McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

1 A - ANO

www.MHEST.com

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

10th Edition

An international reference work in twenty volumes including an index

McGraw-Hill

New York Chicago San Francisco Lisbon London Madrid Mexico City
Milan New Delhi San Juan Seoul Singapore Sydney Toronto

The McGraw-Hill Companies

On the front cover
TraR protein structure of *Agrobacterium tumefaciens*, the causal agent of crown gall disease. (*Image: Argonne National Laboratory*)

Library of Congress Cataloging-in-Publication Data

McGraw-Hill encyclopedia of science & technology—10th ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-07-144143-8 (alk. paper)
1. Science—Encyclopedias. 2. Technology—Encyclopedias.
I. Title: Encyclopedia of science & technology.
II. Title: McGraw-Hill encyclopedia of science and technology.
Q121.M3 2007
503—dc22 2007006137

13-digit ISBN: 978-0-07-144143-8 (set)

10-digit ISBN: 0-07-144143-3 (set)

McGraw-Hill
A Division of The McGraw-Hill Companies

McGRAW-HILL ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY
Copyright © 2007, 2002, 1997, 1992, 1987, 1982, 1977, 1971, 1966, 1960 by The McGraw-Hill Companies, Inc. All rights reserved. Printed in the United States of America. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

1 2 3 4 5 6 7 8 9 0 DOW/DOW 0 1 2 1 0 9 8 7

This book set was printed on acid-free paper.

It was set in Garamond Book and Neue Helvetica Black Condensed by Aptara, Falls Church, Virginia. The art was prepared by Aptara. The book was printed and bound by R. R. Donnelley, Willard, Ohio.

Editorial and Technical Staff

Mark D. Licker, Publisher

Elizabeth Geller, Managing Editor

Jonathan Weil, Senior Staff Editor

David Blumel, Staff Editor

Alyssa Rappaport, Staff Editor

Stefan Malmoli, Staff Editor

Jessa Netting, Staff Editor

Renee Taylor, Editorial Assistant

Charles Wagner, Manager, Digital Content

Editing, Design, & Production Staff

Roger Kasunic, Vice President—Editing, Design, and Production

Joe Faulk, Editing Manager

Frank Kotowski, Jr., Senior Editing Supervisor

Ron Lane, Art Director

Thomas G. Kowalczyk, Production Manager

Suppliers

Art prepared by Aptara, Falls Church, Virginia.

Colorplates printed by Lehigh Press Lithographers, Pennsauken, New Jersey.

Color maps of continents prepared by Mapping Specialists, Ltd., Madison, Wisconsin.

Index prepared by Stephen R. Ingle, WordCo Indexing Services, Norwich, Connecticut.

Consulting Editors

Dr. Milton B. Adesnik. *Department of Cell Biology, New York University School of Medicine, New York.* CELL BIOLOGY.

Prof. Eugene A. Avallone. *Consulting Engineer; Professor Emeritus of Mechanical Engineering, City College of the City University of New York.* MECHANICAL AND POWER ENGINEERING.

Prof. William P. Banks. *Chairman, Department of Psychology, Pomona College, Claremont, California.* PHYSIOLOGICAL AND EXPERIMENTAL PSYCHOLOGY.

Dr. Phillip Barak. *Department of Soil Science, University of Wisconsin–Madison.* CROP SCIENCE.

Prof. Ted Barnes. *Physics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee.* ELEMENTARY PARTICLE PHYSICS.

Dr. Paul Barrett. *Department of Palaeontology, The Natural History Museum, London, United Kingdom.* VERTEBRATE PALEONTOLOGY.

Prof. S. H. Benabdallah. *Department of Mechanical Engineering, Royal Military College of Canada, Kingston, Ontario.* MECHANICAL ENGINEERING.

Prof. Ray Benekohal. *Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign.* TRANSPORTATION ENGINEERING.

Prof. Carrol Bingham. *Department of Physics, University of Tennessee, Knoxville.* NUCLEAR AND ELEMENTARY PARTICLE PHYSICS.

Michael Bosworth. *Vienna, Virginia.* NAVAL ARCHITECTURE AND MARINE ENGINEERING.

Robert D. Briskman. *Technical Executive, Sirius Satellite Radio, New York.* TELECOMMUNICATIONS.

Prof. Richard O. Buckius. *Department of Mechanical & Industrial Engineering, University of Illinois at Urbana-Champaign.* MECHANICAL ENGINEERING.

Dr. Robyn J. Burnham. *Department of Geological Sciences, University of Michigan, Ann Arbor.* PALEOBOTANY.

Dr. Mark Chase. *Molecular Systematics Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey, United Kingdom.* PLANT TAXONOMY.

Prof. Wai-Fah Chen. *Dean, College of Engineering, University of Hawaii at Manoa.* CIVIL ENGINEERING.

Dr. John F. Clark. *Director, Graduate Studies, and Professor, Space Systems, Spaceport Graduate Center, Florida Institute of Technology, Satellite Beach.* SPACE TECHNOLOGY.

Prof. J. John Cohen. *Department of Immunology, University of Colorado School of Medicine, Denver.* IMMUNOLOGY.

Prof. Mark Davies. *Department of Mechanical & Aeronautical Engineering, University of Limerick, Ireland.* AERONAUTICAL ENGINEERING AND PROPULSION.

Prof. Peter J. Davies. *Department of Plant Biology, Cornell University, Ithaca, New York.* PLANT PHYSIOLOGY.

Dr. M. E. El-Hawary. *Associate Dean of Engineering, Dalhousie University, Halifax, Nova Scotia, Canada.* ELECTRICAL POWER ENGINEERING.

Barry A. J. Fisher. *Director, Scientific Services Bureau, Los Angeles County Sheriff's Department, California.* FORENSIC SCIENCE.

Dr. Peter L. Forey. *Department of Palaeontology, The Natural History Museum, London.* ANIMAL SYSTEMATICS AND EVOLUTION.

Dr. Richard L. Greenspan. *The Charles Stark Draper Laboratory, Cambridge, Massachusetts.* NAVIGATION.

Dr. Timothy A. Haley. *Food Science Department, Purdue University, West Lafayette, Indiana.* FOOD SCIENCE.

Dr. Lisa Hammersley. *Department of Geology, California State University, Sacramento.* PETROLOGY.

Prof. John P. Harley. *Department of Biological Sciences, Eastern Kentucky University, Richmond.* MICROBIOLOGY.

Prof. Terry Harrison. *Department of Anthropology, New York University, New York.* ANTHROPOLOGY AND ARCHEOLOGY.

Dr. Ralph E. Hoffman. *Yale New Haven Psychiatric Hospital, New Haven, Connecticut.* PSYCHIATRY.

Dr. Gary Hogg. *Professor of Industrial Engineering, Arizona State University, Tempe.* INDUSTRIAL AND PRODUCTION ENGINEERING.

Prof. A. Gordon L. Holloway. *Chair, Department of Mechanical Engineering, University of New Brunswick, Fredericton, Canada.* FLUID MECHANICS.

Dr. Hong Hua. *Director of 3D Visualization and Imaging System Lab, College of Optical Sciences, The University of Arizona, Tucson.* ELECTROMAGNETIC RADIATION AND OPTICS.

Dr. George Hudler. *Department of Plant Pathology, Cornell University, Ithaca, New York.* PLANT PATHOLOGY.

Dr. S. C. Jong. *Senior Staff Scientist and Program Director, Yeast Genetic Stock Center, American Type Culture Collection, Manassas, Virginia.* MYCOLOGY.

Dr. Edwin Kan. *Associate Professor, School of Electrical and Computer Engineering, Cornell University, Ithaca, New York.* PHYSICAL ELECTRONICS.

Dr. Peter M. Kareiva. *Lead Scientist, Pacific Western Conservation Region, The Nature Conservancy, Seattle, Washington.* ECOLOGY AND CONSERVATION.

Dr. Bryan P. Kibble. *National Physical Laboratory, Teddington, Middlesex, England.* ELECTRICITY AND ELECTROMAGNETISM.

Prof. Robert E. Knowlton. *Department of Biological Sciences, George Washington University, Washington, DC.* INVERTEBRATE ZOOLOGY.

Prof. Cynthia K. Larive. *Department of Chemistry, University of California, Riverside.* ANALYTICAL CHEMISTRY.

Prof. Chao-Jun Li. *Canada Research Chair in Green Chemistry, Department of Chemistry, McGill University, Montreal, Quebec, Canada.* ORGANIC CHEMISTRY.

Prof. Donald W. Linzey. *Wytheville Community College, Wytheville, Virginia.* VERTEBRATE ZOOLOGY.

Dr. Philip V. Lopresti. *Retired; formerly, Engineering Research Center, AT&T Bell Laboratories, Princeton, New Jersey.* ELECTRONIC CIRCUITS.

Dr. Dan Luss. *Cullen Professor of Engineering, Department of Chemical Engineering, University of Houston, Texas.* CHEMICAL ENGINEERING.

Prof. Scott M. McLennan. *Department of Geosciences, State University of New York at Stony Brook.* GEOLOGY (PHYSICAL, HISTORICAL, AND SEDIMENTARY).

Prof. Philip L. Marston. *Department of Physics and Astronomy, Washington State University, Pullman.* ACOUSTICS.

Dr. Ramon A. Mata-Toledo. *Professor of Computer Science, James Madison University, Harrisonburg, Virginia.* COMPUTERS.

Prof. Krzysztof Matyjaszewski. *J. C. Warner Professor of Natural Sciences, Department of Chemistry, Carnegie-Mellon University, Pittsburgh, Pennsylvania.* POLYMER SCIENCE AND ENGINEERING.

Prof. Joel S. Miller. *Department of Chemistry, University of Utah, Salt Lake City.* INORGANIC CHEMISTRY.

Dr. Orlando J. Miller. *Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, Detroit, Michigan.* GENETICS.

Prof. Jay M. Pasachoff. *Director, Hopkins Observatory, Williams College, Williamstown, Massachusetts.* ASTRONOMY.

Prof. David J. Pegg. *Department of Physics, University of Tennessee, Knoxville.* ATOMIC AND MOLECULAR PHYSICS.

Prof. J. Jeffrey Peirce. *Department of Civil and Environmental Engineering, Edmund T. Pratt Jr. School of Engineering, Duke University, Durham, North Carolina.* ENVIRONMENTAL ENGINEERING.

Dr. William C. Peters. *Professor Emeritus, Mining and Geological Engineering, University of Arizona, Tucson.* MINING ENGINEERING.

Prof. Arthur N. Popper. *Department of Biology, University of Maryland, College Park.* NEUROSCIENCE.

Dr. Kenneth Pritzker. *Director, Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario, Canada.* MEDICINE AND PATHOLOGY.

Prof. Justin Revenaugh. *Department of Geology and Geophysics, University of Minnesota, Minneapolis.* GEOPHYSICS.

Dr. Roger M. Rowell. *USDA-Forest Service, Forest Products Laboratory, Madison, Wisconsin.* FORESTRY.

Dr. John L. Safko. *Distinguished Professor Emeritus, Physics and Astronomy, Associated Faculty, School of the Environment, University of South Carolina, Columbia.* CLASSICAL MECHANICS.

Dr. Andrew P. Sage. *Founding Dean Emeritus and First American Bank Professor, University Professor, School of Information Technology and Engineering, George Mason University, Fairfax, Virginia.* CONTROL AND INFORMATION SYSTEMS.

Dr. Alfred S. Schlachter. *Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, California.* ATOMIC AND MOLECULAR PHYSICS.

Prof. Ivan Schuller. *Department of Physics, University of California-San Diego, La Jolla, California.* CONDENSED-MATTER PHYSICS.

Dr. David Sherman. *Department of Earth Sciences, University of Bristol, United Kingdom.* MINERALOGY.

Dr. Steven A. Slack. *Department of Plant Pathology, Cornell University, Ithaca, New York.* PLANT PATHOLOGY.

Dr. Arthur A. Spector. *Professor of Biochemistry and Internal Medicine, University of Iowa, Iowa City.* BIOCHEMISTRY.

Dr. Bruce A. Stanley. *Director, Scientific Programs, Section of Technology Development and Research Resources H093, The Pennsylvania State University College of Medicine, Hershey.* PHYSIOLOGY.

Prof. Anthony P. Stanton. *Carnegie-Mellon University, Pittsburgh, Pennsylvania.* GRAPHIC ARTS AND PHOTOGRAPHY.

Dr. Trent Stephens. *Professor of Anatomy and Embryology, Idaho State University, Pocatello.* DEVELOPMENTAL BIOLOGY.

Prof. John F. Timoney. *Department of Veterinary Science, University of Kentucky, Lexington.* VETERINARY MEDICINE.

Prof. Antonio J. Torres. *Department of Food Science and Technology, Oregon State University, Corvallis.* FOOD SCIENCE.

Dr. Sally E. Walker. *Associate Professor of Geology and Marine Science, University of Georgia, Athens.* INVERTEBRATE PALEONTOLOGY.

Prof. Pao K. Wang. *Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison.* METEOROLOGY AND CLIMATOLOGY.

Prof. Frank M. White. *Professor Emeritus, Department of Mechanical Engineering, University of Rhode Island, Kingston.* FLUID MECHANICS.

Prof. Mary Anne White. *Killam Research Professor in Materials Science, Department of Chemistry, Dalhousie University, Halifax, Nova Scotia, Canada.* MATERIALS SCIENCE AND METALLURGIC ENGINEERING.

Prof. Thomas A. Wikle. *Department of Geography, and Associate Dean, College of Arts and Sciences, Oklahoma State University, Stillwater.* PHYSICAL GEOGRAPHY.

Preface

As we approach the 50th anniversary of the first publication of the *McGraw-Hill Encyclopedia of Science & Technology* in 1960, we are reminded that the objectives expressed in the preface to the first edition remain valid to this day: *"To provide the widest possible range of articles that will be understandable and useful to any person of modest technical training who wants to obtain information outside his [or her] particular field of specialization."* This philosophy has guided the authors and editors of the Encyclopedia in their striving to serve the reference and educational needs of students, professionals, and librarians worldwide by offering reliable, comprehensive information in all major fields of science and engineering.

Advances in science and technology over the five years since the publication of the last edition (2002) have been remarkable in many respects, with some advances even at the center of hotly debated societal issues. Our legislatures, for example, weigh the potential benefits of stem cell research against the ethical concerns raised by some. Ever more sophisticated models running on powerful computing systems point to human-induced changes in global climate over this century in a range that could cause significant societal disruptions. How were these models developed? What technologies are available to mitigate these changes, and at what cost? Advances in information and telecommunications technology are made at an accelerating pace; how are they changing how people communicate, interact, learn, and do business? Dealing with such issues requires authoritative sources of scientific and technical information that are understandable to the nonspecialist, and thus the need for a comprehensive scientific encyclopedia seems as great as ever.

Even though the goal has been a broad revision of the previous edition, much effort was concentrated in certain rapidly advancing areas, particularly in cell and molecular biology; information technology and telecommunications; nanotechnology; the environmental, earth, and climate sciences; materials science; and cosmology, among others. The result of these efforts are the 7100 articles in the tenth edition, many new or rewritten, and others updated as needed, following the recommendations of our distinguished Board of Consulting Editors. Each article was prepared by one or more experts nominated by the Board and coming from universities, industry, and public agencies worldwide.

The Encyclopedia has always been intended to be a work *of*, not *about*, science and technology, with each article reflecting the scope and sophistication of the topic. However, since the Encyclopedia is used primarily by nonspecialists, care has been given to structure, edit, and illustrate the articles such that the reader has ready access to the key concepts of the field. Articles start with a definition and a concise overview of the topic. The subject matter is developed according to a clear outline and concludes with a bibliography of generally available publications for further study (nearly 25,000 bibliographic entries in total). One of the hallmarks of the Encyclopedia is the numerous cross references to related articles—more than 60,000 in this edition—which not only allow the reader to acquire background information, or to study in greater depth, but also help illuminate the context of the material and broader connections among topics.

More than 12,000 illustrations enhance the text and allow the reader to visualize concepts. These include photographic images, two-color graphs, charts, and maps redrawn for uniformity of style; and 88 full-color plates. Over 1400 tables provide useful data, and, where appropriate, chemical structures (900) and reactions (more than 2500) are included as well as mathematical equations (8500). An appendix with tables in volume 20 explains scientific notation and the dual systems for units of measurement—the U.S. Customary System and the International System (SI)—used throughout the text. The reader will also find an extensive table of conversion factors for the measurement systems, symbols used in scientific writing, and chemical nomenclature.

Ease of access to this tremendous amount of information is essential if the Encyclopedia is to fulfill its mission. There are several ways of finding information. The articles are arranged alphabetically, and of course the reader can go directly to an entry in the appropriate volume. To find all of the information available on a particular subject, the reader may use the Analytical Index, which occupies more than 500 pages. The Topical Index lists all of the articles pertaining to a particular discipline, for example, theoretical physics or microbiology. The Study Guides outline the major fields of science and technology and their subdisciplines, and provide for each a list of overview articles. They offer the opportunity not only to find a relevant article but also to systematically learn or review a field. Finally, a comprehensive list of the contributors and their affiliations along with the titles of the contributions allows the reader to find the works of particular authors. All of these indexes may be found in volume 20.

New to the tenth edition is a companion Web site

http://MHEST.com

which contains materials complementary to the printed text, much of it drawn from AccessScience, the McGraw-Hill Encyclopedia of Science & Technology Online: updates of critical articles, "rich media" (such as animations, videos, and audio presentations), commentaries on articles, explorations of topical themes, and more. The content of MHEST.com will be updated periodically, and we invite the reader to visit it often.

The editors of the Encyclopedia are grateful to the international authorship, the consulting editors, reviewers, and many others whose efforts have made this revision possible. As in past editions, we dedicate the product of this massive collaboration to our readers and trust that it will continue to uniquely serve their need for high-quality information as well as the objectives of scientific education and technological literacy.

Mark D. Licker
Publisher

Organization of the Encyclopedia

The *McGraw-Hill Encyclopedia of Science & Technology* presents pertinent information in every field of modern science and technology. The 7100 articles are arranged alphabetically in the 19 text volumes. The range of article titles included in each volume is indicated on the spine and front cover (for example, volume 1 contains articles with titles starting with "Aar" up to "Ano"). Thus the reader may quickly locate an article by its title. The 20th volume contains the indexes and ancillary materials.

Broad survey articles are available for each of the disciplines covered; even readers with little prior knowledge of that discipline will find the basic concepts covered in these articles. From the survey article, the reader may proceed to more specialized articles using the cross-referencing system. These cross references are set in small capital letters for emphasis and are inserted at the relevant points in the text. For example, in a survey article such as **Digital computer**, the reader is directed to numerous other articles such as COMPUTER PERIPHERAL DEVICES, COMPUTER STORAGE TECHNOLOGY, MICROPROCESSOR, and PROGRAMMING LANGUAGES. The references may lead to subjects that have not occurred to the reader. The article **Solvent** has such diverse cross references as COORDINATION CHEMISTRY, HALOGENATED HYDROCARBON, INDUSTRIAL HEALTH AND SAFETY, and WATER POLLUTION. The cross references not only lead to articles of greater specialization but also help illuminate the context of the article and the broader connections among topics. This edition contains more than 60,000 cross references.

The pattern of proceeding from the general to the specific has been employed not only in the plan of the Encyclopedia but within the body of the articles. Each article begins with a definition of the subject, followed by sufficient background material to give a frame of reference and permit the reader to move into the detailed text of the article. Within the text are centered heads and two levels of sideheads that outline the article; they are intended to enhance understanding and can guide the user that prefers to read selectively the sections of a long article.

Alphabetization of article titles is by word, not by letter, with a comma providing a stop in occasional inverted article titles (so that subject matter can be grouped). Two examples of sequence are:

Air	**Earth, age of**
Air-cushion vehicle	**Earth, heat flow of**
Air mass	**Earth crust**
Air-traffic control	**Earth tides**
Aircraft fuel	**Earthquake**

Numerous illustrations, both line drawings and images, contribute to the utility, clarity, and interest of the text. Each illustration (as well as each table) is called out in boldface at its first mention in the text. This emphasis enables the reader to move from an illustration to the point in the text where the illustration is often discussed in detail.

To meet the needs of the Encyclopedia's broad readership, measurements are given in dual systems of units: The U.S. Customary System is used throughout the text along with equivalent measurements in the International System of Units. In particular cases, such as references to measurements in some illustrations or tables, conversion factors may be given for simplicity.

The contributor's full name appears at the end of an article section or an entire article. Each author is identified in an alphabetical Contributors list in volume 20, which cites the university, laboratory, business, or other organization with which the author is affiliated and the titles of the articles written by that contributor.

Most of the articles contain bibliographies citing useful sources. The bibliographies are placed at the ends of articles or occasionally at the ends of major sections in long articles. For additional bibliographies, the reader should refer to related articles as indicated by cross references.

Thus, the alphabetical arrangement of article titles, the text headings, the cross references, and the bibliographies permit the reader to research a particular topic by simply taking a volume from the shelf. However, the reader can also find information in the Encyclopedia by using the Analytical Index and the Topical Index in volume 20. The Analytical Index—over 500 pages in length—contains each important term, concept, and person mentioned throughout the 19 text volumes. It guides the reader to the volume numbers and page numbers concerned with a specific point. The reader wishing to consult everything in the Encyclopedia on a particular aspect of a subject will find that the Analytical Index is the best approach. A broader survey may be made through the Topical Index, which groups all article titles of the Encyclopedia under 90 general headings. For example, under "Atomic and molecular physics," 90 articles are listed, and under "Biochemistry," 147. The Topical Index thus enables the reader quickly to identify all articles in the Encyclopedia in a particular subject area.

The Study Guides in volume 20 provide highly structured outlines of major scientific disciplines and relate groups of Encyclopedia articles to each outline heading. By following a guide, the reader is led through pertinent Encyclopedia articles in a sequence that provides an overall grasp of the discipline.

A useful feature is the section "Scientific Notation" in volume 20. It clarifies usage of symbols, abbreviations, and nomenclature, and is especially valuable in making conversions between the International System, U.S. Customary, and metric measurements.

With the 10th edition, the editors are introducing a new feature to enhance the usefulness of the Encyclopedia: a companion Web site

http://MHEST.com

containing periodically updated collections of articles, graphics, and multimedia content pertaining to a timely theme, as well as selected updates of Encyclopedia articles as developments in science and technology dictate. We encourage readers to visit the site to benefit from this material.

Table of Contents for Colorplates

Colorplates face the page numbers cited. The plates are relevant to the subject matter in the respective articles, but may not be specifically mentioned in the text.

Africa ..... volume 1 page 204
Aircraft testing ..... volume 1 page 372
Andromeda Galaxy ..... volume 1 page 660

Antarctica ..... volume 2 page 10
Antibody ..... volume 2 page 48
Apes ..... volume 2 page 90
Arachnida ..... volume 2 page 116
Asia ..... volume 2 page 244
Aurora ..... volume 2 page 418
Australia ..... volume 2 page 422
Aves ..... volume 2 page 480

Bioeroding sponges ..... volume 3 page 28
Chandra X-ray Observatory ..... volume 3 page 734

Chilopoda ..... volume 4 page 88
Chromosome ..... volume 4 page 140
Climatology ..... volume 4 page 254
Coleoptera ..... volume 4 page 404
Computational fluid dynamics ..... volume 4 page 534
Computer graphics ..... volume 4 page 562
Corallimorpharia ..... volume 4 page 754

Cytochemistry ..... volume 5 page 192
Diffraction ..... volume 5 page 496
Earthquake ..... volume 5 page 820

Eclipse ..... volume 6 page 34
Europe ..... volume 6 page 720

Field-emission microscopy ..... volume 7 page 141
Fluid mechanics ..... volume 7 page 226
Fluvial erosion landforms ..... volume 7 page 246
Galaxy, external ..... volume 7 page 642

Hail ..... volume 8 page 334
Holography ..... volume 8 page 582

Infrared astronomy ..... volume 9 page 180
Integrated circuits ..... volume 9 page 288
Interference of waves ..... volume 9 page 332
Interstellar matter ..... volume 9 page 372
Ionosphere ..... volume 9 page 446
Jupiter ..... volume 9 page 540
Laser ..... volume 9 page 664
Lepidoptera ..... volume 9 page 774

Lightning ..... volume 10 page 38
Magellanic clouds ..... volume 10 page 262
Mars ..... volume 10 page 504
Metallography ..... volume 10 page 770

Meteorite ..... volume 11 page 14
Meteorological optics ..... volume 11 page 30
Mitochondria ..... volume 11 page 272
Moiré pattern ..... volume 11 page 308
Molecular cloud ..... volume 11 page 334
Moon ..... volume 11 page 432
Nebula ..... volume 11 page 646
Neptune ..... volume 11 page 678

Nonlinear optics ..... volume 12 page 32
Nonstoichiometric compounds ..... volume 12 page 64
North America ..... volume 12 page 72
Orion Nebula ..... volume 12 page 586
Paleobotany ..... volume 12 page 706

Particle trap ..... volume 13 page 82
Perciformes ..... volume 13 page 162
Phase equilibrium ..... volume 13 page 316
Photoelasticity ..... volume 13 page 404
Planetary nebula ..... volume 13 page 660

Primates ..... volume 14 page 344

Radar meteorology ..... volume 15 page 22
Radio astronomy ..... volume 15 page 66
Remote sensing ..... volume 15 page 370
Reptilia ..... volume 15 page 424
Rhizostomeae ..... volume 15 page 522

Satellite meteorology ..... volume 16 page 42
Saturn ..... volume 16 page 64
Scanning tunneling microscope ..... volume 16 page 78
Scyphozoa ..... volume 16 page 154
Sintering ..... volume 16 page 520

South America ..... volume 17 page 48
Space flight ..... volume 17 page 114
Space station ..... volume 17 page 172
Spectroscopy ..... volume 17 page 224
Squamata ..... volume 17 page 314
Starburst galaxy ..... volume 17 page 362
Stellar evolution ..... volume 17 page 432
Storm detection ..... volume 17 page 506
Stroboscopic photography ..... volume 17 page 552
Sun ..... volume 17 page 676
Supernova ..... volume 17 page 746

Topographic surveying and mapping ..... volume 18 page 498

Uranium ..... volume 19 page 100
Venus ..... volume 19 page 216
Virus ..... volume 19 page 286
Volcano ..... volume 19 page 344

A

A15 phases — Anoxic zones

A15 phases

A series of intermetallic compounds that have a particular crystal structure and the chemical formula A_3B, where A represents a transition element and B can be either a transition element or a nontransition element. Many A15 compounds exhibit the phenomenon of superconductivity at relatively high temperatures in the neighborhood of 20 K (−424°F) and in high magnetic fields on the order of several tens of teslas (several hundred kilogauss). High-temperature–high-field superconductivity has a number of important technological applications and is a challenging fundamental research area in condensed-matter physics.

Crystal structure. The A15 compounds crystallize in a structure in which the unit cell, the repeating unit of the crystal structure, has the overall shape of a cube. The B atoms are located at the corners and in the center of the cube, while the A atoms are arranged in pairs on the cube faces (see **illus**.). A special characteristic of the A15 crystal structure is that the A atoms form mutually orthogonal linear chains that run throughout the crystal lattice, as shown in the illustration. The intrachain distance between A atoms is the shortest distance between atoms in the A15 crystal structure, and is about 22% less than the smallest interchain distance between A atoms. The extraordinary superconducting properties of the A15 compounds are believed to be primarily associated with these linear chains of transition-element A atoms.

At low temperature, some of the A15 phases undergo a diffusionless transformation from a cubic to a tetragonal structure, in which the length of one of the sides of the unit cell changes with respect to the length of the other two. This structural transformation was first observed in V_3Si at approximately 21 K (−422°F), and has also been found in Nb_3Sn at 43 K (−382°F), V_3Ga at above 50 K (−370°F), and Nb_3Al at 80 K (−316°F). *See* CRYSTAL STRUCTURE.

Superconductivity. Superconductivity is a phenomenon in which a metal, when cooled below its superconducting transition temperature T_c, loses all resistance to the flow of an electric current. The A15 compound Nb_3Ge has a T_c of 23.2 K (−417.9°F), the highest value of the A15 compounds, and, prior to 1986, the highest value ever observed for any known material. Other A15 compounds with high T_c include Nb_3Al (18.8 K or −425.8°F), Nb_3Ga (20.3 K or −423.1°F), Nb_3Si (approximately 19 K or −426°F), Nb_3Sn (18.0 K or −427.3°F), V_3Ga (15.9 K or −431°F), and V_3Si (17.1 K or −428.9°F).

In 1986, evidence of superconductivity with a T_c of approximately 30 K (−405°F) was reported for a material containing the elements lanthanum, barium, copper, and oxygen. Since then, several families of compounds containing copper, oxygen, and

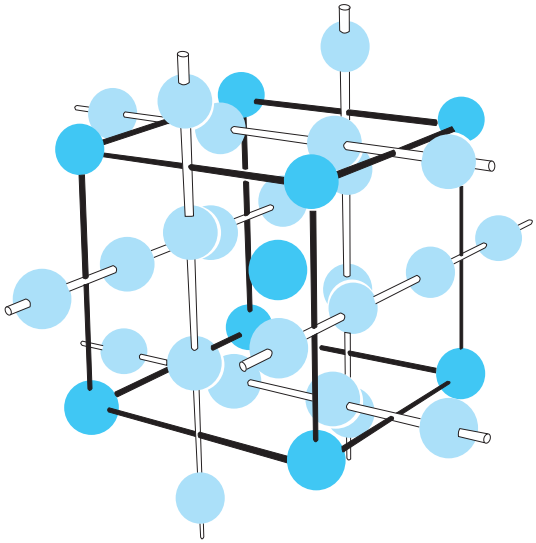

A15 crystal structure of the A_3B intermetallic compounds. The light spheres represent the A atoms; the dark spheres represent the B atoms. The linear chains of A atoms are emphasized.

other elements have been discovered with yet higher superconducting transition temperatures, including a value of 135 K (−217°F) for a compound made of mercury, thallium, barium, calcium, copper, and oxygen. The origin of high-temperature superconductivity in the copper oxide compounds has not yet been established and may be different from that in the conventional superconductors such as the A15 compounds. In a conventional superconductor, the interaction of the negatively charged electrons with the positively charged ions of the crystal lattice produces an attractive interaction between electrons, causing them to form electron pairs, called Cooper pairs, that are responsible for the superconductivity.

The superconductivity of the A15 compounds can be destroyed with a sufficiently high magnetic field, called the upper critical magnetic field, $H_{c2}(T)$, which depends on the temperature T and varies from material to material. As the temperature is lowered, $H_{c2}(T)$ generally increases from zero at T_c to a maximum value, $H_{c2}(0)$, as the temperature approaches absolute zero (0 K or −459.67°F). The A15 compounds also have very high values of $H_{c2}(0)$, the highest of which is 44 teslas (T) for a pseudobinary A15 compound with the composition $Nb_{79}(Al_{73}Ge_{27})_{21}$. Other A15 compounds with high values of $H_{c2}(0)$ include Nb_3Al (32 T), Nb_3Ge (39 T), Nb_3Sn (23 T), V_3Ga (21 T), and V_3Si (25 T). For comparison, the highest value of $H_{c2}(0)$ ever observed for a conventional superconducting material is about 60 T for the Chevrel-phase compound $PbMo_6S_8$, while the values of $H_{c2}(0)$ for the highest-T_c copper oxide superconductors may reach several hundred teslas, values so large that they cannot readily be measured by using the magnetic fields and techniques that are presently available in the laboratory. *See* CHEVREL PHASES; SUPERCONDUCTIVITY.

Technological applications. Because of their high values of T_c and H_{c2}, and critical current density J_c, the A15 compounds have a number of important technological applications. (The critical current density is the highest electric current per unit area that a material can conduct and still remain superconducting in a given magnetic field.) Processes have been developed for preparing multifilamentary superconducting wires that consist of numerous filaments of a superconducting A15 compound, such as Nb_3Sn, embedded in a nonsuperconducting copper matrix. Superconducting wires can be used in electric power transmission lines and to wind electrically lossless coils (solenoids) for superconducting electrical machinery (motors and generators) and magnets. Superconducting magnets are employed to produce intense magnetic fields for laboratory research, confinement of high-temperature plasmas in nuclear fusion research, bending beams of charged particles in accelerators, levitation of high-speed trains, mineral separation, nuclear magnetic resonance imaging, and energy storage. If conductors of the high-temperature, high-field oxide superconductors can eventually be fabricated with high values of J_c, they may revolutionize the technological applications of superconductivity. *See* SUPERCONDUCTING DEVICES.

M. Brian Maple

Bibliography. W. Buckle, *Superconductivity: Fundamentals and Applications*, 1991; P. Fulde (ed.), *Superconductivity of Transition Metals*, Springer Series in Solid State Sciences, vol. 27, 1982; T. H. Geballe and J. K. Hulm, Superconductivity: The state that came in from the cold, *Science*, 239:367–374, 1988; Y. Kao (ed.), *Superconductivity and Its Applications,* 1992; V. Z. Kresin, H. Morawitz, and S. A. Wolf, *Mechanics of Conventional and High T_c Superconductivity*, 1993.

Aardvark

A large piglike mammal in the order Tubulidentata. The order contains a single family, Orycteropodidae, and a single species, *Orycteropus afer*. Aardvarks occur throughout Africa south of the Sahara, wherever suitable habitat exists. In the past, tubulidentates were often considered closely related to ungulates. However, recent data from mitochondrial and nuclear genes support a relationship among aardvarks, elephant-shrews, paenungulates (hyraxes, sirenians, and proboscideans), and golden moles (Chrysochloridae). These ecologically divergent adaptive types probably originated in Africa; the molecular evidence implies that they may have arisen there from a common ancestor that existed in the Cretaceous Period, when Africa was isolated from other continents. This order shows the results of an extreme adaptation for burrowing and feeding on small food items (mainly termites). *See* MAMMALIA; TUBULIDENTATA.

Characteristics. Aardvarks, also known as antbears, resemble a medium-sized to large pig (see **illustration**). The body is massive with an elongate head and snout. The round, blunt, piglike muzzle ends in circular nostrils that possess fleshy tentacles on the nasal septum and dense tracts of white hairs that can serve as a filter and seal the nostrils to prevent soil from entering the lungs. The tough, thick skin is sparsely covered with bristly hair ranging from dull brownish gray to dull yellowish

Aardvark, *Orycteropus afer*. (*Photo by Lloyd Glenn Ingles; © 2001 California Academy of Sciences*)

gray. Numerous vibrissae (hairs with specialized erectile tissues) are present on the face. The waxy, smooth, tubular donkeylike ears are large and can be folded back to exclude dirt when the animal is burrowing, and they can be moved independently of each other. The strong, muscular tail is kangaroolike, tapering to a point. Incisors and canines are lacking in the dentition of adults but are present in the deciduous dentition. Teeth on the sides of the jaw are simple, peglike, open-rooted (not anchored in the jaw), and grow continuously during the animal's lifetime. The teeth are hollow; instead of a pulp cavity, they are transversed by a number of tubules. The dental formula is I 0/0 C 0/0 Pm 2/2 M 3/3 × 2 for a total of 20 teeth. Anal scent glands are present in both sexes. The senses of sight and hearing are acute, but their eyesight appears to be poor. Adult aardvarks have a head and body length of 100–158 cm (39–62 in.), a tail length of 44–71 cm (17–28 in.), and a shoulder height of 60–65 cm (23–25 in.). Most weigh 50–70 kg (110–153 lb), although some individuals may reach 100 kg (219 lb). Aardvarks exhibit a low basal metabolic rate and a low body temperature of about 34.5°C (94.1°F). *See* DENTITION; SCENT GLAND.

The aardvark is one of the world's great diggers. Its short, thick legs (with four toes on the front feet, five on the rear ones) are armed with long, powerful, sharp-edged, spoon-shaped claws that are intermediate between hoof and nail. The aardvark uses its claws to excavate the extensive burrows in which it dwells. Able to dig with amazing speed, it digs shallow holes when searching for food, larger burrows used for temporary shelter, and extensive tunnel systems in which the young are born. Tunnels may be 13 m (43 ft) long with numerous chambers and several entrances. Only mothers and their young share burrows. Burrows dug by aardvarks may be used by a variety of mammals, including warthogs, hyenas, porcupines, jackals, bat-eared foxes, hares, bats, ground squirrels, and civets, as well as monitor lizards and owls. One bird, the ant-eating chat (*Myrmecocicla formicivora*), nests in occupied burrows. In agricultural areas, burrows may cause damage to farming equipment and earthen dams.

Aardvarks are primarily nocturnal and solitary, and occupy a variety of habitats, including grassy plains, bush country, woodland, and savannah. They appear to prefer sandy soils. The main factor in their distribution is the presence of sufficient quantities of termites and ants. The aardvark is of considerable economic importance in keeping the great hordes of termites in check; if termites are not controlled, they do serious damage. The numbers of termites and ants consumed by the big-bodied aardvark in a single night are staggering. Studies have found that more than 50,000 insects may be consumed by an aardvark in one night. The long, round, thin, sticky, protrusible foot-long tongue is broader than that of the typical anteater or pangolin; together with the well-developed salivary glands, it can retrieve more termites in a shorter period of time than the two others. Since aardvarks do not need to chew their food, they have no need for incisors or canine teeth. Instead, their stomachs have a muscular pyloric area that serves a function similar to a gizzard in grinding up the food. They occasionally eat beetles, grasshoppers, and the fruit of the wild cucumber, known as the aardvark pumpkin, apparently as a source of water. *See* ANT; ANTEATER; ISOPTERA.

Breeding and development. A single altricial offspring (a young that is born immature and helpless, thus requiring extended parental care) is born after a gestation period of 7–9 months. Births generally occur from July to November. After remaining in the burrow for about 2 weeks, the young aardvark begins accompanying its mother on her nightly foraging excursions. The young can dig for itself when about 6 months old. Sexual maturity is attained at about 2 years of age. Captive aardvarks have lived up to 10 years.

Threats. Large flesheaters such as the lion, leopard, and ratel (a fierce, badgerlike animal) frequently prey on aardvarks. Pythons also take the young. African natives relish the aardvark's meat. They wear its teeth and claws as bracelets or necklaces as good-luck charms and for warding off illnesses. The tough hide may be made into straps and bracelets. Aardvarks have been greatly reduced in numbers and distribution and are now on Appendix 2 of the Convention on International Trade in Endangered Species (CITES). Termites have caused enormous damage to pasture and cereal crops in areas where aardvarks and other insectivorous animals have been exterminated. *See* ENDANGERED SPECIES.

Donald W. Linzey

Bibliography. D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Abaca

One of the strongest of the hard fibers, commercially known as Manila hemp. Abaca is obtained from the leafstalks of a member of the banana family, *Musa textilis*. The plant resembles the fruiting banana, but is a bit shorter in stature, bears small inedible fruits, and has leaves that stand more erect than those of the banana and are slightly narrower, more pointed, and 5–7 ft (1.5–2 m) long. Relatives of abaca grow wild throughout Southeast Asia, but the plant was domesticated long ago in the southern Philippines. Experiments have succeeded in growing plants yielding fiber in a few other parts of the world, chiefly Middle America, but commercial production has come almost entirely from the Philippines and Borneo. *See* ZINGIBERALES.

Environment and cultivation. Abaca prefers a warm climate with year-round rainfall, high humidity, and absence of strong winds. Soils must always be moist but the plant does not tolerate waterlogging. Abaca grows best on alluvial soils in the southern Philippines and northern Borneo below 1500 ft (457 m) elevation. The plant is best propagated by rootstalk

Appearance of diseased abaca. (*a*) Affected by mosaic; note ragged leaves and chlorotic areas on leaves. (*b*) Bunchy top; plant 1 on left is dead and lower leaves of plant 2 have been killed. (*USDA*)

suckers. There are about 75 varieties grown in the Philippines, grouped into seven categories, each of which varies slightly in height, length, and quality and yield of fiber. Newly planted stock reaches cutting size in 18 months, and a few stalks may be cut from each "hill" every 4 months thereafter for 12–15 years, after which new plantings are made. Plantings on good soils yield 90–150 lb (41–68 kg) of fiber per acre per year from 2 tons (1.8 metric tons) of stalks and the fiber yield ranges 2–4%. On large commercial holdings plants are grown in solid stands with the new stock spaced 15–20 ft (4.5–6 m) apart, but on small farms plants are scattered and mixed with bamboo and fruit trees.

Properties and uses. The fiber ranges 6–14 ft (1.8–4 m) in strand length, is lustrous, and varies from white to dull yellow. Pre-Spanish Filipinos made clothing with the fine fiber, which today has been replaced by softer, more pliable textiles. As one of the longest and strongest plant fibers, resistant to fresh and salt water, abaca is favored for marine hawsers and other high-strength ropes. Abaca is also used in sackings, mattings, strong papers, and handicraft art goods.

Processing and production. The fibers are carried in the outer sections of the leafstalks, and the fleshy material is stripped away mechanically. Traditionally, leafstalks were drawn by hand over a knife blade, under a wooden wedge, and a worker could clean about 15 lb (6.8 kg) of fiber per day. Small power machines strip about 175 lb (79 kg) per worker-day. Large mechanical decorticators save labor and recover more fiber, but the capital outlay is very large.

Diseases. Abaca is affected by several diseases, of which the chief are bunchy top, mosaic, and wilt (see **illus**.). Bunchy top is caused by a virus spread by the banana aphid (*Pentalonia nigronervosa*); leaves become smaller and clustered at the plant top, after which the plant dies. Mosaic is also caused by a virus spread by aphids (chiefly *Rhopalosiphum nymphaeae* and *Aphis gossypii*); the leaves yellow and dry out as the plant dies. Abaca wilt is caused by a soil or water-borne fungus, chiefly attacking plant roots. Filipino small farmers are not easily brought into disease-control cooperative measures, and postwar epidemics of bunchy top and mosaic have caused serious decline in acreage and production. *See* PLANT PATHOLOGY. Elton G. Nelson

Bibliography. R. E. Huke et al., *Shadows on the Land: An Economic Geography of the Philippines*, 1963; W. Manshard, *Tropical Agriculture*, 1975; B. B. Robinson and F. L. Johnson, *Abaca: A Cordage Fiber,* Agr. Monogr. 21, 1954; J. E. Spencer, The abaca plant and its fiber, Manila hemp, *Econ. Bot.*, 7:195–213, 1953; F. L. Wernstedt and J. E. Spencer, *The Philippine Island World: A Physical, Cultural and Regional Geography*, 1967.

Abacus

An early mechanical calculator whose design has evolved through the centuries, with two styles in use today. Both the Chinese and the Japanese styles consist of a frame with a crossbeam. They may be made from many different materials, such as wood

or brass. Rods or wires carrying sliding beads extend vertically through the crossbeam. The Chinese *suan pan* has two beads above the beam on each rod and five beads below. Each rod of the Japanese *soroban* carries one bead above and four below. Similar to the abacus in both construction and use, but much larger, are the counting frames used in elementary schools. Braille versions of the abacus are available for use by those without sight.

Operation. In working with whole numbers, the rightmost rod represents the ones position, with each rod to the left representing the tens, hundreds, thousands, and so forth, respectively. The beads below the crossbeam represent one of that rod's units (that is, a one, a ten, a hundred, and so forth), and those above represent five. Beads are moved from the outer position toward the crossbeam when used to represent a number (**Fig. 1**).

By a series of motions using the thumb and forefinger, numbers may be added and subtracted. Addition and subtraction both work from left to right, in contrast to the usual method, using substantial mental arithmetic. For example, to add 384 + 795, the user begins with the beads positioned to represent 384 (Fig. 1). Then to add 795, one lower bead on the fourth rod from the right is moved up, while simultaneously moving the three beads on the third rod down (700 = 1000 − 300). Next, moving to the tens position, the user adds the 90 by the mental method of "add 100 and subtract 10," thus moving one lower bead up on the hundreds rod and one lower bead down from the center on the tens rod. Finally, to add the 5, the user simply brings one of the upper beads down to the crosspiece. The result of the addition, 384 + 795 = 1179, then appears (**Fig. 2**).

Subtraction is done in a similar manner. In order to subtract 7 from any position, for example, the user subtracts 10 and adds 3. Both multiplication and division can also be performed with the help of the abacus, but these are much more complicated.

Use. The abacus, in contrast to the electronic calculator, is simply an aid to mental computation. A well-developed facility with numbers is required in order to use it effectively. For this reason, it is the calculator of choice for teachers in the Far East. The advent of the electronic calculator, while a boon to the scientific community, may have little impact on shop owners who are comfortable with the abacus. *See* CALCULATORS.

Marjorie K. Gross

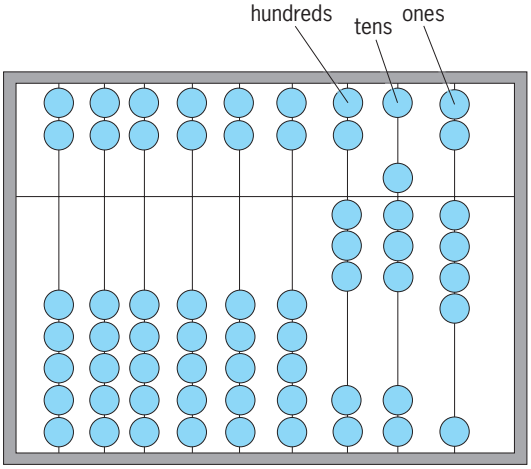

Fig. 1. Beads positioned to represent 384 on a Chinese abacus.

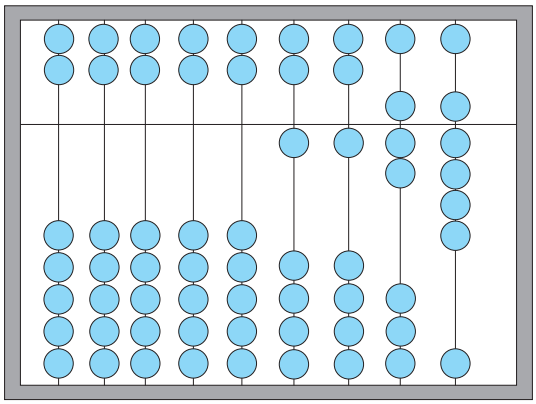

Fig. 2. Beads positioned to represent 384 + 795 = 1179 on a Chinese abacus.

Bibliography. American Printing House for the Blind, Inc., *Abacuses*, June 1999; C. B. Boyer (revised by U. C. Merzbach), *A History of Mathematics*, 3d ed., 1991; J. Dilson, *The Abacus*, St. Martin's Press, 1995.

Abalone

Common name in California for several species of marine gastropod mollusks in the family Haliotidae; also known as ear shells in many tropical and warm-temperate seas, as awabi in Japan, as paua in New Zealand, and as ormers in the Mediterranean region, Spain, and the English Channel.

The haliotid limpets constitute the most successful family of the primitive two-gilled gastropods included in the order Archaeogastropoda. They have a broad muscular sucker foot, an ear-shaped limpet shell (up to 30 cm or 12 in. long in some species) with a row of about seven shell apertures (see **illus.**), and a pair of featherlike gills (aspidobranch ctenidia). They live on rock surfaces in the low intertidal and sublittoral, and are grazing microherbivores, scraping and ingesting algal films with the relatively unspecialized dentition of a rhipidoglossan radula. Haliotids are among the relatively few living gastropods which are zygobranch (with two aspidobranch ctenidia) and diotocardiac (with two auricles in the heart). However, such paired symmetry of pallial, urinogenital, and cardiac structures is almost certainly characteristic of the earliest (ancestral) stocks of true gastropods. In the living forms, which include the rarer pleurotomariids and fissurellids as well as the haliotids, the symmetry of the organs requires a symmetrical respiratory flow through the branchial and pallial structures. The inhalant water current comes in to the mantle cavity from both sides ventrally, and the exhalant current passes out centrally and dorsally. In all of these living zygobranch gastropods

Shell of a typical haliotid limpet, or abalone, showing the characteristic curved series of exhalant apertures.

(and all the more numerous fossil forms) there are shell slits or complex apertures in the shell to ensure that the exhalant current, bearing with it both feces and urinogenital products, does not discharge directly over the head. In the haliotids (see illus.) these form a curved series of six or seven exhalant apertures, maintained during shell growth by the oldest (most apical) aperture being filled from below with shell material while a new aperture is formed initially as an indentation in the margin of the shell. Under the penultimate functioning aperture lies the anus, and thus feces are flushed out of the mantle cavity dorsally and well away from the mouth and sense organs. The sexes are separate, with broadcast spawning (again through the older apertures) and external fertilization. The planktonic larvae are typical veligers.

In all regions where they occur, flesh of the massive foot of these haliotids is prized as food, and the flattened ear-shaped shells, with characteristically green-blue iridescent inner surfaces (termed mother-of-pearl), have been used over many centuries in decorative arts by Maoris, Japanese, Amerindians (Wakashans), and Byzantine Greeks. Collected almost to extinction in some regions for their combined shell beauty and gourmet value, abalones and other haliotids, where protected, have recovered to become again very abundant. *See* GASTROPODA; LIMPET; PROSOBRANCHIA. W. D. Russell-Hunter

ABC lipid transporters

A family of adenosine triphosphate–binding cassette (ABC) transmembrane proteins that use energy to transport various molecules across extracellular and intracellular membranes (cytoplasmic membranes, endoplasmic reticulum, mitochondria, or peroxisomes). ABC lipid transporters consist of two hydrophobic transmembrane domains and two hydrophilic nucleotide-binding folds. These binding folds contain highly conserved sequence motifs (Walker A and B) and are separated by a linker sequence, also known as the signature (C) motif. ABC transporters are organized as full transporters or half transporters depending on the number of transmembrane domains and nucleotide-binding folds. Half transporters must form homodimers or heterodimers to be functional.

ABC transporters are grouped into seven major subfamilies named ABCA to ABCG. The transmembrane domains of the ABC proteins form a chamber within the membrane and act as a "flippase" that gives inner leaflet substrates access to their binding sites. The substrates transported by the ABC proteins can vary from small ions to large polypeptides and polysaccharides.

A great number of the ABC transporters are involved with cellular lipid transport and homeostasis, and the defect or absence of these transporters is often correlated with pathologic phenotypes. The **table** illustrates the major ABC transporters currently associated with lipid transport or homeostasis.

ABCA1. ABCA1 (ABC1; CERP, cholesterol efflux regulatory protein) is a member of the subfamily of ABC full-size transporters, containing two highly conserved ATP-binding cassettes (catalytic regions) and two transmembrane domains. ABCA1 is involved in the formation of high-density lipoprotein (HDL) and the cellular efflux of excess cholesterol and phospholipids. The 149-kb gene is located on the human chromosome 9q22-q31. It comprises 50 exons and codes for a calculated 220-kDa protein that is ubiquitously expressed, with highest expression in placenta, liver, lung, adrenal glands, and testes, and modest expression in small intestine, lung, and adipose. The expression of ABCA1 is highly regulated on a transcriptional as well as posttranslational level by several agents, including cholesterol, fatty acids, liver X receptor agonists, cAMP [cyclic adenosine monophosphate], and cytokines.

ABCA1 has been identified as the molecular defect in Tangier disease, a rare genetic disorder characterized by nearly complete HDL deficiency, high plasma triglyceride levels, orange tonsils, and increased susceptibility to atherosclerosis. Studies in ABCA1 knockout and transgenic mouse models as well as in-vitro studies have led to major advances in our understanding of the ABCA1-mediated formation of plasma HDL and its role in reverse cholesterol transport. Such transport is considered an antiatherogenic process, in which excess cholesterol is transported from peripheral tissues to the liver in the form of HDL, where it can be excreted from the body directly through the bile or indirectly after conversion into bile acids.

ABCA1 is responsible for the initial lipidation of apoA-I, the major apoprotein found on HDL. These nascent or pre-β HDL particles function as acceptors for further lipidation with cholesterol and phospholipids from peripheral tissues, such as lipid-laden foam cells found in atherosclerotic lesions. ABCA1 also interacts with serum amyloid A to form large, dense HDL particles associated with pathological conditions such as acute chronic inflammation.

Major ABC transporters currently associated with lipid transport or homeostasis

Transporter	Function	Disease
ABCA1	Cholesterol and phospholipid transport	Tangler disease
ABCA3	Surfactant metabolism	Neonatal surfactant deficiency
ABCA7	Cholesterol and phospholipid transport	Sjögren's syndrome
ABCG1	Cholesterol and phospholipid transport	Unknown
ABCG2	Drug resistance	Unknown
ABCG4	Cholesterol transport	Unknown
ABCG5	Sterol transport	Sitosterolemia
ABCG8	Sterol transport	Sitosterolemia

Plasma HDL concentrations are regulated primarily by liver-derived (~70%) and intestine-derived (~30%) ABCA1 activity and are inversely correlated with coronary artery disease. Complete ABCA1 deficiency in mice did not alter atherosclerosis, but repopulation of ABCA1-deficient mice with wild-type or ABCA1-overexpressing macrophages led to significantly decreased atherosclerosis. Accordingly, selective inactivation of macrophage ABCA1 in mice results in markedly increased atherosclerosis. However, overexpression of ABCA1 in other tissues, including liver, increased plasma cholesterol concentrations of both the antiatherogenic HDL and the proatherogenic non-HDL, and resulted in both pro- and antiatherogenic effects in different mouse models. Thus, changes in ABCA1 expression alter atherogenesis in a tissue-specific way. Based on these results, upregulation of ABCA1 for therapeutic use may be less desirable, or it may have to target specific tissues to achieve an overall antiatherogenic effect.

ABCA3. The ABCA3 gene maps to chromosome 16p13.3, consists of a 1704-amino-acid polypeptide, and is expressed predominantly in lung. ABCA3 gene mutations cause a fatal deficiency of surfactant in newborns. Pulmonary surfactant is mostly composed of lipids that are densely organized into multilamellar structures. These structures are critical to the transition of the lung from an aqueous to an air environment at birth. In adult lung, surfactants are also critical in preserving its homeostasis. Cellular electron micrographs of ABCA3-deficient human lung cells showed lipid accumulation in the multilamellar vesicles, suggesting that ABCA3 plays an important role in the formation of pulmonary surfactant, probably by transporting phospholipids and cholesterol.

ABCA7. ABCA7 is a highly expressed transporter in spleen, thymus, lymphoid cells, bone marrow, brain, and trachea. It is mapped to chromosome 19 and has 46 introns, and it is a full-size transporter. ABCA7 is highly homologous to ABCA1 and induces an apolipoprotein-mediated assembly of cholesterol similar to that induced by ABCA1, releasing cholesterol and phospholipids to HDL in an apoA-I–dependent mechanism. In mice, it was reported to promote the apoA-I–dependent release of phospholipids, but not cholesterol. Another observation in ABCA7–deficient females, but not male mice, reported less visceral fat and lower total serum and HDL cholesterol levels than wild-type, gender-matched littermates. Also, in these same mice, the lack of ABCA7 did not alter cholesterol and phospholipid efflux.

ABCG1. ABCG1 (ABC8, or human *white* gene) is a member of the subfamily of ABC half-size transporters involved in intracellular lipid transport. In order to create a functional full-size transporter, ABCG1 can homodimerize and most probably can also heterodimerize with other half-size transporters. The human ABCG1 spans ~98 kb and comprises 23 exons, located on chromosome 21q22.3. ABCG1 is expressed in most tissues in mice and humans, with high baseline expression in mouse macrophages, lung, brain, spleen, and other tissues containing macrophages, while in humans the expression is highest in lung, adrenal glands, spleen, heart, placenta, and liver. Several isoforms and splice variants have been found on the RNA level, yet the different protein levels and their functionality remain to be established. Expression of ABCG1 is strongly induced in many tissues by cholesterol loading, by ligands for the nuclear receptors LXRα and β, and by PPARγ-ligands.

ABCG1 is responsible for the control of tissue lipid levels and mediates cholesterol and phospholipid efflux to mature HDL or other phospholipid-containing acceptors, particularly in macrophages under conditions of cholesterol excess. The targeted disruption of ABCG1 in mice on a high-fat/high-cholesterol diet resulted in massive accumulation of neutral lipids and phospholipids in hepatocytes and tissue macrophages; however, no changes in plasma lipids accompanied these findings. In accordance with the regulation of cellular lipid homeostasis, the tissues in mice overexpressing human ABCG1 were protected from diet-induced lipid accumulation.

ABCG1 was found to be localized perinuclear as well as at the plasma membrane of cells. One potential mechanism for the ABCG1-facilitated lipid transfer is by redistributing membrane cholesterol to cell-surface domains that are accessible to removal by HDL. The role of ABCG1 in cholesterol export of lipid-laden cells seems to complement the action of ABCA1, a full-size transporter with similar tissue expression pattern, which mediates the initial lipidation in the formation of HDL. The nascent HDL particles are substrates for ABCG1-mediated lipid efflux, making ABCG1 part of the reverse cholesterol transport, a protective mechanism against the critical step of macrophage lipid overloading in atherogenesis.

ABCG2. Human ABCG2 is also referred to as placenta-specific ABC transporter, breast cancer

resistance protein, or mitoxantrone resistance-associated protein. (Mitoxantrone is a chemotherapy drug used to treat certain cancers.) It was initially identified as a messenger RNA expressed in placenta. The human ABCG2 gene is located on chromosome 4q22 and encodes a half transporter. In normal human tissues, ABCG2 has been found to be expressed on the apical membrane of trophoblast cells in placenta, the apical membrane of enterocytes, the bile canalicular membrane of hepatocytes, and the apical membrane of lactiferous ducts in the mammary gland. In addition, ABCG2 has been demonstrated to be expressed in a wide variety of hematopoietic cells. It is expected that ABCG2 homodimerizes to acquire transporter activity.

ABCG2 transporter functions as a xenobiotic transporter which may play a major role in multidrug resistance. It serves as a cellular defense mechanism in response to mitoxantrone and anthracycline (a chemotherapeutic agent) exposure. Although the function of ABCG2 has been studied extensively in terms of multidrug resistance, the physiological and/or pharmacological functions of this transporter have not been clarified yet. ABCG2 can also transport several dyes, such as rhodamine 123 and Hoechst 33462. Recently, it has been shown to transport steroids such as cholesterol, estradiol, progesterone, and testosterone. It also transports sulfated conjugates of steroids (esterone 3-sulfate, dehydroepiandrosterone sulfate) and certain chlorophyll metabolites.

ABCG4. The ABCG4 gene consists of 14 exons that span 12.6 kb on chromosome 11. ABCG4 gene sequence, protein, regulation, and function are very similar to ABCG1. The ABCG4 protein contains 646 amino acids and shares 72% amino acid sequence identity with ABCG1. However, the tissue distribution of ABCG1 and ABCG4 is not identical: ABCG1 is present in many tissues, including macrophages, lung, brain, liver, and spleen, while ABCG4 is mainly expressed in the brain and spleen. ABCG4 expression in the brain has been suggested to promote cholesterol efflux to HDL-like particles present in cerebrospinal fluids. Low level of ABCG4 has also been found in mouse macrophages, and its expression, similarly to ABCG1, was also upregulated by oxysterols (22-OH cholesterol), retinoids (9-*cis*-retinoic acid), and a (liver X receptor)-specific agonist (T0901713). Because of the different pattern of expression, it seems very unlikely that ABCG1 and ABCG4 act as heterodimers. Instead, evidence that ABCG4 acts as homodimer comes from studies that demonstrate that overexpression of ABCG4 stimulated cholesterol efflux to HDL.

Both ABCG4 and ABCG1 play a major role in the HDL-mediated cholesterol efflux in liver X receptor-induced macrophages. In-vitro studies showed that ABCG4 expression alone increased cholesterol efflux to both HDL-2 and HDL-3, and when both ABCG4 and ABCG1 were overexpressed, the increase in cholesterol efflux was even higher. ABCG4-mediated cholesterol efflux to LDL and cyclodextrin was also observed, but in lesser degree compared with the HDL efflux. ABCG4 suppression reduced the cholesterol efflux to HDL, but also decreased ABCG1 messenger RNA expression, probably due to homology in their sequences. These data indicate that ABCG4 alone or in association with ABCG1 may modulate the intracellular concentration of cholesterol by promoting cholesterol efflux to HDL.

ABCG5 and ABCG8. ABCG5 and ABCG8 are two half transporters that belong to the ABCG subfamily. They were identified as proteins mutated in sitosterolemia, a rare genetic disorder characterized by increased plasma and tissue levels of plant sterols. Patients with sitosterolemia have an increased absorption of dietary sterols and an impaired ability to secrete cholesterol and plant-derived sterols into bile. Many sitosterolemic individuals deposit cholesterol and plant sterols in the skin, tendons, and coronary arteries and suffer premature coronary artery disease.

ABCG5 and ABCG8 genes contain 13 exons and are organized in a head-to-head orientation on chromosome 2 with only 140 base pairs separating their two respective start-transcription sites. In-vivo and in-vitro data indicate that ABCG5 and ABCG8 function as obligate heterodimers to promote sterol excretion into bile. The major regulation of their expression appears to be at the transcriptional level: cholesterol feeding activates the liver X receptor, which increases both ABCG5 and ABCG8 messenger RNA expression. ABCG5/G8 are expressed mainly at the apical surface of hepatocytes and enterocytes, where they limit the accumulation of sterols by promoting sterol excretion into bile and reducing dietary sterol uptake. Overexpression of human ABCG5 and ABCG8 in wild-type mice led to increased cholesterol secretion into bile, and the knockouts of either ABCG5/G8 or ABCG8 alone prevented sterol secretion into bile. However, no changes in liver or plasma cholesterol levels were found. No protection from atherosclerosis was observed in these studies. However, in LDL receptor-deficient mice, which are prone to develop atherosclerosis on a high-fat/high-cholesterol diet, ABCG5/G8 overexpression decreased the absorption of dietary cholesterol, increased the excretion of cholesterol into bile and stool, and decreased cholesterol levels in plasma, leading to a significant reduction in the development of atherosclerotic lesions. *See* ADENOSINE TRIPHOSPHATE (ATP); ATHEROSCLEROSIS; CHOLESTEROL; GENE; LIPID METABOLISM; LIPOPROTEIN; LIVER; PROTEIN.

E. M. Wagner; F. Basso; C. S. Kim; M. J. A. Amar

Bibliography. H. B. Brewer, Jr., et al., Regulation of plasma high-density lipoprotein levels by the ABCA1 transporter and the emerging role of high-density lipoprotein in the treatment of cardiovascular disease, *Arterioscler. Thromb. Vasc. Biol.*, 24:1755–1760, 2004; M. Dean, The Human ATP-Binding Cassette (ABC) Transporter Superfamily, http://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=&rid=mono_001.TOC&depth=2; T. Janvilisri et al., Sterol transport by the human breast cancer resistance protein (ABCG2) expressed in *Lactococcus*

lactis, *J. Biol. Chem.*, 278:20645–20651, 2003; W. Jessup et al., Roles of ATP binding cassette transporters A1 and G1, scavenger receptor BI and membrane lipid domains in cholesterol export from macrophages. *Curr. Opin. Lipidol.*, 17:247–257, 2006; J. Santamarina-Fojo et al., Complete genomic sequence of the human ABCA1 gene: Analysis of the human and mouse ATP-binding cassette A promoter, *Proc. Nat. Acad. Sci. USA*, 97:7987–7992, 2000; K. Takahashi et al., ABC proteins: Key molecules for lipid homeostasis, *Med. Mol. Morphol.*, 38:2–12, 2005; N. Wang et al., ATP-binding cassette transporters G1 and G4 mediate cellular cholesterol efflux to high-density lipoproteins, *Proc. Nat. Acad. Sci. USA*, 101:9774–9779, 2004; K. R. Wilund et al., High-level expression of ABCG5 and ABCG8 attenuates diet-induced hypercholesterolemia and atherosclerosis in Ldlr −/− mice, *J. Lipid Res.*, 45:1429–1436, 2004.

Abdomen

A major body division of the vertebrate trunk lying posterior to the thorax; and in mammals, bounded anteriorly by the diaphragm and extending to the pelvis. The diaphragm, found only in mammals, separates the abdominal or peritoneal cavity from the pleural and pericardial cavities of the thorax. In all pulmonate vertebrates (possessing lungs or lunglike organs) other than mammals, the lungs lie in the same cavity with the abdominal viscera, and this cavity is known as the pleuroperitoneal cavity.

The large coelomic cavity that occupies the abdomen contains the viscera, such as the stomach, liver, gallbladder, spleen, pancreas, and intestinal tract with their associated nerves and blood vessels. The viscera are contained in the peritoneal sac. This structure is adherent to the abdominal wall as the parietal peritoneum and encompasses all parts of the viscera as an outermost coat, the serosa or visceral peritoneum. Connecting sheets of peritoneum from the body wall to the various organs form the mesenteries, which are always double-walled; blood vessels, lymphatics, nerves, and lymph nodes are located between the two walls of the mesenteries. Other folds of the peritoneum form the omenta. At the lesser curvature of the stomach the serosa continues upward to the liver as the lesser omentum, and from the greater curvature it hangs down as a sheet in front of the intestine, forming the greater omentum.

The term abdomen is also applied to a similar major body division of arthropods and other animals. *See* ANATOMY, REGIONAL.

Walter Bock

Aberration (optics)

A departure of an optical image-forming system from ideal behavior. Ideally, such a system will produce a unique image point corresponding to each object point. In addition, every straight line in the object space will have as its corresponding image a unique straight line. A similar one-to-one correspondence will exist between planes in the two spaces.

This type of mapping of object space into image space is called a collinear transformation. A paraxial ray trace is used to determine the parameters of the transformation, as well as to locate the ideal image points in the ideal image plane for the system. *See* GEOMETRICAL OPTICS; OPTICAL IMAGE.

When the conditions for a collinear transformation are not met, the departures from that ideal behavior are termed aberrations. They are classified into two general types, monochromatic aberrations and chromatic aberrations. The monochromatic aberrations apply to a single color, or wavelength, of light. The chromatic aberrations are simply the chromatic variation, or variation with wavelength, of the monochromatic aberrations. *See* CHROMATIC ABERRATION.

Aberration measures. The monochromatic aberrations can be described in several ways. Wave aberrations are departures of the geometrical wavefront from a reference sphere with its vertex at the center of the exit pupil and its center of curvature located at the ideal image point. The wave aberration is measured along the ray and is a function of the field height and the pupil coordinates of the reference sphere (**Fig. 1**).

Transverse ray aberrations are measured by the transverse displacement from the ideal image point to the ray intersection with the ideal image plane.

Some aberrations are also measured by the longitudinal aberration, which is the displacement along the chief ray of the ray intersection with it. The use

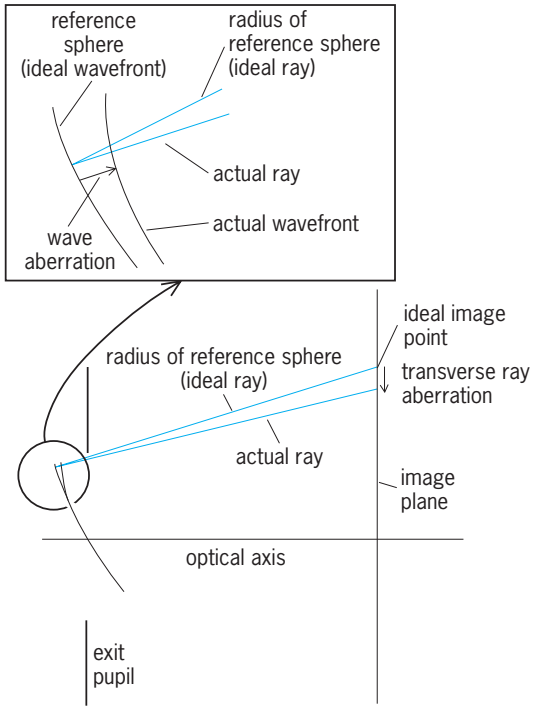

Fig. 1. Diagram of the image space of an optical system, showing aberration measures: the wave aberration and the transverse ray aberration.

of this measure will become clear in the discussion of aberration types.

Caustics. Another aberrational feature which occurs when the wavefront in the exit pupil is not spherical is the development of a caustic. For a spherical wavefront, the curvature of the wavefront is constant everywhere, and the center of curvature along each ray is at the center of curvature of the wavefront. If the wavefront is not spherical, the curvature is not constant everywhere, and at each point on the wavefront the curvature will be a function of orientation on the surface as well as position of the point. As a function of orientation, the curvature will fluctuate between two extreme values. These two extreme values are called the principal curvatures of the wavefront at that point. The principal centers of curvature lie on the ray, which is normal to the wavefront, and will be at different locations on the ray.

The caustic refers to the surfaces which contain the principal centers of curvature for the entire wavefront. It consists of two sheets, one for each of the two principal centers of curvature. For a given type of aberration, one or both sheets may degenerate to a line segment. Otherwise they will be surfaces.

The feature which is of greatest interest to the user of the optical system is usually the appearance of the image. If a relatively large number of rays from the same object point and with uniform distribution over the pupil are traced through an optical system, a plot of their intersections with the image plane represents the geometrical image. Such a plot is called a spot diagram, and a set of these is often plotted in planes through focus as well as across the field.

Orders of Aberrations

The monochromatic aberrations can be decomposed into a series of aberration terms which are ordered according to the degree of dependence they have on the variables, namely, the field height and the pupil coordinates. Each order is determined by the sum of the powers to which the variables are raised in describing the aberration terms. Because of axial symmetry, alternate orders of aberrations are missing from the set. For wave aberrations, odd orders are missing, whereas for transverse ray aberrations, even orders are missing. It is customary in the United States to specify the orders according to the transverse ray designation.

Monochromatically, first-order aberrations are identically zero, because they would represent errors in image plane location and magnification, and if the first-order calculation has been properly carried out, these errors do not exist. In any case, they do not destroy the collinear relationship between the object and the image. The lowest order of significance in describing monochromatic aberrations is the third order.

Chromatic variations of the first-order properties of the system, however, are not identically zero. They can be determined by first-order ray tracing for the different colors (wavelengths), where the refractive

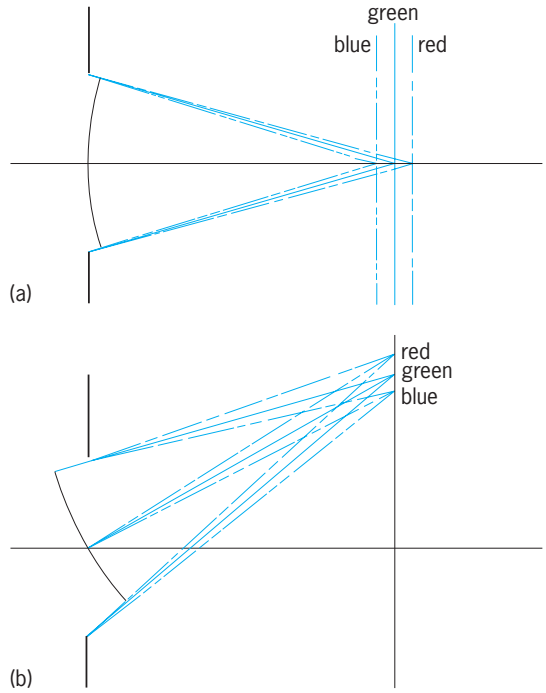

Fig. 2. Chromatic aberration. (*a*) Longitudinal chromatic aberration. (*b*) Transverse chromatic aberration.

indices of the media change with color. For a given object plane, the different colors may have conjugate image planes which are separated axially. This is called longitudinal chromatic aberration (**Fig. 2***a*). Moreover, for a given point off axis, the conjugate image points may be separated transversely for the different colors. This is called transverse chromatic aberration (Fig. 2*b*), or sometimes chromatic difference of magnification.

These first-order chromatic aberrations are usually associated with the third-order monochromatic aberrations because they are each the lowest order of aberration of their type requiring correction.

The third-order monochromatic aberrations can be divided into two types, those in which the image of a point source remains a point image but the location of the image is in error, and those in which the point image itself is aberrated. Both can coexist, of course.

Aberrations of geometry. The first type, the aberrations of geometry, consist of field curvature and distortion.

Field curvature. Field curvature is an aberration in which there is a focal shift which varies as a quadratic function of field height, resulting in the in-focus images lying on a curved surface. If this aberration alone were present, the images would be of good quality on this curved surface, but the collinear condition of plane-to-plane correspondence would not be satisfied.

Distortion. Distortion, on the other hand, is an aberration in which the images lie in a plane, but they are displaced radially from their ideal positions in the image plane, and this displacement is a cubic function of the field height. This means that any straight

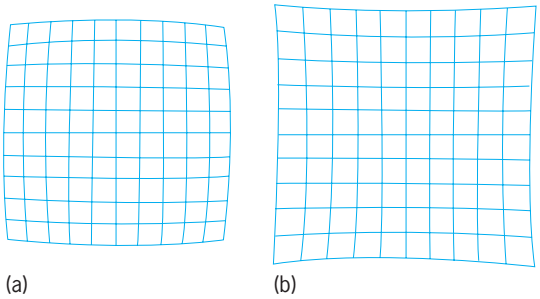

Fig. 3. Distortion. (*a*) Barrel distortion. (*b*) Pincushion distortion.

line in the object plane not passing through the center of the field will have an image which is a curved line, thereby violating the condition of straight-line to straight-line correspondence. For example, if the object is a square centered in the field, the points at the corners of the image are disproportionally displaced from their ideal positions in comparison with the midpoints of the sides. If the displacements are toward the center of the field, the sides of the figure are convex; this is called barrel distortion (**Fig. 3***a*). If the displacements are away from the center, the sides of the figure are concave; this is called pincushion distortion (Fig. 3*b*).

Aberrations of point images. There are three third-order aberrations in which the point images themselves are aberrated: spherical aberration, coma, and astigmatism.

Spherical aberration. Spherical aberration is constant over the field. It is the only monochromatic aberration which does not vanish on axis, and it is the axial case which is easiest to understand.

The wave aberration function (**Fig. 4***a*) is a figure of revolution which varies as the fourth power of the radius in the pupil. The wavefront itself has this wave aberration function added to the reference sphere centered on the ideal image. The rays (Fig. 4*b*) from any circular zone in the wavefront come to a common zonal focus on the axis, but the position of this focus shifts axially as the zone radius increases. This zonal focal shift increases quadratically with the zone radius to a maximum from the rays from the marginal zone. This axial shift is longitudinal spherical aberration. The magnitude of the spherical aberration can be measured by the distance from the paraxial focus to the marginal focus.

The principal curvatures of any point on the wavefront are oriented tangential and perpendicular to the zone containing the point. The curvatures which are oriented tangential to the zone have their centers on the axis, so the caustic sheet for these consists of the straight line segment extending from the paraxial focus, and is degenerate. The other set of principal centers of curvature lie on a trumpet-shaped surface concentric with the axis (**Fig. 5***a*). This second sheet is the envelope of the rays. They are all tangent to it, and the point of tangency for each ray is at the second principal center of curvature for the ray (Fig. 5*b*).

It is clear that the image formed in the presence of spherical aberration (Fig. 4*c*) does not have a well-defined focus, although the concentration of light outside the caustic region is everywhere worse than it is inside. Moreover, the light distribution in the image is asymmetric with respect to focal position, so the precise selection of the best focus depends on the criterion chosen. The smallest circle which can contain all the rays occurs one-quarter of the distance from the marginal focus to the paraxial focus, the image which has the smallest second moment lies one-third of the way from the marginal focus to the paraxial focus, and the image for which the variance of the wave aberration function is a minimum lies halfway between the marginal and paraxial foci.

Coma. Coma is an aberration which varies as a linear function of field height. It can exist only for off-axis field heights, and as is true for all the aberrations, it is symmetrical with respect to the meridional plane containing the ideal image point. Each zone in its wave aberration function (**Fig. 6***a*) is a circle, but each circle is tilted about an axis perpendicular to the meridional plane, the magnitude of the tilt increasing with the cube of the radius of the zone.

The chief ray (Fig. 6*b*) passes through the ideal

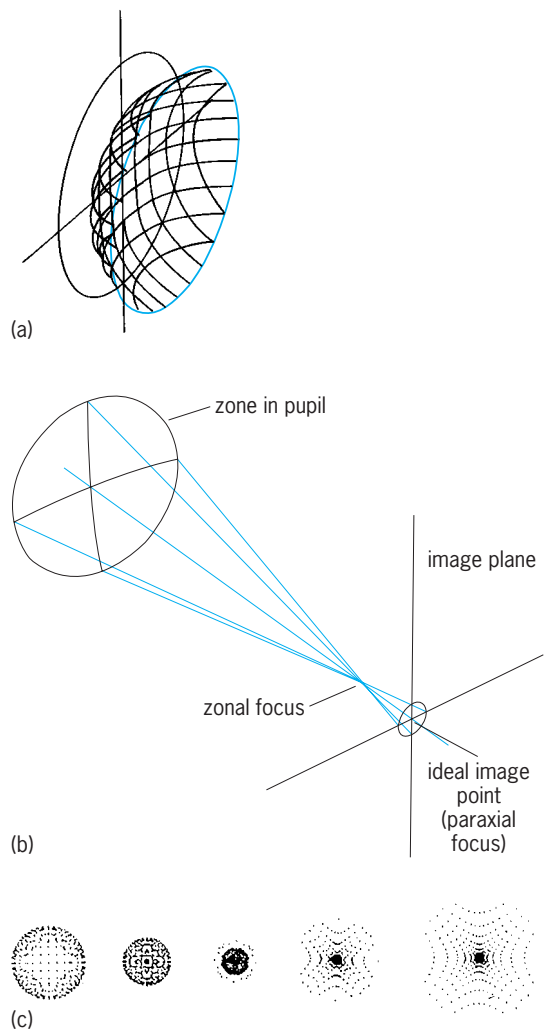

Fig. 4. System with spherical aberration. (*a*) Wave aberration function. (*b*) Rays. (*c*) Spot diagrams through foci showing transverse ray aberration patterns for a square grid of rays in the exit pupil.

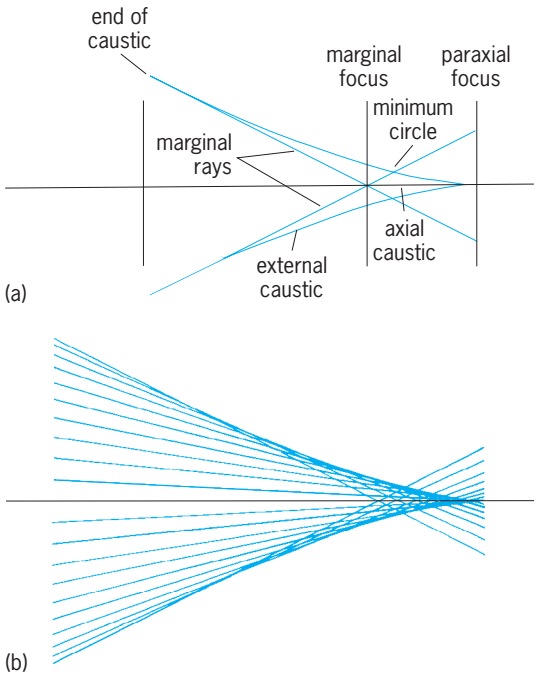

Fig. 5. Caustics in a system with spherical aberration. (*a*) Diagram of caustics and other major features. (*b*) Rays whose envelope forms an external caustic.

image point, but the rays from any zone intersect the image plane in a circle, the center of which is displaced from the ideal image point by an amount equal to the diameter of the circle. The diameter increases as the cube of the radius of the corresponding zone in the pupil. The circles for the various zones are all tangent to two straight lines intersecting at the ideal image point and making a 60° angle with each other. The resulting figure (Fig. 6*c*) resembles an arrowhead which points toward or away from the center of the field, depending on the sign of the aberration.

The upper and lower marginal rays in the meridional plane intersect each other at one point in the circle for the marginal zone. This point is the one most distant from the chief ray intersection with the image plane. The transverse distance between these points is a measure of the magnitude of the coma.

Astigmatism. Astigmatism is an aberration which varies as the square of the field height. The wave aberration function (**Fig. 7***a*) is a quadratic cylinder which varies only in the direction of the meridional plane and is constant in the direction perpendicular to the meridional plane. When the wave aberration function is added to the reference sphere, it is clear that the principal curvatures for any point in the wavefront are oriented perpendicular and parallel to the meridional plane, and moreover, although they are different from each other, each type is constant over the wavefront. Therefore, the caustic sheets both degenerate to lines perpendicular to and in the meridional plane in the image region, but the two lines are separated along the chief ray.

All of the rays must pass through both of these lines, so they are identified as the astigmatic foci. The astigmatic focus which is in the meridional plane is called the sagittal focus, and the one perpendicular to the meridional plane is called the tangential focus.

For a given zone in the pupil, all of the rays (Fig. 7*b*) will of course pass through the two astigmatic foci, but in between they will intersect an image plane in an ellipse, and halfway between the foci they will describe a circle (Fig. 7*c*). Thus, only halfway between the two astigmatic foci will the image be isotropic. It is also here that the second moment is a minimum, and the wave aberration variance is a minimum as well. This image is called the medial image.

Since astigmatism varies as the square of the field height, the separation of the foci varies as the square of the field height as well. Thus, even if one set of foci, say the sagittal, lies in a plane, the medial and tangential foci will lie on curved surfaces. If the field curvature is also present, all three lie on curved surfaces. The longitudinal distance along the chief ray from the sagittal focus to the tangential focus is a measure of the astigmatism.

The above description of the third-order aberrations applies to each in the absence of the other aberrations. In general, more than one aberration will be present, so that the situation is more complicated. The types of symmetry appropriate to each aberration will disclose its presence in the image.

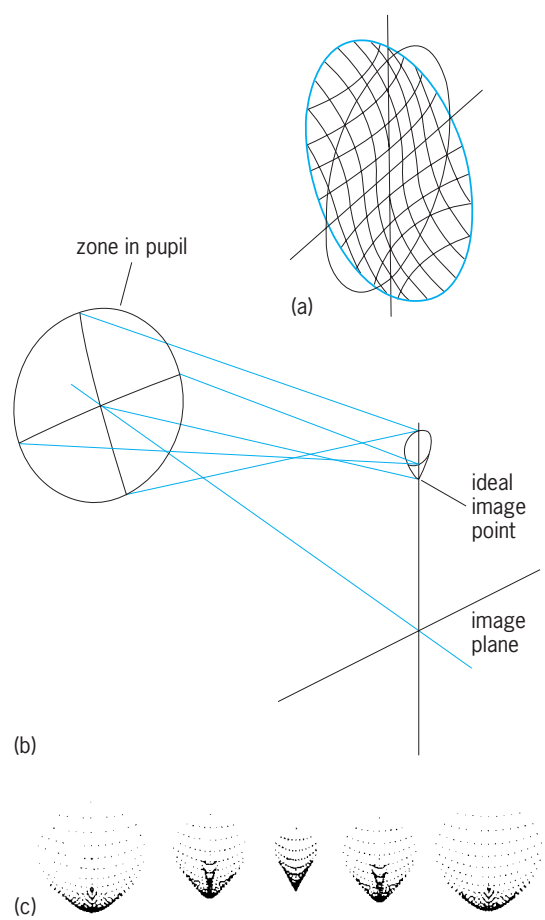

Fig. 6. System with coma. (*a*) Wave aberration function. (*b*) Rays. (*c*) Spot diagrams through foci, showing transverse ray aberration patterns for a square grid of rays in the exit pupil.

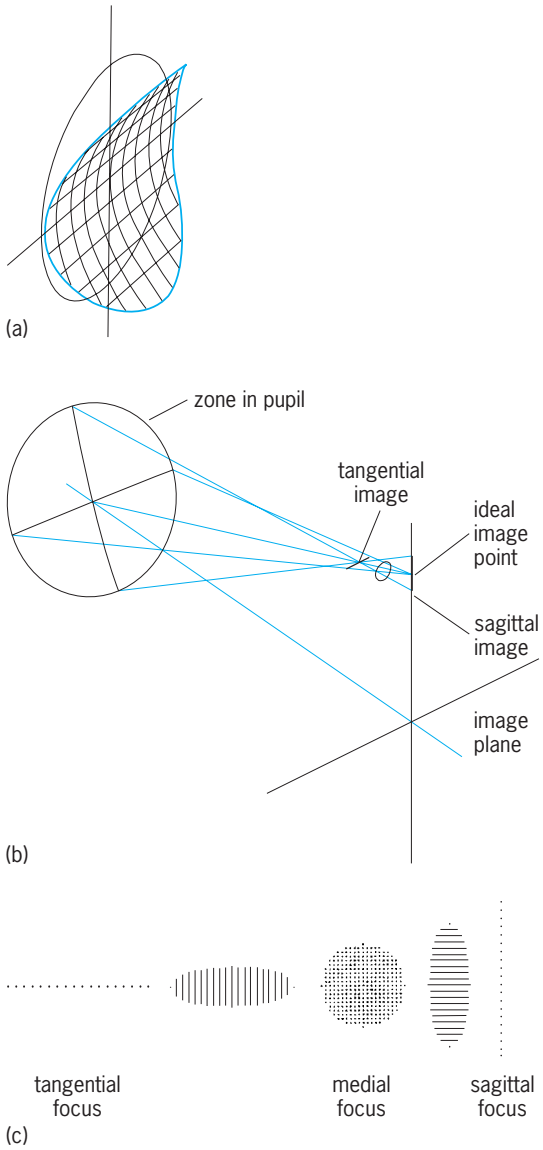

Fig. 7. System with astigmatism. (*a*) Wave aberration function. (*b*) Rays. (*c*) Spot diagrams through foci, showing transverse aberration patterns for a square grid of rays in the exit pupil.

Higher-order aberrations. The next order of aberration for the chromatic aberrations consists of the chromatic variation of the third-order aberrations. Some of these have been given their own names; for example, the chromatic variation of spherical aberration is called spherochromatism.

Monochromatic aberrations of the next order are called fifth-order aberrations. Most of the terms are similar to the third-order aberrations, but with a higher power dependence on the field or on the aperture. Field curvature, distortion, and astigmatism have a higher power dependence on the field, whereas spherical aberration and coma have a higher power dependence on the aperture. In addition, there are two new aberration types, called oblique spherical aberration and elliptical coma. These are not directly related to the third-order terms.

Expansions beyond the fifth order are seldom used, although in principle they are available. In fact, many optical designers use the third order as a guide in developing the early stages of a design, and then go directly to real ray tracing, using the transverse ray aberrations of real rays without decomposition to orders. However, the insights gained by using the fifth-order aberrations can be very useful.

Origin of Aberrations

Each surface in an optical system introduces aberrations as the beam passes through the system. The aberrations of the entire system consist of the sum of the surface contributions, some of which may be positive and others negative. The challenge of optical design is to balance these contributions so that the total aberrations of the system are tolerably small. In a well-corrected system the individual surface contributions are many times larger than the tolerance value, so that the balance is rather delicate, and the optical system must be made with a high degree of precision.

Insight as to where the aberrations come from can be gained by considering how the aberrations are generated at a single refracting surface. Although the center of curvature of a spherical surface lies on the optical axis of the system, it does not in itself have an axis. If, for the moment, the spherical surface is assumed to be complete, and the fact that the entrance pupil for the surface will limit the beam incident on it is ignored, then for every object point there is a local axis which is the line connecting the object point with the center of curvature. All possible rays from the object point which can be refracted by the surface will be symmetrically disposed about this local axis, and the image will in general suffer from spherical aberration referred to this local axis.

A small pencil of rays about this axis (locally paraxial) will form a first-order image according to the rules of paraxial ray tracing. If the first-order imagery of all the points lying in the object plane is treated in this manner, it is found that the surface containing the images is a curved surface. The ideal image surface is a plane passing through the image on the optical axis of the system, and thus the refracting surface introduces field curvature. The curvature of this field is called the Petzval curvature.

In addition to this monochromatic behavior of the refracting surface, the variation in the index of refraction with color will introduce changes in both the first-order imagery and the spherical aberration. This is where the chromatic aberrations come from.

Thus there are fundamentally only three processes operating in the creation of aberrations by a spherical refracting surface. These result in spherical aberration, field curvature, and longitudinal chromatic aberration referred to the local axis of each image. In fact, if the entrance pupil for the surface is at the center of curvature, these will be the only aberrations that are contributed to the system by this refracting surface.

In general, the pupil will not be located at the center of curvature of the surface. For any off-axis object point, the ray which passes through the center of the

pupil will be the chief ray, and it will not coincide with the local axis. The size of the pupil will limit the beam which can actually pass through the surface, and the beam will therefore be an eccentric portion of the otherwise spherically aberrated beam.

Since an aberration expansion decomposes the wave aberration function about an origin located by the chief ray, the eccentric and asymmetric portion of an otherwise purely spherically aberrated wave gives rise to the field-dependent aberrations, because the eccentricity is proportional to the field height of the object point.

In this manner the aberrations which arise at each surface in the optical system, and therefore the total aberrations of the system, can be accounted for. *See* LENS (OPTICS); OPTICAL SURFACES. Roland V. Shack

Bibliography. M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; B. D. Guenther, *Modern Optics*, 1990; R. Kingslake, *Optical System Design*, 1983; K. Narigi and Y. Matsui, *Fundamentals of Practical Aberration Theory*, 1993; W. Smith, *Modern Optical Engineering*, 2d ed., 1990; W. T. Welford, *Aberrations of Optical Systems*, 1986.

Aberration of light

The apparent change in direction of a source of light caused by an observer's component of motion perpendicular to the impinging rays.

To visualize the effect, first imagine a stationary telescope (**illus.** *a*) aimed at a luminous source such as a star, with photons traveling concentrically down the tube to an image at the center of the focal plane. Next give the telescope a component of motion perpendicular to the incoming rays (illus. *b*). Photons passing the objective require a finite time to travel the length of the tube. During this time the telescope has moved a short distance, causing the photons to reach a spot on the focal plane displaced from the former image position. To return the image to the center, the telescope must be tilted in the direction of motion by an amount sufficient to ensure that the photons once again come concentrically down the tube in its frame of reference (illus. *c*). The necessary tilt angle α is given by $\tan \alpha = v/c$, where v is the component of velocity perpendicular to the incoming light and c is the velocity of light. (An analogy illustrating aberration is the experience that, in order for the feet to remain dry while walking through vertically falling rain, it is necessary to tilt an umbrella substantially forward.)

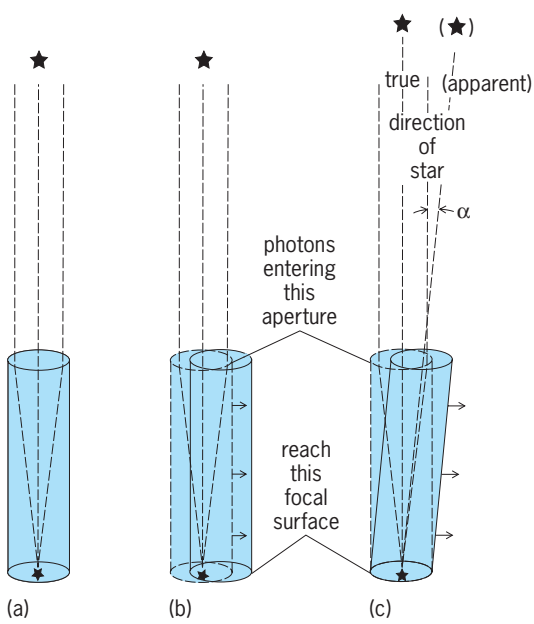

Demonstration of aberration. (*a*) Fixed telescope; photons form image at center of focal plane. (*b*) Moving telescope; image is displaced from center. (*c*) Tilted moving telescope is required to restore image to center.

Aberration of light was discovered by the English astronomer James Bradley in 1725 while searching unsuccessfully for parallax of nearby stars, using the Earth's orbital diameter as baseline. He found instead the necessity to compensate continuously for the Earth's velocity in its elliptical orbit, an effect about a hundred times greater than the parallax of typical nearby stars. This discovery provided the first direct physical confirmation of the Copernican theory.

Annual aberration causes stars to appear to describe ellipses whose major axes are the same length, the semimajor axis (usually expressed in arc-seconds) being the constant of aberration $\alpha = 20.49552''$. The minor axis of the ellipse depends on the star's ecliptic latitude β, and is given by $2\alpha \sin \beta$. The small diurnal aberration arising from the Earth's axial rotation depends on geographic latitude ϕ, and is given by $0.31'' \cos \phi$.

A second important application of aberration has been its clear-cut demonstration that, as is axiomatic to special relativity, light reaching the Earth has a velocity unaffected by the relative motion of the source toward or away from the Earth. This is shown by the fact that the aberration effect is the same for all celestial objects, including some quasars with apparent recessional velocities approaching the speed of light. *See* EARTH ROTATION AND ORBITAL MOTION; LIGHT; PARALLAX (ASTRONOMY); QUASAR; RELATIVITY. Harlan J. Smith

Abrasive

A material of extreme hardness that is used to shape other materials by a grinding or abrading action. Abrasive materials may be used either as loose grains, as grinding wheels, or as coatings on cloth or paper. They may be formed into ceramic cutting tools that are used for machining metal in the same way that ordinary machine tools are used. Because of their superior hardness and refractory properties, they have advantages in speed of operation, depth of cut, and smoothness of finish.

Abrasive products are used for cleaning and machining all types of metal, for grinding and polishing ceramics and glass, for grinding logs to paper pulp, for cutting metals, glass, and cement, and for

manufacturing many miscellaneous products such as brake linings and nonslip floor tile.

Abrasive materials. These may be classified in two groups, the natural and the synthetic (manufactured). The latter are by far the more extensively used, but in some specific applications natural materials still dominate.

The important natural abrasives are diamond (the hardest known material), corundum (a relatively pure, natural aluminum oxide, Al_2O_3), and emery (a less pure Al_2O_3 with considerable amounts of iron). The last of these has been replaced to a great extent by synthetic materials. Other natural abrasives are garnet, an aluminosilicate mineral; feldspar, used in household cleansers; calcined clay; lime; chalk; and silica, SiO_2, in its many forms—sandstone, sand (for grinding plate glass), flint, and diatomite.

The synthetic abrasive materials are silicon carbide, SiC; aluminum oxide, Al_2O_3; titanium carbide, TiC; and boron carbide, B_4C. The synthesis of diamond puts this material in the category of manufactured abrasives. There are other carbides, as well as nitrides and cermets, which can be classified as abrasives but their use is special and limited. The properties of hardness and high melting point are related, and many abrasive materials are also good refractories. *See* DIAMOND; REFRACTORY.

Silicon carbide is still made in much the same way that E. G. Acheson first made it in 1891. The principal ingredients are pure sand, SiO_2; coke (carbon); sawdust (to burn and provide vent holes for the escape of the gaseous products); and salt (to react with the impurities and form volatile compounds). The batch is placed in a troughlike furnace, up to 60 ft (18 m) long, with a graphite electrode at each end. The charge is heated by an electric current, a core of carbon or graphite being used to carry the current. Temperatures up to 4400°F (2400°C) are reached. The net reaction is $SiO_2 + 3C \rightarrow SiC + 2CO$, but the details of the reaction are complicated and not thoroughly understood. After the furnace has cooled, the sides are removed and the usable silicon carbide picked out, crushed, and sized.

Abrasive aluminum oxide is made from calcined bauxite (a mixture of aluminum hydrates) or purified alumina by fusion in an electric-arc furnace with a water-cooled metal shell; the solid alumina that is around the edge of the furnace acts as the refractory.

Various grades of each type of synthetic abrasive are distinguishable by properties such as color, toughness, and friability. These differences are caused by variation in purity of materials and methods of processing.

The sized abrasive may be used as loose grains, as coatings on paper or cloth to make sandpaper and emery cloth, or as grains for bonding into wheels.

Abrasive wheels. A variety of bonds are used in making abrasive wheels: vitrified or ceramic, essentially a glass or glass plus crystals; sodium silicate; rubber; resinoid; shellac; and oxychloride. Each type of bond has its advantages. The more rigid ceramic bond is better for precision-grinding operations, and the tougher, resilient bonds, such as resinoid or rubber, are better for snagging and cutting operations.

Ceramic-bonded wheels are made by mixing the graded abrasive and binder, pressing to general size and shape, firing, and truing or finishing by grinding to exact dimensions. A high-speed rotation test is given to check the soundness of the wheel.

Grinding wheels are specified by abrasive type, grain size (grit), grade or hardness, and bond type; in addition, the structure or porosity may be indicated. The term hardness as applied to a wheel refers to its behavior in use and not to the hardness of the abrasive material itself. The wheel is a three-component system of abrasive, bond, and air; the hardness is a complex function of the type and amount of bond and of the density of the wheel.

Literally thousands of types of wheels are made with different combinations of characteristics, not to mention the multitude of sizes and shapes available; therefore, selecting the best grinding wheel for a given job is not simple. When the abrasive grains of a wheel become slightly dulled by use, the stresses in the grinding operation should increase enough to tear the grain from the wheel to expose a new cutting grain. Thus, too soft a wheel wears too fast, losing grains before they are dulled, whereas too hard a wheel develops a smooth, glazed surface that will not cut. In either case, efficiency drops.

Hardness tests. The hardness of materials is roughly indicated by the Mohs scale, in which 10 minerals were selected, arranged in order of hardness, and numbered from 1 to 10 (softest to hardest); diamond has a hardness number of 10 and talc a hardness number of 1. However, since this scale is not quantitative (for example, aluminum oxide, 9 on the Mohs scale, is not 10% softer than diamond), and since most abrasive materials fall in the region at the top of the scale, other hardness scales have been developed. One is the Knoop indentation test, in which a diamond pyramid of specified shape is pressed under a definite load into the material to be tested; the size of the indentation is taken as a measure of the material's hardness. *See* HARDNESS SCALES.

Ceramic cutting tools are generally made of aluminum oxide by hot-pressing and grinding to final size. *See* CERAMICS. John F. McMahon

Bibliography. S. Krar and E. Ratterman, *Superabrasives: Grinding and Machining with Carbon and Diamond*,1990; J. Wang, *Abrasive Technology: Current Development and Applications I*,1999.

Abscisic acid

One of the five major plant hormones that play an important role in plant growth and development. Abscisic acid (ABA) is ubiquitous in higher plants; the highest levels are found in young leaves and developing fruits and seeds. Abscisic acid is also produced by certain algae and by several phytopathogenic fungi. It has been found in the brains of several mammals; however, animals probably do not produce abscisic acid but acquire it from plants in the diet.

Chemistry and biosynthesis. The chemical structure of (+)-abscisic acid is shown here; naturally

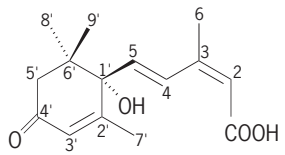

occurring abscisic acid is dextrorotary, hence the (+). It contains 15 carbon atoms and is a weak organic acid, as are two other plant hormones, auxin and gibberellin. *See* AUXIN; GIBBERELLIN.

Abscisic acid is a terpenoid. All chemicals belonging to this class of compounds have mevalonic acid as a common precursor and are built up of five carbon units (isoprenoid units). Other examples of terpenoids in plants are essential oils, gibberellins, steroids, carotenoids, and natural rubber. *See* CAROTENOID; ESSENTIAL OILS; RUBBER; STEROID.

In fungi, abscisic acid is synthesized by the so-called direct pathway, which means that abscisic acid molecules are built up by combining three isoprenoid units. The biosynthetic pathway in higher plants is not clear, but presumably it is an indirect pathway that involves the breakdown of a larger precursor molecule. Several lines of evidence suggest that this precursor may be a carotenoid: mutants lacking carotenoids also fail to produce abscisic acid, and inhibitors that block carotenoid biosynthesis at the same time inhibit abscisic acid formation. Furthermore, labeling studies with $^{18}O_2$ indicate that one ^{18}O atom is rapidly incorporated into the carboxyl group of abscisic acid, whereas the ring oxygens become labeled much more slowly. These observations are in accordance with the hypothesis that abscisic acid is derived from a large precursor molecule in which the ring oxygens are already present and an oxygen atom is introduced into the carboxyl group by oxidative cleavage.

Physiological roles. Abscisic acid was originally discovered by F. T. Addicott as a factor that accelerates abscission of young cotton fruits (hence the original name abscisin), and by P. F. Wareing as a hormone (called dormin) that causes bud dormancy in woody species. However, subsequent work established that abscisic acid has little or no effect on abscission of leaves and fruits and that it is not an endogenous regulator of bud dormancy in woody species. *See* ABSCISSION; DORMANCY.

Several approaches have been used to determine the roles of abscisic acid in plants. Application of abscisic acid results in a multitude of responses, but most of these may be of a pharmacological nature rather than reflect the roles of endogenous abscisic acid. The study of mutants that are deficient in abscisic acid, such as the flacca mutant of tomato and droopy in potato, has been more informative; both mutants synthesize much less abscisic acid than wild-type plants. These mutants wilt readily because they do not close their stomata. When they are treated with abscisic acid, the normal phenotype is restored, thus demonstrating that abscisic acid has a key role in preventing excessive water loss. Abscisic acid also functions in growth of roots and shoots, in heterophylly in aquatic plants, and in preventing premature germination of developing embryos. On the other hand, there is now increasing evidence against a role for abscisic acid in gravitropic curvature of roots. Seedlings in which abscisic acid synthesis is inhibited or mutants in which levels of abscisic acid are reduced, show the same gravitropic response of their roots as control seedlings.

Stress. Biosynthesis of abscisic acid is greatly accelerated under stress conditions. For example, in a wilted leaf the abscisic acid content can increase as much as 40 times the original level over a 4- to 5-h period. Following rehydration of such a leaf, the abscisic acid level returns to that of a turgid leaf after 5 to 6 h. Thus, the water status of the leaf, in particular the loss of turgor, regulates the rate at which abscisic acid is synthesized and degraded. The excess abscisic acid is either oxidized to phaseic acid, which in turn may be converted to dihydrophaseic acid, or it may be conjugated to glucose to give a complex that is stored in the vacuole.

A very rapid response to an increase in abscisic acid in a leaf is the closure of stomata: abscisic acid affects the permeability of the guard cells, which results in loss of solutes and turgor, so that the stomata close and the plant is protected from desiccation.

Abscisic acid is called a stress hormone, because it can ameliorate the effects of various stresses (drought, cold, salinity) to which plants are exposed. Freezing tolerance can be induced in certain plants by application of abscisic acid. *See* COLD HARDINESS (PLANT); PLANT-WATER RELATIONS; PLANTS, SALINE ENVIRONMENTS OF.

Growth. Abscisic acid inhibits the growth of stems and expansion of young leaves by decreasing cell-wall loosening, an effect opposite to that caused by auxin. On the other hand, growth of roots may be promoted by abscisic acid under certain conditions. These opposite effects of abscisic acid on the growth of roots and shoots are advantageous for the survival of plants under water stress conditions. By inhibiting shoot growth, turgor pressure is maintained, whereas a stimulation of root growth by abscisic acid enlarges the root system, thus increasing water uptake. *See* PLANT GROWTH.

Heterophylly. In aquatic plants two distinct types of leaves are produced on the individual. Submerged leaves are highly dissected or linear without stomata, whereas aerial leaves are entire and have stomata. When abscisic acid is added to the medium, aerial leaves with stomata are produced on submerged shoots. It is conceivable that under natural conditions shoot tips emerging from the water experience water stress, which causes production of abscisic acid, which in turn results in a change in leaf morphology.

Embryogenesis, seed maturation, and germination. Following fertilization in higher plants, the ovule develops into the seed with the embryo inside. During

late embryogenesis the seed desiccates and the embryo becomes dormant. However, if young embryos are excised from the seed prior to the onset of dormancy and cultured in the laboratory, they will germinate. This precocious germination of excised, immature embryos can be prevented by application of abscisic acid. There is convincing evidence that abscisic acid produced in the seed plays an important role in preventing precocious germination early in embryogenesis. Mutants deficient in abscisic acid all show the phenomenon of vivipary, which means that the embryo germinates precociously while the seed is still attached to the mother plant. Elegant genetic experiments with seeds of the small crucifer *Arabidopsis thaliana* have demonstrated that it is abscisic acid produced by the embryo, not by the maternal tissues of the seed, that controls dormancy of the embryo. Abscisic acid also induces the accumulation of specific proteins during late embryogenesis, but it is not clear at present that such proteins play a role in the onset of embryo dormancy. *See* PLANT MORPHOGENESIS; SEED.

During germination the reserve foods (carbohydrate, protein, and lipid) in the seed are degraded and mobilized to the young seedling. The hormonal regulation of enzymes involved in the mobilization of these reserves has been studied extensively in cereal grains, particularly in barley. Various hydrolytic enzymes, for example, α-amylase, are synthesized by aleurone cells in response to gibberellin produced by the germinating embryo. It has long been known that applied abscisic acid can inhibit the synthesis of these hydrolytic enzymes. Also, it has been found that abscisic acid produced in germinating seeds during water stress can have the same effect. So, the germinating seed not only has the means to turn on the metabolic machinery to mobilize food reserves but can also turn it off during conditions unfavorable for growth. *See* PLANT PHYSIOLOGY.

Mode of action. Two types of responses to abscisic acid can be distinguished. The rapid response, such as stomatal closure, occurs within a few minutes and can be induced only by naturally occurring (+)-abscisic acid. Since abscisic acid is active on guard cells over a wide pH range, while it is taken up by the cells only at low pH, the site of action is likely on the outer surface of the plasma membrane. Stomatal closure is caused by the rapid release of solutes, in particular the potassium ion (K^+) and malate, from the guard cells. Thus, in the rapid response abscisic acid affects the permeability of membranes.

The slow responses, which take hours or even days to become apparent, involve gene expression and can be induced equally well by (+)-abscisic acid and synthetic (−)-abscisic acid. This implies that the fast and slow responses have different receptor sites. The slow responses involve inhibition of the synthesis of certain enzymes, as well as induction of synthesis of a new group of proteins, the so-called abscisic acid–inducible proteins. These proteins appear during late embryogenesis when the seed begins to desiccate, and during water stress in other organs of the plant. The genes for these proteins are expressed only in the presence of abscisic acid or after water stress. When examined at the level of messenger ribonucleic acid, the effects of abscisic acid and water stress are not cumulative, indicating that both act by way of the same mechanism. Changes in protein pattern induced by water stress persist but are rapidly reversed by rehydration. The function of the abscisic acid–inducible proteins is not known, but it is thought that they have a function in drought tolerance. *See* PLANT HORMONES. Jan A. D. Zeevaart

Bibliography. F. T. Addicott (ed.), *Abscisic Acid*, 1983; J. A. D. Zeevaart and R. A. Creelman, Metabolism and physiology of abscisic acid, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 39:439–473, 1988.

Abscission

The process whereby a plant sheds one of its parts. Leaves, flowers, seeds, and fruits are parts commonly abscised. Almost any plant part, such as very small buds and bracts to branches several inches in diameter, and scales or sheets of bark, may be abscised by some species. However, other species, including many annual plants, may show little abscission, especially of leaves.

Abscission may be of value to the plant in several ways. It can be a process of self-pruning, removing injured, diseased, or senescent parts. It permits the dispersal of seeds and other reproductive structures. It facilitates the recycling of mineral nutrients to the soil. It functions to maintain homeostasis in the plant, keeping in balance leaves and roots, and vegetative and reproductive parts. For example, trees living in climates that have a pronounced dry season are often adapted to abscising all or part of their leaves during the dry period, thus keeping the trees in balance with the water of the environment. Flowers and young fruit are usually produced in much larger numbers than the plant is capable of developing into mature fruit. Abscission of flowers and young fruits is the means whereby the plant keeps the number of fruits in balance with the ability of the leaves to nourish them. The selective abscission of branches and branchlets is an important factor in determining the architecture of a number of tree species.

Mechanisms. In most plants the process of abscission is restricted to an abscission zone at the base of an organ (see **illus**.); here separation is brought about by the disintegration of the walls of a special layer of cells, the separation layer. Less than an hour may be required to complete the entire process for some flower petals, but for most plant parts several days are required for completion of abscission. The portion of the abscission zone which remains on the plant commonly develops into a corky protective layer that becomes continuous with the cork of the stem.

In general, a healthy organ inhibits its own abscission. Only after it is injured, diseased, or senescent

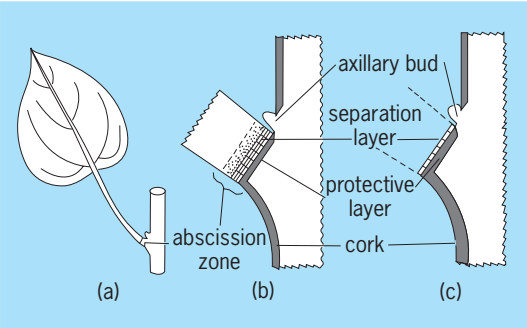

Diagrams of the abscission zone of a leaf. (*a*) A leaf with the abscission zone indicated at the base of the petiole. (*b*) The abscission zone layers shortly before abscission and (*c*) the layers after abscission.

is an organ abscised. In many instances abscission is also a correlation phenomenon, resulting from events elsewhere in the plant. For example, the abscission of flower petals usually follows closely after pollination and fertilization; and in some broad-leaved evergreen trees, such as live oaks, the camphor tree, and *Magnolia grandiflora*, the major flush of leaf abscission follows closely after the appearance of new buds and leaves in the spring.

Abscission is a typical physiological process requiring favorable temperatures, oxygen, and energy-yielding respiration. The energy is required for the synthesis of hydrolytic enzymes. These enzymes function to soften the pectins and other constituents of the cell wall. When the cell walls of the separation layer are sufficiently weakened, the leaf (or other organ) falls from the plant by its own weight.

Abscission takes place in the lower plants (mosses, ferns, algae, and fungi) in the same way as in the higher plants. Enzymes soften and weaken the pectinlike materials that cement cells together. The cementing substances of higher and lower plants are quite similar in that all are composed of branching polysaccharides. Abscission in the lower plants is less conspicuous because of their simpler structure. However, propagules, such as buds and plantlets, are regularly abscised. In the algae, fragmentation by abscission of segments of filaments is a common method of vegetative propagation. Yeast buds, most mold spores, and bacterial colonies also are separated by the dissolution of polysaccharide cementing substances.

Evolution. The fossil record shows that abscission has been taking place for a very long time. The record contains many instances of the abscission of fruits, seeds, and seedlike structures going back millions of years. Leaf abscission became evident soon after leaves evolved in the Devonian Period, some 400 million years ago. Fossils of the oldest known green plants, blue-green algae of the Precambrian Period, 850 million years ago, show that their filaments were abscising (fragmenting), as do present-day blue-green algae. From the evidence summarized above, it is reasonable to conclude that the evolution of abscission began when two cells first separated. *See* FOSSIL SEEDS AND FRUITS; PALEOBOTANY.

Factors affecting abscission. Observations of higher plants show that abscission is affected by a number of environmental and internal factors. It can be initiated or accelerated by extremes of temperature, such as frost; by extremes of moisture, such as drought or flooding; by deficiency of mineral elements in the soil, particularly nitrogen, calcium, magnesium, potassium, and zinc; by the shortening photoperiod in the fall; by oxygen concentrations above 20%; by volatile chemicals, such as ethylene and air pollutants; by moderately toxic chemicals; and by the feeding of some injurious insects. Abscission can be retarded or inhibited by excessive nitrogen from the soil; by high carbohydrate levels in the plant; by oxygen concentrations below 20%; and by application of growth-regulator chemicals, such as naphthaleneacetic acid.

Hormones. Although a number of internal chemical changes regularly precede the abscission of an organ, at most only a few of the changes appear to be directly related to the process of abscission. Of the changes, a decrease in the level of the growth hormone auxin appears to be the most important. Auxin applied experimentally to the distal (organ) side of an abscission zone retards abscission, while auxin applied to the proximal (stem) side accelerates abscission. From this and related evidence, it is considered that the gradient of auxin across the abscission zone is the major internal factor controlling abscission. Further, many other factors affecting abscission appear to act through effects on the auxin gradient. *See* AUXIN.

The gibberellins are growth hormones which influence a number of plant processes, including abscission. When applied to young fruits or to leaves, they tend to promote growth, delay maturation, and thereby indirectly prevent or delay abscission. *See* GIBBERELLIN.

Abscisic acid is a different type of hormone and has the ability to promote abscission and senescence and to retard growth. Like auxin, it is synthesized in fruits and leaves; however, it tends to counteract the effects of the growth-promoting hormones, auxin, gibberellin, and cytokinin. *See* ABSCISIC ACID; CYTOKININS.

Another hormonal chemical is the gas ethylene. Small amounts of ethylene have profound effects on the growth of plants and can distort and reduce growth and promote senescence and abscission. *See* ETHYLENE.

Thus the process of abscission is strongly influenced by interaction of several plant hormones. Frequently the hormones that tend to retard or delay abscission (auxin, gibberellin) are counteracted by the hormones that tend to promote and accelerate abscission (abscisic acid, ethylene). These interactions are further influenced by many environmental and external factors, as mentioned above. Depending on circumstances, any one of several factors may at times be the key factor controlling the entire process of abscission. *See* PLANT HORMONES.

Agricultural applications. In agricultural practice, abscission is delayed or prevented by keeping soil

moisture and soil nitrogen at high levels. High soil moisture helps to keep abscisic acid at relatively low levels, and high soil nitrogen leads to high levels of auxin in the plant. Auxin-type chemical regulators are frequently used to retard abscission. Acceleration of abscission can be obtained by low soil nitrogen (leading to low auxin), by moisture stress (leading to high abscisic acid), by ethylene-releasing chemicals, or by mildly toxic chemicals.

To delay abscission of leaves, fruits, and other plant parts, the auxin-type regulator, naphthaleneacetic acid, or related compounds are commonly used. Some notable applications are the spraying of apple orchards and pear orchards to prevent abscission of nearly mature fruit and the dipping of cut holly to prevent abscission of berries and leaves during shipment.

To accelerate abscission a variety of chemicals find use in one or another practice. Cotton is chemically defoliated to facilitate mechanical harvest. To promote leaf abscission, chemicals such as magnesium chlorate, tributylphosphorotrithioate, and merphos have been widely used. In some regions, young nursery plants are defoliated to facilitate earlier digging and shipping. Certain varieties of apples, peaches, and similar fruits must be thinned to ensure that the remaining fruit can reach marketable size. Naphthaleneacetamide and related compounds are the most frequently used to thin fruit. Although these chemicals are of the auxin type, the dosages applied to the flowers and young fruit are somewhat toxic and lead to abortion and abscission of a portion of the young fruit. The slightly toxic insecticide carbaryl is also used in fruit thinning. Efficient harvesting of mature fruit often depends on some chemical promotion of abscission. Ethaphon and other compounds that can release ethylene are often used, although they sometimes show undesirable side effects of subsequent growth. For the harvesting of oranges in Florida, various combinations of cycloheximide, 5-chloro-3-methyl-4-nitro-1*H*-pyrazole, and chlorothalonil have been very effective in promoting abscission. *See* PLANT PHYSIOLOGY.

Fredrick T. Addicott

Bibliography. F. T. Addicott, *Abscission,* 1982; L. G. Nickell, *Plant Growth Regulators,* 1982.

Absolute zero

The temperature at which an ideal gas would exert no pressure. The Kelvin scale of temperatures is defined in terms of the triple point of water, $T_3 = 273.16^\circ$ (where the solid, liquid, and vapor phases coexist) and absolute zero. Thus, the Kelvin degree is 1/273.16 of the thermodynamic temperature of the triple point. Temperature is measured most simply via the constant-volume ideal-gas thermometer, in which a small amount of gas is introduced (in order to limit the effect of interactions between molecules) and then sealed off. The gas pressure P referenced to its value at the triple point $P(T_3)$ is measured, and temperature T may be inferred from Eq. (1). The

$$T(K) = 273.16 \lim_{P \to 0} \frac{P(T)}{P(T_3)} \tag{1}$$

ideal-gas law applies if the molecules in a gas exert no forces on one another and if they are not attracted to the walls. With helium gas, the constant-volume gas thermometer can be used down to a temperature on the order of 1 K. Absolute zero is the temperature at which the pressure of a truly ideal gas would vanish. *See* GAS THERMOMETRY; TEMPERATURE MEASUREMENT.

In order that a temperature be assigned to a particular specimen, there must exist mechanisms for the free transfer of energy within the system. Measurement of the distribution of speeds of particles can serve to define the temperature through the Maxwell distribution. The most probable speed gets smaller as the temperature decreases. According to classical physics, all motion would cease at absolute zero; however, the quantum-mechanical uncertainty principle requires that there be a small amount of residual motion (zero-point motion) even at absolute zero. *See* KINETIC THEORY OF MATTER; QUANTUM MECHANICS; UNCERTAINTY PRINCIPLE.

The Kelvin scale can be obtained by measurements of the efficiency of an ideal heat engine operating between reservoirs at two fixed temperatures. According to the second law of thermodynamics, the maximum possible efficiency ϵ is given by Eq. (2),

$$\epsilon = 1 - \frac{T_1}{T_2} \tag{2}$$

where T_1 and T_2 are the temperatures of the low- and the high-temperature reservoirs respectively. All comparisons between the gas-thermometer scales and the thermodynamic scales have shown the equivalence of the two approaches. *See* CARNOT CYCLE; THERMODYNAMIC PRINCIPLES.

Temperature can also be defined from the Boltzmann distribution. If a collection of spin 1/2 magnetic ions is placed in a magnetic field, the ratio of the occupancy of the lower to the higher energy state is given by Eq. (3). Here k is Boltzmann's constant, $|\Delta E|$ is the magnitude of the difference in energy between the states, and T is the Kelvin temperature. Thus, at high temperatures the two states have nearly equal occupation probability, while the lower energy state is progressively favored at lower temperatures. At absolute zero, only the lower energy level is occupied. By measuring the populations of the two states, the temperature of the spins can be inferred. This relation allows for the possibility of negative temperatures when the population of the higher energy state exceeds that of the lower state. From the point of view of energy content, negative temperatures correspond to an energy of the spin system that exceeds that of an infinite positive temperature, and thus they are hotter than ordinary

$$\frac{N_L}{N_H} = \exp \frac{|\Delta E|}{kT} \tag{3}$$

temperatures. *See* BOLTZMANN CONSTANT; BOLTZMANN STATISTICS; NEGATIVE TEMPERATURE.

Negative temperatures notwithstanding, the third law of thermodynamics states that the absolute zero of temperature cannot be attained by any finite number of steps. The lowest (and hottest) temperatures that have been achieved are on the order of a picokelvin (10^{-12}K). These are spin temperatures of nuclei which are out of equilibrium with the lattice vibrations and electrons of a solid. The lowest temperatures to which the electrons have been cooled are on the order of 10 microkelvins in metallic systems. *See* LOW-TEMPERATURE PHYSICS; TEMPERATURE.

Jeevak M. Parpia; David M. Lee

Absorption

Either the taking up of matter in bulk by other matter, as in the dissolving of a gas by a liquid; or the taking up of energy from radiation by the medium through which the radiation is passing. In the first case, an absorption coefficient is defined as the amount of gas dissolved at standard conditions by a unit volume of the solvent. Absorption in this sense is a volume effect: The absorbed substance permeates the whole of the absorber. In absorption of the second type, attenuation is produced which in many cases follows Lambert's law and adds to the effects of scattering if the latter is present. *See* ATTENUATION; SCATTERING OF ELECTROMAGNETIC RADIATION.

Energy. Absorption of electromagnetic radiation can occur in several ways. For example, microwaves in a waveguide lose energy to the walls of the guide. For nonperfect conductors, the wave penetrates the guide surface, and energy in the wave is transferred to the atoms of the guide. Light is absorbed by atoms of the medium through which it passes, and in some cases this absorption is quite distinctive. Selected frequencies from a heterochromatic source are strongly absorbed, as in the absorption spectrum of the Sun. Electromagnetic radiation can be absorbed by the photoelectric effect, where the light quantum is absorbed and an electron of the absorbing atom is ejected, and also by Compton scattering. Electron-positron pairs may be created by the absorption of a photon of sufficiently high energy. Photons can be absorbed by photoproduction of nuclear and subnuclear particles, analogous to the photoelectric effect. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; COMPTON EFFECT; ELECTRON-POSITRON PAIR PRODUCTION; PHOTOEMISSION.

Sound waves are absorbed at suitable frequencies by particles suspended in the air (wavelength of the order of the particle size), where the sound energy is transformed into vibrational energy of the absorbing particles. *See* SOUND ABSORPTION.

Absorption of energy from a beam of particles can occur by the ionization process, where an electron in the medium through which the beam passes is removed by the beam particles. The finite range of protons and alpha particles in matter is a result of this process. In the case of low-energy electrons, scattering is as important as ionization, so that range is a less well-defined concept. Particles themselves may be absorbed from a beam. For example, in a nuclear reaction an incident particle X is absorbed into nucleus Y, and the result may be that another particle Z, or a photon, or particle X with changed energy comes out. Low-energy positrons are quickly absorbed by annihilating with electrons in matter to yield two gamma rays. *See* NUCLEAR REACTION.

McAllister H. Hull, Jr.

Matter. The absorption of matter is a chemical engineering unit operation. In the chemical process industries and in related areas such as petroleum refining and fuels purification, absorption usually means gas absorption. This is a unit operation in which a gas (or vapor) mixture is contacted with a liquid solvent selected to preferentially absorb one, or in some cases more than one, component from the mixture. The purpose is either to recover a desired component from a gas mixture or to rid the mixture of an impurity. In the latter case, the operation is often referred to as scrubbing. The gas, once absorbed in the liquid, is in a completely dissolved state and therefore has become a full-fledged member of the liquid phase. In the chemical industry, gas absorption is the second-most common separation operation involving gas-liquid contacting. Only fractional distillation is employed more frequently.

When the operation is employed in reverse, that is, when a gas is utilized to extract a component from a liquid mixture, it is referred to as gas desorption, stripping, or sparging.

In gas absorption, either no further changes occur to the gaseous component once it is absorbed in the liquid solvent, or the absorbed component (solute) will become involved in a chemical reaction with the solvent in the liquid phase. In the former case, the operation is referred to as physical gas absorption, and in the latter case as gas absorption with chemical reaction. Examples of simple, or physical, gas absorption include the absorption of light oil (benzene, toluene, xylenes) from coke oven by-product gases by scrubbing with a petroleum oil, and recovery of valuable solvent vapors, such as acetone or those used in drycleaning processes, from gas streams by washing the gas with an appropriate solvent for the vapors. Examples of gas absorption with chemical reaction include scrubbing the flue gases from coal-fired generating stations with aqueous sodium carbonate solution to remove sulfur dioxide (a potential cause of acid rain), recovery of ammonia from coke oven by-product gases by scrubbing with dilute sulfuric acid, and the purification of natural gas by absorption of hydrogen sulfide in aqueous ethanolamine.

Gas absorption is normally carried out in the same types of columns as those used in fractional distillation, namely tray-type or packed towers, but with the latter used much more frequently than the former in distillation. The operation can be carried out by using cocurrent flow, that is, with both the gas and liquid phases flowing upward in the column, but it is normally countercurrent with the liquid phase flowing

downward and the gas phase rising. Gas absorption, unlike distillation, is not carried out at the boiling point. To the contrary, and because the solubility of gases in liquids increases with decreasing temperature, gas absorption is usually carried out cold, that is, at a temperature not far above the freezing point of the solvent. An evolution of heat occurs during gas absorption, as a result of the absorbed component giving up its latent heat of vaporization upon absorption, just as it would have if it had been condensed. The resulting temperature rise is minimized by a normally large solvent-to-gas ratio used in order to keep low the concentration of absorbed gas in the liquid phase, and therefore to keep high the concentration driving force causing the absorption to occur. Unlike fractional distillation, gas absorption does not require a reboiler, condenser, or reflux. However, it does require a subsequent solvent-recovery operation to separate solvent and solute, so that the solvent can be recycled. The energy requirements for gas absorption and solvent recovery combined are usually substantially less than for fractional distillation. The main basis for solvent selection is that the solvent have a high solubility for the gaseous component to be absorbed, and the lowest possible solubility for the other components of the gas stream—in other words, that the solvent have both a large absorptive capacity and a high degree of selectivity. *See* GAS ABSORPTION OPERATIONS; DISTILLATION; DISTILLATION COLUMN; UNIT OPERATIONS. William F. Furter

Absorption (biology)

The net movement (transport) of water and solutes from outside an organism to its interior. The unidirectional flow of materials into an animal from the environment generally takes place across the alimentary tract, the lungs, or the skin, and in each location a specific cell layer called an epithelium regulates the passage of materials. *See* EPITHELIUM; RESPIRATION.

Structural aspects. The mammalian intestine is a good example of an epithelium which absorbs ions and water, along with a wide range of nutrient materials. The cells lining the gastrointestinal epithelium possess morphological features similar to those of many other epithelia.

Most intestinal epithelia have cells which are columnar in shape, have an extensive microvillar membrane facing the interior (or lumen) of the gut, and are functionally specialized to transport molecules from the gut contents to the blood (**Fig. 1**). The brush border or microvillar apical membrane has an expanded surface to maximize uptake of materials into the cytoplasm. Adjacent cells are attached to one another at the apical surface (facing the lumen) by "tight junctions," which serve as partial barriers to the passage of materials between the lumen and the blood by intercellular paths. Several types of enzymes, including peptidases, carbohydrases, and lipases, are located in the brush border membrane to break down complex peptides, carbohydrates, and fats before they are transported into the epithelial cytoplasm as molecular monomers (single units). *See* DIGESTIVE SYSTEM.

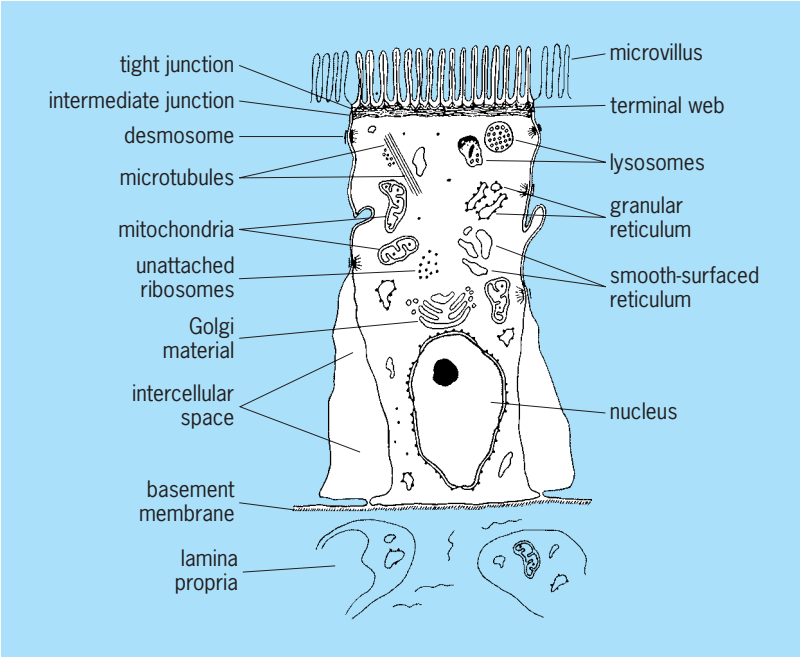

Fig. 1. Diagram of an absorptive epithelial cell from the mammalian small intestine. (*After L. R. Johnson, ed., Physiology of the Gastrointestinal Tract, Raven Press, 1981*)

The sides and bottoms of the cells with respect to the lumen are called the lateral and basal surfaces, or sometimes collectively the basolateral surfaces, and these may vary considerably among different epithelia and among species. While some epithelial cells have long, straight basolateral surfaces, others may be highly interdigitated with neighboring cells to increase their total surface area. In some epithelia these basolateral interdigitations contain abundant mitochondria, which provide energy for transport of materials across membranes. *See* MITOCHONDRIA.

During epithelial absorption, molecules must first pass across the apical membrane (from the mucosal side) into the cytoplasm. This entry process is then followed by ejection of the same molecule across the basolateral border to the blood side (called the serosal side) for eventual distribution to the body via the circulatory system. *See* CELL MEMBRANES.

Kinds of transport. Absorption across epithelia may occur by several different passive and active processes. Simple diffusion is the net movement of molecules from the apical to basolateral surfaces of an epithelium down chemical and electrical gradients without the requirement of cellular energy sources. Facilitated diffusion across the epithelium is similar to simple diffusion in that energy is not required, but in this process, molecular interaction with protein binding sites (carriers) in one or both membranes must occur to facilitate the transfer. Active molecular transport involves the use of membrane protein carriers as well as cellular energy supplies to move a transported molecule up an electrochemical gradient across the epithelium. Endocytosis and phagocytosis are also examples of active

transport because metabolic energy is required, but in these processes whole regions of the cell membrane are used to engulf fluid or particles, rather than to bring about molecular transfer using single-membrane proteins. *See* ENDOCYTOSIS; PHAGOCYTOSIS.

Mechanisms of ion absorption. Although a wide variety of ions are absorbed by different types of epithelial cells, the mechanisms of Na^+ and Cl^- transport in mammalian small intestine are perhaps best known in detail. Transepithelial transport of these two ions occurs in this tissue by three independent processes: active Na^+ absorption, not coupled directly to the flow of other solutes but accompanied indirectly by the diffusional absorption of Cl^-; coupled NaCl absorption; and cotransport of Na^+ with a wide variety of nutrient molecules.

During uncoupled active Na^+ transport, the cation crosses the apical membrane of the small-intestinal epithelial cell down an electrochemical gradient. Transfer across the brush border may be by simple or facilitated diffusion. Efflux of Na^+ up a concentration gradient out of the cell across the basolateral surface occurs by active transport, using a carrier protein called the Na^+/K^+–adenosine triphosphatase (Na^+ pump). Absorption of Na^+ by this two-step process results in the generation of a transepithelial electrical potential difference between the lumen and the blood of about 3–5 millivolts (blood electrically positive to apical surface). This electrical potential difference passively attracts Cl^- from the luminal side, most of this diffusion occurring paracellularly between the epithelial cells.

Two mechanisms have been proposed for coupled NaCl absorption in the mammalian small intestine. In the first model (**Fig. 2***a*), Na^+ and Cl^- bind to common apical membrane carriers and enter the epithelium together, with Na^+ diffusing down an electrochemical gradient and Cl^- moving up an electrochemical gradient. Efflux of Na^+ to the blood occurs by way of the energy-dependent Na^+ pump, while Cl^- moves passively across the basolateral border by facilitated diffusion. In the second model (Fig. 2*b*), Na^+ and Cl^- enter the cell by separate brush border carriers in exchange for intracellular H^+ and HCO_3^-, respectively. Their efflux to the blood occurs as described in Fig. 2*a*.

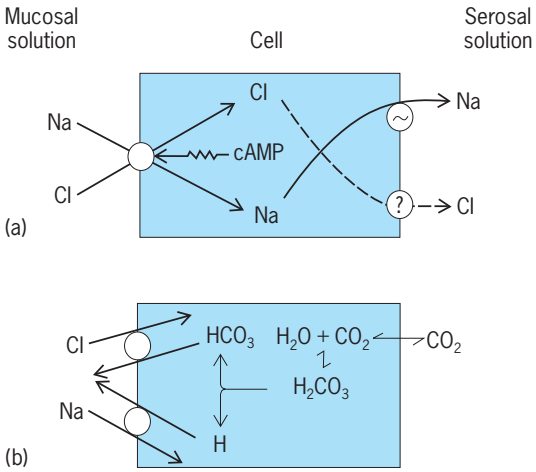

Fig. 2. Two models for electroneutral NaCl absorption by mammalian small-intestinal epithelial cells from the mucosal solution (food side) to the serosal solution (blood side). (*a*) Ions enter together by binding to common apical membrane carriers. (*b*) Ions enter by separate brush border carriers. (*After L. R. Johnson, ed., Physiology of the Gastrointestinal Tract, Raven Press, 1981*)

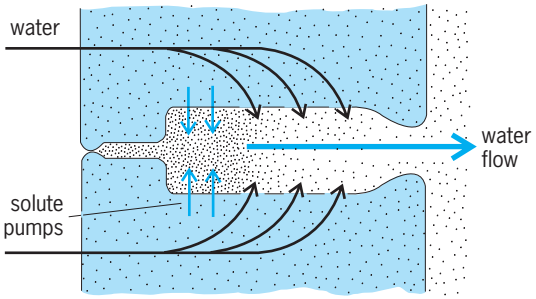

Fig. 3. Mechanism of solute-linked water absorption by epithelial cells. Density of dots refers to relative osmotic pressure in different compartments. (*After R. Eckert and D. Randall, Animal Physiology, Freeman, 1983*)

A final mechanism leading to the net flow of Na^+ across an epithelium is by coupling the entry of this ion into the cell with that of a wide variety of organic solutes. This pathway of Na^+ transport will be discussed below in conjunction with the mechanisms of nutrient absorption.

Mechanisms of water absorption. Net water transport across an epithelium is coupled to net ion transport in the same direction. As shown in **Fig. 3**, Na^+-pump sites are believed to be located along the lateral borders of epithelial cells. Energy-dependent Na^+ efflux from the cells to the intercellular spaces creates a local increase in osmotic pressure within these small compartments. An osmotic pressure gradient becomes established here, with greatest solute concentrations located nearest the tight junctions. Water flows into the cell across the brush border membrane and out the lateral membranes in response to the increased osmotic pressure in the paracellular spaces. Once water is in the intercellular compartment, a buildup of hydrostatic pressure forces the transported fluid to the capillary network. *See* OSMOREGULATORY MECHANISMS.

Mechanisms of nutrient absorption. It is believed that most nutrient molecules enter intestinal epithelial cells in conjunction with Na^+ on a common brush border membrane carrier protein (**Fig. 4**). Several types of proteins may exist on this membrane for each major category of organic solute (for example, a protein for sugars is functionally distinct from one for amino acids), but each carrier type requires Na^+ as a cotransported substrate. In all cases, Na^+ enters the cell down an electrochemical gradient, while the nutrient molecule is accumulated against a concentration gradient within the cytoplasm. The diffusional permeability of the brush border membrane toward organic solutes is very low, so that even though a considerable concentration gradient for the nutrient builds up across the apical cell pole, little passive loss

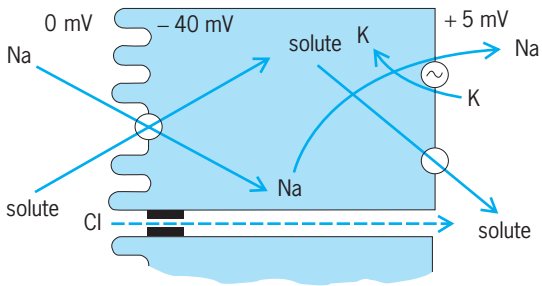

Fig. 4. Model for sodium-coupled absorption of organic solutes (D-hexoses, L-amino acids, dipeptides, tripeptides, water-soluble vitamins, bile salts) by mammalian small-intestinal epithelial cells. Cl^- ions diffuse between epithelial cells, down the electrochemical gradient. (*After L. R. Johnson, ed., Physiology of the Gastrointestinal Tract, Raven Press, 1981*)

to the lumen takes place. Intracellular Na^+ concentration is kept very low due to energy-dependent basolateral Na^+-pump activities transferring this cation up a concentration gradient out of the cell. Accumulated intracellular nutrients are transferred across the basolateral border by facilitated diffusion on a carrier protein which does not have a Na^+ requirement. No cellular metabolic energy is used directly for the absorption of nutrients. Instead, energy resources are used to maintain a steep Na^+ gradient across the brush border membrane to facilitate the accumulation of nutrient across this cell pole as a result of its cotransport with the cation. *See* ION TRANSPORT.

Gregory A. Ahearn

Bibliography. G. Giebisch, D. C. Tosteson, and H. H. Ussing (eds.), *Membrane Transport in Biology*, vols. 3 and 4, 1978; L. R. Johnson (ed.), *Physiology of the Gastrointestinal Tract*, vols. 1 and 2, 1981; H. Murer, U. Hopfer, and R. Kinne, Sodium-proton antiport in brush border membrane vesicles isolated from rat small intestine, *Biochem. J.*, 154:597–604, 1976; H. N. Nellans, R. A. Frizzell, and S. G. Schultz, Coupled sodium-chloride influx across the brush border of rabbit ileum, *Amer. J. Physiol.*, 225:467–475, 1973; S. G. Schultz and P. F. Curran, Coupled transport of sodium and organic solutes, *Physiol. Rev.*, 50:637–718, 1970.

Absorption of electromagnetic radiation

The process whereby the intensity of a beam of electromagnetic radiation is attenuated in passing through a material medium by conversion of the energy of the radiation to an equivalent amount of energy which appears within the medium; the radiant energy is converted into heat or some other form of molecular energy. A perfectly transparent medium permits the passage of a beam of radiation without any change in intensity other than that caused by the spread or convergence of the beam, and the total radiant energy emergent from such a medium equals that which entered it, whereas the emergent energy from an absorbing medium is less than that which enters, and, in the case of highly opaque media, is reduced practically to zero.

No known medium is opaque to all wavelengths of the electromagnetic spectrum, which extends from radio waves, whose wavelengths are measured in kilometers, through the infrared, visible, and ultraviolet spectral regions, to x-rays and gamma rays, of wavelengths down to 10^{-13} m. Similarly, no material medium is transparent to the whole electromagnetic spectrum. A medium which absorbs a relatively wide range of wavelengths is said to exhibit general absorption, while a medium which absorbs only restricted wavelength regions of no great range exhibits selective absorption for those particular spectral regions. For example, the substance pitch shows general absorption for the visible region of the spectrum, but is relatively transparent to infrared radiation of long wavelength. Ordinary window glass is transparent to visible light, but shows general absorption for ultraviolet radiation of wavelengths below about 310 nanometers, while colored glasses show selective absorption for specific regions of the visible spectrum. The color of objects which are not self-luminous and which are seen by light reflected or transmitted by the object is usually the result of selective absorption of portions of the visible spectrum. Many colorless substances, such as benzene and similar hydrocarbons, selectively absorb within the ultraviolet region of the spectrum, as well as in the infrared. *See* COLOR; ELECTROMAGNETIC RADIATION.

Laws of absorption. The capacity of a medium to absorb radiation depends on a number of factors, mainly the electronic and nuclear constitution of the atoms and molecules of the medium, the wavelength of the radiation, the thickness of the absorbing layer, and the variables which determine the state of the medium, of which the most important are the temperature and the concentration of the absorbing agent. In special cases, absorption may be influenced by electric or magnetic fields. The state of polarization of the radiation influences the absorption of media containing certain oriented structures, such as crystals of other than cubic symmetry. *See* STARK EFFECT; ZEEMAN EFFECT.

Lambert's law. Lambert's law, also called Bouguer's law or the Lambert-Bouguer law, expresses the effect of the thickness of the absorbing medium on the absorption. If a homogeneous medium is thought of as being constituted of layers of uniform thickness set normally to the beam, each layer absorbs the same fraction of radiation incident on it. If I is the intensity to which a monochromatic parallel beam is attenuated after traversing a thickness d of the medium, and I_0 is the intensity of the beam at the surface of incidence (corrected for loss by reflection from this surface), the variation of intensity throughout the medium is expressed by Eq. (1), in which α is a

$$I = I_0 e^{-\alpha d} \tag{1}$$

constant for the medium called the absorption coefficient. This exponential relation can be expressed in

an equivalent logarithmic form as in Eq. (2), where

$$\log_{10}(I_0/I) = (\alpha/2.303)d = kd \qquad (2)$$

$k = \alpha/2.303$ is called the extinction coefficient for radiation of the wavelength considered. The quantity $\log_{10}(I_0/I)$ is often called the optical density, or the absorbance of the medium.

Equation (2) shows that as monochromatic radiation penetrates the medium, the logarithm of the intensity decreases in direct proportion to the thickness of the layer traversed. If experimental values for the intensity of the light emerging from layers of the medium of different thicknesses are available (corrected for reflection losses at all reflecting surfaces), the value of the extinction coefficient can be readily computed from the slope of the straight line representing the logarithms of the emergent intensities as functions of the thickness of the layer.

Equations (1) and (2) show that the absorption and extinction coefficients have the dimensions of reciprocal length. The extinction coefficient is equal to the reciprocal of the thickness of the absorbing layer required to reduce the intensity to one-tenth of its incident value. Similarly, the absorption coefficient is the reciprocal of the thickness required to reduce the intensity to $1/e$ of the incident value, where e is the base of the natural logarithms, 2.718.

Beer's law. This law refers to the effect of the concentration of the absorbing medium, that is, the mass of absorbing material per unit of volume, on the absorption. This relation is of prime importance in describing the absorption of solutions of an absorbing solute, since the solute's concentration may be varied over wide limits, or the absorption of gases, the concentration of which depends on the pressure. According to Beer's law, each individual molecule of the absorbing material absorbs the same fraction of the radiation incident upon it, no matter whether the molecules are closely packed in a concentrated solution or highly dispersed in a dilute solution. The relation between the intensity of a parallel monochromatic beam which emerges from a plane parallel layer of absorbing solution of constant thickness and the concentration of the solution is an exponential one, of the same form as the relation between intensity and thickness expressed by Lambert's law. The effects of thickness d and concentration c on absorption of monochromatic radiation can therefore be combined in a single mathematical expression, given in Eq. (3), in which k' is a constant for a

$$I = I_0 e^{-k'cd} \qquad (3)$$

given absorbing substance (at constant wavelength and temperature), independent of the actual concentration of solute in the solution. In logarithms, the relation becomes Eq. (4). The values of the constants k'

$$\log_{10}(I_0/I) = (k'/2.303)cd = \epsilon cd \qquad (4)$$

and ϵ in Eqs. (3) and (4) depend on the units of concentration. If the concentration of the solute is expressed in moles per liter, the constant ϵ is called the molar extinction coefficient. Some authors employ the symbol a_M, which is called the molar absorbance index, instead of ϵ.

If Beer's law is adhered to, the molar extinction coefficient does not depend on the concentration of the absorbing solute, but usually changes with the wavelength of the radiation, with the temperature of the solution, and with the solvent.

The dimensions of the molar extinction coefficient are reciprocal concentration multiplied by reciprocal length, the usual units being liters/(mole)(cm). If Beer's law is true for a particular solution, the plot of $\log(I_0/I)$ against the concentrations for solutions of different concentrations, measured in cells of constant thickness, will yield a straight line, the slope of which is equal to the molar extinction coefficient.

While no true exceptions to Lambert's law are known, exceptions to Beer's law are not uncommon. Such exceptions arise whenever the molecular state of the absorbing solute depends on the concentration. For example, in solutions of weak electrolytes, whose ions and undissociated molecules absorb radiation differently, the changing ratio between ions and undissociated molecules brought about by changes in the total concentration prevents solutions of the electrolyte from obeying Beer's law. Aqueous solutions of dyes frequently deviate from the law because of dimerization and more complicated aggregate formation as the concentration of dye is increased.

Absorption measurement. The measurement of the absorption of homogeneous media is usually accomplished by absolute or comparative measurements of the intensities of the incident and transmitted beams, with corrections for any loss of radiant energy caused by processes other than absorption. The most important of these losses is by reflection at the various surfaces of the absorbing layer and of vessels which may contain the medium, if the medium is liquid or gaseous. Such losses are usually automatically compensated for by the method of measurement employed. Losses by reflection not compensated for in this manner may be computed from Fresnel's laws of reflection. *See* REFLECTION OF ELECTROMAGNETIC RADIATION.

Scattering. Absorption of electromagnetic radiation should be distinguished from the phenomenon of scattering, which occurs during the passage of radiation through inhomogeneous media. Radiant energy which traverses media constituted of small regions of refractive index different from that of the rest of the medium is diverted laterally from the direction of the incident beam. The diverted radiation gives rise to the hazy or opalescent appearance characteristic of such media, exemplified by smoke, mist, and opal. If the centers of inhomogeneity are sufficiently dilute, the intensity of a parallel beam is diminished in its passage through the medium because of the sidewise scattering, according to a law of the same form as the Lambert-Bouguer law for absorption, given in Eq. (5), where I is the intensity of

$$I = I_0 e^{-\tau d} \qquad (5)$$

the primary beam of initial intensity I_0, after it has traversed a distance d through the scattering medium. The coefficient τ, called the turbidity of the medium, plays the same part in weakening the primary beam by scattering as does the absorption coefficient in true absorption. However, in true scattering, no loss of total radiant energy takes place, energy lost in the direction of the primary beam appearing in the radiation scattered in other directions. In some inhomogeneous media, both absorption and scattering occur together. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

Physical nature. Absorption of radiation by matter always involves the loss of energy by the radiation and a corresponding gain in energy by the atoms or molecules of the medium.

The energy of an assembly of gaseous atoms consists partly of kinetic energy of the translational motion which determines the temperature of the gas (thermal energy), and partly of internal energy, associated with the binding of the extranuclear electrons to the nucleus, and with the binding of the particles within the nucleus itself. Molecules, composed of more than one atom, have, in addition, energy associated with periodic rotations of the molecule as a whole and with oscillations of the atoms within the molecule with respect to one another.

The energy absorbed from radiation appears as increased internal energy, or in increased vibrational and rotational energy of the atoms and molecules of the absorbing medium. As a general rule, translational energy is not directly increased by absorption of radiation, although it may be indirectly increased by degradation of electronic energy or by conversion of rotational or vibrational energy to that of translation by intermolecular collisions.

Quantum theory. In order to construct an adequate theoretical description of the energy relations between matter and radiation, it has been necessary to amplify the wave theory of radiation by the quantum theory, according to which the energy in radiation occurs in natural units called quanta. The value of the energy in these units, expressed in ergs or calories, for example, is the same for all radiation of the same wavelength, but differs for radiation of different wavelengths. The energy E in a quantum of radiation of frequency ν (where the frequency is equal to the velocity of the radiation in a given medium divided by its wavelength in the same medium) is directly proportional to the frequency, or inversely proportional to the wavelength, according to the relation given in Eq. (6), where h is a universal constant known as Planck's constant. The value of h is 6.63×10^{-34} joule-second, and if ν is expressed in s^{-1}, E is given in joules per quantum. *See* QUANTUM MECHANICS.

$$E = h\nu \tag{6}$$

The most energetic type of change that can occur in an atom involves the nucleus, and increase of nuclear energy by absorption therefore requires quanta of very high energy, that is, of high frequency or low wavelength. Such rays are the γ-rays, whose wavelength varies downward from 0.01 nm. Next in energy are the electrons nearest to the nucleus and therefore the most tightly bound. These electrons can be excited to states of higher energy by absorption of x-rays, whose range in wavelength is from about 0.01 to 1 nm. Less energy is required to excite the more loosely bound valence electrons. Such excitation can be accomplished by the absorption of quanta of visible radiation (wavelength 700 nm for red light to 400 nm for blue) or of ultraviolet radiation, of wavelength down to about 100 nm. Absorption of ultraviolet radiation of shorter wavelengths, down to those on the border of the x-ray region, excites electrons bound to the nucleus with intermediate strength.

The absorption of relatively low-energy quanta of wavelength from about 1 to 10 micrometers suffices to excite vibrating atoms in molecules to higher vibrational states, while changes in rotational energy, which are of still smaller magnitude, may be excited by absorption of radiation of still longer wavelength, from the short-wavelength radio region of about 1 cm to long-wavelength infrared radiation, some hundredths of a centimeter in wavelength.

Gases. The absorption of gases composed of atoms is usually very selective. For example, monatomic sodium vapor absorbs very strongly over two narrow wavelength regions in the yellow part of the visible spectrum (the so-called D lines), and no further absorption by monatomic sodium vapor occurs until similar narrow lines appear in the near-ultraviolet. The valence electron of the sodium atom can exist only in one of a series of energy states separated by relatively large energy intervals between the permitted values, and the sharp-line absorption spectrum results from transitions of the valence electron from the lowest energy which it may possess in the atom to various excited levels. Line absorption spectra are characteristic of monatomic gases in general. *See* ATOMIC STRUCTURE AND SPECTRA.

The visible and ultraviolet absorption of vapors composed of diatomic or polyatomic molecules is much more complicated than that of atoms. As for atoms, the absorbed energy is utilized mainly in raising one of the more loosely bound electrons to a state of higher energy, but the electronic excitation of a molecule is almost always accompanied by simultaneous excitation of many modes of vibration of the atoms within the molecule and of rotation of the molecule as a whole. As a result, the absorption, which for an atom is concentrated in a very sharp absorption line, becomes spread over a considerable spectral region, often in the form of bands. Each band corresponds to excitation of a specific mode of vibration accompanying the electronic change, and each band may be composed of a number of very fine lines close together in wavelength, each of which corresponds to a specific rotational change of the molecule accompanying the electronic and vibrational changes. Band spectra are as characteristic of the absorption of molecules in the gaseous state, and frequently in the liquid state, as line spectra

are of gaseous atoms. *See* BAND SPECTRUM; LINE SPECTRUM; MOLECULAR STRUCTURE AND SPECTRA.

Liquids. Liquids usually absorb radiation in the same general spectral region as the corresponding vapors. For example, liquid water, like water vapor, absorbs infrared radiation strongly (vibrational transitions), is largely transparent to visible and near-ultraviolet radiation, and begins to absorb strongly in the far-ultraviolet. A universal difference between liquids and gases is the disturbance in the energy states of the molecules in a liquid caused by the great number of intermolecular collisions; this has the effect of broadening the very fine lines observed in the absorption spectra of vapors, so that sharp-line structure disappears in the absorption bands of liquids.

Solids. Substances which can exist in solid, liquid, and vapor states without undergoing a temperature rise to very high values usually absorb in the same general spectral regions for all three states of aggregation, with differences in detail because of the intermolecular forces present in the liquid and solid. Crystalline solids, such as rock salt or silver chloride, absorb infrared radiation of long wavelength, which excites vibrations of the electrically charged ions of which these salts are composed; such solids are transparent to infrared radiations of shorter wavelengths. In colorless solids, the valence electrons are too tightly bound to the nuclei to be excited by visible radiation, but all solids absorb in the near- or far-ultraviolet region. *See* CRYSTAL OPTICS; INTERMOLECULAR FORCES.

The use of solids as components of optical instruments is restricted by the spectral regions to which they are transparent. Crown glass, while showing excellent transparency for visible light and for ultraviolet radiation immediately adjoining the visible region, becomes opaque to radiation of wavelength about 300 nm and shorter, and is also opaque to infrared radiation longer than about 2000 nm in wavelength. Quartz is transparent down to wavelengths about 180 nm in the ultraviolet, and to about 4 μm in the infrared. The most generally useful material for prisms and windows for the near-infrared region is rock salt, which is highly transparent out to about 15 μm. For a detailed discussion of the properties of optical glass *see* OPTICAL MATERIALS.

Fluorescence. The energy acquired by matter by absorption of visible or ultraviolet radiation, although primarily used to excite electrons to higher energy states, usually ultimately appears as increased kinetic energy of the molecules, that is, as heat. It may, however, under special circumstances, be reemitted as electromagnetic radiation. Fluorescence is the reemission, as radiant energy, of absorbed radiant energy, normally at wavelengths the same as or longer than those absorbed. The reemission, as ordinarily observed, ceases immediately when the exciting radiation is shut off. Refined measurements show that the fluorescent reemission persists, in different cases, for periods of the order of 10^{-9} to 10^{-5} s. The simplest case of fluorescence is the resonance fluorescence of monatomic gases at low pressure, such as sodium or mercury vapors, in which the reemitted radiation is of the same wavelength as that absorbed. In this case, fluorescence is the converse of absorption: Absorption involves the excitation of an electron from its lowest energy state to a higher energy state by radiation, while fluorescence is produced by the return of the excited electron to the lower state, with the emission of the energy difference between the two states as radiation. The fluorescent radiation of molecular gases and of nearly all liquids, solids, and solutions contains a large component of wavelengths longer than those of the absorbed radiation, a relationship known as Stokes' law of fluorescence. In these cases, not all of the absorbed energy is reradiated, a portion remaining as heat in the absorbing material. The fluorescence of iodine vapor is easily seen on projecting an intense beam of visible light through an evacuated bulb containing a few crystals of iodine, but the most familiar examples are provided by certain organic compounds in solution—for instance, quinine sulfate, which absorbs ultraviolet radiation and reemits blue, or fluorescein, which absorbs blue-green light and fluoresces with an intense, bright-green color. *See* FLUORESCENCE.

Phosphorescence. The radiant reemission of absorbed radiant energy at wavelengths longer than those absorbed, for a readily observable interval after withdrawal of the exciting radiation, is called phosphorescence. The interval of persistence, determined by means of a phosphoroscope, usually varies from about 0.001 s to several seconds, but some phosphors may be induced to phosphorescence by heat days or months after the exciting absorption. An important and useful class of phosphors is the impurity phosphors, solids such as the sulfides of zinc or calcium which are activated to the phosphorescent state by incorporating minute amounts of foreign material (called activators), such as salts of manganese or silver. So-called fluorescent lamps contain a coating of impurity phosphor on their inner wall which, after absorbing ultraviolet radiation produced by passage of an electrical discharge through mercury vapor in the lamp, reemits visible light. The receiving screen of a television tube contains a similar coating, excited not by radiant energy but by the impact of a stream of electrons on the surface. *See* FLUORESCENT LAMP; PHOSPHORESCENCE.

Luminescence. Phosphorescence and fluorescence are special cases of luminescence, which is defined as light emission that cannot be attributed merely to the temperature of the emitting body. Luminescence may be excited by heat (thermoluminescence), by electricity (electroluminescence), by chemical reaction (chemiluminescence), or by friction (triboluminescence), as well as by radiation. *See* LUMINESCENCE.

Absorption and emission coefficients. The absorption and emission processes of atoms were examined from the quantum point of view by Albert Einstein in 1916, with some important results that have been realized practically in the invention of the maser and the laser. Consider an assembly of atoms undergoing absorption transitions of frequency ν s^{-1} from the

ground state 1 to an excited state 2 and emission transitions in the reverse direction, the atoms and radiation being at equilibrium at temperature T. The equilibrium between the excited and unexcited atoms is determined by the Boltzmann relation $N_2/N_1 = \exp(-h\nu/kT)$, where N_1 and N_2 are the equilibrium numbers of atoms in states 1 and 2, respectively, and the radiational equilibrium is determined by equality in the rate of absorption and emission of quanta. The number of quanta absorbed per second is $B_{12}N_1\rho(\nu)$, where $\rho(\nu)$ is the density of radiation of frequency ν (proportional to the intensity), and B_{12} is a proportionality constant called the Einstein coefficient for absorption. Atoms in state 2 will emit radiation spontaneously (fluorescence), after a certain mean life, at a rate of $A_{21}N_2$ per second, where A_{21} is the Einstein coefficient for spontaneous emission from state 2 to state 1. To achieve consistency between the density of radiation of frequency ν at equilibrium calculated from these considerations and the value calculated from Planck's radiation law, which is experimentally true, it is necessary to introduce, in addition to the spontaneous emission, an emission of intensity proportional to the radiation density of frequency ν in which the atoms are immersed. The radiational equilibrium is then determined by Eq. (7), where B_{21} is

$$B_{12}N_1\rho(\nu) = A_{21}N_2 + B_{21}N_2\rho(\nu) \tag{7}$$

the Einstein coefficient of stimulated emission. The Einstein radiation coefficients are found to be related by Eqs. (8*a*) and (8*b*).

$$B_{12} = B_{21} \tag{8a}$$

$$A_{21} = (8\pi h\nu^3/c^3) \cdot B_{21} \tag{8b}$$

In the past when one considered radiation intensities available from terrestrial sources, stimulated emission was very feeble compared with the spontaneous process. Stimulated emission is, however, the fundamental emission process in the laser, a device in which a high concentration of excited molecules is produced by intense illumination from a "pumping" source, in an optical system in which excitation and emission are augmented by back-and-forth reflection until stimulated emission swamps the spontaneous process. *See* LASER; OPTICAL PUMPING.

There are also important relations between the absorption characteristics of atoms and their mean lifetime τ in the excited state. Since A_{21} is the number of times per second that a given atom will emit a quantum spontaneously, the mean lifetime before emission in the excited state is $\tau = 1/A_{21}$. It can also be shown that A_{21} and τ are related, as shown in Eq. (9),

$$\begin{aligned} A_{21} = 1/\tau &= \frac{(8\pi^2\nu^2e^2)}{mc^3} \cdot f \\ &= 7.42 \times 10^{-22} f\nu^2 \qquad (\nu \text{ in s}^{-1}) \end{aligned} \tag{9}$$

to the f number or oscillator strength for the transition that occurs in the dispersion equations shown as Eqs. (13) to (17). The value of f can be calculated from the absorption integrated over the band according to Eq. (18).

Dispersion. A transparent material does not abstract energy from radiation which it transmits, but it always decreases the velocity of propagation of such radiation. In a vacuum, the velocity of radiation is the same for all wavelengths, but in a material medium, the velocity of propagation varies considerably with wavelength. The refractive index n of a medium is the ratio of the velocity of light in vacuum to that in the medium, and the effect of the medium on the velocity of radiation which it transmits is expressed by the variation of refractive index with the wavelength λ of the radiation, $dn/d\lambda$. This variation is called the dispersion of the medium. For radiation of wavelengths far removed from those of absorption bands of the medium, the refractive index increases regularly with decreasing wavelength or increasing frequency; the dispersion is then said to be normal.

In regions of normal dispersion, the variation of refractive index with wavelength can be expressed with considerable accuracy by Eq. (10), known as

$$n = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} \tag{10}$$

Cauchy's equation, in which A, B, and C are constants with positive values. As an approximation, C may be neglected in comparison with A and B, and the dispersion, $dn/d\lambda$, is then given by Eq. (11). Thus, in

$$\frac{dn}{d\lambda} = \frac{-2B}{\lambda^3} \tag{11}$$

regions of normal dispersion, the dispersion is approximately inversely proportional to the cube of the wavelength.

Dispersion by a prism. The refraction, or bending, of a ray of light which enters a material medium obliquely from vacuum or air (the refractive index of which for visible light is nearly unity) is the result of the diminished rate of advance of the wavefronts in the medium. Since, if the dispersion is normal, the refractive index of the medium is greater for violet than for red light, the wavefront of the violet light is retarded more than that of the red light. Hence, white light entering obliquely into the medium is converted within the medium to a continuously colored band, of which the red is least deviated from the direction of the incident beam, the violet most, with orange, yellow, green, and blue occupying intermediate positions. On emergence of the beam into air again, the colors remain separated. The action of the prism in resolving white light into its constituent colors is called color dispersion. *See* OPTICAL PRISM; REFRACTION OF WAVES.

The angular dispersion of a prism is the ratio, $d\theta/d\lambda$, of the difference in angular deviation $d\theta$ of two rays of slightly different wavelength which pass through the prism to the difference in wavelength $d\lambda$ when the prism is set for minimum deviation.

The angular dispersion of the prism given in Eq. (12) is the product of two factors, the variation,

$$\frac{d\theta}{d\lambda} = \frac{d\theta}{dn} \cdot \frac{dn}{d\lambda} \tag{12}$$

$d\theta/dn$, the deviation θ with refractive index n, and

the variation of refractive index with wavelength, the dispersion of the material of which the prism is made. The latter depends solely on this material, while $d\theta/dn$ depends on the angle of incidence and the refracting angle of the prism. The greater the dispersion of the material of the prism, the greater is the angular separation between rays of two given wavelengths as they leave the prism. For example, the dispersion of quartz for visible light is lower than that of glass; hence the length of the spectrum from red to violet formed by a quartz prism is less than that formed by a glass prism of equal size and shape. Also, since the dispersion of colorless materials such as glass or quartz is greater for blue and violet light than for red, the red end of the spectrum formed by prisms is much more contracted than the blue.

The colors of the rainbow result from dispersion of sunlight which enters raindrops and is refracted and dispersed in passing through them to the rear surface, at which the dispersed rays are reflected and reenter the air on the side of the drop on which the light was incident. *See* METEOROLOGICAL OPTICS; RAINBOW.

Anomalous dispersion. The regular increase of refractive index with decreasing wavelength expressed by Cauchy's equation breaks down as the wavelengths approach those of strong absorption bands. As the absorption band is approached from the long-wavelength side, the refractive index becomes very large, then decreases within the band to assume abnormally small values on the short-wavelength side, values below those for radiation on the long-wavelength side. A hollow prism containing an alcoholic solution of the dye fuchsin, which absorbs green light strongly, forms a spectrum in which the violet rays are less deviated than the red, on account of the abnormally low refractive index of the medium for violet light. The dispersion of media for radiation of wavelengths near those of strong absorption bands is said to be anomalous, in the sense that the refractive index decreases with decreasing wavelength instead of showing the normal increase. The theory of dispersion shows, however, that both the normal and anomalous variation of refractive index with wavelength can be satisfactorily described as aspects of a unified phenomenon, so that there is nothing fundamentally anomalous about dispersion in the vicinity of an absorption band. *See* DISPERSION (RADIATION).

Normal and anomalous dispersion of quartz are illustrated in **Fig. 1**. Throughout the near-infrared, visible, and near-ultraviolet spectral regions (between P and R on the curve), the dispersion is normal and adheres closely to Cauchy's equation, but it becomes anomalous to the right of R. From S to T, Cauchy's equation is again valid.

Relation to absorption. Figure 1 shows there is an intimate connection between dispersion and absorption; the refractive index rises to high values as the absorption band is approached from the long-wavelength side and falls to low values on the short-wavelength side of the band. In fact, the theory of dispersion shows that the complete dispersion curve

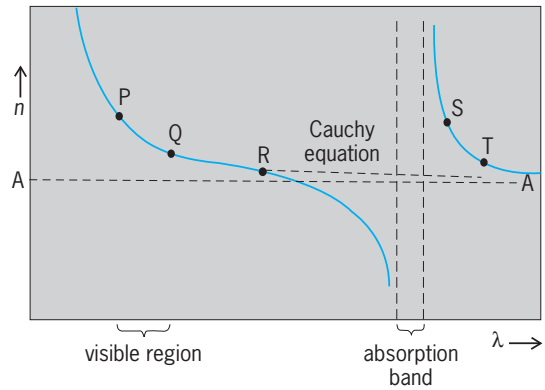

Fig. 1. Curve showing anomalous dispersion of quartz. A is limiting value of *n* as λ approaches infinity. (*After F. A. Jenkins and H. E. White, Fundamentals of Optics, 4th ed., McGraw-Hill, 1976*)

as a function of wavelength is governed by the absorption bands of the medium. In classical electromagnetic theory, electric charges are regarded as oscillating, each with its appropriate natural frequency ν_0, about positions of equilibrium within atoms or molecules. Placed in a radiation field of frequency ν per second, the oscillator in the atom is set into forced vibration, with the same frequency as that of the radiation. When ν is much lower or higher than ν_0, the amplitude of the forced vibration is small, but the amplitude becomes large when the frequency of the radiation equals the natural frequency of the oscillator. In much the same way, a tuning fork is set into vibration by sound waves corresponding to the same note emitted by another fork vibrating at the same frequency. To account for the absorption of energy by the medium from the radiation, it is necessary to postulate that in the motion of the atomic oscillator some frictional force, proportional to the velocity of the oscillator, must be overcome. For small amplitudes of forced oscillation, when the frequency of the radiation is very different from the natural period of the oscillator, the frictional force and the absorption of energy are negligible. Near resonance between the radiation and the oscillator, the amplitude becomes large, with a correspondingly large absorption of energy to overcome the frictional resistance. Radiation of frequencies near the natural frequency therefore corresponds to an absorption band. *See* SYMPATHETIC VIBRATION.

To show that the velocity of the radiation within the medium is changed, it is necessary to consider the phase of the forced vibration, which the theory shows to depend on the frequency of the radiation. The oscillator itself becomes a source of secondary radiation waves within the medium which combine to form sets of waves moving parallel to the original waves. Interference between the secondary and primary waves takes place, and because the phase of the secondary waves, which is the same as that of the atomic oscillators, is not the same as that of the primary waves, the wave motion resulting from the interference between the two sets of waves is different in phase from that of the primary waves incident

on the medium. But the velocity of propagation of the waves is the rate of advance of equal phase; hence the phase change effected by the medium, which is different for each frequency of radiation, is equivalent to a change in the velocity of the radiation within the medium. When the frequency of the radiation slightly exceeds the natural frequency of the oscillator, the radiation and the oscillator becomes 180° out of the phase, which corresponds to an increase in the velocity of the radiation and accounts for the fall in refractive index on the short-wavelength side of the absorption band.

The theory leads to Eqs. (13) through (17) for the refractive index of a material medium as a function of the frequency of the radiation. In the equations the frequency is expressed as angular frequency, $\omega = 2\pi\nu\ \mathrm{s}^{-1} = 2\pi c/\lambda$ where c is the velocity of light. When the angular frequency ω of the radiation is not very near the characteristic frequency of the electronic oscillator, the refractive index of the homogeneous medium containing N molecules per cubic centimeter is given by Eq. (13*a*), where e and m are

$$n^2 = 1 + \frac{4\pi Ne^2}{m} \cdot \frac{f}{\omega_0{}^2 - \omega^2} \qquad (13a)$$

$$n^2 = 1 + 4\pi Ne^2 \sum_i \frac{f_i/m_i}{\omega_i{}^2 - \omega^2} \qquad (13b)$$

the charge and mass of the electron, and f is the number of the oscillators per molecule of characteristics frequency ω_0. The f value is sometimes called the oscillator strength. If the molecule contains oscillators of different frequencies and mass (for example, electronic oscillators of frequency corresponding to ultraviolet radiation and ionic oscillators corresponding to infrared radiation), the frequency term becomes a summation, as in Eq. (13*b*), where ω_i is the characteristic frequency of the ith type of oscillator, and f_i and m_i are the corresponding f value and mass. In terms of wavelengths, this relation can be written as Eq. (14), where A_i is a constant for the medium,

$$n^2 = 1 + \sum_i \frac{A_i\lambda^2}{\lambda^2 - \lambda_i{}^2} \qquad (14)$$

λ is the wavelength of the radiation, and $\lambda_i = c/\nu_i$ is the wavelength corresponding to the characteristic frequency ν_i per second (Sellmeier's equation).

If the medium is a gas, for which the refractive index is only slightly greater than unity, the dispersion formula can be written as Eq. (15).

$$n = 1 + 2\pi Ne^2 \sum_i \frac{f_i/m_i}{\omega_i{}^2 - \omega^2} \qquad (15)$$

So long as the absorption remains negligible, these equations correctly describe the increase in refractive index as the frequency of the radiation begins to approach the absorption band determined by ω_i or λ_i. They fail when absorption becomes appreciable, since they predict infinitely large values of the refractive index when ω equals ω_i, whereas the refractive index remains finite throughout an absorption band.

The absorption of radiant energy of frequency very close to the characteristic frequency of the medium is formally regarded as the overcoming of a frictional force when the molecular oscillators are set into vibration, related by a proportionality constant g to the velocity of the oscillating particle; g is a damping coefficient for the oscillation. If the refractive index is determined by a single electronic oscillator, the dispersion equation for a gas at radiational frequencies within the absorption band becomes Eq. (16).

$$n = 1 + \frac{2\pi Ne^2}{m} \frac{f(\omega_0{}^2 - \omega^2)}{(\omega_0{}^2 - \omega^2)^2 + \omega^2 g^2} \qquad (16)$$

At the same time an absorption constant κ enters the equations, related to the absorption coefficient α of Eq. (1) by the expression $\kappa = \alpha c/2\omega\mu$. Equation (17) shows the relationship. For a monatomic

$$\kappa = \frac{2\pi Ne^2}{m} \frac{f\omega g}{(\omega_0{}^2 - \omega^2)^2 + \omega^2 g^2} \qquad (17)$$

vapor at low pressure, Nf is about 10^{17} per cubic centimeter, ω_0 is about 3×10^{15} per second, and g is about 10^{11} per second. These data show that, when the frequency of the radiation is not very near ω_0, ωg is very small in comparison with the denominator and the absorption is practically zero. As ω approaches ω_0, κ increases rapidly to a maximum at a radiational frequency very near ω_0 and then falls at frequencies greater than ω_0. When the absorption is relatively weak, the absorption maximum is directly proportional to the oscillator strength f. In terms of the molar extinction coefficient ϵ of Eq. (4), it can be shown that this direct relation holds, as seen in Eq. (18). The integration in Eq. (18) is carried out

$$f = 4.319 \times 10^{-9} \int \epsilon d\bar{\nu} \qquad (18)$$

over the whole absorption spectrum. The integral can be evaluated from the area under the curve of ϵ plotted as a function of wave number $\bar{\nu}\ \mathrm{cm}^{-1} = \nu\ (\mathrm{s}^{-1})/c = 1/\lambda$.

The width of the absorption band for an atom is determined by the value of the damping coefficient g; the greater the damping, the greater is the spectral region over which absorption extends.

The general behavior of the refractive index through the absorption band is illustrated by the dotted portions of **Fig. 2**. The presence of the damping term $\omega^2 g^2$ in the denominator of Eq. (17) prevents

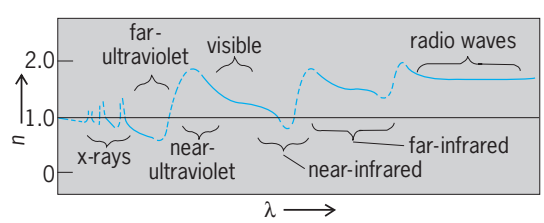

Fig. 2. Complete dispersion curve through the electromagnetic spectrum for a substance. (*After F. A. Jenkins and H. E. White, Fundamentals of Optics, 4th ed., McGraw-Hill, 1976*)

the refractive index from becoming infinite when $\omega = \omega_0$. Its value increases to a maximum for a radiation frequency less than ω_0, then falls with increasing frequency in the center of the band (anomalous dispersion) and increases from a relatively low value on the high-frequency side of the band.

Figure 2 shows schematically how the dispersion curve is determined by the absorption bands throughout the whole electromagnetic spectrum. The dotted portions of the curve correspond to absorption bands, each associated with a distinct type of electrical oscillator. The oscillators excited by x-rays are tightly bound inner elections; those excited by ultraviolet radiation are more loosely bound outer electrons which control the dispersion in the near-ultraviolet and visible regions, whereas those excited by the longer wavelengths are atoms or groups of atoms.

It will be observed in Fig. 2 that in regions of anomalous dispersion the refractive index of a substance may assume a value less than unity; the velocity of light in the medium is then greater than in vacuum. The velocity involved here is that with which the phase of the electromagnetic wave of a single frequency ω advances, for example, the velocity with which the crest of the wave advances through the medium. The theory of wave motion, however, shows that a signal propagated by electromagnetic radiation is carried by a group of waves of slightly different frequency, moving with a group velocity which, in a material medium, is always less than the velocity of light in vacuum. The existence of a refractive index less than unity in a material medium is therefore not in contradiction with the theory of relativity.

In quantum theory, absorption is associated not with the steady oscillation of a charge in an orbit but with transitions from one quantized state to another. The treatment of dispersion according to quantum theory is essentially similar to that outlined, with the difference that the natural frequencies ν_0 are now identified with the frequencies of radiation which the atom can absorb in undergoing quantum transitions. These transition frequencies are regarded as virtual classical oscillators, which react to radiation precisely as do the oscillators of classical electromagnetic theory.

Selective reflection. Nonmetallic substances which show very strong selective absorption also strongly reflect radiation of wavelengths near the absorption bands, although the maximum of reflection is not, in general, at the same wavelength as the maximum absorption. The infrared rays selectively reflected by ionic crystals are frequently referred to as reststrahlen, or residual rays. For additional information on selective reflection. *See* REFLECTION OF ELECTROMAGNETIC RADIATION.

William West

Bibliography. M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; R. W. Ditchburn, *Light*, 3d ed., 1977, reprint 1991; I. S. Grant and W. R. Phillips, *Electromagnetism*, 2d ed., 1991; B. D. Guenther, *Modern Optics*, 1990; E. Hecht, *Optics*, 4th ed., 2001; F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976; E. E. Kriezis, D. P. Chrissoulidis, and A. C. Papagiannakis (eds.), *Electromagnetics and Optics*, 1992; Optical Society of America, *Handbook of Optics*, 2d ed., 4 vols., 1995, 2001; A. Sommerfeld, *Lectures of Theoretical Physics*, vol. 4, 1954.

Abstract algebra

The study of systems consisting of arbitrary sets of elements of unspecified type, together with certain operations satisfying prescribed lists of axioms. Abstract algebra has been developed since the mid-1920s and has become a basic idiom of contemporary mathematics. In contrast to the earlier algebra, which was highly computational and was confined to the study of specific systems generally based on real and complex numbers, abstract algebra is conceptual and axiomatic. A combination of the theories of abstract algebra with the computational speed and complexity of computers has led to significant applications in such areas as information theory and algebraic geometry.

Insight into the difference between the older and the abstract approaches can be obtained by comparing the older matrix theory with the more abstract linear algebra. Both deal with roughly the same area of mathematics, the former from a direct approach which stresses calculations with matrices, and the latter from an axiomatic and geometric viewpoint which treats vector spaces and linear transformations as the basic notions, and matrices as secondary to these. *See* LINEAR ALGEBRA.

Algebraic structures. Abstract algebra deals with a number of important algebraic structures, such as groups, rings, and fields.

An algebraic structure consists of a set S of elements of unspecified nature endowed with a number of finitary operations on S. If r is a positive integer, an r-ary operation associates with every r-tuple $(a_1, a_2, \ldots, a_r)$ of elements a_i in S a unique element $\omega(a_1, a_2, \ldots, a_r)$ in S. It is convenient to consider also nullary operations, that is, the selection of constants from S.

Binary operations. Many of the operations considered in abstract algebra are binary (2-ary) operations, in which case the symbol for the operation is usually written between the elements, or juxtaposition is used to indicate the operation, so that $\omega(a, b)$ is typically represented by $a * b$, $a + b$, or ab, depending on the context. Algebraic structures are classified by properties that the operations of the structure may or may not satisfy. Among the most common properties are:

1. Associativity: A binary operation $*$ of an algebraic structure S is associative if $a * (b * c) = (a * b) * c$ for all elements a, b, c of S.

2. Commutativity: A binary operation $*$ of an algebraic structure S is commutative if $a * b = b * a$ for all elements a, b of S.

3. Existence of an identity: An element e of an algebraic structure S is an identity for the binary

operation $*$ if $a * e = e * a = a$ for all elements a of S.

4. Existence of inverses: Assuming there is an element e of an algebraic structure S that is an identity for the binary operation $*$, an element a of S has an inverse if there is an element a^{-1} of S such that $a * a^{-1} = a^{-1} * a = e$.

5. Distributivity: A binary operation $*$ distributes over a binary operation $+$ of an algebraic structure S if $a * (b + c) = (a * b) + (a * c)$ for all elements a, b, c of S. The property of distributivity thus provides a link between two operations.

Groups. If the structure S is a group, a single binary operation is assumed to satisfy several simple conditions called the group axioms: the associativity of the operation, the existence of an identity element for the operation, and the existence of inverses for the operation. In this case, ab is usually written for $\omega(a,b)$, and the operation is called multiplication; or $a + b$ is written, and the operation is called addition, if the operation ω is commutative. Alternatively, a group may be described as a set with three operations rather than only one. In addition to the associative binary operation mentioned above, it is possible to consider a unary operation providing the inverse of an element, and a nullary operation providing the identity element of the group. Examples of groups include the rigid motions (symmetries) of a square with operation given by function composition, the integers under addition, and the nonzero real numbers under multiplication.

Rings and fields. In the case of a ring R, there are two binary operations, ab and $a + b$, which are assumed to satisfy a set of conditions known as the ring axioms, ensuring the associativity of multiplication and addition, the commutativity of addition, the distributivity of multiplication over addition, and the existence of an additive identity and additive inverses. Examples of rings include the integers, the real numbers, and the continuous real-valued functions of a real variable. *See* RING THEORY.

If axioms asserting the commutativity of multiplication and the existence of a multiplicative identity and multiplicative inverses are adjoined to the ring axioms, the ring is also a field. Examples of fields include the real numbers, the complex numbers, the finite set of integers modulo a prime number, and the rational polynomials with complex coefficients. *See* FIELD THEORY (MATHEMATICS).

Modules. In addition to the intrinsic study of algebraic structures, algebraists are interested in the study of the action of one algebraic structure on another. An important instance is the theory of modules and its special case of vector spaces. A left module over a ring R is defined to be a commutative group M on which R acts on the left in the sense that, given a pair of elements (a,x) where a is in R and x is in M, then the action determines a unique element ax in M. Moreover, the module product ax is assumed to satisfy the module axioms $a(x + y) = ax + ay$, $(a + b)x = ax + bx$, and $(ab)x = a(bx)$, where a and b are arbitrary elements of R, and x and y are arbitrary elements of M. Consideration of the modules admitted by a ring can provide significant information about the structure of the ring. Specializing from an arbitrary ring to a field yields the special type of module called a vector space.

Comparison of similar structures. Along with the study of particular algebraic structures, algebraists are interested in comparing structures satisfying similar axioms or sharing certain properties. Those structures most intimately related to a particular structure are its substructures and quotient structures.

Substructures. A substructure of an algebraic structure S is a subset T that is closed with respect to the operations of S; that is, if ω is any r-ary operation of S and $a_1, a_2, \ldots, a_r$ are elements of the subset T of S, then $\omega(a_1, a_2, \ldots, a_r)$ is again an element of T. In this case, T with operations defined by the operations of S restricted to T becomes an algebraic structure itself. Equivalently S may be referred to as an extension of T. As an example, if S is taken as the commutative group of integers under addition, the subset T of multiples of any particular integer forms a subgroup of S.

Quotient structures. A quotient structure is defined on an algebraic structure S by partitioning the elements of S into a family $\mathbf{Q}$ of (possibly infinitely many) disjoint congruence classes Q_α. As a result, (1) each element of S is a member of one and only one of the subsets Q_α of S, and (2) for any r-ary operation ω of S, no matter what elements $a_1, a_2, \ldots, a_r$ of $Q_{\alpha 1}, Q_{\alpha 2}, \ldots, Q_{\alpha r}$ respectively are chosen, the subset Q_β of S containing $\omega(a_1, a_2, \ldots, a_r)$ is the same. Then the collection of congruence classes $S/\mathbf{Q} = \{Q_\alpha\}$ can be made into an algebraic structure with operations determined by the action of the operations of S on the elements of the congruence classes, called a quotient structure of S. For example, an r-ary operation ω' would act on an r-tuple $(Q_{\alpha 1}, Q_{\alpha 2}, \ldots, Q_{\alpha r})$ by assigning to it the congruence class Q_β if it contains the element $\omega(a_1, a_2, \ldots, a_r)$ for elements $a_1, a_2, \ldots, a_r$ of $Q_{\alpha 1}, Q_{\alpha 2}, \ldots, Q_{\alpha r}$, respectively. An example of a quotient structure is one in which S is the ring of integers under addition and multiplication, and the family of congruence classes $\mathbf{Q}$ consists of the subset E of even integers and the subset O of odd integers. Then the quotient structure $S/\mathbf{Q}$ contains two elements E and O; the operations are given by $E + E = E = O + O$, $E + O = O = O + E$, $E * E = E = E * O = O * E$, and $O * O = O$. These operations symbolically represent the familiar facts that the sum of two even or two odd integers is even, the sum of an even integer and an odd integer is odd, the product of an even integer with either an even or an odd integer is even, and the product of two odd integers is odd.

Comparisons by means of functions. Other comparisons between similar algebraic structures are made by means of functions between the structures that preserve the action of the corresponding operations. Two principal ways of organizing this study are universal algebra and category theory.

1. *Universal algebra*. A considerable part of the study of algebraic structures can be developed in a unified way that does not specify the particular

structure, although deeper results do require specialization to the particular systems whose richness is to a large extent accounted for by their applicability to other areas of mathematics and to other disciplines. The general study of algebraic structures is called universal algebra. A basic notion here is that of a homomorphism of one algebraic structure S to a second one S'. Here there is a one-to-one correspondence $\omega \leftrightarrow \omega'$ between the sets of operations of S and of S' such that ω and ω' are both r-ary for the same $r = 0, 1, 2, 3, \ldots$. Then a homomorphism f from S to S' is a function mapping S to S' such that $\omega(a_1, \ldots, a_r) = \omega'(f(a_1), \ldots, f(a_r))$ for all a_i in S and all the corresponding operations ω and ω'. A substantial part of the theory of homomorphisms of groups and rings is valid in this general setting. Another basic notion of universal algebra is that of a variety of algebraic structures, which is defined to be the class of structures satisfying a given set of identities defined by means of the given compositions (that is, the associative or the commutative laws). Examples are the variety of groups and the variety of rings.

2. *Category theory*. A second approach to the comparison of related structures is that of category theory. The notion of a category is applicable throughout mathematics, for example, in set theory and topology. The notion is formulated in order to place on an equal footing a class of mathematical objects and the basic mappings between these; in contrast is the more elemental focus of universal algebra. A category is defined to be a class of objects (generally mathematical structures) together with a class of morphisms (generally mappings of the structures) subject to certain axioms. An example of a category is the class of all groups as objects and the class of all homomorphisms of pairs of groups as the class of morphisms. Another example is the class of all sets as objects and of mappings of one set into another as the class of morphisms.

As an example of the ever-increasing levels of abstraction available and useful in algebra, once categories have been defined, they can themselves be compared by means of functors, which act as homomorphisms of categories. If C and C' are categories, a functor from C to C' is defined as a mapping F of the objects of C into the objects of C', together with a mapping ϕ of the morphisms of C into the morphisms of C', such that ϕ preserves the compositions and identity morphisms. At this point, functors can be compared via natural transformations, which thus act as homomorphisms of functors.

The concepts of category, functor, and natural transformation appear in many areas of mathematics. Many of the homomorphisms that arise in the study of algebraic structures prove to be instances of natural transformations. The frequency of the occurrences of the naturalness of the homomorphisms has implications for any adequate philosophy of mathematics. Like set theory, category theory can provide a foundation for much of mathematics. Joel K. Haack

Bibliography. M. Artin, *Algebra*, 1995; J. R. Durbin, *Modern Algebra: An Introduction*, 4th ed., 1999; J. Gallian, *Contemporary Abstract Algebra*, 2d ed., 1990; T. W. Hungerford, *Algebra*, rev. ed., 1980, reprint 1993; S. MacLane and G. Birkhoff, *Algebra*, 3d ed., 1987; C. C. Pinter, *A Book of Abstract Algebra*, 2d ed., 1990.

Abstract data type

A mathematical entity consisting of a set of values (the carrier set) and a collection of operations that manipulate them. For example, the Integer abstract data type consists of a carrier set containing the positive and negative whole numbers and 0, and a collection of operations manipulating these values, such as addition, subtraction, multiplication, equality comparison, and order comparison. *See* ABSTRACT ALGEBRA.

Abstraction. To abstract is to ignore some details of a thing in favor of others. Abstraction is important in problem solving because it allows problem solvers to focus on essential details while ignoring the inessential, thus simplifying the problem and bringing to attention those aspects of the problem involved in its solution. Abstract data types are important in computer science because they provide a clear and precise way to specify what data a program must manipulate, and how the program must manipulate its data, without regard to details about how data are represented or how operations are implemented. Once an abstract data type is understood and documented, it serves as a specification that programmers can use to guide their choice of data representation and operation implementation, and as a standard for ensuring program correctness.

A realization of an abstract data type that provides representations of the values of its carrier set and algorithms for its operations is called a data type. Programming languages typically provide several built-in data types, and usually also facilities for programmers to create others. Most programming languages provide a data type realizing the Integer abstract data type, for example. The carrier set of the Integer abstract data type is a collection of whole numbers, so these numbers must be represented in some way. Programs typically use a string of bits of fixed size (often 32 bits) to represent Integer values in base two, with one bit used to represent the sign of the number. Algorithms that manipulate these strings of bits implement the operations of the abstract data type. *See* ALGORITHM; DATA STRUCTURE; NUMERICAL REPRESENTATION (COMPUTERS); PROGRAMMING LANGUAGES.

Realizations of abstract data types are rarely perfect. Representations are always finite, while carrier sets of abstract data types are often infinite. Many individual values of some carrier sets (such as real numbers) cannot be precisely represented on digital computers. Nevertheless, abstract data types provide the standard against which the data types realized in programs are judged.

Example. Data processing problems are rarely so simple that data abstraction is not helpful in their solution. Suppose a program must process data in its

order of arrival but that data may arrive before its predecessors have been completely processed. Data values must wait their turn in storage in a way that preserves their arrival order. Before considering the structure needed to store the data and the algorithms required for creating and maintaining the structure, it is useful to specify an abstract data type for this problem.

What is needed is a well-known abstract data type called a Queue. The carrier set of the Queue abstract data type is the set of all queues, that is, the set of all lists of data values whose elements enter at the rear and are removed from the front. Although many operations might be included in this abstract data type, the following four suffice:

1. The enter operation takes a queue and a data value as parameters and returns the input queue with the new data value added to its rear.
2. The remove operation takes a queue as its parameter and returns the input queue with the data value at its front removed.
3. The front operation takes a queue as its parameter and returns the data value at the front of the queue (the queue is unaltered).
4. The size operation takes a queue as its parameter and returns a whole number greater than or equal to 0 reporting the number of data values in the queue.

These informal operation descriptions should be augmented by specifications that precisely constrain them:

1. For any queue q, if size $(q) = 0$ then remove (q) is invalid. That the operation is invalid is reported as an error.
2. For any queue q, if size $(q) > 0$ then size $(q) =$ size (remove $(q)) + 1$.
3. For any queue q and data value v, size $(q) =$ size (enter $(q,v)) - 1$.
4. For any queue q and data value v, let $q' =$ enter (q,v). Then after an arbitrary sequence of enter and remove operations applied to q' containing exactly size (q) remove operations, and yielding q'', $v =$ front (q'').

The first of these specifications states that nothing can be removed from an empty queue. The second and third state how queues grow and shrink. The final specification captures the crucial property that all data values eventually leave a queue in the order that they enter it. Some reasoning will show that these four specifications completely determine the behavior of queues.

Usefulness. Such specifications of abstract data types provide the basis for their realization in programs. Programmers know which data values need to be represented, which operations need to be implemented, and which constraints must be satisfied. Careful study of program code and the appropriate selection of tests help to ensure that the programs are correct. Finally, specifications of abstract data types can be used to investigate and demonstrate the properties of abstract data types themselves, leading to better understanding of programs and ultimately higher-quality software. *See* COMPUTER PROGRAMMING; SOFTWARE ENGINEERING; SHRINK FIT.

Relation to object-oriented paradigm. A major trend in computer science is the object-oriented paradigm, an approach to program design and implementation using collections of interacting entities called objects. Objects incorporate both data and operations. In this way they mimic things in the real world, which have properties (data) and behaviors (operations). Objects that hold the same kind of data and perform the same operations form a class.

Abstract data values are separated from abstract data type operations. If the values in the carrier set of an abstract data type can be reconceptualized to include not only data values but also abstract data type operations, then the elements of the carrier set become entities that incorporate both data and operations, like objects, and the carrier set itself is very much like a class. The object-oriented paradigm can thus be seen as an outgrowth of the use of abstract data types. *See* OBJECT-ORIENTED PROGRAMMING.

Christopher Fox

Bibliography. J. Guttag, Abstract data types and the development of data structures, *Commun. ACM*, 20(6):396–404, June 1977; P. Thomas, H. Robinson, and J. Emms, *Abstract Data Types*, Oxford University Press, Oxford, 1988.

Acantharea

A class of the phylum Radiozoa in the Actinopoda. The kingdom Protozoa contains 18 phyla. One of the parvkingdoms (a hierarchical classification between kingdom and superphylum that is controversial and not officially recognized as such) is the Actinopoda (originally a class) containing two phyla: Heliozoa and Radiozoa. Within the Radiozoa is the class Acantharea. These marine protozoans possess a nonliving, organic capsular wall surrounding a central mass of cytoplasm. The intracapsular cytoplasm is connected to the extracellular cytoplasm by fine cytoplasmic strands passing through pores in the capsular wall. When viewed with the electron microscope, the capsular wall in some species appears to be made of many layers, each composed of a fine fibrillar network. The skeletons are made of celestite (strontium sulfate) instead of silica. The basic structural elements are 20 rods that pass through the capsule to the center in regular arrangements (polar and equatorial; see **illustration**). An equatorial rod forms a 90° angle with a polar rod, and other groups are arranged with similar exactness. This type of cytoskeleton may be modified by the addition of a latticework, composed of plates, each fused with a skeletal rod. Some genera show a double latticework, concentric with the central capsule. The skeletons do not contribute substantially to the paleontological record in marine sediments since celestite is dissolved by seawater.

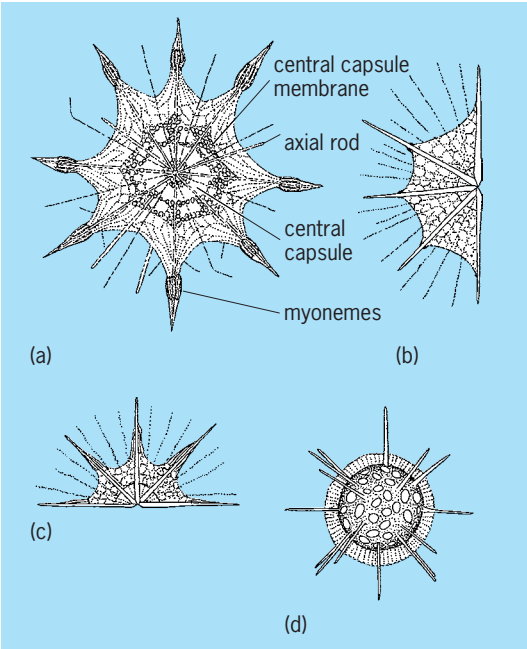

Representatives of the Acantharea; *Acanthometra pellucida*, (*a*) showing axial rods extending through the central capsule, and myonemes (myophrisks) extending from the rods to the sheath; (*b*) contracted myophrisks with expanded sheath; and (*c*) relaxed myophrisks with contracted sheath. (*d*) *Dorotaspis heteropora*, with perforated shell and rods. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)

Dissolution is pronounced below 617 ft (188 m) depth. While the protozoan is still alive, however, the cytoplasm appears to protect the skeleton from dissolution.

Pseudopodia are more or less permanent. A pseudopodium either is composed of an axial rod containing microtubules and a cytoplasmic envelope (forming an axopodium, or small foot; long slender pseudopodia are found in certain protozoa) or is fine-branched (forming a reticulopodium). Mitochondria are always present and contain flattened cristae. No cilia are present in the tropic phase. The central endoplasm contains numerous nuclei and other organelles.

When present, zooxanthellae (algae found within the cytoplasm) are of at least two kinds: Dinophyceae containing trichocysts in the cytoplasm; and a group of algae characterized by numerous discoid plastids, each of which contains an interlamellar pyrenoid. The systematics of the second group is still uncertain.

Myonemes (myophrisks) are significant components of the hydrostatic apparatus and regulate the buoyancy of the Acantharea by expanding or contracting portions of the extracapsular sheath. When observed with the electron microscope, they are composed of microfilaments arranged in a cone around the spine. Microfilaments contain actin, a contractile protein also found in metazoan muscle. Their contraction extends the sheath (illus. *b*); their relaxation flattens it against the rods (illus. *c*).

Although essentially pelagic, Acantharea may move vertically with the help of their hydrostatic apparatus. Symbiont-bearing species (species that contain zooxanthellae) have been reported from depths as great as 13,120 ft (4000 m). Little is known about the ecology or distribution of the Acantharea. The Gulf Stream is rich with these protozoans in the spring and summer but poor during other seasons. When compared to other plankton, Acantharea have a low average percentage of abundance even at peak periods (about 6–7%). Their approximate average abundance in the North Atlantic has been reported as 1–5%. *See* GULF STREAM.

Information about life cycles is very limited. For a few species, flagellate stages, syngamy, and zygotes have been reported. Certain stages of fission have been reported for some primitive types. The taxonomy of the Acantharea is still under review. *See* ACTINOPODEA; CELESTITE; PROTOZOA; SARCODINA; SARCOMASTIGOPHORA; STRONTIUM.

O. Roger Anderson; John P. Harley

Bibliography. T. Cavalier-Smith, Kingdom Protozoa and its 18 phyla, *Microbiol. Rev.*, 57(4):953–994, 1993; M. Grell, in H.-E. Gruner (ed.), *Lehrbuch der Speziellen Zoologie*, 4th ed., Gustav Fischer, Stuttgart, 1980; K. Hausmann and N. Hulsmann, *Protozoology*, Georg Thieme Verlag, New York, 1996.

Acanthocephala

A distinct phylum of helminths, the adults of which are parasitic in the alimentary canal of vertebrates. They are commonly known as the spiny-headed worms. The phylum comprises the orders Archiacanthocephala, Palaeacanthocephala, and Eoacanthocephala. Over 500 species have been described from all classes of vertebrates, although more species occur in fish than in birds and mammals and only a relatively few species are found in amphibians and reptiles. The geographical distribution of acanthocephalans is worldwide, but genera and species do not have a uniform distribution because some species are confined to limited geographic areas. Host specificity is well established in some species, whereas others exhibit a wide range of host tolerance. The same species never occurs normally, as an adult, in cold-blooded and warm-blooded definitive hosts. More species occur in fish than any other vertebrate; however, Acanthocephala have not been reported from elasmobranch fish. The fact that larval development occurs in arthropods gives support to the postulation that the ancestors of Acanthocephala were parasites of primitive arthropods during or before the Cambrian Period and became parasites of vertebrates as this group arose and utilized arthropods for food. *See* ARCHIACANTHOCEPHALA; EOACANTHOCEPHALA; PALAEACANTHOCEPHALA.

Morphology

Adults of various species show great diversity in size, ranging in length from 0.04 in. (1 mm) in some species found in fish to over 16 in. (400 mm) in some mammalian species (**Fig. 1**). Most of the larger

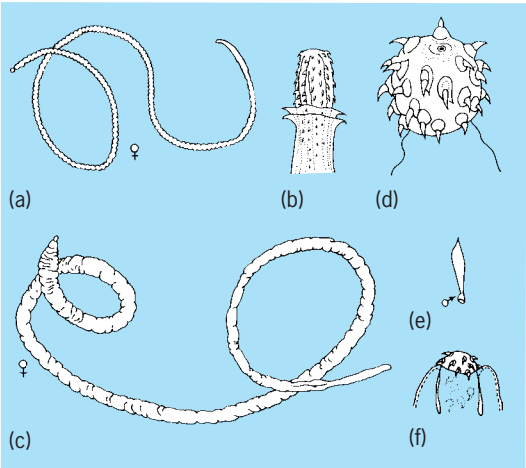

Fig. 1. Various acanthocephalan examples, adult forms drawn to same scale. (*a*) *Moniliformis dubius*, 4–12 in. (10–30 cm) long. (*b*) Proboscis of same. (*c*) *Macracanthorhynchus hirudinaceus*, 10–24 in. (25–60 cm) long. (*d*) Proboscis of same. (*e*) *Oncicola canis*, 0.5–0.8 in. (1–2 cm) long. (*f*) Proboscis of same. (*After A. C. Chandler and C. P. Read, Introduction to Parasitology, 10th ed., Wiley, 1961*)

species are from mammals; however, some mammalian forms are relatively small (*Corynosoma*) even though they parasitize seals and whales. When observed undisturbed in the intestine, the elongate body is distinctly flattened, but on removal to water or saline and after preservation the body becomes cylindrical. Living worms are translucent or milky white, although they may take on a distinctive color from the intestinal contents. The body of both sexes has three divisions: the proboscis armed with hooks, spines, or both; an unspined neck; and the posterior trunk.

Proboscis. The proboscis is the primary organ for attachment of Acanthocephala to the intestinal wall of the host. In some species the proboscis is parallel with the main axis of the body, but frequently it is inclined ventrally. The proboscis may be globular, or elongate and cylindrical in shape and is invariably armed with sharp pointed hooks, spines, or both (Fig. 1). These structures vary in shape and number and are usually arranged radially or in spiral rows. Electron microscopy studies have revealed that proboscis hooks consist of a central core of cytoplasm covered by hard nonliving material. In species with an elongate proboscis the hooks seem to be in longitudinal rows with a quincuncial arrangement.

In most species the proboscis is capable of introversion into a saclike structure, the proboscis receptacle. When in this position, the tips of the hooks are all directed anteriorly and are on the inside of the introverted proboscis. As the proboscis is everted and extended, these hooks engage the mucosa of the host's intestine, with the tips turning posteriorly as they reach their functional position on the exterior of the proboscis. The recurved points of the hooks anchor the worm firmly to the intestinal wall (**Fig. 2**). The proboscis receptacle and neck can be retracted into the body cavity but without inversion.

Within the anterior end of the trunk are found the proboscis receptacle; musculature for retracting the proboscis, receptacle, and neck; and the lemnisci (**Fig. 3**). The proboscis receptacle is attached to the inner surface of the proboscis. Muscles known as inverters of the proboscis are attached to the anterior tip of this structure and pass through the proboscis and receptacle, emerging from the posterior of the receptacle, and continue some distance posteriorly, where they attach to the trunk wall. Usually one inverter is attached dorsally and one ventrally. Contraction of the inverter fibers introverts the proboscis into the receptacle, while further contraction of these fibers pulls the receptacle deep into the body cavity.

Body wall. Electron microscopy studies show that the body wall is covered by a thin mucopolysaccharide known as the epicuticle secreted onto the surface of the worm. Beneath the epicuticle is the thin cuticle with an outer membrane and a homogeneous layer perforated by numerous pores of canals which pass through the striped layer beneath. Many of the canals are filled with an electron-dense material. Below the striped layer is the felt layer composed of many fibrous strands extending in various directions. Mitochondria and numerous vesicles are present in the felt layer. Some vesicles seem to be connected to the canals of the striped layer. The radial layer beneath the felt layer contains fewer fibrous strands, large thin-walled lacunar channels with mitochondria arranged around the channels. Lipids and glycogen are present in the radial layer. Circular and longitudinal muscle layers are found beneath the radial layer separated by plasma membranes. Some investigators consider all body-wall layers beneath the stripped layer as hypodermis.

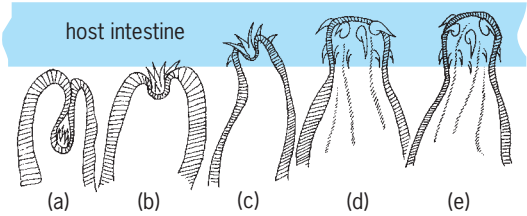

Fig. 2. Diagrams showing mechanics of proboscis attachment of *Neoechinorhynchus emydis* in intestinal wall of a turtle. (*a*) Proboscis fully introverted. (*b*) Most basal hooks in contact with host tissue. (*c–e*) Progressive stages in extroversion of the proboscis. (*After H. J. Van Cleave, Experimental Parasitology, vol. 1, 1952*)

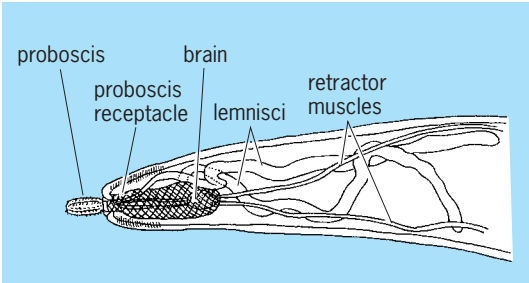

Fig. 3. Anterior end of *Moniliformis dubius*.

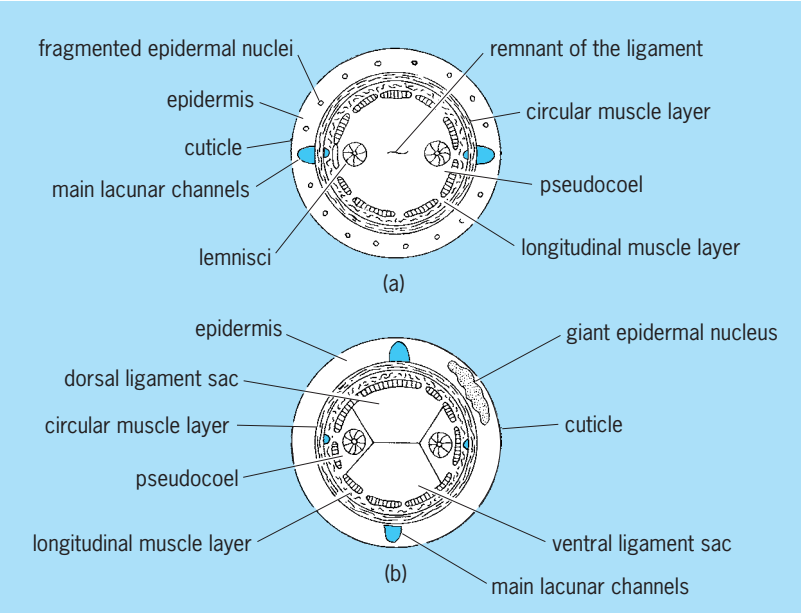

Fig. 4. Cross sections showing body plans of acanthocephalans. (*a*) Palaeacanthocephalan. (*b*) Archiacanthocephalan. (*After L. H. Hyman, The Invertebrates, vol. 3, McGraw-Hill, 1951*)

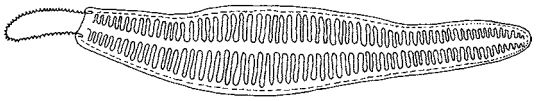

Fig. 5. Lacunar system of *Moniliformis*, dorsal view, showing regular circular branches. (*After L. H. Hyman, The Invertebrates, vol. 3, McGraw-Hill, 1951*)

Giant nuclei. The wall of the trunk has an external cuticula beneath which lies the thick syncytial hypodermis, or subcuticula. The hypodermis contains a relatively small number of giant nuclei which may be round, oval, ameboid, or elongate with a number of lateral branches (**Fig. 4**).

Lacunae. Within the hypodermis is a series of intercommunicating spaces, the lacunar system. Usually the lacunar system is composed of two longitudinal vessels either dorsal and ventral or ventral and lateral. In some species only the dorsal vessel is present. The longitudinal vessels are connected by a series of smaller vessels, the lacunae, which ramify throughout the body. In species exhibiting pseudosegmentation (*Moniliformis*), the regularly spaced lateral lacunae and the body musculature divide the wall into transverse folds (**Fig. 5**). These folds have no effect on the arrangement of internal organs.

Pseudocoele. The body cavity, or pseudocoele, contains all the internal organs, the most conspicuous of which are the reproductive organs enclosed in axial connective-tissue ligament sacs. These ligament sacs are hollow tubes which extend most of the length of the cavity of the trunk. They are single in both males and females of the Palaeacanthocephala, but are divided into dorsal and ventral sacs which communicate anteriorly in females of the other orders. There is no vestige of a digestive system in any stage of the life cycle.

Reproductive system. The reproductive organs of the male consist of a pair of testes and specialized cells, the cement glands. In most species there is a saclike structure behind the cement glands, Saefftigen's pouch, through which run the sperm ducts and the ducts from the cement glands (**Fig. 6*a***). The products of the testes and cement glands are discharged through a penis, which is surrounded by a posteriorly located bell-shaped copulatory bursa which is usually introverted into the posterior extremity of the trunk. At copulation the bursa is extended and applied to the posterior extremity of the female, where it is held firmly in place by the secretions of the cement glands. This material hardens, forming a copulatory cap on the female which is an internal cast of the bursa of the male, and remains attached to the posterior extremity of the female for some time following copulation.

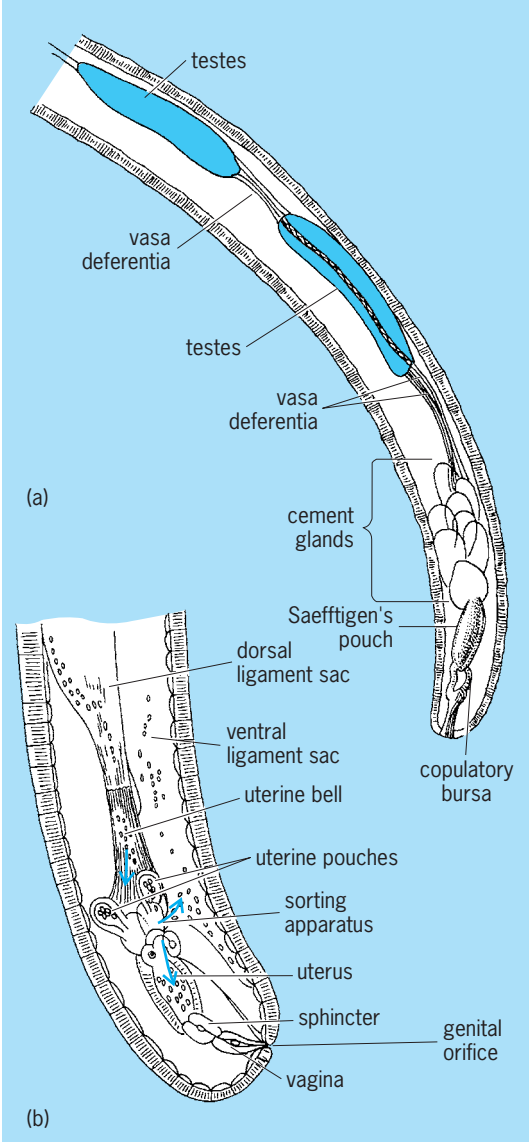

Fig. 6. Reproductive system of *Moniliformis dubius*. (*a*) Posterior end of male. (*b*) Posterior end of female. (*After A. C. Chandler and C. P. Read, Introduction to Parasitology, 10th ed., Wiley, 1961*)

Female Acanthocephala are unique in that the ovary exists as a distinct organ only in the very early stages of development and later breaks up to form free-floating egg balls. The eggs are fertilized as they are released from the egg balls and are retained within the ligament sacs until embryonation is complete. The genital orifice is posterior. A vagina provided with a strong sphincter extends anteriorly from the orifice and a saccular uterus is anterior to the vagina. The anterior end of the uterus is surrounded by a series of guard cells, the selective apparatus. From this extends a funnellike structure, the uterine bell, the broad anterior end of which opens into the body cavity or one of the ligament sacs holding the developing eggs (Fig. 6*b*). During embryonation in the body cavity the eggs acquire a series of membranes and a hard outer shell. Eggs which have not completed embryonation are passed back into the body cavity through special openings in the selective apparatus, whereas eggs which are fully mature are passed into the saccular uterus and eliminated through the vaginal sphincter and genital orifice into the intestinal contents of the host.

Lemnisci. The trunk and presoma are demarcated by an infolding of the cuticula and the presence of two elongate contractile structures, the lemnisci, which arise from the hypodermis and extend posteriorly into the body cavity. The lemnisci are provided with a definite number of nuclei which migrate from the hypodermis into the lemnisci as they are formed during larval development. The function of the lemnisci is unknown. One explanation offered for their function is that they act as reservoirs for the fluid of the lacunar system when the presoma and proboscis are invaginated.

Nervous system. The nervous system is composed of a chief ganglion or brain located within the proboscis receptacle. Most of the nerve trunks are inconspicuous but two of them, the retinacula, associated with muscle fibers, pass through the wall of the proboscis receptacle to innervate the trunk wall.

Excretory system. In members of the Archiacanthocephala modified protonephridial organs are found closely adherent to the reproductive system, but in most species specialized excretory organs are completely lacking. The protonephridial organs consist of a mass of flame bulbs attached to a common stem. The canals unite to form a single canal which joins the sperm duct in the male and the uterus in the female.

Embryology

The Acanthocephala are obligatory parasites throughout their entire life cycle; no known member of this phylum exists as a free-living organism.

Acanthor. The eggs passed in the feces of the host contain a mature embryo, the acanthor, which is surrounded by three membranes (**Fig. 7***a*). The life cycle always involves an intermediate host, which is usually an arthropod. Among these are small crustaceans for parasites of aquatic vertebrates, and grubs, roaches, and grasshoppers for those which parasitize terrestrial vertebrates. The eggs hatch after being ingested by the appropriate arthropod and release the spindle-shaped acanthor, which has a spiny body and is armed with retractile bladelike rostellar hooks used in the penetration of the intestinal wall of the intermediate host (Fig. 7*b*). The central portion of the body of the acanthor contains a column of dense undifferentiated nuclei, the entoblast, from which develop the structures of the ensuing stages. After penetration of the wall of the digestive tract, the acanthor comes to lie beneath the delimiting membrane of the gut wall, becomes quiescent, and begins to grow (Fig. 7*c*). It soon drops free into the host's hemocoele, where it undergoes a gradual transformation into the various larval stages (Fig. 7*d* and *e*).

Acanthella. The term acanthella is applied to the series of stages in which the rudiments of the reproductive organs, lemnisci, proboscis, and proboscis receptacle are formed (Fig. 7*f*). As development progresses, the acanthella becomes elongate, and is surrounded by a hyaline cyst produced by the larva.

Cystacanth. When the proboscis and the proboscis receptacle of the acanthella develop to the point of retractility, the larva becomes infective and is known as the cystacanth (Fig. 7*g* and *h*). The mature or infective cystacanth lies in the hemocoele of the intermediate host with the proboscis and receptacle retracted into the body cavity. The body cavity of the cystacanth contains all of the structures of the adult worm in an immature form.

In the least complicated life cycle the cystacanth remains in the hemocoele of the arthropod until it is ingested by a suitable vertebrate host, in which it excysts and attaches to the intestinal wall by the proboscis and grows to sexual maturity. In some forms the life cycle may be prolonged and complicated by the introduction of one or more transport hosts, usually a vertebrate, in which the cystacanth becomes established as a visceral cyst awaiting ingestion by a suitable vertebrate host in which sexual maturity can be attained.

Cell constancy. The somatic tissues of Acanthocephala are made up of a restricted number of nuclei which is more or less fixed for each tissue. This condition is termed cell constancy or more correctly eutely or nuclear constancy, because the majority of acanthocephalan tissues are syncytial in nature. The nuclei of the early embryonic cells which are destined to become syncytial in nature are large, regular, and spheroidal or elliptical in shape. During the transformation of the acanthella into the cystacanth the nuclei begin to assume the shape and form of the hypodermal nuclei seen in the adult by passing through an integrated series of ovoid, rosette, ameboid, elongate with lateral branches, and arboroid or dendritic forms. These diverse expressions of nuclear shape and modifications are instrumental in maintaining a favorable ratio between nuclear surface and surrounding cytoplasm. At the conclusion of the cleavage process the embryo, or acanthor, contains all of

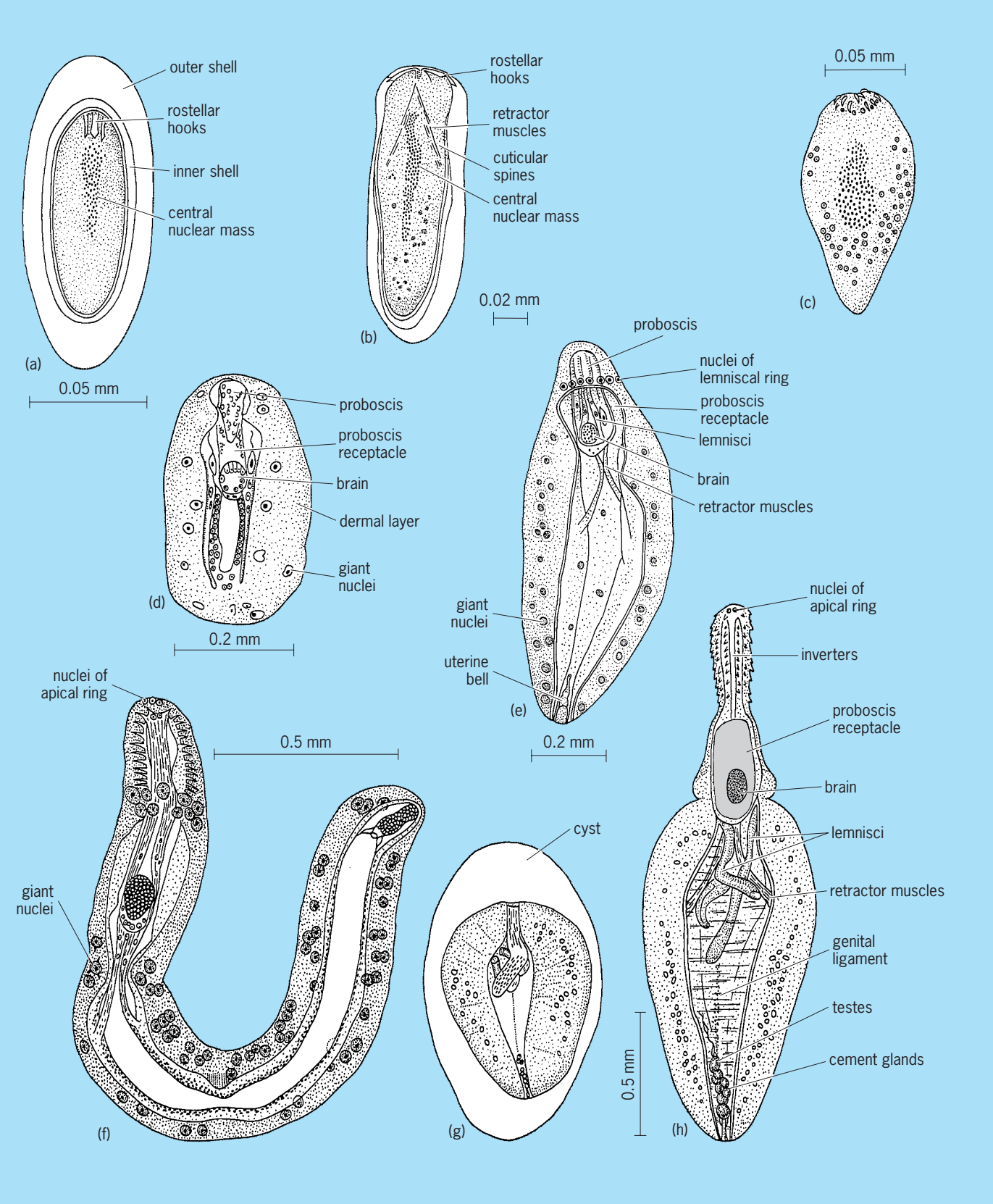

Fig. 7. Life history of *Moniliformis dubius.* (*a*) Mature egg containing acanthor. (*b*) Acanthor in process of escaping from egg shells and membranes. (*c*) Acanthor dissected from gut wall tissue of *Periplaneta americana* 5 days after infection. (*d*) Median sagittal section of larva removed from the body cavity of *P. americana* 29 days after infection. (*e*) Larva removed from body cavity of *P. americana* 42 days after infection. (*f*) Acanthella dissected from its enveloping sheath. (*g*) Encysted cystacanth from body cavity of *P. americana* with proboscis invaginated. (*h*) Cystacanth freed from its cyst and with proboscis evaginated.

the cellular elements of the adult. No mitotic divisions take place, although in the rearrangement of nuclei during larval development certain nuclei may undergo amitotic divisions. The extent of the metamorphoses of the embryonic cells and nuclei varies greatly in the different groups and genera. In the more highly specialized genera, the nuclei of the embryo lose all of their original appearance during larval development. In contrast, the nuclei of the adults of the less specialized genera retain the distinctive embryonic nuclei with little change. Varying changes, intermediate between these two extremes, are observed in other groups. *See* CELL CONSTANCY.

Physiology

Since acanthocephalans lack a gut, they undoubtedly obtain nutrients through the metasomal surface from material in the host's intestinal lumen. The role of the pores in this process is not really known, but it is likely that absorption of nutrients is facilitated by the increased surface area resulting from the plasma membrane lining the pores. It has been suggested that digestion may take place in the region of the presoma, which is in continuous contact with the host tissue. In this area, digestive enzymes are believed to be membrane-bound, allowing products of digestion to be absorbed at the host-parasite interface.

Acanthocephalans contain large quantities of glycogen and in general have a pronounced carbohydrate metabolism. Glycogen is localized in the proboscis, body wall, and muscles. They appear to utilize the Embden-Meyerhoff scheme of phosphorylating glycolysis. Chief waste products of carbohydrate metabolism in *Moniliformis dubius* are succinic, lactic, formic, and acetic acids, along with large quantities of ethanol. It has been suggested that the end products of carbohydrate metabolism excreted by acanthocephalans may be metabolized by the host or may be effective in preventing other helminths from establishing in that area of the host's intestine. *See* CARBOHYDRATE METABOLISM; GLYCOGEN.

A significant Krebs cycle has not been demonstrated in acanthocephalans, but evidence indicates that it may be partially operative in some species. Pyruvate or phosphenalpyruvate formed from glucose by glycolysis may be carboxylated to form malate, which is reduced via fumarate to form succinate. The reduction of fumarate could result in reoxidation of hydrogenated nicotinamide adenine dinucleotide (NADH) under anaerobic conditions, allowing glycolysis and synthesis of adenosine triphosphate (ATP) to be maintained. *See* CITRIC ACID CYCLE.

Donald V. Moore

Bibliography. H. G. Bronn (ed.), *Klassen und Ordnungen des Tierreichs,* vol. 4, 1932–1933; T. Dunagan and D. Miller, Acanthocephala, in *Microscopic Anatomy of Invertebrates*, vol. 4, pp. 299–332, Wiley, 1991; L. H. Hyman, *The Invertebrates*: *Acanthocephala, Aschelminthes, and Entoprocta,* vol. 3, 1951; H. J. Van Cleave, *Acanthocephala of North American Mammals,* 1953.

Acanthodii

A diverse group of early jawed fishes from the Paleozoic Era, usually classified as a distinct class of vertebrates. Most acanthodians have tiny body scales with a cross section like that of an onion, the layers reflecting incremental scale growth. Most acanthodians have strong spines at the leading edges of all fins except for the caudal (tail) fin. Acanthodians mostly were small fishes with large mouths and large eyes, suggesting an active, predatory lifestyle feeding on small prey, including other fishes. The oldest well-preserved examples are of Early Silurian age, and the last surviving examples are of Early Permian age. Primitive forms tended to be from marine environments and to have stouter bodies, larger stouter fin spines, and larger numbers of paired spines in front of the pelvic fins. Later forms are found also in freshwater and estuarine habitats and had more slender bodies with fewer and more slender spines.

Although experts do not currently consider acanthodians to be the most primitive of jawed vertebrates (the Placodermi have that distinction), acanthodians are the earliest jawed vertebrates to be represented in the fossil record by complete specimens. Superficially sharklike in appearance, they are instead currently grouped in the Teleostomi as the closest relatives of the bony fishes (Osteichthyes), with which they share characteristics such as small nasal capsules, large eyes, scales capable of continuous growth (in most species), and (in some specialized forms) three pairs of otoliths (ear stones). *See* OSTEICHTHYES; PLACODERMI; TELEOSTOMI.

Anatomy. Most acanthodians are small fishes, with size ranging from a few centimeters to half a meter. A few species reached lengths of more than 2 m (6.6 ft). The body usually was streamlined, with a blunt head, a terminal mouth, and a long, tapered, upturned tail. The large eyes were set close to the front of the head, behind a pair of small nasal capsules. The head usually was covered by large, flattened scales (tesserae), between which ran sensory canals similar to those in bony fishes. The braincase was ossified in advanced forms but not in primitive forms, where it remains cartilaginous. The ear region included three pairs of growing ear stones in some advanced acanthodians; primitive acanthodians had an open connection (endolymphatic duct) between the inner ear and the external environment, through which tiny sand grains entered and were used for the sense of balance, as in many sharks.

All acanthodians had jaws (sometimes calcified) consisting of a pair each of upper palatoquadrate cartilages and lower meckelian cartilages. In some acanthodians, jawbones, usually with teeth, were attached to these jaw cartilages (**Fig. 1**). In many species, teeth were absent; where present, they were of several types: multiple cusps in a whorl attached to a single base, usually near the front of the mouth; separate teeth in whorl-shaped arrangements along the margins of the mouth; or permanent conical teeth and denticles fixed to the jawbones. Gills were positioned behind the head in a compact gill chamber,

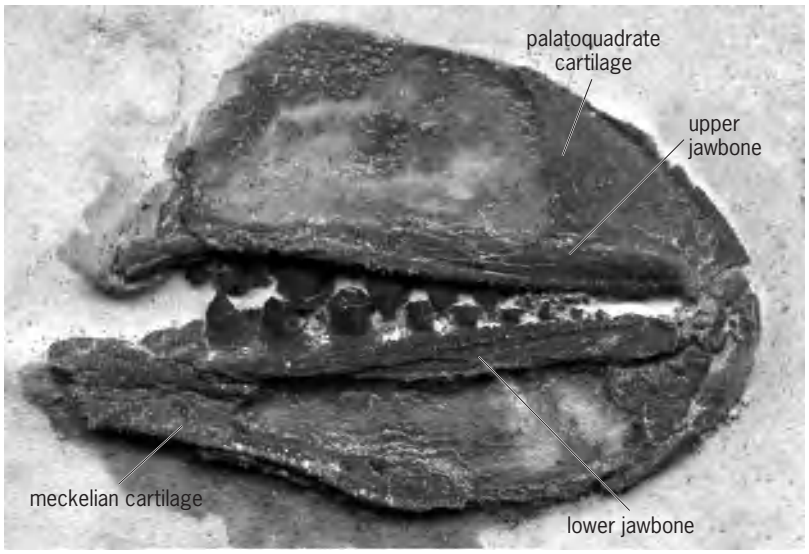

Fig. 1. Jaws of an Early Devonian ischnacanthid from northern Canada. The upper palatoquadrate and lower meckelian cartilages have tooth-bearing jawbones along their facing edges. (*From University of Alberta, Laboratory for Vertebrate Paleontology collection*)

as in bony fishes; a few acanthodians appear to have had multiple external gill slits on each side, but most had ornamented, platelike bony armor that enclosed the gills, leaving only a single pair of gill slits for water flow. *See* GILL.

Primitive acanthodians had two dorsal fins, an anal fin, and paired pectoral and pelvic fins, all with leading-edge spines (**Fig. 2***a–c*). Later forms retained a single dorsal fin (Fig. 2*d*). The tail usually was long and flexible. In primitive forms, a series of paired spines, the prepelvic series, was located along the belly in front of the pelvic fins (Fig. 2*a*). Paired prepectoral spines were present anterior and ventral to the pectoral fins, or were united by bony bases to form a ventral armor between the pectoral fins. These additional paired spines and associated bony plates were reduced or lost in later acanthodians.

Scales of acanthodians were small and continuously growing. Most had an onionskin-like structure internally as successive increments of growth were added to the outside of the scale. The crown of the scale was rhombic in outline, made of dentine, and either smooth or ornamented with ridges, while the bulbous base was composed of cellular or acellular bone. Scales of acanthodians often are abundant and useful for dating and correlating rocks.

Diversity. At one time there was controversy about whether the most primitive acanthodians were those with many fin spines or those with fewer, but more recently it has been recognized that the most primitive acanthodians were those with multiple pairs of stout, heavily ornamented spines, usually classified in the order Climatiiformes. More advanced acanthodians include those in the orders Ischnacanthiformes and Acanthodiformes. Many other acanthodian species are of uncertain relationships because they are known only from isolated scales or fin spines.

Climatiiformes. This order includes the most primitive acanthodians, typically with broad, heavily ridged spines, multiple pairs of prepelvic spines, and either prepectoral spines or well-developed pectoral armor. Teeth, where present, were numerous and arranged in whorl-like sets superficially similar to those of sharks. Examples are the well-known *Climatius, Ptomacanthus, Euthacanthus, Brochoadmones*, and *Lupopsyrus* (Fig. 2*a*). Diplacanthids, including *Diplacanthus*, *Gladiobranchus*, *Uraniacanthus*, and *Tetanopsyrus*, sometimes included as a suborder within Climatiiformes, had a pair of large dermal plates on the side of the head, and had their prepelvic spines reduced to one pair or none (Fig. 2*b*). Climatiiformes appeared in the Silurian; they flourished and then declined in the Devonian. The gyracanthid fishes, sometimes considered to be climatiiform acanthodians, have scales more like those of early chondrichthyans; gyracanthids survived into the Pennsylvanian.

Ischnacanthiformes. These acanthodians had tooth whorls at the front of their mouth as well as tooth-bearing jawbones borne on the upper and lower jaw cartilages (Fig. 1). They had slender fin spines and no prepelvic or prepectoral paired spines. Some ischnacanthiforms were rather large predators. Many species are known only from their toothed jawbones, which often are found separately. Examples include *Poracanthodes* and *Ischnacanthus* (Figs. 1, 2*c*). They are known from the Silurian to the end of the Devonian.

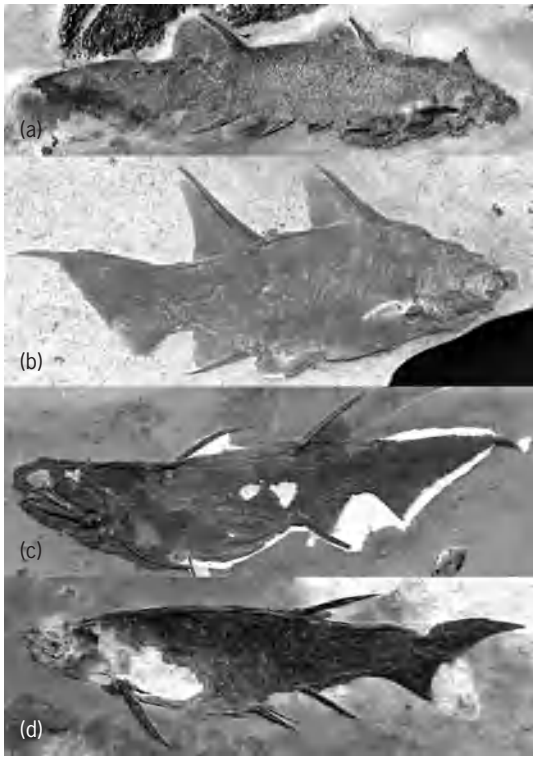

Fig. 2. Examples of acanthodians from the Early Devonian of northern Canada. (*a*) Primitive climatiiform *Lupopsyrus*. (*b*) Diplacanthid *Tetanopsyrus*. (*c*) Ischnacanthiform *Ischnacanthus*. (*d*) Unnamed mesacanthid acanthodiform. (*From the University of Alberta, Laboratory for Vertebrate Paleontology collection*)

Acanthodiformes. These are advanced, streamlined, toothless forms with only one dorsal fin and at most one pair of prepelvic spines. One of the best-known acanthodians of any age is *Acanthodes*, thanks to its ossified internal cranial structures. Later acanthodiforms had long gill rakers and are thought to have been plankton feeders. The group appeared in the Early Devonian and survived until the Early Permian. Examples include *Cheiracanthus* in Cheiracanthidae, *Acanthodes* and *Homalacanthus* in Acanthodidae, and *Triazeugacanthus*, *Melanoacanthus*, and *Mesacanthus* in Mesacanthidae (Fig. 2*d*). Mark V. H. Wilson

Bibliography. R. H. Denison, *Acanthodii: Handbook of Paleoichthyology*, vol. 5, Gustav Fischer Verlag, Stuttgart, 1979; P. Janvier, *Early Vertebrates*, Clarendon Press, Oxford, 1996; J. A. Long, *The Rise of Fishes: 500 Million Years of Evolution*, Johns Hopkins University Press, Baltimore, 1995; J. A. Moy-Thomas and R. S. Miles, *Palaeozoic Fishes*, 2d ed., W. B. Saunders, Philadelphia, 1971.

Acanthometrida

A genus of marine protozoans in the class Acantharea. The kingdom Protozoa contains 18 phyla. One of the parvkingdoms (a hierarchical classification between kingdom and superphylum that is controversial and not officially recognized as such) is the Actinopoda (originally a class) containing two phyla, Heliozoa and Radiozoa. Within the Radiozoa is the class Acantharea, in which the genus *Acanthometrida* is placed. All members are unicellular planktonic marine protozoa with axopodia. Long slender pseudopodia are found in certain protozoans. Mitochondria are always present and contain flattened cristae. No cilia are present in the trophic phase. Members have a skeleton, made up of strontium sulfate (celestite) limited to 20 radially arranged rods that extend from the center, forming a characteristic pattern in which angles are quite exact, as in *Acanthometra* (see **illustration**). Since strontium sulfate is soluble in seawater, no acantharean fossils exist. The central endoplasm contains numerous nuclei and other organelles. The cytoplasm surrounding the spines contains a conical array of contractile microfilaments (myophrisks or myonemes that are Ca^{2+}-activated) containing the contractile protein actin. The microfilaments expand or contract the gelatinous sheath surrounding the cell, a phenomenon apparently responsible for changes in level of flotation. Electron microscopic research with related genera shows that the microfilaments are arranged in two sets of bands: one set runs parallel to the axis of the spine, and the other runs crosswise, thus producing a cross-fibrillar network. Cysts and flagellated swarmers represent two stages of the life cycle, which are not completely known at present. The endoplasm often contains zooxanthellae (small dinoflagellates living in the protozoan's endoplasm). *See* ACANTHAREA; ACTINOPODEA; CELESTITE; PROTOZOA; SARCODINA; SARCOMASTIGOPHORA; STRONTIUM. O. Roger Anderson; John P. Harley

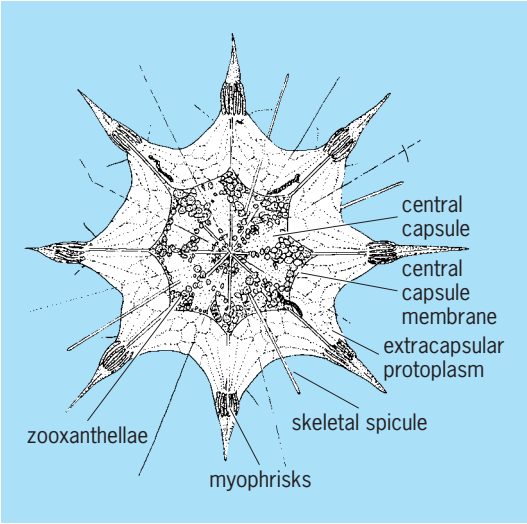

***Acanthometra*. (After L. H. Hyman, The Invertebrates, vol. 1, McGraw-Hill, 1940)**

Bibliography. T. Cavalier-Smith, Kingdom Protozoa and its 18 phyla, *Microbiol. Rev.*, 57(4):953–994, 1993; M. Grell, in H.-E. Gruner (ed.), *Lehrbuch der Speziellen Zoologie*, 4th ed., Gustav Fischer, Stuttgart, 1980; K. Hausmann and N. Hulsmann, *Protozoology*, Georg Thieme Verlag, New York, 1996.

Acanthophractida

An order of Acantharea. In this group of protozoa, skeletons typically include a latticework shell, although the characteristic skeletal rods are recognizable. The latticework may be spherical or ovoid, is fused with the skeletal rods, and is typically concentric with the central capsule. The body is usually covered with a single or double gelatinous sheath through which the skeletal rods emerge. Myonemes (usually a specific number) extend from the gelatinous sheath to each skeletal rod. These marine forms live mostly below depths of 150–200 ft (45–60 m). The order includes *Coleaspis, Diploconus, Dorotaspis*, and many other genera. *See* ACANTHAREA; ACTINOPODEA; PROTOZOA; SARCODINA; SARCOMASTIGOPHORA. Richard P. Hall

Acari

A subclass of Arachnida, the mites and ticks. All are small (0.004–1.2 in. or 0.1–30 mm, most less than 0.08 in. or 2 mm in length), have lost most traces of external body segmentation, and have the mouthparts borne on a discrete body region, the gnathosoma (**Fig. 1**). They are apparently most closely related to Opiliones and Ricinulei.

General characteristics. The chelicerae may be chelate or needlelike. Pedipalps are generally smaller than the walking legs, and are simple or sometimes

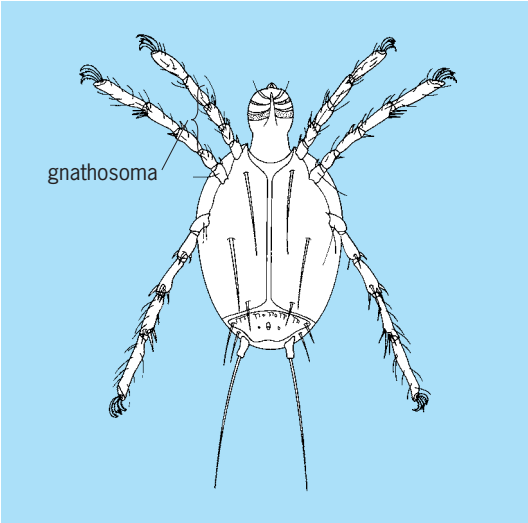

Fig. 1. Ventral view of *Laelaps echidnina*, the spiny rat mite (Parasitiformes), showing major body divisions.

weakly chelate. Legs typically move in the slow walking gait of arachnids; they may, however, be modified by projections, enlarged claws, heavy spines, or bladelike rows of long setae for crawling, clinging to hosts, transferring spermatophores, or swimming. Mites have a variety of integumentary mechanoreceptors and chemoreceptors (often modified setae), and one or two simple eyes (ocelli) frequently occur anterolaterally on the idiosoma (occasionally one anteromedially as well). The ganglia of the central nervous system are coalesced into a single "brain" lying around the esophagus. Excretion is by guanine crystals secreted into the hindgut and voided through a ventral anus. Respiration is by a tracheal network fed by one to four pairs of spiracles that vary in position. A simple dorsal heart is present in a few larger forms.

Life cycle. The typical acarine life cycle consists of six instars: egg; six-legged larva (**Fig. 2**); three nymphal instars (protonymph, deutonymph, and tritonymph); and adult, usually similar to active nymphs. Variations include suppression of the larva, one or more inactive nymphs for internal structural reorganization or phoresy, and change in number of nymphal stages. Male spermatophores may be deposited on the substrate, transferred to the female gonopore by modified legs or chelicerae, or inserted directly by a male copulatory organ. Females are normally oviparous, depositing 4–10,000 eggs singly or in clusters. The generation time varies from 4 days for some terrestrial forms to as much as 2 years in some fresh-water species.

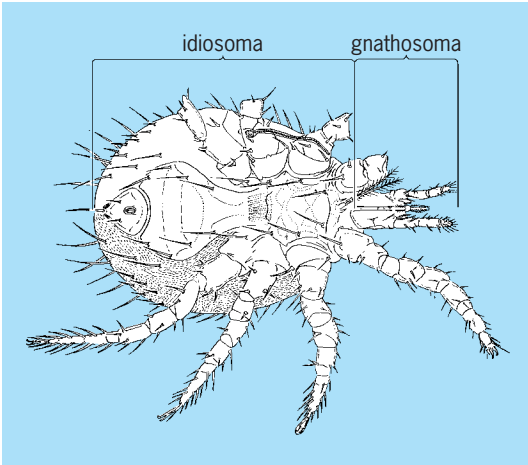

Fig. 2. Ventral view of a *Hygrobates* larva (Acariformes) modified for aquatic life.

Ecology and feeding. Acarine habitats and feeding habits are much more diverse than in other arachnids. Mites are ubiquitous, occurring from oceanic trenches below 13,000 ft (4000 m) to over 21,000 ft (6300 m) in the Himalayas and suspended above 3300 ft (1000 m) in the atmosphere; there are species in Antarctica. Soil mites are taxonomically varied, show the least specialized adaptations to habitat, and are frequently well-sclerotized predators. Inhabitants of stored grains, cheese, and house dust (many causing allergic reactions in humans) are also relatively unspecialized. The dominant forms inhabiting mosses are heavily sclerotized beetle mites (Oribatei), some of which can curl into a ball to reduce moisture loss from the ventral surface. Flowering plant associates include spider mites (Tetranychidae), which suck the contents of leaf epidermal cells, and gall mites (Eriophyidae), which cause a neoplastic growth of host tissue that provides both shelter and food. Fungivores are usually weakly sclerotized inhabitants of moist or semiaquatic habitats, while the characteristic fresh-water mites include sluggish crawlers, rapid swimmers, and planktonic drifters, all with reduced idiosomal setation. Mites associated with invertebrate animals include internal, external, and social parasites as well as inactive phoretic stages of Insecta, Crustacea, Myriapoda, Chelicerata, Mollusca, and Parazoa. A similar diversity of species are parasites or commensals of vertebrates; microhabitats include the nasal mucosa and lungs of reptiles and mammals, burrows in the skin of humans and most other vertebrates (scabies and mange mites), hair follicles (Demodicidae), feathers (both externally and within the feather shaft), fur (bat mites), and host nests and burrows. Some parasitic mites are disease vectors.

Origin and classification. Mites appear in the Devonian among the first true arachnids and thus apparently diverged early from the basic chelicerate stock. Differences in optical activity of the setae and degree of idiosomal segmentation suggest the possibility of a diphyletic origin. Over 30,000 species have been described, and it is estimated that as many as 500,000 may exist. *See* ARACHNIDA. David Barr

Bibliography. D. R. Cook, *Water Mite Genera and Subgenera*, Mem. Amer. Entomol. Inst. 21, 1974; A. Kaestner, *Invertebrate Zoology*, vol. 2: *Arthropod Relatives, Chelicerata, Myriapoda*, 1968; G. W. Krantz, *A Manual of Acarology*, 2d ed., 1978; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; A. E. Treat, *Mites of Moths and Butterflies*, 1975.

Accelerating universe

Cosmic expansion is speeding up. In 1998, astronomers showed evidence from exploding stars that the expansion of the universe is not slowing down because of the gravity of matter in the universe. By observing the light from distant supernovae, astronomers can measure the distance to each explosion. By studying the redshift in the supernova's spectrum, they can measure the expansion of the universe since the light was emitted and, from a set of these measurements, can trace the history of cosmic expansion. Instead of seeing a slowing universe, as expected, observations indicate there is a component of the universe, called the dark energy, that makes cosmic expansion speed up. Dark energy comprises about two-thirds of the energy density of the universe. While it may be related to Albert Einstein's controversial cosmological constant, its true nature remains to be discovered by observation and thought. *See* COSMOLOGICAL CONSTANT; DARK ENERGY; SUPERNOVA.

Astronomers have known since the 1920s that galaxies are separated by millions of light-years and that the space between them is stretching, moving the galaxies apart: the nearest galaxies are moving away from us slowly and distant galaxies are receding more rapidly. This relation was discovered empirically by Edwin Hubble using the 100-in. (2.5-m) telescope at the Mount Wilson Observatory. If our patch of the universe is not unique, the simplest explanation for this observation is that the universe as a whole is expanding in all directions. What is new about the observations made in the 1990s is that the distances and times probed by the supernovae are large enough for the effects of cosmic acceleration to be measurable.

Since light travels at a finite speed, the history of cosmic expansion can be traced by looking at distant objects, whose light left their source hundreds of millions or even billions of years in the past. This procedure makes a telescope into a time machine capable of showing how the expansion of the universe has changed over time. Einstein's general theory of relativity, applied in the cosmic setting, shows that the presence of matter in the universe should lead to the gradual slowing of cosmic expansion. Since the 1950s, astronomers had sought to measure this deceleration of the universe, hoping to use it as a guide to the density of matter in the universe and as a predictor of the future behavior of cosmic expansion—whether the universe would coast outward indefinitely or would be drawn back to a dense state like the big bang from which it emerged 14×10^9 years ago. *See* BIG BANG THEORY; RELATIVITY.

In the 1990s, the perfection of accurate distance-measuring techniques based on supernova explosions and the development of powerful observing techniques made it possible to carry through the astronomers' quest to measure changes in the cosmic expansion rate. Instead of finding that the expansion has slowed over time because of the decelerating effects of gravity, as most expected, the new results from two teams showed the opposite: we live in an accelerating universe. Contrary to all expectation, the universe's expansion is speeding up. Subsequent work on supernovae, on the clustering of galaxies, and on the faint cosmic background light left over from the big bang itself converge on the same solution: the universe today is dominated by a strange dark energy whose role in fundamental physics is certainly important, but of which our present understanding is very limited. *See* COSMIC BACKGROUND RADIATION.

Expanding universe. Before 1920, astronomers thought that the Milky Way Galaxy, of which the Sun is an inconspicuous member, constituted the entire universe. When Einstein applied his recently developed theory of gravitation, the general theory of relativity, to the universe in 1916, he was guided by Willem DeSitter, a leading astronomer, to think of a static system. Einstein found that a static solution to his equations could be constructed if he added in a term that we now call the cosmological constant. This term had the effect of balancing out the attractive force of gravity with a kind of cosmic repulsion to produce a static universe.

Hubble, working at the Mount Wilson Observatory, showed that the distances to what he called the spiral nebulae were much larger than the extent of the Milky Way, and that it was more correct to think of these as independent systems, galaxies, each equivalent to the Milky Way. He did this by measuring the apparent brightness of stars whose intrinsic brightness he knew—Cepheid variables, whose intrinsic brightness can be found by measuring their periods of variation. The fainter a star appears, for a given intrinsic brightness, the more distant it must be. By observing this, a particular type of star, Hubble was able to show that the distances to even the nearest nebulae were a few million light-years, far larger than the size of the Milky Way, whose own size also had to be correctly measured. *See* CEPHEID; GALAXY, EXTERNAL; MILKY WAY GALAXY.

Hubble was also able to use measurements of the spectra of galaxies to show that their spectra have a redshift, the signature that they are receding from us. By combining the measurements of distance and of recession, Hubble discovered that the more distant galaxies are moving away from us more rapidly. The modern interpretation of these observations is cosmic expansion, with the universe stretching out in all directions at a rate of about 1 part in 14,000,000,000 of its size each year. *See* HUBBLE CONSTANT; REDSHIFT.

Naturally, Einstein had to adjust to this new picture of the cosmos. He had inserted his cosmological constant to make a static world, but later observations showed that the world was not static. Einstein was then inclined to banish the cosmological constant, which he knew was not mathematically justified and which he never did like much anyway. But DeSitter was not so sure this was a good idea. He thought it was possible that the universe had started out slowly but had accelerated as a result of the repulsive properties of the cosmological constant, denoted by Λ

Fig. 1. Supernova, 1994D. This type Ia supernova is in a galaxy at a distance of about 5×10^7 light-years in the Virgo cluster of galaxies. For a month, the light from a single exploding white dwarf is as bright as 4×10^9 stars like the Sun. (*P. Challis, Center for Astrophysics/STScI/NASA*)

(lambda). By 1932, Einstein and DeSitter were ready to stop discussing Λ, leaving that for future generations. In a joint paper they said of the cosmological constant, "An increase in the precision of data. . . will enable us in the future to fix its sign and determine its value."

Type Ia supernovae. That future arrived in 1998, 66 years later. The measurements on which the evidence for cosmic acceleration rests are similar to those carried out by Hubble: except that the stars being used are a million times brighter and the distances to the galaxies in which they are located are a thousand times larger. The claim that the universe is accelerating is so extraordinary that it is worth examining the evidence on which it is based: the brightness of exploding stars astronomers call type Ia supernovae (**Fig. 1**).

Type Ia supernovae are a particular type of exploding star that can be recognized by its spectrum. They are found in all types of galaxies and are thought to be the result of a sudden thermonuclear burning wave that propagates through a white dwarf star. This flame burns the carbon and oxygen of the star's interior into radioactive nickel. The subsequent decay into cobalt and iron releases energy gradually as the star is blown apart. For about a month, a type Ia supernova shines as brightly as 4×10^9 stars like the Sun shining simultaneously. This property makes them visible over very large distances, large enough to detect the change in cosmic expansion rate that has taken place while the light was on its way from a type Ia supernova halfway across the universe.

But being bright is only half the story. In order to measure cosmic distances from apparent brightness with good precision, the ideal tool would be a "standard candle" for which all the objects had the same intrinsic brightness. Then a measurement of the apparent brightness would give the distance. Type Ia supernovae are pretty good standard candles: the intrinsic light output at their peak luminosity varies by only a factor of 3 or so. This uniform behavior probably is connected to the theoretical prediction that there is a well-defined upper mass limit for white dwarfs of about 1.4 solar masses that is set by the quantum-mechanical properties of electrons. *See* WHITE DWARF STAR.

Even better precision comes from empirical approaches to type Ia supernovae. It turns out that the brightest supernovae have distinctly slower declines in brightness after they reach their peak, while the dimmer type Ia supernovae decline more rapidly. By measuring the rise and fall in the brightness of a supernova, which does not depend on its distance, one can tell whether an extrabright supernova or an extradim one is being observed, and take that information into account in estimating the distance. When the shape of the light curve is taken into account, the uncertainty in the distance to a single supernova shrinks to under 10%, making this type of supernova the very best distance-measuring tool for surveying the universe (**Fig. 2**). *See* LIGHT CURVES.

Technological advances. The disadvantage of type Ia supernovae is that they are rare. In a typical large galaxy, the rate of type Ia supernova explosions is about one in a century. In our own Milky Way Galaxy, the last event that was probably a type Ia supernova was observed by Tycho Brahe in 1572. So if you want to find and measure many supernovae to survey the history of cosmic expansion, you need either to be very patient or to be able to search many galaxies for the desired events.

This point is where technology has made great strides. The electronic light detectors used in modern telescopes are charge-coupled devices (CCDs), similar to those found in digital cameras. These are

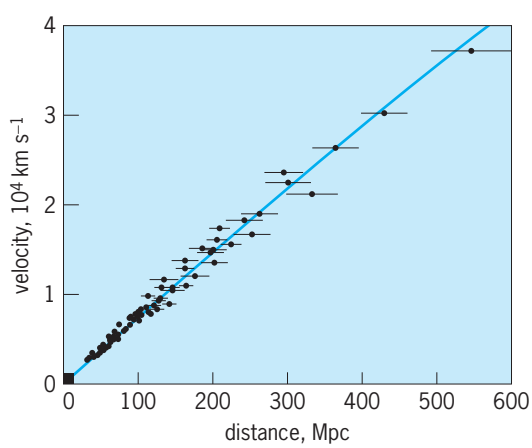

Fig. 2. Modern Hubble diagram for 95 type Ia supernovae. Recession velocity is plotted against distance in megaparsecs (1 Mpc = 3.26×10^6 light-years = 3.09×10^{19} km). The points represent supernovae up to 2×10^9 light-years away, with recession velocities up to one-tenth of the speed of light. The small scatter about the diagonal line shows that type Ia supernovae make excellent standard candles for measuring distances in the universe. This nearby patch does not show the effects of acceleration. (*Based on data from Saurabh Jha*)

more than 100 times more sensitive to light than the photographic plates used in Hubble's time. As the technology for fabricating computer chips has improved, the CCDs available to astronomers have grown larger as well as more sensitive, so that now arrays of 1000 megapixels are being applied to supernova studies and still larger cameras are in devlopment. These large and sensitive devices allow a modern telescope to image many thousands of galaxies in a single exposure. By coming back night after night to the same part of the sky, supernova searches in a single night gather the data needed to find many of these rare events (**Fig. 3**). *See* ASTRONOMICAL IMAGING; CHARGE-COUPLED DEVICE.

But gathering the light is only part of the story of finding supernovae to do cosmology. You also need to detect the new starlike events among the many fuzzy galaxies that do not change from night to night. This is the place where advances in computer hardware and software have turned the task of supernova searching from handicraft into something on an industrial scale. By using a fast computer to subtract digital images obtained last month (or last year) from the ones obtained tonight, it is relatively straightforward to sift out overnight a handful of new objects from the hundreds of thousands that remain unchanged.

Measuring the apparent brightness of these supernovae and the shape of their rise and fall over a month gives enough information to measure the distance to a supernova. Measuring the spectrum of the supernovae can show whether it is of type Ia and measure the redshift. By measuring the redshift at various distances, one can trace the history of cosmic expansion to look for the effects of slowing that results from gravity or acceleration caused by dark energy.

The first attempt at this type of searching was in 1988 by a Danish group. Because only small CCD chips were available, they published data on only one type Ia supernova. But the Supernova Cosmology Project based at Lawrence Berkeley Laboratory developed effective search techniques, and by 1997 they had a preliminary result. Their first indication, based on seven supernovae found in 1994 and 1995, was that the universe was slowing down because of the braking effects of dark matter. In 1998, the High-Z Supernova Search Team published a result indicating the opposite: that the universe was accelerating. Results from the Supernova Cosmology Project published in 1999 showed the same thing: we live in a universe where the rate of cosmic expansion is increasing as the universe grows older.

Subsequent observations have confirmed this latter result, and both teams agree. The samples have grown larger, the span of cosmic time probed by the supernova observations has increased, and methods for dealing with the pernicious effects of absorption by dust have been developed. One especially interesting result comes from a supernova search and follow-up carried out using the *Hubble Space Telescope* both to find and to get the light curves and spectra of very distant supernovae. This work

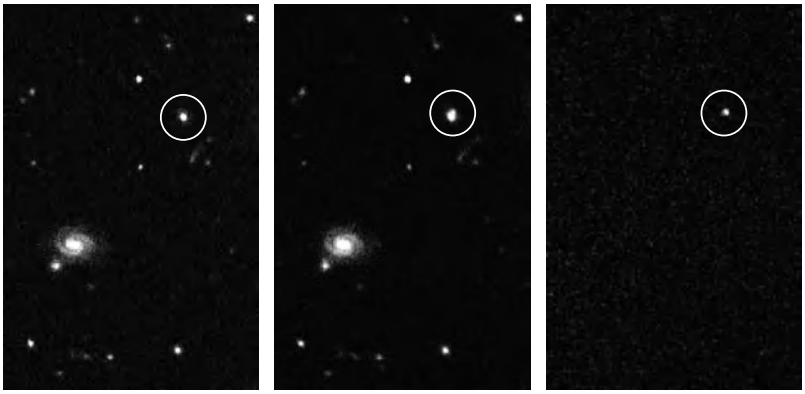

Fig. 3. Subtracting images to find supernovae. The image from a month ago (epoch 1) is subtracted from last night's image (epoch 2) to reveal a new supernova. The image area shown is about 1/1000 of the full area provided by the CCD camera. Rapid processing of dozens of image pairs demands nimble software and a dedicated team of searchers. (*Brian Schmidt, Australian National University*)

showed that cosmic acceleration is a comparatively recent phenomenon, acting over only the last 5×10^9 years or so. Supernovae discovered at larger distances than 5×10^9 light-years show that the universe was probably slowing down when it was younger. So there is evidence that the universe was slowing down when it was young and dense, presumably resulting from the presence of dark matter, but then there came a shift, about 9×10^9 years after the big bang (and 5×10^9 years before the present), from deceleration to acceleration. In the lexicon of elementary physics, change in position is velocity, change in velocity is acceleration, and change in acceleration is called jerk. This program with the *Hubble Space Telescope* has produced evidence for the cosmic jerk. *See* DARK MATTER; HUBBLE SPACE TELESCOPE.

Dark energy? By combining measurements of supernova distances with results from galaxy clustering and from fluctuations in the cosmic microwave background, consensus has emerged that the universe at present has about 73% of its energy density in dark energy, about 23% in dark matter, and only 4% in the form of atoms made of neutrons, protons, and electrons. While it is satisfying that there is a single solution that matches a wide variety of current data, an honest appraisal is that we do not know what the dark energy is or what the dark matter is.

The dark energy could be a form of Einstein's cosmological constant. The cosmological constant has the right property: a tendency to make the universe expand faster over time. This tendency is why DeSitter thought the cosmological constant might be responsible for the cosmic expansion observed by Hubble. But the modern view of the cosmological constant is that it has its origin in the quantum properties of the vacuum. When the same principles that work so well for electromagnetism are applied to gravitation, the result is a quantitative disaster of unprecedented proportions. The observed dark energy is about 10^{120} times smaller than the predicted amount.

Another possibility is that the dark energy is not something that is constant in time, as the cosmological constant would be, but something that has been changing its value as the universe unfolds. This property could be detected by observations of distant supernovae.

The next round of observational programs for distant supernovae is aimed at observing the history of cosmic expansion with enough precision to tell the difference between a dark energy that does not change (like the cosmological constant) or something that varies with time. Large samples of supernovae at moderate distances and moderate samples of supernovae at large distances are being constructed. The goal is to pin down the properties of the dark energy well enough to tell if it is different in any way from a cosmological constant with a very small numerical value. So far, no significant deviation has been found, but the samples published to date are small. Dark energy is so new, so important, and so mysterious that it will be the focus of strenuous efforts in observational astronomy for decades to come. *See* COSMOLOGY; UNIVERSE.

Robert P. Kirshner

Bibliography. D. Goldsmith, *The Runaway Universe*, Perseus Books, Cambridge, MA, 2000; R. P. Kirshner, *The Extravagant Universe: Exploding Stars, Dark Energy, and the Accelerating Cosmos*, Princeton University Press, 2002; L. Krauss, *Quintessence*, Basic Books, New York, 2000.

Acceleration

The time rate of change of velocity. Since velocity is a directed or vector quantity involving both magnitude and direction, a velocity may change by a change of magnitude (speed) or by a change of direction or both. It follows that acceleration is also a directed, or vector, quantity. If the magnitude of the velocity of a body changes from v_1 ft/s to v_2 ft/s in t seconds, then the average acceleration a has a magnitude given by Eq. (1). To designate it fully the direction should be

$$a = \frac{\text{velocity change}}{\text{elapsed time}} = \frac{v_2 - v_1}{t_2 - t_1} = \frac{\Delta v}{\Delta t} \qquad (1)$$

given, as well as the magnitude. *See* VELOCITY.

Instantaneous acceleration is defined as the limit of the ratio of the velocity change to the elapsed time as the time interval approaches zero. When the acceleration is constant, the average acceleration and the instantaneous acceleration are equal.

If a body, moving along a straight line, is accelerated from a speed of 10 to 90 ft/s (3 to 27 m/s) in 4 s, then the average change in speed per second is $(90 - 10)/4 = 20$ ft/s or $(27 - 3)/4 = 6$ m/s in each second. This is written 20 ft per second per second or 20 ft/s^2, or 6 m per second per second or 6 m/s^2. Accelerations are commonly expressed in feet per second per second, meters per second per second, or in any similar units.

Whenever a body is acted upon by an unbalanced force, it will undergo acceleration. If it is moving in a constant direction, the acting force will produce a continuous change in speed. If it is moving with a constant speed, the acting force will produce an acceleration consisting of a continuous change of direction. In the general case, the acting force may produce both a change of speed and a change of direction.

R. D. Rusk

Angular acceleration. This is a vector quantity representing the rate of change of angular velocity of a body experiencing rotational motion. If, for example, at an instant t_1, a rigid body is rotating about an axis with an angular velocity ω_1, and at a later time t_2, it has an angular velocity ω_2, the average angular acceleration $\overline{\alpha}$ is given by Eq. (2), expressed

$$\overline{\alpha} = \frac{\omega_2 - \omega_1}{t_2 - t_1} = \frac{\Delta\omega}{\Delta t} \qquad (2)$$

in radians per second per second. The instantaneous angular acceleration is given by $\alpha = d\omega/dt$.

Consequently, if a rigid body is rotating about a fixed axis with an angular acceleration of magnitude α and an angular speed of ω_0 at a given time, then at a later time t the angular speed is given by Eq. (3).

$$\omega = \omega_0 + \alpha t \qquad (3)$$

A simple calculation shows that the angular distance θ traversed in this time is expressed by Eq. (4).

$$\theta = \overline{\omega} t = \left[\frac{\omega_0 + (\omega_0 + \alpha t)}{2}\right] t = \omega_0 t + {}^1/_2 \alpha t^2 \qquad (4)$$

In the **illustration** a particle is shown moving in a circular path of radius R about a fixed axis through O with an angular velocity of ω radians/s and an angular acceleration of α radians/s^2. This particle is subject to a linear acceleration which, at any instant, may be considered to be composed of two components: a radial component $\mathbf{a}_r$ and a tangential component $\mathbf{a}_t$.

Radial acceleration. When a body moves in a circular path with constant linear speed at each point in its path, it is also being constantly accelerated toward the center of the circle under the action of the force required to constrain it to move in its circular path. This acceleration toward the center of path is called radial acceleration. In the illustration the radial acceleration, sometimes called centripetal acceleration, is shown by the vector $\mathbf{a}_r$. The magnitude of its value is v^2/R, or ω^2/R, where v is the instantaneous linear

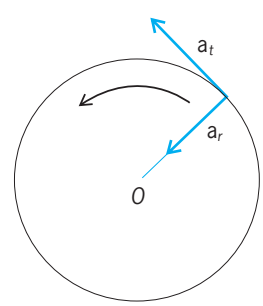

Radial and tangential accelerations in circular motion.

velocity. This centrally directed acceleration is necessary to keep the particle moving in a circular path.

Tangential acceleration. The component of linear acceleration tangent to the path of a particle subject to an angular acceleration about the axis of rotation is called tangential acceleration. In the illustration, the tangential acceleration is shown by the vector $\mathbf{a}_t$. The magnitude of its value is αR. *See* ACCELERATION MEASUREMENT; ROTATIONAL MOTION.

C. E. Howe; R. J. Stephenson

Acceleration analysis

A mathematical technique, often done graphically, by which accelerations of parts of a mechanism are determined. In high-speed machines, particularly those that include cam mechanisms, inertial forces may be the controlling factor in the design of members. An acceleration analysis, based upon velocity analysis, must therefore precede a force analysis. Maximum accelerations are of particular interest to the designer. Although an analytical solution might be preferable if exact maxima were required, graphical solutions formerly tended to be simpler and to facilitate visualization of relationships. Today, for advanced problems certainly, a computer solution, possibly one based on graphical techniques, is often more effective and can also produce very accurate graphical output. *See* FORCE; KINEMATICS; MECHANISM.

Accelerations on a rigid link. On link OB (**Fig. 1***a*) acceleration of point B with respect to point O is the vector sum of two accelerations: (1) normal acceleration $A^n{}_{BO}$ of B with respect to O because of displacement of B along a path whose instantaneous center of curvature is at O, and (2) tangential acceleration of $A^t{}_{BO}$ of B with respect to O because of angular acceleration α.

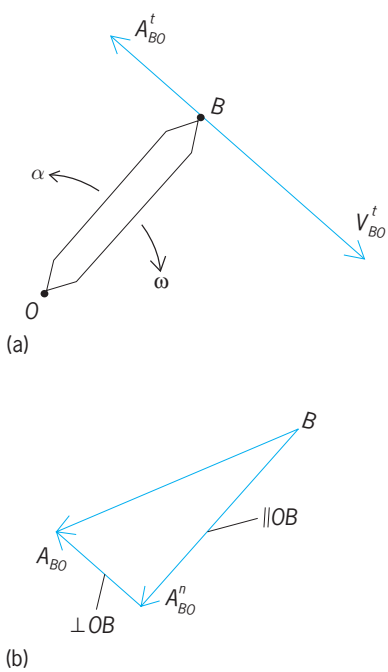

Fig. 1. Elementary condition. (*a*) Point *B* on rigid member rotates about center *O*. (*b*) Vector diagram shows normal, tangential, and resultant accelerations.

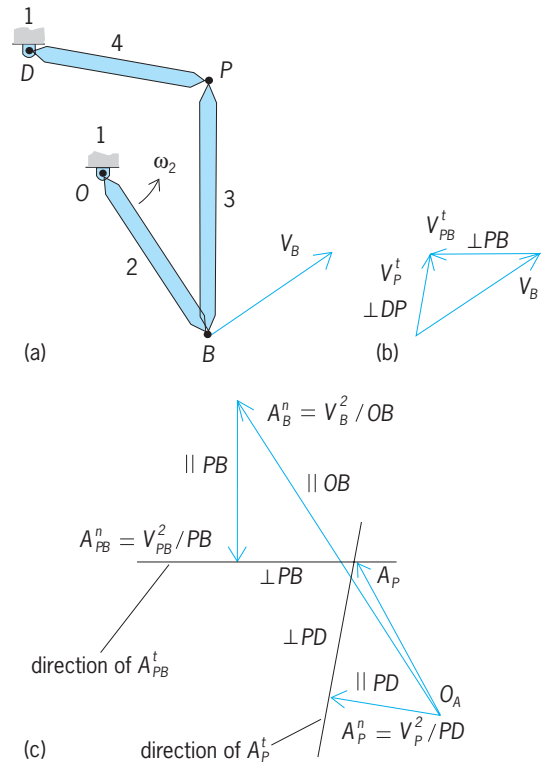

Fig. 2. Four-bar linkage. (*a*) Given are the linkages 2, 3, and 4 between fixed supports 1–1 and the velocity of point *B*, $\omega_2 = 0$. (*b*) A vector polygon is used to determine the velocity at point *P*. (*c*) A graphic solution then finds acceleration at *P*.

For conditions of Fig. 1*a* with link OB rotating about O at angular velocity ω and angular acceleration α, the accelerations can be written as Eqs. (1) and (2). The vector sum or resultant is A_{BO} (Fig. 1*b*).

$$A^n_{BO} = (OB)\omega^2 = V^2_B/(OB) \tag{1}$$

$$A^t_{BO} = (OB)\alpha \tag{2}$$

Accelerations in a linkage. Consider the acceleration of point P on a four-bar linkage (**Fig. 2***a*) with $\alpha_2 = 0$ and hence $\omega_2 = k$, the angular velocity of input link 2. First, the velocity problem is solved yielding V_B and, by Fig. 2*b*, V_{PB}. Two equations can be written for A_P; they are solved simultaneously by graphical means in Fig. 2*c* by using Fig. 2*b*; that is, normal accelerations of P with respect to B and D are computed first. Directions of tangential acceleration vectors A^n_{PB} and A^n_P are also known from the initial geometry. The tip of vector A_P must lie at their intersection, as shown by the construction of Fig. 2*c*. *See* FOUR-BAR LINKAGE.

Explicitly, acceleration A_P of point P is found on the one hand by beginning with acceleration A_B of point B, represented by Eq. (3). To this acceleration

$$A_p = A_B \mapsto A_{PB} \tag{3}$$

is added vectorially normal acceleration A^n_{PB} and tangential acceleration A^t_{PB}, which can be written as Eq. (4). Also for link 2, $\alpha_2 = 0$ and $A^t_B = 0$; and A_B need not be split up.

$$A_P = A_B \mapsto A^n_{PB} \mapsto A^t_{PB} \tag{4}$$

On the other hand, A_P can also be expressed as Eq. (5), or in Fig. 2*c*. The intersection of the two tan-

$$A_P = A^n_P \mapsto A^t_p \tag{5}$$

gential components defines the tip of A_P. This problem illustrates the generalization that any basic acceleration analysis can be thoroughly performed only after completion of the underlying velocity analysis.

Acceleration field. Acceleration is a vector and hence has magnitude and direction but not a unique position. A plane rigid body, such as the cam of **Fig. 3***a*, in motion parallel to its plane will possess an acceleration field in that plane differing from, but resembling, a rigid-body velocity vector field. *See* CAM MECHANISM.

Every acceleration vector in this field, absolute or relative, makes with its instantaneous radius vector the same angle γ (tangent $\gamma = \alpha/\omega^2$) and is proportional in magnitude to the length of this radius vector (Fig. 3*b*). The acceleration field at any instant is thus defined by four parameters: magnitude and direction of accelerations at any two points on the body.

From an acceleration field defined by a_A and a_B (Fig. 3*c*), one can quickly determine the instantaneous center of zero acceleration. From point A', construct vector $A'A''$ parallel and equal to acceleration a_B, represented by vector $B'B$. The relative acceleration a_{AB} is given by the vector of Eq. (6).

$$a_{AB} = a_A - a_B \tag{6}$$

Angle γ between resultant a_{AB} and line AB is common to all radii vectors; therefore, construct radii vectors through A and through B both at γ (properly signed) to their respective acceleration vectors. These lines intersect at Γ, the center of zero acceleration. The geometry contains two similar triangles; $\Delta AB\Gamma$ is similar to $\Delta AA''A'$ because corresponding sides are equally inclined by the angle γ. From these similar triangles comes the generalization that acceleration magnitudes are proportional to radii vectors, just as for the velocity field.

Angle γ ($\tan\gamma = \alpha/\omega^2$) is also used in orientation of acceleration image polygons, especially if the preceding velocity analysis used the image polygon method.

Points not on the same link. Acceleration of follower link 4 of the cam mechanism (Fig. 3*a*) is determined by considering relative accelerations of coincident points not on the same link. Thus point P on link 4, designated P_4, and coincident point P on link 2, designated P_2, have instantaneous velocities, as shown by vectors V_{P4} and V_{P2} (Fig. 3*a*). Relative velocity of P_4 with respect to P_2 is V_{P4P2} as indicated. Accelera-

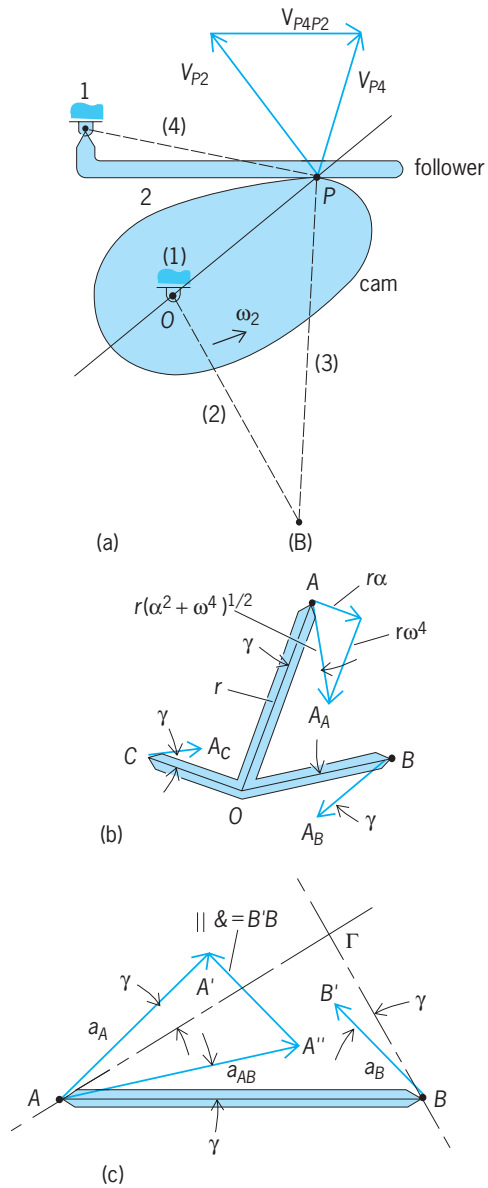

Fig. 3. Accelerations on a cam and follower. (*a*) The cam mechanism. (*b*) The acceleration field about a rigid rotating structure. (*c*) An acceleration field determining an instantaneous center of rotation.

tion of P_4 can now be determined by Eq. (7), where

$$A_{P_4} = A_{P_2} \mapsto A^n_{P_4P_2} \mapsto A^t_{P_4P_2} \mapsto 2\omega_2 V_{P_4P_2} \tag{7}$$

the last term is the Coriolis component. This component results from referring motion of P_4 on body 4 to P_2 on body 2. Serious errors of analysis can result from omission of the Coriolis component. *See* CORIOLIS ACCELERATION.

Occasionally, mechanisms that appear to require the analysis described in the preceding paragraph may also be analyzed by constructing an instantaneously equivalent linkage shown by broken lines (2), (3), and (4) in Fig. 3*a*. The instantaneous equivalent linkage is then analyzed by the method of Fig. 2.

Alternative procedures. Other methods may also be used. It is not always necessary to resolve vectors

into normal and tangential components; they can be solved by conventional orthogonal component techniques. For points A and B in part m, Eq. (8) holds.

$$a_A^{AB} = a_B^{AB} + a_{AB}^{AB} = a_B^{AB} + \omega^2 \cdot AB \qquad (8)$$

A convenient approximate solution is obtained by considering the difference between velocities due to a small angular displacement. This difference in velocities divided by the elapsed short time interval approximates the acceleration. Typically, the angular displacement is taken to be 0.1 radian (0.05 radian on either side of the position under study). *See* ACCELERATION; VECTOR (MATHEMATICS). Douglas P. Adams

Bibliography. J. Angeles, *Spatial Kinematics Chains: Analysis-Synthesis-Optimization,* 1982; J. S. Beggs, *Kinematics,* 1983; C. W. Ham et al., *Mechanics of Machinery,* 4th ed., 1958; D. Lent, *Analysis and Design of Mechanisms,* 2d ed., 1970; A. Sloane, *Engineering Kinematics,* 1966.

Acceleration measurement

The technique of measuring the magnitude and direction of acceleration. Measurement of acceleration includes translational acceleration with tangential and radial components and angular acceleration. *See* ACCELERATION.

Translation acceleration. The components of translational acceleration are the tangential one due to a change in magnitude and the normal one due to a change in direction of the velocity vector. Acceleration measurement by instruments is commonly expressed in terms of g, a dimensionless ratio of acceleration divided by the local value of gravity.

When all points of a rigid body are traveling along parallel straight line paths, the body is said to be in linear motion. For such motion the normal component of acceleration is absent, and if the velocity of the body is changing, the resulting translational acceleration is called linear acceleration. There are many familiar applications of linear acceleration measurement, for example, in land vehicles such as automobiles, in aerospace vehicles, and in moving machinery components.

One of the purposes of acceleration measurement is to determine the force causing the acceleration as manifested by Newton's second law. Also, from acceleration versus time data it is possible by either graphical or numerical integrations to determine the corresponding velocity-time and then the displacement-time histories. Since such techniques are comparatively accurate, it is evident that acceleration measurement is also important in the determination of velocity and displacement.

An acceleration-measuring instrument, or accelerometer, essentially consists of a reference mass mounted by means of a suitable configuration within a housing. A triaxial accelerometer will have three such reference masses so that they respond individually in three mutually perpendicular directions. Accelerometers with a single reference mass respond to accelerations along a predefined axis. To measure translational acceleration whose exact direction is unknown or whose direction varies, either a single triaxial accelerometer or three ordinary accelerometers are required. *See* ACCELEROMETER.

An indirect determination of acceleration is possible by measuring the displacement-time history of a system undergoing linear acceleration. A graphical or a numerical determination of the slope or derivative at selected values of time will produce the corresponding velocity-time relationship, and a repetition of the procedure yields the acceleration history. The accuracy of this means of obtaining acceleration values is poor because the second set of slope determinations by such approximate methods is inherently inaccurate.

The acceleration g due to gravity on the Earth can be determined by various means. A reasonably accurate means of determining g at a given location is by accurately measuring the period T and length L of a simple pendulum on a long cord and subject to a small oscillation in still air. The value of g is then readily determined from the formula $T = 2\pi\sqrt{L/g}$. Another means of determining g is offered by measuring the deflection of a reference mass supported by a highly accurate spring with known characteristics.

Angular acceleration. Technically, a body is said to have angular motion if some fixed reference line between two points rotates or sweeps out an angle. With this definition it is evident that a body can travel along a curved path and still not rotate. For example, the chairs on a moving ferris wheel are not in angular motion with respect to ground. *See* ANGULAR MOMENTUM.

The measurement of angular acceleration is useful in applications such as the angular motion of aerospace vehicles and the components of rotating machinery. Since angular and translational motion are kinematically related, it is possible, when the exact location of the axis of rotation is known, to determine angular acceleration by measuring the corresponding tangential component of translational acceleration.

For determining angular acceleration, a suitably suspended reference mass responds to an angular acceleration about an axis parallel to, or coincident with, its own predefined axis of rotation. As is the case for the translational acceleration measuring instruments, various designs of the angular versions are possible. A direct means of measuring angular acceleration is afforded by the angular accelerometer.

An indirect means of determining angular acceleration is to measure the angular velocity of a body. This technique is applicable in situations where a body is in pure rotation about a fixed axis. The direction of the rotating body can be determined from the output of an electric generator connected to it on the same shaft. A stroboscope can also be used for the same purpose. The rotational speed versus time history can then be graphically or numerically

differentiated once to yield the angular acceleration. *See* ROTATIONAL MOTION; STROBOSCOPE.

Teruo Ishihara; Roger C. Duffield

Bibliography. T. G. Beckwith, R. D. Marangoni, and J. H. Lienhard, *Mechanical Measurements,* 6th ed., 2006; R. S. Figliola and D. E. Beasley, *Theory and Design for Mechanical Measurements*, 4th ed., 2006; R. C. Hibbeler, *Engineering Mechanics: Statics and Dynamics*, 10th ed., 2003; W. F. Riley and L. D. Sturges, *Engineering Mechanics*: *Dynamics*, 2d ed., 1995.

Accelerator mass spectrometry

The use of a combination of mass spectrometers and an accelerator to measure the natural abundances of very rare radioactive isotopes. These abundances are frequently lower than parts per trillion. The most important applications of accelerator mass spectrometry are in archeological, geophysical, environmental, and biological studies, such as in radiocarbon dating by the counting of the rare carbon-14 (radiocarbon; ^{14}C) isotope. *See* MASS SPECTROSCOPE; PARTICLE ACCELERATOR.

The advantage of counting the radioactive atoms themselves rather than their decay products is well illustrated by radiocarbon dating, which requires the measurement of the number of ^{14}C atoms in a sample. This number can be inferred from the beta-particle emission rate or by counting directly the number of radioactive atoms. The long half-life of 5730 years for ^{14}C implies that only 15 beta-particle emissions per minute are observed from 1 g (0.035 oz) of contemporary carbon which contains 6×10^{10} atoms of ^{14}C. Conventional mass spectrometry has not been able to accomplish the task of counting the ^{14}C atoms directly. However, an accelerator mass spectrometer can be used to count the ^{14}C atoms at over 30 per second from a milligram sample of carbon. Consequently, accelerator mass spectrometry can be used to date samples that are a thousand times smaller than those that are dated by using the beta-particle counting method, and the procedure is carried out about 120 times faster. The advantage of accelerator mass spectrometry over beta-particle counting is even greater for very long-lived radioactive isotopes such as aluminum-26 (7×10^5 year half-life) and iodine-129 (1.6×10^7 year half-life). *See* RADIOACTIVITY; RADIOCARBON DATING.

Apparatus. The success of accelerator mass spectrometry results from the use of more than one stage of mass spectrometry and at least two stages of ion acceleration. The layout of an ideal accelerator mass spectrometer for radiocarbon studies is generally divided for convenience into three stages (see **illus.**).

In the first, low-energy, mass spectrometry stage, there is the usual acceleration at the negative-ion source to produce a directed beam of ions with

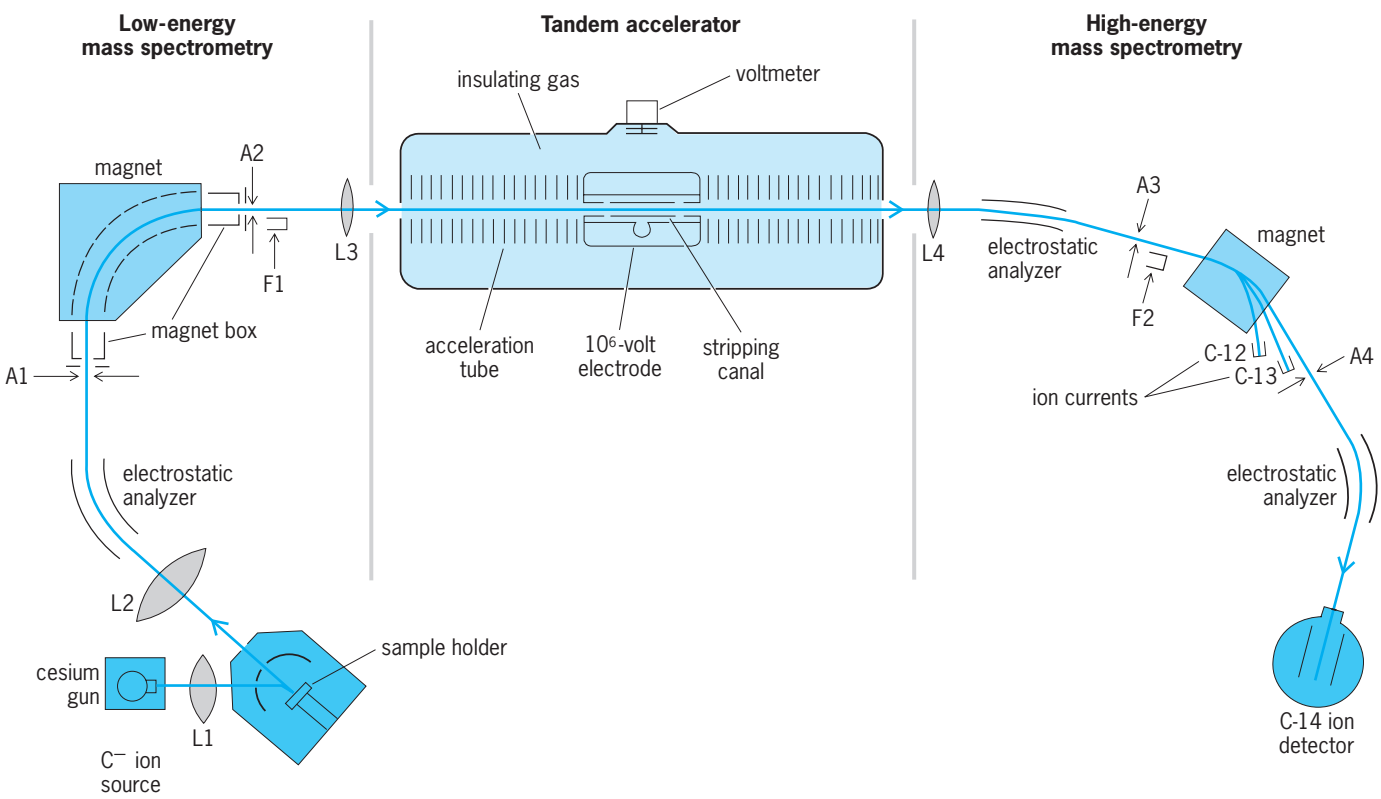

Simplified diagram of an accelerator mass spectrometer used for radiocarbon dating. The equipment is divided into three sections. Electric lenses L1–L4 are used to focus the ion beams. Apertures A1–A4 and charge collection cups F1 and F2 are used for setting up the equipment. The cesium ions from the gun create the negative ions of carbon at the surface of the sample.

an energy of several kiloelectronvolts. This result is achieved by raising the sample holder to the required negative voltage. The first acceleration is followed by an electrostatic lens and an electrostatic analyzer for further ion focusing and energy selection, so that the subsequent mass analysis by a magnetic analyzer can select, uniquely, the ion mass of interest. The first part of the accelerator mass spectrometer is therefore very similar to a conventional mass spectrometer. In this case, the ^{12}C, ^{13}C, and ^{14}C isotopes are analyzed sequentially by varying the voltage on the magnet vacuum box for each isotope in order to analyze each isotope without changing the magnetic field. *See* ION SOURCES.

In the second stage, there is further acceleration by a tandem accelerator. This first accelerates negative ions to the central high-voltage electrode, converts them into positive ions by several successive collisions with gas molecules in a region of higher gas pressure, known as a stripping canal, and then further accelerates the multiply charged positive ions through the same voltage difference back to ground potential. If the acceleration voltage used here is 2×10^6 V, the ions emerging from the accelerator have 8×10^6 eV in energy when they are triply charged in the positive-ion stage.

In the third stage, the accelerated ions are analyzed further by the high-energy mass spectrometer. Following an electrostatic analyzer and a magnetic analyzer, the ion currents of ^{12}C and ^{13}C ions can be measured in the two Faraday cups; additional electrostatic or magnetic analysis then completes the isolation of the rare ionic species. In a typical arrangement (see illus.), the currents of the isotopes ^{12}C and ^{13}C are measured in sequence for a short time, and the much rarer ^{14}C ions then are counted for a longer time in a suitable ion detector. The $^{14}C/^{12}C$ and $^{13}C/^{12}C$ ratios can then be determined from these measurements. Finally, these ratios determine the age of the sample.

Distinguishing features. These three stages of analysis illustrate the features that clearly distinguish accelerator mass spectrometry from conventional mass spectrometry. These are the elimination of molecular ions and isobars from the mass spectrometry.

Elimination of molecular ions. The abundant molecular ions, such as $^{13}CH^-$ and $^{12}CH_2^-$, which have similar masses to the $^{14}C^-$ ions, either must be resolved by a very high-resolution mass spectrometer or must be separated by some other method. A tandem accelerator provides a convenient way of completely eliminating molecular ions from the mass spectrometry because ions of a few megaelectronvolts can lose several electrons on passing through the region of higher gas pressure in the stripping canal. Molecules with more than two electrons missing have not been observed in this case, so that the accelerator mass spectrometry of radiocarbon, utilizing charge-3 ions, is free of molecular interferences. The molecular fragments are removed by the high-energy mass spectrometry.

Elimination of isobars. The use of a negative-ion source, which is necessary for tandem acceleration, can also ensure the complete separation of atoms of nearly identical mass (isobars). In the case of radiocarbon analysis, the abundant stable ^{14}N ions and the very rare radioactive ^{14}C ions, which differ in mass by only 1 part in 86,000, are separated completely because the negative ion of nitrogen is unstable whereas the negative ion of carbon is stable. Other examples of rare radioisotopes that can be completely separated at the negative-ion source are ^{26}Al and ^{129}I. In addition, it is frequently possible to count the ions without any background in the ion detectors because of their high energy. In many cases, it is also possible to identify the ion. For example, ions of the isobars ^{10}Be and ^{10}B can easily be distinguished by their differing ranges in matter in the final ion detection system. Both beryllium and boron form negative ions so they cannot be distinguished at the ion source.

Advantages. As a result of the destruction of molecules in the accelerator and the elimination of the interfering isobars at the negative-ion source or their separation by the ion detection system, the modest mass-resolution requirements of accelerator mass spectrometry can be achieved with high efficiency. This is a great advantage of accelerator mass spectrometry in the measurement of rare radioactive species.

For the study of many rare radioactive atoms, accelerator mass spectrometry also has the important advantage that there can be no background except for contamination with the species being studied. For example, significant interference with the beta-particle counting of radiocarbon from cosmic rays and natural radioactivity occurs for carbon samples about 25,000 years old. In contrast, accelerator mass spectrometer measurements are affected only by the natural contamination of the sample which becomes serious for samples about 50,000 years old.

Applications. Accelerator mass spectrometry has been applied to an increasing variety of problems. In particular, radiocarbon dating is carried out at many laboratories, with thousands of dates being determined each year. These dates approach the accuracy of those produced by the beta-particle counting method. The principal advantage of accelerator mass spectrometry for radiocarbon dating is the ability to date smaller samples of carbon and to do it more quickly. The advantage of the lower background is very useful, and the background limit has been the subject of further studies.

An example of the use of accelerator mass spectrometry can be found in the dating of a small piece of wood, cut by a metal implement, from an archeological site in Newfoundland. This confirmed directly that the site had been occupied by people from Europe about the year 1000. It was presumably one of the early Viking settlements.

There has been an international study of an artifact known as the Shroud of Turin. Tradition holds that this was the burial cloth of Jesus; however, it is now widely believed to be an icon dating from about 1300. Only very small samples from this valuable artifact were made available for dating. *See* ART CONSERVATION CHEMISTRY.

Radiocarbon dating has had an increasing impact on geophysical measurements and is particularly suitable for studying the events associated with the most recent ice age. For example, it is possible to date raised beaches by measuring the radiocarbon from single shells instead of being forced to use heterogeneous mixtures of a large number of shells. Accelerator mass spectrometry is used for oceanographic studies, such as measuring the age of the dissolved carbon dioxide (CO_2) in ocean water, and the age of foraminifera from ocean sediments. Much smaller samples of water or foraminifera can be collected for analysis than would otherwise be possible, facilitating important studies—for example, of global ocean circulation—that are closely related to the understanding of climate.

Radiocarbon has been used extensively as a biomedical tracer in order to detect the effects of toxins, mutagens, carcinogens, or chemotherapeutics in living systems. Radiocarbon is used because it can readily be incorporated into molecules without causing a change in chemical behavior. Accelerator mass spectrometry is important for these studies because of the small sample sizes that can be used, which result in a much lower radiation dose being received by the organism.

Beryllium-10. The long-lived radioactive beryllium isotope ^{10}Be is created in the atmosphere by cosmic rays. Since its half-life is 1.6×10^6 years, it is extremely difficult to study by beta-particle counting, but it was first studied at natural abundances by accelerator mass spectrometry, using a cyclotron as an accelerator and range separation to eliminate the abundant natural isobar ^{10}B. It is now measured routinely by accelerator mass spectrometry in many laboratories by using tandem accelerators. The ^{10}Be atoms in sediments subducted beneath continents and later emitted from volcanoes have been used in studies of the time taken for such subterranean transport. Studies of ^{10}Be in ice cores and sediments are used in determining accumulation rates over millions of years.

Chlorine-36. The long-lived isotope of chlorine ^{36}Cl with a half-life of 350,000 years has also been extensively studied by accelerator mass spectrometry. Unlike ^{10}Be, the ^{36}Cl atoms readily dissolve in water and so are of importance to the science of hydrology. The ^{36}Cl is also created by cosmic rays in the atmosphere and so can be used to study the transport of water from the atmosphere and through underground aquifers. The ^{36}Cl, like ^{26}Al, is also created in surface rocks by cosmic rays, making possible the dating of the exposure of glacial moraines and certain types of rock. For example, the rocks ejected from a meteor crater have been used to estimate that the crater was formed about 30,000 years ago.

Iodine-129. The isotope ^{129}I has a half-life of 1.6×10^7 years and can be studied most easily by accelerator mass spectrometry. The ^{129}I is created in the atmosphere by cosmic rays but is generated mainly by the neutron-induced fission of the uranium isotope ^{235}U. This fission occurs naturally in uranium ores, during the production of nuclear power, and from the fallout following the use of nuclear weapons. This isotope has been used to study the dispersal in the oceans of radioactive wastes from nuclear fuel reprocessing sites and reactors in Europe and North America. The ^{129}I has been measured over a wide area of the North Atlantic, and it can be detected easily in liter quantities of water. It has also been studied in connection with the dispersal of radioactive wastes in Russia and the seas north of Russia.

Many applications of accelerator mass spectrometry will continue to be made in the future. The small sample size required and the high sensitivity for detecting certain long-lived radioactive atoms or very rare stable atoms such as iridium will facilitate their use as environmental and biological tracers. Methods for the detection of very rare atoms in high-purity materials are also being developed. *See* DATING METHODS; GEOCHRONOMETRY; MASS SPECTROMETRY.

Albert E. Litherland

Bibliography. L. Brown, Applications of accelerator mass spectrometry, *Annu. Rev. Earth Planet. Sci.,* 12:39–59, 1984; C. E. Chen et al. (eds.), *Accelerator Mass Spectrometry*, 1992; D. Elmore and F. M. Phillips, Accelerator mass spectrometry for measurement of long-lived radioisotopes, *Science*, 236:543–550, 1987; A. E. Litherland, Fundamentals of accelerator mass spectrometry, *Phil. Trans. Roy. Soc. London,* A323:5–21, 1987.

Accelerometer

A mechanical or electromechanical instrument that measures acceleration. The two general types of accelerometers measure either the components of translational acceleration or angular acceleration. *See* ACCELERATION MEASUREMENT.

Translational accelerometers. Most translational accelerometers fall into the category of seismic instruments, which means the accelerations are not measured with respect to a reference point. Of the two types of seismic instruments, one measures the attainment of a predefined acceleration level, and the other measures acceleration continuously. **Figure 1** shows the first type of instrument, in which a seismic mass is suspended from a bar made of brittle material which fails in tension at a predetermined

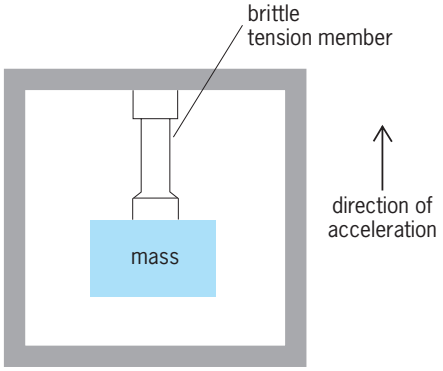

Fig. 1. Brittle-member acceleration-level indicator.

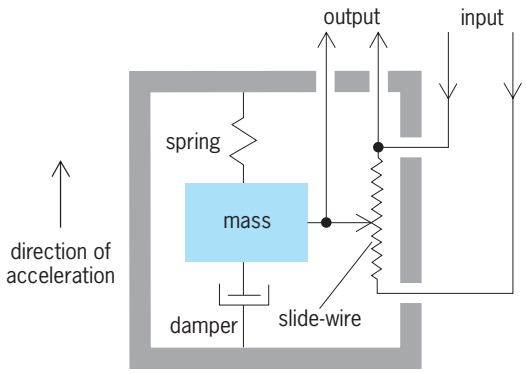

Fig. 2. Slide-wire potentiometer.

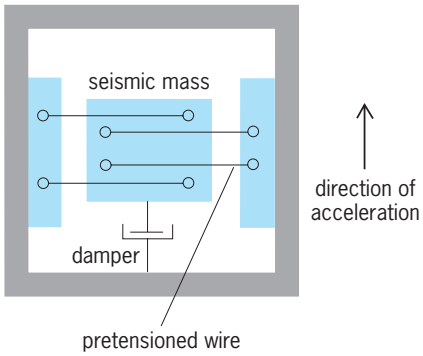

Fig. 3. Unbounded strain gage accelerometer.

acceleration level. *See* SEISMOGRAPHIC INSTRUMENTATION.

Continuously measuring seismic instruments are composed of a damped or an undamped spring-supported seismic mass which is mounted by means of the spring to a housing. The seismic mass is restrained to move along a predefined axis. Also provided is some type of sensing device to measure acceleration. To assure that the deflection of the seismic mass is proportional to the acceleration, such accelerometers are rated so that the excitation frequency associated with the acceleration is only 0.1 to 0.4 of the instrument's resonance frequency, depending on the amount of damping provided. For most accelerometers the damping coefficient varies between 0 and 0.7 of critical damping.

The type of sensing device used to measure the acceleration determines whether the accelerometer is a mechanical or an electromechanical instrument. One type of mechanical accelerometer consists of a liquid-damped cantilever spring-mass system, a shaft attached to the mass, and a small mirror mounted on the shaft. A light beam reflected by the mirror passes through a slit, and its motion is recorded on moving photographic paper. The type of electromechanical sensing device classifies the accelerometer as variable-resistance, variable-inductance, piezoelectric, piezotransistor, or servo type of instrument or transducer.

The sensing device for variable-resistance accelerometers operates on the principle that electrical resistance of an conductor is a function of its dimensions. When the dimensions of the conductor are varied mechanically, as constant current flows through it, the voltage developed across it varies as a function of this mechanical excitation. One type of variable-resistance transducer makes use of a slide-wire potentiometer (**Fig. 2**). This type of accelerometer has low acceleration-measuring ability and low resonance frequency. A more common type of variable-resistance transducer makes use of the fact that the resistance in a wire (called a strain gage) changes when it is strained. Most strain gage accelerometers are either of the unbounded or bounded type. In unbounded strain gage instruments fine pretensioned wires support a fluid-damped seismic mass, and the wires act as both the spring element and the strain gage (**Fig. 3**). The bounded strain gage transducers have the strain gages mounted directly on the spring element, and the gages measure the strain that takes place in the spring due to the relative deflection of the seismic mass. Wire strain gage accelerometers are characterized by limited high resonance frequency, acceleration measurement up to 1000 *g*, and ability to measure acceleration signals down to zero (dc) frequencies. *See* POTENTIOMETER.

Another type of variable-resistance accelerometer uses as the spring a piezoresistive element which is a crystal semiconductor whose resistivity changes with applied force (**Fig. 4**). This type of instrument has much greater sensitivity and resonance frequency than the wire strain gage transducer and is capable of measuring accelerations of 10,000 *g* or less. Variable-resistance instruments are usually arranged electrically to form a Wheatstone bridge. The bridge input excitation is from dry cells or commercial power supplies, and the output which is usually amplified is recorded by means of an oscilloscope or an oscillograph recorder.

A third type of variable-resistance accelerometer employs the Rolamite concept. This accelerometer is composed of two rolling cylinders supported by a thin flexible band, and this low-friction mechanical system is constrained by two parallel guide surfaces (**Fig. 5**). The flexible band has a cutout which provides a negative spring against the displacement of the rollers, which are also the seismic mass. The sensing device consists of a roller making electrical contact across two resistive elements located on one of

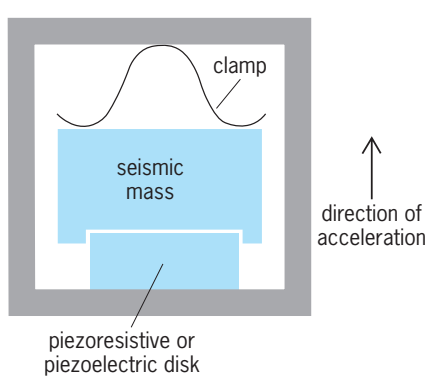

Fig. 4. Piezoresistive or piezoelectric accelerometer.

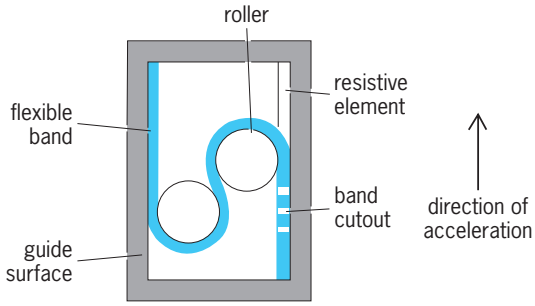

Fig. 5. Rolamite type of accelerometer.

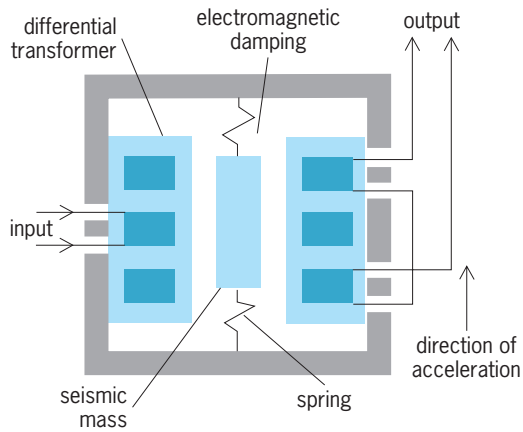

Fig. 6. Variable-inductance accelerometer.

the guide surfaces. The Rolamite accelerometer can be designed to be miniature in size and to measure accelerations in the range 1–500 *g*.

The variable-inductance accelerometer is a differential transformer which makes use of three coils arranged as shown in **Fig. 6**. The center coil is excited from an external ac power source, and the two end coils connected in series opposition are used to produce an ac output which is proportional to the displacement of a seismic mass passing through the coils. This instrument is characterized by fairly low resonance frequencies, acceleration measurements of 50 *g* or less, and high voltage outputs at low frequencies which eliminate the need for preamplifiers.

Piezoelectric accelerometers utilize a piezoelectric crystal which supports the seismic mass in such a way that the crystal acting as a spring is strained in either compression, flexure, or shear (Fig. 4). Many crystals, such as quartz, Rochelle salts, and barium titanate, have these properties. A number of ceramic materials can be used when they contain certain impurities. Piezoelectric instruments are noted for their small size, large acceleration measurement up to 5000 *g*, very high resonance frequency, inability to measure dc frequency acceleration signals, and the need for specialized circuitry. The acceleration signal is usually amplified by means of a charge or voltage amplifier and then recorded.

A piezotransistor accelerometer consists of a seismic mass supported by a stylus which transmits a concentrated force to the upper diode surface of the transistor (**Fig. 7**). Because of this force, localized stresses occur in the upper diode surface which, in turn, cause a very large reversible change in current across the underlying *pn* junction of the transistor. This change of current is a measure of the acceleration. These instruments are characterized by small size, very high resonance frequencies, acceleration measurements of 400 *g* or less, and voltage outputs that are high enough to eliminate need for amplifiers.

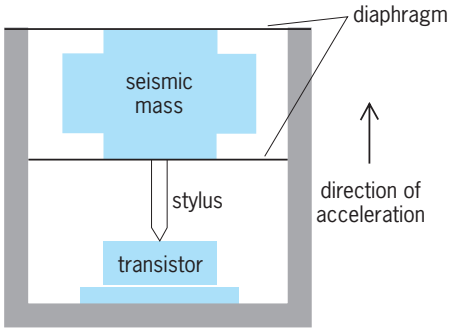

Fig. 7. Piezotransistor accelerometer.

A servo accelerometer is a high-accuracy transducer, containing an electromechanical servomechanism which holds a mass in a fixed position as the instrument is accelerated. The force required to restrain the mass is a measure of the acceleration. This instrument has limited high-frequency response, acceleration outputs of 100 *g* or less, and capability of measuring dc frequency responses. *See* SERVOMECHANISM.

Angular accelerometers. There are several different types of angular accelerometers. In one type the damping fluid serves as the seismic mass. Under angular acceleration the fluid rotates relative to the housing and causes on two symmetrical vanes a pressure which is a measure of the angular acceleration. Another type of instrument has a fluid-damped symmetrical seismic mass in the form of a disk which is so mounted that it rotates about the normal axis through its center of gravity. The angular deflection of the disk, which is restrained by a spring, is proportional to the angular acceleration. The sensing devices used in these instruments to measure the angular acceleration are of the variable-resistance and servo types. Roger C. Duffield; Teruo Ishihara

Bibliography. T. G. Beckwith, R. D. Marangoni, and J. H. Lienhard, *Mechanical Measurements,* 6th ed., 2006; W. Boyes, *Instrumentation Reference Book*, 3d ed., 2002; E. O. Doebelin, *Measurement Systems: Application and Design,* 5th ed., 2003; C. P. Wright, *Applied Measurement Engineering,* 1994.

Acceptor atom

An impurity atom in a semiconductor which can accept or take up one or more electrons from the crystal and become negatively charged. An atom which substitutes for a regular atom of the material but has one less valence electron may be expected to be an acceptor atom. For example, atoms of boron, aluminum, gallium, or indium are acceptors in germanium and silicon (**illus.** *a*), and atoms of antimony

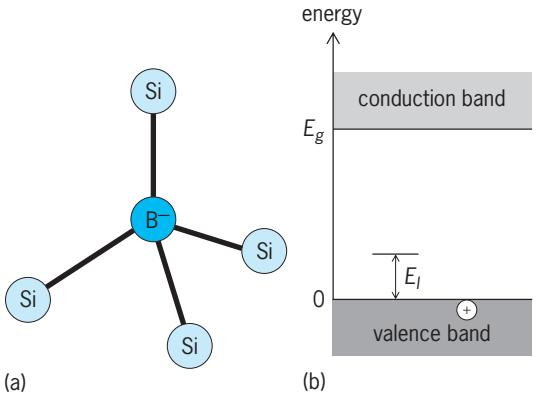

Trivalent acceptor atom, boron (B), in the elemental semiconductor silicon (Si). (*a*) Boron atom in a substitutional position, that is, replacing silicon, a tetravalent host atom, by completing the four tetrahedral covalent bonds with its nearest neighbor silicon atoms. This requires an electron to be accepted from the valence band, thus making boron negatively charged. (*b*) Energy diagram showing that the absence of an electron in the valence band is equivalent to a positive charge carrier, a hole, which is bound to boron via Coulomb attraction with an ionization enerzy E_I. E_g = energy gap separating valence band from conduction band.

and bismuth are acceptors in tellurium crystals. Acceptor atoms tend to increase the number of holes (positive charge carriers) in the semiconductor (illus. *b*). The energy gained when an electron is taken up by an acceptor atom from the valence band of the crystal is the ionization energy of the atom. *See* DONOR ATOM; SEMICONDUCTOR.

H. Y. Fan; Anant K. Ramdas

Bibliography. C. Kittel, *Introduction to Solid State Physics*, 7th ed., Wiley, 1996; A. K. Ramdas and S. Rodriguez, Spectroscopy of the solid-state analogues of the hydrogen atom: Donors and acceptors in semiconductors, *Rep. Prog. Phys.*, 44:1297–1387, 1981.

Acetal

A geminal diether (R_1 = H). Ketals, considered a subclass of acetals, are also geminal diethers (R_1 = C, aliphatic or aromatic). Acetals are (1) independent structural units or a part of certain biological and commercial polymers, (2) blocking or protecting groups for complex molecules undergoing selective synthetic transformations, and (3) entry compounds for independent organic chemical reactions.

Acetals are easily prepared by the reaction of aldehydes with excess alcohol, under acid-catalyzed conditions. This is usually a two-step process, as shown below, in which an aldehyde is treated with an

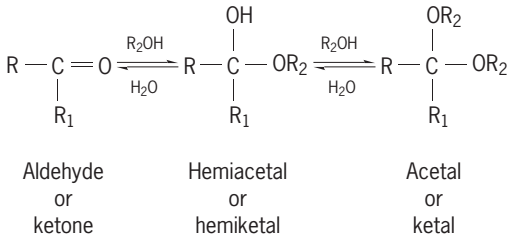

alcohol to yield a less stable hemiacetal, which then reacts with additional alcohol to give the acetal. Protonic or Lewis acids are effective catalysts for acetal formation; dehydrating agents, such as calcium chloride and molecular sieves, can also be used for molecules, such as sugars, where acids may cause problems. Less common acetal preparations are Grignard reagent condensation with orthoformates and mercuric-catalyzed additions of alcohols to acetylenes. *See* GRIGNARD REACTION.

Two well-known hemiacetals (monomers) are α- and β-glucose: the 1,4-polysaccharide from α-glucose is amylose (amylopectin has an acetal 1,6-linkage branch from the linear polymer backbone), and the 1,4-polysaccharide from β-glucose is cellulose. The acetal link connects the monomers incorporated into the polymer in each case, and it is formed by the enzymatic condensation of the hemiacetal functional group (C_1—OH) of one monomer with the (C_4—OH) hydroxyl functional group of another monomer. Several other biopolymers containing similar acetal links between repeating units are chitin, inulin, and pectin. Synthetic polymers with 1,3-diol units, such as polyvinyl alcohol, can be transformed into polymeric acetals, which are stable and useful plastics (for example, formal and butyral). Difunctional aldehydes and ketones can be condensed with certain tetrols (for example, pentaerythritol) for preparation of polyketals, and polyspiroketals, which are highly crystalline materials. Polyoxymethylene, which is the acetal resin homopolymer of formaldehyde, is prepared by cationic polymerization of trioxane, the cyclic acetal of formaldehyde. Homopolymerization of acetaldehyde or acetone, where the acetal or ketal oxygen is also part of the polymer backbone, is possible, but requires specialized conditions. *See* CELLULOSE; POLYACETAL; POLYMER; POLYMERIZATION; POLYSACCHARIDE; STARCH.

Their use as protective groups is attributed to the fact that acetals (or ketals) are readily cleaved by aqueous acids to the respective aldehyde (or ketone) and that they are stable to alkalies and nucleophiles. Cyclic acetals, such as *meta*-dioxolanes and *meta*-dioxanes, are very convenient to work with for blocking–unblocking types of syntheses. During these multistep processes, the carbonyl portion of the molecule is transformed to its acetal (or ketal), while synthetic reactions that do not utilize acid are effected on other functional groups of the molecule. The acetal (or ketal) part of the molecule is then readily hydrolyzed to free the carbonyl functional group. The cyclic acetals can also be used as aprotic solvents and plasticizers. *See* ALDEHYDE; CARBONYL; KETONE; ORGANIC SYNTHESIS.

Acetals undergo a variety of useful reactions such as α-acylation, α-bromination, condensations with enol acetates and silyl enolates, conversion to enol ethers, oxidation to esters, reduction to ethers, and reactions with active hydrogen compounds (for example, diethyl malonate). Acetal additions to vinyl ethers followed by hydrolysis result in a preparation of α,β-unsaturated aldehydes. Dioxolanes (from aldehydes) are also treated with alkyllithiums for the preparation of ketones.

Thioacetals, especially cyclic thioacetals, can be reduced to alkanes with Raney nickel, and 1,3-dithianes, prepared from aldehydes, can be metalated, condensed with a variety of electrophilic reagents, and hydrolyzed to free the functionalized aldehyde. Charles F. Beam

Bibliography. R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; D. S. Kemp and F. Vellaccio, *Organic Chemistry*, 1980; R. W. Lenz, *Organic Chemistry of Synthetic High Polymers*, 1967, reprint 1988; J. March, *Advanced Organic Chemistry: Reactions, Mechanisms, and Structures*, 4th ed., 1992; R. B. Seymour and C. E. Carraher, Jr., *Polymer Chemistry*, 1982; M. P. Stevens, *Polymer Chemistry*, 3d ed., 1998.

Acetic acid

A colorless, pungent liquid, CH_3COOH, melting at 16.7°C (62.1°F) and boiling at 118.0°C (244.4°F). Acetic acid is the sour principle in vinegar. Concentrated acid is called glacial acetic acid because of its readiness to crystallize at cool temperatures. Acetic acid in vinegar arises through an aerobic fermentation of dilute ethanol solutions, such as wine, cider, and beer, with any of several varieties of *Acetobacter*.

Chemical properties. Pure acetic acid is completely miscible with water, ethanol, diethyl ether, and carbon tetrachloride, but is not soluble in carbon disulfide. Freezing of acetic acid is accompanied by a remarkable volume contraction: the molar volume of liquid acetic acid at the freezing point is 57.02 cm^3/mole, but at the same temperature the crystalline solid is 47.44 cm^3/mole. It is a strongly proton-donating solvent with a relatively small dipole moment and a low dielectric constant. In a water solution, acetic acid is a typical weakly ionized acid ($K_a = 1.8 \times 10^{-5}$). *See* CARBOXYLIC ACID.

The vapor density of acetic acid indicates a molecular weight considerably higher than would be expected for a compound with a formula weight of 60.05. The acid probably exists largely as the dimer in the vapor and liquid states.

Acetic acid neutralizes many oxides and hydroxides, and decomposes carbonates to furnish acetate salts, which are used in textile dyeing and finishing, as pigments, and as pesticides; examples are verdigris, white lead, and paris green. *See* ARSENIC; LEAD.

Over two-thirds of the acetic acid manufactured is used in production of either vinyl acetate or cellulose acetate. Acetic anhydride, the key intermediate in making cellulose acetate, is prepared commercially by pyrolysis of acetic acid in the presence of trialkyl phosphate catalyst. Considerable amounts of acetic acid are consumed in the manufacture of terephthalic acid by liquid-phase oxidation of xylene, and in the preparation of esters for lacquer solvents, paints and varnishes, pharmaceuticals, and herbicides.

Production. Acetic acid was formerly manufactured from pyroligneous acid obtained in destructive distillation of wood. These processes are of historical interest because many modern chemical engineering operations developed through the study of acetic acid production. Today acetic acid is manufactured by three main routes: butane liquid-phase catalytic oxidation in acetic acid solvent, palladium–copper salt–catalyzed oxidation of ethylene in aqueous solution, and methanol carbonylation in the presence of rhodium catalyst. Large quantities of acetic acid are recovered in manufacturing cellulose acetate and polyvinyl alcohol. Some acetic acid is produced in the oxidation of higher olefins, aromatic hydrocarbons, ketones, and alcohols. *See* ESTER; OXIDATION PROCESS; SOLVENT; WOOD CHEMICALS. Frank Wagner

Bibliography. W. H. Brown, *Introduction to Organic and Biochemistry*, 4th ed., 1987; T. A. Geissman, *Principles of Organic Chemistry*, 4th ed., 1977; J. D. Roberts and M. C. Caserio, *Basic Principles of Organic Chemistry*, 2d ed., 1977.

Acetone

A chemical compound, CH_3COCH_3; the first member of the homologous series of aliphatic ketones. It is a colorless liquid with an ethereal odor. Its physical properties include boiling point 56.2°C (133°F), melting point −94.8°C (−138.6°F), and specific gravity 0.791. Acetone is an extremely important, low-cost raw material that is used for production of other chemicals.

Acetone is used as a solvent for cellulose ethers, cellulose acetate, cellulose nitrate, and other cellulose esters. Cellulose acetate is spun from acetone solution. Lacquers, based on cellulose esters, are used in solution in mixed solvents including acetone. Acetylene is safely stored in cylinders under pressure by dissolving it in acetone, which is absorbed on inert material such as asbestos. It has a low toxicity.

Production. The principal method of acetone production uses propylene, obtained from the cracking of petroleum. Addition of sulfuric acid to propylene yields isopropyl hydrogen sulfate, which upon hydrolysis yields isopropyl alcohol. Oxidation or dehydrogenation over metal catalysts, such as copper, converts the alcohol to acetone, as shown in reaction (1).

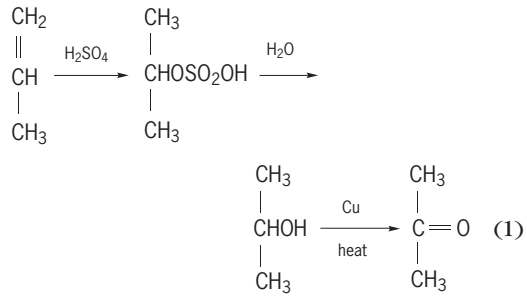

(1)

Acetone is also produced by passage of acetic acid vapor over metallic oxide catalysts at 400–450°C (750–840°F), by partial oxidation of the lower alkane hydrocarbons, and by the decomposition of cumene

hydroperoxide, as shown in reaction (2). Phenol is the other product of this last process.

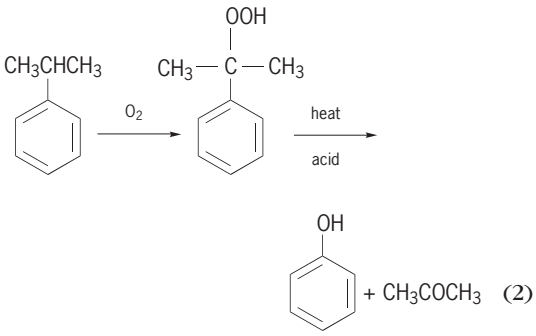

Chemical uses. Pyrolysis of acetone vapor at 700°C (1300°F) produces ketene, which reacts with acetic acid to produce acetic anhydride, as shown in reaction (3).

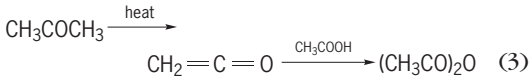

Aldol-type condensation of acetone with $Ba(OH)_2$ yields diacetone alcohol (**1**), mesityl oxide (**2**), and phorone (**3**), shown in reaction (4). Catalytic hydro-

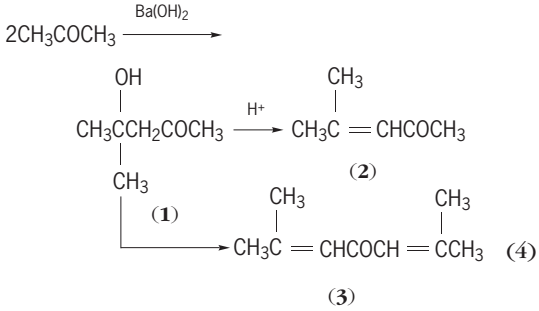

genation of mesityl oxide gives methyl isobutyl ketone. All these products are of commercial use as solvents or as chemical intermediates. *See* KETONE; KETONE.

David A. Shirley

Acetylcholine

A naturally occurring quaternary ammonium cation ester, with the formula $CH_3(O)COC_2H_4N(CH_3)_3^+$, that plays a prominent role in nervous system function.

Neurotransmission. The great importance of acetylcholine derives from its role in physiology as a neurotransmitter for cholinergic neurons. Cholinergic nerves innervate many tissues, including smooth muscle and skeletal muscle, the heart, ganglia, and glands. The effect of stimulating a cholinergic nerve, for example, the contraction of skeletal muscle or the slowing of the heartbeat, results from the release of acetylcholine from the nerve endings. Other types of nerves release other transmitters, for example, norepinephrine.

The release of a neurotransmitter was first clearly shown by Otto Loewi, who demonstrated in 1921 that the perfusate from a frog's heart that was slowed by stimulation of the vagus nerve slowed the heart of a second frog. The active substance in the perfusate was expected to be acetylcholine, which was known to slow the heart; this proved to be correct.

Biosynthesis. Acetylcholine is synthesized at axon endings from acetyl coenzyme A and choline by the enzyme choline acetyltransferase which transfers the acetyl group to choline, reaction (1). Acetylcholine

$$\text{Acetyl coenzyme A} + \text{choline} \longrightarrow \text{coenzyme A} + \text{acetylcholine} \quad (1)$$

is stored at each ending in hundreds of thousands of membrane-enclosed synaptic vesicles. The axon ending is also rich in mitochondria that supply acetyl coenzyme A using pyruvate (from glucose) as the source of the acetyl group, in a reaction catalyzed by the pyruvate dehydrogenase complex, reaction (2).

$$\text{Pyruvate} + \text{coenzyme A} + NAD^+ \longrightarrow \text{acetyl coenzyme A} + CO_2 + NADH, H^+ \quad (2)$$

Mitochondria also supply adenosine triphosphate for the acetylation of coenzyme A by acetate in a two-step reaction (3).

$$\text{Acetate} + \text{adenosine triphosphate} \longrightarrow \text{acetyl adenosine monophosphate} + \text{pyrophosphate}$$
$$\downarrow + \text{coenzyme A}$$
$$\text{adenosine monophosphate} + \text{acetyl coenzyme A} \quad (3)$$

Molecular biology of axon. Neurons, like other cells, have a high internal concentration of potassium ions and a low internal concentration of sodium ions maintained by a sodium–potassium "pump." An axon at rest is permeable mainly to potassium, and so a very small amount of potassium leaves the cell and establishes a membrane potential of about −80 millivolts (the inside negative) that nearly blocks further leakage of potassium. This small potential produces a strong electric field in the membrane because the membrane is thin and has a low dielectric constant. The potential and high external concentration of sodium ions could drive a large inward sodium current, if the membrane were permeable to sodium. *See* ION TRANSPORT.

Axonal membranes contain voltage-gated sodium and potassium channels. These channels are composed of proteins that span the membrane and allow the transient flow of their specific ion, depending on the membrane potential. When an axon is stimulated at some point by slightly increasing the membrane potential to a less negative value, there is a transient opening of sodium channels followed by a transient opening of potassium channels. As a result, sodium ions enter the cell and drive the membrane potential positive. As the sodium current subsides, the potassium channels open and the outward flow of potassium restores the resting potential. This sequence of potential changes, the action potential, is self-propagating as a nerve impulse that moves along

the axon away from the cell body toward the numerous branched endings (toward the cell body in a dendrite). Thus, although the electric currents are perpendicular to the axon, the impulse moves parallel to the axon. *See* BIOPOTENTIALS AND IONIC CURRENTS.

Molecular biology of synapse. When an impulse reaches an axon ending, voltage-gated calcium channels open and calcium, which is extremely low inside the cell, enters the nerve ending. The increase in calcium-ion concentration causes hundreds of synaptic vesicles to fuse with the cell membrane and expel acetylcholine into the synaptic cleft (exocytosis). The acetylcholine released at a neuromuscular junction binds reversibly to acetylcholine receptors in the muscle end-plate membrane, a postsynaptic membrane that is separated from the nerve ending by a very short distance. The receptor is a cation channel. In *Torpedo*, an electric fish, the receptor is composed of five protein subunits of four kinds, $\alpha_2\beta\gamma\delta$. The channel opens when two acetylcholine molecules are bound, one to each α subunit. A sodium current enters the cell and depolarizes the membrane, producing an end-plate potential which triggers an action potential. The action potential, as in the nerve, is self-propagating. The resulting impulse indirectly causes the muscle to contract.

Acetylcholine must be rapidly removed from a synapse in order to restore it to its resting state. This is accomplished in part by diffusion but mainly by the enzyme acetylcholinesterase which hydrolyzes acetylcholine, reaction (4). The enzyme mechanism

$$\text{Acetylcholine} + H_2O \longrightarrow \text{acetic acid} + \text{choline} \qquad (4)$$

involves the transfer of the acetyl group from choline to the hydroxyl group of a specific serine residue of the enzyme located at the active site. The reaction is completed when the acetyl enzyme, itself an ester, hydrolyzes to form acetic acid and free enzyme. As expected, acetylcholinesterase is a very fast enzyme: one enzyme molecule can hydrolyze 10,000 molecules of acetylcholine in 1 s.

Red cells contain acetylcholinesterase and blood serum contains a different cholinesterase, but the function of these enzymes is not clearly known. *See* SYNAPTIC TRANSMISSION.

Poisons, insecticides, and drugs. It is evident that any substance that efficiently inhibits acetylcholinesterase will be extremely toxic. Thus, synthetic organophosphonates such as Sarin, Tabun, and Soman are military nerve gases. These substances, which are volatile liquids, transfer a phosphonyl group to the specific serine side chain of the enzyme to produce a phosphonyl-enzyme derivative analogous to the acetyl enzyme. But unlike the acetyl enzyme, the phosphonyl enzyme hydrolyzes only extremely slowly and is trapped as an inactive derivative. Some organophosphates (similar to phosphonates) are useful: Parathion and Malathion are selectively toxic to insects and are important insecticides. Organophosphate poisoning is treated with atropine and 2-pyridine aldoxime methiodide (2-PAM): atropine antagonizes some of the effects of acetylcholine, and 2-PAM reactivates the inhibited enzyme by removing the phosphoryl group. *See* INSECTICIDE.

Similarly, certain carbamates form carbamyl derivatives of acetylcholinesterase. Carbaryl is an important carbamate insecticide. An increase in the persistence of acetylcholine at synapses is sometimes medically desirable, and carbamates are used to treat a number of medical disorders. Pyridostigmine is extensively used in the treatment of myasthenia gravis, a disease characterized by muscle weakness which is believed to be an autoimune disease, that is, antibodies are produced against the acetylcholine receptor. The proper dosage of the carbamate increases the persistence of acetylcholine at the neuromuscular junction and improves muscle strength. Carbamates, once extensively used in treating glaucoma, usually in conjunction with pilocarpine which activates the acetylcholine receptor now find limited use. Other types of acetylcholinesterase inhibitors are used in managing Alzheimer's disease. *See* MYASTHENIA GRAVIS.

Succinyldicholine, which reacts with the acetylcholine receptor of skeletal muscle, is used to produce paralysis and relaxation of muscle during surgery. This ester is not hydrolyzed by acetylcholinesterase but is hydrolyzed by serum cholinesterase which limits the duration of paralysis. Because a small fraction of the population has a variant cholinesterase, patients who will be given succinyldicholine are tested to be sure their serum will hydrolyze this drug.

There are a number of exotic, naturally occurring poisons that interfere with the nerve impulse or with synaptic transmission. Many of these substances have been very useful in studying the mechanism of these processes. Botulinus toxin prevents the release of acetylcholine, and black widow spider toxin causes the release of acetylcholine. The jungle-dart poison curare binds to the acetylcholine receptor of skeletal muscle and causes paralysis. Bungarotoxin and cobratoxin found in snake venoms have a similar effect. Muscarine from some poisonous mushrooms and arecoline from the betel nut react with the acetylcholine receptor of other tissues. Tetrodotoxin from the Japanese puffer fish, saxitoxin from the bacterium that causes the "red tide" of mussels and clams, and batrachotoxin from a South American frog bind to and block the function of sodium channels, thus preventing the nerve impulse. *See* POISON; TOXICOLOGY; TOXIN.

Electric organs. Some aquatic animals stun or kill their prey by producing a high-voltage electrical discharge. Their electric organs are composed of thousands of flat cells arranged in series. These cells are derived from skeletal muscle and have a conducting membrane on one face which also contains many nerve junctions. The nearly simultaneous excitation of these faces produces action potentials that are summed by the series arrangement to produce hundreds of volts. *See* ELECTRIC ORGAN (BIOLOGY); NERVOUS SYSTEM (VERTEBRATE).

Irwin B. Wilson

Bibliography. B. Alberts et al., *Molecular Biology of the Cell*, 1994; A. G. Gilman et al. (eds.), *The Pharmacological Basis of Therapeutics*, 9th ed., 1996; L. Stryer, *Biochemistry*, 4th ed., 1995.

Acetylene

An organic compound with the formula C_2H_2 or HC≡CH. The first member of the alkynes, acetylene is a gas with a narrow liquid range; the triple point is −81°C (−114°F). The heat of formation (ΔH°_f) is +227 kilojoules/mole, and acetylene is the most endothermic compound per carbon of any hydrocarbon. The compound is thus extremely energy-rich and can decompose with explosive force. At one time acetylene was a basic compound for much chemical manufacturing. It is highly reactive and is a versatile source of other reactive compounds.

Preparation. The availability of acetylene does not depend on petroleum liquids, since it can be prepared by hydrolysis of calcium carbide (CaC_2), obtained from lime (CaO), and charcoal or coke (C) [reaction (1)]. Other preparative methods are based

$$CaO + 3C \longrightarrow CO + CaC_2 \xrightarrow{H_2O} C_2H_2 + Ca(OH)_2 \quad (1)$$

on pyrolysis of some carbon source. In modern practice, methane (CH_4) is passed through a zone heated to 1500°C (2732°F) for a few milliseconds, and acetylene is then separated from the hydrogen in the effluent gas [reaction (2)].

$$2CH_4 \longrightarrow C_2H_2 + 3H_2 \quad (2)$$

Uses. There are processes for the conversion of acetylene to a number of important compounds, including vinyl chloride, vinyl acetate, acrylonitrile, and chloroprene, all of which are monomer intermediates for polymer production [reaction (3)]. Much of this chemistry is applicable also to higher alkynes.

Although acetylene is a versatile and flexible raw material, nearly all of these processes became obsolete as the petrochemical industry developed and ethylene became a more economical feedstock for large-volume vinyl monomers and other C_2 chemicals. A major consideration is that acetylene is a very dangerous substance and handling costs are high. Production of acetylene dropped about 10% per year from a high of 1.1×10^9 lb in 1969 to 3.5×10^8 lb (1.6×10^8 kg) in 1990. *See* ETHYLENE.

The main use of acetylene is in the manufacture of compounds derived from butyne-1,4-diol. The latter is obtained by condensation of acetylene with two moles of formaldehyde and is converted to butyrolactone, tetrahydrofuran, and pyrrolidone. Two additional products, vinyl fluoride and vinyl ether, are also based on acetylene. Cyclooctatetraene, the tetramer of acetylene, is potentially available in quantity if a demand should develop. A superior grade of carbon black that is obtained by pyrolysis of acetylene above 1500°C (2732°F) has good electrical conductivity and is produced for use in dry-cell batteries. *See* PYROLYSIS.

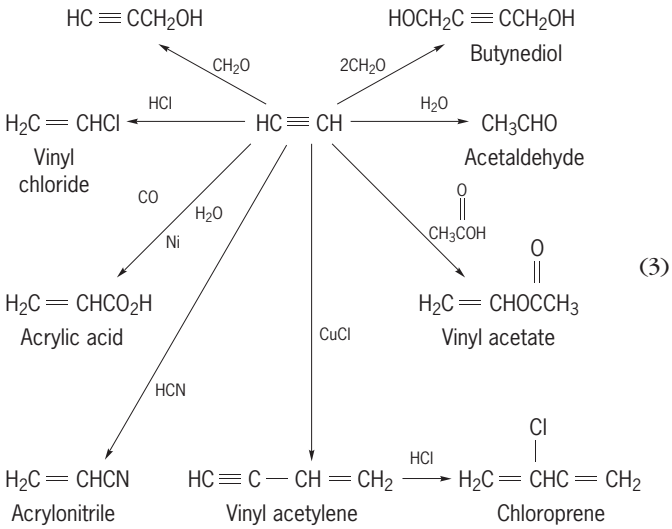

(3)

Because of the very high heat of formation and combustion, an acetylene-oxygen mixture provides a very high temperature flame for welding and metal cutting. For this purpose acetylene is shipped as a solution in acetone, loaded under pressure in cylinders that contain a noncombustible spongy packing. *See* ALKYNE; TORCH.

James A. Moore

Bibliography. *Kirk-Othmer Encyclopedia of Chemical Technology*, vol. 1, 4th ed., 1991; *Ullmann's Encyclopedia of Industrial Chemistry*, vol. A1, 5th ed., 1985.

Acid and base

Two interrelated classes of chemical compounds, the precise definitions of which have varied considerably with the development of chemistry. These changing definitions have led to frequent controversies, some of which are still unresolved. Acids initially were defined only by their common properties. They were substances which had a sour taste, which dissolved many metals, and which reacted with alkalies (or bases) to form salts. For a time, following the work of A. L. Lavoisier, it was believed that a common constituent of all acids was the element oxygen, but gradually it became clear that, if there were an essential element, it was hydrogen, not oxygen. In fact, the definition of an acid, formulated by J. von Liebig in 1840, as "a hydrogen-containing substance which will generate hydrogen gas on reaction with metals" proved to be satisfactory for about 50 years.

Bases initially were defined as those substances which reacted with acids to form salts (they were the "base" of the salt). The alkalies, soda and potash, were the best-known bases, but it soon became clear that there were other bases, notably ammonia and the amines.

Acids and bases are among the most important chemicals of commerce. The inorganic acids are often known as mineral acids, and among the most important are sulfuric, H_2SO_4; phosphoric, H_3PO_4; nitric, HNO_3; and hydrochloric, HCl (sometimes

called muriatic). Among the many important organic acids are acetic, CH_3COOH, and oxalic, $H_2C_2O_4$, acids, and phenol, C_6H_5OH. The important inorganic bases are ammonia, NH_3; sodium hydroxide or soda, NaOH; potassium hydroxide, KOH; calcium hydroxide or lime, $Ca(OH)_2$; and sodium carbonate, Na_2CO_3. There are also many organic bases, mostly derivatives of ammonia. Examples are pyridine, C_5H_5N, and ethylamine, $C_2H_5NH_2$.

Arrhenius-Ostwald theory. When the concept of ionization of chemical compounds in water solution became established, some considerably different definitions of acids and bases became popular. Acids were defined as substances which ionized in aqueous solution to give hydrogen ions, H^+, and bases were substances which reacted to give hydroxide ions, OH^-. These definitions are sometimes known as the Arrhenius-Ostwald theory of acids and bases and were proposed separately by S. Arrhenius and W. Ostwald. Their use makes it possible to discuss acid and base equilibria and also the strengths of individual acids and bases. The ionization of an acid in water can be written as Eq. (1). Qualitatively, an

$$HA = H^+ + A^- \tag{1}$$

acid is strong if this reaction goes extensively toward the ionic products and weak if the ionization is only slight. A quantitative treatment of this ionization of dissociation can be given by utilizing the equilibrium expression for the acid, as shown in Eq. (2), where

$$\frac{[H^+][A^-]}{[HA]} = K_{HA} \tag{2}$$

the brackets mean concentration in moles per liter and the constant K_{HA} is called the dissociation constant of the acid. This dissociation constant is a large number for a strong acid and a small number for a weak acid. For example, at 25°C (77°F) and with water as the solvent, K_{HA} has the value 1.8×10^{-5} for a typical weak acid, acetic acid (the acid of vinegar), and this value varies only slightly in dilute solutions as a function of concentration. Dissociation constants vary somewhat with temperature. They also change considerably with changes in the solvent, even to the extent that an acid which is fully ionized in water may, in some other less basic solvent, become decidedly weak. Almost all the available data on dissociation constants are for solutions in water, partly because of its ubiquitous character, and partly because it is both a good ionizing medium and a good solvent.

Acetic acid has only one ionizable hydrogen and is called monobasic. Some other acids have two or even three ionizable hydrogens and are called polybasic. An example is phosphoric acid, which ionizes in three steps, shown in Eqs. (3), (4), and (5), each with its own dissociation constant.

Ionization reaction	K_{HA}, *25 °C (77 °F)*	
$H_3PO_4 = H^+ + H_2PO_4^-$	1.5×10^{-3}	(3)
$H_2PO_4^- = H^+ + HPO_4^{2-}$	6.2×10^{-8}	(4)
$HPO_4^{2-} = H^+ + PO_4^{3-}$	4×10^{-13}	(5)

A similar discussion can be given for the ionization of bases in water. However, the concentrations of the species H^+ and OH^- in a water solution are not independently variable. This is because water itself is both a weak acid and a weak base, ionizing very slightly according to Eq. (6). For pure water,

$$H_2O = H^+ + OH^- \tag{6}$$

the concentration of H^+ and OH^- are equal. At ordinary temperatures, roughly 2×10^{-7}% of the water is present as ions. As a result of this ionization, the ionic concentrations are related through Eq. (7). At

$$\frac{[H^+][OH^-]}{[H_2O]} = K \tag{7}$$

25°C (77°F) and with concentrations in moles per liter, the product $[H^+][OH^-]$ is equal to 1×10^{-14}.

A major consequence of this interdependence is that measurement of the concentration of either H^+ or OH^- in a water solution permits immediate calculation of the other. This fact led S. P. L. Sørenson in 1909 to propose use of a logarithmic pH scale for the concentration of hydrogen ions in water. Although there are some difficulties in giving an exact definition of pH, it is very nearly correct for dilute solutions in water to write Eq. (8). It then turns out that pH

$$pH = -\log[H^+] \tag{8}$$

values of 0–14 cover the range from strongly acidic to strongly basic solutions. The pH of pure water at ordinary temperature is 7. *See* WATER.

For many situations, it is desirable to maintain the H^+ and OH^- concentration of a water solution at low and constant values. A useful device for this is a mixture of a weak acid and its anion (or of a weak base and its cation). Such a mixture is called a buffer. A typical example is a mixture of sodium acetate and acetic acid. From the treatment just given, it is evident that for this case Eq. (9) can be formed. In the

$$[H^+] = \frac{[CH_3COOH]}{[CH_3COO^-]} \times 1.8 \times 10^{-5} \tag{9}$$

equation $[CH_3COOH]$ and $[CH_3COO^-]$ represent the concentrations of acetic acid and acetate ion, respectively. Thus, if the concentrations of acetic acid and acetate ion are both 0.1 mole per liter, the H^+ concentration will be 1.8×10^{-5} mole per liter and OH^- will be 5.5×10^{-10} mole per liter. The pH of this solution will be about 4.7. Constant acidity is a most important aspect of blood and other life fluids; these invariably contain weak acids and bases to give the necessary buffering action.

Brönsted theory. The Arrhenius, or water, theory of acid and bases has many attractive features, but it has also presented some difficulties. A major difficulty was that solvents other than water can be used for acids and bases and thus need consideration. For many of the solvents of interest, the necessary extensions of the water theory are both obvious and plausible. For example, with liquid ammonia as the solvent, one can define NH_4^+ as the acid ion and NH_2^- as the base ion, and the former can be thought

of as a hydrogen ion combined with a molecule of the solvent. However, for a hydrogenless (aprotic) solvent, such as liquid sulfur dioxide, the extensions are less obvious. Consideration of such systems has led to some solvent-oriented theories of acids and bases to which the names of E. C. Franklin and A. F. D. Germann often are attached. The essence of these theories is to define acids and bases in terms of what they do to the solvent. Thus, one definition of an acid is that it gives rise to "a cation which is characteristic of the solvent," for example, SO^{2+} from sulfur dioxide. These theories have been useful in emphasizing the need to consider nonaqueous systems. However, they have not been widely adopted, at least partly because a powerful and wide-ranging protonic theory of acids and bases was introduced by J. N. Brönsted in 1923 and was rapidly accepted by many other scientists. Somewhat similar ideas were advanced almost simultaneously by T. M. Lowry, and the new theory is occasionally called the Brönsted-Lowry theory.

This theory gives a unique role to the hydrogen ion, and there appeared to be justification for this. One justification, of course, was the historically important role already given to hydrogen in defining acids. A rather different justification involved the unique structure of hydrogen ion. It is the only common ion which consists solely of a nucleus, the proton. As a consequence, it is only about 10^{-14} cm in diameter. All other ordinary ions have peripheral electron clouds and, as a result, are roughly 10^6 times larger than the proton. The small size of the latter makes it reasonable to postulate that protons are never found in a free state but rather always exist in combination with some base. The Brönsted theory emphasizes this by proposing that all acid-base reactions consist simply of the transfer of a proton from one base to another.

The Brönsted definitions of acids and bases are: An acid is a species which can act as a source of protons; a base is a species which can accept protons. Compared to the water theory, this represents only a slight change in the definition of an acid but a considerable extension of the term base. In addition to hydroxide ion, the bases now include a wide variety of uncharged species, such as ammonia and the amines, as well as numerous charged species, such as the anions of weak acids. In fact, every acid can generate a base by loss of a proton. Acids and bases which are related in this way are known as conjugate acid-base pairs, and the **table** lists several examples. By these definitions, such previously distinct chemical processes as ionization, hydrolysis, and neutralization become examples of the single class of proton transfer or protolytic reactions. The general reaction is expressed as Eq. (10). This equation can

$$\text{Acid}_1 + \text{base}_2 = \text{base}_1 + \text{acid}_2 \qquad (10)$$

be considered to be a combination of two conjugate acid-base pairs, and the pairs below can be used to construct a variety of typical acid-base reactions. For example, the ionization of acetic acid in water becomes Eq. (11). Water functions here as a base to

$$\underset{a_1}{CH_3COOH} + \underset{b_2}{H_2O} = \underset{b_1}{CH_3COO^-} + \underset{a_2}{H_3O^+} \qquad (11)$$

form the species H_3O^+, the oxonium ion (sometimes called the hydronium ion). However, water can also function as an acid to form the base OH^-, and this dual or amphoteric character of water is one reason why so many acid-base reactions occur in it.

As the table shows, strengths of acids and bases are not independent. A very strong Brönsted acid implies a very weak conjugate base and vice versa. A qualitative ordering of acid strength or base strength, as above, permits a rough prediction of the extent to which an acid-base reaction will go. The rule is that a strong acid and a strong base will react extensively with each other, whereas a weak acid and a weak base will react together only very slightly. More accurate calculations of acid-base equilibria can be made by using the ordinary formulation of the law of mass action. A point of some importance is that, for ionization in water, the equations reduce to the earlier Arrhenius-Ostwald type. Thus, for the ionization of acetic acid in water Eq. (11) leads to Eq. (12).

$$\frac{[H_3O^+][CH_3COO^-]}{[CH_3COOH][H_2O]} = K \qquad (12)$$

Remembering that the concentration of water will be almost constant since it is the solvent, this can be written as Eq. (13), where K_{HAc} is just the conventional dissociation constant for acetic acid in water.

$$\frac{[H_3O^+][CH_3COO^-]}{[CH_3COOH]} = K[H_2O] \equiv K_{HAc} \qquad (13)$$

One result of the Brönsted definitions is that for a given solvent, such as water, there is only a single scale of acid-base strength. Put another way, the relative strength of a set of acids will be the same for any base. Hence, the ordinary tabulation of ionization constants of acids in water permits quantitative calculation for a very large number of acid-base equilibria.

The Brönsted concepts can be applied without difficulty to other solvents which are amphoteric in the same sense as water, and data are available for many nonaqueous solvents, such as methyl alcohol, formic acid, and liquid ammonia. An important practical point is that relative acid (or base) strength turns out to be very nearly the same in these other solvents as it is in water. Brönsted acid-base reactions can also be studied in aprotic solvents (materials such as

Conjugate acid-base pairs

Strong acids	H_2SO_4	HSO_4^-	Weak bases
	HCl	Cl^-	
	H_3O^+	H_2O	
	HSO_4^-	SO_4^{2+}	
	$HF_{(aq)}$	F^-	
	CH_3COOH	CH_3COO^-	
	HN_4^+	NH_3	
	HCO_3^-	CO_3^{2+}	
Weak acids	H_2O	OH^-	Strong bases
	C_2H_5OH	$C_2H_5O^-$	

hexane or carbon tetrachloride which have virtually no tendency to gain or lose protons), but in this case, both the acid and the base must be added to the solvent.

A fact which merits consideration in any theory of acids and bases is that the speeds of large numbers of chemical reactions are greatly accelerated by acids and bases. This phenomenon is called acid-base catalysis, and a major reason for its wide prevalence is that most proton transfers are themselves exceedingly fast. Hence, reversible acid-base equilibria can usually be established very rapidly, and the resulting conjugate acids (or bases) then frequently offer favorable paths for the overall chemical reaction. The mechanisms of many of these catalyzed reactions are known. Some of them are specifically catalyzed by solvated protons (hydrogen ions); others, by hydroxide ions. Still others are catalyzed by acids or bases in the most general sense of the Brönsted definitions. The existence of this general acid and base catalysis constituted an important item in the wide acceptance of the Brönsted definitions.

Lewis theory. Studies of catalysis have, however, played a large role in the acceptance of a set of quite different definitions of acids and bases, those due to G. N. Lewis. These definitions were originally proposed at about the same time as those of Brönsted, but it was not until Lewis restated them in 1938 that they began to gain wide consideration. The Lewis definitions are: an acid is a substance which can accept an electron pair from a base; a base is a substance which can donate an electron pair. (These definitions are very similar to the terms popularized around 1927 by N. V. Sidgwick and others: electron donors, which are essentially Lewis bases, and electron acceptors, which are Lewis acids.) Bases under the Lewis definition are very similar to those defined by Brönsted, but the Lewis definition for acids is very much broader. For example, virtually every cation is an acid, as are such species as $AlCl_3$, BF_3, and SO_3. An acid-base reaction now typically becomes a combination of an acid with a base, rather than a proton transfer. Even so, many of the types of reactions which are characteristic of proton acids also will occur between Lewis acids and bases, for example, neutralization and color change of indicators as well as acid-base catalysis. Furthermore, these new definitions have been useful in suggesting new interrelations and in predicting new reactions, particularly for solid systems and for systems in nonaqueous solvents.

For several reasons, these definitions have not been universally accepted. One reason is that the terms electron donor and electron acceptor had been widely accepted and appear to serve similar predicting and classifying purposes. A more important reason is unwillingness to surrender certain advantages in precision and definiteness inherent in the narrower Brönsted definitions. It is a drawback of the Lewis definitions that the relative strengths of Lewis acids vary widely with choice of base and vice versa. For example, with the Brönsted definitions, hydroxide ion is always a stronger base than ammonia; with the Lewis definitions, hydroxide ion is a much weaker base than ammonia when reacting with silver ion but is stronger than ammonia when reacting with hydrogen ion. Another feature of the Lewis definitions is that some substances which have long been obvious examples of acid, for example, HCl and H_2SO_4, do not naturally fit the Lewis definition since they cannot plausibly accept electron pairs. Certain other substances, for example, carbon dioxide, are included by calling them secondary acids. These substances, too, tend to have electronic structures in which the ability to accept electron pairs is not obvious, but the more important distinction between them and primary acids is that their rates of neutralization by bases are measurably slow. However, in spite of these difficulties, the use of the Lewis definitions is increasing. Since there does not appear to be any simultaneous tendency to abandon the Brönsted definitions, chemistry seems to be entering a period when the term acid needs a qualifying adjective for clarity, for example, Lewis acid or proton acid.

Hard and soft acids and bases. As pointed out above, one of the drawbacks of such a broad definition as the Lewis one is that it is difficult to systematize the behavior of acids and bases toward each other. Attempts have been made to classify Lewis acids and bases into categories with respect to their mutual behavior. R. G. Pearson in 1963 proposed a simple and lucid classification scheme, based in part on earlier methods, that appears to be promising in its application to a wide variety of Lewis acid-base behavior.

Lewis bases (electron donors) are classified as soft if they have high polarizability, low electronegativity, are easily oxidized, or possess low-lying empty orbitals. They are classified as hard if they have the opposite tendencies. Some bases, spanning the range of hard to soft and listed in order of increasing softness, are H_2O, OH^-, OCH_3^-, F^-, NH_3, C_5H_5N, NO_2^-, NH_2OH, N_2H_4, C_6H_5SH, Br^-, I^-, SCN^-, SO_3^{2-}, $S_2O_3^{2-}$, and $(C_6H_5)_3P$. Acids are divided more or less distinctly into two categories, hard and soft, with a few intermediate cases. Hard acids are of low polarizability, small size, and high positive oxidation state, and do not have easily excitable outer electrons. Soft acids have several of the properties of high polarizability, low or zero positive charge, large size, and easily excited outer electrons, particularly *d* electrons in metals. Some hard acids are H^+, Li^+, Na^+, K^+, Be^{2+}, Mg^{2+}, Ca^{2+}, Sr^{2+}, Mn^{2+}, Al^{3+}, Sc^{3+}, Cr^{3+}, Co^{3+}, Fe^{3+}, As^{3+}, Ce^{3+}, Si^{4+}, Ti^{4+}, Zr^{4+}, Pu^{4+}, $BeMe_2$ (Me is the methyl group), BF_3, BCl_3, $B(OR)_3$, $Al(CH_3)_3$, AlH_3, SO_3, and CO_2. Examples of soft acids are Cu^+, Ag^+, Au^+, Tl^+, Hg^+, Cs^+, Pd^{2+}, Cd^{2+}, Pt^{2+}, Hg^{2+}, CH_3Hg^+, Tl^{3+}, BH_3, $CO(CN)_5^{2-}$, I_2, Br_2, ICN, chloranil, quinones, tetracyanoethylene, O, Cl, Br, I, N, metal atoms, and bulk metals. Intermediate acids are Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Pb^{2+}, Sn^{2+}, $B(CH_3)_3$, SO_2, NO^+, and R_3C^+.

The rule for correlating acid-base behavior is then stated as follows: hard acids prefer to associate with hard bases and soft acids with soft bases. Application, for example, to the problem of OH^- and NH_3, mentioned earlier (and recognizing that OH^- is hard

compared with NH_3), leads to reaction (14), which is unfavorable compared with reaction (15).

$$OH^-(\text{hard}) + Ag^+(\text{soft}) \rightarrow AgOH \quad (14)$$

$$NH_3(\text{soft}) + Ag^+(\text{soft}) \rightarrow Ag(NH_3)^+ \quad (15)$$

However, reaction (16) is unfavorable compared with reaction (17), which is in agreement with experiment.

$$NH_3(\text{soft}) + H^+(\text{hard}) \rightarrow NH_4{}^+ \quad (16)$$

$$OH^-(\text{hard}) + H^+(\text{hard}) \rightarrow H_2O \quad (17)$$

The rule is successful in correlating general acid-base behavior of a very wide variety of chemical systems, including metal-ligand interaction, charge-transfer complex formation, hydrogen bond formation, complex ion formation, carbonium ion formation, covalent bond formation, and ionic bond formation.

It is to be emphasized that the hard-and-soft acid-and-base concept is a means of classification and correlation and is not a theoretical explanation for acid-base behavior. The reasons why hard acids prefer hard bases and soft prefer soft are complex and varied. The already well-developed concepts of ionic and covalent bonds appear to be helpful, however. Hard acids and hard bases with small sizes and high charge would be held together with stable ionic bonds. Conversely, the conditions for soft acids and soft bases would be favorable for good covalent bonding. Existing theories of π-bonding also fit into the scheme.

Usanovich theory. Another comprehensive theory of acids and bases was proposed by M. Usanovich in 1939 and is sometimes known as the positive-negative theory. Acids are defined as substances which form salts with bases, give up cations, and add themselves to anions and to free electrons. Bases are similarly defined as substances which give up anions or electrons and add themselves to cations. Two examples of acid-base reactions under this scheme are reactions (18) and (19). In the first, SO_3 is an acid

$$Na_2O + SO_3 \rightarrow 2Na^+ + SO_4{}^{2-} \quad (18)$$

$$2Na + Cl_2 \rightarrow 2Na^+ + 2Cl^- \quad (19)$$

because it takes up an anion, O^{2-}, to form $SO_4{}^{2-}$. In the second example, Cl_2 is an acid because it takes up electrons to form Cl^-. Using conventional terminology, this second reaction is an obvious example of oxidation-reduction. The fact that oxidation-reduction can also be included in the Usanovich scheme is an illustration of the extensiveness of these definitions. So far, this theory has had little acceptance, quite possibly because the definitions are too broad to be very useful.

Generation of chemical species. The acidic or basic character of a solvent can be used to stabilize interesting chemical species, which would otherwise be difficult to obtain. For example, carbonium ions have been thought to be important intermediates in many organic reactions, but because of their fleeting existence as intermediates, their properties have been difficult to study. Most carbonium ions are very strong bases and would, for example, react, as shown by reaction (20). Accordingly, the equilibrium would

$$R_3C^+ + H_2O \rightarrow R_3COH + H^+ \quad (20)$$

lie far to the right. However, use of a very strongly acidic solvent reverses the reaction, and measurable amounts of carbonium ion are then found. Concentrated sulfuric acid has found use in this connection. The very high acidity of SbF_5 by itself, as a Lewis acid, and in mixtures with other Lewis acids, such as SO_2, or protonic acids, such as HF and FSO_3H, makes possible the study of many otherwise unstable carbonium ions. *See* SUPERACID.

Acidity functions. A very different approach to the definition of acids, or perhaps better, to the definition of acidity, is to base the definition on a particular method of measurement. (As one example, it is probably true that the most nearly exact definition of pH is in terms of the electromotive force of a particular kind of galvanic cell.) It is possible to define various acidity functions in this way, and several have been proposed. One of the earliest and also one of the most successful is the H_0 acidity function of L. P. Hammett. This defines an acidity in terms of the observed indicator ratio for a particular class of indicators, those which are uncharged in the basic form B. Suppose there is available a set of such indicators, and suppose further that the values of the dissociation constants of the acid forms BH^+ are known. Then the h_0 acidity of a solution is defined as Eq. (21),

$$h_0 = K_{BH^+} \frac{[BH^+]}{[B]} \quad (21)$$

where K_{BH^+} is the dissociation constant for the particular indicator employed, and where $[BH^+]/[B]$ is the experimentally observed ratio of concentrations of the conjugate acid and conjugate base forms of the indicator. To have a logarithmic scale (analogous to pH), the further definition is expressed in Eq. (22).

$$H_0 = -\log h_0 \quad (22)$$

The virtues of this scale are that measurements are relatively simple and can be made for concentrated solutions and for solutions in nonaqueous or mixed solvents, situations where the pH scale offers difficulties. A further point is that in dilute aqueous solutions this new acidity becomes identical to pH. Although it has been found that this measure of acidity is fairly consistent within a class of indicators used, different classes can give somewhat different measures of acidity. Hence, caution must be used in interpretation of acidity measured by this technique.

For a discussion of measurement of acidity *see* BASE (CHEMISTRY); BUFFERS (CHEMISTRY); HYDROGEN ION; IONIC EQUILIBRIUM; OXIDATION-REDUCTION; SOLUTION; SOLVENT. Franklin A. Long; Richard H. Boyd

Bibliography. J. F. Coetzee and C. D. Ritchie, *Solute-Solvent Interactions*, vol. 1, 1969, vol. 2, 1976; C. W.

Hand, *Acid-Base Chemistry*, 1986; H. F. Holtzclaw, Jr., W. R. Robinson, and J. D. Odom, *General Chemistry*, 9th ed., 1991.

Acid-base indicator

A substance that reveals, through characteristic color changes, the degree of acidity or basicity of solutions. Indicators are weak organic acids or bases which exist in more than one structural form (tautomers) of which at least one form is colored. Intense color is desirable so that very little indicator is needed; the indicator itself will thus not affect the acidity of the solution.

The equilibrium reaction of an indicator may be regarded typically by giving it the formula HIn. It dissociates into H^+ and In^- ions and is in equilibrium with a tautomer InH which is either a nonelectrolyte or at most ionizes very slightly. In the overall equilibrium shown as reaction (1), the simplifying assump-

$$\underset{(1)}{\mathrm{InH}} \underset{\text{ioniz-ation}}{\rightleftharpoons} \mathrm{HIn} \underset{\text{dissoci-ation}}{\rightleftharpoons} \underset{(2)}{\mathrm{H^+ + In^-}} \qquad (1)$$

tion that the indicator exists only in forms (**1**) and (**2**) leads to no difficulty. The addition of acid will completely convert the indicator to form (**1**), which is therefore called the acidic form of the indicator although it is functioning as a base. A hydroxide base converts the indicator to form (**2**) with the formation of water; this is called the alkaline form. For the equilibrium between (**1**) and (**2**) the equilibrium constant is given by Eq. (2). In a manner similar to

$$K_{\mathrm{In}} = \frac{[\mathrm{H^+}][\mathrm{In^-}]}{[\mathrm{InH}]} \qquad (2)$$

the pH designation of acidity, that is, $\mathrm{pH} = -\log[\mathrm{H^+}]$, the K_{In} is converted to $\mathrm{pK_{In}}$ with the result shown in Eq. (3). It is seen that the pK of an indicator has

$$\mathrm{p}K_{\mathrm{In}} = \mathrm{pH} - \log\frac{[\mathrm{In^-}]}{[\mathrm{InH}]} \qquad (3)$$

a numerical value approximately equal to that of a specific pH level.

Use of indicators. Acid-base indicators are commonly employed to mark the end point of an acid-base titration or to measure the existing pH of a solution. For titration the indicator must be so chosen that its pK is approximately equal to the pH of the system at its equivalence point. For pH measurement, the indicator is added to the solution and also to several buffers. The pH of the solution is equal to the pH of that buffer which gives a color match. Care must be used to compare colors only within the indicator range. A color comparator may also be used, employing standard color filters instead of buffer solutions.

Indicator range. This is the pH interval of color change of the indicator. In this range there is competition between indicator and added base for the available protons; the color change, for example, yellow to red, is gradual rather than instantaneous. Observers may, therefore, differ in selecting the precise point of change. If one assumes arbitrarily that the indicator is in one color form when at least 90% of it is in that form, there will be uncertain color

Common acid-base indicators

Common name	pH range	Color change (acid to base)	p*K*	Chemical name	Structure	Solution
Methyl violet	0–2, 5–6	Yellow to blue violet to violet		Pentamethylbenzyl-pararosaniline hydrochloride	Base	0.25% in water
Metacresol purple	1.2–2.8, 7.3–9.0	Red to yellow to purple	1.5	*m*-Cresolsulfonphthalein	Acid	0.73% in *N*/50 NaOH, dilute to 0.04%
Thymol blue	1.2–2.8, 8.0–9.6	Red to yellow to blue	1.7	Thymolsulfonphthalein	Acid	0.93% in *N*/50 NaOH, dilute to 0.04%
Tropeoline 00 (Orange IV)	1.4–3.0	Red to yellow		Sodium *p*-diphenyl-aminoazobenzene-sulfonate	Base	0.1% in water
Bromphenol blue	3.0–4.6	Yellow to blue	4.1	Tetrabromophenol-sulfonphthalein	Acid	1.39% in *N*/50 NaOH, dilute to 0.04%
Methyl orange	2.8–4.0	Orange to yellow	3.4	Sodium *p*-dimethyl-aminoazobenzene-sulfonate	Base	0.1% in water
Bromcresol green	3.8–5.4	Yellow to blue	4.9	Tetrabromo-*m*-cresolsulfonphthalein	Acid	0.1% in 20% alcohol
Methyl red	4.2–6.3	Red to yellow	5.0	Dimethylaminoazobenzene-*o*-carboxylic acid	Base	0.57% in *N*/50 NaOH, dilute to 0.04%
Chlorphenol red	5.0–6.8	Yellow to red	6.2	Dichlorophenolsulfonphthalein	Acid	0.85% in *N*/50 NaOH, dilute to 0.04%
Bromcresol purple	5.2–6.8	Yellow to purple	6.4	Dibromo-*o*-cresolsulfonphthalein	Acid	1.08% in *N*/50 NaOH dilute to 0.04%
Bromthymol blue	6.0–7.6	Yellow to blue	7.3	Dibromothymolsulfonphthalein	Acid	1.25% in *N*/50 NaOH, dilute to 0.04%
Phenol red	6.8–8.4	Yellow to red	8.0	Phenolsulfonphthalein	Acid	0.71% in *N*/50 NaOH, dilute to 0.04%
Cresol red	2.0–3.0, 7.2–8.8	Orange to amber to red	8.3	*o*-Cresolsulfonphthalein	Acid	0.76% in *N*/50 NaOH, dilute to 0.04%
Orthocresolphthalein	8.2–9.8	Colorless to red		—	Acid	0.04% in alcohol
Phenolphthalein	8.4–10.0	Colorless to pink	9.7	—	Acid	1% in 50% alcohol
Thymolphthalein	10.0–11.0	Colorless to red	9.9	—	Acid	0.1% in alcohol
Alizarin yellow GG	10.0–12.0	Yellow to lilac		Sodium *p*-nitrobenzeneazosalicylate	Acid	0.1% in warm water
Malachite green	11.4–13.0	Green to colorless		*p,p′*-Benzylidene-bis-*N,N*-dimethylaniline	Base	0.1% in water

in the range of 90–10% InH (that is, 10–90% In^-). When these arbitrary limits are substituted into the p*K* equation, the interval between definite colors is shown by Eqs. (4). Thus the pH uncertainty of the

$$\mathrm{pH} = \mathrm{p}K + \log \frac{10}{90} \quad \text{to} \quad \mathrm{pH} = \mathrm{p}K + \log \frac{90}{10} \qquad (4)$$

indicator is from pK + 1 to pK − 1 (approximately), and pK ± 1 is called the range of the indicator. The experimentally observed ranges may differ from this prediction somewhat because of variations in color intensity.

Examples. The **table** lists many of the common indicators, their chemical names, pK values, ranges of pH, and directions for making solutions. Many of the weak-acid indicators are first dissolved in N/50 NaOH to the concentration shown, then further diluted with water. The weak-base indicators show some temperature variation of pH range, following approximately the temperature change of the ionization constant of water. Weak-acid indicators are more stable toward temperature change. The colors are formed by the usual chromophoric groups, for example, quinoid and azo-. One of the most common indicators is phenolphthalein, obtained by condensing phthalic anhydride with phenol. The acid form is colorless. It is converted by OH^- ion to the red quinoid form, as shown by reaction (5).

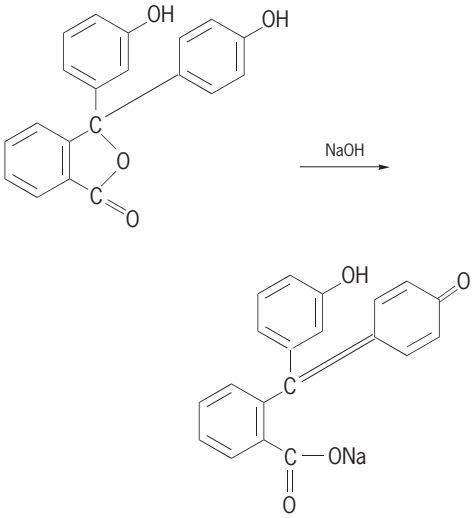

(5)

Other indiators such as methyl orange and methyl red are sodium salts of azobenzene derivatives containing sulfonic and carboxylic groups respectively. *See* ACID AND BASE; HYDROGEN ION; TITRATION.

Allen L. Hanson

Acid anhydride

One of an important class of reactive organic compounds derived from acids via formal intermolecular dehydration; thus, acetic acid (**1**), on loss of water, forms acetic anhydride (**2**).

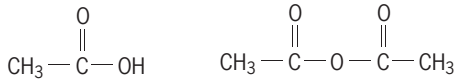

(**1**) (**2**)

Anhydrides of straight-chain acids containing from 2 to 12 carbon atoms are liquids with boiling points higher than those of the parent acids. They are relatively insoluble in cold water and are soluble in alcohol, ether, and other common organic solvents. The lower members are pungent, corrosive, and weakly lacrimatory. Anhydrides from acids with more than 12 carbon atoms and cyclic anhydrides from dicarboxylic acids are crystalline solids.

Preparation. Because the direct intermolecular removal of water from organic acids is not practicable, anhydrides must be prepared by means of indirect processes. A general method involves interaction of an acid salt with an acid chloride, reaction (1).

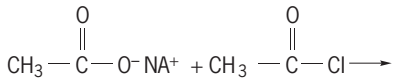

$$\mathrm{NaCl} + (\mathrm{CH_3CO})_2\mathrm{O} \qquad (1)$$

Acetic anhydride, the most important aliphatic anhydride, is manufactured by air oxidation of acetaldehyde, using as catalysts the acetates of copper and cobalt, shown in reaction (2); peracetic acid apparently is an intermediate. The anhydride is separated from the by-product water by vacuum distillation.

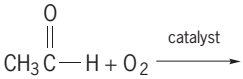

$$\mathrm{CH_3COOH} + \mathrm{CH_3CHO} \longrightarrow (\mathrm{CH_3CO})_2\mathrm{O} + \mathrm{H_2O} \qquad (2)$$

Another important process utilizes the thermal decomposition of ethylidene acetate (made from acetylene and acetic acid), reaction (3).

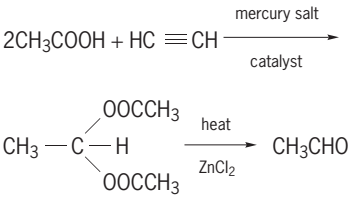

$$\mathrm{CH_3CHO} + (\mathrm{CH_3CO})_2\mathrm{O} \qquad (3)$$

Acetic anhydride has been made by the reaction of acetic acid with ketene, reaction (4).

$$\mathrm{CH_3COOH} + \mathrm{CH_2{=}C{=}O} \longrightarrow (\mathrm{CH_3CO})_2\mathrm{O} \qquad (4)$$

Mixed anhydrides composed of two different radicals are unstable, and disproportionate to give the two simple anhydrides. Direct use is made of this in the preparation of high-molecular-weight anhydrides, as seen in reaction (5). The two simple anhydrides are easily separable by distillation in a vacuum.

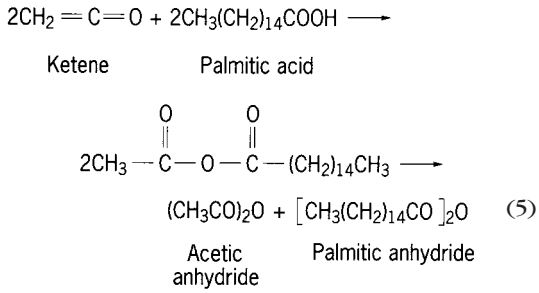

(5)

Cyclic anhydrides are obtained by warming succinic or glutaric acids, either alone, with acetic anhydride, or with acetyl chloride. Under these conditions, adipic acid first forms linear, polymeric anhydride mixtures, from which the monomer is obtained by slow, high-vacuum distillation. Cyclic anhydrides are also formed by simple heat treatment of cis-unsaturated dicarboxylic acids, for example, maleic and glutaconic acids; and of aromatic 1,2-dicarboxylic acids, for example, phthalic acid. Commercially, however, both phthalic (**3**) and maleic (**4**)

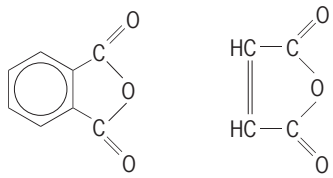

(**3**) (**4**)

anhydrides are primary products of manufacture, being formed by vapor-phase, catalytic (vanadium pentoxide), air oxidation of naphthalene and benzene, respectively; at the reaction temperature, the anhydrides form directly.

Uses. Anhydrides are used in the preparation of esters. Ethyl acetate and butyl acetate (from butyl alcohol and acetic anhydride) are excellent solvents for cellulose nitrate lacquers. Acetates of high-molecular-weight alcohols are used as plasticizers for plastics and resins. Cellulose and acetic anhydride give cellulose acetate, used in acetate rayon and photographic film. The reaction of anhydrides with sodium peroxide forms peroxides (acetyl peroxide is violently explosive), used as catalysts for polymerization reactions and for addition of alkyl halides to alkenes. In Friedel-Crafts reactions, anhydrides react with aromatic compounds, forming ketones such as acetophenone.

Maleic anhydride reacts with many dienes to give hydroaromatics of various complexities (Diels-Alder reaction). Maleic anhydride is used commercially in the manufacture of alkyd resins from polyhydric alcohols. Soil conditioners are produced by basic hydrolysis of the copolymer of maleic anhydride with vinyl acetate.

Phthalic anhydride and alcohols form esters (phthalates) used as plasticizers for plastics and resins. Condensed with phenols and sulfuric acid, phthalic anhydride yields phthaleins, such as phenolphthalein; with *m*-dihydroxybenzenes under the same conditions, xanthene dyes form, for example, fluorescein. Phthalic anhydride is used in manufacturing glyptal resins (from the anhydride and glycerol) and in manufacturing anthraquinone. Heating phthalic anhydride with ammonia gives phthalimide, used in Gabriel's synthesis of primary amines, amino acids, and anthranilic acid (*o*-aminobenzoic acid). With alkaline hydrogen peroxide, phthalic anhydride yields monoperoxyphthalic acid, used along with benzoyl peroxide as polymerization catalysts, and as bleaching agents for oils, fats, and other edibles.

Anhydrides react with water to form the parent acid, with alcohols to give esters, and with ammonia to yield amides; and with primary or secondary amines, they furnish *N*-substituted and *N*,*N*-disubstituted amides, respectively. *See* ACID HALIDE; ACYLATION; CARBOXYLIC ACID; DIELS-ALDER REACTION; ESTER; FRIEDEL-CRAFTS REACTION. Paul E. Fanta

Bibliography. F. Bettelheim and J. March, *Introduction to Organic and Biochemistry*, 1997; R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998.

Acid halide

One of a large group of organic substances possessing the halocarbonyl group

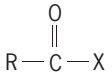

in which X stands for fluorine, chlorine, bromine, or iodine. The terms acyl and aroyl halides refer to aliphatic or aromatic derivatives, respectively.

The great inherent reactivity of acid halides precludes their free existence in nature; all are made by synthetic processes. In general, acid halides have low melting and boiling points and show little tendency toward molecular association. With the exception of the formyl halides (which do not exist), the lower members are pungent, corrosive, lacrimatory liquids that fume in moist air. The higher members are low-melting solids.

Preparation. Acid chlorides are prepared by replacement of carboxylic hydroxyl of organic acids by treatment with phosphorus trichloride, phosphorus pentachloride, or thionyl chloride [reactions (1)–(3)].

$$3RCOOH + PCl_3 \rightarrow 3RCOCl + H_3PO_3 \quad (1)$$

$$RCOOH + PCl_5 \rightarrow RCOCl + POCl_3 + HCl \quad (2)$$

$$RCOOH + SOCl_2 \rightarrow RCOCl + SO_2 + HCl \quad (3)$$

Although acid bromides may be prepared by the above methods (especially by use of PBr_3), acid iodides are best prepared from the acid chloride treatment with either CaI_2 or HI, and acid fluorides from the acid chloride by interaction with HF or antimony fluoride.

Reactions and uses. The reactivity of acid halides centers upon the halocarbonyl group, resulting in substitution ofthe halogen by appropriate

structures. Thus, as shown by reactions (4), with

$$\mathrm{RCOCl} + \begin{cases} \mathrm{HOH} \rightarrow \mathrm{HCl} + \mathrm{RCOOH} \\ \mathrm{C_2H_5OH} \rightarrow \mathrm{HCl} + \mathrm{RCOOC_2H_5} \\ \mathrm{2NH_3} \rightarrow \mathrm{NH_4Cl} + \mathrm{RCONH_2} \\ \mathrm{2HNR_2} \rightarrow \mathrm{NH_2R_2Cl} + \mathrm{RCONR_2} \end{cases} \quad (4)$$

substances containing active hydrogen atoms (for example, water, primary and secondary alcohols, ammonia, and primary and secondary amines), hydrogen chloride is formed together with acids, esters, amides, and *N*-substituted amides, respectively.

The industrially prepared acetyl and benzoyl chlorides are much used in reactions of the above type, particularly in the acetylating or benzoylating of amines and amino acids, and alcohols, especially polyalcohols such as glycerol, the sugars, and cellulose, to form amides, esters, and polyesters, respectively. In reactions (4) the by-product hydrogen chloride must be neutralized. With aliphatic acid halides, this must be done by using an excess of the amine (as shown); aromatic acid chlorides, being insoluble in water, can be used in the presence of aqueous sodium hydroxide (Schotten-Baumann reaction). An alternative technique uses the tertiary amine, pyridine, both as solvent and as neutralizing agent.

Interaction of acid chlorides with sodium salts of organic acids furnishes a general method for the preparation of acid anhydrides, as shown by reaction (5). Analogously, aromatic acyl peroxides, used

$$\mathrm{RCOCl} + \mathrm{RCOONa} \rightarrow \mathrm{NaCl} + \mathrm{(RCO)_2O} \quad (5)$$

as bleaching agents for flour, fats, and oils, and as polymerization catalysts, may be prepared from sodium peroxide and the acid halide [reaction (6)].

$$\mathrm{2C_6H_5COCl} + \mathrm{Na_2O_2} \rightarrow \mathrm{2NaCl} + \underset{\text{Benzoyl peroxide}}{\mathrm{C_6H_5COOOOCC_6H_5}} \quad (6)$$

See PEROXIDE.

Acid halides are used to prepare ketones, either by reaction with alkyl or aryl cadmium reagents [reaction (7)] or by the aluminum chloride catalyzed Friedel-Crafts reaction [reaction (8)].

$$\mathrm{RCOCl} + \mathrm{RCH_2CdCl} \rightarrow \mathrm{RCOCH_2R} + \mathrm{CdCl_2} \quad (7)$$

$$\mathrm{RCOCl} + \mathrm{C_6H_6} \xrightarrow{\mathrm{AlCl_3}} \mathrm{RCOC_6H_5} + \mathrm{HCl} \quad (8)$$

Reduction of acid halides is easily effected. In the Rosenmund method, used mainly for aromatic acid halides, hydrogen and a poisoned palladium catalyst are employed [reaction (9)]. The product here

$$\mathrm{RCOCl} + \mathrm{H_2} \xrightarrow[\text{poisoned}]{\mathrm{Pd}} \mathrm{HCl} + \mathrm{RCHO} \quad (9)$$

is the aldehyde. For reduction to the alcohol stage, the vigorous reagent lithium aluminum hydride reduces both aliphatic and aromatic acid halides [reaction (10)].

$$\mathrm{4RCOCl} + \mathrm{LiAlH_4} \rightarrow \mathrm{LiAlCl_4} + \mathrm{(RCH_2O)_4LiAl} \xrightarrow{\mathrm{4HCl}} \mathrm{4RCH_2OH} + \mathrm{LiAlCl_4} \quad (10)$$

Direct substitution of halogen (chlorine fastest, bromine more slowly) into acid aliphatic halides is relatively easy, compared with substitution in the parent acid, and takes place on the carbon next to the carbonyl group. The product is an α-halo acid halide, RCHXCOX. These compounds interact with carboxylic acids, via an equilibrium, to form an α-halo acid and an acid halide [reaction (11)].

$$\mathrm{RCHXCOX} + \mathrm{RCH_2COOH} \rightleftharpoons \mathrm{RCHXCOOH} + \mathrm{RCH_2COX} \quad (11)$$

Thus, if a small amount of acyl halide is added to a large amount of carboxylic acid, the latter can be chlorinated or brominated to completion by treatment with either chlorine or bromine (Hell-Volhard-Zelinski reaction). *See* ACYLATION; CARBOXYLIC ACID; FRIEDEL-CRAFTS REACTION. Paul E. Fanta

Bibliography. N. L. Allinger et al., *Organic Chemistry*, 2d ed., 1976; W. H. Brown, *Introduction to Organic and Biochemistry*, 4th ed., 1987.

Acid rain

Precipitation that incorporates anthropogenic acids and acidic materials. The deposition of acidic materials on the Earth's surface occurs in both wet and dry forms as rain, snow, fog, dry particles, and gases. Although 30% or more of the total deposition may be dry, very little information that is specific to this dry form is available. In contrast, there is a large and expanding body of information related to the wet form: acid rain or acid precipitation. Acid precipitation, strictly defined, contains a greater concentration of hydrogen (H^+) than of hydroxyl (OH^-) ions, resulting in a solution pH less than 7. Under this definition, nearly all precipitation is acidic. The phenomenon of acid deposition, however, is generally regarded as resulting from human activity. *See* PRECIPITATION (METEOROLOGY).

Sources. Theoretically, the natural acidity of precipitation corresponds to a pH of 5.6, which represents the pH of pure water in equilibrium with atmospheric concentrations of carbon dioxide. Atmospheric moisture, however, is not pure, and its interaction with ammonia, oxides of nitrogen and sulfur, and windblown dust results in a pH between 4.9 and 6.5 for most "natural" precipitation. The distribution and magnitude of precipitation pH in the United States (**Fig. 1**) suggest the impact of anthropogenic rather than natural causes. The areas of highest precipitation acidity (lowest pH) correspond to areas within and downwind of heavy industrialization and urbanization where emissions of sulfur and nitrogen oxides are high. It is with these emissions that the most acidic precipitation is thought to originate. *See* ACID RAIN.

Atmospheric processes. The transport of acidic substances and their precursors, chemical reactions, and deposition are controlled by atmospheric processes. In general, it is convenient to distinguish between physical and chemical processes, but it must

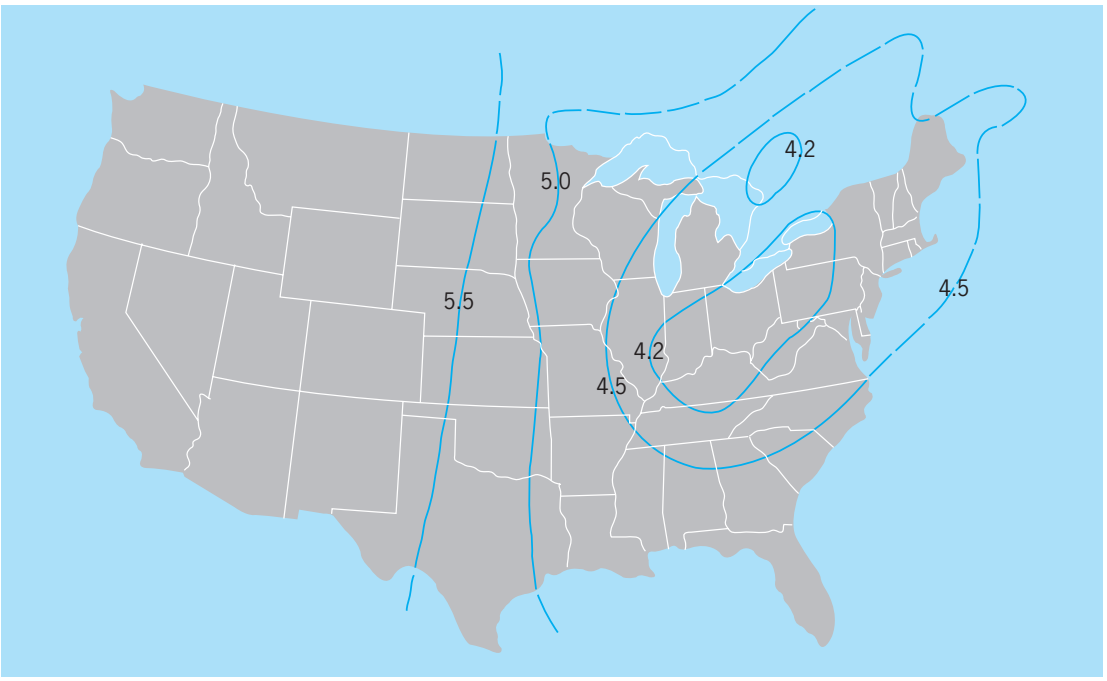

Fig. 1. Distribution of rainfall pH in the eastern United States.

be realized that both types may be operating simultaneously in complicated and interdependent ways. The physical processes of transport by atmospheric winds and the formation of clouds and precipitation strongly influence the patterns and rates of acidic deposition, while chemical reactions govern the forms of the compounds deposited.

In midlatitude continental regions, such as eastern North America, most precipitation arises from cyclonic storms. As shown by the schematic representation of a typical storm (**Fig. 2**), rain tends to form along the surface cold and warm fronts that define the so-called warm sector of the storm. This characteristic structure, arising from prevailing north-south temperature gradients, simultaneously sets the stage for chemical transformations of pollutants and for incorporation of the compounds into precipitation. The motion of the cold front toward the southeast forces the moist warm-sector air to stream along the cold front toward the low-pressure region (L), as shown by the broad arrow in Fig. 2. At the same time, the air is gradually lifted out of the surface layer and into the colder, upper parts of the storm, which allows the supply of water vapor to condense out, forming the cloud and precipitation. *See* CYCLONE; FRONT; STORM.

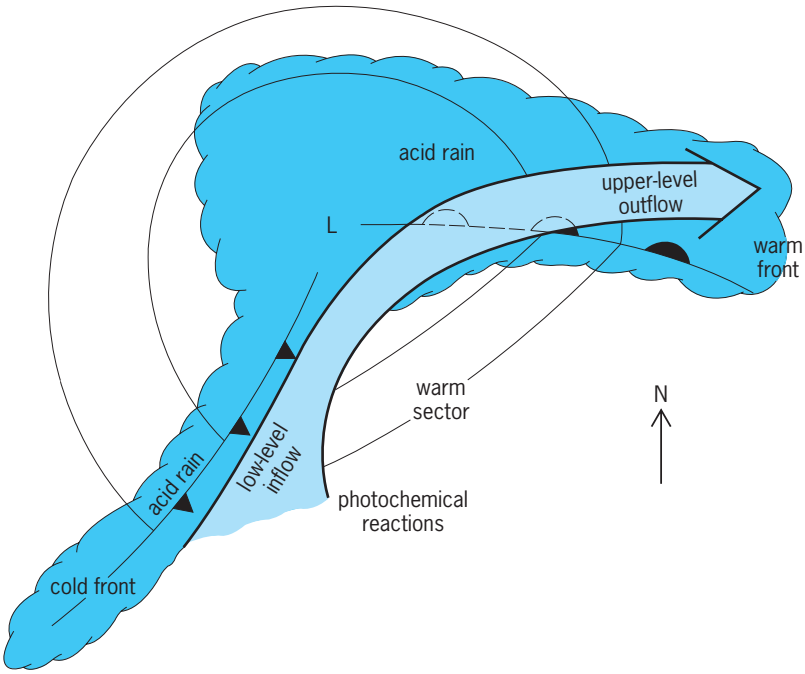

Fig. 2. Schematic view of a typical cyclonic storm in eastern North America with a low-pressure region (L).

There are a number of chemical pathways by which the primary pollutants, sulfur dioxide (SO_2) from industry, nitric oxide (NO) from both industry and automobiles, and reactive hydrocarbons mostly from trees, are transformed into acid-producing compounds. Some of these pathways exist solely in the gas phase, while others involve the aqueous phase afforded by the cloud and precipitation. As a general rule, the volatile primary pollutants must first be oxidized to more stable compounds before they are efficiently removed from the atmosphere. Ironically, the most effective oxidizing agents, hydrogen peroxide (H_2O_2) and ozone (O_3), arise from photochemical reactions involving the primary pollutants themselves. *See* AIR POLLUTION.

All of the ingredients needed to form the strong mineral acids of sulfur and nitrogen [sulfuric acid (H_2SO_4) and nitric acid (HNO_3)], which constitute most of the acids found in rain, exist in the warm sector. Especially in summertime, stagnant air conditions trap the pollutants under clear skies for several

days, permitting photochemical reactions to initiate other gas-phase reactions. Nitrogen oxides (NO_x) in combination with ultraviolet light from the Sun, reactive hydrocarbons, atmospheric oxygen, and water vapor simultaneously give rise to HNO_3 vapor, odd hydrogen radicals (particularly OH), and the strong oxidants H_2O_2 and O_3. Meanwhile, slow oxidation of sulfur dioxide (SO_2), initiated by reaction with the OH radical, leads to the gradual buildup of sulfate ($SO^{-2}{}_4$) aerosol, and hazy skies result from the highly condensable nature of H_2SO_4. While some dry deposition does occur under these conditions, the warm sector of midlatitude cyclones is a region conducive to the gradual accumulation of primary pollutants, oxidants, and acid-forming compounds in the lower atmosphere. *See* SMOG.

As the cold front approaches any given location in the warm sector, the airborne acids and their precursors are drawn into the circulations of the cyclone, particularly into the stream of moist air associated with the fronts. By this large-scale meteorological process, the pollutants become integrally associated with the frontal cloud systems, and attach to individual cloud particles by a variety of microphysical processes. Since most atmospheric sulfate particles are very soluble, they act as good centers (nuclei) for cloud drop formation. Thus, the process of cloud formation is itself a mechanism for scavenging particulate pollutants from the air. As the cloud drops grow by condensation, soluble gases will be absorbed by the cloud water. It is during this phase of the overall process that additional SO_2 will be oxidized to H_2SO_4 by the previously formed strong oxidants, particularly the H_2O_2.

The actual removal of pollutants from the atmosphere by wet deposition requires the formation of precipitation within the clouds. Without cloud elements greater than about 100 micrometers in diameter, the pollutant mass remains in the air, largely in association with the relatively small cloud drops, which have negligible rates of descent. Since most precipitation in midlatitude storms is initiated by the formation of ice particles in the cold, upper reaches of the clouds, pollutant fractionation between water phases inhibits the transfer of cloud water acidity to large particles, which do precipitate readily. Partly because of such microphysical phenomena and partly because of nonuniform distributions of pollutants within the clouds, the acidity of precipitation tends to be substantially less than that of the cloud water that remains aloft. The acidity of the precipitation appears to be acquired largely through collisions of the melted ice particles with the relatively concentrated cloud drops in the lower portions of the storm clouds. The rate of wet deposition is thus governed by a complicated set of meteorological and microscale processes working in harmony to effect a general cleansing of the atmosphere. *See* ATMOSPHERIC CHEMISTRY; CLOUD PHYSICS.

Terrestrial and aquatic effects. The effect of acid deposition on a particular ecosystem depends largely on its acid sensitivity, its acid neutralization capability, the concentration and composition of acid reaction products, and the amount of acid added to the system. As an example, the major factors influencing the impact of acidic deposition on lakes and streams are (1) the amount of acid deposited; (2) the pathway and travel time from the point of deposition to the lake or stream; (3) the buffering characteristics of the soil through which the acidic solution moves; (4) the nature and amount of acid reaction products in soil drainage and from sediments; and (5) the buffering capacity of the lake or stream.

In many ecosystems, except for foliar effects, the impacts of acid precipitation on aquatic and terrestrial ecosystems overlap. This is because soils are the key intermediate. They provide the root environment for terrestrial vegetation, and also control the water quality of runoff and soil drainage, which supplies most of the water to the aquatic system. A number of acid-consuming reactions occur in soil, which lessen the impact of acid additions on both the soil and soil drainage waters. In soils, indigenous or agriculturally amended carbonates (CO_3^{2-}) react with acid precipitation to raise the pH of soil drainage waters, while maintaining soil pH. Also, soils exhibit a cation exchange capacity that can serve to neutralize acid precipitation. At neutral soil pH (6.0–8.0), most cations on the exchange are calcium (Ca) and magnesium (Mg). When acids are added, H^+ from solution exchanges with the adsorbed Ca^{2+} and Mg^{2+}. Although this reduces the acidity of soil drainage waters, the soil acidity is increased. Eventually, when the neutralizing carbonates and exchangeable calcium and magnesium supplies are exhausted, soil minerals react with the acid. At a soil pH of about 5.2 or less, substantial aluminum (Al) can be solubilized from the dissolution of clay minerals and coatings on soil particles. *See* SOIL CHEMISTRY.

Acid deposition directly into lakes and streams causes chemical reactions analogous to those in the soil system. However, instead of soil materials, the carbonate-bicarbonate (HCO_3^-) system buffers the solution. As with soils, waters of low pH and buffering capacities, in this case bicarbonate concentrations, are most sensitive. At solution pH of 4.5, bicarbonate is essentially depleted, and subsequent acid additions that reduce pH also mobilize metals from suspended solids and the lake or stream bed. Generally, lakes and streams that drain acid-susceptible soils are most susceptible to direct acid deposition as well. Shallow soils with low pH, low cation-exchange capacity, and high permeability not only are least able to neutralize acid flows but also are poor sources of the bicarbonate waters needed to buffer the aquatic system. *See* BUFFERS (CHEMISTRY).

While the study of acid precipitation effects on terrestrial ecosystems is relatively new, soil acidification is a naturally occurring process in humid climates and has long been the subject of research, whose findings suggest acid precipitation effects. The generally accepted impact of soil acidification on the productivity of terrestrial plants is summarized as follows. As soil becomes more acidic, the basic cations (Ca, Mg) on the soil exchange are replaced by

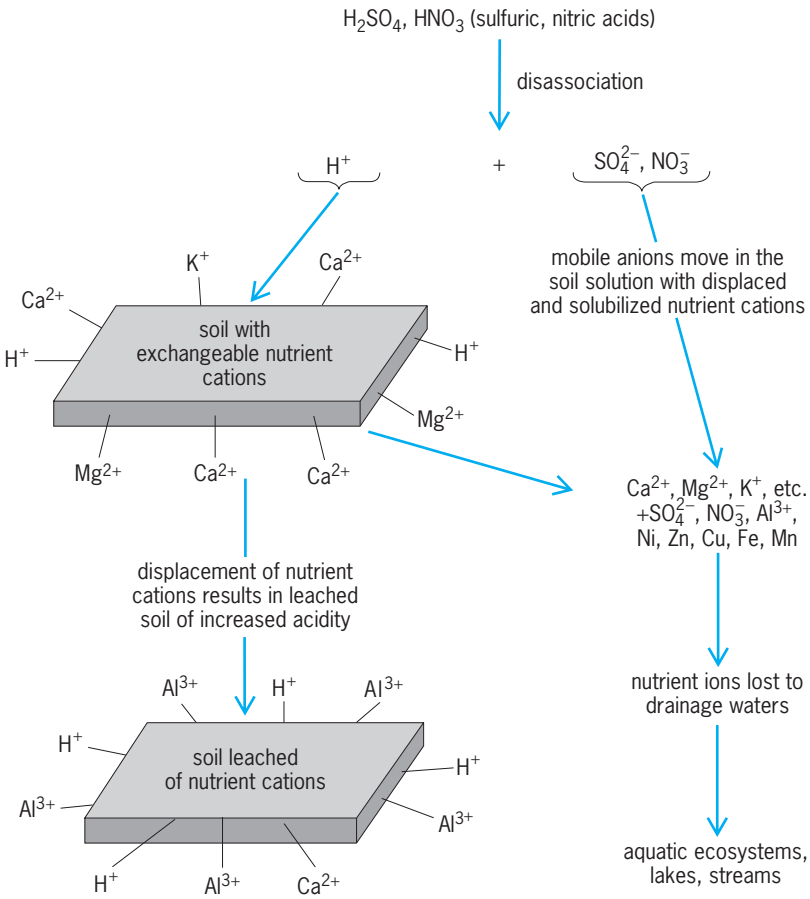

Fig. 3. Soil acidification and loss of soil nutrients following acid inputs.

hydrogen ions or solubilized metals. The basic cations, now in solution, can be leached through the soil (**Fig. 3**). As time progresses, the soil becomes less fertile and more acidic. Resultant decreases in soil pH cause reduced, less-active populations of soil microorganisms, which in turn slow decomposition of plant residues and cycling of essential plant nutrients. The plant availability of phosphorus, now composing mostly aluminum and iron (Fe) phosphates, decreases, while the plant availability of trace metals [aluminum, copper (Cu), iron, zinc (Zn), boron (B), manganese (Mn)] increases, sometimes to phytotoxic levels.

Acid precipitation may injure trees directly or indirectly through the soil. Foliar effects have been studied extensively, and it is generally accepted that visible damage occurs only after prolonged exposure to precipitation of pH 3 or less (for example, acid fog or clouds). Measurable effects on forest ecosystems will then more likely result indirectly through soil processes than directly through exposure of the forest canopy. The characteristics that make a soil sensitive to acidification are extremely variable from soil to soil. Accordingly, forest decline will not occur regionally across a deposition gradient, but it will occur in localized areas with sensitive soils. The occurrence of both acid-sensitive and acid-tolerant soils in areas receiving high levels of acid deposition and in areas receiving low levels of acid deposition makes cause-and-effect relationships between acid deposition and forest decline difficult to establish. For example, 10 years of monitoring indicated that hardwood forests in Ontario were on average healthy and stable. Still, tree condition was poorer in areas with both moderately high pollution deposition and acid-sensitive soils. The clearest cause-and-effect relationship exists for red spruce and white birch exposed to acid clouds and fog. Areas of declining red spruce and white birch coincide with areas of acid clouds and fog, and the extent of decline is related to the duration of exposure.

Many important declines in the condition of forest trees have been reported in Europe and North America during the period of increasing precipitation acidity. These cases include injury to white pine in the eastern United States, red spruce in the Appalachian Mountains of eastern North America, and many economically important species in central Europe. Since forest trees are continuously stressed by competition for light, water, and nutrients; by disease organisms; by extremes in climate; and by atmospheric pollutants, establishing acid deposition as the cause of these declines is made more difficult. Each of these sources of stress, singly or in combination, produces similar injury. However, a large body of information indicates that accelerated soil acidification resulting from acid deposition is an important predisposing stress that in combination with other stresses has resulted in increased decline and mortality of sensitive tree species and widespread reduction in tree growth. *See* FOREST ECOSYSTEM; TERRESTRIAL ECOSYSTEM.

Aquatic biology effects. Aquatic ecosystems are sustained by intricate exchanges of energy and material among organisms at all levels. Acidic deposition impacts aquatic ecosystems by harming individual organisms and by disrupting flows of energy and materials through the ecosystem. The effect of acid deposition is commonly assessed by studying aquatic invertebrates and fish. Aquatic invertebrates live in the sediments of lakes and streams and are vitally important to the cycling of energy and material in aquatic ecosystems. These small organisms, rarely larger than 20 mm (1 in.), break down large particulate organic matter for further degradation by microorganisms, and they are an important food source for fish, aquatic birds, and predatory invertebrates.

Individual aquatic organisms are harmed through direct physiological effects caused by low pH, high metal concentrations, and the interaction between pH and metals. Low pH causes lesions to form on gills and erodes gill tissue. Gill damage compromises respiration, excretion, and liver function, which in turn leaves organisms susceptible to further health problems. As pH decreases, the aluminum concentration of streams and lakes increases. Aluminum enhances the toxic effects of low pH. Together these phenomena result in severe acidosis and cause gills to secrete excessive amounts of mucus, further

interfering with respiration by reducing oxygen diffusion across gills. Additionally, gill membranes become more permeable, reducing sodium uptake and increasing sodium loss. The most likely physiological cause of fish death at pH 4–5 is the inability to maintain a proper salt balance (ion regulation).

Metals such as aluminum liberated from acidified soils may accumulate externally on the gill surfaces of invertebrates, and this accumulation also impairs respiration. Iron has been shown to accumulate on the gut membranes of invertebrates, inhibiting food absorption. It should be noted that the formation of metallo-organic complexes reduces the toxicity of metals in low-pH waters. Acid bogs and marshes support healthy populations of aquatic organisms because soluble metals and humic compounds form such complexes.

Currently, there are concerns that acid deposition is causing the loss of fish species, through the physiological damage discussed above and by reproductive impairment. In order to find water with a suitable pH, fish may be forced to spawn in areas that are less than optimal for egg development, with a consequent reduction in the number of, or total loss of, offspring. Female fish need high serum calcium levels for normal ovary formation and egg laying. Female fish in acidified waters maintain lower than normal serum calcium levels, and consequently egg laying may be inhibited, again resulting in fewer or no offspring. Early life stages of fish (eggs and fry) are usually more sensitive than adults to low pH, and acute mortality of these life stages may be high during periods of rapid declines in pH following rainfall. The death of fish embryos is thought to result from the corrosion of skin cells, which interferes with respiration and osmoregulation. While fish die from acidification, their numbers and diversity are more likely to decline from a failure to reproduce.

The effects of acid deposition on individuals described above in turn elicit changes in the composition and abundance of communities of aquatic organisms. The degree of change depends on the severity of acidification, and the interaction of other factors, such as metal concentrations and the buffering capacity of the water. The pattern most characteristic of aquatic communities in acidified waters is a loss of species diversity, and an increase in the abundance of a few, acid-tolerant taxa. At moderately affected sites, acid-sensitive species are frequently replaced by more acid-tolerant forms, while the total number or mass of organisms may not change. For example, ephemeropterans (mayflies) are quite sensitive to reductions in pH and their numbers may decline, while plecopterans (stone flies) and dipterans (true flies) become relatively more dominant following acidification. Invertebrates in the feeding group that shred leaves also become relatively more dominant than those that graze on diatoms in acid-impacted streams. As pH decreases further and toxic metals increase, both the number of species and the number of organisms decrease sharply (**Fig. 4**). Mollusks and crustaceans are particularly sensitive to acidification, since both groups require high levels of calcium to form their shells. Calcium is scarce in the poorly buffered systems that are most affected by acidic deposition.

Community-level effects may occur indirectly, as a result of changes in the food supply and in predator-prey relations. Reduction in the quality and amount of periphyton may decrease the number of herbivorous invertebrates, which may in turn reduce the number of organisms (predatory invertebrates and fish) that feed upon herbivorous invertebrates. The

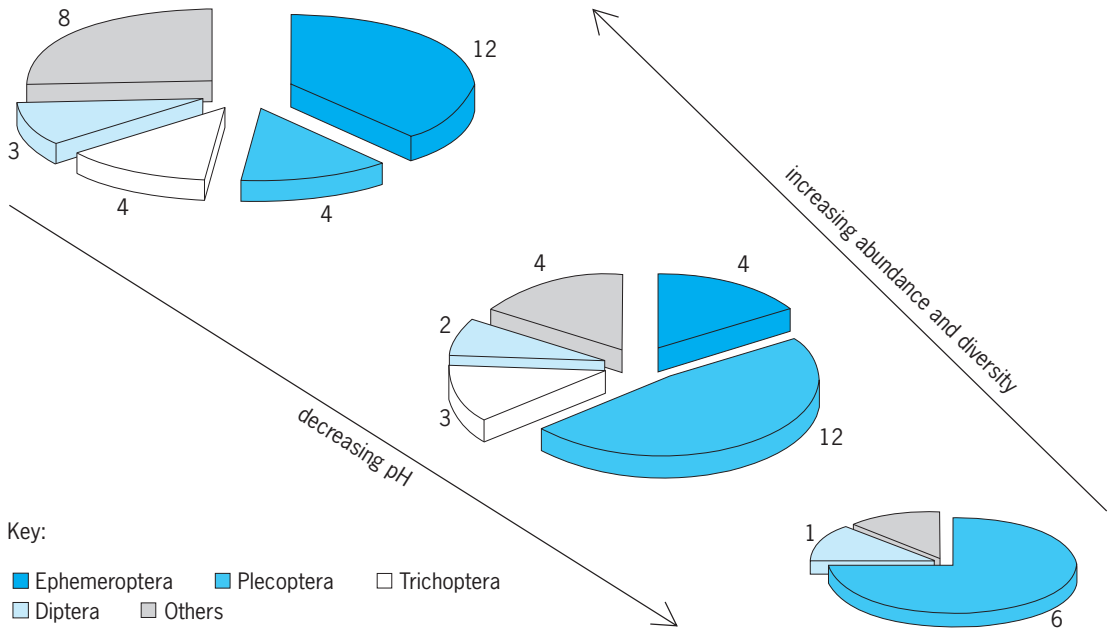

Fig. 4. General pattern of aquatic invertebrate abundance and diversity with a reduction in pH. Total abundance is indicated by disk sizes. Numbers of species within each order are indicated.

disappearance of fish may result in profound changes in plant and invertebrate communities. Dominant fish species function as keystone predators, controlling the size distribution, diversity, and numbers of invertebrates. Their reduction alters the interaction within and among different levels of the food web and the stability of the ecosystem as a whole.

Acid deposition affects the survival of individual organisms and the relationship of communities within aquatic ecosystems. Early life stages are generally more sensitive to low pH then adult organisms. Directs effects of pH and aluminum toxicity affect survival more than the indirect effects of food availability or predation.

The rate and magnitude of acidification of aquatic systems depend upon the factors discussed above. Predictable changes in water chemistry, as the pH of surface waters decreases, include decreased alkalinity, decreased buffering capacity, and increased concentrations of aluminum, magnesium, and iron, among other elements. Such changes in water chemistry generally result in decreased diversity of aquatic species and reduced productivity.

The impact of acid deposition on terrestrial and aquatic ecosystems is not uniform. While increases in acid deposition may stress some ecosystems and reduce their stability and productivity, others may be unaffected. The degree and nature of the impact depend on the acid input load, organismal susceptibility, and buffering capacity of the particular ecosystem. *See* BIOGEOCHEMISTRY; ECOSYSTEM.

Ronald R. Schnabel; Dennis Lamb; Harry B. Pionke; Dennis Genito

Bibliography. *Acid Deposition: Atmospheric Processes in Eastern North America*, National Academy Press, 1983; *Acid Rain and Transported Air Pollutants: Implications for Public Policy*, U.S. Congress, Office of Technology Assessment, 1984; R. J. Charlson and H. Rohde, Factors controlling the acidity of natural rainwater, *Nature*, 195:683–685, 1982; F. M. D'Itri (ed.), *Acid Precipitation: Effects on Ecological Systems*, 1982; D. L. Godbold and A. Hüttermann (eds.), *Effects of Acid Rain on Forest Processes*, Wiley-Liss, New York, 1994; R. M. Heard et al., Episodic acidification and changes in fish diversity in Pennsylvania headwater streams, *Trans. Amer. Fisheries Soc.*, 126:977–984, 1997; *Interim Assessment: The Causes and Effects of Acidic Deposition*, National Acid Precipitation Assessment Program, Washington, D.C., 1987; W. G. Kimmel et al., Macroinvertebrate community structure and detritus processing rates in two southwestern Pennsylvania streams acidified by atmospheric deposition, *Hydrobiologia*, 124:97–102, 1985; W. P. Robarge and D. W. Johnson, The effects of acidic deposition on forested soils, *Adv. Agron.*, 47:1–83, 1992; A. D. Rosemond et al., The effects of stream acidity on benthic invertebrate communities in the south-eastern United States, *Freshwater Biol.*, 27:193–209, 1992; J. N. Wordman and E. B. Cowling, Airborne chemicals and forest health, *Environ. Sci. Technol.*, 21:120–126, 1987.

Acipenseriformes

An order of actinopterygian fishes in the subclass Chondrostei, which probably originated from palaeonisciform ancestors during the Triassic or Jurassic periods. Acipenseriformes, the only extant order of about 10 orders of chondrosteans, are characterized by a largely cartilaginous skeleton, heterocercal tail, fins with more than one ray per pterygiophore, spiral valve intestine, myodome and preopercle reduced or absent, and gular plate absent. The cartilaginous skeleton of Acipenseriformes, long regarded as primitive, was taken as indicative of an alliance with the Chondrichthyes; however, cartilage is secondary and the true relationship is with bony fishes. *See* ACTINOPTERYGII; CHONDROSTEI; OSTEICHTHYES.

Acipenseriformes includes two suborders: the fossil Chondrosteoidei, which became extinct in the Mesozoic, and the Acipenseroidei, with two families, Acipenseridae and Polyodontidae, both with little apparent change from the Cretaceous.

Acipenseridae (sturgeons). Sturgeons are characterized by a robust body armed with five longitudinal rows of bony plates; a pronounced snout bearing sensitive barbels on the underside; an inferior mouth, which is protrusible and, in adults, toothless; spiracles present or absent; gill rakers fewer than 50; a large swim bladder; and pectoral fins, the leading edge of which consists of fused rays.

The family includes four extant genera and about 25 species in two subfamilies. The subfamily Acipenserinae comprises three genera: *Acipenser* (see **illustration**) is distinguished by having spiracles and a subconical snout, and is known from North America, Europe, and Asia. In *Pseudoscaphirhynchus*, which is endemic to the Aral Sea basin, spiracles are absent, the snout is depressed, and the caudal peduncle is short, depressed, and incompletely armored. *Scaphirhynchus* is similar to *Pseudoscaphirhynchus* but differs in having a long, depressed, and completely armored caudal peduncle; it is endemic to the Mississippi basin of North America.

The subfamily Husinae has two species: *Huso dauricus* (kaluga), which is native to the Amur River in Russia and China, and *Huso huso* (beluga), which occurs in the Black and Caspian seas and rarely in the Adriatic Sea. The beluga holds the distinction of being the largest freshwater fish in the world. Just how large is uncertain, as some records are in terms of length, others by weight. Hearsay has belugas reaching 6–8 m (20–26 ft) in length, but their weights are unknown. If such giants ever existed, it would have been well over 100 years ago. FishBase

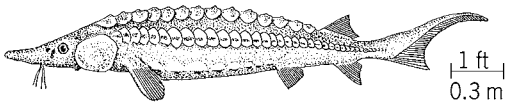

Lake sturgeon (*Acipenser fulvescens*).

reports a 5-m (16.5-ft) 2072-kg (4558-lb) specimen estimated to be 118 years old.

Sturgeons are sluggish, slow-growing, long-lived fishes found in Eurasia and North America. All spawn in freshwater, but several species spend most of their lives in the sea. Valued commercially for their flesh and as the source of caviar, sturgeons have long been seriously depleted in North America and most of Europe, but a substantial fishery is maintained in Western and Central Asia. *See* STURGEON.

Family Polyodontidae (paddlefishes). Paddlefishes have a long, paddlelike snout. There are no large plates and scutes as in sturgeons, but patches of small ganoid scales; minute barbels on the underside of the snout; spiracles present; minute teeth; and greatly extended gill covers. There are several fossil genera known primarily from the Cretaceous Period in China and North America (Montana and Wyoming) and two living species, *Psephurus gladius* from China and *Polyodon spathula* from the United States.

Polyodon differs from *Psephurus* in having long gill rakers that number in the hundreds and are used as a filter in feeding on plankton, in contrast to gill rakers that are shorter and fewer in number in *Psephurus*; and scales mostly on the caudal pecuncle and caudal fin in contrast to scales mostly on the trunk. *Psephurus* is the larger, reaching 3 m (10 ft) in total length. The largest *Polyodon* of recent record, weighing 65 kg (143 lb), was caught in the Missouri River in Montana in 1973.

It is interesting to speculate on the utility of the paddle of the paddlefish. A common belief of nineteenth and early twentieth century fishers and professional naturalists was that the paddle is used to stir the bottom sediments to dislodge food, which then would be filtered from the water. Digging utility is not likely because the process would damage the numerous sensory and electroreceptors on the paddle. It is more likely that the paddle is used in detecting food. Another interesting hypothesis is that the paddle functions as a stabilizer to prevent dipping of the head that might occur as the fish cruises through water with its large mouth agape.

Herbert Boschung; Reeve M. Bailey

Bibliography. W. E. Bemis, E. K. Findeis, and L. Grande, An overview of Acipenseriformes, *Environ. Biol. Fishes*, 48:25–71, 1997; W. E. Bemis and B. Kynard, Sturgeon rivers: An introduction to Acipenseriformes biogeography and life history, *Environ. Biol. Fishes*, 48:167–183, 1997; V. J. Birstein and W. E. Bemis, How many species are there within the genus *Acipenser*?, *Environ. Biol. Fishes*, 48:157–163, 1997; V. J. Birstein, P. Doukakis, and R. DeSalle, Molecular phylogeny of Acipenseridae: Nonmonophyly of Scaphirhynchinae, *Copeia*, 2002(2):287–301, 2002; A. Choudhury and T. A. Dick, The historical biogeography of sturgeons (Osteichthyes: Acipenseridae): A synthesis of phylogenetics, paleontology and palaeogeography, *J. Biogeog.*, 25:623–640, 1998; R. Froese and D. Pauly (eds.), FishBase, version 05/2005; L. Grande and W. E. Bemis, Interrelationships of Acipenseriformes with comments on "Chondrostei," pp. 85–115 in M. L. J. Stassny, L. R. Parenti, and G. D. Johnson (eds.), *Interrelationships of Fishes*, Academic Press, San Diego, 1996; L. Grande and W. E. Bemis, Osteology and phylogenetic relationships of fossil and recent paddlefishes (Polyodontidae) with comments on the interrelationships of Acipenseriformes, *J. Vert. Paleontol.*, vol. 11, Mem. 1, 1991; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Acoela

An order of marine platyhelminthes of the class Turbellaria, generally regarded as primitive, and lacking protonephridia, oviducts, yolk glands, a permanent digestive cavity, and strictly delimited gonads. The nervous system is epidermal in the most primitive representatives and resembles that of cnidarians, but it is generally submuscular and consists of three to six pairs of longitudinal strands without anterior concentrations recognizable as a brain. The anterior or median ventral mouth opens into a simple pharynx which lacks elaborate musculature and leads directly to the syncytial vacuolated mass of endoderm forming the spongy intestinal tissue and lacking any true lumen.

Eyes are often absent; when present they are of the usual paired turbellarian pigment-cup type, with one to a few pigment and retinal cells. A single (sometimes double) statocyst and sensory bristles (modified cilia) are generally present anteriorly, and clusters of frontal glands open by one or more pores at the anterior tip. The epidermis, typically, is uniformly ciliated, and locomotion is accomplished by means of rapid, smooth ciliary gliding or swimming.

The acoels (see **illus.**) are mostly small (one to several millimeters in length), of simple contour, and plump and broad- to elongate-oval in form. Their

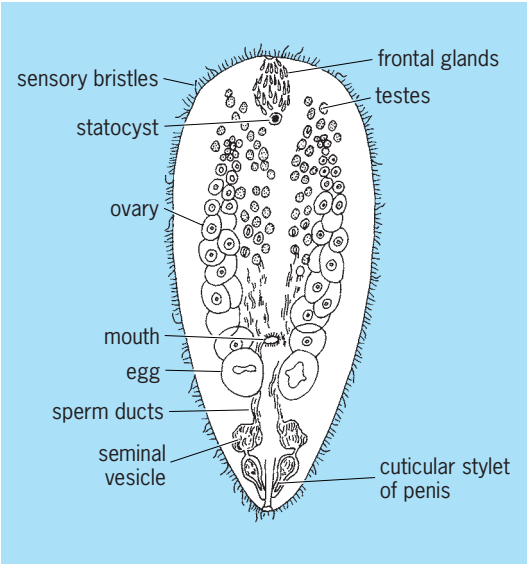

Childia spinosa.

coloration is generally drab, or inconspicuous grayish-white, unless the body is colored brown or green by symbiotic algae. They are virtually worldwide in distribution, living beneath stones, among algae, on the bottom mud, or interstitially, from the littoral zone to deeper waters. A few are pelagic (in tropical and subtropical zones), and others have become symbiotic, living entocommensally in various invertebrates but mainly in echinoderms.

Free-living species feed on diatoms, protozoans, and other microscopic organisms by partially extruding the digestive syncytium through the mouth and engulfing the food, ameboid pseudopodial fashion, in temporary vacuoles in the syncytium. Larger organisms (crustacean larvae, adult copepods, and the like) are seized by the anterior body (which may be expanded or curved dorsoventrally); the grip is facilitated by sticky mucus from the frontal glands, and the prey is thrust through the mouth into a larger temporary digestive cavity. Digestion is rapid and effected by acid and alkaline proteases, carbohydrases, and lipases acting in sequence and secreted from the syncytium. Entocommensal species feed and digest in the same manner, the presence of appropriate cosymbiotic organisms permitting retention of the basic nutritional physiology characteristic of their free-living relatives.

The Acoela, as primitive Turbellaria, are of considerable interest phylogenetically since they show features reminiscent of the cnidarian grade of organization as well as others foreshadowing those of more elaborate turbellarians. In particular, their structure supports the hypothesis that ancient planula or planuloid larvae of hydrozoan cnidarians, by developing gonads and an oral aperture, gave rise to an acoeloid ancestral stock from which all modern flatworms, rhynchocoelans, and the coelomate Metazoa have subsequently evolved. *See* CNIDARIA; PLATYHELMINTHES; TURBELLARIA. J. B. Jennings

Aconchulinida

An order of the subclass Filosia comprising a small group of naked amebas which form filamentous pseudopodia (filopodia). The genus *Penardia*, possibly the only valid one in the order, includes filopodia-forming amebas with variable shape, an outer zone of clear ectoplasm, and inner cytoplasm often containing many vacuoles and sometimes zoochlorellae (see **illus.**). Pseudopodia are sometimes extended primarily at the poles of the organism, but also may fuse laterally to form small sheets of clear cytoplasm. A single nucleus and one water-elimination vesicle are characteristic of these amebas. Size of described species ranges from less than 10 to about 400 micrometers. *See* PROTOZOA; RHIZOPODEA; SARCODINA; SARCOMASTIGOPHORA. Richard P. Hall

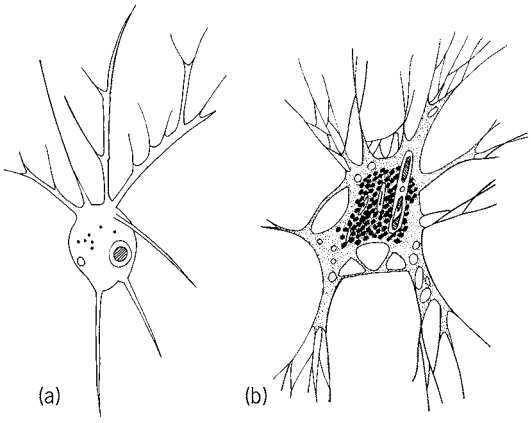

Aconchulinida. (*a*) *Penardia cometa*. (*b*) *Penardia mutabilis*.

Acorales

An order of monocotyledonous angiosperms composed of a single genus, *Acorus* (sweet flag or sweet calamus), with two species widespread in the Northern Hemisphere. Formerly, *Acorus* was included in the aroid family, Araceae, but several lines of evidence, including deoxyribonucleic acid (DNA) sequences, have firmly established it as the first-branching lineage of the monocotyledons. These species are emergent aquatics with peculiar inflorescences, and they are unusual among the monocots in having characters otherwise confined to the magnoliid dicots, such as the anther formation and the presence of ethereal oils. These oils are the basis of the species' frequent use as a medicine to treat toothache and dysentery, and as an ointment in religious ceremonies. *See* MAGNOLIOPHYTA. Mark W. Chase

Acoustic emission

A method of nondestructive testing and materials characterization that uses mechanical waves moving through materials. It is similar to seismology, except in being concerned with the scale of engineering structures, such as aircraft, bridges, and chemical tanks. When a structure is subjected to external force (or stress), a defect (for example, a crack or welding flaw) on the structure is activated and enlarged dynamically, and thus generates waves, which spread through materials at a certain speed. Such waves, known as acoustic emission signals, are detected by sensors attached on the surfaces of the structure. Mechanical vibration due to acoustic emission signals is weak and requires high-sensitivity sensors and electronic amplification before it can be analyzed. *See* SEISMOLOGY.

In nondestructive testing of structures, acoustic emission signals are typically evaluated in order to know if the failure of a structure is imminent; if cracks and other defects, presumed to be present in any structure, are active; the positions of such active defects; and whether a structure with such defects can be safely operated. In evaluating materials behavior and quality, acoustic emission is used to assess how a material (such as a steel or an aluminum alloy) responds to mechanical stress, that is, when and how it changes shape permanently and how it proceeds to

eventual fracture; how an alloy withstands repeated application of stress (known as fatigue); the level of stress and corrosive environment that lead to failure of a material; and the types of microscopic failure processes that arise in a material under stress. *See* METAL, MECHANICAL PROPERTIES OF; PLASTIC DEFORMATION OF METAL; STRESS AND STRAIN.

Methodology. Acoustic emission signals emanating from their sources contain information about the source, such as the direction and speed of crack opening. For example, the high-speed cracking of brittle materials (such as high-strength steels and ceramics) produces short, fast-varying acoustic emission signals (**Fig. 1**), which are typically plotted against time measured in microseconds. In contrast, slow-growing defects in plastics result in longer, slowly varying signals (**Fig. 2**), which are typically plotted on a time scale of milliseconds. Because of the distortion of waves during the propagation through a complex structure and detection by a sensor, however, much of the information is lost. Thus, the presence of detectable acoustic emission signals is the most important clue in assessing the integrity of the structure.

By detecting one such signal at multiple sensor positions, the location of its source can be determined from the timing of signal arrivals. The basic principle of triangulation is the same as practiced in seismology, except that the differences in signal arrival times are of the order of microseconds to milliseconds. The speed of wave propagation is a material constant, determined by the stiffness and mass density of the propagating medium. Typically, the wave speed in solids is several kilometers per second. (By comparison, the sound speed in air is about 330 m or 1080 ft per second.) *See* SOUND.

An example of a sensor arrangement is shown in **Fig. 3** for a chemical tank. Here, the sensors are typically separated by several meters to 10 m (33 ft), and cover the entire surface of the tank with blocks of roughly triangular shape. The position of an acoustic emission source within each block is determined from the arrival times at the three surrounding sensors. A typical plot of acoustic emission source positions is shown in **Fig. 4**, indicating a concentration of sources.

The detection of acoustic emission signals requires a sensor possessing high sensitivity. A typical sensor uses a piezoelectric ceramic element, which converts mechanical vibration into an electrical signal, which can be amplified 1000 to 10,000 times. Various electrical measurement techniques are used to characterize and analyze the signals received. It is common to obtain and record several features of acoustic emission signals. These form the basis of real-time analysis and decision-making. In laboratory studies, the entire waveforms are also recorded for detailed analysis after testing. *See* PIEZOELECTRICITY.

Applications. By considering observed acoustic emission signal strength, correspondence to possible defective sites, and comparison to previous acoustic emission observations at the same position, the seriousness of the presumed defect is estimated, sometimes resulting in immediate shut-down of the entire plant. The sign of imminent failure is a rapid increase in the number of acoustic emission signals from a few concentrated sources. However, the value of the acoustic emission method lies in its pointing out potential trouble spots long before danger signs are generated. In this type of structural monitoring, the acoustic emission method can be applied without stopping normal operation (that is, without emptying the content of a tank or pipeline). Acoustic emission testing is applied to almost all tanks and vessels when new, and periodically during their service life.

Other applications include detecting leakage of fluid and gas from pressurized tanks, pipes, and valves; monitoring key structural elements of bridges, dams, and tunnels; detecting damage in bearings and other rotating components, such as motors and compressors; detecting failure of cutting tools in automatic machining operation; and finding fissures and cracks in geothermal wells. In most

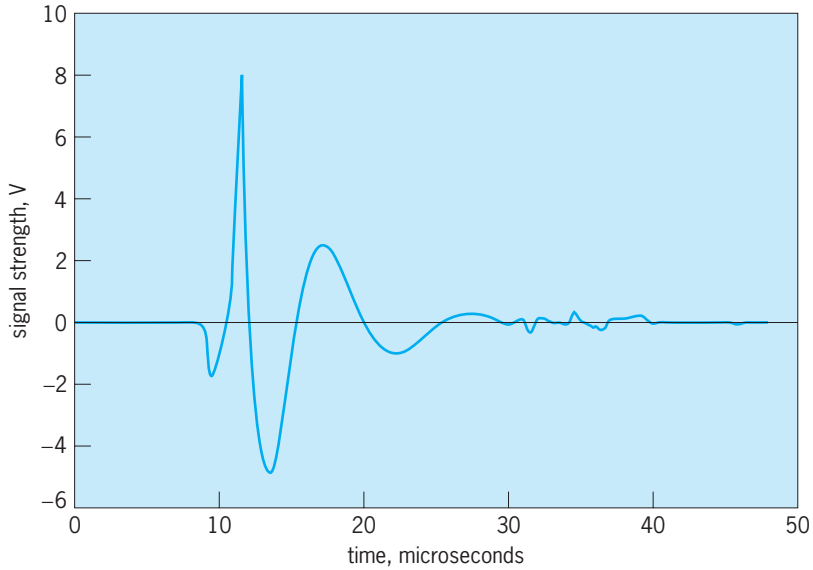

Fig. 1. Waveform of rapidly varying acoustic emission signal from a brittle material.

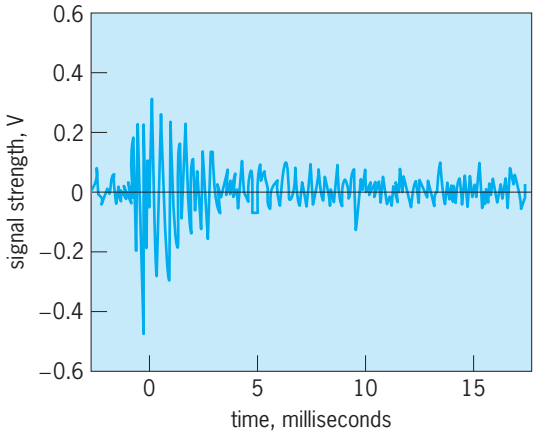

Fig. 2. Waveform of slowly varying acoustic emission signal from a plastic.

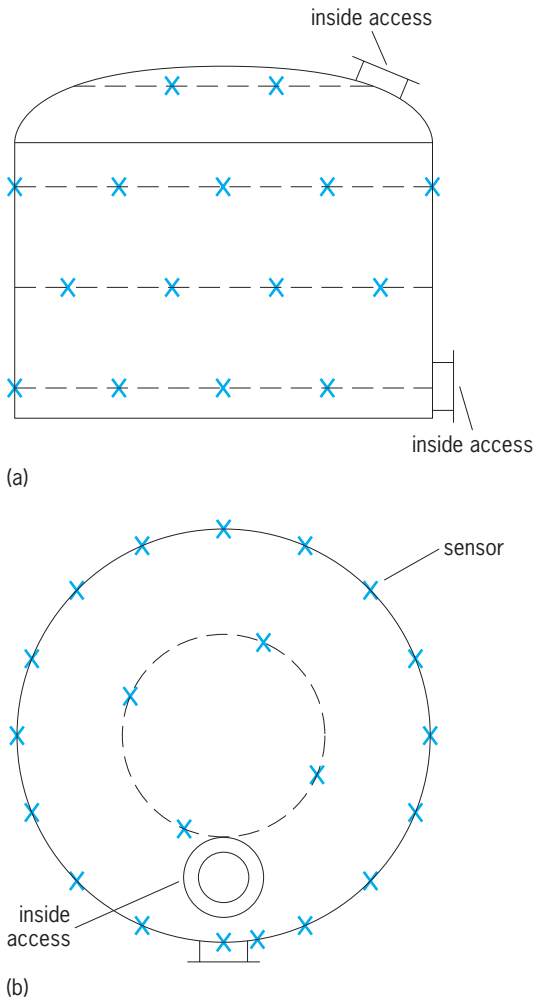

Fig. 3. Sensor arrangement for a chemical tank. (*a*) Side view. (*b*) Top view.

instances, the phenomena of acoustic emission generation are complex, and the evaluation of the results is largely learned from experience.

Acoustic emission methods are also valuable in the laboratory. Acoustic emission signals are emitted when materials are permanently deformed (stretched, pressed, bent, or twisted). Materials can exhibit several characteristic acoustic emission behaviors. One type of material emits many acoustic emission signals when the deformation begins, followed by a decreasing trend as more deformation occurs. Another type is characterized by the near absence of acoustic emission. This behavior is found when the material has been previously deformed heavily. These changes reflect the variation in the internal composition of materials, and acoustic emission provides key information in understanding deformation characteristics.

Acoustic emission signals are also produced when materials are stressed to a level nearing fracture. Material fracture results from the accumulation of microscopic internal damage. Such damage develops either gradually or rapidly. While acoustic emission cannot detect slow damage formation (such as forming voids), rapid damage processes that form cracks are serious defects and can be easily detected. One goal of materials engineering is to develop a strong and tough alloy with excellent fracture resistance, and acoustic emission provides valuable insights regarding the basic mechanical damage processes that are relevant to this endeavor.

An example of damage accumulation is shown in **Fig. 5**. Here, stress is applied to a high-strength steel specimen. As the applied stress is increased beyond a limit, the number of acoustic emission signals starts to rise and continues to rise quickly. Each such signal is produced in this steel by the fracture of a very hard and brittle carbide particle embedded in the steel matrix. Near the final fracture point, some acoustic emission signals result from microscopic cracks. In this study, carbide particles are shown to be the main source of damage. Consequently, the strength of the steel is improved by 20% or more by reducing the size of carbide particles. *See* STEEL.

In addition to the observation of deformation and fracture, acoustic emission is utilized in various areas of materials studies. These include abrasion, casting,

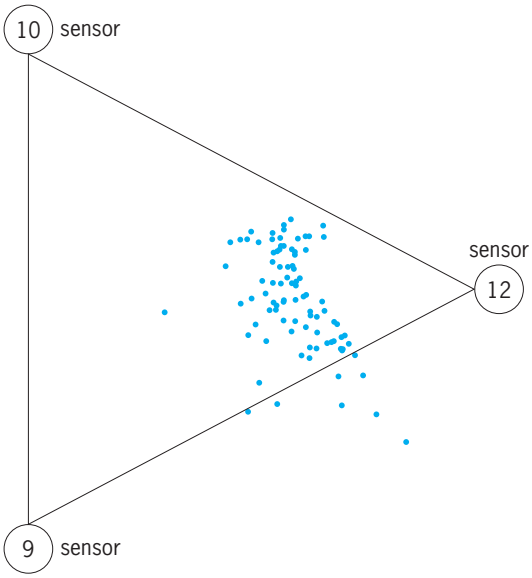

Fig. 4. Typical plot of acoustic emission source positions, indicating a concentration of sources.

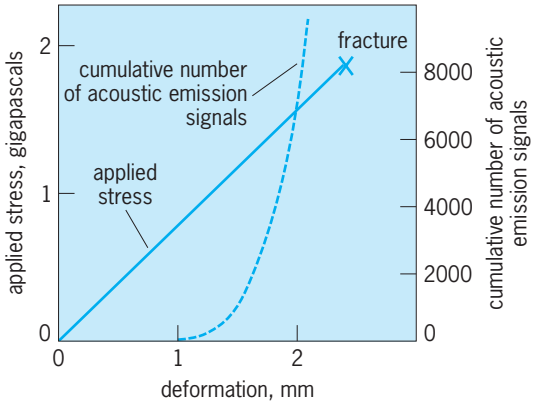

Fig. 5. Behavior of a high-strength tool steel specimen as stress is applied.

film coating and plating, corrosion, curing of plastics, friction, magnetization, oxidation, and hardening of steels. *See* NONDESTRUCTIVE EVALUATION.

Kanji Ono

Bibliography. Acoustic emission, in *Metals Handbook*, 9th ed., vol. 17: *Nondestructive Testing and Quality Control*, pp. 278–294, ASM International, Materials Park, OH, 1989; R. K. Miller and P. McIntire (eds.), Acoustic emission testing, *Nondestructive Testing Handbook*, 2d ed., vol. 5, American Society for Nondestructive Testing, Columbus, OH, 1987; K. Ono (ed.), *J. Acous. Emission*, 1982– ; K. Ono (comp.), *Progress in Acoustic Emission IX*, 1998 [Proceedings of the 14th International Acoustic Emission Symposium: Transitions in AE for the 21st century]; I. G. Scott, *Basic Acoustic Emission*, Gordon and Breach, New York, 1991.

Acoustic impedance

The ratio of the complex amplitude of spatially averaged acoustic pressure over a surface to the complex amplitude of volume velocity on that surface. The unit is the newton-second/meter^5, or the mks acoustic ohm. In the cgs system the unit is the dyne-second/centimeter^5. The volume velocity is the integral over the surface of the particle velocity normal to the surface. Acoustic impedance measures the degree to which an acoustic medium impedes the fluid motion created by an imposed pressure. It was originally defined by analogy to electrical systems. *See* ELECTRICAL IMPEDANCE; SOUND PRESSURE; UNITS OF MEASUREMENT.

In cases where the acoustic pressure p and volume velocity Q vary sinusoidally with time t, it is convenient to express them (as with quantities in alternating-current circuits) as the real parts of complex numbers that are rotating in the complex plane. This is done in Eqs. (1), where $\tilde{p}$ and $\tilde{Q}$ are complex

$$\begin{aligned} p(t) &= \mathrm{Re}\,\{\tilde{p}\exp(i\omega t)\} \\ Q(t) &= \mathrm{Re}\,\{\tilde{Q}\exp(i\omega t)\} \end{aligned} \tag{1}$$

numbers, Re means "real part of," $i = \sqrt{-1}$, and ω is the angular frequency. The acoustic impedance is given by Eq. (2). It is generally complex-valued be-

$$Z = \frac{\tilde{p}}{\tilde{Q}} \tag{2}$$

cause the amplitudes $\tilde{p}$ and $\tilde{Q}$ are generally complex-valued. The phase of Z represents the phase difference between pressure and volume velocity. For a fixed pressure, a large volume velocity is associated with a small magnitude of acoustic impedance. *See* ALTERNATING-CURRENT CIRCUIT THEORY; COMPLEX NUMBERS AND COMPLEX VARIABLES.

Specific acoustic impedance z is defined as the acoustic impedance multiplied by the area, so given by Eq. (3), where $\tilde{V}$ is the complex amplitude of the

$$z = AZ = \frac{\tilde{p}}{\tilde{V}} \tag{3}$$

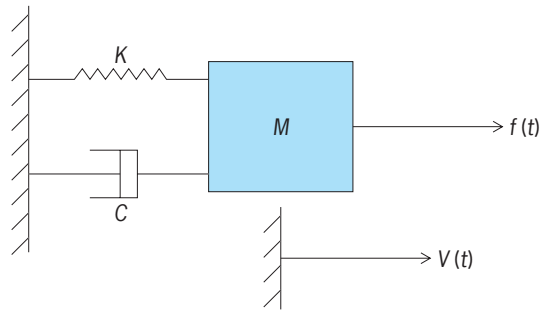

Mechanical system used to interpret acoustic impedance.

average velocity normal to the surface. The unit is the newton-second/meter^3, or the mks rayl. In the cgs system the unit is the dyne-second/centimeter^3, or the rayl. The characteristic acoustic impedance is the specific acoustic impedance for a plane wave propagating through an acoustic medium and is equal to ρc, where ρ is the mass density and c is the sound speed. The characteristic impedance of water is approximately 1.5×10^6 mks rayls, and the characteristic impedance of air at 0°C (32°F) is approximately 400 mks rayls.

Interpretation of real and imaginary parts. Acoustic impedance is often expressed as the sum of real and imaginary parts, as in Eq. (4), where R is the acous-

$$Z = R + iX \tag{4}$$

tic resistance and X is the acoustic reactance. These quantities may be interpreted by analogy to the mechanical system shown in the **illustration**, where the mechanical impedance is defined by Eq. (5). A

$$Z_{\mathrm{mech}} = \frac{\tilde{f}}{\tilde{V}} = C + i\left(\omega M - \frac{K}{\omega}\right) \tag{5}$$

comparison of Eqs. (4) and (5) indicates that the resistance is analogous to a mechanical dashpot (a damping element in which the applied force is proportional to velocity) and the reactance is analogous to either a mass or a spring, depending on its sign.

A more general interpretation of the resistance comes from Eq. (6) for the time-averaged power flow-

$$P_{\mathrm{ave}} = \tfrac{1}{2}R|\tilde{Q}|^2 \tag{6}$$

ing through the surface, which indicates that the power flow is proportional to resistance for a fixed volume velocity. The reactance is associated with an energy storage.

Circular aperture in planar baffle. One widely used example of acoustic impedance arises when the surface is chosen as a circular aperture in a planar rigid baffle of infinite extent with an acoustic fluid on one side. The velocity normal to the aperture is assumed constant over the aperture. The acoustic fluid has mass density ρ and sound speed c. When the radius of the circle is much less than an acoustic wavelength,

the resistance and reactance are given by Eqs. (7),

$$R_p = \rho c \frac{1}{2} \frac{(ka)^2}{\pi a^2} \qquad \text{for} \quad (ka)^4 \ll 1$$
$$X_p = \rho c \frac{8}{3\pi} \frac{ka}{\pi a^2} \qquad \text{for} \quad (ka)^2 \ll 1 \tag{7}$$

where $k = \omega/c$ is the acoustic wavenumber.

Acoustic reflection from a fluid interface. A common use of acoustic impedance arises in acoustic reflection problems. Reflected and transmitted waves are created when a plane wave is normally incident on a boundary between two fluids. If the incident wave propagates in a fluid denoted by 1 and the other fluid is denoted by 2, the complex-valued reflection coefficient is the ratio of reflected to incident wave amplitudes. It is related to the acoustic impedances of the fluids by Eq. (8).

$$R = \frac{Z_2 - Z_1}{Z_2 + Z_1} \tag{8}$$

This result confirms the intuitive expectation that a relatively high impedance for fluid 2 will create a reflection coefficient similar to that of a rigid boundary ($R = 1$). Conversely, a relatively low impedance for fluid 2 will create a reflection coefficient similar to that of a pressure-release boundary ($R = -1$). *See* SOUND.

J. Gregory McDaniel

Bibliography. L. L. Beranek, *Acoustics*, 1948, reprint, 1989; M. C. Junger and D. Feit, *Sound, Structures, and Their Interaction*, 1972, reprint 1993; H. F. Olsen, *Dynamical Analogies*, 1958; A. D. Pierce, *Acoustics: An Introduction to its Physical Principles and Applications*, 1981, reprint, 1989.

Acoustic interferometer

An instrument that is sensitive to the interference of two or more acoustic waves. It provides information on acoustic wavelengths that is useful in determining the velocity and absorption of sound in samples of gases, liquids, and materials, and it yields information on the nonlinear properties of solids.

Operating principles. In its simplest form, an acoustic interferometer for use in liquids has a fixed piezoelectric crystal (acting as a transmitter) tuned to the frequency of interest and a parallel reflector at a variable distance from it. Driven by an oscillating electrical voltage, the piezoelectric crystal generates a sound wave, which in turn is reflected by the reflector. The acoustic pressure amplitude on the front face of the crystal depends on the velocity amplitude at the face and the distance to the reflecting surface. The amplitude ratio (radiation impedance) of the acoustic pressure to the velocity and the relative phase shift between the two oscillating quantities depend solely on the distance to the reflecting surface. If the reflector acts as a rigid surface, this amplitude ratio is ideally zero whenever the net round-trip distance between the crystal and the reflector is an odd number of half-wavelengths because the reflected wave is then exactly out of phase with the incident wave at the crystal's location. The crystal then draws the maximum current since the oscillations are unimpeded. *See* ACOUSTIC IMPEDANCE; PIEZOELECTRICITY; WAVE MOTION; WAVELENGTH.

Wavelength and sound-speed measurement. During operation, the current drawn by the crystal is monitored as the reflector is gradually moved away from the crystal. Whenever the reflector position is such that the crystal is at a pressure antinode (place of maximum pressure in a standing wave), there is a strong dip in the current drawn due to the relatively high radiation impedance presented by the standing wave to the crystal face. Consecutive antinodes are a half-wavelength apart. For a given frequency f, a measured distance L between the location of any one antinode and that of its nth successor yields the wavelength $2L/n$ and the speed of sound $c = 2Lf/n$. An acoustic interferometer based on this principle can achieve a precision of 0.01%. Since the current drawn by the crystal is relatively insensitive to the frequency for a given radiation impedance, the sound speed can also be determined by keeping the distance between the crystal and the reflector fixed and gradually sweeping the frequency. With electronic feedback to lock the interferometer on a particular mode, signal measurement is enhanced.

Absorption measurement. The pressure nodes and antinodes correspond to the local maxima and minima, respectively, in the current drawn. The peak of the current amplitude decreases with the distance traversed by the reflector. If the separation distance is sufficiently large that the exponential decrease associated with absorption dominates any spreading losses, the absorption coefficient for the medium can be derived by measurement of the ratios of current amplitudes at two successive points where the current drawn is a local maximum. *See* SOUND; SOUND ABSORPTION; ULTRASONICS.

George S. K. Wong; Allan D. Pierce; Sameer I. Madanshetty

Bibliography. E. Carr Everbach and R. E. Apfel, An interferometric technique for B/A measurement, *J. Acous. Soc. Amer.*, 98(6):3428–3438, 1995; L. Cuscó and J. P. M. Trusler, Identification of environmentally acceptable low-sound speed liquids, *Int. J. Thermophys.*, 16(3):675–685, 1995; J. Tapson, Stochastic resonance in a mode locked acoustic interferometer, *Ultrasonics*, 36:415–419, 1998; M. S. Zhu et al., Sound velocity and ideal-gas specific heat gaseous 1,1,1,2-tetrafluoroethane (R134a), *Int. J. Thermophys.*, 4(5):1039–1050, 1993.

Acoustic levitation

The use of intense acoustic waves to hold a body that is immersed in a fluid medium against the force of gravity without obvious mechanical support.

In general, levitation involves the use of a fundamental (noncontact) force to balance gravity, and it is usually employed to determine some basic characteristic associated with the force in question. The near levitation of charged oil drops was used by

Waterdrop acoustically levitated in air between a vibrating surface (below the drop) and a reflecting surface (above the drop). The drop is located slightly below the pressure antinode of a 21-kHz standing wave. The equilibrium shape of the drop is oblate because of the balance between acoustic radiation pressure and capillary stresses.

R. A. Millikan to measure the quantum of electrical charge, and this work has been extended by using electromagnetically levitated, superconducting niobium spheres to search for the hypothetical fractionally charged quarks of elementary particle theory. The levitation of particles by light beams has also been demonstrated. *See* ELECTRON; QUARKS; RADIATION PRESSURE.

Levitation is not restricted to electromagnetic and optical means, but can also occur in the presence of fluid flow, including the back-and-forth fluid flow produced by the passage of an acoustic wave. Such acoustically generated forces are extremely small in common experience. But intense acoustic waves are nonlinear in their basic character and, therefore, may exert a net acoustic radiation pressure on an object sufficient to balance the gravitational force and thus levitate it (see **illustration**). Since sound is a mechanical wave, the force on the object will depend on the degree to which the mechanical properties (density and elasticity) of the levitated object differ from those same properties of the surrounding medium. *See* ACOUSTIC RADIATION PRESSURE.

The applications of acoustic levitation in air or other gas include an acoustic positioning module that has been carried in the space shuttle and used in fundamental studies of the oscillation and fission of spinning drops. An acoustic levitation furnace has been designed to study the possibility of containerless solidification of molten materials. This could result in materials of commercial interest, and lead to the bulk processing of materials in space. *See* SPACE PROCESSING.

Levitation in the normal gravitational field of the Earth can occur in the presence of sufficiently intense acoustic fields. Objects can also be levitated by ultrasound in a liquid. Not only is the required force smaller in a liquid than in air because of buoyancy, but also much stronger acoustic standing waves can be established in a liquid than in a gas on account of the liquid's cohesive strength.

Applications of the levitation of objects in liquids have included measurements of the ultimate tensile strengths of liquids, mechanical characterization of superheated and supercooled liquids, the measurement of properties of biological materials (including human red blood cells). The shape oscillations and interfacial properties of drops levitated in air and in liquids have been studied. The evaporation of charged arrays of drops levitated electroacoustically has been characterized.

Closely related to acoustic levitation is the trapping of cold and heavy gases near the pressure nodes of stationary ultrasonic fields. Aerosol particles in air can also be localized and concentrated by stationary ultrasonic fields. Ultrasonic fields in liquids are also used to trap single isolated bubbles in such a way that a light pulse is emitted during each acoustic oscillation. *See* SONOLUMINESCENCE; SOUND; ULTRASONICS.

Robert E. Apfel; Philip L. Marston

Acoustic microscope

An instrument that utilizes focused acoustic waves to produce images of surface and subsurface features in materials, and to measure elastic properties on a microscopic scale. It has been used to image and measure local elastic properties in metals, ceramics, semiconductor integrated circuits, polymeric materials, and biological materials including individual cells.

Scanning acoustic microscope. The most important component of a scanning acoustic microscope is the acoustic lens (**Fig. 1**). It is usually made from a cylinder of single-crystal sapphire. At one end of the cylinder a spherical cavity is ground and polished, while to the other end a high-frequency (100–2000 MHz) piezoelectric transducer is attached. The sample is placed below the spherical cavity near the focal point with a drop of water between the lens and sample. An electrical signal consisting of a narrow-width tone burst at a very high radio frequency is applied to the transducer. The transducer converts the electrical signal into plane-wave acoustic pulses, which propagate through the cylindrical rod and impinge

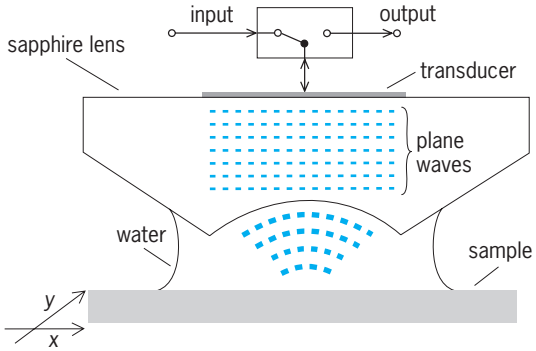

Fig. 1. Acoustic lens.

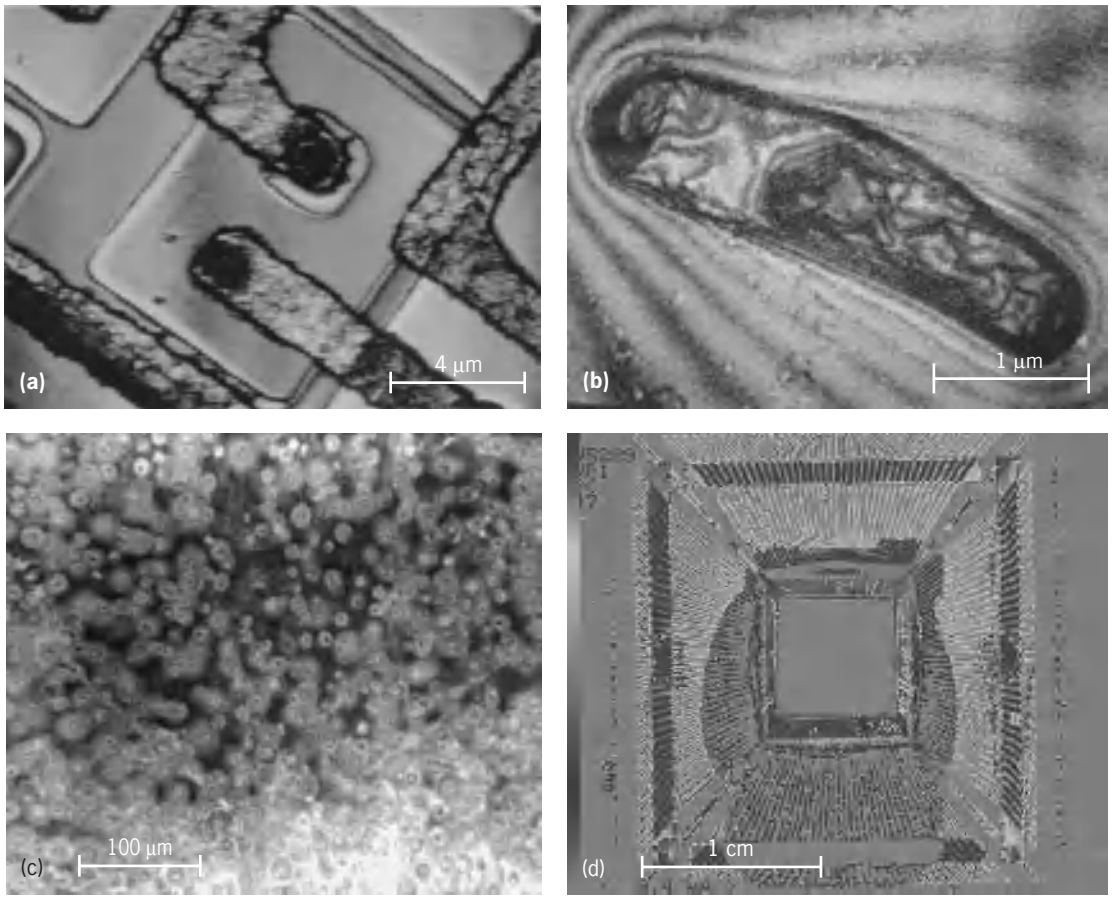

Fig. 2. Acoustic images. (*a*) Bipolar transistor on silicon integrated circuit and (*b*) myxobacterium, both images taken in superfluid helium at 0.2 K (*from J. S. Foster and D. Rugar, Low-temperature acoustic microscopy, IEEE Trans., SU-32, 139–151, 1985*). (*c*) Delaminations at a polymer-metal interface (*courtesy of S. Sathish*). (*d*) Delamination in an electronic ceramic package (*courtesy of G. Pfannschmidt*).

on the spherical cavity of the lens. At the interface between lens and water, the acoustic waves converge and focus to a diffraction-limited spot on the surface of the sample. A part of the incident acoustic energy is reflected by the sample, while the rest propagates into the sample. The reflected signal propagates back into the water and reaches the transducer after propagating through the lens. The transducer converts the acoustic signal into an electrical signal, which is electronically amplified and digitized, and the amplitude is stored in a computer. The lens is mechanically raster-scanned across the surface of the sample, and the reflected amplitude at each location is acquired and displayed on the monitor of a computer as an acoustic image. Since the acoustic image is built with the help of reflected amplitude of the signal, the microscope is called a reflection scanning acoustic microscope. *See* PIEZOELECTRICITY; TRANSDUCER.

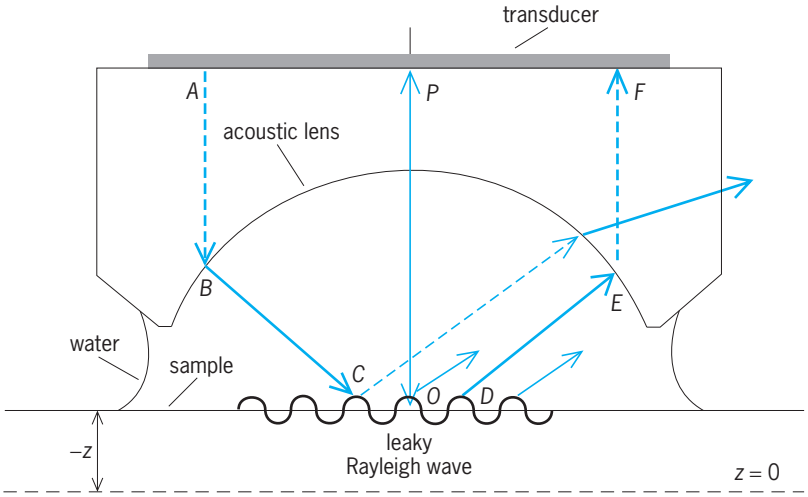

Fig. 3. Acoustic rays interacting to produce contrast enhancement in imaging at defocus and to generate *V(z)* curve.

Resolution and contrast. The diffraction-limited acoustic spot size and hence the resolution of an acoustic microscope increases with the increasing frequency of the acoustic waves. A resolution of 0.5 micrometer can be routinely achieved at a frequency of 2 GHz with water as coupling fluid. This is comparable to the resolution of an optical microscope. By employing liquid helium as the coupling fluid, a resolution of 20 nanometers has been achieved. The contrast in acoustic images arises from the local variation in the elastic properties of the materials and is related to the local acoustic refractive index, which is defined as the ratio of the velocity of sound in the material to the velocity of sound in the coupling fluid. *See* ELASTICITY; LOW-TEMPERATURE ACOUSTICS.

Defocused operation. Acoustic images can be obtained not only with lens focused on the surface of the sample but also with the lens defocused so that the focal point is inside the material (a mode of operation known as imaging at a negative defocus). Acoustic waves focused on the surface produce images of the surface features (**Fig. 2***a*, *b*). However, when the lens is defocused, acoustic waves penetrating into the sample are brought to focus in the interior, and this produces images of the interior features of materials. This feature of an acoustic microscope has been used to image materials that are opaque to light. It has found applications in the analysis of flaws and defects in several industrial materials, including electronic components (Fig. 2*c*, *d*). *See* NONDESTRUCTIVE EVALUATION.

When a large-aperture acoustic lens is used for imaging at a negative defocus, the contrast and the resolution do not degrade as observed in other types of microscopes. Instead, spectacular enhancement in the contrast occurs. The enhancement of the contrast is due to the ability of solids to support surface acoustic waves, and the capability of an acoustic lens to transmit and receive them. The origin of the contrast enhancement can be explained by the interaction of waves inside an acoustic lens (**Fig. 3**). Using a ray picture of acoustic waves, in an acoustic lens, as the angle of incidence increases, a critical angle is reached when total internal reflection occurs. At the critical angle, Rayleigh surface acoustic waves are generated on the surface of the sample, and propagate along the interface between water and the sample. The surface acoustic waves, while propagating at the interface, reradiate energy back into the water, adding an extra contribution to the signal reaching the transducer. These surface acoustic waves are very sensitive to surface and near-surface defects and to the anisotropic elastic properties of materials. This mode of operation of the acoustic microscope is helpful for observing microstructure in polished and unetched metallic materials, and for the detection of surface microcracks. Acoustic images obtained at several different defocus distances show reversal of contrast in many grains.

Quantitative acoustic microscopy. Recording the variation of the received signal amplitude as the lens-to-sample distance is reduced produces a $V(z)$ curve, which can be used to calculate the elastic properties in a microscopic region of the sample. The most important feature of the curve is a periodic oscillation in the amplitude. The origin of this oscillatory behavior is explained using the acoustic rays propagating between the transducer, the acoustic lens, the water, and the sample (Fig. 3). The directly reflected ray (*PO*) and the Rayleigh ray (*AB-BC-CD-DE-EF*) interfere at the face of the transducer and produce a periodic variation in the amplitude. The period of the oscillations can be used to calculate the Rayleigh wave velocity, which, in turn, provides information about elastic properties.

Measuring elastic anisotropy. In an acoustic microscope operating with a spherical lens, the signal returning to the transducer is an average over the diffraction-limited spot. Hence, it is not sensitive to the elastic anisotropy of materials. Acoustic microscopes sensitive to elastic anisotropy have been developed using a cylindrical lens that focuses the acoustic beam along a line. With a cylindrical lens, the Rayleigh surface acoustic waves propagate perpendicular to the cylindrical axis. A measurement of the $V(z)$ curve with such a lens will provide the Rayleigh wave velocity along a line. Such lenses have been used to investigate the elastic anisotropy of materials.

Ultrasonic force microscope. The development of scanning probe microscopy has led to a new type of acoustic microscope capable of imaging elastic properties at nanometer resolution. In the ultrasonic force microscope, an acoustic wave is generated by a piezoelectric transducer attached to one face of the sample, and the tip of an atomic force microscope is used to measure the amplitude of this wave when it has propagated through the sample to the other side. By scanning the tip in contact with the sample, an acoustic image with elastic stiffness information can be produced. The resolution of the microscope depends on the diameter of the tip rather than the acoustic wavelength in the material. *See* MICROSCOPE; SCANNING ELECTRON MICROSCOPE.

Shamachary Sathish

Bibliography. A. Briggs, *Acoustic Microscopy*, Clarendon Press, Oxford, 1992; A. Briggs, *Advances in Acoustic Microscopy*, vol. 1, Plenum Press, New York, 1995; A. Briggs and W. Arnold, *Advances in Acoustic Microscopy*, vol. 2, Plenum Press, New York, 1996; B. T. Khuri-Yakub and C. F. Quate (eds.), *Selected Papers in Scanning Acoustic Microscopy*, SPIE, Bellingham, WA, 1992.

Acoustic mine

A mine that either passively listens to a target's sound noises, or periodically interrogates its environment by actively emitting acoustic pulses that may return echoes if prospective targets come within range. A mine is an underwater weapon consisting of a shell case which contains high explosives. Mines can be planted by airplanes, surface ships, or submarines. Submarines are used for deployment when secrecy is important. There are three types of mines: drifting, moored, and bottom. Drifting mines float freely at the sea surface. Moored mines are positively buoyant and are held at a given depth by a cable anchored to the bottom. Bottom mines rest on the sea floor. A mine can be activated by various means, such as by actually contacting a target, by sensing a target's magnetic field, by listening to the acoustic noises that emanate from a target, by sensing the excess pressure field that may be induced on the mine's sensor by a target passing above it, by the reception of acoustic echoes that a target may return to the mine's sonar after it has sent out its interrogating pings, or by a combination of several of these. *See* SONAR.

Classification mechanisms. Some mines have sophisticated classification mechanisms that allow them to distinguish between various types of nearby

objects. Only when a true target (as opposed to a false alarm) is identified is the mine actually triggered. Simpler mines detonate when they sense any large object nearby, which is always assumed to be enemy. More advanced mines consist of cases with a torpedo inside, and they rest near the sea bottom. The acoustic sensor in the outer capsule constantly explores its environment with its sonar. When a true target is detected and identified, the inner torpedo is fired. It then tracks the prospective target with its own sensor and guidance system until it hits the target. *See* ACOUSTIC TORPEDO.

Any mine can be set so that it will allow the passage of one or more ships and then explode when the desired next one is detected as it comes within range. The acoustic classification phase of any mine mechanism requires considerably more advanced signal processing techniques, and more favorable environmental conditions, than the initial detection phase; hence, it is less reliable.

Countermeasures. Closely related to acoustic mines is the area of acoustic mine countermeasures. Minesweeping vessels also try to find and neutralize mines by means of sonar. Bottom mines are the hardest to find. Often they become buried beneath the bottom sediment, which dampens their possible vibrations, making them nearly impossible to detect. The acoustic detection, classification, and neutralization of mines lying close to—not necessarily beneath—the sea bottom is the most challenging problem in mine countermeasures, and the research required to adequately address these problems is in its infancy.

Guillermo C. Gaunaurd

Bibliography. G. K. Hartmann, *Weapons That Wait*, U.S. Naval Institute, 1979.

Acoustic noise

Unwanted sound. Noise control is the process of obtaining an acceptable noise environment for people in different situations. These definitions and the words "unwanted" and "acceptable" suggest that criteria need to be established to determine when noise from different sources is unwanted and that these criteria could or should be used to decide on acceptable noise limits. Understanding noise and its control, then, requires a knowledge of the major sources of noise, sound propagation, human response to noise, and the physics of methods of controlling noise. The continuing increase in noise levels from many different human activities in industrialized societies led to the term noise pollution. Different governments have passed legislation and created regulations to control noise.

Noise as an unwanted by-product of an industrialized society affects not only the operators of machines and vehicles, but also other occupants of buildings in which machines are installed, passengers of vehicles, and most importantly the communities in which machines, factories, and vehicles are operated.

Propagation of Sound

Sound is a three-dimensional wave motion in the atmosphere. The acoustic waves travel from the source with a speed independent of their amplitude (unless the latter is very large). This speed is then dependent only on the acoustic medium and is proportional to the square root of the absolute temperature for any given medium. For air at 68°F (20°C), the speed of sound c is 1117 ft/s (343 m/s). *See* ACOUSTICS; WAVE MOTION.

Pure tones. The sources of sound may consist either of vibrating solid bodies (such as loudspeakers or vibrating metal panels on machines) or of random motion of air particles (such as when an air jet mixes with the atmosphere). If a loudspeaker cone is made to vibrate with simple harmonic motion at a given frequency f (in hertz or cycles per second), it gives rise to a sinusoidal disturbance in the atmosphere. This sound is known as a pure tone, of frequency f. At any instant there will be a sinusoidal variation of the atmospheric pressure with distance away from the source. The sound pressure p is defined as the difference in the pressure from the undisturbed pressure. For a pure tone, wave crests or wave troughs (maximum and minimum values of pressure) are separated by a distance called the wavelength λ. The relationship between the speed of sound c, wavelength λ, and frequency f is given by Eq. (1). Of course, at

$$\lambda = c/f \tag{1}$$

some point in space the sound pressure p also varies sinusoidally with time. The period T between pressure maxima is related to the frequency by Eq. (2) for a pure tone (**Fig. 1**).

$$T = 1/f \tag{2}$$

Many sound sources, such as a singing voice or a musical instrument, contain strong pure-tone components. These sounds contain several simultaneous pure tones with a fundamental (or lowest-frequency tone) which is usually the strongest. Examples of noise sources which contain pure tones include fans, engine exhausts, pumps, compressors, gears, bearings, and electric motors.

Random noise. If a sound source is made to vibrate with a random motion, it also creates sound waves. However, the sound pressure–time history at some point in space then resembles **Fig. 2** rather than

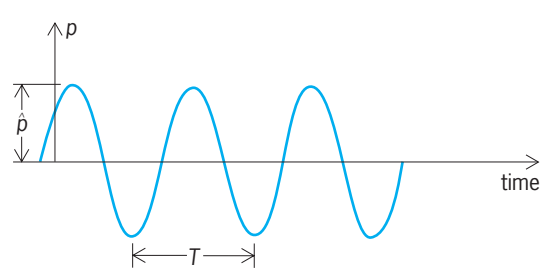

Fig. 1. Sound pressure p at one point in space, varying sinusoidally with time, in response to a source vibrating with simple harmonic motion.

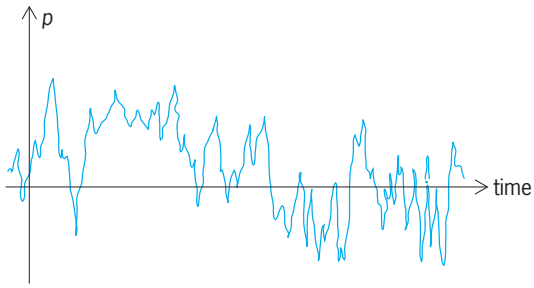

Fig. 2. Sound pressure *b* at one point in space, varying at random with time (random noise), in response to a source vibrating with random motion.

Fig. 1. This sound is called random noise, because it contains all frequencies instead of just one. Many noise sources, for example, jet engine exhausts, fans, flow noise in pipes, waterfalls, and wind, contain mostlyrandom noise.

Although the strength of a pure tone can be described by its amplitude or peak pressure $\hat{p}$ (Fig. 1), a random noise cannot, because the amplitude is never constant. Thus, noise signals are normally described by their time-averaged (effective or root-mean-square) values, P_{rms}, given by Eq. (3), where T

$$p_{\text{rms}} = (1/T)\int_0^T p^2 dt \tag{3}$$

is the averaging time, which is allowed to approach infinity. Only in the case of pure tone does $p_{\text{rms}} = \hat{p}/\sqrt{2}$.

Sound pressure level. The value of effective sound pressure p_{rms} increases about 10^6 times from a pin drop to a thunderclap. Because of these large variations in sound magnitudes, and because the human hearing sensation seems to vary in a logarithmic way, logarithms are used in measurement of sound. The sound pressure level L_p of a sound of effective pressure p_{rms} is given by Eq. (4), where the units of L_p

$$L_p = 10\log(p_{\text{rms}}^2/p_{\text{ref}}^2) \tag{4}$$

are decibels (dB). The reference pressure is internationally accepted to be $p_{\text{ref}} = 2 \times 10^{-4}$ dyne/cm.

Sound power level and intensity. If a source is constantly pulsating, it will radiate acoustic energy at a rate that is called the acoustic power and is expressed as W, in watts (joules/s). Again, because acoustic powers vary from very small to very large values, logarithms are used. The sound power level L_w (in decibels) of such a source is given by Eq. (5), where

$$L_w = 10\log(W/W_{\text{ref}}) \tag{5}$$

the reference sound power is $W_{\text{ref}} = 10^{-12}$ W. In an ideal source, if it is assumed that there is no energy dissipation as the sound energy radiates away from the source, then the intensity I (power/unit area) is

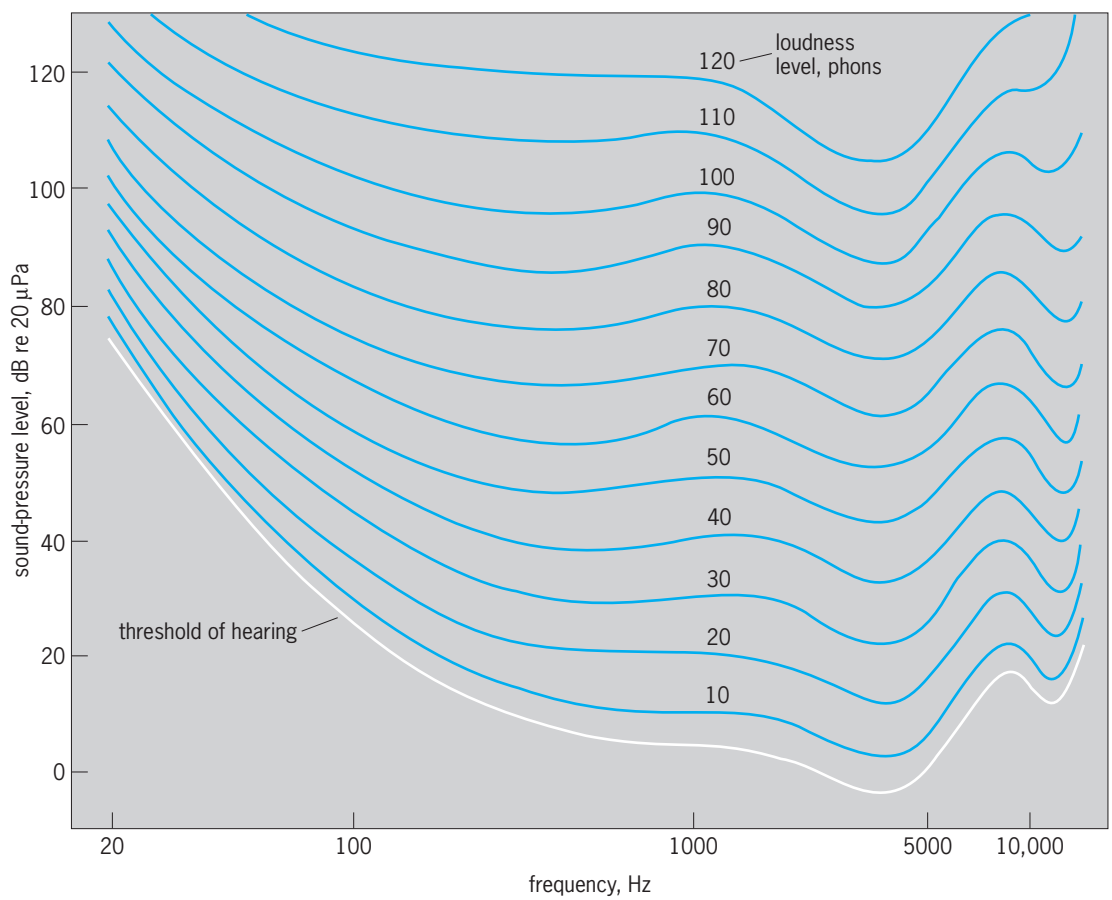

Fig. 3. Equal-loudness contours for pure tones. (*After D. W. Robinson and R. S. Dadson, A redetermination of the equal loudness relations for pure tones, Brit. J. Appl. Phys., 7:166, 1956*)

reduced by a factor of $(1/r^2)$ or $^1/_4$ every time the distance r from the source is doubled. This is the inverse-square law, Eq. (6). Because it is easily shown

$$I \propto 1/r^2 \tag{6}$$

that the intensity I is proportional to p^2_{rms}, doubling the distance from the source reduces the sound pressure level L_p [Eq. (4)] by $10 \log (2)^2 = 20(0.30) \approx 6$ dB, provided there are no reflections. One of the simplest noise control measures thus becomes evident; noise sources should be placed as far away from receivers as possible.

Frequency analysis. In order to determine the frequency distribution of a noise, the intensity or sound pressure level in different frequency bands is measured. The frequency analysis may be done by using electrical filters or fast Fourier transform (FFT) analyzers. The bands most commonly used are constant percentage bands of width one octave or one-third of one octave. (An octave is a doubling in frequency.) Constant-bandwidth bands such as 5, 20, or 50 Hz are also used, to obtain information on pure tones contained in a noise signal. The constant-percentage filter rapidly increases in bandwidth as the center frequency is raised. The range of hearing for most people is from about 20 to 16,000 Hz, and can be covered with 10 octave band filters with center frequencies: 31.5, 63, 125, 250, 500, 1000, 2000, 4000, 8000, and 16,000 Hz. The one-octave band has a bandwidth of 73% of the center frequency, while the one-third-octave band has a bandwidth of 23% of the center frequency. *See* HARMONIC ANALYZER; NOISE MEASUREMENT.

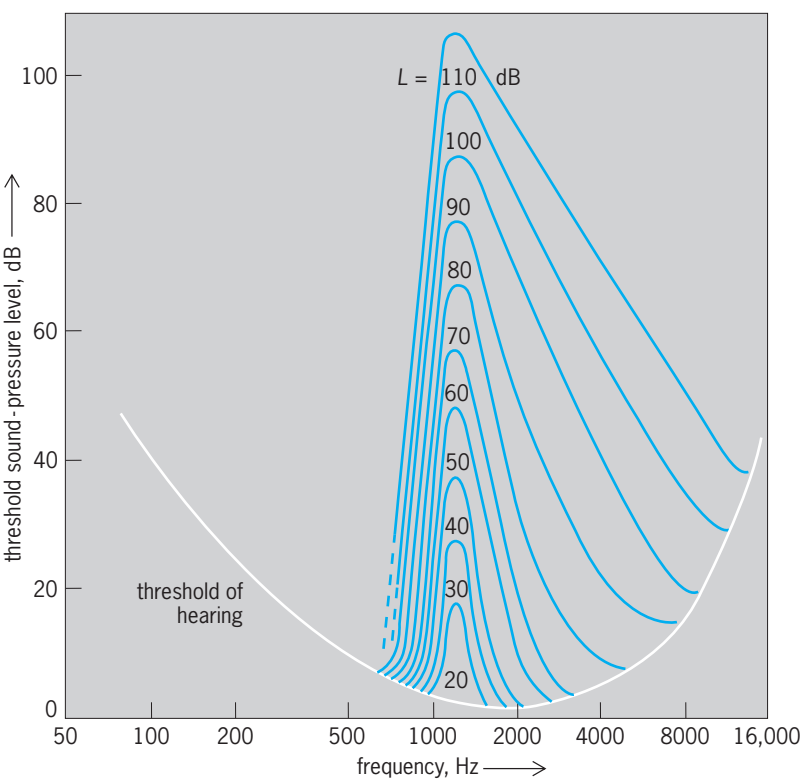

Fig. 4. Masking effect of narrow-band noise with a center frequency of 1200 Hz. Parameter *L* is root-mean-square value of sound pressure level of noise band. Curves show sound pressure level at threshold of hearing in presence of noise band. (*Data from E. Zwicker*)

Human Response to Noise

Noise interferes with some human activities and if sufficiently intense can permanently damage the ear.

Hearing mechanism. The ear is a complicated transducer which converts acoustical energy into mechanical vibrations of the eardrum, the three auditory ossicles (bones), and the cochlear membrane, and eventually into electrical energy which triggers nerve impulses in the auditory nerve. The outer ear canal and the eustachian tube are filled with air, while the inner ear (the cochlea) is filled with fluid. *See* EAR (VERTEBRATE); HEARING (HUMAN); PHYSIOLOGICAL ACOUSTICS.

Loudness. Individual responses to pure tones of different frequencies vary a little. Equal-loudness level contours (**Fig. 3**) represent pure tones which appear equally loud to most people at various frequencies. The lowest curve in Fig. 3 represents the threshold of hearing or the softest sounds that can be heard. The top curve approximately represents the threshold of "in" or "feeling." The equal-loudness level curves are given units of phons P, and are numerically equal to their sound pressure level value at 1000 Hz. Sounds must be increased about 10 phons in level before they double in loudness, and because of this nonlinear subjective response of the ear, a linear loudness scale with units of sones S, defined by Eq. (7), is used. Thus a sound level of 50 phons has

$$S = 2^{(P-40)/10} \tag{7}$$

a loudness of 2 sones; a sound of 60 phons, 4 sones; and so on.

Equal-loudness curves for bands of noise resemble those shown in **Fig. 4** for pure tones, but with some differences. It follows from Fig. 4 that the low-frequency sounds must be much more intense than mid-frequency and high-frequency sounds in order to seem equally loud. The greatest hearing sensitivity is in the range of about 1000 to 4000 Hz, which corresponds to the middle of the speech range.

Equal-loudness level contours have been used in acoustical instrumentation to produce useful single-number measures of the loudness or disturbing effect of noise. A-, B-, and C-weighting filters (**Fig. 5**), corresponding approximately to the inverse of the 40-, 70-, and 100-phon curves respectively, have been built into most sound level meters. The sound readings obtained using these filters are known as the A-weighted, B-weighted, or C-weighted sound levels. Although originally intended for low-level sounds, the A-weighted sound level is used for monitoring both low-level and intense sounds from almost all machine and vehicle noise sources. A-weighted levels are sometimes abbreviated as dB(A). *See* LOUDNESS; PSYCHOACOUSTICS.

Hearing damage. Immediate permanent hearing damage, normally to the eardrum or ossicles, can result from very intense sounds, above about 140

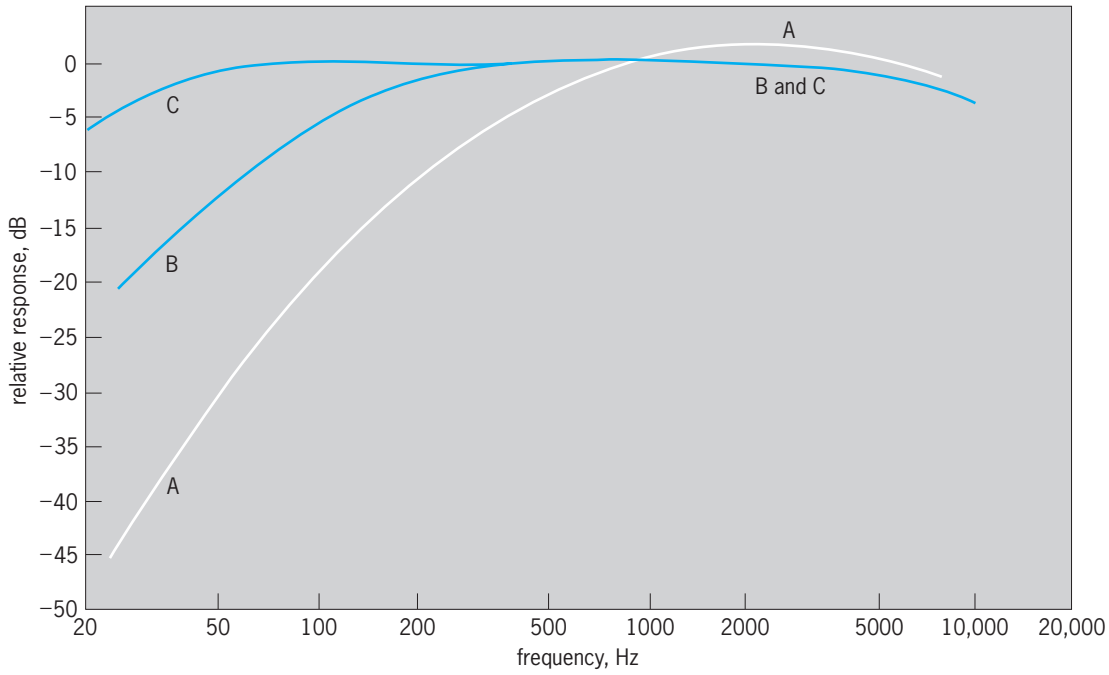

Fig. 5. A-, B-, and C-frequency weightings used with many sound-level meters.

or 150 dB, such as those associated with nearby gunfire or explosions. Lower-level continuous intense noises about 90 to 110 dB(A) can cause temporary hearing loss, from which a person recovers after a period of rest in a quiet environment. However, if such lower-level intense noise is experienced every day over a period of years, permanent hearing damage occurs. The amount of hearing loss depends upon frequency and sound pressure level of noise, bandwidth of noise, duration of exposure each day, and number of years of exposures. A general criterion for specifying tolerable exposures to noise is that the noise should not exceed levels or durations that will cause the average person (after 10 years of exposure) a measurable loss in understanding normal conversation. The damage risk criteria given in **Fig. 6** are based primarily on industrial surveys, and many people can be exposed to greater intensities and durations than these criteria, particularly if the total duration is less than 8 hours per day. *See* HEARING IMPAIRMENT.

Masking and speech interference. When two sounds contain the same frequency components, the more intense sound can normally be heard and it is difficult or impossible to detect the other. This phenomenon is known as masking. If the sounds are of different frequencies, the effect is more complicated and depends on the bandwidth and frequency separation of the two sounds. Figure 4 shows how the hearing threshold is changed for different levels of narrow band of noise centered at 1200 Hz. High-frequency sounds are masked easily by lower-frequency noise, while low-frequency sounds are not masked very easily by high-frequency noise.

Masking noise, sometimes called acoustic perfume, is supplied by loudspeakers in some buildings such as open-plan offices to try to mask unwanted sound such as conversation between other individuals at a distance. Noise in offices or workshops can become loud enough to mask wanted speech sounds, which normally have a sound pressure level of about 65 dB at the listener's ear. People do not become used to hearing speech above masking noises. One measure of the interference with speech of a steady broadband background noise is the preferred speech interference level (PSIL), the arithmetic average of the sound pressure levels of the background noise in the three octave bands centered at 500, 1000, and 2000 Hz (**Table 1**). *See* MASKING OF SOUND.

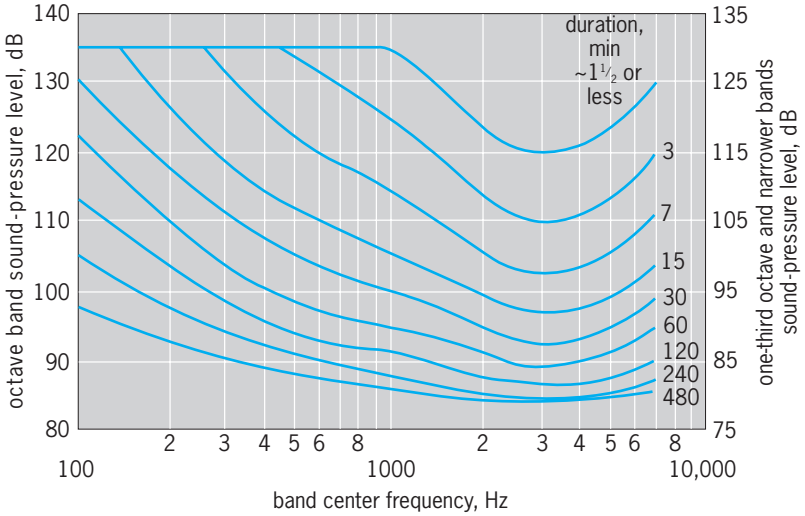

Fig. 6. Damage-risk contours for one exposure per day to one-octave (left-hand ordinate) and one-third-octave or narrower (right-hand ordinate) hands of noise. This graph can be applied to individual band levels which are present in broadband noise. (*After K. D. Kryter et al., Hazardous exposures to intermittent and steady state noise, J. Acous. Soc. Amer., 39:451–463, 1966*)

TABLE 1. Preferred speech-interference levels (500–2000 Hz) that barely permit reliable conversation

Distance between talker and listener, ft (m)	Speech-interference level, dB			
	Normal vocal effort	Raised vocal effort	Very loud vocal effort	Shouting
0.5(0.15)	72	77	83	89
1 (0.3)	66	71	77	83
2 (0.6)	60	65	71	77
3 (0.9)	56	61	67	73
4 (1.2)	54	59	65	71
5 (1.5)	52	57	63	69
6 (1.8)	50	55	61	67
12 (3.7)	44	49	55	61

Annoyance. The louder the noise, the more annoying it tends to be. With continuing exposure to a noise, adaptation occurs as long as the noise is accepted as a part of the environment. Because of adaptation and the difficulty of separating noise annoyance from the effects of other environmental factors, it has not been possible to determine an acceptable annoyance criterion for noise.

Sleep interference. Sufficiently intense noise will awaken a person; less intense noise usually arouses a person from deep sleep to more shallow sleep. However, as with annoyance, people tend to adapt to noise during sleep, making it difficult to specify a noise criterion for sleep interference.

Work performance. Studies to find the effects (if any) of noise on work productivity, efficiency, concentration, incidents of errors and accidents, and so forth, have been inconclusive. Some researchers state that after a period of adaptation noise has little or no effect on work performance, provided it is not sufficiently intense to interfere with speech communication, while others claim that noise can interfere particularly with those engaged in intellectual tasks. *See* INDUSTRIAL HEALTH AND SAFETY.

Community reaction. The effect of noise on whole communities rather than individuals or relatively small groups has also been studied (**Fig. 7**). Several physical measures have been used to characterize the noise environment, including effective perceived noise level (EPNL), composite noise rating (CNR), noise and number index (NNI), and noise exposure forecast (NEF). Most of these measures were originally created to rate aircraft noise while others, such as equivalent sound level L_{eq} and day-night level L_{dn}, were developed to characterize other sources such as traffic, factory, and construction noise.

Noise Reduction Methods

Most noise problems can be modeled as source–path–receiver systems. It is most desirable to reduce the strength or number of the sources. For example, the noise from the impact of two metal machine parts in a punch press might be reduced by replacing one or both of the metal contacts with softer material such as nylon or strong durable plastic. However, it is sometimes difficult to reduce the noise at a source without extensive redesign. For example, changes in some metal-cutting machines may affect the cutting process adversely. In such cases, it may be possible to reduce the source strength by substituting a quieter machine or using a different process.

When all possible ways of reducing the source strength have been tried, ways of reducing noise propagation along paths from the sources to the receivers should be investigated. As a last resort, the receiver, which in most cases is the human ear, can be protected by using earplugs or earmuffs or by enclosure in a booth. However, the ear devices can interfere with communication and be uncomfortable, and booths can be inconvenient and expensive. In cases of community noise problems caused by traffic or aircraft, receiver noise control usually becomes socially unacceptable, extremely expensive, or even virtually impossible. It is usually necessary to control path noise propagation.

Planning. It is often possible to use distance and source directivity with advantage in noise reduction. Equation (6) shows that outdoors the sound pressure level theoretically decreases by 6 dB for each doubling of distance. Indoor sound behavior is described by Eq. (8), sometimes known as the room

$$L_p = L_w + 10\log\left(\frac{Q}{4\pi r^2} + \frac{4}{R}\right) \tag{8}$$

equation. Here, L_w is the sound power level of the noise source [Eq. (5)], Q is the directivity index of the noise source, r is the distance from source to receiver in meters, and R is the room constant defined as $R = S\overline{\alpha}/(1 - \overline{\alpha})$, where $\overline{\alpha}$ is the average absorption coefficient of the room walls of area S, in m^2. The sound pressure level L_p (in decibels) theoretically decreases at 6 dB per doubling of distance near to the source (where the term $Q/4\pi r^2$ dominates) and at a lower rate in the reverberant field (where the $4/R$ term and reflections dominate). Outdoors, one may assume no reflections, so that $\overline{\alpha} = 1$, $4/R = 0$, and Eq. (8) reduces to Eq. (6). If a source is omnidirectional (radiates equal sound energy to all directions), $Q = 1$. However, many noise sources are directional, making it advantageous in outdoor situations to position the receiver at locations relative to the source where the directivity index Q is small (directions in which the noise source radiates little energy). The same procedure is effective indoors but only close to the source, since the $4/R$ term dominates Eq. (8) in the reverberant field. *See* ARCHITECTURAL ACOUSTICS; DIRECTIVITY; REVERBERATION.

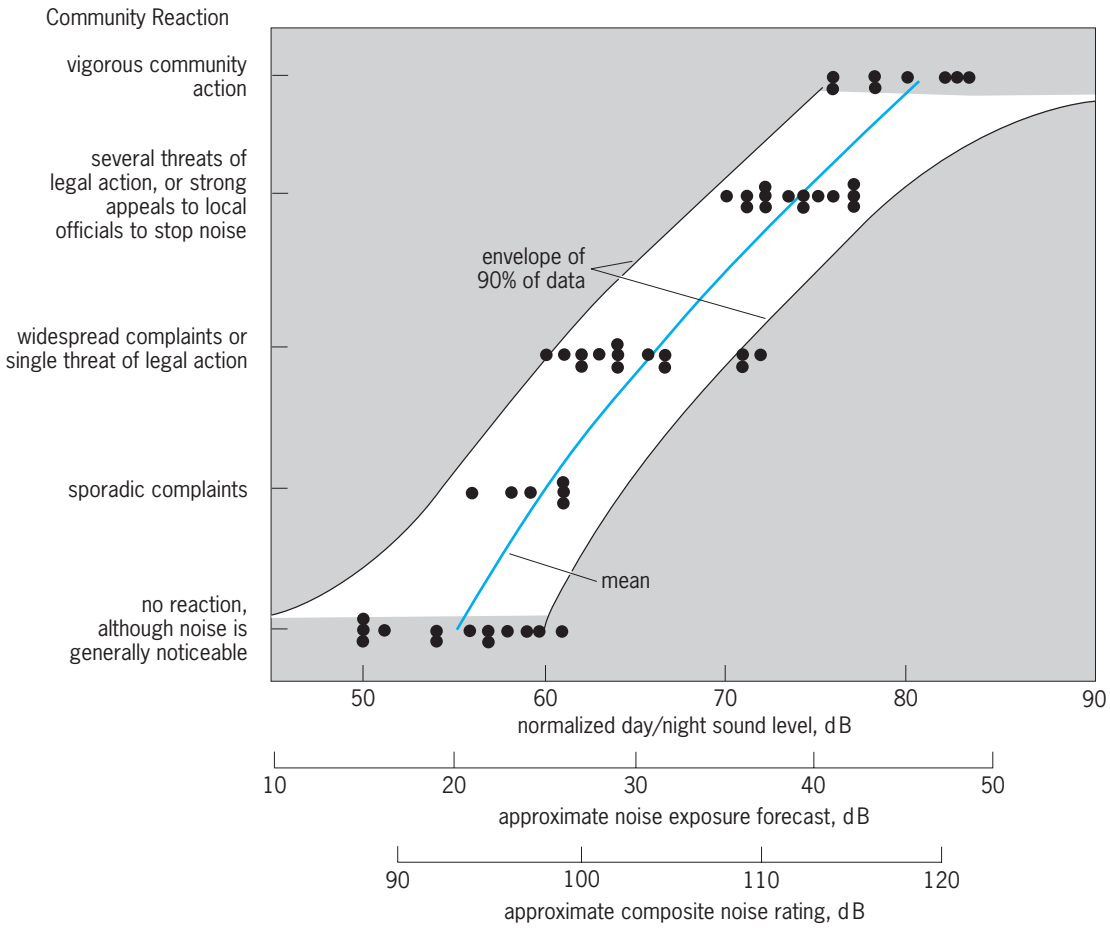

Fig. 7. Community reaction to many types of instrusive noise. Data normalized to residential urban noise, some prior exposure, windows partially open, no pure tones or impulses. (*After K. M. Eldred, Assessment of community noise, Noise Control Eng., 3(2):88–95, 1974*)

Absorbing materials. Although Eq. (8) is not accurate in rooms with low ceilings (such as factories), it can still be used for qualitative guidance with noise problems in irregularly shaped rooms. For example, in close proximity to a noise source, there is no reduction in sound pressure level L_p if the absorption coefficient of the materials on the walls is increased since the term $Q/4\pi r^2$ dominates; thus, it is not feasible to help the operator of a machine by using absorbing materials. However, far from a noise source in a room, a reduction of the reverberant noise level is achieved by increasing the absorption of the room walls. The absorption coefficient of a material α is defined as the fraction of incident acoustic intensity which is absorbed. Acoustic absorbing materials are usually made of porous materials such as fiberglass or open-celled foams. In environments where oil, water, or dirt can clog the pores of an absorbing material, a very thin impervious sheet of plastic may be placed over the absorbing material without substantially altering its sound-absorbing properties. The sheet (**Fig. 8**) increases the absorption at low frequencies but reduces it athigh frequencies.

Enclosures. Noise may be reduced by enclosure of the source. If it is essential to have continuous access to a noise source such as a machine used in mass production, or if cooling is necessary for a machine, a partial enclosure must be used. However, in the latter case noise leakage can be minimized by providing cooling vents, built from bent ducts, lined with absorbing material, and supplied with air from cooling fans, if necessary.

Transmission loss. The transmission less TL in dB of a partition is defined by Eq. (9), where τ is the fraction of incident acoustic intensity transmitted by the

$$\mathrm{TL} \approx 10 \log(1/\tau) \tag{9}$$

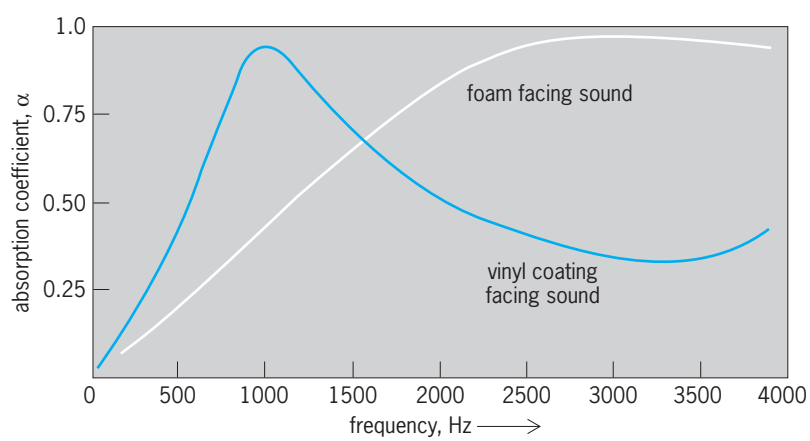

Fig. 8. Normal incidence absorption coefficient of a $^3/_4$-in.-thick (19-mm) porous foam material with vinyl coating facing sound and with foam facing sound.

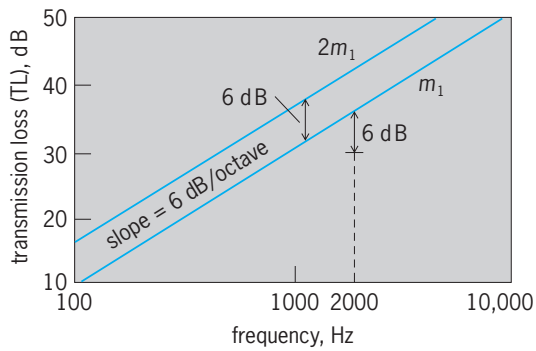

Fig. 9. Random incidence transmission loss for walls with mass per unit area m_1 and $2m_1$.

wall. For random incidence sound, the transmission loss in dB of an enclosure wall, a wall in a building, or even an airplane cabin wall is given empirically by Eq. (10), where m is the wall mass per unit area

$$\mathrm{TL} \approx 20 \log mf - 48 \tag{10}$$

in kg/m^2. **Figure 9** shows a plot of Eq. (10). Enclosure walls are almost "transparent" to low-frequency sound, but for every doubling of frequency the transmission loss increases by 6 dB. Transmission loss also increases by 6 dB at any given frequency if mass per unit area is doubled. These increases are not always observed in practice; an increase of about 4 to 5 dB per octave is often found. If massive walls are used in buildings to obtain a high transmission loss, care must be used to prevent leaks since these will seriously reduce the sound insulation of the wall. Two walls with an air gap usually have a greater transmission loss than one wall of the same total mass.

Other measures of performance. Sometimes measures other than transmission loss are used to rate the effectiveness of enclosures, walls, or even mufflers. The noise reduction (NR) of an enclosure is the sound pressure level L_{p1} measured inside minus the sound pressure level L_{p2} at some point outside the enclosure, Eq. (11). The insertion loss (IL) of the enclosure

$$\mathrm{NR} = L_{p1} - L_{p2} \tag{11}$$

is the difference in sound pressure levels at some point with $[L_p(w)]$ and without $[L_{p(wo)}]$ the presence of the enclosure, Eq. (12).

$$\mathrm{IL} = L_{p(wo)} - L_{p(w)} \tag{12}$$

Although Eqs. (10), (11), and (12) have been defined for enclosures, very similar definitions can be used for walls in buildings and mufflers. In general, transmission loss, noise reduction, and insertion loss are not equal. Sometimes transmission loss is averaged throughout the important audible frequency range 125–4000 Hz to give a single-number rating of a wall. A more complicated scheme in which the transmission loss–against–frequency curve is compared with a set of standard curves yields a single number rating called sound transmission class (STC).

Vibration isolation. If a machine is rigidly attached to the floor or supporting structure of a building or vehicle, the structure or floor can act in a manner similar to the sounding board of a musical instrument, and radiate large amounts of noise. This problem can be overcome by placing vibration isolators between machines and their supports. Such isolators may be metal or elastomeric springs.

Forces on machines at discrete frequencies can be caused by magnetic forces (usually at 120 Hz and integer multiples in the United States), and out-of-balance forces in rotating machines such as engines, motors, pumps, and fans. Vibration isolators should be designed so that the natural frequency of the machine on the isolators is very much less than any such forcing frequency. Such isolators will transmit much smaller forces to a support than those which act on the machine itself. However, isolators with this property generally must be very soft, and the machine may undergo static deflections when placed on them, resulting in misalignment of parts.

When a machine is started up, the forcing frequency increases and coincides briefly with the natural frequency before exceeding it. When the frequencies coincide, the machine will have a large vibration amplitude and transmit a large force to its support. To reduce this problem, during start-up, and during stopping, some damping is usually provided in isolators. Vibration and consequently noise can also be reduced by the application of viscous damping materials to the structures. *See* MECHANICAL VIBRATION; VIBRATION; VIBRATION DAMPING; VIBRATION ISOLATION.

Barriers. Barriers are widely used in industry and alongside roads and railways to shield receivers from noise sources. For a given geometry, the insertion loss of a barrier depends on frequency, and can be calculated approximately from **Fig. 10** by first determining by Fresnel number N, given by Eq. (13). Here λ = sound wavelength [Eq. (1)], $a + b$ = shortest path length of the sound wave over the barrier (Fig. 10), and d = straight-line distance between source S and receiver R. The positive sign is for the receiver in the shadow zone (receiver able to see source) and the negative sign for the receiver in the bright zone (receiver unable to see source). Figure 10 shows the theoretical insertion loss of a barrier as a function of N for a point source, under the assumption that the barrier is infinitely long and impervious to sound waves. If the sound source is composed of a line of sources (for example, closely spaced road vehicles), then the insertion loss is changed somewhat.

$$N = \pm(2/\lambda)(a + b - d) \tag{13}$$

Obviously barriers for any given geometry are more effective at high frequencies, when λ is small, and are ineffective at low frequencies, when λ is large, and thus N is small. For any particular frequency, the insertion loss of the barrier can be made larger by increasing $a + b - d$, or equivalently by placing the barrier as near the source or receiver R as possible. If barriers are used in low factory-type buildings, the barrier effectiveness can be reduced by sound reflections from the ceiling and walls near the barrier, but this can be mitigated by adding absorbing materials to such ceilings and walls. Barriers

will not reduce the noise on the receiver side, but will increase it, unless the barrier is also covered in absorbing material. If the source must be seen, a transparent barrier (made of Plexiglas) can be placed between a machine and operator.

Mufflers. Mufflers are used to reduce the sound from systems containing a noise source connected to a pipe or duct system such as air-conditioning systems, fans, and industrial blowers, gasoline and diesel engines, compressors, and jet engine inlets and exhausts. There are two main types of mufflers, reactive and dissipative. Reactive mufflers are usually composed of several chambers of different volumes and shapes connected together with pipes, and tend to reflect the sound energy back to the source. They are essentially sound filters. Dissipative mufflers are usually composed of ducts or chambers which are lined with acoustic absorbing materials that absorb the acoustic energy and turn it into heat. Some mufflers are a combination of reactive and dissipative types. The type of muffler selected for any particular application will depend upon the noise source to be silenced and several environmental factors.

Reactive mufflers. Reactive types (**Fig. 11**) are most useful when the noise source to be reduced contains pure tones at fixed frequencies or when there is a hot, dirty, high-speed gas flow. Reactive mufflers for such purposes can be made quite inexpensively and require little maintenance. Such mufflers lose their effectiveness when used with large-diameter ducts and at high frequencies (that is, if 0.8 of the sound wavelength is less than the greatest diameter or lateral dimension of the muffler) due to the formation of lateral waves or cross modes in the muffler.

Dissipative mufflers. Dissipative types are useful when the source produces noise in a broad frequency band. They are particularly effective at high frequencies, but special precautions must be taken if the gas stream has a high speed and temperature and if it contains particles or is corrosive. If the speed of the gas stream is above about 15 m/s, the absorbing material (fiberglass or rock wool) should have surface bonding to prevent damage. At high speeds (up to 300 ft/s or 100 m/s), facing materials such as wire screens or perforated metal sheets are needed to prevent erosion of the absorbing material. If the gas stream has a high temperature (up to 1000°F or 550°C), materials such as Corten with special paints or stainless steel may be used for the facing material. Contamination of the absorbing material with oil, water, and dirt may be prevented by using a very thin surface sheet of impervious plastic material such as Mylar. The thin sheet causes a slight loss of acoustic absorption at high frequency but an increase at low frequency. Both parallel baffle and blocked line-of-sight mufflers are used. The latter tend to be better particularly at high frequency, although construction is more complicated.

Measures of performance. As with enclosures, the insertion loss, transmission loss, and noise reduction of a muffler can be defined and measured. The first two quantities are usually more useful. In general, the three quantities are not necessarily equal, although, in special cases, insertion loss can equal transmission

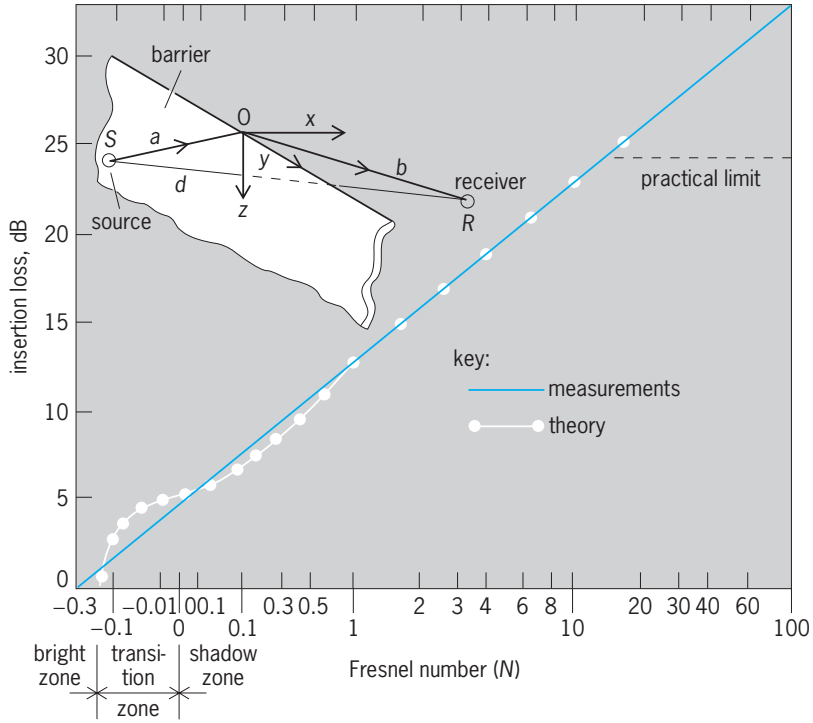

Fig. 10. Insertion loss of sound from a point source by a rigid barrier as a function of Fresnel number *N*. Negative *N* refers to the case where the receiver is able to see the source. (*After L. L. Beranek, Noise and Vibration Control, McGraw-Hill, 1971*)

loss. Insertion loss and transmission loss are sometimes loosely described interchangeably as attenuation. *See* MUFFLER.

Ear protectors. When all possible ways of reducing noise source strength and propagation along paths have been exhausted, the last resort is to protect the ear (the receiver). If a limited number of people are involved, this may sometimes be an attractive and economic approach (for example, to protect the sole operator of a large noisy machine). In many cases it is undesirable or impossible (for example, to protect thousands of residents living around airports).

There are three main types of ear protectors: earplugs that fit into the ear canal; earmuffs that fit over the external ear; and rigid helmets under which earmuffs or earplugs or both may be worn. There is some advantage in using more than one of these types simultaneously, although usually the attenuation is not strictly additive. Fortunately, both continuous unwanted sound and wanted sound are reduced by the same amount so that the ability to

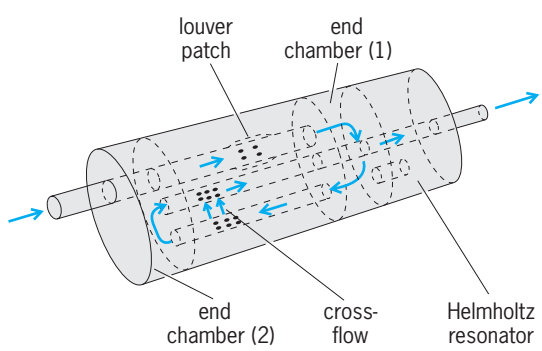

Fig. 11. Typical reverse flow automobile muffler.

hear warning signals and speech is not adversely affected. This is because the signal-to-noise ratio is not changed.

In cases where only one or two people are involved, as in monitoring instrumentation near a large noisy machine, it may be desirable to build a special acoustic enclosure around the operators. The booth should be built as far as possible from the machine, vibration-isolated and acoustically sealed, with forced ventilation if necessary. *See* EAR PROTECTORS.

Noise Sources

The principal sources of noise may be classified as surface transportation noise, aircraft noise, industrial noise, noise in the community from industrial and construction sites, and noise at home.

Surface transportation noise. There are several sources of this noise: road traffic, railroads, off-road recreational vehicles, ships, and hovercraft.

Road traffic has become the dominant source of noise annoyance in most industrialized countries. The power/weight ratio of road vehicles has been continually increased to permit higher payloads, greater acceleration, and higher cruising speeds, resulting in more powerful engines which are usually more noisy than lower-power ones. The main effect of traffic noise is annoyance caused by interference with speech sounds. Traffic noise can also interfere with sleep, although (except for heavy truck traffic) it tends to decrease at night and people tend to adapt to it. Attempts have been made to devise special noise measures such as the traffic noise index and noise pollution level to account for the annoying effects of fluctuations in level, in addition to the overall level, but these measures are not universally accepted.

In many countries, there are government regulations for car and truck noise. In the United States there are federal regulations for the maximum sound levels permitted from trucks. In addition, some cities and states have regulations for maximum sound levels for automobiles. For new vehicles, peak A-weighted sound level limits at a distance of 50 ft (15 m) from the center of the road during maximum acceleration are normally specified. These limits have been progressively reduced.

The major sources of community noise from vehicles are exhaust, cooling fan, engine, and, at high speed [above about 50 mi/h (80 km/h)], tires. Exhaust and fan noise can be reduced fairly easily; engine and tire noise are more difficult to reduce. Heavy diesel engine trucks have become the major source of traffic noise in the United States, producing about four times as much acoustic energy each day as the total fleet of automobiles. In planning new cities or new highways, care should be taken to route heavy traffic far from residential areas and to locate light and heavy industry nearer to such highways. The double glazing of windows and "soundproofing" homes in some European cities are probably equally as cost-effective as reducing the vehicle noise. Barriers can also shield residential areas from highway and railroad noise.

The interior noise for passengers of cars and rail systems is not normally a major problem. However, some truck drivers are subject to noise which is potentially hazardous if experienced for long periods.

Aircraft noise. Aircraft noise is a much more localized problem than surface transportation noise, since it occurs only around major airports. In the United States and many other industrialized countries, only about 25% as many people are seriously disturbed by aircraft noise as by road traffic noise. However, because it is localized and more easily identifiable, it seems to receive more attention. Most of the noise is produced by scheduled airlines; the contribution from the large numbers of light general-aviation aircraft is relatively small.

Jet airliners produce more community noise than did the early propeller airliners, particularly at high frequency. However, jet airliners using fanjets, which bypass a considerable proportion of the inlet air past the main compressor and combustion chambers, are about 10 or 15 dB quieter than the earlier pure-jet airliners because their design results in a lower exhaust velocity. Exhaust noise is proportional to the eighth power of velocity, so that reducing this velocity results in much lower exhaust noise. Radiation through the fan inlet and noise from the exhaust of the fan and the compressor are still problems in fanjet airliners. However, extensive use of sound absorptive liners in the fan intake and exhaust ducts has effectively decreased this noise also.

Measures such as perceived noise level (PNL) and A-weighted sound level [dB(A)] are usually used to monitor the noise of individual aircraft, and since 1969 the United States has had federal regulations which set noise limits for new passenger airliners. In order to describe the effect of aircraft noise near an airport, more complicated measures have been created to allow not only for the noise of each event (takeoff and landing) but also for the number of aircraft movements and the time of day when each occurs. In the United States the noise exposure forecast (NEF) is widely used; in Britain, the noise and number index (NNI) is used. The International Civil Aviation Organization (ICAO) has recommended the use of the equivalent continuous perceived noise level (ECPNL).

Interior noise levels in early jet aircraft were sometimes very high. These resulted from transmission of the engine noise exhaust and from fuselage wall excitation by boundary-layer pressure fluctuations. Modern fanjet airliners have much lower interior noise levels. However, interior noise levels of commuter propeller-driven passenger airliners remain intolerably high in many cases. These interior levels can be reduced by using larger slower-rotating propellers and locating them further from the fuselage.

Industrial noise. Industrial noise is a widespread problem in the United States and other industrialized countries. However, unlike traffic and aircraft noise which are mainly annoyance problems, exposure to industrial noise each day or over a period of years can cause permanent hearing damage. Probably about 5 million people in the United States have varying degrees of such damage.

TABLE 2. Permissible noise exposures for occupational noise in the United States

Duration per day, h	Sound level, dB(A)
8	90
6	92
4	95
3	97
2	100
$1\frac{1}{2}$	102
1	105
$\frac{1}{2}$	110
$\frac{1}{4}$ or less	115

In 1969 the U.S. government created industrial noise regulations under the Walsh-Healey Act, and in 1971 these were extended to cover almost all workers under the Occupational Safety and Health Act (OSHA). For every 5-dB increase in A-weighted sound level, a halving in exposure time is allowed (**Table 2**). In most other countries, using energy considerations, the halving in exposure time occurs for a 3-dB(A) increase. In the United States a maximum peak overall level of 140 dB is allowed for impulsive noise.

Most metal-cutting, metal-forming, and woodcutting machines produce intense noise; many manufacturing industries are noisy. Noise reduction methods of enclosure, absorption, and vibration isolation and damping described previously can be used in many different cases; however, the cost of achieving significant noise reduction through such engineering means is very high.

Community site noise. In countries where cities are compact, factories have been built close to residential communities. Although community industrial noise can be reduced by some of the engineering methods already described, careful planning and zoning of new factories and cities constitute the best solution to such problems. There are no federal regulations governing exterior industrial noise in the United States, although some cities and states do have regulations. Some other countries have such regulations at the national level.

Construction noise is dissimilar to exterior industrial noise in that it is more temporary in nature and sometimes continues evenings, nights, and weekends. Several phases of construction are usually involved, including site clearing; demolition; excavation; placing of foundations; erection of floors, frames, walls, windows, and pipes; and finishing, filling, and paving. Often the initial and final phases of construction are the noisiest. Two different methods have been proposed to control construction noise: setting up acceptable noise limits at site or nearest residential boundaries; and specifying acceptable noise limits for each piece of equipment used. Both approaches are used in different countries, and a mixed approach is probably desirable in the control of both exterior industrial noise and construction noise.

Noise at home. Appliances used in and around houses can cause annoyance not only to the users but to others in adjoining rooms, apartments, and even separate houses. A few of the appliances that can cause noise include furnaces, plumbing, air conditioners, fans, water heaters, pumps, dishwashers, refrigerators, vacuum cleaners, blenders, mixers, electric razors, hair driers, saws, drills, sanders, typewriters, and lawnmowers. The problem can be reduced by manufacturers producing, and consumers purchasing, quieter appliances; by more carefully constructed houses which provide better sound insulation between walls and apartments; and by the use of acoustic tiles, drapes, and carpets.

Malcolm J. Crocker

Flow Noise

It is the ear's sensitivity to pressure variation that causes the sensation of sound. When those variations propagate through the air at the speed of sound, they are acoustical waves. But when the pressures heard by the ear are unsteady only because the ear is near to and buffeted by rough flow, then the ears sense flow noise, distinct from sound in its inability to propagate away from its turbulent source.

Unsteady flow always involves pressure fluctuations whose gradients produce the forces that accelerate the fluid particles. At large scale the motions tend to evolve slowly, as in the weather, but in smaller flows they have a shorter life, particularly when the flow is fast. The ratio of flow speed to the length scale of the eddying flow is the characteristic frequency of flow noise. The flow in a round jet of speed u and diameter L is very rough and unsteady, the associated pressure variations being flow noise. In fact, that noise is most energetic at frequency $0.3u/L$, and the buffeting effects on nearby parts of an aircraft structure can cause structural failure by acoustical fatigue.

A body moving through the air carries with it a thin boundary layer where the adjacent air tends to move with the body. Often that boundary layer is turbulent, and the rough unsteady pressures within and near that turbulence are another instance of flow noise. Such unsteady pressures cause vibration of the body structure, and that in turn causes noise in the interior, which can be a problem for passengers close to the fuselage surface in an aircraft; they hear the flow noise of the boundary layer. Similarly, the surface of a submarine is buffeted by the flow noise of its surrounding boundary layer, and that interferes with the sensitive sonar equipment used for underwater navigation. The same kind of flow noise is heard inside an automobile when the windows are open, or even on a rapidly moving bicycle as the wind causes flow noise at the ear. *See* BOUNDARY-LAYER FLOW; TURBULENT FLOW.

The term flow noise is sometimes used for all noises generated by flow processes, even when that noise is organized in sound waves that propagate away from the flow. Indeed it is often very difficult to distinguish between the local field where unsteady pressure is simply a reflection of the local unsteady flow, and the acoustic elements that escape. Sometimes the term pseudosound is used for flow noise, on account of its inability to propagate as sound. Whenever the flow speed u is comparable with the sound speed c, the distinction is very hard to

maintain; but in many common flows of small characteristic Mach number, that is, where u is much smaller than the speed of sound, the coupling between flow and sound is very weak. The sound proper is then very small, with most of the buffeting noise being flow noise whose characteristics are only weakly influenced by acoustic properties. *See* MACH NUMBER.

A microphone in a wind senses flow noise that is localized to the windy region. The compressibility of air is unimportant to that noise, and flow noise does not signify any energy or power transmission. Acoustic motions are ones that depend crucially on compressibility, and this is possibly the clearest distinction that can be made between the two kinds of noise. *See* SOUND. John E. Ffowcs-Williams

Bibliography. J. Anderson and M. Brates-Anderson, *Noise*: *Its Measurement, Analysis, Rating, and Control*, 1993; A. Barber (ed.), *Handbook of Noise and Vibration Control*, 6th ed., 1992; L. L. Beranek (ed.), *Noise and Vibration Control Engineering*, 1992; M. J. Crocker, A. J. Price, and F. M. Kessler (eds.), *Noise and Noise Control*, vol. 1, 1975, vol. 2, 1982; A. P. Dowling and J. E. Ffowcs Williams, *Sound and Sources of Sound*, 1983; C. M. Harris (ed.), *Handbook of Acoustical Measurements and Noise Control*, 4th ed., 1998; R. S. Jones, *Noise and Vibration Control in Buildings*, 1984; M. P. Norton, *Fundamentals of Noise and Vibration Analysis for Engineers*, 1990; H. K. Pelton, *Noise Control Management*, 1992.

Acoustic phonetics

The discipline of acoustic phonetics can be narrowly defined as the study of the acoustic output of the vocal tract for speech, but ultimately it encompasses much more. Acoustic phonetics makes direct contact with, and in many cases is the foundation of, areas of study such as speech synthesis, automatic (computer) recognition of speech, speech perception, phonological analysis and theory, and speech pathology.

Source-filter theory. There is a well-articulated theory of acoustic phonetics, the source-filter theory, where the source is the input to the system and the filter is the resonator. It is easiest to introduce the theory by first discussing vowels, the kinds of speech sound for which the theory is best known and for which it makes the most precise predictions.

Source. The source is the energy produced by the vibrating vocal folds (**Fig. 1**), which are set into motion by a combination of muscular forces that close the vocal folds as well as raise the pressure in the trachea relative to the pressure above the vocal folds. When the tracheal pressure overcomes the resistance offered by the closed vocal folds, the latter are blown apart and set into oscillatory motion that will continue, by virtue of a delicate interplay of aerodynamic and mechanical forces, as long as an adequate pressure differential exists. The nearly periodic vibration of the vocal folds can be characterized in either the time domain (as a waveform) or frequency domain (as a spectrum); most models of the vibration focus on parameters of the time-domain representation and their relationship to the spectral composition of vibratory energy. Vibration of the vocal folds generates a consecutive-integer harmonic spectrum, with greatest energy at the fundamental frequency (F0), or lowest-frequency component, and decreasing energy across the higher harmonics at the frequencies 2F0, 3F0, 4F0, 5F0, . . . ,nF0. The time-domain models show how subtle changes in the temporal parameters result in systematic frequency-domain changes. In certain cases, the periodic output of the vibratory motion may be mixed with variable degrees of aperiodic energy, as in more breathy types of voice. Thus the frequency-domain representation, also called the glottal source spectrum, can be associated very directly with the perception of voice quality (encompassing harshness, breathiness, and so forth). *See* HARMONIC (PERIODIC PHENOMENA); VIBRATION.

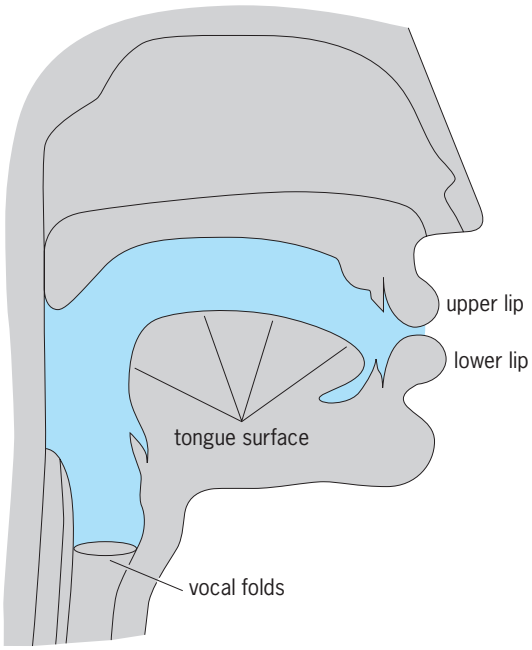

Fig. 1. Midsagittal schematic view of vocal folds (source) and vocal tract (filter, colored area).

Filter. The glottal source spectrum can be considered as the input to the filter, which in anatomical terms is the vocal tract. The vocal tract is the length of airway running from the vocal folds, in the larynx, to the lips (Fig. 1). The vocal-tract filter has the characteristics of a resonating tube, with one end closed. The closed end is at the vocal folds because the acoustic energy generated there excites the vocal-tract filter each time the vocal folds snap shut during consecutive vibratory cycles. The open end is at the lips because vowel sounds are produced with the mouth open to varying degrees, depending on the vowel. A tube resonator with a closed end and uniform cross-sectional area from end to end has multiple resonant frequencies determined by the tube's length. For such a tube having a length of 17 cm (6.7 in.),

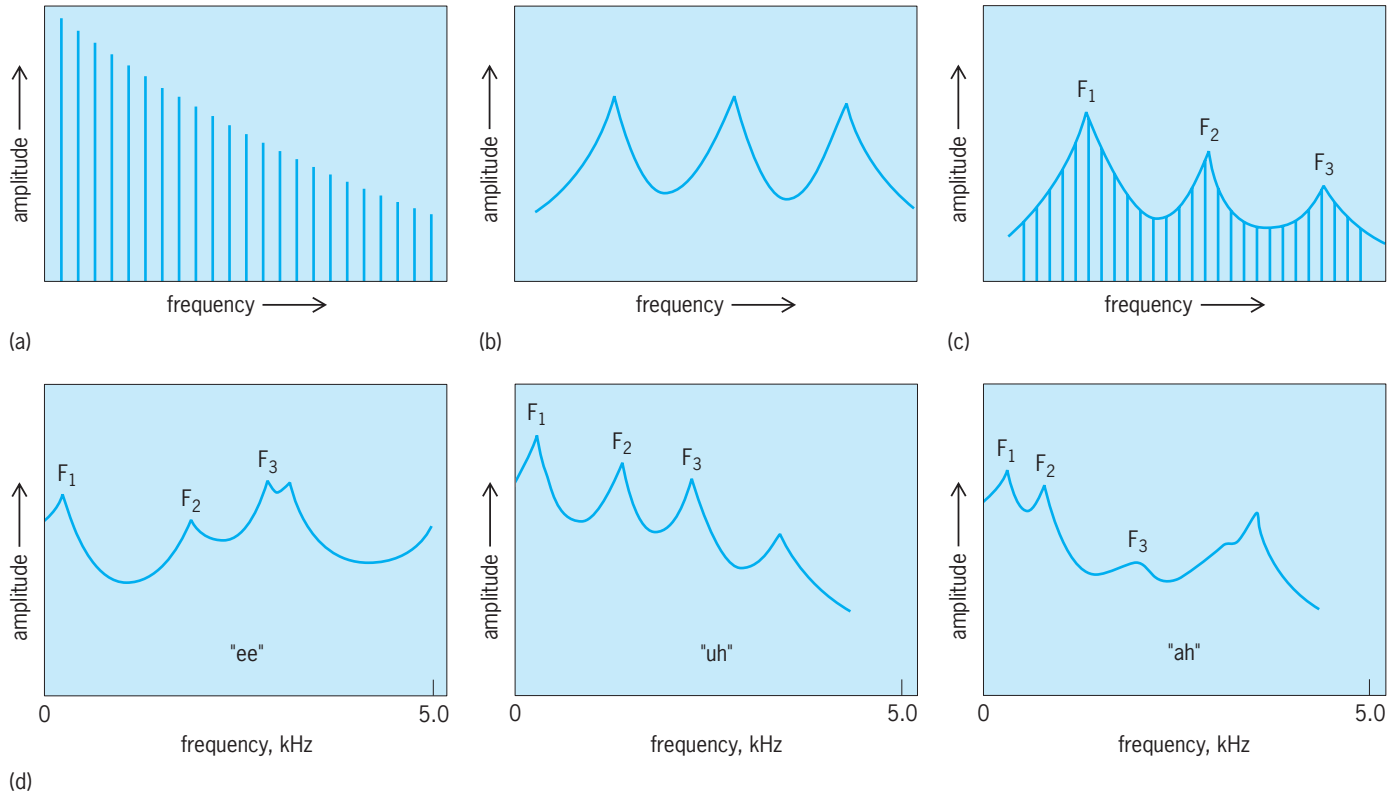

Fig. 2. Production of output spectrum of the vocal tract. (*a*) Source spectrum produced by vibrating vocal folds. (*b*) Resonance curve (filter) of vocal-tract tube having uniform cross-sectional area. (*c*) Output spectrum from combined source and filter. (*d*) Output spectra for vowels "ee," "uh," and "ah."

a typical length for the adult male vocal tract, the resonant frequencies are determined by the quarter-wavelength rule, given by the equation below,

$$f_r = \frac{(2n - 1)\,c}{4l}$$

where f_r is a resonant frequency, $n = 1, 2, 3, \ldots$, c is the constant speed of sound in air (35,400 cm/s), and l is the length of the tube. This formula gives an odd-integer series of resonant frequencies at approximately 520 Hz, 1560 Hz, 2600 Hz, and so forth. These resonant frequencies can be thought of as peaks in a theoretical spectrum (where "theoretical" means calculated from theory, rather than being measured empirically), which have noninfinite amplitude as a result of both energy loss within the vocal tract and sound radiating from the lips. The overall shape of the theoretical spectrum can be thought of as the resonance curve or filter function of the vocal tract.

Output signal. The filtering effect of the vocal tract can be thought of in another way, that of the vocal-tract filter function shaping the source harmonic energy generated by the input, or vibrating vocal folds. The output signal thus represents the combined effect of the source and filter (**Fig. 2***a–c*). The peaks in the output spectrum are referred to as formants, and the frequency locations of the first three formants (F1, F2, and F3, the first three resonances of the tube) are a function of the configuration of the vocal-tract tube, which is to say of the vowel articulation. The resonant frequencies predicted by the quarter-wavelength rule—when the tube has a uniform cross-sectional area from end to end—are changed in systematic ways that depend on the location of constrictions relative to the standing-wave patterns of pressure and velocity within the tube. Thus a vowel such as the "ee" in the word "heat," for which the vocal tract tube shape is constricted toward the front (that is, toward the lips) and wide open toward the back (that is, toward the vocal folds) is associated with a resonance curve, and therefore with an output spectrum, substantially different from the case of the tube without any constrictions (as for the first vowel "uh" in the word "about"); on the other hand, a vowel like the "ah" in "father" constricts the tube toward the back and leaves the front wide open, yielding a different kind of deviation from the quarter-wavelength resonances of an unconstricted tube (Fig. 2*d*).

Although a tube resonator has, in theory, an infinite series of resonance peaks, only the first three formant frequencies have critical phonetic importance in the sense of distinguishing acoustically among the vowels of a language. Not surprisingly, these same three formant frequencies have been shown to be of critical importance in the perception of vowel categories, although in some cases the auditory system may not treat the individual formant frequencies as separate pieces of acoustic information. *See* SPEECH PERCEPTION.

Consonants and nasals. The model of a source spectrum shaped by a vocal-tract filter or resonator also

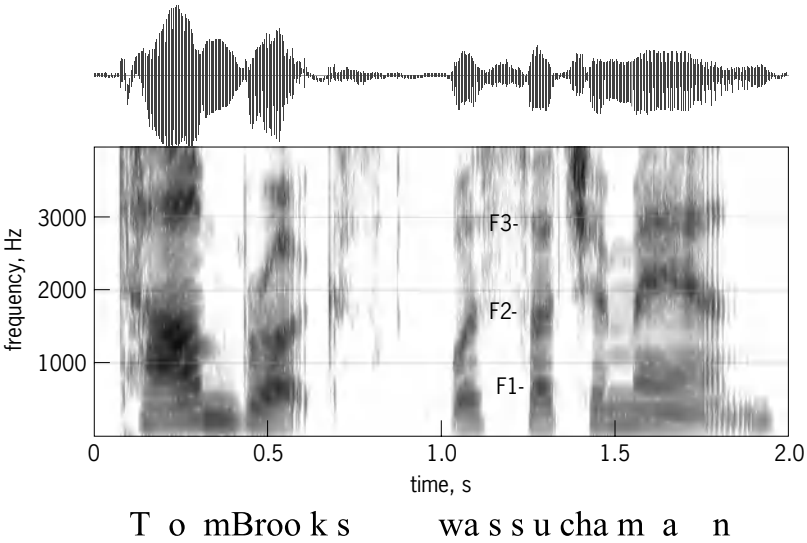

Fig. 3. Spectrogram of the utterance, "Tom Brooks was such a man." Intensity is coded by the gray scale. The speech waveform (amplitude over time) is at the top of the plot. Letters beneath the plot show the approximate locations of the sounds. First three formants, F1, F2, and F3, are indicated for the vowel "uh" in "such."

applies to consonants such as stops (such as p, b), fricatives (such as s, z), and affricates (such as ch, j), but in these cases the sources are likely to involve aperiodic (inharmonic) spectra generated by impulse or turbulent air flows. For some sounds (such as b, z, j) aperiodic sources within the vocal tract are combined with the periodic vocal fold source described above. Another class of sounds, the nasals (such as m, n), require a theory for parallel resonating tubes where one of the tubes has a dead end and therefore traps energy at its resonant frequencies. In this case the theoretical and empirical spectrum will have reverse peaks or antiresonances, indicating frequencies at which the combination of source and filter results in little or no output energy. Antiresonances are also found in the spectra of the stops, fricatives, and affricates.

Properties of speech sounds. The last half-century has seen an enormous amount of research on the temporal and spectral properties of speech sounds. This work has been critical to the design of speech synthesizers and algorithms for computer recognition of speech, and in some cases to the design of hearing aids and signal-processing schemes for cochlear implants; the data have also proven useful in understanding speech disorders. The standard representation of speech sounds in acoustic phonetics research is the spectrogram (**Fig. 3**), which shows the formant frequencies of the vocal-tract signal as a function of time, with intensity coded in a gray or color scale. Measurement conventions have been established for the correspondence between temporal intervals and sound categories, as well as for the spectral characteristics of speech sounds. A large amount of data has been generated for naturally spoken vowel and consonant durations and spectra, as well as the characteristics of the vocal fold vibratory source, and these have been used to design high-quality speech synthesizers. The very high intelligibility of sophisticated speech synthesizers, whose algorithms are based on these measurements, is strong evidence for the high quality of acoustic-phonetics research. *See* ACOUSTIC SIGNAL PROCESSING; HEARING IMPAIRMENT; SPEECH RECOGNITION; VOICE RESPONSE.

Factors affecting speech sounds. A predictable characteristic of any spectral measurement of vocal-tract output is its dependence on the age and gender of the speaker. Because the vocal tract resonates like a tube closed at one end with tube length being a scaling factor for the resonant frequencies, the formant frequencies of a particular sound spoken by men, women, and children will be quite different as a result of their general size differences, which are mirrored by differences in vocal-tract length. More generally, the acoustic characteristics of any speech sound, including their durations and spectra, are affected by many factors including (but not limited to) age, gender, speaking rate, speaking style, dialect, linguistic stress (for example, *re*bel versus re*bel*), and phonetic context (the sounds surrounding the sound of interest). Thus, the acoustic characteristics of a given sound cannot be considered as fixed or templatelike. One challenge of research in acoustic phonetics is to account for the factors that can affect the acoustic representation of speech sounds, and to relate these factors to speech synthesis, automatic speech recognition, speech perception, phonological theory, and clinical application in the diagnosis and understanding of speech disorders. Sophisticated models of the vocal tract are available to complement models of vocal vibration, allowing careful experimentation on the specific factors that contribute to the acoustic variation of speech sounds in natural speech. *See* SOUND; SPEECH.

Gary Weismer

Bibliography. G. Fant, *Acoustic Theory of Speech Production*, rev. ed., Mouton de Gruyter, The Hague, 1970; K. Forrest et al., Statistical analysis of word-initial voiceless obstruents: Preliminary data, *J. Acous. Soc. Amer.*, 84:115–123, 1988; J. Hillenbrand et al., Acoustic characteristics of American English vowels, *J. Acous. Soc. Amer.*, 97:3099–3111, 1995; D. H. Klatt and L. C. Klatt, Analysis, synthesis, and perception of voice quality variation among female and male talkers, *J. Acous. Soc. Amer.*, 87:820–857, 1990; K. N. Stevens, *Acoustic Phonetics*, MIT Press, Cambridge, 1998; B. H. Story, A parametric model of the vocal tract area function for vowel and consonant simulation, *J. Acous. Soc. Amer.*, 117:3231–3254, 2005.

Acoustic radiation pressure

The net pressure exerted on a surface or interface by an acoustic wave. One might presume that the back-and-forth oscillation of fluid caused by the passage of an acoustic wave will not exert any net force on an object, and this is true for sound waves normally encountered. (The sound power of a normal speaking voice is less than one-millionth of the electric power

of a 100-W light.) Intense sound waves, however, can exert net forces in one direction of sufficient magnitude (proportional to the sound intensity) to balance gravitational forces and thus levitate an object in air. These acoustic forces, generated by acoustic radiation pressure, are generally not of the magnitude to lift a table off the floor, although such is not inconceivable.

The nature of the acoustic radiation pressure forces was clarified in a simple experiment by G. Hertz and H. Mende in which an acoustic transducer in water sends a beam of sound at a water–carbon tetrachloride interface. These mutually immiscible liquids are chosen because their properties are matched in such a way as to permit the passage of the sound beam across the interface with minimal reflection. But as the sound intensity is increased, an observer will notice that the interface statically deforms toward the water. Moreover, if the transducer is immersed in the carbon tetrachloride and pointed at the interface, the interface will again deform toward the water, suggesting that the difference in acoustic energy densities between the two liquids—a direction-independent quantity—is the relevant factor.

The problem is somewhat more complex when considering the force on a small object rather than an interface, not only because the problem is three-dimensional, but also because the scattering from the object and the oscillating and translatory motion must be considered. Such a force is relatively easy to generate by using a standing acoustic wave field, as is found when sound of a certain resonance frequency bounces back and forth in an organ pipe. An acoustic standing wave field is characterized by regions of acoustic pressure maxima and minima. For rigid objects in air or liquid, the sound field pushes the object toward minima. But objects that have some significant compressibility as compared to the surrounding medium (such as some liquid drops, or small bubbles held in liquid) are pushed toward maxima.

Forces due to acoustic radiation pressure have been used to calibrate acoustic transmitters, to deform and break up liquids, to collect like objects or to separate particles (including biological cells) based on mechanical properties, and to position objects in a sound field, sometimes levitating the sample so that independent studies of the object's properties can be performed. Single bubble sonoluminescence phenomena depend on acoustic radiation forces to maintain a bubble in a zone while its substantial radial oscillations take place.

Radiation pressure of sound in air and in liquids has been used to suppress the capillary instability of cylindrical liquid surfaces that ordinarily causes long liquid columns to break into drops. Experiments in the 1970s demonstrated that low-frequency oscillations in fluid systems could be generated by modulating acoustic radiation pressure. This technique has been applied to imaging the low-frequency elastic response of tissue in the human body. *See* ACOUSTIC LEVITATION; SONOLUMINESCENCE; SOUND; ULTRASONICS.

Robert E. Apfel; Philip L. Marston

Acoustic radiometer

A device to measure the acoustic power or intensity of a sound beam by means of the force or torque that the beam exerts on an inserted object or interface. The underlying theory involves the concept of radiation pressure, which is the time-independent part of the pressure associated with a nominally sinusoidal acoustic disturbance. Such pressure occurs, for example, when a plane sound wave is partially reflected at an interface between two materials, with the nonlinear interaction between the incident and reflected waves giving rise to a steady pressure on the interface. If a narrow beam is incident on the interface and the transmitted wave is fully absorbed by the second material, the magnitude of the radiation force F (area integral of radiation pressure) equals a constant times W/c, where W is the power of the sound beam and c is the sound speed. The multiplicative constant is of the order of unity and depends on the thermodynamic properties of the fluid. The force acts in the vector direction of propagation of the sound wave. When the inserted object is of small size, the force is not given by as simple an expression, but it can nevertheless be predicted from basic principles. Thus, acoustic radiometers do not require external calibration and are themselves sometimes used in the calibration of transducers.

Vane radiometer. A modern design of an acoustic radiometer, used to measure the total power of an ultrasonic sound beam in water and other liquids, employs a vane suspended in the fluid in such a manner that its displacement in a direction normal to its face is proportional to the net force pushing on its front face. The vane is ideally of dimensions somewhat larger than the incident beam's diameter, so that the encountered force is associated with the entire incident beam. To eliminate the possibility of sound being reflected toward the transmitting transducer, the vane is oriented at 45° to the incident sound

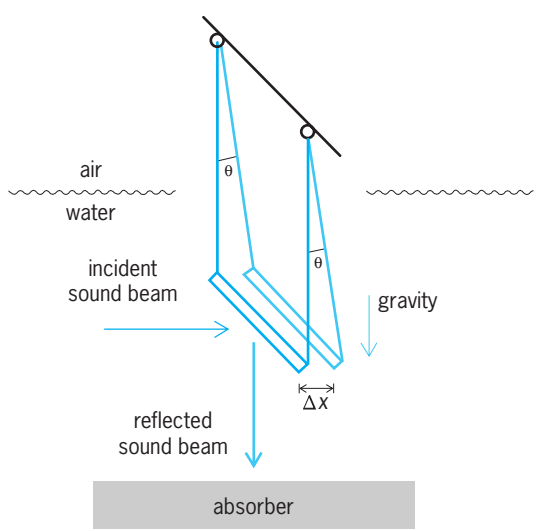

Acoustic radiometer, used to measure the acoustic power of a sound beam in water. The vane is suspended from a pendulum whose rotation is opposed by gravity. Deflection Δx of the vane is proportional to the force on the vane.

beam. The vane's horizontal displacement is made to be proportional to the imposed force by fastening the vane at one end of a long pendulum (see **illus.**) whose rotation from the vertical is opposed by the effect of gravity, such that the apparent spring constant for displacement in a direction at 45° to the face is approximately Mg/L, where M is the apparent mass of the vane (corrected for the presence of water), g is the acceleration of gravity, and L is the length of the pendulum. A nonlinear acoustics theory for such a circumstance yields a proportionality relation between the net horizontal radiation force on the vane and the acoustic power associated with the incident beam. Because the deflection of the vane is proportional to the radiation force, the acoustic power can be determined. *See* BUOYANCY; PENDULUM.

Rayleigh disk. The concept of the vane device evolved from that of the Rayleigh disk, which was a circular disk that could rotate about its diameter and whose deflection from a nominal 45° orientation was opposed by a torsional spring. The Rayleigh disk was taken to have a radius much smaller than the wavelength, and its use ideally yielded a measurement of the local acoustic intensity that would have existed at the center of the disk were the disk not present. This is in contrast to the vane device, which is intended for use in circumstances when the vane's surface dimensions are large compared to a wavelength and the incident beam is entirely captured by the vane's surface. *See* ACOUSTIC RADIATION PRESSURE; SOUND; SOUND INTENSITY; SOUND PRESSURE.

Allan D. Pierce; Sameer I. Madanshetty

Acoustic signal processing

A discipline that deals generally with the extraction of information from acoustic signals in the presence of noise and uncertainty. Acoustic signal processing has expanded from the improvement of music and speech sounds and a tool to search for oil and submarines to include medical instrumentation; techniques for efficient transmission, storage, and presentation of music and speech; and machine speech recognition. Undersea processing has expanded to studying underwater weather and long-term global ocean temperature changes, mammal tracking at long ranges, and monitoring of hot vents. These techniques stem from the rapid advances in computer science, especially the development of large, inexpensive memories and ever-increasing processing speeds.

Sound can be said to be low and slow compared to computers. The audio frequencies are low, spanning roughly the range from 15 to 25,000 hertz, with seismic frequencies below this range and frequencies relevant to medical ultrasonics above. All of these frequencies are well within the limits of analog-to-digital input samplers and digital-to-analog output converters that form the bridges between continuous-time physical waveforms and the streams of digits handled by a computer. The speed of sound is slow, about 335 m/s (1100 ft/s) in air and 1500 m/s (5000 ft/s) in water, compared to light and radio waves at 3×10^8 m/s (10^9 ft/s), the upper limit for electrical signals in computer circuits. A computer can carry out thousands of elementary computations between input samples. *See* ANALOG-TO-DIGITAL CONVERTER; DIGITAL-TO-ANALOG CONVERTER; SOUND.

Time and frequency domains. Signal processors, both human and computer, "think" in two domains, two pictures of the signal world. One domain is the pictures of waveforms, the sound pressure as it changes in time. This time-domain picture is $s(t)$; t is time, and s might be the sound pressure or a microphone voltage. The other picture is the complex magnitude at every important frequency during an interval of time. This frequency-domain or spectrum picture is $S(f \mid T_n)$; f is the frequency in hertz, and T_n is the nth time interval. J. Fourier (1768–1830) formalized his series version of this picture and showed that $s(t)$ could be calculated from $S(f)$, and vice versa. Mathematicians have developed other versions of spectral transforms, each with its special area of application. *See* FOURIER SERIES AND TRANSFORMS; INTEGRAL TRANSFORM.

Each domain picture provides the same information, but sometimes it is easier to think, or compute, using one domain rather than the other. The computer version used is the DFT (discrete Fourier transform). Transforms were formerly time-consuming computations; for N points in one domain the time needed was proportional to N squared. In the 1960s a layered algorithm was developed to speed up the DFT. The emphasis is on sample sizes which are integer powers of 2. For example, 1024 is the tenth power of 2; it uses 10 layers, and the time factor dropped from 1024∗1024 to 10∗1024, over 100 times faster. These fast Fourier transforms have had a major impact on spectral processing.

Medical applications. Ultrasonic imaging of the unborn fetus is commonly used to observe growth and gender, as well as the positioning of multiple fetuses, and to warn of delivery problems. Similar techniques are used to view the heart in action, and to carry out breast examinations for cancerous nodules. Ultrasonic imaging devices typically use a pulse transmitter surrounded by a two-dimensional matrix of microphones. The array can be focused about a given depth by selecting the travel time (the time it takes the pulse to go from the transmitter to a specific location and back to a receiver). Tiny changes in travel time are manifested as signal phase changes, and at each depth the processor focuses on each point in the plane by adjusting the received signal phases to add constructively. The Doppler effect is used to observe blood flow: blood flowing toward the sensors causes the reflected frequency to be higher than transmitted, and blood flowing away lowers the frequency. Doppler imaging needs long pulses to measure slight changes in frequency, so wideband frequency-modulated signals may be used to achieve the time resolution of short broadband pulses and the frequency resolution of longer-duration signals. *See* BIOMEDICAL ULTRASONICS; MEDICAL ULTRASONIC TOMOGRAPHY; ULTRASONICS.

The cochlea in the inner ear is nature's time-domain–to–frequency-domain transformer, exciting

a line of nerve endings in response to something like the short-time spectrum of the input. For people in whom this natural mechanism is dysfunctional but the cochlear nerve is intact, a technology for restoring hearing is developing based on modest DFT analysis of sound near an ear, feeding perhaps 32 channels to an implanted microconnector in the cochlear nerve bundle. *See* HEARING (HUMAN); HEARING AID; PHYSIOLOGICAL ACOUSTICS.

Hearing aids can make use of miniaturized electronics to provide tailored spectral gain and overall dynamic gain adjustment to keep the total signal amplitude at a suitable level. *See* HEARING AID.

A binaural "eyeglasses sonar" for blind persons who have normal hearing was developed in New Zealand in the 1970s. With a transmitter in the bridge and receivers on either side, the beam moves naturally with head motion. The signal is a periodic, octave-wide, ultrasonic frequency sweep. The signal processing removes the non-information-bearing sounds due to the periodicity of the signal, and frequency translates the signal to the normal hearing frequencies. It is a sensitive motion detector. The intelligent signal processing is done by the wearer, whose head movement and walking pace cause the auditory sensations that contain the information about the wearer's surroundings that would be provided by vision for a sighted person.

Applications to rooms and halls. Audio filters are the hardware that adjust the gain across the spectrum to emphasize some frequencies and deemphasize (notch) others. Analog filters required heavy bulky components for low frequencies. Digital filters are programs in computer chips of negligible weight. Digital "equalizers" allow the user to adjust the spectral gain in 3 to 30 adjacent frequency bands to suit the listener. In churches and lecture halls that echo and emphasize resonant frequencies, fixed-frequency notch filters may be used; a dynamic system may use a fast Fourier transform to watch for frequencies that are too loud and apply appropriately deep notches, then remove the notches when the bothersome frequencies are no longer excited. For halls with distributed speakers, the signal can be written into a short memory and read out at desired delay times to different speakers; this is a natural solution for long narrow churches. *See* DIGITAL FILTER; ELECTRIC FILTER; EQUALIZER; SOUND FIELD ENHANCEMENT.

Transmission and storage applications. Part of information theory deals with removing all the redundancy in a processor input so that only essential, information-bearing bits remain, just enough to restore the original signal. Lossless algorithms yielding 2:1 compression for text and data are in common use. Lossy compression tolerates minor errors in reconstruction and may further reduce the number of needed bits. With the interest in sending speech, music, and video as well as written text over the Internet, algorithms were developed yielding 12:1 compression on such material. *See* DATA COMPRESSION; INFORMATION THEORY.

Speech recognition. One of the most difficult signal processing areas is machine recognition of spoken speech. Vocabulary size and number of speakers are key parameters. The fundamental problems are: (1) words are not spoken separately, but in streams of connected sound, (2) a phrase is seldom said the same way every time, (3) the vocabulary is enormous, and (4) there is a huge variety of speaker accents and rhythms. The signal processing must segment an utterance into phonemes, words, or phrases; pick out key features; or compare the whole segment with a library for likely matches.

Many telephone information systems now accept single digits spoken in English and perhaps Spanish. This avoids the first three problems. Also in use are user-specific algorithms that are trained separately with each user, with a specific library of commands. User-specific dictation algorithms with large vocabularies are under development. *See* SPEECH RECOGNITION.

Sonar. The ocean is transparent to sound and nearly opaque to light and radio waves. Even for sound, the attenuation increases exponentially as the frequency squared, so low audio frequencies must be used to transmit signals as distances increase. The speed of sound varies with temperature and pressure, causing sound to travel in paths that oscillate from near the surface to kilometers deep in the oceans. In most underwater situations the propagation is by multiple paths or by multiple modes, with many interfering sources. *See* SOUND ABSORPTION; UNDERWATER SOUND.

Input processing. The receiving array is the initial signal processor. A large number of hydrophones in a three-, two-, or one-dimensional array is used to focus a beam in both vertical and horizontal angles. Each beam former adds the inputs from every hydrophone after each is time-delayed to synchronize all inputs from the desired angle, and after differences in individual gains are compensated to form a clean beam. The gain in average signal-to-noise ratio is proportional to the number of hydrophones; however, ships, biological organisms, and storms are localized sources of acoustic noise, and therefore may be suppressed more than average by the beam pattern.

Postinput processing. Signal detection and estimation theory guide the processing of noise, interference, and their power levels as well as the uncertainty as to their exact description (stochastic and mechanistic) and the uncertainty in the received signal (which may also be part stochastic and part mechanistic). The ultimate goal in processing is to include the total reception as input, together with the propagation possibilities, noise, and interference, and to determine how likely each combination of these would be to cause the reception just observed. Signal detection, parameter estimation, and propagation learning proceed together, not sequentially. *See* ESTIMATION THEORY; STOCHASTIC PROCESS.

This approach was impractical in the past unless one met or assumed specific statistics. Now the speed, lower cost, and increased memory of computers make this viewpoint ever more applicable; what had been considered "random" may now be considered "complicated but manageable." This opens the

way to vastly improved performance at low input-signal–to–noise ratios. *See* SONAR.

Theodore G. Birdsall

Acoustic torpedo

An autonomous undersea vehicle that can be launched from submarines, surface ships, or aircraft to attack enemy submarines and surface ships. An acoustic torpedo is a sophisticated weapon. Its main components are a guidance and control system, a power plant to provide propulsive and electrical energy, a propulsor to control speed and direction, and a warhead. The launching platform performs the function of determining the approximate location of the target and launching the torpedo in the proper direction. The torpedo typically utilizes an acoustic sensor in its nose and is controlled by an on-board computer. During its operation, the torpedo searches the volume of the ocean determined by the launch platform. It progresses through the following phases: detection (an object is present), classification (the object is a target of interest), homing (steer at the object), and detonation of the warhead. *See* GUIDANCE SYSTEMS; HOMING.

Acoustics is the primary means of propagation in the undersea environment. The acoustic sensor enables the detection, classification, and homing functions. An electromechanical device converts acoustic (pressure) waves in the ocean to electrical signals that are processed. Similarly, electrical signals produced by the torpedo are converted to acoustic waves for active sonar transmissions. Because of their small size, torpedoes utilize high frequencies of operation. This results in short detection ranges, perhaps several hundred meters. *See* ACOUSTIC MINE; TRANSDUCER.

Acoustic torpedoes have two modes of sonar operation: passive and active. In passive sonar the torpedo listens for sounds emitted by the target. If the emissions are detected and classified relative to other sounds in the ocean, the torpedo enters into homing on the target. In active sonar the torpedo transmits acoustic pulses that reflect off objects in the environment. These reflections are processed on board the torpedo for detection, classification, and homing. *See* SONAR.

Limitations on torpedo performance are provided by three mechanisms: background noise provided by the flow of water over the sensor face; reverberation (the backscattering of acoustic energy from the ocean surface, bottom, and particles in the water column); and acoustic devices that are designed to deliberately produce noise to mask the target or to provide a decoy to divert the torpedo. These limitations are addressed through a combination of processing algorithms, strategies for tactical employment, and logic commands that are inserted into the on-board computer. *See* ACOUSTIC SIGNAL PROCESSING; ANTISUBMARINE WARFARE; ELECTRONIC WARFARE; REVERBERATION; UNDERWATER SOUND.

Frank W. Symons

Acoustical holography

The recording of sound waves in a two-dimensional pattern (the hologram) and the use of the hologram to reconstruct the entire sound field throughout a three-dimensional region of space. Acoustical holography, which first appeared in the 1960s in studies in ultrasonics, is an outgrowth of optical holography, invented by Dennis Gabor in 1948. The wave nature of both light and sound make holography possible. The objective of optical holography is to observe (reconstruct) three-dimensional images of the sources of reflected light (visible or nonvisible). Acoustical holography involves reconstruction of the sound field that arises due to radiation of sound at a boundary, such as the vibrating body of a violin, the fuselage of an aircraft, or the surface of a submarine. This reconstruction represents a solution to an inverse wave propagation problem explained heuristically using Huygens' principle. Both acoustical holography and optical holography rely on the acquisition of an interferogram, a two-dimensional recording at a single frequency of the phase and amplitude of an acoustic or electromagnetic field, usually in a plane. Gabor called this interferogram a hologram. *See* HOLOGRAPHY; HUYGENS' PRINCIPLE; INVERSE SCATTERING THEORY.

Farfield and nearfield holography. Two distinct forms of acoustical holography exist. In farfield acoustical holography (FAH), the hologram is recorded far removed from the source (in the Fresnel or Fraunhofer zones). This form of acoustical holography is characterized by the fact that the resolution of the reconstruction is limited to a half-wavelength. This resolution restriction is removed, however, when the hologram is recorded in the acoustic nearfield, an important characteristic of nearfield acoustical holography (NAH), invented by E. G. Williams and J. D. Maynard in 1980.

In all forms of holography the reconstruction of the field is based theoretically on a mathematically rigorous inversion of the Helmholtz integral equation. In practice, however, most applications compromise much of this mathematical rigor, especially in farfield acoustical holography. Nearfield acoustical holography treats the inverse problem much more rigorously than farfield acoustical holography, due to the inclusion of evanescent waves in the reconstruction process. Evanescent waves decay exponentially from the source surface, where they are generated. In order to capture them the hologram must be recorded very close to the source surface, that is, in the acoustic nearfield. The reconstruction provided by the solution of the inverse problem can be thought of as a backtracking of the sound field from the measurement surface to the source. This backtracking process is carried out in a computer which is able to increase exponentially the amplitude of the evanescent waves until their initial value at the source surface is reached. It is these waves that contain the high-resolution spatial information about the source. The most accurate reconstructions are obtained when the two-dimensional hologram is

larger than the object (**Fig. 1**) or when the hologram is a closed surface conformal to the surface of the vibrating object.

Using multiple reconstructions, a three-dimensional image of the sound field is obtained (**Fig. 2**). With this volume mapping, it is possible to derive the particle velocity, acoustic intensity, energy densities, and farfield directivity patterns. Besides the reconstruction of the pressure, nearfield acoustical holography is able to reconstruct the normal velocity on the surface of the object, providing detailed information about its vibrations—a key to unraveling its vibration and radiation mechanisms. Nearfield acoustical holography can also reconstruct sound fields within boundaries (for example, the noise inside an aircraft). In that case, the hologram is acquired in the interior, and multiple reconstructions provide the complete acoustic field in the interior as well as the velocity on the fuselage.

Nearfield acoustical holography has been used in the automotive industry to study interior noise and tire noise, in musical acoustics to study vibration and radiation of violin-family instruments, and in the aircraft industry to study interior cabin noise and fuselage vibrations. Applications are also found in underwater acoustics, especially in studies of vibration, radiation, and scattering from ships and submarines. *See* ACOUSTIC NOISE; MUSICAL ACOUSTICS; UNDERWATER SOUND.

Data acquisition and reconstruction. Typically, temporal acoustic data are acquired by measurement of the acoustic pressure with a single microphone or hydrophone, which scans an imaginary two-dimensional surface. This surface is most often a planar or cylindrical one, but surfaces of varied shapes have been used. In some cases, an array of microphones is used and the pressure is measured instantaneously by the array.

Fig. 1. Underwater hologram measurement system for the study of vibration, radiation, and scattering from air-filled steel shells. Diagram shows vibrating source and array of points marking measurement locations in the hologram. Hydrophone and and axial scanner are not shown.

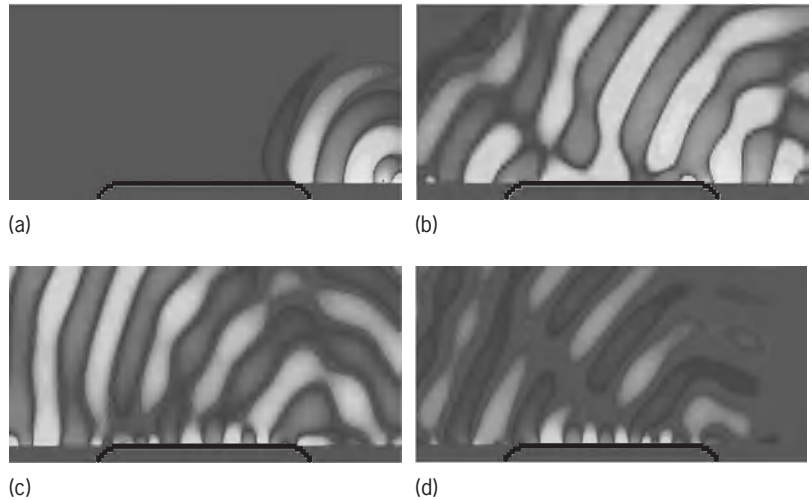

Fig. 2. Reconstruction, by means of nearfield acoustical holography, of the pressure field at four instants in time during the scattering of a wave pulse (traveling from right to left). Different grays indicate positive pressure, negative pressure, and ambient pressure. (*a*) 50 μs. (*b*) 350 μs. (*c*) 650 μs. (*d*) 950 μs.

Figure 1 shows an example of the measurement of a cylindrical hologram used for a coherent vibrator and scatterer underwater. In this case, the pressure field is sensed by a single hydrophone attached to an axial scanning device. In the circumferential direction, data are obtained by rotation of the vibrating object about its axis. At each data point in the scan, the pressure is recorded in a computer as it varies with time, along with a reference signal which provides a phase reference and amplitude calibration. Typical reference signals might be the force applied by a force generator exciting the object, the acceleration at a fixed point on the body, or the incident field pressure (for a scattering hologram).

The measured data are processed in a computer to reconstruct the pressure at the surface of the object as well as the vibration of the surface. In this processing, the measured time data is Fourier-transformed into the frequency domain, creating a set of holograms, one for each frequency bin in the transform. In the inversion process, each hologram is broken up into a set of waves or modes whose propagation characteristics are known from basic principles. Each wave or mode is then back-propagated to the source surface by multiplication by the known inverse propagator, and the field is then recomposed by addition of all these waves or modes. Evanescent modes beyond the dynamic range of the system are filtered out and are not included in this recomposition. *See* FOURIER SERIES AND TRANSFORMS.

Two kinds of inverse propagators exist, one which reconstructs velocity and the other which reconstructs pressure. This whole process is carried out for each hologram (each frequency). If desired, the reconstructed holograms are inverse-transformed from frequency back to time. An example is shown in Fig. 2 for a hologram generated by the scattering

of a wave pulse (traveling from right to left) off a cylindrical, air-filled, steel shell submerged in a water tank. The four time snapshots show the reconstructed pressure field not only at the surface of the shell (drawn at the bottom of each snapshot) but also at 32 concentric cylindrical surfaces of increasing radius (32 separate reconstructions) put together to generate a planar view. Earl G. Williams

Bibliography. J. W. Goodman, *Introduction to Fourier Optics*, 2d ed., McGraw-Hill, New York, 1996; E. G. Williams, *Fourier Acoustics, Sound Radiation and Nearfield Acoustical Holography*, Academic Press, London, 1999.

Acoustics

The science of sound, which in its most general form endeavors to describe and interpret the phenomena associated with motional disturbances from equilibrium of elastic media. An elastic medium is one such that if any part of it is displaced from its original position with respect to the rest, as for example by an impact, it will return to its original state when the disturbing influence is removed. Acoustics was originally limited to the human experience produced by the stimulation of the human ear by sound incident from the surrounding air. Modern acoustics, however, deals with all sorts of sounds which have no relation to the human ear, for example, seismological disturbances and ultrasonics.

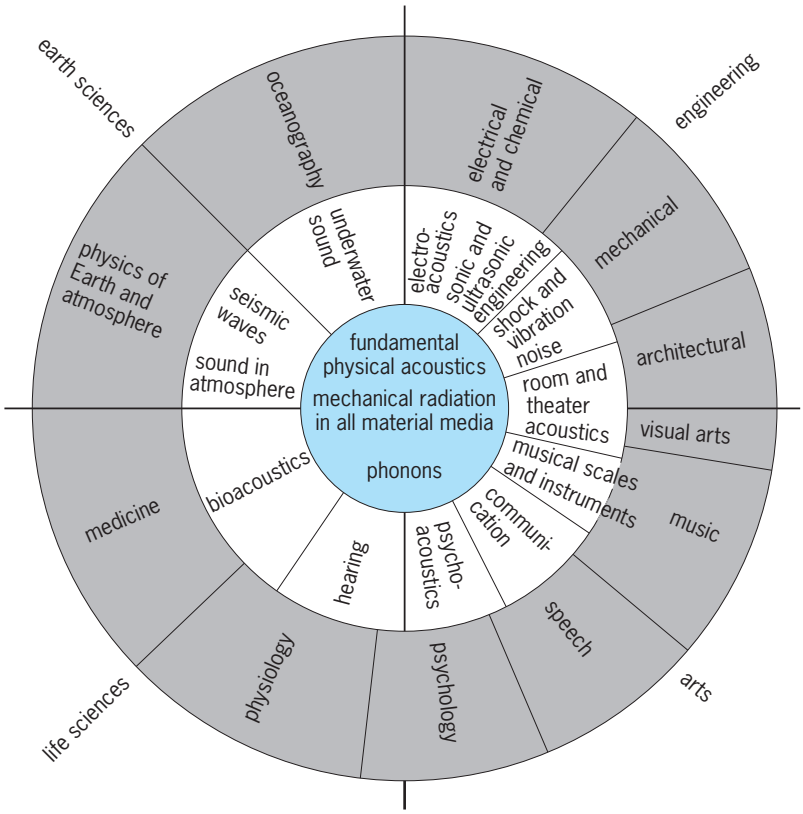

Relations among the various branches of acoustics and related fields of science and technology.

Basic acoustics may be divided into three branches, namely, production, transmission, and detection of sound.

Sound production. Any change of stress or pressure producing a local change in density or a local displacement from equilibrium in an elastic medium can serve as a source of sound. The human vocal mechanism is an obvious example. Other illustrations are provided by struck solids (as a drum, or violin or piano string), flow of air in a jet, and underwater explosions.

Modern acoustics as a science and technology has experienced enormous development by the invention of sources of sound which can be precisely controlled with respect to both frequency and intensity. The most important of these are called acoustic transducers, namely, devices by which any form of energy can be transformed into sound energy and vice versa. The most useful type is the electroacoustic transducer, in which electrical energy is transformed into the mechanical energy of sound. An example is the electrodynamic loudspeaker used in sound-reproducing systems, public address systems, and radio and television. Piezoelectric and magnetostrictive transducers are widely used in scientific and industrial applications of acoustics. An example of a nonelectric acoustic transducer is the siren, in which interrupted fluid flow results in the production of sound.

Sound transmission. Transmission of sound takes place through an elastic medium by means of wave motion. A wave is the motion through the medium of a disturbance as distinguished from the motion of the medium as a whole. An obvious example is a surface wave on water. Most sound waves of importance are transmitted compressional disturbances, that is, disturbances in which the pressure or density at any point in the medium is caused to vary from its equilibrium value. The change is then propagated through the medium with a velocity which depends on the type of wave in question, the nature of the medium, and the temperature. For example, the velocity of sound in still air of normal atmospheric composition at 0°C (32°F) is 331.45 m/s (1087.43 ft/s). Another important property of sound transmission is intensity, measured by the average rate of flow of energy in the wave per unit of time and per unit area perpendicular to the direction of propagation. The intensity of all practical sound waves diminishes with distance from the source, a property known as attenuation. This is due either to the spreading of the sound-wave energy over larger and larger surfaces or to actual absorption of the energy by the medium with transformation into heat.

The most important sound waves are harmonic waves, defined as waves for which the propagated disturbance at any point in its path varies sinusoidally with time with a definite frequency or number of complete cycles per second (the unit being the hertz). For a harmonic wave the disturbance at any instant repeats itself in magnitude and phase at intervals along the direction of propagation equal to

the so-called wavelength, which is, therefore, equal to the sound velocity divided by the frequency.

Acoustics deals with waves of all frequencies, but not all frequencies are audible by human beings, for whom the average range of audibility extends from 20 to 20,000 Hz. Sound below 20 Hz is referred to as infrasonic, and that above 20,000 Hz is called ultrasonic.

Sound detection. The detection of sound is made possible by the incidence of transmitted sound energy or an appropriate acoustic transducer. For human beings with so-called normal hearing, the most important transducer is the ear, a remarkably sensitive organ able to detect a sound intensity as low as 10^{-16} W/cm^2. For modern applied acoustics, transducers such as the microphone, based on the piezoelectric effect, are widely used. Generally speaking, any transducer used as a source of sound is also available as a detector, though the sensitivity varies considerably with the type.

Applications. The practical applications of acoustics are multifarious. They include architectural acoustics, or the study of sound waves in closed rooms arranged for the satisfactory production and reception of speech and music; underwater acoustics, dealing specifically with all aspects of sound in the sea and its use for the detection of vessels and the exploration of the sea bed; engineering acoustics, including the whole technology of sound reproduction and recording, sound motion pictures, and radio and television, as well as the study of vibrations of solids and their control and the use of high-intensity ultrasonics in industrial processing (for example, in metallurgy). Noise control is also an important case of engineering acoustics. Further examples are provided by physiological and psychological acoustics, dealing with hearing in humans and animals. Communication acoustics considers the production and transmission of speech. Musical acoustics deals with the physics of musical instruments. Bioacoustics and medical acoustics take up the use of sound in medical diagnosis and therapy as well as its use in the study of the overall behavior of animals in general.

All applications of acoustics are based on the fundamental physical properties of sound waves. The extensive inderdisciplinary is exhibited in the **illustration**. *See* ACOUSTIC NOISE; ARCHITECTURAL ACOUSTICS; ATMOSPHERIC ACOUSTICS; LOUDSPEAKER; MECHANICAL VIBRATION; MICROPHONE; MUSICAL ACOUSTICS; PHYSIOLOGICAL ACOUSTICS; PSYCHOACOUSTICS; SOUND; SOUND ABSORPTION; SOUND RECORDING; SOUND-REINFORCEMENT SYSTEM; SOUND-REPRODUCING SYSTEMS; ULTRASONICS; UNDERWATER SOUND; VIBRATION; VIBRATION DAMPING; VIBRATION ISOLATION. R. Bruce Lindsay

Bibliography. D. T. Blackstock, *Fundamentals of Physical Acoustics*, 2000; F. Fahy, *Foundations of Engineering Acoustics*, 2001; L. E. Kinsler et al., *Fundamentals of Acoustics*, 4th ed., Wiley, 2000; A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications*, 1989.

Acoustooptics

The field of science and technology that is concerned with the diffraction of visible or infrared light (usually from a laser) by high-frequency sound in the frequency range of 50–2000 MHz. The term "acousto" is a historical misnomer; sound in this frequency range should properly be called ultrasonic. Such sound cannot be supported by air, but propagates as a mechanical wave disturbance in amorphous or crystalline solids, with a sound velocity ranging from 0.6 to 6 km/s (0.4 to 4 mi/s) and a wavelength from 3 to 100 μm. *See* LASER; ULTRASONICS.

The sound wave causes a displacement of the solid's molecules either in the direction of propagation (longitudinal wave) or perpendicular to it (shear wave). In either case, it sets up a corresponding wave of refractive-index variation through local dilatation or distortion of the solid medium. It is this wave that diffracts the light by acting as a three-dimensional grating, analogous to x-diffraction in crystals. The fact that the grating is moving is responsible for shifting the frequency of the diffracted light through the Doppler effect. *See* DIFFRACTION GRATING; DOPPLER EFFECT; REFRACTION OF WAVES; WAVE MOTION; X-RAY DIFFRACTION.

Since the 1960s, acoustooptics has moved from a scientific curiosity to a relevant technology. This evolution was initially driven by the need for fast modulation and deflection of light beams, and later by demands for more general optical processing. It was made possible by the invention of lasers, the development of efficient ultrasonic transducers, and the formulation of realistic models of sound-light interaction. *See* TRANSDUCER.

Bragg diffraction. When the sound propagates in a very wide, parallel beam that approximates a plane wave, the analogy with x-ray diffraction is most pronounced. The phenomenon is then called Bragg diffraction and is conventionally illustrated by the wave-vector diagram (**Fig. 1**).

The diagram shows an isosceles triangle, the sides of which are formed by the wave vectors $\mathbf{k}_0$ and $\mathbf{k}_1$ of the incident and diffracted light, and $\mathbf{K}$ of the sound. The directions of the wave vectors show which way sound and light propagate, while their lengths are proportional to the momenta of the corresponding field quanta, that is, photons and phonons. In quantum-mechanical terms, the wave-vector diagram represents conservation of momentum in a photon-phonon collision. The incident light photon ($\mathbf{k}_0$) collides with a sound phonon ($\mathbf{K}$) to create a

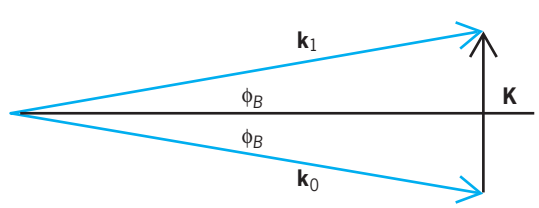

Fig. 1. Wave-vector diagram of acoustooptic interaction.

diffracted photon ($\mathbf{k}_1$). *See* CONSERVATION OF MOMENTUM; PHONON; PHOTON; QUANTUM MECHANICS.

Conservation of momentum in the collision requires that the arrows line up to form a closed triangle. This determines the incident angle ϕ_B, the Bragg angle. Its magnitude equals approximately $\lambda/2\Lambda$, where λ and Λ are the wavelengths of light and sound. Typical Bragg angles are small, ranging from 0.1 to 10°.

In addition to momentum, energy must be conserved in the collision. The energy of the colliding phonon is added to that of the incident photon to create the diffracted photon. As the energy of a quantum is proportional to its frequency, this process results in the frequency of the diffracted photon being upshifted by the sound frequency. This is the quantum-mechanical interpretation of the Doppler effect. *See* CONSERVATION OF ENERGY.

Bragg cell. In the wave-vector diagram, each arrow represents a plane wave carrying photons or phonons. In actuality, however, both sound and light propagate as beams of finite width rather than as infinitely wide plane waves. The finite size of the sound beam relaxes the requirement that the incident angle precisely equal the Bragg angle. In this way a tolerance about the Bragg angle is provided that makes practical acoustooptic devices possible, such as the Bragg cell (**Fig. 2**).

The acoustic beam of width L is generated by applying an electrical signal at frequency F to the acoustic transducer. The transducer is made of a thin layer of piezoelectric material, often crystalline, that generates mechanical vibrations when activated electrically. *See* PIEZOELECTRICITY.

Frequency shifting. The incident light beam (called the zeroth order) has a frequency f, a wavelength λ, and a width d. It is incident at an angle ϕ_B as required by the wave-vector diagram. The diffracted beam (called the first order) propagates at the deflection angle $\phi_d = 2\phi_B$. This beam has a frequency $f + F$, where F is the frequency of the acoustic signal. Thus the first obvious application of a Bragg cell is as a frequency shifter. The frequency F can be varied within limits without having to change the incident angle, because of the tolerance about the Bragg angle. *See* OPTICAL ISOLATOR.

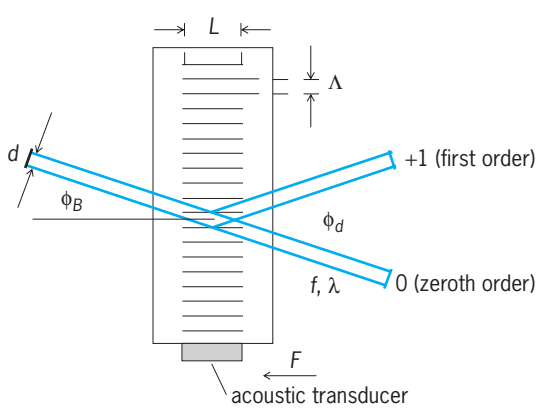

Fig. 2. Basic acoustooptic Bragg cell.

Light modulation. A second application is a light modulator. Here use is made of the fact that the intensity of the diffracted light is proportional to the intensity of the sound wave. Thus, if the electrical signal at frequency F is modulated in amplitude, the result is a similar modulation of the diffracted light. *See* OPTICAL MODULATORS.

Light deflection. Another application of the Bragg cell is as an electronic light deflector. Varying the sound frequency changes the sound wavelength and hence the value of the Bragg angle. This in turn changes the deflection angle in proportion. (Again, because of the tolerance about the Bragg angle, the input angle need not be changed.) Thus, by sweeping the acoustic frequency, the Bragg cell can be made to produce a scanning, diffracted beam of light. For random-access applications the frequency change is discontinuous so as to direct the diffracted beam quickly in any desired direction.

It is customary to focus the diffracted beam upon a focal plane by use of a suitable lens. When the beam scans, the focused spot traces out a line. The number of resolvable spots, N, that can be addressed along this line is given by the product of the maximum frequency swing $\Delta F_{\max}$ and the transit time τ of the sound through the light beam. Typical values are $\Delta F_{\max} = 40$ MHz, $\tau = 10\ \mu$s, $N = 400$. The transit time τ also determines the shortest time in which the beam can change direction.

Spectrum analysis. If a spectrum of discrete frequencies, $F_1, F_2, F_3, \ldots$, is applied simultaneously to the transducer, each frequency produces its own diffracted beam and hence its own spot in the focal plane. Thus, in yet another application, the Bragg cell acts as a spectrum analyzer by making the spectral distribution of the sound signal visible in the focal plane. The obtainable frequency resolution equals $1/\tau$. *See* SPECTRUM ANALYZER.

Imaging. If the light beam is wide enough to fill the sound cell, a substantial portion of the sound signal is "addressed" by the incident light. The cell then generates in diffracted light a "moving image" of the amplitude modulation of the sound wave. Such an image may be further optically processed to enhance key features or to extract specific information. Alternatively, if the amplitude modulation represents successive lines of a television image, that image may be projected upon a screen, using appropriate dynamic optics to immobilize it. Such an acoustooptic television display was in fact developed in 1936 in Great Britain, long before the invention of the laser.

Width of sound beam. The efficiency of a Bragg cell increases by increasing the width L of the sound beam. Unfortunately the tolerance about the Bragg angle (and therefore the acoustic frequency range and the acceptance angle) are inversely proportional to this length. The design of a Bragg cell is thus a compromise between efficiency and bandwidth. Furthermore, if the width of the sound beam is too small, the beam does not resemble a plane wave, and undesirable higher-order light beams are generated by multiple diffraction. The cell is then said to work in

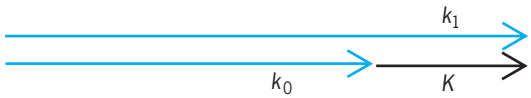

Fig. 3. Wave-vector diagram showing collinear interaction in an anisotropic crystal.

the Raman-Nath region (named after two of the first investigators of the phenomenon), rather than the Bragg region.

Optical filters. An important application of the Bragg cell is to electronically tunable optical filters. Here a narrow tolerance about the Bragg angle is required to minimize the window of allowable incident wavelengths clustered around a central one. When the sound frequency is changed, a different central wavelength will satisfy the Bragg angle condition, and the window moves along the optical spectrum.

Use of anisotropic crystals. In isotropic materials a tunable acoustooptic filter requires a wide sound beam, which is not always feasible. In practice, anisotropic crystals are used for this application. In such crystals the light velocity depends upon direction of propagation and polarization. The wave-vector diagram (Fig. 2) must then be modified to account for different polarizations of incident and diffracted light. This results in the light wave vectors $\mathbf{k}_0$ and $\mathbf{k}_1$ having unequal lengths. By collapsing the wave-vector triangle to three parallel lines (**Fig. 3**), a configuration is obtained in which the sound and the light propagate in the same direction. This collinear interaction maximizes the interaction length and obviates the need for a wide sound beam. *See* CRYSTAL OPTICS; POLARIZED LIGHT.

Another advantage of using anisotropic crystals is that some of them have sound velocities an order of magnitude smaller than isotropic materials such as glass or fused quartz. (In tellurium dioxide, TeO_2, for example, one of the possible shear waves has a sound velocity of 617 m/s or 2024 ft/s). Since the efficiency of a Bragg cell—often summarized by a material figure of merit called M_2—is inversely proportional to the cube of the sound velocity, it is clear that crystals such as tellurium dioxide offer potential design advantages.

Two-dimensional devices. It has proven possible to construct two-dimensional acoustooptic configurations. Here, acoustic surface waves interact with light propagating in optical waveguides made of thin films, deposited on a suitable substrate. Such devices are of obvious importance when combined with integrated electronic circuits. *See* INTEGRATED CIRCUITS; INTEGRATED OPTICS; LIGHT; SOUND.

Adrian Korpel

Bibliography. A. P. Goutzoulis and D. R. Pape (eds.), *Design and Fabrication of Acousto-optic Devices*, Marcel Dekker, New York, 1994; A. Korpel, *Acousto-optics*, 2d ed., Marcel Dekker, New York, 1997; A. Korpel (ed.), *Selected Papers on Acousto-optics*, SPIE, Bellingham, 1990; A. Sliwinski, B. B. J. Linde, and P. Kwiek (eds.), Acousto-optics and applications, *Proc. SPIE*, vol. 3581, 1998.

Acquired immune deficiency syndrome (AIDS)

A viral disease of humans caused by the human immunodeficiency virus (HIV), which attacks and compromises the body's immune system. Individuals infected with HIV proceed through a spectrum of stages that ultimately lead to the critical end point, acquired immune deficiency syndrome. The disease is characterized by a profound progressive irreversible depletion of T-helper-inducer lymphocytes (CD4+ lymphocytes), which leads to the onset of multiple and recurrent opportunistic infections by other viruses, fungi, bacteria, and protozoa (see **table**), as well as various tumors (Kaposi's sarcoma, lymphomas). HIV infection is transmitted by sexual intercourse (heterosexual and homosexual), by blood and blood products, and perinatally from infected mother to child (prepartum, intrapartum, and postpartum via breast milk).

HIV infection. In 1983, a virus recovered from the blood of a patient with persistent generalized lymphadenopathy attacked and killed T lymphocytes but could not be well characterized because of the inability to grow it in sufficient amounts. In 1984, T cells were identified that supported growth of the virus and were able to at least partially withstand its cytopathic effects. This retrovirus was variously named LAV (lymphadenopathy-associated virus) and HTLV-III (human T-cell lymphotropic virus, type III), but is now universally named HIV-1 (human immunodeficiency virus, type 1). *See* RETROVIRUS.

The availability of large quantities of virus and the resultant diagnostic reagents led to the finding that HIV-1 was the primary etiological agent of this syndrome. A majority of those infected will ultimately develop AIDS. Virtually all individuals with AIDS, as well as those with the generalized

Agents of major opportunistic infections in patients with AIDS

Type	Agent
Viruses	*Cytomegalovirus*
	Herpes simplex
	Herpes zoster
	Papovavirus (progressive multifocal leukoencephalopathy)
Fungi	*Candida* species
	Cryptococcus neoformans
	Histoplasma capsulatum
	Coccidioides immitis
Bacteria	*Mycobacterium avium* complex
	Mycobacterium tuberculosis
	Streptococcus pneumoniae
	Haemophilus influenzae
	Salmonella typhimurium
	Shigella flexneri
	Campylobacter species
	Treponema pallidum
	Rochalimaea henselae (bacillary angiomatosis)
Protozoa	*Pneumocystis carinii*
	Toxoplasma gondii
	Isospora belli
	Cryptosporidia
	Microsporidia

lymphadenopathy syndrome, tested seropositive for HIV-1. This important breakthrough also permitted the introduction of serological testing of blood donors, vastly decreasing the incidence of transfusion-associated HIV infection. The time interval from initial HIV infection until HIV antibodies can be detected by using commercially available assays ranges from 3 weeks to 6 months; however, HIV infection can precede the appearance of clinical disease by 10 years or more. Serological testing allows for the detection of viral infection before any clinical manifestations appear.

Since retroviruses such as HIV-1 integrate their genetic material into that of the host cell, infection is generally lifelong and cannot be eliminated easily. Therefore, medical efforts have been directed toward preventing the spread of virus from infected individuals. The best method for preventing transmission of HIV infection is to avoid high-risk behaviors, such as unprotected sexual intercourse or intravenous drug abuse.

Clinical disease. Approximately 50–70% of individuals with HIV infection experience an acute mononucleosis-like syndrome approximately 3–6 weeks following primary infection. In the acute HIV syndrome, symptoms include fever, pharyngitis, lymphadenopathy, headache, arthralgias, myalgias, lethargy, anorexia, nausea, and erythematous maculopapular rash. These symptoms usually persist for 1–2 weeks and gradually subside as an immune response to HIV is generated.

Asymptomatic stage. Although the length of time from initial infection to development of the clinical disease varies greatly from individual to individual, a median time of approximately 10 years has been documented for homosexual or bisexual men, depending somewhat on the mode of infection. Intravenous drug users experience a more aggressive course than homosexual men and hemophiliacs because their immune systems have already been compromised and the virus is acting on a weakened system.

Early symptomatic disease. As HIV replication continues, the immunologic function of the HIV-infected individual declines throughout the period of clinical latency. At some point during that decline (usually after the CD4+ lymphocyte count has fallen below 500 cells per microliter), the individual begins to develop signs and symptoms of clinical illness, and sometimes may demonstrate generalized symptoms of lymphadenopathy, oral lesions (thrush, hairy leukoplakia, aphthous ulcers), herpes zoster (shingles), and thrombocytopenia.

Secondary opportunistic infections. Secondary opportunistic infections are a late complication of HIV infection, usually occurring in individuals with less than 200 CD4+ lymphocytes per microliter. They are characteristically caused by opportunistic organisms such as *Pneumocystis carinii* and cytomegalovirus that do not ordinarily cause disease in individuals with a normally functioning immune system. However, the spectrum of serious secondary infections that may be associated with HIV infection also includes common bacterial pathogens, such as *Streptococcus pneumoniae*. Secondary opportunistic infections are the leading cause of morbidity and mortality in persons with HIV infection. Therefore, HIV-infected individuals are administered protective vaccines (pneumococcal) as well as prophylactic regimens for the prevention of infections with *P. carinii, Mycobacterium tuberculosis*, and *M. avium* complex. *See* MYCOBACTERIAL DISEASES; PNEUMOCOCCUS; STREPTOCOCCUS.

Tuberculosis has become a major problem for HIV-infected individuals. Those with pulmonary tuberculosis cannot be distinguished from those with community-acquired pneumonias, as well as *P. carinii* pneumonia. Any HIV-infected individual with pneumonia should be isolated immediately from other persons and health care staff members until a diagnosis of tuberculosis is ruled out. Health care settings should implement the use of negative-pressure sputum induction booths, as well as negative-pressure isolation rooms. Further complicating the dual epidemic of HIV disease and tuberculosis has been the emergence of multidrug-resistant tuberculosis, with its therapeutic and prophylactic dilemmas. *See* DRUG RESISTANCE; OPPORTUNISTIC INFECTIONS; TUBERCULOSIS.

Treatments. Antiretroviral treatment with deoxyribonucleic acid (DNA) precursor analogs—for example, azidothymidine (AZT), dideoxyinosine (ddI), and dideoxycytidine (ddC)—has been shown to inhibit HIV infection by misincorporating the DNA precursor analogs into viral DNA by the viral DNA polymerase. Nevertheless, these agents are not curative and do not completely eradicate the HIV infection.

Abe M. Macher; Eric P. Goosby

Acquired immunological tolerance

An induced state in which antigens originally regarded as foreign become regarded as self by the immune system. Tolerance can be induced (tolerization) in all of the cells of the immune system, including T cells (also known as T lymphocytes), the antibody-forming B cells (also known as B lymphocytes), and natural killer cells. Artificially induced immunological tolerance can be helpful in a number of clinical settings.

Autoimmune disease. The major function of the normal mammalian immune system is distinguishing foreign antigens (usually derived from microorganisms) from self-antigens so that only foreign antigens will be the targets of immune attack. To ensure the ability to make this distinction, developing T cells and B cells normally undergo selection processes based on the antigen receptors they express. Thereby, only cells with potential reactivity to foreign antigens, and not those with the ability to recognize self-antigens (autoantigens), contribute to the lymphocyte repertoire. However, this selection is incomplete, and additional mechanisms are required to help prevent autoreactive lymphocytes from mounting harmful antiself immune responses (termed autoimmunity). Regulatory processes, however, sometimes go awry, resulting in autoimmune disease, in which a variety of tissues (for example,

the insulin-producing beta cells of the pancreas in diabetes and the thyroid gland in autoimmune thyroiditis) may be targeted. *See* ANTIBODY; ANTIGEN; CELLULAR IMMUNOLOGY; IMMUNITY.

Inducing self-tolerance in the immune system could be an approach to curing autoimmune diseases. Some autoimmune diseases are believed to be initiated by autoreactive T lymphocytes. These may also be associated with autoantibody production by B cells that are dependent on help from autoreactive T cells. For these diseases, tolerization of T cells should, in theory, be sufficient to correct the disease. Other autoimmune diseases may be B cell–autonomous, so tolerization of autoreactive B cells would be essential for an effective cure. *See* AUTOIMMUNITY; IMMUNOLOGICAL DEFICIENCY.

Organ transplantation. Tolerization can also be used to facilate organ transplantation. Immunosuppressive drug therapy, which globally suppresses immune responses, is normally taken continuously by organ transplant recipients in an effort to prevent rejection in the graft. Despite improvements in immunosuppressive drug therapy, teaching the immune system to regard a set of foreign antigens presented by the organ graft as self (that is, tolerance induction) has become an important goal for several reasons: (1) It would eliminate the need for chronic immunosuppressive therapy, which is associated with lifelong increased risks of infection and malignancy, and other side effects. (2) It would prevent chronic rejection (a major problem even with immunosuppressive therapy), which often leads to late graft loss. (3) It presents a less toxic alternative to the unacceptably high levels of nonspecific immunosuppressive therapy that would likely be required to prevent rejection of xenografts (grafts from a donor of another species).

With the exception of blood group antigens, which are usually matched before transplanting organs between allogeneic donors and recipients (donors and recipients of the same species), most individuals do not have preexisting natural antibodies against organs from other humans. Since natural killer cells do not appear to make a significant contribution to the rejection of allogeneic organs (other than bone marrow), induction of T cell tolerance would be sufficient to overcome the problems of allotransplantation. In contrast, T cell–, natural killer cell–, and B cell–mediated immune responses to xenografts are generally greater than those to allografts, so in such cases the induction of B cell– and natural killer cell tolerance is of interest in addition to T cell tolerance. For example, although pigs are considered the most promising xenogeneic species for organ donation to humans, they express a carbohydrate antigen [Galα1,3Gal (Gal)] on most of their tissues that is recognized by natural antibodies present in the sera of all humans. These antibodies can bind to the Gal antigen expressed on the cells lining the blood vessels of pig organs and initiate a very rapid rejection process called hyperacute rejection. Although studies in primates indicate that hyperacute rejection can be prevented by removing the antibodies from the recipient serum prior to transplant, the antibodies rapidly return to the circulation and cause other, more delayed forms of rejection. Tolerization of the B cells that produce these natural antibodies presents a possible solution to this problem. Natural killer cells seem to respond more strongly to xenografts than to allografts, perhaps because xenogeneic tissues do not express molecules that can be recognized by inhibitory receptors on natural killer cells that normally prevent the natural killer cells from killing autologous cells. Although little is known about the process by which natural killer cells are educated to distinguish self from nonself, bone marrow transplantation studies in mice suggest that acquired tolerance can be induced among natural killer cells. *See* TRANSPLANTATION (BIOLOGY).

Central tolerance. Many strategies for inducing immunological tolerance involve reproducing the mechanisms involved in natural central tolerance—the phenomenon by which self-tolerance is maintained among immature lymphocytes developing in the central lymphoid organs. For developing T cells, tolerance occurs in the thymus, the central organ for T cell development. For B cells, development occurs in the bone marrow, and an encounter with self antigens can induce tolerance there among the immature cells. *See* IMMUNOLOGICAL ONTOGENY.

Clonal deletion. Clonal deletion (or negative selection) is a major natural mechanism of central tolerance. Developing lymphocytes with uniquely rearranged receptors (T cell receptor in the case of T cells; immunoglobulin receptor in the case of B cells) that recognize self antigens may be selected against if they encounter these self antigens on an appropriate cell in the thymic (for T cells) or marrow (for B cells) environment. Developing T cells encountering such antigens will undergo apoptotic death (programmed cell death). Developing B cells in the marrow or immature B cells in the spleen that recognize self-antigens may die or undergo additional gene rearrangements that produce new immunoglobulin receptors with different recognition specificities (receptor editing).

Hematopoietic cell transplantation. The transplantation of bone marrow or other sources of hematopoietic (blood cell–producing) stem cells provides a very powerful means of inducing T cell central tolerance. One reason is that cells derived from the donor stem cells (for example, dendritic cells) enter the thymus and present donor antigens to developing thymocytes, which are killed if they bind strongly to an antigen. If a constant supply of these donor cells is provided by a successful hematopoietic stem cell graft, T cells developing subsequently will be tolerized to the marrow donor by the same mechanism through which self-tolerance is normally achieved (death of alloreactive cells). The clinical potential of bone marrow or other types of hematopoietic cell transplantation for the induction of transplantation tolerance in humans has not yet been realized, however, because the methods traditionally used to prepare a recipient for a hematopoietic cell transplantation (for the treatment of leukemia, for example) are far too toxic to be appropriate in this setting.

In addition, major problems due to host-versus-graft and graft-versus-host immune reactivity occur with extensively mismatched marrow transplants.

More recently, animal studies have indicated that less toxic but more specific treatments to target the cells responsible for this immune reactivity can allow hematopoietic cell transplantation to be performed with minimal toxicity. Since the host marrow is not subjected to ablative treatments (that is, treatments that eliminate host hematopoietic cells), a state of mixed chimerism (the coexistence of blood cells derived from both the donor and the host) results when donor hematopoietic stem cell engraftment is achieved. Thus, hematopoietic cell transplantation may have clinical applicability for the induction of allograft tolerance. *See* CHIMERA.

Hematopoietic cell transplantation has been shown in animal models to induce T cell central tolerance even across xenogeneic barriers when the donor and host species are very closely related. T cells are made tolerant to the donor by similar mechanisms to those described above for allogeneic marrow. Natural antibody-producing B cells (both preexisting and newly developing cells) are made tolerant to the donor antigens when mixed chimerism is achieved. Therefore, successful achievement of marrow engraftment across highly disparate species barriers has the potential to prevent both T cell– and natural antibody–mediated rejection of xenografts. In the most disparate species combinations, however, nonimmunological barriers such as species differences in the proteins important for the function of donor hematopoietic stem cell grafts make it difficult to achieve engraftment. Donor cells may also be destroyed by cells of the innate immune system (for example, macrophages).

Peripheral tolerance. Peripheral tolerance comprises mechanisms to prevent immune responses among mature lymphocytes in the peripheral tissues.

Anergy. One major mechanism of T cell and B cell peripheral tolerance is anergy, in which the cells cannot be fully activated by encounter with the antigens that their receptors recognize. Numerous methods of inducing T and B cell anergy have been described. For B cells, the induction of anergy versus activation might be dependent on antigen concentration, with exposure to low antigen levels leading to an anergic state. However, T cells may be anergized by an encounter with antigen in the absence of costimulatory signals from the antigen-presenting cell that are required for T cell activation, or by encounter with an antigenic variant for which it has low affinity. Many experimental approaches to inducing transplantation tolerance, such as the use of costimulatory blockade, may induce a state of anergy among donor-reactive T cells. Costimulatory blockade may also lead to peripheral deletion of T cells encountering alloantigens (antigens that stimulate responses in another member of the some species) of a graft.

Suppression. Another major mechanism of peripheral tolerance is suppression, in which both B cells and T cells may be rendered tolerant of a specific antigen through the activity of substances or cells that actively suppress the lymphocyte's function. Numerous means of inducing B cell and T cell suppression have been described, and convincing evidence implicates cells with the ability to suppress T cell alloresponses (immune responses to alloantigens) in transplantation models. Recently, T cells expressing the molecules CD4 and CD25 have been implicated in the tolerance induced with costimulatory blockade and other strategies used in rodent models. However, although a variety of mechanisms resulting in suppression have been described, the mechanisms operative in vivo are incompletely understood. A better understanding of these complex phenomena may lead to the ability to apply some of these strategies from animal models to clinical transplantation. *See* IMMUNOSUPPRESSION.

Megan Sykes

Bibliography. W. E. Paul (ed.), *Fundamental Immunology*, Lippincott-Raven, Philadelphia, 1999; M. Sykes and S. Strober, Mechanisms of tolerance, in E. D. Thomas, K. G. Blume, and S. J. Forman (eds.), *Hematopoietic Cell Transplantation*, pp. 264–286, Blackwell Science, Malden, 1999.

Acrothoracica

An order of the subclass Cirripedia. All members burrow into shells of mollusks and thoracican barnacles, echinoderm tests, polyzoans, dead coral, and limestone. The entrance to the burrow is slitlike. In thin shells the burrow and animal are strongly compressed (see **illus.**). The mantle lacks calcareous

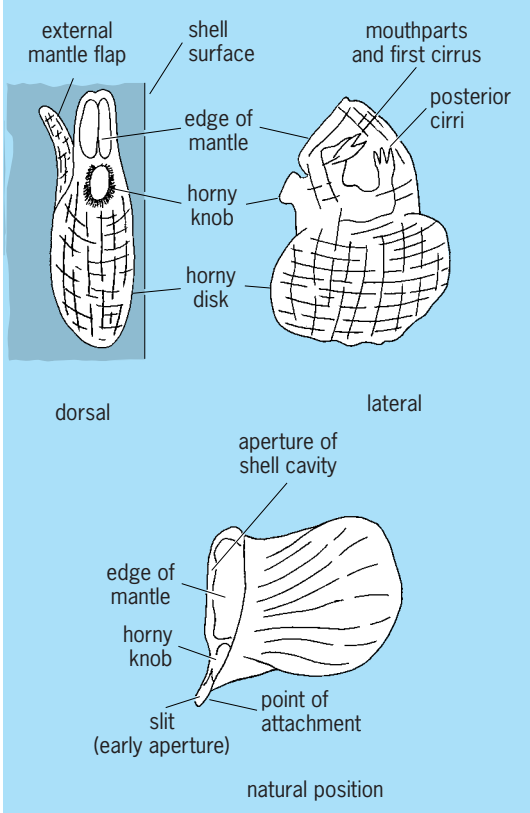

***Trypetesa lateralis* Tomlinson. (*After J. T. Tomlinson, A burrowing barnacle of the genus Trypetesa, order Acrothoracica, J. Wash. Acad. Sci., 43(11):373–381, 1953*)**

plates and is attached to the burrow by cement from glands in the dorsal part of the mantle. The attached area becomes thickened by the buildup of successive layers of chitin which cannot be shed at molting. The normal three pairs of cirriped mouthparts are present: mandibles, maxillules, and maxillae. Four to six pairs of cirri occur, often greatly reduced in size, the first pair close to the oral appendages and the remainder packed together at the posterior end of the body. The sexes are separate; dwarf males are found on the mantle or wall of the burrow of the female. Naupliar larval stages may be omitted, but a cypris larva always ccurs in the life cycle. Three families, about eight genera, and 40 species are recognized. Detailed examination of shell debris, especially from warmer seas, has revealed many new species. Doubtless many more exist. Fossils, known from burrows only, occur from the Carboniferous to Recent. *See* CIRRIPEDIA. H. G. Stubbings

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; J. T. Tomlinson, The burrowing barnacles (Cirripedia: Order Acrothoracica), *Bull. U.S. Nat. Mus.*, 296:1–162, 1969.

Acrotretida

An order of inarticulated brachiopods consisting of a group of sessile, suspension-feeding, marine, benthic, epifaunal bivalves, with representatives occurring throughout the Phanerozoic Era (from the Early Cambrian to the present). Extant members of this group originated in the Ordovician Period, when the order achieved maximum diversity.

Most shells are circular or subcircular in outline. The dorsal (brachial) valve is usually larger than the ventral (pedicle) valve, and is convex and often conical. The ventral valve is flattened and disklike. The valves are separate and do not articulate about a hinge but possess a complex arrangement of muscles. Acrotretids have two pairs of anterior and posterior adductor muscles, two pairs of oblique muscles, an elevator, and three pairs of minor muscles: the lophophore protractors, retractors, and elevators.

The tentacular feeding organ (lophophore) occupies the mantle cavity. The digestive system consists of a mouth, pharynx, esophagus, stomach, digestive diverticula, pylorus, and intestine, terminating in a functional anus. Four digestive diverticula in *Discinisca* and two in *Neocrania* open through ducts to the stomach. The excretory system consists of a pair of ciliated funnels (metanephridia) which during spawning act as gonoducts and allow the discharge of gametes from the coelom into the mantle cavity. Some solid waste may also be ejected through the nephridiopores, enmeshed in mucus, while the main excretory product, ammonia, is voided through the tissues of the mantle and lophophore.

Acrotretids possess an open circulatory system of blood vessels and coelomic canals, containing a fluid that is coagulable and has free cellular inclusions consisting of blood cells and coelomocytes. Acrotretids have a central nervous system containing unsheathed nerves, but no differentiated sense organs. Members of this order have been reported only as gonochoristic.

Two suborders are recognized: (1) Acrotretidina have a short pedicle emerging through a foramen in the ventral valve, chitinophosphatic shells, and planktotrophic larval development. (2) Craniidina lack a pedicle and are cemented to the substrate by the ventral valve, have shells composed of calcite and scleroproteins, and produce lecithotrophic larvae. *See* BRACHIOPODA; INARTICULATA. Mark A. James

Bibliography. G. A. Cooper, *Jurassic Brachiopods of Saudi Arabia*, 1989; M. J. S. Rudwick, *Living and Fossil Brachiopods*, 1970.

Acrylonitrile

An explosive, poisonous, flammable liquid, boiling at 77.3°C (171°F), partly soluble in water. The formula $CH_2{=}CH{-}C{\equiv}N$ indicates it may be regarded as vinyl cyanide, and its systematic name is 2-propenonitrile. Acrylonitrile is prepared by ammoxidation of propylene, according to the reaction below, over various sorts of catalysts, chiefly metallic

$$2CH_3{=}CHCH_3 + 2NH_3 + 3O_2 \longrightarrow 2CH_2{=}CHC{\equiv}N + 6H_2O$$

oxides. Older processes, such as catalytic addition of hydrogen cyanide to acetylene, and the catalytic reaction of propylene with nitric oxide, are little used today.

Most of the acrylonitrile produced is consumed in the manufacture of acrylic and modacrylic fibers. Substantial quantities are used in acrylonitrile-butadiene-styrene (ABS) resins, in nitrile elastomers, and in the synthesis of adiponitrile by electrodimerization. The adiponitrile is subsequently hydrogenated to hexamethylenediamine, a constituent of nylon. Smaller amounts of acrylonitrile are used in cyanoethylation reactions, in the synthesis of drugs, dyestuffs, and pesticides, and as co-monomers with vinyl acetate, vinylpyridine, and similar monomers.

Acrylonitrile undergoes spontaneous polymerization, often with explosive force. It polymerizes violently in the presence of suitable alkaline substances. *See* NITRILE; POLYMERIZATION. Frank Wagner

Bibliography. R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; P. A. Smith, *Open-Chain Nitrogen Compounds*; *Chemistry of Non-Cyclic Nitrogen-Containing Organic Functional Groups*, 2d ed., 1982.

Actiniaria

An order of the cnidarian subclass Zoantharia (Hexacorallia) known as the sea anemones, the most widely distributed of the anthozoans. They have even been discovered in frigid waters. Usually they are solitary animals which live under the low-tide mark attached to some solid object by a basal expansion

or pedal disk. They feed on prey such as copepods, mollusks, annelids, crustaceans, and fishes. The burrowing species, like *Edwardsia*, *Halcampella*, and *Harenactis*, lack a pedal disk and bury their elongated bodies in the soft sediment of oceans. The actinians move rather actively. *Gonactinia* and *Boloceroides* can swim by using their tentacles.

Morphology. The freely retractile, skeletonless polyp has a cylindrical body, with a thick, tough, rough column wall often bearing rugae, verrucae, tubercles, or suckers (**Fig. 1**). The body is often encrusted with sand grains, pebbles, and other detritus. *Actinia*, *Diadumene*, *Metridium*, and other species have smooth, thin walls. The acrorhagi or marginal sphaerules, small rounded bodies covered with nematocysts, are arranged in a circlet on the margin which is the junction between the oral disk and column. When lacking nematocysts, they are termed fronds or pseudoacrorhagi. Nematocysts discharge a toxic substance, tetramethylammonium hydroxide or tetramine; however, the human skin is seldom affected by this. The junction between the pedal disk and column is termed the limbus. In *Tealia*, *Anemonia*, *Phellia*, *Ilyanthus*, and many other species, the upper part of the column is folded to form a collar or parapet, which divides the body into an upper

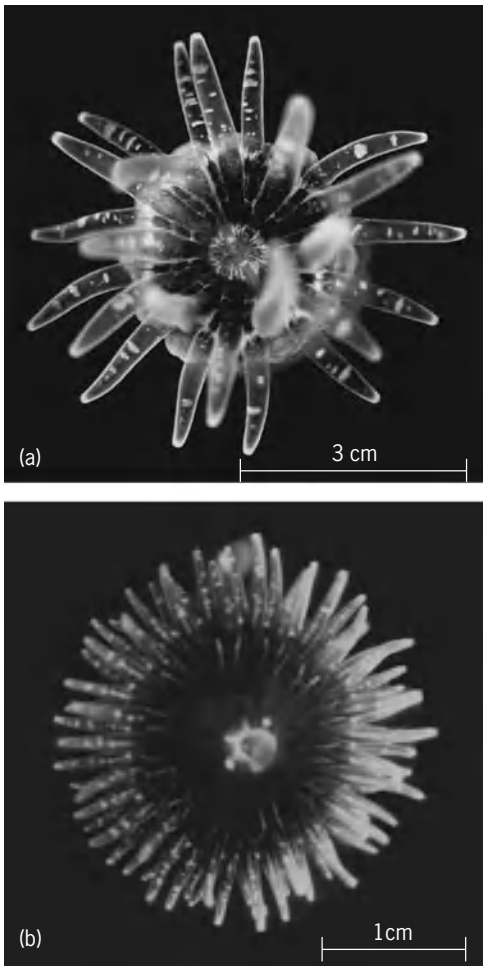

Fig. 1. Sea anemones, oral view. *Anthopleura* sp. (*a*) Young. (*b*) Adult.

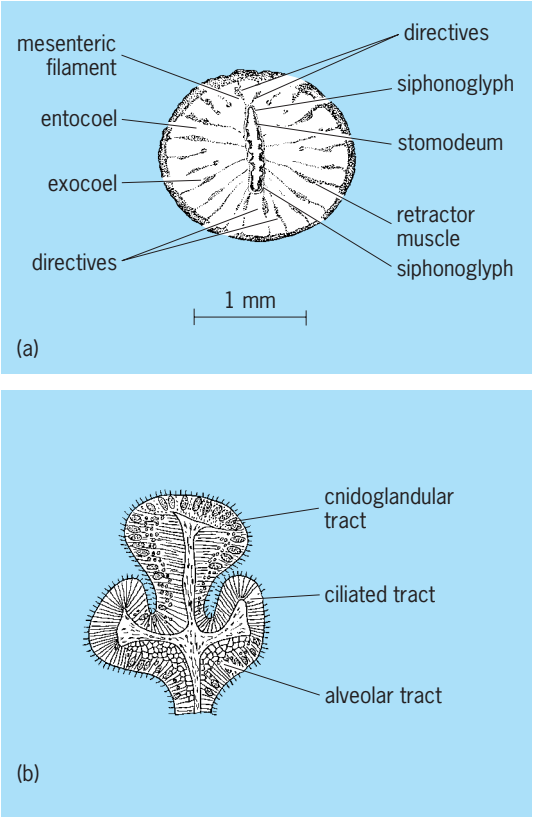

Fig. 2. Some of the morphological features of the actinians. (*a*) Cross-sectional view of *Anthopleura* sp. (*b*) Three-tracted mesenteric filament of *Phymanthus crucifer.*

capitulum and a lower scapus. Normally anemones have two siphonologlyphs, the sulcus and sulculus, and are termed diglyphic (**Fig. 2***a*); however, monoglyphic and occasionally tri- or tetraglyphic species such as *Diadumene* and *Metridium* occur. The colors vary with species and many variations occur even among the same species. *Diadumene luciae* has been divided into four races according to differences in color.

The tentacles increase in number regularly and are arranged in several cycles. There are 6 primary, 6 secondary, 12 tertiary, 24 quaternary, and so forth in the hexameroustype (Fig. 1). The paired mesenteries appearing first in couples show a bilateral arrangement. Their retractor muscles face each other, except the sulcal and asulcal pairs or directives which face away from each other. The space between each set of the paired mesenteries is an entocoele while its neighboring space is an exocoele, in which additional pairs of mesenteries are added (Fig. 2*a*). All the spaces communicate with each other either by oral or labial stomata; marginal or parietal ones, or both, may be present. The mesenteric filament is composed of three tracts, the cnidoglandular, ciliated, and alveolar (Fig. 2*b*). In Diadumenidae, Metridiidae, and Sagartiidae whitish threads or acontia protrude from small pores or cinclides of the column. These acontia are continuations of the septal filaments.

The musculature is the most highly developed in

the coelenterates. Generally, the longitudinal muscle is ectodermal, and the circular muscle, endodermal. The principal muscles are the longitudinal retractors, which form characteristic muscle bands; circular muscles of the column; marginal and tentacular sphincters, which serve to contract the body, oral disk, and tentacles respectively; the basilar muscle in the pedal disk; and, near this, the parietobasilar.

Reproduction. Most actinians are dioecious. The egg is comparatively large and contains a great amount of yolk. It is covered with numerous short processes (**Fig. 3***a*). The spermatozoon has a long flagellum (Fig. 3*b*). Developing larvae (**Fig. 4**) pass through the *Edwardsia* stage with eight complete mesenteries, then the *Halcampoides* stage with six pairs of protocnemes (primary mesenteries). Longitudinal fission frequently occurs as well as budding (**Fig. 5***a*, *b*, and *c*). Sometimes new individuals result from laceration (Fig. 5*d* and *e*). *See* CNIDARIA; REPRODUCTION (ANIMAL); HEXACORALLIA. Kenji Atoda

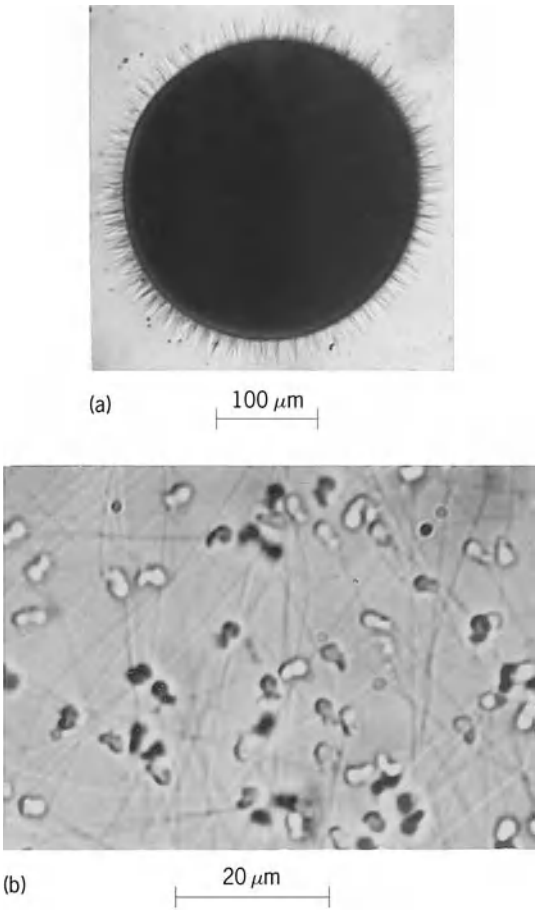

Fig. 3. Gametes of actinians. (*a*) Egg of *Anthopleura stella*. (*b*) Spermatozoa of *A. xanthogrammica*.

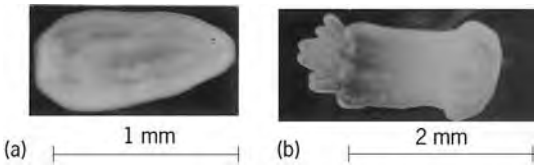

Fig. 4. Larval stages of *Anthopleura* sp. (*a*) Ciliated larva. (*b*) Larva just after extrusion.

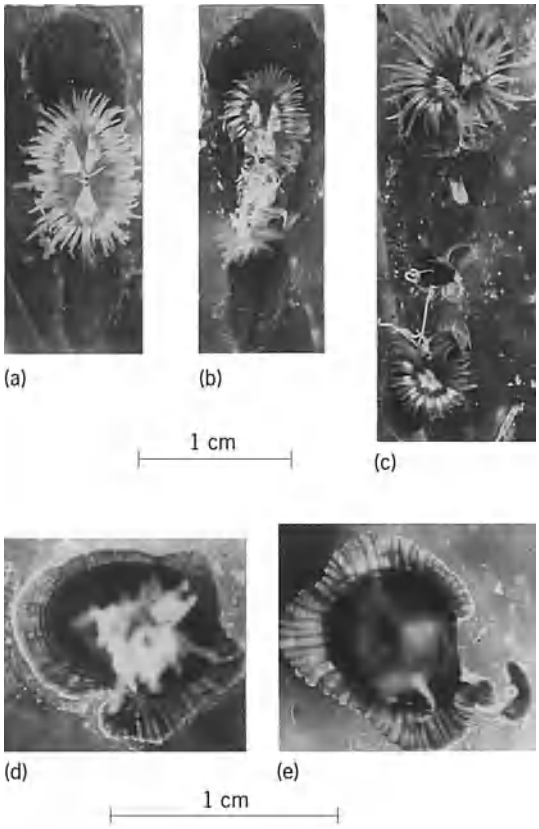

Fig. 5. Division of the pedal disk in *Diodumene luciae*. (*a*) Beginning of fission. (*b*) After 9½ h. (*c*) Polyp dividing into three unequal portions. (*d*) Beginning of laceration. (*e*) A small piece separated.

Bibliography. J. M. Shick, *A Functional Biology of Sea Anemones*, Chapman and Hall, 1991.

Actinide elements

The series of elements beginning with actinium (atomic number 89) and including thorium, protactinium, uranium, and the transuranium elements through the element lawrencium (atomic number 103). These elements, chemically similar, have a strong chemical resemblance to the lanthanide, or rare-earth, elements of atomic numbers 57 to 71. Their atomic numbers, names, and chemical symbols are 89, actinium (Ac), the prototype element, sometimes not included as an actual member of the actinide series; 90, thorium (Th); 91, protactinium (Pa); 92, uranium (U); 93, neptunium (Np); 94, plutonium (Pu); 95, americium (Am); 96, curium (Cm); 97, berkelium (Bk); 98, californium (Cf); 99, einsteinium (Es); 100, fermium (Fm); 101, mendelevium (Md); 102, nobelium (No); 103, lawrencium (Lr).

Studies of the chemical and physical properties of actinide and lanthanide elements and their compounds have indicated that their electronic structure must be similar; an inner electron shell of fourteen 5*f* electrons in the case of the actinides, and fourteen

Oxidation states of actinide elements in aqueous solution

Oxidation state		Elements*
II		Cf, Es, Fm, Md, No
III		Ac, Pa, U, Np, Pu, Am, Cm, Bk, Cf, Es, Fm, Md, No, Lr
IV		Th, Pa, U, Np, Pu, Bk
V		Pa
V	(as MO_2^{+} ions)	U, Np, Pu, Am
VI	(as MO_2^{2+} ions)	U, Np, Pu, Am
VII	(as MO_5^{3-} ions)	Np, Pu

*The most stable states are underscored.

4f electrons in the case of the lanthanides, is filled in progressing across the series. Except for thorium and uranium, the actinide elements are not present in nature in appreciable quantities. The transuranium elements were discovered and investigated as a result of their synthesis in nuclear reactions. All are radioactive, and except for thorium and uranium, weighable amounts must be handled with special precautions.

The uranium isotopes ^{233}U and ^{235}U and the plutonium isotope ^{239}Pu undergo nuclear fission with slow neutrons with the liberation of large amounts of energy. Thorium can be converted to ^{233}U, and the isotope ^{238}U to ^{239}Pu by neutron irradiation; hence thorium and natural uranium can be used indirectly as nuclear fuels in breeder reactors.

Ion-exchange chromatography has been an important experimental technique in the study of the chemistry of the actinide elements. This method together with the analogy between corresponding actinides and lanthanides was the key to the discovery of the transcurium elements.

The actinide elements are very similar chemically. Most have the following in common: trivalent cations which form complex ions and organic chelates; soluble sulfates, nitrates, halides, perchlorates, and sulfides; and acid-insoluble fluorides and oxalates.

The characteristic oxidation state in aqueous solution is the III state, with higher oxidation states prominent in the early and the lower II state appearing in the latter part of the series (see **table**).

Many solid compounds including hydrides, oxides, and halides have been prepared by dry chemical methods. Binary compounds with carbon, nitrogen, silicon, and sulfur are of interest because of their stability at high temperatures. *See* ACTINIUM; AMERICIUM; ATOMIC STRUCTURE AND SPECTRA; BERKELIUM; CALIFORNIUM; CURIUM; EINSTEINIUM; FERMIUM; LAWRENCIUM; MENDELEVIUM; NEPTUNIUM; NOBELIUM; PERIODIC TABLE; PLUTONIUM; PROTACTINIUM; THORIUM; TRANSURANIUM ELEMENTS; URANIUM.

Glenn T. Seaborg

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; S. Cotton, *Lanthanide and Actinide Chemistry*, 2d ed., 2006; S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; J. J. Katz, G. T. Seaborg, and L. R. Morse (eds.), *The Chemistry of the Actinide Elements*, 2 vols., 2d ed., 1986; G. Meyer and L. R. Morse, *Synthesis of Lanthanide and Actinide Compounds*, 1991.

Actinium

A chemical element, Ac, atomic number 89, and atomic weight 227.0. Actinium was discovered by A. Debierne in 1899. Milligram quantities of the element are available by irradiation of radium in a nuclear reactor. Actinium-227 is a beta-emitting element whose half-life is 22 years. Six other radioisotopes with half-lives ranging from 10 days to less than 1 minute have been identified.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

The relationship of actinium to the element lanthanum, the prototype rare earth, is striking. In every case, the actnium compound can be prepared by the method used to form the corresponding lanthanum compound with which it is isomorphous in the solid, anhydrous state. *See* ACTINIDE ELEMENTS; LANTHANUM; NUCLEAR REACTION; RADIOACTIVITY.

Sherman Fried

Bibliography. S. Cotton, *Lanthanide and Actinide Chemistry*, 2d ed., 2006; J. J. Katz, G. T. Seaborg, and L. R. Morse (eds.), *The Chemistry of the Actinide Elements*, 2 vols., 2d ed., 1986; G. Meyer and L. R. Morse, *Synthesis of Lanthanide and Actinide Compounds*, 1991.

Actinobacillus

A genus of gram-negative, immotile and nonspore-forming, oval to rod-shaped, often pleomorphic bacteria which occur as parasites or pathogens in mammals (including humans), birds, and reptiles. They are facultatively aerobic, capable of fermenting carbohydrates (without production of gas) and of reducing nitrates. Most species are oxidase- and catalase-positive. Some cultures tend to stick on the surface of agar media, particularly on primary isolation. The genome deoxyribonucleic acid contains between 40 and 47 mol % guanine plus cytosine. The actinobacillus group shares many biological properties with the genus *Pasteurella*. At least two of the following features or combinations of features differentiate members of the *Actinobacillus* group

from *Pasteurella*: hemolysis, delayed or lacking fermentation of D-galactose or D-mannose, fermentation of inositol, positive reactions for both urease and β-galactosidase, hydrolysis of salicin or esculin, and fermentation of maltose together with negative reactions for trehalose fermentation and ornithine decarboxylase. *See* PASTEURELLA.

Actinobacillus lignieresii, the type species, causes granulomatous lesions predominantly in the upper alimentary tract of cattle (in particular, "wooden tongue"), and putrid skin and lung lesions in sheep. *Actinobacillus equuli* causes suppurative lesions, mainly in the kidneys and joints of foals. *Actinobacillus suis* is the etiologic agent of various lesions in piglets, and *A. pleuropneumoniae* that of epizootic necrotizing pleuropneumoniae in swine. *Actinobacillus capsulatus* was isolated from joint disease in rabbits. *Actinobacillus* (*Pasteurella*) *ureae* and *A. hominis* occur in the respiratory tract of healthy humans and may be involved in the pathogenesis of sinusitis, bronchopneumonia, pleural empyema, and meningitis.

Actinobacillus actinomycetemcomitans occurs in the human oral microflora, and together with anaerobic or capnophilic organisms may cause endocarditis and suppurative lesions in the upper alimentary tract. A specific role of *A. actinomycetemcomitans* in the pathogenesis of localized juvenile periodontitis has been suggested. *Pasteurella pneumotropica, P. haemolytica, Haemophilus paragallinarum*, and a variety of other groups of organisms that occur in healthy or diseased mammals, birds, and reptiles have been located with, or in the close vicinity of, the genus *Actinobacillus* on the basis of genetic relatedness.

Actinobacilli are susceptible to most antibiotics of the β-lactam family, aminoglycosides, tetracyclines, chloramphenicol, and many other antibacterial chemotherapeutics. *See* ANTIBIOTIC; MEDICAL BACTERIOLOGY. Walter Mannheim

Bibliography. R. G. E. Murray, N. R. Krieg, and J. G. Holt (eds.), *Bergey's Manual of Systematic Bacteriology,* vol. 1, 1984; C. C. Tsai et al., Serum neutralizing activity against *Actinobacillus actinomycetemcomitans* leukotoxin in juvenile periodontitis, *J. Clin. Periodont.,* 8:338–348, 1981.

Actinomycetes

A heterogeneous collection of bacteria with diverse ecologies and physiologies. The initial definition was based on their branching filamentous cellular morphology during a stage of the growth cycle. The name actinomycete means "ray fungus," and for many years the actinomycetes were erroneously considered to be fungi, or at least closely related to the fungi. True fungi, such as the common bread mold, also grow as branching filaments (hyphae) that are several micrometers in diameter, whereas those of the actinomycetes are about 1 μm in diameter. The genetic material of the eukaryotic true fungi consists of DNA associated with proteins, and is enclosed in a membrane (the nucleus), whereas the genetic material of the prokaryotic actinomycetes consists of free DNA in direct contact with the cell sap (cytoplasm). In general, actinomycetes do not have membrane-bound organelles, whereas the fungi have an array of organelles such as mitochondria and vacuoles. The cell walls of fungi are made of chitin or cellulose, whereas the cell walls of the actinomycetes are made of a cross-linked polymer containing short chains of amino acids and long chains of amino sugars. Actinomycetes are susceptible to a wide range of antibiotics used to treat bacterial diseases, such as penicillin and tetracycline. *See* AMINO ACIDS; AMINO SUGAR; ANTIBIOTIC; DEOXYRIBONUCLEIC ACID (DNA); FUNGI.

Diversity and distribution. The actinomycetes occupy virtually all environmental niches, with some general tendency for members of one genus to have a more limited distribution. Not unexpected for such a diverse collection of prokaryotic microorganisms, the boundaries of the actinomycetes are not sharply defined. To add to the difficulty in selecting delineating characteristics, the genus *Actinomyces* is, in most respects, an exceptional taxon within the actinomycetes, or Actinomycetales. The actinomycetes include plant pathogens, animal pathogens, commensals in the intestines of invertebrates, and free-living bacteria in soil and in fresh and marine water. Only a few actinomycetes are pathogenic for plants, but *Streptomyces scabies* is one of the exceptions. For the most part, the aerobic *Streptomyces* are found in soil with a pH range of 5–6.5 with 10% organic matter. *Micromonospora* and *Streptosporangium* are found in soil with a pH of 4–5, with an organic content of 4–7%. *Nocardia* and *Actinomadura* are best isolated from soil with a pH of 4–5 and an organic content of 10%. The unique *Actinoplanes* with motile spores are found in aquatic habitats. The microaerophilic *Actinomyces* tend to colonize mammals and are often found in the mouth and gastrointestinal tract. When displaced from their normal sites within the mouth or gastrointestinal tract, *Actinomyces* may cause diseases in humans such as lung abscesses and appendicitis, and lumpy jaw in cattle. The animal pathogens are distributed throughout several genera, including *Mycobacterium*, *Corynebacterium*, *Nocardia*, *Actinomadura*, and *Actinomyces* (see **illustration**). *See* BACTERIA; BACTERIAL GROWTH; MEDICAL BACTERIOLOGY; PATHOGEN; SOIL MICROBIOLOGY.

Metabolism. The actinomycetes display a wide range of metabolic activities. Physiologically the actinomycetes have extraordinary ability to break down polysaccharides and hydrocarbons. In general, actinomycetes grow more slowly than "true" bacteria such as *Bacillus*, enteric bacteria, and *Pseudomonas*, in part because they are best adapted to the temperature of soil and water (70°F or 21°C). Members of the genus *Actinomyces* do not require oxygen for growth and are sometimes referred to as anaerobic bacteria. It is actually a requirement for elevated levels of carbon dioxide rather than a

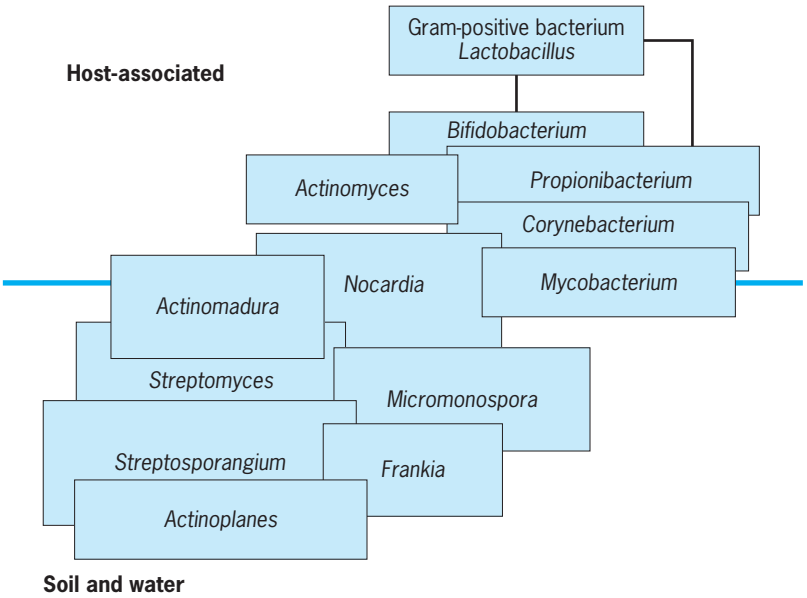

Relationships among some better-known actinomycete genera. The actinomycetes are gram-positive bacteria that bridge through *Bifidiobacterium* and *Propionibacterium* to *Lactobacillus*. The genera above the colored line tend to live in association with hosts, but infrequently as pathogens. The genera below the colored line are found in soil and water.

negative effect of oxygen. *Actinomyces* is cultivated in the laboratory at body temperature (98°F or 37°C) in a rich nutrient medium, and frequently in an atmosphere that is enriched with carbon dioxide. *Actinomyces* is moderately slow-growing, requiring 1–2 weeks to achieve visible growth. Members of the genus *Nocardia* require oxygen for growth and have the ability to digest and transform many organic chemicals, including petroleum products and industrial products, intermediates, and by-products. *Nocardia* strains have been used commercially to modify chemical compounds related to cholesterol to make biologically active steroids. *Nocardia* inhaled into the lungs from soil may cause a disease in humans resembling tuberculosis. The streptomycetes are aerobic bacteria that have the ability to thrive with a wide range of nitrogen and carbon sources, including cellulose. The aroma of fresh soil and newly dug potatoes is actually due to streptomycetes. Streptomycetes do not produce disease in humans and are best known for their prolific production of antibiotics. Members of the genus *Frankia* symbiotically fix atmospheric nitrogen in the root nodules of the alder tree (*Alnus rugosa*) and the bayberry shrub (*Myrica pennsylvanica*). *See* TUBERCULOSIS.

Identification and classification. The actinomycetes were brought together as a group based upon morphological considerations, and morphological and developmental characteristics remain important in identification and classification. Members of the genus *Actinomyces* initially grow as branching filaments and then fragment into short rods. *Micromonospora* produces solitary spores in the vegetative hyphae. Streptomycetes persist as filamentous vegetative hyphae and form chains of spores in aerial hyphae. *Nocardia* does not produce aerial hyphae. The hyphae of some *Nocardia* species form branching filaments only for a brief time, whereas other species have persistent long, branching vegetative hyphae. There are differences in the chemical composition of actinomycete cells, especially the cell walls, which can be characterized by staining or chemical analysis. *Mycobacterium* cells are difficult to stain, and the films on microscope slides must be heated for the dye to penetrate. Once stained, the mycobacterium cells tenaciously retain the dye, and even treatment with dilute acid does not remove the dye; hence the name acid-fast bacteria. The actinomycete genera differ in the amounts of meso-diaminopimelic acid, arabinose, galactose, and long-chain mycolic acids in their cell walls. Streptomycete cell walls contain LL-diaminopimelic acid, whereas those of *Nocardia* contain meso- and/or DD-diaminopimelic acid. Analysis of menaquinones and lipid profiles of cell extracts have been used to demonstrate relatedness between actinomycete strains. For example, dihydromenaquinones are found in *Corynebacterium* and *Mycobacterium*, whereas tetrahydromenaquinones are found in *Nocardia*. Analysis of various nucleic acid preparations has become a widely used and accepted means for identifying and demonstrating relationships among actinomycetes. Initially, the guanine and cytosine (G+C) content of DNA was used to establish unrelatedness but could not establish relatedness. The ability of denatured DNA to reanneal became the standard method for measuring relatedness. Currently one of the popular nucleic acid techniques involves the analysis of the 16S fraction of ribosomes. As the capability to sequence the nucleotides in DNA samples has improved, direct comparison of genome sequences has become the definitive methodology. *See* NUCLEIC ACID; NUCLEOTIDE.

Molecular biology. Genetically the actinomycetes have DNA with high proportions of the purine guanine and the pyrimidine cytosine (55–72% G+C), and have rather large genome sizes (three to eight megabases). The high G+C content of the DNA affords some degree of protection of the genome for those species living in the soil and water. The smaller genomes of *Corynebacterium*, *Mycobacterium*, and *Rhodococcus* are organized into circular chromosomes, whereas the larger chromosomes of *Streptomyces* and *Nocardia asteroids* are usually linear. The *Streptomyces coelicolor* genome contains 8,667,507 base pairs, organized into a linear chromosome containing an estimated 7825 genes. The central core of the *S. coelicolor* chromosome contains considerable similarity to the whole chromosome of *Mycobacterium tuberculosis* and *Corynebacterium diphtheriae*. The actinomycetes have the ability to transfer genetic material among closely related strains by several different mechanisms, including conjugation and plasmid transfer. There are limited data indicating that plasmid transfer of genetic material occurs in nature as well as the laboratory. Methodology for the transfer of genes to or from actinomycetes to taxonomically distant bacteria, such as *Escherichia coli*, has been accomplished. There is growing evidence

that limited gene transfer has occurred between actinomycetes and eukaryotic organisms. *See* CHROMOSOME. S. Gaylen Bradley

Bibliography. S. Donadio et al., Microbial technologies for the discovery of novel bioactive metabolites, *J. Biotech.*, 99:187–198, 2002; T. Kieser et al., *Practical Streptomyces Genetics*, John Innes Foundation, Norfolk, UK, 2000; E. M. Miguelez, C. Hardisson, and M. B. Manzanal, Streptomycetes: A new model to study cell death, *Int. Microbiol.*, 3:153–158, 2000.

Actinomyxida

An order of the protozoan class Myxosporidea (subphylum Cnidospora). It is characterized by the production of trivalved spores with three polar capsules and one to many sporoplasms. The spore membrane may be extended into anchor-shaped processes, which may have bifurcate tips. These protozoan parasites are found in the body cavity or in the intestinal lining of fresh-water annelids and in marine worms of the phylum Sipunculoidea. *See* INVERTEBRATE PATHOLOGY.

The life cycle for most species is not well known. Uninucleate sporoplasms may pair and fuse (plasmogamy) to form binucleate amebulas. Each amebula may produce a number of small and large cells through repeated binary fission. A large and small cell will pair and fuse (anisogamy) into a single cell (zygote) which becomes the sporoblast. The sporoblast divides repeatedly by binary fission, producing cells that will form the spore. Thus, in *Triactinomyxon legeri* (see **illus.**) parasitic in the gut of tubificid annelids, three cells form the polar capsules, three form the valves of the membrane, and one cell becomes the sporoplasm. The nucleus of the sporoplasm divides repeatedly by mitosis until 27 nuclei are formed, of which 24 become the sporoplasmic nuclei and 3 become the residual somatic nuclei. *See* CNIDOSPORA; MYXOSPORIDEA; OLIGOCHAETA; PROTOZOA. Ross F. Nigrelli

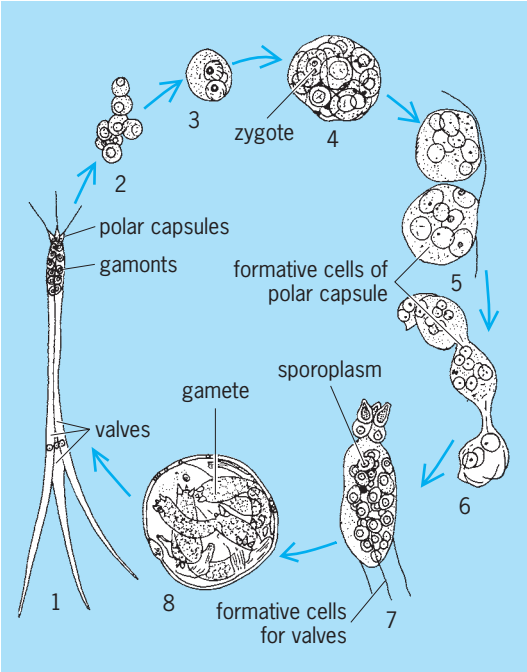

Life cycle of *Triactinomyxon:* 1, mature spore; 2, liberated gamonts; 3, gamonts pairing; 4, zygote formation; 5, spore formation from zygote; 6–7, later stages in spore formation; 8, cyst in host tissue filled with young spores.

Actinophage

Any of a number of bacteriophages that infect and lyse certain bacteria (as originally classified in the order Actinomycetales; see **table**). Actinophages of particular interest are those that include in their host range any organisms of the genus *Streptomyces,* since most of the therapeutically useful antibiotics require culture of the particular *Streptomyces* species on a large scale. Contamination of a culture with a specific actinophage may result in lysis and destruction of the bacteria, which obviously halts antibiotic production. The elimination of actinophages from the environment is sometimes difficult, and the problem of phage contamination is often solved by the isolation and use of a phase-resistant mutant of the antibiotic-producing bacterium. *See* ANTIBIOTIC; BACTERIOPHAGE; LYSOGENY. Lane Barksdale

Host ranges of some actinophages*

Actinophage	Sensitive organism†		
	Streptomyces griseus	*Nocardia braziliensis*	*Mycobacterium rhodocrous*
MNP 6	+	+	−
MSP 1	−	+	−
MNP 8	−	−	+

*After L. A. Jones and S. G. Bradley, *Mycologia*, 56:505–513, 1964.
†Here + = lysed by bacteriophage, − = not lysed.

Actinophryida

An order of Heliozoia. These protozoans are characterized by stiff arms radiating from the central cell mass. Each arm is supported internally by an array of microtubules which terminate on the surface of nuclei. The organisms are found in fresh-water, marine, and soil habitats. They feed on other protozoans which collide with and stick to the arms. Well-studied genera are *Actinophrys* (**illus.** *a*), with a single central nucleus, *Actinosphaerium*, and *Echinosphaerium* (illus. *b*); the latter two have many nuclei lying just internal to a peripheral layer of vacuoles. The body size ranges from 50 to several hundred micrometers. Actinophryids form cysts within which a type of uniparental sexual reproduction

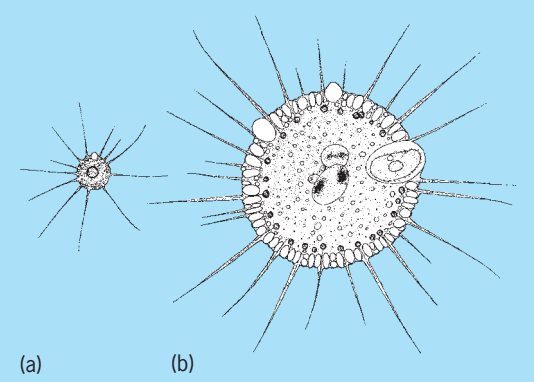

Actinophryida. (*a*) *Actinophrys sol*. (*b*) The much larger *Echinosphaerium nucleofilum*.

occurs. *See* ACTINOPODEA; HELIOZOIA; PROTOZOA; SARCODINA. David J. Patterson

Bibliography. J. J. Lee, S. H. Hutner, and E. C. Bovee, *Illustrated Guide to the Protozoa*, Society of Protozoologists, 1984; D. J. Patterson, On the organization and classification of the protozoon *Actinophrys sol* Ehrenberg, 1830, *Microbios*, 26:165–208, 1979.

Actinopodea

A class of Sarcodina; or in some modern schemes a superclass (Actinopoda) within the phylum Sarcomastigophora. Some of these protozoans have more or less permanent pseudopodia, composed of a central shaft of microtubules surrounded by a thin cytoplasmic envelope (axopodia); others have delicate and often radially arranged filopodia (filamentous, flexible pseudopodia) or filoreticulopodia (filamentous, branched pseudopodia) without axial microtubules. Although some are stalked and attached to the substratum (sessile), most are floating types. There are four subclasses: Radiolaria (Polycystinea and Phaeodarea), Acantharia, Heliozoia, and Proteomyxidia.

Radiolaria are all marine planktonic protozoans with a characteristic morphology consisting of a central capsule surrounded by a perforated capsular wall enclosed by the frothy or weblike mass of extracapsular cytoplasm. Solitary and colonial (multicellular) forms occur in some groups, and many species possess a characteristic, often highly ornamented, siliceous skeleton that persists after death of the protozoon and contributes to the sedimentary deposits in the ocean floor. The skeleton may be spherical, composed of concentric lattice shells, dome-shaped, or consist of a variety of solid geometric shapes, often with an esthetically pleasing architecture. Some species are surface-dwelling, and others occur from near surface to great depths in the ocean. *See* RADIOLARIA.

Acantharia, marine planktonic organisms resembling the Radiolaria, possess a central capsule, but the skeleton is composed of strontium sulfate. The skeleton is composed of rods arranged in a strict dimensional pattern, usually following the plan of a set of cartesian axes. In some species the skeleton is augmented with latticelike plates. The skeletal rods originate typically from a central point of attachment within the capsule and radiate outward through the capsular wall. Upon death of the acantharian, the strontium sulfate skeleton dissolves; hence no record is deposited in the ocean sediments. Algae, presumably symbionts, have been observed in the extracapsular cytoplasm or sometimes within the capsule of some species. Axopodia are characteristic. An apparently contractile apparatus (myoneme) attaches the cytoplasmic sheath to the skeleton and may regulate the shape of the peripheral cytoplasm, changing the hydrostatic properties of the organism and adjusting its buoyancy. *See* ACANTHAREA; ACANTHOMETRIDA.

Heliozoia include both fresh-water and marine organisms. The pseudopodia may be either filopodia or axopodia radially arranged around the central cytoplasm. Skeletal elements are either absent or simpler than observed in Acantharia or Radiolaria, and may consist of scales, spines, or in certain genera a perforate continuous exoskeleton (test). A centroplast (microtubular organizing center) is present in some genera, but lacking in others. Reproduction as exemplified by *Actinophrys* and *Actinosphaerium* is by fusion of gametes produced by the mother cell (autogamy). A zygocyst is formed and gives rise to the mature floating radiate form bearing radial pseudopodia. In *Actinophrys sol*, the sexual process begins with formation of a cyst followed by division within the cyst and formation of the haploid gametes that fuse to yield the zygote. *See* HELIOZOIA.

Proteomyxidia lack skeletons, but may gather detritus or mineral matter around the central cell body to form a loose protective coat. Filopodia or filoreticulopodia are characteristic and emerge either from one or both poles of the central body, which in some forms becomes spindle-shaped, especially during fission. Food of some species in the form of bacteria or small detrital particles is gathered by pseudopodial streaming from the surrounding environment. A number of species invade algae and, rarely, higher plants. During periods unfavorable for growth, some species may form resting cysts which resume active growth when the environment improves. *See* PROTEOMYXIDIA.

Little is known about the reproduction of many Actinopodea. *See* PROTOZOA; SARCODINA; SARCOMASTIGOPHORA. O. Roger Anderson

Actinopterygii

A class of teleostom fishes commonly known as the ray-finned fishes comprising the subclasses Cladistia (Polypteriformes and fossil orders), Chondrostei (Acipenseriformes and fossil orders), and Neopterygii (the remaining actinopterygian orders). The Neopterygii, minus the Lepisosteiformes,

Amiiformes, and several fossil taxa, comprise the Teleostei (see **illustration**). The Actinopterygii comprise about half of all vertebrate species and about 96% of all currently existing fishes. It is considered a nonmonophyletic group that is derived from more than one lineage when tetrapods are excluded. The Actinopterygii and the Sarcopterygii, minus the Tetrapoda, comprise what was previously called Osteichthyes (the bony fishes). The class includes 44 orders, 453 families, 4289 genera, and about 27,000 described extant species. Many species are known to science but are not yet described; further, in regions such as the Amazon and Congo basins, species are probably becoming extinct before they are discovered and described in the scientific literature.

Characteristics. Actinopterygii, being an extremely large and heterogenous group of fishes, are not identifiable by a unique character common to all taxa; suffice it to characterize the class as follows: The scales are dermal. Ganoid-type scales (scales covered by a hard mineral substance called ganoin or ganoine) are present in the Cladistia, Chondrostei, and holosteans (Lepisosteiformes, Amiiformes) and in several fossil orders. The Neopterygii (less the holosteans) have bony-ridge scales, which lack dense enamel or ganoin and are marked with bony ridges (circuli) that alternate with valleylike depressions. Bony-ridge scales are thin and overlapping and are typically cycloid or ctenoid, or highly modified as scutes, semirigid body rings (pipefishes and seahorses), rigid encasements resembling a turtle shell (trunkfishes and cowfishes), armature of bony plates (for example, the catfish families, Callichthyidae and Loricariidae; and poachers, Agonidae), and lancets (bladelike structures, one on each side of the caudal peduncle) of surgeonfishes (Acanthuridae). Some neopterygians lack scales entirely (for example, most catfishes and most eels). A sensory canal is present in the dentary bone, which is usually enclosed. Dermal bones of the skull are completely ossified in most groups. Radial bones of the pectoral girdle are attached to the scapulocoracoid complex. An interopercular bone and branchiostegal rays are usually present. Typically there is on each side one otolith in each of the three membranous sacs of the labyrinth of the inner ear. External nares (nostrils) are usually paired and end in a blind sac; internal nares are absent; and spiracles are usually absent.

About 40% of living actinopterygian species live exclusively or almost exclusively in freshwater. The rest inhabit mostly marine, brackish, or combination environments. The group has diversified through time to occupy an enormous range of habitats, from oceanic regions over 11,000 m (36,000 ft) deep (numerous species) to mountain streams up to 5200 m (17,000 ft) above sea level (homalopterid river loaches of Tibet), and from subfreezing water at −1.98°C (28.4°F) [Antarctic icefishes, Notothenioidei] to hot springs at 43.8°C (110.8°F) [pupfishes, *Cyprinodon*, and livebearers, *Gambusia*].

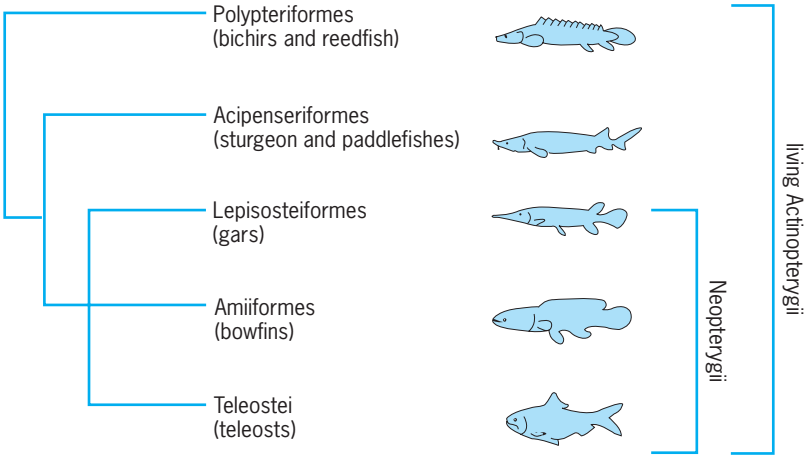

Tree illustrating some basal groups of living Actinopterygii. The arrangement reflects current understanding of their relationships, and the lack of clear resolution among certain neopterygian groups.

Some species move up on land to feed (mudskippers, *Periophthalmodon*), while others may take to the air by gliding on wind currents (flying fishes, Exocoetidae) or fly short distances using their large, well-muscled pectoral fins as wings (Gastropelecidae). While most species are herbivores (plant eaters) or piscivores (fish eaters), some are parasitic (candiru, Trichomycteridae) of South America. Some actinopterygians are ram-filter feeders (paddlefish, *Polyodon*), and some live within the internal organs of other organisms (pearlfishes, Carapidae) in possibly symbiotic, parasitic, or mutualistic relations.

Actinopterygians are extremely diverse morphologically and include the smallest known adult vertebrates [for example, a goby called the stout infant fish, *Schindleria brevipinguis*, from the Great Barrier Reef of Australia reaching only 7 mm (0.28 in.) in length]. The longest actinopterygian is the giant oarfish (*Regalecus*), which reaches reported lengths of over 8 m (26 ft).

Fossil record. The fossil record indicates that actinopterygians are at least as old as the Late Silurian (about 420 million years before present). Fossil actinopterygians are species-rich and extremely abundant, making up the majority of vertebrate fossils that are known by complete skeletons. Many major radiations of early actinopterygians, such as pycnodonts, semionotiforms, and palaeonisciforms, have been extinct for tens of millions of years. Other early actinopterygian groups, such as the Cheirolepiformes, have been extinct for hundreds of millions of years. Based on the fossil record, the most major differentiation of the group began in the late Mesozoic. *See* SEMIONOTIFORMS.

Living forms. Living Actinopterygii (see illustration) comprise Polypteriformes (bichirs and reedfishes, represented by at least 10 species, all in freshwaters of Africa); Acipenseriformes (sturgeons and paddlefishes, represented by at least 24 anadromous or freshwater species,

widespread within the Northern Hemisphere); Lepisosteiformes (gars, represented by at least seven species, mostly in freshwater but also rarely in brackish or even marine water, found in eastern North America, Central America, and Cuba); Amiiformes (bowfins, represented by one living species in the freshwaters of eastern North America); and Teleostei (teleosts, represented by about 24,000 species, widespread in nearly all aquatic environments, and including about 99.8% of all living actinopterygians). The gars, bowfins, and teleosts are grouped together in the Neopterygii. *See* ACIPENSERIFORMES; AMIIFORMES; POLYPTERIFORMES.

Phylogeny. There are two major theories to explain how the three basal extant lineages of Neopterygii are related. The older theory recognizes Holostei (a group containing gars and bowfins but excluding teleosts) as a natural (monophyletic) group. The other theory, resulting from detailed anatomical studies of the 1970s and 1980s, recognizes Halecostomi (a group containing bowfins and teleosts but excluding gars) as the natural group. Most systematic ichthyologists since the mid-1970s have favored recognition of Halecostomi over Holostei. However, with the advent of molecular evolutionary studies in the 1990s, the controversy has become even more complicated, because the most detailed morphological data now indicate Halecostomi as the natural group but molecular data favor Holostei. Thus, gars, bowfins, and teleosts are presented in the illustration as an unresolved trichotomous branch with inconclusive evidence for unambiguous phylogenetic resolution. *See* CHONDROSTEI; HOLOSTEI; OSTEICHTHYES; TELEOSTEI. Herbert Boschung; Lance Grande

Bibliography. L. Grande and W. E. Bemis, A Comprehensive Phylogenetic Study of Amiid Fishes (Amiidae) Based on Comparative Skeletal Anatomy: An Empirical Search for Interconnected Patterns of Natural History, *Soc. Vert. Paleontol. Mem.*, no. 4, suppl. to *J. Vert. Paleontol.*, vol. 18, no. 1, 1998; G. V. Lauder and K. Liem, The evolution and interrelationships of the actinopterygian fishes, *Bull. Mus. Comp. Zool.*, 150:95–107, 1983; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; J. R. Paxton and W. N. Eschmeyer (eds.), *Encyclopedia of Fishes*, University of New South Wales Press, 1994; M. Stiassny, L. Parenti, and G. D. Johnson (eds.), *Interrelationships of Fishes*, Academic Press, San Diego, 1996.

Action

Any one of a number of related integral quantities which serve as the basis for general formulations of the dynamics of both classical and quantum-mechanical systems. The term has been associated with four quantities: the fundamental action S, for general paths of a dynamical system; the classical action S_C, for the actual path; the modified action S', for paths restricted to a particular energy; and action variables, for periodic motions.

Least-action principles. A dynamical system can be described in terms of some number N of coordinate degrees of freedom that specify its configuration. As the vector q whose components are the degrees of freedom $q_1, q_2, \ldots, q_N$ varies with time t, it traces a path $q(t)$ in an N-dimensional space. The fundamental action S is the integral of the lagrangian of the system taken along any path $q(t)$, actual or virtual, starting from a specified configuration q_1 at a specified time t_1, and ending similarly at configuration q_2 and time t_2. The value of this action $S[q(t)]$ depends on the particular path $q(t)$. The actual path $q_C(t)$ which is traversed when the system moves according to newtonian classical mechanics gives an extremum value of S, usually a minimum, relative to the other paths. This is Hamilton's least-action principle. The extremum value depends only on the end points and is called the classical action $S_C(q_1, q_2; t_1, t_2)$. *See* DEGREE OF FREEDOM (MECHANICS); LAGRANGE'S EQUATIONS.

An important variant of Hamilton's principle applies when the virtual paths $q(t)$ are restricted to motions all of the same energy E, but no longer to a specific time interval, $t_1 - t_2$. The modified action $S' = S - E(t_1 - t_2)$ obeys a modified least-action principle, usually called Maupertuis' principle, namely, that the classical path gives again an extremal value of S' relative to all paths of that energy. Maupertuis' principle is closely related to Fermat's principle of least, time in classical optics for the path of light rays of a definite frequency through a region of inhomogeneous refractive index. *See* HAMILTON'S PRINCIPLE; LEAST-ACTION PRINCIPLE; MINIMAL PRINCIPLES.

Action variables. When some degree of freedom q_i executes a periodic or oscillatory motion at energy E, its contribution S_i to the action S' is given by Eq. (1),

$$S_i = \int p_i(q_i, E) dq_i \tag{1}$$

where p_i is the momentum conjugate to q_i. A new quantity, the action variable $J_i(E)$, can be defined for that degree of freedom, as the amount that the integral S_i increases for each cycle of motion of the coordinate q_i, that is, $J_i(E)$ is given by Eq. (2), where

$$J_i(E) = \oint p_i(q_i, E) dq_i \tag{2}$$

$\oint$ represents integration over one cycle of motion. Action variables have the important property of adiabatic invariance, namely, that they do not change during a process in which parameters λ, representing either external or internal features of the system, change sufficiently slowly, but not necessarily by a small amount, from λ to λ'. Although quantities, including the energy, change in general (say, from E to E'), J_i remains unchanged, so that $J_i(E; \lambda) = J_i(E'; \lambda')$. Two examples of such adiabatic invariants are those associated with a simple mechanical oscillator, and with a charged particle moving in a magnetic field.

Simple mechanical oscillator. For this case, λ may represent internal parameters such as the mass m or the

stiffness constant k. The single action variable $J(E)$ is the ratio $2\pi E/\omega$, where $\omega =$ is the angular frequency. The adiabatic invariance of J means that if the parameters change slowly enough that the relative change of frequency during each cycle is small ($\Delta\omega/\omega \ll 1$), the energy will change in proportion to the angular frequency. *See* HARMONIC OSCILLATOR.

Charged particle in magnetic field. For this case, there are several adiabatic invariants in situations where the field changes sufficiently slowly in space and time as to be nearly constant over distances of the order of the cyclotron orbit radius of the motion around the field lines. The principal adiabatic invariants are $J_\perp$ and $J_{//}$, which are associated with cyclotron-orbit motion perpendicular to the field line and the motion of the orbit "guiding center" along the field line, respectively. Their adiabatic invariance underlies the basic features of confinement of charged particles in magnetic-bottle field configurations such as the Van Allen belt in the Earth's magnetosphere. The value of $J_\perp$ is $(2\pi mc/e)\mu_z$, where μ_z is the orbital magnetic moment parallel to the field. This moment is diamagnetic, thus tending to confine the particles to weaker fields, say, away from the Earth's polar regions, where they then execute periodic motion along the field lines. Under time-varying magnetic fields, induction effects will accelerate the particles and change their energy. For slow enough rates of variation, the energy changes are determined by the condition of near-invariance of the longitudinal action variable $J_{//}[E(t); \lambda(t)]$. *See* DIAMAGNETISM; MAGNETIC MOMENT; MAGNETOSPHERE; PLASMA (PHYSICS); VAN ALLEN RADIATION.

Use in old quantum theory. In the early period of development of quantum theory, action variables were recognized as being candidates from within classical mechanics which could possibly be assigned quantized values. Their invariance under slowly changing external fields would then maintain the observed discreteness of energy levels in atoms. For example, the invariance of E/ω led to Eq. (3) for the energy levels

$$E_n = n\hbar\omega \tag{3}$$

of an oscillator in the so-called old quantum theory, where $\hbar$ is Planck's constant divided by 2π, and n is any positive integer.

Significance in quantum mechanics. In quantum mechanics, as originally formulated by E. Schrödinger, the state of particles is described by wave functions which obey the Schrödinger wave equation. States of definite energy in, say, atoms are described by stationary wave functions, which do not move in space. Nonstationary wave functions describe transitory processes such as the scattering of particles, in which the state changes. Both stationary and nonstationary state wave functions are determined, in principle, once the Schrödinger wave propagator (also called the Green function) between any two points q_1 and q_2 is known. In a fundamental restatement of quantum mechanics, R. Feynman showed that all paths from q_1 to q_2, including the virtual paths, contribute to the wave propagator. Each path contributes a complex phase-term $\exp i(\phi[q(t)])$, where the phase ϕ is proportional to the action, $S[q(t)]$, for that path, being given by Eq. (4). The resulting sum over paths, appropriately

$$\phi[q(t)] = \frac{S[q(t)]}{\hbar} \tag{4}$$

defined, is the path integral (or functional integral) representation of the Schrödinger wave propagator. The path integral has become the general starting point for most formulations of quantum theories of particles and fields. The classical path $q_C(t)$ of least action now plays the role in the wave function as being the path of stationary phase $\phi_{\text{stationary}}$, which is equal to $S_C(q_1, q_2; t_1, t_2)/\hbar$. *See* GREEN'S THEOREM; PROPAGATOR (FIELD THEORY).

Semiclassical approximation. The functional integration can be performed explicitly for only very few systems. An important case where it can be carried out explicitly, but approximately, is when large quantum fluctuations Δq of the paths $q(t)$ away from the extremal path $q_C(t)$, on the scale of $\hbar^{1/2}$, are neglected. This results in the very useful semiclassical approximation. The only surviving phase factor is now the classical stationary phase, so that the semiclassical approximation also establishes the classical least-action principle in the limit that $\hbar$ approaches 0, that is, the short-wavelength limit of the quantum theory.

WKB approximation. The semiclassical approximation for the propagator at a definite energy is similar, but is based on the phase $S'/\hbar$ and Maupertuis' least-action principle. Its most common use is for approximating stationary states of motions that are restricted to one coordinate dimension. Here it is known as the Wentzel-Kramers-Brillouin (WKB) approximation, and bound-state energies E_n are determined by the condition $J_i\ (E_n) = 2\pi(n + \delta)$. The factor δ represents a correction to the old quantum theory due to wavelike behavior at the end points of the motion. For the simple oscillator discussed above, δ has the value $^1/_2$ and changes the energies to those given by Eq. (5).

$$E_n = \left(n + {}^1/_2\right)\hbar\omega \tag{5}$$

See HAMILTON-JACOBI THEORY; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; WENTZEL-KRAMERS-BRILLOUIN METHOD.

Bernard Goodman

Bibliography. W. Dittrich and M. Reuter, *Classical and Quantum Dynamics: From Classical Paths to Path Integrals*, 3d ed., 2001; H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; R. J. Goldston and P. H. Rutherford, *Introduction to Plasma Physics*, 1995; L. D. Landau and E. M. Lifshitz, *Quantum Mechanics, Nonrelativistic Theory,* 3d ed., 1977; L. S. Schulman, *Techniques and Applications of Path Integration,* 1981, reprint 2005.

Activated carbon

A powdered, granular, or pelleted form of amorphous carbon characterized by very large surface area per unit volume because of an enormous number of fine pores. Activated carbon is capable of collecting gases, liquids, or dissolved substances on the surface of its pores. For many gases and liquids, the weight of adsorbed material approaches the weight of the carbon. *See* ADSORPTION.

Adsorption on activated carbon is selective, favoring nonpolar over polar substances and, in a homologous series, generally improving with increasing boiling point. Adsorption is also improved with increased pressure and reduced temperature. Reversal of the physical adsorptive conditions (temperature, pressure, or concentration) more or less completely regenerates the carbon's activity, and frequently allows recovery of both the carrier fluid and adsorbate. Compared with other commercial adsorbents, activated carbon has a broad spectrum of adsorptive activity, excellent physical and chemical stability, and ease of production from readily available, frequently waste materials.

Large-pore decolorizing carbons are used in liquid-phase work. Applications include improving the color of manufactured chemicals, oils, and fats, as well as controlling odor, taste, and color in potable water supplies, beverages, and some foods. Gas-adsorbent carbons are generally harder, higher-density, finer-pore types useful in gas separations, recovering solvent vapors, air conditioning, gas masks, respirators, carbon canisters in automobiles, and supporting metal salt catalysts, particularly in the production of vinyl-resin monomers.

Almost any carbonaceous raw material can be used for the manufacture of activated carbon. Wood, peat, and lignite are commonly used for the decolorizing materials. Bone char made by calcining bones is used in large quantity for sugar refining. Nut shells (particularly coconut), coal, petroleum coke, and other residues in either granular, briqueted, or pelleted form (see **illus.**) are used for adsorbent products.

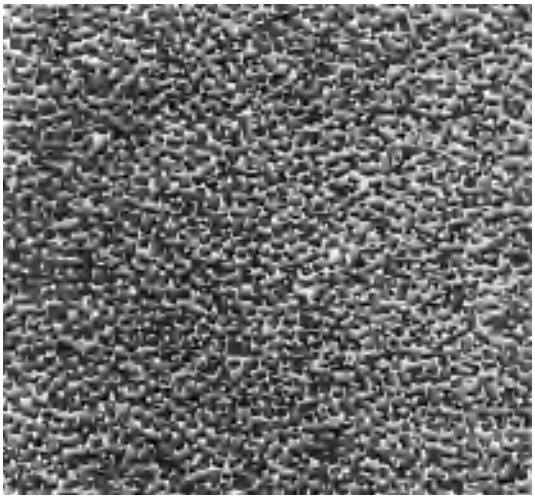

Activated carbon pellets, 4- to 6-mesh.

Activation. This is the process of treating the carbon to open an enormous number of pores in the 1.2- to 20-nanometer-diameter range (gas-adsorbent carbon) or up to 100-nanometer-diameter range (decolorizing carbons). After activation, the carbon has the large surface area (500–1500 m^2/g) responsible for the adsorption phenomena. Carbons that have not been subjected previously to high temperatures are easiest to activate. Selective oxidation of the base carbon with steam, carbon dioxide, flue gas, or air is one method of developing the pore structure. Other methods require the mixing of chemicals, such as metal chlorides (particularly zinc chloride) or sulfides or phosphates, potassium sulfide, potassium thiocyanate, or phosphoric acid, with the carbonaceous matter, followed by calcining and washing the residue. The economics of the latter process requires recovery of the chemical agent.

Tests. Tests to describe activated carbon's ability to perform are designed to simulate operating conditions. Carbon tetrachloride activity shows a gas-adsorbent carbon's capacity for vapors, and is the percentage by weight of carbon tetrachloride adsorbed at 25°C (77°F) from dry air saturated at 0°C (32°F). Retentivity, which correlates with a carbon's ability to remove low concentrations of vapor from a gas stream, is then determined by blowing dry air through the saturated carbon. Iodine activity indicates a carbon's ability to remove iodine from a standard stock solution and is used in specifications for liquid purification carbon. Quantitative liquid decolorizing evaluations are frequently determined by adding varying amounts of carbon to a standard series of solution aliquots and plotting (Freundlich isotherm) on logarithmic paper the concentration of adsorbate remaining in solution against the ratio of adsorbed material to weight of carbon. Minute service is applied to gas mask carbons, and represents the length of time during which a thin bed of activated carbon will completely adsorb chloropicrin gas. The break point occurs when penetration of gas through the bed activates a detection device downstream. Hardness, or strength, of the coarser activated carbons is calculated from the change in screen analysis experienced after mechanically abrading the carbon. Other quality tests, such as moisture and ash content, bulk or apparent density, and screen analysis, are also used. *See* CARBON; CHARCOAL; DESTRUCTIVE DISTILLATION; RESPIRATOR.

H. Burnham Allport

Bibliography. American Water Works Assoc., *Design and Use of Granular Activated Carbon*, 1989; H. Janowska, A. Swiatowski, and J. Choma, *Active Carbon*, 1991.

Activation analysis

A technique in which a neutron, charged particle, or gamma photon is captured by a stable nuclide to produce a different, radioactive nuclide which is then measured. The technique is specific, highly sensitive, and applicable to almost every element

in the periodic table. Because of these advantages, activation analysis has been applied to chemical identification problems in many fields of interest.

Neutron method. Neutron activation analysis (NAA) is the most widely used form of activation analysis. In neutron activation analysis the sample to be analyzed is placed in a nuclear reactor where it is exposed to a flux of thermal neutrons. Some of these neutrons are captured by isotopes of elements in the sample; this results in the formation of a nuclide with the same atomic number, but with one more mass unit of weight. A prompt gamma ray is immediately emitted by the new nuclide, hence the term (n,γ) reaction, which can be expressed as reaction (1), where z refers to the atomic number and

$$ {}^{z}_{m}A(n,\gamma) \rightarrow {}_{m+1}^{z}A + \gamma \qquad (1)$$

m the atomic weight. Usually the product nuclide (A^{z}_{m+1}) is radioactive, and by measuring its decay products one can identify and quantify the amount of target element in the sample. The basic activation equation is given by Eq. (2), and enables one

$$A = Nf\sigma\left(1 - e\frac{-0.693t}{t_{1/2}}\right) \qquad (2)$$

where A = activity of product nuclide (disintegrations per second)
N = atoms of target element
f = flux of neutrons (neutrons per cm^2-s)
σ = cross section of the target nuclide (cm^2)
$t_{1/2}$ = half-life of induced radioactive nuclide
t = time of irradiation

to calculate the number of atoms of an unknown target element by measuring the radioactivity of the product. These radioactive products usually decay by emission of a beta particle (negative electron) followed by a gamma ray (uncharged). Cross sections and half-lives are well known, and neutron fluxes can be measured by irradiation and measurement of known materials. When such techniques are used, the method is known as absolute activation analysis. A simple (and older) method is to simultaneously irradiate and compare with the unknown sample a standard containing known amounts of the elements in question. This is called the comparator technique.

Measurement techniques. Measurement of the induced radioactivities is the key to activation analysis. This is usually obtained from the gamma-ray spectra of the induced radionuclides. Gamma rays from radioactive isotopes have unique, discrete energies, and a device that converts such rays into electronic signals that can be amplified and displayed as a function of energy is a gamma-ray spectrometer. It consists of a detector [germanium doped with lithium, GeLi, or sodium iodide doped with thallium, NaI(Tl)] and associated electronics. Gamma rays interact in the detector to form photoelectrons, and these are then amplified and sorted according to energy. Peaks in the resulting gamma-ray spectra are called gamma-ray photo-peaks. By taking advantage of the different half-lives and different gamma-ray energies of the induced radionuclides, positive identification of many elements can be made. A computer, or provision for recording the spectrometric data in computer-compatible form, is almost a necessity. Calibration of counting conditions with a particular detector enables the activation analyst to relate the area under each gamma-ray photo-peak to an absolute disintegration rate; this supplies the A in Eq. (2). Such techniques are multielement, instrumental, and absolute. Where the sought element captures a neutron to produce a non-gamma-emitting nuclide, the analyst must make chemical separations and then beta-count the sample, or must go to another technique. *See* GAMMA-RAY DETECTORS.

Charged-particle method. Activation analysis can also be performed with charged particles (protons or He^{3+} ions, for example), but because fluxes of such particles are usually lower than reactor neutron fluxes and cross sections are much smaller, charged-particle methods are usually reserved for special samples. Charged particles penetrate only a short distance into samples, which is another disadvantage. A variant called proton-induced x-ray emission (PIXE) has been highly successful in analyzing air particulates on filters. Here the samples are all similar, and are low in total mass, and many of the elements of interest such as sulfur, calcium, iron, zinc, and lead are not easy to determine by neutron activation analysis. The protons excite prompt x-rays characteristic of the element, and these are measured. Prompt gamma rays have been used for measurement in some neutron activation analysis studies, but that method has had only limited success. Photon activation, using photons produced by electron bombardment of high-z targets such as tungsten, is another special variant of rather minor interest. Neutron sources other than reactors are sometimes used: Cf^{252} is an element that emits neutrons as it decays, and can be used for neutron activation analysis; and small accelerators called 14-MeV neutron generators produce a low flux of high–energy neutrons which are used primarily for determination of oxygen and nitrogen.

Applications. Activation analysis has been applied to a variety of samples. It is particularly useful for small (1 mg or less) samples, and one irradiation can provide information on 30 or more elements. Samples such as small amounts of pollutants, fly ash, very pure experimental alloys, and biological tissue have been successfully studied by neutron activation analysis. Of particular interest has been its use in forensic studies; paint, glass, tape, and other specimens of physical evidence have been assayed for legal purposes. In addition, the method has been used for authentication of art objects and paintings where only a small sample is available. Activation analysis services are available in numerous university, private, commercial, and United States government laboratories, and while it is not an inexpensive method, for many special samples it is the best and cheapest method of acquiring necessary data. *See* FORENSIC CHEMISTRY;

NUCLEAR REACTION; PARTICLE DETECTOR; RADIOISOTOPE.

W. S. Lyon

Bibliography. Z. B. Alfassi (ed.), *Activation Analysis*, vols. 1 and 2, 1990; W. S. Lyon and H. H. Ross, Nucleonics, *Anal. Chem. (Annu. Rev.)*, 54(4):227R, 1982; 1976 International Conference on Modern Trends in Activation Analysis, *J. Radioanal. Chem.*, vols. 37–39, 1977; 1981 International Conference on Modern Trends in Activation Analysis, *J. Radioanal. Chem.*, vols. 69–71, 1982.

Active sound control

The modification of sound fields by using additional sound transmitted from loudspeakers. The signals from the loudspeakers are controlled electronically by using digital signal processing techniques.

Physical principles. The physical principle underlying the active control of sound is interference. That is, the sound pressure fluctuations produced at a given point in space by two sources of sound simply add together at each point in time. Thus if a secondary source of sound is made to produce the opposite pressure fluctuation to a given primary source at a given point in space, the net result will be silence at that point. *See* INTERFERENCE OF WAVES.

In 1877 Lord Rayleigh described an experiment in which he used electromagnets driven from the same electrical source to synchronize the vibrations of two tuning forks, placed about 9 m (30 ft) apart, at a frequency of 250 Hz. With one ear closed, he was able to delineate regions of silence, and he observed that moving his head about 2.5 cm (1 in.) was "sufficient to produce a marked revival of sound." **Figure 1** shows the results of a computer simulation of Rayleigh's experiment. The two sources of sound radiate spherical waves at the frequency of 250 Hz, and the fields superpose to give regions of both destructive and constructive interference. The region of complete silence detected by Rayleigh lies along the vertical line bisecting that connecting the two sources. Although it is indeed possible to cancel the field produced by a primary source by introducing a secondary source, the quiet zone produced is restricted in dimensions: the size of the zone is proportional to the wavelength of the sound. It is therefore possible, in general, to produce larger quiet zones at lower frequencies, where the wavelength of the sound is longer.

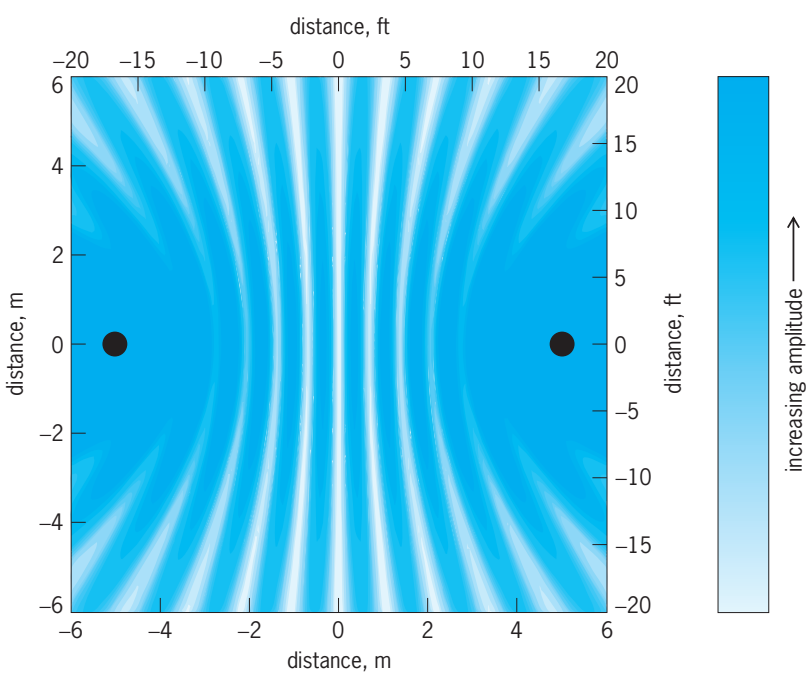

Fig. 1. Results of a computer simulation of the sound field produced by two sources of spherical sound waves in air that have identical source strengths and are separated by a distance of 10 m. The sources are out of phase and radiate sound at a frequency of 250 Hz. The amplitude of the resulting pressure fluctuation is shown.

It is also possible to interfere destructively with the sound field radiated by a primary source of single-frequency sound by placing a secondary source of the same frequency at a distance from the primary source which is very much less than the wavelength of the sound. In this case, if the secondary source is driven out of phase with the primary source, the amplitude of the sound is very much reduced over all space. The secondary source does not "destroy the energy" of the sound radiated by the primary source but merely prevents its radiation in the first place. If the secondary source is placed within one-tenth of a wavelength of the primary source, the amplitude of the sound is reduced everywhere (except very close to the sources) by about 10 decibels on average. This amounts roughly to a halving of the loudness of the sound. Since the wavelength of sound in air is about 3.4 m at 100 Hz (and 0.34 m at 1000 Hz), it is clear that the application of this approach is limited to localized primary sources radiating sound at the lowest frequencies of practical interest.

Application of digital signal processing. The developments in modern electronics during the latter part of the twentieth century enabled the practical implementation of the active control of sound with much greater ease than was possible in Lord Rayleigh's time (**Fig. 2**). A detection microphone is first used to sense the primary sound. The electrical signal from this microphone is passed through an analog low-pass filter, then sampled and converted into digital format. This resulting sequence of numbers is then passed through a digital filter prior to being converted back into an analog signal and fed to the loudspeaker comprising the secondary source. The value of the output from the digital filter at a given time is typically calculated from both the current and a number of previous values of the input sequence. These values are multiplied by a series of numbers comprising the filter coefficients before being added together to produce the output value at that particular time. These arithmetic operations can be carried out extremely fast on a special-purpose microprocessor that is used to implement the digital filter. An error microphone is then used to sense the degree of interference between the primary and secondary sound. *See* DIGITAL FILTER; MICROPROCESSOR.

The coefficients of the digital filter are chosen to ensure that the waveform of the sound radiated from the loudspeaker is aligned in time in order to

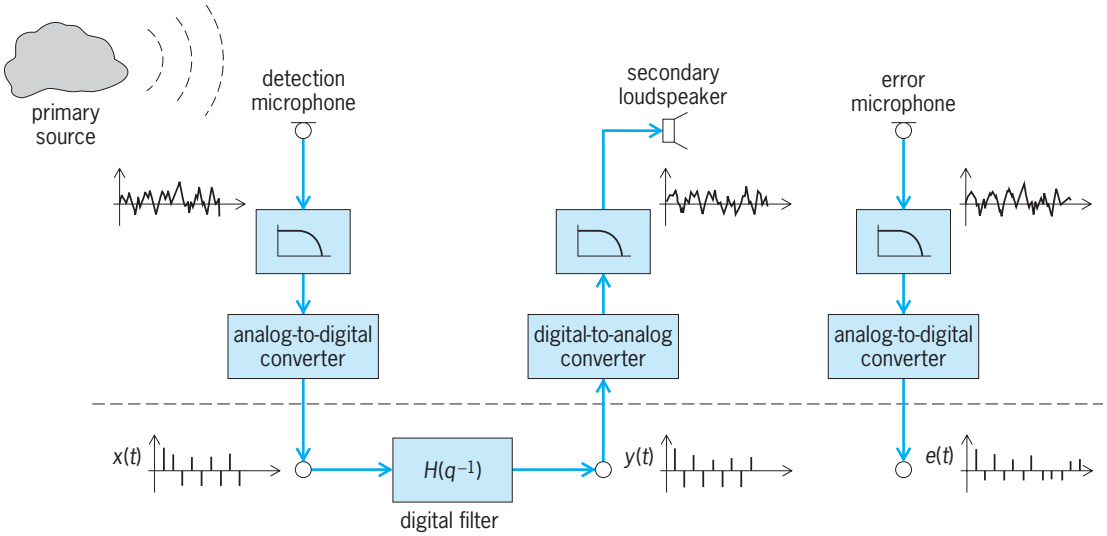

Fig. 2. Feed-forward active noise control system. Elements of the system below the broken line are entirely digital, the filter *H* operating on the input sequence *x* to produce the output sequence *y*. The characteristics of the filter are designed by using the error sequence *e*.

be (as far as possible) opposite to the waveform produced at the error microphone by the primary source. When the waveform of the primary sound is unpredictable, it is important to ensure that the time taken to process and transmit the secondary sound is sufficiently short that it arrives soon enough at the error microphone to cancel the primary sound. This basic approach to the problem is often referred to as feed-forward active control, since it makes use of the advanced warning of the arrival of unwanted sound. This is provided by the finite time required for sound to propagate from the detection microphone to the error microphone.

The design of the filter is often based on an approach that was developed during the 1940s by Norbert Wiener. This technique can be used to design a filter that, for example, minimizes the average over time of the squared value of the signal from the error microphone. Furthermore, the digital filter can be made adaptive so that the coefficients of the filter are rapidly updated at every time sample to ensure that the time-averaged value of the error signal is minimized. An algorithm for accomplishing this is a variant of the LMS algorithm developed by Bernard Widrow and Marcian (Ted) Hoff in the 1960s. Such a strategy can be very useful when, for example, the characteristics of transmission path from the primary source to the error microphone change with time.

Sound from the secondary source may also feed back into the detection microphone. Such a problem often arises when active control systems are used to reduce the amplitude of sound propagating in tubes or ducts (such as those in air-conditioning systems). It is possible to compensate for this effect by modeling the feedback path electronically with a further filter whose output is then subtracted from that of the detection microphone. *See* ADAPTIVE SIGNAL PROCESSING.

Multichannel systems. Some physical limitations of the technique can be overcome simply by increasing the number of secondary sources used to control the field. For example, if 10 loudspeakers are used to control a pure-tone sound field, it is possible in principle to produce 10 points of silence in the field. Of course, depending upon the geometrical arrangement and the acoustical environment of the primary and secondary sources, increases in level may well be produced at other positions. Another strategy is to drive the secondary sources in order to minimize the sound level at a number of error microphones which is larger than the number of secondary sources. An example of the practical application of this approach is in the control of propeller noise inside aircraft (**Fig. 3**). In this case the dominant source of unwanted noise is at the blade passage frequency of the propellers, typically below about 100 Hz. In addition, the interiors of such aircraft are usually relatively confined and the sound field produced by the rotation of the propellers excites a relatively small number of resonant modes of the aircraft cabin. Under these circumstances it is possible to produce widespread reductions in sound level. **Figure 4** shows the sound field both before and after when 16 secondary loudspeaker sources are used to minimize the sum of

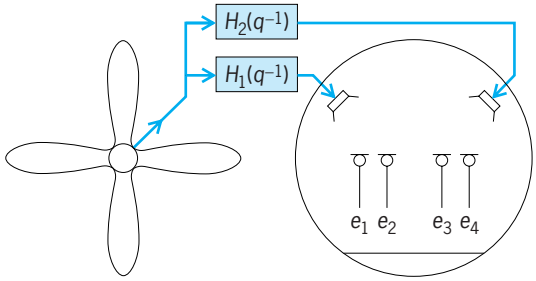

Fig. 3. Active control system for the suppression of propeller-induced aircraft cabin noise. The digital filters H_1 and H_2 are designed to minimize the sum of the squared error signals e_1–e_4.

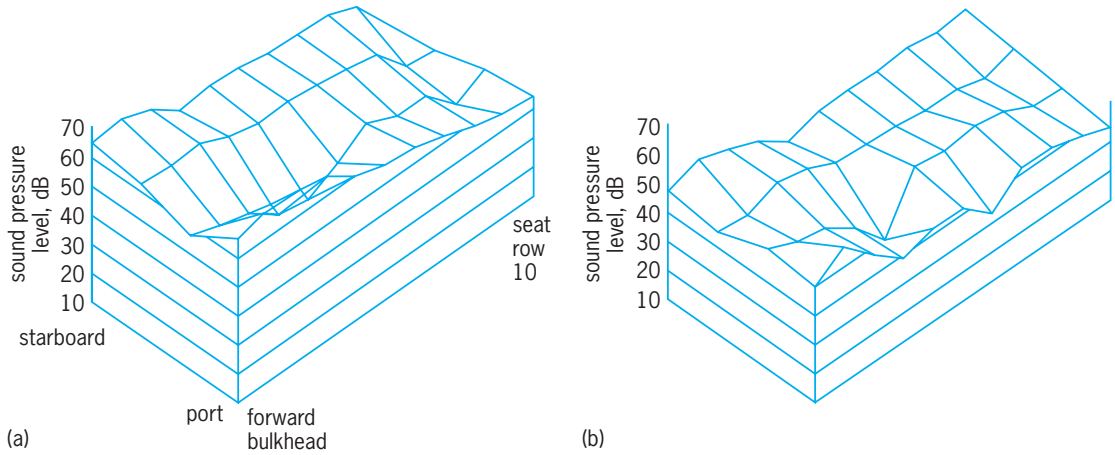

Fig. 4. Distribution of sound pressure level (in decibels on an arbitrary scale) at head height in a 48-seat propeller aircraft (*a*) before and (*b*) after the use of a multichannel active noise control system consisting of 16 loudspeakers and 32 microphones. The results are at a frequency of 88 Hz and show the reduction in the contribution due to the port propeller only. (*After S. J. Elliott et al., In-flight experiments on the active control of propeller-induced cabin noise, J. Sound Vib., 140:219–238, 1990*)

the squared pressure amplitudes at 32 error microphones distributed at head height through the cabin of a 48-seat passenger aircraft. Since the basic source of sound is due to the rotation of the propellers, all that is required is the detection of the frequency of the primary source (88 Hz in this example), and this is readily provided by filtering a signal from a tachometer on one of the engines. This reference signal is then sampled and passed through a digital filter associated with each secondary source. (This signal is uncorrupted by feedback from the secondary source outputs.) The coefficients of these filters are then adjusted adaptively in response to the signals sampled at the error microphones by using a multichannel generalization of the LMS algorithm. This ensures the minimization of the sum of the squared outputs from the error microphones and thus the suppression of the sound field.

Many acoustic waveforms of practical interest are highly unpredictable (unlike the propeller noise referred to above). Examples include the waveforms of the noise generated inside an automobile due to the airflow past the passenger cabin and the vibrations generated by the contact of the tires with the road. In addition, there may be multiple primary sources of such unwanted sound. In these cases it becomes more difficult to find reference signals that give a prior indication of the acoustic pressure fluctuations well before their arrival at the ears of the occupants. **Figure 5** shows a matrix of digital filters whose inputs are from a number of vibration sensors placed on the body of the automobile and whose outputs are used to drive the secondary sources. It is possible to find the optimal matrix of digital filters that ensures the minimization of the sum of the time-averaged signals from a number of error microphones. **Figure 6** shows the sound pressure level (in decibels) at a position close to the ear of a driver of a small passenger car both before and after the implementation of active control. The filter inputs were provided by reference signals from six vibration sensors placed at carefully chosen positions on the car body. Two secondary loudspeakers were used, each being fed by a combination of all the digitally filtered reference signals in order to minimize the sum of the time-averaged signals from two error microphones. It is clear that some reductions in sound level can be produced, but these are again restricted to frequencies below about 250 Hz. *See* ACOUSTIC NOISE; CONTROL SYSTEMS; SOUND.

Philip A. Nelson

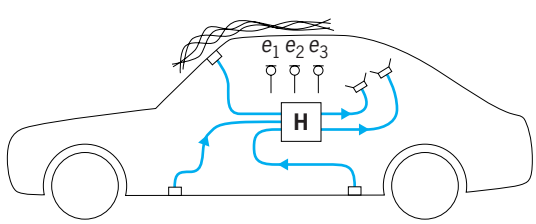

Fig. 5. Multichannel feed forward active control system for suppressing noise from multiple primary sources.

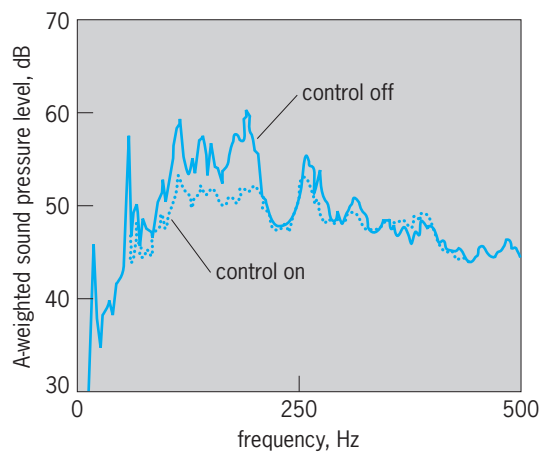

Fig. 6. Performance of a multichannel feed forward system for the active suppression ofroad-induced noise in a small passenger car at a speed of 60 km/h (37 mi/h). Reference signals were taken from six vibration sensors on the vehicle suspension. The A-weighted sound pressure level is shown before and after the application of active control. (*After T. J. Sutton et al., Active control of road noise inside vehicles, Noise Control Eng. J., 42:137–148, 1994*)

Bibliography. S. Elliott, *Signal Processing for Active Control*, Acadmic Press, 2000; P. A. Nelson and S. J. Elliott, *Active Control of Sound*, Academic Press, 1992; A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Practical Applications*, McGraw-Hill, 1981, reprint, Acoustical Society of America, 1989; L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, Prentice Hall, 1975; B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice Hall, 1985.

Activity (thermodynamics)

The activity of a substance is a thermodynamic property that is related to the chemical potential of that substance. Activities are closely related to measures of concentration, such as partial pressures and mole fractions, and are more convenient to use than chemical potentials. The conditions that hold in chemical reaction equilibrium and in phase equilibrium can be expressed in terms of activities of the species involved.

The activity a_i of chemical species i in a phase (a homogeneous portion) of a thermodynamic system is defined by Eq. (1), where μ_i is the chemical potential of i in that phase, μ_i° is the chemical potential of i in its standard state, R is the gas constant, and T is the absolute temperature. Equation (1) shows that the activity a_i depends on the choice of standard state for species i. Since μ_i is an intensive quantity that depends on the temperature, pressure, and composition of the phase, a_i is an intensive function of these variables. a_i is dimensionless. When $\mu_i = \mu_i^\circ$, then $a_i = 1$. The degree of departure of a_i from 1 measures the degree of departure of the chemical potential from its standard-state value. The chemical potential is defined by $\mu_i \equiv \partial G/\partial n_i$, where G is the Gibbs energy of the phase, n_i is the number of moles of substance i in the phase, and the partial derivative is taken at constant temperature, pressure, and amounts of all substances except i. *See* CHEMICAL THERMODYNAMICS; FREE ENERGY.

$$a_i \equiv \exp[(\mu_i - \mu_i^\circ)/RT] \tag{1}$$

From Eq. (1), it follows that μ_i is given by Eq. (2).

$$\mu_i = \mu_i^\circ + RT \ln a_i \tag{2}$$

The chemical potentials μ_i are key thermodynamic quantities of a phase, since all thermodynamic properties of the phase can be found if the chemical potentials in the phase are known as functions of temperature, pressure, and composition. Activities are more convenient to work with than chemical potentials, because the chemical potential of a substance in a phase goes to minus infinity as the amount of that substance in the phase goes to zero; also, one can determine only a value of μ_i relative to its value in some other state, whereas one can determine the actual value of each a_i.

Gases. For a pure gas i or a component i of any gas mixture, the standard state is defined as pure ideal gas i at the temperature T of the pure gas or mixture and at a standard pressure P° of 1 bar (100 kilopascals). Formerly, 1 atm was used as the standard-state pressure. Since real gases do not exhibit ideal behavior at 1 bar, the standard state of a gas is a hypothetical state.

For a component of an ideal gas mixture (one whose density is low enough that intermolecular forces are negligible), the chemical potential can be shown to be given by Eq. (3), where μ_i° is a function

$$\mu_i = \mu_i^\circ + RT \ln(P_i/P^\circ) \tag{3}$$

of T, and the partial pressure P_i of gas i is defined by $P_i \equiv x_i P$, where x_i is the mole fraction of i in the mixture and P is the mixture's pressure. Comparison with Eq. (2) shows that $a_i = P_i/P^\circ$ for a component of an ideal gas mixture. The motivation for the Eq. (1) definition of activity is to produce an expression for the chemical potential that closely resembles the expression for μ_i in an ideal system, so that the degree of departure from ideal behavior can be readily assessed. Thus, the general expression Eq. (2) for μ_i resembles Eq. (3) for an ideal gas mixture. For a component i of a real gas mixture, the fugacity f_i is defined as $f_i \equiv a_i P^\circ$, so that $a_i = f_i/P^\circ$ in a real gas mixture, as compared with $a_i = P_i/P^\circ$ for an ideal gas mixture. The fugacity coefficient ϕ_i of component i of a gas mixture is defined by $\phi_i \equiv f_i/P_i$; so $a_i = \phi_i P_i/P^\circ$. In an ideal gas mixture, each ϕ_i equals 1.

For a gas mixture at temperature and pressure T and P', the activity a_i of component i can be found by using Eq. (4) to find the fugacity coefficient ϕ_i, where

$$\ln \phi_i = \int_0^{P'} \left(\frac{\overline{V}_i}{RT} - \frac{1}{P} \right) dP \tag{4}$$

the integration is at constant T and composition and $\overline{V}_i$ is the partial molar volume of component i in the mixture. A reliable equation of state is often used for the mixture to find an approximate expression for $\overline{V}_i$ to use in Eq. (4).

Pure solids and pure liquids. For a pure solid or liquid, the standard state is defined as the pure substance at a pressure of 1 bar (100 kPa) and at the temperature T of the substance. For most pure solids or liquids at pressures below 20 bar (2000 kPa), a_i is quite close to 1.

Solutions. Several possible choices exist for standard states of the components of a liquid or solid solution.

One choice (called the symmetrical convention) is to treat all components on the same footing and take the standard state of each component i as pure substance i at the temperature and pressure of the solution. This choice is commonly made for solutions of miscible liquids. With this choice, the composition of the solution is expressed using mole fractions.

For an ideal solution (one where the components of the solution resemble each other so closely in molecular structure that the differences in molecular

size and shape and in intermolecular forces are negligible), one can show that the chemical potentials are given by Eq. (5), where μ_i^* is the chemical potential

$$\mu_i = \mu_i^* + RT \ln x_i = \mu_i^\circ + RT \ln x_i \qquad (5)$$

of pure liquid i at the temperature and pressure of the solution and x_i is the mole fraction of i in the solution. Thus, $a_i = x_i$ for an ideal-solution component.

For a component of a nonideal solution where the symmetrical convention is used, the activity is defined by Eq. (1) and the activity coefficient f_i of component i (not to be confused with the symbol used for fugacity in gases) is defined by $f_i \equiv a_i/x_i$, so $a_i = f_i x_i$. For an ideal solution, each f_i equals 1. In a nonideal solution, f_i measures the degree of departure from ideality and f_i goes to 1 as x_i goes to 1.

In the unsymmetrical convention, one component of the solution is designated as the solvent and the other components are called the solutes. The standard state of the solvent is chosen as pure solvent at the temperature and pressure of the solution. If the solution composition is specified using mole fractions, the standard state of solute i is chosen as the hypothetical state with $x_i = 1$ and solute i behaving as it would in an ideally dilute solution. An ideally dilute solution is one so extremely dilute that molecules of i interact only with solvent molecules. If the solution composition is specified using molalities m_i for solutes, then the standard state of solute i is the hypothetical state with $m_i = m^\circ \equiv 1$ mol/kg and i behaving as if in an ideally dilute solution. If molar concentrations c_i are used, the standard state has $c_i = c^\circ \equiv 1$ mol/L and i exhibiting ideally dilute behavior. These three choices and Eq. (1) give three different solute activities, $a_{x,i}$, $a_{m,i}$, and $a_{c,i}$. Three different solute activity coefficients, $\gamma_{x,i}$, $\gamma_{m,i}$, and $\gamma_{c,i}$, are defined to satisfy $a_{x,i} = \gamma_{x,i} x_i$, $a_{m,i} = \gamma_{m,i} m_i/m^\circ$, and $a_{c,i} = \gamma_{c,i} c_i/c^\circ$. All these solute-$i$ activity coefficients go to 1 as the solvent mole fraction goes to 1. *See* CONCENTRATION SCALES.

For an electrolyte solute C_cA_a, where c and a are the numbers of cations and anions, respectively, in the formula, Eq. (2) can be applied to the solute as a whole and also to each individual ion. The molality scale is usually used for the solute. The activity a_i of the electrolyte solute is related to the activities a_+ and a_- of its cation and anion, respectively, by $a_i = a_+^c a_-^a$. Since properties of individual ions are not experimentally determinable, a mean molal activity coefficient $\gamma_\pm$ is defined by the relation $(\gamma_\pm)^{c+a} = (\gamma_+)^c(\gamma_-)^a$. Values of $\gamma_\pm$ can be found from vapor-pressure measurements or galvanic-cell data.

Some workers take the standard-state pressure for each solution component as the fixed pressure of 1 bar.

Reaction equilibrium. The equilibrium constant for a chemical reaction can be expressed in terms of the equilibrium values of the activities of the reactants and products. *See* CHEMICAL EQUILIBRIUM.

Phase equilibrium. For a solution in equilibrium with its vapor (assumed to be ideal), the equilibrium condition that the chemical potential of substance i must be equal in each phase leads to the relation $P_i = f_i x_i P_i$, where P_i is the partial pressure of substance i in the vapor in equilibrium with the solution, P_i° is the vapor pressure of pure i at the temperature of the solution, x_i is the mole fraction of i in the solution, and f_i is the symmetrical-convention activity coefficient of i in the solution. This relation allows the activity coefficient and hence the activity of i in the solution to be found. If the vapor departs significantly from ideality, pressures are replaced by fugacities in this relation. Ira N. Levine

Bibliography. K. Denbigh, *The Principles of Chemical Equilibrium*, Cambridge University Press, 4th ed., 1981; I. N. Levine, *Physical Chemistry*, 4th ed., McGraw-Hill, 1995.

Activity-based costing

A specific system or procedure for determining accurate manufacturing costs in order to achieve profitability goals. Traditional accounting systems allocate overhead costs by using a volume-oriented base, such as direct labor hours or direct material dollars. The cost allocation bases of direct labor or production quantity were designed primarily for inventory valuation. As a consequence, traditional cost accounting methods are fully effective only when direct labor (or direct materials) is the dominant cause of cost.

Cost models. While traditional standard cost systems were effective in the past, changes in manufacturing technologies, such as the just-in-time philosophy, robotics, and flexible manufacturing systems, made traditional cost models somewhat obsolete. Rapid technological advancement has resulted in the restructuring of manufacturing cost patterns (for example, the direct labor and inventory costs are decreasing, while those of technology depreciation, engineering, and data processing are increasing). Because of the changing nature of these types of costs, existing (that is, traditional) cost accounting systems often do not adequately support the objectives of advanced manufacturing. *See* FLEXIBLE MANUFACTURING SYSTEM.

In automated manufacturing facilities, distortion in product costs results from volume-based allocations such as those based on direct labor, because the bases used to allocate overhead do not cause the costs. As a result, product cost distortion occurs because of high overhead rates that are inflated by many costs that should be directly traceable to the product rather than arbitrarily allocated on the basis of direct labor or direct materials content. Several components of a product's cost that should be traced to the product include hidden overhead costs, such as for material movement, order processing, process planning, rework, warranty maintenance, production planning and control, and quality assurance. *See* MATERIALS HANDLING.

Key concepts. Activity-based costing systems track direct and indirect costs to the specific activities that

TABLE 1. Manufacturing costs for Models A and B (current year)

Category	Model A	Model B
Direct material cost	$2,800,600	$1,500,000
Direct labor cost	350,000	250,000
Direct labor hours	35,000	25,000
Total manufacturing overhead costs		$12,000,000

cause them and thus attempt to provide a more reliable estimate of product cost. In this approach, four key concepts differentiate activity-based costing from traditional costing systems: activity accounting, cost drivers, direct traceability, and non-value-added costs. These four basic concepts are embodied in activity-based costing systems and lead to more accurate costing information. In addition, activity-based costing systems provide more flexibility than conventional (for example, absorption-based) costing systems because they produce a variety of information useful for technology accounting, product costing, and design-to-cost.

Activity accounting. In an activity-based system, the cost of the product is the sum of all costs required to manufacture and deliver the product. The activities that a firm pursues consume its resources; and resource availability and usage create costs. Activity accounting decomposes an organization into its basic business processes and their associated activities in order to provide a cause-and-effect rationale for how fundamental objectives create costs and result in outputs.

Cost drivers. A cost driver is an activity that causes costs to be incurred. Familiar cost drivers include the number of machine setups, the number of engineering change notices, and the number of purchase orders. By controlling the cost driver, unnecessary costs can be eliminated, resulting in improved product cost.

Direct traceability. Direct traceability involves attributing costs to those products or processes that consume resources. Many hidden overhead costs can be effectively traced to products, thus providing more accurate information upon which to base decisions such as product costing, make-versus-buy, and investment justification.

Non-value-added costs. In manufactured products, customers may perceive that certain features add no value to the product. Through identification of cost drivers, a firm can pinpoint these sources of unnecessary cost. In fact, an activity-based cost system attempts to identify and place a cost on the manufacturing processes performed (either value-adding or non-value-adding) so that management can determine desirable changes in resource requirements for these activities. This approach is in contrast to traditional cost systems that accumulate costs by budgetary line items and by functions.

Application. An example will illustrate the key concepts of activity-based costing. A hypothetical firm, the XYZ Company, makes pagers for the telecommunications industry. Currently, the company is making two models of pagers with equal annual production volume of 50,000 each. Ever since its inception, the company has been using an absorption-based traditional costing system for product costing. Thus, the company uses direct labor hours as the allocation base for all its manufacturing overhead costs (**Table 1**).

Recently, the company installed an activity-based costing system and subsequently identified six main activities to be responsible for most of the manufacturing overhead costs. The overhead costs were distributed to these six activities through activity accounting. Also, six corresponding cost drivers and their budget consequences for the current year were established (**Table 2**). Besides establishing the activity-based costing data, the activity-based costing system also measured the levels of each activity (or cost driver rate) required by the two products (**Table 3**). For postaudit purposes the company ran a cost comparison between the two costing systems to arrive at product costs for two models, A and B. The two sets of computations involved the traditional (absorption-based) costing system and the activity-based costing system.

For the cost computation using the traditional method, the computation involved determining the overall manufacturing overhead cost rate, that is, total overhead costs divided by total direct hours. When the total cost per unit for each model was computed, the value for Model A was $203 and the value for Model B was $135.

A very different cost per unit resulted from cost computation using activity-based costing. In this method, the first step was to establish the applied activity rate of the six activities; the applied activity rate equals the activity cost divided by the budgeted cost driver rate in Table 2. In the second step, the activity cost is traced to each pager model, that is, the cost driver rate from Table 3 multiplied by the

TABLE 2. Activity-based costing data

Activity	Costs	Cost driver	Budgeted rate
Production	$8,000,000	Number of machine hours	200,000 hours
Engineering	1,000,000	Number of engineering change orders	40,000 orders
Material handling	1,000,000	Number of material moves	60,000 moves
Receiving	800,000	Number of batches	500 batches
Quality assurance	800,000	Number of inspections	20,000 inspections
Packing and shipping	400,000	Number of products	100,000 products

TABLE 3. Cost driver rates for Models A and B

Activity	Model A	Model B
Production	50,000 hours	150,000 hours
Engineering	15,000 orders	25,000 orders
Materials handling	20,000 moves	40,000 moves
Receiving	150 batches	350 batches
Quality assurance	6,000 inspections	14,000 inspections
Packing and shipping	50,000 products	50,000 products

TABLE 4. Product cost comparison

Costing system	Model A	Model B
Traditional	$203.00	$135.00
Activity-based	130.78	207.24

applied activity rate from Table 1. When the total cost per unit model was determined by this method, the cost per unit was $130 for Model A and $207 for Model B. There is a significant difference in the cost per unit as computed by these two different methods.

Thus the two costing systems produced different cost estimates (**Table 4**). Using absorption-based (traditional) costing, Model A costs more to make than Model B. However, based on activity-based costing, Model A is less expensive to make than Model B.

Traditional costing bases its cost estimates on the assumption that Model A is responsible for more overhead costs than Model B, using direct labor hours as an allocation base. However, because of differences in the number of transactions that affect indirect (overhead costs), the reality is that Model B actually incurs more overhead costs than Model A. This is shown by the analysis carried out with activity-based costing. As this example demonstrates, a traditional costing method generates distorted product costs that do not accurately reflect the actual amount of overhead costs incurred by the product. Companies making decisions based on distorted costs may unknowingly price some of their products out of the market while selling others at a loss. Or they may make key design decisions, based on competitor target costs, that do not reflect the true costs of manufacturing their new or revised products. *See* INDUSTRIAL ENGINEERING. William G. Sullivan

Bibliography. J. A. Brimson, *Activity Accounting: An Activity-Based Costing Approach*, 1997; R. Cooper and R. S. Kaplan, Measure costs right: Make the right decisions, *Harvard Bus. Rev.*, September-October 1988; R. Cooper and R. S. Kaplan, *The Design of Cost Management Systems*, 2d ed., 1998; H. T. Johnson, Activity-based management: Past, present, and future, *Eng. Econ.*, 36(3):219–238, Spring 1991; R. S. Kaplan, New approaches to measurement and control, *Eng. Econ.*, 36(3):201–218, Spring 1991; J. K. Shank, Strategic cost management: New wine, or just new bottles, *Manag. Acc. Res.*, 1:47–65, Fall 1989.

Acylation

A process in which a hydrogen atom in an organic compound is replaced by an acyl group (R—CO where R = an organic group). The reaction involves substitution by a nucleophile (electron donor) at the electrophilic carbonyl group (C═O) of a carboxylic acid derivative. The substitution usually proceeds by an addition-elimination sequence [reaction]. Two common reagents, with the general formula RCOX, that bring about acylation are acid halides (X = Cl,Br) and anhydrides (X = OCOR). There are also other acylating reagents. The carboxylic acid (X = OH) itself can function as an acylating agent when it is protonated by a strong acid catalyst as in the direct esterification of an alcohol. Typical nucleophiles in the acylation reaction are alcohols (ROH) or phenols (ArOH), both of which give rise to esters, and ammonia or amines (RNH_2), which give amides. *See* ACID ANHYDRIDE; ACID HALIDE; AMIDE; AMINE; ELECTROPHILIC AND NUCLEOPHILIC REAGENTS.

Carbon acylation. Acylation at a carbon atom can take several forms. Reaction of an acid halide with an organometallic compound such as a Grignard reagent (RMgX) is a rapid and efficient route to ketones. A widely used reaction is acylation of an aromatic ring with an acid chloride or anhydride in a Friedel-Crafts reaction. In acylation of the α-CH_2 group of a ketone or ester, hydrogen usually is first removed with a base to give the enolate, and the acyl group is then introduced. An acid chloride can be used, or since the enolate is a strong nucleophile, another ester can serve. *See* FRIEDEL-CRAFTS REACTION; GRIGNARD REACTION.

Esters are not generally useful as acylating agents unless the nucleophile is a strong base such as an enolate or the reaction is catalyzed by strong acid. However, thioesters are significantly more reactive acylating agents. In living systems, acylation occurs in several metabolic pathways; a thioester, acetyl coenzyme A, is the key intermediate in such reactions. *See* ESTER.

Another type of acylation is the reaction of one ester with a different alcohol to give a new ester. In this case the original ester is the acylating agent; the process is called transesterification and is exemplified by methanolysis of a fat. This reaction, which is usually carried out with an acid catalyst, is the first step in analysis of the fatty acid composition of fats.

Other carboxyl activation methods. Acylation can be thought of as the transfer of an acyl group from some reactive form of a carboxylic acid to a nucleophilic acceptor. As organic synthesis has become more refined and targets more complex, improved ways of activating carboxylic acids have been developed. These enable acylation to be carried out with a variety of acids, more selectively and under milder

conditions, than is possible with traditional acid chlorides or anhydrides.

One approach is to convert an acid to a mixed carboxylic-carbonic anhydride. These derivatives are prepared from the acid, ethyl chloroformate, and a tertiary amine at 0°C (32°F). They are more reactive than acid chlorides for acylation of amines or diazomethane.

Formation of an amide bond between two protected amino acids is an important process. This acylation is called peptide coupling, and it is the crucial reaction in the synthesis of peptides and proteins. Many methods have been developed for this purpose. One of the most enduring is the use of carbodiimide reagents, which effect activation of the acid and reaction with the amine in one operation. *See* PEPTIDE; PROTEIN.

James A. Moore

Bibliography. W. Carruthers and I. Coldham, *Modern Methods of Organic Synthesis*, 4th ed., 2004; R. J. Sundberg and F. A. Carey, *Advanced Organic Chemistry*, *Part B: Reaction and Synthesis*, 4th ed., 2001; B. M. Trost and I. Fleming (eds.), *Comprehensive Organic Synthesis*, vol. 6, 1991.

Adaptation (biology)

A characteristic of an organism that makes it fit for its environment or for its particular way of life. For example, the Arctic fox (*Alopex lagopus*) is well adapted for living in a very cold climate. Appropriately, it has much thicker fur than similar-sized mammals from warmer places; measurement of heat flow through fur samples demonstrates that the Arctic fox and other arctic mammals have much better heat insulation than tropical species. Consequently, Arctic foxes do not have to raise their metabolic rates as much as tropical mammals do at low temperatures. This is demonstrated by the coati (*Nasua narica*), which lives in Panama and has a body mass similar to Arctic foxes and about the same metabolic rate at comfortable temperatures. When both animals are cooled, however, the coati's metabolic rate starts to rise steeply as soon as the temperature falls below 68°F (20°C), while that of the Arctic fox begins to rise only below −22°F (−30°C). The insulation is so effective that Arctic foxes can maintain their normal deep-body temperatures of 100°F (38°C) even when the temperature of the environment falls to −112°F (−80°C). Thus, thick fur is obviously an adaptation to life in a cold environment. *See* THERMOREGULATION.

In contrast to the clear example above, it is often hard to be sure of the effectiveness of what seems to be an adaptation. For example, tunas seem to be adapted to fast, economical swimming. Their swimming muscles are kept warmer than the environment. The body has an almost ideal streamlined shape, rounded in front and tapering behind to a very narrow caudal peduncle immediately in front of the tail fin. The tail fin itself is tall and narrow, the shape that should enable it to propel the fish at least energy cost. Experiments with young tunas, however, have failed to show that they are faster or more economical than apparently less well adapted relatives.

A structure that evolved as an adaptation for one function may later come to serve a different function. This phenomenon is known as exaptation. For example, feathers seem to have evolved in the ancestors of birds as an adaptation for heat insulation, but in the evolution of wings feathers were exapted for flight. The wings in turn became an exaptation for swimming, when the penguins evolved.

Evolution by natural selection tends to increase fitness, making organisms better adapted to their environment and way of life. It might be inferred that this would ultimately lead to perfect adaptation, but this is not so. It must be remembered that evolution proceeds by small steps. For example, squids do not swim as well as fish. Fish swim by beating their tails, and squids by jet propulsion. Comparison of the swimming performance of a trout and a similar-sized squid showed that the fish could swim faster and with less energy cost. The squid would be better adapted for swimming if it evolved a fish-like tail instead of its jet propulsion mechanism, but evolution cannot make that change because it would involve moving down from the lesser adaptive summit before climbing the higher one. *See* ANIMAL EVOLUTION; ORGANIC EVOLUTION.

R. McNeill Alexander

Bibliography. R. McNeill Alexander, *Optima for Animals*, 2d ed., 1996; U. Dieckmann et al., *Adaptive Speciation*, 2004; S. H. Orzack and E. Sober, *Adaptationism and Optimality*, 2001; D. Schluter, *The Ecology of Adaptive Radiation*, 2000; T. Shanahan, *The Evolution of Darwinism*, 2004; E. R. Weibel, C. R. Taylor, and L. Bolis (eds.), *Principles of Animal Design*, 1998.

Adaptive control

A special type of nonlinear control system which can alter its parameters to adapt to a changing environment. The changes in environment can represent variations in process dynamics or changes in the characteristics of the disturbances. *See* NONLINEAR CONTROL THEORY.

A normal feedback control system can handle moderate variations in process dynamics. The presence of such variations is, in fact, one reason for introducing feedback. There are, however, many situations where the changes in process dynamics are so large that a constant linear feedback controller will not work satisfactorily. Control of a supersonic aircraft is a typical example. The dynamics of the airplane changes drastically with Mach number and dynamic pressure. A flight control system with constant parameters will not work well for an aircraft which operates over wide ranges of speeds and altitudes. *See* FLIGHT CONTROLS.

Adaptive control is also useful for industrial process control. In a given operating condition, most

processes can be controlled well with regulators with fixed parameters. Since delay and holdup times depend on the production, it would, however, be desirable to retune the regulators when there is a change in production. Adaptive control can also be used to compensate for changes due to aging and wear. Typical examples are variations in catalyst activity in chemical reactions and slow changes in heat transfer due to sediments. Wear in valves and mechanical systems are other examples. *See* PROCESS CONTROL.

Gain scheduling. It is sometimes possible to find auxiliary variables in a system which correlate well with the changes in process dynamics. It is then possible to eliminate the influences of parameter variations by changing the parameters of the regulator as functions of the auxiliary variables (**Fig. 1**). This method of eliminating variations in process dynamics is called gain scheduling. Gain scheduling could be considered an extension of feed-forward compensation. It can be seen from Fig. 1 that there is no way to correct for an incorrect schedule.

There is a controversy in nomenclature as to whether or not a system with gain scheduling should be considered an adaptive system. Gain scheduling is, nevertheless, a very useful technique for eliminating parameter variations. It is, in fact, the predominant method used to handle parameter variations in flight control systems. In that case, the Mach number and the dynamic pressure are measured by air data sensors and used as scheduling variables. The parameters of the flight control system are then determined by table look-up and interpolation.

One drawback of systems based on gain scheduling is that their design is time-consuming. The controllers must be designed for each operating condition. The interpolation method and the safe operation of the system must also be verified by extensive simulations. It is sometimes possible to obtain the gain scheduling by using normalized dimension-free parameters. The auxiliary measurements are used together with the process measurements to obtain the normalized variables. The normalized control variable is calculated as the output of a linear constant-coefficient system driven by the normalized measurements. The control variable is retransformed before it is applied to the process.

Model reference adaptive systems. In model reference adaptive systems (MRAS) the dynamic specifications are given in terms of a reference model which tells how the process output ideally should respond to the command signal. The reference model is part of the control system (**Fig. 2**). The regulator can be thought of as consisting of two loops. The inner loop is an ordinary control loop composed of the process and a regulator. The parameters of the regulator can be adjusted, and the adjustments are made in the outer loop, which attempts to drive the regulator parameters in such a way that the error between the model output y_m and the process output y becomes small. The outer loop thus also looks like a regulator loop. The key problem is to determine the adjustment mechanism so that a stable system which brings the error to zero is obtained. This problem is nontrivial. It is easy to show that it cannot be solved with a simple linear feedback from the error to the controller parameters.

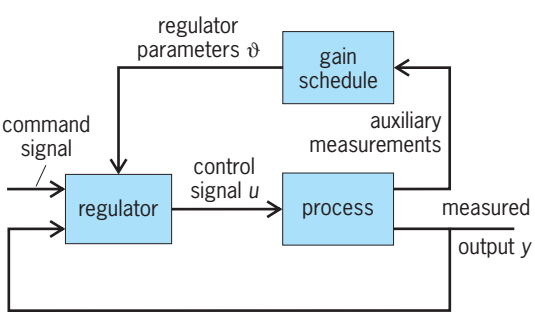

Fig. 1. Block diagram of a system where parameter variations are eliminated by gain scheduling.

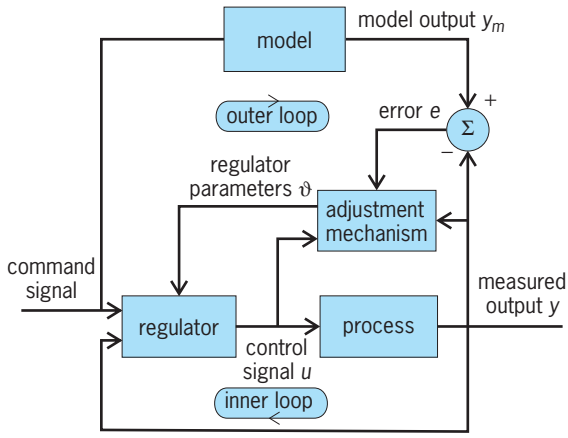

Fig. 2. Block diagram of model reference adaptive system (MRAS).

The parameter adjustment mechanism of Eq. (1),

$$\frac{d\vartheta_i}{dt} = -k\frac{\partial e}{\partial \vartheta_i} \cdot e \quad i = 1, \ldots, n \tag{1}$$

called the MIT rule, was used in the original MRAS. The variables $\vartheta_1, \ldots, \vartheta_n$ are the adjustable regulator parameters, $e = y_m - y$ is the error, $\partial e/\partial \vartheta_i$ are the sensitivity derivatives, and k is a parameter which determines the adaptation rate. Equation (1) represents a parameter adjustment mechanism which is composed of three parts: a linear filter for computing the sensitivity derivatives from process inputs and outputs, a multiplier, and an integrator. This configuration is typical for many adaptive systems.

The MIT rule can unfortunately give an unstable closed-loop system. The rule can be modified using Lyapunov or Popov stability theory. But it was only in the late 1970s that real progress was made in the theory of stability for adaptive systems.

Self-tuning regulators. The self-tuning regulator (STR; **Fig. 3**) can be thought of as composed of two loops. The inner loop consists of the process and an ordinary linear feedback regulator. The parameters of the regulator are adjusted by the outer loop, which is composed of a recursive parameter estimator and

a design calculation. There are many variants of the self-tuners because there are many combinations of design and parameter estimation schemes.

The regulator shown in Fig. 3 is called a regulator based on identification of an explicit process model. It is sometimes possible to reparametrize the process in such a way that the process is expressed in terms of the regulator parameters. The self-tuning regulator is then considerably simplified, because the design calculations are eliminated. Such a self-tuner is called an algorithm, based on estimation of an implicit process model. It is very similar to a model reference adaptive system because the parameter estimator can be interpreted as an adjustment mechanism for the regulator parameters, as can be seen by comparison of Figs. 2 and 3.

The self-tuning regulators can be used in several different ways. Since the regulator becomes an ordinary constant-gain feedback if the parameter estimates are kept constant, it can be used as a tuner whose purpose is to adjust the parameters of a control loop. In this case, the self-tuner is connected to the process, which is run until satisfactory performance is obtained. The self-tuner is then disconnected, and the system is left with the constant regulator parameters obtained. Since the tuning is done automatically, it is possible to use control algorithms with many adjustable parameters. The self-tuner can also be used to build up a gain schedule. The system is then run with the self-tuner at different operating points. The controller parameters obtained are stored. In this way, it is possible to obtain suitable regulator settings for different operating conditions. The self-tuner can, of course, also be used as a true adaptive controller for systems with varying parameters.

Stochastic adaptive control. Regulator structures such as MRAS and STR are based on heuristic arguments. It would be appealing to obtain the regulators from a unified theoretical framework. This can, in principle, be done by using nonlinear stochastic control theory. The system and its environment are then described by a stochastic model. The criterion is formulated so as to minimize the expected value of a loss function which is a scalar function of states and controls. *See* STOCHASTIC CONTROL THEORY.

The problem of finding a control which minimizes the expected loss function is difficult. Useful explicit conditions for existence are not known in general.

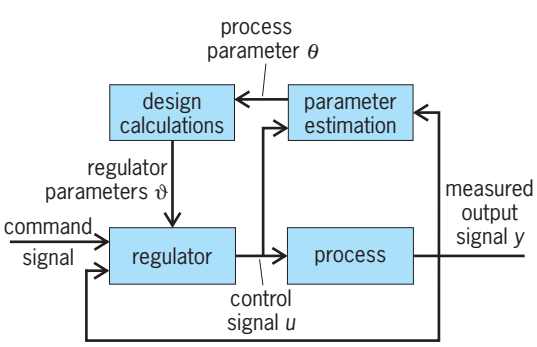

Fig. 3. Block diagram of a self-tuning regulator (STR).

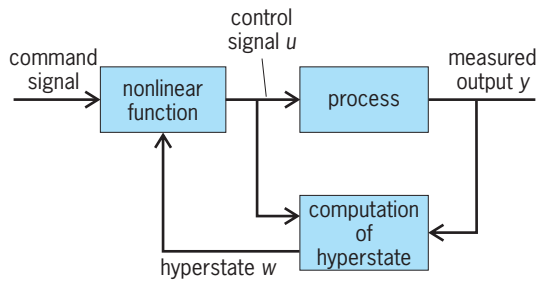

Fig. 4. Block diagram of an optimal nonlinear stochastic controller. The hyperstate is generated from a dynamical system using *u* and *y* as inputs. The regulator is a static nonlinear function which gives the control variable as a function of the hyperstate and the command signal.

Under the assumption that a solution exists, a functional equation for the optimal loss function can be derived by using dynamic programming. This equation, which is called the Bellman equation, can be solved numerically only in very simple cases. Nevertheless, the approach is of interest because it gives an insight into the structure of the optimal controller (**Fig.** 4). The controller can be thought of as composed of two parts, an estimator and a feedback regulator. The estimator generates the conditional probability distribution of the state from the measurements. This distribution is called the hyperstate of the problem. The feedback regulator is a nonlinear function which maps the hyperstate into the space of control variables. To solve a nonlinear stochastic control problem, it is necessary to determine the estimator, that is, the formula for updating the hyperstate, and to solve the Bellman equation. The structural simplicity of the solution is obtained at the price of introducing the hyperstate, which is a quantity of very high dimension. For a problem where the state space is R^4 (four-dimensional euclidean space), the hyperstate is, for example, a distribution over R^4. *See* ESTIMATION THEORY; OPTIMAL CONTROL THEORY.

Since it is difficult to solve the Bellman equation, approximative solutions are of considerable interest. A simple example will be used to illustrate some common approximations. Consider a process described by Eq. (2), where u is the control, y the

$$y(t+1) = y(t) + bu(t) + e(t) \tag{2}$$

output, e white noise, and b a constant parameter. Equation (2) can be interpreted as a sampled-data model of an integrator with unknown gain. Let the criterion be to minimize expression (3), where E de-

$$\lim_{N\to\infty} E\frac{1}{N}\sum_{1}^{N} y^2(t) \tag{3}$$

notes the mathematical expectation.

If the parameter b is known, the control law which minimizes (3) is given by Eq. (4). If the parameter

$$u(t) = \frac{1}{b}y(t) \tag{4}$$

b has a gaussian prior distribution, it follows that the conditional distribution of b, given inputs and outputs up to time t, is gaussian with mean $\hat{b}(t)$

and variance $P(t)$. The hyperstate of the problem can then be characterized by the triple $[y(t), \hat{b}(t), P(t)]$. In this simple case, the Bellman equation can be solved numerically. The control law obtained cannot be characterized in a simple way. It has, however, a very interesting property. Not only will the control signal attempt to bring the output close to zero; but also, when the parameters are uncertain, the regulator will inject signals into the system to reduce the uncertainty of the parameter estimates. The optimal control law will give the right balance between keeping the control errors and the estimation errors small. This is called dual control.

The control law given by Eq. (5) is an approximative solution. This control is called the certainty equivalence control. It is obtained simply by solving the control problem in the case of known parameters and substituting the known parameters with their estimates. The self-tuning regulator can be interpreted as a certainty-equivalence control.

$$u(t) = -\frac{1}{\hat{b}(t)} y(t) \tag{5}$$

The control law given by Eq. (6) is another approximation, called cautious control because it hedges and uses low gain when estimates are uncertain. The cautious control law minimizes the criterion given by expression (7), but it is not optimal for Eq. (4).

$$\begin{aligned} u(t) &= -\frac{\hat{b}(t)}{\hat{b}^2(t) + P(t)} y(t) \\ &= -\frac{1}{\hat{b}(t)} \cdot \frac{\hat{b}^2(t)}{\hat{b}^2(t) + P(t)} y(t) \end{aligned} \tag{6}$$

$$E[y^2(t)/y(t-1), y(t-2), \ldots] \tag{7}$$

Status and prospects. The phrase adaptive control is unfortunately used in many different ways and its precise meaning is subject to controversy. The field is nevertheless in a state of rapid development. The availability of microprocessors, which make it possible to implement the controllers economically, has been a strong driving force.

The theory of adaptive control is still in its infancy. Major steps have been taken toward an understanding of the stability problem, but much research is needed to develop the appropriate conceptual framework and the key theoretical problems. *See* CONTROL SYSTEMS.

Karl Johan Åström

Bibliography. K. J. Åström and B. Wittenmark, *Adaptive Control*, 2d ed., 1994; C. C. Hang, T. H. Lee, and W. K. Ho, *Adaptive Control*, 1993; R. Isermann et al., *Adaptive Digital Control Systems*, 1991; P. P. Kokotovic, *Foundations of Adaptive Control*, 1991; K. S. Narendra, R. Ortega, and P. Dorato (eds.), *Advances in Adaptive Control*, 1991.

Adaptive management

An approach to management of natural resources that emphasizes how little is known about the dynamics of ecosystems and that as more is learned management will evolve and improve. Natural systems are very complex and dynamic, and human observations about natural processes are fragmentary and inaccurate. As a result, the best way to use the available resources in a sustainable manner remains to be determined. Furthermore, much of the variability that affects natural populations is unpredictable and beyond human control. This combination of ignorance and unpredictability means that the ways in which ecosystems respond to human interventions are unknown and can be described only in probabilistic terms. Nonetheless, management decisions need to be made. Adaptive management proceeds despite this uncertainty by treating human interventions in natural systems as large-scale experiments from which more may be learned, leading to improved management in the future.

Experimental work. For example, consider the case of a newly developed fishery. Managers must decide how much can be harvested from a given fish population in order to maximize profits without jeopardizing the viability of the population and the sustainability of the fishery. The adequate rate of harvest will depend on the effect that reducing the size of the reproductive population has on the subsequent number of juveniles recruiting annually into the fishery. All fisheries management hinges on identifying the stock size (number of reproductives in the population) that can produce a maximum sustained harvest and also ensure against overexploitation or chance extinction. In order to identify this target stock size, how the number of recruits (new fish added to the population by reproduction each year) varies with stock size must be determined, which requires collecting additional data (see **illus.**). Using adaptive management allows heavy harvesting, thereby lowering the stock size. Doing this management experiment should allow the identification of the appropriate recruitment function, and the progression toward a better-informed management

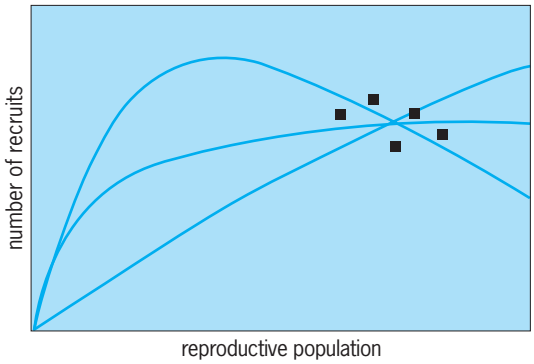

Plausible curves relating number of recruits (squares) to stock size. The level of harvest that a fish population can sustain depends on the relationship between the number of reproductive individuals and the subsequent number of recruits. Several hypotheses are consistent with the available data, each of which would lead to a different management policy. A large-scale management experiment involving harvesting down the population and observing recruitment responses would help to discriminate among the alternative models.

policy. Of course, in the real world this effort is further clouded by the fact that recruitment varies enormously from year to year due to the vagaries of the environment, quite independent of stock size. Delaying action to collect more information will not help, because in order to acquire more information the system must be moved to states at which the different models predict different responses.

This example illustrates a general problem that managers and environmental policy makers frequently face: The information available about natural systems is normally consistent with several competing hypotheses, each of which may dictate a different best course of action. Experimental work and basic science may help to postulate alternative hypotheses about possible consequences of management actions. However, it is generally difficult to extrapolate results from experiments conducted at small temporal and spatial scales directly to the actual ecosystem. Intervention in the real world and conducting large-scale management experiments are ways to learn about methods of management. But to be effective, adaptive management has to go far beyond a simple trial-and-error approach: it must involve the design and implementation of a formal plan for action, a plan by which management decisions are determined based on available knowledge and relevant information gaps, population responses are monitored, and future management is refined based on what is learned from the effects of previous management actions.

Stages. A key first step in the development of an adaptive management program is the assessment of the problem. During this stage, existing knowledge and interdisciplinary experience is synthesized and formally integrated by developing a dynamic model of the system. This modeling exercise helps to identify key information gaps and to postulate hypotheses about possible system responses to human intervention consistent with available information. Different management policies have to be screened in order to narrow down the alternatives to a few plausible candidates. Possible outcomes, including biological, economical, and societal variables of interest, have to be anticipated as much as possible based on the information gathered. The assessment of the problem can be accomplished through a series of workshops with participation of the different stakeholders together with the scientists and managers. Because effective management requires cooperation from stakeholders, their involvement during the initial planning stages is essential.

The second stage involves the formal design of a management and monitoring program. This usually involves the use of models to evaluate several alternative management policies in terms of their ability to meet management goals. To the extent that new information can result in improved future management, adaptive management programs may include large-scale experiments deliberately designed to accelerate learning. Some management actions may be more effective than others at filling the relevant information gaps. In cases where spatial replication is possible (such as small lakes, patches of forest, and reefs), policies that provide contrasts between different management units will be much more informative about the system dynamics than those that apply the same rule everywhere. For example, to control a pest population a predator is released into the field. To learn what level of control is most effective, a number of spatially discrete populations may be selected and different numbers of predators may be released into these units. Without experimental management (such as a single level of control) and without monitoring, little or no information will be gained to improve future management. This works much in the same way as treatment and control work in experimental science. In small-scale research experiments, however, it is much easier to control most sources of variability and to replicate the treatments so that results from the experiments can be unequivocally attributed to the factors tested. In the real world, most sources of variability cannot be controlled, and the effects of changes in environmental conditions tend to be confounded with those of the management actions. In the fishery problem discussed earlier, recruitment might improve as the reproductive population is harvested down; but if the policy is applied to a single population without controls, it will not be possible to determine if that is a consequence of management and not the effect of a more favorable environment. The lack of replicates can be a major impediment to learning through experimental management.

There are other barriers to the implementation of large-scale management experiments. Experiments usually have associated costs; thus, in order to be worthwhile, benefits derived from learning must overcompensate short-term sacrifices. Choices may be also restricted by social concerns or biological constraints, or they may have unacceptably high associated risks. There may be a well-justified reluctance to experiment with high-value, unique ecosystems. Risks can be especially great when management decisions have irreversible consequences, for example, the introduction of an exotic species as a biological control agent or for rearing purposes, which may prove harmful to the environment. The design of adaptive management programs involves anticipating possible consequences and evaluating trade-offs between expected costs and gains associated with the different candidate policies.

Once a plan for action has been chosen, the next stage is to implement the program in the field. This is one of the most difficult steps, because it involves a concerted and sustained effort from all sectors involved in the use, assessment, and management of the natural resources. Beyond the implementation of specific initial actions, putting in place an adaptive management program involves a long-term commitment to monitoring the compliance of the plan, evaluating the effects of management interventions, and adjusting management accordingly.

No matter how thorough and complete the initial assessment and design may have been, systems may always respond in manners that could not be

foreseen at the planning stage. Ecosystems exhibit long-term, persistent changes at the scale of decades and centuries; thus, recent experience is not necessarily a good basis for predicting future behavior. The effects of global climatic change on the dynamics of ecosystems, which are to a large extent unpredictable, will pose many such management challenges. Adaptive management programs have to include a stage of evaluation and adjustment. Outcomes of past management decisions must be compared with initial forecasts, models have to be refined to reflect new understanding, and management programs have to be revised accordingly. New information may suggest new uncertainties and innovative management approaches, leading to another cycle of assessment, design, implementation, and evaluation.

Applications. Adaptive management, although it has always been practiced in a more or less ad hoc manner, was formally developed in the 1970s by C. S. Holling and coworkers at the University of British Columbia, Canada. Since then, it has been applied to a variety of ecological problems involving the management of fisheries and forests, the control of introduced species in conservation problems, the recovery of essential habitat for threatened populations, the regulation of river flows, and the restoration of large-scale ecosystems, among others. Management of natural resources requires a much more extensive application of adaptive management approaches than is typical of current practices. Unless a wider application of adaptive management is begun, inadequate conservation of resources will continue. *See* CONSERVATION OF RESOURCES; ECOSYSTEM; FISHERIES ECOLOGY. Ana M. Parma

Bibliography. A. M. Parma et al., What can adaptive management do for our fish, forests, food, and biodiversity, *J. Integrative Biol.*, 1:16–26, 1998; C. J. Walters, *Adaptive Management of Renewable Resources*, Macmillan, New York, 1986; C. J. Walters, Challenges in adaptive management of riparian and coastal ecosystems, *Conserv. Ecol.* [online], 1:1–9, 1997; C. J. Walters and R. Green, Valuation of experimental management options for ecological systems, *J. Wildlife Manag.*, 1999.

Adaptive optics

The science of optical systems in which a controllable optical element, usually a deformable mirror, is used to optimize the performance of the system, for example, to maintain a sharply focused image in the presence of wavefront aberrations. A distinction is made between active optics, in which optical components are modified or adjusted by external control to compensate slowly changing disturbances, and adaptive optics, which applies to closed-loop feedback systems employing sensors and data processors, operating at much higher frequencies. *See* CONTROL SYSTEMS.

The practical development of adaptive optics started in the late 1960s. Its main applications have been to compensate for the effects of atmospheric turbulence in ground-based astronomical telescopes and to improve the beam quality of high-power lasers. Adaptive optics is now used routinely at several astronomical observatories.

In a typical adaptive optics system (**Fig. 1**), the distorted light beam to be compensated is reflected from the deformable mirror and is sampled by a beam splitter. The light sample is analyzed in a wavefront sensor that determines the error in each part of the beam. The required corrections are computed and applied to the deformable mirror whose surface forms the shape necessary to flatten the reflected wavefront. The result is to remove the optical error at the sampling point so that the light passing through the beam splitter may be focused to a sharp image. Nonlinear optical devices are also capable of performing some adaptive optics functions; these devices operate at high optical power levels. *See* ABERRATION (OPTICS); GEOMETRICAL OPTICS; NONLINEAR OPTICAL DEVICES.

Astronomical applications. There are two main limitations to the performance of large ground-based astronomical telescopes: (1) aberrations in the optical components, particularly the primary mirror, due to temperature variations and mechanical sagging at different pointing angles; and (2) atmospheric turbulence, which distorts the light waves received from distant objects. The angular resolution of a telescope is theoretically λ/D, where λ is the wavelength of the radiation and D is the diameter of the aperture. A 4-m (157-in.) telescope operating at visible wavelengths (0.5 μm) should ideally be capable of separating objects less than 0.03 arc-second apart. In practice, atmospheric turbulence limits the angular resolution of uncompensated telescopes to about 1 second of arc, regardless of their size. Modern telescopes make use of both active and adaptive optics to overcome these limitations.

Active optics. Active optics is employed to maintain the accuracy of the primary mirror surface to a fraction of the wavelength of the radiation being observed. Such active mirrors use either thin monolithic faceplates supported on an array of actuators, or segmented faceplates in which a number of individual, stiff panels are precisely positioned by control of their supporting points. Examples of the monolithic approach are the Very Large Telescope of the European Southern Observatory, which uses four 8.2-m (323-in.) faceplates, each supported by 150 actuators, and the Subaru telescope operated by the National Astronomical Observatory of Japan, which has a single 8.3-m (327-in.) faceplate controlled by 261 actuators. Segmented primary mirrors are used in the two 10-m (400-in.) Keck telescopes, each of which consists of 36 hexagonal elements adjustable in tip, tilt, and piston.

Another important application of active optics is in long-baseline interferometers, in which the optical path lengths of two or more arms must be precisely controlled. This is achieved by a fringe-tracking detector coupled to an actively controlled optical delay line.

Turbulence compensation. Adaptive optics is used in astronomical telescopes to compensate for the optical effects of atmospheric turbulence. Because of the large mass of primary mirrors, it is not feasible to drive them at the high speed necessary for turbulence compensation, which requires measuring and correcting the wavefront at rates between 100 and 1000 times per second. The adaptive optics systems used for this purpose employ small high-speed deformable mirrors, together with real-time wavefront sensors and data processors.

The data to drive the deformable mirror are provided by a very sensitive high-speed wavefront sensor that operates by using the radiation from a star or similar object. There are two fundamental limitations to the use of adaptive optics for turbulence compensation: the photon limit and the size of the isoplanatic patch. To make useful wavefront measurements, it is necessary to collect at least 100 photons from the reference source within the atmospheric change time. At visible wavelengths the limiting magnitude of the reference star is about the 10th visual magnitude while for infrared observations it may be 15th magnitude. The isoplanatic patch is the field of view around the reference star over which the wavefront measurement is valid. Because of turbulent layers high in the atmosphere, this angle is extremely small, often only a few seconds of arc at visible wavelengths. The result of these limitations is that turbulence compensation is possible only in the immediate vicinity of a few bright stars or similar objects.

The spatial characteristics of turbulence are a critical factor in the design of adaptive optics systems. Turbulence strength is specified in terms of the coherence length, r_0, which is defined statistically as the aperture diameter over which the wavefront has a mean square error of 1 radian squared; as the turbulence increases, the value of r_0 gets smaller. The significance of r_0, which is typically 10–20 cm (4–8 in.) at visible wavelengths, is that it determines the size of the compensation elements and consequently the number of actuators required for a given telescope aperture. At visible wavelengths, the number of actuators required to compensate a 4-m (157-in.) telescope is between 500 and 1000. The value of r_0 varies as the 6/5 power of wavelength, so that adaptive optics is much easier to implement at infrared wavelengths. Not only are fewer actuators required, but the turbulence changes at a lower rate and the isoplanatic angle is larger. At wavelengths above 10 μm, adaptive optics is no longer necessary. *See* COHERENCE; INFRARED ASTRONOMY; TURBULENT FLOW.

Laser beacons. The only method of obtaining good sky coverage at shorter wavelengths is the use of laser beacons, which may be created in the upper atmosphere either by backscatter from air molecules (Rayleigh scattering) at 10–40 km (6–25 mi) altitude or by stimulating fluorescence in the sodium layer at around 90 km (55 mi). *See* FLUORESCENCE; SCATTERING OF ELECTROMAGNETIC RADIATION.

The operating sequence (**Fig. 2**) is generally as follows. Laser pulses are projected through the telescope to produce a focused spot of length about 1.5 km (0.9 mi) and diameter about 1 m (3 ft). The backscattered light passes through the turbulent atmosphere, is collected by the telescope, and is

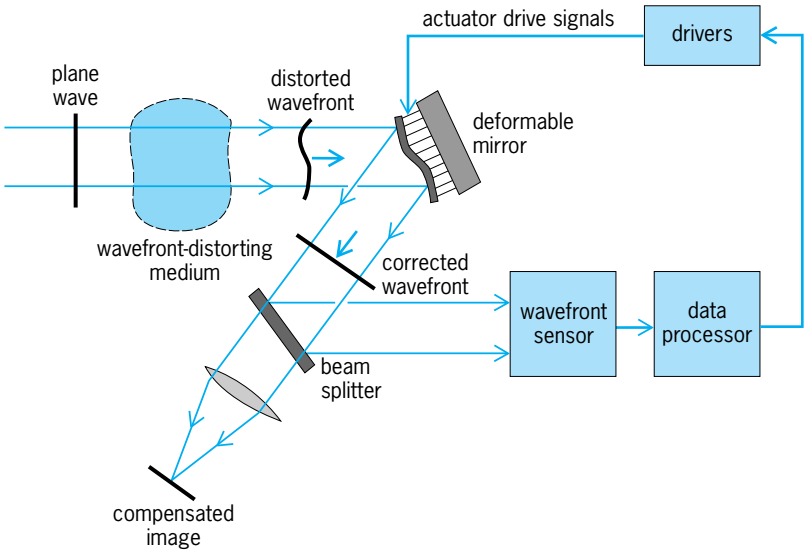

Fig. 1. Typical adaptive optics system using discrete components.

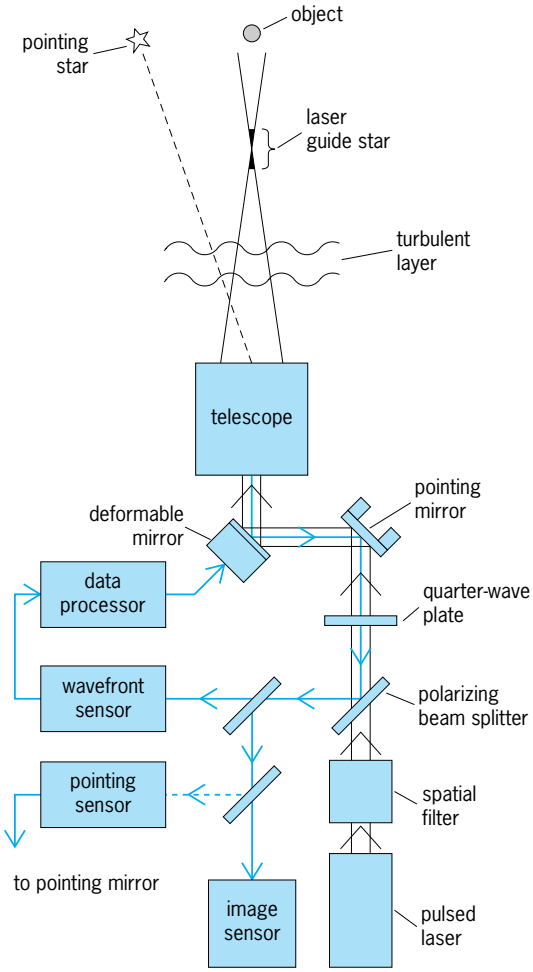

Fig. 2. Adaptive optical system using a laser beacon (guide star).

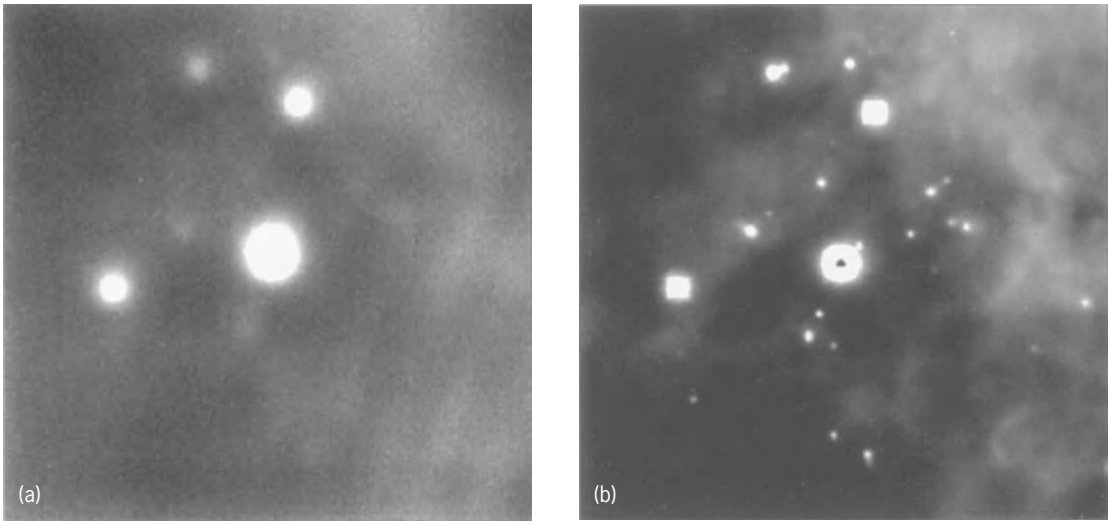

Fig. 3. Images of the Trapezium region in the Orion Nebula, obtained by using adaptive optics with a laser beacon reference source and a 241-actuator deformable mirror on a 1.5-m (59-in.) telescope. Fields of view are 40 arc-seconds square. (*a*) Uncorrected image. (*b*) Real-time image correction with laser beacon adaptive optics. (*Robert Q. Fugate, USAF Phillips Laboratory, Starfire Optical Range*)

directed to the wavefront sensor, which is range-gated to select light scattered from the required altitude. The static focus error of the received wavefront (due to the finite altitude of the laser beacon) is removed, and the residual wavefront errors are measured. The drive signals required to null the wavefront errors are then computed and applied to the deformable mirror. The imaging sensor then records a compensated image of the object (**Fig. 3**).

A laser beacon projected through the telescope cannot provide absolute pointing information, so it is necessary to use a natural star within the field of view as a direction reference; stars as faint as 20th magnitude may be used.

A single laser beacon samples a partial (conical) volume of the telescope beam, resulting in an additional error known as focal anisoplanatism. The magnitude of this error depends on the altitude of the spot and is most serious with lower-altitude Rayleigh scattering.

Many telescopes have been fitted with adaptive optics, including the Mount Wilson 2.5-m (100-in.), the Palomar 5-m (200-in.), the Canada-France-Hawaii 3.6-m (144-in.), the European Southern Observatory 3.6-m (142-in.), and the Lick Observatory 3-m (120-in.), instruments. In addition, adaptive optics will be installed in almost all of the current generation of 6–10-m (240–400-in.) giant telescopes, including the European Southern Observatory Very Large Telescope, Keck 2, Single-Mirror MMT, Gemini, and Subaru instruments. *See* TELESCOPE.

Laser applications. Adaptive optics is used to improve the quality of beams generated in laser cavities, and also to compensate for thermal effects and turbulence in the propagation path. The aberrations can be cleaned up by using an adaptive optics system similar to that discussed above. For this application adequate optical power is usually available to make the wavefront measurement. *See* LASER.

Compensation of laser beams transmitted through a turbulent medium such as the Earth's atmosphere relies on the principle of optical reciprocity. A low-power reference source, at or adjacent to the target, transmits its light through the turbulent medium to the laser aperture, where the wavefront aberrations are measured. A correction for these aberrations is applied to a deformable mirror before the laser is fired. The outgoing beam is predistorted by the deformable mirror and retraces the path taken by the light from the reference source. When it arrives at the target, its predistortion will have been exactly canceled by the atmosphere and so it will be precisely focused.

The isoplanatic limitation also applies to such a system: it is necessary for the reference beam to be aligned with the intended laser path; otherwise the measured wavefront will not be correct. If the target is moving, it is necessary to correct for the finite velocity of light by making the wavefront measurement at a point ahead of the target.

Technology. Adaptive optical systems may be broadly classified into conventional systems employing discrete optical and electronic components, and unconventional systems using integrated devices in which the basic functions are performed by physical processes within a nonlinear medium.

Conventional systems employ three main components; a wavefront sensor, a data processor, and a wavefront compensator.

Wavefront sensors. Most practical wavefront sensors measure the gradient of the wavefront over an array of subapertures in the incoming beam. These local gradient measurements are then reconstructed into a map of the wavefront error over the whole aperture. The two main techniques for gradient sensing use the imaging Hartmann sensor and the shearing interferometer. The imaging Hartmann sensor divides the optical beam into an array of contiguous subapertures, each containing a lens that brings the light from the reference source to a separate focus.

Displacement of the focused spots reveals the direction and magnitude of the two-dimensional wavefront gradient.

The principle of shearing interferometry is to split the wavefront to be measured into two replicas which are mutually displaced by a small distance and then recombined. Lateral displacement or shear is most often used, in which case the intensity of the interference pattern is proportional to the wavefront gradient in the direction of shear. *See* INTERFERENCE OF WAVES; INTERFEROMETRY.

Curvature sensing is an alternative technique that measures the local wavefront curvature (a scalar quantity) rather than the wavefront slope (a vector quantity). The curvature measurements can be applied directly to a deformable mirror, employing bimorph actuators that produce local curvature of the faceplate.

Data processors. Reconstruction of the gradient measurements of wavefront sensors into an error map involves the solution of a set of simultaneous equations, or equivalently the inversion of a matrix. The computation is usually implemented by a digital processor using an array of multiplier-accumulator devices. For large apertures the computational load is considerable. *See* DIGITAL COMPUTER; MATRIX THEORY.

Wavefront compensators. In principle, wavefront compensators may either shift the optical phase of a light beam or change the optical path length. Phase-shift devices such as electrooptical crystals have a limited range of correction and are often spectrally limited. Deformable mirrors employ mechanical displacement of the reflecting surface to control the optical path length. Advantages of deformable mirrors include high reflection efficiency, wide spectral bandwidth, and large correction range. *See* ELECTROOPTICS.

Three main types of deformable mirror have been developed: the monolithic deformable mirror, the thin-plate mirror using discrete actuators, and the segmented mirror. The monolithic mirror employs a solid block of piezoelectric ceramic such as lead zirconium titanate, with an array of electrodes at the top surface, to which is bonded a thin sheet of glass forming the reflecting surface. The thin-plate mirror employs an array of discrete multilayer piezoelectric actuators mounted on a rigid baseplate. Segmented mirrors are composed of many separate panels, each of which is normally supported on three actuators that provide tip, tilt, and piston adjustments. The panels are relatively small and easier to fabricate than one large mirror. However, maintaining the alignment of the segments to maintain a properly phased surface is quite difficult, and a separate wavefront sensor is usually employed for this purpose. To reduce the number of components in the optical train, some compensated telescopes are designed with adaptive secondaries, which employ thin deformable faceplates controlled by several hundred electromagnetic actuators. *See* PIEZOELECTRICITY.

Nonlinear optics. Nonlinear optical devices function as conjugate mirrors in which the output beam is a replica of the input except that it travels in exactly the reverse direction. If a plane wave passes through an aberrating medium, is reflected by a conjugate mirror, and then retraces its path, it will emerge in its original state, once again as a plane wave. The basic mechanism of nonlinear phase conjugation is the formation within a cell of zones of different optical refractivity that form a partially reflecting three-dimensional grating structure. There are two main types of nonlinear devices: (1) those based on stimulated Brillouin scattering and the related stimulated Raman scattering, which are produced in suitable transparent media by the self-pumping effect of high-intensity, high-power laser beams; and (2) three-wave and four-wave mixing devices, in which auxiliary high-power pump beams set up the grating structure, allowing considerable gain to be obtained. *See* NONLINEAR OPTICS; OPTICAL PHASE CONJUGATION; RAMAN EFFECT.

John W. Hardy

Bibliography. D. G. Crowe (ed.), *Selected Papers on Adaptive Optics and Speckle Imaging*, SPIE, vol. MS93, 1994; M. A. Ealey and F. Merkle (eds.), Adaptive Optics in Astronomy, *Proc. SPIE*, vol. 2201, 1994; J. W. Hardy, *Adaptive Optics for Astronomical Telescopes*, Oxford University Press, New York, 1998; J. E. Pearson (ed.), *Selected Papers on Adaptive Optics for Atmospheric Compensation*, SPIE, vol. MS92, 1994.

Adaptive signal processing

Signal processing is a discipline that deals with the extraction of information from signals. The devices that perform this task can be physical hardware devices, specialized software codes, or combinations of both. In recent years the complexity of these devices and the scope of their applications have increased dramatically with the rapidly falling costs of hardware and software and the advancement of sensor technologies. This trend has made it possible to pursue sophisticated signal-processing designs at relatively low cost. Some notable applications, in areas ranging from biomedical engineering to wireless communications, include the suppression of interference arising from noisy measurement sensors, the elimination of distortions introduced when signals travel through transmission channels, and the recovery of signals embedded in a multitude of echoes created by multipath effects in mobile communications.

Statistically based systems. Regardless of the application, any functional system is expected to meet certain performance specifications. The requirements, as well as the design methodology, vary according to the nature of the end application. One distinctive design methodology, which dominated much of the earlier work in the information sciences, especially in the 1950s and 1960s, is based on statistical considerations. This framework assumes the availability of information in advance about the statistical nature of the signals involved, and then proceeds to design systems that optimize some statistical

criterion. The resulting optimal designs are, in general, complex to implement. Only in special, yet important cases have they led to successful breakthroughs culminating with the Wiener and Kalman filters. *See* ESTIMATION THEORY; OPTIMIZATION.

Moreover, in many situations a design that is motivated by statistical considerations may not be immediately feasible, because complete knowledge of the necessary statistical information may not be available. It may even happen that the statistical conditions vary with time. It therefore may be expected, in these scenarios, that the performance of any statistically based optimal design will degrade, the more so as the real physical application deviates from the modeling assumptions.

Adaptive systems. Adaptive systems provide an attractive solution to the problem. They are devices that adjust themselves to an ever-changing environment; the structure of an adaptive system changes in such a way that its performance improves through a continuing interaction with its surroundings. Its superior performance in nonstationary environments results from its ability to track slow variations in the statistics of the signals and to continually seek optimal designs.

Adaptive signal processing deals with the design of adaptive systems for signal-processing applications. Related issues arise in control design, where the objective is to alter the behavior of a system, and lead to the study of adaptive control strategies; the main issue is the stability of the system under feedback. *See* ADAPTIVE CONTROL.

Operation. The operation of an adaptive system can be illustrated with a classical example in system identification. The **illustration** shows a plant (or system) whose input-output behavior is unknown and may even be time variant. The objective is to design an adaptive system that provides a good approximation to the input-output map of the plant. For this purpose, the plant is excited by a known input signal, and the response is taken as a reference signal. Moreover, a structure is chosen for the adaptive system, say a finite-impulse response structure of adequate length, and it is excited by the same input signal as the plant. At each time instant the output of the finite-impulse response system is compared with the reference signal, and the resulting error signal is used to change the coefficients of the finite-impulse response configuration. This learning process is continued over time, and the output of the adaptive system is expected to provide better tracking of the plant output as time progresses, especially when the structure of the plant is time invariant or varies only slowly with time.

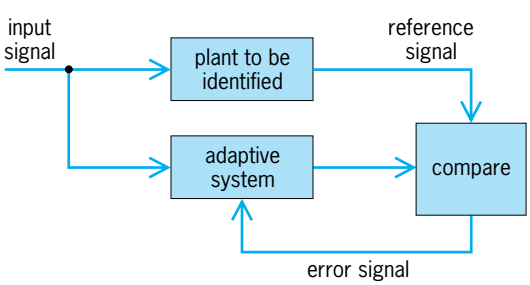

Adaptive system identification.

Characteristics. Apart from emphasizing one particular application for adaptive systems, the above example also highlights several characteristics that are common to most adaptive designs:

1. An adaptive system adjusts itself automatically to a changing environment. This is achieved by changing the parameters of its internal structure in response to an error measure. Several adaptive structures have been used in practice, but the most frequent ones are linear combiners, finite-impulse response filters, infinite-impulse response filters, and linear combiners that are followed by nonlinearities such as sigmoid functions or nonlinear decision devices. *See* ELECTRIC FILTER.

2. The interaction of an adaptive system with its environment takes place through an input signal and a reference signal. The reference signal is used to evaluate the performance of the adaptation process through the computation of an error signal. The adaptive system responds to the error signal in such a way as to minimize a predetermined cost function that is computed either from the error signal directly or from a filtered version of it.

3. An adaptive system is inherently nonlinear and time variant. For this reason, it is in general considerably more difficult to analyze its performance than that of a linear time-invariant system. Nevertheless, adaptive systems offer more possibilities than conventional nonadaptive designs, and many analysis methods have been developed that offer reasonable approximate methods for performance evaluation.

4. An adaptive system can be trained to perform certain tasks. This usually involves a learning phase where the adaptive system is exposed to typical input and reference data and is left to adjust itself to them. At the end of the learning procedure, the system can be exposed to new data and will be expected to provide a reasonable response. Typical examples of this scenario abound in neural-network applications and in the equalization of communication channels. In the latter application, the outputs of communication channels (at the receiver) are equalized to compensate for the distortion of signals sent over long distances, and to thereby estimate the input (or transmitted) signals. *See* NEURAL NETWORK.

Adaptive algorithms. The performance of an adaptive system is critically dependent not only on its internal structure but also on the algorithm used to automatically adjust the parameters that define the structure. In general, the more complex the internal structure of the adaptive system, the more complex the algorithm that is needed for its adaptation. Moreover, several algorithms may exist for the same adaptive structure, and the choice of which algorithm to use is dictated by several factors that include the following:

1. The speed with which the algorithm learns in a stationary environment; that is, how fast the algorithm converges to the optimal statistical solution under stationarity assumptions. This factor determines the convergence rate of the algorithm.

2. The speed with which the algorithm tracks statistical variations. This factor determines the tracking capability of an algorithm and is a major motivation for the use of adaptive schemes, especially in recent mobile communications applications where adaptive equalizers are used to compensate for (that is, adapt to) changes in time-variant transmission channels, and thereby give good estimates of the transmitted signal from the received signal despite these changes.

3. The manner in which the performance of an adaptive scheme, operating in steady-state conditions, deviates from the performance of a statistically optimal design. This factor measures the so-called misadjustment of an adaptive scheme and serves to compare its performance with that of an optimal design in a statistical framework.

4. The amount of computational effort required to carry out each adjustment of the parameters of the adaptive system structure. Applications that require a large number of adaptive parameters tend to dictate a preference for computationally fast algorithms at the expense of other performance factors.

5. The reliability of an algorithm in finite-precision implementations. This factor is concerned with the numerical behavior of an algorithm when implemented in finite-precision arithmetic, and whether numerical effects might lead to erroneous behavior.

6. The robustness of an adaptive system to disturbances and unmodeled dynamics. This factor determines whether small disturbances can result in large errors and therefore compromise the performance of an adaptive scheme.

The above factors are often competing requirements, so that it is usually necessary to seek a compromise solution that best suits a particular application.

Derivation procedures. There have been many procedures for the derivation of adaptive algorithms, but the most frequent, at least in having had the most applications, are procedures that are based either on the method of stochastic approximation or on the least-squares criterion. In both cases, and especially for finite-impulse response adaptive structures, each criterion has led to several different variants that address in one way or another the above performance factors.

In stochastic-approximation algorithms, a recursive procedure is devised for the minimization of a predetermined cost function by iteratively approximating its minimum through a gradient descent procedure (a standard procedure in optimization theory). While this class of algorithms generally suffers from slow convergence rates, it still enjoys widespread use due to its simplicity, low computational requirements, and often observed robustness under different operating conditions. The most prominent member of the stochastic-approximation algorithms is the least-mean-squares algorithm, which is undoubtedly the most widely used adaptive filtering algorithm. Other members include filtered-error variants that are useful in active noise control and infinite-impulse response system identification problems, as well as several frequency-domain adaptive schemes. *See* ACTIVE SOUND CONTROL; STOCHASTIC CONTROL THEORY.

In least-squares algorithms, a recursive procedure is devised for the minimization of a quadratic cost function. This family of algorithms is based on the least-squares criterion, which was developed in the late eighteenth century in work on celestial mechanics. Since then, the least-squares criterion has enjoyed widespread popularity in many diverse areas as a result of its attractive computational and statistical properties. Notably, for linear data models, least-squares solutions can be explicitly evaluated in closed form, can be recursively updated as more input data are made available, and are optimal maximum likelihood estimators in the presence of gaussian measurement noise. *See* LEAST-SQUARES METHOD.

Many recursive least-squares algorithms have been developed. Several of them are computationally more demanding than least-mean-squares-type algorithms, but variants exist that are computationally competitive, although more complex. They have better convergence properties but are less robust to disturbances.

Array algorithms. Following a trend initiated in the late 1960s in Kalman filtering methods, least-squares adaptive schemes are currently more often implemented in convenient array algorithms. These algorithms are closely related to the QR method, a numerically stable algorithm for solving systems of linear equations, and have the properties of better conditioning (that is, lower sensitivity to errors in the initial data), reduced dynamical range (which favors better conditioning), and orthogonal transformations, which typically lead to better numerical performance in finite-precision arithmetic. In the array form, an algorithm is described as a sequence of elementary operations on arrays of numbers. Usually, a prearray of numbers has to be triangularized by a rotation, or a sequence of elementary rotations, in order to yield a postarray of numbers. The quantities needed to form the next prearray can then be read off from the entries of the postarray, and the procedure can be repeated. The explicit forms of the rotation matrices are not needed in most cases. Such array descriptions are more truly algorithms in the sense that they operate on sets of numbers and provide other sets of numbers, with no explicit equations involved. *See* ALGORITHM.

Ali H. Sayed; Babak Hassibi; Thomas Kailath

Bibliography. S. Haykin, *Adaptive Filter Theory*, 3d ed. 1996; J. G. Proakis et al., *Advanced Digital Signal Processing*, 1992; B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, 1985.

Adaptive wings

A wing is the primary lift-generating surface of an aircraft. Adaptive wings—also known as smart, compliant, intelligent, morphing, controllable, and reactive wings—are lifting surfaces that can change their shape in flight to achieve optimal performance at different speeds, altitudes, and ambient conditions. There are different levels of sophistication, or intelligence, that can be imbued in a particular design.

Whether the wing is rigidly attached to the fuselage for fixed-wing airplanes or rotating for helicopters, a primary design objective of such a lifting surface is to maximize the lift-to-drag ratio, which is achieved by controlling the airflow around the wing. Other design objectives for the wing include improving maneuverability and minimizing vibrations and flow-induced noise. The wing can have a set design optimized for specific flight conditions, or it can change shape to conform to a variety of conditions. Chosen judiciously, minute dynamic changes in the wing's shape can, under the right circumstances, greatly affect the airflow and thus the aircraft's performance. *See* AERODYNAMIC FORCE; AIRFOIL.

Flying insects and birds, through millions of years of evolution, can change the shape of their flapping wings, subtly or dramatically, to adapt to various flight conditions. The resulting performance and agility are unmatched by any airplane. For example, the dragonfly can fly forward and backward, turn abruptly and perform other supermaneuvers, hover, feed, and even mate while aloft (**Fig. 1**). Undoubtedly, its prodigious wings contributed to the species' survival for over 250 million years. *See* FLIGHT.

Among human-produced flyers, the Wright brothers changed the camber of the outboard tip of their aircraft's wings to generate lateral or roll control (combined with simultaneous rudder action for banked turn), thus achieving in 1903 the first heavier-than-air, controlled flight. The R. B. Racer built by the Dayton Wright Airplane Company in 1920 allowed the pilot to change the wing camber in flight using a hand crank. The wings of today's commercial aircraft contain trailing-edge flaps and leading-edge slats to enhance the lift during the relatively low speeds of takeoff and landing, and ailerons for roll control, all engaged by the pilot via clumsy, heavy, and slow servomechanisms. To equalize the lift and prevent rolling in forward flight, the rotary wings of most helicopters are cyclically feathered to increase the pitch on the advancing blade and decrease it on the retreating blade. *See* AILERON; ELEVATOR (AIRCRAFT); FLIGHT CONTROLS; HELICOPTER.

Fig. 1. Male and female Cardinal Meadowhawk dragonflies following airborne mating. The male has towed the just-inseminated female to a pond and is dipping her tail in the water so she can deposit her eggs. (*Reprinted with permission, The Press Democrat, Santa Rosa, CA*)

While bird wings are quite smart, human-designed ones are not very intelligent. With few exceptions, the level of autonomous adaptability sought in research laboratories is some years away from routine field employment. Intelligent control of the wing shape involves embedded sensors and actuators with integrated control logic; in other words, the wing's skin is made out of smart materials. The sensors detect the state of the controlled variable, for example the wall shear stress, and the actuators respond to the sensors' output based on a control law to effect the desired in-flight metamorphosing of the wing. For certain control goals, for example skin-friction drag reduction, required changes in the wing shape can be microscopic. For others, for example morphing the wing for takeoff and landing, dramatic increases in camber may be needed.

Smart materials. Adaptive wing design involves adding smart materials to the wing structure and using these materials to effect flow changes. Smart materials are those that undergo transformations through physical interactions. Such materials sense changes in their environment and adapt according to a feedforward or feedback control law. Smart materials include piezoelectrics, electrostrictors, magnetostrictors, shape-memory alloys, electrorheological and magnetorheological fluids, and optical fibers. For no rational reason, several other types of sensors and actuators that fall outside those categories are not usually classified as constituting elements of smart structures.

The piezoelectric effect is displayed by many noncentrosymmetric ceramics, polymers, and biological systems. The direct effect denotes the generation of electrical polarization in the material in response to mechanical stress; the poled material is then acting as a stress or strain sensor. The converse effect denotes the generation of mechanical deformation upon the application of an electrical charge; in this case the poled material is acting as an actuator. The most widely used piezoceramic and piezopolymer are, respectively, lead zirconate titanate (PZT) and polyvinylidene fluoride (PVDF). Piezoelectrics are the most commonly used type of smart materials and the only ones that can be used readily as both sensors and actuators. *See* PIEZOELECTRICITY.

Electrostrictive materials are dielectrics that act similarly to piezoelectric actuators, but the relation between the electric charge and the resulting deformation in this case is nonlinear. Examples of such materials are lead magnesium niobate (PMN) compounds, which are relaxor ferroelectrics. Magnetostrictive materials, such as Terfenol-D, are actuators that respond to a magnetic field instead of an electric field. *See* ELECTROSTRICTION; MAGNETOSTRICTION.

Shape memory alloys, such as a nickel–titanium alloy known as Nitinol, are metal actuators that can sustain large deformation and then return to their original shape by heating without undergoing plastic deformation. Electrorheological and magnetorheological fluids rapidly increase in viscosity—by several orders of magnitude—when placed in, respectively, electric or magnetic fields. Both kinds of fluids can provide significant actuation power and are therefore considered for heavy-duty tasks such as shock absorbing for large-scale structures. Finally, optical fibers are sensors that exploit the refractive properties of light to sense acoustical, thermal, and mechanical-strain perturbations. *See* FIBER-OPTIC SENSOR; SHAPE MEMORY ALLOYS.

Outstanding issues to be resolved before smart materials for aircraft wings become routine include cost; complexity; computer memory, speed, and software; weight penalty; maintenance; reliability; robustness; and the integrity of the structure on which the sensors and actuators are mounted. Sensors and actuators that have length scales between 1 and 1000 micrometers constitute a special domain of smart materials that in turn is a cornerstone of micro-electro-mechanical systems (MEMS). *See* MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS).

Flow control. An external wall-bounded flow, such as that developing on the exterior surfaces of a wing, can be manipulated to achieve transition delay, separation postponement, lift increase, skin-friction and pressure drag reduction, turbulence augmentation, mixing enhancement, and noise suppression. These objectives are interrelated (**Fig. 2**). If the boundary layer around the wing becomes turbulent, its resistance to separation is enhanced and more lift can be obtained at increased incidence. On the other hand, the skin-friction drag for a laminar boundary layer can be as much as an order of magnitude less than that for a turbulent one. If transition is delayed, lower skin friction and lower flow-induced noise are achieved. However, a laminar boundary layer can support only very small adverse pressure gradients without separation and, at the slightest increase in angle of attack or some other provocation, the boundary layer detaches from the wing's surface and subsequent loss of lift and increase in form drag occur. Once the laminar boundary layer separates, a free-shear layer forms, and for moderate Reynolds numbers transition to turbulence takes place. Increased entrainment of high-speed fluid due to the turbulent mixing may result in reattachment of the separated region and formation of a laminar separation bubble. At higher incidence, the bubble breaks down, either separating completely or forming a longer bubble.

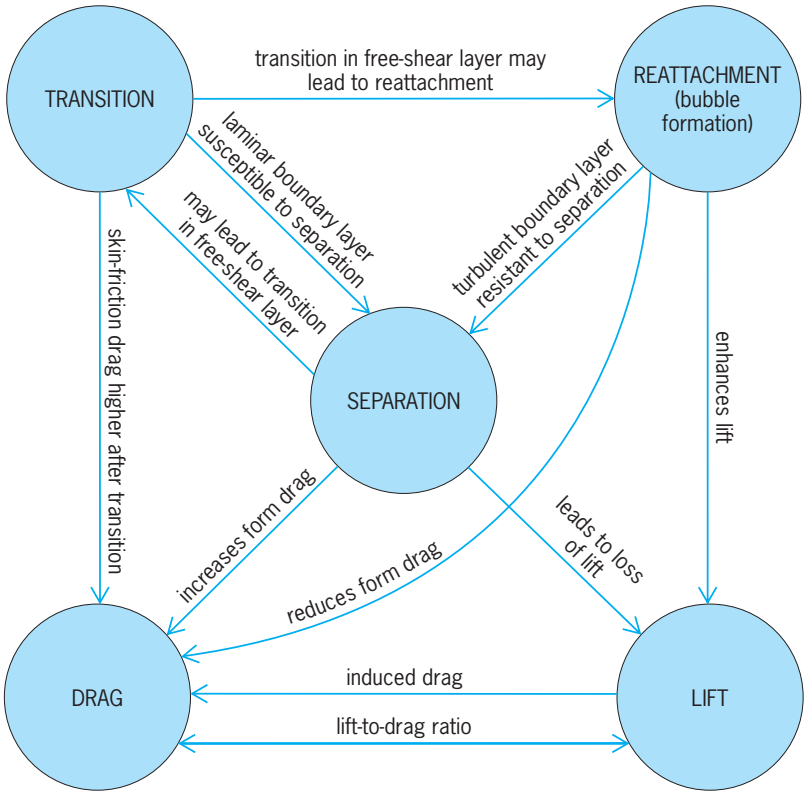

Fig. 2. Partial representation of the interrelation among flow control goals. (*After M. Gad-el-Hak, Flow Control: Passive, Active and Reactive Flow Management, Cambridge University Press, London, 2000*)

In either case, the form drag increases and the lift curve's slope decreases. The ultimate goal of all this is to improve the airfoil's performance by increasing the lift-to-drag ratio. However, induced drag is caused by the lift generated on a wing with a finite span. Moreover, more lift is generated at higher incidence, but form drag also increases at these angles. *See* BOUNDARY-LAYER FLOW; FLUID MECHANICS; LAMINAR FLOW; REYNOLDS NUMBER; TURBULENT FLOW.

Flow control is most effective when applied near the transition or separation points, where conditions are near those of the critical flow regimes where flow instabilities magnify quickly. Therefore, delaying or advancing the laminar-to-turbulence transition and preventing or provoking separation are easier tasks to accomplish. Reducing the skin-friction drag in a nonseparating turbulent boundary layer, where the mean flow is quite stable, is a more challenging problem. Yet, even a modest reduction in the fluid resistance to the motion of, for example, the worldwide commercial airplane fleet is translated into annual fuel savings estimated to be in billions of dollars. Newer ideas for turbulent flow control focus on targeting coherent structures, which are quasiperiodic, organized, large-scale vortex motions embedded in a random, or incoherent, flow field (**Fig. 3**).

Future systems. Future systems for control of turbulent flows in general and turbulent boundary layers in particular could greatly benefit from the merging of the science of chaos control, the technology of microfabrication, and the newest

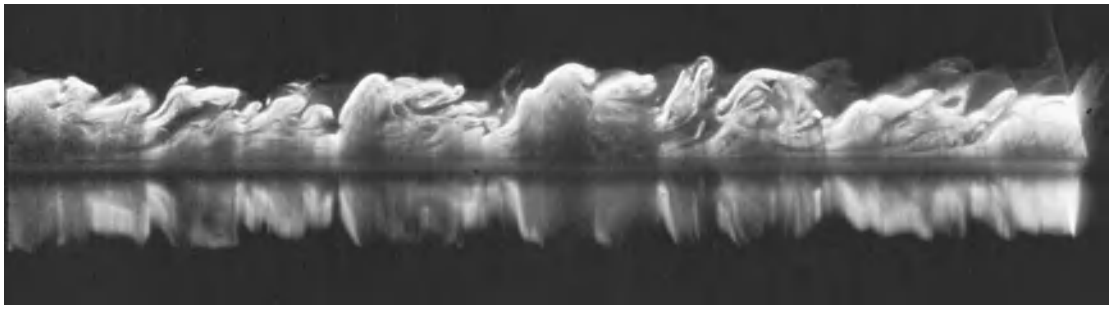

Fig. 3. Large-eddy structures (one type of coherent structure) in a turbulent boundary layer. The side view is visualized using a sheet of laser light and a fluorescent dye. (*From M. Gad-el-Hak, R. F. Blackwelder, and J. J. Riley, On the interaction of compliant coatings with boundary layer flows, J. Fluid Mech., 140:257–280, 1984*)

computational tools collectively termed soft computing. Control of chaotic, nonlinear dynamical systems has been demonstrated theoretically as well as experimentally, even for multi-degree-of-freedom systems. Microfabrication is an emerging technology which has the potential for mass-producing inexpensive, programmable sensor–actuator chips that have dimensions of the order of a few micrometers. Soft computing tools include neural networks, fuzzy logic, and genetic algorithms. They have advanced and become more widely used in the last few years, and could be very useful in constructing effective adaptive controllers. Such futuristic systems are envisaged as consisting of a colossal number of intelligent, interactive, microfabricated wall sensors and actuators arranged in a checkerboard pattern and targeted toward specific organized structures that occur quasirandomly (or quasiperiodically) within a turbulent flow. Sensors would detect oncoming coherent structures, and adaptive controllers would process the sensors' information and provide control signals to the actuators that in turn would attempt to favorably modulate the quasiperiodic events. A finite number of wall sensors perceives only partial information about the flow field. However, a low-dimensional dynamical model of the near-wall region used in a Kálmán filter can make the most of this partial information. Conceptually all of that is not too difficult, but in practice the complexity of such control systems is daunting and much research and development work remain. *See* CHAOS; COMPUTATIONAL INTELLIGENCE; ESTIMATION THEORY; FUZZY SETS AND SYSTEMS; GENETIC ALGORITHMS; MICROSENSOR; NEURAL NETWORK.

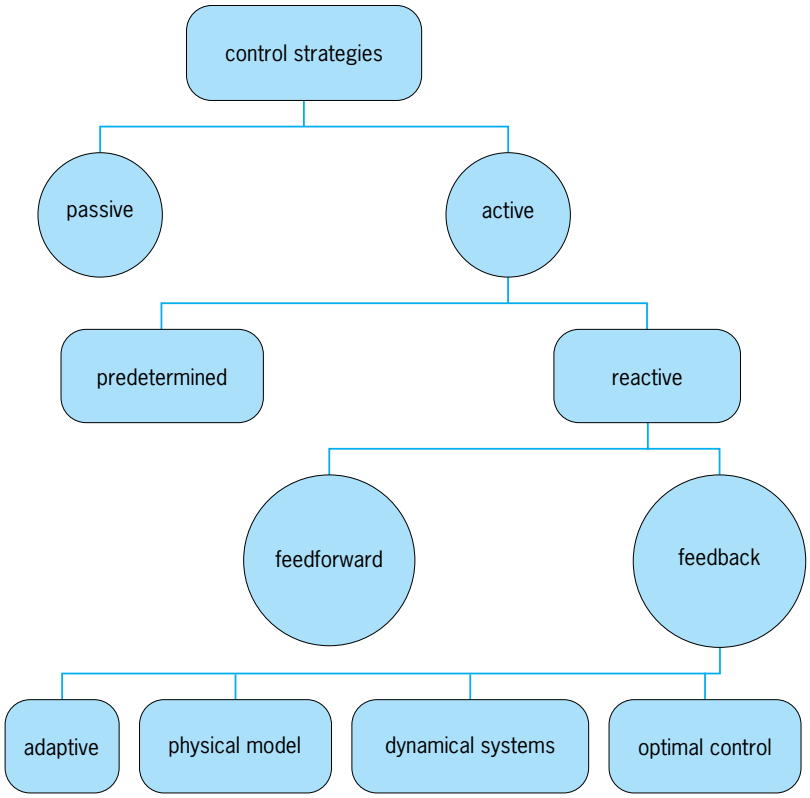

Fig. 4. Classification of control strategies. (*After M. Gad-el-Hak, Flow Control: Passive, Active and Reactive Flow Management, Cambridge University Press, London, 2000*)

Control strategies. Different levels of intelligence can be imbued in a particular control system (**Fig. 4**). The control can be passive, requiring no auxiliary power and no control loop, or active, requiring energy expenditure. Manufacturing a wing with a fixed streamlined shape is an example of passive control. Active control requires a control loop and is further divided into predetermined or reactive. Predetermined control includes the application of steady or unsteady energy input without regard to the particular state of the system—for example, a pilot engaging the wing's flaps for takeoff. The control loop in this case is open, and no sensors are required. Because no sensed information is being fed forward, this open control loop is not a feedforward one. This subtle point is often confused, blurring predetermined control with reactive, feedforward control.

Reactive, or smart, control is a special class of active control where the control input is continuously adjusted based on measurements of some kind. The control loop in this case can be either an open, feedforward one or a closed, feedback loop. Achieving that level of autonomous control (that is, without human interference) is the ultimate goal of smart-wing designers. In feedforward control, the measured variable and the controlled variable are not necessarily the same. For example, the pressure can be sensed at an upstream location, and the resulting signal is used together with an appropriate control law to actuate a shape change that in turn influences the shear stress (that is, skin friction) at

a downstream position. Feedback control, on the other hand, necessitates that the controlled variable be measured, fed back, and compared with a reference input. Reactive feedback control is further classified into four categories: adaptive, physical model–based, dynamical systems–based, and optimal control. An example of reactive control is the use of distributed sensors and actuators on the wing surface to detect certain coherent flow structures and, based on a sophisticated control law, subtly morph the wing to suppress those structures in order to dramatically reduce the skin-friction drag.

Prospects. It is anticipated that the first field applications of adaptive wings will probably take place in micro-air-vehicles (MAV) and unmanned aerial vehicles (UAV). The next generations of cruise missiles and supermaneuverable fighter aircraft will probably have adaptive wings. In the 2010s or 2020s, commercial aircraft may have smart wings, fuselage, and vertical and horizontal stabilizers. *See* DRONE; WING.

Mohamed Gad-el-Hak

Bibliography. H. T. Banks, R. C. Smith, and Y. Wang, *Smart Material Structures: Modeling, Estimation and Control*, Wiley, New York, 1996; M. Gad-el-Hak, *Flow Control: Passive, Active and Reactive Flow Management*, Cambridge University Press, London, 2000; M. Gad-el-Hak (ed.), *The MEMS Handbook*, vols. 1–3, 2d ed., Taylor & Francis/CRC Press, Boca Raton, FL, 2006; M. Gad-el-Hak, R. F. Blackwelder, and J. J. Riley, On the interaction of compliant coatings with boundary layer flows, *J. Fluid Mech.*, 140:257–280, 1984; M. Schwartz (ed.), *Encyclopedia of Smart Materials*, vols. 1 and 2, Wiley-Interscience, New York, 2002.

Addictive disorders

Addictive disease disorders are characterized by the chronic use of a drug (such as heroin, cocaine, or amphetamines), alcohol, or similar substance. These disorders usually result in (1) the development of tolerance for the substance, with the need for increasing amounts to achieve the desired effect; (2) physical dependence, characterized by a sequence of well-defined signs and physiological symptoms, such as the withdrawal or abstinence syndrome on cessation of use of the substance; and (3) compulsive drug-seeking behavior, with chronic, regular, or intermittent use, despite possible harm to self or others. Since the early 1960s, research has been increasing in the biology of addictive diseases, and emphasis has shifted from only psychological, sociological, and epidemiological studies to investigations of the metabolic, neurobiological, and molecular bases of addiction.

The four major addictive diseases are alcoholism, narcotic (or opiate) addiction, cocaine and other stimulant addiction, and nicotine addiction. Drug addiction may also occur after chronic use of other types of agents such as barbiturates, benzodiazepines, and marijuana. Abusive use of other, licit and illicit drugs and other chemical agents is common. However, many types of drug abuse, such as hallucinogen abuse, do not necessarily result in drug addiction but only in drug-using behavior. *See* ALCOHOLISM; BARBITURATES; COCAINE; NARCOTIC; OPIATES.

Opiate receptors and endogenous opioids. The conclusive identification of opiate receptors (cell structures that function as an intermediary between the opioid and the physiological response) in mammals (including humans) in 1973 has been extremely important for many aspects of physiology and pathology. Subsequently, it has been determined by use of selective chemical ligands that there are at least three separate types of opioid receptors—mu receptors, delta receptors, and kappa receptors—all of which meet the accepted criteria of being specific opiate receptors. Opioid (opiatelike) drugs of different types bind to these receptors. This binding may be displaced by a specific opioid (or narcotic) antagonist such as naloxone, and this binding leads to some demonstrable effect. The genes encoding each of these receptor types were cloned for the first time in 1992, beginning with the delta opioid receptor.

Subsequent to the discovery of specific opioid receptors, endogenous ligands which bind to these receptors, the so-called endogenous (originating from within) opioids were discovered. Opioids include substances that are produced endogenously (such as the enkephalins, endorphins, and dynorphins) and may be produced synthetically. Exogenous synthetic opioids are used extensively in the treatment of pain.

The first of these endogenous opioids to be found were met-enkephalin and leu-enkephalin, which are now known to be derived from a single gene product, preproenkephalin. This propeptide is further processed to yield at least five active endogenous opioids. It is produced in the adrenal medulla as well as in the brain and in possibly many other peripheral sites.

The second class of endogenous opioids to be delineated was beta endorphin. This has been shown to be biotransformed from beta lipotropin in equimolar amounts with the release of adrenocorticotropic hormone (ACTH; stimulates the adrenal cortex). Each of these peptides is derived from a single precursor peptide, proopiomelanocortin (POMC). These peptides are found also predominantly in the anterior pituitary. However, the POMC gene is also expressed in the hypothalamus and in other regions of the brain, as well as at peripheral sites.

The third class of endogenous opioids to be delineated were the dynorphins. They consist of multiple active compounds coming from a precursor peptide, preprodynorphin, found in the brain, the spinal cord, and other sites. *See* ENDORPHINS.

Role of endogenous opioids. It is not known to what extent the endogenous opioids play a role in addictive diseases. This is especially true for narcotic (primarily heroin) addiction and alcoholism, and more recently for cocaine addiction. It has been suggested that narcotic addiction may be a disorder characterized by (1) a relative or absolute deficiency of

endogenous opioids; (2) an end organ or receptor failure to respond to normal or possibly increased levels of endogenous opioids; (3) genetic or acquired (for example, short-acting opiate or stimulant drug–induced) abnormalities in the feedback control of the synthesis, release, and processing, or degradation of one or more types of the endogenous opioids; or (4) genetic variations in the opioid receptors. It has also been documented that acute or chronic use of short-acting narcotics, such as heroin or morphine, results in a suppression of levels of beta endorphin as well as levels of ACTH. These two peptides are released together from their precursor peptide. Cortisol (a steroid hormone produced by the adrenal cortex) levels are also lowered and their diurnal rhythm is blunted during use of short-acting narcotics. Conversely, during chronic long-term administration of the long-acting narcotic methadone, in the maintenance treatment of addiction in humans, normalization of these abnormalities occurs, with restoration of normal levels of beta endorphin, ACTH, and cortisol. Also, the normal circadian rhythm of these hormones is restored. Such normalization occurs at a time when there is no longer any drug-seeking behavior, although a high degree of tolerance to mu opioid receptor agonists and of physical dependence exists. When the beta endorphin–ACTH–cortisol levels become normalized, other physiological alterations caused by chronic short-acting narcotic use, such as suppression of luteinizing hormone (secreted by the anterior pituitary, and participating in the menstrual cycle) and peripheral reproductive steroid hormones, also become normalized. Menstrual function and fertility return to normal in women under chronic methadone treatment. Studies have shown that immune function disrupted during heroin addiction may normalize during methadone maintenance treatment. *See* ENDOCRINE MECHANISMS.

It has been found that full tolerance does not develop to one well-documented physiological effect of the acute and subacute administration of any narcotic—that is, the effect of exogenous opioid on release of prolactin (a hormone of the anterior pituitary that stimulates lactation in females). Even during chronic long-term steady-state methadone maintenance treatment in which sustained plasma levels of methadone are achieved, full tolerance does not develop to the prolactin-releasing effect of peak levels of opioid. The peak plasma levels of prolactin occur around the time of peak plasma levels of methadone, that is, around 2–4 h after an oral dose of methadone, rather than in the early morning hours, the normal time of peak levels of prolactin. However, the actual elevations in levels of prolactin caused by the opioids do not, in most cases, exceed the upper limits of prolactin normally observed during spikes in prolactin levels. Since prolactin release is normally under the tonic inhibitory control by dopamine or dopaminergic mechanisms, these findings suggest that even during chronic steady-state treatment methadone suppresses dopaminergic activity, and that tolerance does not develop to this effect during chronic use. It is also of interest that several psychotropic (affecting the mind) drugs, especially the neuroleptic agents (substances that act on the nervous stystem such as antipsychotics), have similar effects on prolactin release, and tolerance apparently does not develop to this effect of neuroleptics. The mechanism underlying the prolactin release by neuroleptics may be similar to the mechanism of action of the exogenous opioids. Both exogenous and endogenous kappa opioid receptor–preferring ligands (such as dynorphins) may modulate prolactin release, as well as exogenous and endogenous mu opioid receptor–directed ligands (such as beta endorphin).

The possible role of endogenous opioids in alcoholism is not clear. It has been shown that a specific opioid antagonist (inhibits the effect of the drug), naloxone, may reverse or ameliorate some of the signs and symptoms and physical abnormalities of the acute alcohol intoxication syndrome. However, this apparent beneficial effect of an opioid antagonist may not be related to the addictive disease per se, but may be due to the counteracting of the possible acute release of large amounts of endogenous opioids in response to excessive acute alcohol ingestion. Theories with some experimental data have further linked alcoholism and narcotic addiction. It has been shown that acetaldehyde, the first major metabolite of ethyl alcohol, can condense in the body with endogenous amines, particularly some specific neurotransmitters, to form opiatelike compounds. Laboratory studies have also shown that a very large amount of ethanol will displace some endogenous opioids from binding to the opioid receptor. Several recent studies of these specific opioid antagonists, naltrexone and nalmefene, may be beneficial in treating chronic alcoholism. Laboratory-based clinical studies have suggested that part of these beneficial effects of opioid antagonists in the management of alcoholism may be mediated through the effects of the endogenous opioid system in modulating the stress-responsive hypothalamic-pituitary-adrenal axis. Whether there are some common mechanisms with a genetic, metabolic, or purely behavioral basis underlying these two addictive diseases has not yet been defined.

Treatment. There are three approaches to the management of opiate (primarily heroin) addiction: pharmacological treatment, drug-free treatment following detoxification, and incarceration. Statistical data over the last century reveal that after release from prison, completion of drug-free treatment, or discontinuation of chronic pharmacological treatment with methadone or other mu agonists, partial agonists, or antagonists, fewer than 30% of former long-term heroin addicts are able to stay drug-free. (Long-term addiction is defined as more than 1 year of daily use of several doses of heroin, with tolerance, physical dependence, and drug-seeking behavior.)

Methadone. Since the 1960s, methadone maintenance treatment has been documented to be medically safe and effective, and the most effective available treatment for heroin addiction. Unlike heroin, which has a 3-min half-life in humans, and its major

metabolite, morphine, which has a half-life of 4–6 h, methadone has a 24-h half-life. Therefore, when given orally on a daily basis, methadone prevents the signs and symptoms of narcotic abstinence and also prevents drug hunger, without causing any narcotic-induced euphoria or high during the 24 h between doses. Steady moderate to high doses of methadone can be used in the treatment of addiction over long periods of time without tolerance developing to these desired effects. Because of the high degree of cross-tolerance developed for other narcotics, a patient receiving methadone maintenance treatment does not experience any narcotic high or other effects after self-administration of a dose of short-acting narcotic such as heroin. Through this cross-tolerance mechanism, methadone blocks any illicit narcotic effect.

Over 60–90% of heroin addicts who have entered into methadone maintenance treatment will stay in treatment voluntarily for 1 year or more. Illicit opiate abuse drops to less than 15% during such treatment when adequate doses of methadone (usually 60–150 mg per day) are administered. Cocaine or polydrug abuse or alcohol abuse may persist in 20–30% of patients even in excellent methadone programs, since methadone maintenance is specific for treatment of narcotic dependency.

Chronic methadone treatment also brings about the normalization of the multiple physiological alterations caused by chronic heroin abuse. Methadone-maintained patients, who are not chronic abusers of other drugs or alcohol, are able to work, to attend school, and to take part in normal socialization, including restoration of family life. Long-term follow-up studies have shown that methadone is medically safe. All these effects are due to the long-acting pharmacokinetic properties of methadone, which yields no euphoria or high, yet does not allow any withdrawal or sick period during a 24-h dosing interval. However, when maintenance treatment is discontinued following slow or rapid dose reduction and elimination, using methadone alone or combined with other agents such as clonidine, over 70% of all patients will return to opiate abuse within 2 years. Few studies focusing on physiological and medical effects have been conducted, but the longer-acting congener (related substance) of methadone, l-alpha-acetylmethadol (LAAM), has been found to be similarly effective and medically safe in the treatment of heroin addiction.

Naltrexone. A second type of pharmacological treatment of narcotic addiction is chronic treatment with a specific narcotic antagonist, naltrexone. This drug also prevents any narcotic effect from any illicitly administered heroin, but does so by a different mechanism than methadone: by way of direct opioid receptor blocking, with displacement of endogenous or exogenous opioids from receptor binding. Naltrexone will precipitate narcotic withdrawal symptoms in narcotic-dependent individuals, and thus may be given only after detoxification from heroin has been completed. Naltrexone will not prevent drug hunger. Naltrexone treatment has limited acceptance by unselected opiate addicts; only 15–30% of former narcotic addicts entering treatment with this agent remain in treatment for 6 months or more. Therefore, although it may be a worthwhile treatment for some small defined special populations, it does not seem to be an effective treatment for the majority of unselected long-term heroin addicts. Naltrexone does not allow the normalization of all physiological functions disrupted by chronic heroin use. However, it may be useful in the treatment of early, sporadic or intermittent narcotic abusers.

Drug-free treatment. Drug-free treatment involving either long-term institutionalization in a drug-free residence or, less frequently, attendance at an outpatient resource, has resulted in long-term success in approximately 10–30% of all unselected heroin addicts who enter treatment. Many addicts remain in treatment only a very short time, and unfortunately sometimes are not included in calculations of treatment outcome. Most drug-free treatments include a 1–3-year residential period followed usually by a 1–3-year nonresidential follow-up period. Only a very small percentage of those who initially enter treatment (though a larger percentage complete this 2–5 years of treatment) will remain drug-free 1 year or more after discharge from treatment.

Alcoholism. Alcoholism has been difficult to treat over a long-term period. There is no specific pharmacological replacement treatment for alcoholism. Disulfiram (Antabuse) is an agent which blocks the metabolism of acetaldehyde, the major metabolite of ethanol. When a person treated with Antabuse drinks alcohol, there is a rapid buildup of acetaldehyde and a severe physiological syndrome ensues, which frequently prevents or modifies further immediate drinking behavior. Most former alcoholics stop the use of disulfiram prior to returning to alcohol use.

The most widely used treatment of alcoholism is self-help groups, such as Alcoholics Anonymous (AA), where mutual support is available on a 24-h basis, along with group recognition of chronic problems with alcohol. However, only around 20–40% of severe chronic alcoholics are able to stay alcohol-free for more than 2 years even under such management. Early studies of the use of an opioid antagonist naltrexone or nalmefene for the treatment of chronic alcoholism have shown that about 50% of the subjects had reduced numbers and magnitudes of relapse events. Ethanol, unlike heroin and other narcotics, has many well-defined severe specific toxic effects on several organs, including the liver, brain, and reproductive system, so even relatively short drug-free intervals will decrease exposure to this toxin. Therefore, it has been suggested that success in treatment of alcoholism be measured in part by increasing lengths of alcohol-free intervals, and not just by permanent restoration of the alcohol-free state.

Cocaine addiction. Cocaine addiction may also involve disruption of the endogenous opioid system in addition to the well-known primary effect of cocaine in blocking reuptake of dopamine by the synaptic

dopamine transporter protein. This effect results in the accumulation of dopamine in the synapse and similar actions at the serotonin and norepinephrine reuptake transporter. Demonstrated changes in the mu and kappa endogenous opioid system increase the complexity of the effect of cocaine and may contribute to its resultant reinforcing properties leading to addiction, as well as to the drug craving and relapse. This may explain in part the resultant difficulties in developing a pharmacotherapeutic approach for the treatment of cocaine addiction. Mary Jeanne Kreek

Bibliography. C. Bond et al., Single nucleotide polymorphism in the human mu opioid receptor gene alters beta-endorphin binding and activity: Possible implications for opiate addiction, *Proc. Nat. Acad. Sci.*, 95:9608–9613, 1998; M. J. Kreek, Clinical update of opioid agonist and partial agonist medications for the maintenance treatment of opioid addiction, *Sem. Neurosci.*, 9:140–157, 1997; M. J. Kreek, Opiate and cocaine addictions: Challenge for pharmacotherapies, *Pharm. Biochem. Behav.*, 57:551–569, 1997; M. J. Kreek, Opiates, opioids and addiction, *Mol. Psych.*, 1:232–254, 1996; M. J. Kreek, Opioid receptors: Some perspectives from early studies of their role in normal phsyiology, stress responsivity and in specific addictive diseases, *J. Neurochem. Res.*, 21:1469–1488, 1996; J. H. Lowinson et al. (eds.), *Substance Abuse: A Comprehensive Textbook*, 3d ed., Williams & Wilkins, 1997; C. P. O'Brien and J. H. Jaffe (eds.), *Addictive States*, Raven Press, 1992; R. A. Rettig and A. Yarmolinsky (eds.), *Federal Regulation of Methadone Treatment*, National Academy Press, 1994.

Addition

One of the four fundamental operations of arithmetic and algebra. The symbol + of addition is thought to be a ligature for "et," used in a German manuscript of 1456 to denote addition. Its first printed appearance is in Johann Widman's *Behennede und hüpsche Rechnung*, Leipzig, 1489. As a symbol of operation, the plus sign appeared in algebra before arithmetic, and now the term addition, together with its symbol, is applied to many kinds of objects other than numbers. For example, two vectors $\mathbf{x}$, $\mathbf{y}$ are added to produce a third vector $\mathbf{z}$ obtained from them by the "parallelogram" law, and two sets A, B are added to form a third set C consisting of all the elements of A and of B. *See* CALCULUS OF VECTORS.

As an operation on pairs of real or complex numbers, addition is associative, Eq. (1), and commutative, Eq. (2); and multiplication is distributive over addition, Eq. (3). There are important mathematical structures in which an addition is defined that lacks one or more of these properties. Although addition is frequently a primitive concept (defined only by properties assumed for it), it is explicitly defined in Peano's postulates for the natural numbers, in terms of the primitive operation "successor of." When this is denoted by $'$, for any two natural numbers a and b, $a + 1 = a'$ and $a + b' = (a + b)'$. *See* ALGEBRA; DIVISION; MULTIPLICATION; NUMBER THEORY; SUBTRACTION. Leonard M. Blumenthal

$$a + (b + c) = (a + b) + c \tag{1}$$

$$a + b = b + a \tag{2}$$

$$a(b + c) = a \cdot b + a \cdot c \tag{3}$$

Adenohypophysis hormone

Of the endocrine glands, the anterior pituitary, or adenohypophysis, occupies the prime place because, through the secretion of various hormones, it controls the functioning of certain other endocrine glands, namely, the adrenal cortex, the thyroid, and the gonads. In addition, hormones from the anterior pituitary influence the growth and metabolism of the organism through direct action on skeletal, muscular and other tissues. The pituitary maintains control over the various target organs by a type of feedback mechanism which is sensitive to circulating levels of hormones from the target organs. When, for example, the level of thyroxine falls below that required for the maintenance of metabolic homeostasis, the pituitary secretes thyroid-stimulating hormone which induces the thyroid gland to produce additional thyroxine. Pituitary hormones are also released in response to other metabolic conditions which they help to control. One factor influencing the secretion of growth hormone, for example, is the level of blood glucose. *See* HORMONE.

Composition. There are eight anterior pituitary hormones whose existence has been firmly established for some time. They include the two gonadotropic hormones, interstitial-cell stimulating hormone (ICSH, or luteinizing hormone, LH) and follicle-stimulating hormone (FSH); thyrotropic hormone (thryoid-stimulating hormone, TSH); lactogenic hormone (prolactin); growth hormone (GH, or somatotropin, STH); adrenocorticotropic hormone (ACTH, or adrenocorticotropin, or corticotropin); and the two melanocyte-stimulating hormones (α-MSH and β-MSH). In addition, two peptides have been isolated which are structurally related to ACTH and the MSHs. These peptides have been designated β-lipotropic hormone (β-LPH) and γ-lipotropic hormone (γ-LPH).

The hormones of the adenohypophysis are composed of amino acids in peptide linkage and are therefore classed as either polypeptides or proteins, depending on their size. In addition, three of the hormones, TSH, ICSH, and FSH, contain carbohydrate and are therefore categorized as glycoproteins. *See* AMINO ACIDS; PROTEIN.

Gonadotropic hormones. In the female, FSH initiates the development of ovarian follicles. ICSH, acting synergistically with FSH, is necessary for the final stages of follicular maturation and the production of estrogen. ICSH also stimulates the development of the corpus luteum. In the male, FSH stimulates spermatogenesis through its action on the germinal epithelium of the testis, and ICSH primarily activates

the Leydig cells which produce androgen.

ICSH has been isolated from sheep, beef, human, and pork pituitaries. The amino acid composition of ICSH from each of these sources is known, but a complete amino acid sequence has not yet been determined. As is generally true regarding species variation of a given protein hormone, the ICSHs isolated from two closely related species have greater chemical similarities than do ICSHs from distantly related species. Ovine and bovine ICSHs, for example, are more alike in amino acid composition and carbohydrate content than are ovine and human ICSHs. Ovine ICSH consists of 202 amino acids and contains carbohydrate in the form of hexose and hexosamine. The molecule has a weight of approximately 30,000 and under acidic conditions can be dissociated into two subunits having molecular weights of approximately 15,000. The subunits, however, are not identical, having different amino acid and carbohydrate compositions. Each subunit separately has no ICSH activity. Under suitable chemical conditions, however, the two subunits can be recombined to form the active hormone. Thus far ICSH is the only pituitary hormone known to have a subunit structure. *See* CARBOHYDRATE.

FSH has also been isolated from a variety of sources, but highly purified quantities of FSH have proven more difficult to obtain than ICSH, since FSH seems to be present in smaller quantities in the pituitary and is more subject to inactivation during isolation. Ovine FSH has an approximate molecular weight of 25,000.

Metabolic hormones. TSH stimulates the growth of the thyroid gland and the secretion of thyroid hormones. Although TSH has been isolated from beef and human pituitary glands, difficulties in purifying the hormone have hindered studies on its chemical properties. At one time it was believed bovine TSH had a rather low molecular weight of approximately 10,000, but studies now indicate the TSHs from the above-mentioned sources have molecular weights in the range of 25,000. *See* THYROID GLAND.

Lactogenic hormone. Through a process which requires other hormones as well, lactogenic hormone stimulates the mammary gland to secrete milk. There is evidence that in some species of mammals lactogenic hormone also plays a role in maintaining the corpus luteum in the ovary. Lactogenic hormone has been isolated from beef and sheep pituitaries. The chemistry of ovine lactogenic hormone has been substantially investigated, and the entire amino acid sequence of the molecule has been elucidated. The ovine molecule is a chain of 198 amino acids, beginning at the NH_2 terminus with threonine and ending at the COOH terminus with half-cystine. Three disulfide linkages cause the formation of three loops in the chain. In contrast to all nonprimate species investigated, a distinct lactogenic hormone has never been isolated from the pituitaries of the monkey or the human. In these two species a hormone containing lactogenic activity can indeed be isolated from the pituitary, but it also has growth-promoting activity and in major respects corresponds to growth hormones isolated from nonprimate species. *See* LACTATION.

Growth hormones. Thus, in nonprimates, a distinct growth hormone, having no lactogenic activity, can be isolated from the pituitary gland. In the primates, however, growth-promoting and lactogenic activities are present in the same molecule. In spite of its dual activities, the primate hormone is usually referred to simply as growth hormone, which promotes an increase in body size. It stimulates the growth of bones, muscles, and other tissues and enhances the effects of other pituitary hormones on their target organs. Although certain cellular effects of growth hormone are known, such as increasing incorporation of amino acids into muscle protein, the biochemical mechanisms whereby this hormone exerts its effects at the cellular level remain a mystery. The first highly purified growth hormone preparation was obtained from beef pituitaries. Although bovine growth hormone is active in the rat and mouse, it was found to have no effect on human growth when attempts were made to use it for the treatment of human dwarfism. When growth hormones were isolated from human and monkey pituitaries, it was found that only these primate growth hormones were active in humans. Human growth hormone has a molecular weight of 21,500, and the complete sequence of its 191 amino acids has been determined. Also, since the protein is not glycosylated, human growth hormone has been successfully produced via recombinant DNA technology. *See* ANIMAL GROWTH.

Adrenocorticotropin. The hormone ACTH stimulates the growth of the adrenal cortex and the secretion of cortisol and other cortical hormones. ACTHs have been isolated in pure form from sheep, beef, and hog pituitaries, and the amino acid sequences of these molecules have been determined (see **illus.**). Slight variations exist in the structures, but all of the molecules consist of a peptide chain containing 39 amino acids, having an NH_2-terminal serine and COOH-terminal phenylalanine and a molecular weight of about 4500. In 1960–1961 three groups of investigators synthesized peptides having ACTH activity. These peptides represented partial sequences of the ACTH molecule; one, for example, being a nonadecapeptide corresponding to the first 19 amino acids from the NH_2 terminus of ACTH. The entire ACTH molecule was subsequently synthesized, but the synthesis of partial sequences showed that the entire molecule was not necessary for adrenocorticotropic activity. An interesting aspect of the ACTH molecule is that it contains in part a sequence found in the melanocyte-stimulating hormones and lipotropic hormones (see below). In accordance with their chemical similarities, it is not surprising that all of these hormones exhibit both melanocyte-stimulating and lipolytic activities. The complete structural requirements for adrenocorticotropic activity, however, are not present in the MSHs and LPHs, since these hormones do not significantly increase the output of corticoid hormones when assayed in various adrenal-stimulating assays.

Intermedins. The melanocyte-stimulating hormones are also called intermedins, since they can be isolated from the intermediate lobe of the pituitary in those animals which have a distinct intermediate lobe.

Structure of bovine ACTH

```
Ser-Tyr-Ser-Met-Glu-His-Phe-Arg-Try-Gly-Lys-
 1   2   3   4   5   6   7   8   9  10  11
Pro-Val-Gly-Lys-Lys-Arg-Arg-Pro-Val-Lys-Val-
12  13  14  15  16  17  18  19  20  21  22
                                        NH2
                                         |
Tyr-Pro-Asp-Gly-Glu-Ala-Glu-Asp-Ser-Ala-Glu-
23  24  25  26  27  28  29  30  31  32  33
                    Ala-Phe-Pro-Leu-Glu-Phe
                    34  35  36  37  38  39
```

Species	*Amino acid residue in position*
	25 26 27 28 29 30 31 32 33
Pig	Asp-Gly-Ala-Glu-Asp-Glu(NH2)-Leu-Ala-Glu
Sheep	Ala-Gly-Glu-Asp-Asp-Glu-Ala-Ser-Glu(NH2)
Beef	Asp-Gly-Glu-Ala-Glu-Asp-Ser-Ala-Glu(NH2)

Amino acid composition and chemical structure of bovine ACTH compared with ACTH of pig, sheep, and beef. Abbreviations are: Ala, alanine; Arg, arginine; Asp, aspartic acid; Gly, glycine; Glu, glutamic acid; His, histidine; Leu, leucine; Lys, lysine; Met, methionine; Phe, phenylalanine; Pro, proline; Ser, serine; Try, typtophan; Tyr, tyrosine; and Val, valine.

Two types of MSHs have been isolated. α-MSH isolated from pig, beef, and horse consists of 13 amino acids and has an acetylated NH_2-terminal serine and a COOH-terminal valine amide. With the exception of the NH_2-acetyl group, α-MSH is identical to the first 13 residues in ACTH. β-MSH from pig, beef, and horse consists of 18 amino acids, and the sequence from residues 7–13 is identical to the sequence from residues 4–10 in α-MSH and ACTH. Human β-MSH differs from the other known β-MSHs in having 22 amino acids, the 4 additional ones being attached to the NH_2 terminus. Melanocyte-stimulating activity refers to the ability of these hormones to cause dispersion of pigment granules in melanocytes, producing a darkening of the skin, an effect that is usually measured on frog skins. α-MSH is the most potent melanocyte-stimulating hormone as determined by frog skin assay. β-MSH has about 50% of the activity of α-MSH, while ACTH and the LPHs have about 1%.

Lipotropic hormones. In 1965–1966 two new substances, β-lipotropic hormone and γ-lipotropic hormone, were isolated from the pituitaries of sheep and were chemically characterized. The name lipotropic refers to the lipolytic activity of these substances on adipose tissue, which was the initially studied biological action of these hormones. Subsequently, it was learned that they also possessed melanocyte-stimulating activity. Lipolytic activity refers to the ability of certain hormones to stimulate the breakdown of lipid in adipose tissue to free fatty acids and glycerol. As has been mentioned, ACTH and the MSHs are also lipolytic hormones, and are, in fact, of greater potency in this regard than the LPHs. β-LPH has been isolated from beef, pork, and human pituitaries, as well as sheep, and the molecular weight of β-LPH from all of these sources appears to be about 10,000. The amino acid sequence of the ovine β-LPH molecule, which consists of 90 amino acids, has also been determined. In the middle of the chain the entire sequence of β-MSH is present, and therefore it is not surprising that the two molecules exhibit the same activities. γ-LPH, isolated thus far only from sheep pituitaries, consists of the first 58 amino acid residues of β-LPH, ending at the COOH terminus with the complete sequence of β-MSH. The discovery of the LPHs has raised speculation concerning the manner in which the pituitary gland synthesizes a number of substances having common sequences and suggests the possibility that even more peptides related to ACTH and MSH may be present in the pituitary gland. *See* PITUITARY GLAND. Choh Hao Li

Adenosine triphosphate (ATP)

A nucleotide composed of three subunits: ribose (5-carbon sugar), adenine (a nitrogenous base composed of two carbon-nitrogen rings), and a triphosphate group (a chain of three phosphate groups, designated α, β, and γ) [see **illus.**]. ATP is a vital energy-rich chemical found in all living cells. The oxidation of carbohydrates, fats, and proteins provides chemical energy that is used to synthesize ATP, which then serves as the source of the chemical energy used for biosynthesis, ion transport, and muscle contraction. The idea that ATP serves as the common currency of energy exchange in all cells is a cornerstone of biology that was first described by Fritz Lipmann in 1941.

Turnover. ATP can be broken down to adenosine diphosphate (ADP), which contains ribose and adenine but just two phosphate groups, α and β, and inorganic phosphate (P_i) with the liberation of chemical energy. Alternatively, ATP can be broken down to adenosine monophosphate (AMP) and inorganic

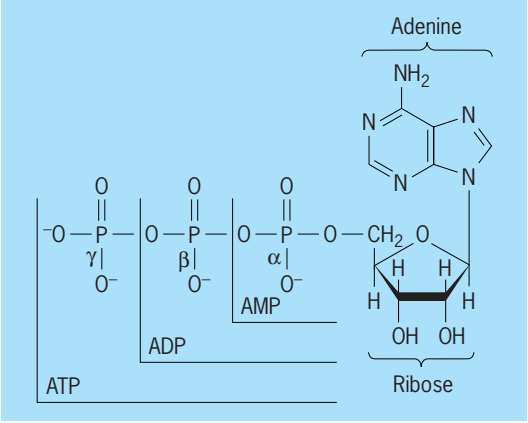

Structure of ATP, ADP, and AMP.

pyrophosphate (PP_i) with the liberation of energy. The magnitude of the synthesis and breakdown of ATP in humans is remarkable. A 110-lb (50-kg) adult female on a 2000-calorie (8360-kilojoule) diet breaks down and resynthesizes about 80 moles of ATP daily. This corresponds to 90 lb (41 kg) of ATP, and each molecule of ATP is broken down and resynthesized about every 1.5 min. *See* BIOLOGICAL OXIDATION; ENERGY METABOLISM; METABOLISM.

Synthesis. ATP is synthesized in cells via three general mechanisms: substrate-level phosphorylation, oxidative phosphorylation, and photophosphorylation.

Substrate-level phoshorylation. In substrate-level phosphorylation, ATP is generated from ADP and P_i as a consequence of energy-yielding oxidation-reduction reactions mediated by nonmembrane-bound enzymes. For example, during the breakdown of glucose, the sequence of reactions involved in the oxidation of glyceraldehyde 3-phosphate to 3-phosphoglyceric acid is coupled to the uptake of P_i and the formation of ATP. The breakdown of fat (or fatty acids) also generates large amounts of ATP. *See* CARBOHYDRATE METABOLISM.

Oxidative phosphorylation. During oxidative phosphorylation, ATP is generated while electrons are transported from organic compounds to oxygen by an electron transport chain. Electron transport produces an increase in the proton concentration outside the inner mitochondrial membrane in animals and plants or outside the cell membrane of oxidative bacteria. Protons, also called hydrogen ions (H^+), bear a positive charge, and the generation of a concentration gradient across a membrane leads to an electrochemical potential difference between the two sides. The protons return to the mitochondria or bacterial cell through a membrane-bound ATP synthase, and this energetically favorable process drives the energetically unfavorable conversion of ADP and phosphate to ATP. *See* ELECTRON-TRANSFER REACTION; MITOCHONDRIA.

Photophosphorylation. In photosynthetic organisms, ATP is generated as a result of photochemical reactions during photophosphorylation. Light energy generates reductants that donate electrons to a photosynthetic electron transport chain. As a result of electron transport, the proton concentration on one side of the inner mitochondrial membrane is increased, and protons pass down their concentration gradient through a membrane-bound ATP synthase resulting in ATP formation. *See* CYTOCHROME; PHOTOSYNTHESIS.

Functions. ATP is a powerful donor of phosphate groups to suitable acceptors because of the energy-rich nature of the bonds between the α- and β-phosphates and between the β- and γ-phosphates (see illus.). For instance, in the phosphorylation of glucose, which is an essential reaction in carbohydrate metabolism, the enzyme hexokinase catalyzes the transfer of the terminal phosphoryl group from ATP to glucose (a carbohydrate) as shown in reaction (1).

$$\text{ATP} + \text{glucose} \rightarrow \text{glucose 6-phosphate} + \text{ADP} \qquad (1)$$

Biosynthesis reactions. ATP is involved directly in many biosynthetic reactions. It is required for reactions that lead to the formation of the bases that occur in deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) and the activation of amino acids, a necessary step in protein synthesis. ATP is also used for the synthesis of amino acids, the building blocks of proteins, and nitrogen fixation in plants. ATP generated during photosynthetic phosphorylation is used to convert carbon dioxide to carbohydrate. *See* NITROGEN FIXATION.

ATP supports many other biosynthetic pathways indirectly. It converts guanosine diphosphate (GDP) to guanosine triphosphate (GTP), which is used in protein synthesis and in many other cellular processes. ATP also converts uridine diphosphate (UDP) to uridine triphosphate (UTP), which is used in carbohydrate biosynthesis, and cytidine diphosphate (CDP) to cytidine triphosphate (CTP), which is used in lipid biosynthesis. The enzyme nucleoside diphosphate catalyzes each of these transphosphorylation reactions, in which a phosphoryl group is transferred from ATP to a nucleoside diphosphate (GDP, UDP, or CDP) to produce a nucleoside triphosphate (GTP, UTP, or CTP) and ADP. These nucleoside triphosphates, along with ATP, are required for RNA synthesis. *See* LIPID METABOLISM; NUCLEIC ACID; PROTEIN; RIBONUCLEIC ACID (RNA).

Enzyme pump. In animals, a considerable amount of ATP is expended to maintain ion gradients using the sodium/potassium ATPase. This enzyme pumps three sodium ions out of cells into the extracellular compartment, and it pumps two potassium ions into cells for each ATP that is hydrolyzed to ADP and P_i.

Muscle metabolism. ATP serves as the source of energy for the mechanical work performed by muscle. Myosin, the chief muscle protein, is the site of energy transduction, by which the chemical energy of ATP is converted into the mechanical energy of muscle contraction. Myosin possesses ATPase activity; it catalyzes the reaction of ATP and water to produce ADP, P_i, and mechanical energy. In addition to oxidative phosphorylation and the breakdown of glucose (substrate-level phosphorylation), ATP synthesis takes place in muscle and the brain during brief periods of energy need through the use of creatine phosphate. Creatine phosphate possesses an energy-rich linkage and is able to transfer the terminal phosphoryl group to ADP to form ATP, which can be used to energize muscle contraction when ATP would otherwise be depleted. Creatine kinase is the enzyme that mediates the reaction (2). When

$$\text{Creatine phosphate} + \text{ADP} \rightleftharpoons \text{ATP} + \text{creatine} \qquad (2)$$

the energy stores of muscle return, ATP mediates the rephosphorylation of creatine to form creatine phosphate.

The adenylate kinase, or myokinase, reaction mediates the transfer of a phosphoryl group from one ADP to another, generating a molecule of ATP that can be used for myosin ATPase and a molecule of

adenylic acid (AMP) as shown in reaction (3). This

$$\text{ADP} + \text{ADP} \rightleftharpoons \text{ATP} + \text{adenylic acid} \qquad (3)$$

reaction releases energy from ADP to produce ATP in times of energy need. When the energy stores of muscle return, the reaction runs in reverse—ATP donates its terminal phosphoryl group to adenylic acid to form two molecules of ADP. *See* MUSCLE; MUSCLE PROTEINS.

Robert Roskoski, Jr.

Bibliography. J. M. Berg, J. L. Tymoczko, and L. Stryer, *Biochemistry*, 5th ed., 2001; C. Mathews, K. Van Holde, and K. Ahern, *Biochemistry*, 3d ed., 2000; D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 3d ed., 2000; D. G. Nicholls and S. J. Ferguson, *Bioenergetics 3*, 2002; R. Roskoski, Jr., *Biochemistry*, 1996.

Adeno-SV40 hybrid virus

A type of defective virus particle in which part of the genetic material of papovavirus SV40 is encased within an adenovirus protein coat (capsid). Human adenoviruses require human cells for their propagation; however, papovavirus SV40 can serve as a "helper," enabling human adenoviruses to replicate in monkey cells.

Characteristics. After numerous continuous passages of adenovirus and SV40 together in monkey cells, adenovirus progeny may be obtained that possesses properties different from those of the original parent adenovirus. It then produces tumors in newborn hamsters; it replicates in monkey cell cultures; and although it does not produce infectious SV40, it causes monkey cells in culture to produce a new cellular antigen known to be specifically induced by SV40—the SV40 tumor, or T, antigen. The new virus stock therefore behaves like a hybrid. It has proved to be a population of two distinct kinds of virus particles. One kind of particle is a true adenovirus. The other particle is the adeno-SV4 hybrid, which has an adenovirus coat, but whose genetic material appears to consist of defective adenovirus type 7 DNA, representing about 85% of the adenovirus genome, covalently linked to a portion (about 50%) of an SV40 genome.

In this virus population the particle carrying the SV40 genetic material has been termed PARA (particle aiding replication of adenovirus). The PARA is considered an unconditionally defective virus, since under no known conditions can it reproduce itself except in a cell coinfected with adenovirus. The adenovirus is considered conditionally defective, since it can reproduce independently in human cells but not in monkey cells (**illus.** *a* and *b*).

When this hybrid virus stock is cultivated serially in human cell cultures, only the true adenovirus particles reproduce. The SV40 genetic determinants are no longer present in the progeny, and the new generations of virus, if inoculated into monkey cells, cannot reproduce themselves. On the other hand, if the hybrid stock population is cultivated in monkey cells, the two kinds of particles complement each other, and new generations contain both kinds of particles. Thus there is a partnership in which the adeno-SV40 hybrid particle supplies genetic information that allows the human adenovirus to reproduce in these foreign cells, and the adenovirus supplies genetic information for synthesis of the protein coat that endows the hybrid particle with the power to attach to cells and to infect them.

Transcapsidation. Type 7 and a number of other adenovirus types can act as helpers for the replication of PARA. In each case the PARA progeny takes on the type-specific protein coat, or capsid, of the helper adenovirus. This change in the capsid of PARA from one type of adenovirus to another is termed transcapsidation (illus. *c*). Some human adenoviruses are able by themselves to induce formation of tumors in hamsters, but other adenoviruses are nononcogenic. After transcapsidation of the PARA particle to some of these previously nononcogenic adenoviruses, they acquire the ability to induce tumors in newborn hamsters. When the tumors are analyzed, both SV40 tumor antigen and adenovirus tumor antigen can be detected in the tumor cells,

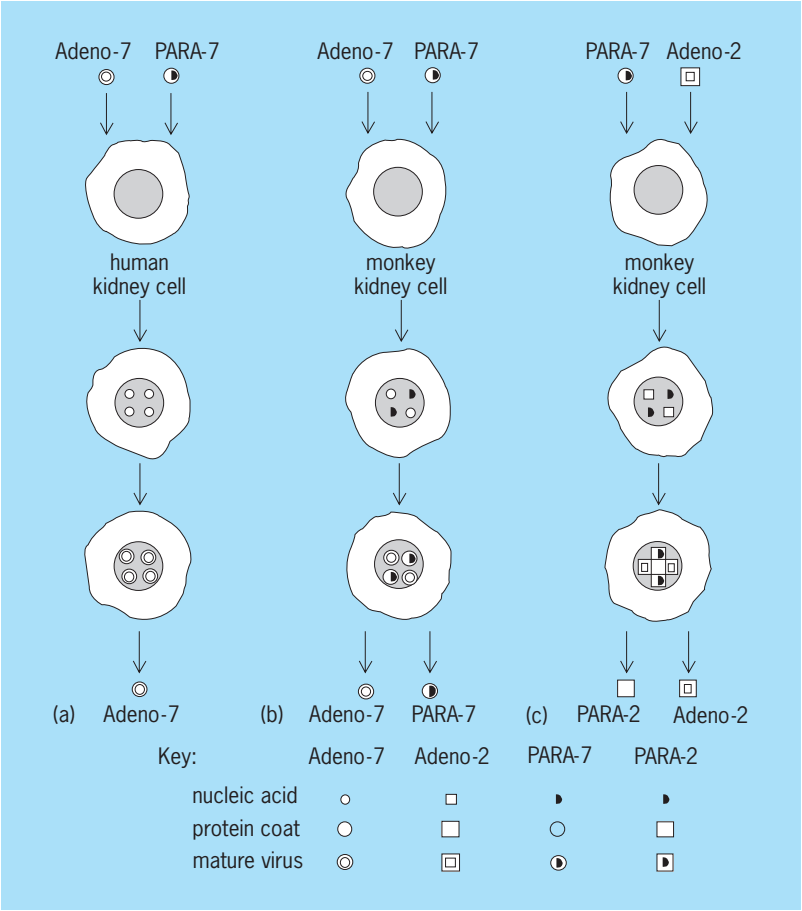

Interactions between PARA-7 particles and adenoviruses. (*a, b*) Mutual dependence between adeno particles and PARA in hybrid populations. The synthesis of SV40 tumor antigen in *b* by a hybrid requires the multiplication of both adeno particles and PARA. (*c*) Transcapsidation. PARA-7 particles (from an SV40-adenovirus type 7 hybrid) grown in the presence of adenovirus type 2 acquire the protein coat of the helper adenovirus. This results in an SV40-adenovirus type 2 hybrid population containing both pure adeno-2 and PARA-2 particles.

and the tumor-bearing animals develop antibodies to both of these antigens. The tumors are of three main types: those with histopathology characteristic for adenovirus tumors, those typical of SV40 tumors, and those with a mixture of both types of pathology.

A second type of hybrid, the $Ad2^{+}ND$ viruses, consists of a series of nondefective adenovirus type 2 isolates carrying different amounts (5–44%) of the SV40 genome. Hybrids of still another type, designated $Ad2^{++}$, contain the entire SV40 genome and are still defective, but they can be propagated in cells coinfected with wild adenovirus type 2.

MAC adenovirus hybrid. Other stocks of human adenovirus have been found that are able to replicate in monkey cells when neither SV40 nor PARA is present. Such stocks contain two types of particles similar to the adenovirus-plus-PARA populations described in the foregoing. The "hybrid" particle containing the nonadenovirus genetic material which makes the adenovirus virulent for monkey cells is known as MAC (monkey adapting component); like PARA, it can be transcapsidated to other types of adenovirus. The source from which this foreign genetic material originally was picked up by the adenovirus is unknown.

Significance. The discovery and exploration of the properties of these hybrids have shown that one virus can usurp the protein coat of another and, while concealed from the usual viral identity tests, can endow a previously mild virus with new powers—which can include the power to produce virulent infection in some cases or cancer in other instances.

These hybrid viruses have been useful in the genetic analysis of SV40. Through the use of physical mapping techniques, the precise amount of SV40 information has been determined for each hybrid. Through comparison of the SV40 antigens (T, U, TSTA) expressed in cells infected by each of the hybrids, a map has been constructed for SV40 genes that code for the different antigens. *See* ADENOVIRIDAE; ANIMAL VIRUS; VIRUS, DEFECTIVE.

Joseph L. Melnick

Bibliography. T. J. Kelly, Jr., et al., *Cold Spring Harbor Symp. Quant. Biol.*, 39:409–417, 1975; F. Rapp and J. L. Melnick, The footprints of tumor viruses, *Sci. Amer.*, 214:34–41, 1966.

Adenoviridae

A family of viral agents. They have been associated with pharyngoconjunctival fever (types 3, 4, and 7, in particular), acute respiratory disease (ARD), epidemic keratoconjunctivitis (type 8), and febrile pharyngitis in children (types 1, 2, and 5). A number of types have been isolated from tonsils and adenoids removed from surgical patients. Although most of the illnesses caused by adenoviruses are respiratory, adenoviruses are frequently excreted in stools, and certain adenoviruses have been isolated from sewage. Distinct serotypes of mammalian (genus *Mastadenovirus*) and avian (genus *Aviadenovirus*) are known. These genera contain 87 and 14 species, respectively. *See* ANIMAL VIRUS.

Infective virus particles, 70 nanometers in diameter, are icosahedrons with shells (capsids) composed of 252 subunits (capsomeres). No outer envelope is known. The genome is double-stranded deoxyribonucleic acid (DNA), with a molecular weight of $20-25 \times 10^6$. Three major soluble antigens are separable from the infectious particle by differential centrifugation. These antigens—a group-specifc antigen common to all adenovirus types, a type-specific antigen unique for each type, and a toxinlike material which also possesses group specificity—represent virus structural protein subunits that are produced in large excess of the amount utilized for synthesis of infectious virus.

The known types of adenoviruses of humans total at least 33, and previously unrecognized types continue to be isolated. The serotypes are antigenically distinct in neutralization tests, but they share a complement-fixing antigen, which is probably a smaller soluble portion of the virus.

Morphologic and antigenic subunits. Adenoviruses contain four complement-fixing antigens—A, B, C, and P—whose characteristics are described in the **table**. The hexons, each surrounded by six neighboring capsomeres, constitute the majority of the capsomeres of the adenocapsid (240); they are tube-like and are 8 nm in diameter. The pentons, found at the 12 vertexes of the capsid, are also 8 nm in diameter. Associated with the penton base antigen is a toxinlike activity which causes tissue culture cells to detach from the surface on which they are growing. Attached to each penton base is a fiber antigen. For adenovirus type 5, the fiber antigen consists of a rod (2×20 nm) attached to the penton, with a 4-nm knob at the end. However, the dimensions of the fiber antigen vary according to the adenovirus type. The P antigen of adenoviruses is an internal antigen which is released upon virus disruption and is very unstable. The hemagglutinating activity of adenovirus is associated with the penton and fiber antigens. Excess pentons produced in cells infected with some types of adenovirus (3, 4, 7, 9, 11, and 15) form dodecahedral aggregates of 12 pentons (dodecons) which have hemagglutinating activity.

Protein components. Polyacrylamide gel electrophoresis of purified adenovirus and adenovirus subunits after disruption by detergent (sodium dodecyl sulfate) and urea has revealed nine virus-specific polypeptides, three of which are associated with the DNA-containing viral core. The hexon capsomere is composed of three to six molecules of a single type (component II) which makes up about 50% of the total virion. Groups of these hexon capsomeres ("groups of nine") are associated with two additional minor polypeptides (components VIII and IX). The penton base is composed of a single type of peptide (component III), as is the fiber antigen (component IV). The internal core released by treatment of the virion with 5 *M* urea contains the viral DNA in association with three arginine-rich peptides (components V, VI, and VII) which constitute about

Table showing comparative data on adenovirus type 2 morphologic and antigenic subunits and protein components

	Morphologic subunits				Antigenic subunits		Protein components
	Appearance	Name	Number per virion	Molecular weight	Antigen	Specificity	
	virion	DNA		23,000,000			
		protein		150,000,000			
		hexon	240	210,000 400,000 320,000 360,000	A	group	II
		hexons	20	3,600,000			II, VIII, IX
		penton	12	280,000 1,100,000			III, IV
		penton base	12	210,000	B	subgroup	III
		fiber	12	70,000	C	type	IV
	core	DNA	1	23,000,000	P		
		protein		29,000,000			V, VI, VII
		protein		13,000			VIII, IX
		protein		7,500			X

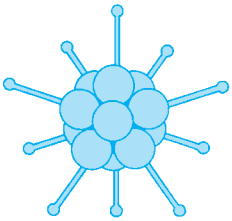

dodecon: hemagglutinin made up of 12 pentons with their fibers

20% of the total virion protein. In addition, there is a very small polypeptide (component X) whose role and location in the virion remain unknown.

Subgroups. A hemagglutinin has been used to separate the human adenoviruses into subgroups. Group A (types 3, 7, 11, 14, 16, 20, 21, 25, and 28) agglutinates rhesus but not rat erythrocytes; group B (8, 9, 10, 13, 15, 17, 19, 22, 23, 24, 26, 27, 29, and 30) agglutinates rat cells but not (or hardly) rhesus cells; group C (1, 2, 4, 5, and 6) fails to agglutinate rhesus cells and only partially agglutinates rat cells. Types 12, 18, and 31 do not usually agglutinate, but some strains partially agglutinate rat cells. The hemagglutination inhibition test is being used for type-specific identification. Some cross-reactions have been noted, however.

Pathogenicity and tumors. The virus does not commonly produce acute disease in laboratory animals but is cytopathogenic, that is, destroys cells, in cultures of human tissue. Certain human adenovirus serotypes produce cancer when injected into newborn hamsters.

In cells derived from other species the human adenoviruses undergo an abortive replicative cycle. Adenovirus tumor antigen, messenger RNA (mRNA), and DNA are all synthesized, but no capsid proteins or infectious progeny are produced. Although complete, infective virus cannot be recovered from adenovirus-induced hamster tumors, a new antigen induced by the virus can be detected by complement fixation or immunofluorescence techniques. This antigen, called tumor or T antigen because of its association with tumor or transformed cells, can also be detected in the cytolytic cycle of the virus. Hamster cells can be transformed in tissue cultures by the oncogenic human adenoviruses. These cells contain the adenovirus tumor antigen, do not contain infectious virus, and produce tumors when inoculated into adult hamsters. Adenovirus mRNA can be detected in the cytoplasm and nucleus of both transformed and tumor cells.

Base ratio determinations have revealed three distinct groups of adenoviruses: those with a low guanine plus cytosine (G + C) content (48–49%); those with an intermediate G + C content (50–53%); and those with a high G + C content (56–60%). The strongly oncogenic adenovirus types 12, 18, and 31 are the only members of the group with low G + C, and certain adenoviruses in the intermediate group (types 3, 7, 14, 16, and 21) are mildly oncogenic. The adenovirus mRNA observed in transformed and tumor cells has a G + C content of 50–52% in the DNA. This suggests that viral DNA regions containing 47–48% G + C are integrated into the tumor cells or that such regions are preferentially transcribed. However, the mRNA from tumor cells induced by one

subgroup such as the highly oncogenic adenoviruses (types 12 and 18) do not hybridize with DNA from the other two subgroups. Apparently, different viral-coded information is involved in carcinogenesis by the three different groups of adenoviruses.

With simian adenovirus 7 (SA7), the intact genome, as well as the heavy and light halves of the viral DNA, is capable of inducing tumors when injected into newborn hamsters. Extensive studies have failed to demonstrate adenovirus DNA or viral-specific mRNA in human tumors.

Adenoviruses, especially types 3, 4, 7, 17, and 21, cause by far the largest number of disabling respiratory diseases among military recruits. In seasoned troops, adenovirus disease is not a serious problem; but the rate in recruits is 33 times higher than in the older group. Presumably, seasoned troops have acquired immunity to the common adenoviruses either before or during their basic training.

Live virus vaccines against type 4 and type 7 have been developed and used extensively in military populations. When both are administered simultaneously, vaccine recipients respond with neutralizing antibodies against both virus types. *See* ADENO-SV40 HYBRID VIRUS; ANTIGEN; COMPLEMENT-FIXATION TEST; NEUTRALIZATION REACTION (IMMUNOLOGY); VIRUS CLASSIFICATION.

Joseph L. Melnick; M. E. Reichmann

Bibliography. E. Everitt et al., Structural proteins of adenoviruses, X: Isolation and topography of low molecular weight antigens from the virion of adenovirus type 2, *Virology,* 52:130–147, 1973; H. Fraenkel-Conrat and R. R. Wagner (eds.), *Comprehensive Virology,* vol. 3: *Reproduction, DNA Animal Viruses,* pp. 143–228, 1974; G. G. Jackson and R. L. Muldoon, Viruses causing common respiratory infections in man, IV: Reoviruses and adenoviruses, *J. Infect. Dis.,* 128:811–866, 1973; B. A. Rubin and H. Tint, The development and use of vaccines based on studies of virus substructures, *Progr. Med. Virol.,* 21:144–157, 1975.

Adhesive

A material capable of fastening together two other materials by means of surface attachment. The terms glue, mucilage, mastic, and cement are synonymous with adhesive. In a generic sense, an adhesive is any material capable of fastening by means of surface attachment, and thus includes inorganic materials such as portland cement and solder. Practically, however, adhesives are a broad set of materials composed of organic, primarily polymeric materials that can be used to fasten two other materials together. The materials being fastened are often called adherends, and the resulting assembly is called an adhesive joint or adhesive bond. *See* ADHESIVE BONDING.

Theories of adhesion. Adhesion is the physical attraction between two surfaces. This attraction is the result of intermolecular interactions that provide the attraction between atoms and molecules, such as electrostatic, van der Waals forces (dipole-dipole interactions, dipole–induced dipole interactions, and dispersion forces), hydrogen bonding, and covalent bonding. With the exception of electrostatic, all of these interactions are strong only at short (nanometer-range) distances. This leads directly to the primary theory of adhesion, the wettability-adsorption theory, in which an adhesive must come into intimate contact with or wet the adherend in order for it to adhere. In this theory, surface energetics and the measurement and interpretation of contact angles play an important role. The diffusion theory of adhesion is used to describe the situation in which an adhesive and adherend are soluble in one another. Here the solubility parameter or the χ-parameter plays an important role. Covalent bonding at interfaces is usually not necessary for a strong adhesive joint. However, it has been found that interfacial covalent bonding is important when the adhesive bond is exposed to adverse environmental conditions such as temperature and humidity. *See* ADSORPTION; CHEMICAL BONDING; INTERFACE OF PHASES; INTERMOLECULAR FORCES; SOLUTION; SURFACE TENSION.

Even though the strength of an adhesive joint depends fundamentally on adhesion, the strength of an adhesive joint is not equal to the force of adhesion. Rather, the strength of an adhesive joint depends in a complex way on many factors, including adhesion, the design of the adhesive joint, and the physical properties of the adhesive and the adherend. *See* JOINT (STRUCTURES).

Types of adhesives. Most polymeric materials have been evaluated as adhesives. Materials that are recognized as adhesives come into intimate contact with the adherends as liquids and then solidify to form the adhesive bond.

Adhesives have been used since ancient times, with the earliest adhesives coming from flora (tree resins and vegetable oils), fauna (casein, blood, and collagen), or a combination of these from prehistoric times, such as bitumen and tar. Generically, these adhesives are called bioadhesives or adhesives of natural origin. These materials are still used. Protein-based adhesives are used to make interior-grade plywood. Some protein-based adhesives are used as tissue sealants during surgery. Cellulosic and starch-based adhesives are used in paper binding. *See* BIOPOLYMER.

The most significant growth in the development and use of adhesives came with the development of synthetic (organic) polymers. The first of these, based on the reaction of phenol and formaldehyde, were phenolic resins. Broadly, organic polymers are classified as thermosets and thermoplastics. Thermoplastics become soft or liquid upon heating and are soluble. Upon heating or curing, thermosets become insoluble and infusible. Both polymer types are used as adhesives. Depending on the ambient temperature, thermoplastics and cured thermosets can be glassy or elastomeric. *See* PHENOLIC RESIN; POLYMER.

Pressure-sensitive adhesives. Pressure-sensitive adhesives are primarily thermoplastic elastomers but may

have some thermoset character. These materials exhibit tack, the property of an adhesive that enables it to form a bond of measurable strength immediately after the adherend and adhesive are brought into contact under low pressure. Thus, when an adhesive is on a backing (as in the case of a pressure-sensitive adhesive tape), just touching the adhesive to a surface is enough to provide measurable adhesive bond strength.

The performance of a pressure-sensitive adhesive results from its viscoelasticity; that is, the adhesive displays both liquidlike and solidlike properties that depend upon temperature and rate. The silicone-based polymer Silly Putty® provides a good example of viscoelasticity. If one places Silly Putty on a surface, after sufficient time (low rate of application of force) it will flow like a viscous liquid. However, if one rolls the Silly Putty into a ball and bounces it against a surface (high rate of application of force), it responds like a solid or elastically.

Pressure-sensitive adhesives are made from a wide range of elastomeric materials, including natural rubber, acrylic resins, poly(vinyl ethers), styrene-butadiene copolymers, and silicone, to name some. Pressure-sensitive adhesives are primarily used for adhesive tapes, which come in various types depending upon the backing and the adhesive. Uses range from paper mending tape, to packaging tape, to removable notes, to adhesive bandages. *See* COPOLYMER; POLYACRYLATE RESIN; POLYVINYL RESINS; RUBBER; SILICONE RESINS.

Semistructural adhesives. Semistructural adhesives are able to support a small load (50–500 lb/in.^{2}; 0.3–3 MPa) for a long time. Semistructural adhesives are either elastomer-based or hot-melt. The elastomer-based adhesives are usually partially thermoset materials formulated with various additives in solvent. These adhesives are used for applying ceramic tile on walls, adhering carpeting to floors, laminating furniture, and many other applications. Elastomers often used in this type of adhesive are chloroprene and nitrile rubber.

Hot-melt adhesives are primarily thermoplastic in character. They are solid materials formulated or synthesized to provide desired physical properties. The melted (liquid) adhesive wets the surface. Upon cooling, the adhesive solidifies and provides adhesive strength. In recent years, the development and use of hot-melt adhesives has become popular because these adhesives do not require the use of solvents for their formulation or application. Hot-melt adhesives, therefore, are environmentally friendly. The materials most often used in hot-melt adhesives are the semicrystalline thermoplastics such as ethylene–vinyl acetate co-polymers and other polyolefins, but polyamides and polyesters are also used. The uses of hot-melt adhesives range from paper and book binding, to laminating, to craft and product assembly. *See* POLYAMIDE RESINS; POLYESTER RESINS; POLYOLEFIN RESINS.

Structural adhesives. Structural adhesives are primarily thermosets and, once cured, have the ability to bond high-strength solids such as wood and metal to create adhesive joints that can bear a significant load (>1000 psi; 7 MPa) for a long time. These adhesives are liquids or become liquids during application. The adhesive solidifies (cures) by means of a chemical reaction that can be initiated by heat or some other means.

Structural adhesives are chemically quite varied, ranging from phenolics and epoxies to polyurethanes and acrylics and encompassing cyanates and polyimides. The chemistry of the adhesive governs the application and use conditions. Epoxy adhesives can be cured at temperatures ranging from room temperature to 170°C (338°F) and can provide strength to temperatures as high as 200°C (392°F). Polyimides cure at high temperatures (in excess of 170°C or 338°F) and provide strength to 350°C (662°F). Acrylic adhesives provide high strength and rapid cure at room temperature. *See* EPOXIDE; HETEROCYCLIC POLYMER; POLYETHER RESINS; POLYURETHANE RESINS.

A well-known structural adhesive is the cyanoacrylate Super-Glue™. Recently, a cyanoacrylate-based adhesive has been approved for medical applications such as wound closure. The primary use of structural adhesives is in load-bearing structures such as aircraft, automobiles, and building materials (such as plywood, pressboard, and strand board). *See* COMPOSITE MATERIAL; WOOD ENGINEERING DESIGN.

Alphonsus V. Pocius

Bibliography. A. J. Kinloch, *Adhesion and Adhesives*, 1987; J. D. Minford, *Treatise on Adhesion and Adhesives*, vol. 7, 1991; R. L. Patrick (ed.), *Treatise on Adhesion and Adhesives*, vols. 1–6, 1967–1990; A. V. Pocius, *Adhesion and Adhesives Technology: An Introduction*, 2d ed., 2002; A. V. Pocius, D. A, Dillard, and M. K. Chaudhury (eds.), *Adhesion Science and Engineering*, vols. 1 and 2, 2002; D. Satas (ed.), *Handbook of Pressure Sensitive Adhesives*, 3d ed., 1999; I. Skeist (ed.), *Handbook of Adhesives*, 3d ed., 1990.

Adhesive bonding

The process of using an adhesive to manufacture an assembly. In this article, the adhesive-bonded assembly will be called an adhesive joint, and the materials to which the adhesive adheres will be called the adherends. The adhesive bonding process consists of adhesive joint design, adherend surface preparation, adhesive application, adhesive joint assembly, solidification of the adhesive (if necessary), adhesive joint evaluation, and use of the assembly. *See* ADHESIVE.

Adhesive joint design. As shown in **Fig. 1**, there are three primary ways in which a load can be applied to an adhesive joint: tension, shear, and cleavage or peel. In general, adhesive joints are designed to perform under a shear load, although adhesive joints can be loaded in all three modes. As with all materials, adhesives are weaker in cleavage than they are in shear or tension. Adhesive joints are designed to minimize the possibility of cleavage loading. *See* SHEAR.

Adhesive joints are designed by first determining the loads that are to be supported. Adherends and adhesives are chosen according to the needs of the application. Properties such as stiffness, toughness (fracture resistance), and elongation are considered. Mechanical engineering principles are applied to ensure that the joint can support the necessary load. A properly designed adhesive joint will provide for adherend failure rather than adhesive failure unless the joint is designed to be reworked or reused. Usually, the design is subjected to a test protocol before production.

Adherend surface preparation. Adhesive joints are made by means of surface attachment; thus the condition of the adherend surface must be taken into account. This is particularly important when the adhesive joint is to be exposed to adverse environmental conditions such as temperature and humidity. In general, surface preparation is done to remove weak boundary layers (such as oils and greases), increase the adherend surface energy, and provide a surface with enough mechanical roughness to "key" the adhesive into the surface of the adherend.

Metal adherend surface preparation methods include abrasion (for example, sanding or grit blasting) and electrochemical methods such as etching for most metals or anodization for metals such as aluminum, titanium, and magnesium. Other inorganic adherends, such as glass and ceramics, can be mechanically abraded or etched. Polymer surface preparation methods include abrasion, chemical etching, corona and flame treatment, and plasma treatment. Bicyclists are familiar with the use of sandpaper or other abrasive to "roughen up" the surface of an inner tube in order to get better adhesion of a patch. In general, a wood surface is not prepared except for sanding the surfaces to be joined in order to ensure proper mating.

In some cases, a primer is applied to the adherend before applying the adhesive. For plastics, a primer can sometimes take the place of a surface preparation. For inorganic materials, a primer can assist formation of covalent chemical bonds between the adherend and the adhesive. Silanes are commonly used to achieve interfacial covalent bonding between an inorganic surface and an organic adhesive. For metals, primers often have other functions such a corrosion protection. *See* CHEMICAL BONDING; ORGANOSILICON COMPOUND; SILICON.

Adhesive application and joint assembly. For a proper adhesive joint, the adhesive must come into intimate contact with or wet the adherend. As a guideline, the adhesive must have a liquid surface energy less than the critical wetting tension of the adherend. If the adherend's surface has been properly prepared or primed, this is usually achievable. Alternatively, the adhesive can be chosen such that intimate contact is achievable. *See* INTERFACE OF PHASES; SURFACE TENSION.

The adhesive can be applied to the adherend by a number of means. The adhesive can be troweled into place as is often done in the application of ceramic tile or floor tile. It can be roller-coated or sprayed as is often done in lamination of surfacing materials onto furniture. The adhesive can be caulked into place. Adhesives in film form are applied as a solid material that is melted and then cooled (solidified). An adhesive can be coated on a backing, such as a pressure-sensitive adhesive tape, and applied by pressure with no further processing steps required.

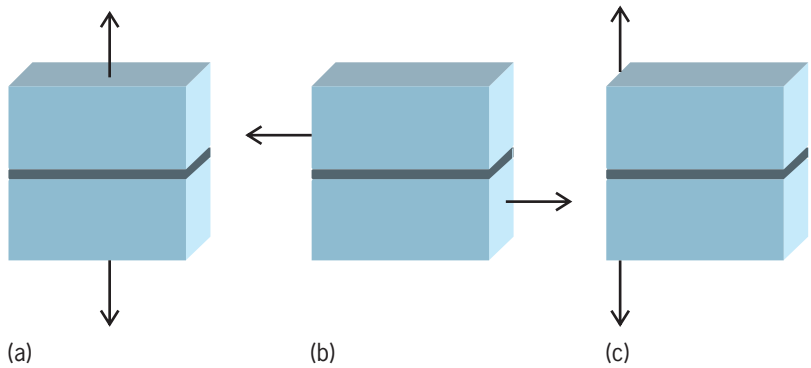

Fig. 1. Modes in which adhesive bonds can be loaded: (*a*) tension, (*b*) shear, and (*c*) cleavage. The layer in the middle of the specimens is the adhesive. In general, adhesive bonds are strongest in shear and weakest in cleavage.

Joint assembly is an important consideration in adhesive bonding. In many cases, the adhesive has a set time, a period in which the adhesive has little or no strength until some solidification takes place. During the solidification process, the adherends must be kept in place by some means. This can be done with materials as crude as simple weights, but often the adherends are kept in place by means of clamps. In some applications, such as aerospace adhesive bonding, the adhesive joint is constructed in a specially made tool that holds the parts in place during solidification. Care must be taken that the force used to hold the adherends in place does not squeeze out significant quantities of the adhesive or the strength of the joint could be adversely affected.

Adhesive solidification. Pressure-sensitive adhesives usually require no processing to solidify, as they are already viscoelastic solids; that is, they have both liquidlike and solidlike character. Adhesives such as rubber-based adhesives and contact bond adhesives require the evaporation of solvent or water to solidify. Thus, some provision must be made to remove the solvent. Adhesives may undergo a chemical reaction to solidify. For example, two-part epoxy adhesives must be properly mixed in order to effect the solidification (cure) of the adhesive. Some adhesives require the application of heat to cure the adhesive, while other adhesives are cured by ultraviolet or visible light. Hot-melt adhesives are applied in the liquid state and solidify upon cooling. *See* POLYETHER RESINS; POLYMER; RUBBER.

Adhesive evaluation. In the majority of cases, adhesive joints cannot be tested for strength without actually breaking the joint. Evaluation of adhesive bonding processes is usually done by coupon tests. That is, model adhesive joints are generated on the same processing line with the actual joints and then tested to failure. If the model joints have strength that is in line with specifications, it can be assumed that the

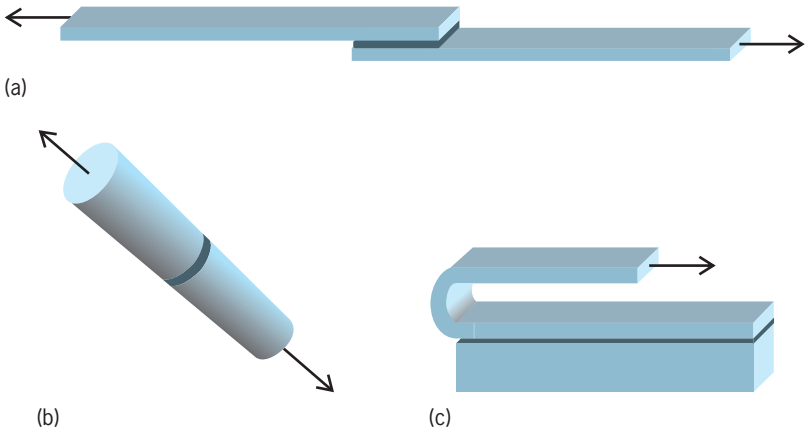

Fig. 2. Examples of adhesive bond test specimens. (*a*) Lap shear test specimen (ASTM D1002). (*b*) Butt tensile specimen (ASTM D2095). (*c*) 180° peel specimen (ASTM D903).

actual joints will perform properly. Model adhesive test joints, as considered in the American Society for Testing and Materials (ASTM), include the lap shear specimen, tensile specimens, and various peel specimens. **Figure 2** shows several of the model adhesive joints that are used to evaluate adhesives and adhesive bonding processes. The lap shear joint is used to test the shear properties of the adhesive, while peel tests are used to evaluate the fracture resistance (toughness) of the adhesive.

Adhesive uses. Adhesives have a wide range of uses. Many types of electronic assemblies use adhesives. For example, silicon chips (dies) are attached to lead frames by means of a die-attach adhesive. This type of adhesive is usually heat-cured, and the cured adhesive is thermally conductive. Adhesives are also used to rigidize parts on a circuit board.

In the automotive industry, adhesives are used to attach decorative features such as body side moldings. They also are used to attach brake linings to brake shoes and friction surfaces in transmissions. Many parts of the automobile are bonded in a hem flange arrangement, in which an inner panel is bonded to an outer skin using an adhesive. The outer skin is then crimped over the inner panel.

Adhesives are used extensively in furniture construction, with rubber-based adhesives used to adhere fabric over foam cushions. Rubber-based adhesives also are used to laminate Formica™ to base wood to produce counter tops. Floor tile and wall tile are attached by means of rubber-based adhesives. Plywood manufacturers are one of the biggest users of adhesive. *See* PLYWOOD.

Aircraft manufacturers extensively use adhesive bonding because of several favorable attributes of adhesives. These include their ability to bond and seal simultaneously, the low specific gravity of polymeric adhesives, and the ability of the adhesive to dampen vibrations so that the adhesive joint is more fatigue-resistant than a welded or mechanically fastened joint.

Adhesives are also used in many noncritical applications such as paper binding, carton sealing (hot-melt adhesives), and envelope sealing. Pages in books have been mended with pressure-sensitive adhesive-backed tapes. Information is transmitted using Post-It™ notes.

Adhesives are being used more often in medicine. In addition to the adhesive bandage that uses a pressure-sensitive adhesive, adhesives are used as tissue sealants during surgeries, and have recently received permission to be used as wound closure materials. Adhesive bonds are also used in transdermal drug delivery systems (such as the nicotine patch). *See* DRUG DELIVERY SYSTEMS. Alphonsus V. Pocius

Bibliography. A. J. Kinloch, *Adhesion and Adhesives*, Kluwer Academic, 1987; J. D. Minford (ed.), *Treatise on Adhesion and Adhesives*, vol. 7, Marcel Dekker, 1991; R. L. Patrick (ed.), *Treatise on Adhesion and Adhesives*, vol. 6, Marcel Dekker, 1990; A. V. Pocius, *Adhesion and Adhesives Technology: An Introduction*, 2d ed., Hanser Gardner, 2002; A. V. Pocius, D. A. Dillard, and M. K. Chaudhury (eds.), *Adhesion Science and Engineering*, vols. 1 and 2, Elsevier Health Sciences, 2002; D. Satas (ed.), *Handbook of Pressure Sensitive Adhesives*, 3d ed., Satas & Associates, 1999; I. Skeist (ed.), *Handbook of Adhesives*, 3d ed., Kluwer Academic, 1990.

Adiabatic demagnetization

The removal or diminution of a magnetic field applied to a magnetic substance when the latter has been thermally isolated from its surroundings. The process concerns paramagnetic substances almost exclusively, in which case a drop in temperature of the working substance is produced (magnetic cooling). *See* PARAMAGNETISM.

Thermodynamic principles. When a specimen is magnetized in a magnetic field **H** and a magnetic moment **M** is produced, there results a contribution to the energy of the specimen of magnitude $-\mathbf{M}\cdot\mathbf{H}$. By restricting the discussion to isotropic media for simplicity, and taking the ambient pressure to be modest and the compressibility of the substance to be very small, the thermodynamics of the situation may be taken over from the conventional results entirely by substituting $-MH$ for $+PV$ (where P is the pressure and V is the volume) everywhere. One particular result is the relation below, where T is the

$$(\partial T/\partial H)_S = (\partial M/\partial S)_H = -(T/C_H)(\partial M/\partial T)_H$$

temperature, S is the entropy, and C_H is the heat capacity at constant magnetic field. As C_H is always positive, the sign of the magnetocaloric effect is determined by (and is opposite to the sign of) the quantity $(\partial M/\partial T)_H$. Where the latter is negative, as is the case for paramagnetics, $(\partial T/\partial H)_S$ is positive and magnetization produces heating, and vice versa. The maximum theoretical effect will not be achieved in practice unless the process is truly isentropic. In addition, it is obviously desirable that nonmagnetic contributions to C_H be kept small. In nearly all cases, this requires restricting the process to a region far

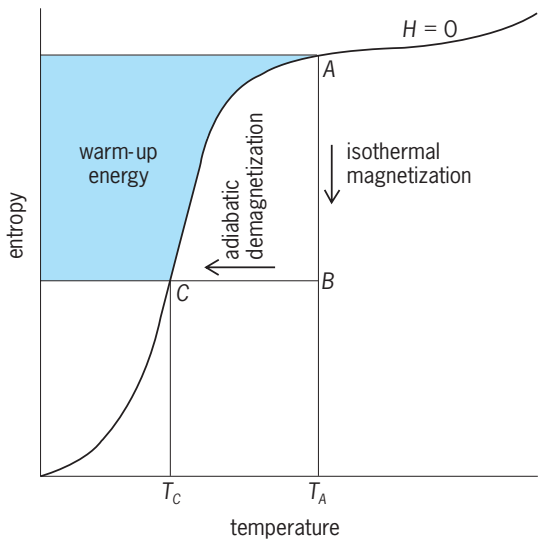

Fig. 1. Entropy–temperature diagram of a paramagnetic substance, showing process of adiabatic demagnetization.

below room temperature, where the lattice specific heat is very small in comparison with that of the assembly of magnetic ions (the spin system). Adiabatic demagnetization was employed from 1933 on to produce temperatures below those readily obtainable by using only liquid helium (that is, below 1 K). *See* LIQUID HELIUM; MAGNETOCALORIC EFFECT.

The process is most perspicuously discussed in terms of entropy rather than energy or heat capacity (**Fig. 1**). The atomic magnets (electronic or nuclear), of moment μ, interact weakly with each other, but at sufficiently low temperatures ($T \sim U/k$, where k is Boltzmann's constant) the interaction energy U will restrict the motions of the atomic magnets and cause a drop in entropy S (Fig. 1). Well above this temperature region, the orientation of μ is practically random over the possible $(2J + 1)$ orientations (where J is the quantum number appropriate to the situation under discussion), and the "infinite temperature" entropy of $R \cdot \ln (2J + 1)$ per mole is approached, where R is the gas constant. At still higher temperatures, the lattice entropy begins to rise to nonnegligible values. Consider a starting condition represented by a state A, the temperature being T_A. Application of a magnetic field isothermally will reduce the entropy to a value corresponding to a state B, say, governed by the thermodynamic relation $(\partial S/\partial H)_T = (\partial M/\partial T)_H$. (The energy which is released by the magnetization process is carried away by exchange gas, helium at low pressure, or a heat switch.) After reaching B, the substance is isolated from its surroundings and H is reduced to zero. Now, no energy exchange with the surroundings can occur, the system returns to the zero-field entropy curve along an isentrope (horizontal line BC), and it arrives at a state C, where the temperature T_C is much lower than temperature T_A. *See* ENTROPY; THERMODYNAMIC PRINCIPLES.

Limiting temperatures. The thermodynamic properties are found to be functions of the quantity $(\mu H/kT)$. The situation may be thought of as a competition between the magnetic (restraining) energy μH and the thermal (disruptive) energy kT. If the interaction U is thought of as arising from an internal magnetic field b, then b at a temperature T_C and H at T_A are equally effective in holding the system at the diminished entropy represented by the isentrope CB. Hence, given that S and, therefore, $(\mu H/kT)$ are constant, the equations below provide a rough approximation. In order to produce a very low temperature, T_C, it is necessary that b and T_A be small and H large. Traditionally, low-temperature physicists have employed electronic paramagnets (paramagnetic salts) to reach limiting temperatures of 2 to 30 millikelvins with $T_A \sim 1$ K and H lying in the range 10–60 kiloersteds or 0.8–4.8 MA/m (magnetic inductions of 10–60 kilogauss or 1–6 teslas). The lower the final temperature T_C, the more difficult it is to maintain the sample in that very low-temperature region, given a fixed background heat influx. The "warm-up energy" (Fig. 1) will be obviously very small if the steep fall in S occurs at a very low temperature. Compromises may be effected by working with different salts or, if feasible, carrying out the adiabatic demagnetization to a nonzero final value of H, which results in a higher final temperature but an enhanced heat capacity.

$$\mu H/T_A = U/T_C = \mu b/T_C$$

$$T_C/T_A = U/\mu H = b/H$$

Nuclear adiabatic demagnetization. Nuclear magnetic moments are one or two thousand times smaller than their ionic (that is, electronic) counterparts, and the characteristic temperature U/k of their mutual interaction lies in the microkelvin rather than millikelvin region. Successful experiments in nuclear adiabatic demagnetization date from the mid-1950s. First, as will be readily deduced from the above discussion, the total thermal energy associated with a loss (or restoration) of a fixed amount of entropy at microkelvin temperatures is very small; hence special precautions for thermal isolation are required. It has, in fact,proved possible to routinely reduce heat leaks from the 0.1 erg s^{-1} (10^{-8} J $\cdot$ s^{-1}) level to 0.1 erg min^{-1} (1.7×10^{-10} J $\cdot$ s^{-1}) in a typical ensemble. Second, as the entropy leverage in the basic process is a function of $\mu H/T_A$, very intense magnetic fields or very low starting temperatures (or a combination of both) must be utilized to compensate for the thousandfold diminution in μ. As there are severe practical limits to substantially increasing H—a maximum induction of 15 teslas, say, using modern superconducting-solenoid technology—salvation is mainly sought in reducing T_A. A factor of 100 may be gained here, that is, from 1 K down to 0.01 K; the latter may be reached by a conventional electronic stage in a two-stage process (**Fig. 2**) or by other modern techniques, such as Pomeranchuk cooling or ^{3}He-^{4}He dilution refrigeration.

In moving from electronic to nuclear magnetic cooling, many difficulties and subtle features of design must be dealt with. Thermal conductivities, heat transfer coefficients, intersystem relaxation times,

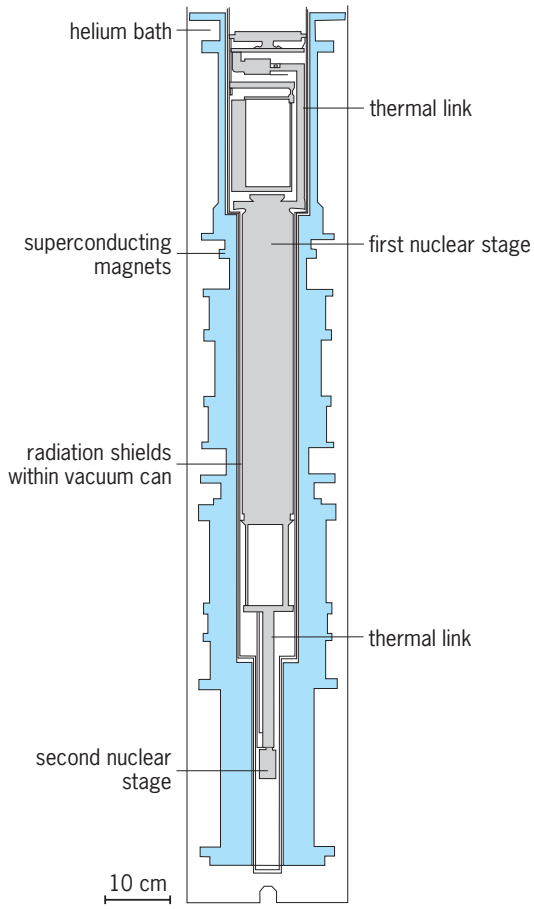

Fig. 2. Low-temperature part of two-stage nuclear refrigerator with superconducting magnets. Demagnetization of the second stage from an initial magnetic field of 9 teslas gave measured electronic temperatures of 10–12 μK. *(After F. Pobell, Matter and Methods at Low Temperatures, Springer, 1992)*

and irreversibilities must be understood and taken into account. (It may, for example, require several hours, rather than minutes, to remove the initial heat of magnetization.) Refinements have also entered through the utilization of the very intense, naturally occurring magnetic fields which make it possible to substitute more subtle and more powerful techniques for the straightforward "brute force" technique discussed so far. Through the nuclear hyperfine interaction, the atomic nuclear moments may be subject to enormous magnetic fields arising from the same (or neighboring) atoms' electronic moments. In general, these hyperfine fields are constant and thus cannot be used for magnetic cooling. In certain systems, however, the electronic magnetic moment is suppressed, or quenched, by crystal electric field effects, and may be partially restored by an external magnetic field, as a result of quantum-mechanical "mixing" of excited states and the ground state of the atom. This induced hyperfine field may correspond to an enhancement of the external magnetic field by a factor of about 10 in praseodymium compounds and about 100 in thulium compounds, for example. The fact that these substances are metals also has practical advantages in energy transmission at very low temperatures. *See* ABSOLUTE ZERO; CRYOGENICS; HYPERFINE STRUCTURE; LOW-TEMPERATURE PHYSICS.

Ralph P. Hudson

Bibliography. R. P. Hudson, *Principles and Application of Magnetic Cooling,* 1972; O. V. Lounasmaa, *Experimental Principles and Methods below 1K,* 1974; F. Pobell, *Matter and Methods at Low Temperatures*, 2d ed., 1996.

Adiabatic process

A thermodynamic process in which the system undergoing the change exchanges no heat with its surroundings. An increase in entropy or degree of disorder occurs during an irreversible adiabatic process. However, reversible adiabatic processes are isentropic; that is, they take place with no change in entropy. In an adiabatic process, compression always results in warming, and expansion always results in cooling. *See* ENTROPY; ISENTROPIC PROCESS.

By the first law of thermodynamics, the change of the system's internal energy U in any process is equal to the sum of the heat Q gained and the work W done *on* the system, as expressed in Eq. (1).

$$\Delta U = Q + W \tag{1}$$

For an adiabatic process, which involves no heat flow to or from the system, the change in internal energy when the system goes from state 1 to state 2 is equal to the external work performed on the system (which brings about the change). In Eq. (2), if U_2 is less than U_1, then W is negative and $-W$ is the work done *by* the system.

$$U_2 - U_1 = W \tag{2}$$

From the ideal gas equation of state, it can be shown that Eq. (3) is followed during an adiabatic

$$P_1 V_1^{\gamma} = P_2 V_2^{\gamma} = \text{constant} \tag{3}$$

isentropic process. Here, γ is defined in terms of the isobaric and isochoric heat capacities, C_P and C_V, respectively, by Eq. (4), where R is the gas constant

$$\gamma = \frac{C_P}{C_V} = \frac{C_V + nR}{C_V} = 1 + \frac{nR}{C_V} \tag{4}$$

and n is the number of moles in the system. *See* GAS; GAS CONSTANT; HEAT CAPACITY; ISOBARIC PROCESS.

During an adiabatic process, temperature changes are due to internal system fluctuations. For example, the events inside an engine cylinder are nearly adiabatic because the wide fluctuations in temperature take place rapidly, compared to the speed with which the cylinder surfaces can conduct heat. Similarly, fluid flow through a nozzle may be so rapid that negligible exchange of heat between fluid and nozzle takes place. The compressions and rarefactions of a sound wave are rapid enough to be considered adiabatic. *See* NOZZLE; SOUND; THERMODYNAMIC PROCESSES.

Philip E. Bloomfield

Bibliography. D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, 5th ed., 1996; J. R. Howell and R. O. Buckius, *Fundamentals of Engineering Thermodynamics*, 2d ed., 1992; G. J. Van Wylen, R. E. Sonntag, and C. Borgnakke, *Fundamentals of Classical Thermodynamics*, 4th ed., 1994; M. W. Zemansky, M. M. Abbott, and H. C. Van Ness, *Basic Engineering Thermodynamics*, 2d ed., 1975.

Adipose tissue

A type of connective tissue that is specialized for the storage of neutral fats (lipids). Adipose cells originate mainly from fibroblasts. In the mammalian embryo, development of adipose cells starts at an early stage. The two types of adipose cells have names reflecting their gross physical appearance: white fat, which can be yellowish if the animal's diet is rich in carotenoids (as found in tomatoes), and brown fat, containing vascularization and respiratory pigments.

White fat. White fat is the more common type of adipose cell. These cells are found in a wide variety of locations in the mammalian body, and their function varies from location to location. For example, they may act to store food reserves and to provide thermal and physical insulation. Locations of stored food reserves include abdominal areas (such as the greater omentum) and around the kidneys and other internal organs. Because lipids are poor conductors of heat, layers of fat beneath the skin serve as heat barriers. This is especially significant in aquatic mammals, such as seals and whales, residing in polar waters. Adipose tissue in the palms of the hands, soles of the feet, and around the eyes acts as a physical cushion.

The number of white adipose cells and the amount of fat in a cell are regulated by various factors, including genetics, hormones, diet, innervation, and physical activity. Studies of identical twins revealed similar fat location but sometimes different amounts of fat tissue. The sympathetic nervous system and norepinephrine are involved in fat mobilization, especially during times of reduced caloric intake or in cold environments. Some hypothalamic neurosecretions play a role in increased body fat. Many animals, especially migratory and hibernating mammals, greatly increase their fat reserves in preparation for travel or for hibernation.

A single white fat cell (an adipocyte; **Fig. 1**) contains a large, single lipid droplet, not bound by membranes, occupying 90% of the cell's volume. The nucleus and other organelles are peripheral at one side of the droplet, with only a thin ring around the remainder of the droplet. While fairly constant in size, the droplet is constantly undergoing a turnover in lipid content. Normal histologic tissue preparation removes the lipid, leaving a clear, open area with the nucleus on one side. Special techniques, such as osmium tetroxide fixation (making fat droplets appear black) or cryostatic sections (frozen sections), are needed to observe lipids intact.

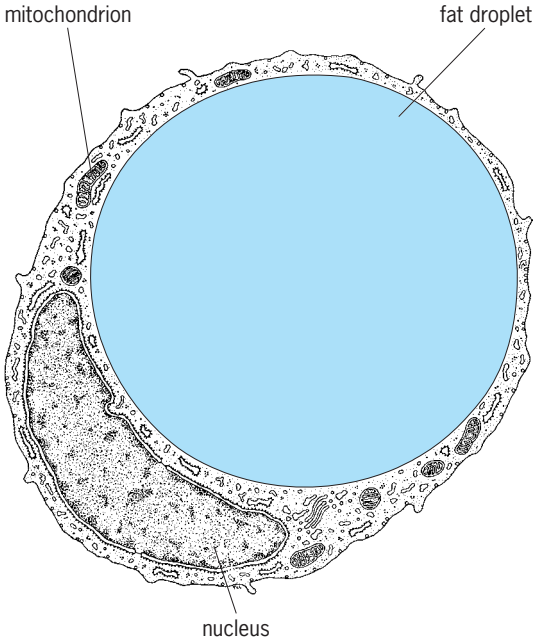

Fig. 1. White adipose cell with a large single fat droplet. (*After T. L. Lentz, Cell Fine Structure: An Atlas of Drawings of Whole-Cell Structure, W. B. Saunders, 1971*)

Brown fat. Brown adipose tissue is more limited in anatomical distribution. It is mainly found in subscapular, interscapular, and mediastinal areas. Brown adipose cells differ from white cells in that they have a centrally located nucleus, many small fat droplets, and abundant mitochondria (**Fig. 2**). These

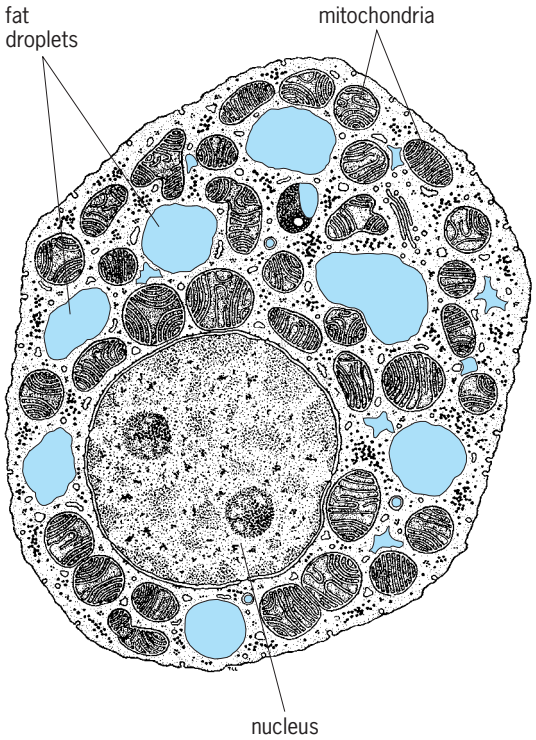

Fig. 2. Brown adipose cell with many small lipid droplets. (*After T. L. Lentz, Cell Fine Structure: An Atlas of Drawings of Whole-Cell Structure, W. B. Saunders, 1971*)

structural differences reflect the thermogenic (heat production) function of these cells, mainly associated with hibernating and newborn mammals. Brown adipose tissue is also highly vascular, aiding in distribution of heat and innervated by the sympathetic nervous system. Brown fat in nonhibernating animals gradually disappears with age and increased size. *See* CONNECTIVE TISSUE; LIPID. Todd Georgi

Bibliography. G. A. Bray, Progress in understanding the genetics of obesity, *J. Nutr.*, 127(5):940S–942S, 1997; R. E. Frisch (ed.), *Adipose Tissue and Reproduction*, 1990; S. Klaus, Functional differentiation of white and brown adipocytes, *Bioessays*, 19(3):215–219, 1997; M. H. Ross, L. J. Romrell, and G. Kaye, *A Text and Atlas*, 1995.

Admittance

The ratio of the current to the voltage in an alternating-current circuit.

In terms of complex current I and voltage V, the admittance of a circuit is given by Eq. (1), and is related to the impedance of the circuit Z by Eq. (2). Y is a complex number given by Eq. (3). G, the real part

$$Y = \frac{I}{V} \tag{1}$$

$$Y = \frac{1}{Z} \tag{2}$$

$$Y = G + jB \tag{3}$$

of the admittance, is the conductance of the circuit, and B, the imaginary part of the admittance, is the susceptance of the circuit. The units of admittance are called siemens or mhos (reciprocal ohms). *See* CONDUCTANCE; SUSCEPTANCE.

The modulus of the admittance is given by Eq. (4),

$$|Y| = \sqrt{G^2 + B^2} \tag{4}$$

which is the ratio of the maximum current to the maximum voltage. The phase angle ψ of the admittance is given by Eq. (5) and is the angle by which

$$\psi = \tan^{-1} \frac{B}{G} \tag{5}$$

the current leads the voltage. The power factor of the circuit is $\cos \psi$, given by Eq. (6).

$$\cos \psi = \frac{G}{\sqrt{G^2 + B^2}} \tag{6}$$

From Eq. (2) it follows that Eqs. (7), (8), and (9) are valid, where R, the resistance, and X, the

$$|Y| = \frac{1}{|Z|} \tag{7}$$

$$G = \frac{R}{R^2 + X^2} \tag{8}$$

$$B = \frac{-X}{R^2 + X^2} \tag{9}$$

reactance, are the real and imaginary parts of the circuit impedance. *See* ALTERNATING-CURRENT CIRCUIT THEORY; ELECTRICAL RESISTANCE; REACTANCE.

J. O. Scanlan

Adrenal gland

A complex endocrine organ in proximity to the kidney. Adrenal gland tissue is present in all vertebrates from cyclostomes to placental mammals. The adrenal consists of two functionally distinct tissues: steroidogenic cells of mesodermal origin and neural crest–derived catecholamine-secreting cells. While "adrenal" refers to the gland's proximity to the kidney, significant variation exists among vertebrates in its anatomic location as well as the relationship of the two endocrine tissues which make up the gland. In mammals, steroidogenic cells are separated into distinct zones that together form a cortex. This cortical tissue surrounds the catecholamine-secreting cells, constituting the medulla. In most other vertebrates, this unique anatomic cortical-medullary relationship is not present. In species of amphibians and fish, adrenal cells are found intermingling with kidney tissue, and the steroidogenic cells are often termed interrenal tissue.

Cortex and medulla. In mammals, the steroid secretory products of the cortex include glucocorticoids, such as cortisol and corticosterone, and mineralocorticoids, primarily aldosterone. Secretion of glucocorticoids is principally regulated by a protein hormone from the pituitary gland, adrenocorticotropin (ACTH). This hormone has been found to regulate steroidogenic secretion in virtually all vertebrates. In contrast, mineralocorticoid secretion is regulated by several factors, including the renin–angiotensin system and the plasma concentration of potassium. Glucocorticoids and mineralocorticoids act on target cells by binding to steroid-specific, intracellular receptors. In some mineralocorticoid-responsive tissues (such as the kidney), target cells contain an enzyme, 11β-hydroxysteroid dehydrogenase, that metabolizes any glucocorticoid entering these cells to inactive by-products. The adrenal cortex is essential for life; glucocorticoids have diverse functions and are especially important in carbohydrate, protein, and water metabolism, while mineralocorticoids are of fundamental importance in controlling sodium balance and extracellular fluid volume. *See* STEROID.

The adrenal medulla contains specialized neurons lacking axons that release catecholamine hormones into circulation. Treatment of medullary cells with potassium dichromate or chromic acid results in a yellow-brown staining pattern termed the chromaffin reaction. Consequently, medullary cells are termed chromaffin cells, although other cells in the body show similar staining. The adrenal medulla is really a ganglion (junction between pre- and postganglionic nerves) in the sympathetic nervous system. The chromaffin cells are unique in that they have lost their axons and instead are specialized for hormone secretion. Two major endocrine cells make up the

medulla, and contain and secrete primarily either epinephrine (adrenaline) or norepinephrine (noradrenaline). Small amounts of dopamine are also secreted. The medulla also includes the nerve terminals of preganglionic neurons that form synapses with the chromaffin cells. Catecholamines are released in response to a variety of stressful events and emergencies, and their hormones help the organism respond to such challenges by mobilizing necessary nutrients, increasing the metabolic rate, increasing alertness, and enhancing activity of the cardiovascular system.

Development. The adrenal gland forms from two primordia: cells of mesodermal origin which give rise to the steroid-secreting cells, and neural cells of ectodermal origin which develop into the catecholamine-secreting tissue. In higher vertebrates, mesenchymal cells originating from the coelomic cavity near the genital ridge proliferate to form a cluster of cells destined to be the adrenal cortex. Another proliferation of cells from the coelomic mesothelium occurs later in fetal life and is added to the original group; this latter group of cells eventually develops into the definitive or adult cortex, while the former proliferation forms the fetal or primary cortex, a structure which disappears in human infants within the first year of life. During the second month of human development, cells of the neural crest migrate to the region of the developing adrenal and begin to proliferate on its surface. The expanding cortical tissue encapsulates the neural cells forming the cortex and medulla. In mammals, three distinct zones form within the cortex: the outermost zona glomerulosa, the middle zona fasiculata, and the inner zona reticularis. The glomerulosa cells contain an enzyme, aldosterone synthase, which converts corticosterone to aldosterone, the principal steroid (mineralocorticoid) secreted from this zone. In addition, new cortical cells originate in the zona glomerulosa, supplying the inner two zones. These inner zones (fasiculata and reticularis) primarily secrete glucocorticoids and large amounts of sex steroid precursors. In many lower vertebrates, the two tissues form from similar primordia but migrate and associate in different ways to the extent that in some cases the two tissues develop in isolation from each other.

Comparative anatomy. While the paired adrenals in mammals have a characteristic cortical-medullary arrangement with distinct zonation present in the cortex, such distinctions are lacking in nonmammalian species. In more primitive fishes, chromaffin cells form in isolation from steroidogenic tissue. A general trend is present, however, throughout vertebrates for a closer association of chromaffin and steroidogenic tissues. Zonation in steroidogenic tissue is largely confined to mammals, although suggestions of separate cell types have been postulated in birds and in some other species.

Cyclostomes (hagfish, lamprey). Cells of adrenal nature that are presumably steroidogenic are found scattered along the walls of the cardinal veins and near the kidneys. Chromaffin cells are similarly located in clumps in these sites.

Elasmobranchs (sharks, rays). Steroidogenic cells are located in a single gland between the kidneys that is termed the interrenal gland. The gland assumes a variety of shapes in different species, ranging from rod-like to horseshoe-shaped. Chromaffin cells are found in paired clusters along the length of the kidney, not associated with steroidogenic cells.

Teleosts (bony fish). Steroidogenic cells are found concentrated in cords or clusters in the region near the posterior cardinal vein and the anterior portion of the kidneys. This structure is termed the head kidney, and additionally contains lymphoid cells. In most teleosts, chromaffin cells are found scattered in the region of the head kidney.

Amphibians. The steroidogenic cells of the adrenal are distributed in a variety of ways in this class, but in many species they are located in islets or nodules near the ventral surface of the kidneys. In some cases, this tissue forms interrenal glands on the ventrum of the kidney. This interrenal tissue is often found intermingling with renal tubules and interstitium. Chromaffin cells are found diffusely intermingling with steroidogenic cells. Another cell, called the Stilling or summer cell, is found in interrenal tissue of certain frogs, and appearance of this cell varies with season, being evident in summer and regressing in winter.

Reptiles. In most reptiles, the adrenals are present as paired, elongated glands anterior to the kidneys (suprarenal). The adrenals, which are covered by a distinct capsule, comprise a mixture of chromaffin and steroidogenic cells. In some reptilian species, chromaffin cells tend to concentrate in masses or clumps, often in the periphery of the gland. In lizards and some snakes, this chromaffin zone almost surrounds the steroidogenic cells, resulting in a gland with an anatomic arrangement opposite to that in mammals.

Birds. In the majority of birds, the adrenals are paired structures near the anterior end of the kidney. The glands consist of a mixture of chromaffin cells and steroidogenic cells, the former present in islets, strands, or larger masses within the gland; the latter as tortuous cords of cells radiating from the center of the gland outward. Evidence for rudimentary zonation of the steroidogenic cells is apparent in some species.

Mammals. The adrenals or suprarenal glands are located as paired structures near the anterior poles of the kidneys. The gland is distinctly separated into two segments, the cortex containing steroidogenic cells concentrically arranged in three zones and the inner medulla containing chromaffin tissue.

Comparative endocrinology. Hormones are secreted from the cells of both the medulla and the cortex.

Chromaffin cells. In all vertebrates, chromaffin cells secrete catecholamines into circulation. These substances are aminogenic derivatives of the amino acid tyrosine. In most species, the major catecholamine secreted is epinephrine, although significant amounts of norepinephrine are released by many animals. Some dopamine is also secreted. No phylogenetic trend is obvious to explain or predict

the ratio of epinephrine to norepinephrine secreted in a given species. Lampreys, for example, release more norepinephrine than epinephrine; the reverse is true in several teleosts and amphibians. Even in mammals the ratios vary: Cats secrete mainly norepinephrine while dogs and humans predominantly secrete epinephrine. A given species may release the two catecholamines in different ratios, depending on the nature of the stimulus. The great majority of the norepinephrine in circulation actually originates from that which is released from non-adrenal sympathetic nerve endings and leaks into the bloodstream. Epinephrine is formed from norepinephrine by the action of the enzyme phenylethanolamine-*N*-methyltransferase. This enzyme is induced in mammalian chromaffin cells by the high levels of glucocorticoids entering the tissue from drainage of blood from the cortex. Animals showing low activity of this enzyme tend to secrete more norepinephrine than epinephrine; it is possible that the physical separation of chromaffin from steroidogenic cells accounts in part for this finding. In addition to catecholamines, chromaffin cells secrete an array of other substances, including proteins such as chromogranin A and opioid peptides. *See* EPINEPHRINE.

Biologic effects of catecholamines are mediated through their binding to two receptor classes, α- and β-adrenergic receptors. Further examination of these receptors has revealed that subclasses of each type exist and likely account for the responses on different target tissues. In general, biologic responses to catecholamines include mobilization of glucose from liver and muscle, increased alertness, increased heart rate, and stimulation of metabolic rate.

Steroid hormones. In broad terms, most steroids secreted by adrenal steroidogenic cells are glucocorticoids, mineralocorticoids, or sex hormone precursors. However, these classes have been established largely on the basis of differential actions in mammals. The principal glucocorticoids are cortisol and corticosterone, while the main mineralocorticoid is aldosterone. This division of action holds for mammalian species and likely for reptiles and birds. In other vertebrates, such as fish and amphibians, steroids from the interrenal tissue do not show such specialized actions; instead, most show activities of both glucocorticoid and mineralocorticoid type. Mammals, birds, reptiles, and amphibians secrete cortisol, corticosterone, and aldosterone. The ratios of the two glucocorticoids vary across species; in general, corticosterone is the more important product in nonmammalian species. Even within mammals, a large variation exists across species, due to the relative ratio of cortisol to corticosterone from the adrenal cortex. Interestingly, the dominant steroid found in teleosts is cortisol. In cyclostomes and elasmobranchs, corticosterone, cortisol, and 11-desoxy-steroids appear to be the major steroid products. In elasmobranchs, a unique steroid, 1-α-hydroxycorticosterone, has been reported. In cyclostomes, elasmobranchs, and teleosts, 18-hydroxylation of corticosterone may not occur (no aldosterone synthase), thus accounting for the lack of aldosterone in these species. Consequently, in these species corticosterone and other glucocorticoids have dual roles in activating both glucocorticoid-type and mineralocorticoid-type responses.

Effects of adrenal-derived steroids in lower vertebrates involve a diverse array of actions, including control of distribution and availability of metabolic fuels such as glucose, and regulation of sodium and extracellular fluid volume. Most studies concerning the role of adrenal steroids in nonmammalian species have centered on their effects on salt and water transport, actions mediated by aldosterone and other mineralocorticoids in mammals. In nonmammalian vertebrates, corticosterone, cortisol, and aldosterone possess mineralocorticoid effects. Some fish can survive in both salt- and fresh-water environments, and other species of vertebrates adapt well to arid climates. Adrenal steroids play important roles in such animals to help regulate salt and water movement inside versus outside the body and between tissue compartments. Targets for these electrolyte- and water-regulating actions of steroids in nonmammalian species include the kidney, gills, muscle, skin, urinary bladder, and alimentary tract. In reptiles and birds, steroids act on the salt (nasal) glands to help regulate excretion of electrolytes. In birds, these glands are located in the orbit above the eyes. Birds that ingest salt water eliminate the excess salt by excretion through these glands; adrenal steroids are necessary for this process. Other hormones, such as prolactin, vasopressin (antidiuretic hormone), and thyroid hormones, act in concert with steroids to accomplish salt and water regulation. Other areas where adrenal steroids likely contribute to biologic processes include control of protein, fat, and carbohydrate balance; reproduction; and growth and development.

Robert J. Kemppainen

Bibliography. I. Chester-Jones, P. M. Ingleton, and J. G. Phillips, *Fundamentals of Comparative Vertebrate Endocrinology*, Plenum Press, New York, 1987; W. Hanke and W. Kloas, Comparative aspects of regulation and function of the adrenal complex in different groups of vertebrates, *Hormone Metab. Res.*, 27:389–397, 1995; D. O. Norris, *Vertebrate Endocrinology*, 2d ed., Lea & Febiger, Philadelphia, 1985; G. P. Vinson, B. Whitehouse, and J. Hinson, *The Adrenal Cortex*, Prentice Hall, Englewood Cliffs, NJ, 1992.

Adrenal gland disorders

Malfunctions of the paired adrenal glands. The adrenal glands can have congenital defects and can be damaged by infections and destructive tumors. Anencephaly, that is, gross underdevelopment of the brain in fetal life, leads to hypoplasia of the adrenal cortex. The cortex is also, although rarely, invaded by acute bacterial diseases, for example, septicemia due to the colon bacillus or meningococcus, which can result in failure of the adrenal cortex followed by shock and death (Waterhouse-Friderichsen

syndrome). Chronic infection of the cortex by the tubercle bacillus or a fungal infection such as histoplasmosis can cause primary adrenal deficiency. Cancer may spread to the adrenal glands, and there are uncommon primary tumors of the adrenal that do not secrete hormones. The most important disorders of the adrenal cortex and medulla are those characterized by abnormal hormone secretion. *See* ADRENAL GLAND.

Despite their anatomic proximity, the two elements of the mammalian adrenal gland are embryologically distinct: the cortex is mesodermal, while the medulla is ectodermal in origin. They also secrete different types of hormones: the cortex secretes steroids and the medulla secretes amines. The cortex is controlled by hormones from the anterior pituitary and the kidney; the medulla is directly controlled by the nervous system. The cortex and medulla are also functionally separate, except that both contribute to the mammalian body's response to outside stimuli, especially stress. *See* PITUITARY GLAND; STEROID.

Adrenal cortex. There are a number of disorders of the adrenal cortex, including primary adrenal cortical insufficiency, secondary cortical insufficiency, adrenal cortical hyperfunction, hyperaldosteronism, secondary hyperaldosteronism, Cushing's syndrome, and virilizing syndromes.

Primary adrenal cortical insufficiency. Primary adrenal cortical insufficiency, or Addison's disease, is rare, but if it is not treated it always results in death. Onset of the disease is subtle and insidious and can mimic malingering or psychoneurosis. The disease is usually caused by autoimmune destruction of the adrenals (cause unknown), by tuberculosis, or less often by chronic fungal diseases. Symptoms include lassitude, muscular weakness, weight loss, prostration during minor illnesses, low blood pressure, and brown pigmentation of the skin. In individuals with this disorder, acute intermittent diseases, surgery, or trauma precipitate Addisonian crisis; shock is reversible only by intravenous administration of water, salt, and large doses of adrenal cortical steroids. The water and electrolyte loss associated with this condition results from deficiency of the adrenal hormone, aldosterone; lack of steroids (for example, cortisol) accounts for the skin color and the extreme susceptibility to trivial stress. Maintenance therapy with small daily doses of hydrocortisone and synthetic, salt-retaining steroids permits a normal life, including the ability to do vigorous physical exercise, to become pregnant, and to have normal longevity. *See* ALDOSTERONE; TUBERCULOSIS.

Secondary adrenal cortical insufficiency. Secondary adrenal cortical insufficiency sometimes accompanies pituitary failure due to tumors, vascular accidents, chronic infectious diseases, or granulomas. The clinical picture differs from Addison's disease: weak, listless, and intolerant of minor illness, individuals also are pale and usually show neurologic signs related to the pituitary lesion. Since the thyroid gland and gonads, like the adrenal cortex, depend on the pituitary gland, hypothyroidism and testicular or ovarian deficiency also are often present. Treatment consists of cortisol, thyroid hormone, and sex steroids. *See* THYROID GLAND.

A more common form of secondary adrenal failure, which involves a feedback mechansim, follows long-term administration of high doses of semisynthetic steroids for treatment of such diseases as rheumatoid arthritis, asthma, and acute lymphoblastic leukemia. Individuals who are treated with these steroids have the general symptoms of spontaneously occurring Cushing's syndrome; however, hormonal tests show that the adrenal gland secretes very little hydrocortisone, not too much as in the natural disease. This deficiency requires the administration of large amounts of cortisol during surgical procedures or severe acute illness.

Adrenal cortical hyperfunction. Manifestations of adrenal cortical hyperfunction, or hyperadrenocorticism, can result from excessive secretion of steroid hormones normally secreted by the adrenal cortex. The four classes of adrenal cortical hormones include mineralocorticoids, such as aldosterone, which regulate salt retention; glucocorticoids, such as cortisol, that affect carbohydrate and protein metabolism, muscle function, and blood pressure; androgens, such as dehydroepiandrosterone; and estrogens, such as estradiol. Estrogens are secreted by very rare adrenocortical tumors (usually malignant) that are most easily recognized in men by signs of feminization (breast enlargement, loss of male body hair, and impotence). *See* ANDROGENS; ESTROGEN.

Hyperaldosteronism. Excess secretion of aldosterone produces hyperaldosteronism. The classic primary form arises from a small unilateral tumor known as an adenoma of the adrenal cortex, from bilateral nodular overgrowth of the cortex, or sometimes from overfunction of the kidney's renin-angiotensin system which stimulates the cortex to secrete too much aldosterone (Bartter's syndrome). The affected individuals have high blood pressure (they make up 1% or less of the hypertensive population), muscle weakness, and decreased levels of potassium in the blood. Diagnosis is facilitated by modern methods of radiologic examination, including computerized tomography and magnetic resonance imaging, together with hormone measurements. The tumors can be removed surgically. In cases where there is no tumor, the individual can be treated with diuretics that act by physiologically opposing the action of aldosterone on the kidney.

Secondary hyperaldosteronism. Secondary hyperaldosteronism occurs in very severe hypertension and in advanced liver disease (cirrhosis), kidney disease (nephrotic syndrome), and heart failure when the body retains excess salt and water with swelling.

Cushing's syndrome. Excessive secretion of cortisol produces hypercortisolism, or Cushing's syndrome. Clinical manifestations include obesity of the trunk with thin arms and legs, round red face, thin skin, brittle bones, high blood pressure, diabetes mellitus, and, in women, virilism (masculinity). The disease must be distinguished from ordinary obesity and other virilizing disorders of the adrenal gland

or ovary. If not treated, Cushing's syndrome is fatal within five years because of complications of hypertension or infections. Hormonal testing and radiographic imaging techniques are needed to determine cause. Small unilateral adenomas of the cortex are easily removed by surgery, and the individual will be completely cured. Carcinomas of the cortex (often virilizing) are malignant; survival after surgery averages 2 years even with chemotherapy. More and more cases are being found where the cause is secretion of adrenocorticotropic hormone by a cancer (such as of the lung, pancreas, and thymus), which stimulates oversecretion of cortisol by the adrenals. In these malignant disorders, treatment of the hormonal disorder is attempted by chemotherapy with drugs that oppose the secretion or action of cortisol. *See* CANCER (MEDICINE).

The most common cause of Cushing's syndrome is overactivity of the pituitary gland, often due to a small tumor that secretes excess adrenocorticotropic hormone. The disease is difficult to treat; methods include surgical removal of the pituitary tumor, x-ray or proton-beam treatment of the pituitary region, chemotherapy, or surgical removal of both adrenals. If the excess cortisol secretion is controlled, the individual can lead a normal life, although sometimes requiring hormone replacement.

Virilizing syndromes. Virilizing syndromes may accompany benign or malignant androgen-secreting tumors of the adrenal cortex. The most common and most important virilizing conditions result from congenital adrenocortical hyperplasia, which is often recognizable at birth or soon after. The cause is deficiency of one of the several adrenocortical enzymes involved in steroid biosynthesis; normal adrenal hormones are not produced in sufficient quantities, and androgenic hormones are synthesized instead. The most common defect, 21-hydroxylase deficiency, is not rare, but it is compatible with a normal life-span and good health; however, the great oversecretion of androgens impairs fertility in both sexes, and females show obvious masculinization. Other forms are accompanied by genital malformation, high blood pressure, or severe salt loss by way of the kidney; the exact nature of the clinical syndrome depends on which adrenal enzyme is lacking. Treatment involves suppressing the pituitary-adrenal axis with cortisol and its derivatives. In special cases, salt-active hormones must also be given.

Adrenal medulla. There is no recognized hormone deficiency of the adrenal medulla in humans. Excess secretion of norepinephrine and epinephrine accompanies pheochromocytoma, a rare adrenal medullary tumor that may be single or multiple, benign or malignant. Sometimes the tumor occurs in multiple endocrine adenomatosis, type II (Sipple's syndrome), characterized clinically by pheochromocytoma, tumors of the parathyroid glands with high blood calcium, and carcinoma of the thyroid. Tumors of the medulla account for fewer than 0.1% of cases of hypertension and are difficult to detect unless the blood pressure rises intermittently, in sudden attacks accompanied by obvious symptoms of excess epinephrine. Diagnosis can be made with imaging methods and by measuring urinary catecholamines, that is, metabolic products of the medullary amines. Treatment is difficult, and consists of careful surgical removal of the tumor. *See* EPINEPHRINE.

Neuroblastoma (in children), ganglioneuroblastoma, and ganglioneuroma are malignant tumors of the adrenal medulla or sympathetic ganglia; they may secrete amines that have only a minor effect on blood pressure. The clinical importance of these tumors arises from their extreme malignancy, but some are known to regress spontaneously. *See* ENDOCRINE MECHANISMS.

Nicholas P. Christy

Bibliography. N. P. Christy (ed.), *The Human Adrenal Cortex*, 1971; H. Cushing, The basophil adenomas of the pituitary body and their clinical manifestations (pituitary basophilism), *Bull. Johns Hopkins Hosp.*, 50:137–195, 1932; L. J. DeGroot (ed.), *Endocrinology*, 3rd ed., 1994; E. M. Gold, The Cushing syndromes: Changing views of diagnosis and treatment, *Ann. Intern. Med.*, 90:829–844, 1979; J. D. Wilson and D. W. Foster (eds.), *Williams Textbook of Endocrinology*, 9th ed., 1998.

Adsorption

A process in which atoms or molecules move from a bulk phase (that is, solid, liquid, or gas) onto a solid or liquid surface. An example is purification by adsorption where impurities are filtered from liquids or gases by their adsorption onto the surface of a high-surface-area solid such as activated charcoal. Other examples include the segregation of surfactant molecules to the surface of a liquid, the bonding of reactant molecules to the solid surface of a heterogeneous catalyst, and the migration of ions to the surface of a charged electrode.

Adsorption is to be distinguished from absorption, a process in which atoms or molecules move into the bulk of a porous material, such as the absorption of water by a sponge. Sorption is a more general term that includes both adsorption and absorption. Desorption refers to the reverse of adsorption, and is a process in which molecules adsorbed on a surface are transferred back into a bulk phase. The term adsorption is most often used in the context of solid surfaces in contact with liquids and gases. Molecules that have been adsorbed onto solid surfaces are referred to generically as adsorbates, and the surface to which they are adsorbed as the substrate or adsorbent. *See* ABSORPTION.

Process. At the molecular level, adsorption is due to attractive interactions between a surface and the species being adsorbed. The magnitude of these interactions covers approximately two orders of magnitude (8–800 kilojoules/mole), similar to the range of interactions found between atoms and molecules in bulk phases. Traditionally, adsorption is classified according to the magnitude of the adsorption forces. Weak interactions (<40 kJ/mol) analogous to those between molecules in liquids give rise to what is called physical adsorption or physisorption.

Strong interactions (>40 kJ/mol) similar to those found between atoms within a molecule (for example, covalent bonds) give rise to chemical adsorption or chemisorption. In physisorption the adsorbed molecule remains intact, but in chemisorption the molecule can be broken into fragments on the surface, in which case the process is called dissociative chemisorption.

The extent of adsorption depends on physical parameters such as temperature, pressure, and concentration in the bulk phase, and the surface area of the adsorbent, as well as on chemical parameters such as the elemental nature of the adsorbate and the adsorbent. Low temperatures, high pressures, high surface areas, and highly reactive adsorbates or adsorbents generally favor adsorption.

A thermodynamic description of adsorption quantifies the driving force for adsorption in terms of these parameters. An important quantity that can be defined in this regard is the Gibbs free energy of the surface, that is, the surface free energy. For a one-component system, surface free energy is also called surface tension, and it corresponds to the so-called reversible work required to create a surface. Surface free energy is the two-dimensional analog of pressure, and whereas pressure is measured as force per unit area, surface free energy is given as force per unit length. Common units for surface free energy are dynes/cm or, equivalently, ergs/cm^2. Easily created surfaces such as those for liquids and layered solids (for example, graphite and mica) have low surface free energies of 0–500 ergs/cm^2. Clean surfaces of tough, refractory materials have high free energies of up to several thousand ergs per square centimeter. At constant temperature and pressure, spontaneous processes lower the free energy of a system, and adsorption generally lowers the free energy of a surface; therefore, adsorption tends to occur most readily on materials with high surface free energies.

A common way to portray the results of adsorption studies on solid surfaces is in the form of an adsorption isotherm. Such a diagram gives the amount of adsorbed material per surface area at a constant temperature as a function of pressure or concentration in the bulk phase. Adsorption often initially increases with pressure and then saturates at a value that corresponds to one layer (a monolayer) of adsorbed molecules. The adsorbed monolayer has a lower free energy than the clean surface, so the driving force for adsorption is decreased upon completion of the monolayer. At still higher pressures, however, the free energy of the molecules in the gas is high enough that there is measurable adsorption of molecules on top of the monolayer and multiple layers or multilayers can be condensed onto the surface.

An adsorption isotherm represents a chemical equilibrium in which molecules are simultaneously adsorbing to and desorbing from the surface. The isotherm shows the extent of the adsorbed layer that is the net result of these two competing processes. For many adsorption systems, particularly those involving physical adsorption, the potential energy decreases monotonically to a minimum as the adsorbate approaches the surface so that there is no energy barrier for adsorption, and equilibrium is readily achieved. Other systems involve activated adsorption, and the adsorbate encounters a potential energy barrier as it approaches the surface. This phenomenon is commonly encountered in dissociative chemisorption; the process can be depicted schematically in a one-dimensional potential diagram (**Fig. 1**). The height of the potential energy barrier in such a system is determined by the extent to which the molecular fragments bond to the surface as the bond within the molecule is being broken. *See* CHEMICAL THERMODYNAMICS.

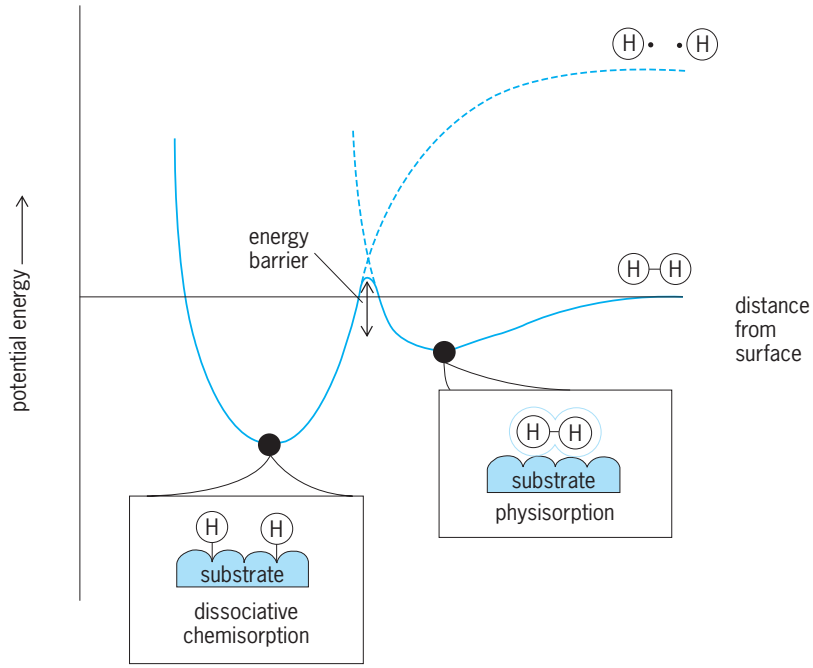

Fig. 1. One-dimensional potential energy curves typical for hydrogen molecules (H_2) and atoms (H) as they approach the surface of a metal substrate. The solid curve, which involves a crossing from the molecular curve to the atomic curve, shows the type of potential profile expected for dissociative chemisorption of H_2 on the surface.

Detection and study. Studying adsorption, particularly in the monolayer regime, is difficult because of the small numbers of molecules involved. For 1 cm^2 of surface area, the number of surface atoms or adsorbed molecules in a monolayer is of the order of 10^{15}, which is only $\sim 10^{-9}$ mole. Often, however, adsorbents have surface areas of hundreds of square centimeters per gram, and adsorption can be quantified simply by measuring the change in pressure or concentration of the bulk phase. Adsorption can also be quantified by measuring the change in mass of the adsorbent and by monitoring changes in the properties of the surface of the adsorbent such as the surface tension and the surface potential. By determining the number of molecules in an adsorbed monolayer, constructing an isotherm, and knowing the size of the adsorbed molecule, the area of the surface can be determined. Such a procedure is the basis for the Brunauer-Emmett-Teller (BET) method for determining surface areas.

Spectroscopies are commonly employed to study

directly the structure and bonding of atoms or molecules in adsorbed monolayers. Spectroscopies that measure the absorption of light have the advantage of high resolution but the disadvantage of a low absorption cross section. High-intensity light sources (lasers and synchrotrons) and high-sensitivity detectors are often required to compensate for low absorption cross sections. If, however, the adsorbent does not interact strongly with the light, optical spectra of adsorbates can be obtained by using high-surface-area adsorbents.

Higher-sensitivity detection of adsorbed monolayers can be achieved by using electrons, ions, and atoms as probes. An advantage of these techniques is that because the probing particles do not significantly penetrate the solid, the information that they provide is surface-specific. Interactions of the probing particles with gas phase and liquid molecules, however, limit these types of studies to vacuum conditions where the particles can be scattered from adsorbed monolayers without interference from the gas phase. *See* MONOMOLECULAR FILM.

The application of spectroscopies to study adsorbed monolayers has formed the basis for what is known as surface science. A wide variety of spectroscopic and diffraction techniques that utilize various combinations of electrons, photons, atoms, and ions as probes have been developed as part of surface science. These include Auger electron spectroscopy (AES), x-ray photoelectron spectroscopy (XPS), electron energy loss spectroscopy (EELS), ion scattering spectroscopy (ISS), and electron or atom diffraction. *See* AUGER EFFECT; ELECTRON DIFFRACTION; SPECTROSCOPY; X-RAY SPECTROMETRY.

Another vacuum-based technique that is frequently applied to probe the strength of surface-adsorbate bonds and to study chemical reactions in adsorbed monolayers is temperature-programmed desorption (TPD). In a temperature-programmed desorption experiment, a substrate covered with an adsorbed monolayer is heated, and the species that desorb from the surface are detected (typically with a mass spectrometer) as a function of surface temperature. The resulting temperature-programmed desorption spectrum consists of peaks that are proportional to the amount of material desorbed, and the peak temperatures are related to the strengths of the adsorbate-surface bonds—higher temperatures corresponding to stronger bonds. If a reaction occurs in the adsorbed monolayer during heating, a temperature-programmed desorption spectrum shows the products that are evolved from the surface and indicates the rate at which they are produced.

Many surface science studies are carried out under ultrahigh-vacuum conditions (pressure = 10^{-8} to 10^{-11} torr, or 10^{-3} to 10^{-6} pascal). Ultrahigh-vacuum conditions make it possible to prepare and maintain atomically clean surfaces of materials that react with air. Often in these studies, samples of single crystals are utilized, and surfaces that are atomically smooth over the range of thousands of atoms can be prepared. These well-defined surfaces can serve as model systems for understanding the molecular details of adsorption.

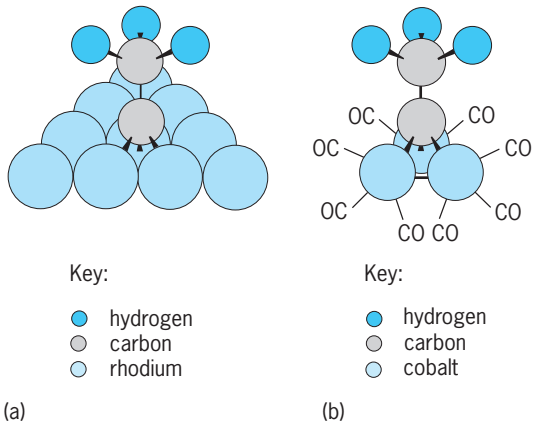

Fig. 2. Analogy between the bonding of (*a*) the ethylidyne ligand (CCH_3) to a rhodium (Rh) single crystal surface and (*b*) to a cobalt carbonyl cluster [$H_3CC(Co)_2(CO)_7$]. The bond lengths and angles, as determined by diffraction techniques, are virtually identical in the two systems.

A general finding from surface science studies of adsorption on single crystal surfaces is that, even for strongly chemisorbed layers, the adsorbate bonding to the surface is directly analogous to ligand bonding in discrete molecular compounds. An example is the ethylidyne ligand (CCH_3) bound to a rhodium surface and to a molecular cobalt complex (**Fig. 2**). *See* INTERMOLECULAR FORCES.

An exciting frontier in surface science studies of adsorption is the application of scanning tunneling microscopy (STM). This technique can provide an atomic resolution image of the electronic structure at the surface (**Fig. 3**). *See* SCANNING TUNNELING MICROSCOPE.

Applications. Direct applications of adsorption in processes such as filtration and detergent action are well known. Adsorption is also the basis for a series of vacuum pumps known as getter pumps. In these pumps, molecules are removed from the gas phase either by physisorption on high-surface-area materials at low temperatures (cryopumps and sorption pumps) or by chemisorption on highly reactive metal surfaces (ion pumps and titanium sublimation pumps).

Adsorption also plays an important role in processes such as heterogeneous catalysis, electrochemistry, adhesion, lubrication, and molecular recognition. In heterogeneous catalysis, gas or solution-phase molecules adsorb onto the catalyst surface, and reactions in the adsorbed monolayer lead to products which are desorbed from the surface. In electrochemistry, molecules adsorbed to the surface of an electrode donate or accept electrons from the electrode as part of oxidation or reduction reactions. In adhesion and lubrication, the chemical and mechanical properties of adsorbed monolayers play a role in determining how solid surfaces behave when in contact with one another. In biological systems, the adsorption of atoms and molecules onto the surface of a cell membrane is the first step in

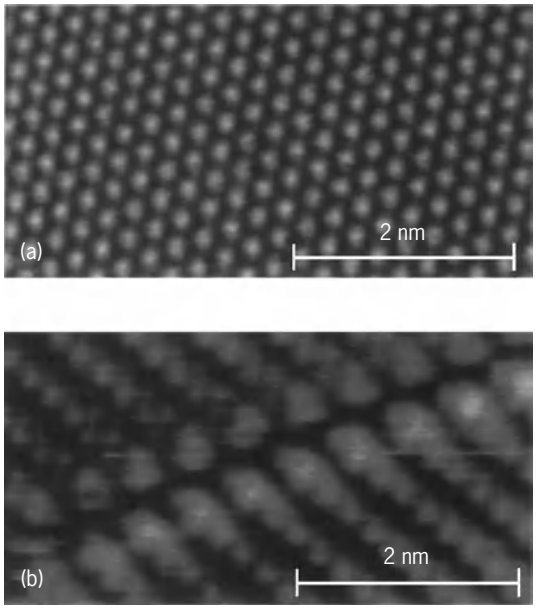

Fig. 3. Scanning tunneling micrographs of 5× 2.25 nanometer areas of (*a*) the basal plane of a graphite surface and (*b*) the same surface after being covered with a liquid drop of 1-triacontanol. The bright spots in *a* correspond to every other carbon atom on the graphite surface. The lumps along the fibrous strands in *b* can be correlated with every other CH_2 group in the 30-carbon chain of 1-triacontanol ($C_{30}H_{61}OH$), which is adsorbed as an ordered monolayer on the graphite surface. (*Courtesy of B. Venkataraman, J. Breen, and G. W. Flynn*)

molecular recognition. *See* ELECTROCHEMISTRY; HETEROGENEOUS CATALYSIS; MOLECULAR RECOGNITION.

Because of the many important technological applications of adsorption, studies of adsorption are important in disciplines ranging from solid-state physics and physical chemistry to materials science and molecular biology. *See* SURFACE AND INTERFACIAL CHEMISTRY; SURFACE PHYSICS. Brian E. Bent

Bibliography. A. W. Adamson, *Physical Chemistry of Surfaces,* 6th ed., 1997; A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications,* 1980; I. Langmuir, The adsorption of gases on plane surfaces of glass, mica, and platinum, *J. Amer. Chem. Soc.,* 40:1361, 1918; J. E. Lennard-Jones, Processes of adsorption and diffusion on solid surfaces, *Trans. Faraday Soc.,* 28:333, 1932; G. A. Somorjai, *Introduction to Surface Chemistry and Catalysis,* 1994.

Adsorption operations

Processes for separation of gases based on the adsorption effect. When a pure gas or a gas mixture is contacted with a solid surface, some of the gas molecules are concentrated at the surface due to gas-solid attractive forces, in a phenomenon known as adsorption. The gas is called the adsorbate and the solid is called the adsorbent.

Adsorption can be either physical or chemical. Physisorption resembles the condensation of gases to liquids, and it may be mono- or multilayered on the surface. Chemisorption is characterized by the formation of a chemical bond between the adsorbate and the adsorbent.

If one component of a gas mixture is strongly adsorbed relative to the others, a surface phase rich in the strongly adsorbed species is created. This effect forms the basis of separation of gas mixtures by gas adsorption operations. Gas adsorption has become a fast-growing unit operation for the chemical and petrochemical industries, and it is being applied to solve many different kinds of gas separation and purification problems of practical importance. *See* ADSORPTION; UNIT OPERATIONS.

Adsorption equilibrium. The extent of adsorption is measured in terms of the surface excess (n^e), which is the actual amount adsorbed (n) minus the amount of gas that would be in the adsorbed phase if the solid exerted no intermolecular forces. The n^e is approximately equal to n when the gas pressure (P) is low. Both n and n^e are expressed in specific terms, that is, in number of moles of gas per unit mass of adsorbent. The relationship between n and P for a pure gas is called the adsorption isotherm. Its shape depends on the nature of the gas and the solid, as well as on the temperature.

Practical considerations require that the adsorbents have a high surface area so that a large amount of gas can be adsorbed in a small volume of the solid. This is achieved by using porous adsorbents which provide a large surface area on the internal pore walls. **Figure 1** shows an example of adsorption isotherms on such a porous solid. These gas loadings are for equilibrium conditions; the actual amounts adsorbed may be less if the contact time is too short. The equilibrium adsorption reaches a maximum with increasing pressure because the total adsorption space is limited by the pore volume of the solid. The type I adsorption isotherms shown on Fig. 1 are the most common, but other shapes are also possible.

The adsorption process is exothermic, and therefore the amount adsorbed decreases with increasing temperature at any given pressure. The heat of adsorption can be defined in various ways, the most practical of which is the isosteric heat of adsorption

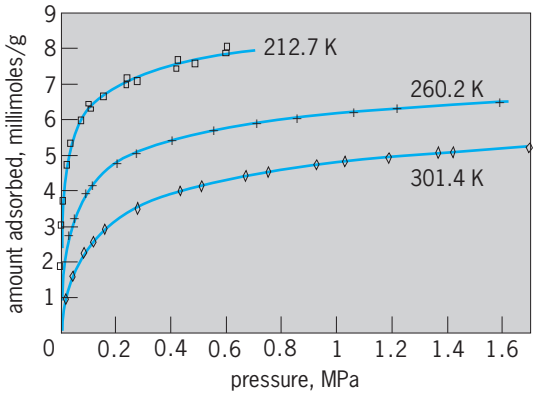

Fig. 1. Adsorption isotherms for ethylene gas on activated carbon. Points are experimental data and solid lines are Eq. (6).

(q) given by Eq. (1). The value for q can be calculated

$$q = RT^2 \left(\frac{\partial \ln P}{\partial T} \right)_n \tag{1}$$

from isotherms measured at different temperatures using Eq. (1).

The amount adsorbed of component i (n_i) from a multicomponent gas mixture of composition y_i depends on P, y_i, and T. A practical means of representing the competitive adsorption of component i with respect to component j of a mixture is in terms of the selectivity (s_{ij}) parameter [Eq. (2)].

$$s_{ij} = \frac{n_i y_j}{n_j y_i} \tag{2}$$

The selectivity, which is a function of P, y_i, and T, forms the basis of separation of a gas mixture by adsorption. Component i is selectively adsorbed over component j if s_{ij} is greater than 1. The higher the value of s_{ij}, the better the separation of the components by adsorption.

Model isotherms. The simplest model for type I isotherms of a pure gas is the Langmuir equation (3), where the value for K is derived from Eq. (4).

$$n = mKP(1 + KP)^{-1} \tag{3}$$

$$K = K_0 \exp \frac{q}{RT} \tag{4}$$

In Eq. (3), the term m is the saturation adsorption capacity and K is a temperature-dependent constant. The two-parameter (m, K) Langmuir isotherm is applicable to homogeneous adsorbents for which the heat of adsorption (q) is independent of coverage (n).

Most commercial adsorbents are, however, energetically heterogeneous, as indicated by the variation of q with n. A common approach to account for heterogeneity is to assume that the adsorbent consists of a distribution of homogeneous patches and that the overall amount adsorbed can be obtained by integration of the contributions of each patch [Eq. (5)], where n_H is the homogeneous isotherm

$$n(P) = \int n_H \lambda(q)\, dq \tag{5}$$

(such as Langmuir's) on a patch defined by the energy q, and $\lambda(q)$ is the probability density function of q on the adsorbent surface.

One such heterogeneous isotherm is that of Toth [Eq. (6)], where the constants m, b, and t are functions of temperature. The Toth equation has been successful in describing type I isotherms of pure gases as shown in Fig. 1. Other three-parameter heterogeneous isotherms have been derived by different choices of n_H and $\lambda(q)$ in Eq. (5).

$$n = mP(b + P^t)^{-1/t} \tag{6}$$

Multicomponent equilibria. Many theories are available for calculating adsorption from a multicomponent gas mixture. The simplest of these for a porous adsorbent is the Langmuir equation (3), written for mixtures [Eq. (7), where K_i is the Langmuir constant for the ith pure gas at the same temperature].

$$n_i = mK_iPy_i(1 + \Sigma K_iPy_i)^{-1} \tag{7}$$

Equation (7) incorrectly predicts that the isothermal selectivity s_{ij} is constant and therefore is not very accurate for heterogeneous adsorbents. Other methods based on adsorption thermodynamics such as ideal adsorbed solution theory are used in practice.

Adsorption kinetics. The actual physisorption process is generally very rapid. However, the measured rate of uptake of a gas by a porous adsorbent particle may be limited by its rate of transfer to the adsorption sites. Slow mass transfer may be caused by the gas-film resistance outside the particle; an anisotropic skin of binding material at the surface of the particle; and internal macro- and micropore diffusional resistances.

The rate of uptake may also be limited by the rate of dissipation of heat liberated during the sorption process. Thus, unlike the equilibrium isotherms, which can be defined without specifying the adsorbent structure or the mechanisms for mass and heat transfer, mathematical models are needed to describe the sorption kinetics. Kinetic models based upon linearized driving forces for heat and mass transfer, or Fickian diffusion of adsorbates into the pores of the solid, are well developed. *See* DIFFUSION.

Direct measurement of the intracrystalline self-diffusion of gases in zeolitic adsorbents can be achieved by the nuclear magnetic resonance pulsed-field gradient technique. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

Column dynamics. Most separations and purifications of gas mixtures are carried out in packed columns (that is, columns filled with solid adsorbent particles). The dynamics of column adsorption can be complex, and it has been the subject of numerous studies.

Adsorption. For adsorption of a single adsorbate of mole fraction y_0 from an inert carrier gas, the process consists of flowing the feed mixture through an isobaric and adiabatic column which has been previously saturated with the pure carrier gas at the temperature (T_0) and pressure (P_0) of the feed gas. Two types of behavior may be observed. (1) In type I behavior, two pairs of heat- and mass-transfer zones are formed in the column as shown in **Fig. 2**. The column ahead (section 1) of the front zones (section 2) remains saturated with carrier gas at the initial conditions. The column (section 3) between the front and rear (section 4) zones is equilibrated with a gas mixture of composition y^* ($y_0 > y^*$) at temperature T^* ($T^* > T_0$). The column (section 5) behind the rear zones is equilibrated with the feed gas at feed conditions. (2) In type II behavior, a pure heat-transfer zone is formed, followed by a pair of mass- and heat-transfer zones as shown in **Fig. 3**. The adsorbate is absent in sections 1 to 3 in this case. The column

behind the rear zones remains equilibrated with the feed gas at feed conditions.

In both cases, the zones propagate through the column after their formation as more feed gas is introduced. The front zones move faster than the rear zones. The shapes and sizes of the zones can be measured by monitoring the $y(t)$ and $T(t)$ profiles in the effluent gas from the column. The effluent profiles are known as breakthrough curves, which are mirror images of the profiles within the column. The zones may expand (proportionate pattern) or retain their shape (constant pattern) as they move through the column.

Constant-pattern front zones for type I systems and constant-pattern rear zones for type II systems are usually formed in a long column when the isotherm of the adsorbate is sufficiently concave toward the pressure axis. Approximate criteria for the formation of type I and type II systems are given in Eqs. (8), where n_0 is the adsorbate capacity at feed

$$\frac{n_0}{y_0} < \frac{c_s}{c_g} \qquad \text{(type I)} \qquad (8a)$$

$$\frac{n_0}{y_0} > \frac{c_s}{c_g} \qquad \text{(type II)} \qquad (8b)$$

conditions and c_s $[J/(g - K)]$ and c_g $[J/(\text{mol} - K)]$ are, respectively, the heat capacities of the adsorbent and the gas mixture. Type I behavior is common for adsorption of bulk adsorbate systems, while type II is favored when the adsorbate is dilute ($y_0 \ll 1$) and strongly adsorbed.

Multiple transfer zones and equilibrium sections are formed in multicomponent adsorption, which is also characterized by roll-over effects, where the more strongly adsorbed species displaces the weaker

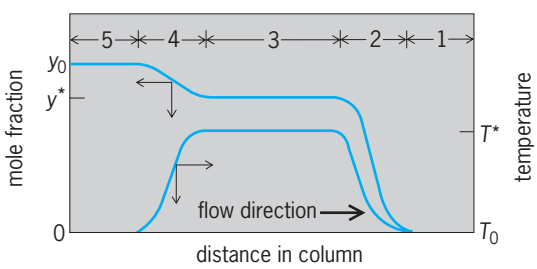

Fig. 2. Composition and temperature profiles during adsorption of type I systems in packed columns. Arrows at the curves indicate the relevant axes.

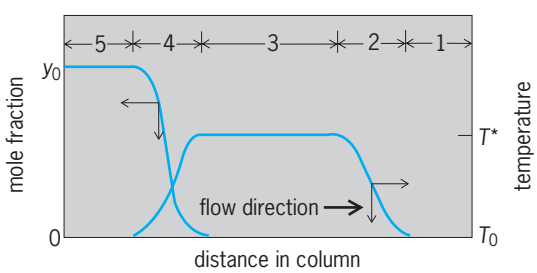

Fig. 3. Composition and temperature profiles during adsorption of type II systems in packed columns. Arrows at the curves indicate the relevant axes.

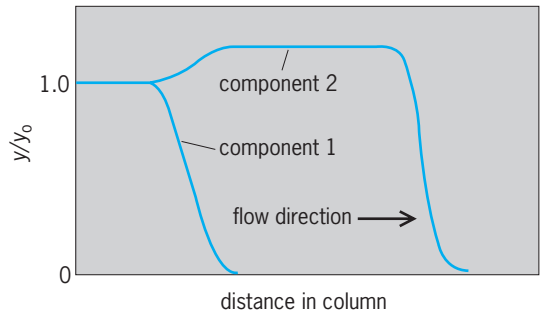

Fig. 4. Roll-over effect showing displacement of more weakly adsorbed component 2 by component 1 during isothermal adsorption from an inert carrier gas.

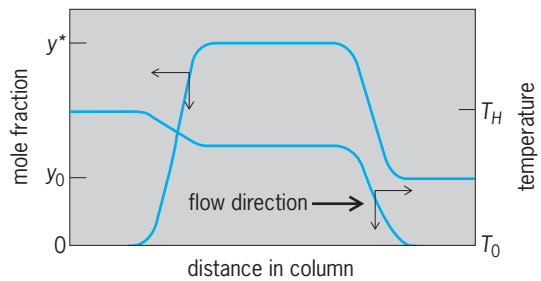

Fig. 5. Composition and temperature profiles during thermal desorption in packed columns. Arrows at the curves indicate the relevant axes.

ones as the zones propagate. **Figure 4** shows, for example, y profiles for an isothermal column adsorbing two substances from an inert gas; component 1 is more strongly adsorbed.

The general behavior of a column can be estimated by the simultaneous solution of the partial differential equations describing the mass and heat balances for each adsorbate species present in the column, in conjunction with their kinetic and equilibrium adsorption properties. Numerical solutions requiring considerable computation time are generally needed. Other factors influencing the dynamics are gas maldistribution, channeling, and nonadiabatic columns.

Desorption. Desorption of adsorbates from a column is usually accomplished by heating the column with a hot, weakly adsorbed gas; lowering the column pressure; purging the column with a weakly adsorbed gas; or combinations of these methods.

Like adsorption, well-defined transfer zones and equilibrium sections can be formed during desorption by the methods of heating and purging as shown in **Figs. 5** and **6**, respectively, for the case of a single adsorbate. Two pairs of mass- and heat-transfer zones are formed in both cases. The sections ahead of the front zones remain equilibrated at the initial column conditions. The middle equilibrium sections, however, attain temperatures (T^*) and compositions (y^*) greater or less than T_0 and y_0 depending on whether the desorption is by the heating or the purging method, respectively. The section behind the rear transfer zones is equilibrated at the conditions of the purge gas.

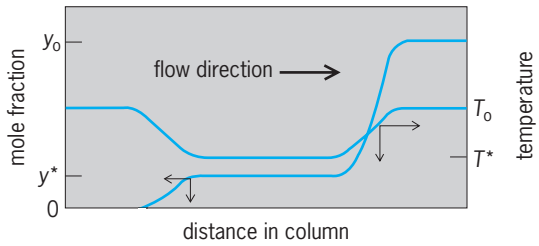

Fig. 6. Composition and temperature profiles during purge desorption in packed columns. Arrows at the curves indicate the relevant axes.

Adsorbents. Microporous adsorbents like zeolites, activated carbons, silica gels, and aluminas are commonly employed in industrial gas separations. These solids exhibit a wide spectrum of pore structures, surface polarity, and chemistry which makes them specifically selective for separation of many different gas mixtures. Separation is normally based on the equilibrium selectivity. However, zeolites and carbon molecular sieves can also separate gases based on molecular shape and size factors which influence the rate of adsorption. *See* ACTIVATED CARBON; MOLECULAR SIEVE; ZEOLITE.

Industrial applications. By far the most frequent industrial applications of gas adsorption have been the drying of gases, solvent vapor recovery, and removal of impurities or pollutants. The adsorbates in these cases are present in dilute quantities. These separations use a thermal-swing adsorption process whereby the adsorption is carried out at a near-ambient temperature followed by thermal regeneration using a portion of the cleaned gas or steam. The adsorbent is then cooled and reused. Many variations of this theme are practiced to reduce the energy consumption and the adsorbent inventory.

Pressure-swing adsorption processes are employed for separation of bulk gas mixtures. The adsorption is carried out at an elevated pressure level to give a product stream enriched in the more weakly adsorbed component. After the column is saturated with the strongly adsorbed component, it is regenerated by depressurization and purging with a portion of the product gas. The cycle is repeated after raising the pressure to the adsorption level. Numerous combinations of these cyclic steps in conjunction with internal recycling of gases have been patented. Key examples of pressure-swing adsorption processes are production of enriched oxygen and nitrogen from air; production of ultrapure hydrogen from various hydrogen-containing streams such as steam-methane reformer off-gas; and separation of normal from branched-chain paraffins. Both thermal-swing and pressure-swing adsorption processes use multiple adsorbent columns to maintain continuity, so that when one column is undergoing adsorption the others are in various stages of regeneration modes. The thermal-swing adsorption processes typically use long cycle times (hours) in contrast to the rapid cycle times (minutes) for pressure-swing adsorption processes. A. L. Myers; S. Sircar

Bibliography. S. J. Gregg and K. S. W. Sing, *Adsorption, Surface Area and Porosity*, 1982; A. L. Myers and G. Belfort (eds.), *Fundamentals of Adsorption*, 1984; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; D. M. Ruthven, *Principles of Adsorption and Adsorption Processes*, 1984; D. M. Young and A. D. Crowell, *Physical Adsorption of Gases*, 1962.

Aerial photography

A photograph of a portion of the Earth's surface taken from an aircraft or from a satellite. Most often, these photographs are taken sequentially and overlap each other, to give complete coverage of the area of interest. Thus they may be viewed stereoscopically to give a three-dimensional view of the Earth's surface (**Fig. 1**). Although the camera in the aircraft may be pointed obliquely to the side of the line of flight, by far the most common type of aerial photograph is taken with the camera pointed vertically downward beneath the plane. *See* STEREOSCOPY.

The entire United States has been photographed, generally more than once, and most of the rest of the world has also been covered. In the United States, the federal government is the major holder and seller of aerial photographs, although private companies also make and sell them. Federal agencies such as the Geological Survey, the Department of Agriculture, NASA, and others routinely use aerial photographs.

Aerial photographs may be made with a variety of films and filters. By far the most common is black-and-white panchromatic film used with a yellow filter. Other films are black-and-white infrared, color, and color infrared. The color and color infrared films are usually flown by companies on contract with the user and are rarely available from the government. The scales of aerial photographs are generally between 1 to 2500 and 1 to 80,000, and the majority have been taken at a scale of about 1 to 20,000. On a 1-to-20,000 photograph, 1 mi on the ground is about 3 in. on the photograph (1 km = 5 cm). As the most frequently available picture size is 9 in. by 9 in. (23 cm by 23 cm), most available federal aerial photographs cover about 9 mi^2 (23 km^2) each.

Photogrammetry. Aerial photographs have two main uses, the making of planimetric and topographic maps, and interpretation or data-gathering in a variety of specialized fields. All modern topographic (contour) maps are made from stereoscopic aerial photographs, usually black-and-white panchromatic vertical photographs. The science of making accurate measurements and maps from aerial photographs is called photogrammetry. Because vertical aerial photographs are central perspective views of the Earth (seen through a single small camera lens), and because the camera may not be precisely vertical when the picture is taken, the individual aerial photograph is not an exact map or geometric representation of the Earth's surface. The geometric discrepancy between the distribution of features seen on the photograph and their actual distribution on the

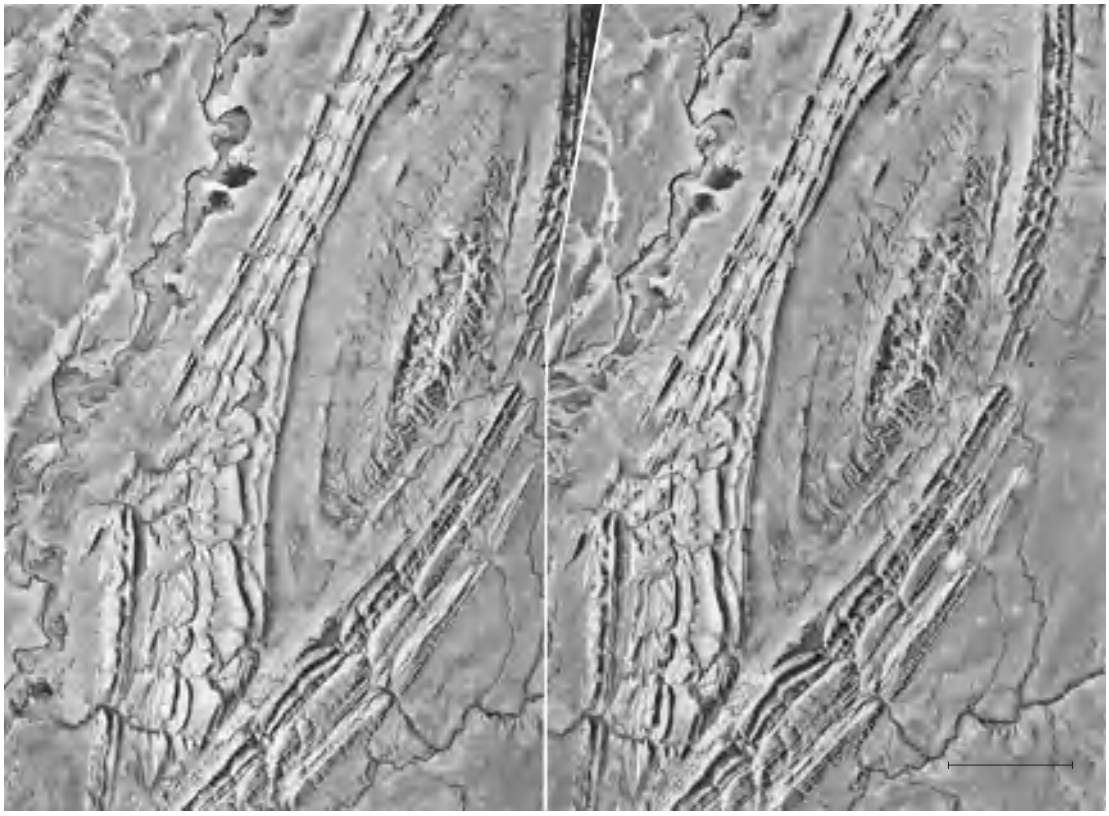

Fig. 1. Stereogram of sedimentary rocks outcropping around eroded structural dome in Wyoming. Such photographs supply much of the information needed in preparing geologic maps. Structures of this type are commonly important in localizing petroleum accumulations. To obtain a three-dimensional image, a standard lens-type stereoscope may be used. *(USGS)*

Earth's surface depends on several variables, the most important of which is the relief of the surface. The error becomes greater as relief increases. Photogrammetry deals with these errors and allows highly accurate and complete map production. Topographic maps are made from one or more stereoscopic pairs of aerial photographs by means of optical-mechanical plotting machines or mathematical methods (analytical photogrammetry). In addition to topography, many other features such as vegetation and structures made by people can be transferred from the photographs to the maps. There is not now available any method of map production which can compete in versatility, speed, and completeness with aerial photographs. *See* CARTOGRAPHY.

Photo interpretation. The other major use of aerial photographs is photo interpretation, utilizing all types of film. Photo interpretation is the attempt to extract information about the Earth's surface (and sometimes, subsurface) from aerial photographs. It is usually carried out by specialists in a subject area. Some users of photo interpretation methods are geologists, foresters, hydrologists, agricultural specialists, soil scientists, land-use planners, environmentalists, military intelligence personnel, engineers, and archeologists. The systematic study of aerial photographs, particularly when they are viewed stereoscopically, gives these specialists the ability to gather rapidly, record, map, and interpret a great deal of information. In addition, an aerial photograph is a record at a moment in time; photographs taken at intervals give an excellent record of the change with time in surface features of the Earth. In some specialties this is of vital importance. The use of different types of film, such as black-and-white infrared or color infrared, allows the interpreter to see the Earth's surface by reflected light energy beyond the range of the human eye. Thus features normally difficult to see may be made apparent.

Applicability. Some examples of the use of aerial photographs may make clear their value. Geologists use the photographs primarily for geologic mapping. It is often possible to see different types of rocks and trace their distribution. This, combined with fieldwork, enormously increases the efficiency and completeness of geologic mapping. But, in addition, the synoptic view of the Earth's surface given by the photograph enables the geologist to map fractures in the Earth's surface, subtle tonal variation in the soil, and other features difficult or impossible to see on the ground. The land-use planner obtains a synoptic view of the surface and of human activities. This, coupled with aerial photographs taken at various times, gives a spatial and temporal database vital to planning. The forester has learned to recognize various tree types and calculate tree sizes, thus enabling rapid and accurate forest inventories to be done. Importantly, various types of film allow rapid recognition of diseased trees and mapping of the

Fig. 2. Aerial view over Sinai Peninsula from *Gemini II*, looking northeast. Eastern desert of Egypt is in foreground, with Israel and Jordan in background. Note detail in which topographic and drainage features are shown. (*NASA*)

extent of the disease. Agricultural specialists can recognize and map various crops, thus making crop inventories in much less time and with much less expense than could be done purely by ground methods. Disease, drought, and other crop stresses can be recognized and mapped. Archeologists can see ancient occupation sites, travel routes, and other features on aerial photographs which are very difficult to see on the ground. Environmental studies are greatly aided by the enlargement of the aerial photograph, and environmental monitoring gains much from the study of changes seen on aerial photographs taken at different times. The use of aerial photographs for military intelligence gathering, bomb damage assessment, and other uses is well known. This short, incomplete list of the wide applicability of photo interpretation to many fields of study indicates the great value of aerial photographs. *See* AGRICULTURE; LAND-USE PLANNING; TOPOGRAPHIC SURVEYING AND MAPPING; VEGETATION AND ECOSYSTEM MAPPING.

Remote sensing. Photo interpretation depends on two things: the appearance and distribution (geometry) of familiar objects, and their spectral reflectance characteristics. The latter category is of increasing importance and is based on the fact that objects on the Earth's surface reflect different amounts of energy (light) in different parts of the electromagnetic spectrum. Thus by careful selection of films and filters with different light sensitivity in different wavelengths of the spectrum, individual "target" objects may be emphasized. This technique is of considerable importance in the interpretation of satellite imagery, such as that transmitted from *Landsat*. In this case the original output is not a photograph but is digital and may be treated mathematically before being made into a photographic image. The field of remote sensing is much concerned with these techniques. Aerial photographs, as well as satellite imagery and airborne radar imagery, are remote-sensed images which, taken as a group, provide a view of the Earth's surface that was never before available (**Fig. 2**). *See* REMOTE SENSING. L. H. Lattman

Bibliography. J. Ciciarelli, *A Practical Guide to Aerial Photography*, 1991; R. Graham, *Digital Aerial Survey: Theory and Practice*, 2002; H. Lloyd, *Aerial Photography: Professional Techniques and Applications*, 1990; E. M. Mikhail et al., *Introduction to Modern Photogrammetry*, 2001; D. P. Paine and J. D. Kiser, *Aerial Photography and Image*, 2003.

Aerodynamic force

The force exerted on a body whenever there is a relative velocity between the body and the air. There are only two basic sources of aerodynamic force: the pressure distribution and the frictional shear stress distribution exerted by the airflow on the body surface. The pressure exerted by the air at a point on the surface acts perpendicular to the surface at that point; and the shear stress, which is due to the frictional action of the air rubbing against the surface, acts tangentially to the surface at that point. The pressure and shear stress act at each point on the body surface, and hence the distribution of pressure and shear stress represent a distributed load over the surface. The net aerodynamic force on the body is due to the net imbalance between these distributed loads as they are summed (integrated) over the entire surface. *See* BOUNDARY LAYER FLOW; FLUID FLOW; WIND STRESS.

Lift and drag definitions. For purposes of discussion, it is convenient to consider the aerodynamic force on an airfoil (**Fig. 1**). The net resultant aerodynamic force R acting through the center of pressure

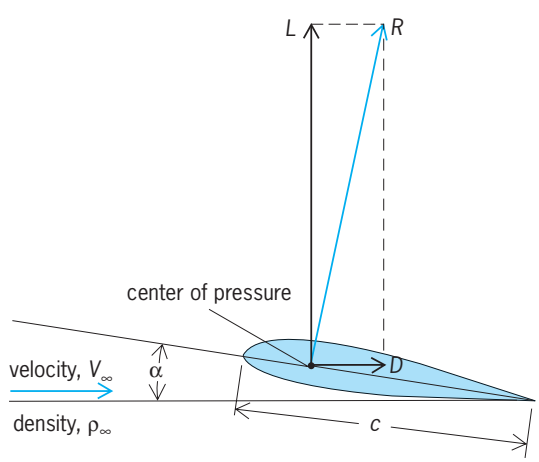

Fig. 1. Resultant aerodynamic force (*R*), and its resolution into lift (*L*) and drag (*D*) components.

on the airfoil represents mechanically the same effect as that due to the actual pressure and shear stress loads distributed over the body surface. The velocity of the airflow V_∞ is called the free-stream velocity or the free-stream relative wind. By definition, the component of R perpendicular to the relative wind is the lift, L, and the component of R parallel to the relative wind is the drag D. *See* AIRFOIL.

Force coefficients. The orientation of the body with respect to the direction of the free stream is given by the angle of attack, α. The magnitude of the aerodynamic force R is governed by the density ρ_∞ and velocity of the free stream, the size of the body, and the angle of attack. An index of the size of the body is given by some characteristic area S of the body. If the body is a winglike shape, a suitable choice for S is the planform area (the area seen when the wing is viewed directly from above). If the body is shaped like a projectile, a suitable choice for S is the maximum cross-sectional area of the body. Since S is used simply as an index of the size of the body, the area chosen is somewhat arbitrary.

The aerodynamic force on a given body shape moving through the air at speeds considerably less than the speed of sound is directly proportional to the density, the area, and the square of the velocity; hence, the same proportionality holds for lift and drag. The proportionality constants are denoted by $C_L/2$ and $C_D/2$, respectively, so that the lift and drag are given by Eqs. (1) and (2). Here, C_L is defined as

$$L = {}^1/_2\, \rho_\infty V_\infty^2 S C_L \tag{1}$$

$$D = {}^1/_2\, \rho_\infty V_\infty^2 S C_D \tag{2}$$

the lift coefficient and C_D as the drag coefficient.

These coefficients vary with angle of attack. At higher speeds, especially near and above the speed of sound, C_L and C_D also become functions of the free-stream Mach number M_∞, defined as $M_\infty = V_\infty/a_\infty$, where a_∞ is the speed of sound in the free stream. Also, in flows where friction is an important aspect of the aerodynamic force, C_L and C_D become functions of the Reynolds number, defined as $\mathrm{Re} = \rho_\infty V_\infty c/\mu_\infty$, where c is some characteristic length of the body (such as the chord length c; Fig. 1), and μ_∞ is the coefficient of viscosity. Therefore, in general, the functional dependence of the lift and drag coefficients is given by Eqs. (3) and (4).

$$C_L = f_1(\alpha, M_\infty, \mathrm{Re}) \tag{3}$$

$$C_D = f_2(\alpha, M_\infty, \mathrm{Re}) \tag{4}$$

See MACH NUMBER; REYNOLDS NUMBER.

Separation and stall. The lift varies linearly with angle of attack, up to a point where the curve that graphs the lift as a function of angle of attack bends over (**Fig. 2**). Beyond this point the lift decreases with increasing α. In this region, the wing is said to be stalled. On the linear portion of the lift curve, the flow over the airfoil is smooth and attached (Fig. 2). In contrast, in the stall region the flow has separated from the top surface of the wing, creating a type of slowly recirculating dead-air region, which decreases the lift and substantially increases the drag. An important measure of aerodynamic efficiency is the ratio of lift to drag, L/D. The higher the value of L/D, the more efficient is the lifting action of the body. The value of L/D reaches a maximum, denoted by $(L/D)_{\max}$ (see **table**), at a relatively low angle of attack (Fig. 2). *See* FLIGHT CHARACTERISTICS.

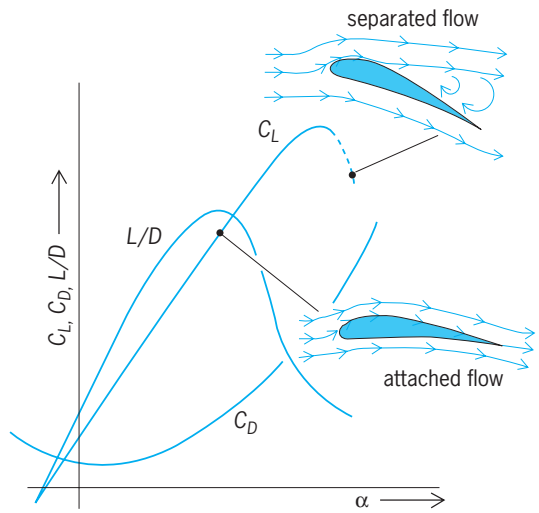

Fig. 2. Typical variation of lift coefficient (C_L), drag coefficient (C_D), and lift-to-drag ratio (L/D) for an airfoil as functions of angle of attack (α). Insets show physical nature of the flow field over the wing at the indicated points on the lift curve.

Maximum lift-to-drag ratios of representative aircraft

Aircraft	$(L/D)_{\max}$
Wright flyer	5.7
Sopwith Camel	7.7
Ryan NYP (*Spirit of St. Louis*)	10.1
Douglas DC-3	14.7
Boeing B-52	21.5
F-4 fighter	8.6

Drag components. The total aerodynamic drag on a body is often dissected into contributing parts.

Skin friction drag. This is due to the net integrated effect of the surface shear stress exerted over the body.

Pressure drag due to flow separation. When the flow separates from the body, the change in the surface pressure distribution in the dead-air separated region causes an increase in drag. Although the aerodynamic mechanism that causes the flow to separate from the surface involves the effect of friction in the flow, this drag is purely a pressure effect. This drag is sometimes called form drag.

Parasite and profile drag. The sum of skin-friction drag and pressure drag due to flow separation is denoted for a wing as profile drag, and for a complete airplane as parasite drag.

Induced drag. This is associated with the vortices that are produced at the tips of wings and trail downstream from the wings (**Fig. 3**). A wing produces lift because the pressure on its bottom surface is

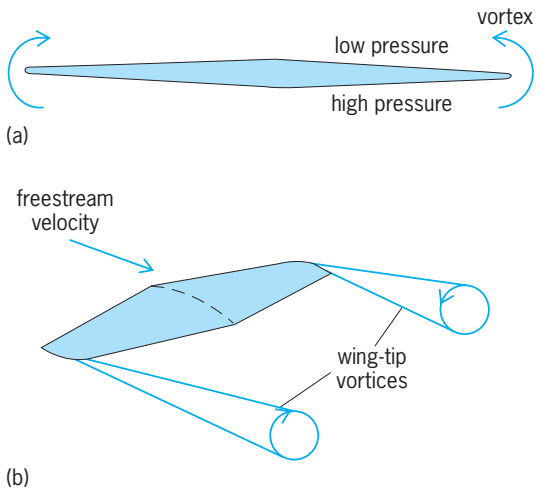

Fig. 3. Origin of wing-tip vortices on a finite wing as shown in (*a*) front view and (*b*) oblique view of wing.

higher than that on the top. At a wing tip, the higher-pressure air curls around the tip from bottom to top, setting up a circulatory motion which is superimposed on the main flow velocity and creates vortices which trail downstream from each tip. In turn, the presence of these vortices changes the pressure distribution over the wing surface in comparison to that which would exist if the wing had no tips (that is, if the wing were infinite in span). This change in the pressure distribution is in such a direction that it results in an increase in drag. This increment in drag is called induced drag; it is purely a pressure effect. Since the wing-tip vortices, which cause induced drag, in turn result from the pressure differential on the wing surfaces, and this pressure differential is also responsible for the creation of lift, induced drag is frequently called drag due to lift. *See* VORTEX.

Induced drag is the penalty paid for the production of lift. For an aircraft in steady, level flight, the lift produced must equal the weight of the aircraft. The induced drag caused by this lift must be overcome by increased thrust from the engines.

Wave drag. This is associated with the presence of shock waves on the aerodynamic vehicle. It is a pressure drag that is caused by severe increases in pressure that occur across shock waves. *See* AERODYNAMIC WAVE DRAG.

Drag coefficient expressions. In essence, the total drag on a body can be visualized as the sum of parasite drag, induced drag, and wave drag. For a wing at subsonic speeds, the total drag coefficient C_D is expressed as the sum of the profile drag coefficient $c_{d,p}$ and induced drag coefficient C_{Di}, as in Eq. (5), where the induced drag coefficient is given by Eq. (6). Here,

$$C_D = c_{d,p} + C_{Di} \tag{5}$$

$$C_{Di} = \frac{C_L^2}{\pi e \mathrm{AR}} \tag{6}$$

AR is the wing aspect ratio, defined as b^2/S, where b is the span of the wing (distance from one wing tip to the other), and e is a span effectiveness factor having to do with the shape of the planform of the wing; values of e typically range from 0.85 to 1.0. As seen in Eq. (6), induced drag increases as the square of the lift. *See* ASPECT RATIO.

For a complete airplane at subsonic speed, the total drag coefficient is the sum of the parasite drag coefficient C_{DP} and induced drag coefficient, as in Eq. (7). Since the pressure drag on an airplane depends on angle of attack, the total drag for an airplane is frequently expressed as a sum of the zero-lift drag and the drag due to lift. In coefficient form, the total airplane drag coefficient is given by Eq. (8). Here,

$$C_D = C_{DP} + \frac{C_L^2}{\pi e \mathrm{AR}} \tag{7}$$

$$C_D = C_{D,0} + \frac{C_L^2}{\pi e_1 \mathrm{AR}} \tag{8}$$

$C_{D,0}$ is the parasite drag coefficient at zero lift, and $C_L^2/\pi e_1 \mathrm{AR}$ is the drag coefficient due to lift, which is physically the sum of the induced drag coefficient plus the extra increment in parasite drag coefficient when the airplane is pitched to an angle of attack to produce lift. This extra parasite drag is reflected in the factor e_1, called Oswald's efficiency factor, whose value is typically of the order of 0.55–0.7. Equation (8) is called the drag polar for the airplane, and contains all the practical aerodynamic information needed to analyze the airplane's aerodynamic performance.

John D. Anderson, Jr.

Bibliography. J. D. Anderson, Jr., *Fundamentals of Aerodynamics*, 4th ed., 2005; J. D. Anderson, Jr., *Introduction to Flight*, 5th ed., 2005; J. J. Bertin, *Aerodynamics for Engineers*, 4th ed., 2002; S. F. Hoerner, *Fluid-Dynamic Drag*, 1965; S. F. Hoerner and H. V. Bortst, *Fluid-Dynamic Lift*, 2d ed., 1985.

Aerodynamic sound

Sound that is generated by the unsteady motion of a gas and its interaction with surrounding surfaces. Aerodynamic sound or noise may be pleasant, such as the sound generated by a flute, or unpleasant, such as the noise of an aircraft on landing or takeoff, or the impulsive noise of a helicopter in descending or forward flight. The typical human ear can detect a great range of sounds. A logarithmic scale, called the decibel (dB) scale, is used to describe the sound pressure level (SPL). The sound pressure level is given by the equation below, where p_{rms} = root-mean-square

$$\mathrm{SPL} = 20 \log_{10} \left(\frac{p_{\mathrm{rms}}}{p_{\mathrm{ref}}} \right) \mathrm{dB}$$

pressure level of the sound, and reference pressure $p_{\mathrm{ref}} = 2 \times 10^{-5}\ \mathrm{N/m^2}$. Zero decibels corresponds to the quietest whisper, whereas the threshold of pain occurs at 130–140 dB. *See* LOUDNESS; SOUND PRESSURE.

Basic principles. Sources of aerodynamic sound may be classified according to their multipole order. Sources associated with unsteady mass addition to

the gas are called monopoles. These could be caused by the unsteady mass flux in a jet exhaust or the pulsation of a body. The sound power radiated by monopoles scales with the fourth power of a characteristic source velocity. Sources related to unsteady forces acting on the gas are called dipoles. The singing in telephone wires is related to the nearly periodic lift variations caused by vortex shedding from the wires. Such sources, called dipoles, generate a sound power that scales with the sixth power of the characteristic source velocity. *See* KÁRMÁN VORTEX STREET.

A turbulent fluid undergoes local compression and extension as well as shearing. These events are nearly random on the smallest scales of turbulent motion but may have more organization at the largest scales. The earliest theories of aerodynamic noise, called acoustic analogies, related these unsteady stresses in the fluid to the noise that they would generate in a uniform ambient medium with a constant speed of sound. Such sources are called quadrupoles. The sound power that they radiate scales with the eighth power of the characteristic source velocity. *See* TURBULENT FLOW.

Subsequent extensions of these theories allowed for the motion of these sources relative to a listener. This may result in a Doppler shift in frequency and a convective amplification of the sound if the sources are moving toward the listener. In addition, if the sources are embedded in a sheared flow, such as the exhaust plume of a jet engine, the sound is refracted away from the jet downstream axis. As sound propagates away from the source region, it experiences attenuation due to spherical spreading and real-gas and relaxational effects. The latter effects are usually important only for high-amplitude sound or sound propagation over large distances. *See* ATMOSPHERIC ACOUSTICS; DOPPLER EFFECT; SOUND ABSORPTION.

Aircraft noise. The main sources of aircraft noise are the jet engine and the airframe. Propeller-driven aircraft and helicopters have additional noise sources.

Engine noise. A modern commercial jet engine has a high bypass ratio, that is, the ratio of the mass flow rates through the fan and core (**Fig. 1**). An alternative design mixes the fan and core streams internally and has a single exhaust nozzle. The sound power generated by the jet exhaust plume is proportional to the eighth power of the exit velocity and the square of the exit area. The high bypass ratio reduces the amount of thrust generated by the high-speed flow and replaces it with thrust delivered by the relatively low-speed fan flow. This reduces the overall radiated noise power. For jets with a subsonic exit velocity, turbulent mixing generates stresses in the exhaust that in turn generate the radiated noise. The directivity of the jet noise peaks at a relatively small angle to the jet downstream direction. This is a result of the convective amplification and refraction effects. At high jet exit velocities, the turbulence may convect supersonically with respect to the ambient speed of sound. This results in highly directional and efficient noise radiation called Mach wave radiation. These are all examples of jet mixing noise. *See* AIRCRAFT ENGINE PERFORMANCE.

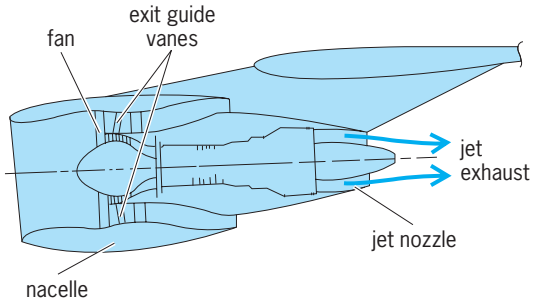

Fig. 1. Separate-flow high-bypass-ratio jet engine.

If the jet is operating off-design, that is, the jet exit static pressure is different from the ambient pressure, a shock cell structure forms in the jet exhaust. Shock cells are diamond-shaped regions of alternating high and low pressure in the jet exhaust through which the jet exit pressure adjusts to the ambient pressure. The interaction between the turbulence in the jet and the shock cell structure generates broadband shock-associated noise and screech. The former noise radiates primarily in the upstream direction, is broadband, and peaks at frequencies higher than the jet mixing noise. Screech is a feedback phenomenon in which sound generated by the turbulence interacting with the shock cells propagates upstream and triggers turbulence at the jet exit. This convects downstream, interacting again with the shock cells and completing the cycle. The feedback cycle generates an intense tone known as jet screech.

Noise is generated by the fan in the high-bypass-ratio engine due to interactions between the wakes shed by the fan blades and the exit guide vanes. This interaction generates spinning acoustic modes that may propagate in the fan duct, either upstream toward the engine inlet or downstream through the fan exhaust duct. (Spinning acoustic modes are the natural acoustic waveforms in a circular duct. They consist of pressure disturbances with wave fronts that follow helical paths.) The noise consists of a tonal component proportional to the fan's blade passage frequency and a broadband component associated with the turbulent wake interactions as well as turbulence ingested into the fan duct. The turbulent interactions cause unsteady loading on the fan blades and the exit or outlet guide vanes. The reaction forces acting on the fluid generate noise. Acoustic liners, usually made of perforated plates with a honeycomb backing, can attenuate the duct modes. In addition, the choice of the relative number of fan blades to exit guide vanes, as well as the tip speed of the fan, affects the noise generation efficiency. *See* AIRCRAFT ENGINE PERFORMANCE; JET PROPULSION; TURBOFAN; WAKE FLOW.

Airframe noise. Airframe noise is generated by the nonpropulsive components of an aircraft. With continued reduction in engine noise, airframe noise can be a significant component of the overall aircraft noise, especially during landing. The major sources

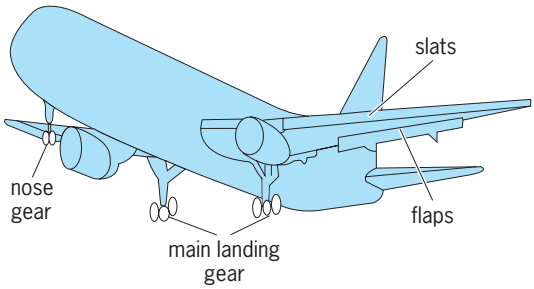

Fig. 2. Sources of airframe noise.

of airframe noise are associated with the high-lift devices, namely the trailing-edge flaps and leading-edge slats, and the landing gear—both the nose and main gear (**Fig. 2**). *See* FLIGHT CONTROLS; LANDING GEAR.

Flap systems may consist of multiple elements. Also, they may not extend over the entire wingspan. Gaps or gates may be used to limit engine exhaust impingement on the flaps. Vortices are shed from the flap side edges. The strength of the vortices depends on the loading distribution on the flap. The vortex moves over the flap side edge, inducing unsteady surface pressures on the flap, resulting in dipole noise radiation. In addition, the turbulence in the vortex generates noise that can scatter from the side edge corners or flap trailing edge. Since the loading noise scales with the sixth power of the local Mach number, and corners and edges scatter noise that scales with the fifth power, the latter mechanism can dominate at low Mach number. *See* MACH NUMBER; VORTEX.

Noise is generated at the trailing edge of the leading-edge slat. This can be tonal and is believed to be associated with vortex shedding from the slat trailing edge. This vortex shedding may be reinforced by a resonant feedback in which the radiated noise reflects from the main element of the wing. The vortex shedding depends on the slat trailing-edge thickness as well as the boundary layer thicknesses on the upper and lower surfaces of the slat.

The landing gear acts as a bluff body and generates noise associated with unsteady loading in a similar manner to the vortex-shedding noise of a circular cylinder. If any sharp edges are in the vicinity of the landing gear, such as a door that opens with the gear, acoustic scattering may occur, which increases the efficiency of the noise radiation. In addition, small-scale fittings, such as hydraulic lines attached to the gear, generate noise at a higher frequency, due to their smaller scale, than that generated by the wheels and the main gear components.

Propeller and rotorcraft noise. The noise generated by propellers contains a broadband and a harmonic component. The harmonic component occurs at multiples of the blade passage frequency. It consists of thickness noise and loading noise. Thickness noise is caused by the displacement of the air by the rotating propeller blade. It is a monopolelike source acting at the blade surface. Loading noise is related to the lift and drag forces acting on the blade. The source is of dipole type. For a propeller not in forward motion, the noise behind the plane of the propeller blades is greater than that ahead of the propeller. Forward flight causes convective amplification so that noise levels are greater when the aircraft is approaching a stationary observer.

As is the case for fan noise, broadband propeller noise is associated with turbulence ingestion and blade self-noise. The latter sources include the surface pressure fluctuations caused by the turbulent boundary layer on the blade and vortex shedding from the blade trailing edge. If the propeller tip speed is supersonic, sound generated at several points on the blade surface can reach a listener at the same time. The reason is that the source region is moving faster relative to the listener than the local speed of sound. The listener hears an intense, crackling sound during a short fraction of the propeller's rotation period. *See* PROPELLER (AIRCRAFT).

The same harmonic and broadband sources arise in helicopter noise generation. The noise levels for a helicopter in hover are greater below the plane of the rotor. In addition, in descending flight the rotor blades may encounter tip vortices shed from the passage of previous blades. This gives a transient loading to the blade and an impulsive noise radiation called blade vortex interaction (BVI) noise. The characteristic signatures of this type of noise differ, depending on whether the interaction occurs on the advancing or retreating sides of the rotor plane. In forward flight, the flow at the tip of the advancing rotor can become locally supersonic. The spatial signature of the shock that closes the supersonic region can radiate to the acoustic far field. This is referred to as high speed impulsive (HSI) noise. *See* HELICOPTER; SOUND.

Philip J. Morris

Bibliography. H. H. Hubbard, *Aeroacoustics of Flight Vehicles*, vols. 1 and 2, Acoustical Society of America, 1995.

Aerodynamic wave drag

The force retarding an airplane, especially in supersonic flight, as a consequence of the formation of shock waves. Although the physical laws governing flight at speeds in excess of the speed of sound are the same as those for subsonic flight, the nature of the flow about an airplane and, as a consequence, the various aerodynamic forces and moments acting on the vehicle at these higher speeds differ substantially from those at subsonic speeds. Basically, these variations result from the fact that at supersonic speeds the airplane moves faster than the disturbances of the air produced by the passage of the airplane. These disturbances are propagated at roughly the speed of sound and, as a result, primarily influence only a region behind the vehicle.

Causes of wave drag. The primary effect of the change in the nature of the flow at supersonic speeds is a marked increase in the drag, resulting from the formation of shock waves about the configuration. These strong disturbances, which may extend for many miles from the airplane, cause significant

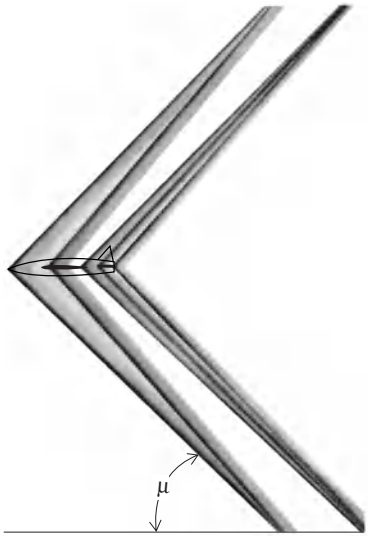

Fig. 1. Shock waves about an airplane at supersonic speeds. μ = angle of obliqueness.

energy losses in the air, the energy being drawn from the airplane. At supersonic flight speeds these waves are swept back obliquely, the angle of obliqueness decreasing with speed (**Fig. 1**). For the major parts of the shock waves from a well-designed airplane, the angle of obliqueness is equal to $\sin^{-1}(1/M)$, where M is the Mach number, the ratio of the flight velocity to the speed of sound. *See* SHOCK WAVE; SUPERSONIC FLIGHT.

The shock waves are associated with outward diversions of the airflow by the various elements of the airplane. This diversion is caused by the leading and trailing edges of the wing and control surfaces, the nose and aft end of the fuselage, and other parts of the vehicle. Major proportions of these effects also result from the wing incidence required to provide lift.

Magnitude of wave drag. Usually the aerodynamic forces acting on a vehicle, such as drag, are considered as coefficients. The wave-drag coefficient for a vehicle depends on many parameters, including the thickness-to-chord ratios and shapes of the airfoil sections of wings and fins, the planform shapes of such surfaces, the length-to-diameter ratio of the body, and the shape of the engine housing. It also depends on the Mach number. For an unswept surface with a thin airfoil whose shape is designed for wave-drag reduction (as discussed below), the coefficient of wave drag is given approximately by Eq. (1), where t/c is

$$C_d = \frac{4}{\sqrt{M^2 - 1}}\left[(t/c)^2 + \alpha^2\right] \qquad (1)$$

the airfoil thickness-to-chord ratio and α is the angle of attack in radians. For surfaces with more complex airfoil shapes and planforms, the computation of wave drag is more complex.

For complete vehicles, the wave drag at zero lift is given by the area rule. For supersonic flight, the cross-sectional areas used are obtained in planes inclined at the Mach angle. At the lower supersonic speeds, the wave drag at the zero lift condition is usually more significant than the drag due to wing incidence. But when the Mach number is increased, the relative magnitude of wave drag at the zero lift condition gradually decreases, and the drag associated with wing incidence progressively increases, so at the higher supersonic speeds wave drag due to lift is usually more important than the zero lift value. For a well-designed vehicle, wave drag is usually roughly equal to the sum of the basic skin friction and the induced drag due to lift. *See* AERODYNAMIC FORCE; AIRFOIL; TRANSONIC FLIGHT.

Reduction of wave drag. The wave drag at the zero lift condition is reduced primarily by decreasing the thickness-chord ratios for the wings and control surfaces and by increasing the length-diameter ratios for the fuselage and bodies. Also, the leading edge of the wing and the nose of the fuselage are made relatively sharp (**Fig. 2**). With such changes, the severity of the diversions of the flow by these elements is reduced, with a resulting reduction of the strength of the associated shock waves. Also, the supersonic drag wave can be reduced by shaping the fuselage and arranging the components on the basis of the area rule. *See* WING; WING STRUCTURE.

The wave drag can also be reduced by sweeping the wing panels (**Fig. 3***a*). Some wings intended for supersonic flight have large amounts of leading-edge sweep and little or no trailing-edge sweep (Fig. 3*b*). Such planforms are referred to as delta or modified delta. For a simple infinite-span, constant-chord airfoil, the effective Mach number determining the aerodynamic characteristics is the component of the flight Mach number normal to the swept elements (**Fig. 4**). This Mach number is defined by

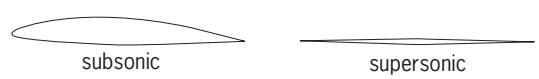

Fig. 2. Comparison of airfoil sections for subsonic flight and supersonic flight.

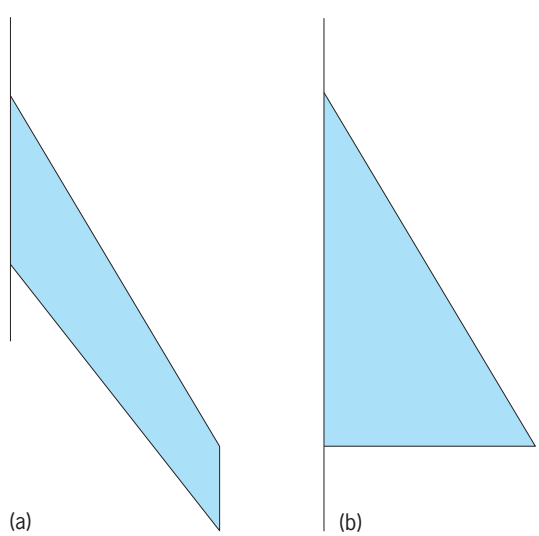

Fig. 3. Wing panels. (*a*) Sweep-back. (*b*) Delta.

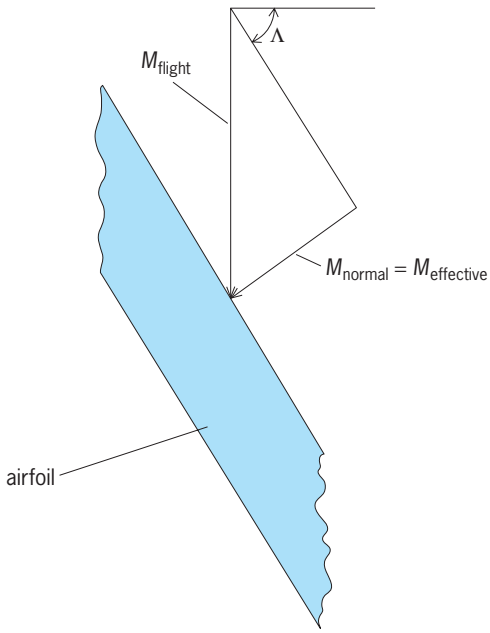

Fig. 4. Geometry that determines the effective Mach number for an infinite-span swept airfoil.

Eq. (2), where Λ is the angle of the sweep. If the

$$M_{normal} = M_{flight} \times \cos \Lambda \qquad (2)$$

sweep is such that M_{normal} is less than the subsonic Mach number at which the initial onset of wave drag occurs for the unswept airfoil, the swept airfoil will have no wave drag for the flight Mach number. For flight Mach numbers for which the normal component is greater than that for the onset of a shock on the unswept airfoil, the wave drag is not eliminated but is significantly reduced. For practical finite-span swept or delta wings, the reductions of wave drag are usually less than for the ideal infinite-span case. These reductions can be substantially improved by the proper chordwise and spanwise variations of thickness, camber, and incidence. The shape changes required are now determined using very complex fluid-dynamic relationships and supercomputers. *See* COMPUTATIONAL FLUID DYNAMICS.

When the speed is increased to supersonic values, an airplane at a given attitude and altitude experiences large increases in drag, in addition to those associated with the different nature of the flow, because of the higher dynamic pressure at these higher speeds. To offset this effect, supersonic airplanes usually fly at considerably higher altitudes than subsonic vehicles. For example, for efficient flight at Mach 2, an airplane must fly at an altitude of about 60,000 ft (18,000 m).

A major problem associated with supersonic flight, particularly at the higher supersonic speeds, is that of taking air into the engines. This air must be decelerated from the flight velocity to a relatively low speed at the compressor of the engine, without excessive energy losses. With a simple inlet, such as that used on subsonic and transonic airplanes, a strong normal shock wave forms ahead of the forward face at supersonic speeds. This shock causes severe loss of energy in the air reaching the engine, and consequent losses of engine performance. In addition, the drag of the airplane is increased. To reduce these losses, special inlets and diffusers which decelerate the airflow to the engine by a series of weak disturbances are used. *See* SUPERSONIC DIFFUSER.

Richard T. Whitcomb

Bibliography. J. D. Anderson, Jr., *Fundamentals of Aerodynamics*, 4th ed., 2005; J. J. Bertin, *Aerodynamics for Engineers*, 4th ed., 2002; A. M. Kuethe and C.-Y. Chow, *Foundations of Aerodynamics: Bases of Aerodynamic Design*, 5th ed., 1997.

Aerodynamics

The applied science that deals with the dynamics of airflow and the resulting interactions between this airflow and solid boundaries. The solid boundaries may be a body immersed in the airflow, or a duct of some shape through which the air is flowing. Although, strictly speaking, aerodynamics is concerned with the flow of air, now the term has been liberally interpreted as dealing with the flow of gases in general.

Depending on its practical objectives, aerodynamics can be subdivided into external and internal aerodynamics. External aerodynamics is concerned with the forces and moments on, and heat transfer to, bodies moving through a fluid (usually air). Examples are the generation of lift, drag, and moments on airfoils, wings, fuselages, engine nacelles, and whole airplane configurations; wind forces on buildings; the lift and drag on automobiles; and the aerodynamic heating of high-speed aerospace vehicles such as the space shuttle. Internal aerodynamics involves the study of flows moving internally through ducts. Examples are the flow properties inside wind tunnels, jet engines, rocket engines, and pipes. In short, aerodynamics is concerned with the detailed physical properties of a flow field and also with the net effect of these properties in generating an aerodynamic force on a body immersed in the flow, as well as heat transfer to the body. *See* AERODYNAMIC FORCE; AEROTHERMODYNAMICS.

Types of aerodynamic flow. Aerodynamics can also be subdivided into various categories depending on the dominant physical aspects of a given flow. In low-density flow the characteristic size of the flow field, or a body immersed in the flow, is of the order of a molecular mean free path (the average distance that a molecule moves between collisions with neighboring molecules); while in continuum flow the characteristic size is much greater than the molecular mean free path. More than 99% of all practical aerodynamic flow problems fall within the continuum category. *See* RAREFIED GAS FLOW.

Continuum flow can be subdivided into viscous flow, which is dominated by the dissipative effects of viscosity (friction), thermal conduction, and mass

diffusion; and inviscid flow, which is, by definition, a flow in which these dissipative effects are negligible. Both viscous and inviscid flows can be subdivided into incompressible flow, in which the density is constant, and compressible flow, in which the density is a variable. In low-speed gas flow, the density variation is small and can be ignored. In contrast, in a high-speed flow the density variation is keyed to temperature and pressure variations, which can be large, so the flow must be treated as compressible. *See* COMPRESSIBLE FLOW; FLUID FLOW; INCOMPRESSIBLE FLOW; VISCOSITY.

Compressible flow regimes. In turn, compressible flow is subdivided into four speed regimes: subsonic flow, transonic flow, supersonic flow, and hypersonic flow. These regimes are distinguished by the value of the Mach number, which is the ratio of the local flow velocity to the local speed of sound (**Fig. 1**). *See* MACH NUMBER.

Subsonic flow. A flow is subsonic if the Mach number is less than 1 at every point. Subsonic flows are characterized by smooth streamlines with no discontinuity in slope (Fig. 1*a*). Disturbances (caused by, say, the sudden deflection of the trailing edge of the airfoil) propagate both upstream and downstream and are felt throughout the entire flow field. The flow over light, general-aviation airplanes is subsonic. *See* SUBSONIC FLIGHT.

Transonic flow. A transonic flow is a mixed region of locally subsonic and supersonic flow. The flow far upstream of the airfoil can be subsonic (Fig. 1*b*), but as the flow moves around the airfoil surface it speeds up, and there can be pockets of locally supersonic flow over both the top and bottom surfaces of the airfoil. If the flow far upstream of the airfoil is at a low supersonic value, say with a Mach number of 1.1 (Fig. 1*c*), a bow shock wave is formed in front of the body. Behind the nearly normal portion of the shock wave, the flow is locally subsonic. This subsonic flow subsequently expands to a low supersonic value over the airfoil. *See* SHOCK WAVE; TRANSONIC FLIGHT.

Supersonic flow. In a supersonic flow, the local Mach number is greater than 1 everywhere in the flow. Supersonic flows are frequently characterized by the presence of shock waves. Across shock waves, the flow properties and the directions of streamlines change discontinuously (Fig. 1*d*), in contrast to the smooth, continuous variations in subsonic flow. If a disturbance is introduced into a steady supersonic flow at some point, the effects of this disturbance are not felt upstream. Instead, these disturbances pile up along a front just in front of the nose, creating a shock wave in the flow. *See* SUPERSONIC FLIGHT.

Hypersonic flow. This is a regime of very high supersonic speeds. A conventional rule is that any flow with a Mach number equal to or greater than 5 is hypersonic (Fig. 1*e*). Examples include the space shuttle during ascent and reentry into the atmosphere, and the flight of the X-15 experimental vehicle. The kinetic energy of many hypersonic flows is so high that, in regions where the flow velocity decreases, kinetic energy is traded for internal energy of the gas, creating high temperatures. Aerodynamic heating is a particularly severe problem for bodies immersed in a hypersonic flow. *See* ATMOSPHERIC ENTRY; HYPERSONIC FLIGHT.

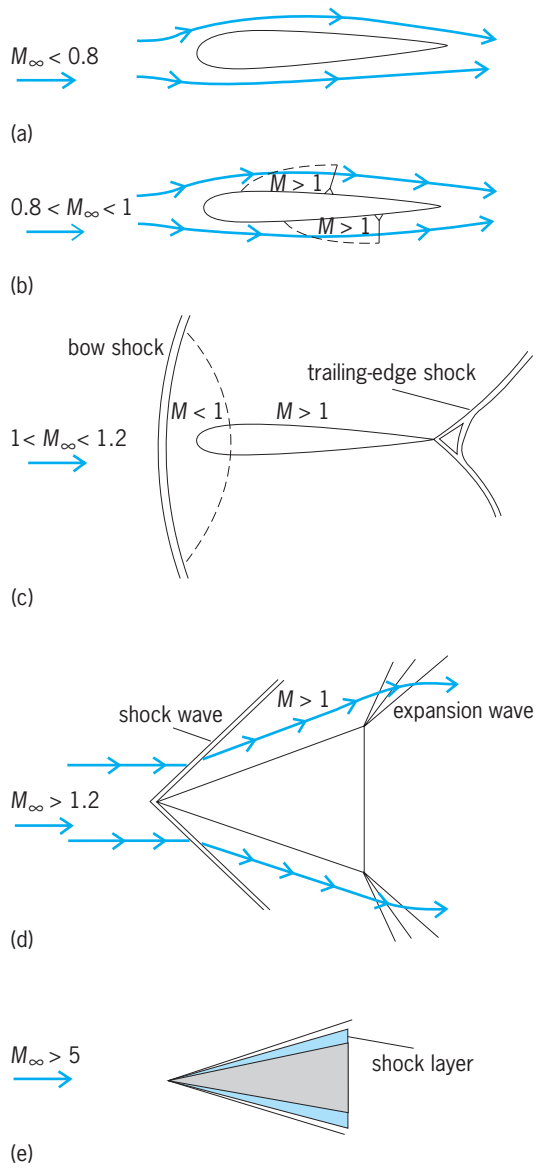

Fig. 1. Regimes for aerodynamic flows; M_∞ = Mach number for flow far upstream of the airfoil: (*a*) subsonic flow; (*b*) transonic flow with $M_\infty < 1$; (*c*) transonic flow with $M_\infty > 1$; (*d*) supersonic flow; and (*e*) hypersonic flow, where the thin, hot shock layer has viscous interaction and chemical reactions.

Nonaeronautical applications. Applications of aerodynamics cut across various disciplines. For example, automobile aerodynamics has taken on new importance in light of the incentive for energy-efficient vehicles. External aerodynamics is important in the reduction of aerodynamic drag on an automobile. With wind-tunnel testing (**Fig. 2**) and computational aerodynamic analyses, this drag has gradually been reduced by up to 70%, and similar improvements have been reported for trucks and trailers. Internal aerodynamics is important in understanding flows through engine and exhaust systems, thereby contributing to the

Fig. 2. Aerodynamic flow over an automobile. (*Ford Motor Co.*)

design of engines with improved gas mileage. In another field, environmental aerodynamics focuses on the study of airflows in heating and air-conditioning systems, and on the circulation patterns of air inside buildings and homes. Studies of the external airflow around buildings, especially under hurricane conditions, are of great interest. *See* COMPUTATIONAL FLUID DYNAMICS; GAS DYNAMICS; WIND TUNNEL.

John D. Anderson, Jr.

Bibliography. J. D. Anderson, Jr., *Fundamentals of Aerodynamics*, 4th ed., 2005; J. D. Anderson, Jr., *Introduction to Flight*, 5th ed., 2005; J. J. Bertin, *Aerodynamics for Engineers*, 4th ed., 2002; A. M. Kuethe and C. Y. Chow, *Foundations of Aerodynamics*, 5th ed., 1997.

Aeroelasticity

The branch of applied mechanics which deals with the interaction of aerodynamic, inertial, and structural forces. It is important in the design of airplanes, helicopters, missiles, suspension bridges, power lines, tall chimneys, and even stop signs. Variations on the term aeroelasticity have been coined to denote additional significant interactions. Aerothermoelasticity is concerned with effects of aerodynamic heating on aeroelastic behavior in high-speed flight. Aeroservoelasticity deals with the interaction of automatic controls and aeroelastic response and stability.

The primary concerns of aeroelasticity include flying qualities (that is, stability and control), flutter, and structural loads arising from maneuvers and atmospheric turbulence. Methods of aeroelastic analysis differ according to the time dependence of the inertial and aerodynamic forces that are involved. For the analysis of flying qualities and maneuvering loads wherein the aerodynamic loads vary relatively slowly, quasi-static methods are applicable, although autopilot interaction could require more general methods. The remaining problems are dynamic, and methods of analysis differ according to whether the time dependence is arbitrary (that is, transient or random) or simply oscillatory in the steady state.

Lift distribution and divergence. The redistribution of airloads caused by structural deformation will change the lifting effectiveness on the aerodynamic surfaces from that of a rigid vehicle. The simultaneous analysis of equilibrium and compatibility among the external airloads, the internal structural and inertial loads, and the total flow disturbance, including the disturbance resulting from structural deformation, leads to a determination of the equilibrium aeroelastic state. If the airloads tend to increase the total flow disturbance, the lift effectiveness is increased; if the airloads decrease the total flow disturbance, the effectiveness decreases. In a limiting case of increasing lift effectiveness, there is a critical speed at which the rate of change of airload with deformation equals the rate of change of structural reaction, and no statically stable equilibrium condition exists; at a higher speed the deformation will increase to a point of structural failure. This critical speed is called the divergence speed.

The effectiveness of a lifting surface influences the selection of planform size. Optimum design for strength requires design for the actual loads applied, taking due account of the redistribution of loading resulting from flexibility. The extent of changes in effectiveness and load distribution on a straight wing depends on its torsional stiffness and the chordwise distance between the aerodynamic center of the airfoil and its center of twist. Divergence can occur on a straight wing if the airfoil aerodynamic center is forward of the center of twist. Sweep has a significant influence on effectiveness and redistribution of loads to an extent also determined by the surface bending stiffness, while the effectiveness of stabilizers is also influenced by the stiffness of the fuselage. The bending of a swept-back wing has a substantial stabilizing effect on divergence to the extent that only a slight amount of sweep is required to eliminate the possibility of divergence. The redistribution of wing loading has a slight effect on the induced drag, but if the wing is swept it may have a more significant effect on static longitudinal stability. If the wing is swept-back, the redistribution of loading usually shifts the spanwise center of pressure location inboard, and correspondingly the chordwise center of pressure location moves forward. This forward movement of center of pressure reduces the static stability. The same effect occurs on swept-back stabilizing surfaces and reduces effective tail lengths. This can be an extremely critical problem on a vertical stabilizer since it provides the primary source of static directional stability, whereas it is the combination of wing and horizontal stabilizer that determines the static longitudinal stability. The bending of a wing will also introduce a change in dihedral that may affect the dynamic lateral-directional stability of the aircraft. *See* WING STRUCTURE.

On each of the attempts to attain crewed, powered flight in October and December of 1903, a wing of Samuel Langley's aerodome collapsed, and the failures may have been caused by torsional divergence. The first successful flight of a power-driven heavier-than-air machine was achieved the week following

Langley's second failure, when the Wright brothers' biplane flew at Kitty Hawk, North Carolina, on December 17. The Wright biplane experienced no aeroelastic difficulties and, in fact, made use of an aeroelastic control system in which warping of the wing gave rise to the necessary controlling aerodynamic forces. The Wright brothers had also noted a reduction in thrust due to twisting in their experimental work on propeller blades, although it had no significant effect on the blades' performance. During World War I the Fokker D-8 had just been put into service when three failures occurred in high-speed maneuvers. As a result of static strength and deflection measurements, it became apparent that the failures were due to torsional divergence; the load resulting from the air pressure in a steep dive would increase faster at the wing tips than at the middle.

Control effectiveness and reversal. The airloads induced by means of a control-surface deflection also induce an aeroelastic loading of the entire system. Equilibrium is determined as above in the analysis of load redistribution. Again, the effectiveness will differ from that of a rigid system, and may increase or decrease depending on the relationship between the net external loading and the deformation. In a situation of decreasing control-surface effectiveness, the speed at which the effectiveness vanishes is called the reversal speed. At this speed the control-surface–induced air loads are exactly offset by the resulting aeroelastic loads. At higher speeds the aeroelastic loads will exceed the control-surface loading, and the resultant load will act in the reverse direction from the control-surface loading, causing the control system to act in a direction opposite to that desired.

Control effectiveness determines the selection of control-surface size and maximum travel required. It may determine the need for all-movable horizontal or vertical stabilizers. Loss in control-surface effectiveness on a straight planform depends on the torsional stiffness of the forward primary surface, to some extent on the torsional stiffness of the control surface as well as its actuation system, and on the distance between the airfoil aerodynamic center and the aerodynamic center of the loading induced by the control surface. Sweepback, which is favorable in the case of divergence, has the unfavorable effect of decreasing control-surface effectiveness. The effectiveness of a tail or canard control surface is also influenced by fuselage bending and, in the case of a vertical stabilizer, torsion of the fuselage. Control reversal is an extremely critical problem with ailerons, and may involve severe weight penalties in providing sufficient torsional stiffness and, in the case of a swept-back wing, bending stiffness. The problem is less critical with elevators and rudders and, of course, is eliminated by the use of all-movable stabilizers for control. Unlike divergence, the phenomenon of reversal is not destructive in itself, but it can lead to structural failure when a marginally stable vehicle cannot be controlled. *See* AILERON; AIRCRAFT RUDDER; ELEVATOR (AIRCRAFT).

Aileron reversal first became a crucial problem for World War II fighters. High roll rates at high speeds were essential in combat maneuvers, and any loss in rolling ability put a fighter pilot at a serious disadvantage. The designers of the Spitfire, the FW-190, and the P-51 sought very high roll rates, and this led to shorter wingspans, aerodynamically balanced ailerons for larger deflection at high speeds, and higher torsional stiffness to increase the reversal speed well beyond the performance limits of the aircraft. The Japanese Zero, on the other hand, was designed for lightness and minimum-radius turns which resulted in a low reversal speed. The U.S. Navy pilots soon discovered this weakness and avoided circling combat but established high-speed, single-pass techniques, where their superior dive speeds and high roll rates could not be followed by the Zero operating close to its aileron reversal speed.

Flutter. A self-excited vibration is possible if a disturbance to an aeroelastic system gives rise to unsteady aerodynamic loads such that the ensuing motion can be sustained. At the flutter speed a critical phasing between the motion and the loading permits extraction of an amount of energy from the airstream equal to that dissipated by internal damping during each cycle and thereby sustains a neutrally stable periodic motion. At lower speeds any disturbance will be damped, while at higher speeds, or at least in a range of higher speeds, disturbances will be amplified. The simplest type of flutter occurs when a single motion induces an aerodynamic force having a component in the direction of the motion and in phase with the velocity. This is described as a case of negative aerodynamic damping or single-degree-of-freedom flutter. An example is a power line that "gallops" as a result of ice formation which gives the cross section of the cable something of an airfoil shape. The term classical flutter is used to denote the more complicated instability that typically arises from a critical coupling of two or more modes of motion, each of which is stable by itself.

The mechanism for classical flutter may be described in terms of the bending deflection, the angle of twist, and the disturbance in lift. If the structural characteristics (that is, stiffness and inertia) are such that an upward incremental lift, corresponding to a leading-edge upward twist, occurs when the airfoil is bending upward, and, when the motion reverses, the incremental lift acts downward as the airfoil is bending downward, energy will be extracted from the airstream. If the energy extracted in each cycle exceeds that dissipated by internal structural damping, the motion will oscillate with an increasing amplitude until a structural failure occurs.

The parameters of the flutter problem are those of the divergence problem with the addition of the dynamic inertial characteristics. The flutter speed is basically proportional to the fundamental torsion frequency so that torsional stiffness is a primary design variable. A forward center of gravity is stabilizing because it tends to drive the bending and twisting motions out of phase, which results in the incremental lift acting in the opposite direction from the bending motion. Mass balance is therefore

another primary design variable. Flutter speed is usually reduced at high altitudes because aerodynamic damping decreases faster than aerodynamic stiffness (that is, lift effectiveness) with air density. Flutter is essentially a resonance phenomenon between two motions, so that another design consideration is separation of natural vibration frequencies. The effect of sweepback is not as significant on flutter characteristics as it is with divergence, although increasing the sweepback angle does increase the flutter speed. An interesting design possibility is to suppress flutter by automatic control systems, and development of such systems has been undertaken. In order to control the aerodynamic forces, either an additional control surface or a rapidly responding actuator for an existing control surface is required, and an example configuration is a primary surface with both leading- and trailing-edge control surfaces. For practical applications the automatic control system will need multiple redundancies since a component failure could result in violent flutter.

The first known case of aerodynamic flutter occurred during World War I on the horizontal tail of a Handley-Page 400, twin-engine bomber. The flutter apparently was caused by coupling of the fuselage torsion mode and the antisymmetrical rotations of the independently actuated right and left elevators. The coupling was eliminated by connecting the elevators to a common actuating torque tube. Toward the end of the war, tail flutter difficulties were overcome by the same redesign, and the attachment of elevators to the same torque tube became a standard feature of subsequent designs. Soon after the war, static mass balancing of ailerons about their hinge lines was found to be an effective means of avoiding wing-aileron flutter. This technique has proved to be both a design and redesign procedure of permanent value, for, ever since, there have been occasional instances of coupled control-surface flutter involving both wings and tail surfaces that did not lead to destruction but could be eliminated merely by increasing the static mass balance of the control surface. *See* FLUTTER (AERONAUTICS).

Gust response. Transient meteorological conditions such as wind shears, vertical drafts, mountain waves, and clear air or storm turbulence impose significant dynamic loads on aircraft. So does buffeting during flight at high angles of attack or at transonic speeds. The response of the aircraft determines the stresses in the structure and the comfort of the occupants. Aeroelastic behavior makes a condition of dynamic overstress possible; in many instances, the amplified stresses can be substantially higher than those that would occur if the structure were much stiffer. *See* CLEAR-AIR TURBULENCE; LOADS, DYNAMIC; TRANSONIC FLIGHT.

Structural design considerations for gusts include both strength and fatigue life. Atmospheric turbulence and buffeting are random phenomena and can be described only statistically, but wind shears and mountain waves take on the appearance of discrete gusts. Since statistical data on all atmospheric conditions are limited, recourse to arbitrary design criteria is necessary. Strength has generally been determined for response to a discrete gust with a specified profile and maximum velocity normal to the flight path, while fatigue life is estimated by statistical methods based on the power spectral characteristics of the atmosphere. However, strength is frequently also calculated by the statistical methods since static strength may be regarded as the limiting case of fatigue under a single loading.

The gust loading condition is a critical one in the design of the wings of large aircraft, and the amount of sweep and the location of wing-mounted engines influence the dynamic response. The basic parameters that determine the response are the longitudinal dynamic flight characteristics of the aircraft, primarily its short-period frequency and damping, and the bending stiffness and frequency of the wing. These same factors, coupled with fuselage flexibility, also determine the aircraft ride qualities. Both gust response loads and ride qualities have been improved by use of automatic control systems.

The gustiness of the atmosphere has obviously been a concern throughout the history of flight. The first report of the National Advisory Committee for Aeronautics [now the National Aeronautics and Space Administration (NASA)], published in 1915, was an experimental and theoretical investigation of aircraft response to gusts. The first practical applications of automatic controls for gust alleviation and structural dynamic stability augmentation were made to the B-52 and B-70 bombers. Successful demonstrations of the load alleviation systems were begun with B-52 flights in late 1967 and B-70 flights in 1968. *See* AERODYNAMICS; FLIGHT CHARACTERISTICS; SUBSONIC FLIGHT.

William P. Rodden

Bibliography. R. L. Bisplinghoff and H. Ashley, *Principles of Aeroelasticity*, 2d ed., 1962, reprint 2002; E. H. Dowell (ed.), *A Modern Course in Aeroelasticity*, 4th ed., 2004; Y.-C. B. Fung, *An Introduction to the Theory of Aeroelasticity*, 1955, reprint 2002; W. P. Jones (ed.), *AGARD Manual on Aeroelasticity*, 1960–1968; D. McRuer, I. Ashkenas, and D. Graham, *Aircraft Dynamics and Automatic Control*, 1973.

Aeromonas

A bacterial genus in the family Vibrionaceae comprising oxidase-positive, facultatively anaerobic, monotrichously flagellated gram-negative rods. The mesophilic species are *A. hydrophila, A. caviae* and *A. sobria*; the psychrophilic one is *A. salmonicida*. Aeromonads are of aquatic origin and are found in surface and wastewater but not in seawater. They infect chiefly cold-blooded animals such as fishes, reptiles, and amphibians and only occasionally warm-blooded animals and humans.

Human wound infections may occur following contact with contaminated water. Septicemia has been observed mostly in patients with abnormally low white blood counts or liver disease. There is evidence of intestinal carriers. The three mesophilic species are also associated with diarrheal disease (enteritis and colitis) worldwide. A heat-labile

enterotoxin and cytotoxin(s) are produced only by *A. hydrophila* and *A. sobria*, but experimental diarrhea could not be produced in humans. *See* DIARRHEA.

A related lophotrichous genus, *Plesiomonas* (single species, *P. shigelloides*), is also known as an aquatic bacterium and is associated with diarrhea chiefly in subtropical and tropical areas. It is also found in many warm-blooded animals. Systemic disease in humans is rare. *See* MEDICAL BACTERIOLOGY.

Alexander von Graevenitz

Bibliography. E. H. Lennette et al. (eds.), *Manual of Clinical Microbiology*, 4th ed., 1985.

Aeronautical engineering

That branch of engineering concerned primarily with the special problems of flight and other modes of transportation involving a heavy reliance on aerodynamics or fluid mechanics. The main emphasis is on airplane and missile flight, but aeronautical engineers work in many related fields such as hydrofoils, which have many problems in common with aircraft wings, and with such devices as air-cushion vehicles, which make use of airflow around the base to lift the vehicle a few feet off the ground, whereupon it is propelled forward by use of propellers or gas turbines. *See* AERODYNAMICS; AIRPLANE; FLUID MECHANICS; HYDROFOIL CRAFT.

Aeronautical engineering expanded dramatically after 1940. Flight speeds increased from a few hundred miles per hour to satellite and space-vehicle velocities. The common means of propulsion changed from propellers to turboprops, turbojets, ramjets, and rockets. This change gave rise to new applications of basic science to the field and a higher reliance on theory and high-speed computers in design and testing, since it was often not feasible to proceed by experimental methods only. *See* JET PROPULSION; PROPULSION; ROCKET PROPULSION; SPACE TECHNOLOGY.

Aeronautical engineers frequently serve as system integrators of important parts of a design. For example, the control system of an aircraft involves, among other considerations, aerodynamic input from flow calculations and wind-tunnel tests; the structural design of the aircraft (since the flexibility and strength of the structure must be allowed for); the mechanical design of the control system itself; electrical components, such as servo-mechanisms; hydraulic components, such as hydraulic boosters; and interactions with other systems that affect the control of the aircraft, such as the propulsion system. The aeronautical engineer is responsible for ensuring that all of these factors operate smoothly together.

The advent of long-range ballistic missiles and spacecraft presented new challenges for aeronautical engineering. Flight velocities increased up to 35,000 ft/s (10,668 m/s), and the resulting heating problems required new kinds of materials and cooling-system concepts to protect the spacecraft or reentry vehicle. For example, materials were developed by ground testing in high-temperature electric arcs and rocket exhausts and then were tested by flights of crewless models at full speed. The resulting heat shields were successfully used to return astronauts from the Moon and to protect the warheads of ballistic missiles while maintaining missile accuracy. *See* BALLISTIC MISSILE.

The techniques that aeronautical engineers had developed for aircraft were applied to help develop such devices as the air-cushion vehicle, which is of much interest to the military for landing troops on a beach. Helicopters and vertical takeoff and landing aircraft are other developments in which aeronautical engineering has played an important role. Hydrofoils are basically boats that fly on wings under the water. Because the water is so dense, these wings are usually small compared to aircraft wings, but the same principles apply. The control of such wings is very sensitive, and careful attention must be paid to the control system used lest the high forces involved cause severe damage to the wings or to the boat itself. *See* AIR-CUSHION VEHICLE; HELICOPTER; VERTICAL TAKEOFF AND LANDING (VTOL).

Aerodynamics has always been one of the main tools of aeronautical engineering. As noted above, aerodynamics interacts closely with control design. For example, as an aircraft approaches the ground to land, the effect of the solid barrier of the ground affects the aerodynamics of the aircraft. This effect can be simulated in a computer model by placing another wing in a mirror image under the ground. The resulting flow pattern corresponds to the real wing and the solid ground. In any but the simplest cases, this type of calculation cannot be carried out except with the aid of high-speed computers, which have made possible the calculation of entire flow fields around aircraft and spacecraft as they pass through the atmosphere. Wind-tunnel testing of scale models is another important source of aerodynamic data. *See* COMPUTATIONAL FLUID DYNAMICS; WIND TUNNEL.

Aircraft and missile structural engineers have raised the technique of designing complex structures to a level never considered possible before the advent of high-speed computers. Structures can now be analyzed in great detail and the results incorporated directly into computer-aided design (CAD) programs. These programs are now so complete that in some cases no drafting work with paper and pencil is used at all and all drawings are printed directly from the design devised by the engineer working only on the computer screen. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

These examples illustrate the scope of aeronautical engineering. In general, it overlaps many other disciplines, and the aeronautical engineer may concentrate on aerodynamics, structures, control systems, propulsion systems, or the overall design of the vehicle.

John R. Sellars

Bibliography. J. E. Allen, *Aerodynamics,* 1982; W. Biddle, *Barons of the Sky: From Early Flight to Strategic Warfare,* 1991; P. A. Hanle, *Bringing Aerodynamics to America: A History of Aerodynamics Research,* 1982; F. E. Weick, *From the Ground Up: The Autobiography of an Aeronautical Engineer,* 1988.

Aeronautical meteorology

The branch of meteorology that deals with atmospheric effects on the operation of vehicles in the atmosphere, including winged aircraft, lighter-than-air devices such as dirigibles, rockets, missiles, and projectiles. The air which supports flight or is traversed on the way to outer space contains many potential hazards. Nine major hazards are considered in this article, which is principally concerned with commercial and general aviation.

Low visibility. Poor visibility caused by fog, snow, dust, and rain is a major cause of aircraft accidents and the principal cause of flight cancellations or delays. For the average private pilot with a minimum of equipment in the aircraft, an unexpected encounter with low-visibility conditions can be a serious matter, although technology is bringing aviation closer to the hitherto elusive goal of all-weather flying.

The weather conditions of ceiling and visibility required by regulations for crewed aircraft during landing or takeoff are determined by electronic and visual aids operated by the airport and installed in the aircraft. There are three categories of airports as related to the installation of various levels of specialized aids to navigation. At category 1 airports, landings and takeoffs are permitted when the ceiling is 200 ft (60 m) or more and the runway visual range (RVR) is 1800 ft (550 m) or more. The pilot operates the controls at touchdown, although the initial approach may be controlled by instruments. At category 2 airports, landings and takeoffs are authorized down to a ceiling height of about 150 ft (45 m) and runway visual range of about 500 ft (150 m). At category 3 airports, sophisticated aids to navigation, such as Global Positioning System (GPS), provide for safe, fully automated, and electronically coupled landings under "zero, zero" conditions. Airports smaller than the major city airports may have higher ceiling requirements, as may airports with tall buildings or mountains nearby. *See* AIR NAVIGATION; AIRCRAFT INSTRUMENTATION; INSTRUMENT LANDING SYSTEM (ILS).

The accurate forecasting of terminal conditions is critical to flight economy, and to safety where sophisticated landing aids are not available. Improved prediction methods are under continuing investigation and development, and are based on mesoscale and microscale meteorological analyses, electronic computer calculations, radar observations of precipitation areas, and observations of fog trends. *See* MESOMETEOROLOGY; MICROMETEOROLOGY; RADAR METEOROLOGY.

Turbulence and low-altitude wind shear. Atmospheric turbulence is principally represented in vertical currents and their departures from steady, horizontal airflow. When encountered by an aircraft, turbulence produces abrupt excursions in aircraft position, sometimes resulting in discomfort or injury to passengers, and sometimes even structural damage or failure. Major origins of turbulence are (1) mechanical, caused by irregular terrain below the flow of air; (2) thermal, associated with vertical currents produced by heating of air in contact with the Earth's surface; (3) thunderstorms and other convective clouds (**Fig. 1**); (4) mountain wave, a regime of disturbed airflow leeward of mountains or hills, often comprising both smooth and breaking waves formed when stable air is forced to ascend over the mountains; and (5) wind shear, usually variations of horizontal wind in the vertical direction, occurring along air-mass boundaries, temperature inversions (including the tropopause), and in and near the jet stream.

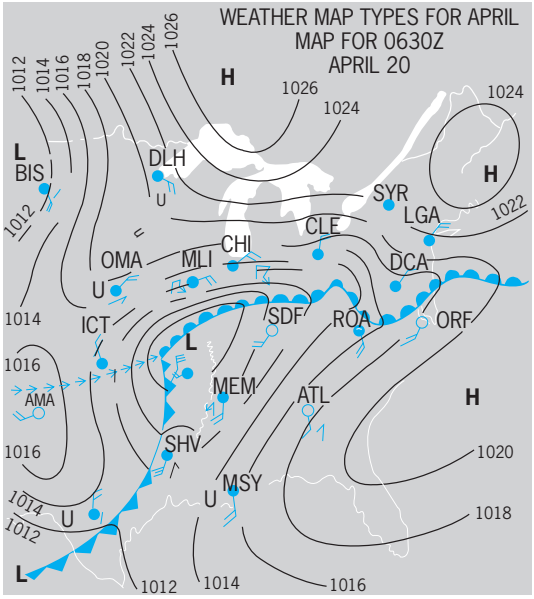

Fig. 1. Severe warm-front thunderstorm situation over the eastern United States. Open or solid circles with three letters represent reporting stations, such as OMA for Omaha and CHI for Chicago. Contour lines express millibar pressure. (*United Air Lines*)

Vertical and horizontal wind gusts in the upper troposphere and lower stratosphere, as encountered by jet transport aircraft that frequent these altitudes, are called clear-air turbulence. Turbulence encounters are the principal cause of nonground impact incidences in commercial aviation. The distinction between clear air and thunderstorm turbulence becomes uncertain as aircraft are routinely operated at altitudes above 30,000 ft (9 km) and encounter thunderstorms growing to more than 60,000 ft (18 km).

While encounters with strong turbulence anywhere in the atmosphere represent substantial inconvenience, encounters with rapid changes in wind speed and direction at low altitude can be catastrophic. Generally, wind shear is most dangerous when encountered below 1000 ft (300 m) above the ground, where it is identified as low-altitude wind shear. Intense convective microbursts, downdrafts usually associated with thunderstorms (**Fig. 2**), have caused many aircraft accidents often resulting in a great loss of life. The downdraft emanating from convective clouds, when nearing the Earth's surface, spreads horizontally as outrushing rain-cooled air. When entering a microburst outflow, an aircraft first meets a headwind that produces increased

performance by way of increased airspeed over the wings. Then within about 5 s, the aircraft encounters a downdraft and then a tailwind with decreased performance. A large proportion of microburst accidents, both after takeoff and on approach to landing, are caused by this performance decrease, which can result in rapid descent. *See* THUNDERSTORM.

Turbulence and low-altitude wind shear can readily be detected by a special type of weather radar, termed Doppler radar. By measuring the phase shift of radiation backscattered by hydrometeors and other targets in the atmosphere, both turbulence and wind shear can be clearly identified. It is anticipated that Doppler radars located at airports, combined with more thorough pilot training regarding the need to avoid microburst wind shear, will provide desired protection from this dangerous aviation weather phenomenon. *See* DOPPLER RADAR; METEOROLOGICAL RADAR.

Hail. Fabric-covered aircraft are particularly susceptible to hail damage. The advent of all-metal aircraft reduced the probability of damage from small (0.2 in. or 5 mm diameter) hail, but larger hail (1–2 in. or 25–50 mm diameter) can be very destructive—shattering the wind screen, dimpling skin surfaces, and cracking radomes. Avoidance of thunderstorms is the best preventive when in flight; on-board radar is a valuable detection device, especially at night and for storms embedded in widespread clouds. *See* HAIL.

Turbulence. Turbulence encounters in and near thunderstorms are increasing with attendant increasing reports of injury to passengers. Part of the increase is because more aircraft fly above 30,000 ft (9 km) and sky space is more crowded. Thunderstorms extend above 60,000 ft (18 km) with their upper portions and anvils spreading out for tens of miles (or kilometers). These storms also compete for space with air traffic. Turbulence at these levels is often located in the regions between the high reflectivity volumes—that is, on the edges of the larger scale vertical motion cores—and in areas where lightning activity is less frequent (charge separation tends to be destroyed by turbulence). Severe turbulence is encountered as much as 20 miles (37 km) from the thundersorm cores.

Turbulence associated with mountains and wind shears is frequently encountered in clear air, with little forewarning to the pilot. The unexpected nature of the clear-air turbulence adds to its danger, especially with respect to passenger safety. The intensity of clear-air turbulence depends on the size and speed of the aircraft. Generally, the accelerations are larger at faster speeds. Clear-air turbulence exists even at the 49,000–65,000 ft (15,000–20,000 m) flight levels of supersonic transport aircraft, and the paths of such aircraft must be planned to avoid areas where strong jet streams are located over tall mountains.

Since clear-air turbulence is a small-scale phenomenon (patches of turbulence frequently measure only a few miles in diameter and a few hundred to a few thousand feet in thickness) the upper-air sounding network is insufficient for locating or predicting

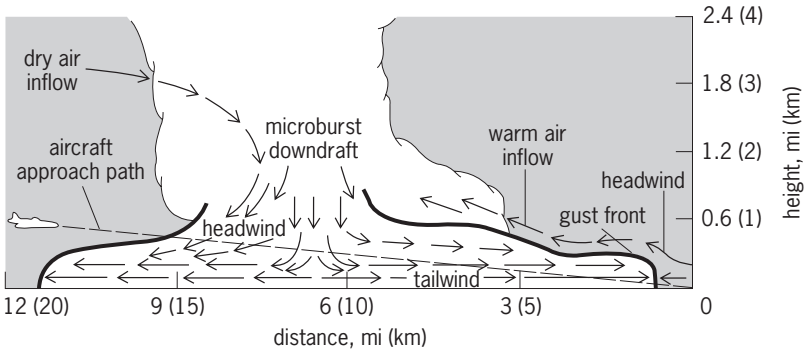

Fig. 2. Schematic of a thunderstorm intersecting the approach path to an airport. Arrows indicate wind-flow pattern. Broken line is 3° glide slope approach. Heavy lines outline the leading edge of the thunderstorm's cold air outflow (gust front). Note rapid change in wind in center of microburst downdraft.

the precise locations of clear-air turbulence zones. Remote sensors, such as the Doppler lidar and the infrared radiometer, are able to detect zones of wind shear and temperature inversions ahead of the aircraft, permitting evasive actions to be taken or passenger warning signs to be activated. *See* CLEAR-AIR TURBULENCE; LIDAR.

Upper winds and temperature. Since an aircraft's speed is given by a propulsive component plus the speed of the air current bearing the aircraft, there are aiding or retarding effects depending on wind direction in relation to the track flown. Wind direction and speed vary only moderately from day to day and from winter to summer in certain parts of the world, but fluctuations of the vector wind at middle and high latitudes in the troposphere and lower stratosphere can exceed 200 knots ($100\ m \cdot s^{-1}$). Careful planning of long-range flights is completed with the aid of prognostic upper-air charts. These charts are prepared through the use of computers to solve equations that model atmospheric motion. The selection of flight tracks and cruising altitudes having relatively favorable winds and temperature results in lower elapsed flight times, in some cases even with considerable added ground distance. The role of the aeronautical meteorologist is to provide accurate forecasts of the wind and temperature field, in space and time, through the operational ranges of each aircraft involved. For civil jet-powered aircraft, the optimum flight plan must always represent a compromise among wind, temperature, and turbulence conditions. *See* UPPER-ATMOSPHERE DYNAMICS; WIND.

Jet stream. This is a meandering, shifting current of relatively swift wind flow which is embedded in the general westerly circulation at upper levels. Sometimes girdling the globe at middle and subtropical latitudes, where the strongest jets are found, this band of strong winds, generally 180–300 mi (300–500 km) in width, has great operational significance for aircraft flying at cruising levels of 4–9 mi (6–15 km). It causes some serious flight-schedule delays, and unexpected encounters may force unscheduled fuel stops.

Average speed in the core of a well-developed jet near the tropopause at middle latitudes in winter is

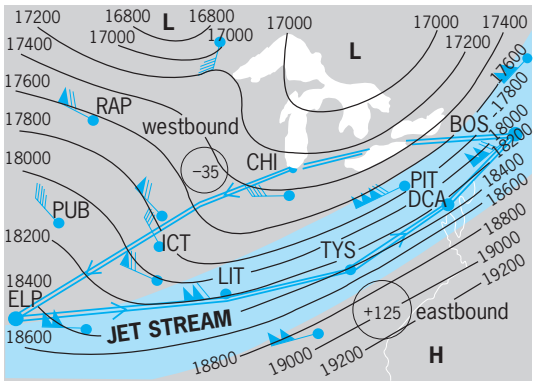

Fig. 3. Use of minimal time track in flight planning over the eastern United States. The solid circles represent weather-reporting places, landing places, or both, such as CHI for Chicago and BOS for Boston. The plus or minus figures shown in the large circles indicate the relation of air speed to ground speed (mi/h) to be expected in the direction of flight. Contour lines express height of the 500-millibar pressure surface in feet. *(United Air Lines)*

close to 100 knots (50 m · s^{-1}), but speeds nearly twice that are fairly common, and extremes may be three times larger. The jet stream challenges the forecaster and the flight planner to utilize tailwinds to the greatest extent possible on downwind flights and to avoid retarding headwinds as much as practicable on upwind flights (**Fig. 3**). As with considerations of wind and temperature, altitude and horizontal coordinates are considered in flight planning for jet-stream conditions. Turbulence in the vicinity of the jet stream is also a forecasting problem.

There is also the "low-level" jet, a relatively narrow (100-nautical-mile or 185-km) stream of air having wind speeds of more than 40 knots (20 m · s^{-1}) through a depth of around 1000 ft (300 m) at an altitude of 1000–3000 ft (300–900 m) above the surface. The northward-flowing jet is usually found from Texas on into the Great Lakes Region. When it occurs, the jet brings moist air rapidly northward and greatly influences the formation of convective storms. *See* JET STREAM.

Tropopause. This boundary between troposphere and stratosphere is generally defined in terms of the zone in the vertical air-mass soundings where the temperature lapse rate becomes less than 3°F/mi (2°C/km). It is sometimes defined as a discontinuity in the wind shear associated with horizontal temperature gradients. Sometimes the tropopause appears as a layer of irregular temperature variations, caused by overlapping "tropical" and "polar" leaves.

The significance of the tropopause to aviation arises from its frequent association with mild forms of clear-air turbulence and change of vertical temperature gradient with altitude. In midlatitudes, the tropopause is commonly located near 6–7 mi (10–12 km), which is about the cruising altitude of most jet aircraft, while in tropical regions it is around 9–10 mi (16–17 km), that is, a little below the cruising altitude of supersonic aircraft. The tropopause marks the usual vertical limit of clouds and storms; however, thunderstorms are known to punch through the surface to heights near 12 mi (20 km) at times, and lenticular clouds showing wave motions strongly developed over mountains have been observed at similar heights. *See* TROPOPAUSE.

Lightning. An electrical discharge to or from an aircraft is experienced as a blinding flash and a muffled explosive sound, usually audible above the roar of the engines. Structural damage to the metal portions of the craft is commonly limited to melting of small spots in the outer skin at the point of entry or exit, and fusing of antenna wires. Nonmetallic surfaces such as radomes and composite tail surfaces may suffer punctures or delaminations. Atmospheric conditions favorable for lightning strikes follow a consistent pattern, characterized by solid clouds or enough clouds for the aircraft to be flying intermittently on instruments; active precipitation of an icy character; and ambient air temperature near or below 32°F (0°C). Saint Elmo's fire, radio static, and choppy air often precede the strike. However, the charge separation processes necessary for the production of strong electrical fields is destroyed by strong turbulence. Thus turbulence and lightning usually do not coexist in the same space.

An aircraft usually triggers the lightning stroke, and evasive action by the pilot usually consists of air-speed reduction, change of course as suggested by radar echoes, or change of altitude. Pilots unable to avoid a typical strike situation should turn up cockpit lights and avert their eyes in order to decrease the risk of being blinded. As metal becomes replaced by composite semiconducting materials in aircraft, there is increased risk of damage from lightning, both to the composite section itself and to inadequately shielded microelectric components. *See* ATMOSPHERIC ELECTRICITY; LIGHTNING; SAINT ELMO'S FIRE.

Icing. Modern aircraft operation finds icing to be a major factor in the safe conduction of a flight both in the takeoff and landing phases as well as the in-flight portion of the trip. Ice and frost formation on an aircraft result in increased drag, decreased lift due to disruption of airflow over the wing and tail surfaces, and increased weight. Icing usually occurs when the air temperature is near or below freezing (32°F or 0°C) and the relative humidity is 80% or more. Clear ice is most likely to form when the air temperature is between 32 and −4°F (0 and −20°C) and the liquid water content of the air is high (large drops or many small drops). As these drops impinge on the skin of an aircraft, the surface temperature of which is 32°F (0°C) or less, the water freezes into a hard, high-density solid. When the liquid water content is small and when snow or ice pellets may also be present, the resulting rime ice formation is composed of drops and encapsulated air, producing an ice that is less dense and opaque in appearance.

The accumulation of ice can occur in a very short time interval. In flight, the use of deicing boots (inflatable rubber sleeves) or heated air of the leading edge of the wing or tail surfaces can reduce the problem. On the ground, ice and snow can be removed either by mechanical means or by application of

deicing solutions. Commercial operations use deicing solutions before takeoff to remove and temporally retard the formation of these increased drag-producing and decreased lift-producing hazards.

Carburetor-equipped aircraft may also experience carburetor icing. The sudden cooling of moist air caused by the vaporization of fuel and the expansion of air as it passes through the carburetor can be as much as 60°F (33°C) in a few seconds. Water thus condensed is deposited as frost or ice inside the carburetor, thereby restricting the airflow. The result is a power decrease and in some cases even engine failure. Carburetor heating judicially used can reduce this problem. Icy or deeply snow-covered airport runways offer operational problems for aviation, but these factors are reduced in importance through the use of modern snow removal equipment. However, accurate forecasts and accurate delineation of freezing conditions are essential for safe aircraft operations. J. T. Lee; J. McCarthy

Bibliography. D. R. MacGorman and W. D. Rust, *The Electoral Nature of Storms*, 1998.

Aeronomy

Aeronomy concentrates on the regions above the tropopause or upper part of the atmosphere. The region of the atmosphere below the tropopause is the site of most of the weather phenomena that so directly affect all life on the planet; this region has primarily been the domain of meteorology.

Aeronomy describes the chemical and physical properties of planetary atmospheres and the changes that result from external and internal forces. The results of this interaction impact all altitudes and global distributions of atoms, molecules, ions, and electrons, both in composition and in density. Dynamical effects are seen in vertical and horizontal atmospheric motion, and energy is transferred through radiation, chemistry, conduction, convection, and wave propagation. Thus aeronomy is a truly interdisciplinary field encompassing chemistry, physics, electrical engineering, and mathematics and dealing with all aspects of theory, simulations, modeling, and laboratory and field experiments. Research may involve the use of specialized ground-based optical and radar facilities; satellite, rocket, and balloon instrumentation; and state-of-the-art computational facilities, locally, regionally, and globally. Since aeronomy is not restricted to earth science, optical and radar remote-sensing techniques are also used to study the atmospheres of other planets in the solar system. *See* METEOROLOGICAL RADAR; REMOTE SENSING.

The atmosphere of the Earth is separated into regions defined by the variation of temperature with height (see **illustration**). Above the surface, temperature minima occur near 10 km (6 mi; the tropopause) and 90 km (54 mi; the mesopause). In between these two minima, the maximum in temperature near 50 km (30 mi) is called the stratopause. The region below the tropopause is the troposphere; the regions between the minimum and maximum values are the stratosphere and mesosphere. Above the mesopause the region is called the thermosphere. The temperature gradually increases above the mesopause throughout the thermosphere to a constant value near 300–500 km (180–300 mi). Above this altitude the region is called the exosphere, which extends out to the limit of the detectable terrestrial atmosphere. The region between the tropopause and the mesopause (roughly 10–100 km or 6–60 mi) is called the middle atmosphere, and the region above (>100 km or 60 mi) is called the upper atmosphere. *See* MESOSPHERE; METEOROLOGY; STRATOSPHERE; TROPOSPHERE.

Middle atmosphere. In the middle atmosphere, that region of the atmosphere between the tropopause and the mesopause (10–100 km or 6–60 mi), the temperature varies from 250 K (−9.7°F) at the tropopause to 300 K (80°F) at the stratopause and back down to 200 K (−100°F) at the mesopause (see illus.). These temperatures are average values, and they vary with season and heat and winds due to the effect of the Sun on the atmosphere. Over this same height interval the atmospheric density varies by over five orders of magnitude. Although there is a constant mean molecular weight over this region, that is, a constant relative abundance of the major atmospheric constituents of molecular oxygen and nitrogen, there are a number of minor constituents that have a profound influence on the biosphere and an increasing influence on the change in the atmosphere below the tropopause associated with the general topic of global change. These constituents, called the greenhouse gases (water vapor, ozone, carbon dioxide, methane, chlorine compounds,

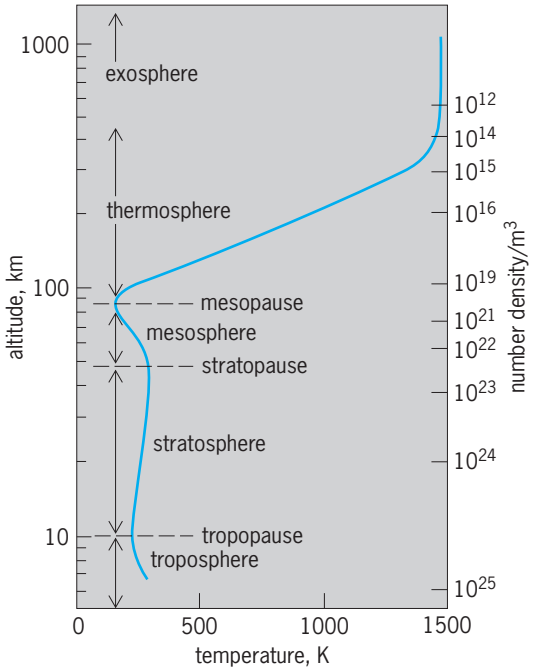

Typical atmospheric temperature variation with altitude and constituent number density (number of atoms and molecules per cubic meter) at high solar activity. °F = (K × 1.8) − 459.67. 1 km = 0.6 mi.

nitrogen oxides, chlorofluorocarbons, and others), are all within the middle atmosphere. To fully understand and model this region of the atmosphere requires consideration of the solar variability, horizontal and vertical wind systems, and the reaction rates and time constants involved with all the chemical and radiative processes that control these constituent distributions. This information needs to be obtained simultaneously at all altitudes and all global locations. The magnitude of this task is huge, and thus aeronomy requires national and international coordinated programs. *See* GREENHOUSE EFFECT.

Solar radiation. Solar radiation is the primary energy source that drives all of the various processes in middle-atmospheric aeronomy. Although the solar spectrum can at times be approximated by a body radiating at a temperature around 6000 K (10,300°F), a detailed study of the spectrum as received at different altitudes in the middle atmosphere shows that absorption by the constituents of the atmosphere modifies the solar radiation at specific wavelengths. Since ozone in the stratosphere strongly absorbs solar radiation below 300 nanometers, the amount of that radiation reaching the ground is very dependent on the total amount of ozone. The harmful relationship between radiation below 300 nm and the Earth's biosystems is the driving concern about the destruction of ozone through the addition of new chemicals that will modify the historical chemical balance of the stratosphere. *See* SOLAR RADIATION; STRATOSPHERIC OZONE.

Dynamics. Understanding the aeronomy of the middle atmosphere requires the study of the physical motion of the atmosphere. The particular composition and chemistry at any given time, location, or altitude depends on how the various constituents are transported from one region to another. Thus, global circulation models for both horizontal and vertical motions are needed to completely specify the chemical state of the atmosphere. In understanding the dynamics of the middle atmosphere, internal gravity waves and acoustic gravity waves play significant roles at different altitudes, depending on whether the particle motion associated with the wave is purely transverse to the direction of propagation or has some longitudinal component. These waves originate primarily in meteorological events such as wind shears, turbulent storms, and weather fronts; and their magnitude can also depend on orographic features on the Earth's surface. One example of coupling between regions in the middle atmosphere is seen in the events known as stratospheric warmings, which are produced by gravity waves generated from the large tropospheric waves that warm the stratosphere and cool the mesosphere. In the polar regions the winter circulation tends to be very stable and circumpolar, and it is usually referred to as the polar vortex. In the Antarctic this region is so stable that photolysis of many of the middle atmospheric species does not occur, thus allowing the buildup of species that would otherwise never develop significant concentrations. Upon reexposure to sunlight during the spring, there is a dramatic photolytic conversion to high concentrations of other species that combine in the presence of polar stratospheric clouds to produce the observed spring ozone depletion (ozone hole) in the lower stratosphere. In the summer polar regions the mesopause temperature becomes very low (<140 K or −208°F) and leads to the formation of very high altitude (84 km or 50 mi) clouds. Historically seen from the ground since 1885 and known as noctilucent clouds, they have now been seen from satellites and are also known as polar mesospheric clouds. *See* ATMOSPHERIC CHEMISTRY; MIDDLE-ATMOSPHERE DYNAMICS.

Upper atmosphere. The upper atmosphere is that region above the middle atmosphere that extends from roughly 100 km (60 mi) to the limit of the detectable atmosphere of the planet. This region is characterized by an increasing temperature until it reaches a constant exospheric temperature. There is a slow transition from the region of constant mean molecular weight associated with the middle atmosphere to that of almost pure atomic hydrogen at high altitudes of the exosphere. This is also the region of transition between transport dominated by collision and diffusion, and transport influenced by plasma convection in the magnetic field. The neutral density varies by over ten orders of magnitude from one end to the other and is dominated by molecular processes in the high-density region and an increasing importance of atomic, electron, and ion processes as the density decreases with altitude.

The ionosphere is formed primarily by solar ultraviolet radiation at wavelengths below that produced by the hydrogen atom emission line (Lyman alpha) at 121.6 nm. This radiation ionizes a fraction, variable with altitude, of the atoms and molecules of the upper atmosphere. Basic studies are concerned with energy sources and sinks, dynamics as a function of altitude, chemistry and particle precipitation leading to optical emissions from the atmospheric constituents, coupling between the low altitudes and the higher altitudes, and the impact of heating by both ions and electrons and their reaction to variable electric fields (related to solar activity) and subsequent variation in the Earth's magnetic field. *See* IONOSPHERE; ULTRAVIOLET RADIATION.

Dynamics. Studies of the movement of the neutral atmosphere are complicated because of the variation in atmospheric density and various other influences. The major structure of the upper-atmospheric wind field is produced by the pressure gradient caused by the difference in solar heating on the day and night sides of the Earth. Comparison of observations and this type of simple model shows that the interaction of the ionosphere with the neutral atmosphere needs to be considered on a global basis, even though most of the major effects are associated with particle precipitation and effects of the configuration of the electric field in the polar regions. Only when the impact of ion drag (the slowing down of moving ions because of collisions with atmosphere gases) is taken into consideration do the models begin to reproduce the observations. Thus topics including Joule

heating, particle precipitation, magnetospheric electric fields, conductivity, plasma convection, and gravity tidal- and planetary-wave activity need to be included in the coupled Ionospheric-thermospheric models. *See* UPPER-ATMOSPHERE DYNAMICS.

Dayglow. While the visible solar light plays an important role in the heating of the middle atmosphere during the daytime, the incident solar far-ultraviolet light at wavelengths below 200 nm ionizes, dissociates, and excites to radiative states the various atmospheric molecules and atoms that make up the upper atmosphere. The light given off by these atmospheric constituents makes up what is called the dayglow spectrum of the atmosphere. *See* AIRGLOW.

Nightglow. As the Sun sets on the atmosphere, dramatic changes due to the loss of the solar radiation occur. The excitation of the atmosphere declines, and chemical reactions between the constituents become more and more dominant as the night progresses. When observed with very sensitive instruments, the night sky appears to glow at various colors of light. Prominent green and red atomic emissions due to oxygen at 555.7 and 630 nm, respectively, and yellow sodium light at 589 nm appear, while molecular bands of the hydroxyl radical and molecular oxygen even further in the red collectively contribute most of the total intensity of the nightglow spectrum.

Aurora. The aurora that appears in the southern and northern polar regions is the optical manifestation of the energy loss of energetic particles precipitating into the atmosphere. The region of highest probability of occurrence is called the auroral oval. At high altitudes, electrons and ions present in the magnetosphere are accelerated along magnetic field lines into the atmosphere at high polar latitudes. The energy of most of the auroral particles (<15 keV) is lost before penetration very far into the atmosphere; hence most of the energy is deposited between 100 and 120 km (60 and 72 mi), although optical emissions can be detected throughout the 80–300-km (48–180-mi) region. In the visible part of the spectrum the most intense feature is the 557.7-nm atomic oxygen line. The bands of molecular nitrogen ions fill the region between 380 and 530 nm, and the molecular oxygen and nitrogen bands occur above 590 nm. However, optical emissions from the various atmospheric constituents occur at discrete wavelengths over the whole region from 30 nm to the far-infrared (15,000 nm). The primary auroral energy resides in the flux of electrons; however, at times some regions of the auroral oval are dominated by energetic proton precipitation, with some evidence of significant amounts of energetic oxygen ions and helium ions at other times. *See* AURORA; MAGNETOSPHERE.

Other planets. Most of the topics that are of concern in middle- and upper-atmospheric aeronomy apply to atmospheres of other planets as well. The remote-sensing techniques applied to the study of the Earth can also be used to study various aspects of the aeronomy of other planets. Mars and Venus have atmospheres dominated by carbon dioxide, traces of oxygen, some molecular nitrogen, and very little water. However, there is very little ozone, so without the heat input from absorption of solar radiation as in the Earth's atmosphere, the middle atmospheres of these planets remain cold. The outer planets have atmospheres mainly of hydrogen and helium, and thus they do not absorb much radiation and remain very cold. However, some ionization does take place, and ionospheres form and vary with altitude much like that on Earth. Energetic particle precipitation on Jupiter and Saturn and their satellites Io and Titan produce aurora similar to those on Earth, except with emissions related to the characteristics of their particular atmospheres and magnetic fields. *See* ATMOSPHERE; JUPITER; MARS; SATURN; VENUS.

Gerald J. Romick

Bibliography. S. J. Bauer and H. Lammer, *Planetary Aeronomy: Atmosphere Environments in Planetary System*, 2004; G. Brasseur and S. Solomon, *Aeronomy of the Middle Atmosphere: Chemistry and Physics of the Stratosphere and Mesosphere*, 3d ed., 2005; R. M. Goody and Y. L. Yung, *Atmospheric Radiation: Theoretical Basis*, 2d ed., 1995; M. C. Kelley, *The Earth's Ionosphere, Plasma Physics and Electrodynamics*, 1989; R. R. Meier, Ultraviolet spectroscopy and remote sensing of the upper atmosphere, *Space Sci. Rev.*, 58:1–185, 1991; M. H. Rees, *Physics and Chemistry of the Upper Atmosphere*, 1989.

Aerosol

A suspension of small particles in a gas. The particles may be solid or liquid or a mixture of both. Aerosols are formed by the conversion of gases to particles, the disintegration of liquids or solids, or the resuspension of powdered material. Aerosol formation from a gas results in much finer particles than disintegration processes (except when condensation takes place directly on existing large particles). Dust, smoke, fume, haze, and mist are common terms for aerosols. Dust usually refers to solid particles produced by disintegration, while smoke and fume particles are generally smaller and formed from the gas phase. Mists are composed of liquid droplets. These special terms are helpful but are difficult to define exactly.

Aerosol particles range in size from molecular clusters on the order of 1 nanometer to 100 micrometers (**Fig. 1**). The stable clusters formed by homogeneous nucleation and the smallest solid particles that compose agglomerates have a significant fraction of their molecules in the surface layer.

Role in science and technology. Aerosols are important in the atmospheric sciences and air pollution; inhalation therapy and industrial hygiene, manufacture of pigments, fillers, and metal powders; and fabrication of optical fibers. Atmospheric aerosols influence climate directly and indirectly. They directly affect radiation transfer on global and regional scales. Indirect effects result from their role as cloud condensation nuclei in changing droplet-size distributions that affect the optical properties of clouds and

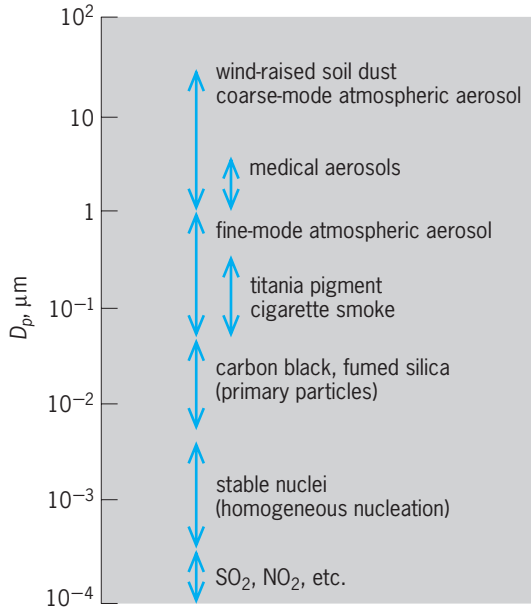

Fig. 1. Examples of aerosol particle-size ranges.

precipitation. There is evidence that the stratospheric aerosol is significant in ozone destruction. *See* METEOROLOGICAL OPTICS; RADIATIVE TRANSFER.

The atmospheric aerosol consists of material emitted directly from sources (primary component) and material formed by gas-to-particle conversion in the atmosphere (secondary component). The secondary component is usually the result of chemical reactions in either the gas or aerosol phases. Contributions to the atmospheric aerosol come from both natural and anthropogenic sources. The effects of the atmospheric aerosol are largely determined by the size and chemical composition of the individual particles and their morphology (shape or fractal character). For many applications, the aerosol can be characterized sufficiently by measuring the particle-size distribution function and the average distribution of chemical components with respect to particle size. The chemical composition of the atmospheric aerosol can be used to resolve its sources, natural or anthropogenic, by a method based on chemical signatures. Particle-to-particle variations in chemical composition and particle structural characteristics can also be measured; they probably affect the biochemical behavior and nucleating properties of aerosols.

Aerosol optical properties depend on particle-size distribution and refractive index, and the wavelength of the light. These are determining factors in atmospheric visibility and the radiation balance.

It is possible to generate aerosols composed of particles that are nearly the same size. These monodisperse (homogeneous) aerosols are normally used for the calibration of aerosol instruments. In industry and nature, aerosols are usually composed of particles of many different sizes; that is, they are polydisperse. The most important physical characteristic of polydisperse aerosols is their particle-size distribution. It is defined, for spherical particles of diameter D_p, as $dN = n(D_p)dD_p$. Here dN is the concentration of particles in a volume element of gas in the particle size range D_p to $D_p + dD_p$, and $n(D_p)$ is the particle-size distribution function. Aerosol size distribution data are often shown by plotting the volume distribution, a weighted function of $n(D_p)$ as a function of particle diameter. Volume distributions for the atmospheric aerosol are frequently bimodal (**Fig. 2**). The lower mode is usually composed of the products of atmospheric gas-to-particle conversion processes, principally ammonium sulfate and nitrate, and of combustion products such as soot and organics. The upper mode includes soil dust and coarse fly-ash particles.

Effects of the atmospheric aerosol on human health have led to the establishment of ambient air-quality standards by the United States and other industrialized nations. Adverse health effects have stimulated many controlled studies of aerosol inhalation by humans and animals. There is much uncertainty concerning the chemical components of the atmospheric aerosol that produce adverse health effects detected in epidemiological studies. *See* AIR POLLUTION.

Aerosols containing pharmaceutical agents have long been used in the treatment of lung diseases such as asthma. Current efforts are directed toward systemic delivery of drugs, such as aerosolized insulin, which are transported across the alveolar walls into the blood. Three main types of medical aerosol inhalers have been developed: (1) nebulizers in which aqueous solutions are aerosolized, (2) pressurized metered-dose inhalers that use chlorofluorocarbon propellant to aerosolize a suspension of the drug, and (3) dry-powder inhalers (DPIs), which depend on inspiration by the user for dispersion. The phase-out of chlorofluorocarbons has led to increased use of dry-powder inhalers; however, a serious technical problem is the need to generate a fine aerosol from a powder that is increasingly cohesive as the particle size decreases.

Aerosol processes are used routinely in the manufacture of fine particles. Aerosol reaction engineering refers to the design of such processes, with the goal of relating product properties to the properties of

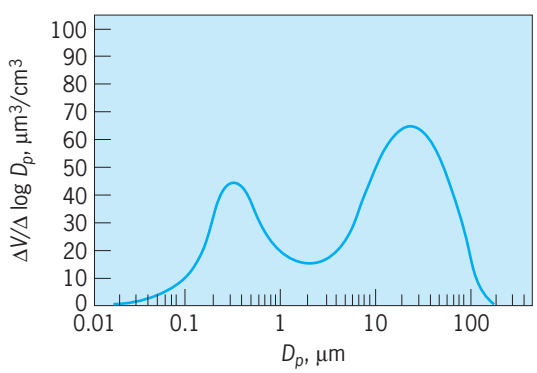

Fig. 2. Bimodal volume distribution of a type frequently found for atmospheric and combustion aerosols. Such distributions usually result from two different generation processes, the smaller mode from molecular condensation and the larger from breakup or redispersion.

the aerosol precursors and the process conditions. The most important large-scale commercial systems are flame reactors for production of pigments and powdered materials such as titania and fumed silica. Optical fibers are fabricated by an aerosol process in which a combustion-generated silica fume is deposited on the inside walls of a quartz tube a few centimeters in diameter, along with suitable dopant aerosols to control refractive index. The original 1-m (3.3-ft) tube is then pulled out to produce a 100-μm fiber 50–100 km (30–60 mi) long. Pyrolysis reactors are used in carbon black manufacture. Micrometer-size iron and nickel powders are produced industrially by the thermal decomposition of their carbonyls. Large pilot-scale aerosol reactors are operated using high-energy electron beams to irradiate flue gases from fossil fuel combustion. The goal is to convert sulfur oxides and nitrogen oxides to ammonium sulfate and nitrate that can be sold as a fertilizer.

Chemical composition. Atmospheric aerosols and aerosols emitted from industrial sources are normally composed of mixtures of chemical compounds. Each chemical species is distributed with respect to particle size in a way that depends on its source and past history; hence, different substances tend to accumulate in different particle-size ranges. This effect has been observed for emissions from pulverized coal combustion and municipal waste incinerators, and it undoubtedly occurs in emissions from other sources. Chemical segregation with respect to size has important implications for the effects of aerosols on public health and the environment, because particle transport and deposition depend strongly on particle size.

Aerosol composition is used routinely to determine the sources of atmospheric aerosols. For example, the chemical mass-balance method permits quantitative estimates of sources based on measured atmospheric aerosol chemical compositions at a given site and a set of known (or assumed) source compositions. There are many variations on this approach, which has been adopted by pollution control agencies.

Transport and deposition. An understanding of particle transport is basic to the design of particle collection equipment and aerosol sampling instruments as well as to the scavenging of particulate matter from the atmosphere. For particles smaller than a few tenths of a micrometer, at normal temperatures and pressures the controlling mechanism of particle transport is Brownian diffusion. Other important driving forces for particle transport are the temperature gradient and, for charged particles, electrical potential gradients. Thermophoresis, which refers to the motion of particles in a temperature gradient, plays a key role in the manufacture of optical fibers. Particles are always driven from high to low temperatures; for particles smaller than the mean free path of the gas, the thermophoretic migration velocity is nearly independent of particle size and chemical composition. Electrical migration is the controlling mechanism for particle deposition in electrostatic precipitators, widely used in industry for gas cleaning. The electrical migration velocity normally passes through a minimum with respect to particle diameters in the range 0.1–1.0 μm. *See* BROWNIAN MOVEMENT; ELECTROSTATIC PRECIPITATOR.

Deposition from flowing gases. Diffusional transport in flowing gases is called convective diffusion. Particle deposition on surfaces by this mechanism can be predicted from the usual mass-transfer correlations, provided that the particle size is small compared with the characteristic length of the collecting elements. The tendency for deposition by diffusion to decrease with increasing particle size is countered by the fact that the particles need to diffuse to only one particle radius from the surface on which they will be deposited. This effect, direct interception, results in an increase in particle deposition with size, usually for particle diameters in the range 0.3–3 μm. *See* DIFFUSION.

Particles unable to follow the motion of an accelerating gas because of their inertia may deposit on surfaces by a process known as inertial deposition or impaction. Inertial deposition is often the controlling removal mechanism for the collection of particles larger than a few micrometers. Inertial deposition rates can be calculated from a force balance on a particle, and depend on the dimensionless group known as the Stokes number, which is related to the particle density, the gas velocity, and the gas viscosity. Inertial deposition often increases sharply over a narrow range of values for the Stokes number.

The efficiency of particle removal from a gas flowing over a set of collecting elements such as those in a filter or packed bed is a function of particle size for a fixed gas velocity (**Fig. 3**). The removal efficiency is high for very small particles because of their intense Brownian motion: it declines until the direct interception mechanism becomes operative, and it continues to rise as impaction and sedimentation become important. The minimum in the efficiency usually occurs in the particle-size range near 0.3 μm—the range normally used to test filter performance.

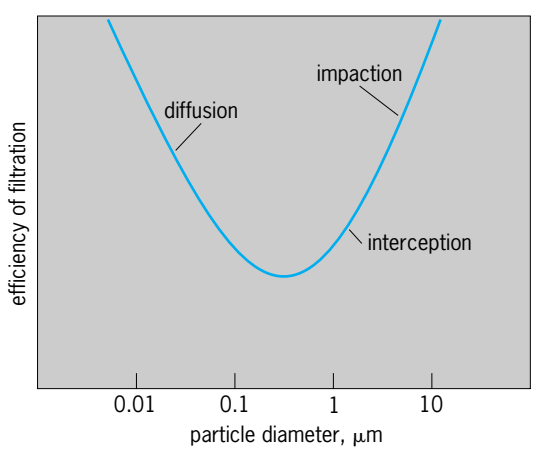

Fig. 3. Variation in the collection efficiency of a fibrous filter with particle size, showing the mechanisms controlling deposition in various particle-size ranges. Leaks or pinholes in the filter material lead to markedly different behavior.

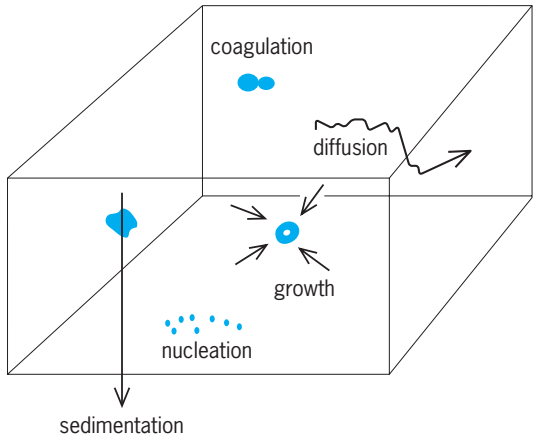

Fig. 4. Processes shaping the particle-size distribution function in a small volume element of gas. Diffusion and sedimentation involve transport across the walls of the element. Coagulation, nucleation, and growth take place within the element.

Aerosol dynamics. The aerosol-size distribution function changes as a gas flows through industrial process equipment or the atmosphere. Aerosol dynamics, which refers to the study of the change in the particle-size distribution function with position and time, gives a distinctive theoretical and experimental character to aerosol science and technology. As an example, consider a small volume element of gas in which the aerosol size distribution is modified by various processes (**Fig. 4**). Particle collisions lead to coagulation. Cooling or chemical reaction may convert molecules in the gas to the aerosol phase by the formation of many new particles smaller than 1–10 nm in diameter or by condensation on existing particles. The formation of new particles, homogeneous nucleation, leads to a particle-size distribution quite different from the one resulting when the same amount of material deposits on existing particles (heterogeneous condensation).

Both gas-to-particle conversion and coagulation take place within the volume element and do not change the total mass of material within the element. Particles may be transported across the element boundary by diffusion, inertia, or an external force field. These processes also change the particle-size distribution function; the change in the distribution function is the sum of the rates for the internal and external processes. A general dynamic equation can be set up for the particle-size distribution function. Also known as a population balance equation, it serves as the starting point for the analysis of aerosol behavior in industrial reactors, pollution sources, and the atmosphere.

Coagulation is a key internal process shaping the aerosol size distribution. The change in the size distribution function resulting from coagulation is given by the Smoluchowski equation,

$$\left(\frac{\partial n}{\partial t}\right)_{\text{coag}} = \frac{1}{2}\int_0^v \beta(\tilde{v}, v-\tilde{v})n(\tilde{v})n(v-\tilde{v})\,d\tilde{v} - \int_0^\infty \beta(v,\tilde{v})n(v)n(\tilde{v})\,d\tilde{v}$$

in which coagulation is related to formation by collision of smaller particles and loss by collision with other particles. Collision kernels have been derived for Brownian coagulation of particles and for many other collision mechanisms, where the particle-size distribution, $n(v,\vec{r},t)$ is a function of position $\vec{r}$, time t, and particle volume v. The collision kernel $\beta(v,\tilde{v})$ is the collision frequency function for particles of volumes v and $\tilde{v}$.

In classical coagulation theory, it is assumed that colliding particles coalesce instantaneously to form larger spherical particles. A striking aspect of Brownian coagulation is that the particle-size distribution tends to approach an asymptotic form independent of the initial size distribution. This asymptotic distribution is referred to as self-preserving, because it does not change its shape when plotted in reduced form as a function of the ratio of particle volume to the average particle volume.

Agglomerates. Aerosol agglomerates are dendritic structures composed of small solid particles that collide but do not coalesce. Industrial aerosols such as pyrogenic silica, titania, and carbon black are often produced in this form.

The individual subunits (primary particles) composing the agglomerate structures are usually 2–100 nm in diameter, much larger than the stable clusters that form in homogeneous nucleation. Indeed, for many metal salt or oxide particles the vapor pressure is so low that the individual molecules serve as stable nuclei. Under such circumstances, the size of the primary particles is determined by the relative rates of the collisions among the particles and their coalescence (or sintering). Coalescence may result from solid-state diffusion or viscous flow (for liquid particles). Both are activated processes and depend sensitively on the temperature of the system.

When coalescence is rapid, collisions produce larger spherical particles. As the temperature falls, coalescence ceases, resulting in solid particles that continue to collide and assemble as agglomerate structures. The size of the primary particles depends on the maximum temperature in the system, decreasing rapidly as the temperature increases. Since agglomerate structures have much larger collision diameters than spheres of the same mass, agglomerates collide much more rapidly than coalescing spheres, leading to the formation of large flocs of low concentration and mass density.

Measurement methods. Aerosol measurement systems are used to monitor atmospheric pollution or industrial gas streams or to test gas-cleaning equipment such as filters and scrubbers. Most instruments used for measuring aerosol properties depend on particle transport or light scattering. Theory provides useful guidelines for instrument design, but it is rarely possible to predict performance from first principles; it is necessary to calibrate the instruments by using monodisperse aerosols. *See* LIGHT-SCATTERING TECHNIQUES.

Quantities often measured include the mass concentration, number concentration, various ranges of the particle-size distribution function, extinction coefficient, average chemical composition, and chemical composition over certain size ranges. There have been rapid developments in the design of instruments for the measurement of size distribution down to the single-particle level.

The aerosol instruments selected for a particular application depend on several factors. Most important is the type of information sought. Other factors include cost, portability, and reliability under the conditions of operation. Stack-gas monitoring poses particularly difficult demands because of extreme conditions of temperature and humidity. In the case of measurement systems designed for routine monitoring, maintenance is an important consideration.

Sampling. In some cases, optical methods can be used to measure aerosol characteristics in the original gas stream, particularly the light extinction, without withdrawing a sample. In most cases, however, it is necessary to sample from a flowing gas through a tube into an instrument to characterize the aerosol. The sampling stream intake should be designed to minimize preferential withdrawal of particles with respect to size. Deposition on the inside walls of the sampling tube and subsequent reentrainment must be minimized or taken into account. Precautions are also necessary to prevent condensation and other gas-to-particle conversion processes. This problem is particularly acute in the sampling of hot, humid process gases.

In general, redispersing deposited aerosols is not a good way to determine the original particle-size distribution. Unless special care is taken in sampling, deposited particulate matter loses its original morphological properties. For microscopic observations, particles can be deposited on electron micrograph grids and optical microscope slides, taking precautions to avoid sampling artifacts.

Physical properties. Special methods have been developed for measuring physical properties of aerosols. Condensation particle counters are used to determine the total particle concentration. The particles entering the counter are mixed with an alcohol vapor, then cooled to produce condensation on the particles. The particle concentration can be determined from light-scattering measurements. Such counters usually respond to particles larger than 5–10 nm. Smaller particles may not be detected because of enhanced vapor pressure (Kelvin effect).

Another integral aerosol property, the mass density, can be determined by weighing the material that accumulates in a filter or, on-line, by measuring beta-ray attenuation. Sampling artifacts include the adsorption and reaction of polar gases in the filter and evaporative losses from the deposited particles.

There is no instrument that is capable of measuring size distributions over the entire size range of practical interest. The single-particle optical counter is widely used for on-line measurement of particles ranging about 0.1–10 μm. Such instruments consist of a light source (which may be a laser), a sensitive volume in which the individual particle and the light interact, and a sensor for the light scattered, usually in some special configuration designed to ensure a unique relationship between the signal and the particle diameter. For particles much smaller than about 0.1 μm, the signal is difficult to distinguish from instrument noise.

In the size range 0.01–0.1 μm, particle-size distributions can be measured with an electrostatic classifier. The particles are charged in an ionizer of which there are several types, by mixing with negative ions generated by a corona discharge. The aerosol then flows through the annular space between a tube and a collecting rod down its center, with an electrical potential between the tube and rod. The charged aerosol enters near the outer circumference of the annular space, and particles with a narrow range in electrical mobility exit through a slit near the end of the rod. The distribution in so-called mobility diameters is determined by measuring the number concentration exiting the slit (usually with a condensation particle counter) as a function of the applied potential. Simultaneous operation of the mobility analyzer and optical particle counter has provided the most complete data available on aerosol size distributions.

Chemical properties. The average aerosol chemical composition is usually determined by collecting particles on a filter and applying a variety of analytical techniques. The chemical speciation and valence states for the collected aerosol may differ significantly from the aerosol in its airborne state because of chemical changes during filtration or storage. Variations in average chemical composition with particle size can be determined by separating the particles into different size fractions. The cascade impactor makes use of inertial deposition to collect particles of progressively smaller aerodynamic diameter on a series of stages. The material on each stage is then analyzed chemically. The method gives the average composition in each size range, but it does not give information on variations in composition from particle to particle; methods based on mass spectrometry have been developed for nearly real-time measurement of the chemical composition of individual particles. *See* CHEMOMETRICS; MASS SPECTROMETRY.

Sheldon K. Friedlander

Bibliography. P. A. Baron and K. Willeke (eds.), *Aerosol Measurement: Principles, Techniques, and Applications*, 2d ed., 2001; Environmental Protection Agency, Office of Research and Development, *Air Quality Criteria for Particulate Matter*, vols. 1 (EPA/600/P-99/002aF) and 2 (EPA/600/P-99/002bF), 2004; S. K. Friedlander, *Smoke, Dust, and Haze: Fundamentals of Aerosol Dynamics*, 2d ed., 2000; W. C. Hinds, *Aerosol Technology: Properties, Behavior, and Measurement of Airborne Particles*, 2d ed., 1999; T. T. Kodas and M. J. Hampden-Smith, *Aerosol Processing of Materials*, 1999; S. E. Pratsinis, The role of aerosols in materials processing, *J. Aerosol Sci. Technol.*, special issue, vol. 19, no. 4, 1993; R. K. Wolff (ed.), Advances in medical aerosol delivery and diagnostics, *Aerosol Sci. Technol.*, vol. 22, 1995.

Aerospace medicine

The special field of medicine that deals with humans in environments encountered beyond the surface of the Earth. It includes both aviation medicine and space medicine and is concerned with humans, their environment, and the vehicles in which they fly. Its objective is to ensure human health, safety, well-being, and effective performance through careful selection and training of flight personnel, protection from the unique flight environment and its physiological and psychological effects, and understanding of the flight vehicle and humans' interaction with it.

The environment encountered in air or space flight is very different from that on Earth. Consequently aerospace medicine must deal with the physics of the atmosphere and space and with the conditions and influences introduced by the flight vehicle. Aerospace medicine is therefore concerned with the physiological effects of changes in barometric pressure, atmospheric constituents, toxic substances, acceleration, weightlessness, noise, vibration, ionizing radiation, and thermal and other environmental stresses, as well as psychological stresses and their effects on behavior and performance.

Environment of flight. The flight conditions of advanced supersonic aircraft introduce extreme variations in temperature, extensive and sudden changes in pressure, and rapid acceleration. The unaided human body cannot compensate for these conditions and must rely upon equipment to survive.

Three gas laws apply directly to physiological problems at altitude. Boyle's law, which states that the volume of a mass of gas is inversely proportional to the pressure exerted on it, providing the temperature remains constant, explains expansion of intestinal gases, and the earaches that can be caused by trapped gases in the middle ear, sinuses, and mastoid cells. Henry's law, that the weight of a given gas dissolved in a liquid varies directly with the partial pressure of the gas, explains how exposure to the low pressure at altitudes may release nitrogen gas normally held in the blood in solution and thus produce bends symptoms. Dalton's law states that in a mixture of gases the part of the total pressure or partial pressure exerted by each gas is proportional to its volume percent. As altitude increases, oxygen continues to constitute 21% of the atmosphere, but the decreased pressure of the upper atmosphere and consequently of the oxygen results in reduced oxygenation of the blood. Physical well-being and the ability to think and to reason are therefore inherently dependent upon the pressure of oxygen breathed. *See* BOYLE'S LAW; DALTON'S LAW; GAS.

The physiological problems of acceleration are an integral part of aerospace medicine because humans are exposed to forces of acceleration almost constantly throughout flight that are different from the 1-*g* environment of Earth. Acceleration is expressed in *g* units, where 1 *g* represents the acceleration of falling bodies due to the force of gravity at 32 ft/s^2 (9.8 m/s^2). Positive *g* means increased acceleration and is accompanied by a physical feeling of heaviness, while negative *g* means reduced acceleration and imparts a sensation of reduced weight. When positive *g* forces are applied to the body, the blood is forced downward away from the head and heart, and blackout and unconsciousness may occur. Under negative *g* conditions, the blood is forced upward so that the blood vessels of the head are engorged. This may result in redout, a condition in which the visual field reddens due to engorged eye blood vessels, and unconsciousness. *See* ACCELERATION.

Effects of weightlessness. All of the environmental hazards incumbent in flight in aircraft are present in space flight, but space also has unique hazards, of which weightlessness is paramount. The equilibrium of many biological systems is disrupted by extended exposure to weightlessness.

Biomedical data collected on humans exposed to long and short periods of weightlessness on space shuttle, Skylab, and Salyut missions have provided information on the time course of physiological acclimation to space and readaptation to Earth. The susceptibility, rate of change, and reversibility of the physiological effects vary with each affected physiological system. Thus neurovestibular effects associated with space motion sickness occur during the first few days in orbit. At the same time there is a shift in body fluids toward the head, and faces become puffy. This is followed by a loss of fluids and electrolytes. With decreased fluids, red blood cell mass slowly decreases for about 60 days into flight. Cardiovascular deconditioning also occurs within the first month. These systems in general appear to acclimate to the weightless environment in 4–6 weeks.

Postflight symptoms include orthostatic intolerance associated with shifts in body fluid and resultant cardiopulmonary neuroreceptor reflex responses. Difficulties in postural equilibrium and occasionally motion sickness accompany neurovestibular readaptation. A return to preflight levels of body function is reached within 1–3 months after return to Earth in a time course similar to the initial adjustment of the various systems to 0 *g*. This is not so with the musculoskeletal system. Many believe there is a gradual progressive loss of calcium and lean body mass regardless of flight duration. While Y. Romanenko spent 430 days in flight, including one flight of 326 days, and 16 other cosmonauts have spent more than 100 days in space, there is as yet no documented residual pathology in astronauts or cosmonauts. In fact, the Russians have reported that, with the use of their current countermeasures, bone loss reaches a plateau or is prevented, while antigravity muscles are reduced by only 10–15% after 237 days of flight. Yet it is still surmised that the normal morphology and function of bone and possibly muscle are not fully regained after continuous exposure of over 6 months to weightlessness.

Extensive data collected on the physiological effects of space flight on humans have led to a number of hypotheses to explain the responses, but the underlying mechanisms are still generally unknown. The various countermeasures used to prevent adverse effects of space flight, coupled with the small

sample size of individuals who have flown in space, the marked variation in mission profiles, and the limited capabilities for scientific observation, have compromised scientific investigations. *See* WEIGHTLESSNESS.

Countermeasure. The United States and Russia use many similar countermeasures including physical, psychological, pharmacological, and nutritional means to prevent or control deleterious physiological responses to space flight. Specifically, these include exercise, especially of the lower extremities, by using unique bicycles and treadmills; a vacuum suit that applies negative pressure to the lower body to stress the cardiovascular system; salt water loading on the last day of flight to increase blood volume and prevent orthostatic intolerance; and nutritional supplements including calcium. Pharmacologic agents are used by the United States to curb space sickness symptoms and by Russia for radiation protection.

Radiation hazards. Radiation hazards have been circumvented in both the United States and Russian space programs through the relatively short duration of flights in carefully selected low-Earth orbits in the absence of solar flare activity. This will not be possible when longer space flights and interplanetary missions will expose spacecraft to increased radiation levels. Space radiation includes heavy ions, electrons, protons, and neutrons. The different types of radiation produce different amounts of biological damage. Charged particles such as heavy ions [abbreviated HZE, for high charge (Z) and energy] and low-energy protons, with high rates of energy loss per unit length of path (high linear energy transfer), are more deleterious than electrons and high-energy protons with low rates of energy loss. *See* LINEAR ENERGY TRANSFER (BIOLOGY).

Radiation is monitored on United States flights with passive dosimeters for precise radiation information and active dosimeters worn by the crew to determine current radiation danger. Dosimetry data from United States flights showed a variation from 11 mrad/day (0.11 milligray/day) for *Gemini 4* to almost 90 mrad/day (0.9 mGy/day) for *Skylab 4*. *See* DOSIMETER.

The potential biological effects of galactic cosmic radiation are categorized as either early, including damage to bone marrow and lymphopoietic, intestinal, and gonadal tissues, or late, including infertility, cancer induction, and heritable effects. Both shielding and radioprotective chemicals afford some protection from space radiation. Total shielding is impossible because of the excess weight it would impart to the spacecraft and the ability of heavy ions to penetrate even heavy shielding. Space medicine consequently has the responsibility to identify appropriate protective procedures, define exposure limits, and develop therapeutic measures. *See* RADIATION BIOLOGY; RADIATION INJURY (BIOLOGY); SPACE BIOLOGY.

Thora Waters Halstead

Bibliography. S. Bonting (ed.), *Advances in Space Biology and Medicine*, vol. 1, 1991; R. L. DeHart (ed.), *Fundamentals of Aerospace Medicine*, 2d ed., 1996; F. J. Del Vecchio, *Physiological Aspects of Flight*, 1985; D. Lorr and V. Garshnek (eds.), *Working in Orbit and Beyond: The Challenges for Space Medicine*, 1989; A. Nicogossian, C. Leach-Huntoon, and S. Pool (eds.), *Space Physiology and Medicine*, 3d ed., 1993.

Aerospike engine

The aerospike engine (**Fig. 1***a*) is an advanced liquid-propellant rocket engine with unique operating characteristics and performance advantages over conventional rocket engines. It combines a contoured axisymmetric plug nozzle (**Fig. 2**), an annular torus-shaped combustion chamber, conventional turbopumps, a turbine exhaust system that injects the turbine drive gases into the base of the plug nozzle, and a simple combustion tap-off engine cycle. The aerospike is one-quarter the length of a conventional rocket engine, yet it delivers comparable performance (efficiency) at high altitude and superior performance at low altitude. The low-altitude performance advantage is primarily due to the fact that the plug nozzle compensates for altitude whereas the nozzle of a conventional rocket engine does not. While the plug nozzle and its benefits are not new to the field of air-breathing propulsion, the aerospike represents the first application of this type of nozzle to the field of rocket propulsion. Typical propellants are liquid hydrogen (fuel) and liquid oxygen (oxidizer). A variation of the aerospike engine is the linear aerospike engine.

Altitude compensation. For a given set of operating conditions in the combustion chamber (pressure, temperature, and mass-flow rate) and ambient pressure, the thrust of a rocket engine is governed by the nozzle exit area, or the nozzle area ratio if the exit area is divided by the nozzle throat area. When the nozzle area ratio is specified such that the pressure at the nozzle exit is equal to ambient pressure, the flow through the nozzle is optimally expanded and the engine is operating at its maximum efficiency. This is the nozzle design point.

When the area ratio is too large, the nozzle exit pressure is lower than ambient pressure and the flow through the nozzle is overexpanded. This situation is accompanied by a decrease in thrust (efficiency). More thrust could be produced if the area ratio of the nozzle were reduced, yielding a higher exit pressure. When the area ratio is too small, the nozzle exit pressure is greater than ambient pressure and the nozzle is underexpanded. Again, thrust is reduced relative to the optimally expanded condition. Under this condition, a nozzle with a higher area ratio would produce more thrust.

Because rocket engines are required to operate over a range of ambient pressure, or altitude, and because engine efficiency is greatest at only a single altitude, selection of the nozzle area ratio (that is, the nozzle design point) for a particular application introduces a compromise. Operation at high altitude where ambient pressure is lower calls for a high area ratio to extract maximum thrust from the

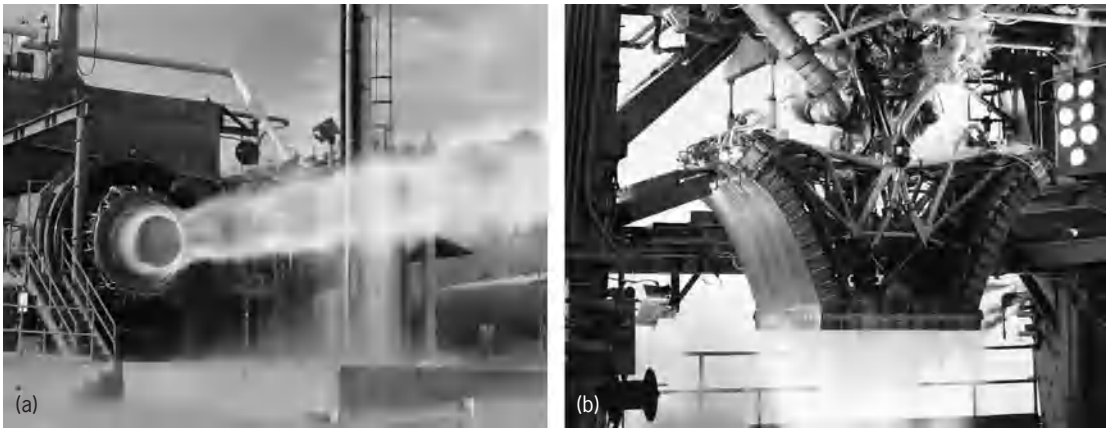

Fig. 1. Static firing tests of (*a*) aerospike engine with 250,000 pounds (1,112,000 newtons) of thrust, and (*b*) linear aerospike engine with 125,000 pounds (556,000 newtons) of thrust. Both engines use hydrogen/oxygen propellants. *(Boeing Company, Rocketdyne Division)*

high-pressure, hot combustion gases, while a lower area ratio is better during low-altitude operation to avoid the reduction in thrust caused by overexpansion. Typically, the selected area ratio falls between these two limits.

Of course, if the geometry of the nozzle were not fixed, and area ratio were allowed to increase with increasing altitude, optimum performance would be obtained from the rocket engine throughout the entire flight trajectory, resulting in dramatic gains in overall vehicle performance. Since the nozzle area ratio is dependent on the nozzle exit area, changes in area ratio could be effected by simply varying nozzle exit area. This effect is referred to as altitude compensation.

Comparison with conventional engines. In a very general sense, the aerospike engine is similar to a conventional liquid-propellant rocket engine. In both cases, turbopumps are used to pressurize liquid propellants. These high-pressure propellants are then injected into a combustion chamber, where they rapidly mix and chemically react to form high-temperature, high-pressure combustion gases. Finally, these hot gases are accelerated to high velocities by expanding them through a nozzle to produce thrust.

In more specific terms, however, several key differences distinguish the aerospike engine from the conventional rocket engine. The most striking difference is the nozzle. The conventional rocket engine uses a converging-diverging nozzle, sometimes referred to as a bell nozzle owing to its shape. The hot combustion gases expand inside the nozzle in the general direction of the nozzle axis, completely contained by the nozzle wall (**Fig. 3***a*). Because the wall of a conventional nozzle forms a physical boundary between the expanding combustion gases and the surrounding atmosphere, the expansion process in the nozzle is isolated from the effect of ambient pressure. Consequently, the combustion gases will not necessarily be optimally expanded, but will be under- or overexpanded depending upon the ambient pressure. *See* NOZZLE.

The aerospike engine features a plug nozzle that is characterized by the combustion gases expanding on the outside of the nozzle (Fig. 3*b*). Initially, the gas flow is directed radially inward, toward the axis or center of the nozzle. The expansion process occurs about the engine cowling, controlled by the plug wall on the inside and the effect of ambient pressure on the outer jet boundary. Because the amount of expansion is controlled by ambient pressure, the aerospike engine has the ability to altitude-compensate in the absence of variable nozzle geometry. Variations in area ratio are achieved automatically by the influence of ambient pressure acting on the jet boundary, causing the effective flow area at the nozzle exit to change with altitude. Thus, ambient pressure effectively provides passive control of the expansion process in the aerospike engine, preventing both under- and overexpansion and the associated loss in performance.

The plug nozzle of the aerospike engine is obtained by truncating a full-length spike nozzle. Since

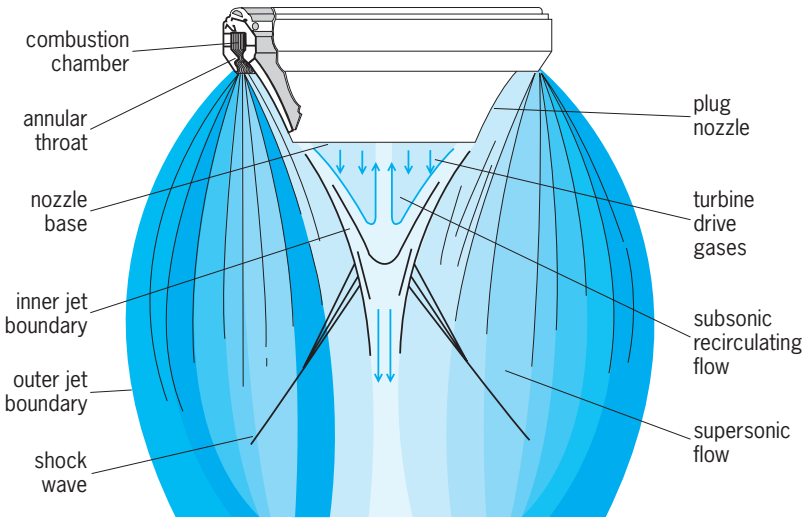

Fig. 2. Aerospike engine with plug nozzle, torus-shaped combustion chamber, and combustion gases expanding outside the nozzle. *(Boeing Company, Rocketdyne Division)*

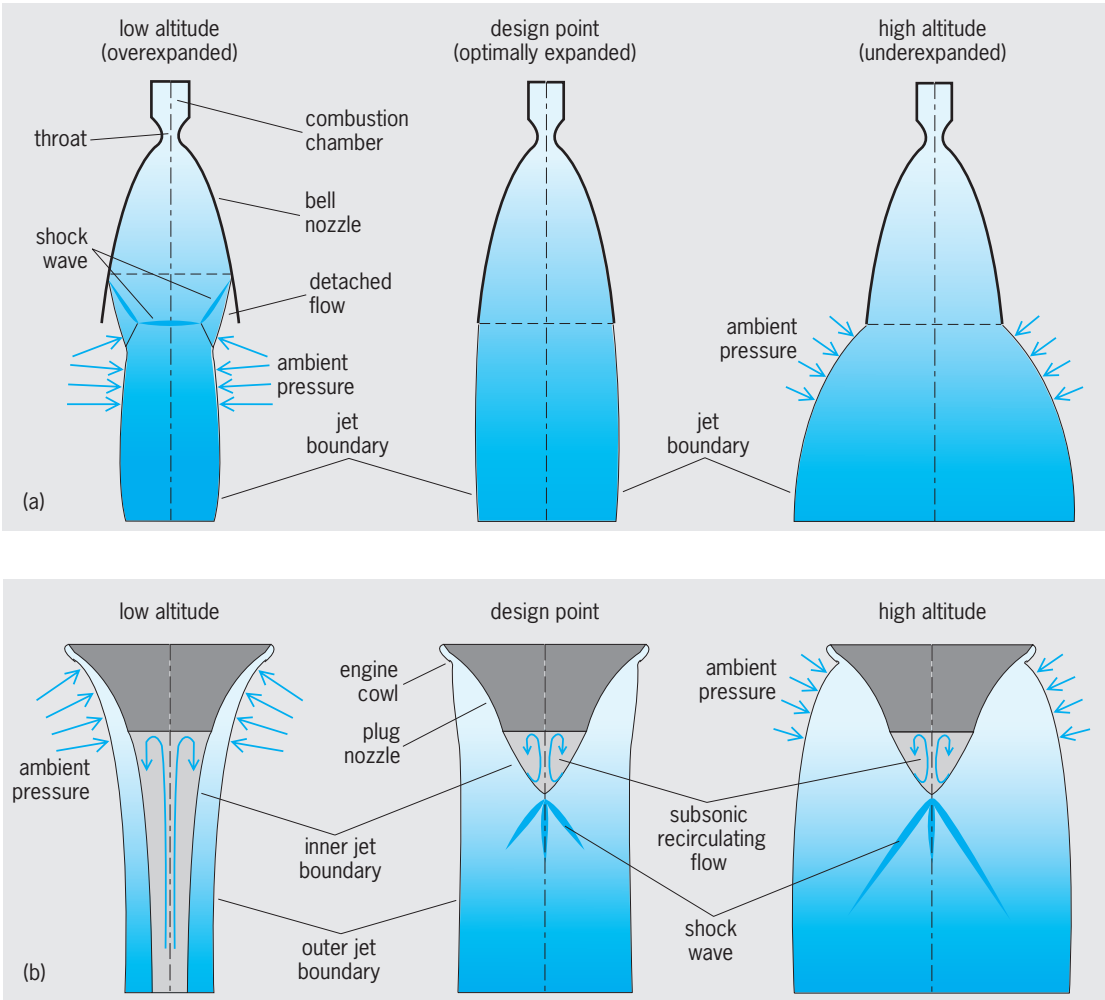

Fig. 3. Nozzle flow comparison. (*a*) Conventional rocket engine (bell nozzle). (*b*) Aerospike engine (plug nozzle). ***(After R. A. O'Leary and J. E. Beck, Nozzle design, Threshold no. 8, pp. 34–43. Rockwell International Corporation, Rocketdyne Division, Spring 1992)***

the amount of thrust produced by the spike nozzle drops off very rapidly toward the aft end, a significant portion of the spike nozzle can be eliminated without undue loss in thrust. By injecting the turbine drive gases into the resulting base (Fig. 2), an aerodynamic spike (thus the name aerospike) is formed.

In the combustion tap-off cycle, the turbine that powers the pumps is driven by a small amount of hot gases that are tapped from the combustion chamber. This procedure eliminates the need for an additional gas generator. Once expanded through the turbine, the gases are exhausted into the blunt base of the plug nozzle, increasing the base pressure, thrust, and overall engine efficiency. This arrangement contrasts with that of a conventional rocket engine, where the turbine drive gases are usually ejected through a small secondary nozzle to produce additional thrust.

To match the unique shape of the plug nozzle, the shape of the combustion chamber and the arrangement of the injectors into the aerospike engine are very different from those of the conventional rocket engine. In the conventional engine the combustion chamber is typically cylindrical and the injectors are arranged in the circular forward end of the chamber. In the aerospike engine the combustion chamber is torus shaped (Fig. 2) and the injectors are arranged in an annular pattern.

Advantages. The aerospike engine offers performance, operational, and configurational advantages over the conventional rocket engine equipped with a bell nozzle. First, altitude compensation yields higher performance at low altitude while producing comparable high performance at high altitude. Second, for the same thrust the aerospike is considerably shorter, by as much as 75%. This allows for a shorter-length, lighter-weight vehicle or, for the same-length vehicle, greater volume for higher-capacity propellant tanks or payload. For reusable winged or lifting-body vehicles, the short length moves the center of gravity forward, improving flight characteristics. Third, the altitude-compensating feature of the plug nozzle permits the safe operation of a nozzle with high area ratio at low altitude (sea level) without undesirable flow detachment in the nozzle (Fig. 3*a*) and the accompanying asymmetric lateral forces, or side loads, on the nozzle. Fourth, the aerospike integrates well into the base of a conventional cylindrical launch vehicle with high aspect ratio, reducing

vehicle base drag and eliminating the need for a base heat shield, since the engine fills the entire base. Finally, the thrust load of the aerospike is distributed at the maximum diameter of the engine instead of concentrated at a single point (the gimbal), eliminating the need for a heavy thrust structure to distribute the load through the vehicle, and resulting in a lighter vehicle.

Linear aerospike. A variation is the linear aerospike engine (Fig. 1*b*). This rocket engine concept offers the same performance advantages as the annular aerospike while offering some unique configurational advantages owing to its linear shape. The combustion chamber is made up of a series of modular chamber segments, and the gas generator engine cycle is used in place of the combustion tap-off cycle.

Advanced launch vehicles . During the 1990s, interest has been renewed in single-stage-to-orbit reusable launch vehicles. Numerous studies have shown that reduced launch costs will be best achieved through the development of a fully reusable single-stage-to-orbit vehicle.

Unlike multistage launch vehicles that depend upon one rocket propulsion system for boost and others for high-altitude operation, future single-stage-to-orbit vehicles will be dependent on a single rocket propulsion system from boost to orbit insertion. While each rocket engine of a multistage vehicle can be individually tailored (for example, the nozzle area ratio) to meet the requirements of its portion of the trajectory, rocket engines for single-stage-to-orbit vehicles must provide high performance over the entire flight trajectory. Thus, advanced rocket propulsion technologies that further increase the performance of liquid-propellant rocket engines will be required. The aerospike engine is one of these advanced propulsion concepts. *See* ROCKET PROPULSION; SPACECRAFT PROPULSION. James E. Beck

Bibliography. R. I. Baumgartner and J. D. Elvin, *Lifting Body: An Innovative RLV Concept*, AIAA Pap. 95-3531, 1995; T. K. Mattingly, A simple ride into space, *Sci. Amer.*, 277(4):120–125, October 1997; I. Parker and F. Colucci, Inside-out engine, *Space*, 8(4):38–40, August–September 1992; G. P. Sutton, *Rocket Propulsion Elements*, 6th ed., 1986.

Aerothermodynamics

Flow of gases in which heat exchanges produce a significant effect on the flow. Traditionally, aerodynamics treats the flow of gases, usually air, in which the thermodynamic state is not far different from standard atmospheric conditions at sea level. In such a case the pressure, temperature, and density are related by the simple equation of state for a perfect gas; and the rest of the gas's properties, such as specific heat, viscosity, and thermal conductivity, are assumed constant. Because fluid properties of a gas depend upon its temperature and composition, analysis of flow systems in which temperatures are high or in which the composition of the gas varies (as it does at high velocities) requires simultaneous examination of thermal and dynamic phenomena. For instance, at hypersonic flight speed the characteristic temperature in the shock layer of a blunted body or in the boundary layer of a slender body is proportional to the square of the Mach number. These are aerothermodynamic phenomena.

Two problems of particular importance require aerothermodynamic considerations: combustion and high-speed flight. Chemical reactions sustained by combustion flow systems produce high temperatures and variable gas composition. Because of oxidation (combustion) and in some cases dissociation and ionization processes, these systems are sometimes described as aerothermochemical. In high-speed flight the kinetic energy used by a vehicle to overcome drag forces is converted into compression work on the surrounding gas and thereby raises the gas temperature. Temperature of the gas may become high enough to cause dissociation (at Mach number ≥ 7) and ionization (at Mach number ≥ 12); thus the gas becomes chemically active and electrically conducting. *See* COMBUSTION; HYPERSONIC FLIGHT; JET PROPULSION; MACH NUMBER; ROCKET PROPULSION.

In order to describe aerothermodynamic problems more fully, three specific phenomena are discussed in the following sections: internal flow, external flow, and aerodynamic heating.

Internal flow. In internal flow the gas is confined by the walls of a duct. Aerothermodynamic effects in this case are caused either by gases, such as air at high temperatures, or by combustion. The internal flow of gases at high temperature is a phenomenon largely confined to laboratory equipment, such as the hypersonic wind tunnel, the shock tube, the hot-shot tunnel, or the plasma jet used to simulate flight conditions for testing models. The rocket, ramjet, and the turbojet engine also involve combustion processes in which aerothermodynamic effects are important.

The shock tube consists basically of a long pipe, divided by a diaphragm into two compartments and closed at both ends. Gases at different pressures are placed in the two sections of pipe, the diaphragm is ruptured, and a shock wave propagates into the quiescent low-pressure gas. The gas behind the shock wave is accelerated, compressed, and heated to a high temperature. This region of high-velocity and high-temperature air can be used to simulate high-speed flight conditions. The radiation from the hot gas, chemical kinetics, heat transfer to simple shapes, forces on simple shapes, and boundary-layer transition can all be studied by this experimental device. *See* SHOCK TUBE; SHOCK WAVE.

For the propulsion units normally used in various vehicles that rely on ambient air as a source of oxygen and as a working fluid, the free-stream tube of air entering an inlet must be decelerated to a low subsonic velocity before entering the combustion chamber. In practical applications the deceleration of a supersonic stream (flow traveling faster than the local speed of sound) is not possible without the formation of discontinuities, or shock waves. The formation of shock waves in a stream always results in

an increase in entropy; that is, the available energy in the stream is diminished as the flow proceeds across the shock wave. *See* SUPERSONIC DIFFUSER.

Deceleration of the supersonic airstream to subsonic speed before it enters the combustion chamber is accomplished most efficiently by an oblique shock diffuser. The oblique shock diffuser, which consists of a cone or a wedge, produces a shock wave system of one or more oblique shock waves followed by one normal shock wave. Because the Mach number of the air following a normal shock wave is always subsonic, the downstream air is at the required low velocity. As the upstream Mach number increases, the subsonic Mach number following the normal shock decreases.

In combustion engines aerothermodynamic effects are important in several respects. First there is the diffusion and mixing process of the fuel and oxidizer. When fuel is introduced in liquid form, as in liquid-propellant rocket engines, the combustors of ramjet and turbojet or fanjet engines, and afterburners, the process involves the breakup of the fuel spray, evaporation of the liquid drops, and mixing of fuel and oxidizer. *See* AFTERBURNER; ROCKET ENGINE; SUPERCHARGER; TURBOFAN; TURBOJET.

In the combustion process itself chemical reactions are complex and turbulence is high. One important problem is to stabilize the flame. To produce useful work, the high-temperature products of combustion are expanded, either in a turbine or in a nozzle. *See* GAS TURBINE.

As the flight speed increases into the high supersonic and hypersonic regimes, the diffusion process in the inlet becomes very inefficient. The flow that is slowed down by the normal shock causes an increase in both static temperature and static pressure. As a result of the high static temperature, air begins to dissociate appreciably, causing a substantial energy loss. In addition, both high static temperatures and static pressures may, in turn, require a heavier structure. These problems can be considerably minimized if the combustion process is allowed to occur in the supersonic stream. However, the energy losses due to heat addition processes in a supersonic stream can be appreciable. *See* SUPERSONIC FLIGHT.

In a turbine the design problems are twofold: the analysis of flow over airfoils in cascade, and heat transfer to the turbine blades. Flow is usually treated by ideal-gas techniques. Because the turbine blades are constantly exposed to hot combustion gases, it is necessary to develop materials able to withstand higher temperatures and to devise ways of cooling the turbine blades. *See* TURBINE.

In the nozzle, high-temperature and high-pressure products of combustion are expanded to a high velocity as the temperature and pressure decrease. The flow is complicated by the complex chemical composition of the gas. The efficiency of the nozzle is affected by its contour, the amount of heat lost through its wall, the degree of incomplete combustion, and the presence of nonequilibrium thermal and chemical states, flow separation, shock waves, and turbulence. Heat transfer to the nozzle walls, especially at the throat, is important because of its effect on the structural integrity of the walls. *See* NOZZLE.

External flow. Aerothermodynamic effects in the external flow about bodies occur during high supersonic or hypersonic flight. Any real body has a finite radius of curvature at the leading edge or nose. This radius results in a bow shock, which is detached and nearly normal to the flow at the stagnation region. The air between the normal shock and the stagnation region of the body is compressed to a high temperature. This compressed region is the primary source of high-temperature gas in the external flow about a body. For the purpose of discussion and analysis, the external flow field about a body (between the shock wave and the body surface) is divided into three regions, namely, the inviscid external flow, the viscous boundary layer, and the wake or jet in the base region.

Both the compression and the temperature rise that a gas experiences as it passes through the bow shock increase with Mach number and shock wave angle. The high-temperature gas transfers heat from the boundary layer to the body. A large amount of heat is transferred in the stagnation region of a nose (see **illustration**) and on the leading edge of a wing, not only because the gas pressure is maximum, but also because the boundary layer has a minimum thickness in such regions. *See* NOSE CONE.

In designing a vehicle, it is necessary to apply information obtained from the solution of the equations of the external flow field to such problems as structural design (in the form of applied forces and heat input), guidance and control (in the form of applied forces), and communication and detection (in the form of electromagnetic disturbances). The applied forces are usually determined by solving the aerodynamic equations for an ideal inviscid gas and correcting for real-gas effects. Heat transfer, on the other hand, is estimated by a consideration of the thin viscous boundary layer.

At high temperatures the gas in the shock layer emits electromagnetic radiation. While this phenomenon increases heat transfer, it is more important

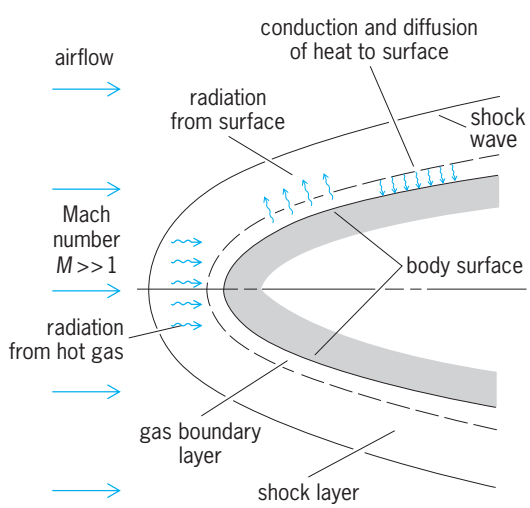

Heat transfer in stagnation region of blunted nose cap.

as a source of electromagnetic noise, which can also adversely affect communication with the vehicle. At still higher temperatures the gas ionizes, producing electromagnetic forces in addition to the radiation. *See* MAGNETOHYDRODYNAMICS.

Aerodynamic heating. Aerodynamic heating is a severe problem at high flight speeds and at all altitudes at which air forms a continuum, that is, up to approximately the altitude at which the mean free path is of the same order of magnitude as the diameter of the vehicle's nose. Such an altitude might be about 350,000 ft (107 km) for a nose having a diameter of about 1 ft (0.3 m).

To investigate aerodynamic heating on a missile, aircraft, spacecraft, or other flying object, one needs detailed information about flight profiles, such as the altitude, velocity, angle of attack, and bank-angle history of the vehicle as a function of time. For most supersonic aircraft, the thermal design can be based on the fixed-equilibrium skin temperature (the temperature at which heat input equals heat output). For a missile, equilibrium is not reached, and the transient nature of missile flight trajectories complicates the temperature analysis. For an atmospheric-entry spacecraft, skin temperature will in some cases reach equilibrium because of its longer flight time, but the equilibrium skin temperature may change with time because of changes in altitude, speed, angle of attack, and bank angle. *See* ATMOSPHERIC ENTRY.

The atmospheric-entry characteristics of spacecraft and of ballistic missiles differ in both the total heat load experienced and in maximum heating rates. Total heat load is greater for the spacecraft, whereas the maximum heating rate is greater for the ballistic missile because of its steeper entry path. For spacecraft with a high lift-to-drag ratio ($L/D \approx 3$), the severity of heating can be reduced by trajectory selection, configuration design, and radiation cooling. For spacecraft with medium and low lift-to-drag ratios ($L/D \approx 2$ or less) and for the ballistic entry vehicle, ablation cooling techniques have been used extensively to protect interior structure. Changing the nose configuration alone cannot completely solve the structural heating problem for a ballistic missile. *See* SPACECRAFT STRUCTURE.

In general, heating analysis is conducted by dividing the vehicle into components; appropriate heat-transfer formulas are then applied to each component to estimate its heating rate. Analysis can be further complicated by boundary-layer transition, boundary-layer interaction, separation flow, shock impingement, mass transfer, nonequilibrium phenomena, gap effects, and rough or wavy surfaces.

Accurate estimates of aerodynamic heating also require knowledge of fluid properties of the gas, such as composition, thermal and transport properties, reaction rates, relaxation times, and radiation characteristics within the fluid around the hypersonic vehicle. Under standard conditions air is composed chiefly of diatomic molecules. Motions of translation and rotation are excited at room temperature. As the temperature of air increases, vibration begins, increasing in amplitude as temperature increases, until eventually the intramolecular bond is broken. The molecule is then said to be dissociated. Still higher energy levels are excited as the temperature increases further; eventually, electrons are separated from their parent atoms and the gas becomes ionized.

Shih-Yuan Chen

Bibliography. J. D. Anderson, Jr., *Hypersonic and High Temperature Gas Dynamics*, 1989, reprint, 2000; J. D. Anderson, Jr., *Modern Compressible Flow*, 3d ed., 2003; J. J. Bertin, *Hypersonic Aerothermodynamics*, 1994; G. Oates, *Aerothermodynamics of Gas Turbine and Rocket Propulsion*, 3d ed., 1997; C. Park, *Nonequilibrium Hypersonic Aerothermodynamics*, 1990.

Affective disorders

A group of psychiatric conditions, also known as mood disorders, that are characterized by disturbances of affect, emotion, thinking, and behavior. Depression is the most common of these disorders; about 10–20% of those affected by depression also experience manic episodes (hence this condition is known as manic-depression or bipolar affective disorder). The affective disorders are not distinct diseases but psychiatric syndromes that likely have multiple or complex etiologies.

Clinical Syndromes

Affective disorders are distinguished from states of normal sadness or elation by several features. Their clinical syndromes consist of characteristic changes in brain and psychological function that affect sleep, appetite, speed of movement, energy, concentration, self-esteem, motivation, and hedonic (pleasure) capacity. During an episode, these features are persistent, occurring day after day for at least several weeks. Affective syndromes also differ from normal mood states by the degree of associated impairment of functioning. The most severe episodes include hallucinations or delusions (unshakable false beliefs of obvious personal significance, typically reflecting themes that are congruent with the mood disturbance).

A natural parallel to the clinical syndrome of depression is the state of grief. Grief is associated with many of the symptoms of depression, although the changes of mood and behavior that are associated with grief are less persistent and more circumscribed. Conversely, new romantic love or an infatuation is a natural parallel of a hypomanic episode (that is, a mild episode of mania), although again the normal "affliction" is less persistent and pervasive and should not cause functional impairment. Unresolved grief can progress into clinical depression, however, and there is a pathological form of delusional romantic love that is called erotomania.

Aside from the death of a loved one, other types of losses, including divorce, the breakup of a romantic relationship, or setbacks in vocational pursuits, commonly precede the onset of an initial depressive episode. A transitional diagnosis of an adjustment disorder is used when the duration or severity of

depressive symptoms has not crossed the threshold necessary for diagnosis of a depressive disorder. A majority of adjustment disorders, like normal grief, remit as the individual adapts to the stressor.

Major depressive disorders. The most common form of affective disorder is a major depressive episode. The episode is defined by a pervasively depressed or low mood (experienced most of the day over a period of 2 weeks or longer) and at least four associated symptoms affecting sleep, appetite, hedonic capacity, interest, and behavior. Major depressive episodes have several clinical forms.

Melancholia is a severe episode characterized by anhedonia (inability to experience pleasure), markedly diminished appetite (anorexia; note the anorexia of depression is *not* related to anorexia nervosa), with weight loss, early morning awakening, observable motor disturbances (extreme slowing, or retardation, or pacing and stereotypic agitated behaviors), and diurnal mood variation (mood is worse in the morning). Some episodes of melancholia are associated with delusions, typically of guilt, sin, illness, punishment, or nihilism.

Common among young patients, especially women, is a milder syndrome historically referred to as atypical depression. Atypical depression is characterized by a less pervasive mood disturbance, with mood reactivity preserved (that is, the patient's spirits go up or down in response to day-to-day events) and by so-called reverse symptoms: oversleeping, overeating, or gaining weight. Significant anxiety symptoms, including phobias and panic attacks, also are common in atypical depression.

Both atypical and other milder forms of major depressive disorder were formerly referred to as reactive or neurotic because psychosocial factors often appeared to play a prominent role in their onset. Melancholia, by contrast, is more autonomous and less clearly linked to life stress. In fact, melancholia was formerly called endogenous depression because the cause was presumed to be within the brain. Current thinking on the roles of life events and brain-related processes is less simplistic: it is likely that both sets of factors are relevant to all forms of affective disorder.

Dysthymia. A more chronic, insidious form of depression known as dysthymia "smolders" at a subsyndromal level (that is, there are three or four daily symptoms rather than five or more) for at least 2 years. Dysthymia often begins early in life and, historically, has been intertwined with atypical and neurotic characteristics of depression. There may be a depressive personality or temperament that underlies childhood-onset dysthymia. Most people with untreated dysthymia eventually develop major depressive episodes and, despite its milder symptoms, dysthymia causes considerable impairment.

Mania and bipolar affective disorder. About 10–20% of people who experience major depressive episodes will also suffer from abnormal periods of elation, which are referred to as manias (or in less severe form, hypomanias). A manic episode is heralded by euphoric or irritable mood and at least four of the following symptoms: increased energy, activity, self-esteem, or speed of thought; decreased sleep; poor judgment; and increased risk-taking. The energy, drive, and optimism of a hypomanic individual can seem captivating to others. About one-half of manic episodes are psychotic. The delusions of mania typically reflect grandiose or paranoid themes.

Almost everyone who experiences manic episodes also will suffer from recurrent depressive episodes. Since there is a 50-50 chance that an episode of depression may precede the onset of the first mania, changes in diagnosis are common.

The term bipolar affective disorder has largely replaced the former term manic-depression, although both names convey the cyclical nature of this illness. The classical presentation (which includes full-blown manic episodes) is known as type I disorder. The diagnosis of bipolar type II disorder is used when there are recurrent depressive episodes and at least one hypomania. A third diagnosis, cyclothymia, is used when neither hypomanias nor depressions have reached syndromal levels. For some, bipolar II or cyclothymic disorders subsequently evolve into the more severe type I form of illness.

Several important variations of bipolar disorder are increasingly recognized.

Mixed episode. A mixed episode is diagnosed when the symptoms of mania and depression coexist. Mixed episodes are diagnostically and therapeutically challenging, and their existence suggests that mania and depression are not caused by polar-opposite abnormalities of brain functioning. Mixed episodes are often associated with sleep deprivation or drug and alcohol use.

Rapid cycling. The term rapid cycling is used when there have been four or more episodes within a time frame of 1 year. Cases of ultrarapid cycling (mood cycles occur within hours or days) also have been reported. Rapid cycling is more common among women, and may be provoked by antidepressant therapies. Thyroid disease and the bipolar II subtype also increase the likelihood of rapid cycling.

Seasonal affective disorder. A number of affective disorders follow a seasonal pattern. A pattern of recurrent fall/winter depressions (also known as seasonal affective disorder) has generated considerable interest because it may be treated with bright white (full-spectrum) light, which artificially lengthens the photoperiod.

Disorders associated with general medical conditions. Both depression and mania can be caused by general medical illnesses and medications that affect brain function (such as antihypertensives, hormonal therapies, steroids, and stimulant drugs). The diagnosis "mood disorder associated with a general medical condition" is applied to these conditions. Recognition that the mood disorder is linked to another general medical illness can have important treatment implications.

Pathophysiology

Affective disorders have diverse biopsychosocial underpinnings that result, at least in part, in extreme or distorted responses of several neurobehavioral systems. The neurobehavioral systems of greatest

relevance regulate a person's drives and pursuits, responses to acute stress, and capacity to dampen or quiet pain or distress. The system that regulates the daily rhythms of sleeping and waking and associated patterns of hormonal secretion likewise helps to ensure well-being. The coordinated function of these response and homeostatic systems conveys important and, from an evolutionary standpoint, ancient advantages for survival. Despite obvious overlap with how stress can adversely affect other primates and lower mammals, the affective disorders are distinctly human conditions. They may be thought of as one disadvantageous consequence of the complexity of affective, cognitive, and social functioning that distinguishes humans from their phylogenetic relatives.

Although there is considerable evidence that affective disorders are heritable, vulnerability is unlikely to be caused by a single gene. Bipolar disorder is clearly more commonly inherited than major depressive disorders, and among the major depressive disorders the more severe, psychotic, and early-onset forms are more heritable than the remainder. Such risks are relative, however, and even identical twins have no more than a 50–60% risk of concordance of bipolar disorder. It is likely that some combination of genes conveys greater risk and, like an amplifier, distorts the neural signals evoked by stress and distress. *See* BEHAVIOR GENETICS; HUMAN GENETICS.

Research permits several firm conclusions about brain neurochemistry in stress and depression. Acute stress affects the release of three vital brain monoamines—serotonin, norepinephrine, and dopamine—as well as glucocorticoids such as cortisol. Cortisol release is "driven" by the neuropeptide corticotrophin releasing hormone, which is released from neurons in both the hypothalamus and limbic cortex. Sustained unresolvable stress eventually depletes the neurotransmitters (cortisol levels remain high), inducing a behavioral state that has been called "learned helplessness." In severe recurrent episodes of depression, especially those with psychotic features, cortisol elevations approach those seen in Cushing's disease. Sustained elevations of cortisol have been shown to inhibit the growth of new neurons (neurogenesis) and ultimately may cause atrophy of some regions of the brain. *See* ADRENAL CORTEX HORMONE; BRAIN; STRESS (PSYCHOLOGY).

As described earlier, both depression and mania are characterized by alterations in the sleep-wake cycle, including changes in nocturnal body temperature, growth hormone secretion, and the electroencephalographic patterns that define the stages of sleep. In depression, core body temperature tends to be elevated, growth hormone release is blunted, and there are increased awakenings. The brain's patterns of electrical activity are altered, with decreased deep (slow-wave) sleep and increased rapid eye movement (REM) sleep in individuals with mania and depression. These changes suggest increased nocturnal arousal, perhaps because of decreased inhibitory serotoninergic tone. *See* SLEEP AND DREAMING.

Positron emission tomography (PET) studies of brain function in depression document decreased prefrontal cortex metabolism and increased activity of deeper, paralimbic structures such as the amygdala. The former disturbance may explain the cognitive abnormalities of severe depression, whereas the latter may reflect intense affective arousal.

Mania, by virtue of the increased activity and characteristic changes in judgment and cooperativeness, has proved more difficult to study. In the prevailing view, it is a state of increased catecholamine activity, perhaps superimposed on a deficit of inhibitory serotoninergic input. Repeated studies of patients with bipolar affective disorder suggest that cerebral blood flow and glucose metabolism increase in prefrontal regions following a shift from depression to hypomania.

Psychosocial and neurobiologic vulnerabilities, no doubt, intersect. For example, harsh early maltreatment, neglect, or other abuses can have lasting effects on both self-concept and brain responses to stress. Low self-esteem, in turn, can amplify the impact of a particular adverse event for an individual. Conversely, blunted hedonic capacity, perhaps mediated by low serotonin and catecholamine levels, decreases coping responses and the willingness to seek out new rewards. Similarly, chronic hyperarousal undoubtedly reinforces caution and more avoidant and dependent behaviors.

Epidemiology

The lifetime rates of affective disorders are increasing, with an earlier age of onset in more recent generations. The risks in the United States are for major depressive disorder, 10–15% of individuals; dysthymia, 3–5%; bipolar disorder type I, 1%; and bipolar disorder type II, 1–3%. Among these conditions, only bipolar disorder type I has a comparable incidence among men and women. For the remainder, female preponderance (1.5 or 2 to 1 ratios) is observed in most western cultures, which probably reflects both psychosocial and biological vulnerabilities. Minor depressive symptoms, including insomnia, constitute a major longitudinal risk factor. Other risk factors include divorced or single marital status and concomitant psychiatric, general medical, or substance-abuse disorders.

The onset of major depression most often occurs in those in their late 20s to mid-30s; dysthymia and bipolar disorder typically begin about a decade earlier. However, no age group is immune to an affective disorder. Vulnerability is not strongly related to social class or race, although the affluent are more likely to receive treatment.

Worldwide, the affective disorders are underrecognized and undertreated. In the United States, less than one-third of depressed people are receiving proper treatment, and only 60% of those with bipolar disorder are treated at any given time. Untreated depressive episodes tend to last months longer, increasing the risk of chronicity. Once chronicity is established, spontaneous remission is uncommon and episodes may last decades.

About 70% of the first-onset major depressive episodes eventually become recurrent. After three or more episodes, the risk of further episodes (without

preventive treatment) exceeds 90%. This apparently increasing risk of recurrence is presumed to result from changes in brain stress responses referred to as "kindling."

The affective disorders, on average, cause more day-to-day disability than arthritis, diabetes, and most other common chronic illnesses. Depression's effects on the quality of life essentially match that of congestive heart failure or chronic obstructive pulmonary disease. The World Health Organization ranks depression as the world's fourth greatest health problem and estimates a second-place ranking in 2020. Bipolar disorder is ranked as the eighth greatest cause of illness burden. The lifetime risk of suicide is about 8–10% for the depressive disorders and 10–15% for bipolar disorder.

Treatment

Most episodes of dysthymia and major depressive disorder respond to treatment with either psychotherapy or antidepressant medication, either singly or in combination. Many experts now recommend the newer forms of psychotherapy, including cognitive-behavioral therapy and interpersonal therapy, because they have been better studied than more traditional psychoanalytic therapies and because they have been found to be as effective as medications.

Nearly 30 antidepressant medications are available worldwide, with most falling into three classes: tricyclic antidepressants (TCAs), selective serotonin reuptake inhibitors (SSRIs), and monoamine oxidase reuptake inhibitors (MAOIs). Most classes of antidepressants enhance the efficiency of serotonin or norepinephrine neurotransmission. Antidepressants are not habit-forming and have no mood-elevating effects for nondepressed people. *See* MONOAMINE OXIDASE; NORADRENERGIC SYSTEM; SEROTONIN.

The classes of antidepressants are not interchangeable, and a poor response or intolerable side effects with one medication do not preclude a good response to another. The SSRIs [in the United States, fluoxetine (Prozac®), sertraline (Zoloft®), paroxetine (Paxil®), citalopram (Celexa®), and escitalopram (Lexapro®)], venlafaxine (Effexor®), and bupropion (Wellbutrin®) are most often prescribed as first-line treatments because of their tolerability, safety, and relative simplicity. It is hard to predict which depressed patient will respond best to which particular antidepressant. When treatment is effective, an improvement should be apparent within 4–6 weeks. An effective antidepressant should be maintained for 6–9 months to reduce the risk of relapse. Even longer prophylactic treatment is recommended for prevention of highly recurrent depressions. *See* PSYCHOPHARMACOLOGY; PSYCHOTHERAPY.

There has been recent controversy about the possibility that antidepressants might paradoxically increase the likelihood of suicidality, especially early in the course of therapy. Although a cause-and-effect relationship has not been established, it is true that a small percentage of people who begin taking antidepressants (on the order of 1–3%) do feel worse and report either the onset, or worsening, of suicidal thoughts or actions. Accordingly, regulatory authorities such as the U.S. Food and Drug Administration have required that the manufacturers of antidepressants include warnings about the possibility of treatment-emergent suicidality and the need for careful monitoring. A related issue concerns the use of antidepressants in children and teenagers. In brief, younger depressed people may be at greater risk for treatment-emergent suicidality and, in youth, the therapeutic benefits of antidepressant medications are not as well established.

Acute manic episodes are usually treated with lithium salts, divalproex sodium, or one of the newer atypical antipsychotic medications, singly or in combination. The newer class of antipsychotic medication includes olanzapine (Zyprexa®), risperidone (Risperdal®), quetiapine (Seroquel®), Ziprasidone (Geodon®), and aripiperazole (Ability®). The potent sedative benzodiazepines, such as lorazepam or clonazepam, also are used to control agitation and insomnia. Several other drugs, notably carbamazepine, are used when the primary agents are ineffective.

Successful treatment of a manic episode should usually lead to preventive therapy. In addition to the medications used to treat mania, the anticonvulsant lamotrigine (Lamictal®) is used for preventive care. Relapses that occur despite preventive treatment may require the use of more complex, multidrug regimens. Although psychotherapy does not have a major role in the acute treatment of mania, it may help people come to terms with their illness, cope more effectively with stress, or curb minor depressive episodes. Antidepressants also are used to treat "breakthrough" major depressive episodes.

When pharmacotherapies are not effective, the oldest proven treatment of the affective disorders, electroconvulsive therapy (ECT), still provides a powerful alternative. Today, ECT is a highly modified and carefully monitored treatment that has little in common with its depictions in the movies. Nevertheless, confusion and transient amnesia are still common short-term side effects. Vigorous pharmacotherapy is needed after successful treatment to lessen the risk of relapse. *See* ELECTROCONVULSIVE THERAPY.

Michael E. Thase

Bibliography. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed. (DSM-IV), 1994; American Psychiatric Association, Practice guideline for the treatment of patients with major depressive disorder (revision), *Amer. J. Psychiat.*, 157 (4 suppl.):1–45, 2000; F. K. Goodwin and K. R. Jamison, *Manic-Depressive Illness*, Oxford University Press, 1990; R. M. A. Hirschfeld et al., Practice guideline for the treatment of patients with bipolar disorder (revision), *Amer. J. Psychiat.*, 159(suppl. 4): 4:4–50, 2002; S. D. Hollon, M. E. Thase, and J. C. Markowitz, Treatment and prevention of depression, *Psychol. Sci. Public Interest*, 3(2):39–77, 2002; M. E. Thase, R. Jindal, and R. H. Howland, Biological aspects of depression, in I. H. Gotlib and C. L. Hammen (eds.), *Handbook of Depression*, pp. 192–218, Guilford Press, New York, 2002.

Aflatoxin

Any of a group of secondary metabolites produced by the common molds *Aspergillus flavus* and *A. parasiticus* that cause a toxic response in vertebrates when introduced in low concentration by a natural route. The group constitutes a type of mycotoxin. Discovered in 1960 after a massive poisoning of turkey poults fed moldy peanut meal, aflatoxin has become the focus of a highly interdisciplinary research field involving chemists, mycologists, veterinarians, agriculturalists, toxicologists, and other basic and applied scientists. *See* TOXIN.

Chemistry. The naturally occurring aflatoxins are identified in physicochemical assays as intensely blue (aflatoxins B_1 and B_2) or blue-green (aflatoxins G_1 and G_2) fluorescent compounds under long-wave ultraviolet light. The common structural feature of the four major aflatoxins is a dihydrodifurano or tetrahydrodifurano group fused to a substituted coumarin group (see **illus.**). The relative proportions of the four major aflatoxins synthesized by *Aspergillus* reflect the genetic constitution of the producing strain and the parameters associated with fungal growth. In addition, derivative aflatoxins are produced as metabolic or environmental products.

Aflatoxins are formed through a polyketide pathway involving a series of enzymatically catalyzed reactions. In laboratory cultures, aflatoxins are biosynthesized after active growth has ceased, as is typical for secondary metabolites. By using blocked mutants and metabolic inhibitors, many of the intermediates have been identified as brightly colored anthraquinones. Working with these colored metabolites has made studies of aflatoxin biosynthesis into a model system for polyketide research and has also facilitated molecular genetics. Several genes encoding enzymes that catalyze steps in the pathway or affect more global regulatory mechanisms have been cloned by using complementation studies in which transformants are selected based on pigmentation.

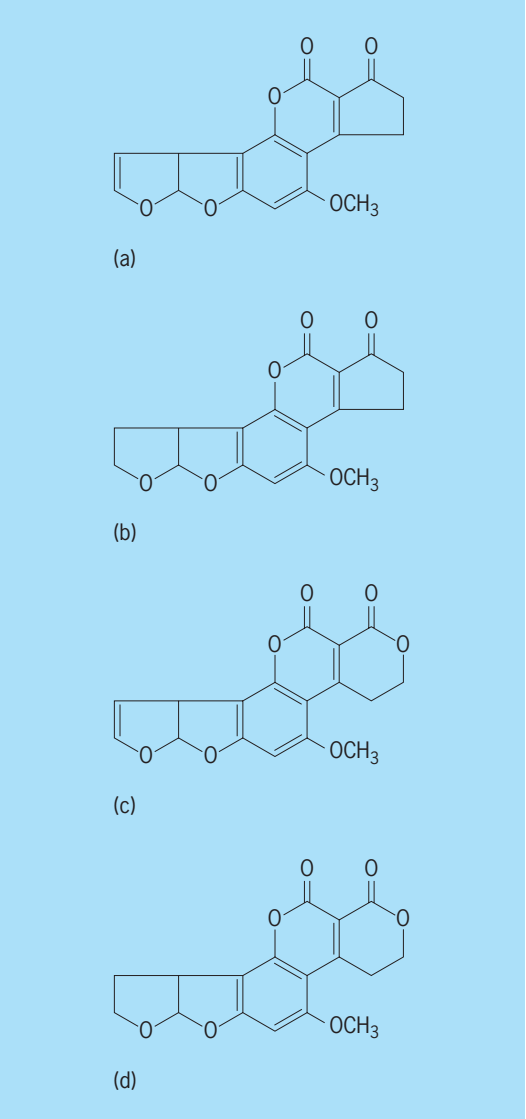

Structures of major naturally occurring aflatoxins. (*a*) B_1. (*b*) B_2. (*c*) G_1. (*d*) G_2.

Biological effects. Aflatoxins are potent molecules with many biological effects. They are toxigenic, carcinogenic, mutagenic, and teratogenic in various animal species. Aflatoxin B_1 is usually the most abundant naturally occurring member of the family, and most studies on the pharmacological activity of aflatoxin have been conducted with this congener.

Interspecific and interstrain variation is the rule for both acute and chronic effects. Trout and ducklings are among the most sensitive species known. The toxic properties manifest themselves in different ways depending on the test system used, the dose given, and the duration of exposure. In addition, age, nutritional status, and sex affect the toxicological response; in general, the young of a species are the most sensitive.

In veterinary medicine, aflatoxicosis is often difficult to detect. Inhibited growth, weight loss, and immune response breakdown are observed, and subtle synergies with both infectious and nutritional diseases complicate the diagnostic profile. The main target organ for aflatoxin is the liver. Aflatoxin B_1 is the most potent hepatocarcinogenic agent known, although the liver by no means is the only organ susceptible to aflatoxin carcinogenesis.

Epidemiological studies in Africa and Asia link ingestion of mold-contaminated food with incidence of primary liver cancer. Moreover, independent studies on liver cancer from southern Africa and Qidong, China, revealed guanine to thymine transversions in a tumor suppressor gene. Identification of this molecular hot spot constitutes some of the most convincing evidence that a natural product in a nonindustrial setting can induce a specific form of cancer. Aflatoxin is listed as a probable human carcinogen by the International Agency for Research on Cancer. *See* BOTULISM; LIVER DISORDERS; MUTAGENS AND CARCINOGENS; PLANT PATHOLOGY.

Economics. Aflatoxins are a major agricultural problem. Contamination can occur in the field, during harvest, or in storage and processing. Corn, rice, cottonseed, and peanuts are the major crops

regularly displaying high levels of aflatoxin contamination. Since *A. flavus* and *A. parasiticus* are nearly ubiquitous in the natural environment, numerous other grain, legume, nut, and spice crops, as well as coffee and cocoa, have been reported to contain aflatoxins. Given the potential of aflatoxins as human carcinogens and their known activity as toxins in animal feeds, many international regulatory agencies monitor aflatoxin levels in susceptible crops.

Contaminated agricultural commodities typically contain low amounts of aflatoxins (in the parts per billion or parts per million range) so numerous methods have been devised for extraction, concentration, and purification. Since agricultural commodities are diverse, no single extraction procedure can be applied uniformly to every product. Most of the original analytical methods employ thin-layer chromatography or high-performance liquid chromatography. More recently, immunological assays have become popular because they require fewer extractions and less complicated equipment.

Prevention is the main line of defense against aflatoxins entering the food chain. Moisture, temperature, and composition of the substrate are the chief factors affecting fungal growth and toxin production. In the field, insect damage is often involved. Good agronomic techniques that minimize damage to crops, storage under conditions unfavorable to mold growth, and screening of crops in order to exclude those which have been contaminated are the major methods for protecting the food supply. In some cases, portions of crops may be salvaged by mechanical separation. Detoxification is a last line of defense. Several commercially feasible methods of ammoniation have been developed for reducing levels of aflatoxin contamination in animal feeds.

Safety. Laboratory workers handling aflatoxins, especially in purified forms, should maintain strict safety precautions. Aflatoxin-contaminated foods should not be eaten or fed to domesticated animals. *See* AGRONOMY; MYCOTOXIN. J. W. Bennett

Bibliography. J. W. Bennett and M. A. Klich (eds.), *Aspergillus: Biology and Industrial Applications*, 1992; R. J. Cole and R. H. Cox, *Handbook of Toxic Fungal Metabolites*, 1981; J. E. Smith and M. S. Moss, *Mycotoxins: Formation Analysis, and Significance*, 1985.

Africa

A continent that straddles the Equator, extending between 37°N and 35°S. It is the second largest continent, exceeded by Eurasia. The area, shared by 55 countries, is 11,700,000 mi^2 (30,300,00 km^2), approximately 20% of the world's total land area. Despite its large area, it has a simple geological structure, a compact shape with a smooth outline, and a symmetrical distribution of climate and vegetation.

Structure and geology. The structural makeup is partially explained by the continental drift theory, in which all of the Earth's continents were previously assembled into one supercontinent called Pangaea (or Pangea). Over time, Pangaea broke into two landmasses, Laurasia and Gondwana, and subsequently into several continents. Africa occupied the core of Pangaea, and as neighboring landmasses drifted away, the present structure and configuration of the continent evolved. *See* CONTINENTAL DRIFT; CONTINENTS, EVOLUTION OF; PLATE TECTONICS; SUPERCONTINENT.

Three major rock formations can be identified. The first is a massive crystalline shield (Precambrian) that underlies most of the continent, with occasional outcrops in about one-third of the area. In some places, these rocks have been subject to metamorphic processes, resulting in a rich variety of mineral resources. The second set of significant rocks is the Karoo series in southern and eastern Africa. These are sedimentary rocks varying in age from the Carboniferous to the lower Jurassic Period. They are presently vital sources of coal. A third set of rock formations comprises marine sediments stretching in a discontinuous belt from the northwest to the Gulf of Guinea. These were formed during the Cretaceous Period, when parts of the continent were submerged in water. The sediments deposited at the time contain hydrocarbons that now account for oil production in Libya, Algeria, Nigeria, Gabon, and Angola. *See* CARBONIFEROUS; CRETACEOUS; JURASSIC; PRECAMBRIAN.

Africa has few inlets or natural harbors and a small number of offshore islands that are largely volcanic in origin. Madagascar is the largest island, with an area of 250,000 mi^2 (650,000 km^2). It is separated from the mainland by the Mozambique Channel, which is about 250 mi (400 km) wide. Close to Madagascar is the Mascarene, a volcanic ridge submerged in the Indian Ocean. Its highest peaks are exposed above the sea surface to form the islands of Seychelles, Mauritius, and Réunion. North of Madagascar, there is another group of volcanic islands, the Comoros. The largest of these is Njazide (Grand Comore), with an active volcano, Mount Kartala (7677 ft; 2340 m). *See* INDIAN OCEAN.

Surface features. Africa is primarily a high interior plateau bounded by steep escarpments. These features show evidence of the giant faults created during the drift of neighboring continents. The surface of the plateau ranges from 4000–5000 ft (1200–1500 m) in the south to about 1000 ft (300 m) in the Sahara. These differences in elevation are particularly apparent in the Great Escarpment region in southern Africa, where the land suddenly drops from 5000 ft (1500 m) to a narrow coastal belt. Although most of the continent is classified as plateau, not all of its surface is flat. Rather, most of its physiographic features have been differentially shaped by processes such as folding, faulting, volcanism, erosion, and deposition. *See* ESCAPEMENT; FAULT AND FAULT STRUCTURES; PLATEAU.

Basin surfaces. In several areas the plateau has sagged under the weight of accumulated riverine sediments to form large structural basins. These include the Sudan, Chad, and Djouf basins in the north, Congo (Zaire) basin in central Africa, and the Kalahari basin

in the south. These basins play a significant role in the hydrography of the continent. *See* BASIN; HYDROGRAPHY.

Rift valley system. The old crustal surfaces in the eastern part of the continent have been disrupted by a major structural feature known as the East African rift valley system that consists of a complex set of troughs formed by the downward displacement of land blocks against one another. The origin of the rift valley system can be traced to movements associated with separation of the Arabian plate from the African plate. The resulting landscape emanates from the Red Sea and passes through the Ethiopian highlands, along the course of the Awash river and a number of lakes (including Zwai, Shala, Abaya, and Turkana) into western Kenya. From there, it cuts across the Kenyan highlands and a series of lakes from Lake Banyo to Lake Magadi. It then enters Tanzania via Lake Natron, after which it loses its trough-like appearance as a result of the accumulation of sediments and volcanic ash. The system reappears in its original form as it passes through Lake Malawi and the Shire river into the lower Zambezi river and finally enters the Indian Ocean. *See* RED SEA.

The western arm of the rift valley starts north of Lake Albert (Lake Mobutu) in Uganda and passes through Lakes Edward (Amin), Kivu, and Tanganyika into Lake Nyasa, with a trench running southwest as the Luangwa valley of Zambia. Lake Tanganyika is the continent's deepest lake (5100 ft; 1500 m) with a bottom that is roughly 0.5 mi (800 m) below sea level.

The rift valley system is one of the most striking features of the African landscape. Sliding blocks have created wide valleys 20–50 mi (30–80 km) wide bounded by steep walls of variable depth and height. Within the eastern and western branches of the system, there is a large but shallow depression occupied by Lake Victoria. Between Lake Edward and Lake Albert on the western extension is the ice-capped Ruwenzori range, a block mountain, that rises over 16,732 ft (5100 m). These areas, along with other parts of the continent, occasionally experience earthquakes; the largest earthquake in the twentieth century occurred near the southern end of Lake Tanganyika. *See* RIFT VALLEY.

Volcanic landforms. Several volcanic features are associated with the rift valley system. The most extensive of these are the great basalt highlands that bound either side of the rift system in Ethiopia. These mountains rise over 10,000 ft (3000 m), with the highest peak, Ras Dashan, reaching 15,158 ft (4500 m). There are also several volcanic cones, including the most renowned at Mount Elgon (14,175 ft; 4321 m); Mount Kenya (17,040 ft; 5194 m); and Mount Kilimanjaro, reaching its highest point at Mount Kibo (19,320 ft; 5889 m). Mounts Kenya and Kilimanjaro are permanently snowcapped. *See* BASALT.

There are spectacular volcanic landforms in other parts of the continent such as the Ahaggar (Hoggar) mountains in central Sahara, which has over 300 lava plug domes rising above 1000 ft (305 m). There are also huge craters or calderas scattered in various places such as Trou au Natron in the Tibesti mountains and the Ngorongoro in Tanzania. Present-day active volcanoes include Oldoinyo Lengai in Tanzania; Nyamlagira and Nyriragongo in the Virunga range on the Congo (Zaire)–Uganda border; Erta Ale in the Afar Triangle of Ethiopia and on several offshore islands such as San Juan on Las Palmas islands; Fogo on Cape Verde islands; Kartala on Comoros islands; and Piton de la Fournaise on Réunion island. The largest active volcano is Mount Cameroon (13,350 ft; 4069 m). Although it lies on the mainland, it is part of a chain of volcanoes extending into the Atlantic Ocean. The islands of Fernando Po, Principe, and São Tomé are also part of this system. *See* OCEANIC ISLANDS; VOLCANO.

Fold mountains. The northwestern and southern margins of the continent have been subject to folding and subsequent erosion, resulting in a complex chain of mountains and valleys. The Atlas ranges in the northwest extend in a belt up to 200 mi (320 km) wide from southern Morocco to Tunisia, a distance of 1400 mi (2250 km). The system consists of a series of parallel mountains built by folding and uplifting processes during the Tertiary Period. Geologically they are an extension of the Alpine system of Europe, now separated by the Mediterranean Sea. The two principal chains are found along the coast and in the interior. The coastal chain starts from Tangiers in Morocco as the Rif, extends eastward into Algeria as the Tell, and ends in Tunisia. The interior chain starts west of Morocco as the High Atlas and Anti-Atlas. It continues eastward into Algeria as the Sahara Atlas and terminates in Tunisia. Between the two chains lies the Plateau of Shott, a salty shallow basin with internal drainage. Another plateau, the Moroccan Meseta, is found in the western extreme, sloping toward the Atlantic Ocean. *See* TERTIARY.

The Cape ranges in southern Africa are less complex, consisting of folded sedimentary rocks that correspond in age and topography to Appalachia in the eastern United States. The mountains run in a north-south direction along the west coast for about 150 mi (240 km) and then make right angles to run east-west for another 600 mi (900 km). *See* OROGENY; SEDIMENTARY ROCKS.

Climatic factors. Since the Equator transects the continent, the climatic conditions in the Northern Hemisphere are mirrored in the Southern Hemisphere. Nearly three-quarters of the continent lies within the tropics and therefore has high temperatures throughout the year. Frost is uncommon except in mountainous areas or some desert areas where nighttime temperatures occasionally drop below freezing. These desert areas also record some of the world's highest daytime temperatures, including an unconfirmed record of 136.4°F (58°C) at Azizia, Tripoli. *See* EQUATOR.

Africa is largely under the influence of subtropical high-pressure belts separated by the equatorial low-pressure zone known as the Inter-Tropical Convergence Zone (ITCZ). The position of this zone across the tropics shifts with the seasonal movement of the Sun. In July, it moves to North Africa along the

margins of the Sahara. In January, it moves south, taking up a clearly defined position in West Africa. Its position in eastern and southern Africa is not well defined because of the elevational variations there, but it tends to slope along the eastern margins of the Democratic Republic of Congo (formerly Zaire) and eastward through Mozambique.

The Inter-Tropical Convergence Zone plays a fundamental role in the circulation of the airstreams and subsequent climatic conditions within the continent. The associated air masses are the tropical maritime and the tropical continental air with notable differences in humidity due to their source areas. The tropical maritime air is warm and moist, originating from the Atlantic Ocean. The tropical continental air is warm and dry because it develops over the desert areas. In January, following the southern movement of the Inter-Tropical Convergence Zone, most of the Northern Hemisphere falls under the influence of tropical continental air, bringing in stable dry weather and dusty haze known as the harmattan in West Africa. In July, the shifting Inter-Tropical Convergence Zone and the tropical maritime air take over the region, bringing in heavy rainfall. *See* AIR MASS; CLIMATOLOGY; TROPICAL METEOROLOGY.

The wind reversals associated with the Inter-Tropical Convergence Zone produce distinct monsoonal seasons along the coastal margins of eastern and western Africa. The conditions observed along the eastern margins (including the island of Madagascar) are part of the larger Asian monsoonal system. During the wet summer monsoon, the northern shift of the Inter-Tropical Convergence Zone results in moist onshore winds from the warm tropical Indian Ocean. The atmospheric conditions are reversed in January, when dry offshore winds prevail following the southern shift of the Inter-Tropical Convergence Zone. *See* MONSOON METEOROLOGY.

The West African monsoon is smaller in magnitude than the Asian monsoonal system observed in eastern Africa. It dominates within 400 mi (650 km) of the coastline, particularly along the Gulf of Guinea. The causal factors are similar, however: the shifting wind belts produce alternate wet and humid conditions in the summer and dry and dusty conditions in winter.

Finally, the convergence of surface winds along the Inter-Tropical Convergence Zone is also responsible for the Easterly Wave, which occurs when the Inter-Tropical Convergence Zone is displaced away from the Equator, usually between 5° and 20° latitude. The weak wave is an area of low pressure that is propelled westward within the trade wind belt. Given the appropriate surface and weather conditions, these systems can develop into powerful, well-organized tropical disturbances that eventually grow into hurricanes. Most of the tropical storms and hurricanes in the northern Atlantic are linked to the Easterly Wave originating off the western coast of Africa. However, not all storms grow into hurricanes. Hurricanes typically develop where surface water temperatures are warm, 79°F (26°C) or higher, over a large area with sustained winds exceeding 74 mi/h (119 km/h).

Regional distribution of climate, vegetation, and soils. Africa can be classified into broad regions based on the climatic conditions and their associated vegetation and soil types. The tropical rainforest climate starts at the Equator and extends toward western Africa. The region has rainfall up to 200 in. (500 cm) per year and continuously high temperatures averaging 79°F (26°C). The eastern equatorial region does not experience these conditions because of the highlands and the presence of strong seasonal winds that originate from southern Asia.

The tropical rainforest vegetation consists of tall broadleaf trees that are categorized into different layers. The topmost layer, or emergents, consists of tall, straight trees that average 165 ft (50 m) in height. The middle layer, with trees of 80–155 ft (25–35 m), forms a dense interlacing canopy that allows little sunlight to penetrate to the forest floor. The lowest layer consists of young trees that average 50 ft (15 m) in height. The forest floor is typically damp and relatively open except for dead leaves, fallen branches, logs, and woody vines that connect the the tree trunks to the canopy.

The areal extent of the African rainforest region (originally 18%) has dwindled to less than 7% as a result of high rates of deforestation. Despite these reductions, the region is still one of the most diverse ecological zones in the continent. It has not only supported the timber industry in several countries but also served as an excellent reservoir for wildlife and medicines for traditional herbalists. Estimates from the International Union for Conservation of Nature show that the forested regions along the West African coast and within the central African basin contain more than 8000 plant species, including a large variety of hard and soft woods. This region also has about 84% of Africa's primate species and 68% of its passerine birds. Most of the mammals are arboreal, but there are several ground-dwelling animals such as rodents, wild pigs, and deer. Insects generally outnumber all other animal species combined. Typical are lateritic soils or oxisols with limited plant nutrients due to accelerated rates of decomposition and severe leaching. *See* RAINFOREST.

The tropical rainforest gradually merges northward and southward into the tropical savanna. This savanna encompasses a more extensive area and is best characterized by its distinct wet and dry seasons. Annual rainfall is lower (40 in.; 100 cm), and the mean annual temperature is higher than the equatorial zone. Vegetation includes savanna woodlands and shrubs interspersed with coarse tall grasses.

Unlike the forested zones, there is no continuous tree canopy. The woodlands consist of open stands of trees that are at least 16 ft (5 m) tall. The two major zones are the Guinea savanna region in West Africa and the Zambezian region in eastern Angola, the Democratic Republic of Congo, and Zambia. Dominant tree species include the African fan palm (*Borassus aethiopicum*), silk cotton tree (*Bombax petandrum*), boabab (*Adansonia digitata*), African

rubber (*Landolphia* spp.), and mopane and miombo tree species. In moving farther north or south of the Equator, tree size and overall density gradually decrease, with an increasing coverage of herbaceous plants and grassland.

Extensive savanna grasslands are found along the Sudanian zone of West Africa, within the Zambezian region and the Somalia-Masai plains. Large areas such as the Serengeti plains in the Somalia-Masai plains are home to a diverse range of wild animals. The soils are well developed and rich in nutrients because they are not subject to severe erosion. However, in some areas, especially where the topsoil has been removed, they have developed into thick crusts or pans during the dry seasons. The volcanic regions in eastern Africa are an exception, having productive soils.

The tropical steppe forms a transition zone between the humid areas and the deserts. This includes the area bordering the south of the Sahara that is known as the Sahel, the margins of the Kalahari basin, and the Karoo grasslands in the south. Average temperatures range from a daily high of 87°F (31°C) in August to 107°F (42°C) in April. Annual rainfall is highly variable, averaging 10–20 in. (25–50 cm). Inadequate rainfall coupled with poor soil chemistry (such as lack of nutrients) limits plant growth and species diversity. Nonetheless, this region has been used extensively for pastoral farming and for growing staples such as millet and sorghum. It is the most vulnerable ecological zone in the continent. In particular, the Sahelian zone, stretching from Mauritania and Senegal in the west to southern Sudan and Ethiopia in the east, has gained worldwide attention as a result of a series of devastating droughts that have affected millions of lives since the 1970s. The underlying causes for these problems are complex and intricately tied to both physical environmental and human-made factors.

Climatically, the most significant cause of desertification is the variability and marked reduction in rainfall beginning in the 1970s. This has been partly attributed to global warming trends. Specifically, higher sea surface temperatures over the Atlantic tend to hinder the northward penetration of the southwest monsoonal winds, thus reducing the likelihood for precipitation. An alternative explanation links the observed drought conditions to dust storms that frequently occur in the Sahara. These tend to heat up the atmosphere and prevent the upward flow of air needed for precipitation to occur. *See* DESERTIFICATION.

Aside from physical factors, the desertlike conditions in the Sahel are partly the result of specific human activities that degrade land. These activities include the removal of stabilizing natural vegetation for fuel wood and other land uses, overcultivation of the marginal areas, and overgrazing. Also contributing are a host of socioeconomic and political problems in the affected countries.

As one moves farther away from these marginal areas, the steppe trends into the desert landscape in the southwestern and northern parts of the continent. The Namib desert borders the Atlantic Ocean in southern Africa. Inland, it merges into the Kalahari, more accurately described as a semidesert instead of a true desert. In the north the world's largest desert, the Sahara, covers about 30% of the continent. The desert consists of mostly flat, dry plains and plateau with a few mountains such as the Ahaggar, the Air, and the Tibesti. Unlike most deserts around the world, more than 20% of the desert surface is covered with ergs (rolling plains of sand) such as the Libyan Erg (close to 200,000 mi^2 or 80,940 km^2) and the eastern and western ergs of Algeria (each about 20,000 mi^2 or 8094 km^2). Given this vast areal extent, the landscape takes on an undulating appearance with ripples of sand occasionally interrupted by small groups of dunes. In other areas, the desert landscape is characterized by regs (plains covered by pebbles and boulders), hammadas (polished rock surfaces), and basins of interior drainage (some containing salt pans). *See* DESERT.

Average annual rainfall in the desert is consistently below 10 in. (25 cm) with extreme variability from year to year. The relative humidity is very low (under 5% and rarely above 20%). There are wide annual temperature ranges and extreme diurnal ranges. The most common weather occurrences are sandstorms, particularly in the winter and early spring. Depending on the velocity and the volume of sand being transported, these storms can be disastrous for crops, livestock, and the inhabitants of this region. Extensive studies conducted in the last few decades have focused on the characteristics of these storms in relation to the materials deposited elsewhere (such as the Americas, Asia, and Europe). Specifically, airborne dust from the Sahara negatively impacts the ecology and climate within the region itself but provides valuable loess deposits in areas westward across the Atlantic Ocean, northward over the Mediterranean Sea, and eastward across the Red Sea. *See* LOESS.

Land-use patterns are extremely variable in the deserts. They range from extensive nomadic grazing of livestock to intensive cultivation around oases. These activities, however, depend on the local variation in climate, soil moisture, and the availability of water through traditional irrigation systems. The most important staple is the date palm. Fruits, vegetables, and cereals are also grown.

The northwestern and southwestern margins of the continent have a mediterranean climate, with dry summers and cool wet winters. The latter are caused by the polar air mass which brings relatively low temperatures and rainfall. Annual rainfall is about 30 in. (75 cm) or less. Mostly drought-resistant trees and scrubs are found here. The soils are variable, depending on the underlying bedrock, but can be generally classified as acidic and less fertile. *See* PLANT GEOGRAPHY; SOIL.

Drainage systems. The structural evolution of the continent has much to do with the drainage patterns. Originally, most of the rivers did not drain into the oceans, and many flowed into the large structural basins of the continent. However, as the continental

drift occurred and coasts became more defined, the rivers were forced to change courses, and flow over the escarpments in order to reach the sea. Several outlets were formed, including deep canyons, waterfalls, cataracts, and rapids as the rivers carved out new drainage patterns across the landscape. Most of the rivers continue to flow through or receive some of their drainage from the basins, but about 48% of them now have a direct access into the surrounding oceans. The major rivers are the Nile, Congo (Zaire), Niger, and Zambezi. *See* RIVER.

Nile. This is the only perennial river that enters the Mediterranean Sea from Africa. With an overall length of 4157 mi (6690 km), it is the longest river. The water originates from three major sources. About 56% comes from the Blue Nile, 22% from the Atbara River, and the rest from the White Nile. At Khartoum the White Nile joins the Blue Nile, and the plain between the two rivers, the Gezira, is the site of a large irrigation project. The Nile is controlled by several dams. The largest is the Aswan High Dam, which created Lake Nasser, over 300 mi (480 km) long. *See* DAM; MEDITERRANEAN SEA.

Congo. The Congo (Zaire), which is 3000 mi (4800 km) long, is Africa's largest river in terms of discharge. Its flow is stable throughout the year because it receives water from tributaries that lie on both sides of the Equator. The river originates near Lake Tanganyika and descends over a series of rapids and falls on its way down the Great Escarpment to the Atlantic Ocean. The Inga Dam, built between Kinshasa and Matadi, was completed in 1972. Large sections of the river are also navigable, particularly its middle course, which provides a 1085-mi (1746-km) stretch from Kinshasa to Kisangani. *See* ATLANTIC OCEAN.

Niger. The Niger rises in the Guinean highlands, an area of heavy rainfall in West Africa. After a total course of 2600 mi (4160 km), it enters the Atlantic through a large delta in Nigeria. It was dammed in 1969 at Kainji, Nigeria, creating a large artificial lake. Other large rivers in this region include the Senegal and the Volta. *See* DELTA.

Zambezi. The Zambezi begins in Angola, and its 2200-mi (3540-km) course drains most of the area south of the Congo system into the Indian Ocean. There are several obstructions in its middle course, including Victoria Falls. Here the Zambezi plunges 350 ft (107 m) from the inland Kalahari basin onto the edge of hard resistant rocks, cutting deep canyons in its path. A dam was built downstream from the falls in 1959, creating Lake Kariba. Also in Mozambique, the Caborra Bassa dam was completed in 1976 at the point where the river leaves the Kebrabasa gorge on its way to coastal lowlands. Other major rivers within southern Africa include the Limpopo and the Orange.

Fauna. The tremendous diversity in wildlife continues to be one of the primary attractions of this continent. Africa is one of the few remaining places where one can view game fauna in a natural setting. There is a tremendous diversity in species, including birds, reptiles, and large mammals. Wildlife are concentrated in central and eastern Africa because of the different types of vegetation which provide a wide range of habitats. Until recently, these regions were the least disrupted areas. Human population densities were relatively low except within highland and fertile volcanic areas. However, the situation is changing rapidly as the animals face constant threats, including disease, frequent drought, poaching, and population encroachment. Efforts to save them include the development of game reserves, bans on hunting or poaching, controlling exports of wildlife (or their products), and wildlife management. *See* CONSERVATION OF RESOURCES.

Population and environmental hazards. Africa is not a densely populated continent. With an estimated population of 743,000,000, its average density is 64 per square mile (26 per square kilometer). However, some areas have large concentrations, including the Nile valley, the coastal areas of northern and western Africa, the highland and volcanic regions of eastern Africa, and parts of southern Africa. These are mostly areas of economic or political significance.

In several regions, the human population is exposed to a variety of hazards such as riverine flooding, drought, disease, and pests. These hazards are largely meteorological in origin, such as the long periods of drought in the Sahel and eastern and southern Africa. The climate also provides a habitat for harmful insects and other vectors of endemic diseases such as malaria, yellow fever, trypanosomiasis, onchocerciasis, and bilharzia. Crops are occasionally attacked by swarms of locusts and other pests. *See* CONTINENT; EPIDEMIOLOGY.

Florence Lansana Margai

Bibliography. S. Aryeetey-Attoh (ed.), *Geography of Sub-Saharan Africa*, Prentice Hall, 1997; A. Grainger, *The Threatening Desert: Controlling Desertification*, Guersney Press, 1982; I. L. Griffiths, *An Atlas of African Affairs*, University of Witwatersrand Press, 1994; J. Gritzner, *The West African Sahel*, University of Chicago, Committee on Geographic Series, 1988; International Union for Conservation of Nature, *The Conservation Atlas of Tropical Forests in Africa*, Simon and Schuster, 1992; L. A. Lewis and L. Berry, *African Environments and Resources*, 1988; A. B. Mountjoy and D. Hilling, *Africa: Geography and Development*, Barnes and Noble, 1987; S. E. Nicholson, Climatic variations in the Sahel and other African regions during the past five centuries, *J. Arid Environ.*, 1:3–24, 1987; J. M. Pritchard, *Landform and Landscape in Africa*, Arnold, 1979; P. Richards, The environmental factor in African studies, *Prog. Human Geog.*, 4:589–600, 1980; C. Stager, Africa's Great Rift, *Nat. Geog.*, pp. 2–14, May 1990.

African horsesickness

A highly fatal insect-borne viral disease of horses and mules, and a mild subclinical disease in donkeys and zebras. It normally occurs in sub-Saharan Africa but occasionally spreads to North Africa, the Iberian Peninsula, and Asia Minor.

Etiologic agent. The African horsesickness virus is an orbivirus (family Reoviridae) measuring 68–70 nanometers in diameter. The outer layer of the double-layered protein shell is ill defined and diffuse and is formed by two polypeptides. The highly structured core consists of five structural proteins arranged in icosahedral symmetry. The viral genome is composed of 10 double-stranded ribonucleic acid (RNA) segments (genes) ranging in size from 240,000 to 2,530,000 daltons. Nine distinct serotypes which can be distinguished by neutralization tests are known. The virus can be cultivated in various cell cultures, in the brains of newborn mice, and in embryonated hen eggs by intravascular inoculation. *See* ANIMAL VIRUS.

Transmission. African horsesickness is a noncontagious disease that can readily be transmitted by the injection of infective blood or organ suspensions. In nature, the virus is biologically transmitted by midges of the genus *Culicoides*, such as *C. imicola*. The disease has a seasonal incidence in temperate regions (late summer to autumn), and its prevalence is influenced by climatic conditions favoring insect breeding (for example, warm, moist conditions in low-lying areas). Mechanical transmission by large biting flies is possible, but plays a much smaller role than biological transmission in the epidemiology of this disease.

Clinical signs. These indications result from selective replication of the virus in vascular endothelial cells, leading to increased vascular permeability in specific organs or regions of the body. Clinically, four forms of the disease are recognized. (1) In the peracute, pulmonary form, an incubation period of 3–5 days is followed by a febrile reaction and signs of respiratory distress. The breathing accelerates (75 breaths per minute or more), and the horse stands with its neck extended and the nostrils fully dilated. Profuse sweating is common, and terminally spasmodic coughing may be observed with frothy fluid exuding from the nostrils. Death usually occurs within hours after the first respiratory signs are observed. (2) In the subacute, edematous or cardiac form, the incubation period is about 7–14 days. The febrile reaction is followed by characteristic swellings of the hollows above the eyes, the eyelids, the lower lip, and the neck region. Swelling of the limbs is never observed. Some animals may show signs of colic, and death occurs 4–8 days after the onset of the febrile reaction. (3) Most horses show the acute or mixed form, which represents a combination of the pulmonary and edematous forms. (4) African donkeys, zebras, and partially immune horses may exhibit the fever form, where a mild febrile reaction is the only clinical sign.

The mortality rate in fully susceptible horses varies from about 70 to 95% and is about 50% in mules. In European and Asian donkeys, the fatality rate may be about 10% while African donkeys and zebras never succumb to the disease.

Pathology. The gross autopsy findings depend largely on the clinical form of disease manifested by the animal before death. In the pulmonary form, large amounts of fluid are found in the lung tissue or in the thoracic cavity, and death results from suffocation. Other lesions commonly observed are intense inflammation of the glandular fundus of the stomach (which can easily be confused with poisoning) and hemorrhage or swelling of the large and small intestines.

In the subacute form, there is extensive infiltration of subcutaneous and intermuscular tissues with yellow gelatinous fluid. In mild cases just the head and neck are involved, but in more severe cases the edema also involves the brisket and shoulders. Accumulation of fluid in the pericardial sac is a constant and very significant finding. Extensive hemorrhages are found on the surface of the heart as well as in the left ventricle. The lymph nodes are generally enlarged and edematous, and lesions in the gastrointestinal tract are largely similar to those found in the pulmonary form except that submucosal edema of the large intestine tends to be more conspicuous.

In the mixed form the lesions represent a combination of those observed in the pulmonary and edematous forms.

Treatment and control. There is no specific treatment for this disease. Infected animals should be given complete rest as the slightest exertion may result in death.

Stabling of horses at night during the African horsesickness season reduces exposure to insect vectors and hence the risk of disease transmission. Prophylactic vaccination is the most practical and effective control measure. In epidemic situations outside Africa, the causal virus should be serotyped as soon as possible, allowing the use of a monovalent vaccine. However, in endemic regions it is imperative to use a polyvalent vaccine, which should render protection against all nine serotypes of African horsesickness virus. *See* DISEASE; EPIDEMIC; VACCINATION.

Baltus Erasmus

Afterburner

A device in a turbojet aircraft engine, between turbine and nozzle, which improves thrust by the burning of additional fuel (**Fig. 1**). To develop additional

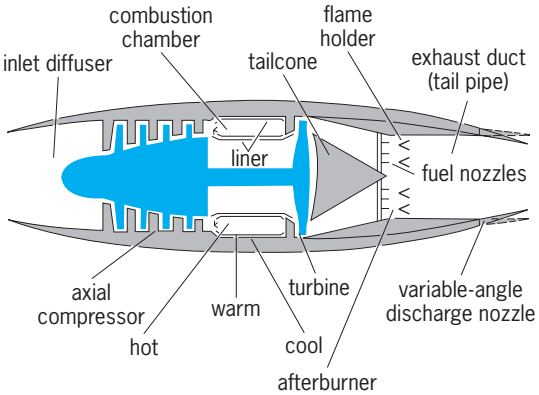

Fig. 1. Diagram of turbojet engine showing afterburner.

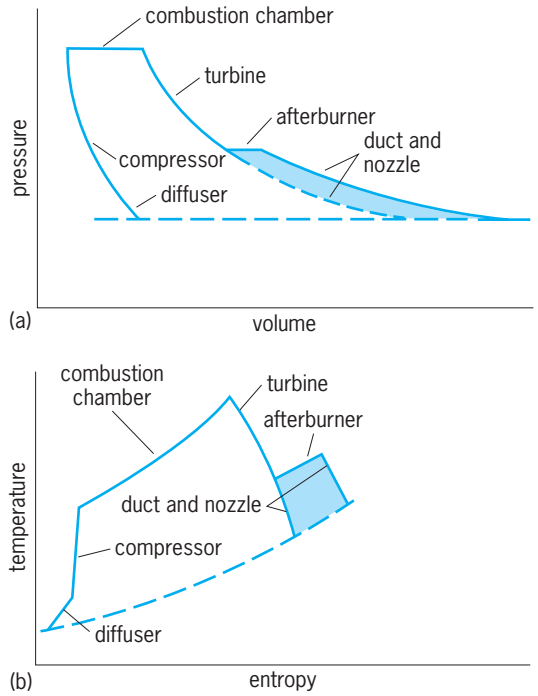

Fig. 2. Thermocycle diagrams for turbojet engine. (*a*) Pressure vs volume. (*b*) Temperature vs entropy. Shaded areas show increased output produced by afterburner.

thrust for takeoff and climb and for periods of dash of military aircraft, it is advantageous to augment the engine thrust. This is done by afterburning, also called reheating, tail-pipe burning, or postcombustion. **Figure 2** shows the effect on exhaust-nozzle gas pressure and temperatures. The augmentation of thrust obtained by afterburning may be well over 40% of the normal thrust and at supersonic flight can exceed 100% of normal thrust.

When a turbojet is developing the maximum thrust for which it is built, the compressor is operating at its maximum compression ratio and the gases enter the turbine at maximum allowable temperature. The most feasible means to increase the thrust further is to reheat the gases after they leave the turbine. In a conventional turbojet the gases discharge from the turbine at approximately 1500°R (800 K) and with sufficient oxygen to permit burning liquid hydrocarbon fuel. By afterburning, the gases may be heated to approximately 3500°R (2000 K) before they enter the discharge nozzle.

The augmented thrust depends directly on the increase in temperature and also on aircraft speed, afterburning being more effective at supersonic speeds. At subsonic speeds specific fuel consumption is approximately doubled so that the technique is suitable only for brief periods. At supersonic speeds the use of continuous afterburning is feasible. An engine with an afterburner also requires a variable-area nozzle because of the large difference in nozzle area required during the afterburning and nonafterburning conditions. *See* AIRCRAFT ENGINE; TURBOJET. Benjamin Pinkel

Agar

A major constituent of the cell walls of certain red algae, especially members of the families Gelidiaceae and Gracilariaceae. Extracted for its gelling properties, it is one of three algal polysaccharides of major economic importance, the others being alginate and carrageenan.

Agar is composed of two similar fractions, agarose and agaropectin, in which the basic unit is galactose, linked alternately α-1,3-(D-galactose) and β-1,4-(α-L-galactose). Agaropectin has substitutions by sulfate and pyruvic acid. In the closely related carrageenans, all of the galactose is in the D form and usually at least one unit of each pair is sulfated. The gel strength of the agar is proportional to its agarose content. Agar is insoluble in cold water, but swells by absorbing up to 20 times its weight of water. Dissolved in boiling water and cooled, it sets to a firm gel at concentrations as low as 0.5%. A 1% solution solidifies when cooled to 90–102°F (32–39°C), while the gel does not melt at a temperature below 185°F (85°C).

Agar is prepared by boiling the algae (agarophytes) in water, after which the filtered solution is cooled, purified, and dried. It is an amorphous, translucent material that is packaged in granules, flakes, bricks, or sheets. Before World War II, Japan monopolized the industry, but important production facilities have subsequently been established in Spain, Portugal, Argentina, and China.

Agar derives its name from the Malayan agar-agar, a traditional algal jelly now known to be a carrageenan. One of its chief uses is as a gelling agent in media for culturing microorganisms, a use begun by Robert Koch in 1882. Earlier it had been used (and continues to be used) in making confections. It is also used as an emulsifier in cosmetics and food products, as a sizing agent, as an inert carrier of drugs in medicine, and as a laxative. *See* CULTURE; RHODOPHYCEAE. Paul C. Silva; Richard L. Moe

Agate

A variety of chalcedonic quartz that is distinguished by the presence of color banding in curved or irregular patterns (**Fig. 1**). Most agate used for ornamental purposes is composed of two or more tones or intensities of brownish-red, often interlayered with white, but is also commonly composed of various shades of gray and white. Since agate is relatively porous, it can be dyed permanently in red, green, blue, and a variety of other colors. The difference in porosity of the adjacent layers permits the dye to penetrate unevenly and preserves marked differences in appearance between layers. The properties of agate are those of chalcedony: refractive indices 1.535 and 1.539, hardness 6.5 to 7, and specific gravity about 2.60.

The term agate is also used with prefixes to describe certain types of chalcedony in which banding is not evident. Moss agate is a milky or almost transparent chalcedony containing dark inclusions in a

Fig. 1. Section of polished agate showing the characteristic banding. (*Field Museum of Natural History, Chicago*)

Fig. 2. Moss agate cut as gems, showing the dendritic patterns. (*The Mineralogist*)

dendritic pattern (**Fig. 2**). Iris agate exhibits an iridescent color effect. Fortification, or landscape, agate is translucent and contains inclusions that give it an appearance reminiscent of familiar natural scenes. Banded agate is distinguished from onyx by the fact that its banding is curved or irregular, in contrast to the straight, parallel layers of onyx. *See* CHALCEDONY; GEM; QUARTZ.

Richard T. Liddicoat, Jr.

Agglutination reaction

A reaction in which suspended particles are aggregated or clumped. It occurs upon the admixture of another type of particle, a change in the composition of the suspending fluid, or the addition of a soluble agent that acts as a bridge between two or more particles. The reaction is a secondary one in that the process resulting in agglutination occurs after the primary antigen-antibody linkage has taken place.

The particles undergoing agglutination may be either unicellular or microscopic multicellular organisms (such as bacteria and parasites), individual cells of multicellular organisms (such as erythrocytes and lymphocytes), or artificial particles (such as beads of plastic, glass, or polysaccharide). The immunological specificity of agglutination depends upon the uniqueness of the reaction between a marker substance on one type of particle and a receptor on either another type of particle or a specific antibody in solution. The marker can be a usual biological component of the surface of the particle (for example, a specific polysaccharide on bacteria) or blood group substance on red cells. It can be an enzymatically or a chemically modified chemical group on the surface of biological particles. It can also be an adsorbed or a chemically attached substance. The attachment can be to biological particles (for example, bacteria, cells, and subcellular particle components) or artificial ones (for example, spheres of plastic such as latex, glass, or polysaccharide such as agarose). The receptor can be a biological component of the particle, an attached antibody, or antibody in solution. A reverse reaction is one in which the antibody is attached to a particle (say, latex) and the addition of the antigen (such as bacterial product) causes the mixture to clump. The addition of substances, typically polymeric (for example, polyethylene glycol, polyvinylpyrrolidone, and polyions), that affect the mobility or charge of the particles tends to increase the propensity of a mixture to agglutinate. Increased adherence can also occur if charged groups are removed by treatment of biological particles (such as erythrocytes) with enzymes (bromelin, papain, neuraminidase, and so on). Inhibition of agglutination can also be used to test for antigens, especially of low molecular weight, in a manner similar to that for agglutination itself. *See* IMMUNOASSAY.

Alexander Baumgarten

Bibliography. C. P. Engelfriet et al., *Blood Transfusion in Clinical Medicine*, 10th ed., 1997; N. R. Rose, E. C. De MacArio, and J. D. Folds, *Manual of Clinical Laboratory Immunology*, 5th ed., 1997.

Agglutinin

A substance that will cause a clumping of particles such as bacteria or erythrocytes. Of major importance are the specific or immune agglutinins, which are antibodies that will agglutinate bacteria containing the corresponding antigens on their surfaces. Agglutinin activity is frequently displayed by purified antibody preparations that also precipitate or give other serological reactions. Agglutinins are readily determined, and their presence is of diagnostic value to indicate present or past host contact with the microbial agent sufficient to result in antibody formation. *See* AGGLUTINATION REACTION; ANTIBODY; RICKETTSIOSES.

Analogous reactions involve erythrocytes and their corresponding antibodies, the hemagglutinins. Hemagglutinins to a variety of erthyrocytes occur in many normal sera, and their amounts may be increased by immunization. The blood group isoagglutinins of humans and animals are important special cases which must be considered in all proposed blood transfusions lest transfusion reactions result. Some, but not all, hemagglutinins may display further lytic activities when complement is added. *See* BLOOD GROUPS; COMPLEMENT.

Certain agglutinating systems agglutinate weakly, if at all, unless a conglutinin is added. Bacteria and erythrocytes may also, at times, be nonspecifically agglutinated by other substances, such as acids and plant extracts. Unfortunately, these latter agents are also frequently termed agglutinins. *See* CONGLUTINATION.

Henry P. Treffers

Bibliography. E. L. Cooper, *General Immunology*, 1982; D. M. Weir (ed.), *Handbook of Experimental Immunology*, 5th ed., 1996.

Aggression

Behavior that is intended to threaten or inflict physical injury on another person or organism; a broader definition may include such categories as verbal attack, discriminatory behavior, and economic exploitation. The inclusion of intention in defining aggression makes it difficult to apply the term unequivocally to animals in which there is no clear means of determining the presence or absence of intention. As a result, animal violence is usually equated with aggression.

Causes. There are four main approaches to understanding the causes or origins of human aggression. First, the basis may be differences among people, due either to physiological difference or to early childhood experiences. Second, there are sociological approaches which seek the causes of aggression in social factors such as economic deprivation and social (including family) conflicts. Third, causes may be found in the power relations of society as whole, where aggression arises as a function of control of one group by another. Fourth, aggression may be viewed as an inevitable (genetic) part of human nature; this approach has a long history and has produced extensive arguments. Earlier suggestions that aggression operated as a drive (for example, by S. Freud) and was generated in part by increases in drive level due to deprivation do not have any support. Some support for the role of individual genetic factors derives from breeding experiments which have shown that selective breeding for animal aggression is possible. *See* DEVELOPMENTAL PSYCHOLOGY.

Given the wide variation in aggressive behavior in different societies and the occasional absence of such behavior in some groups and some individuals, a general human genetic factor is unlikely. However, some genetic disposition to react with force when an individual is blocked from reaching a goal may provide an evolutionary basis for the widespread occurrence of violence and aggression. The existence of different kinds of aggression suggests that different evolutionary scenarios need to be invoked and that aggression is not due to a single evolutionary event. All of these approaches may have some degree of validity, and as with many explanations of complex human behavior, it is likely that aggression is multidetermined and rarely, if ever, due to a single factor. *See* BEHAVIOR GENETICS.

Types. Aggressive behavior can occur in a large number of different situations, and is often identified by the occasions when it occurs as well as by the behavior that is produced. Aggression in humans ranges through fear-induced aggression, parental disciplinary aggression, maternal aggression, and sexual aggression. E. Fromm identified one clearly biologically adaptive type, defensive aggression, which occurs when fight responses are mobilized in defense of an organism's vital interests, such as obtaining food or the protection of its young. The aim of defensive aggression is not destruction but the preservation of life. Thus, aggression can serve both destructive and constructive purposes. Among animals, the varieties of aggression include most of the human types as well as predatory aggression, territorial defense, and sexually related aggression in competition for a mate. *See* REPRODUCTIVE BEHAVIOR.

Humans and other animals. In most animals, aggression and violence is stimulus- and response-specific. That is, the appearance of a threat or a hunting object produces the aggressive response, but responses such as specific attack behavior are very delimited. In contrast, aggression in humans can be produced by a wide variety of situations, and aggression may incorporate any number of physical, social, or psychological forms. In addition, humans produce much more intraspecific aggression than other animals. It was once believed that only humans could show evidence of intraspecific aggression and that all seriously aggressive behavior in other animals was directed toward other species. The fighting and dominance behavior shown by the males of many species is usually placed in a different category, since it very rarely results in the death or severe incapacitation of either participant. However, it has been shown that intraspecific aggression also occurs in some animal species, although much less frequently than in humans.

Cultural factors. The incidence of aggression in humans varies widely as a function of cultural background and societal factors in general. Aggression in different cultures varies, for example, from some of the Yanomamo people of the Amazon Basin, who consider murder and rape natural and acceptable occurrences, to the Semai Senoi of West Malaysia, who are embarrassed by shows of violence and successfully teach their children to avoid violent or aggressive behavior. Among industrial societies, also, the incidence of violence and aggression varies widely.

Physiology. A number of different physiological bases of aggression have been proposed. Among them are a prevalence of parasympathetic over sympathetic nervous system activity, leading to very low levels of arousal and thrill-seeking aggressive behavior; left cerebral hemisphere dysfunction, which leads to a loss of control over impulsive behavior; and a high level of the hormone testosterone, which is related to some indices of aggression. *See* EMOTION.

George Mandler

Bibliography. E. Fromm, *The Anatomy of Human Destructiveness,* 1974; R. G. Geen, *Human Aggression,* 1990; G. Mandler, *Mind and Body: Psychology of Emotion and Stress,* 1984; A. Montagu, *The Nature of Human Aggression,* 1976.

Aging

Definitions of aging differ between biologists and behavioral scientists. The biologists regard aging as reflecting the sum of multiple and typical biological

decrements occurring after sexual maturation; the behavioral scientists view it as reflecting regular and expected changes occurring in genetically representative organisms advancing through the life cycle under normal environmental conditions. It is difficult to define normal aging, since many changes observed in older adults and previously perceived as concomitants of normal aging are now recognized as effects of disease in later life. The behavioral science view allows for incremental as well as decremental changes with aging. Senescence is not always equated with aging; it is viewed as the increasing vulnerability or decreasing capacity of an organism to maintain homeostasis as it progresses through its life span. Gerontology refers to the study of aging. Geriatrics refers to the clinical science that is concerned with health and illness in the elderly.

Demography. The dramatic growth of the older age group is a rather recent phenomenon. In the United States, for example, in 1900 approximately 3 million people (4% of the total population) were 65 years of age or older. By the mid-1980s, more than 28 million (12%) were in this age group. Moreover, by 2050, this age group is expected to increase to over 60 million (over 20% of the total population). The group 85 years of age and older is actually the fastest-growing segment of the population; it is expected to multiply more than sixfold between 1985 and 2050, to more than 16 million.

Not only is life expectancy generally increasing, but it is also increasing from age 65. In 1900 the average life expectancy in the United States was 47 years; by the mid-1980s it was 74.5. In Rome during the height of the Roman empire, average life expectancy was only in the low 20s (because of war, disease, famine, and childbirth). In contemporary America, average life expectancy from age 65 is approaching 20 years.

Scientific progress in gerontology and geriatrics has been substantial since the mid-1960s, but scientists are confronting a potential dilemma: progress in understanding the mechanisms of aging leading to increased longevity should not outpace progress in improving the quality of life in later years.

Aging versus illness. There are many myths and stereotypes about what changes in later life reflect normal aging as opposed to illness. For example, it was long thought that senility was the inevitable outcome of growing old and that most people over age 65 ended up in nursing homes. But research has revealed that only about 6% of the group 65 and older develop Alzheimer's disease, the most common form of "senility'' or dementia (marked impairment of memory and intellectual functioning). Under conditions of generally good health, older people maintain a high level of mental functioning. This is not to say that decrements do not normally occur in later life; some do, such as reaction time, which diminishes with aging. If the person is healthy, however, some increments also occur with aging. Vocabulary, for example, has been shown to continue to expand in studies of individuals followed well into their 80s. *See* ALZHEIMER'S DISEASE.

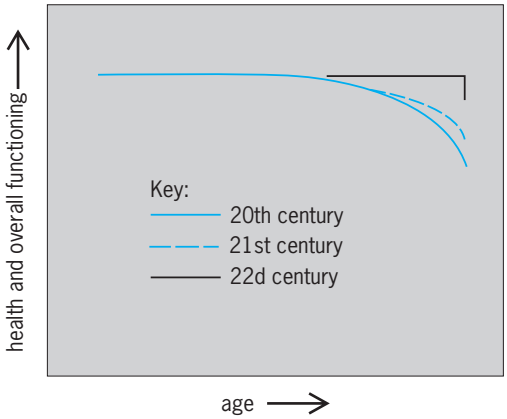

Hypothetical "curve" illustrating impact of reduced disability of the aging population in the twentieth century. The goal is to keep the interval of disability prior to death as high (rectangular) as possible at the end of the life cycle.

Similarly, most older people continue to maintain a high level of independent functioning. Only about 5% of the elderly are in nursing homes or hospitals; the remainder reside in the community. More and more people are maintaining stable functioning longer into the life cycle, resulting in the rectangularization of the curve representing their level of functioning as a factor of age (see **illus.**). The years of good health have been extended for these individuals, while the interval of disability has been reduced, in comparison with individuals of the same age at the beginning of the twentieth century. But there is another group, small in percentage but large in number, for whom increased longevity has brought increased years of disability.

Theories of aging. Inquiries into why organisms age involve both the purpose of aging and the process of aging. There are theories, but no conclusive evidence, in both areas. The most common explanations of the purpose of aging are based on theories that aging is adaptive for a species: aging leading to death is a mechanism by which species keep from overcrowding their environments; and finite life spans permit greater turnover of generations and thereby greater opportunities for a species to adapt, in a Darwinian sense. However, findings of very high accidental mortality among various species, resulting in few of that species having the opportunity to grow old, suggest that aging may not be necessary in an evolutionary sense. Still, the adaptive theories of why people age remain popular.

Theories about the process of aging concern how people age. These biological theories address two sets of factors—intrinsic and extrinsic to the organism. Intrinsic factors are influences operating on the body from within, such as the impact of genetic programming. Extrinsic factors are influences on the body from the environment, such as the impact of cumulative stresses. There are at least seven different major theories to explain the biological basis of aging: (1) Genetic programming: the organism's deoxyribonucleic acid (DNA) contains an intrinsic, built-in genetic program for aging leading to death. (2) DNA damage or genetic mutation: extrinsic

factors from the environment (such as radiation, environmental toxins, and viruses) damage the organism's DNA or cause the genetic material to mutate, bringing about aging. (3) Free radical: highly reactive chemicals (known as free radicals) within the body, produced by exposure to various extrinsic or intrinsic influences, interfere with normal cellular events, resulting in aging. (4) Cross linkage: improper diet or excessive caloric intake results in abnormal protein cross linkages between various cells of the body, resulting in age-associated changes like cataracts or skin problems. (5) Senescent substances: certain naturally produced substances (such as the pigment lipofuscin) or a protein (such as statin) in cells build up over time, altering the function of the cell and inducing age-related changes. (6) Immunologic: a combination of intrinsic or extrinsic factors cause cumulative alterations in the immune system over time, leading to problems associated with aging. (7) Hormones as clocks of aging: certain hormones may function as pacemakers or clocks of aging, affecting development and death. Despite an abundance of theories to explain the process of aging, its mechanisms remain a mystery.

Life extension. It is important to differentiate between life expectancy and life span. Life expectancy is the average number of years of life in a given species; it is significantly influenced by factors beyond aging alone, such as famine and disease. Life span is the maximum number of years of life possible for that species; it is more fundamentally linked to the process of aging itself. Over the centuries, life expectancy has increased (due to improved sanitation and health care practices); life span has not. Approximately 115 years appears to be the upper limit of life span in humans.

Life extension efforts really aim to increase life span. But just as the causes of aging have been elusive, sure formulas for life extension in humans have not been found. Efforts to achieve life extension have involved a wide diversity of approaches, including special diets and dietary restriction, exercise programs, vitamins and minerals, and drugs.

Many drugs have been used to reduce the age pigment lipofuscin, antioxidants have been used to combat free radicals, and hormones have been used to improve immune functioning. None of these approaches has had an unequivocal impact on life span in humans. At the same time, several of them, such as improvements in diet and exercise and the cessation of consumption of certain substances (for example, tobacco) and drugs (for example, excessive alcohol) have led to increase in life expectancy.

Expectations. Long overlooked but now increasingly recognized are the capacities for change, developmental growth, and creativity throughout the life cycle. This is apparent both in the response of aging brain neurons to environmental challenge and in the behavior of aging individuals. For example, Verdi composed operas in his 80s, Picasso painted into his 90s, and some people embark on a second career after initial retirement. *See* DEATH; HUMAN GENETICS.

Gene D. Cohen

Bibliography. R. H. Binstock and E. Shanas, *Handbook of Aging and the Social Sciences*, 1990; J. E. Birren, S. Birren, and K. Warren (eds.), *Handbook of the Psychology of Aging*, 1990; G. D. Cohen, *The Brain in Human Aging*, 1990.

Agnosia

An impairment in the recognition of stimuli in a particular sensory modality. True agnosias are associative defects, where the perceived stimulus fails to arouse a meaningful state. An unequivocal diagnosis of agnosia requires that the recognition failure not be due to sensory-perceptual deficits, to generalized intellectual impairment, or to impaired naming (as in aphasia). Because one or more of these conditions frequently occur with agnosia, some clinical scientists have questioned whether pure recognition disturbances genuinely exist; but careful investigation of appropriate cases has affirmed agnosia as an independent entity which may occur in the visual, auditory, or somesthetic modalities. *See* APHASIA.

Visual agnosias. The patient with visual object agnosia, though quite able to identify objects presented auditorily or tactually, cannot name or give other evidence of recognizing visually presented objects. Perceptual intactness is indicated by retained ability to draw such objects or to perform match-to-sample tasks. Because visual object agnosia is a rather rare disorder, knowledge of its underlying neuropathology is incomplete. Most reported cases have shown bilateral occipital lobe lesions, with the lesion extending deep into the white matter and often involving the corpus callosum. In the few cases where the reported lesion has been unilateral, the left occipital lobe has been the focus. *See* BRAIN.

Prosopagnosia is the inability to recognize familiar faces. Persons well known to the individual before onset of the condition, including members of the immediate family, are not recognized. In many instances, individuals fail to recognize picture or mirror images of themselves. Some individuals with prosopagnosia utilize nonvisual cues, such as voice, in order to identify a person. Specific visual features (a beard, facial scar, or hairstyle, for example) may also be used. The ability to utilize specific facial details for recognition, even though the overall facial configuration remains elusive, suggests that prosopagnosia might be understood as a highly specific breakdown in holistic visual perception (considered to be a function of the right hemisphere) though the ability to utilize perceived detail (a left-hemisphere function) remains intact. This view is consistent with the association of prosopagnosia with right posterior lesions, though bilateral lesions are common. *See* HEMISPHERIC LATERALITY; VISION.

Isolated impairment of reading is frequently considered to be an exotic form of aphasia. Logically, however, it may be considered as a visual-verbal agnosia (also referred to as pure word blindness or alexia without agraphia). Individuals with this disorder show a marked reduction in their ability to read

the printed word, though their writing and other language modalities remain essentially intact. Particularly dramatic is the failure of these persons to read material which they have written. The specific visual nature of the deficit is indicated by retained ability to decode words spelled aloud. Visual-verbal agnosia is typically associated with right homonymous hemianopsia (blindness in one-half of the visual field), indicating damage to the left occipital lobe. This lesion is typically coupled with damage to the corpus callosum, which connects the right and left visual association zones. The disorder has been interpreted as a disconnection syndrome. The patient's general visual functioning is good because of an intact right visual cortex. But in order for linguistic materials to be decoded, activity of the language-dominant left hemisphere is required, and this is precluded by the callosal lesion.

Auditory agnosias. The term auditory agnosia is most often used to indicate failure to recognize nonverbal acoustic stimuli despite adequate hearing sensitivity and discrimination. The individual fails to recognize a telephone by its ringing sound, for example, though it is identified immediately by sight or touch. In most well-documented cases of agnosia for sounds, the subjects have had bilateral temporal lobe lesions. In those instances where the damage has been unilateral, the right temporal lobe has been involved, suggesting a special role for this brain region in the recognition of nonverbal acoustic stimuli.

Auditory-verbal agnosia (or pure word deafness) is a disturbance in comprehension of spoken language, in the presence of otherwise intact auditory functioning and essentially normal performance in other language modalities. The person's speech expression is remarkably intact in comparison with the gross impairment in understanding speech. Like its visual analog, visual-verbal agnosia, this is a disconnection syndrome. It is produced by damage to the left primary auditory cortex (or the tracts leading to it) coupled with a lesion to the corpus callosum. The individual is thus able to carry out most auditory functions because of an intact right hemisphere; but auditory-linguistic material cannot be interpreted because the information fails to reach the dominant left hemisphere.

Phonagnosia is a disturbance in the recognition of familiar voices. The person has good comprehension of what is spoken, but the speaker cannot be identified. Phonagnosia, like its visual analog, prosopagnosia, is associated with right-hemisphere damage; in fact, these two disorders frequently coexist. *See* HEARING (HUMAN); PSYCHOACOUSTICS.

Astereognosis. Astereognosis (or tactile agnosia) is the inability to recognize objects by touch and manipulation. One or both hands may be affected. The deficit is at the associative level and is not explainable on the basis of impaired somesthetic sensation or perception. A diagnosis of astereognosis thus requires that the patient show reasonably intact sensitivity and discrimination for somesthetic stimuli. In addition, the perceptual features of the object (such as size, shape, or texture) should be recognizable. The classical view concerning localization is that the lesion responsible for the impairment is in the somesthetic association cortex opposite the affected side of the body. The evidence for this is tenuous, however, because astereognosis is so often complicated by sensory-perceptual impairment that few unequivocal cases have been satisfactorily studied. *See* PERCEPTION; SENSATION.

Gerald J. Canter

Bibliography. G. J. Beaumont, *Understanding Neuropsychology*, 1988; K. M. Heilman and E. Valenstein (eds.), *Clinical Neuropsychology*, 3d ed., 1993; M. H. Johnson (ed.), *Brain Development and Cognition: A Reader*, 1993.

Agonomycetes

Fungi which usually produce neither sexual (meiotic) nor asexual (mitotic) spores. These fungi have been denoted by various names. The most common names are Mycelia Sterilia (having sterile mycelium), Agonomycetales (used when classifying these fungi at the level of order), or Agonomycetes (used when assigning their rank at the level of class). These fungi make up an artificial (nonphylogenetic) group consisting of both ascomycetes and basidiomycetes. A few species have associated sexual states, but these sexual states are infrequently produced. Although true spores are mostly lacking in this group, some species produce sporelike structures which function effectively in survival and dispersal of the species. Some of these sporelike structures mimic true spores so well in form and function that the genera producing them (for example, *Beverwykella*, *Cancellidium*, and *Tretopileus*) have occasionally been classified with spore-forming fungi. *See* ASCOMYCOTA; BASIDIOMYCOTA; DEUTEROMYCOTINA; FUNGI.

Morphology and function. The most common sporelike structures are allocysts, bulbils, chlamydospores, papulospores, and sclerotia. These structures are not formed in the same manner as sexual spores (via meiosis), nor are they formed like true asexual spores, because the sporelike structures are simply modified hyphae. Chlamydospores originate from terminal or intercalary hyphae, have thickened walls, and are primarily a survival structure; they tend to be unicellular but often occur in pairs or chains. Allocysts resemble chlamydospores but store metabolic by-products and do not germinate. Bulbils are compact, multicellular, and internally undifferentiated propagules (somatic structures of propagation or survival); they resemble papulospores except that the latter have a core of large cells and a sheath of smaller cells. There are propagules intermediate between bulbils and papulospores as well. Sclerotia are larger, multicellular aggregations of hyphae and are often differentiated into layers at maturity. Outer layers may have dark, thick-walled cells and the inner layers may have thin-walled cells, but such differentiation is not universal. Sclerotia occur in a range of sizes and shapes, depending on the species forming them, and the smaller ones are termed microsclerotia. These various structures formed from hyphae

may also be present in species of fungi producing true spores (sexual and/or asexual). Chlamydospores and sclerotia are especially common. In general, the larger propagules serve for survival more than dissemination, but any propagule may be transported in soil particles by machinery or tools, movement of animals or humans, irrigation water, or strong winds. Sclerotia and chlamydospores may also be disseminated with plant seed, vegetative plant parts, and nursery stock.

Identification, pathology, and economic significance. The absence of true spores and structures producing true spores renders identification difficult. However, morphology-based keys to genera have been constructed using such characters as whether hyphae are simple or aggregated, the manner in which hyphal aggregations are formed, the presence of dolipore septa or clamp connections (indicating affiliation with basidiomycetes), and the manners in which chlamydospores, bulbils, sclerotia, papulospores, or other propagules are formed and positioned. Commercially produced kits are available for identification of *Rhizoctonia* and *Sclerotinia* (by ELISA [enzyme-linked immunosorbent assay] or PCR [polymerase chain reaction]), and experimental identification of selected genera of Agonomycetes has been published using techniques such as PCR-RFLPs [restriction fragment length polymorphisms], RAPDs [random amplifications of polymorphic DNA], and DNA (deoxyribonucleic acid) sequence analyses.

Several members of the Agonomycetes are of extreme importance as plant pathogens. *Rhizoctonia* species (which have infrequently produced sexual states in *Ceratobasidum* or *Thanatephorus*) cause root diseases. *Rhizoctonia solani* is a pathogen with an exceptionally broad host range encompassing dozens of genera, but is especially important on beans, carrots, cabbage, and other vegetables, and *R. cerealis* is important on wheat and other small grains. Interestingly, some *Rhizoctonia* species are beneficial mycorrhizal symbionts with orchid roots. In general, a nonsporulating fungus can be identified as *Rhizoctonia* by the thickness, septation, growth rate, and branching patterns of its hyphae. However, the taxonomy and identification of *Rhizoctonia* as to species and important subspecific groups is complex, and depends to considerable extent on anastomosis groupings, that is, how pairs of isolates behave when confronted with each other on artificial laboratory media.

Also important is the form-genus *Sclerotium*, species of which produce sclerotia. Some *Sclerotium* isolates will produce sexual states in *Athelia* (a basidiomycete) or other genera. The most important plant pathogens include *S. rolfsii* (corresponding to *Athelia rolfsii*) which is a pathogen of many crops, especially vegetables, and *S. cepivorum*, the causal agent of white rot of onion and garlic. A fungus sometimes classified and keyed with the Mycelia Sterilia is *Sclerotinia sclerotiorum*. Many isolates of this fungus reproduce sexually in nature or under proper, carefully controlled laboratory conditions, and some also produce small, nongerminating asexual spores, but on artificial media in the laboratory isolates often produce only sclerotia. *Sclerotinia sclerotiorum* has an exceptionally broad host range but is especially important on vegetables. *Sclerotinia trifoliorum* and *S. minor* are important on legumes and lettuce, respectively. Prior to the advent of molecular genetics, and still today in routine diagnoses, species of *Sclerotinia* are separated on the basis of characters of the colonies and sclerotia on artificial media. Once established in soil, sclerotia may survive for years, even decades, rendering disease control extremely difficult.

Other notable members of the Agonomycetes are *Papulaspora byssina*, troublesome in commercial mushroom operations, and *Cenococcum geophilum*, an ectomycorrhizal fungus symbiotic with many species of higher plants. *See* PLANT PATHOLOGY.

Frank M. Dugan

Bibliography. H. L. Barnett and B. B. Hunter, *Illustrated Genera of Imperfect Fungi*, 1998; E. Kiffer and M. Morelet, *The Deuteromycetes: Mitosporic Fungi, Classification and Generic Keys*, 2000; J. A. von Arx, *Plant Pathogenic Fungi: Beihefte zur Nova Hedwigia*, vol. 87, 1987.

Agouti

A large rodent that resembles the rabbit or hare in size and shape as well as in the elongated hindlegs, which make them well adapted for speed (see **illustration**). The agouti and the closely related smaller acouchi are inhabitants of clearings in forested

Agouti (*Dasyprocta aguti*), a rodent found in Mexico, South America, and the West Indies. (*Courtesy of Lloyd Glenn Ingles; © California Academy of Sciences*)

areas of the Amazon region. Some range into Central America and as far as the Guianas.

Thirteen species of agouti have been described, the common one being *Dasyprocta aguti*. Some authorities are of the opinion that all of these are varieties or subspecies of *D. aguti*. The acouchi is represented by two species, the green acouchi (*Myoprocta pratti*) and the red acouchi (*M. acouchy*). Behaviorally the agouti and acouchi have many similarities; thus, some authorities believe that the acouchi may be a dwarf variety of the agouti.

Dentition is typical of the cavimorph rodents, and the dental formula is I 1/1 C 0/0 Pm 1/1 M 3/3. Like most rodents, the agouti is a prolific breeder, and may breed at any time during the year. These animals feed primarily on vegetation, such as roots, berries, and fruits. Ecologically they form a part of the food chain for a variety of carnivorous animals, including humans, because they are the basic available animals common to the area. *See* DENTITION; FOOD WEB; RODENTIA.

Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999.

Agricultural aircraft

Aircraft designed, or adapted from a general utility airframe, for use in agriculture and forestry and for control of insect vectors of human, animal, and plant diseases. Agricultural aircraft have become an indispensable tool for high-productivity agriculture and have contributed to the worldwide crop production revolution.

Agricultural aircraft are widely used in developed countries and, increasingly, in developing countries, particularly on high-value export crop and forest products. Aircraft use covers a wide range of agricultural crop and pest applications, including control of competing weeds and unwanted brush and trees, control of insect and disease pests, application of plant nutrients, and broadcast seeding of many crops. Other, less widely used applications include defoliation to aid in cleaner crop harvesting, frost protection, rain removal from sensitive horticultural crops, and control of insect vectors or disease carriers affecting plants, animals, or humans. Large-area applications, frequently with twin- and four-engine transport as well as smaller-type aircraft, are used to control outbreaks of locust and grasshoppers and reduce or eradicate specific insect pests, such as fruit flies. Releases of sterile male insects and predacious spiders are made by aircraft, and other nonchemical biopesticides, such as bacteria, fungi, and various insect hormones that interfere with mating or metamorphosis, are applied as capsules, as baits, or in liquid carriers. *See* DEFOLIANT AND DESICCANT; FERTILIZER; PESTICIDE.

The principal advantages of either fixed- or rotary-wing aircraft for the treatment of crops lies with their ability to rapidly cover large crop acreages and to travel over rough terrain, irrigation structures, and wet fields. This timeliness factor is often considered critical to optimum pest control and effective crop protection. The primary disadvantage is their inability to direct the released material onto the target crop with the precision that can be accomplished with ground-based applications. This lack of precision, together with the losses in the form of drift or movement away from the target area, constitutes an area of concern. Increasing limitations are being placed on aircraft use, especially when highly toxic crop chemicals are to be applied to small fields and where sensitive nontarget crops are grown nearby.

Types. The two basic types, fixed- and rotary-wing aircraft, are both used for all of the applications noted above. In the United States, there are approximately 6000 aircraft in use, of which about 8% (480) are helicopters, flying an estimated total of 1.875×10^6 h per year. Each aircraft averages approximately 150 acres (60 hectares) per flight hour; thus, approximately 300×10^6 crop acres (120×10^6 hectares) are treated each year. There has been a gradual change to larger-capacity aircraft powered by engines of 1000–1500 horsepower (750–1125 kW), 10–15% of which are more powerful gas turbine-propeller types. Average fixed-wing aircraft carry 250–400 gallons (946–1514 liters), with the largest capable of lifting 800 gallons (3028 liters). Helicopters in agricultural use carry loads of about one-half these amounts.

While larger, faster-flying (up to 140 mi/h or 225 km/h) aircraft can accomplish greater hourly productivity, other factors such as size and location of crop fields, distance to fields from suitable landing strips, and time lost in field turns also affect field productivity. Helicopters, which can turn quickly and can land and be serviced close to crop fields or even on landing docks on top of the service trucks, can thus gain back some of the lowered productivity of these aircraft due to smaller load capacity and lower field speeds. However, it is the greater application precision and downblast from their rotary wing (at reduced forward speeds) that has generated the increased use of helicopters, in spite of their greater initial and operating costs.

Dispersion wake. The aerodynamic wake or air dispersion produced by the action of the wing or helicopter rotor and the mass of air displaced by the aircraft in motion are utilized to aid in spreading either liquid or dry materials. The interaction of the wing or rotor length and the length of the spray boom, along with the strength of the wake and height of flight, controls the usable swath width laid down by the aircraft.

Application equipment. The application equipment for either fixed- or rotary-wing aircraft is customarily an integral part of the aircraft. The spray equipment for a fixed-wing aircraft (**Fig. 1**) may be detached from the bottom of the tank or hopper, and a ram-air-type spreader for dry materials may then be brought into place. Spray equipment for helicopters (**Fig. 2***a*) is similar to that for fixed-wing aircraft. However, since helicopters are usually operated at lower speeds (60–90 mi/h or 96–145 km/h), windmill-powered pumps and ram-air spreaders do

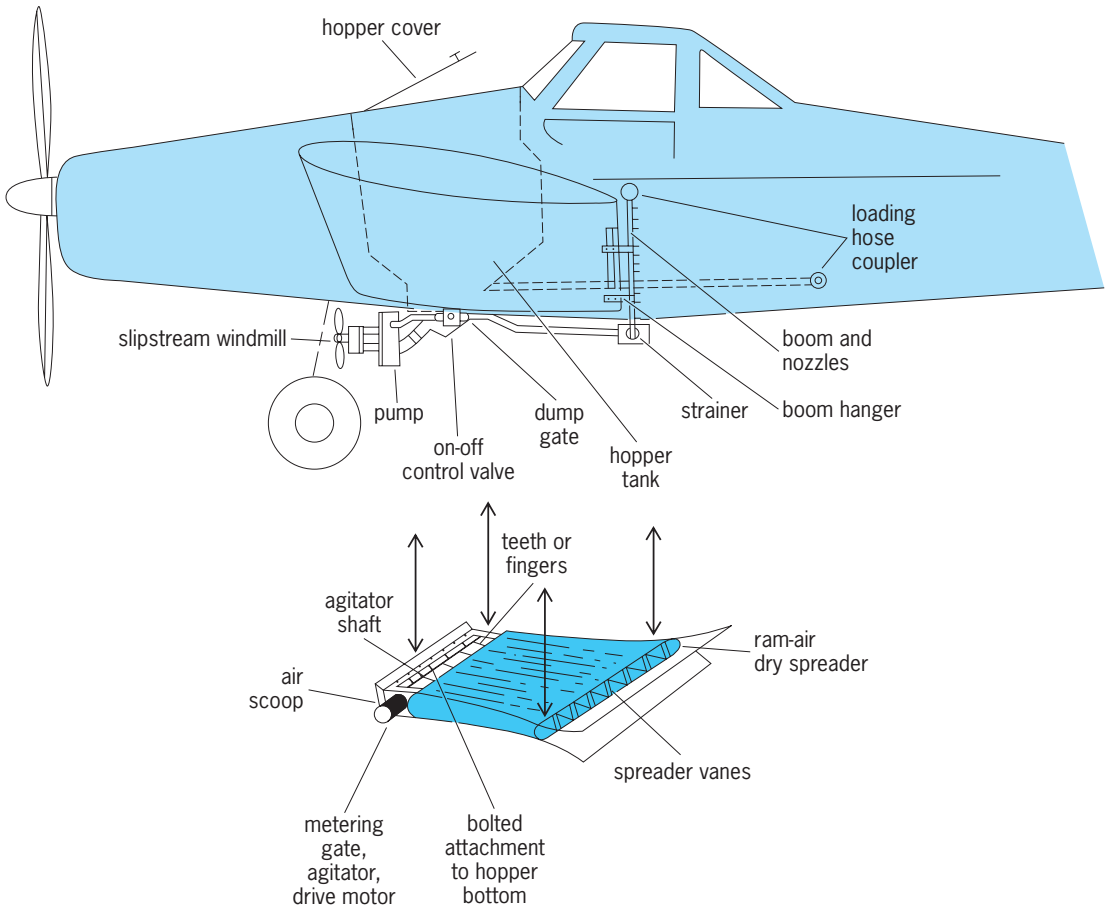

Fig. 1. Application equipment for fixed-wing agricultural aircraft. The spraying equipment is shown on the aircraft, and a ram-air-type spreader for dry materials is shown below the aircraft.

not function well on these aircraft. Another primary difference between aircraft types is the requirement for lateral as well as fore-and-aft balance on the helicopter. This necessitates the use of split hoppers, one on each side of the helicopter cockpit. A large crossover pipe connects the bottoms of the two hoppers. A pump moves the liquid formulation from its connection at the crossover to the spray boom. For dry materials, each hopper has a controlled feeder and usually a spinning device (Fig. 2*b*) to distribute the dry granules, seeds, fertilizers, or baits. Another helicopter system for applying both spray and dry materials is the sling unit, which can be suspended below the helicopter as a single hopper. Application materials from this unit are distributed in much the same way as from the hopper used on fixed-wing aircraft. Power for the pump or dry-materials spinner may be supplied by electric or hydraulic motors, these in turn being powered from generators or hydraulic pumps on the main engine. Auxiliary gasoline engines may also be used to furnish power directly or through electric or hydraulic coupling to the pumps or spinners.

Spraying equipment. The spraying equipment for a fixed-wing aircraft (Fig. 1) consists of a hopper bottom assembly bolted to the hopper, which supports the pump, the slipstream windmill with brake, and the on-off control valve. The spray boom is customarily attached to the trailing edge of the wing by hangers fastened to the wing frame. Nozzles for atomizing the spray along with their diaphragm cut-off valves are placed along the boom on each wing. Pump pressure forces liquid against the diaphragms, which open and permit flow to the atomizers for a hollow-cone nozzle. Stationary whirl plates behind the nozzle orifices spin the liquid, causing it to form circular or conical sheets on discharge. A liquid-tight cover is located on the top of the hopper tank, but loading is customarily done by means of hoses that are quick-coupled to loading couplers located either at the end of the boom or at the end of a special pipe leading directly into the hoppers. *See* ATOMIZATION; NOZZLE.

Centrifugal pumps are customarily used for high volume rates of 50–100 gal/min (190–387 liters/min) at relatively low pressures of 40–60 lb/in.^{2} (276–415 kilopascals). On fixed-wing aircraft these may be powered by a windmill (Fig. 1), or by electric or hydraulic motors powered though electric generators or hydraulic pumps directly driven from the propulsion engine. The latter approach is customarily usedfor the helicopter.

Spreading dry materials. The ram-air spreader is used on the fixed-wing plane for applying dry materials such as dusts or granules (Fig. 1). The liquid system is unbolted and dropped, and the ram-air dry

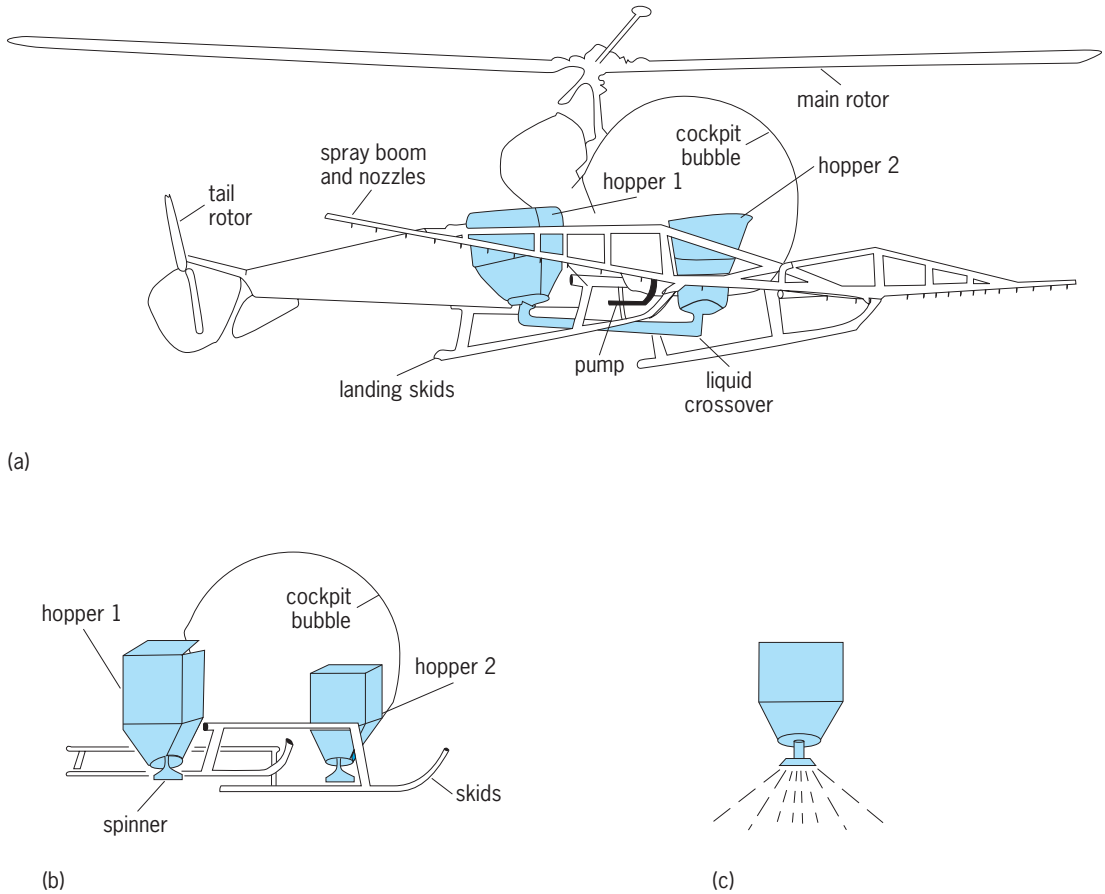

Fig. 2. Application equipment for rotary-wing agricultural aircraft. (*a*) Helicopter with spray equipment attached. (*b*) Dry equipment. (*c*) Detail of hopper and spinner.

spreader is then attached. This operates on the principle of a venturi device with air gathered by a large frontal opening then forced through a narrow section directly beneath the hopper and metering gate, which then widens by a factor of 1.5–2 at the discharge end. Longitudinal vanes inside the spreader guide the flow of material and help to direct the material outward from the aircraft. An agitator, either windmill- or motor-driven, may be located in the throat of the spreader, just below the hopper. On helicopters, detachable spinner-type spreaders for dry materials can be placed below each of the twin hoppers (Fig. 2*c*). A metering valve is placed between each hopper outlet and the spinner to control the flow of material to each spinner.

Spray applications. The characteristic spray swath laid down by the aircraft can be determined by catching and analyzing the amount of spray caught on plastic or paper tapes or sheets laid perpendicular to the aircraft flight pattern. A rapid spray deposit evaluation technique uses a fiber string or plastic monofilament stretched across the flight line to catch the spray. Analysis is performed by feeding the string through a special washing unit which strips the string of tracer dye and evaluates the amount of collected spray caught by the string. Another system uses a video camera and dyed spray to visually indicate high- and low-density portions of the spray pattern. The amount of drift loss, however, can only be measured by air samples.

A single pass of a fixed-wing aircraft can spread deposits over a width of about 180 ft (55 m). However, the useful or flagged swath width that can produce a uniformly deposited pattern from successive swaths across a field is considerably less, depending on spray drop size.

The spray pattern from a helicopter application is much the same as that from a fixed-wing aircraft. The swath width is a function of the rotor width, which creates much the same type and strength of downwash as is produced by a fixed-wing aircraft of similar wing span.

A typical spray drop spectrum produced by hollow cone nozzles on the spray boom of a fixed-wing aircraft, directed with a 100-mi/h (160-km/h) slipstream, has a volume median diameter (50% of the spray volume in drops above the volume mean diameter, and 50% below this size) of 454 micrometers. The drop size for 0–10% by volume is 207 μm, and the size for 0–90% by volume is 896 μm. These values indicate that drop sizes ranging from 20 to 1100 μm in diameter are produced by this atomizer.

Dry applications. The swath patterns from applications of dry materials such as granules, seeds, baits, or other dry particles can be patterned by collecting the materials in buckets placed across the aircraft

flight pattern. The graphs produced will appear similar to the spray distribution patterns, whether the materials are applied by ram-air or rotary spreaders, on fixed- or rotary-wing aircraft.

Airborne transport or drift. The movement of particles or spray drops less than 100 μm in diameter constitutes the burden of the contamination or drift problem that has plagued the aircraft application industry since its inception. After release from the aircraft, the distances that these small particles can travel are dependent on atmospheric conditions, which control their rate of diffusion or dispersion in the downwind atmosphere. Principal among these dispersion factors is atmospheric stability, or the rate of air mixing, vertically and horizontally. A temperature inversion with warmer air overhead at 100–200-ft (30–60-m) height will prevent vertical mixing, and minimal winds will reduce lateral diffusion or spreading of the pesticide-laden cloud.

The first step taken to reduce and control the drift loss of pesticide, sprays, or dry particles is to increase the size of the particles or spray drops as much as is practical. Spray releases with a volume median diameter over 1000 μm will minimize the percentage of drops present below 100 μm diameter, reducing the volume of these to 1–2% of the total. This control will reduce drift losses beyond displaced swath distances of 200–300 ft (60–90 m) but will not eliminate all drift loss. However, sprays with such large drop sizes are suitable only for translocated-type herbicides and crop defoliants, or for applications that target the top surfaces of crops or are applied to the soil surface. For fungicides and insecticides to be effective, the median drop size must be reduced to 200–400 μm; the airborne or drift losses will then increase to 25% or more of the applied active pesticide. *See* AGRICULTURE; AIRPLANE; HELICOPTER. Norman B. Akesson

Bibliography. B. N. Devisetty, D. G. Chasin, and P. D. Berger (eds.), *Pesticide Formulation and Application Systems*, ASTM STP 1146, American Society for Testing and Materials, 1993; C. G. McWhorter and M. R. Gebhardt (eds.), *Methods of Applying Herbicides*, Weed Science Society of America, 1987; H. R. Quantick, *Handbook for Agricultural Pilots*, 4th ed., 1985.

Agricultural buildings

All buildings used on the farm for housing livestock and the farm family and for storing grain, livestock feed, and farm machinery. In the United States, today's farms are larger, more specialized, and more mechanized than those of the past, and these changes are reflected in the buildings. For example, the two-story barns with haylofts are seldom built any more, in favor of one-story buildings without interior columns and with trussed roofs. The building frames are usually constructed of steel or wood and are covered with prepainted metal sheets. This type of construction is low in cost and has a useful life of about 25 years, which results in a nearly simultaneous wearout and obsolescence of the building. Some buildings have an open front facing the south or east. Others are completely enclosed and insulated so that a warm environment may be maintained to protect young animals or to maximize animal growth and production.

Housing of livestock. Livestock farms usually contain one or two species of animals. Buildings are designed for specific uses which depend upon the species and age of animal, climatic conditions, and the management practices followed. Buildings are the shelter portion of the housing system, which also includes a ventilation system and feed, produce, and manure-handling equipment.

Dairy cattle. These animals are housed in two types of barn. Stall barns have individual stalls in which the cows are tied. Feed is conveyed to the manger at the front of the stalls. The manure is mechanically scraped out of the barn. The barn temperature is held above 45°F (7°C) by conserving the body heat given off by the cows through the use of insulation and thermostatic controlling of the rate of ventilation. Milking is performed in the stalls; therefore, the barn and the cows must be kept very clean. The cows are let out of the barn daily for exercise or pasturing.

The second type, the free-stall barn (**Fig. 1**), also provides a separate stall for each animal to rest; however, they are not tied and are free to leave the stall at any time. A mechanical feeder is installed in the feeding area. The manure is collected and transported by mechanical scrapers, by flushing with water, or by passing through slots in the floor. The cows are milked in a separate milking parlor with elevated stalls. The milk is weighed and conveyed through stainless steel pipes into a milk cooler located in the milk room. *See* DAIRY CATTLE PRODUCTION.

Swine. Swine housing changed dramatically between the late 1950s and 1980s. Better facilities were provided to reduce labor and to allow for farrowing of piglets throughout the year. Thus a more uniform supply of pork is available over time.

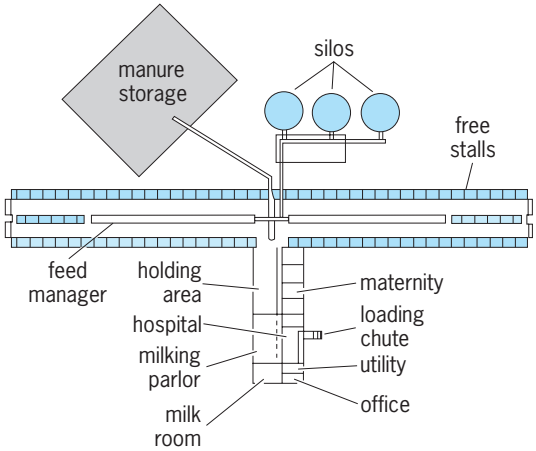

Fig. 1. Typical arrangement of a 200-cow free-stall dairy barn. Cows are held in the holding area before entering the double-8 herringbone milking parlor. (*After Midwest Plan Service, Structures and Environment Handbook, 11th ed., 1983*)

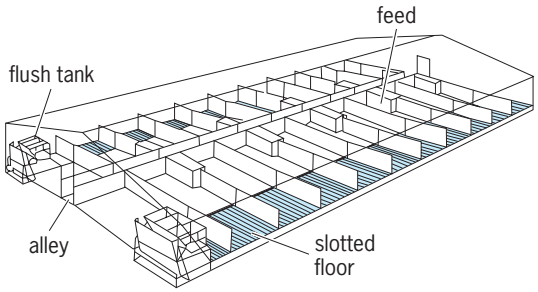

Fig. 2. Environmentally controlled swine growing-finishing house. (*After Midwest Plan Service, Structures and Environment Handbook, 11th ed., 1983*)

There are special buildings for each age group so that an optimum arrangement and thermal environment can be provided for good health and growth. Two types of systems are used to house swine. With the pasture system, minimal buildings are utilized and the pigs are raised on pasture during the growing season. The confinement system confines the pigs to a number of specialized buildings.

Farrowing houses shelter the piglets from birth to about 4 weeks of age. These buildings are heated and well insulated to provide an air temperature of 95°F (35°C) for the piglets and 60°F (16°C) for the sows. After weaning, the pigs are moved to a nursery building for about 8 weeks, where the room temperature is first held at 85°F (29°C) and gradually lowered to 65°F (18°C). A growing house is used next for about 6 weeks until the pigs reach about 150 lb (68 kg) in weight with the room temperature continued at about 65°F (18°C). Lastly, the pigs are moved to a finishing house and kept there until they reach a market weight of about 220 lb (100 kg). Some farmers use a combination of these buildings (for example, nursery plus growing, or growing plus finishing; **Fig. 2**). Additional buildings are required for the gestating sows and boars.

Much labor-saving equipment is used in swine production: conveyors transport the feed from storage to the feeders; automatic waterers supply drinking water; and thermostatically controlled fans regulate the barn temperature. Slotted floors are commonly used for all ages to allow the waste to pass into an underfloor tank where the manure can be easily pumped out into large tank wagons for field spreading. *See* SWINE PRODUCTION.

Beef cattle. These animals require minimal shelter for survival and growth. In the western states, the cattle are kept on the open range or in an open feed lot with only a windbreak for protection. In the midwestern states, more shelter is required because of higher amounts of precipitation during cold weather. The housing consists of an open front shed with access to a yard. The shed contains a layer of bedding and the yard is for exercise and feeding. Mounds of earth are often built in the yards to give well-drained resting areas (**Fig. 3**). A less common type, confinement housing, confines the animals in pens in an open-front building with a slotted or flume-type floor for the handling of manure. Mechanical feeders are commonly used in all kinds of housing systems. *See* BEEF CATTLE PRODUCTION.

Sheep. Sheep require minimum shelter, similar to beef cattle, to minimize lamb losses and increase labor efficiency. An open-front barn with an outdoor exercise yard or a confinement building can be used. Labor-saving equipment is available for feeding, sorting, treating, and shearing animals and for manure disposal. *See* SHEEP.

Poultry. Poultry buildings for egg production usually confine the birds in cages arranged in single to triple decks. Feeding, watering, and egg collection are completely mechanized. The buildings are insulated and ventilated to control the thermal environment. Broiler production also utilizes enclosed buildings where the birds are raised on a floor covered with litter such as straw or wood shavings. *See* POULTRY PRODUCTION.

Other specialized buildings are used to shelter such animals as horses, goats, or rabbits. *See* GOAT PRODUCTION; HORSE PRODUCTION.

Livestock feed and commercial grain. Grass and corn silage is a major feed for dairy and beef cattle. Silos provide an oxygen-deficient environment for the fermentation of grass or corn with a moisture content of 45–65%. The end product of the fermentation process is silage.

Vertical silos are recommended for small and medium-size farms. They consist of cylinders 14–30 ft (4.3–9.1 m) in diameter and up to 90 ft (27.4 m) in height. Silos can be of the conventional type, in which the top surface is exposed to the atmosphere with some spoilage occurring, or the sealed type, with an oxygen-free environment throughout with no spoilage. Mechanical unloaders installed on the top or the bottom of the silo remove the silage and drop it into a conveyor leading to a mechanical feeding system.

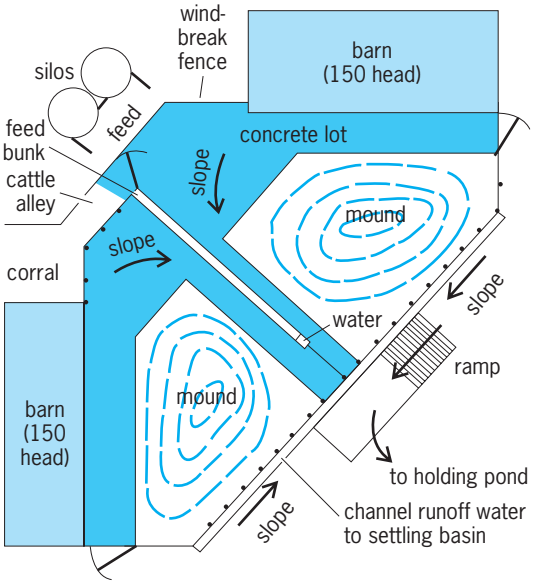

Fig. 3. Beef cattle barn with open front, mechanical bunk feeder, and mounds in the lot. (*After Midwest Plan Service, Structures and Environmental Handbook, 11th ed.,1983*)

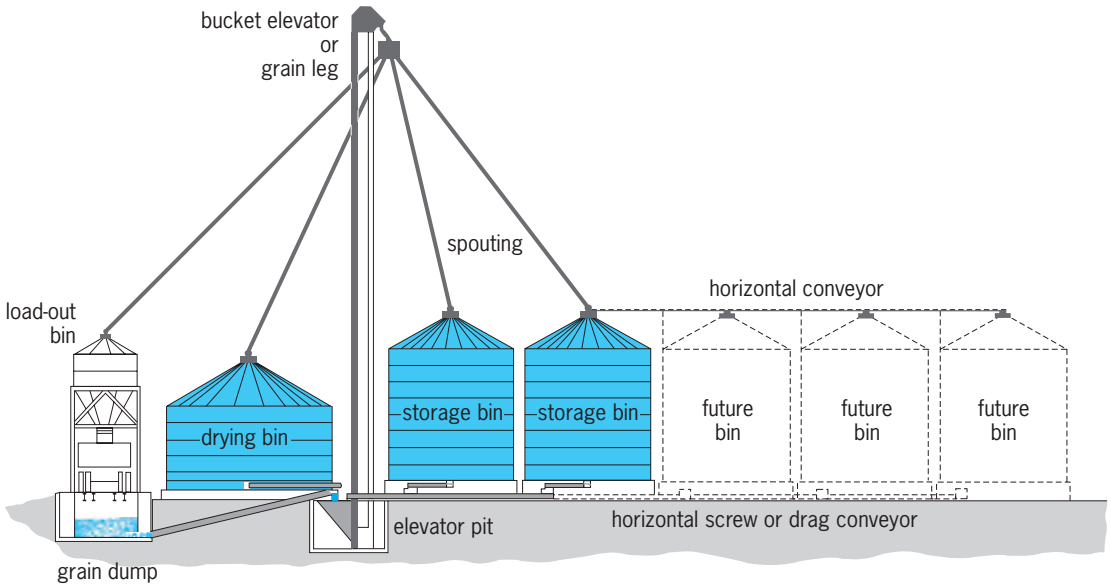

Fig. 4. Farm grain center with drying, conveying, and storage facilities. (*After Midwest Plan Service, Structures and Environment Handbook, 11th ed., 1983*)

Horizontal silos are more economical for large farms and consist of concrete-lined bunkers. A large top surface is exposed and is usually covered with plastic sheet to limit spoilage. The silage is generally removed with front-end loaders on tractors.

Grain fed to livestock need not be dried and is often stored as high-moisture grain in silos. Grain produced for sale must be dried and stored in a dry, insect-free environment to maintain its quality. A grain-handling system consists of a number of components for transporting, drying, and storage. These components include a grain dump to catch the grain discharged from a truck, and conveyors or grain legs to elevate and move it to the dryers and then to storage. Round metal bins and special buildings, called flat storages, are used for grain storage. An aeration system of fans is needed to blow air through the grain and keep it from spoiling. A farm grain center is shown in **Fig. 4**.

Machinery storage. Buildings to house the field machinery are provided on most farms. These buildings must have at least 14-ft-high (4-m) side walls, large doors, and clear span roofs to accommodate the large machines needed on today's farms. A heated machinery repair shop is often located in a portion of this building. *See* AGRICULTURAL MACHINERY.

Farm houses. These homes have all the modern conveniences of urban houses and are constructed of similar materials. The house design includes special features to accommodate the farm activities. Near the rear entrance are facilities to allow the workers to clean up and to store their farm clothing. An office is also desirable for record keeping.

Farmstead planning. The various buildings, fences, and roads composing the farmstead are laid out for efficient use. The farm home, the center of family living, is located to provide a view of the yards. The other buildings are located so as to maintain a high quality of living around the house and efficient operation. The large livestock buildings are at the greatest distance from the house and on its downwind side so as to minimize the effect of objectionable odors and insects. Machinery storage and repair and grain storage buildings are located closer to the house. *See* AGRICULTURE. C. O. Cramer

Bibliography. American Society of Agricultural Engineers, *Agricultural Engineers Yearbook*, 1983–1984; Midwest Plan Service, *Structures and Environment Handbook*, 11th ed., 1983; Midwest Plan Service, *Swine Housing and Equipment Handbook*, 4th ed., 1983; J. H. Whitaker, *Agricultural Buildings and Structures*, 1979.

Agricultural chemistry

The science of chemical compositions and changes involved in the production, protection, and use of crops and livestock. As a basic science, it embraces, in addition to test-tube chemistry, all the life processes through which humans obtain food and fiber for themselves and feed for their animals. As an applied science or technology, it is directed toward control of those processes to increase yields, improve quality, and reduce costs. One important branch of it, chemurgy, is concerned chiefly with utilization of agricultural products as chemical raw materials.

Scope of field. The goals of agricultural chemistry are to expand understanding of the causes and effects of biochemical reactions related to plant and animal growth, to reveal opportunities for controlling those reactions, and to develop chemical products that will provide the desired assistance or control. So rapid has progress been that chemicalization of agriculture has come to be regarded as a twentieth century revolution. Augmenting the benefits of mechanization (a revolution begun in the mid-nineteenth

century and still under way), the chemical revolution has advanced farming much further in its transition from art to science.

Every scientific discipline that contributes to agricultural progress depends in some way on chemistry. Hence agricultural chemistry is not a distinct discipline, but a common thread that ties together genetics, physiology, microbiology, entomology, and numerous other sciences that impinge on agriculture. Chemical techniques help the geneticist to evolve hardier and more productive plant and animal strains; they enable the plant physiologist and animal nutritionist to determine the kinds and amounts of nutrients needed for optimum growth; they permit the soil scientist to determine a soil's ability to provide essential nutrients for the support of crops or livestock, and to prescribe chemical amendments where deficiencies exist. *See* FERTILIZER; SOIL.

Chemical materials developed to assist in the production of food, feed, and fiber include scores of herbicides, insecticides, fungicides, and other pesticides, plant growth regulators, fertilizers, and animal feed supplements. Chief among these groups from the commercial point of view are manufactured fertilizers, synthetic pesticides (including herbicides), and supplements for feeds. The latter include both nutritional supplements (for example, minerals) and medicinal compounds for the prevention or control of disease. *See* HERBICIDE; PESTICIDE.

Important chemicals. Chemical supplements for animal feeds may be added in amounts as small as a few grams or less per ton of feed or in large quantities, for example, urea, which is used as an economical nitrogen for ruminant animals. The tremendous tonnage of processed feeds sold, coupled with the high unit value of some of the chemical supplements, makes this a large market. *See* ANIMAL FEEDS.

Of increasing importance since their commercial introduction have been chemical regulators of plant growth. Besides herbicides (some of which kill plants through overstimulation rather than direct chemical necrosis), the plant growth regulators include chemicals used to thin fruit blossoms, to assist in fruit set, to defoliate plants as an aid to mechanical harvesting, to speed root development on plant cuttings, and to prevent unwanted growth, such as sprouting of potatoes in storage. *See* DEFOLIANT AND DESICCANT.

Striking effects on growth have been observed in plants treated with gibberellins. These compounds, virtually ignored for two decades after they were first isolated from diseased rice in Japan, attracted widespread research attention in the United States in 1956; first significant commercial use began in 1958. The gibberellins are produced commercially by fermentation in a process similar to that used to manufacture penicillin. New product classes include the cytokinins and the polyamines. *See* GIBBERELLIN.

In the perennial battle with insect pests, chemicals that attract or repel insects are increasingly important weapons. Attractants (usually associated with the insect's sexual drive) may be used along with insecticides, attracting pests to poisoned bait to improve pesticidal effectiveness. Often highly specific, they are also useful in insect surveys; they attract specimens to strategically located traps, permitting reliable estimates of the extent and intensity of insect infestations. Additional approaches involve the use of antifeedants, juvenile hormones, and other life-cycle-altering chemicals, and research focused on natural compounds that are more pest-specific and biodegradable. *See* INSECT CONTROL, BIOLOGICAL.

Repellents have proved valuable, especially in the dairy industry. Milk production is increased when cows are protected from the annoyance of biting flies. Repellents also show promise as aids to weight gain in meat animals and as deterrents to the spread of insect-borne disease. If sufficiently selective, they may protect desirable insect species (bees, for instance) by repelling them from insecticide-treated orchards or fields.

Agricultural chemistry as a whole is constantly changing. It becomes more effective as the total store of knowledge is expanded. Synthetic chemicals alone, however, are not likely to solve all the problems humans face in satisfying food and fiber needs. Indeed, many experts are coming to the view that the greatest hope for achieving maximum production and protection of crops and livestock lies in combining the best features of chemical, biological, and cultural approaches to farming, for example, integrated pest management. *See* AGRICULTURAL SCIENCE (ANIMAL); AGRICULTURAL SCIENCE (PLANT); AGRICULTURE.

Rodney N. Hader

Agricultural engineering

A discipline concerned with solving the engineering problems of providing food and fiber for the people of the world. These problems include designing improved tools to work the soil and harvest the crops, as well as developing water supplies for agriculture and systems for irrigating and draining the land where necessary. Agricultural engineers design buildings in which to house animals or store grains. They also work on myriad problems of processing, packaging, transporting, and distributing the food and fiber products.

Agricultural engineering combines the disciplines of mechanical, civil, electrical, and chemical engineering with a basic understanding of biological sciences and agricultural practices. Concern for the properties of agricultural materials and the way they influence engineering designs characterizes agricultural engineers.

Some agricultural engineers work directly with farmers. Most, however, work with the companies that manufacture and supply equipment, feeds, fertilizers, and pesticides. Others work for companies that provide services to farmers, such as developing irrigation and drainage systems or erecting buildings and facilities. Still others work with food-processing companies.

Training. Some engineers, trained in mechanical, civil, electrical, or chemical engineering, have become agricultural engineers through experience in working on agricultural problems. Since 1906, however, special agricultural engineering curricula have been available. The United States has 49 universities offering such curricula; most of those are approved by the Accreditation Board for Engineering and Technology. Several of the universities also offer advanced education at the master of science and doctorate level.

The American Society of Agricultural Engineers, founded in 1907, provides continuing education in the form of local, regional, and national conferences, workshops, and technical meetings. It establishes standards for the profession and publishes technical information.

Applications. A few examples will serve to illustrate the broad general classes of problems with which agricultural engineers work, and the almost infinite variety of special problems they encounter.

Crop production. Soil and water are the basic resources of agriculture, and agricultural engineers design farming systems that include methods of preparing land for cultivation and of controlling water. Land-grading equipment has been designed to follow a beam of laser light. These machines can produce surfaces free of standing pools of water and with slopes so gentle that runoff water will not erode the soil. *See* IRRIGATION (AGRICULTURE); LAND DRAINAGE (AGRICULTURE).

The tillage needed for seedbed preparation depends on the soil type, the crop, and the climate, requiring agricultural engineers to produce a variety of special tools. Chisels deep-rip the subsoil. Bed shapers with rotary tillers produce fine seedbeds for small seeds. Planters with special furrow openers plant directly through the trash from the previous crop for no-till planting. *See* AGRICULTURAL SOIL AND CROP PRACTICES.

Equipment for applying chemicals for insect and weed control must apply the materials uniformly, minimize drift to neighboring regions, and not damage the crop. For weeds which grow taller than the crop, a wick-type applicator has been developed which wipes herbicides onto the weeds. *See* HERBICIDE.

Harvesting equipment is particularly dependent on the nature of the crop. The combined harvester, which cuts and threshes cereal grains, has been adapted to handle corn and fine grass seeds. For horticultural crops, agricultural engineers work closely with plant breeders to find combinations of machine principles and special cultivars that can be harvested mechanically. Processing of tomatoes provides a success story. Plant breeders found varieties for which most of the fruit ripened at one time, and then bred these varieties for high skin strength. Agricultural engineers designed machines to cut and elevate the plants, remove the tomatoes from the vines, sort the red fruit from the green, and deliver the fruit to bulk containers for transport to a cannery. *See* AGRICULTURAL MACHINERY; AGRICULTURAL SCIENCE (PLANT).

Processing. The value of agricultural products depends on their quality and location with respect to time. Wheat is cleaned to eliminate chaff, weed seeds, and lightweight wheat kernels. It is dried for long-term storage and stored with ventilation to remove heat of respiration. Conveyors, hopper-bottom trailers, and barges are used to transport it. It is cracked, rolled, or ground for animal feed; it is milled and packaged for human food. Agricultural engineers design the processes and equipment, the storage structures, and the handling systems. The design requirements are different for each commodity. *See* FOOD ENGINEERING.

Animal production. Growth rate and productivity of animals depend to a large extent on the environment in which they are housed. Air temperature, humidity, speed of air movement, light level, day length, and available floor space are all important. Agricultural engineers control these factors in designs of buildings. *See* AGRICULTURAL BUILDINGS.

Fresh feed and clean water are automatically supplied to laying hens in special cages. Eggs roll out of the cages automatically to collection points. *See* POULTRY PRODUCTION.

In some dairies, data on milk production are stored in computers. Transponders hung from the necks of the cows signal the identity of each cow as it enters the milking parlor. The computer measures an amount of feed based on each cow's production record. The milk is weighed and the computer record is updated—all automatically. Heat recovered from cooling the milk is used to heat water for washing the cows and sterilizing the milking equipment. *See* AGRICULTURAL SCIENCE (ANIMAL); DAIRY CATTLE PRODUCTION.

Waste management. When animals graze the range, their waste material is recycled, and few problems result. When cows are brought to the barn for milking or chickens are raised in confinement, agricultural engineers are faced with the problems of designing equipment, facilities, and processes for managing the concentrated wastes. Processing plants and even field crops produce wastes, and the problems of collecting, treating, storing, and transporting the wastes need specific solutions.

New directions. As agriculture adopts new products, agricultural engineers encounter a new array of production, harvesting, handling, processing, storage, and waste management problems. Greenhouse culture of horticultural and floricultural crops requires structural design, environmental control, irrigation systems, and handling schemes. Forest culture requires equipment for seed collection, seedling culture and transplanting, control of forest weeds and stand thinning, and harvest system design. The culture and harvest of aquatic animals and plants offers new opportunities for agricultural engineers. Processes and systems for controlling water quality are crucial. Feeding, harvesting, and processing equipment are also needed. *See* AQUACULTURE; FLORICULTURE; FOREST AND FORESTRY; HORTICULTURAL CROPS.

Computer technology is rapidly being adopted by

agriculture. The primary bottleneck to fuller utilization of the technology is a shortage of low-cost, reliable sensors to provide data about such factors as soil moisture, plant growth and health, weather, and product quality. Agricultural engineers will need to develop these sensors.

It is reasonable to expect that robots capable of performing some agricultural operations will be important developments of agricultural engineers. The lack of well-controlled conditions and well-defined operations in agriculture suggests, however, that the human operator will remain the key controller in most systems. Agricultural engineers will design equipment and systems more to enhance human capabilities than to replace them. *See* AGRICULTURE; AGRONOMY.

Roger E. Garrett

Bibliography. American Society of Agricultural Engineers, *Agricultural Engineers Yearbook*; S. M. Henderson and R. L. Perry, *Agricultural Process Engineering*, 3d ed., 1976; R. A. Kepner, R. Bainer, and E. L. Barger, *Principles of Farm Machinery*, 3d ed., 1978; Midwest Plan Service, *Structures and Environment Handbook,* 11th ed., 1983; L. O. Roth, F. R. Crow, and G. W. A. Mahoney, *An Introduction to Agricultural Engineering*, 2d ed., 1991; R. E. Stewart, *Seven Decades that Changed America*, American Society of Agricultural Engineers, 1979.

Agricultural machinery

Mechanized systems of food and fiber production used in agriculture. These systems extend from initial tillage of the soil through planting, cultural practices during the growing season, protection from pests, harvesting, conditioning, livestock feeding, and delivery for processing. The trend has been to larger self-propelled special-purpose machines, except for tillage where the trend has been to large four-, six-, or eight-wheel or crawler tractors used in conjunction with high-capacity plows, disks, and deep rippers.

The use of hydraulic power has made possible highly specialized mechanisms to perform intricate operations. Hydraulic power offers the advantages of being easily controlled and automated. Sophisticated technology is used to increase the precision needed in modern agriculture: lasers for laying out fields for surface irrigation systems; microprocessors for sensing and controlling intricate operations, such as controlling feed mixtures for dairy cows and grading fruits and vegetables; and electronic devices in the automation of many harvesters.

Tillage equipment. Primary and secondary tillage equipment, such as plows, disks, and harrows, are designed to prepare the seedbed and root zones for crop production. Multipurpose machines are used where a high degree of precision and specialization is needed. **Figure 1** shows a machine that may be used to simultaneously rototill the soil, form beds and irrigation furrows, incorporate a herbicide, and plant the seed in one tripacross the field.

Laser land-leveling machinery allows the farmer to prepare a field for surface irrigation by grading the field to an exact slope. A laser beam, transmitted from a rotating command post at a preset slope, is received on a mobile scraper that generates a signal to automatically keep the scraper blade on the desired slope. These techniques allow land to be leveled with greater accuracy than can be achieved by traditional methods. Irrigation water savings of 20–30% have been reported after laser land leveling. *See* IRRIGATION (AGRICULTURE).

Planting and cultivation equipment. Agricultural planting equipment is commonly used for the planting of raw uncoated seed. Some of the more expensive or irregularly shaped seeds can be coated to form a uniformly shaped object that can be handled and planted with extreme accuracy and efficiency. For some crops, other types of planting equipment may be used to plant tubers, such as potatoes, and transplants, such as tomatoes.

No-till planters have been developed for the planting of seed into undisturbed soil; no-till farming eliminates all primary and secondary tillage normally conducted between cropping years, and therefore the soil remains undisturbed. Conditions for no-till planting are not at all the same as for conventional tillage farming in that the soil is covered with residue or sod

Fig. 1. Multipurpose machine used for precision tillage, planting, bed shaping, and fertilizing in one pass, and for later bed shaping, cultivating, and fertilizing. (*Johnson Farm Machinery Co.*)

Fig. 2. Agricultural aircraft applying insecticide on a field.

Fig. 3. Portable feed grinder-mixer delivering processed feed to feed bunk. (*Sperry New Holland, Division of Sperry Rand Corp.*)

and is likely wetter, firmer, and rougher than conventionally prepared soils. Generally, a no-till planter is built heavier and stronger than conventional planters and is normally equipped with a cutting disk in front of each planting unit. The no-till farming concept is widely used in the Corn Belt states and generally benefits farmers in terms of soil and water conservation, lower production costs, and greater production efficiencies.

Crop cultivation is accomplished primarily to rid the crop of competing weeds. Cultivation equipment is normally quite simple and is designed to cut, slice, bury, or rip out weeds.

Chemical application equipment. Crop chemicals, such as fertilizers and pesticides, are routinely used in agricultural production. Application equipment for crop chemicals should apply a designated chemical in the proper amount, at the correct time, and to the correct target. Liquid pesticides are commonly applied with high- or low-pressure sprayers mounted on tractors, trailers, trucks, and aircraft. Air-blast or orchard-type sprayers utilize a high-speed air stream to carry the liquid chemical to the surface being treated.

Dry chemicals, such as fertilizers, are applied with either broadcast or band-type applicators. Broadcast applicators spread the chemical over the entire field surface at a prescribed rate, while band-type applicators apply the chemical at a prescribed rate in strips or bands across the field. *See* FERTILIZING.

Fig. 4. Hay cuber compressing alfalfa hay into 1.6-in.-square (4-cm-square) cubes. (*University of California*)

Fig. 5. Self-propelled combined harvester-thresher. (*Sperry New Holland, Division of Sperry Rand Corp.*)

The use of aircraft has revolutionized many farming operations. Rice is sown in flooded fields, and fertilizers and pesticides are applied from the air (**Fig. 2**). Herbicides and insecticides are also applied by air on other crops, such as grains, vegetables, and tree fruits. *See* AGRICULTURAL AIRCRAFT; HERBICIDE; INSECTICIDE.

Farmstead machinery and equipment. Mechanized operations around farmsteads vary from very little on subsistence-type farms to nearly complete on larger commercial farms. These operations are for either crop conditioning or materials handling. The equipment is powered by electric motors if it is stationary, and by tractor hydraulics or power takeoff shafts if it is portable (**Fig. 3**). Stationary equipment includes conveyors and grinders, and portable equipment includes feed wagons, mixers, and manure spreaders.

Hay, grain, and cotton harvesting equipment. Self-propelled hay cubers compress hay into small cubes to make it easy to handle with conveyors (**Fig. 4**). Practically all small grains are harvested by self-propelled combines (**Fig. 5**) that cut the crop and deliver the grain to a truck. While some corn for cattle and dairy cows is chopped for silage (stalk and grain), most of it is harvested as grain by harvesters similar to combines but fitted with harvester heads that strip the ears from six or more rows at a time and shell them. Cotton harvesting is done by stripping the fiber-containing bolls from the plant or by rotating spindles that collect the fibers with a twisting action and release them in an airstream that carries them into a hopper.

Fruit and vegetable harvesting. The biggest demand for hand labor has been in the harvest of fruits and vegetables. Much research has been done on breeding of plants and on machine development to reduce hand labor, primarily during harvesting, since that is the time of peak labor demand. Most successful attempts have utilized shaking of trees or vines to remove fruits. For some vegetables, such as cabbage and celery, the head or stalk is cut off by disks or knives. Most of these machines are nonselective, that is, all of the crop is harvested in one operation whether it is mature or not. Selection for color, maturity, and size is usually done by people on machines or at grading stations. Electronics are used for color maturity sorting on tomato harvesters (**Fig. 6**); each tomato passing through the harvester is inspected individually at a high rate of speed by an optical sensor. The optical sensor determines whether the tomato is red (mature) or green (not mature); if the tomato is not mature a mechanical system is triggered that removes it from the harvester.

There are several selective harvesters in use, but on a very limited basis. For example, an asparagus harvester has been developed that selects only spears of sufficient length, a cucumber harvester removes only the mature fruits, and a lettuce harvester electronically selects and removes those heads that are

Fig. 6. This tomato harvester cuts the vines, removes the fruit, and then loads it into a truck after the rejected fruit has been removed by the sorters. (*Blackwelder Mfg. Co.*)

Fig. 7. Pea pods being harvested and threshed by a self-propelled pea viner machine. (*FMC Co.*)

Fig. 8. These grapes are being harvested by beaters which impact the vines. The berries and bunches are collected on conveyors and then delivered to a dump trailer. (*Up-Right Inc.*)

large or dense enough. Harvesters for root crops such as potatoes, carrots, turnips, beets, and parsnips cut the tops and then perform a lifting action with a blade. This is followed by an open-chain conveyor that retains the roots while sifting out the soil. Sizing on the machine is usually done mechanically, while grading for quality is done by individuals who ride on the machines. Some of these harvesters have color sorters that remove some rejects as well as clods. While dried beans have been mechanically harvested for many years, it was only in the late 1970s that a green bean havester was developed. Harvesters for canning peas (**Fig. 7**) and peas for freezing are also available.

Some fruits and vegetables for the fresh produce market, such as citrus fruits, melons, grapes, eggplant, broccoli, and cauliflower, are still harvested by hand.

Machinery development. The increased mechanization in agriculture has resulted from economic and social factors. Economic forces have lowered profits per unit of production. In the free-market situation,

the only way for a farmer to improve total income has been to produce more units. This has led to larger farms and fewer farmers. The uncertainty of available seasonal or migrant labor for performing farm tasks has also speeded the farmers' steady move to mechanization.

Precision has received maximum emphasis in the development of new machinery. Plant scientists and design engineers have worked together to produce plant–machine compatibility in production and harvesting. For example, grain breeders have produced hybrid varieties that have their heads at uniform height and are easily threshed, and grape producers have trained the vines so that mechanical harvesters can clip or shake off the berries and retrieve them on conveyors (**Fig. 8**). *See* AGRICULTURAL SOIL AND CROP PRACTICES; AGRICULTURE; DAIRY MACHINERY.

James W. Rumsey; Michael O'Brien

Bibliography. D. Hunt, *Farm Power and Machinery Management*, 9th ed., 1995; R. A. Kepner et al., *Principles of Farm Machinery*, 1978; M. O'Brien et al., *Principles and Practices in Harvesting and Handling Fruits and Nuts*, 1980.

Agricultural meteorology

A branch of meteorology that examines the effects and impacts of weather and climate on crops, rangeland, livestock, and various agricultural operations. The branch of agricultural meteorology dealing with atmospheric-biospheric processes occurring at small spatial scales and over relatively short time periods is known as micrometeorology, sometimes called crop micrometeorology for managed vegetative ecosystems and animal biometeorology for livestock operations. The branch that studies the processes and impacts of climatic factors over larger time and spatial scales is often referred to as agricultural climatology. *See* CLIMATOLOGY; MICROMETEOROLOGY.

Agricultural meteorology, or agrometeorology, addresses topics that often require an understanding of biological, physical, and social sciences. It studies processes that occur from the soil depths where the deepest plant roots grow to the atmospheric levels where seeds, spores, pollen, and insects may be found. Agricultural meteorologists characteristically interact with scientists from many disciplines.

Role of meteorologists. Agricultural meteorologists collect and interpret weather and climate data needed to understand the interactions between vegetation and animals and their atmospheric environments. The climatic information developed by agricultural meteorologists is valuable in making proper decisions for managing resources consumed by agriculture, for optimizing agricultural production, and for adopting farming practices to minimize any adverse effects of agriculture on the environment. Such information is vital to ensure the economic and environmental sustainability of agriculture now and in the future. *See* WEATHER OBSERVATIONS.

Agricultural meteorologists also quantify, evaluate, and provide information on the impact and consequences of climate variability and change on agriculture. Increasingly, agricultural meteorologists assist policy makers in developing strategies to deal with climatic events such as floods, hail, or droughts and climatic changes such as global warming and climate variability.

Agricultural meteorologists are involved in many aspects of agriculture, ranging from the production of agronomic and horticultural crops, trees, and livestock to the final delivery of agricultural products to market. They study the energy and mass exchange processes of heat, carbon dioxide, water vapor, and trace gases such as methane, nitrous oxide, and ammonia, within the biosphere on spatial scales ranging from a leaf to a watershed and even to a continent. They study, for example, the photosynthesis, productivity, and water use of individual leaves, whole plants, and fields. They also examine climatic processes at time scales ranging from less than a second to more than a decade.

Crop micrometeorology. This branch of agricultural meteorology deals with the study of mass, radiation, and energy exchanges between plants and their surrounding physical environment. Fundamental is an understanding of the physical laws that govern the interactions between vegetation and radiation in the various wavelength regions of the solar spectrum—including ultraviolet radiation, photosynthetically active radiation, and near-infrared radiation, and thermal infrared radiation. Likewise, the laws that govern the conduction and convection of heat to plants, soil, and the air along with the latent heat exchange and the turbulent transport of water vapor and carbon dioxide between the plants and the atmosphere are central to crop micrometeorology.

Crop micrometeorologists, for example, cooperate with other scientists to change plant and canopy architecture and improve management practices to optimize radiation use efficiency, defined as the amount of plant material produced per unit of solar radiation intercepted; and to increase water use efficiency, defined as the amount of plant material produced per unit of water used. They often work with agronomists and biological engineers to develop methods to promote the efficient use of irrigation and to enhance the capture and storage of precipitation. They also assist in the study and development of practices such as windbreaks and frost protection schemes to modify the plant environment to benefit crop production.

Crop ecology and phenology. The specific types of crops and crop cultivars that survive and grow well in a given area are determined primarily by climatic factors such as daylength, precipitation during the growing season, maximum and minimum temperatures, and the length of the frost-free period. Agricultural meteorologists develop an understanding of the relationships between the climate and the growth, yield, and phenological (growth stage) development of crops. They then use such relationships to evaluate areas for sustaining the growth and survival of specific crop types.

Relationships between the accumulation of thermal time (also referred to as heat units or growing degree days) and the phenological stages of various crops have been developed. These relationships provide information valuable for planning of critical farm operations during the proper growth stage. For example, by using thermal time to predict the period of corn pollination, a farmer might irrigate to provide adequate soil water to prevent water stress during this critical growth stage.

Crop production and growing season. Crop productivity is affected by the weather. The weather or climate of an area determines the crop growing season, defined as the period of time when climatic conditions are favorable for plant growth. Typically the growing season is defined as the length of time between the first and last killing frost in an area, but in some locations the length of the growing season is set by the period when adequate water is available to the crops. During the growing season the most important variables affecting crop production are precipitation, temperature, and solar radiation. Wind and the carbon dioxide concentration in the air are also important. Sometimes it is possible to adopt practices that will alter the microclimate to maintain a more favorable atmospheric environment for crop production. Agricultural meteorologists have helped develop practices such as irrigation and the establishment of windbreaks, which modify and improve the crop microclimate.

Important climatic factors that characterize the growth environment for crops are the length of the growing season; the amount, timing, and intensity of rainfall; and the patterns of plant water use (evapotranspiration). Agricultural meteorologists often collaborate with agronomists, horticulturists, and agricultural economists to evaluate the effects of changing and variable climatic patterns on crop production and the economic value of crops. For example, the estimated levels of crop production at various growth stages made by agricultural meteorologists are often used to plan worldwide food reserves and to estimate future commodity prices. Such information can be used to influence crop planting and production patterns. *See* AGRICULTURAL SOIL AND CROP PRACTICES.

Climate change and agricultural production. Any change in climate will affect agricultural production, either beneficially or deleteriously. Because agricultural meteorologists study the interactions between climatic factors and crop growth and yield, they are able to make a major contribution to understanding the impact of potential climate change on agriculture. For several decades the concentration of carbon dioxide and other radiatively active gases has been increasing, due primarily to the combustion of fossil fuels and secondarily to the clearing of forests for agricultural purposes. Many climatologists believe that these increasing concentrations of carbon dioxide, methane, nitrous oxide, and the chloro- fluorohydrocarbons and other radiatively active trace gases are increasing the global temperature and altering regional precipitation patterns. There is considerable disagreement among scientists about the magnitude of the change and the impact that these changes will have on agriculture. Many believe that the increasing concentration of carbon dioxide will increase crop production and improve the water use efficiency of crops. Changing temperature patterns may cause a shift in areas where certain crops are grown. Agricultural meteorologists face numerous challenges in assembling and interpreting the facts about climatic change to assess the future real impacts on agriculture and to help agriculturists adapt to climate change. *See* CLIMATE HISTORY.

Water use efficiency. Agricultural meteorologists conduct research aimed at improving the water use efficiency of crops. Throughout much of the world, crop productivity is limited by inadequate precipitation. Irrigation has been widely used around the world to supplement natural rainfall. However, crops require large amounts of water, and in many areas agriculture has competed for increasingly limited water resources. Agricultural meteorologists play an important role in ensuring that water is used as efficiently as possible by agricultural producers. They work with plant breeders to select crops that improve plant water use efficiency or, in other words, that are more conservative in the use of water. They develop methods for providing reliable and accurate estimates of the amount of water consumed by crops. This information is needed to determine when to schedule irrigation and how much water to apply. They also assist in the development of other practices to promote more efficient use of water. These practices include selection of optimum row widths, use of mulches, and adoption of minimum tillage methods. *See* IRRIGATION (AGRICULTURE).

Plant and animal diseases and pests. The total productivity of agriculture is often considerably reduced by plant and animal diseases and pests. Outbreaks of diseases and insect infestations are controlled in large measure by climatic variables, the most important of which are wind, temperature, humidity, and dew. In some cases, insects and disease-causing organisms are carried long distances by the wind, so synoptic weather patterns are studied to predict where disease or pest outbreaks are expected.

In general, there is sufficient inoculum for plant pathogens and adequate numbers of harmful insects always present in the agricultural environment to cause severe economically damaging outbreaks if favorable microclimatic conditions occur. Agricultural meteorologists work with plant and animal pathologists and entomologists to construct mathematical models to describe the dynamic processes of disease and pest outbreaks. They develop models that describe the conditions favorable for the development of beneficial organisms useful in the biological control of some pests and diseases. Accurate prediction of disease and pest outbreaks facilitates the ability of agriculturists to control diseases and pests through proper use of cultural practices, pesticides, and biological control methods. Teams of agricultural meteorologists, entomologists, and plant pathologists will increasingly find ways to reduce the destructive

influence of diseases and insects on agriculture, while reducing pesticide use. *See* ENTOMOLOGY, ECONOMIC; INSECT CONTROL, BIOLOGICAL; PLANT PATHOLOGY.

Pesticide application. In addition to predicting the outbreaks of diseases and insects, agricultural meteorologists play an important role in defining the environmental conditions for the application of pesticides for maximum benefit. They also describe the environmental conditions when pesticides should or should not be applied. Information on the effects of wind and temperature gradients on the airborne drift of a pesticide after it is released from an aircraft is required for its safe aerial application. Agricultural meteorologists in cooperation with weed scientists, entomologists, aircraft operators, and pesticide manufacturers help define conditions for effective pesticide application. *See* AGRICULTURAL AIRCRAFT.

Agricultural meteorologists help to determine the fate of pesticides and fertilizers. Many chemicals break down in the soil or are taken up by plants and do not harm in any way the environment. However, some chemicals are carried to the ground water or into water bodies, and others volatize and are carried into the atmosphere by wind and turbulence. The fate of many agricultural chemicals is dependent on climatic factors. For example, excessive rains can carry chemicals from the soil into streams, rivers, ponds, lakes, oceans, or ground water. Winds and high temperature may act on agricultural chemicals to cause air-pollution problems. By defining clearly the role that climate plays on the fate of chemicals, agricultural meteorologists, working with other scientists, can assist in developing appropriate management strategies to reduce water and air pollution. *See* AIR POLLUTION; FERTILIZER; GROUND-WATER HYDROLOGY; PESTICIDE; WATER POLLUTION.

Frost protection. Production of some agronomic and horticultural crops may be severely limited by frost, when frost comes before the crops are ready to be harvested. Protecting agronomic crops from frost damage can rarely be economically justified. However, agricultural meteorologists can define optimal planting dates for a given site that will reduce the chance of frost damage.

By determining weather conditions that cause frost, agricultural meteorologists can suggest the best frost protection methods for a particular type of frost and discourage the use of the least successful methods. Effective forecasting of frost events and the use of effective methods for altering the microclimate that reduce the occurrence or intensity of a frost can help prevent millions of dollars in crop losses. For high-value crops such as vegetables and fruits, especially citrus fruits, agricultural meteorologists play important roles in predicting the occurrence of frost and in developing methods that protect against frost damage. The best method of frost protection is proper site selection. Agricultural meteorologists can help in selecting the most appropriate site for growing high-value crops by defining the microclimate of certain areas and evaluating the effects of large bodies of water and topographic position on these microclimates. They can determine the appropriate sites on hillsides for orchards to take advantage of a warm layer of air that often exists on hillsides. They can also determine sites where cold air pools as it drains to lower elevations. *See* FROST.

Wind damage and erosion. Wind sometimes causes mechanical damage to trees and crops, and often contributes to an environment that promotes plant stress. Windbreaks are an effective way to reduce mechanical damage by wind and to provide a modified microclimate that reduces plant water use, improves plant water use efficiency, and increases crop productivity. Agricultural meteorologists play an important role in designing windbreaks that are most effective in bringing about the desired microclimate modification. They help to determine and quantify the climatic parameters that have been modified.

Wind causes soil erosion, which results in very serious damage and depletes a vital natural agricultural resource. Working with soil scientists, agricultural meteorologists develop equations to describe the amount of soil movement under given wind and surface conditions. They also suggest approaches to reduce erosion, such as roughing the soil surface, leaving crop residue on the surface, and constructing windbreaks. *See* SOIL CONSERVATION.

Simulation models. Since the late 1970s, agricultural meteorologists have assisted in the development of mathematical models of crop and livestock behavior based on various climatological and physiological parameters. The earliest models were statistical and were developed by correlating climatic data with yields of various crops or with livestock performance. These empirical models generally included temperature and precipitation as the most important parameters. They work reasonably well, except when the weather deviates significantly from normal. Of course, estimates of the deviation of yields are most valuable for those years with abnormal temperature and precipitation patterns. Since the advent of computers, more mechanistic models have been developed. Mechanistic models are based on sound physical and physiological theories, and they perform quite satisfactorily. Because these models account for changing weather during a growing season, they can help improve management practices for crop and livestock production. Expert systems and artificial intelligence are useful for suggesting management practices that achieve increased crop and animal production. *See* ARTIFICIAL INTELLIGENCE; CLIMATE MODELING; EXPERT SYSTEMS; SIMULATION.

Livestock production. The performance of livestock is influenced by weather and climate. In many cases, the microclimate has been modified and controlled by confining animals to buildings. Certain agricultural meteorologists, known as biometeorologists, work with animal scientists to design buildings with an environment favorable for efficient feed utilization and improved production and efficiency. They also develop indices to evaluate the degree of stress that animals experience under conditions of high heat and humidity. With the use

of weather forecasts, these indices alert livestock producers to initiate stress-relieving procedures and reduce production losses by avoiding shipment of animals under these conditions. Agricultural meteorologists also study the effects of cold stress on animals in confined environments. Results may suggest, for example, the need to establish windbreaks for animals to use for shelter to reduce cold stress in unconfined environments. *See* AGRICULTURAL SCIENCE (ANIMAL).

Developing strategies and policy. Agricultural meteorologists work with governmental officials and other policy makers to develop, formulate, and adopt reasonable strategies and policies for dealing with extreme climatic events such as drought and floods. Agricultural meteorologists, working with social scientists, also determine the likely impacts of extreme climatic events and climate change on agriculture and all elements of society and identify rational response options to mitigate those impacts. *See* METEOROLOGY. Blaine L. Blad

Bibliography. B. J. Barfield and J. F. Gerber (eds.), *Modification of the Aerial Environment of Crops*, 1979; G. S. Campbell and J. M. Norman, *An Introduction to Environmental Biophysics*, 1988; D. M. Gates, *Biophysical Ecology*, 1980; I. F. Griffiths (ed.), *Handbook of Agricultural Meteorology*, 1994; W. P. Lowry, *Atmospheric Ecology for Designers and Planners*, 1988; J. L. Monteith (ed.), *Vegetation and the Atmosphere*, vols. 1 and 2, 1977; N. J. Rosenberg, B. L. Blad, and S. B. Verma, *Microclimate: The Biological Environment*, 1983; D. A. Wilhite (ed.), *Drought Assessment, Management and Planning: Theory and Case Studies*, 1993.

Agricultural science (animal)

The science which deals with the selection, breeding, nutrition, and management of domestic animals for economical production of meat, milk, eggs, wool, hides, and other animal products. Horses for draft and pleasure, dogs and cats for pets, rabbits for meat production, and bees for honey production may also be included in this group. *See* BEEF CATTLE PRODUCTION; BEEKEEPING; DAIRY CATTLE PRODUCTION; GOAT PRODUCTION; HORSE PRODUCTION; MULE PRODUCTION; POULTRY PRODUCTION; SHEEP; SWINE PRODUCTION.

When primitive societies first domesticated animals, they were kept as means of meeting the immediate needs for food, transportation, and clothing. Sheep probably were the first and most useful animals to be domesticated, furnishing milk and meat for food, and hides and wool for clothing.

As chemistry, physiology, anatomy, genetics, nutrition, parasitology, pathology, and other sciences developed, their principles were applied to the field of animal science. Since the beginning of the twentieth century, great strides have been made in livestock production. Today, farm animals fill a highly important place in human existence. They convert raw materials, such as pasture herbage, which are of little use to people as food, into animal products having nutritional values not directly available in plant products.

Ruminant animals (those with four compartments or stomachs in the fore portion of their digestive tract, such as cattle and sheep) have the ability to consume large quantities of roughages because of their particular type of digestive system. They also consume large tonnages of grains, as well as mill feeds, oil seed meals, industrial and agricultural byproducts, and other materials not suitable for human food.

Products of the animal industry furnish raw materials for many important processing industries, such as meat packing, dairy manufacturing, poultry processing, textile production, and tanning. Many services are based on the needs of the animal industry, including livestock marketing, milk deliveries, poultry and egg marketing, poultry hatcheries, artificial insemination services, feed manufacturing, pharmaceutical industry, and veterinary services. Thus, animal science involves the application of scientific principles to all phases of animal production, furnishing animal products efficiently and abundantly to consumers. Products from animals are often used for consumer products other than food, for example, hides for leather, and organ meats for preparation of drugs and hormones.

Livestock breeding. The breeding of animals began thousands of years ago. During the last half of the nineteenth century, livestock breeders made increasing progress in producing animals better suited to the needs of humans by simply mating the best to the best. However, in the twentieth century animal breeders began to apply the scientific principles of genetics and reproductive physiology. Some progress made in the improvement of farm animals resulted from selected matings based on knowledge of body type or conformation. This method of selection became confusing due to the use of subjective standards which were not always related to economic traits. This error was corrected, however, and many breeders of dairy cattle, poultry, beef cattle, sheep, and swine in the mid-1970s made use of production records or records of performance. Some of their breeding plans were based on milk fat production or egg production, as well as on body type or conformation. The keeping of poultry and dairy cow production records began in a very limited way late in the nineteenth century. The first Cow-Testing Association in the United States was organized in Michigan in 1906. Now over 1,500,000 cows are tested regularly in the United States. *See* BREEDING (ANIMAL).

Many states have production testing for beef cattle, sheep, and swine, in which records of rate of gain, efficiency of feed utilization, incidence of twinning, yield of economically important carcass cuts, and other characteristics of production are maintained on part or all of the herd or flock. These records serve as valuable information in the selection of animals for breeding or sale.

Breeding terminology. A breed is a group of animals that has a common origin and possesses

characteristics that are not common to other individuals of the same species.

A purebred breed is a group that possesses certain fixed characteristics, such as color or markings, which are transmitted to the offspring. A record, or pedigree, is kept which describes their ancestry for five generations. Associations were formed by breeders primarily to keep records, or registry books, of individual animals of the various breeds. Purebred associations take active roles in promoting and improving the breeds.

A purebred is one that has a pedigree recorded in a breed association or is eligible for registry by such an association. A crossbred is an individual produced by utilizing two or more purebred lines in a breeding program. A grade is an individual having one parent, usually the sire, a purebred and the other parent a grade or scrub. A scrub is an inferior animal of nondescript breeding. A hybrid is one produced by crossing parents that are genetically pure for different specific characteristics. The mule is an example of a hybrid animal produced by crossing two different species, the American jack, *Equus asinus*, with a mare, *E. caballus*.

Systems of breeding. The modern animal breeder has genetic tools such as selection and breeding, and inbreeding and outbreeding. Selection involves the retaining or rejecting of a particular animal for breeding purposes, based largely on qualitative characteristics. Inbreeding is a system of breeding related animals. Outbreeding is a system of breeding unrelated animals. When these unrelated animals are of different breeds, the term crossbreeding is usually applied. Crossbreeding is in common use by commercial swine producers. About 80–90% of the hogs produced in the Corn Belt states are crossbred. Crossbreeding is also used extensively by commercial beef and sheep producers.

Grading-up is the process of breeding purebred sires of a given breed to grade females and their female offspring for generation after generation. Grading-up offers the possibility of transforming a nondescript population into one resembling the purebred sires used in the process. It is an expedient and economical way of improving large numbers of animals.

Formation of new breeds. New breeds of farm animals have been developed from crossbred foundation animals. Montadale, Columbia, and Targhee are examples of sheep breeds so developed. The Santa Gertrudis breed of beef cattle was produced by crossing Brahman and Shorthorn breeds on the King Ranch in Texas. In poultry, advantage has been taken of superior genetic ability through the development of hybrid lines. *See* GENETICS.

Selection of animals for breeding. The evaluation or selection of animals for breeding purposes is of importance to the commercial as well as to the purebred breeder. In selecting animals for breeding, desirable conformation or body type is given careful attention. The animals are also examined carefully for visible physical defects, such as blindness, crooked legs, jaw distortions, and abnormal udders. Animals known to be carriers of genes for heritable defects, such as dwarfism in cattle, should be discriminated against.

When they are available, records of performance or production should be considered in the selection of breeding animals. Some purebred livestock record associations now record production performance of individual animals on pedigrees.

Genetic engineering is a new frontier in animal breeding. For example, mature eggs are collected and fertilized in test tubes, and then a fertilized egg is placed in a uterus where it develops into a fetus and eventually an offspring. There are other techniques, largely experimental, that may be practical, such as embryo splitting for the production of identical twins. Identical twins are exceedingly important to the researcher, since experiments can be conducted without involving a large number of animals. They are also important when saving rare and endangered species. Another genetic engineering technique is fusing embryos of different species, and this method can also be used to save the genetic material of endangered species. *See* GENETIC ENGINEERING.

Artificial insemination. In this process spermatozoa are collected from the male and deposited in the female genitalia by instruments rather than by natural service. In the United States this practice was first used for breeding horses. Artificial insemination in dairy cattle was first begun on a large scale in New Jersey in 1938. Freezing techniques for preserving and storing spermatozoa have been applied with great success to bull semen, and it is now possible for outstanding bulls to sire calves years after the bulls have died. The use of artificial insemination for beef cattle and poultry (turkeys) is common practice. There are also drugs which stimulate beef cow herds to come into heat (ovulate) at approximately the same time. This permits the insemination of large cow herds without the individual handling and inspection which is used with dairy cattle. Although some horses are bred by artificial insemination, many horse breed associations allow only natural breeding.

Livestock feeding. Scientific livestock feeding involves the systematic application of the principles of animal nutrition to the feeding of farm animals. The science of animal nutrition has advanced rapidly since 1930, and the discoveries are being utilized by most of those concerned with the feeding of livestock. The nutritional needs and responses of the different farm animals vary according to the functions they perform and to differences in the anatomy and physiology of their digestive systems. Likewise, feedstuffs vary in usefulness depending upon the time and method of harvesting the crop, the methods employed in drying, preserving, or processing, and the forms in which they are offered to the animals. *See* ANIMAL FEEDS.

Chemical composition of feedstuffs. The various chemical compounds that are contained in animal feeds are divided into groups called nutrients. These include proteins, fats, carbohydrates, vitamins, and mineral matter. Proteins are made up of amino acids. Twelve amino acids are essential for all nonruminant animals and must be supplied in their diets. Fats and

carbohydrates provide mainly energy. In most cases they are interchangeable as energy sources for farm animals. Fats furnish 2.25 times as much energy per pound as do carbohydrates because of their higher proportion of carbon and hydrogen to oxygen. Thus the energy concentration in poultry and swine diets can be increased by inclusion of considerable portions of fat. Ruminants cannot tolerate large quantities of fat in their diets, however. *See* CARBOHYDRATE; LIPID; PROTEIN.

Vitamins essential for health and growth include fat-soluble A, D, E, and K, and water-soluble vitamins thiamine, riboflavin, niacin, pyrodoxine, pantothenic acid, and cobalamin. *See* VITAMIN.

Mineral salts that supply calcium, phosphorus, sodium, chlorine, and iron are often needed as supplements, and those containing iodine and cobalt may be required in certain deficient areas. Zinc may also be needed in some swine rations. Many conditions of mineral deficiency have been noted by using rations that were not necessarily deficient in a particular mineral but in which the mineral was unavailable to the animal because of other factors in the ration or imbalances with other minerals. For example, copper deficiency can be caused by excess molybdenum in the diet.

By a system known as the proximate analysis, developed prior to 1895 in Germany, feeds have long been divided into six fractions including moisture, ether extract, crude fiber, crude protein, ash, and nitrogen-free extract. The first five fractions are determined in the laboratory. The nitrogen-free extract is what remains after the percentage sum of these five has been subtracted from 100%. Although proximate analysis serves as a guide in the classification, evaluation, and use of feeds, it gives very little specific information about particular chemical compounds in the feed.

The ether extract fraction includes true fats and certain plant pigments, many of which are of little nutritional value.

The crude fiber fraction is made up of celluloses and lignin. This fraction, together with the nitrogen-free extract, makes up the total carbohydrate content of a feed. *See* CELLULOSE.

The crude protein is estimated by multiplying the total Kjeldahl nitrogen content of the feed by the factor 6.25. This nitrogen includes many forms of nonprotein as well as protein nitrogen. *See* NITROGEN.

The ash, or mineral matter fraction, is determined by burning a sample and weighing the residue. In addition to calcium and other essential mineral elements, it includes silicon and other nonessential elements.

The nitrogen-free extract (NFE) includes the more soluble and the more digestible carbohydrates, such as sugars, starches, and hemicelluloses. Unfortunately, most of the lignin, which is not digestible, is included in this fraction. *See* HEMICELLULOSE.

A much better system of analysis was developed at U.S. Department of Agriculture laboratories for the crude fiber fraction of feedstuffs. This system separates more specifically the highly digestible cell-soluble portion and the less digestible fibrous portion of plant cell walls.

Digestibility of feeds. In addition to their chemical composition or nutrient content, the nutritionist and livestock feeder should know the availability or digestibility of the different nutrients in feeds. The digestibility of a feed is measured by determining the quantities of nutrients eaten by an animal over a period of time and those recoverable in the fecal matter. By assigning appropriate energy values to the nutrients, total digestible nutrients (TDN) may be calculated. These values have been determined and recorded for a large number of feeds.

Formulation of animal feeds. The nutritionist and livestock feeder finds TDN values of great use in the formulation of animal feeds. The TDN requirements for various classes of livestock have been calculated for maintenance and for various productive capacities. However, systems have been developed for expressing energy requirements of animals or energy values of feeds in units which are more closely related to the body process being supported (such as maintenance, growth, and milk or egg production). Tables of feeding standards have been published using the units of metabolizable energy and net energy, which are measurements of energy available for essential body processes. Recommended allowances of nutrients for all species of livestock and some small animals (rabbit, dogs, and mink) are published by the National Academy of Sciences–National Research Council.

Nutritional needs of different animals. The nutritional requirements of different classes of animals are partially dependent on the anatomy and physiology of their digestive systems. Ruminants can digest large amounts of roughages, whereas hogs and poultry, with simple stomachs, can digest only limited amounts and require more concentrated feeds, such as cereal grains. In simple-stomached animals the complex carbohydrate starch is broken down to simple sugars which are absorbed into the blood and utilized by the body for energy.

Microorganisms found in the rumen of ruminant animals break down not only starch but the fibrous carbohydrates of roughages, namely, cellulose and hemicellulose, to organic acids which are absorbed into the blood and utilized as energy. Animals with simple stomachs require high-quality proteins in their diets to meet their requirements for essential amino acids. On the other hand, the microorganisms in ruminants can utilize considerable amounts of simple forms of nitrogen to synthesize high-quality microbial protein which is, in turn, utilized to meet the ruminant's requirement for amino acids. Thus, many ruminant feeds now contain varying portions of urea, an economical simple form of nitrogen, which is synthesized commercially from nonfeed sources. Simple-stomached animals require most of the vitamins in the diet. The microorganisms in the rumen synthesize adequate quantities of the water-soluble vitamins to supply the requirement for the ruminant animal. The fat-soluble vitamins A, D,

and E must be supplied as needed to all farm animals. Horses and mules have simple stomachs but they also have an enlargement of the cecum (part of the large intestine), in which bacterial action takes place similar to that in the rumen of ruminants. The requirements for most nutrients do not remain the same throughout the life of an animal but relate to the productive function being performed. Therefore, the requirements are much higher for growth and lactation than they are for pregnancy or maintenance.

Livestock judging. The evaluation, or judging, of livestock is important to both the purebred and the commercial producer. The purebred producer usually is much more interested in show ring judging, or placings, than is the commercial producer. Because of the short time they are in the show ring, the animals must be placed on the basis of type or appearance by the judge who evaluates them. The show ring has been an important influence in the improvement of livestock by keeping the breeders aware of what judges consider to be desirable types. The shows have also brought breeders together for exchange of ideas and breeding stock and have helped to advertise breeds of livestock and the livestock industry. The demand for better meat-animal carcasses has brought about more shows in which beef cattle and swine are judged, both on foot and in the carcass. This trend helps to promote development of meat animals of greater carcass value and has a desirable influence upon show-ring standards for meat animals. The standards used in show rings have shifted toward traits in live animals which are highly related to both desirable carcass traits and high production efficiency.

Grading of market animals. The grading on foot of hogs or cattle for market purposes requires special skill. In many modern livestock markets, hogs are graded as no. 1, 2, 3, or 4 according to the estimated values of the carcasses. Those hogs grading no. 3 are used to establish the base price, and sellers are paid a premium for better animals. In some cases the grade is used also to place a value on the finished market product. For example, the primal cuts from beef cattle are labeled prime, choice, or good on the basis of the grade which the rail carcass received.

Livestock disease and pest control. The numerous diseases of farm livestock require expert diagnosis and treatment by qualified veterinarians. The emphasis on intensive animal production has increased stresses on animals and generally increased the need for close surveillance of herds or flocks for disease outbreaks.

The transmission of infectious diseases is a three-stage process involving the reservoir of the disease organisms, the mode of transmission, and the susceptible animal. A disease can be thwarted at each stage: at the reservoir by isolation and quarantine, at the transmission stage by good hygiene, and in the susceptible animal by immunization and antibiotics. The most efficacious point for disease control varies with the particular disease etiology.

Some diseases are common to several classes of livestock. For example, the following diseases may affect cattle, sheep, and goats: actinomycosis, anthrax, blackleg, brucellosis, leptospirosis, listeriosis, mastitis, pinkeye, and shipping fever. Other diseases generally affect only a particular species of livestock.

Both external and internal parasites are common afflictions of livestock but can be controlled by proper management of the animals. Sanitation is of utmost importance in the control of these pests, but under most circumstances sanitation must be supplemented with effective insecticides, ascaricides, and fungicides. *See* FUNGISTAT AND FUNGICIDE; INSECTICIDE; PESTICIDE.

Internal parasites. Internal parasites, such as stomach and intestinal worms in sheep, cannot be controlled by sanitation alone under most farm conditions. They are a more critical problem under intensive management systems and in warm, humid climates. For many years the classic treatment for sheep was drenching with phenothiazine and continuous free choice feeding of one part phenothiazine mixed with nine parts of salt. Drugs have been developed which more effectively break the life cycle of the worm and have a broader spectrum against different classes of parasites.

Control of gastrointestinal parasites in cattle can be accomplished in many areas by sanitation and the rotational use of pastures. In areas of intensive grazing, animals, especially the young ones, may become infected. Regular and timely administration of antiparasite drugs is the customary means of controlling the pests. Otherwise their effects will seriously decrease the economic productivity of animals. *See* MEDICAL PARASITOLOGY.

The use of drugs to control gastrointestinal parasites and also certain enteric bacteria in hogs is commonplace. Control is also dependent on good sanitation and rotational use of nonsurfaced lots and pastures. Similarly, because of intensive housing systems, the opportunities for infection and spread of both parasitism and disease in poultry flocks are enhanced by poor management conditions. The producer has a large number of materials to choose from in preventing or treating these conditions, including sodium fluoride, piperazine salts, nitrofurans, and arsenicals and antibiotics such as penicillin, tetracyclines, and hygromycin.

External parasites. Control of horn flies, horseflies, stable flies, lice, mange mites, ticks, and fleas on farm animals has been changed with insecticides. Such compounds as methoxychlor, toxaphene, lindane, and malathion were very effective materials for the control of external parasites. The use of these materials has been restricted to certain conditions and classes of animals by law. Reliable information should be obtained before using these materials for the control of external parasites.

Control of cattle grubs, or the larvae of the heel fly, may be accomplished by dusting the backs of the animals with powders or by spraying them under high pressure. Systemic insecticides for grub control are approved if they are used according to the manufacturer's recommendation.

Fungus infections. Actinomycosis is a fungus disease commonly affecting cattle, swine, and horses. In cattle this infection is commonly known as lumpy jaw. The lumpy jaw lesion may be treated with tincture of iodine or by local injection of streptomycin in persistent cases. Most fungus infections, or mycoses, develop slowly and follow a prolonged course. A veterinarian should be consulted for diagnosis and treatment.

General animal health care. Numerous other disease organisms pose a constant threat to livestock. Although many of these organisms can be treated therapeutically, it is much more advisable economically to establish good preventive medicine and health care programs under the guidance of a veterinarian.

Ronald R. Johnson; W. A. Williams; R. Albaugh

Bibliography. D. Acker and M. Cunningham, *Animal Science and Industry*, 5th ed., 1997; J. R. Campbell and J. F. Lasley, *The Science of Animals That Serve Mankind*, 3d ed., 1985; J. R. Campbell and R. T. Marshall,*The Science of Providing Milk for Man*, 1975; M. E. Ensminger, *Beef Cattle Science*, 6th ed., 1987; M. E. Ensminger, *Horses and Horsemanship*, 7th ed., 1998; M. E. Ensminger, *Sheep and Goat Science*, 1998; J. L. Krider and W. E. Carroll, *Swine Production*, 5th ed., 1982; M. O. North, *Commercial Chicken Production Manual*, 3d ed., 1984; J. F. Timoney et al., *Hagan and Bruner's Microbiology and Infectious Diseases of Domestic Animals*, 8th ed., 1988.

Agricultural science (plant)

The pure and applied science that is concerned with botany and management of crop and ornamental plants for utilization by humankind. Crop plants include those grown and used directly for food, feed, or fiber, such as cereal grains, soybeans, citrus, and cotton; those converted biologically to products of utility, such as forage plants, hops, and mulberry; and those used for medicinal or special products, such as digitalis, opium poppy, coffee, and cinnamon. In addition, many plant products such as crambe oil and rubber are used in industry where synthetic products have not been satisfactory. Ornamental plants are cultured for their esthetic value. *See* FLORICULTURE; ORNAMENTAL PLANTS.

The ultimate objective of plant agriculture is to recognize the genetic potential of groups of plants and then to manipulate and utilize the environment to maximize that genetic expression for return of a desirable product. Great advancements in crop culture have occurred by applying knowledge of biochemistry, physiology, ecology, morphology, anatomy, taxonomy, pathology, and genetic engineering. Contributions of improved plant types by breeding, and the understanding and application of principles of atmospheric science, soil science, and animal and human nutrition, have increased the efficiency and decreased the risks of crop production. *See* GENETIC ENGINEERING.

Domestication of crop plants. All crops are believed to have been derived from wild species. However, cultivated plants as they are known today have undergone extensive modification from their wild prototypes as a result of the continual efforts to improve them. These wild types were apparently recognized as helpful to humans long before recorded history. Desirable plants were continually selected and replanted in order to improve their growth habit, fruiting characteristics, and growing season. Selection has progressed so far in cases such as cabbage and corn (maize) that wild ancestors have become obscure.

Centers of origins of most crop plants have been determined to be in Eurasia, but many exceptions exist. This area of early civilization apparently abounded with several diverse plant types that led to domestication of the crops known today as wheat, barley, oats, millet, sugarbeets, and most of the cultivated forage grasses and legumes. Soybeans, lettuce, onions, and peas originated in China and were domesticated as Chinese civilization developed. Similarly, many citrus fruits, banana, rice, and sugarcane originated in southern Asia. Sorghum and cowpeas are believed to have originated in Africa. Crops which were indigenous to Central and South America but were brought to North America by migrating peoples include corn, potato, sweet potato, pumpkin, sunflower, tobacco, and peanut. Thus, prior to 1492 there was little, if any, mixing of crop plants and cultural practices between the Old and the New Worlds. Most of the major agricultural crops of today in the United States awaited introduction by early settlers, and later by plant explorers. *See* ORIENTAL VEGETABLES.

During domestication of crop plants the ancient cultivators in their geographically separated civilizations must have had goals similar to present-day plant breeders. Once valuable attributes of a plant were recognized, efforts were made to select the best types for that purpose. Desirable characteristics most likely included improved yield, increased quality, extended range of adaptation, insect and disease resistance, and easier cultural and harvesting operations. About 350,000 species of plants exist in the world, yet only about 10,000 species can be classified as crops by using the broadest of definitions. Of these, about 150 are of major importance in world trade, and only 15 make up the majority of the world's food crops. On a world basis, wheat is grown on the most acreage, followed by rice, but rice has a higher yield per area than wheat so their total production is about equal. Other major crops are corn, sorghum, millet, barley; sugarcane and sugarbeet; potato, sweet potato, and cassava; bean, soybean, and peanut; and coconut and banana.

Redistribution of crop plants. Crop distribution is largely dictated by growth characteristics of the crop, climate of the region, soil resources, and social habits of the people. As plants were domesticated by civilizations in separate parts of the world, they were discovered by early travelers. This age of exploration led to the entrance of new food crops into European

agriculture. The potato, introduced into Spain from the New World before 1570, was to become one of the most important crops of Europe. When introduced into Ireland, it became virtually the backbone of the food source for that entire population. Corn was introduced to southern Europe and has become an important crop. Rice has been cultivated in Italy since the sixteenth century. Tobacco was also introduced to European culture, but remained a major source of export of the early colonial settlements in America.

European agriculture reciprocated by introducing wheat, barley, oats, and several other food and feed crops into the New World. In the new environment, plants were further adapted by selection to meet the local requirements. Cultural technology regarding seeding, harvesting, and storing was transferred along with the crops. This exchange of information oftentimes helped allow successful culture of these crops outside of their center of origin and domestication. Today the center of producton of a crop such as wheat in the United States and potato in Europe is often markedly different from its center of origin.

The United States recognized early the need for plant exploration to find desirable types that could be introduced. Thomas Jefferson wrote in 1790, "The greatest service which can be rendered to any country is to add a useful plant to its culture." Even today the U.S. Department of Agriculture conducts plant explorations and maintains plant introduction centers to evaluate newly found plants. Explorations are also serving a critical need for preservation of germplasm for plant breeders, as many of the centers of origin are becoming agriculturally intensive, and wild types necessary to increase genetic diversity will soon be extinct.

Adaptation of crop plants. Crop plants of some sort exist under almost all environments, but the major crops on a world basis tend to have rather specific environmental requirements. Furthermore, as crop plants are moved to new locations the new environment must be understood, and cultural or management changes often must be made to allow best performance. Varieties and strains of crops have been specifically developed that can better cope with cold weather, low rainfall, diseases, and insects to extend even further the natural zone of adaptation.

Temperature. Temperature has a dominant influence on crop adaptation. The reason is that enzyme activity is very temperature-dependent and almost all physiological processes associated with growth are enzymatically controlled. Crops differ widely in their adapted temperature range, but most crops grow best at temperatures of 59–88°F (15–32°C). Optimum day temperature for wheat, however, is 68–77°F (20–25°C), while for corn it is about 86°F (30°C) and for cotton about 95°F (35°C).

The frost-free period also influences crop adaptation by giving an indication of the duration of the growing season. Thus a growing season for corn or soybeans might be described as one with mean daily temperatures between 64 and 77°F (18 and 25°C), and with an average minimum temperature exceeding 50°F (10°C) for at least 3 months. Small grains such as wheat, barley, and oats tolerate a cooler climate with a period of only 2 months when minimum temperatures exceed 50°F (10°C). Because of optimum growth temperatures and frost-free periods, it is easily recognized why the spring Wheat Belt includes North Dakota and Montana, the Corn Belt Iowa and Illinois, and the Cotton Belt Misssissippi and Alabama. Farmers use planting dates and a range in maturity of varieties to match the crop to the growing season. *See* PLANT GROWTH.

In many winter annual (the plants are sown in fall, overwinter, and mature in early summer), biennial, and perennial crops, cold temperatures also influence distribution. The inherent ability of the crop to survive winter limits distribution of crops. As a generalized example, winter oats are less able to survive cold winters than winter barley, followed by winter wheat and then winter rye. Thus oats in the southern United States are mostly winter annual type, while from central Arkansas northward they are spring type. The dividing line for barley is about central Missouri and for wheat about central South Dakota. Winter rye can survive well into Canada. *See* COLD HARDINESS (PLANT).

A cold period may be required for flowering of winter annual and perennial crops. This cold requirement, termed vernalization, occurs naturally in cold environments where the dormant bud temperature is near 32°F (0°C) for 4 to 6 weeks during winter. Without this physiological response to change the hormonal composition of the terminal bud, flowering of winter wheat would not occur the following spring. Bud dormancy is also low-temperature-induced and keeps buds from beginning spring growth until a critical cold period has passed. The flower buds of many trees such as peach, cherry, and apple require less chilling than do vegetative buds, and therefore flower before the leaves emerge. The intensity and duration of cold treatment necessary to break dormancy differs with species and even within a species. For example, some peaches selected for southern areas require only 350 h below −12°F (8°C) to break dormancy, while some selected for northern areas may require as much as 1200 h. This cold requirement prevents production of temperate fruit crops in subtropical regions, but in temperate climates its survival value is clear. A physiological mechanism that prevents spring growth from occurring too early helps decrease the possibility of cold temperature damage to the new succulent growth. *See* DORMANCY; VERNALIZATION.

Temperature of the crop environment can be altered by date of planting of summer annuals, by proper site selection, and by artificial means. In the Northern Hemisphere the south- and west-facing slopes are usually warmer than east- or north-facing slopes. Horticulturists have long used mulches to control soil temperature, and mists and smoke for short-term low-temperature protection.

Water. Water is essential for crop production, and natural rainfall often is supplemented by irrigation. Wheat is grown in the Central Plains states of the

United States because it matures early enough to avoid the water shortage of summer. In contrast, corn matures too late to avoid that drought condition and must be irrigated to make it productive. In the eastern United States where rainfall is higher, irrigation is usually not needed for good yields. *See* IRRIGATION (AGRICULTURE).

Crops transpire large amounts of water through their stomata. For example, corn transpires about 350 lb of water for each pound of dry weight produced. Wheat, oats, and barley transpire about 600 lb, and alfalfa about 1000 lb for each pound of dry weight produced. Fallowing (allowing land to be idle) every other season to store water in the soil for the succeeding crop has been used to overcome water limitations and extend crops further into dryland areas. *See* PLANT-WATER RELATIONS.

Light. Light intensity and duration also play dominant roles. Light is essential for photosynthesis. The yield of a crop plant is related to its efficiency in intercepting solar radiation by its leaf tissue, the efficiency of leaves in converting light energy into chemical energy, and the transfer and utilization of that chemical energy (usually sugars) for growth of an economic product. Crops differ markedly in the efficiency of their leaves. *See* PHOTOSYNTHESIS.

Crops also differ markedly in leaf area and leaf arrangement. Humans greatly influence the photosynthetic area by altering number of plants per area and by cutting and pruning. Corn producers have gradually increased plant population per area about 50% since about 1950 as improved varieties and cultural methods were developed. This has increased the leaf area of the crop canopy and the amount of solar energy captured. Defoliation and pruning practices also influence solar energy capture and growth rate. Continued defoliation of pastures by grazing may reduce the photosynthetic area to the point that yield is diminished. In these perennial plants, carbohydrates are stored in underground organs such as roots, rhizomes, and stem bases to furnish food for new shoots in spring and following cutting. Availability of water and nitrogen fertilizer also influences the amount of leaf area developed.

Photoperiodism, the response of plants to day length, also has a dramatic effect on plant distribution. Such important adaptive mechanisms as development of winter hardiness, initiation and maintenance of bud dormancy, and floral initiation are influenced by photoperiod. Plants have been classified as long-day, those that flower when day lengths exceed a critical level; short-day, which is opposite to long-day; and day-neutral, those that are not affected by day length. Photoperiod also has a controlling influence on formation of potato tubers, onion bulbs, strawberry runners, and tillers of many cereal grains and grasses. Farmers select varieties bred for specific areas of the country to ensure that they flower properly for their growing season. Horticulturists, in a more intensive effort, provide artificial lighting in greenhouses to lengthen the photoperiod, or shorten it by shading, to induce flowering and fruit production at will; for example, to ready poinsettias for the Christmas holiday trade or asters for Easter. Natural photoperiods are important in plant breeding when varieties of day-length-sensitive crops must be developed for specific localities. Soybeans are day-length-sensitive and have been classified into several maturity groups from north to south in latitude. *See* PHOTOPERIODISM.

Pathogens. Pathogens of plants that cause diseases include fungi, bacteria, viruses, and nematodes. These organisms are transmitted from plant to plant by wind, water, and insects and infect the plant tissue. Organisms infect the plant and interfere with the physiological functions to decrease yield. Further, they infect the economic product and decrease its quality. Pathogens are most economically controlled by breeding resistant varieties or by using selective pesticides. Insects decrease plant productivity and quality largely by mechanical damage to tissue and secretion of toxins. They are also usually controlled by resistant varieties or specific insecticides. *See* PLANT PATHOLOGY.

Soil. The soil constitutes an important facet of the plant environment. Soil physical properties such as particle size and pore space determine the water-holding capacity and influence the exchange of atmospheric gases with the root system. Soil chemical properties such as pH and the ability to supply nutrients have a direct influence on crop productivity. Farmers alter the chemical environment by addition of lime or sulfur to correct acidic or basic conditions, or by addition of manures and chemical fertilizers to alter nutrient supply status. Soil also is composed of an important microbiological component that assists in the cycling of organic matter and mineral nutrients. *See* SOIL.

Management. During and following domestication of crop plants, many cultural or management practices have been learned that enhance production or quality of the crop. This dynamic process is in operation today as the quest continues to improve the plant environment to take advantage of the genetic potential of the crop. These practices have made plant agriculture in the United States one of the most efficient in the world. Some major changes in technology are discussed in the following sections.

Mechanization. In 1860 an average United States farm worker produced enough food and fiber for fewer than 5 other persons. In 1950 it was enough for 25, and today exceeds 50 other persons. Mechanization, which allowed each farm worker to increase the area managed, is largely responsible for this dramatic change.

A model of a grain reaper in 1852 demonstrated that nine farmers with the reaper could do the work of 14 with cradles. By 1930 one farmer with a large combine had a daily capacity of 20–25 acres (8–10 hectares), and not only harvested but threshed the grain to give about a 75-fold gain over the cradle-and-flail methods of a century earlier. In 1975 one farmer with a faster-moving combine could harvest 50–75 acres (20–30 ha) per day. The mechanical cotton picker harvests a 500-lb (227-kg) bale in 75 min,

40–50 times the rate of a hand picker. The peanut harvester turns out about 300 lb (136 kg) of shelled peanuts per hour, a 300-h job if done by hand labor. Hand setting 7500 celery plants is a day's hard labor for one person. However, a modern transplanting machine with two people readily sets 40,000, reducing labor costs by 67%. Today crops such as cherries and tomatoes are machine-harvested. When machines have been difficult to adapt to crops such as tomatoes, special plant varieties that flower more uniformly and have tougher skins on the fruit have been developed. *See* AGRICULTURAL MACHINERY.

Fertilizers and plant nutrition. No one knows when or where the practice originated of burying a fish beneath the spot where a few seeds of corn were to be planted, but it was common among North American Indians when Columbus discovered America and is evidence that the value of fertilizers was known to primitive peoples. Farm manures were in common use by the Romans and have been utilized almost from the time animals were first domesticated and crops grown. It was not until centuries later, however, that animal fertilizers were supplemented by mineral forms of lime, phosphate, potassium, and nitrogen. Rational use of these substances began about 1850 as an outgrowth of soil and plant analyses.

Justus von Liebig published *Chemistry and Its Application to Agriculture and Physiology* in 1840, which led to concepts on modifying the soil by fertilizers and other amendments. These substances soon became the center of crop research. In 1842 John Bennet Lawes, who founded the famous Rothamsted Experiment Station in England, obtained a patent for manufacture of superphosphates and introduced chemical fertilizers to agriculture. As research progressed, crop responses to levels of phosphate, lime, and potassium were worked out, but nitrogen nutrition remained puzzling. The nitrogen problem was clarified in 1886, when H. Hellriegel and H. Wilfarth, two German chemists, reported that root nodules and their associated bacteria were responsible for the peculiar ability of legume plants to use atmospheric nitrogen. These findings collectively allowed rational decision-making regarding nutrition of crop plants and were understandably very significant in increasing crop productivity.

By the early twentieth century 10 elements had been identified as essential for proper nutrition. These were carbon, hydrogen, and oxygen, which are supplied by the atmosphere; and nitrogen, potassium, phosphorus, sulfur, magnesium, calcium, and iron, supplied by the soil. The first 40 years of the twentieth century witnessed the addition of manganese, boron, copper, zinc, molybdenum, and chlorine to the list of essential mineral nutrients. These 6 are required only in very small amounts as compared with the first 10 and have been classified as micronutrients. From a quantitative view they are truly minor, but in reality they are just as critical as the others for plant survival and productivity. Many early and puzzling plant disorders are now known to be due to insufficient supplies of micronutrients in the soil, or to their presence in forms unavailable to plants. *See* PLANT MINERAL NUTRITION.

An important result of discoveries relating to absorption and utilization of micronutrients is that they have served to emphasize the complexity of soil fertility and fertilizer problems. In sharp contrast to early thoughts that fertilizer practices should be largely a replacement in the soil of what was removed by the plant, it is now recognized that interaction and balance of mineral elements within both the soil and plant must be considered for efficient crop growth. Usage of chemical fertilizers has become more widespread with time. *See* FERTILIZER; FERTILIZING.

Pesticides. Total destruction of crops by swarms of locusts and subsequent starvation of many people and livestock have occurred throughout the world. Pioneers in the Plains states suffered disastrous crop losses from hordes of grasshoppers and marching army worms. An epidemic of potato blight brought hunger to much of western Europe and famine to Ireland in 1846–1847. The use of new insecticides and fungicides has done much to prevent such calamities. Various mixtures, really nothing more than nostrums (unscientific concoctions), were in use centuries ago, but the first really trustworthy insect control measure appeared in the United States in the mid-1860s, when paris green was used to halt the eastern spread of the Colorado potato beetle. This was followed by other arsenical compounds, culminating in lead arsenate in the 1890s.

A major development occurred during World War II, when the value of DDT (dichlorodiphenyltrichloroethane) for control of many insects was discovered. Although this compound was known to the chemist decades earlier, it was not until 1942 that its value as an insecticide was definitely established and a new chapter written in the continual contest between humankind and insects. Three or four applications of DDT gave better season control of many pests at lower cost than was afforded by a dozen materials used earlier. Furthermore, DDT afforded control of some kinds of pests that were practically immune to former materials. In the meantime other chemicals have been developed that are even more effective for specific pests, and safer to use from a human health and ecological viewpoint.

Dusts and solutions containing sulfur have long been used to control mildew on foliage, but the first highly effective fungicide was discovered accidentally in the early 1880s. To discourage theft, a combination of copper sulfate and lime was used near Bordeaux, France, to give grape vines a poisoned appearance. Bordeaux mixture remained the standard remedy for fungus diseases until the 1960s, when other materials with fewer harmful side effects were released.

New pesticides are constantly being developed by private industry and are carefully monitored by several agencies of the federal government. Besides evaluation for its ability to repel or destroy a certain pest, influence of the pesticide on physiological processes in plants, and especially long-range

implications for the environment and human health, are carefully documented before federal approval for use is granted. *See* PESTICIDE.

Herbicides. Because they are readily visible, weeds were recognized as crop competitors long before microscopic bacteria, fungi, and viruses. The time-honored methods of controlling weeds have been to use competitive crops or mulches to smother them, or to pull, dig, hoe, or cultivate them out. These methods are still effective and most practical in many instances. However, under many conditions weeds can be controlled chemically much more economically. A century ago regrowth of thistles and other large weeds was prevented by pouring salt on their cut stubs. Salt and ashes were placed along areas such as courtyards, roadsides, and fence rows where all vegetation needed to be controlled. However, until the 1940s selective weed control, where the crop is left unharmed, was used on a very limited scale. The first of these new selective herbicides was 2,4-D (2,4-dichlorophenoxyacetic acid), followed shortly by 2,4,5-T (2,4,5-trichlorophenoxyacetic acid) and other related compounds. This class of herbicides is usually sprayed directly on the crop and weeds and kills susceptible plants by upsetting normal physiological processes, causing abnormal increases in size, and distortions that eventually lead to death of the plant. As a group, these herbicides are much less toxic to grasses than to broad-leaved species, and each one has its own character of specificity. For instance, 2,4-D is more efficient for use against herbaceous weeds, whereas 2,4,5-T is best for woody and brushy species.

Today's herbicides have been developed by commercial companies to control a vast array of weeds in most crop management systems. Some herbicides are preemergence types that are sprayed on the soil at time of corn planting. This type of herbicide kills germinating weed seeds by interfering with their photosynthetic system so they starve to death. Meanwhile metabolism of the resistant corn seedlings alters the chemical slightly to cause it to lose its herbicidal activity. Such is the complexity that allows specificity of herbicides. The place that herbicides have come to occupy in agriculture is indicated by the fact that farmlands treated in the United States have increased from a few thousand acres in 1940 to over 150,000,000 acres (60,000,000 ha), not including large areas of swamp and overflow lands treated for aquatic plant control and thousands of miles of treated highways, railroad tracks, and drainage and irrigation ditches. *See* HERBICIDE.

Growth regulators. Many of the planters' practices since the midnineteenth century may be classified as methods of regulating growth. Use of specific substances to influence particular plant functions, however, has been a more recent development, though these modern uses, and even some of the substances applied, had their antecedents in century-old practices in certain parts of the world.

Since about 1930 many uses have been discovered for a considerable number of organic compounds having growth-regulating influences. For instance, several of them applied as sprays a few days to several weeks before normal harvest will prevent or markedly delay dropping of such fruits as apples and oranges. Runner development in strawberries and sucker (tiller) development in tobacco can be inhibited by sprays of maleic hydrazide. Another compound called CCC (2-chloroethyl trimethyl ammonium chloride) is used to shorten wheat plants in Europe to allow higher levels of fertilizer. Many greenhouse-grown flowers are kept short in stature by CCC and other growth regulators. Striking effects from gibberellins and fumaric acid have also been reported. The first greatly increase vegetative growth and the latter causes dwarfing. In practice, growth-inhibiting or growth-retarding agents are finding wider use than growth-stimulating ones. In higher concentration many growth-retarding agents become inhibiting agents. *See* GIBBERELLIN.

There are marked differences between plant species and even varieties in their response to most plant growth regulators. Many, if not most, growth regulators are highly selective; a concentration of even 100 times that effective for one species or variety is necessary to produce the same response in another. Furthermore, the specific formulation of the substances, for example, the kind or amount of wetting agent used with them, is important in determining their effectiveness. In brief, growth regulators are essentially new products, though there are century-old instances of the empirical use of a few of them. Some influence growth rate, others direction of growth, others plant structure, anatomy, or morphology. With the discovery of new ones, indications are that it is only a matter of time before many features of plant growth and development may be directly or indirectly controlled by them to a marked degree. Applications of these substances in intensive agriculture are unfolding rapidly, and their use is one of the many factors making farming more of a science and less of an art. *See* PLANT HORMONES.

Plant improvement. From earliest times humans have tried to improve plants by selection, but it was the discovery of hybridization (cross-mating of two genetically different plants) that eventually led to dramatic increases in genetic potential of the plants. Hybridization was recognized in the early 1800s, well before Mendel's classic genetic discoveries, and allowed the combination of desirable plants in a complementary manner to produce an improved progeny. Plant breeding had a dramatic flourish in the early twentieth century following the rediscovery of Mendel's research and its implications, and has had much to do with the increased productivity per area of present-day agriculture.

Corn provides the most vivid example of how improvement through genetic manipulation can occur. Following the commercial development of corn hybrids about 1930, only a few acres were planted, but by 1945 over 90% of the acreage was planted to hybrids, and today nearly 100% is planted. It has been conservatively estimated that hybrids of corn have 25% more yield potential than old-style varieties. Subsequently, plant breeders have utilized hybridization

for development of modern varieties of most major crop species.

While very significant changes through crop breeding have occurred in pest resistance and product quality, perhaps the character of most consequence was an increase in the lodging (falling over) resistance of major grain crops. Corn, wheat, and rice have all been bred to be shorter in stature and to have increased stem resistance to breaking. These changes have in turn allowed heavier fertilization of crops to increase photosynthetic area and yield. The impact of this was recognized when Norman Borlaug was awarded the Nobel Peace Prize in 1970 for his breeding contribution to the "green revolution." His higher-yielding wheats were shorter and stiff-strawed so they could utilize increased amounts of nitrogen fertilizer. They were also day-length-insensitive and thus had a wide adaptation. New varieties of rice such as IR-8 developed at the International Rice Research Institute in the Philippines are shorter, are more responsive to nitrogen fertilizer, and have much higher potential yield than conventional varieties. With those new wheat and rice varieties many countries gained time in their battle between population and food supply. *See* BREEDING (PLANT).

Often in tree crops and high-value crops it is not feasible to make improvements genetically, and other means are utilized. Grafting, or physically combining two or more separate plants, is used as a method of obtaining growth control in many fruit trees, and also has been used to provide disease resistance and better fruiting characteristics. This technique is largely limited to woody species of relatively high value. Usually tops are grafted to different rootstocks to obtain restricted vegetative growth, as in dwarf apple and pear trees which still bear normal-sized fruit. Alternatively, junipers and grapes are grafted to new rootstocks to provide a better root system. In both cases the desirability of the esthetic or economic portion warrants the cost and effort of making the plant better adapted to environmental or management conditions. Curtis L. Nelson

Seed production. Seed production is the process of sexual plant propagation in which genetic recombination occurs, and it is an important step in crop improvement and distribution. Developing more productive crops through plant hybridization and genetic engineering is a primary method of increasing agricultural production. Improvement comes from greater crop yield, better quality, and improved pest resistance. Commercial seed production is the major method of propagating a new, improved variety. Genetic change must not occur during seed multiplication.

A prescribed system of maintenance and seed multiplication is established for individual crops. The methods used to increase a new variety will depend on the species and the breeding system. In the United States, state agencies supervise the multiplication to ensure that adequate safeguards have been used to protect the genetic purity through a certification program. Special tags are used to identify certified seed. Privately owned varieties are often produced and marketed under the supervision of the organization's quality control department, without public agency certification.

Seed production is a major economic enterprise of international scope. Regions with a soil and a climate favorable to seed production specialize in providing high-quality seed, and are often located far from the consuming area. For example, the dry climate in the irrigated regions in the western United States favors high seed yield, and harvest is not affected by rain; therefore much of the vegetable and forage legume seed is produced there. Cool season grass species are adapted to the mild climate of the Pacific Northwest, a major source of forage grass seed. When rapid seed multiplication is needed, stock seed from the Northern Hemisphere crops is shipped to the Southern Hemisphere to allow the harvest of two seed crops in 12 months.

Seed crop species are self- or cross-pollinated. Isolation distance and crop management are determined by pollination characteristics. Self-pollinated species include wheat, barley, soybeans, and lettuce. These crop varieties are homozygous when released, and the progeny will be true to type. Once a wheat variety has been released, for example, the breeder will maintain the variety by selecting heads which are threshed separately and grown in headrows. A row which fails to conform to the varietal type is discarded. Approved rows are harvested, bulked, and classified as breeder seed for further multiplication. *See* POLLINATION.

Cross-pollinated species are fertilized by pollen carried by wind or insects. Wind-borne pollen is important in crops such as corn, sugarbeet, and ryegrass. Insect pollination is valuable in the fertilization of alfalfa, clovers, and mustards. Cross-pollinated crops require isolation from other varieties of the same species for protection of purity.

Superior plant vigor is developed from F1 hybrid (first-generation cross) seed produced by crossing two or more inbred lines. Hybrid vigor is the increased performance over parental types. Although seed of hybrid varieties are more difficult to produce, they are important in commercial production of crops such as corn which is grown exclusively from hybrid varieties. Other important hybrid crops include cabbage, cantaloupe, sugarbeet, cucumber, onion, petunia, tomato, squash, and watermelon. Seed from the F1 hybrid will not produce the same yield or crop quality of the parent plants; therefore new hybrid seed must be purchased each year, creating a large demand for planting seed. Some hybrid seed crops such as corn are grown in large fields, while certain flower hybrids such as petunias require greenhouse culture.

Seed production has become a very specialized enterprise serving scientific agriculture. *See* REPRODUCTION (PLANT); SEED. Harold Youngberg

Bibliography. J. A. Barden, R. G. Halfacre, and D. J. Parrish, *Plant Science*, 1987; R. J. Delorit, L. J. Greub, and H. L. Ahlgren, *Crop Production*, 4th ed., 1974; W. R. Fehr, *Principles of Cultivar*

Development, 1987; J. R. Harlan, *Crops and Man*, 2d ed., 1992; J. Janick et al., *Plant Science: An Introduction to World Crops*, 3d ed., 1981; J. H. Martin et al., *Principles of Field Crop Production*, 3d ed., 1989.

Agricultural soil and crop practices

The techniques and methods used in plowing, harrowing, planting, tilling, harvesting, drying, storing, and processing of agricultural crops.

Plowing

Cutting and turning a furrow with a plow is usually the first and most effective operation of tillage. There are two types of plow, the moldboard (**Fig. 1**) and the disk.

Moldboard plow. The working part of the moldboard plow is the plow bottom. The plow furrow is cut or broken loose on the land side by the shin of the plow or by a coulter or jointer attached to the plow beam, and on the bottom by the point and edge of the plow share. The advancing plow moldboard exerts pressure on the furrow slice, pulverizing it and causing it to move upward and to the side in a spiraling path so that it is partially or completely inverted. The shape of the moldboard may be described as a section cut from a warped cylinder, on the inside of which the soil spirals. The selection of design of the moldboard depends upon the physical reaction of the kind of soil and sod cover on which it is to be used. Because precise information of soil physical reaction is inadequate and the amount of pulverization and placement, or throw, of the furrow slice depends upon the speed of plowing, plows have been designed largely by the cut-and-try method. This has resulted in a great variety of shapes. Moldboard plows have largely been adapted to power farming by the selection of bottoms suitable for certain soils and combination of these bottoms into gangs sufficiently large to use the power of the tractor at the approximate speed for which the bottoms were designed. The power required at normal speed of 4 mi/h (6.5 km/h) varies from 2–3 lb/in.^{2} (14–21 kilopascals) of cross section of furrow slice for sand to 20 lb/in.^{2} (140 kPa) for tough clay. Modifications of shape have permitted a speed of 5–6 mi/h (8–9.5 km/h).

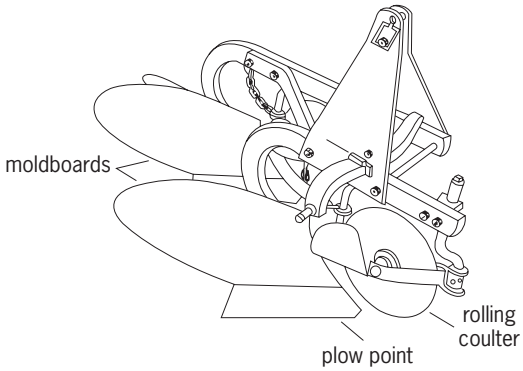

Fig. 1. One type of moldboard plow. (*Tractor and Implement Operations–North America, Ford Motor Co.*)

Moldboard plow bottoms are designated as right-hand or left-hand bottoms, depending upon the direction of throw. Both may be combined on one frame so that the plow can be pulled back and forth, always throwing the furrow downhill, or in one direction. Sizes range from 8 to 20 in. (20 to 50 cm) in width of cut; the 10–12-in. (25–30-cm) size is suitable for use with animal power, and the 14–16-in. (35–40-cm) size is largely used with power equipment. Although a plow may be used to plow at different depths, most have been designed to work at a depth of approximately one-half the width of the furrow.

The effectiveness of plowing may be materially increased by attachments. A jointer, or a jointer and coulter combination, which cuts a small furrow in front of the plow shin, permits complete coverage of sod. Where there are large amounts of rubbish or crop residue, a weed chain or wire can be used to drag the debris into the furrow and hold it there until covered. A furrow wheel may reduce land-side friction, and a depth gage on the beam helps secure uniform depth. A modified form of plow bottom, called a lister, is in effect a right and a left moldboard joined at the shin so as to throw soil both to the right and to the left. This produces a furrow or trough, called a list, in which seed is planted. Because it concentrates rainfall in the furrow, this method is used largely in areas of light rainfall.

Disk plow. The disk plow consists of a number of disk blades attached to one axle or gang bolt. This plow is used for rapid, shallow plowing. In fields where numerous rocks and roots are present, the disk plow, which rolls over obstacles, is substituted for the moldboard. The disk is also used for sticky soils that will not scour on a moldboard. The disk plow is manufactured in widths of from $2^1/_2$ to 20 ft (0.7 to 6 m). The disks are commonly spaced 8–10 in. (20–25 cm) apart. The angle between the gang bolt and the direction of travel is usually adjustable from 35 to 55°.

Harrowing

Soil preparation for planting usually involves the pulling of an implement called a harrow over the plowed soil to break clods, level the surface, and destroy weeds. A wide variety of implements are classified as harrows; the most common kinds are the disk harrow, the spike-tooth harrow, the spring-tooth harrow, and the knife harrow. Previously the function of seedbed preparation was performed almost entirely by the implements classified as harrows. With the introduction of power, farming is now performed in large part by field cultivators, rod weeders, rotary hoes, or treaders, subsurface packers, and various designs of rollers. Power-driven rotary tillers perform the function of both plowing and harrowing.

Kinds of harrows. The spike-tooth is the oldest form of commercial harrow and consists of spikes or teeth (usually adjustable) extending downward from

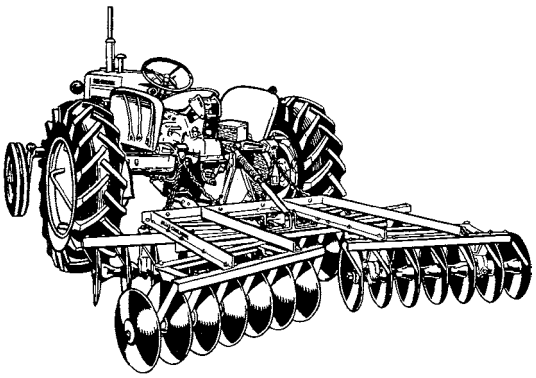

Fig. 2. One type of disk harrow. (*Allis-Chalmers*)

a frame. The teeth extend into the soil, and when the harrow is pulled forward, they cut through clods and break them. The teeth also stir and level the soil surface and kill weeds. This type of harrow has light draft and is built in sections, many of which may be joined together so that large areas can be covered quickly. This implement is most effective if used before clods dry; it is frequently attached behind the plow.

The spring-tooth harrow is similar to the spike-tooth type but has long curved teeth of spring steel. The spring action renders it suitable for rough or stony ground. It is particularly useful in bringing clods to the surface, where they can be pulverized. It is also used to bring the roots of weeds and obnoxious grasses to the surface for destruction, and to renovate and cultivate alfalfa fields. The knife harrow consists of a frame holding a number of knives which scrape and partly invert the surface to smooth it and destroy small weeds.

The disk harrow is probably the most universally used type (**Fig. 2**). It cuts clods and trash effectively, destroys weeds, cuts in cover crops, and smooths and prepares the surface for other farming operations. The penetration of the disk harrow depends largely upon weight. The disk blades are commonly 16–24 in. (40–60 cm) in diameter and are spaced 6–10 in. (15–25 cm) apart in gangs of 3–12 disks. Disk harrows can be obtained in widths up to 20 ft (6 m). A single-acting disk harrow has two opposed gangs throwing soil outward from the center; a tandem or double-acting disk has two additional gangs which throw the soil back toward the center. An important advancement in the design of the disk harrow is the offset disk. A right-hand offset disk harrow has a gang in front which throws to the right and a rear gang which throws to the left. It may be adjusted to pull to one side and to the rear of the tractor so as to harrow beneath the low limbs of orchard trees.

Other soil-preparation equipment. The field cultivator is used to perform many of the jobs of harrows before planting. It usually consists of a number of adjustable standards with sweeps or scrapes attached to tool bars in such a fashion that the soil is stirred from underneath, killing the weeds and creating a surface mulch for moisture conservation. The rod weeder is a power-driven rod, usually square in cross section, which also operates beneath the surface of loose soil, killing weeds and maintaining the soil in a loose mulched condition. It is adapted to large operations and is used in dry areas of the Northwest. A variety of rollers and packing wheels and clod crushers have been designed.

Planting

The practice of placing seed or vegetative propagating material in soil for multiplication through growth and reproduction is usually a seasonal operation. Its success depends upon soil preparation and placing the seed in an environment favorable to growth. The seed, which is an embryonic plant enclosed in a protective membrane, usually contains enough nutritional material to start growth. It must have suitable temperature, adequate air, and sufficient moisture to overcome its dormant condition and induce vigorous growth. In general, the seeding process consists of opening a furrow in properly prepared soil to the correct depth, metering and distributing the seed or planting material, depositing the seed in the furrow, and covering and compacting the soil around the seed to a degree suitable to the crop. Fertilizer is usually placed in the soil sufficiently near the seed that it will be available to the young plants after germination.

There are five general methods of planting based on five special types of machinery: (1) broadcasters; (2) seed drills used for small seed and grains; (3) planters for cultivated row crops such as corn or cotton; (4) special planters for parts of plants used for propagation, such as potato planters; and (5) transplanters used to set out small plants that have been grown in beds from small seed. The last is commonly used for tobacco, sweet potatoes, cabbage, trees, and many horticultural crops.

Broadcasters. Small grains, grasses, and clovers are planted by broadcasting or drilling. The broadcaster is usually a rotating fanlike distributor which throws the seed over a wide area by centrifugal force. Like hand seeding, this method requires the absence of gusty wind for most effective distribution. Under proper conditions, broadcasting can be done from airplanes. Broadcasting is especially suited to sowingseed in another crop without unduly disturbing the soil, such as sowing clover seed in wheat.

Drills. The grain drill opens a furrow and places the seed in it. Attachments, such as covering chains and wheels to press seed into the soil, are commonly used (**Fig. 3**). The seed is metered by a special apparatus into rows 6–14 in. (15–35 cm) apart. Several types of furrow openers adapted to different soil and crop conditions are available. Grain drills are also commonly equipped for fertilizer distribution and grass seeding.

Row-crop planters. Such crops as corn and cotton are planted with special planters in rows to simplify cultivation. Because yield may be greatly affected by the stand's being too thick or too thin, precision planting is important to avoid the cost of thinning or interplanting. Delinting of cotton-seed and sizing of seed corn and other seeds are important to precision

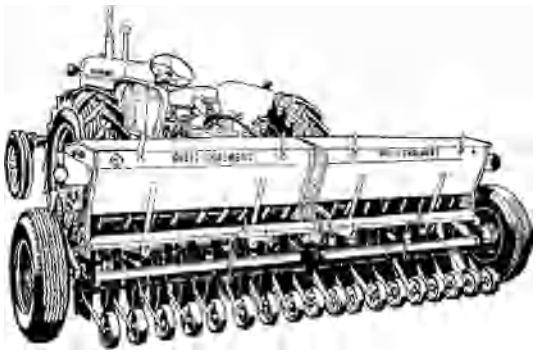

Fig. 3. All-crop drill with rubber press wheels to push seeds into soil. (*Allis-Chalmers*)

planting. Planters usually are equipped for dropping in hills or drilling in rows. The hills may be check-rowed, that is, spaced equally apart on the corners of squares so that the crop can be cultivated in two directions, thus avoiding hoeing. The precision necessary for this type of planting is secured by opening valves in the planter shank at measured intervals by means of buttons on a check wire.

Transplanters. Special kinds of equipment designed for the planting of cuttings or small plants are known as transplanters. Such machines usually transport one or more workers who assist the action of the machine in placing the plants in a furrow and properly covering them. Transplanters commonly supply a small quantity of water to each plant.

Tillage

The mechanical manipulation of the soil to improve its physical condition as a habitat for plants is called tillage. It includes plowing, inversion, loosening, harrowing, pulverization, packing, and rolling the soil, all to improve aeration and temperature conditions and to produce a firm seedbed (**Fig. 4**). Subsurface tillage is the loosening of soil by sweeps or blades pulled beneath the surface without inversion of the soil. This practice, especially adapted to dry areas, fragments the soil and leaves a mulch of stubble or other plant residues on the soil surface to conserve water and help control erosion.

Effective tillage eliminates competitive vegetation, such as weeds, and stimulates favorable soil microbiological activities. Natural forces of heating and cooling, swelling and shrinkage, wetting and drying, and freezing and thawing account for the major pulverization of soil and assist in the production of a favorable crumb structure. Wise practice dictates the avoidance of tillage when soil is so wet and plastic that its crumb structure is easily destroyed, as well as the use of those operations which put the soil in the most favorable condition for natural forces to act. This results in the minimum amount of time and power for soil preparation. Manipulation of the soil by machinery is an essential part of soil management, which includes such soil-building practices as grass and legume rotations, fertilization, and liming.

Mark L. Nichols

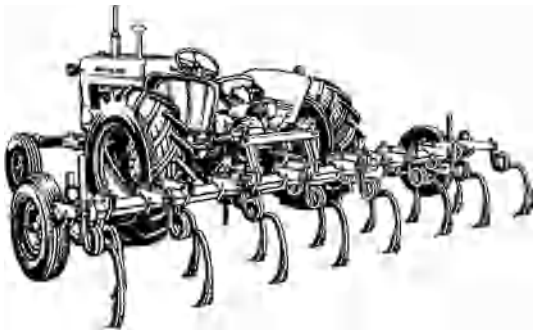

Fig. 4. Coil-shank field cultivator. (*Allis-Chalmers*)

Reduced tillage and no-tillage. These are land preparation practices that were used on over 31% of the cropland in the United States in 1992. These soil management systems involve limited or no physical disturbance of the soil prior to planting. Weed growth is controlled either by application of herbicides or by midseason cultivation. Residue from the previous crop remains on or near the soil surface; with more conventional clean tillage, essentially all crop residues are incorporated into the soil, usually by plowing or disking. The no-till planter is a multifunction device fitted with a disk that cuts through the stubble of a prior year's crops, making a furrow, a planter that drops a seed into the furrow, and wheels that close the furrow.

Advantages. Reduced and no-tillage practices are commonly used on highly erodible soils because they provide soil erosion control and greatly reduce fuel, time, and equipment costs, compared to clean tillage. Also, crops can be planted when the soil is wetter. Because of these advantages, the crop can be planted closer to the optimum planting date, potentially increasing crop yields. Leaving crop residues on the surface reduces water losses to runoff and evaporation, increasing infiltration rate and soil water storage by 1–2 in. (25–50 mm), which also improves crop yield potential. Improved soil erosion control maintains soil fertility and quality, thereby maintaining crop yield potential. Consequently, reduced-till and no-till production methods may result in greater net income because of lower production costs, greater yields, and diminished offsite damage.

Reduced-till and no-till systems have been used extensively only since about 1970, when suitable herbicides and equipment became widely available. Early research showed that when weed growth was controlled by herbicides and crop residues were left on the soil surface, soil water storage was improved by reduced weed competition and soil water evaporation rate. Also, soil productivity was maintained because erosion losses were controlled. These findings stimulated research to develop economical herbicide technology and suitable planting equipment. To a large extent, this objective was achieved during the 1970s and 1980s, permitting the widespread use of reduced-till and no-till management systems in agricultural production. *See* HERBICIDE.

Because of the presence of surface residues, the surface of no-till soil is moister and cooler than that

of bare soils. Surface residues also provide soluble carbon as an energy source for soil microorganisms. Consequently, because of the more favorable environment, the populations of all classes of microbes are greater on the surface of no-till soils than on bare soils. This results in more sustained and prolonged turnover of nitrogen, phosphorus, and other nutrients, which is often at a slower rate because of lower temperatures. Soil aggregation and soil porosity are usually improved by the enhanced biological activity resulting from the more favorable soil environment created by no-tillage. Greater earthworm activity in no-till soils also provides large-diameter soil pores that enhance infiltration and reduce runoff of water. Collectively, these advantages of reduced and no-tillage systems provide United States farmers with a useful tool for meeting the soil erosion limits on highly erodible land imposed by the conservation compliance requirements of the 1990 Farm Bill. As a consequence, acreage of reduced and no-tillage has increased greatly in recent years.

Disadvantages. Problems associated with reduced-till and no-till systems include herbicide costs, reduced soil temperature, inadequate weed control, poor drainage, soil compaction, disease and insect problems, and residual damages from herbicides. While these tillage systems require less investment in equipment, fuel, and labor than does clean tillage, they do require a substantial cash investment in herbicides. Each herbicide has its own spectrum of plant species for which it is effective; herbicide selection is therefore dictated by the crops to be produced and by the dominant weed species likely to be present. Most herbicides are degraded through microbiological activity or by volatilization. In warm wet years, the herbicide may rapidly degrade and later-germinating weeds may survive and grow, requiring additional expense for weed control. When the herbicide degrades slower than anticipated (cool or dry years), residual herbicidal activity may damage the following crop.

Special situations. Reduced-till and especially no-till production methods are poorly suited for slowly drained soils because no-tillage increases soil water content, thereby intensifying drainage problems. This is especially true in more northern climates where the lower soil temperature plus greater soil water content associated with no-tillage may result in delayed crop establishment. These problems can be offset to some extent by using ridge-till systems, wherein soil ridges 6–10 in. (15–25 cm) high are built during cultivation of the previous crop. At planting time, the top few inches of the ridge are leveled off, providing a well-drained, warm, weedfree, flat zone on the ridge top for the crop seedlings. Weeds between crop rows are controlled by tillage. Ridge-till systems have become very popular in the midwestern and southeastern United States.

Use of no-till systems may reduce need for phosphorus fertilizers because much of the organic phosphorus in crop residues is made available to the next crop through microbial processes. Fertilizer nitrogen requirements for no-tillage may be slightly greater than for clean tillage because the improved soil water supply may provide a greater crop growth potential. More nitrogen may also be temporarily tied up in the microbial biomass of no-till soils. This nitrogen often becomes available to the crop later in the growing season because the surface residues for no-till soil keep the surface soil moist longer, allowing microbial activity to continue longer into dry summer months. Although these systems conserve more soil water and increase water infiltration, the quantity of soluble nitrogen in no-till and reduced-till soils is usually less than in clean, cultivated soils. Thus, these tillage practices have little effect upon the leaching of nitrates to ground-water aquifers. *See* FERTILIZER.

J. F. Power

Soil and tillage interactions. Tillage is the use of a tool to change the physical condition of the soil. Many times the tool is part of a machine powered by a tractor, but it can be as simple as a hoe or shovel. A tillage tool affects the soil by cutting, shattering, packing, mixing, or inverting it. Normally, the objective of tillage is to improve the condition of the soil or to control weeds so that crop seeds will germinate quickly after planting and the resulting crop plants will produce to their potential.

Tillage can create short-term improvements in a soil condition by lowering bulk density by cutting and shattering the soil mass, but that may lead to long-term effects such as loss of stable structure. For example, a soil that has not been tilled recently often has a more stable structure than soil in the tilled field next to it. Better structure means that the basic particles (sand, silt, and clay) are grouped into aggregates with space around them for the movement of air, water, and roots. Stable structure means that the structure will resist forces such as tillage or wetting and drying without having the aggregates break into smaller pieces. The soil that is tilled may form a crust with these smaller pieces after a rain that will cause the next rain to run off or impede air transfer or plant growth. A fence row might maintain an open structure, allowing adequate air flow and water between its stable aggregates.

The use of a tillage tool depends on the soil type and the condition of the soil at the time it is tilled. Maps have been produced which cover much of the crop-producing areas of the United States and other developed countries. The areas of soil that are similar in properties are grouped and identified by names that specify texture as well as other attributes that tend not to change over time. Soil textures range from coarse (for example, sandy soil) to very fine (clay soil).

Soil tilth is a qualitative term describing the physical condition of soils at any time. It indicates the ease of tillage, seedbed preparation, seedling emergence, and root growth. It may also indicate the ability of the soil to resist erosion. The best condition for seeding wheat may not be the best condition for growing cranberries. Soil quality is dependent upon the ability of a soil to function in its immediate environment for a particular use and interact positively with the general environment.

A proposed Tilth Index describes the physical condition of the soil by combining numerical values for the density of the soil, the resistance of the soil to penetration by a steel cone, the organic matter content of the soil, the slipperiness of the soil when it is wet, and the size distribution of aggregates of soil particles. The Tilth Index combines soil physical condition indicators by multiplication after converting them to numbers between 0 and 1.

An ideal management system would use the Tilth Index, a soil quality index, or some other quantitative measure of the soil condition as a benchmark for deciding whether tillage is necessary. By using this approach, the condition of the soil could be determined at any time prior to planting, and tillage would be done only when benefits outweighed direct costs, such as fuel, and indirect costs, such as increased soil erosion potential. Tillage that has been done to change the physical condition of the soil may have been more important for its impact on the crop residue cover of the soil.

There is evidence that it may be more important to manage the residue cover than to change physical condition of the soil. In the northern regions of the United States corn belt, maintaining crop residues on the soil surface can reduce soil temperatures and negatively affect growth and development of subsequent crops. In Iowa and Minnesota, it has been difficult to prevent lower yields of corn in a rotation when a moldboard plow was not used to incorporate crop residue. In Ohio, on well-drained soils, there has been a yield advantage to leaving the crop residue on the surface without tillage prior to planting. The increased potential for harboring and transmitting plant pathogens in surface residues is a concern that is being evaluated through research programs because of the probability of increased area of crops produced with prior crop residue on the soil surface in the middle to late 1980s. This change is due to government efforts to reduce soil erosion in the United States.

The force required to perform tillage is often influenced by soil type and condition and by implement speed. When this force is provided as a pull by animals or tractors, it is often called draft. A plow equipped with high-speed moldboards, coulters, and a landside working at 4 mi/h (6.5 km/h) might require a 4-lb/in.^{2} (2.5-newton/cm^2) cross section of cut sandy loam soil. This cross section is also called the furrow slice. This would be compared to 6 lb/in.^{2} (4 N/cm^2) in loam or 13 lb/in.^{2} (9 N/cm^2) in silty clay. The draft decreases approximately 10% for each 1% increase in moisture content. Increasing soil strength by increasing the density by compaction or by planting a crop such as alfalfa, because of tough roots, will increase the draft. Increasing the speed of the plow by 50% will increase the draft by 25% in loam soil.

The draft of a disk plow is not affected by soil texture as much as that of a moldboard plow. The draft of subsoilers, chisel plows, and field cultivators depends on the depth and speed of the operation and is normally expressed as a value per shank. The reason is that the shanks may be spaced at different widths and the soil is not worked to the same depth across the width of the machine. The draft of a disk harrow is related to implement weight as well as soil texture. A disk harrow weighing 120 lb/ft (177 kg/m) and operating in sandy loam soil requires a draft force of 96 lb/ft (1380 N/m) of width. This requirement would increase by 50% if the disk were used in silt loam, or by 90% in clay.

Fuel requirements for tillage vary with implement draft and traction conditions for the tractor. Heavy draft primary-tillage tools, such as plows or subsoilers, require more fuel per unit of land area (acre or hectare) than other tillage tools. Fuel requirements for secondary tillage tools (such as disking) are influenced by soil conditions created by primary tillage in addition to soil type. If the tractor is on loose soil, it experiences more motion resistance than if it is operating on a firm soil. Properly matching tractor power to load size is important in conserving fuel. It is important to conserve fuel in crop production, but more energy is usually required in corn production to manufacture nitrogen-rich fertilizer and the fuel used for drying the crop and for irrigating some soils than the combined fuel used directly by machines to till, plant, and harvest.

Matching the tractor to the load includes consideration of whether the tractor is two- or four-wheel drive or uses tracks. The weight of the tractor as it is prepared for the field can influence slippage. Slippage of more than 10–15% with a wheeled tractor usually wastes fuel and can be corrected by adding weights to the tractor. Too much weight can shorten the productive life of the tractor. *See* SOIL.

Thomas S. Colvin

Harvesting

The practice of severing and reaping the plant or any of its parts is called harvesting.

Crops harvested for grain. The process of gathering such crops as corn, sorghums, wheat, oats, barley, rye, buckwheat, and rice is called grain harvesting. Ear corn is harvested by means of a corn picker (**Fig. 5**). The ears are snapped off by specially designed rollers which pass around the standing stalks. The husks are removed by a husking bed consisting of rolls of various types, over which the ears are passed.

Fig. 5. Corn picker. (*New Idea Farm Equipment Co.*)

Fig. 6. Forage harvester chopping haylage for livestock feed. (*Sperry New Holland, Division of Sperry Rand Corp.*)

Fig. 7. Mowing machine. (*Massey Ferguson Co.*)

Shelled corn is harvested by a picker-sheller or by the combine harvester. The picker-sheller snaps the ears from the stalks in the same manner as the picker, but the husking bed is replaced by a shelling unit. A trailing shelling unit attached to the picker can also be used.

Corn is also harvested by a self-propelled combine harvester. The header can be removed and replaced with a snapping unit, or the header can remain in place with the whole plant passing through the machine. Grain sorghums and cereals are harvested largely with the combine harvester, a machine that severs the standing crop, shells the grain, separates grain from straw, and removes chaff and trash in one operation. Sometimes these crops are severed and windrowed, allowed to dry, and threshed later.

Grain crops harvested for ensilage. This operation is used for corn, sweet sorghums, and cereals such as oats, wheat, barley, and rye.

Row crops, such as corn and some sorghums, are harvested with a forage harvester equipped with a row-crop attachment. High-moisture corn may be shelled by a picker-sheller and stored as ensilage.

Drilled crops and some row crops are harvested with the forage harvester equipped with a sickle-bar attachment (**Fig. 6**). Rotary-type harvesters are also used. The plants are severed at or near the soil surface and cut or shredded into short lengths.

Crops for silage, soilage, and hay. This type of harvesting is used for legumes (other than edible-podded legumes) and grasses (excluding corn, sorghums, and cereals). The sickle-bar or rotary-type forage harvester is used. It severs the crop near the ground surface and chops it into desired lengths for silage or soilage. It also chops hay from the windrow. The crop may be wilted slightly in the windrow to reduce moisture for silage preservation. The conventional mower, or the mower-crusher designed to speed up drying, is used to harvest crops for hay (**Fig. 7**).

Legumes and grasses for seed. Legumes and grasses are harvested largely by the combine harvester, either by direct or windrow methods (**Fig. 8**). Windrowing becomes necessary when the crop fails to ripen evenly. Because some seeds are lighter than cereal grains, machine adjustments differ widely. To increase overall efficiency, two combines may be hooked together in tandem. All straw and chaff from the lead combine passes through the rear one.

Podded legumes which are harvested include soybeans, dry edible beans, and peas. Soybeans are harvested exclusively by the combine-harvester direct method. Peas and beans may be harvested by the combine harvester or by bean threshers with multiple shelling cylinders (**Fig. 9**). In many cases, beans or peas may be removed or severed from the soil and windrowed prior to threshing. To prevent the cracking of seeds, cylinder speeds are reduced and concave clearance increased. Rubber-covered rolls, placed ahead of the cylinder, may be used to squeeze beans from pods.

Fig. 8. Combine-harvester. (*International Harvester Co.*)

Fig. 9. Bean thresher. (*C. B. Hay Co.*)

Root crops. Harvested root crops include sugarbeets, potatoes, and peanuts. Sugarbeets are gathered by special harvesters. One type tops the beets in place, after which the beets are lifted by specially designed blades or fingers. Another type lifts the beets by gripping the tops or by impaling the beets on a revolving spiked wheel (**Fig. 10**). The beets are then topped in the machine. An elevator conveys the beets to trucks for bulk handling.

Potatoes are harvested by several methods. They may be (1) dug with a one- or two-row digger and picked up by hand; (2) dug, sorted, and placed into containers by machine (vines, trash, and clods are removed mechanically and by workers riding the machine); (3) harvested by a fully mechanized procedure which includes digging, sorting, removal of vines, trash, and clods, and loading into bulk trucks (**Fig. 11**); and (4) dug with a standard digger, windrowed for drying, and picked up later by an indirect harvester. Sweet potatoes are harvested largely by the first method.

Peanuts are harvested by the following methods: (1) the pole-stack method, in which the peanuts are dug and hand-shaken, stacked around poles, and later picked; (2) a method in which they are dug with a one- or two-row digger, windrowed, and harvested later with a peanut picker (**Fig. 12**); and (3) the once-over method, in which all operations are accomplished with the same machine.

Crops harvested for fiber. Cotton is harvested by two methods: (1) pulling lint from the bolls by means of the cotton picker, a method which requires several pickings and is accomplished by broached spindles revolving rearward on which the lint is wound; and (2) pulling the entire boll from plants by a cotton stripper (**Fig. 13**), a once-over method accomplished by rolls of various types (steel, rubber paddle, or brush) which strip the bolls from the plants.

Special crops. There are several crops requiring special methods to be harvested. These include tobacco, castor beans, and sugarcane. Tobacco is harvested by two general methods. (1) Leaves may be primed from the stalks as they mature, in several primings starting with the lower, more mature, leaves; or (2) the stalks may be severed near the base (**Fig. 14**), upended, and speared, after which laths are inserted to handle the plants and support them in the curing barn. Machines have been developed

Fig. 10. Detail of spiked-wheel harvester gathering sugarbeets. (*Blackwelder Manufacturing Co.*)

Fig. 11. Potato harvester. (*USDA Agricultural Research Service, Red River Valley Potato Research Center*)

Fig. 12. Peanut digger-shaker-windrower. (*Department of Agricultural Engineering, North Carolina State College*)

Fig. 13. Cotton stripper. (*Allis-Chalmers*)

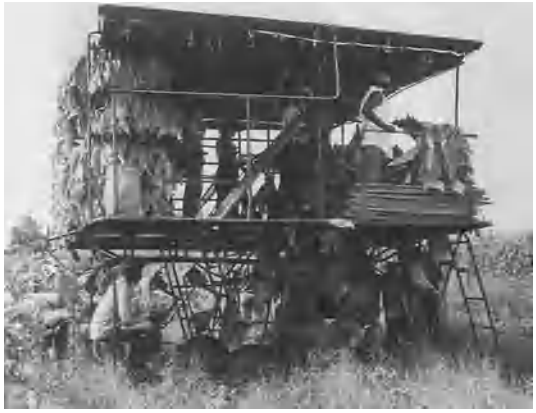

Fig. 14. Tobacco harvester. (*Department of Agricultural Engineering, North Carolina State College*)

Fig. 15. Sugarcane harvester. (*Agricultural Engineering Department, Louisiana State University*)

to speed up the priming process; workers ride rather than walk.

Castor beans are harvested by special machines that straddle the rows, with revolving beaters that strip the beans from the standing stalks.

Sugarcane is harvested by self-propelled harvesters which sever the cane at or slightly below the soil surface (**Fig. 15**). Additional knives top the cane. Tops and trash are removed by fans and distributed over the soil. Conveyors move the cane into heap rows or directly into trucks or wagons. Some machines cut the cane into short lengths for easier handling and processing. Edward A. Silver; R. B. Musgrave

Drying and Storage

Farm crops may be harvested at the most desirable stage of maturity and stored for weeks or months if properly dried or preserved. Field drying is an inexpensive procedure. However, in cool, humid areas a fine crop may deteriorate to a low-value feedstuff if it is damaged by rain during field drying; loss of quality may also occur as a result of mold or spontaneous heating in storage. To reduce such losses in forage crops, many farmers partially cure grasses or legumes in the field and then finish drying them in hay mows, bins, special wagons, or drying buildings by passing heated or unheated air through the forage with a power-driven fan unit attached to an air-duct system (**Fig. 16**).

Forage. Forage containing 70–80% moisture at harvest time is field-dried to 30–40% and finish-dried to 22–25% for safe storage as chopped or long hay, or to 20% as baled hay. Hay processed in this manner is superior to field-dried material in color, carotene content, and leafiness.

Because hay quality is to some extent a function of the rapidity of drying, heated air usually produces the best product. Very rapid drying can be accomplished with dehydrating equipment which can dry material from 75 to 10% moisture in 20 min or less. The quality of a dehydrated product is high; however, costs of labor and fuel are also high. Alfalfa to be used in mixed feed is the most frequently dehydrated crop (**Fig. 17**).

Field drying can be accelerated by the use of crushing machines which crack or shred the freshly cut forage as it passes through one or more pairs of crushing rollers (**Fig. 18**). Overall drying time is shortened because the stems, if crushed, will dry almost as fast as the leaves which, if the hay was improperly dried, often shatter and drop off.

Small grains and shelled corn. Small grains and shelled corn are dried to 12% moisture or less in either continuous or batch driers and stored in bins. Ear corn stored in an open crib must be dried to 16% moisture if mold growth in storage is to be prevented. Generally, temperatures of drying should not exceed 100°F (37°C) for seed and malting grains, 130°F (54°C) for milling corn, and 200°F (93°C) for feedstuffs (**Fig. 19**). Frequently rice drying is carried on in two stages to prevent cracking.

Ensiling. The anaerobic fermentation process of ensiling is used to preserve immature green corn, legumes, grasses, and grain plants. The crop is chopped and packed while at about 70–80% moisture and put into cyclindrical tower-type silos, horizontal trenchlike structures, or other containers to exclude the air. This tightly packed, juicy material is preserved by proper bacterial fermentation. Desirable microorganisms in grass and legume silage can be encouraged by field-wilting the crop to 70% moisture, or by adding chemical preservatives, sugar, and starch materials, such as ground corn or small grain.

Fig. 16. Special wagon rack attached to portable oil-burning crop drier for finish-drying chopped forage.

Fig. 17. Portable alfalfa-dehydrating equipment.

Fig. 18. Forage crusher for cracking or shredding stems to accelerate field drying rate of forage crops.

Fig. 19. Perforated metal bin for drying small grain and shelled corn with heated air from portable crop dryer.

Shelled or chopped ear corn is occasionally stored under similar anaerobic conditions. Hjalmar D. Bruhn

Processing Crops

Crop processing involves such operations as shelling, cleaning, separating, sorting, washing, treating, scarifying, testing, grinding, and ginning.

Shelling. The separation of corn kernels from the cob or the removal of the shell from nuts such as peanuts, walnuts, and hickory nuts is called shelling. This can be done with two types of machines, the spring type in which the kernels are rubbed from the ears by a notched metal bar called a rag iron, and the power-driven cylinder-type machine into which the ears are fed between a revolving shelling cylinder and stationary bars called concaves (**Fig. 20**). The kernels are rubbed off and separated before the cobs are removed. Proper shelling is obtained by control of the rate of feeding, tension of the shelling bar, choice of shelling concave, cleaning-air control, and cob-outlet control. The practice of picking high-moisture corn in the field and shelling it with a picker-sheller or combine-sheller is increasing.

Cleaning and separation. These procedures include the removal of foreign material, such as weed seeds, chaff, dead insects, and broken stems. The fanning mill, consisting of two vibrating screens and an air blast, is used on the farm for cleaning (**Fig. 21**). The most common methods of cleaning are by size, using a screen; by length, using a cylinder or disk with indented pockets to accept only short, small grain; by specific gravity, using a vibrating screen or inclined deck through which air is blown to remove the light material from the top; and by brine solutions of such density as to float light material and permit heavy material to settle. Seeds which become sticky when wet are separated from other seeds (for example, buckhorn seed from cloverseed) by moistening of the seed surfaces with wet iron filings or sawdust. The wetted seeds and iron filings or sawdust stick together, forming large clumps which are then removed by screening. Buckhorn seeds are removed in this manner from clover seeds. Seed shape can also be used as a means of separation because the round seeds roll and the flat seeds slide. Smooth and rough materials may be separated by using velvet- or velveteen-fabric-covered rolls, or by air in which the materials with irregular and rough surfaces are removed from those with smooth surfaces. Fruits may be cleaned by brushing, wet or dry, and by flailing.

Sorting. The separation of products into prescribed standards is called sorting. Grading is sorting to meet state and federal agency regulations. Grading, particularly of fruits and vegetables, is done by machinery to obtain the proper size or weight. Size grading is practiced to obtain desired diameter by

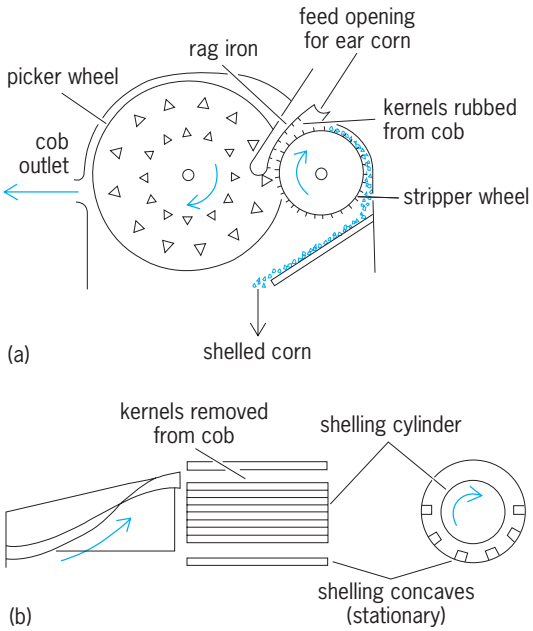

Fig. 20. Operation of two types of corn shellers. (*a*) Spring. (*b*) Cylinder.

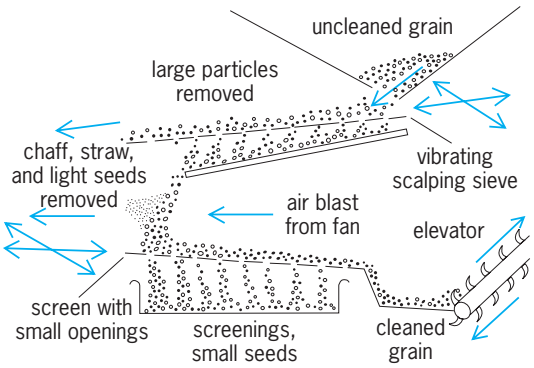

Fig. 21. Fanning mill operation.

using a belt with holes, a link chain, a diversion belt, diversion rollers, or spool units. Apples, peaches, potatoes, onions, lemons, and oranges are commonly sorted by size. The weight method is sometimes used to place products in appropriate grades. For beans, coffee, and lemons, color sorting is accomplished by scanning with an electric-eye device which rejects the materials that possess undesirable light-reflecting properties.

Washing. Foreign materials are removed from fruits, seeds, nuts, and vegetables by washing, often with a detergent and chlorine solution. Washing is done by soaking in tanks, spraying with high-pressure nozzles, or moving the materials through a washing solution in a rotary cylinder.

Treating. Seeds are treated for disease or growth control. Treating was formerly performed by heating the seeds in water, but now it is usually done with a chemical. The treating material may be applied as a chemical slurry, dust, liquid, or vapor. Mercury compounds are often used, but seed treated with mercury cannot be used for animal or human feed. Legume seeds are often inoculated with bacteria which convert the nitrogen of the soil air into a form which is suitable for use by the plant.

Scarifying. This process, usually preceded by hulling, is an operation in which hard seed, such as that of legumes, is scratched or scarred to facilitate water absorption and to speed germination.

Testing. Usually testing is performed on batches of seeds to determine the purity, moisture content, weight, and germination potential. The major problem in testing is to obtain a representative sample. The moisture content of grains and seeds indicates the keeping quality in storage and is a basis for estimating their commercial value. Weight per bushel is an important commercial criterion used for grain and is known as the test weight.

Grinding. Reduction in size of the material to improve utilization by animals and to make handling easier is called grinding. Measurement of the fineness of the ground material is designated by fineness modulus numbers from 1 (fine) to 7 (coarse). Units for size reduction include the hammer mill, a rotating high-strength beater which crushes the material until it will pass through a screen mounted above, below, or around the rotating hammer; the burr or attrition mill, in which two ribbed plates or disks rub or crush the material between them; and the roller mill, in which grain passes between pairs of smooth or corrugated rolls (**Fig. 22**). The last is used extensively for flour manufacture. The crimper-roller is used on farms to crush grains. The roller-crusher is also used to reduce the size of ear corn before it is fed into the hammer mill or burr mill.

Ginning. Separation of lint from cottonseed is called ginning. The cotton gin cleans the cotton and seed in addition to separating them. The saw gin, which consists of about 100 saws per gin stand mounted about $^5/_8$ in. (1.5 cm) apart on a shaft, is the most commonly used. The 12-in.-diameter (30-cm) saws are made up of spikes on wheels which

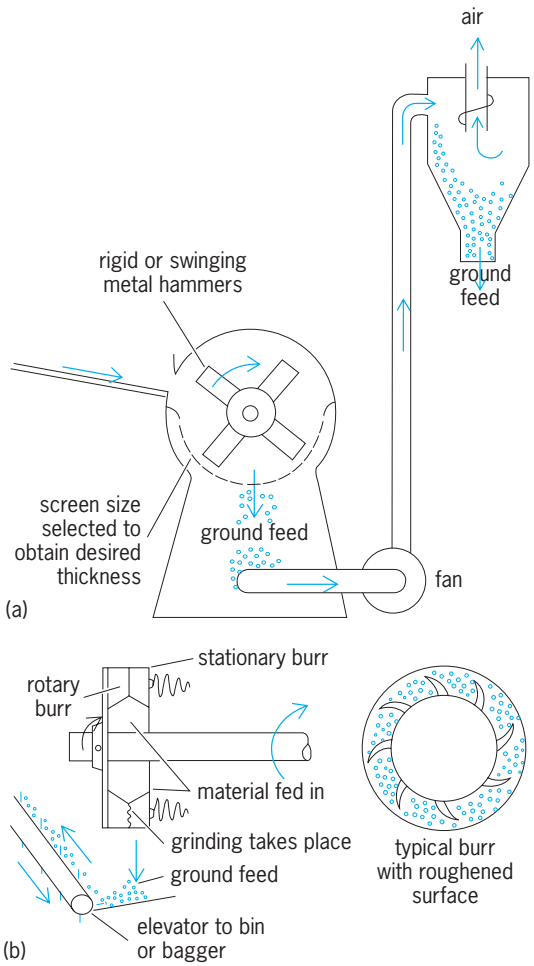

Fig. 22. Operation of two machines for grinding. (*a*) Hammer mill. (*b*) Burr mill.

pull the cotton through huller ribs, but the seed cotton (cotton containing seed) is not pulled through. The lint is removed from the saw blades with air or with a high-speed brush. *See* AGRICULTURAL ENGINEERING; AGRICULTURAL MACHINERY; AGRICULTURAL SCIENCE (PLANT); COTTON; FOOD MANUFACTURING.

Carl W. Hall

Sustainable Agriculture

Sustainable agriculture represents an integration of traditional, conservation-minded farming techniques with modern scientific advances. This system maximizes use of on-farm renewable resources instead of imported and nonrenewable resources, while earning a return that is large enough for the farmer to continue farming in an ecologically harmless, regenerative way. Agriculture in the United States is mostly conventional, and thus heavily dependent on the use of fossil-fuel-based inputs— inorganic fertilizers, pesticides, and labor-saving but energy-intensive farm machinery.

Conventional farming systems have greatly increased crop production and labor efficiency, but serious questions are being raised about their energy-intensive nature and their adverse effects on soil productivity, environmental quality, farm profitability, and human and animal health. This concern has

led to an increasing interest in sustainable farming systems, because they have the potential to reduce some of the negative effects of conventional agriculture.

Sustainable agriculture emphasizes use of natural, on-farm, nonchemical approaches to pest control that do not pose health threats to humans or decimate populations of earthworms and other beneficial life forms. Compared to conventional agriculture, sustainable agriculture improves wildlife habitat by reducing the possibility of pesticide poisoning and by increasing crop diversity. The use of crop rotations and green-manure legume crops in sustainable agriculture on sloping, erodible soils generally produces much less erosion than conventional agriculture.

Sustainable agriculture also emphasizes use of farm profits to finance the ownership and operation of the farm, rather than depending heavily on borrowed capital. It increases net income by reducing inputs, rather than by increasing output. Conventional farmers wanting to become sustainable farmers do not merely terminate pesticide and fertilizer use; they replace petrochemical inputs with a variety of practices, such as crop rotations, growing legumes or cover crops, use of green manures, and livestock enterprises.

More sustainable agriculture technologies are being developed and adopted, featuring reduced dependence on chemical technologies as well as making greater use of soil-conserving methods of growing crops. Farmers who consider their farming methods to represent organic, regenerative, reduced-input, biological, or other sustainable approaches are employing some of these increasingly popular farming technologies, in addition to a number of less conventional methods. Many of these farmers report that their costs of production are lower than those of their neighbors who use conventional methods. Some farmers report lower yields than their neighbors, but the sacrifice in yield is often more than offset by cost reductions.

J. P. Reganold; R. I. Papendic; J. F. Parr

Solarization

Soil solarization is a hydrothermal process involving solar radiation that disinfests soil of plant pathogens and weed pests and also effects other changes in the physical and chemical properties of soil that lead to improved growth and development of plants.

Solarization of soil is accomplished by placing a sheet of plastic on moist soil for 4–6 weeks during the warm summer months. The plastic sheeting must be laid close to the surface of the soil to reduce the insulating effect of an air layer under the plastic; the edges of the plastic are buried or anchored in the soil. Solar radiation then heats the soil, and the resulting heat and the moisture in the soil are conserved by the plastic cover. Over time, the heat and moisture bring about changes in soil microbial populations, release of soil nutrients, and changes in the soil structure that are beneficial to plants subsequently established there.

Technology. Soil solarization is a mulching process, and it may seem to resemble the practice of using plastic mulches to warm soils to increase seed germination and seedling growth in early spring months. However, soil solarization differs from such use of plastic in that killing soil temperatures (approximately 97°F or 37°C or higher) are attained for many soil-borne organisms, seeds and youg plants or seedlings.

Soil solarization has been especially useful in regions with high solar radiation where daily temperatures in summer months often reach 90°F (32°C) or more; it is a nonchemical, easily applied method that is of little or no danger to workers; it embodies the objectives of integrated pest management and reduces the need for toxic chemicals for plant disease and pest control.

Soil preparation. The area or field to be solarized should be disked, rototilled, or worked by hand to provide a smooth surface so that the plastic sheeting will lie flat. For crops grown on-the-flat or for solarization of areas such as lawns, the edges of the plastic sheeting are anchored or buried about 8 in. (20 cm) deep. For bedded-up soil, the plastic sheeting is laid over the length of the bed, and the edges are buried in the furrows. For large-scale use of solarization on flat or level soil, after the first sheet is laid and anchored in the soil, additional sheets are glued together on one side and anchored in the soil on the other side. The ends of the plastic sheeting are also anchored in the soil.

Soil moisture. Soil solarization is best accomplished when the soil is moist (at least 70% of field capacity), a condition that increases the thermal sensitivity of soil-borne pathogens and pests and also enhances heat conduction within the soil. When conditions permit, plastic sheeting may be laid on dry soil, and then irrigation water can be run under the plastic in order to wet the soil to a depth of at least 24 in. (60 cm). If the field or area is level, shallow furrows made by tractor wheels during the laying of plastic sheeting provide convenient courses for the water. If the area or field is not level, the soil should be preirrigated and the plastic sheeting then laid within a few days, or as soon as the soil will support machinery without compaction. When soils are bedded up before solarization, the plastic is laid over the bed and anchored on each side in the furrows. Irrigation water is then run in the furrows, and the moisture subs into the raised beds (moves by capillarity).

Plastic film. Plastic sheeting used in soil solarization provides a barrier to losses of moisture and heat from soil due to evaporation and heat convection. Clear or transparent polyethylene plastic sheeting is more effective for heating soil than black plastic, since the transparent plastic allows greater transmission of solar energy; it gives maximum transmittancy of radiation from wavelengths of 0.4 to 36 micrometers. Other plastics are also effective, and in some instances superior to polyethylene. Colored transparent plastics tend to reduce the formation of water droplets on the underside or soil side of the plastic sheeting that may act as reflectors of radiation.

Polyvinyl chloride films also have been used effectively for soil solarization in greenhouses.

The thickness of the polyethylene sheeting also affects the heating of the soil; film 25-μm thick results in temperature increases of soil several degrees higher than thicker films that reflect more solar energy. Polyethylene sheeting used for soil solarization requires the presence of an ultraviolet inhibitor to prevent its deterioration by sunlight.

Soil temperature. Day length, air temperature, cloud cover, and the darkness of soils all have marked effects on the heating of soil during solarization. The ideal time is during the summer months, when the days are longest and air temperatures are highest. Maximum temperatures in the upper 6–12 in. (15–30 cm) of soil are reached 3–4 days after solarization begins. Depending on soil depth, maximum temperatures of solarized soil are commonly between 108 and 131°F (42 and 55°C) at 2 in. (5 cm) depth and range from 90 to 97°F (32 to 36°C) at 18 in. (46 cm) depth. Dark soils reach higher temperatures during solarization than lighter-colored soils because of their greater absorption of solar radiation.

Modes of action. Soil solarization encompasses changes in the biological, chemical, and physical properties of soil, changes that are often interdependent, dynamic, and evident for two or more growing seasons. Reduction in plant diseases and pests, coupled with better soil tilth and increases of mineral nutrients in solarized soil, contributes to the improved growth response of plants and increases in the quantity and quality of crop yields.

Biological effects. Soil microorganisms whose growth and development are inhibited at temperatures of about 86°F (30°C) or higher include most plant pathogens and some pests. The death rates of these organisms in moist soil at temperatures of about 97°F (36°C) or higher increase linearly with temperature and time. In contrast, other microorganisms in soil are thermotolerant, and their populations decrease less during solarization. Thermotolerant organisms include Actinomycetes and *Bacillus* species and certain mycorrhizal fungi; especially in the absence of soil-borne pathogens and pests, they are often highly beneficial to plant growth. In some soils, solarization causes biological changes that suppress the reestablishment of pathogens.

Chemical effects. Among the chemical changes that occur in soils during solarization are significant increases in soluble nutrients, such as nitrogen, calcium, and magnesium, which are associated with the increased growth response of plants grown in solarized soils.

Physical effects. The physical effects of soil solarization are largely unknown, but marked improvement in soil tilth occurs. During solarization, soil moisture tends to rise to a cooler soil surface at night, whereas during the day the moisture tends to move toward a cooler temperature zone deeper in the soil. This diurnal cycling of soil moisture is believed to leach sodic layers and thus accounts for the reduction of salinity caused by solarization in some soils.

Applications. Soil solarization is a safe, nonchemical, and effective method for plant disease and pest control. When used in conjunction with agricultural chemicals and biological control agents, soil solarization can be extended to regions where environmental conditions are less suitable for it. Use of solarization is increasing in orchard, field, and truck crop farming as well as in greenhouse and nursery operations; it is also finding increased use for garden and landscape improvement. *See* IRRIGATION (AGRICULTURE); MYCORRHIZAE; SOIL; SOIL ECOLOGY; SOIL MICROBIOLOGY; SOLAR RADIATION. James E. DeVay

Bibliography. D. Baize, *Soil Science Analysis: A Guide to Current Use,* 1993; Deere and Company Staff (eds.), *Soil Management: FBM Farm Business Management,*1993; J. E. DeVay and J. J. Stapleton (eds.), *Soil Solarization,*1991; R. E. Phillips and S. A. Phillips, *No-Tillage Agriculture: Principles and Practices,*1984; K. A. Smith, *Soil Analysis: Physical Methods,*1990; M. A. Sprague and G. B. Triplett, *No-Tillage and Surface-Tillage Agriculture,*1986.

Agriculture

The art and science of crop and livestock production. In its broadest sense, agriculture comprises the entire range of technologies associated with the production of useful products from plants and animals, including soil cultivation, crop and livestock management, and the activities of processing and marketing. The term agribusiness has been coined to include all the technologies that mesh in the total inputs and outputs of the farming sector. In this light, agriculture encompasses the whole range of economic activities involved in manufacturing and distributing the industrial inputs used in farming; the farm production of crops, animals, and animal products; the processing of these materials into finished products; and the provision of products at a time and place demanded by consumers. The proportion of the United States economy dedicated to the production of food and fiber products is about one-sixth of the total national product.

Many different factors influence the kind of agriculture practiced in a particular area. Among these factors are climate, soil, topography, nearness to markets, transportation facilities, land costs, and general economic level. Climate, soil, water availability, and topography vary widely throughout the world. This variation brings about a wide range in agricultural production enterprises. Certain areas tend toward a specialized agriculture, whereas other areas engage in a more diversified agriculture. As new technology is introduced and adopted, environmental factors are less important in influencing agricultural production patterns. Continued growth in the world's population makes critical the continuing ability of agriculture to provide needed food and fiber.

Agriculture in the United States

Agriculture is the means by which the resources of land, water, and sunlight are converted into useful

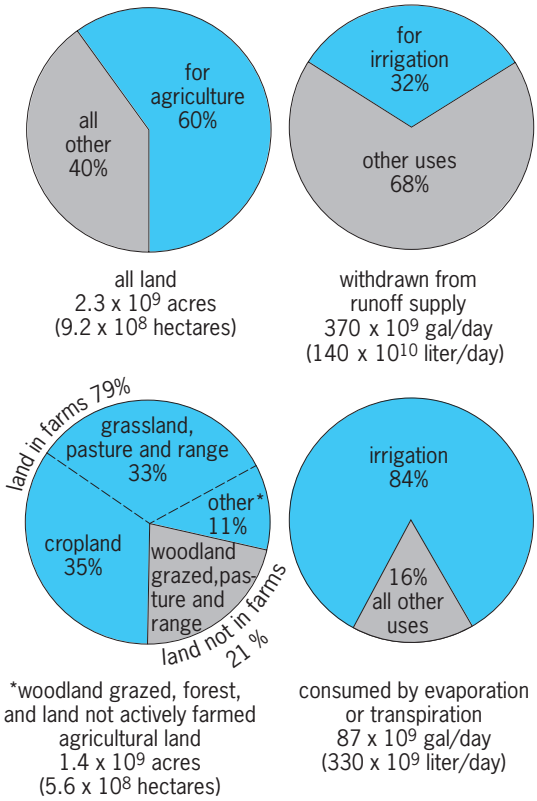

Fig. 1. Use of land and water resources in the United States. (*USDA*)

products. Land resources must be carefully managed to maintain and enhance the productivity of areas suitable for crops and for pasture and grazing, as well as to ensure effective use of rainfall so that the water needs of the nation will be met (**Fig. 1**). Of the 2.3×10^9 acres of land in the United States, about 1.3×10^9 acres (5.3×10^8 hectares) is used in agriculture, of which about 27% is cropland, 36% pasture and rangeland, and 25% forest. However, land-use categories are not static; pasture and rangeland may be shifted to forestry use, and as crop yields increase there is a net loss from cropland to less intense agricultural use. A significant portion of the land in pasture, range, and forestry can be considered as potential cropland, depending on agricultural demand and prices.

The United States has developed the most efficient agricultural system in the world, and this system continues to improve through scientific discoveries. Agriculture is increasingly a high-technology industry. There is now a tremendous array of powerful tools of modern plant science for the improvement of traditional crops through genetic manipulation and for finding alternative means of controlling pests (which could replace or greatly reduce pesticides that have environmental negatives).

The nation has evolved from a predominantly agricultural to a highly industrialized society (see **table**). The farm population grew in actual numbers up to 1940, but the percentage of the total population engaged in agricultural production steadily decreased as American farmers became more productive. A large population is engaged in supporting agriculture by providing agriculture inputs: tractors and trucks, machinery and equipment, petroleum products, fertilizers, pesticides and other agricultural chemicals, and other essential materials. A still larger group of people is engaged in agricultural industries that store, process, distribute, and sell farm products to the consumer. The number of people employed by these supporting industries is much greater than the number working on farms. The United States continues to depend on its lands and waters for the well-being of its people. In addition, the United States has become the world's greatest producer of foodstuffs for export to other nations.

Farm output has more than kept pace with the population growth of the nation. The output per worker-hour of farm workers has continued to increase because of the greater use of power and machinery, more effective uses of fertilizers to enhance crop productivity, use of higher-yielding crop varieties, more efficient feeding and management of livestock, and similar advances.

The output of farm products has increased more rapidly than inputs owing to the application of new knowledge and skills in farm operations. Labor costs have been reduced by the use of power and machinery; there have been great improvements in the use of fertilizers and other items such as chemical pesticides to control insects, diseases, and weeds, and in the use of livestock feed supplements. Total farm output and the output per unit of the various products that constitute inputs are closely related. Increases in American agricultural productivity are due to increased farmer efficiency and skill in applying agricultural science and technology. *See* AGRICULTURAL MACHINERY; AGRICULTURAL SCIENCE (ANIMAL); AGRICULTURAL SCIENCE (PLANT); FERTILIZER; FOOD MANUFACTURING; PESTICIDE.

American agriculture has provided a great diversity of food and fiber stuffs at lower costs than are enjoyed by most other nations.

Characteristics and trends. The skills and knowledge of the successful agricultural producer have changed drastically with the advent of tractors and trucks to replace horses, the development of more effective machines to replace much of the previous hand labor, the breeding of more productive

Chronology of American agriculture

Year	Total population, 10^6	Farm population, 10^6	Percent of total
1790	3.9	3.5	90+
1820	9.6	6.9	72
1850	23.2	11.7	50
1880	50.1	24.5	49
1910	92.0	28.5	31
1940	131.8	30.8	23.4
1950	151.0	25.0	16.5
1960	180.0	15.6	8.7
1970	204.0	9.7	4.8
1980	225.6	7.2	3.2
1990	247.8	4.6	1.9

crops and livestock, the development of chemical fertilizers to supplement native soil fertility, and the advent of chemical pesticides to control insects, diseases, and weeds. Increased processing has saved food products from spoilage and losses, and converted them into goods more readily utilized in meeting human needs. These agricultural industries now constitute a large segment of the national system for channeling agricultural products to consumers.

Nature of agricultural products. The primary agricultural products consist of crop plants for human food and animal feed and livestock products. The crop plants can be divided into 10 categories: grain crops (wheat, for flour to make bread, many bakery products, and breakfast cereals; rice, for food; maize, for livestock feed, syrup, meal, and oil; sorghum grain, for livestock feed; and oats, barley, and rye, for food and livestock feed); food grain legumes (beans, peas, lima beans, and cowpeas, for food; and peanuts, for food and oil); oil seed crops (soybeans, for oil and high-protein meal; and linseed, for oil and high-protein meal); root and tuber crops (principally potatoes and sweet potatoes); sugar crops (sugarbeets and sugarcane); fiber crops (principally cotton, for fiber to make textiles and for seed to produce oil and high-protein meal); tree and small fruits (apples, peaches, oranges, lemons, prunes, plums, cherries, grapes, and strawberries); nut crops (walnuts, almonds, and pecans); vegetables (melons, sweet corn, cabbage, cauliflower, lettuce, celery, tomatoes, eggplant, peppers, onions, and many others); and forages (for support of livestock pastures and range grazing lands and for hay and silage crops). The forages are dominated by a wide range of grasses and legumes, suited to different conditions of soil and climate.

Forest products include wood for construction material and panel products, fiber for pulp and paper, and extractives such as turpentine. Forests also serve the environment as a source of clean water and air, support of wildlife, and recreational and pleasant surroundings. Forestry, the culture of forestlands for human purposes, is often considered separately from agriculture, although the distinction between them is decreasing. It is clear that forest ecology interacts with all human condition and activities. *See* FOREST TIMBER RESOURCES.

Livestock products include cattle, for beef, tallow, and hides; dairy cattle, for milk, butter, cheese, ice cream, and other products; sheep, for mutton (lamb) and wool; pigs, for pork and lard; poultry (chiefly chickens but also turkeys and ducks) for meat and eggs; and horses, primarily for recreation.

In the United States, 65–90% of all feed units for cattle, sheep, and goats is provided by forage crops (grazing lands, hay, and silage), and the remainder from feed grains and other concentrated feeds. In most of the less-developed countries, such livestock subsist solely on forages.

Livestock play a vital role in using lands not suited for crops, to produce animal products that are indispensable to humans. However, commercially produced pigs and poultry are dependent on supplies of grains and concentrates. *See* BEEF CATTLE PRODUCTION; DAIRY CATTLE PRODUCTION; GOAT PRODUCTION; POULTRY PRODUCTION; SHEEP; SWINE PRODUCTION.

Exports. In the 1970s the world market for bulk commodities (maize, soybeans, and wheat) increased dramatically. In the 1980s the growth in the world market was for value-added products consisting of alternative and processed products, while the demand for bulk commodities, less than 10% of the exports from the United States, was flat. The United States is the world's largest exporter of agricultural products, principally grains and feed, oil seed and products, fruits and vegetables, cotton, and tobacco.

In addition to commercial sales of agricultural products, the United States has had a major role in providing foods and other supplies for relief and rehabilitation in developing countries. Such supplies have been sent as outright grants or as long-term loans without interest and with long grace periods. However, the reduction of grain surpluses in the United States reduced the nation's capacity to participate in large-scale food aid. Food give-away programs have a disastrous impact on the agriculture of developing countries and are, therefore, deemphasized except for famine conditions.

World Agricultural Conditions

The developing countries have had a mixed record of increasing their total food production in line with their increases in population growth. In South America increases in food production have kept up with population increases, while in Asia food production per capita increased as a result of the green revolution, the increased agricultural technology spearheaded by the introduction of improved wheat and rice. India, China, and Indonesia have made remarkable progress. Food production in Africa, however, has not increased, and per-capita production has actually declined in some countries, leading to grave problems. The application of science and technology to improve agriculture in Africa has great potential, but to be successful it requires changes in population control and political stability. The acquired immune deficiency syndrome (AIDS) epidemic in Africa is a complication that has affected long-term development there. In contrast, the developed countries in North America, western Europe, and Australia have continued to produce food surpluses.

In addition to the United Nations Development Program and the Food and Agricultural Organization (FAO) of the United Nations, there are various international agricultural research centers that are dedicated to solving the urgent agricultural problems in the tropics and subtropics. Agricultural development programs are also supported by the World Bank and by regional banks in Latin America and Asia. The bilateral assistance programs of the U.S. Agency for International Development is paralleled by bilateral programs of western European countries, Canada, and Japan. These programs deal not only with agricultural production and marketing but also with the

health, education, and institutional development of developing countries.

Environmental Concerns

A number of serious concerns impinged upon agriculture in the latter half of the twentieth century: rapid population increases, environmental deterioration, and energy price destabilization. These problems were shown to be part of a single issue—the distribution of the Earth's renewable and nonrenewable resources. Fear of emerging world food shortages caused by an unchecked population increase in the developing world did not lead to the dire consequences predicted earlier, largely because agricultural production increased in most, but not all, countries of the world. Populations in the developed world have reached a steady state.

Concerns over environmental safety as well as the impact of synthetic chemicals on human health dampened enthusiasm for chemical solutions to agricultural problems. This increasing concern resulted in enormous expense in the registration of new chemical materials because the agrochemical industry is forced to prove that substances are safe, a much more difficult task than to prove that a new substance is not harmful. Agricultural science has responded to worldwide concern by developing an integrated pest management (IPM) system which seeks a more rational basis for pest control through a combination of biological pest control strategies and a minimum use of pesticides, restricting use to those which are environmentally safe. Reducing agriculture's dependency on chemical solutions created new opportunities for genetic changes through both traditional and nontraditional techniques. However, the new genetics, known as genetic engineering, came under attack because of the unknown effects of foreign genes. Transfer of genes to crop species from other plant species has been accepted, but there is still considerable reluctance to accept the transfer of genes from animals or bacteria to crop plants.

Energy Costs

The United States food system uses about one-sixth of the energy in the nation, with food production accounting for one-fifth of this total. Thus, primary food production accounts for only about 3% of the nation's total energy demand. As shown in **Fig. 2**, it takes almost three times as much energy in home food preparation as in primary farm production. Notwithstanding, increasing concern over energy has generated various input/output accounting to calculate energy efficiency in agriculture. This has resulted in equating the energy in coal or oil with the energy in human labor or of food products—the results of which are clearly absurd.

The dramatic increases in energy costs will have large effects on the economics of various agricultural practices. A third of the energy used in actual agricultural production is related to the use of pesticides and fertilizers; other uses include field machinery, 23%; transportation, 18%; and irrigation, 13%

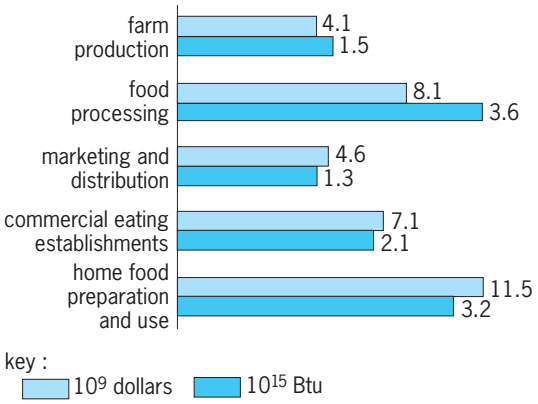

Fig. 2. Energy use in United States food system. (*USDA*)

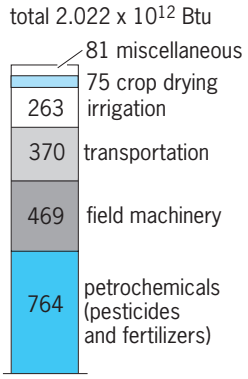

Fig. 3. Energy use in United States agricultural production. (*USDA*)

(**Fig. 3**). The cost/benefit ratio of many energy-dependent practices will change if energy costs increase out of proportion to other costs. This is clearly seen in the greenhouse industry, producer of flowers and some food crops. Greenhouse growers throughout the world have had to resort to various strategems to conserve energy, but present production has not declined, and the higher energy costs so far have been passed on to the consumer.

Jules Janick

Subsistence Agriculture

More than 500 million smallholders in the tropics farm under rain-fed conditions in diverse and risk-prone environments. In a constant struggle to survive, farm communities have developed innumerable ways of obtaining food and fiber from plants and animals. A wide range of different farming systems have evolved, each adapted to the local ecological conditions and inextricably entwined with the local culture. These traditional subsistence farming systems changed during the 1970s, 1980s, and 1990s primarily as a result of the expression of local creativity.

Subsistence farming. Subsistence farms comprise a patchwork of small fields containing mixtures of crops. Small farmers employ intricate farming systems to adjust to seasonal environmental changes,

marketing conditions, and the availability of family labor. More than 10% of the production is consumed directly on the farm's premises; there is little selling or trading. Subsistence farmers typically produce a great variety of crops and animals. For example, in Asian monsoon areas with more than 60 in. (1500 mm) of annual rainfall, it is not unusual to find as many as 20–30 tree crops, 30–40 annual crops, and 5–6 animal species on a single farm. Such multiple cropping methods are estimated to provide as much as 15–20% of the world's food supply. Throughout Latin America, farmers grow 70–90% of their beans with maize, potatoes, and other crops. Maize is intercropped on 60% of the region's maize-growing area.

Subsistence farming is characterized by diverse labor requirements. Having evolved to produce food year-round, the system provides continuous employment for unskilled labor to tend crops and livestock. *See* MULTIPLE CROPPING.

Typically, the subsistence farmer plants some commercial crops that entail relatively high risks. The farmer hedges against this risk by growing several less valued but more reliable crops, such as cassava or beans. In monsoon climates, with pronounced alternations of wet and dry seasons, the subsistence farmer ensures a stable production with long-duration root crops, tree crops, and animals. Although subsistence farming lacks the potential for producing a marketable surplus and thus supporting a higher standard of living for the farm family, it can often be adapted to increase productivity on more highly developed farms that are attempting to overcome resource limitations.

Uncultivated areas. Not only is diversity maintained within the area cultivated, but many peasants maintain uncultivated areas (such as forests, lakes, grasslands, streams, and swamps) in or adjacent to their fields, thus providing valuable products (including food, construction materials, medicines, organic fertilizers, fuels, and religious items). In humid, tropical conditions, collecting resources from primary and secondary forests is very intensive. In the Uxpanapa region of Veracruz, Mexico, peasants utilize about 435 wild plant and animal species, of which 229 are eaten. In many semiarid areas, gathering enables peasant and tribal groups to maintain their nutritional standards even during periods of drought.

Smallholder farming systems. Conventional agricultural development projects have demonstrated that although external inputs such as pesticides, fertilizers, and irrigation water can increase the productivity of peasant agriculture, such inputs can be maintained only at high financial and environmental cost. Moreover, they depend on the uninterrupted availability of commercial inputs—which is simply not viable given the level of impoverishment in most rural areas of the Third World. An agricultural strategy based on traditional cropping systems can bring moderate to high levels of productivity; and depending on prevalent socioeconomic and agroecological conditions, such diversified systems are sustainable at a much lower cost and for a longer period of time. *See* AGROECOSYSTEM.

Miguel A. Altieri

Remote Sensing

Information requirements of those who seek to manage the world's agricultural resources fall into five categories: the type of crop growing in each field; the state of health or vigor of the crop; the location and identity of crop-damaging agents or organisms; the crop yield likely to be obtained per unit of field area; and the total area of each field. Timely information in such areas would permit better crop management and increased production.

Agriculturists have used remote sensing to a significant extent since the mid-1930s. The only type of remote-sensing data ordinarily available until the late 1940s, however, consisted of small-scale, black-and-white aerial photos, most of which had been recently obtained for mapping purposes. Such photos were used by agriculturists during that period primarily in regional surveys for the annual identification of crops, the locating of fields, and the accurate delineation of field boundaries so that field acreages could be more accurately determined. In the 1950s it was found that early-warning evidence of certain crop diseases could be obtained through the use of a photographic film sensitive to radiation in the near infrared. Most of these applications continue to be made both in the United States and in many other parts of the world. *See* AERIAL PHOTOGRAPHY; PHOTOGRAMMETRY.

Space-acquired data. As operational space vehicles emerged, it became apparent that photos taken from a spacecraft in orbit at an altitude of 100 mi (160 km) or more above the Earth's surface had agricultural applications. The Johnson Manned Spacecraft Center conducted some highly successful photographic experiments from various crewed spacecraft, culminating in the *Apollo 9* flight in 1969. Furthermore, such vehicles offered the possibility of cost-effective monitoring of crop growth through the availability of repeated coverage of vast agricultural areas several times during each growing season.

Parallel developments in computer technology prompted consideration of computer-assisted analyses of the vast numbers of remote-sensing data. Computer applications required that some attribute of crops as imaged on a space photograph satisfy two conditions: the attribute must be of great diagnostic value in relation to the identification and monitoring of agricultural crops, and it must lend itself to being readily and unambiguously digitized.

Many photo image attributes had proved to be useful in the analysis by humans of agricultural crops on conventional aerial photographs. These included size, shape, shadow, pattern, tone, topographic site, and texture. However, only tone satisfied these two criteria. Furthermore, tone was the least degraded when remote sensing of agricultural crops was accomplished from a very high flying spacecraft rather than from a relatively low-flying aircraft.

During the 1970s progress was made in the development of a special kind of "camera" known as

a line scanner. As seen in **Fig. 4**, this device employs a very narrow-angle rotating prism which, as the remote-sensing vehicle advances, permits scene brightness to be observed and recorded for each small spot in an agricultural field. This digital record of tone can be telemetered directly and instantaneously to suitably located receiving stations on the ground and recorded there, pixel by pixel, on magnetic tape. Later, when the record is played back, it can be used to construct a complete swath of photo-like imagery of the entire scene that was overflown by the spacecraft, as shown in **Fig. 5**. Alternatively, the data can be fed directly into a computer which has been programmed to identify, pixel-by-pixel and field-by-field, the crop types and crop condition classes.

For this computer-assisted analysis to be successfully employed, there must be a unique tone or scene brightness associated with each crop type and condition class that is to be identified. A typical line scanner recognizes and correspondingly digitizes as many as 256 gray levels, each indicative of a different scene brightness so that even very subtle but unique differences in scene brightness can be utilized. Many crops have been found to exhibit similar digitized signatures, at least throughout the vegetative stage of their growth. This problem is remedied in part by use of a line scanner system known as a multispectral scanner, which simultaneously acquires the digital records of scene brightness in each of several wavelength bands. More crop types and condition classes are computer-identifiable from the resulting digital multiband signature than from any single band. Through use of the proper combination of bands, an object's subatomic, atomic, molecular, and macromolecular properties can all be sensed.

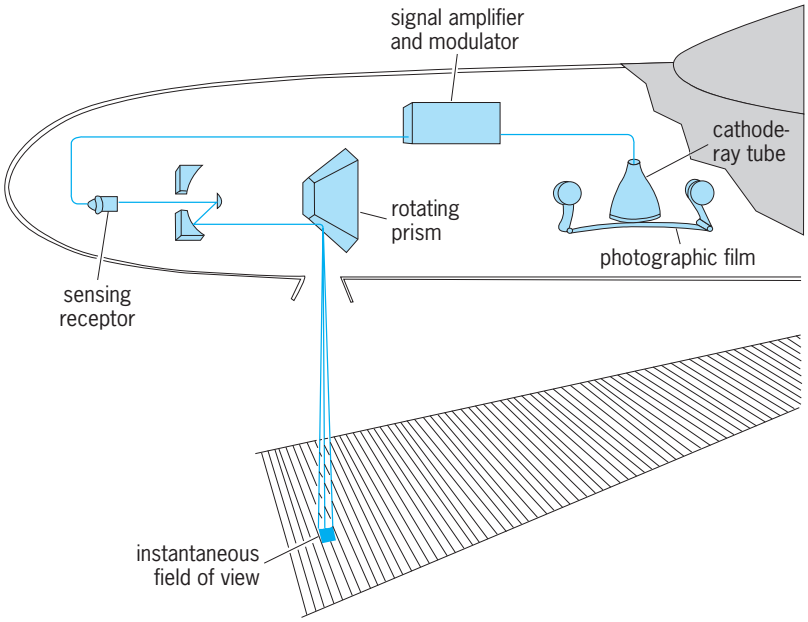

Fig. 4. Diagram of the operation of a multispectral scanning system.

In an effort to realize all of the benefits that might come from space-acquired remote-sensing data, NASA developed a crewless spacecraft for the specific purpose of repeatedly acquiring multispectral data about agricultural crops and other features on the surface of the Earth. This spacecraft (*Landsat*) was so programmed that, from its flight altitude of approximately 570 mi (912 km), all of the Earth's surface between about 80°N and 80°S latitudes would be systematically covered, once every 18 days (absence of cloud cover permitting) by its multispectral scanner. Still further progress toward the automation of agricultural remote sensing data analysis was made

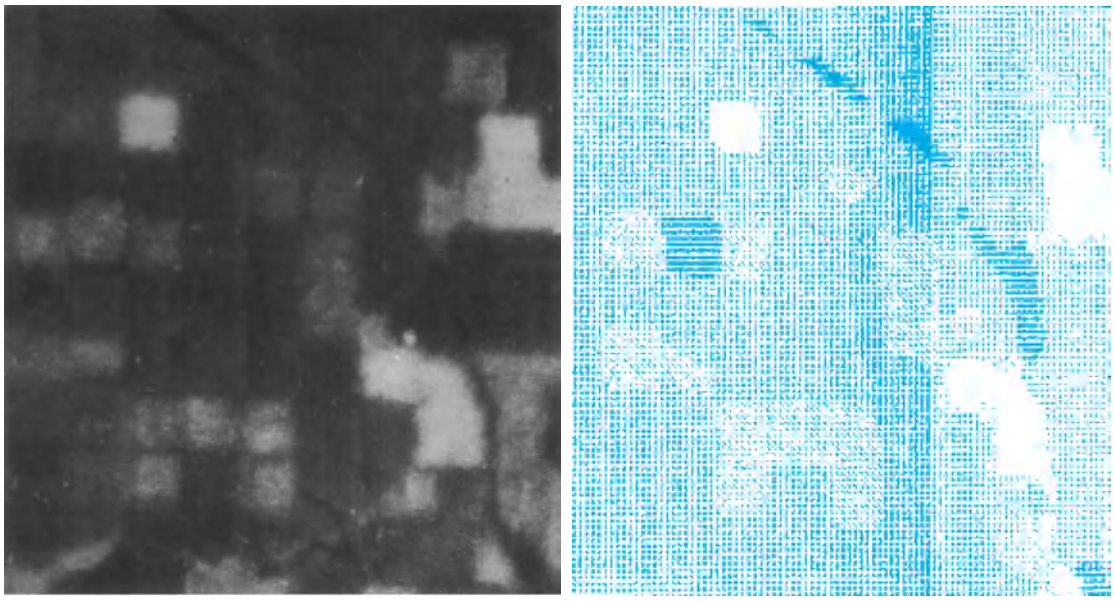

(a) (b)

Fig. 5. Field-by-field images of an agricultural area, acquired as in Fig. 4. (*a*) Image multispectral lines obtained with a *Landsat* multispectral scanner from an altitude of 570 mi (912 km) [enlarged 20 times]. (*b*) Computer printout of the same area made from the line scanner's digital data as telemetered to Earth.

possible through the use of multidate records, typically acquired at 9- or 18-day intervals (cloud cover permitting) throughout the entire growing season from Sun-synchronous crewless satellites that are in the *Landsat* series.

The accuracy and timeliness of remote sensing–derived information about agricultural crops could be improved through the use of pertinent weather data as acquired by meteorological satellites with all-weather and day-and-night capabilities. Such capability steadily improves as a more extensive historical database develops that correlates crop production to the multidate weather data as acquired at several times throughout the growing season.

AgRISTARS. Undoubtedly the most substantial research and development activities relative to agricultural applications of remote sensing was the multiagency program of the United States government known as AgRISTARS (Agricultural Resource Inventories Through Aerospace Remote Sensing). This was a follow-on to the completed program, under similar sponsorship, known as LACIE (Large Area Crop Inventory Experiment). The overall goal of this multiagency effort was to determine the usefulness, costs, and extent to which aerospace remote sensing data can be integrated into existing or future USDA information systems to improve the objectivity, reliability, timeliness, and adequacy of information required to carry out USDA missions.

The phase known as early-warning/crop condition assessment provides a capability for the USDA to respond in a timely manner to factors which affect the quality and production of agricultural crops. Among the potentially adverse factors for which a remote sensing–based early-warning capability is needed are diseases, insects, untimely freezes, winds, floods, and droughts. Another phase of AgRISTARS known as foreign commodity production forecasting was to develop and evaluate remote sensing–based technology for making improved forecasts of agricultural production in foreign areas and to determine the suitability of that technology for integration into USDA crop information systems.

The *Landsat*-based multispectral scanner system, and certain derivatives of it, served as the primary data source for evaluation in relation to the AgRISTARS objectives. Listed below are some of the primary attributes of the *Landsat* vehicle and its multispectral scanner system that have value in relation to achieving those objectives. (At present, no other vehicle-sensor systems provide this important combination of characteristics.)

1. With multispectral capability, these scanner systems sense for the optimum wavelength bands needed to take inventory and monitor most types of crops and related agricultural resources, and provide high spectral fidelity within each such band.
2. With multitemporal capability, the systems provide multiple "looks" for monitoring the seasonal development of agricultural crops.
3. The constant repetitive observation point facilitates change detection by the matching of multitemporal images.
4. Sun synchronicity (nearly constant Sun angle) ensures nearly uniform lighting and uniform image tone or color characteristics for use in identification of crops and related features from their multiband tone signatures.
5. A narrow angular field characterizes the sensors. Because of the 570-mi (912-km) altitude and only 115-mi (184-km) swath width of a typical *Landsat* system, tone or color fall-off at the edges of a swath is minimized, thereby increasing still further the uniformity of image tone or color characteristics.
6. Systematic coverage provides a nearly uniform look at all agricultural areas of the entire Earth.
7. Computer-compatible products are provided directly, facilitating automatic data processing.
8. Potential minimum delay in data availability to users permits real-time analysis and thus facilitates the making of timely agricultural inventories, whether on a local, regional, national, or global basis.
9. Spatial resolution is optimum for the first-stage look and, in relation to concerns about military security, is politically palatable, both domestically and internationally.
10. The data can be routinely placed in the public domain for benefit of all peoples.

In addition to these attributes of *Landsat*, there is a capability for the satellite to receive crop-related data as broadcast from ground-based data platforms (such as periodic data on air, soil temperature, soil moisture, wind direction, and wind velocity), process the data with on-board computers, and relay the resulting information back to ground-based receiving stations for integration with data acquired directly from *Landsat*.

Livestock management. Most of the agricultural crops for which management can be improved and production increased through proper inventory and monitoring are of value, primarily, in providing the food and fiber products that are used directly by humans. Other plants, however, are valuable because they provide feed for livestock, leading to the production of meat, hides, and other animal products. Among plants in the latter category are those occupying the pasturelands and rangelands that are grazed by cattle, sheep, and a few other kinds of domesticated animals. Also included are various feed crops such as certain varieties of corn, oats, barley, rye, sorghum, and alfalfa.

On rangelands, there are likely to be wide variations from one year to another in temperature, precipitation, and other climatic factors that control the growth rate of plants and require monitoring by remote sensing from aircraft and spacecraft.

These climatic variations lead to variations not only in the time of range readiness for grazing but also in both the amount and quality of forage produced per unit area—the two primary factors that govern the animal-carrying capacity. Consequently, in any given year, the range manager, equipped with timely, accurate multidate remote sensing–derived information, is better able to determine how many grazing animals, and what kind, should be permitted to graze those lands, and at what time during the grazing season.

Despite the limited spatial resolution of aerospace photography, attempts have been made to take remote sensing–based inventories on rangelands as to the numbers of livestock present, area by area, by livestock type and also by vigor class. For such tasks to be performed, the remote-sensing data must exhibit a sufficiently high degree of spatial resolution. The quality of *Landsat* imagery and other space-acquired photography has not been sufficiently high to permit the making of such determinations. Robert N. Colwell

Bibliography. M. A. Altieri, *Agroecology: The Scientific Basis of Alternative Agriculture,* 1987; M. A. Altieri and S. B. Hecht, *Agroecology and Small Farm Development,* 1991; E. C. Barrett and L. F. Curtis, *Introduction to Environmental Remote Sensing of Soils and Vegetation,* 1990; P. J. Curran, K. Kandratyev, and V. Kozogerov, *Remote Sensing of Soils and Vegetation,* 1990; C. Reinjtes, B. Haverkort, and A. Winters-Bayer, *Farming for the Future,* 1992.

Agroecosystem

A model for the functionings of an agricultural system, with all inputs and outputs. An ecosystem may be as small as a set of microbial interactions that take place on the surface of roots, or as large as the globe. An agroecosystem may be at the level of the individual plant–soil–microorganism system, at the level of crops or herds of domesticated animals, at the level of farms or agricultural landscapes, or at the level of entire agricultural economies.

Agroecosystems differ from natural ecosystems in several fundamental ways. First, the energy that drives all autotrophic ecosystems, including agroecosystems, is either directly or indirectly derived from solar energy. However, the energy input to agroecosystems includes not only natural energy (sunlight) but also processed energy (fossil fuels) as well as human and animal labor. Second, biodiversity in agroecosystems is generally reduced by human management in order to channel as much energy and nutrient flow as possible into a few domesticated species. Finally, evolution is largely, but not entirely, through artificial selection where commercially desirable phenotypic traits are increased through breeding programs and genetic engineering.

Energy flux. Agroecosystems are usually examined from a range of perspectives including energy flux, exchange of materials, nutrient budgets, and population and community dynamics. Each is a valid approach, but some approaches may be more appropriate in a particular case than others. For example, questions regarding production efficiency would likely require analyses of energy flux and nutrient budgets, while questions regarding crop–pest dynamics would likely necessitate a population or community approach.

Solar energy influences agroecosystem productivity directly by providing the energy for photosynthesis and indirectly through heat energy that influences respiration, rates of water loss, and the heat balance of plants and animals. At the level of individual plants, plant breeders have tried unsuccessfully to improve the effeciency of photosynthesis and reduce respiration losses. These basic metabolic pathways seem resistant to artificial selection. Breeders have successfully altered plant architecture, such as by changing leaf angles in wheat, so that the plant intercepts more sunlight. Similarly, animal breeders have selected livestock strains that are better able to withstand temperature extremes. *See* BIOLOGICAL PRODUCTIVITY; PHOTOSYNTHESIS.

The percentage of solar energy that is used by fields of crops also depends on the types of crops and spacing patterns. Densely planted crops such as wheat intercept more sunlight than more sparsely planted crops such as many vegetables. Comparisons of yields of different crops in similar fields show that net primary production at crop levels is a function of species-specific physiology and architecture and planting practices. For example, between 1930 and 1985 the average yield for maize in the United States increased from approximately 38.2 to 118 bushels/acre (3325 to 10,271 liters/hectare), and soya bean yields increased from 21.7 to 34.1 bu/acre (1915 to 2959 liters/ha). These yield differences reflected inherent species differences, increased planting densities, improved and more intense production inputs, and selection for new varieties. *See* ECOLOGICAL ENERGETICS.

Nutrient resources. Nutrient uptake from soil by crop plants or weeds is primarily mediated by microbial processes. Some soil bacteria fix atmospheric nitrogen into forms that plants can assimilate. Other organisms influence soil structure and the exchange of nutrients, and still other microorganisms may excrete ammonia and other metabolic by-products that are useful plant nutrients. There are many complex ways that microorganisms influence nutrient cycling and uptake by plants. Some microorganisms are plant pathogens that reduce nutrient uptake in diseased plants. Larger organisms may influence nutrient uptake indirectly by modifying soil structure or directly by damaging plants. When farms contain both domestic livestock and crops, animal manure may provide additional crop nutrients. *See* NITROGEN CYCLE; SOIL MICROBIOLOGY.

In most commerical farm operations, crop nutrients are greatly supplemented by fertilizer application to enhance yields. At the crop level, the long-term addition of fertilizer to soils may change soil pH, increase mineral salt concentration, and alter the

normal microbe–soil–plant interactions. *See* FERTILIZER.

Population and community interactions. Although agroecosystems may be greatly simplified compared to natural ecosystems, they can still foster a rich array of population and community processes such as herbivory, predation, parasitization, competition, and mutualism. Crop plants may compete among themselves or with weeds for sunlight, soil nutrients, or water. Cattle overstocked in a pasture may compete for forage and thereby change competitive interactions among pasture plants, resulting in selection for unpalatable or even toxic plants. Indeed, one important goal of farming is to find the optimal densities for crops and livestock. *See* HERBIVORY; POPULATION ECOLOGY.

However, other factors may influence a farmer's decision to plant at densities above or below this optimum. Risk of some mortality due to density-independent factors (such as from local variation in soil moisture) may lead to planting densities above the optimum. If the farmer is concerned about density-dependent mortality (for example, from a fungal pathogen whose spores are spread from plant to plant), crops may be planted more sparsely. Other factors of production, such as the cost of fertilizer, may also influence planting density. Optimal planting densities for a particular crop are determined with respect to a given set of production inputs, such as fertilizer levels, pesticides, irrigation, and tillage. Thus, at any given locale and planting year, a particular crop may have a range of optimal planting densities that depends on the amount and type of production factors used, and on the environmental risk factors that may concern the farmer.

The presence of weeds among crops can reduce crop yields through interspecfic competition. However, under many conditions the presence of weeds may not affect crop yields, and in certain circumstances weeds may actually enhance long-term production. Whether or not weeds will affect crop yields depends on the particular weed-crop species. Many weeds have little effect on yields and can be largely ignored, while other weeds, such as johnsongrass, can severely affect crop yields. At the appropriate densities, some weeds can improve crop production, particularly of orchard and vine crops, by providing ground cover, providing a habitat for beneficial insects, and improving soil nutrients. Weeds are also important for restoring soil nutrients and for protection against erosion during fallow periods. *See* WEEDS.

Biotechnology. Human influence on agroecosystems is increasing as advances are made in research geared toward artifically manipulating agricultural populations. New developments in pesticides and genetic engineering have had particularly strong impacts.

Pesticides. Since World War II, widespread use of synthetic chemical pesticides has bolstered farm production worldwide, primarily by reducing or eliminating herbivorous insect pests. Traditional broad-spectrum pesticides such as DDT, however, can have far-ranging impacts on agroecosystems. For instance, secondary pest outbreaks associated with the use of many traditional pesticides are not uncommon due to the elimination of natural enemies or resistance of pests to chemical control. The Colorado potato beetle, a common pest of many crops, quickly developed resistance to scores of pesticides; it examplifies the way in which long-term pesticide use may alter agroecosystem relationships. In addition, broad-spectrum pesticides may alter the soil community in an agroecosystem or kill pollinators and other beneficial organisms.

In 1996 the U.S. Congress passed the Food Quality Protection Act, which mandates a more restrictive means of measuring pesticide residues on crop plants. This legislation, borne primarily out of environmental and public health concerns, has resulted in the severe reduction and in some cases the banning of many broad-spectrum, traditional pesticides. As a result, growers and pesticide developers in temperate regions have begun to focus on alternative means of control. Pesticide developers have begun producing selective pesticides, which are designed to target only pest species and to spare natural enemies, leaving the rest of the agroecosystem community intact. Many growers are now implementing integrated pest management programs that incorporate the new breed of biorational chemicals with cultural and other types of controls. *See* FOREST PEST CONTROL; INSECT CONTROL, BIOLOGICAL; PESTICIDES.

Genetic engineering. The last few decades have seen tremendous advances in molecular approaches to engineering desirable phenotypic traits in crop plants. Traits such as resistance to insect attack, drought, frost, pathogens, and herbicides are among many that have been genetically spliced into numerous crop plants and made commercially available in the global market. Although artificially modifying crop plants is nothing new, the techniques used in genetic engineering allow developers to generate new varieties an order-of-magnitude faster than traditional plant breeding. In addition, genetic engineering differs from traditional breeding in that the transfer of traits is no longer limited to same-species organisms. Many of these genetically modified organisms are being designed with an eye toward increased marketability (for example, better-tasting crops and consistent ripening) as well as better environmental protection (for example, reduced reliance on pesticide sprays). At the same time, scientists are still assessing the effects that the widespread deployment of these traits may have on agroecosystems and natural ecosystems. There is some concern, for instance, that engineered traits may escape, via genes in pollen transferred by pollinators, and become established in weedy populations of plants in natural ecosystems, in some cases creating conservation management problems and new breeds of superweeds. As with pesticides, there is evidence

Maize production practices in the United States and Guatemala

Production factors	Guatemala	United States
Cost	Personal labor accounts for nearly all factors of production.	Fertilizers and pesticides amount to more than one-third of total production costs.
Planting technique	Maize is planted with other crops (intercropping) such as beans and squash. Total combined yield is greater than crops grown as monocultures in separate fields.	Maize is planted as a monoculture; chemical inputs set planting densities; sometimes maize is rotated with other crops based on market demand.
Distribution	Maize is used for human consumption; small livestock is fed on surplus grain and crop residue. Livestock is allowed to forage in fields after harvest.	Maize is used mostly for animal feeds or as additives (oils and meal) to prepared foods for human consumption.
Yield	Relatively low.	High.
Environmental factors	Land pressures force steep-slope cultivation, causing significant erosion. The chemical control of weeds can lead to pollution.	High environmental costs include soil erosion and chemical contamination of soil and water.

that insects are already becoming resistant to some more widespread traits used in transgenic plants, such as the antiherbivore toxin produced by the bacterium *Bacillus thuringiensis*. *See* BIOTECHNOLOGY; GENETIC ENGINEERING.

Comparative agroecosystems. A contrast between extreme forms of farm production illustrates the different principles of agroecosystem structure and function (see **table**). Commercial maize production in the United States corn belt can be contrasted with traditional maize cultivation in the highlands of Guatemala. In both regions, maize is a major part of the farm economy.

In highly commercial farms, the agroecosystem is more open in that soil loss from erosion and nutrient and chemical leaching are high; and in that yields are sold off the farm. Because fertilizer, pesticides, and other inputs are readily available, the weaknesses of the system are more than compensated by the large amounts of inputs. In the more traditional system, the agroecosystem is more closed with a greater cycling of nutrients taking place within the farm. Chemical inputs are low because the complexity of the fields offers some protection against pests and because the combination of maize and legumes may provide a mutualistic benefit. Bacteria associated with the legume roots fix atmospheric nitrogen, possibly providing the maize with a source of nitrogen. The vine legumes benefit from using the maize as a trellis.

The specter of incorporating biotechnological advances into commercial farms has increased hopes of augmenting on farm productivity while also reducing the use of environmentally damaging chemical inputs. *See* AGRICULTURAL SOIL AND CROP PRACTICES; AGRICULTURE; ECOSYSTEM; SOIL CONSERVATION.

C. Ronald Carroll; Carol A. Hoffman; John E. Banks

Bibliography. C. R. Carroll, J. H. Vandermeer, and P. M. Rosset (eds.), *Agroecology*, 1990; S. R. Gliessman, *Agroecology: Ecological Processes in Sustainable Agriculture*, 1998; S. R. Gliessman, *Agroecology: Researching the Ecological Basis for Sustainable Agriculture*, 1990; S. R. Gliessman, *Agroecosystem Sustainability: Developing Practical Strategies*, CRC Press, 2001; O. Koul and G. S. Dhaliwal (eds.), *Predators and Parasitoids* (Advances in Biopesticide Research Series), Taylor & Francis, 2003; S. Krimsky and R. Wrubel, *Agricultural Biotechnology and the Environment*, 1996; National Research Council, *Alternative Agriculture*, 1989; National Research Council, *Environmental Effects of Transgenic Plants: The Scope and Adequacy of Regulation*, National Academy Press, 2002.

Agronomy

The science and study of crops and soils. Agronomy is the umbrella term for a number of technical research and teaching activities (see **illus.**): crop physiology and management, soil science, plant breeding, and weed management frequently are included in agronomy; soil science may be treated separately; and vegetable and fruit crops generally are not included. Thus, agronomy refers to extensive field cultivation of plant species for human food, livestock and poultry feed, fibers, oils, and certain industrial products. *See* AGRICULTURE.

Agronomic studies include some basic research, but the specialists in this field concentrate on applying information from the more basic disciplines (see illus.), among them botany, chemistry, genetics, mathematics, microbiology, and physiology. Agronomists also interact closely with specialists in other applied areas such as ecology, entomology, plant pathology, and weed science. The findings of these collaborative efforts are tested and recommended to farmers through agricultural extension agents or commercial channels to bring this knowledge into practice. Agronomy is an integrative science that brings together principles and practices from other fields into combinations that are useful to farmers. This critical area is now focused on the efficiency of resource use, profitability of management practices, and minimization of the impact of farming on the immediate and the off-farm environment.

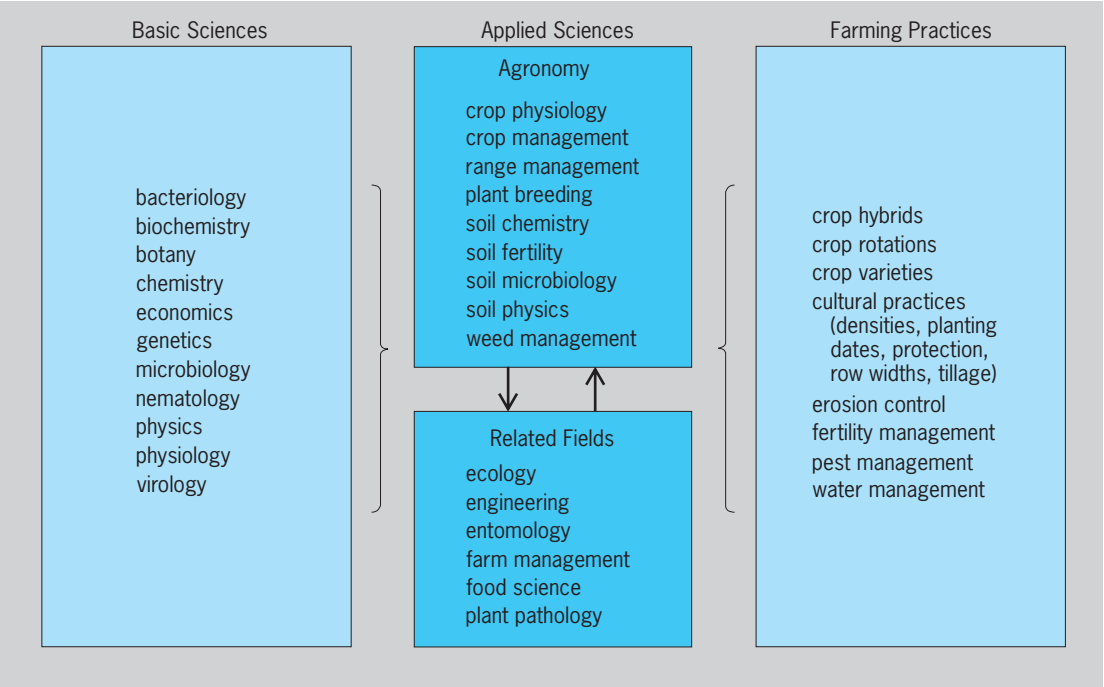

Relationship of agronomy to basic sciences, related fields, and practical farming methods.

History of cropping. Agronomy has its roots in the prerecorded history of human settlements. The first agronomists were no doubt women who nurtured small plants in garbage areas near campsites and then began to consciously bring back plants and seeds and tend them to useful maturity in these primitive gardens. The most vigorous and desirable plants were used, and seeds were selected for the next planting. Such trial-and-error planting led to the development of more permanent settlements based on agriculture to complement gathering, fishing, and hunting. The more organized farming systems with crops in rows and intensive management are barely more than one to two centuries old. This so-called agricultural revolution and the subsequent application of machines during the industrial revolution brought increased efficiency to farming and reduced much physical drudgery. Agronomy now brings practical science and the basic results of research to bear on increased productivity and efficiency. *See* AGRICULTURAL MACHINERY.

Science has made major contributions to agriculture through agronomy in the latter half of the twentieth century. Hybrid corn and sorghum, new soybean and wheat varieties, manufactured or enhanced chemical fertilizers, and a range of pesticides contributed greatly to labor productivity and crop yield during this time. In some developing countries in the 1960s and 1970s, a so-called green revolution based on new plant varieties, fertilizer and pesticide usage, and improved water control brought new hope to areas facing severe population and food crises. The green revolution was effective primarily on fertile, mechanized farms where infrastructure was well developed, inputs and markets were available, and substantial capital was invested in agriculture. *See* BREEDING (PLANT); FERTILIZER; PESTICIDE.

Technical specialists. Agronomists promote the use of improved technology through research, teaching, and extension activities. In the United States, much of this is in the state and federal systems. Commercial employment includes soil testing and recommendations, scouting for insect, weed, and disease problems, and consulting on fertilizer, pesticide, and irrigation water use. Field representatives and salespeople for chemical and fertilizer companies, as well as technical seed production specialists, have training in agronomy. Many farmers and consultants for large commercial farms have formal college or technical school courses in the specific details of science as related to agriculture.

Future directions. The positive and negative impacts of current farming practices are becoming more apparent. High-productivity-per-unit labor has been achieved through large investments of fossil-fuel-derived inputs, such as fertilizers, pesticides, and diesel fuel for drying crops and pumping water for irrigation. In addition, the amounts and off-farm effects of nitrates and pesticide residues in surface waters and underground aquifers can be measured with greater precision. Agriculture in some areas has been identified as a major nonpoint source of pollution of the environment. Agronomists have been called upon to provide solutions in these areas. *See* IRRIGATION (AGRICULTURE).

Principles of plant ecology and the biological processes and interactions in natural communities are being studied to gain insight into the design of future cropping systems. Crop rotations, reduced tillage, and contour strip cropping can help prevent the loss

of soil and nutrients from sloping lands or reduce the need for added nitrogen before the planting of each cereal crop. Rotations break the reproductive cycles of weeds and insects and reduce pesticide requirements. Plant breeders can develop new hybrids and varieties that are more tolerant to diseases and insects. Commercial applications of biotechnology will result in products that are biologically derived, are used in ultralow volume, or have a more specific effect on the target species. The result is a less negative impact on desirable species and on the total environment. Agronomists in the public and private sectors integrate these new technologies with current farm practices to create a more resource-efficient, profitable, and sustainable agriculture. *See* AGRICULTURAL SOIL AND CROP PRACTICES; AGROECOSYSTEM; CONSERVATION OF RESOURCES; SOIL CONSERVATION.

Charles A. Francis

Bibliography. J. A. Barden, R. G. Halfacre, and D. J. Parrish, *Plant Science*, 1987; W. R. Fehr, *Principles of Cultivar Development*, 1987; M. S. Kipps, *Production of Field Crops: A Textbook of Agronomy*, 6th ed., 1970; J. H. Martin, W. H. Leonard, and D. L. Stamp, *Principles of Field Crop Production*, 3d ed., 1989; N. C. Stoskopf, *Cereal Grain Crops*, 1985; S. L. Tisdale and W. L. Nelson, *Soil Fertility and Fertilizers*, 6th ed., 1998.

Aharonov-Bohm effect

The predicted effect of an electromagnetic vector or scalar potential in electron interference phenomena, in the absence of electric or magnetic fields on the electrons.

The fundamental equations of motion for a charged object are usually expressed in terms of the magnetic field $\vec{B}$ and the electric field $\vec{E}$. The force $\vec{F}$ on a charged particle can be conveniently written as in Eqs. (1), where q is the particle's charge, $\vec{v}$ is

$$\vec{F} = q\vec{E} \qquad \vec{F} = q\vec{v} \times \vec{B} \tag{1}$$

its velocity, and the symbol $\times$ represents the vector product. Associated with $\vec{E}$ is a scalar potential V defined at any point as the work W necessary to move a charge from minus infinity to that point, $V = W/q$. Generally, only the difference in potentials between two points matters in classical physics, and this potential difference can be used in computing the electric field according to Eq. (2), where d is the distance

$$\vec{E} = \frac{V_1 - V_2}{\vec{d}} \tag{2}$$

between the two points of interest. Similarly, associated with $\vec{B}$ is a vector potential $\vec{A}$, a convenient mathematical aid for calculating the magnetic field; the magnetic field is the curl of the vector potential, Eq. (3).

$$\vec{B} = \nabla \times \vec{A} \tag{3}$$

See CALCULUS OF VECTORS; ELECTRIC FIELD; POTENTIALS.

In quantum mechanics, however, the basic equations that describe the motion of all objects contain $\vec{A}$ and V directly, and they cannot be simply eliminated. Nonetheless, it was initially believed that these potentials had no independent significance. In 1959, Y. Aharonov and D. Bohm discovered that both the scalar and vector potentials should play a major role in quantum mechanics. They proposed two electron interference experiments in which some of the electron properties would be sensitive to changes of $\vec{A}$ or V, even when there were no electric or magnetic fields present on the charged particles. The absence of $\vec{E}$ and $\vec{B}$ means that classically there are no forces acting on the particles, but quantum-mechanically it is still possible to change the properties of the electron. These counterintuitive predictions are known as the Aharonov-Bohm effect.

Effect for vector potentials. According to quantum mechanics, any object with mass can behave as both a particle and a wave. Classical interference phenomena, such as the interference pattern produced when a point source of light is imaged by two slits, are still observed when this light source is replaced by an electron beam, demonstrating the electron's wave-like nature. The phase δ of the quantum-mechanical wave function for the electron can be written as in Eq. (4), where δ_0 is an arbitary phase factor, e is

$$\delta = \delta_0 + \frac{2\pi e}{h}\int \vec{A}\cdot\vec{dl} + \frac{2\pi e}{h}\int V(t)dt \tag{4}$$

the charge of the electron, h is Planck's constant, the integral of the dot product of the vector potential $\vec{A}$ with $\vec{dl}$ is taken over the electron path, and

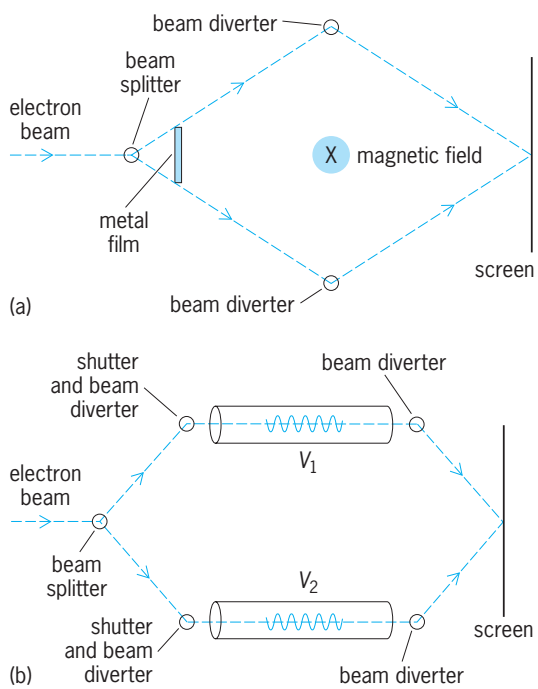

Fig. 1. Aharonov and Bohm's proposed experiments to demonstrate the effects of electromagnetic potentials in electron interference phenomena. (*a*) Experiment to verify the effect of a vector potential. (*b*) Experiment to demonstrate the effect of a scalar potential.

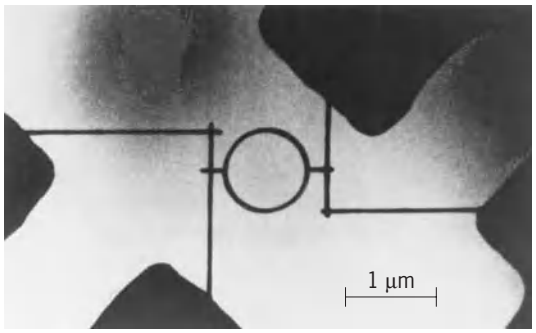

Fig. 2. Very small ring fabricated from a gold film 38 nm thick. LInes forming the ring are 40 nm wide, and the average diameter is 0.82 μm. The electrical resistance is measured by applying a constant current through two of the leads connecting the sample and measuring the voltage difference between the two other leads. (*From R. A. Webb et al., Observation of h/e Aharanov-Bohm oscillations in normal-metal rings, Phys. Rev. Lett., 54:2696–2699, 1985*)

the scalar potential $v(t)$ can vary with time t. *See* ELECTRON DIFFRACTION; INTERFERENCE OF WAVES; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

In Aharonov and Bohm's first proposed experiment (**Fig. 1***a*), a beam of electrons traveling in a vacuum is split such that half of the beam travels clockwise around a region containing a magnetic field and the other half counterclockwise. The integral of the vector potential over the electron path given in Eq. (4) is identically equal to the magnetic flux Φ (the magnetic field times the area over which the field is nonzero) enclosed by the two beams. When the beams are recombined, the intensity of the resulting beam will oscillate periodically as the enclosed magnetic field is changed. This oscillation, with a period in Φ of $\Phi_0 = h/e$, represents patterns of constructive and destructive interference occurring between the two beams. If the magnetic field (produced, for example, by an infinite solenoid) is confined to the interior of the region within the electron paths, the recombined beam will still oscillate periodically as the field is varied because $\vec{A}$ is not zero in the region where the electrons travel; only $\vec{B}$ is.

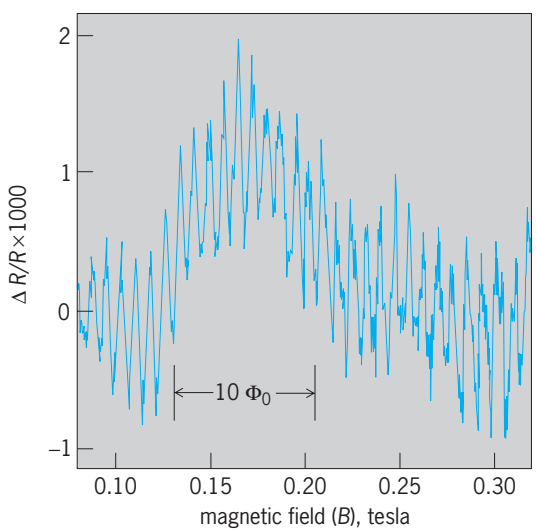

Fig. 3. Resistance oscillations observed in the ring of Fig. 1 at temperature of 0.01 K over a very small magnetic field range. Vertical scale measures deviation ΔR of resistance from its mean value R. The period of the h/e Aharonov-Bohm oscillations in magnetic field is 0.0076 tesla (76 gauss). The arrows show the field scale for 10 flux quanta (10 Φ_0) in the average area of the ring.

Within a year of this prediction, evidence for the Aharonov-Bohm effect and the physical reality of the vector potential was obtained in electron-beam experiments, but only in 1986 was a conclusive demonstration obtained.

Effect for scalar potentials. Aharonov and Bohm's second interference experiment (Fig. 1*b*) is designed to demonstrate that a scalar potential can also change the phase of the electron wave function. An electron beam is again split in half; each half passes through a long, cylindrical, high-conductivity metal tube; and the beam is recombined. Two electrical shutters chop the beam into segments that are much shorter than the tube length but much longer than the de Broglie wavelength. If a different scalar potential (V_1 and V_2 in Fig. 1*b*) is applied to each metal tube after the chopped beam enters the tube and is turned off before the beam leaves the tube, the intensity of the recombined beam will be altered and will oscillate periodically as the potential difference between the tubes ($V_1 - V_2$) is changed. Since a potential applied to a highly conducting metal tube produces no electric field inside the tube, no classical forces act on these electrons during the experiment. While this prediction has not been conclusively demonstrated, evidence for its validity has been obtained in electron-beam experiments and in electrical measurements on normal metals. *See* ELECTROSTATICS.

Effect for normal metals. Surprisingly, the Aharonov-Bohm effect plays an important role in understanding the properties of electrical circuits whose wires or transistors are smaller than a few micrometers. The electrical resistance in a wire loop oscillates periodically as the magnetic flux threading the loop is increased, with a period of h/e, the normal-metal flux quantum. In single wires, the electrical resistance fluctuates randomly as a function of magnetic flux. Both these observations reflect an Aharonov-Bohm effect. They have opened up a new field of condensed-matter physics because they are a signature that the electrical properties are dominated by quantum-mechanical behavior of the electrons, and that the rules of classical physics are no longer operative.

These experiments were made possible by advances in the technology for fabricating small samples. The metallic ring shown in **Fig. 2** is about 100 atoms wide and 100 atoms thick, and contains only about 10^8 atoms. As the magnetic field is varied, the electrical resistance of the ring oscillates with a characteristic period of $B = 0.0076$ tesla (**Fig. 3**). From measurements of the average area S of the ring, the period is equal to the normal-metal flux quantum $\Phi_0 = h/e = BS$. These data demonstrate that the Aharonov-Bohm effect exists in resistive samples.

Richard A. Webb

Bibliography. M. Peshkin and A. Tonomura, *The Aharonov-Bohm Effect*, 1989.

Aileron

The hinged rear portion of an aircraft wing, moved differentially on each side of the aircraft to obtain lateral or roll control moments. The angular settings of the ailerons are controlled by the human or automatic pilot through the flight control system. Typical flap- and spoiler-type ailerons are shown in the **illustration**. *See* FLIGHT CONTROLS.

Operating principles. The operating principles of ailerons are the same as for all trailing-edge hinged control devices. Deflection of an aileron changes the effective camber, or airfoil curvature relative to the wing chord, of the entire wing forward of the aileron. With the trailing edge deflected upward, reduced low flow velocities are produced on the upper wing surface, and increased local flow velocities are produced on the lower wing surface. By Bernoulli's law, this results in a reduction of lift over the portion of the wing forward of the aileron, and on the aileron itself. Conversely, trailing-edge down deflection of a flap-type aileron increases the lift in the same areas. Ailerons are located as close as possible to the wing tips, to maximize rolling moment by increasing the moment arm of the force due to the change in wing lift. In the case of flap-type ailerons, when the trailing edge is raised on one wing, say the left, the trailing edge of the aileron on the opposite or right wing is lowered by about the same amount. The decrease in lift on the left wing is accompanied by a lift increase on the right wing. While the net wing lift remains about the same, a rolling moment or torque about the aircraft's fore-and-aft axis develops in a left, or counterclockwise, direction as seen by the pilot. *See* BERNOULLI'S THEOREM.

Aileron effectiveness is reduced at supersonic speeds, since trailing-edge devices cannot affect pressures on the wing surface forward of themselves when the flow velocity exceeds the speed of pressure propagation (speed of sound). Lift changes thus occur only on the ailerons themselves at supersonic speeds, except for end effects. *See* SUPERSONIC FLIGHT.

Flight maneuvers. Ailerons are used to perform rolling or banking maneuvers. Banked turns, in which one wing is lower than the other, are used to change aircraft heading. Ailerons are deflected to establish a rolling moment and hence a rolling acceleration. The rolling acceleration becomes a steady rolling velocity usually less than 2 s after the ailerons have been deflected, as a result of the heavy damping in roll (rolling moment opposing rolling velocity) of the wings. Neutralizing the ailerons stops the rolling velocity, thus allowing the rolling motion to be stopped at the desired bank angle. Aileron control effectiveness is evaluated relative to given levels of rolling velocity, or alternatively to banking a given amount, such as 30 degrees, in a specified time. Other functions of the ailerons are the maintenance of lateral balance under conditions of aerodynamic, inertial, or thrust asymmetries, crosswind takeoffs and landings, deliberate sideslips, and lateral gusts. Ailerons have become the primary spin-recovery control for aircraft with low rolling inertia relative to pitching inertia; this characteristic is typical of modern fighter aircraft with thin wings of short span.

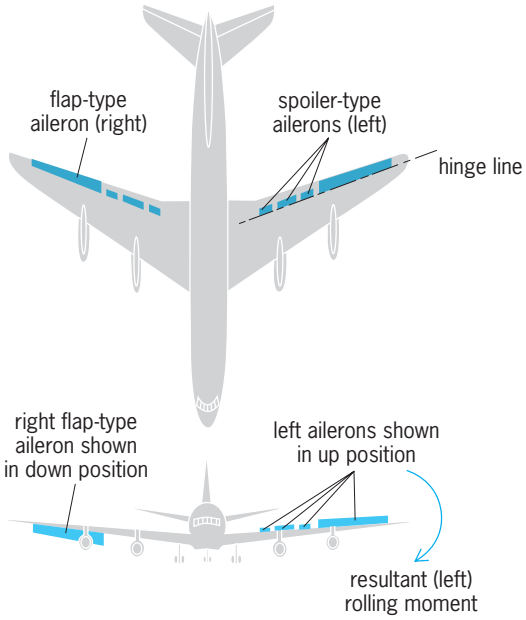

Flap and spoiler-type ailerons on jet transport airplane.

Spoiler ailerons. Flap-type ailerons are replaced or supplemented by spoiler-type ailerons for a variety of reasons. Spoiler ailerons are usually installed forward of the landing flaps on commercial jet transports, in order to supplement aileron effectiveness during landing approaches, when the landing flaps are extended. Greatly reduced takeoff and landing speeds can be obtained by devoting the trailing edge of the entire wing to high-lift flaps. This is made possible by substituting spoilers for flap-type ailerons.

Spoilers have the additional advantages of reduced wing twisting moments due to chordwise aerodynamic loading and of improved effectiveness at transonic speeds. Spoilers designed for lateral control at transonic or supersonic speeds are often located on the inboard portion of the wings. This further reduces wing twisting due to chordwise aerodynamic loading.

A major disadvantage of spoiler ailerons is that a practical means of self-balancing the operating torques with other aerodynamic torques has not been found. All current spoilers are hinged at the leading edge and must be forced open against large positive aerodynamic pressures acting on their outer surfaces. *See* AERODYNAMICS; WING.

Malcolm J. Abzug

Bibliography. J. D. Anderson, *Introduction to Flight*, 4th ed., 1999; E. H. Pallett, *Automatic Flight Control*, 4th ed., 1994.

Air

A predominantly mechanical mixture of a variety of individual gases enveloping the terrestrial globe to form the Earth's atmosphere. In this sense air is one of the three basic components, air, water, and land (atmosphere, hydrosphere, and lithosphere), that form the life zone at the face of the Earth. *See* AIR POLLUTION; AIR PRESSURE; AIR TEMPERATURE; ATMOSPHERE; ATMOSPHERIC CHEMISTRY; ATMOSPHERIC GENERAL CIRCULATION; METEOROLOGY.

Radio, radar, rockets, satellites, and interplanetary probes are stimulating investigation of the upper atmosphere and the transition zone to outer space. *See* IONOSPHERE.

Some aspects of air important in engineering are well known. For example, pneumatic equipment commonly transmits force and energy in a variety of mining machines and drills, in air brakes, automotive air suspension, and other devices. The study of forces and moments exerted on aircraft in motion through the air is a special field of aerodynamics. *See* AERODYNAMIC FORCE; AERODYNAMICS; COMPRESSOR; WIND STRESS. Charles V. Crittenden

Bibliography. M. Allaby, *Air*, 1992.

Air armament

That category of weapons which are typically delivered on target by fixed or rotary-wing aircraft, with the exception of nuclear weapons. Specifically included are guns and ammunition, rockets, free-fall bombs, cluster weapons that consist of a dispenser and submunitions, air-to-air and air-to-surface guided weapons, mines, and antiradiation missiles. Nuclear weapons such as the air-launched cruise missile and nuclear bombs, which are delivered by aircraft, are considered strategic weapons, a specific class of air armament. Related to, but not included as, armament are carrying racks and ejector mechanisms, fire-control computers, interface electronics, and head-up displays, which are critical to delivering the unguided weapons accurately. *See* ATOMIC BOMB; HYDROGEN BOMB.

Aircraft guns and cannon. Since World War II, all major nations have practiced optimization, or fitting the weapon to the job, in attempting to obtain maximum effectiveness against anticipated targets with minimum weight and volume systems. The most widely used aircraft guns outside the former Communist countries are the French DEFA and British Aden 30-mm guns, followed by the United States 20-mm M61 Vulcan.

Since World War II, the Russians have employed a variety of calibers. Their dominant weapons are all designed to fire the same 23-mm round, although .50-caliber guns are still seen on some helicopters and 30 mm on some fixed-wing aircraft.

The most modern aircraft guns are the 27-mm Mauser developed for multirole use on the German, Italian, and British multirole combat aircraft (MRCA) Tornado, and the large 30-mm GAU-8 and GAU-13 antiarmor guns which carry high-explosive incendiary (HEI) and armor-piercing incendiary (API) ammunition. The GAU-8 is carried inside the United States A-10 aircraft, for which it was developed. The GAU-13, developed for the A-7/F-4, can be mounted in the GPU-5A gunpod, outside most modern tactical fighters.

A 20-mm aircraft gun system for air-to-air combat has been developed that greatly expands the gun engagement envelope by employing telescoped ammunition. This gun was tailored for the F-22 and Joint Strike Fighter (JSF) aircraft. With this type of ammunition, the muzzle velocity of the round is increased to over 5000 ft/s (1500 m/s), compared with 3300 ft/s (1000 m/s) for conventional ammunition. This feature, combined with a lead-computing, optical gunsight, gives the pilot an all-aspect engagement capability.

Aircraft rockets. Free rockets, or unguided rockets, developed during World War II, are more accurate than bombs but less accurate than guns. Their warheads are generally larger than shells that can be fired from aircraft guns but smaller than most bombs. Most rocket usage is air-to-surface, where reasonable accuracy is attainable. Although rocket systems are often considered obsolete, virtually all major powers maintain one or more in their arsenal. Most current systems are fin-stabilized (as opposed to spin-stabilized), and vary from about 2.4 in. (60 mm) to over 8 in. (200 mm) in diameter. They are designed to be employed at ranges from 3000 to 30,000 ft (1 to 10 km). *See* ROCKET.

Bombs. Aircraft-delivered bombs were considered to be the ultimate weapon until the end of World War II. Conventional bombs employing TNT or TNT-based mixtures have been made in many types and sizes. Armor-piercing bombs were designed for use against concrete fortifications and submarine pens. Demolition bombs have thin steel cases and maximum explosive loads. General-purpose bombs are a compromise between armor-piercing and demolition. Bombs have been made to deliver virtually anything against any type of target. Some examples are incendiary bombs to ignite wooden buildings; napalm or jellied gasoline for use against anything that can be damaged by flame; underwater bombs for use against hydroelectric dams; cluster bombs, which separate and disperse to saturate an area; and leaflet bombs to deliver propaganda.

General-purpose, demolition, and armor-piercing bombs are made in a variety of sizes up to about 2000 lb (900 kg). The advent of nuclear weapons made larger bombs obsolete. Smaller, more accurate, conventional bombs are preferred to destroy targets while minimizing collateral damage. Examples are the large number of smart weapons employed during Operation Desert Storm in 1991 and subsequent air strikes against Iraq to enforce the United Nations sanctions.

The majority of bombs (and other stores) are carried on ejector racks, which forcibly eject the bomb through aerodynamic interference to ensure that it

leaves the aircraft. Pylons, racks, and ejectors have become a special field of aircraft armament. *See* AERODYNAMICS.

Cluster weapon. An entirely new class of aircraft armament known as cluster weapons has evolved since the 1950s. It is partly an outgrowth of cluster bombs from World War II and partly a result of the realization that many targets, notably personnel and light vehicles, could be destroyed more efficiently with several small bombs than with one large bomb. The primary advantage of cluster weapons over unitaries is their large footprint or area covered, which compensates for delivery errors and target uncertainty errors incurred by unguided weapons.

Both captive dispensers, which stay with the aircraft, and droppable stores, which sometimes fly preprogrammed maneuvers, are used. The means of dispersing the munitions are varied. The munitions may be shot out of their recesses, ejected by air or gas pressure, separated after ejection by various aerodynamic means, or even separated by centrifugal force from a spinning canister. The trend is toward dispensers that deploy a wide pattern of munitions from aircraft at treetop height.

The 950-lb (430-kg) combined effects munition (CBU-87/B, where CBU stands for cluster bomb unit) can be delivered at supersonic speeds and is used to attack armor, vehicles, and personnel over a wide area. Each free-fall dispenser carries 202 submunitions, weighing 3.4 lb (1.5 kg) apiece, which consist of a steel fragmentation case, a shaped charge that can defeat armored vehicles, and fire-starting materials. Fusing can be selected as a function of time for low-level dispensing or as a function of ground proximity when the weapon is released at higher altitudes. Adjustable dispenser-spin and submunition-dispersion characteristics allow the weapon to be tailored to the situation, such as vehicles parked in an area or strung out along a moving column.

The CBU-89/B or Gator mine system carries a combination of 72 antitank/antivehicle mines and 22 antipersonnel mines. This 710-lb (320-kg) free-fall cluster weapon is the first and only scatterable mine system in the U.S. Air Force inventory, and can be used against tanks, trucks, personnel, and armored vehicles. After release from an aircraft, the dispenser is opened by either a timed fuse or a ground-proximity sensor, depending on the aircraft altitude. This altitude determines how widely the mines are scattered. The mines detonate when either a preprogrammed self-destruct time-out is reached, the battery voltage in the mine falls below a certain level, or a target is detected. The antitank/antivehicle mines sense the proximity of the target with a magnetic sensor. The antipersonnel variety explodes if any one of four triplines is disturbed or someone tries to move it.

An even more sophisticated cluster weapon, the sensor-fused weapon, CBU-97/B, is designed to carry 10 submunition delivery vehicles (SDVs), each containing four warheads. After dispersion, the SDVs orient, stabilize, and drop by parachute to the optimal altitude as determined by the on-board altimeter. The 40 sensor-fused warheads are then dispersed, each searching for targets by using individual infrared sensors. When a target is sensed, the warhead fires a high-velocity, self-forging projectile at the target vehicle.

The cluster bomb units, CBU-87/B, CBU-89/B, and CBU-97/B, greatly improved in accuracy with the addition of a fin control kit containing an inertial guidance system, which achieved initial operational capability in 1999. With the added kit, the cluster weapons are known as wind-corrected munitions dispensers, CBU-103/B, CBU-104/B, and CBU-105/B. The need for such kits became apparent during the high-altitude bombing of Iraq during Desert Storm, where wind effects made the basic cluster weapons ineffective.

Fire control. Although not normally included as part of the field of aircraft armament, fire control is the term that covers the sighting, aiming, and computation which enables the pilot or aircrew to hit the target. Basically, it examines the conditions of the engagement and indicates when to release the armament in order to obtain hits.

The most primitive system, the fixed sight, is a mechanical or optical device that is aligned to coincide with the gun-bore line some fixed distance in front of the aircraft. To use such a sight requires estimating range, lead, deflection, and a number of lesser variables. *See* GUNSIGHTS.

The first degree of sophistication, the lead computing optical sight, uses gyroscopic sensors to measure own-vehicle motion and, based on the assumption that both aircraft are flying similar maneuvers, predicts the future position of the target and lags the reticle, or pipper, so that when the reticle is on the target the correct lead has been generated. This system is also known as a disturbed-reticle sight because the reticle is in constant motion in a maneuvering encounter.

The ultimate fire-control system, the director, measures all conditions, own-vehicle motion, velocity, air density, angle of attack, and other deterministic variables; tracks the target with radar or electrooptical sensors; measures the range to target; computes in real time the ballistics of the round; predicts where it will pass the target at target range; and displays a reticle on the head-up display in the pilot's field of view. If the pilot pulls the trigger when this reticle is on the target, hits will likely result.

Guided weapons. Guided weapons is a generic term which applies to any of the previously described ballistic systems when they are deliberately perturbed from their ballistic path after launch in order to increase the probability of hitting a target. There are three fundamental problems in guided weapons: determining where the target is or will be, determining where the weapon is, and correcting the weapon's location to coincide with the target's location at the time of closest encounter. The first two problems are called guidance; the third, control. Control is the most straightforward problem and is usually accomplished by aerodynamic-control-surface deflection, although it may also be achieved

by rocket-thrust-vector control, gas thrusters, explosive squib control, or any other device to alter the velocity vector of the weapon. *See* AUTOPILOT; FLIGHT CONTROLS; GUIDANCE SYSTEMS.

There are five fundamental concepts of guidance, namely inertial guidance, command guidance, active guidance, semiactive guidance, and passive guidance, and many different implementations of each.

Inertial guidance. This requires that the initial conditions at launch be precisely known, that is, distance and heading to target, and velocity and orientation of the launch vehicle. On-board sensors (gyroscopes and accelerometers) measure all changes in initial conditions and, using computers, direct the guidance elements (autopilot) to maneuver the weapon to the target. This guidance concept is not very accurate because it is open-loop and is generally used with more precise terminal guidance which is closed-loop.

Aided inertial guidance using an inertial navigation system and the constellation of Global Positioning System (GPS) satellites has become an affordable way to achieve accurate terminal guidance where the targets are not moving and the locations of the targets in the GPS coordinate system are known. GPS provides accurate position and velocity information to the inertial system, thereby bounding the inertial system's errors, while the inertial system provides short-term filtering of the guidance information between GPS updates. The GPS system errors [typically, 15 m (50 ft) spherical error probable] and the target location uncertainty limit the accuracy achievable. *See* CONTROL SYSTEMS; INERTIAL GUIDANCE SYSTEM; SATELLITE NAVIGATION SYSTEMS.

Command guidance. In command guidance, the weapon system operator or on-board sensors observe the relative location of weapon and target and direct trajectory corrections. In the simplest application, the operator may visually observe the weapon and target and remotely maneuver the weapon. A second degree of refinement is a television camera in the nose of the missile, through which the operator observes the target and guides the missile. In the most refined state, sensors on the launch aircraft track both the target and missile, and command the flight trajectory of the weapon without human intervention.

Active guidance. In this type of system, electromagnetic emissions, for example, radar, microwave, or laser, are transmitted from the weapon to the target, and the return energy reflections are measured to determine range and angle to the target. This return is continuously processed to provide control signals. Depending upon the sophistication of the weapon's seeker, the reflected energy may be processed in a manner that will produce imagery of the target to increase the likelihood of hitting a particular aim point. *See* LIDAR; RADAR.

Semiactive guidance. This resembles active guidance except that the illumination of the target is provided by a designator not located on the weapon. It may be on the launch aircraft, another aircraft, or the ground. When illumination is provided by a source other than the launch platform, the term bistatic guidance is used.

Passive guidance. This uses the natural emissions radiating from targets (such as infrared, visible, radio-frequency, or acoustic), which are usually referred to as the target's characteristic signature, to uniquely acquire a target and subsequently guide the weapon to the target.

Smart weapons. The ultimate objective of a weapon system is to be completely autonomous. An example of an autonomous system is one combining inertial and passive or active terminal guidance, and the appropriate algorithms to acquire the target after launch without operator intervention. In this case, the weapon is launched into an area where targets are known to exist, and, upon reaching the area, the weapon searches and finds its own target, homes in on it, and destroys it. The trend toward weapons that can autonomously acquire targets allows weapons to be built that have a substantial standoff capability which increases the survivability of the launch platform, improves accuracy, increases proficiency, and reduces the logistical burden.

There are two classes of smart weapons. The first class consists of those guided weapons that possess some form of terminal guidance and home in on the target. Weapon systems in this class include laser-guided bombs such as Paveway II and Paveway III, imaging systems such as the GBU-15, AGM-130, and Maverick where television or infrared imagery is used, and hotspot tracking systems such as the air-to-air AIM-9S Sidewinder missile and the high-speed antiradiation missile (HARM).

Laser-guided bombs. The Paveway III (GBU-24, GBU-27) low-level laser-guided bomb can alter its course from midpoint to impact by homing in on the reflections off the target from a laser illuminator. The aircraft using the weapon flies low enough to avoid falling victim to air-defense weapons. The weapon can be delivered in any of three modes—level, lofted, or diving—adapting automatically to the mode employed. Target illumination is either continuous or delayed until after the weapon is well along in its flight path, and provided by the releasing aircraft, another aircraft, or supporting ground forces.

As a result of the buried command-and-control bunkers encountered during the Desert Storm conflict, a 4700-lb (2130-kg) variant of the 2000-lb (900-kg) GBU-24 was produced in 28 days. Coincidentally, the GBU-28 was designed to defeat just such buried command-and-control sites, and successfully accomplished this, ending the Desert Storm conflict one day after being deployed in theater and used effectively only one time.

Imaging systems. The Maverick, AGM-65D, an air-to-surface missile, uses infrared imagery to develop the terminal guidance information, and must be locked on to the target prior to launch. Subsequent guidance is completely autonomous until target impact.

The GBU-15 air-to-surface guided glide weapon is similar in some respects to the Maverick and uses either television (visible) or imaging infrared guidance and a data link. Information supplied by the data link

is used to acquire the target and perform midcourse maneuvers. Once the target is acquired, the missile's seeker can be locked on and the aimpoint selected. Terminal guidance can then be performed on-board the weapon automatically. A longer-range version of the GBU-15 is the AGM-130, where a 600-lb (270-kg) rocket motor has been added to the GBU-15 airframe to increase standoff. Also, as a consequence of experience gained during Desert Storm, the AGM-130 is fitted with a GPS system to reduce the workload of the weapons officer during the midcourse phase of the weapon's flight.

Hot-spot tracking systems. In the case of the AIM-9S Sidewinder air-to-air missile, lock-on to the target must occur before launch. Terminal guidance is performed on-board by the missile's electronics and seeker following launch.

Brilliant weapons. The second class of smart weapons includes those that autonomously acquire the target after launch, and are usually termed lock-on-after-launch, fire-and-forget, or brilliant weapons. Among such weapons are the high-speed antiradiation missile, the advanced medium-range air-to-air missile (AMRAAM), and developmental weapons like the Autonomous Guided Conventional Weapon (AGCW), the Joint Direct Attack Munition (JDAM), the Joint Standoff Weapon (JSOW), and the Joint Air-to-Surface Standoff Missile (JASSM) for air-to-surface application.

The AGM-88 high-speed antiradiation missile has the mission of suppressing air defense threats. It can be employed in a prebriefed, self-protect, or target-of-opportunity mode. After launch, using a selfcontained, broadband, radar-emission homing seeker and a sophisticated signal processor, the missile autonomously searches for and acquires the target, and terminally guides itself to strike the radiating source.

The AGM-120 advanced medium-range air-to-air missile operates in a similar manner against aircraft targets beyond visual range. Its guidance consists of an active radar seeker in conjunction with an inertial reference unit, allowing the missile to be almost independent of the aircraft's fire control system. Once the missile closes in on the target, its active radar seeker autonomously locks on and guides it to intercept. This allows the pilot to simultaneously aim and fire several missiles at multiple targets, and then disengage from the attack and turn away before they hit.

The reduction of nuclear arms in the arsenals of the United States and Russia has resulted in the conversion of a number of cruise missiles from their nuclear role to that of conventional weapons. The Conventional Air Launched Cruise Missile (CALCM), AGM-86C, was developed to provide B52H bombers with a precision strike capability just in time for Desert Storm. The cruise missile was modified to use a GPS-aided inertial navigation system and carry a 3000-lb (1350-kg) blast fragmentation warhead.

Advanced weapons. Developmental weapons are all categorized as brilliant weapons and do not rely on a human in the loop, provide pinpoint accuracy, and use warheads tailored to maximize the probability of defeating the target with minimal collateral damage.

An improvement to the Paveway class of laser-guided weapons is the Autonomous Guided Conventional Weapon, which is designed to use infrared imagery to autonomously acquire high-value targets. The weapon is first loaded with navigation algorithms, target data, and the location of the intended launch point. After launch, the weapon is designed to autonomously detect, acquire, and attack specific targets by using an imaging infrared seeker, inertial

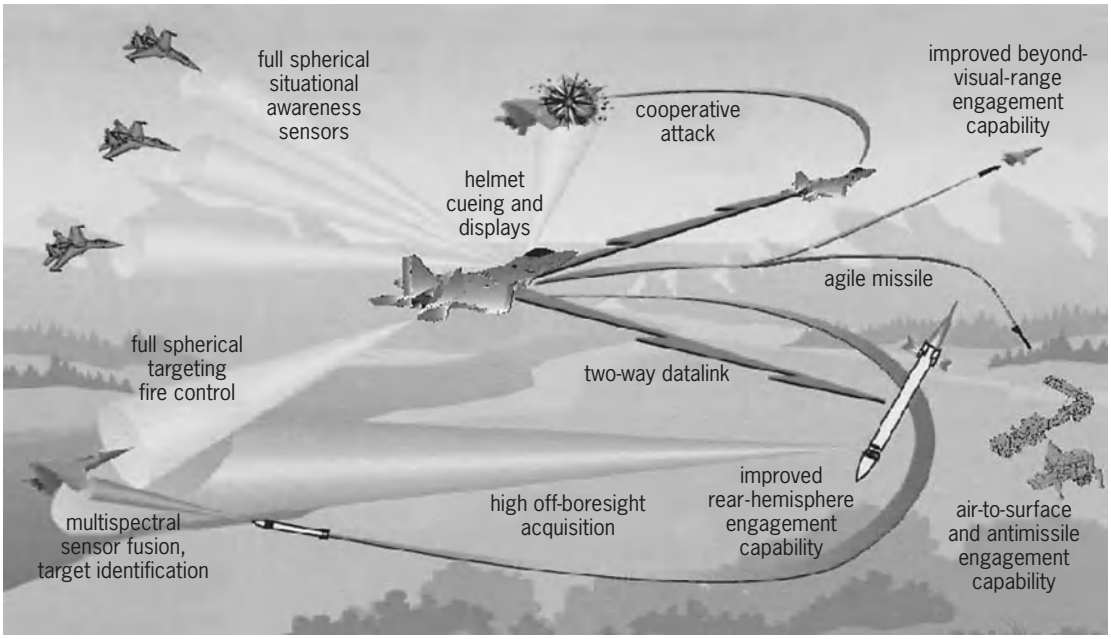

Fig. 1. Air Superiority Missile Concept.

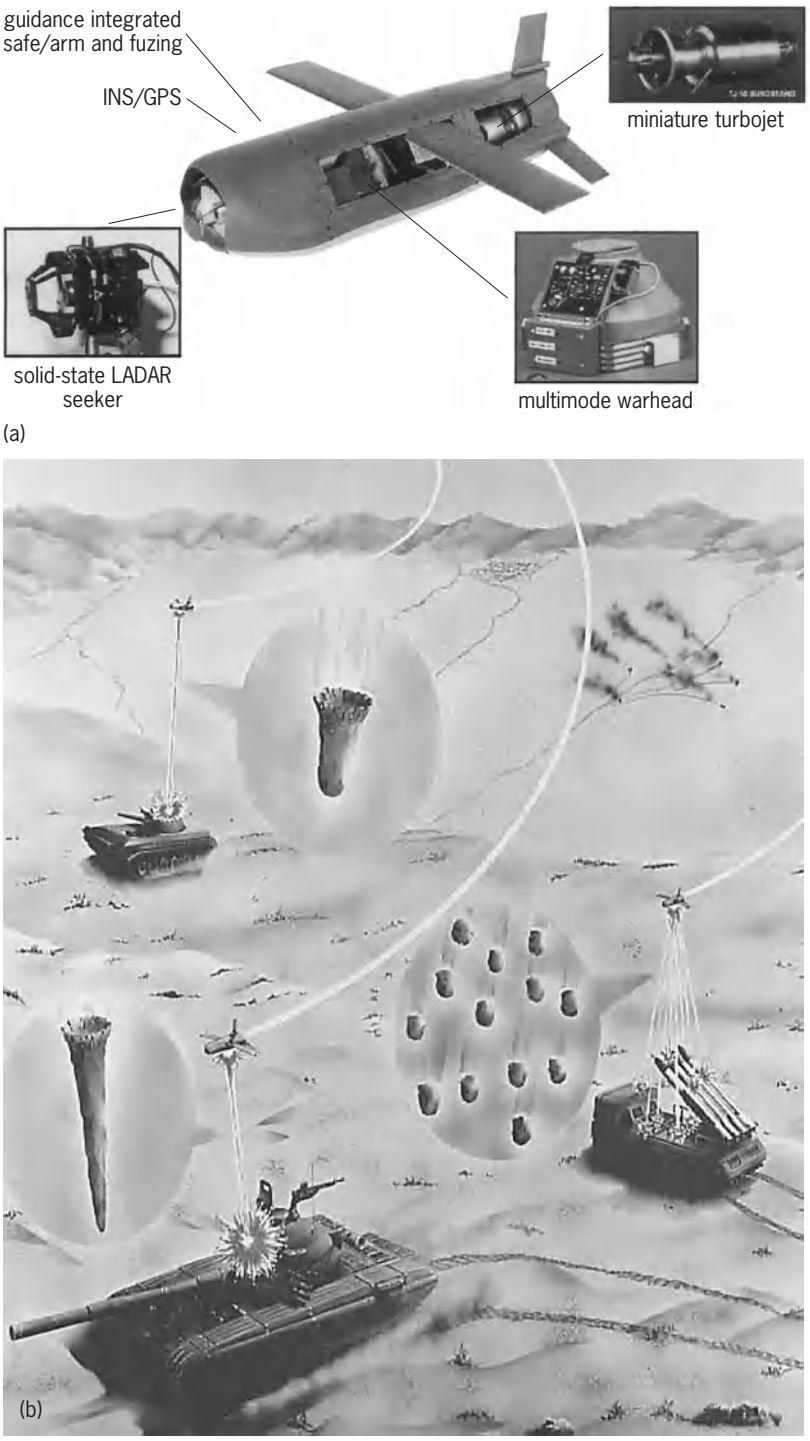

Fig. 2. Low-Cost Autonomous Attack System (LOCAAS; formerly Low-Cost Antiarmor Submunition). (*a*) Autonomous submunition, with length of 30 in. (0.75 m), wingspan of 40 in. (1.0 m), and weight of 90–100 lb (40–45 kg). (*b*) Various warhead modes.

navigation, and signal processing for target acquisition and aimpoint selection. Although this weapon appears to resemble a Paveway III, it is truly autonomous and does not require a human in the loop.

The Joint Direct Attack Munition is an autonomous precision strike weapon with an all-weather capability and consists of a conventional MK-84 general-purpose bomb and an inexpensive tail kit. The tail kit uses a GPS-aided inertial navigation system to derive guidance information, and controls the weapon's trajectory until target impact.

For air targets, the AIM-9X missile is being developed to replace the AIM-9S. It has increased maneuverability and a high degree of countermeasure resistance since it makes use of an imaging seeker. This missile is a stepping stone to the Air Superiority Missile Concept (**Fig. 1**), which will incorporate an advanced seeker and be extremely agile. Moreover, it will be capable of engaging targets anywhere in the vicinity of the launching aircraft and will rely on a helmet-mounted sight, or cueing system, in which the weapon or other sensors are controlled to automatically look in the same direction that the pilot looks.

Weapon systems for defeating mobile targets are evolving around the developmental Low-Cost Autonomous Attack System (LOCAAS, formerly the Low-Cost Anti Armor Submunition). This is an autonomous submunition (one of several similar units which are packaged together to form the munition or weapon) with a laser radar (LADAR) seeker capable of autonomously searching for targets, acquiring the target, performing target classification, and guiding the weapon to the target. The LADAR seeker develops an image of the target using high-resolution range measurements of the area in front of the submunition's flight path. Classification of the target is performed to enhance the probability of kill of the warhead by matching the mode of the warhead to the type of target encountered (**Fig. 2**).

The Joint Air-to-Surface Standoff Missile, in development, is designed to attack hard, soft, relocatable, and area-type targets at long standoff ranges comparable to those typical of a cruise missile scenario, that is, well beyond area defenses, so as not to risk launch platforms and crews. *See* ARMY ARMAMENT; GUIDED MISSILE; NAVAL ARMAMENT. Samuel Lambert

Bibliography. W. Armstrong, Jr., Replacing the Sidewinder, *Armed Forces J. Int.*, pp. 30–33, January, 1999; B. Gunston, *The Illustrated Encyclopedia of Aircraft Armament*, Orion Books, New York, 1988; E. Ulsamer, "Brilliant" weapons gather momentum, *Air Force Mag.*, pp. 74–80, August, 1988.

Air brake

A friction type of energy-conversion mechanism used to retard, stop, or hold a vehicle or other moving element. The activating force is applied by a difference in air pressure. With an air brake, a slight effort by the operator can quickly apply full braking force. *See* FRICTION.

The air brake, operated by compressed air, is used in buses; heavy-duty trucks, tractors, and trailers; and off-road equipment. The air brake is required by law on locomotives and railroad cars. The wheel-brake mechanism is usually either a drum or a disk brake. The choice of an air brake instead of a mechanical, hydraulic, or electrical brake depends partly on the

availability of an air supply and the method of brake control.

For example, in a bus in which compressed air actuates doors, air may also actuate the brakes. On railroad cars, compressed air actuates the brakes so that, in the event of a disconnect between cars, air valves can automatically apply the brakes on all cars. Air brakes can also be applied and held on mechanically, even if no compressed air is available and the vehicle is not in use. Regulations require alternate and fail-safe methods of applying air brakes.

The force with which air brakes are applied depends on the area of an air-driven diaphragm as well as on air pressure. As a result, a large activating force is developed from moderate pressure. On passenger trains, trucks, and buses, air brakes operate at about 110 psi (760 kilopascals).

System operation. In a motor vehicle, the air-brake system consists of three subsystems: the air-supply, air-delivery, and parking/emergency systems.

Air-supply system. This system includes the compressor, reservoirs, governor, pressure gage, low-pressure indicator, and safety valve. The engine-driven compressor takes in air and compresses it for use by the brakes and other air-operated components. The compressor is controlled by a governor that maintains air compression within a preselected range. The compressed air is stored in reservoirs. If the pressure in an air reservoir becomes too high, a safety valve allows air to escape. A pressure gage indicates the pressure within the system. A low-pressure indicator turns on a warning lamp or sounds a buzzer when the air pressure drops below a safe minimum for normal operation. *See* SAFETY VALVE.

Air-delivery system. This system includes a foot-operated brake valve, one or more relay valves, the quick-release valve, and the brake chambers. The system delivers compressed air from the air reservoirs to the brake chambers, while controlling the pressure of the air. The amount of braking is thereby regulated.

The brake valve meters the delivery of air pressure to the front brake chambers and the rear relay valve, relative to the distance that the brake-valve foot pedal is depressed. The rear relay valve is constantly supplied with full-system pressure from the primary reservoir. As the brake valve meters the pressure to the relay valve, the relay valve controls and delivers the air pressure to the rear brake chambers.

In the brake chambers, the air pressure is converted into a mechanical force to apply the brakes. As the pressure increases in each brake chamber, movement of the diaphragm pushrod forces the friction element against the rotating surface to provide braking. When the driver releases the brake valve, the quick-release valve and the relay valve release the compressed air from the brake chambers.

Parking/emergency system. This system includes a parking-brake control valve and spring brake chambers. These chambers contain a strong spring to mechanically apply the brakes (if the brakes are properly adjusted) when air pressure is not available. During normal vehicle operation, the spring is held compressed by system air pressure acting on a diaphragm.

For emergency stopping, the air-brake system is split into a front brake system and a rear brake system. If air pressure is lost in the front brake system, the rear brake system will continue to operate. However, the supply air will be depleted after several brake applications. Loss of air pressure in the rear brake system makes the front brake system responsible for stopping the vehicle, until the supply air is depleted.

In a semitrailer (a vehicle consisting of a tractor pulling a trailer), a tractor protection valve, or breakaway valve, automatically closes to apply the trailer brakes and preserve the tractor air supply if the trailer should break away and snap the air lines. The valve also closes if severe air leakage develops in the tractor or trailer.

Combined brake systems. A straight air-brake system, described above, may be combined with other brake systems. In an air-assisted hydraulic brake system, the hydraulic system is actuated by an air-hydraulic cylinder or power unit. The driver's effort is combined in the power unit with force from the air-brake cylinder piston or chamber diaphragm to displace fluid under pressure for actuation of the brakes. Other combined brake systems include the use of an air-powered hydraulic master cylinder actuated by an air-brake cylinder or chamber.

An antilock-braking system may also be used with air brakes to improve braking performance, reduce stopping distance, and help prevent trailer jackknifing. Thus, the vehicle can retain its directional stability and steerability even during severe braking on slippery surfaces. Another benefit is longer tire life resulting from reduced wear caused by tire skidding.

Railroad air brake. On railroad trains, a pressurized pipe runs the length of the train. Release of air pressure in the pipe reduces the pressure to the triple valve, which connects the compressor, auxiliary reservoir, and brake cylinder. The triple valve then admits compressed air from each car's auxiliary reservoir to the car's brake cylinders, which actuate the brake rods, or pushrods, to apply the brakes.

Vacuum brake. This alternate form of air brake maintains a vacuum in the actuating chamber or cylinder. By opening one side of the cylinder to the atmosphere, atmospheric pressure moves the diaphragm to apply the brakes.

Aircraft air brake. A different kind of air brake is used on aircraft. Energy of momentum is transferred to the air as heat by an air brake that consists of a flap or other device for producing drag. When needed, the braking device is extended from an aircraft wing or fuselage into the airstream. *See* AUTOMOTIVE BRAKE; BRAKE.

Donald L. Anglin

Bibliography. A. K. Baker, *Vehicle Braking*, 1986; R. Limpert, *Brake Design and Safety*, 2d ed., 1999; *Ford Truck Shop Manual*, annually; E. J. Schulz, *Diesel Equipment I*, 1982.

Air conditioning

The control of certain environmental conditions including air temperature, air motion, moisture level, radiant heat energy level, dust, various pollutants, and microorganisms.

Comfort air conditioning refers to control of spaces to promote the comfort, health, or productivity of the inhabitants. Spaces in which air is conditioned for comfort include residences, offices, institutions, sports arenas, hotels, factory work areas, and motor vehicles. Process air-conditioning systems are designed to facilitate the functioning of a production, manufacturing, or operational activity. There are many examples. Heat-producing electronic equipment in an airplane cockpit must be kept cool to function properly, while the occupants of the cockpit are maintained in comfortable conditions. A multicolor printing press requires accurate registration of the colors so the environment must be maintained at a constant relative humidity to avoid paper expansion or shrinkage, and press heat and ink mists must be removed as potential health hazards for personnel. Maintenance of conditions within surgical suites of hospitals and in "clean" or "white" rooms of manufacturing plants, where an atmosphere almost free of microorganisms and dust must be maintained, is a specialized subdivision of process air conditioning. *See* AUTOMOTIVE CLIMATE CONTROL.

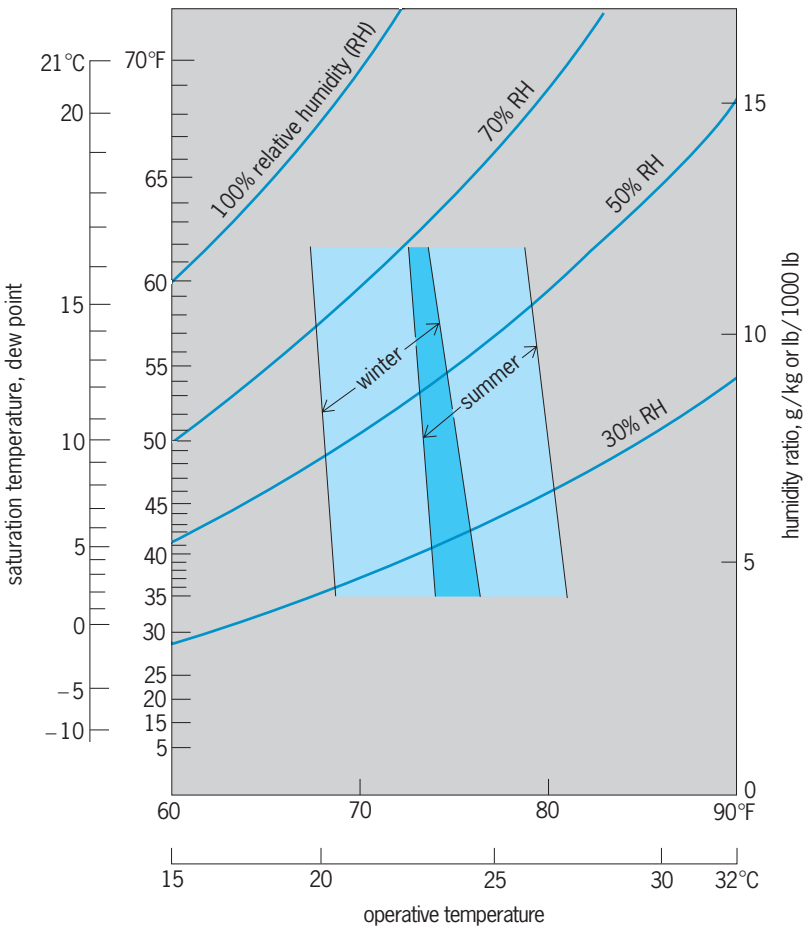

Fig. 1. Ranges of operative temperature and relative humidity acceptable to at least 80% of sedentary persons in typical summer or winter clothing and doing light activities. In the overlap area, people with summer clothing might feel slightly cool, while those in winter clothing might feel slightly warm.

Physiological principles. A comfort air-conditioning system is designed to help maintain body temperature at its normal level without undue stress and to provide an atmosphere which is healthy to breathe.

The human body produces heat at various rates, depending upon the person's weight and degree of activity. Normally more heat is generated than is required to maintain body temperature at a healthful level. Hence, proper air environment is required to permit normal cooling, even in winter. Heating the room air does not heat people; rather it affects the rate at which they lose heat and thereby affects their comfort and health. Since all people do not have the same metabolic rate or wear comparable clothing, some people in a room may feel comfortable while others do not. The acceptable ranges of operative temperatures and relative humidities shown in shaded areas for winter and summer in **Fig. 1** were developed by the American Society of Heating, Refrigerating and Air-Conditioning Engineers.

The comfort chart (Fig. 1) describes average responses to given environmental conditions. Preferences vary considerably from person to person. For instance, in office buildings women may complain of chill or drafts under conditions where men, in different attire, are comfortable. Metabolism rates vary with individuals. Whenever a building budget permits, an engineer designs an air-conditioning system that is flexible enough to allow adjustment by individual occupants of both temperature and air motion in the air-conditioned space. Conditions that would satisfy all persons within the same space have yet to be achieved. *See* PSYCHROMETRICS.

Control of body temperature is accomplished by control of the emission of energy from the body by radiation, by convection to air currents that impinge on the skin or clothing, by conduction of clothing and objects that are contacted, and by evaporation of moisture in the lungs and of sweat from the skin. Radiant emission from the body is a function of the amount of clothing worn (or blankets used) and the temperature of the surrounding air and objects, so that a body may experience a net loss by radiation to its surroundings or a net gain. Consequently, air-conditioning system operation is affected substantially by the basic construction of the building thermal envelope; cold outer walls induce more body heat loss for the same average room temperature than do warmer walls. Evaporation and convection heat losses are functions of air temperature and velocity. Evaporation is a function, in addition, of relative humidity.

When the amount and type of clothing and the temperature, velocity, and humidity of the air are such that the heat produced by the body is not dissipated at an equal rate, blood temperature begins to rise or fall and discomfort is experienced in the form of fever or chill, in proportion to the departure of body temperature from the normal 98.6°F

(37°C) [for individuals this standard may be slightly low or high]. Hence, space conditions to maintain comfort depend upon the degree of human activity in the space, the amount and type of clothing worn, and, to a certain extent, the physical condition of the occupants, because old age or sickness can impair the body's heat-producing and heat-regulating mechanisms.

The heat-dissipating factors of temperature, humidity, air motion, and radiant heat flow must be considered simultaneously. Within limits, the same amount of comfort (or, more objectively, of heat-dissipating ability) is the result of a combination of these factors in an enclosure. Conditions for constant comfort are related to the operative temperature. The perception of comfort is related to one's metabolic heat production, the transfer of this heat to the environment, and the resulting physiological adjustments and body temperature. The heat transfer is influenced by the environmental air temperature, thermal radiation, air movement, relative humidity, and the personal effects of activity and clothing. The effect of clothing is evaluated by its thermal resistance, or clo value; clo = 0.155 m^2 K/W (0.88 ft^2 h °F/Btu).

For some years, the effective temperature was used to indicate the uniform temperature of a radiantly black enclosure at 50% relative humidity in which an occupant would feel the same comfort, physiological strain, and heat exchange as in the actual environment with the same air motion. The preferred indicator is the operative temperature, the uniform temperature of a radiantly black enclosure in which an occupant would exchange the same amount of heat by radiation plus convection as in the actual nonuniform environment. Operative temperature is the average of the air and mean radiant temperatures weighted by the heat transfer coefficients for convection and radiation. At mean radiant temperatures less than 120°F (50°C) and air speeds of 80 ft/m (0.4 m/s) or less, the operative temperature is approximately the simple average of the air and the mean radiant temperature, and is equal to the adjusted dry bulb temperature.

In practice, most air-conditioning systems for offices, schools, and light-work areas are designed to maintain temperature in the range 68–75°F (20–24°C). In hot weather, in order to minimize the effects of sudden temperature differences when people leave temperature-controlled spaces to go outdoors, higher temperatures are maintained than in cold weather. Lower cold-weather temperatures are used when energy conservation is a factor. Depending on climatic characteristics, in hot weather the relative humidity is reduced to below 50%, while in cold weather it is increased above the 10–15% relative humidity that may occur when no humidification is provided. However, great care must be used in establishing relative humidity limits for each building design and occupancy so that condensation of moisture on and within building components does not create deteriorating conditions, many of which may not be detectable until performance problems occur.

Calculation of loads. Engineering of an air-conditioning system starts with selection of design conditions; air temperature and relative humidity are principal factors. Next, loads on the system are calculated. Finally, equipment is selected and sized to perform the indicated functions and to carry the estimated loads.

Design conditions are selected on the bases discussed above. Each space is analyzed separately. A cooling load will exist when the sum of heat released within the space and transmitted to the space is greater than the loss of heat from the space. A heating load occurs when the heat generated within the space is less than the loss of heat from it. Similar considerations apply to moisture.

Heat generated within the space consists of body heat, heat from all electrical appliances and lights, and heat from other sources such as cooking stoves and industrial ovens. Heat is transmitted through all parts of the space envelope, which includes walls, floor, slab on ground, ceiling, doors, and windows. Whether heat enters or leaves the space depends upon whether the outside surfaces are warmer or cooler than the inside surfaces. The rate at which heat is conducted through the building envelope is a function of the temperature difference across the envelope and the thermal resistance of the envelope (R-value). Overall R-values depend on materials of construction and their thickness along the path of heat flow, and air spaces with or without reflectances and emittances, and are evaluated for walls and roofs exposed to outdoors, and basements or slab exposed to earth. In some cases, thermal insulations may be added to increase the R-value of the envelope.

Solar heat loads are an especially important part of load calculation because they represent a large percentage of heat gain through walls, windows, and roofs, but are very difficult to estimate because solar irradiation is constantly changing. Intensity of radiation varies with the seasons [it rises to 457 Btu/h ft^2 (2881 W/m^2) in midwinter and drops to 428 Btu/h ft^2 (2698 W/m^2) in midsummer]. Intensity of solar irradiation also varies with surface orientation. For example, the half-day total for a horizontal surface at 40° north latitude on January 21 is 353 Btu/h ft^2 (2225 W/m^2) and on June 21 it is 1121 Btu/h ft^2 (7067 W/m^2), whereas for a south wall on the same dates comparable data are 815 Btu/h ft^2 (5137 W/m^2) and 311 Btu/h ft^2 (1961 W/m^2), a sharp decrease in summer. Intensity also varies with time of day and cloud cover and other atmospheric phenomena. *See* SOLAR RADIATION.

The way in which solar radiation affects the space load depends also upon whether the rays are transmitted instantly through glass or impinge on opaque walls. If through glass, the effect begins immediately but does not reach maximum intensity until the interior irradiated surfaces have warmed sufficiently to reradiate into the space, warming the air. In the case of irradiated walls and roofs, the effect is as if the outside air temperature were higher than it is. This apparent temperature is called the sol-air temperature, tables of which are available.

In calculating all these heating effects, the object is proper sizing and intelligent selection of equipment; hence, a design value is sought which will accommodate maximums. However, when dealing with climatic data, which are statistical historical summaries, record maximums are rarely used. For instance, if in a particular locality the recorded maximum outside temperature was 100°F (38°C), but 95°F (35°C) was exceeded only four times in the past 20 years, 95°F may be chosen as the design summer outdoor temperature for calculation of heat transfer through walls. In practice, engineers use tables of design winter and summer outdoor temperatures which list winter temperatures exceeded more than 99% and 97.5% of the time during the coldest winter months, and summer temperatures not exceeded 1%, 2.5%, and 5% of the warmest months. The designer will select that value which represents the conservatism required for the particular type of occupancy. If the space contains vital functions where impairment by virtue of occasional departures from design space conditions cannot be tolerated, the more severe design outdoor conditions will be selected.

In the case of solar load through glass, but even more so in the case of heat transfer through walls and roof, because outside climate conditions are so variable, there may be a considerable thermal lag. It may take hours before the effect of extreme high or low temperatures on the outside of a thick masonry wall is felt on the interior surfaces and space. In some cases the effect is never felt on the inside, but in all cases the lag exists, exerting a leveling effect on the peaks and valleys of heating and cooling demand; hence, it tends to reduce maximums and can be taken advantage of in reducing design loads.

Humidity as a load on an air-conditioning system is treated by the engineer in terms of its latent heat, that is, the heat required to condense or evaporate the moisture, approximately 1000 Btu/lb (2324 kilojoules/kg) of moisture. People at rest or at light work generate about 200 Btu/h (586 W). Steaming from kitchen activities and moisture generated as a product of combustion of gas flames, or from all drying processes, must be calculated. As with heat, moisture travels through the space envelope, and its rate of transfer is calculated as a function of the difference in vapor pressure across the space envelope and the permeance of the envelope construction.

Years ago engineers and scientists contributed their experiences with moisture migration rates into walls of ordinary houses in ordinary climates; this led to defining a limiting rate of moisture flow that was designated a perm (from permeance). Permeance is the rate of moisture flow through a material or construction as used, whereas permeability is a property of a material that expresses moisture flow rate for a unit thickness. A unit perm was considered the highest rate of water vapor flow into a wall (or roof) that would differentiate the probability of moisture problems from the probability of no moisture problems. As stated above, excessive moisture adversely affects the thermal performance of constructions, and thereby the condition of air for comfort and health of people; in addition, it affects energy waste that could be avoided.

While moisture migrates by diffusion, building designs and the air-conditioning designs must also consider the effects of air infiltration and exfiltration, especially from wind in a particular location.

Another load-reducing factor to be calculated is the diversity among the various spaces within a building or building complex served by a single system. Spaces with east-facing walls experience maximum solar loads when west-facing walls have no solar load. In cold weather, rooms facing south may experience a net heat gain due to a preponderant solar load while north-facing rooms require heat. An interior space, separated from adjoining spaces by partitions, floor, and ceiling across which there is no temperature gradient, experiences only a net heat gain, typically from people and lights. Given a system that can transfer this heat to other spaces requiring heat, the net heating load may be zero, even on cold winter days.

Air-conditioning systems. A complete air-conditioning system is capable of adding and removing heat and moisture and of filtering airborne substitutes, such as dust and odorants, from the space or spaces it serves. Systems that heat, humidify, and filter only, for control of comfort in winter, are called winter air-conditioning systems; those that cool, dehumidify, and filter only are called summer air-conditioning systems, provided they are fitted with proper controls to maintain design levels of temperature, relative humidity, and air purity.

Design conditions may be maintained by multiple independent subsystems tied together by a single control system. Such arrangements, called split systems, might consist, for example, of hot-water baseboard heating convectors around a perimeter wall to offset window and wall heat losses when required, plus a central cold-air distribution system to pick up heat and moisture gains as required and to provide filtration for dust and odor. *See* HOT-WATER HEATING SYSTEM.

Air-conditioning systems are either unitary or built-up. The window or through-the-wall air conditioner (**Fig. 2**) is an example of a unitary summer air-conditioning system; the entire system is housed in a single package which contains heat removal, dehumidification, and filtration capabilities. When an electric heater is built into it with suitable controls, it functions as a year-round air-conditioning system. Unitary air conditioners are manufactured in capacities as high as 100 tons (1 ton of air conditioning equals 12,000 Btu/h or 76,000 W/m^2) and are designed to be mounted conveniently on roofs, on the ground, or other convenient location, where they can be connected by ductwork to the conditioned space.

Built-up or field-erected systems are composed of factory-built subassemblies interconnected by means such as piping, wiring, and ducting during final assembly on the building site. Their capacities range up to thousands of tons of refrigeration and millions of Btu per hour of heating. Most large buildings are so conditioned.

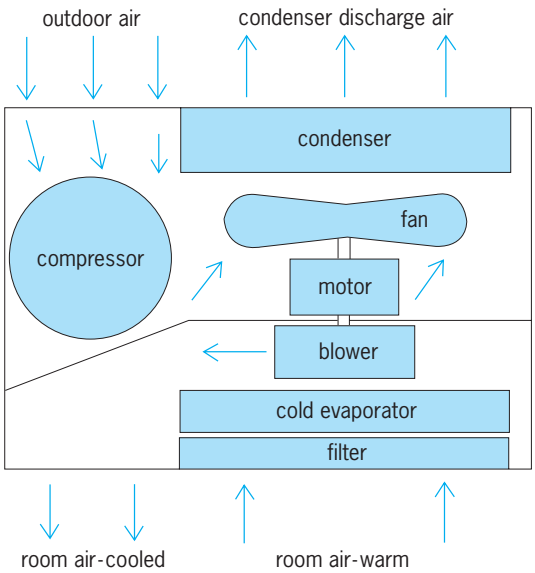

Fig. 2. Schematic of room air conditioner. (*After American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Inc., Guide and Data Book, 1967*)

Another important and somewhat parallel distinction can be made between incremental and central systems. An incremental system serves a single space; each space to be conditioned has its own, self-contained heating-cooling-dehumidifying-filtering unit. Central systems serve many or all of the conditioned spaces in a building. They range from small, unitary packaged systems serving single-family residences to large, built-up or field-erected systems serving large buildings. *See* CENTRAL HEATING AND COOLING.

When many buildings, each with its own air-conditioning system which is complete except for a refrigeration and a heating source, are tied to a central plant that distributes chilled water and hot water or steam, the interconnection is referred to as a district heating and cooling system. This system is especially useful for campuses, medical complexes, and office complexes under a single management. *See* DISTRICT HEATING.

Conditioning of spaces. Air temperature in a space can be controlled by radiant panels in floor, walls, or ceiling to emit or absorb energy, depending on panel temperature. Such is the radiant panel system. However, to control humidity and air purity, and in most systems for controlling air temperature, a portion of the air in the space is withdrawn, processed, and returned to the space to mix with the remaining air. In the language of the engineer, a portion of the room air is returned (to an air-handling unit) and, after being conditioned, is supplied to the space. A portion of the return air is spilled (exhausted to the outdoors) while an equal quantity (of outdoor air) is brought into the system and mixed with the remaining return air before entering the air handler. *See* PANEL HEATING AND COOLING.

Typically, the air-handling unit contains a filter, a cooling coil, a heating coil, and a fan in a suitable casing (**Fig. 3**). The filter removes dust from both return and outside air. The cooling coil, either containing recirculating chilled water or boiling refrigerant, lowers air temperature sufficiently to dehumidify it to the required degree. The heating coil, in winter, serves a straightforward heating function, but when the cooling coil is functioning, it serves to raise the temperature of the dehumidified air (to reheat it) to the exact temperature required to perform its cooling function. The air handler may perform its function, in microcosm, in room units in each space as part of a self-contained, unitary air conditioner, or it may be a huge unit handling return air from an entire building. *See* AIR COOLING; AIR FILTER; HUMIDITY CONTROL.

There are three principal types of central air-conditioning systems: all-air, all-water, and air-processed in a central air-handling apparatus. In one type of all-air system, called dual-duct, warm air and chilled air are supplied to a blending or mixing unit in each space. In a single-duct all-air system, air is supplied at a temperature for the space requiring the coldest air, then reheated by steam or electric or hot-water coils in each space.

In the all-water system the principal thermal load is carried by chilled and hot water generated in a central facility and piped to coils in each space; room air then passes over the coils. A small, central air system supplements the all-water system to provide dehumidification and air filtration. The radiant panel system, previously described, may also be in the form of an all-water system.

In an air-water system, both treated air and hot or chilled water are supplied to units in each space. In winter, hot water is supplied, accompanied by cooled, dehumidified air. In summer, chilled water is supplied with warmer (but dehumidified) air. One supply reheats the other.

All-air systems preceded the others. Primary motivation for all-water and air-water systems is their capacity for carrying large quantities of heat energy in small pipes, rather than in larger air ducts. To accomplish the same purpose, big-building all-air

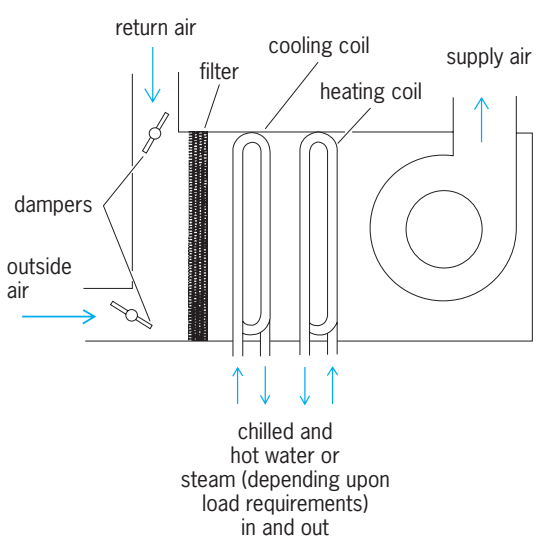

Fig. 3. Schematic of central air-handling unit.

systems are designed for high velocities and pressures, requiring much smaller ducts.

Richard L. Koral; E. C. Shuman

Bibliography. B. C. Langley, *Fundamentals of Air Conditioning Systems*, 2d ed., 2000; C. McPhee (ed.), *ASHRAE Pocket Guide for Air-Conditioning, Heating, Ventilation, and Refrigeration*, 2005; F. C. McQuiston, J. D. Parker, and J. D. Spitler, *Heating, Ventilating and Air Conditioning Analysis and Design*, 2004.

Air cooling

Lowering of air temperature for comfort, process control, or food preservation. Air and water vapor occur together in the atmosphere. The mixture is commonly cooled by direct convective heat transfer of its internal energy (sensible heat) to a surface or medium at lower temperature. In the most compact arrangement, transfer is through a finned (extended surface) coil, metallic and thin, inside of which is circulating either chilled water, antifreeze solution, brine, or boiling refrigerant. The fluid acts as the heat receiver. Heat transfer can also be directly to a wetted surface, such as waterdroplets in an air washer or a wet pad in an evaporative cooler. *See* COMFORT TEMPERATURES; HEAT TRANSFER.

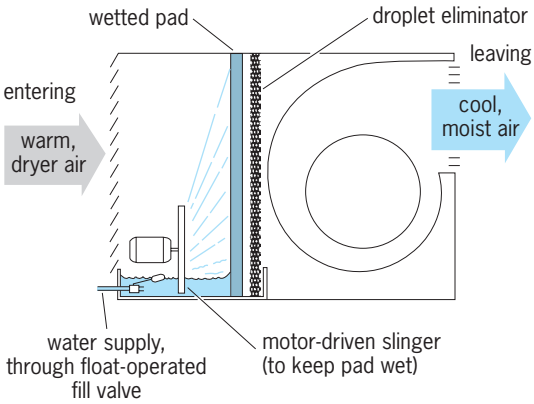

Fig. 1. Schematic view of simple evaporative air cooler.

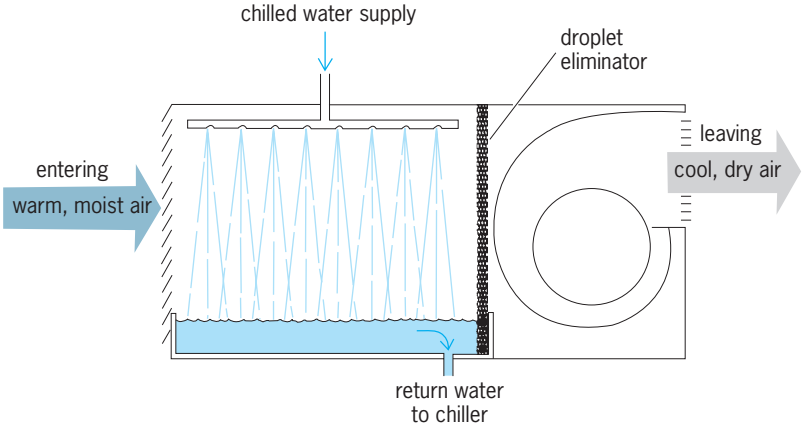

Fig. 2. Schematic of air washer.

Evaporative cooling. For evaporative cooling, nonsaturated air is mixed with water. Some of the sensible heat transfers from the air to the evaporating water. The heat then returns to the airstream as latent heat of water vapor. The exchange is thermally isolated (adiabatic) and continues until the air is saturated and air and water temperatures are equal. With suitable apparatus, air temperature approaches within a few degrees of the theoretical limit, the wet-bulb temperature. Evaporative cooling is frequently carried out by blowing relatively dry air through a wet mat (**Fig. 1**). The technique is employed for air cooling of machines where higher humidities can be tolerated; for cooling of industrial areas where high humidities are required, as in textile mills; and for comfort cooling in hot, dry climates, where partial saturation results in cool air at relatively low humidity. *See* HUMIDITY.

Air washer. In the evaporative cooler the air is constantly changed and the water is recirculated, except for that portion which has evaporated and which must be made up. Water temperature remains at the adiabatic saturation (wet-bulb) temperature. If water temperature is controlled, as by refrigeration, the leaving air temperature can be controlled within wide limits. Entering warm, moist air can be cooled below its dew point so that, although it leaves close to saturation, it leaves with less moisture per unit of air than when it entered. An apparatus to accomplish this is called an air washer (**Fig. 2**). It is used in many industrial and comfort air-conditioning systems, and performs the added functions of cleansing the airstream of dust and of gases that dissolve in water, and in winter, through the addition of heat to the water, of warming and humidifying the air.

Air-cooling coils. The most important form of air cooling is by finned coils, inside of which circulates a cold fluid or cold, boiling refrigerant (**Fig. 3**). The latter is called a direct-expansion (DX) coil. In most applications the finned surfaces become wet as condensation occurs simultaneously with sensible cooling. Usually, the required amount of dehumidification determines the temperature at which the surface is maintained and, where this results in air that is colder than required, the air is reheated to the proper temperature. Droplets of condensate are entrained in the airstream, removed by a suitable filter (eliminator), collected in a drain pan, and wasted.

In the majority of cases, where chilled water or boiling halocarbon refrigerants are used, aluminum fins on copper coils are employed. Chief advantages of finned coils for air cooling are (1) complete separation of cooling fluid from airstream, (2) high velocity of airstream limited only by the need to separate condensate that is entrained in the airstream, (3) adaptability of coil configuration to requirements of different apparatus, and (4) compact heat-exchange surface.

Defrosting. Wherever air cooling and dehumidification occur simultaneously through finned coils, the coil surface must be maintained above 32°F (0°C) to prevent accumulation of ice on the coil. For this reason, about 35°F (1.7°C) is the lowest-temperature

air that can be provided by coils (or air washers) without ice accumulation. In cold rooms, where air temperature is maintained below 32°F (0°C), provision is made to deice the cooling coils. Ice buildup is sensed automatically; the flow of cold refrigerant to the coil is stopped and replaced, briefly, by a hot fluid which melts the accumulated frost. In direct-expansion coils, defrosting is easily accomplished by bypassing hot refrigerant gas from the compressor directly to the coil until defrosting is complete.

Cooling coil sizing. Transfer of heat from warm air to cold fluid through coils encounters three resistances: air film, metal tube wall, and inside fluid film. Overall conductance of the coil, U, is shown in the equation below, where K_o is film conductance of the out-

$$\frac{1}{U} = \frac{1}{K_o} + r_m + \frac{R}{K_i}$$

side (air-side) surface in Btu/(h)(ft^2)(F); r_m is metal resistance in (h)(ft^2)(F)/Btu, where area is that of the outside surface; K_i is film conductance of the inside surface (water, steam, brine, or refrigerant side) in Btu/(h)(ft^2)(F); U is overall conductance of transfer surface in Btu/(h)(ft^2)(F), where area again refers to the outside surface; and R is the ratio of outside surface to inside surface.

Values of K_o are a function of air velocity and typically range from about 4 Btu/(h)(ft^2)(F) at 100 feet per minute (fpm) to 12 at 600 fpm. If condensation takes place, latent heat released by the condensate is in addition to the sensible heat transfer. Then total (sensible plus latent) K_o increases by the ratio of total to sensible heat to be transferred, provided the coil is well drained.

Values of r_m range from 0.005 to 0.030 (h)(ft^2)(F)/Btu, depending somewhat on type of metal but primarily on metal thickness.

Typical values for K_i range from 250 to 500 Btu/(h)(ft^2)(F) for boiling refrigerant. In 40°F (4.4°C) chilled water, values range from about 230 Btu/(h)(ft^2)(F) when water velocity is 1 foot per second (fps) to 1250 when water velocity is 8 fps.

Use of well water. Well water is available for air cooling in much of the world. Temperature of water from wells 30 to 60 ft (9 to 18 m) deep is approximately the average year-round air temperature in the locality of the well, although in some regions overuse of available supplies for cooling purposes and recharge of ground aquifers with higher-temperature water has raised well water temperature several degrees above the local normal. When well water is not cold enough to dehumidify air to the required extent, an economical procedure is to use it for sensible cooling only, and to pass the cool, moist air through an auxiliary process to dehumidify it. Usually, well water below 50°F (10°C) will dehumidify air sufficiently for comfort cooling. Well water at these temperatures is generally available in the northern third of the United States, with the exception of the Pacific Coast areas.

Ice as heat sink. For installations that operate only occasionally, such as some churches and meeting

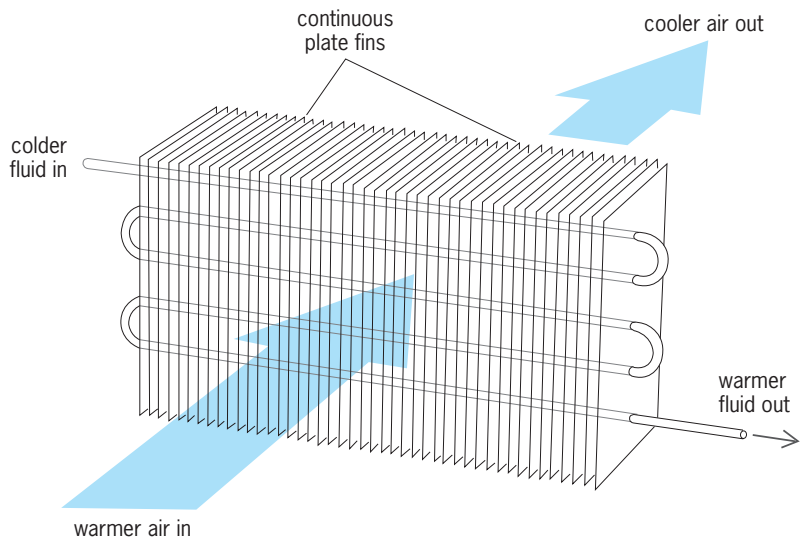

Fig. 3. Typical extended-surface air-cooling coil.

halls, water recirculated and cooled over ice offers an economical means for space cooling (**Fig. 4**). Cold water is pumped from an ice bunker through an extended-surface coil. In the coil the water absorbs heat from the air, which is blown across the coil. The warmed water then returns to the bunker, where its temperature is again reduced by the latent heat of fusion (144 Btu/lb) to 32°F (0°C). Although initial cost of such an installation is low, operating costs are usually high.

Refrigeration heat sink. Where electric power is readily available, the cooling function of the ice, as described above, is performed by a mechanical refrigerator. If the building complex includes a steam plant, a steam-jet vacuum pump can be used to cause the water to evaporate, thereby lowering its temperature by the latent heat of evaporation (about 1060 Btu/lb, depending on temperature and pressure). High-pressure steam, in passing through a primary ejector, aspirates water vapor from the evaporator, thereby maintaining the required low pressure that causes the water to evaporate and thus to cool itself (**Fig. 5**). *See* REFRIGERATION.

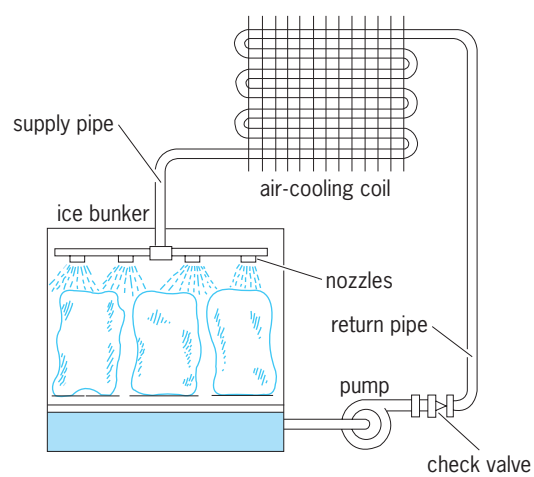

Fig. 4. Air cooling by circulating ice-cooled water.

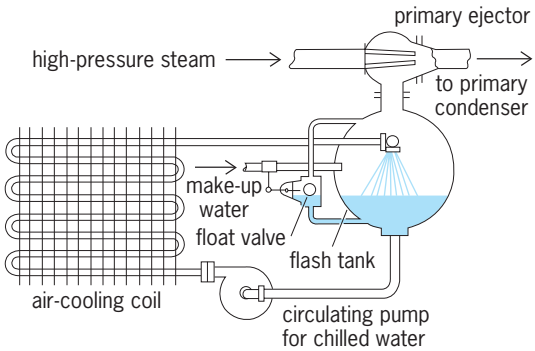

Fig. 5. Air cooling by circulating water that is cooled, in turn, by evaporation in flash tank.

Where electric power is costly compared to low-temperature heat, such as by gas, absorption refrigeration may be used. Two fluids are used: an absorbent and a refrigerant. The absorbent is chosen for its affinity for the refrigerant when in vapor form; for example, water is used as the absorber with ammonia as the refrigerant. Concentrated ammonia water is pumped to a high pressure and then heated to release the ammonia. The high-pressure ammonia then passes through a condenser, an expansion valve, and an evaporator, as in a mechanical system, and is reabsorbed by the water. The cycle cools air circulated over the evaporator. *See* AIR CONDITIONING.

Richard L. Koral

Bibliography. Air Conditioning and Refrigeration Institute, *Refrigeration and Air Conditioning*, 3d ed., 1998; A. Althouse et al., *Modern Refrigeration and Air Conditioning,* 18th ed., 2003; American Society of Heating and Air Conditioning Engineers, *Guide and Data Book,* annual; L. Jeffus, *Refrigeration and Air Conditioning: An Introduction to HVAC*, 4th ed., 2003.

Air-cushion vehicle

A transportation vehicle, also called a hovercraft, that rides slightly above the Earth's surface on a cushion of air. The air is continuously forced under the vehicle by a fan, generating the cushion that greatly reduces friction between the moving vehicle and the surface. The air is usually delivered through ducts and injected at the periphery of the vehicle in a downward and inward direction (**Fig. 1**). The design of the vehicle's underside, combined with seals or skirts attached below the hull around the perimeter, restrains the air, creating the cushion. Because the vehicle is not in contact with the surface, it has six dimensions of motion. *See* DEGREE OF FREEDOM (MECHANICS); DUCTED FAN.

Generally, an air-cushion vehicle is an amphibious aerostatic craft capable of slight vertical lift regardless of forward speed. This type of air-cushion vehicle can operate equally well over ice, water, marsh, or relatively level land. A variation is the surface-effect ship (SES), which has rigid side hulls that ride in the water like a catamaran, containing the air cushion and reducing air loss. Flexible air seals across the forward and aft ends of the surface-effect ship reduce drag, while the side hulls and the use of screw propellers or waterjets increase drag. These also limit operation to water and make the surface-effect ship nonamphibious. *See* AERODYNAMIC FORCE.

Design and construction. The air-cushion vehicle is basically a load-carrying hull fitted with a seal or skirt system to contain the air cushion, a lift system, a propulsion system, and a control system. One or more engines supply power to provide lift and propulsion. Some air-cushion vehicles have gas turbine engines which drive free or ducted air propellers through gearboxes. Depending on size, other air-cushion vehicles have diesel engines, automotive engines, or combination power sources such as a lawnmower engine for lift and a snowmobile engine for propulsion. *See* GAS TURBINE.

Many air-cushion vehicles have a length-to-beam ratio of approximately 2:1. Speed generally is in the 30–100-mi/h (50–160-km/h) range. Cushion pressure usually is in the range of 30–50 lb/ft^2 (1.50–2.40 kilopascals), although some air-cushion vehicles have used a pressure of approximately 100 lb/ft^2 (4.80 kPa). Cushion depth varies from about 8 in. (20 cm) on recreational vehicles carrying two or three passengers to 9 ft (3 m) on large vehicles capable of carrying hundreds of passengers. A greater cushion depth usually allows an increase in the maximum wave height over which the vehicle can operate.

Applications. Commercial uses of air-cushion vehicles include transportation, supply, and pipeline and cable inspection and maintenance. In scheduled service, large air-cushion vehicles have been used to ferry cars and passengers across the English Channel (**Fig. 2**). These craft displace 336 short tons (305 metric tons), and have a maximum speed of 65 knots (120 km/h), with propulsion provided by swivel-mounted air propellers. Military missions of aircushion vehicles, which can be lightly armed, include patrolling, and transporting troops and equipment from ship to shore and across the beach. Small air-cushion vehicles are available in single-seat, two-seat, and four-seat versions. Most are personal sport craft for recreational use on land or over calm water. Air-cushion vehicles of various sizes and maximum speeds are employed by government agencies in law enforcement, fire fighting, and disaster relief. *See* NAVAL SURFACE SHIP.

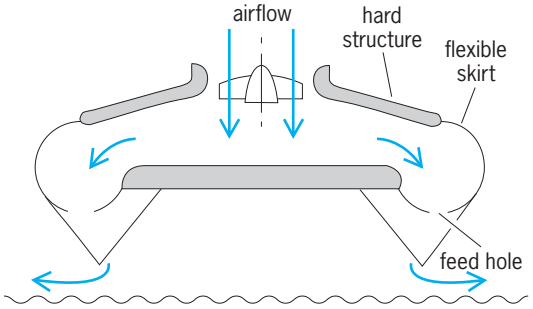

Fig. 1. Basic construction of an aerostatic air-cushion vehicle.

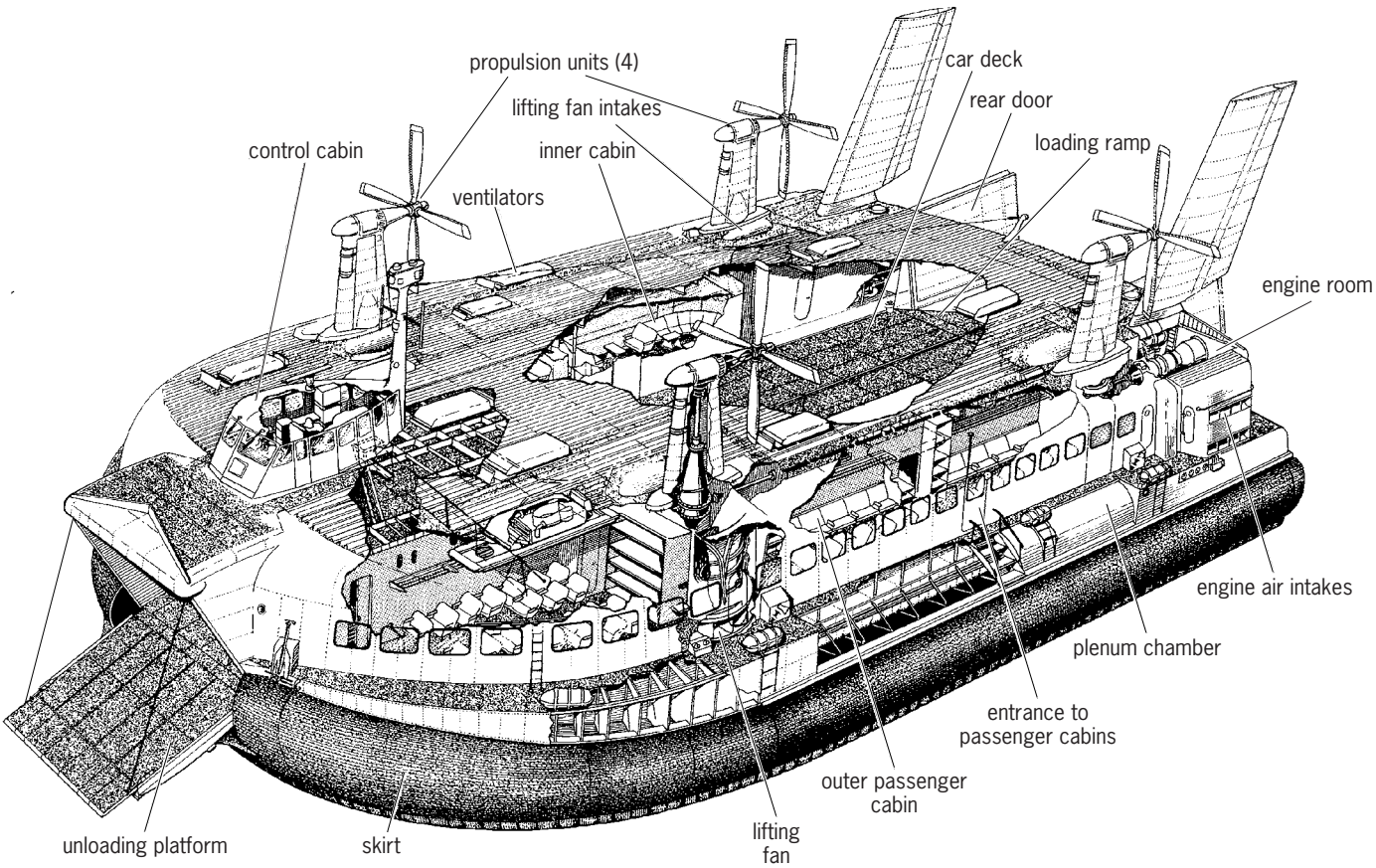

Fig. 2. Air-cushion vehicle used to transport cars and passengers across the English Channel. (*British Hovercraft Corp.*)

Wing-in-ground-effect craft. Another type of vehicle relying on a cushion of air for support is the wing-in-ground-effect (WIG) craft, also called the wing-in-surface-effect (WISE) craft. Similar in design to an aircraft, the ground-effect craft is an aerodynamic vehicle that generally travels only over water and requires forward speed to provide lift. It is designed to take advantage of two characteristics of an airfoil moving through air while close to the surface: (1) Air pressure builds up under the wing, producing greater lift than if the wing were in conventional flight; and (2) the induced drag of the wing is less than in conventional flight because wingtip vortices are weaker and downwash angle decreases. The result is an increase in the ratio of lift to drag when a craft is flying in ground effect. *See* AERODYNAMIC FORCE; AERODYNAMICS; AIRCRAFT; AIRCRAFT DESIGN; AIRFOIL; AIRPLANE.

A ground-effect vehicle can achieve greater fuel economy and range than an aircraft in conventional flight, since flying at a height of about 10% of wingspan reduces induced drag by about 50%. However, the ceiling and maneuverability of the ground-effect vehicle are limited, although some vehicles can fly out of ground effect but only with great loss in efficiency. Cruising speed of most ground-effect vehicles has been in the range of 70–170 mi/h (110–270 km/h), with some models cruising at 310–340 mi/h (500–550 km/h).

Beginning in the 1960s, the Russians built and flew a variety of wing-in-ground-effect craft, called ekranoplans, which were developed primarily for military applications. One ekranoplan displaced up to 595 short tons (540 metric tons), making it probably the largest vehicle to ever lift itself from the water and into the air under its own power. The Russians and other designers of ground-effect craft have used ducted fans, deflected exhaust, and separate turbofan engines to blow air or exhaust under the wings to augment takeoff power. *See* JET PROPULSION; TURBOFAN.

Several types and sizes of wing-in-ground-effect vehicles are available (**Fig. 3**). Some designs utilize

Fig. 3. Wing-in-ground-effect craft used over water as a commuter vehicle. (*FlareCraft Corp.*)

composite construction and are comparable in performance with hydrofoils. The first wing-in-ground-effect vehicle in commercial service served as a water taxi. *See* HYDROFOIL CRAFT. Donald L. Anglin

Bibliography. *Jane's High Speed Marine Transportation*, Jane's Information Group Ltd., London, 1999; Y. Liang and A. Biaut, *Theory and Design of Air Cushion Craft*, Wiley, New York, 1999; Special edition: Modern ships and craft, *Nav. Eng. J.*, 97(2):1–336, February 1985.

Air filter

A component of most systems in which air is used for industrial processes, for ventilation, or for comfort air conditioning. The function of an air filter is to reduce the concentration of solid particles in the airstream to a level that can be tolerated by the process or space occupancy purpose. Degrees of cleanliness required and economics of the situation (life cycle costs) influence the selection of equipment. *See* AIR; AIR CONDITIONING; VENTILATION.

Solid particles in the airstream range in size from 0.01 micrometer (the smallest particle visible to the naked eye is estimated to be 20 micrometers in diameter) up to things that can be caught by ordinary fly screens, such as lint, feathers, and insects. The particles generally include soot, ash, soil, lint, and smoke, but may include almost any organic or inorganic material, even bacteria and mold spores. This wide variety of airborne contaminants, added to the diversity of systems in which air filters are used, makes it impossible to have one type that is best for all applications.

Three basic types of air filters are in common use today: viscous impingement, dry, and electronic. The principles employed by these filters in removing airborne solids are viscous impingement, interception, impaction, diffusion, and electrostatic precipitation. Some filters utilize only one of these principles; others employ combinations. A fourth method, inertial separation, is finding increasing use as a result of the construction boom throughout most of the Middle East.

Viscous impingement filters. The viscous impingement filter is made up of a relatively loosely arranged medium, usually consisting of spun glass fibers, metal screens, or layers of crimped expanded metal. The surfaces of the medium are coated with a tacky oil, generally referred to as an adhesive. The arrangement of the filter medium is such that the airstream is forced to change direction frequently as it passes through the filter. Solid particles, because of their momentum, are thrown against, and adhere to, the viscous coated surfaces. Larger airborne particles, having greater mass, are filtered in this manner, whereas small particles tend to follow the path of the airstream and escape entrapment.

Operating characteristics of viscous impingement filters include media velocities ranging from 300 to 500 ft/min (1.5–3 m/s), with resistance to airflow about 0.10 in. (2.5 mm) of water gage, or w.g., when clean and 0.50 in. (13 mm) w.g. when dirty. A glass-fiber filter is thrown away when the final resistance is reached; the metal type is washed, dried, recoated with adhesive, and reused. Automatic types, employing rolls of glass-fiber material in blanket form, or an endless belt of metal plates which pass through an oil-filled reservoir to remove the accumulated dirt, are often chosen to minimize maintenance labor.

Dry filters. Dry-air filters are the broadest category of air filters in terms of the variety of designs, sizes, and shapes in which they are manufactured. The most common filter medium is glass fiber. Other materials used are cellulose paper, cotton, and polyurethane and other synthetics. Glass fiber is used extensively because of its relatively low cost and the unique ability to control the diameter of the fiber in manufacture (in general, the finer the fiber diameter, the higher will be the air-cleaning efficiency of the medium). *See* MANUFACTURED FIBER; NATURAL FIBER.

Dry filters employ the principles of interception, in which particles too large to pass through the filter openings are literally strained from the airstream; impaction, in which particles strike and stick to the surfaces of the glass fibers because of natural adhesive forces, even though the fibers are not coated with a filter adhesive; and diffusion, in which molecules of air moving in a random pattern collide with very fine particles of airborne solids, causing the particles to have random movement as they enter the filter media. It is this random movement which enhances the likelihood of the particles coming in contact with the fibers of filter media as the air passes through the filter. Through the process of diffusion a filter is able to separate from the airstream particles much smaller than the openings in the medium itself. *See* MOLECULAR ADHESION.

Most dry filters are of the extended surface type, with the ratio of filtered area to face area (normal to the direction of air travel) varying from 7.5:1 to as much as 50:1. The higher the filter efficiency, the greater the ratio of areas, in order to hold down the air friction loss. Clean resistance through dry filters can be as low as 0.10 in. (2.5 mm) w.g. or, for very-high-efficiency filters, as much as 1 in. (25 mm) w.g. These filters are generally allowed to increase 0.5–1 in. (13–25 mm) in pressure drop before being changed. Face velocities from 400 ft/min (2 m/s) to 600 ft/min (3 m/s) are common for this class of filter, although newer energy-conscious design favors the lower portion of this range.

Electronic air cleaners. Limited primarily to applications requiring high air-cleaning efficiency, these devices operate on the principle of passing the airstream through an ionization field where a 12,000-V potential imposes a positive charge on all airborne particles. The ionized particles are then passed between aluminum plates, alternately grounded and connected to a 6000-V source, and are precipitated onto the grounded plates.

The original design of the electronic air cleaner utilizes a water-soluble adhesive coating on the plates, which holds the dirt deposits until the plates require cleaning. The filter is then deenergized, and the dirt and adhesive film are washed off the plates. Fresh adhesive is applied before the power is turned on again. Other versions of electronic air cleaners are designed so that the plates serve as agglomerators; the agglomerates of smaller particles are allowed to slough off the plates and to be trapped by viscous impingement or dry-type filters downstream of the electronic unit.

Designs of most electronic air cleaners are based on 500 ft/min (2.5 m/s) face velocity, with pressure losses at 0.20 in. (5 mm) w.g. for the washable type and up to 1 in. (25 mm) w.g. for the agglomerator type using dry-type after-filters. Power consumption is low, despite the 12,000-volt ionizer potential, because current flow is measured in milliamperes. *See* ELECTRIC CHARGE.

Inertial separators. These are devices for utilizing the principle of momentum to separate larger particulate, primarily that above 10 μm in diameter, from a moving airstream. The larger particles tend to keep going in the same direction in which they entered the device while the light-density air changes direction. Commonly offered in manifolded multiples of V-shaped pockets with louvered sides or in small-diameter cyclones, the inertial separators are used most frequently in hot, arid, desertlike regions where winds generate significant airborne dust and sand particles. A secondary bleed-air fan is used to discharge the separated particulate to the outside of the building.

Testing and rating. ASHRAE Test Standard 52 is becoming universally accepted as the optimum procedure for the testing and rating of all types of air filters and for comparing the performance of the products of competing manufacturers. An arrestance test, which measures the ability of a filter to remove the weight component of airborne particulate, is used for measuring the effectiveness of inertial separators, all viscous impingement filters, and some of the less effective dry filters. An efficiency test (dust spot efficiency using atmospheric air) provides ratings for dry filters and for electronic air cleaners. Enforcement of the certification procedures, under which manufacturers may claim certified performance, is the responsibility of the Air Conditioning and Refrigeration Institute (ARI) under their Standard 680.

The ASHRAE arrestance test measures the ability of a filter to remove a specified dust sample from the airstream which passes through the test filter. Constituents of the dust sample are carefully controlled to ensure reproducibility of test results. In the dust spot efficiency test, samples of unfiltered atmospheric air and filtered air (downstream of the test filter) are drawn off through chemical filter paper. A photoelectric cell scans the blackness of the dust spots that appear on the two pieces of filter paper; the times required to achieve equal discoloration of both pieces of filter paper, one in the unfiltered air and the other in the filtered air, are translated into a mathematical ratio, which becomes the efficiency of the test filter. *See* GALVANOMETER.

The development of ultrahigh-efficiency (HEPA) filters for special purposes (such as protection from radioactive and bacterial contaminants) led to the development of the dioctylphthalate (DOP) test. Of all the methods used to evaluate air filters, the DOP test procedure is the most sophisticated; it is used only for testing filters to be employed in critical air-cleaning applications. In a homogeneous smoke of controlled particle size (0.3-μm particles are used in this test), the DOP vapor passes through the filter. A light-scattering technique counts the number of particles entering the test filter, and a similar device counts those that emerge from the filter. Certain applications call for filters of such high efficiency that only 0.001% of incident 0.3-μm particles is permitted to emerge.

Morton A. Bell

Air heater

A component of a steam-generating unit that absorbs heat from the products of combustion after they have passed through the steam-generating and superheating sections. Heat recovered from the gas is recycled to the furnace by the combustion air and is absorbed in the steam-generating unit, with a resultant gain in overall thermal efficiency. Use of preheated combustion air also accelerates ignition and promotes rapid burning of the fuel.

Air heaters frequently are used in conjunction with economizers, because the temperature of the inlet air is less than that of the feedwater to the economizer, and in this way it is possible to reduce further the temperature of flue gas before it is discharged to the stack. *See* BOILER ECONOMIZER.

Air heaters are usually classed as recuperative or regenerative types. Both types depend upon convection transfer of heat from the gas stream to a metal or other solid surface and upon convection transfer of heat from this surface to the air. In the recuperative type, exemplified by tubular- or plate-type heaters, the metal parts are stationary and form a separating boundary between the heating and cooling fluids, and heat passes by conduction through the metal wall (**Fig. 1**). In rotary regenerative air heaters

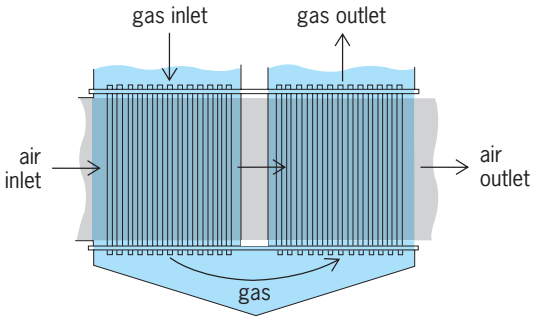

Fig. 1. Tubular air heater, two-gas single-air pass.

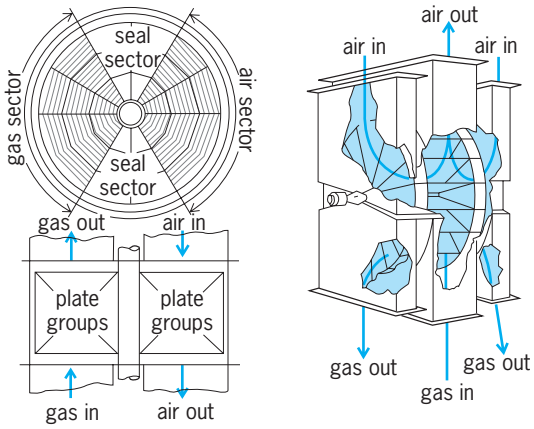

Fig. 2. Two types of rotary regenerative air heaters.

(**Fig. 2**) heat-transferring members are moved alternately through the gas and air streams, thus undergoing repetitive heating and cooling cycles; heat is transferred to or from the thermal storage capacity of the members. Other forms of regenerative-type air heaters, which seldom are used with steam-generating units, have stationary elements, and the alternate flow of gas and air is controlled by dampers, as in the refractory stoves of blast furnaces; or they may employ, as in the pebble heaters used in the petroleum industry for high-temperature heat exchange, a flow of solid particles which are alternately heated and cooled. *See* CONVECTION (HEAT).

In convection heat-transfer equipment, higher heat-transfer rates and better utilization of the heat-absorbing surface are obtained with a counterflow of gases through small flow channels. The rotary regenerative air heater readily lends itself to the application of these two principles and offers high performance in small space. However, leakage of air into the gas stream necessitates frequent maintenance of seals between the moving and stationary members, and fly ash often is transported into the combustion air system. These problems are not experienced with recuperative air heaters of the tubular type. *See* STEAM-GENERATING UNIT. George W. Kessler

Bibliography. American Society of Heating, *Refrigerating and Air Conditioning Engineers, 2004 ASHRAE Handbook: Heating, Ventilating, and Air-Conditioning: Systems and Equipment*, 2004; F. C. McQuiston, J. D. Parker, and J. D. Spitler, *Heating, Ventilating and Air Conditioning Analysis and Design*, 6th ed., 2004.

Air mass

In meteorology, an extensive body of the atmosphere which is relatively homogeneous horizontally. An air mass may be followed on the weather map as an entity in its day-to-day movement in the general circulation of the atmosphere. The expressions air mass analysis and frontal analysis are applied to the analysis of weather maps in terms of the prevailing air masses and of the zones of transition and interaction (fronts) which separate them.

The relative horizontal homogeneity of an air mass stands in contrast to sharper horizontal changes in a frontal zone. The horizontal extent of important air masses is reckoned in millions of square miles. In the vertical dimension an air mass extends at most to the top of the troposphere, and frequently is restricted to the lower half or less of the troposphere. *See* FRONT; METEOROLOGY; WEATHER MAP.

Development of concept. Practical application of the concept to the air mass and frontal analysis of daily weather maps for prognostic purposes was a product of World War I. A contribution of the Norwegian school of meteorology headed by V. Bjerknes, this development originated in the substitution of close scrutiny of weather map data from a dense local network of observing stations for the usual far-flung international network. The advantage of air-mass analysis for practical forecasting became so evident that during the three decades following World War I the technique was applied in more or less modified form by nearly every progressive weather service in the world. However, the rapid increase of observational weather data from higher levels of the atmosphere during and since World War II has resulted in a progressive tendency to drop the careful application of air-mass analysis techniques in favor of those involving the kinematical or dynamic analysis of upper-level air flow, usually involving numerical methods supported by large computers.

Origin. The occurrence of air masses as they appear on the daily weather maps depends upon the existence of air-mass source regions, areas of the Earth's surface which are sufficiently uniform so that the overlying atmosphere acquires similar characteristics throughout the region. *See* ATMOSPHERIC GENERAL CIRCULATION.

Weather significance. The thermodynamic properties of air mass determine not only the general character of the weather in the extensive area that it covers, but also to some extent the severity of the weather activity in the frontal zone of interaction between air masses. Those properties which determine the primary weather characteristics of an air mass are defined by the vertical distribution of water vapor and heat (temperature). On the vertical distribution of water vapor depend the presence or absence of condensation forms and, if present, the elevation and thickness of fog or cloud layers. On the vertical distribution of temperature depend the relative warmth or coldness of the air mass and, more importantly, the vertical gradient of temperature, known as the lapse rate. The lapse rate determines the stability or instability of the air mass for thermal convection and consequently, the stratiform or convective cellular structure of the cloud forms and precipitation. The most unstable moist air mass, in which the vertical lapse rate may approach 1°F/170 ft (1°C/100 m), is characterized by severe turbulence and heavy showers or thundershowers. In the most stable air mass there is observed an actual increase (inversion) of

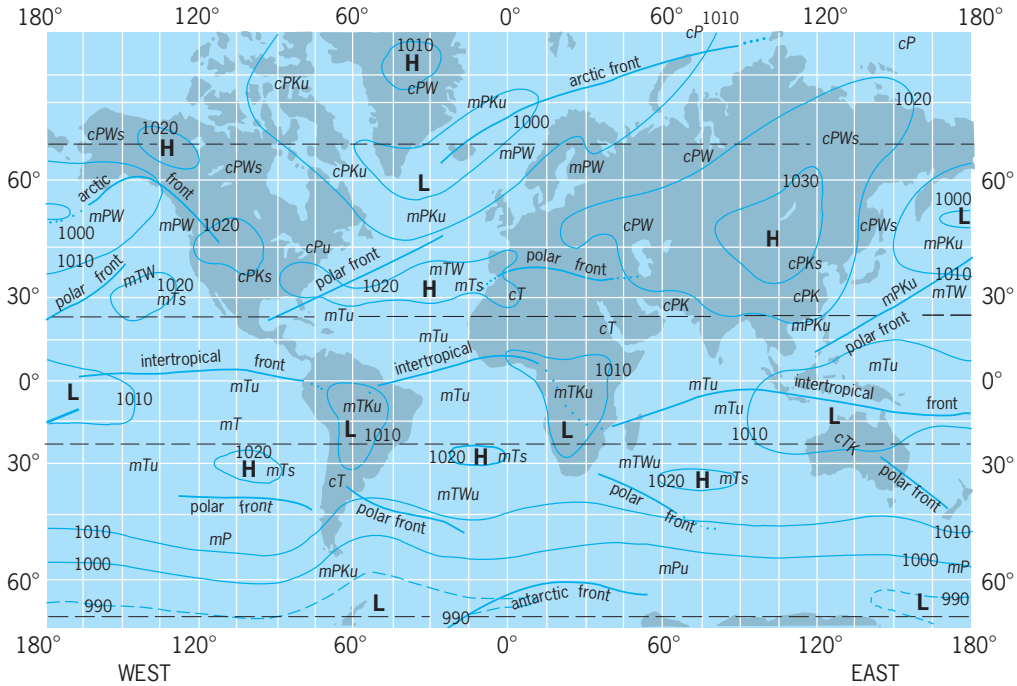

Fig. 1. Air-mass source regions, January. High- and low-atmospheric-pressure centers are designated *H* and *L* within average pressure lines numbered in millibars (such as 1010); 1 millibar = 10^2 Pa. Major frontal zones are labeled along heavy lines. (*After H. C. Willett and F. Sanders, Descriptive Meteorology, 2d ed., Academic Press, 1959*)

temperature with increase of height at low elevations. With this condition there is little turbulence, and if the air is moist there is fog or low stratus cloudiness and possible drizzle, but if the air is dry there will be low dust or industrial smoke haze. *See* TEMPERATURE INVERSION.

Classification. A wide variety of systems of classification and designation of air masses was developed by different weather services around the world. Most systems of air-mass classification are based on a designation of the character of the source region and the subsequent modifying influences to which the

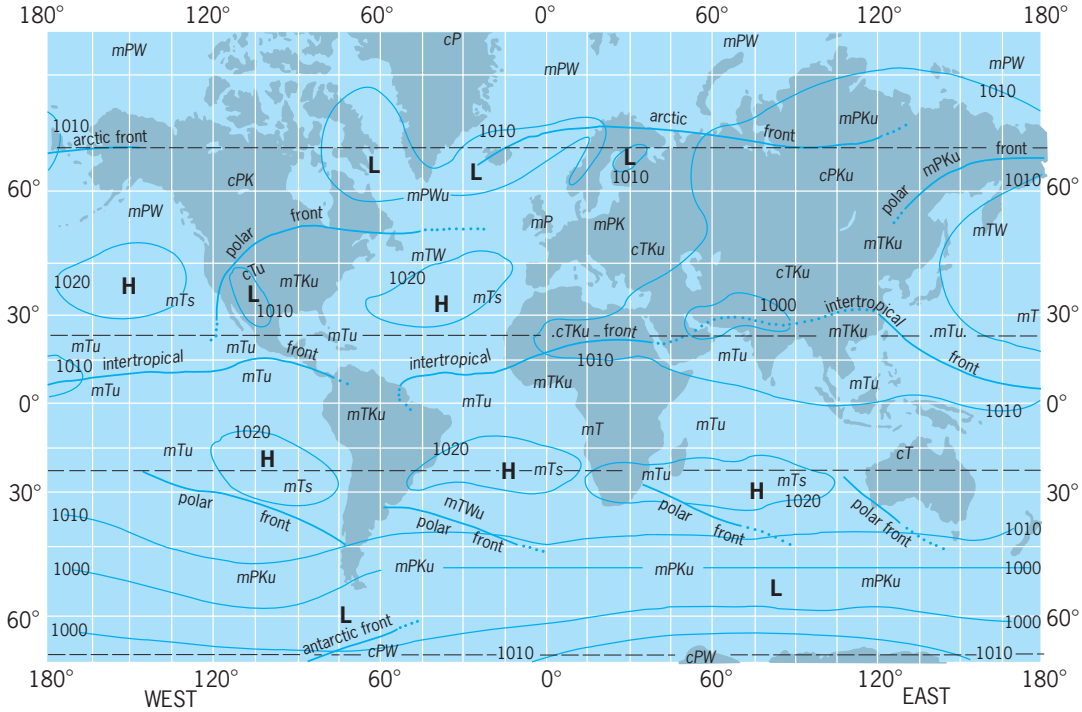

Fig. 2. Air-mass source regions, July. The symbols which are used in this figure are the same as those for Fig. 1. (*After H. C. Willett and F. Sanders, Descriptive Meteorology, 2d ed., Academic Press, 1959*)

air mass is exposed. Probably the most effective and widely applied system of classification is a modification of the original Norwegian system that is based on the following four designations.

Polar versus tropical origin. All primary air-mass source regions lie in polar (*P* in **Figs. 1** and **2**) or in tropical (*T*) latitudes. In middle latitudes there occur the modification and interaction of air masses initially of polar or tropical origin. This difference of origin establishes the air mass as cold or warm in character.

Maritime versus continental origin. To be homogeneous, an air-mass source region must be exclusively maritime or exclusively continental in character. On this difference depends the presence or absence of the moisture necessary for extensive condensation forms. However, a long trajectory over open sea transforms a continental to a maritime air mass, just as a long land trajectory, particularly across major mountain barriers, transforms a maritime to a continental air mass. On Figs. 1 and 2, *m* and *c* are used with *P* and *T* (*mP*, *cP*, *mT*, and *cT*) to indicate maritime and continental character, respectively.

Heating versus cooling by ground. This influence determines whether the air mass is vertically unstable or stable in its lower strata. In a moist air mass it makes the difference between convective cumulus clouds with good visibility on the one hand and fog or low stratus clouds on the other. Symbols *W* (warm) and *K* (cold) are used on maps—thus, *mPK* or *mPW*.

Convergence versus divergence. Horizontal convergence at low levels is associated with lifting and horizontal divergence at low levels with sinking. Which condition prevails is dependent in a complex manner upon the large-scale flow pattern of the air mass. Horizontal convergence produces vertical instability of the air mass in its upper strata (*u* on maps), and horizontal divergence produces vertical stability (*s* on maps). On this difference depends the possibility or impossibility of occurrence of heavy air-mass showers or thundershowers or of heavy frontal precipitation.

Examples of the designation of these tendencies and the intermediate conditions for maritime polar air masses are *mPWs*, *mPW*, *mPWu*, *mPs*, *mPu*, *mPKs*, *mPK*, and *MPKu*. Hurd C. Willett; Edwin Kessler

Bibliography. S. Ackerman and J. A. Knox, *Meteorology: Understanding the Atmosphere*, 2002; R. Anthes, *Meteorology*, 7th ed., 1996; L. J. Battan, *Fundamentals of Meteorology*, 2d ed., 1984; J. M. Moran and M. D. Morgan, *Meteorology: The Atmosphere and the Science of Weather*, 5th ed., 1996; R. B. Stull, *Meteorology for Scientists and Engineers*, 2d ed., 1999.

Air navigation

A discipline that combines the means to know the location and heading of an aircraft with respect to some map or coordinate system, with guidance which means steering the aircraft along some desired path toward a destination. The construction of the desired path is done with respect to the same a map or coordinate system used for location. In civil aviation, the coordinate system that is becoming a de facto standard is the World Geodetic System of 1984 (WGS-84). This is largely due to the fact that positioning in the Global Positioning System (GPS) is based on this mapping and coordinate system.

In the aircraft, the process of planning and directing its progress between selected geographic points or over a selected route is the main navigation task. The primary tasks are planning the flight path or route, guiding the aircraft safely along the desired route, conforming with the rules of flight and with special procedures such as noise abatement, maximizing fuel efficiency, departure takeoff, and landing. Landing may require precision navigation and guidance if flight instruments are used.

Modern air navigation makes use of four dimensions: latitude, longitude, altitude, and time. Time becomes a factor whenever time constraints have to be met. Examples are arrival at specific waypoints or at the destination at a precisely specified time.

The simplest form of air navigation is pilotage, in which the pilot directs the aircraft by visual reference to objects on the ground. Pilotage is the most basic skill, and despite the accuracy of modern air navigation instruments, it can be drawn upon in the case of an emergency where navigational aids experience failure. More complex methods rely upon navigational aids external to the aircraft or upon self-contained, independent systems. *See* PILOTAGE.

Dead reckoning is the process of estimating one's current position by measuring the change in position since the last accurate position fix. The simplest form of dead reckoning requires an airspeed indicator, outside-air-temperature gage, clock, compass, and chart (map). Using these tools and a few simple computations, the pilot can follow a compass heading from a point of departure for a specified time and arrive at a dead-reckoned position. Dead reckoning is periodically validated by position information from a source external to the aircraft, such as a ground sighting, radio bearing, celestial observation, or a direct position solution.

All navigation instruments that propagate the-break position based on position increments, velocity measurement, or acceleration measurement are dead-reckoning systems. Examples are the Doppler velocity systems (DVS) and inertial navigation systems (INS). *See* DEAD RECKONING.

Navigation systems that derive a position fix from range measurements or from range and angle measurements are direct position solutions. These navigation instruments measure range or angles between the aircraft and known points. Examples are distance measurement equipment (DME) and very high frequency (VHF) omnidirectional radio range (VOR), respectively. Depending on the geometry of the range and angle measurements, the position solution may not be unique. Examples of nonunique position

solutions are DME-DME; examples of unique position solutions are GPS position fixes. Whenever available, unique position solutions are regarded as the last accurate position fix.

Flight Planning

Air navigation begins with a flight plan. For a flight in visual meteorological conditions (VMC), a simple plan may start with drawing a course line on a specially designed aeronautical chart between the point of departure and the destination. From the chart and the course line, the pilot, using a protractor, determines the true course (in degrees clockwise from true north), magnetic variation, and distance to be flown, usually in nautical miles. The pilot also checks the chart for obstacles or hazards that may be on or adjacent to the planned route.

The three minimum considerations that must be accounted for in the flight plan are:

1. The pilot must properly calculate the course and fuel consumption.
2. The pilot must be able to accurately fly the aircraft along the intended route without hazard.
3. The pilot must obtain the weather conditions, particularly the winds aloft, from the weather service.

The first two considerations are under the pilots control and can be planned. The third is not but it establishes the VMC.

During flight planning, the pilot will apply corrections to the true course for magnetic variation (the angular difference between true north and magnetic north), deviation (the angular difference between the nominal magnetic heading and the actual aircraft compass heading), and computed wind drift.

A preflight briefing is customary and may be obtained from a flight service station (FSS). It will contain the expected weather conditions for the flight and pertinent NOTAMS (Notices to Airmen), and can be obtained in person, by telephone, or by a direct user access terminal (DUAT) system, requiring a personal computer and modem. Following the preflight weather briefing, the pilot has the necessary facts to prepare a simple flight plan: wind direction and velocity, which were included in the briefing; true course, which was determined from the course line on the chart; and true airspeed, which is calculated by applying the corrections for temperature and altitude (about +2% for each 1000 ft or 300 m above sea level) to the indicated airspeed. With the wind, true course, and true airspeed, the pilot can construct a wind triangle to determine the effect of the wind on speed and heading. The wind triangle can be constructed manually by using a protractor and ruler, or it may be set up on an electronic calculator or a manually operated navigational computer. The wind triangle computation provides groundspeed (speed over the surface) and wind drift angle, in degrees left or right of the course (**Fig. 1**). During flight, a correction counter to the drift angle (called wind correction angle or CRAB) will be applied to the aircraft heading to prevent drifting off course. With the planned groundspeed and the distance to be flown

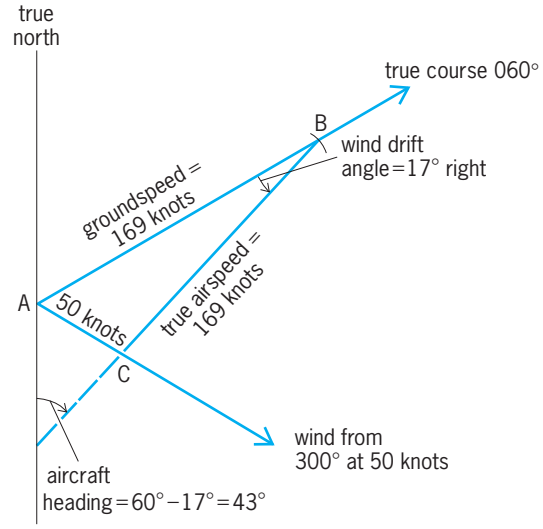

Fig. 1. Typical wind triangle, drawn to scale. 1 knot = 1 nmi/h = 0.514 m/s.

in miles, the pilot can compute the flight time and the fuel required. Extra fuel is added to assure an adequate reserve upon landing.

The above description applies to the formation of a simple, fair-weather flight plan. The process assumes greater complexity for a higher-performance aircraft, for a longer flight, or for a flight in adverse weather.

Weather forecasts for the departure airport, for the route of flight, and for the destination airport must be considered in formulating the plan. Severe storms or turbulence can sometimes be avoided by taking a different route or flying at a different altitude. More favorable winds may also be found by selecting a different route or altitude. *See* AERONAUTICAL METEOROLOGY.

Air navigation is three-dimensional, and selection of a flight altitude is an important part of the planning process. Light, piston-engine aircraft are usually operated at altitudes of 10,000 ft (3000 m) or less, while jet aircraft commonly operate at much higher altitudes, where they may take advantage of the jet's increased fuel economy. During flight planning for large turboprop and jet aircraft, the time en route and the fuel used at various altitudes or over different routes are often compared to select the optimum altitude and route.

Aircraft flying in the lower altitudes must select altitudes that will safely clear the terrain, including artificial obstacles such as television towers and tall buildings. Instrument flights are required to fly at least 1000 ft (300 m) above the highest obstacle in nonmountainous areas or 2000 ft (600 m) above the highest obstacle in designated mountainous areas.

The selection of a flight altitude is further constrained by requirements for separation from other aircraft. Generally, flights proceeding in a northerly or easterly direction fly at odd altitudes (3000, 5000, 7000 ft, and so forth, equivalent to 900, 1500, 2100 m), while flights proceeding in a southerly or westerly direction fly at even altitudes (2000, 4000, 6000 ft, and so forth, equivalent to 600, 1200,

1800 m). Flights operating under visual flight rules (VFR) add 500 ft (150 m) to the altitudes above. Above 29,000 ft (8800 m), somewhat different rules apply because of increased vertical separation required between aircraft.

Because of the cost of fuel and the large quantities used, fuel conservation is important. Flying at higher altitudes and reducing a jet's speed slightly will usually produce significant fuel savings with little increase in time; consequently, many operators regularly fly at the optimal, most fuel-efficient speed. *See* AIRCRAFT FUEL.

When the destination weather is expected to be poor, an alternative airport must be designated and additional fuel carried to fly there should a diversion be necessary.

Navigation Charts

All navigation requires special maps or charts such as topographic maps intended for pilotage or radio navigation charts intended to be used with ground radio aids. Aeronautical maps such as sectional charts show terrain contours and the elevation above mean sea level of significant terrain and other obstacles, such as mountain peaks, buildings, and television towers.

With a few exceptions, most charts for aeronautics use the Lambert conformal projection. Significant advantages of the Lambert projection are that a straight line drawn between two points is a great circle, which is the shortest distance between those points, and distances are uniformly scaled across the entire chart.

Mercator projection charts draw lines of longitude and latitude in straight lines. A straight line on a Mercator chart is called a rhumb line. It defines a constant heading path between departure and destination points and is not in general a great circle path. Distances are not uniformly scaled across the chart from north to south. *See* MAP PROJECTIONS.

Some special tasks, such as polar navigation, utilize a grid chart whose meridians and parallels are equally spaced, resembling lines on graph paper. *See* POLAR NAVIGATION.

Operational Air Navigation

The operational phase of air navigation commences with the preflight cockpit check of clocks, radios, compasses, flight management systems, and other flight and navigation equipment. Some equipment such as inertial navigation requires certain initialization data to be inserted before the units can operate properly. An air-traffic control clearance for the planned flight is obtained from the controlling authorities during preflight.

The flight management systems (FMS) can automate a very significant portion of the cockpit workload, allowing the flight crew to manage the aircraft systems instead of interacting with the various systems. However, the FMS must be set up properly, including entry of the flight path data bases, vertical navigation performance parameters and policies, weather information (if not done so automatically), and any other pertinent data. The FMS can ensure optimal fuel burn and arrival at the destination on time while providing horizontal and vertical guidance to the flight director or autopilot, achieving true four-dimensional navigation. *See* AUTOPILOT.

Departure. During taxi to the takeoff runway, the pilot follows a detailed airport chart and instructions received by radio from the airport ground controller. After a takeoff in visual flight conditions, the pilot turns to a heading to intercept the planned course and climbs to the proper altitude for the direction of flight.

The departing pilot may be guided by a standard instrument departure (SID) chart for the airport or may be radar-guided by instructions from the radar departure controller. SIDs are used to automate and completely specify the departure route. Aircraft that can follow SID procedures have the advantage of quick and efficient dispatch. However, SIDs are complex in general and are typically executed by an FMS.

En route. The pilot's primary navigational task is to assure that the aircraft stays on course. If navigation is by pilotage, the pilot frequently checks the aircraft's position by reference to topographic features on the aeronautical chart being used, and the aircraft's heading is corrected as necessary to follow the course. Ground speed is determined from the time required to fly the distance from the departure airport to the first and subsequent fixes. Fuel is checked to verify that the quantity used thus far agrees with the planned use of fuel.

The pilot estimates the time that the aircraft is expected to be over the next fix, based on the speed thus far made good and the speed anticipated over the next segment. This process will continue throughout the flight, with the pilot repeatedly assessing speed and fuel used versus the flight plan and making alterations as needed.

An FMS automates this task. It collects the positioning information from the various navigation sensors such as GPS, INS, VOR, and DME, and produces the best probable position fix. It may employ sophisticated techniques such as Kálmán filtering and hierarchical logic to produce the position solution. Simultaneously, the FMS will check flight policies, check altitude and noise restrictions, and provide guidance to the autopilot, and in addition it can monitor fuel consumption.

VOR navigation. The most common form of en route navigation over landmasses in the Western world uses a network of VOR stations. VOR transmitters are spaced at strategic intervals, averaging about 100 mi (160 km) between stations within the United States but sometimes much farther apart. Designated routes between VOR stations are referred to as airways.

VOR transmitters emit a special signal that, when received, indicates heading toward a VOR station. The pilot can then determine the direction from the ground station to the aircraft, measured in degrees magnetic and referred to as a radial. Distance to the

station is displayed when a DME transmitter is colocated with the VOR. Such an installation is referred to as a VOR/DME. When the DME function is provided by a colocated Tacan transmitter, the installation is referred to as a VORTAC. *See* DISTANCE-MEASURING EQUIPMENT; TACAN.

An airway is defined by a series of VOR stations along an airway route. Navigation charts show the location of the VORs, their radio frequencies, and the radials (courses) between them which make up the airways. The pilot refers to a chart for the frequency of the VOR and the radial to be followed, which are then set on the appropriate cockpit equipment. After following the outbound radial to the midpoint or designated changeover point between stations, the next station on the airway is tuned in and the inbound radial to that station is followed. When the aircraft is equipped with an FMS, the FMS performs the tuning function automatically, leaving the pilot to monitor the flight and make necessary adjustments as succeeding stations are selected along the route.

It is not necessary to follow charted airways in order to use VOR stations for navigation. Flights are often cleared by air-traffic control to proceed "direct" from one VOR to another in order to detour around adverse weather or to change routes. *See* RHO-THETA SYSTEM; VOR (VHF OMNIDIRECTIONAL RANGE).

Arrival. Arrival routes and procedures at major airports are often so complex that it is necessary to detail them on standard terminal arrival route (STAR) charts. Details include navigational aids, routes, required altitudes, and communications frequencies.

A high-altitude jet starts descending miles from the landing airport. A jet flying at 39,000 ft (11,900 m) might start to descend more than 100 nautical miles (185 km), or 15–20 min, from the airport. Unpressurized aircraft carrying passengers fly much lower, but their descent must be more gradual to minimize passenger discomfort in adjusting to the change in atmospheric pressure.

Regardless of the aircraft type, the pilot estimates as accurately as possible the time and point to start descending. This is particularly important in jet-type aircraft since an early descent causes a significant increase in fuel consumption while a late descent introduces other problems.

Every instrument approach for landing has its own chart, whether it uses an instrument landing system (ILS), the Space-Based Augmentation System (SBAS), VOR, nondirectional beacon (NDB), radar, or some other aid. The approach chart contains such things as the radio frequency and identification of the aids to be used, altitudes to be observed at various points during the approach, heading of the final approach course, minimum weather conditions required for the approach, missed approach instructions, and airport altitude.

The preferred aid for instrument landing use is ILS. It consists of two separate radio beams: one called the localizer is aligned very precisely with the runway centerline, and the other called the glideslope is projected upward at a slight angle from the landing touchdown point on the runway. By following the two beams very closely on the cockpit instruments, the pilot can descend safely for landing, regardless of weather.

ILS accuracy permits aircraft to safely descend to lower weather minimums than are possible with other aids such as VOR or NDB. Descent for landing is commonly authorized to 200 ft (60 m) above ground, with 0.5-mi (0.8-km) visibility (Category I or Cat-I). Most ILS approaches in small aircraft are controlled manually, but many large aircraft are equipped with automatic pilots that can be coupled to follow the ILS. Aircraft with more sophisticated approach and landing capability may be permitted to land on specially designated runways when weather is less than Category I. Category III is the most stringent weather condition, and properly equipped aircraft may be authorized to make autopilot-coupled ILS approaches and automatic landings on those runways when ceiling and visibility are zero or near zero. *See* AUTOMATIC LANDING SYSTEM; INSTRUMENT LANDING SYSTEM (ILS).

SBAS approaches have now been introduced and are capable of providing an ILS CAT-I equivalent service. Approach procedures based on SBAS are called localizer with precision vertical (LPV) approaches. Such approaches require no airport ground aides, only an SBAS/GPS avionics receiver. The LPV approach charts are similar to ILS charts, the major difference being that an LPV approach requires an SBAS area navigation (RNAV) capability, typically with addition of an FMS. It is expected that SBAS will replace ILS CAT-I approaches.

After landing, the pilot taxis the aircraft to the parking area following the airport chart and instructions received by radio from the ground controller. Some major airports are equipped with special radar for following taxiing aircraft during very low visibility. *See* SURVEILLANCE RADAR.

Altimetry

Aircraft altitude may be obtained from a barometric altimeter, a radio altimeter, or GPS. Although all of these instruments produce an altitude measurement, they are not the same kind of altitude measurement and they cannot be used interchangeably. The reason is that their reference systems are different. In general, baro-altimeters produce mean sea level altitude (for corrected low altitudes) or an altitude with respect to a standard atmospheric altitude; GPS receivers produce altitude with respect to the WGS-84 coordinate system; and radar altitude is simply the distance between the aircraft and the closest terrain beneath it (including mountains).

Mean sea level (MSL) altitude. The MSL is an average sea level surface that defines the zero altitude and is used to define the altitude reference in all navigation and mapping charts. The MSL zero altitude is determined by measuring the Earth's gravity with satellites. The zero altitude surface cannot be directly observed due to tidal actions, changes in the Earth's

weather, melting of glaciers, and other effects. *See* GEODESY.

Barometric altitude. The primary altitude instrument in the aircraft is the baro-altimeter. The baro-altimeter measures air pressure and coverts it to a displayed altitude in meters or feet. *See* ALTIMETER.

Barometric altimeters employ three different operating modes: by setting altitude above mean sea level, as reported by a nearby weather station (QNH); by using a standard setting (QNE) of 29.92 in. of mercury (1013.2 millibars) as the sea level or zero MSL altitude; and by setting the altimeter to read zero when on the airport surface (QFE), thereby showing height above the airfield when in flight. A small number of airlines use QFE when approaching the destination airport, but most use QNH.

At and above a transition altitude, all aircraft use the standard QNE altimeter setting, and altitudes are termed flight levels (FL). For example, 29,000 ft becomes FL290. The transition altitude in United States airspace is 18,000 ft (FL180) or 5500 m. Elsewhere, transition altitudes are determined by the controlling governmental agency. The QNE barometric altitude has the same definition (and conversion from air pressure to altitude) for all aircraft.

Radar altitude. All radar altimeters measure the smallest range between the terrain immediately below the aircraft and the aircraft. The radar altimeter is intended for instrument approaches that require accurate height above the landing runway threshold. Most radar altimeters operate only below 2500 ft (750 m) above ground.

GPS altitude. GPS receivers provide accurate WGS-84 altitude information. GPS altitude may be converted to MSL altitude using an MSL model such as the Earth Gravitational Model of 1996 (EGM96), developed by the National Imagery and Mapping Agency, the NASA Goddard Space Flight Center, and Ohio State University.

GPS altitude is unaffected by the errors found in barometric altimetry. It has an accuracy of approximately 35 ft (10 m). The accuracy of standard GPS positioning is much improved when augmented by the Space-Based Augmentation System.

SBAS altitude. SBAS receivers provide improved GPS altitude that can be used for performing LPV landing approaches equivalent to ILS CAT-I approaches with 200-ft (60-m) decision height. LPV approach plates (single-page aeronautical charts that provide the pilot with the detailed information needed to perform a precision approach) reference MSL; however, the final approach segment (FAS) is specified in WGS-84 coordinates.

Reduced Vertical Separation Minima (RVSM)

East-west traffic has resulted in a heavily congested airspace. The vertical separation required of aircraft between FL290 and FL410 has been 2000 ft (600 m). Special RVSM rules were brought in to allow the reduction of this vertical separation to 1000 feet (300 m). The result is an additional six flight levels between FL290 and FL410, allowing more aircraft to fly efficient time-fuel profiles. RVSM-qualified aircraft must have higher-accuracy baro-altimeters (operating in QNE mode) and other surveillance equipment such as TCAS-II, and must meet altitude-keeping performance standards. *See* AIRCRAFT COLLISION AVOIDANCE SYSTEM.

There are several airspaces to which RVSM applies. In the United States, such airspace is referred to as Domestic RVSM or DRVSM. When flying in an RVSM airspace, the pilot must abide by special rules governing that particular airspace.

Additional Navigation Aids and Systems

The **table** lists the electronic systems discussed, gives their principal frequency bands, and indicates which systems are internationally standardized.

Flight management system (FMS). An FMS integrates aircraft communication, navigation, flight guidance and control, and aircraft performance management functions. The navigation function may be as simple as the area navigation described below, but today it is more likely to utilize inputs from several sensors: VOR/DME, GPS/SBAS, Loran-C, and INS. Typically, the FMS will select the navigation sensors that produce the best possible position, favoring SBAS and then proceeding down to less accurate sources of navigation. Navigation sources may be integrated through a Kálmán filter algorithm. The Kálmán filter models and estimates the behavior of dynamical systems such as an INS, and it maintains the dynamical system error (the variance). In air navigation, Kálmán filters are typically used to integrate the inertial reference systems (IRS) with GPS/SBAS, taking advantage of the self-contained short-term navigation performance of IRS and combining it with the long-term accuracy and integrity provided by SBAS. *See* ESTIMATION THEORY.

Extensive aircraft performance and regularly updated regional or worldwide navigation databases are stored in and utilized by the FMS. Commonly used routes can be stored in the FMS and called up when needed. In addition to lateral navigation (LNAV), a flight management computer can manage vertical

Electronic navigation systems

System	Frequency band
VOR*	108-118 MHz
DME*, Tacan	960-1215 MHz
ILS localizer*	108-112 MHz
ILS glideslope*	329-335 MHz
Doppler radar	10-20 GHz
Loran-C	90-110 kHz
MLS	5.0-5.25 GHz
GPS	1227, 1575 MHz
SBAS	1575 MHz
ADF/NDB*	190-1600 kHz
Weather radar	5, 9 GHz
Radar altimeter	4.2 GHz
Doppler velocity radar	13.325 GHz

*Internationally standardized systems.

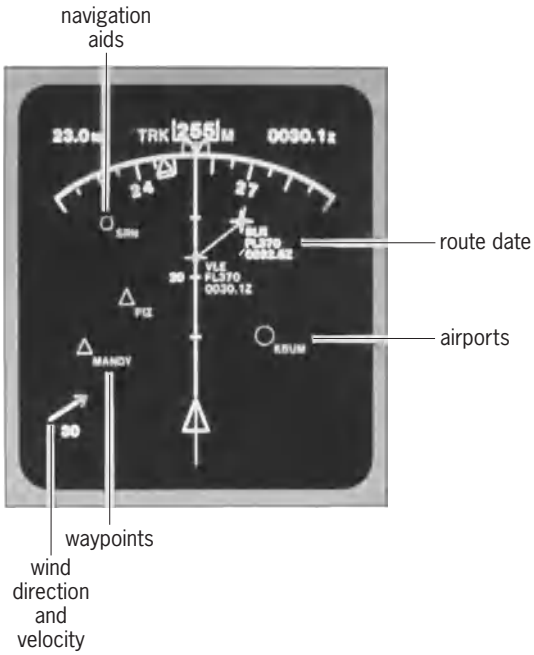

Fig. 2. Electronic map display developed by a flight management computer on the pilot's instrument panel of a Boeing 767 airplane. (*United Airlines*)

navigation (VNAV), that is, climbs and descents on the proper profiles.

The FMS can also control aircraft speed to achieve required time of arrival at specific waypoints along the intended path of flight. The time management function of the FMS can be used in conjunction with air-traffic control management to achieve precise arrival control of the aircraft for improved airspace and runway utilization. The combination of lateral and vertical navigation with management of speed and time is referred to as 4-D RNAV. Besides coupling with the autopilot and autothrottle, in some installations the flight management computer is coupled with an electronic map display which can be used for primary navigation guidance (**Fig. 2**). *See* AIRCRAFT INSTRUMENTATION.

Current FMS also computes a cost performance index (CPI). The CPI is an indicator of the aircraft flight performance with respect to an optimal path. The optimal path is developed by an operator, or airline, according to its policies. These policies may require a strict time of arrival, a minimum fuel burn, or any combination of policies and requirements. Variable factors such as winds, air-traffic control directives, and aircraft weight have an impact on the actual CPI.

Global Navigation Satellite System (GNSS). The Global Positioning System, sometimes called NAVSTAR, is a part of the Global Navigation Satellite System, and is operated by the U.S. Department of Defense. GPS is a space-based positioning, radionavigation, and time-distribution system designed for worldwide military use. (A comparable system called GLONASS developed by the former Soviet Union is still in operation.) GPS has been made available for worldwide civil and military use.

The Wide Area Augmentation System (WAAS) was designed in the United States by the Federal Aviation Administration to provide GPS-independent real-time GPS integrity, differential corrections, and ranging signals. The WAAS complements GPS and extends GPS capability as a primary-means-of-navigation system and as a landing approach system.

The WAAS has been adopted internationally by the International Civil Aviation Organization (ICAO) and has become the Space Based Augmentation System (SBAS). Several nations are developing SBAS systems: in Europe as the European Geostationary Navigation Overlay System (EGNOS), in India as GAGAN, in China as BEIDOU, and in Japan as MSAS. All of these systems are interoperable and conform to the ICAO SBAS standards. SBAS provides GPS integrity, real-time differential corrections, and ranging signals similar to GPS. SBAS will eventually cover the Earth except for latitudes exceeding 75° due to the fact that SBAS satellites are parked in geosynchronous orbits. All continents with the exception of Antarctica will be covered by SBAS.

Future GNSS development includes the European Galileo satellite navigation system, and GPS and SBAS modernization. Galileo is similar to GPS in function; however, its signal transmission formats differ from GPS. Equipment that can receive GPS, Galileo, and GLONASS signals is being developed. Future GPS and SBAS will provide a second civil frequency called L5, centered at 1176 MHz, as well as adding a civil signal to the L2 frequency at 1227 MHz. Furthermore, the civil C/A codes on the L1 band at 1575 MHz will be augmented by new civil codes with less cross-correlation that support CAT-IIIb functionality. The resulting benefits are projected to support CAT-II, and possibly CAT-III, with GPS, SBAS, and Galileo.

Airborne GPS/SBAS navigation receivers continuously process GPS and SBAS satellite data to update the real-time position, latitude, longitude, and altitude in WGS-84 coordinates. The position may be displayed as a digital readout of present position, sent as an input to the course deviation indicator on the pilot's instrument panel, or sent as a sensor input to a navigation computer such as the flight management system or inertial navigation system.

All certified, airborne GPS/SBAS receivers are distinguished by providing integrity with the position solution. Integrity is an estimate of the position fix error, assuming that a satellite signal is erroneous or misleading. The integrity algorithms are referred to as either the Receiver Autonomous Integrity Monitor (RAIM) or the Fault Detection and Exclusion (FDE) algorithms. In one case, the SBAS network determines the integrity of each GPS satellite in real time, and relays it to the SBAS receivers and to the avionics displays and enunciators within 6 seconds of the total elapsed time since detection.

Terrestrial enhancements of GNSS. Two terrestrial enhancements of GNSS have been proposed. The first is the Local Area Augmentation System (LAAS) that has been internationally adopted by ICAO as the

Ground Based Augmentation System (GBAS). It is based on the same principles and techniques used in differential GPS (DGPS) satellite surveying to determine the satellite positioning error. The identified errors and corrections are valid only within the local area.

GBAS requires precisely located ground reference stations which receive satellite signals, compute a position based upon that data, and compare it with the reference stations' known position to compute corrections to the satellite signals. The resultant corrections are transmitted to all GBAS receivers via a standard VHF data broadcast (VDB) link. The defined operational radius of the VDB link is 23 nautical miles (43 km).

GBAS is intended to provide the required accuracy for precision instrument approaches, including Category III instrument landings. Due to the effects of spatial decorrelation, cross-correlation effects of the GPS C/A positioning code, and the lack of an additional radio-frequency channel, achieving a certified CAT-II/III capability has been elusive. It is believed that with a dual-frequency GPS constellation, and perhaps with Galileo, GBAS will achieve certified CAT-II/III operational capability. *See* SATELLITE NAVIGATION SYSTEMS.

The second proposed terrestrial enhancement is the Ground Regional Augmentation System (GRAS), developed by Australia. It intends to provide SBAS functionality, but by using specially modified GBAS ground stations. The VDB link remains with a 23-nautical-mile radius.

Area navigation (RNAV). An area navigation system provides the pilot with navigational guidance along random routes which are not restricted to airways or along direct routes between ground-based navigation aids such as VOR or NDB.

VORTAC-based RNAV relies upon precise DME distance from two or more VORTACs or VOR/DMEs (rho/rho mode) to continuously fix aircraft position. SBAS-based RNAV uses the position fix capability provided by an SBAS avionics receiver.

To plot an RNAV course, a direct course line is drawn on a navigation chart between the departure and destination points. The pilot then locates several VOR/DMEs or VORTACs along the proposed route and draws lines from those facilities to intersect the direct course. Those intersections will serve as waypoints. This task is automated by an FMS. SBAS position is always available and requires no planning such as with VOR/DME.

Prior to departure or en route, the pilot enters the waypoints into the RNAV course-line computer (CLC), which will compute direct navigational guidance to the waypoints. The cockpit equipment displays the course to be followed, distance remaining, and time-to-go to the next waypoint. The course to be followed may be coupled to an autopilot or the course deviation indicator (CDI).

RNAV is particularly useful in locating otherwise hard-to-find airports which are not served by a navigational facility. More advanced area navigation (random route) capability is available by using Loran-C, an inertial navigation system, or SBAS.

Inertial navigation system (INS). An INS is a self-contained system. It relies on highly accurate gyroscopes and accelerometers to produce an aircraft position solution, true heading, groundspeed, track, and attitude. The INS does this by continuously integrating accelerometer and angular rate sensor data into its position and attitude solution. The INS is a dead-reckoning sensor, and its position and attitude solution is subject to drift and error growth as well as to initialization error. An INS must be initialized before it is used for navigation.

Some INS installations have provisions for automatically updating (correcting) position information with input from GPS/SBAS, Loran-C, or DME stations via a Kálmán filter. Lacking such updates or data augmentation, INS position accuracy degrades in a predictable way with operating time. The INS is not considered sufficiently accurate to be used as an approach aid; however, an INS may be required for executing missed approaches, depending on terrain.

Prior to the introduction of FMS, INS provided steering information used by the autopilot to maintain the computed great circle route of flight. An INS was commonly used on long, overwater flights and over large, undeveloped land areas, and frequently used for point-to-point direct navigation in more developed regions. In modern installations, inertial reference systems (IRSs) are used in combination with FMSs and GPS/SBAS. The IRS is a sensor that measures orientation and acceleration, whereas an INS includes the computations required to convert these measurements into a navigation solution. *See* INERTIAL GUIDANCE SYSTEM.

Loran-C. Loran-C is a low-frequency, hyperbolic navigation system which operates by measuring the difference in arrival times of synchronized radio signals from transmitters located hundreds of miles apart. Inexpensive Loran-C receiver technology, system reliability, accuracy, and coverage have greatly increased the system's availability and acceptance at all levels of civil aviation. Stations transmitting on a common frequency make up a so-called chain which is composed of a master and two or more secondary transmitters. **Figure 3** shows a chain of stations which serves the southeastern United States and is typical of Loran-C installations.

Chain coverage is determined by the transmitted power from each station, their geometry including distance between stations, and their orientation. Within the nominal coverage range, 1000 nautical miles (1850 km), propagation is affected by the physical condition of the Earth's surface and the atmosphere.

In the United States, Loran-C is authorized by the Federal Aviation Administration (FAA) for use as a supplemental area navigation system over much of the 48 conterminous states. Because of the popularity of Loran-C as a highly accurate (to within 0.25 nautical mile or 0.5 km) navigation system,

the FAA has designed the National Aerospace System (NAS) to include Loran-C and nonprecision approaches at selected airports with adequate Loran-C coverage. Despite the extensive availability of Loran-C, however, it is not feasible to provide worldwide coverage because of range limits on the signals and the requirement that transmitters be land-based. Loran-C is not typically found on commercial air transport aircraft but on general aviation aircraft. *See* LORAN.

Microwave landing system (MLS). Like the ILS, MLS transmits two signals: one for runway alignment (localizer) and one for glidepath guidance. Unlike ILS, MLS allows multiple paths, including curved approaches to the runway, and variable glideslope angle selection. Those qualities make it particularly suited for airports located in mountain valleys or in noise-sensitive areas. It allows reduced intervals between aircraft to increase runway acceptance rates, and facilitates short-field operations for short and vertical takeoff and landing (STOL and VTOL) aircraft. MLS has not been accepted or deployed for civil aviation; its main use has been in military aircraft. *See* MICROWAVE LANDING SYSTEM (MLS); SHORT TAKEOFF AND LANDING (STOL); VERTICAL TAKEOFF AND LANDING (VTOL).

Nondirectional beacon (NDB). Low-frequency, nondirectional beacons are used in many parts of the world and are primarily associated with instrument approaches to airport runways. In less developed regions, they are also used as en route radio aids, which have a longer range and are less expensive to install and maintain than VORs but are less accurate. The NDB relies upon the directional properties of the aircraft's automatic direction finder (ADF) loop antenna, but bearings taken on an NDB are relatively unreliable when the station is more than a short distance away. This is especially true in the presence of heavy buildup of static electricity, which is common in the vicinity of thunderstorms. *See* DIRECTION-FINDING EQUIPMENT.

Radar. Navigation of civil aircraft by radar is largely limited to instructions provided by ground-based radar controllers. A common application is the use of radar vector headings to guide departing and arriving aircraft. Using radar, air-traffic controllers monitor airways, following all aircraft operated under instrument flight rules, and call deviations to the pilot's attention. Radar controllers also simplify long-range navigation across the United States by providing radar-guided direct routes to distant airports. *See* AIR-TRAFFIC CONTROL.

Surveillance radar may be used to provide instructions for a nonprecision approach to a specific runway. Using a radar map overlay and an identified return from the aircraft being controlled, the controller issues headings and recommended altitude instructions to the pilot so as to place the aircraft in a position from which a landing can be completed by visual reference to the ground.

Airborne weather radar has limited ground mapping capability which can be used as a backup to other navigation systems, particularly over coastal areas. Approach capabilities using airborne weather radar in the ground mapping mode have been established for helicopter operations at offshore oil-drilling platforms. *See* AIRBORNE RADAR.

Fig. 3. Loran-C chain GRI 7980 in the southeastern United States, comprising master station M and slave stations W, X, Y, and Z. A position fix is shown at point A, where the lines of position 7980-Y-45000 and 7980-Z-61000 intersect.

Tacan. TACAN (Tactical Air Navigation) is a primary aid to navigation for the U.S. Department of Defense, NATO, and a small number of civil users. It is designated for en route and terminal use, including naval shipboard, and operates on radio frequencies in the ultrahigh-frequency (UHF) band. It provides azimuth and distance information with reference to the station. The instrument indications are virtually the same as those provided by VOR; distance information is the same as that provided by a VORTAC. TACAN stations for en route use are usually colocated with VORs so that military aircraft which are unable to receive VORs can navigate on the same airway structure as civil aircraft.

Doppler velocity system (DVS) radar. Doppler velocity radar navigation is a self-contained dead-reckoning navigation system. It has been largely replaced on civil airplanes by more modern systems such as INS and GPS, but is still in used on military aircraft. The DVS can produce excellent velocity measurements, but it does track moving water currents. While this may be desirable in helicopter sea rescue, it is also the cause of navigation system error. *See* DOPPLER RADAR; ELECTRONIC NAVIGATION SYSTEMS; NAVIGATION.

John Studenny

Bibliography. D. J. Clausing, *Aviator's Guide to Flight Planning*, 1989; D. J. Clausing, *Aviator's Guide to Navigation*, 3d ed., 1997; A. Helfrick, *Principles of Avionics*, 3d ed., 2004; M. Kayton and W. Fried, *Avionics Navigation Systems*, 2d ed., 1997;

C. F. Lin, *Modern Navigation, Guidance, and Control Processing*, 1991; M. S. Nolan, *Fundamentals of Air Traffic Control*, 4th ed., 2003; Radio Technical Commission for Aeronautics, *Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System Airborne Equipment*, RTCA/DO-229D, 2006; J. Sanderson, *Private Pilot Manual*, 1993; A. T. Wells, *Flight Safety*, 1992.

Air pollution

The presence in the atmosphere of natural and artificial substances that affect human health or well-being, or the well-being of any other specific organism. Air pollution also applies to situations where contaminants affect structures and artifacts or esthetic sensibilities (such as visibility or smell). Most artificial impurities are injected into the atmosphere at or near the Earth's surface. The lower atmosphere (troposphere) cleanses itself of some of these pollutants in a few hours or days as the larger particles settle to the surface and soluble gases and particles encounter precipitation or are removed through contact with surface objects. Removal of some particles (such as sulfates and nitrates) by precipitation and dry deposition results in acid deposition, which may cause serious environmental damage. Also, mixing of the pollutants into the upper atmosphere may dilute the concentrations near the Earth's surface, but can cause long-term changes in the chemistry of the upper atmosphere, including the ozone layer. *See* ATMOSPHERE; TROPOSPHERE.

Characteristics

All particulate matter and contaminant gases exist in the atmosphere in variable amounts. Typical natural sources of particles are salt particles from the oceans, volcanoes, and soil. Natural sources of gases are from active volcanoes, vegetation, and soils. Typical artificial contaminants are waste smokes and gases formed by industrial, municipal, transportation, and residential processes. Pollens, spores, and rusts are natural aerosols augmented artificially by human land-use practices. *See* ATMOSPHERIC CHEMISTRY; SMOKE.

There are a number of air pollutants that are regulated by the U.S. Environmental Protection Agency (EPA). The Clean Air Act, which was last comprehensively amended in 1990, requires the EPA to set National Ambient Air Quality Standards (NAAQS) for pollutants considered harmful to public health and the environment. The act established two types of national air quality standards. Primary standards set limits to protect public health, including the health of "sensitive" populations such as asthmatics, children, and the elderly. Secondary standards set limits to protect public welfare against decreased visibility and damage to animals, crops, vegetation, and buildings.

The EPA set National Ambient Air Quality Standards for six principal pollutants, which are called criteria pollutants (**Table 1**). Units of measure for the standards are parts per million (ppm) by volume, milligrams per cubic meter of air (mg/m^3), and micrograms per cubic meter of air ($\mu g/m^3$) at 25°C (77°F).

In addition to the criteria pollutants, certain fluorocarbon compounds (gases) significantly affect the ozone layer in the stratosphere. Also, trace gases and metals (such as dioxin, toxaphene, and mercury) that are emitted as by-products of industrial operations or refuse burning are suspected of being toxic at low concentrations. And some substances, such as benzene, are emitted by burning gasoline.

There are some gases that do not directly affect human health at ambient concentrations but pose a threat due to their capacity to absorb radiation. They

TABLE 1. National Ambient Air Quality Standards

Pollutant	Standard value	Standard type
Carbon monoxide (CO)		
8-h average	9 ppm (10 mg/m^3)	Primary
1-h average	35 ppm (40 mg/m^3)	Primary
Nitrogen dioxide (NO_2)		
Annual arithmetic mean	0.053 ppm (100 $\mu g/m^3$)	Primary and secondary
Ozone (O_3)*		
8-h average	0.08 ppm (157 $\mu g/m^3$)	Primary and secondary
Lead (Pb)		
Quarterly average	1.5 $\mu g/m^3$	Primary and secondary
Particulate <10 μm (PM_{10})		
Annual arithmetic mean	50 $\mu g/m^3$	Primary and secondary
24-h average	150 $\mu g/m^3$	Primary and secondary
Particulate <2.5 μm ($PM_{2.5}$)*		
Annual arithmetic mean	15 $\mu g/m^3$	Primary and secondary
24-h average	65 $\mu g/m^3$	Primary and secondary
Sulfur dioxide (SO_2)		
Annual arithmetic mean	0.03 ppm (80 $\mu g/m^3$)	Primary
24-h average	0.14 ppm (365 $\mu g/m^3$)	Primary
3-h average	0.50 ppm (1300 $\mu g/m^3$)	Secondary

*The 8-h average standard for ozone, and the standards for PM 2.5, have been remanded by the D.C. Circuit Count of Appeals for reconsideration by the agency.
†Parenthetical value is an approximately equivalent concentration.

TABLE 2. Global emissions for several important pollutants*

Pollutant	Anthropogenic sources	Natural sources	Annual emissions, 10^{12} oz/yr (10^{12} g/yr)	
			Anthropogenic	Natural
SO_2	Fossil fuel combustion	Volcanoes, biogenic processes	7.48 (212)	0.71 (20)
CO	Auto exhaust, general combustion	Forest fires, photochemical reactions	25 (700)	74 (2100)
NO, NO_2	Combustion	Biogenic processes in soil, lightning	2.7 (75)	6.3 (180)
CO_2	Combustion	Biological processes	776 (22,000)	10^4 (10^6)
Nonmethane hydrocarbons	Combustion	Biogenic processes in soil and vegetation	1.4 (40)	705 (20,000)

*Based on 1978 emissions figures.

are known as greenhouse gases and include carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), as well as many of the fluorocarbon compounds. Greenhouse gases absorb radiation emitted from the Earth's surface and, in increasing concentrations, are capable of increasing the atmosphere's temperature.

History. Air pollution has been a significant problem since the industrial revolution, when the intensive burning of coal and oil in centralized locations began. The problem was compounded because the population of the Earth had also been rapidly growing, and the growth in use of motor vehicles in the twentieth century added to the problem.

Air pollution regulations were enacted after the occurrence of three notorious episodes associated with light winds and reduced vertical mixing that persisted for several days. Many deaths were recorded in 1930 in the Meuse Valley in Belgium, in 1948 in Donora, Pennsylvania, and in 1952 in London. Smoke and sulfur dioxide (SO_2) concentrations measured during the episode in London were on the order of 10 times the current air quality standards, resulting in the passage of the British Clean Air Act. Through elimination of some sources and controls on others, the smoke concentration in London was cut by more than one-half in the decade following the Clean Air Act.

Types of sources. Sources may be characterized in a number of ways. A distinction may be made between natural and anthropogenic sources. Another frequent classification is in terms of stationary (power plants, incinerators, industrial operations, and space heating) and moving (motor vehicles, ships, aircraft, and rockets) sources. Another classification describes sources as point (a single stack), line (a line of sources), or area (city).

Different types of pollution are specified in various ways: gaseous, such as carbon monoxide, or particulate, such as smoke, pesticides, and aerosol sprays; inorganic, such as hydrogen fluoride, or organic, such as mercaptans; oxidizing substances, such as ozone, or reducing substances, such as oxides of sulfur and oxides of nitrogen; radioactive substances, such as iodine-131; inert substances, such as pollen or fly ash; or thermal pollution, such as the heat produced by nuclear power plants.

Air contaminants are produced in many ways and come from many sources. It is difficult to identify all the various producers. For some pollutants, such as carbon dioxide and methane, natural emissions sometimes far exceed anthropogenic emissions (**Table 2**).

Both anthropogenic and natural emissions are variable from year to year, depending on fuel usage, industrial development, climate, and volcanic activity. In some countries where pollution control regulations have been implemented, emissions have been significantly reduced. In dry regions, natural emissions of nitrogen oxides (NO_x), carbon dioxide (CO_2), and hydrocarbons can be greatly increased during a season with high rainfall and above-average vegetation growth. *See* NITROGEN OXIDES.

Most estimates of the methane budget put the anthropogenic component at about two-thirds. Ruminant production and emissions from rice paddies are regarded as anthropogenic because they result from human agricultural activities. Increases in carbon dioxide since the industrial revolution are also principally the result of human activities. These emissions have not yet equilibrated with the rest of the carbon cycle and so have had a profound effect on atmospheric levels, even though emissions from fossil fuel combustion are dwarfed by natural emissions.

An overlooked source of air pollution is the interior of homes, offices, and industrial buildings. Indoor air pollution has received increased attention since the discovery that concentrations of radon gas (naturally emitted from soils and structures) and household chemicals in indoor air can reach 5–10 times the levels in the dirtiest outside air. These high concentrations are caused by inadequate ventilation. As an example of controlling indoor air pollution, bans on cigarette smoking have proliferated since the late 1980s. *See* RADON.

The possibility of accidental release of toxic gases into the atmosphere was studied a great deal following the 1984 industrial accident in Bhopal, India, where over 2000 people died. Since then, the types of chemicals and source scenarios that might occur have been better defined.

Effects. In order to design regulations for air pollution, it is first necessary to estimate the effects on

humans and animals, vegetation, materials, and the atmospheric system. This assessment is not easy at low air pollution levels, where effects are subtle, occur gradually over long periods of time, or are the result of several confounding factors occurring simultaneously.

Humans and animals. The major concern with air pollution relates to its effects on humans. Since most people spend most of their time indoors, there has been increased interest in air pollution concentrations in homes, workplaces, and shopping areas. Much of the early information on health effects came from occupational health studies completed prior to the implementation of general air-quality standards.

Air pollution principally injures the respiratory system, and health effects can be studied through three approaches: clinical, epidemiological, and toxicological. Clinical studies use human subjects in controlled laboratory conditions, epidemiological studies assess human subjects (health records) in real-world conditions, and toxicological studies are conducted on animals or simple cellular systems. Epidemiological studies are the most closely related to actual conditions, but they are the most difficult to interpret because of the lack of control and the subsequent problems with statistical analysis. Another difficulty arises because of differences in response among different people. *See* EPIDEMIOLOGY.

The response of humans and animals to air pollution also depends on the exposure time (**Fig. 1**). It is well known that survival at much higher concentrations is possible if exposure time is only a few seconds. These short averaging times are of interest for accidental release of toxic chemicals. Conversely, low levels of air pollution can be harmful if there is continual exposure to them over a period of a year or more. *See* BIOMETEOROLOGY.

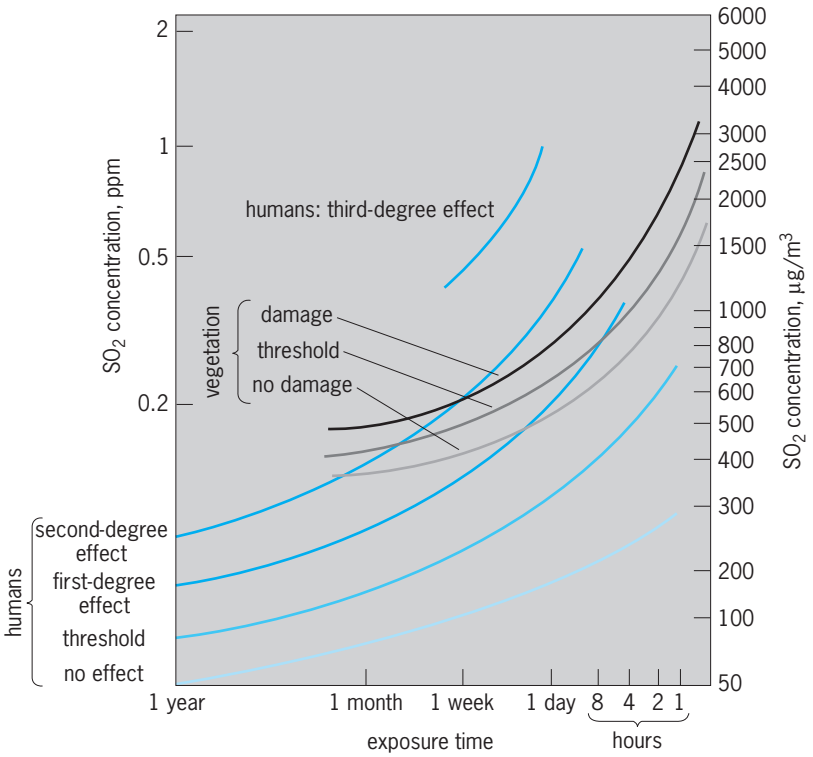

Fig. 1. Effects of sulfur dioxide on humans and vegetation. (*After L. J. Brasser et al., Sulphur Dioxide: To What Level Is It Acceptable?, Res. Inst. Pub. Health Eng. [Delft, Netherlands] Rep. G300, 1967*)

Vegetation. Damage to vegetation by air pollution is of many kinds. Sulfur dioxide may damage field crops such as alfalfa and trees such as pines, especially during the growing season (Fig. 1). Both hydrogen fluoride (HF) and nitrogen dioxide (NO_2) in high concentrations have been shown to be harmful to citrus trees and ornamental plants, which are of economic importance in central Florida. In addition, ozone and ethylene can cause damage to certain kinds of vegetation.

Since the early 1970s, the long-term effect of acid rain (rain with low pH caused by regional sulfate and nitrate pollution) on vegetation and fish has been recognized. Injury to conifer trees in Scandinavia and eastern North America has been documented, as well as fish kills in lakes in those regions. However, in most cases the damage can be blamed on other causes, such as climate change.

Materials. Corrosion of materials by atmospheric pollution is a major problem. Damage occurs to ferrous metals; to nonferrous metals, such as aluminum, copper, silver, nickel, and zinc; to building materials; and to paint, leather, paper, textiles, dyes, rubber, and ceramics. *See* CORROSION.

Atmospheric system. Gases emitted into the atmosphere can affect the dynamics of the atmosphere through changes in longwave and shortwave radiation processes. Particles can absorb or reflect incoming shortwave solar radiation, reducing the amount reaching the Earth's surface during the day. Greenhouse gases can absorb longwave radiation emitted by the Earth's surface and atmosphere.

Carbon dioxide, methane, fluorocarbons, nitrous oxides, ozone, and water vapor are important greenhouse gases that selectively absorb longwave radiation. This effect (the greenhouse effect) warms the temperature of the Earth's atmosphere and surface higher than would be found in the absence of an atmosphere. Because the amount of greenhouse gases in the atmosphere is rising, it is likely that the temperature of the atmosphere will gradually rise, possibly resulting in a general warming of the global climate over time. The carbon dioxide increase is primarily due to fossil fuel burning. It is very difficult to model the magnitude and distribution of the possible warming accurately because of complicated atmospheric responses and unknown natural sources and sinks, but it is estimated that the global temperature may rise as much as 5°C (8°F) by 2040. Because the warming would melt sea and land ice, the polar regions are believed to be warmed the most. As ice melts, the exposed land surface will be able to absorb more solar radiation than when it was ice covered, further exacerbating warming in that region. The warming would result in significant environmental effects, such as a rise in mean sea level of a few meters due to the

melting of ice sheets and a shift in vegetation patterns. *See* CLIMATE HISTORY; GREENHOUSE EFFECT.

Researchers are also concerned with pollution of the stratosphere (10–50 km or 6–30 mi above the Earth's surface) by aircraft and by broad surface sources. The stratosphere is important, because it contains the ozone layer, which absorbs part of the Sun's shortwave radiation and keeps it from reaching the surface. If the ozone layer is significantly depleted, an increase in skin cancer in humans is expected. Each 1% loss of ozone is estimated to increase the skin cancer rate 3–6%. *See* STRATOSPHERE.

Visibility is reduced as concentrations of aerosols or particles increase. The particles do not just affect visibility by themselves but also act as condensation nuclei for cloud or haze formation. In each of the three serious air pollution episodes discussed above, smog (smoke and fog) was present with greatly reduced visibility. Before emissions controls were imposed in the United States, visibility in many urban areas was often less than 1 km (0.6 mi). Even with the reductions in emissions in urban areas, a study by the U.S. Department of the Interior found that visibility in the eastern United States has been cut more than 50% in 40 years. Visibility is also an issue in pristine areas where the air historically has been very clean, with visibilities normally greater than 100 km (60 mi). If a power plant plume impacts such an area, the plume can be visible over the entire region. Likewise, the broad plumes from urban areas can significantly reduce visibilities in pristine regions, such as the Grand Canyon National Park, many hundreds of kilometers downwind. *See* SMOG.

Meteorology. The dynamics of the atmosphere will affect the concentration of air pollutants. It is often important to understand the physical processes leading to an observed concentration at a given point, or to model the expected concentration at the given point based on future emission scenarios. In both cases, it is necessary to estimate the fraction of the total emissions at the source that arrives at the receptor location in question. The physical processes relating to air quality are described below.

Wind transport. For distance scales up to regional scales (about 1000 km or 600 mi), the key question in understanding air quality variations in the transport wind is in what direction and how fast is the wind blowing. In general, for near-surface emissions, the surface air pollution concentrations will be greatest at low wind speeds. With elevated stacks, there is a trade-off between the diluting influence of plume rise (which is greater with lower wind speeds) and the diluting influence of greater ventilation (which is greater with higher wind speeds).

To estimate a source's contribution to a receptor, it is necessary to measure or estimate the wind flow patterns at and downwind of the source region. It is best if winds are measured at several locations between the sources and the receptor, since the wind field is often highly variable over distances of 10 km (6 mi) or more. But the wind field can also be extrapolated or interpolated from limited measurements, or can be modeled by using fundamental meteorological laws.

Trajectories are sometimes used to simulate the motion of the plume over multiple hours or days. Backward trajectories are used to identify the probable sources of observed pollutants; forward trajectories estimate the likely path of released pollutants. *See* WIND.

Stability. The vertical dispersion or spread of air pollution near the ground can be very rapid on sunny afternoons (unstable conditions), or can be nearly nonexistent on calm clear nights (stable conditions). In the first case the temperature decreases with height at a rate slightly greater than 1°C/100 m (0.55°F/100 ft), and in the second case the temperature can decrease with height at a slower rate than 1°C/100 m or even may increase with height (inversion). Neutral stability occurs during high-wind, cloudy conditions, when the temperature decreases with height at a rate of 0.98°C/100 m (0.53°F/100 ft). Specific patterns of plume dispersion are associated with each condition (**Fig. 2**). The terms stable and unstable refer to whether a parcel of air displaced upward adiabatically (with no exchange of heat with its environment) will tend to return to its original level or will be accelerated upward.

If the wind is blowing, the air near the ground will be turbulent even at night, but the turbulence will drop off rapidly with height so that the atmosphere has little turbulence at heights above 100 m (330 ft). During the day, the well-mixed turbulent layer adjacent to the ground is usually capped by an abrupt inversion at a height of above 1000 m (3300 ft), marking the tops of the thermals driven by surface heating. The atmosphere above this height shows little diurnal variability and is, in general, slightly stable. The depth of the turbulent air adjacent to the ground is commonly called the mixing depth, and is important because it marks the maximum vertical extent of diffusion in the mixed layer.

Dispersion. Dispersion is defined as the spread of pollutants caused by atmospheric turbulence, or

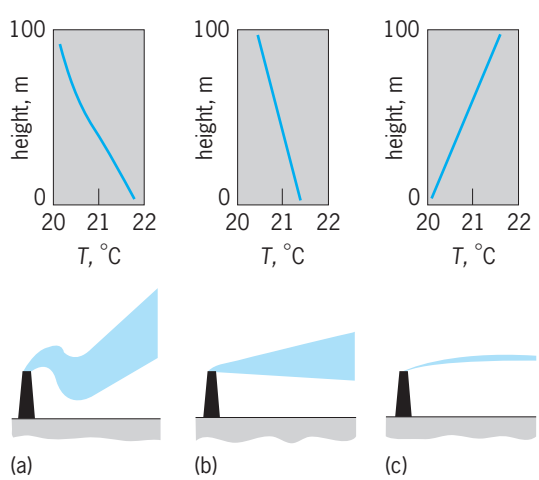

Fig. 2. Typical temperature variation with height and plume behavior for (*a*) unstable, (*b*) neutral, and (*c*) stable conditions. 1 m = 3.3 ft. °F = (°C × 1.8) + 32.

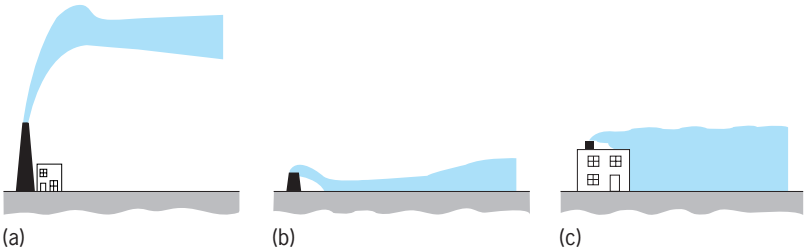

Fig. 3. Examples of the behavior of (*a*) a buoyant plume, (*b*) a dense gas plume, and (*c*) a plume being drawn into the aerodynamic wake of a building.

random fluctuations in wind velocity. Over the course of an hour during the daytime, the fluctuations in wind speed are about 20% of the average speed of the wind. These fluctuations are caused by random whorls or eddies of air that include scales of motion from about 0.1 mm to a few kilometers for vertical eddies and up to hundreds of kilometers for horizontal eddies. The rate of dispersion is faster when the size of the plume is smaller than the size of the eddies affecting the plume.

In the United States, the EPA offers a series of regulatory models for the purpose of estimating the impact of specific sources or groups of sources on downwind receptors. These models assume an averaging time of an hour or longer, and that the dispersion within the hour is random so the plume can be described as a straight line with concentration decreasing from the centerline. The rate of decrease of concentration laterally and vertically from the centerline of the plume is described by a statistically "normal" (Gaussian) distribution with standard deviations that increase with downwind distance at a rate dependent on the atmospheric stability.

Turbulence and dispersion are enhanced over rough surfaces, such as forests, hills, or urban areas, because of their larger surface friction. For a given wind speed, dispersion is about four times as great over a rough urban surface as over a smooth water surface. Dispersion is even more strongly influenced by stability, and vertical plume spread can increase by a factor of 10 or more from night to day.

There has been much research to improve dispersion estimates in situations where maximum concentrations are observed for pollutants emitted from tall stacks. For example, highest short-term (1 h average) ground-level concentrations in flat terrain around tall stacks are observed during light-wind, sunny, daytime conditions when vigorous eddies frequently bring the plume to the ground near the stack. Another serious condition occurs in mountainous terrain where plumes during stable conditions may affect the terrain directly, leading to high concentrations in a narrow belt on the mountainside.

In the case around tall stacks during light-wind, sunny, daytime conditions, it has been observed that, on the average, the plume will slowly be moved toward the ground as the areas of downward-moving air are larger than the areas of more rapidly upward-moving thermals. Depending on the height of the tall stack, plumes will affect the surface with a greater probability than is predicted by using the Gaussian plume simulation techniques.

Air pollution variability. The atmosphere is characterized by intense turbulence over a wide range of space or time scales. Just as instantaneous measurements of the wind speed will vary randomly about their long-term mean, so will air pollution measurements. This variability is greatest for instantaneous measurements and slowly decreases as averaging time increases. For a given hourly average wind speed and stability, the hourly average concentration from a source could vary by a factor of 2 from one day to the next because of turbulence. This natural variability imposes a limit on the accuracy of any transport and dispersion model.

Source effects. Before atmospheric dispersion has a chance to act on a plume, there are sometimes significant source effects due to differences between the density and velocity of the effluent and the ambient air. Most stack plumes are much warmer and hence less dense than the ambient air, and also usually have a significant upward velocity (on the order of 10 m/s or 33 ft/s) out of the stack. In this case, the plume can rise as much as several hundred feet above the stack before leveling off (**Fig. 3*a***). For constant source conditions, plume rise decreases as ambient stability and wind speed increase.

Sometimes the effluent is denser than the ambient air (for example, cold plumes or plumes containing gases that are denser than air), in which case the plume sinks toward the ground. Other dense gas sources may occur as a result of accidental rupturing of gas tanks. For example, a dense gas plume hugs the ground until it becomes sufficiently dilute to be dispersed by ambient turbulence (Fig. 3*b*).

At many short industrial stacks, the possibility exists that the plume may be drawn into the aerodynamic wake caused by the obstruction of the airflow by a nearby building. This may happen if the stack height is less than about 2.5 times the height of the structure (Fig. 3*c*). The ground-level concentration of air pollutants will increase because the plume is "downwashed" to the ground, but the average concentration in the plume is less because of the increased ambient turbulence.

Chemistry. Air pollution can be divided into primary and secondary compounds, where primary pollutants are emitted directly from sources (for example, carbon monoxide and sulfur dioxide) and secondary pollutants are produced by chemical reactions between other pollutants and atmospheric gases and particles (for example, sulfates and ozone). Most of the chemical transformations are best described as oxidation processes. In many cases, these secondary pollutants can have significant environmental effects, such as acid rain and smog.

Photochemistry. Los Angeles smog is a well-known example of secondary pollutants formed by photochemical processes, as a result of primary emissions of nitric oxide (NO) and reactive hydrocarbons from anthropogenic sources such as transportation and industry as well as natural sources. Energy from the

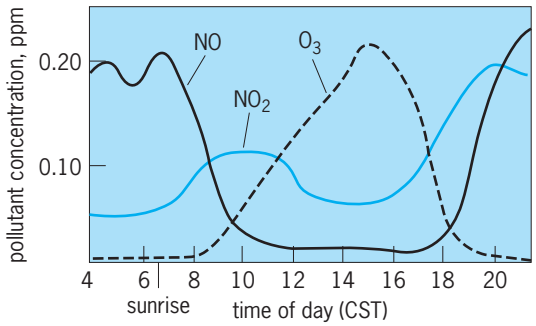

Fig. 4. The strong dependence of pollutant concentration on the intensity of sunlight. Note that ozone (O_3) concentration is a maximum in the afternoon, and zero at night. (*After A. C. Stern et al., Fundamentals of Air Pollution, 2d ed., Academic, 1984*)

Sun causes the formation of nitrogen dioxide, ozone (O_3), and peroxyacetalnitrate, which cause eye irritation and plant damage (**Fig. 4**).

Acid rain. The study of acid rain grew dramatically in the 1980s; prompted by the discovery of damage to sensitive lakes in areas where acidity of rain was the strongest. It has been shown that when emissions of sulfur dioxide and nitrogen oxide from tall power plant and other industrial stacks are carried over great distances and combined with emissions from other areas, acidic compounds can be formed by complex chemical reactions. In the absence of anthropogenic pollution sources, the average pH of rain is around 5.6 (slightly acidic). In the eastern United States, acid rain with a pH less than 5.0 has been measured and consists of about 65% dilute sulfuric acid, 30% dilute nitric acid, and 5% other acids (**Fig. 5**).

The chemistry of acid rain involves phase changes, catalysts, and aerosol formation, and is greatly influenced by cloud processes. A major component of sulfate deposition onto the surface is due to dry deposition, which occurs during all periods when precipitation is absent. Dry deposition is caused by impaction of particles or absorption of gases by the ground, water, or vegetation surface. It is possible that the chemical reactions are such that sulfate deposition is linearly related to emissions, implying that a 50% cut in industrial sulfur emissions would result in a 50% decrease in sulfate deposition. *See* ACID RAIN.

Particles. Particles in the atmosphere are characterized by size. The EPA has designated particles as supercoarse, coarse, fine, and ultrafine (**Table 3**). These distinctions generally represent differing sources and varying potentials for the particle to behave in the atmosphere and to affect human health. Supercoarse particles generally come from natural sources, including windblown sands and soil, volcanic activity, sea salt, and pollen. They are considered less of a human health threat since they have an appreciable mass and tend to fall out of the atmosphere quickly upon release. Moreover, the nose is reasonably efficient in removing these particles before they can reach the lungs.

Particulate matter less than approximately 10 micrometers (PM_{10}) can penetrate the lower respiratory tract. Course particles are a mix of natural and anthropogenic sources.PM_{10} is regulated as a specific

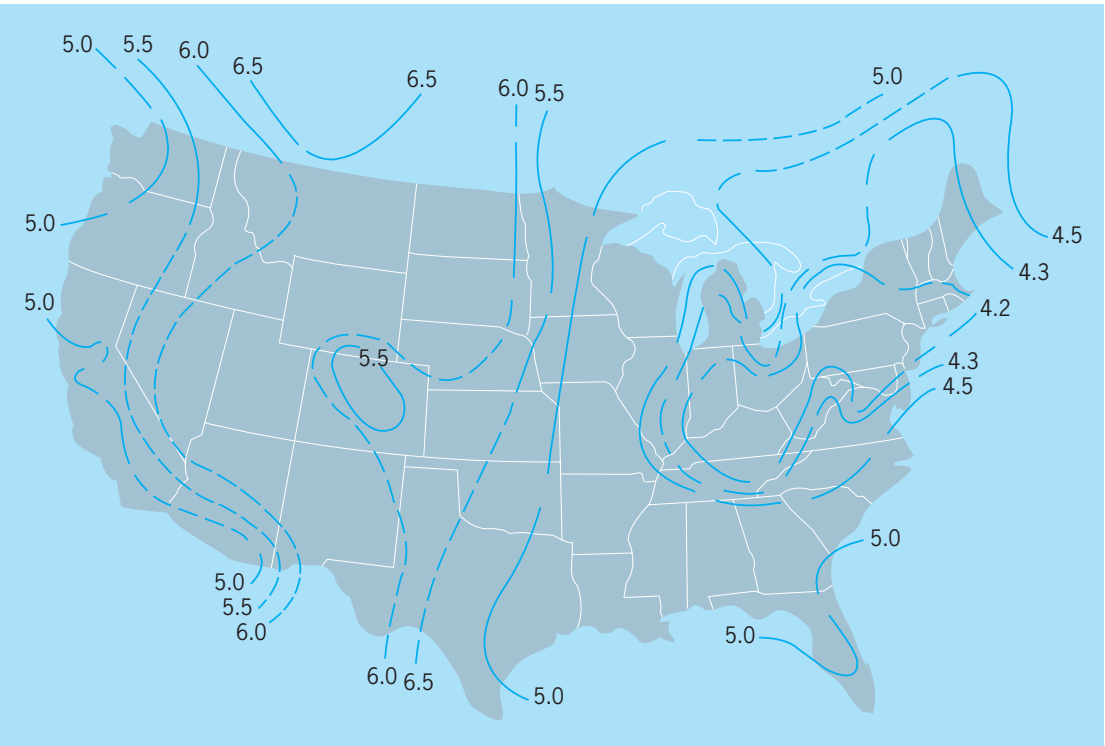

Fig. 5. Average pH of rain in the United States, showing the broad area of low pH in the northeastern United States, downwind of the large power plants in the Ohio Valley. A change in pH of 1.0 indicates a factor-of-10 change in acidity. (*After J. Wisniewski and E. L. Keitz, Acid rain deposition patterns in the continental United States, Air Soil Pollut., 19:327–339, 1983*)

TABLE 3. Terminology for particle sizes

Description	Particle size
Supercoarse	$D_p > 10\ \mu m$
Coarse	$2.5\ \mu m < D_p < 10\ \mu m$
Fine	$0.1\ \mu m < D_p < 2.5\ \mu m$
Ultrafine	$D_p < 0.1\ \mu m$

type of pollutant because this size range is considered respirable. The particle-size range between 0.1 and 10 μm is especially important in air pollution studies. A major fraction of the particulate matter generated in some industrial sources is in this size range.

Fine particles tend to be of anthropogenic origin and the result of particle growth from the ultrafine particles. Particles of this size, also referred to as the accumulation mode, grow through a cycle that can include condensation of flue gases, coagulation of smaller particles, and processes involved in the formation and evaporation of cloud droplets. Some of these particles have chemical characteristics that attract water vapor (hygroscopic) and will form "cloud condensation nuclei"—the basis for cloud droplet growth. Chemical reactions inside the cloud droplets and the collision and coalescence of the droplets with subsequent evaporation are believed to speed the formation of these size particles.

EPA chose 2.5 μm as the partition between fine and coarse particulate matter. Particles less than approximately 2.5 μm are regulated as $PM_{2.5}$. Air emission testing and air pollution control methods for $PM_{2.5}$ particles are different from those for coarse and supercoarse particles.

$PM_{2.5}$ particles settle quite slowly in the atmosphere, relative to coarse and supercoarse particles. Normal weather patterns can keep $PM_{2.5}$ particles airborne for several hours to several days and enable these particles to cover hundreds of miles. $PM_{2.5}$ particles can cause health problems due to their potentially long airborne retention time and the inability of the human respiratory system to defend itself against particles of this size.

In addition, the chemical makeup of $PM_{2.5}$ particles is quite different from coarse and supercoarse particles. EPA data indicate that $PM_{2.5}$ particles are composed primarily of sulfates, nitrates, organic compounds, and ammonium compounds. The EPA also determined that $PM_{2.5}$ particles often contain acidic materials, metals, and other contaminants believed to be associated with adverse health effects.

Particles less than 1 μm in diameter are called ultrafine (or submicrometer) particles and can be the most difficult size to collect. Particles in the range 0.2–0.5 μm are common in many types of combustion, waste incineration, and metallurgical sources. Particles in the range 0.1–1.0 μm are important because they can represent a significant fraction of the particulate emissions from some types of industrial sources and because they are relatively hard to collect.

Particles can be much smaller than 0.1 μm. In fact, particles composed of as little as 20 to 50 molecules clustered together can exist in a stable form. Some industrial processes, such as combustion and metallurgical sources, generate particles in the range 0.01–0.1 μm. These sizes are approaching the size of individual gas molecules, which are in the range 0.0002–0.001 μm. However, particles in the size range 0.01–0.1 μm tend to agglomerate rapidly to yield particles in the >0.1-μm range. Accordingly, very little of the particulate matter entering an air pollution control device or leaving the stack remains in the size range 0.01–0.1 μm.

Visibility reduction. Particles in the atmosphere are distributed over a wide range of sizes and shapes. They can be divided into Aiken particles (less than 0.2 μm), fine particles (0.2–2 μm), and large particles (greater than 2 μm). The Aiken particles are, in general, the result of condensation of hot gases within plumes. They often coagulate onto the fine particles through Brownian motion. The fine particles are too small to be removed from the atmosphere through sedimentation, and this size becomes the point of accumulation for particles in the atmosphere. Sources for the large particles can be either anthropogenic or natural, but the ratio of natural particles is highest for this size category.

The large particles have a size range that overlaps with that of the wavelengths of visible light and as such are very efficient at scattering light. Consequently, large particles play a significant role in reducing visibility. Research has shown the existence of large regions of low-visibility air that can be tracked from day to day as it moves around the eastern United States in the summertime. These regions are thought to be largely the result of the chemical production of sulfate particles from sulfur dioxide that is exacerbated by the presence of water vapor.

In the western United States, where the visibility is generally better than in the eastern half, individual plumes have been shown to degrade visibility in specific National Parks on specific days. However, the problem of plume impaction by aggregates of plumes from urban areas in the summertime in National Parks such as the Grand Canyon and Yosemite is significant and more difficult to regulate.

Air Quality Index. In an effort to create a more accessible measure of air quality, the EPA developed the Air Quality Index (AQI). The AQI indicates how clean or polluted the air is, as well as the associated health effects. The AQI focuses on health effects experienced within a few hours or days after breathing polluted air. The EPA calculates the AQI for five major air pollutants regulated by the Clean Air Act: ground-level ozone, particle pollution (also known as particulate matter), carbon monoxide, sulfur dioxide, and nitrogen dioxide. For each of these pollutants, the EPA has established national air quality standards to protect public health (Table 1).

The AQI can be considered a yardstick that runs from zero to 500. The higher the AQI value, the greater the level of air pollution and the greater the

health concern. For example, an AQI value of 50 represents good air quality with little potential to affect public health, while an AQI value over 300 represents hazardous air quality.

An AQI value of 100 generally corresponds to the national air quality standard for the pollutant, which is the level EPA has set to protect public health. AQI values below 100 are generally thought of as satisfactory. When AQI values are above 100, air quality is considered to be unhealthy—at first, for certain sensitive groups of people, then for everyone as AQI values get higher.

Stratospheric ozone layer. The stratospheric ozone layer absorbs much of the Sun's ultraviolet radiation, and any depletion of this ozone could have serious effects on animals and vegetation. Several types of air pollutants may attack the ozone layer, including fluorocarbons and nitrogen oxides. Photochemical reactions with hydrocarbons, nitrogen oxides, and ozone may lead to a new chemical balance in the stratosphere. The possibility of these changes was thought to have diminished because of a reduction in fluorocarbon use during the 1970s. However, in the mid-1980s observations of ozone at great elevations in the atmosphere suggested that there was a steady depletion, with major losses over Antarctica. *See* STRATOSPHERIC OZONE.

Accidental chemical releases. The chemical thermodynamics of some accidental releases are quite complicated. For example, if liquid chlorine is stored in a tank at high pressure and a nozzle ruptures, the releases will consist of two phases: liquid chlorine and gaseous chlorine. The evaporation process can lead to extreme cooling of the cloud close to the source.

Steven R. Hanna; Perry J. Samson

Controls

Air pollution controls have been designed for both mobile sources (mainly motor vehicles) and stationary sources.

Mobile sources. Air pollution control technologies for mobile sources involve vehicles fueled by gasoline and by diesel fuel. In addition, the fuel quality is an important factor.

Gasoline-fueled vehicles. A wide variety of emissions-reduction technologies has been developed for gasoline-fueled vehicles. Before controls were required, engine crankcases were vented directly to the atmosphere. Crankcase emissions controls, which basically consist of closing the crankcase vent port, were included in designs for automotive vehicles manufactured in the United States beginning in the early 1960s, with the result that control of these emissions is no longer considered a serious technical concern.

Exhaust emissions of hydrocarbons, carbon monoxide, and nitrogen oxides are related to the air–fuel mixture inducted, the peak temperatures and pressures in each cylinder, and other engine design parameters. Variations in these parameters are, therefore, capable of causing significant increases or decreases in emissions. *See* GASOLINE.

Dilution of the incoming charge has been shown to reduce peak cycle temperature by slowing flame speed and absorbing some heat of combustion. Recirculating a portion of the exhaust gas back into the incoming air–fuel mixture, thereby lowering peak cycle temperature, is used to lower the concentration of nitrogen oxides in the emissions. Improvements in mixture preparation, air-fuel intake systems, and ignition systems can increase dilution tolerance. It can also be increased by increasing the burn rate or flame speed of the air-fuel charge.

Both electronic control and exhaust aftertreatment devices are utilized to reduce air pollution.

1. *Electronic controls.* With so many interrelated engine design and operating variables playing an increasingly important role in the modern automotive engine, the control system has taken on increased importance. Modifications in spark timing must be closely coordinated with air–fuel ratio changes and degrees of exhaust gas recirculation to prevent significant decreases in fuel economy or performance which may result from emissions reductions, or to prevent an increase of NO_x emissions as carbon monoxide (CO) emissions decrease. In addition, controls that are selective in responding to engine load or speed have been found beneficial in preventing widespread adverse impacts resulting from emissions of pollutants. Therefore, electronic controls have replaced more traditional mechanical controls. For example, electronic control of ignition timing has demonstrated an ability to optimize timing under all engine conditions, and has the added advantage of reduced maintenance requirements and improved durability compared with mechanical systems. When it is coupled with electronic control of exhaust gas recirculation, NO_x emissions can be reduced without affecting fuel economy, in some cases resulting in improvements. *See* IGNITION SYSTEM.

2. *Exhaust aftertreatment devices.* When stringent exhaust emissions standards (especially for hydrocarbons or nitrogen oxides) are mandated, exhaust aftertreatment devices such as catalytic converters tend to be used to supplement engine modifications. An oxidation catalyst is a device that is placed on the tailpipe of a car and that, if the chemistry and thermodynamics are correct, will oxidize almost all the hydrocarbons and carbon monoxide in the exhaust stream. Three-way catalysts (which can lower levels of hydrocarbons, carbon monoxide, and nitrogen oxides simultaneously) were introduced in the United States in 1977, and subsequently became widely used when the United States standard for nitrogen oxides was lowered to 1.0 g per mile. To work effectively, these catalysts require precise control of air–fuel mixtures. As a result, three-way systems have indirectly fostered improved air-fuel management systems, such as throttle body fuel injection systems as well as electronic controls. One unique advantage of catalysts is their ability to selectively eliminate some of the more hazardous compounds in vehicle exhaust such as aldehydes, reactive hydrocarbons, and polynuclear

hydrocarbons. *See* AUTOMOBILE; CATALYSIS; CATALYTIC CONVERTER; FUEL INJECTION; INTERNAL COMBUSTION ENGINE.

Diesel-fueled vehicles. While the major technical problems associated with reducing emissions from gasoline-fueled cars have been solved, it is apparent that reductions from these vehicles alone are not sufficient to solve the air pollution problems in many areas. Therefore, diesel trucks and buses have received attention as significant sources of particulates and nitrogen oxides. Degradation of emissions is very low for diesel-fueled vehicles compared to gasoline-fueled ones; and diesel vehicles have inherently low hydrocarbon emissions, although these hydrocarbons are higher in molecular weight and thus of a different character from those emitted by gasoline engines. Further, uncontrolled diesel engines emit objectionable exhaust odors, which are a frequent source of complaints from the public. In the United States, standards have been adopted for these vehicles, which foster technological developments similar to those that have already been achieved for gasoline-fueled vehicles.

Smoke emissions from diesel engines are composed primarily of unburned carbon particles from the fuel. This type of emission usually results when there is an excess amount of fuel available for combustion. This condition is most likely to occur under high engine load conditions, such as acceleration and engine lugging, when the engine needs additional fuel for power. Failure to clean or replace a dirty air cleaner, a common maintenance error, may produce a large quantity of smoke emissions; a dirty air cleaner can choke off available air to the engine, resulting in a lower-than-optimum air–fuel mixture. The manner in which the vehicle is operated can also be important, since smoke emissions from diesel engines are minimized by selection of the proper transmission gear to keep the engine operating at the most efficient speeds. Moderate accelerations and fewer changes in highway cruising speed as well as reduced speed for hill climbing also minimize smoke emissions.

Basic approaches to diesel engine emission control fall into three major categories: (1) engine modifications, including combustion chamber configuration and design, fuel-injection timing and pattern, turbocharging, and exhaust gas recirculation; (2) exhaust aftertreatment, including traps, trap oxidizers, and catalysts; and (3) fuel modifications, including control of fuel properties, fuel additives, and alternative fuels. *See* COMBUSTION CHAMBER; TURBOCHARGER.

Design strategies for controlling emissions of nitrogen oxides in diesel-fueled vehicles include variable injection timing and pressure, charge cooling (cooling of the incoming air to make it more dense), and exhaust gas recirculation. Retarding injection timing, while a well-known method of reducing formation of nitrogen oxides, can lead to increases in fuel consumption and emissions of particulates and hydrocarbon. These problems can be mitigated by varying the injection timing with engine load or speed. Also, high-pressure injection can reduce these problems. If this injection is coupled with electronic controls, emissions of nitrogen oxides can be reduced significantly with a simultaneous improvement in fuel economy (although not as great as could occur if electronics were added without any emission requirements).

With relatively lenient particulate standards, engine modifications are generally sufficient to lower engine-out emission levels (emissions out of the engine but before they would go into some type of aftertreatment device; in this case, synonymous with tailpipe emission levels). The necessary modifications include changes in combustion chamber design, fuel injection timing and spray pattern, turbocharging, and the use of exhaust gas recirculation. Further particulate controls are possible through greater use of electronically controlled fuel injection. With such a system, signals proportional to fuel rate and piston advance position are measured by sensors and are electronically processed by the electronic control system to determine the optimum fuel rate and timing.

One of the important factors in diesel control design is that emissions of nitrogen oxides, particulates, and hydrocarbons are closely interdependent. For example, retarded fuel timing within certain ranges decreases emission of nitrogen oxides and particulates but can lead to increases in emission of hydrocarbons. As technology has advanced, these potential trade-offs have diminished. For example, engine designs such as modifications in the combustion chamber and electronically controlled fuel injection have resulted in simultaneous reduction in emissions of hydrocarbons, particulates, and nitrogen oxides.

Exhaust aftertreatment methods include traps, trap oxidizers, and catalysts. Prototypes of trap oxidizer systems have achieved 70–90% reductions from engine-out particulate emissions rates and, with proper regeneration, have demonstrated the ability to achieve these rates for high milage. Basically, all these designs rely on trapping a major portion of the engine-out particles and consuming them before they accumulate sufficiently to saturate the filter and cause problems in fuel economy and performance or other difficulties. *See* DIESEL ENGINE.

Improved fuel quality. Improvements in fuel quality are a significant factor in control of air pollution from mobile sources. Modifications to diesel fuel composition are a quick and cost-effective means of reducing emissions from existing vehicles. The two modifications that show the most promise are reductions in sulfur content and in the fraction of aromatic hydrocarbons in the fuel. In the United States, the EPA has reduced sulfur content in diesel fuel to a maximum of 0.05% by weight.

Lowering the sulfur content of diesel fuel not only causes a direct reduction in emissions of sulfur dioxide (SO_2) and sulfate particles but also reduces the indirect formation of sulfate particles from SO_2 in the atmosphere. In Los Angeles, it is estimated that each pound of SO_2 emitted results in roughly 2.2 kg (1 lb) of fine particulate matter in the atmosphere.

In this case, therefore, the indirect particulate emissions due to SO_2 from diesel vehicles are roughly equal to their direct particulate emissions. Conversion of SO_2 to particulate matter is highly dependent on local meteorological conditions, so the effects could be greater or less in other cities.

A reduction in the aromatic hydrocarbon content of diesel fuel may also help to reduce emissions, especially where fuel aromatic levels are high. For existing diesel engines, a reduction in aromatic hydrocarbons from 35% to 20% by volume would be expected to reduce transient particulate emissions by 10–15% and emissions of nitrogen oxides by 5–10%. Hydrocarbon emissions, and possibly the mutagenic activity of the particulate soluble organic fraction, would also be reduced. Aromatic reductions of this magnitude can often be obtained through alterations in methods of diesel fuel production and blending strategy without a need for major new investments in additional processing capacity.

Reduced aromatic content in diesel fuels would have other environmental and economic benefits. The reduced aromatic content would improve the fuel's ignition quality, facilitating cold starting and idling performance and reducing engine noise. The reduction in the use of catalytically cracked blending stocks should also have a beneficial effect on deposit-forming tendencies in the fuel injectors, and would reduce maintenance costs. On the negative side, the reduced aromatics might result in some impairment of cold flow properties, due to the increased paraffin content of the fuel.

Detergent additives in diesel fuel have the ability to prevent and remove injector tip deposits, thus reducing smoke levels. The reduced smoke probably results in reduced particulate emissions as well, but this has not been demonstrated as clearly because of the great expense of particulate emissions tests on in-use vehicles. Cetane-improving additives can also produce some reduction in emissions of hydrocarbons and particulates in marginal fuels. *See* CETANE NUMBER.

Alternative fuels. The possibility of substituting cleaner-burning fuels for diesel fuel has drawn attention since the mid-1980s. The advantages provided by this substitution include conservation of oil products and the security of readily available energy sources, as well as the reduction or elimination of particulate emissions and visible smoke. The most promising of the alternative fuels are natural gas, methanol made from alcohol, and, in limited applications, liquefied petroleum gas (LPG).

1. *Natural gas.* Natural gas has many desirable qualities as an alternative to diesel fuel in heavy-duty vehicles. Clean burning, cheap, and abundant in many parts of the world, it already plays a significant role as an energy source for vehicles in a number of countries. The major disadvantage is its gaseous form at normal temperatures. Pipeline-quality natural gas is a mixture of several different gases, but the primary constituent is methane, which typically makes up 90–95% of the total volume. Methane is a nearly ideal fuel for Otto cycle (spark ignition) engines. As a gas under normal conditions, it mixes readily with air in any proportion, eliminating cold-start problems and the need for cold-start enrichment. It is flammable over a fairly wide range of air–fuel ratios. With a research octane number of 130 (the highest of any commonly used fuel), it can be used with engine compression ratios as high as 15:1 (compared to 8–9:1 for gasoline), thus giving greater efficiency and power output. The low lean flammability limit permits operation with extremely lean air-fuel ratios, having as much as 60% excess air. However, its high flame temperature tends to result in high emissions of nitrogen oxides, unless very lean mixtures are used.

Because of its gaseous form and poor self-ignition qualities, methane is a poor fuel for diesel engines. Since diesels are generally somewhat more efficient than Otto cycle engines, natural gas engines are likely to use somewhat more energy than the diesels they replace. The high compression ratios achievable with natural gas limit this efficiency penalty to about 10% of the diesel fuel consumption, however. *See* LIQUEFIED NATURAL GAS (LNG); METHANE; NATURAL GAS; OTTO CYCLE.

2. *Liquefied petroleum gas.* Liquefied petroleum gas is widely used as a vehicle fuel in the United States, Canada, the Netherlands, and elsewhere. As a fuel for spark ignition engines, it has many of the same advantages as natural gas, with the additional advantage of being easier to carry aboard the vehicle. Its major disadvantage is the limited supply, which would rule out any large-scale conversion to liquefied petroleum gas. *See* LIQUEFIED PETROLEUM GAS (LPG).

3. *Methanol.* Methanol has many desirable combustion and emissions characteristics, including good lean combustion characteristics, low flame temperature (leading to low emissions of nitrogen oxides), and low photochemical reactivity. As a liquid, methanol can either be burned in an Otto cycle engine or be injected into the cylinder in a diesel. With a fairly high octane number of 112 and excellent lean combustion properties, methanol is a good fuel for lean-burn Otto cycle engines. Its lean combustion limits are similar to those of natural gas, while its low energy density results in a low flame temperature compared to hydrocarbon fuels, and thus lower emissions of nitrogen oxides. Methanol burns with a sootless flame and contains no heavy hydrocarbons. As a result, particulate emissions from methanol engines are very low, consisting essentially of a small amount of unburned lubricating oil. Methanol's high octane number results in a very low cetane number, so that methanol cannot be used in a diesel engine without some supplemental ignition source. The use of ignition-improving additives, spark ignition, glow plug ignition, or dual injection with diesel fuel in converted heavy-duty diesel engines has been demonstrated. *See* ALCOHOL FUEL; METHANOL; OCTANE NUMBER.

Methanol combustion does not produce soot, so particulate emissions from methanol engines are limited to a small amount derived from lubricating oil. Methanol's flame temperature is also lower than that

for hydrocarbon fuels, resulting in emissions of nitrogen oxides that are typically 50% lower. Emissions of carbon monoxide are generally comparable to or somewhat greater than those from a diesel engine (except for stoichiometric Otto cycle engines, for which emissions of carbon monoxide may be much higher). These emissions can be controlled with a catalytic converter, however.

The major pollution problems associated with methanol engines come from emissions of unburned fuel and formaldehyde. Methanol (at least in moderate amounts) is relatively innocuous—it has low photochemical reactivity—and, while acutely toxic in large doses, displays no significant chronic toxicity effects. Formaldehyde, the first oxidation product of methanol, is much less benign. A powerful irritant and suspected carcinogen, it also displays very high photochemical reactivity. Although all internal combustion engines produce some formaldehyde, some early designs of methanol engines exhibited greatly increased formaldehyde emissions as compared to those from diesel engines. The potential for large increases in emissions of formaldehyde with the widespread use of methanol vehicles has raised considerable concern about what would otherwise be a very benign fuel from an environmental standpoint. Formaldehyde emissions can be reduced through changes in the design of the combustion chamber and the injection system; they are also readily controllable through the use of catalytic converters, at least under warmed-up conditions. *See* FORMALDEHYDE. Michael P. Walsh

Stationary sources: volatile and hazardous pollutants. Volatile organic compounds and hazardous air pollutants come from many sources. In the United States, hazardous air pollutants are identified in Title III of the Clean Air Amendments Act of 1990, as pollutants that present, or may present, through inhalation or other routes of exposure, a threat of adverse human health effects. Volatile organic compounds react in the atmosphere to form photochemical oxidants (including ground-level ozone) that affect health, damage materials, and cause crop and forest losses. In the United States, the EPA has established a 120-parts-per-billion ambient ozone standard (1-h average, daily maximum) and an 80-parts-per-billion standard (8-h average). The 120-ppb standard applies only to areas that were designated nonattainment when the 80-ppb standard was adopted in July 1997. Millions of people live in areas where the ozone standards are routinely exceeded. Hazardous air pollutants can also cause detrimental effects on human health.

Controlling the emissions of volatile organic compounds and hazardous air pollutants is important but complex, because there are many sources that usually emit mixtures of compounds. Because organic compounds may undergo complex reactions in the atmosphere, there are a number of technologies designed to control emissions. Often, it is preferable to prevent the formation and emission of organic compounds rather than attempting to control them. Organic emissions can be reduced or prevented by using alternative chemicals, by modifying chemical processes, and by changing the products being used. In addition to preventing pollution, such measures can save energy and reduce costs.

Conventional control technologies. These include thermal incinerators, catalytic oxidizers, adsorbers, flares, boilers and industrial furnaces, and condensers.

1. *Thermal incinerators.* Thermal incinerators raise the temperature of emission streams so that organic compounds are destroyed; destruction efficiencies can exceed 99%. Thermal incinerators are typically applied to emission streams that contain dilute mixtures of volatile organic compounds and hazardous air pollutants. The effectiveness of a thermal incinerator is largely dependent on the temperature in the combustion chamber. Supplementary fuels (natural gas or fuel oil) must be used to maintain desired combustion temperatures. Thermal incinerators are available in many sizes.

2. *Catalytic oxidizers.* Catalytic oxidizers use a catalyst to speed up the reaction of organic compounds at temperatures lower than those required by thermal incinerators, thus reducing requirements for supplementary fuel. Destruction efficiencies are in the 95% range, but they can be higher if more catalyst and supplementary fuel are used. The catalyst may be fixed (fixed bed) or mobile (fluid bed). Catalytic oxidation is not as widely used as thermal incineration because some materials may damage the catalyst. Even with proper usage, the catalyst must be replaced periodically. Catalytic oxidizers are available in the same size range as thermal incinerators.

3. *Adsorbers.* Adsorbers employ a bed packed with an adsorbent material that captures volatile organic compounds and hazardous air pollutants. The most common adsorbent is carbon, but other materials, such as silica gel, alumina, and polymers, are also used. Removal efficiencies up to 99% can be obtained in adsorbers, and they provide the option of reclaiming the adsorbed material for reuse. There are many important organic compounds for which carbon adsorption is quite effective. Some examples are toluene, benzene, and methyl ethyl ketone. Carbon adsorbers are sensitive to excessive humidity, emission stream temperatures above 54°C (130°F), and inlet organic compound concentrations of 10,000 ppm and above. *See* ADSORPTION.

4. *Absorbers.* Absorbers are control devices that provide contact between the emission stream and a carefully selected solvent in which the volatile organic compounds and hazardous air pollutants readily dissolve. When the solvent has absorbed the pollutants, the solvent must be regenerated or disposed of. Regeneration of the solvent can be expensive, and disposal can be a problem because of the potential for surface-water or ground-water contamination. Thus, absorbers are often used in conjunction with other control devices such as thermal incinerators. *See* ABSORPTION.

5. *Flares.* Flares are emission control devices that destroy volatile organic compounds and hazardous air pollutants by combustion. Often, supplemental fuel is added to maintain the combustion process

and to destroy the pollutants. Flares are often chosen as emission control devices for volatile organic compounds and hazardous air pollutants when emission stream flow is uncertain or intermittent, as in process upsets and emergencies.

6. *Boilers and industrial furnaces.* In-place boilers and industrial furnaces can control emissions of volatile organic compounds and hazardous air pollutants. Not only can these units control an emission stream that is the main fuel supply, but they can also control a waste stream that has a small flow rate in comparison to the flow rate of the fuel–air mixture used to fire the unit. Since proper functioning of such systems is essential to the proper operation of a plant, their use as control devices for volatile organic compounds and hazardous air pollutants must be carefully monitored to avoid adverse impacts on their performance and reliability.

7. *Condensers.* Condensers remove volatile organic compounds and hazardous air pollutants from emission streams by lowering the temperature of the emission stream, thus condensing these substances. Condensers are widely used to reduce the concentrations of volatile organic compounds and hazardous air pollutants in emission streams before they enter other control devices such as thermal incinerators, catalytic oxidizers, and adsorbers.

Alternative control technologies. These include biofilters, separation systems, ultraviolet oxidizers, corona destruction, and methods involving plasmas.

1. *Biofilters.* Biofilters use microorganisms to destroy volatile organic compounds and hazardous air pollutants. In the biofilter, the emission stream comes in contact with microorganisms that feed on these materials. Biofilters may have cost advantages over other control technologies, but significant issues such as proper matching of microorganisms to specific emission streams, long-term operational stability, and waste disposal may also be factors.

2. *Separation systems.* Separation technologies have been used extensively to control emissions of volatile organic compounds and hazardous air pollutants. Absorption technology has utilized such improvements as high-efficiency packing and trays, highly selective solvents, and advanced process configurations. Hydrophobic molecular sieves and superactivated carbons have enhanced the efficiency and applicability of adsorption technology. Improved membranes to separate and recover organic compounds may also serve to control emissions. *See* ACTIVATED CARBON; CHEMICAL SEPARATION (CHEMICAL ENGINEERING); MEMBRANE SEPARATIONS; MOLECULAR SIEVE.

3. *Ultraviolet oxidizers.* Ultraviolet radiation has the potential to destroy some volatile organic compounds and hazardous air pollutants. When organic compounds absorb ultraviolet light energy, they are activated; and they may decompose (photodissociation) or react with other compounds (radical oxidation). This technology may be combined with the use of a catalyst to increase reaction rates.

4. *Corona destruction.* There are two principal types of corona destruction reactors: packed bed corona and pulsed corona. The packed bed reactor is packed with high dielectric material, and high voltage (15,000–20,000 V) is applied to electrodes at both ends of the bed. High-energy electric fields are created in the spaces between the packing material, and volatile organic compounds and the hazardous air pollutants in an emission stream passed through the bed are destroyed. The pulsed corona reactor has a wire electrode in the center, and the wall of the reactor serves as the other electrode. When high voltages are applied to the two electrodes at nanosecond intervals, high-energy fields are established, and pollutants passing through the reactor are destroyed. *See* CORONA DISCHARGE.

5. *Plasmas.* Processes involving plasma technology include mechanisms for creating very high temperatures. The volatile organic compounds and hazardous air pollutants in incoming waste streams are destroyed as a result of the high temperatures and the interaction with other chemical species that exist at these temperatures. Wade H. Ponder

Stationary sources: particulates. Particulate control devices remove particulate matter from a gas stream prior to its discharge to the atmosphere. Generally these devices are in the form of mechanical collectors, wet scrubbers, electrostatic precipitators, and fabric filters. Occasionally two or more of these mechanisms are used in one device. Control of fine particles from large industrial sources predominantly employs electrostatic precipitators and fabric filters.

Mechanical collectors. These devices range from simple settling chambers, which make use of gravity, to rotary-fan collectors, to cyclone and multiclone inertial devices. In general, these devices can collect particles of about 5 μm in diameter with efficiencies of 90% or more. They are ineffective for the submicrometer particles whose control is required by modern air pollution regulations. Mechanical collectors are useful for applications in which the particulate matter consists of only large particles. A second use is to remove most of the particulate matter from a gas stream before it enters a high-efficiency control device in order to minimize material handling. *See* CENTRIFUGATION.

Wet scrubbers. This generic term covers many types of control devices for capturing particles within a gas stream by bringing them into contact with a liquid. The capture mechanisms can be inertial impaction, direct interception, or diffusion resulting from Brownian motion of very fine particles. Inertial impaction and direct interception depend on the number of waterdroplets and the relative velocities of the particles and droplets. This implies high energy to the scrubber supplied through either the gas or liquid streams or by a mechanical device such as an impeller. Diffusion capture of small particles is enhanced by adding energy to increase the quantity of small liquid droplets, which limits the length of the diffusion path required for capture. To meet modern air pollution requirements for control of fine particles, a high-energy input is needed, usually as a high-pressure drop across the device. In addition, liquid waste treatment is required for the scrubbing

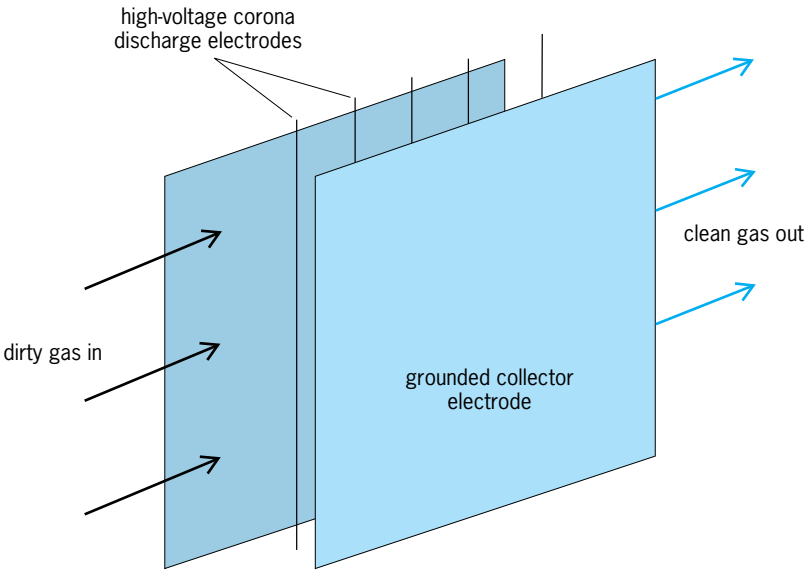

Fig. 6. Single lane of one section of an electrostatic precipitator.

process. However, a potential side benefit of scrubbing is the simultaneous ability to collect, by mass transfer, gaseous pollutants in the gas stream.

To help offset the high power requirement that results from the pressure drop across the scrubber, it is possible that electrostatic attractive forces can be introduced to improve capture. Energy needs can also be reduced by causing condensation upon the particles in the saturated atmosphere of the scrubber.

Electrostatic precipitators. These devices give particles an electrical charge and subject them to an electrical field that exerts a force to move them to the collecting electrode. In the basic process, particles are charged and collected simultaneously (**Fig. 6**). Incoming particles are charged by ions that are generated by the corona, which surrounds the high-voltage discharge electrodes. The high voltage also establishes the electric field to the grounded electrodes, which collect the particles. Periodically, the collected particles are knocked from the collector electrodes by mechanical rapping, and they fall into a hopper located beneath them.

An industrial-size electrostatic precipitator, such as at an electric utility, has collector electrodes 7.5–13.5 m (25–45 ft) in height and 1.5–4.5 m (5–15 ft) in length. A section might have 30 to 60 or more adjacent lanes with widths of 200–450 mm (8–18 in.). The unit might also have three to eight consecutive sections. Gas flow through the electrostatic precipitator could range from 250,000 to 5,000,000 m^3/h (150,000 to 3,000,000 ft^3/min) at temperatures from 150 to 343°C (300 to 650°F). The voltage applied to the corona discharge electrode ranges from 30 to more than 80 kV. The power required for operation is predominantly electrical. The pressure drop through the units is considerably less than for scrubbers.

Variations to the basic electrostatic precipitator process improve performance. Gas conditioning (addition of chemical additives to a gas stream) and pulse energization (intense pulses of direct-current high voltage supplied to the electrostatic precipitator) help overcome the detrimental effects of high-resistivity particle matter. Particles having a high electrical resistivity are more difficult to charge and collect than are particles of lower electrical resistivity. Electrostatic precipitators with two stages separate the charging and collecting functions, allowing both to be optimized for maximum efficiency. Continuous flushing away of the collected particles with water allows simultaneous collection by mass transfer of gaseous pollutants in the gas stream. However, operating the unit in the wet mode introduces a wastewater treatment requirement.

Electrostatic evaporators can be designed to specific efficiency requirements. Computer performance prediction models help in the engineering design of these devices. *See* ELECTROSTATIC PRECIPITATOR.

Fabric filtration. Filtration by collection of particles on a fabric makes use of inertial impaction, direct interception, and diffusion capture. The filtration fabric may be configured as bags, envelopes, or cartridges. For use with processes having high volumes of emissions, bag filters are the most important. The fabric bags are contained in a baghouse, fitted with automatic cleaning mechanisms to remove the collected particles, making it unnecessary to remove and replace dirty bags.

There are two basic methods for arranging and cleaning filter bags in baghouses. These are inside-to-outside flow (**Fig. 7***a*) in which the particulate matter is collected on the interior surface, and outside-to-inside flow (Fig. 7*b*) in which the particulate matter is collected on the outside of the bags. Inside-to-outside bags are cleaned by reversing the gas flow momentarily, shaking the bags, or a combination of both methods. Outside-to-inside flow bags are cleaned by directing a jet of air down the center of the bag to

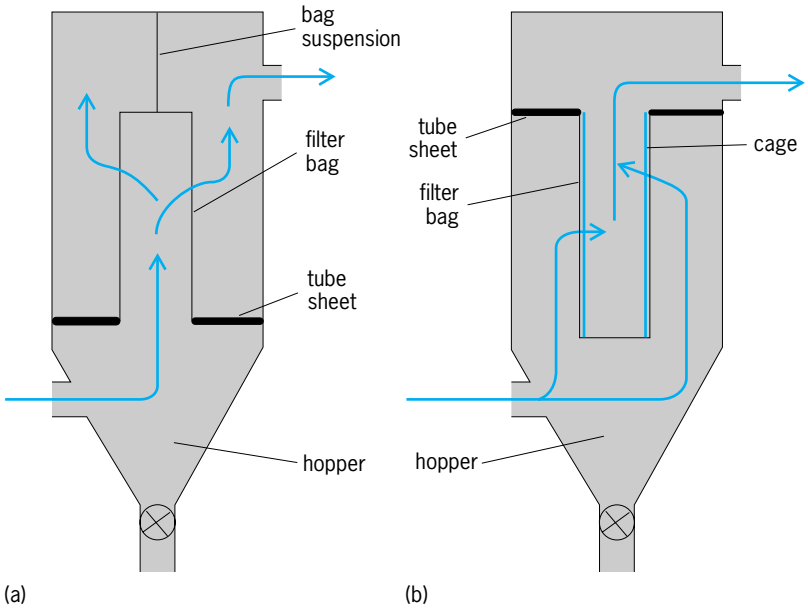

Fig. 7. Basic configurations of fabric filtration baghouses. (*a*) Inside-to-outside flow. (*b*) Outside-to-inside flow. Arrows indicate direction of gas flow.

dislodge the deposited dust. In both cases the particulate matter falls into a hopper, from which it is removed. The bags are attached to a tube sheet that separates the clean side of the baghouse from the dirty side. Outside-to-inside flow bags contain an internal cage to prevent their collapse during filtration. Inside-to-outside bags, in industrial applications, are as large as 10.8 m (36 ft) in length with a diameter of 12 in. (300 mm). Outside-to-inside bags range from 2.5 to 7.2 m (8 to 24 ft) in length with diameters from 100 to 200 mm (4 to 8 in.). Depending upon the size of the process to be controlled, a baghouse might contain from fewer than 100 bags to several thousand bags. Power required for operation is needed predominantly to move the gas through the device. The pressure drop for filtration is greater than for electrostatic precipitators but less than for scrubbers.

Inside-to-outside bags are usually made from woven fibers, whereas outside-to-inside bags use felted fibers. The temperature and chemical characteristics of the gas stream determine the type of fabric material used. Performance prediction models are utilized in the design; however, data from field testing are usually required for their use. For example, one focus has been on the use of electrostatics with fabric filtration to increase fine-particle capture and decrease pressure drop. *See* AIR FILTER; DUST AND MIST COLLECTION; FILTRATION. Norman Plaks

Hybrid electrostatic precipitator–fabric filter concepts. At least three hybrid systems are being developed to reduce fine PM emissions: the Compact Hybrid Particulate Collector (COHPAC), developed by the Electric Power Research Institute; the Advanced Hybrid Particulate Collector (AHPC), developed by the Energy and Environmental Research Center of North Dakota under U.S. Department of Energy (DOE) sponsorship; and the Electrostatically Stimulated Fabric Filter (ESFF), being developed by Southern Research Institute under EPA sponsorship. The COHPAC I concept adds a small pulse-jet fabric filter baghouse immediately downstream of an electrostatic precipitator or replaces the latter stages within the precipitator structure (COHPAC II). Residual particle charge is reported to improve the fabric filter performance, so that very low emissions have been experienced.

The AHPC system combines reduced-size electrostatic precipitator and fabric filter compartments within a single vessel and introduces proprietary synergistic principles to achieve reduction of reentrainment and recollection of dust caused by close bag spacing, reduction of chemical attack on bagfilters, and ultrahigh collection efficiencies.

The ESFF system consists of dust precharging at the baghouse inlet and vertical charging wires placed at the centerline of each four-bag array that induce an electrostatic field between the wire and the grounded cage supporting each bagfilter. Modeling based on small-scale research indicates up to an order of magnitude reduction of submicrometer particles and two orders of magnitude supermicrometer particle reduction at a pressure loss of only 20–30% compared to that of conventional fabric filtration.

Stationary sources: nitrogen and sulfur oxides. The major gaseous pollutants emitted from stationary emission sources in the United States include sulfur oxides (SO_x) and nitrogen oxides (NO_x) from fuel combustion, waste combustion, metallurgical processes, and ore reduction. Sulfur dioxide (SO_2) and sulfur trioxide (SO_3) are emitted from fuel combustion. Sulfur dioxide is the dominant pollutant in most applications, but there is also a concern with the emission of acid gases, notably hydrochloric acid (HCl) from waste combustion. Adverse health effects from exposure to both SO_x and NO_x have been documented and have been the focus of most major air pollution legislation, along with particulate matter, in the United States. For example, the Clean Air Act of 1970 established National Ambient Air Quality Standards for emissions of SO_2 and NO_x from electric power plants. The Clean Air Act Amendments of 1990 identify SO_2 and NO_x as the major contributors to acid rain, and specify major reductions in emissions from the largest emitting sources—coal-fired electric power plants. *See* ACID RAIN.

Sulfur oxides are formed during fuel combustion by the oxidation of naturally occurring sulfur compounds in both mineral and organic portions of fossil fuels. Fuel nitrogen oxides are formed by oxidation of nitrogen compounds in fuel, and thermal nitrogen oxides are formed by the reaction of nitrogen and oxygen in air within a high-temperature flame. When wastes are burned, chlorinated compounds in synthetic materials are converted to HCl gas, while sulfur- and nitrogen-bearing wastes are oxidized to SO_x and NO_x. Many metallurgical processes use ores that occur naturally as sulfides: copper, zinc, iron, lead, and mercury are examples. Roasting these ores liberates fairly high concentrations of SO_x compared to fossil fuel combustion.

Control of SO_x, NO_x, and HCl may occur prior to combustion by pretreatment of fuels or wastes (precombustion), within the combustion zone furnace, or may take place after combustion. For solid fuels, the precombustion technologies are generally lower in cost and lower in efficiency, and the postcombustion technologies are higher in cost and higher in efficiency. For liquid and gaseous fuels, fuel pretreatment is generally more effective and cheaper than postcombustion technology. The strategies for control vary with the type and size of the parent process and the stringency of regulations.

Conventional technologies. These include precombustion controls, combustion controls, and postcombustion controls.

1. *Precombustion controls.* Fuel or waste pretreatment can reduce the sulfur, nitrogen, or chlorine content and subsequent emissions from the combustion process. Examples are gas sweetening, hydrotreating of liquid fuels, physical coal cleaning plants, and sorting of wastes prior to incineration.

• Gas sweetening. Natural gas and gases derived from oil, oil shale, bitumen, and coal may contain significant amounts of hydrogen sulfide, which when combusted forms SO_x. Absorbers that use regenerable sorbent liquids are used to remove and

concentrate the hydrogen sulfide gas into a feed stream for a sulfur plant, which converts gaseous hydrogen sulfide into elemental sulfur liquid.

• Oil desulfurization. Petroleum liquids and crude oil may be treated by injection of hydrogen gas, vacuum distillation, and partial oxidation to liberate and remove sulfur and nitrogen impurities. When burned, the cleaned fuels emit significantly lower concentrations of SO_x and NO_x than untreated fuels.

• Coal cleaning. A major industry within coal mining and distribution is that of physically cleaning coal to upgrade the coal heating value. Undesirable ash is physically removed by a number of washing processes. In the case of coals with higher sulfur content that are found in the eastern and midwestern United States, a significant amount of sulfur is also removed in the coal cleaning process, typically 20–30% of the total sulfur. This directly results in comparable reduction of SO_x when coal is burned.

2. *Combustion controls.* Within the combustion process, a number of techniques have been developed to minimize the formation of NO_x. Attempts to premix chemical sorbents with fuel or inject sorbents into the combustor for control of SO_x have not been as successful. Preferred controls for combustion NO_x vary according to fuel and combustor type. Major combustion sources of NO_x include steam boilers, gas-fired turbines, internal combustion engines, and waste combustors.

For steam boilers, minimizing the combustion air (low excess air) is commonly practiced. Minimizing the air-to-fuel ratio in the combustor reduces NO_x by 5–10% over uncontrolled combustion. Staging burners are devices that either lower oxygen in the flame to render it fuel-rich or lower peak flame temperature to render it air-rich; either strategy lowers NO_x formation. Staging the air or fuel to lower the peak temperature and availability of oxygen for NO_x formation can reduce emissions of NO_x by 10–45%. Special staging burners (low NO_x burners) are used in large utility boilers to reduce emissions of NO_x up to 45%. Supplemental, low-nitrogen fuels, such as natural gas or distillate oil, can be added via a special fuel-staging system, known as fuel reburning, which reduces NO_x by up to 70% in large coal-fired boilers.

Techniques for reducing flame temperature, and hence NO_x, in other NO_x sources include water injection in turbines and special ignition systems for diesel-fuel and natural-gas-fuel engines.

3. *Postcombustion controls.* These controls reduce sulfur oxide, nitrogen oxide, and hydrogen chloride.

• Sulfur oxide. For control of SO_x, the dominant system is the wet flue-gas desulfurization system, or wet scrubber, which contacts flue gas in an absorber vessel with a slurry or solution spray of water and alkaline sorbent (**Fig. 8**). Depending on the source of the SO_x, the economic choice of sorbent may be lime, limestone, or sodium carbonate. The SO_x reacts to form a salt in the scrubber liquid, which exits the bottom of the scrubber. A water treatment system separates the solids from water, and the wet solids are disposed of in ponds or landfills. In the case of large coal-fired utility boilers, the limestone scrubber is dominant, and the resulting solid salts or sludge create a considerable disposal problem. Use of a special oxidation step can convert the sludge to

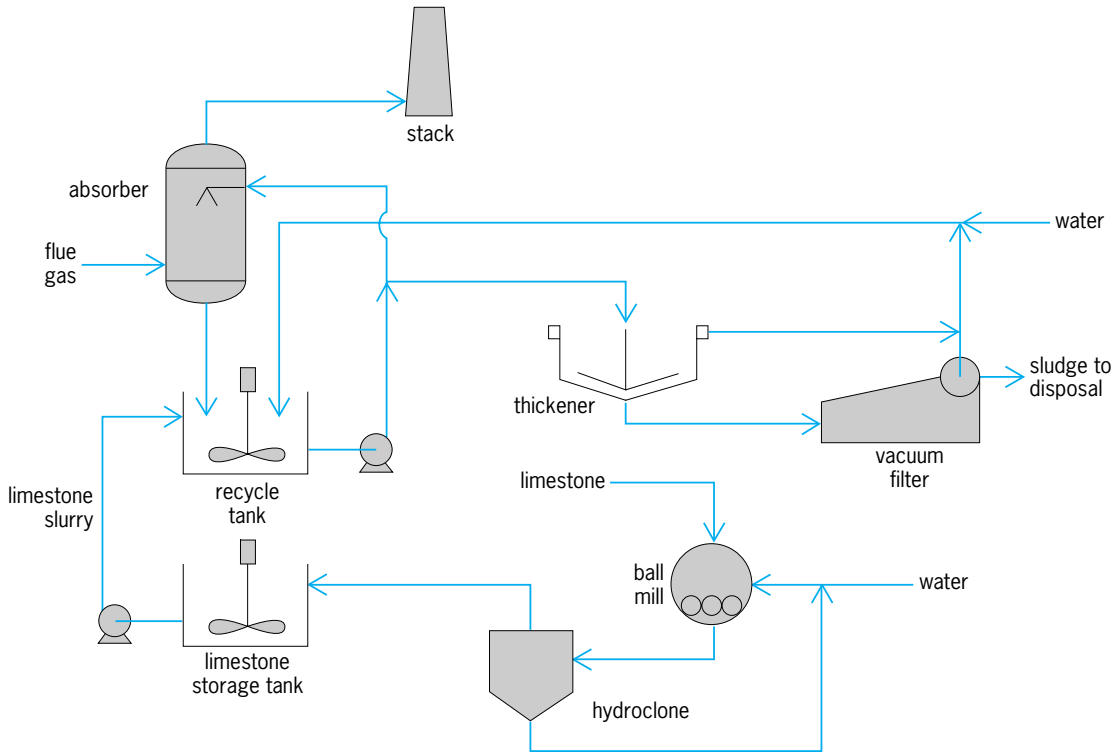

Fig. 8. Wet scrubber SO_x control system.

gypsum, which has commercial value in construction. A special subset of scrubbing includes the use of reduced water scrubbers (spray dryer absorbers). The process is sometimes known as dry scrubbing (which is a misnomer since the scrubbing is wet but the product is collected dry). Spray dryers are common in smaller industrial boilers, cogeneration boilers, and waste combustors. The advantages of lime spray drying are reduced water usage, no water–solids separation, and lower capital cost. Whereas halogenated acids such as hydrochloric acid (HCl) adversely affect wet scrubber operations in mass transfer, solids separation, and materials, the presence of chlorides in spray dryers enhances mass transfer and collection of fine particles. Sulfuric acid mist is also more readily collected in the spray dryer–dust collector system. Disadvantages are a dry waste of fly ash and spent sorbent which is difficult to market, usually ending up in a landfill; and limited application to low-to-moderate sulfur fuels (95% SO_2 removal on $<2\%$ sulfur fuel is the upper limit for most applications).

Another type of gas absorber is the circulating-bed scrubber using hydrated lime to absorb acid gases. The system operates upstream of an existing dust collector, much like the spray dryer, but is essentially a vertical pipe with internals that encourage solids–gas mixing and flash drying of moisture, within a fraction of the gas residence time of its competitors. With no lime slurrying or spray nozzles, many of the operating problems of dry scrubbers are avoided. With high solids recirculation, 98% removal of sulfur oxides has been achieved in European applications, and the system has been applied to higher-sulfur (4% sulfur) fuels. The disadvantages of this system are dry solids disposal, upper size limits on single-module operation, and larger solids loadings on the dust collector.

In ore roasting, the concentrations of SO_x are generally an order of magnitude higher than for fuel burning. These high concentrations of SO_x make recovery of SO_x by conversion to sulfuric acid somewhat profitable. Therefore, many primary smelters generate by-product sulfuric acid from SO_x in conventional acid plants. *See* SULFURIC ACID.

• Nitrogen oxide. Postcombustion control of NO_x consists of two techniques that use ammonia-based liquids injected into the flue gas. With selective noncatalytic reduction, the ammonia liquid is injected in the cooler regimes of the combustor to reduce NO_x to nitrogen and oxygen, achieving reductions of NO_x in the range 30–60%. Injecting ammonia in front of a catalyst system designed for control of NO_x can remove up to 90% of NO_x. This selective catalytic reduction control system is applicable to practically all sources of NO_x, and it is generally the highest-capital-cost system for such control.

Oxidation–absorption of NO_x is another alternative where a gas absorber is already installed or simultaneous SO_x/NO_x control is desirable. In this technique, an oxidant is added to the flue gas, such as ozone or chloric acid, or absorber liquor (sodium chlorate, for example) that causes NO_x to be oxidized to water-soluble nitric acid or nitrous acid vapor and absorbed in the scrubber liquor.

• Hydrogen chloride. Control of HCl has been achieved by adapting postcombustion SO_x control systems to combined HCl/SO_x removal systems. Wet scrubbers, spray dryers, and dry sorbent injection are applicable to waste combustors. In the United States, spray dryer scrubbers are a popular system for controlling HCl, typically removing 90–95% HCl and 70–90% SO_x.

Toxic gases and vapors (mercury). In the December 1997 report on mercury to Congress, coal-fired boilers represented over 50% of the total anthropogenic mercury emissions in the United States, estimated at 144 megagrams per year. The control of mercury is complicated by several factors, the most prominent being the very low concentrations in flue gas and the species present. Mercury vapor occurs in coal combustor flue gas at concentrations ranging from 3 to 11 $\mu g/nm^3$ as elemental (Hg^0) form or ionic mercuric chloride ($HgCl_2$) or mercuric oxide (HgO) form. Elemental mercury is insoluble in water and therefore not easily captured by wet or semidry scrubbing technology. $HgCl_2$ is water-soluble and readily absorbed in alkaline scrubbing liquor. Little is known about the behavior of HgO in flue gases.

Activated carbon has been used in the United States to reduce mercury emissions from municipal waste incinerators, but has not been commercially applied to coal-fired boilers. Substantial differences in flue gas species and concentration make transfer of technology impractical. Wet and dry flue gas desulfurization (FGD) units have been shown to be effective in removing oxidized mercury species.

Sulfuric acid mist. Since 1998, fossil fuel–fired power plants are required to report annual release inventories of nearly 600 toxic chemicals covered under the national Toxic Release Inventory, established under Section 313 of the Emergency Planning and Community Right-to-Know Act of 1986. This generally applies to any release exceeding 11,340 kg (25,000 lb) per year. Unwittingly power plants have become manufacturers of several toxic agents by virtue of their formation in the combustion of fuel and subsequent cooling and treating of exhaust flue gas. One recent study estimates that of a potential 17 metal species, 12 organic compounds, and 13 other acids and oxidants, a 650-MWe power plant burning a coal of 2.5 lb/MBtu sulfur content can be expected to report on 10 of these, with major components being hydrochloric acid (controls discussed above) and sulfuric acid (condensed sulfur trioxide). Primary particle collectors operating above the acid dew point can remove solid sulfates that have reacted with alkaline compounds such as lime, sodium, or ammonia; therefore any injection of these compounds for other reasons can directly reduce sulfuric acid (H_2SO_4) emissions. Some control is likely to occur by sorption or condensation of acids on collectible fly ash particles. Upstream NO_x controls may also influence H_2SO_4 emissions—catalytic systems may oxidize SO_2 to SO_3 and increase potential emissions, while ammonia slip from SCR and SNCR will likely react with SO_3

to form ammonium sulfate and bisulfate particles, effectively reducing H_2SO_4 emissions at the stack. The wet scrubber for SO_2 control ironically exacerbates SO_3 problems because (1) the particle collector is upstream and sees SO_3 as a vapor; (2) the scrubber cools the gas, converting SO_3 vapor to H_2SO_4 aerosol; and (3) the scrubber is designed for gas absorption only, allowing the submicrometer H_2SO_4 particles to exit, largely uncontrolled. One study indicates that only 15% of the SO_3 entering the FGD system is removed by the scrubbing process. Conversely, the so-called dry systems that scrub acid gases upstream of the particle collector seem particularly adept at converting SO_3 to solids that are effectively collected in the dust collector as particles. In severe cases of wet FGD-induced H_2SO_4 stack plumes, a wet electrostatic precipitator has been installed between the FGD system and stack, or the mist eliminator inside the FGD has been replaced by a wet electrostatic precipitator, to reduce plume opacity to acceptable levels.

Multipollutant controls. The concept of multipollutant controls is to develop a compact system that removes many pollutants in a minimum of process steps, and is upgradable to meet future emission codes. One approach is to convert all pollutant vapors and gases to particles upstream of the particle collector, thereby collecting all gaseous pollutants as condensed or sorbed particles. The class of circulating bed absorbers offers such an opportunity for multiple-pollutant control. Special sorbents, such as activated carbon, zeolites, or calcium silicates for sequestering toxic metals and organics, may be easily added to the base lime or sodium stream, and the large effective solids residence time allows for better utilization of sorbent. Oxidants for converting NO_x into more reactive acid vapors, which would cause potential wet chemistry problems in wet scrubbers, would not present such problems in dry solids absorption; hence simultaneous removal of SO_x, NO_x, and toxic vapors by a multipollutant sorbent or mix of sorbents introduced into a rapidly cooling gas stream, typical of that in a circulating bed, followed by an efficient particle collector, would be possible.

Alternative technologies. In the United States, the focus was upon clean energy in the late 1980s and early 1990s. The U.S. Clean Coal Technology Program administered by the Department of Energy emphasized the concept of repowering as opposed to construction of conventional coal-fired boilers. The concept of repowering includes various configurations of advanced power cycles that use coal gasification, fluidized bed combustion, gas turbines, and heat recovery boilers to achieve higher efficiencies in conversion of coal to electric power. These configurations avoid conditions favorable for the formation of NO_x normally associated with coal combustion. The amount of sulfur oxides is also much reduced, since coal gasification converts fuel sulfur (sulfur oxides formed by the oxidation of sulfur compounds in fuel) to hydrogen sulfide, 99% of which is typically removed by absorption and conversion to liquid sulfur.

Charles B. Sedman

Legal issues

Industry often has little economic incentive to voluntarily install pollution controls, and private litigation has often not addressed what is normally a communitywide problem. Thus, many countries have established schemes of legislative and administrative regulation. In the United States, the Clean Air Act directs the EPA to establish national ambient air quality standards that limit the permissible concentration of air pollutants at any outdoor point to which the public has access. There are two types of ambient air quality standards: primary and secondary (Table 1). Primary standards must protect the public health with an adequate margin of safety. Secondary standards must safeguard public welfare from known or anticipated adverse effects. (Generally, as is shown in Table I, primary and secondary standards have been set at the same levels.) The EPA may not take compliance costs into account in setting ambient air quality standards. In 1997, the EPA tightened the ambient air quality standards for particulate matter and ozone. New standards have been set to control particulate matter with an effective aerodynamic diameter of 2.5 μm or less ($PM_{2.5}$). The ozone standard was revised to regulate concentrations over an 8-h averaging time to reflect scientific findings that health effects are associated with exposures over an extended time.

Each state is required to submit for EPA approval a State Implementation Plan (SIP). This must consist of measures, such as limitations on emissions from individual sources of pollution, sufficient to demonstrate that the state will attain and maintain the primary standards by specific deadlines and the secondary standards as expeditiously as practicable. Each state has a great deal of discretion in allocating the reduction burden among sources as long as the state shows that its SIP plan will result in timely attainment and maintenance of the standards. In addition, each state's plan must be adequate to prevent the transport of air pollution across state lines that would significantly interfere with another state's ability to meet the standards. The EPA has become more aggressive in implementing this prohibition.

Each SIP must include programs for the review of the proposed construction or modification of major sources of air pollution, both in areas in which the ambient air quality standards are violated (nonattainment areas) and in clean-air areas. Under both programs, a source may receive a permit to construct or modify only if it uses the best pollution controls available. The program in nonattainment areas requires that the permit applicant furnish offsets—decreases in air pollution that go beyond what is required by the SIP—to compensate for the new or modified source's effect on air quality. The program in clean-air areas, known as prevention of significant deterioration, establishes increments that, in effect, cap the growth of concentrations of sulfur dioxide,

particulate matter, and nitrogen oxide in clean areas. These increments are especially stringent in national parks and wilderness areas. There is much controversy about which changes at existing sources constitute modifications that are subject to the programs.

Some kinds of sources are regulated directly by the federal government. For example, the federal government establishes emission limits for new cars and trucks and regulates the content of automotive and truck fuel. These regulations have virtually eliminated lead emissions in the United States. The federal government also regulates sources of hazardous air pollutants that pose a risk of cancer or other life-threatening disease. The 1990 Amendments to the Clean Air Act list 189 substances as air pollutants, subject to addition or deletion by the EPA. The EPA must establish emissions standards that require the use of maximum available control technology by major sources of such pollutants, and must set additional, stricter standards for a substance if the agency finds that, even with maximum available control technology, a major source poses a lifetime risk of more than one in a million to the most exposed individual. These additional standards must protect health with an ample margin of safety.

The 1990 Amendments established an innovative program to decrease emissions of sulfur dioxide, which turn into sulfates that degrade visibility, increase concentrations of health-threatening fine particles, and cause acid rain. The Amendments require a 10-million-ton reduction in annual emissions of sulfur dioxide and establish annual allowances for individual sources. These sources may trade allowances among themselves so that the reduction can be accomplished in the most cost-effective way. The EPA is encouraging states to use the same approach in its new Clean Air Implementation Rule. This rule requires 28 eastern and midwestern states to further reduce sulfur dioxide and nitrogen oxide emissions that cross state lines and interfere with the attainment of the new air quality standards for ozone and particulate matter. Craig N. Oren

Bibliography. R. W. Boubel et al., *Fundamentals of Air Pollution*, 3d ed., 1994; W. T. Davis (ed.), *Air Pollution Engineering Manual*, 2d ed., 2000; R. P. Donovan, *Fabric Filtration for Combustion Sources*, 1985; M. Lippmann (ed.), *Environmental Toxicants: Human Exposures and Their Health Effects*, 2d ed., 2005; L. R. Paulson (ed.), *Human Exposure Assessment for Airborne Pollutants: Advances and Opportunities*, 1991; K. R. Parker (ed.), *Applied Electrostatic Precipitation*, 1996; K. C. Schifftner, *Air Pollution Control Equipment Selection Guide*, 2002; J. H. Seinfeld and S. N. Pandis, *Atmospheric Chemistry and Physics: From Air Pollution to Climate*, 2d ed., 2006; J. D. Spengler et al., *Indoor Air Quality Handbook*, 2000; L. K. Wang et al. (eds.), *Air Pollution Control Engineering*, 2004; K. Wark et al., *Air Pollution: Its Origin and Control*, 3d ed., 1997; A. Wellburn, *Air Pollution and Climate Change*, 2d rev. ed., 1994; Y. Zhang, *Indoor Air Quality Engineering*, 2004.

Air pollution, indoor

The presence of gaseous and particulate contaminants in the indoor environment. Most pollution is due to human sources, although natural sources do exist, including plants, animals, and other living organisms and bodies of water that release various chemical aerosols. Contamination can occur from infiltration indoors of atmospheric pollutants generated outdoors, and thus the indoor environment is affected by meteorological conditions.

Contamination of indoor air is considered very important because of the amount of time that people spend indoors (65–95%), and the effects of indoor pollutants on humans, pets, and materials. It was not until the twentieth century that serious scientific measurements were taken and studies of air pollution done. Later developments included investigations of indoor sources and effects. Another concern has been the amount of human cancer caused by airborne toxins and radon from both indoor and outdoor sources. Research involving energy-saving measures, human activities, and consumer products has led to evaluation of the health effects of the indoor environment specifically, as well as of the total air pollution.

Sources. Natural sources of air pollution are soils and water that release radon decay products, volatile organic compounds, fungi, and so on, and living organisms that release allergens (for example, dander from pets). Human-produced (anthropogenic) chemical releases (emissions) originate from various types of appliances, combustion sources (including vehicle combustion engines), and building materials. There has been an increase in the home use of secondary combustion heating sources, such as wood-burning and gas and kerosene space heaters (often unvented) due to increased costs of regular heating, and they are very important contributors to high levels of indoor air pollutants. Carbon dioxide (CO_2) is generated both by combustion and by living organisms. Other pollutants include molds, microorganisms (such as bacteria and viruses), and pollen from outdoor and indoor plants (**Table 1**). Human activity has been noted to increase indoor air contamination by all of these. Some pollutants are generated by individual behavior, such as tobacco smoking (environmental tobacco smoke, or ETS), or through use of consumer products.

Classification and behavior of pollutants. The chemicals in pollutants can be divided into inorganic and organic, and can be further classified on the basis of their thermal and hydroscopic characteristics. Particles are characterized also by functional size (usually mass-median aerodynamic diameter), density, chemical identity, water content, and affinity for absorption and adsorption. The mass-median aerodynamic diameter is the measurement of size and density of particles (including viable and nonviable organic particles); it is a major characteristic of particles and determines their behavior (dispersal, deposition), especially their likelihood of inhalation and

TABLE 1. Major sources and types of indoor pollutants

Sources	Pollutants
Fossil fuel combustion (natural gas and kerosene appliances); wood combustion (wood stoves, fireplaces)	Particulates (PM); nitrogen oxides (NO_x); carbon monoxide and dioxide (CO, CO_2); lead and other trace metals; polynuclear aromatic hydrocarbons (PNAs) and other volatile organic compounds (VOCs)
Tobacco smoking	PM; CO; CO_2; NO_x; PNA; VOCs; radon progeny
Building and furnishing materials	PNAs (especially aldehydes) and other VOCs; PM; radon progeny; also act as reservoirs for molds and other allergens
Water reservoirs (fixtures; air conditioning/cleaning/treating)	Mold; *Bacillus* sp. and other bacteria
Consumer products	PNAs and other VOCs; trace metals
Other product examples	Cleaners, solvents, disinfectants, cosmetics, air purifiers, etc.
Animals (pets and opportunistic dwellers) and plants	Allergens; CO_2
Infiltration	PM; NO_x; sulfur oxides (SO_x); pollen; mold

deposition in different parts of the respiratory system. For instance, environmental tobacco smoke produces very small (fine) particles, and intense combustion can produce ultrafine particles, both of which are more deeply inhaled than larger particles. Sometimes, gaseous pollutants are further divided into vapor and nonvapor phase, and degree of volatility. *See* PARTICULATES.

Nitrogen oxide (NO_x) can be produced as a gas [nitric oxide (NO) or nitrogen dioxide (NO_2)], as particles [nitrate (NO_3^-)], or in aerosol form [such as nitrous acid (HNO_2)]. Radon decay products are more important when adsorbed on particles, as are aldehydes, other volatile organic compounds, and compounds derived from environmental tobacco smoke; all of these are normally gaseous, but all except radon progeny can occur in the form of particles. *See* NITROGEN OXIDES; RADON.

The generation and behavior of pollutants in enclosed environments are also affected by most meteorological factors. Indoor environments are often sealed to the extent that the weather is cold or hot. Infiltration of outdoor pollutants is affected by climate, temperature, humidity, barometric pressure, and wind speed and direction. The small particles and gases that infiltrate most—NO_2 and volatile organic compounds—are affected by outdoor concentrations, tightness of the indoor space (being sealed with little air exchange), operation of heating and cooling systems, convection currents, full growth (or lack thereof) of local trees and shrubs, and so forth. Indoor environments also possess individual characteristics of ventilation, dispersion, and deposition. Some gases, for example sulfur dioxide (SO_2) and ozone (O_3), infiltrating the indoor environment are absorbed readily if absorbent materials (such as carpets, drapes, cloth-covered furniture, clothes) are present. These gases thus exist in high concentrations only when the outdoor concentrations are very high or such materials are not present. *See* OZONE.

Exposure assessment. Chemical-physical mass-balance models are used to estimate pollutant concentrations, as are statistical models. Assessments of exposures provide information on the concentration and spatial-temporal distribution of each pollutant and are related to spatial-temporal activities of the people assessed. Personal exposure factors include time spent in different indoor environments (**Table 2**), activities and behavior, breathing rates, and so forth. These factors indicate that individuals experience pollution levels very different from those measured at a nearby fixed monitoring station. The pollutant benzene provides a good example (**Fig. 1**). Personal exposure must also take into account the major indoor sources, such as environmental tobacco smoke, which produces particulate matter leading to concentrations that frequently occur in excess of ambient standards (set by law in the United States). Lower socioeconomic status groups not only live in areas with higher ambient exposures, but also live in poorly ventilated homes and use more inappropriate combustion sources having higher smoking levels.

Biological markers of exposures to pollutants have been developed for a variety of toxic materials in the air. Lead is measured by well-established blood assays. These tests have indicated declines of lead in the body that parallel the declines in amount of lead in the atmosphere, the amount of lead in gasoline, and the amount of lead in paint and dust that were mandated by law in the United States. Cotinine, a nicotine metabolite, is the marker used for environmental tobacco smoke. Using metal biomarkers as estimates of exposures to elemental metals (air exposures mostly from smaller particles) in the National Human Exposure Assessment Survey (NHEXAS), researchers find that children have higher levels of

TABLE 2. Time activity patterns*

Location	Percentage of time spent by age group, years		
	<5	6–65	>65
Home (inside)	79	63	78
Work (inside)	—	11	1
Other (inside)	8	7	5
Home (outside)	6	5–6	8
Work (outside)	—	3	1
Other (outside)	3	4–5	3
In transit	4	6	4

*From Arizona National Human Exposure Assessment Survey.

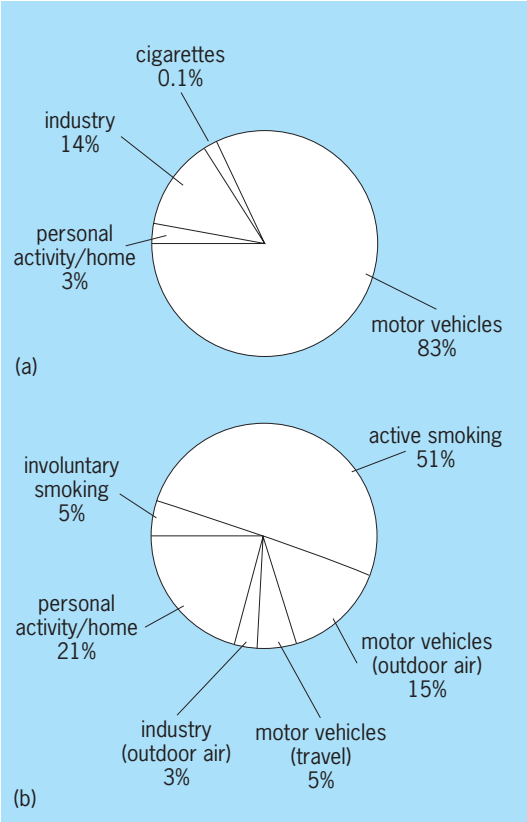

Benzene emissions versus exposures. (*a*) Emissions. (*b*) Exposures. (*From L. Wallace, Enviromental Protection Agency, 1989; National Research Council, Human Exposure Assessment for Airborne Pollutants, 1991*)

cadmium, nickel, chromium, and lead if they have been exposed to environmental tobacco smoke. This survey found also that children had higher exposures than adults to pesticides.

Indoor monitoring has specific requirements related to the type of pollutants and enclosed spaces monitored. Monitoring protocols are developed with regard to the availability, practicality, and expense of continuous or integrated sampling methods to measure the pollutant over the periods of interest. These monitors are usually fixed or portable. Personal monitors have been developed for some pollutants. Good surrogate measurements are often utilized to represent more complex exposures (for example, total hydrocarbons for all organics).

Effects. Characterization of the effects of air pollution requires determination of concentration/exposure, exposure/dose (for humans and animals), and exposure/dose/response relationships. These studies have involved humans, animals, and materials.

Humans. Effects experienced by humans include acute and chronic symptoms (morbidity), increases in acute respiratory illness, declines in lung function and other physiological and immunological changes, and deaths, especially in those with chronic diseases. The risks of multiple exposures may contribute additively or synergistically to health consequences. There are several indoor pollutants that have similar, and possibly synergistic, effects: particulate matter, NO_2, and formaldehyde (HCHO) produce irritation of eyes and mucous membranes; various types of particulate matter, HCHO, and allergens produce allergic symptoms; NO_2, particulate matter, and HCHO produce changes in pulmonary function; carbon monoxide (CO), NO_2, and HCHO affect cognitive skills; CO, NO, NO_2, nitrates, and methylene chloride (CH_2Cl_2) change carboxyhemoglobin and methemoglobin levels and affect the heart and brain; and radon progeny, especially when associated with tobacco smoke, can promote lung cancer. Lead has been shown to have neurological effects at concentrations that occurred frequently in the past, especially in infants and children. Benzene has been shown to produce a form of leukemia, and both radon progeny and asbestos have been shown to produce lung cancer in humans. Other types of cancers may occur with exposures to carcinogens found in occupational settings. *See* ASBESTOS; BENZENE; MUTAGENS AND CARCINOGENS; RESPIRATORY SYSTEM DISORDERS.

Aeroallergens are also of great importance, as many people have allergic problems, including asthma. It is likely that aeroallergens and interactive exposures with other pollutants can aggravate allergies and produce respiratory diseases. Thus, a large amount of acute disability is related to indoor pollutant exposures. *See* ALLERGY; ASTHMA.

Animals. Sudies of pollutant effects on animals have shown effects similar to those found in humans. These studies used mortality, lung pathology, and changes in pulmonary physiology and immunology as criteria. Animal studies of cancer have provided the evidence to protect humans from those carcinogens.

Materials. Indoor air pollution produces destruction by corrosion. In addition, outdoor pollutants can infiltrate and interact with indoor pollutants to produce new compounds that are damaging to materials; an example is O_3 interacting with indoor volatile organic compounds to produce corrosive compounds affecting electronic equipment. Economically significant damage has occurred in various metals, ferrous (iron-based) and nonferrous (copper, silver, nickel, aluminum, zinc). Extensive damage to rubber, paint and dyes, leather and textiles, and ceramics has been demonstrated. *See* CORROSION.

Controls. For some pollutants, control measures used to protect health are based on regulatory law. Unfortunately, the experience gained in developing and implementing strategies for the outdoor environment is not very applicable to indoor environments, and new strategies are needed. Only a few countries have regulations for indoor pollutants, specifically radon progeny, formaldehyde, and CO_2. Controls should include, where necessary, practices, labeling requirements, building codes, guidelines, and recommendations.

ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) standards are used in the United States. Some European countries and Japan also have such standards, and World Health Organization (WHO) publications have recommendations for control as well. Primary control measures

include removal of sources (including water, plants, pets, and certain building materials) or reductions in source emissions through manufacturing processes, changes in building materials, prevention of water leakage, and source usage recommendations. Source removal is the first and most logical form of control. Reductions in source emissions have been achieved for several products through changes in manufacturing process or construction. Controls have included recommendations for source placement, use, and ventilation (including the amount of fresh air intake and filtration). Reduction of exposure also includes standard direct methods such as filtration and cleaning, as well as good building and source design. Appropriate building maintenance is necessary. Ventilation that produces adequate exfiltration of contaminated air is very efficient. Filtration is useful but can be expensive. Active charcoal can be used for filtering most gases, and HEPA filters can be used to filter most particle matter, but they need to be monitored and maintained appropriately.

Governmental agencies can control product manufacturing, promote ventilation in workplaces, use public health education to help people minimize risks and promote safe indoor environments, and develop regulations (such as banning tobacco smoking in workplaces and public buildings). Laws can be passed to prevent exposures in public locations. Finally, individuals can limit their exposures, especially in residences, by several measures, including avoidance. *See* AIR POLLUTION; ENVIRONMENTAL TOXICOLOGY.

Michael D. Lebowitz

Bibliography. American Society for Heating, Refrigeration and Air Conditioning Engineers (ASHRAE), Standards for heating, cooling and ventilation, ASHRAE Web site, National Human Exposure Assessment Survey (NHEXAS), *J. Expos. Anal. Environ. Epidemiol.*, vol. 9, 1999; R. Bertollini et al. (eds.), *Environmental Epidemiology: Exposure and Disease*, CRC/Lewis, London, 1996; M. Lippmann (ed.), *Environmental Toxicants*, 2d ed., Wiley, New York, 2000; National Research Council (NRC), *Human Exposure Assessment for Airborne Pollutants*, National Academy Press, Washington, DC, 1991; NRC, *Indoor Allergens*, National Academy Press, Washington, DC, 1993; NRC, *Policies and Procedures for Control of Indoor Air Quality*, National Academy Press, Washington, DC, 1987.

Air pressure

The force per unit area that the air exerts on any surface in contact with it, arising from the collisions of the air molecules with the surface. It is equal and opposite to the pressure of the surface against the air, which for atmospheric air in normal motion approximately balances the weight of the atmosphere above, about 15 lb/in.^{2} at sea level. It is the same in all directions and is the force that balances the weight of the column of mercury in the torricellian barometer, commonly used for its precise measurement. *See* BAROMETER.

Units. The units of pressure traditionally used in meteorology are based on the bar, defined as equal to 1,000,000 dynes /cm^2. One bar equals 1000 millibars or 100 centibars.

In the meter-kilogram-second or International System of Units (SI), the unit of force, the pascal (Pa), is equal to 1 newton/m^2. One millibar equals 100 pascals. The normal pressure at sea level is 1013.25 millibars or 101.325 kilopascals.

Also widely used in practice are units based on the height of the mercury barometer under standard conditions, expressed commonly in millimeters or in inches. The standard atmosphere (760 mmHg) is also used as a unit, mainly in engineering, where large pressures are encountered. The following equivalents show the conversions between the commonly used units of pressure, where $(\text{mmHg})_n$ and $(\text{in. Hg})_n$ denote the millimeter and inch of mercury, respectively, under standard (normal) conditions, and where $(\text{kg})_n$ and $(\text{lb})_n$ denote the weight of a standard kilogram and pound mass, respectively, under standard gravity.

$$\begin{aligned}
1\ \text{kPa} &= 10\ \text{millibars} = 1000\ \text{N/m}^2 \\
&= 7.50062\ (\text{mmHg})_n \\
&= 0.95300\ (\text{in. Hg})_n \\
1\ \text{millibar} &= 100\ \text{Pa} = 1000\ \text{dynes/cm}^2 \\
&= 0.750062\ (\text{mmHg})_n \\
&= 0.0295300\ (\text{in. Hg})_n \\
1\ \text{atm} &= 101.325\ \text{kPa} = 1013.25\ \text{millibars} \\
&= 760\ (\text{mmHg})_n = 29.9213\ (\text{in. Hg})_n \\
&= 14.6959\ (\text{lb})_n/\text{in.}^2 \\
&= 1.03323\ (\text{kg})_n/\text{cm}^2 \\
1\ (\text{mmHg})_n &= 1\ \text{torr} = 0.03937008\ (\text{in. Hg})_n \\
&= 1.333224\ \text{millibars} \\
&= 133.3224\ \text{Pa} \\
1\ (\text{in. Hg})_n &= 33.8639\ \text{millibars} \\
&= 25.4\ (\text{mmHg})_n \\
&= 3.38639\ \text{kPa}
\end{aligned}$$

Variation with height. Because of the almost exact balancing of the weight of the overlying atmosphere by the air pressure, the latter must decrease with height, according to the hydrostatic equation (1),

$$dP = -g\ \rho\ dZ \tag{1}$$

where P is air pressure, ρ is air density, g is acceleration of gravity, Z is altitude above mean sea level, dZ is infinitesimal vertical thickness of horizontal air layers, and dP is pressure change which corresponds to altitude change dZ. Integration of Eq. (1) yields Eq. (2), where P_1 is pressure at altitude Z_1, and P_2

$$P_1 - P_2 = \int_{Z_1}^{Z_2} \rho g\ dZ \tag{2}$$

is pressure at altitude Z_2. The expressions on the right-hand side of Eq. (2) represent the weight of the column of air between the two levels Z_1 and Z_2.

In the special case in which Z_2 refers to a level above the atmosphere where the air pressure is nil, one has $P_2 = 0$, and Eq. (2) yields an expression for air

pressure P_1 at a given altitude Z_1 for an atmosphere in hydrostatic equilibrium.

By substituting in Eq. (1) the expression for air density based on the well-known perfect gas law and by integrating, one obtains the hypsometric equations for dry air under the assumption of hydrostatic equilibrium, Eqs. (3), valid below about 54 mi (90 km),

$$\log_e\left(\frac{P_1}{P_2}\right) = \frac{M}{R}\int_{Z_1}^{Z_2}\frac{g}{T}dZ \qquad (3a)$$

$$Z_2 - Z_1 = \frac{R}{M}\int_{P_2}^{P_1}\frac{T}{g}\frac{dP}{P} \qquad (3b)$$

where g is the gravitational acceleration; M is the gram-molecular weight, 28.97 for dry air; R is the gas constant for 1 mole of ideal gas, or 8.315×10^7 erg/(mole)(K); and T is the air temperature in K.

Equation (3) may be used for the real moist atmosphere if the effect of the small amount of water vapor on the density of the air is allowed for by replacing T by T_v, the virtual temperature given by Eq. (4), in which e is partial pressure of water vapor

$$T_v = T\left[1 - \left(1 - \frac{M_w}{M}\right)\frac{e}{P}\right]^{-1} \qquad (4)$$

in the air, M_w is gram-molecular weight of water vapor (18.016 g/mole), and $(1 - M_w/M) = 0.3780$.

Equation (3a) is used in practice to calculate the vertical distribution of pressure with height above sea level. The temperature distribution in a standard atmosphere, based on mean values in middle latitudes, has been defined by international agreement. The use of the standard atmosphere permits the evaluation of the integrals of Eqs. (3a) and (3b) to give a definite relation between pressure and height. This relation is used in all altimeters which are basically barometers of the aneroid type. The difference between the height estimated from the pressure and the actual height is often considerable; but since the same standard relationship is used in all altimeters, the difference is the same for all altimeters at the same location, and so causes no difficulty in determining the relative position of aircraft. Mountains, however, have a fixed height, and accidents have been caused by the difference between the actual and standard atmosphere. *See* ALTIMETER.

Horizontal and time variations. In addition to the large variation with height discussed in the previous paragraph, atmospheric pressure varies in the horizontal and with time. The variations of air pressure at sea level, estimated in the case of observations over land by correcting for the height of the ground surface, are routinely plotted on a map and analyzed, resulting in the familiar "weather map" representation with its isobars showing highs and lows. The movement of the main features of the sea-level pressure distribution, typically from west to east in middle and high latitudes, and from east to west in the tropics, produces characteristic fluctuations of the pressure at a fixed point, which vary by a few percent within a few days.

Smaller-scale variations of sea-level pressure, some even too small to appear on the ordinary weather map, are also present. These are associated with various forms of atmospheric motion, including small-scale wave motion and turbulence. Relatively large variations in short distances are found in hurricanes and in intense winter storms, and in and near thunderstorms; the most intense is the low-pressure region in a tornado. The pressure drop within tornadoes and hurricanes can be extreme, about 10% of normal pressure. *See* ISOBAR (METEOROLOGY); WEATHER MAP.

It is a general rule that in middle latitudes at localities below 3280 ft (1000 m) in height above sea level, the air pressure on the continents tends to be slightly higher in winter than in other seasons; whereas at considerably greater heights on the continents and on the ocean surface, the reverse is true.

Various maps of climatic averages indicate certain regions where systems of high and low pressure predominate. Over the oceans there tend to be areas or bands of relatively high pressure, most marked during the summer, in zones centered near latitude 30°N and 30°S. The Asiatic landmass is dominated by a great high-pressure system in winter and a low-pressure system in summer. Deep low-pressure areas prevail during the winter over the Aleutian, the Icelandic-Greenland, and Antarctic regions. These and other centers of action produce offshoots which may travel for great distances before dissipating.

Thus during the winter, spring, and autumn in middle latitudes over the land areas, it is fairly common to experience the passage of a cycle of low- and high-pressure systems in alternating fashion over a period of about 6–9 days in the average, but sometimes in as little as 3–4 days, covering a pressure amplitude which ranges on the average from roughly

Mean atmospheric pressure and temperature in middle latitudes, for specified heights above sea level*

Altitude above sea level†			
Standard geopotential meters (m′)	Meters at latitude 45°32′40″	Air pressure, millibars	Assumed temperature, K‡
0	0	1.01325×10^3	288.15
11,000	11,019	2.2632×10^2	216.65
20,000	20,063	5.4747×10^1	216.65
32,000	32,162	8.6798×10^0	228.65
47,000	47,350	1.1090×10^0	270.65
52,000	52,429	5.8997×10^{-1}	270.65
61,000	61,591	1.8209×10^{-1}	252.65
79,000	79,994	1.0376×10^{-2}	180.65
88,743	90,000	1.6437×10^{-3}	180.65

*Approximate annual mean values based on radiosonde observations at Northern Hemisphere stations between latitudes 40 and 49° N for heights below 105,000 ft (32,000 m) and on observations made from rockets and instruments released from rockets. Some density data derived from searchlight observations were considered. Values shown above 105,000 ft (32,000 m) were calculated largely on the basis of observed distribution of air density with altitude. In correlating columns 1 and 2, the acceleration of gravity, G, is taken to be 98,066.5 cm^2/s^2 per standard geopotential meter (m′). Data on first three lines are used in calibration of aircraft altimeters. 1 m = 3.281 ft.

†Above 295,000 ft (90,000 m) there occurs an increase of temperature with altitude and a variation of composition of the air with height, resulting in a gradual decrease in molecular weight of air with altitude.

‡Temperature (°F) = 1.8 [temperature (K)] − 459.67.

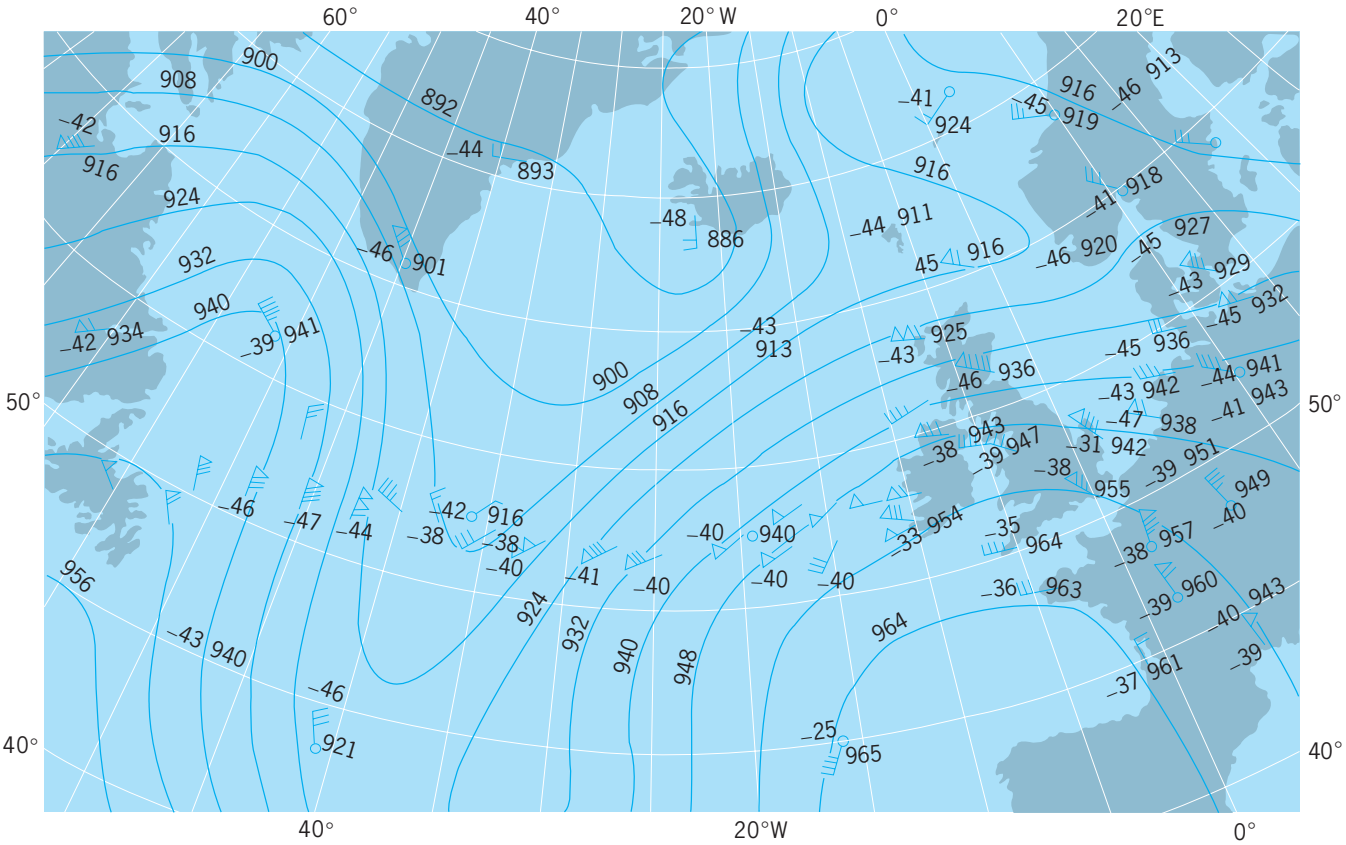

Contours of 300-millibar surface, in tens of meters, with temperature in °C [°F = (°C × 1.8) + 32°] and measured winds at the same level. Winds are plotted with arrow pointing in direction of the wind, with each bar of the tail representing 10 m/s (33 ft/s or 19.4 knots). Triangle represents 50 m/s (165 ft/s or 97 knots).

15–25 millibars less than normal in the low-pressure center to roughly 15–20 millibars more than normal in the high-pressure center. During the summer in middle latitudes the period of the pressure changes is generally greater, and the amplitudes are less than in the cooler seasons (see **table**).

Within the tropics, where there are comparatively few passages of major high- and low-pressure systems during a season, the most notable feature revealed by the recording barometer (barograph) is the characteristic diurnal pressure variation. In this daily cycle of pressure at the ground there are, as a rule though with some exceptions, two maxima, at approximately 10 A.M. and 10 P.M., and two minima, at approximately 4 A.M. and 4 P.M., local time.

The total range of the diurnal pressure variation is a function of latitude as indicated by the following approximate averages (latitude N): 0°, 3 millibars; 30°, 2.5 millibars; 35°, 1.7 millibars; 45°, 1.2 millibars; 50°, 0.9 millibar; 60°, 0.4 millibar. These results are based on the statistical analysis of thousands of barograph records for many land stations. Local peculiarities appear in the diurnal variation because of the influences of physiographic features and climatic factors. Mountains, valleys, oceans, elevations, ground cover, temperature variation, and season exert local influences; while current atmospheric conditions also affect it, such as amount of cloudiness, precipitation, and sunshine. Mountainous regions in the western United States may have only a single maximum at about 8–10 A.M. and a single minimum at about 5–7 P.M., local time, but with a larger range than elsewhere at the same latitudes, especially during the warmer months (for instance, about 4 millibars difference between the daily maximum and minimum). *See* ATMOSPHERIC TIDES.

At higher levels in the atmosphere the variations of pressure are closely related to the variations of temperature, according to Eq. (3*a*). Because of the lower temperatures in higher latitudes in the lower 6 mi (10 km), the pressures at higher levels tend to decrease toward the poles. The **illustration** shows a typical pattern at approximately 6 mi (10 km) above sea level. As is customary in representing pressure patterns at upper levels, the variation of the height of a surface of constant pressure, in this case 300 millibars, is shown rather than the variation of pressure over a horizontal surface. *See* AIR TEMPERATURE.

Besides the latitudinal variation, the illustration also shows a wave pattern typical of the pressure field, and the midlatitude maximum in the wind field known as the jet stream, with its "waves in the westerlies." In the stratosphere the temperature variations are such as to reduce the pressure variations at higher levels, up to about 50 mi (80 km), except that in winter at high latitudes there are relatively large variations above 6 mi (10 km). At altitudes above 50 mi (80 km) the relative variability of the pressure increases again. Although the pressure and density at these very high levels are small, they are important

for rocket and satellite flights, so that their variability at high altitudes is likewise important.

Relations to wind and weather. The practical importance of air pressure lies in its relation to the wind and weather. It is because of these relationships that pressure is a basic parameter in weather forecasting, as is evident from its appearance on the ordinary weather map.

Horizontal variations of pressure imply a pressure force on the air, just as the vertical pressure variation implies a vertical force that supports the weight of the air, according to Eq. (1). This force, if unopposed, accelerates the air, causing the wind to blow from high to low pressure. The sea breeze is an example of such a wind. However, if the pressure variations are on a large scale and are changing relatively slowly with time, the rotation of the Earth gives rise to geostrophic or gradient balance such that the wind blows along the isobars. This situation occurs when the pressure variations are due to the slow-moving lows and highs that appear on the ordinary weather map, and to the upper air waves shown in the illustration, in which the relationship is well illustrated. *See* CORIOLIS ACCELERATION; GEOSTROPHIC WIND.

The wind near the ground, in the lowest few hundred meters of the atmosphere, is retarded by friction with the surface to a degree that depends on the smoothness or roughness of the surface. This upsets the balance mentioned in the previous paragraph, so that the wind blows somewhat across the isobars from high to low pressure.

The large-scale variations of pressure at sea level shown on a weather map are associated with characteristic patterns of vertical motion of the air, which in turn affect the weather. Descent of air in a high heats the air and dries it by adiabatic compression, giving clear skies, while the ascent of air in a low cools it and causes it to condense and produce cloudy and rainy weather. These processes at low levels, accompanied by others at higher levels, usually combine to justify the clear-cloudy-rainy marking on the household barometer. Raymond J. Deland; Edwin Kessler

Bibliography. R. A. Anthes, *Meteorology*, 7th ed., 1996; R. G. Fleagle, *An Introduction to Atmospheric Physics*, 2d ed., 1980; D. D. Houghton (ed.), *Handbook of Applied Meteorology*, 1985; A. Miller and R. A. Anthes, *Meteorology*, 4th ed., 1980.

Air register

A device attached to an air-distributing duct for the purpose of discharging air into the space to be heated or cooled. These openings are referred to as registers, diffusers, supply outlets, or grills. By common acceptance, a register is an opening provided with means for discharging the air in a confined jet, whereas a diffuser is an outlet which discharges the air in a spreading jet (see **illus.**). Both registers and diffusers may be placed at a number of locations in a room, including the floor, baseboard, low or high on the sidewall, window sill,or ceiling.

For heating, the preferred location is in the floor, at the baseboard, or at the low sidewall of the outside wall, preferably under a window. For cooling, the preferred location is high on the inside wall or the ceiling. For year-round air conditioning in homes, a compromise location is the floor, baseboard, or low sidewall at the exposed wall, especially if adequate air velocity in an upward direction is provided at the supply outlet.

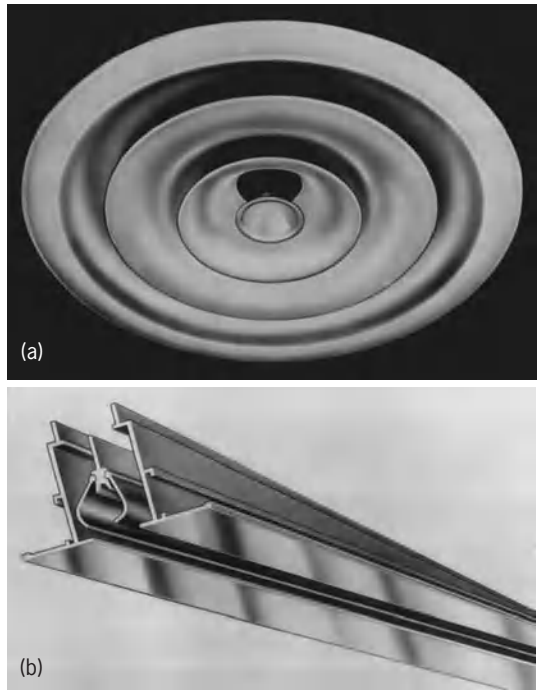

Some of the more common diffusers. (*a*) Round ceiling diffuser. (*b*) Single-slot ceiling diffuser.

A well-designed register effectively conceals the hole at the end of the duct, throws or projects the air in the direction and at the distance desired, limits the velocity usually to 1000 ft/min (300 m/min) or slower, and deflects the air away from walls and obstructions. The register also adjusts the direction of airflow to provide on-the-spot manipulation of the airstream, and adjusts the airflow rate to lesser amounts. It should accomplish these functions without producing dust streaks on nearby walls and ceilings, excessive air noise, or large pressure losses. Many registers, diffusers, slots, and air panels are commercially available and satisfy a majority of these qualifications. *See* OIL BURNER; WARM-AIR HEATING SYSTEM. Seichi Konzo

Bibliography. American Society of Heating, Refrigerating, and Air Conditioning Engineers, *Handbook and Product Directory*: *Equipment*, 1979; N. C. Harris, *Modern Air Conditioning Practice*, 3d ed., 1983.

Air separation

Separation of atmospheric air into its primary constituents. Nitrogen, oxygen, and argon are the primary constituents of air. Small quantities of neon, helium, krypton, and xenon are present at constant concentrations and can be separated as products. Varying quantities of water, carbon dioxide,

Composition of dry air

Component	Percent by volume	Component	Parts per million by volume
Nitrogen	78.084	Carbon dioxide	350–400
Oxygen	20.946	Neon	18.2
Argon	0.934	Helium	5.2
		Krypton	1.1
		Xenon	0.09
		Methane	1–15
		Acetylene	0–0.5
		Other hydrocarbons	0–5

hydrocarbons, hydrogen, carbon monoxide, and trace environmental impurities (sulfur and nitrogen oxides, chlorine) are present depending upon location and climate. Typical quantities are shown in the **table**. These impurities are removed during air separation to maximize efficiency and avoid hazardous operation. *See* AIR; ARGON; HELIUM; KRYPTON; NEON; XENON.

Three different technologies are used for the separation of air: cryogenic distillation, ambient temperature adsorption, and membrane separations. The latter two have evolved to full commercial status. Membrane technology is economical for the production of nitrogen and oxygen-enriched air (up to about 40% oxygen) at small scale. Adsorption technology produces nitrogen and medium-purity oxygen (85–95% oxygen) at flow rates up to 100 tons/day. The cryogenic process can generate oxygen or nitrogen at flows of 2500 tons/day from a single plant and make the full range of products.

Cryogenic air separation. The process of cryogenic distillation takes place at liquid-air temperature and has several steps: air compression; purification to remove water, carbon dioxide, and hydrocarbons; heat exchange between cooling air and warming products; distillation; and turbine expansion to provide refrigeration (**Fig. 1**). Products may be obtained as cryogenic liquids by incorporating additional refrigeration equipment. *See* CRYOGENICS; LIQUEFACTION OF GASES.

Air compression and purification. Air is compressed in a multistage centrifugal compressor to about 90 lb/in.^{2} absolute pressure (psia; 6.2 bars). The heat of compression is removed by water cooling. Condensed water is removed in a separator vessel. Water vapor, carbon dioxide, and hydrocarbons except for methane are then removed by adsorption onto solid adsorbent pellets of alumina or molecular sieves. These impurities would otherwise condense and accumulate in the cold process equipment. The adsorbent beds are regenerated at intervals by purging them with dry, hydrocarbon-free, waste gas from the distillation process.

Cooling and separation. The purified air is cooled to near liquefaction temperature in a heat exchanger, where cold product and waste streams are simultaneously rewarmed. The cold air enters the bottom of a multiple-column distillation system, where it separates into oxygen, nitrogen, and argon products and anitrogen-rich waste stream. The vapor in

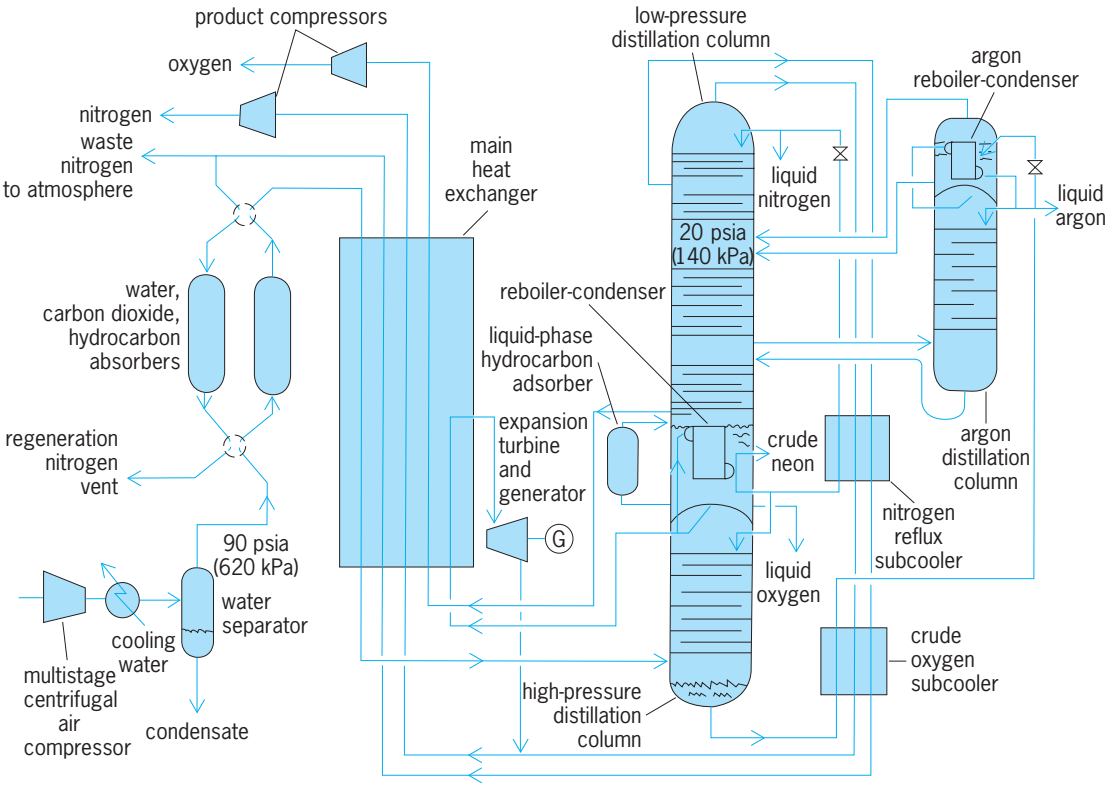

Fig. 1. Cryogenic air separation process. Air is compressed, purified, and cooled to liquefaction temperature; it is then distilled in a series of three columns to separate oxygen, nitrogen, and argon products.

equilibrium with a boiling liquid-air mixture contains an increased proportion of nitrogen. This principle is used to progressively enrich nitrogen in vapor and oxygen in liquid by successively contacting the streams in a series of stages. Thus as the air vapor rises through a sequence of liquid contact stages in the lower distillation column, it is depleted in oxygen until a pure nitrogen gas is obtained.

A small part of this nitrogen flows to the main heat exchanger, where it warms and expands through a turbine to provide process refrigeration. The rest of the nitrogen condenses to liquid in the reboiler-condenser, where oxygen is simultaneously boiled in the upper distillation column.

The condensed nitrogen divides into two streams. One stream passes to the upper column, and the other flows down the lower column and is enriched to an oxygen content of approximately 35%. This oxygen-rich stream is reduced in pressure and partly vaporized in a second reboiler-condenser that condenses argon column overhead vapor. It is then fed to the upper distillation column, which operates at a pressure of about 20 psia (1.4 bars). The difference in pressure between the upper and lower columns is determined by the temperature difference (about 2°F or 1°C) between the condensing nitrogen vapor and boiling oxygen liquid. In the upper column, the oxygen and nitrogen are separated to produce liquid oxygen at the bottom (purity > 99.5%) and nitrogen gas at the top.

The oxygen stream is boiled in the reboiler-condenser to provide both product gas and the reboil vapor for the upper column. A small stream of the liquid oxygen is circulated continuously through a silica gel adsorption bed to remove traces of acetylene, which might otherwise accumulate to hazardous levels in the oxygen reboiler. Nitrogen at the column top is purified by the reflux liquid nitrogen stream from the lower column. A second (waste) nitrogen stream containing a small percentage of oxygen is removed a few trays below. Argon has a boiling temperature between that of oxygen and nitrogen, and thus it accumulates in the middle of the distillation column. The argon-enriched stream (10% argon) is removed and is distilled in a third column to a purity of 96–99% argon. Further refining can provide high purity. While this stream has a small flow (<1% of feed air), it has a high value. Neon may be recovered from a noncondensed vent stream leaving the reboiler-condenser in the upper column; and a mixture of krypton and xenon from liquid oxygen at the bottom of the upper column.

The distillation columns have cylindrical shells containing numerous perforated metal plates that cause vapor to bubble through the liquid. An important innovation has been to replace the plates with high-efficiency structured packing. This modification reduces pressure loss in each stage and allows the use of more stages and lower operating pressure, which in turn reduces the energy requirement for air compression. There are many variations of the process arrangement to suit specific product flows, pressures, and purity. For example, to produce nitrogen alone, the upper columns are not required, and the enriched oxygen stream may be vaporized and expanded directly to provide refrigeration. The heat exchangers in the process have very large contact areas to provide high thermal efficiency. *See* DISTILLATION; HEAT EXCHANGER.

Adsorptive air separation. Adsorptive separation occurs when one component of a gas mixture preferentially accumulates onto a solid surface, thus depleting its concentration in the gas phase. Zeolite molecular sieves preferentially adsorb nitrogen over oxygen and argon at ambient temperature and pressure, and they are used in the vacuum-swing adsorption (VSA) or pressure-swing adsorption (PSA) process to produce oxygen (**Fig. 2**). Air is compressed by a centrifugal blower to a pressure of about 20 psia (1.4 bars) and then is cooled, and any water condensate is separated. The air passes through a bed of zeolite on which nitrogen is adsorbed. The oxygen product contains about 5% argon and up to 5% nitrogen. After a period (usually less than 1 min) the adsorbent bed is saturated with nitrogen and must be regenerated. At this time the airflow is switched to a second zeolite bed to maintain continuous oxygen production. *See* MOLECULAR SIEVE; ZEOLITE.

The saturated bed is evacuated to a pressure below about 5 psia (0.35 bar) to remove nitrogen. It is then repressurized to atmospheric pressure by using a part of the oxygen product from the on-line bed. This step ensures that the bed is clean before repeating the nitrogen adsorption step. Two or more beds may be used in the process with various pressure equalization steps to achieve maximum process efficiency. Oxygen may be obtained from this process at an energy consumption comparable to that required for cryogenic distillation (240 kWh/ton oxygen) but at a lower purity.

Nitrogen may also be produced by an adsorption process using carbon molecular sieves. Oxygen is adsorbed faster than nitrogen on such materials because of smaller molecular diameter. Air is compressed to about 125 psia (8.6 bars) and cooled, and water is separated. The dehydrated air then passes through the adsorbent bed, where oxygen is removed. The process again uses two or more beds operating in sequence to obtain a continuous nitrogen flow at purities to 0.1% oxygen. When high-purity nitrogen is required, residual oxygen may be reduced to part-per-million levels by catalytic reaction with added hydrogen. Once a bed is saturated with oxygen (<1 min), it is regenerated by depressurizing to atmospheric pressure. Process efficiency is enhanced by partial repressurization with some of the product nitrogen. In both adsorption processes, a computerized automatic control system is used to determine valve switching and optimize production. *See* ADSORPTION.

Membrane separation. Oxygen permeates faster than nitrogen through many organic polymers. This characteristic may be used for air separation. A typical membrane separator contains small hollow polymer fibers 100–500 micrometers in diameter and 1–3 m (3–10 ft) in length. These are assembled in

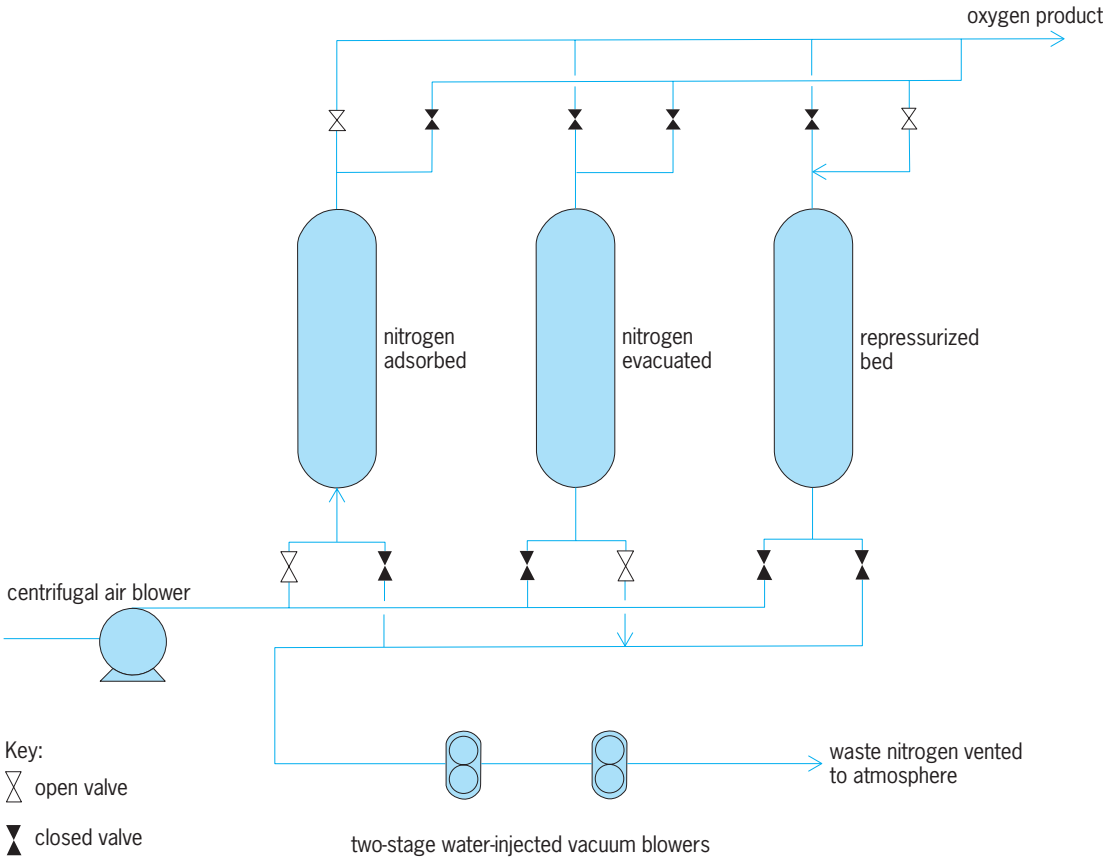

Fig. 2. Process for oxygen production by vacuum swing adsorption.

bundles of 0.1–0.25 m (0.3–0.8 ft) diameter. Polymer fibers used in commercial separators have very thin dense polymer layers as small as 35 nm that are supported on thicker porous walls. Air is compressed to about 160 psia (11 bars) and cooled, and condensed water is separated. The air then flows across one side of the membrane fibers. Oxygen permeates selectively through the fiber wall and is vented to the atmosphere. Some nitrogen permeates with the oxygen, depending upon the selectivity of the polymer used. Commercial polymers have permeation rate selectivities of about 6 for oxygen over nitrogen. Examples of polymers in use are polysulfone, polycarbonate, and polyimides.

Nitrogen is produced directly at pressures up to 150 psia (10.3 bars), with oxygen content down to 0.5% (**Fig. 3**). Higher nitrogen purity may be obtained and process efficiency improved by using additional membrane stages. The low oxygen content of the feed to the second stage produces a nitrogen-rich permeate that can be beneficially recycled to the feed to the first stage. Improved polymers will reduce the energy consumption required for gas separation and may allow the commercial production of oxygen as well as nitrogen. Small membrane units are used in oxygen enrichment for medical applications. *See* MEMBRANE SEPARATIONS; NITROGEN; OXYGEN.

Use for products. Air separation is a major industry. Nitrogen and oxygen rank second and third in the scale of production of commodity chemicals; and air is the primary source of argon, neon, krypton, and xenon. Oxygen is used for steel, chemicals manufacture, and waste processing. Important uses are in integratedgasification combined

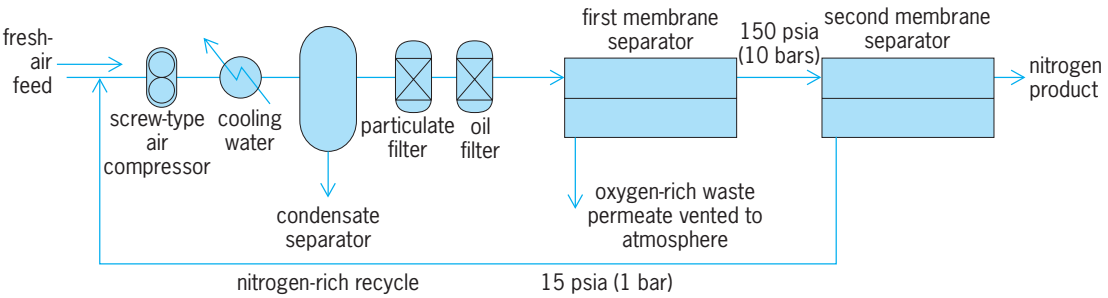

Fig. 3. Membrane process for nitrogen separation.

cycle production of electricity, waste water treatment, and oxygen-enriched combustion. Nitrogen provides inert atmospheres for fuel, steel, and chemical processing and for the production of semiconductors. Robert M. Thorogood

Bibliography. T. Flynn, *Cryogenic Engineering*, 2d ed., 2004; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers Handbook*, 7th ed., 1997; R. Rousseau (ed.), *Handbook of Separation Process Technology*, 1987; D. M. Ruthven, *Principles of Adsorption and Adsorption Processes*, 1984; *Separation and Purification Technology* (journal).

Air temperature

The temperature of the atmosphere represents the average kinetic energy of the molecular motion in a small region, defined in terms of a standard or calibrated thermometer in thermal equilibrium with the air. *See* TEMPERATURE.

Measurement. Many different types of thermometer are used for the measurement of air temperature, the more common depending on the expansion of mercury or alcohol with temperature, the variation of electrical resistance with temperature, or the thermoelectric effect (thermocouple). The electrical methods are especially useful for the automatic recording of temperature. The basic problems of ensuring that the temperature of the thermometer be as close as possible to that of the air are the same for all methods of measurement. For the atmosphere, probably the most serious difficulty is the heating or cooling of the thermometer by radiation to and from other bodies at different temperatures, the most obvious being the Sun. The representativeness of temperature measurements, meaning the degree to which they provide information about the temperature of the air over a region much larger than the thermometer, is also an important practical requirement. The well-known standard meteorological measurement of the air temperature in a louvered shelter, about 6.5 ft (2 m) above a natural ground surface, with a mercury-in-glass thermometer that averages the temperature over a period of about 1 min because of its thermal inertia, is designed to satisfy the above requirements. *See* METEOROLOGICAL INSTRUMENTATION; TEMPERATURE MEASUREMENT; THERMOMETER.

Causes of variation. The temperature of a given small mass of air varies with time because of heat added or subtracted from it, and also because of work done during changes of volume, according to alternate Eqs. (1) and (2). Here h represents heat added,

$$\frac{dT}{dt} = \frac{1}{C_v}\left(\frac{dh}{dt} - P\frac{d\alpha}{dt}\right) \tag{1}$$

$$\frac{dT}{dt} = \frac{1}{C_p}\left(\frac{dh}{dt} + \frac{1}{\rho}\frac{dP}{dt}\right) \tag{2}$$

P the pressure, ρ the density and α its reciprocal, and C_v and C_p the specific heat at constant volume and constant pressure, respectively. The heat added or subtracted may be due to many different physical processes, of which the most important are absorption and emission of radiation, heat conduction, and changes of phase of water involving latent heat of condensation and freezing. In the upper atmosphere, above about 12 mi (20 km), photochemical changes are also important; for example, those that occur when ultraviolet radiation dissociates oxygen molecules to atomic oxygen, which then recombines with molecular oxygen to form ozone. Because of the variation of air pressure with height, rising and sinking of air causes expansion and contraction and thus temperature changes due to the work of expansion, explicitly represented by the second term in parentheses in Eq. (1).

A spectacular example of temperature rise due to sinking of air from higher levels is the chinook, a warm wind that sometimes blows down the eastern slope of the Rocky Mountains in winter. The slower seasonal temperature changes are mainly due to a combination of radiational heat exchange and conduction to and from the ground surface, whose temperature itself changes in response to the varying radiational exchange with the Sun and the atmosphere. On a shorter time scale the diurnal variation of temperature throughout the day is caused by the same processes. *See* AIR PRESSURE; ATMOSPHERIC GENERAL CIRCULATION; CHINOOK.

The rate at which the temperature changes at a particular point, that is, as measured by a fixed thermometer, depends on the movement of air as well as the physical processes discussed above. Large changes of air temperature from day to day are mainly due to the horizontal movement of air, bringing relatively cold or warm air masses to a particular point, as large-scale pressure-wind systems move across the weather map. *See* AIR MASS.

Temperature near the surface. Temperatures are read at one or more fixed times daily, and the day's extremes are obtained from special maximum and minimum thermometers, or from the trace (thermogram) of a continuously recording instrument (thermograph). The average of these two extremes, technically the midrange, is considered in the United States to be the day's average temperature. The true daily mean, obtained from a thermogram, is closely approximated by the mean of 24 hourly readings, but may differ from the mid-range by 1 or 2°F (0.6 or 1°C), on the average. In many countries temperatures are read daily at three or four fixed times, chosen so that their weighted mean closely approximates the true daily mean. These observational differences and variations in exposures complicate comparison of temperatures from different countries and any study of possible climatic changes.

Averages of daily maximum and minimum temperature for a single month for many years give mean daily maximum and minimum temperatures for that month. The average of these values is the mean monthly temperature, while their difference is the mean daily range for that month. Monthly means, averaged through the year, give the mean annual temperature; the mean annual range is the difference between the hottest and coldest mean monthly values.

The hottest and coldest temperatures in a month are the monthly extremes; their averages over a period of years give the mean monthly maximum and minimum (used extensively in Canada), while the absolute extremes for the month (or year) are the hottest and coldest temperatures ever observed. The interdiurnal range or variability for a month is the average of the successive differences, regardless of sign, in daily temperatures.

Over the oceans the mean daily, interdiurnal, and annual ranges are slight, because water absorbs the insolation and distributes the heat through a thick layer. In tropical regions the interdiurnal and annual ranges over the land are small, because the annual variation in insolation is relatively small. The daily range also is small in humid tropical regions, but may be large (up to 40°F or 22°C) in deserts. Interdiurnal and annual ranges increase generally with latitude, and with distance from the ocean; the mean annual range defines continentality. The daily range depends on aridity, altitude, and noon sun elevation.

Extreme temperatures arouse much popular interest and often are cited uncritically, despite their possible instrumental, exposure, and observational errors of many kinds. The often given absolute maximum temperatures of 134°F (57°C) for the United States in Death Valley, California (July 10, 1913), and 136°F (58°C) for the world in Azizia, Tripoli (September 13, 1922) are both questionable; in the subsequent years, Death Valley's hottest reading has been only 127°F (53°C), and the Azizia reading was reported by an expedition, not a regular weather station. Lowest temperatures in the Northern Hemisphere are −90°F (−68°C) at Verkhoyansk (−89.7°F or −67.7°C on February 5 and 7, 1982) and Oimekon (−89.9°F or −67.7°C on February 6, 1933), Siberia; −87°F (−66°C) at Northice, Greenland (January 9, 1954); −81°F (−63°C) at Snag, Yukon Territory, Canada (February 3, 1947); −70°F (−57°C) at Rogers Pass, Montana (the current United States record, on January 20, 1954). The first winter at Vostok, 78°27′S, 106°52′E, encountered a minimum temperature of −125°F (−87.2°C) on August 25, 1958, and the third winter a minimum of −127°F (−88.3°C) on August 24, 1960, much lower than the lowest at the United States station at the South Pole. At Vostok on July 21, 1983, a new global minimum temperature record for Earth's surface was recorded, −129°F (−89.4°C). *See* ANTARCTICA.

Vertical variation. The average vertical variation of temperature in the atmosphere is shown in the **illustration.** The atmosphere is seen to consist of layers, each of which has a characteristic variation of temperature with height. The decrease of temperature with height in the lowestlayer, the troposphere, is basically due to the presence of a heat source resulting from the solar radiation absorbed at the Earth's surface, giving an excess of heat that is carried away from the surface mainly by convection currents and lost to space by reradiation. Heating a compressible fluid such as air from below results in a decrease in temperature with height because rising masses of air cool as they expand, according to Eq. (2). This is the main reason for the decrease of temperature with height in the troposphere. *See* ATMOSPHERE; HEAT BALANCE, TERRESTRIAL ATMOSPHERIC.

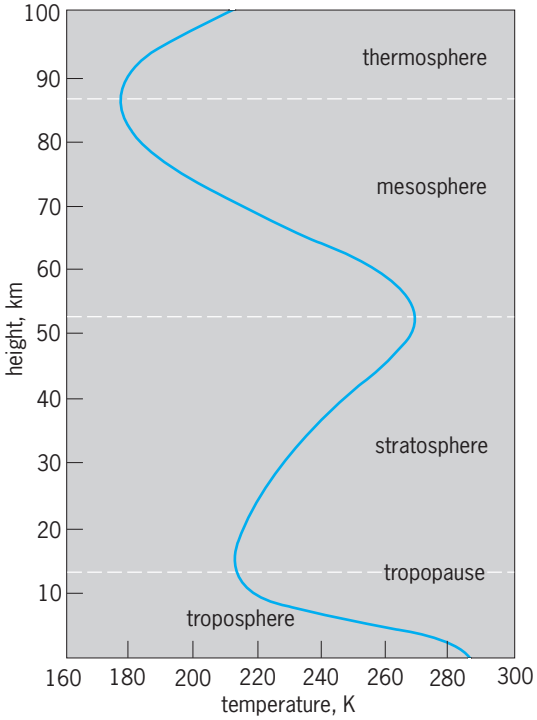

Average temperature distribution with height, from the International Reference Atmosphere. Note that the nomenclature for stratosphere and mesosphere varies, the upper part of the stratosphere being considered part of the mesosphere by some authorities. (*After R. G. Fleagle and J. A. Businger, An Introduction to Atmospheric Physics, Academic Press, 1963*)

In the mesosphere and higher layers the exchange of heat energy between layers of air by emission and absorption of infrared radiation is the most important factor determining the distribution of temperature with height. *See* INSOLATION; RADIATION.

Raymond J. Deland; Edwin Kessler

Bibliography. R. A. Anthes, *Meteorology*, 7th ed., 1996; R. G. Fleagle, *An Introduction to Atmospheric Physics*, 2d ed., 1980; D. D. Houghton (ed.), *Handbook of Applied Meteorology*, 1985; A. Miller and R. A. Anthes, *Meteorology*, 4th ed., 1980; H. Riehl, *Introduction to the Atmosphere*, 3d ed., 1978.

Air-traffic control

A service to promote the safe, orderly, and expeditious flow of air traffic. Safety is principally a matter of preventing collisions with other aircraft, obstructions, and the ground; assisting aircraft in avoiding hazardous weather; assuring that aircraft do not operate in airspace where operations are prohibited; and assisting aircraft in distress. Orderly and expeditious flow assures the efficiency of aircraft operations along the routes selected by the operator. It is provided through the equitable allocation of system resources to individual flights, generally on a first-come-first-served basis.

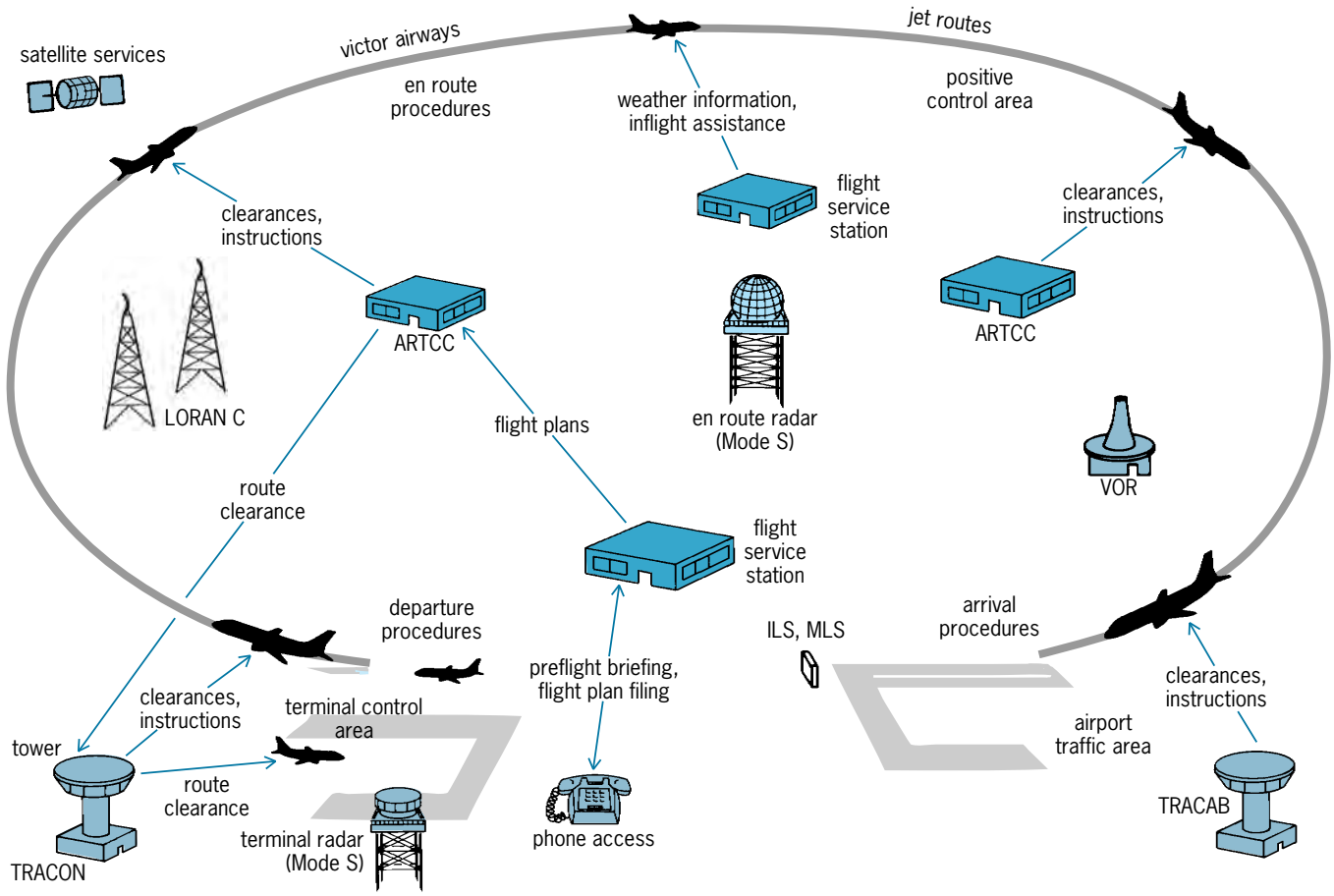

Fig. 1. Elements of air-traffic control.

In the United States, air-traffic control (ATC) is the product of the National Airspace System (NAS), comprising airspace; air navigation facilities and equipment; airports and landing areas; aeronautical charts, information, and publications; rules, regulations, and procedures; technical information; and personnel (**Fig. 1**).

Rules

The rules governing air-traffic control in the United States are published in Federal Aviation Regulations (FARs); two principal categories of rules are visual flight rules (VFR) and instrument flight rules (IFR).

Visual flight rules. These govern the procedures for conducting flight where the visibility, the ceiling, and the aircraft distance from clouds are equal to or greater than established minima. Ceiling is the height above the Earth's surface of the lowest layer of clouds or obscuring phenomenon that significantly restricts visibility. The minima for operation under visual flight rules vary by airspace. In controlled airspace, the ceiling must be at least 1000 ft (305 m) and the visibility must be at least 3 statute miles (4830 m). The aircraft must remain clear of clouds, at least 500 ft (150 m) below, 1000 ft (305 m) above, and 2000 ft (610 m) horizontally.

Instrument flight rules. When the meteorological conditions expressed in terms of visibility, distance from clouds, and ceiling are less than the minima specified for visual flight rules, that is, less than the minima for visual meteorological conditions (VMC), instrument meteorological conditions (IMC) prevail. Instrument flight rules govern flight under such conditions. To operate under these rules, the pilot must pass an instrument flight examination and have an adequately instrumented aircraft.

Flight plans. Instrument flight rules and visual flight rules also refer to types of flight plans. A flight plan is filed with the authority providing air-traffic control services [in the United States, the Federal Aviation Administration (FAA)] to convey information about the intended flight of the aircraft. Both types of flight plans contain essentially the same information, that is, aircraft identification number, make and model, and color; planned true airspeed and cruising altitude; origin and destination airports; planned departure time and estimated time en route; planned route of flight, fuel, and number of people on board; pilot's name and address; navigation equipment on board; and the aircraft's radio call sign, if different from the aircraft identification number.

Generally, a flight plan is not required for a flight under visual flight rules. However, if a flight plan is filed for such a flight and the aircraft is overdue at its destination, search and rescue procedures will be initiated. Hence the flight plan under visual flight rules provides a significant safety benefit. An IFR flight plan

is required for operation in controlled airspace when instrument meteorological conditions prevail.

Aircraft operating under visual flight rules (VFR aircraft) maintain separation from other aircraft visually; that is, the flight crew visually acquires other aircraft in the vicinity, and the pilot maneuvers the aircraft to maintain separation (a procedure called see-and-avoid). IFR aircraft in controlled airspace operate in accordance with clearances and instructions provided by air-traffic controllers for the purpose of maintaining separation and expediting the flow of traffic. (Formally, an air-traffic control clearance is an authorization by air-traffic control, for the purpose of preventing collisions between aircraft, for an aircraft to proceed under specified traffic conditions within controlled airspace. An air-traffic control instruction is a directive issued by air-traffic control for the purpose of requiring a pilot to take specific actions, for example, "turn left to heading two five zero.") Flight crews operating under instrument flight rules are responsible for seeing and avoiding other aircraft, but the air-traffic control clearances they receive provide substantial added assurance of safe separation. Consequently, flight crews often will operate under instrument flight rules even though the weather satisfies visual meteorological conditions.

Airspace

The two principal categories are controlled and uncontrolled airspace. In controlled airspace some or all aircraft are required to operate in accordance with air-traffic control clearances in order to assure safety, to meet user needs for air-traffic control, or to accommodate high volumes of traffic. Air-traffic control services including air-to-ground communications and navigation aids are provided in controlled airspace.

Two specific examples of controlled airspace are class A (the positive control area or PCA) and class B (the terminal control area or TCA). The positive control area is, with a few exceptions, the airspace within the conterminous 48 states and Alaska extending from 18,000 to 60,000 ft (5490 to 18,290 m) above mean sea level. All aircraft in a positive control area must operate in accordance with instrument flight rules and carry prescribed equipment.

Terminal control areas are centered on primary airports and extend from the surface to specified altitudes. An air-traffic control clearance and prescribed equipment are required prior to operating within a terminal control area regardless of weather conditions. Uncontrolled airspace simply is airspace that has not been designated as controlled; air-traffic control services may not be available in such airspace.

Equipment

An array of sophisticated equipment is required for providing air-traffic control services. Principal categories are air-to-ground communications, radio navigation aids, surveillance systems, and air-traffic control automation.

Air-to-ground communications. Two-way air-to-ground voice communications between civil pilots and air-traffic controllers are conducted in the very high frequency (VHF) band. In the United States there are approximately 2400 communications facilities providing access to air-traffic control services from terminal and en route facilities as well as from flight service stations (FSSs). In addition, certain radio navigation aids [nondirectional radio beacons (NDBs) and very high frequency omnidirectional range (VOR) stations] can provide one-way communications from controllers to aircraft. These channels generally are used to broadcast weather and aeronautical information to pilots. *See* RADIO SPECTRUM ALLOCATIONS.

Air-to-ground data communications (that is, data link) increasingly are used to transfer information to and from the cockpit. A principal advantage is that many of the communications errors associated with humans incorrectly reading, speaking, and hearing text are eliminated by communications protocols that detect errors in data transmissions, by electronically displaying the information received in the cockpit and the air-traffic control facility, and by storing the received information so that it can be recalled and reviewed by pilots and controllers. Data link also permits large quantities of data to be exchanged between ground-based and airborne computers. Civil aviation is exploiting three data-link media: some VHF voice channels have been converted to data-link channels; Mode S provides a data-link capability, both air to ground and air to air; and communications satellites provide data transmission capabilities.

Radio navigation aids. Radio navigation aids are used to determine the plan position of the aircraft (that is, the position in the horizontal plane) in coordinates referenced either to the navigation aid or to the Earth (that is, latitude and longitude). For most operations, the aircraft vertical position is determined by sensing atmospheric pressure on board and converting this pressure to altitude, based on a standard model of the atmosphere. For the landing phase of flight, precision landing aids provide horizontal and vertical position referenced to the runway. *See* ALTIMETER.

VOR. This is a principal system used for determining plan position, with approximately 1000 ground stations nationwide. The system provides the magnetic azimuth from the VOR station to the receiving aircraft accurate to $\pm 1^\circ$. Position determinations can be obtained from the intersection of radials from VORs with overlapping coverage volumes. A radial is a line of position extending from the VOR station at the indicated magnetic azimuth of the aircraft. With the addition of distance-measuring equipment at a VOR station, it is possible to obtain a position determination from a single station. *See* DISTANCE-MEASURING EQUIPMENT; RHO-THETA SYSTEM; VOR (VHF OMNIDIRECTIONAL RANGE).

Nondirectional radio beacon. This is an older technology, with few installations remaining. The system radiates a continuous signal from which direction-finding receivers can determine the azimuth to the ground station. *See* DIRECTION-FINDING EQUIPMENT.

Loran C. This is a pulsed system, with chains of ground stations each consisting of one master station and at least two secondary stations organized to transmit their signals at precisely defined time epochs. Loran C coverage in the United States includes the conterminous 48 states and southern Alaska. *See* LORAN.

Landing systems. In order to conduct approaches and landings in low-visibility conditions, it is necessary that an electronic glideslope (or glidepath) be provided as a reference for controlling the descent of the aircraft to the runway. In addition, a stable guidance signal is required to align the aircraft with the runway centerline.

The instrument landing system (ILS) has been the standard means for providing precision landing guidance to the runway, and is installed on approximately 1000 runways in the United States. The localizer antenna transmits the lateral (left and right) guidance signal over a 20° sector, 10° on both sides of the extended runway centerline. The glideslope antenna transmits the elevation guidance signal over a 1.4° sector, 0.7° on both sides of the glidepath, which is normally 3.0° above the horizontal. *See* INSTRUMENT LANDING SYSTEM (ILS).

A new standard system for providing precision approach guidance, the microwave landing system (MLS) has been designed to eliminate limitations of the instrument landing system. It utilizes scanning-beam technology to provide proportional landing guidance over 80° in azimuth (40° on both sides of the extended runway centerline) and 15° in elevation. In combination with a precision version of distance-measuring equipment, the system can provide three-dimensional landing guidance within the scanned volume, thereby permitting curved approaches and approaches at higher glideslope angles than those available from the instrument landing system. In the United States, the development of MLS for civilian applications has been terminated. A mobile version of MLS (MMLS) continues to be used for military operations. *See* MICROWAVE LANDING SYSTEM (MLS).

Global Positioning System. The constellation of Global Positioning System (GPS) satellites provides a highly accurate worldwide position determination and time transfer capability. In the horizontal plane, the position determined by a GPS receiver is within approximately 108 ft (33 m) of the true receiver position at least 95% of the time, averaged over the globe. The vertical position is accurate to within approximately 240 ft (73 m) on the same 95% probability basis. In addition, the receiver provides Coordinated Universal Time (UTC) with an accuracy of 50 nanoseconds (95% probability), averaged over the globe. During times of significant ionospheric activity, the timing accuracy degrades to approximately 120 nanoseconds at the worst location. Coordinated Universal Time is an internationally accepted time standard that never differs from Greenwich Mean Time by more than 1 s. The principal advantages of GPS are its accuracy and worldwide coverage. GPS is approved for oceanic, en route, area navigation (RNAV), and nonprecision approach operations. GPS receivers that incorporate correction signals from the Wide Area Augmentation System (WAAS) have accuracies of approximately 5–7 ft or 1.5–2 m (95% probability) in both horizontal and vertical dimensions. GPS/WAAS is approved for aircraft guidance down to 200 ft (61 m) above the runway surface for localizer performance with vertical guidance (LPV) approaches. LPV is an ILS-like capability without the need for ground equipment at the airport. *See* AIR NAVIGATION; ELECTRONIC NAVIGATION SYSTEMS; SATELLITE NAVIGATION SYSTEMS.

Surveillance systems. Air-traffic controllers use radar to monitor the positions of aircraft in their sectors of responsibility and to monitor areas of heavy precipitation. The radar information is used to develop clearances and instructions for separating aircraft operating under instrument flight rules, and to provide traffic advisories to IFR aircraft and to VFR aircraft receiving the traffic advisory service. Traffic advisories assist a pilot in seeing and avoiding other aircraft. They provide the ranges, bearings, and altitudes of aircraft in the pilot's immediate vicinity. The pilot is responsible for visually acquiring and avoiding any traffic that may be a collision threat. Two principal types of radar are used in civil air-traffic control: secondary, or beacon, radar and primary radar. *See* RADAR.

Secondary radar. This is an interrogate-respond system. The rotating directional antenna of the ground station transmits a pulse pair to the transponder in the aircraft. The pulse spacing encodes one of two messages, "transmit your altitude" (the Mode C interrogation) or "transmit your identity" (the Mode A interrogation). The aircraft transponder transmits an encoded pressure-altitude reply in response to the first interrogation and a four-digit identity code, assigned by air-traffic control and entered into the transponder by the pilot, in response to the second. The aircraft is shown on the controller's plan view display at the azimuth corresponding to the pointing direction of the antenna and the range corresponding to the round-trip time between transmission of the interrogation and receipt of the reply. Air-traffic control computers receive the encoded reply data from radar sites and place corresponding information in data blocks next to the symbols depicting the aircraft positions on the display. The identity code assigned by air-traffic control is correlated with the flight-plan database in the computer to display the radio call sign in the data block. The aircraft pressure altitude is displayed in hundreds of feet.

Primary radar. This operates by transmitting high-power, radio-frequency pulses from a rotating directional antenna. The energy is reflected from any aircraft in the directional beam and received by the antenna. The aircraft is displayed at the azimuth corresponding to the pointing direction of the antenna and the range corresponding to the round-trip time between pulse transmission and receipt of the reflected signal.

Comparison of radar systems. Primary radar has the advantage that aircraft without air-traffic control transponders can be detected, and energy reflected from heavy precipitation indicates to the controller areas of potentially hazardous weather. Secondary radar ground stations interrogate at 1030 MHz, and the transponder replies at 1090 MHz. The use of different frequencies eliminates extraneous returns (clutter) from surrounding buildings and terrain that can reduce the effectiveness of primary radar in detecting aircraft. Secondary radar also has the advantage that aircraft identity and altitude can be determined. At most air-traffic control radar sites, the secondary radar antenna is mounted on the primary radar antenna, and they are turned by a common drive system. In the United States, there are 220 radar systems operated in terminal areas and 116 long-range radars in the en route environment.

Mode S. The secondary radar system has been improved through the addition of Mode S, which employs more sophisticated signaling formats than Modes A and C. Each aircraft transponder is permanently assigned a unique address and interrogations therefore can be addressed to individual aircraft. Two benefits accrue. For a number of technical reasons, the accuracy and reliability of the surveillance data are improved. Second, it is possible to establish direct data communications with individual aircraft. The data link can be used to provide air-traffic control clearances to aircraft and to convey a wide variety of aeronautical and weather information in response to requests from the flight crew.

Oceanic surveillance. In the oceanic environment, the ground-based surveillance systems described above obviously cannot be used. Oceanic operations are now based on rigid procedures and high-frequency (HF) communications that sometimes are unreliable. With the advent of commercially available mobile satellite communication systems, the development of a technique called automatic dependent surveillance (ADS) has been undertaken to provide real-time position information from aircraft over the ocean. In the operation of this system, the position of the aircraft, as determined from on-board navigation sensors, is communicated to air-traffic control facilities when requested by satellite relay. This position information can be displayed to controllers as though it had been determined by a radar system. With automatic dependent surveillance and reliable pilot-controller communications via satellite, a number of improvements in oceanic procedures are possible, with corresponding benefits for safety and aircraft operating efficiency.

Automation. The principal elements of the controller's workstation are the plan view display, a track ball or mouse, the data-entry keyboard, printed flight strips showing the flight plans of aircraft for which the controller is responsible, and interfaces with communications facilities linking the controller with aircraft and with other controllers and air-traffic control facilities. The plan view display shows two principal types of data, map data and radar data (**Fig. 2**). Map data include the locations of airports and their runways, navigation aids, airways, obstructions, and the geographical limits of the facility's airspace. Radar data comprise the positions of aircraft, including their altitudes, ground speeds, and radio callsigns, as well as areas of precipitation. The

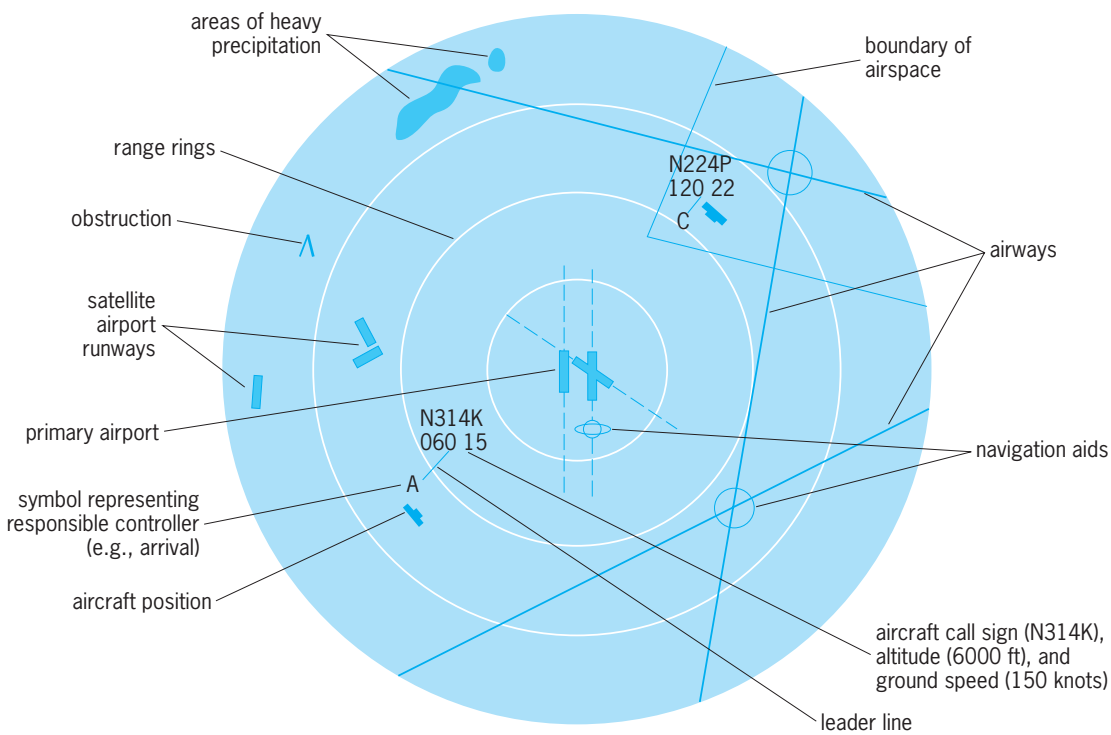

Fig. 2. Terminal radar controller's plan view display. Aircraft positions are shown by radar returns. Display is centered on primary airport with parallel and intersecting runways. Range rings are spaced 5 mi (9 km) apart. Airspace within boundary and above 10,000 ft (3050 m) belongs to the center.

data-entry keyboard allows the controller to modify data stored in the automation system, including flight plans. Extensive automation (computer) equipment is used in maintaining the flight-plan databases and processing radar data.

A number of automation aids have been developed to assist controllers in separating aircraft as well as in sequencing and metering aircraft into and out of busy terminal areas. The objective of these aids is to enhance the safety and efficiency of flight operations.

Flight management computer systems are installed in aircraft for the purpose of guiding the aircraft along its planned route of flight while minimizing operating costs by selecting optimum speeds and altitudes. Extensive databases are stored in the flight management computer system (FMCS), including the current flight plan, wind velocities and air temperatures along the planned route of flight, and the positions and operating frequencies of the radionavigation aids to be used. Interfaces with the FMCS for air-to-ground data communications permit changes to be made to the databases in flight and allow information to be extracted, such as automatic dependent surveillance position reports and estimated times of arrival at specific points along the planned route of flight. The integration of FMCS operations with ground-based air-traffic control operations is a principal theme in the further development of the art of air-traffic control. *See* AIRCRAFT INSTRUMENTATION.

Airborne collision avoidance systems. These systems are installed in aircraft to provide ground-independent protection from midair collisions, as a backup to the conventional air-traffic control system. Within the United States, the system is known as the Traffic Alert and Collision Avoidance System (TCAS). The TCAS equipment in the aircraft interrogates the secondary surveillance radar transponders in proximate aircraft and processes the replies to determine if any aircraft is on a collision course. Traffic advisories are displayed to the pilot to portray the range, bearing, and relative altitude of any aircraft that penetrates a protection volume around the TCAS-equipped aircraft. The pilot's response to a traffic advisory is to visually search for the intruder and to maintain visual separation once the intruder is acquired. If the intruder continues to close on the TCAS-equipped aircraft, a resolution advisory will be displayed to tell the pilot how to maneuver to avoid a collision. *See* AIRCRAFT COLLISION AVOIDANCE SYSTEM.

Airways and Procedures

Two fixed-route systems have been established for air navigation. From 1200 ft (360 m) above the surface up to but not including 18,000 ft (5490 m) above mean sea level, there are designated airways based on VORs and nondirectional beacons. The most prevalent are the so-called victor (V) airways defined by VORs. For example, in western New York, airway V115 extends from the Buffalo VOR to the Jamestown VOR. Jet (J) routes are defined from 18,000 to 45,000 ft (5490 to 13,710 m) above mean sea level, based solely on VORs.

Aircraft may navigate from one terminal area to another by following fixed routes from one VOR to the next. In this case, the route shown on the flight plan would list the airways to be flown by their V and J designations.

There are three principal categories of procedures: departure procedures for leaving terminal areas, arrival procedures for entering terminal areas, and en route procedures. Departure procedures prescribe the process for route clearance delivery to an aircraft, for providing takeoff runway and taxi instructions, and for defining or placing limitations on the climb-out route of the aircraft to the en route environment. Generally, pilots of IFR aircraft call the clearance delivery controller for their route clearance prior to taxiing. The route in the clearance may differ from the filed route because of system restrictions such as excess traffic, facility outages, and weather.

En route procedures deal principally with reporting aircraft flight progress to air-traffic control (position reporting) when the aircraft is outside radar coverage or is operating in holding patterns.

Arrival procedures prescribe the process for making the transition from the en route structure to the terminal area, for approaching the landing runway, and for executing a missed approach when a landing cannot be accomplished. An instrument approach procedure is a series of predetermined maneuvers by reference to flight instruments for the orderly transfer of an aircraft from an initial approach fix to a landing or to a point from which a landing can be made visually. Several procedures, using different navigation and approach aids, may be established for an airport.

Facilities

Air-traffic control facilities include flight service stations, air-route traffic control centers (ARTCCs), and terminal facilities.

Flight service stations. These provide preflight briefings for pilots, accept flight plans, broadcast aviation weather information, assist lost aircraft and aircraft in distress, and monitor the operation of radio navigation aids. The preflight briefing provides weather and aeronautical information pertinent to the proposed flight, including any facility outages that may be relevant and any known air-traffic control delays that may affect the flight. There are 75 flight service stations in the United States.

Air-route traffic control centers. These monitor all IFR aircraft not under the control of military or terminal facilities. In particular, they control all traffic operating in positive-control airspace. They assure separation of IFR aircraft by issuing clearances and instructions as necessary and issuing traffic advisories to aid see-and-avoid, provide weather advisories, accept amendments to flight plans from flight crews, and assist aircraft in distress.

Flight plans submitted to flight service stations usually are transmitted to the parent air-route

traffic control center, where they are processed and the route clearance is generated. This clearance is transmitted to the clearance delivery position at the departure airport about 30 min prior to the estimated departure time shown in the flight plan.

A typical center is responsible for more than 100,000 mi^2 (250,000 km^2) of airspace and hundreds of miles of airways. This geographical area is divided into 30 or more control sectors each with a controller team assigned. Each sector is defined by boundaries in the horizontal plane as well as by altitude bounds. There are 20 centers in the conterminous 48 states plus one each in Alaska, Hawaii, Puerto Rico, and Guam.

Terminal facilities. There are three principal positions staffed by tower controllers. The ground controller position is responsible for all ground traffic not on active runways (runways in use for takeoffs and landings). The local controller has jurisdiction over the active runways and the airspace close to the airport used by arrivals and departures. The clearance delivery position may be covered by the ground controller when traffic activity is low. The tower environment is essentially a visual environment. Controllers visually acquire and track aircraft and direct their movements by using radio or, when an aircraft has no operating radio, signal lights. In some locations, radar indicator equipment is installed in the tower to electronically display traffic that is being tracked by the local air-traffic control radar.

A radar terminal control facility serving a major airport and its satellite airports is frequently located in the tower building below the tower cab in the terminal radar approach control (TRACON) facility. At some low-activity airports, the radar displays for arrival and departure control are in the tower cab. This arrangement is known as the terminal radar approach cab (TRACAB). Terminal airspace where radar service is provided is divided into sectors which may include a number of departure sectors and arrival sectors each staffed by a team of one or more controllers. When activity is low, sectors may be combined to reduce personnel requirements.

There are approximately 5400 airports in the United States open to the public. Approximately 550 receive scheduled commercial passenger service, and 462 have towers staffed by the FAA. *See* AIR TRANSPORTATION.

Clyde A. Miller; Richard L. Greenspan

Bibliography. *Federal Aviation Regulations/Aeronautical Information Manual*, Federal Aviation Administration, published annually; *Federal Radionavigation Plan*, U.S. Department of Transportation, revised every 2 years; *Global Positioning System Standard Positioning Service Performance Standard*, Assistant Secretary of Defense for Command, Control, Communications, and Intelligence, October 2001; M. Kayton and W. R. Fried, *Avionics Navigation Systems*, 2d ed., Wiley-Interscience, New York, 1997; *National Air-space System Architecture*, Federal Aviation Administration, updated annually; M. S. Nolan, *Fundamentals of Air Traffic Control*, 4th ed., Brooks/Cole, Belmont, CA, 2003; C. D. Wickens et al., *Flight to the Future: Human Factors in Air Traffic Control*, vols. I and II, National Academy Press, 1997.

Air transportation

The movement of passengers and cargo by aircraft such as airplanes and helicopters. Air transportation has become the primary means of common-carrier traveling. Greatest efficiency and value are obtained when long distances are traveled, high-value payloads are moved, immediate needs must be met, or surface terrain prevents easy movement or significantly raises transport costs. Although the time and cost efficiencies obtained decrease as distance traveled is reduced, air transport is often worthwhile even for relatively short distances. Air transportation also provides a communication or medical link, which is sometimes vital, between the different groups of people being served.

Elements of the air system. The provision, continuance, and improvement of air transport services of all types require a complex integrated system whose main parts are the aircraft operator, the aircraft manufacturer, the airport terminal, and the air-traffic control system.

Aircraft operator. The aircraft operator provides the basic flight service to the customer, usually maintains the aircraft, and provides any needed customer services. The operator may be an airline providing for-hire service to customers on a commercial basis, or may operate the aircraft entirely for personal use. Scheduled and nonscheduled for-hire services are provided by a wide variety of airlines offering different classes of service. The type of route system served by the airline is often used to describe broad groups of carriers (that is, majors, nationals, large regionals, and medium regionals). More specialized airlines haul only cargo or operate only helicopters. Industry trade groups represent the combined interests of these carriers. Within the United States, the Air Transport Association represents the 21 largest American airlines from among approximately 93 airlines that provided scheduled service in 1987. Internationally, many of the major airlines are members of the International Air Transport Association, an organization primarily concerned with international tariff levels but also active in legal and safety matters. The International Civil Aviation Organization handles international air transportation matters, including technical standards and practices; provides an international language, including standard chart and identification codes; and collects and disseminates statistics and other data on the world's airlines.

Aircraft manufacturer. Aircraft manufacturers include the airframe manufacturer and the engine manufacturer, who together design and produce the airplane, fabricate the spare parts necessary to continue operation, and provide continuing technical help to the aircraft operator. The design, development, and fabrication of a major aircraft constitute a project of

tremendous scope which requires billions of dollars to complete. The airframe and engine manufacturer may be supported by thousands of subcontractors who produce different components of the airplane. The entire aircraft design, production, and assembly process may stretch across several nations. Governments may at times participate by providing direct financial support or by helping develop a steady stream of advanced technology which can be used to improve future aircraft. Such advances may come through spinoffs from advanced military aircraft or through research supported or accomplished by civil government agencies such as the National Aeronautics and Space Administration. In the United States, the transport airframe and engine manufacturers are represented by the Aerospace Industries Association of America, Inc., the general aviation industry by the General Aviation Manufacturers Association, and the helicopter industry by the Helicopter Association International.

Airport terminal. The airport terminal, characterized primarily by a runway and a passenger or cargo service facility, provides a central location for the airplane to take off and land and for passengers to begin and end the air portion of their journey. Major parts of the airport include the runways, taxiways, control tower, and buildings to service passengers, cargo, aircraft, and administrative needs. The largest airports have runways in excess of 8000 ft (2500 m) and are generally publicly owned and operated. Only one major airport has been built in the United States since 1970 and new construction has been confined to lengthening airport runways and improving the quality of smaller airports. Construction costs of a major new airport easily exceed several billion dollars, and land acquisition is both difficult and expensive. Local citizen opposition to airport construction is sometimes intense, with opponents citing objections to aircraft noise, ground congestion, land-use changes, and potential accidents. Airports may serve a wide variety of purposes or may be specialized (that is, passengers, cargo, general aviation, private, or military). Most airports of any size in the United States are paved, but in less developed parts of the world the vast majority of a nation's airports may have unpaved runways. *See* AIRPORT.

Air-traffic control system. Much of the air-traffic control (ATC) system is located at the airport terminal. Within the United States the ATC system is operated by a government agency, the Federal Aviation Administration (FAA; part of the Department of Transportation). The ATC system controls the aircraft's movements both on the ground and during flight, and provides a means of allocating available runways and air space between user aircraft. Equipment includes radios, radars, light signals, displays, and other types of electronic instrumentation. The airport control tower directs air traffic in the vicinity of the airport and on the runways and taxiways. Voice communication, navigational aid, weather information, and other services are also provided by the ATC operators to the crew flying the aircraft. The United States has over 17,000 public and private airports (including those with unpaved runways), of which over 400 have FAA towers. *See* AIR-TRAFFIC CONTROL.

Evolution and characteristics. The United States is the world's leading user and operator of aircraft, and has also produced most of the commercial airplanes used throughout the world. The foundation for this was laid with the Air Commerce Act of 1926, which began the development of commercial aviation in the United States. Regularly scheduled transcontinental service originated with a single route between New York and Los Angeles in 1930 (**Fig. 1**). To help

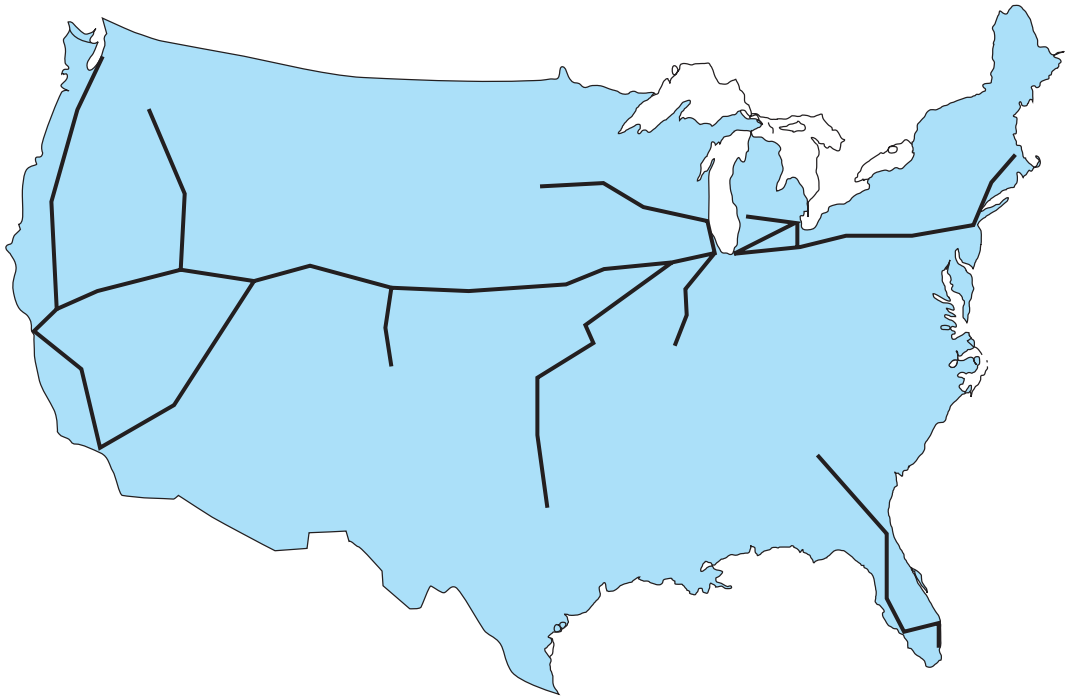

Fig. 1. Early airline route structure in the United States.

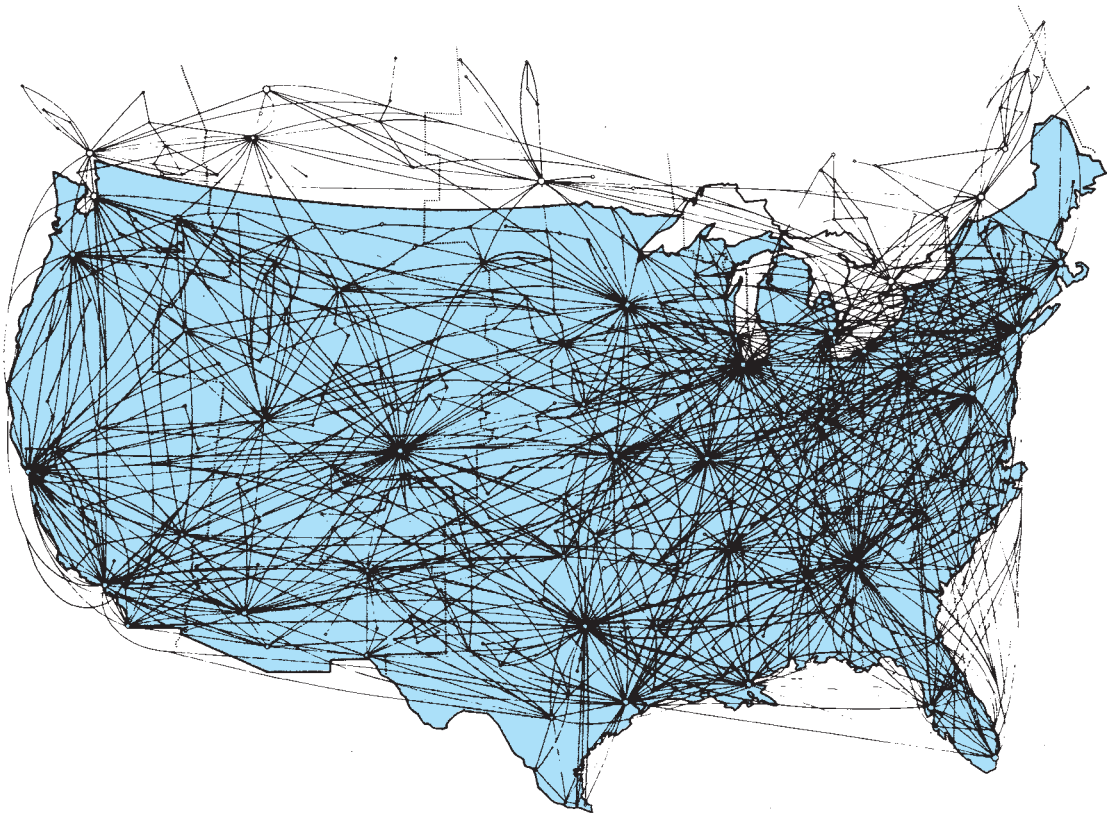

Fig. 2. Present airline route structure in the United States.

establish widespread air service, the government built and maintained navigational routes and provided subsidies to the airline operators through Post Office contracts to carry mail and through direct cash subsidies paid to cover losses generated between various city pairs. Subsidy payments to airlines were discontinued as the lines became able to support themselves.

Since the late 1940s, the number of passengers traveling by air between the United States and the rest of the world has exceeded the number of passengers traveling by sea. In domestic service, the revenue passenger-miles moved by air has exceeded those by rail since 1956. (A revenue passenger-mile is defined as one fare-paying passenger transported 1 statute mile or 1.6 km.) Air transportation is the dominant means of intercity common-carrier passenger service, moving over 92% of this traffic. The complex airline route system (**Fig. 2**) is a dramatic change from the original system (Fig. 1). Figure 2 also illustrates the so-called hub cities in the United States; air traffic is funneled from the smaller cities to the larger hub cities. Airlines are primarily carriers of people, obtaining over 87% of their revenue from passenger fares. Fares charged by airlines have historically increased at a considerably lower rate than consumer prices as a group.

Passenger composition. For a long time, airline service was oriented toward the business traveler, who made up the great bulk of demand. Over the years, however, the market has changed to the point where the business traveler now represents only about half of all passengers carried, with the pleasure or private traveler accounting for the other half.

Aerospace industry. The aerospace industry (aviation and missiles) is the leading United States exporter of manufactured products, consistently making an important contribution to the nation's balance of payments. Other countries have recognized the significance of this industry to their national economy, and some now produce aircraft similar in quality to those of the United States.

Safety. Safety has always been an overriding air transportation concern. Considerable media attention is given to an air crash of almost any sort, and such unfavorable publicity can be financially disastrous to a commercial carrier perceived as negligent. In 1958 Congress vested safety regulation in the Federal Aviation Administration. Before a new type of aircraft can be introduced to service, the FAA must certify the ability of the aircraft to meet airworthiness standards. Accident investigations are handled by the National Transportation Safety Board (NTSB). United States scheduled airlines have improved their safety performance with time: from the mid-1950s to the mid-1980s the fatality rate dropped from approximately 0.50 to 0.05 fatality per 100,000,000 passenger-miles. Much of this improvement is due to the introduction of the jet airplane and its almost complete replacement of propeller aircraft, as discussed below. Air transportation safety is comparable to that of other public modes of transportation and superior to that of the automobile.

Fig. 3. Advanced aircraft. (*a*) Boeing B-747-400 (*Boeing Company*). (*b*) McDonnell-Douglas MD-11 (*McDonnell-Douglas Corp.*).

Air fleet. The subsonic jet aircraft was introduced to United States commercial service in 1958 and revolutionized the air transport industry as well as the composition of the airline fleet. Within 10 years, jet aircraft had almost completely replaced propeller-driven airplanes. The advantages of the jet transport include superior economics, speed, range, capacity, comfort, and safety, which, taken together, forced the replacement of most propeller aircraft. Propeller aircraft are still used in certain types of airline service mainly for short-haul, low-passenger-capacity, and lower-altitude flights. *See* AIRCRAFT PROPULSION; JET PROPULSION.

Major transport airplanes include the B-767/757/747/737/727, the MD-80 series, and the DC-10/9/8 (**Fig. 3**). Such aircraft generally cruise at speeds of approximately 540 mi/h or 870 km/h (240 m/s) and at altitudes high enough that aircraft pressurization is required. Principal characteristics of three selected transport aircraft are given in the **table**.

Technological advances. Since its introduction, the transport airplane has steadily increased in speed, size, and range. Successive technology advances have continuously improved the quality and cost effectiveness of the air fleet. The most important of these advances have been the development of the jet and fan-jet engine, the high-bypass-ratio jet engine, major increases in aircraft size, and improvements in aerodynamics, materials, design techniques, and manufacturing. Many years (even decades) of research may be required before a new advance can be safely and economically incorporated into a production airplane. Despite the continuous improvements in technology, commercial aircraft are long-lived vehicles, and are often used for periods exceeding 20 years.

Supersonic aircraft. In the late 1970s, a new type of high-speed airplane, the Concorde, entered service on certain international air routes. Although it is capable of flight at speeds more than twice that of sound, its relatively high costs, limited range, small payload, and restricted flight capabilities resulted in a total of only 16 aircraft built. Widespread use of supersonic commercial aircraft therefore requires the development of a more technologically advanced airplane able to operate economically with a much greater range and payload capability, possibly over land as well as over water, and with acceptable environmental characteristics.

General aviation fleet. The general aviation fleet is far larger in numbers than the fleet operated by the commercial airlines. These aircraft are used for almost every conceivable transport purpose (personal, commercial, recreational, agricultural, and training). Many of these aircraft operate over short distances and do not require the elaborate communications, navigation, and safety equipment found in the typical commercial transport. The flexibility and utility of general aviation vehicles, however, are so great that the fleet has experienced major growth. Helicopters are also used for an increasing number of purposes, with industrial uses such as construction, mining, offshore oil drilling, and laying pipelines joining the traditional use of transporting passengers. *See* GENERAL AVIATION; HELICOPTER.

Development of new aircraft. New aircraft conceptual designs incorporating varying degrees of advanced technology are continually evaluated by the manufacturers of both transport and general aviation aircraft. Leading airlines work closely with transport manufacturers during the development stage to determine exactly what specifications and features the proposed aircraft should have. The transport manufacturers produce a new aircraft approximately every 12 years, and also supplement their product line with variations of the basic airplane called derivatives. Derivative aircraft are usually produced by "retching" the original airplane fuselage to permit the fabrication of airplanes of different length and passenger size. Sometimes the original airplane is "shortened" or "shrunk" for the same reason. Manufacturers thus increase the number of units eventually built and achieve economies of scale. General aviation manufacturers are not burdened by the

Characteristics of three transport aircraft

Characteristic	Boeing 747-400	McDonnell-Douglas MD-11	McDonnell-Douglas MD-81
Maximum weight: lb	870,000	602,500	140,000
kg	394,625	273,289	63,503
Passenger capacity (approx.)	600	320	155
Range: nmi	6600	6880	1650
km	12,222	12,741	3056

tremendous development costs of major transport aircraft, and are able to produce new aircraft more often than do the major manufacturers. Thus, despite the relatively limited resources of the general aviation manufacturers, they have been able to introduce new-technology aircraft before the major manufacturers.

CRAF. The airline fleet can also support the military logistic capability of the U.S. Air Force by its participation in the Civil Reserve Air Fleet (CRAF). These aircraft are able to carry certain types of military equipment and can be pressed into such service should the need arise. The CRAF fleet consists of several hundred aircraft, many of which are modern, long-range, high-payload vehicles.

Outlook. Air transportation is expected to experience steady growth in demand. In the developed countries, increases in economic growth and disposable income have brought air travel within the reach of large numbers of people and greatly lessened the industry's former dependence on the business traveler. In the developing countries, air transportation is being used to provide rapid development of transportation capabilities at low initial cost, and also to promote understanding and cooperation between segments of a country's population previously isolated from one another.

Air transportation, however, is also beset by many increasingly severe problems. These problems will require both institutional change and more use of advanced technology if the expected growth in passenger demand is to be satisfied while still providing adequate service. Though it is always difficult to establish which problems will be most important in the future, concerns now include energy, noise pollution, ground and air congestion, and the economic deregulation of the industry (required by the Airline Deregulation Act of 1978).

Energy. The importance of energy was first brought to public attention by the Arab oil embargo of 1973, which severely disrupted the air transportation system. Government regulatory authorities took strong actions aimed at curtailing fuel usage, including fuel allocations and approval of route capacity agreements aimed at reducing duplicate airline flights. Large numbers of commercial flights were canceled, airplanes sat idle, fuel prices increased by an order of magnitude, several major airlines went bankrupt, and some were forced to sell aircraft at bargain prices to foreign competitors. As a result of this experience, many actions were taken to improve the fuel performance of aircraft. Airlines increased the number of passengers carried on each flight, instituted drag cleanup control measures, flew more fuel-efficient flight profiles, and began many other fuel-use control efficiencies. Manufacturers developed a new series of fuel-conserving airframes and engines, mainly utilizing advanced aerodynamics, materials, and engine technologies, which permits more than a 35% improvement in the airplane miles obtained per gallon of fuel. Airlines purchased hundreds of these new-generation aircraft, which began appearing in 1982. Future airplanes will show even greater long-term fuel efficiency improvements as increased government and industry research efforts in most areas of aeronautics pay off. *See* AIRCRAFT ENGINE; AIRFRAME; COMPOSITE MATERIAL.

Aircraft noise. Concern over aircraft noise and its impact on people living near airports led to major improvements in the noise characteristics and the operation of jet aircraft. Advances include "quiet" acoustic nacelle liners, improved internal and external engine and nacelle designs, and operational procedures on takeoff and approach that minimize the impact of the aircraft on the community surrounding the airport. Some of these advances were retrofitted to the existing fleet. Other advances are an integral part of the design of aircraft introduced to service since 1982. Such aircraft are capable of meeting much more restrictive government noise standards, and should result in the confinement of excessive aircraft noise to the boundaries of the airport in most areas.

Congestion. The congestion problem (both ground and air) has become increasingly severe. It has long been recognized that ground congestion reduced the time and cost savings obtained by air transport, and airport officials have continuously tried to improve access and egress to airport facilities. The air congestion problem, however, has been largely overshadowed by other problems. In the early 1970s, declining traffic growth rates, introduction of high-capacity wide-body aircraft, and a diffusion of service caused by a partial shift from the predominant hub-spoke system (where regional air traffic is funneled through a major airport) to more direct point-to-point air service temporarily masked the disruption that can result from congested air service. In the later 1970s and 1980s, air traffic surged, causing long delays, "stacking" of airplanes over airports, crowding, and sometimes chaotic conditions at the busiest terminals. Airport curfews (used as a noise control measure) also contributed to the congestion problem. Maintaining adequate air service becomes progressively more difficult as traffic continues to increase and the construction of new airports lags over long periods of time. Long-term solutions have not as yet been identified, let alone implemented. Near-term approaches include the use of additional cities as hubs.

Economic deregulation. Economic deregulation has had a far-reaching impact on the air transportation system. With deregulation (established in 1978) fully effective since 1985, major airlines either left or reduced service to the smaller cities, added capacity between larger cities (where such carriers could better utilize their existing fleet), offered significant discount fares subject to certain conditions, and, in general, increased demand for air service. *See* AIRPLANE; AVIATION.

Dal V. Maddalon

Bibliography. Aerospace Industries Association of America, Inc., *Aerospace Facts and Figures,* annually; G. L. Donohue and A. G. Zellweger (eds.), *Air Transportation Systems Engineering*, 2001; R. M. Kane, *Air Transportation*, 14th ed., 2002.

Airborne radar

Radar equipment carried by commercial and military aircraft. These aircraft use airborne radar systems to assist in weather assessment and navigation. Military systems also provide other specialized capabilities such as targeting of hostile aircraft for air-to-air combat, detection and tracking of moving ground targets, targeting of ground targets for bombing missions, and very accurate terrain measurements for assisting in low-altitude flights. Airborne radars are also used to map and monitor the Earth's surface for environmental and topological study.

Airborne radars present unique design challenges, mainly in the severe nature of the ground echo received by the radar and in the installation constraints on the size of the radar. The peculiar clutter situation governs the nature of the signal processing, and the installation limitations influence the antenna design and the radio frequency to be used (the two being strongly related) as well as the packaging of the rest of the radar. Similar considerations influence the design of space-based radars as well.

Weather assessment. A particularly valuable use of airborne radar is weather assessment (**Fig. 1**), contributing to safer and smoother navigation of severe weather areas. Two of the most serious threats to aircraft operations are turbulence and hail, both products of thunderstorms. Radars generally operating in the C or X bands (around 6 GHz or around 10 GHz, respectively) permit both penetration of heavy precipitation, required for determining the storm's extent, and sufficient reflection from less intense precipitation (stronger for the shorter wavelengths of the higher frequencies). Regions of varying but heavy precipitation extending more than 15 mi (24 km) can be clearly displayed, even if they include areas of very heavy rain (greater than 2.4 in./h or 60 mm/h). The experienced pilot can recognize structures to avoid, such as hail shafts.

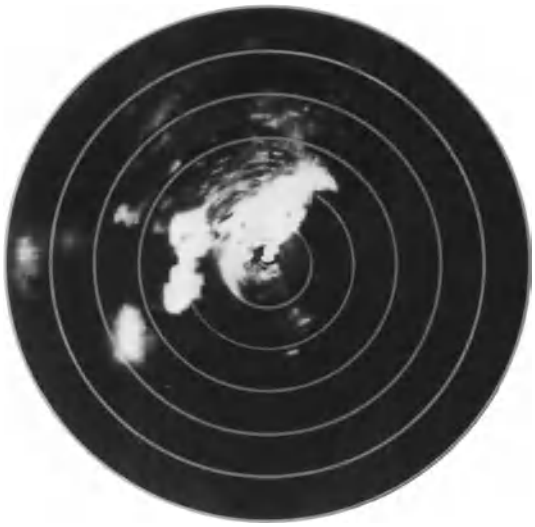

Fig. 1. Radar presentation showing a squall line extending from northwest to southeast. The range marks, indicating 5-mi (8-km) units, give estimation of extent of disturbance. (*RCA Corp.*)

It is important in such radars to detect the varying density of rain since the greatest turbulence is to be expected in the regions of the greatest rate of change or gradient of radar reflectivity, where there is a large change of reflectivity (and hence precipitation density) in a very short distance. Processing to enhance the reflectivity gradient and the resulting displays of this property, as well as the use of color coding to alert the operator, are valuable features of these radars. *See* METEOROLOGICAL RADAR; RADAR METEOROLOGY.

Radar navigation. Another basic and valuable airborne radar function is altimetry. The aircraft's altitude can be continuously measured, using (generally) C-band frequencies (around 6 GHz), low-power transmission, and a downward-oriented antenna beam. Sometimes, information from additional beams (looking somewhat forward, for example) is combined with measurements of the Doppler shift of the ground echo received to further aid in navigation. *See* ALTIMETER; DOPPLER EFFECT.

Another type of radar used in navigation is the radar beacon, in which a ground-based receiver detects an interrogation pulse from the aircraft and sends back a so-called reply on a different frequency, to which the receiver on the aircraft is tuned. The transmission of the reply at a different frequency than the interrogating signal avoids the reception of clutter echo from weather and the ground. Examples of this type of system are distance-measuring equipment (DME) and air-traffic control radar beacon systems (ATCRBS). The DMEs are used in conjunction with the network of very high frequency omnirange (VOR) stations, at each of which a transponder (transmitter-responder) has been located. By interrogating it with an inexpensive interrogator, an aircraft determines its distance from the station; a series of such measurements permits determination of the aircraft's position by trilateration. *See* DISTANCE-MEASURING EQUIPMENT; RHO-THETA SYSTEM; VOR (VHF OMNIDIRECTIONAL RANGE).

The ATCRBS involves interrogators at ground stations, generally at the sites of the major air-traffic control radars. The antennas of the two systems are aligned and rotate together to facilitate combining the radar and beacon data. When a target is detected by the radar, an interrogate pulse can be sent to the aircraft. The transponder in the aircraft sends a signal not only confirming its location but also containing other information (such as altitude and identifying number) of value to the air-traffic management system. The military "identification, friend or foe" (IFF) system operates in a similar way. These beacon systems operate in the L band (using 1.03 and 1.09 GHz frequencies for interrogation and response, respectively) and are called secondary surveillance radar (SSR) to distinguish them from ordinary two-way reflective radar, which is called primary. *See* AIR NAVIGATION; AIR-TRAFFIC CONTROL; SURVEILLANCE RADAR.

High-resolution mapping. Airborne radars are used effectively to provide high-resolution mapping of Earth's (or other planetary) surface, with a technique

Fig. 2. High-resolution radar map of a golf course near Stockbridge, New York. (*M.I.T.-Lincoln Laboratory*)

called synthetic aperture radar (SAR) [**Fig. 2**]. The processing uses the fact that surface objects produce a Doppler shift (due to the aircraft's flight) unique to their position as the aircraft passes by; this Doppler history is indicative of the scatterer's lateral, or cross-range, position at the particular range determined by the usual echo timing. With very stable radars and well-measured flight characteristics (and other focusing methods), picture cells (pixels) of 1 ft × 1 ft (0.3 m × 0.3 m) can be formed in the processed images from radars tens or hundreds of miles away. The resolution is somewhat like that possible had the flight path itself been used as a huge antenna, the synthetic aperture.

Modern SARs provide image quality that rivals aerial photography while being far less subject to conditions restricting optical visibility. For varying reasons, SARs use a wide range of frequencies, some in even the very high frequency (VHF) band, to be more sensitive to certain targets of interest; many use the familiar frequencies of 1 to 10 GHz, and some of the more resolute operate at Ka band (around 33 GHz). In some SARs, analysis of the polarization of the return (polarimetry) assists in target or surface texture (vegetation) classification. In others, interferometric analysis of the signals from two slightly different positions reveals topological (height) information for each pixel, useful in important environmental surveying and even in earthquake prediction. *See* REMOTE SENSING.

Target detection and tracking. Airborne radar provides military aircraft with an essential capability, the detection of targets, strategic and tactical, to be engaged. Strategic targets are those associated with an enemy's ability to wage war, such as munitions factories, shipyards, railroads, and other resources for delivering essential products. Tactical targets include the enemy's weapons and weapon systems brought to battle, such as fighter and bomber aircraft, warships, tanks, and other battlefield elements.

Strategic targets. Strategic targeting is accomplished by using high-resolution ground-mapping radar systems, such as those described above, to augment conventional strategic target location systems. Radar provides the ability to generate required imagery during either day or night and under all weather conditions. Typically, the strategic targets of interest for mapping radars are those that are movable (for example, deployed military equipment such as aircraft on the ground at air fields) or have been built after the most recent surveillance information was generated.

Tactical targets. Two major types of military airborne radar that deal with tactical warfare are the airborne early warning (AEW) radars and the airborne intercept (AI) radars. Of the former, the U.S. Air Force Airborne Warning and Control System (AWACS) is a good example; this S-band (around 3 GHz) system uses a large disk-shaped rotating antenna mounted atop the E-3A aircraft. AWACS provides from its typical flight altitudes a large-radius view of surface and air traffic, comprising targets of great tactical significance, and needs, therefore, excellent antenna performance (low sidelobes) and Doppler-sensitive signal processing (to detect targets moving only slightly relative to the heavy surface clutter background). The U.S. Navy's aircraft-carrier-based airborne early warning radar is borne by the E-2C Hawkeye aircraft, uses a similar but smaller top-mounted "rotodome" antenna, and operates at ultrahigh frequencies (UHF; around 400 MHz) to reduce the intensity of the surface clutter somewhat; it represents a design older than the AWACS, but has served very well in extended-area sea-surface tactical surveillance. *See* INSULATION RESISTANCE TESTING; MILITARY AIRCRAFT.

In higher-performance tactical aircraft, the radar antenna cannot possibly be mounted outside the aircraft body. Consequently, airborne intercept radars, used generally for targeting other aircraft in air-to-air combat, operate with very limited antenna dimensions (a flat disk antenna mounted behind a radome at the nose of the aircraft, for example). Therefore, to achieve useful narrow beamwidths for angle-measurement accuracy, they typically use the shorter wavelengths (about 3 cm or less) of the X band (around 10 GHz) for which the ground clutter is quite intense, placing great dependence upon sophisticated Doppler-sensitive signal processing.

Missiles are often guided by radar seekers mounted in the missile. Antenna size, limited to the diameter of the missile, suggests the use of shorter wavelengths. The X band and above (about 10 GHz or more; wavelengths less than 3 cm) are common for such radars to make best use of such small antennas. Some use continuous-wave (cw) waveforms rather than pulses, emphasizing angle sensing only, and some only receive the transmissions from a cooperating "illuminator" and are called semiactive seekers (bearing no transmitter in the missile). *See* GUIDED MISSILE.

Terrain following. For certain missions, military aircraft and attack missiles attempt to remain hidden from enemy radar defenses by flying at very low altitudes. In general, the lower the altitude, the better the shielding from the ground-based radars; but there is a limit below which the pilot's or the guidance system's capabilities may place the craft at

risk. Therefore, before the mission, speed and altitude are carefully chosen, considering the defenses at hand, foreknowledge of the terrain, and the dynamics of the aircraft. Then, during the flight, radar altimetry and a forward-looking radar are both used to establish angles to obstructions and the flight's relationship to an offset terrain (a line parallel to the terrain at the required clearance altitude); and with allowance for the flight dynamics involved, the system gives either directions to the pilot in an aircraft or inputs to the autopilot in a crewless missile. *See* GUIDED MISSILE.

Robert T. Hill

Bibliography. S. A. Hovanession, *Radar System Design and Analysis*, Artech House, Dedham, MA, 1984; W. C. Morchin, *Airborne Early Warning Radar*, Artech House, Norwood, MA, 1990; M. I. Skolnik (ed.), *Radar Handbook*, 2d ed., McGraw-Hill, New York, 1990; G. J. Sonnenberg, *Radar and Electronic Navigation*, 6th ed., 1988; G. Stimson, *Introduction to Airborne Radar*, 2d ed., SciTech Publishing, Mendham, NJ, 1998.

Aircraft

Any vehicle which carries one or more persons and which navigates through the air. The two main classifications of aircraft are lighter than air and heavier than air.

The term lighter-than-air is applied to all aircraft which sustain their weight by displacing an equal weight of air, for example, blimps and dirigibles. The weight of such aircraft is sustained by buoyant forces similar to the forces which sustain a ship in water. *See* AIRSHIP; ARCHIMEDES' PRINCIPLE; BLIMP.

Heavier-than-air craft are supported by giving the surrounding air a momentum in the downward direction equal to the weight of the aircraft. Aircraft with fixed wings impart a small downward momentum to a large quantity of air because they have a large forward velocity. Aircraft with rotating wings, such as helicopters, are also sustained by the downward momentum which they impart to the surrounding air. Because they operate at slower speeds, the momentum change is imparted to a smaller mass of air but the change in velocity is correspondingly greater. *See* AIRPLANE; BERNOULLI'S THEOREM; HELICOPTER; VERTICAL TAKEOFF AND LANDING (VTOL).

In general, aircraft have a means of support, a propulsion system to impart forward velocity, and a means of directional control so that navigation can be accomplished. *See* AIR NAVIGATION; AIRCRAFT COMPASS SYSTEM; AIRCRAFT PROPULSION.

Richard G. Bowman

Aircraft collision avoidance system

A device that reduces the risk of midair collision by providing advisories to the flight crew. It is known in the United States as a Traffic Alert and Collision Avoidance System (TCAS), and operates independently of the ground-based air-traffic control system to provide a safety net should normal separation standards be jeopardized. *See* AIR-TRAFFIC CONTROL.

TCAS advisories are of two types: traffic advisories (TAs) which aid the pilot in visually acquiring other aircraft, and resolution advisories (RAs) which recommend a vertical escape maneuver. TCAS II provides both traffic advisories and resolution advisories, and is mandated worldwide on aircraft with more than 30 passenger seats or more than 15,000 kg in weight. TCAS I provides only traffic advisories and is required by the U.S. Federal Aviation Administration (FAA) for aircraft with 10–30 passenger seats.

TCAS consists of three major subsystems: surveillance, collision avoidance system (CAS) logic, and pilot displays.

Surveillance. TCAS surveillance is based on the aircraft transponders carried by all commercial passenger and most private aircraft. The reception of a Mode A interrogation from an FAA ground radar will cause the transponder to transmit a Mode A reply that contains a 12-bit identity (squawk) code (**Fig. 1**). Mode C replies can also be solicited that contain automatically encoded 12-bit barometric altitude expressed in hundreds of feet. *See* SURVEILLANCE RADAR.

Newer Mode S transponders and ground interrogators also exist and are compatible with Mode A/C air and ground equipment, but provide a unique 24-bit address for every aircraft and an integral (that is, part of or built into) data link. TCAS interrogates both kinds of transponders to develop a surveillance picture once each second.

For Mode A/C aircraft, TCAS transmits only Mode C interrogations to obtain altitude. Range is estimated by measuring the roundtrip interrogation/reply time, and relative bearing is estimated using a small (approximately 100-cm-square or 40-in.-square) antenna. In this way, a three-dimensional picture of all nearby Mode A/C aircraft is developed each second.

Since TCAS antennas are small and have an omni or four-quadrant pattern, a "whisper-shout" interrogation sequence is used to avoid receiving multiple, overlapping replies from Mode A/C transponders at similar ranges. In this technique, TCAS first transmits a low-power P_1P_3 (Fig. 1). interrogation, resulting in replies from aircraft that have the most sensitive roundtrip link—a combination of transmit power, range, cable losses, antenna gains, and receiver sensitivities. TCAS then transmits a suppression pulse pair, P_1P_2, to suppress the transponders that previously replied, immediately followed by another P_1P_3 interrogation at a slightly higher power. The first set of aircraft will then be suppressed and not respond to the second interrogation. This sequence of suppressions and interrogations is repeated multiple times, ending in a full-power interrogation. The result is that the replies are separated, that is, do not overlap, even in dense airspace.

In contrast, Mode S surveillance interrogations are directed to individual aircraft. TCAS begins by passively listening for Mode S squitters, which are replies that are spontaneously transmitted each second by Mode S transponders and include the aircraft's unique 24-bit address. Once an aircraft's squitters are

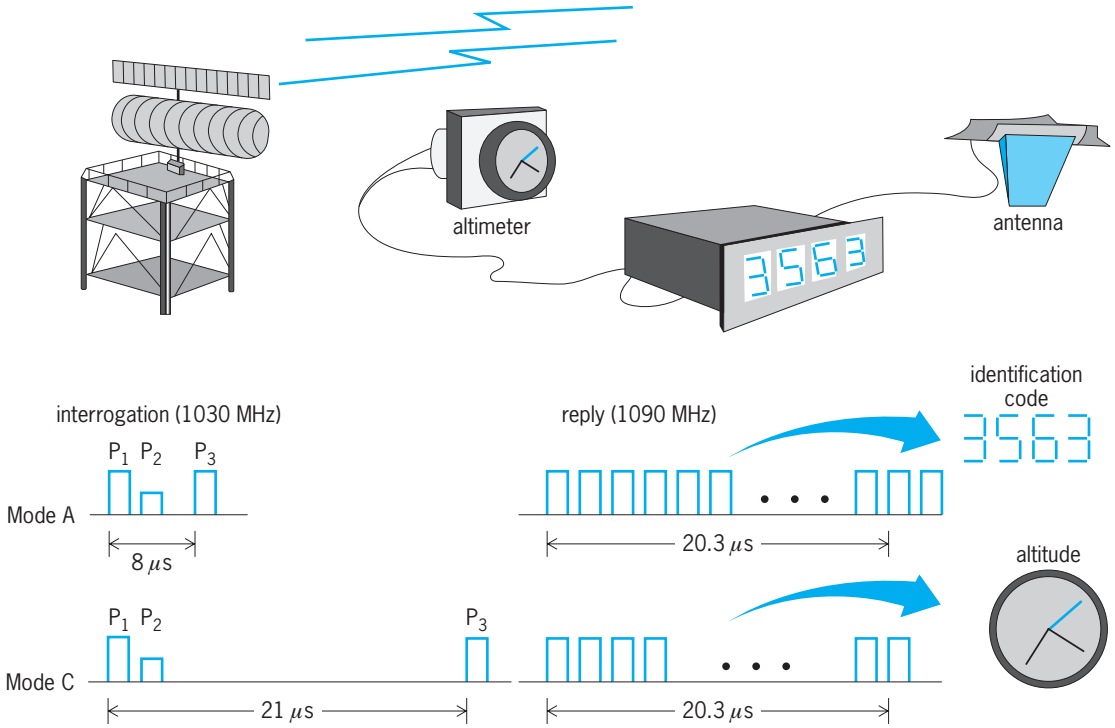

Fig. 1. Surveillance with Mode A/C transponders. (*Reprinted with permission of MIT Lincoln Laboratory, Lexington, MA*)

reliably received, TCAS begins regular interrogations of that aircraft using its address.

TCAS must provide reliable surveillance to at least 14 nautical miles (26 km) so that a flight crew can safely respond to TCAS resolution advisories for aircraft that are approaching as fast as 1200 knots (2200 km/h). In practice, surveillance is routinely achieved out to more than 50 nautical miles (93 km) in en route flight altitudes and 20 nautical miles (37 km) in the terminal area. TCAS avoids degrading the performance of air-traffic control ground interrogators by reducing the TCAS power and operating range in busy airspace and for distant intruder aircraft.

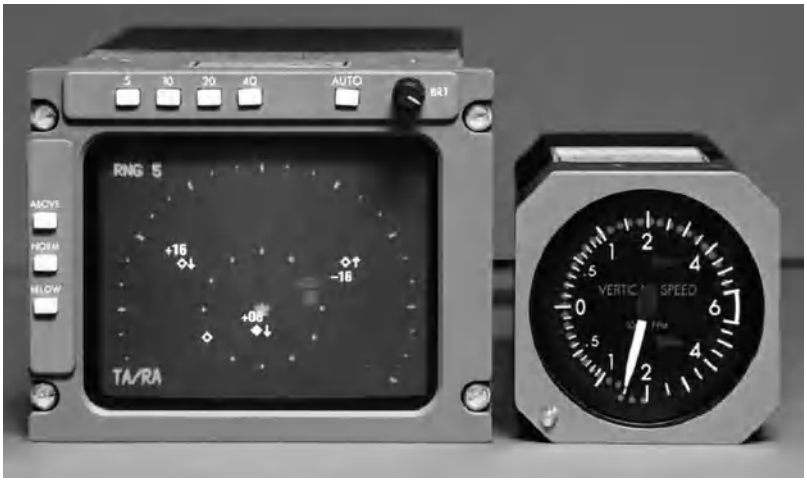

(a) (b)

Fig. 2. TCAS pilot displays. (*a*) Traffic advisory (TA) display. (*b*) Vertical resolution advisory (RA) display. (*Reprinted with permission of MIT Lincoln Laboratory*)

Collision avoidance system (CAS) logic. The CAS logic examines the position of each intruder aircraft provided by the surveillance system to determine the time to closest point of approach (CPA) for that intruder. This time is referred to as tau and is computed, in its most simple form, as range divided by range rate. Tau thresholds are based on the altitude of the encounter—the lower the altitude, the smaller the threshold. If tau is "small" (typically less than 25 seconds), TCAS computes the projected vertical separation at the closest point of approach. If the separation is also "small" (typically less than 750 ft or 230 m), TCAS issues a resolution advisory, or vertical maneuver. Traffic advisories, or warnings, are determined by the same process as above but use larger tau thresholds, typically 40 seconds.

Should a resolution advisory be deemed necessary, the least disruptive maneuver that will achieve adequate vertical separation is selected for display to the flight crew. If the encounter involves another TCAS-equipped aircraft, air-to-air coordination takes place using the Mode S data link to ensure compatible maneuvers.

The encounter is reevaluated each second to determine whether the advisory should be strengthened, sustained, weakened, or canceled. In addition, if the encounter geometry has changed and the advisory is no longer predicted to provide adequate separation, a sense reversal—climb instead of descend, or vice versa—may occur. This is most likely if a

non-TCAS-equipped intruder maneuvers unexpectedly after the resolution advisory has been issued, or if a TCAS-equipped intruder maneuvers contrary to its resolution advisory.

Since TCAS does not know the current flight plan or recent changes initiated by the flight crew or air-traffic controllers, "nuisance" advisories can occur. Early test flights and operational use led to a variety of CAS logic enhancements designed to minimize nuisance alarms while not compromising the collision avoidance performance. Pilots following a TCAS resolution advisory are instructed to notify air traffic control as soon as is practical and return to their clearance when the conflict is resolved.

Pilot displays. TCAS installations require two types of displays (**Fig. 2**). The traffic advisory display aides the flight crew in visually acquiring intruders and other nearby aircraft. Aircraft are tagged to show relative altitude and vertical rate trend, and color-coded to indicate degree of threat. The resolution advisory display indicates the vertical maneuver needed to achieve separation from the intruder. In Fig. 2*a*, the intruder at 2 nautical miles (3.7 km) and 2 o'clock is a threat, and the resolution advisory (Fig. 2*b*) tells the pilot to descend at 1500 ft (450 m) per minute. Resolution advisories may also be shown on vertical speed tapes; as pitch cues on the Primary Flight Display (PFD), using Flight Director guidance; or on a heads-up display (HUD). *See* AIRCRAFT INSTRUMENTATION.

Aural annunciations accompany both traffic advisories and resolution advisories. The annunciation for traffic advisories is "traffic, traffic." Annunciations for resolution advisories range from the basic ("climb, climb" and "descend, descend") to phrases that advise vertical rate increases or decreases, sense reversals, and altitude crossings.

Development areas. Automatic Dependent Surveillance-Broadcast (ADS-B), developed since the mid-1990s, provides for self-reporting of aircraft position, flight identification, and other aircraft-derived information at ranges up to 100 nautical miles (185 km). A Mode S extended squitter has been defined to broadcast ADS-B information and is compatible with TCAS receivers.

A method for allowing TCAS to make use of ADS-B data and still retain its independence as a safety system is referred to as TCAS Hybrid Surveillance. It allows use of extended squitter data for target surveillance only if those data have been validated by comparison with data from an active interrogation.

As a result of a midair collision near Namibia, Africa, and a heightened awareness of military safety concerns, the Department of Defense is equipping military transport aircraft with civil safety technologies, including TCAS. Special military versions of TCAS have also been developed to aid in tanker rendezvous and formation operations. The latest of these military versions make use of TCAS Hybrid Surveillance. TCAS Hybrid Surveillance is expected to make its way into the civil TCAS fleet in the future.

Raymond LaFrey

Bibliography. Federal Aviation Administration, *Aeronautical Information Manual*, FAA DoT, 4-4-10:4-4-15, 2004, published annually; International Civil Aviation Organization, *International Standards and Recommended Practices, Annex 10 to the Convention on International Civil Aviation, Aeronautical Telecommunications*, vol. IV: *Surveillance Radar and Collision Avoidance Systems*, 3d ed., 2002; Radio Technical Commission for Aeronautics, *Minimum Operational Performance Standards for Air Traffic Control Radar Beacon System/Mode Select (ATCRBS/Mode S) Airborne Equipment*, RTCA/DO-181C, 2001; Radio Technical Commission for Aeronautics, *Minimum Operational Performance Standards for Traffic Alert and Collision Avoidance System II (TCAS II) Airborne Equipment*, RTCA/DO-185A, 1997; T. Williamson and N. A. Spencer, Development and operation of the Traffic Alert and Collision Avoidance System, *Proc. IEEE*, 77(11):1735–1744, 1989.

Aircraft compass system

An instrument that indicates the bearing, or angle of the direction in which an aircraft is pointing in the horizontal plane. A compass may indicate magnetic heading or bearing, bearing referenced to a radio signal source, or bearing with respect to an inertially maintained line of position.

Inertial type. In modern commercial jet transport or military fighter aircraft, the primary sensor of the aircraft compass system is the inertial reference system (IRS). This system provides a gyroscopically derived reference to an inertial reference axis by sensing the linear and angular accelerations of the aircraft and continuously integrating these values to provide angular and linear velocities. These velocities are, in turn, integrated to develop estimates of the position and attitude of the aircraft. *See* GYROSCOPE; INERTIAL GUIDANCE SYSTEM.

The inertial reference system derives its position and attitude in an inertial frame referenced to true north. This approach represents an increase in precision over the original compass systems, which developed their lines of position referenced to magnetic north. Magnetic north is located at approximately longitude 104° west and latitude 74° north. The difference between the inertially derived north reference and the magnetically derived reference is called the magnetic declination, or magnetic variation. The magnetic variation is subject to local magnetic anomalies. Additionally, at extreme latitudes the magnetic variation can become very large, so as to make the magnetic reference unusable. *See* GEOMAGNETISM.

As navigation charting is usually referenced to magnetic values except near the poles, the aircraft converts from true to magnetic reference by accessing a magnetic variation model at the current aircraft position. As the magnetic variation value changes with time, this model must be revised approximately every 10 years.

Display. The magnetic and true bearings of the aircraft are presented on computer-generated cathode-ray tubes or flat-panel displays. These displays present aircraft heading in the form of a compass rose on both the primary flight display and the navigation display. The compass information can be presented to the flight crew in terms of either magnetic or true heading. The display can be oriented to present the data in a north-up or in a flight track-up mode (that is, north or the flight path of the aircraft points to the top of the display). References to radio navigation sources are also presented on the navigation display.

Magnetic type. The magnetic compass remains a simple, inexpensive, and reliable instrument for indicating the aircraft bearing. Limitations in the accuracy of this compass are due to the accelerations and vibrations of the aircraft, the local induced magnetic field of the aircraft, and lack of knowledge of local magnetic variation.

The magnetic compass is a secondary sensor of bearing on jet transports and most military aircraft, although it remains the primary instrument for many small, general aviation aircraft. This compass is typically a stand-alone instrument consisting of a card indicating the bearing installed in a liquid-filled case. The liquid serves to dampen rapid aircraft movements or oscillations. Magnets within the compass card align themselves with the magnetic field of the Earth. Compensators are installed to remove errors due to nearby instruments. A compass deviation card allows the navigator to correct for local magnetic variation. *See* MAGNETIC COMPASS.

In some aircraft, the magnetic compass has been coupled with a gyroscopic element to provide the gyroslaved, or gyrosynchronized, magnetic compass. The gyrosynchronized compass provides a means to overcome one limitation of the basic magnetic compass: the error in the presence of accelerations and maneuvers caused by the effect of gravity on the compass magnetic element. The gyrosynchronized compass uses a directional gyro to determine the local horizon and to sense accelerations and thus correct for the gravity effects on the compass output. *See* AIRCRAFT INSTRUMENTATION.

Robert W. Schwab

Bibliography. M. Kayton, and W. R. Fried (eds.), *Avionics Navigation Systems 2d ed.,* 1997; E. H. J. Pallett, *Aircraft Instruments, Principles and Applications,* 2d ed., 1981; G. M. Siouris, *Aerospace Avionics Systems: A Modern Synthesis,* 1993.

Aircraft design

The process of designing an aircraft, generally divided into three distinct phases: conceptual design, preliminary design, and detail design. Each phase has its own unique characteristics and influence on the final product.

Design Phases

Conceptual design activities are characterized by the definition and comparative evaluation of numerous alternative design concepts potentially satisfying an initial statement of design requirements. Conceptual design engineering is primarily analytical, making use of the body of knowledge from past aircraft designs and of computer programs for conceptual analysis based on correlations of prior design data. Experimental work, such as wind-tunnel testing, may be initiated during conceptual design if unique concepts are being considered. The conceptual design phase is iterative in nature. Design concepts are evaluated, compared to the requirements, revised, reevaluated, and so on until convergence to one or more satisfactory concepts is achieved. During this process, inconsistencies in the requirements are often exposed, so that the products of conceptual design frequently include a set of revised requirements. *See* AIRCRAFT TESTING; WIND TUNNEL.

During preliminary design, one or more promising concepts from the conceptual design phase are subjected to more rigorous analysis and evaluation in order to define and validate the design that best meets the requirements. Extensive experimental efforts, including wind-tunnel testing and evaluation of any unique materials or structural concepts, are conducted during preliminary design. The end product of preliminary design is a complete aircraft design description including all systems and subsystems. All major configuration features, size, shape, weight, and systems definitions are established and are not expected to change significantly during the final detail design. Performance and cost data are generated in sufficient depth to support the decision to proceed to the relatively lengthy and costly detail design phase.

During detail design the selected aircraft design is translated into the detailed engineering data required to support tooling and manufacturing activities.

The major design decisions determining the ultimate capabilities, production cost, and operating costs of the new aircraft are made during the conceptual design and early preliminary design phases.

Requirements

The requirements used to guide the design of a new aircraft are established either by an emerging need or by the possibilities offered by some new technical concept or invention.

Needs-based requirements, in the cases of military and commercial (airline) aircraft, are ordinarily established by the customer or user of the projected aircraft in a process which involves much interaction with aircraft companies in order to assure technical and economic feasibility of the requirements. These interactions also provide the basis for early conceptual design studies by aircraft companies. Needs-based requirements in the case of general aviation (private and business aircraft) are developed by classical market analysis techniques which involve economic projections and surveys of past and potential customers. *See* AIR TRANSPORTATION; GENERAL AVIATION; MILITARY AIRCRAFT.

Innovation-based requirements emerge from some new concept or invention which offers capabilities

not previously possible. This is illustrated by the development of the turbojet engine after World War II. This propulsion technology made it possible to design aircraft which flew higher and faster than previous generations, and at the same time provided engines which were inherently simpler and easier to maintain than the reciprocating engines which they replaced. *See* AIRCRAFT PROPULSION; JET PROPULSION; RECIPROCATING AIRCRAFT ENGINE; TURBOJET.

Requirements can be divided into two general classes: technical requirements (speed, range, payload, and so forth) and economic requirements (costs, maintenance characteristics, and so forth). A first-order list of aircraft design requirements is as shown below:

- **Technical**
  - Flight performance
    - Speed-altitude envelope
    - Range-payload
    - Takeoff and landing distances
    - Maneuvering
  - System requirements
    - Aircraft systems
    - Military mission systems
    - Environmental constraints
  - Airworthiness
    - Flying qualities
    - Structural integrity
    - Reliability
  - Military
    - Survivability
- **Economic**
  - Procurement cost
    - Aircraft
    - Spares
    - Support equipment
    - Training systems
  - Operating costs
    - Fuel usage
    - Maintainability
    - Training
  - Design life

Technical requirements. Flight performance requirements define the characteristics of speed, altitude, and payload carriage that will largely determine the size, aerodynamic configuration, and propulsion concept of the aircraft. Military combat aircraft have, in addition, stringent maneuvering requirements (for example, to permit a fighter to turn more rapidly than its opponent) that also strongly influence the size and shape of the aircraft.

The requirements for on-board aircraft systems, such as communication and navigation systems, are either defined in functional terms or may be specified in terms of selected existing systems. In the case of military aircraft, the system requirements are considerably more complex because of the demands of target detection, secure communications, special offensive and defensive avionics systems, and weapons capabilities. System requirements include the description of any constraints set by the aircraft's operating environment.

Airworthiness requirements comprise those issues having to do with safety. Flying-qualities requirements are concerned with the dynamic motions of the aircraft and its tendency to maintain a state of equilibrium, in response to pilot control inputs or to disturbances due to gusty or turbulent air. Structural-integrity requirements deal with the ability of the aircraft structure to withstand the maximum expected airloads, the life of the structure under repeated loads (fatigue life), and related issues such as corrosion resistance. Reliability requirements focus on the probabilities and consequences of failure of the many individual aircraft systems—mechanical, hydraulic, electrical, or electronic. Special military aircraft requirements include vulnerability and survivability criteria under specified hostile conditions. *See* AIRFRAME; FLIGHT CHARACTERISTICS; RELIABILITY, AVAILABILITY, AND MAINTAINABILITY.

The task of designing a set of technical requirements for a new aircraft is made easier by the existence of many requirements specification documents established over the years by government agencies. In the United States, civil aircraft are covered by various Federal Aviation Administration (FAA) specifications, and counterpart military specifications are issued by the Department of Defense. These specifications contain the basic criteria which must be met in order for a new aircraft to be permitted to operate, that is, to be certified by the FAA in the case of civil aircraft or approved for production by the military in the case of military aircraft.

Economic requirements. These are not as susceptible to quantitative statements as technical requirements but are more often stated as cost goals or as requirements related to cost, such as maintainance worker-hours per flight hour. Both procurement and operating cost requirements usually include the costs of any special or nonstandard support equipment for the new aircraft design, and also the costs of equipment and systems that may be required to train flight crew or maintenance personnel. Both commercial and military operators may also invoke some form of mission success criteria in the area of operating costs. Commercial operators, for example, might state requirements for dispatch reliability, while the military operator might require some overall measure of probability of mission success. Requirements for design life influence both procurement costs (since longer life costs more but is amortized over a longer period) and operating costs (since longer life results in more scheduled maintenance and overhaul actions).

Design Process

The design phases discussed above all involve aerodynamic, propulsion, and structural design, and the design of aircraft and mission systems. Each of these areas will be discussed, and the detail design phase will then be considered.

Aerodynamic design. Initial aerodynamic design centers on defining the external geometry and

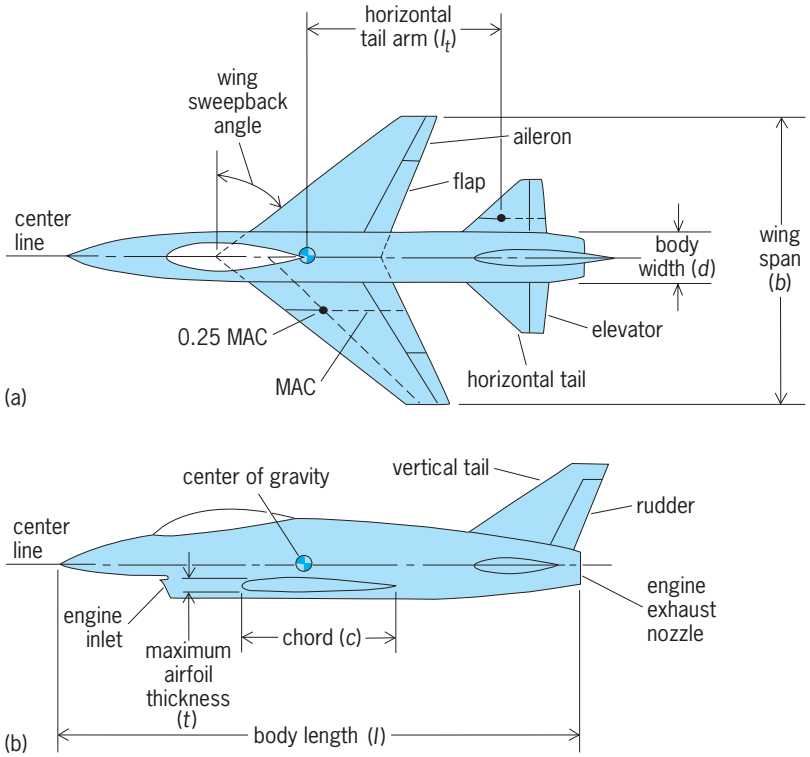

Fig. 1. Aircraft geometry and nomenclature. (*a*) Plan (top) view of aircraft. (*b*) Side view of aircraft. The terms MAC and 0.25 MAC point are explained in the text. Other important geometric parameters are the wing area S_w, the wetted area S_{wet}, and the wing aspect ratio, also defined in the text.

general aerodynamic configuration of the new aircraft. **Figure 1** defines the principal geometric parameters involved in aerodynamic design.

Lift and drag. The aerodynamic forces that determine aircraft performance capabilities are drag and lift. The basic, low-speed drag level of the aircraft is conventionally expressed as a term at zero lift composed of friction and pressure drag forces plus a term associated with the generation of lift, the drag due to lift or the induced drag. Since wings generally operate at a positive angle to the relative wind (angle of attack) in order to generate the necessary lift forces, the wing lift vector is tilted aft, resulting in a component of the lift vector in the drag direction (**Fig. 2**). *See* AERODYNAMIC FORCE; AIRFOIL; WING.

The magnitude of the friction drag force is directly related to the total surface area of the configuration exposed to the airflow, the wetted area S_{wet}. The level of pressure drag forces is determined by the relative slenderness of the body or fuselage ($1/d$ in Fig. 1) and by the relative thickness of wing and tail airfoils (t/c). The magnitude of the drag due to lift is influenced by wing area S_w, wing thickness (t/c), airfoil shape, and wing aspect ratio. [The wing area S_w is the total area of both wings, projected on the plan (top) view, and measured to the center line. The wing aspect ratio is the quantity b^2/S_w, where b is the wing span.] All of these influence the lifting efficiency at a given angle of attack and thus the size of the drag-due-to-lift vector. *See* ASPECT RATIO.

Aircraft that fly near or above the speed of sound must be designed to minimize aerodynamic compressibility effects, evidenced by the formation of shock waves and significant changes in all aerodynamic forces and moments. Because of the influence of compressibility effects, aerodynamic characteristics are usually referenced to the Mach number, the ratio of the airspeed to the speed of sound. Flight at the speed of sound is at Mach = 1.0 or sonic speed. Lower speeds or higher speeds are correspondingly termed subsonic or supersonic. Compressibility effects are mediated by the use of thin airfoils, wing and tail surface sweepback angles, and detailed attention to the lengthwise variation of the cross-sectional area of the configuration (**Fig. 3**). Smoothness of the area variation and the lowest possible peak value are crucial. *See* COMPRESSIBLE FLOW; MACH NUMBER; SHOCK WAVE; SUBSONIC FLIGHT; SUPERSONIC FLIGHT.

Those design parameters, S_w, t/c, wing aspect ratio, and airfoil shape, which were noted above as important to lifting efficiency (that is, in developing lift with minimum drag due to lift), are also the parameters which govern maximum lift capabilities. **Figure 4** illustrates two of these effects: that due to aspect ratio and that due to one of the airfoil shape parameters, camber, or the curvature of the airfoil center line (mean line).

Wing area S_w is strongly influenced by the selection of high-lift devices, wing components located in the leading-edge or trailing-edge regions or both. They augment lift by increasing the wing area or by changing the effective wing camber. **Figure 5** illustrates some typical effects of high-lift devices.

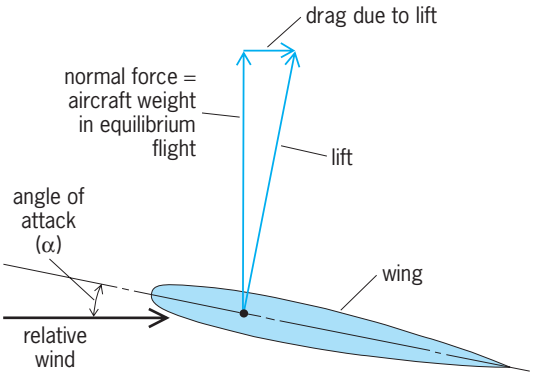

Fig. 2. Wing lift and drag due to lift.

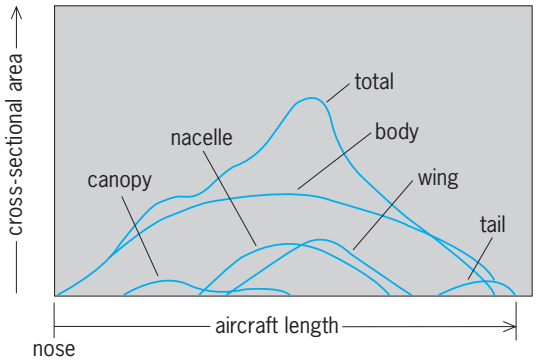

Fig. 3. Typical lengthwise variation of the cross-sectional area of an aircraft and its components.

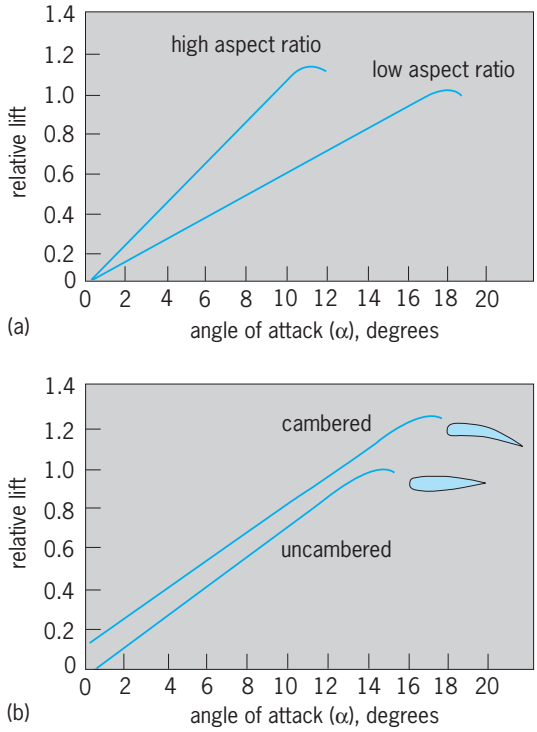

Fig. 4. Effects of wing geometry on lift. (*a*) Aspect ratio effects. (*b*) Camber effects.

Stability and control. The size and location of vertical and horizontal tail surfaces are the primary parameters that determine aircraft stability and control characteristics. In considering the conditions for longitudinal stability and control (relating to nose-up and nose-down motions), it is useful to introduce the concept of the 0.25 MAC point (Fig. 1). The mean aerodynamic chord (MAC) is defined to be the theoretical wing (or tail) chord line which divides the wing (or tail) into two areas having equal lift. The 0.25 MAC point is the intersection of the mean aerodynamic chord with the 25% chord line of the wing (or tail). Wing lift can be analyzed as a single force which typically acts at this point. If the wing is located so that this point is aft of the aircraft center-of-gravity location, then the aircraft is said to be statically stable, tail-off (without any tail contribution to stability). That is, if the aircraft is upset in the nose-up direction, then the increased wing lift (due to increased angle of attack) provides a moment about the center of gravity in the direction to restore equilibrium (that is, nose-down), and the converse is true for a nose-down upset. In this case the horizontal tail area and tail moment arm (l_t in Fig. 1) about the center of gravity are sized by the requirements for trim (to balance the moments about the center of gravity to zero in equilibrium flight) and for control to provide the required maneuvering capabilities. If the wing is located so that its lift center is forward of the center-of-gravity location, then the aircraft is statically unstable, tail-off, and sufficient horizontal tail area must be provided to restore static stability, that is, to make the tail moment about the center of gravity larger than the destabilizing wing moment. Analogous considerations apply for the case of directional (nose left–nose right) stability and control and the sizing and location of the vertical tail.

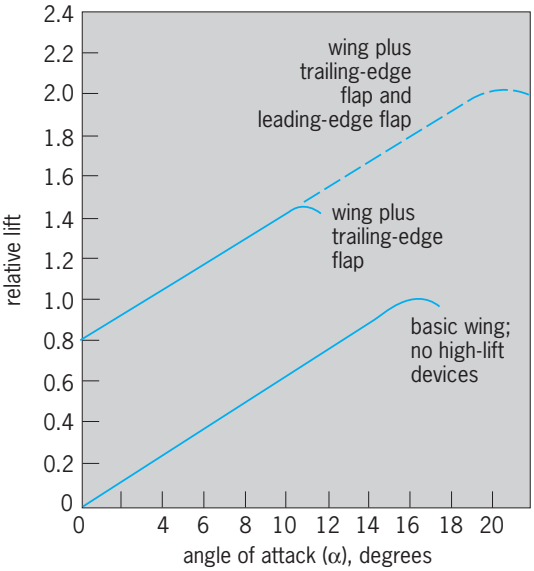

Fig. 5. Effects of high-lift devices.

Developments in digital computing and flight-control technologies have made the concept of artificial stability practical. Artificial stability systems continuously measure the aircraft's motions and activate its controls to provide inputs to change these motions to a desired state. *See* STABILITY AUGMENTATION.

To achieve the required control power, the designer must also decide how to provide for variation of the tail-surface lift forces. This is typically done by the use of the trailing-edge control surface (elevator or rudder) or by making the entire tail surfaces movable. Controls to effect rolling motions must also be provided, for example, wing trailing-edge surfaces (ailerons) or spoilers (wing upper-surface panels which are raised to deliberately spoil the lift on one wing). *See* AILERON; ELEVATOR (AIRCRAFT).

Design activities. The aerodynamic geometric parameters discussed above must be selected early in the design process and to reasonable accuracy. By the end of the conceptual design phase, these parameters should be known to within about 10%. This is necessary not only because the basics of size and performance are crucial to the technical and economic feasibility of the postulated designs, but also because the other design functions—structures, systems, and so forth—cannot proceed into much detail without accurate knowledge of aircraft size and shape constraints.

During preliminary design, the aerodynamic design activities focus on more rigorous methods of analysis together with wind-tunnel testing to refine all of the geometric parameters, to establish aerodynamic loads and pressure distributions for the structural designer, and to extend the analysis of flight performance capabilities. Through the use of a variety of analytical approximations and mathematical computational methods, aerodynamic estimates

of excellent accuracy are possible, and, with modern supercomputers, are practical in terms of time and cost. These methods have begun to replace the traditional use of wind-tunnel testing to provide aerodynamic data. However, until computational methods are developed to the point where they can provide solutions for arbitrary combinations of aircraft configuration and flight conditions, large-scale wind-tunnel testing will continue to be necessary to provide credence to the aerodynamic data.

Propulsion design. Propulsion design comprises the selection of an engine from among the available models and the design of the engine's installation on or in the aircraft.

Engine selection. Selection of the best propulsion concept involves choosing from among a wide variety of types ranging from reciprocating engine-propeller power plants through turboprops, turbojets, turbofans, and ducted and unducted fan engine developments. The selection process involves aircraft performance analyses comparing flight performance with the various candidate engines installed. In the cases where the new aircraft design is being based on a propulsion system which is still in development, the selection process is more complicated. Then the basic engine size is subject to variation, as are certain internal engine components and engine control systems, so that some degree of tailoring is possible to provide the best performance match between engine and airframe. Analysis of these expanded possibilities requires much interaction between aircraft and engine manufacturer design teams. *See* AIRCRAFT ENGINE PERFORMANCE; TURBOFAN; TURBOPROP.

Engine inlet design. Once an engine has been selected, the propulsion engineering tasks are to design the air inlet for the engine, and to assure the satisfactory physical and aerodynamic integration of the inlet, engine, and exhaust nozzle or the engine nacelles with the rest of the airframe.

Figure 6 illustrates the design parameters that govern the aerodynamic efficiency of a typical turbine engine inlet system. The major parameters to be chosen include the throat area, the diffuser length and shape, and the relative bluntness of the inlet lips. The throat area is the minimum cross-sectional area in the inlet. This area is chosen to match the engines' airflow demand with the flow of incoming air at a selected flight condition, for example, the flight condition at which the aircraft will spend the bulk of its flight time (for example, the cruise speed and altitude for a commercial airliner). At speeds or engine settings where airflow demand is either more or less than at the match point, drag penalties occur due to the nonuniformity of the incoming airflow at the inlet entrance.

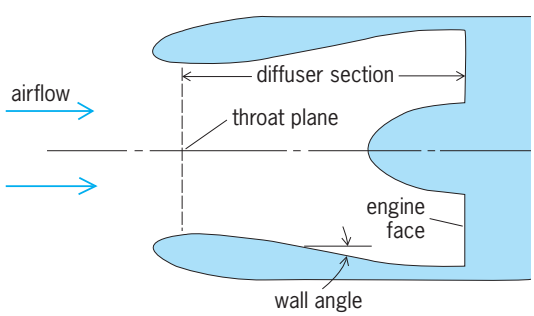

Fig. 6. Cross section of engine inlet, showing geometry and nomenclature.

Pressure distortions at the engine face adversely affect the inlets' pressure recovery (ratio of total pressure at the engine face to the total pressure available in the incoming stream). This, in turn, reduces engine thrust and in extreme cases can affect engine performance sufficiently to stall the engine. Inlet lips with a blunt leading-edge radius provide smoother airflow than sharper lip designs and operate well through a larger range of incoming flow angles (corresponding to a larger range of aircraft flight attitudes) than do sharp lip designs. The blunter lip designs, however, suffer a drag penalty at high speeds.

Diffuser geometry is governed by inlet airstream speed considerations. Turbine engines operate most efficiently with airflow velocities at the engine face of around 30–40% of the speed of sound (Mach = 0.3–0.4). In aircraft that fly at speeds much above this, the inlet must be designed to reduce the inlet stream velocity to match the engines' needs. The diffuser section accomplishes this; the airflow passing through the throat section (the minimum cross-sectional area) slows down in proportion to the increase in duct area through the diffuser to the engine face. Diffuser length becomes a function of the magnitude of aircraft design speed since the rate of increase of duct cross-sectional area is limited. Duct wall angles cannot exceed about 6° or the flow will separate (pull off) from the wall, causing unacceptable flow distortion at the engine face. *See* BOUNDARY-LAYER FLOW; DIFFUSER.

Supersonic inlet design is much more complex. The configuration shown in Fig. 6 can be varied to function as a supersonic inlet up to Mach numbers of 1.3–1.5. At higher Mach numbers, inlet pressure recovery is degraded to the point that more complicated inlet designs are necessary. In a supersonic version of the configuration of Fig. 6, the inlet lips are relatively sharper (than for subsonic design) to minimize high-speed drag, and the diffuser relatively longer because of the greater speed difference between throat and engine face conditions. At aircraft speeds of Mach = 1.0 and above, a shock wave forms just ahead of the inlet entrance. This shock wave serves to provide the initial slowdown of the incoming stream from supersonic to high subsonic speeds at the inlet entrance. Inlets designed for higher maximum Mach numbers require some form of variable geometry such as movable internal surfaces ahead of and at the throat to vary the throat area and set up multiple shock-wave systems at the inlet entrance, in order to provide the necessary subsonic flow at the diffuser entrance. *See* SUPERSONIC DIFFUSER.

Activities. As is the case with the aerodynamic design of the aircraft, the major propulsion design parameters must be established to within about 10% by the end of the conceptual design. During preliminary

design the propulsion design parameters are refined, supported by inlet model wind-tunnel tests or complete aircraft and inlet (or nacelle) model wind-tunnel tests. These tests frequently include electrically driven model engines to provide the most complete possible simulation of all airflow conditions.

Structural design. Structural design begins when the first complete, integrated aerodynamic and propulsion concept is formulated. The process starts with preliminary estimates of design airloads and inertial loads (loads due to the mass of the aircraft being accelerated during manuevers). These design loads, called static loads, are calculated at peak or maximum loading cases (combinations of speed, altitude, maneuvering condition, and aircraft weight) which are critical for a particular structural component.

Conceptual design activities. During conceptual design, the structural design effort centers on a first-order structural arrangement which defines major structural components and establishes the most direct load paths through the structure that are possible within the constraints of the aerodynamic configuration. An initial determination of structural and material concepts to be used is made at this time, for example, deciding whether the wing should be constructed from built-up sheet metal details, or by using machined skins with integral stiffeners, or from fiber-reinforced composite materials. These initial decisions are based on the new aircraft's design criteria and on weight and cost considerations from prior experiences. *See* WING STRUCTURE.

At this point there is sufficient information for mass properties analysis to begin. Initial estimates of weight and balance (center-of-gravity location) are made based on statistical correlations of data from past aircraft designs. These data support calculations to assess flight performance versus requirements. The initial weight estimates also provide a rough indicator of aircraft cost, and the initial balance estimates are critical to the accurate sizing of wing and tail areas.

Preliminary design activities. During preliminary design, the structural design effort expands into consideration of dynamic loads, airframe life, and structural integrity. Dynamic loading conditions arise from many sources: landing impact, flight through turbulence, taxiing over rough runways, and so forth. In the case of military aircraft, additional dynamic loads occur due to abrupt maneuvers and to gun firing and forced ejection of bombs and other weapons. The general effect of various dynamic load conditions is to add to material thickness requirements in localized regions. However, these loads may be severe enough to change the basic structural arrangement or choice of materials. For example, Navy aircraft designed to operate on aircraft carriers experience dynamic loadings during catapult launch and arrested landing which dominate the design of landing gears and fuselage structure. *See* LOADS, DYNAMIC.

Airframe life requirements are usually stated in terms of desired total flight hours or total flight cycles. (One takeoff, flight, and landing equals one cycle.) To the structural designer this translates into requirements for airframe fatigue life. Fatigue life measures the ability of a structure to withstand repeated loadings without failure. Design for high fatigue life involves selection of materials and the design of structural components that minimize concentrated stresses. Fatigue damage starts with a crack which propagates to the point that the affected structure cannot carry its design load. Particulars of the metal alloy selected influence resistance to crack propagation, as do the details of material processing (for example, heat treatment). These factors also influence the susceptibility to corrosion. Corrosion damage is one of the mechanisms that can initiate structural cracks (and can also influence airframe life in other ways, for example, loss of strength in a joint due to fastener corrosion). One of the benefits of fiber-reinforced composite materials for aircraft is the inherent resistance to both crack propagation and corrosion damage. Structural design to minimize stress concentrations includes the avoidance of structural discontinuities (for example, access doors cut in load-bearing structures such as wing skins) and the use of structures with multiple load paths. *See* COMPOSITE MATERIAL; CORROSION; METAL, MECHANICAL PROPERTIES OF; PLASTIC DEFORMATION OF METAL.

Structural integrity design activities impose requirements for damage tolerance, the ability of the structure to continue to support design loads after specified component failures. Fail-safe design approaches are similar to design for fatigue resistance: avoidance of stress concentrations and spreading loads out over multiple supporting structural members. Typical concepts are multispar wings, and the use of multiple bolts at splice joints rather than one or two large fasteners. Military aircraft have additional structural requirements in order to be able to survive specified damage due to hostile gunfire or missile firings.

As these structural design refinements proceed during preliminary design, the structural load and stress analysis techniques used become increasingly more detailed and accurate. The structural arrangement and selection of structural concepts are completed in sufficient detail to permit accurate sizing of structural members (**Fig. 7**). Weight and balance estimates for the structure with a statistical basis are replaced by estimates based on direct calculation of sized structural components.

As preliminary design of the aircraft structure progresses, it may become desirable to conduct laboratory testing of any unique structural components or unusual combinations of materials to verify estimated strengths and to provide confidence to proceed to detailed design. *See* STRUCTURAL DESIGN.

Aircraft systems design. Aircraft systems include all of those systems and subsystems required for the aircraft to operate. The major systems are power systems, flight-control systems, navigation and communication systems, crew systems, the landing-gear system, and fuel systems.

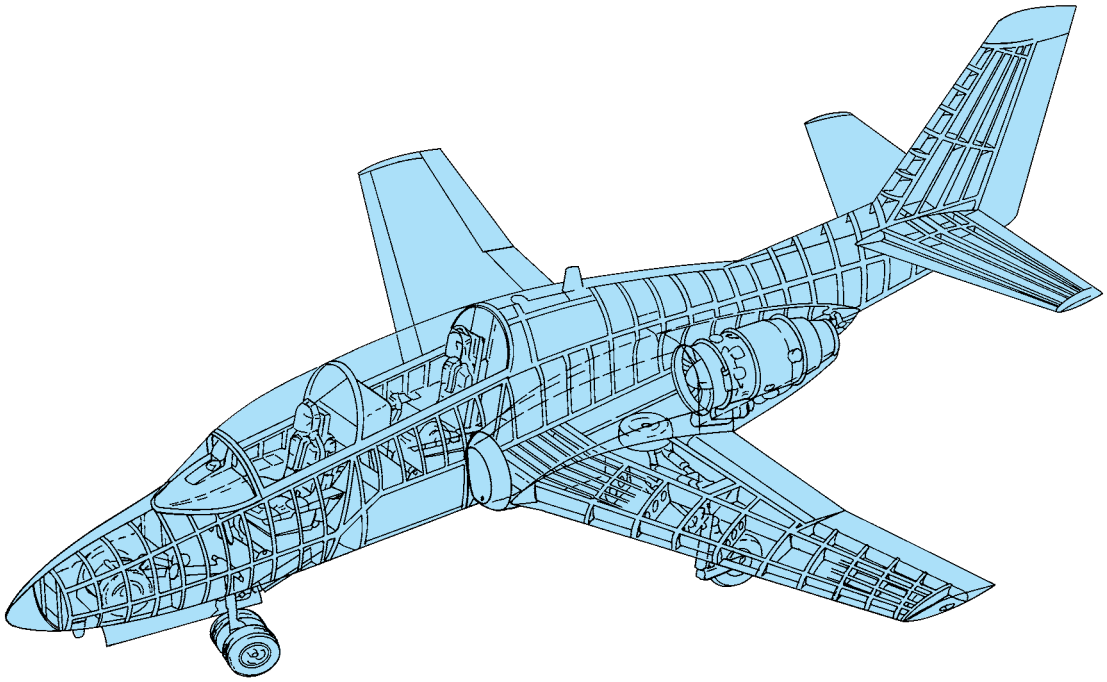

Fig. 7. Perspective drawing of a hypothetical two-place trainer aircraft, illustrating the structural detail developed during preliminary design.

Power systems comprise the electrical, hydraulic, or pneumatic power generating and power distribution subsystems to serve the other aircraft systems.

Flight-control systems comprise the subsystems which operate movable control surfaces, and provide direct (mechanical) or indirect (electrical or electronic) transmissions of pilot control signals to control surfaces. The design of autopilots and stability augmentation subsystems is also part of flight-control system design activities. *See* AUTOPILOT; FLIGHT CONTROLS.

Navigation and communication systems comprise the subsystems which provide the pilot or flight crew with information as to aircraft position and directive information to the planned destination, and the subsystems to communicate with ground stations and other aircraft as desired. *See* AIR NAVIGATION.

Crew systems comprise the cockpit or flight-deck design, including the human-factors engineering aspect of controls and displays, seat design, the subsystems which provide the crew with information as to aircraft flight and system status, and so forth. *See* AIRCRAFT INSTRUMENTATION.

The landing-gear system comprises the landing-gear structure, wheels and tires, and the means to activate the gear. *See* LANDING GEAR.

Fuel systems comprise the tanks and plumbing to distribute the fuel, pumps, and devices that measure the fuel condition.

Conceptual design. Design of these major subsystems must begin relatively early in the conceptual design phase, because they represent large dimensional and volume requirements which can influence overall aircraft size and shape or because they interact directly with the aerodynamic concept (as in the case of flight-control systems) or propulsion selection (as in the case of power systems). Early estimates must be made, to guide the aircraft design, of aircraft system features such as the size of nose-mounted radar dishes, the volume and location of fuel tanks, and the landing-gear design and location. Since many of the subsystems and much of the equipment in the various aircraft systems are purchased from suppliers, an important activity during conceptual design is preparation of technical and cost comparisons of available subsystems and equipment.

Preliminary design. During preliminary design, the aircraft system definition is completed to include the additional subsystems not yet defined. Examples include environmental-control subsystems which provide conditioned air to crew and passengers, fire detection and suppression subsystems, and various system-monitoring and alerting subsystems which detect system failures and alert the crew. The installation of the many aircraft system components and the routing of tubing and wiring through the aircraft are complex tasks which are often aided by the construction of partial or complete aircraft mock-ups. These are full-scale models of the aircraft, made of inexpensive materials, which aid in locating structural and system components.

A major systems design tool is the flight simulator. This system typically comprises a cockpit with instruments, displays, and controls; and a means for projecting an outside world view to the pilot in the cockpit. A digital computer mechanizes these hardware items and provides for simulation of arbitrary aircraft performance and system characteristics. The systems designer, by deliberately varying system characteristics, can develop optimal designs.

Selection of subsystems and equipment should be essentially complete at the end of preliminary design,

along with system mechanization concepts, installation drawings, and system reliability estimates.

Mission systems design. Mission systems are those additional systems and subsystems peculiar to the role of military combat aircraft. Offensive systems include target detection and location, weapon carriage, weapon arming and release, weapon guidance, and so forth. Defensive systems comprise various electronic countermeasures designed to confuse hostile tracking and fire-control systems, and subsystems which provide information about the location and activities of threatening aircraft or ground systems. The design process for military systems is similar to that for aircraft systems. *See* AIR ARMAMENT; ELECTRONIC WARFARE.

Detail design. Detail design engineering verifies and completes the technical definition of the aircraft configuration developed during the conceptual and preliminary design phases; translates the aircraft design data into the formats required to accomplish tooling, fabrication, and assembly processes; and, finally, provides support in resolving any difficulties that may arise in fabrication or assembly of the aircraft.

Aerodynamic and propulsion design activities center on completion of a final set of aerodynamic loads as a basis for completion of structural and flight-control system design. A final round of large-scale, powered-model wind-tunnel testing is usually conducted.

Structural design work peaks during detail design. Final analyses of static and dynamic loads, including all fatigue-life and fail-safe design considerations, are completed.

Verification of structural integrity is accomplished by extensive laboratory testing of structures, including static testing of a complete airframe during which it is loaded to simulate various critical load conditions. Additional structural integrity verifications include drop testing to demonstrate the integrity of landing gear and gear attachment structures under dynamic impact loads, and fatigue-life demonstration. These latter tests, which may involve millions of repeated load applications, usually extend from late in the detail design phase to well beyond the first deliveries of production aircraft.

Detail part design is completed, and all structural component and assembly drawings are completed and released for manufacture. The application of computer aided design and manufacturing (CAD/CAM) to this phase of the design process has grown rapidly. These technologies replace the use of drawings and various intermediate steps between drawing release and actual parts fabrication with the direct translation of engineering digital data describing a component or subassembly into software instructions to machining or assembly tools of various kinds. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Aircraft and mission systems detail design efforts include preparation of technical specifications used for procurement of purchased subsystems and equipment, and preparation of final system installation drawings or computer databases, as appropriate. Piloted simulation studies generally continue into this detail design phase with real equipment items progressively tied into the simulation in place of software representations. This provides improved verification of end-to-end performance of the simulated systems.

Detail design normally is complete with the release of all information for production. Subsequent events may still cause further design activity. These range from design changes to facilitate manufacturing improvements and changes due to problems uncovered in flight tests, to problems that may emerge late in the aircraft's service life such as premature fatigue failures. *See* AIRPLANE. Peter L. Marshall

Bibliography. G. Corning, *Supersonic and Subsonic, CTOL and VTOL, Airplane Design*, 4th ed., revised, 1979; H. D. Curtis, *Fundamentals of Aircraft Structural Analysis*, 1997; J. P. Fielding, *Introduction to Aircraft Design*, 1999; L. M. Nicolai, *Fundamentals of Aircraft Design*, 1984; C. D. Perkins and R. E. Hage, *Airplane Performance*, *Stability and Control*, 1949; D. P. Raymer, *Aircraft Design: A Conceptual Approach*, 1999; F. K. Teichmann, *Fundamentals of Aircraft Structural Analysis*, 1968; K. D. Wood, *Aircraft Design*, 3d ed., 1968.

Aircraft engine

A component of an aircraft that develops either shaft horsepower or thrust and incorporates design features most advantageous for aircraft propulsion. An engine developing shaft horsepower requires an additional means to convert this power to useful thrust for aircraft, such as a propeller, a fan, or a helicopter rotor. It is common practice in this case to designate the unit developing shaft horsepower as the aircraft engine, and the combination of engine and propeller, for example, as an aircraft power plant. In case thrust is developed directly as in a turbojet engine, the terms engine and power plant are used interchangeably. *See* TURBOJET.

The characteristics primarily emphasized in an aircraft engine are high takeoff thrust and low specific weight, low specific fuel consumption, and low drag of the installed power plant at the aircraft speeds and altitudes desired. Reliability and durability are essential, as is emphasis on high output and light weight, so that a premium is placed on quality materials and fuels, as well as on design and manufacturing skills and practices. *See* AIRCRAFT ENGINE PERFORMANCE; AIRCRAFT FUEL.

Air-breathing types of aircraft engines use oxygen from the atmosphere to combine chemically with fuel carried in the vehicle, providing the energy for propulsion, in contrast to rocket types in which both the fuel and oxidizer are carried in the aircraft. Air-breathing engines suffer decreased power or thrust output with altitude increase, due to decreasing air density, unless supercharged. *See* AIRCRAFT PROPULSION; INTERNAL COMBUSTION ENGINE; JET PROPULSION; PROPELLER (AIRCRAFT); PROPULSION;

RECIPROCATING AIRCRAFT ENGINE; ROCKET PROPULSION; TURBINE PROPULSION. Ronald Hazen

Aircraft engine performance

The power or thrust, the specific fuel consumption (mass of fuel consumed per hour per unit of delivered thrust or power), and any other important operating parameters of an aircraft engine, reported in graphical, tabular, or computerized format as a function of a thrust (or power) selection parameter, of ambient air pressure and temperature, of flight speed, and of other environmental, operational, and installation variables. *See* POWER; SPECIFIC FUEL CONSUMPTION; THRUST.

Characterization. The ultimate function of an aircraft engine is to supply the thrust for aircraft propulsion and lift. An engine's performance is most conveniently reported as the net thrust force that the engine supplies. However, in certain types of engines, such as reciprocating engines, turboshaft engines, and turboprop engines, the propulsor, that is, the thrust-producing device (for example, the rotor of a helicopter or the propeller of an aircraft), is treated as a separate system. For these cases, the aircraft engine's performance is reported as the power provided by the engine to drive the propulsor, rather than as the thrust provided by the total propulsive system. Any residual thrust from the engine exhaust is reported separately, or converted to an equivalent amount of power. For turbojet engines, where the thrust is supplied by reaction forces in the basic power unit, or for turbofan engines, where there is a propulsor integrated into the engine, the net thrust of the engine is reported directly, without reference to the power generated within the engine. *See* AIRCRAFT PROPULSION; HELICOPTER; PROPELLER (AIRCRAFT); RECIPROCATING AIRCRAFT ENGINE; TURBINE PROPULSION; TURBOFAN; TURBOJET; TURBOPROP.

Operating parameters. Thrust selection is categorized by various thrust or power ratings. These may include special high ratings whose usage is limited to short duration or under specified extraordinary situations, such as maximum contingency or automatic power reserve, invoked when one engine becomes inoperative during takeoff in a multiengine aircraft. Use of these ratings may entail consequent engine inspection or overhaul. Another category of high ratings, which includes takeoff, maximum climb, and military or combat or intermediate in a military aircraft, may be limited to a specified duration or cumulative usage in order to conserve the life of the engine. Still other ratings such as maximum cruise, normal rated, maximum continuous, and cruise are permitted for extended use and are designed to afford long life and minimum fuel consumption. Additional ratings termed flight idle and ground idle are specified to accommodate operation where power or thrust and fuel consumption are to be minimized, while maintaining the engine in a self-sustaining condition where it is available for rapid application of power.

Other reported operating parameters may be starting time, acceleration time, engine rotative speed (or speeds for multirotor engines), and pressures and temperatures within the engine that may be monitored. Engine operation is affected by inlet and exhaust system losses, ambient humidity, power extracted from the engine to run aircraft accessories, and compressed air extracted from the engine for use in aircraft systems.

Power producer operation. At the heart of any reciprocating or gas turbine aircraft engine is a thermodynamic device that converts the energy from the combustion of fuel with air to useful mechanical energy. The fuel is usually a petroleum fraction such as aviation gasoline for reciprocating engines, and kerosine or jet fuel for gas turbine engines. A typical heating value of petroleum fuels is 18,650 Btu/lb (43,350 kilojoules/kg). If 100% conversion to mechanical energy were feasible, 1 lb/h of fuel would give 7.3 horsepower or a specific fuel consumption of 0.14 lb/(h)(hp), and 1 kg/h of fuel would give 12 kW or specific fuel consumption of 0.083 kg/(h)(kW). Basic thermodynamic principles imply that no more than 35 to 40% of this theoretical value of useful power is actually achievable, implying a lower limit on specific fuel consumption of about 0.4 lb/(h)(hp) [0.24 kg/(h)(kW)]. Inefficiencies in the components of any real power producer include leakage of high-pressure working fluid (air or products of combustion); leakage of heat from the high-temperature working fluid to the lower-temperature engine bay; pressure losses in the passages through which the working fluid moves from component to component; diversion of a portion of the air to cool high-temperature parts; and inefficiencies in the compression, combustion, and expansion components of the engine. *See* AIRCRAFT FUEL.

The basic principle on which all engines operate is first to compress the air drawn into the engine inlet, then to heat the compressed air by burning fuel in it, and then to extract work from the high-pressure, high-temperature gas, to drive the compression device and then to provide useful thrust for aircraft propulsion.

Reciprocating engine. In a reciprocating engine (the Otto cycle, **Fig. 1**), the movement of working fluid through the engine is a batch process, with each cylinder accepting a slug of air once every two revolutions of the shaft. The compression is accomplished by the piston reducing the volume of the slug of air trapped in the cylinder. The pressure rise afforded by compression in the cylinder may be supplemented by a supercharger, a compressor mounted in the engine's air inlet which is usually driven by a turbine mounted in the engine's exhaust stream. Combustion is accomplished at approximately constant volume with the air fully compressed. Work extraction is accomplished by allowing the hot compressed gas to expand against the retreating piston. The cycle is the same as that of an automobile engine. *See* INTERNAL COMBUSTION ENGINE; OTTO CYCLE; SUPERCHARGER.

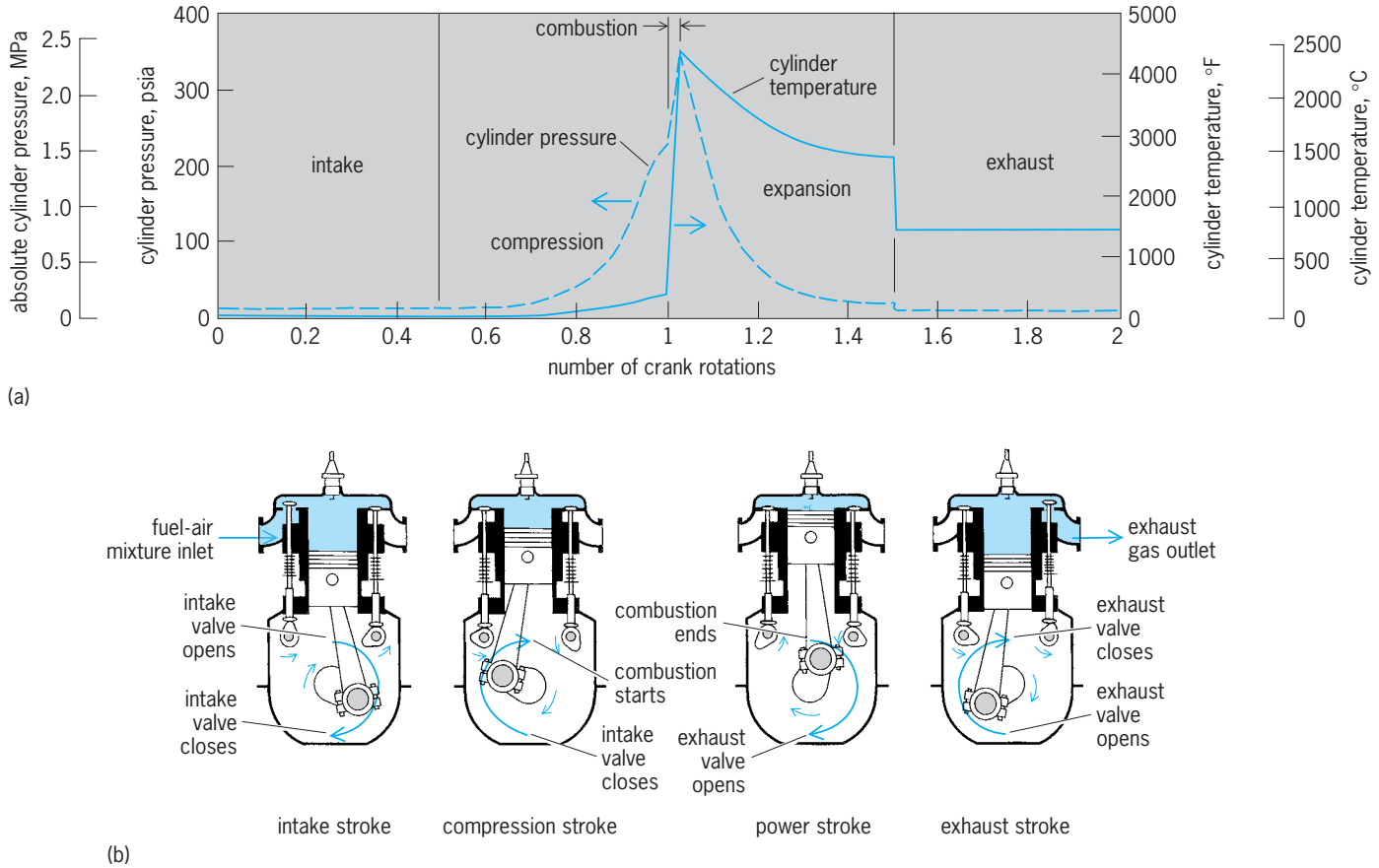

Fig. 1. Cycle of reciprocating engine. (*a*) Cylinder temperature and pressure during cycle of typical engine. (*b*) Engine configuration at five stages of the cycle. (*After R. D. Bent and J. L. McKinley, Aircraft Powerplants, 5th ed., McGraw-Hill, 1985*)

Gas turbine engine. In a gas turbine engine (the Brayton cycle, **Fig. 2**), the process is one of steady flow of the air through the components. A compressor increases the pressure of the continuous flow of air. The pressure rise may be supplemented by the propulsor (the propeller or fan) if the air passes through the propulsor on its way to the compressor. In order to achieve very high pressure ratios, some

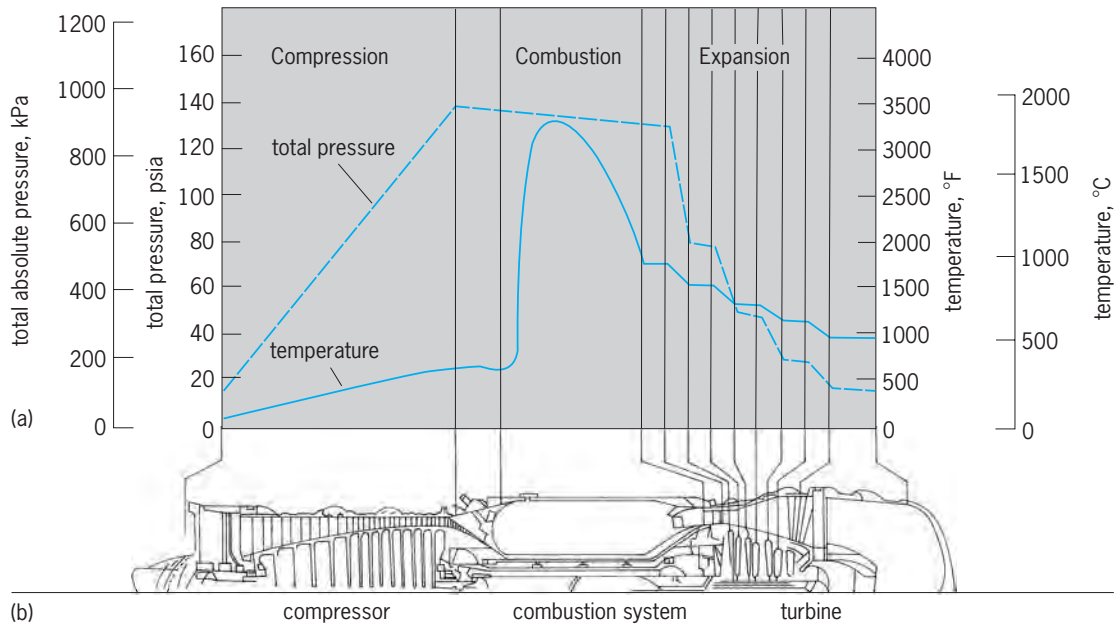

Fig. 2. Cycle of typical gas turbine (turbojet) engine. (*a*) Plot of temperature and pressure of air flowing through engine. Positions on horizontal scale are identified with locations on (*b*) diagram of engine. (*After I. E. Treager, Aircraft Gas Turbine Technology, 2d ed., McGraw-Hill, 1979*)

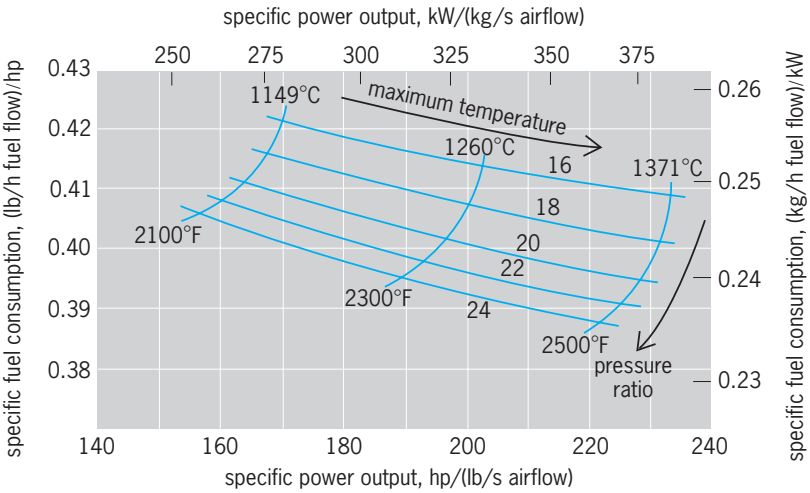

Fig. 3. Influence of cycle pressure ratio and maximum cycle temperature on performance (specific fuel consumption and specific power output) of a gas turbine (turboshaft) engine.

gas turbines use two compressors in tandem, each with its own driving turbine and each turning at its own rotative speed. Heat addition is accomplished at constant pressure in a steady-flow combustion chamber. Work extraction to drive the compressor is accomplished by steady-flow expansion through a turbine. *See* BRAYTON CYCLE; GAS TURBINE.

Ram effect. The pressure rise provided by the engine's compression device is supplemented by the ram effect of slowing down the high-velocity air stream (as seen on the aircraft) into the engine's inlet. This effect can almost double the pressure and the density of the air brought aboard at transonic flight speed, and results in considerably greater increases at supersonic flight speed.

Water injection. In certain applications, the power of the engine can be augmented for short durations by injection of water into the engine inlet. The evaporation of the water cools the compressed air to allow more fuel addition and increases the effective mass flow rate of the working fluid.

Power producer performance. The performance of a power producer is generally measured and assessed in terms of two parameters, the specific fuel consumption and the specific power. The specific fuel consumption is the mass rate of fuel flow per unit of thrust or power produced. A very low specific fuel consumption is desirable in order to minimize the cost of operation and maximize the range and performance of the aircraft by minimizing the weight of fuel it must carry. The specific power is the power produced per unit of airflow into the engine. A very high specific power is desirable since the bulk of the engine (its weight, length, and drag-inducing frontal area) is strongly dependent on the airflow of the engine. For a given required power, the bulk is minimized by increasing the specific power. These parameters are a primary function of two key variables that are selected by the engine's designers: the overall compression ratio of the cycle; and the maximum temperature of the products of combustion of the fuel and air.

Effects of cycle temperature. High maximum cycle temperature is of primary importance in achieving high specific power, and is of moderate importance in achieving low specific fuel consumption at low values of the cycle pressure ratio (**Fig. 3**). Modern gas turbines have maximum cycle temperatures between 2000 and 3000°F (1090 and 1650°C). Since the gas turbine is a continuous flow process, and the turbine is continuously bathed in air at these high temperatures, and since these temperatures often exceed the melting temperatures of the materials of which the engine is built, there is a considerable challenge to develop efficient means of cooling the structural parts, and to develop new materials that will withstand these high temperatures and still provide the durability and integrity demanded by the engine user. In a reciprocating engine, the peak cycle temperature is experienced only instantaneously one time in each cylinder for every two rotations of the crankshaft. Cycle temperatures of 4000–5000°F (2200–2760°C) are therefore feasible.

Effects of pressure ratio. High cycle pressure ratios are of considerable importance in achieving low specific fuel consumption, and have a moderately deleterious effect in reducing specific power (Fig. 3). Modern gas turbine power-producer sections, particularly those used in subsonic aircraft that do not benefit from the large ram pressure ratio experienced at supersonic flight speed, may have cycle pressure ratios of 40:1 or more, requiring considerable ingenuity and skill in the design of the compressor. Reciprocating engines may have pressure ratios of 10:1 or 11:1.

Effects of fuel flow. Thrust or power selection is usually accomplished by manipulating the fuel flow to the engine. The power producer responds to a reduction in fuel flow by reducing its rotative speed, which reduces the airflow into the engine and the power produced. In a gas turbine engine, this reduction in rotative speed is accompanied by a reduction in the pressure ratio produced by the compressor, and a considerable increase in specific fuel consumption at low power. On the other hand, the compression ratio built into the cylinder and piston of a reciprocating engine does not vary with the rotative speed of the engine, and it does not experience the steep increase in specific fuel consumption experienced by the gas turbine at low power.

Effects of ambient conditions. Ambient pressure and temperature and aircraft flight speed also have major effects on aircraft engine performance. Average ambient pressure and temperature vary considerably with altitude (**Fig. 4**) and also vary with latitude, with time of year, and from day to day. Decreasing ambient pressure at increasing altitude is accompanied by proportionate reductions in air density. Since engine geometry generally accommodates a fixed volumetric flow of air, the reduced density implies a reduction in mass flow of air and a consequent reduction in the power produced. An increase in ambient temperature also reduces air density, but more importantly, high temperature of the inlet air has a magnified effect in increasing the temperature of the compressed air, limiting the amount of fuel that may

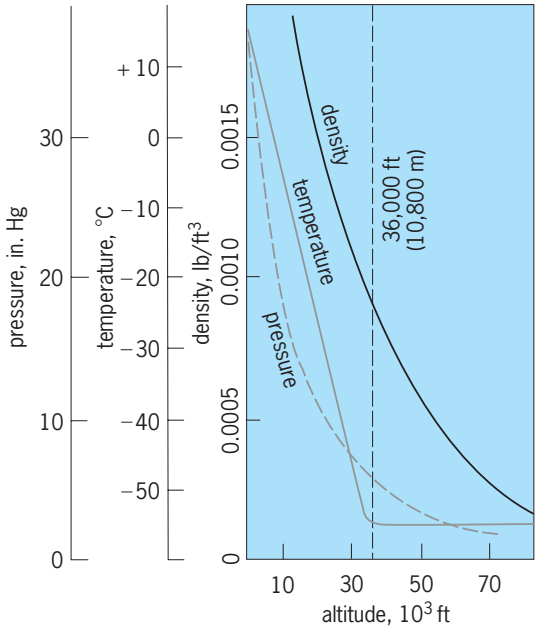

Fig. 4. Variation of ambient pressure, temperature, and density as functions of altitude. 1 in. Hg = 3.4 kPa; °F = (°C × 1.8) + 32; 1 lb/ft³ = 16.0 kg/m³; 10 M ft = 10,000 ft = 3000 m. (*After I. E. Treager, Aircraft Gas Turbine Technology, 2d ed., McGraw-Hill, 1979*)

be introduced without exceeding the limiting temperature of the turbine. Consequently, aircraft engine size must often be specified to accommodate the limiting condition of achieving the required takeoff power on the hottest day anticipated at the highest-altitude airport that might be used. Otherwise, an aircraft must have its cargo or passenger capacity restricted when these adverse conditions are encountered. *See* ATMOSPHERE.

Effects of flight speed. Aircraft flight speed also affects power producer performance, primarily by way of the ram pressure and ram temperature rise in the inlet air. In addition to the ram effects on the power producer, there are important flight speed effects on the conversion of the power to propulsive thrust, as discussed below.

Propulsor operation. After the expansion process, where work has been extracted to power the compression process—in the retreating piston of the reciprocating engine, or the turbine of the gas turbine engine—there is a surplus of energy residing in the high-pressure, high-temperature gases that is exploited to provide propulsion in either or both of two additional processes. (1) The high-pressure, high-temperature gases can be exhausted through a jet nozzle to produce thrust, as in the simple jet engine. The high-pressure gases, acting on the engine structure, provide a forward force (or thrust). The rearward force that would normally balance the forward force in a closed pressure vessel is absent in the jet nozzle by virtue of the nozzle orifice. (2) The high-pressure, high-temperature gases may be expanded in the gas turbine engine's second turbine or the reciprocating engine's cylinder to provide mechanical energy to an engine shaft which can then power a separate propulsor—a propeller or a helicopter rotor. *See* JET PROPULSION.

The turbofan exploits both of these processes. The high-pressure, high-temperature gases from the power producer are partially expanded through a turbine to power the fan (propulsor), and the residual energy is sent through a jet nozzle to produce thrust.

Afterburner. The thrust of a jet engine may be augmented by as much as 50% by burning additional fuel in a special auxiliary combustion chamber (that is, an afterburner) located just upstream of the jet nozzle. The increase in temperature of the exhaust gases to the range of 3200°F (1760°C) causes a considerable increase in exhaust jet velocity and, as discussed below, a consequent increase in thrust. Although afterburners are a relatively simple, lightweight, and inexpensive means of thrust augmentation, their specific fuel consumption is considerably higher than the basic jet engine's fuel consumption, and their use is generally limited to combat or short-duration portions of an aircraft's mission. *See* AFTERBURNER.

Newton's second law. The two propulsion processes discussed above are actually two aspects of one principle, Newton's second law of motion. That law, applied to aircraft propulsion (**Fig. 5**), provides a direct relationship between the net thrust developed by a propulsive device and the change in velocity imparted to the flow of air through the device. In Eq. (1), F is the resultant force or net thrust of the

$$F = M_e V_e - M_i V_i \tag{1}$$

propulsor; M_e is the mass flow rate and V_e is the the velocity of the gas stream exiting from the propulsor or the engine; M_i is the mass flow rate and V_i is the velocity (that is, the flight speed) of the gas stream entering the engine or propulsor; and the forces and velocity components are measured in the same direction. For engines like turbofans that have more than one exhaust stream, the jet momenta of the several streams must be added. For isolated propulsors like propellers or helicopter rotors, thrust is usually assessed by focusing on the left side of Eq. (1), by summing up the pressure forces on the blades of the propulsor. For jet engines and turbofans, it is more usual and convenient to infer the forces by

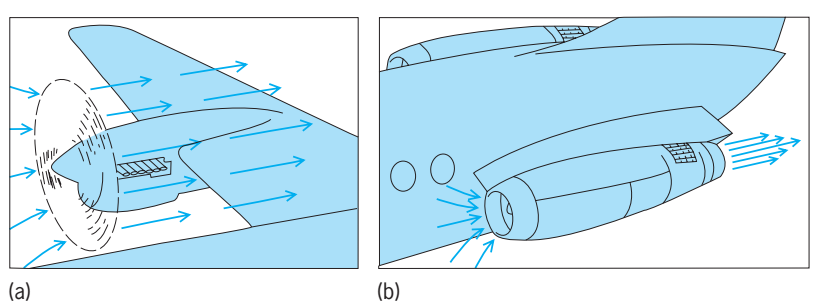

Fig. 5. Newton's second law of motion applied to two classes of aircraft propulsion. (*a*) Propeller gives a small acceleration to a large mass of air. (*b*) Turbojet engine gives a large acceleration to a small mass of air. (*After I. E. Treager, Aircraft Gas Turbine Technology, 2d ed., McGraw-Hill, 1979*)

evaluating the right side of Eq. (1). *See* FORCE; NEWTON'S LAWS OF MOTION.

Equation (1) also illustrates another important aspect of aircraft engine performance: the effect of flight speed on net thrust. The first term on the right side of Eq. (1), called the gross thrust, is the hypothetical thrust that would be developed at zero flight speed. The second term on the right side of Eq. (1), called the ram drag or induced drag, is the drag on the engine induced by the ram pressure developed on the front face of the engine. For supersonic and hypersonic aircraft, the effect is considerable, so that net thrust is a very small fraction of gross thrust, and the whole system's effectiveness is very sensitive to small changes in engine efficiency.

Propulsor performance. To complete an assessment of the performance of the total aircraft propulsion system, it is necessary to supplement the evaluation of the power producer with an evaluation of the propulsor.

Propulsive efficiency versus size. If the fuel flow rate is assumed negligible with respect to the airflow rate (it is indeed just a few percent), then, from Eq. (1), the specific thrust is given by Eq. (2), where a is the

$$\frac{F}{M_i} = V_i(a - 1) \tag{2}$$

ratio of exhaust velocity to flight speed ($a = V_e/V_i$).

Multiplication of the net thrust from Eq. (1) by the flight speed provides a measure of the useful energy developed by the engine. Comparison of this energy with the energy which the engine has imparted to the airstream provides a measure of the propulsive efficiency η_p of the device, which is given by Eq. (3).

$$\eta_p = \frac{2}{1 + a} \tag{3}$$

Equations (2) and (3) embody the basic dilemma of propulsor design. Equation (2) indicates that, for positive thrust, the velocity ratio a must be greater than 1.0. The greater the exhaust velocity, the less will be the airflow required for a given thrust, which implies increasingly lighter and smaller propulsors. On the other hand, Eq. (3) indicates that the greatest propulsion efficiency is obtained with the lowest exhaust velocity. An exhaust velocity just equal to the flight speed ($a = 1$) gives 100% propulsive efficiency but requires an infinitely large propulsor. Different propulsion systems have therefore evolved with a balanced compromise between propulsive efficiency and size.

Bypass ratio. An important approach to achieving this compromise is encompassed in the concept of the bypass ratio, β, given by Eq. (4) where M_b is that

$$\beta = \frac{M_b}{M_p} \tag{4}$$

portion of the airflow through the propulsion system which bypasses the power producer, and M_p is that portion of this airflow which goes through the power producer.

Typical bypass ratios and operational flight speed ranges of various aircraft engines

Engine type	Typical bypass ratio	Operational flight speed (Mach number)
Ramjet	0	Higher than 2
Turbojet	0	0.8–2
Low-bypass turbofan	0.1–2	0.8–1.8
High-bypass turbofan	3–8	0.7–0.9
Ultrahigh-bypass turbofan	9–12	0.5–0.85
Propfan or Unducted Fan (UDF*)	25–40	0.4–0.8
Conventional turboprop (including propeller)	50–150	0.2–0.6
Helicopter turboshaft (including helicopter rotor)	500–1700	0–0.2

*UDF is a registered trademark.

For a power producer of given size, generating a given amount of energy, the size of the propulsor may be varied. The bypass ratio can thereby be varied to optimize the propulsive efficiency for the flight speed regime of the aircraft which the engine is to power. A wide range of bypass ratios are in use (see **table**). At one extreme, for very high supersonic flight speeds (Mach numbers much greater than 1.0), where high exhaust velocities are required, turbojets are used which bypass none of the inlet airflow (that is, the bypass ratio is 0). At the other extreme, for very low-speed aircraft such as helicopters, which fly at zero speed (hover) to 0.2 Mach number, the propulsor (the main rotor used for lift and propulsion) may pass several hundred times the airflow of the turboshaft engine which powers it. As the bypass ratio increases through this range, the airflow and engine diameter also increase, while the specific fuel consumption and specific thrust both decrease. *See* MACH NUMBER.

Integrated performance. Because of its inherent mechanical, thermodynamic, and aerodynamic limitations, an aircraft engine is generally limited to operation over a flight envelope, most often specified

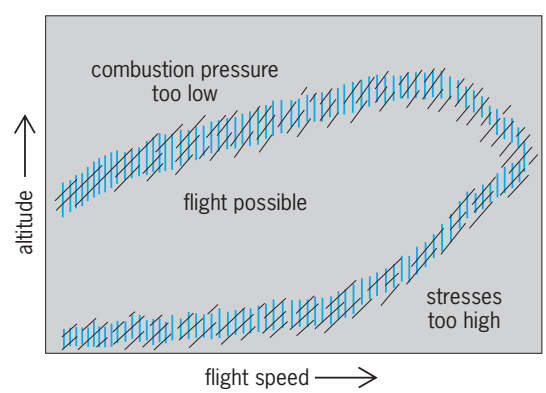

Fig. 6. Typical flight envelope for an aircraft engine. (Numerical values of quantities depend on particular engine.) (*After P. J. McMahon, Aircraft Propulsion, Barnes and Noble, 1971*)

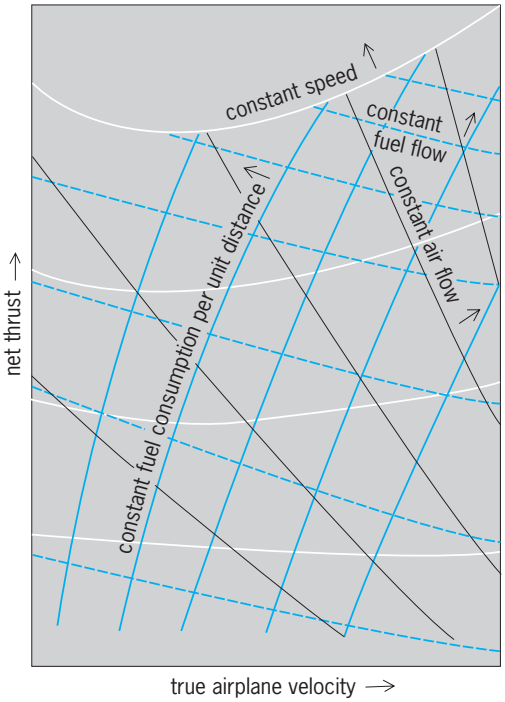

Fig. 7. Typical performance chart for a turbojet engine. (*After W. Kent, Mechanical Engineers' Handbook, vol. 4, 12th ed., Wiley, 1950*)

as the combinations of altitude and flight speed that are available for operation (**Fig. 6**). For a given engine type, of a given detailed design, it is then possible to assemble a comprehensive prediction of the engine's performance—its power or thrust and its fuel consumption—over the entire flight envelope by using a combination of data obtained from tests of the individual engine components; tests of the engine itself in a test cell which has the facility for simulating different flight speeds and ambient conditions; data from similar components or engines which may be scaled or otherwise modified to reflect the specific design being analyzed; and theoretical or analytical simulations of engine components. An excerpt from a typical display of the performance of a turbojet engine is shown in **Fig. 7**. Such performance summaries, more often in computerized format, are essentialelements in the design of aircraft, the prediction of the aircraft's performance, and the specification of the aircraft's flight plan to best accomplish its specified mission. *See* AIRCRAFT ENGINE; AIRCRAFT TESTING; GENERAL AVIATION.

Prospects. Future engines, particularly those intended for very high-speed supersonic or hypersonic flight, may be designed to use fuels with a higher ratio of hydrogen to carbon molecules, such as methane or pure hydrogen gas, to exploit their higher heating value and consequent lower required fuel weight. These gaseous fuels would be cryogenically liquefied for use as aircraft engine fuel. In engines for such high-speed aircraft, the ram pressure rise is so great that it is sufficient to provide an efficient thermodynamic cycle without being supplemented by a mechanical compressor. Such a compressorless jet engine is called a ramjet. Since they are ineffective at low flight speeds, ramjet-powered aircraft must be launched at high speed or require additional propulsion systems to take off and accelerate to high speed.

There has been considerable interest in applying the regenerator or recuperator to the gas turbine engine. This involves the use of a heat exchanger to transfer waste heat from the exhaust to the high-pressure air exiting the compressor, thereby reducing the fuel required to heat the air to the peak temperature required by the thermodynamic cycle.

Much experimental and development effort has been directed at variable-cycle, hybrid, and convertible engines, which contain mechanical devices for varying the bypass ratio and for converting turboshaft engines to turbofan engines, to power versatile airplanes which are required to perform complex missions that would otherwise require more than one type of engine.

In a completely separate class are human-powered aircraft, where the propeller or helicopter rotor is driven by bicycle pedals. *See* AIRCRAFT ENGINE; GENERAL AVIATION.

Fredric F. Ehrich

Bibliography. P. Hill and C. Peterson, *Mechanics and Thermodynamics of Propulsion*, 2d ed., 1992; J. L. Kerrebrock, *Aircraft Engines and Gas Turbines*, 2d ed., 1992; M. J. Kroes et al., *Aircraft Powerplants*, 7th ed., 1995; J. D. Mattingly, W. H. Heiser, and D. H. Daley, *Aircraft Engine Design*, 1987; G. C. Oates (ed.), *Aerothermodynamics of Gas Turbine and Rocket Propulsion*, 3d ed., 1997; Rolls Royce PLC, Technical Publications Department, *The Jet Engine*, 1986; I. E. Treager, *Aircraft Gas Turbine Engine Technology*, 3d ed., 1995.

Aircraft fuel

The source of energy required for the propulsion of airborne vehicles. This energy is released in the form of heat and expanding gases that are products of a combustion reaction that occurs when fuel combines with oxygen from ambient air. The exhaust gases are water vapor formed from hydrogen in the fuel, carbon dioxide formed from carbon in the fuel, traces of carbon monoxide and nitrogen oxides, and heated but uncombusted components of the intake air. Aircraft fuel is burned with ambient air and is thereby distinct from rocket propellants, which carry both fuel and oxidant. An important criterion for aircraft fuel is that its energy density, or heat of combustion per unit of weight, be high. This allows reasonable expenditures of fuel during takeoff, efficient performance in flight, and long range of flight duration. *See* PROPELLANT.

There are two general types of aircraft fuels in conventional use: gasolines for reciprocating (piston) engines, and kerosinelike fuels (called jet fuels) for turbine engines.

Piston engine fuels. Piston engine fuels, or aviation gasolines, are special blends of gasoline stocks and additives that produce a high-performance fuel that is graded by its antiknock quality. The gasoline blending stocks are virgin (uncracked) naphtha,

alkylate, and catalytically cracked gasoline. Naphthas are mixtures of hydrocarbons distilled directly from crude oil; alkylates are branched paraffin compounds synthesized by refining processes; and catalytically cracked gasolines contain ring compounds called aromatics (such as benzene). In general, the chemical composition of aviation gasoline can be approximated as $C_xH_{1.9x}$, where the number of carbon atoms x is between 4 and approximately 10. Tetraethyllead (Tel) is a common additive used in concentrations of up to 4 ml/gal (1.057 ml Tel/liter) of fuel to increase the antiknock quality of the fuel. *See* GASOLINE; NAPHTHA; PETROLEUM PROCESSING AND REFINING.

Antiknock quality. In reciprocating aircraft engines the fuel-to-air ratio can be varied from lean (for maximum economy) to rich (for maximum power). Fuel combustion is more likely to detonate—knock—under fuel-lean conditions. Knocking, if allowed to occur extensively, can harm an engine. Aviation gasolines, like automotive gasolines, are rated according to their antiknock quality as compared with a reference fuel (isooctane or isooctane plus specified amounts of Tel). *See* OCTANE NUMBER.

Aviation gasolines are rated by the (minimum observed) octane number for both lean and rich conditions. The American Society for Testing and Materials (ASTM) and the American National Standards Institute (ANSI) have specified standards and testing procedures for determining lean and rich knock values. For example, a fuel rated as 115/145 has a minimum octane rating of 115 when tested under fuel-lean conditions and 145 under fuel-rich conditions. Often, aviation gasolines are graded and designated by their lean octane number of 80 and a rich octane number of 87; grade 100 and grade 100LL have a lean octane number of 100 and a rich octane rating at least that of isooctane plus 1.28 ml Tel/gal (0.338 ml Tel/liter). Grade 80 may contain up to 0.5 ml Tel/gal (0.132 ml Tel/liter) and is dyed red. Grade 100 may contain up to 4 ml Tel/gal (1.057 ml Tel/liter) and is dyed yellow. Grade 100LL may contain up to 2 ml Tel/gal (0.528 ml Tel/liter) and is dyed blue.

Volatile. Aviation gasolines must be sufficiently volatile to evaporate quickly and blend with air in the engine manifolds and must be distributed evenly among all cylinders. They cannot be too volatile, however, or the fuel will boil in the tanks or lines. Gasolines boiling over a range of about 100°F (43°C) to 325°F (163°C) meet these requirements, and all grades of aviation gasolines have identical ANSI/ASTM distillation specifications. The tendency of a gasoline to boil is characterized by its Reid vapor pressure (RVP), which is approximately the absolute pressure that the gasoline will exert at 100°F (37.778°C). The RVP is between 5.5 and 7.0 lb/in.^{2} (38 and 48 kilopascals) for all grades of aviation gasoline. Aviation gasolines must also have low freezing points to be stable in storage; the ANSI/ASTM specification is −72°F (−58°C).

Heat of combustion. The heat of combustion of all grades of aviation gasoline is about 18,700 Btu/lb (43.5 megajoules/kg). This is the net or low heating value at which all combustion products are gaseous. A gallon of aviation gasoline weighs about 6.1 lb, and thus has an energy content of about 114,000 Btu (1 liter weighs about 0.73 kg and has an energy content of about 31.8 MJ).

Turbine engine fuels. Turbine engine fuels are distillate hydrocarbon fuels, like kerosines, used to operate turbojet, turbofan, and turboshaft engines. While all piston engine fuels have the same volatility but differ in combustion characteristics, jet fuels differ primarily in volatility; differences in their combustion qualities are minor. The volatility characteristics of several grades of jet fuel are shown in **Table 1**. For fuels in which the RVP is too low for accurate measurement, the flash point is given. This is the temperature to which a fuel must be heated to generate sufficient vapor to form a flammable mixture in air. The characteristics listed for Jet A and Jet B are the 1978 ANSI/ASTM standard specifications. *See* KEROSINE.

Fuel JP-1, which is no longer used, was the kerosine first used by the military and is substantially the same as Jet A or Jet A-1, which is now the most widely used commercial jet fuel. JP-4 and Jet B are military and commercial fuels, respectively, with nearly the same specifications. They are sufficiently volatile that explosive mixtures are present at most ground storage conditions and many flight conditions. JP-4 is used by the Air Force in subsonic aircraft, but Jet B has seen little commercial use in the United States. JP-5, the least volatile of the turbine fuels, is the Navy service fuel. JP-6 is the Air Force fuel for supersonic aircraft. *See* JET FUEL.

Composition. Production of distillate turbine fuel uses up to about 5% of crude oil input to a refinery. This percentage could be increased at added incremental costs and with a concurrent reduction in the

TABLE 1. Volatility characteristics of jet fuels

Jet fuel grade	Distillation range, °F (°C)	RVP, lb/in.^{2}, absolute (kPa, absolute)	Flash point, °F (°C)
JP-1	325–450 (163–230)	—	120 (49)
JP-3	100–500 (38–260)	6 (41)	—
JP-4	150–500 (65–260)	2.5 (17)	—
JP-5	350–500 (177–260)	—	150 (65)
JP-6	300–500 (149–260)	—	100 (38)
Jet A	—	—	100 (38)
Jet B	—	—	100

TABLE 2. Comparison of cryogenic fuels and Jet A fuel

Fuel	Chemical composition	Heat of combustion, Btu/lb (MJ/kg)	Boiling point, °F (°C)	Density, lb/ft^3 (kg/liter)	Specific heat capacity, Btu/lb-°F(kJ/kg-°C)
Jet A	$CH_{1.9}$	18,400 (42.8)	572 (300)*	51 (0.82)	0.47 (1.97)
Liquid hydrogen	H_2	51,600 (120.0)	−423 (−253)	4.4† (0.070)	2.3 (9.6)
Liquid methane	CH_4	21,500 (50.0)	−258 (−161)	26* (0.12)	0.84 (3.52)

*Final maximum boiling point.
†At normal boiling point.

output of motor gasoline and diesel fuel. Turbine fuel contains aromatic hydrocarbons; limits are placed on this content owing to concerns about smoke and coke formation. For military jet fuels the limit on aromatics is 25% by volume, and for commercial fuel the limit is 20% (except by mutual agreement between supplier and purchaser, in which case the content may not exceed 25% for Jet A or 22% for Jet A-1 or Jet B). Smoke can be an atmospheric pollutant, but its formation does not represent an appreciable loss in combustion efficiency. Coke is a carbonaceous deposit that adheres to the internal parts of the combustor and can reduce engine life.

High-temperature stability. An important requirement is to provide a fuel that is stable at relatively high temperatures. In subsonic jets the fuel is used to cool the engine lubricant, and the temperature of the fuel can be raised by about 200°F (110°C). In supersonic jets the fuel is used as a heat sink for the engine lubricant, for cabin air conditioning, and for cooling the hydraulic systems. For very high speed flight, the fuel may be used to cool additional engine components and critical air frame areas, such as the leading edges of wings. Therefore, depending on flight speeds and aircraft design, turbine fuels can be heated from 300°F (150°C) to 500°F (260°C) before they are burned. When they are heated to this degree, small amounts of solids may form, and foul the heat exchangers and clog the filters and fuel injectors. There are specifications to indicate the temperature at which solids are first formed and the amount of solids formed with time. In a specification test, fuel is preheated and passed through a heated filter for 5 h. No significant amount of solids may form in the preheater, and the pressure drop across the filter must stay within limits. For JP-5 and Jet A, A-1, and B, the test temperatures are 300°F (148.9°C) for the preheater and 400°F (204.4°C) for the filter. For JP-6, these temperatures are 425°F (218.3°C) and 525°F (273.9°C), respectively.

Freezing point. Turbine fuels must have low freezing points: −40°F (−40°C) for Jet A and −58°F (−50°C) for Jet A-1 and Jet B. There is also a limit on sulfur content: 0.3% by weight.

Heat of combustion. The heat of combustion of all jet fuels is about 18,400 Btu/lb (42.8 MJ/kg). This is the net low heating value. A gallon of turbine fuel weighs about 6.7 lb and thus has an energy content of about 123,000 Btu (1 liter weighs about 0.80 kg and has an energy content of about 34.4 MJ).

Alternative fuels. Alternative fuels, made from coal, oil shale, or solar or nuclear energy plus a suitable raw material, have been under consideration by the National Aeronautics and Space Administration (NASA) and several aircraft manufacturers.

Liquid hydrogen and liquid methane. Two alternative fuels are cryogenic liquid hydrogen and liquid methane. **Table 2** compares some pertinent properties of these fuels with those of the conventional Jet A fuel. As indicated in the table, liquid hydrogen has an energy density (heat of combustion) 2.8 times that of Jet A. Its volume is relatively large, and for the same energy content, the volume of liquid hydrogen would be four times that of Jet A. This, however, is of less importance than the high energy density, which is the predominant benefit of liquid hydrogen. Methane, with properties intermediate between those of Jet A and liquid hydrogen, could be more attractive than hydrogen on the basis of cost.

Hydrogen can be produced from water through electrolysis, with commercially available electrolyzers. For such production to make sense in terms of energy utilization, the electric power for operating the electrolyzers should be generated with a nonpetroleum energy source like falling water or perhaps nuclear heat. Both hydrogen and methane can be made from coal. *See* COAL GASIFICATION; HYDROGEN; METHANE.

Liquid hydrocarbons from oil shale or coal. Another option is to produce a synthetic crude or mixture of liquid hydrocarbons from oil shale or coal. When followed by refining steps, these liquefaction processes might be used to produce a synthetic turbine fuel similar to Jet A, as well as other fuels and chemicals. *See* AIRCRAFT PROPULSION; COAL LIQUEFACTION; JET PROPULSION; METAL-BASE FUEL; OIL SHALE; PROPULSION.

John B. Pangborn

Aircraft icing

Aircraft icing encompasses a range of conditions during which frozen precipitation forms on an aircraft. It is usually separated into two broad classifications, ground icing and in-flight icing. Icing can compromise flight safety by affecting the performance, stability, and control of the aircraft, and as a result the ability of the pilot to maintain the desired flight path. The primary effect of icing is its adverse impact on the aerodynamics of the airplane. Ice accretion results in

increased drag and reduced maximum lift which reduce the performance and safety of the flight. While ice does add weight to the aircraft, the amount is usually a very small percentage of the aircraft gross weight, and its effects are insignificant compared to the aerodynamic effects.

Ice accretion. Ground icing occurs when the aircraft is on the ground, and becomes significant when it affects the aircraft's ability to take off. This form of icing occurs when ice, snow, or freezing rain collects on the upper surfaces of the aircraft or when frost forms on the aircraft.

The formation of in-flight icing is much more complex. In-flight icing forms when an aircraft flies through a cloud of supercooled precipitation. Icing clouds usually contain water droplets with diameters of 2–50 micrometers and have concentrations, or liquid water content, of up to 2.5 grams of water per cubic meter of air. Water droplets approach the aircraft, approximately following the air streamlines. Near the surface of the aircraft, large changes in velocity exist. The droplets, because of their inertia, cannot change velocity rapidly enough to follow the air around the aircraft, and strike or impinge on the aircraft surface. Because of the role played by the aerodynamic velocity gradients and the droplet size in this process, more droplets impinge when they are large and the forward-facing aircraft surface has a small leading-edge radius. Ice forms on the leading or forward-facing edges of the wings, tail, antennas, windshield, radome, engine inlet, and so forth. *See* CLOUD PHYSICS.

Once the droplets have impinged, the freezing process determines the actual size and shape of the ice accretion. The two most common types of ice accretions are rime ice (**Fig. 1***a*) and glaze ice (Fig. 1*b*). Rime ice occurs at low temperature, low liquid water content, and low flight velocity so that the droplets freeze on impact. As a result, rime ice grows into the direction of the incoming droplets and forms an opaque, streamlined shape. Glaze ice occurs at temperatures near freezing with high liquid water content and high flight velocity. In this type of accretion, the impinging water droplets do not freeze on impact with the surface. Liquid water forms as a film or as hemispherical beads on the surface and eventually freezes, resulting in ice that is clear in color. The formation of beads, and sometimes larger ice roughness due to water flowing on the surface, augments the droplet impingement and accretion process, causing the large ice protrusions referred to as horns (Fig. 1*b*). Mixed ice accretions contain features of both rime and glaze ice, with glaze accretion occurring near the leading-edge stagnation point (where the air is stationary with respect to the wing), and rime ice accretion occurring farther back on the airfoil.

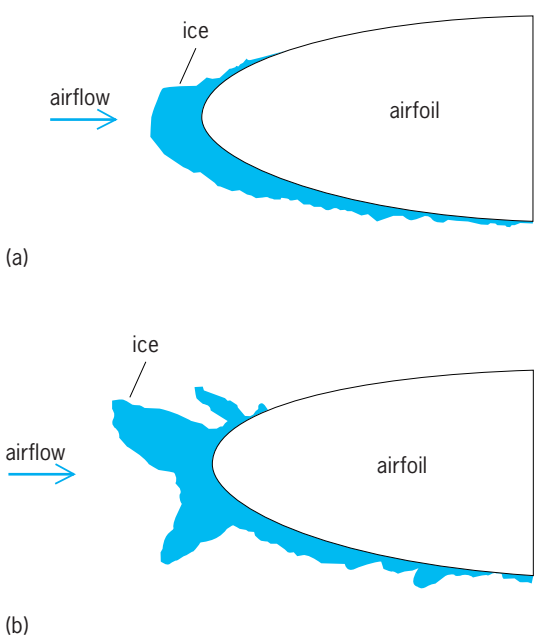

Fig. 1. Ice accretion on an airfoil leading edge. (*a*) Rime ice. (*b*) Glaze ice. *(NASA)*

Droplets much larger than 50 μm may exist in some meteorological situations, including freezing drizzle and freezing rain. These larger drops are referred to as supercooled large droplets. Ice formations from these droplets can extend very far back on the wing or other aircraft components and can pose serious safety hazards. *See* AERONAUTICAL METEOROLOGY.

Aerodynamics. Icing affects the propulsion system by reducing the thrust of the engine and thus the ability of the aircraft to climb and maintain speed or altitude. This thrust reduction is usually due to the restriction of airflow into the engine by ice formation in the carburetor of piston-engine-powered light aircraft or ice formation on the inlet of larger jet aircraft. On a propeller-driven aircraft, ice formation on the propeller can also result in reduced thrust.

Probably the most dangerous way that ice acts on an aircraft is through its effect on the aerodynamics, which results in degraded performance and control. Small amounts of ice or frost add roughness to the airplane surfaces. The roughness increases the friction of the air over the surface; this is called skin friction. Higher-than-usual levels of skin friction cause additional drag and also affect the lift (**Fig. 2**). The effect of small ice accretions or frost on lift is a reduction in the maximum lift coefficient and a reduction in the airplane angle of attack at which it is reached. Higher drag is also seen, which reduces the ability of the aircraft to climb and reduces its maximum speed. Reduced maximum lift coefficient increases the stall speed of the aircraft and thus increases the takeoff distance. Safety is compromised if ice is present and landing and takeoff speeds are not increased, since the margin between flight speed and stall speed is reduced. *See* AERODYNAMIC FORCE.

Large accretions can drastically alter the shape of the wing (Fig. 1). Then, in addition to skin friction, flow separation results in a further reduction in aerodynamic performance of the aircraft. For larger ice accretions, there is a larger reduction in maximum lift coefficient and a larger increase in drag coefficient (Fig. 2). Large ice accretions may also reduce the lift coefficient at angles of attack below stall.

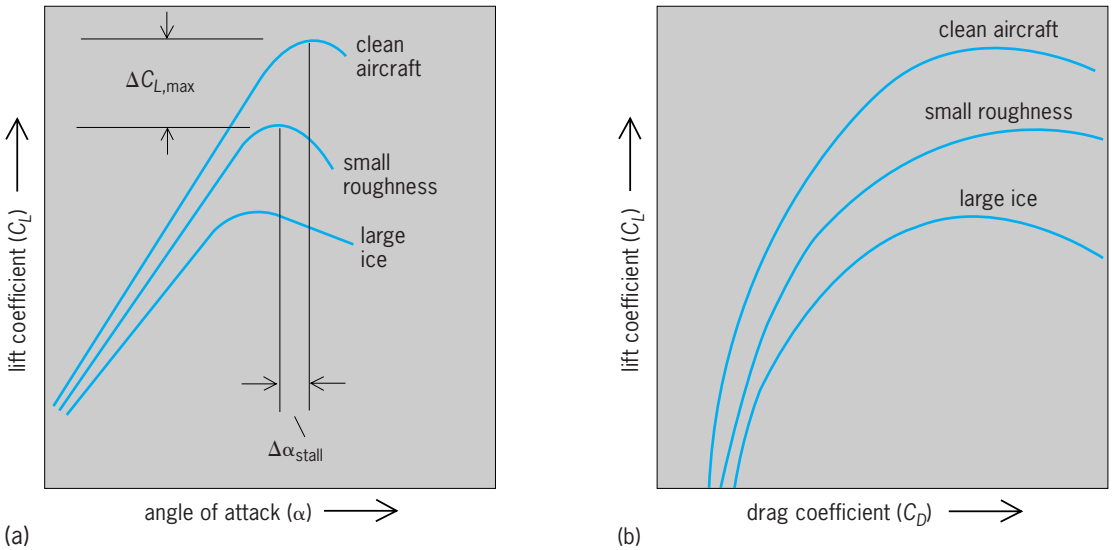

Fig. 2. Effect of ice on lift and drag coefficients. (*a*) Lift coefficient (C_L) versus angle of attack (α). $\Delta C_{L,max}$ is the reduction in maximum lift coefficient. $\Delta\alpha_{stall}$ is the reduction in the angle of attack at which stall occurs. Both are shown for small roughness case only. (*b*) Lift coefficient versus drag coefficient (C_D). (*After Effects of Adverse Weather on Aerodynamics, Proceedings of the AGARD Fluid Dynamics Specialists Meeting, AGARD-CP-496, 1991*)

Aircraft control can be seriously affected by ice accretion. Ice accretion on the tail can lead to reduced elevator effectiveness, reducing the longitudinal control (nose up and down) of the aircraft. In some situations, the tail can stall or lose lift prematurely, resulting in the aircraft pitching nose down. Similarly, ice on the wing ahead of the aileron can result in roll upset. Ice is often asymmetric and affects one wing more than another. This leads to unexpected roll, which can be difficult to control. For aircraft with unpowered controls, the pilot directly opposes the hinge moment produced by aerodynamic forces through the force applied to the control wheel. The ice-induced flow separation causes a redistribution of the air pressure on the wing or tail surface, which results in a large change in hinge moment and, therefore, in the control force required to maintain the desired flight condition. Both tail stall and roll upset are thought to be the cause of recent aircraft icing accidents. *See* AERODYNAMICS; FLIGHT CONTROLS.

Ice protection. All large aircraft and many light aircraft are equipped with in-flight ice protection systems to reduce the effect of ice. Ice protection systems are classified as de-ice or anti-ice systems. De-ice systems allow some ice to accrete, and then they periodically remove the ice. The most common system is the pneumatic boot, in which inflatable tubes are placed on the wing leading edge or other surfaces. The inflating boot expands, breaking the bond between the ice and the surface. The ice fractures into small pieces, which are carried downstream by aerodynamic forces and away from the aircraft. Anti-ice systems prevent ice from forming either by heating the surface above 0°C (32°F) or through the use of freezing-point depressants. The most common systems are electrothermal, used to protect small surfaces, and hot-air systems, used on most jet aircraft where engine bleed air is available. Wing and tail leading edges, engine inlet lips, propellers, air data system components, and windshields are typically ice protected.

Ground icing is usually dealt with by ground-based de-icing systems. Freezing-point depressant fluids are applied to the upper surface of the aircraft to remove ice and frost, and a coating remains to prevent the formation of additional ice. The fluid is effective for some period of time after application, referred to as holdover time, which depends on the type of fluid, atmospheric temperature, and rate of precipitation. The fluid is designed to flow off the aircraft wing due to aerodynamic shear during takeoff so that the fluid itself does not act as surface roughness and reduce the performance and safety of the flight. *See* ANTIFREEZE MIXTURE.

Supercooled large droplets. Until very recently, aircraft ice protection systems were tested and certified only for icing cloud droplet sizes between 2 and 50 μm. The area of the aircraft to be protected was set by the impingement location of these drops. However, during the investigation of a commuter aircraft accident which occurred in 1994, the probable cause was determined to be the resulting aerodynamic effect of ice accretion due to supercooled large droplets, which have diameters from 50 to 500 μm. Supercooled large droplets impinge much farther back from the leading edge and can cause ice to accrete aft of the ice protection systems. Research is under way to characterize the atmosphere, improve computer models, understand aerodynamic effects, and improve ice protection systems for the supercooled-large-droplet icing environment.

Michael B. Bragg

Bibliography. M. B. Bragg et al., Effect of underwing frost on a transport aircraft airfoil at flight Reynolds number, *J. Aircraft*, 31:1372–1379, 1994; A. Heinrich et al., *Aircraft Icing Handbook*, DOT/FAA/CT-88/8-1, 1991; A. Khodadoust and M. B.

Bragg, Aerodynamics of a finite wing with simulated ice, *J. Aircraft*, 32:137–144, 1995; *Proceedings of the FAA International Conference on Aircraft Inflight Icing*, DOT/FAA/AR-96/81, vol. 2, 1996.

Aircraft instrumentation

A coordinated group of instruments that provides the flight crew with information about the aircraft and its subsystems. Together with the controls, aircraft instruments form the human–machine interfaces that enable the flight crew to operate the aircraft in accordance with the flight plan. These instruments provide flight data, navigation, power plant performance, and aircraft auxiliary equipment operating information to the flight crew, air-traffic controllers, and maintenance personnel. While not considered as instrumentation, communication equipment is, however, directly concerned with the instrumentation and overall indirect control of the aircraft.

Situation information on the operating environment, such as weather reports and traffic advisories, has become a necessity for effective flight planning and decision making. The prolific growth and multiplicity of instruments in the modern cockpit and the growing need for knowledge about the aircraft's situation are leading to the introduction of computers and advanced electronic displays as a means for the pilot to better organize and assimilate this body of information.

Types of Instruments

Instrumentation complexity and accuracy are dictated by the aircraft's performance capabilities and the conditions under which it is intended to operate. Light aircraft may carry only a minimum set of instruments: an airspeed indicator, an altimeter, an engine tachometer and oil pressure gage, a fuel quantity indicator, and a magnetic compass. These instruments allow operation by a pilotage technique, that is, operation where weather conditions and visibility permit visual reference to a horizon for attitude control and navigation by topographic observation in relation to a map or the pilot's knowledge of local terrain. Operation at low altitude and under visual flight rules (VFR) allows use of very basic instruments that sense a parameter and convert that sensory information through direct mechanical means to a desired instrument reading (for example, static air pressure is converted to an altitude reading). *See* AIRSPEED INDICATOR; ALTIMETER; PILOTAGE; RATE-OF-CLIMB INDICATOR.

Operation under low visibility and under instrument flight rules (IFR) requires this same information for a more precise form and also requires attitude and navigation data. An attitude-director indicator (ADI) presents an artificial horizon, bank angle, and turn-coordination data for attitude control without external visual reference. The attitude-director indicator may contain a vertical gyro within the indicator, or a gyro may be remotely located as a part of a flight director or navigational system.

Flying through a large speed range at a variety of altitudes is simplified if the indicated airspeed is corrected to true airspeed for navigation purposes and the Mach number (M) is also shown on the ADI for flight control and performance purposes. Mach number is the ratio of the aircraft speed to the speed of sound under the existing temperature and pressure conditions. Rate-of-climb is provided by an instantaneous vertical-speed indicator (IVSI). Heading data are provided by a directional gyro or data derived from an inertial reference system. The basic flight instruments are usually located directly in front of the pilots in the form of a T. The type of information and the arrangement of the indicators on a "full panel" have evolved over the years as the airplane has increased in performance capability and complexity. *See* GYROSCOPE; INERTIAL GUIDANCE SYSTEM.

Navigation instruments. Navigation instruments primarily relate to the position of the aircraft with respect to specific locations on the Earth. Navigational aids include very high-frequency omnidirectional radio ranges (VOR) that transmit azimuth information for navigation at specified Earth locations; distance-measuring equipment (DME) that indicates the distance to radio aids on or near airports or to VORs; automatic direction finders (ADF) that give the bearing of other radio stations (generally low-frequency); low-range radio altimeters (LRRA) which by radar determine the height of the aircraft above the terrain at low altitudes; and instrument landing systems (ILS) that show vertical and lateral deviation from a radio-generated glide-path signal for landing at appropriately equipped runways. Some inertial navigation systems include special-purpose computers that provide precise Earth latitude and longitude, ground speed, course, and heading. Integration of this equipment with a coupled automatic pilot allows automatic controlled flight and landings under low visibility conditions. The pilot may eventually control the airplane manually with augmented control and advanced displays, even through landing in fog conditions. Today, however, the pilot must see outside to land manually or to monitor the automatic landing. *See* AIR NAVIGATION; AIR-TRAFFIC CONTROL; AUTOPILOT; DIRECTION-FINDING EQUIPMENT; DISTANCE-MEASURING EQUIPMENT; ELECTRONIC NAVIGATION SYSTEMS; INSTRUMENT LANDING SYSTEM (ILS); MICROWAVE LANDING SYSTEM (MLS); VOR (VHF OMNIDIRECTIONAL RANGE).

Engine instruments. Engines require specific instruments to indicate limits and efficiency of operation. For reciprocating engines, instruments may display intake and exhaust manifold pressures, cylinder head and oil temperatures, oil pressure, and engine speed. For jet engines, instruments display engine pressure ratio (EPR), exhaust gas temperature (EGT), engine rotor speed, oil temperature and pressure, and fuel flow. Vibration monitors on both types of engines indicate unbalance and potential trouble. Engine instruments tend to be clustered together. In large aircraft they are usually located in the center front instrument panel between, and visible to, both pilots. Special-purpose computers may also be

Fig. 1. Cockpit displays and flight engineer's instruments in Boeing 747. (*Boeing Co.*)

used to compute and indicate engine performance limits for the existing environmental conditions and aircraft flight mode. *See* JET PROPULSION; RECIPROCATING AIRCRAFT ENGINE; TACHOMETER.

Auxiliary instruments. Depending on the complexity of the aircraft and the facilities that are provided, there is also an assortment of instruments and controls for the auxiliary systems. Pressurized cabins require cabin altitude, cabin differential pressure, cabin rate-of-climb or descent, and cabin temperature. Secondary power system instruments and controls are provided for electrical, hydraulic, and pneumatic power and distribution systems. Fuel flow rates, fuel remaining in specific tanks, and total fuel remaining are also shown on other indicators. Many of these system instruments and control switches are strategically located on a schematic diagram on the instrument panel to provide an overview of the system configuration and operational status. Engine instruments are also included when auxiliary ground power or in-flight emergency power is provided by an auxiliary power unit (APU). Indicators show the positions of functional equipment such as landing gear, flaps, ailerons, spoilers, elevator, rudder, and trim devices. Other indicators show aircraft condition; these include fire detection and warning, acceleration monitors, overspeed deviation devices, stall warning indicators, approach-speed deviation indicators, and configuration and altitude deviation devices. These devices may provide indication through round dial or vertical tape indicators, warning lights, or audible signals. Modern computerized ground proximity warning systems (GPWS) provide warning of altitudes too close to the terrain under different flight regimes by computer-synthesized voice signals.

Large commercial transport aircraft also have flight recorders, cockpit voice recorders, and maintenance recorders. The more modern aircraft also include special built-in test equipment (BITE) that when activated on the ground conducts a prescheduled test and evaluates the response of a system or component in order to isolate problems to a specific item that is replaceable as a line replaceable unit (LRU). **Figure 1** shows the flight deck of a Boeing 747 jet transport. This flight deck is typical of those commercial jet transports built in the 1960s which used two pilots and a flight engineer.

Applications of Electronic Technology

Electronic technology developments during the 1960s and 1970s included ring laser gyros, strapdown inertial reference systems, microprocessor digital computers, color cathode-ray tubes (CRT), liquid crystal displays (LCD), light-emitting diodes (LED), and digital data buses. Application of this technology allowed a new era of system integration and situation information on the aircraft flight deck and instrument panels. Commercial jet transports developed during the 1980s use digital electronics to improve safety, performance, economics, and passenger service. The concept of an integrated flight management system (FMS) includes automatic flight control, electronic flight instrument displays, communications, navigation, guidance, performance management, and crew alerting to satisfy the

Fig. 2. Flight deck mock-up of Boeing 767 configuration for crew of three. (*Boeing Co.*)

requirements of the current and future air-traffic and energy-intensive environment. **Figure 2** shows the flight deck mock-up for a crew of three in a Boeing 767 jet transport. The following discussion is representative of the concepts used on this generation of aircraft instrumentation and controls. *See* ELECTRONICS; INTEGRATED CIRCUITS.

Flight management system. To be effective in the role of flight manager, the pilot must have ready access to relevant flight information and suitable means to accomplish aircraft control within reasonable work-load bounds. The extensive data-processing capabilities and integrated design of a flight management system provide the pilot with access to pertinent information and a range of control options for all flight phases. The elements of an integrated flight management system are shown in **Fig. 3**.

The avionics may be subdivided into three groups: sensors, computer subsystems, and cockpit controls/displays. The cockpit controls operate the sensors and computer subsystems, and the displays are supplied with raw and processed data from them. The cockpit-mounted elements of the system include caution-advisory display, attitude-director indicator (ADI), horizontal-situation indicator [HSI, including map and weather radar (WXR)], and controls for the sensors and computer subsystems.

Especially notable among the avionics systems for innovation and special benefits are the inertial reference system (IRS), the flight management computers (FMC), the electronic flight instrument system

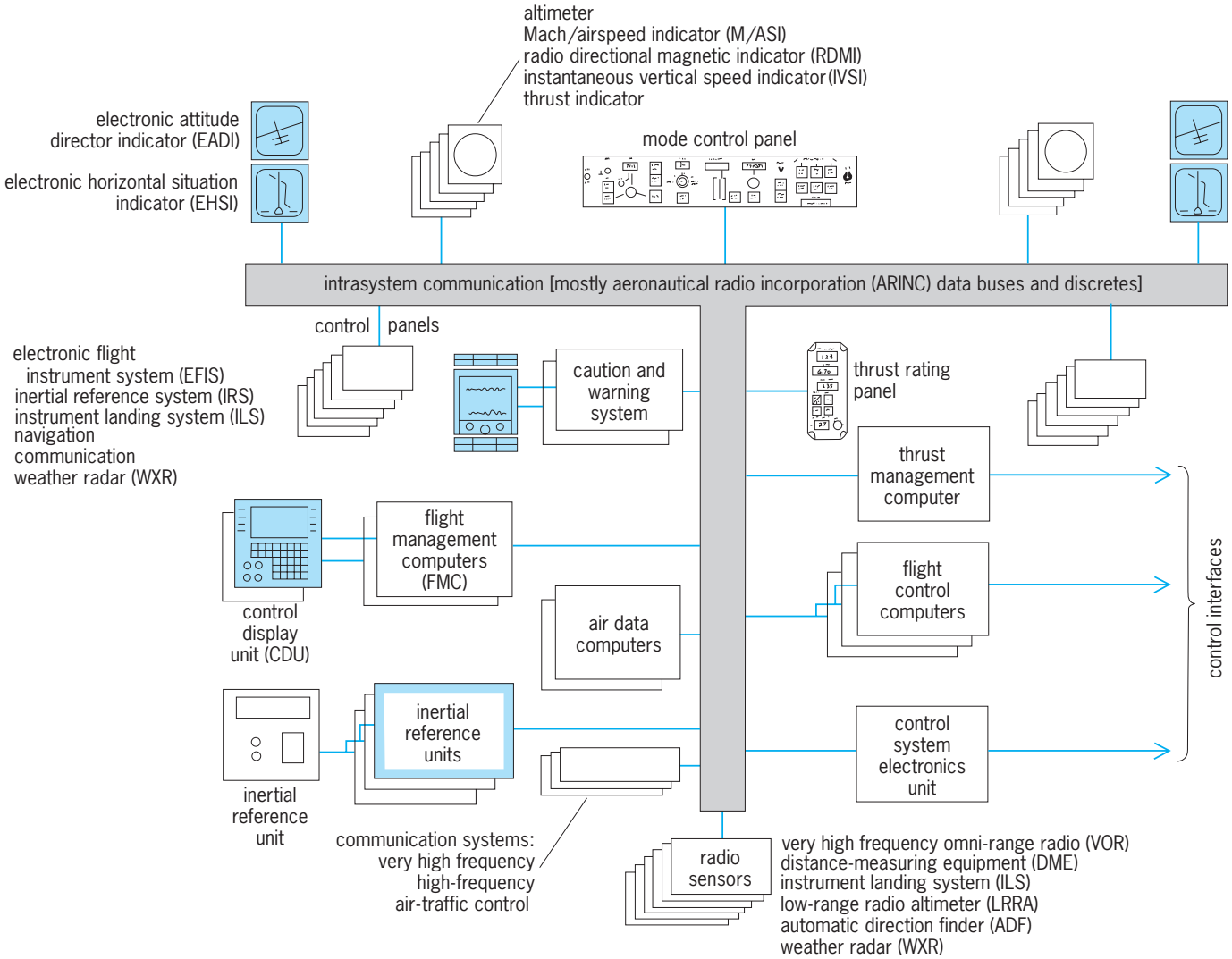

Fig. 3. Block diagram of Boeing 757/767 flight management system. (*Boeing Co.*)

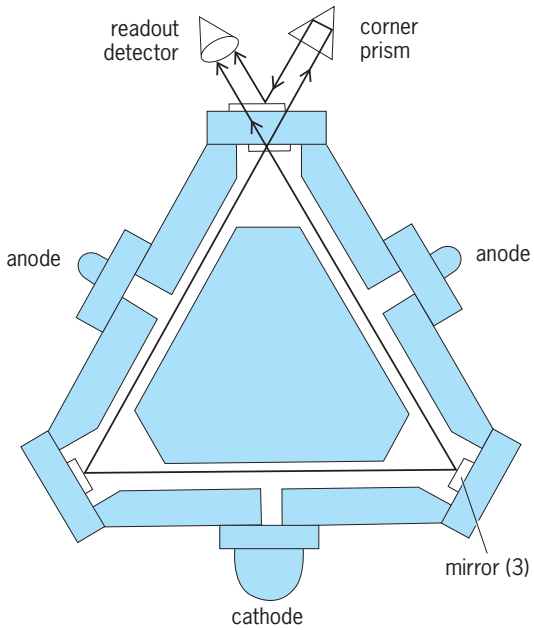

Fig. 4. Ring laser gyro operation. (*Honeywell, Inc.*)

(EFIS), and the caution and warning system.

Inertial reference system. Inertial navigation systems have benefited from two technologies developed in the 1970s: strap-down inertial techniques and the ring laser gyro. Strap-down inertial techniques eliminate the costly and bulky gimbaled stable platform previously used in high-accuracy inertial navigation systems. The laser gyro is unconventional since it does not have a spinning wheel. It detects and measures angular rates by measuring the frequency difference between two contrarotating laser beams. **Figure 4** shows how the two laser beams circulate in a triangular cavity simultaneously.

Flight management computers. Flight management computers integrate the functions of navigation, guidance, and performance management. Accurate path guidance is dependent upon display of the aircraft's current position and velocity to the crew. Position and velocity are determined by combining data from the inertial reference system with range and bearing from VOR/DME stations. The appropriate VOR/DME stations are normally automatically tuned as the aircraft progresses through its flight, but can be manually tuned at the pilot's option.

The control-display unit (CDU) is the interface between the pilot and the flight management computer. It provides the means for manually inserting system control parameters and selecting modes of operation. The control-display unit incorporates an alphanumeric keyboard for data entry and dedicated mode keys that select navigation and performance functions. In addition, it provides computer readout capability and enables pilot verification of data entered into storage. Flight plan, performance, and advisory data are continuously available for display on the control-display unit. The data are displayed on a cathode-ray tube, and control of the unit is provided by a microprocessor.

Electronic flight instrument system. Effective flight management is closely tied to providing accurate and timely information to the pilot. The nature of the pilot's various tasks determines the general types of data which must be available. The key is to provide these data in a form best suited for use. If the pilot is not required to accomplish extensive mental processing before information can be used, then more information can be presented and less effort, fewer errors, and lower training requirements can be expected. Computer-generated displays offer significant advances in this direction.

Color cathode-ray-tube displays have replaced the conventional attitude-director indicator and horizontal-situation indicator. The high-resolution, sunlight-readable, electronic displays selectively display more information with less clutter than is possible with electromechanical instruments. *See* CATHODE-RAY TUBE; ELECTRONIC DISPLAY.

An electronic attitude-director indicator (EADI; **Fig. 5**) provides a multicolor cathode-ray-tube display of information such as that found on previous attitude-director indicators. This gives attitude information showing the airplane's position in relation to the instrument landing system or a very high-frequency omnirange station. In addition, the EADI indicates the mode in which the automatic flight control system is operating and presents the readout from the radio altimeter. Ground speed is displayed digitally at all times near the airspeed indicator.

The electronic horizontal-situation indicator (EHSI; **Fig. 6**) provides an integrated multicolor map display of the airplane's position, plus a color weather radar display. The scale for the radar and map can be selected by the pilots but is always the same for both. Wind direction and velocity for the airplane's present position and altitude, provided by the inertial reference system, are shown at all times. Both the horizontal situation of the airplane and its deviation from the planned vertical path are also provided, thus making it a multidimensional situation indicator.

The EHSI operates in three primary modes—as a map display, a full compass display, and a VOR mode

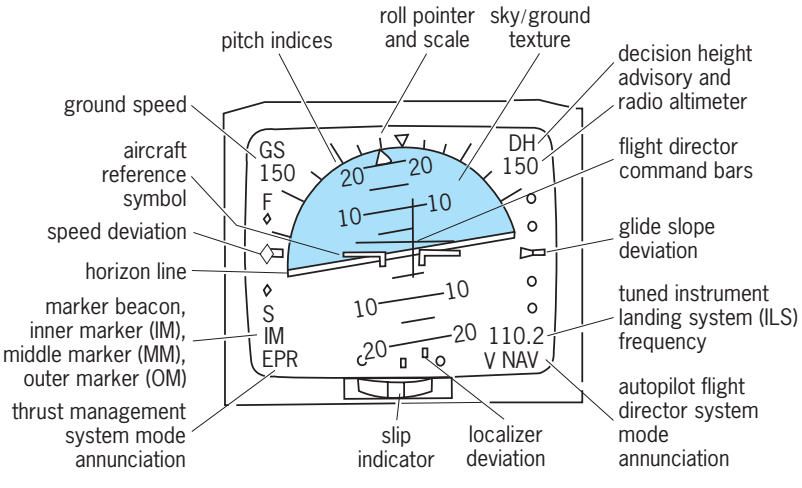

Fig. 5. Electronic attitude-director indicator (EADI). (*Boeing/Collins*)

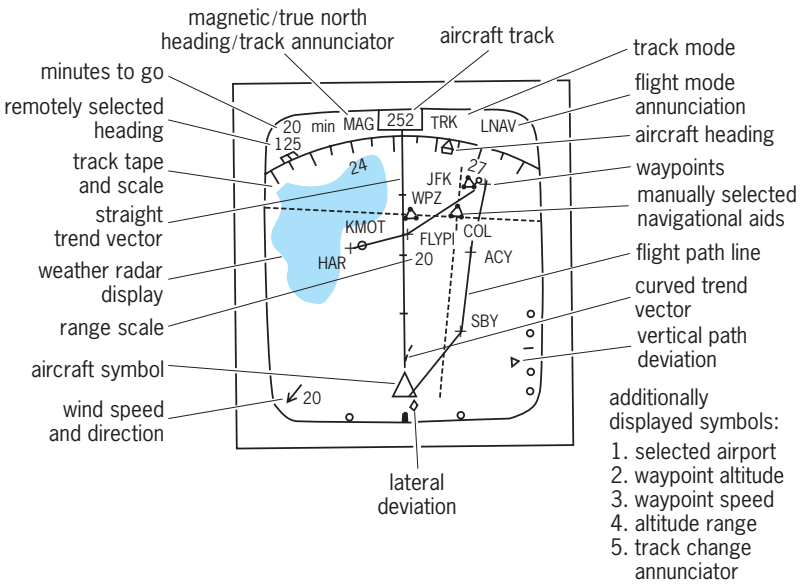

Fig. 6. Electronic horizontal-situation indicator (EHSI) map mode. (*Boeing/Collins*)

that displays a full or partial compass rose. The map displays are configured to present basic flight-plan data, including such parameters as the route of flight, planned way points, departure or arrival runways, and tuned navigational aids. Additional data such as navigation stations, airports, way-point altitude and speed targets, intersections, and weather radar returns can be called up by the pilot as the need arises, using any of several display-data push buttons.

A powerful display tool is the presentation of predictive information in the context of the basic display. The map display features two such predictions. One combines current ground speed and lateral acceleration into a prediction of the path over the ground to be followed by the airplane over the next 30, 60, and 90 s. The second prediction, an altitude range arc used for climb or descent, shows where the airplane will be when the target altitude is reached. This feature allows the pilot to quickly assess whether or not a target altitude will be reached before a particular location over the ground.

The capabilities of cathode-ray-tube displays for application to future aircraft generations include computer-generated runway imagery on the EADI, combined with other information, which may permit manually controlled landings without seeing outside at all. Also, other pertinent traffic may be portrayed on the EHSI which might allow the pilot to assure safe separation from other aircraft, control the spacing in trailing and merging situations, and avoid conflicts by early corrective action. *See* COMPUTER GRAPHICS.

Caution and warning system. The essential display elements of a typical alerting system for airplanes with a crew of three are portrayed in **Fig. 7**. The alert message display is a cathode-ray tube with multicolor capability located on the pilot's forward main engine instrument panel. Two colors are used: warnings (emergency operational or aircraft system conditions that require immediate corrective or compensatory action by the crew) are presented in red alphanumerics; cautions (conditions that require immediate crew awareness and eventual corrective or compensatory action) and advisories are presented with amber alphanumerics. The advisories (conditions that require crew awareness and may require action) are indented to distinguish them from cautions. Cancel and recall switches enable the pilot to control the density of information on the alert message display and redisplay all messages that are still valid.

The discrete warning display is a matrix of dedicated red lights that serves as a backup to the cathode-ray-tube display and provides a display of warnings when only standby electric power is available. The discrete caution display consists of

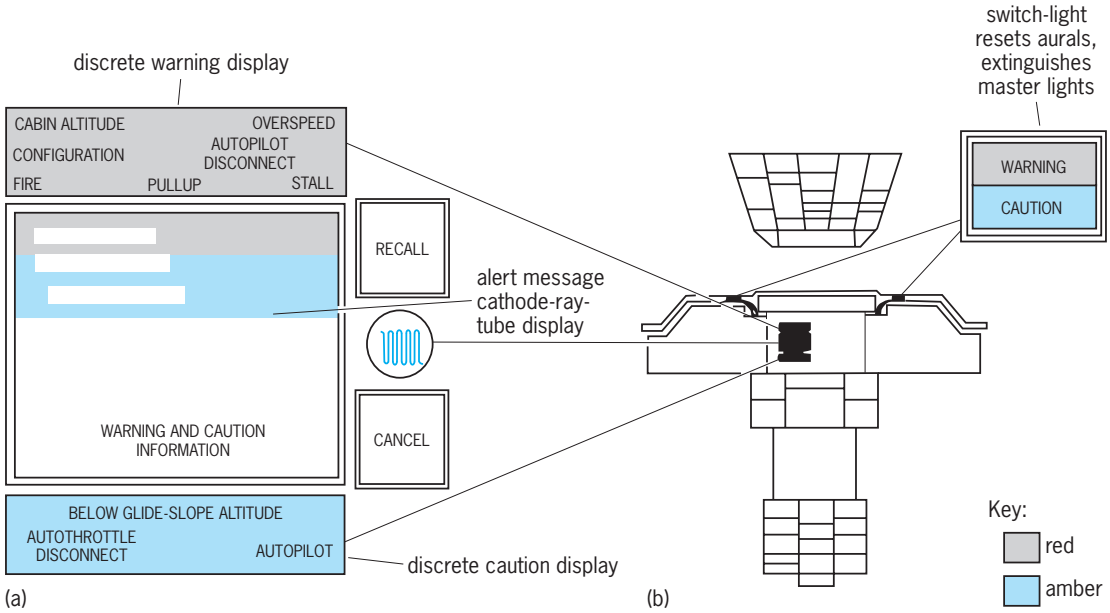

Fig. 7. Caution and warning display. (*a*) Configuration of elements. (*b*) Location of elements on aircraft instrument panel. (*Boeing Co.*)

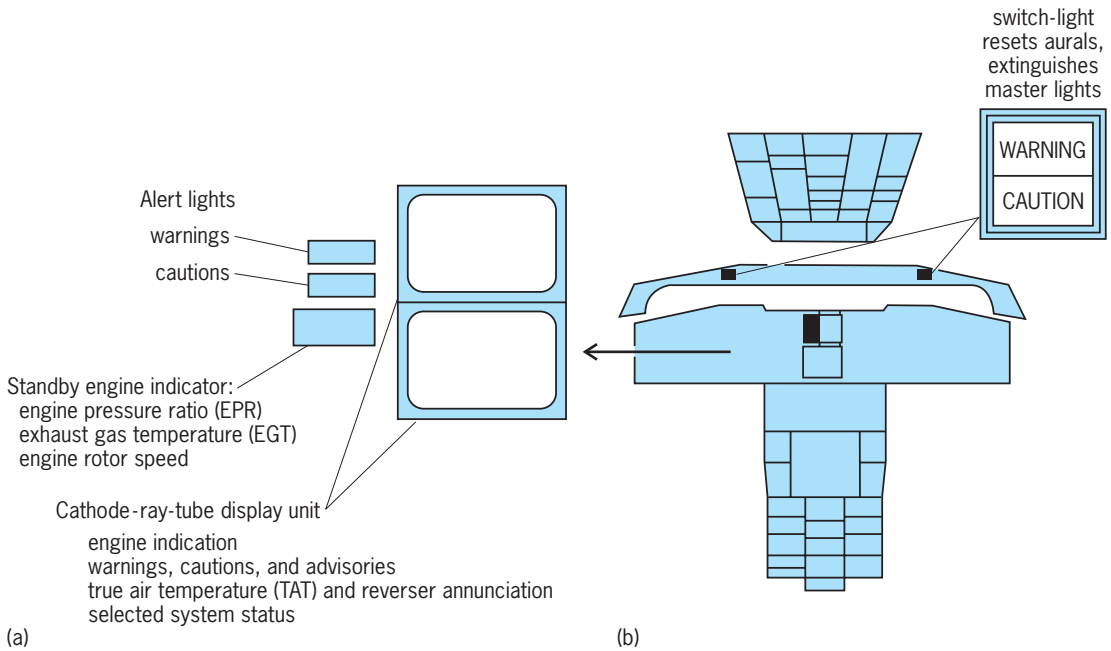

Fig. 8. Two-crew engine-indication and crew-alerting system (EICAS). (*a*) Configuration of the elements. (*b*) Location of the elements as they occur on an aircraft instrument panel. (*Boeing Co.*)

miscellaneous, not duplicated, amber lights normally located on the pilot's front panel. Master warning/master caution push-button lights illuminate whenever any warning or caution situation occurs. Both lights can be extinguished by pushing the light. Pushing the master warning light also cancels certain warning-related aural alerts.

Engine indication. An engine-indication and crew-alerting system (EICAS) is a careful consolidation of many formerly dedicated displays of propulsion parameters and subsystem indications, plus the full caution and subsystem indications, plus the full caution and warning system described above (**Fig. 8**). It provides additional automation for propulsion control and subsystem monitoring on airplanes with a crew of two.

The EICAS consolidates information on centrally located cathode-ray-tube and other displays and automatically monitors engine and subsystem parameters. During normal flight, the crew can perform any required operation with only two or three primary thrust parameters on display. Should a problem develop, the out-of-tolerance parameter is automatically called up on the display, together with other related parameters, to assist the crew in assessing the problem. A high exhaust-gas temperature, for example, will call up the closely related fuel flow indication. Any problem with engine oil pressure, quantity, or temperature will call all three of these parameters to the display. The particular item which is out of tolerance is identified because its normally white pointers and numerals are shown in amber or red, depending on the extent of the problem. These automatically displayed parameters cannot be turned off by the crew until the problem is alleviated. At any time, however, the crew can call up all engine indications via the EICAS display select panel.

A portion of one cathode-ray tube is reserved for caution and warning indications. This area is not "time-shared" by any other function; thus the appearance of any information in this area is an immediate alert to the crew.

The EICAS also features status and maintenance formats. The status format provides information concerning the airplane's status for dispatch, such as hydraulic quantity and control surface positions, and a listing of any equipment failures which could affect dispatch or routing of the flight. The maintenance format contains those parameters of use only to the maintenance crew and is not available in flight. It lists all equipment failures regardless of whether or not they affect dispatch. Bernard C. Hainline

Military Aircraft

Two types of military instrumentation requirements exist, one essentially for the instrumentation described above and one for special mission needs. Many display and control requirements for military aircraft are the same as those for commercial aircraft, including flight control, altimetry, navigation, engines, and auxiliary/systems management instruments. Additionally, a large percentage of military aircraft has much the same noncombat flight regime as the same aircraft types found in the civil sector. Accordingly, such aircraft feature much the same primary instrumentation.

Distinctive military aircraft instrumentation relates to the military mission and weapons operations. As management of these operations becomes more complex, instrumentation requirements increase in order to present the added information and control parameters. However, this results in increased workloads and attention has then focused on automation, on the simplification of display and control requirements, on the integration of information,

and on simplified display formats for easier, more accurate use.

Use of electronic displays and controls in aircraft weapon systems has led to a variety of changes in the presentation formats, which are now tailored to specific aircraft mission requirements. Situation comprehension requirements change with the mission and with emerging avionics systems, navigation aids, weapons operation, data-processing capability, and battle-management requirements. Key uses of electronic displays in this context are (1) the change of formats during the mission according to need (time sharing); (2) the tailoring of complex information into more interpretable and usable formats; and (3) the prioritizing of information (for example, caution-warning) according to critical, present need.

Types of display. Two major forms of electronic display were used initially in military systems. In one, the head-up display, a collimating lens is used in the forward field of view. (The effects of collimation are similar to those of far-focusing eyeglasses.) The resulting flight symbology is thus focused at infinity and is mixed with the infinity focus of out-the-window vision, so that near-to-far vision accommodation problems are avoided. The other type, a radar map display, presents radar reflections of ground imagery and targeting information. Two types of displays that were subsequently introduced are infrared displays, based on object temperatures, and multifunction displays, with changing, time-shared formats that present data on, for example, aircraft systems, weapons, battle situations, and integrations of tactical and strategic information. *See* AIRBORNE RADAR; INFRARED IMAGING DEVICES.

Examples. Current military instrumentation (**Fig. 9**) ranges from somewhat antiquated round-dial instruments through the sophisticated use of data-processing and electronic display and control

Fig. 9. Military aircraft instrumentation. (*a*) Soviet Mikoyan MiG-29 cockpit. (*b*) United States Navy/Marine Corps McDonnell Douglas F/A-18 cockpit (*parts a and b from Aviat. Week Space Technol., 129(12):32, September 19, 1988*). (*c*) Operators' stations for on-board mission management system of advanced strategic aircraft (*Boeing Co.*). (*d*) Advanced multirole, multifunction cockpit system (*Boeing Co.*).

systems to integrate and combine information.

For example, the cockpit of the Soviet Mikoyan MiG-29 (Fig. 9*a*) features round-dial instruments for engine and flight data and numerous stick- and throttle-mounted actuators. Primary flight-control data are on the left and weapons management information on the right. (Traditionally, primary flight information has been on the centerline.) The aircraft has a radar, an infrared display, and a head-up display (to the top of the figure, not shown). The complexity of the conventional approaches is evident, and Soviet pilots have indicated that integrating weapons operation is, in fact, difficult.

Alternatively, the United States Navy/Marine Corps McDonnell Douglas F/A-18, developed in the same period as the MiG-29, incorporates one of the most advanced cockpits of United States fighter attack aircraft. It has a head-up display, three format-interchangeable cathode-ray tube displays, and small round-dial instruments for backup. Primary flight control and aiming data are presented on the head-up display. Communication and navigation information is directly below, on the control panel. The lowest central display is a moving map, an electronic map of the area moving below the airplane. The other displays present weapons management and radar data. Format integration and color coding of the electronic displays improve pilot performance time, accuracy, and workload.

Figure 9*c* shows the work stations of an on-board mission management system for the offensive and defensive operations of an advanced strategic aircraft. This system employs artificial intelligence and expert systems software to recommend tactics and courses of action, and full-color, high-resolution graphic displays with programmable touch panels. Mapping, targeting, weapons management, flight reference data, and control modes are shown. *See* ARTIFICIAL INTELLIGENCE; EXPERT SYSTEMS.

Development has been undertaken of an advanced multirole, multifunction cockpit system that features a head-up display with advanced graphic pictorials (Fig. 9*d*). Lower displays show threat zones ahead, and an overall situation display of area, threat, path, and target information. The weapons management displayed on the left represents stores, location, and ready status. To the right is an integrated engine and energy management display. Donald L. Parks

Aircraft propulsion

Aircraft generally derive their propulsion from fuel-fed heat engines whose power is fed to a propulsor. The propulsor accelerates a stream of air through the engine (as in the case of turbojets, turbofans, and ramjets) or around the aircraft (as in the case of helicopter rotors and propeller rotors) in a direction rearward of the flight direction. The integrated pressure forces impacting on the surfaces of the propulsive machinery required to accelerate the propulsive stream ultimately react on the aircraft to propel it in the flight direction. This is an application of Newton's second and third laws: The force required to accelerate a mass flow is proportional to the quantity of the mass flow multiplied by the rate of its acceleration, and for every action (of accelerating the mass flow) there is an equal and opposite reaction (on the engine and ultimately on the aircraft). *See* NEWTON'S LAWS OF MOTION.

Types of propulsion systems. The two most common types of heat engines used as power producers in aircraft propulsion are (1) a reciprocating- or piston-engine (Otto cycle) power producer, often mechanically or turbo-supercharged, which, produces mechanical energy to drive a propulsor; and (2) a gas turbine (Brayton cycle) power producer (compressor/combustor/turbine), where the energy produced is in the form of high-pressure, high-temperature products of combustion which are fed into a power turbine to drive a propulsor or into a jet nozzle which serves as a propulsor. A wide variety of propulsors are driven by these power producers, tailored to the aircraft range of speed and its unique mission requirements. In the case of rocket engines, the heat engine is simply a combustion chamber into which fuel and oxidizer (or a single monopropellant) are fed and burned, and the resultant stream of high-temperature, high-pressure air is fed to a propulsion nozzle. The most common types of aircraft propulsion systems are listed in **Table 1**, and diagrams of basic types are shown in **Fig. 1**. *See* AIRCRAFT ENGINE PERFORMANCE; HELICOPTER; PROPELLER (AIRCRAFT); RECIPROCATING AIRCRAFT ENGINE; ROCKET PROPULSION; SCRAMJET; TURBOFAN; TURBOJET; TURBOPROP.

Other propulsion systems have been proposed and studied and have undergone partial development, but have played a lesser role in the development of the airplane. For instance:

1. The pulsejet engine involves intermittent, cyclic combustion in a cylindrical chamber fitted with a propulsive jet nozzle and with a grid of flapper valves (acting as check valves) at its forward and aft ends to permit entry of air and exit of heated, pressurized gas. The pulsejet engine was initially developed by the Germans for use on the V-2 missile during World War II. It has no long-term utility for aircraft because of its extreme noise, internally destructive vibrations, limited flight speed, low efficiency, and short life.

2. The use of nuclear reactors instead of chemical fuels as the source of energy for aircraft engines was thought to offer the unique and attractive capability of providing spectacularly long unrefueled flight endurance. However, efforts to develop nuclear aircraft engines were terminated because of the unacceptable hazard to life and crops posed by the release of radioactive material in the event of an aircraft crash.

3. The turboramjet engine is a hybrid engine that includes a ramjet for high-speed flight and a turbojet for takeoff and acceleration to high speed. The complexity of the engine and the failure to identify important missions that might utilize its capabilities relegated it to a marginal, historical role in aircraft propulsion development.

4. The propfan is composed of a gas turbine power producer whose discharge is fed to a power turbine

TABLE 1. Types of propulsion systems

Engine type	Power producer, propulsor, and usage	Typical cruise speed*
Reciprocating- or piston-engine–driven helicopter rotor	Reciprocating- or piston-engine (Otto cycle) power producer that drives a helicopter rotor; used in small systems only	Low subsonic
Turboshaft-engine–driven helicopter rotor	Gas turbine power producer whose discharge is fed to a power turbine that drives a helicopter rotor; in general usage	Low subsonic
Reciprocating- or piston-engine–driven propeller	Reciprocating- or piston-engine (Otto cycle) power producer that drives a propeller; used in small systems only	Medium subsonic
Turboprop engine	Gas turbine power producer whose discharge is fed to a power turbine that drives a propeller; in general usage	Medium subsonic
Ultrahigh- or high-bypass turbofan engine or afterburning turbofan engine	Gas turbine power producer whose discharge is fed to a power turbine that drives a fan, a much larger turbocompressor that passes 2 to 10 times the airflow of the power producer (that is, a bypass ratio of 2 to 10) and is generally mounted in front of the power producer; The high-bypass turbofan is in general usage. The ultrahigh-bypass turbofan has been the subject of experimental demonstration and development.	High subsonic to transonic, $M < 1.2$
Medium- or low-bypass turbofan engine or afterburning turbofan engine	Gas turbine power producer whose discharge is fed to a power turbine that drives a fan, a somewhat larger turbocompressor that passes up to two times the airflow of the power producer and is generally mounted in front of the power producer. The engine may include an afterburner to augment the engine thrust for takeoff, combat, or transonic acceleration. In general usage.	Transonic to low supersonic, $M < 2.5$
Turbojet engine or afterburning turbojet engine	Gas turbine power producer whose discharge is fed to a jet nozzle to produce thrust. The engine may include an afterburner to augment the engine thrust for takeoff, combat, or transonic acceleration. In general usage.	Transonic to low supersonic, $M < 3$
Ramjet or scramjet engine	Engine with the same (Brayton) cycle as a jet engine, but with the compression produced by diffusion of the supersonic inlet airstream without the use of turbomachinery. The scramjet includes combustion of the airstream without prediffusion to subsonic speed. Crewless missile and experimental usage only.	Medium supersonic, $3 < M < 8$
Rocket	Combustion chamber and exhaust jet nozzle with not only the fuel but also the oxidizer carried onboard (rather than using captured air from the surrounding environment as oxidizer). Missile and extraterrestrial usage. Historical usage for assisted takeoff in aircraft.	High supersonic

*Where "sonic" is the velocity of sound or a Mach number (M) of 1.0 and the speed is 762 mi/h (1226 km/h) at an ambient temperature of 60°F (15.6°C) and 669 mi/h (1077 km/h) at an ambient temperature of −60°F (−51.1°C).

which drives a set of coaxial counterrotating high-speed propellers, mounted aft of the power producer. The specific fuel consumption at high subsonic flight speed is significantly lower than that of the high-bypass turbofan, and approaches that of the turboprop. In 1988, a flight demonstration was conducted of a unique version of this type of engine (which was called the unducted fan or UDF), involving a counterrotating power turbine that obviated the need for a gearbox to drive the propellers. In the same time period, a geared prop fan was demonstrated. Neither engine attracted enough interest to justify further development, primarily because of the difficulty of mounting the large-diameter propellers under the wing, and the unfavorable impact on aircraft weight balance in mounting the engines at the aft end of the fuselage. *See* TURBOPROP.

Propulsion system installation arrangements. In transport aircraft, the most usual approach to mounting the engines on the aircraft is enclosing each of the engines in a aerodynamically refined pod, called a nacelle, which is suspended from the aircraft on a pylon. **Figure 2** shows some typical arrangement for two-, three-, and four-engine transport aircraft.

In modern military combat aircraft, it is most usual to blend the engine enclosure into the underside of the aircraft's fuselage, with the engine's inlet just forward of the aircraft wing's leading edge. **Figure 3** illustrates such an arrangement for one- and two-engine military combat aircraft.

The major issues involved in the mounting of engines on aircraft are: (1) locating the engine inlet so that boundary layer or separated flow from the aircraft's fuselage, wings, or empennage (particularly in extreme flight maneuvers) is not ingested into the engine; (2) locating the engine inlet so that, during takeoff roll, sand, dust, and debris will not be ingested into the engine; (3) in any case where the engine is mounted in a nacelle hung by a pylon from the wing or fuselage of the aircraft, keeping appropriate separation between the nacelle and the wing or fuselage so that the nacelle does not engender deleterious interference drag; (4) locating the center of gravity of the engine or engines so that the center of gravity of the aircraft (to which the engines' weight makes a major contribution) is as close as possible to the center of lift of the aircraft's wings; and (5) in a multiengine application, keeping the engines' center lines as close as possible to the aircraft's center line so that, in an emergency one-engine-out situation, unbalanced yawing moments of the operative engines are minimized.

Propulsion system technical requirements. The requirements that may be placed upon the propulsion system for a particular aircraft may include any or all of the following:

Performance. Performance requirements include a definition of the flight regime or flight envelope as a function of altitude and flight speed in terms of Mach number (**Fig. 4**). Also to be specified are the thrust or power delivered by the engine and the rate of fuel consumed over the entire flight regime as a function of throttle setting (maximum, takeoff, climb, cruise, idle, and so forth), flight speed,

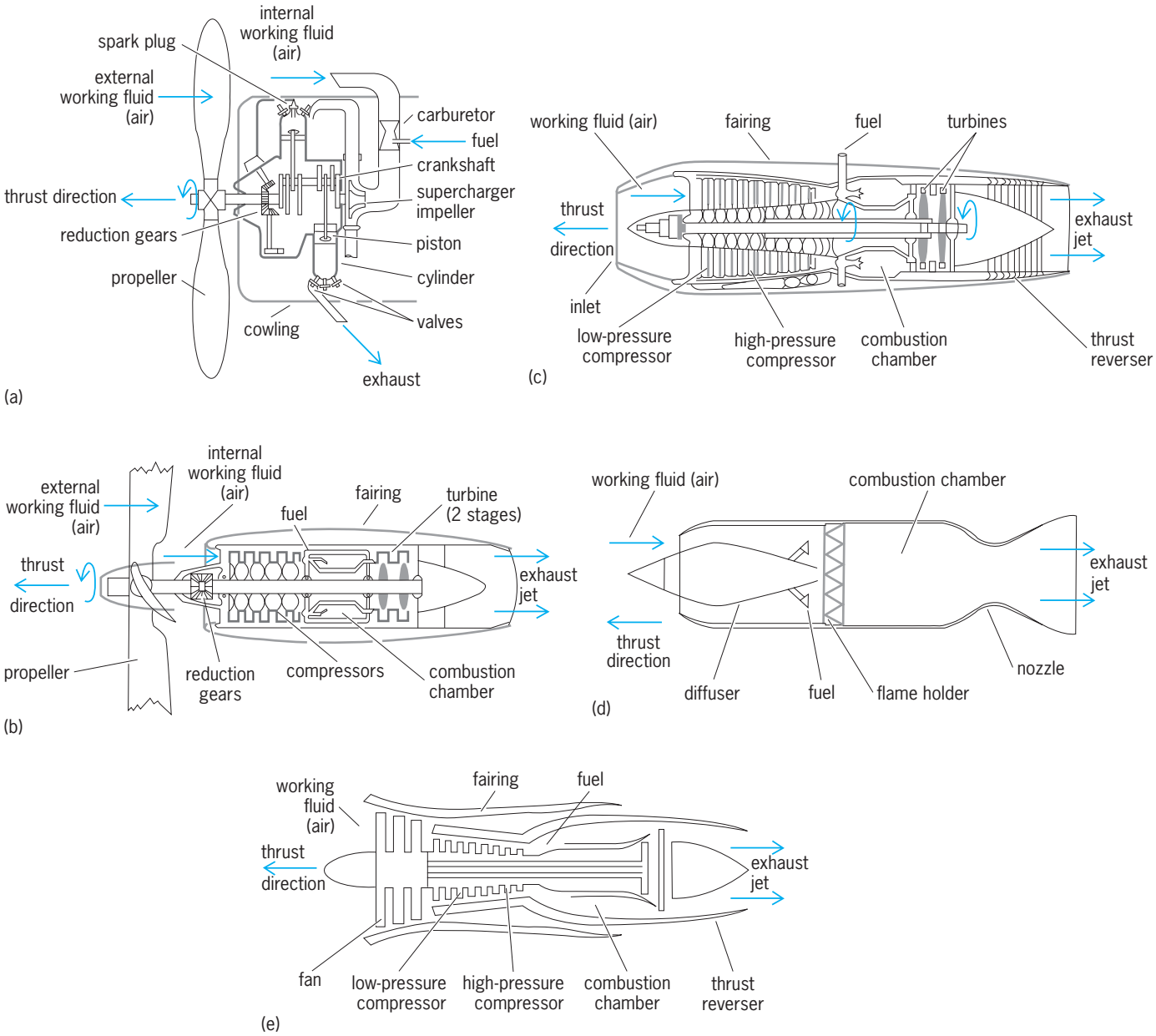

Fig. 1. Operational diagrams of the basic types of airplane engines. (*a*) Reciprocating. (*b*) Turboprop. (*c*) Turbojet. (*d*) Supersonic ramjet. (*e*) Fanjet (turbofan).

altitude, ambient pressure, temperature, humidity, and so forth; a description of additional modes of operation such as afterburning-augmented operation and reverse thrust; time limits on the usage of the high-power settings imposed to prolong the life of the engine; the maximum elapsed time allowed for transients in engine power to accommodate emergencies such as the sudden need for a burst of power

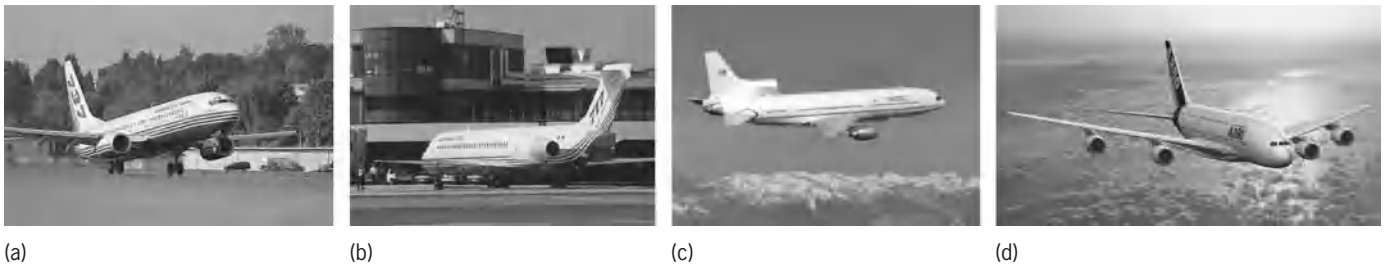

Fig. 2. Typical installation arrangements for multiengine transport aircraft with pylon/nacelle-mounted engines. (*a*) Boeing 737 with two-engine, under-the-wing mounted installation (*Boeing*). (*b*) Boeing 717 with two-engine rear-fuselage mounted installation (*Boeing*). (*c*) Lockheed Tristar with three-engine installation with tail-mounted middle engine (*NASA*). (*d*) Airbus 380 with four-engine under-the-wing installation (*Airbus*).

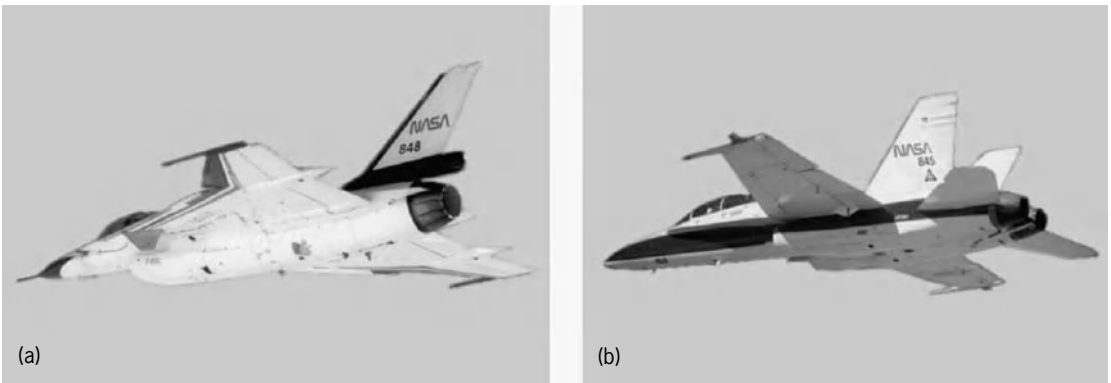

Fig. 3. Typical installation arrangements for one- and two-engine military combat aircraft. (*a*) Lockheed Martin F-16 with one-engine fuselage-mounted installation. (*b*) Boeing F-18 with two-engine fuselage-mounted installation. (*NASA*)

in the event of an aborted landing; the time for an engine to reach idle in the course of a starting sequence; the maximum deterioration of thrust or power that one might encounter in the course of the engine's life; the amounts of engine mechanical and electrical auxiliary power available to drive aircraft accessories, and the compressed air available from the engine (as well as the effects of their extraction on engine performance); and fuel types and limiting pressures and temperatures of the fuel delivered to the engine. *See* AIRCRAFT FUEL.

Weight. An aircraft's drag, and hence its thrust and fuel requirements, is strongly dependent on its weight since the two quantities are closely related by the characteristic lift/drag ratio of the wings. In virtually all aircraft systems, the engines' weight constitutes a major fraction of the total aircraft system weight, so the engines' weight is a critical parameter to be specified. The location of the engine's center of gravity may also be specified. *See* AERODYNAMIC FORCE.

Mechanical installation. The description of the mechanical installation includes the physical envelope of the engine, the mounting points of the engine where it is structurally mated to the aircraft's structure, and the details of the mount (whether

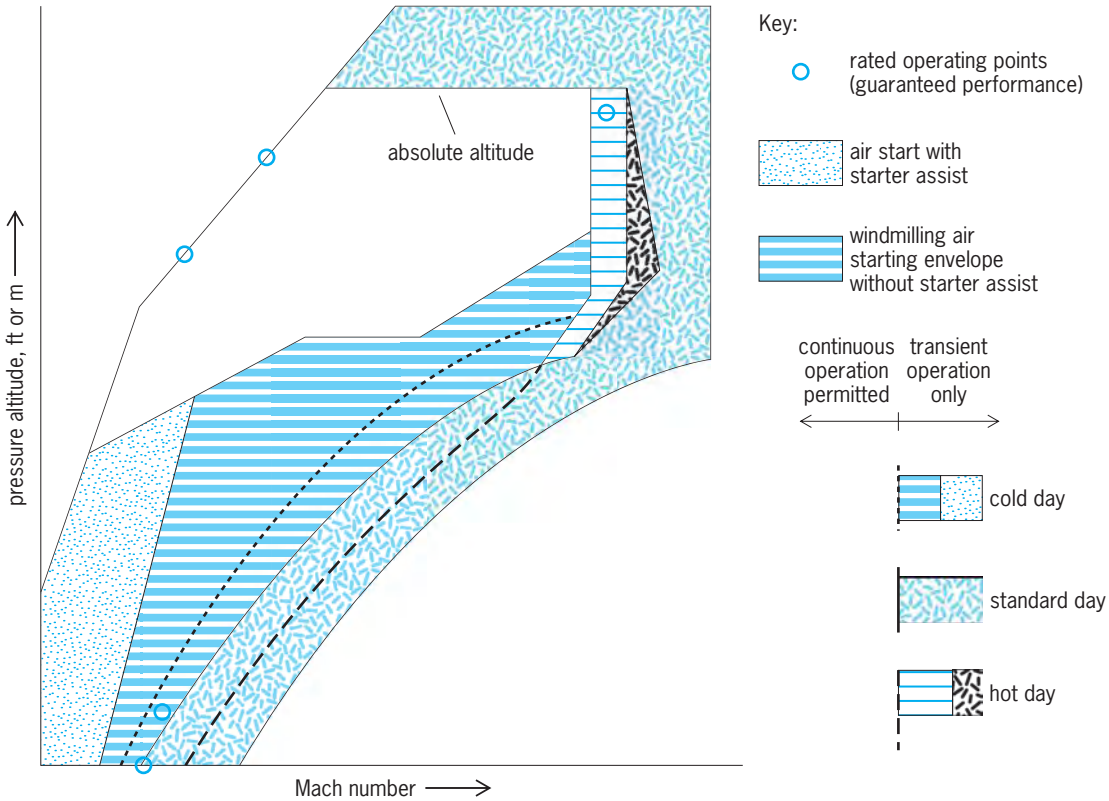

Fig. 4. Typical flight envelope of an aircraft in terms of the aircraft propulsion system limitations. Cold day, standard day, and hot day refer to U.S. Standard Atmosphere (1976).

hard-mounted or mounted in vibration and shock isolators, and any restrictions on the direction of the forces to be transmitted). Any limitations on the accelerations and gyroscopic moments associated with aircraft maneuvers or hard landings also must be identified. The interconnections between the aircraft and the engine—such as instrumentation and engine control information channels; fuel connections; pneumatic, electrical, and hydraulic system connections; and so forth—may be characterized. Surface temperatures of the engine must be specified to enable the design of the surrounding aircraft structure.

Environmental impact limitations. In particular, these include limitations on propulsion system noise and propulsion system exhaust emissions. In addition to the environmental impact of these emissions, the effect of noise and vibration on the aircraft and its passengers must be limited to acceptable values.

Engine durability, reliability, and safety. The engine durability in such terms as time between overhauls must be specified and the criteria for parts replacement identified. A schedule of periodic inspections may be specified. The engine's safety in the event of ingestion of hail and ice, birds, or other foreign objects must be ensured by specifying amounts of remaining thrust or power in the event of a minor incident or assurance of damage containment in the event of a major incident. The amount of pressure and temperature distortion in the inlet airflow to the engine that can be tolerated by the engine without its stalling or surging must be specified. Limitation must be placed on the leakage of flammable fluids, and any fire-suppression or inhibition devices (such as fire-walls) may be specified.

Engine accessibility and maintainability. The ease of engine removal, the simplicity of making engine inspections and adjustments while on-the-wing, and the capability of making on-the-wing change-out of accessories are important features of any engine design.

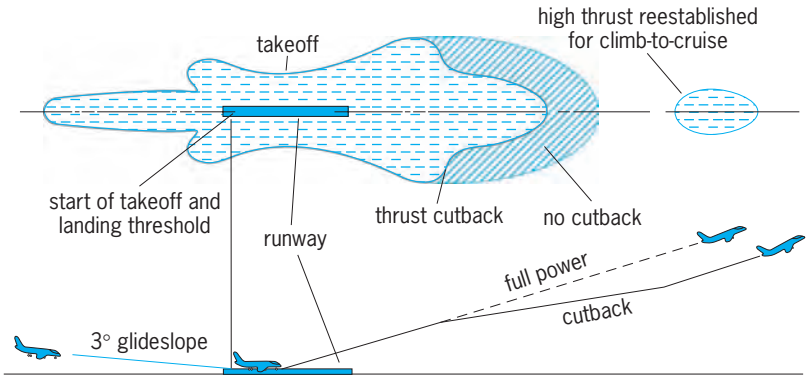

Fig. 5. Typical aircraft composite noise footprint for landing, full-power climb, and climb with power cutback. (*After M. J. T. Smith, Aircraft Noise, 2d ed., Cambridge University Press, 2004*)

TABLE 2. Aircraft/propulsion system noise emissions

Noise category	Cause	Type of noise generated	Mitigation or remediation
Jet noise	Random turbulence generated by the mixing layer at the periphery of high-velocity exhaust jets from jet engines and from low- and medium-bypass turbofan engines	Broadband or "white" noise (roaring or hissing sound)	Exhaust mixer nozzles
Combustor and afterburner noise	Random turbulence generated by mixing and flame instability in the combustion zones of the combustor or afterburner	Broadband or "white" noise (roaring or hissing sound)	In afterburner engines, limiting usage in populated areas
Combustor screech	Afterburner acoustic instability	Pure-tone noise	Perforated screech liner in the afterburner
Duct and turbomachinery blade surface noise	Random turbulence generated by flow over the surfaces of engine ducts, struts, and turbomachinery blading	Broadband or "white" noise (roaring or hissing sound)	Minimization of surface velocities where possible and necessary
Rotating blade and blade interaction noise	Interaction of rotating pressure fields with static parts or of static pressure fields with rotating turbomachinery blade rows at engine inlet and at engine exhaust	Discrete frequency or pure-tone noise at rotor-blade-passing frequency or its integer multiples	(1) Minimization of blade row-to-row interaction of blading pressure fields upstream of blade rows and trailing-edge wakes downstream of blade rows by maximizing spacing between succeeding blade rows; (2) selection of blade number to achieve "cutoff" for maximum acoustic attenuation; (3) employment of duct liners in inlet and exhaust ducts
Fan rotor "buzz-saw" noise	Upstream propagation of oblique shock from the tip of the fan rotor blade whose rotating velocity with respect to the incoming flow is supersonic	Discrete frequency or pure-tone noise at rotor-blade-passing frequency or its integer multiples	Employment of duct liners in inlet and exhaust ducts
Propulsor (that is, helicopter rotor or propeller) noise	Interaction of rotating propulsor pressure fields with static parts or static pressure fields with rotating propulsor blades	Discrete frequency or pure-tone noise at rotor-blade-passing frequency or its integer multiples	—
Sonic boom	Generated by the oblique shock cone emanating from the forward tip and wing leading-edge of the aircraft traveling at transonic and supersonic speed	Large pressure pulse as heard by the stationary observer on the ground	(1) Constraining aircraft to operation at subsonic speed over populated areas; (2) constraining aircraft to high enough altitude that cone shock is sufficiently attenuated when it reaches the ground

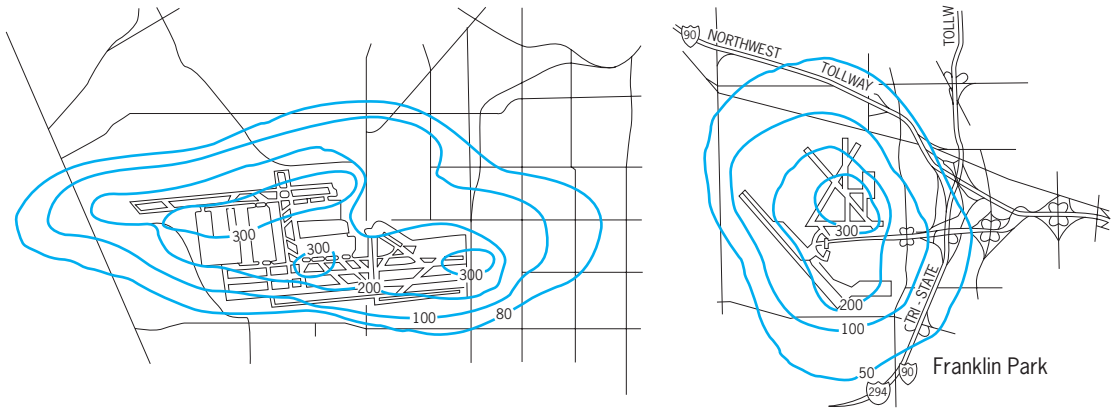

(a) (b)

Fig. 6. Typical contour maps of NO_X concentration (measures in micrograms per cubic meter) in major airports—annual averages for 1980. (*a*) Los Angeles International Airport. (*b*) O'Hare Chicago International Airport. (*After U.S. Environmental Protection Agency, Aircraft Emissions: Impact on Air Quality and Feasibility of Control, Report to Congress, undated*)

Unique military requirements. The engine requirements may also include the capability of ingestion from ordnance gun gas or rocket exhaust, minimization of radar cross section and infrared emissions, invulnerability to and survivability from weapons damage, hardening of electronics to damage from nuclear radiation, and so forth.

Propulsion system noise emissions. The noise generated by the propulsion system is a major consideration in the design and installation of the engine and propulsor. **Table 2** summarizes some of the most important constituents in the generated noise, the specific causes, the type of noise generated, and the measures that are taken to mitigate or remediate their impact on the public. *See* AERODYNAMIC SOUND; SONIC BOOM; TURBINE ENGINE SUBSYSTEMS.

Operation of the aircraft can have a profound effect on the noise "footprint" of the aircraft. In the case in **Fig. 5**, the power of the engine is cut back from full power in the midst of its climb after takeoff to reduce the noise heard by the public under the flight path, substantially reducing the footprint. But an additional area of intrusive noise is included at a further point where high thrust climb is reestablished for continuing the full-power climb to cruise altitude. *See* ACOUSTIC NOISE; AIRCRAFT NOISE.

Propulsion system exhaust emissions. A second major impact on the environment of heat engines fueled by fossil fuels for aircraft propulsion is in the engines' emission of products of combustion into the atmosphere. Although the mechanism of production of pollutants is similar to that in all vehicular and stationary power plants, the specific consequences and issues are unique to aircraft propulsion because of the aircrafts' use in the troposphere and stratosphere and the intense concentration of their use on and near the ground at airports and their environs (**Fig. 6**). The major constituents of aircraft engine exhaust pollution emissions are listed in **Table 3** along with the engine operation at which they are encountered, their impact, and approaches to their mitigation or remediation. *See* AIR POLLUTION.

TABLE 3. Propulsion system exhaust emissions

Emission category	Primarily associated with engine operation at	Impact	Mitigation or remediation
Smoke (soot)	Takeoff and climb-out	Visibility nuisance around airports	Lean primary combustion zone fuel–air mixtures; more effective fuel-air mixing within the primary zone
Unburned hydrocarbons	Low power, especially ground idle	Contribution to urban smog burdens	Improvement of fuel atomization at low power; enhancing primary zone fuel-air mixing quality; minimizing fuel impingement on the combustor liner; reduction of idle power usage
Carbon monoxide (CO)	Low power, especially ground idle	Contribution to urban CO burdens	Improvement of fuel atomization at low power; enhancing primary zone fuel-air mixing quality; minimizing fuel impingement on the combustor liner; reduction of idle power usage
Nitrogen oxides (NO_X)	All high power, including cruise	Contribution to urban smog burdens in environs of airports	Reduction of peak flame temperatures; leaner primary zone mixtures at high power
Subsonic aircraft engines		Possible contribution to global warming	
Future supersonic aircraft engines		Stratospheric ozone depletion	
Carbon dioxide (CO_2)	All power settings	Contribution to global warming	Improvement of engine fuel efficiency
Sulfur oxides (SO_X)	All power settings	Contribution to urban smog burdens	Use of low-sulfur fuel

The disparate requirements of improving combustor operation at both very low power conditions and at very high power conditions have dictated that more complex combustion systems with variable geometry or variability in the fuel injection system have had to be employed. Typically, a modern aircraft gas turbine engine combustion system might include two sets of fuel injection nozzles and employ a single set at low power settings and both sets at high power settings.

Limitations on emissions. Limitations on aircraft noise emissions and on aircraft exhaust emissions have been promulgated by the International Civil Aviation Authority (ICAO). These emissions are regulated in the United States by the Federal Aviation Administration (FAA), in the United Kingdom by the Civil Aviation Authority (CAA), and in Europe by the Joint Aviation Authority (JAA). Fredric F. Ehrich

Bibliography. M. Barrett, *Aircraft Pollution—Environmental Impacts and Future Solutions*, WWF Research Paper, August 1991; R. D. Bent and J. L. McKinley, *Aircraft Powerplants*, 5th ed., 1985; J. L. Kerrebrock, *Aircraft Engines and Gas Turbines*, 2d ed., 1992; M. J. Kroes and T. W. Wild, *Aircraft Powerplants*, 7th ed., 1995; J. D. Mattingly, W. H. Heiser, and D. T. Pratt, *Aircraft Engine Design*, 2d ed., AIAA, Reston, VA, 2002; G. C. Oates, *Aircraft Propulsion Systems: Technology and Design*, 1989; Rolls-Royce PLC, *The Jet Engine*, Derby, United Kingdom, 5th ed., 2005; M. J. T. Smith, *Aircraft Noise*, 2d ed., Cambridge University Press, 2004.

Aircraft rudder

The hinged rear portion of an airplane's vertical tail. The vertical tail is composed of the vertical stabilizer and the rudder. The vertical stabilizer is mounted to the fuselage and is fixed relative to it. The rudder is hinged to the rear of the vertical stabilizer (see **illus.**) and moves to the left or right in response to control inputs from the rudder pedals or from an automatic stability and control system.

The rudder provides an aerodynamic moment about the aircraft's center of gravity for the purpose of yaw control. When the rudder turns clockwise, for example, as viewed from above, its trailing edge moves to the left, effectively adding camber to the vertical tail. The result is that an aerodynamic side force is produced on the vertical tail to the right. This force, in turn, produces a counterclockwise yawing moment about the airplane's center of gravity, resulting in a turn to the left. Conversely, to turn to the right, a pilot pushes on the right rudder pedal, causing the rear of the rudder to swing to the right and the airplane to turn in that direction. *See* AIRFOIL.

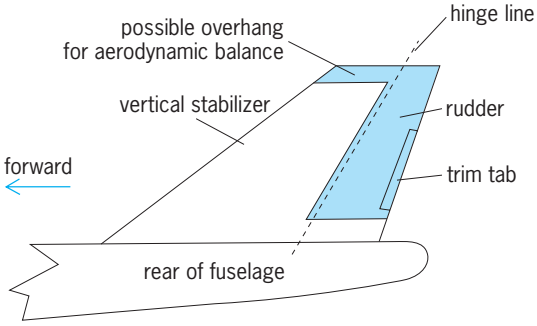

Vertical tail of an airplane, showing location of the rudder.

An additional, small, movable surface, known as a trim tab, may be hinged to the rudder. When deflected to a fixed position, the tab causes the rudder to deflect to, and hold, a desired angle. Thus a steady yawing moment is produced which can relieve the pilot from having to make a sustained push to offset any disturbing moments such as from a propeller.

For some supersonic applications, the vertical tail is a one-piece configuration, with the entire tail rotating to provide yaw control. For vee-tails, the rudder and elevator control surfaces are one and the same. For pitch control (which is provided by the elevators in a conventional configuration), the hinged surfaces on either side move together. For yaw control, they move differentially; that is, one of the surfaces moves up while the other one moves down. The result is a net side force to one side or the other while the net vertical force is zero. *See* ELEVATOR (AIRCRAFT); TAIL ASSEMBLY.

Some airplanes employ twin, or even multiple, vertical tails. For this configuration, rudders are usually placed on each vertical tail and move in unison to the pilot's control input to yaw to the left or right. The total effectiveness of multiple rudders is equal closely to the effectiveness of one rudder multiplied by the number of vertical tails. If the tails are close together, there is some amount of interaction which can be determined from aerodynamic analyses. *See* FLIGHT CONTROLS. Barnes W. McCormick, Jr.

Aircraft testing

Subjecting a complete aircraft or its components (such as wings, engines, or electronics systems) to simulated or actual flight conditions in order to measure and record physical phenomena that indicate operating characteristics. Testing is essential to the design, development, and acceptance of any new aircraft.

Aircraft and their components are tested to verify design theories, obtain empirical data where adequate theories do not exist, develop maximum flight performance, demonstrate flight safety, and prove compliance with performance requirements. Testing programs originate in laboratories with the evaluation of new design theory; progress through extensive tests of components, subsystems, and subsystem assemblies in controlled environments; and culminate with aircraft tests in actual operational conditions. *See* AIRCRAFT DESIGN.

Laboratory Testing

Instrument testing, in controlled conditions of environment and performance, is used extensively during the design performance assessment of new

Fig. 1. Scale model of fighter aircraft being mounted inside test section of large wind tunnel. (*NASA*)

aircraft to avoid the costly and sometimes dangerous risks of actual flight. Each method of testing contributes to the development of an efficient aircraft.

Wind tunnel tests. A wind tunnel is basically an enclosed passage through which air is forced to flow around a model of a structure to be tested, such as an aircraft. Wind tunnels vary greatly in size and complexity, but all of them contain five major elements: an air-drive system, a controlled stream of air, a model, a test section, and measurement instruments. The drive system is usually a motor and one or more large fans that push air through the tunnel at carefully controlled speeds to simulate various flight conditions. A scale model of an actual or designed aircraft is supported inside the test section (**Fig. 1**), where instruments, balances, and sensors directly measure the aerodynamic characteristics of the model and its stream of airflow.

Wind tunnel tests measure and evaluate airfoil (wing) and aircraft lift and drag characteristics with various configurations, stability and control parameters, air load distribution, shock wave interactions, stall characteristics, airflow separation patterns, control surface characteristics, and aeroelastic effects (**Fig. 2**). Tunnels are also used to calibrate airflow-sensing and pressure-sensing devices, engine inlet performance, engine fan airflow–nacelle (engine housing) interference, weapon drop trajectories, and engine performance tests. *See* AIRFOIL.

Test speeds are generally divided into subsonic (Mach numbers 0–0.95), transonic (Mach 0.95–1.3), supersonic (Mach 1.3–5), and hypersonic (Mach 5 and above). The hypersonic range includes the transatmospheric vehicle concept that merges aeronautics and space technologies across the speed ranges involved in the conventional ground takeoff of the vehicle to Earth-orbital velocities and the vehicle's return to ground landing. *See* HYPERSONIC FLIGHT; MACH NUMBER; SUBSONIC FLIGHT; SUPERSONIC FLIGHT; TRANSONIC FLIGHT.

Different kinds of wind tunnels simulate the airflow of these speed ranges, including two- and three-dimensional, return and nonreturn flow, intermittent and continuous flow, atmospheric and variable density, and vertical free flight. Hypervelocity facilities have been developed to better simulate conditions at hypersonic speeds, including shock tubes and tunnels, hotshot and impulse tunnels, ballistic ranges, and light-gas guns.

A relatively new design of transonic wind tunnel uses cryogenic (supercold) nitrogen gas as a test medium (instead of high-pressure air) in a continuous-flow, pressurized, closed circuit. Liquid nitrogen is vaporized inside the tunnel at temperatures as low as -149°C (-300°F) to reduce the viscosity and increase the density of the gas moving through the tunnel. The cryogenic concept provides almost exact simulations of the flight conditions of advanced design aircraft that will fly in or through the transonic speed range.

Several aircraft configurations generally evolve during preliminary design, and the best candidates are chosen for testing. Scale models of each configuration are installed in a wind tunnel for short tests to determine if a model's aerodynamic shape produces the desired performance. Many tests may be made of various configurations before the most satisfactory aerodynamic design is selected for further development.

After a general configuration has been established, a more detailed design is developed, and a complex series of wind tunnel tests is begun. Aircraft lift and drag curves are determined throughout the design speed and Mach number range. Aircraft static stability forces and moments are measured for expected variations in flight conditions. Wing air-load distributions are found by measuring the static air pressure

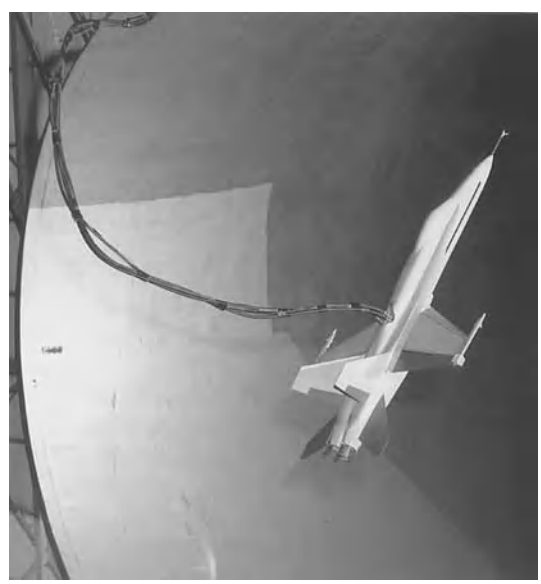

Fig. 2. Free-flying scale model aircraft in wind tunnel; attached cable contains sensors and other measurement instruments. (*NASA*)

on the upper and lower wing surfaces, fuselage, and tail surfaces, at various points along the aircraft's length and width.

Patterns of airflow are made visible through several techniques, including neutrally buoyant luminescent bubbles of helium, smoke filaments, and tufts of thread. Schlieren photography or shadowgraphs can visualize large changes in air density, such as shock patterns. Two techniques make use of lasers. (1) A laser hologram uses relatively simple optics to split the light from a powerful laser into two beams, one of which passes through a wind tunnel test section. When recombined on a photographic plate, the beams form a hologram (an interference pattern) that captures the pattern of density gradients within the test section. (2) A laser velocimeter operates like a radar to acquire detailed measurements of flow velocity. A laser beam is split into two beams that cross at adjustable points in a tunnel, allowing two velocity components in two different planes to be determined at the same time. Free-flight vertical wind tunnels are used to evaluate an aircraft's dynamic spin characteristics. Flexible models simulate the structural flexibility of an aircraft to evaluate aeroelasticity characteristics. *See* AEROELASTICITY; FLOW MEASUREMENT; HOLOGRAPHY; SCHLIEREN PHOTOGRAPHY; SHADOWGRAPH; WIND TUNNEL.

Integration testing. Aircraft components are integrated into subsystems, system elements, and complete operational systems to help resolve interface problems. Tests establish functional and operational capability and evaluate complete system compatibility, operation, maintenance, safety, reliability, and best possible performance (**Fig. 3**).

Typical examples of integration testing are tests of control systems and avionics (aviation electronics) systems. Testing the integration of a control system involves a completely rigged primary control system that is installed on a structural frame similar to that of an actual aircraft. The arrangement can test friction, force feel, fatigue wear, and response, while subjecting the system to flightlike air loads. An automatic flight control system checks compatibility, and pilots are included in tests to obtain significant results. Avionics systems tests can range from a simple antenna test with a receiver to such complex tests as integrating a system of electrical power, communications and guidance equipment, and mission avionics, which includes airborne warning radar and command and control equipment. *See* AIRBORNE RADAR; CONTROL SYSTEMS; FLIGHT CONTROLS.

Dynamic ground tests. Rocket-propelled sled tests evaluate crew ejection escape systems for high-performance aircraft. A fuselage section, mounted on a sled, is propelled by rockets along fixed tracks. When a desired speed is reached, the ejection mechanism is automatically triggered, firing rockets that propel crew seats (containing instrumented mannequins) clear of the fuselage, and activating parachutes to limit the free-flight trajectory of the mannequins and allow safe descent to the ground. Water-propelled sled tests study landing gear systems and runway surface materials. High-pressure water can propel sled-mounted test equipment at speeds comparable to those of high-speed aircraft (**Fig. 4**). The sled technique provides accurately controlled test conditions that are documented with high-speed photography. Results are studied to verify that all parts of a test system are within human or mechanical tolerances.

Fig. 3. Scale model of E-7 advanced fighter aircraft used for tests of experimental component called augmenter (rectangular opening in wing with open flap). (*NASA*)

Other dynamic ground tests include acceleration and arresting tests of aircraft fuel system venting,

Fig. 4. Test sled propelled down runwaylike track by high-pressure water to test aircraft components during simulated landing. (*NASA*)

transfer, and delivery, which are evaluated while the system is subjected to flightlike forces and attitudes. Engines are operated in highly instrumented test stands that use aircraft mountings, controls, and inlet and exhaust cowlings containing thrust reversers. Tests substantiate satisfactory cooling, vibration, and installed engine performance. *See* AIRCRAFT ENGINE PERFORMANCE.

Ground vibration tests are done before the first flight of a new model aircraft, subjecting the entire aircraft to vibrations induced from electromagnetic exciters. Structural response is analyzed with magnetic pickups whose outputs are read on oscilloscopes. The tests verify that fuselage structures and empennage (aft control structures) are not responsive to vibrations that will be encountered by engines and equipment during gust air loads anticipated in actual flight.

Static load tests. Proof load tests of actual aircraft are usually done on one or more of the first airframes built. An airframe, mounted in a laboratory, is fitted with thousands of strain gages, the outputs of which are recorded on an automatic data-recording system. Simulated air and inertia loads are applied to airframe components, which are loaded simultaneously, in specified increments, to simulate loads encountered during takeoff, maneuvering flight, and landing. Structural stress and deflection data are recorded at each load level, starting with loadings equal to 80% of design limit. Loadings are increased to design limit and then to ultimate failure to locate possible points of excessive yield. Flight air loads, including those encountered during gust upsets, are repeatedly applied to determine the structure's fatigue life. Critical parts may be redesigned to withstand required loads and ensure satisfactory service life. *See* AIRFRAME; STRAIN GAGE.

Environmental tests. Component parts and system subassemblies are tested with various loadings while operating under expected extremes of temperature, humidity, and vibration to determine service life. Scale models containing communication antennas are tested for radiation properties to verify transmission performance. Mockups that simulate electrical and avionic components are operated under load to verify that their installation is free from electromagnetic interference. Environmental chambers verify the operation of rain-removal, de-ice, and anti-ice equipment. Maurice Parker

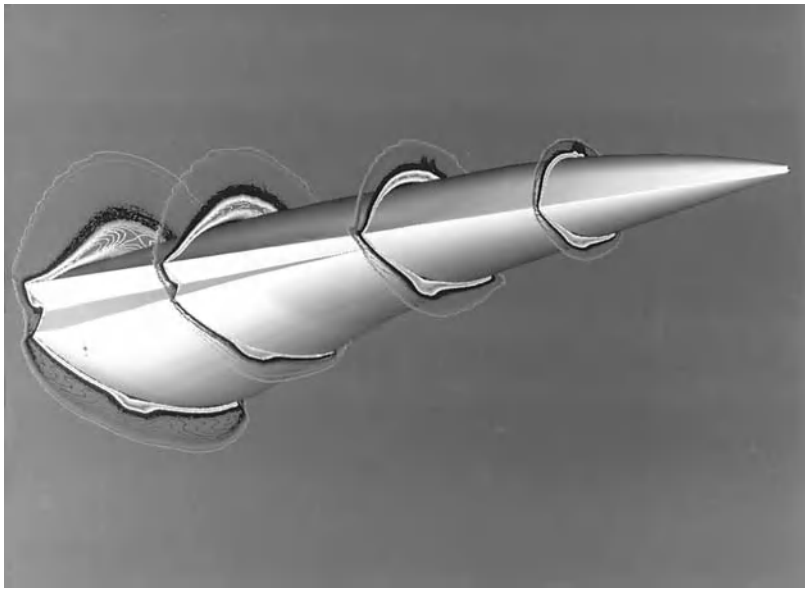

Fig. 5. Supercomputer-aided study of hypersonic lifting body, showing computer-generated lines of constant pressure during Mach 19 flight. (*NASA*)

Flight Simulation

Aircraft flight and systems characteristics are represented with varying degrees of realism for research, design, or training simulation purposes. The representation is usually in the form of analytic expressions programmed on a digital computer. Flight simulation may be performed with or without a human pilot in the loop. The pilot imposes additional constraints on the simulator such as requiring a means of control in a manner consistent with the means provided in the aircraft being simulated. Flight simulation requires representation of the environment to an extent consistent with the purposes of the simulation, and it further requires that all events in the simulator occur in real time. Real time is a term which is used to indicate that all time relationships in the simulator are preserved with respect to what they would be in the airplane in flight. *See* REAL-TIME SYSTEMS. Frank Cardullo

Simulators range in size and complexity from actual aircraft, outfitted with special flight decks that can be reconfigured to test different systems, to desktop simulators that can test individual or integrated components.

Use of computers. Highly sophisticated computers are the primary element in most aircraft simulation tests. Analog computers provide accurate and repeatable discrete simulations, such as systems analysis and human–machine flight simulations. Hybrid computers combine analog and digital computers to work on a problem simultaneously, with each computer doing the task best suited to its special capability. Realistic simulation testing provides flight crews with exact flight deck representations (on either fixed- or moving-base platforms), visual displays, and functioning instruments. *See* ANALOG COMPUTER; DIGITAL COMPUTER.

Supercomputers, large, powerful scientific computers that contain memories of more than 250 million words and can perform 1 billion computational operations each second, are accelerating the rate of progress in computationally intensive disciplines such as aerodynamics, structures, materials, and chemistry, particularly for hypersonic and transatmospheric flight vehicles (**Fig. 5**). A relatively new testing technique, made possible by the increasing availability of supercomputers, is computational fluid dynamics. It provides powerful analytical, simulated, and predictive tools to study the basic physics of aerodynamic flow fields, especially to better understand the complex flow environment of advanced aircraft configurations, and allow the aerodynamic

Fig. 6. Controls of transport-systems research-vehicle simulator, used to develop automated pilot aids, being operated by engineer. (*NASA*)

perfection of new designs. Computations are made of three-dimensional viscous compressible flows that contain separated flow regions over wing and fuselage combinations. An example is computation of hypersonic wing–body surface pressure contours for a flight environment of Mach 25, a 5° angle of attack, and a 3000°R (1667 K) wall temperature. *See* SUPERCOMPUTER.

Applications. Simulators may be classified by their use in research, design, or training. Research simulators are usually employed to determine patterns of human behavior under various workloads or in response to different flight instrument display configurations or different aircraft dynamic characteristics. An example of a simulator-tested system is the takeoff performance monitoring system (**Fig. 6**), used to develop automatic pilot aids. Inside the flight deck of a transport-system research-vehicle simulator, information displayed on electronic screens as symbols and numbers is superimposed onto a scaled out-the-window runway scene. Predictions and advisories are updated in real time, based on sensed conditions and performance during a takeoff roll. The system predicts where on a runway important takeoff events (such as rotation or stopping) will occur and provides advisory information on whether the takeoff should be continued or aborted. The system should enhance takeoff safety by providing pilots with previously unavailable information. Maurice Parker

Design simulators are used to conduct tradeoff studies to evaluate different design approaches in the aircraft. Virtually all aircraft manufacturers in the United States employ simulators to a large extent in the aircraft design process. While the investment is substantial, many hours of flight tests are saved, and designs can be evaluated much earlier in the aircraft production schedule, resulting in considerable savings in retrofitting hardware. Simulators are also used by aircraft manufacturers to demonstrate aircraft characteristics to potential customers well before the first flight of the airplane.

The most pervasive use of flight simulators is for training operators of the aircraft and its systems and maintenance personnel. The flight simulator was patented in 1929 by E. A. Link to train pilots in instrument flying; it was responsible for training a half million pilots during World War II. The current flight simulator is quite different from Link's Blue Box and is used to train pilots of virtually every military and commercial aircraft in the United States and Europe.

Training simulators can be broken down into five categories. Operational flight trainers are used to train pilots, copilots, flight engineers, and navigators, as appropriate, in all aspects of flying a modern aircraft, either with a full crew or alone. A second type of device is a weapons systems trainer, employed on tactical and strategic aircraft. These devices add sophisticated avionics and weapons systems simulation to an operational flight trainer. A part task trainer is a device which provides simulation of part of the operational flight trainer or weapons systems trainer, such as the situations confronted by a navigator in a strategic aircraft or an antisubmarine warfare systems operator in a naval patrol aircraft. Cockpit procedures trainers are employed to train pilots in the procedures of operating the systems of an aircraft. Finally, maintenance trainers are employed to train mechanics and technicians in the diagnosis of problems and the procedures of routine maintenance on various aircraft systems such as engine, flight controls, and radar.

Training simulators have been used extensively by NASA for training astronauts in programs from Project Mercury to the space shuttle. They were also used during the *Apollo 13* mission to verify procedures and flight program modifications that helped to avert possible disaster after an inflight explosion partially disabled the spacecraft. *See* SPACE FLIGHT.

That training simulators save substantial amounts of money in fuel costs, weapons expenditure, and so forth is obvious. However, the simulator has much greater benefit in that it is in many cases a better training device than the aircraft. This is true because of the safety, versatility, and speed with which critical maneuvers may be performed. For example, in practicing landing in the aircraft, a student pilot may make an approach to either a touch-and-go landing or a full-stop landing or perhaps a missed approach. The student then must contend with air traffic and weather to reenter the landing pattern for another attempt. This takes considerable time, whereas in the simulator, after landing, the simulated aircraft may reset to the approach configuration at the top of the glide slope, ready to go again in seconds. The simulator also allows the student pilot to learn to deal with malfunctions in a safe environment. Practicing the

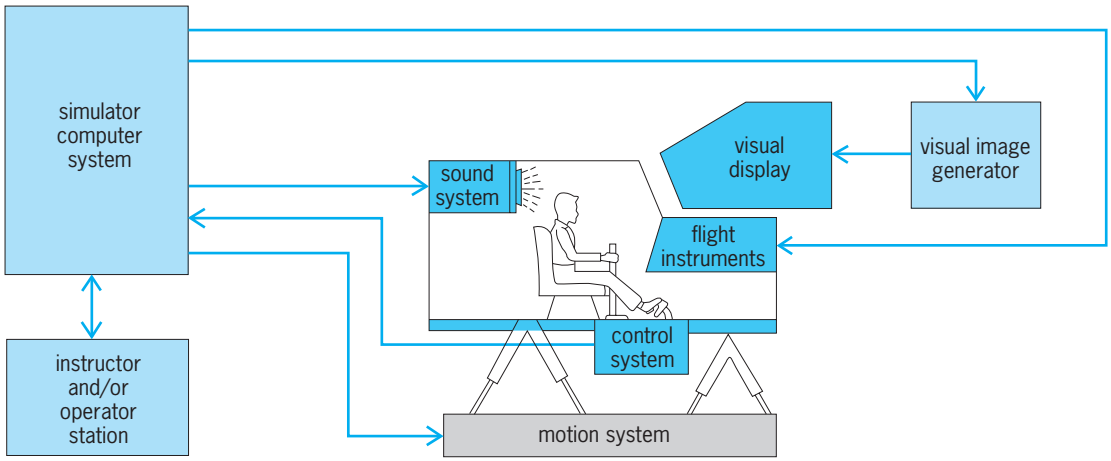

Fig. 7. Diagram of flight simulator illustrating major components and their interconnections.

engine-out approach and landing is very useful but also very dangerous to perform in the airplane. The simulator provides an excellent training medium for this and many other malfunctions.

Components. **Figure 7** illustrates the major components of a piloted flight simulator. The primary component is the cockpit with flight controls and instruments. The other elements are the visual display system, which replicates the out-the-window scene; the motion cuing system, which provides the bodily sensations of the motion of the aircraft; the sound system, which provides the normal and abnormal sounds of flight; the instructor or operator station, which provides the control of the device; and the computer system, which drives the entire complex.

The pilot of a flight simulator manipulates the controls in the simulator cockpit in the same manner as in the aircraft. Sensors in the control system measure this control manipulation and transmit the analog information to the simulator computer system, whereupon the appropriate dynamics of the aircraft are computed by the algorithms in the computer system. These dynamics then are used by the cockpit instrument drive algorithms to provide the appropriate commands to each of the instruments indicating the state of the simulated vehicle, including its engines and other systems. The vehicle state information is also used by the simulator visual system to display the resulting out-the-window imagery and by the simulator motion systems to provide the appropriate motion cues. The sound system is driven by the vehicle and system state information. Algorithms in the computer also determine the proper force which should be felt by the pilot due to the aerodynamic effects on the control surfaces. The calculated force is then transmitted to a hydraulic force-feel system at the cockpit which alters the forces experienced by the pilot.

Software. A key element in this simulation process is the software which controls and performs the simulation. **Figure 8** illustrates the software found in a typical simulator and also shows the hardware interfaces. The flight simulator software comprises three parts: the system software, the instructional or operator software, and the mathematical models which simulate the various aircraft functions. The system software includes the real-time operating system, input-output processing, task scheduling, and so forth. The instructional software is utilized only by training simulators and provides the features necessary for training such as performance monitoring, recording and playback of various scenarios, and maneuver demonstration. The operator software allows the various modes of the simulator to be exercised, such as freeze, which stops the action in place; reset, which allows the operator or instructor to position the aircraft at predetermined locations in space; and fast time–slow time, which allows for faster than or slower than real-time operation.

Mathematical models are required of the aircraft dynamics, the engines, the control system, the avionics systems, the weapon systems, the atmosphere, and aircraft systems such as electrical, hydraulic, and fuel management. A mathematical model is a set of mathematical equations that describe the behavior of a physical system. Figure 8 illustrates the type of mathematical models typically found in a flight simulator and the information flowing among the various modules. The broken lines enclose the portion of the simulator which is implemented in software. The area to the right represents hardware at the flight station, and the area to the left is the instructor or operator station.

The vehicle dynamics simulation comprises the mathematical models contained within the innermost area of Fig. 8. The flight-controls module senses the pilot's control activity and interprets it as control surface deflections or engine commands. These parameters are then passed to the appropriate modules, such as the engine module or the aerodynamics module. The engine module contains the simulation of the engine that is installed in the particular aircraft being simulated. The output of this module is the thrust of each engine, which ultimately contributes to the forces and moments acting on the airplane, as well as to the revolutions per minute, fuel consumption, and other engine parameters which are displayed in the cockpit.

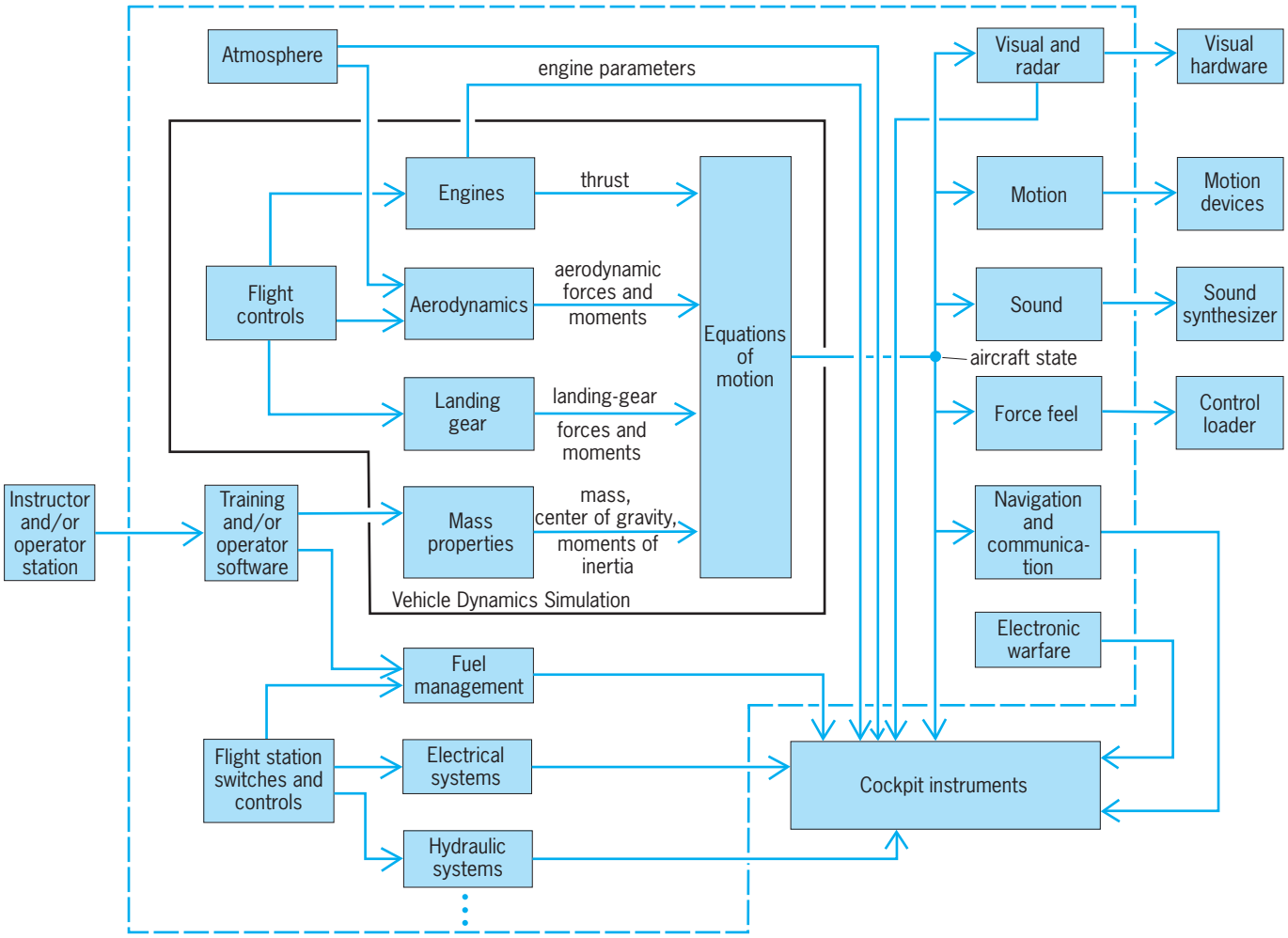

Fig. 8. Simulator block diagram indicating information flow among the software modules, within the broken line, and the interfaces to major hardware elements.

The aerodynamics module computes all the aerodynamic forces and moments acting on the aircraft as a result of control actions by the pilot and the aircraft's interaction with its environment. The forces are drag, side force, and lift. The moments are rolling moment, pitching moment, and yawing moment.

The landing gear module computes the forces and moments acting on the aircraft as a consequence of the interaction of the airplane and the ground in taxiing, landing, and taking off. The mass-properties module computes the vehicle mass and location of the center of gravity as well as the moments of inertia.

The outputs of these four modules are used by the equations-of-motion module to compute the simulated aircraft state vector, which is the aircraft's location in space (latitude, longitude, and altitude) and its orientation relative to the ground (pitch, roll, and heading). The aircraft state vector includes velocity and acceleration along the flight path as well as velocity and acceleration components in the above-mentioned six degrees of freedom. This computation is accomplished by summing the forces and moments due to engines, aerodynamics, landing gear, and weight; computing the acceleration components; integrating these components to obtain velocity components; and integrating again to obtain the spatial coordinates and orientation of the vehicle relative to the ground.

As is illustrated in Fig. 8, the simulated aircraft state vector is passed to various other software modules which utilize it to simulate other aircraft systems or to drive cuing systems such as visual, motion, and sound. The navigation and communication module simulates the aircraft's navigation aids, landing aids, and various communication devices. The electronic warfare simulation may actually be composed of several modules distributed between offensive systems and defensive systems, which may contain mathematical models of threats such as other aircraft or missiles. *See* ELECTRONIC WARFARE.

As shown in Fig. 8, these software modules interface with hardware systems that provide cues to enable the pilot to control the simulated aircraft as if it were a real aircraft. In most cases, the cuing devices synthesize the cues; therefore, the corresponding software modules contain algorithms to effect that synthesis.

Visual simulation. The visual and radar module, in modern flight simulators, formats the aircraft state

vector and environmental information, which is then passed to the image generator. The image generator employs sophisticated computer graphics techniques either to generate imagery for a radar display or to create the view the pilot would have out the windows of the cockpit. The technology employed for generating these scenes has in the past included film, camera–model, scanned transparency, and video disk. However, the most advanced technique is now computer-generated imagery. In addition to image generation, projectors and screens of some sort are required to present the scene properly to the pilot of the simulator. The projectors include cathode-ray-tube and light-valve projectors. The screens include flat or curved screens, dome projection surfaces, and virtual-image displays composed of large, complex arrays of optics. **Figure 9** shows a simulator used to train pilots of the E-3A (AWACS) aircraft. The exterior of the projection screen is at the front of the simulator cab. *See* COMPUTER GRAPHICS.

Motion simulation. Also shown in Fig. 9 below the cab is a platform supported by six hydraulic jacks which are controlled by the motion system software to provide cues to the pilot of the motion environment. While these motion systems are capable of providing motion in all six degrees of freedom in which the aircraft moves, that motion is constrained by the physical limitations of the actuation size and geometry. Therefore, highly sophisticated drive algorithms are required in the software to provide usable cues to the pilot and eliminate false cues. Flight simulators frequently use ancillary motion cuing devices such as seat vibrators to simulate stall buffet and other vibratory motion. In addition, g-seats are sometimes employed to provide sustained cues by stimulating pressure sensations across the back and buttocks of the pilot in the simulator. The anti-g suits commonly found in high-performance aircraft are frequently utilized in simulators to assist the pilot in withstanding the stress of high acceleration, and these have been found to produce very effective cues.

Fig. 9. Training simulator for the Boeing E-3A (AWACS) aircraft. Cabin containing flight deck, exterior of visual display system, and supporting six-degree-of-freedom motion system are shown. (*Rediffusion Simulation*)

Sound simulation. Modern flight simulators often include sound simulation, which is achieved by software control of digital sound synthesizers. These systems are capable of reproducing all aircraft sounds such as engine sounds, aerodynamic noise, and sounds from landing gear and flap deployment. In addition, messages can be given to pilots from simulated air-traffic controllers or on-board warning systems, or maneuver critiques from a surrogate controller.

Cockpit instrument simulation. All or some of the myriad of cockpit instruments are simulated, depending on the purpose of the simulator. In the past, simulator cockpit instrument panels were cluttered with numerous dials, indicators, switches, lights, and other devices by which the status of the simulated aircraft's systems could be monitored and controlled, reflecting the aircraft being simulated. Each of these devices was under computer control. However, modern aircraft have replaced many of these instruments with video display units, and the buttons, knobs, and switches by computer key pads. Hence, the simulator has followed suit, thereby simplifying the cockpit hardware substantially. *See* AIRCRAFT INSTRUMENTATION.

Validation. For simulators to be effective research, design, or training devices, their performance must be validated against the aircraft being simulated to the extent required for the mission of the simulator. This is frequently accomplished by ascribing tolerances to the simulator performance with respect to aircraft flight tests for the vehicle dynamics simulation and with respect to other test data for other systems. *See* SIMULATION.

Frank Cardullo

Flight Testing

Flight testing can be considered the final step in the proving of a flight vehicle or system as capable of meeting its design objectives. This definition applies whether the concept is a complete vehicle, a vehicle subsystem, or a research concept. Flight testing can be categorized as research, development, and operational evaluation. These categories apply both to aircraft and to spacecraft and missiles.

Research testing. The purpose of this form of flight testing is to validate or investigate a new concept or

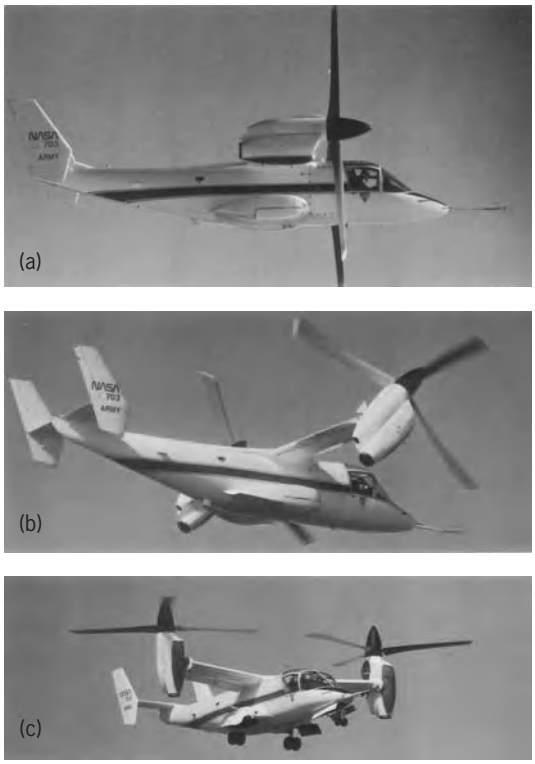

Fig. 10. XV-15 tilt-rotor research aircraft in (*a*) airplane, (*b*) conversion, and (*c*) helicopter mode. (*NASA*)

method with the goal of increasing the researchers' knowledge. Many times, the vehicle used is a one- or two-of-a-kind article designed specifically for the concept being investigated. One such concept aircraft is the XV-15 tilt-rotor (**Fig. 10**), of which two were made. This aircraft takes off with the rotors in a helicopter configuration and then undergoes a transition into a propeller configuration for high-speed forward flight. This concept and its testing proved so successful that it was adopted by the military for a new aircraft, the V-22 Osprey. *See* VERTICAL TAKEOFF AND LANDING (VTOL).

Research testing may also be done with an existing production aircraft to which modifications are made incorporating advanced technology. Such an aircraft is the F-111/AFTI, in which a production F-111 was modified to fly a mission-adaptable wing. Knowledge from research testing is used by designers in producing future generations of systems and is, in general, not applied to current in-service aircraft.

Development testing. A new vehicle or subsystem enters this phase of testing after it has been designed and the basic concepts proven in research flight testing. During this phase of testing, problems with the design are uncovered and solutions are developed for incorporation in the production aircraft.

Development testing involves the use of several preproduction aircraft, each assigned to specific testing tasks. These tasks can include stability and control, engine tests, performance evaluations, flutter and vibration investigations, air loads, avionics, weapons integration, and operational evaluation. These aircraft are highly instrumented, with specialized sensors designed to measure events associated with the task assigned. Testing is performed on all the aircraft simultaneously to expedite testing and field the system more rapidly.

Initial development testing of an aircraft is approached with caution: only after the aircraft has been shown to be controllable are maneuvering tests performed to prove structural or functional integrity to the limits of the design environment. These tests are performed carefully in a step-by-step process in which data from one step indicate if it is safe to proceed to the next step.

Aeroelastic stability testing includes both quasi-steady (divergence and control surface reversal) and dynamic (flutter and vibration) considerations. Quasi-steady aeroelastic effects are usually associated with loss of control effectiveness (the control surface provides less force for the same incidence angle) and a reduction in maneuverability (pitch, roll, or yaw) with higher speeds.

Dynamic flutter is a self-excited, often destructive response of the vehicle structure involving structural rigidity, inertias, and aerodynamic forces. This can be detected by an increase in strain or accelerometer readings as the flight condition for flutter is approached. In this way, with careful airspeed increases and close attention to the sensors, the condition of flutter can be detected before loss of the vehicle occurs. *See* FLUTTER (AERONAUTICS).

Vibration detection and reduction is an important consideration in testing since vibration affects not only the fatigue life of the structure but the reliability and accuracy of the on-board subsystems as well. Included in the fatigue considerations are the gross structural vibrations caused by such sources as power plants and aerodynamic forces and also the low-amplitude effects of acoustic fields. Frequency spectra from on-board accelerometers, strain gages, and microphones are used to determine how to reduce the forcing functions or where to modify the structure to withstand the vibratory forces and enhance fatigue life. *See* VIBRATION.

Air-load data are obtained from strain gages and are used in determining the adequacy of the structure and the accuracy of the designers' predicted loads and load paths. Such data are usually obtained during the initial development flight testing and determine the safe operational flight envelope for the airframe.

Performance testing. Once the final configuration is relatively fixed, performance testing is done to develop aircraft characteristic data. Such data include takeoff and landing distances, rate of climb, best-range airspeed, maximum level flight and limit speeds, and time to climb. Installed engine performance is also evaluated. These data are used to prepare the flight manuals for the vehicle and to verify that the design performance requirements were met.

Avionics testing. The term avionics is a contraction of aviation electronics and refers to the electronic devices and computers used in a flight system. With advances in electronics and computer science, the so-called black box systems have become more sophisticated and in some cases can include millions

of lines of code. Thorough testing of the avionics becomes increasingly important with the complex and information-intensive environment in which the pilot and vehicle have to perform. In general, avionics systems handle communication, identification, navigation, guidance and control, countermeasures, data processing, and information display. In most cases, each of these systems is linked through digital data buses with other systems, and they must interact with each other as well as with other aircraft subsystems such as controls and weapons.

The ultimate aim of avionics flight testing is to develop each system to effectively fulfill definite operational requirements and reliably interface with other systems to create a reliable and viable vehicle avionics system. The electronic subsystem is first tested in a laboratory environment to determine its standalone capabilities and correct any deficiencies found. It is then integrated into the vehicle and flight tested as a complete system to detect any deficiencies in interface design.

System demonstration is performed after the integration testing to determine if the applicable specifications and standards have been met. At this stage, not only is electronic behavior tested to determine correctness of programming, but the behavior of the electronics in the presence of environmental effects is tested as well. The durability and ability of the electronics to withstand the vibration, temperature, humidity, and electromagnetic environment of the operational envelope are determined and any limits are assessed.

Operational testing. This test phase involves customer participation to evaluate the capability of the fully equipped vehicle to meet its intended mission objectives. Testing is performed to determine system reliability, define maintenance requirements, and evaluate special support equipment. Military vehicles are also tested to determine weapon delivery techniques and effectiveness, including target acquisition capabilities, ability to perform in all weather conditions, operational behavior in battlefield conditions, and, in the case of naval aircraft, carrier suitability. Commercial aircraft are tested for blind landing-approach systems, passenger services, baggage and cargo loading, noise levels, and safety provisions. Crew training simulators, handbooks, and procedures are also tested in this phase to demonstrate the ability to maintain and operate the aircraft effectively.

Concurrent testing. An increase in the time necessary to fully test a system is required as the flight systems being fielded become more sophisticated and more electronic. This means that the traditional sequence of testing can delay fielding a system to the point that it is technologically outmoded by the time it is in service. One method of testing that is designed to save time in fielding a system is called concurrent testing. Concurrent testing involves starting production of a system before the development or operational testing is completed, with the idea that any modifications generated by the testing will be made as field upgrades to in-service systems.

Instrumentation. The quantities measured in flight testing have remained basically unchanged in character since the Wright brothers first flew at Kitty Hawk, North Carolina. Aircraft attitudes (such as pitch, roll, and yaw) and velocities (airspeed, pitch rate, roll rate, and yaw rate) are still the basic measurements. However, the accuracy, sophistication, and quantity

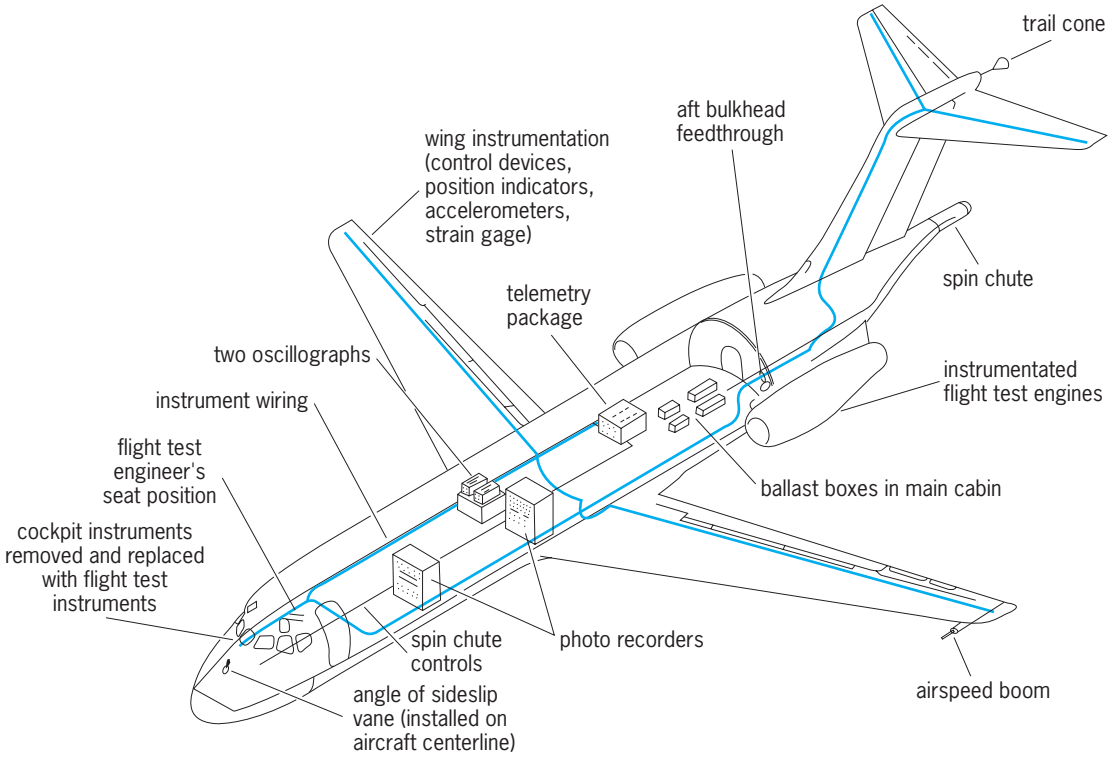

Fig. 11. Test instrumentation installed in a DC-9 aircraft for flight testing.

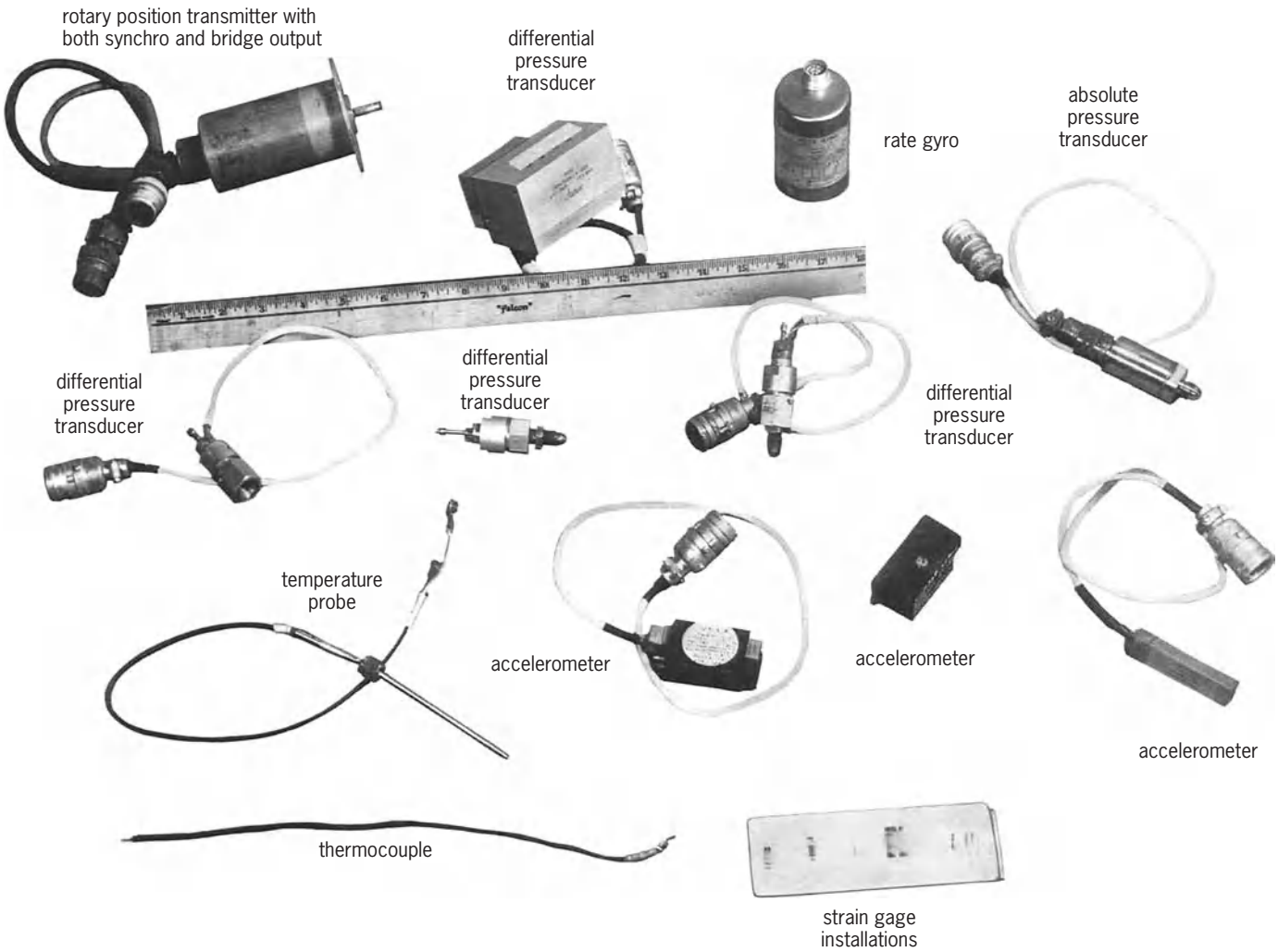

Fig. 12. Examples of miniaturized transducers used in flight testing.

of test data have made tremendous advances with the coming of the digital era. It is not uncommon for millions of bits of information to be present on data buses which transfer digital data from aircraft systems at high speed.

Special-purpose test equipment. To control conditions and to acquire specific data during test flights require the design, development, and installation of test equipment tailored to the constraints of the particular aircraft (**Fig. 11**). Hydraulically actuated aerodynamic vanes are installed at the wing tips to provide sinusoidal motions at controlled frequencies for flutter testing; a mechanically or ballistically deployed parachute is attached to the tail cone to ensure spin recovery during aerodynamic maneuvers; trailing cone airspeed static sources and wing- or nose-mounted airspeed booms are used to obtain distortion-free air samples for airspeed calibration; and dumpable water ballast systems are installed in large aircraft to control weight and center of gravity during flight tests. Specially designed pressure rakes measure inlet duct recovery efficiencies, and motion picture and television cameras document structural motions, information displays, and other test phenomena. Rain and icing environments are provided by flying the test aircraft in the path of water discharged from a tanker aircraft at altitudes where icing temperatures exist.

Test ranges. Test instrumentation can include equipment performing ground measurements as well as on-board systems. Ground instrumentation includes tracking systems (such as radar, laser, and optical) which give the location of the vehicle with reference to the ground. Such ground systems become extremely important in missile testing, in which the trajectory is as important as the vehicle state. *See* OPTICAL TRACKING SYSTEMS.

Transducers. The basic device used to measure phenomena is the transducer. Transducers provide a voltage output, change in impedance, or electric current proportional to the physical event that is being measured. In many cases this output is converted from analog (continuous in time) form to digital (comprising discrete time samples) form at the transducer. Common transducers are accelerometers, pressure sensors, rate gyros, attitude gyros, potentiometers, strain gages, and thermocouples. In general, these are electromechanical devices designed to measure such physical quantities as accelerations, airspeed, surface air pressure, altitude, angular rates about aircraft axes, aircraft attitudes, control and control surface positions, structural loads, and temperatures.

Because of severe weight and space constraints, most instrumentation used in flight testing has been miniaturized (**Fig. 12**).

Most electrical transducers employ the bridge principle, where the sensing element is in one or more legs of the bridge. Strain gages, for example, are strained through mechanical means, resulting in an unbalancing of the bridge with the current or voltage output from the bridge being proportional to the mechanical input. *See* BRIDGE CIRCUIT.

Transducers supply either analog or digital electrical outputs. An analog output consists of one signal which is continuous in time and amplitude. In a digital system the output is a sequence of discrete signals with finite differences in amplitude between them.

Data recording. Signals from the transducers either are transmitted to the ground for real-time data processing and recording or are recorded on on-board tape recorders or other instruments, or both. The real-time data processing and display include the traditional strip chart recorder, in which an ink pin or thermal print head is moved in response to an electrical signal on a moving strip of paper, giving a permanent record of the signals with time. Oscilloscopes can be used to capture events with a photographic record being taken. By far the method of choice is computer analysis and display of digital data. This method yields higher frequency responses and increased resolution. Additionally, digital analysis allows real-time reduction of the incoming data and display of derived parameters to aid the test director in evaluating how well the aircraft is achieving the desired conditions. The real-time ability saves repeating costly test flights to obtain hard-to-get data points, because the test director can ascertain if the data are good while the aircraft is in the air and does not have to wait for processing of the data until after the landing. *See* ANALOG-TO-DIGITAL CONVERTER; GRAPHIC RECORDING INSTRUMENTS; OSCILLOSCOPE; TRANSDUCER.

Oscillographs. The recording oscillograph is an automatic device using miniature galvanometers containing a small mirror that projects a light beam on photosensitive paper or film. As the galvanometer rotates in response to current output from the sensing device, the light beam is deflected across the photosensitive paper. Simultaneously, the paper is transported past an exposure slit at a fixed rate. The processed record thus contains a trace (line) that is a function of time with its deflection proportional to the physical stimulus. Timing lines, correlation counter readings, and a trace numbering system are also provided (**Fig. 13**).

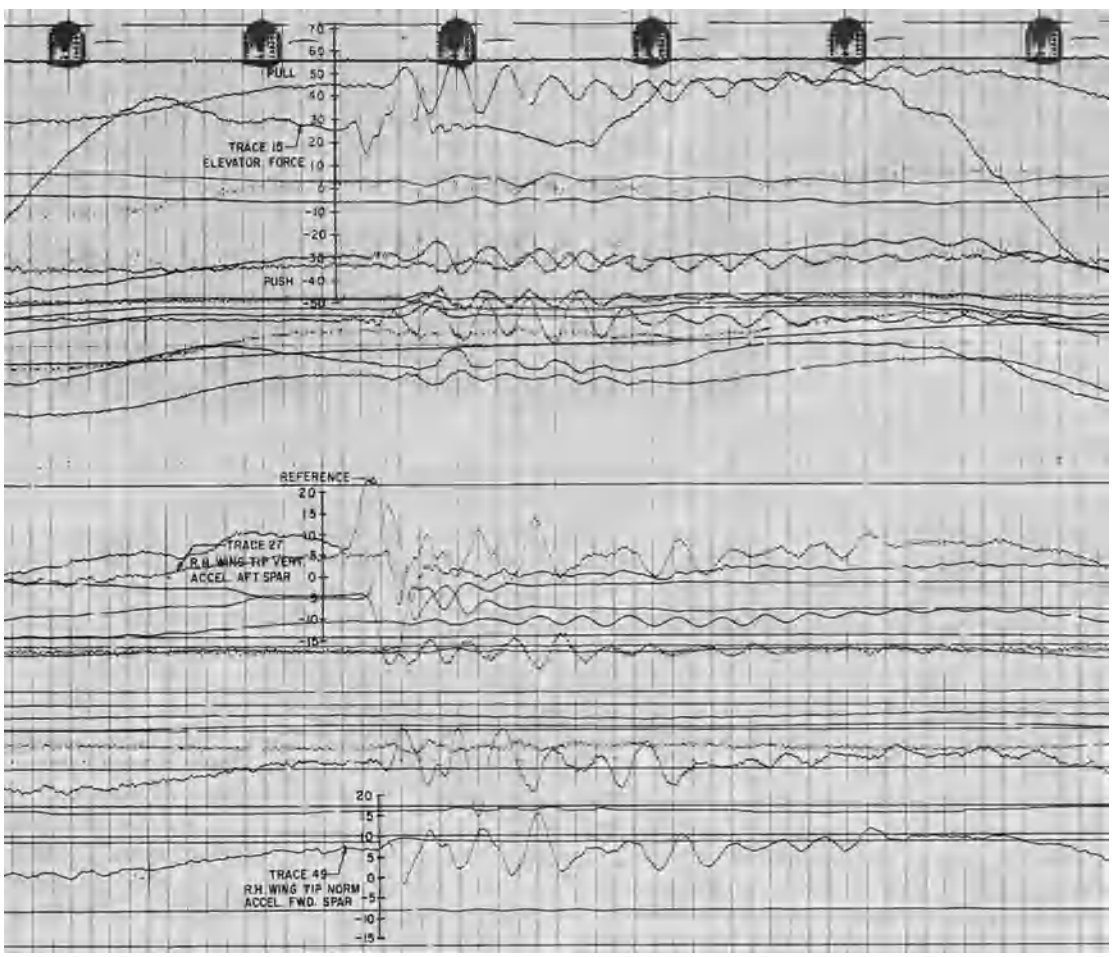

Fig. 13. Typical multichannel flight testing oscillograph.

Two types of recording oscillographs are available. One uses a tungsten lamp, requiring a closed magazine and wet chemical processing for the photosensitive paper. The other uses a mercury vapor lamp (direct write) producing paper lines that are brought into view by exposure to direct sunlight. This process may be accelerated by using a fluorescent latensifier. This oscillograph, because of its dry process and immediate delivery of data, is especially useful for airborne applications.

Oscillographs can record from 1 to 50 different signals. As many as nine recording oscillographs may be installed in one large aircraft.

Magnetic tape recorders. Magnetic tape recorders use the input signals from transducers to excite magnetic tape as a function of time. During playback, electronic equipment decodes the signals and produces a visual display. Alternatively, data from the tape can be processed automatically at high speed by using digital computers and special analyzers. *See* MAGNETIC RECORDING.

Analog magnetic tape recording systems are widely used both in the laboratory and in test aircraft. When a ground-based system is used, data are telemetered via one or more radio-frequency channels to a ground station recorder for future playback. Airborne systems record data signals directly, and the tape record is reproduced on the ground after the aircraft has completed its tests. Often airborne and ground tape recorders are used simultaneously to allow for limited analysis of the data while the aircraft is in flight and to provide a means of saving the data should the test vehicle be lost. Data are recorded in three different formats: direct analog, frequency modulation (FM), and pulse code modulation (PCM). Each format has its particular advantages, and for many tests all three techniques are used simultaneously. A single tape can simultaneously record output signals from several data systems through the use of multiple recording tracks, using a separate tape track for each input. Most instrumentation tape recorders (25-mm or 1-in. tape) contain 14 recording channels, while some have 28 recording tracks. *See* MODULATION.

The direct recording system preserves the varying amplitude input signal in the form available from the transducer. A recording amplifier provides the gain required to produce proper recording head currents. Frequency response of this technique is 100–100,000 Hz. A limitation is that each input signal requires a separate tape track.

FM recording involves converting the input amplitude signals to changes in a carrier frequency. Conversion is accomplished by an oscillator capable of frequency shifts induced by input amplitudes. Zero input signal causes the oscillator to assume a center frequency; plus and minus amplitude inputs produce frequencies above and below the center frequency, respectively. This technique thus results in recording a frequency-varying signal proportional to the amplitude of the input signal. Direct-current response (steady-state input signal) is achieved, and upper-frequency limits of 20,000 Hz can be obtained. As many as 12 separate signals can be recorded on one track of tape by using oscillators with different center frequencies; signals are then separated by filters during playback. Because the signals are in the form of a frequency analog, this system is free from errors caused by amplitude variations, but the technique requires exceptionally close speed regulation of the record and playback tape transport units. An adaptation to the system (called pulse amplitude modulation) is obtained by commutating many input channels into one frequency-modulated oscillator. This multiplexing provides more input channels, but at a drastic reduction in frequency response of each channel. *See* FREQUENCY MODULATION.

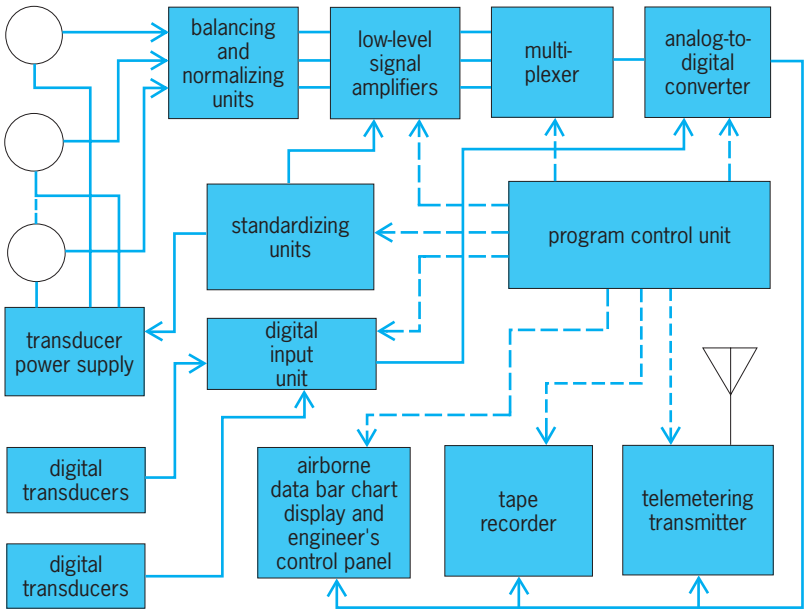

Fig. 14. Simplified functional diagram showing system for digital recording of data in aircraft and for PCM transmission to ground station during flight tests.

The airborne digital data system utilizing the PCM technique (**Fig. 14**) contains all equipment necessary to input, condition, format, record, and transmit data to the ground station. Low-level signal amplifiers multiply the millivolt output of each transducer for input to the analog-to-digital converter. Balancing and normalizing units provide for adjusting the electrical output to zero when zero physical stimulus is applied, and for attenuating the signal to the maximum acceptable input of the amplifier. Standardizing units enable the recording system and the transducer power supply to be calibrated for zero drift and sensitivity during flight. The multiplexor consists of electronic switches that gate transducer outputs in time sequence to the analog-to-digital converter. The analog-to-digital converter functions as an accurate voltage comparator and converts the transducer analog signal to a binary digital signal for recording on magnetic tape. The program control unit is the clock that time-sequences and coordinates all functions.

Sampling and digitizing rates of 1,000,000 data bits per second and tape recording densities of 6250 data bits per inch per track have been attained by using

parallel recording techniques, in which each binary bit level is recorded on a separate track. Developments in tape transport technology permit serial recording of the multiplexed data stream on one recorder track with packing densities up to 14,000 bits per inch. The remaining tracks of a standard 14-track tape transport can be used to record outputs of several data systems or to record direct analog, or FM, data. Digital data may be telemetered by pulse code modulation and recorded on the ground without loss of accuracy. *See* PULSE MODULATION.

Digital systems are characterized by a high resolution and accuracy. Because data are digitized as they are sensed, data contamination is held to a minimum, and once recorded in digital form, no errors are produced in recovery of the information. The system is highly flexible with capacities of 1 to 10,000 channels and frequency response up to 20,000 Hz. Digital systems have the advantage over analog systems that data are recorded in a form that is easily transcribed for entry into digital computers. These systems are normally designed with data-editing and -transcribing facilities as part of an integrated overall data acquisition and processing system.

Aircraft test data processing. Most final test data are processed and analyzed automatically in general-purpose digital computers. A major part of any test effort is the preparation of data for entry into the computer. When a photographic recorder is used, this preparation consists of reading and tabulating the instrument indications and keying them into a computer.

Semiautomatic means may be used to transcribe data from an oscillogram. These devices normally consist of an analog reading device and an analog-to-digital converter. The operator establishes a zero reference on the record and then moves cross hairs connected to potentiometers to intersect the recorded trace at the desired point. The electrical output of these potentiometers, proportional to the displacement of the cross hairs, is fed to the analog-to-digital converter, and the corresponding digital value is automatically entered into the computer. An average operator can prepare 400–600 data values per hour for computer entry by using this equipment.

The major advantage of the all-automatic magnetic tape recording system is the rate at which the desired data can be selected and written on computer magnetic tape for direct entry into the computer. In these systems it usually requires no more time to transcribe thousands of data values from a given test than it does a few. Typical transcription rates are in the order of 5000 data values per second. Thus these systems considerably reduce the cost of data reduction and processing.

Advanced cathode-ray tubes (CRT) for automatic data processing reduce the overall time of final data analysis by permitting a person-machine interface during either real time (telemetry) or post-flight reduction of airborne recorded magnetic tapes. Rather than requiring individual editing, transcribing, and preprogrammed computer data processing, this technique provides a single, high-speed process. The CRT displays, in conjunction with computer interactive hardware and software, provide direct communication with the tape signals by using the digital computer and peripheral processing equipment to receive and decode the incoming data stream. The data engineer controls display format through the use of control keyboards and photoelectric pens. These devices control the path of the data processing by switching to various routines within the computer program, and entering numeric or alphabetic information, such as channel number, scale ranges, and titles. Multiple CRT displays and interactive hardware allow simultaneous processing of more than one flight tape and hasten the data processing on any one flight tape or telemetry stream. As each graph or tabular listing is completed, a reproducible hard-copy print-out can be obtained through a slave CRT copier for a permanent record. These records are available for inclusion in formal test reports. *See* AIRPLANE; CATHODE-RAY TUBE.

Michael Watts; Donald W. Douglas, Jr.

Bibliography. *NASA Aeronautics*, annually; J. M. Rolfe and K. J. Staples (eds.), *Flight Simulation*, 1986; D. T. Ward, *Introduction to Flight Test Engineering*, 2d ed., 1998.

Airfoil

The cross section of a body that is placed in an airstream in order to produce a useful aerodynamic force in the most efficient manner possible. The cross sections of wings, propeller blades, windmill blades, compressor and turbine blades in a jet engine, and hydrofoils on a high-speed ship are examples of airfoils. *See* COMPRESSOR; PROPELLER (AIRCRAFT); TURBINE PROPULSION; WIND POWER; WING.

Geometry. The mean camber line of an airfoil (**Fig. 1**) is the locus of points halfway between the upper and lower surfaces as measured perpendicular to the mean camber line itself. The most forward and rearward points of the mean camber line are the leading and trailing edges, respectively. The straight line connecting the leading and trailing edges is the chord line of the airfoil, and the distance from the leading to the trailing edge measured along the chord line is simply designated the chord of the airfoil, represented by c. The thickness of the airfoil is the distance from the upper to the lower surface, measured perpendicular to the chord line, and varies with distance along the chord. The maximum thickness, and where it occurs along the chord, is an important design feature of the airfoil. The camber is the maximum distance between the mean camber line and the chord line, measured perpendicular to the chord line. Both the maximum thickness and the camber are usually expressed in terms of a percentage of the chord length; for example, a 12% thick airfoil has a maximum thickness equal to $0.12c$.

Aerodynamic forces. The airfoil may be imagined as part of a wing which projects into and out of the page, stretching to plus and minus infinity. Such a wing, with an infinite span perpendicular to the page, is called an infinite wing. The aerodynamic force on the airfoil, by definition, is the force exerted

on a unit span of the infinite wing. For this reason, airfoil data are frequently identified as infinite wing data.

The flow of air (or any fluid) over the airfoil results in an aerodynamic force (per unit span) on the airfoil, denoted by R (**Fig. 2**). The relative wind is the magnitude and direction of the free-stream velocity far ahead of the airfoil. The angle between the chord line and relative wind is defined as the angle of attack of the airfoil, denoted by α. By definition, the component of R perpendicular to the relative wind is the lift, L; similarly, the component of R parallel to the relative wind is the drag, D.

The airfoil may be visualized as being supported by an axis perpendicular to the airfoil, and taken through any point on the airfoil. The point that is one-quarter of the chord distance from the leading edge, the so-called quarter-chord point, is frequently chosen (Fig. 2). The airfoil has a tendency to twist about this axis; that is, there is an aerodynamic moment exerted on the airfoil. By definition, the moment is positive or negative if it tends to increase or decrease respectively the angle of attack (that is, if it tends to pitch the airfoil up or down, respectively). The moment about the quarter-chord point is designated $M_{c/4}$.

Airplane flight. Airplane wings are made up of airfoil sections. Airplanes are sustained in the air by the production of lift on the wings. The physical manner by with an airfoil generates lift is based on the fact that an increase in velocity of air is always accompanied by a decrease in pressure. This is frequently called the Bernoulli effect and is quantified by the differential equation $dp = -\rho V dV$, where ρ is the fluid density, V is the fluid velocity, dV is an infinitesimally small change of velocity, and dp is the corresponding infinitesimally small change in pressure. This equation is simply a statement of Newton's second law (force = mass × acceleration) applied to a moving element of the air. If the velocity increases (dV is positive), the pressure decreases (dp is negative). When an airfoil is producing lift (Fig. 2), the average flow velocity is higher over the top surface and lower over the bottom surface. In turn, the average pressure is lower over the top surface and higher over the bottom surface. The net difference between these pressures results in a net upward force, or lift. For most conventional airplanes in cruise, relative to the freestream pressure the average pressure over the top surface of the wing changes more than the average pressure over the bottom surface; hence typically about 70% of the lift of an airfoil is due to the lower pressure on the top surface. *See* COMPRESSOR.

The reasons that the flow over the top of the airfoil moves faster than that over the bottom are based on another fundamental principle, mass conservation. The lifting airfoil presents a type of obstruction to the flow, and the individual streamtubes of air that curl around the leading edge and flow downstream over the top are squashed, causing the flow in these streamtubes to speed up in order to preserve a constant-mass flow. This occurs even when the airfoil is a flat plate. Indeed, for a flat plate at an angle of attack, the streamtubes that curl around the leading edge are especially squashed, leading to a very large increase in velocity in that region. This is why paper airplanes can fly. However, for such a flat plate at a moderate angle of attack, the flow separates over the top surface and causes a high drag. The smooth, streamlined shape of an airfoil is designed to avoid flow separation; hence airfoils produce lift with only a small amount of drag.

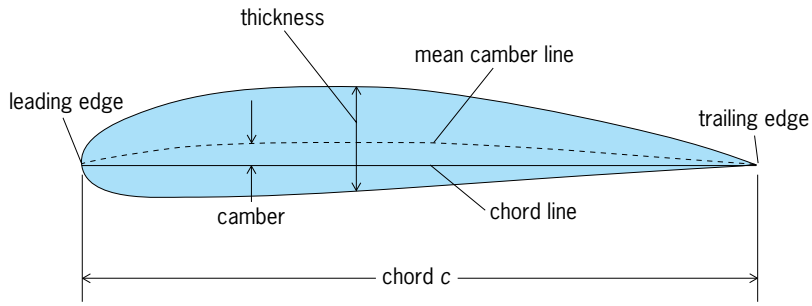

Fig. 1. Airfoil nomenclature. The shape shown is an NACA 4415 airfoil.

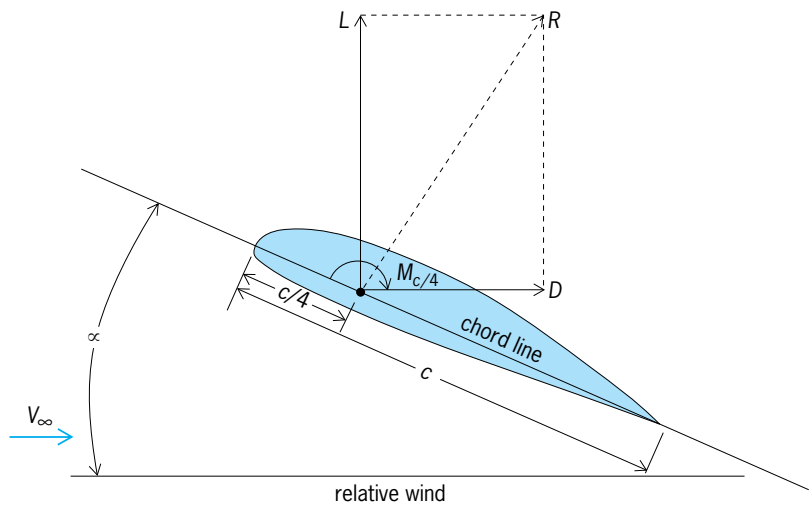

Fig. 2. Definitions of lift, drag, moments, angle of attack, and relative wind for an airfoil.

Dimensionless coefficients. The dynamic pressure q_∞ is defined by Eq. (1), where ρ_∞ and V_∞ denote

$$q_\infty = {}^1/_2\, \rho_\infty V_\infty^2 \tag{1}$$

the free-stream density and velocity. With L and D as the lift and drag per unit span, the airfoil lift and drag coefficients c_l and c_d are defined by Eqs. (2) and (3). Similarly, with M as the moment per unit span, the moment coefficient c_m is defined by Eq. (4).

$$c_l = \frac{L}{q_\infty c} \tag{2}$$

$$c_d = \frac{D}{q_\infty c} \tag{3}$$

$$c_m = \frac{M}{q_\infty c^2} \tag{4}$$

See AERODYNAMIC FORCE; AERODYNAMICS.

The aerodynamic characteristics of airfoils, as couched in terms of the lift, drag, and moment coefficients, are quite different depending on whether the flow is low speed (somewhat less than Mach 1) or high speed (near or above Mach 1). The reason is that the physical characteristics of subsonic flow are quite different from those of transonic or supersonic flow.

In low-speed, subsonic flow over a typical airfoil the lift coefficient c_l varies linearly with the angle of attack α at low values of α. At higher angles of attack, the lift coefficient reaches a maximum value denoted by $c_{l,\max}$, and then decreases as the angle of attack is further increased. In this region, where c_l the lift coefficient rather dramatically decreases, the airfoil is said to be stalled. Positive lift exists at zero angle of attack, and the airfoil must be pitched to some negative angle of attack for zero lift (called the zero-lift angle of attack). This behavior is characteristic of all positively cambered airfoils; positive camber occurs when the camber line lies above the chord line (Fig. 1).

The lift coefficient also depends on the Reynolds number Re, defined by Eq. (5), where μ_∞ is the co-

$$\mathrm{Re} = \frac{\rho_\infty V_\infty c}{\mu_\infty} \tag{5}$$

efficient of viscosity in the free stream. In the linear region, the Reynolds number has virtually no effect. However, at higher angles of attack, the value of $c_{l,\max}$ is dependent on the Reynolds number; higher values of $c_{l,\max}$ are achieved at higher Reynolds numbers. The drag coefficient c_d is also sensitive to both the angle of attack and the Reynolds number. *See* REYNOLDS NUMBER.

Compressibility effects. High-speed flow over an airfoil, but with a subsonic free stream, depends on the free-stream Mach number, which is the ratio of the velocity to the speed of sound in the free stream. When the free-stream Mach number is higher than 0.3 but still less than 1.0, compressibility effects occur which influence both the lift and drag coefficients, and the variation with Mach number becomes a primary consideration. For subsonic flow, the lift coefficient increases with the free-stream Mach number. A simple, and historically the first, compressibility correction to the lift coefficient is the Prandtl-Glauert rule: If $c_{l,0}$ denotes the low-speed value of the lift coefficient, then the value of the lift coefficient for the same airfoil at the same angle of attack, but with a free-stream Mach number Ma_∞, is given by Eq. (6). The rule holds reasonably accurately up

$$c_l = \frac{c_{l,0}}{\sqrt{1 - \mathrm{Ma}_\infty^2}} \tag{6}$$

to Mach numbers of about 0.7. *See* GAS DYNAMICS; SUBSONIC FLIGHT.

Transonic flow. Airfoil characteristics at high subsonic Mach number are associated with what is known as the transonic flight regime. When the free-stream Mach number approaches 1, the flow field over the airfoil begins to exhibit some dramatic changes. An index for the threshold of these changes is the critical Mach number. When the flow expands over the top surface of the airfoil, the local velocity increases above that for the free stream. Hence, even though the free-stream Mach number is subsonic, if it is close enough to unity there can be regions of locally supersonic flow over the airfoil. By definition,

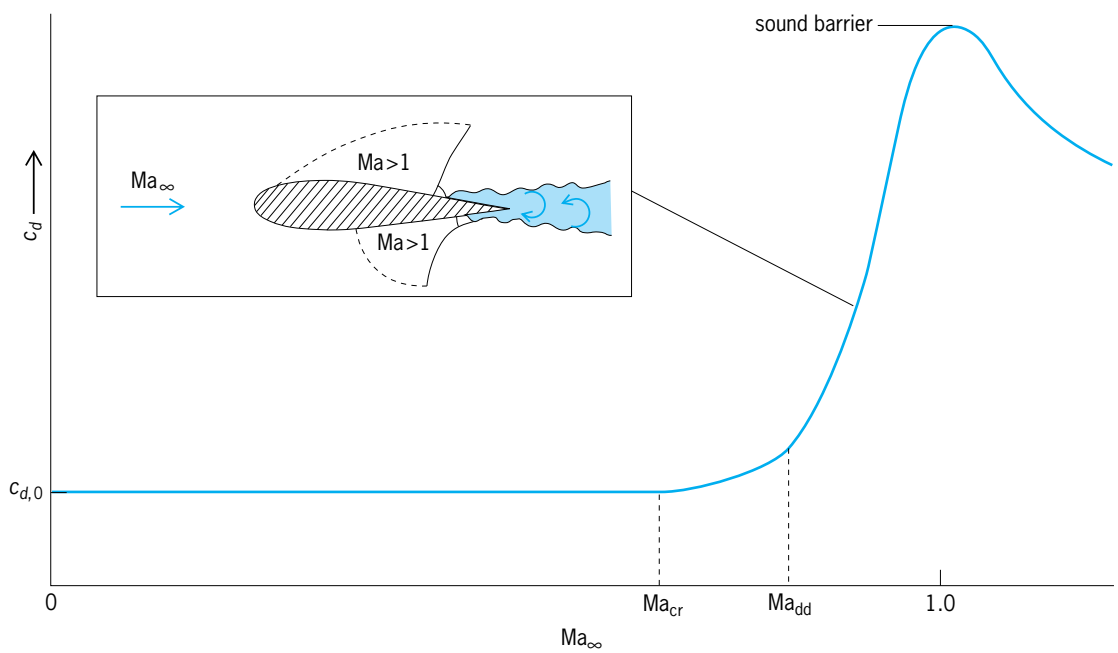

Fig. 3. Variation of the profile drag coefficient with the free-stream Mach number Ma_∞, illustrating the critical Mach number Ma_{cr} and drag-divergence Mach numbers Ma_{dd} and showing the large drag rise near Mach 1. Inset shows the flow field at velocities above that of the drag-divergence Mach number, with large regions of locally supersonic flow (Ma > 1).

for a given airfoil at a given angle of attack, that free-stream Mach number at which sonic flow is first encountered at a point on the airfoil is the critical Mach number. When the free-stream Mach number is increased slightly above the critical Mach number, whole regions of locally supersonic flow occur over the airfoil (**Fig. 3**). These supersonic regions are terminated by a shock wave. If the shock wave is strong enough, its interaction with the boundary layer separates the flow at the shock impingement point. Both the shock wave itself and the flow separation caused by the shock contribute to a massive increase in drag and a decrease in lift.

The onset of these phenomena is best described by considering the measurement of the drag coefficient of a given airfoil at a fixed angle of attack in a wind tunnel, as a function of Mach number (Fig. 3). The drag coefficient remains relatively constant as the Mach number is increased from an initial low value all the way up to the critical Mach number. As the Mach number is increased slightly above the critical value, a finite region of supersonic flow appears on the airfoil. With a further small increase in Mach number, a point is encountered where the drag coefficient suddenly starts to increase. The value of the Mach number at which this sudden increase in drag starts is defined as the drag-divergence Mach number. Beyond the drag-divergence Mach number, the drag coefficient can become very large, typically increasing by a factor of 10 or more. The peak value of the drag coefficient was so daunting to aeronautical engineers in the 1930s and early 1940s that it gave rise to the concept of the sound barrier, that is, that airplanes could not fly faster than sound (faster than Mach 1). Of course, there is no such barrier to supersonic flight. The supersonic transport Concorde cruises at Mach 2.2, and some high-speed military airplanes can fly at Mach 3 and above.

The advent of jet propulsion in the 1940s ushered in the era of high-speed flight and focused attention on designing airfoils with high critical Mach numbers, thus trying to delay the transonic drag divergence to higher subsonic Mach numbers. One well-known approach was to use thin airfoils on high-speed aircraft; everything else being equal, a thin airfoil has a higher critical Mach number than a thicker airfoil. In the 1960s, a new airfoil design philosophy was developed, the supercritical airfoil. For the same free-stream Mach number, the supercritical airfoil is shaped so as to produce a smaller region of supersonic flow with a lower local supersonic Mach number than the more conventional airfoil shape. Because of this behavior, the supercritical airfoil has a larger gap between the critical and the drag-divergence Mach numbers; that is, it can penetrate farther into that part of the free-stream Mach number regime above the critical Mach number before encountering the drag-divergence phenomena. *See* SUPERCRITICAL WING; TRANSONIC FLIGHT.

Supersonic flow. When the free-stream Mach number is greater than 1, the physical aspects of the flow over an airfoil are quite different from those for a free-stream Mach number less than 1. In supersonic flow, shock waves and expansion waves dominate the flow pattern. Consequently, the pressure distribution over an airfoil in supersonic flow is quite different than that for subsonic flow, giving rise to a major component of drag called wave drag. Supersonic airfoil shapes are characterized by sharp leading edges and thin profiles. Both features reduce the strength of the shock waves and hence reduce the wave drag. *See* SUPERSONIC FLIGHT. John D. Anderson, Jr.

Bibliography. I. H. Abbott and A. E. von Doenhoff, *Theory of Wing Sections*, 1949; J. D. Anderson, Jr., *Aircraft Performance and Design*, 1999; J. D. Anderson Jr., *Fundamentals of Aerodynamics*, 4th ed., 2005; J. D. Anderson, Jr., *A History of Aerodynamics*, 1997; J. D. Anderson, Jr., *Introduction to Flight*, 5th ed., 2005.

Airframe

The structure consisting of the wings, fuselage, and fin and tail plane (vertical and horizontal stabilizer) of an airplane. This structure has a clearly defined role: to resist all the extreme loads that can be experienced with no danger to the passengers, whether the aircraft be civil or military. The temptation therefore could be to overdesign the airframe and "play safe." But an overweight civil aircraft would be commercially disastrous, and an overweight military aircraft could be outperformed. To put this matter into perspective, a civil airframe can be up to 40% of the total dry weight, whereas the payload may be as low as 20%. Thus a structure which is overdesigned by only 10% could reduce the payload by 20%. This represents the profit margin on many routes.

Design loads. An airframe has to withstand two different types of loading: a maneuver and a sudden gust. A maneuver is intentional, and when a pilot increases the aircraft incidence by applying a download on the tail plane the main wing experiences a sudden increase in lift and accelerates upward. The resulting inertia forces experienced by all the masses can be visualized as balancing the extra lift (**Fig. 1**). This extra inertia is conventionally referred to as the number of g, where 1 g is the normal weight due to gravity. For an agile combat military aircraft the specification may call for a factor of 9 g or more, depending on the assistance given to the pilot by pressurized suits and whether the maneuver lasts for

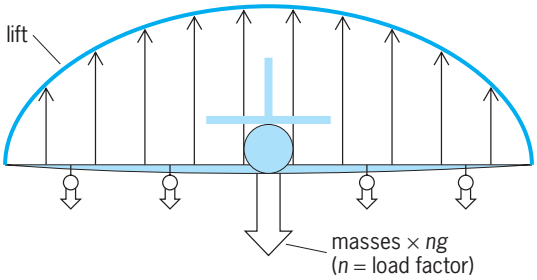

Fig. 1. Forces acting on the main wing during a maneuver.

many seconds. For a civil aircraft the current allowable acceleration is $+2.5\,g$ or $-1\,g$ "downwards." *See* AERODYNAMIC FORCE.

Civil aircraft also have to be designed to resist sudden gusts. A vertical upward gust of velocity u will effectively increase the incidence of the aircraft by u/V radians, where V is the aircraft velocity. Currently the maximum specified gust is 20 m/s (or 45 mi/h). This figure represents the maximum ever recorded in flight. However, all maneuver and gust loadings are multiplied by 1.5 as a safety factor when designing civil aircraft. This somewhat arbitrary figure has withstood the test of time. Structural failures are almost unheard of, provided the specified maintenance is followed. There have been a very small number of cases where pilots have overloaded an aircraft, but nowadays a computer between the pilot and the controls will impose software limitations on control forces. Certain parts of the airframe may be designed to withstand rather special loads, such as the undercarriage in heavy landings.

Materials behavior. To discuss airframe design, it is first necessary to describe the behavior of airframe materials. The simplest way to test a material (say, a simple rod) is to pull it and measure the extension due to the applied force. However, a better measure of the actual forces on the material molecules or crystals is the stress (σ), that is, force divided by cross-sectional area. Similarly, a better measure of the crystal deformation is the strain (ε) or fractional change in length of the rod. A final important property is the material's stiffness (E), which is the ratio of stress to strain for small strains. Some typical stress/strain curves for aircraft materials are shown in **Fig. 2**. *See* STRESS AND STRAIN.

The most common structural material in use worldwide is mild steel for which the strain is proportional to stress (a linear curve) until "yield," when the crystals start to move along their boundaries or dislocations. This material then deforms at virtually constant stress, and then when unloaded will do so linearly as shown in Fig. 2, leaving a permanent deformation of, say, 7%. This particular property is extremely valuable since it allows thin steel sheets to be pressed into everything from car bodies to domestic cutlery.

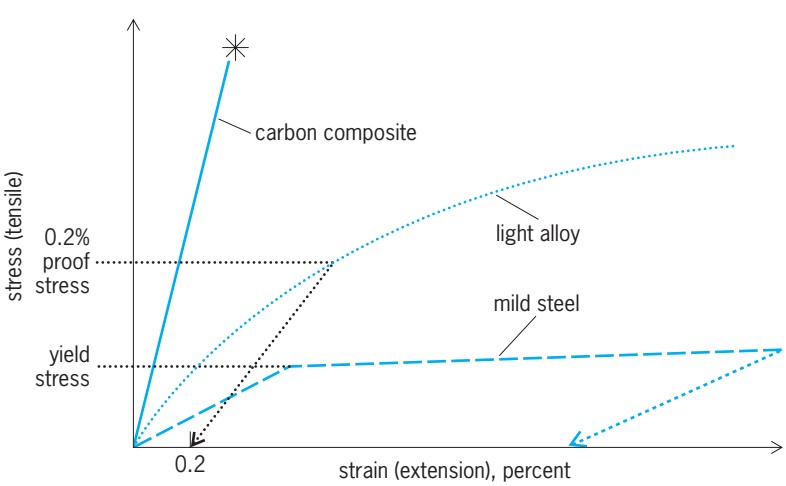

Fig. 2. Stress/strain curves for various types of material.

Fig. 3. Effect of wing box strains of 0.7%.

A similar behavior accurs in aluminum, but by adding small amounts of copper, zinc, magnesium, and so forth, the dislocations can be inhibited at key points. The maximum stress can therefore be increased as shown in Fig. 2. Light aluminum alloys have been the most common airframe material from the early days. To define the maximum allowable stress (in the absence of a true yield), the aircraft sector has concentrated on the allowable strain when the material is unloaded. (Clearly an aircraft should return from flight with the same shape as it took off.) The atypical allowable is a strain of 0.2%, and the corresponding stress is called the 0.2% proof stress. Although these strains may seem very small, an aircraft wing, for example, is not a very efficient structure in bending, and if the strains reached 0.7% the deformations would look something like **Fig. 3**. *See* METAL, MECHANICAL PROPERTIES OF; PLASTIC DEFORMATION OF METAL.

To obtain higher failure stresses, by avoiding dislocations and extending the linear range, it is necessary to turn to brittle materials like glass, carbon, boron, and various ceramics, all of which have a much purer crystalline structure. However, unlike ductile metals, these materials are very susceptible to flaws and minor cracks. In fact, the standard way that a glazier will cut glass is by inscribing a shallow surface scratch and then propagating it. One way of overcoming this 'notch sensitivity' is to manufacture extremely fine fibers and then embed millions of them in a resin matrix. Fibers still have surface flaws, but typically the distance between flaws is about 15–25 diameters for carbon fibers. Thus, when a single fiber breaks, the load is diffused onto the surrounding fibers since the chances of all fibers having cracks at the same location are zero. A common way of using carbon fiber composites is to make a thin lamina of about 0.13 mm (0.005 in.) thickness in which all the thousands of fibers lie in one direction. A simple laminar (or ply) is extremely stiff and strong in the direction of these fibers but very weak in a direction at right angles to them (like balsa wood). A composite wing skin, for example, is then built by stacking a large number of plies in various directions to suit the nature of the internal stress field. Composite structures are consequently "tailored," unlike structures of homogeneous metals.

Carbon composite structures have dominated combat military aircraft since the 1980s, since they are exceptionally stiff and strong. The **table** shows

Properties of some materials used in aircraft structures

Material	E, GPa	σ, MPa	ρ, Mg/m^3	E/ρ	σ/ρ
Aluminum alloy	71	480	2.8	25	172
Titanium	110	1000	4.4	25	226
High-tensile steel	207	1720	7.8	26	221
Carbon composite	200	1500	1.55	130	970
Glass composite	40	840	1.80	22	460
Kevlar composite	82	1500	1.39	60	1080

the values of strength (ultimate tensile strength, σ) of stiffness (E), and of density (ρ) of some materials used in aircraft structures. The specific stiffness and strength (ratios of stiffness and strength, respectively, to density) are the crucial measures since there is no point in having a superior material that is too heavy.

The values in the last two columns of the table indicate that the specific stiffnesses of the three metals are comparable, but clearly carbon composites are superior. Carbon also has better specific strength, matched only by Kevlar. This artificial aramid fiber is cheaper than the other structural materials but is competitive only in tension; it has a very poor compressive strength. *See* COMPOSITE MATERIAL; MANUFACTURED FIBER.

Wing design. An aircraft wing is not a very efficient structure in resisting bending due to gust or maneuver loadings. Clearly, aerodynamic efficiency is the driver, and hence a wing has a small thickness/chord ratio, as low as 1/20 for high-speed military aircraft. When a wing bends due to aerodynamic lift, the top surface goes into compression, the lower into tension. For maximum efficiency, most of the structural material should go to these surfaces, as far as possible away from the center of the section (the neutral axis), where the bending strains are zero. The ubiquitous I section, used in buildings and so forth, is a good example. However, a wing also has to resist torsion, and should twist as little as possible to avoid changing the effective angle of incidence. An I section is hopeless at resisting torsion. For maximum torsional stiffness, a thin-walled structure should be a closed "tube" with as large an enclosed area as possible. In a wing, a fuel tank also fulfills this specification.

A typical wing section, with a front and rear spar, is shown in **Fig. 4**. The trailing edge also contains the flaps and ailerons, and the leading edge may contain deployable slats for increased lift. The upper skin, when in compression, has to be designed to resist buckling. The stress at which a thin plate buckles is proportional to $(t/b)^2$, where t is the plate thickness and b is the width. This property explains the large number of stiffeners shown diagrammatically in

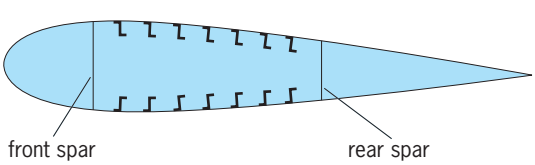

Fig. 4. Typical wing box section.

Fig. 5. Wing rib for taking engine mounting on Airbus 380 aircraft. (*Courtesy of Airbus UK*)

Fig. 4 and intended to reduce the effective width b. The buckling stress is also proportional to $(1/l)^2$, where this time the plate length l is involved and is reduced by inserting ribs into the wing box. These ribs are quite light since their only function is to stop the outer skin's buckling and to maintain their shape for aerodynamic reasons; a large aircraft may have 40 or 50 of them along the wing. Some ribs are heavily loaded, however; for example, the engine loads have to be diffused into the wing box. **Figure 5** shows an engine mounting rib (for the Airbus 380), which is clearly heavy and has an array of stiffeners (machined from the solid) that inhibit buckling under the high shear forces coming from the engine. Where the top edge of the rib is flanged (to be bolted to the wing skin), the cut-outs are seen to allow the wing skin stiffeners to pass through without interruption. *See* AILERON.

The fin (vertical stabilizer) and the tail plane (horizontal stabilizer) have very similar structures to the main wing, although they are much lighter. Actually, their structures should be exceptionally light since they are at the extreme end of the aircraft and if too heavy would adversely move the aircraft's center of gravity. For this reason all Airbus (and now Boeing) fins and tail planes have been made of carbon composites.

A final hazard for the wing designer is the interaction between the wing deflections and the aerodynamic forces: an aeroelastic effect. If the wing lacks sufficient torsional stiffness, it may twist and thereby increase the aerodynamic lift due to the increase in incidence. At a critical divergence speed, the twist may grow beyond limit and lead to structural failure. Another phenomenon, called flutter, occurs when the balance between aerodynamics, structural stiffness, inertia forces, and damping leads to oscillations that may grow without limit. In the absence of aerodynamic forces, the vibration of any structure will gradually subside due to natural material damping. However, the aerodynamic lift and pitching moment are proportional to the surface velocity and result in negative damping. One solution is to introduce mass balancing for controls like the aileron and rudder, so that the extra inertia forces dampen the

oscillations. On a main wing, the forward-placed engines fill this role and ensure a sufficiently high flutter speed. *See* AEROELASTICITY; AIRCRAFT RUDDER; FLIGHT CONTROLS; FLUTTER (AERONAUTICS); WING; WING STRUCTURE.

Fuselage design. The fuselage has a much larger cross section than a wing and is consequently much better able to resist bending and torsion. A fuselage skin may be only 1–2 mm (0.04–0.08 in.) thick, compared to wing skin thicknesses in excess of 20 mm (0.8 in.). Parts of the fuselage also experience bending compression during landing and takeoff, for example. Therefore, as with wings, there is a need for stiffeners and ribs, although the "ribs" are really frames of depth 10–20 mm (0.4–0.8 in.). Heavier frames are needed to take the wing loads into the fuselage, for example, or at undercarriage pickup points. At the rear of the fuselage, the fin or tail plane is often connected to the curved bulkhead, designed to take the cabin pressure loading.

Although the working stresses in a fuselage can be low, the design case is really one of fatigue. On every flight, the external pressure at altitude falls to a very low value, while the cabin pressure is maintained at about half an atmosphere for comfort. Thus, in an operating lifetime the fuselage stress field will be cycled thousands of times. All metallic materials can fail at a relatively low stress after a sufficiently large number of cycles. The response of carbon composites is much better since the loading and unloading seems to be reversible with no material damage even at the crystal scale.

The presence of windows and doors can exacerbate the fatigue problem by introducing local stress concentrations, but designers now know how to design window and door frames to minimize these concentrations. Unfortunately, this was not the case in the 1950s when the Comet jetliner, the first civil aircraft to cruise at altitudes in excess of 10,000 m (30,000 ft), had fuselage fatigue failures in three aircraft before it was grounded. *See* FUSELAGE.

Composite structures. The superiority of carbon composite materials has already been mentioned, that is, their strength, stiffness, and fatigue performance. In spite of their basic advantages, there are two disadvantages that have to be overcome, and indeed mostly have been.

1. Carbon composite materials are susceptible to stress concentrations. Laminated thin plates are strong in their own plane but have little resistance to peeling apart of the individual plies. They can easily delaminate at the matrix resin strength. Ideally, all laminates should be in a two-dimensional stress field. But stresses will be three-dimensional near any discontinuity of geometry or load. Careful design and local reinforcements can eliminate this problem. The Airbus 380, for example, has a totally composite center wing box where the two wings intersect the fuselage. This is a region full of joints and discontinuities, and the design is a tribute to the skill of manufacturers and designers who are able to model this problem and analyze the stress fields.

2. Carbon composites can be expensive. Carbon and boron fibers are costly to produce in small numbers. For a long time, the total output of carbon fiber for sporting goods exceeded that needed by the aerospace industry. Only military aircraft could disregard this expense since in that case a poor performer demanded rejection.

Not only was the basic material costly, but so were the manufacturing methods. Hand lay-ups of many thin plies (possibly up to 80 in military aircraft) are costly. This cost has gradually been reduced by replacing hand lay-ups with computer-controlled automatic machines, which lay a large number of overlapping tapes with great precision. Another breakthrough has been resin transfer molding, in which a dry fabric can be sandwiched between two faces of a mold and then the resin injected, before cure, under pressure or sucked through under vacuum.

The Boeing 787 is an impressive example of a civil aircraft with composite wings and fuselage, having overcome all these problems. Possibly the only threat to the economics of mass-produced airframes is the future legal requirement that manufacturers recycle carbon structures when their operating life is over.

Virtual testing. To satisfy either military or civil airworthiness authorities, aircraft have to be tested to ultimate design loads. Prior to this, the manufacturer will have made many component tests on panels, ribs, undercarriages, and so forth. These very expensive and time-consuming tests have been traditional since no theories were available for highly three-dimensional and detailed components. However, since the 1980s remarkable progress has been made in the ability to simulate in a computer model the behavior of structures loaded to failure. Basically there are only three sorts of equations to satisfy in theoretical structures: (1) the stresses everywhere have to be in equilibrium with themselves and the applied forces; (2) the strains have to be related to the displacement field (purely a geometrical argument); and (3) the two are linked by a stress/strain law based on material tests. The main problem has been the enormous geometrical complexity of aircraft structures at the detailed level, but the advent of the digital computer (and later

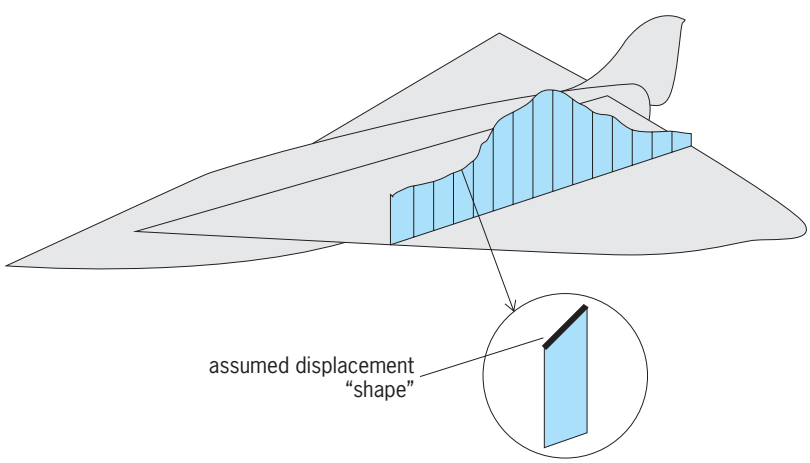

Fig. 6. Finite element approximation.

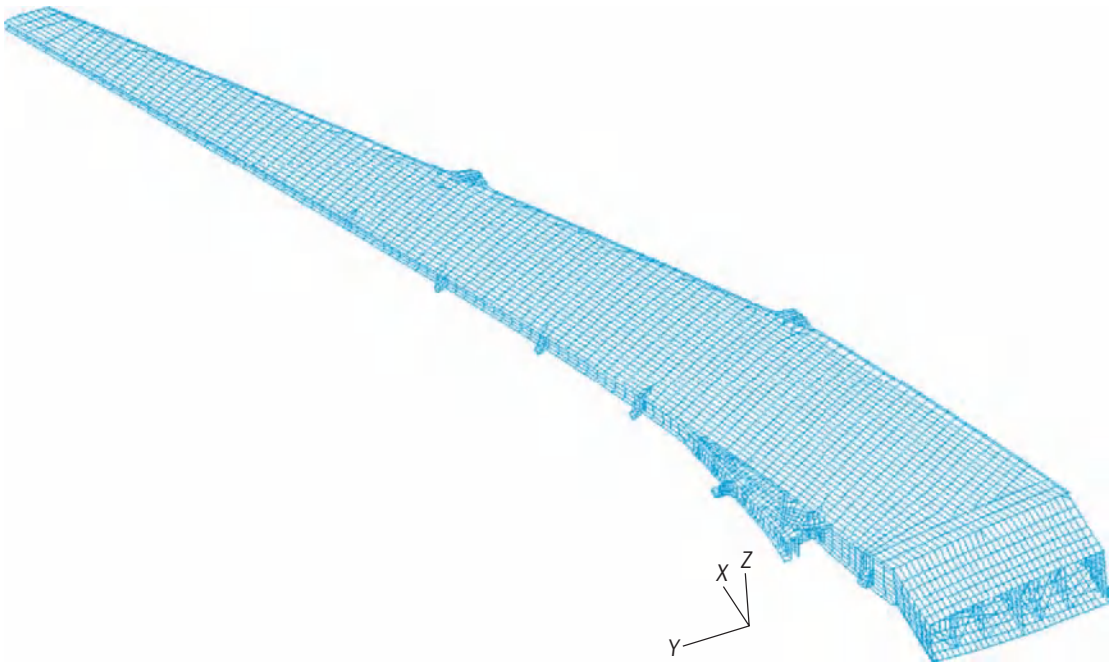

Fig. 7. Finite element model of a wing box. (*Courtesy of Airbus UK*)

computer-aided-design software) has changed this forever. The finite element method is conceptually simple. A typical displacement field across a wing from leading to trailing edge is illustrated in **Fig. 6**. The structure is then divided up into finite elements which are so small that a reasonable approximation can be made for the shape of the variation across any element. In Fig. 6 a linear variation is implied. Then, all that remains is the magnitude of the displacements at, say, the edges of the elements. The strains can then be expressed in terms of these magnitudes, the strain/stress law implemented, and finally the equilibrium equations. The result is a large number of equations giving the many displacements in terms of the applied forces. The number of displacement unknowns and equations can run into hundreds of thousands, but modern computers are powerful and inexpensive enough to cope with problems of this magnitude. A typical finite element display is shown in **Fig. 7**. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING; FINITE ELEMENT METHOD.

Commercial computer codes have been available since the 1970s, but only recently has it been feasible to simulate the failure process, which is nonlinear as damage propagates and which modern codes can simulate. Virtual testing is now a reality and more informative than field testing since the internal stress fields are modeled in detail—for example, when a bird strikes an aircraft or is sucked into an engine fan. *See* AIRCRAFT TESTING. Glyn A. O. Davies

Bibliography. H. D. Curtis, *Fundamentals of Aircraft Structural Analysis*, Times-Mirror HE Group, Los Angeles, 1997; M. Davies (ed.), *The Standard Handbook for Aeronautical and Astronautical Engineers*, McGraw-Hill, New York, 2003; T. H. Megson, *Aircraft Structures for Engineering Students*, 3d ed., Edward Arnold, London, 1999.

Airglow

Visible, infrared, and ultraviolet emissions from the atoms and molecules in the atmosphere above 30 km (20 mi), generally in layers, and mostly between 70 and 300 km (45 and 200 mi).

The airglow, together with the ionosphere, is found in the uppermost parts of the atmosphere that absorb the incoming energetic radiations from the Sun. While the airglow consists of spectral features similar to those of the aurora, it is mostly uniform over the sky; and it is caused by the absorption of solar ultraviolet and x-radiations, rather than energetic particles. *See* AURORA.

Dayglow. The daytime airglow (dayglow) is caused mainly by fluorescence processes as molecules and atoms are photodissociated and photoionized. The photoelectrons that are produced in the ionization processes are a further source of airglow in their collisions with other atoms and molecules. Emissions produced by recombination of previously produced neutral species and ions are a minor source of dayglow, and resonant and fluorescent scattering of sunlight also contributes. *See* FLUORESCENCE.

Only during twilight is the blue sky glow that is caused by sunlight scattered on the lower atmosphere sufficiently weak that some dayglow emissions can be observed from the ground. Thus, twilight offers an opportunity to observe resonant scattering of sunlight on layers such as those of the alkali atoms sodium, lithium, and potassium. As the Earth's shadow scans through the layers, the changes of intensity allow their heights (near 90 km or 55 mi) to be measured. The atoms can be observed by lidar (laser radar) with similar results but more detail. *See* ALKALI EMISSIONS; LIDAR; SCATTERING OF ELECTROMAGNETIC RADIATION.

Nightglow. The nighttime airglow (nightglow) is predominantly due to recombination emissions. The ionospheric plasma recombines near the bottom of the F region (150–200 km or 90–120 mi) where the densities and thus collision frequencies are higher, producing bright atomic oxygen (O) spectral lines in the red (at 630 and 636 nanometers) and a weaker green (558-nm) line. The recombination of the neutral radicals that are generated in daytime photodissociation proceeds fastest in the mesosphere, where the densities are higher at the bottom of the region of production.

Atomic oxygen recombination forming excited dioxygen (O_2) molecules leads to several characteristic spectral-band systems. The recombination energy can alternatively be transferred to an oxygen atom in collision and can become the source of a strong green (558-nm) emission. Excited oxygen atoms that could produce the red (630- and 636-nm) lines in the mesosphere are nearly always rapidly quenched in another collision. *See* ATOMIC STRUCTURE AND SPECTRA.

Strong infrared hydroxyl (OH) emission arises in the mesosphere from the reaction of atomic hydrogen (H) with ozone (O_3), and the H is regenerated by the OH reacting with O to produce O_2, with the net result being the recombination of O and O_3 and the recycling of the H. Similar cyclic reactions excite sodium atoms. A weak continuum in the green is due to the recombination of nitric oxide (NO). The total of all the visible nightglow emissions, together with zodiacal light and scattered starlight, can be seen with the naked eye as the faint light between stars. From space, the edge-on view of airglow is of a band of light above the horizon. *See* ATMOSPHERIC OZONE; ZODIACAL LIGHT.

Observations. Observations made by spacecraft are free of the constraint of the blue sky glow, and can also detect the ultraviolet emissions that cannot be seen from the ground. Such observations of resonant ultraviolet scattering of oxygen ions (O^+) and helium ions (He^+) can be used to map the distribution of ions in the plasmasphere to distances to four times the Earth radius. Scattering in the Lyman series maps the so-called geocorona of atomic hydrogen out to 15 times the Earth radius.

A variety of remote-sensing techniques from the ground and from space utilize airglow to determine the composition and structure not only of the Earth's upper atmosphere but also of the atmospheres of other planets. Spectroscopic observations of relative intensities and Doppler profiles identify the composition, temperature, winds, and energy inputs, as well as ionospheric layer heights and ion concentrations. Images of extended areas of terrestrial airglow show wave structures caused by atmospheric gravity waves. These waves originate deep in planetary atmospheres, and their amplitude grows exponentially with height; their dissipation is an important source of heating of upper atmospheres.

Images made of the ionospheric recombination emissions, looking toward the magnetic equator from low-latitudes sites on Earth, show large-scale structures due to plasma instabilities.

The airglows from other planets are remarkably different from that of Earth. The airglows from Mars and Venus show strong emissions of carbon-oxygen species (CO, CO^+, CO_2^+), O, and H; from Titan emissions of molecular nitrogen (N_2) and atomic nitrogen (N); and from Jupiter and Saturn, emissions of molecular hydrogen (H_2) and H. These emissions confirm that CO_2, N_2, and H_2 are the dominant constituents in each case. Brian A. Tinsley

Bibliography. J. W. Chamberlain, *Physics of the Aurora and Airglow*, 1961; J. W. Chamberlain and D. M. Hunten, *Theory of Planetary Atmospheres*, 2d ed., 1987; G. Paschmann et al., *Auroral Plasma Physics*, 2003; M. H. Rees et al., *Physics and Chemistry of the Upper Atmosphere*, 1989.

Airplane

A heavier-than-air vehicle designed to use the pressures created by its motion through the air to lift and transport useful loads. Although airplanes exist in many forms adapted for diverse purposes, they all employ power to overcome the aerodynamic resistance, termed drag, thereby achieving forward motion through the air. The air flowing over specially designed wing surfaces produces pressure patterns which are dependent upon the shape of the surface, angle at which the air approaches the wing, physical properties of the air, and velocity. These pressure patterns acting over the wing surface produce the lift force necessary for flight. *See* AIRFOIL.

To achieve practical, controllable flight, an airplane must consist of a source of thrust for propulsion, a geometric arrangement to produce lift, and a control system capable of maneuvering the vehicle within prescribed limits. Further, to be satisfactory, the vehicle should display stable characteristics, so that if it is disturbed from an equilibrium condition, forces and moments are created which return it to its original condition without necessitating corrective action on the part of the pilot. Efficient design will minimize the aerodynamic drag, thereby reducing the propulsive thrust required for a given flight condition, and will maximize the lifting capability per pound of airframe and engine weight, thereby increasing the useful, or transportable, load.

Air propulsion systems. Because the airplane moves through a fluid medium, its propulsion system must produce a change in the momentum of its working fluid to generate the thrust force required to overcome aerodynamic drag. Reciprocating and turboprop engines produce power by converting chemical energy of fuel into rotation of a shaft; a propeller fitted to this shaft accelerates the air passing through its rotational area, thereby changing engine power into useful thrust. Turbojet and ramjet engines produce thrust directly by adding heat to the air passing through the engines and ejecting the resulting hot exhaust gases at high velocity. A fanjet (turbofan) engine is a hybrid between the turboprop and the pure turbojet that combines a number of features common to both. The turbojet, ramjet, and fanjet en-

gines use the surrounding air both as primary working fluid and as oxidizer to support combustion. *See* JET PROPULSION; PROPELLER (AIRCRAFT); RAMJET; RECIPROCATING AIRCRAFT ENGINE; TURBOFAN; TURBOJET; TURBOPROP.

Thrust. The engine-propeller combinations, whether reciprocating or turboprop, are effective at low forward speeds, but their thrust falls off rapidly as flight speeds approach the speed of sound (Mach 1), largely because of the tremendous increase in power required to turn the blades when shock waves form at their tips. The turbojet thrust increases with speed because of aerodynamic compression, which becomes more marked at higher speeds. The fanjet, by combining many of the characteristics of the propeller and the turbojet, provides efficient propulsion in the range of speeds between those where the thrust of the propeller falls off rapidly owing to the compressibility drag rise at the tips, and the thrust of the turbojet picks up. The limit of turbojet operation is set largely by the temperature limitations of the materials used to fabricate the turbine and by shock patterns which may produce unsteady flow and separations at high speeds, interfering with efficient aerodynamic compression, creating difficulties with the intake system, and causing the engine to stall and flame out. *See* SUPERSONIC DIFFUSER.

The ramjet cannot produce thrust at low forward speeds because it depends upon an appreciable aerodynamic (ram) compression. It therefore requires a booster system to raise the airplane speed sufficiently that useful thrust can be produced. Because of improved aerodynamic compression, as with the turbojet, the thrust produced increases as the velocity of flight increases. Since a ramjet is designed for maximum performance at a given flight condition, its intake design may be such that its range of operating speeds is small.

Weight and fuel consumption. Ramjet engines, although light, have large fuel consumptions, whereas the heavy reciprocating engine has the lowest consumption, and turboprop, fanjet, and turbojet engines are intermediate in weight and fuel consumption (see **table**). Requirements of speed, range, size, and cost must be balanced to select the engine best suited for the specific airplane design. *See* AIRCRAFT ENGINE PERFORMANCE; SPECIFIC FUEL CONSUMPTION.

Aerodynamic resistance. Aerodynamic resistance is composed of four parts: (1) skin friction drag, which arises from the viscosity of the air; (2) pressure drag, which arises from the pressure field about the body; (3) wave drag, which arises from the compressible nature of air, and which occurs as flight speeds approach and exceed the speed of sound; and (4) induced drag, which is a resistance produced by the generation of lift.

Skin friction. The magnitude of skin friction drag depends upon the surface area of the airplane and the nature of the flow layer that is in immediate contact with it. If the boundary-layer flow is smooth and steady (laminar), the skin friction generated is much less than when this layer becomes turbulent. As a laminar boundary layer proceeds along a surface, depending upon the smoothness and shape of the surface, as well as the density, viscosity, and velocity of the air, the fluid momentum of the flow decreases. Simultaneously, disturbances in the flow or the surface produce motions perpendicular to the surface; these motions tend to grow as the fluid momentum decreases. Eventually a transition condition is reached at which these motions cause the laminar boundary layer to change to turbulent flow. Hence, unless special devices are employed, such as suction slots to remove the low momentum boundary-layer air before it becomes turbulent, the laminar flow will break down after it has passed over a sufficient length of surface. For this reason, on those portions of the airplane where flow initiates, such as the nose and wing leading edges, extreme smoothness to minimize viscous losses is desirable, while further aft, where the flow is turbulent, greater surface roughness can be tolerated. *See* BOUNDARY-LAYER FLOW.

Pressure drag. If sufficient momentum is lost in the boundary layer, the main flow breaks away from the surface, forming a region of separation. Separation results in considerably lower surface pressures than would result in the ideal streamline case. As a consequence, suction pressure on the downstream portions of the body produces pressure drag. This drag is minimized by shaping the body geometrically to reduce the separation region to a minimum by giving it a streamlined shape (**Fig. 1**). In modern high-speed aircraft, pressure drag has been reduced to a small percentage of the total drag. *See* STREAMLINING.

Wave drag. As flight speed of the airplane increases, there comes a point where the flow passing over

Fuel consumption and weight of principal types of airplane engine

Engine	Fuel consumption, lbm fuel/(h)(lbf thrust)*	Weight, lbm engine/lbf thrust†
Ramjet	1.7–2.6	0.4
Turbojet	0.75–1.1‡	0.8§
Fanjet (turbofan)	0.3–0.5	0.7§
Turboprop	0.4	1.2
Reciprocating	0.2	2.4

*Minimum values. Value for ramjet is for Mach 2; values for turbojet and fanjet engines are based on static ground tests; values for reciprocating and turboprop engines are for speeds below Mach 0.3. 1 lbm/(h)(lbf) = 0.1 kg/(h)(N).
†Based on total installed nacelle and power-plant weight, and on thurst at 375 mi/h (168 m/s) at sea level. 1 lbm/lbf = 0.1 kg/N.
‡Value for Olympus 593 engine aboard Concorde is 1.19 at Mach 2.
§Approximate cruise value.

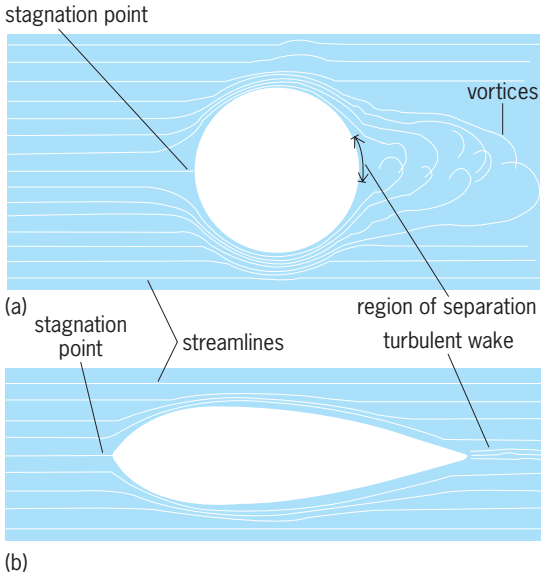

Fig. 1. Effect of streamlining. (*a*) Body without streamlining. (*b*) Body with streamlining, minimizing the region of separation.

a surface reaches the local speed of sound. At this speed (Mach 1) pressure waves propagate through the air. This wave propagation forms shock patterns that interact with the boundary layer to cause premature separation and a sudden drag rise. Because, to create low pressures necessary to generate lift, local velocities next to the upper surface of the wing can be appreciably higher than the flight speed, this separation usually takes place over the wing. Separation over a wing causes an abrupt increase in drag and loss of lift. As the speed continues to increase, shock patterns become steady, the bow and stern shocks changing location and becoming much stronger until they become analogous to the surface waves that are created by a ship. As with such waves, at supersonic speeds the generation of these shock waves represents the major contribution to the drag of the vehicle. *See* AERODYNAMIC WAVE DRAG.

Early experimenters discovered that it was possible to delay and alleviate somewhat the drag rise associated with the shock-induced separations that occur when local sonic speed is first reached at some point on the wing surface. They swept back the leading edge of the wing with respect to the flow direction. The velocity past the wing can be considered as comprising a component in the spanwise direction and a component normal to the span (**Fig. 2**). Only this latter component affects lift and is subjected to local accelerations produced by the airfoil shape. As a consequence, the free-stream velocity can be considerably increased over the value that would produce force divergence for a straight wing. When the flight Mach number becomes sufficiently large, the effect of sweep tends to vanish.

Other attempts to delay the onset and lessen the magnitude of shock-induced separations have concentrated on airfoil design. If the region in which the air passing over the upper surface exceeds the speed of sound can be distributed over more of the surface at a given flight speed, the magnitude of both the shock and the resulting separation can be reduced. So-called supercritical airfoils offer the possibility either of increasing the Mach number (speed of flight/speed of sound) by the order of 15% or, holding the Mach number constant, of increasing the thickness of the wing by as much as 40%. *See* SUPERCRITICAL WING.

Area rule. Studies, at first theoretical in nature but later confirmed by experiment, demonstrated that the drag rise associated with the initial formation and subsequent growth of the shock wave pattern about an aircraft traveling at transonic and supersonic speeds was a function of the rate of change of aircraft cross-sectional area. Furthermore, these studies showed that, if there were an abrupt change in the cross-sectional area perpendicular to the line of flight, the drag rise at a given speed would be greater than if the change were tailored to be more gradual. For this reason aircraft are now designed with the fuselage shape, as well as engine nacelle location and shape, so arranged to produce optimum variation of cross-sectional area for the given design speed; the design principle is termed the area rule. (**Fig. 3**). *See* TRANSONIC FLIGHT.

Induced drag. The final major type of aerodynamic resistance that must be considered arises from the fact that in creating lift the wing must impart a downward momentum to the air flowing past. This downward component of local velocity, termed downwash, rotates the resultant force of the wing in such a manner that a force component is directed in the downstream direction. This component, which depends directly upon the magnitude of the lift force, is termed the induced drag.

Lift of an airplane wing depends upon its shape, speed of flight, angle the wing makes with the free

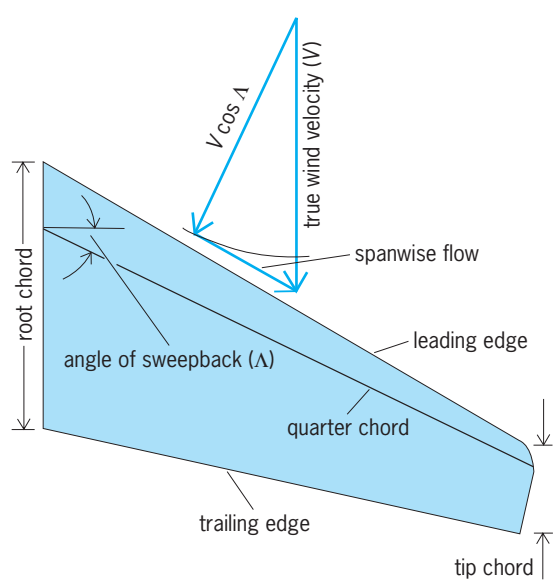

Fig. 2. Geometry of an airplane wing. Wind velocity component perpendicular to the wing span ($V \cos \Lambda$, where V is the true wind velocity and Λ is the angle of sweepback) produces lift.

stream, and properties of the air. These factors are all grouped in the nondimensional coefficient of lift C_L, which is related to the lift L by Eq. (1), where ρ and

$$L = {}^1\!/_2 \rho V^2 S C_L \qquad (1)$$

V are the density and speed of the air, and S is the wing planform area. The coefficient of lift is generally

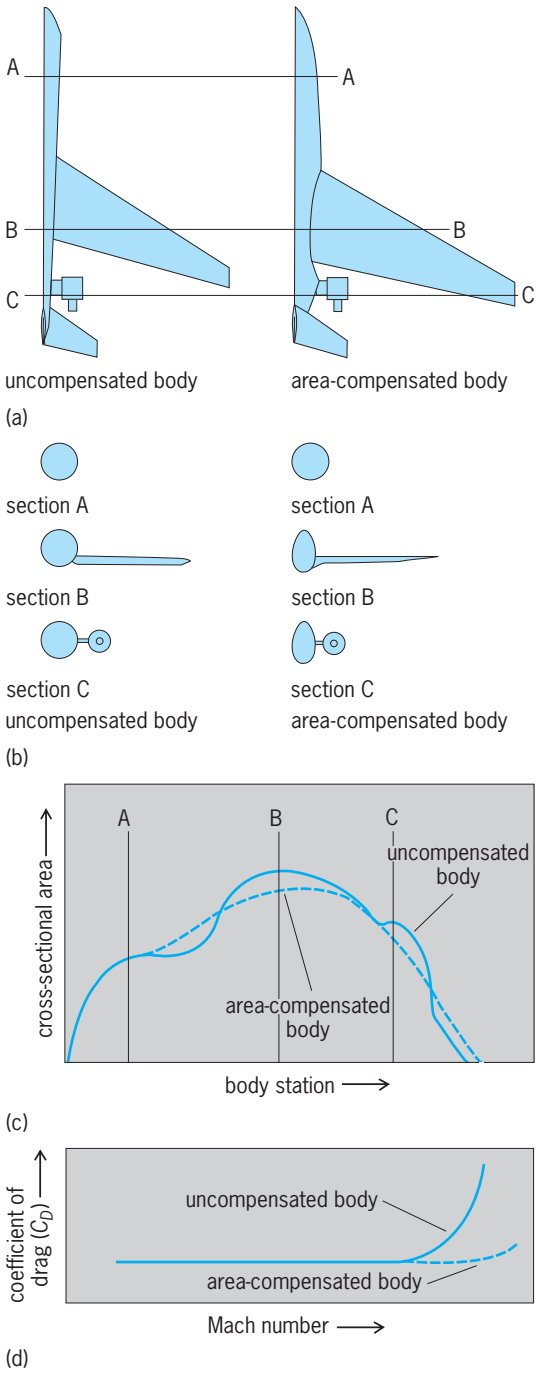

Fig. 3. Illustration of the area rule. (*a*) Half-plan views of airplanes with uncompensated body and area-compensated body. (*b*) Cross sections of these airplanes. (*c*) Variation of cross-sectional areas of these airplanes with body station (longitudinal distance along aircraft). (*d*) Resulting variation of drag coefficients C_D of these airplanes with Mach number.

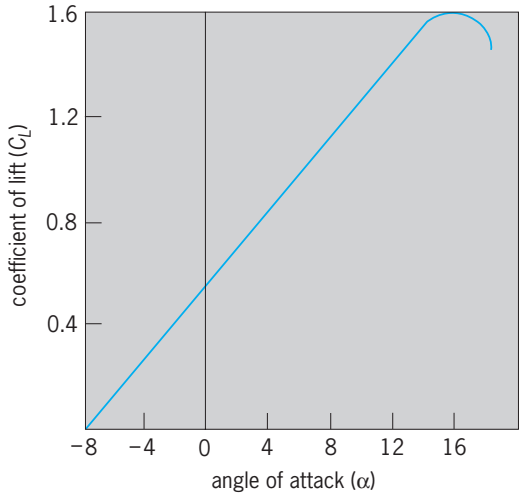

Fig. 4. Variation of coefficient of lift with the angle of attack, showing increase up to the stall angle.

presented as a function of the angle of attack (**Fig. 4**). To maintain level flight, the wing must produce a lift force equal to the weight of the airplane. For a given altitude this implies a low C_L at high speeds but, as the speed is decreased, C_L must increase if level flight is to be maintained. The pilot actually controls the speed of the airplane by changing the angle of attack.

There is a maximum angle of attack beyond which no lift increase is experienced (Fig. 5). As this angle is approached, the air in the boundary layer next to the wing surface experiences such energy losses that it can no longer flow over the wing, but separates, resulting in a loss of suction pressures and a decrease in lift. When there is no further increase in lift as the angle of attack increases, the wing is said to be stalled.

Because induced drag depends upon the magnitude of the lift coefficient, it is of concern primarily in those flight regimes where a large lift coefficient is encountered: relatively low speed or high altitudes. The magnitude of the angle through which the resultant force of the wing is rotated is a function of the aspect ratio, the ratio of the span to the chord of the wing. The higher this ratio, the smaller the angle. Hence airplanes designed for flight at high altitude or for cruising at relatively low speeds will have long, slender wings. Planes designed for higher speeds will be less affected by the induced drag and will generally have much smaller spans to minimize the other forms of resistance.

A measure of efficiency of the airframe is the ratio of lift to drag that it can produce. To achieve very low speeds, special devices such as flaps or slots are used. Such devices, as well as high-lift boundary-layer control, where separation of the boundary layer is delayed either by sucking the low-momentum air next to the surface away through slots or holes in the surface or by adding momentum to the layer by blowing into it through slots, increase the lift that can be achieved before the wing stalls. These devices are thus useful for landing and takeoff maneuvers. The increase of lift achieved by their deployment is

usually accompanied by an appreciable increase of drag and hence does not increase airframe efficiency. *See* AERODYNAMIC FORCE; AERODYNAMICS.

Range. The distance that an airplane can travel with a given amount of fuel is its range, expressed by Eq. (2), where C_t is specific fuel consumption in

$$\text{Range} = \int \frac{V}{C_t}\frac{L}{D}\frac{dw}{w} \qquad (2)$$

pounds of fuel consumed per pound of thrust per hour, V is flight velocity, L/D is lift to drag ratio, and dw/w is a measure of the rate at which total airplane weight changes during the flight.

If the total difference in weight during the flight is produced by fuel consumption, the term $(V/C_t)(L/D)$ should be maximized. In general, it is desirable to have low specific fuel consumption, high speed, and high L/D. Unfortunately, these factors do not maximize at the same point. With reciprocating engine-propeller combination aircraft, maximum range could be achieved at the speed for maximum L/D. With the advent of other forms of propulsion, however, the speed for maximum range has increased. With some aircraft it is possible to achieve high range even at supersonic speeds, for although L/D is reduced, the increase in flight velocity tends to compensate.

Airplanes traveling at supersonic speeds create shock waves which extend outward like the wake of a ship. As the speed and size of aircraft increase, the impact of the shock waves becomes increasingly noticeable and possibly destructive in the form of a sonic boom. Commercial airplanes designed to fly at supersonic speeds must, therefore, operate to reduce shock waves when flying over populated areas. The airplanes must fly either at extremely high altitudes or at speeds below the speed of sound. A significant amount of flight time of supersonic aircraft is spent flying at speeds below Mach 1. The L/D of typical optimum supersonic configurations with highly swept wings and low aspect ratio is low at subsonic speeds. The result of the combination of low L/D and low speed is a substantial reduction in range. To overcome this disadvantage, airplane configurations capable of changing the sweep of the wings in flight have been developed. The variable-geometry compromise provides both high-aspect-ratio, nearly straight wings at low speed and low-aspect-ratio, highly swept wings at supersonic speeds. *See* SONIC BOOM; SUPERSONIC FLIGHT.

Stability. A necessary element of a successful airplane is stability, which is the tendency of forces and moments to be set up that restore the vehicle to its equilibrium position if it is disturbed from this position. For an airplane to be in equilibrium, the sum of the forces acting at its center of gravity must equal zero, as must the sum of the moments. *See* FLIGHT CHARACTERISTICS.

To examine the condition of stability, suppose that the equilibrium of the airplane is suddenly disturbed in a manner that causes it to accelerate. This means that its lift coefficient decreases (**Fig. 5**). A stable airplane would immediately be subjected to a nose-up moment which would tend to increase the lift coefficient and reduce the speed. Similarly, a reduction in the speed of a stable airplane should produce a nose-down moment.

Stability is achieved by balancing the force and moment contributions of various portions of the airplane so that the summation is stable. In conventional configurations the major stabilizing contribution arises from the tail, but by careful design tailless or tail-first (Canard) configurations are possible. Because the balance is achieved for moments about the center of gravity, the location of this point is important. Its position depends upon the load condition and can change in flight as fuel is consumed or as disposable load is dropped. The airplane must be designed so that it reacts in a stable manner throughout the entire range of center-of-gravity locations encountered in service.

Controls. An airplane must be controllable. Major aerodynamic control surfaces include the elevators, ailerons, and rudder (Fig. 5*a*). The angle of attack, and consequently the speed in level flight, is controlled by the elevators. Deflecting these surfaces downward increases the lift of the horizontal tail (just as flaps increase the lift of the wing) and hence causes the airplane to nose downward. An upward deflection produces the opposite effect. *See* ELEVATOR (AIRCRAFT).

Lateral control is achieved by differential action of ailerons. One aileron goes down, increasing wing lift on that side, and the other goes up, reducing lift on that side. The effects couple to produce a rolling moment that tilts the wing. Ailerons also serve to turn the airplane by tilting the wing lift and producing a component normal to the flight path tending to curve it (Fig. 5*b*). *See* AILERON.

Occasionally ailerons are replaced by a spoiler, a device which projects upward on the downgoing

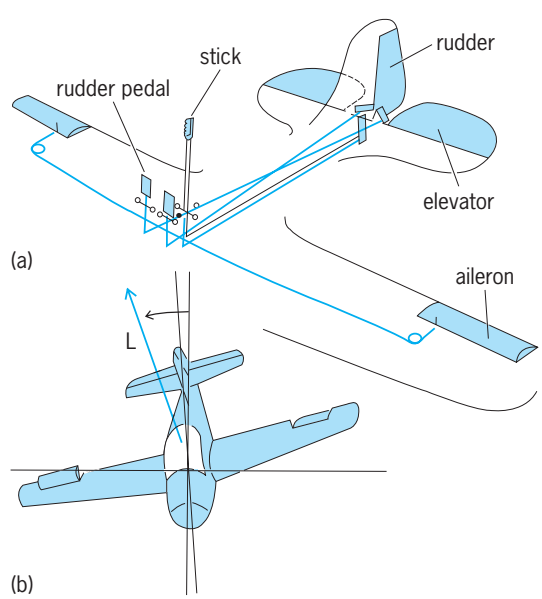

Fig. 5. Aerodynamic control of an airplane. (*a*) Basic controls and control surfaces. (*b*) Effect of tilting the airplane in the direction of the wing lift *L*.

wing, thus producing separation that reduces lift. The spoiler may be used for several reasons. It reduces the twisting moment, which for an aileron can be so great that the torsional resistance of the wing is exceeded, causing it to rotate and thereby changing its local angle of attack. The upgoing aileron, instead of reducing lift, can produce such a change in angle of attack that the control action is reversed. A spoiler is also used to overcome adverse yaw, the tendency of the higher lift on the upgoing wing to produce an increased induced drag on that side. This drag tends to turn the airplane in the wrong direction.

The vertical tail provides directional stability. The rudder prevents the airplane from slewing, or picking up too large an angle of yaw during maneuvers. Although the airplane could be turned with the rudder, such a maneuver would be uncomfortable, and the ailerons are always used. If the vertical tail were large enough to prevent slewing by itself, the rudder could be eliminated and a two-control machine built. *See* AIRCRAFT RUDDER; FLIGHT CONTROLS.

Augmentation systems. While it is both possible and desirable to build aircraft that are inherently stable at relatively low flight speeds, as these speeds increase into the transonic and supersonic ranges the magnitudes and nature of the pressure fields about the aircraft change, resulting in two effects. The first arises from the airloads becoming so great that not only do the forces required to move the controls exceed the pilot's strength but the aircraft, normally thought of as a rigid body, distorts and twists, changing its aerodynamic characteristics (aeroelastic effects). The second effect results from the change in the nature of the distribution of the loads over the airframe and control surfaces (compressibility effects). In combination the two effects alter both the stability and controllability characteristics of the airplane. *See* AEROELASTICITY.

To handle such cases it is often necessary to provide systems that not only supply the necessary power to move the control surfaces but also augment the stability characteristics of the basic aircraft by sensing motions and automatically applying compensating control inputs independent of pilot action. A common application of such a system is to augment yaw damping at high subsonic speeds. In many modern designs, particularly military ones, the pilot is no longer physically connected to the control surfaces. Instead, pilot control motions are interpreted by a computer that superimposes pilot commands on those control motions required to maintain stability. In some cases the airplane has been designed to be unstable if unaugmented in order to achieve high maneuverability. *See* STABILITY AUGMENTATION.

Structures. Airplane structures must provide the required strength and rigidity, within the shape envelope dictated by the aerodynamic requirements of the airplane, for the lowest possible weight. Other factors such as manufacturing ease, corrosion resistance, fatigue life, and cost are important in selecting both the materials and the nature of the structural system to be employed, but the strength-to-weight ratio is of critical importance.

Most materials used in aircraft construction display the characteristic that stress is essentially proportional to strain up to a point called the proportional limit. Removing the load before this point is reached returns the material to its original unstrained shape, while exceeding the proportional limit results in permanent distortion. Airplane structures are designed so that no stress exceeds the proportional limit if the airplane is maneuvered over its full range of speed and design accelerations: the ultimate stress is not exceeded even if the structure is subject to $1^1/_2$ times its design load. *See* AIRFRAME; FUSELAGE.

The major structural problem is design of the wing. Aerodynamic considerations demand that wing thickness be kept small with respect to wing chord. The air load is distributed along the wing, and the magnitude and configuration of the load are a function of the maneuver under consideration. The wing may thus be considered as a long, slender beam subjected to loads that during accelerations may total several times the weight of the aircraft. Because the depth of the beam (airfoil thickness) is small, the outer fiber stresses are large. *See* WING; WING STRUCTURE.

One solution to the problem of high bending stresses is the biplane, where the upper and lower wings, which are connected by struts and wires, act as the flanges of a beam having a depth equal to the separation between the wings. Another solution is the use of a pin connection at the wing root and an external strut.

High-performance airplanes cannot afford the drag penalties of such solutions and must employ cantilever wings. Because of the high bending and torsional stresses encountered, a cantilever structure implies a stressed skin, or outer covering of the wing. The skin may be thin, supported by many stringers, or relatively thick with fewer longitudinal members. Ribs along the span aid in distributing loads and in reducing column action in the stringers.

High flight speeds demand smooth skin surfaces. These are sometimes produced by using skin milled from metal blocks rather than sheet. Two thin sheets bonded to a low-density core form a sandwich, a type of construction frequently used.

Aircraft flying at high speeds encounter high temperatures due to the high aerodynamic compression that occurs at the nose and the leading edges of the wings. At high Mach numbers these temperatures become sufficiently high to reduce the strength of aluminum and its alloys, so other materials, such as stainless steel or titanium, are sometimes used in these areas. When flight speeds become so high that the temperature rise affects the characteristics of the construction material, it may be necessary to employ cooling in sensitive areas. *See* AEROTHERMODYNAMICS.

David C. Hazen

Bibliography. J. D. Anderson, Jr., *Introduction to Flight*, 5th ed., 2005; R. Jackson (ed.), *Jane's All the World's Aircraft*, revised periodically; M. J. H. Taylor and D. Mondey, *Guinness Book of Aircraft*, 6th ed., 1992.

Airport engineering

The planning, design, construction, and operation and maintenance of facilities providing for the landing and takeoff, loading and unloading, servicing, maintenance, and storage of aircraft. Facilities at airports are generally described as either airside, which commences at the secured boundary between terminal and apron and extends to the runway and to facilities beyond, such as navigational or remote air-traffic-control emplacements; or landside, which includes the terminal, cargo-processing, and land-vehicle approach facilities.

The design of airports involves many diverse areas of engineering. Their significant cost has resulted in the development of sophisticated techniques for planning, construction, and operation of airports.

Design considerations. Airport design provides for convenient passenger access, efficient aircraft operations, and conveyance of cargo and support materials. Airports provide facilities for changing transportation modes, such as people transferring from cars and buses to aircraft, cargo transferring from shipping containers to trucks, or regional aircraft supplying passengers and cargo for intercontinental aircraft. In the United States, engineers utilize standards from the Federal Aviation Administration (FAA), Transportation Security Administration (TSA), aircraft performance characteristics, cost benefit analysis, and established building codes to prepare detailed layouts of the essential airport elements. These elements include airport site boundaries, runway layout, terminal-building configuration, support-building locations, roadway and rail access, and supporting utility layouts. Airport engineers constantly evaluate new mechanical and computer technologies that might increase throughput of baggage, cargo, and passengers.

Design choices must enhance user orientation and comfort. Because airports and airlines are key transportation and economic facilities for major cities worldwide, airport engineers frequently integrate a wide variety of secondary design considerations, including air- and water-quality enhancement, noise concerns, citizen participation, wildlife impacts, and esthetic features.

Site selection for new or expanded airports. Site selection factors vary somewhat according to whether (1) an entirely new airport is being constructed or (2) an existing facility is being expanded. Few metropolitan areas have relatively undeveloped acreage within reasonable proximity to the population center to permit development of new airports. For those airports requiring major additional airfield capacity, and hence an entirely new site, the following factors must be evaluated for each alternative site: proximity to existing highways and major utilities; demolition requirements; contamination of air, land, and water; air-traffic constraints such as nearby smaller airports facilities; nearby mountains; numbers of households affected by relocation and noise; political jurisdiction; potential lost mineral or agricultural production; and cost associated with all these factors. Some governments have elected to create sites for new airports using ocean fills. The exact configuration of the artificial island sites is critical due to the high foundation costs, both for the airport proper and for the required connecting roadway and rail bridges.

New airports offer significantly increased runway capacity over airport reconstruction projects due to the larger space available for more runways, longer runways, and more widely spaced runways. New airports are rare: Dallas/Fort Worth International Airport opened in 1972, and Denver International opened in 1995. In 1998, two new airports constructed at Air Force bases opened: the Austin, Texas, airport; and MidAmerica, at Scott Air Force Base in Illinois. Killeen Airport in Texas opened as a joint-use facility with Robert Gray Army airfield in 2004. Denver capitalized on its geographic location in the midsection of continental America, its historic hub airline operations, and undeveloped land adjacent to the metropolitan area to build an unusually large airport (**Fig. 1**). This location is farther from the Rocky Mountains than the old airport, a significant aeronautical advantage. The number of citizens affected by the FAA's regulated noise level was significantly reduced by the new airport's location.

Denver's master plan and environmental approvals were completed between 1986 and 1989. Design, construction, training, testing, and airport certification were completed between 1989 and 1995.

For commercial airport sites being expanded or military bases being revamped for commercial use, many site factors must be considered. The possibility of acquiring adjacent commercial or residential properties requires extensive analysis of social and economic impacts and development costs. Mitigation measures of noise impact constitute a key element of any proposed airport expansion. Such measures can include restrictions on types of aircraft using particular runways, avoidance of noise-sensitive areas by air- traffic routing, acquisition of additional land adjacent to the airport as a noise buffer, improvements to existing occupied structures to dampen noise transmission, or construction of noise berms for ground operations.

Airfield configuration. Since the runways and taxiways constitute the largest portion of the airport's land mass, their layout is generally one of the first steps in the airport planning process. The optimum runway configuration, or layout, is based on long-range forecasts of the estimated numbers of aircraft landings and departures, airport elevation, local prevailing wind direction, and regional and national air-traffic patterns. A paved runway surface 12,000 ft (3660 m) long and 150 ft (45 m) wide is suitable for most applications. However, runway length requirements vary according to the type of aircraft, temperature, and elevation. A parallel taxiway is generally constructed 600 ft (180 m) from the runway (measured centerline to centerline). It is connected by shorter high-speed taxiways to allow arriving aircraft to leave the runway surface quickly so that another aircraft can land. This combination of paved

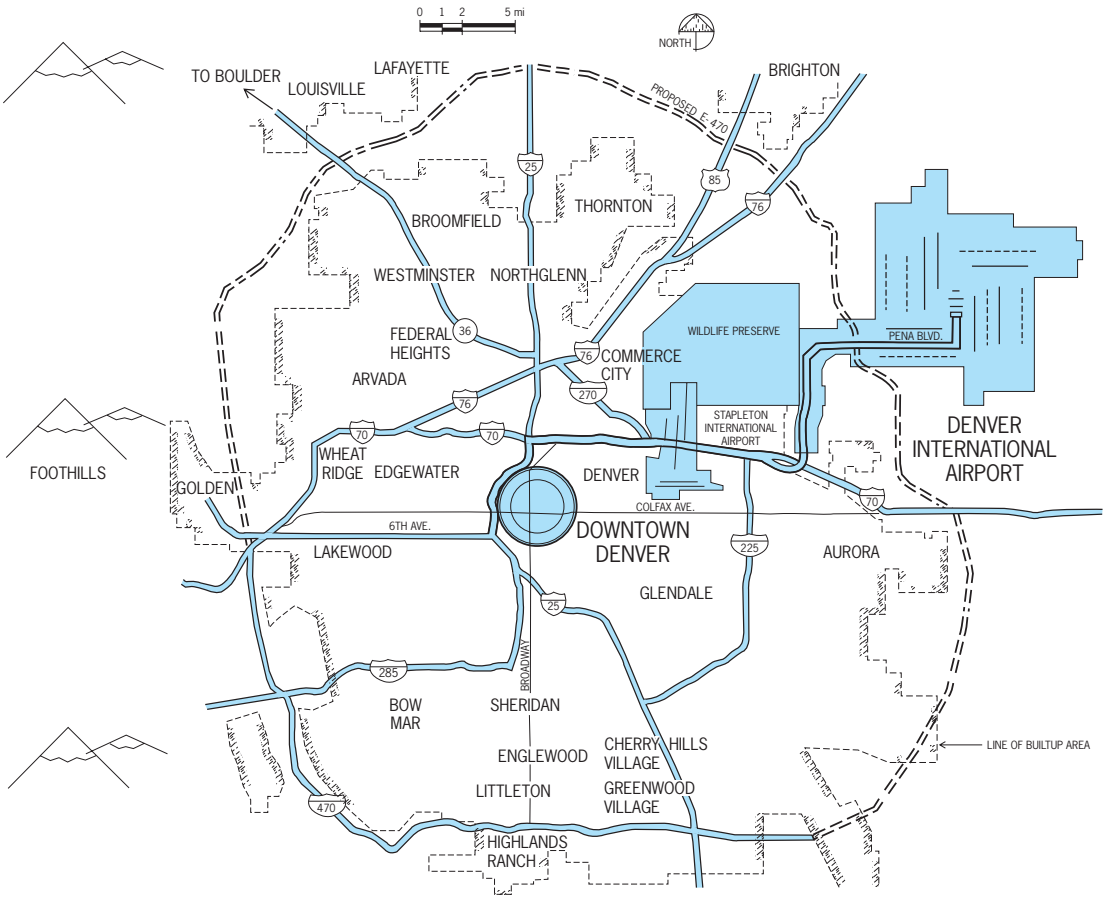

Fig. 1. Regional map of the Denver International Airport.

surfaces is generally referred to as a runway-taxiway complex.

Airports provide runway layouts so that incoming and departing aircraft can maintain required minimum separation for safe, simultaneous aircraft operations. For parallel runways, thresholds would be slightly staggered to avoid wake turbulence interference between incoming aircraft. Staggered thresholds might also be used to minimize crossing of active runways by taxiing aircraft. Each runway crossing is a potential aircraft delay and a safety hazard.

When airports have sufficiently high-velocity crosswinds or tailwinds from more than one direction, crosswind runways are located at some angle to the primary runway as dictated by a wind rose analysis. *See* WIND ROSE.

Alternative airfield configurations are generally analyzed under various operating conditions by using computer simulations. Typically, these simulations assume a certain number of aircraft arrivals and departures, based on actual airline schedules for an entire year, weather conditions based on historical data, and various combinations of runway usage by the air-traffic controllers; and then the required time for each aircraft operation is calculated (**Fig. 2**). The configuration with the least amount of arrival time and taxi time on an annual basis is preferred. Justification for additional runways would accrue as the construction and financing costs of the runway are compared to the operational savings of reduced aircraft delays.

Some runways, widely spaced so that the aircraft operations are independent, are used primarily for instrument flight rules (IFR) conditions, when visibility is low. In visible flights rules (VFR) conditions, FAA rules allow independent operations with more closely spaced runways. The result is shorter taxi distances after the aircraft lands. The IFR runways, which are farther from the terminal, might also be used during peak periods to increase the airport's capacity. If excess runway capacity is available, FAA air-traffic controllers have more discretion to assign aircraft landings and departures to minimize noise exposure and taxi distances.

In the United States, the FAA defines standard techniques for calculating the hourly capacity, annual service volumes, delays, average taxiing distances, and annual runway crossings.

After the runway configuration is determined, detailed construction drawings are prepared that dictate the runway profile and cross section, pavement depth and materials, subsurface remediation, drainage controls, lighting, and paint markings. A cost-effective design balances grading materials from the airport site to minimize trucking costs and the associated impact on air quality and traffic congestion, and specifies locally available materials for the paving mix. *See* AIR POLLUTION; PAVEMENT.

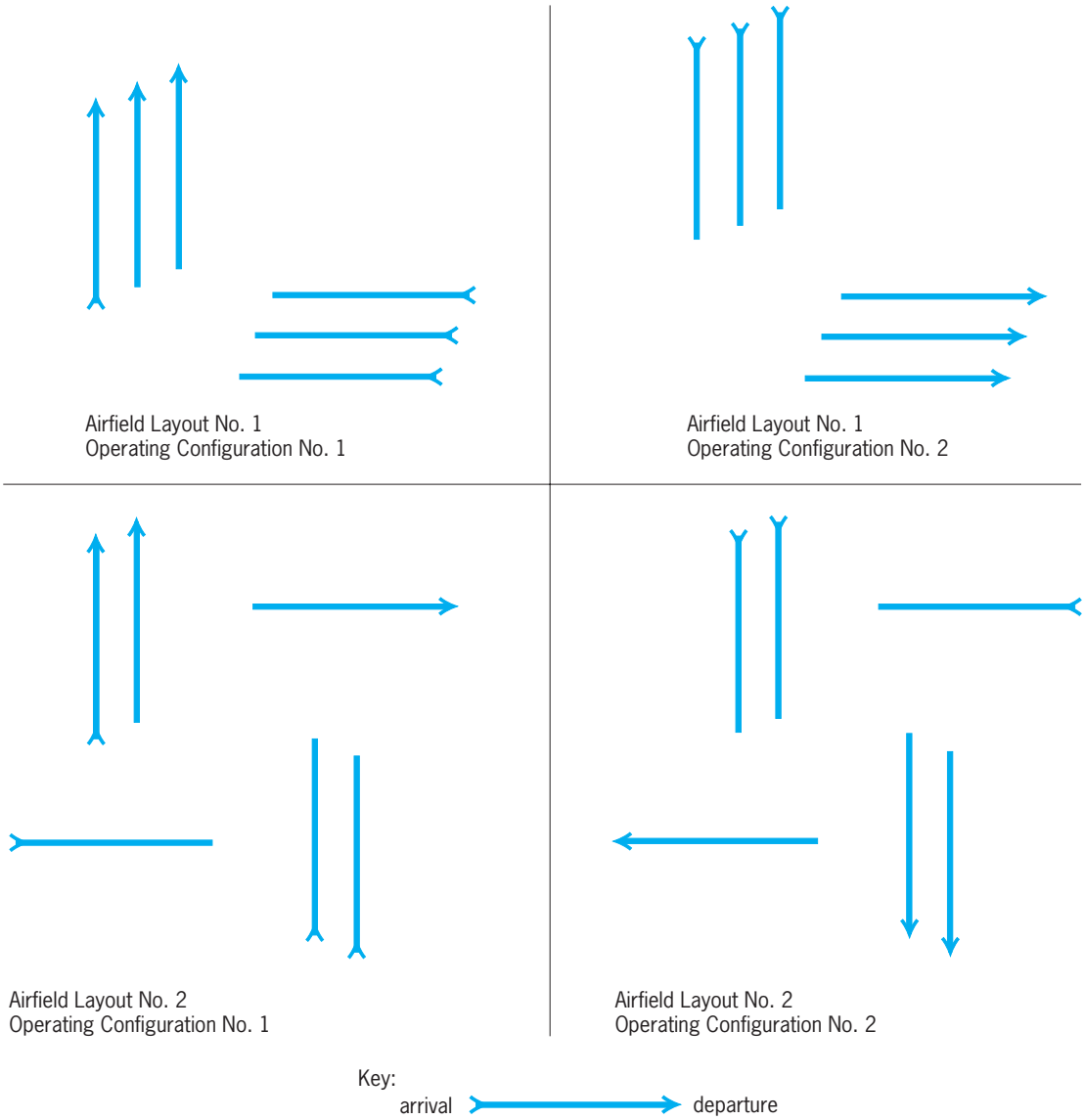

Fig. 2. Combinations of arrivals and departures on the same runway configuration.

Runways are paved with concrete, asphalt, concrete-treated base, or some combination of layers of these materials (**Fig. 3**). Runways for larger aircraft require thicker, more expensive pavement sections. Engineers design these pavements for long-term durability. The expected life of a concrete runway can be increased from 20 to 40 years, based on enhanced mix designs and sections. *See* CONCRETE.

The longitudinal profile of a runway requires an iterative design process. The objective is to minimize grade changes for the smoothest aircraft operations possible, while minimizing the total quantities of earthwork (cut-and-fill) to construct the runway.

The availability of computerized controls and improved lighting fixtures is increasing the use of technology known as surface movement guidance systems for pilots. These advanced airfield lighting systems provide positive control for taxiing aircraft to prevent aircraft collisions. Centerline lights provide positive guidance along predetermined routes for the pilot to follow into gates during low-visibility conditions. Stop bars consisting of rows of red lights prevent planes and vehicles from traveling onto an active runway without positive control from the air-traffic control tower. The controller has positive identification of the aircraft location through the use of ground pressure sensors and airport surface detection radar equipment. Design of these systems must mirror the airport's operating procedures. As an example, the lighting circuit design reflects certain groupings of taxiway segments and runways used for various operating configurations (**Fig. 4**).

A system of vehicle service roads must be provided around the perimeter of the airfield both for access to the runways and for security patrols of the perimeter fencing. Airfield security fencing with a series of access gates is monitored with patrols and, increasingly, a remote camera surveillance system.

Terminal configuration. The terminal building houses the ticketing, baggage claim, and transfer

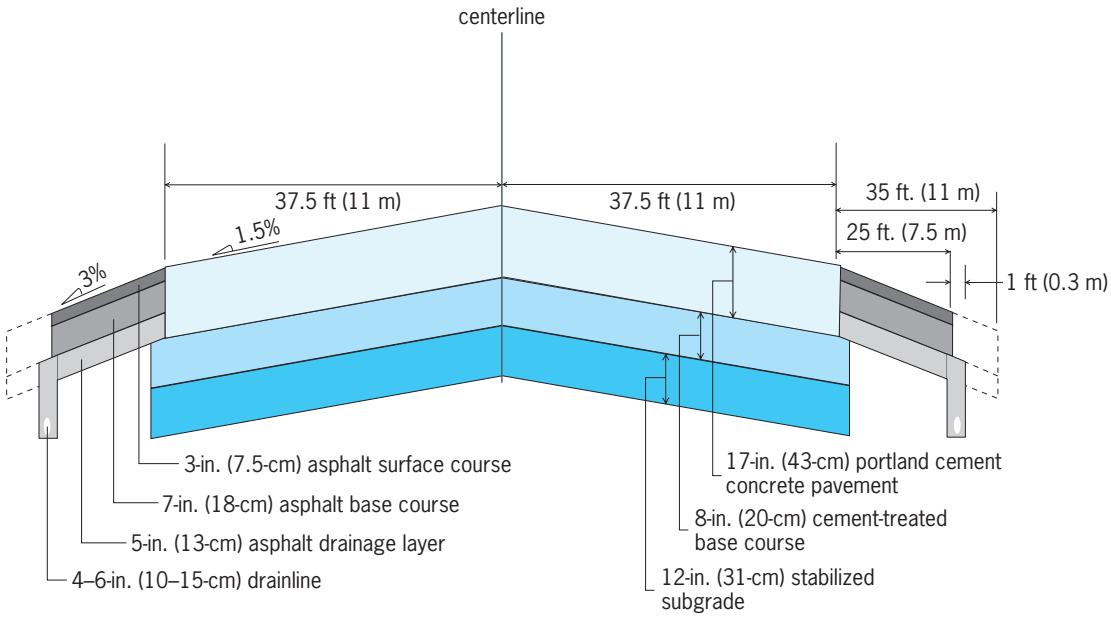

Fig. 3. Cross section of a runway-taxiway pavement.

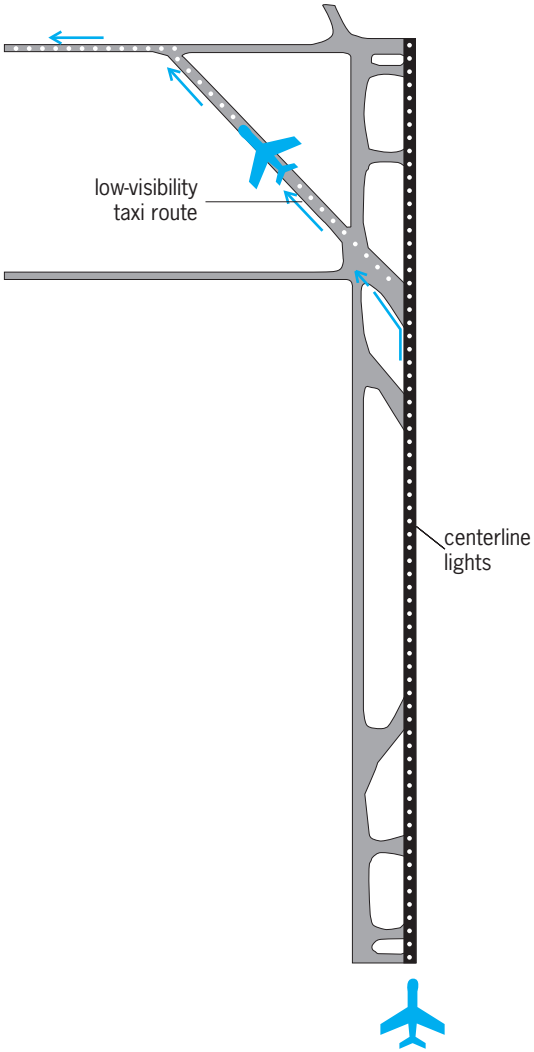

Fig. 4. Surface movement and guidance and control system plan for a runway and the associated low-visibility taxi route.

to ground transportation facilities. The concourse is generally the combination of facilities for passenger-boarding of aircraft, sorting baggage according to flight, and uploading cargo carried in commercial aircraft. The most popular terminals with passengers are those that offer amenities such as a hotel integrated into the terminal main area, extensive retail shops, and vibrant integrated artwork. Airport terminal and concourse configurations generally fall into three categories (**Fig. 5**): (1) a terminal contiguous with concourse satellite extensions (known as piers or fingers) used for boarding aircraft, such as at Miami International Airport; (2) unit terminals, which serve as transfer points both from the ground transportation modes into the building and from the building into the aircraft, such as at New York's John F. Kennedy Airport; (3) and a detached terminal and concourses, sometimes referred to as a landside and airside terminals, connected by a people-mover train system such as at Atlanta's Hartsfield International Airport, or an underground walkway such as at Chicago's O'Hare Airport, or a surface transport vehicle such as at Washington's Dulles Airport.

Contiguous. The contiguous terminal-concourse has a relatively low cost since no trains or special airport vehicles are used. Passengers walk through the complex, aided by moving sidewalks. Expansion of additional aircraft gates is convenient by lengthening one of the piers or adding a pier. However, expansion is fairly limited since continued lengthening of the concourse adds walking distance and can infringe on required airfield operational space. Another limitation is the maneuverability of aircraft between the piers. Aircraft can be delayed while waiting for an aircraft to push back from an adjacent or opposite gate.

Unit. Unit terminals are the most convenient for local passengers, allowing short distances from the

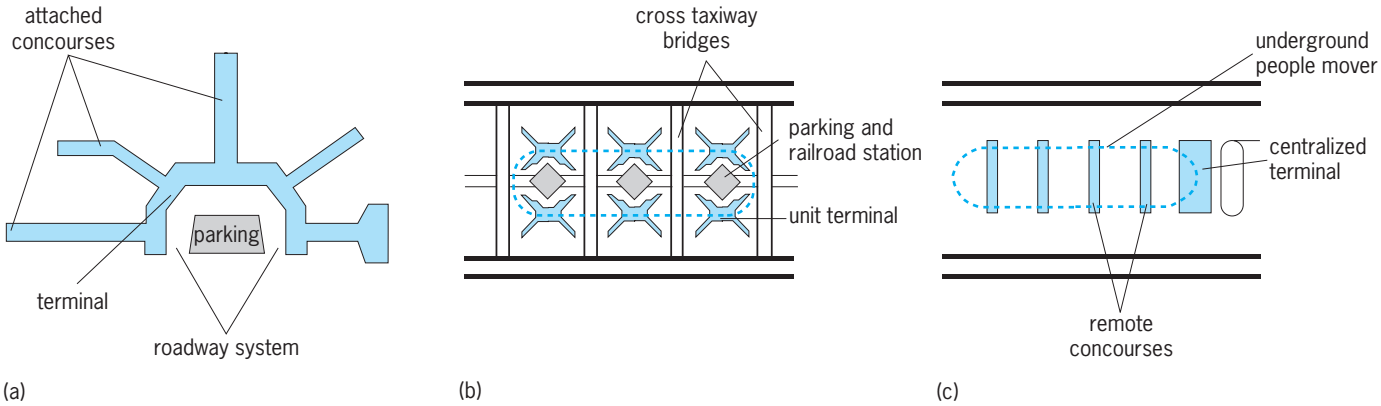

Fig. 5. Alternative terminal configurations. (*a*) Contiguous terminal and concourses. (*b*) Unit terminal. (*c*) Detached terminal and concourses.

curbside roadway system to the aircraft boarding area. This arrangement requires a more extensive investment in the curbside roadway infrastructure. Hubbing passengers, those transferring from one plane to another at the same airport, are disadvantaged in this configuration because they have to move from one building to another.

Detached. Detached terminal and concourse configurations allow airfield facilities to be widely spaced for more efficient aircraft movement, while keeping passenger walk distances short by providing automatic transport systems. Landside and airside building can be expanded independently as different functions require. This configuration allows for a single security screen pavilion at the entrance to the aircraft gate area, which is an advantage for both security and cost.

Shapes. A number of shapes can be used for the airside concourse, depending on the available space. The most common is a long, linear concourse with an underground train station in the middle. This concept was originally developed for Atlanta's Hartsfield International Airport, and it proved to be the most efficient shape for aircraft movement. This linear shape mirrors the parallel pattern of aircraft movement required for highly efficient operations. The hydraulic analogy is laminar flow, as opposed to turbulent flow, which is inefficient. The master plan for Denver International Airport recognized the efficiencies of the Atlanta-style concept and adapted to it. The Denver airport's linear concourses are more widely spaced than the Atlanta concourses, allowing simultaneous taxiing between concourses. Denver's linear concourses were widened to allow wide moving sidewalks. An expanded central core was also added to allow for additional food and retail services for passengers. This master plan concept has now been adopted for the new terminal at Salt Lake International Airport and the new Terminal 5 at London's Heathrow International Airport.

At the Pittsburgh airport, a remote concourse has been constructed in an X shape. The large middle houses the station for the underground people mover. A great advantage is that the average passenger walk distance is halved in comparison to the linear concourse. If each concourse contains 40 gates, a passenger on the linear must walk the length of 20 gates to reach the farthest gate. The farthest a passenger in an X concourse would walk is the length of 10 gates. Gates vary in size according to the type of aircraft scheduled for use. Wider wingspans require wider (longer) gates. Aircraft with higher seat counts must have gates with a commensurate amount of passenger seating.

Smaller facilities. Smaller terminal facilities are required for general aviation, charter companies, and commuter or regional aircraft. These are generally less system-intensive. Selection of locations must take into account both the number of passengers that connect from these terminals to the air carrier concourses and the preferred mode of transport between the terminals.

Dimensional requirements. Airports are among the most specifically designed and configured public facilities. An important part of airport engineering is optimizing the trade-off between providing sufficient dimensional space requirements for such functions as passenger queues, aircraft movement, seating areas, and circulation without requiring overly large spaces that drive up the cost of construction and operation. A detailed dimensional requirements study is prepared that determines the width of ticket lobby, size of train stations, mix of various-sized aircraft gates, width of concourse, number of roadway lanes, number of short-term parking spaces, number and size of food service outlets, linear feet of ticket counter, and space programs for support functions; and that otherwise defines as many elements the airport as possible. Reasonable delay criteria are defined for almost every step of the trip through the airport. These criteria are then translated into speed and size requirements.

Terminal subsystems. Because of the high volume of passengers and baggage in an airport terminal, mechanized subsystems are used extensively. Moving sidewalks shorten walk distances and passenger connect times. Double or extrawide moving sidewalks are desirable, provided the building is sufficiently wide, to allow some passengers to rest while others move briskly. Increasing passenger

preference to carry small items of luggage and shopping parcels has also required engineers to widen moving sidewalks.

Two devices are used to move passengers from one level in an airport building to another. Banks of elevators are required for every level change in order to accommodate all possible passenger needs. Escalators are required in order to move the peak numbers of passengers down from ticketing to a people mover station or from a concourse down to baggage claim level. These two types of devices are complementary and are always used in tandem. Their numbers and dimensions are determined by accessing the passenger demand and required hourly capacity. One level change, for example, might require one elevator, two down escalators, and one up escalator.

Use of electric trains to transport passengers within an airport building or between buildings is the most expensive alternative. Extensive structural requirements and fire code provisions add to the base system cost. Extensive signage and voice messages are required to ensure that passengers feel comfortable using such a system to move from gate to gate or concourse to concourse. This system is, however, the most comfortable for passengers.

Mechanized sortation systems are used extensively at airports for mail, cargo, parts, and baggage. They significantly decrease the time to sort and transport the aircraft load. They generally use tunnels and basements to move these materials into the aircraft. Each system has special requirements (such as maximum and minimum dimensions, heavy structural loading, and maintenance access requirements) that must be considered in the building design.

Security requirements. After the events of September 11, 2001, new security regulations enforced by the Transportation Security Administration have resulted in many changes to airport facilities and operational procedures. They include 100% screening of all checked baggage, matching all checked baggage to passengers, closer inspection of carry-on luggage, and prohibition of certain objects in the aircraft cabin, as well as secure cockpit doors, more rigorous background checks of employees, and more explosives detection. New airport security initiatives focus on gathering and analyzing information about passengers, cargo, vehicles, employees, and even the planes, both on the ground and in the air. An entirely new security-focused information management infrastructure is being created.

Information management/information technology. In recent years, computers and computerized control systems have been developed for virtually every component of airport design and operations. Therefore, it is critical to the success of every airport project that these systems be defined and planned at the earliest stage. Airports have an unusually high number of tenants and users. Each requires a different set of systems for up-to-date information and business controls.

Generally, a fiber-optic cabling system is provided for transmission of data and voice signals by all tenants. This has become an essential utility. For radio transmission, some airports are managing the high usage of radio with another utility system, a leaky coaxial cable, which picks up low-intensity transmissions by individual radios and transmits them for amplification for the end-receiver of the message. This avoids high-intensity transmissions, which cause interference between users.

Airports generally have a set of life-safety systems that include the fire alarm system, public address system, smoke evacuation, security system, and some aspects of transport systems such as underground train systems. Each of these systems must perform its role in communicating to the passengers and airport employees and controlling devices such as exhaust fans, doors, and trains.

Airport operations require a paging system, card access security, controlled-circuit TV security, ground-to-ground radio, ground radar, parking and ground transportation revenue and control system, broadcast TV, and Doppler radar weather system.

Some level of integration of these systems is essential to minimize the time required to communicate critical information. For instance, if a weather detection system senses a tornado or wind shear condition, the integrated control system can preprogram audible alarms.

Airports also have a set of maintenance systems that are best managed by an integrated control system. These systems manage devices such as airfield in-pavement lights, pavement temperature sensors, moving sidewalk controls and alarms, irrigation controls, fuel system monitoring data, environmental monitoring data, work orders, and spare parts inventory.

For engineering purposes, airports may use a computer-assisted drafting (CAD) database of all airports facilities. Information is generally stored in layers, with all structural and demising walls in one layer, utilities and electrical wiring in another, door hardware in another, and lease information in yet another.

Airlines are increasingly providing their own radar to gain up-to-date information on incoming aircraft locations for better management of gates. They use both air-to-air and air-to-ground radio systems. Each airline has a master computer system that controls reservations, ticketing, and flight management. This master computer system may control other computerized systems, such as flight information displays or baggage systems.

Some changes are occurring in the areas of responsibility for these various systems. Some airports are installing Common User Terminal Equipment, which provides complete passenger data-handling infrastructure to different airlines. Airports are also increasingly installing their own Flight Information Display Systems, so that multiple carriers' flights can be displayed on the same devices. Another change is that systems historically installed and operated by the FAA may transfer to airports. One example is approach lightning systems, which indicate the distance to the runway threshold to pilots on approach. As these systems are turned over to airports,

another integration opportunity is presented. Integration controls for both approach lights and inpavement lights would decrease the opportunity for delay or error and allow the controller to focus on communication with pilots.

A gradual change from radio technology to use of digital transmissions will affect many airport systems. Receipt of a radio transmission to a pilot cannot be verified by the air-traffic controller. However, receipt of a digitized message can be verified and can be printed in the cockpit. As older aircraft are replaced, new aircraft are equipped to receive such messages.

Industry groups encourage cooperation between airlines and airport authorities. They are also encouraging development of system standards and interfaces between airline and airport authority computer systems. Recent technology developments include machine-readable boarding passes and tags, which allow better security and faster passenger processing; self-service passenger check-in kiosks; and software that allows printing of boarding passes at home. These innovations decrease passenger waiting times on lines and reduce airline personnel costs.

Movement of cargo requires communication between different agencies such as airlines, customs agents, cargo shippers, cargo receivers, and trucking companies. Guidelines for standardizing these types of system are also being recommended by the industry groups.

Ground transportation facilities. Airline passengers use terminal and parking facilities to transfer to private cars, commercial vans and buses, taxis, rental cars, or rail transport. Large numbers of these vehicles must be able to park, process business transactions, and load passengers. Demand for short walk distances requires that many of these functions be incorporated into the front area of the terminal building.

Remote staging facilities are often constructed so that large commercial vehicles do not dwell in front of the terminal. Commercial vehicles are mobilized to the curbside by radio dispatch, loaded with passengers, and moved away from the terminal as quickly as possible, thereby maximizing the number of vehicles that can load in a peak hour, and minimizing the exposure of waiting passengers to vehicle emissions.

Rail connections are best provided by having a station incorporated directly into the terminal or conveniently accessible from the terminal. Stations providing airline passengers access to regional transit are becoming more common. Airports with such facilities include New York JFK, Atlanta Hartsfield, Washington Reagan National, St. Louis Lambert, Portland (Oregon), and Baltimore-Washington. Several airports are now planning such links, including Washington Dulles, Los Angeles, and Phoenix Sky Harbor.

Support buildings. The primary types of support buildings required by the airlines for their airport operations are flight kitchens to prepare meals for passenger, hangars to service aircraft, and ground support equipment buildings to service vehicles such as tugs, baggage carts, and service trucks. The high number of trips for support vehicles to travel from these buildings to load or service aircraft requires that the buildings be located in reasonable proximity to the aircraft gates. However, the buildings should be sufficiently distant to allow the concourses to be expanded without requiring demolition of these support facilities. Careful forecasting of growth, analysis of operational costs, and dimensional analysis are required to obtain the optimum location.

Airline companies may also require parts storage buildings, employee facilities, and administrative offices. Airlines frequently have an exclusive cargo building to support delivery and sortation of cargo carried in excess space on commercial flights.

An airport requires fire equipment to provide extremely fast primary and secondary responses to every runway. Locating the aircraft rescue and firefighting stations to meet these responses times stipulated by the FAA requires careful positioning with respect to the taxiway systems. The exact building layout must recognize the need for the fire equipment doors to exit onto the most direct response route. *See* FIRE TECHNOLOGY.

Other types of support buildings include vehicle maintenance buildings for snow removal and airport vehicles, sand storage buildings, roadway revenue plaza offices, and training facilities.

Fuel and deicing facilities. Economics of scale and safety considerations generally encourage the implementation of large, centralized common systems for aircraft fuel. The large storage tanks required to ensure adequate reserves of fuel are located in remote areas of the airport, generally in aboveground facilities. Underground distribution piping transports the fuel to hydrant pits or truck fueling stations close to aircraft operations. This system, like most utilities, is designed with backup capacity by looping piping around each service area. If a problem occurs in one section of pipe, valves are automatically closed and the supply direction is reversed. Fuel tanks require extensive structural, mechanical, and electrical design. These tanks are widely spaced to avoid the transmission of fire and to allow room for a surface detention area to store burning fuel.

Airlines use either propylene glycol or ethylene glycol, mixed with water, to deice aircraft and to provide coatings that retard development of ice accumulation on aircraft. Facilities are required to store the glycol, mix it, and sometimes heat it before distribution. It can then be loaded onto deicing trucks for spraying onto the aircraft or distributed to deicing gantries or booms that are fixed to a base but have an extendable arm with an operator cab on top from which to spray the fluid.

Glycols have serious water-quality impacts on receiving streams and treatment plants. Increasingly, airports are making provisions to grade the ramp areas where deicing operations occur so that the spent glycol can be captured. Once it is segregated from the normal storm collection system, it can be

stored for metering in treatments facilities, reprocessed for reuse off airport, or otherwise disposed of. Ginger Sunday Evans

Bibliography. A. Graham, *Managing Airports: An International Perspective*, 2d ed., 2003; R. M. Horonjeff, F. X. McKelvey, and B. Sproule, *Planning and Design of Airports*, 5th ed., 2007; R. de Neufville and A. Odoni, *Airport Systems: Planning, Design, and Management*, 2002; A. T. Wells, *Airport Planning and Management*, 4th ed., 2000.

Airport noise

The unwanted sound from airport operations, primarily from aircraft of all types. It affects neighbors both adjacent to and farther from the airport. It can be controlled using several approaches, and assessed using a number of tools. Airport traffic and construction, which also create noise, will not be discussed here.

Types of operations. Aircraft using an airport cause noise during several types of operations. The noisiest is when an aircraft takes off. The plane is fully loaded with passengers, cargo, and fuel and uses maximum thrust to take off from the runway. Neighbors behind and to the side of the runway experience the first noise from a take-off. At later times, people along the take-off path experience the noise from the plane as it passes overhead.

When the plane approaches its destination airport, it again causes noise for the people along its path. It is now much lighter, having used up fuel enroute. However, at United States airports, the planes are also much lower as they line up to land on the designated runway, coming in on a 3° glide slope. When the aircraft lands, the pilot applies the brakes and/or reverses the engines. This operation also causes noise for the neighbors adjacent to the airport.

Aircraft also produce noise when they taxi to or from the runways, gates, parking areas, or maintenance areas. In this case, the amount of thrust needed is much smaller than the amount needed to be airborne, and the associated noise is much less.

Finally, aircraft can cause noise when they perform maintenance run-ups. After aircraft maintenance, the Federal Aviation Administration (FAA) requires that the system involved be tested before flight. This may require running the engines to partial or full thrust. The run-up is done on airport property but without the plane actually taking off. A full-power run-up is nearly as noisy as a take-off.

Mitigation. Noise control can be applied at the source, along the sound path, or at the receiver. In the case of airport noise, the source is the aircraft and the receiver is the neighbor. Noise control at the source is preferred, since it reduces the total amount of sound released into the environment.

Source noise control of jets has progressed with time. The first jet aircraft, called stage 1, had no noise control on their engines. Federal law now prohibits use of stage 1 in the United States. Simple muffling was added to the jet engines in the 1970s to give stage 2 aircraft. At the same time, improvements were made to the engine design, resulting in about 10 decibels of noise reduction for a given aircraft type. This type of engine powers the stage 3 aircraft. As of January 1, 2000, all large turbojet aircraft operating in the United States are stage 3, by federal mandate. In anticipation of the stage 3 mandate, owners of stage 2 aircraft could sometimes use an engine retrofit, called a hushkit, to meet the stage 3 noise criteria, prolonging the useful life of the aircraft. Additional improvements, starting in 1992, have continued to reduce aircraft noise emissions. In June 2001, the International Civil Organization (ICAO) adopted the next quieter level, called Chapter 4 of their regulations (Annex 16). It effectively eliminates hushkitted older aircraft from operating in countries which adopt these rules. *See* AIRCRAFT ENGINE PERFORMANCE; JET PROPULSION; MUFFLER; TURBOJET.

Airports can reduce the noise from backing up, repositioning, or taxiing by having the aircraft towed to or from the gate. The noise from towing is much less than that from the aircraft engines. In addition, aircraft may be asked to use only the engine away from the neighbors, to cut the noise source level by 3–6 dB depending on the number of engines on the aircraft. Finally, aircraft can be connected to airport electrical service at the gates instead of running their auxiliary power units or they can use ground power units, especially at night.

The source noise during flight can be reduced by using lower thrust settings when planes fly over noise-sensitive areas. So, if neighbors are far from the airport, the aircraft can use maximum thrust to get as high as possible and lower thrust farther away. This method, encouraged by the FAA, is used at a number of United States airports.

The most common path noise control is by use of barriers or berms. For effective noise reduction (at least 5 dB), the top of the barrier must intercept the line-of-sight between the aircraft engine and the observer. Since the elevation of aircraft engines can be as high as 30 ft (9.1 m), noise reduction occurs only when the neighbors are close to the ground, when the aircraft are on the ground, and when the barriers are high. Noise from taxiing and run-ups can be reduced in this way. At some airports, buildings are constructed near the active taxiways to provide effective barriers.

A second type of path noise control is to increase the distance between the source and receiver. The sound level is decreased 6 dB for every doubling of distance. Airports, working with the FAA, can develop noise abatement departure or arrival paths to maximize the distance between the source and receiver or to take advantage of open space. The airport can also move or "buy out" the affected residents. The area left vacant can either be redeveloped as compatible land use, which may be commercial or industrial, or left undeveloped.

Path noise control of aircraft run-ups is achieved by use of a hush-house, a hangarlike structure in which the aircraft can perform its engine test. This partially

or fully enclosed structure reduces the amount of noise emitted to the environment. These facilities are used at military facilities, overseas, and at a small number of United States commercial airports.

Receiver noise control is usually by sound-insulating the structure. The windows are replaced with double or triple glazing, and doors may be replaced with solid core versions. Typical improvement in noise reduction is 5–25 dB depending on the original condition of the structure. Air conditioning may also be added. This approach, which is effective only when the neighbor is inside the structure, has been used in the United States since 1980.

Time-of-day considerations. Most airports receive complaints from their neighbors about aircraft noise. The complaints include not being able to hear the television, constant noise, sleep deprivation, inability to work, and vibrations. United States airports must (by federal law) be open 24 hours per day, 7 days per week, unless there is another facility which can take the traffic. Complainants may also be affected by through flights or other airports' operations.

To mitigate this situation, an airport can direct nighttime operations to occur on a runway which affects the minimum number of people, whenever operationally possible. Limitations can also be placed on runways which affect many neighbors or are very close to them, if other runways are available. Altitude restrictions may be added at night to maximize the distance between the aircraft and the neighbors.

At major airports, where the maximum hourly number of arrivals and departures may be larger than 100, there will be several runways which can be used for aircraft operations. In this case, an airport can request that the FAA alternate among them, consistent with operational requirements, to relieve communities from constant noise.

Assessment. Airport noise is measured with noise monitors, consisting of microphones and meters. Most large airports have a distributed network of permanent noise monitors which record observed noise levels and accumulate the data for future transfer to the airport's computer. Small airports are most likely to have only portable measuring equipment or to use consultants. *See* NOISE MEASUREMENT.

Observations of aircraft operations are either direct or automated. With an automated system, an airport may have access to the FAA's data or to passive radar data. These data give the exact location, identity, and time of every operation in the area. Direct observations of aircraft operations are very reliable but are labor-intensive.

The radar data can be used to determine which operation affected a given resident at a given date and time. It can also be accumulated to determine the number of operations in a year by aircraft type on each of the airport's runways by time of day. The information can be used by the airport for assessing conformance with established noise abatement paths and procedures and for verifying operational reports from the airlines.

Combined with information on airline engine types and typical performance data on the aircraft type, the radar data can be used in airport noise prediction models, such as the FAA's Integrated Noise Model, to determine contours of noise exposure. These contours are averages of day-night equivalent sound levels for the year and can be compared with measured averages of the same type for locations for which the airport has noise monitors. These contours, in turn, can be used by the airport for planning purposes or for obtaining federal funds for noise mitigation programs.

If an airport has regulations, the above information may be used for enforcement. If the regulations are noise-level-based, microphone data will be used for this purpose. At airports where the regulations are operations-based, observations by airport personnel are often used to ensure compliance with local regulations. Complaints from the community may assist an airport in identifying potential violations. The FAA also assists the airport in ensuring compliance with noise regulations.

In deciding to adopt new noise mitigation at an airport, the FAA's regulations in Part 150 of the Code of Federal Regulations govern. A study of existing noise impacts is prepared, and mitigation strategies such as those discussed above are proposed. The FAA reviews the document and may approve it for federal funding of the proposed projects. Any new regulations would also need to be approved by the FAA. In addition, many states have environmental reporting requirements for their airports. Measurements of the noise impacts and new initiatives are properly filed with the appropriate authority. *See* ACOUSTIC NOISE.

Nancy S. Timmerman

Bibliography. M. J. Crocker (ed.), *Encyclopedia of Acoustics*, Wiley, New York, 1997; C. M. Harris, *Handbook of Acoustical Measurements*, 3d ed., Acoustical Society of America, Sewickley, PA, 1998; C. M. Harris, *Handbook of Noise Control*, 2d ed., McGraw-Hill, New York, 1979; H. H. Hubbard, *Aeroacoustics of Flight Vehicles: Theory and Practice*, Acoustical Society of America, Woodbury, NY, 1995.

Airport surface detection equipment

A ground mapping system that uses analog radar equipment to provide surveillance of aircraft and other surface vehicles on an airport surface. It is used by air-traffic controllers in airport control towers to monitor and control the movement of aircraft and vehicles. A situation display of the targets includes a map identifying the runways and taxiways and a visual map of the airport features, created through the contrast on the radar display resulting from the absence of returns from smooth concrete surfaces and the ground-clutter returns from grassy areas. An important safety function of the airport surface detection equipment (ASDE) is to determine whether or not a runway is clear for the next departure or arrival operation. The prevention of runway incursions by aircraft or by service vehicles has become

more urgent because of several incidents since 2003. This runway clearance determination is aided by the ASDE's capability to display an image of the aircraft in which the target's extremities (nose, tails, and wing tips), especially for large aircraft, are evident to the eye. *See* AIRPORT.

Operation. The ASDE antenna revolves at 60 revolutions per minute, providing a rapid update of target movements. It is located on a high vantage point, typically on top of the tower cab or on a special remote tower, that provides line-of-sight coverage of the desired runway and taxiway areas. Older ASDE systems have several inherent weaknesses, especially equipment failures, and excessive clutter during precipitation conditions, and are typically operated at a fixed frequency (for example, the ASDE-2 operates at 24 GHz).

Modern ASDEs utilize digital processing and an interrogation technique called frequency agility, which permits the transmission of up to 16 different Ku-band frequencies within a cycle. These techniques optimize target detection in the presence of ground and rain clutter and the rejection of false targets. In the digital processing, different thresholds for rejecting clutter are used on specified areas of the airport surface, such as on paved surfaces versus grassy areas. These systems are designed to detect all aircraft and vehicles having a radar cross section of 30 ft^2 (3 m^2) or greater, and will resolve two closely spaced targets when separated by 40 ft (12 m) in range or 80 ft (24 m) in azimuth at a range of 12,000 ft (3650 m). The basic coverage is the entire airport surface out to 24,000 ft (7300 m) in range and 200 ft (60 m) in altitude. Areas where coverage is not desired are ignored for detection and blanked out on the situation display.

Automation functions. The digital processing in modern ASDEs also provides the basis for automation functions that process target data to provide controllers with improved information. Principal automation features include target tracking, the application of runway safety logic functions to the tracks, and subsequent automatic alerting of the controller to dangerous situations that are determined by the logic. Advanced airport surface surveillance systems will enhance the data already provided by the ASDE, permitting more advanced automation functions. The design requirements for these systems include providing data that allows determination of target identity and target intent from flight-plan information or from data-link messages between the control tower and aircraft. *See* ELECTRONIC NAVIGATION SYSTEMS; MOVING-TARGET INDICATION; RADAR.

Robert A. Bales

Airship

A propelled and steered aerial vehicle, dependent on the displacement of air by a lighter gas for lift. An airship, or dirigible balloon, is composed primarily of a streamlined hull, usually a prolate ellipsoid which contains one or more gas cells, fixed and movable tail surfaces for control, a car or cabin for the crew or passengers, and a propulsion system.

Past designs. Two fundamentally different designs have been successfully used for past airships, the nonrigid and the rigid. A third type, the semirigid, is

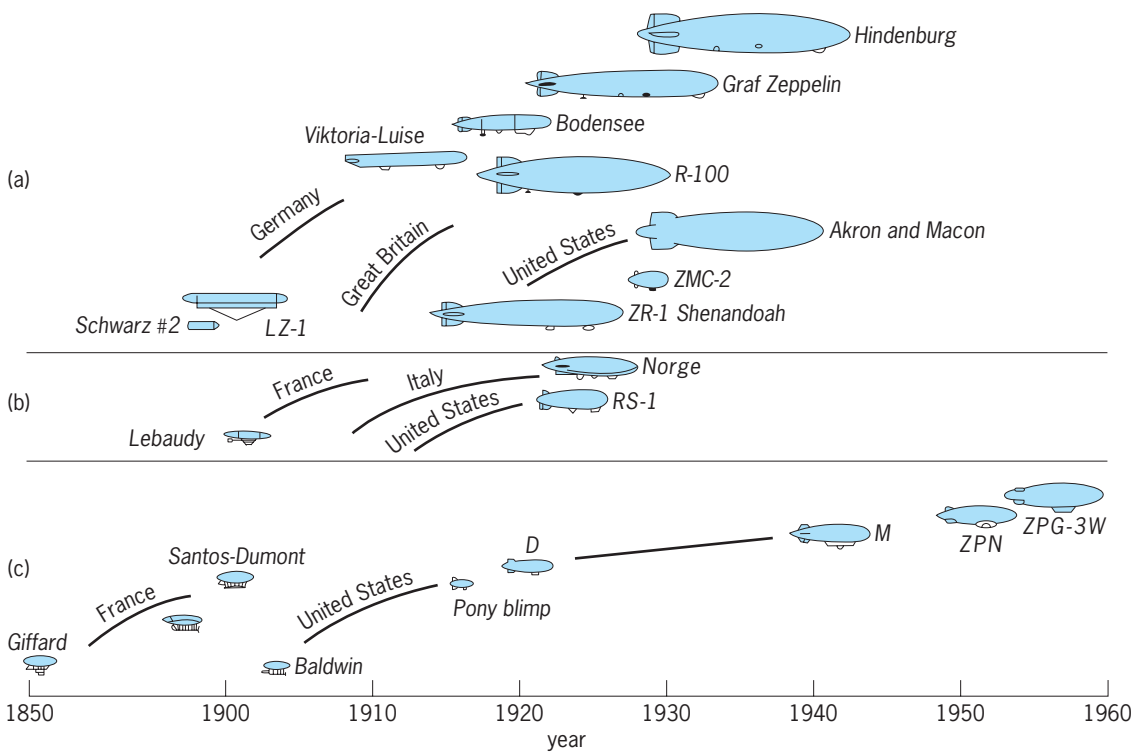

Fig. 1. History of airship development. (*a*) Rigid airships. (*b*) Semirigid airships. (*c*) Nonrigid airships.

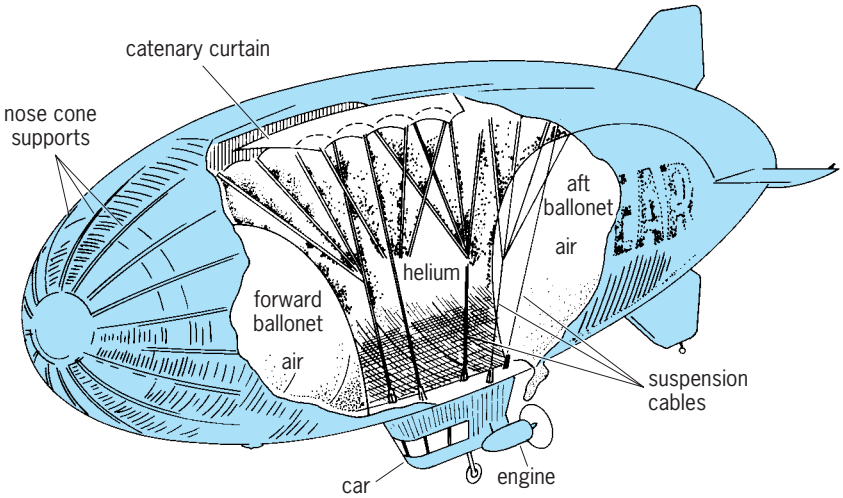

Fig. 2. Typical nonrigid airship.

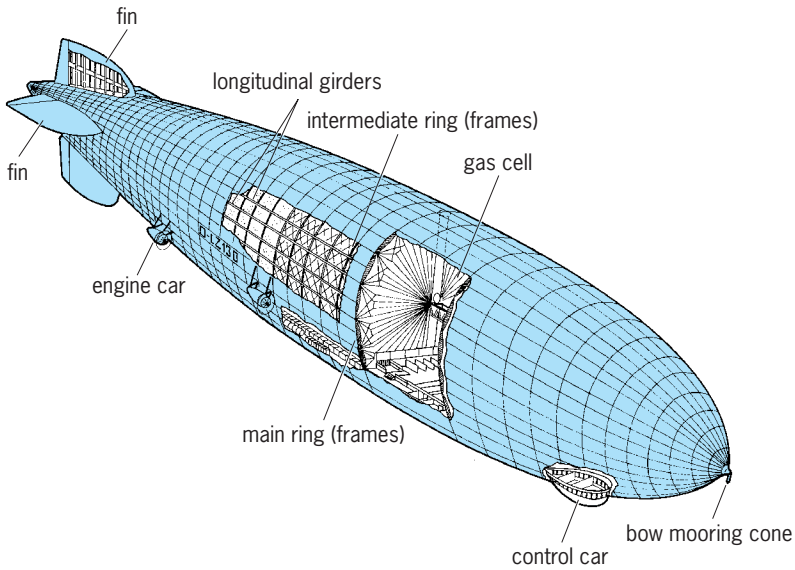

Fig. 3. Typical rigid airship.

Fig. 4. *Akron*, rigid airship of the U.S. Navy.

essentially a variant of the nonrigid type, differing by the addition of rigid keel. The principal development trends of the three types of conventional airships are depicted in **Fig. 1**.

Nonrigid. A typical nonrigid airship, or blimp (**Fig. 2**), consists of a flexible envelope, usually fabric, filled with lifting gas that is slightly pressurized. Internal air compartments (ballonets) expand and contract to maintain the pressure in the envelope as atmospheric pressure and temperature vary. Ballonet volume is controlled by ducting air from the prop wash or by electric blowers. The weights of the car structure, propulsion system, and other concentrated loads are supported by catenary systems attached to the envelope.

The nonrigid airships are historically significant for two reasons. First, a nonrigid airship was the first aircraft of any type to achieve controllable flight, in the 1850s. Second, nonrigid airships were the last type to be used on an extensive operational basis; the U.S. Navy decommissioned the last of its nonrigid airship fleet in the early 1960s. During the many years the Navy operated them, a high degree of availability and reliability was achieved. Most of these nonrigid airships were built by Goodyear, and a few, based on a modified pre-World War II Goodyear design, are used today for advertising by that company. *See* BLIMP.

Rigid. The other major type of airship is classified rigid because of its structure (**Fig. 3**). This structure was usually an aluminum ring-and-girder frame. An outer covering was attached to the frame to provide a suitable aerodynamic surface. Several gas cells were arrayed longitudinally within the frame. These cells were free to expand and contract, thereby allowing for pressure and temperature variations. Thus, despite their nearly identical outward appearance, rigid and nonrigid airships were significantly different in their construction and operation.

The rigid airship was developed primarily by the Zeppelin Company of Germany, and, in fact, rigid airships became known as zeppelins. Even the small percentage of rigid airships not built by this company was based, for the most part, on the Zeppelin concept. The rigid airships of the Zeppelin Company recorded some historical firsts in air transportation, including inaugurating the first passenger air service. The culmination of zeppelin development were the *Graf Zeppelin* and *Hindenburg* airships—unquestionably outstanding engineering achievements in their day. All of the rigid airships produced in the United States were for military purposes; none was in operation at the outbreak of World War II.

The last rigid airships built and operated in the United States were the *Akron* (**Fig. 4**) and *Macon*, built in 1931 and 1933. These airships were in many respects typical of the large rigid type but also had some unique design features. They contained just under 7,000,000 ft^3 (200,000 m^3) of helium, were a little over 800 ft (240 m) long, and had a gross lifting capability of about 450,000 lb (200,000 kg). Among the unique features of these airships were their eight engines. The propellers could be swiveled

in any vertical or fore and aft direction, which made the *Akron* and *Macon* two of the first aircraft to use vectored thrust propulsion. The panels on the side of the airship that looked like windows were actually condensers which collected water vapor out of the engine exhaust to replace weight on board the airship as fuel was burned. The internal rigid structure of the *Akron* and *Macon* was made of aluminum alloy. In fact, the use of aluminum alloy in aerospace structures was pioneered by rigid airships. The role of the *Akron* and *Macon* was to serve as flying aircraft carriers and scout vehicles. They carried five small airplanes on board to extend their range of surveillance, protect the airship, and serve as scouts for the surface fleet.

Commercial operations. The only significant past commercial airship operations were those of the Zeppelin Company and its subsidiary DELAG. This organization began carrying passengers on flights within Germany in 1910, and from 1929 to 1937 provided a unique transatlantic air service with the *Graf Zeppelin* and *Hindenburg*. Prior to the burning of the *Hindenburg* in 1937, there had been no fatal accidents in commercial service. However, none of these commercial operations can be considered a financial success, and most were subsidized by the German government.

Modern vehicle concepts. **Figure 5** shows a representative sample of the many modern airship vehicle concepts which have been proposed. The fully buoyant conventional concepts are modern versions of the classical, fully buoyant, rigid and nonrigid airship concepts. Fully buoyant means all the lift is provided by displacement. These airships would make extensive use of modern aircraft structural materials, propulsion systems, control systems, and electronics.

The other four concepts in Fig. 5 are partially buoyant, or hybrid, designs in which the buoyant lift is substantially augmented by aerodynamic or propulsive lift. Thus, these vehicles are partly heavier than air and partly lighter than air. Two of the hybrids are short takeoff and landing (STOL) vehicles, and two have vertical takeoff and landing (VTOL) capability. Generally speaking, the partially buoyant concepts would have higher cruise speed, higher fuel consumption, and higher structural weight as compared with the classical concepts. It will therefore depend on the specific applications as to which is the best concept. An important advantage of the hybrids is that they promise to alleviate the costly ground-handling requirement of past airship designs. *See* SHORT TAKEOFF AND LANDING (STOL); VERTICAL TAKEOFF AND LANDING (VTOL).

The Dynairship is representative of lifting-body concepts in which the hull is shaped to be more aerodynamically efficient than the conventional airship shape. This particular concept has a delta platform. The vehicle is flown at positive angle of attack in cruise to generate aerodynamic lift. The Megalifter concept is a typical example of a winged airship. The Helipsoid is intended to be a compromise between the ellipsoidal and deltoid shapes. It will have a better

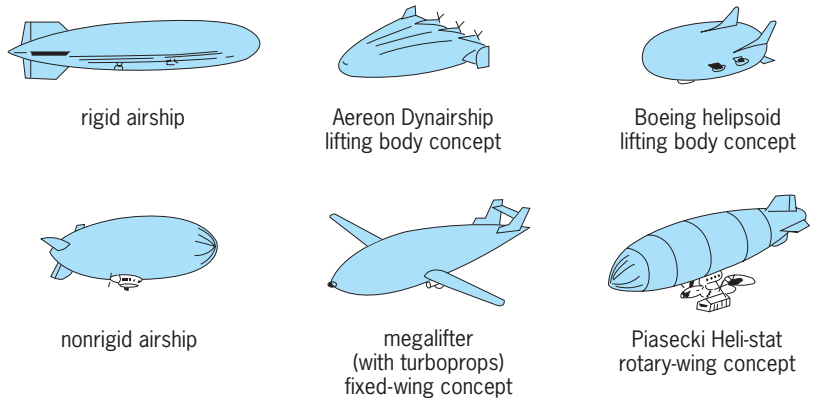

Fig. 5. Modern airship concepts.

surface-area/volume ratio than the Dynairship, at the expense of the degraded stability and control characteristics. Finally, the Heli-stat, which is representative of the rotary-wing hybrids, combines an airship hull with helicopters or helicopter rotor systems. *See* BALLOON.

Mark R. Ardema

Airspeed indicator

A device that computes and displays speed of an aircraft relative to the air mass in which the aircraft is flying. Three common types are the indicated airspeed meter, true airspeed indicator, and Machmeter.

The commonest type is the indicated airspeed meter, which measures differential pressure between the total ram pressure from the Pitot system and the total static pressure; it then converts this difference into units of speed (miles per hour or knots) under standard conditions, 29.9212 in. Hg (101.325 kilopascals) absolute and 59°F (15°C; see **illus.**). Although the indicated values are incorrect above zero altitude, the relationship to the aircraft handling remains essentially unchanged, thus providing a measure of the flyability of the aircraft. These instruments

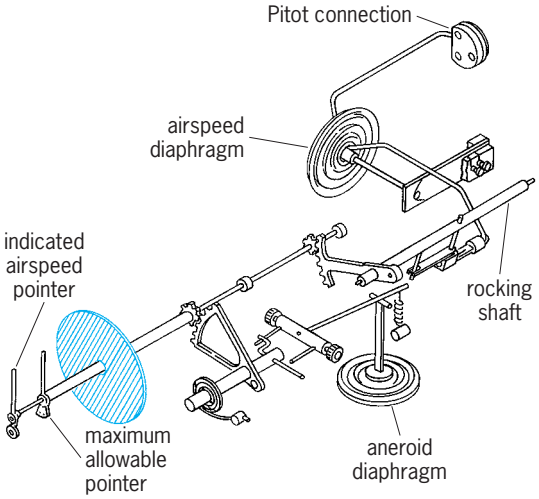

Components of indicated airspeed meter.

are frequently supplemented with an auxiliary mechanism, driving a second pointer which computes and indicates a value known as the maximum allowable speed for the type of aircraft in which the instrument is installed.

True airspeed indicators are similar to indicated airspeed meters but include a more complex mechanism that also senses both the absolute pressure and temperature, and compensates for the change of density of the air mass, thus obtaining true airspeed. This indication is of value in computing course information; hence the true airspeed indicator is a navigation instrument whereas the indicated airspeed indicator is a flight instrument.

For those aircraft that reach higher speeds (transonic and supersonic), the ratio of the actual speed to the local speed of sound is used. Devices that compute this value are known as Machmeters. Relative Mach number is computed by determining the ratio of total ram pressure divided by total static pressure and converting the ratio into Mach number, according to the applicable formulas. Certain instruments combine airspeed, Mach number, and maximum allowable speed in one indicator to provide a central speed-data display. *See* MACH NUMBER; PITOT TUBE.

James W. Angus

Aistopoda

An order of extremely elongate, limbless fossil amphibians in the subclass Lepospondyli from Permo-Carboniferous rocks of North America and the British Isles. The order includes three families: Lethiscidae (*Lethiscus*), Ophiderpetontidae (*Coloraderpeton*, *Ophiderpeton*), and Phlegethontiidae (*Aornerpeton, Phlegethontia, Sillerpeton*). All genera are monotypic with the exception of *Ophiderpeton* (six species) and *Phlegethontia* (possibly three species). In seeming contrast to their derived state, aistopods are among the very oldest of all known fossil tetrapods.

The skulls of aistopods are fenestrated and exhibit a reduced number of bony elements. This fenestration is extreme in phlegethontiids, whose skulls may have been kinetic.

The vertebral centra are holospondylous (single-pieced), hourglass-shaped, and fused to their neural arches. Vertebrae can exceed 200 in number and frequently bear foramina for passage of spinal nerves. Ribs typically bear a unique process, which gives them a K-shape in some species.

Lethiscids and ophiderpetontids possess a ventral armor of fusiform, bony elements arranged in a chevron pattern. Such armor in phlegethontiids consists of more slender, threadlike elements.

Most aistopods were presumably aquatic, although rib specializations, and the more gracile proportions of phlegethontiids, suggest a rather snake-like, terrestrial habit. *See* AMPHIBIA; LEPOSPONDYLI.

Carl F. Wellstead

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988.

Albedo

A term referring to the reflecting properties of a surface. White surfaces have albedos close to 1; black surfaces have albedos close to 0.

Bond albedo and normal reflectance. Several types of albedos are in common use. The Bond albedo (A_B) determines the energy balance of a planet or satellite and is defined as the fraction of the total incident solar energy that the planet or satellite reflects to space. The "normal albedo" of a surface, more properly called the normal reflectance (r_n), is a measure of the relative brightness of the surface when viewed and illuminated vertically. Such measurements are referred to a perfectly white Lambert surface—a surface which absorbs no light and scatters the incident energy isotropically—usually approximated by magnesium oxide (MgO), barium sulfate ($BaSO_4$), or some other bright material. *See* PHOTOMETRY; PLANET.

Bond albedos for solar system objects range from 0.9 for Saturn's icy satellite Enceladus and Neptune's Triton to values as low as 0.01–0.02 for dark objects such as the satellites of Mars (**Table 1**). Cloud-shrouded Venus has the highest Bond albedo of any planet (0.76). The value for Earth is typically 0.35, but varies with the extent of cloud and snow cover. The Bond albedo is defined over all wavelengths, and its

TABLE 1. Bond albedos and visual geometric albedos of selected solar system objects

Object	Bond albedo (A_B)	Visual geometric albedo (p_v)
Mercury	0.12	0.14
Venus	0.76	0.59
Earth	0.35	0.37
Mars	0.24	0.15
Jupiter	0.34	0.45
Saturn	0.34	0.46
Uranus	0.34	0.48
Neptune	0.28	0.50
Pluto	0.5(?)	0.61
Moon	0.12	0.14
Phobos (M1)	0.02	0.05
Deimos (M2)	0.02	0.06
Io (J1)	0.56	0.63
Europa (J2)	0.58	0.68
Ganymede (J3)	0.38	0.43
Callisto (J4)	0.13	0.17
Amalthea (J5)	0.02	0.06
Mimas (S1)	0.60	0.75
Enceladus (S2)	0.90	1.20
Tethys (S3)	0.60	0.80
Dione (S4)	0.45	0.55
Rhea (S5)	0.45	0.65
Titan (S6)	0.20	0.20
Ariel (U1)	0.21	0.39
Umbriel (U2)	0.10	0.21
Titania (U3)	0.15	0.27
Oberon (U4)	0.12	0.23
Miranda (U5)	0.18	0.32
Triton (N1)	0.90	0.75
Nereid (N2)	0.07	0.14
Proteus (N8)	0.03	0.06
Ceres	0.03	0.06
Vesta	0.12	0.23
Comet Halley (nucleus)	0.01	0.03

TABLE 2. Normal reflectances of materials*

Material	Albedo
Lampblack	0.02
Charcoal	0.04
Carbonaceous meteorites	0.05
Volcanic cinders	0.06
Basalt	0.10
Iron meteorites	0.18
Chondritic meteorites	0.29
Granite	0.35
Olivine	0.40
Quartz	0.54
Pumice	0.57
Snow	0.70
Sulfur	0.85
Magnesium oxide	1.00

* Powders; for wavelengths near 0.5 micrometer.

value therefore depends on the spectrum of the incident radiation. For objects in the outer solar system not yet visited by spacecraft (such as Pluto), the values of A_B in Table 1 are estimates derived indirectly, since for these bodies it is impossible to measure the scattered radiation in all directions from Earth.

Normal reflectances of some common materials are listed in **Table 2**. The normal reflectances of many materials are strongly dependent on wavelength, a fact that is commonly used in planetary science to infer the composition of surfaces remotely. While the Bond albedo cannot exceed unity, the normal reflectance of a surface can do so if the material is more backscattering at opposition than the reference surface.

Geometric albedo. In the case of solar system objects, a third type of albedo, the geometric albedo (p), is commonly defined. It is the ratio of incident sunlight reflected in the backscattering direction (zero phase angle or opposition) by the object, to that which would be reflected by a circular disk of the same size but covered with a perfectly white Lambert surface. Objects like the Moon, covered with dark, highly textured surfaces, show uniformly bright disks at opposition (no limb darkening); for these, $p = r_n$. At the other extreme, a planet covered with a Lambert-like visible surface (frost or bright cloud) will be limb-darkened as the cosine of the incidence angle at opposition, and $p = {}^2/_3 r_n$. Table 1 lists the visual geometric albedo p_v, which is the geometric albedo at a wavelength of 550 nanometers.

Phase integral. For solar system objects, the ratio A_B/p is called the phase integral and is denoted by q. Here p is the value of the geometric albedo averaged over the incident spectrum, and does not generally equal p_v given in Table 1. Values of q range from near 1.5 for cloud-covered objects (1.4 for Saturn and 1.3 for Venus) to 0.2 for the very dark and rugged satellites of Mars: Phobos and Deimos. The value for the Moon is about 0.6.

Other albedo types. The quantities AB, p, and q quantify the reflection of light averaged over an entire hemisphere of a planet. They are most appropriately used when details of the distribution of the reflectance on a planet's surface are beyond the resolution of the observations. It is possible to define other types of albedos based on the geometry and nature (diffuse or collimated) of the incident beam and scattered beams. For such bidirectional albedos or reflectances, the angles of incidence and scattering as well as the angle between these two directions (phase angle) must be specified. Finally, in determining the energy balance of surfaces, one is often interested in the analog for a flat surface of the Bond albedo, that is, the fraction of the incident energy that the surface reflects. For a plane parallel beam incident at an angle i, this quantity, $A(i)$, is generally close to the value A_B, but can be significantly larger than A_B for some surfaces, for large values of i.

Role in planetary astronomy. The role that these albedos play in planetary astronomy is illustrated by the discovery of a distant new planet in the solar system. The distance of the planet from the Earth and Sun is accurately known from its motion; determination of its orbit requires only a small number of measurements of its position relative to the stars over a few months. The total brightness of the planet is the product of its area and (geometric) albedo. For a distant or small object whose size cannot be resolved with even the largest telescopes, an albedo may be assumed, allowing its size to be estimated from its brightness alone. What is known about the planet can be improved if thermal radiation from the planet is detected at infrared wavelengths. The quantity of incident solar energy absorbed by a planet is $(1 - A_B)$, that is, energy that is not reflected is absorbed. In this way, the Bond albedo determines the temperature of a planet. The absorbed sunlight heats the planet until it reaches an equilibrium temperature where the total energy it radiates at infrared wavelengths equals the total solar energy absorbed. The infrared brightness depends on the area of the planet and its temperature, so if both the infrared (thermal emission) and visual brightness (reflected sunlight) are measured, both the size and the albedo of a distant unresolved planet can be determined unambiguously. Using this technique, the sizes and albedos of thousands of asteroids and other objects that populate the solar system are now known. *See* HEAT RADIATION; INFRARED ASTRONOMY; RADIOMETRY.

Joseph Veverka; Jay Goguen

Bibliography. J. K. Beatty, C. C. Petersen, and A. Chaikin, *The New Solar System*, 4th ed., Sky Publishing, Cambridge, MA, 1999; B. Hapke, *Theory of Reflectance and Emittance Spectroscopy*, Cambridge University Press, 1993; M. Harwit, *Astrophysical Concepts*, 4th ed., Springer, 2006.

Albite

A sodium-rich plagioclase feldspar mineral whose composition extends over the range $Ab_{100}An_0$ to $Ab_{90}An_{10}$, where Ab (= albite) is $NaAlSi_3O_8$ and An (= anorthite) is $CaAl_2Si_2O_8$ (see **illus.**). Albite occurs in crustal igneous rocks as a major component of pegmatites and granites, in association with quartz, mica (usually muscovite), and potassium

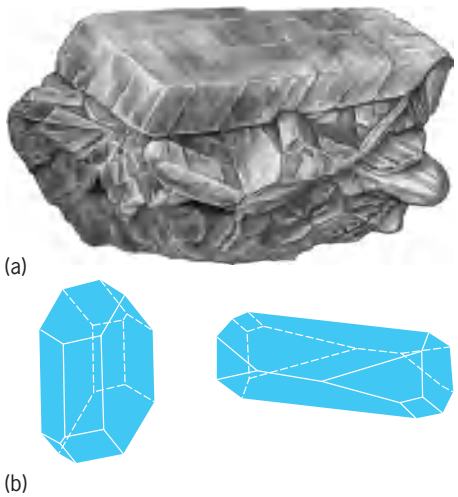

Albite. (*a*) Crystals, Amelia Court House, Virginia (*specimen from Department of Geology, Bryn Mawr College*). (*b*) Crystal habits (*after D. S. Hurlbut, Jr., Dana's Manual of Mineralogy, 17th ed., Wiley, 1959*)

feldspar (orthoclase or microcline). Sodium and potassium feldspars usually occur as distinct mineral grains, sizes varying from millimeter to meter scale. However, they are frequently intergrown, having exsolved from a single phase at high temperatures. If the intergrowth is visually observable in a hand specimen, the composite material is known as macroperthite; if visible only in a microscope, microperthite; and if submicroscopic in scale, cryptoperthite. In metamorphic rocks albite is found in granitic gneisses, and it may be the principal component of arkose, a feldspar-dominant, sedimentary rock. Cleavelandite, a platy variety, is sometimes found in lithium-rich pegmatites. *See* ARKOSE; GNEISS; PERTHITE.

Its hardness is 6 on the Mohs scale, specific gravity 2.62, and melting point 1118°C (2044°F). If observed at temperatures above 985°C (1805°F), albite has monoclinic symmetry, provided that the distribution of aluminum (Al) and silicon (Si) among the tetrahedral sites of its framework structure is random (disordered). Below 958°C (1805°F) the symmetry becomes triclinic as the $[AlSi_3O_8]_\infty^{1-}$ tetrahedral framework collapses around the small sodium atom, primarily through the mechanism of intertetrahedral angle bending. With annealing at lower temperatures, Al and Si become completely ordered, forming what is called low albite. If the An component exceeds 2–3 mole %, this feldspar exsolves into two distinct, submicroscopic phases, one of nearly pure albite (An_{0-1}) composition, and the other of composition between An_{15} and An_{25}; the resultant lamellar intergrowth of these two phases is called peristerite, with optical interference colors resembling those on the neck feathers of a pigeon. Such material, especially that from Hybla, Ontario, Canada, is valued as a semiprecious gemstone and may be cut and polished into cabochons. Albite is used in ceramic materials, as a source of soda (Na_2O) and alumina (Al_2O_3) in glasses, and in certain, relatively soft abrasives. *See* FELDSPAR; GEM; IGNEOUS ROCKS; PEGMATITE.

Paul H. Ribbe

Albuliformes

An order of actinopterygian fishes in the subdivision Elopomorpha, along with Elopiformes, Anguilliformes, and Saccopharyngiformes, all of which have a leptocephalous larval stage. Albuliformes, comprising three families, eight genera, and 30 species, are distinguished from other Elopomorpha by an open mandibular sensory canal in the dentary and angular bones. Two suborders are recognized, Albuloidei and Notacanthoidei, the latter of which was recognized previously as the order Notacanthiformes. *See* ACTINOPTERYGII; ANGUILLIFORMES; ELOPIFORMES; NOTACANTHOIDEI.

Albuloidei, consisting of the family Albulidae, are characterized by a fusiform body; the tip of the snout overhanging an inferior mouth; a mouth bordered by premaxillae; a gular plate having a thin median splint or being entirely absent; an infraorbital lateral-line canal extending onto the premaxillae; and caudal fin rays supported by six hypural bones. Albulidae contains two subfamilies, Albulinae with one genus (*Albula*) and three species and Pterothrissinae with one genus (*Istieus*) and two species. *Albula* (see **illustration**) differs from *Istieus* (formerly *Pterothrissus*) in having 16–21 dorsal fin rays versus 55–65; a small gular plate versus no gular plate; and the maxillae toothless versus maxillae toothed. *Albula* occurs in tropical seas, rarely entering brackish and fresh water, and attains a length of 105 cm (41 in.).

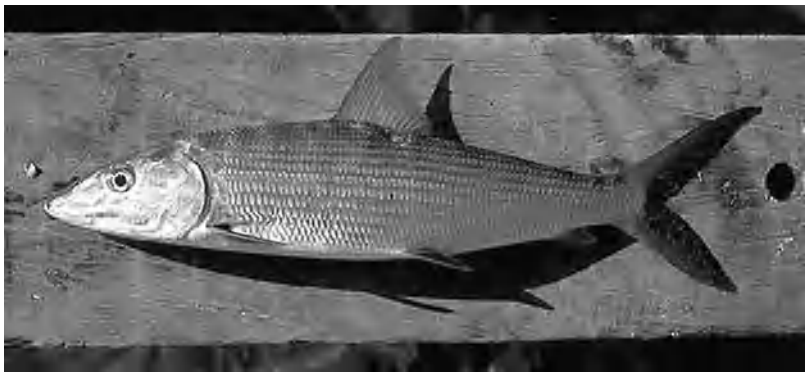

***Albula vulpes*. (Courtesy of C. Q. Jessen)**

Istieus occurs in the eastern Atlantic from Mauritania to Namibia and in the western Pacific from China and Japan, and probably does not exceed 40 cm (16 in.) in length. *Albula* can tolerate low oxygen tension by "inhaling" air into its lunglike air bladder. It is benthic in habitat and feeds on invertebrates. *Istieus* is also benthic but usually occurs in deeper waters of continental slopes. Herbert Boschung

Bibliography. J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; D. G. Smith, Albulidae (pp. 683–684), Halosauridae (685–687), Notacanthidae (688–689), in K. E. Carpenter (ed.), *The Living Marine Resources of the Western Central Atlantic*, FAO Species Identification Guide for Fishery Purposes, vol. 2, FAO, Rome, 2002.

Albumin

A plasma protein produced by the liver that maintains fluid balance in the blood and transports fatty acids in the plasma and interstitial fluid. It is the most abundant protein in human serum, and one of the first discovered and earliest studied proteins. Serum albumin was precipitated from urine in 1500 by Paracelsus, and was crystallized in 1894 by A. Gürber. Probably no other protein has been studied as extensively as serum albumin, and our knowledge of its structure and interactions with its ligands has come from many researchers, using a great variety of experimental approaches. Its ability to bind many different ligands, most of which are hydrophobic anions, and several molecules of the same ligand (fatty acid) is well documented. Fatty acids, bilirubin (an orange-yellow bile pigment formed in the breakdown of red blood cells), and hematin (a blue to blackish-brown compound formed in the decomposition of hemoglobin) represent the endogenous ligands of albumin with highest affinity. *See* PROTEIN.

Structure. A major breakthrough in albumin research came in the early 1990s, when x-ray diffraction studies of crystalline human albumin provided the first accurate images of this remarkable protein. The high-resolution structure revealed an overall heart shape, both in the defatted form and with bound fatty acids (see **illustration**). The protein comprises three domains, each with two subdomains, as predicted by many studies. The fatty acids are bound throughout the protein in structurally distinct sites with similar features: the carboxyl group of each fatty acid forms a salt bridge or a hydrogen bond with basic and polar amino acid side chains, and the hydrocarbon chain of the fatty acid fits into a hydrophobic cavity between helices.

Physiological function. In human plasma, albumin is the major buffer for pH variations and is essential for maintaining the plasma colloidal osmotic pressure, causing fluid to remain in the blood instead of leaking out into the tissues. Albumin is the main transport vehicle for fatty acids for utilization in various tissues. As albumin circulates through the vascular system, fatty acids desorb rapidly from the protein in spite of their high affinity for the binding sites (see illustration) and diffuse through the adjacent cell membranes to enter cells. Albumin also binds fatty acids that are released into the plasma from cells such as fat cells for transport to other tissues, where the fatty acids serve as fuel or substrates for lipid synthesis. *See* BLOOD; LIPID METABOLISM.

Role in disease. Low plasma levels of albumin are associated with several diseases, such as diabetes, liver disease, and cardiovascular disease. Undetectably low or very low plasma levels of albumin are found in an extremely rare inherited disorder known as analbuminemia. There is inadequate compensation for albumin's role in maintaining plasma colloidal osmotic pressure, which is very low in affected individuals; however, the plasma lipoproteins

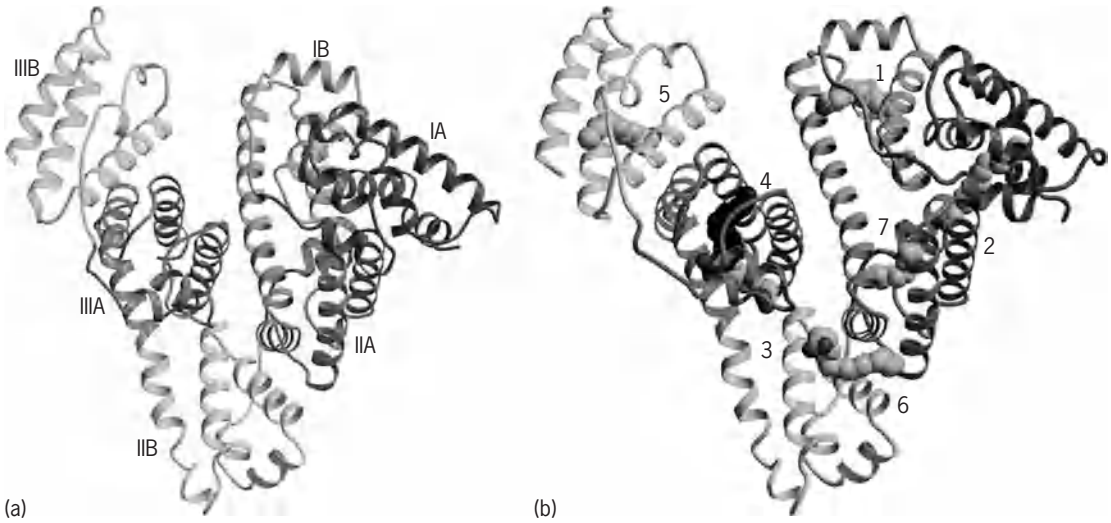

Human serum albumin. (*a*) Domain structure of the defatted protein and (*b*) location of eight myristic acid (a 14-carbon fatty acid) ligands in different binding sites. In *a*, the secondary structure of the protein is shown schematically, and the domains are designated I, II, III. The subdomains are A and B. Fatty acids with different chain lengths are present in the plasma, and the binding affinity increases with increasing chain length of the fatty acid, up to a length of 20 carbons. In *b*, showing the binding of the myristic acid, the atom type is differently shaded to show carbon and oxygen. Myristic acid in site 4 is darker to distinguish it from myristic acid in site 3.

take over the role of transport of fatty acids and bilirubin. Analbuminemia is not a lethal disorder, although edema, fatigue, and lipodystrophy (a disturbance of fat metabolism in which the subcutaneous fat disappears over some regions of the body but is unaffected in others) are common complaints. The high concentration of albumin in the general population, therefore, is not essential for life but appears to be important for optimal health.

Drug transport. Human serum allumin (HSA) has two principal drug-binding sites (sites I and II in the classical literature), which have been localized in the crystal in subdomains IIA and IIIA (see illustration). Site I binds bulky heterocyclic molecules with a delocalized negative charge in the center of a predominantly hydrophobic structure, including warfarin, phenylbutazone, thyroxine, benzodiazepines, antidiabetic agents, and tryptophan. Site II binds aromatic carboxylic acids that are ionized at physiological pH and contain a negative charge at one end of the molecule, including propofol, diazepam, ibuprofen, acetylsalicylic acid (aspirin), bilirubin, and thyroxine. Solubilization of these hydrophobic compounds by albumin contributes to a more homogeneous distribution of the drug in the body, reduces the volume of distribution, lowers the rate of clearance, and increases the plasma half-life of the drug. *See* ASPIRIN; BILIRUBIN; HETEROCYCLIC COMPOUNDS; THYROXINE.

James A. Hamilton

Bibliography. J. R. Brown and P. Shockley, Serum albumin: Structure and characterization of its ligand binding sites, Wiley, New York, 1982; J. K. Choi et al., Interactions of very long-chain saturated fatty acids with serum albumin, *J. Lipid Res.*, 43:1000–1010, 2002; S. Curry et al., Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites, *Nat. Struc. Biol.*, 5:827–835, 1998; C. Daniels, N. Noy, and D. Zakim, Rates of hydration of fatty acids bound to unilamellar vesicles of phosphatidylcholine or to albumin, *Biochemistry*, 24:3286–3292, 1985; J. A. Hamilton, Fatty acid transport: Difficult or easy, *J. Lipid. Res.*, 39:467–481, 1998; X. M. He and D. C. Carter, Atomic structure and chemistry of human serum albumin, *Nature*, 358:209–215, 1992 [published erratum appears in *Nature*, 364(6435):362, July 22, 1993]; T. Peters, Jr., *All About Albumin*, Academic Press, San Diego, 1996; A. A. Spector, Fatty acid binding to plasma albumin, *J. Lipid Res.* 16:165–179, 1995.

Alcohol

A member of a class of organic compounds composed of carbon, hydrogen, and oxygen. They can be considered as hydroxyl derivatives of hydrocarbons produced by the replacement of one or more hydrogens by one or more hydroxyl (—OH) groups.

Classification. Alcohols may be mono—, di—, tri—, or polyhydric, depending upon the number of hydroxyl groups they possess. They are classified as primary (RCH_2OH), secondary (R_2CHOH), or tertiary (R_3COH), depending on the number of hydrogen atoms attached to the carbon atom bearing the hydroxyl group. Alcohols can also be characterized by the molecular configuration of the hydrocarbon portion (aliphatic, cyclic, heterocyclic, or unsaturated). There are two systems in use for alcohol nomenclature, the common naming system and the IUPAC (International Union of Pure and Applied Chemistry) naming system. The common name is sometimes associated with the natural source of the alcohol or with the hydrocarbon portion (for example, methyl alcohol, ethyl alcohol). The IUPAC method is a systematic procedure with agreed-upon rules. The name of the alcohol is derived from the parent hydrocarbon which corresponds to the longest carbon chain in the alcohol. The final "e" in the hydrocarbon name is dropped and replaced with "ol"; and a number before the name indicates the position of the hydroxyl. Examples of these two systems are given in the **table**.

Reactions. Oxidation of primary alcohols produces aldehydes (RCHO) and carboxylic acids (RCO_2H); oxidation of secondary alcohols yields ketones ($RCOR'$). Dehydration of alcohols produces alkenes and ethers (ROR). Reaction of alcohols with carboxylic acids results in the formation of esters ($ROCOR'$), a reaction of great industrial importance. The hydroxyl group of an alcohol is readily replaced by halogens or pseudohalogens. *See* ALDEHYDE; ALKENE; CARBOXYLIC ACID; ESTER; ETHER; KETONE.

Uses. Industrially, the monohydric aliphatic alcohols are classified according to their uses as the lower alcohols (1–5 carbon atoms), the plasticizer-range alcohols (6–11 carbon atoms), and the detergent-range alcohols (12 or more carbon atoms). The lower alcohols are employed as solvents, extractants, and antifreezes. Esters of the lower alcohols are employed extensively as solvents for lacquers, paints, varnishes, inks, and adhesives. The plasticizer-range alcohols find their primary use in the form of esters as plasticizers and also as lubricants in high-speed applications such as jet engines. The detergent-range alcohols are used in the form of sulfate and ethoxysulfate esters in detergents and surfactants.

Ethanol and methanol are volatile, combustible substances, but their use as motor fuel is limited by economic considerations. In the United States, ethanol made by fermentation of agricultural products is blended with gasoline to produce a motor fuel called gasohol, but this constitutes an insignificant fraction of the energy supply. In Brazil, where sugarcane is a major agricultural product, the large-scale production of ethanol by fermentation has been promoted by the government, and ethanol is a significant and rapidly growing component of the automotive fuel supply. *See* ETHYL ALCOHOL; METHANOL.

Sources. Alcohols are derived either from natural-product processing, such as the fermentation of carbohydrates and the reductive cleavage of natural fats and oils, or by chemical synthesis based on the hydrocarbons derived from petroleum or the synthesis gas from coal.

Alcohols and their formulas

Name		
Common	IUPAC	Formula
Methyl alcohol	Methanol	CH_3OH
Ethyl alcohol	Ethanol	CH_3CH_2OH
n-Propyl alcohol	1-Propanol	$CH_3CH_2CH_2OH$
Isopropyl alcohol	2-Propanol	$(CH_3)_2CHO$
n-Butyl alcohol	1-Butanol	$CH_3(CH_2)_2CH_2OH$
sec-Butyl alcohol	2-Butanol	$CH_3CH_2CHOHCH_3$
tert-Butyl alcohol	2-Methyl-2-propanol	$(CH_3)_3COH$
Isobutyl alcohol	2-Methyl-1-propanol	$(CH_3)_2CHCH_2OH$
n-Amyl alcohol	1-Pentanol	$CH_3(CH_2)_3CH_2OH$
n-Hexyl alcohol	1-Hexanol	$CH_3(CH_2)_4CH_2OH$
Allyl alcohol	2-Propen-1-ol	$CH_2{=}CHCH_2OH$
Crotyl alcohol	2-Buten-1-ol	$CH_3CH{=}CHCH_2OH$
Ethylene glycol	1,2-Ethanediol	$HOCH_2CH_2OH$
Propylene glycol	1,2-Propanediol	$CH_3CHOHCH_2OH$
Trimethylene glycol	1,3-Propanediol	$HOCH_2CH_3CH_2OH$
Glycerol	1,2,3-Propanetriol	$CH_2OHCHOHCH_2OH$

The fermentation of sugars and starches (carbohydrates) to produce alcoholic beverages has been employed at least since history has been recorded. The industrial fermentation process is the biological transformation of a carbohydrate by a highly specialized strain of yeast to produce the desired product, such as ethanol, or 1-butanol and acetone. Fermentation is no longer the major source of 1-butanol, but still accounts for all potable ethanol and a large proportion of the ethanol used industrially worldwide. As sources of hydrocarbons based on petroleum continue to be depleted, fermentation processes based on renewable raw materials are likely to become more important. *See* DISTILLED SPIRITS; FERMENTATION; INDUSTRIAL MICROBIOLOGY.

Syntheses. The major processes employed industrially to produce aliphatic monohydric alcohols are the reduction of synthesis gas, the hydration and oxonation of hydrocarbons, the condensation of aldehydes followed by reduction, the Ziegler process, and the reduction of animal fats and vegetable oils.

Reduction. Methanol, the simplest alcohol, is produced almost entirely by the catalytic reduction of synthesis gas (carbon monoxide and hydrogen), as shown in reaction (1).

$$\underset{\text{Carbon monoxide}}{CO} + \underset{\text{Hydrogen}}{2H_2} \longrightarrow \underset{\text{Methanol}}{CH_3OH} \qquad (1)$$

Hydration. Treatment of lower unsaturated hydrocarbons with water in the presence of an acidic catalyst is employed commercially to produce ethanol and 2-propanol, as shown in reactions (2) and (3).

$$\underset{\text{Ethylene}}{CH_2{=}CH_2} + H_2O \longrightarrow \underset{\text{Ethanol}}{CH_3CH_2OH} \qquad (2)$$

$$\underset{\text{Propylene}}{CH_3CH{=}CH_2} + \underset{\text{Water}}{H_2O} \longrightarrow \underset{\text{2-Propanol}}{CH_3\underset{\underset{OH}{|}}{C}HCH_3} \qquad (3)$$

Most of the synthetic ethanol in the United States is made by this method.

Oxonation. A modification of the Fischer-Tropsch process, known as the oxo reaction, involves the treatment of unsaturated hydrocarbons with synthesis gas at high temperatures under high pressure in the presence of a cobalt catalyst; it is a major source of monohydric aliphatic alcohols ranging from 1-propanol to the tridecanols. A mixture of isomeric alcohols is generally obtained by this process, shown in reaction (4).

$$\underset{\text{Ethylene}}{CH_2{=}CH_2} + \underset{\text{Carbon monoxide}}{CO} + \underset{\text{Hydrogen}}{H_2} \longrightarrow$$

$$\underset{\text{Propionaldehyde}}{CH_3CH_2CHO} \xrightarrow{H_2} \underset{\text{1-Propanol}}{CH_3CH_2CH_2OH} \qquad (4)$$

See FISCHER-TROPSCH PROCESS.

Condensation. Some of the higher aliphatic alcohols, such as 2-ethylhexanol, are synthesized by the condensation of aldehydes followed by dehydration and reduction. This base-catalyzed condensation, known as the aldol reaction, is shown in reaction (5).

$$\underset{\textit{n}\text{-Butyraldehyde}}{2CH_3CH_2CHO} \xrightarrow{\text{base}} \underset{\text{Butyraldol}}{CH_3CH_2CH_2\underset{\underset{CH_2CH_3}{|}}{\overset{\overset{OH}{|}}{C}HCHCHO}} \xrightarrow[\text{(2) } H_2]{\text{(1) } - H_2O}$$

$$\underset{\text{2-Ethylhexanol}}{CH_3CH_2CH_2CH_2\underset{\underset{CH_2CH_2}{|}}{C}HCH_2OH} \qquad (5)$$

Ziegler process. Formerly, the major source of detergent range alcohols was hydrogenolysis of animal fats and vegetable oils, but the Ziegler process, using aluminum alkyls as the catalyst and ethylene from the cracking of petroleum as the feedstock, has become the major source of industrial supply. In this sequence of reactions, a trialkyl aluminum such as

triethyl aluminum [$Al(C_2H_5)_3$] is alkylated with ethylene to give the desired higher trialkyl aluminum [AlR_3], which is then oxidized to the trialkoxy aluminum [$Al(OR)_3$], which is hydrolyzed to give a mixture of primary alcohols in the desired molecular weight range. *See* STEREOSPECIFIC CATALYST.

Ethylene glycol and propylene glycol are the most industrially significant polyhydric alcohols. They are made from ethylene, as shown in reaction (6), and

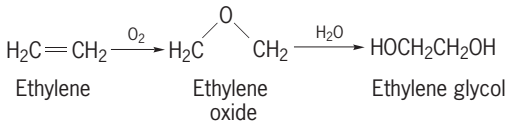

(6)

propylene, by means of their intermediate oxides. The oxides are synthesized by catalytic oxidation of the olefins. *See* ESTER; PHENOL; POLYOL. Paul E. Fanta

Bibliography. *Kirk-Othmer Encyclopedia of Chemical Technology*, vol. 1, 4th ed., 1991.

Alcohol fuel

Any alcohol burned as a fuel. Generally, the term alcohol refers to ethanol, the first organic chemical produced by humans. If available, any alcohol may be used as a fuel, but ethanol and methanol are the only ones which are sufficiently inexpensive. Alcohols are useful fuels, even though the oxygen atom in any alcohol molecule reduces its heating value, because the molecular structure increases the combustion efficiency. *See* COMBUSTION; GASOLINE.

Ethanol. Alcohol (ethanol) was used as a motor fuel, either alone or in mixtures with gasoline (gasohol), in many countries during the early 1930s, and it became an important alternative fuel for automobiles again in the 1970s and 1980s. It burns well in engines designed for gasoline and has a high octane rating. However, because of the high cost of the raw materials required, or of their conversion, it costs much more than gasoline. Gasohol is produced by dissolving 5–15% absolute water-free alcohol in gasoline. The 5% water fraction in 95% alcohol causes phase separation because of its insolubility, making 95% alcohol unsuitable for use in gasohol. This is first seen as cloudiness, then as small droplets of a water-alcohol solution, and finally as two distinct layers. Such phase separation also occurs because alcohol is hygroscopic; that is, it tends to absorb water from the air and cause dilution, either before or after mixing with gasoline.

Gasoline, alcohol, and water have individual solubility effects on the metals, coatings, and plastic parts of a car's carburetor and other parts of its fuel system. Usually, gasohol has about the same effects on these materials as gasoline; and in the absence of absorbed water, gasohol causes no difficulties in usual servicing and operation of automobiles. Because alcohol increases the octane rating, and gives other improvements in combustion, gasohol usually performs better and gives an increase in octane number higher than would be expected on the basis of the small fraction of alcohol in the fuel. *See* OCTANE NUMBER.

Specially modified engines and parts in contact with water-free (absolute or neat) alcohol account also for the considerably greater efficiency obtained in its use. Neat alcohol is not used to any extent as a fuel in the United States, but is widely used as a 95% impure grade in Brazil in cars fitted with special parts.

Ethanol has many advantages, but it does not compete in most countries with other liquid fuels on an economic basis for two reasons. First, raw materials are much more valuable for food or other uses, if there are no political constraints or artificial pricing. Second, basically, alcohol always requires more energy to produce than it will deliver in any use. Thus it will always be inefficient and expensive from a thermal standpoint. Nevertheless, alcohol will find use as fuel for a number of reasons: (1) In many countries there may be no other indigenous raw material from which to make motor fuels. (2) The plants for its product can be quite simple, small, and inexpensive compared to the very large, highly capital-intensive plants required for making other fluid fuels. (3) Technology is simple, well known, and established worldwide. (4) Highly labor-intensive programs are required with simple and familiar rural duties for most of the labor force. *See* ETHYL ALCOHOL.

Methanol. Carbon in any material can be converted to methanol. Natural gas during the 1970s and 1980s was favored as a raw material, and its energy may be delivered at a much lower cost if the gas is converted to methanol than if it is piped directly for 4000–5000 mi (6400–8000 km), even less and with much greater safety than if the gas is converted to liquefied natural gas, shipped for 2000–3000 mi (3200–4800 km), and stored at its cryogenic temperature. *See* LIQUEFIED NATURAL GAS (LNG); NATURAL GAS.

As gas prices increase, otherwise unsalable solid fuels will be used increasingly as feedstocks for methanol. Low-grade coal, lignite, or peat may be too far from market or have too much sulfur, chemically bound water, or ash to be viable fuels. Plants built alongside deposits of such fossil fuels may pipe or ship liquid methanol to market.

Methanol has been synthesized in the United States since 1925, first from synthesis gas (also known as syngas, a mixture of CO and H_2) manufactured from coal and, more recently, almost entirely from natural gas. While the basic chemistry is still the same, the many improvements in processing, flow sheets, catalysts, and construction materials now allow single units to produce up to 1.5×10^6 gal/day (5.7×10^6 liters/day) of fuel-grade methanol. Such so-called world-size plants are necessary if methanol is to be economical for fuel use. One such unit using, for example, low-grade lignite, will fuel automobiles for as many miles as the gasoline from 55,000 barrels (8700 m^3) of crude oil a day. *See* LIGNITE.

In producing synthesis gas, the feedstock (for example, the fossil fuel), after any necessary preparation, is gasified to hydrogen, carbon monoxide, and carbon dioxide. The gasifier should operate at as high a pressure as possible to minimize the cost of compression of the synthesis gas produced up to the operating conditions of the converter: 100–300 atm ($1–3 \times 10^7$ pascals) pressure and 390–750°F (200–400°C). Correspondingly large is the air separation plant to supply 5 tons (4.5 metric tons) per minute of oxygen for the partial oxidation of a coal feedstock, and to give the oxygen in the large volume of methanol produced. The steps of air separation, syngas generation, balancing of reaction to produce the desired amount of hydrogen, and the separation of CO_2 and H_2S (then sulfur) from the synthesis gas require a vast complex of massive equipment. Purified synthesis gas passes through catalyst chambers and heat exchangers which, by correct interrelation and proper control of temperatures and space velocities, give 97% methanol, 1% water, and 1–2% higher alcohols. Modern processing (Wentworth) delivers this fuel-grade methanol without any refining being necessary for immediate use as motor fuel and with a higher energy content of the original fuel and a much lower plant cost and operating cost than earlier processes used to make the chemical grade.

Methanol for fuel use has a heating value of about 10,000 Btu/lb (23 megajoules/kg). Fifteen gallons (0.57 m^3) of methanol give 1×10^6 Btu (1×10^9 joules). To make it requires an input of from 1.5 to 2 times as much thermal energy based on the higher heating value of a solid fuel, or somewhat less if it is made from petroleum gas or liquid feedstocks. Thus the overall energy conversion from solid fuels is in the range of 50 to 65%, and the lower efficiencies in this range are for poorer coals or lignites.

Methanol is the most versatile and cheapest liquid fuel which can be made. It is also less flammable than gasoline; accidental fires are extinguished with water instead of being spread as flaming films. When it combusts, it yields neither particulates (soot) nor sulfur oxides, and yields lower quantities of nitrogen oxides than any other fuel. In an accident at sea, there is no fire or oil slick because the methanol dissolves in water. When produced from natural gas or solid fuel, it can always be regasified to give substitute natural gas (SNG) with only a small loss of energy.

Diesel engines in trucks, railroad locomotives, and ships may use a small amount of their regular diesel fuel for start-up. Then methanol may be used for at least 90% of the total fuel.

Steam may be generated in almost any furnace which can be fired by oil or gas, by changing the nozzle which feeds fuel to the combustion chamber so as to use methanol, which gives no emission problems. However, methanol's greatest potential use in electric power plants is by internal combustion in gas turbines, where it is a superior fuel. It may be used most efficiently in a combined cycle with a steam boiler and turbine on the low-temperature side. Compared with a coal-fired, steam-powered plant with its modern accessories, methanol gives a capital cost reduced by almost 50% and a thermal efficiency over a third greater (46% versus 32%).

Methanol has been preferred in automobile engines, for example, in racing cars where efficiency, performance, and safety are more important than cost. With its high octane number (110), 10% can be added to improve the performance of unleaded gasoline, along with other beneficial additives amounting to 30% or more.

Methanol is used much better neat or mixed with just a few percent of common liquids as additives to eliminate problems related to cold start, materials of construction, lubricity, and so on. Lean mixtures may be burned at compression ratios of between 12 and 16 to 1, to give 50 to 80% higher thermal efficiencies than gasoline, even though the heating value of methanol is only about one-half. Even when expensive chemical-grade methanol is used in automobiles with relatively insignificant changes, the mileage costs are less than when gasoline is used as a fuel. When methanol from a world-size fuel-grade plant is used, the fuel cost per mile will be less than half. Pollution from emissions is nonexistent. *See* METHANOL. Donald F. Othmer

Bibliography. D. F. Othmer, *Chem. Eng*. (London), January 1981; J. H. Perry, *Methanol*: *Bridge to a Renewable Energy Future*, 1990.

Alcoholism

The continuous or excessive use of alcohol (ethanol) with associated pathologic results. Alcoholism is characterized by constant or periodic intoxication, although the pattern of consumption varies markedly. Individuals admitted for the first time to an alcoholism treatment center typically have been consuming 3–4 oz (80–100 g) of pure alcohol per day, corresponding to seven to nine drinks or bottles of beer or glasses of wine. Studies have shown that problem drinking in these populations starts at about 2 oz/day (60 g/day), that is, four to five drinks per day, and that these are consumed in rapid succession, leading to intoxication on three or more days per week. Individuals who consume these levels of alcohol have a greater-than-average risk of developing alcoholic liver cirrhosis. However, the levels should not be taken as absolute, since they can vary greatly in different individuals, according to body weight and other factors.

The symptoms and consequences associated with severe alcohol consumption also vary greatly; that is, in some individuals only a few may be present. These may consist of the development of physical dependence manifested as a state of physical discomfort or hyperexcitability (tremors or shakes) that is reduced by continued consumption; the development of tolerance to the effects of alcohol, which leads individuals to increase their consumption; accidents while intoxicated; blackouts, characterized by loss of memory of events while intoxicated; work problems, including dismissal; loss of friends and

family association; marital problems, including divorce; financial losses, including bankruptcy or continual unemployment. Medical problems can include gastric ulcers, pancreatitis, liver disease, and brain atrophy. The last is often associated with cognitive deficiencies, as shown by the inability to comprehend relatively simple instructions or to memorize a series of numbers. *See* COGNITION.

Individuals seeking an early treatment for their alcohol problems have very good probabilities of recovery. The lesser the number of presenting problems described above, the better the chances of favorable outcome, and so an early identification of problem drinking by family, friends, employers, or physicians becomes very important. Employee assistance programs have become an important factor in identification and referral and rehabilitation of individuals with alcohol problems in the United States, Canada, and many other countries. The types of intervention vary greatly, progressing from self-monitoring techniques, to intensive outpatient and inpatient programs, to Alcoholics Anonymous groups.

Absorption of alcohol. Alcohol is absorbed most rapidly from solutions of 15–30% (30–60 proof) and less rapidly from beverages containing below 10% and over 30%. This is to be expected at the lower concentrations, since the rate of absorption depends on the concentration gradient across the mucosal surface. At higher concentrations, ethanol abolishes the rhythmic opening of the pylorus, thus preventing the passage to the intestine, where absorption is faster. The presence of food in the stomach is also known to delay gastric emptying and thus to slow absorption. Once in the bloodstream, alcohol is distributed evenly in all tissues, according to their water content. As a rule of thumb, for a 150-lb (68-kg) individual two standard drinks (1.5 oz of a distilled beverage or 13.6 g per standard drink) will yield blood alcohol levels of about 0.06 g per 100 mL of blood. It was previously believed that all ethanol consumed was absorbed into the bloodstream. However, studies have shown that a small part of the ethanol ingested is degraded directly in the stomach without entering the blood. *See* DISTILLED SPIRITS; MALT BEVERAGE; WINE.

Effects on the nervous system. The exact mechanisms of the pharmacological actions of alcohol are not known. Alcohol can act as a stimulant at lower doses, and as a depressant at higher doses. Even at very low doses, alcohol can impair the sensitivity to odors and taste. Also, low doses are known to alter motor coordination and time and space perception, important aspects of car driving (about 50% of all fatal traffic accidents are caused by intoxicated drivers). Some effects are already seen at levels of 0.05%. Pain sensitivity is diminished with moderate doses. In some individuals, alcohol is known to diminish feelings of self-criticism and to inhibit fear and anxiety, effects which are probably related to an alcohol-induced sociability. These effects act, no doubt, as psychological reinforcers for the use of alcoholic beverages.

It is generally accepted that alcohol affects the nerve cell by preventing the production and propagation of electric impulses along a network consisting of axons and synapses. The brain functions much as an electronic system in which one nerve cell, acting as a current generator, communicates information to many other cells, which in turn receive impulses from many other areas. Some impulses are enhanced, others are blunted. Memory and conditioning appear to play an important role in integrating the impulses which are finally expressed as behaviors. Studies in the United States and England have shown that when alcohol becomes dissolved in the membrane of the cells, it fluidizes or disorganizes the membrane, which in turn leads to changes in the physical and biochemical characteristics of the latter. Chronic exposure to alcohol alters the composition of the membrane and its rigidity, so that alcohol becomes less of a disorganizing agent. The new membrane composition also appears to partly exclude the alcohol molecules from dissolving in it. These changes are likely to be related to tolerance, when more alcohol is required to produce the same effect. Studies have shown that learning and conditioning also play a role in the latter phenomenon. In general, the knowledge gained from these studies is that the state of organization (fluidity–rigidity) of the membrane is correlated with the state of excitability of the brain; the increased rigidity of the membrane that follows chronic alcohol consumption might account for the alcohol withdrawal hyperexcitability (physical dependence) and the fact that continued ingestion of alcohol is required to alleviate the physical dependence. Studies in rodents and monkeys have addressed the question of what factors can lead to alcohol consumption in these animals, which normally reject alcohol. It has been found that small amounts of alcohol-derived molecules, such as the tetrahydroisoquinolines, when injected into the brain, produce a marked increase in the preference for alcohol of the animals. These studies raise the possibility that these substances could be formed in greater amounts in individuals who consume alcohol heavily.

Studies have shown that lipid molecules in the membranes of brain cells of animals fed alcohol for a prolonged time can confer alcohol resistance to brain cell membranes of animals that have never received alcohol. Not only brain lipids are modified but also some brain proteins. Thus, it is clear that brain chemistry is modified by chronic alcohol exposure. This effect remains for days or weeks after alcohol administration has been discontinued. Brain cells communicate with each other and with cells involved in sensation by means of chemical messengers, or neurotransmitters, released by these cells. Once a neurotransmitter is externally recognized by a cell, the information is passed to the interior of the cell by second messengers, a system that involves G-proteins. The continued presence of alcohol makes the nerve cells produce less of these proteins, and as a consequence an external chemical message is not well communicated to the interior of the cell.

Studies using white blood cells of alcoholics as sensors of this deficiency have shown that as much as 75% of the external message is not transmitted to the interior of the cell. These changes are of importance in that they may explain the reduction in cognitive efficiency, in relation to learning and to processing external information, which starts at levels of alcohol consumption that are considered moderate social drinking.

A major finding in the mid-1980s was that some of the effects of alcohol can be quickly reversed by new experimental drugs. Studies have shown that alcohol enhances the actions of an inhibitory brain neurotransmitter referred to as gamma-aminobutyric acid (GABA). Benzodiazepines, such as diazepam, are anxiety-reducing and sedative drugs which also enhance the effects of GABA. These effects can be reduced by experimental antagonist molecules, which interact in the brain in the same regions where GABA is found. *See* CELL MEMBRANES; NERVOUS SYSTEM (VERTEBRATE); SYNAPTIC TRANSMISSION.

Effects on the liver. The liver is responsible for about 80% of the metabolism of alcohol. In the liver, alcohol is first oxidized to acetaldehyde and then to acetate, which is metabolized in many tissues, including the brain, heart, and muscles. A 150-lb (68-kg) person metabolizes approximately 0.4 oz (10 g) of pure alcohol per hour (about 1 oz of a distilled beverage per hour) or, if alcohol is continuously present in the bloodstream, about 8–10 oz (190–240 g) of pure alcohol per day, equivalent to 1300–1600 calories per day. Since alcoholic beverages contain negligible levels of essential nutrients, these calories are called empty calories. Many alcoholics show malnutrition due to the fact that an important part of their caloric intake is alcohol. Alcohol also impairs the absorption and the metabolism of some essential nutrients. *See* ABSORPTION (BIOLOGY); LIVER; MALNUTRITION.

In the presence of alcohol, about 80% of oxygen consumed by the liver is devoted to the metabolism of alcohol; as a consequence, other substances such as fats, normally oxidized by the liver, are not metabolized, leading to fat accumulation in the liver. Continued consumption of alcohol has been reported to increase the capacity of the liver to oxidize alcohol, such that some alcoholics can metabolize up to 16 oz (400 g) of pure alcohol per day. Since all the hepatic metabolism of alcohol, by any known route, requires oxygen, there is also an increased uptake of oxygen by the liver. *See* LIPID METABOLISM.

Alcoholic liver disease and liver cirrhosis rank among the 10 leading causes of mortality in the United States and Canada. It was initially believed that cirrhosis in the alcoholic was exclusively of nutritional origin. However, baboons fed 50% of their caloric intake as alcohol develop cirrhosis in 1–4 years, even when the alcohol is given with a nutritious diet. Baboons fed the same diet, but with carbohydrates given instead of alcohol, did not develop cirrhosis. It is not known, however, if nutrient supplementation would have reduced the damage in its initial stages. Some investigators have proposed that in humans alcoholic liver disease is produced by increased utilization of oxygen by the liver, when the extra utilization is not matched by an adequate supply of oxygen to the organ. Studies in Canada have shown that the antithyroid drug propylthiouracil, which reduces the increased oxygen demand induced by alcohol in the liver, greatly protects against the mortality of alcoholic liver disease. In a 2-year treatment study, deaths in patients treated with propylthiouracil were 50–60% lower than in patients given placebo capsules.

The mortality due to liver disease can be drastically reduced if the individuals abstain from alcohol. Since it is unlikely for liver scar tissue in cirrhosis to reverse to normal, studies suggest that complications accompanying cirrhosis may be more damaging than the cirrhosis itself.

Alcoholic liver disease is characterized by two conditions: failure of the liver to detoxify noxious substances and to produce essential products, and increased resistance to blood flow through the liver. The latter condition (portal hypertension) results in the opening of collateral vessels (esophageal varices) as an alternative for the passage of portal blood. The expanded vessels can burst, leading to internal bleeding. Also, noxious substances bypass the cleansing action of the liver. This condition is responsible for about one-half of the deaths in alcoholic liver disease. Studies have demonstrated that chronic alcohol consumption results in enlargement of the liver cells, which leads to compression of the minute intrahepatic vessels that feed the liver cells. In turn, compression of these vessels increases the resistance to blood passage and to portal hypertension. This effect is quickly reversible upon abstinence from alcohol, and may be responsible for the reduction in mortality from liver disease in abstinent alcoholics. *See* CIRRHOSIS.

Effects on sexual and reproductive functions. Alcohol consumption produces a striking reduction in the hormone testosterone in males. This appears to be due to an effect in brain areas which are responsible for specific signals for the maintenance of normal testosterone levels and to a metabolic effect in the testicular cells that produce testosterone. In addition, alcoholics have increased rates of testosterone destruction by the liver. Feminization signs are often seen in male alcoholics, while impotence and testicular atrophy are also found frequently.

Studies in France and the United States in the early 1970s were the first to report a distinct pattern of physical and behavioral anomalies in children of alcoholic women who drank heavily during pregnancies. These anomalies which characterize the fetal alcohol syndrome include facial and cranial deformities such as drooping eyelids (ptosis), small eyes (microophthalmia), and underdeveloped midface (midfacial hypoplasia). A 10-year follow-up of these children in the 1980s indicated that all were borderline for intelligence, while half of them presented IQs of 20 to 60, in the retarded range. The prevalence of this condition in the general population has not been fully established, but it ranges between 0.5 and 3 births in 1000. In the United States the risk of fetal alcohol

syndrome is about seven times higher in the black race than in the white population. Studies so far do not indicate that there is a safe level of habitual alcohol consumption, below which there is no effect on the unborn; and the role of specific drinking patterns, including a single bout, needs to be determined. In animals with short gestation periods, a single massive intoxication results in damage to the newborn. Social drinking during pregnancy has been associated with lower scores on tests of spoken English and verbal comprehension in the children. *See* FETAL ALCOHOL SYNDROME.

Early identification of alcohol abuse. Several laboratory tests are used to detect high levels of alcohol consumption. Among these are an elevated serum gamma-glutamyl transferase and the presence of enlarged red cells. In addition, rib and vertebral fractures, as seen on routine x-rays, are found 15 times more frequently in alcoholics than in normal populations. Combinations of these tests allow diagnosis of heavy alcohol consumption with a good degree of accuracy. A simple litmus-paper-like test to detect alcohol in saliva, serum, and urine consists of a strip of paper with special reagents which, when impregnated by any of the fluids, indicates the amount of alcohol present by changing color. Another test to determine intoxication involves the eyealyzer, which in 15 seconds senses alcohol vapors emanating from the lacrimal fluid normally covering the eye surface. This noninvasive device is 95% accurate and is not affected by residual alcohol in the mouth, as is the breathalyzer, and it can be used in persons that are unconscious since the eyelid can be held open.

Genetic factors. There is abundant evidence that tendency to alcoholism can be of familial origin, due to environmental, cultural, and genetic factors. A Swedish study demonstrated that identical twins are twice as likely to have a common alcoholic problem as fraternal twins. In an American-Danish study, it was shown that children of alcoholic parents are more likely to develop alcoholism (18%) than children of nonalcoholic parents (4%) when both groups of children were adopted by nonrelatives within 6 weeks of birth.

In a Swedish study, similar findings were observed where male adoptees whose biological fathers were severely alcoholic had a 20% occurrence of alcohol abuse, as compared with 6% for adoptee sons of nonalcoholic fathers. All adoptions in this study had occurred with the first 3 years of life and in most cases in the first few months. A subsequent Swedish-American study revealed the existence of two types of genetic predisposition to alcoholism. The more common type, mileu-limited alcoholism (type I), is associated with mild manifestations and adult-onset alcohol abuse in one of the biological parents. In this group, a significant contribution to alcoholism predisposition was low socioeconomic status. A second type, male-limited alcoholism (type II), accounts for about 25% of all male alcoholics and is unaffected by the environment. This type of alcoholism, often with teenage onset, was associated with severe alcoholism in the biological father but not in the mother. The biological father also tended to have a record of more serious criminality and an earlier onset of alcoholism. Alcohol abuse was nine times more frequent in adoptees of type II alcoholic fathers, regardless of the environment into which they were adopted. Brain electrical activity in response to external stimuli is markedly different in the male children of type II alcoholics, at an age as early as 12 years. It should be noted that about 60–80% of sons of alcoholics do not become alcoholic and thus environmental factors have a marked influence, for example in type I alcoholism. In identical twins, where genetic factors are constant, alcoholism occurs in both individuals in only 25% of the cases.

Studies from Germany and Japan demonstrate that some genetic factors can protect against the development of alcoholism. About one-half of the population of Japan lack a specific aldehyde dehydrogenase to rapidly metabolize the toxic acetaldehyde produced in the liver in the first reaction of alcohol oxidation. As a consequence, blood acetaldehyde levels build up, leading to facial flushing and to unpleasant reactions which may include nausea. These individuals appear not to develop alcoholism, as shown by the finding that an extremely low proportion of individuals lacking this specific aldehyde dehydrogenase seek treatment of alcoholism.

A number of studies also showed that special genetic strains of rodents can be developed which either prefer alcohol to water or reject alcohol. Some strains of alcohol-preferring rats drink to intoxication, develop physical dependence manifested as hyperexcitability upon withdrawal from alcohol, and have been shown to work (conduct specific tasks) to receive alcohol. Other genetic strains of mice can be shown to differ markedly in sensitivity to the effects of ethanol in that after receiving similar doses of ethanol, one strain (LS or long-sleep) sleeps five times longer than another strain (SS or short-sleep). These differences have been found to result from differences in the brain of these animals. *See* BEHAVIOR GENETICS.

Statistics. In the United States there are an estimated 10 million problem drinkers. Official statistics indicate that in 1986 1.2 million people were admitted for alcoholism treatment in private and government facilities. Alcohol-related deaths are estimated at 100,000 per year. Accidents and cirrhosis are the leading causes of mortality associated with excessive alcohol use, followed by hypertensive diseases, infections, and cancer. Eighty percent of all cirrhosis and about one-half of all accidental deaths, suicides, and homicides are alcohol-related. The average per-capita consumption (over 14 years of age) is 2.65 gallons (10 liters) of pure alcohol per year. Heavy drinkers, constituting about 10% of the drinking population, account for one-half of all alcohol consumed in the United States. It is estimated that alcohol creates problems for 18 million persons 18 years old or more. Both the level of consumption and the number of deaths attributed to alcohol declined during the 1980s. Women drink significantly less than men; however, there has been

an increase in drinking among women aged 35 to 64. People over age 65 consume less alcohol than younger adults and have lower prevalence of alcohol abuse.

Studies in the Scandinavian countries and Canada have shown that the price of alcoholic beverages (relative to the disposable income) is inversely related to the total amount of alcohol consumed by the population. Cirrhosis in different countries has been shown to correlate almost perfectly with the per-capita consumption, with the United States showing about one-third of the incidence found in France. Some studies have shown that the frequency of coronary heart disease is lower in countries with higher per-capita consumption of alcohol. However, others suggest that this is largely due to the fact that populations in these countries also have dietary patterns that by themselves lead to less coronary heart disease. *See* ADDICTIVE DISORDERS; BEHAVIORAL TOXICOLOGY. Yedy Israel

Pharmacotherapy. Pharmacotherapy for alcohol rehabilitation has been gaining wider acceptance. Specific pharmacotherapies which have received the most research attention utilize naltrexone, acamprosate, and disulfiram. Other pharmacological interventions for specific indications are SSRIs, (selective serotonin reuptake inhibitors) and buspirone.

Naltrexone. Naltrexone is an opiate receptor antagonist which blocks the effects of endogenous opioids in the brain. Research from animal studies suggests that alcohol activates endogenous opioid systems and may contribute to the pleasurable effects produced by alcohol consumption. Consequently, naltrexone might reduce the reinforcing effects of alcohol consumed by people and decrease their incentive to drink. The efficacy of naltrexone treatment has been examined by several investigators in double-blind, placebo-controlled studies. These studies have demonstrated that naltrexone (50 mg per day) reduces the frequency of alcohol consumption, resumption of heavy drinking, and intensity of alcohol craving with relatively few significant side effects. Relative to placebo controls, relapse rates were reduced significantly when patients took naltrexone. Naltrexone is well tolerated; about 15% of patients suffer from nonserious side effects, primarily nausea. These findings contributed to the 1995 decision by the U.S. Food and Drug Administration (FDA) to approve naltrexone for use in alcoholism treatment.

Acamprosate. The mechanism of action of acamprosate, calcium-acetyl-homotaurinate, involves primarily the restoration of a normal *N*-methyl-D-aspartate (NMDA) receptor tone in glutamatergic systems. It decreases postsynaptic potentials in the neocortex and diminishes voluntary alcohol intake in alcohol-preferring rats. It was proposed that one of acamprosate's actions is suppressing conditioned withdrawal craving by its effects on NMDA receptors and calcium channels. Acamprosate was investigated in 20 controlled clinical trials with about 5000 patients up to early 2006. Acamprosate (1998 mg per day) lengthens time to relapse, reduces drinking days, and increases complete abstinence among alcohol-dependent patients. Adverse events tended to be mild and transient, primarily involving the gastrointestinal tract with diarrhea and abdominal discomfort in about 10% of the patients. In 2004, acamprosate was approved in the United States.

Combined treatment. It was expected that naltrexone would specifically attenuate "reward craving" (anticipating hedonic effects of alcohol), while acamprosate would diminish relief craving (anticipating unpleasant effects associated with the absence of alcohol). The combination of these two drugs would act simultaneously on two different aspects of craving and might have a more incisive effect on the risk of relapse than either treatment alone. Both drugs are well tolerated and have no marked propensity for potentially dangerous drug–drug interactions. Now preclinical and clinical data exist to demonstrate a good tolerability of combined medication and, moreover, provide some hints for a significant improvement in outcome regarding the duration of abstinence following alcohol withdrawal.

Disulfiram. Disulfiram is a drug which causes an inhibition of the enzyme aldehyde dehydrogenase. Aldehyde dehydrogenase breaks down the first metabolite of alcohol, acetaldehyde. After taking disulfiram for 3 or more days, a person who drinks alcohol usually will experience an increase in acetaldehyde blood levels. This rise will produce nausea, vomiting, tachycardia, difficulty in breathing, and changes in blood pressure leading to hypotension. These effects typically last up to 3 h. When the drug is prescribed at the standard dose (250 mg per day), dangerous side effects are minimized, although the side-effect profile of disulfiram must be considered carefully for each patient.

Given its pharmacological effects, disulfiram acts as a deterrent to future alcohol consumption by making intake an aversive experience. Alcohol use cannot be resumed safely until 6–7 days after the last disulfiram dose. Despite the compelling rationale for its potential effectiveness, studies of disulfiram's efficacy are equivocal, providing only modest evidence for reducing drinking frequency without significantly enhancing abstinence rates. While some professionals endorse the effectiveness of disulfiram treatment, others advocate more controlled clinical trials to determine its relative costs and benefits before it is used routinely in alcoholism treatment.

Several new drug therapies for alcoholism rehabilitation are under investigation. Buspirone, a nonbenzodiazepine antianxiety agent, may decrease anxiety symptoms associated with a protracted alcohol withdrawal syndrome, thus reducing alcohol relapse potential. SSRIs also seem to reduce alcohol intake, especially in clinically depressed patients suffering from a co-morbid alcohol dependence. These medications require further investigation to determine their effectiveness as pharmacotherapeutic agent, in the treatment of alcoholism.

Pharmacotherapy approaches for alcohol rehabilitation are beginning to show promise as clinical interventions assisting patients in their efforts to recover from alcoholism. Concurrent behavioral and

supportive psychotherapeutic treatments, self-help program participation, and supervision of prescribed medication regimens may further optimize pharmacotherapy outcomes. Determining the best combination of pharmacotherapy approaches, psychotherapeutic interventions, and patient characteristics remains a challenge to the alcoholism treatment field. Falk Kiefer; Steve Martino

Bibliography. H. Bogleites and B. Kissin (eds.), *The Genetics of Alcoholism*, 1994; Combine Study Group, Testing combined pharmacotherapies and behavioral interventions for alcohol dependence (the COMBINE study): A pilot feasibility study, *Alcohol Clin. Exp. Res.*, 27(7):1123–1131, 2003; J. C. Garbutt et al., Pharmacological treatment of alcohol dependence: A review of the evidence, *JAMA*, 281:1318–1325, 1999; D. B. Goldstein, *Pharmacology of Alcohol*, 1983; D. W. Goodwin, *Alcoholism*, 2d ed., 1994; J. C. Hughes and C. C. Cook, The efficacy of disulfiram: A review of outcome studies, *Addiction*, 92:381–395, 1997; F. Kiefer et al., Comparing and combining naltrexone and acamprosate in relapse prevention of alcoholism: A double-blind, placebo-controlled study, *Arch. Gen. Psychiat.*, 60:92–99, 2003; R. Z. Litten and J. P. Allen, Pharmacotherapies for alcoholism: Promising agents and clinical issues, *Alcoholism Clin. Exp. Res.*, 15:620–633, 1991; C. P. Rivers, *Alcohol and Human Behavior: Theory, Research, and Practice*, 1994; J. R. Volpicelli et al., Effect of naltrexone on alcohol "high" in alcoholics, *Amer. J. Psychiat.*, 152:613–615, 1995.

Alcyonacea

An order of the cnidarian subclass Alcyonaria (Octocorallia). Alcyonacea, the soft corals (see **illus.**), are littoral anthozoans, which form massive or dendriform colonies with yellowish, brown, or olive colors. Most attach to some solid substratum; however, some remain free in sandy or muddy places. The only skeletal structures are small, elongated, spindle-shaped or rodlike, warty sclerites which are scattered over the mesoglea. The colony is supple and leathery. The polyp body is embedded in the coenenchyme, from which retractile anthocodia protrude in *Alcyonium*. In *Xenia* and *Heteroxenia*, anthocodia are nonretractile. The polyp base is protected by many sclerites and is termed a calyx.

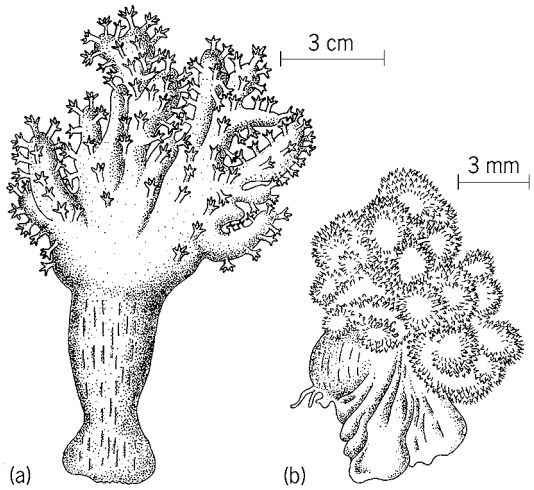

Alcyonaceans. (*a*) *Alcyonium palmatum*. (*b*) *Dendronephytha* sp.

The aciniform gonads develop in the six ventral mesenteries and hang inside the gastrovascular cavity. *Anthomastus*, *Sarcophyton*, and *Lobophytum* are dimorphic and the siphonozooid is fertile only in *Anthomastus*. New polyps arise asexually also, from the solenial network. *See* OCTOCORALLIA (ALCYONARIA); CNIDARIA. Kenji Atoda

Aldebaran

A cool red giant star, prominently located in the constellation Taurus. With its red color, it appropriately represents the eye of the Bull. At a distance of 18.5 parsecs (60 light-years), Aldebaran, or α Tauri, is among the nearest (and brightest) giant stars to the Sun. It appears to be situated near the middle of the Hyades, but Aldebaran is actually less than half the distance of this nearby star cluster. The star is an example of a K-type giant, a very common type of evolved star that derives its energy from the thermonuclear burning of helium in a core surrounded by a thin, hydrogen-burning shell. Its spectral type of K5III corresponds to an effective temperature of 6700°F (4000 K) and a radius of about 40 times the Sun. It is nearly 150 times more luminous than the Sun and, as is typical for K giants, its brightness varies by a modest amount. Aldebaran is accompanied in a long-period binary-star system by a cool dwarf companion star some 100,000 times fainter than the giant. *See* BINARY STAR; GIANT STAR; HYADES; SPECTRAL TYPE; STELLAR EVOLUTION; TAURUS; VARIABLE STAR.

Because it is among the nearest and brightest members of its class, Aldebaran has been studied in considerable detail. A thin and very cool shell of dust surrounds the star out to 500 stellar radii. The dust grains originate in the outer atmosphere of Aldebaran and are then gently blown away by radiation pressure. Aldebaran's angular diameter of about 20 milli-arc-seconds (1/180,000 degree) has been measured by several interferometric methods, including a measurement with a two-telescope interferometer that determined the diameter to an accuracy of less than 1%. Observations from the *Hubble Space Telescope* detected the emission from ionized atoms in the chromosphere of Aldebaran, where the turbulent gases appear to be churning at speeds of more than 12 mi/s (20 km/s). *See* STAR.

Harold A. McAlister

Bibliography. A. Frankoi, D. Morrison, and S. C. Wolff, *Voyages Through the Universe*, 3d ed., Brooks/Cole, 2004; J. B. Kaler, *The Hundred Greatest Stars*, Copernicus Books, 2002; J. M. Pasachoff and A. Filippenko, *Astronomy in the New*

Millennium, 3d ed., Brooks/Cole, 2007; A. Quirrenbach et al., Angular diameter measurements of cool giant stars in strong TiO bands and in the continuum, *Astrophys. J.,* 406:215–219, 1993.

Aldehyde

One of a class of organic chemical compounds represented by the general formula RCHO. Formaldehyde, the simplest aldehyde, has the formula HCHO, where R is hydrogen. For all other aldehydes, R is a hydrocarbon radical which may be substituted with other groups such as halogen or hydroxyl (see **table**). Aldehydes are closely related to ketones, since both types of compounds contain the carbonyl group

$$>C=O$$

In the case of ketones, the carbonyl group is attached to two hydrocarbon radicals. Because of their high chemical reactivity, aldehydes are important intermediates for the manufacture of resins, plasticizers, solvents, dyes, and pharmaceuticals. *See* KETONE.

Physical properties. At room temperature formaldehyde is a colorless gas, which is handled commercially in the form of a 37% aqueous solution or as its polymer, paraformaldehyde $(CH_2O)_x$. The other low-molecular-weight aldehydes are colorless liquids having characteristic, somewhat acrid odors. The unsaturated aldehydes acrolein and crotonaldehyde are powerful lacrimators.

Chemical properties. The important reactions of aldehydes include oxidation, reduction, aldol condensation, Cannizzaro reaction, and reactions with compounds containing nitrogen.

Oxidation. Aldehydes are oxidized to carboxylic acids by a wide variety of oxidizing agents, such as air, oxygen, hydrogen peroxide, permanganate, peracids, silver oxide, and nitric acid, as in reaction (1).

$$RC\overset{O}{\overset{\|}{H}} + (O) \longrightarrow R\overset{O}{\overset{\|}{C}}-OH \qquad (1)$$

The catalytic oxidation of acetaldehyde is used for the production of three industrially important compounds, as in reaction (2), where the product

$$CH_3CHO \xrightarrow[\text{catalyst}]{O_2\ (\text{air})} \underset{\text{Acetic acid}}{CH_3COOH},\ \underset{\text{Acetic anhydride}}{(CH_3CO)_2O},\ \underset{\text{Peracetic acid}}{CH_3\overset{O}{\overset{\|}{C}}OOH} \qquad (2)$$

depends on the conditions of the reaction.

Reduction. Aldehydes are reduced to primary alcohols cheaply by reaction with hydrogen in the presence of a platinum or nickel catalyst, or in the laboratory by treatment with aluminum isopropoxide or oxide or with complex metal hydrides such as lithium aluminum hydride ($LiAlH_4$) or sodium borohydride ($NaBH_4$), as in reaction (3).

$$RCHO + (H) \longrightarrow RCH_2OH \qquad (3)$$

Aldehydes react with hydrogen in the presence of ammonia and a catalyst to form amines. This process [reaction (4)], known as reductive amination, is used

$$\underset{\text{Aldehyde}}{RCH{=}O} + NH_3 + H_2 \xrightarrow{\text{Pt or Ni}} \underset{\text{Amine}}{RCH_2NH_2} + H_2O \qquad (4)$$

commercially to produce many important amines.

Aldol condensation. Two molecules of an aldehyde react in the presence of alkaline, acidic, or amine catalysts to form a hydroxyaldehyde. This reaction proceeds only when the carbon adjacent to the carbonyl group has at least one hydrogen atom [reaction (5)].

$$\underset{\text{Aldehyde}}{RCH_2CHO} + RCH_2CHO \xrightarrow{\text{catalyst}} \underset{\text{Aldol (a hydroxy aldehyde)}}{RCH_2\overset{OH}{\overset{|}{C}}H\underset{R}{\underset{|}{C}}HCHO} \qquad (5)$$

Aldols are readily dehydrated to form unsaturated aldehydes, and the latter can be hydrogenated to form saturated, primary alcohols. This reaction sequence is the basis of the commercial preparation of crotonaldehyde and 1-butanol (*n*-butyl alcohol) from acetaldehyde [reaction (6)].

$$2CH_3CHO \xrightarrow[\text{catalyst}]{\text{alkaline}} \underset{\text{Acetaldol}}{CH_3CHOHCH_2CHO} \xrightarrow{\text{heat}}$$

$$\underset{\text{Crotonaldehyde}}{CH_3CH{=}CH{-}CHO} \xrightarrow{H_2} \underset{\text{1-Butanol}}{CH_3CH_2CH_2OH} \qquad (6)$$

Aldehydes and their formulas

Compound	Formula
Formaldehyde	HCHO
Acetaldehyde	CH_3CHO
Acetaldol	$CH_3CHOHCH_2CHO$
Propionaldehyde	C_2H_5CHO
n-Butyraldyde	$CH_3(CH_2)_2CHO$
Isobutyraldehyde	$(CH_3)_2CHCHO$
Acrolein	$CH_2{=}CHCHO$
Crotonaldehyde	$CH_3CH{=}CHCHO$
Chloral	CCl_3CHO
Chloral hydrate	$CCl_3CH(OH)_2$
Benzaldehyde	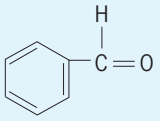
Cinnamaldehyde	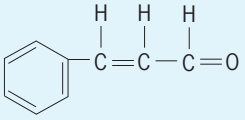

Cannizzaro reaction. On treatment with strong alkali, aldehydes which lack a hydrogen attached to the carbon adjacent to the carbonyl group undergo a mutual oxidation and reduction, yielding one molecule of an alcohol and one molecule of the carboxylic acid salt. The reaction of benzaldehyde with aqueous potassium hydroxide is typical [reaction (7)].

$$2C_6H_5CHO + KOH \xrightarrow{H_2O} C_6H_5CH_2OH + C_6H_5COOK \quad (7)$$

Acetaldehyde and formaldehyde in the presence of calcium hydroxide react by an aldol condensation followed by a "crossed" Cannizzaro reaction to give the industrially useful polyhydroxy compound pentaerythritol [reaction (8)].

$$\underset{\text{Formaldehyde}}{CH_3CHO} + 4CH_2O + Ca(OH)_2 + H_2O \longrightarrow \underset{\text{Pentaerythritol}}{HOCH_2-C(CH_2OH)_2-CH_2OH} + \underset{\text{Calcium formate}}{(HCOO)_2Ca} \quad (8)$$

Reaction with ammonia and amines. The addition of ammonia to the carbonyl group of an aldehyde is reversible, and the resulting aminoalcohol usually cannot be isolated [reaction (9)].

$$RCHO + NH_3 \rightleftharpoons RCH(OH)NH_2 \quad (9)$$

However, aromatic aldehydes such as benzaldehyde react with primary amines such as aniline with elimination of a molecule of water to give stable and well-characterized imines, also known as Schiff bases [reaction (10)].

$$C_6H_5CHO + N_2NC_6H_5 \longrightarrow C_6H_5CH{=}NC_6H_5 + H_2O \quad (10)$$

Oximes, phenylhydrazones, and semicarbazones. Aldehydes react readily with hydroxylamine to eliminate water and form oximes [reaction (11)].

$$\underset{\text{Aldehyde}}{RCHO} + \underset{\text{Hydroxylamine}}{H_2NOH} \longrightarrow \underset{\text{Oxime}}{RCH{=}NOH} + H_2O \quad (11)$$

Phenylhydrazine and semicarbazide react in a similar manner to produce phenylhydrazones and semicarbazones, respectively [reactions (12) and (13)].

$$\underset{\text{Aldehyde}}{RCHO} + \underset{\text{Phenylhydrazine}}{C_6H_5NHNH_2} \longrightarrow \underset{\text{Phenylhydrazone}}{RCH{=}NNHC_6H_5} + H_2O \quad (12)$$

$$\underset{\text{Aldehyde}}{RCHO} + \underset{\text{Semicarbazide}}{H_2NCONHNH_2} \longrightarrow \underset{\text{Semicarbazone}}{RCH{=}NNHCONH_2} + H_2O \quad (13)$$

Since these aldehyde derivatives are usually crystalline compounds with sharp melting points, they are frequently employed for the characterization and recognition of aldehydes. *See* HYDRAZINE; OXIME.

Cyanohydrin formation. Aldehydes combine with hydrogen cyanide to form hydroxynitriles, known as cyanohydrins, as shown in reaction (14) for benzaldehyde. This particular cyanohydrin, known as mandelonitrile, occurs in nature in the form of a glycoside derivative, amygdalin, which is responsible for the characteristic flavor of bitter almonds and the seeds of peaches and apricots.

$$C_6H_5CHO + HCN \longrightarrow C_6H_5CH(OH)CN \quad (14)$$

Hemiacetal-acetal formation. Alcohols react with aldehydes in the presence of acidic catalysts to form hemiacetals [reaction (15)].

$$\underset{\text{Aldehyde}}{RCHO} + \underset{\text{Methanol}}{CH_3OH} \rightleftharpoons \underset{\text{Hemiacetal}}{RCH(OCH_3)(OH)} \quad (15)$$

Hemiacetals may react further with alcohols in the presence of acidic catalysts, but not basic catalysts, to form acetals [reaction (16)].

$$\underset{\text{Hemiacetal}}{RCH(OCH_3)(OH)} + \underset{\text{Methanol}}{CH_3OH} \longrightarrow \underset{\text{Acetal}}{RCH(OCH_3)_2} + H_2O \quad (16)$$

Hemiacetals are generally too unstable to permit isolation, but acetals are very stable under basic conditions, and they are therefore useful as "protecting groups," preventing reaction of the carbonyl group while transformations are carried out elsewhere in the molecule.

Condensation with active methylene compounds. Aldehydes condense with a variety of active methylene compounds in the presence of a base with elimination of a molecule of water to give unsaturated acids or acid derivatives. The reactions of benzaldehyde shown in reactions (17)–(19) are typical.

$$C_6H_5CHO + \underset{\text{Malonic acid}}{CH_2(COOH)_2} \longrightarrow \underset{\text{Benzylidenemalonic acid}}{C_6H_5CH{=}C(COOH)_2} \quad (17)$$

$$C_6H_5CHO + \underset{\text{Acetic anhydride}}{(CH_3CO)O} \longrightarrow \underset{\text{Cinnamic acid}}{C_6H_5CH{=}CHCOOH} \quad (18)$$

$$C_6H_5CHO + \underset{\text{Ethyl acetate}}{CH_3CO_2C_2H_5} \longrightarrow \underset{\text{Ethyl cinnamte}}{C_6H_5CH{=}CHCO_2C_2H_5} \quad (19)$$

Wittig reaction. Aldehydes react with phosphorus ylides to form olefins, as shown in reaction (20). The

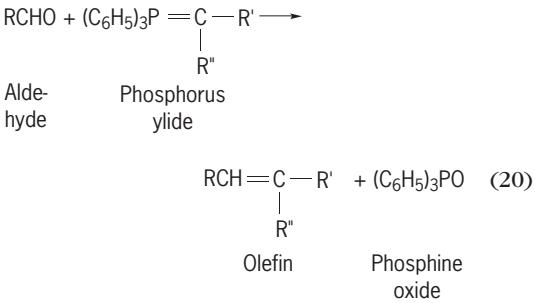

aldehyde may be aliphatic or aromatic. It may contain double or triple bonds and functional groups such as hydroxyl, alkoxyl, nitro, halo, acetal, and ester. Thus the reaction is very general, and the position of the double bond is always certain.

Reformatsky reaction. Aldehydes react with bromo esters and zinc to form hydroxy esters [reaction (21)].

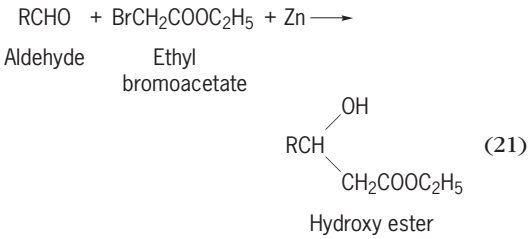

Benzoin condensation. Aromatic aldehydes undergo bimolecular condensation in the presence of alkali cyanide catalyst [reaction (22)].

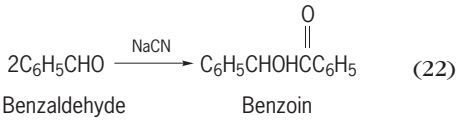

Tests for aldehydes. Tollens' reagent, a solution of silver nitrate in ammoniacal sodium hydroxide, oxidizes aldehydes, with the formation of a mirror of metallic silver. Schiff's reagent (fuchsin aldehyde reagent) reacts with aldehydes to give a violet-purple color. Fehling's solution (a blue, alkaline cupric tartrate complex) is especially valuable for the characterization of reducing sugars. A positive test is indicated by the formation of a red precipitate of cuprous oxide. In the infrared absorption spectrum, the aldehyde group has a characteristic strong band at 1660–1740 cm^{-1} due to the carbonyl stretching vibration accompanied by two weak bands at 2700–2900 cm^{-1} due to the aldehydic carbon-hydrogen stretching. In the proton magnetic resonance spectrum the aldehydic protons have a unique absorption region at 0.0–0.6 τ.

Synthesis. Because of the importance of aldehydes as chemical intermediates, many industrial and laboratory syntheses have been developed. The more important of these methods are illustrated by the following examples.

Catalytic dehydrogenation of primary alcohols. Formaldehyde is produced on a large scale industrially by the catalytic dehydrogenation of methanol. Acetaldehyde is produced similarly by the dehydrogenation of ethanol over a copper catalyst at 250–300°C (480–570°F) with formation of hydrogen as a by-product.

Oxidation of primary alcohols. Acetaldehyde is produced on a large scale industrially by passing a mixture of ethanol, oxygen, and steam over silver gauze at 480°C (896°F) to give an overall yield of 85–90% of the aldehyde [reaction (23)].

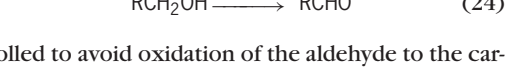

On a laboratory scale, oxidizing agents such as chromic acid or manganese dioxide have been used to produce aldehydes from primary alcohols, as in reaction (24). The reaction must be carefully con-

trolled to avoid oxidation of the aldehyde to the carboxylic acid.

Oxidation of olefins. Ethylene is oxidized directly to acetaldehyde by the Wacker process, employing a palladium–cupric chloride catalyst as shown in reaction (25).

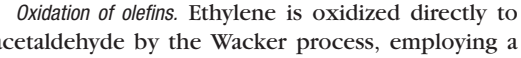

Acrolein, prepared industrially by the oxidation of propylene over a copper oxide catalyst at 350°C (660°F), as in reaction (26), is used in the manufac-

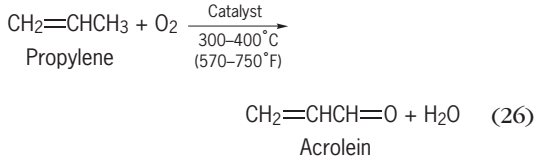

ture of glycerol and acrylic acid. *See* GLYCEROL.

In the laboratory, olefins are treated with ozone to form an ozonide which decomposes with water to form aldehydes and hydrogen peroxide [reaction (27)].

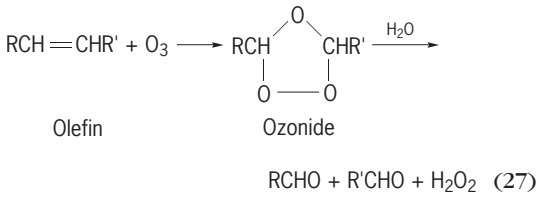

Hydroformylation of olefins. Olefins may be hydroformylated by reaction with carbon monoxide and hydrogen in the presence of a catalyst, usually cobalt carbonyl, at 180°C (360°F) and 6000 lb/in.^{2} (41 megapascals) as shown in reaction (28). This reaction is

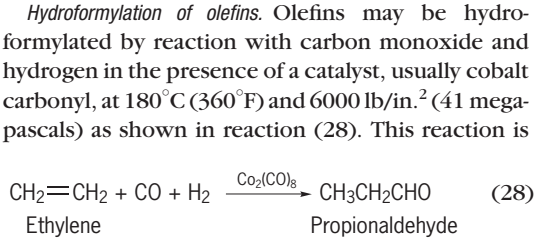

also known as the oxo process.

Chlorination of a methyl group attached to a benzene ring. The chlorination of toluene in the presence of strong light (which serves as a catalyst) proceeds stepwise

to form benzyl chloride and benzal chloride at 135–175°C (275–347°F), by reactions (29) and (30), respectively.

$$\underset{\text{Toluene}}{C_6H_5CH_3} + Cl_2 \longrightarrow \underset{\text{Benzyl chloride}}{C_6H_5CH_2Cl} + HCl \qquad (29)$$

$$C_6H_5CH_2Cl + Cl_2 \longrightarrow \underset{\text{Benzal chloride}}{C_6H_5CHCl_2} + HCl \qquad (30)$$

Hydrolysis of benzal chloride gives benzaldehyde, as in reaction (31). This is one industrial process for

$$C_6H_5CHCl_2 + H_2O \longrightarrow \underset{\text{Benzaldehyde}}{C_6H_5CHO} + 2HCl \qquad (31)$$

the production of benzaldehyde.

Oxidation of a methyl group attached to a benzene ring. Certain substituted benzaldehydes, such as *p*-nitrobenzaldehyde, are readily prepared by oxidation of the corresponding toluene derivatives with chromic acid. Acetic anhydride is used to esterify the aldehyde groups and prevent further oxidation. Subsequent hydrolysis of the diacetate produces the aldehyde [reaction (32)].

$$\underset{\text{p–Nitrotoluene}}{O_2NC_6H_4CH_3} + O_2 + \underset{\text{Acetic anhydride}}{2(CH_3CO)_2O} \longrightarrow$$

$$\underset{\text{Diacetate}}{O_2NC_6H_4CH(OOCCH_3)_2} \xrightarrow{H_2O} \underset{\text{p–Nitrobenzaldehyde}}{O_2NC_6H_4CHO} + \underset{\text{Acetic acid}}{2HOOCH_3} \qquad (32)$$

One commercial synthesis of benzaldehyde involves the passing of a mixture of toluene and air over a heated metallic oxide catalyst at temperatures above 500°C (930°F) [reaction (33)].

$$\underset{\text{Toluene}}{C_6H_5CH_3} + O_2 \xrightarrow{\text{Catalyst}} \underset{\text{Benzaldehyde}}{C_6H_5CHO} + HCl \qquad (33)$$

Reduction of acid chlorides. Acid chlorides react with hydrogen in the presence of a specially prepared palladium catalyst to form aldehydes and hydrogen chloride, as in reaction (34). This reaction, known

$$\underset{\text{Acid chloride}}{RCOCl} + H_2 \xrightarrow{Pd} \underset{\text{Aldehyde}}{RCHO} + HCl \qquad (34)$$

as the Rosenmund synthesis, is a useful laboratory method, since the acid chlorides can be prepared from the corresponding carboxylic acids.

Reaction of Grignard reagents. Aldehydes are formed by the reaction of alkylmagnesium halides (Grignard reagents) with ethyl orthoformate or diethyl formamide [reactions (35) and (36)].

$$RMgX + HC(OCH_2CH_3)_3 \longrightarrow$$

$$RCH(OCH_2CH_3)_2 \xrightarrow{H_3O^-} RCHO \qquad (35)$$

$$RMgX + HC(=O)N(CH_3)_2 \longrightarrow$$

$$RCH(OMgX)N(CH_3)_2 \xrightarrow{H_3O^+} RCHO \qquad (36)$$

Reimer-Tiemann synthesis. When phenols are treated with chloroform and strong alkali, *o*-hydroxybenzaldehydes (salicylaldehydes) are formed [reaction (37)].

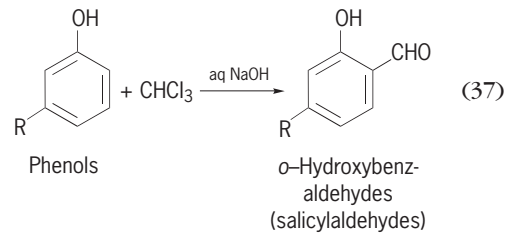

(37)

Furfural production. Cornstalks, corncobs, grain hulls, and similar farm wastes produce furfural upon heating with dilute sulfuric acid. This reaction (38), em-

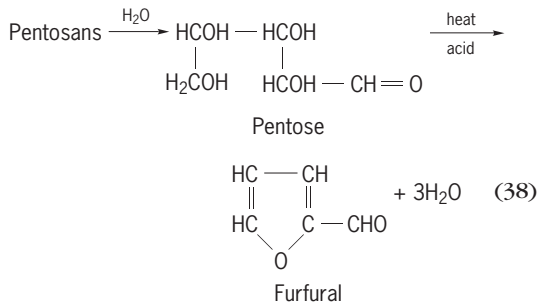

(38)

ployed on a large scale industrially, proceeds from pentosans found in the agricultural residues.

Chloral production. Chloral is an important intermediate for the production of DDT. It is formed by chlorination of anhydrous ethanol [reaction (39)].

$$\underset{\text{Ethanol}}{2CH_3CH_2OH} + \underset{\text{Chlorine}}{8Cl_2} \longrightarrow$$

$$\underset{\text{Chloral}}{2Cl_3CCHO} + \underset{\text{Hydrogen chloride}}{10HCl} \qquad (39)$$

See CARBOHYDRATE; FORMALDEHYDE. Paul E. Fanta

Bibliography. R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; T. W. G. Solomons, C. B. Fryhle, and T. G. Solomons, *Organic Chemistry,* 7th ed., 1999.

Alder

Any of the deciduous shrubs and trees of the genus *Alnus* in the birch family (Betulaceae). There are about 30 species, 10 in the United States, widespread in cool north temperate regions, and southward in the Andes of South America. They have a smooth gray bark, elliptical or ovate saw-toothed leaves in three rows, male flowers in long catkins mostly in early spring, and clusters of several dry hard ellipsoid blackish fruits 0.5–1 in. (1.2–2.5 cm) long. The fruits are conelike and present throughout the year. Alders are common in wet soils, such as stream borders. They often form thickets even beyond treelines and are pioneers on bare areas such as landslides, roadsides, and after fire or logging. As in legumes, the roots bear swellings or nodules containing nitrogen-fixing bacteria which enrich the soil. *See* FAGALES; LEGUME; NITROGEN FIXATION.

Red alder (*Alnus rubra*) fruit and leaf.

Red alder (*Alnus rubra*; see **illus.**) is a small to large tree of the Pacific Northwest along the coast from southeast Alaska south to central California, and local in mountains of northern Idaho. It is the most common and most important hardwood in that region of conifers, occurring in pure stands and mixed with other species. This fast-growing, short-lived tree becomes 100–120 ft (30–36 m) tall on good sites. The wood is reddish or yellowish brown (whitish when freshly cut), moderately lightweight, moderately soft, and fine-textured. Principal uses are pulpwood, furniture, and cabinetwork. *See* FOREST AND FORESTRY; TREE.

Elbert L. Little, Jr.

Aldosterone

The principal mineralocorticoid (a steroid hormone that controls electrolyte and fluid balance) in humans. An adrenocortical steroid, aldosterone was first isolated in 1953 by S. A. Simpson and J. F. Tait in London. This hormone is synthesized from cholesterol in a process involving four different enzymes in the zona glomerulosa (outer zone) of the adrenal cortex. Aldosterone contains an equilibrium mixture of the aldehyde (see **illustration**) and the hemiacetal, the equilibrium favoring the latter. Normal serum concentrations range from 5 to 25 mg/dL. *See* ANDRENAL CORTEX; CHOLESTEROL; HORMONE; STEROID.

Biosynthesis. Direct stimulators of aldosterone biosynthesis include angiotensin II and III, ACTH, alpha-MSH, prolactin, vasopressin, potassium, hydrogen, ammonium, serotonin, histamine, and selected prostaglandins. The two major stimuli for aldosterone secretion are angiotensin II and potassium. Aldosterone synthesis is primarily regulated by the renin-angiotensin system. This system is activated by low sodium balance or decreased blood flow to the kidneys, where renin is produced. Renin is a proteolytic enzyme that converts the inactive angiotensinogen globulin to angiotensin I, which is then converted to angiotensin II by the angiotensin converting enzyme. *See* KIDNEY; THIRST AND SODIUM APPETITE.

Physiologic activity. Aldosterone acts primarily at the renal distal convoluted tubule promoting the excretion of potassium and retention of sodium, thereby influencing extracellular volume homeostasis and blood pressure. *See* POTASSIUM; SODIUM.

Alterations in Na^+/K^+ flux are seen between 30 and 90 minutes after aldosterone administration. This delayed response in activity is a nongenomic effect and represents binding to cytosolic steroid receptors, translocation to the nucleus, interaction with DNA, and finally genomic transcription and translation of effector proteins. Aldosterone stimulates the synthesis of epithelial sodium channels at the apical epithelial membrane of the epithelial cells in the kidney. Aldosterone increases the activity of the Na^+/K^+-ATPase. Mineralocorticoid receptors (the cytosolic steroid receptors) are present in epithelial cells in the kidney, bladder, gastrointestinal tract, sweat and salivary glands, smooth muscle, vascular endothelium, brain, and myocytes (muscle cells).

Increased aldosterone secretion contributes to some forms of arterial hypertension, disorders of sodium retention and edema, and syndromes of potassium wasting (observed when intake of potassium is less than output in the urine). Chronic aldosterone excess in the presence of high salt intake causes cardiac fibrosis in experimental animals. Patients with chronic heart failure had a significant reduction in morbidity and mortality when given a mineralocorticoid receptor antagonist, spironolactone, in addition to traditional therapy.

There are rapid (nongenomic) effects of aldosterone in smooth muscle, skeletal muscle, colonic

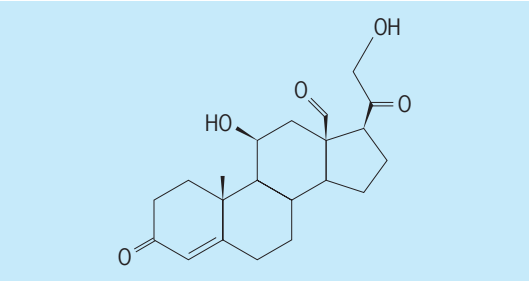

Molecular structure of aldosterone, 11β,21-dihydroxy-3,20-dioxo-pregn-4en-18al.

epithelial cells, and myocardial cells. These nongenomic effects have been linked to the development of increased systemic vascular resistance and therefore might participate in human hypertension and cardiovascular disease. *See* CARDIOVASCULAR SYSTEM; HYPERTENSION. Maria I. New; Alejandro Diaz

Bibliography. J. W. Funder, Aldosterone action, *Annu. Rev. Physiol.*, 55:115–130, 1993; D. N. Orth and W. J. Kovacs, The adrenal cortex, in J. D. Wilson et al. (eds.), *William Textbook of Endocrinology*, 9th ed., W. B. Saunders, Philadelphia, pp. 517–664, 1998; J. S. Willians and G. H. Willians, 50th aniversary of aldosterone, *J. Clin. Endocrinol. Metab.*, 88:2364–2372, 2003.

Alfalfa

The world's most valuable forage legume, *Medicago sativa*, also known also as lucerne, and less often as purple medic, medica, snail clover, median herb, Burgundy hay or clover, Chilean clover, and Burgoens Hoy. It is often referred to as the queen of forages. Alfalfa is produced worldwide on more than 32,000,000 hectares (80,000,000 acres). Seven countries—United States, Russia, Argentina, France, Italy, Canada, and China—account for 87% of the production area (see **table**). *See* LEGUME FORAGES; ROSALES.

Estimated hectarage in selected countries growing alfalfa

Continent and country	Year	Hectarages*
Europe		
Austria	1982	12,630
Bulgaria	1982	399,000
Czechoslovakia	1983	200,000
East Germany	1982	190,000
France	1983	566,000
Greece	1980	198,700
Hungary	1982	337,500
Italy	1982	1,300,000
Poland	1981–1983	258,000
Romania	1981	400,000
Soviet Union (European)	1971	3,375,000
Spain	1981	332,600
Switzerland	1983	6,000
West Germany	1983	31,000
Yugoslavia	1984	337,000
North America		
Canada	1981	2,544,300
Mexico	1982	245,000
United States	1981	10,559,025
South America		
Argentina	1981	7,500,000
Chile	1983	60,000
Peru	1981	120,000
Asia		
Iran	1977	270,000
Soviet Union (Siberian)	1971	1,125,000
Turkey	1969	73,700
Africa		
Republic of South Africa	1985	300,000
Oceania		
Australia	1981–1982	111,500
New Zealand	1984	101,200

*1 hectare = 2.5 acres.

Description. Alfalfa is a herbaceous perennial legume. The seed, 0.04–0.08 in. (1–2 mm) long and wide, and 0.04 in. (1 mm) thick, is kidney-shaped and yellow to olive-green or brown. Plants produce a deep taproot that may have several branches. Two to twenty-five stems are borne from a single crown. Trifoliolate leaves are produced alternately on the stem. The flower is a simple raceme with several complete and perfect florets (**Fig. 1**), which may be various shades of purple, variegated, cream, yellow, and white. Pods are usually spirally coiled. Alfalfa is considered to be cross-pollinated. Under production conditions, pollination is by honeybees (*Apis mellifera*), leafcutter bees (*Megachile*), and alkali bees (*Nomia*), depending on production location.

Origin and domestication. Alfalfa originated in the Near East, in the area extending from Turkey to Iran and north into the Caucasus. Some scientists believe that alfalfa had two distinct centers of origin. The first center was in the mountainous region of Transcaucasia, and Asia Minor and the adjoining areas of northwestern Iran. The second area was in Central Asia. Domestication of alfalfa probably corresponded with the domestication of the horse. The oldest records trace to Turkey between 1400 and 1200 B.C., where alfalfa was fed to domestic animals during the winter and was regarded as a highly nutritious feed. The exact path of movement of alfalfa into the rest of the world is difficult to document. However, maritime trade was well developed in the eastern Mediterranean prior to 1400 B.C. and could have contributed to the spread of the crop. The Romans are known to have had alfalfa as early as the second century B.C. and are given credit for the development of advanced production practices. During this same period of time a Chinese expedition passing into Turkistan to obtain Iranian horses also obtained seed of alfalfa. With the fall of the Roman empire, alfalfa almost disappeared from Europe. During the eighteenth century alfalfa was taken from Europe to New Zealand, Australia, and the New World. There are two primary avenues of introduction of alfalfa into the United States. Spanish missionaries probably brought alfalfa from Chile and Mexico to Texas, New Mexico, Arizona, and California. The Chilean types were nonhardy, and thus climatic conditions prevented their movement into the northern United States and Canada. The second avenue of introduction was into Minnesota in 1857. A immigrant named Wendelin Grimm brought a few pounds of alfalfa seed (*M. media* = *M. sativa* × *M. falcata*) from Germany. Although his initial plantings were not very successful, by harvesting the seed from surviving plants he eventually developed a hardy type that became known as Grimm alfalfa.

Establishment, management, and harvesting. Ideally, alfalfa should be planted in deep, well-drained, medium-textured soil with a pH of about 6.8. Because alfalfa is grown from a single planting for 3 or more years, land preparation is extremely important to successful establishment and production. In areas where irrigation is practiced, special attention must be paid to slope, levee construction, and drainage.

Soils should be able to supply the following kilograms of nutrients per megagram of hay produced (1 kg/Mg = 2 lb/ton): potassium, 40; calcium, 36; magnesium, 7; phosphorus, 6; sulfur, 5; and iron, chlorine, manganese, boron, zinc, copper, molybdenum, less than 0.5. In addition, the crop requires 123 lb (56 kg) of nitrogen, which is largely derived by nitrogen fixation, by *Rhizobium meliloti*. Seed should be inoculated with these bacteria, if there is any question of their presence in the soil.

Planting rates vary between 16 and 45 kg/ha (14 and 40 lb/acre) depending on soil type and planting method. The higher rates are more common on heavy soils and when seeding is done by aircraft. Optimum soil temperatures for germination and seeding development are between 68 and 86°F (20 and 30°C). Dormant cultivars have lower optimum temperatures than nondormant cultivars. Forage should be harvested at one-tenth bloom or with 0.8–1.2 in. (2–3 cm) of regrowth (**Fig. 2**). In northern production areas, one to three harvests are taken annually, whereas in areas such as the Imperial Valley of California up to ten harvests may be taken in a year. *See* NITROGEN CYCLE.

Genetics and plant breeding. Alfalfa is an autotetraploid with a basic genomic number of 8 ($2n = 4x = 32$). However, diploid, hexaploid, and $n = 7$ species are represented in the genus. The frequency of quadrivalents is low, and so the frequency of double reduction is near zero. In the United States, prior to 1940 most of the improvement of alfalfa was by importation of germplasm adapted to production areas. Since the discovery of bacterial wilt disease (caused by *Clavibacter michiganese* ssp. *insidiosum*) considerable effort has been placed on the genetic improvement of the crop, with particular emphasis on the development of multiple pest and disease resistance. Phenotypic recurrent selection is the most common breeding method used for the improvement of alfalfa, but mass selection and various forms of genotypic recurrent selection are used. Cultivars released before 1960 were combinations of one to three germplasm sources, obtained from such areas as India, Chile, Peru, Africa, and Belgium, and primarily developed through public breeding programs. Subsequent to 1960, emphasis in public programs shifted to the development of breeding methods and germplasm. Most cultivars released since 1960 are populations (synthetics), proprietary, and combinations of several germplasm sources. A few commercial hybrids have been developed, and there is currently considerable interest in complementary strain crossing to produce new cultivars. *See* BREEDING (PLANT).

Seed production and standards. Seed production is most efficient in hot and dry environments such as the San Joaquin Valley of California. Seed fields are usually planted on beds at 2–4 kg/ha (1.8–3.6 lb/acre). Certified seed is produced under strict standards, which are monitored by state member organizations of the International Crop Improvement Association. These standards include the geographic area for seed production, the number of generations and number of years of seed increase, isolation, freedom from volunteer plants and weeds, viability, and mechanical purity.

Larry R. Teuber

Fig. 1. Alfalfa stems at flowering stage. (*Iowa State University Photo Service*)

Diseases. Alfalfa is susceptible to at least 75 diseases caused by fungi, bacteria, mycoplasma-like

Fig. 2. Alfalfa of wide crown type at late bud or early bloom stage. (*Iowa State University Photo Service*)

Fig. 3. Damaging effect of bacterial wilt of alfalfa (left), contrasted with healthy plant (right).

organisms, viruses, and nematodes. Most of these occur sporadically, and rarely in epidemic proportions; however, several are responsible for appreciable losses in forage yield and seed production. The best way to reduce disease losses in alfalfa is to grow locally adapted disease-resistant varieties.

Root and crown diseases. Bacterial wilt is one of the best-known and most damaging diseases of alfalfa (**Fig. 3**). The bacteria enter the plants principally through wounds in the roots and through cut ends of newly mowed stems. They are spread in the field by surface water and by tillage and harvesting machinery. Bacterial wilt causes stunting and reduced stand longevity. Resistant varieties have been developed for all alfalfa-growing regions of the United States.

Fig. 4. Fruiting structures of the Verticillium wilt fungus. Conidia (spores) are borne singly at the tips of conidiophores. (*From J. H. Graham et al., A Compendium of Alfalfa Diseases, American Phytopathological Society, 1979*)

Verticillium wilt (**Fig. 4**), first identified in the United States in 1976, has been a damaging disease in Europe since the 1930s. It occurs in the northwestern United States and adjacent areas in Canada. Fusarium wilt, another fungus disease, occurs worldwide but is most severe in warm regions.

Phytophthora root rot, caused by a water mold, occurs worldwide. It is favored by waterlogged soils resulting from excessive irrigation, abundant rainfall, or inadequate soil drainage. Resistant varieties are available.

Other root and crown diseases include Sclerotinia crown rot, active during cool moist weather; Stagonospora root rot and leaf spot; and Fusarium and Rhizoctonia crown and root rots, more active during warm periods. Frequently, crown and root rots cannot be ascribed to a single pathogen, so the term disease complex is used to characterize the pathogens in combinations with various interacting agents (**Fig. 5**). These agents include stress factors such as winter damage, improper cultural and management practices, and root insects. No variety of alfalfa resistant to this complex of disorders has been developed so far.

Stem and leaf diseases. Many microorganisms attack the stems and foliage of alfalfa. Some, such as the fungi that cause anthracnose and spring black stem, spread into the crown and upper part of the taproot, weakening or killing the plant. Anthracnose, favored by warm humid weather, is associated with the summer decline of alfalfa in the eastern United States. Resistance to anthracnose has been incorporated into some varieties; however, many of these varieties are susceptible to a second race of the anthracnose fungus which was discovered in 1978. Spring black stem occurs in most alfalfa-growing areas of the world and is particularly damaging in cooler humid regions. The fungus can attack any part of the plant, but the shiny black streaks on stems and irregularly shaped black spots on leaves are most conspicuous. Summer black stem is another disease of stems and leaves; its reddish to smoky-brown lesions appear when temperatures increase during summer.

Several leaf diseases occur widely and cause losses through defoliation and a general weakening of the plant. Common leaf spot, yellow leaf blotch, Leptosphaerulina leaf spot, and Stemphylium leaf spot may develop so abundantly that they cause serious loss of leaves. Forage quality and yield are lowered. Some varieties have low levels of resistance. Early harvest may help reduce defoliation as well as spread of leaf and stem diseases.

Some less prevalent but occasionally locally important foliar diseases are downy mildew, leaf rust, and bacterial leaf spot.

Virus diseases. The alfalfa mosaic virus (AMV) complex consists of many strains that differ in virulence

Fig. 5. Crown rot of alfalfa. It may be caused by several fungi in combination with stress factors such as winter injury. (*From J. H. Graham et al., A Compendium of Alfalfa Diseases, American Phytopathological Society, 1979*)

and symptoms. The classic symptom is yellow mottling of leaves. Other symptoms are stunting, contortions of leaves and petioles, and, in some instances, death of plants. However, most AMV-infected plants in an alfalfa stand never show symptoms. Exact information on disease losses is lacking. Because the virus is seedborne, seed produced in areas of low virus incidence should be used. Other viruses of alfalfa include alfalfa enation, transient streak, and alfalfa latent viruses. *See* PLANT VIRUSES AND VIROIDS.

Mycoplasmalike diseases. Witches'-broom, long believed to be a leafhopper-transmitted virus, is now thought to be caused by a mycoplasmalike microorganism. The disease is of minor importance in the United States but is widespread in parts of semiarid Australia.

Nematodal diseases. Plant-parasitic nematodes of alfalfa not only damage plants directly but also aid in the transmission and entry of other pathogens. The alfalfa stem nematode is most damaging, especially in the western United States. Infected crown buds and stems are swollen and stunted and may be killed. Reuse of irrigation water is a major pathway for dissemination. Some resistant varieties have been developed. The root-knot nematodes (*Meloidogyne* spp.) and the root-lesion nematodes (*Pratylenchus* spp.) are widespread and attack many crop plants and weeds. There is usually less damage to the alfalfa crop than to more susceptible crops that follow in a rotation.

Noninfectious diseases. Alfalfa frequently shows disease symptoms not associated with pathogenic or parasitic organisms. These noninfectious or abiotic diseases are caused by the lack or excess of plant nutrients, extreme temperatures, air pollutants, pesticides, harvesting machinery, and other factors. Symptoms caused by these agents can be confused with those caused by infectious (biotic) agents. In addition, abiotic agents often influence the severity of diseases caused by biotic agents. *See* PLANT PATHOLOGY. J. H. Graham

Bibliography. D. K. Barnes et al., Alfalfa germplasm in the United States: Genetic vulnerability, use improvement and maintenance, *USDA Tech. Bull.* 1571, 1977; C. H. Hanson (ed.), *Alfalfa Science and Technology*, American Society of Agronomy, 1972; A. A. Hanson, R. R. Hill, and D. K. Barnes (eds.), *Alfalfa and Alfalfa Improvement*, American Society of Agronomy, 1988; D. Smith, *Forage Management in the North*, 1975.

Alfvén waves

Propagating oscillations in electrically conducting fluids or gases in which a magnetic field is present.

Magnetohydrodynamics deals with the effects of magnetic fields on fluids and gases which are efficient conductors of electricity. Molten metals are generally good conductors of electricity, and they exhibit magnetohydrodynamic phenomena. Gases can be efficient conductors of electricity if they become ionized. Ionization can occur at high temperatures or through the ionizing effects of high-energy (usually ultraviolet) photons. A gas which consists of free electrons and ions is called a plasma. Most gases in space are plasmas, and magnetohydrodynamic phenomena are expected to play a fundamental role in the behavior of matter in the cosmos. *See* MAGNETISM; PLASMA (PHYSICS).

Waves are a particularly important aspect of magnetohydrodynamics. They transport energy and momentum from place to place and may, therefore, play essential roles in the heating and acceleration of cosmical and laboratory plasmas. A wave is a propagating oscillation. If waves are present, a given parcel of the fluid undergoes oscillations about an equilibrium position. The parcel oscillates because there are restoring forces which tend to return it to its equilibrium position.

In an ordinary gas, the only restoring force comes from the thermal pressure of the gas. This leads to one wave mode: the sound wave. If a magnetic field is present, there are two additional restoring forces: the tension associated with magnetic field lines, and the pressure associated with the energy density of the magnetic field. These two restoring forces lead to two additional wave modes. Thus there are three magnetohydrodynamic wave modes. However, each restoring force does not necessarily have a unique wave mode associated with it. Put another way, each wave mode can involve more than one restoring force. Thus the usual sound wave, which involves only the thermal pressure, does not appear as a mode in magnetohydrodynamics. *See* SOUND; WAVE MOTION IN FLUIDS.

The three modes have different propagation speeds, and are named fast mode (F), slow mode (S), and intermediate mode (I). The intermediate mode is sometimes called the Alfvén wave, but some

scientists refer to all three magnetohydrodynamic modes as Alfvén waves. The intermediate mode is also called the shear wave. Some scientists give the name magnetosonic mode to the fast mode.

Basic equations. The magnetohydrodynamic wave modes are analyzed by using the magnetohydrodynamic equations for the motion of a conducting fluid in a magnetic field, combined with Maxwell's equations and Ohm's law. *See* MAXWELL'S EQUATIONS.

In studies of waves, the most important of the magnetohydrodynamic equations is called the momentum equation. It expresses how fluid parcels accelerate in response to the various forces acting on the fluid. The essential forces come from the thermal pressure in the fluid, and from the tension and pressure forces associated with the magnetic field. The magnetic field can exert forces on the fluid only when an electric current flows in the fluid; thus magnetohydrodynamic waves can exist only in an electrically conducting fluid. *See* FLUID FLOW.

Two other equations express the conservation of mass of the fluid and the conservation of entropy. However, entropy is conserved only when viscosity, heat conduction, and electrical resistivity are ignored. *See* CONSERVATION OF MASS; ENTROPY.

The fourth equation is Ohm's law, which expresses the relationship between the electric current and the electric field in the fluid. If the fluid has no electrical resistivity, any nonzero electric field would lead to an infinite current, which is physically unacceptable. Thus the electric field "felt" by any fluid parcel must be zero. However, detailed analysis of the motions of electrons and ions reveals that this statement is valid only if the wave frequency is much less than the cyclotron frequency of the plasma ions; thus magnetohydrodynamic waves are very low-frequency waves. *See* OHM'S LAW; PARTICLE ACCELERATOR; RELATIVISTIC ELECTRODYNAMICS.

It is possible to combine Ohm's law with Faraday's law of induction. The resultant equation is called the magnetohydrodynamic induction equation, which is the mathematical statement of the "frozen-in" theorem. This theorem states that magnetic field lines can be thought of as being frozen into the fluid, with the proviso that the fluid is always allowed to slip freely along the field lines. The coupling between the fluid and the magnetic field is somewhat like the motion of a phonograph needle in the grooves of a record: the needle is free to slip along the grooves (the field lines) while otherwise being constrained from crossing the grooves. It is this coupling between the fluid and the magnetic field which makes magnetohydrodynamic waves possible. The oscillating magnetic field lines cause oscillations of the fluid parcels, while the fluid provides a mass loading on the magnetic field lines. This mass loading has the effect of slowing down the waves, so that they propagate at speeds much less than the speed of light (which is the propagation speed of waves in a vacuum). *See* ELECTROMAGNETIC RADIATION; FARADAY'S LAW OF INDUCTION; LIGHT.

The final equation is Ampère's law, which states that an electric current always produces a magnetic field. *See* AMPÈRE'S LAW; BIOT-SAVART LAW.

Thus there are six equations for six unknowns (the pressure, density, and velocity of the fluid; the electric current; and the electric and magnetic fields). Six equations are in principle sufficient to determine six unknowns.

Linearization of equations. Unfortunately, the six equations are too difficult to be of much use because some of them are nonlinear; that is, they contain products of the quantities for which a solution is sought. Nonlinear magnetohydrodynamics is still only in its infancy, and only a few specialized solutions are known. In order to get solvable equations, scientists accept the limitation of dealing with small-amplitude waves and linearize the equations, so that products of the unknowns are removed. Fortunately, much can still be learned from this procedure. The linearization consists of four steps: (1) All quantities are split into two parts: a background part (denoted by the subscript 0) which is constant in space and time, and a wave part (denoted by the prefix δ) which oscillates in space and time; for example, $\mathbf{B} = \mathbf{B}_0 + \delta\mathbf{B}$ and $\mathbf{v} = \mathbf{v}_0 + \delta\mathbf{v}$, where $\mathbf{B}$ denotes the vector magnetic field and $\mathbf{v}$ denotes fluid flow velocity. (2) These definitions are inserted into the six fundamental equations. (3) It is assumed that the waves are of such small amplitude that products of the wave parts can be ignored. (4) It is assumed that there is no background flow. (This last step simplifies the mathematics without loss of generality.)

The resulting equations have solutions which are harmonic in time and space, that is, which vary as $\exp(i\mathbf{k} \cdot \mathbf{r} - i\omega t)$. Here $\mathbf{r}$, is spatial position, ω is (angular) frequency, and $\mathbf{k}$ is called the wave vector. The magnitude of the wave vector represents how rapidly the wave oscillates in space, and its direction represents the direction of propagation. The magnitude of the angular frequency represents how rapidly the wave oscillates in time. Since the equations are linear, more general solutions can be constructed by adding together an arbitrary number of harmonic solutions. *See* HARMONIC MOTION.

Dispersion relation. The linearized equations have solutions only when ω and $\mathbf{k}$ are related to one another in a special way. The equation which expresses this relationship between ω and $\mathbf{k}$ is called the dispersion relation. It is a fundamental relationship in the theory of waves. *See* DISPERSION RELATIONS.

The dispersion relation for magnetohydrodynamic waves is given by Eq. (1). The quantity M_0^2 is given by Eq. (2). The quantity $k_\parallel$ is the component of $\mathbf{k}$

$$(\omega^2 - k_\parallel^2 v_A^2)(k_\perp^2 + M_0^2) = 0 \tag{1}$$

$$M_0^2 = \frac{(k_\parallel^2 c_s^2 - \omega^2)(k_\parallel^2 v_A^2 - \omega^2)}{(c_s^2 + v_A^2)(k_\parallel^2 c_T^2 - \omega^2)} \tag{2}$$

parallel to $\mathbf{B}_0$ while $k_\perp$ is the component perpendicular to $\mathbf{B}_0$. Equation (2) contains several velocities which are fundamental velocities in magnetohydrodynamics. The quantity v_A is called the Alfvén speed. The Alfvén speed is equivalent to the speed at which a wave travels along a taut string, except that the tension of the string is replaced by the tension in

the magnetic field lines. The quantity c_s is the usual speed of sound in a gas, and c_T is called the tube speed, given by Eq. (3).

$$c_T^2 = c_s^2 v_A^2/(c_s^2 + v_A^2) \tag{3}$$

Equation (1) is a cubic equation for ω^2. The three roots represent the three magnetohydrodynamic wave modes. For each of the three values of ω^2, there are two values of ω (one positive and one negative) which correspond to waves propagating in opposite directions.

A particularly useful quantity is the propagation speed, or phase speed, of the wave, given by Eq. (4).

$$v_{\text{ph}} = \omega/k \tag{4}$$

This quantity can be obtained by solving Eq. (1). The three modes are designated fast, intermediate, and slow, according to their phase speeds (**Fig. 1**). The phase speed depends on the angle θ between the wave propagation direction and the direction of $\mathbf{B}_0$. Thus θ satisfies Eq. (5). The graphs of phase speed

$$\tan\theta = k_\perp/k_\parallel \tag{5}$$

change character depending on whether c_s/v_A is less than or greater than 1. *See* PHASE VELOCITY.

All three modes can propagate along the magnetic field, while only the fast mode can propagate across the magnetic field. For propagation along the field ($\theta = 0$), two of the modes have $v_{\text{ph}} = v_A$, while the third has $v_{\text{ph}} = c_s$; the latter mode is essentially a sound wave unaffected by the magnetic field, while the other two modes are intrinsically magnetic. The intermediate mode has $v_{\text{ph}} = v_A \cos\theta$; this mode is always magnetic in character, and thus the sound speed does not affect its phase speed. In all other cases the fast and slow modes have a mixed character, and all three restoring forces combine in a complicated way. The essential properties of the modes are summarized below.

Intermediate mode. The velocity and magnetic field fluctuations in this mode are perpendicular to the plane containing $\mathbf{k}$ and $\mathbf{B}_0$. This means that the motions are pure shears. There is no compression of the plasma. This is why the sound speed does not appear in the dispersion relation. The tension in the magnetic field lines is the only restoring force involved in the propagation of the wave. This mode is therefore closely analogous to the propagation of waves on a string.

The dispersion relation is given by Eq. (6). The

$$\omega^2 = k_\parallel^2 v_A^2 \tag{6}$$

quantity $k_\perp$ does not appear because the motions on neighboring field lines do not communicate with one another. This is a consequence of the fact that the motions are shears, so that neighboring field lines never bump together.

The fluctuating parts of the fluid flow velocity and the magnetic field, $\delta\mathbf{v}$ and $\delta\mathbf{B}$, are closely related (**Fig. 2**) and in fact satisfy Eq. (7). (The symbol sgn means the algebraic sign of the quantity in parentheses.) Equation (7) has been used to identify the presence of the intermediate mode in the solar wind. There, spacecraft-borne instruments have measured the magnetic field and plasma velocity directly. The fluctuations have been found to obey Eq. (7) fairly closely. It appears that intermediate waves are rather copious in the solar wind. Further investigations have revealed that the waves usually propagate away from the Sun. The Sun may thus be an emitter of these waves. However, the waves that are found in the solar wind also exhibit many properties of turbulence; thus, nonlinear effects, such as interactions between different waves, are important. *See* SOLAR WIND; SUN.

$$\delta\mathbf{v}/v_A = -\text{sgn}(\mathbf{k}\cdot\mathbf{B}_0)\delta\mathbf{B}/\mathbf{B}_0 \tag{7}$$

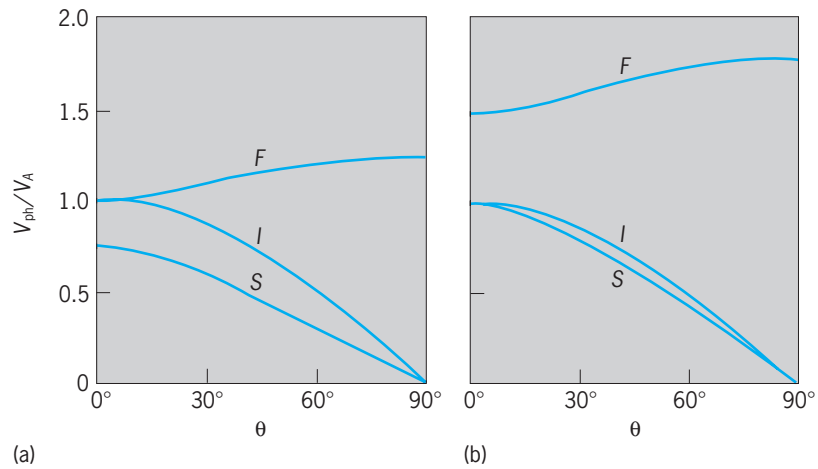

Fig. 1. Phase speed as a function of propagation direction for the fast (F), slow (S), and intermediate (I) modes. The ordinate is phase velocity V_{ph}, normalized to the Alfvén speed V_A. The abscissa is the angle θ between the wave propagation of the direction of the magnetic field B_0. (*a*) $C_s/V_A = 0.75$. (*b*) $C_s/V_A = 1.5$.

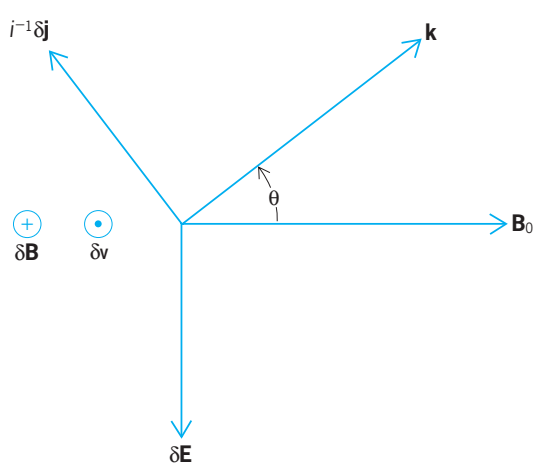

Fig. 2. Vector relationship between the various fluctuating wave quantities for the intermediate mode. Symbols next to $\delta\mathbf{B}$ and $\delta\mathbf{v}$ indicate that these quantities are pointing into and out of the paper, respectively. The electric current $\delta\mathbf{j}$ must be multiplied by i^{-1}(or $1/\sqrt{-1}$) because the current is 90° out of phase with the other quantities.

Waves propagate energy. Since the intermediate mode does not involve the thermal properties of the plasma, its energy flux density is given by the time average of the electromagnetic Poynting flux. For the intermediate mode, the wave energy flux density is always parallel or antiparallel to $\mathbf{B}_0$. Thus the intermediate mode always channels energy along the magnetic field. In fact, the energy flux density of the wave consists of the wave's kinetic energy density plus the wave's magnetic energy density, propagating along the magnetic field at speed v_A. *See* POYNTING'S VECTOR.

Because these waves channel energy along magnetic fields, they may be responsible for the observed fact that cosmical plasmas are strongly heated in the presence of magnetic fields. For example, in the solar wind the wave energy is observed to decrease at greater distances from the Sun, and the amount of wave energy that is lost is about the same as the heat energy that is gained by the solar-wind plasma. Thus, the solar wind is an example of a cosmical plasma that is heated by waves. Waves may also accelerate the solar wind through an effect which is closely analogous to the radiation pressure exerted by light waves. They may also heat the solar atmosphere. Results from the *Solar and Heliospheric Observatory (SOHO)* indicate that protons and heavier ions are strongly heated close to the Sun, in the region where the solar wind is formed. The heating is strongest in the directions perpendicular to $\mathbf{B}_0$. The high-frequency intermediate mode (also called the ion-cyclotron wave) is capable of producing this kind of heating. However, it has not yet been proven that the waves are present at the Sun. It is even less certain whether waves heat the atmospheres of other stars, or other astrophysical objects. Intermediate waves have also been proposed as a means of heating plasmas in controlled thermonuclear fusion devices. *See* NUCLEAR FUSION.

Fast mode. The fast mode is difficult to analyze. However, in many cosmical and laboratory plasmas, v_A^2 is much greater than C_S^2. This is the strong-magnetic-field case. In that case the fast mode is more easily understood.

From Eq. (1), the fast-mode dispersion relation is approximately given by Eq. (8). The direction of $\mathbf{k}$

$$\omega^2/k^2 \approx v_A^2 \tag{8}$$

does not appear in Eq. (8). Thus the waves propagate isotropically.

Fast waves are compressive, and the magnetic field strength fluctuates as well (**Fig. 3**). Thus fast waves are governed by the two restoring forces associated with the tension and pressure in the magnetic field.

Because the magnetic field is strong, the wave energy flux density is mainly due to the time average of the Poynting flux. For this mode, it is evident that the energy flux is along $\mathbf{k}$. Thus the fast mode can propagate energy across the magnetic field.

Slow mode. Like the fast mode, the slow mode is difficult to study in general, and the discussion will again be confined to strong magnetic fields, so that v_A^2 is much greater than C_S^2. The slow mode in a strong field is equivalent to sound waves which are guided along the strong magnetic field lines. The strong magnetic field lines can be thought of as a set of rigid pipes which allow free fluid motion along the pipes, but which restrict motion in the other two directions. The motions on the individual pipes are not coupled together, and thus the slow mode is analogous to the sound waves on a set of independent organ pipes. The slow mode channels energy along the magnetic field. Because the sound speed is small, by assumption, the slow mode transmits energy less effectively than the fast or intermediate modes.

Nonlinear effects. Only small-amplitude waves have been considered. Real waves have finite amplitude, and nonlinear effects can sometimes be important. One such effect is the tendency of waves to steepen, ultimately forming magnetohydrodynamic shock waves and magnetohydrodynamic discontinuities. There is an abundance of magnetohydrodynamic discontinuities in the solar wind. Nonlinear studies have shown these are the result of the nonlinear steepening of the intermediate waves. Nonlinear studies have also shown that the intermediate mode will naturally reach a state where the magnetic field can have large fluctuations of its direction, but maintain a constant field strength. Thus the tip of the magnetic field vector moves on the surface of a sphere. This state of spherical polarization is in fact commonly observed in the solar wind. *See* NONLINEAR ACOUSTICS; SHOCK WAVE.

It is also possible that waves can degenerate into turbulence. There are indications that this too happens in the solar wind. The observed power spectra show that fluctuations are present with a very broad range of frequencies and wave-vector magnitudes, which is a property of turbulence. At greater distances from the Sun, the outward-propagating waves are less dominant; they are apparently getting scrambled because of turbulent processes. *See* TURBULENT FLOW.

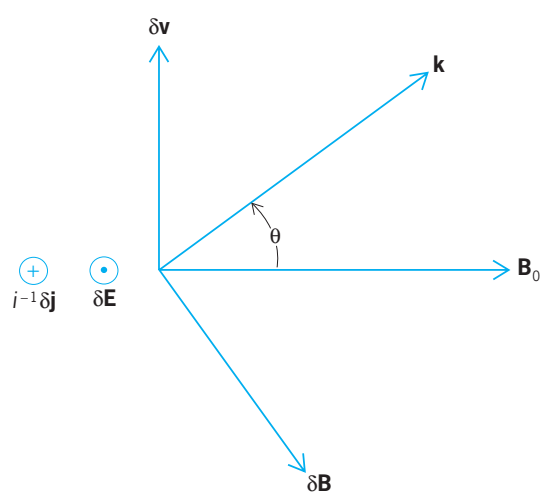

Fig. 3. Vector relationships between the various fluctuating wave quantities for the fast mode in a strong magnetic field, where v_A^2 is much greater than c_A^2.

Dissipation. Viscosity, heat conduction, and electrical resistivity have all been ignored in the above discussion. These effects will all tend to dissipate the wave energy and heat the plasma. *See* CONDUCTION (HEAT); ELECTRICAL RESISTIVITY; SOUND ABSORPTION; VISCOSITY.

In some plasmas, the electrons and ions almost never collide with one another. In such nearly collisionless plasmas, the fast and slow modes can dissipate by the process of Landau damping. Shock formation and turbulence also lead to wave dissipation.

Surface waves. Only waves in a spatially uniform background have been considered. While the analysis of magnetohydrodynamic waves in a nonuniform background is complicated, it is possible to consider an extreme limit, in which the background is uniform except at certain surfaces where it changes discontinuously. Surfaces can support magnetohydrodynamic waves, which are in some respects similar to waves on the surface of a lake. These waves may play important roles in heating cosmical and laboratory plasmas. *See* MAGNETOHYDRODYNAMICS; WAVE MOTION IN LIQUIDS. Joseph V. Hollweg

Bibliography. R. L. Carovillano and J. M. Forbes (eds.), *Solar-Terrestrial Physics*, 1983; F. F. Chen, *Introduction to Plasma Physics*, 1984; J. V. Hollweg, *Computer Phys. Rep.*, 12:205–232, 1990; E. R. Priest, *Solar Magnetohydrodynamics*, 1982; R. Schwenn and E. Marsch (eds.), *Physics of the Inner Heliosphere*, vol. 2, 1991.

Algae

An informal assemblage of predominantly aquatic organisms that carry out oxygen-evolving photosynthesis but lack specialized water-conducting and food-conducting tissues. They may be either prokaryotic (lacking an organized nucleus) and therefore members of the kingdom Monera, or eukaryotic (with an organized nucleus) and therefore members of the kingdom Plantae, constituting with fungi the subkingdom Thallobionta. They differ from the next most advanced group of plants, Bryophyta, by their lack of multicellular sex organs sheathed with sterile cells and by their failure to retain an embryo within the female organ. Many colorless organisms are referable to the algae on the basis of their similarity to photosynthetic forms with respect to structure, life history, cell wall composition, and storage products. The study of algae is called algology (from the Latin *alga*, meaning sea wrack) or phycology (from the Greek *phykos*, seaweed). *See* BRYOPHYTA; PLANT KINGDOM; THALLOBIONTA.

General form and structure. Algae range from unicells 1–2 micrometers in diameter to huge thalli [for example, kelps often 100 ft (30 m) long] with functionally and structurally distinctive tissues and organs. Unicells may be solitary or colonial, attached or free-living, with or without a protective cover, and motile or nonmotile. Colonies may be irregular or with a distinctive pattern, the latter type being flagellate or nonmotile. Multicellular algae form packets, branched or unbranched filaments, sheets one or two cells thick, or complex thalli, some with organs resembling roots, stems, and leaves (as in the brown algal orders Fucales and Laminariales). Coenocytic algae, in which the protoplast is not divided into cells, range from microscopic spheres to thalli 33 ft (10 m) long with a complex structure of intertwined siphons (as in the green algal order Bryopsidales).

Classification. Sixteen major phyletic lines (classes) are distinguished on the basis of differences in pigmentation, storage products, cell wall composition, flagellation of motile cells, and structure of such organelles as the nucleus, chloroplast, pyrenoid, and eyespot. These classes are interrelated to varying degrees, the interrelationships being expressed by the arrangement of classes into divisions (the next-higher category). Among phycologists there is far greater agreement on the number of major phyletic lines than on their arrangement into divisions.

Superkingdom Prokaryotae
- Kingdom Monera
  - Division Cyanophycota (= Cyanophyta, Cyanochloronta)
    - Class Cyanophyceae, blue-green algae
  - Division Prochlorophycota (= Prochlorophyta)
    - Class Prochlorophyceae

Superkingdom Eukaryotae
- Kingdom Plantae
  - Subkingdom Thallobionta
    - Division Rhodophycota (= Rhodophyta, Rhodophycophyta)
      - Class Rhodophyceae, red algae
    - Division Chromophycota (= Chromophyta)
      - Class: Chrysophyceae, golden or golden-brown algae
        - Prymnesiophyceae (= Haptophyceae)
        - Xanthophyceae (= Tribophyceae), yellow-green algae
        - Eustigmatophyceae
        - Bacillariophyceae, diatoms
        - Dinophyceae, dinoflagellates
        - Phaeophyceae, brown algae
        - Raphidophyceae, chloromonads
        - Cryptophyceae, cryptomonads
    - Division Euglenophycota (= Euglenophyta, Euglenophycophyta)
      - Class Euglenophyceae
  - Division Chlorophycota (= Chlorophyta, Chlorophycophyta)
    - Class: Chlorophyceae, green algae
      - Charophyceae, charophytes
      - Prasinophyceae

Placing more taxonomic importance on motility than on photosynthesis, zoologists traditionally have considered flagellate unicellular and colonial algae as protozoa, assigning each phyletic line the rank of order. *See* BACILLARIOPHYCEAE; CHAROPHYCEAE; CHLOROPHYCEAE; CHRYSOPHYCEAE; CRYPTOPHY-

CEAE; CYANOPHYCEAE; DINOPHYCEAE; EUGLENOPHYCEAE; EUKARYOTAE; EUSTIGMATOPHYCEAE; PHAEOPHYCEAE; PRASINOPHYCEAE; PROCHLOROPHYCEAE; PROKARYOTAE; PROTOZOA; PRYMNESIOPHYCEAE; RAPHIDOPHYCEAE; RHODOPHYCEAE; THALLOBIONTA; XANTHOPHYCEAE.

Cell covering. Although some unicellular algae are naked or sheathed by mucilage or scales, most are invested with a covering (wall, pellicle, or lorica) of diverse composition and construction. These coverings consist of at least one layer of polysaccharide (cellulose, alginate, agar, carrageenan, mannan, or xylan), protein, or peptidoglycan that may be impregnated or encrusted with calcium carbonate, iron, manganese, or silica. They are often perforated and externally ornamented (**Fig. 1***a, d, i, j,* and *k*). Diatoms have a complex wall composed almost entirely of silica. In multicellular and coenocytic algae,

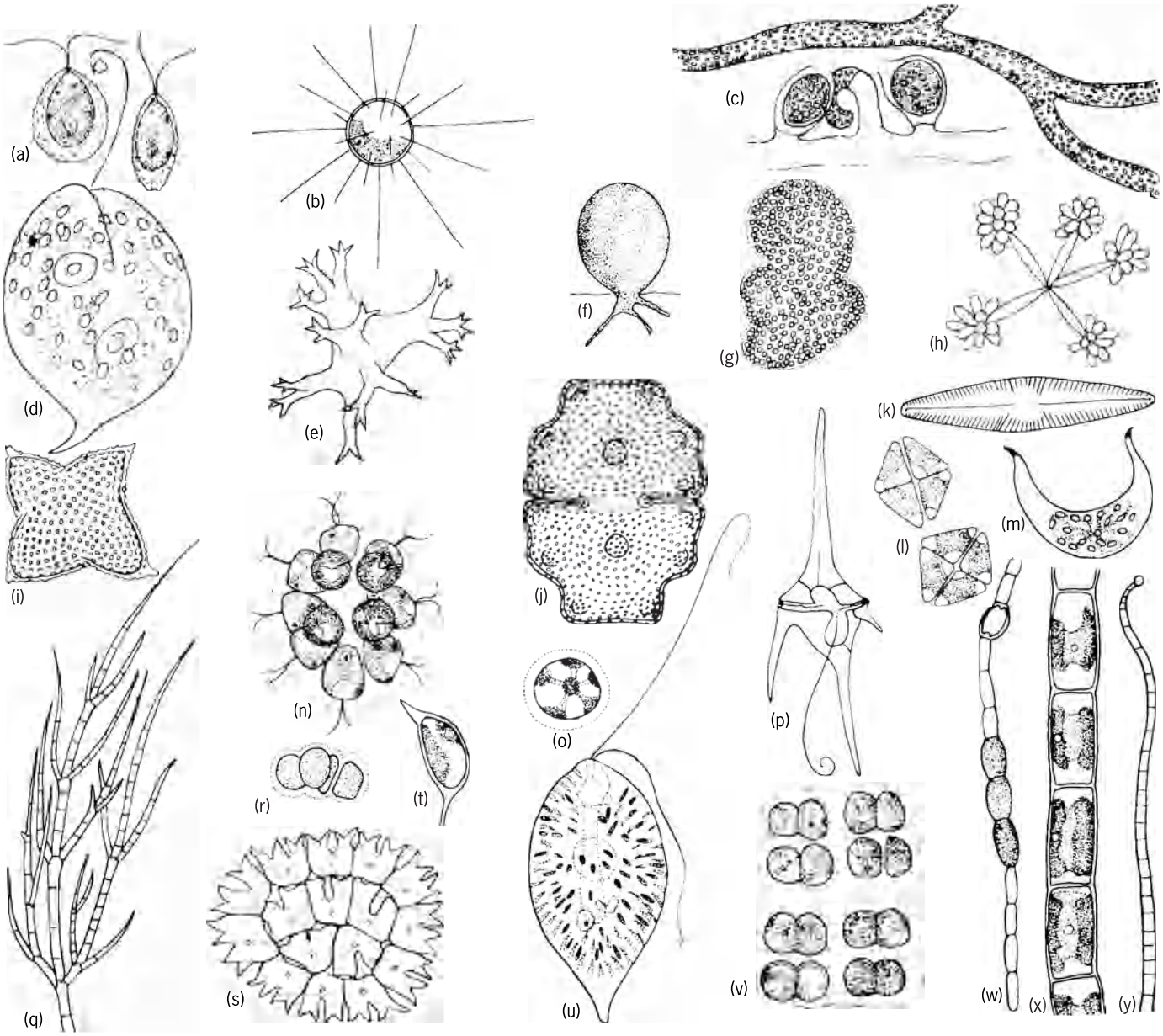

Fig. 1. Representative fresh-water algae. ***(a) Phacotus*** **(Chlorophyceae), motile unicell with a lorica;** ***(b) Golenkinia*** **(Chlorophyceae), nonmotile planktonic unicell;** ***(c) Vaucheria*** **(Xanthophyceae), siphonous coenocyte;** ***(d) Phacus*** **(Euglenophyceae), uniflagellate unicell;** ***(e) Tetraedron*** **(Chlorophyceae), nonmotile planktonic unicell;** ***(f) Botrydium*** **(Xanthophyceae), terrestrial spheroidal coenocyte;** ***(g) Microcystis*** **(Cyanophyceae), coccoid colony;** ***(h) Actidesmium*** **(Chlorophyceae), irregular colony;** ***(i) Tetragoniella*** **(Xanthophyceae), nonmotile unicell;** ***(j) Euastrum*** **(Chlorophyceae), desmid showing semicells;** ***(k) Navicula*** **(Bacillariophyceae), unicell with siliceous wall;** ***(l) Crucigenia*** **(Chlorophyceae), regular nonmotile colony;** ***(m) Cystodinium*** **(Dinophyceae), nonmotile cystlike dinoflagellate:** ***(n) Pandorina*** **(Chlorophyceae), regular motile colony;** ***(o) Asterococcus*** **(Chlorophyceae), unicell in mucilaginous sheath;** ***(p) Ceratium*** **(Dinophyceae), motile dinoflagellate;** ***(q) Stigeoclonium*** **(Chlorophyceae), branched filament;** ***(r) Chroococcus*** **(Cyanophyceae), unicell, solitary or in irregular colony;** ***(s) Pediastrum*** **(Chlorophyceae), regular nonmotile colony;** ***(t) Characium*** **(Chlorophyceae), epiphytic unicell;** ***(u) Gonyostomum*** **(Raphidophyceae), biflagellate unicell;** ***(v) Merismopedia*** **(Cyanophyceae), platelike colony;** ***(w) Wollea*** **(Cyanophyceae), one trichome from a gelatinous colony;** ***(x) Ulothrix*** **(Chlorophyceae), unbranched filament;** ***(y) Oscillatoria*** **(Cyanophyceae), trichome lacking sheath.**

most reproductive cells are naked, but vegetative cells have walls whose composition varies from class to class. *See* CELL WALLS (PLANT).

Cytoplasmic structure. Prokaryotic algae lack membrane-bounded organelles. Eukaryotic algae have an intracellular architecture comparable to that of higher plants but more varied. Among cell structures unique to algae are contractile vacuoles in some freshwater unicells, gas vacuoles in some planktonic blue-green algae, ejectile organelles in dinoflagellates and cryptophytes, and eyespots in motile unicells and reproductive cells of many classes. Nuclei of different taxonomic groups vary in the degree of nuclear breakdown during division, the presence of histone around the deoxyribonucleic acid (DNA), details of spindle formation and elongation, and chromosome behavior. Chromosome numbers vary from $n = 2$ in some red and green algae to $n \geq 300$ in some dinoflagellates. The dinoflagellate nucleus is in some respects intermediate between the chromatin region of prokaryotes and the nucleus of eukaryotes and is termed mesokaryotic. Some algal cells characteristically are multinucleate, while others are uninucleate. Chloroplasts, which always originate by division of preexisting chloroplasts, have the form of plates, ribbons, disks, networks, spirals, or stars and may be positioned centrally or along the cell wall. Photosynthetic membranes (thylakoids) are arranged in distinctive patterns and contain pigments diagnostic of individual classes. Chloroplast DNA may be dispersed throughout the organelle or (in Chromophycota) distributed in a ring. Pyrenoids—proteinaceous bodies around which carbohydrate is stored—occur in the chloroplasts of at least some species of all classes. They have been shown to contain the same carboxylating enzyme that is present in higher plants with C_3 metabolism. *See* CELL (BIOLOGY); CELL PLASTIDS; CHROMOSOME; DEOXYRIBONUCLEIC ACID (DNA); PHOTOSYNTHESIS; PLANT CELL.

Movement. In all classes of algae except Prochlorophyceae, there are cells that are capable of movement. The slow, gliding movement of certain blue-green algae, diatoms, and reproductive cells of red algae presumably results from extracellular secretion of mucilage. Ameboid movement, involving pseudopodia, is found in certain Chrysophyceae and Xanthophyceae. An undulatory or peristaltic movement occurs in some Euglenophyceae.

The fastest movement is produced by flagella, which are borne by unicellular algae and reproductive cells of multicellular algae representing all classes except Cyanophyceae, Prochlorophyceae, and Rhodophyceae. Most motile cells have two flagella, inserted apically or laterally, but some have three, four, eight, or a crown of flagella. Those that appear to be uniflagellate probably have a second, nonemergent flagellum. The basic structural feature of a flagellum—two central microtubules surrounded by nine doublets—is constant throughout the algae (as in all other plants and animals) except in male gametes of diatoms, in which the central pair is absent. Flagella may be decorated with hairs, spines, or scales, the details being of diagnostic value at various taxonomic levels. Details of the flagellar basal region are extremely variable and are considered to be of phylogenetic significance. Biflagellate cells may be isokont (with identical flagella) or heterokont (with morphologically different flagella). Flagellate cells with eyespots are phototactic, moving into areas of optimal light intensity.

Internal movement also occurs in algae in the form of cytoplasmic streaming and light-induced orientation of chloroplasts. *See* CELL MOTILITY; CILIA AND FLAGELLA.

Cell division and growth. Cell division is the process by which unicellular algae multiply and multicellular algae increase in size and complexity. Even within a single class, such as the green algae, details of cell division vary with respect to involvement and interaction of nuclei, microtubules, vesicles, and plasmalemma in wall formation. Random cell division results in diffuse growth, as in a filament or sheet. Thalli with differentiated organs, however, require the localization of cell division in discrete areas or tissues (meristems), which may be apical, basal, marginal, or intercalary. Growth in coenocytic algae entails expansion or elongation of the wall and repeated division of the nucleus and other organelles, but without formation of internal walls. *See* CELL DIVISION.

Reproduction and life histories. Sexual reproduction is unknown in prokaryotic algae and in three classes of eukaryotic unicells (Eustigmatophyceae, Cryptophyceae, and Euglenophyceae), in which the production of new individuals is by binary fission. In sexual reproduction, which is found in all remaining classes, the members of a copulating pair of gametes may be morphologically indistinguishable (isogamous), morphologically distinguishable but with both gametes motile (anisogamous), or differentiated into a motile sperm and a relatively large nonmotile egg (oogamous). Gametes may be formed in undifferentiated cells or in special organs (gametangia), male (antheridia) and female (oogonia). (The highly modified gametangia of red algae are called spermatangia and carpogonia, respectively.) When gametangia are multicellular, every cell becomes fertile, sterile protective cells being absent. Sex hormones (attractants and determinants) have been demonstrated in a few brown and green algae. Sexual reproduction may be replaced or supplemented by asexual reproduction, in which special cells (spores) capable of developing directly into a new alga are formed in undifferentiated cells or in distinctive organs (sporangia). Spores may be motile or nonmotile, and may germinate immediately upon release or may secrete a thick wall to protect themselves during a period unfavorable to germination. Many multicellular algae reproduce vegetatively by fragmentation or by the formation of propagules. *See* REPRODUCTION (PLANT).

Life histories range from binary fission in some unicells to an elaborate sequence of four multicellular somatic phases in many red algae, involving

highly specialized reproductive systems. The different life history patterns are defined in terms of the position of syngamy and meiosis in relation to somatic phases. The three basic types of life history described below may be characterized as having zygotic meiosis, sporic meiosis, and gametic meiosis, respectively. In exceptional cases, meiosis regularly occurs in a somatic tissue, resulting in chimeralike thalli and a complicated life history.

In fresh-water algae, sexual reproduction is generally correlated with the onset of adverse environmental conditions (drying up of the habitat in summer or freezing in winter). The zygote resulting from gametic union secretes a thick, protective wall, transforming itself into a resting stage (zygospore). At the return of favorable growing conditions, the zygote germinates by undergoing meiosis and giving rise usually to four haploid spores, each of which produces a new thallus. The habitat of marine algae, by contrast, undergoes seasonal changes more gradually, and the zygote usually develops immediately into a diploid somatic phase. Adverse conditions of light and temperature are endured, not by thick-walled spores, but by filamentous or crustose stages that receive protection from predation and physical stress by nestling in crevices.

In green and brown marine algae, the diploid somatic phase is autonomous, but in red algae, which have two successive diploid phases, only the second is autonomous, the first remaining an integral part of the parent gamete-producing plant (gametophyte). The autonomous diploid plant (sporophyte) in turn produces spores by meiosis. The haploid spores (meiospores) then give rise directly to unisexual (dioecious) or bisexual (monoecious) gametophytes, which produce gametes to complete the cycle. This life history, which involves an alternation or sequence of somatic phases correlated with changes in ploidy level, is often inappropriately called alternation of generations. The haploid and diploid somatic phases may be similar to one another (isomorphic) or dissimilar in varying degrees (heteromorphic). Extreme cases are presented by kelps, in which the sporophyte often reaches a length of 100 ft (30 m) while the gametophytes are microscopic.

A few algae (Fucales, some Bryopsidales) have a life history in which there is only one somatic phase (a bisexual plant) or two expressions (male and female) of one somatic phase, which are diploid. The formation of gametes entails meiosis, and the zygote develops directly into a new diploid thallus.

Culture studies supplemented by field studies have revealed that some algae are extremely plastic with regard to life history, being able to bypass one or another phase by recycling (repeating), apomeiosis (short-circuiting of sexual reproduction), pedogenesis, parthenogenesis, and vegetative multiplication, and to produce special microscopic (usually filamentous) stages in response to seasonal, long-term, or sporadic changes in the environment. Some species of benthic marine algae have different life histories in different parts of their geographic range, determined probably by temperature and light regimes.

Metabolism. Most algae are autotrophic, obtaining energy and carbon through photosynthesis. All photosynthetic algae liberate oxygen and use chlorophyll *a* as the primary photosynthetic pigment. Secondary (accessory) photosynthetic pigments, which capture light energy and transfer it to chlorophyll *a*, include chlorophyll *b* (Prochlorophyceae, Euglenophyceae, Chlorophycota), chlorophyll *c* (Chromophycota), fucoxanthin among other xanthophylls (Chromophycota), and phycobiliproteins (Cyanophyceae, Rhodophyceae, Cryptophyceae). Other carotenoids, especially β-carotene, protect the photosynthetic pigments from oxidative bleaching. Except for different complements of accessory pigments (resulting in different action spectra), photosynthesis in algae is identical to that in higher plants. Carbon is predominantly fixed through the C_3 pathway. *See* CAROTENOID; CHLOROPHYLL.

Obligately or facultatively heterotrophic algae occur in most classes. These algae are parasites (on other algae, higher plants, and a broad spectrum of animals including humans) or free-living osmotrophs or phagotrophs.

The chemical products of photosynthesis in most algae are stored as one of two polymers of glucose—starch similar to that in higher plants, in which the glucose units have α-1,4-linkages, or laminaran, chrysolaminaran, and paramylon, with β-1,3-linkages. Other storage products (sugars, alcohols, and lipids) are less common, but may be abundant in certain taxonomic groups. Phosphorus in the form of polyphosphate and nitrogen in the form of proteins or peptides are stored by some algae.

The source of carbon for most photosynthetic algae is carbon dioxide (CO_2), but some can use bicarbonate. Many photosynthetic algae are also able to use organic substances (such as hexose sugars and fatty acids) and thus can grow in the dark or in the absence of CO_2. Colorless algae obtain both energy and carbon from a wide variety of organic compounds in a process called oxidative assimilation.

Most algae obtain nitrogen for protein synthesis from nitrate or ammonium, but some unicells can use organic sources such as urea or amino acids. Some blue-green algae—with few exceptions only those with heterocysts—can meet their requirements by fixing atmospheric nitrogen. *See* NITROGEN FIXATION.

Algae, like higher plants, require a certain complement of elements for successful growth and reproduction. In addition to the elements required by higher plants, at least some algae require vanadium, silicon, iodine, or sodium. Unlike higher plants, many algae need an external source of certain vitamins (vitamin B_{12}, thiamin, and biotin, singly or in combination). These vitamins are usually available in natural waters as a result of the metabolic activity of bacteria, but they may be depleted during phytoplankton blooms. *See* PLANT METABOLISM.

Numerous substances are liberated into water by living algae, often with marked ecological effects.

These extracellular products include simple sugars and sugar alcohols, wall polysaccharides, glycolic acid, phenolic substances, and aromatic compounds. Some secreted substances inhibit the growth of other algae and even that of the secreting alga. Some are toxic to fishes and terrestrial animals that drink the water. Nitrogen-fixing blue-green algae release a large proportion of their fixed nitrogen into the water. As part of the plankton tug-of-war, some algae secrete vitamins while others secrete vitamin-inactivating substances. *See* ALLELOPATHY.

Occurrence. Algae are predominantly aquatic, inhabiting fresh, brackish, and marine waters without respect to size or degree of permanence of the habitat. They may be planktonic (free-floating or motile) or benthic (attached). Benthic marine algae are commonly called seaweeds. Substrates include rocks (outcrops, boulders, cobbles, pebbles), plants (including other algae), animals, boat bottoms, piers, debris, and less frequently sand and mud. Some species occur on a wide variety of living organisms, suggesting that the hosts are providing only space. Many species, however, have a restricted range of hosts and have been shown to be (or are suspected of being) at least partially parasitic. Certain algae parasitize fishes while various unicells are characteristic symbionts of protozoa, coelenterates, mollusks, and flatworms. All reef-building corals contain dinoflagellates, without which their calcification ability is greatly reduced. Different phases in a life history may have different substrate preferences. The depth at which algae can grow is determined by the availability of light (water transparency). In the ocean, transparency is affected by latitude (angle of incident sunlight), distance from terrestrial runoff, and physical nature of the bottom. In exceptionally clear areas, such as the Mediterranean and off the east coast of Florida, algae grow at depths of at least 330 ft (100 m). *See* PHYTOPLANKTON.

Many fresh-water algae have become adapted to a nonaquatic habitat, living on moist soil, masonry and wooden structures, and trees. A few parasitize higher plants (expecially in the tropics), producing diseases in such crops as tea, coffee, and citrus. There are several specialized habitats. In deserts, algae (chiefly blue-greens) live on the undersurface and in cracks of rocks. Representatives of several classes (mostly unicells) grow on snow and ice, which they tint green or red depending upon the amount of protective reddish-orange pigment present in the cells. Thermophilic algae (again, chiefly blue-greens) live in hot springs at temperatures up to 163°F (73°C), forming a calcareous deposit known as tufa. *See* TUFA.

One of the most remarkable adaptations of certain algae (blue-greens and greens) is their coevolution with fungi to form a compound organism, the lichen. Most lichens are subaerial, being especially common on the bark and dead twigs of trees and on rocks exposed to moisture-laden winds, but a few are marine. *See* LICHENS.

Geographic distribution. Fresh-water algae, which are distributed by spores or fragments borne by the wind or by birds, tend to be widespread if not cosmopolitan, their distribution being limited by the availability of suitable habitats. Certain species, however, are characteristic of one or another general climatic zone, such as cold-temperate regions or the tropics. Marine algae, which are spread chiefly by water-borne propagules or reproductive cells, often have distinctive geographic patterns. Along temperate shores, individual species tend to have fairly limited ranges, the end points being determined primarily by water temperature. Thus, the marine floras of Japan, Pacific North America, South Africa, southern Australia, and New Zealand all have a high proportion of endemic species. Distributional patterns of marine algae are correlated with patterns of general oceanic circulation, the main currents not only providing a means and determining the direction of transport of propagules but also controlling the temperature. Major floristic provinces include the North Pacific (Japan eastward through Baja California, Mexico, established by the Kuroshio Current System), the North Atlantic (Cape Hatteras, North Carolina, northeastward through Atlantic Europe, established by the Gulf Stream System), the tropical Atlantic (established by the Atlantic Equatorial Current System), the Arctic (related to the Arctic Ocean), the Antarctic/Subantarctic (related to the Antarctic Convergence and the circumpolar current), and the Indo-Pacific (transcending modern current systems and related to ancient positions of land masses). The marine algal flora at any given latitude on the Pacific coast of America is at least 90% different from that at the same latitude on the Atlantic coast. *See* POPULATION DISPERSAL; POPULATION DISPERSION.

Many taxonomic groups are widely distributed, but others are characteristic of particular climatic zones or geographic areas. The large brown seaweeds called kelps are mostly confined to cold waters of the Northern Hemisphere, although certain genera occur only in the Southern Hemisphere. *Sargassum*, also a brown alga, is extremely common in warm waters everywhere. Coenocytic green algae, many of which have calcareous thalli, are characteristic of all tropical areas except the eastern Pacific. Crustose coralline algae are common in all oceans, but individual genera have distribution patterns related to climatic zones and hemispheres. There is a large pantropical floristic element comprising red, brown, and green seaweeds. *See* PLANT GEOGRAPHY.

Vertical distribution. On coasts that are subjected to tidal fluctuation and offer the protection of regular fog or cloud cover, an intertidal belt of seaweeds may be found. When this belt is at least 3 ft (1 m) in breadth, it is often subdivided into horizontal bands that are determined by the ability of particular species to withstand desiccation and wave exposure and by the availability of suitable substrate. Intertidal zonation of seaweeds is correlated with that of invertebrates. Much less is known about patterns of vertical subtidal zonation, where the determining factors appear to be water movement (surge and currents), light intensity, substrate, and predation. Thus, some species grow only at relatively great depths or

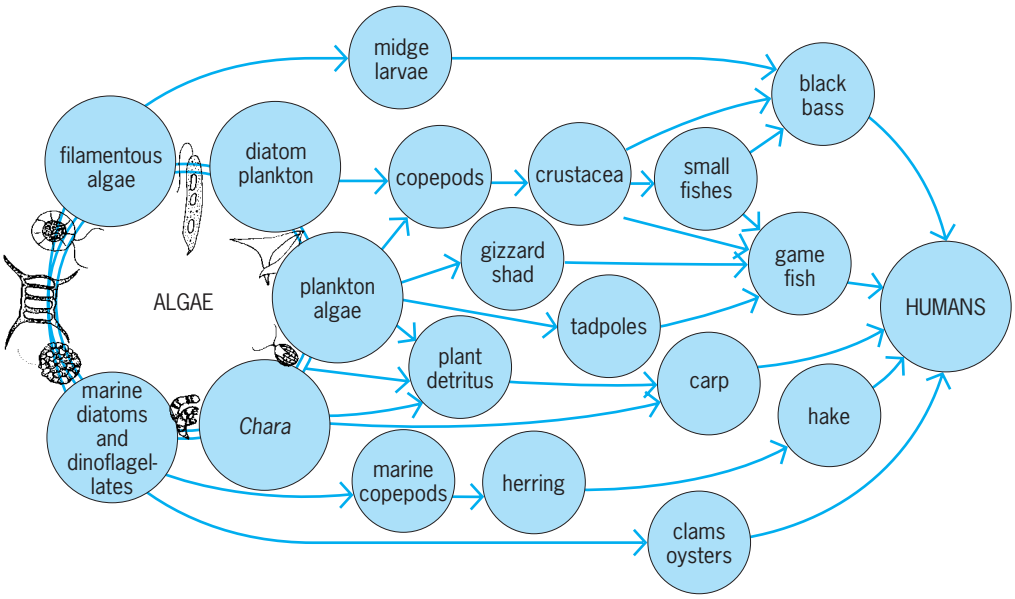

Fig. 2. Food chains from algae to humans. (*After H. C. Sampson et al., eds., Textbook of Botany, 3d ed., Harper, 1966*)

only on vertical rock faces or only on mollusk shells on quiet sandy bottoms. Algae growing on shores of lakes with wave action (such as Lake Michigan) exhibit small-scale vertical zonation.

Economic importance. Planktonic algae, as the primary producers in oceans and lakes, support the entire aquatic trophic pyramid and thus are the basis of the fisheries industry (**Fig. 2**). Concomitantly, their production of oxygen counteracts its uptake in animal respiration. The ability of certain planktonic algae to assimilate organic nutrients makes them important in the treatment of sewage. *See* FOOD WEB; SEWAGE TREATMENT.

Numerous red, brown, and green seaweeds as well as a few species of fresh-water algae are consumed by the peoples of eastern Asia (especially Japan), Indonesia, Polynesia, and the North Atlantic. *Porphyra* (nori), *Laminaria* (kombu), and *Monostroma* (aonori) are produced by intensive mariculture in Japan and China. Large brown seaweeds may be chopped and added to poultry and livestock feed or applied whole as fertilizer for crop plants. The purified cell-wall polysaccharides of brown and red algae (alginate, agar, carrageenan) are used as gelling, suspending, and emulsifying agents in numerous industries, including food processing, cosmetics, pharmaceuticals, textiles, paper, and printing, as well as in dentistry, medicine, and biological research. Some seaweeds have specific medicinal properties, such as effectiveness against worms. The blue-green alga *Spirulina*, a traditional food in parts of Mexico and central Africa, is being grown commercially and marketed as a high-protein dietary supplement. Nitrogen-fixing blue-green algae growing in rice paddies provide significant nitrogen enrichment. Petroleum is generally believed to result from bacterial degradation of organic matter derived primarily from planktonic algae. Diatomaceous earth and diatomite (fossil diatom deposits) have many industrial applications. Unicellular algae, because of the ease with which they can be grown in pure culture and thus provide quantities sufficient for experimentation, have been used traditionally in certain biochemical, biophysical, and physiological studies. *Chlamydomonas, Chlorella, Ankistrodesmus*, and *Euglena* have thus been the subject of thousands of investigations, yielding fundamental biological information. *See* AGAR; ALGINATE.

On the negative side, algae can be a nuisance by imparting tastes and odors to drinking water, clogging filters, and making swimming pools, lakes, and beaches unattractive. Sudden growths (blooms) of planktonic algae can produce toxins of varying potency. In small bodies of fresh water, the toxin (usually from blue-green algae) can kill fishes and livestock that drink the water. In the ocean, toxins produced by dinoflagellate blooms (red tides) can kill fishes and render shellfish poisonous to humans. Ciguatera poisoning in tropical areas has been traced to toxin produced by benthic dinoflagellates and accumulated in reef-dwelling fishes. Aquarium fishes may die from an infection by a parasitic dinoflagellate, while humans may suffer from prototheco-sis, a gastrointestinal disease caused by a colorless chlorophycean unicell (*Prototheca*). In warm humid regions, leaves of crop plants (tea, coffee, citrus) may be badly damaged by a parasitic green alga. A few marine algae have spread far beyond their original home and become noxious weeds, damaging shellfish plantings, clogging the propellers of small boats, and contributing a large biomass to odoriferous beach wrack. *See* EUTROPHICATION; TOXIN.

Fossil algae. At least half of the classes of algae are represented in the fossil record, usually abundantly, in the form of siliceous, calcareous, or organic remains, impressions, or indications. Blue-green algae were among the first inhabitants of the Earth, appearing in rocks at least as old as 2.3 billion years.

Their predominance in shallow Precambrian seas is indicated by the extensive development of stromatolites, laminated reeflike structures produced by algal mats that entrap detrital sediments and sometimes deposit calcium carbonate. Modern stromatolites (as at Shark Bay, Western Australia) implicate blue-green algae, although bacteria may have been at least partly responsible for some of the ancient formations. *See* STROMATOLITE.

All three classes of seaweeds (reds, browns, and greens) were well established by the close of the Precambrian, 600 million years ago (mya). Shallow-water reef deposits were formed in various geological periods by calcareous red and green seaweeds, and the process continues today. The chlorophycean order Dasycladales, in which the entire thallus may be calcified, has left a rich record of more than 150 genera dating from the Precambrian. The calcified zygospores (gyrogonites) of Charophyceae are found in fresh-water deposits as old as the Devonian (400 mya). The relatively low diversity among living dasyclads and charophytes clearly indicates that they are relicts. Among red algae, coralline algae date from the Devonian, while two frequently encountered groups (Solenoporaceae and Gymnocodiaceae) are known only as fossils. *See* GEOLOGIC TIME SCALE.

By far the greatest number of fossil taxa belong to classes whose members are wholly or in large part planktonic. Siliceous frustules of diatoms and endoskeletons of silicoflagellates, calcareous scales of coccolithophorids, and highly resistant organic cysts of dinoflagellates contribute slowly but steadily to sediments blanketing ocean floors, as they have for tens of millions of years. Cores obtained in the Deep Sea Drilling Project have revealed an astounding chronology of the appearance, rise, decline, and extinction of a succession of species and genera. From this chronology, much can be deduced about the climate, hydrography, and ecology of particular geological periods. More than 350 genera of dinoflagellates and an even greater number of genera of coccolithophorids (Prymnesiophyceae) have been described from strata younger than the Permian (about 250 mya). Centric diatoms appear in the Cretaceous (about 100 mya), with pennate diatoms apparently evolving in the Paleocene (65 mya). Because of their abundance, planktonic algae have been important rock builders in shallow seas. Mesozoic coccolithophorids resulted in the accumulation of extensive chalks and limestones (such as the white cliffs of Dover) and there are massive Tertiary deposits of diatoms (diatomaceous earth or diatomite). *See* PALEOBOTANY.

Paul C. Silva; Richard L. Moe

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae*: *Structure and Reproduction*, 1978; V. J. Chapman, *Seaweeds and Their Uses*, 1970; J. D. Dodge, *The Fine Structure of Algal Cells*, 1973; E. Gantt (ed.), *Handbook of Phycological Methods*: *Developmental and Cytological Methods*, 1980; J. A. Hellebust and J. S. Craigie (ed.), *Handbook of Phycological Methods*: *Physiological and Biochemical Methods*, 1978; C. S. Lobban and W. J. Wynne (ed.), *The Biology of Seaweeds*, 1981; F. E. Round, *The Ecology of Algae*, 1981; J. R. Stein (ed.), *Handbook of Phycological Methods*: *Culture Methods and Growth Measurements*, 1973; W. D. P. Stewart (ed.), *Algal Physiology and Biochemistry*, 1974; H. Tappan, *The Paleobiology of Plant Protists*, 1980; O. R. Zaborsky (ed.), *CRC Handbook of Biosolar Resources*, vol. 1, pt. 1, 1982.

Algebra

The branch of mathematics dealing with the solution of equations. These equations involve unknowns, or variables, along with fixed numbers from a specified system. The origins of algebra were based on the need to develop equations that modeled real-world problems. From this came a very extensive theory based on the need to find the values that can be successfully used in the equations.

Number systems. Classical algebra is conducted in one of several number systems. The most basic is the natural numbers, consisting of the counting numbers: 1, 2, 3, 4. . . . The natural numbers along with their negatives and 0 form the set of integers: . . . , −2, −1, 0, 1, 2,. . . . Any number which can be expressed as a quotient a/b of two integers a and b is called a rational number. If the decimal expansion of a number either is repeating or terminates in a string of zeros, it is always possible to find integers a and b such that the quotient a/b gives the original number. Numbers where the decimal expansion cannot be written as a quotient of integers are the so-called irrational numbers. Early mathematicians did not recognize the existence of such numbers until geometric considerations showed that they must exist as the lengths of sides of right triangles.

The sets of rational numbers and irrational numbers have no elements in common, and together they make up the real-number system. It is within the system of real numbers that algebraic tasks are most often performed. However, there is a still larger set of numbers, known as the complex numbers, which serve much more completely in solving equations. The complex numbers consist of numbers of the form $a + bi$, where i is a symbol representing $\sqrt{-1}$. The symbol i is commonly called the imaginary root of −1, and the numbers of the form bi are known as the set of imaginary numbers. In a complex number $a + bi$, the real number a is called the real part while the real number b is the imaginary part. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

Real numbers are often represented on a number line (**Fig. 1***a*). In this method of depiction, there is a one-to-one correspondence between the points on the line and the real numbers. There is no such representation of the complex numbers since the number line implies an ordering notion which does not easily extend beyond the real numbers. However, the complex numbers can be represented on a

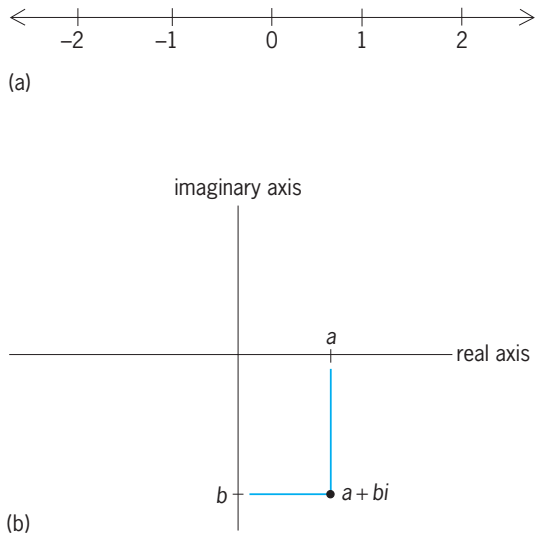

Fig. 1. Representations of number systems. (*a*) Of real numbers on a number line. (*b*) Of complex numbers on a two-dimensional plane.

two-dimensional plane (Fig. 1*b*), by plotting the real part on the horizontal axis and the imaginary part on the vertical axis.

Operations. Within any number system there is a set of operations which can be performed on some or all of the members of the system. Operations take a pair of numbers in the system and produce a single new number from this pair. The most standard operations are addition (+), multiplication (·), subtraction (−), and division (÷). A set is closed under an operation if the application of the operation to any pair of numbers in the set results in another number in the same set. For example, the set of natural numbers is closed under + but not under − (since, for example, $1 - 2 = -1$ is not a natural number). The rational numbers, the real numbers, and the complex numbers are closed under all four of the basic operations (except in the case where division by 0 may occur). *See* ARITHMETIC.

Other operations arise from the four basic ones. Exponentiation is derived from the process of repeatedly multiplying a number by itself. If n is a positive integer, the symbol a^n indicates that the number a is multiplied by itself n times. If $n > 0$ is an integer, then a^{-n} refers to the quotient $1/a^n$. Fractional exponents can also be considered by defining $a^{1/n}$ to be the number b such that $b^n = a$. This notion is further extended by letting $a^{m/n} = (a^{1/n})^m$. For real numbers, some fractional exponents may not exist. For example, $(-1)^{1/2}$ is not a real number. The complex numbers, however, are closed with respect to exponentiation.

Algebraic expressions. In algebra, symbols are frequently used to designate unknown values. For example, letters such as x and y can stand for any one of a set of numbers satisfying certain conditions. These symbols are combined with numbers and the basic operations to form algebraic expressions such as (1).

$$3x^2 + 5y^3x - 6 \qquad \frac{(x^5 - 1)^{3xy}}{y^2 + 2^x} \tag{1}$$

The symbols are usually called variables, while the numbers (or constants) multiplied by them are the coefficients of the expression. The coefficients are all assumed to come from a designated number system. Algebraic expressions take on specific values when each of the variables involved is assigned a numerical value.

Algebraic equations. A statement is a relationship between several algebraic expressions. This relationship can be equality (=) or an inequality (< or >), but in any case it puts restrictions on the values of the variables. For example, each of expressions (2)

$$x^2 + 3xy = y - 3 \qquad 2xy > y^5 - 1 \tag{2}$$

represents a relationship between the variables x and y, while Eq. (3) relates the three variables, x, y, and z.

$$4z - 2xy = 5 + yz^2 \tag{3}$$

If, when specific values from an appropriate number system are substituted for the variables in an equation, the resulting statement is true, then these numbers are said to be a solution to the equation. The solutions of a particular equation are dependent on the number system in use. For instance, the equation $x^4 = 2$ has no solutions in the set of rational numbers, two solutions ($x = 2^{1/4}$ and $x = -2^{1/4}$) in the real numbers, and four solutions in the complex field.

Algebraic equations are used to describe many phenomena in the real world. Their application runs through almost all disciplines in the sciences, social sciences, and business. Once an equation is established as a good model of some situation, it is necessary to determine the solutions to the equation, as these represent instances when the conditions described by the equation are actually met.

One of the most basic kinds of equations is the linear equation. If such an equation involves the variables $x_0, x_1, \ldots, x_n$, then it can be written in the form of Eq. (4), where $a_0, a_1, \ldots, a_n, b$ are coefficients

$$a_0x_0 + a_1x_1 + \cdots + a_nx_n = b \tag{4}$$

from some number system.

In the simplest form, where only two variables, x and y, are involved, a linear equation represents a straight line in two dimensions. The equation of such a line can be written in the form $y = mx + b$, where m is the slope of the line (the rate of change of the y variable with respect to the x variable) and b is the y intercept, or the y value corresponding to $x = 0$.

Another important algebraic equation is the quadratic equation, which has the standard form $y = ax^2 + bx + c$ for coefficients a, b, c. The values of x that produce the value $y = 0$ are found through the quadratic formula, Eq. (5). These x values are known

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{5}$$

as roots. In the complex number system, there are always either one or two roots to a quadratic equation. However, in the real number system, the number of roots of a particular quadratic equation is

given by the discriminant, $b^2 - 4ac$. There are two real roots to a quadratic equation if the discriminant is positive, one real root if it is 0, and no real roots if it is negative. For example, the equation $y = 3x^2 - 2x + 1$ has the discriminant given by Eq. (6), and

$$b^2 - 4ac = (-2)^2 - 4(3)(1) = -8 \qquad (6)$$

hence has no real roots. The complex roots are x_1 and x_2 given by Eqs. (7). The ability to always find

$$x_1 = \frac{-(-2) + \sqrt{(-2)^2 - 4(3)(1)}}{2(3)} = \frac{1 + \sqrt{2}i}{3} \qquad (7a)$$

$$x_2 = \frac{-(-2) - \sqrt{(-2)^2 - 4(3)(1)}}{2(3)} = \frac{1 - \sqrt{2}i}{3} \qquad (7b)$$

roots of quadratic equations (as well as much more complicated ones) in the complex numbers is an indicator of the usefulness of this number system.

Functions. Some of the most fundamental algebraic equations are those involving two variables. When one of these variables, say y, can be expressed in terms of the other variable, say x, in such a way that each value of x produces exactly one value of y, then the relationship described by the equation is known as a function. More general relationships which may or may not produce unique values of y from each value of x are called relations. In a function such as has been described, the variable x is called the independent variable while the resulting variable y is called the dependent variable. The possible values that can be used by the independent variable make up the domain of the function (or, more generally, the relation). The resulting values of the dependent variable compose the range of the function.

It is most common and quite practical to write a variable y which is dependent on a variable x in the form $y = f(x)$. This notation, known as function notation, indicates that every y value is determined by an x value. An equation such as $y = x^2 + 3$ would be written in functional notation as $f(x) = x^2 + 3$. If $x = a$, then the corresponding value of y is given by $y = f(a)$. In the particular function $f(x) = x^2 + 3$, the value of y associated with $x = -1$ is $y = f(-1) = (-1)^2 + 3 = 4$.

Suppose an equation involves the variables x and y. It is said that the ordered pair (a, b) is part of the function (or relation) relating x and y whenever the value $x = a$ produces the value $y = b$. The ordered pair $(-1, 4)$ is part of the function $f(x) = x^2 + 3$, since $f(-1) = 4$. Equations involving three variables, say x, y, z, produce ordered triples (a, b, c) describing corresponding values of the variables. This process can readily be extended so that an equation with n variables would give rise to a set of ordered "n-tuples" where the terms for the corresponding variables can be substituted successfully into the equation.

Graphs. Relations, and more specifically functions, are frequently represented effectively with graphs. If an equation involves the variables x and y, then a two-dimensional coordinate system can be used to depict the values that correspond in the equation. If

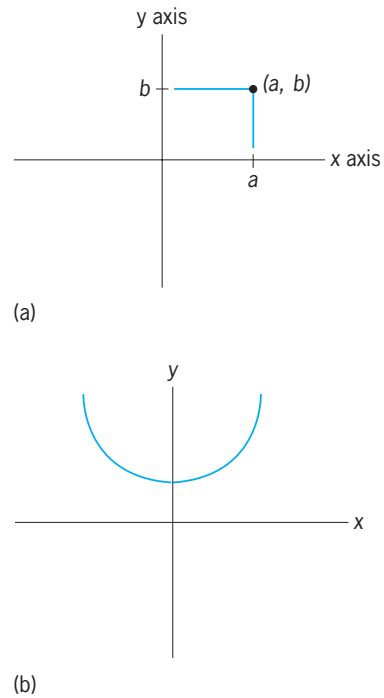

Fig. 2. Representation of relations by graphs. (*a*) Graphic depiction of an ordered pair (*a, b*) given by an equation. (*b*) Graph of the function $f(x) = x^2 + 3$.

the ordered pair (a, b) is given by the equation, then it can be graphically depicted (**Fig. 2***a*). All of the ordered pairs given by the equation combine to give a complete graphic representation of the relationship described by the equation (Fig. 2*b*).

Functions are readily distinguished among more general relations by their graphical representations. Since each x value from the domain of a function produces only one y value in the range, and since all of the points on a vertical line have the same x value, it must be the case that each vertical line intersects the graph of a function in at most one place. For example, from the vertical line test it is readily seen that the relation described in Eq. (8) is a function

$$y = 2x^2 - 4x + 5 \qquad (8)$$

(**Fig. 3***a*) while the relation described by Eq. (9) is not (Fig. 3*b*).

$$y^2 = x - 1 \qquad (9)$$

One of the most important properties of a function is given by the places where the dependent variable takes on the values 0. These correspond to the points where the graph crosses, or touches, the horizontal or x axis. The values of x for which $f(x) = 0$ are called zeros (or roots) of the function. *See* ANALYTIC GEOMETRY; GRAPHIC METHODS.

Polynomials. There are many special functions which are commonly used to describe real situations. One of the most basic is the polynomial, which is a function of the form (10). Each coefficient a_0,

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \qquad (10)$$

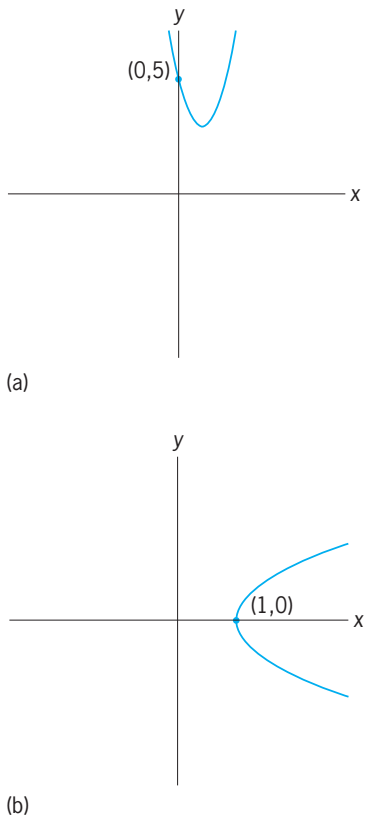

Fig. 3. Graphs of the relations described by the equations (*a*) $y = 2x^2 - 4x + 5$ and (*b*) $y^2 = x - 1$. The vertical-line test shows that the first relation is a function while the second is not.

$a_1, \ldots, a_n$ is a real number, and the exponents of the variables are nonnegative integers. The largest exponent of x is called the degree of the polynomial. When graphed, a polynomial of degree n will reverse directions no more than $n - 1$ times. For example, the graph of the degree-4 polynomial in Eq. (11) [**Fig. 4**] changes direction three times.

$$f(x) = 3x^4 - 4x^3 - 12x^2 + 8 \qquad (11)$$

A polynomial of degree n will have no more than n roots in the real number system. Sometimes it may have fewer. For example, the function $f(x) = x^2 + 1$ has no real roots, while the function $f(x) = x^2 - 1$ has the roots $x = -1$ and $x = 1$. The function $f(x) = x^2$ has the single root $x = 0$. Whenever $x = a$ is a root of a polynomial $f(x)$ of degree n, then the polynomial can be factored as $f(x) = (x - a)\, g\,(x)$, where $g\,(x)$ is a polynomial of degree $n - 1$. If the term $(x - a)$ can be successively factored from a polynomial k times, then $x = a$ is said to be a root of multiplicity k. The equation $y = (x - 1)(x - 1)$ has only one root, $x = 1$, but this root is of multiplicity 2.

It is in the consideration of roots of an equation that the complex number system becomes particularly important. If a polynomial is formed with coefficients from the complex number system, then the sum of the multiplicities of the roots (from the complex numbers) of the polynomial coincides with the degree of the polynomial. Consequently, a complex number polynomial $f(x)$ can be factored completely as $f(x) = (x - a_1)(x - a_2) \ldots (x - a_n)$, where $a_1, a_2, \ldots, a_n$ are the complex number roots and n is the degree of the polynomial. This is not always possible for real polynomials. For example, the polynomial $f(x) = x^2 + 1$ cannot be factored further by using polynomials with real coefficients. It is called irreducible over the real numbers. As a complex polynomial, it can be factored as $x^2 + 1 = (x + i) \cdot (x - i)$.

Factoring of polynomials. Even in the real numbers, a polynomial often may be factored into other polynomials of lesser degree. This allows for analysis of the polynomial by consideration of the simpler factors. Some of the most common and useful factoring patterns are given in Eqs. (12).

$$x^2 - a^2 = (x - a)(x + a) \qquad (12a)$$

$$x^2 + 2ax + a^2 = (x + a)(x + a) \qquad (12b)$$

$$x^2 - 2ax + a^2 = (x - a)(x - a) \qquad (12c)$$

$$x^3 + 3ax^2 + 3a^2x + a^3 = (x + a)(x + a)(x + a) \qquad (12d)$$

$$x^3 + 3ax^2 - 3a^2x - a^3 = (x - a)(x - a)(x - a) \qquad (12e)$$

$$x^3 - a^3 = (x - a)(x^2 + ax + a^2) \qquad (12f)$$

$$x^3 + a^3 = (x + a)(x^2 - ax + a^2) \qquad (12g)$$

Rational functions. Functions which are of the form $f(x) = g(x)/h(x)$, where $g(x)$ and $h(x)$ are polynomials, are called rational functions. They frequently appear in instances where comparisons of two polynomials are necessary. A rational function is defined only at points where the denominator is nonzero. The roots of a rational function agree with

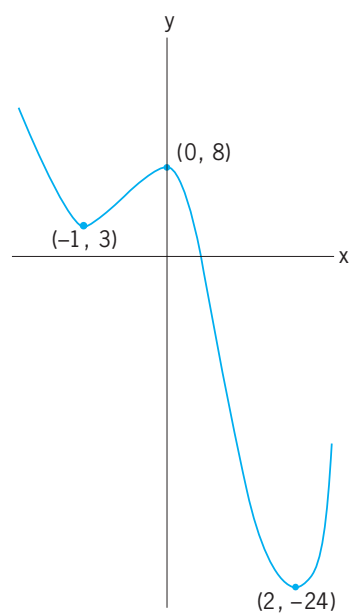

Fig. 4. Graph of the degree-4 polynomial $f(x) = 3x^4 - 4x^3 - 12x^2 + 8$, which changes direction three times.

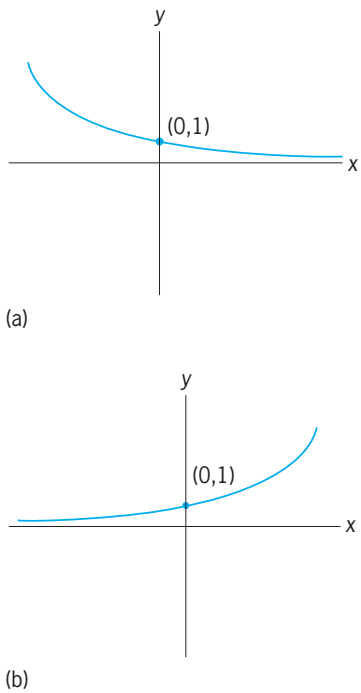

Fig. 5. Graphs of exponential functions. (*a*) Decreasing graph of $y = (\frac{1}{2})^x$. (*b*) Increasing graph of $y = 2^x$.

the roots of the polynomial that is found in the numerator.

Exponential functions. In general, an exponential function is one which has the form $f(x) = b^x$ for some constant b in the number system. The graphs of these functions take on one of two basic types. The graph is decreasing if $0 < b < 1$ (**Fig. 5***a*), and it is increasing if $b > 1$ (Fig. 5*b*). The case of $b = 1$ is trivial, since all values of the function are 1.

Exponential functions are of special importance in describing a situation where the rate of growth is proportional to the amount present. Among the situations that can be described by exponential functions are those involving the calculation of interest, radioactive decay, and population growth.

Calculation with exponential functions is based on the basic equalities given in Eqs. (13).

$$b^{x+y} = b^x b^y \tag{13a}$$

$$b^{x-y} = \frac{b^x}{b^y} \tag{13b}$$

$$(ab)^x = a^x b^x \tag{13c}$$

$$(b^x)^y = b^{xy} \tag{13d}$$

Logarithmic functions. Another function that plays an especially important role in algebra is the logarithm. In general, if $b > 1$, the logarithm to the base b of a value x is denoted by Eq. (14). The logarithm

$$y = \log_b x \tag{14}$$

function is defined in terms of exponentials by reversing the roles of the x and y variables. It follows that if $y = \log_b$, then y is the exponent which when applied to b produces x, as in Eqs. (15). The loga-

$$y = \log_b x \leftrightarrow x = b^y \tag{15}$$

rithm function, like the exponential function, has a characteristic graph (**Fig. 6**).

Basic rules that are necessary for calculation with logarithms are given in Eqs. (16). For example, since

$$\log_b(xy) = \log_b x + \log_b y \tag{16a}$$

$$\log_b\left(\frac{x}{y}\right) = \log_b x - \log_b y \tag{16b}$$

$$\log_b x^r = r(\log_b x) \tag{16c}$$

$$\log_b b^x = x \tag{16d}$$

$$b^{\log_b x} = x \tag{16e}$$

logarithms are exponents, the exponent for the product in Eq. (16*a*) is the sum of the exponents for the factors. *See* LOGARITHM.

Systems of linear equations. Frequently, a situation is described by a system of linear equations which must all be satisfied simultaneously. For example, the system of Eqs. (17) involves three variables and three equations.

$$\begin{aligned} 2x + 3y - 4z &= 2 \\ 4x - 5y - z &= 0 \\ -3x + y - 2z &= 1 \end{aligned} \tag{17}$$

More general systems involve n variables and m equations and take the form of Eqs. (18). A solution

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{18}$$

to the system is a set of numbers $c_i, c_2, \ldots, c_n$ which satisfies each equation.

The equations in a linear system can be multiplied by constants, where a single number is multiplied by each term in the sum. They can also be added together, where the coefficients of the corresponding variables are summed. A linear combination of one or more equations in the system is formed by multiplying each equation by a constant and then

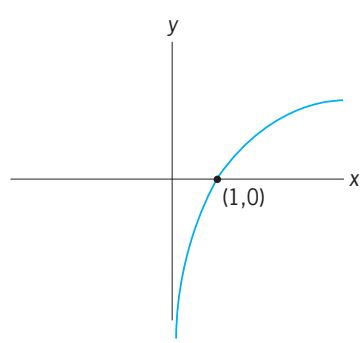

Fig. 6. Graph of the logarithm function $y = \log_b x$.

summing the resulting equations. When one equation in a system is a linear combination of one or more of the other equations, the system is called dependent. The solutions to a dependent system are the same as the solutions to the system formed by deleting the equation that is a linear combination of other equations in the system. In order to find solutions to a system, it is frequently expedient to first try to eliminate equations that are linear combinations of others in the system. *See* LINEAR SYSTEMS OF EQUATIONS.

Wayne B. Powell

Bibliography. D. T. Christy, *College Algebra*, 2d ed., 1993; C. H. Edwards and D. Penney, *Calculus and Analytic Geometry*, 4th ed., 1994.

Algebraic geometry

The study of zero sets of polynomial equations. Examples are the parabola $y - x^2 = 0$, thought of as sitting in the (x, y)-plane, and the locus of all points (t, t^2, t^3, t^4), which is defined by Eqs. (1), in the

$$y^2 = xz \qquad yx = wz \qquad wy = x^2 \tag{1}$$

coordinate space with coordinates w, x, y, z. Another interesting example is an elliptic curve, typically defined by an equation like Eq. (2), where A and B are

$$y^2 = x(x - A)(x - B) \tag{2}$$

constants. Objects such as these are called affine algebraic sets. They exist in affine n-space, denoted $\mathbf{A}^n$, which is defined to be the coordinate space with coordinates $x_1, \ldots, x_n$. The coordinate ring $k[V]$ of an affine algebraic set V is the set of functions on V obtained by restriction from polynomial functions on the ambient affine space. These functions can be added, subtracted, and multiplied (that is, they form a ring). For example, the coordinate ring of the parabola has in it functions y and x satisfying $y = x^2$. This relation can be used to eliminate all references to y, so that the coordinate ring of the parabola is identified with the ring of polynomials in x. *See* ANALYTIC GEOMETRY; POLYNOMIAL SYSTEMS OF EQUATIONS; RING THEORY.

The possibility of studying a question geometrically via the zero loci of polynomials or algebraically via the algebra of the coordinate ring gives the subject much of its power and flavor. This has led to an amazing growth in applications to other disciplines. For example, the integers $\ldots, -2, -1, 0, 1, 2, \ldots$ form a ring that is algebraically similar to the coordinate ring of the affine line, and algebrogeometric methods have come to play a central role in number theory. In studying the path space of strings and the partition function, modern physicists have made the moduli space of curves a central object of research. Finally, a number of differential equations in engineering and physics can best be studied algebrogeometrically, although their solutions are not algebraic functions. *See* DIFFERENTIAL EQUATION; NUMBER THEORY; SUPERSTRING THEORY.

Basic numerical invariants. An algebraic set V is irreducible if $fg = 0$ in $k[V]$ implies that either f is 0 or g is 0. For example, the algebraic set defined by $xy = 0$ is not irreducible. An irreducible algebraic set is called an algebraic variety.

Dimension. The procedure of defining the dimension of an algebraic variety V begins with the intuition that affine n-space has dimension n. Geometrically, if V is defined by the vanishing of several polynomials in n variables, so that V is contained in affine n-space, the coordinates on $\mathbf{A}^n$ are changed in what is said to be a general way and the projection given by Eqs. (3) is considered. It turns out that for

$$\mathbf{A}^n \to \mathbf{A}^d \qquad (x_1, \ldots, x_n) \to (x_1, \ldots, x_d) \tag{3}$$

one value of $d \le n$, this mapping when restricted to V will be surjective with finite fibers, that is, there will exist at least one and at most a finite number of points of V of the form $(a_1, \ldots, a_d, \ldots)$ for given $a_1, \ldots, a_d$. This value of d is the dimension of V. Algebraically, if $k[V]$ denotes the coordinate ring of V, the function field $k(V)$ is considered, consisting of fractions f/g with f and g in $k[V]$ and g not zero. [The word "field" refers to the fact that in $k(V)$ it is possible to add, subtract, multiply, and divide.] There will exist, in $k(V)$, d elements $z_1, \ldots, z_d$ satisfying no polynomial relations among themselves, but having the property that, for any w in $k(V)$, there is a nonzero polynomial $P(x_1, \ldots, x_{d+1})$ with $P(z_1, \ldots, z_d, w) = 0$. *See* FIELD THEORY (MATHEMATICS).

Degree. The degree of an algebraic set is a central notion of enormous depth and power. Intuitively, if V is a variety of dimension d in $\mathbf{A}^n$, the degree of V is the number of points of intersection of V with $\mathbf{A}^{n-d}$. There are a number of pitfalls here. So far, there has been no specification of where the functions take values. If the coordinates are real numbers, then the parabola $y - x^2 = 0$ meets the line $y = 1$ in two points, but it meets the line $y = -1$ in zero points (**Fig. 1**). This problem is solved by taking coordinates in the field $\mathbf{C}$ of complex numbers. (More generally, it is possible to work over any algebraically closed

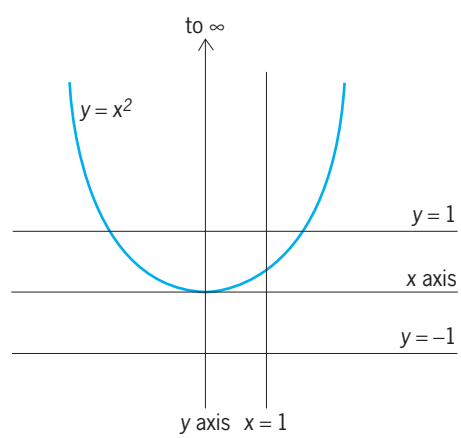

Fig. 1. Parabola in affine 2-space, $\mathbf{A}^2$. While it would seem that lines in $\mathbf{A}^2$ can meet the parabola in 0, 1, or 2 points, with the introduction of algebraically closed fields of coordinates and projective varieties, all such intersections have two points, counting multiplicity.

field. A field F is said to be algebraically closed if any polynomial with coefficients in F has a root in F. The complex numbers have this property.) It can also be argued that the parabola meets the line $y = 0$ in just one point, but this is a point of tangency and therefore must be counted twice. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

A more serious difficulty is how to arrange for the parabola to meet the lines $x =$ constant in two points. A glance at Fig. 1 suggests that one point of intersection should be thought of as lying at ∞ (infinity). The affine (x, y) plane containing the parabola can be imagined positioned as the locus $Z = 1$ inside $\mathbf{A}^3$ with coordinates X, Y, Z. The set of all lines through the origin in $\mathbf{A}^3$ meeting the parabola (**Fig. 2**) is contained in the locus $X^2 - YZ = 0$, but there is one line, defined by $Z = X = 0$, lying in this locus but not meeting the parabola. The lines through the origin meeting the line $x = c$ lie in the plane $X - cZ = 0$ in $\mathbf{A}^3$, and the common zeros, satisfying Eq. (4),

$$X - cZ = 0 = X^2 - YZ \tag{4}$$

consist of two lines, one being the line that joins (0,0,0) with $(c, c^2, 1)$ and the other being the line $Z = X = 0$. This motivates defining projective n-space, $\mathbf{P}^n$, to be the set of all lines through the origin in $\mathbf{A}^{n+1}$. For example, the Riemann sphere $\mathbf{P}^1$ is equivalent to $\mathbf{A}^1$ together with the point ∞ at infinity. Projective algebraic sets in $\mathbf{P}^n$ are defined by the vanishing of homogeneous polynomials in the coordinates on $\mathbf{A}^{n+1}$. For many purposes, it is preferable to work with projective algebraic sets, which can be thought of as completions or compactifications of the affine algebraic sets defined above. In the above example, the collection of lines through the origin lying on the homogeneous conic $X^2 - YZ = 0$, which comprises the lines that meet the parabola $x^2 = y$ together with the line $X = Z = 0$, is the projective completion of the parabola. If V is a projective algebraic variety of dimension n in $\mathbf{P}^N$ and L is a linear space of dimension $N - n$ such that the intersection of V and L is finite, then the number of points of intersection of V and L, counted with multiplicity, is independent of L and is called the degree of V. *See* PROJECTIVE GEOMETRY.

Classification. The problem of classifying projective algebraic varieties typically leads to a finite

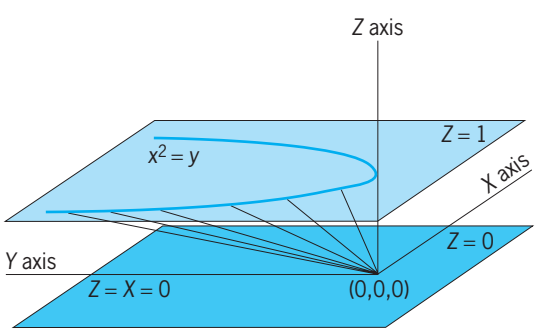

Fig. 2. Projective completion of the parabola of Fig. 1, consisting of lines through the origin lying on the homogeneous conic $X^2 - YZ = 0$, including lines meeting the parabola $x^2 = y$ together with the line $Z = X = 0$.

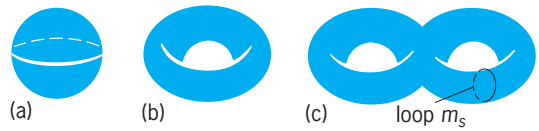

Fig. 3. Riemann surfaces of low genus. (*a*) Genus 0. (*b*) Genus 1. (*c*) Genus 2.

number of discrete invariants together with a continuous parameter space, the moduli space of the problem, which is itself an algebraic variety. The classification of projective varieties of dimension 1 (algebraic curves) and dimension 2 (algebraic surfaces) will be considered.

Algebraic curves. To help visualize the classification of algebraic curves, the coordinates will be chosen to be complex numbers, and the curves will be assumed to be nonsingular, that is, to have a well-defined tangent at each point. In this case, the algebraic curves are topological surfaces (frequently referred to as Riemann surfaces), and, by topology, they can all be realized as (hollow) doughnuts with $g = 0, 1, 2, \ldots$ holes (**Fig. 3**). The number of holes is the genus of the curve and is the only discrete invariant.

Algebraically, a polynomial in two variables $P(x, y)$ can be written as in Eq. (5). After projective comple-

$$a_0(x)y^n + a_1(x)y^{n-1} + \cdots + a_n(x) = 0 \tag{5}$$

tion and elimination of singular points (by a process called normalization), the locus of zeros of P yields an algebraic curve C. For a general value of x, there are n solutions for y, and therefore C can be visualized as an n-sheeted covering of the projective line $\mathbf{P}^1$. The function y has poles on C, but it is single-valued and satisfies the given polynomial relation. Simple, but by no means trivial, examples are the elliptic and hyperelliptic equation (6). Values of x (possibly including

$$y^2 = f(x) \tag{6}$$

$x = \infty$) for which there are fewer than n solutions in y are called branch points. For a general polynomial $P(x, y)$ there will be $n - 1$ solutions in y over branch points. In this case, the number b of branch points is even, and the genus of C is given by the Hurwitz formula (7).

$$g = 1 - n + (b/2) \tag{7}$$

See CONFORMAL MAPPING; TOPOLOGY.

Curves of a given genus form a continuous family parametrized by the moduli space $\mathbf{M}_g$. The only curve of genus zero is the Riemann sphere, and therefore $\mathbf{M}_g$ is a point. Curves of genus 1 (elliptic curves) are complex tori, which can be viewed as quotients of the complex plane modulo a lattice L_z (called the period lattice) consisting of elements of the form $m + nz$ for fixed complex number z and all integers m and n. Scaling the lattice by a complex factor does not change the quotient curve, and therefore, if a, b, c, and d are integers satisfying Eq. (8) and

w is given by Eq. (9), then L_w and L_z give the same curve. [Indeed, L_z is related to L_w by Eq. (10).] Al-

$$ad - bc = \pm 1 \tag{8}$$

$$w = (az + b)(cz + d)^{-1} \tag{9}$$

$$L_z = (cz + d)L_w \tag{10}$$

gebraic functions on $\mathbf{M}_1$ are meromorphic functions $f(z)$ defined for $z = x + y\sqrt{-1}$ and $y \neq 0$, satisfying $f(z) = f(w)$ with any w as above, and not growing too rapidly as $y \to \infty$. These coincide with rational functions $P[j(z)]/Q[j(z)]$ in the classical j-function given by Eq. (11). In fact, $\mathbf{M}_1 = \mathbf{A}^1$ with parameter

$$j(z) = q^{-1}(1 + 744q + 196884q^2 + \cdots) \qquad q = e^{2\pi iz} \tag{11}$$

j. For $g > 1$, $\mathbf{M}_g$ is an algebraic variety of dimension $3g - 3$. Although not complete, it may be completed to a projective variety by admitting points corresponding to certain singular curves (stable curves).

A divisor D on a curve C is a formal sum $n_1p_1 + n_2p_2 + \cdots + n_rp_r$ with n_i an integer and each p_i a point of C. The degree of D equals $n_1 + n_2 + \cdots + n_r$. To an algebraic function f on C is associated a divisor (f) of degree 0 by associating to each singular point its multiplicity as a zero or pole. By definition, the jacobian $J(C)$ is the group of divisors of degree 0 modulo the subgroup of divisors (f) associated to functions. The jacobian $J(C)$ is a g-dimensional complex torus, whose period lattice is given by integrating a basis of g algebraic differential one-forms over closed loops. In fact, $J(C)$ is a projective algebraic variety (called a principally polarized abelian variety), and the classical Torelli theorem states that the association $C \to J(C)$ embeds $\mathbf{M}_g$ into the Siegel moduli space $\mathbf{A}_g$ of principally polarized abelian varieties of dimension g. The Schottky problem of identifying the image of $\mathbf{M}_g$ in $\mathbf{A}_g$ has been solved.

Algebraic surfaces. The classification of projective varieties of dimension 2 (algebraic surfaces) hinges on a careful analysis of line bundles and linear systems. To map a variety V to $\mathbf{P}^n$ means to associate to each point of V a line through the origin in $\mathbf{A}^{n+1}$, but a modern algebraic geometer would take a dual view and think of the space of linear functions on the line. The collection of these dual lines associated to points of V forms a line bundle. A linear function f on $\mathbf{A}^{n+1}$ induces an element f_v in the line over v in V, and hence a section of the line bundle. The set of v where the f_v vanishes is a Cartier divisor on V. A two-dimensional subspace of linear functions gives rise to a linearly varying family of Cartier divisors called a pencil on V.

This correspondence between line bundles, linear systems, and Cartier divisors lies at the heart of the classification of algebraic surfaces. The key difficulty in calculating the dimension of the space of all sections of a given line bundle is usually overcome via the Riemann-Roch theorem. The space of sections of line bundle L is denoted $H^0(V, L)$, but there are also defined higher cohomology groups $H^1(V, L)$ and $H^2(V, L)$. These are finite-dimensional vector spaces. The Riemann-Roch theorem calculates the Euler-Poincaré characteristic given by Eq. (12).

$$\chi(L) = \text{dimension of } H^0(V, L) - \text{dimension of } H^1(V, L) + \text{dimension of } H^2(V, L) \tag{12}$$

Among the interesting kinds of surfaces are (1) rational surfaces, whose function field $k(V)$ is generated by two independent transcendental elements ($\mathbf{P}^2$ is a rational surface); (2) ruled surfaces, with a one-parameter family of curves of genus 0; (3) elliptic surfaces, with a pencil of elliptic curves; and (4) K3 surfaces, a family of surfaces including nonsingular surfaces of degree 4 in $\mathbf{P}^3$. These K3 surfaces are of interest to physicists as providing a plausible model for the universe.

Applications to differential equations. The linear Picard-Fuchs equations arise very naturally. If $[X_s]$ is a family of algebraic curves (or indeed of varieties of any dimension), depending on a parameter s, then the de Rham cohomology of a curve X_s can be thought of as the space of closed differential forms modulo exact forms, that is, as the space of integrands to be integrated over closed loops m_s on X_s (Fig. 3). These integrands can themselves be differentiated with respect to the parameter s. Since they form a finite-dimensional space, there will be some operator having the form given by Eq. (13)

$$D = a_0(s)(d/ds)^n + a_1(s)(d/ds)^{n-1} + \cdots + a_n(s) \tag{13}$$

[Picard-Fuchs equation] which vanishes identically. A solution may be obtained by choosing a de Rham cohomology class w_s and a loop m_s varying continuously with s. The integral in Eq. (14) satisfies Eq. (15). As s describes a loop in the parameter

$$F(s) = \int w_s \tag{14}$$

$$DF(s) = \int Dw_s = 0 \tag{15}$$

space, m_s does not necessarily return to its original position, and therefore $F(s)$ is in general a multiple-valued function of s. For example, the hypergeometric equation (16) can be studied in this way.

$$s(1 - s)(d^2f/ds^2) + [c - (a + b + 1)s](df/ds) - abf = 0 \tag{16}$$

See HYPERGEOMETRIC FUNCTIONS.

There has also been much interest in certain nonlinear partial differential equations such as the Korteweg–de Vries equation (17). The coordinate

$$\partial f/\partial t = \partial^3 f/\partial s^3 + 6f\partial f/\partial s \tag{17}$$

ring R of an affine algebraic curve is embedded inside a large ring D of differential operators of the form

$D = a_0(s)(d/ds)^n + \cdots$ mentioned above. This embedding depends on the choice of a line bundle L. When L is made to vary with respect to a second parameter t, the Korteweg–de Vries equation and related equations arise as equations of motion for a fixed element of R inside D.

Applications to number theory. Three major conjectures have been solved by using algebrogeometric techniques. The Weil conjecture concerned numbers of solutions of equations mod p. A key idea was to reinterpret solutions modulo p as fixed points of the Frobenius endomorphism given by Eq. (18). The

$$(x_1, \ldots, x_n) \to (x_1^p, \ldots, x_n^p) \tag{18}$$

Lefschetz fixed-point theorem in algebraic topology counts the number of fixed points of an endomorphism. This theorem was ported to algebraic geometry via étale cohomology and used to help solve the conjecture.

The Mordell conjecture stated that a curve of genus greater than 1 can have only a finite number of points whose coordinates are rational numbers. This implies, for example, that many polynomials in two variables, like the Fermat equation (19), have

$$x^n + y^n - 1 = 0 \qquad (n \geq 4) \tag{19}$$

only finitely many rational solutions. The proof of the Mordell conjecture involved an arithmetic generalization of the concept of degree, which was then applied to study the geometry of the Siegel moduli space.

Finally, it seems likely that the Fermat conjecture itself, which states that Eq. (19) has no rational solutions with $xy \neq 0$, has been proved. Thus it can be said that Eq. (20) has no solution in integers with n

$$a^n + b^n + c^n = 0 \tag{20}$$

prime and $abc \neq 0$. Central to the proof is the Frey elliptic curve, given by Eq. (21), which is associated

$$y^2 = x(x - a^n)(x + b^n) \tag{21}$$

to a solution of Eq. (20).

These problems illustrate the power of geometry in dealing with seemingly algebraic problems like the Mordell and Fermat conjectures which admit no plausible, purely algebraic attack. Even the study of mathematics at the school and college level, where the algebraic manipulation of formulas tends to be emphasized, can benefit from greater attention to deeper and more powerful geometric insights. *See* GEOMETRY.

Spencer J. Bloch

Bibliography. S. S. Abhyanker, *Algebraic Geometry for Scientists and Engineers*, 1992; J. Harris, *Algebraic Geometry: A First Course*, 1992; R. Hartshorne, *Algebraic Geometry,* rev. ed., 1991; I. R. Shafarevich, *Basic Algebraic Geometry,* 1977, paper, 1990.

Alginate

A major constituent (10–47% dry weight) of the cell walls of brown algae. Extracted for its suspending, emulsifying, and gelling properties, it is one of three algal polysaccharides of major economic importance, the others being agar and carrageenan. *See* AGAR; CARRAGEENAN.

Alginate is a linear (unbranched) polymer in which regions have a predominance of either D-mannuronic acid with β-1,4 linkages or L-guluronic acid with α-1,3 linkages. The ratio of the monomers varies greatly among different species, in the same species at different seasons, and in different parts of the same plant. Regions rich in mannuronic acid have a ribbonlike conformation, while those rich in guluronic acid have kinks with a marked affinity for calcium ions. Alginate gels are not formed by cooling, as with agar and carrageenan, but by the addition of calcium ions, which causes separate polymeric chains to cohere along surfaces rich in guluronic acid.

The chief sources of alginate are members of the family Fucaceae (rockweeds) and the order Laminariales (kelps), harvested from naturally occurring stands on North Atlantic and North Pacific shores. The supply has been significantly increased through mariculture, especially in China. Processing is carried out chiefly in the United Kingdom, United States, Norway, Canada, France, and China. In the extraction process, washed and macerated plants (fresh or dried) are digested with alkali to solubilize the alginate, and the filtrate is precipitated with a concentrated solution of calcium chloride. The calcium alginate is converted to alginic acid by dilute hydrochloric acid. After further purification and chemical treatment, the product is marketed as sodium, potassium, or ammonium alginate, usually in powdered form.

Because of its colloidal properties, alginate finds numerous industrial applications, especially in the food, textile, paper, printing, paint, cosmetics, and pharmaceutical industries. About half of the consumption is in the making of ice cream and other dairy products, in which alginate prevents the formation of coarse ice crystals and provides a smooth texture. As an additive to paint, it keeps the pigment in suspension and minimizes brush marks. An alginate gel is used in making dental impressions. *See* ICE CREAM; MILK; PHAEOPHYCEAE.

Paul C. Silva; Richard L. Moe

Algorithm

A well-defined procedure to solve a problem. The study of algorithms is a fundamental area of computer science. In writing a computer program to solve a problem, a programmer expresses in a computer language an algorithm that solves the problem, thereby turning the algorithm into a computer program. *See* COMPUTER PROGRAMMING.

Classic algorithms. Efficient algorithms for solving certain classic problems were discovered long before

the advent of electronic computers.

Sieve of Eratosthenes. The algorithm known as the Sieve of Eratosthenes dates back to the 3d century B.C. It solves the problem of finding all the prime numbers between 1 and a positive integer N. It works as follows: First write the numbers from 1 to N, and cross out 1. Then cross out the multiples of 2 except 2 itself. Then find the first number not yet crossed out after 2. The number is 3. Cross out the multiples of 3 except 3 itself. Then find the first number not yet crossed out after 3. The number is 5. Cross out the multiples of 5 except 5 itself. And so on. This process is repeated until no more numbers can be crossed out.

For example, the case $N = 20$, carried through the crossing out of multiples of 3, gives:

$$\cancel{1}\ 2\ 3\ \cancel{4}\ 5\ \cancel{6}\ 7\ \cancel{8}\ \cancel{9}\ \cancel{10}\ 11\ \cancel{12}$$
$$13\ \cancel{14}\ \cancel{15}\ \cancel{16}\ 17\ \cancel{18}\ 19\ \cancel{20}$$

Consideration of the numbers 5, 7, and so on will not cross out any more numbers. The numbers that have not been crossed out (2, 3, 5, 7, 11, 13, 17, 19) are the prime numbers between 1 and 20. *See* NUMBER THEORY.

Euclid's algorithm. Euclid's algorithm, detailed in his treatise *Elements*, solves the problem of finding the greatest common divisor of two positive integers. It makes use of the fact that the greatest common divisor of two numbers is the same as the greatest common divisor of the smaller of the two and their difference. It works as follows: First find the larger of the two numbers and then replace it by the difference between the two numbers. Then find the larger of the resultant two numbers and then replace it by the difference between the two numbers. And so on. When the two numbers become equal, their value is the greatest common divisor and the algorithm terminates (finishes).

The steps of the algorithm when it is used to find the greatest common divisor of 48 and 18 are:

Replace 48 by 30 (= 48 − 18)	~~48~~	18
Replace 30 by 12 (= 30 − 18)	~~30~~	18
Replace 18 by 6 (= 18 − 12)	12	~~18~~
Replace 12 by 6 (= 12 − 6)	~~12~~	6
The answer is 6	6	6

Operation. An algorithm generally takes some input (for example, the value N in the Sieve of Eratosthenes and the two given integers in Euclid's algorithm), carries out a number of effective steps in a finite amount of time, and produces some output. An effective step is an operation so basic that it is possible, at least in principle, to carry it out using pen and paper. In computer science theory, a step is considered effective if it is feasible on a Turing machine or any of its equivalents. A Turing machine is a mathematical model of a computer used in an area of study known as computability, which deals with such questions as what tasks can be algorithmically carried out and what cannot. *See* AUTOMATA THEORY; RECURSIVE FUNCTION.

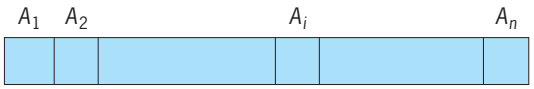

Fig. 1. An array of *n* elements.

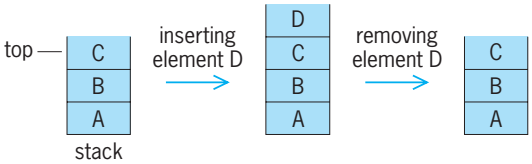

Fig. 2. Inserting and deleting an element from a stack.

Data structures. Many computer programs deal with a substantial amount of data. In such applications, it is important to organize data in appropriate structures so as to make it easier or faster to process the data. In computer programming, the development of an algorithm and the choice of appropriate data structures are closely intertwined, and a decision regarding one often depends on knowledge of the other. Thus, the study of data structures in computer science usually goes hand in hand with the study of related algorithms. Commonly used elementary data structures include records, arrays, linked lists, stacks, queues, trees, and graphs.

A record is a group of related components. For example, a record representing a date may have three components: month, day, and year. An array is a sequence of indexed (subscripted) components, A_1, $A_2, \ldots, A_n$, stored in the memory of a computer in such a manner that each component A_i can be accessed by specifying its index (subscript) i (**Fig. 1**). The components of an array can be accessed individually and independently. In contrast, a linked list is a sequence of components stored in such a manner that the components can be accessed only sequentially. That is, the first component must be accessed before the second, the second before the third, and so on. A stack is a sequence of components that allows component access, addition, and removal only at one end of the sequence called the top of the stack (**Fig. 2**). Adding an element to the top of a stack is

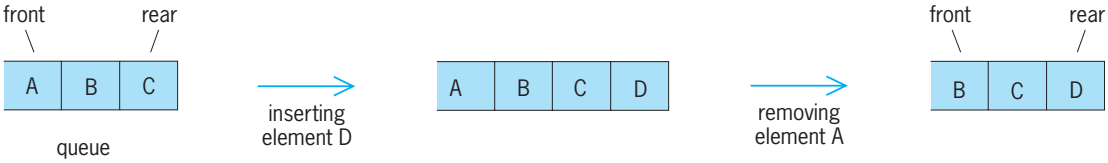

Fig. 3. Inserting and removing an element from a queue.

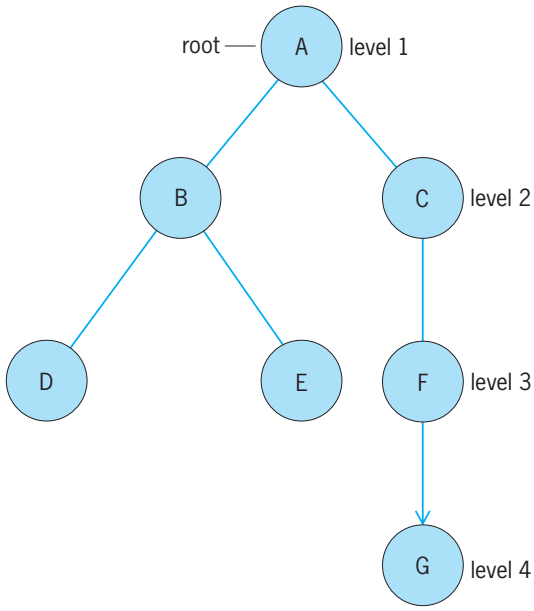

Fig. 4. An example of a tree.

generally called a push operation. Removing an element from a stack is called a pop operation. A queue is a sequence of components that allows component access and removal at one end (called the front of a queue) and addition of new components at the other end of the sequence (called the rear of a queue) [**Fig. 3**].

A tree is a data structure whose components are arranged in a hierarchy of levels. The topmost level has only one component called the root of the tree, but other levels can have many components. Each component on some level (except the root) is a child of some component on the level above. For example, each component on the second level is a child of the root, and each component on the third level is a child of some component on the second level, and so on. In summary, every component of a tree has one parent located on the level above (except the root, which does not have a parent) and possibly several children located on the level below. The number of levels in such a hierarchy is called the height of a tree. For example, the components *A*, *B*, *C*, *D*, *E*, *F*, and *G* will form a tree of height 4 if they are related in the following manner: A is the parent of *B* and *C*, *B* is the parent of *D* and *E*, *C* is the parent of *F*, and *F* is the parent of *G* (**Fig. 4**). A binary tree is a tree in which each component may have up to two children.

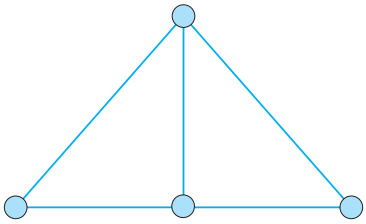

Fig. 5. A graph with four vertices and five edges.

A graph is similar to a tree in that it represents a collection of components and their relationship. However, the relationship among its components is not necessarily hierarchical, and any component can be related to any number of components. A component is called a vertex or node, and a relation between two vertices is called an edge. A convenient way to show a graph is to draw a picture in which a vertex is represented by a dot and an edge is represented by a line connecting two vertices (**Fig. 5**). If the edges have a direction assigned to them, the graph is called a directed graph. *See* GRAPH THEORY.

Elementary algorithms. Many algorithms are useful in a broad spectrum of computer applications. These elementary algorithms are widely studied and considered an essential component of computer science. They include algorithms for sorting, searching, text processing, solving graph problems, solving basic geometric problems, displaying graphics, and performing common mathematical calculations.

Sorting arranges data objects in a specific order, for example, in numerically ascending or descending orders. Internal sorting arranges data stored internally in the memory of a computer. Simple algorithms for sorting by selection, by exchange, or by insertion are easy to understand and straightforward to code. However, when the number of objects to be sorted is large, the simple algorithms are usually too slow, and a more sophisticated algorithm, such as heap sort or quick sort, can be used to attain acceptable performance. External sorting arranges data records stored on disk or tape.

Searching looks for a desired data object in a collection of data objects. Elementary searching algorithms include linear search and binary search. Linear search examines a sequence of data objects one by one. Binary search adopts a more sophisticated strategy and is faster than linear search when searching a large array. A collection of data objects that are to be frequently searched can also be stored as a tree. If such a tree is appropriately structured, searching the tree will be quite efficient.

A text string is a sequence of characters. Efficient algorithms for manipulating text strings, such as algorithms to organize text data into lines and paragraphs and to search for occurrences of a given pattern in a document, are essential in a word processing system. A source program in a high-level programming language is a text string, and text processing is a necessary task of a compiler. A compiler needs to use efficient algorithms for lexical analysis (grouping individual characters into meaningful words or symbols) and parsing (recognizing the syntactical structure of a source program). *See* SOFTWARE ENGINEERING; WORD PROCESSING.

A graph is useful for modeling a group of interconnected objects, such as a set of locations connected by routes for transportation. Graph algorithms are useful for solving those problems that deal with objects and their connections—for example, determining whether all of the locations are connected, visiting all of the locations that can be reached from a

given location, or finding the shortest path from one location to another.

Geometric objects such as line segments, circles, and polygons are fundamental in modeling many physical objects, such as a building or an automobile. Geometric algorithms are useful for solving many problems in designing and analyzing a geometric model of a physical object. Basic geometric algorithms include those for determining whether two line segments intersect, and whether a given point is within a given polygon. A related area, computer graphics, is the study of methods to represent and manipulate images, and to draw them on a computer screen or other output devices. Basic algorithms in the area include those for drawing line segments, polygons, and ovals. *See* COMPUTER GRAPHICS.

Mathematical algorithms are of wide application in science and engineering. Basic algorithms for mathematical computation include those for generating random numbers, performing operations on matrices, solving simultaneous equations, and numerical integration. Modern programming languages usually provide predefined functions for many common computations, such as random number generation, logarithm, exponentiation, and trigonometric functions.

Recursion. An algorithm for solving a problem can often be naturally expressed in terms of using the algorithm itself to solve a simpler instance of the problem. Euclid's algorithm makes use of the fact that the greatest common divisor of two different integers is the same as the greatest common divisor of the smaller number of the two and their difference. By recognizing that the problem of finding the greatest common divisor of the smaller number and the difference is a simpler instance of the given problem, the algorithm can be stated recursively: If the two given numbers are equal, their value is the answer; otherwise, (this is the recursive part) find the greatest common divisor of the smaller number and their difference. The following shows the steps of this recursive algorithm when used to find the greatest common divisor of 48 and 18.

a. Since 18 is smaller, find the greatest common divisor of 18 and 30 (= 48 − 18).
b. Since 18 is smaller, find the greatest common divisor of 18 and 12 (= 30 − 18).
c. Since 12 is smaller, find the greatest common divisor of 12 and 6 (= 18 − 12).
d. Since 6 is smaller, find the greatest common divisor of 6 and 6 (= 12 − 6).
e. The greatest common divisor is 6.

Since 6 is the greatest common divisor of 6 and 6, it is also the greatest common divisor of 6 and 12, of 18 and 12, of 18 and 30, and of the given numbers 48 and 18.

Analysis of algorithms. A problem to be solved by an algorithm generally has a natural problem size, which is usually the amount or the magnitude of data to be processed, and the number of operations that the algorithm performs depends predominantly on the problem size. A common approach to characterize the time efficiency of an algorithm is to count the number of operations that the algorithm performs in the worst case. For example, the problem of searching an array $A_1, A_2, \ldots, A_n$ for a value V has the problem size n. The linear search algorithm successively compares the desired value V with the array components $A_1, A_2, \ldots, A_n$ until it finds that a component is equal to V or until it has compared V with all of the components without finding one that is equal to V. The worst case obviously requires the algorithm to compare V with all n components. Therefore, the number of operations performed by linear search is essentially proportional to n, and linear search is said to have a running time of $O(n)$.

If the array is sorted, binary search can be used instead of linear search. Binary search first compares V with the middle component A_m, where m is $(1 + n)/2$. If A_m is equal to V, it terminates; otherwise, the search continues with at most one-half of the array components remaining to be considered. (Depending on whether V is greater than or less than A_m, either the components $A_1 \ldots A_{m-1}$ or the components $A_{m+1} \ldots A_n$ still need to be considered in the search, since the array is sorted.) A comparison of V with the middle component of the part of the array remaining to be considered either will lead to success of the search and termination of the algorithm, or again will reduce the number of elements to be searched by at least one-half of those remaining to be considered, and so forth. In summary, each comparison either leads to termination of the algorithm or reduces the number of components remaining to be considered in the search by at least one-half. In the worst case, about $\log_2 n$ comparisons are needed to reduce the number of components remaining in the search to one (by then the search is trivial). The number of operations is essentially proportional to $\log_2 n$, and the algorithm is said to have a running time of $O(\log n)$. When the problem size n is large, $\log_2 n$ is much smaller than n, indicating that binary search will be much faster than linear search when searching a large array.

An open question in the study of computational complexity concerns a class known as NP-complete problems. The class includes such problems as the traveling salesperson problem (minimizing the distance traveled in a tour of a number of cities) and finding a hamiltonian path in a graph (a path that passes through each vertex once). Although the correctness of proposed solution to these problems can be checked in polynomial time (a running time that is a polynomial function of the problem size), no polynomial-time algorithms to solve these problems have been found. (Known algorithms to solve them have an exponential running time, which makes them impractical when the problem size is large.) Furthermore, if any NP-complete problem can be solved in polynomial time, so can all others in the

class. Whether these problems can be solved in polynomial time remains an open question. *See* CRYPTOGRAPHY.

Randomized algorithms. By using a randomizer (such as a random number generator) in making some decisions in an algorithm, it is possible to design randomized algorithms that are relatively simple and fast. Since the output of a randomizer may vary from run to run, the output and the running time of a randomized algorithm may differ from run to run for the same input. There is a nonzero probability that the output of a randomized algorithm may be incorrect and execution of the algorithm may be slower than desired. A design goal of randomized algorithms is to keep such a probability small. However, there are critical systems that cannot tolerate a nonzero, albeit small, probability of incorrect output or a longer running time.

Adaptive systems. In many applications, a computer program needs to adapt to changes in its environment and continue to perform well. An approach to make a computer program adaptive is to use a self-organizing data structure, such as one that is reorganized regularly so that those components most likely to be accessed are placed where they can be most efficiently accessed. A self-modifying algorithm that adapts itself is also conceivable. For developing adaptive computer programs, biological evolution has been a source of ideas and has inspired evolutionary computation methods such as genetic algorithms. *See* GENETIC ALGORITHMS.

Parallel algorithms. Certain applications require a tremendous amount of computation to be performed in a timely fashion. An approach to save time is to develop a parallel algorithm that solves a given problem by using a number of processors simultaneously. The basic idea is to divide the given problem into subproblems and use each processor to solve a subproblem. The processors usually need to communicate among themselves so that they may cooperate. The processors may share memory, through which they can communicate, or they may be connected by communication links into some type of network such as a hypercube. *See* CONCURRENT PROCESSING; MULTIPROCESSING; SUPERCOMPUTER.

Samuel C. Hsieh

Bibliography. F. M. Carrano, P. Helman, and R. Veroff, *Data Abstraction and Problem Solving with C++*, Addison Wesley Longman, 1998; E. Horowitz, S. Sahni, and S. Rajasekaran, *Computer Algorithms/C++*, Computer Science Press, 1996; B. R. Preiss, *Data Structures and Algorithms with Object-Oriented Design Patterns in C++*, Wiley, 1999.

Alicyclic hydrocarbon

An organic compound that contains one or more closed rings of carbon atoms. The term alicyclic specifically excludes carbocyclic compounds with an array of π-electrons characteristic of aromatic rings. Compounds with one to five alicyclic rings of great variety and complexity are found in many natural products such as steroids and terpenes. By far the majority of these have six-membered rings. *See* AROMATIC HYDROCARBON; STEROID; TERPENE.

Structures and nomenclature. The bonding in cyclic hydrocarbons is much the same as that in open-chain alkanes and alkenes. An important difference, however, is the fact that the atoms in a ring are part of a closed loop. Complete freedom of rotation about a carbon-carbon bond (C—C) is not possible; the ring has faces or sides, like those of a plate.

Simple monocyclic hydrocarbons are usually represented as bond line structures; for example, cyclopropane (**1***a*) is usually represented as structure (**1***b*) and methylcycloheptane as structure (**2**). These

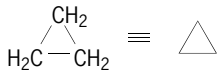

(**1***a*) (**1***b*)

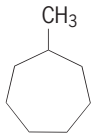

(**2**)

hydrocarbons are named by adding the prefix cyclo to the stem of the alkane corresponding to the number of atoms in the ring. When two or more substituents are attached to the ring, the relative positions and orientation must be specified: *cis* on the same side and *trans* on the other, as in *cis*-1,2-dimethylcyclopentane (**3**) and trans-1,3-dimethylcyclopentane (**4**).

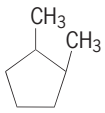

(**3**)

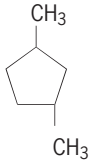

(**4**)

In bicyclic compounds the rings can be joined in three ways: spirocyclic, fused, and bridged, as illustrated in the structures for spiro[4.5]decane (**5**), bicyclo[4.3.0]nonane (**6**), and bicyclo[2.2.1]heptane (**7**). In each case the name indicates the total

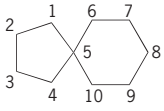

(**5**)

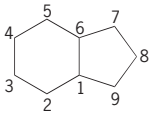

(**6**)

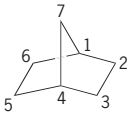

(**7**)

number of carbon atoms,and the number of atoms in each bridge. Atoms are numbered as shown.

Any of these bicyclic systems can be transformed to a tricyclic array by introduction of another bond between nonadjacent carbons, as in structure (**8**), or an additional ring, as in structure (**9**). Cyclic structure

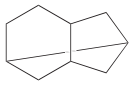

(**8**)

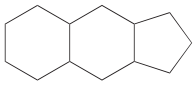

(**9**)

gives rise to the possibility of compounds made up of molecular subunits that are linked mechanically rather than chemically. In rotaxanes (**10**), bulky groups are introduced at the ends of a long chain that is threaded through a large ring ($>C_{30}$). Cyclization of the ends leads to a catenane (**11**). Several

(**10**)

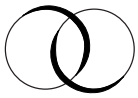

(**11**)

examples of compounds with these structures have been prepared.

Ring strain and conformation. In a planar ring of n carbon atoms, the internal bond angles are given by

$$180\left(\frac{n-2}{n}\right)$$

and the angles (α) for three-, four-, and five-membered rings are 60°, 90°, and 108°, respectively. The angle between bonds of a tetrahedral carbon atom is 109°. Thus in the smaller rings, particularly cyclopropane, a significant distortion of the tetrahedral bond angle is required, and the ring is strained. (**Fig. 1**). *See* BOND ANGLE AND DISTANCE.

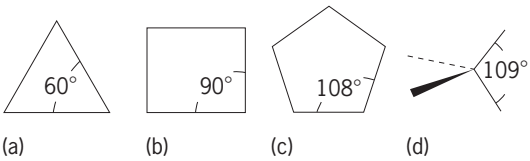

Fig. 1. Internal bond angles (α) for (*a–c*) planar rings and (*d*) a tetrahedral carbon atom.

Six-membered and larger rings can have normal bond angles by adopting nonplanar puckered conformations. For cyclohexane, two nonplanar arrangements are called chair and boat (**Fig. 2**). *See* CONFORMATIONAL ANALYSIS.

Angle strain is not the only consideration in the conformation of alicyclic compounds. Another factor is the interaction that occurs when bonds on adjacent carbons are aligned (eclipsing interaction; Fig. 2), as in cyclopropane or the boat form of cyclohexane. The rings in cyclobutane and cyclopentane are bent slightly to reduce these eclipsing interactions; cyclohexane exists almost entirely in a somewhat twisted chair conformation. In medium rings of 8–11 carbon atoms, another destabilizing effect is the interaction of atoms situated across the ring from each other.

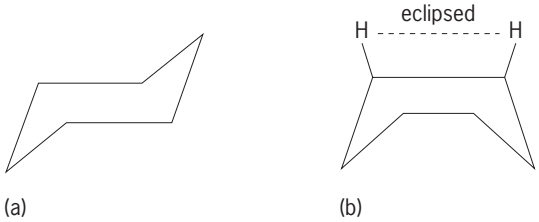

Fig. 2. Nonplanar (*a*) chair and (*b*) boat structures of cyclohexane.

The presence of a *cis* double bond in small rings increases the ring strain somewhat, but cyclobutene and even cyclopropene can be prepared. A cyclopropene ring is present in the naturally occurring sterculic acid (**12**). Alicyclic rings with a *trans* double bond or a triple bond are possible when the ring contains eight or more carbons, as in *trans*-cyclooctene (**13**).

(**12**)

(**13**)

Properties. The boiling points, melting points, and densities of cycloalkanes are all higher than those of their open-chain counterparts, reflecting the more compact structures and greater association in both liquid and solid. Geometrical constraints in the smaller rings have significant effects on reactions of cycloalkanes and derivatives. Because of ring strain, ring-opening reactions of cyclopropane, such as isomerization to propene, take place under conditions that do not affect alkanes or larger-ring cycloalkanes [reaction (1)]. Comparison of reaction rates of

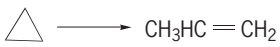

(1)

compounds with different ring size or *cis-trans* configuration has provided important insights about reaction mechanism and conformational analysis. *See* ORGANIC REACTION MECHANISM.

Ring-forming reactions. A number of useful methods have been devised that lead to alicyclic rings. These reactions are of three types: C-C bond formation between atoms in an open-chain precursor,

cycloaddition or cyclooligomerization, and expansion or contraction of a more readily available ring. In cycloaddition, two molecules react with formation of two bonds; in cyclooligomerization, three or more molecules combine to form three or more bonds.

Three-membered rings. Cyclopropane can be prepared by dehalogenation of 1-bromo-3-chloropropane (**14**) with zinc (Zn), as in reaction (2). A more general

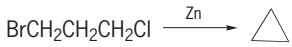

(2)

(**14**)

approach is cycloaddition of a carbene or carbenoid reagent to a double bond, as in reactions (3).

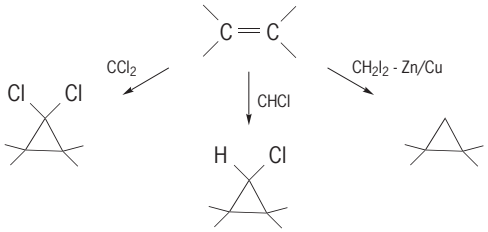

(3)

Four-membered rings. One of the important routes to cyclobutanes is cycloaddition of two double bonds. These reactions are facile with allenes, fluoroalkenes, and other alkenes in which a free-radical intermediate can be stabilized, as in reaction (4). 2+2 Cycload-

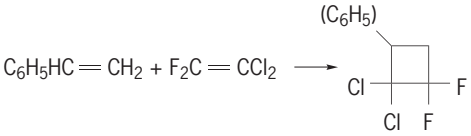

(4)

ditions (in which a two-atom molecule combines with a two-atom partner to give a four-atom ring) also occur photochemically [reaction (5)]. Four-

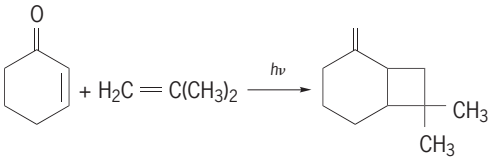

(5)

membered rings can also be obtained by ring contraction reactions of cyclopentanes.

Five- and six-membered rings. Compounds with these rings are available in quantity, and there are many ways to interconvert them. Intramolecular reactions between groups separated by four or five atoms are highly favorable; an example is the cationic cyclization in the biosynthesis of terpenes and steroids. Another very general route to cyclohexane derivatives is Diels-Alder cycloaddition. *See* BIOSYNTHESIS; DIELS-ALDER REACTION.

Seven-membered rings. Cycloheptane derivatives can be obtained by ring expansion reactions such as those in reactions (6).

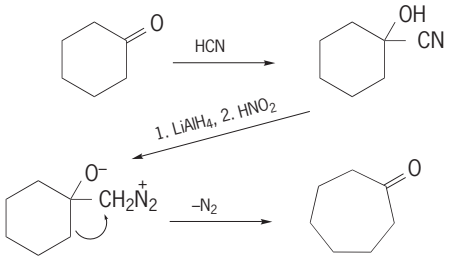

(6)

Eight-membered rings. The compound of central importance in this series is cyclooctatetraene (**15**), which is derived from four molecules of acetylene by cyclooligomerization with a nickel (Ni) catalyst [reaction (7)]. Similarly, dimerization of 1,3-dienes

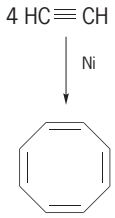

(7)

(**15**)

gives cycloocta-1,5-dienes [reaction (8)] and trimerization leads to cyclododecatrienes [reaction (9)].

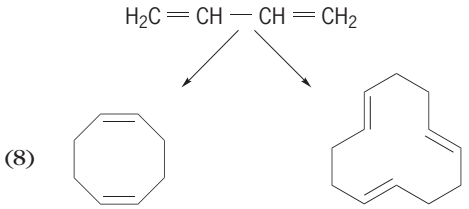

Cyclooctatetraene is a highly reactive, nonaromatic polyene that exists in a tub conformation. In several reactions, the products are those arising from the bicyclic valence tautomer (an isomer formed rapidly and reversibly by changes in valence bonds), as in reaction (10).

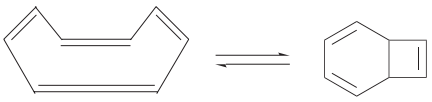

(10)

See TAUTOMERISM.

Larger rings. Ring closure between ester or nitrile groups at the ends of a chain by base-catalyzed condensation can be used to obtain rings of 14 atoms or more, as in reaction (11), where Et = C_2H_5. Very

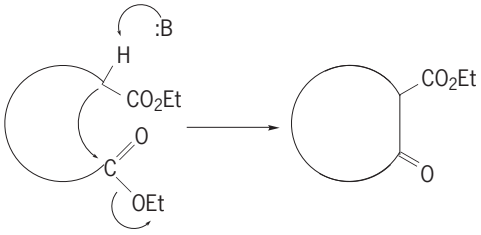

(11)

dilute solutions must be used in order to minimize

intermolecular reaction. The yields of medium rings (9–12 ring atoms) by these reactions are negligible, however, because the geometry necessary for such an enolate addition to the ester is unfavorable. A useful method for all rings is known as the acyloin condensation of a diester on molten sodium (Na), as in reaction (12), where R = an organic group. In this

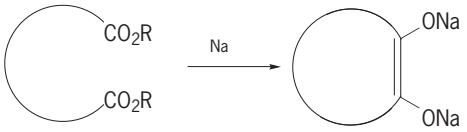

(12)

process, the ends of the chain can be aligned by moving on the sodium surface.

Major compounds. Two alicyclic compounds are manufactured in large volume; both are products of the petroleum industry. Cyclopentadiene (**16**) is formed from various alkylcyclopentanes in naphtha fractions during the refining process. It is a highly reactive diene and spontaneously dimerizes in a $4 + 2$ cycloaddition [reaction (13)]; it is used as a copolymer in several resins.

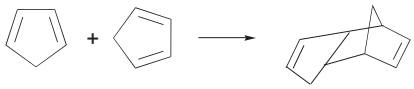

(13)

(**16**)

Cyclohexane is produced in large quantity by hydrogenation of benzene. The principal use of cyclohexane is conversion by oxidation in air to a mixture of cyclohexanol and the ketone, which is then oxidized further to adipic acid for the manufacture of nylon. *See* ORGANIC SYNTHESIS.

James A. Moore

Bibliography. F. A. Carey and R. J. Sundberg, *Advanced Organic Chemistry: Structure and Mechanisms (Part A)*, 4th ed., 2004; M. Grossel, *Alicyclic Chemistry*, 1997; Z. Rappaport and S. Patai (eds.), *Chemistry of Alkanes and Cycloalkanes*, 1992.

Alismatales

A small order of flowering plants, division Magnoliophyta (Angiospermae), which gives its name to the subclass Alismatidae of the class Liliopsida (monocotyledons). It consists of three families (Alismataceae, Butomaceae, and Limnocharitaceae) and less than a hundred species. They are aquatic and semiaquatic herbs with a well-developed, biseriate perianth that is usually differentiated into three sepals and three petals, and with a gynoecium of several or many, more or less separate carpels. Each flower is usually subtended by a bract. *Butomus umbellatus* (flowering rush) and species of *Sagittaria* (arrowhead, family Alismataceae) of this order are sometimes cultivated as ornamentals.

The Alismatales are usually regarded as the most primitive existing group of monocotyledons, but because of their trinucleate pollen and nonendospermous seeds they cannot be considered as directly ancestral to the rest of the group. Instead they appear to form a near-basal side branch which has undergone its own sort of specialization while remaining relatively primitive in gross floral structure.

The Alismatales and some related orders have often been treated as a single order Helobiae or Helobiales, embracing most of what is here treated as the subclass Alismatidae. *See* ALISMATIDAE; LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM.

Arthur Cronquist

Alismatidae

A relatively primitive subclass of the class Liliopsida (monocotyledons) of the division Magnoliophyta (Angiospermae), the flowering plants, consisting of 4 orders, 16 families, and less than 500 species. Typically they are aquatic or semiaquatic, with apocarpous flowers and nonendospermous seeds. They have trinucleate pollen, and the stomates usually have two subsidiary cells. The orders Alismatales, Hydrocharitales, and Najadales are closely related among themselves and have often been treated as a single order, Helobiae or Helobiales. The Triuridales differ from the other orders in being terrestrial and mycotrophic, without chlorophyll, and in having abundant endosperm in the seeds. *See* ALISMATALES; HYDROCHARITALES; TRIURIDALES; MAGNOLIOPHYTA; NAJADALES; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Alkali

Any compound having highly basic properties, strong acrid taste, and ability to neutralize acids. Aqueous solutions of alkalies are high in hydroxyl ions, have a pH above 7, and turn litmus paper from red to blue. Caustic alkalies include sodium hydroxide (caustic soda), the sixth-largest-volume chemical produced in the United States, and potassium hydroxide. They are extremely destructive to human tissue; external burns should be washed with large amounts of water. The milder alkalies are the carbonates of the alkali metals; these include the industrially important sodium carbonate (soda ash) and potassium carbonate (potash), as well as the carbonates of lithium, rubidium, and cesium, and the volatile ammonium hydroxide. Sodium bicarbonate is a still milder alkaline material. *See* ACID AND BASE; PH.

Soda ash and potash washed from the ashes of wood or other biomass fires were among the first manufactured chemicals, and they produced crude soap when reacted with fats. Today, the manufacture of soap and synthetic detergents still requires large amounts of alkalies. *See* DETERGENT; SOAP.

Synthetic soda ash had been made from salt since about 1840, with the Solvay process, using limestone and ammonia, predominating since 1950. Beginning about 1940, large sources of mineral deposits of impure soda ash crystallized from ancient lakes were mined and purified in increasing quantities; however, in the 1980s, production of synthetic and natural soda ash became about equal.

Until about 1900, all caustic soda was made chemically from soda ash by action with slaked lime, and this method was used for most of the soda ash manufactured until about World War II. Meanwhile, the availability of inexpensive electric power had allowed the commercial development of salt electrolysis. This process yields (at the cathode) a 10–12% sodium hydroxide solution and hydrogen gas and chlorine gas (at the anode). Since demand for chlorine increased rapidly after 1940, large quantities of its chemical equivalent, caustic soda, were being produced and sold. Subsequent industrial problems caused wide market swings involving the replacement of soda ash by caustic soda, and the marketing and pricing of soda ash, caustic soda, and chlorine.

About 50% of the caustic soda produced goes into making many chemical products, about 16% into pulp and paper, 6.5% each into aluminum, petroleum, and textiles (including rayon), with smaller percentages into soap and synthetic detergents, and cellophane. For soda ash, about 50% goes to react mainly with sand in making glass, 25% to making miscellaneous chemicals, 6.5% each to alkaline cleaners and pulp and paper, and a few percent to water treatment and other uses. *See* ALKALI METALS; ELECTROCHEMICAL PROCESS; GLASS; HYDROXIDE; PAPER; SOAP; TEXTILE CHEMISTRY.

Donald F. Othmer

Alkali emissions

Light emissions in the upper atmosphere from elemental lithium, potassium, and especially sodium. These alkali metals are present in the upper atmosphere at altitudes from about 50 to 62 mi (80 to 100 km) and are very efficient in resonant scattering of sunlight. The vertical column contents (number of atoms per square meter) of the alkali atoms are easily deduced from their respective emission intensities. First detected with ground-based spectrographs, the emissions were observed mainly at twilight since they tend to be overwhelmed by intense scattered sunlight present in the daytime. A chemiluminescent process gives rise to so-called nightglow emissions at the same wavelengths. The development of lidars (laser radars) that are tuned to the resonance lines have enabled accurate resolution of the concentrations of these elements versus altitude for any time of the day. *See* AERONOMY; CHEMILUMINESCENCE.

There is little doubt that the origin of these metals is meteoritic ablation. Rocket-borne mass spectrometers have found that meteoritic ions are prevalent above the peaks of neutral atoms with a composition similar to that found in carbonaceous chondrites, a common form of meteorites. The ratio of the concentrations of ions to neutral atoms rises rapidly with altitude above 55 mi (90 km). *See* METEORITE.

Sodium emission. Sodium (Na) is the most abundant alkali metal in meteorites. The sodium D-lines, a doublet at 589 and 589.6 nanometers, were first detected in the nightglow in the late 1920s and at twilight a decade later. The nominal peak concentration of sodium is 3×10^9 atoms m^{-3} near 55 mi (90 km) where the total gas concentration of the atmosphere is 7×10^{19} atoms (or molecules) m^{-3}. The sodium concentration declines by half at 3 mi (5 km) above and below the peak. The altitude of the maximum concentration for the sodium layer is in accord with the fact that the bulk of the micrometeorites ablate at this height, since their mean velocity is about 22 mi/s (35 km/s).

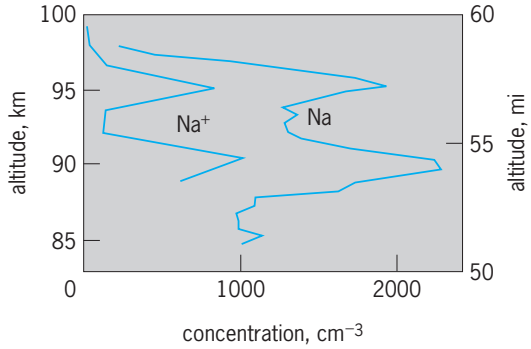

Fig. 1. Summer observations of sodium (Na) and sodium ion (Na^+) at high latitudes. (*After W. Swinder, Sodium chemistry: A brief review and two new mechanisms for sudden sodium layers, Planet. Space Sci., 40: 247–253, 1992*)

Mesospheric dynamics commonly distorts the sodium profile. Lidar measurements have provided the best evidence of these distortions and will allow scientists to achieve a better understanding of the dynamics of the mesosphere. *See* MICROMETEORITE.

The discovery of the occasional appearance of so-called sudden sodium layers (**Fig. 1**) is indicative of the complex dynamics and chemistry in the upper mesosphere. While some fresh meteor trails may conceivably contribute to such narrow layers of neutral meteor atoms, it is likely that the principal mechanism for sudden sodium layers originates with the neutralization of meteoritic ions, which can be layered by a process involving the winds and the Earth's magnetic field.

Lidar has established that the vertical column content of sodium, about 3×10^{13} atoms m^{-2} at mid-latitudes during equinox, varies very little diurnally. Twilight spectroscopic data, confirmed by the more recent lidar data, have long indicated that there is a seasonal variation in sodium, with more present in winter. The variation increases with latitude, about a factor of 2 near 20° latitude and a factor of 5 near 55° (**Fig. 2**). Data for polar latitudes are scarce. This pattern is thought to relate to the chemistry. Thus, the main depletion process for sodium, reaction (1),

$$Na + O_2 + M \rightarrow NaO_2 + M \qquad (1)$$

where M is either oxygen (O_2) or nitrogen (N_2), proceeds more slowly as temperature increases. The temperature at mesospheric altitudes is greater in winter than in summer. *See* LIDAR.

Other factors may be more or equally important, like the seasonal variation of atomic oxygen. The decline of sodium below its peak would be even

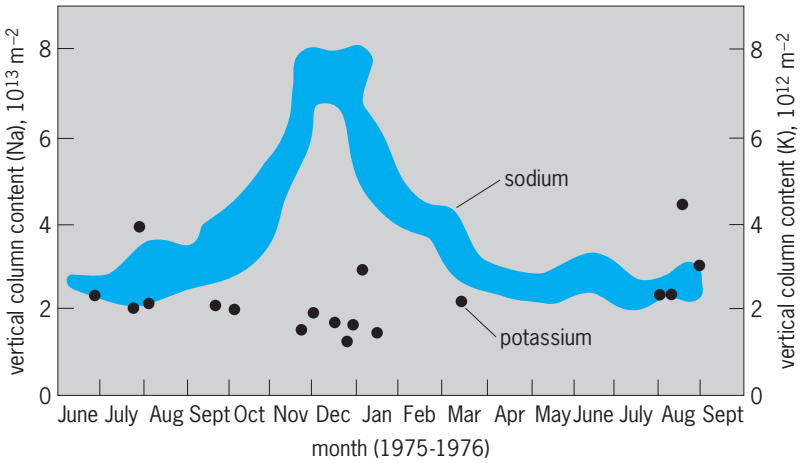

Fig. 2. Seasonal patterns of the vertical column contents of sodium and potassium at 44° latitude. (*After W. Swider, Chemistry of mesospheric potassium and its different seasonal behavior as compared to sodium, J. Geophys. Res., 92(D5):5621–5626, 1987*)

sharper if it were not for reactions (2*a*) and (2*b*).

$$NaO_2 + O \rightarrow NaO + O_2 \qquad (2a)$$

$$NaO + O \rightarrow Na + O_2 \qquad (2b)$$

Process (2*b*) can leave sodium in an excited˜state. Relaxation of the excited atoms to the ground state produces the sodium D-lines. The father of aeronomy, S. Chapman, proposed this nightglow mechanism a half century ago. Current models of the nightglow yield emission intensities compatible with observations.

Potassium emission. Potassium (K) is 15 times less abundant than sodium in meteorites. The ratio of the potassium and sodium column contents in the mesosphere ranges from 1/10 to 1/100 because, unlike sodium, the column content of potassium, 2×10^{12} m^{-2}, is fairly constant with season (Fig. 2). Once thought related to a possible additional source for sodium like sea spray, this variation may have a chemical origin. Although specific reactions may be equivalent, the particular rate coefficients may differ. For example, the rate coefficient for the formation of potassium superoxide (KO_2) is about twice that for sodium superoxide (NaO_2), and is less strongly dependent on temperature. The potassium doublet at 767 and 770 nm is estimated to have a nightglow intensity near the night sky background, 50 times smaller than the typical intensity of the sodium nightglow. The potassium nightglow has never been detected.

Lithium emission. Lithium (Li) is 35 times less abundant in meteorites than potassium, and 500 times less abundant than sodium. The meteoric source of lithium can be swamped at times by other sources, such as nuclear explosions, volcanic eruptions, and deliberate releases of lithium for aeronomic investigations of the upper atmosphere. Nevertheless, the lithium emission at 671 nm has been observed at twilight by spectrometers and at night by lidars, which indicate that the vertical content, 10^{10} m^{-2} at equinox, increases in winter like sodium. The lithium nightglow emission is undetectable. *See* ALKALI METALS; ATMOSPHERIC CHEMISTRY; MESOSPHERE.

William Swider

Bibliography. P. P. Batista et al., Characteristics of the sporadic sodium layers observed at 23°S, *J. Geophys. Res.*, 94:15349–15358, 1989; J. M. C. Plane, The chemistry of meteoric metals in the Earth's upper atmosphere, *Int. Rev. Phys. Chem.*, 10:55–106, 1991; W. Swider, Sodium chemistry: A brief review and two new mechanisms for sudden sodium layers, *Planet. Space Sci.*, 40:247–253, 1992.

Alkali metals

The elements of group 1 in the periodic table. Of the alkali metals, lithium differs most from the rest of the group, and tends to resemble the alkaline-earth metals (group 2 of the periodic table) in many ways. In this respect lithium behaves as do many other elements that are the first members of groups in the periodic table; these tend to resemble the elements in the group to the right rather than those in the same group. Francium, the heaviest of the alkali-metal

Isotopes of the alkali metals

Element	Normal at. wt	Mass no.	Radio-active	Half-life*
Lithium, Li	6.939	5	Yes	10^{-21} s
		6	No	Stable (7.5)
		7	No	Stable (92.5)
		8	Yes	0.83 s
		9	Yes	0.17 s
Sodium, Na	22.9898	20	Yes	0.23 s
		21	Yes	23.0 s
		22	Yes	2.6 y
		23	No	Stable (100)
		24	Yes	15.0 h
		25	Yes	60 s
Potassium, K	39.102	37	Yes	1.2 s
		38	Yes	7.7 m
		39	No	Stable (93.1)
		40	Yes	1.2×10^9 y
		41	No	Stable (6.9)
		42	Yes	12.4 h
		43	Yes	22 h
		44	Yes	27 m
Rubidium, Rb	85.47	81	Yes	4.7 h
		82	Yes	6.3 h
		83	Yes	80 d
		84	Yes	23 m
		85	No	Stable (72.2)
		86	Yes	19 d
		87	Yes	6.2×10^{10} y (27.8)
		88	Yes	18 m
		89	Yes	15 m
		90	Yes	2.7 m
		91	Yes	14 m
Cesium, Cs	132.905	127–132	Yes	Short
		133	No	Stable (100)
		134	Yes	3×10^6 y
		135	Yes	2.3 y
		136	Yes	1.3 d
		137	Yes	37 y
		138–145	Yes	Short
Francium, Fr		223	Yes	21 m

*Figures in parentheses indicate the percentage occurrence in nature.

elements, has no stable isotopes and exists only in radioactive form (see **table**). *See* PERIODIC TABLE.

In general, the alkali metals are soft, low-melting, reactive metals. This reactivity accounts for the fact that they are never found uncombined in nature but are always in chemical combination with other elements. This reactivity also accounts for the fact that they have no utility as structural metals (with the possible exception of lithium in alloys) and that they are used as chemical reactants in industry rather than as metals in the usual sense. The reactivity in the alkali-metal series increases in general with increase in atomic weight from lithium to cesium. *See* CESIUM; ELECTROCHEMICAL SERIES; FRANCIUM; LITHIUM; POTASSIUM; RUBIDIUM; SODIUM. Marshall Sittig

Bibliography. H. V. Borgstedt and C. K. Matthews, *Applied Chemistry of the Alkali Metals*, 1987; J. Bowser, *Inorganic Chemistry*, 1993; A. S. Kertes and C. A. Vincent (eds.), *Alkali Metal, Alkaline-Earth Metal and Ammonium Halides in Amide Solvents*, 1980.

Alkaline-earth metals

Usually calcium, strontium, and barium, the heaviest members of group 2 of the periodic table (excepting radium). Other members of the group are beryllium, magnesium, and radium, sometimes included among the alkaline-earth metals. Beryllium resembles aluminum more than any other element, and magnesium behaves more like zinc and cadmium. The gap between beryllium and magnesium and the remainder of the elements of group 2 makes it desirable to discuss these elements separately. Radium is often treated separately because of its radioactivity.

J. J. Berzelius first reduced the three alkaline-earth metals to the elementary state, but obtained them as amalgams by electrolysis. Humphry Davy in 1808 isolated the metals in the pure state by distillation of amalgams produced electrolytically. Today industrial preparation of these elements involves electrolysis of their molten chlorides or reduction of their oxides with aluminum.

The alkaline earths form a closely related group of highly metallic elements in which there is a regular gradation of properties. The metals, none of which occurs free in nature, are all harder than potassium or sodium, softer than magnesium or beryllium, and about as hard as lead. The metals are somewhat brittle, but are malleable, extrudable, and machinable. They conduct electricity well; the specific conductivity of calcium is 45% of that of silver. The oxidation potentials of the triad are as great as those of the alkali metals.

The alkaline earths exist as large divalent cations in all their compounds, in which the elements are present in the 2+ oxidation state. The metals have a gray-white luster when cut but tarnish readily in air. They burn brilliantly in air when heated, and form the metal monoxide, except for barium, which forms the peroxide. A certain amount of nitride is formed simultaneously, especially with calcium. All the metals dissolve readily in acid. Whereas calcium reacts smoothly with water to yield hydrogen, the heavier members react as violently as sodium does. All the metals are soluble in liquid ammonia, yielding strongly reducing, electrically conducting, blue solutions. The order of solubility in water for most salts is calcium > strontium > barium, except that the order is reversed for the fluorides, hydroxides, and oxalates. All three elements unite directly with hydrogen to form hydrides and with nitrogen to form nitrides, but whereas ease of formation of a nitride increases with atomic number, ease of formation of a hydride decreases.

The elements and their compounds find important industrial uses in low-melting alloys, deoxidizers, and drying agents and as cheap sources of alkalinity. *See* BARIUM; BERYLLIUM; CALCIUM; MAGNESIUM; PERIODIC TABLE; RADIUM; STRONTIUM. Reed F. Riley

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; T. R. Dickson, *Introduction to Chemistry*, 7th ed., 1995; K. Mackay, *Introduction to Modern Inorganic Chemistry*, 6th ed., 2002.

Alkaloid

A cyclic organic compound that contains nitrogen in a negative oxidation state and is of limited distribution among living organisms. Over 10,000 alkaloids of many different structural types are known; and no other class of natural products possesses such an enormous variety of structures.

Therefore, alkaloids are difficult to differentiate from other types of organic nitrogen-containing compounds.

Simple low-molecular-weight derivatives of ammonia, as well as polyamines and acyclic amides, are not considered alkaloids because they lack a cyclic structure in some part of the molecule. Amines, amine oxides, amides, and quaternary ammonium salts are included in the alkaloid group because their nitrogen is in a negative oxidation state (the oxidation state designates the positive or negative character of atoms in a molecule). Nitro and nitroso compounds are excluded as alkaloids. The almost-ubiquitous nitrogenous compounds, such as amino acids, amino sugars, peptides, proteins, nucleic acids, nucleotides, prophyrins, and vitamins, are not alkaloids. However, compounds that are exceptions to the classical-type definition (that is, a compound containing nitrogen, usually a cyclic amine, and occurring as a secondary metabolite), such as neutral alkaloids (colchicine, piperine), the β-phenyl-ethylanines, and the purine bases (caffeine, theophylline, theobromine), are accepted as alkaloids.

Nomenclature. The nomenclature of alkaloids has not been systematized. Most alkaloids bear the suffix -ine and their names are derived in various ways: from the generic name of the plant, for example hydrastine from *Hydrastis canadensis*; from the specific plant name, for example cocaine from *Erythroxylum coca*; from the common name of the drug yielding

them, for example, ergotamine from ergot; or from a particular physiological activity exhibited, for example, morphine from Morpheus, Greek god of dreams. Only the pelletierine group is named for a person, P. J. Pelletier, who discovered a number of alkaloids in the first third of the nineteenth century.

Occurrence. Alkaloids are derived from both plants and animals, including marine organisms.

Plant-derived alkaloids. Alkaloids often occur as salts of plant acids such as malic, meconic, and quinic acids. Some plant alkaloids are combined with sugars, for example, solanine in potato (*Solanum tuberosum*) and tomatine in tomato (*Lycopersicum esculentum*). Others occur as amides, for example, piperine from black pepper (*Piper nigrum*), or as esters, for example, cocaine from coca leaves (*Erythroxylum coca*). Still other alkaloids occur as quaternary salts or tertiary amine oxides.

Although about 40% of all plant families contain at least one alkaloid-bearing species, only about 9% of the over 10,000 plant genera produce alkaloids. Among the angiosperms, alkaloids occur in abundance in certain dicotyledons and especially in the families Apocynaceae (dogbane, quebracho), Asteraceae (groundsel, ragwort), Berberidaceae (European barberry), Fabaceae (broom, gorse, lupine), Loganiaceae (*Strychnos* species), Menispermaceae (moonseed), Papaveraceae (poppies, chelidonium), Ranunculaceae (aconite, larkspur), Rubiaceae (cinchona bark, ipecac), Rutaceae (citrus, fagara), and Solanaceae (tobacco, deadly nightshade, tomato, potato, thorn apple). Rarely are they present in cryptogamia, gymnosperms, or monocotyledons. Among the monocotyledons, the Amaryllidaceae (amaryllis, narcissus) and Liliaceae (meadow saffron, veratrums) are families that bear alkaloids.

Examples of well-known alkaloids are morphine and codeine from the opium poppy (*Papaver somniferum*), strychnine from *Strychnos* species, quinine (**1**) from *Cinchona* bark and various *Cinchona*

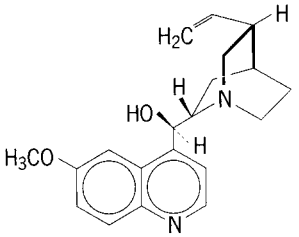

(1)

species, and coniine from poison hemlock (*Conium maculatum*). Other important alkaloids are colchicine from the autumn crocus (*Colchicum autumnale*), caffeine from coffee, tea, kola, and maté, and nicotine from the tobacco plant (*Nicotiana tabacum*) and Aztec tobacco (*N. rustica*). *See* CAFFEINE; COFFEE; COLA; COLCHICINE; NICOTINE ALKALOIDS.

Animal-derived alkaloids. While most alkaloids have been isolated from plants, a large number have been isolated from animal sources. They occur in mammals, anurans (frogs, toads), salamanders, arthropods (ants, millipedes, ladybugs, beetles, butterflies), marine organisms, mosses, fungi, and certain bacteria. Samandarine has been isolated from the European fire salamander (*Salamandra maculosa*). Very toxic steroidal alkaloids occur in certain frogs and toads; for example, the Columbian arrow-poison frog (*Phyllobates aurotaenia*) produces a highly lethal venom containing batrachotoxin. The Canadian beaver (*Castor fiberi*) produces castoramine, and the scent gland of the musk deer (*Moschus moschiferus*) produces muscopyridine. The Southern fire ant (*Solenopsis invicta*) secretes a powerful venom containing a series of 2,6-dialkylpiperidines. The European millipede (*Glomeris marginata*) when provoked discharges a defensive secretion containing glomerine, a quinazolone. When molested, the European ladybug secretes hemolymph at the joints to afford protection against predators; this hemolymph contains precoccinelline and its oxide, coccinelline. Certain species of butterflies elaborate pheromones derived from exogenous pyrrolizidine alkaloid precursors; the pyrrolizidine alkaloid occurs in the secretions of *Lycorea ceres ceres* and *Danaus gilippus berenice*, where it elicits olfactory-receptor responses in female antennae and serves as a chemical messenger that induces mating behavior. *See* PHEROMONE.

Marine-derived alkaloids. Many alkaloids have been isolated from marine organisms, both plant and animal. Of great interest are the alkaloids occurring in certain toxic dinoflagellates, such as *Gonyaux tamarensis* in the North Atlantic and *G. cantanella* in the South Atlantic. The toxic principles include the alkaloids saxitoxin, gonyautoxin-II, and gonyautoxin-III. These paralytic alkaloids also occur in the Alaska butter clam (*Saxidomus giganteus*) and toxic mussels (*Mytilus californianus*). Tetrodotoxin, present in the ovaries and liver of the Japanese puffer fish (*Spheroides rubripes* and *S. vermicularis*), in the Taiwanese goby fish (*Gobius cringer*), and in the California newt (*Taricha torosa*), is one of the most lethal low-molecular-weight toxins known.

Isolation and purification. Since few plant species produce only a single alkaloid, there is the problem of separating complex mixtures. The alkaloid mixture is usually extracted from the finely powdered plant material by percolation with some solvent, such as methanol or aqueous ethanol. Evaporation leaves a residue, which is then dissolved in an organic solvent such as chloroform, methylene chloride, or toluene, and filtered. The solution is extracted repeatedly first with dilute sulfuric or hydrochloric acid and, after basification, with an organic solvent. These extracts when evaporated yield different mixtures of alkaloids which may be further separated by column chromatography over alumina or silica gel, vacuum liquid chromatography, countercurrent distribution, centrifugally accelerated radial thin-layer chromatography, or preparative-layer chromatography. Eventually, it is possible to obtain compounds in a high state of purity. *See* CHEMICAL SEPARATION TECHNIQUES; CHROMATOGRAPHY.

Biosynthesis. Of great interest to chemists is the mode of alkaloid synthesis in plants. The

biosynthesis of most alkaloids can be derived hypothetically, by using a few well-known chemical reactions, from such relatively simple precursors as phenylalanine, tyrosine, tryptophan, histidine, methionine, acetate units, terpene units, and amino acids such as anthranilic acid, lysine, and ornithine. A number of simple alkaloids have been synthesized from amino acid derivatives under physiological conditions: pelletierine is derived from the amino acid lysine, while cocaine and nicotine are derived from ornithine (**2**). Many alkaloids, such as

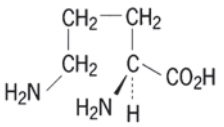

(**2**)

morphine, berberine, and emetine, incorporate a tetrahydroisoquinoline unit derived from tyrosine; and others, such as strychnine, quinine, reserpine, and camptothecin, are derived from tryptophan. The terpenoid and steroidal alkaloids, such as the diterpenoid alkaloid aconitine and the steroidal alkaloid tomatidine from the tomato, are likely derived from mevalonic acid lactone. *See* BIOSYNTHESIS; STEROID; TERPENE.

Function. The exact function of alkaloids in plants is not well understood, but they are often regarded as by-products of plant metabolism. They are sometimes considered to be reservoirs for protein synthesis; as protective agents to discourage attack by animals, insects, or microbes; as regulators of growth, metabolism, and reproduction; and as end products of detoxification of substances whose accumulation might be injurious to the plant.

Medicinals. Many alkaloids exhibit marked pharmacological activity, and some find important uses in medicine. Atropine, the optically inactive form of hyoscyamine, is used widely in medicine as an antidote to cholinesterase inhibitors such as physostigmine and insecticides of the organophosphate type; it is also used in drying cough secretions. Morphine and codeine are narcotic analgesics, and codeine is also an antitussive agent, less toxic and less habit forming than morphine. Colchicine, from the corms and seeds of the autumn crocus, is used as a gout suppressant. Caffeine, which occurs in coffee, tea, cocoa, and cola, is a central nervous system stimulant; it is used as a cardiac and respiratory stimulant and as an antidote to barbiturate and morphine poisoning. Emetine, the key alkaloid of ipecac root (*Cephaelis ipecacuanha*), is used in the treatment of amebic dysentery and other protozoal infections. Epinephrine (**3**), or adrenaline, produced in most

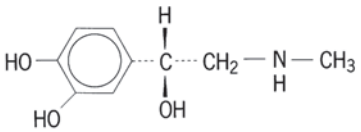

(**3**)

animal species by the adrenal medulla, is used as a bronchodilator and cardiac stimulant and to counter allergic reactions, anesthesia, and cardiac arrest.

Pilocarpine, from *Pilocarpus jaborandi*, in hydrochloride or nitrate form is cholinergic and is used topically in the treatment of glaucoma; it is also used orally or subcutaneously to stimulate secretion of saliva in individuals who undergo therapy with ganglionic blocking agents. Quinine, used in many areas of the world to treat malaria, is used in the United States to treat chloroquine-resistant falciparum malaria. Quinidine, a stereoisomer of quinine, is an antiarrhythmic agent used to control auricular fibrillation. Reserpine, from *Rauwolfia serpentina*, produces sedation, tranquilization, and depression of blood pressure; it is sometimes used to decrease high blood pressure and as a tranquilizer in cases of nervousness, hysteria, and stress. Taxol, from the Pacific yew tree (*Taxus brevifolia*), is used in the treatment of ovarian and breast cancer. Tubocurarine chloride (curare, from *Strychnos* and *Chondodendron* species) is employed as a skelet al muscle relaxant during surgery and to control convulsions resulting from strychnine poisoning and tetanus. Vinblastine and vincristine, from the Madagascar periwinkle plant (*Catharanthus roseus*), are potent anticancer drugs; the former is used to treat generalized Hodgkin's disease and choriocarcinoma, while the latter is used to treat acute leukemia. *See* CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; EPINEPHRINE; MORPHINE ALKALOIDS; PHARMACOGNOSY; QUININE; STRYCHNINE ALKALOIDS.

S. William Pelletier

Bibliography. A. Brossi and G. A. Cordell (eds.), *The Alkaloids: Chemistry and Physiology*, vol. 42, 1992; G. A. Cordell, *Introduction to Alkaloids: A Biogenetic Approach*, 1981; D. R. Dalton, *The Alkaloids: The Fundamental Chemistry—A Biogenetic Approach*, 1979; K. Mothes, H. R. Shutte, and M. Luckner, *Biochemistry of Alkaloids*, 1985; S. W. Pelletier (ed.), *Alkaloids: Chemical and Biological Perspectives*, vols. 1–9, 1983–1995; S. W. Pelletier, The nature and definition of an alkaloid, *Alkaloids: Chemical and Biological Perspectives*, vol. 1, 1983; T. Robinson, *The Biochemistry of Alkaloids*, 2d ed., 1981; I. W. Southon and J. Buckingham (eds.), *Dictionary of Alkaloids*, 1989.

Alkane

An organic compound with the general formula C_nH_{2n+2}. Alkanes are open-chain (aliphatic or noncyclic) hydrocarbons with no multiple bonds or functional groups. They consist of tetrahedral carbon atoms, up to 10^5 carbons or more in length. The C-C σ bonds are formed from sp^3 orbitals, and there is free rotation around the bond axis. *See* MOLECULAR ORBITAL THEORY; STRUCTURAL CHEMISTRY.

Alkanes provide the parent names for all other aliphatic compounds in systematic nomenclature. Alkanes are designated by the ending -ane appended to a stem denoting the chain length. The straight-chain isomer is designated by the prefix *n*(normal); other isomers are named by specifying the size of the branch and its location (see **table**). The number of isomers increases enormously in larger molecules;

Nomenclature and properties of alkanes

Name	Structure	Boiling point, °C (°F)	Heat of formation (ΔH_f°), kJ
Methane	CH_4	−162 (−260)	−74.5
Ethane	CH_3CH_3	−89 (−128)	−83.45
Propane	$CH_3CH_2CH_3$	−42 (−44)	−104.6
n-Butane	$CH_3(CH_2)_2CH_3$	−0.5 (33)	−125.7
2-Methylpropane (isobutane)	$(CH_3)_2CHCH_3$	−11.7 (10.9)	−134.2
n-Pentane	$CH_3(CH_2)_3CH_3$	36.1 (97.0)	−146.87
2-Methylbutane (isopentane)	$(CH_3)_2CHCH_2CH_3$	29.9 (85.8)	−153.7
2,2-Dimethyl propane (neopentane)	$(CH_3)_4C$	9.4 (49)	−167.9
n-Hexane	$CH_3(CH_2)_4CH_3$	68.7 (156)	−167.0

thus there are 75 isomers of $C_{10}H_{22}$ and over 4 billion for $C_{30}H_{62}$.

Properties. Alkanes with four or fewer carbons are gases at atmospheric pressure. Higher *n*-alkanes are liquids or, above about 20 carbons, solids known as paraffin wax. Alkanes have densities lower than that of water and have very low water solubility; other properties depend on the degree of branching (see table). The heat of formation (ΔH°_f) is a measure of the energy content of a compound relative to the component elements in standard states. For example, comparison of heat of formation values for three alkane isomers with the molecular formula C_5H_{12} indicates that the relative energy content of the three pentane isomers decreases with increased branching; that is, the branched isomer is thermodynamically most stable. *See* PARAFFIN.

Sources. Alkanes are the major components of natural gas and petroleum, which are the only significant sources. Much smaller amounts of alkanes have been produced from coal at various times and locations, either indirectly by the Fischer-Tropsch process or by direct liquefaction. *See* COAL LIQUEFACTION; FISCHER-TROPSCH PROCESS; NATURAL GAS; PETROLEUM.

Certain long-chain alkanes and also some alicyclic hydrocarbons are very widely distributed both in sedimentary soils and shales and among living organisms. Phytane and pristane are derived from phytol, and these alkanes are of significance as biological markers in tracing diagenesis. Both insects and plants secrete alkanes, which function as moisture barriers and protective coatings. Cuticular waxes on fruit and leaves contain predominantly the *n*-alkanes with 29, 31, and 33 carbons; the preponderance of odd-carbon chains reflects the origin by loss of carbon monoxide (CO) or carbon dioxide (CO_2) from even-carbon precursors. *See* DIAGENESIS.

Individual lower alkanes can be separated from the more volatile distillate fractions of petroleum, but beyond the C_7–C_8 range the alkanes obtained are mixtures of many isomers. Compounds of a specific structure can be prepared on a laboratory scale by chemical synthesis. Various C-C bond-forming steps such as coupling or condensation are carried out to build up the desired carbon skeleton. The final step is usually removal of a functional group by some type of reduction. Several insect pheromones, for example 15,19,23-trimethylheptatriacontane ($C_{40}H_{82}$) from the female tsetse fly, have been synthesized for biological study. *See* PHEROMONE.

Chemistry. Much of the chemistry of alkanes begins at the petroleum refinery, where several reactions are carried out to adjust the hydrocarbon composition of crude oil to that needed for a constantly changing set of applications. Major reactions are (1) isomerization of straight-chain alkanes to branched compounds; (2) cracking to produce smaller molecules; (3) alkylation, for example, combination of propylene (**1**) and butane (**2**) to give 2,3-dimethylpentane (**3**), as in reaction (1); and

$$\underset{(\mathbf{1})}{CH_3CH{=}CH_2} + \underset{(\mathbf{2})}{CH_3CH_2CH_2CH_3} \longrightarrow \underset{(\mathbf{3})}{CH_3CH(CH_3)CH(CH_3)CH(CH_3)CH_2CH_3} \quad (1)$$

(4) cyclodehydrogenation (platforming), in which aromatizations occur. An important objective in some of these processes is to increase the yield of highly branched alkanes in the C_6–C_8 range needed for gasoline. *See* ALKYLATION (PETROLEUM); AROMATIZATION; CRACKING; GASOLINE.

By far the most important end use of alkanes is combustion as fuel to provide heat and electric or motive power. In most cases, complete oxidation is not achieved, and varying amounts of incompletely oxidized fragments, carbon monoxide, and elemental carbon are produced.

Controlled partial oxidation is possible if all the C-H bonds in an alkane are equivalent or if one C-H bond is significantly weaker than all the others. An example of the latter situation is isobutane, which is converted to the hydroperoxide on industrial scale for the manufacture of *t*-butyl alcohol. *See* AUTOXIDATION; COMBUSTION.

Alkanes have been referred to as paraffin hydrocarbons to indicate their low affinity or reactivity. They contain no unshared electron pairs or accessible empty bonding orbitals, and they are unaffected by many reagents that attack π bonds or other functional groups. One type of reaction that does occur is substitution by a radical chain process. Examples are chlorination and vapor-phase nitration, involving the odd-electron species Cl· and NO_2·, respectively. Neither reaction is selective; when two or more types of C-H bonds are present in the alkane, mixtures of products are usually obtained. Thus

propane gives rise to 1- and 2-chloropropanes as well as dichloro compounds. Nitration of propane leads to a mixture of 1- and 2-nitropropane, and also nitromethane and nitroethane by C-C bond cleavage. *See* HALOGENATED HYDROCARBON; HALOGENATION; NITRATION.

Since alkanes are the most abundant class of organic compounds available, much effort has been directed to the development of methods for converting them selectively to other compounds. Contributions to this objective have come from several diverse sources. Biodegradation of petroleum hydrocarbons is a well-known phenomenon and is useful in decontamination of oil leaks and spills. Closely related is utilization of the metabolic apparatus of individual microorganisms to carry out a specific reaction, such as hydroxylation of the C-H bond at the 11 position of a steroid. Many such biochemical oxidations occur by means of a cytochrome P-450 enzyme system in which an oxygen atom is transferred from dioxygen (O_2) to an alkane at the iron atom in a heme protein. Synthetic metalloporphyrins and other iron complexes have been studied as model systems for these oxidations. *See* BIODEGRADATION; CYTOCHROME; PORPHYRIN. James A. Moore

Activation by organometallic compounds. Another contribution to the activation of alkanes is from the field of organometallic chemistry. One way that C-H bonds have been cleaved involves the formation of a vacant site on a metal in a low oxidation state that contains electron-donating ligands. The vacant site provides room for the incoming hydrocarbon, and the electron-rich metal readily gives up its electrons in a process called oxidative addition. The net reaction results in the use of two electrons from the C-H bond plus two electrons from the metal to form two metal-carbon bonds, each with two electrons. The vacant site is most easily generated by irradiation of the precursor complex with ultraviolet light, a technique that makes the ligands fall off easily at ambient temperatures. If the irradiation is carried out in a pure hydrocarbon solvent, the metal oxidatively adds to the solvent, resulting in the formation of a complex with metal-hydrogen and metal-carbon bonds. The C-H bond has been completely cleaved, and further chemistry can now be envisioned with the fragments attached to the metal. Reaction (2) shows several specific metal complexes that will undergo this type of oxidative addition chemistry with saturated and unsaturated hydrocarbons such as methane, ethane, hexane, and benzene. In reaction (2), M = rhodium (Rh) or iridium (Ir); Me = methyl group (CH_3); Me_3P = trimethylphosphine; L = PMe_3 or carbonyl (C═O); and R = CH_3, C_2H_5 (ethyl), C_6H_{13} (hexyl), or C_6H_5 (phenyl). With linear alkanes such as hexane, there is a preference of cleavage of the less hindered primary (terminal) C-H bonds at the end of the molecule over the more hindered secondary (internal) C-H bonds of the molecule.

In one application based on this reactivity, a complex has been used to convert alkanes into alkenes plus hydrogen by using light energy. The complex chlorocarbonyl bis-(trimethylphosphine) rhodium

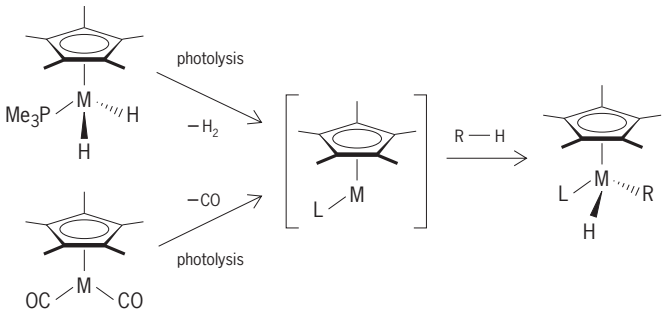

(2)

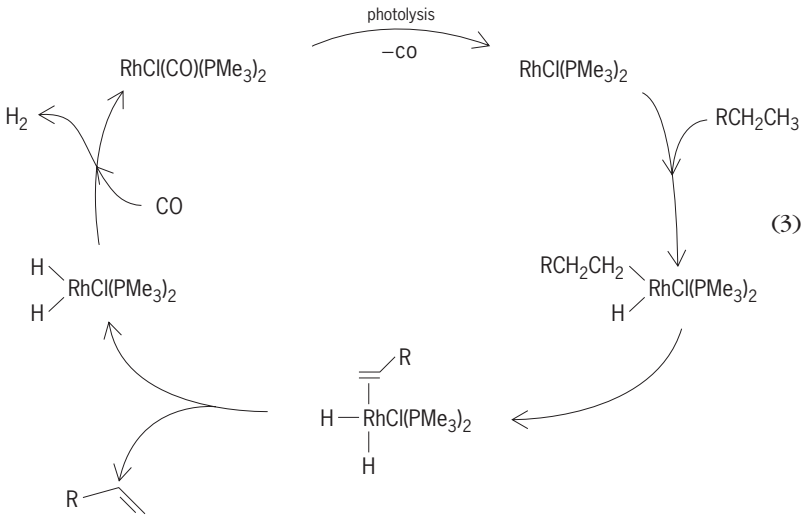

(3)

[$RhCl(CO)(PMe_3)$] reacts with alkanes to give many thousands of turnovers of alkenes upon photolysis, as in reactions (3). Beginning with the rhodium complex $RhCl(CO)(PMe_3)_2$, the fundamental step involving activation of the C-H bond can be seen in the catalytic cycle following the loss of carbon monoxide (CO). The loss of a hydrogen atom on the carbon that is adjacent to the one bound to the metal is a common reaction in organometallic chemistry, and it leads to the production of the alkene product still bound to the rhodium. The olefin is then released from the complex. Finally, hydrogen is displaced by reaction with CO to regenerate the initial metal complex. Light energy is required in this cycle to remove CO from the metal. *See* ORGANOMETALLIC COMPOUND. William D. Jones

Bibliography. J. A. Davies et al. (eds.), *Selective Hydrocarbon Activation,* 1990; R. J. Fessenden and J. S. Fessenden, *Organic Chemistry,* 6th ed., 1998.

Alkene

A hydrocarbon that contains a carbon-carbon double bond (C═C); dienes or polyenes contain two or more double bonds. The synonym olefin is often used, especially in connection with industrial-scale production and applications. The term unsaturated distinguishes compounds such as fats that contain double bonds from saturated molecules that do not.

Structure and names. The carbon atoms joined by a double bond have three planar sp^2 orbitals and a p orbital with lobes above and below. The double

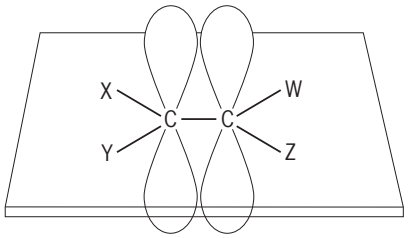

Fig. 1. Molecular orbital representation of a carbon double bond; W, X, Y, and Z represent attached groups.

bond is a combination of a single σ bond between the nuclei and a π bond resulting from an electron pair in overlapping *p* orbitals (**Fig. 1**). As a result, the attached groups *w, x, y*, and *z* are coplanar, and rotation around the C-C axis is possible only by disrupting the π bond. A double bond is shorter (0.133 nanometer) and stronger (610 kilojoules) than a typical C-C bond (0.154 nm, 347 kJ). *See* CHEMICAL BONDING; MOLECULAR ORBITAL THEORY.

The names of alkenes are formed by changing the -ane ending of the corresponding alkane to −ene, with a number to indicate the double-bond location. Since the double bond is a barrier to rotation, an alkene such as 2-butene has two stereoisomeric forms. These geometric isomers are designated as *E* or *trans* (similar groups on opposite side) or *Z* (*cis*, similar groups on same side), as in the structures below. For the simplest alkenes (C_2–C_4) the ending

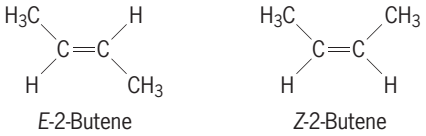

-ylene is often used. Two other common names are vinyl for the group —HC═CH_2 and allyl for the group —CH_2HC═CH_2. *See* ALKANE.

An alkene can undergo isomerization by shifts of the double bond, such as by treatment with an acid catalyst. Isomers with the double bond in different positions or geometries have different internal energies and therefore different stabilities. In most cases a *trans* isomer is favored over the *cis*, because larger groups are farther apart. Moreover, isomers in which the double bond is in an internal position on the chain are usually more stable than those with the double bond at the end. For the three straight-chain butene isomers, the equilibrium composition is shown in notation (1).

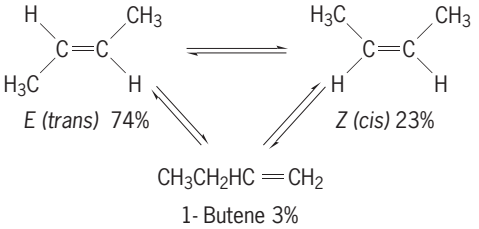

(1)

See ISOMERIZATION.

Source and preparation. Ethylene (H_2C═CH_2) and the C_3–C_5 alkenes are produced in very large volume by catalytic cracking of higher hydrocarbons or dehydrogenation of alkanes during petroleum refining. Alkenes in the $C_{10\text{-}20}$ range are important starting materials in several large-scale industrial applications. These are prepared by controlled oligomerization of ethylene with a triethylaluminum catalyst [reaction (2), where $n = 5$–10].

$$n\,HC_2{=}CH_2 \xrightarrow{(CH_3CH_2)_3Al} CH_3CH_2(CH_2CH_2)_{n-2}HC{=}CH_2 \qquad (2)$$

See CRACKING; DEHYDROGENATION.

Several general methods for preparing alkenes involve elimination reactions. Two long-known examples are removal of a hydrogen-halogen group (HX) from a halide with strong base, and loss of water from alcohols. Both reactions can often give more than one alkene, and acid-catalyzed dehydration may lead to rearrangement. In the latter case, conversion of the alcohol to an ester and pyrolysis may be preferable [reaction (3)]. Dehydration of a β-

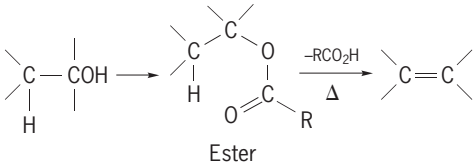

(3)

hydroxycarbonyl system is the second step in many carbonyl condensation reactions [reaction (4)].

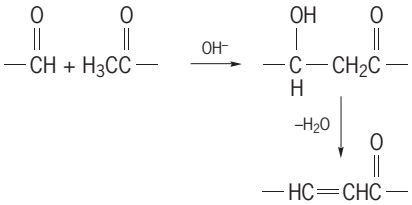

(4)

A widely used preparative method for alkenes is the Wittig reaction. In this process a phosphorane adds to a carbonyl compound, and the double bond is formed by loss of triphenylphosphine oxide from a phosphetane [reaction (5)]. Another approach to

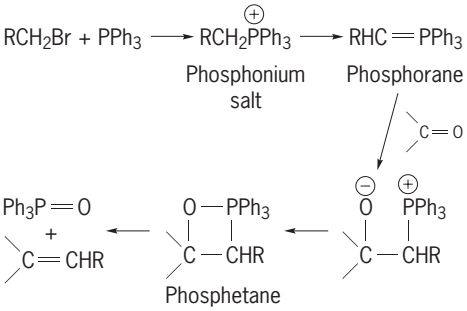

(5)

the formation of a double bond is reduction of a triple bond. *See* ALKYNE; ORGANOPHOSPHORUS COMPOUND.

Chemical activity. The double bond is a highly reactive and versatile functional group. Because of the readily accessible electrons in the π bond, alkenes react with a variety of Lewis acids, or electrophiles.

A characteristic reaction is addition, with formation of two new σ bonds in place of the π bond.

Hydrogenation. Addition of hydrogen occurs on the surface of a finely divided metal catalyst, usually nickel, palladium, or platinum. This reaction is useful for several purposes; one application is the partial hydrogenation of unsaturated vegetable oils to remove some of the double bonds and raise the melting temperature. *See* FAT AND OIL (FOOD).

Proton acids. Several types of addition products can arise from the reaction of alkenes with acids. In a simple case hydrochloric acid (HCl) or hydrobromic acid (HBr) add to propene to give the 2-halopropane; with aqueous sulfuric acid the elements of water (H_2O) add, and 2-propanol is obtained [reaction (6),

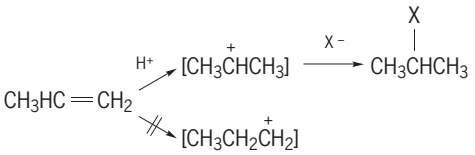

(6)

where X= chlorine (Cl), bromine (Br), or hydroxyl (OH)]. These additions occur by way of an intermediate carbenium ion that rapidly reacts with a nucleophile such as halide ion or water. The structure of the product depends on the fact that the cation with the larger number of alkyl groups attached to the electron-deficient carbon is more stable and more easily formed. *See* REACTIVE INTERMEDIATES.

The carbenium ion formed from an alkene may undergo other reactions as well. Under suitable conditions it can react as an electrophile with another molecule of the alkene to give branched dimers, trimers, and so forth [reaction (7), where R= CH_3

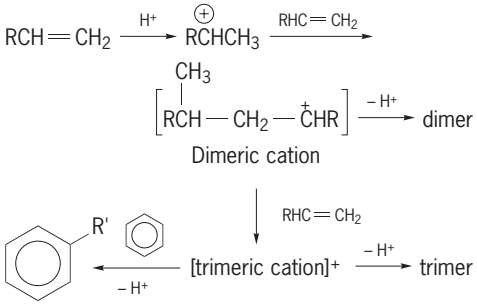

(7)

or C_2H_5 and R′ = C_9H_{18} or $C_{12}H_{24}$]. An alkene with acid can also be used to alkylate a benzene ring in a variant of the Friedel-Crafts reaction. Both processes are used in the production of alkylbenzenes for detergents and other uses. *See* FRIEDEL-CRAFTS REACTION.

Halogenation. Bromine or chlorine adds very rapidly to alkenes; bleaching of the red color of Br_2 has long been recognized as a test for a double bond. The reaction of ethylene with chlorine to give 1,2-dichloroethane is the first step in the manufacture of vinyl chloride.

In the addition of halogen to a double bond [shown in reaction (8) for Br_2], a cyclic cation is initially formed. In the second step, bromide ion attacks this intermediate at a C-Br bond. The result is *anti* addition, with the Br atoms bonded to opposite sides of the original double bond. When the reaction is carried out in an aqueous or alcoholic solution of the halogen, water or alcohol is the nucleophile in the second step and the product is the bromo alcohol or ether.

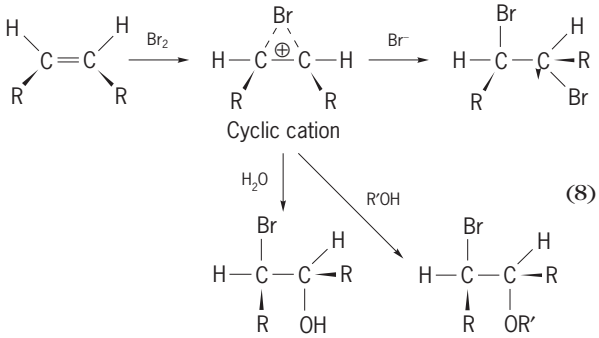

(8)

See HALOGENATED HYDROCARBON.

Oxidation. Addition of oxygen to a double bond gives an epoxide. This reaction is carried out on an industrial scale by passing an alkene and air over a silver (Ag) catalyst. In smaller-scale work, epoxides can be obtained by treating an alkene or other unsaturated compound with a peroxide or peracid [reaction (9)].

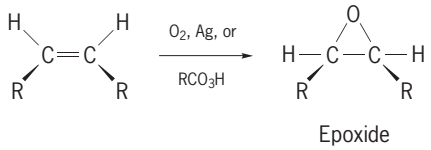

(9)

See EPOXIDE.

Alkenes can also be oxidized to diols. This reaction takes place when a transition-metal oxide such as permanganate (MnO_4) or osmium tetroxide adds to the double bond. A cyclic intermediate results, and this is hydrolyzed to give the diol [reaction (10)].

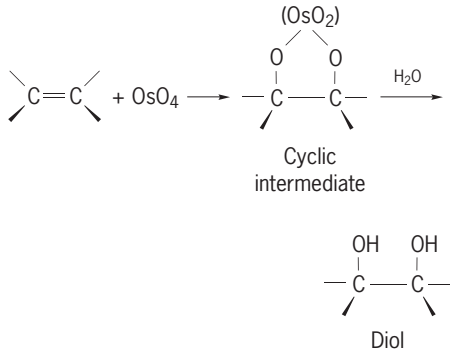

(10)

Hydroboration. Alkenes combine readily with borane (BH_3). The resulting alkyl boranes are highly reactive compounds and are useful and versatile reagents in organic synthesis. *See* HYDROBORATION.

Free radicals. Double bonds undergo addition reactions by attack of free radicals, X·. Some of the important radicals are obtained from HX compounds such as hydrogen bromide (HBr), hydrogen sulfide (H_2S), a thiol (HSR), and chloroform ($HCCl_3$). The process is initiated by homolysis of HX, often with air [reaction (11), where X=HBr, HSR, or $HCCl_3$]. A chain reaction then begins (propagation), with X· adding to the π bond [reaction (12)]. Subsequent reaction of

the intermediate radical with another HX molecule gives a product and a new X· radical to continue the chain [reaction (13)]. The direction of addition

Initiation:

$$HX \xrightarrow{O_2} X\cdot + HO_2\cdot \qquad (11)$$

Propagation:

$$RHC{=}CH_2 + X\cdot \longrightarrow R\dot{C}HCH_2X \qquad (12)$$

$$R\dot{C}H\,CH_2X + HX \longrightarrow RCH_2CH_2X + X\cdot \qquad (13)$$

is due to stabilization of the electron-deficient radical by alkyl substituents, just as in the case of a carbenium ion [reaction (6)]. When X = Br, it will be noted that addition by the radical mechanism [reactions (12) and (13)] occurs in the direction opposite to that in the ionic mechanism. The observation and clarification of this reversal of the addition of HBr was a milestone in the understanding of how reactions occur. *See* CHAIN REACTION (CHEMISTRY).

The free-radical intermediate in reaction (12) can react with HX as in reaction (13), but it can also react with another alkene molecule, or several molecules, before being capped. This possibility is shown in reaction scheme (14) for the free-radical addition of

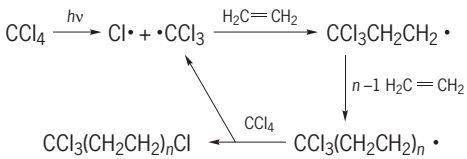

(14)

carbon tetrachloride (CCl_4) to ethylene. This reaction results in a mixture of products; the main products have the following yields; $n = 1$, 9%; $n = 2$, 57%; and $n = 3$, 24%. The overall process is called telomerization. *See* FREE RADICAL.

Reactions of metal complexes. Among the most important properties of alkenes is the formation of π complexes by donation of an electron pair to the coordination shell of a transition metal. This is a very general reaction, and the complexes are central to some of the most important chemistry, including industrial processes, involving alkenes. Chief among these is the production of polyethylene and polypropylene by polymerization of the alkenes with Ziegler-Natta catalysts, in which monomer units are incorporated into the chain at a coordination site on a transition metal such as titanium(III). *See* COORDINATION COMPLEXES; ORGANOMETALLIC COMPOUND; POLYMERIZATION; STEREOSPECIFIC CATALYST.

In complex formation, electron density shifts from the π bond to the metal, and the coordinated alkene is therefore susceptible to addition of nucleophiles such as an OH or NH group. Palladium (Pd) is a particularly versatile metal in several reactions. In the Wacker process for the production of acetaldehyde, ethylene is complexed with Pd^{II} [reaction (15)]. After addition of water, the complex undergoes β-elimination of enol, and palladium is thereby reduced

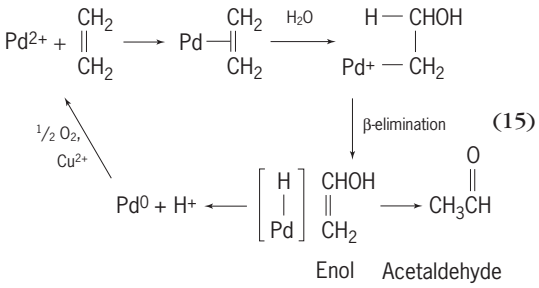

(15)

to Pd^0. It is reoxidized to Pd^{II} by copper ion (Cu^{2+}), which is in turn reoxidized by oxygen (air). The net reaction is thus a two-electron oxidation of ethylene, with both palladium and copper required in only catalytic amounts. Analogous reactions in which alcohol or acetic acid are used rather than water produce vinyl ether and vinyl acetate, respectively. Extension of this reaction to higher 1-alkenes provides a convenient and practical method to prepare methyl ketones [reaction (16)].

$$RHC{=}CH_2 \xrightarrow[H_2O]{Pd^{II},\ air} R\overset{O}{\overset{\|}{C}}CH_3 \qquad (16)$$

Methyl ketone

Another example of the utility of palladium π complexes is the substitution of an aryl (Ar) or vinyl group for a vinyl hydrogen in an alkene [reaction (17)]. In this reaction the aryl group is transferred

$$ArX + R\overset{H}{C}{=}CR_2 \xrightarrow{Pd^{II}} R\overset{Ar}{C}{=}CR_2 + HX \qquad (17)$$

from the metal to the complexed alkene, followed by β-elimination of the product.

Hydroformylation is a major industrial reaction in which an alkene is converted to an aldehyde containing an additional carbon. The alkene is complexed to cobalt (Co) in the form of $HCo(CO_3)_3$. Transfer of hydrogen from the metal is followed by insertion of carbon monoxide (CO) and finally reduction of a Co-CO bond [reaction (18)]. *See* HYDROFORMYLATION.

A different type of organometallic chemistry underlies a process known as olefin metathesis, in which the groups attached to the double bonds in two alkenes are scrambled. The reaction proceeds by way of metal carbene complexes of molybdenum or tungsten (W), and the reversible formation of metallacyclobutanes [reaction (19); L = CO]. In this way an alkene with a double bond in the center of a chain can be converted by reaction with ethylene to two new terminal alkenes.

Allylic compounds. The position adjacent to a double bond is termed allylic, and a group or atom at that position is strongly affected by the π bond. Thus in a radical or an ionic intermediate, the electron deficiency or excess is delocalized by conjugation with the p orbitals (**Fig. 2**). As a consequence, these species are lower in energy and more easily formed than those in a saturated molecule. For this reason, several reactions of alkenes occur by

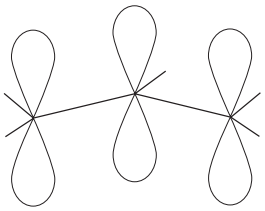

Fig. 2. Molecular orbital representation of a reactive intermediate, with delocalization by conjugation with the *p* orbitals.

substitution of the allylic position rather than addition at the double bond. Examples are the catalytic oxidation of propene to acrylic acid or acrylonitrile, or free-radical chlorination to allyl chloride [reaction (20)]. *See* DELOCALIZATION.

Reactions such as hydrolysis of an allylic type chloride often occur with rearrangement, since the allylic carbenium ion can be solvated at either end of the allylic system. An example is 1-chloro-2-butene, which gives mainly 1-buten-3-ol and a small amount of 2-buten-1-ol. The other chloride gives the same mixture of alcohols [reaction (21), where the positive charge and the π bond in the intermediate are delocalized over two carbon atoms].

Allylic compounds readily form π complexes with palladium, nickel, and other metals. These π-allyl complexes, in which the metal is complexed to three carbons, provide selective activation of the allylic system for C-C bond formation with a nucleophilic carbon, as in reaction (22) where Nu represents the nucleophile and X = Cl, RO, or ROC = O.

Dienes. Two double bonds, for example in a pentadiene, can exist in three arrangements: cumulative or allenic, conjugated, and nonconjugated, as in 2,3-, 1,3-, and 1,4-pentadiene, respectively. The 2,3-isomer is an allene, for example, 2,3-pentadiene ($CH_3HC{=}C{=}CHCH_3$); 1,3-pentadiene ($H_2C{=}CH\text{-}CH{=}CHCH_3$) is conjugated; 1,4-pentadiene ($H_2C{=}CHCH_2CH{=}CH_2$) is nonconjugated. In 2,3-pentadiene the central carbon has *sp* hybrid orbitals and is susceptible to attack by either electrophilic or nucleophilic reagents. A disubstituted allene such as 2,3-pentadiene is a chiral molecule and exists in enantiomeric forms (**Fig. 3**).

The characteristic feature of a conjugated diene is 1,4-addition. In the reaction of 1,3-pentadiene with bromine, electrophilic attack gives an allylic cation which can lead to either 1,4- or 1,2-addition [reaction (23)]. The major product generally results from the former.

Another important aspect of conjugated diene chemistry is a group of reactions that take place simply by shifts of σ and π bonds. These include ring opening of cyclohexadienes, Diels-Alder cycloadditions, and double-bond (sigmatropic) rearrangements, referred to collectively as pericyclic reactions. *See* CONJUGATION AND HYPERCONJUGATION; DIELS-ALDER REACTION; PERICYCLIC REACTION; WOODWARD-HOFFMANN RULE.

In nonconjugated dienes the double bonds can react independently, but with a 1,5- or 1,6-diene electrophilic attack at one double bond may be

Fig. 3. Enantiomeric forms of the disubstituted allene 2,3-pentadiene. The bold dot represents the central carbon atom.

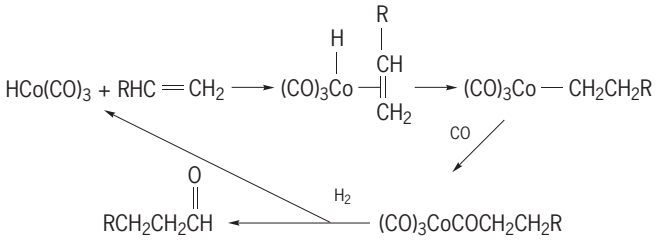

(18)

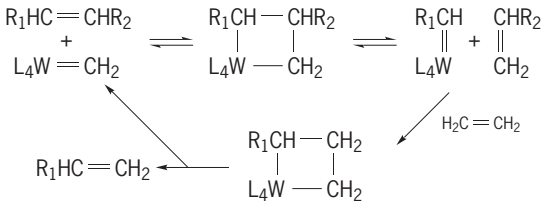

(19)

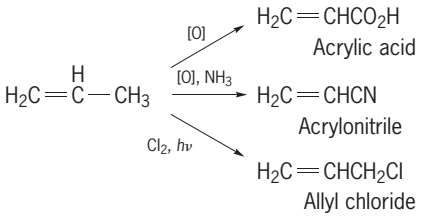

(20)

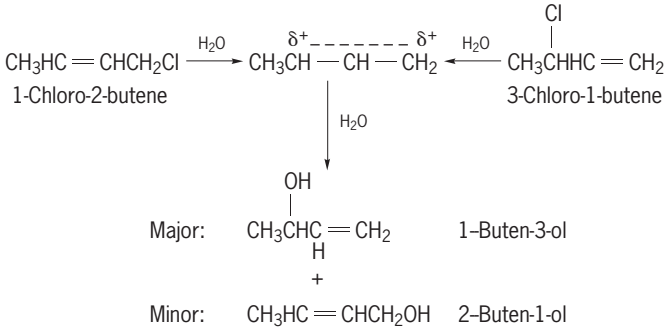

(21)

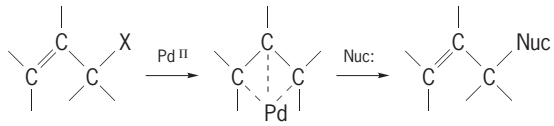

(22)

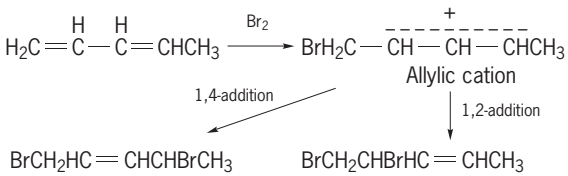

(23)

accompanied by cyclization [reaction (24), where E

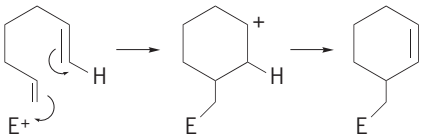

(24)

represents the electrophile]. This process frequently occurs in the biosynthesis of steroids and terpenoids.

Dienes form complexes of several kinds with transition metals. With Pd^{II} salts, butadiene forms π-allyl complexes. Complexes of iron with conjugated dienes are coordinated to four atoms. Typical compounds are linear or cyclic diene–iron tricarbonyl complexes. The elusive cyclobutadiene forms stable complexes with iron and nickel; the bonding is delocalized to achieve an 18-electron count around the metal, as in metallocenes. *See* METALLOCENES.

Metal complexes are involved in the oligomerization of dienes to cyclic dimers and trimers. Polymerization of 1,3-dienes occurs by 1,4-addition and can lead to either *cis*- or *trans*-linked chains [reaction (25),where R═CH_3, the 1,3-diene is isoprene,

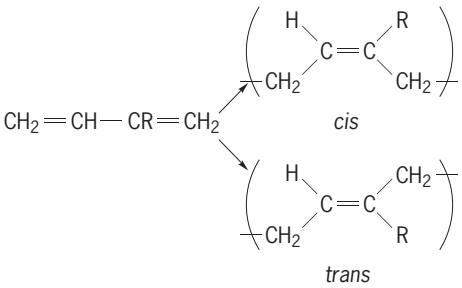

(25)

and where R═Cl, the 1,3-diene is chloroprene]. The stereochemistry can be controlled by the choice of catalysts. These polymers represent the structures of elastomeric materials. *Cis*-1,4-(poly)isoprene obtained by polymerization with R_3Al-$TiCl_4$ catalysts is the basic unit of natural rubber from *Hevea* species. The *trans* polyisoprene is obtained with other R_3Al catalysts and is the repeating unit of gutta-percha. Addition of small amounts of *cis*-1,4(poly)butadiene significantly improves the wear resistance of natural rubber. Neoprene, the polymer of 2-chlorobutadiene (chloroprene), has several advantageous features. *See* RUBBER.

Polyenes. Compounds with three or more double bonds are found in numerous naturally occurring sources. The long-chain (C_{12}–C_{20}) carboxylic acids present as esters in fats and other lipids typically contain two or three nonconjugated double bonds. The most conspicuous polyenes are carotenoids, the red-yellow pigments that occur abundantly in plants and many other organisms. These compounds are typically C_{40} molecules with an extended system of conjugated *trans* double bonds. Vitamin A is a pentaene derived in the organism by cleavage of β-carotene. Reversible *cis-trans* isomerization of one double bond plays a central role in color vision. *See* CAROTENOID; VITAMIN A.

A very long polyene chain is obtained by polymerization of acetylene with an organometallic catalyst. This material can be treated to give an electrically conducting polymer. *See* ORGANIC CONDUCTOR.

James A. Moore

Bibliography. A. Streitwieser, C. Heathcock, and E. M. Kosower, *Introduction to Organic Chemistry*, 4th ed. 1992.

Alkylation (petroleum)

In the petroleum industry, a chemical process in which an olefin (ethylene, propylene, and so forth) and a hydrocarbon, usually 2-methylpropane, are combined to produce a higher-molecular-weight and higher-carbon-number product. The product has a higher octane rating and is used to improve the quality of gasoline-range fuels. The process was originally developed during World War II to produce high-octane aviation gasoline. Its current main application is in the production of unleaded automotive gasoline. *See* GASOLINE; OCTANE NUMBER.

Reactions. The alkylation reaction is initiated by the addition of a proton (H^+) to the olefin [reaction (1)]. The protonated olefin (carbonium reactions) then reacts with the isobutane by abstraction of a proton from the isobutane to produce the *t*-butyl carbonium reactions [reaction (2)]. Reaction of this tertiary carbonium reactions with the olefin proceeds by combination of the two species [reaction (3)] to produce a more complex six-carbon carbonium reation which yields a stabilized product by abstraction of a proton from another molecule of isobutane [reaction (4)]. The reaction progresses to more product by reaction of the *t*-butyl and carbonium reactions, as already shown in reaction (3). *See* REACTIVE INTERMEDIATES.

$$H_2C{=}CH_2 + H^+ \rightarrow H_3C{-}CH_2^+ \qquad (1)$$

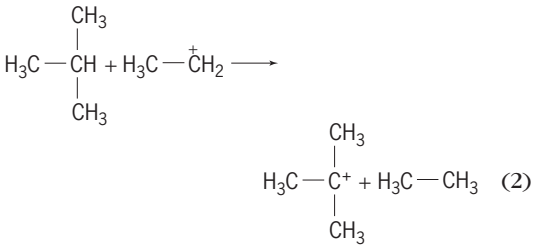

(2)

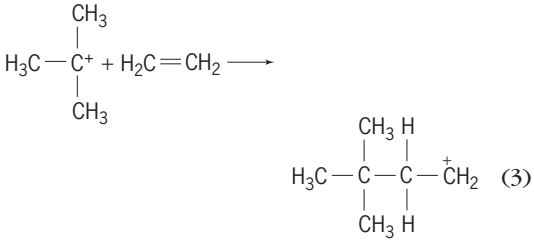

(3)

Processes. In actual refinery practice, the feedstock for the alkylation process is isobutane (or an isobutane-enriched stream) recovered from refinery

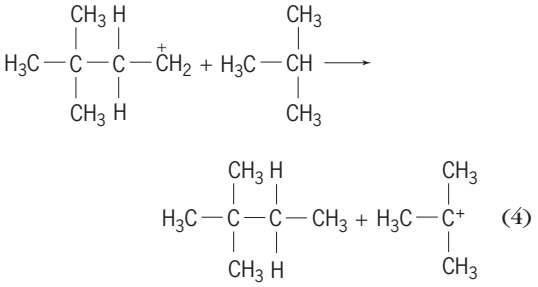

(4)

Alkylate

gases or produced by isobutane isomerization. The olefin for the feedstock is derived from the gas production of a catalytic cracker. *See* CRACKING; ISOMERIZATION.

The acid catalyst is sulfuric acid (H_2SO_4), hydrofluoric acid (HF), or aluminum chloride ($AlCl_3$). When sulfuric acid is employed, the reaction must be maintained at 35–45°F (2–8°C) to reduce unnecessary side reactions; with hydrogen fluoride and with aluminum chloride, refrigeration is not necessary, and temperatures up to 115°F (45°C) may be employed.

Sulfuric acid is used with propylene and higher-carbon-number olefins but not with ethylene because of the tendency to produce the acid-wasting ethyl hydrogen sulfate ($C_2H_5HSO_4$). Hydrogen fluoride is more generally used, and has the advantage of being more readily separated from the products because of a boiling point below that of sulfuric acid. Aluminum chloride is used to a lesser extent than both of the acids and requires injection of water to produce the hydrogen chloride promoter. *See* CATALYSIS.

Thermal alkylation is a noncatalyzed process that requires temperatures on the order of 950±25°F (about 510°C) and pressures in the range 3000–8000 lb/in.^{2} (18–55 megapascals). The reaction is believed to proceed by way of a (neutral) free-radical intermediate and is much less specific than the acid-catalyzed process. The starting hydrocarbon is usually isobutane, and the olefin is ethylene, propylene, or higher olefins such as butenes. The nature of the reactants influences the quality of the alkylate. For example, alkylation of 1-butene with isobutane gives an alkylate with a research octane rating of about 93, whereas using 2-butene as the olefin yields an alkylate having a research octane number of about 98. *See* ALKANE; ALKENE; FREE RADICAL; PETROLEUM PROCESSING AND REFINING. James G. Speight

Bibliography. S. Glasstone, *Energy Deskbook*, 1983; J. McKetta, *Petroleum Processing Handbook*, 1992; J. G. Speight, *The Chemistry and Technology of Petroleum*, 3d ed., 1999.

Alkyne

A hydrocarbon that contains a triple carbon bond, —C≡C—. The first compound in the series is acetylene. Names of alkynes are formed by changing the -ane ending of the corresponding alkane to -yne and numbering as needed, for example, 2-pentyne (structure 1), the 2 indicating that the triple bond is at the second carbon atom of the chain. Some compounds can be named as a substituted acetylene, for example, diphenylacetylene (**2**); or the ethynyl group (—C≡CH) can be named as a substituent, for example, 1-ethynylcyclopentanol (**3**).

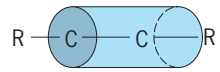

Fig. 1. Geometry of a linear arrangement of triply bonded carbon atoms and two attached groups.

$CH_3CH_2C{\equiv}CCH_3$ (**1**)

$C_6H_5C{\equiv}CC_6H_5$ (**2**)

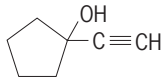

(**3**)

See ACETYLENE; ALKANE.

Structure and properties. A triple bond consists of a σ bond and two π bonds between two *sp* carbon atoms. The carbons are held closer together (0.121 nanometer) and more strongly (836 kilojoules/mole) than those in a double bond. The π bonds are orthogonal, and the electron density is distributed in a cylindrical sheath around the C-C axis. The resulting geometry is a linear arrangement of the triply bonded carbons and the two attached atoms (**Fig. 1**). *See* MOLECULAR ORBITAL THEORY.

The energy difference between a 1-alkyne and 2-alkyne (20 kJ) is larger than that between 1- and 2-alkenes, and a 1-alkyne can be completely isomerized with a base catalyst [reaction (1)]. An intermediate in this reaction is the allene isomer, which is only slightly lower in energy than the 1-alkyne. Cycloalkynes can be prepared, but the requirement of a linear four-atom system for a triple bond places a limit on the minimum size of a cycloalkane ring. Cyclooctyne, with an eight-membered ring, can be isolated; seven- and six-membered cyclic alkynes can be trapped as transient intermediates. In cycloalkynes with medium rings (8–11 carbons) the amount of allene isomer is substantial, since only three atoms are collinear and the cyclic allene in this case is significantly more stable than the cycloalkyne. *See* ISOMERIZATION; REACTIVE INTERMEDIATES.

$$RCH_2C{\equiv}CH \rightleftharpoons RCH{=}C{=}CH_2 \longrightarrow RC{\equiv}CH_3$$

A distinctive property of acetylene and other 1-alkynes is the acidity of the —C≡C—H bond. Because the electrons in an *sp* carbon orbital are closer to carbon than those in sp^2 or sp^3 orbitals, the anion —C≡C:$^-$ is stabilized. The acidities of a few relevant compounds are given in the **table**. The terminal C-H

Comparison of acidity values of several compounds

Compound	Acidity, pK_A ($-\log K_A$)
Ethane ($H_3C{-}CH_3$)	50
Ethylene ($H_2C{=}CH_2$)	44
Ammonia (NH_3)	35
Acetylene ($HC{\equiv}CH$)	25
Ethanol (CH_3CH_2OH)	16

in acetylene (and other alkynes) is far more acidic (about 10^{20}–10^{25} times) than hydrogen in alkenes or alkanes. Hydrogen can be removed from an alkyne by reaction with sodium amide, a reaction that is conveniently carried out in liquid ammonia [NH_3; reaction (2)]. Many other bases, such as alkyllithiums, can

$$\underset{\text{Alkyne}}{RC{\equiv}CH} + NaNH_2 \xrightarrow{NH_3} \underset{\text{Sodium acetylide}}{RC{\equiv}CNa} + NH_3 \qquad (2)$$

be used to prepare the acetylide anion ($-C{\equiv}C{:}^{\ominus}$). *See* PK.

Preparation. One general method for introducing a triple bond into a carbon chain is by elimination reactions, for example, removal of HX (X = halogen) from a vinyl halide [$H_2C{=}CHX$] with very strong base, or removal of two HBr molecules from a dibromide [reaction (3)].

$$\begin{array}{l} -H_2C-CBr_2- \\ -HBrC-CHBr- \end{array} \xrightarrow{-HBr} -HC{=}CBr- \xrightarrow{-HBr} -C{\equiv}C- \qquad (3)$$

Another starting point for elimination is an enol ester such as a trifluoromethylsulfonate (**4**) or an enol phosphate (**5**), as in reactions (4). The latter

$$\underset{\text{Carbonyl compound}}{-CH_2\overset{O}{\overset{\|}{C}}-} \xrightarrow{(CF_3SO_2)_2O} \begin{array}{l} -CH{=}C(OSO_2CF_3)- \quad (\mathbf{4}) \\ -CH{=}C(OPO_3H)- \quad (\mathbf{5}) \end{array} \longrightarrow -C{\equiv}C- \qquad (4)$$

has been suggested as the precursor of the triple bonds in naturally occurring acetylenes.

Alkynes are also obtained by thermal elimination of neutral molecules, as in the preparation of cycloalkynes by oxidizing the hydrazones of 1,2-diketones. Another reaction is elimination of carbon monoxide (CO) from a cyclopropenone (**6**) [reaction (5)] or cyclobutanedione (**7**) [reaction

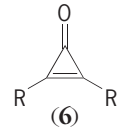

$$\xrightarrow[\Delta]{-CO} RC{\equiv}CR \qquad (5)$$

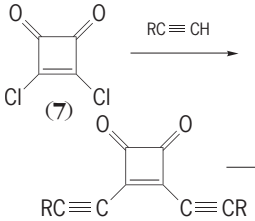

$$\xrightarrow[\Delta]{} R-(C{\equiv}C)_3-R \qquad (6)$$

(6)]. The latter reaction can be used to build up chains containing several triple bonds.

A second general approach for introducing a triple bond is using reactions involving an acetylide anion in either a substitution or addition reaction. The acetylide anion is an excellent nucleophile, similar in reactivity to cyanide ion ($N{\equiv}C{:}^{\ominus}$). Any alkyl halide or sulfonate ester that is a practical substrate for substitution can be converted to an alkyne. With a dihalide (**8**) and excess acetylide, a diacetylene (**9**) is available [reaction (7)].

$$\underset{(\mathbf{8})}{Br(CH_2)_nBr} + \underset{\text{Sodium acetylide}}{2HC{\equiv}CNa} \longrightarrow \underset{(\mathbf{9})}{HC{\equiv}C(CH_2)_nC{\equiv}CH} \qquad (7)$$

Addition of acetylides to aldehydes, ketones, or epoxides leads to the corresponding alkynols. An acetylide can be acylated with an acid chloride or carbonated to give the ketone or carboxylic acid, respectively.

Addition reactions. In general, electrophilic additions are slower with triple bonds than double bonds because π electrons in the latter are more accessible. However, several addition reactions of alkynes are useful. Hydrogenation of a triple bond can be carried out with a palladium (Pd) catalyst that is deactivated ["poisoned" with the addition of lead (Pb)] to permit selective reduction to the *cis*-alkene. Reduction to the *trans*-alkene can be accomplished with sodium metal (Na) in ammonia (NH_3) containing alcohol [reaction (8)].

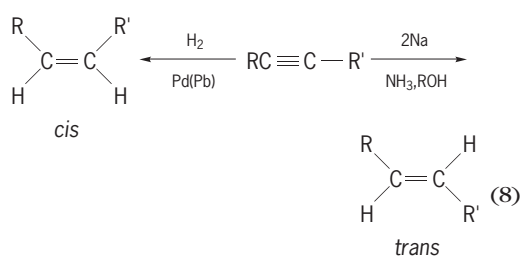

(8)

Addition of water to a terminal triple bond occurs in the presence of a mercuric (Hg^{II}) salt, leading

initially to an enol and thence to a ketone. Reaction of the alkyne with borane (B_2H_6) gives a vinyl borane that can be oxidized in the presence of water (H_2O) and hydrogen peroxide (H_2O_2) to the isomeric enol of the aldehyde [reactions (9)]. Dialkyl aluminum

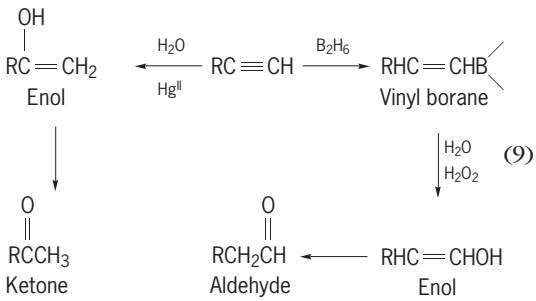

reagents add to alkynes. The alkenylaluminum products can be converted stereoselectively to alkenes or to alkenyl halides or allylic alcohols.

Cyclization reactions. In early work benzene was observed among the products formed on heating acetylene. Substituted acetylenes are quite stable thermally, but metal-catalyzed thermal trimerization gives trisubstituted benzenes, with the 1,2,4-isomer greatly predominating over the 1,3,5-isomer [reaction (10)].

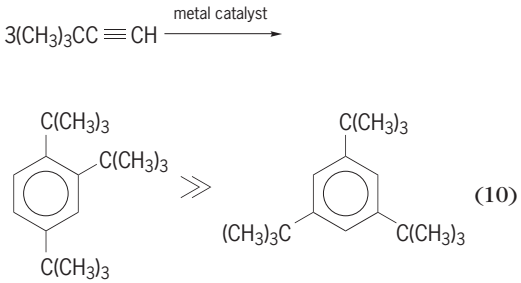

This bias for the 1,2,4 pattern indicates a mechanism that is not a simple head-to-tail cyclization, and may involve some type of metal-oriented dimeric complex. The reaction is a route to benzenes with adjacent bulky groups.

The formation of cyclooctatetraene in high yield by reaction of acetylene in the presence of a nickel (Ni^{II}) catalyst was a seminal finding in organometallic chemistry. A variety of other cyclizations of alkyne-metal complexes have been observed. In some of these reactions carbon monoxide molecules are also involved, leading to cyclopentadienones and tropones. The cocylization of an alkyne with an alkene and carbon monoxide, catalyzed by dicobalt octacarbonyl [$Co_2(CO)_8$], is a general and flexible method for cyclopentenone preparation. *See* ORGANOMETALLIC COMPOUND.

Diynes and polyynes. 1-Alkynes can be coupled by mild oxidation to the conjugated diacetylene. The reaction is carried out by air oxidation of the cuprous (Cu^I) acetylide or oxidation of the alkyne with cupric (Cu^{II}) ion in the presence of an amine. For coupling of two different alkynes, the reaction is carried out with the copper derivative of one alkyne and a different 1-bromoalkyne. The yields in these couplings are usually good, and a variety of functional groups, including additional triple bonds, can be present. Oxidative coupling of terminal dialkynes leads to macrocyclic polyacetylenes [reaction (11)]. This is

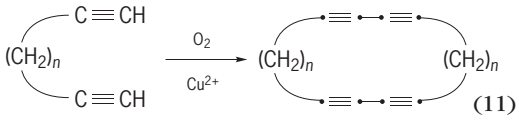

a very direct way to prepare large-ring compounds, although rings with different numbers of diyne units and also polymeric products are formed. *See* MACROCYCLIC COMPOUND.

Polyynes occur in many species of higher plants, particularly in the families Compositae, Umbelliferae, and Asteraceae as well as in Basidiomycetes and other fungi. Many hundreds of compounds have been isolated and characterized; nearly all of these are structurally and biogenetically related to fatty acids with unbranched C_{10}–C_{20} chains. A number of compounds are derived from a series of C_{13} hydrocarbons that contain a combination of six triple and double bonds. A common feature is the occurrence of thiophenes arising from addition of the biochemical equivalent of hydrogen sulfide (H_2S) to two adjacent triple bonds [reaction (12)]. Many other functional

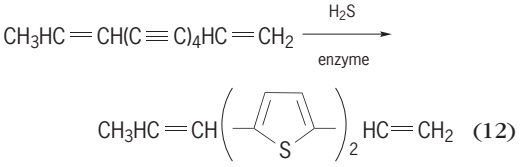

groups such as hydroxyl, epoxy, furan, and pyran rings are also present.

A group of antibiotics that are potent inhibitors of tumor growth have as a common feature a macrocyclic enediyne system. The mechanism of action of these compounds involves rearrangement of the multiple bonds to a benzene diradical which then interferes with deoxyribonucleic acid (DNA) synthesis. *See* ANTIBIOTIC.

Compounds containing an extended sequence of conjugated triple bonds with the general formula —(C≡C)$_n$— have been detected in interstellar matter as well as in plant sources. Polyyne chains with up to 16 triple bonds (**Fig. 2***a*), and also molecules such

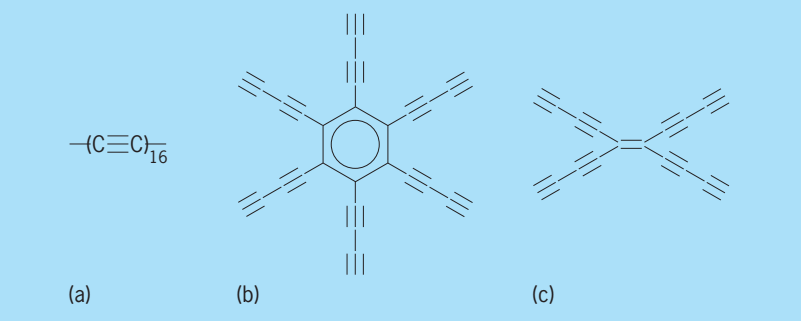

Fig. 2. Compounds containing an extended sequence of conjugated triple carbon bonds: (*a*) polyyne chain with 16 triple bonds; (*b*) hexa(butadiynyl)benzene; and (c) tetrabutadiynyl ethylene.

as hexa(butadiynyl)benzene (Fig. 2*b*) and tetrabutadiynyl ethylene (Fig. 2*c*), have been prepared. These compounds are stable with a bulky group such as trialkylsilyl at the ends of the polyyne chains; but with terminal hydrogen, they are extremely reactive and may ignite in air or explode. Compounds with extended polyyne systems are of interest as precursors of carbon rods or two-dimensional carbon networks, both of which have potential as new materials. James A. Moore

Bibliography. L. Brandsma and H. D. Verkruijsse, *Syntheses of Acetylenes, Allenes and Cumulenes*, 1981; F. Diedrich and Y. Rubin, Synthetic approaches to molecular carbon allotropes, *Angew. Chem. Int. Ed.*, 31:1101–1123, 1992; H. G. Viehe (ed.), *Chemistry of Acetylenes*, 1969.

Allantois

A fluid-filled sac- or sausagelike, extraembryonic membrane lying between the outer chorion and the inner amnion and yolk sac of the embryos of reptiles, birds, and mammals. The allantois eventually fills up the space of the extraembryonic coelom in most of these animals. It is composed of an inner layer of endoderm cells, continuous with the endoderm of the embryonic gut, or digestive tract, and an outer layer of mesoderm, continuous with the splanchnic mesoderm of the embryo. It arises as an outpouching of the ventral floor of the hindgut and dilates like a filling balloon into a large allantoic sac which spreads throughout the extraembryonic coelom. The allantois remains connected to the hindgut by a narrower allantoic stalk which runs through the umbilical cord. *See* AMNION; CHORION; GERM LAYERS.

The allantois eventually fuses with the overlying chorion to form the compound chorioallantois, which lies just below the shell membranes in reptiles and birds. The chorioallantois is supplied with an extensive network of blood vessels and serves as an important respiratory and excretory organ for gaseous interchange. The allantoic cavity also serves as a reservoir for kidney wastes in some mammals, in reptiles, and in birds. In the latter two groups the allantois assists in the absorption of albumin. In some mammals, including humans, the allantois is vestigial and may regress, yet the homologous blood vessels persist as the important umbilical arteries and veins connecting the embryo with the placenta. *See* FETAL MEMBRANE; PLACENTATION. Nelson T. Spratt, Jr.

Bibliography. B. I. Balinsky, *Introduction to Embryology,* 5th ed., 1981.

Allele

Any of a number of alternative forms of a gene. Allele is a contraction of allelomorph, a term which W. Bateson used to designate one of the alternative forms of a unit showing mendelian segregation. New alleles arise from existing ones by mutation. The diversity of alleles produced in this way is the basis for hereditary variation and evolution. *See* GENE; MENDELISM; MUTATION.

The genetic material is a coded set of instructions in chemical form (deoxyribonucleic acid) for making a living cell. Each gene is like an essential word in the code. The different alleles of a given gene determine the degree to which the specific hereditary characteristic controlled by that gene is manifested. The particular allele which causes that characteristic to be expressed in a normal fashion is often referred to as the wild-type allele. Mutations of the wild-type allele result in mutant alleles, whose functioning in the development of the organism is generally impaired relative to that of the wild-type allele. In this analogy, the wild-type allele corresponds to the correct word in the genetic code and the mutant alleles correspond to mistakes in the spelling of that word. *See* DEOXYRIBONUCLEIC ACID (DNA); GENETIC CODE.

An allele occupies a fixed position or locus in the chromosome. In the body cells of most higher organisms, including humans, there are two chromosomes of each kind and hence two alleles of each kind of gene, except for the sex chromosomes. Such organisms and their somatic cells are said to carry a diploid complement of alleles. A diploid individual is homozygous if the same allele is present twice, or heterozygous if two different alleles are present. Let A and a represent a pair of alleles of a given gene; then A/A and a/a are the genetic constitutions or genotypes of the two possible homozygotes, while A/a is the genotype of the heterozygote. Usually the appearance or phenotype of the A/a individuals resembles that of the A/A type; A is then said to be the dominant allele and a the recessive allele. In the case of the sex chromosomes, one sex (usually the male in most higher animals, with the exception of birds) has only one X chromosome, and the Y lacks almost all of the genes in X. The male thus carries only one dose of X-linked genes and is said to be hemizygous for alleles carried on his X chromosome. As a result, if a male inherits a recessive mutant allele such as color blindness on his X chromosome, he expresses color blindness because he lacks the wild-type allele on his Y chromosome. *See* CHROMOSOME; SEX-LINKED INHERITANCE.

The wild-type allele of a gene is symbolized by attaching a superscript plus sign to the gene symbol; for example, a^+ is the wild-type allele of the recessive mutant gene a. When no confusion arises, the symbol $+$ is used instead of a^+. This means that $a/+$ is equivalent to a/a^+. In some cases a mutant is dominant to the wild-type allele, and then the mutant symbol is usually capitalized.

In a population of diploid individuals, it is possible to have more than two alleles of a given gene. The aggregate of such alleles is called a multiple allelic series. Since genes are linear sequences of hundreds or even thousands of nucleotide base pairs, the potential number of alleles of a given gene which can arise by base substitution alone is enormous. For example, a gene coding for a protein with only 100 amino acids would be considered a very small gene, and would have 300 base pairs in its coding region; yet, since each base pair can exist in four different chemical

forms, 4^{300} possible mutant alleles could arise within the coding region alone, or approximately 10^{180} alleles—a number vastly larger than the estimated number of particles in the known universe. Actually, many other types of mutant alleles of a gene do occur, including duplications, deletions, and inversions of DNA within the noncoding as well as coding regions of the gene. Indeed, many spontaneous mutations in at least some higher organisms, such as yeast, maize, and *Drosophila,* turn out to be insertions of large segments of DNA into either the noncoding or coding regions. These insertions, known as transposons, appear to be virallike elements that have infected the germ line. *See* TRANSPOSONS.

Mutations of independent origin that involve changes at the same nucleotide positions in a gene are known as homoalleles, while mutations that involve changes at different nucleotide positions are called heteroalleles. Recombination between alleles of the same gene does not occur or tends to be extremely rare in higher organisms. Also, recombination within the gene in such cases tends to be different in outward properties from ordinary crossing over and is known as gene conversion. *See* RECOMBINATION (GENETICS).

It is often difficult to distinguish operationally between a series of multiple alleles of a single gene and a cluster of closely linked genes with similar effects. Such a cluster, originally known as a pseudoallelic series, is now termed a gene complex, gene cluster, or multigene family, except that, in the latter case, the genes of the family are not necessarily closely linked. In bacteria, gene clusters are usually coordinately regulated and are termed operons. Gene cloning has verified the existence of gene clusters in higher organisms as well as in bacteria. *See* BACTERIAL GENETICS; COMPLEMENTATION (GENETICS); GENETICS; OPERON.

Edward B. Lewis

Allelopathy

The biochemical interactions among all types of plants, including microorganisms. The term is usually interpreted as the detrimental influence of one plant upon another but is used more and more, as intended originally, to encompass both detrimental and beneficial interactions.

Forms and occurrence. At least two forms of allelopathy are distinguished: (1) the production and release of an allelochemical by one species inhibiting the growth of only other adjacent species, which may confer competitive advantage for the allelopathic species; and (2) autoallelopathy, in which both the species producing the allelochemical and unrelated species are indiscriminately affected.

The term allelopathy, frequently restricted to interactions among higher plants, is now applied to interactions among plants from all divisions, including algae. Even interactions between plants and herbivorous insects or nematodes in which plant substances attract, repel, deter, or retard the growth of attacking insects or nematodes are considered to be allelopathic.

Interactions between soil microorganisms and plants are important in allelopathy. Fungi and bacteria may produce and release inhibitors or promoters. Through the secretion of siderophores they can immobilize iron, which may deprive harmful rhizobacteria and pathogens of this important trace element. Siderophores from pseudomonad bacteria, however, may reduce iron uptake by plants, whereas siderophores from agrobacteria may enhance it. Some bacteria enhance plant growth through fixing nitrogen, others through providing phosphorus. The activity of nitrogen-fixing bacteria may be affected by allelochemicals, and this effect in turn may influence ecological patterns. Allelochemicals, released from plant litter by microbial action, are often also removed from the rhizosphere by microorganisms. The rhizosphere must be considered the main site for allelopathic interactions. *See* NITROGEN FIXATION; RHIZOSPHERE.

Chemicals produced and secreted by microorganisms that inhibit other microorganisms are called antibiotics. In principle, this is a form of allelopathy. Vascular plants react to microbial infection as well as to abiotic elicitors by synthesizing phytoalexins, which are low-molecular-weight, broad-spectrum antimicrobial compounds, another form of allelopathy. In contrast to phytoalexins, allelochemicals are produced without the action of an elicitor. *See* ANTIBIOTIC; PHYTOALEXINS; SOIL MICROBIOLOGY; TOXIN.

Allelopathy is clearly distinguished from competition: in allelopathy a chemical is introduced by the plant into the environment, whereas in competition the plant removes or reduces such environmental components as minerals, water, space, gas exchange, and light. In the field, both allelopathy and competition usually act simultaneously, and assessing their individual influence and function is a continuing challenge. In controlled laboratory experiments with extracts or exudates from plants suspected of allelopathy, bioassays often yield significant allelopathic effects that are not manifest in nature. Soil and weather conditions, and the influence of surrounding plants and microbial interactions, carefully controlled in the laboratory, may mask allelopathic effects in the field. The unequivocal demonstration of allelopathy in the field as a result of the action of chemical compounds showing allelopathic activity in a laboratory bioassay is difficult because of the complexity and dynamic nature of the environmental interactions.

Allelopathy was first observed in ancient times: Pliny (A.D. 23–79) related the inhibition of plant growth around walnut trees. The black walnut tree *(Juglans nigra)* produces a potent allelochemical, juglone (5-hydroxynaphthoquinone; structure **1**); that inhibits growth of many annual

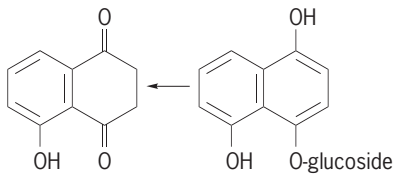

(**1**) (**2**)

herbs. Tomato and alfalfa under or near black walnut trees wilt and die. Alder *(Alnus glutinosa),* planted as nurse trees in black walnut orchards, decline and die from allelopathy within 8–13 years. Water leachate from leaves and extracts from the soil under a walnut tree contain juglone. Adding juglone to plant beds inhibits or kills the plants. Juglone is not, however, a general toxin for all plants; Kentucky bluegrass *(Poa pratensis)* and brambles such as wild raspberries and blackberries *(Rubus* spp.) grow under and around black walnut trees. In the cells of walnut trees, juglone is present in a safe form as 1,4,5-trihydroxynaphthalene glucoside (**2**), which is not inhibitory. Upon leaching from the leaves and bark by rain or release from the roots, the glucoside is hydrolyzed and the free hydroxynaphthalene is oxidized to the active allelochemical, juglone. The release of inhibitors, for example, hydrocyanic acid (HCN), from harmless precursors, such as cyanogenic glycosides, occurs frequently in allelopathy.

Desert plants. For plants growing in habitats with extreme climates, such as a desert, competition for the limited resources is critical, and allelopathy may have survival value. Desert shrubs are often surrounded by a bare zone; thus, all the moisture of that zone remains available to the shrub and is not shared with other plants. In the Mojave Desert of California incienso *(Encelia farinosa)* inhibits the growth of desert annuals, which will grow only around dead incienso shrubs. From the decomposing leaf litter, 5-acetyl-1-2-methoxybenzaldehyde (**3**) is released and

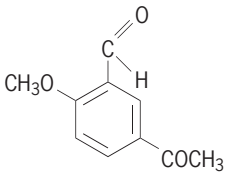

(**3**)

persists in the desert soil, functioning as an allelochemical. Incienso is apparently not affected by its own toxin. *See* DESERT.

Chaparral. Important allelochemicals in chaparrals are terpenoids and water-soluble compounds, such as organic acids and phenols.

Terpenoids. The obvious and striking vegetation pattern of bare ground and stunted herbs and grasses around the shrub thickets of sagebrush *(Salvia leucophylla)* and wormwood *(Artemisia californica)* in the Californian chaparral is caused by an allelopathic effect of the volatile terpenes produced by these shrubs. *Salvia* and *Artemesia* are surrounded by a zone of bare soil 3–6 ft wide (1–2 m), followed, toward the periphery, by a zone of annual herbs and grasses with stunted growth and limited development and, finally, thriving grassland with wild oats *(Avena fatua),* brome grass *(Bromus rigidus),* and other grass species. Careful analyses have eliminated shade, nutrient deficiency, water availability, root density, slope of ground, and animal feeding and running activity as the causes for growth inhibition. While grazing by small mammals, birds, and insects—all taking advantage of the shrub for shelter—helps to keep down growth, grazing does not initiate or usually maintain the zones of inhibition. *See* HERBIVORY.

The uphill and downhill inhibition zones of *Salvia* shrubs are about equal in size, indicating that water-soluble inhibitors are unlikely in the dry climate of the chaparral. This suggests the effect of volatile compounds. Volatile terpenes from *Salvia* and *Artemisia* and other shrubs and their surrounding soil were identified and found to be inhibitory to seed germination and plant growth. Camphor (**4**) and 1,8-cineole (**5**) are the most effective inhibitors; α-pinene (**6**), β-pinene (**7**), camphene (**8**), and thujone (**9**) are

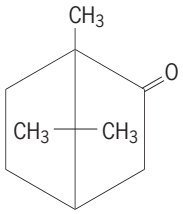

(**4**)

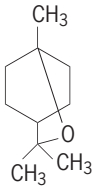

(**5**)

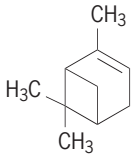

(**6**)

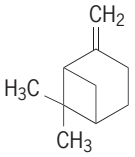

(**7**)

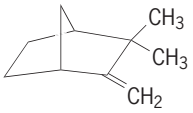

(**8**)

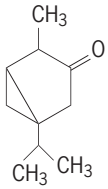

(**9**)

also active volatile allelochemicals. Growth inhibition is most obvious around shrubs growing in the fine-textured Zaca clay, which absorbs the terpenes and thus facilitates contact with the roots.

Spectacular fires, which sweep the chaparral approximately every 25 years, volatilize the soil-bound terpenes and destroy the shrubs. The burned sites quickly turn green with annual herbs and grasses that dominate for approximately 6 years until the shrubs grow back from surviving underground parts. With the accumulation of terpenes in the soil, the zones of bare ground and stunted growth reappear and remain apparent until the next fire.

Terpenoids are common plant products and might be expected to serve as allelochemicals in other plant species. Indeed, for eucalyptus trees and sassafras, which do not inhabit the chaparral, terpenoids have been identified as effective allelochemicals; in other

plants, especially conifers, terpenoids are suspected of acting as such also. *See* TERPENE.

Water-soluble allelochemicals. Chamise *(Adenostema fasciculatum)* and manzanita *(Arctostaphylos glandulosa),* two common chaparral shrubs species, lack volatile terpenes, but because growth around them is also inhibited, other allelochemicals may be involved. Ferulic acid (**10**), *p*-hydroxybenzoic acid (**11**), and vanillic acid (**12**) are found in their leaves

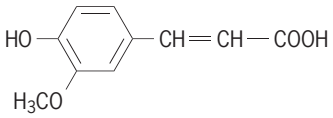

(**10**)

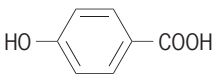

(**11**)

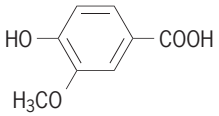

(**12**)

as well as in the surrounding soil and act as strong inhibitors of seed germination and growth. Other phenols and phenolic acids are present in the leaves but are not always detected in the soil. Rainfall in the chaparral is only moderate. The coastal fog, which prevails throughout the year, however, supplies sufficient moisture to leach the phenolic compounds into the soil, rendering them effective allelochemicals. *See* CHAPARRAL.

Old-field succession. Allelopathy has been studied in the four stages of old-field succession of the Oklahoma-Kansas prairie. The pioneer weed stage, lasting only 2–3 years, is characterized by sturdy weeds with low mineral needs, such as sunflower *(Helianthus annuus)*, ragweed *(Ambrosia psilostachya)*, Johnson grass *(Sorghum halepense)*, crab grass *(Digitaria sanguinalis)*, and spurge *(Euphorbia supina)*. Autoallelopathy of the pioneer weeds accelerates their own replacement by the dominant species of the second stage: namely the triple-awn grass *(Aristida oligantha)*, which lasts about 9–13 years. Tolerant of the pioneer weed-produced allelochemicals, such as chlorogenic acid (**13**), *p*-coumaric acid (**14**), gallic acid (**15**), and *p*-hydroxybenzaldehyde (**16**), e-awn grass also

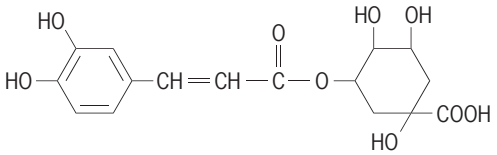

(**13**)

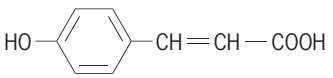

(**14**)

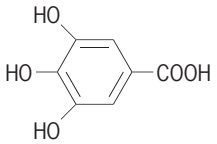

(**15**)

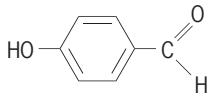

(**16**)

(**17**)

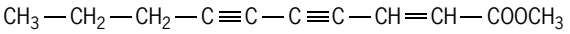

(**18**)

has low mineral requirements. Allelochemicals produced from triple-awn grass, such as gallic acid and tannic acid, inhibit free-living and symbiotic nitrogen-fixing bacteria, thus giving the grass a selective advantage over invading species of the third stage that require higher nitrogen concentrations. Inhibition of the nitrifying bacteria by allelochemicals produced by the invading species, in turn, controls the rather slow succession of the third and fourth stages. Gradually, ammonia in the soil increases to such a level that species of the third stage, perennial bunchgrasses—lasting some 30 years—and finally, the climax prairie, can slowly claim the field.

In Japanese urban wasteland, weed succession, which is similar to that of old fields, appears to be controlled by polyacetylene-type allelochemicals. Matricaria esters (**17**) and lachnophyllum ester (**18**) in the pioneer weeds, *Solidago altissima* and *Erigeron* species, have also been identified in the soil, suggesting ecological significance. *See* ECOLOGICAL SUCCESSION.

Aqueous habitats. Allelopathy is not restricted to terrestrial habitats but occurs also in fresh water and marine environments. The spikerush (*Eleocharis coloradoensis*) is an aquatic plant that has been shown to produce an allelochemical effect. Allelopathic substances in saltwater are of special interest with regard to the red tide, that is, blooms of microorganisms that discolor the water and discharge toxins. Algae-produced allelochemicals may be able to limit the growth of red tide producers. *Nannochloris*, an alga, synthesizes a cytolytic agent affecting a red tide–causing dinoflagellate. Other algae have been found to secrete fatty acids that inhibit algal growth.

Agriculture. The reduction of crop yields by weed competition is aggravated by the allelopathic effect of weeds on crops. It was already noted in 1832 that thistles *(Cirsium)* diminished the yield of oats, and spurge *(Euphorbia)* diminished the yield of flax. The growth inhibition of wheat by couch grass *(Agropyron repens)* results from impaired phosphate uptake,

even in the presence of adequate phosphate supply, and results in severe crop losses. Leachates from Italian ryegrass *(Lolium multiflorum)*, its litter, and the soil in which it grows inhibit the growth of oat, clover, lettuce, and bromegrass. Foxtail species *(Setaria)* reduce the growth of corn by as much as 35% through allelopathy. The combined effects of foxtail allelopathy and competition can result in 90% growth reduction.

Crop plants may inhibit their own growth and reduce the yield of subsequent crops, an observation well known to farmers as soil sickness long before the investigations of allelopathy. Crop rotation cures soil sickness only when the subsequent crop is not affected by the accumulated allelochemicals of the previous crop or when they have been detoxified by soil microorganisms. Cabbage, rich in thiocyanates, inhibits grass species, tobacco, and grapes. Woad *(Isatis tinctoria)* contains high levels of glucosinolates (mustard oil glycosides), especially glucobrassicin (**19**), and its planting is there-

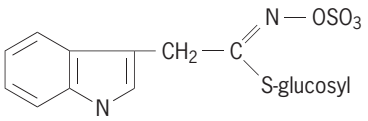

(**19**)

fore often avoided because of its ill effects on succeeding crops. Straw from wheat, corn, rye, rice, or sorghum decomposing in the field yields phenolic compounds, such as vanillic acid (**12**), ferulic acid (**10**), *p*-hydroxybenzoic acid (**14**), *p*-coumaric acid (**13**), *o*-hydroxyphenylacetic acid, and protocatechuic acid (**20**), all of which exert allelopathic ef-

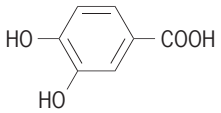

(**20**)

fects. In addition, an increase in the soil microflora in response to the crop residue generally results in nitrogen immobilization and, hence, in reduced growth of the succeeding crop. As usual in nature, competition and allelopathy occur simultaneously, and the relative effect of each on plant growth and yield is difficult to assess.

Soil microorganisms have a critical role in the allelopathy responsible for clover soil sickness. Clover roots exude toxic isoflavonoids, which do not cause allelopathy but are converted by microorganisms to salicylic acid (**21**), *p*—methoxybenzoic acid (**22**), 2,4-hydroxybenzoic acid (**23**), and other phenolic com-

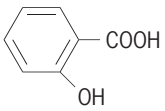

(**21**)

(**22**)

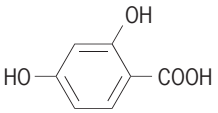

(**23**)

pounds that are responsible for clover soil sickness.

Soil microorganisms may also prevent allelopathy. In sandy soils, sorghum autoallelochemicals reduce the yield of the first and subsequent sorghum crops. In clay-rich soils, sorghum autoallelopathy does not become manifest because fungi degrade the allelochemicals before growth of the new seedlings starts.

In the complex interaction between crops and weeds and among crops lie exciting possibilities for utilizing the allelopathic properties of crop plants for weed control. Allelopathic weed control with crop plants has long been practiced by farmers and horticulturists. The challenge is twofold: to minimize the negative impact of allelochemicals on crop growth and yield and to exploit allelopathic mechanisms for pest control and crop growth regulation strategies. Allelochemicals in new crop cultivars may provide naturally occurring pesticides that can limit or suppress weeds as well as prevent insect and nematode attack and damage. The biotechnology resources for the production of herbicide-resistant crops could then be channeled into the engineering of other desired crop qualities. Allelochemicals may furnish an entirely new generation of naturally produced weed-controlling compounds, replacing synthetic herbicides and pesticides with nonaccumulating, easily degradable substances. *See* HERBICIDE.

The natural occurrence of many plant species in pure stands that apparently cannot be invaded by other species may result, in part, from allelopathic mechanisms that could be exploited to achieve weed-free crop stands. Barley, considered a smother crop, was thought to inhibit weeds through competition with its extensive root system. However, even in the absence of competition, barley inhibits weeds, such as chickweed (*Stellaria media*), shepherd's purse (*Capsella bursa-pastoris*), and tobacco, but not wheat. Gramine (**24**), an alkaloid, which occurs

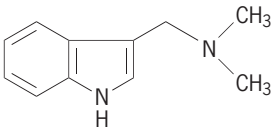

(**24**)

in barley and inhibits *Stellaria* growth, may function as an allelochemical. Allelochemicals in seeds, including alkaloids (for instance, in coffee beans), may protect seeds from microbial attack and destruction while they remain viable and buried in the soil for years. Agricultural practice uses cover crops, which are either killed by frost or by desiccant sprays, and the allelochemicals of the crop residues suppress weed growth. Oats, corn, sunflower, sorghum, and fescue appear promising in allelopathic weed control. *See* ALKALOID.

Not all weed-crop interactions, however, are necessarily detrimental for the crop. Corn cockle (*Agrostemma githago*), growing in wheat fields in former Yugoslavia, increases the wheat yield but suppresses its own growth. Agrostemin and gibberellin have been identified as the allelochemicals of corn cockle. Triacontanol, $CH_3(CH_2)_{28}CH_2OH$, an allelochemical from alfalfa, promotes growth of tomato, cucumber, and lettuce among other species. *See* GIBBERELLIN.

Parasitism. Infestation of vascular plants by vascular plant parasites, often treated as an oddity, is an allelopathic relationship. The seeds of the parasite witchweed (*Striga asiatica*) germinate in response to root exudate from host plants. While the active compounds have not yet been identified, strigol (**25**) induces witchweed seed germination and the formation of a haustorium (a nutrient-absorbing structure) that attaches to the host. Strigol is beneficial to the parasite by stimulating its growth and facilitating recognition of the host, but is detrimental to the plant producing it, when the parasite finds the producer and attaches itself. Purple false foxglove (*Agalinis purpurea*), which parasitizes legumes, responds to xenognosin A (**26**) with haustorium formation.

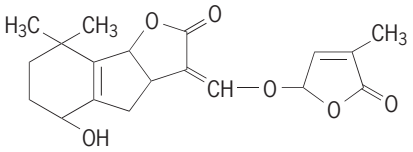

(**25**)

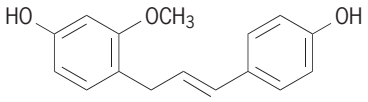

(**26**)

Forestry. As in agriculture and horticulture, allelopathy along with other ecologically important factors affects the pattern of forest growth and development and therefore has strong economic implications. While examples of tree allelopathy, such as that of the black walnut, were well known, the general significance of allelopathy for forestry has been recognized only in recent times.

Oak trees left as seed trees may impede reforestation by retarding the growth of the tree seedlings. Salicyclic acid (**21**) has been identified as one of the allelochemicals, but strong competition from oak roots and physical suppression of seed germination by oak litter also take part in retarding growth.

Bare areas under hackberry (*Celtis laevigata*) result from accumulation of the same phenolic allelochemicals that affect many other plant communities. Synergistic action of these allelochemicals, while often suspected, has been observed with hackberry toxins. Seasonality of allelopathy in a forest of sycamore (*Platanus occidentalis*), rough-leaved hackberry (*C. occidentalis*), and white oak (*Quercus alba*) may indicate that certain allelochemicals are effective upon release, whereas others may require chemical change by microorganisms before they become active. Production and release of toxins must exceed decomposition for allelopathy to occur. *Grevillia robusta*, a stately tree in the subtropical rainforest of Australia, does not form pure stands, and, not surprisingly, does not grow in monoculture beyond 10–12 years, when it declines because of strong autoallelopathy.

Bracken stands or bracken-grassland communities surrounded by forest often remain free of trees. Phenolic compounds, including caffeic acid (**27**) and ferulic acid (**10**), appear to be the controlling allelochemicals. A dense ground cover of bracken, wild oat grass (*Danthonia compressa*), goldenrod (*Solidago rugosa*), and flat-topped aster (*Aster umbellatus*), does not allow a forest in the Allegheny Plateau of northwestern Pennsylvania to grow back, and half a century after clear-cutting, only a few scattered black cherry (*Prunus serotina*) and red maple (*Acer rubrum*) grow. Other woodland ferns, such as the hay-scented fern (*Dennstaedtia punctilobula*) and New York fern (*Thelypteris noveboracensis*), also contribute to the allelopathic effect.

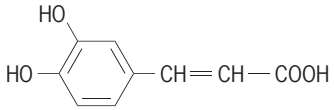

(**27**)

Release of allelochemicals from bracken appears to be timed for the most effective suppression of surrounding vegetation. In the tropics, toxins are released from green bracken leaves year-round, while in southern California and in the Pacific Northwest, toxins are released from dead bracken fronds, litter, and roots when seeds germinate, either at the beginning of the rainy season or in the spring, respectively.

Allelopathic inhibition of forest regrowth may be indirect by affecting the mycorrhizal fungi (fungi which grow mutualistically in association with roots of vascular plants). Shrub litter, heather (*Calluna vulgaris*), and prairie grasses inhibit Norway spruce (*Picea abies*) development by preventing the establishment of mycorrhiza. *See* FOREST AND FORESTRY; MYCORRHIZAE.

Action mechanisms of allelochemicals. While numerous allelochemicals have been identified, most studies on allelopathy deal with unidentified compounds. Allelochemicals are secondary plant products and belong to a few major classes of chemical compounds: terpenoids, phenolic compounds, phenylpropane derivatives, flavonoids, long-chain fatty acids, organic cyanides, probably also alkaloids and purines, and perhaps some steroids. In the field, allelopathy involves a complex of compounds; a single specific compound does not appear to produce the observed allelochemical effects on neighboring plants. This complicates the investigation of the action mechanism of allelochemicals.

There is no single set of physiological reactions controlled by allelochemicals that result in allelopathy; the action of allelochemicals is diverse and affects a large number of physiological functions and biochemical reactions. At the whole-plant level, allelochemicals interfere with internode elongation, leaf expansion, cell division, dry-weight accumulation, as well as seed germination. Growth responses mediated by natural plant growth regulators, such as auxin or gibberellin, are often impaired. In field experiments, allelochemicals from some crop residues promote bacterial and fungal lesions on roots, enhancing diseases of the test seedlings. *See* PLANT HORMONES.

Common to many allelochemicals is their effect on membrane permeability. Ion uptake can be enhanced or reduced; certain phenolic acids inhibit the uptake of phosphate and potassium. Fescue leachates impair uptake of phosphorus and nitrogen in sweetgum. Some flavonoids, phenolic acids, and juglone inhibit membrane adenosine triphosphatase activity that is involved in the control of membrane transport. Sphagnum acidifies its habitat by releasing hydrogen ions, which may make conditions less suitable for competing plants.

Significant effects on photosynthesis by allelochemicals have been observed. Allelochemicals from pigweed and foxtail litter affect photosynthesis in corn and soybeans and, in addition, alter the biomass partitioning into the leaves compared with that of the total plant. Stomatal movement, which regulates gas exchange and water status, and net photosynthesis may also be inhibited. At the cellular and subcellular level, kaempferol (**28**), a flavonol, interferes with the

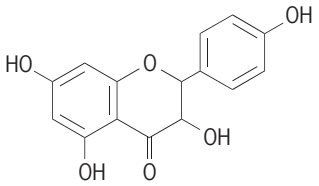

(**28**)

photosynthesis of isolated chloroplasts by inhibiting coupled electron transport and photophosphorylation (adenosine triphosphate synthesis). In isolated mitochondria, kaempferol inhibits oxidative phosphorylation. Respiration was also inhibited by juglone (**1**) and other allelochemicals. The biosynthesis and metabolism of proteins, lipids, and organic acids are also affected by allelochemicals. While it is relatively easy and straightforward to isolate and identify allelochemicals, to test them in bioassays, and to measure their physiological and biochemical reactions, in the field it is difficult to prove directly that the same compounds move from the producing plant to the neighboring plants and function as allelochemicals in the field. Manfred Ruddat

Bibliography. W. Y. Garner and J. Harvey, Jr. (eds.), *Chemical and Biological Control in Forestry*, ACS Symp. Ser. 238, 1984; M. B. Greene and P. A. Hedin (eds.), *Natural Resistance of Plants to Pests: Role of Allelochemicals*, ACS Symp. Ser. 296, 1986; J. B. Harbone, *Introduction to Ecological Biochemistry*, 4th ed., 1994; M. M. Harlin, Allelochemistry in marine macroalgae, *CRC Crit. Rev. Plant Sci.*, 5:237–249, 1987; A. R. Putnam and Chung-Shih Tang (eds.), *The Science of Allelopathy*, 1986; E. L. Rice, *Allelopathy*, 2d ed. 1984; A. C. Thompson (ed.), *The Chemistry of Allelopathy: Biochemical Interactions among Plants*, ACS Symp. Ser. 268, 1985; G. R. Waller (ed.), *Allelochemicals: Role in Agriculture and Forestry*, ACS Symp. Ser. 330, 1987.

Allergy

Altered reactivity in humans and animals to allergens (substances foreign to the body that cause allergy) induced by exposure through injection, inhalation, ingestion, or skin contact. The most common clinical manifestations of allergy are hay fever, asthma, hives, atopic (endogenous) eczema, and eczematous skin lesions caused by direct contact with allergens such as poison ivy or certain chemicals.

Allergens. A large variety of substances may cause allergies: pollens, animal proteins, molds, foods, insect venoms, foreign serum proteins, industrial chemicals, and drugs. Most natural allergens are proteins or polysaccharides of moderate molecular size (molecular weights of 10,000 to 200,000). Chemicals or drugs of lower molecular weight (haptens) have first to bind to the body's own proteins (carriers) in order to become fully effective allergens.

For the development of the hypersensitivity state underlying clinical allergies, repeated contact with the allergen is required. Duration of the sensitization period is usually dependent upon the sensitizing strength of the allergen and the intensity of exposure. Some allergens (for example, saliva, urine, and hair proteins of domestic animals) are more sensitizing than others. In most instances, repeated contact with minute amounts of allergen is required: several annual seasonal exposures to grass pollens or ragweed pollen usually occur before an overt manifestation of hay fever. On the other hand, allergy to cow milk proteins in infants can develop within a few weeks. When previous contacts with allergens have not been apparent (for example, antibiotics in food), an allergy may become clinically manifest even upon the first conscious encounter with the offending substance.

Besides the intrinsic sensitizing properties of allergens, individual predisposition of the allergic person to become sensitized also plays an important role. Clinical manifestations, such as hay fever, allergic asthma, and atopic (endogenous) dermatitis, occur more frequently in some families: the inheritance of a capacity to develop these forms of allergy has been called atopy or the atopic state. These terms indicating a disease "not at its proper place" are misnomers but are sanctioned by historical usage. In other clinical forms of allergy, genetic predisposition, though possibly present as well, is not as evident.

Mechanisms. Exposure to sensitizing allergens may induce several types of immune response,

and the diversity of immunological mechanisms involved is responsible for the various clinical forms of allergic reactions which are encountered in practice. Three principal types of immune responses are encountered: the production of IgE antibodies, IgG or IgM antibodies, and sensitized lymphocytes. *See* ANTIBODY; IMMUNOGLOBULIN.

IgE antibodies. IgE antibodies belong to a peculiar class of serum immunoglobulins which is responsible for the majority of allergies of the so-called immediate or anaphylactic type. IgE antibodies are present in blood in very small amounts (normally less than 0.2 microgram per milliliter), but have the capacity to bind strongly to the membrane of tissue mast cells and blood basophils. These cells are the major effector components in an immediate allergic reaction. They contain a wide variety of preformed inflammatory mediators, such as histamine and serotonin. They also have the capacity to form, upon interaction with allergen, further substances causing inflammatory changes in tissue. The latter include increased vascular permeability that results in swelling and is manifested by redness and wheals as observed in hives. Among such mediators formed after cell stimulation, leukotrienes, the platelet-activating factor (PAF), and several other substances which attract other blood cells, such as eosinophils, to the site of reaction have gained the most attention. Mast cells and basophils may be compared to powder kegs which will be fired upon the interaction of allergen with its corresponding cell surface–bound IgE antibody. Understandably, such a reaction proceeds quite rapidly in a sensitized individual following renewed contact with allergen. Within minutes, hay fever patients develop symptoms after inhaling grass pollen, as do individuals allergic to bee venom after being stung. Generally, individuals possessing a genetic predisposition to atopic diseases are extremely prone to develop IgE antibodies to a variety of inhaled or ingested allergens. The IgE antibodies may also bind, although with lesser affinity, to some mononuclear cells and in particular to Langerhans cells in the skin. This phenomenon may be involved in the mechanism of some eczematous skin lesions, as seen in the frequent allergic condition termed atopic dermatitis. *See* HISTAMINE; SEROTONIN.

IgG or IgM antibodies. The production of IgG of IgM antibodies requires, as a rule, more massive exposure to allergens than that leading to the formation of IgE antibodies. Upon renewed contact with sufficient amounts of allergen, immune complexes may form with serum antibodies and be deposited in various tissues where they cause an acute inflammatory reaction. Preferential sites for such reactions, depending upon the mode of entry and distribution of the allergen, are alveolar walls of the lungs, glomeruli of kidneys, synovial membranes of joints, and walls of small blood vessels of the skin. Immune complex deposition below the basement membrane of blood vessels induces a cascade activation of complement (a multicomponent system of interacting blood proteins) which generates further active mediators of inflammation. The most important include the fragments of complement factors C3 and C5 (denominated C3a and C5a or anaphylatoxins) which are chemotactic (attracting) for white blood cells, especially polymorphonuclear leukocytes, attracting them to infiltrate the site of the reaction. Once arrived there, these cells proceed to destroy excess allergen and immune complexes, and by various mechanisms enhance the inflammatory reaction. The process takes several hours, and reactions of this type (Arthus reactions) become clinically apparent and peak only 6 to 12 h following contact with allergen. *See* IMMUNE COMPLEX DISEASE; INFLAMMATION.

Sensitized lymphocytes. Sensitized lymphocytes are responsible for the clinical manifestations of so-called delayed hypersensitivity. If an individual has been exposed to allergen and has developed sensitized lymphocytes capable of reacting with it, renewed exposure will trigger the lymphocytes to produce an array of glycoproteins called lymphokines. Lymphokines are also mediators capable of causing considerable inflammation in the surrounding tissues. Although produced at first only by the relatively small proportion of lymphocytes which specifically recognize and interact with allergen, lymphokines have amplifying effects. They stimulate, indiscriminately, other cell types, such as the macrophages and monocytes, attract these cells to the site of the reaction, and cause the formation of an inflammatory infiltrate. In the skin, such a reaction manifests itself by redness, swelling, and the formation of an inflammatory papule. A typical example is the inflammatory nodular reaction experienced upon injection in the skin of protein allergens produced by tubercle bacilli (tuberculin reaction). Such a reaction occurs only in people who were once infected with tubercle bacilli, and who have developed an immune response to these microorganisms.

Skin reactions occurring upon contact withbreak sensitizing agents such as poison ivy or industrial chemicals also are caused by the same immune mechanism. Since production of lymphokines by lymphocytes following stimulation by allergens takes several hours and since the reaction becomes clinically apparent only when a minimal number of infiltrating cells (mostly monocytes) have accumulated at the reaction site, the time elapsed between allergen contact and clinical reaction is usually 24 to 72 h. Such reactions are therefore termed delayed reactions, in contrast to the immediate (anaphylactic) reactions which occur within minutes after contact with allergen C. *See* HYPERSENSITIVITY.

Clinical forms. Among the clinical manifestations of allergy are allergic rhinitis, bronchial asthma, food allergy, occupational allergy, skin allergy, and anaphylactic shock.

Allergic rhinitis. The most frequent manifestation of allergy is undoubtedly allergic rhinitis. Allergic rhinitis was originally misnamed hay fever because it was believed that hay causes the disorder, and "fever" was a term loosely applied to many ailments, even when not accompanied by fever. However, in addition to nasal symptoms (watery nasal discharge, sneezing, runny eyes, and itchy nose, throat, and roof of the

mouth), the individual may experience fatigue, irritability, and loss of appetite.

In industrialized countries, the frequency of allergic rhinitis has been estimated to be 5 to 8% of the population and appears to be increasing. Allergic rhinitis may be seasonal, as the traditional hay fever, or perennial, occurring year-round. The seasonal form is usually caused by pollens from grasses, weeds, or trees. Related symptoms appear only during that time of the year when the pollens to which the individual is sensitive occur in the air. When allergens are present all the time, allergic rhinitis may occur year-round. Frequent causes are allergens present in the house or workplace, such as dust, molds, and animal danders.

Chronic or recurrent rhinitis, however, is not always related to allergy; in many instances it may be caused by an abnormal irritability of the nose (so-called intrinsic or vasomotor rhinitis) attributable to still ill-defined factors, such as hormonal factors or stress.

Bronchial asthma. The allergic disease which by its frequency and severity causes the most problems is bronchial asthma. Although an attack of asthma may be easily described, since it involves acutely occurring shortness of breath and wheezing, a general accurate definition, encompassing all possible causes and mechanisms of the disease, is still not available. Constriction of tiny bronchi in the lung airways, accompanied by an augmentation of bronchial secretions and the swelling of bronchial mucosa, contributes to obstruct the free passage of air through the lungs, producing wheezing. However, wheezing is not always indicative of asthma. Furthermore, in numerous instances, the cause of asthma is not related to allergy against some airborne allergens. It is customary to classify asthma as extrinsic, that is, due to contact with an external allergen, or intrinsic, that is, due to endogenous causes which are still poorly defined, such as stress or bronchial hyperreactivity owing to hormonal and other factors. However, even if the causes of disease may be very dissimilar, the mechanisms and particularly the mediators responsible for allergic inflammation in the lung may be the same, leading to almost indistinguishable clinical pictures.

The same airborne allergens which cause allergic rhinitis may also be responsible for seasonal or year-round asthma. Indeed, allergic rhinitis and asthma are often associated in the same individual. In addition, ingestion of allergens in foods and inhalation of drugs by aerosol or of sensitizing industrial chemicals at the workplace are frequent causes of allergic asthma.

Allergic asthma is considered a major public health problem owing to its severity, and to the possibility of death during an acute and prolonged asthma attack (status asthmaticus), and also because of the possible evolution to severe chronic respiratory insufficiency with invalidism in neglected cases. Available statistics indicate that in most countries 2–4% of the population is affected by asthma. In view of the large variety of possible causes of asthma, different mechanisms may be operating. In classical cases of allergic asthma due to airborne allergens, the presence of IgE antibodies probably plays a major role, but it is not excluded that other forms of immune reactions also occur. *See* ASTHMA.

Food allergy. This may occur as a reaction to natural products (for example, celery, milk, and egg proteins) as well as to chemicals or preservatives added to foods (for example, dyestuffs and antioxidants). Allergy to foods may manifest itself at the primary site of contact with allergen, namely in the gastrointestinal tract, causing symptoms such as diarrhea, abdominal pain, nausea, and vomiting. However, it may also take more general forms once the allergen has been absorbed, causing symptoms in the upper airways (rhinitis and asthma) and in the skin (hives), and, in cases of extreme hypersensitivity, generalized anaphylactic shock.

In infants from allergic families, the development of allergy to cow milk is quite frequent; strict breast feeding during the first months of life is therefore increasingly advocated. The breast-feeding mother should abstain from ingesting large quantities of potential allergens such as milk and eggs, since the food allergens can pass into the mother's milk and sensitize the child. After a few months, the susceptibility of infants to sensitization by ingestion appears to decrease markedly.

Occupational allergies. Such allergies become increasingly frequent in industrialized countries. Chemicals used in plastic and pharmaceutical industries or in metal refining and construction work are frequent offenders. However, natural organic materials, such as baker's flour, animal products, castor bean, silk, and furs, may be involved as well. Here, too, the most frequent manifestations involve the upper respiratory tract and the skin.

A peculiar form of occupational lung disease, hypersensitivity pneumonitis, is characterized by repeated episodes of cough, fever, chills, and shortness of breath, occurring 4–10 h after massive exposure to some airborne allergen. In contrast to allergic asthma, the lesions do not occur at the level of the bronchi but of the alveolar walls in the lung. The inflammatory reaction taking place there may have severe consequences when occurring repeatedly, since it leads to progressive fibrosis of the lung, impairment of gas exchanges with blood, and respiratory insufficiency. The mechanisms of this type of allergic reaction are not entirely elucidated, but it is known that IgE antibodies are not involved. IgG antibodies are frequently present and possibly cause the formation of immune complexes with allergen. However, delayed hypersensitivity mechanisms are probably at play as well. Among the most frequent offending allergens are molds sometimes found in hay (causing "farmer's lung") and in various other locations, such as air conditioners, cellars, and logs used by woodworkers.

Skin allergy. This condition may take several forms. Atopic dermatitis is a chronic superficial itching

inflammation of the skin (eczema) which occurs in 0.5–0.7% of the population and is often of long duration. Atopic dermatitis usually occurs early in life together with cow milk allergy (infantile eczema). Its evolution is unpredictable, and the disease may seriously affect psychological health and working ability of the individual. IgE antibodies frequently are notably increased. This disease exhibits the same genetic predisposition ("atopy") as allergic rhinitis and asthma.

Allergic contact dermatitis is characterized by a red rash, swelling, intense itching, and sometimes blisters at the site of contact with the allergen, which may be of plant origin (for example, poison ivy or turpentine) or may be an industrial chemical. Lesions are caused by sensitized lymphocytes and belong to the delayed hypersensitivity type of allergy.

Urticaria (hives) is a very common manifestation of skin allergy. When evoked by sensitization to some allergen, IgE antibodies are present. Foods, drugs such as the penicillins, and, more exceptionally, natural inhalation allergens such as pollens are the most frequently involved. However, utricaria, resulting from the release of inflammatory mediators (particularly histamine) by skin mast cells, is even more frequently caused by nonallergic mechanisms. Some foods (for example, crustaceans or strawberries) contain substances which, when ingested in sufficient quantities, can directly trigger mast cells without prior sensitization and formation of IgE antibodies.

Anaphylaxis. The most feared allergic reaction is an acute event known as anaphylactic shock. This is an IgE-mediated reaction with sudden onset: symptoms develop within minutes after exposure to the triggering allergen. The victim may experience a feeling of great anxiety; symptoms include swelling and skin redness often accompanied by hives, vomiting, abdominal cramps and diarrhea, life-threatening breathing difficulties due to swelling of the throat, and severe bronchospasm. A drastic fall in blood pressure may cause fatal shock within a few minutes. Insect stings and drugs such as the penicillins are now the most frequent causes of anaphylactic shock; in earlier times the therapeutic injection of animal serum proteins (for example, horse antitetanus serum) was often responsible. Some acute reactions, called anaphylactoid reactions, may cause similar symptoms but are not due to IgE-mediated sensitization. Such reactions are encountered mostly in people intolerant to some drugs, such as aspirin, or to contrast media used in diagnostic radiology. *See* ANAPHYLAXIS.

Diagnosis. Diagnosis of allergic diseases encompasses several facets. The large variety of clinical forms of allergic reactions and the multitude of allergens possibly involved frequently require prolonged observation and in-depth investigation of the individual's life habits. Since many clinical manifestations of allergy are mimicked by nonallergic mechanisms, it is usually necessary to use additional diagnostic procedures to ascertain whether the person has developed an immune response toward the incriminated allergen. Such procedures primarily consist of skin tests, in which a small amount of allergen is applied on or injected into the skin. If the individual is sensitized, a local immediate reaction ensues, taking the form of a wheal (for IgE-mediated reactions), or swelling and redness occur after several hours (for delayed hypersensitivity reactions). The blood may also be analyzed for IgE and IgG antibodies by serological assays, and sensitized lymphocytes are investigated by culturing them with the allergen.

Since the discovery of the responsible allergens markedly influences therapy and facilitates prediction of the allergy's outcome, it is important to achieve as precise a diagnosis as possible. Most of the tests described above indicate whether the individual is sensitized to a given allergen, but not whether the allergen is in fact still causing the disease. Since in most cases the hypersensitive state persists for many years, it may well happen that sensitization is detected for an allergen to which the individual is no longer exposed and which therefore no longer causes symptoms. In such cases, exposition tests, consisting of close observation of the individual after deliberate exposure to the putative allergen, may yield useful information.

Therapy. Therapy of allergic reactions may be performed at several levels. Whenever feasible, the most efficient treatment, following identification of the offending allergen, remains elimination of allergen from the person's environment and avoidance of further exposure. This treatment is essential for allergies caused by most household and workplace allergens.

Inflammatory reactions and corresponding symptoms caused by allergic mediators may be relieved by several drugs acting as pharmacological antagonists to these mediators, for example, antihistaminics or steroids. Other drugs act at a more central level and influence the ease with which mast cells or basophils release their mediators; this is apparently the case for chromones. Since the state of hypersensitivity usually remains present over many years, especially when contact with allergen is periodically renewed (as is the case with pollens), such an approach to allergy therapy may be only palliative.

Attempts to influence the hypersensitive state itself in IgE-mediated allergy are more decisive. Repeated injections of increasing doses of allergen progressively diminish the degree of hypersensitivity. This procedure, which must usually be pursued over several years, is referred to as hyposensitization or immunotherapy. *See* IMMUNOLOGY; IMMUNOTHERAPY.

A. L. de Weck

Bibliography. R. Davies and S. Ollier, *Allergy: The Facts*, 1989; H. F. Krause, *Otolaryngic Allergy and Immunology,* 1989; L. M. Lichtenstein and A. S. Fauci, *Current Therapy in Allergy, Immunology, and Rheumatology*, 5th ed., 1996; G. Melillo and J. A. Nadel (eds.), *Respiratory Allergy, Advances in Clinical Immunology and Pulmonary Medicine*, 1993; D. Reinhardt and E. Schmidt (eds.), *Food Allergy,* 1988; I. M. Roitt, *Essential Immunology,* 1991.

Alligator

A large aquatic reptile of the family Alligatoridae. Common usage generally restricts the name to the two species of the genus *Alligator*. The family also includes three genera (five species) of caiman. With the crocodiles and gharial (also spelled gavial), these are survivors of archosaurian stock and are considered close to the evolutionary line which gave rise to birds.

Alligator species. The two living species have a disjunct subtropical and temperate zone distribution. The American alligator (*A. mississippiensis*) ranges throughout the southeastern United States from coastal North Carolina (historically from southeastern Virginia) to the Rio Grande in Texas, and north into southeastern Oklahoma and southern Arkansas (see **illustration**).

Poaching and unregulated hunting for the valuable hide decimated the alligator populations until the animal was placed on the U.S. Endangered Species List in 1967. The species has responded well to protection and has become abundant in many areas in its range, particularly in Florida and parts of Georgia, Louisiana, and Texas, where it is now a common sight in the freshwater habitats, including swamps, marshes, lakes, rivers, and even roadside ditches.

The second species is the Chinese alligator (*A. sinensis*), restricted to the region of the Yangtze River valley, where it inhabits burrows in the floodplains and riverbanks. This species is considered "critically endangered" by international conservation organizations. It is believed that less than 200 individuals remain in the wild.

The American alligator is by far the larger of the two species, exceeding 15 ft (4.5 m). The average length of *A. sinensis* is 4–5 ft (1.2–1.5 m).

Caiman species. There are five species of caimans, once often sold as "baby alligators." Caimans differ from alligators in technical details of their internal anatomy and scale characteristics.

The black caiman (*Melanosuchus niger*) of the Amazon Basin of South America strongly resembles the American alligator in superficial appearance and may reach a length approaching 16 ft (5 m). The spectacled caiman (*Caiman crocodilus*), with its several subspecies, is sold in the pet trade. It is the most widely distributed of the caimans, ranging from southwestern Mexico to Paraguay and Bolivia. Spectacled caimans reach a length of 8 ft (2.5 m) and can be distinguished from alligators by the presence of a ridge across the snout between their eyes, the nose bridge of the "spectacles."

The broad-nosed caiman (*C. latirostris*) occurs from southern Brazil to Paraguay and northern Argentina and reaches a length of 9 ft (3 m). The two species of smooth-fronted caiman, *Paleosuchus trigonatus* and *P. palpebrosus*, are the smallest of the family, reaching a maximum length of 7 ft (2 m). They occur over most of northern South America, south to Peru, Bolivia, and southeastern Brazil.

Anatomical features. Alligators, including caimans, are generally distinguished from crocodiles by their broader, rounded, and more massive snout, the arrangement of their teeth, and other technical anatomical details. As in all crocodilians, the teeth are conical and sharp, equivalent in shape (homodont), and replaced throughout life. Moreover, as in all crocodilians, alligators possess a functionally four-chambered heart, a secondary palate, a septum or "diaphragm" that separates the lung and peritoneal regions, a compressed tail, webbed feet, and other adaptations for an aquatic existence.

Nesting. All alligators build a conical nest of available materials near the water's edge, where they deposit 15–60 leathery-shelled eggs. Incubation takes 30–60 days; females of at least some species may guard the nest, assist in the hatching process, and remain with the young for as long as 3 years. In some cases, females carry young in their mouth to nearby water.

Habitat and behaviors. Alligators are generalized carnivores, feeding on any invertebrates, fishes, reptiles and amphibians, birds, or mammals that can be caught and overpowered. They play a dominant role in the energy and nutrient cycling in their habitats, and through their construction of "gator holes," which often retain water in times of drought, are considered important to the survival of many others species.

Alligators have also been found to have a complex social system that relies heavily on communication. Visual, tactile, auditory, and even subauditory communications have been described in alligator populations. Males roar during the breeding season to attract females, and females bellow and growl to communicate receptiveness to males. Using trunk muscles and low-frequency sound, alligators create rapid vibrations on the water surface to communicate. The vibrations coupled with the low-frequency sound can be heard at great distances (perhaps up to a kilometer). Males use a diversity of sounds as a threat warning, including a "cough" and hiss. Juveniles employ distress calls when they feel threatened. These calls attract the attention of the nearby female. *See* ARCHOSAURIA; CARDIOVASCULAR SYSTEM; CROCODILE; CROCODYLIA; DENTITION; REPTILIA.

W. Ben Cash; Howard W. Campbell

Half-submerged American alligator (*Alligator mississippiensis*). (*Photo courtesy of W. Ben Cash*)

Bibliography. S. Grenard and W. Loutsenhizer, *Handbook of Alligators and Crocodiles*, Krieger,

Melbourne, FL, 1995; S. B. Hedges and L. L. Poling, A molecular phylogeny of reptiles, *Science*, 283:998–1001, 1999; D. W. Linzey, *Vertebrate Biology*, McGraw-Hill, New York, 2001; F. H. Pough et al., *Herpetology*, 3d ed., Prentice Hall, Upper Saddle River, NJ, 2004; K. A. Vliet, Social dynamics of the American alligator (*Alligator mississippiensis*), *Amer. Zool.*, 29:1019–1031, 1989; G. R. Zug, L. J. Vitt, and J. P. Caldwell, *Herpetology*, Academic Press, San Diego, 2001.

Allometry

The study of changes in the characteristics of organisms with body size. Characteristics such as body parts or timing of reproductive events do not necessarily change in direct proportion to body size, and the ways in which they change relative to body size can often provide insights into organisms, construction and behavior. Large organisms are often not merely magnified small ones.

Variables. Many characteristics, ranging from brain size and heart rate to life span and population density, change consistently with body size. These relationships normally fit a simple power function given by Eq. (1), where y is the variable under study,

$$y = km^b \tag{1}$$

m is body mass, k is the allometric constant, and b is the allometric exponent. The use of logarithms makes the equation easier to visualize—the exponent becomes the slope of a straight line when the logarithm of the variable (y) is plotted against the logarithm of body mass (m) [Eq. (2)]. Unless the allo-

$$\log(y) = \log(k) + b\log(m) \tag{2}$$

metric exponent (b) equals 1, the ratio of y/m varies with m. The terms isometry, positive allometry, and negative allometry are sometimes used for $b = 1$, $b > 1$, and $b < 1$, respectively. If the variable of interest is the mass of an organ, under isometry organ mass is a fixed proportion of body mass; under positive allometry larger organisms have disproportionately large organs; and under negative allometry larger organisms have disproportionately small organs. *See* LOGARITHM.

Proportions. The mass of the brain as a proportion of body mass from adults of different mammalian species is an example of negative allometry. Proportions vary among related species of different sizes (that is, interspecific allometry); for example, the brain is about 5% of body mass in a mouse but only 0.05% in a whale (**Fig. 1**). Proportions often vary within a species as well. Individuals change shape during growth, a phenomenon known as ontogenetic or growth allometry; for example, among humans the head is 25% of a baby's length but only 15% of an adult's height. Within species, small and large adults of the same sex may be isometric with respect to some bodily proportions, but there may also be static intraspecific allometry. The different

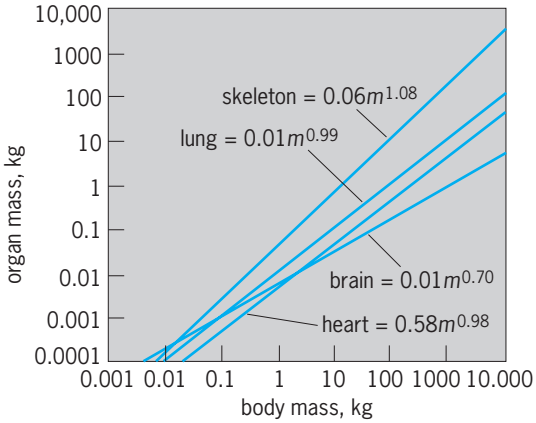

Fig. 1. Allometric lines for the mass of some organs across mammal species. (These are best-fit lines: individual species may lie above or below them.) Note that the skeleton is disproportionately heavy in large species (positive allometry), whereas the brain is disproportionately light (negative isometry). Heart and lungs scale approximately isometrically.

types of allometry (**Fig. 2**) often show a character scaling in different ways, reflecting different underlying causes.

Even if small and large organisms are the same shape, size is significant because surface area and volume scale differently with body length. Consider what happens to a cube when the length of each side is doubled: the surface area increases fourfold, the volume (and hence mass) eightfold. The same relations hold for any shape: the surface area is proportional to the length squared and the mass is proportional to the length cubed. It follows, therefore, that larger organisms have less surface area per unit mass. This relationship has profound consequences, helping to explain, for instance, why shrews and hummingbirds are the smallest endotherms—any organism much smaller would lose heat too quickly through its relatively large surface. Size is also

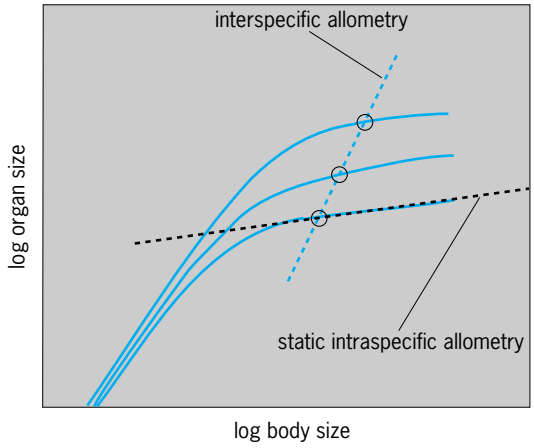

Fig. 2. Three different types of allometry. Each curve shows the growth allometry of organ size in a different species. The adults of each species show static intraspecific allometry. The circles indicate a particular stage of the life history (for example, sexual maturity). The line of interspecific allometry links points from different species.

significant for rates; for example, longer limbs can take longer strides, so larger animals need fewer strides to cover a given distance. Across mammal species, limb length scales as $m^{0.35}$, so the frequency of strides of two animals running at the same speed will differ according to their differences in $m^{0.35}$.

Allometric lines. In themselves, allometric lines are rough and probably simplified descriptions of the combined effects of the many (usually unknown) selective forces correlated with size; it still must be explained why the exponent takes the value that it does. For instance, in the skeletons of geometrically similar land animals, mass is proportional to length cubed while the cross-sectional area of any given bone, and hence its strength, scales as the square of its length. Beyond some critical size, these animals would collapse under their own weight. In fact, skeletal mass is proportional to $m^{1.08}$ for mammals and birds because bones face static (compressing) stress when the animal is stationary and elastic (buckling) stress during locomotion. Mechanical principles can be used to ascertain whether skeletons are adapted to prevent fracture by compression or by buckling. *See* BONE.

The leg bones of a quadruped with a vertical posture are like tall, thin, cylinders. To withstand compression equally at all body weights—a condition known as static stress similarity—their cross-sectional area must increase as m^1 and their length scales as $m^{1/3}$, so bone mass should scale as $(m^1)(m^{1/3}) = m^{4/3}$. But bone mass is proportional to $m^{1.08}$, so skeletons do not show static stress similarity.

Bones may instead be adapted to withstand buckling equally at all body sizes—the condition known as elastic stress similarity. The limb bones of hoofed mammals fit the model of a column, but those of other mammals do not, perhaps because their limbs are less pillarlike and so face different stresses. Elastic stress similarity is also found further afield. Limb bones of bipedal theropod dinosaurs fit the theory. Tree trunks and limbs are also thin columns faced with buckling stresses: tree trunks retain elastic stress similarity as they grow, and branch thickness scales with length as predicted. *See* STRESS AND STRAIN.

Many allometries are less well understood. Across mammal species, for example, small mammals clearly get enough oxygen for their maximal needs, so large species would seem to have more lung than they need.

Deviations from the line. Allometric lines are often used as baselines for comparison: if it is assumed that a line shows how organ mass has to change with body mass in order to keep doing the same job, organisms above or below the line have larger or smaller organs than expected for their body size. The question then arises as to why particular groups deviate from the line. Finding other variables that are correlated with this body-size-independent variation allow for powerful tests of hypotheses of adaptation. *See* ADAPTATION (BIOLOGY).

For example, across mammal species the scaling of brain mass (proportional to $m^{0.7}$) is similar to the scaling of metabolic rate (proportional to $m^{0.75}$). The similarity prompted the suggestion that brain size of newborn mammals (which is linearly related to adult brain size) scales as $m^{0.75}$ because it is somehow constrained by the mother's metabolic rate. However, another hypothesis also predicts an exponent near 0.7; surface area scales as $m^{2/3}$, as discussed earlier. If processing requirements are set by the amount of peripheral sensory input, brain size might scale as $m^{2/3}$ too. Choosing between these hypotheses without very good estimates of the exponent seems impossible. However, deviations from the line provide a critical test. The metabolic rate argument predicts that species with fast metabolisms for their size will also have relatively large brains; however, data do not support this prediction. Once deviations are understood, they can be used to make inferences. For instance, eye-socket size scales allometrically among extant primate species, but nocturnal species lie above the line. This knowledge leads to the inference of the activity patterns of species known only from fossils.

Using allometry in this way assumes that the part under study has the same function throughout the group being considered. If function differs, so may the scaling. For example, wing length scales as $m^{1/3}$ across most bird species. Hummingbird wings, however, must support the bird during hovering as well as more conventional flight, and hovering requires relatively larger wings, so hummingbird wing length scales as $m^{2/3}$.

Models. Allometry is perhaps most powerful when known allometric relations are coupled with theory to produce new hypotheses that can be tested. For example, mammal species with large bodies grow more slowly, mature later, have fewer but larger young after longer gestation periods, and live longer than do small species; these traits remain correlated with each other among deviations from the allometric lines. Allometric models predict not only the allometric exponents but also the correlations among the deviations. The models even provide a framework for asking questions vital to understanding scaling, such as what factors affect the evolution of body size itself. *See* BIOPHYSICS. Andy Purvis; Paul H. Harvey

Bibliography. R. McN. Alexander, *Optima for Animals*, 1982; E. L. Charnov, *Life History of Invariants*, 1993; P. H. Harvey and J. R. Krebs, Comparing brains, *Science*, 249:140–146, 1990; P. H. Harvey and M. D. Pagel, *The Comparative Method in Evolutionary Biology*, 1991; T. A. McMahon and J. T. Bonner, *On Size and Life*, 1983; K. Schmidt-Nielsen, *Scaling: Why Is Animal Size So Important?*, 1984.

Allosteric enzyme

Any one of the special bacterial enzymes involved in regulatory functions. End-product inhibition is a bacterial control mechanism whereby the end product of a biosynthetic pathway can react with the first enzyme of the pathway and prevent its activity. This end-product inhibition is a device through which

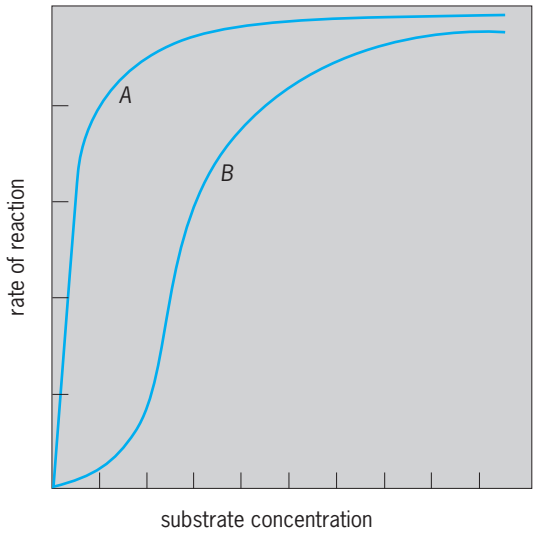

Fig. 1. Rate of reaction of an enzyme in relation to substrate concentration. Curve A obtained with ordinary enzymes, curve B obtained with allosteric enzymes.

a cell conserves its economy by shutting off the synthesis of building blocks when too many are present. *See* BACTERIAL PHYSIOLOGY AND METABOLISM.

Kinetics. Enzymes are protein substances which act as biological catalysts in converting one substance, the substrate, to another, the product. To understand the nature of the special allosteric enzymes, they must first be compared with ordinary enzymes as to their kinetics, that is, the way they work with increasing amounts of substrate. With ordinary enzymes the rate of reaction increases as the amount of substrate is increased until eventually a saturation level is obtained and no further increase occurs. The curve obtained is described as a hyperbola (**Fig. 1**). From this curve it can be said that the substrate molecules tend to occupy more and more binding sites until all the sites are in use. The allosteric enzymes, when similarly plotted, instead of showing a hyperbolic shape, often show a sigmoid, or S-shaped, curve. This means that there is a cooperative effect between substrate molecules so that, when the first molecules are bound to the enzyme, the subsequent ones are bound more readily. A clue to this concept of cooperative substrate interaction was obtained by the study of how the hemoglobin of the blood reacts with oxygen. Here, too, the curve describing the saturation of hemoglobin with oxygen traces a sigmoid shape.

Inhibition. Another characteristic of enzymes is that their action can be prevented by inhibitors. In ordinary enzymes the inhibitors are substances which closely resemble the structure of the substrate. Because of this structural similarity, they can jam the site that the substrate normally occupies. They can thus be likened to a false coin which fits the slot of the machine but does not operate it. The surprising feature about the inhibitors of allosteric enzymes is their lack of structural similarity with the substrate. Hence, they cannot be expected to work by competing with the substrate for a common slot. Instead, they apparently have their own special site of attachment, and when they become attached there, they somehow alter the site where the substrate fits. The substrate kinetics of the inhibited reaction supports this concept. Furthermore, S-shaped curves are also obtained when the inhibition rates are plotted against amounts of the inhibitor. This means that there is cooperative interaction between inhibitor molecules in the same way as there is cooperative interaction between substrate molecules.

Activation. The initial idea of allosteric inhibition was used as an explanation for the negative feedback control shown by the regulatory enzymes. There are also times when the control must have a positive effect, and these are brought about by allosteric activators. Consider, for example, two parallel production lines of metabolic activity leading to two different end products that must be built into a common structural unit. If there is too much of one of the building blocks, then the regulatory mechanisms can react in two ways. In one case the end product can shut off its own production by negative feedback control, and in the other case it can speed up the other parallel line by positive control. In this way the activity of allosteric enzymes may be either decreased by inhibitors or increased by activators. Again, if the substrate curves are studied with and without the activator, it is found that the allosteric activator converts the S-shaped curve into the ordinary hyperbolic type.

Desensitization. It is possible to treat allosteric enzymes in various ways so that they lose their sensitivity to the allosteric inhibitor. The enzymes, though desensitized with respect to inhibition, retain their activity with the substrate and, indeed, this activity is often increased to the same extent as is obtained with an activator. The S-shaped kinetics is changed to the normal variety. The agents that can cause this desensitization include heat, acid, heavy metals such as mercury, and agents which react with sulfhydryl (SH) groups on the proteins. These agents are known to be able to separate complex proteins into subunits. Allosteric enzymes are therefore considered to be complex proteins made up of identical monomers. The single unit, the monomer, is active with normal hyperbolic kinetics and is not affected by the regulatory molecules (inhibitors and activators). The multiple structure, several units combined, is the form that shows the properties of allosterism, S-shaped kinetics, and alteration by the regulators.

Allosteric model. The above considerations have led to the development of a model to explain the action of allosteric enzymes (**Fig. 2**). It suggests that the enzyme molecule is a complex consisting of identical subunits, each of which has a site for binding the substrate and another for binding the regulatory substance. These subunits interact in such a way that two conformational forms may develop. One form is in a relaxed condition (R state) and has affinity for the substrate and activator; the other form is in a constrained or taut condition (T state) and has affinity for the inhibitor. The forms exist in a state of equilibrium, but this balance can be readily tipped

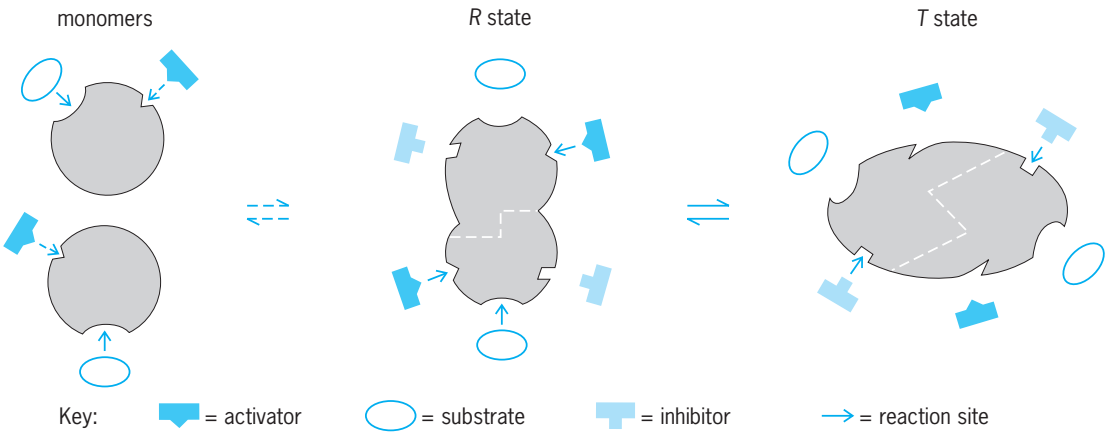

Fig. 2. Model of allosteric enzyme composed of two subunit monomers. Regulatory effects are obtained by shifting back and forth between the two states, relaxed (R state) and taut (T state). In the R state it reacts with substrate and activators, and in the T state it reacts with inhibitors. (*After J. Monod, J. Wyman, and J.-P. Changeux, On the nature of allosteric transitions: A plausible model, J. Mol. Biol., 12:88, 1965*)

by binding one of the reactants. The substrate and activator are bound by the relaxed form; when this happens, the balance is tipped in favor of that state. Conversely, the inhibitor will throw the balance toward the constrained state. The balance is thus tipped one way or the other, depending on the relative concentrations of substrate and inhibitor. Since the two states require subunit interaction for their maintenance, it can be seen why dissociation of the subunits leads to a simple monomeric enzyme which no longer exhibits allosteric effects. The model also shows how the binding sites may interact in either a cooperative or antagonistic manner. *See* ENZYME.

Joseph S. Gots

Bibliography. J.-P. Changeux, Control of biochemical reactions, *Sci. Amer.*, 212(4):36–45, 1965; B. D. Davis, *Microbiology*, 3d ed., 1980; J. Monod, J. Wyman, and J.-P. Changeux, On the nature of allosteric transitions: A plausible model, *J. Mol. Biol.*, 12:88, 1965.

Allotheria

One of the four subclasses of Mammalia, containing a single order, the Multituberculata. The Allotheria first appeared in the Late Jurassic and survived well into the Cenozoic, a period of at least 100,000,000 years. Fossils are known from North America, Europe, and Asia.

The diagnostic features of the subclass are in the dentition, which is very specialized. There is a pair of enlarged incisors above and below; reduced lateral incisors may persist in the upper jaw. Canines are absent, leaving a diastema between incisors and cheek teeth. Premolars are variable and often reduced. The lower molars have five or more cusps in two parallel longitudinal rows, the upper molars two or three parallel rows of cusps; hence the molars are multituberculate. *See* DENTITION; MAMMALIA; MULTITUBERCULATA.

D. Dwight Davis; Frederick S. Szalay

Allowance

An intentional difference in sizes of two mating parts. With running or sliding fits, in which mating parts move relative to each other, allowance is a clearance, usually for a film of oil. In this sense, allowance is the space "allowed" for motion between parts. To avoid binding between parts, a minimum clearance is critical for a running fit.

With force or shrink fits, in which mating parts, once assembled, are fixed in position relative to each other, allowance is an interference of metal; that is, a portion of metal in one part tends to occupy the same space as the adjacent portion of metal in the mating part. In this sense, allowance is the interference "allowed" to produce pressure between parts. To avoid breaking the external part, a maximum interference is critical for a drive, force, or shrink fit.

Fits are classed as running or sliding fits for parts that move freely against each other. Location fits provide accurate orientation of mating parts. Force fits require appreciable assembly pressure and produce more or less permanent assembly of parts. *See* FORCE FIT; LOCATION FIT; PRESS FIT; RUNNING FIT; SHRINK FIT; TOLERANCE.

Paul H. Black

Alloy

A metal product containing two or more elements as a solid solution, as an intermetallic compound, or as a mixture of metallic phases. This article will describe alloys on the basis of their technical applications. Alloys may also be categorized and described on the basis of compositional groups. *See* INTERMETALLIC COMPOUNDS; METAL; SOLID SOLUTION.

Except for native copper and gold, the first metals of technological importance were alloys. Bronze, an alloy of copper and tin, is appreciably harder than copper. This quality made bronze so important an alloy that it left a permanent imprint on the

civilization of several millennia ago now known as the Bronze Age. Today the tens of thousands of alloys involve almost every metallic element of the periodic table. *See* BRONZE; PERIODIC TABLE; PREHISTORIC TECHNOLOGY.

Alloys are used because they have specific properties or production characteristics that are more attractive than those of the pure, elemental metals. For example, some alloys possess high strength, others have low melting points, others are refractory with high melting temperatures, some are especially resistant to corrosion, and others have desirable magnetic, thermal, or electrical properties. These characteristics arise from both the internal and the electronic structure of the alloy. For a discussion of the physical structures of alloys and the interatomic forces in alloys. *See* ALLOY STRUCTURES.

Bearing alloys. These alloys are used for metals that encounter sliding contact under pressure with another surface; the steel of a rotating shaft is a common example. Most bearing alloys contain particles of a hard intermetallic compound that resist wear. These particles, however, are embedded in a matrix of softer material which adjusts to the hard particles so that the shaft is uniformly loaded over the total surface. The most familiar bearing alloy is babbitt metal, which contains 83–91% tin (Sn); the remainder is made up of equal parts of antimony (Sb) and copper (Cu), which form hard particles of the compounds SbSn and CuSn in a soft tin matrix (**Fig. 1**). Other bearing alloys are based on cadmium (Cd), copper, or silver (Ag). For example, an alloy of 70% copper and 30% lead (Pb) is used extensively for heavily loaded bearings. Bearings made by powder metallurgy techniques are widely used. These techniques are valuable because they permit the combination of materials which are incompatible as liquids, for example, bronze and graphite. Powder techniques also permit controlled porosity within the bearings so that they can be saturated with oil before being used, the so-called oilless bearings. *See* ANTIFRICTION BEARING; COPPER ALLOYS; POWDER METALLURGY; WEAR.

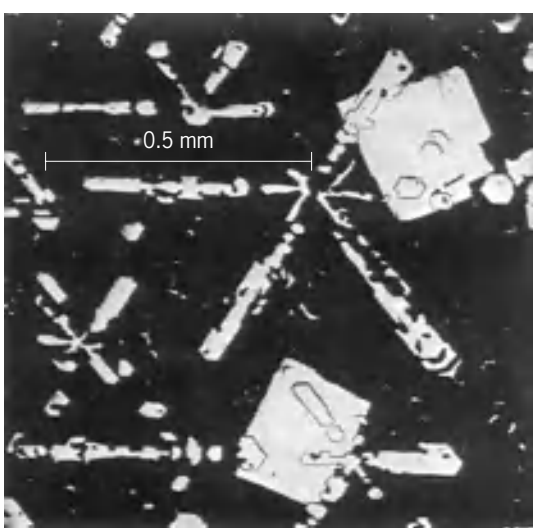

Fig. 1. Babbitt metal (84% Sn–7% Cu–9% Sb). The white areas are crystals of CuSn (star dendrites) and SbSn (nearly square); these resist wear. The dark matrix is a soft, tin-rich alloy, which adjusts to the hard particles to give uniform loading. (*From R. M. Brick et al., Structure and Properties of Alloys, 3d ed., McGraw-Hill, 1965*)

Corrosion-resisting alloys. Certain alloys resist corrosion because they are noble metals. Among these alloys are the precious-metal alloys, which will be discussed separately. Other alloys resist corrosion because a protective film develops on the metal surface. This passive film is an oxide which separates the metal from the corrosive environment. Stainless steels and aluminum alloys exemplify metals with this type of protection. Stainless steels are iron alloys containing more than 12% chromium (Cr). Steels with 18% Cr and 8% nickel (Ni) are the best known and possess a high degree of resistance to many corrosive environments. Aluminum (Al) alloys gain their corrosion-deterring characteristics by the formation of a very thin surface layer of aluminum oxide (Al_2O_3), which is inert to many environmental liquids. This layer is intentionally thickened in commercial anodizing processes to give a more permanent Al_2O_3 coating. Monel, an alloy of approximately 70% nickel and 30% copper, is a well-known corrosion-resisting alloy which also has high strength. Another nickel-base alloy is Inconel, which contains 14% chromium and 6% iron (Fe). The bronzes, alloys of copper and tin, also may be considered to be corrosion-resisting. *See* ALUMINUM ALLOYS; CORROSION; IRON ALLOYS; NICKEL ALLOYS; STAINLESS STEEL.

Dental alloys. Amalgams are predominantly alloys of silver and mercury (Hg), but they may contain minor amounts of tin, copper, and zinc (Zn) for hardening purposes, for example, 33% silver, 52% mercury, 12% tin, 2% copper, and less than 1% zinc. Liquid mercury is added to a powder of a precursor alloy of the other metals. After being compacted, the mercury diffuses into the silver-base metal to give a completely solid alloy. Gold-base dental alloys are preferred over pure gold (Au) because gold is relatively soft. The most common dental gold alloy contains gold (80–90%), silver (3–12%), and copper (2–4%). For higher strengths and hardnesses, palladium and platinum (up to 3%) are added, and the copper and silver are increased so that the gold content drops to 60–70%. Vitallium [an alloy of cobalt (65%), chromium (5%), molybdenum (3%), and nickel (3%)] and other corrosion-resistant alloys are used for bridgework and special applications. *See* AMALGAM; DENTISTRY; GOLD ALLOYS; SILVER ALLOYS.

Die-casting alloys. These alloys have melting temperatures low enough so that in the liquid form they can be injected under pressure into steel dies. Such castings are used for automotive parts and for office and household appliances which have moderately complex shapes. This processing procedure eliminates the need for expensive machining and forming operations. Most die castings are made from zinc-base or aluminum-base alloys. Magnesium-base alloys also find some application when weight reduction is paramount. Low-melting alloys of lead and tin are not common because they lack the necessary

strength for the above applications. A common zinc-base alloy contains approximately 4% aluminum and up to 1% copper. These additions provide a second phase in the metal to give added strength. The alloy must be free of even minor amounts (less than 100 ppm) of impurities such as lead, cadmium, or tin, because impurities increase the rate of corrosion. Common aluminum-base alloys contain 5–12% silicon (Si), which introduces hard-silicon particles into the tough aluminum matrix. Unlike zinc-base alloys, aluminum-base alloys cannot be electroplated; however, they may be burnished or coated with enamel or lacquer. Advances in high-temperature die-mold materials have focused attention on the die-casting of copper-base and iron-base alloys. However, the high casting temperatures introduce costly production requirements, which must be justified on the basis of reduced machining costs. *See* METAL CASTING.

Eutectic alloys. In certain alloy systems, a liquid of a fixed composition freezes to form a mixture of two basically different solids or phases. An alloy that undergoes this type of solidification process is called a eutectic alloy. A typical eutectic alloy is formed by combining 28.1% of copper with 71.9% of silver. A homogeneous liquid of this composition on slow cooling freezes to form a mixture of particles of nearly pure copper embedded in a matrix (background) of nearly pure silver.

The advantageous mechanical properties inherent in composite materials such as plywood composed of sheets or lamellae of wood bonded together and fiber glass in which glass fibers are used to reinforce a plastic matrix have been known for many years. Attention is being given to eutectic alloys because they are basically natural composite materials. This is particularly true when they are directionally solidified so as to yield structures with parallel plates of the two phases (lamellar structure) or long fibers of one phase embedded in the other phase (fibrous structure). Directionally solidified eutectic alloys are being given serious consideration for use in fabricating jet engine turbine blades. For this purpose eutectic alloys that freeze to form tantalum carbide (TaC) fibers in a matrix of a cobalt-rich alloy have been heavily studied. *See* EUTECTICS; METAL MATRIX COMPOSITE.

Fusible alloys. These alloys generally have melting temperatures below that of tin (450°F or 232°C), and in some cases as low as 120°F (50°C). Using eutectic compositions of metals such as lead, cadmium, bismuth, tin, antimony, and indium achieves these low melting temperatures. These alloys are used for many purposes, for example, in fusible elements in automatic sprinklers, forming and stretching dies, filler for thin-walled tubing that is being bent, and anchoring dies, punches, and parts being machined. Alloys rich in bismuth (Bi) were formerly used for type metal because these low-melting metals exhibited a slight expansion on solidification, thus replicating the font perfectly for printing and publication.

High-temperature alloys. Energy conversion is more efficient at high temperatures than at low; thus the need in power-generating plants, jet engines, and gas turbines for metals which have high strengths at high temperatures is obvious. In addition to having strength, these alloys must resist oxidation by fuel-air mixtures and by steam vapor. At temperatures up to about 1380°F (750°C), the austenitic stainless steels (18% Cr–8% Ni) serve well. An additional 180°F (100°C) may be realized if the steels also contain 3% molybdenum. Both nickel-base and cobalt-base alloys, commonly categorized as superalloys, may serve useful functions up to 2000°F (1100°C). Nichrome, a nickel-base alloy containing 12–15% chromium and 25% iron, is a fairly simple superalloy. More sophisticated alloys invariably contain five, six, or more components; for example, an alloy called René-41 contains approximately 19% chromium, 1.5% aluminum, 3% titanium (Ti), 11% cobalt (Co), 10% molybdenum, 3% iron, 0.1% carbon (C), 0.005% boron (B), and the balance nickel. Other alloys are equally complex. The major contributor to strength in these alloys is the solution-precipitate phase of Ni_3 (Ti,Al), γ'. It provides strength because it is coherent with the nickel-rich γ phase. Cobalt-base superalloys may be even more complex and generally contain carbon which combines with the tungsten and chromium to produce carbides that serve as the strengthening agent (**Fig. 2**). In general, the cobalt-base superalloys are more resistant to oxidation than the nickel-base alloys are, but they are not as strong. Molybdenum-base alloys have exceptionally high strength at high temperatures, but their brittleness at lower temperatures and their poor oxidation resistance at high temperatures have limited their use. However, coatings permit the use of such alloys in an oxidizing atmosphere, and they are finding increased application. A group of materials called cermets, which are mixtures of metals and compounds

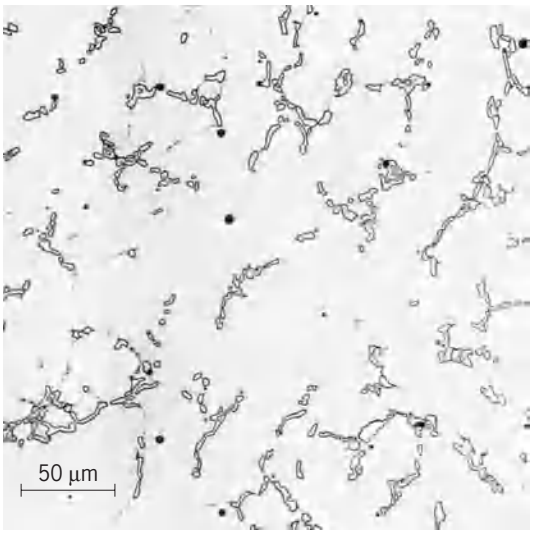

Fig. 2. Cobalt-base superalloy HS31. The dispersed phase is a carbide, $(CoCrW)_6C$, which strengthens the metal. (*Haynes Stellite Co.*)

such as oxides and carbides, have high strength at high temperatures, and although their ductility is low, they have been found to be usable. One of the better-known cermets consists of a mixture of titanium carbide and nickel, the nickel acting as a binder or cement for the carbide. *See* CERMET.

Joining alloys. Metals are bonded by three principal procedures: welding, brazing, and soldering. Welded joints melt the contact region of the adjacent metal; thus the filler material is chosen to approximate the composition of the parts being joined. Brazing and soldering alloys are chosen to provide filler metal with an appreciably lower melting point than that of the joined parts. Typically, brazing alloys melt above 750°F (400°C), whereas solders melt at lower temperatures. A 57% copper–42% zinc–1% tin brass is a general-purpose alloy for brazing steel and many nonferrous metals. A silicon-aluminum eutectic alloy is used for brazing aluminum, and an aluminum-containing magnesium eutectic alloy brazes magnesium parts. The most common solders are based on lead-tin alloys. The prevalent 60% tin–40% lead solder is eutectic in composition and is used extensively for electrical circuit production, in which temperature limitations are critical. A 35% tin–65% lead alloy has a range of solidification and is thus preferred as a wiping solder by plumbers. *See* BRAZING; SOLDERING; WELDING AND CUTTING OF METALS.

Light-metal alloys. Aluminum and magnesium (Mg), with densities of 1.6 and 1.01 oz/in.^{3} (2.7 and 1.75 g/cm^3), respectively, are the bases for most of the light-metal alloys. Titanium (2.6 oz/in.^{3} or 4.5 g/cm^3) may also be regarded as a light-metal alloy if comparisons are made with metals such as steel and copper. Aluminum and magnesium must be hardened to receive extensive application. Age-hardening processes are used for this purpose. Typical alloys are 90% aluminum–10% magnesium, 95% aluminum–5% copper, and 90% magnesium–10% aluminum. Ternary (three-element) and more complex alloys are very important light-metal alloys because of their better properties. The aluminum-zinc-magnesium system of alloys, used extensively in aircraft applications, is a prime example of one such alloy system. *See* ALUMINUM; MAGNESIUM; TITANIUM.

Low-expansion alloys. This group of alloys includes Invar (64% iron–36% nickel), the dimensions of which do not vary over the atmospheric temperature range. It has special applications in watches and other temperature-sensitive devices. Glass-to-metal seals for electronic and related devices require a matching of the thermal-expansion characteristics of the two materials. Kovar (54% iron–29% nickel–17% cobalt) is widely used because its expansion is low enough to match that of glass. *See* THERMAL EXPANSION.

Magnetic alloys. Soft and hard magnetic materials involve two distinct categories of alloys. The former consists of materials used for magnetic cores of transformers and motors, and must be magnetized and demagnetized easily. For alternating-current applications, silicon-ferrite is commonly used. This is an alloy of iron containing as much as 5% silicon. The silicon has little influence on the magnetic properties of the iron, but it increases the electric resistance appreciably and thereby decreases the core loss by induced currents. A higher magnetic permeability, and therefore greater transformer efficiency, is achieved if these silicon steels are grain-oriented so that the crystal axes are closely aligned with the magnetic field. Permalloy (78.5% nickel–21.5% iron) and some comparable cobalt-base alloys have very high permeabilities at low field strengths, and thus are used in the communications industry. Ceramic ferrites, although not strictly alloys, are widely used in high-frequency applications because of their low electrical conductivity and negligible induced energy losses in the magnetic field. Permanent or hard magnets may be made from steels which are mechanically hardened, either by deformation or by quenching. Some precipitation-hardening, iron-base alloys are widely used for magnets. Typical of these are the Alnicos, for example, Alnico-4 (55% iron–28% nickel–12% aluminum–5% cobalt). Since these alloys cannot be forged, they must be produced in the form of castings. Hard magnets are being produced from alloys of cobalt and the rare-earth type of metals. The compound RCo_5, where R is samarium (Sm), lanthanum (La), cerium (Ce), and so on, has extremely high coercivity.

Precious-metal alloys. In addition to their use in coins and jewelry, precious metals such as silver, gold, and the heavier platinum metals are used extensively in electrical devices in which contact resistances must remain low, in catalytic applications to aid chemical reactions, and in temperature-measuring devices such as resistance thermometers and thermocouples. The unit of alloy impurity is commonly expressed in karats, when each karat is 1/24 part. The most common precious-metal alloy is sterling silver (92.5% silver, with the remainder being unspecified, but usually copper). The copper is very beneficial in that it makes the alloy harder and stronger than pure silver. Yellow gold is an Au-Ag-Cu alloy with approximately a 2:1:1 ratio. White gold is an alloy which ranges from 10 to 18 karats, the remainder being additions of nickel, silver, or zinc, which change the color from yellow to white. The alloy 87% platinum (Pt)–13% rhodium (Rh), when joined with pure platinum, provides a widely used thermocouple for temperature measurements in the 1830–3000°F (1000–1650°C) temperature range. *See* GOLD; SILVER.

Shape memory alloys. These alloys have a very interesting and desirable property. In a typical case, a metallic object of a given shape is cooled from a given temperature T_1 to a lower temperature T_2 where it is deformed so as to change its shape. Upon reheating from T_2 to T_1, the shape change accomplished at T_2 is recovered so that the object returns to its original configuration. This thermoelastic property of the shape memory alloys is associated with the fact that they undergo a martensitic phase transformation (that is, a reversible change in crystal structure

that does not involve diffusion) when they are cooled or heated between T_1 and T_2.

For a number of years the shape memory materials were essentially scientific curiosities. Among the first alloys shown to possess these properties was one of gold alloyed with 47.5% cadmium. Considerable attention has been given to an alloy of nickel and titanium known as nitinal. The interest in shape memory alloys has increased because it has been realized that these alloys are capable of being employed in a number of useful applications. One example is for thermostats; another is for couplings on hydraulic lines or electrical circuits. The thermoelastic properties can also be used, at least in principle, to construct heat engines that will operate over a small temperature differential and will thus be of interest in the area of energy conversion. *See* SHAPE MEMORY ALLOYS.

Thermocouple alloys. These include Chromel, containing 90% nickel and 10% chromium, and Alumel, containing 94% nickel, 2% aluminum, 3% chromium, and 1% silicon. These two alloys together form the widely used Chromel-Alumel thermocouple, which can measure temperatures up to 2200°F (1204°C). Another common thermocouple alloy is constantan, consisting of 45% nickel and 55% copper. It is used to form iron-constantan and copper-constantan couples, used at lower temperatures. For precise temperature measurements and for measuring temperatures up to 3000°F (1650°C), thermocouples are used in which one metal is platinum and the other metal is platinum plus either 10 or 13% rhodium. *See* LASER ALLOYING; STEEL; THERMOCOUPLE. Lawrence H. Van Vlack; Robert E. Reed-Hill

Prosthetic alloys. As discussed here, prosthetic alloys are alloys used in internal prostheses, that is, surgical implants such as artificial hips and knees. External prostheses are devices that are worn by patients outside the body; alloy selection criteria are different from those for internal prostheses. In the United States, surgeons use about 250,000 artificial hips and knees and about 30,000 dental implants per year.

Alloy selection criteria for surgical implants can be stringent primarily because of biomechanical and chemical aspects of the service environment. Mechanically, an implant's properties and shape must meet anticipated functional demands; for example, hip joint replacements are routinely subjected to cyclic forces that can be several times body weight. Therefore, intrinsic mechanical properties of an alloy, for example, elastic modulus, yield strength, fatigue strength, ultimate tensile strength, and wear resistance, must all be considered. Likewise, because the pH and ionic conditions within a living organism define a relatively hostile corrosion environment for metals, corrosion properties are an important consideration. Corrosion must be avoided not only because of alloy deterioration but also because of the possible physiological effects of harmful or even cytotoxic corrosion products that may be released into the body. (Study of the biological effects of biomaterials is a broad subject in itself, often referred to as biocompatibility.) The corrosion resistance of all modern alloys stems primarily from strongly adherent and passivating surface oxides, such as titanium oxide (TiO_2) on titanium-based alloys and chromium oxide (Cr_2O_3) on cobalt-based alloys.

The most widely used prosthetic alloys therefore include high-strength, corrosion-resistant ferrous, cobalt-based, or titanium-based alloys: for example, cold-worked stainless steel; cast Vitallium; a wrought alloy of cobalt, nickel, chromium, molybdenum, and titanium; titanium alloyed with aluminum and vanadium; and commercial-purity titanium. Specifications for nominal alloy compositions are designated by the American Society for Testing and Materials (ASTM).

Prosthetic alloys have a range of properties. Some are easier than others to fabricate into the complicated shapes dictated by anatomical constraints. Fabrication techniques include investment casting (solidifying molten metal in a mold), forging (forming metal by deformation), machining (forming by machine-shop processes, including computer-aided design and manufacturing), and hot isostatic pressing (compacting fine powders of alloy into desired shapes under heat and pressure). Cobalt-based alloys are difficult to machine and are therefore usually made by casting or hot isostatic pressing. Some newer implant designs are porous-coated, that is, they are made from the standard ASTM alloys but are coated with alloy beads or mesh applied to the surface by sintering or other methods. The rationale for such coatings is implant fixation by bone ingrowth.

Some alloys are modified by nitriding or ion-implantation of surface layers of enhanced surface properties. A key point is that prosthetic alloys of identical composition can differ substantially in terms of structure and properties, depending on fabrication history. For example, the fatigue strength approximately triples for hot isostatically pressing versus as-cast Co-Cr-Mo alloy, primarily because of a much smaller grain size in the microstructure of the former.

No single alloy is vastly superior to all others; existing prosthetic alloys have all been used in successful and, indeed, unsuccessful implant designs. Alloy selection is only one determinant of performance of the implanted device. *See* METAL, MECHANICAL PROPERTIES OF; PROSTHESIS. John Brunski

Superconducting alloys. Superconductors are materials that have zero resistance to the flow of electric current at low temperatures. There are more than 6000 elements, alloys, and compounds that are known superconductors. This remarkable property of zero resistance offers unprecedented technological advances such as the generation of intense magnetic fields. Realization of these new technologies requires development of specifically designed superconducting alloys and composite conductors. An alloy of niobium and titanium (NbTi) has a great number of applications in superconductivity; it becomes superconducting at 9.5 K (critical superconducting temperature, T_c). This alloy is preferred because of its ductility and its ability to carry large

amounts of current at high magnetic fields, represented by $J_c(H)$ [where J_c is the critical current and H is a given magnetic field], and still retain its superconducting properties. Brittle compounds with intrinsically superior superconducting properties are also being developed for magnet applications. The most promising of these are compounds of niobium and tin (Nb_3Sn), vanadium and gallium (V_3Ga), niobium and germanium (Nb_3Ge), and niobium and aluminum (Nb_3Al) which have higher T_c (15 to 23 K) and higher $J_c(H)$ than NbTi.

Superconducting materials possess other unique properties such as magnetic flux quantization and magnetic-field-modulated supercurrent flow between two slightly separated superconductors. These properties form the basis for electronic applications of superconductivity such as high-speed computers or ultrasensitive magnetometers. Development of these applications began using lead or niobium (T_c of 7 K and 9 K) in bulk form, but the emphasis then was transferred to materials deposited in thin-film form. Lead-indium (PbIn) and lead-gold (PbAu) alloys are more desirable than pure lead films, as they are more stable. Improved vacuum deposition systems eventually led to the use of pure niobium films as they, in turn, were more stable than lead alloy films. Advances in thin-film synthesis techniques led to the use of the refractory compound niobium nitride (NbN) in electronic applications. This compound is very stable and possesses a higher T_c (15 K) than either lead or niobium. *See* REFRACTORY.

Novel high-temperature superconducting materials may have revolutionary impact on superconductivity and its applications. These materials are ceramic, copper oxide–based materials that contain at least four and as many as six elements. Typical examples are yttrium-barium-copper-oxygen (T_c 93 K); bismuth-strontium-calcium-copper-oxygen (T_c 110 K); and thallium-barium-calcium-copper-oxygen (T_c 125 K). These materials become superconducting at such high temperatures that refrigeration is simpler, more dependable, and less expensive. Much research and development has been done to improve the technologically important properties such as $J_c(H)$, chemical and mechanical stability, and device-compatible processing procedures. It is anticipated that the new compounds will have a significant impact in the growing field of superconductivity. *See* CERAMICS; SUPERCONDUCTIVITY. D. Gubser

Bibliography. G. S. Brady, H. R. Clauser, and J. A. Vaccari, *Materials Handbook*, 15th ed., 2002; W. D. Callister, *Materials Science and Engineering: An Introduction*, 6th ed., 2002; J. Frick (ed.), *Woldman's Engineering Alloys*, 9th ed., 2001; J. A. Helsen and H. J. Breme, *Metals as Biomaterials*, 1998; D. P. Henkel and A. Pense, *Structures and Properties of Engineering Materials*, 5th ed., 2001; J. Mordike and P. Haasen, *Physical Metallurgy*, 3d ed., 1996; C. T. Lynch (ed.), *Handbook of Materials Sciences*, vol. 2, 1975; L. H. Van Vlack, *Elements of Material Science and Engineering*, 6th ed., 1989.

Alloy structures

Metals in actual commercial use are almost exclusively alloys, and not pure metals, since it is possible for the designer to realize an extensive variety of physical properties in the product by varying the metallic composition of the alloy. As a case in point, commercially pure or cast iron is very brittle because of the small amount of carbon impurity always present, while the steels are much more ductile, with greater strength and better corrosion properties. In general, the highly purified single crystal of a metal is very soft and malleable, with high electrical conductivity, while the alloy is usually harder and may have a much lower conductivity. The conductivity will vary with the degree of order of the alloy, and the hardness will vary with the particular heat treatment used. For commercial applications of alloys and other information *see* ALLOY

The basic knowledge of structural properties of alloys is still in large part empirical, and indeed, it will probably never be possible to derive formulas which will predict which metals to mix in a certain proportion and with a certain heat treatment to yield a specified property or set of properties. However, a set of rules exists which describes the qualitative behavior of certain groups of alloys. These rules are statements concerning the relative sizes of constituent atoms for alloy formation, and concerning what kinds of phases to expect in terms of the valence of the constituent atoms. The rules were discovered in a strictly empirical way, and for the most part, the present theoretical understanding of alloys consists of rudimentary theories which describe how the rules arise from the basic principles of physics.

Prior to a discussion of the rules proposed by W. Hume-Rothery concerning the binary substitutional alloys, some information concerning alloy composition diagrams will be presented. Robb M. Thomson

Phase Diagrams

Alloys are classified as binary alloys, composed of two components; as ternary alloys, composed of three components; or as multicomponent alloys. Most commercial alloys are multicomponent. The composition of an alloy is described by giving the percentage (either by weight or by atoms) of each element in it.

A phase diagram is a graphical description of the kinds and amounts of the phases that can be expected in an alloy as a function of its composition, temperature, and pressure when it has reached thermodynamic equilibrium. The phases may be liquid, vapor, or solid with various ordered and disordered crystal structures. A phase diagram does not provide information about how rapidly equilibrium can be reached; when a phase diagram is determined experimentally, it is necessary not only to find out what phases are present but also to assure that the alloy is in the stable equilibrium state. Equilibrium is reached when the Gibbs free energy of the system has reached its lowest possible value. The thermodynamic principle of minimum Gibbs energy imposes

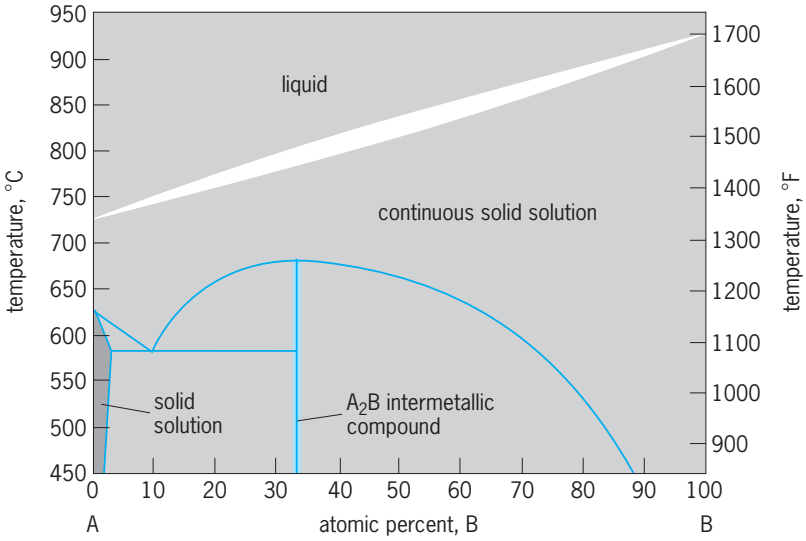

Fig. 1. Phase diagram at fixed pressure for a hypothetical alloy composed of elements A and B.

restrictions on the possible forms a phase diagram can take, and even a complicated diagram with many phases is composed of only a few geometrical features. *See* FREE ENERGY; PHASE EQUILIBRIUM; THERMODYNAMIC PRINCIPLES.

Figure 1 shows a phase diagram for a hypothetical alloy system composed of elements A and B. As with most phase diagrams, it is drawn at fixed pressure. Hence it is a two-dimensional figure with composition and temperature as the axes. There are four equilibrium phases in the A-B system: the liquid, two solid phases with crystal structures of the pure components, and an intermetallic compound. The intermetallic compound is an orderly arrangement of two A atoms for every B atom on a crystal lattice. The other solid phases are disordered solutions of A and B atoms on different crystal lattices, one with a complete range of solubility and one with only a small solubility of B in A. The phase diagram for this system maps the melting and solidification temperatures, and the solubilities of the elements in each other and in the intermetallic compound. In the shaded regions the labeled phases are the only stable states of the alloys: in these regions the alloys are homogeneous liquid, solid solution, or A_2B compound. In the other regions of the diagram, two phases of different composition can coexist in two-phase, or heterogeneous, equilibrium. Two key features of the A-B diagram are the two-phase equilibrium and the three-phase reaction, and these are illustrated in more detail in **Figs. 2** and **3**. *See* SOLID SOLUTION; SOLID-STATE CHEMISTRY; THERMODYNAMIC PROCESSES.

Two-phase equilibrium. Figure 2 illustrates the concept of two-phase equilibrium using the specific example of the melting behavior of the A-B system. The pure elements A and B melt completely; each has a single melting-point temperature. The alloys, however, begin melting at one temperature but do not finish melting until they reach a higher temperature. In between, the alloy is slushy, and liquid and solid coexist in equilibrium with each other. Upon cooling the liquid alloy, the first small amount of solid to precipitate as the alloy solidifies is richer in element B than the original liquid. This phenomenon is called segregation of the two constituents. The lens region of the diagram is the region of the two-phase equilibrium between liquid and solid. The upper line where the liquid begins to solidify is the liquidus, and the lower line where the solid first begins to melt is the solidus.

For an alloy of overall composition x, the compositions of liquid and solid, x_l and x_s, in equilibrium at a particular temperature are found by reading off the compositions of the liquidus and solidus at that temperature. From the overall composition x, the amounts of liquid and solid, f_l and f_s, can also be calculated. For a two-phase alloy, the overall composition is obtained by averaging the compositions of liquid and solid, weighted by the amounts present. Since the overall composition is not changed during segregation, Eqs. (1) are valid. These equations are

$$\begin{aligned} x &= f_l x_l + f_s x_s \\ f_l + f_s &= 1 \end{aligned} \tag{1}$$

a statement of the lever rule, one of the important rules for reading phase diagrams.

Liquid and solid equilibrium for fixed pressure over a range of temperatures does not occur in pure substances but only in alloys. The phase diagram can be used to predict how molten A-B alloys will solidify in different conditions. If the temperature of a liquid alloy is brought below the liquidus, a small amount of solid will form with a composition different from that of the liquid. How the solidification proceeds now depends on the rate of cooling. If the alloy is cooled very slowly through the two-phase region, equilibrium can be reached at each temperature and the compositions of the liquid and solid can continuously adjust to the values on the phase diagram. If the cooling rate is too high for equilibrium to be reached at each temperature, which is typically the

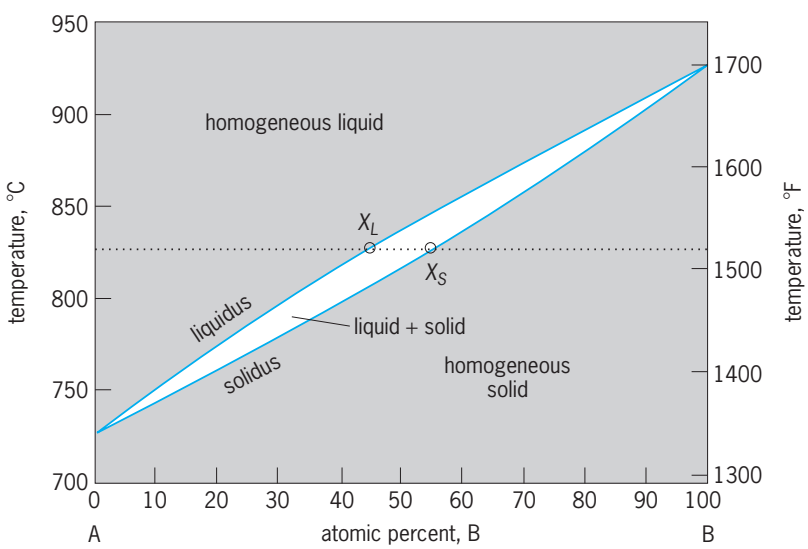

Fig. 2. Detail of the A-B system showing the liquidus and solidus.

case, the phase diagram can still predict what can happen. The first solid to form is rich in element B compared to the liquid. The solid does not have time to readjust its composition by diffusion of element B back out of the solid, so the liquid becomes more and more enriched in element A. The solid ends up inhomogeneous in composition, and its microstructure depends on the phase diagram and the cooling rate.

The liquid-solid phase equilibria shown schematically in Fig. 2 make up the complete phase diagram for many systems, for example, silver-gold. Silver and gold are completely miscible in both the liquid and solid states and form no intermetallic compounds because of the close similarity of their chemical properties and atom sizes.

Three-phase reaction. The concept of the three-phase reaction is illustrated in Fig. 3 by the solidification behavior of the silver-copper alloy system. Silver and copper are not as similar to each other as are silver and gold. In the solid state, silver and copper do not mix in all proportions, and it is possible to have inhomogeneous alloys with two different solid phases in equilibrium, as well as solid phases in equilibrium with the liquid. Molten alloys can solidify to form copper-rich or silver-rich solids. The three two-phases fields intersect at a horizontal line, the eutectic horizontal. If a silver-rich liquid is slowly cooled, it separates into a silver-rich solid and a copper-rich liquid. However, below the eutectic temperature, the liquid is no longer a stable phase of the system. The silver-rich solid and the copper-rich liquid react to the eutectic temperature to form a new phase, solid copper. The reaction continues at the eutectic temperature until all the liquid is consumed, leaving a two-phase mixture of silver and copper. *See* EUTECTICS.

Minimization of Gibbs energy. The concepts of the two-phase equilibrium and the three-phase reaction describe all the ingredients of the phase diagrams of binary alloys. No more than three phases can ever coexist in equilibrium in a binary system at constant pressure. No matter how many complicated phases appear in the phase diagram, the regions are either single-phase or two-phase; horizontal lines represent the three-phase temperatures.

This property of binary phase diagrams stems from the thermodynamic principle of the minimization of the Gibbs energy. In alloy systems, each phase has a Gibbs energy function of composition and temperature at fixed pressure. For a pure substance, composition does not enter the picture, but there are several Gibbs energy functions of temperature for the different phases. Whichever function is lower at the temperature of interest is the stable phase. In a binary alloy system, the composition variable provides a new way for the system to lower the Gibbs free energy. **Figure 4** shows the Gibbs energy functions for liquid and solid for a hypothetical system. At the composition x, the solid phase has a lower Gibbs energy than the liquid and so is more stable. However, if the alloy segregates into a liquid of composition x_l and a solid of composition x_s, the Gibbs energy is lower yet than for either a uniform solid or a uniform liquid. The only constraint on the way the alloy can split up is that the slopes of the Gibbs energies of the two phases in equilibrium must be equal. The slope of the Gibbs energy regulates the composition exchange taking place between the two phases, and at equilibrium the compositions must be constant. The compositions can be determined by drawing a common tangent to the two Gibbs energy functions; the points of tangency determine x_l and x_s. This method of finding a two-phase equilibrium on the phase diagram is called the common tangent construction.

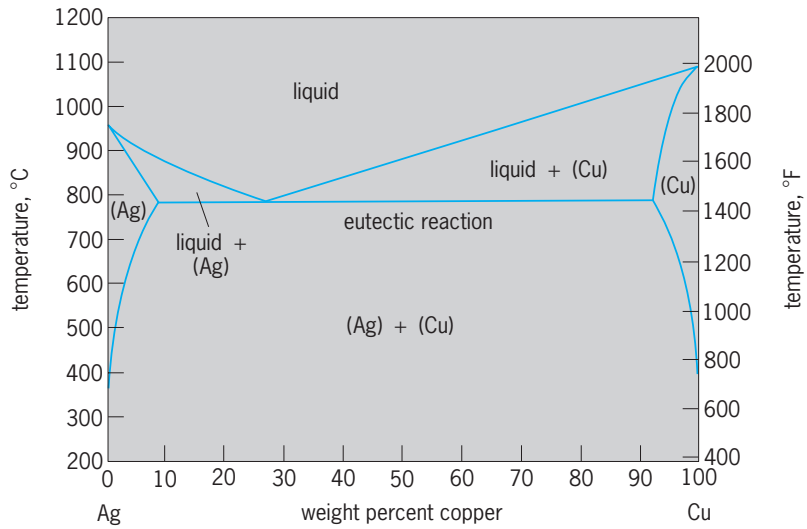

Fig. 3. Phase diagram of the silver-copper system, which exhibits the eutectic reaction. Silver-rich and copper-rich phases are denoted by (Ag) and (Cu), respectively.

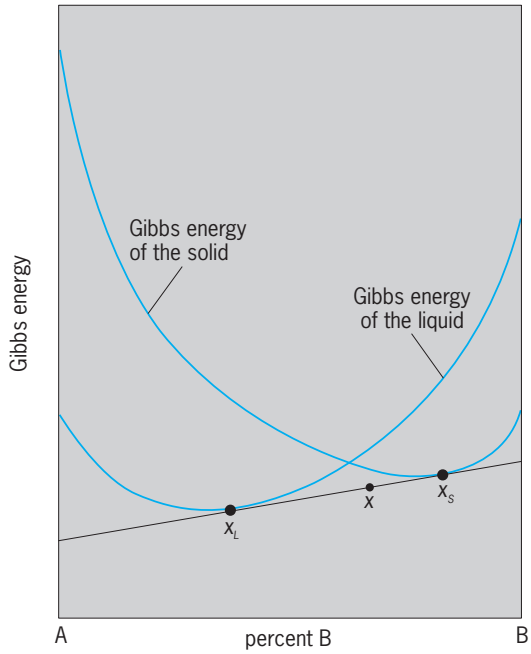

Fig. 4. Gibbs energies at fixed pressure and temperature as a function of concentration, and common tangent construction, for a hypothetical system.

The phase diagram is mapped out by determining the positions of common tangents to the Gibbs energy functions at various temperatures. If two Gibbs energy functions intersect, it is always possible to draw a common tangent. This is why the two-phase fields can extend over a range of temperature. To have three phases in equilibrium, however, three Gibbs energy functions must line up along a common tangent. This can only happen at one temperature, above and below which the Gibbs energies will move back out of line. Four phases cannot all be lined up. This is a statement of the Gibbs phase rule for binary systems. *See* PHASE RULE.

Ternary and multicomponent systems. In a ternary system such as silver-gold-copper, the composition is represented by two variables, for example, the percentage of silver and the percentage of gold. The temperature is also still a variable. The phase diagram must be drawn in three dimensions. The Gibbs energy functions are surfaces, and the common tangent is a tangent plane rather than a line. It is possible to find a plane tangent to three surfaces over a range of temperatures, so that for ternary systems the maximum number of phases in equilibrium is four, and increases by one for each element added. The same principles of construction of the three-dimensional phase diagram work for ternary systems, with one to four phases in equilibrium in each region, but the representation of the diagrams can become complicated.

Nonequilibrium processes. The state of a real alloy depends not only on the equilibrium shown on the phase diagram, but also on how far away from equilibrium the system has been brought by the temperature treatments it has been subjected to and on how close to equilibrium the system can approach in the time given it. The system always behaves in accordance with the thermodynamics and Gibbs energy functions that underlie the equilibrium phase diagram. If an equilibrium phase is not able to nucleate and grow, then a metastable equilibrium might be achieved between two other competing phases. A phase can sometimes be taken so far from its equilibrium range that it becomes very unstable and can rapidly transform to another structure by one of various nonequilibrium transformation processes. The materials scientist uses thermodynamic ideas to read from the equilibrium phase diagrams the possibilities of these processes.

In short, the phase diagram is a road map that guides the materials scientist through problems encountered in alloy design, design of processing techniques, failure analysis, and behavior of alloys in performance. Moreover, phase diagrams are not used only for metallurgical systems; the behavior of ceramics and minerals, and even mixtures of polymers, is also described by phase diagrams.

Joanne L. Murray; Kirit J. Bhansali

Size and Valence Factors

Comparison studies of a large number of alloys have led to the rules for alloy formation formulated by W. Hume-Rothery and others. These rules have recognized two general types of criteria regarding alloy formation. One relates to the relative size of the atomic constituents, the other to the valence, or the electronic behavior, of the two metals. These criteria have generally been treated independently of one another, although in reality some interaction can be expected. *See* VALENCE.

Size factor. The empirical rules for alloy formation state that, for substitutional alloys, the size of the constituents must be approximately the same, whereas one of the constituents of an interstitial alloy must be small compared to the other. Simply stated, the size factor recognizes the importance of choosing two metals which can fit together in a lattice structure. In order to define the concept of size, one must refer to the general ideas of the band theory of solids. According to this theory, the valence electrons of the metal atoms are detached from the immediate vicinity of the atom and contribute to the conduction band. In the conduction band the electrons have many of the characteristics of free particles similar to the atoms of a gas. The conduction electrons are spread over the metal, filling the interstices between the remaining ion cores of the metal atoms. The electrons of the inner shells remain tightly bound to the individual ions, however. *See* BAND THEORY OF SOLIDS; FREE-ELECTRON THEORY OF METALS.

It is assumed in any discussion of atomic size that the free electrons cause a rather weak attractive force drawing the metal together, while the ion cores strongly repel one another when one core overlaps a neighbor. The strength and form of the attractive force obviously depends on the crystal type; however, the ion cores are more nearly independent of their surroundings. To the extent that the size of the atom in the metal is given by the radius of the ion core, the concept of size is well defined. However, in those cases where some of the repulsive force is contributed by the structure-dependent free electrons, the size of an atom in an alloy will vary from one alloy to another.

To the extent that the size of an ion core can be rigorously defined, the alloy will contain large misfit stresses unless the ion-core sizes of the two metal atoms have a favorable ratio.

In the primary or terminal solid solution parts of the composition diagram, there are two simple possibilities. The atoms of the minority category (the solute atoms) are able to supplant the solvent atoms on the solvent atom lattice, or they may fit into the open spaces of the solvent lattice. The first case is known as a substitutional alloy, while the second is called an interstitial alloy.

Interstitial alloys. Examples of interstitial alloys are some of the alloys of the transition elements and, most familiarly, those of iron. The iron lattice is face-centered cubic at medium temperatures. If the face-centered lattice is considered to be made up of hard spheres of radius a_0, the largest atom which can be fitted into an interstitial position (the cube center) has a radius $0.59a_0$. There are only four neutral atoms which have smaller radii than this value for the transition-metal group. They are hydrogen, carbon,

nitrogen, and boron. Carbon is actually an exception to the rule for iron, as it has a radius of $0.63a_0$ for the iron lattice. Actually, these four elements are not metals in their normal state, even though they do form metallic alloys with the transition metals. *See* CRYSTAL DEFECTS; CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Substitutional alloys. In the substitutional primary solutions the solute atom takes the place of one of the solvent atoms. In this case the size of the solvent and solute must be nearly the same. It is possible to calculate in a crude way the maximum permissible difference in size. The size difference is reflected in a lattice distortion, which is a contribution to the internal energy of the system.

In general, the formation of a homogeneous alloy is possible if the change in the free energy of the alloy relative to the two separate elements is negative, as in Eq. (2). Here F is the Helmholtz free energy, U is

$$\Delta F = \Delta U - T\Delta S \tag{2}$$

the internal energy, T is the absolute temperature, and S is the entropy. The distortion energy increases ΔU. On the other hand, there is an entropy of mixing when the solute atoms are distributed in a random fashion on the solvent lattice. Thus, if the distortion energy is not too large to overbalance the entropy change, the alloy can form. An estimate of the distortion energy as a function of size suggests that the size of the solute atom must be within about 15% of the size of the solvent atom. This number agrees with the empirically derived Hume-Rothery rule. *See* ENTROPY.

The substitutional alloys are the commonest type of alloy structure and have received the most study. Cu-Au is a good example of the substitutional alloy. The atomic sizes are very close, and the electronic structures of Cu and Au are very similar. As a result, this system forms a single primary solid solution system from one end of the composition diagram to the other. The only complexities are due to ordering phenomena, which will be discussed later.

Valence factor. In addition to the effects of the size of the ions on the formation of alloys, the electronic structure of the atoms involved also plays a role. It is the electronic configuration of the atoms which determines why mixtures of some elements form metallic alloys, whereas some form insulating compounds. Qualitatively, it is found that alloys are formed from the atoms of the middle of the electrochemical series of the elements. The reason is that there is always a tendency for the ions of the metallic state to polarize with respect to their neighbors and form an ionic solid instead of the metallic one. If the tendency is strong, no metallic alloy state can be formed; hence, only elements from the middle of the series, where there is little change in the ionization potential from one atom to another, can form successful alloys. *See* ELECTROCHEMICAL SERIES.

Stoichiometry. Closely related to the polarization effect is the reason why the alloys are nonstoichiometric (that is, why they form mixtures which do not consist of small-number ratios of one element to the other) and form homogeneous phases over wide ranges of composition. For example, copper and gold mix homogeneously for any composition. This result is in striking contrast to the behavior of the ionic compounds such as sodium chloride, NaCl. *See* STOICHIOMETRY.

In NaCl an excess Na ion is bound in the crystal with an energy less by about 1 eV than a normal ion, and this energy discrepancy is large enough so that, when the crystal becomes far from stoichiometric, the homogeneous compound is no longer formed. The excess constituent forms a separate phase, either as Na metal crystals embedded in the matrix or as bubbles of Cl_2. The energy discrepancy in the case of NaCl is primarily due to the excess of charge of one sign or the other when the wrong ion is on a lattice site. Since the lattice is completely ionized, the excess charge due to the excess ion amounts to the complete ionic charge. *See* IONIC CRYSTALS.

The situation for a metal is completely different. The charge of the ionic core of a metal is balanced by the charge of the free electrons which each atom contributes to the conduction band. Thus, in a perfect metal, the volume occupied in the lattice by each atom is neutral. If a single gold atom is placed substitutionally in a copper lattice, there will be a tendency for the charge in the gold lattice volume to become unbalanced because the copper and gold atoms have different ionization energies. Effectively, the copper and gold atoms in the metal have different affinities for the free electrons of the valence band; hence there will be a small charge imbalance induced around the gold lattice site. However, in the case of Au-Cu, the chemical valence is equal, and the amount of polarization inducted is a small fraction of a unit charge.

Screening length. A further effect is the screening of the polarization around the gold atom by the conduction electrons themselves. In this case, when the region around the gold is polarized, a voltage difference is also generated. Hence, the electrons of the metal tend to rush into the affected region to even out the discrepancy. However, it is not possible to redress the balance completely, and a region of the order of the size of the gold atomic volume itself remains polarized. The size of this region is indicated by the screening length of the electrons of the metal, which are here listed for several metals in units of lattice spacing a_0: Cu, 1.1; Al, 1.4; Tl, 1.0; Fe, 3.2; and Ni, 4.8.

The effect of replacing a copper atom with a gold atom in Cu-Au is thus seen to be very much less than the effect of replacing a Cl ion with an Na ion in NaCl. For this reason the copper lattice can accept an unlimited number of exchanges without breaking up its structure, and the alloy forms over the entire composition range. Cu-Au is a rather special example to pick, since the ionization potentials of Cu and Au are very close, and the size factor is very favorable. The discussion, however, does illustrate the reason why these metals do not form compounds in the chemical sense.

Much the same reasoning would be applicable for showing why covalent crystals should also form stoichiometric compounds, where dangling covalent bonds would have much the same effect on the energy of the crystal as does an excess charge in an ionic lattice. *See* CHEMICAL BONDING.

Electron Compounds

Even though no metal alloy forms a compound in the chemical sense, there is a characteristic behavior when the alloying constituents have differing valences, and these alloys go under the name of electron compounds. In the usual alloy the intermediate region of the phase diagram is covered by several intermediate phases, with more than the one such phase illustrated in Fig. 1. The term electron compound refers to the fact that there is considerable regularity in the types of intermediate phases which are observed.

Structure and electron density. It appears that in numerous cases there is a tendency for a particular type of crystal structure to correspond to a particular electron density in the conduction band. Thus, when zinc is alloyed to copper, near the 50% composition there is a phase change from the face-centered-cubic (fcc) lattice of copper to a body-centered-cubic (bcc) lattice. If it is assumed that all the valence electrons of all the atoms are contributed to the conduction band, there will be an average of 1.5 electrons per atom of the alloy in this band, since the valence of copper is 1 and the valence of zinc is 2. For the alloys Cu_3Al and Cu_5Sn, the crystal structure is also bcc for the same average free electron concentration. T. B. Massalsky has listed (see **table**) the intermediate phases which center about the characteristic electron concentrations shown. Note that there are three different compounds centered about the electron concentration 1.5 in the table. There are also two other well-developed electron compounds, one at about concentration 1.6 and one at about concentration 1.75. Also, there are two places where the hexagonal close-packed (hcp) structures occur, one at 1.5 and one at 1.75. For example, in the case of Cu-Si, the hcp structure is interrupted at 1.6 by a change to the phase called the γ-brass structure, which is a very complex crystal type with 52 atoms per unit cell.

Brillouin zones. The striking correlations of the various electron compounds in the empirical tabulation of the table has been interpreted on the basis of the free-electron theory of metals. H. Jones has used the free-electron theory of metals modified by perturbation theory to derive an expression for the difference in energy of the electrons of the fcc and bcc crystals. His result is that as the Brillouin zone is filled, at first the energy is very nearly equal for the two crystals. The energy of the fcc crystal then drops below that of the bcc crystal at an electron concentration of about 1 electron per atom; but at a concentration close to 1.5, the situation reverses and the bcc crystal becomes more stable. Similar discussions have been given for the hexagonal lattices; these show that certain distortions of the symmetry of the perfect hexagonal lattice which occur are due to electronic effects related to filling the Brillouin zone. *See* BRILLOUIN ZONE.

Binary systems with intermediate phases at concentrations of 1.5, 1.6, and 1.75 electrons per atom

1.5			1.6	1.75
Body centered cubic	Hexagonal close-packed	β-Manganese	$\gamma\gamma$ Brass	Hexagonal close-packed
Cu-Be	Cu-Ag	Cu-Si	Cu-Zn	Cu-Zn
Cu-Zn	Cu-Ge	Ag-Hg	Cu-Cd	Cu-Cd
Cu-Al	Cu-Si	Ag-Al	Cu-Hg	Cu-Sn
Cu-Ga	Ag-Al	Au-Al	Cu-Al	Cu-Si
Cu-In	Ag-In	Co-Zn	Cu-Ga	Ag-Zn
Cu-Si	Ag-Sn		Cu-In	Ag-Cd
Cu-Sn	Ag-Sb		Cu-Si	Ag-Al
Ag-Mg	Ag-Cd		Cu-Sn	Au-Zn
Ag-Zn	Au-Sn		Ag-Zn	Au-Cd
Ag-Cd	Au-In		Ag-Cd	Au-Sn
Ag-Al			Ag-Hg	Au-Al
Ag-In			Ag-In	
Au-Mg			Au-Zn	
Au-Zn			Au-Cd	
Au-Cd			Au-In	
Fe-Al			Mn-Zn	
Co-Al			Fe-Zn	
Ni-Al			Co-Zn	
Ni-In			Ni-Be	
Pd-In			Ni-Zn	
			Ni-Cd	
			Rh-Zn	
			Pd-Zn	
			Pt-Be	
			Pt-Zn	
			Na-Pb	

However, in spite of the seeming success of the theory in confirming the electron compound concept in terms of the interaction of the electrons with the boundaries of the Brillouin zone, the Jones treatment must be considered a very limited theory. First of all, Jones assumed, when he adopted the free-electron picture, that the role of the ions in the alloy is a minor one. On the other hand, when a zinc atom is placed in a copper lattice, it is necessary that the vicinity of the zinc atom be highly polarized, according to the discussion of the preceding section. In a 50% alloy of Cu-Zn, the atomic volume of a zinc ion will have only 1.5 electrons to cancel the charge of a double ionized core. Thus, the lattice cell of the zinc atom has a net charge of $+^1/_2$. Hence, the valence electrons from zinc atoms cannot be freely contributed to the free-electron cloud of the crystal. There must be a considerable clumping of this charge around the various different ions of the alloy in such a way that the excess charge of the ions is screened. In addition, there is some evidence from the experiments on nuclear magnetic resonance in dilute alloys that the electron density in the conduction band does not change as the divalent ion is added, in direct contradiction to the simple free-electron picture.

In view of these considerations, it is surprising that such simple results as those obtained by Jones are correct. In later work J. Friedel has been able to show that, although the description of the electrons in the conduction band is far from the free-electron picture

of Jones, the so-called rigid band model is valid for dilute alloys and leads to results similar to those of Jones. The theory of electron compounds is not yet firmly established.

Order-Disorder in Alloys

In the preceding discussion on the formation of alloys, it was tacitly assumed that the atoms of the two constituents were randomly distributed on the lattice of the alloy. In very dilute primary solutions the random distribution holds because the free energy of the system in the completely random state is lower than that of a more symmetric state. To be more specific, in a dilute solution of metal B in metal A, if there is no gain of internal energy by ordering B (placing the atoms of B in a regular arrangement on the lattice sides of A), the free energy, $U - TS$, is a minimum when B is completely randomized.

On the other hand, if an appreciable percentage of the lattice sites are occupied by B and if there is a difference in the interatomic forces between A-B nearest-neighbor atom pairs and those between A-A or B-B pairs, the situation is different. The internal energy of the alloy is then given by Eq. (3), where

$$U = N_{AB}E_{AB} + N_{AA}E_{AA} + N_{BB}E_{BB} \tag{3}$$

N_{AB} is the number of A-B nearest-neighbor pairs in the alloy, E_{AB} is the energy of interaction between an A-B pair, and so on. It is assumed that the alloy atoms interact appreciably only when they are nearest neighbors. If the interaction between A-B pairs is greater than that between B-B or A-A pairs, then the state of lowest energy is that in which the number of like pairs is minimized. This state is also a completely ordered state of the crystal.

For any given temperature the equilibrium state of the crystal is that for which the free energy is minimized, and a balance is struck between the contradictory tendencies of the U and TS terms of the free energy. At low temperatures the U term can be expected to predominate, and ordering will usually occur, while at higher temperatures the entropy term will predominate, and a transition to a disordered state is usual.

Long- and short-range order. The ordering of a crystal is described in terms of two different order parameters. One is called local or short-range order, and the other is called long-range order. For the initial discussion of the long- and short-range order parameters, a 50% A-B alloy will be considered. It will also be supposed that the lattice structure is bcc and that the unlike pair interaction is the stronger one. The advantage of the bcc lattice for the discussion is that, in this lattice, all the neighbors of any given atom can be made unlike atoms.

In the completely ordered state every atom of the crystal is surrounded by unlike nearest neighbors, and the crystal becomes disordered when some of the A atoms exchange with B atoms. The long-range order parameter Σ is a measure of the number of such interchanges and expresses the number of sites

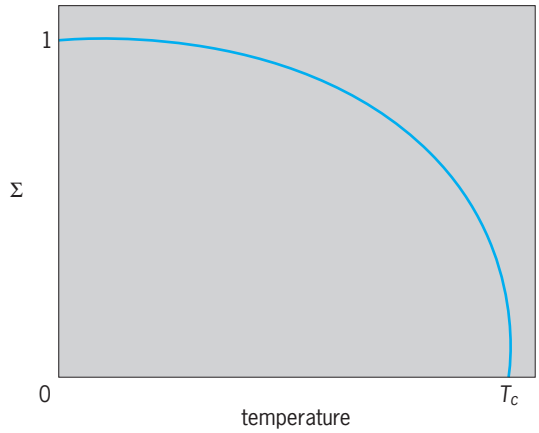

Fig. 5. Long-range order Σ of a crystal as a function of temperature. The order parameter has a rather precipitous drop at the transition temperature and is zero above that point.

occupied by the wrong atom, as in Eq. (4). Here n_a

$$\Sigma = \frac{2n_a}{N} - 1 \tag{4}$$

is the number of A sites occupied by A atoms, and N is the total number of A lattice sites. The quantity Σ varies from -1 to $+1$, and the state of complete disorder corresponds to $\Sigma = 0$.

The short-range order expresses the fact that the neighbors surrounding any given A atom will have a tendency to be all B atoms. If this tendency is averaged over the lattice, the short-range order σ is defined in Eq. (5). Here σ is the local-order parameter, and q is the number of A-B pairs in the crystal divided by the total number of atom pairs in the crystal. The parameter σ has the range -1 to $+1$, with complete disorder at $\sigma = 0$.

$$\sigma = 2(q - {}^1/_2) \tag{5}$$

Variation with temperature. A good qualitative notion of the behavior of the order of a crystal is obtained by observing the variation of the order parameter of a crystal as a function of temperature. Theory predicts the curves shown in **Figs. 5** and **6**. The long-range order decreases from a state of complete order at absolute zero temperature. At a critical transition

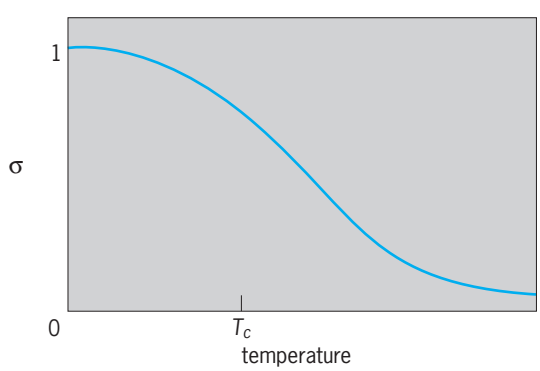

Fig. 6. Short-range order σ as a function of temperature. It varies more slowly than long-range order. Considerable order still exists above the transition temperature.

temperature which depends on the strength of the A-B bonds relative to the A-A bonds, the long-range order drops to zero. The local order also starts at absolute zero with a maximum value, but decreases more slowly than the long-range order does and remains finite for all temperatures.

The discussion up till now has been based on the model of a 50% alloy in a simple lattice. The theory for a material in which the composition is allowed to vary over the entire composition diagram is exceedingly complicated. The complication is due to the difficulty of specifying the types of order configurations which are possible, and computing their entropy. However, the general results are still comparable with the simple case. Long- and short-range order parameters can still be defined, with a transition temperature at the point where the long-range order disappears. The short-range order again persists to a considerable degree even above the transition temperature.

Transition temperature. The existence of the transition temperature amounts to a change of phase, because there is an attendant singularity in the specific-heat curve of the material. The phase change may be of the first or second order, depending on the type of singularity present in the long-range order at the transition temperature. The curve of Fig. 5 corresponds to a second-order transition; however, if there is a discontinuous drop of the order to zero at the transition point, the transition is first order. A general rule seems to be that for lattices in which the completely ordered state has only neighbors of A-B bonds, the transition is second order. The bcc lattice is such a case, as has already been explained. However, in close-packed lattices such as the fcc lattice, it is not possible for all the bonds of a 50% alloy to be of the sort A-B, and in such cases, it is found experimentally that the transition is first order. *See* PHASE TRANSITIONS.

Detection of order. The presence of order in a crystal and the transition from the ordered to the disordered state are detected by a variety of techniques. The resistance of the alloy at low temperatures varies with the order of the crystal because the electron waves of the crystal are sensitive to irregularities in the crystal, so that as disorder increases, the resistivity increases also. X-ray and neutron scattering of the lattice are also functions of the degree of order for the same reason. The specific heat of the crystal varies with the order, since as the crystal loses its order, energy must be supplied to the lattice to form the A-A and B-B bonds which have higher energy. At the transition temperature the specific heat rises to a sharp peak, which is easily detected. *See* NEUTRON DIFFRACTION; SPECIFIC HEAT OF SOLIDS; X-RAY DIFFRACTION.

Superlattices. The x-ray and neutron scattering of a completely ordered crystal have an interesting peculiarity. In the Bragg scattering processes new lines appear, called superlattice lines. The name is derived from the fact that the lines correspond to the appearance of a secondary crystalline structure in the lattice. Their explanation is clear in the light of the discussion of the ordered lattice. For a 50% A-B lattice in the bcc form, the A atoms form the corners of the cube, while the B atoms are at the center positions. The crystal is then composed of two interpenetrating simple cubic crystals, and is called a superlattice. The x-ray lines for two simple cubic structures appear instead of those for a bcc crystal. *See* INTERMETALLIC COMPOUNDS.

Robb M. Thomson

Bibliography. W. Hume-Rothery and R. E. Smallman, *The Structure of Metals and Alloys*, 1988; A. G. Khachaturyan, *Theory of Structural Transformations in Solids*, 1983; W. G. Moffatt, *Handbook of Binary Phase Diagrams*, 5 vols., 1981; W. F. Smith, *Structure and Properties of Engineering Alloys*, 2d ed., 1992.

Allspice

The dried, unripe fruits of a small, tropical, evergreen tree, *Pimenta officinalis*, of the myrtle family (Myrtaceae). This species (see **illus.**) is a native of the

Allspice. (*a*) Branch with fruit. (*b*) Flowers.

West Indies and parts of Central and South America. The spice, alone or in mixtures, is much used in sausages, pickles, sauces, and soups. The extracted oil is used for flavoring and in perfumery. Allspice is so named because its flavor resembles that of a combination of cloves, cinnamon, and nutmeg. *See* MYRTALES; SPICE AND FLAVORING.

Perry D. Strausbaugh; Earl L. Core

Almanac

A book that contains astronomical or meteorological data arranged according to days, weeks, and months of a given year and may also include diverse information of a nonastronomical character. This article is restricted to astronomical and navigational almanacs.

Development. The earliest known almanac material was computed by the Egyptians in A.D. 467, and astronomical ephemerides appeared irregularly thereafter. The first regular almanac appears to have been produced in Germany from 1475 to 1531. The first almanac to be published regularly by a national government was the *Connaissance des Temps* (1679) by France. In 1767 the British began publishing the *Nautical Almanac* to improve the art of navigation. The *Berliner Astronomisches Jahrbuch* was first published in Germany for 1776, and the *Efemerides Astronomicas* was published in Spain in 1791. In the United States *The American Ephemeris and Nautical Almanac* has been published since 1855. Beginning with the issue for 1981, the two series of publications titled *The Astronomical Ephemeris,* which had previously replaced *The Nautical Almanac and Astronomical Ephemeris,* and *The American Ephemeris and Nautical Almanac* were continued with the title of *The Astronomical Almanac.*

Over the years cooperation has developed between the organizations responsible for the preparation and publication of the almanacs. Under the auspices of the International Astronomical Union (IAU), agreements are reached concerning the bases and constants to be used in the publications and the exchange of data in different forms.

Astronomical Almanac. *The Astronomical Almanac* contains ephemerides, which are tabulations, at regular time intervals, of the orbital positions and rotational orientation of the Sun, Moon, planets, satellites, and some minor planets. It also contains mean places of stars, quasars, pulsars, galaxies, and radio sources, and the times for astronomical phenomena such as eclipses, conjunctions, occultations, sunrise, sunset, twilight, moonrise, and moonset. This volume contains the fundamental astronomical data needed by astronomers, geodesists, navigators, surveyors, and space scientists. The theory and methods on which *The Astronomical Almanac* is based are provided in the *Explanatory Supplement to the Astronomical Almanac. See* ASTRONOMICAL COORDINATE SYSTEMS; EPHEMERIS.

Navigational almanacs. While *The Astronomical Almanac* is basically designed for the determination of positions of astronomical objects as observed from the Earth, *The Nautical Almanac* and *The Air Almanac* are designed to determine the navigator's position from the tabulated position of the celestial object. The ground point of the celestial object is used with the North and South poles and the observer's assumed position to establish a spherical triangle known as the navigational triangle. By means of spherical trigonometry or navigational tables, the navigator determines a computed altitude and the direction of the ground point (azimuth) of the celestial object with respect to his or her position. The computed altitude and azimuth enable the navigator to plot a line of position that goes through or near this assumed position. The computed altitude is compared to the observed altitude of the celestial object. From the combination of two or more observations the navigator is able to determine position. *See* CELESTIAL NAVIGATION.

The Nautical Almanac contains hourly values of the Greenwich hour angle and declination of the Sun, Moon, Venus, Mars, Jupiter, and Saturn and the sidereal hour angle and declination of 57 stars for every third day. Monthly apparent positions are tabulated for an additional 173 navigational stars. The positions are tabulated to an angular accuracy of 0.1 minute of arc, which is equivalent to 0.1 nautical mile (0.2 km). Since tabular quantities must be combined to derive the navigational fix, this tabular accuracy is sufficient to produce a computed position with an error no greater than 0.3 to 0.4 nmi (0.6 to 0.7 km).*The Nautical Almanac* also contains the times of sunrise, sunset, moonrise, moonset, and twilight for various latitudes. A diary of astronomical phenomena, predictions of lunar and solar eclipses, visibility of planets, and other information of interest to the navigator are also included.

The Air Almanac is arranged with two pages of data back to back on a single sheet for each day; the positions of the Sun, first point of Aries, three planets, and the Moon are given at 10-min intervals. As necessary, information is adjusted so that the tabulated data at any given time can be used during the interval to the next entry, without interpolation, to an accuracy sufficient for practical air navigation. The times of sunrise, sunset, twilight, moonrise, and moonset are also given daily. Also provided are star recognition charts; a sky diagram; rising, setting, and depression graphs; standard times; and correction tables necessary for air navigation. While designed for air navigators, *The Air Almanac* is used by mariners who accept the reduced accuracy in exchange for its greater convenience compared with *The Nautical Almanac.* A number of countries publish in their languages air and nautical almanacs that are based on, or similar to, the English language versions.

Other almanacs. *The Astronomical Phenomena,* published annually, contains the times for phenomena such as eclipses, conjunctions, occultations, sunrise, sunset, moonrise, moonset, and moon phases, seasons, and visibility of planets. Also, the dates of civil and religious holidays are included.

For surveyors, the Royal Greenwich Observatory publishes *The Star Almanac,* designed to permit the surveyor to determine a geographical position from celestial observations.

Ephemerides of Minor Planets is prepared annually by the Institute of Applied Astronomy, St. Petersburg, and published by the Academy of Sciences of Russia. This volume contains the elements, opposition dates, and opposition ephemerides of all numbered minor planets. *See* ASTEROID.

Computer-based almanacs. The *Almanac for Computers* was published annually by the U.S. Naval Observatory from 1977 to 1991 to provide the coefficients for power series and Chebyshev polynomials for computation of any specific kind of astronomical data. The Royal Greenwich Observatory began publishing in 1981 *Compact Data for Navigation and Astronomy.* The Naval Observatory introduced the *Floppy Almanac* as an annual source of astronomical data by means of widely used personal-computer software. The latest source of high-precision astronomical data is the *Multi-Year Interactive Computer Almanac (MICA),* which provides data on compact disks. Thus, in addition to being able to compute the almanac data as published, the user is able to compute data for a particular location and time. The Naval Observatory also prepares a *Satellite Almanac,* which provides accurate positions for the planetary satellites.

The Naval Observatory provides software that can be used for computation of apparent places and transformation of coordinates. Astronomical data in the form of star catalogs are available from astronomical data centers in Strasbourg, France, and Goddard Space Flight Center in Maryland.

Publication and distribution. *The Astronomical Almanac, The Nautical Almanac, The Air Almanac*, and *Astronomical Phenomena* are cooperative publications of the U.S. Naval Observatory and Rutherford Appleton Laboratory, and are available from the U.S. Government Printing Office and the Stationery Office. These publications are used in many countries, and similar references, generally based on the same basic data, are published in various languages by the Spanish, Chinese, Russian, Japanese, German, French, Indian, Argentinian, Brazilian, Danish, Greek, Indonesian, Italian, Korean, Mexican, Norwegian, Peruvian, Philippine, and Swedish governments. Examples of publications comparable to *The Astronomical Almanac* are the *Astronomical Almanac* by Russia, the *Efemerides Astronomicas* by Spain, the *Chinese Astronomical Ephemeris,* the *Indian Ephemeris,* and the *Japanese Ephemeris.*

P. K. Seidelmann

Almond

A small deciduous tree, *Prunus amygdalus* (also known as *P. dulcis* or *Amygdalus communis*), closely related to the peach and other stone fruits and grown widely for its edible seeds. The fruit, classified botanically as a drupe (see **illus.**), is analogous to that of the peach except that the outer fleshy (mesocarp) layer does not enlarge in size during the latter part of the fruit growth-curve. Instead it splits (dehisces) as it nears maturity and separates from the shell (endocarp) and then dries. The nut with the kernel, or seed, within is then released. *See* PEACH; ROSALES.

Origin. Numerous wild almond species are found in the mountainous and desert regions extending from southwestern Europe to Afghanistan, Turkistan, and western China, sometimes occurring as extensive forests and thickets. The cultivated almond species apparently originated in central Asia (Iran) from hybridization among native species followed by local seedling selection by native peoples. Its cultivation spread with civilization along the shores of the Mediterranean Sea into North Africa and to Italy, Spain, southern France, and Portugal. From there seeds and scions were used to introduce the crop to other parts of the world, including the United States, Australia, South Africa, and South America.

Production. In the United States, commercial production is limited to California; one-half or more of this production is exported. Spain is the second leading producer, but the amount produced is about one-half that of the United States. Italy has historically been a leading producer, primarily from the Bari and Sicily areas, but production has declined sharply.

Cultivation. Because of the early spring flowering habit, commercial culture of the almond is limited to areas with mild winters and few late spring frosts. The trees thrive under moderate to high summer temperatures.

Trees can withstand considerable drought due to natural adaptation to arid conditions. Thus traditional older culture in European and Asian areas has been largely restricted to hillside and otherwise marginal areas without irrigation. However, the almond responds so well to irrigation, improved nutrition, and intensification of culture (two- to threefold increases in yield) that modern culture systems in California, and now in other parts of the world, utilize the best soil areas, supplemental irrigation, fertilization, disease and insect control, herbicide application, and all other aspects of advanced agriculture.

Original establishment of almonds was by seeds, but modern cultivation includes selection of cultivars budded or grafted to rootstocks. Many cultivars exist, and in general these have arisen as local or chance seedlings from a particular production area, such as California, Spain, Sicily, Bari, and Australia.

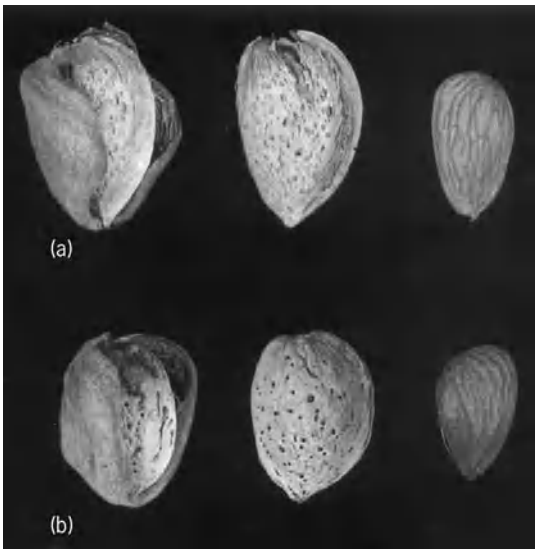

Nut characteristics of the two major almond cultivars of California: (*a*) nonpareil, (*b*) mission. Left: Mature nut with dehisced hull; middle: nut with shell; right: kernels.

Most cultivars are self-incompatible, and orchards require provision for combination of cultivars to provide cross pollination. Pollinizer insects, primarily honeybees, are required to accomplish cross pollination. Some self-fertile cultivars have been discovered or produced by plant breeding. *See* BREEDING (PLANT).

Harvesting. In California, essentially all almond orchards are mechanical-harvested. Nuts are knocked to the ground when they (particularly in the center of the tree) are dry. They are then swept into windrows and taken by machines after careful soil preparation. Hulls are removed from the nuts by machines in a central huller and delivered to the processor and handler. There nuts are fumigated for worm control and stored. Most nuts are shelled, except for a few special varieties. Kernels are sorted electronically and graded into various shapes and sizes. *See* AGRICULTURAL MACHINERY.

In some parts of the world where culture is less intensive, almonds are knocked to the ground by mallets or poles onto canvases spread under the trees. Nuts are then collected by hand into boxes or bags, or may be lifted onto almond "boats" or "sleds."

Use. Almonds are used in a variety of products. Some are roasted whole and salted to be used as snacks. Others are blanched (the skin is removed) by steam and subjected to slicing, dicing, or halving. These may be roasted, and go into products such as candy bars, bakery products, ice cream, and almond paste, among many other uses.

Almonds are of two general types: the bitter type is a source of prussic acid and flavoring extracts, and the sweet type has various food uses as described above. The almond kernels contain approximately 50% fat or oil, 20% protein, 20% carbohydrate, and a variety of minerals and vitamins. *See* NUT CROP CULTURE. D. E. Kester

Alpaca

A member of the camel family, Camelidae, which has been domesticated for more than 2000 years. The Camelidae belong to the mammalian order Artiodactyla, the even-toed ungulates. The alpaca (*Lama pacos*), more restricted in its range than other species of this family, is found at elevations above 12,000 ft (3600 m) along the shores of Lake Titicaca on the boundaries of Peru and Bolivia (see **illustration**).

As in other members of this family, the alpaca's neck and head are elongate and the upper lip has a deep cleft. The long, slender legs terminate in two toes; the feet are digitigrade, that is, the animals walk on the toes and not on the entire foot or the tip of the digits. Although the majority of the artiodactyl animals characteristically possess horns, the alpaca and other members of this family do not. The alpaca has 36 teeth, dental formula I 1/3 C 1/1 Pm 3/3 M 3/3.

The long, fine repellent hair, or wool, ranges from black to white and is highly prized for manufacturing cloth, particularly the white wool. Like many breeds of domesticated animals, the alpaca has been bred to produce pure strains for the wool. Although the alpaca is raised chiefly for its wool, its flesh is palatable.

Alpacas (*Lama pacos*). (*Courtesy of Brent Huffman/Ultimate Ungulate Images*)

The known extinct and extant species of the camel family provide an interesting example of discontinuous distribution. The family originated in North America and migrated centuries ago, so that the living species are now widely separated geographically. Among the living species, affinities and relationships are further complicated and obscured by their long history of domestication. *See* ARTIODACTYLA; BREEDING (ANIMAL); CAMEL; DENTITION; LLAMA; MAMMALIA; NATURAL FIBER. Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999.

Alpha Centauri

The third brightest star in the sky, apparent magnitude −0.3, and the brightest in the southern constellation Centaurus. It is the closest star to the Sun at a distance of 1.35 parsecs 2.59×10^{13} mi (or 4.16×10^{13} km), and its light takes more than 4 years to reach the Earth. It has an unusually large proper motion across the sky of 3.7 seconds of arc per year. *See* CENTAURUS.

Alpha Centauri is in reality a triple system, the two main components orbiting each other with period of nearly 80 years. They were discovered as a telescopic visual pair in 1689, and have now been followed for more than three complete revolutions. The orbit is eccentric, and their mean separation is approximately 23 astronomical units. α Cen A and B, as they are also known, are main-sequence stars of spectral types G2 and K1, respectively, and the brightest of the two is very similar to the Sun in mass (A = 1.16

and B = 0.97 solar masses) luminosity, and effective temperature. Chemical analyses indicate that the relative abundance of elements heavier than helium in α Cen A and B is almost twice as large as in the Sun. *See* ASTRONOMICAL UNIT; BINARY STAR; SPECTRAL TYPE.

The third component of this system, known as Proxima Centauri, is a faint reddish star of 11th magnitude, approximately 2.2° away in the sky. It was discovered in 1915, and has nearly the same proper motion and distance as the other two stars. Proxima is actually slightly closer to the Sun, which makes it the Earth's nearest neighbor in space. The linear separation from α Cen A and B is estimated to be about 13,000 astronomical units. Although the probability of such a close association in space and in projected motion occurring by chance among field stars is very small, measurements of the velocity of Proxima relative to the close binary cannot yet determine whether it is gravitationally bound. If its velocity exceeds the escape velocity of the system at Proxima's distance from the center of mass, it cannot be in orbit and is merely traveling together with the other stars in space. However, if it is indeed dynamically bound, the orbital period is probably on the order of 10^6 years. *See* STAR. David W. Latham

Alpha fetoprotein

A glycoprotein that is normally present in significant amounts only in the serum of the fetus. It is produced in the yolk sac, the liver, and other tissues of the gastrointestinal tract. Its role is unknown, but alpha fetoprotein may function as a carrier (or modulator of the concentration) of a small ligand, as an immunosuppressive, as a modulator of intracellular transport of unsaturated fatty acids, as a factor in estrogen transport, or as a means of binding retinoic acid.

Levels. Peak concentration of alpha fetoprotein in human fetal serum occurs at 13 weeks of gestation, and at that time it also reaches maximum levels in the amniotic fluid. Concentration of alpha fetoprotein in maternal serum peaks at approximately week 30–32 of gestation, reflecting the effect of both the production rate in the fetus, which is already declining at that age, and the mass of the growing fetal liver.

In pregnant women, there are large differences in the concentration of maternal serum alpha fetoprotein among individuals. The differences are attributable to the inbred differences in the production of alpha fetoprotein among fetuses, the sex of the fetus (males produce more alpha fetoprotein), the permeability of the intervening tissues, the weight of the mother (which reflects her blood volume), and maternal diabetes mellitus (which lowers the concentration). Ethnic differences in maternal serum alpha fetoprotein concentrations are most distinguishable between whites and blacks, the latter having a 15% higher concentration. There is a difference of only a few percent, if that, between whites and Asian or Indian populations.

Abnormal levels. Substantially increased levels of alpha fetoprotein were first observed in association with certain tumors, especially of the liver. Subsequently, increased concentrations were noted in the amniotic fluid and maternal serum of pregnant women carrying a fetus affected by an open-neural-tube defect. The two major types of neural tube defects are anencephaly, which results in a failure of the forebrain to form, and spina bifida, which is a failure of a variable part of the spinal cord to close. Both defects result in abnormalities in the overlying skin that allow excessive amounts of serum to leak into the amniotic fluid, and then into maternal blood. Thus, screening of maternal serum is now routine.

While the majority of pregnant women having a high concentration of serum alpha fetoprotein experience normal pregnancies, others manifest various problems, including multiple fetuses and placental abnormalities, or defects, such as gastroschisis, strictures of the fetal gastrointestinal or genitourinary tract, the Finnish type of nephrosis, occasional fetal tumors, and skin defects; these may result in fetal death. Further, about one out of three pregnancies in which alpha fetoprotein is increased has an adverse outcome, including fetal growth retardation, low birth weight, and increased perinatal mortality. In pregnancies where the fetus is affected by trisomy 21 (Down syndrome), there is lower concentration of alpha fetoprotein than usual. *See* DOWN SYNDROME.

Other genetic defects, including tyrosinemia, are also associated with increased alpha fetoprotein concentration. More frequently, acquired illness is responsible for the elevation. The most common cause of slight-to-moderate increases in serum alpha fetoprotein concentration is cirrhosis of the liver; the resumed production of alpha fetoprotein occurs in regenerating liver tissue. Higher concentrations are also associated with tumors, especially of the liver, gonads, and pancreas. A. Baumgarten

Bibliography. G. J. Mizejewski and H. I. Jacobson (eds.), *Biological Activities of Alpha-Fetoproteins*, vol. 1, 1989.

Alpha particles

Helium nuclei, which are abundant throughout the universe both as radioactive-decay products and as key participants in stellar fusion reactions. Alpha particles can also be generated in the laboratory, either by ionizing helium or from nuclear reactions. They expend their energy rapidly as they pass through matter, primarily by taking part in ionization processes, and consequently have short penetration ranges. Numerous technological applications of alpha particles can be found in fields as diverse as medicine, space exploration, and geology. Alpha particles are also major factors in the health concerns associated with nuclear waste and other radiation hazards.

The helium nucleus, or alpha particle (α), with mass 4.00150 atomic mass units (u) and charge +2,

is a strongly bound cluster of two protons (p) and two neutrons (n). Its stability is evident from mass-energy conservation in the hypothetical fusion reaction (1). The product mass (=4.00150 u) is less

$$2p + 2n \rightarrow \alpha \tag{1}$$

than the reactant mass (=2 × 1.00728 u + 2 × 1.00866 u) by 0.03038 u. By using Einstein's relation $E = mc^2$ (where c is the speed of light), this decrease in mass m (the alpha-particle binding energy) is equivalent to 28.3 MeV of energy E. The enormous magnitude of this energy is reflected in the fact that the fusion transformation of hydrogen into helium is the main process responsible for the Sun's energy. *See* CONSERVATION OF ENERGY; ENERGY; HELIUM; NUCLEAR BINDING ENERGY; NUCLEAR FUSION; PROTON-PROTON CHAIN; STELLAR EVOLUTION.

Discovery. Early absorption and deflection experiments with so-called rays from radioactive minerals revealed three distinct components, of which the least penetrating (named alpha rays by E. Rutherford in 1899) was shown to consist of positively charged particles, with a charge-to-mass ratio about half that of hydrogen. These particles were shown to be helium nuclei when helium gas was identified spectroscopically in a thin-walled glass vessel within which the rays had been trapped. *See* BETA PARTICLES; GAMMA RAYS.

Noting that a few alpha particles were strongly backscattered by thin foils while most passed directly through, Rutherford concluded that the occasional deflections were caused by coulombic repulsion between the alpha particles and compact positive entities which occupied only a small fraction of the total volume. These observations led Rutherford to propose his so-called planetary model of the atom in 1911. With subsequent modifications and refinements, Rutherford's visual concept of atomic structure has evolved to become the model of the atom accepted today. *See* ATOMIC STRUCTURE AND SPECTRA.

Alpha radioactivity. Coulombic repulsion between the protons within a nucleus leads to increasingly larger ratios of neutron number N to proton number Z for stable nuclei, as the mass numbers A ($= Z + N$) increase. Neutron-deficient nuclei can improve their N/Z ratios by means of alpha decay. The decay occurs because the parent nucleus has a total mass greater than the sum of the masses of the daughter nucleus and the alpha particle. The energy converted from mass energy to kinetic energy is called the Q value and is given by the Einstein mass-energy equivalency relation, Eq. (2), where c is the velocity of light in vacuum and the M is atomic mass. If the masses are given in atomic mass units, then the c^2 can be replaced with the conversion factor 931.501 MeV/u. The kinetic energy released is shared between the daughter nucleus and the alpha particle in accordance with the conservation of momentum, thus giving Eq. (3) for

$$Q = [M(\text{Parent}) - M(\text{Daughter}) - M(^4\text{He})]c^2 \tag{2}$$

$$\text{KE}_\alpha = Q\left[1 - \frac{M(^4\text{He})}{M(\text{Daughter})}\right] \tag{3}$$

the kinetic energy of the alpha particle, KE_α, in terms of the Q value and the masses. Thus, each radioactive alpha-emitting nuclide emits the alpha with a characteristic kinetic energy, which is one fingerprint in identification of the emitter. *See* NUCLEAR REACTION.

For example, in reaction (4), where the notation

$$^{235}_{92}\text{U}_{143} \rightarrow {}^{231}_{90}\text{Th}_{141} + {}^{4}_{2}\text{He}_2 \tag{4}$$

$^A{}_Z\text{U}_N$, and so forth, is used, an alpha particle (that is, a $^4{}_2\text{He}_2$ nucleus) is emitted from uranium-235, leaving thorium-231. For this decay, Q = (235.0439231 − 231.0362971 − 4.0026032)(931.501) MeV = 4.679 MeV, and thus the energy of the alpha is KE_α = 4.679 MeV (1 − 4/231) = 4.598 MeV.

An alpha particle interacts with the nucleus through the strong nuclear force, which has a very short range (~10^{-15} m) and is attractive, and the repulsive electrical (Coulomb) force. In alpha decay the Coulomb force is primarily responsible for accelerating the alpha to its final energy, but before that can happen the alpha must escape from the hold of the nuclear force. Relative to a potential energy of 0 at large distances, the alpha particle inside a heavy nucleus has a large positive component (~50 MeV) due to the Coulomb force, but a much larger negative component (~−100 MeV) due to the nuclear force. Just outside the nuclear surface, the part due to the nuclear force drops to zero, while that due to the Coulomb force is still large (~25 MeV). The total energy (potential and kinetic) of the typical decay alpha particle (4–10 MeV) is smaller than the potential energy at distances slightly larger than the nuclear radius. Thus, the alpha particle must penetrate this potential energy barrier (Coulomb barrier) in order to escape from the nucleus. Such barrier penetration is forbidden in classical physics but is permitted in quantum-mechanical theory. In quantum mechanics there is a straightforward procedure to calculate the probability that the alpha can penetrate the Coulomb barrier; the penetration probability is sensitively dependent on the energy of the alpha and the height of the barrier. Other factors that influence the half-life are the orbital angular momentum of the emitted alpha particle, which in quantum theory adds to the height of the Coulomb barrier, and the similarity of the alpha-emitting nucleus and the resulting daughter nucleus. *See* NONRELATIVISTIC QUANTUM THEORY; RADIOACTIVITY.

There are three major natural series, or chains, through which isotopes of heavy elements decay by successions of alpha decays. Within these series, and with all reaction-produced alpha emitters as well, each isotope decays with a characteristic half-life and emits alpha particles of particular energies and intensities (see **table**).

The presence of these radioactive nuclides in nature depends upon either a continuous production mechanism, for example the interaction of cosmic

Energies and intensities of alpha-particle standard sources

Source	Half-life	E_α*	I_α†
^{147}Sm	1.06×10^{11} years	2.234 ± 0.003	
^{232}Th	1.41×10^{10} years	4.013 ± 0.003	77
		3.954 ± 0.008	23
^{238}U	4.468×10^{9} years	4.197 ± 0.005	77
		4.150 ± 0.005	23
^{235}U	7.04×10^{8} years	4.400 ± 0.002	55
^{230}Th	7.54×10^{4} years	4.6877 ± 0.0015	76.3
		4.6212 ± 0.0015	23.4
^{239}Pu	2.411×10^{4} years	5.1566 ± 0.0004	73.2
		5.1438 ± 0.0008	15.1
^{210}Po	138.38 days	5.30438 ± 0.00007	
^{241}Am	432.2 years	5.48560 ± 0.00012	85.2
		5.44290 ± 0.00013	12.8
^{238}Pu	87.7 years	5.49907 ± 0.00020	71.6
		5.4563 ± 0.0002	28.3
^{242}Cm	162.8 days	6.11277 ± 0.00008	74.0
		6.06942 ± 0.00012	25.0
^{253}Es	20.4 days	6.63257 ± 0.00005	89.8

*Alpha-particle energy in MeV.
†Alpha-particle intensity in number of alpha particles per 100 disintegrations of source.

rays with the atmosphere, or extremely long half-lives of heavy radioactive nuclides produced in past cataclysmic astrophysical events, which accounts for uranium and thorium ores in the Earth. The relative abundances of uranium-238, uranium-235, and their stable final decay products in ores of heavy elements can be used to calculate the age of the ore, and presumably the age of the Earth. *See* GEOCHRONOMETRY.

In addition to the study of alpha-particle emitters that appear in nature, alpha decay has provided a useful tool to study artificial nuclei, which do not exist in nature due to their short half-lives. Alpha decay is a very important decay mode for nuclei far from stability with a ratio of protons to neutrons that is too large to be stable, especially for nuclei with atomic mass greater than 150 u. Because of the ease of detecting and interpreting decay alpha particles, their observation has aided tremendously in studying these nuclei far from stability, extending the study of nuclei to the very edge of nuclear existence. The measured energy of the alpha particle can be used in Eqs. (2) and (3) to determine the mass difference of the parent and daughter nuclei, and if one of the masses is known by previous experiments, the absolute value for the other mass can be determined. The half-life, the measured time for the alpha-particle rate to be reduced by a factor of 2, is used to calculate a decay rate, which is a product of the barrier penetration probability and nuclear structure factors. By dividing out the barrier penetration probability, a number (the reduced width) is obtained which is relatively large if the parent and daughter nuclei have the same spins and parities and similar nuclear wave functions, and is smaller for decays requiring a change in spin or nuclear wave functions. Systematic measurement of reduced widths over a range of nuclei enables scientists to trace changes in nuclear structure as a function of proton and neutron number. Nuclear structure information for more than 400 nuclides has been obtained in this way (**Fig. 1**). In addition, fine structure peaks

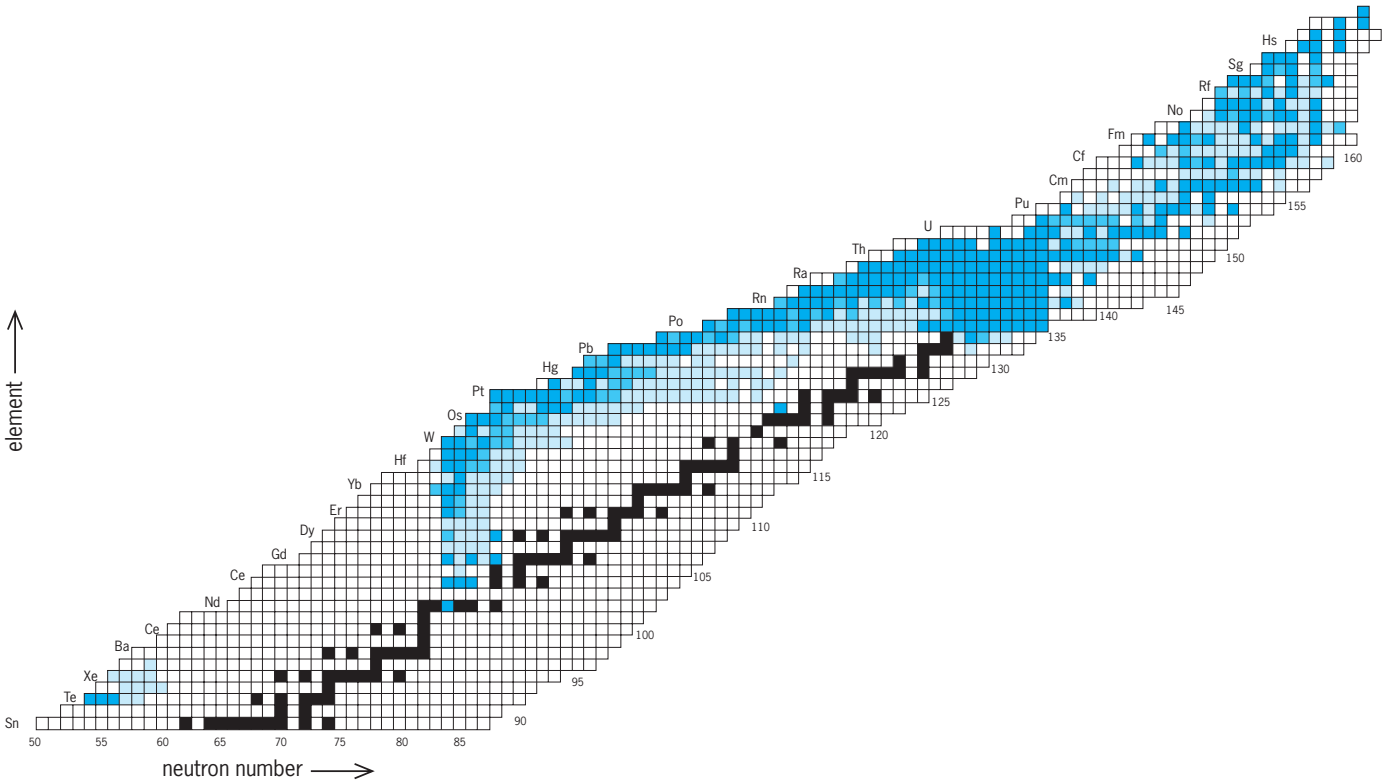

Fig. 1. Locations of known alpha emitters on the chart of nuclides are shown in color. Boxes are included for all nuclides with proton number $Z > 49$ which have been observed. The fraction of the time that a particular nuclide decays by alpha-particle emission is indicated by the shade of the colored square. Black boxes indicate the stable nuclides.

appear in the alpha-particle spectra for many of these nuclides; each such fine structure peak gives similar information about an excited state in the daughter nucleus.

Interactions with matter. By virtue of their kinetic energy, double positive charge, and large mass, alpha particles follow fairly straight paths in matter, interacting strongly with atomic electrons as they slow down and stop. These electrons may be excited to higher energy states in their host atoms, or they may be ejected, forming ion pairs in which the initial host atom becomes positively charged and the electron leaves. The more energetic ejected electrons, known as delta electrons, cause considerable secondary ionization, which accounts for 60–80% of the total ionization. A cascade of processes occurs along the alpha particle's track, leading to tens of thousands of disruptive events per alpha particle. *See* DELTA ELECTRONS; RADIATION DAMAGE TO MATERIALS.

The amount of energy expended by an alpha particle to form a single ion pair in passing through a medium is nearly independent of the alpha particle's energy, but it depends strongly on the absorbing medium. While it takes about 35 eV in air and 43 eV in helium to form an ion pair, an energy of only 2.9 eV is required in germanium and 3.6 eV in silicon. The energies expended in gases are roughly correlated to their ionization potentials. For germanium, silicon, and other semiconductors, the lower ion pair energy is, effectively, the amount required to raise an electron to the conduction band. *See* IONIZATION POTENTIAL; SEMICONDUCTOR.

The distance (or range) that an alpha particle travels before it stops depends both on the energy of the particle and on the absorbing medium. A 5.15-MeV alpha particle from plutonium-239 creates approximately 150,000 ion pairs, and travels about 1.5 in. (3.7 cm) in air before stopping. This same 5.15-MeV alpha particle has a range of about 0.0009 in. (0.0023 cm) in aluminum, 0.0014 in. (0.0035 cm) in water, and 0.0010 in. (0.0025 cm) in bone.

The passage of alpha particles through silicon is a particularly important example (**Fig. 2**). The semiconductor industry now produces chips so small that alpha particles from contaminants in the packaging materials can disrupt the memory-array areas of the chips, a serious problem which has been researched in considerable detail. About 1985, successful barriers were designed that could reduce the problem significantly, but there is concern that alpha particles might be a limiting factor in the increasing miniaturization of chips. The daily number of alpha particles to which these chips are exposed is typically less than 15/in.^{2} (2.3/cm^2). *See* INTEGRATED CIRCUITS; RADIATION HARDENING.

As alpha particles expend energy and slow down, the frequency of their interactions per unit distance increases. The rate at which an alpha particle loses energy is called its stopping power or specific ionization. A Bragg curve shows the stopping power of an alpha particle as a function of the distance to the end of its path (**Fig. 3**). The energy expended per

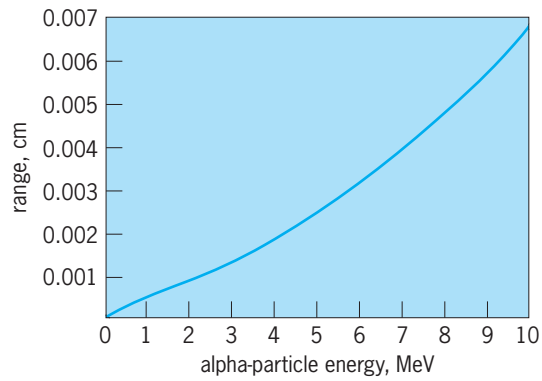

Fig. 2. Range-energy relationship for alpha particles in silicon. 1 cm = 0.4 in.

interval is seen to increase to the large, pronounced Bragg peak near the end of the path, and then to fall off to a small tail at the very end, due to range straggling.

An approximate relation for the range of alpha particles in air is given by Eq. (5). Here R is the

$$R = 0.309 \times E^{3/2} \tag{5}$$

range in centimeters (1 cm = 0.4 in.) and E is the alpha particle's initial energy in MeV. For an absorbing medium other than air, only the coefficient 0.309 would be different. According to another relationship, known as the Bragg-Kleeman rule, the relative stopping power per atom is approximately proportional to the square root of the mass number A.

In biological systems, the ionization and excitation produced by alpha particles can damage or kill cells. By rupturing chemical bonds and forming highly reactive free radicals, alpha particles can be far more destructive than other forms of radiation which interact less strongly with matter. *See* CHARGED PARTICLE BEAMS; RADIATION BIOLOGY; RADIATION CHEMISTRY; RADIATION INJURY (BIOLOGY).

Detectors. Alpha particles are detected by means of the ionization and excitation that they produce

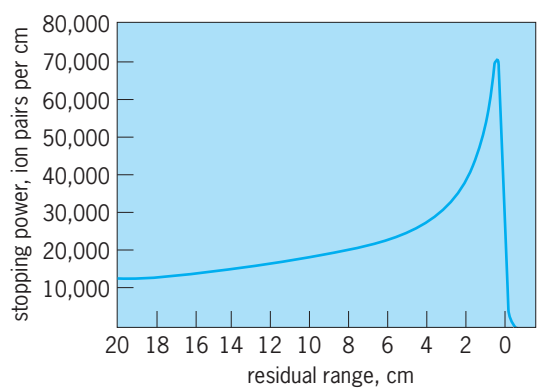

Fig. 3. Bragg curve for single alpha particle in air. The stopping power (specific ionization) in ion pairs per centimeter of air is plotted against residual range (distance from end of alpha particle's path) in centimeters. 1 cm = 0.4 in.

in matter. The first detectors were photographic plates, which turned dark when alpha particles sensitized and reduced grains of silver halide in the emulsions. In magnetic spectrometers, calibrated photographic emulsions have been used to determine alpha-particle energies and intensities.

In ionization chambers, the relationship between the number of ion pairs formed and the amount of charge collected is used in the detection process. In cloud chambers, tracks of droplets appear whenever alpha particles leave ion pairs to act as condensation centers. Finally, in present-day semiconductor detectors, the excitation of electrons from a valence band to the conduction band is detected electrically, and is then converted into precise energy and intensity values. *See* IONIZATION CHAMBER; JUNCTION DETECTOR; PARTICLE DETECTOR.

Applications. In the promising medical field of charged-particle radiotherapy, alpha particles are useful in the treatment of inaccessible tumors and vascular disorders. It is clear from the Bragg curve (Fig. 3) that the ionizing power of alpha particles is concentrated near the ends of their paths. Thus they can deliver destructive energy to a tumor while doing little damage to nearby healthy tissue. With proper acceleration, positioning, and dosage, the energy can be delivered so precisely that alpha-particle radiotherapy is uniquely suited for treating highly localized tumors near sensitive normal tissue (for example, the spinal cord). *See* RADIOLOGY.

The element-specific energies of backscattered (Rutherford-scattered) alpha particles are used in remote probes to analyze the mineral composition of geological formations. In particular, alpha particles scattered by light elements transfer more energy than those scattered by heavy elements. In another alpha-particle device, the energy from ^{238}Pu alpha decay is reliably harnessed in batteries based on the Brayton cycle, and used to power scientific equipment left on the Moon. Large power systems of this type are contemplated for use in space stations. These examples are just a few of the widespread applications of alpha particles. *See* BRAYTON CYCLE; ION-SOLID INTERACTIONS; NUCLEAR BATTERY; SPACE POWER SYSTEMS.

Hazards. Since an energetic alpha particle usually loses its energy by ionizing the atoms it encounters, it causes severe damage to living cells in its path. However, because of its short range, an alpha-particle source external to the body does not generally pose a problem. The major hazard for humans results if an alpha-particle emitter is ingested and becomes a long-term component of the body. Great care must be taken when working with alpha emitters to prevent such ingestion.

Research on indoor air quality in homes and buildings has brought attention to the long-standing health hazard associated with radon. Small amounts of this alpha-emitting gas, which is generated in the decay of heavy elements in granite and building materials, can diffuse within buildings, to be ingested by occupants. This problem is exacerbated by the fact that increasing numbers of modern buildings are tightly sealed, concentrating the radon. Environmental sources of alpha-particle exposure also include radon in mines, polonium in cigarette smoke, and thorium in coal. *See* ENVIRONMENTAL RADIOACTIVITY; RADON.

Along with the many other dangerous components, alpha-emitting isotopes in nuclear waste present serious health hazards. Efforts to achieve safe disposal of the growing accumulations of nuclear waste are receiving increased attention. *See* RADIOACTIVE WASTE MANAGEMENT.

Carrol Bingham

Bibliography. Y. A. Akovali, Review of alpha-decay data from doubly-even nuclei, *Nucl. Data Sheets*, 84:1–113, 1998; M. Eisenbud and T. F. Gesell, *Environmental Radioactivity: From Natural, Industrial, and Military Sources*, 4th ed., Academic Press, 1997; K. S. Krane, *Introductory Nuclear Physics*, Wiley, 1987; S. S. M. Wong, *Introductory Nuclear Physics*, 2d ed., Wiley, 1998; F. Yang and J. H. Hamilton, *Modern Atomic and Nuclear Physics*, McGraw-Hill, 1996.

Alpine vegetation

Plant growth forms characteristic of upper reaches of forests on mountain slopes. In such an environment, trees undergo gradual changes that, though subtle at first, may become dramatic beyond the dense forest as the zone of transition leads into the nonforested zone of the alpine tundra. In varying degrees, depending on the particular mountain setting, the forest is transformed from a closed-canopy forest to one of deformed and dwarfed trees interspersed with alpine tundra species (**Fig. 1**). This zone of transition is referred to as the forest-alpine tundra ecotone. The trees within the ecotone are stunted, often shrublike, and do not have the symmetrical shape of most trees within the forest interior. The classic image is one of twisted, stunted, and struggling individual trees clinging to a windswept ridge (**Fig. 2**). The ecotone in which these trees exist is visually one of the most striking vegetational transition areas known.

Fig. 1. Forest-alpine tundra near Hyndman Cirque, Idaho. The trees change in density and form within the ecotone beyond the boundary of the timberline upslope toward the alpine tundra. The distribution of timberline, treeline, and the ecotone is irregular because of climatic, geomorphic, and topographic controls. *(Courtesy of T. Crawford)*

Fig. 2. Whitebark pine (*Pinus albicaulis*) showing the classic dwarfed and contorted image of trees within the windswept forest-alpine tundra ecotone of the Beartooth Plateau, Wyoming. *(Courtesy of T. Crawford)*

These trees are often referred to as krummholz, a German term meaning crooked wood. In the correct scientific usage, krummholz refers only to those ecotonal trees of the European Alps that have inherited their deformed shape genetically. Trees that have not been proven to have the genetic deformation should be referred to as environmental dwarfs, cripples, elfin wood, or (in reference to their precise form) flag, flag-mat, and mat tree species. Whether or not the deformation is inherited, the ultimate cause is the presence of severe environmental conditions.

The forest-alpine tundra ecotone is a mosaic of both tree and alpine tundra species; and it extends from timberline (the upper limit of the closed-canopy forest of symmetrically shaped, usually evergreen trees) to treeline (the uppermost limit of tree species) and the exposed alpine tundra. With elevational increases, tree deformation is magnified, tree height is reduced, and the total area occupied by trees becomes smaller as the alpine shrub, grass, and herbaceous perennials become more dominant. *See* PLANTS, LIFE FORMS OF.

Distribution. The forest-alpine tundra ecotone is generally at its highest elevation in the tropics and lowest in the polar regions. It is also higher in continental areas than in marine locations. The precise distribution is controlled by a combination of climatic, topographic, and geomorphic processes. The effect of these active processes can be seen by an irregular distribution of timberlines and treelines (Fig. 1). Coniferous trees on many steep slopes are eliminated by avalanches; and they often are replaced by flexible deciduous species or herbaceous meadow cover, leaving trees on the ridges where, undisturbed, they may attain great ages. *See* ALTITUDINAL VEGETATION ZONES; PLANT GEOGRAPHY.

Environmental controls. The environment in which these tenacious individuals survive is harsh and involves a complex interaction of many factors, with the major controlling factor often being climate. The climate is characterized by a short growing season, low air temperatures, frozen soils, drought, high levels of ultraviolet radiation, irregular accumulation of snow, and strong winds. The interaction of all these factors produces varying levels of stress within the trees.

The ultimate cause of the tree deformations and of the eventual complete cessation of tree growth lies in the inability of the tissues of the shoots and the needles to mature and prepare for the harsh environmental conditions. As the length of the growing season decreases with elevation, new needles often do not mature; they have thinner cuticles (the wax-like covering on the needles that protects against desiccation and wind abrasion), and they are less acclimated against low air temperatures. Factors that particularly affect the length of the growing season include air and soil temperatures, and the depth and distribution of snow.

Low air temperatures restrict growth by limiting net uptake and assimilation of carbon dioxide, affecting the ratio of photosynthesis to respiration. With a certain duration of heat, enzymes accumulate and initiate cell division and growth. With an adequately long growing season, the tissues of the trees ripen enough to tolerate the seasonally adverse conditions. If the season is inadequate, the tissues are not well prepared.

Snow serves as the primary source of moisture for the trees; and it insulates the soil, keeping temperatures somewhat moderate. If an adequate amount of snow accumulates around the tree prior to the arrival of the low air temperatures of winter, the soil temperatures may remain above 32°F (0°C), providing liquid water. Availability of water for root uptake and an adequately developed cuticle are critical for preventing winter desiccation, considered a prime cause of tissue loss, deformation, and eventual death. Late-lying snow restricts the warming of soil in early summer, slowing the initiation of growth. Late-lying snow also provides a habitat for a parasitic snow fungus that kills needles covered by snow.

Wind is a major factor in sculpturing the trees in the ecotone. Mechanical abrasion by wind-transported snow, rock, and ice particles pits needle surfaces and breaks stems. The exposure of pitted tissue to a dry, winter atmosphere guarantees desiccation, which is usually lethal. Wind can blow snow away from the trees, exposing them to desiccation,

low air temperatures, and high amounts of radiation. *See* WIND.

Winter air temperatures become critical, depending on the timing and the status of the tree tissues. When adequately prepared, the tree tissues can tolerate temperatures as low as −40°F (−40°C). This survival mechanism, known as cold-hardy acclimation, is the seasonal transition from the tender, frost-susceptible condition to the hardy, non-frost-susceptible one. If the tissues have not adequately acclimated because of a poor growing season, the cells are lethally sensitive to temperatures above the minimums tolerated by well-prepared tissues. *See* ADAPTATION (BIOLOGY); AIR TEMPERATURE.

Radiation is critical, because the greater amounts of ultraviolet at higher elevations are magnified by the reflective snow cover. Deactivation of the chlorophyll may occur, resulting in the weakening or death of needle tissue. Large amounts of radiation can be detrimental to photosynthesis, again restricting the growth of viable tissue tolerant of the seasonal climates.

Tree forms. Changes in tree form reflect increases in severity of climate with elevational rise and topographic exposure. A common progression of form changes, upslope from timberline to treeline, is from a flag, to a flag-mat, to a mat form. The trees with a flag form have branching mainly on the leeward, protected side of the upper part of the trunk. The flag branching is exposed above the winter snow cover, and growth on the windward side is eroded away by mechanical abrasion. Snow often accumulates to great depths around the base of these trees, inducing the parasitic snow fungus there. The combination results in a distorted flag form with barren branches on the lower trunk (**Fig. 3**). The trees of flag-mat form support scrawny trunks that are flagged. Close to the ground, a mat, or shrublike growth, survives protected under the winter snowpack. At the highest elevations within the ecotone, the mat (cushion-tree) form exists. The tree is dwarfed to a mat. Rarely are vertical branches or upright leaders found. Each mat has a size and shape that is adapted to its specific topographic environment. The wind side of the mat is a contorted mass of broken needles and distorted branches. Survival of the mat is dependent on winter snow cover, as shoots that project above the snow are abraded mechanically.

Fig. 3. Flagged tree, mainly supporting branching on the leeside of the upper trunk; needles have been lost from the lower trunk to the parasitic snow fungus. ***(Courtesy of T. Crawford)***

Importance of trees. The environmentally dwarfed and contorted trees of the forest-alpine tundra ecotone are important for many reasons. Snow is maintained in the ecotone by low air temperatures and tree shading, providing a water source late into the summer season. The trees afford shade, wind protection, and the moisture of snowbanks, critical for the establishment and survival of seedlings. Many of the oldest trees in the world have been found in the ecotone. These offer yearly historical records of events (such as climate changes, volcanic eruptions, pollution, and fire), as recorded in tree rings. Finally, the ecotone offers a unique habitat for other plants and animals to thrive as the two communities combine resources.

Ecosystems of the forest-alpine tundra ecotone are increasingly subjected to disturbances from human activities. Large numbers of visitors enter mountain areas and forest-tundra ecotones in search of recreation and wilderness experience. The impact can be large. Vegetation is trampled, water often becomes polluted, and soil is eroded along trails, contributing to a progressive degradation of ecotones. Trees that grow only a few millimeters each year are sometimes stripped of their wood for campfires or even for decorative items. Then much time is needed to replenish the growth; and in many cases regrowth may not be possible under the present climate. Strategies for management must respond to the human use within these ecosystems, as well as to the impacts of acid rain and additional climatic variability (due to increasing amounts of carbon dioxide and other chemicals in the atmosphere and to depletion of ozone). A combination of good management and better understanding of the processes and dynamics of the forest-alpine tundra ecotone ecosystem would provide a safeguard against potential environmental degradation. *See* ACID RAIN; FOREST ECOSYSTEM; TREE; TUNDRA; VEGETATION AND ECOSYSTEM MAPPING. Katherine J. Hansen

Bibliography. J. D. Ives (ed.), *Mountains*, 1994; J. D. Ives and R. G. Barry (eds.), *Arctic and Alpine Environments*, 1974; L. W. Price, *Mountains and Man: A Study of Process and Environment*, 1981; W. Tranquillini, *Physiological Ecology of the Alpine Timberline*, 1979.

Alternating current

Electric current that reverses direction periodically, usually many times per second. Electrical energy is ordinarily generated by a public or a private utility organization and provided to a customer, whether industrial or domestic, as alternating current. *See* ELECTRIC POWER GENERATION.

One complete period, with current flow first in one direction and then in the other, is called a cycle, and 60 cycles per second (60 Hz) is the

customary frequency of alternation in the United States and in the rest of North America. In Europe and in many other parts of the world, 50 Hz is the standard frequency of alternation. On aircraft a higher frequency, often 400 Hz, is used to make possible lighter electrical machines.

When the term alternating current is used as an adjective, it is commonly abbreviated to ac, as in ac motor. Similarly, direct current as an adjective is abbreviated dc.

Advantages. The voltage of an alternating current can be changed by a transformer. This simple, inexpensive, static device permits generation of electric power at moderate voltage, efficient transmission for many miles at high voltage, and distribution and consumption at a conveniently low voltage. With direct (constant) current it is not possible to use a transformer to change voltage. On a few power lines, electrical energy is transmitted for great distances as direct current, but the electrical energy is generated as alternating current, transformed to a high voltage, rectified to direct current and transmitted, and then changed back to alternating current by an inverter, to be transformed down to a lower voltage for distribution and use. *See* DIRECT-CURRENT TRANSMISSION.

In addition to permitting efficient transmission of energy, alternating current provides advantages in the design of generators and motors, and for some purposes gives better operating characteristics. Certain devices involving chokes and transformers could not be operated on direct current. Also, the operation of large switches (called circuit breakers) is facilitated because the instantaneous value of alternating current automatically becomes zero twice in each cycle, and an opening circuit breaker need not interrupt the current but only prevent current from starting again after its instant of zero value. *See* ALTERNATING-CURRENT GENERATOR; ALTERNATING-CURRENT MOTOR; CIRCUIT BREAKER.

Sinusoidal form. An alternating current waveform is shown diagrammatically in **Fig. 1**. Time is measured horizontally (beginning at any arbitrary moment) and the current at each instant is measured vertically. In this diagram it is assumed that the current is alternating sinusoidally; that is, the current i is described by Eq. (1), where I_m is the maximum

$$i = I_m \sin 2\pi f t \qquad (1)$$

instantaneous current, f is the frequency in cycles

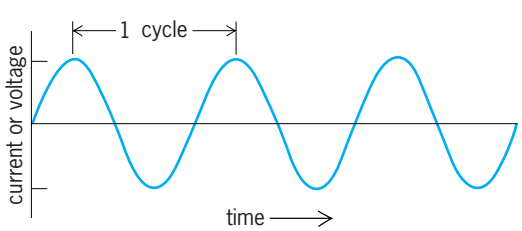

Fig. 1. Diagram of sinusoidal alternating current.

per second (hertz), and t is the time in seconds. *See* SINE WAVE.

Measurement. Quantities commonly measured by ac meters and instruments are energy, power, voltage, and current. Other quantities less commonly measured are reactive volt-amperes, power factor, frequency, and demand (of energy during a given interval such as 15 min).

Energy is measured on a watt-hour meter. There is usually such a meter where an electric line enters a customer's premises. The meter may be single-phase (usual in residences) or three-phase (customary in industrial installations), and it displays on a register of dials the energy that has passed, to date, to the system beyond the meter. The customer frequently pays for energy consumed per the reading of such a meter. *See* ELECTRICAL ENERGY MEASUREMENT; WATT-HOUR METER.

Power is measured on a wattmeter. Since power is the rate of consumption of energy, the reading of the wattmeter is proportional to the rate of increase of the reading of a watt-hour meter. The same relation is expressed by saying that the reading of the watt-hour meter, which measures energy, is the integral (through time) of the reading of the wattmeter, which measures power. A wattmeter usually measures power in a single-phase circuit, although three-phase wattmeters are sometimes used. *See* ELECTRIC POWER MEASUREMENT; WATTMETER.

Current is measured by an ammeter. Power absorbed by an element is the product of current through the element, voltage across it, and the power factor between the current and the voltage, as shown in Eq. (5). With unidirectional (direct) current, the amount of current is the rate of flow of electricity; it is proportional to the number of electrons passing a specified cross section of a wire per second. This is likewise the definition of current at each instant of an alternating-current cycle, as current varies from a maximum in one direction to zero and then to a maximum in the other direction (Fig. 1). An oscilloscope will indicate instantaneous current, but the value of instantaneous current is not often useful for analysis purposes. A dc (d'Arsonval-type) ammeter will measure average current, but this is useless in an ac circuit, for the average of sinusoidal current is zero. A useful measure of alternating current is found in the ability of the current to do work, and the amount of current is correspondingly defined as the square root of the average of the square of instantaneous current, the average being taken over an integer number of cycles. This value is known as the root-mean-square (rms) or effective current. It is measured in amperes. It is a useful measure for current of any frequency. The rms value of direct current is identical to its dc value. The rms value of sinusoidally alternating current is $I_m/2$, where I_m is the maximum instantaneous current. [See Fig. 1 and Eq. (1).] *See* AMMETER; CURRENT MEASUREMENT; OSCILLOSCOPE.

Voltage is measured by a voltmeter. Voltage is the electrical pressure. It is measured between one point and another in an electric circuit, often between the

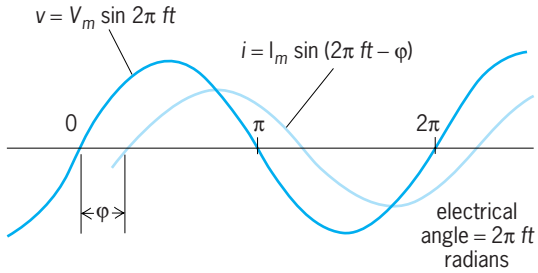

Fig. 2. Phase angle φ.

two wires of the circuit. As with current, instantaneous voltage in an ac circuit reverses each half cycle and the average of sinusoidal voltage is zero. Therefore the rms or effective value of voltage is used in ac systems. The rms value of sinusoidally alternating voltage is $V_m/2$, where V_m is the maximum instantaneous voltage. This rms voltage, together with rms current and the circuit power factor, is used to compute electric power, as in Eqs. (4) and (5). *See* VOLTAGE MEASUREMENT; VOLTMETER.

The ordinary voltmeter is connected by wires to the two points between which voltage is to be measured, and voltage is proportional to the current that results through a very high electrical resistance within the voltmeter itself. The voltmeter, actuated by this current, is calibrated in volts.

Phase difference. Phase difference is a measure of the fraction of a cycle by which one sinusoidally alternating quantity leads or lags another. **Figure 2** shows a voltage waveform v which is described in Eq. (2) and a current i which is described in Eq. (3).

$$v = V_m \sin 2\pi f t \tag{2}$$

$$i = I_m \sin (2\pi f t - \varphi) \tag{3}$$

The angle φ is called the phase difference between the voltage and the current; this current is said to lag (behind this voltage) by the angle φ. It would be equally correct to say that the voltage leads the current by the phase angle φ. Phase difference can be expressed as a fraction of a cycle, as an angle in degrees, or as an angle in radians, as shown in Eq. (3). If there is no phase difference, and $\varphi = 0$, voltage and current are in phase. If the phase difference is a quarter cycle, and $\varphi = \pm 90^\circ$, the quantities are said to be in quadrature.

Power factor. Power factor is defined in terms of the phase angle. If the rms value of sinusoidal current from a power source to a load is I and the rms value of sinusoidal voltage between the two wires connecting the power source to the load is V, the average power P passing from the source to the load is shown as Eq. (4). The cosine of the phase angle,

$$P = VI \cos \varphi \tag{4}$$

$\cos \varphi$, is called the power factor. Thus the rms voltage, the rms current, and the power factor are the components of power.

The foregoing definition of power factor has meaning only if voltage and current are sinusoidal. Whether they are sinusoidal or not, average power, rms voltage, and rms current can be measured, and a value for power factor is implicit in Eq. (5). This

$$P = VI \text{ (power factor)} \tag{5}$$

gives a definition of power factor when V and I are not sinusoidal. If the voltage across the element and the current through it are in phase (and of the same waveform), power factor equals 1. If voltage and current are out of phase, power factor is less than 1. If voltage and current are sinusoidal and in quadrature, then the power factor is zero.

The phase angle and power factor of voltage and current in a circuit that supplies a load are determined by the load. Thus a load of pure resistance, such as an electric heater, has unity power factor. An inductive load, such as an induction motor, has a power factor less than 1 and the current lags behind the applied voltage. A capacitive load, such as a bank of capacitors, also has a power factor less than 1, but the current leads the voltage, and the phase angle φ is negative.

If a load that draws lagging current (such as an induction motor) and a load that draws leading current (such as a bank of capacitors) are both connected to a source of electric power, the power factor of the two loads together can be higher than that of either one alone, and the current to the combined loads may have a smaller phase angle from the applied voltage than would currents to either of the two loads individually. Although power to the combined loads is equal to the arithmetic sum of power to the two individual loads, the total current will be less than the arithmetic sum of the two individual currents (and may, indeed, actually be less than either of the two individual currents alone). It is often practical to reduce the total incoming current by installing a bank of capacitors near an inductive load, and thus to reduce power lost in the incoming distribution lines and transformers, thereby improving efficiency. This process is called power-factor correction.

Three-phase system. Three-phase systems are commonly used for generation, transmission, and distribution of electric power. A customer may be supplied with three-phase power, particularly if a large amount of power is needed or if the customer wishes to use three-phase loads. Small domestic customers are usually supplied with single-phase power.

A three-phase system is essentially the same as three ordinary single-phase systems with the three voltages of the three single-phase systems out of phase with each other by one-third of a cycle (120°), as shown in **Fig. 3**. The three voltages may be written

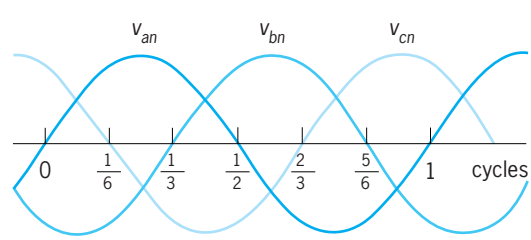

Fig. 3. Voltages of a balanced three-phase system.

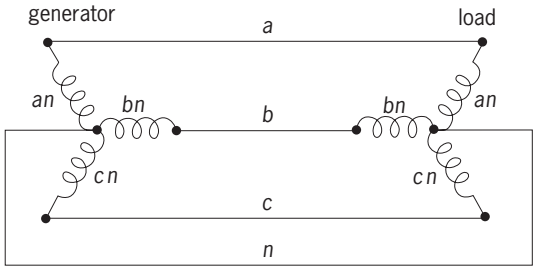

Fig. 4. Connections of a simple three-phase system.

as Eqs. (6), (7), and (8), where $V_{an(\max)}$ is the maxi-

$$v_{an} = V_{an(\max)} \sin 2\pi ft \tag{6}$$

$$v_{bn} = V_{bn(\max)} \sin 2\pi (ft - {}^1/_3) \tag{7}$$

$$v_{cn} = V_{cn(\max)} \sin 2\pi (ft - {}^2/_3) \tag{8}$$

mum value of voltage in phase *an*, and so on. The three-phase system is balanced if relation (9) holds,

$$V_{an(\max)} = V_{bn(\max)} = V_{cn(\max)} \tag{9}$$

and if the three phase angles are equal, one-third cycle each as shown.

If a three-phase system were actually three separate single-phase systems, there would be two wires between the generator and the load of each system, requiring a total of six wires. In fact, however, a single wire can be common to all three systems, so that it is only necessary to have three wires for a three-phase system (**Fig. 4***a–c*) plus a fourth wire *n* serve as a common return or neutral conductor. On some systems the Earth is used as the common or neutral conductor.

Each phase of a three-phase system carries current and conveys power and energy. If the three loads on the three phases of the three-phase system are identical and the voltages are balanced, then the currents are balanced also. Figure 2 can then apply to any one of the three phases. It will be recognized that the three currents in a balanced system are equal in rms (or maximum) value and that they are separated one from the other by phase angles of one-third cycle and two-thirds cycle. Thus the currents (in a balanced system) are themselves symmetrical. Note, however, that the three currents will not necessarily be in phase with their respective voltages; the corresponding voltages and currents will be in phase with each other only if the load is pure resistance and the phase angle between voltage and current is zero. Otherwise some such relation as that of Fig. 2 will apply to each phase.

It is significant that, if the three currents of a three-phase system are balanced, their sum is zero at every instant. In practice, the three currents are not usually exactly balanced, and one of two situations arises. Either the common neutral wire *n* is used, in which case it carries little current (and may be of high resistance compared to the other three line wires), or else the common neutral wire *n* is not used, only three line wires being installed, and the three phase currents are thereby forced to add to zero even though this requirement results in some imbalance of phase voltages at the load.

It is also significant that the total instantaneous power from generator to load is constant (does not vary with time) in a balanced, sinusoidal, three-phase system. Power in a single-phase system that has current in phase with voltage is maximum when voltage and current are maximum, and is instantaneously zero when voltage and current are zero; if the current of the single-phase system is not in phase with the voltage, the power will reverse its direction of flow during part of each half cycle. However, in a balanced three-phase system, regardless of the phase angle, the flow of power is unvarying from instant to instant. This results in smoother operation and less vibration of motors and other ac devices.

Three-phase systems are almost universally used for large amounts of power. In addition to providing smooth flow of power, three-phase motors and generators are more economical than single-phase machines. Polyphase systems with two, four, or other numbers of phases are possible, but they are little used except when a large number of phases, such as 12, is desired for economical operation of a rectifier. H. H. Skilling

Symmetrical (0, 1, 2) components. When three coils are equally spaced around the periphery of the stator of a generator, a properly shaped magnetic field rotating in a forward direction at uniform velocity will induce voltages in these coils. The three voltages may be written as Eqs. (6), (7), and (8). For analytical purposes it is more convenient to express the voltages as phasors, as in Eqs. (10). For a discus-

$$\begin{aligned} \mathbf{V}_{a1} &= \mathbf{V}_{a1} \\ \mathbf{V}_{b1} &= \mathbf{V}_{a1}\, e^{-j2\pi/3} \\ \mathbf{V}_{c1} &= \mathbf{V}_{a1}\, e^{+j2\pi/3} \end{aligned} \tag{10}$$

sion of phasor notation and complex representation. *See* ALTERNATING-CURRENT CIRCUIT THEORY.

In order to simplify notation, it has become accepted practice to introduce a standard set of unit three-phase phasors as in Eqs. (11), so that Eqs. (10)

$$1 = e^{j0} \quad \mathbf{a} = e^{j2\pi/3} \quad \mathbf{a}^2 = e^{-j2\pi/3} \tag{11}$$

may be written more succinctly in matrix format as in Eq. (12).

$$\begin{bmatrix} \mathbf{V}_{a1} \\ \mathbf{V}_{b1} \\ \mathbf{V}_{c1} \end{bmatrix} = (\mathbf{V}_{a1}) \begin{bmatrix} 1 \\ \mathbf{a}^2 \\ \mathbf{a} \end{bmatrix} \tag{12}$$

See MATRIX THEORY.

When the voltages proceed in time-phase in the order $\mathbf{V}_a$, $\mathbf{V}_b$, $\mathbf{V}_c$, as reflected in Eqs. (6), (7), and (8), the set of three-phase voltage is called a positive-sequence set of voltages and is identified by the subscript 1, as in Eqs. (10) and (12). When three coils are equally spaced around the periphery of the stator of a generator, a properly shaped magnetic field rotating in a backward direction at uniform velocity will induce voltages in these coils which proceed in time-phase in the order $\mathbf{V}_a$, $\mathbf{V}_c$, $\mathbf{V}_b$, and the set of

three-phase voltages is called a negative-sequence set of voltages and identified by the subscript 2. In terms of phasor notation, the voltages are written as in Eqs. (13).

$$\begin{bmatrix} \mathbf{V}_{a2} \\ \mathbf{V}_{b2} \\ \mathbf{V}_{c2} \end{bmatrix} = (\mathbf{V}_{a2}) \begin{bmatrix} 1 \\ \mathbf{a} \\ \mathbf{a}^2 \end{bmatrix} \tag{13}$$

If coils b and c are placed in the same slot with coil a, then $\mathbf{V}_a = \mathbf{V}_b = \mathbf{V}_c$, and the set of three-phase voltages is called a zero-sequence set of voltages and identified by the subscript 0. In terms of phasor notation the voltages are written as in Eqs. (14).

$$\begin{bmatrix} \mathbf{V}_{a0} \\ \mathbf{V}_{b0} \\ \mathbf{V}_{c0} \end{bmatrix} = (\mathbf{V}_{a0}) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{14}$$

The set of positive-, negative-, and zero-sequence components are collectively referred to as symmetrical components.

If the voltages $\mathbf{V}_a$, $\mathbf{V}_b$, and $\mathbf{V}_c$ are written as in Eqs. (15), and, in matrix format, as in Eq. (16), then

$$\begin{aligned} \mathbf{V}_a &= \mathbf{V}_{a0} + \mathbf{V}_{a1} + \mathbf{V}_{a2} = \mathbf{V}_{a0} + \mathbf{V}_{a1} + \mathbf{V}_{a2} \\ \mathbf{V}_b &= \mathbf{V}_{b0} + \mathbf{V}_{b1} + \mathbf{V}_{b2} = \mathbf{V}_{a0} + \mathbf{a}^2\mathbf{V}_{a1} + \mathbf{a}\mathbf{V}_{a2} \\ \mathbf{V}_c &= \mathbf{V}_{c0} + \mathbf{V}_{c1} + \mathbf{V}_{c2} = \mathbf{V}_{a0} + \mathbf{a}\mathbf{V}_{a1} + \mathbf{a}^2\mathbf{V}_{a2} \end{aligned} \tag{15}$$

$$\begin{bmatrix} \mathbf{V}_a \\ \mathbf{V}_b \\ \mathbf{V}_c \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \mathbf{a}^2 & \mathbf{a} \\ 1 & \mathbf{a} & \mathbf{a}^2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_{a0} \\ \mathbf{V}_{a1} \\ \mathbf{V}_{a2} \end{bmatrix} \tag{16}$$

by selecting complex values for $\mathbf{V}_{a0}$, $\mathbf{V}_{a1}$, and $\mathbf{V}_{a2}$, respectively, in an arbitrary manner, it is possible to establish an infinite number of unbalanced phase voltages $\mathbf{V}_a$, $\mathbf{V}_b$, and $\mathbf{V}_c$. Conversely, and more important, since the inverse relationship given by Eq. (17)

$$\begin{bmatrix} \mathbf{V}_{a0} \\ \mathbf{V}_{a1} \\ \mathbf{V}_{a2} \end{bmatrix} = (1/3) \begin{bmatrix} 1 & 1 & 1 \\ 1 & \mathbf{a} & \mathbf{a}^2 \\ 1 & \mathbf{a}^2 & \mathbf{a} \end{bmatrix} \begin{bmatrix} \mathbf{V}_a \\ \mathbf{V}_b \\ \mathbf{V}_c \end{bmatrix} \tag{17}$$

exists, for any given unbalanced set of three-phase voltages, it is always possible to determine a corresponding unique set of zero-, positive-, and negative-sequence voltages. Equally well, for any given unbalanced set of three-phase currents, it is always possible to determine a corresponding unique set of zero-, positive-, and negative-sequence currents.

Three-phase power system components, such as transmission lines, transformers, loads, motors, and generators, may, in many practical instances, be considered geometrically symmetrical among the three phases. In this event, it can be shown that when a zero-, positive-, or negative-sequence voltage is applied to such a component, the resultant current will be a zero-, positive-, or negative-sequence component, respectively. The associated zero-, positive-, and negative-sequence impedances can be readily identified.

Consequently, because of the symmetrical nature of each of the symmetrical-component voltages and currents, it is necessary to consider only one of the three phases of the three-phase power system in an analysis. By accepted convention, phase "a" is the selected phase.

Perhaps the most important virtue of symmetrical components lies in the fact that if a common type of unsymmetrical fault (single-line-to-ground fault, line-to-line fault, one open conductor, two open conductors, and so forth) occurs at one point in an otherwise symmetrical three-phase network, a relatively simple interconnection occurs between the symmetrical component networks at the fault location.

M. Harry Hesse

Power and information. Although this article has emphasized electric power, ac circuits are also used to convey information. An information circuit, such as telephone, radio, or control, employs varying voltage, current, waveform, frequency, and phase. Efficiency is often low, the chief requirement being to convey accurate information even though little of the transmitted power reaches the receiving end. For further consideration of the transmission of information. *See* ELECTRICAL COMMUNICATIONS; RADIO; TELEPHONE; WAVEFORM.

An ideal power circuit should provide the customer with electrical energy always available at unchanging voltage of constant waveform and frequency, the amount of current being determined by the customer's load. High efficiency is greatly desired. *See* CAPACITANCE; CIRCUIT (ELECTRICITY); ELECTRIC CURRENT; ELECTRIC FILTER; ELECTRICAL IMPEDANCE; ELECTRICAL RESISTANCE; INDUCTANCE; JOULE'S LAW; NETWORK THEORY; OHM'S LAW; RESONANCE (ALTERNATING-CURRENT CIRCUITS).

H. H. Skilling

Bibliography. J. Arrillaga and C. P. Arnold, *Computer Analysis of Power Systems*, 1990; O. I. Elgerd, *Electric Energy Systems Theory*, 2d ed., 1982; J. J. Grainger and W. D. Stevenson, Jr., *Power System Analysis*, 1994; C. P. Paul, S. A. Nasar, and L. E. Unnewehr, *Introduction to Electrical Engineering*, 2d ed., 1992.

Alternating-current circuit theory

The mathematical theory for the analysis of electric circuits when the currents and voltages are alternating functions of time.

In an electric circuit carrying direct current only (**Fig. 1**), the current I (amperes) flowing through a circuit of total resistance R (ohms) when a steady voltage V (volts) is applied is given by Ohm's law, Eq. (1).

$$I = \frac{V}{R} \tag{1}$$

See CIRCUIT (ELECTRICITY); DIRECT-CURRENT CIRCUIT THEORY; OHM'S LAW.

In many practical situations, electrical energy is generated, transmitted, and distributed in a form in which the currents and voltages are not constant but are sinusoidally varying or "alternating" functions of time (**Fig. 2**). Public electricity supply systems are

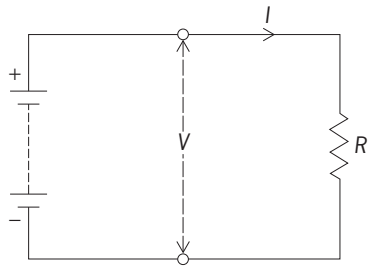

Fig. 1. Direct-current circuit.

of this type, and many telecommunications systems use electrical signals of sinusoidal form as carriers of information. It is necessary to have a comprehensive theory to analyze such systems and to provide the basis for their design.

This is the object of alternating-current circuit theory. It differs from direct-current theory in that there are circuit elements other than resistors (particularly inductors and capacitors) which respond to (impede) the flow of current: these are known generally as impedances. Also, it is possible to couple energy from one circuit to another by means of mutual inductances or transformers without the need for a linking conductor. Whereas in the dc case the sizes of voltages, currents, and resistors can all be described by single numbers, in the ac case pairs of numbers are needed to describe amplitude and phase or, for an impedance, amplitude ratio and phase shift. Mathematically, the properties and interrelations of pairs of numbers can be dealt with succinctly by using complex algebra. Alternating-current circuit theory is treated using simple time-domain analysis of waveforms, illustrated by phasor diagrams. It will be shown how the use of complex algebra provides a more powerful and elegant theory.

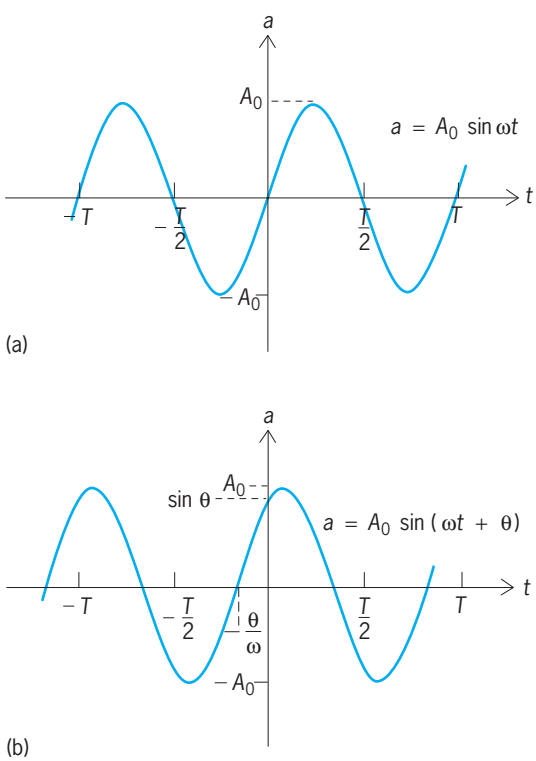

Fig. 2. Sinusoidal waveforms. (*a*) Simple sinusoid without phase shift (*b*) Sinusoid with phase shift.

Waveforms. A sinusoidal (or alternating) function a of time t is shown in Fig. 2*a*. It is described by Eq. (2), where A_0 is the maximum value (or amplitude) of a and ω is known as the angular or radian frequency. The period T of the function is the time after which the function repeats itself. During one period the function takes on all its possible values and is said to go through one cycle. The frequency f of the sinusoid as usually defined (for example, 60 Hz) is the number of cycles that occur in 1 s, so $f = 1/T$. Since $\sin(x + 2\pi) = \sin x$, $\omega T = 2\pi$, and hence $\omega = 2\pi f$.

$$a(t) = A_0 \sin \omega t \tag{2}$$

It is necessary to consider waveforms which do not cross the zero level at $t = 0$ but which include a phase shift. Figure 2*b* shows the waveform of Eq. (3), which waveform is identical in shape with the waveform described by Eq. (2) but shifted in time so that it crosses the zero level at a time θ/ω before $t = 0$. This is referred to as a phase advance; the corresponding waveform with θ negative would illustrate a phase delay.

$$a(t) = A_0 \sin(\omega t + \theta) \tag{3}$$

Since $\sin(x + \pi/2) = \cos x$, Eq. (3) could equally well be written as Eq. (4), where the phase angle is now $\psi = \theta - \pi/2$ radians. Henceforth, either of the forms in Eqs. (3) and (4) will be used, as convenient, to represent an alternating quantity. An alternating voltage is written as Eq. (5), and an alternating current as Eq. (6), where v and i represent the instantaneous values, and V_0 and I_0 the maximum values or amplitudes, of voltage and current respectively. *See* ALTERNATING CURRENT.

$$\begin{aligned} a &= A_0 \sin\left(\omega t + \frac{\pi}{2} + \theta - \frac{\pi}{2}\right) \\ &= A_0 \cos(\omega t + \psi) \end{aligned} \tag{4}$$

$$v = V_0 \sin(\omega t + \theta) \text{ or } v = V_0 \cos(\omega t + \psi) \tag{5}$$

$$i = I_0 \sin(\omega t + \theta) \text{ or } i = I_0 \cos(\omega t + \psi) \tag{6}$$

Phasor diagrams. It is helpful to visualize an alternating waveform as the projection of a rotating vector. This may be considered as corresponding in some sense to the generation of an alternating current by the rotation of a coil in a magnetic field.

In **Fig. 3***a*, OR is a vector of length A_0 rotating in the anticlockwise direction about the origin of the x-y plane with angular velocity ω radians per second. Then the projection of the vector OR on the x axis is $OX = A_0 \cos \omega t$, and the projection on the y axis is $OY = A_0 \sin \omega t$. The rotating vector can therefore be thought of as generating a sinusoidal waveform.

If, instead of the vector OR, the vector OR' in Fig. 3*b* is considered, displaced from OR by an angle ϕ and rotating with it, then $OX' = \cos(\omega t + \phi)$ and

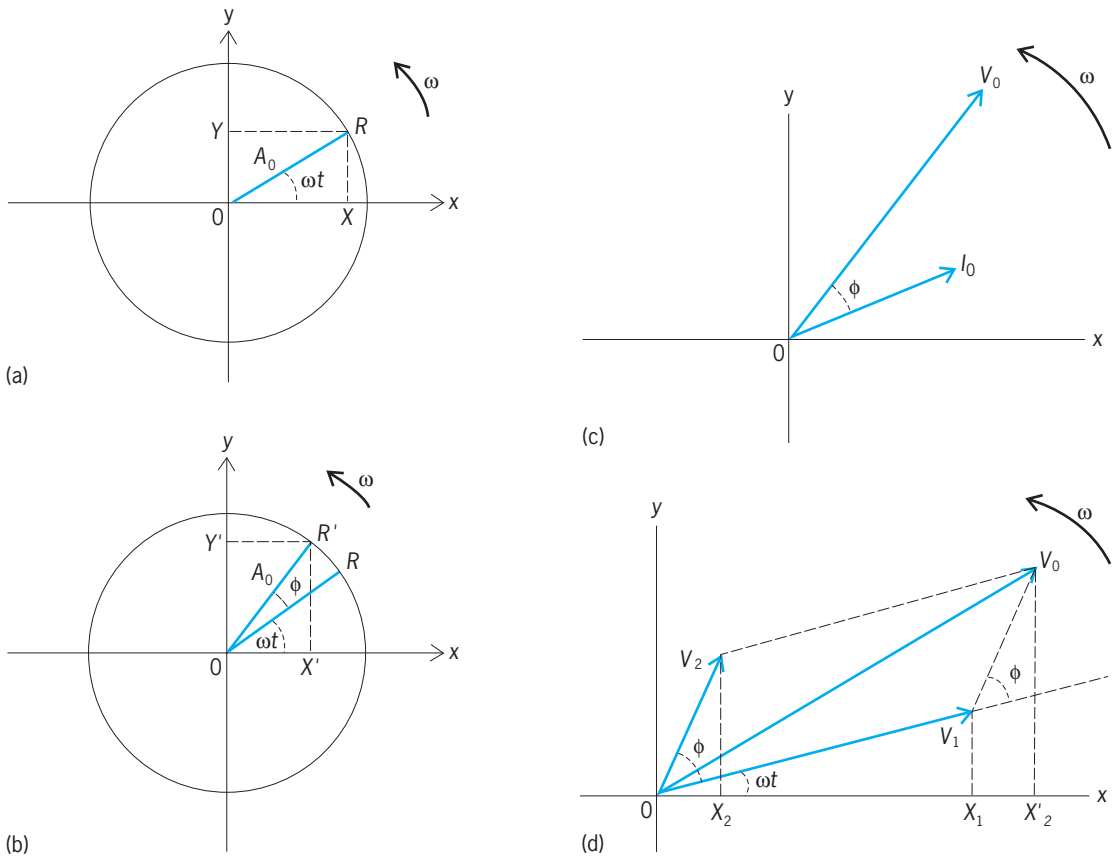

Fig. 3. Phasors. (*a*) Simple phasor. (*b*) Phasor with phase shift. (*c*) Voltage and current phasors. (*d*) Phasor for the sum of two voltages.

$OY' = \sin(\omega t + \phi)$, so that rotation of the phasor (the name given to the vector in this type of diagram) corresponds to a phase shift in the waveform.

In a linear circuit, the voltage and current waveforms have related amplitudes and often a phase shift between them, but both are of the same frequency. They can thus be represented, as in Fig. 3*c*, as the projections of phasors of amplitudes V_0 and I_0 with an angle ϕ between them, rotating together in the $x-y$ plane. Circuit theory enables the ratio V_0/I_0 phase shift ϕ to be calculated from the circuit parameters.

If there are two voltages (or two currents) of different amplitudes and with a phase shift between them, say $\upsilon_1 = V_1 \cos \omega t$ and $\upsilon_2 = V_2 \cos(\omega t + \phi)$, represented as the projection on the x-axis of two phasors of amplitudes V_1 and V_2, it is easy to see from Fig. 3*d* that their sum can be represented as the projection of the resultant phasor V_0, the vector sum of V_1 and V_2. For $OX_1 = \upsilon_1$, $X_1 X'_2 = OX_2 = \upsilon_2$, and therefore $OX'_2 = \upsilon_1 + \upsilon_2$.

Complex representation. While a phasor diagram helps to visualize the relationship between current and voltage in an ac circuit, it does little to help with the mathematical calculations. These are greatly simplified if complex algebra is used to describe the current, voltage, and circuit parameters. Only the formulas of complex algebra essential for circuit theory are reviewed here. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

A complex number z is defined by a pair of ordinary numbers. It can be written as in Eq. (7), where

$$z = x + jy \tag{7}$$

x and y are real numbers and $j = -1$. (Mathematicians often use i for this quantity, but in electrical work the symbol i is reserved for electric current.) Here x is called the real part of z, written Re (z), and y the imaginary part, written Im (z). If x and y are regarded as coordinates in a (complex) plane, z defines a point in the plane, as in **Fig. 4***a*. The distance r from the origin O to z is known as the modulus of z, written $|z|$; and the angle θ between the line from O to z and the real (x) axis is known as the argument of z. These quantities are given by Eqs. (8) and (9), and x and y are expressed in terms of them by Eqs. (10).

$$|z| = r = \sqrt{x^2 + y^2} \tag{8}$$

$$\theta = \tan^{-1} y/x \tag{9}$$

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \end{aligned} \tag{10}$$

De Moivre's theorem states that Eq. (11) is valid,

$$e^{j\theta} = \cos\theta + j\sin\theta \tag{11}$$

so an alternative way of writing z is Eq. (12),

and in giving numerical values it is often convenient to write Eq. (13).

$$z = re^{j\theta} \tag{12}$$

$$z = r\angle\theta \tag{13}$$

The complex conjugate of z, z^* is defined by Eq. (14). This is illustrated in Fig. 4*b*.

$$z^* = x - jy = re^{-j\theta} \tag{14}$$

Multiplication by j is equivalent to rotation through $\pi/2$ radians, or 90°, as Fig. 4*c* illustrates, for

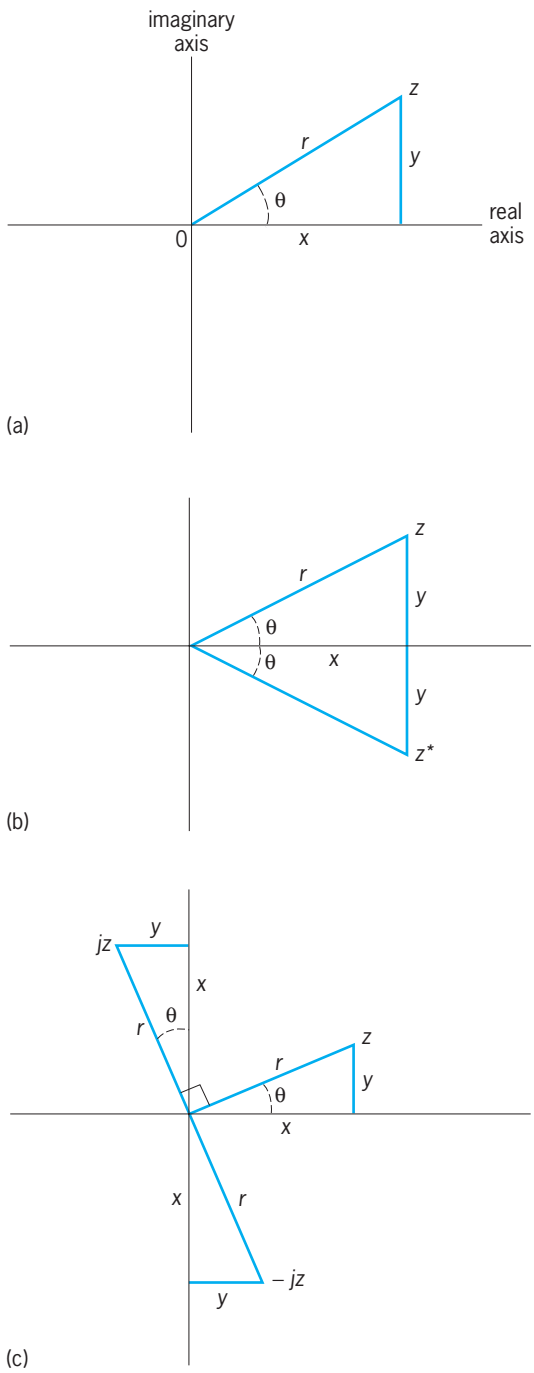

Fig. 4. Complex numbers. (*a*) The number $z = x + jy = re^{j\theta}$. (*b*) Its complex conjugate $z^* = x - jy = re^{j\theta}$. (*c*) Multiplication of z by $\pm j$, equivalent to rotation through 90°.

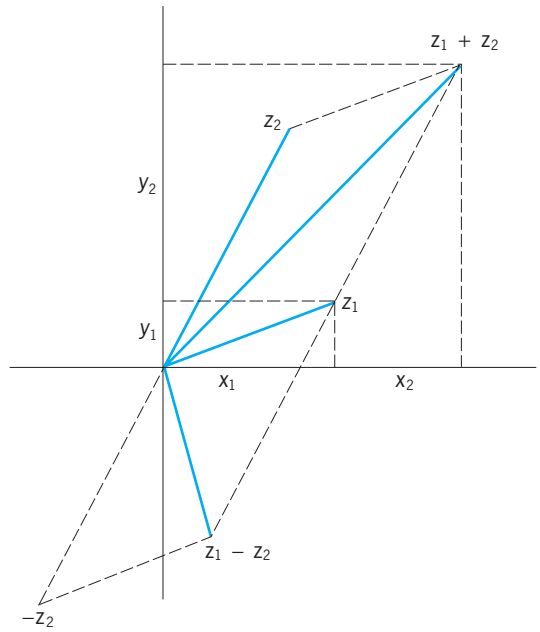

Fig. 5. Addition and subtraction of complex numbers.

Eq. (15) is valid since $j^2 = -1$, and the real and imaginary parts of the number interchange as shown.

$$jz = j(x + jy) = -y + jx \tag{15}$$

Similarly, multiplication by $-j$ corresponds to rotation through 90° in the opposite (clockwise) direction. From this it is apparent that $j = e^{j\pi/2}$.

In circuit analysis it is necessary to combine complex numbers in various ways. The rules of complex algebra for combining the numbers $z_1 = x_1 + jy_1$ and $z_2 = z_2 + jy_2$ are as follows.

The rule for addition is given by Eq. (16). This can be interpreted geometrically by saying that $z_1 + z_2$ is represented by the vector sum of the two vectors Oz_1 and Oz_2, as shown in **Fig. 5**.

$$z_1 + z_2 = (x_1 + x_2) + j(y_1 + y_2) \tag{16}$$

The rule for subtraction is given by Eq. (17). Similarly, $z_1 - z_2$ can be represented as the vector sum of the two vectors Oz_1 and $O(-z_2)$, as in Fig. 5.

$$z_1 - z_2 = (x_1 - x_2) + j(y_1 - y_2) \tag{17}$$

The rule for multiplication is given by Eq. (18), and that for division by Eq. (19). The process

$$\begin{aligned} z_1 z_2 &= (x_1 + jy_1)(x_2 + jy_2) \\ &= (x_1x_2 - y_1y_2) + j(x_1y_2 + x_2y_1) \\ &= r_1r_2e^{j(\theta_1+\theta_2)} \end{aligned} \tag{18}$$

$$\begin{aligned} \frac{z_1}{z_2} &= \frac{x_1 + jy_1}{x_2 + jy_2} \\ &= \frac{(x_1 + jy_1)(x_2 - jy_2)}{(x_2 + jy_2)(x_2 - jy_2)} \\ &= \frac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2} + j\frac{x_2y_1 - x_1y_2}{x_2^2 + y_2^2} \\ &= (r_1/r_2)e^{j(\theta_1-\theta_2)} \end{aligned} \tag{19}$$

illustrated in Eq. (19) for converting the quotient of two complex numbers into a number of form $a + jb$ by multiplying both numerator and denominator by the complex conjugate of the denominator is known as rationalization.

A comparison of Figs. 3*a* and 4*a* shows that alternating waveforms can be represented by the real or imaginary parts of complex numbers of argument ωt, as in Eqs. (20).

$$\begin{aligned} A_0 \cos \omega t &= A_0 \text{ Re } e^{j\omega t} \\ A_0 \sin \omega t &= A_0 \text{ Im } e^{j\omega t} \end{aligned} \tag{20}$$

A sinusoid of arbitrary phase may be written as in Eq. (21), where B is given by Eq. (22) [by De

$$\begin{aligned} A_0 \cos(\omega t + \theta) &= {}^1/_2 A_0 e^{j(\omega t+\theta)} + {}^1/_2 A_0 e^{-j(\omega t+\theta)} \\ &= Be^{j\omega t} + B^* e^{-j\omega t} \end{aligned} \tag{21}$$

$$B = {}^1/_2 A_0 e^{j\theta} \tag{22}$$

Moivre's theorem]. B contains all the information needed in circuit analysis, since A_0 and θ can be calculated from B, and the frequency is given by ω. It is therefore possible to formulate a circuit analysis in which the driving waveforms are of the form given by Eqs. (23), where V and I are complex quantities.

$$\begin{aligned} v &= Ve^{j\omega t} = V_0 e^{j\theta} e^{j\omega t} \\ i &= Ie^{j\omega t} = I_0 e^{j\theta} e^{j\omega t} \end{aligned} \tag{23}$$

The circuit parameters are also in complex form, and from these a response in similar form can be calculated which can be directly interpreted as a sinusoidal waveform.

Circuit elements. With constant currents, the only circuit element of significance is resistance. Voltage, current, and resistance are related by Ohm's law: circuit theory is needed to calculate the effective resistance of various series and parallel combinations of individual resistors. With alternating currents, inductance and capacitance have also to be taken into account. In these elements the rate of change of current or voltage is important. The relation between alternating current and voltage in the various types of circuit elements will be discussed. For purposes of theory, these are regarded as ideal elements of one kind only. In practice it is impossible to make an inductor, for example, without some residual resistance, but this will be ignored in the discussion below.

Resistance. As in the dc case, the instantaneous voltage v across a resistance R and the current i through it are related by Ohm's law, Eq. (24). If v is given by Eq. (25), then i is given by Eq. (26). Current and

$$v = iR \tag{24}$$

$$v = V_0 \cos \omega t \tag{25}$$

$$i = \frac{V_0}{R} \cos \omega t \tag{26}$$

voltage are in phase and can be represented in a phasor diagram as in **Fig. 6**. *See* ELECTRICAL RESISTANCE; RESISTOR.

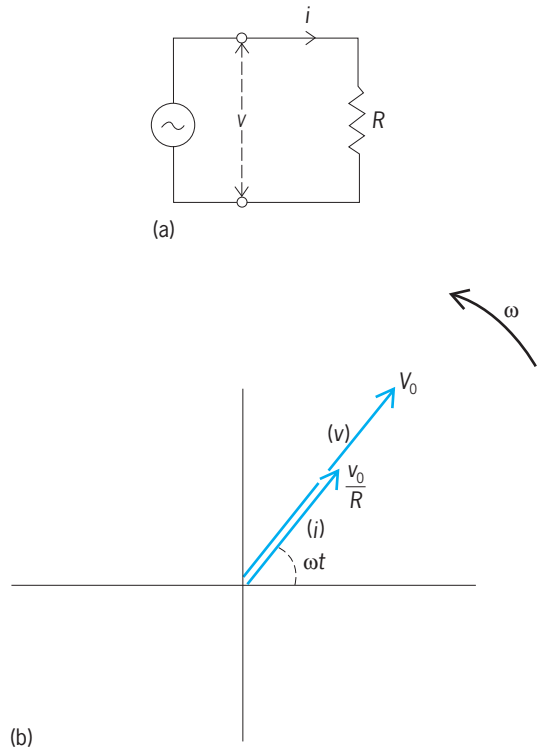

Fig. 6. Resistance. (*a*) Circuit diagram of resistor and source of alternating voltage. (*b*) Phasor diagram of voltage and current through a resistance.

Self-inductance. An inductor typically takes the form of a coil of wire, with or without a core of high-permeability magnetic material depending on the value of inductance required and the frequency of operation. The essential property of a single, or self-, inductance (so called to distinguish it from mutual inductance involving two coils, discussed below) is that the voltage across its terminals is related to the rate of change of current through it by Eq. (27). The

$$v = L\frac{di}{dt} \tag{27}$$

inductance L is measured in henrys when the voltage and current are in volts and amperes.

If the current is alternating and given by Eq. (28), the voltage is then given by Eq. (29). Thus the ratio

$$i = I_0 \cos \omega t \tag{28}$$

$$v = LI_0(-\omega \sin \omega t) = \omega L I_0 \cos\left(\omega t + \frac{\pi}{2}\right) \tag{29}$$

of the amplitudes of voltage and current is ωL and the voltage waveform is in advance of the current by a phase shift of $\pi/2$, or 90° (**Fig. 7**).

If the current is instead represented by a complex quantity as in Eq. (30), then Eq. (31) is obtained.

$$i = Ie^{j\omega t} \tag{30}$$

$$v = j\omega L I e^{j\omega t} \tag{31}$$

This represents the same result, since multiplying by j is equivalent to a positive (anticlockwise) rotation through a right angle. *See* INDUCTANCE; INDUCTOR.

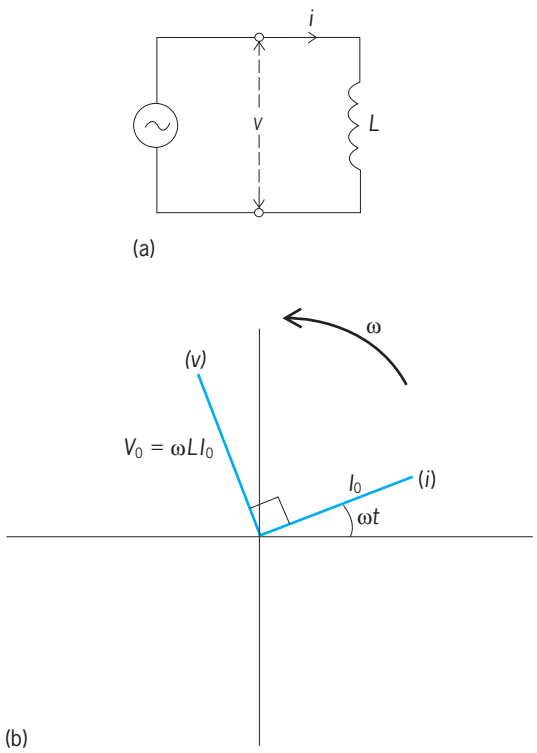

Fig. 7. Inductance. (*a*) Circuit diagram of inductor and source of alternating voltage. (*b*) Phasor diagram of voltage and current through an inductance.

Capacitance. An inductor stores energy in the magnetic field associated with an electric current. A capacitor stores energy in the electric field produced by the separation of charge in a dielectric medium. Typically a capacitor consists of two conducting plates separated by a dielectric. The current flowing in is related to the rate of change of voltage across the capacitor by Eq. (32), where C is the capacitance in farads.

$$i = C\frac{dv}{dt} \tag{32}$$

If the voltage is given by Eq. (33), then the current is given by Eq. (34). So in this case the ratio of volt-

$$v = V_0 \cos \omega t \tag{33}$$

$$i = CV_0(-\omega \sin \omega t) = \omega C V_0 \cos\left(\omega t + \frac{\pi}{2}\right) \tag{34}$$

age to current amplitudes is $1/\omega C$, and the voltage waveform lags behind the current by $\pi/2$ (**Fig. 8**).

In complex form, if the voltage is given by Eq. (35), the current is given by Eq. (36), representing the

$$v = Ve^{j\omega t} \tag{35}$$

$$i = j\omega CVe^{j\omega t} \tag{36}$$

same result. *See* CAPACITANCE; CAPACITOR.

Mutual inductance. If two coils are linked by a common magnetic flux, the voltage v_2 across the second coil is related to the rate of change of current i_1 in the first coil by Eq. (37), where M is the mutual inductance, measured in henrys.

$$v_2 = M\frac{di_1}{dt} \tag{37}$$

If i_1 is given by Eq. (38), then v_2 is given by Eq. (39), and the voltage is 90° out of phase with the primary current.

$$i_1 = I_1 e^{j\omega t} \tag{38}$$

$$v_2 = j\omega L I_1 e^{j\omega t} \tag{39}$$

For further discussion of mutual inductance. *See* COUPLED CIRCUITS.

Circuit analysis. The main problem of conventional ac circuit theory is the calculation of the relationships between currents and voltages in circuits containing various combinations of the circuit elements described above.

In order to analyze any physical electric circuit, it is necessary to construct a mathematical model. Such a model is an idealization of the real circuit, but the behavior of the model may be made to approximate that of the real circuit sufficiently closely so that analysis of the model yields results which may be applied with confidence to the real circuit. In the vast majority of cases the circuits to be analyzed are linear. In a linear circuit the most important property is that if the amplitude of the source voltages or currents is changed, then the amplitudes of all other currents and voltages are changed in the same proportion. In the case of sinusoidal currents and voltages every current and voltage in the circuit has

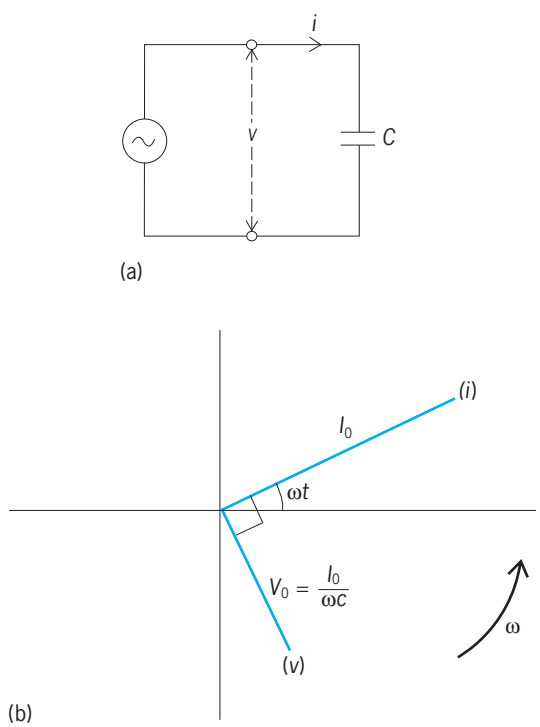

Fig. 8. Capacitance. (*a*) Circuit diagram of capacitor and source of alternating voltage. (*b*) Phasor diagram of voltage and current through a capacitance.

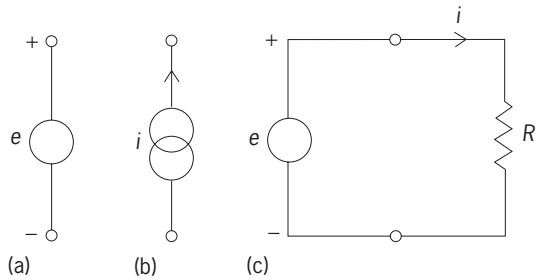

Fig. 9. Sources and signs. (*a*) Voltage source. (*b*) Current source. (*c*) Sign convention.

the same frequency, ω. In the rest of this article only linear circuits are discussed.

In an electric circuit the elements are connected together in some given fashion. Sources (or generators) of external current or voltage are applied to the circuit, and the problem is to calculate the resulting current and voltage associated with each circuit element. **Figure 9***a* and *b* illustrates the symbols used for voltage and current sources respectively. Although voltage and current change direction every half-cycle, it is necessary to have a sign convention to relate voltage and direction of current flow, compatible of course with that used in dc circuits. This is illustrated in Fig. 9*c* which is to be understood in the sense that, during the half-cycle in which the upper terminal of the source in the diagram carries the positive voltage, current flows through the resistor in the direction shown, from top to bottom of the diagram. In the next half-cycle both voltage and current reverse. The power dissipated in the resistor is vi, so the convention ensures that the power is positive in each half-cycle. *See* GENERATOR.

The analysis of any circuit is based on Kirchhoff's laws, which state that:

1. The sum of the instantaneous currents entering the junction of two or more circuit elements is zero.
2. The sum of the instantaneous voltages around any closed loop formed by two or more circuit elements is zero.

In both cases, proper account must be taken of the convention regarding positive and negative quantities. *See* KIRCHHOFF'S LAWS OF ELECTRIC CIRCUITS.

Analysis in the time domain. Consider the simple circuit shown in **Fig. 10**, which consists of a voltage source, a resistor, and an inductor connected in series. (Components are in series when the same current flows through each.) Kirchhoff's law indicates that the sum of the voltages around the loop is zero; that is, Eq. (40) holds. Now the voltage source e

$$v_R + v_L - e = 0 \tag{40}$$

is $E_0 \sin(\omega t + \theta)$, and the (unknown) current is assumed to be $I_0 \sin(\omega t + \beta)$, so that Eq. (41) is valid, and must be solved to find I_0 and β.

$$RI_0 \sin(\omega t + \beta) + L\omega I_0 \cos(\omega t + \beta) - E_0 \sin(\omega t + \theta) = 0 \tag{41}$$

When $\omega t + \theta = 0$, Eq. (41) reduces to Eq. (42) or Eq. (43).

$$R \sin(\beta - \theta) + L\omega \cos(\beta - \theta) = 0 \tag{42}$$

$$\tan(\beta - \theta) = -\frac{L\omega}{R} \tag{43}$$

When $\omega t + \theta = \pi/2$, Eq. (41) reduces to Eq. (44) or (45). Substituting from Eq. (43) for $\sin(\beta - \theta)$ and $\cos(\beta - \theta)$ yields Eqs. (46) or (47).

$$RI_0 \sin\left(\frac{\pi}{2} + \beta - \theta\right) + L\omega I_0 \cos\left(\frac{\pi}{2} + \beta - \theta\right) = E_0 \tag{44}$$

$$I_0[R\cos(\beta - \theta) - L\omega \sin(\beta - \theta)] = E_0 \tag{45}$$

$$I_0\left(\frac{R^2}{\sqrt{R^2 + L^2\omega^2}} + \frac{L^2\omega^2}{\sqrt{R^2 + L^2\omega^2}}\right) = E_0 \tag{46}$$

$$I_0 = \frac{E_0}{\sqrt{R^2 + L^2\omega^2}} \tag{47}$$

Hence the voltages across the resistor and the inductor are given by Eqs. (48) and (49). The phasor diagram for this circuit is given in Fig. 3*b*. The voltages across the resistor and the inductor are $\pi/2$

$$v_R = \frac{E_0 R}{\sqrt{R^2 + \omega^2 L^2}} \cdot \sin\left(\omega t + \theta - \tan^{-1}\frac{\omega L}{R}\right) \tag{48}$$

$$v_L = \frac{E_0 \omega L}{\sqrt{R^2 + \omega^2 L^2}} \cdot \sin\left(\omega t + \theta - \tan^{-1}\frac{\omega L}{R} + \frac{\pi}{2}\right) \tag{49}$$

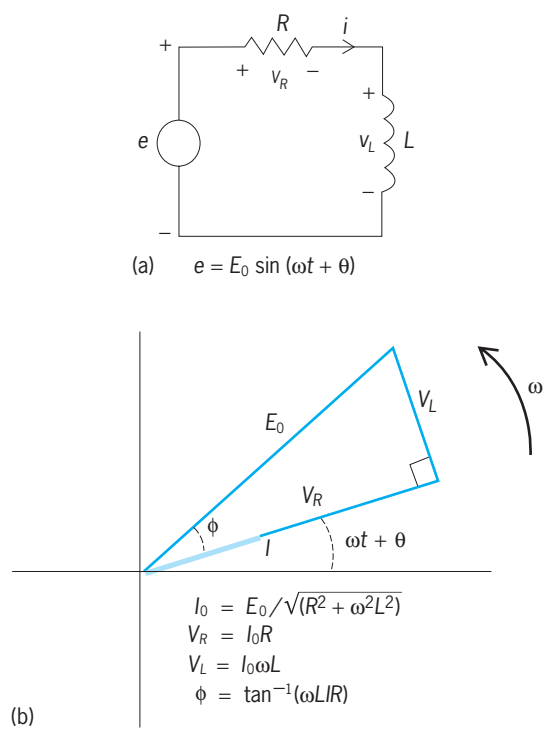

Fig. 10. Circuit with resistance and inductance. (*a*) Circuit diagram. (*b*) Phasor diagram.

radians out of phase, and the voltage across the circuit leads the current through it by the phase angle $\tan^{-1}(\omega L/R)$.

The analysis of any other circuit proceeds in the same way. The necessary equations are written by using Kirchhoff's laws and the relationships of Eqs. (24), (29), and (34) as appropriate. The solutions for the unknown currents or voltages are sinusoids with the same frequency as the applied voltages and currents but unknown amplitudes and phase angles. By evaluating the equations at particular values of time, as in the above examples, the unknown amplitudes and phase angles are then determined. The phase angles are found as differences relative to the phase angle of the applied generator. In the above example, $\beta - \theta$ is found, but not β as such. This is because one phase angle (here θ) must be taken as a reference angle, and the only important quantity is the value of other phase angles relative to this. Another way of looking at this is to note that the choice of origin on the time axis is quite arbitrary since all voltages and currents extend from $t = -\infty$ to $t = \infty$, which is an idealization of the real situation where the generator must start at some definite time. The choice of the reference phase angle corresponds to a particular choice of time origin.

While this method can, in principle, be used to analyze any circuit, it becomes extremely cumbersome for a circuit of even moderate complexity. Simplification of the analysis by the use of complex algebra will be discussed, but first the dissipation of power in an ac circuit must be considered.

Power. Consider the current and voltage associated with an arbitrary circuit (**Fig. 11**). The instantaneous power p is equal to vi. Let the voltage and current be given by Eqs. (50) and (51). Then the instantaneous

$$v = V_0 \sin(\omega t + \theta) \tag{50}$$

$$i = I_0 \sin(\omega t + \beta) \tag{51}$$

power is given by Eq. (52), which is the sum of a

$$\begin{aligned} p &= V_0 I_0 \sin(\omega t + \theta)\sin(\omega t + \theta) \\ &= {}^1/_2 V_0 I_0 [\cos(2\omega t + \theta + \beta) + \cos(\beta - \theta)] \end{aligned} \tag{52}$$

sinusoid at radian frequency 2ω and a constant independent of t. The instantaneous power is generally of much less interest than the average power P. The average value of a sinusoid is zero so that the average power is given by Eq. (53), where $\hat{V} = V/\sqrt{2}$ and

$$P = {}^1/_2 V_0 I_0 \cos(\beta - \theta) = \hat{V}\hat{I}\cos(\beta - \theta) \tag{53}$$

$\hat{I} = I/\sqrt{2}$ are called the root-mean-square (rms) values of v and i respectively, and $\cos(\beta - \theta)$ is called the power factor of the circuit. Thus the power depends not only on the amplitudes of the current and voltage but also on the phase angle between them. This is illustrated by the cases of the resistor, given by Eqs. (54); the inductor, given by Eqs. (55); and the capacitor, given by Eqs. (56). For both the inductor and the capacitor the average power is zero.

$$\left.\begin{aligned} V_0 &= RI_0,\ \beta - \theta = 0 \\ P &= \hat{V}\hat{I} = R\hat{I}^2 = \hat{V}^2/R \end{aligned}\right\} \text{resistor} \tag{54}$$

$$\left.\begin{aligned} V_0 &= L\omega I_0,\ \beta - \theta = -\tfrac{\pi}{2} \\ P &= 0 \end{aligned}\right\} \text{inductor} \tag{55}$$

$$\left.\begin{aligned} I_0 &= C\omega V_0,\ \beta - \theta = \tfrac{\pi}{2} \\ P &= 0 \end{aligned}\right\} \text{capacitor} \tag{56}$$

It follows therefore that in an ac circuit, power is dissipated only in resistive elements, not in inductors or capacitors. The voltage across these latter elements is 90° out of phase with the current through them. The total voltage across a series circuit is made up of the vector sum of the phasors representing the individual element voltages. It is common to speak of in-phase components of the voltage (having the same phase as the current) and out-of-phase or quadrature components with a 90° phase difference. Because no power is dissipated in these components, the current through an inductor or capacitor is sometimes referred to as a wattless current.

In electric power systems the values quoted for current and voltage are usually the rms values since the power available is related directly to these quantities. However, it must be borne in mind that they must be multipled by 2 to obtain the actual amplitudes of current and voltage.

Analysis using complex impedances. The method of time domain analysis described above is quite general, and any circuit may be solved with those techniques. However, the method is cumbersome and does not give much insight into circuit behavior. Further, and perhaps more important, there are a wide variety of circuit theorems developed for use in dc circuit analysis which can both simplify analysis and provide considerable insight, but which cannot be used in ac circuits if one is restricted to the methods of time domain analysis. These considerations led to the development of a method of analysis based on complex notation which makes these techniques available in ac circuit analysis.

In an ac circuit every current and voltage is a sinusoid of the same frequency ω, and each is characterized by two quantities, namely the amplitude and the phase angle relative to some reference angle. Thus two independent qualities are needed to describe each current and voltage. As discussed above, an alternating current or voltage can be represented by

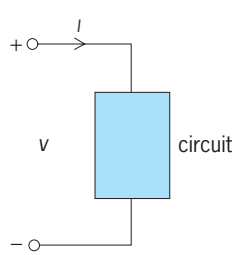

Fig. 11. Generalized circuit.

an expression of the form $Ie^{j\omega t}$ or $Ve^{j\omega t}$, and this representation can be applied to circuit analysis.

In the previous example of Fig. 3, Eq. (40) leads to Eq. (57) or (58), from which the modulus and argument of I are given by Eqs. (59) and (60). Thus if the voltage source is given by Eq. (61), then the current is given by Eq. (62), which is the result previously

$$RIe^{j\omega t} + jL\omega Ie^{j\omega t} - Ee^{j\omega t} = 0 \tag{57}$$

$$I = \frac{E}{R + jL\omega} \tag{58}$$

$$|I| = \frac{E}{\sqrt{R^2 + L^2\omega^2}} \tag{59}$$

$$\text{Argument } I = \text{Argument } E - \tan^{-1}\frac{L\omega}{R} = \beta \tag{60}$$

$$e = E_0 \sin(\omega t + \theta) \tag{61}$$

$$i = |I| \sin\left(\omega t + \theta - \tan^{-1}\frac{L\omega}{R}\right) \tag{62}$$

found. However, now the analysis is very simple by using the complex quantities.

In the case of the resistor, inductor, and capacitor and in the example, equations are obtained where every term is multiplied by $e^{j\omega t}$, and this is true in general. Thus, this multiplier could just as well be omitted in every case and equations could be written involving only the complex quantities V and I. Once V and I are found, the answer for sinusoidal generators can be found immediately. Thus, for a resistor, $V = RI$; for an inductor, $V = j\omega LI$; and for a capacitor, $V = I/j\,C\omega$. For the complex voltages and currents corresponding to a frequency ω, an inductor behaves as if it were a resistor of value $j\omega L$, and a capacitor behaves as if it were a resistor of value $1/j\omega C$. The problem of analysis is thereby reduced to the analysis of dc circuits, except that all currents and voltages are complex numbers and some "resistors" are real while others are imaginary. *See* NETWORK THEORY.

When these complex currents and voltages are used, the ratio V/I in a circuit is called the (complex) impedance of the circuit (usually written Z; **Fig. 12**), and the ratio I/V is called the (complex) admittance of the circuit (usually written Y). The units of impedance are ohms, and of admittance are siemens (mhos). The impedance and admittance are expressed in terms of their real and imaginary parts by Eqs. (63) and (64) and are related to each other by Eq. (65) or (66). Similarly, Eq. (67) is valid.

$$\frac{V}{I} = Z = R + jX \tag{63}$$

$$\frac{I}{V} = Y = G + jB \tag{64}$$

$$Y = \frac{1}{Z} \tag{65}$$

$$G + jB = \frac{1}{R + jX} = \frac{R}{R^2 + X^2} - j\frac{X}{R^2 + X^2} \tag{66}$$

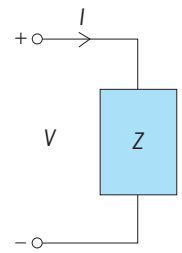

Fig. 12. General impedance.

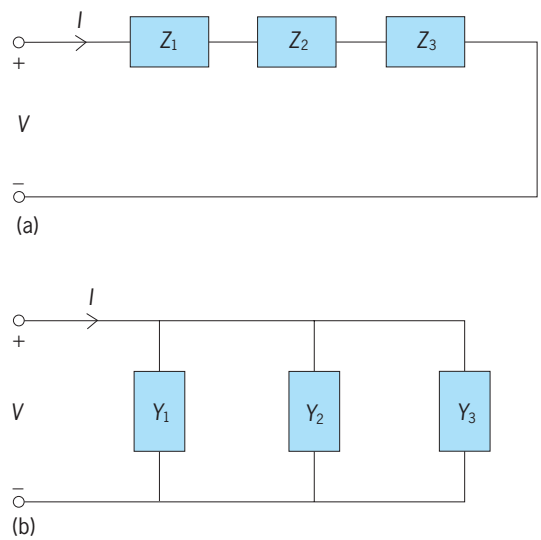

Fig. 13. Combination of impedances and admittances. (*a*) Impedances in series. (*b*) Admittances in parallel.

$$R + jX = \frac{1}{G + jB} = \frac{G}{G^2 + B^2} - j\frac{B}{G^2 + B^2} \tag{67}$$

Impedances, like resistances, can be combined in series and parallel. **Figure 13*a*** shows three impedances in series. By definition the voltage across each impedance is ZI, and application of Kirchhoff's law to the loop yields Eq. (68) or (69), so the equivalent impedance Z_{eq} is the sum of the impedances in series.

$$Z_1 I + Z_2 I + Z_3 I - V = 0 \tag{68}$$

$$\frac{V}{I} = Z_1 + Z_2 + Z_3 = Z_{eq} \tag{69}$$

By definition the current in each admittance is YV, and application of Kirchhoff's law yields Eq. (70) or Eq. (71). Thus the reciprocal of the equivalent impedance Z_{eq} is the sum of the admittances in parallel.

$$Y_1 V + Y_2 V + Y_3 V - I = 0 \tag{70}$$

$$\frac{V}{I} = \frac{1}{Y_1 + Y_2 + Y_3} = Z_{eq} \tag{71}$$

The real and imaginary components of impedance and admittance are given special names. For impedance, $Z = R + jX$, the real part R is resistance

in the usual way; the imaginary part X (ωL for an inductor or $-1/\omega C$ for a capacitor) is reactance.

For admittance, $Y = G + jB$, the real part G is called conductance, and the imaginary part B, susceptance.

In a series circuit having resistance R and reactance X, the conductance is not simply the reciprocal of the resistance but, from Eq. (66), a quantity involving both the resistance and the reactance, namely $R/(R^2 + X^2)$.

Occasionally, it is useful to be able to refer to either impedance or admittance without specifying which in particular is meant. The general term immittance is used to cover both. For obvious reasons it is not possible to specify units for it. *See* ADMITTANCE; CONDUCTANCE; ELECTRICAL IMPEDANCE; IMMITTANCE; REACTANCE; SUSCEPTANCE.

Power dissipated. As discussed above, the power entering a circuit, such as Fig. 11, is given by Eq. (72),

$$P = {}^1/_2 V_0 I_0 \cos(\theta - \beta) \tag{72}$$

where V_0 and I_0 are the amplitudes of the voltage and current sinusoids, and $\theta - \beta$ is the angular phase difference between them. This power can also be expressed in terms of the complex quantities V and I.

If V and I are written as in Eqs. (73) and (74), then Eqs. (75) and (76) are valid. Thus, the power is given by Eq. (77).

$$V = V_0 e^{j\theta} \tag{73}$$

$$I = I_0 e^{j\beta} \tag{74}$$

$$\begin{aligned} VI^* &= V_0 I_0 e^{j(\theta-\beta)} \\ &= V_0 I_0 [\cos(\theta - \beta) + j\sin(\theta - \beta)] \end{aligned} \tag{75}$$

$$V^* I = V_0 I_0 [\cos(\theta - \beta) - j\sin(\theta - \beta)] \tag{76}$$

$$P = {}^1/_4 (VI^* + V^* I) = {}^1/_2 V_0 I_0 \cos(\theta - \beta) \tag{77}$$

If $V = ZI$, then the power is also given by Eq. (78), and if $I = YV$, then it is also given by Eq. (79).

$$P = {}^1/_4 (ZII^* + Z^* II^*) = {}^1/_4 I_0^2 (Z + Z^*) \tag{78}$$

$$P = {}^1/_4 (YVV^* + Y^* VV^*) = {}^1/_4 V_0^2 (Y + Y^*) \tag{79}$$

These equations give alternative ways of calculating the power from the voltage or current and circuit parameters.

Resonant circuit. Consider the circuit of **Fig. 14*a*** with a voltage generator driving a capacitor in parallel with an inductor and resistance in series. The generator produces a voltage $e = E_0 \sin(\omega t + \theta)$. Suppose it is necessary to calculate the currents in the capacitor, resistor, and inductor, and the power dissipated in the circuit.

The circuit can be redrawn as in Fig. 14*b* Y_1 as the admittance of the capacitor and $Y_2 = 1/Z_2$, where Z_2 is the impedance of the resistor and inductor in series. Now Z_2 is given by Eq. (80), so Y_2 is given by Eq. (81) and Y_1 is given by Eq. (82). The equivalent impedance of the circuit, Z_{eq}, is given by Eq. (83), and

$$Z_2 = R + jL\omega \tag{80}$$

$$Y_2 = \frac{1}{R + jL\omega} = \frac{R}{R^2 + L^2\omega^2} - j\frac{L\omega}{R^2 + L^2\omega^2} \tag{81}$$

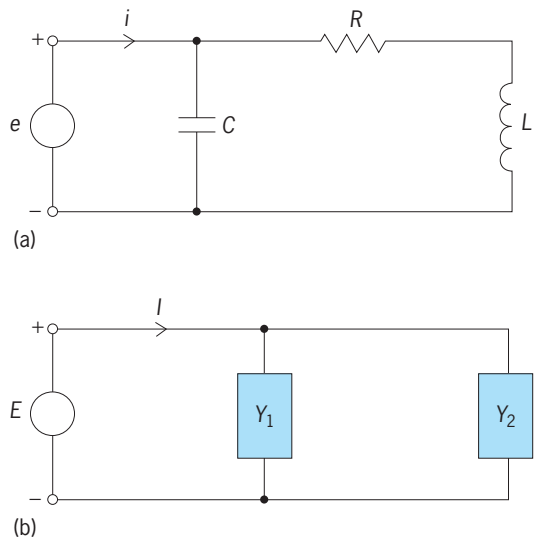

Fig. 14. Resonant circuit. (*a*) Diagram showing components. (*b*) Diagram showing admittances.

$$Y_1 = jC\omega \tag{82}$$

$$Z_{eq} = \frac{1}{Y_1 + Y_2} \tag{83}$$

the generator voltage is related to the total current by Eq. (84) or Eq. (85) so that the total current is given by Eq. (86).

$$E = Z_{eq} I \tag{84}$$

$$I = E/Z_{eq} \tag{85}$$

$$\begin{aligned} I &= E(Y_1 + Y_2) \\ &= E\left[\frac{R}{R^2 + L^2\omega^2} + j\left(C\omega - \frac{L\omega}{R^2 + L^2\omega^2}\right)\right] \end{aligned} \tag{86}$$

The current through the capacitor is $Y_1 E = j\omega CE$, while the current through the series combination of the resistor and inductor is $Y_2 E$, given by Eq. (87).

$$Y_2 E = E\left(\frac{R}{R^2 + L^2\omega^2} - j\frac{L\omega}{R^2 + L^2\omega^2}\right) \tag{87}$$

The power dissipated in the circuit is found, by using Eq. (79), to be given by Eq. (88).

$$\begin{aligned} P &= {}^1/_4 E_0^2 (Y_{eq} + Y_{eq}^*) \\ &= {}^1/_4 E_0^2 (Y_1 + Y_2 + Y_1^* + Y_2^*) \\ &= {}^1/_2 E_0^2 \frac{R}{R^2 + L^2\omega^2} \end{aligned} \tag{88}$$

This circuit exhibits the phenomenon of resonance. Suppose the frequency ω is varied. Then in Eq. (86) the imaginary part becomes zero when Eq. (89) is satisfied, which gives Eq. (90) for ω^2. This

$$C\omega = \frac{L\omega}{R^2 + L^2\omega^2} \tag{89}$$

$$\omega^2 = \frac{1}{LC} - \left(\frac{R}{L}\right)^2 = \omega_0^2 \tag{90}$$

is called the resonant frequency, and at this value of ω the power dissipated is given by Eq. (91). The quan-

$$P_0 = {}^1/_2 E_0^2 \frac{RC}{L} \tag{91}$$

tity $L\omega_0/R$ is called the quality factor or Q-factor of the circuit and is given by Eq. (92). Since Q is often very large, it is usually approximated by Eq. (93).

$$Q = \sqrt{\frac{L}{R^2C} - 1} \tag{92}$$

$$Q = \frac{1}{R}\sqrt{\frac{L}{C}} \tag{93}$$

Having found these various complex voltages and currents, the corresponding sinusoidal quantities could be written immediately, but this is rarely done in practice since all the information is already contained in the complex quantities. *See* RESONANCE (ALTERNATING-CURRENT CIRCUITS).

Practical circuit elements. Alternating-current circuit theory is, of course, applicable at higher, radio frequencies so long as the circuit elements can be considered as "lumped," that is, so long as their dimensions are small compared with a wavelength of the current and its phase is the same at all points in the component. But particularly at the higher frequencies, the idealized representations of circuit elements which were given earlier need to be replaced by representations which more accurately reflect the properties of actual components.

An inductor, for example, in the form of a wire-wound coil, will have resistance as well as inductance, so it should be shown as a resistance and inductance in series, with impedance $R + j\omega L$. For a well-designed coil, R is much smaller than ωL at the frequency of use. There will also be capacitance between the turns of the coil which becomes important at the higher frequencies. As discussed above, this is a resonant circuit, and there is for any coil a frequency at which it is self-resonant; above this frequency its effective impedance is capacitative rather than inductive. *See* CORE LOSS; SKIN EFFECT (ELECTRICITY).

A capacitor will generally suffer some power loss in the dielectric medium, which can be represented by a parallel high resistance, and the connecting leads will have low resistance and inductance which may be significant at very high frequencies. *See* PERMITTIVITY.

Other stray impedances may also have to be taken into account. At high frequencies it is particularly important to remember that even lengths of connecting wire may have significant inductance and capacitance to ground.

Microwave circuits. At microwave frequencies the assumption of lumped components whose size is small compared with the wavelength breaks down. The basic principles of ac circuit theory can still be applied, but it is necessary to assume that the properties of resistance, inductance, and capacitance are no longer localized but are distributed throughout the circuits, and to take account of the finite time of travel of a wave from one part of the circuit to another. The concept of impedance then has to be generalized. *See* MICROWAVE; TRANSMISSION LINES.

Active circuits. In addition to the passive circuits discussed in this article, circuit theory can be extended to cover the cases of amplification and feedback. Again the concept of impedance may have to be extended: it is possible, for example, to produce circuits whose effective impedance is that of a negative capacitor, or which do not correspond to any possible combination of simple circuit elements. *See* AMPLIFIER; CONTROL SYSTEMS; FEEDBACK CIRCUIT.

Nonsinusoidal waveforms. The above analysis of circuits has been entirely in terms of the response to single-frequency sinusoidal waveforms. There are many cases where it is necessary to deal with other waveforms. Two methods are available:

1. When it is necessary to find the response to a repetitive waveform or perhaps a single pulse whose Fourier transform can easily be calculated as $F(\omega)$, the impedance or transfer function (circuit property relating input and output voltage or current) can be calculated in the usual way as $Z(\omega)$, and then the inverse transform of $F(\omega) \cdot Z(\omega)$ can be extracted to give the response waveform as a function of time. *See* FOURIER SERIES AND TRANSFORMS.

2. An alternative method, particularly useful in dealing with transient waveforms, is (for example) to take the Laplace transform of the input waveform, $L(s)$, and to multiply it by $Z'(s)$, given by the impedance or transfer function calculated in the usual way but with $j\omega$ replaced by s. The output is then the inverse Laplace transform of $L(s) \cdot Z'(s)$. *See* LAPLACE TRANSFORM. J. O. Scanlan; A. E. Bailey

Bibliography. C. K. Alexander and M. N. O. Sadiku, *Fundamentals of Electric Circuits*, 3d ed., McGraw-Hill, 2007; J. W. Nilsson and S. A. Riedel, *Electric Circuits w/PSpice*, 7th ed., Prentice Hall, 2005.

Alternating-current generator

A machine that converts mechanical power into alternating-current electric power. Almost all electric power is produced by alternating-current (ac) generators that are driven by rotating prime movers. Most of the prime movers are steam turbines whose thermal energy comes from either fossil or nuclear fuel. Combustion turbines are often used for the smaller units and in cases where gas or oil is the available fuel. Where water power is available from dams, hydroelectric ac generators are powered by hydraulic turbines. Small sites may also use diesel or gasoline engines to drive the generator, but these units are usually used only for standby generation or to provide electric power in remote areas. *See* DIESEL ENGINE; GAS TURBINE; HYDRAULIC TURBINE; HYDROELECTRIC GENERATOR; INTERNAL COMBUSTION ENGINE; STEAM ELECTRIC GENERATOR; STEAM TURBINE.

Alternating-current generators are used because ac power can easily be stepped up in voltage, by using transformers, for more efficient transmission of power over long distances and in larger amounts.

Similarly, transformers step the voltage down again at the utilization site to safer and more convenient levels. *See* DIRECT-CURRENT GENERATOR; ELECTRIC POWER SYSTEMS; TRANSFORMER.

Principles of operation. Most ac generators are synchronous machines, that is, the rotor is driven at a speed that is exactly related to the rated frequency of the ac network. Generators of this type have a stationary armature with three windings that are displaced at regular intervals around the machine to produce three-phase voltages. These machines also have a field winding that is attached to the rotor. This winding provides magnetic flux that crosses the air gap and links the stator coils to produce a voltage according to Faraday's law. The field winding is supplied with direct current, usually through slip rings. *See* ARMATURE; SLIP RINGS; WINDINGS IN ELECTRIC MACHINERY.

To understand the action of an ac generator, it is helpful to visualize a rotating magnetic flux-density wave in the air gap of the machine (**Fig. 1**). This wave links the stator winding, causing each coil to experience an alternating flux. This is the mechanism for inducing an alternating voltage.

It is usually assumed that the flux density in the air gap has a sinusoidal distribution, which may be written as Eq. (1), where θ is the angular position mea-

$$B = B_{\max} \cos \frac{p\theta}{2} = B_{\max} \cos \theta_e \qquad (1)$$

sured around the air gap of the machine in a certain direction. The angle θ_e is defined by Eq. (2), where p

$$\theta_e = \frac{p}{2}\theta \quad \text{electrical radians} \qquad (2)$$

is the number of poles. This angle is defined so that 360 electrical degrees corresponds to an angle that passes through flux defining a north pole and then a south pole. *See* ELECTRICAL DEGREE.

The total flux linking the coil is given by Eq. (3), where the differential area is given by Eq. (4), and

$$\phi_c = \int\int B dA \qquad (3)$$

$$dA = LrD\theta = \frac{2Lr}{p} d\theta_e \qquad (4)$$

L is the coil length, r is the radius of the air gap in the cylindrical geometry of the machine, and θ is the angle defined above. The generator shaft rotates at synchronous speed with velocity given by Eq. (5), where ω_e, defined by Eq. (6), is the speed in electri-

$$\omega_s = \frac{2\pi f}{p/2} = \frac{2}{p}\omega_e \quad \text{radians per second} \qquad (5)$$

$$\omega_e = 2\pi f \qquad (6)$$

cal radians per second and f is the system frequency.

The flux density may be regarded as a traveling wave, given by Eq. (7). Substituting the expressions

$$B(\theta, t) = B_{\max} \cos \frac{p}{2}(\theta - \omega_s t)$$
$$= B_{\max} \cos (\theta_e - \omega_e t) \qquad (7)$$

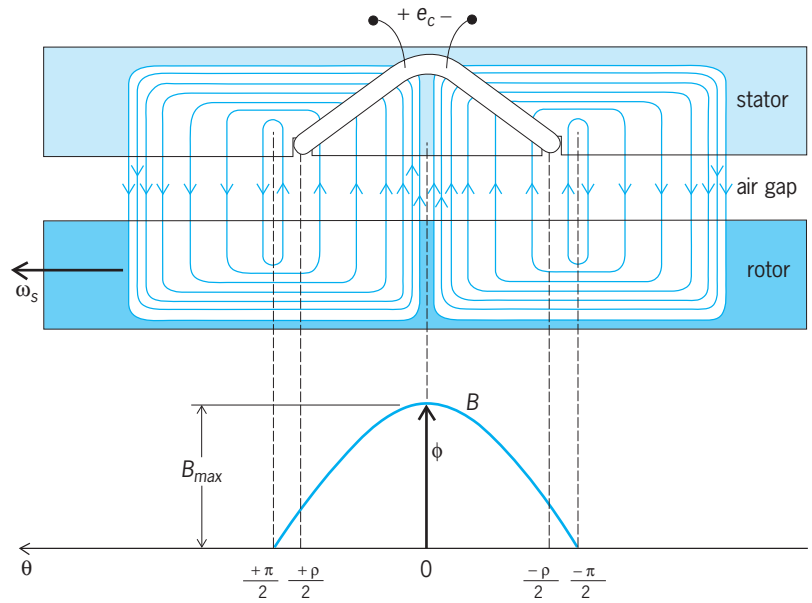

Fig. 1. End view of one coil of an alternating-current generator, linked by the moving airgap flux.

in Eqs. (4) and (7) for dA and B into the integrand of Eq. (3) and evaluating between the limits $\pm\rho/2$ yields Eq. (8). Then by using the identity of Eq. (9),

$$\phi_c = \int_{-\rho/2}^{\rho/2} \frac{2Lr}{p} [B_{\max} \cos (\theta_e - \omega_e t)] \, d\theta_e \qquad (8)$$

$$\cos (x - y) = \cos x \cos y - \sin x \sin y \qquad (9)$$

the integration is readily performed. The flux for one coil is found to be given by Eq. (10), where k_p and ϕ_p are defined by Eqs. (11) and (12).

$$\phi_c = k_p \phi_p \cos \omega_e t \qquad (10)$$

$$k_p = \text{ pitch factor } = \sin \frac{\rho}{2} \qquad (11)$$

$$\phi_p = \text{flux per pole} = \frac{4B_{\max} Lr}{p} \qquad (12)$$

The induced electromotive force (emf) for the coil is computed from Faraday's law, which states that the emf is equal to the rate of change of flux linkages; that is, Eq. (13) is valid where N_c is the number of turns in

$$e_c = -\frac{d\lambda_c}{dt} = -\frac{d(N_c \phi_c)}{dt} = \omega_e N_c k_p \phi_p \sin \omega_e t \qquad (13)$$

the coil. It is convenient to write the coil voltage as in Eq. (14), where E_c is the root-mean-square (rms)

$$e_c = \sqrt{2} E_c \sin \omega_e t \qquad (14)$$

value of the coil voltage. The total pitch of the coil ($\pi - \rho$) is less than one pole pitch (π). This difference has the effect of reducing harmonics more than it reduces the fundamental component of voltage. This reduction is expressed in terms of the pitch factor. *See* ELECTROMAGNETIC INDUCTION.

The voltage e_c is the induced voltage in only one coil. (Fig. 1). Usually, the voltages of a group of coils are added together by connecting the coil ends in

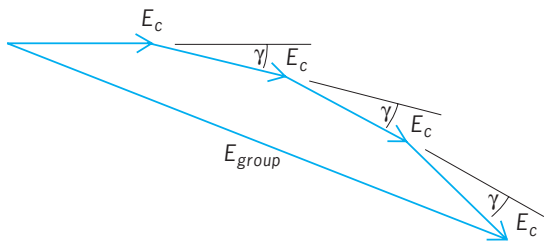

Fig. 2. Phasor diagram for the total root-mean-square emf of a group of coils, E_{group}.

series. The coils in the group are not all in the same slots, however, but are displaced by the slot pitch γ. Therefore, the voltage induced in the individual coils will be out of phase by this angle. This means that the addition of the voltages is not a simple arithmetic addition, but is usually performed as a phasor addition to compute the total rms emf of the group of coils, E_{group}. This addition is shown in **Fig. 2**, where the number of coils, n, in the group is assumed to be four. *See* ALTERNATING-CURRENT CIRCUIT THEORY.

From the geometry of Fig. 2, Eq. (15) may be com-

$$E_{group} = nE_c \frac{\sin \frac{n\gamma}{2}}{n \sin \frac{\gamma}{2}} = nE_c k_d \tag{15}$$

puted, where a new constant k_d is defined for convenience. Finally, the total phase voltage is composed of p groups in series and is thus given by Eq. (16),

$$E_{phase} = pE_{group} \tag{16}$$

where p is the number of poles. For steam-turbine-driven generators p is usually 2 or 4. Hydroelectric generators may have a much larger number of poles, depending on the shaft speed.

Three-phase generators produce three-phase voltages that are equal in magnitude but displaced in phase by 120 electrical degrees. The phasor diagram for the voltages of a three-phase ac generator is shown in **Fig. 3**.

Armature reaction. Armature reaction is a demagnetizing effect that limits the output of an ac generator. The magnetomotive forces (mmf's) of the coils combine to produce a rotating mmf that opposes the field flux. Thus, more excitation is required to maintain the flux magnitude as the armature current is increased. The mmf of one coil per pole is a square wave that has a sine-wave fundamental component with a maximum value given by Eq. (17), where I_c is the rms armature current in the

$$F_{max} = \sqrt{2}\frac{4}{\pi} N_c I_c k_p \tag{17}$$

coil. The mmf factor for all coils in one phase is $F_{max}k_d$.

To compute the resultant rotating mmf for a polyphase winding, the single alternating wave is divided into two equal components of one-half value each, rotating in opposite directions. For balanced phase currents, the forward components combine and the reverse components cancel. Thus, for three-phase windings, the resultant rotating wave has a maximum of $^3/_2$ the single phase value and is usually referred to as F_1, where the subscript 1 indicates the fundamental component of the sine wave. *See* ARMATURE REACTION.

Characteristics. Certain constants define the operation of the ac generator on the system to which it is connected. One important set of machine constants is the reactances of the windings. There are several different reactances that are derived to describe the machine behavior under different conditions. *See* REACTANCE.

The synchronous reactance x_d is the unsaturated steady-state reactance, including leakage and armature reaction, and is used to describe steady-state generator performance. The transient reactance x'_d is useful in determining how the machine responds to sudden changes, after the first few cycles. This reactance is often used to compute the fault current for circuit-breaker opening. The subtransient reactance x''_d determines the initial fault current, which is a current that the machine windings and connecting bus-work must be able to withstand mechanically.

The generator operates in synchronism with the system to which it is connected. As load is increased, the shaft angle moves ahead by the angle δ. The power delivered is given by Eq. (18), where the sub-

$$P = \frac{E_s E_g}{x_s + x_d} \sin \delta \tag{18}$$

script s stands for system quantities and the subscript g stands for generator quantities. The voltage E_g for steady-state conditions is the rated excitation voltage, and x_g is the synchronous reactance.

To calculate system swings due to sudden changes, x_d can be taken as the transient reactance and E_d as the effective voltage behind that reactance. This

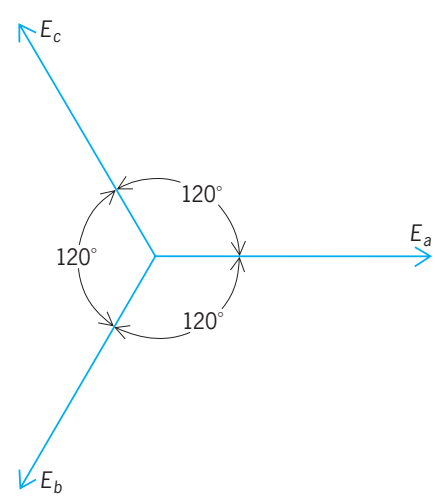

Fig. 3. Phasor diagram of the balanced three-phase generated voltages, E_a, E_b, and E_c, for a synchronous alternating-current generator.

voltage changes only slowly but can be corrected for the action of the excitation system, which affects the control of the voltage and plays an important role in the dynamic performance of the machine.

The calculation of dynamic swings during a short circuit can be performed step by step or by computer simulations, the latter method being the usual choice for systems containing more than one machine. The computation gives the change in power output of the machine and the resulting acceleration and angle, taking into account the moment of inertia of the turbine-generator combination.

Growth in size. The size of ac generators increased rapidly up to the mid-1970s. This increase resulted in considerable economy in terms of the cost per megawatt of machine rating and a concurrent improvement in turbine efficiency. The maximum ratings doubled about every 10 years for many years, which is about the same rate at which the total electrical load in North America was increasing. The largest units now manufactured are rated at about 1800 megavolt-amperes. This rapid growth in unit size was a major challenge to machine designers. Better rotor materials were required to withstand higher centrifugal forces associated with larger rotor diameters. Major changes in the method of cooling were required, resulting first in hydrogen cooling and later in water-cooled stator windings. These refinements were necessary to keep the efficiency of the generator at about 99%. *See* VOLT-AMPERE.

Other generator types. The foregoing discussion has dealt entirely with the class of generators called synchronous machines. These are by far the most common, but are not the only type of ac generators.

Induction generators have been used in remote applications where maintenance of the excitation system is a problem. These units are essentially like induction motors, but are driven by a prime mover at speeds slightly above synchronous speed, forcing the unit to generate power due to the reverse slip. Induction generator units draw reactive power from the system and are not as efficient as synchronous generators. *See* INDUCTION MOTOR.

High-frequency single-phase generators have been built as induction alternators, usually with twice as many stator poles (teeth) as rotor poles, and with a constant air-gap flux supplied from a homopolar field coil in the center of the machine, pushing flux into the stator at one end and out at the other. Their effectiveness is lower than that of the synchronous machine because the flux is a pulsating unidirectional field, rather than an alternating field. *See* ALTERNATING CURRENT; ELECTRIC ROTATING MACHINERY; GENERATOR.

Lee A. Kilgore; Paul M. Anderson

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; A. E. Fitzgerald, C. Kingsley, and S. D. Umans, *Electric Machinery*, 6th ed., 2003; G. McPherson and R. D. Laramore, *An Introduction to Electrical Machines and Transformers*, 2d ed., 1990.

Alternating-current motor

An electrical machine that converts alternating-current (ac) electric energy to mechanical energy. Alternating-current motors are widely used because of the general availability of ac electric power and because they can be readily built with a variety of characteristics and in a large range of sizes, from a few watts to many thousands of kilowatts. They can be broadly classified into three groups—induction motors, synchronous motors, and ac series motors:

- **Induction motor**
  - Single-phase
    - Split-phase
    - Capacitor-start
    - Capacitor-run
  - Polyphase
- **Synchronous motor**
  - Single-phase
    - Permanent-magnet (PM)
    - Reluctance
    - Hysteresis
  - Polyphase
    - Wound-field
    - Permanent-magnet (PM)
    - Reluctance
- **AC series or universal motor (single-phase)**

See ALTERNATING CURRENT; DIRECT-CURRENT MOTOR.

Induction motor. The most common type of ac motor, both in total number and in total power, is the induction motor. In larger sizes these machines employ a polyphase stator winding, which creates a rotating magnetic field when supplied with polyphase ac power. The speed of rotation depends upon the frequency of the supply and the number of magnetic poles created by the winding; thus, only a discrete number of speeds are possible with a fixed frequency supply.

Currents are induced in the closed coils of the rotor for any rotor speed different from the speed of the rotating field. The difference in speed is called the slip speed, and efficient energy conversion occurs only when the slip speed is small. These machines are, therefore, nearly constant-speed machines when operated from a constant-frequency supply. They are, however, routinely started from zero speed and accelerated through the inefficient high-slip-speed region to reach operating speed. *See* SLIP (ELECTRICITY).

Single-phase induction machines are generally inferior in conversion efficiency and have zero torque at stand-still (zero speed). Auxiliary starting means usually involving a second winding and some technique for creating a time phase difference in the second winding current must be employed. Capacitors are commonly used to create a large phase shift, although extra resistance in the starting winding (in the split-phase motor) can be used if reduced starting

torque is acceptable. Typically, the second winding is switched off by a rotation-speed-actuated switch, but in some low-starting-torque applications the second winding with a series capacitor is left connected to improve the running condition (in the capacitor-run motor). Single-phase machines are widely used in situations where only single-phase ac power is available, especially for appliances in homes and small commercial installations. *See* INDUCTION MOTOR.

Synchronous motor. In contrast to an induction motor, the rotor of a synchronous motor runs exactly at the rotating field speed and there are no induced rotor currents. Torque is produced by the interaction of the rotating field with a direct-current (dc) field created by injected dc rotor current or permanent magnets, or with a rotor magnetic structure that has an easy direction for magnetization (in the reluctance motor). Since for any frequency of excitation there is only one speed for synchronous torque, synchronous machines have no starting torque unless the frequency is variable. When the motor is used in fixed-frequency applications, an induction-machine winding is also placed on the rotor to allow starting as an induction motor and running as a synchronous motor. *See* SYNCHRONOUS MOTOR.

Large synchronous motors for industrial applications have dc windings on the rotor and always have polyphase windings on the stator. These machines characteristically have high efficiency and a power factor that is adjustable via the rotor dc field current. Both leading and lagging power factors are possible. Smaller polyphase synchronous motors using permanent magnets on the rotor are widely used in self-synchronous, variable-frequency drive systems (brushless dc motors and electronically commutated motors). In these drives the rotor position is detected and used to determine the frequency and phase of the excitation. The advantages include high efficiency, fast response, and very good control characteristics.

In smaller sizes, synchronous reluctance motors are used where precise speed control is required, such as in timing motors or where many small motors must be operated at exactly the same speed. Reluctance machines generally have lower efficiency and power factor but are inexpensive. Improvements in design may make the reluctance machine competitive with induction machines in variable-speed applications using variable-frequency excitation. *See* RELUCTANCE MOTOR.

For special, low-power applications requiring precise speed control and very smooth torque, the hysteresis motor is used. These machines use a special permanent-magnet rotor material which is remagnetized each time the motor is started. Starting torque results from the lag between magnetizing force and flux density and is essentially independent of rotor speed. At synchronism the hysteresis motor runs as a permanent-magnet synchronous motor. *See* HYSTERESIS MOTOR.

AC series motor. A dc motor with the armature and field windings in series will run on ac since both magnetic fields reverse when the current reverses. Since these machines run on ac or dc, they are commonly called universal motors. The speed can be controlled by varying the voltage, and these machines are therefore widely used in small sizes for domestic appliances that require speed control or higher speeds than can be attained with 60-Hz induction motors. *See* ELECTRIC ROTATING MACHINERY; MOTOR; UNIVERSAL MOTOR; WINDINGS IN ELECTRIC MACHINERY.

Donald W. Novotny

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; A. E. Fitzgerald, C. Kingsley, and S. D. Umans, *Electric Machinery*, 6th ed., 2003; G. R. Slemon, *Electric Machines and Drives*, 1992.

Alternative fuels for vehicles

Conventional fuels such as gasoline and diesel are gradually being replaced by alternative fuels such as gaseous fuels (natural gas and propane), alcohol (methanol and ethanol), and hydrogen. Conventional fuels can also be modified to a reformulated gasoline to help reduce toxic emissions. Technological advances in the automotive industry (such as in fuel cells and hybrid-powered vehicles) are helping to increase the demand for alternative fuels.

Two key issues associated with the use of conventional fuels for vehicles prompt interest in alternative fuels: (1) gasoline- and diesel-powered vehicles are considered to be a significant source of air pollution; and (2) conventional fuels are produced from crude oil, a nonrenewable energy resource. The success of alternative fuels in the marketplace will depend on numerous factors, including public understanding and consumer awareness; economics; automobile performance; availability of fuels, vehicles, and distribution and marketing systems; and changes in technology. *See* AIR POLLUTION.

Gaseous fuels. Vehicle emissions from natural gas and propane are expected to be lower and less harmful to the environment than those of conventional gasoline. Because natural gas and propane are less complex hydrocarbons, the levels of volatile organic compounds and ozone emissions should be reduced. Both of these fuels are introduced to the engine as a gas under most operating conditions and require minimal fuel enrichment during warm-up. Leaner burning fuels, they also achieve lower carbon dioxide and carbon monoxide levels than gasoline. However, because they burn at higher temperatures, emissions of nitrogen oxide are higher. An important property of gaseous fuels is their degree of resistance to engine knock. Because of their higher-octane value relative to gasoline, there is less of a tendency for these fuels to knock in spark-ignition engines. To achieve the optimal performance and maximum environmental benefits of natural gas and propane, technological advancements must continue to reduce the costs of dedicated vehicles to be competitive with conventional vehicles, and the necessary fueling infrastructure must be ensured.

Natural gas. Natural gas is a colorless, odorless hydrocarbon that is neither carcinogenic nor corrosive.

Once processed at a gas plant, natural gas is composed of about 97% methane (CH_4). Being lighter than air, it does not pool on the ground when leaked but dissipates into the atmosphere, reducing the hazard of fire. One drawback as a transportation fuel is its very low energy density relative to gasoline or diesel fuel. To eliminate bulkiness, the gas is compressed and stored in cylindrical tanks. At a pressure of 3000 pounds per square inch (20 megapascals), the volumetric energy density of a typical natural gas cylinder is only 20% that of a similar volume of gasoline. Thus, the driving range of a natural gas vehicle carrying the same volume of fuel as a gasoline-fueled vehicle will be less. Larger vehicles, carrying two or three natural gas tanks, can achieve an adequate driving range for most personal or fleet purposes, but at the cost of space and the additional weight of the cylinders. *See* METHANE.

Natural gas engine technology is fairly well developed, including heavy-duty applications. Vehicle refueling appliances have reduced the constraints associated with the lack of distribution and refueling infrastructure; however, the associated costs in addition to those related to conversion tend to limit natural gas usage to high-mileage vehicles. Significant advances have been made with regard to lighter and stronger fuel cylinders, and there is more variety of natural gas vehicles produced by original equipment manufacturers than for any other alternative fuel. Still required for market competitiveness of natural gas is a refueling infrastructure with widespread access and low-cost fast-fill capabilities. *See* LIQUEFIED NATURAL GAS (LNG); NATURAL GAS.

Propane. Propane (C_3H_8), a hydrocarbon produced from crude oil and natural gas, has been one of the more successful alternative transportation fuels. A cleaner-burning fuel than gasoline or diesel fuel, it can result in lower maintenance costs, requiring fewer spark plug and oil changes, and less wear on such components as piston rings and bearings. Since propane has a higher octane rating than gasoline, propane-fueled vehicles can use higher engine compression ratios, resulting in more power and better fuel efficiency.

Propane has a lower energy density than gasoline by volume but a higher energy density by weight; therefore, the weight of a tank full of propane is similar to that of one filled with gasoline. Compared to ethanol, methanol, or natural gas, propane has a higher energy content by weight and volume, so it can take a vehicle farther on an average-capacity tank than a similar tank of any of the other alternative fuels. Refueling time for propane is similar to conventional fuels.

Several factors related to customer convenience and conversion costs have been restrictive to automotive propane market development. The physical characteristics of propane require modifications to conventional vehicles by either the automotive manufacturer at the factory or after-market conversions; either way, the cost of the modifications adds to the price paid by the consumer for a gasoline or diesel vehicle. In order to maintain a liquid state, propane must be stored under pressures of up to 200 psi (1.3 MPa). Because of this fact and the lower energy density of propane relative to gasoline, propane-powered vehicles must be fitted with storage tanks that impose weight and space penalties in order for the vehicles to travel the same distance as gasoline counterparts. Also, the vapor pressure of propane varies in relation to the atmospheric temperature. At temperatures of -43°C (-45°F) and colder, the pressure in a fuel tank is essentially zero, the product becomes liquid, and propane fuel systems simply do not operate. Another inconvenience arises due to the potential hazard of vapor accumulation at ground level; for this reason, some city and municipality bylaws forbid propane vehicles from parking in underground facilities.

On the basis of a grams per unit of distance traveled, propane full fuel cycle emissions compare favorably with gasoline, except for nitrogen oxides. With further development of propane vehicles by original equipment manufacturers, the establishment of certification standards for propane conversions, and continued technological improvements in emission control systems, the environmental impact of propane could prove even greater. The potential economic benefits from conversion to propane are greatest in high-fuel-consumption fleets such as taxis, where payback occurs in 1–2 years.

Alcohol fuels. The most significant advantage of alcohol fuels over gasoline is their potential to reduce ozone concentrations and to lower levels of carbon monoxide. Another important advantage is their very low emissions of particulates in diesel engine applications. In comparison with hydrocarbon-based fuels, the exhaust emissions from vehicles burning low-level alcohol blends (such as gasohol containing 10% alcohol by volume) contain negligible amounts of aromatics and reduced levels of hydrocarbons and carbon monoxide but higher nitrogen oxide content.

Exposure to aldehydes, in particular formaldehyde which is considered carcinogenic, is an important air-pollution concern. The aldehyde fraction of unburned fuel, particularly for methanol, is appreciably greater than for hydrocarbon-based fuels; therefore, catalytic converters are required on methanol vehicles to reduce the level of formaldehyde to those associated with gasoline. *See* ALCOHOL FUEL.

Methanol. Methanol (CH_3OH) is a clear, colorless, high-performance liquid fuel that can be used in both spark-ignition and diesel engines. It can be burned as a neat (100% methanol) or near-neat fuel in internal combustion engines. A blend of 85% methanol and 15% gasoline, M85, is the most common form of methanol fuel used in light-duty vehicles. M100, or neat methanol, is used in some heavy-duty trucks and buses.

The appeal of methanol as an alternative to gasoline stems from its being a relatively inexpensive clean-burning liquid fuel. A constraint is the fact that it has only half the energy density of gasoline, compensated to some degree by its better thermal efficiency. Safety concerns relate to the fact that

methanol is toxic and can be absorbed through the skin. Also, neat methanol burns with an invisible flame, making methanol fires without a colorant difficult to detect.

Besides being used as a neat or near-neat transportation fuel, methanol has been used to produce methyl tertiary butyl ether (MTBE), an oxygenate which, if blended with gasoline up to 10%, adds one octane number to the fuel. Concerns related to MTBE leaking and contaminating ground water in some areas in the United States resulted in a reexamination of the use of this oxygenate, and the phase-out of its use in California by December 31, 2002. Methanol is also being tested in the transportation industry as a source of the hydrogen in hybrid fuel cell vehicles. *See* ETHER; METHANOL.

Ethanol. Ethanol (C_2H_5OH) produced from biomass is a renewable fuel source currently being marketed in neat, near-neat, and low-level premium fuel blends. In addition to providing environmental benefits, use of ethanol fuel produced from fermentable starch or sugar crops provides economic advantages to the agriculture sector. However, ethanol production costs are very high relative to gasoline and must be reduced substantially if ethanol is to compete on an economic basis.

To date, in the fuel industry ethanol has been used mainly as an additive in low-level gasohol blends. These blends have been successfully used in unmodified gasoline vehicles with warranty coverage provided for automobiles sold in North America. Neat ethanol, generally used in heavy-duty applications, is less common and requires costly engine modifications. Flexible fuel vehicles can operate on a mixture of gasoline and up to 85% ethanol (E85). Ethanol has also been used in the production of ethyl tertiary butyl ether (ETBE), but to date it has not been economically competitive with MTBE.

The oxygen content of ethanol fuels used in properly modified vehicles results in increased energy efficiency and engine performance. Extended spark-plug life and lower carbon deposits are also expected because ethanol burns cleaner than gasoline.

A significant drawback of ethanol is its much lower energy density than gasoline (approximately 34% less). To travel distances similar to a gasoline-powered vehicle without refueling would require a much bigger fuel tank; this requirement is slightly offset by the greater energy efficiency of ethanol. *See* ETHYL ALCOHOL.

Hydrogen. Hydrogen is the lightest and most abundant element in the universe. It can be produced from a number of feedstocks in a variety of ways. The production method thought to be most environmentally benign is the electrolysis of water, but probably the most common source of hydrogen is the steam reforming of natural gas. Once produced, hydrogen can be stored as a gas, liquid, or solid and distributed as required. Liquid storage is currently the preferred method, but it is very costly. Metal hydride and compressed storage are also being investigated. *See* METAL HYDRIDES.

Hydrogen-powered vehicles can use internal combustion engines or fuel cells. They can also be hybrid vehicles of various combinations. When hydrogen is used as a gaseous fuel in an internal combustion engine, its very low energy density compared to liquid fuels is a major drawback requiring greater storage space for the vehicle to travel a similar distance to gasoline. Although hybrid vehicles can be more efficient than conventional vehicles and result in lower emissions, the greatest potential to alleviate air-pollution problems is thought to be in the use of hydrogen-powered fuel cell vehicles. Though currently very expensive, fuel cells are more efficient than conventional internal combustion engines. They can operate with a variety of fuels, but the fuel of choice is gaseous hydrogen since it optimizes fuel cell performance and does not require on-board modification.

Numerous hydrogen fuel cell vehicle models are being developed and tested with the objective of having such vehicles available for sale to the public in the early years of the twenty-first century. However, for gaseous hydrogen to become a fuel of choice in the transportation industry, numerous technical and economic constraints related to hydrogen production, on-board and site storage, and marketing and distribution systems must be overcome. Safety issues and perception must also be addressed. *See* FUEL CELL; HYDROGEN.

Reformulated fuels. Conventional gasoline and diesel fuels are complex mixtures of many different chemical compounds. Over time these fuels have undergone reformulation to improve their value to the transportation industry; this reformulation ranged from mild, like the backing-out of butane to reduce the volatility, to severe adjustments that substantially change the composition.

The U.S. Clean Air Act Amendments (CAAA) have served to increase interest in using regulated changes to motor fuel characteristics as a means of achieving environmental goals. The reformulated gasoline (RFG) program was designed to resolve ground-level ozone problems in urban areas. Under this program, compared to conventional gasoline, the amount of heavy hydrocarbons is limited in reformulated gasoline, and the fuel must include oxygenates and contain fewer olefins, aromatics, and volatile organic compounds.

Diesel-powered heavy-duty trucks and buses are said to account for about a quarter of the nitrogen oxides (NO_x) that result in smog formation and well over half of the particulates produced by mobile sources. Particulates, unburned fuel particles, can stick to the lungs if inhaled, causing increased susceptibility to respiratory illness. Research is under way to develop diesel fuel that will reduce both the particulate and NO_x levels. These fuels will have higher cetane ratings than conventional diesel fuel and lower sulfur and aromatics. Michelle Heath

Bibliography. R. L. Bechtold, *Alternative Fuels Guidebook: Properties, Storage, Dispensing, and Vehicle Facility Modifications*, 1997; R. L. Busby,

Hydrogen and Fuel Cells: A Comprehensive Guide, 2005; Society of Automotive Engineers, *State of Alternative Fuel Technologies 2000*, 2000.

Altimeter

Any device which measures the height of an aircraft. The two chief types are the pressure altimeter, which measures the aircraft's distance above sea level, and the radar altimeter, which measures distance above the ground.

Pressure Altimeter

A pressure altimeter precisely measures the pressure of the air at the level an aircraft is flying and converts the pressure measurement to an indication of height above sea level according to a standard pressure-altitude relationship. In essence, a pressure altimeter is a highly refined aneroid barometer since it utilizes an evacuated capsule whose movement or force is directly related to the pressure on the outside of the capsule (**Fig. 1**). Various methods are used to sense the capsule function and cause a display to respond such that the pilot sees the altitude level much as one looks at a watch.

Because altitude measured in this manner is also subject to changes in local barometric pressure, altimeters are provided with a barosetting that allows the pilot to compensate for these weather changes, the sea-level air pressure to which the altimeter is adjusted appearing in a window of the dial. Flights below 18,000 ft (5486 m) must constantly contact the nearest traffic center to keep the altimeters so updated. Flights above 18,000 ft and over international waters utilize a constant altimeter setting of 29.92 in. Hg, or 1013.2 millibars (101.32 kilopascals), so that all high-flying aircraft have the same reference and will be interrelated, providing an extra margin of safety. *See* AIR NAVIGATION.

Electromechanical computers are provided that correct the sensed static pressure for the static systems defect error, which is an inherent aircraft error due to the effect of high-speed flying on the air mass immediately around the aircraft fuselage in the vicinity of the static ports. The corrected altitude information is usually transmitted to electromechanical indicators for the pilot's display. Much work has been done to improve the readability of the display, so that pilots may make rapid readings with a low probability of reading error. The present accuracy of altimeters is on the order of 0.5% of the indicated altitude for all mechanical indicators and 0.2% of the indicated altitude for computer systems. *See* BAROMETER.

James W. Angus

Radar Altimeter

A radar altimeter is a low-power radar that measures the distance of an aircraft (or other aerospace vehicle) above the ground. It differs from a pressure altimeter, which measures distance above sea level when properly adjusted for atmospheric pressure variations.

Radar altimeters are often used in aircraft during bad-weather landings. They are an essential part of many blind-landing and automatic navigation systems and are used over mountains to indicate terrain clearance. Special types are used in surveying for quick determination of profiles. Radar altimeters are used in bombs, missiles, and shells as proximity fuses to cause detonation or to initiate other functions at set altitudes. Radar altimeters have been used on

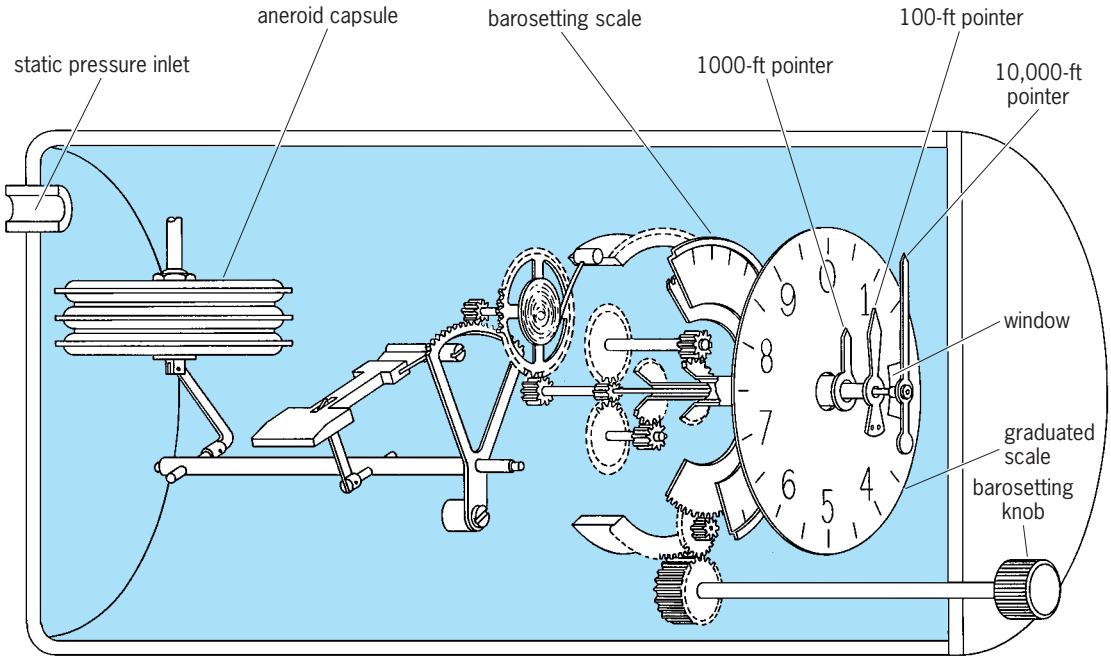

Fig. 1. Schematic arrangement of sensitive pressure altimeter. 1 ft = 0.3 m.

various spacecraft, starting with *Skylab* in 1973, to measure the shape of the geoid and heights of waves and tides over the oceans. Notable spacecraft altimeters over the Earth were those aboard the *Seasat* and *Geosat* American satellites, the American-French *TOPEX/POSEIDON,* and the European *ERS 1.* Other spacecraft altimeters provide topographic information on other planets, particularly Venus. *See* AUTOMATIC LANDING SYSTEM; GROUND PROXIMITY WARNING SYSTEM.

Principle of operation. Like other radar devices, the altimeter measures distance by determining the time required for a radio wave to travel to and from a target, in this case the Earth's surface. If the Earth were a perfectly flat horizontal plane or smooth sphere, the signal would come only from the closest point (**Fig. 2***a*), and would be a true measure of altitude. Actually, the Earth is not smooth, and energy is scattered back to the radar from all parts of the surface illuminated by the transmitter (Fig. 2*b*). For the radar to measure distance to the ground accurately, it must distinguish between the energy from points near the vertical and that from more distant points.

Radar altimeters can seldom have highly directive antennas, for they must be able to function regardless of the aircraft attitude. Thus the antenna cannot discriminate against off-vertical signals. The wide beam width means the antenna gain is small, so the power required is much greater than for a comparable directive radar working against a large, close target. Spaceborne altimeters used to measure distances to the ocean and ice sheets, however, are mounted on stable platforms that allow them to use very narrow antenna beams. *See* ANTENNA (ELECTROMAGNETISM).

Most radio altimeters use either pulse or frequency modulation, the former being more popular for high altitudes, and the latter for low altitudes. *See* FREQUENCY MODULATION; PULSE MODULATOR.

Pulse altimeters. In a typical pulse altimeter (**Fig. 3***a*) the radio-frequency carrier is modulated with short pulses (under 0.25 microsecond). The short pulse permits measurements, even at low altitudes, of the time delay between the leading edge of the transmitted pulse and that of the pulse returned from the ground. Early pulse altimeters displayed the received signal on a cathode-ray tube with circular sweep, allowing the pilot to determine the leading-edge position of the echo signal. Modern pulse altimeters use a tracking gate system. One gate is kept close to the leading edge by a servo system that adjusts the position of the gate to the optimum delay point. A simple single-gate system can be used (Fig. 3*a*), but most pulse altimeters use two or three gates to achieve better distance measurement in the presence of noise and fading. The time delay is displayed for the pilot as a range-dial reading or digital readout. The delay is usually converted to a digital signal that is delivered to the autopilot and flight-control computer.

The return power for a pulse altimeter (the mean of many pulses; Fig. 2*c*) is a combination of scattered returns from the entire illuminated area and specular return from the nearest point to which the beam is perpendicular. The specular contribution is important only for very smooth water, pavements, or desert playas. Usually the maximum value of the return power varies with height h as $1/h^2$ to $1/h^3$. Interference between different scattering centers causes fading, so the pulses must be averaged over a considerable time to get a precise indication. *See* INTERFERENCE OF WAVES.

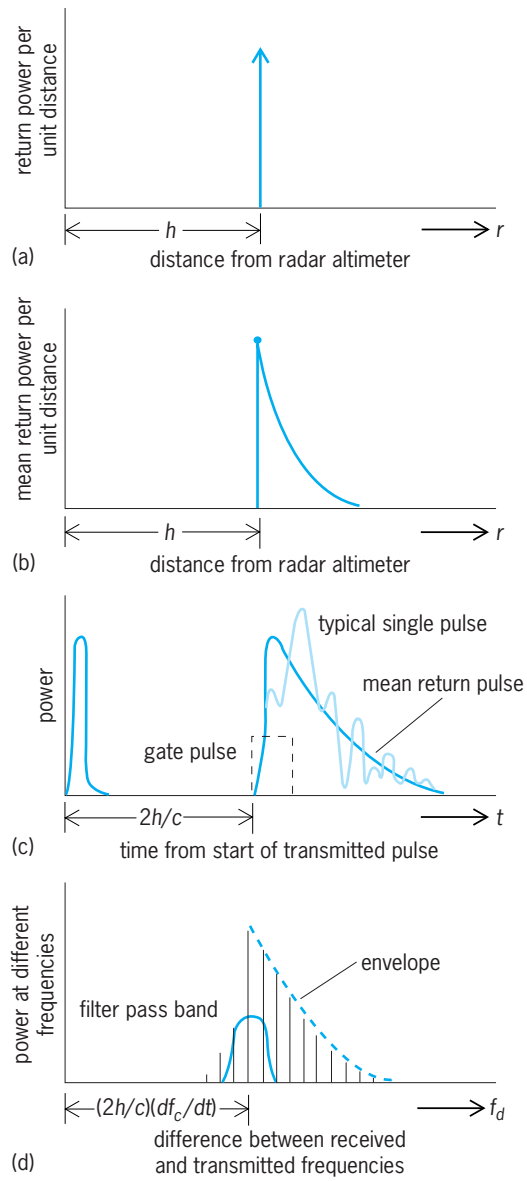

Fig. 2. Radar altimeter principles. (*a*) Ideal return power distribution, which would be observed if the signal came only from the closest point to the radar. (*b*) Typical distribution of average returned power as a function of distance from the radar. (*c*) Pulse shapes in a wide-beam, narrow-pulse altimeter with a single tracking gate. (*d*) Typical FM altimeter difference-frequency spectrum. Spacing between lines is the sweep repetition frequency.

FM altimeters. In a frequency-modulated (FM) altimeter (Fig. 3*b*), the frequency of a continuous carrier is swept in some manner, usually to give a triangular frequency-time curve. The difference in frequency between that received from the ground (but transmitted earlier) and that being transmitted is

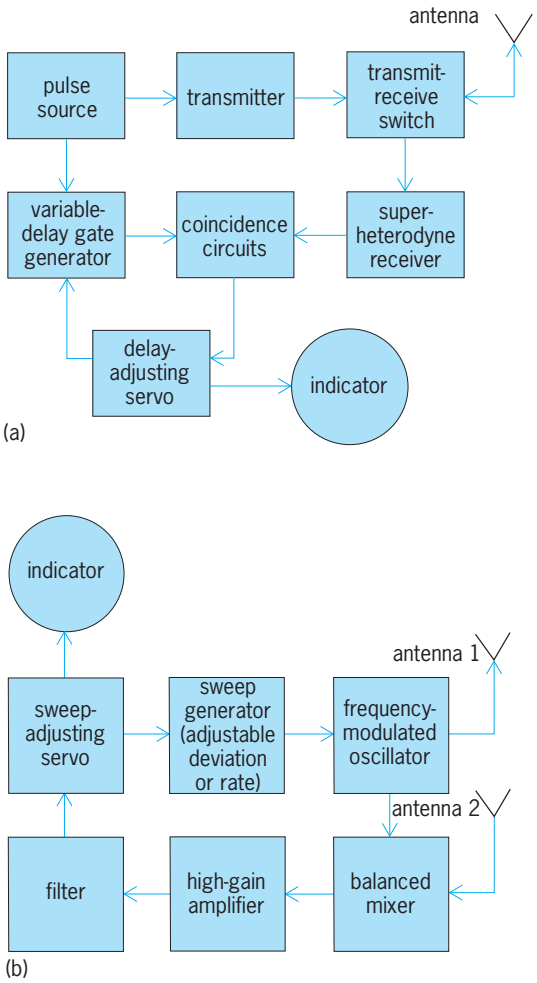

Fig. 3. Block diagrams of typical radar altimeters. (*a*) Automatic pulse altimeter. Control, tracking, and output circuits may all be digital. (*b*) Servo-adjusted-sweep FM altimeter.

a measure of the time delay (Fig. 2*d*). This difference frequency is given by the equation below, where r

$$f_d = \frac{df_c}{dt}(2r/c)$$

is the distance from the radar, c is the speed of light, and f_c is the carrier frequency. Some radars with fixed df_c/dt (lacking a servo and fixed filter) use electronic frequency meters calibrated in range as indicators. Because the return comes from many ranges besides the altitude, an unsophisticated radar of this type may read an effective range frequency more than 10% high in some cases. Other types (Fig. 3*b*) use a relatively narrow filter and keep the minimum difference frequency (altitude frequency) centered in this filter by a servo which adjusts the sweep rate df_c/dt. The sweep rate is indicated, as range, on a dial and also digitized for supply to autopilot and flight computer.

The frequency spectrum (Fig. 2*d*) is composed of individual lines, spaced by the sweep frequency. Suitable signal processing permits recovery of the envelope of the spectrum, which is analogous to the mean return pulse of a pulse altimeter (Fig. 2*c*). Because of problems with short pulses at short ranges, FM altimeters are normally used where altitude measurements must be below about 100 ft (30 m).

Other radio devices may also serve as altimeters. The proximity fuse is a crude device for measuring altitude by field strength alone. Various correlation devices have been proposed, some of which employ noise modulation of either amplitude or frequency.

Spaceborne altimeters. Pulse altimeters are used on spacecraft to observe properties of the ocean and continental ice sheets (Greenland and Antarctica). For these applications, the exact location of the spacecraft must be determined by accurate tracking from land. The distance from altimeter to surface then becomes a measure of the variation of the surface itself. The altimeters must achieve great accuracy. To do so, they transmit pulses with very large bandwidths, thereby achieving resolutions of only a few feet. The actual distance measurement accuracy can be better than 5 cm (2 in.), provided suitable corrections are made for slight variations in the refractive index of (and therefore velocity of propagation through) the atmosphere. *TOPEX/POSEIDON* uses two frequencies to aid in correcting for the minute effects of the ionosphere.

To achieve this great precision without the use of extremely short pulses that would need excessive peak power, pulse-compression techniques are employed. Normally these methods use a binary phase-coded pulse.

These spaceborne systems have allowed major improvements in knowledge of the shape of the geoid over the oceans, where surveying methods that work with great precision over land perform poorly or cannot be used at all. Moreover, with the precision now available, oceanographers can locate and track changes in the major current systems in the ocean by the slight bulge associated with the current flow. Glaciologists use the altimeter data taken over the continental ice sheets to detect growth and decay of the ice, which can be an indication of global warming or cooling. Since the ice sheets contain most of the fresh water on Earth, significant melting resulting from global warming could result in major catastrophes along coastlines, making monitoring of glaciers an important concern. *See* GEODESY; GLACIOLOGY; OCEAN CIRCULATION.

The same altimeters also can measure wave height at points along the ocean surface beneath the satellite (the subsatellite track). When ocean waves are high, the leading edge of the average pulse (Fig. 2*c*) is stretched out, since the signal returns first from a relatively small area near the crests of the waves and then builds up as the entire surface is illuminated. The altimeter samples the pulse at close intervals (typically about 10 nanoseconds), and the averaged samples are used to determine the shape of the received pulse. This can be compared with theoretical shapes to achieve good estimates of the root-mean-square wave height beneath the spacecraft.

Altimeters can also be used as scatterometers to estimate the wind speed beneath the spacecraft. This is possible because the received signal decreases as the wind increases. Since meteorologists require wind

measurements over wide areas, this subsatellite wind measurement is less useful than that from spaceborne radar scatterometers looking for hundreds of miles to the side of the spacecraft. Moreover, the wind sensitivity of radar scatterometers beyond an angle of 20° from the vertical is much greater than that at vertical incidence. Nevertheless, the altimeters have proved useful during intervals when no spaceborne scatterometer was active. *See* REMOTE SENSING.

Richard K. Moore

Bibliography. D. E. Barrick, Analysis and interpretation of altimeter sea echo, *Adv. Geophys.,* 37:60–98, 1985; B. C. Douglas and R. E. Cheney, Geosat: Beginning of a new era in satellite oceanography, *J. Geophys. Res.,* 95:2833–2836, 1990; Seasat Special Issue II: Scientific Results, *J. Geophys. Res.,* 88(C3):1531–1745, 1983; M. I. Skolnik, *Introduction to Radar Systems,* 2d ed., 1980; G. I. Sonnenberg, *Radar and Electronics Navigation,* 6th ed., 1988; F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing,* vol. 2, 1982.

Altitudinal vegetation zones

Intergrading regions on mountain slopes characterized by specific plant life forms or species composition, and determined by complex environmental gradients.

Along an altitudinal transect of a mountain, there are sequential changes in the physiognomy (growth form) of the plants and in the species composition of the communities. This sequential zonation of mountain vegetation has been recognized for centuries. Vertical zonation was fully developed as an ecological concept by the work of C. H. Merriam with the U.S. Biological Survey of 1889. He described a series of life zones on the slopes of the San Francisco Peaks in Arizona, based on characteristic species of the flora and fauna. Other patterns of plant physiognomic and community zonation have now been cataloged in mountain ranges throughout the world. *See* LIFE ZONES.

Merriam associated his life zones with temperature gradients present along mountain slopes. Later research on patterns of altitudinal zonation has centered on the response of species and groups of species to a complex of environmental gradients. Measurements of a species along a gradient, for example, the number of individuals, biomass, or ground coverage, generally form a bell-shaped curve. Peak response of a species occurs under optimum conditions and falls off at both ends of the gradient. The unique response of each species is determined by its physiological, reproductive, growth, and genetic characteristics.

Zones of vegetation along mountain slopes are formed by intergrading combinations of species that differ in their tolerance to environmental conditions. Zones are usually indistinct entities rather than discrete groupings of species. However, under some conditions of localized disjunctions, very steep sections of gradients, or competitive exclusion, discontinuities in the vegetation can create discrete communities. *See* ECOLOGICAL COMMUNITIES.

Vegetation zones are often defined by the distributions of species having the dominant growth form, most frequently trees. For example, in the Cascade Mountains of Washington and Oregon, tundra above treeline has been called the sedge–grass zone, but successively lower zones are spruce–fir, arborvitae–hemlock, Douglas-fir, ponderosa pine, and finally several fescue–wheatgrass–sagebrush zones in which no trees grow. Another set of zones could be designated if other criteria were used. The boundaries of the middle elevation zones might be different if they had been defined instead by distributions of the dominant shrubs.

Environmental gradients in mountains. Altitudinal vegetation zonation, therefore, is an expression of the response of individual species to environmental conditions. Plants along an altitudinal transect are exposed, not to a single environmental gradient, but to a complex of gradients, the most important of which are solar radiation, temperature, and precipitation. Although these major environmental gradients exist in most mountain ranges of the world, the gradients along a single altitudinal transect are not always smooth because of topographic and climatic variability. This environmental variation can result in irregular vegetation zones.

The solar energy received by mountain surfaces increases with altitude, associated with decreases in air density and the amount of dust and water vapor. Global radiation levels in the European Alps have been shown to be 21% greater at 10,000 ft (3000 m) than at 650 ft (200 m) under a cloudless sky. An overcast sky is more efficient at reducing short-wave energy reaching low elevations and can increase the difference in energy input to 160%. However, more frequent clouds over high elevations relative to sunnier lower slopes commonly reduces this difference. *See* SOLAR RADIATION.

Vegetation patterns are also strongly influenced by the decline in air temperature with increasing altitude, called the adiabatic lapse rate. Lapse rates are generally between 1.8°F to 3.6°F per 1000 ft (1°C to 2°C per 300 m), but vary with the amount of moisture present; wet air has a lower lapse rate. Thus, plants occurring at higher elevations generally experience cooler temperatures and shorter growing periods than low-elevation plants. Variation in the temperature gradient can be caused by differences in slope, aspect, radiation input, clouds, and air drainage patterns. *See* AIR TEMPERATURE.

The precipitation gradient in most mountains is the reverse of the temperature gradient: precipitation increases with altitude. Moist air from low elevations is forced upward by the blocking action of the mountains into regions of cooler temperatures. Air holds less moisture at low temperatures than at warmer ones. When the atmosphere reaches the point of saturation (the dew point), condensation occurs, forming clouds or precipitation. This moisture gradient is most pronounced on windward slopes and in ranges close to warm and moist oceanic

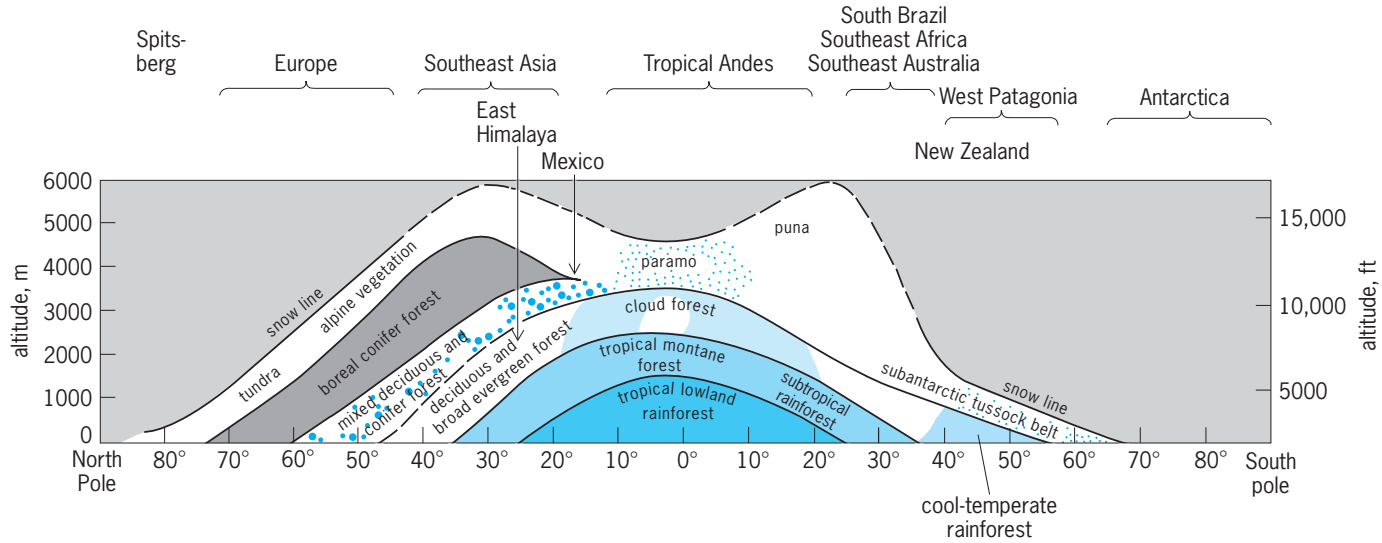

Fig. 1. Generalized patterns of altitudinal vegetation zonation in the Northern and Southern hemispheres. (*After L. W. Price, Mountains and Man, University of California Press, 1981*)

winds. As the moving air passes over the peaks of the mountain range and moisture in the air is depleted, precipitation is reduced, creating a rain "shadow" on the lee side of the range. *See* PRECIPITATION (METEOROLOGY).

Temperate, tropical, and high latitudes. General changes in vegetation with increases in altitude include reduction in plant size, slower growth rates, lower production, communities composed of fewer species, and less interspecific competition. However, many regional exceptions to these trends exist. In western North America, for example, the lowest zone is often a treeless shrubland or prairie, so plant size initially increases with altitude as trees become important. Above this zone, the trend toward smaller size prevails.

Characteristics of vegetation zones also vary with latitude (**Fig. 1**). Mountains at higher latitudes have predominantly seasonal climates, with major temperature and radiation extremes between summer and winter. Equatorial and tropical mountains have a strong diurnal pattern of temperature and radiation input with little seasonal variation. The upper altitudinal limit of trees, and the maximum elevation of plant growth generally, decreases with distance from the Equator, with the exception of a depression near the Equator.

The following examples of altitudinal zonation in temperate, tropical, and high-latitude regions illustrate patterns that exist in each region. These patterns are not comprehensive, and many others exist locally and worldwide.

Temperate. The vegetation of the Santa Catalina Mountains in Arizona illustrates the variety of zones found in a temperate mountain range (**Fig. 2**). Typically, precipitation increases with altitude, while air and soil temperatures decrease. Community diversity, the number of species present, decreases with altitude in all zones except the scrub desert, where arid conditions support fewer species. Biomass production is less at low elevations than on higher slopes, the reverse of the general trend, because of limitations to growth under arid conditions.

Hot and dry conditions at lower elevations support a treeless desert shrub and grassland, dominated by such shrubs as palo verde (*Cercidium microphyllum*) and mesquite (*Prosopis juliflora*). With increasing elevation, more moisture allows growth of broadleaf deciduous and evergreen oak trees (*Quercus* spp.). Needle-leaf evergreen trees, including ponderosa pine (*Pinus ponderosa*) and border limber pine (*P. strobiformis*), become increasingly important in the mid-elevation vegetation zones where temperatures are cooler, moisture more readily available, and the growing season longer. Broadleaf trees decrease in importance in these zones. Finally, in the upper zones, where the coolest temperatures and highest moisture levels prevail, conifers that commonly extend farther north including Douglas-fir (*Pseudostuga menziesii*) and Engelmann spruce (*Picea engelmanni*) replace species better adapted to lower-elevation environments. At higher latitudes and where mountains are high enough, an alpine zone of grasses, forbs, cushion plants, and dwarf shrubs occurs above treeline. A similar sequential replacement of species in the altitudinal zones is repeated in both the herb and shrub layer.

The Southern Hemisphere has fewer mountains in the temperate regions than areas north of the Equator. The oceanic climate of these mountains is cooler and moister because of the small land mass relative to the nearby oceans. A prolonged cold season is absent. In New Zealand, broadleaf evergreen trees such as evergreen beech (*Nothofagus* spp.) dominate moist mountain forests, with needle-leaf evergreens such as *Podocarpus* spp. being of lesser importance. Zones at all elevations have a richer, more diverse flora than their Northern Hemisphere counterparts. Within each zone, shrubs are more luxurious, and epiphytes and hanging vines are common.

Tropical. Vegetation zones at all elevations in tropical mountains are dominated by a diurnal climate

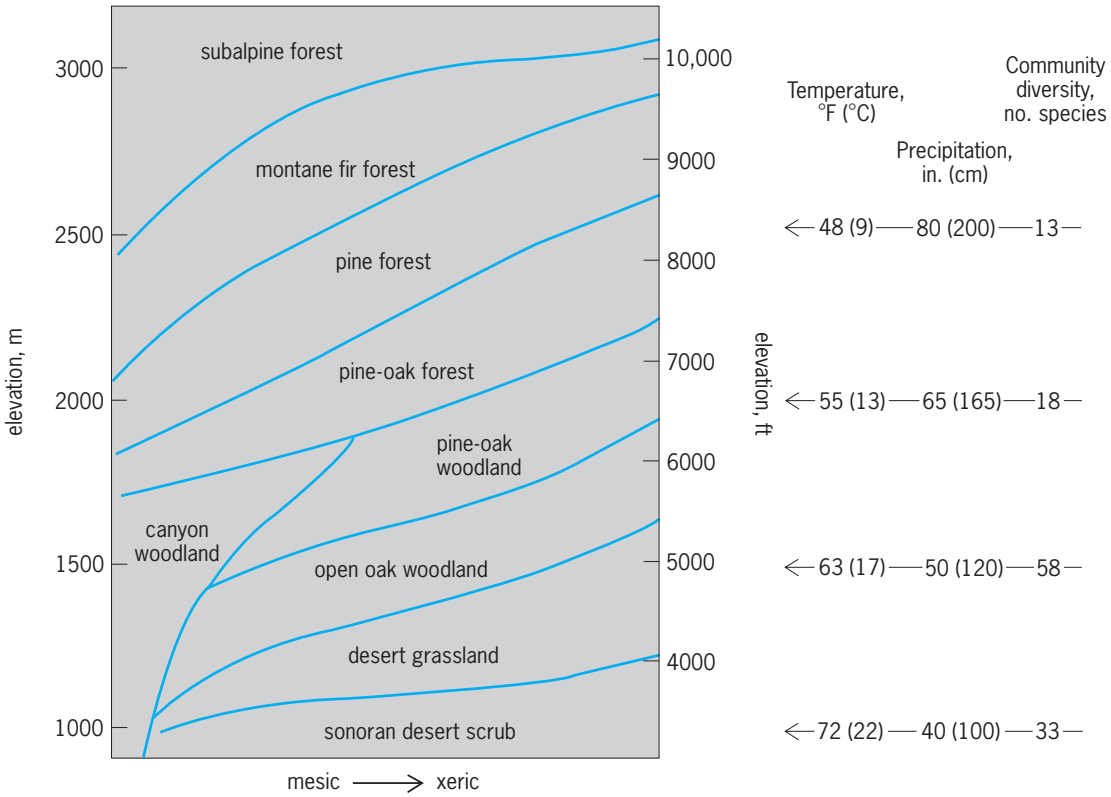

Fig. 2. Vegetation zones and gradients of mean annual soil temperature, annual precipitation, and community diversity on the south slopes of the Santa Catalina Mountains, Arizona, a temperate mountain range. (*After R. H. Whittaker, and W. A. Niering, Vegetation of the Santa Catalina Mountains: A gradient analysis of the south slope, Ecology, 46: 429–451, 1965*)

with little seasonality, because sun angles are high throughout the year. Growing seasons are determined by moisture availability rather than by temperature. At lower elevations in the Andes, tropical rainforests with a species-rich multilayered and luxurious canopy of broadleaf evergreen trees cover mountain slopes. These trees are accompanied by a lush mixture of hanging vines, epiphytes, shrubs, and flowering herbs. In the submontane and montane zones at higher elevations, the complexity of the forest is reduced. Trees are shorter and are adapted to drier conditions, and the canopy has fewer layers. Vines, epiphytes, and shrubs are present in fewer numbers, but mosses and lichens become more common.

The subalpine zone is an elfin woodland or cloud forest of short, stunted trees covered with an abundance of mosses and lichens. Vegetation is adapted to the higher moisture and lower light levels created by clouds, which almost continuously envelop the zone. Plant communities in this zone are less complex than those lower on the mountain; fewer species are present and canopy structure is simpler. In the northern Andes of Colombia, the alpine zone consists of a low vegetation called paramo, which is dominated by arborescent members of the sunflower family, Compositae. Growth forms consist of tall tussock grasses, dwarf shrubs, herbs, and two forms unique to the tropics of the Southern Hemisphere: dendroid or tufted-leaf stemmed plants, and woolly candlelike plants. *See* PARAMO.

Farther from the Equator in the central and southern Andes, the climate is drier and lower zones are dominated by forests of deciduous trees or grassland savannas. The alpine zone consists of puna, a vegetation of smaller tussock grasses and many cushion and rosette plants. *See* PUNA; SAVANNA.

High latitude. At high latitudes in both Northern and Southern hemispheres, fewer vegetation zones exist. Short growing seasons permit no tree growth, and shrubs and herbs are limited at higher elevations by a permanent cover of ice or snow, called the nival zone. Plant communities consist of fewer species than in temperate or tropical regions. In the Olgivie Mountains, Yukon Territory, Canada, valley bottoms are covered by willow (*Salix* spp.) and birch (*Betula* spp.) shrubs with a ground layer of grasses and forbs in drier sites, and tussock tundra dominated by cottongrass (*Eriophorum* spp.) in wet sites. At higher elevations, tall woody shrubs are absent, and this zone consists of a tundra of dwarf heath shrubs including mountain heather (*Cassiope* spp.) and mountain avens (*Dryas* spp.), grasses, forbs, and plants with cushion or rosette growth forms. *See* TERRESTRIAL ECOSYSTEM.

John S. Campbell

Bibliography. E. W. Beals, Vegetational change along altitudinal gradients, *Science*, 165:981-985, 1969; W. Lauer, The altitudinal belts of the vegetation in the central Mexican highlands and their climatic condition, *Arctic Alpine Res.*, 5(pt. 2):A99-114, 1973; L. W. Price, *Mountains and Man*, 1981; C. Troll (ed.), *Geoecology of the Mountainous Regions of the Tropical Americas*, 1968; H. Walter, *Vegetation of the Earth and Ecological Systems of the Geo-Biosphere*, 1979; R. H. Whittaker and W. A. Niering,

Vegetation of the Santa Catalina Mountains: A gradient analysis of the south slope, *Ecology*, 46:429-451, 1965.

Alum

A colorless to white crystalline substance which occurs naturally as the mineral kalunite and is a constituent of the mineral alunite. Alum is produced as aluminum sulfate by treating bauxite with sulfuric acid to yield alum cake or by treating the bauxite with caustic soda to yield papermaker's alum. Other industrial alums are potash alum, ammonium alum, sodium alum, and chrom alum (potassium chromium sulfate). Major uses of alum are as an astringent, styptic, and emetic. For water purification alum is dissolved; it then crystallizes out into positively charged crystals that attract negatively charged organic impurities to form an aggregate sufficiently heavy to settle out. Alum is also used in sizing paper, dyeing fabrics, and tanning leather. With sodium bicarbonate it is used in baking powder and in some fire extinguishers. *See* ALUMINUM; COLLOID. Frank H. Rockett

Aluminum

A metallic chemical element, symbol Al, atomic number 13, atomic weight 26.98154, in group 13 of the periodic system. Pure aluminum is soft and lacks strength, but it can be alloyed with other elements to increase strength and impart a number of useful properties. Alloys of aluminum are light, strong, and readily formable by many metalworking processes; they can be easily joined, cast, or machined, and accept a wide variety of finishes. Because of its many desirable physical, chemical, and metallurgical properties, aluminum has become the most widely used nonferrous metal. *See* PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Aluminum is the most abundant metallic element on the Earth and Moon but is never found free in nature. The element is widely distributed in plants, and nearly all rocks, particularly igneous rocks, contain aluminum in the form of aluminum silicate minerals. When these minerals go intosolution, depending upon the chemical conditions, aluminum can be precipitated out of the solution as clay minerals or aluminum hydroxides, or both. Under such conditions bauxites are formed. Bauxites serve as principal raw materials for aluminum production. *See* BAUXITE; CLAY MINERALS; IGNEOUS ROCKS; WEATHERING PROCESSES.

Physiochemical properties of pure aluminum

Property	Value
Atomic number	13
Atomic weight	26.98154
Crystal structure	Face-centered cubic
Density, g/cm^3 at 25°C	2.698
Melting point	660.37°C (1220.7°F)
Boiling point	2447°C (4436.6°F)
Latent heat of fusion, cal/g	94.9
Latent heat of vaporization, cal/g	2576
Heat of combustion, cal/g at 25°C	7420
Specific heat, cal/g/°C at 25°C	0.215
Thermal conductivity, cal/cm/°C/s at 25°C	0.566
Coefficient of thermal expansion, 10^{-6}/°C	23
Thermal diffusivity, cm^2/s at 25°C	0.969
Electrical resistivity, microhm-cm at 25°C	2.7
Hardness, Mohs	2–2.9
Reflectivity	85–90%

Aluminum is a silvery metal. Naturally occurring aluminum consists of a single isotope, $^{27}_{13}$Al. Aluminum crystallizes in the face-centered cubic structure with edge of the unit lattice cube of 4.0495 angstroms. Aluminum is known for its high electrical and thermal conductivities and its high reflectivity. A summary of some of the important properties of pure aluminum is given the **table**.

The electronic configuration of the element is $1s^2 2s^2 2p^6 3s^2 3p^1$. Aluminum exhibits a valence of +3 in all compounds, with the exception of a few high-temperature monovalent and divalent gaseous species.

Aluminum is stable in air and resistant to corrosion by seawater and many aqueous solutions and other chemical agents. This is due to protection of the metal by a tough, impervious film of oxide. At a purity greater than 99.95%, aluminum resists attack by most acids but dissolves in aqua regia. Its oxide film dissolves in alkaline solutions, and corrosion is rapid. *See* CORROSION.

Aluminum is amphoteric and can react with mineral acids to form soluble salts and to evolve hydrogen.

Molten aluminum can react explosively with water. The molten metal should not be allowed to contact damp tools or containers.

At high temperatures aluminum reduces many compounds containing oxygen, particularly metal oxides. These reactions are used in the manufacture of certain metals and alloys.

Applications in building and construction represent the largest single market of the aluminum industry. Millions of homes use aluminum doors, siding,

windows, screening, and down-spouts and gutters. Aluminum is also a major industrial building product. Transportation is the second largest market. In automobiles, aluminum is apparent in interior and exterior trim, grilles, wheels, air conditioners, automatic transmissions, and some radiators, engine blocks, and body panels. Aluminum is also found in rapid-transit car bodies, rail cars, forged truck wheels, cargo containers, and in highway signs, divider rails, and lighting standards. In aerospace, aluminum is found in aircraft engines, frames, skins, landing gear, and interiors. The food packaging industry is a fast-growing market.

In electrical applications, aluminum wire and cable are major products. Aluminum appears in the home as cooking utensils, cooking foil, hardware, tools, portable appliances, air conditioners, freezers, and refrigerators.

There are hundreds of chemical uses of aluminum and aluminum compounds. Aluminum powder is used in paints, rocket fuels, and explosives, and as a chemical reductant. *See* ALUMINUM ALLOYS.

Allen S. Russell

Bibliography. F. A. Cotton, et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004.

Aluminum (metallurgy)

The separation, extraction, and purification of alumina from ores, followed by the production of aluminum. Aluminum is the most abundant metallic element on the Earth and Moon but is never found free in nature. It makes up more than 8% of the solid portion of the Earth's surface. Seawater contains an average of only 0.5 ppm aluminum. The element is widely distributed in plants, where it may be present in significant concentrations, particularly in vegetation in marshy places and acid soils.

Nearly all rocks, particularly igneous rocks, contain aluminum in the form of aluminum silicate minerals. When subjected to weathering under the right conditions, these minerals go into solution; then, depending upon the chemical conditions, aluminum can be precipitated out of the solution as clay minerals or aluminum hydroxides, or both. Any quartz in the parent rock is relatively unattacked by the weathering processes and remains as sand in the deposit. Under conditions which precipitate clay minerals and leach out iron and the alkali and alkaline-earth metals, deposits of high-grade kaolin clay can result. Sometimes natural elutriation or mechanical separation by washing occurs, which separates the clay from the sand and enhances the purity of the deposit. Under conditions which precipitate aluminum principally as hydroxide minerals and leach out alkali and alkaline-earth metals completely and iron to varying degrees, bauxites are formed. It is also possible for clay minerals to be desilicified to aluminum hydroxides and for aluminum hydroxides to be resilicified to clay minerals. *See* BAUXITE; CLAY MINERALS; IGNEOUS ROCKS; WEATHERING PROCESSES.

Bauxites usually consist of mixtures of the following minerals in varying proportions: gibbsite, also known as hydrargillite [$Al(OH)_3$]; boehmite and diaspore [$AlO(OH)$]; clay minerals such as kaolinite [$Al_2Si_2O_5(OH)_4$]; quartz [SiO_2], and anatase, rutile, and brookite [TiO_2]. Small amounts of magnetite [Fe_3O_4], ilmenite [$FeTiO_3$], and corundum [Al_2O_3] are sometimes present. In addition to the above minerals, bauxites usually contain traces of other insoluble oxides such as zirconium, vanadium, gallium, chromium, and manganese. The term alumina trihydrate is often applied to the mineral gibbsite, and the term alumina monohydrate to the minerals boehmite and diaspore. However, the minerals are not hydrates in the true sense of the word. However, for processing purposes, bauxites are often classified as trihydrate bauxites and monohydrate bauxites, depending upon their content of the principal mineral or minerals containing the extractable alumina. In general, European bauxites are of the monohydrate type and are geologically older than bauxites in tropical countries, which occur generally as the trihydrate type. This, however, is not an absolute rule. Some European bauxites contain gibbsite along with boehmite and diaspore, and Caribbean bauxites contain small-to-medium amounts of boehmite along with the gibbsite. Typical bauxites which are used for aluminum production contain 40–60% total alumina (Al_2O_3), 1–15% total silica (SiO_2), 7–30% total Fe_2O_3, 1–5% titania (TiO_2), and 12–30% combined H_2O.

The suitability of a bauxite as a raw material for aluminum production depends not only on its alumina content but also on its content of combined silica, which is usually in the form of the mineral kaolinite. Kaolinite not only contains aluminum that cannot be extracted in the Bayer process, but also reacts with the sodium–aluminate solution to cause a loss of sodium hydroxide. In some bauxites, the clay minerals are concentrated in the fine-particle size range. In this case, if the economics are favorable, the bauxite can be beneficiated by washing it on screens to remove substantial amounts of the clay and thus make a product containing higher alumina and lower combined silica contents than the original bauxite (**Fig. 1**).

History. H. C. Oersted probably prepared the first metallic aluminum (impure) in 1824 by reducing aluminum chloride with potassium amalgam. F. Wöhler is generally credited with the discovery of the metal by virtue of the first isolation of less impure aluminum in 1827 and the first description of a number of its properties. The metal remained a laboratory curiosity until 1854, when Henri Sainte-Claire Deville improved the earlier method of preparation by employing sodium as reductant and was successful in producing larger quantities of relatively pure metal. The first electrolytic preparation of aluminum, which also occurred in 1854, was accomplished by electrolysis of fused sodium aluminum chloride independently in the laboratories of Deville in France and of R. Bunsen in Germany. *See* ELECTROLYSIS.

The modern industrial electrolytic method of production was discovered simultaneously and independently by Charles Martin Hall in the United States

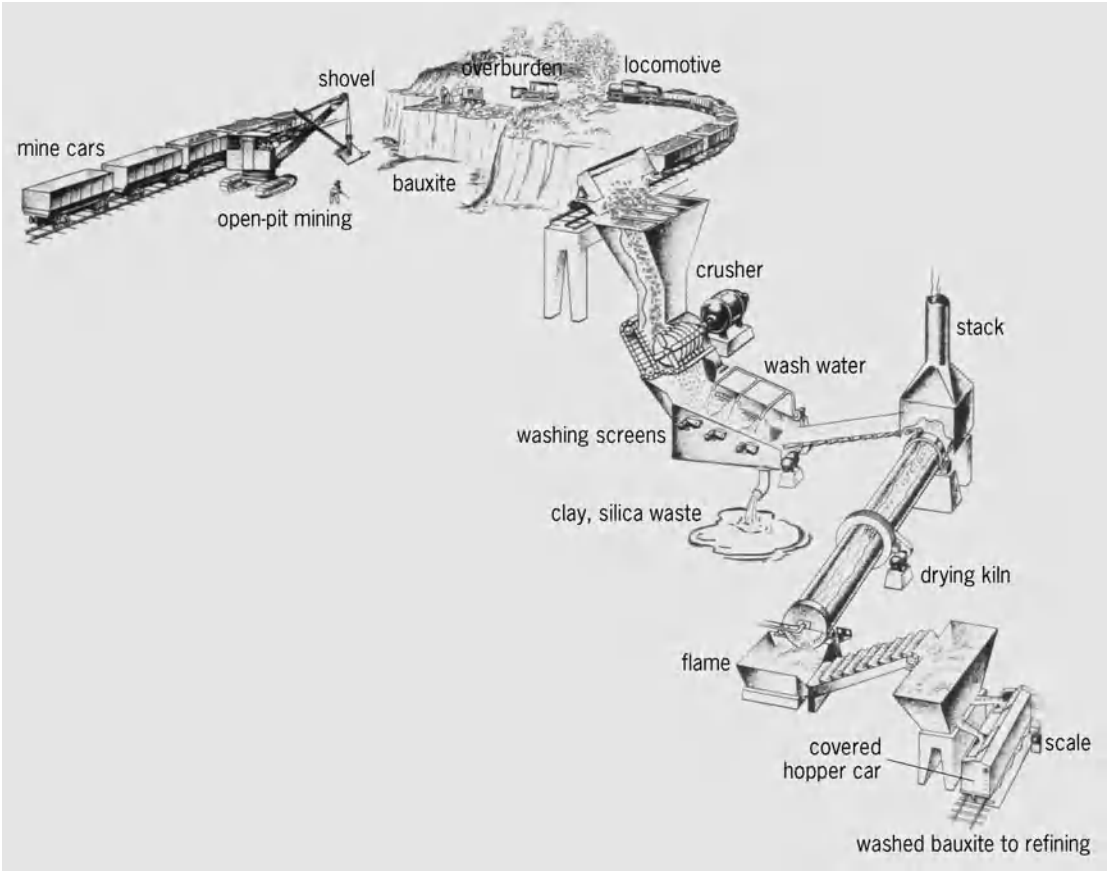

Fig. 1. Bauxite washing. (*Aluminum Company of America*)

and Paul-Louis Héroult of France in 1886. The essentials of their discovery remain the basis for today's aluminum industry. Alumina dissolved in a fluoride fusion (largely cryolite, Na_3AlF_6) is electrolyzed with direct current. Carbon dioxide is discharged at the anode, while the metal is deposited on molten aluminum lying on the carbon lining at the bottom of the cell. The technology for extracting aluminum from its ores was further improved in 1888, when Karl Josef Bayer patented in Germany a method for making pure alumina from bauxite.

Production. The most widely used technology for producing aluminum involves two steps: extraction and purification of alumina from ores, and electrolysis of the oxide after it has been dissolved in fused cryolite. Bauxite is by far the most used raw material. Technically feasible processes exist for extracting alumina from other raw materials such as clays, anorthosite, nepheline syenite, and alunite; however, these processes have not been competitive with the Bayer process in the United States. With the present rapidly changing economics of bauxite production, some of these processes may become competitive in the future. A few of these alternate raw materials are used as a source for alumina in Europe and Asia.

Alumina extraction. A general flow sheet for the Bayer process is shown in **Fig. 2.** In the process, bauxite is crushed and ground, then digested at elevated temperature (280–450°F or 140–230°C) and pressure in a strong solution of caustic soda (80–110 g Na_2O/liter or 0.7–0.9 lb Na_2O/gal). For monohydrate-type bauxites, in which the alumina occurs in forms which are more difficult to dissolve than in trihydrate-type bauxites, stronger solutions (up to 220 g Na_2O/liter or 1.8 lb Na_2O/gal), higher temperatures (up to 570°F or 300°C) and pressures (as high as 150 atm or 1.52×10^7 pascals), and sometimes longer digestion times are required. The gibbsite, boehmite, or diaspore in the bauxites reacts with the sodium hydroxide to form soluble sodium aluminate. The residue, known as red mud, contains the insoluble impurities and the sodium aluminum silicate compound, referred to as desilication product, formed by the reaction of clay minerals with the sodium aluminate–sodium hydroxide solution. The red mud is separated from the solution by countercurrent decantation and filtration. After cooling, the solution is supersaturated with respect to alumina. It is seeded with recycled synthetic gibbsite (alumina trihydrate) and agitated. A large part of the alumina in solution thus crystallizes out as gibbsite. This gibbsite is classified into product and seed, the seed being recycled and the product washed. Wash water and spent liquor, after concentration by evaporation, are recycled to the digestion system. For metal production, the product is calcined at temperatures up to 2400°F (1300°C) to produce alumina containing about 0.3–0.8% soda, less than 0.1% iron oxide plus silica, and trace amounts of other oxides. Some of the precipitated gibbsite is sold for use in the chemicals industry. Specially precipitated pigment-grade gibbsite is used for rubber,

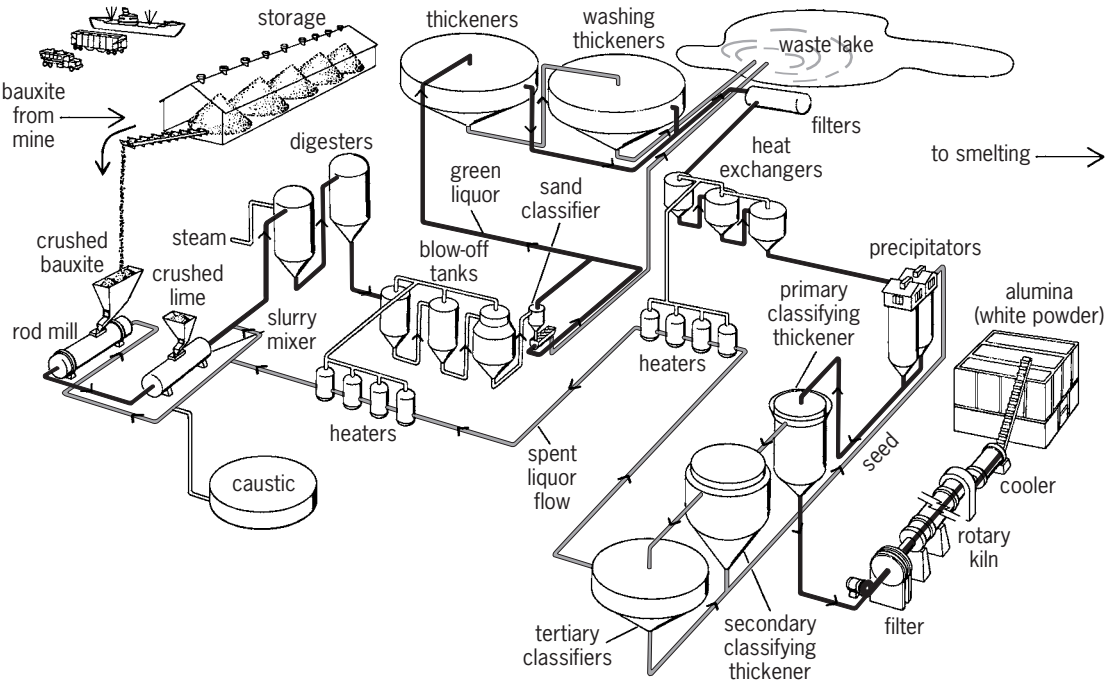

Fig. 2. Processing of alumina from bauxite. (*Aluminum Company of America*)

plastic and paper fillers, and paper coating. Activated aluminas having high internal surface area are made from the gibbsite by calcination at low-to-moderate temperatures. They are used as desiccants and catalysts. Fully calcined aluminas are used in the production of ceramics, abrasives, and refractories.

Alumina can be recovered from low-grade bauxites containing high concentrations of combined silica by the so-called combination or lime–soda sinter process. The process consists of treatment by the Bayer process, following which the red mud is mixed with calcium oxide and sodium carbonate and calcined. This treatment forms sodium aluminate from the aluminous phases in the red mud. The sodium aluminate is then leached out with water, and alumina is recovered from the resulting solution.

Processes for treating other ores have been developed which consist of calcination or fusion of the ore with limestone to make calcium aluminate, from which alumina is leached out with sodium carbonate solution and recovered. Such a process is used in Russia for treating nepheline syenite to yield alumina, and cement as a by-product.

High-iron bauxites have been smelted with limestone in the Pedersen process to yield pig iron and a calcium aluminate slag that can be leached with sodium carbonate solution to recover alumina.

Electrolytic reduction (smelting). Although unchanged in principle, the smelting process of today differs in detail and in scale from the original process discovered by Hall and Héroult. Technology has effected substantial improvements in equipment, materials, and control of the process, and has lowered the energy and labor requirements and the final cost of primary metal.

In a modern smelter (**Fig. 3**), alumina is dissolved in cells (pots)—rectangular steel shells lined with carbon—containing a molten electrolyte (bath) consisting mostly of cryolite. The bath usually contains 2–8% alumina. Excess aluminum fluoride and calcium fluoride are added to lower the melting point and to improve operation. Carbon anodes are hung from above the cells with their lower ends extending to within about 1.5 in. (3.8 cm) of the molten metal, which forms a layer under the molten bath. The heat required to keep the bath molten is supplied by the electrical resistance of the bath as current passes through it. The amount of heat developed with a given current depends on the length of the current path through the electrolyte, that is, anode–cathode distance, which is adjusted to maintain the desired operating temperature, usually 1760–1780°F (960–970°C). A crust of frozen bath, 1–3 in. (2.5–7.5 cm) thick, forms on the top surface of the bath and on the sidewalls of the cell. Alumina is added to the bath or on the crust, where its sorbed moisture is driven off by heat from the cell. While preheating on the crust, the alumina charge serves as thermal insulation. Periodically, the crust is broken and the alumina is stirred into the bath to maintain proper concentration. *See* CRYOLITE.

The passage of direct current through the electrolyte decomposes the dissolved alumina. Metal is deposited on the cathode, and oxygen on the gradually consumed anode. About 0.5 lb (0.23 kg) of carbon is consumed for every pound of aluminum produced. The smelting process is continuous. Alumina is added, anodes are replaced, and molten aluminum is periodically siphoned off without interrupting current to the cells.

Current efficiencies in the industrial electrolytic process are about 85–92%, and the energy efficiency is about 40%. The voltage at the cell terminals is 4–6 V, depending on the size and condition of the

cell. Voltage is required to force the current through the entire cell, and the corresponding power (voltage × amperage) is largely converted into heat in the bath. The amount of power required to maintain the temperature is a smaller proportion of the total power input in large cells than in small ones because of the lower ratio of surface to volume. Thus, power consumed per pound of metal is somewhat less in large than in small cells. Consumption of 6–8 kWh/lb (13–18 kWh/kg) of aluminum produced includes bus-bar, transformer, and rectifier losses.

A potline may consist of 50–200 cells with a total line voltage of up to 1000 V at current loads of 50,000–225,000 A. Electric power is one of the most costly raw materials in aluminum production. Aluminum producers have continually searched for sources of cheap hydroelectric power, but have also had to construct facilities that produce power from fossil fuels. In the past half century, technological advances have significantly reduced the amount of electrical energy necessary to produce a pound of aluminum.

Current is led out of the cell to the anode bus-bar by a number of carbon block anodes suspended in parallel rows on vertical conducting rods of copper or aluminum. Because impurities in the anodes dissolve in the bath as they are consumed, pure carbon (calcined petroleum coke or pitch coke) is used as raw material. The ground coke is mixed hot with enough coal tar or petroleum pitch to bond it into a block when pressed in a mold to form the "green anode." This is then baked slowly at temperatures up to 2000–2200°F (1100–1200°C). In a cavity molded in the top of each block, a steel stub is embedded by casting molten iron around it or by using a carbonaceous paste; the conducting bar is bolted to this stub. Such an electrode is termed a prebaked anode to distinguish it from the Soderberg anode, in which the electrode (single large anode to a cell) is formed in place from a carbonaceous paste that is baked by heat from the pot as it gradually descends into the electrolyte.

The steel shell of the pot is thermally insulated and lined with carbon. The carbon bottom, covered with molten aluminum, serves as cathode of the cell. Electrical connection is made to the carbon cathode by steel bars running through the base of the cell and embedded in the carbon lining.

Molten aluminum is siphoned from the smelting cells into large crucibles. From there the metal may be poured directly into molds to produce foundry ingot, or transferred to holding furnaces for further refining or alloying with other metals to form fabricating ingot. As it comes from the cell, primary aluminum averages about 99.8% purity.

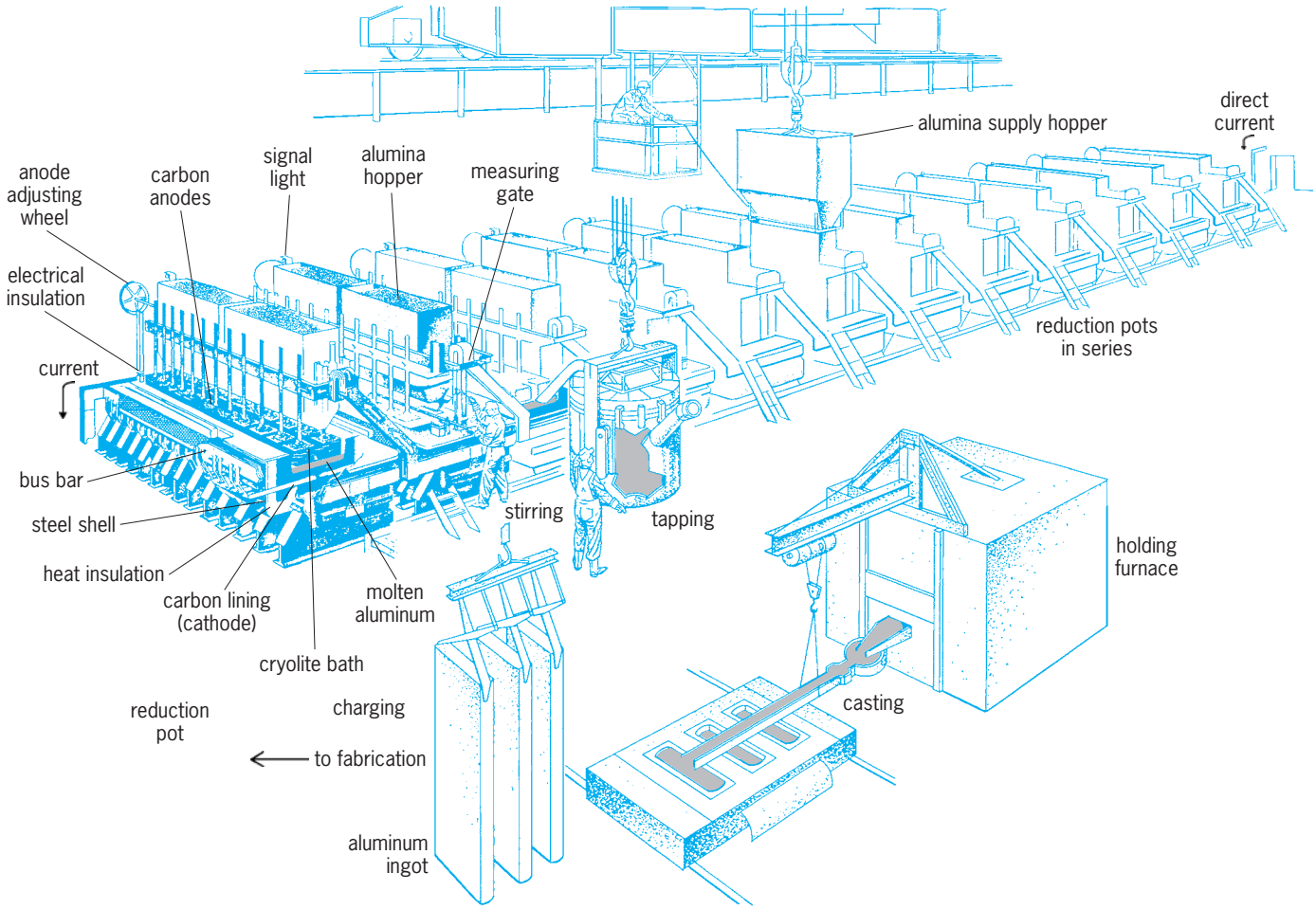

Fig. 3. Production of aluminum metal. (*Aluminum Company of America*)

See ELECTROMETALLURGY; PYROMETALLURGY, NONFERROUS.

Melting. In plants not adjacent to a smelter, it is necessary to remelt charges, usually consisting of mill scrap or returned scrap with enough primary metal to provide composition control. For casting ingot for fabrication, gas- or oil-fired reverberatory furnaces are commonly used. Molten metal is generally transferred from the melting furnace to a reverberatory holding furnace or to a ladle for subsequent casting. Metal for foundry use in producing sand, permanent mold, and die castings is usually remelted in gas- or oil-fired crucible furnaces, although in large-scale operations reverberatory furnaces may be used.

Scrap recycling has grown through the years to considerable economic importance, and provides a valuable source of aluminum at a much lower energy expenditure than for primary metal. For many years, secondary producers have purchased scrap and reclaimed it, generally into foundry ingot for remelting. Primary producers have also purchased large amounts of mill scrap, as well as discarded cans, and recycled it into ingot used to fabricate sheet for more cans. This closed-circuit type of recycling saves energy, helps to preserve the environment, and conserves metal to the greatest degree.

After remelting, the molten aluminum alloy must be treated to remove dissolved hydrogen, inclusions such as oxides formed during remelting or from the surface of the charge material, and undesirable trace elements, such as sodium, in some alloys. Classically, the practice has been to bubble chlorine or a mixture of nitrogen and chlorine through the melt by means of graphite tubes. Modern practice makes use of a wide variety of in-line systems in which the metal is treated by filtering or fluxing, or both, during transfer between the holding furnace and the casting unit. These processes are more efficient than furnace fluxing, and their use minimizes air pollution, improves metal quality, and saves production time.

Although many ingot casting methods have been used historically, today most ingots for fabricating are cast by the direct chill process or some modification thereof. This process is accomplished on a semicontinuous basis in vertical casting and on a continuous basis in horizontal casting. The direct chill process consists of pouring the metal into a short mold. The base of the mold is a separate platform that is lowered gradually into a pit while the metal is solidifying. The frozen outer shell of metal retains the still-liquid portion of the metal when the shell is past the mold wall. Water is used to cool the mold, and water impingement on the ingot shell also cools the ingot as it is lowered. Ingot length is limited only by the depth to which the platform can be lowered. In horizontal casting, the same principles are applied, but the mold and base are turned so that the base moves horizontally, removing the constraint of pit depth. Much longer ingot can be practically cast by the horizontal method. By sawing the ingot while casting is progressing, the process can be made continuous.

Fabrication methods. Aluminum alloys are fabricated by all the methods known to metalworking and are commercially available in the form of castings; in wrought forms produced by rolling, extruding, forging, or drawing; in compacted and sintered shapes made from powder; and in other monolithic forms made by a wide variety of processes. Some of the common fabrication processes are represented in **Fig. 4**. In addition, pieces produced separately can be assembled by common joining procedures to build up complex shapes and structures. *See* ALUMINUM ALLOYS; METAL FORMING.

Casting. In casting aluminum alloys, molten metal is poured, or forced by pressure, into a mold where it solidifies. The temperature is usually 1150–1400°F (620–760°C) depending on the alloy composition and on the type of casting. For die casting and permanent mold casting, the molds are metal and are used repeatedly. Other processes use expendable molds made of sand or plaster. *See* METAL CASTING.

Rolling. Rolling is the process of reducing material by passing it between pairs of rolls. Aluminum alloy rolled products include plate (0.250 in. or 0.6 cm or more thick), sheet (0.006–0.249 in. or 0.02–0.632 cm thick), and foil (less than 0.006 in. thick). The starting product is ingot, which may be up to 360 in. (914 cm) long, 72 in. (182 cm) wide, and 26 in. (66 cm) thick. Initial reduction is hot (800–1100°F, or 426–593°C, depending on alloy) and is done in a reversing mill. Subsequent rolling operations may be in multistand continuous mills or single-stand mills, and may be hot or cold depending on the thickness and properties required in the final product. Other aluminum alloy rolled products include rod and bar, which are passed through grooved rolls. *See* METAL ROLLING.

Extruding. In extruding aluminum alloys, an ingot or billet is forced by high pressure to flow from a container through a die opening to form an elongated shape or tube, usually at metal temperatures between 550 and 1050°F (287 and 565°C). Hydraulic presses with capacities of 500 to 14,000 tons (4.5×10^6 to 1.2×10^8 newtons) are used. The extrusion process is capable of producing sections with weights of a few ounces to more than 200 lb/ft (300 kg/m), with thicknesses from a few tenths of an inch to about 10 in. (25 cm), with circumscribing circle diameters of 0.25 in. (0.64 cm) to about 3 ft (0.9 m), and lengths in excess of 100 ft (30 m). Tubing in diameters of 0.25–33 in. (0.64–0.84 cm) and pipe up to 20 in. (50 cm) are also produced. Stock that is produced by extrusion is also fabricated to tubing by drawing and to forgings by subsequent metalworking operations. *See* EXTRUSION.

Forging. Forgings are produced by pressing or hammering aluminum alloy ingots or billets into simple rectangular or round shapes on flat dies and into complex forms in cavity dies. Hydraulic presses capable of forces up to 75,000 tons (6.60×10^8 N) and mechanical presses with capacities up to 16,000 tons are used. Forging hammers have ram weights of 500–110,000 lb (227–50,000 kg). Metal temperatures of 600–880°F (315–471°C) are generally used,

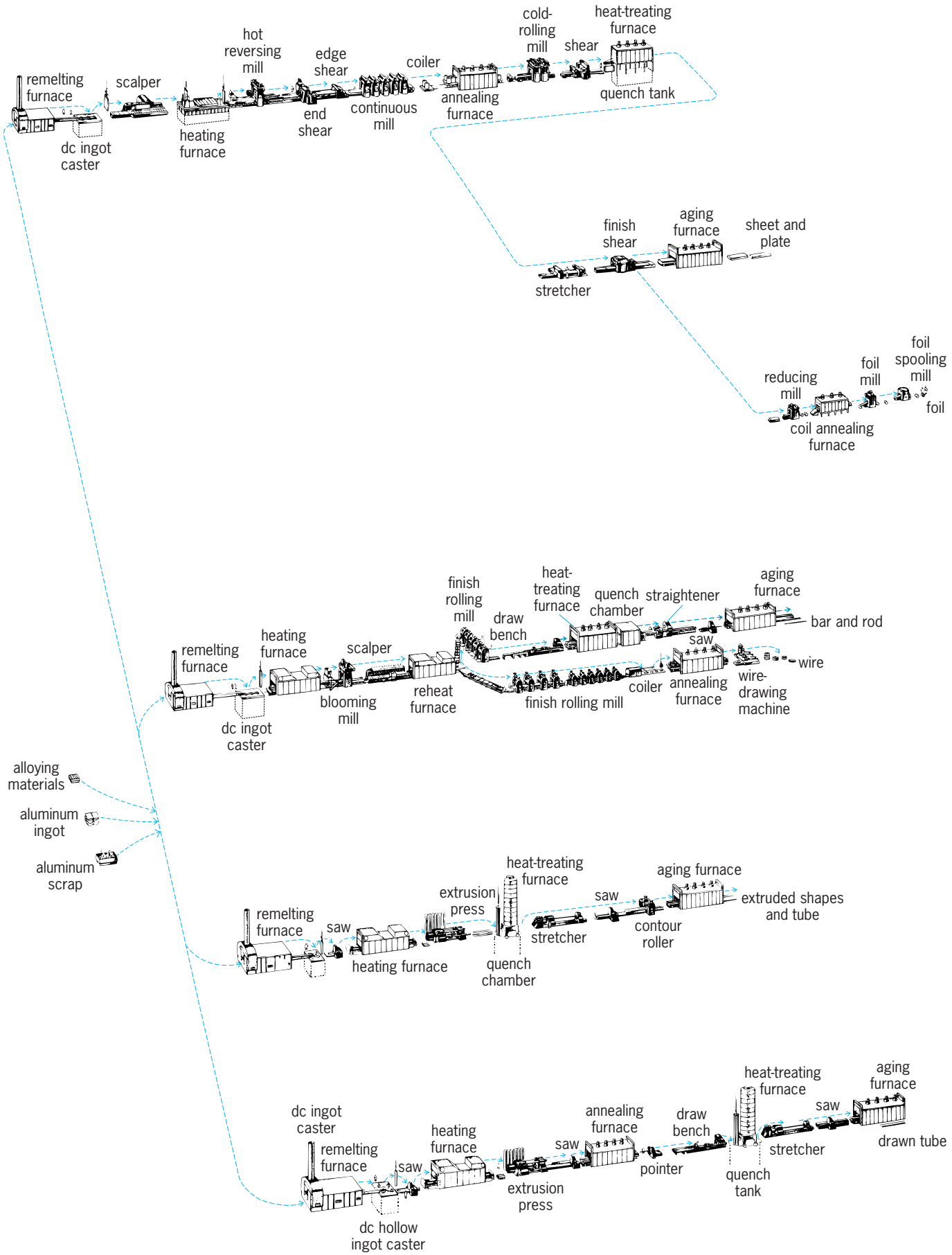

Fig. 4. Fabrication operations. (*Aluminum Company of America*)

depending on alloy and type of forging. *See* FORGING.

Drawing. Aluminum alloys can be made into deep-drawn shapes, of which pots and pans and beverage cans are common examples, by forcing sheet or foil into deep holes or cavities of the desired shape. Other drawn products, such as wire and tubing, are produced by pulling rolled or extruded stock through dies. *See* DRAWING OF METAL.

Compacting and sintering. Aluminum powders produced by atomizing can be compacted, either alone or mixed with powders of other elements, in dies of complex shape. The compact is then heated to a high temperature to promote additional bonding between the particles to result in products with desired shapes. *See* POWDER METALLURGY.

Other processes. Aluminum products are machined to final shape and dimension on lathes, joiners, routers, drill presses, shears, grinders, and many other high-speed metal-cutting and metalworking machines. Other fabricating techniques include impact extrusion, forming, stamping, embossing, coining, bending, rotary swaging, and cold heading.

Joining. Aluminum products formed into shapes by any of the above processes can be joined by welding, brazing, soldering, and adhesive bonding, and by mechanical means such as crimping, seaming, screwing, bolting, and riveting.

Products and uses. Applications in building and construction represent the largest single market of the aluminum industry. Millions of homes use aluminum doors, siding, windows, screening, and downspouts and gutters, which require little maintenance and provide a long life. Aluminum is also a major industrial building product. Excellent weather resistance makes it ideal for all climates and locations. Selected aluminum alloys are suitable near the seacoast, where salt spray may be deleterious to other metals. Colored and given additional protection by an electrolytic process called anodizing, aluminum appears in the curtain wall construction of many of the world's tallest buildings.

Transportation is the second largest market. In automobiles, aluminum is apparent in interior and exterior trim, grilles, wheels, and air conditioners. Another major use is in automatic transmissions. Some auto radiators, engine blocks, and body panels are made of aluminum to hold down weight and improve fuel economy. Aluminum is also found in rapid-transit car bodies, engine parts for diesel locomotives, rail freight and tank cars, bus and truck engines, forged truck wheels, cargo containers, and in highway signs, divider rails, and lighting standards. Aluminum's light weight and corrosion resistance are responsible for applications in pleasure-boat hulls, fishing boats, tanks for ships transporting liquefied natural gas, and deck houses for naval vessels.

In aerospace, aluminum is found in aircraft engines, frames, skins, landing gear, and interiors. Toughness, lightness, and heat-reflective characteristics have made it the preferred material for satellites and other spacecraft. *See* AIRFRAME; SPACECRAFT STRUCTURE.

In the packaging industry, aluminum helps in the preparation of foods and beverages and keeps them pure during distribution and storage. The fast-cooling, easy-opening, lightweight, and recyclable features of the all-aluminum can have resulted in the use of billions of aluminum beverage containers. Foil pouches and bags, twist-off closures, and easy-open ends revolutionized the food and beverage packaging industries. *See* FOOD ENGINEERING.

In electrical applications, aluminum wire and cable are major products. Underground electrical cables require large amounts of aluminum. Aluminum wiring is also used in residential, commercial, and industrial buildings. *See* TRANSMISSION LINES; WIRING.

Consumer use began before the 1900s with aluminum cooking utensils. In modern homes, aluminum appears as cooking foil, hardware, tools, portable appliances, air conditioners, freezers, and refrigerators.

There are hundreds of chemical uses of aluminum and aluminum compounds. Aluminum powder is used in paints, rocket fuels, and explosives, and as a chemical reductant. Allen S. Russell

Bibliography. Aluminum Corporation of America, *Melting and Casting Aluminum*, 1995; G. S. Brady, H. R. Clauser, and J. A. Vaccari, *Materials Handbook*, 15th ed., 2002; J. E. Hatch (ed.), *Aluminum: Properties and Physical Metallurgy*, 1984; G. E. Totten and D. S. MacKenzie (eds.), *Handbook of Aluminum*, vol. 1: *Physical Metallurgy and Processes*, 2003.

Aluminum alloys

Substances formed by the addition of one or more elements, usually metals, to aluminum. The principal alloying elements in aluminum-base alloys are magnesium (Mg), silicon (Si), copper (Cu), zinc (Zn), and manganese (Mn). In wrought products, which constitute the greatest use of aluminum, the alloys are identified by four-digit numbers of the form NXXX, where the value of N denotes the alloy type and the principal alloying element(s) as follows: 1 (Al; at least 99% aluminum by weight), 2 (Cu), 3 (Mn), 4 (Si), 5 (Mg), 6 (Mg + Si), 7 (Zn), 8 (other). *See* COPPER; MAGNESIUM; MANGANESE; SILICON; ZINC.

Iron and silicon are commonly present as impurities in aluminum alloys, although the amounts may be controlled to achieve specific mechanical or physical properties. Minor amounts of other elements, such as chromium (Cr), zirconium (Zr), vanadium (V), lead (Pb), and bismuth (Bi), are added to specific alloys for special purposes. Titanium additions are frequently employed to produce a refined cast structure. *See* IRON; TITANIUM.

Aluminum-base alloys are generally prepared by making the alloying additions to molten aluminum, forming a liquid solution. As the alloy freezes, phase separation occurs to satisfy phase equilibria requirements and the decrease in solubility as the temperature is lowered. The resultant solidified structure consists of grains of aluminum-rich solid solution

TABLE 1. Nominal composition and forming processes for common wrought aluminum alloys

Alloy	Form*	Si	Cu	Mg	Mn	Zn	Other
1100	b–d	—	—	0.1	—	—	99.00 Al (min.)
1350	b–d	—	—	—	—	—	99.5 Al (min.)
2011	b–d	—	5.5	—	—	—	0.5 Pb, 0.5 Bi
2014	b–e	0.8	4.4	0.4	0.8	—	—
2024	b–e	—	4.4	1.5	0.6	—	—
2036	b	—	2.6	0.45	0.25	—	—
2219	b, e	—	6.3	—	0.3	—	0.1 V, 0.1 Zr
3003	b–d	—	—	—	1.2	—	—
3004	b–d	—	—	1.0	1.2	—	—
5052	b–d	—	—	2.5	—	—	0.25 Cr
5083	b, e	—	—	4.5	0.7	—	0.15 Cr
5086	b–d	—	—	4.0	0.5	—	0.15 Cr
5182	b	—	—	4.5	0.35	—	—
5657	b, c	—	—	0.8	—	—	—
6061	b–e	0.6	0.25	1.0	—	—	0.25 Cr
6063	b–e	0.4	—	0.7	—	—	—
7005	b–e	—	—	1.5	0.5	4.5	0.15 Cr, 0.14 Zr
7050	b, e	—	2.4	2.3	—	6.2	0.12 Zr
7075	b–e	—	1.6	2.5	—	5.6	0.25 Cr

*b = rolled; c = drawn; d = extruded; and e = forged.

and crystals of intermetallic compounds. Elements which lower the freezing point of aluminum, such as copper, magnesium, and silicon, tend to segregate to the portions of the grains which freeze last, such as the cell boundaries. Elements which raise the freezing point, such as chromium or titanium (Ti), segregate in the opposite manner. *See* EUTECTICS; SOLID SOLUTION.

A decrease in solubility with falling temperature also provides the basis for heat treatment of solid aluminum alloys. In this operation, the alloy is held for some time at a high temperature to promote dissolution of soluble phases and homogenization of the alloy by diffusion processes. The limiting temperature is the melting point of the lowest melting phase present. The time required depends both on temperature and on the distances over which diffusion must occur to achieve the desired degree of homogenization. Times of several hours can be necessary with coarse structures such as sand castings. Only a few minutes may be adequate, however, for rapidly heated thin sheet. *See* HEAT TREATMENT (METALLURGY).

The solution heat treatment is followed by a quenching operation in which the article is rapidly cooled, for example, by plunging it into cold or hot water or by the use of an air blast. This produces a supersaturated metallic solid solution that is thermodynamically unstable at room temperature. In several important alloy classes, such as 2XXX, 6XXX, and 7XXX, the supersaturated solution decomposes at room temperature to form fine, submicroscopic segregates or precipitates that are precursors to the equilibrium phases predicted by phase diagrams. The precipitation phenomenon, occurring over periods of days to years, produces substantial increases in strength. Additional precipitation strengthening can be obtained by heating the alloy at temperatures in the range 250–450°F (121–232°C), the time and temperature varying with alloy composition and the objectives with respect to mechanical properties and other characteristics such as corrosion resistance.

Wrought alloys. **Table 1** lists the nominal compositions of a number of commercially important wrought alloys and the type of products for which they are used. The alloys are generally classified in two broad categories depending upon their response to heat treatment as described in the preceding paragraph. Those having no or minor response are identified as non-heat-treatable alloys and include the 1XXX, 3XXX, and 5XXX compositions. Those that do respond are known as heat-treatable alloys and include the 2XXX, 6XXX, and 7XXX compositions.

Included in the non-heat-treatable group is 1350, a special grade used for electrical conductor products. Alloy 1100 is a grade of 99.0% minimum aluminum content with particular controls on iron (Fe), silicon, and copper contents, available in a variety of product forms such as sheet, foil, wire, rod, and tube, used for packaging, fin stock, and a variety of sheet metal applications. The manganese-containing alloy 3003 is a moderate-strength, very workable alloy for cooking utensils, tube, packaging, and lithographic sheet applications. The stronger aluminum-manganese-magnesium alloy 3004 is used for architectural applications, for storage tanks, and especially for drawn and ironed beer and beverage containers.

Alloy 5052 is a workable, corrosion-resistant aluminum-magnesium alloy for many metalworking and metal-forming purposes and marine applications. Where higher strength is required, the higher-Mg-content, weldable 5086 or 5083 alloys may be used. The latter is employed in construction of welded tanks for liquefied gas (cryogenic) transport and storage and for armor plate in military vehicles. Alloy 5182 is also a high-strength aluminum-magnesium alloy that is employed primarily in a highly strain-hardened condition for beverage can ends. Alloy 5657 is a lower-strength material

TABLE 2. Nominal composition and casting procedure for common aluminum casting alloys

Alloy	Form*	Si	Cu	Mg
413.0	D	12		
B†443.0	B, C	5.3	.15 max	
F†332.0	C	9.5	3.0	1.0
355.0	B, C	5.0	1.3	0.5
356.0	B, C	7.0		0.3
380.0	D	8.5	3.5	
390.0	C, D	17	4.5	0.55

* B = sand casting; C = permanent mold casting; and D = die casting.
† The letter indicates modifications of alloys of the same general composition or differences in impurity limits, from alloys having the same four-digit numerical designations.

produced with a bright anodized finish for automobile trim and other decorative applications.

The heat-treatable alloys 2014, 2024, and 7075 have high strengths and are employed in aircraft and other transportation applications. Modifications of these basic alloys, such as 2124 and 7475, were developed to provide increased fracture toughness. High toughness at high strength levels is achieved in thick-section products with 7050. Where elevated temperatures are involved, the 2XXX alloys are preferred. One such alloy, 2219, also has good toughness at cryogenic temperatures and is weldable. This alloy was prominently employed in the fuel and oxidizer tanks serving as the primary structure of the Saturn space vehicle boosters.

Alloy 6061 is used for structural applications where somewhat lower strengths are acceptable. For example, 6061 may be used for trailers, trucks, and other transportation applications. Alloy 6063 is a still-lower-strength, extrudable, heat-treatable alloy for furniture and architectural applications. The use of lead and bismuth in 2011 produces a heat-treatable, free-machining alloy for screw-machine products. Alloy 2036 is a moderate-strength alloy with good workability and formability, employed as body sheet in automotive applications.

Casting alloys. **Table 2** shows the nominal compositions of several important casting alloys. The major alloying addition is silicon, which improves the castability of aluminum and provides moderate strength. Other elements are added primarily to increase the tensile strength. Most die castings are made of alloy 413.0 or 380.0. Alloy 443.0 has been very popular in architectural work, while 355.0 and 356.0 are the principal alloys for sand castings. Number 390.0 is employed for die-cast automotive engine cylinder blocks, while alloy F332.0 is used for pistons for internal combustion engines.

Casting alloys are significant users of secondary metal (recovered from scrap for reuse). Thus, casting alloys usually contain minor amounts of a variety of elements; these do no harm as long as they are kept within certain limits. The use of secondary metal is also of increasing importance in wrought alloy manufacturing as producers take steps to reduce the energy that is required in producing fabricated aluminum products. *See* ALLOY; ALLOY STRUCTURES; METAL CASTING.

Allen S. Russell

High-strength alloys. Since aluminum comprises 70–80% of the weight of an airframe, metallurgists have been pursuing aluminum alloy development programs directed toward producing materials which would be characterized by stronger, stiffer, and lighter-weight properties. In addition, the titanium alloys of aircraft gas turbines are prime targets for replacement by lighter-weight alloys. Researchers have improved aluminum alloys by using techniques that add lithium and by using three processing technologies comprising powder metallurgy, mechanical alloying, and a process which involves blending aluminum alloy powders and silicon carbide fibers to form a composite.

Aluminum-lithium alloys. Lithium is the lightest metal in existence. For each weight percent of lithium added to aluminum, there is a corresponding decrease of 3% in the alloy's weight. Therefore, aluminum-lithium alloys offer the attractive property of low density. A second beneficial effect of lithium additions is the increase in elastic modulus (stiffness). Also, as the amount of lithium is increased, there is a corresponding increase in strength due to the presence of very small precipitates which act as strengthening agents to the aluminum. As the precipitates grow during heat treatment, the strength increases to a limit, then begins to decrease. Aluminum-lithium alloys therefore come under the classification of precipitation-strengthening alloys. Since the size and distribution of the precipitates can be controlled by heat treating, these alloys are also referred to as heat-treatable.

Since the addition of lithium to aluminum has also been shown to result in an alloy with unacceptable levels of ductility, a necessary engineering property for many aerospace applications, other elements such as copper, magnesium, and zirconium have been added to offset the loss in ductility. However, these alloy additions, particularly copper, increase the density. Consequently, schemes of alloy development have focused on balancing the various positive and negative attributes of the different elements, to arrive at a composition with suitable properties. **Figure 1** is a schematic relating yield strength and

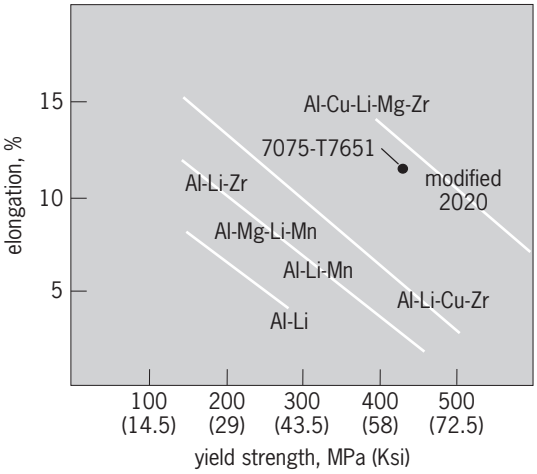

Fig. 1. Strength–ductility relationships for various Al-Li-X alloys. The alloy 7075-T7651 is included as a reference point.

elongation for a variety of aluminum-lithium alloys. The alloy 7075–T7651, commonly used for such applications, is included as a reference point. In order to be competitive, the new alloys should have strength-ductility relations equal to or better than 7075. *See* LITHIUM.

Powder metallurgy. As early as 1947, investigators showed interest in the development of high-modulus, high-temperature aluminum alloys. Elements such as iron, cobalt, and nickel have been added to aluminum in large quantities to produce alloys which have relatively stable structures at elevated temperatures and higher elastic moduli than conventional aluminum alloys.

The titanium in the fan sections of gas turbines has to withstand operating temperatures up to 450°F (235°C). The substitution of a suitable aluminum alloy for titanium could result in a 31% reduction in weight for that part. In order to develop aluminum alloys with sufficient strength to withstand the operating temperatures of a gas turbine and remain thermally stable, researchers have considered alloy additions to aluminum which move slowly (have low diffusivities) in it, so that the characteristics of that structure do not change with time at the service temperature. Since the elements which have low diffusivities in aluminum also have very limited solubilities, processes other than conventional ingot casting must be utilized.

An alternative for producing a candidate alloy is the powder metallurgy process, which is a rapid solidification process. Rapid solidification processing involves the transformation of finely divided liquid either in the form of droplets or thin sheets which become solid at high solidification rates (on the order of 10^4 K/s). Solidification occurs as heat is removed from the molten metal. The primary mechanisms of heat transfer can be divided into two broad categories—conductive and convective. Certain rapid-solidification-process forms such as ribbons and fibers are cooled primarily by conduction. The heat is transferred away by a substrate—the surface on which the liquid is cast. The substrate is generally made out of copper and can be either a rotating drum or a rotating wheel. On the other hand, another rapid-solidification-process form, powder, is cooled primarily by convection. The heat is transferred away from the fine liquid droplets in a gas stream. Regardless of the mechanisms of heat transfer, the resulting structures are more homogeneous and are on a finer scale than those formed by conventional ingot-casting techniques. The rapidly solidified particulates are then consolidated by forging or extrusion into a final shape. Even after consolidation, the various rapid-solidification-process forms are more homogeneous and have finer microstructures than a corresponding ingot-cast and worked product.

Research centered on utilizing the natural benefit of rapid-solidification processing has produced metastable microstructures which decompose and precipitate during consolidation and working. Thus elements which have limited solubility and low diffusivity in aluminum (such as iron, cobalt, and nickel—the transition metals), in conjunction with elements such as molybdenum or cerium, can be used to maximum advantage. However, the use of these alloying elements necessitates control of thermal history of the particulate both in its formation and upon subsequent processing. For example, in aluminum–transition metal alloys, the precipitates can form directly from the melt rather than from the solid. Consequently, the precipitates in these systems tend to be coarser than those in aluminum-lithium systems.

As in the case of aluminum-lithium alloys, the yield strength is affected primarily by the interparticle spacing and the fineness of the precipitates which form. The ductility and fracture toughness are affected by macroparticles that concentrate stress, thereby reducing the toughness. Since ductility and fracture toughness are important properties, rapid-solidification-process alloys will be limited until a control of the volume fraction of coarse particles can be achieved.

Rapid-solidification-process aluminum alloys containing iron and cerium appear to result in alloys which have refined particle-size distribution and improved high-temperature stability. These powder metallurgy alloys appear competitive with titanium up to 375°F (191°C). *See* METALLIC GLASSES; POWDER METALLURGY.

Mechanical alloying. This process circumvents the limitations of conventional ingot casting. Blends of powders are mixed in a ball mill. A drum is mounted horizontally and half-filled with steel balls and blends of elemental metal powders. As the drum rotates, the balls drop on the metal powder. The degree of homogeneity of the powder mixture is determined by the size of the particles in the mixture. If the powder's particles are too coarse, the different constituent elements in the blend will not interdiffuse during consolidation. Consequently, high rates of grinding must be used.

In a high-energy mill the particles of metal are flattened, fractured, and rewelded. Every time two balls collide, they trap powder particles between them. The impact creates atomically clean fresh surfaces. When these clean surfaces meet, they reweld. Since such surfaces readily oxidize, the milling operation is generally conducted in an inert atmosphere. During the initial stages of the mechanical alloying process, the particles are layered composites of the constituent powders and the starting constituents are easily identifiable within the composite particle. The composition and shape of the individual particles are different. As alloying progresses, the particles are further fractured and rewelded. The intimate mixture of the starting constituents decreases the diffusion distance, and the particle tends to become more homogeneous. Metastable and stable phases are beginning to form in the individual particles. Diffusion of the elements is accelerated by the strain introduced during the milling. During the final stage of processing, the individual particles became similar in composition and similar to one another. Completion of the process occurs when the tendency to fracture is balanced by the tendency to

TABLE 3. Data for several loading fractions of silicon carbide (SiC)

Tensile properties of SiC/2024 Al-T4 composites[*]				
Vol. % SiC:	0	15	20	25
Ultimate tensile strength, MPa (Ksi)[†]	400–434 (58–64)	NA[‡]	455–524 (66–76)	552–641 (80–93)
Elongation, GPa (Msi)[§]	73 (10.6)	89–97 (13–14)	97–117 (14–17)	117–151 (17–22)
Elongation to failure, %	19–20	1–2	1–2	1–2

Tensile properties of SiC/2024 Al composites, as extruded[¶]			
Vol. % SiC:	0	20	25
Ultimate tensile strength, MPa (Ksi)[†]	186 (27)	331–372 (48–54)	400–448 (58–65)
Elongation, GPa (Msi)[§]	73 (10.6)	97–117 (14–17)	117–151 (17–22)
Elongation to failure, %	20–22	1–2	1–2

[*]18 specimens. [†]Ksi = thousand pounds per square inch. [‡]Not available. [§]Msi = million pounds per square inch. [¶]20 specimens.

weld and the particle size distribution is constant. The relationship between the thickness and layers within the composite is schematically illustrated in **Fig. 2.**

Aluminum–silicon carbide composites. A composite is a material in which two or more constituents are combined to result in a material which has properties different from those of either constituent. Typical composites are from materials in which one of the components has very high strength and modulus and the other has high ductility. Their properties generally follow a rule of mixtures. For example, if elastic modulus is the property of interest, the elastic modulus of the composite is approximately the weighted sum of the elastic moduli of the constituents.

In one example, silicon carbide whiskers are mixed with aluminum alloy powder, compacted under pressure at elevated temperatures, and extruded or forged into a final product. The very high-modulus silicon carbide is incorporated in the ductile aluminum alloy matrix. The resulting properties depend on the volume fraction of silicon carbide, and to a large degree on the fabricating methods. The primary advantage of an aluminum–silicon carbide composite is the high elastic modulus and strength (**Table 3**). *See* ALUMINUM; COMPOSITE MATERIAL.

Tom H. Sanders, Jr.

Bibliography. American Society for Metals, *Source Book on Selection and Fabrication of Aluminum Alloys*, 1978; A. P Divecha, S. G. Fishman, and S. D. Karmarkar, Silicon carbide reinforced aluminum: A formable composite, *J. Metals*, 33:12–17, September 1981; P. S. Gilman and J. S. Benjamin, Mechanical alloying, *Annu. Rev. Mater. Sci.*, 13: 279–300, 1983; F. King, *Aluminum and Its Alloys*, 1987; S. L. Langenbeck, Elevated temperature aluminum alloy development, *Interim Technical Report for Period April–September 1981*, U.S. Airforce Rep. LR 29977, October 1981; E. A. Starke, Jr., T. H. Sanders, Jr., and I. G. Palmer, New approaches to alloy development in the Al-Li system, *J. Metals*, 33:24–33, August 1981.

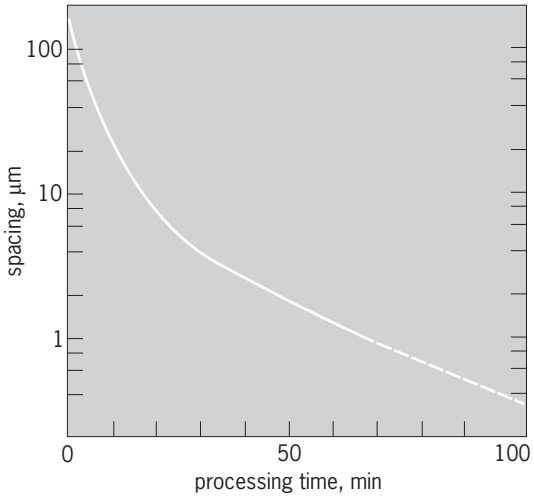

Fig. 2. Effect of processing time on the spacing between the layers of mechanically alloyed powder. The smaller the spacing, the more homogeneous the powder particle.

Alunite

A mineral of composition $KAl_3(SO_4)_2(OH)_6$. Alunite occurs in white to gray rhombohedral crystals or in fine-grained, compact masses. It has a hardness of 4 and a specific gravity of 2.6–2.8.

Alunite is produced by sulfurous vapors on acid volcanic rocks and also by sulfated meteoric waters affecting aluminous rocks. The mineral is generally found associated with quartz and kaolinite.

Alunite is used as a source of potash or for making alum. Alum has been manufactured from the well-known alunite deposits at Tolfa, near Civita Vecchia, Italy, since the mid-fifteenth century. In the United States alunite is widespread in the West. Large deposits occur in Mineral and Hindsdale counties, Colorado, and in Piute County, Utah. *See* ALUM; FERTILIZER; POTASSIUM.

Edward C. T. Chao

Alzheimer's disease

A disease of the nervous system characterized by a progressive dementia that leads to profound impairment in cognition and behavior. Dementia occurs in a number of brain diseases where the impairment in cognitive abilities represents a decline from prior levels of function and interferes with the ability to perform routine daily activities (for example, balancing a checkbook or remembering appointments). Alzheimer's disease is the most common form of dementia, affecting 5% of individuals over age 65. The onset of the dementia typically occurs in middle to late life, and the prevalence of the illness increases with advancing age to include 25–35% of individuals over age 85. *See* AGING.

Clinical characteristics. Memory loss, including difficulty in remembering recent events and learning new information, is typically the earliest clinical feature of Alzheimer's disease. As the illness progresses, memory of remote events and overlearned information (for example, date and place of birth) declines together with other cognitive abilities. Orientation to time (for example, date, day of week, or season) is often impaired, followed by a decline in orientation to personal information and immediate environment. Language deficits usually begin with difficulty in naming objects and in word-finding ability and gradually progress to a loss of all language skills. The ability to use visuospatial skills to draw simple figures and shapes is increasingly impaired. Declines in abstract reasoning, judgment, and problem solving interfere with the ability to perform mathematical calculations and to make appropriate decisions in social situations. In the later stages of Alzheimer's disease, there is increasing loss of cognitive function to the point where the individual is bedridden and requires full-time assistance with basic living skills (for example, eating and bathing). *See* MEMORY.

Although Alzheimer's disease typically has a slowly progressing course, the rate of cognitive decline is variable among individuals. The average duration from the onset of symptoms to death is 8–10 years. In individuals reaching the end stage of dementia, the cause of death is often related to secondary medical conditions (for example, respiratory infections). Behavioral disturbances that can accompany Alzheimer's disease include agitation, aggression, depressive mood, sleep disorder, and anxiety. Changes in personality, increased social withdrawal, and irritability are observed in some individuals. Delusions and hallucinations can occur. The onset of clinical symptoms can vary from middle to advanced age. Although a distinction between individuals with an early onset (65 years of age and younger) and with a late onset (older than 65) of dementia has been made, the same neuropathological disease process occurs in both groups.

Neuropathological features. The major neuropathological features of Alzheimer's disease include the presence of senile plaques, neurofibrillary tangles, and neuronal cell loss. Although the regional distribution of brain pathology varies among individuals, the areas commonly affected include the association cortical and limbic regions. Senile plaques are extracellular and contain a core of an amyloid beta peptide that is derived from the cleavage of a much larger beta amyloid precursor protein. Although the amyloid precursor protein occurs normally in neural cell membranes, the reason why individuals with Alzheimer's disease produce excessive amounts of the amyloid beta peptide is unknown. Neurofibrillary tangles represent intracellular pathology in which paired helical filaments accumulate within neurons. The basic component of neurofibrillary tangles is a hyperphosphorylated tau protein producing a highly insoluble cytoskeletal structure that is associated with cell death. Prominent loss of pyramidal neurons from cerebral cortical layers III and V has been observed in Alzheimer's disease. Other microscopic features include neuropil threads, Hirano bodies, granulovascular degeneration, and inflammation around plaque formations. The brain regions of the medial temporal lobes are commonly affected early in the course of Alzheimer's disease. In vivo neuroimaging studies that measure functional brain activity by assessing cerebral blood flow and metabolism have shown that the temporal, parietal, and frontal association cortices appear most consistently and severely affected in the early to middle stages of the illness.

Alzheimer's disease has been associated with dysfunction of several neurotransmitter systems. These neurochemical systems are responsible for the transfer of information between neurons via the synapse and are important for cognitive and behavioral functions. In Alzheimer's disease, disruption of information transfer between neurons can occur with loss of presynaptic elements and reduction of the recognition sites at which the neurotransmitter acts. It is also thought that loss of or abnormalities in the receptor-operated ion channels and second messenger systems within cells may be related to neuronal dysfunction in Alzheimer's disease. Deficits in cholinergic, serotonergic, noradrenergic, and peptidergic (for example, somatostatin) neurotransmitters have been demonstrated. Dysfunction of the cholinergic neurotransmitter system has been specifically implicated in the early occurrence of memory impairment in Alzheimer's disease, and it has been a target in the development of potential therapeutic agents. The deficiency in choline acetyltransferase, an enzyme important for the synthesis of acetylcholine, has been associated with dementia severity. The cortical projection of the cholinergic system is especially concentrated in key structures such as the hippocampus in the medial temporal lobes, brain structures that are thought to be important in aspects of learning and memory. *See* ACETYLCHOLINE; NEUROBIOLOGY; NORADRENERGIC SYSTEM.

Diagnosis. A definite diagnosis of Alzheimer's disease is made only by direct examination of brain tissue obtained at autopsy or by biopsy to determine the presence of senile plaques and neurofibrillary tangles. A clinical evaluation, however, can provide

a correct diagnosis in more than 80% of cases. The clinical diagnosis of Alzheimer's disease requires a thorough evaluation to exclude all other medical, neurological, and psychiatric causes of the observed decline in memory and other cognitive abilities. This evaluation begins with a clinical history obtained from the patient and caregivers concerning the onset and course of symptoms and whether there are other medical conditions, head injury, excessive use of alcohol, or use of medications that could cause or exacerbate the cognitive decline. Neurological and psychiatric examinations are performed to characterize the cognitive deficits and to evaluate the presence of other disorders, such as Parkinson's disease, Huntington's disease, and cerebrovascular disease, that can cause dementia. The presence of behavioral disturbances such as depression and psychosis that can compromise cognitive function must also be evaluated. Laboratory tests are used to screen for medical conditions, including diabetes, renal dysfunction, hypothyroidism, infection, and vitamin deficiencies. A detailed neuropsychological evaluation is made to quantify and describe the cognitive deficits and to assist in staging the severity of dementia. Neuroimaging studies of the brain are performed to help rule out the presence of stroke, brain tumor, or other central nervous system abnormalities.

Risk factors. Although the cause of Alzheimer's disease is unknown, a number of factors that increase the risk of developing this form of dementia have been identified. Age is the most prominent risk factor, with the prevalence of the illness increasing twofold for each decade of life after age 60. Research in molecular genetics has shown that Alzheimer's disease is etiologically heterogeneous. Gene mutations on several different chromosomes are associated with familial inherited forms of Alzheimer's disease. Mutations of the amyloid precursor protein gene located on chromosome 21 have been linked to early-onset Alzheimer's disease in a small number of families. Additionally, virtually all individuals with Down syndrome (where there is an extra copy of chromosome 21) develop the brain pathology of Alzheimer's disease by the fourth decade of life. Gene defects on chromosomes 1 and 14 have also been implicated in familial early-onset forms of the illness, and they are thought to account for the vast majority of familially inherited cases of Alzheimer's disease. In the more common sporadic, late-onset form of Alzheimer's, specific allele types of the apolipoprotein gene on chromosome 19 act as a susceptibility factor that influences the age of onset of the disease.

Treatment. A major strategy for the treatment of Alzheimer's disease has focused on the relation between memory impairment and dysfunction of the acetylcholine neurotransmitter system. Tetrahydroaminoacridine (Tacrine) and donepezil hydrochloride (Aricept) were approved by the U.S. Food and Drug Administration for the treatment of Alzheimer's disease. These medications serve to increase the amount of acetylcholine available in the synapse by preventing the breakdown of the neurotransmitter that naturally occurs with presence of the enzyme acetylcholinesterase. Other treatment strategies to delay or diminish the progression of Alzheimer's disease are being explored, including efforts to modulate multiple neurotransmitter systems, to reduce the accumlation of amyloid in the brain, and to diminish the inflammatory response. Behavioral and pharmacological interventions are also available to treat the specific behavioral disturbances that can occur in Alzheimer's disease, such as depression, agitation, and sleep disorder. Gene Alexander

Bibliography. M. M. Esiri and J. H. Morris (eds.), *The Neuropathology of Dementia*, Cambridge University Press, 1997; B. A. Lawlor (ed.), *Behavioral Complications in Alzheimer's Disease*, American Psychiatric Press, 1995; R. D. Terry, R. Katzman, and K. C. Bick (eds.), *Alzheimer's Disease*, Raven Press, 1994; W. Wasco and R. E. Tanzi (eds.), *Molecular Mechanisms of Dementia*, Humana Press, 1997.

Amalgam

An alloy of mercury. Practically all metals will form alloys or amalgams with mercury, with the notable exception of iron. Amalgams are used as dental materials, in the concentration of gold and silver from their ores, and as electrodes in various industrial and laboratory electrolytic processes.

Amalgams used in dental work require the following composition: silver, 65% minimum; copper, 6% maximum; zinc, 2% maximum; and tin, 25% minimum. These amalgams are prepared by the dentist as needed, and harden within 3–5 min, but may be shaped by carving for 15 min or so. *See* ALLOY; GOLD; MERCURY (ELEMENT); SILVER. E. Eugene Weaver

Amaranth

An annual plant (seldom perennial) of the genus *Amaranthus* (family Amaranthaceae), distributed worldwide in warm and humid regions. Amaranths are botanically distinguished by their small chaffy flowers, arranged in dense, green or red, monoecious or dioecious inflorescences (spikes; see **illus**.), with zero to five perianth segments and two or three styles and stigmata, and by their dry membranous, indehiscent, one-seeded fruit. *See* FLOWER; FRUIT.

Physiological, genetic, and nutritional studies have revealed their potential economic value. Of particular interest are high rate of productivity as a rapidly growing summer crop, the large amounts of protein in both seed and leaf with high lysine, the overall high nutritional value, and the water use efficiency for the C_4 photosynthetic pathway. Amaranths are important in the culture, diet, and agricultural economy of the people of Mexico, Central and South America, Africa, and northern India. Genetic, ethnobotanical, and agronomic research has been undertaken to develop amaranths as an important food plant in modern agriculture.

Classification and species relationships. Up to 200 or more species of *Amaranthus* have been

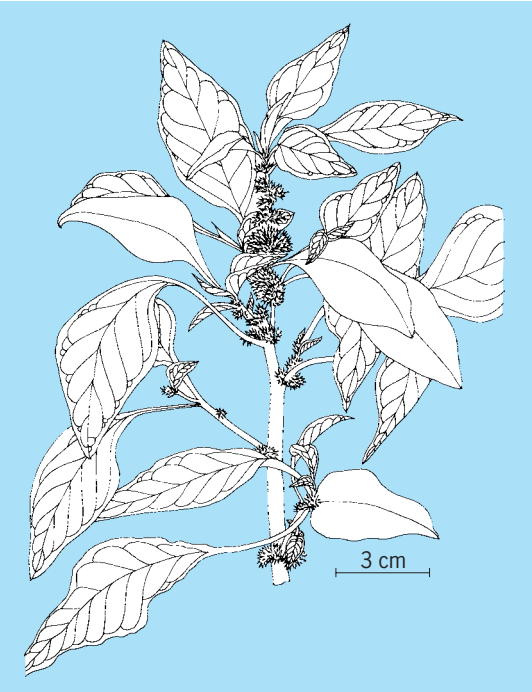

Flowers on branch of *Amaranthus tricolor*.

described, of which nearly 50 are distributed in the New World. Some of the more widespread weedy species are red amaranth or pigweed (*A. hybridus*), thorny amaranth (*A. spinosus*), tumbleweed (*A. albus*), and wild beet (*A. hybridus* ssp. *hypochondriacus*); Joseph's coat (*A. tricolor*) and love-lies-bleeding (*A. caudatus*) are recognized for their ornamental spikes and foliage colors.

Two sections of the genus are *Amaranthotypus*, with large terminal inflorescence, five tepals and stamens, and fruit opening circularly; and *Blitopsis*, with axillary cymes, often two to four tepals and stamens, and fruits opening irregularly. The cultivated grain species (*Amaranthus cruentus, A. caudatus*, and *A. hypochondriacus* = *A. leucocarpus*) with ornamental and crop forms belong to the section *Amaranthotypus*, whereas two important cultivated vegetable species (*A. tricolor* and *A. lividus*) belong to the section *Blitopsis*. Several Asian natives, such as *A. tricolor, A. gangeticus*, and *A. melancholicus*, have been grown as pot herbs or ornamentals but never as a grain crop.

The cultivated grain species are largely pale-seeded, whereas the ornamentals and weedy forms are black-seeded, presumably an important trait used by early peoples in distinguishing them. Interspecific hybridization, reported in various regions, is considered to be a primary source of evolutionary diversity. The grain crop forms and their weedy relatives include both 32- and 34-chromosome species; much of the preliminary biosystematic and cytogenetic research has not yet provided conclusive species relationships. However, the weedy and cultivated species of *Amaranthotypus* are separable into two distinct groups on the basis of taxonomic and protein-variation studies.

Domestication and cultivation. Amaranths are of ancient New World origin and were important as grains in the Aztec, Incan, and Mayan civilizations where wheat and rice were not grown. Archeological and ethnobotanical evidence suggests that the earliest domestication took place in South America at least 4000 years ago. About A.D. 500, grain amaranths were a widely cultivated food crop in Mexico. Some of the ornamental forms moved from the New World to Africa and Asia in the early 1800s, and may have spread from India to the East Indies, China, Manchuria, and Polynesia as a leaf vegetable. Ceremonial sacrifice and use in rituals drew the attention of the Spaniards, who suppressed the Aztec culture and amaranth cultivation.

Amaranths are a widely grown vegetable in tropical regions, including southern India, West and Equatorial Africa, and New Guinea. Amaranths have been found to be the most popular vegetable in southern Benin, in West Africa. Varieties such as Chinese spinach, Ceylon spinach, and African spinach suggest use as a salad ingredient; however, somewhat high oxalic acid content make the leaves unpleasant for this purpose.

Food value. Leaf amaranths are rich in provitamin A, vitamin C, iron, calcium, and protein, with lysine constituting as much as 5.9% of protein (equal to soymeal, and more than some of the best maize strains); glutamic acid constitutes as much as 10.8% of protein in *A. edulis*. Amaranth grains, with starch of very high quality and quantity, high protein (up to 17%), and high digestibility, rate better in nutritive value than all other cereals. They have been used in Mexico in modern times to make alegrias (cakes of popped seed), flour (of popped seed) used in the preparation of pinole, and paste used in the preparation of chuale; the black-seeded and variegated forms have been used to make amaranth milk (atole). In India, amaranth seed is used in alegrias and in confectioneries (flour) and as a cereal (popped seed). *See* CARYOPHYLLALES. Subodh K. Jain

Bibliography. R. Bressani, *Amaranth Newsletters*, Arch. Latino Americanos de Nutricion, Guatemala, 1984–1997; G. J. H. Grubben, *The Cultivation of Amaranth as a Tropical Leaf Vegetable*, Roy. Trop. Inst. Commun. 67, 1976; S. K. Jain, P. A. Kulakow, and I. Peters, Genetics and breeding of grain amaranth: Some research issues and findings, *Proceedings of the 3d Amaranth Conference, 1984*, pp. 174–191, 1986; National Research Council, *Amaranth*: *Modern Prospects for an Ancient Crop*, 1984; R. M. Saunders and R. Becker, *Amaranthus*: A potential food and feed resource, *Adv. Cereal Sci. Technol.*, 6:358–396, 1984.

Amateur radio

Two-way radio communications by individuals as a leisure-time activity. Amateur, or "ham," radio is defined by international treaty as a "service of self-training, intercommunications, and technical investigation carried on by amateurs; that is, by duly

authorized persons interested in radio technique solely with a personal aim and without pecuniary interest."

Activities. The government allows amateur operators many privileges because the hobby is partially based on service to the general public, and hams can be relied on to assist during emergencies. Groups of amateur operators meet annually to practice handling emergency communications in the field and to compete against other groups nationwide in performing certain emergency-related tasks. Amateur operators may set up warning and relief networks during the hurricane and tornado seasons, and handle communication when phone lines or cellular links are damaged by disasters.

In addition to public service activities, amateurs enjoy many recreational activities, including DXing (where the objective is to contact amateurs in as many foreign countries as possible), contesting (where the amateurs compete for the maximum number of contacts in a given time span), and fox-hunting (where the objective is to use radio skills to locate a hidden transmitter).

Since high-frequency (HF) signals (below 30 MHz) are reflected from the Kennelly-Heaviside ionosphere layers, amateurs are commonly able to carry out international communication. Very high frequencies (VHF; 30–300 MHz) and ultrahigh frequencies (UHF; 300–3000 MHz) have been subject to exploration by amateurs at the leading edge of communications technology. Amateurs bounce signals off the Moon or ionized meteor trails, and communicate through amateur operator-built earth satellites called OSCAR (orbiting satellites carrying amateur radio).

Equipment. For reliable, versatile, long-distance communications, some operators use large, directional beam antennas, but many have much simpler wire antennas strung between trees. Many operators have inexpensive equipment consisting of a wire antenna and just one transceiver (a radio that both receives and transmits) that functions on five or more frequency bands. *See* ANTENNA (ELECTROMAGNETISM).

Unlike most radio services, amateurs may design and build their equipment as long as it meets FCC requirements for spectral purity and bandwidth. They may assemble it from kits or may purchase it ready to go. As a result of the proliferation of commercial equipment, operators generally build only accessory items or experiment with modifications to equipment they have purchased. A few operators design equipment for frequency bands where commercial choices are limited, or experiment with new devices originally designed for the commercial frequencies.

Privileges. In the United States, the Federal Communications Commission (FCC) issues five classes of license, progressively allowing more frequencies and greater privileges. The first license level is the Novice Class, followed, in order of increasing privileges, by the Technician Class, General Class, Advanced Class, and Amateur Extra Class.

In 1991 the FCC eliminated the Morse code requirement from the Technician Class license, which shortly thereafter became the license of choice for most entry-level amateurs. Technician Class licensees have operating privileges on all amateur radio frequencies above 30 MHz.

Novice Class licensees may use voice, Morse code, and digital communications on a limited number of frequencies that allow worldwide HF voice communications and short-distance VHF voice communications on frequencies where readily available handheld radios allow mobility and portability.

Higher-class licensees have more frequency privileges and are permitted 1500 W peak envelope power output on most bands. Higher-class licensees may use any of the following modes: code (CW); amplitude-modulated (AM) and single-sideband (SSB) suppressed-carrier voice; frequency modulation (FM); radioteletype (RTTY), using either the Baudot, ASCII, or AMTOR (amateur teletype over radio) digital codes; facsimile; television; or pulse modulation. *See* AMPLITUDE-MODULATION RADIO; FREQUENCY-MODULATION RADIO; PULSE MODULATION; SINGLE SIDEBAND; TELETYPEWRITER; TELEVISION.

Amateur operators with General Class and higher licenses can communicate on portions of or all of the following frequency bands, depending on their specific license level:

1.8–2.0 MHz	1240–1300
3.5–4.0	2300–2310
7.0–7.3	2390–2450
10.10–10.15	3300–3500
14.00–14.35	5650–5925
18.068–18.168	10,000–10,500
21.000–21.450	24,000–25,250
24.890–24.990	48,000–50,000
28.0–29.7	71,000–76,000
50.0–54.0	165,000–170,000
144.0–148.0	240,000–250,000
222–225	Others above
420–450	300,000 MHz
902–928	

See RADIO SPECTRUM ALLOCATIONS.

Requirements. Questions found on each license class examination involve basic radio theory, rules that the FCC requires amateur radio operators to abide by, and a knowledge of international Morse code (except the Technician Class license). In each license class exam the questions get progressively more difficult to cover topics that befit the accompanying privileges of that license. For example, because the General Class license allows RTTY, there are questions covering RTTY on that exam but not on the Novice Class exam.

The Novice Class license requires the ability to send and receive Morse code at the rate of five words per minute (WPM); the General Class license requires 13 WPM; and the Amateur Extra, 20 WPM.

Technical developments. The open policies toward amateur radio by the government have made pioneering possible in several areas. In 1923, for instance, amateurs led the way in the use of short

waves after an amateur in Connecticut talked with a ham in France on the wavelength of 110 m (360 ft). The accepted theory had been that only wavelengths above 200 m (660 ft) were really useful for reliable communications.

By the early 1930s communications between points on opposite sides of the globe had become commonplace on wavelengths of 80–20 m (264–66 ft; frequencies of 3.5–14.4 MHz). Many hams then turned their attention to higher frequencies. Again amateurs shattered previous conceptions, especially those concerning so-called line-of-sight limitations on communications at 56 MHz and above. These pioneers discovered means of radio propagation which reached far beyond the usual horizon: reflections from the aurora borealis, from meteor trails, from sporadic E layers (patches of ionized particles about 70 mi or 110 km above the Earth), and from the Moon; bending of radio waves through layers of stable air; and a phenomenon called transequatorial scatter, by which stations on one side of the Equator may communicate with stations on the other side over distances of more than 1000 mi (1600 km) at times when such communication would otherwise be considered impossible. This mode and several of the others were subjects of an International Geophysical Year study undertaken by amateurs through the American Radio Relay League. *See* RADIO-WAVE PROPAGATION; TROPOSPHERIC SCATTER.

Other technical accomplishments and experimentation by amateurs include computer communications multiplexed on common radio channels, or packet radio; computer station control and interfacing as a result of the popularity and availability of personal computers; and spread spectrum techniques, enabling hundreds of simultaneous conversations to be carried on one frequency band without any mutual interference. Amateurs have installed relay repeater stations on mountaintops in many areas and use them to extend the range of communication on the VHF and UHF bands from tens to hundreds of miles. Under the direction of the Radio Amateur Satellite Corporation (AMSAT), amateurs have built and launched as secondary payloads on scheduled space shots a series of highly advanced satellites, some into low earth orbit and others into high elliptical orbits, containing refirable hydrazine engines, master on-board computers, and high-power VHF-UHF transponders. These satellites act as relay stations, again increasing the range of possible communications on the frequencies of the transponders. *See* SPREAD SPECTRUM COMMUNICATION.

Associations. The developments in the hobby over the years required a clearinghouse, an information exchange, which is embodied in the American Radio Relay League (ARRL), an association of radio amateurs in North America founded in 1914. It publishes a monthly magazine, a handbook, and other publications and operator aids covering various aspects of the hobby. ARRL also serves to represent amateurs in government regulatory matters; presents new technical developments and sponsors operating contests and other activities; and organizes networks to handle messages and coordinates emergency communications training. The League is also Secretariat for the International Amateur Radio Union, composed of over 125 similar national radio societies around the world. *See* RADIO.

Steve Mansfield

Bibliography. American Radio Relay League: *Advanced Class License Manual, Extra Class License Manual, The FCC Rule Book, Radio Amateur's Handbook, General Class License Manual,Now You're Talking, Radio Amateur's Handbook*, all annually.

Amber

Most commonly, a generic name for all fossil resins, although it has been restricted by some to refer only to succinite, the mineralogical species of fossil resin making up most of the Baltic Coast deposits. Resins generally are complex mixtures of mono-, sesqui-, di-, and triterpenoids; however, some resins contain aromatic phenols or are even predominantly composed of these compounds. Resins are synthesized in appreciable quantity by about 10% of present-day plant families. Among the plants, primarily trees, that produce copious amounts of resin that may fossilize to become amber, about two-thirds are tropical or subtropical. Members of the families Pinaceae, Araucariaceae, and Taxodiaceae are the most prominent copious producers among the conifers, and Leguminosae, Dipterocarpaceae, Burseraceae, and Hamamelidaceae among the angiosperms. *See* RESIN.

Fossilization. Under natural forest conditions resins harden with varying degrees of rapidity, and can become fossilized if they have the requisite chemical composition for polymerization, are sufficiently stable to withstand both oxidative and microbial degradation, and appropriate depositional conditions are available. The fossilization process is still not completely understood but appears primarily to involve polymerization, with some modification of the original constituents, especially due to the temperature during the maturation of the resin in the sediments.

Occurrence. Although ambers occur throughout the world in deposits from Carboniferous to Pleistocene in age, they have been reported most commonly from Cretaceous and Tertiary strata and often are associated with coal or lignites. Amber may contain beautifully preserved insects, spiders, flowers, leaves, and even small animals (**Fig. 1**). The most extensively studied deposits are those from the Baltic Coast, Alaska, Canada, North American Atlantic coast, southeast Asia, Dominican Republic, and Mexico (**Fig. 2**).

Uses. Since the earliest stages of human social development, ambers have had esthetic appeal, have been used to ward off evil powers, and have been thought to cure certain illnesses. The attribution of these special powers to amber may result partially from the negative electrical properties exhibited when ambers are rubbed. Thales recognized

Fig. 1. ***Pseudosphegina carpenteri*** **(Diptera), syrphid fly in Baltic amber, Oligocene. (*Courtesy of F. M. Carpenter*)**

these properties and, in fact, the term electricity evolved from *electron,* the Greek name for amber. *See* ELECTRIC CHARGE; ELECTRICITY.

Carvings and beads of Baltic amber found in caves and tombs indicate its significance in human activities since the Stone Age. Amber provided a distinctive and imperishable barter item for trade routes that crossed Europe from the Baltic to the Adriatic and Black seas to the Bronze and Iron ages as well as during Roman and Greek periods. When used for jewelry, it usually is transparent yellow, reddish-brown, or "amber" color. Translucent or semitranslucent amber is used for pipe stems, decorating small boxes, and other ornamental purposes. Specific gravity varies from 1.05 to 1.10, and hardness from 1 to 3 on Mohs scale.

Large amounts of amber occur in some coal deposits, which are being modified for some industrial uses, such as jet fuel. Furthermore, fossil resins, probably from members of the Dipterocarpaceae, have made marked contributions to petroleum in numerous southeast Asian sites.

Chemistry and origin. At one time, chemical studies of amber were mineralogically oriented because the purpose was to describe and classify amber as a semiprecious gem. However, phytochemical studies comparing fossil and present-day resins, employing such techniques as infrared spectroscopy, x-ray diffraction, nuclear magnetic resonance, and pyrolysis gas chromatography–mass spectrometry are providing information regarding the botanical origins of ambers. For example, amber from Chiapas, Mexico, and from the Dominican Republic has been shown to be derived from the leguminous genus *Hymenaea*; that from Sumatra from *Shorea* (Dipterocarpaceae); that from Ecuador from members of the Burseraceae; that from Australia and New Zealand

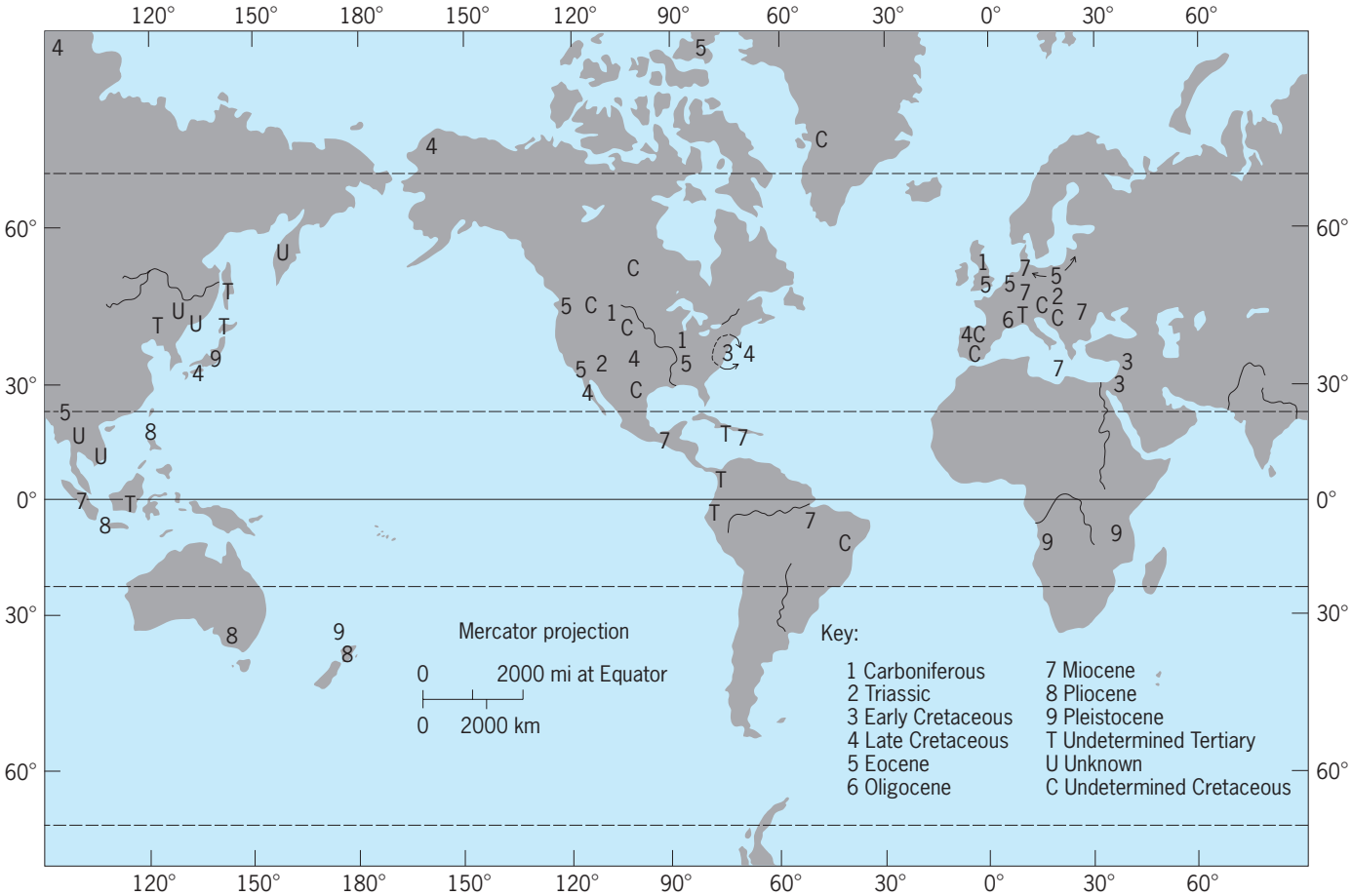

Fig. 2. Important amber deposits of the world. (*After J. H. Langenheim, Amber: A botanical inquiry, Science, 163:1157–1169, 1969*)

from *Agathis* (Araucariaceae); and that from Alaska probably from genera, such as *Metasequoia*, in the Taxodiaceae. The origin of the amber from the extensive Baltic deposits remains a mystery as the chemistry of the resin is more closely related to *Agathis* than to *Pinus* (Pinaceae), which early was considered to be the source from parts of the plant found in the amber. Then *Agathis* was considered because it has many chemical similarities to Baltic amber; however, it lacks succinic acid, a characteristic component, as well as plant remains of it in the amber. Recently amber from a Canadian Arctic site that contains succinic acid has been found within cone scales of *Pseudostrobus* (Pinaceae). As with *Agathis*, no parts of *Pseudostrobus* have been found in the amber. Although some chemical aspects differ from Baltic amber, the presence of succinic acid in the *Pseudostrobus* amber may provide a clue for future searches regarding the enigmatic source of Baltic amber. The predominantly tropical or subtropical occurrence of amber-producing plants through geologic time has led to evolutionary studies of the natural purpose of resins and their possible defensive role for trees against injury and disease inflicted by the high diversity of insects and fungi in tropical environments. *See* ARCHEOLOGICAL CHEMISTRY.

Jean H. Langenheim

Bibliography. D. A. Grimaldi, *Amber: Window to the Past*, Harry Abrams, New York, 1996; J. H. Langenheim, Biology of amber-producing trees: Focus on case studies of *Hymenaea* and *Agathis*, in K. B. Anderson and J. C. Crelling (eds.), *Amber, Retinite and Fossil Resins*, ACS Ser., no. 617, 1995; G. O. Poinar, Jr., and R. Poinar, *The Amber Forest*, Princeton University Press, 1999; P. C. Rice, *Amber: The Golden Gem of the Ages*, Kosciusko Foundation, New York, 1987.

Ambergris

A fatty substance formed in the intestinal tract of the sperm whale (*Physeter catodon*). There is a question as to whether the origin of ambergris is normal or pathological, but it does serve as protection from the horny indigestible portions of the squid and cuttlefish that constitute much of the whale's diet.

Ambergris contains acids, alkaloids, and a fatty substance, its main constituent, called ambrein. Although fresh ambergris is soft and black and has an offensive odor, it hardens into pleasantly fragrant gray or yellow masses when exposed to the air, sun, and sea. Being lighter than water, it is found in lumps, weighing from $^1/_2$ oz (15 g) to over 100 lb (45 kg) floating on tropical seas or cast up on the shores. It is also gathered directly from the abdomens of dead or captured whales when they are slaughtered. Collecting grounds for ambergris are principally on the shores of China, Japan, Africa, the Americas, tropical islands, and the Bahamas.

Known and highly prized since the earliest times, ambergris was used chiefly for medicinal purposes. In modern times this rare commodity is valued in the manufacture of perfumes. The ambergris is ground and used in the form of a tincture, dissolved in a dilute solution of alcohol, which when added to perfume acts as a fixative, increasing the duration of the fragrance while adding its own sweet, earthy scent.

Sybil P. Parker

Amblygonite

A lithium aluminum phosphate mineral of basic formula $LiAl(PO_4)(F)$. The structure of amblygonite consists of phosphate (PO_4) groups of tetrahedra and AlO_6 groups of octahedra. Each PO_4 tetrahedron is connected to an AlO_6 octahedron. Corner-sharing octahedra form zig-zag chains along the b axis. Lithium (Li) is in fivefold coordination, and lies between the PO_4 tetrahedra and nearest AlO_6 octahedra. The dominant substitution in this mineral structure is hydroxyl (OH) for fluorine (F). This substitution gives rise to the amblygonite-montebrasite [$LiAlPO_4(OH)$] solid solution series. When OH is greater than F, the mineral is known as montebrasite. Appreciable amounts of sodium substitute for lithium in the five-coordinated polyhedra. The sodium-rich varieties such as natromontebrasite and hedronite are rare.

Amblygonite crystallizes in the triclinic system. Its color is commonly white or gray with tints of blue, green, and yellow. Amblygonite is transparent to translucent and has a vitreous to pearly luster. It often exhibits polysynthetic twinning on both a hand specimen and a microscopic scale. Two cleavage directions are prominent. Cleavage fragments of amblygonite-montebrasite may be confused with potassium feldspar, but they are distinguished by their higher density (specific gravity = 3.0–3.1) and cleavage angles that are not at 90°.

Members of the amblygonite-montebrasite series typically occur as coarse, cleavable nodules in the inner zones and quartz cores of zoned lithium-rich, granitic pegmatites. Quartz, spodumene, petalite, and feldspar are the most common minerals associated with amblygonite. Minerals such as lepidolite, beryl, cassiterite, tourmaline, and niobium-tantalum oxides are commonly associated with amblygonite-bearing mineral assemblages in smaller abundances. These large masses of amblygonite-montebrasite are magmatic in origin and generally contain 4–7 wt % fluorine. Clear, equant crystals are rare and occur in vugs in pegmatites. Crystals associated with vugs and altered (secondary) magmatic amblygonite contain 0.3–4 wt % fluorine. A single mass of amblygonite with dimensions of 7.62 m × 2.44 m × 1.83 m (25 ft × 8 ft × 6 ft) has been documented in the Hugo granitic pegmatite in the Black Hills of South Dakota. Larger masses have been reported. The best-known occurrences of amblygonite are in Montebras, France; the Black Hills of South Dakota; the White Picacho District in Arizona; pegmatite districts in Maine; the Tanco pegmatite in Manitoba, Canada; and Portland, Connecticut. While

amblygonite has been mined as an ore of lithium, it is not a major ore. *See* PHOSPHATE MINERALS.

Charles K. Shearer

Bibliography. C. Klein and C. S. Hurlbut, Jr., *Manual of Mineralogy,* 21st ed., 1999; P. B. Moore, *Pegmatite Minerals of P (V) and B (III): Granitic Pegmatities in Science and Industry,* 1982.

Amblypygi

The tailless whip scorpions or whip spiders, an order of the class Arachnida. There are about 80 species in the tropics and subtropics. They are flattened, red to brown, and range from 5 to 45 mm (0.2 to 1.8 in.) in body length. All are nocturnal, hiding under bark or stones, among leaves, or in caves during the day.

The cephalothorax (prosoma) and segmented abdomen (opisthosoma) are narrowly attached. There is no tail. The cephalothoracic shield (carapace) has three simple eyes on each side and a pair of eyes anterior. The chelicerae are like the jaws of a spider. The raptorial pedipalps are strong and have long, hard spines used to surround and crush insect prey. The long, whiplike first legs are sense organs used to locate prey. There are booklungs on the underside of the abdomen. There are no venom or repellent glands.

During courtship the male uses his long first legs to stroke the female; he may tremble while stroking. After some time he deposits a spermatophore (a package of sperm) on the ground. He pulls or guides the female over the spermatophore by using his chelicerae or first legs. The female picks up the spermatophore with her gonopore. Up to 60 eggs are laid in a flexible sac secreted by the female reproductive organs; the sac remains attached to the underside of the female's abdomen. On hatching, the young climb up on their mother's abdomen and are carried until their next molt. *See* ARACHNIDA.

H. W. Levi

Ameba

Any protozoon moving by means of protoplasmic flow. In their entirety, the ameboid protozoa include naked amebas, those enclosed within a shell or test, as well as more highly developed representatives such as the heliozoians, radiolarians, and foraminiferans. Ameboid movement is accomplished by pseudopods—cellular extensions which channel the flow of protoplasm. Pseudopods take varied forms and help distinguish among the different groups. A lobe-shaped extension or lobopod is perhaps the simplest type of pseudopod. The shapelessness and plasticity of these locomotory organelles impart an asymmetric, continually changing aspect to the organism. Other, more developed, representatives have pseudopodial extensions containing fibrous supporting elements (axopods) or forming an extensive network of anastomosing channels (reticulopods). Though involved in locomotion, these organelles are also functional in phagocytosis—the trapping and ingesting of food organisms (usually bacteria, algae, or other protozoa) or detritus. *See* FORAMINIFERIDA; HELIOZOIA; PHAGOCYTOSIS; RADIOLARIA.

Amebas are found in a variety of habitats, including fresh-water and marine environments, soil, and as symbionts and parasites in body cavities and tissues of vertebrates and invertebrates. Because of their manner of locomotion, amebas typically occur on surfaces, such as the bottom of a pond, on submerged vegetation, or floating debris. In soil, they are a significant component of the microfauna, feeding extensively on bacteria and small fungi. Amebas in marine habitats may be found as planktonic forms adapted for floating at the surface (having oil droplets to increase bouyancy and projections to increase surface area), where they feed upon bacteria, algae, and other protozoa. Several species of amebas may be found in the human intestinal tract as harmless commensals (for example, *Entamoeba coli*) or as important parasites responsible for amebic dysentery (*E. histolytica*).

Amebas range from small soil organisms, such as *Acanthamoeba* (20 micrometers), to the large fresh-water forms *Amoeba proteus* (600 μm; see **illus.**) and *Pelomyxa* (1 mm, or more). Some types, such as *Amoeba,* are uninucleate; others are multinucleate. Reproduction is by mitosis with nuclear division preceding cytoplasmic division to produce two daughters. Multinucleate forms have more unusual patterns of division, since nuclear division is not immediately nor necessarily followed by cytoplasmic division. Transformation of the actively feeding ameba into a dormant cyst occurs in many species, particularly those found in soil or as symbionts. The resting stages allow survival over periods of desiccation, food scarcity, or transmission between hosts. *See* REPRODUCTION (ANIMAL).

Much attention has focused on the mechanism of ameboid movement in *Amoeba proteus,* a favorite research species because of its large size and ease of maintenance in the laboratory. Under the light microscope, the organism is seen to contain a granular endoplasmic core in a sol (fluid) state, enclosed by a gel-like ectoplasm. Endoplasm flows forward as the organism progresses; as it reaches the front end of the advancing pseudopod, the endoplasm is deflected sideways and rearward, where it converts to ectoplasmic gel. This pattern has been termed fountain streaming. At the posterior end of the organism, the reverse occurs, with ectoplasm reverting to a forward-flowing endoplasm. The ameba resembles an ectoplasmic cylinder enclosing a fluid endoplasmic core. It is generally accepted that movement is brought about by the contraction of protoplasm at the anterior end of the pseudopod, pulling the ameba along the substrate. The contraction involves cytoplasmic microfilaments, adenosine triphosphate (ATP) as the energy source, and divalent

Phase-contrast photomicrograph of *Amoeba proteus*, a large fresh-water ameba. The organism is seen moving by means of a single lobose pseudopod.

cations, for example, Ca^{2+} and Mg^{2+}. *See* CELL MOTILITY.

Cytoplasm of *Amoeba* contains, besides food vacuoles, a contractile vacuole functioning for osmoregulation—a microkidney, which expells water. Energy for this, as well as other functions, is derived from numerous mitochondria present in the cytoplasm, which may also contain crystals and endosymbiotic bacteria. *See* MITOCHONDRIA; OSMOREGULATORY MECHANISMS; VACUOLE.

Ameboid protozoa are believed to have evolved from flagellated protozoan ancestors through loss of the flagellar apparatus. Some ameboid organisms, the ameboflagellates, can change from creeping amebas to swimming flagellates during their life cycle, providing support for this evolutionary pathway. Among certain specialized ameboid protozoa, production of flagellated gametes prior to sexual reproduction may occur. *See* AMOEBIDA; PROTOZOA.

Frederick L. Schuster

Americium

A chemical element, symbol Am, atomic number 95. The isotope ^{241}Am is an alpha emitter with a half-life of 433 years. Other isotopes of americium range in mass from 232 to 247, but only the isotopes of mass 241 and 243 are important. The isotope ^{241}Am is routinely separated from "old" plutonium and sold for a variety of industrial uses, such as 59-keV gamma sources and as a component in neutron sources. The longer-lived ^{243}Am (half-life 7400 years) is a precursor in ^{244}Cm production.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

In its most prominent aqueous oxidation state, 3+, americium closely resembles the tripositive rare earths. The formal analogy to the rare earths is also marked in anhydrous compounds of both tripositive and tetrapositive americium. Americium is different in that it is possible to oxidize Am^{3+} to both the 5+ and 6+ states.

Americium metal has a vapor pressure markedly higher than that of its neighboring elements and can be purified by distillation. The metal is nonmagnetic and superconducting at 0.79 K. Under high pressure the metal has been compressed to 80% of its room-temperature volume and displays the α-uranium structure. Americium-241 is used as an alpha-particle source in ionization-type smoke detectors. *See* ACTINIDE ELEMENTS; ALPHA PARTICLES; BERKELIUM; CURIUM; FIRE DETECTOR; NUCLEAR REACTION; PERIODIC TABLE; TRANSURANIUM ELEMENTS.

R. A. Penneman

Bibliography. N. M. Edelstein, J. O. Navratil, and W. W. Schulz, *Americium and Curium Chemistry and Technology*, 1985; S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; J. J. Katz, G. T. Seaborg, and L. R. Morss (eds.), *The Chemistry of the Actinide Elements*, 2 vols., 2d ed., 1986; W. W. Schulz, *The Chemistry of Americium*, ERDA Crit. Rev. Ser. Rep. TID-26971, 1976.

Amethyst

The transparent purple to violet variety of the mineral quartz. Although quartz is perhaps the commonest gem mineral known, amethyst is rare in the deep colors that characterize fine quality. Amethyst is usually colored unevenly and is often heated slightly in an effort to distribute the color more evenly. Heating at higher temperatures usually changes it to yellow or brown (rarely green), and further heating removes all color. The principal sources are Brazil, Arizona, Uruguay, and Russia. Amethyst is often cut in step or brilliant shapes, and drilled or carved for beads. Carvings are made both from transparent and nontransparent material. *See* GEM; QUARTZ.

Richard T. Liddicoat, Jr.

Amide

A derivative of a carboxylic acid with general formula (**1**), where R is hydrogen or an alkyl or

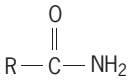

(**1**)

aryl radical. Amides are divided into subclasses, depending on the number of substituents on nitrogen. The simple, or primary, amides are considered to be derivatives formed by replacement of the carboxylic hydroxyl group by the amino group, NH_2. They are named by dropping the "-ic acid" or "-oic acid" from the name of the parent carboxylic acid and replacing it with the suffix "amide," as shown in examples (**2–5**). In the secondary and tertiary amides, one or

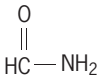

Formamide (**2**)

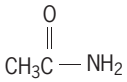

Acetamide (**3**)

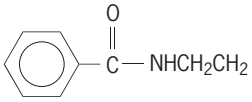

N-Ethylbenzamide (**4**)

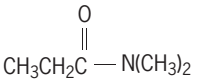

N,N-Dimethylpropanamide (**5**)

both hydrogens are replaced by other groups. The presence of such groups is designated by the prefix capital *N* (for nitrogen), as shown in the examples.

Except for formamide, all simple amides are relatively low-melting solids, stable, and weakly acidic. They are strongly associated through hydrogen bonding, and hence soluble in hydroxylic solvents, such as alcohol. Because of ease of formation and sharp melting points, amides are frequently used for the identification of organic acids and, conversely, for the identification of amines.

Formation and properties. Commercial preparation of amides involves thermal dehydration of ammonium salts of carboxylic acids. Thus, slow pyrolysis of ammonium acetate, $CH_3COO^-NH_4^+$, forms water and acetamide. *N,N*-dimethylacetamide may be similarly prepared from dimethylammonium acetate:

$$CH_3COO^-\overset{+}{N}H_2(CH_3)_2$$

Acid anhydrides react with ammonia or with primary or secondary amines to form amides. The preparation of acetanilide [reaction (1)] is an example.

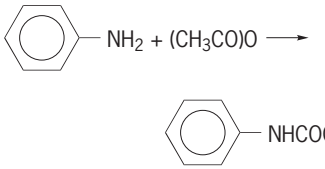

$$\text{—}NH_2 + (CH_3CO)O \longrightarrow \text{—}NHCOCH_3 + CH_3COOH \quad (1)$$

The reaction of amines with acid chlorides in the presence of aqueous sodium hydroxide (Schotten–Baumann reaction) is often used [reaction (2)].

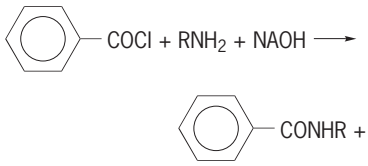

$$\text{—}COCl + RNH_2 + NAOH \longrightarrow \text{—}CONHR + NaCl + H_2O \quad (2)$$

When urea, the diamide of carbonic acid, is heated with a carboxylic acid, a new amide is formed by an exchange reaction. The other product of the reaction, carbamic acid, decomposes to permit the reaction to go to completion [reaction (3)].

$$CO(NH_2)_2 + CH_3COOH \rightleftharpoons CH_3CONH_2 + H_2NCOOH \longrightarrow CO_2\uparrow + NH_3\uparrow \quad (3)$$

The partial hydrolysis of nitriles also affords a convenient synthesis of a variety of amides [reaction (4)]. The reaction is catalyzed by both acid and base.

$$CH_3CH_2CH_2CH_2C{\equiv}N + H_2O \longrightarrow CH_3CH_2CH_2CH_2CONH_2 \quad (4)$$

Uses. Amides are important chemical intermediates since they can be hydrolyzed to acids, dehydrated to nitriles, and degraded to amines containing one less carbon atom by the Hofmann reaction. In pharmacology, acetaminophen (**6**) is a popular analgesic. *N,N*-Dimethylformamide (DMF; **7**) is a useful aprotic, highly polar solvent.

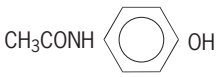

(**6**)

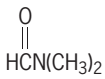

(**7**)

The amides of the straight-chain fatty acids having 12, 14, 16, or 18 carbon atoms are especially useful industrially because of the variety of properties

that can be obtained by substitution on the nitrogen atoms. Such products find extensive use as waterproofing agents, lubricant additives, detergents, emulsifiers, and wetting agents.

Amides are very prevalent in nature, since all peptides and proteins are polymers of the natural α-amino acids, as represented by (**8**).

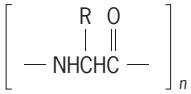

(**8**)

See PEPTIDE; PROTEIN.

Nylon is the generic term for any synthetic, long-chain polymer with an amide linkage as a recurring, integral part of the chain. Nylon-6 (**9**) is one of the most useful.

$$[-NH(CH_2)_5CO-]_n$$

(**9**)

See ACID ANHYDRIDE; ACID HALIDE; AMINE; CARBOXYLIC ACID; NITRILE; POLYAMIDE RESINS.

Paul E. Fanta

Bibliography. R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998.

Amiiformes

An order of actinopterygian fishes in the subclass Neopterygii. Amiiformes plus several related fossil orders comprise the Halecomorphi, a well-developed group known from the middle Mesozoic. The Amiiformes comprise several families that persisted into the Cenozoic era; but only one, the Amiidae, survived to the present and it consists of a single species, *Amia calva*, the bowfin. The fossil amiids are from freshwater deposits, the oldest of which are of Jurassic age. The other halecomorphs were marine species.

Amia calva (see **illustration**) is characterized by an abbreviated heterocercal tail; fusiform body; large mouth equipped with sharp teeth; an elongate dorsal fin, with each ray supported by a single pterygiophore; scales with a ganoin (enamel-like) surface but thin and overlapping; no spiracles; a vascular swim bladder that can act as a lung; and a large gular plate. The maximum length of the bowfin is about 90 cm (35 in.). Its usual habitat is still waters with an abundance of rooted vegetation in lowlands of eastern North America from the St. Lawrence River to the Gulf Slope and from the Mississippi Basin to the Atlantic Slope, avoiding Appalachian streams. *See* ACTINOPTERYGII; OSTEICHTHYES; SWIM BLADDER; TELEOSTEI.

Herbert Boschung

***Amia calva*, the bowfin. (*Courtesy of Frank Teigler, Hippocampus Bildarchiv*)**

Bibliography. L. Grande and W. E. Bemis, A comprehensive phylogenetic study of amiid fishes (Amiidae) based on comparative skeletal anatomy, *J. Vert. Paleontol.*, Spec. Mem. 4 (suppl. to vol. 18), 1998; J. G. Maisey, *Santana Fossils: An Illustrated Atlas*, T. F. H. Publications, Neptune City, NJ, 1991; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Amine

A member of a group of organic compounds which can be considered as derived from ammonia by replacement of one or more hydrogens by organic radicals. Generally amines are bases of widely varying strengths, but a few which are acidic are known.

Amines constitute one of the most important classes of organic compounds. The lone pair of electrons on the amine nitrogen enables amines to participate in a large variety of reactions as a base or a nucleophile. Amines play prominent roles in biochemical systems; they are widely distributed in nature in the form of amino acids, alkaloids, and vitamins. Many complex amines have pronounced physiological activity, for example, epinephrine (adrenaline), thiamin or vitamin B_1, and Novocaine. The odor of decaying fish is due to simple amines produced by bacterial action. Amines are used to manufacture many medicinal chemicals, such as sulfa drugs and anesthetics. The important fiber nylon is an amine derivative.

Amines are classified according to the number of hydrogens of ammonia which are replaced by radicals. Replacement of one hydrogen results in a primary amine (RNH_2), replacement of two hydrogens results in a secondary amine (R_2NH), and replacement of all three hydrogens results in a tertiary amine (R_3N). The substituent groups (R) may be alkyl, aryl, or aralkyl. In another group of amines the nitrogen forms part of a ring (heterocyclic amines).

Examples of such compounds are nicotine (**1**), which is obtained commercially from tobacco for use as an insecticide, and serotonin (**2**), which plays

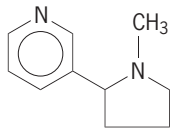

(**1**)

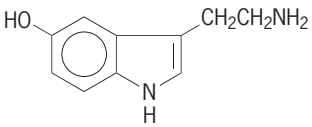

(**2**)

a key role as a chemical mediator in the central nervous system.

Many aromatic and heterocyclic amines are known by trivial names, and derivatives are named as substitution products of the parent amine. Thus, $C_6H_5NH_2$ is aniline and $C_6H_5NHC_2H_5$ is *N*-ethylaniline. For a discussion of definitive rules for naming amines *see* ORGANIC CHEMISTRY

According to the Brönsted-Lowry theory of acids and bases, amines are basic because they accept protons from acids. In water the equilibrium shown in reaction (1) lies predominantly to the left. The ex-

$$RNH_2 + H_2O \rightleftharpoons RNH_3^+ + OH^- \quad (1)$$

tent to which the amine is successful in picking up a proton from water is given by Eq. (2), where the

$$K_b = \frac{[RNH_3]^+[OH]^-}{RNH_2} \quad (2)$$

quantities in brackets signify concentrations of the species given. For short-chain aliphatic amines the basic ionization constant K_b lies near 10^{-4}; for aromatic amines $K_b < 10^{-9}$; for ammonia $K_b = 1.8 \times 10^{-5}$. Stable salts suitable for the identification of amines are in general formed only with strong acids, such as hydrochloric, sulfuric, oxalic, chloroplatinic, or picric.

Reactions and identification. Several test-tube reactions for recognition and characterization of amines are known: the Schotten-Baumann reaction, the Hinsberg test, the carbylamine reaction, and the action of nitrous acid.

The Schotten-Baumann reaction involves treatment of an amine with benzoyl chloride in basic solution. It serves to distinguish tertiary amines from primary and secondary amines by the formation of substituted benzamides from the primary and secondary amines [reactions (3) and (4)]. Tertiary

$$RNH_2 + C_6H_5COCl + OH^- \longrightarrow C_6H_5CONHR + Cl^- + H_2O \quad (3)$$

$$R_2NH + C_6H_5COCl + OH^- \longrightarrow C_6H_5CONR_2 + Cl^- + H_2O \quad (4)$$

amines do not react with benzoyl chloride. The substituted benzamides are generally insoluble in water, solid, and easily purified, and they have characteristic melting points. Thus they serve to identify the amines.

The more reactive acylating agents, acetic anhydride and acetyl chloride, give substituted acetamides without added base. This reaction gives the same type of information as the Schotten-Baumann reaction.

Closely related to the Schotten-Baumann reaction is the Hinsberg test. This has the added advantage of distinguishing between primary and secondary amines. It involves reaction of an amine with benzenesulfonyl chloride in alkaline solution. Both it and the Schotten-Baumann test are applicable to both aliphatic and aromatic amines with the exception of those amines which are substantially nonbasic in character. Primary amines give sulfonamides that are soluble in basic solutions; secondary amines give insoluble derivatives; and tertiary amines with no replaceable hydrogen do not react with the reagent. In general, the sulfonamides are solids and are useful for identification of the amines.

Carbylamines (isocyanides) possess a very unpleasant, nauseating odor and are formed by the reaction of any primary amine with chloroform in basic solution [reaction (5)].

$$RNH_2 + CHCl_3 + 3OH^- \longrightarrow R{-}N{\equiv}C + 3Cl^- + 3H_2O \quad (5)$$

Reaction with nitrous acid serves as a further method for distinguishing between various classes of amines. Primary aliphatic amines evolve nitrogen, whereas primary aromatic amines give diazonium compounds, which may be recognized by dye formation on coupling with a suitable second component. Secondary amines of both series give nitrosamines, generally as yellow oils. Tertiary aliphatic amines do not react with nitrous acid, and mixed aliphatic aromatic tertiary amines undergo nuclear nitrosation. Nitrosamines have been identified as potent carcinogenic substances. *See* DIAZOTIZATION.

In the infrared absorption spectrum, amines exhibit characteristic bands due to N-H stretching and bending, as well as C-N stretching vibrations. In the proton magnetic resonance spectrum, the appearance of the proton on nitrogen is complicated by the rate of exchange and the electrical quadrupole of the ^{14}N nucleus.

In alkaline solutions tertiary amines are oxidized by hydrogen peroxide to amine oxides which, although still basic, are not strong bases as are the quaternary ammonium types [reaction (6)].

$$R_3N + H_2O_2 \longrightarrow R_3NO + H_2O \quad (6)$$

Aromatic amines undergo halogenation and sulfonation on the ring. However, because of their susceptibility to oxidation, nitration cannot be accomplished without prior protection of the labilizing amino group. This is commonly done by acetylation.

Exhaustive methylation is one of a number of reactions of amines associated with the name of A. W. Hofmann. It involves a sequence of reactions terminating in the thermal decomposition of a quaternary ammonium hydroxide to yield a tertiary amine, usually trimethylamine, water, and an olefin. It has been widely used as a tool in the determination of the structures of complex compounds and, to a lesser extent, for the synthesis of olefins. The sequence of reactions is shown in (7).

$$RCH_2CH_2NH_2 \xrightarrow{3CH_3I} RCH_2CH_2\overset{+}{N}(CH_3)_3L^- \xrightarrow{AgOH} RCH_2CH_2\overset{+}{N}(CH_3)_3OH^- \xrightarrow{heat} RCH{=}CH + (CH_3)_3N + H_2O \quad (7)$$

See ALKALOID.

Preparation. Commercial preparation of aliphatic amines can be accomplished by direct alkylation of ammonia (Hofmann method, 1849) or by catalytic alkylation of amines with alcohols at elevated temperatures. Reduction of various nitrogen functions carrying the nitrogen in a higher state of oxidation also leads to amines. Such functions are nitro, oximino, nitroso, and cyano. For the preparation of pure primary amines, Gabriel's synthesis and Hofmann's hypohalite reaction are preferred methods. The Bucherer reaction is satisfactory for the preparation of polynuclear primary aromatic amines.

Gabriel's synthesis. This is a method for the synthesis of pure primary aliphatic amines by the hydrolysis of an *N*-alkyl phthalimide. The *N*-alkyl phthalimides are prepared by reaction of potassium phthalimide with an alkyl halide (preferably a bromide) [reaction (8)].

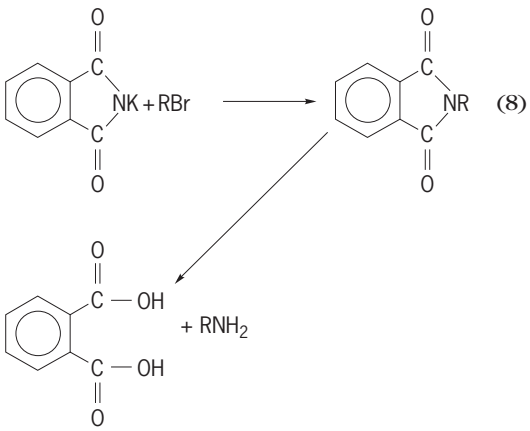

Hofmann hypohalite reaction. This is another reaction associated with Hofmann, which furnishes a convenient method for the preparation of pure primary amines of either the aliphatic or aromatic series. In the overall sense, it involves conversion of an acid amide to an amine with loss of one carbon atom. The reaction proceeds through the stages shown in reactions (9)–(11). The amine arises by hydrolysis of

$$RCONH_2 + NaOH \longrightarrow RCONHX + NaOH \quad (9)$$

$$RCONHX + NaOH \longrightarrow R{-}N{=}C{=}O + NaX + H_2O \quad (10)$$

$$R{-}N{=}C{=}O + 2NaOH \longrightarrow RNH_2 + Na_2CO_3 \quad (11)$$

the isocyanate formed by migration of R from carbon to nitrogen in reaction (10) . Sodium hypobromite is the common laboratory reagent, but the cheaper calcium hypochlorite is used in commercial applications, such as the manufacture of anthranilic acid from phthalimide. *See* AMINO ACIDS; QUATERNARY AMMONIUM SALTS. Paul E. Fanta

Bibliography. N. L. Allinger et al., *Organic Chemistry*, 2d ed., 1976; W. H. Brown and E. P. Rogers, *General, Organic, and Biochemistry*, 3d ed., 1987.

Amino acid dating

Determination of the relative or absolute age of materials or objects by measurement of the degree of racemization of the amino acids present. With the exception of glycine, the amino acids found in proteins can exist in two isomeric forms called D- and L-enantiomers. Although the enantiomers of an amino acid rotate plane-polarized light in equal but opposite directions, their other chemical and physical properties are identical. Amino acid handedness or homochirality is one of the most distinctive features of terrestrial life. It was discovered by L. Pasteur around 1850 that only L-amino acids are generally found in living organisms, but scientists still have not formulated a convincing reason to explain why life on Earth is based on only L-amino acids. *See* AMINO ACIDS; ENANTIOMER.

Racemization. Under conditions of chemical equilibrium, equal amounts of both enantiomers are present ($D/L = 1.0$); this is called a racemic mixture. Living organisms maintain a state of disequilibrium through a system of enzymes that selectively utilize only the L-enantiomers. Once a protein has been synthesized and isolated from active metabolic processes, the L-amino acids are subject to a racemization reaction that converts them into a racemic mixture. The racemization reaction can be written as (1), where k_i is the first-order rate constant for the

$$\text{L-amino acid} \underset{k_i}{\overset{k_i}{\rightleftharpoons}} \text{D-amino acid} \quad (1)$$

interversion of the enantiomers. The rate of racemization is equal to $2k_i$. The kinetic equation of the racemization reaction is given by (2), where D/L is

$$\ln\left(\frac{1+(D/L)}{1-(D/L)}\right) - \ln\left(\frac{1+(D/L)}{1-(D/L)}\right)_{t=0} = 2k_i t \quad (2)$$

the ratio of the enantiomers of a certain amino acid at a particular time t. The $t = 0$ term is necessary to account for some racemization that occurs during sample preparation.

Since racemization is a chemical process, the extent of racemization is dependent not only on the time that has elapsed since the L-amino acids were synthesized but also on the exposure temperature: the higher the temperature, the faster the rate of racemization. The rate of racemization is also different for most of the various amino acids. The half-life (for example, the time required to reach a D/L ratio of 0.33) at neutral pH for aspartic acid, an amino acid with one of the fastest racemization rates, is approximately 3500 years at 25°C (77°F) but is only 35 days at 100°C (212°F). Amino acids with the slowest racemization rates have half-lives roughly 10 times greater than those of aspartic acid. *See* HALF-LIFE.

A variety of analytical procedures can be used to separate amino acid enantiomers; gas chromatography and high-performance liquid chromatography are the most widely used. Since these techniques have sensitivities in the parts-per-billion range, only a few hundred milligrams of sample material are

normally required. Samples are first hydrolyzed in hydrochloric acid to break down the proteins into free amino acids, which are then isolated by cation-exchange chromatography. *See* CHROMATOGRAPHY.

Since the late 1960s, the geochemical and biological significance of amino acid racemization has been extensively investigated. Geochemical uses of amino acid racemization include the dating of fossils or, in the case of known age specimens, the determination of their temperature history. Fossil types such as bones, teeth, and shells have been studied, and racemization has been found to be particularly useful for dating specimens that were difficult to date by other methods. Racemization has also been observed in the metabolically inert tissues of living mammals. Racemization can be studied in certain organisms and used to assess the biological age of a variety of mammalian species; in addition, it may be important in determining the biological lifetime of certain proteins. *See* RACEMIZATION.

Dating fossils. Fossils have been found to contain both D- and L-amino acids, and the extent of racemization generally increases with geologic age. The amount of racemization in a fossil is determined by the value of k_i and its age. The magnitude of k_i in a fossil from a particular locality is primarily a function of a value known as the integrated exposure temperature, although protein diagenetic processes (for example, geochemical degradation and alteration) are also important. The rate constant k_i can be estimated by using a calibration procedure, wherein the *D/L* ratio is measured in a fossil of known age from the study area. This *D/L* ratio and the calibration age of the sample are substituted into Eq. (2); an in-place k_i value is thus calculated. This calibrated rate constant integrates variations in exposure temperature and other environmental parameters of the locality over the time period represented by the age of the calibration sample. Following calibration, the k_i value can be used, with certain limitations, to date other samples from the surrounding region. The principal limitation is that the calibration sample should have a similar amino acid composition and content, and should be from roughly the same geologic period, as the sample being dated.

The Olduvai Gorge region in the north-central Tanzanian Rift Valley offers an excellent opportunity to study the racemization reaction of amino acids in fossil bones and teeth over a time period extending back several million years. Since the geology of the region has been extensively studied as a result of the discovery of numerous hominid fossils and artifacts, the relative and absolute ages of the various stratigraphic units in the Olduvai region are well established. Thus, amino acid racemization can be investigated over several geologic periods.

Aspartic acid racemization can be used to date only the Holocene and perhaps upper Pleistocene deposits in the Olduvai region. The k_i values determined for aspartic acid at Olduvai suggest that bones and teeth older than about 60,000 years should contain only racemic aspartic acid, for example, *D/L* aspartic acid = 1.0. However, fossil bones from the older stratigraphic units at Olduvai have been found to contain nonracemic aspartic acid. This indicates that these bones contain secondary aspartic acid, which probably is incorporated into fossil bone by percolating ground waters. Aspartic acid racemization of bones is of limited use in dating the important anthropological deposits in the Olduvai Gorge region because it racemizes rapidly and is prone to contamination problems.

In contrast, the epimerization of L-isoleucine, which yields the nonprotein amino acid D-alloisoleucine, provides a much better chronological tool for dating the deposits in the Olduvai region. Isoleucine has two centers of asymmetry, and thus the reaction is termed epimerization rather than racemization; the equilibrium ratio of D-alloisoleucine to L-isoleucine is 1.3, and so Eq. (2) must be slightly modified to account for this. In fossil bones and teeth ranging in age from the upper Pleistocene to the Pliocene, the aile/ile ratio in tooth enamel steadily increases until an approximate equilibrium ratio of 1.3 is attained in the oldest deposits. This indicates that isoleucine epimerization in teeth is effectively a closed system with respect to the introduction of secondary isoleucine for a period of 3–4 million years. Isoleucine may be less susceptible to contamination than aspartic acid because the peptide bonds involving this amino acid are more stable with respect to hydrolysis than are those containing aspartic acid.

By using the enamel sample from the base of Olduvai Bed I for calibration, isoleucine epimerization ages for the middle and lower Pleistocene deposits are obtained that are in good agreement with stratigraphic and radiometric age determinations. Some bones also yield reasonable ages. For example, the isoleucine epimerization age for a bone from the lower unit of the Masek beds is about 500,000 years, which is consistent with the age estimated from other evidence. Bones from older stratigraphic units, however, have been found to have aile/ile ratios less than those measured in tooth enamel. This implies that isoleucine in bones is more susceptible to contamination than is isoleucine in enamel. Isoleucine epimerization ages of bones from the older deposits should probably be viewed with caution and should be considered minimum age estimates. Contamination should consist primarily of L-amino acids; thus, incorporation of secondary amino acids into a fossil would lower the *D/L* ratio and yield an age estimate that is too young.

The k_i values for isoleucine epimerization in the upper Pleistocene deposits are about four times greater than those determined by using the lower Pleistocene Bed I sample. The faster rate of isoleucine epimerization in the younger deposits may be due to several factors, such as different exposure temperatures and the complex process of protein diagenesis. The upper Pleistocene k_i value should be used only to date upper Pleistocene samples; it is not applicable to dating the older stratigraphic units.

In general, the ages based on racemization and epimerization at Olduvai Gorge are consistent with those estimated from other evidence. The best results are obtained by using tooth enamel. On the basis of these results, the racemization and epimerization reactions in bones and teeth have been investigated at many localities throughout the world. Several hominid remains have been directly dated, and the technique has been used to evaluate dates derived from geologic, radiometric, and faunal evidence.

Other racemization studies have utilized both terrestrial and marine fossil shells to estimate the age of various archeological sites as well as uplifted marine terrace deposits. In these cases, the limit for racemization dating is in the 100,000-year range, although at high latitudes where temperatures are cold the age range is extended nearly 1 million years. *See* FOSSIL.

Racemization in living mammals. Enamel, dentine, and the proteins present in the nucleus of the eye lens are metabolically inert and are thus incubated at approximately 37°C (98.6°F) throughout a mammal's life span. In mammals that live for decades, detectable racemization of aspartic acid has been found to take place in these tissues. Although the extent of racemization is quite small (for example, *D/L* ratios are in the range 0.02–0.10), it can be easily and precisely measured with available analytical techniques. Only aspartic acid racemization has been detected, which is expected since this amino acid has the fastest racemization rate.

The racemization of amino acids thus can be used as a biochronological tool. This aging method is particularly useful for determining the ages of humans and certain other mammals, whose ages are not well documented. An interesting application of the racemization technique has been the determination of the ages of various marine mammals. For example, the ages of narwhals (*Mondon monoceros*), Arctic cetaceans that are characterized by the long (about 6.5 ft or 2 m) counterclockwise-spiraling tusk of the male, have been deduced by using the extent of aspartic acid racemization in their teeth. The method is particularly useful for determining the ages of female narwhals whose ages previously were largely unknown. The aspartic acid measurements indicate that both female and male narwhals can live more than 50 years.

The method has also been used to estimate the ages of bowhead whales that live in the Arctic. During the recent hunting of these animals by Arctic peoples, some whales were found to have embedded in their skin stone harpoon tips, which were no longer used in the Arctic after around 1880. This finding suggested that bowhead whales might have unusually long life spans and, in order to help verify this surmise, a systematic study of the racemization in the eye lens nucleus was carried out. Extensive racemization was observed in the eye lens nucleus of some animals, and based on these results their ages were estimated to be in excess of 100 years. In the case of one whale, the calculated age was close to 200 years. These findings suggest that bowhead whales may be some of the most aged mammals in the world.

The extent of aspartic acid racemization in enamel and dentine can also be used to estimate the death age of humans from ancient burials. The main requirement in this application is that either the burial should be fairly recent or the burial environmental temperature should be relatively low, so that the extent of postmortem racemization is small compared with that which occurred during the lifetime of the individual. The racemization method has been used to calculate the age at death of mummified Alaskan Eskimos, of individuals in a medieval cemetery in Czechoslovakia, and of corpses in forensic cases. The racemization ages have greater accuracy (about ±5%) than those estimated by other means, and the method is especially useful in determining the ages of older individuals, whose age at death is often difficult to estimate from anatomical evidence. *See* ARCHEOLOGICAL CHEMISTRY; GEOCHRONOMETRY; RADIOCARBON DATING; ROCK AGE DETERMINATION.

Other applications of amino acid racemization. Amino acid racemization provides a general indicator of the level of preservation of biomolecules in geological specimens. The extent of racemization of amino acids correlates with the level of overall degradation of labile molecules. One example deals with the preservation of ancient DNA. It has been found that only in specimens in which the extent of racemization is minimal ($D/L < 0.1$) are fragments of original DNA preserved. This has proved to be a valuable index for judging if DNA in fossils is original or derived from contamination. *See* DEOXYRIBONUCLEIC ACID (DNA).

The reaction has implications with respect to protein turnover in mammalian tissues. It has been shown that there is an increasing level of water-insoluble protein present in the eye lens nucleus, of long-lived mammals with increasing age. This is correlated with increasing racemization in the eye lens nucleus, and it has been suggested that structural changes induced by racemization result in protein insolubility. Amino acid racemization has also been found to take place during the processing of various food materials. This is caused by both heat exposure during cooking and the treatment of food components with alkali in order to make them more palatable. The racemization induced by these treatments may have deleterious nutritional effects because it has been found that proteins containing racemized amino acids are less digestible.

The racemization of amino acids has implications in the search for life beyond Earth. It is assumed that life elsewhere would require homochiral amino acids, although whether it would be based on L- or D-amino acids is considered to be mainly a matter of a chance selection event that occurred either at the time of the origin of life or during early evolution. Detection of this amino acid homochirality signature would depend on whether life still existed on another world or, if it became extinct, whether racemization has erased the signature. In the case of

Mars, using racemization rates determined in terrestrial environments, the survival time of an amino acid homochiral signature has been estimated for the conditions relevant to the surface of Mars throughout the planet's history. Homochiral amino acids associated with a Martian biota that became extinct during the early history of the planet would still be preserved today in dry, cold (<250 K) Martian environments. In contrast, if liquid water was present for extended periods (several millions of years) after life on Mars became extinct, racemization would have converted any originally homochiral amino acids into a racemic mixture. These considerations are important in the search for evidence of life on Mars and other solar system bodies during future space missions designed to search for evidence of extraterrestrial life. *See* MARS.

Jeffrey L. Bada

Bibliography. J. L. Bada, Amino acid racemization dating of fossil bones, *Annu. Rev. Earth Planet. Sci.*, 13:241–268, 1985; J. L. Bada, *In vivo* racemization in mammalian proteins, *Meth. Enzymol.*, 106:98–115, 1984; J. L. Bada and G. D. McDonald, Amino acid racemization on Mars: Implications for the preservation of biomolecules from an extinct Martian biota, *Icarus*, 114:139–143, 1995; J. C. George et al., Age and growth estimates of bowhead whales (*Balaena mysticetus*) via aspartic acid racemization, *Can. J. Zool.*, 77:571–580, 1999; H. Y. Goksu, M. Oberhofer, and D. Regulla (eds.), *Scientific Dating Methods*, 1991; W. C. Mahaney (ed.), *Quaternary Dating Methods*, 1984; H. N. Poinar et al., Amino acid racemization and the preservation of ancient DNA, *Science*, 272:864–866, 1996; M. R. Zimmerman and J. L. Angel (eds.), *Dating and Age Determination of Biological Materials*, 1986.

Amino acids

Organic compounds possessing one or more basic amino groups and one or more acidic carboxyl groups. Of the more than 80 amino acids which have been found in living organisms, about 20 serve as the building blocks for the proteins.

All the amino acids of proteins, and most of the others which occur naturally, are α-amino acids, meaning that an amino group (—NH_2) and a carboxyl group (—COOH) are attached to the same carbon atom. This carbon (the α carbon, being adjacent to the carboxyl group) also carries a hydrogen atom; its fourth valence is satisfied by any of a wide variety of substituent groups, represented by the letter R in **Fig. 1**.

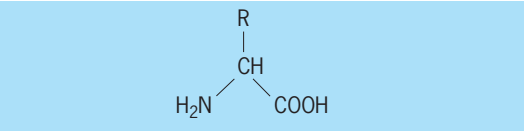

Fig. 1. Structural formula for an amino acid.

Amino acids of proteins, grouped according to the nature of R

Amino acids*	R
Glycine	Hydrogen
Alanine, valine, leucine, isoleucine	Unsubstituted aliphatic chain
Serine, threonine	Aliphatic chain bearing a hydroxyl group
Aspartic acid, glutamic acid	Aliphatic chain terminating in an acidic carboxyl group
Asparagine, glutamine	Aliphatic chain terminating in an amide group
Arginine, lysine	Aliphatic chain terminating in a basic amino group
Cysteine, cystine, methionine	Sulfur-containing aliphatic chain
Phenylalanine, tyrosine	Terminates in an aromatic ring
Tryptophan, proline, histidine	Terminates in a heterocyclic ring

* See articles on the individual amino acids listed in the table.

In the simplest amino acid, glycine, R is a hydrogen atom. In all other amino acids, R is an organic radical; for example, in alanine it is a methyl group (—CH_3), while in glutamic acid it is an aliphatic chain terminating in a second carboxyl group (—CH_2—CH—COOH). Chemically, the amino acids can be considered as falling roughly into nine categories based on the nature of R (see **table**).

Occurrence of conjugated amino acids. Amino acids occur in living tissues principally in the conjugated form. Most conjugated amino acids are peptides, in which the amino group of one amino acid is linked to the carboxyl group of another. This type of linkage is known as a peptide bond; a molecule of water is split out when a peptide bond is formed, and a molecule of water must be added when a peptide bond is broken, as shown in the reaction below.

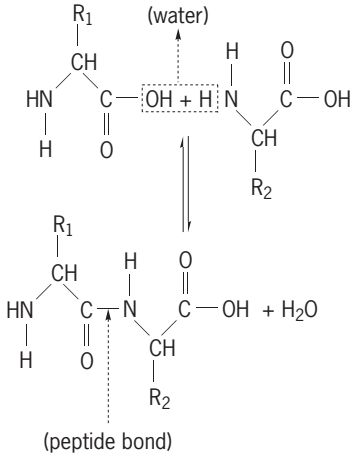

Since each amino acid possesses both an amino group and a carboxyl group, the acids are capable of linking together to form chains of various lengths, called polypeptides. Proteins are polypeptides ranging in size from about 50 to many thousand amino acid residues. The process by which peptides are

formed from free amino acids actually cannot be as simple as pictured in the equation, for a considerable amount of energy is required. This process is discussed later in this article.

Although most of the conjugated amino acids in nature are proteins, numerous smaller conjugates occur naturally, many with important biological activity. The line between large peptides and small proteins is difficult to draw, with insulin (molecular weight = 7000; 50 amino acids) usually being considered a small protein and adrenocorticotropic hormone (molecular weight = 5000; 39 amino acids) being considered a large peptide. In addition to their role as hormones, peptides often occur in coenzymes (such as folic acid and glutathione), bacterial capsules (the polyglutamic acid capsule which contributes to the pathogenicity of *Bacillus anthracis*), fungal toxins (the tomato wilt toxin of *Fusarium* and the toxins of the poisonous mushroom *Amanita phalloides*), and antibiotics (chloramphenicol, penicillin, bacitracin, and polymixins). Elucidation of the structure of bacterial cell walls has shown that they are composed in part of a series of cross-linked peptides, the cross-linking providing the wall with a large part of its rigidity. The action of penicillin in inhibiting this cross-linking reaction accounts for its antibiotic activity. Finally, a considerable part of the phospholipid fraction of any organism contains serine linked by phosphoester bond to glycerol phosphate. *See* ANTIBIOTIC; COENZYME; TOXIN.

Occurrence of free amino acids. Free amino acids are found in living cells, as well as the body fluids of higher animals, in amounts which vary according to the tissue and to the amino acid. The amino acids which play key roles in the incorporation and transfer of ammonia, such as glutamic acid, aspartic acid, and their amides, are often present in relatively high amounts, but the concentrations of the other amino acids of proteins are extremely low, ranging from a fraction of a milligram to several milligrams per 100 g wet weight of tissue. In view of the fact that amino acid and protein sythesis go on constantly in most of these tissues, the presence of free amino acids in only trace amounts points to the existence of extraordinarily efficient regulation mechanisms. Each amino acid is ordinarily synthesized at precisely the rate needed for protein synthesis. The regulation mechanism has been found most often to be one of feedback control; each amino acid acts as an inhibitor of its own biosynthesis. If any amino acid is formed in excess of that required for protein synthesis, the biosynthesis of that amino acid is slowed down until the excess has been used.

In addition to the amino acids of protein, a variety of other free amino acids occurs naturally. Some of these are metabolic products of the amino acids of proteins; for example, γ-aminobutyric acid occurs as the decarboxylation product of glutamic acid. Others, such as homoserine and ornithine, are biosynthetic precursors of the amino acids of protein. However, the origin and role of many unusual free amino acids is not yet known.

General properties. The amino acids are characterized physically by the following: (1) the pK_1, or the dissociation constant of the various titratable groups; (2) the isoelectric point, or pH at which a dipolar ion does not migrate in an electric field; (3) the optical rotation, or the rotation imparted to a beam of plane-polarized light (frequently the D line of the sodium spectrum) passing through 1 decimeter of a solution of 100 grams in 100 milliliters; and (4) solubility. *See* IONIC EQUILIBRIUM; ISOELECTRIC POINT; OPTICAL ACTIVITY.

At ordinary temperatures, the amino acids are white crystalline solids; when heated to high temperatures, they decompose rather than melt. They are stable in aqueous solution, and with few exceptions can be heated as high as 120°C (248°F) for short periods without decomposition, even in acid or alkaline solution. Thus, the hydrolysis of proteins can be carried out under such conditions with the complete recovery of most of the constituent free amino acids. The exceptions are as follows: Acid hydrolysis of protein destroys most of the tryptophan and some of the serine and threonine, oxidizes cysteine to cystine, and deamidates glutamine and asparagine; alkaline hydrolysis destroys serine, threonine, cystine, cysteine, and arginine, and also causes deamidations.

Enantiomorphs. Since all of the amino acids except glycine possess a center of asymmetry at the α carbon atom, they can exist in either of two optically active, mirror-image forms or enantiomorphs. All of the common amino acids of proteins appear to have the same configuration about the α carbon; this configuration is symbolized by the prefix L-. The opposite, generally unnatural, form is given the prefix D-. Some amino acids, such as isoleucine, threonine, and hydroxyproline, have a second center of asymmetry and can exist in four stereoisomeric forms. The prefix allo- is used to indicate one of the two alternative configurations at the second asymmetric center; thus, isoleucine, for example, can exist in the L, L-allo, D, and D-allo forms. *See* STEREOCHEMISTRY.

Unlike chemical syntheses, which lead to mixtures of D and L forms, biosynthetic processes invariably produce optically active amino acids. For most amino acids, only the L isomer occurs naturally; but in a few cases, the D isomer is found also. For example, the cell walls of certain bacteria contain D-alanine and D-glutamic acids, and the D isomers of phenylalanine, leucine, serine, and valine occur in some antibiotic peptides.

Ionic state. Another important general property of all amino acids is their ionic state (**Fig. 2**). The basic amino group can bind a proton from solution and become a cation; the acidic carboxyl group can release a proton into solution and become an anion. At the isoelectric point (the pH at which the molecule has no net charge), amino acids exist as dipolar ions or zwitterions, while in strong acid solution, the carboxyl group exists in the undissociated form, and the molecule becomes a cation. If such an acidic solution is titrated with strong alkali, two

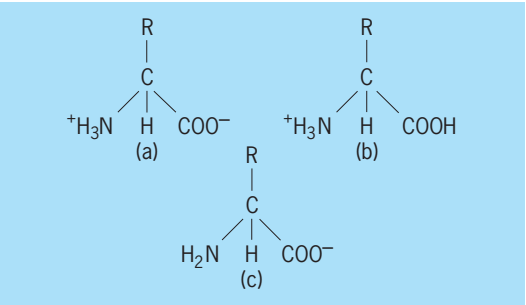

Fig. 2. Ionic states of an amino acid. (*a*) Zwitterion. (*b*) Cation. (*c*) Anion.

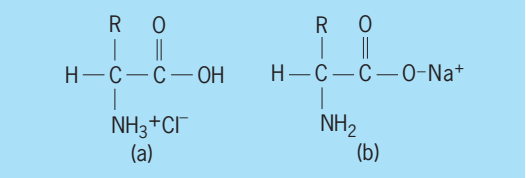

Fig. 3. Salts of amino acids. (*a*) Amino acid hydrochloride. (*b*) Sodium amino acid.

dissociations of protons are observed. The carboxyl group, having the weakest affinity for its proton, dissociates at a fairly low pH; its p*K* (the pH at which half of the molecules are dissociated) in most cases is close to 2.0. As more alkali is added, the proton on the amino group begins to dissociate; p*K* values for this dissociation are generally found close to 9.5. When sufficient alkali has been added to pull off all the dissociable protons, the amino acid exists as an anion.

Since the amino acids are ions, they can be prepared as their salts. For example, the titration of an amino acid solution with hydrochloric acid (HCl) leads to formation of the amino acid hydrochloride, while titration with sodium hydroxide (NaOH) forms the sodium salt (**Fig. 3**).

When the R radical contains an ionizable group, the amino acid will have correspondingly more ionic forms. Those amino acids whose radicals contain carboxyl groups (aspartic and glutamic acids) are known as the acidic amino acids, since a solution of the zwitterion will be strongly acidic. Similarly, histidine, lysine, and arginine are known as the basic amino acids, and the rest as the neutral amino acids. (It is important to note that a solution of the zwitterion of a neutral amino acid will in fact be slightly acidic.)

The salts are, in general, more soluble in water or alcohol than the corresponding zwitterions.

Isolation and determination. Since most amino acids occur in conjugated form, their isolation usually requires their prior release in free form by acid or alkaline hydrolysis. Hydrolysates of proteins or other polypeptides, or crude extracts of plants, animal, or microbial materials, serve as the starting point for the isolation in pure form of single amino acids. Prior to the application of chromatography in the early 1940s, the isolation of amino acid depended on slight differences in the solubilities of amino acid salts in various solvents and at different pH values. For example, the isolation of aspartic acid was accomplished by adding an excess of calcium hydroxide to an aqueous solution of amino acids and then precipitating the calcium aspartate with alcohol.

Such methods, although used successfully to isolate each of the common amino acids of protein, are difficult as well as tedious, and require relatively large amounts of starting material. Chromatography, on the other hand, is simple, rapid, and capable of isolating amino acids even when they are present in microgram quantities. Thus chromatography has been the method of choice for amino acid isolation ever since its first application by A. J. P. Martin and R. L. M. Synge in 1941. *See* CHROMATOGRAPHY.

Chromatography is carried out by using either cylindrical glass tubes (columns) packed with a porous solid or by using sheets of filter paper. In the former method, the column is packed with any of a variety of substances, such as starch, powdered cellulose, or cation-exchange resin, and is saturated with the chosen solvent. A solution of amino acids is allowed to percolate into the top of the column,

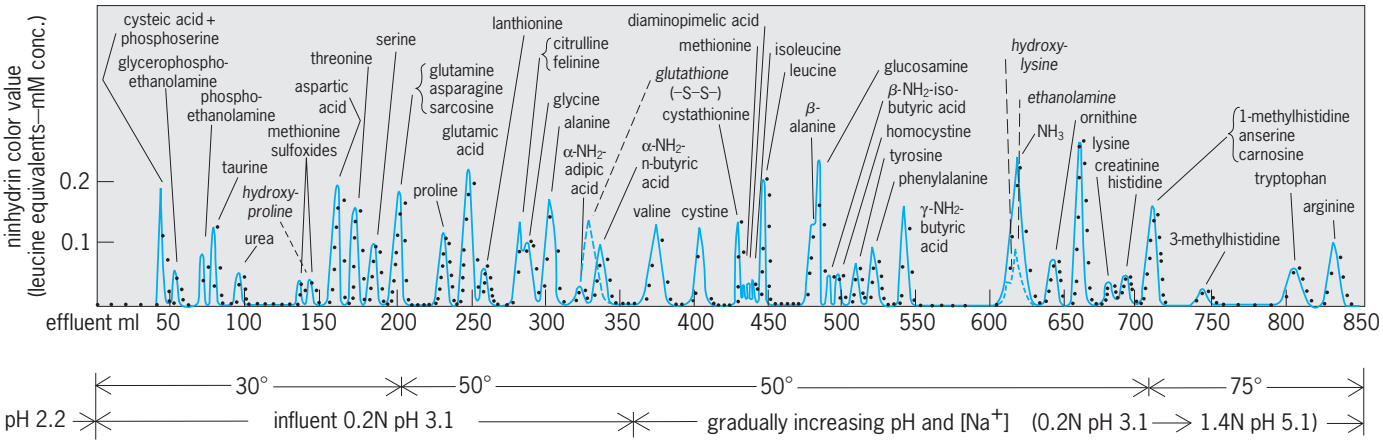

Fig. 4. Typical protein fractionation into amino acids shown in a cation-exchange resin column chromatogram. (*After S. Moore and W. H. Stein, J. Biol. Chem., 211:893–906, 1954*)

and then solvent is forced through the column at a controlled rate. The solvent is allowed to flow out through an opening at the bottom of the column, and the eluate is caught in a series of test tubes.

A given amino acid will have been eluted from the column at a time depending on its own characteristic rate of movement, and will have been caught in one or a few tubes; a different amino acid will have been eluted into a separate set of tubes. To detect their presence, as well as to determine their exact quantity, a substance which reacts with amino acids to give a visible color is then added to each tube. The best such reagent is ninhydrin, which reacts with amino acids to produce carbon dioxide, ammonia, and aldehyde, and forms a purple compound with the liberated ammonia. The amount of color which develops under standard conditions can then be measured in a photoelectric colorimeter, and the precise amount of amino acid determined by comparison with a standard curve based on reactions with known quantities. **Figure 4** shows a typical separation by passing water adjusted to different pH values and temperatures through a column of cation-exchange resin, collecting fractions, and determining the amount of amino acid in each with ninhydrin.

The column method described above is capable of giving the most precise quantitative data, and can handle relatively large quantities of amino acids. For the analysis of mixtures containing only a few micrograms of each amino acid, however, paper chromatography is the most simple and rapid procedure, and can also be made quantitative. In this procedure, the mixture of amino acids is applied as a drop of solution to a spot close to one corner of a sheet of paper. The sheet is then placed in a vapor-tight chamber with one edge of the paper dipping into a chosen, water-saturated organic solvent. The solvent flows through the paper by capillarity, water becoming bound to the paper and the organic solvent flowing past it. The amino acids travel through the paper as discrete spots, exactly as described above for their travel through a column. When the solvent has traveled to the opposite edge of the paper, the sheet is removed, dried, and sprayed with a reagent such as ninhydrin. The position of each amino acid is then revealed by a colored spot which appears when the sheet is heated.

If the solvent used brings two or more amino acids to the same position, two-dimensional chromatography is employed. The dried sheet is not sprayed, but is rotated 90° and is placed with the edge along which the amino acids are located in a second solvent. This solvent flows through the paper at right angles to the direction taken by the first solvent, and if correctly chosen, will separate those amino acids which stayed together in the first solvent. The sheet is then dried and sprayed to locate the amino acids, and then the spots can be cut out with scissors and eluted with water in separate tubes for colorimetric determination. A two-dimensional chromatogram of the amino acids of a protein is shown in **Fig. 5**.

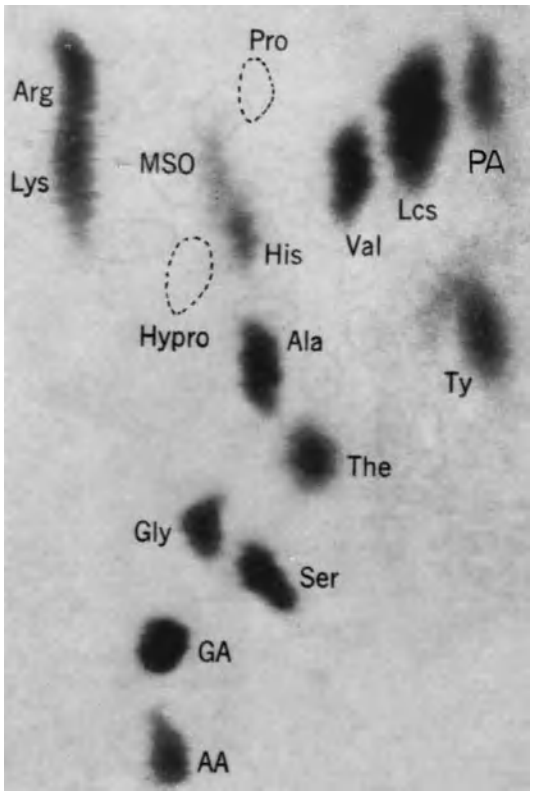

KEY:

Arg = arginine	His = histidine
Lys = lysine	Ala = alanine
MSO = methionine sulfoxide	The = threonine
Hypro = hydroxyproline	Se = serine
Gly = glycine	Val = valine
GA = glutamic acid	Lcs = leucines
AA = aspartic acid	Ty = tyrosine
Pro = proline	PA = phenylalanine

Fig. 5. Two-dimensional chromatogram of a protein; the solvents were phenol and lutidine. (*From R. J. Block, E. L. Durrum, and G. Zweig, A Manual of Paper Chromatography and Paper Electrophoresis, 2d ed., Academic Press, 1958*)

The determination of total α-amino acids in extracts of natural materials can be carried out by using the ninhydrin method, described above, or the Van Slyke method, which measures the amount of nitrogen gas given off on treatment with nitrous acid. Specific determinations of a given amino acid were previously carried by means of sensitive microbiological assays which measured the growth of a microorganism dependent on the amino acid as a function of the concentration of the unknown material. Nowadays this rather difficult and cumbersome method has been largely replaced by ion-exchange chromatography.

There are relatively specific color reactions for many amino acids; while still used to some extent as quantitative assays, these now find most widespread

use as qualitative tests for the presence of the amino acids.

Amino Acid Metabolism

Although amino acids and some other charged molecules can enter a cell passively by means of simple diffusion, there are present in all cells so far examined special systems for concentrating such small molecules inside the cell. These systems, called permeases, are localized in the cell membrane. They are protein complexes, probably containing at least two parts, which utilize metabolic energy to transport small molecules against the concentration gradient. Bacterial amino acid permeases are capable of achieving a concentration inside a cell 1000 times greater than that outside the cell.

The criteria for a permease, or active transport system, are (1) it must require energy, (2) it must be relatively specific, and (3) it must concentrate the transported substance against a gradient. Most permeases obey the classical Michaelis-Henry enzyme kinetics. Portions of the system which concentrates valine, isoleucine, and leucine in *Escherichia coli* have been purified; thus the mechanism of action of this permease may soon be clear.

Biosynthesis. Since amino acids, as precursors of proteins, are essential to all organisms, all cells must be able to synthesize those they cannot obtain from their environment. The selective advantage of being able rapidly to shift from endogenous to exogenous sources of these compounds has led to the evolution in bacteria and many other organisms of very complex and precise methods of adjusting the rate of synthesis to the available level of the compound. These regulatory mechanisms can be divided according to whether they require a short or a long time to take effect.

The immediately effective control is that of feedback inhibition. As Figs. 9–14 show, the biosynthesis of amino acids is relatively complicated and usually requires at least three enzymatic steps. In most cases so far examined, the amino acid end product of the biosynthetic pathway inhibits the first enzyme to catalyze a reaction specific to the biosynthesis of that amino acid. This inhibition is extremely specific; the enzymes involved have special sites for binding the inhibitor. This inhibition functions to shut off the pathway in the presence of transient high levels of the product, thus saving both carbon and energy for other biosynthetic reactions. When the level of the product decreases, the pathway begins to function once more.

The one exception to this general rule is the growth factor requirement which results from the presence in the environment of a metabolic inhibitor. For example, a certain strain of bacterium is very sensitive to inhibition by valine, and this inhibition can be overcome by isoleucine. In the presence of valine, then, this strain requires isoleucine for growth. There are many such antagonisms between amino acids, with the result that organisms which require several amino acids must receive them in balanced amounts; any one in excess may prove inhibitory.

If a microorganism is grown for several generations in the presence of an amino acid, the levels of the enzymes of the biosynthetic pathway decrease considerably. This phenomenon is called enzyme repression, and it comes about because the synthesis of the enzymes is decreased in the presence of the end product. However, the enzyme already present in the cell is stable; therefore several generations are required before it is diluted to its lowest level by being apportioned among daughter cells at each division. (The level of such enzymes never reaches zero.) If at any time during this process the amino acid ceases to be available to the microorganism, synthesis of the enzyme immediately begins and continues until the proper intracellular concentration of the amino acid is reached. At this point the biosynthesis of the enzyme slows down and an equilibrium is reached such that the level of the enzyme remains constant until there is another alteration in the exogenous level of the amino acid. In contrast to feedback inhibition, which requires only milliseconds to act, repression and derepression require from one-half to several generations to reach a new equilibrium.

The actual metabolic pathways by which amino acids are synthesized are presented in diagrammatic form in Figs. 9–14. These pathways generally are found to be the same in all living cells investigated, whether microbial or animal. Biosynthetic mechanisms thus appear to have developed soon after the origin of life and to have remained unchanged through the divergent evolution of modern organisms. The major exception is lysine, which is formed from aspartic acid via diaminopimelic acid in bacteria, but from α-ketoglutaric acid in the fungi. Indeed, the occurrence of diaminopimelic acid as a precursor of lysine, or as a constituent of proteins, or both, is a major taxonomic property of the bacteria and the related blue-green algae.

Formation and transfer of amino groups. The biosynthetic pathway diagrams reveal only one quantitatively important reaction by which organic nitrogen enters the amino groups of amino acids: the reductive amination of α-ketoglutaric acid to glutamic acid by the enzyme glutamic acid dehydrogenase. All other amino acids are formed either by transamination (transfer of an amino group, ultimately from glutamic acid) or by a modification of an existing amino acid. An example of the former is the formation of valine by transfer of the amino group from glutamic acid to α-ketoisovaleric acid; an example of the latter is the reduction and cyclization of glutamic acid to form proline.

Two other direct conversions of inorganic nitrogen to amino acid nitrogen are known: the reductive amination of pyruvic acid to alanine and the addition of ammonia to fumaric acid to form aspartic acid. However, there is no evidence that either of these reactions is quantitatively important in amino nitrogen formation. In any case, ammonia is the only form

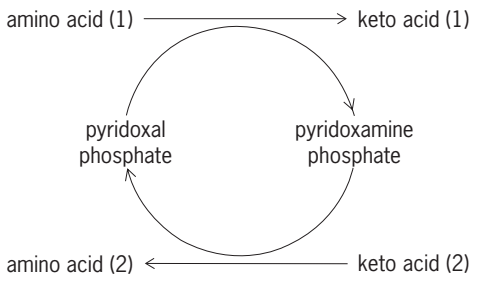

Fig. 6. Graphic representation of transamination mechanism by true transaminase.

of inorganic nitrogen which has been clearly shown to enter organic compounds directly; nitrate (NO_3^-), nitrite (NO_2^-), and nitrogen gas (N_2) are probably reduced to free intracellular ammonia before being converted to organic form in those plants and microorganisms which can use them as a nutritional source of nitrogen.

The principal mechanism of amino group transfer, transamination, is extremely important in many phases of nitrogen metabolism. Although many transamination reactions are known (all or most naturally occurring amino compounds probably participate in transamination in one tissue or another), the actual number of transaminases involved is uncertain. The few transaminases which have been highly purified all catalyze amino group exchange between more than just one pair of amino acids.

A true transaminase uses pyridoxal phosphate or pyridoxamine phosphate as coenzyme; the amino group is transferred to the former, which then gives it up to the keto acid acceptor (**Fig. 6**). A different mechanism of amino group transfer occurs in the biosyntheses of arginine and of adenylic acid; here, aspartic acid is added to a keto group to form a stable intermediate; a second enzyme then cleaves the intermediate to fumaric acid plus the new amino compound (see Fig. 10, showing the pathway of arginine biosynthesis).

One other important route by which ammonia enters organic compounds is by way of the amide group of glutamine. This group is formed by the direct addition of ammonia to glutamic acid, the necessary energy coming from the breakdown of adenosine triphosphate (ATP), first to adenosine diphosphate (ADP), then to inorganic phosphate. Once formed, amide nitrogen can be transferred to suitable acceptors to form precursors of the purine and histidine rings, as well as to hexose 6-phosphate to form glucosamine 6-phosphate.

Asparagine is another important amide, but the mechanism of asparagine formation is still in some doubt, and the only known product of the asparagine amide group is free ammonia. Glutamine is also readily deamidated to ammonia; both glutamine and asparagine serve as important storage forms of ammonia in higher plants and animals, as well as being constituent amino acids of proteins.

Degradation. Most organisms are capable of degrading some amino acids, and metabolic pathways leading to degradation to CO_2 and H_2O are known for each of the common amino acids. These pathways are detailed in the articles on the individual amino acids. There are, however, general features (**Fig. 7**) characteristic of degradative pathways which are discussed below.

The first step in the degradation of all amino acids, with the exception of tyrosine and phenylalanine, is the labilization of one of the four groups on the α-carbon atom. The labilization always involves an enzyme containing pyridoxal phosphate as a cofactor, except in the case of oxidative deamination where a flavoprotein is involved. Pyridoxal phosphate acts by forming a Schiff's base (shown as **1** in Fig. 7) with the amino acid, as shown in Fig. 7. By a rearrangement of electrons and protons (a tautomerization) the double bond moves to the position shown in **2**, with a concomitant alteration in the electron distribution in the pyridine ring of pyridoxal phosphate and the loss of the hydrogen on the α-carbon to the solvent. Intermediate **2** may now be hydrolyzed to yield the α-keto acid (**3**) corresponding to the amino acid, or it may return to intermediate **1**. In the latter case there is often a racemization, since **2** is a symmetric compound and the reversal of the tautomerization is not always carried out in an asymmetric manner.

If, instead of the hydrogen atom, the carboxyl group is labilized by donating its electrons to form the new double bond, intermediate **2***a* is formed. A reversal of the tautomerization cannot regenerate an amino acid in this case; the reaction is a decarboxylation and the product is an amine. Similarly, the organic radical R may be lost (**2***b*) whereupon the reversal of the tautomerization yields glycine. The last reaction is possible only when the radical is substituted with a hydroxyl group on the β-carbon (adjacent to the α-carbon). A final variation on the basic reactions of pyridoxal phosphate takes place when the R group contains an electronegative substituent on the β-carbon (**4**). In this case it is possible to expel the substituent, leading to an allyl amino acid, intermediate **5**, analogous to **1**. Intermediate **5** is now hydrolyzed to yield the extremely unstable intermediate **6** which rearranges to give the keto acid, ammonia, and X^- from the amino acid.

Transamination. Transamination is accomplished by hydrolysis of intermediate **2** to yield the keto acid and pyridoxamine phosphate. The latter compound reacts with another keto acid, yielding pyridoxal phosphate and the new amino acid. Transaminase enzymes are relatively nonspecific, reacting with groups of amino acids with similar R groups (for example, the aromatic amino acids or the branched-chain amino acids). The equilibrium constant of a transaminase reaction is usually very close to 1, reflecting the similarity in the free energies of formation of keto acids from amino acids. Transamination is much commoner than deamination in nature, since a deamination without a subsequent transfer to the amino group to another keto acid renders the coenzyme inactive. Most transaminase reactions involve

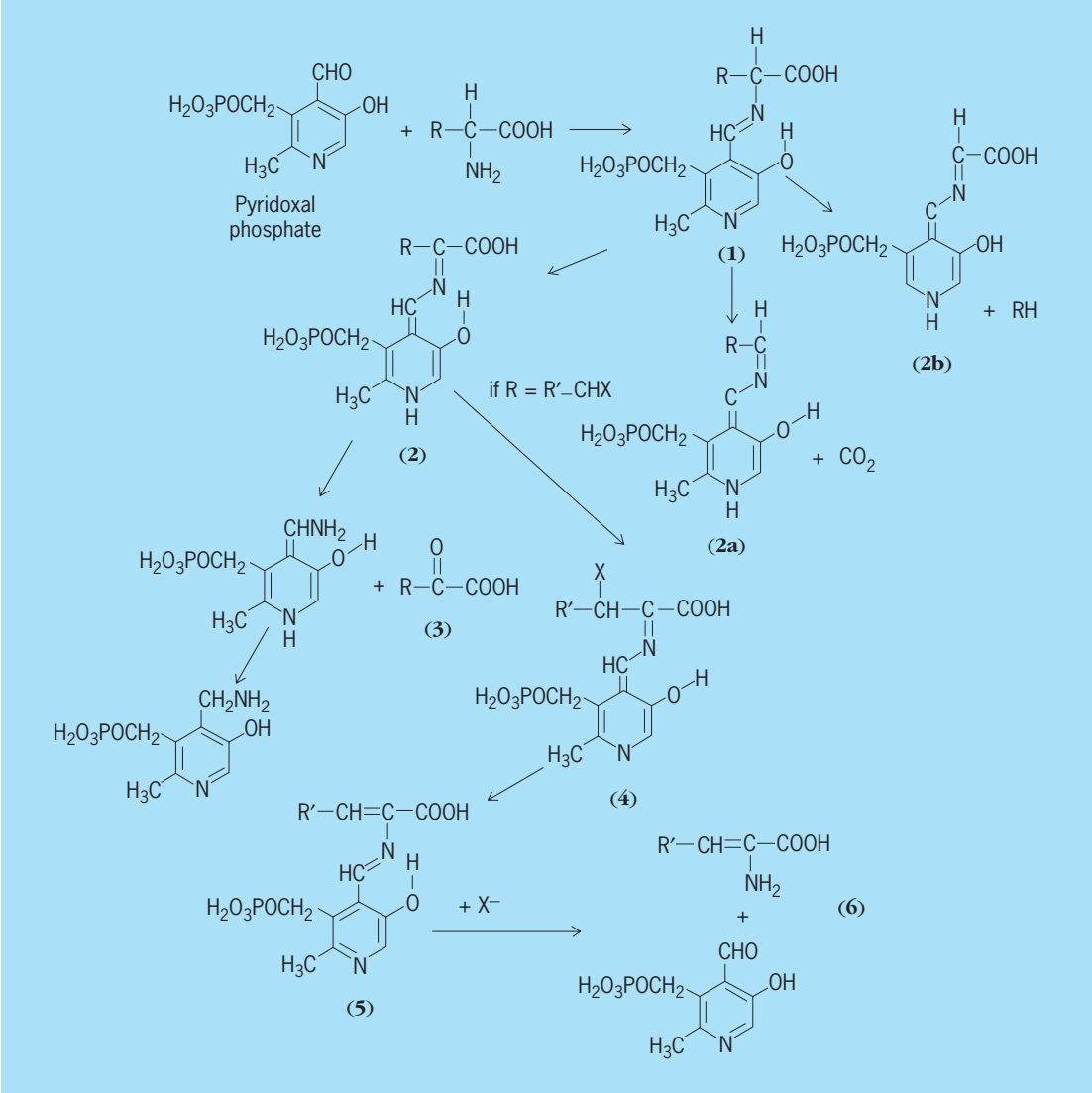

Fig. 7. General features of degradative pathways for amino acids.

aspartate or glutamate as one partner, with alanine also being rather common.

Decarboxylation. Decarboxylations are carried out by rather specific enzymes, and their products are often biologically active substances. Histamine, tryptamine (serotonin), and dopamine are all examples of decarboxylation products which are very active in animal tissues.

β Elimination. These reactions result in the loss of an electronegative group, such as OH^- or SH^-, from the carbon adjacent to the α carbon. Usually the net result is a deamination as well, since α,β unsaturated amino acids rapidly tautomerize to α,β saturated imino acids, which hydrolyze to yield the corresponding keto acid. Sometimes, however, there is addition to the α,β double bond of the pyridoxal phosphate–amino acid complex, so that the net result is a β substitution. This type of reaction is important in the synthesis of tryptophan.

β Cleavage. Instead of the α hydrogen or the carboxyl group, the β carbon can be eliminated, giving rise to glycine. These reactions only occur when the β substituent is electronegative, as in serine or threonine.

Conjugation. Until 1956 the mechanism of protein synthesis was totally unknown. Since then most of the details have been elucidated. A brief outline of the steps involved follows. The brevity of the outline precludes mention of many of the details.

The structure of a protein is determined by the arrangement of bases in the portion of the deoxyribonucleic acid (DNA) making up the gene for that protein. This sequence of DNA is transcribed into a complementary molecule of ribonucleic acid (RNA). The specific RNA, called messenger RNA (mRNA), binds to ribosomes, which are complex subcellular particles composed of a different sort of RNA, ribosomal RNA (rRNA), and protein. The ribosomes serve as the sites where conjugation of the amino acids takes place.

The amino acids are activated by reaction with a molecule of ATP to form an amino-acyl adenylate (**Fig. 8**) that remains temporarily bound to the enzyme which catalyzes the activation, an

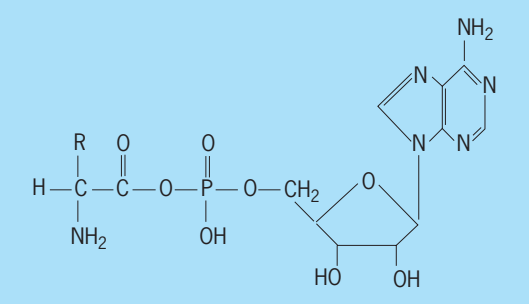

Fig. 8. Structural formula for amino-acyl adenylate.

amino-acyl-tRNA synthetase (tRNA indicates transfer RNA). The amino-acyl moiety is then transferred to a specific molecule of RNA, that is, to tRNA. The linkage of the amino acid to the tRNA takes place by esterification to the 2′-hydroxyl of the ribose moiety of the 3′-terminal end. The tRNA molecule has a complex secondary structure which exposes a sequence of three bases in the interior of the molecule. These three bases are called the anticodon. Their nature and order are specific for the amino acid involved. The amino-acyl-tRNA synthetase contributes the specificity which ensures, with very high certainty, that the amino acid is attached to a tRNA with the proper anticodon.

The amino-acyl-tRNA interacts with the ribosome messenger complex when there is a sequence of three bases (the codon) on the messenger complementary, in the Watson-Crick sense, to the anticodon. The tRNA is aligned on the ribosome in such a way that the nascent protein chain, held by a C-terminal amino acid still linked to its tRNA, is brought into proximity with the free α-amino group of the amino acid. A series of enzymes effects formation of the peptide bond, and the nascent protein chain has been elongated by one amino acid. The chain is now held to the ribosome by the tRNA of the latest residue added, and it is ready to accept the amino acid specified by the next codon. Thus the ribosome progresses along the mRNA, "reading" it, until all the amino acids specified by the original DNA gene have been added. It then reaches a codon which signifies termination of the chain, and a soluble protein is released.

The nascent protein chain begins in bacteria with a special methionyl-tRNA complex, in which the amino group of methionine is blocked with a formyl group. This is the first amino-acyl-tRNA to bind to the ribosome, and apparently all genes begin with the codon specific for this tRNA. In the first peptide bond formed, the *N*-formyl-methionyl-tRNA plays the part of the nascent protein chain.

For smaller peptides, such as glutathione and the mucopeptide of bacterial cell walls, the conjugation process is catalyzed by specific enzymes and involves different intermediates. Necessarily this means separate enzymes for each such peptide, in contrast to protein synthesis where the same machinery (except for the mRNA) serves for all chains of whatever sequence.

Amino Acids in Nutrition

The nutritional requirement for the amino acids of protein can vary from zero, in the case of an organism which synthesizes them all, to the complete list, in the case of an organism in which all the biosynthetic pathways are blocked. There are 8 or 10 amino acids required by certain mammals; most plants synthesize all of their amino acids, while microorganisms vary from types which synthesize all, to others (such as certain lactic acid bacteria) which require as many as 18 different amino acids. *See* NUTRITION; PROTEIN METABOLISM.

It seems likely that, when life originated, amino acids were taken from the rich organic medium which the oceans then offered, and biosynthetic abilities evolved only slowly as the supply of exogenous materials became depleted. A stage must have eventually been reached, however, at which all the amino acids were being synthesized metabolically, and none was required nutritionally. As evolution progressed, food chains developed, and some forms of life became adapted to obtain many of their organic nutrients at the expense of other living forms, either directly or indirectly. In these dependent types, mutations had occurred, causing the loss of specific biosynthetic enzymes and hence the gain of nutritional requirements. It is easy to duplicate this process in the laboratory: A microorganism with full biosynthetic ability can be induced to undergo random mutations, and selective methods can then be used to isolate mutants requiring amino acids, vitamins, or other normal metabolites. In every case, it is found that a given mutation deprives the cell of a single biosynthetic enzyme, blocking the reaction which that enzyme catalyzes and thus the entire pathway of which that reaction is a part. *See* PREBIOTIC ORGANIC SYNTHESIS.

In summary, the nutrition of many organisms must include the provision of growth factors, which are defined as organic compounds which the organism requires for its growth but which it cannot synthesize. Growth factor requirements reflect the heritable loss of biosynthetic enzymes, as the result of gene mutations. Amino acids are typical growth factors for many organisms.

Graphic Presentation of Amino Acid Biosynthesis

The amino acids are grouped into families on the basis of their common biosynthetic origins (Figs. 9–13). Lysine is shown in two families, because its biosynthesis in bacteria differs from that in fungi. Intermediates which are hypothetical are shown in brackets. The notation −2H or +2H refers to the removal or addition of two electrons and two hydrogen ions with the aid of either diphosphopyridine nucleotide (DPN) or triphosphopyridine nucleotide (TPN), both of which are coenzymes of hydrogen transfer. Symbols used are ~Ac, coenzyme A–bound acetate; PRRP, phosphoribosyl pyrophosphate; CAP, carbamyl phosphate; and ATP, adenosine triphosphate. An arrow between two compounds in the diagrams does not necessarily imply a single

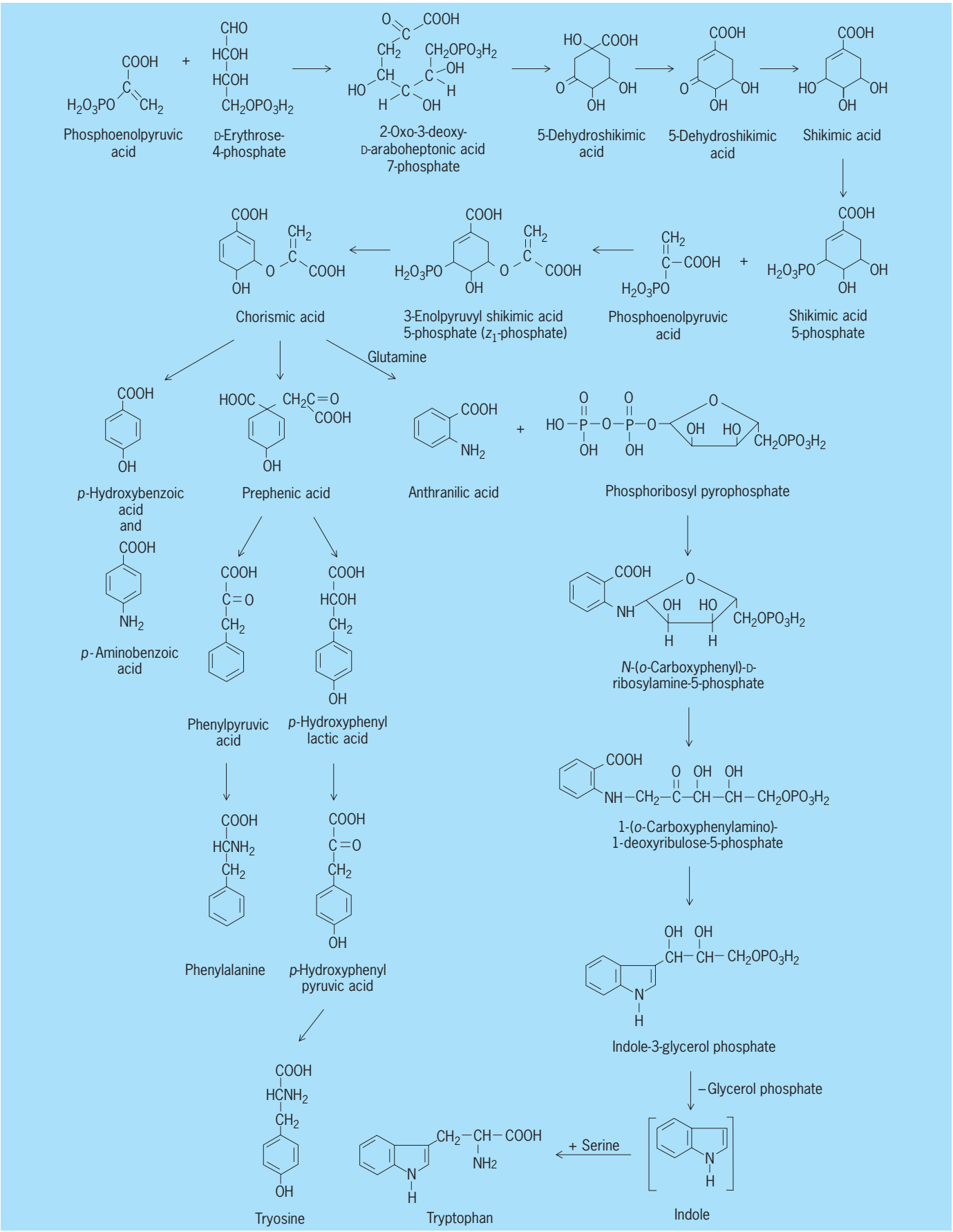

Fig. 9. Metabolic pathways for the aromatic family of amino acids.

Fig. 10. Metabolic pathways for the α-ketoglutaric family of amino acids.

Fig. 11. Metabolic pathways for the aspartic acid family of amino acids.

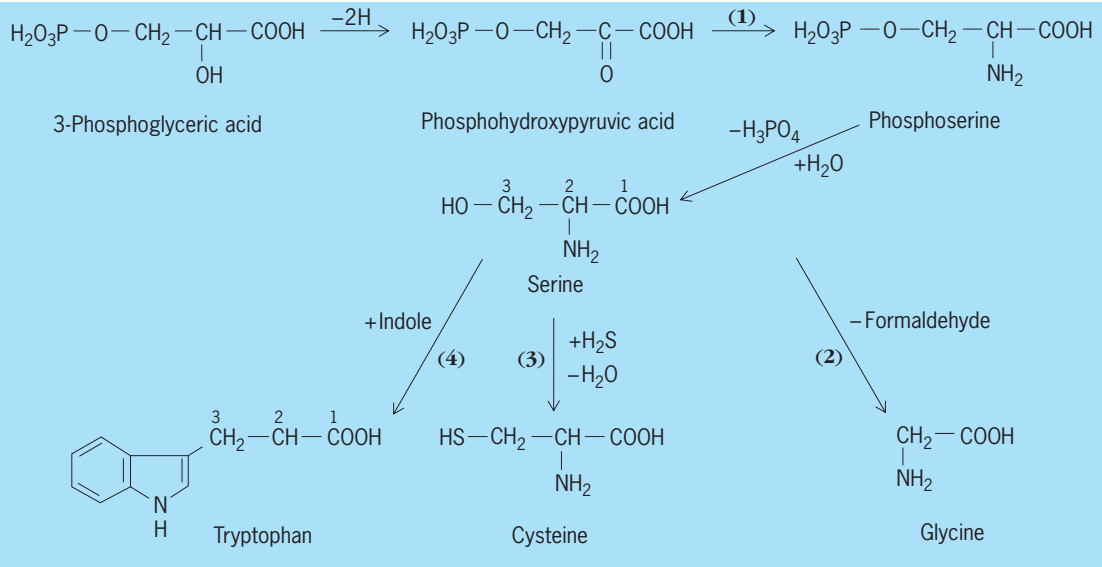

Fig. 12. Metabolic pathways for the serine family of amino acids.

enzymatic reaction. In many cases, the arrow represents a sequence of reactions for which the intermediates are unknown.

Edward A. Adelberg; Paul T. Magee

Aromatic family. This family is composed of phenylalanine, tyrosine, tryptophan, and two other important metabolites, *p*-aminobenzoic acid and *p*-hydroxybenzoic acid. The initial precursors for the biosynthesis of these amino acids, phosphoenol pyruvate and D-erythrose 4-phosphate, are metabolites of glucose catabolism.

Although the intermediates shown in **Fig. 9** have all been isolated and identified, not all of the enzymes (particularly those involved in chorismic acid metabolism) have been studied. It is probable that *p*-hydroxybenzoic acid and prephenic acid are synthesized directly from chorismic acid by a single enzymatic reaction, and *p*-aminobenzoic

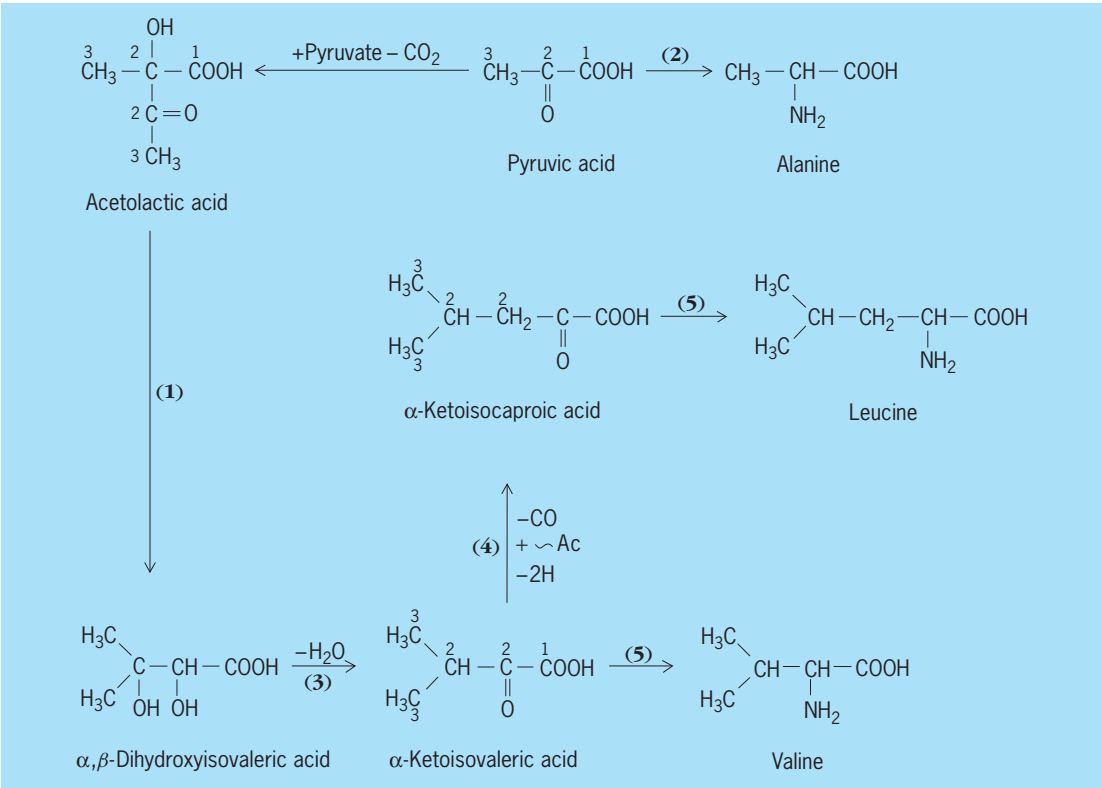

Fig. 13. Metabolic pathways for the pyruvic acid family of amino acids.

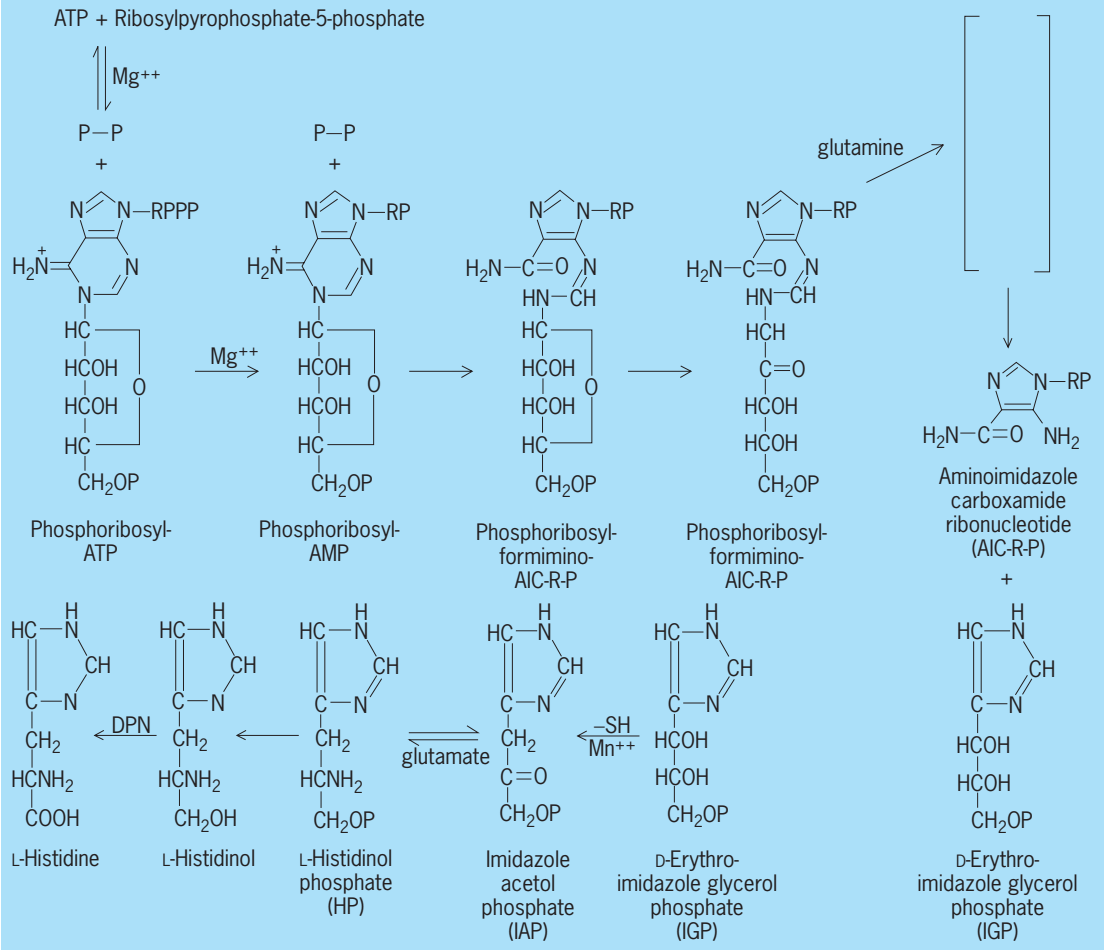

Fig. 14. Metabolic pathways in histidine biosynthesis.

acid may also be a direct metabolite. It is certain that other unidentified intermediates exist in the pathway between chorismic acid and anthranilic acid.

Two different enzymes for the conversion of chorismic acid to prephenic acid have been demonstrated in microorganisms. One of these enzymes is controlled by the pool size of tyrosine, while the other is controlled by phenylalanine.

The glycerol phosphate side chain of indoleglycerol phosphate derived from phosphoribosyl pyrophosphate can be exchanged directly for serine without the formation of free indole as an intermediate. In the absence of serine, the enzyme liberates free indole from indoleglycerol phosphate, and the same enzyme will condense indole with serine to form tryptophan.

There is some evidence for the existence of an anthranilic acid–tryptophan cycle in microorganisms. Formylkynurenine produced by the action of tryptophan pyrrolase on tryptophan can regenerate anthranilic acid by the combined action of kynureninase and kynureinine formamidase. R. G. Martin

α-Ketoglutaric acid family. This family is composed of glutamic acid, proline, lysine, and arginine. The numbered items refer to **Fig. 10**.

1. In yeast and other fungi, lysine is formed from α-ketoglutaric acid plus a C_2 fragment derivable from acetate. Lysine is formed by a different pathway in bacteria (see following section on aspartic acid).

2. Presumably by transamination.

3. This reductive amination is the main source of organic nitrogen for most microorganisms.

4. The cyclization takes place spontaneously.

5. The acetylation of glutamic acid prevents cyclization at the next step and permits the eventual formation of ornithine. This mechanism has been demonstrated in *Escherichia coli*, but does not take place in the fungus *Neurospora*; the fungus appears able to form ornithine via the nonacetylated intermediates.

6. Transamination.

7. Carbamyl phosphate (CAP).

Aspartic acid family. This family is composed of aspartic acid, lysine, threonine, methionine, and isoleucine. The numbered items refer to **Fig. 11**.

1. Aspartate arises principally by the transamination of oxaloacetate. In plants and in some microorganisms, it is formed by the direct amination of fumaric acid.

2. The compound formed by the transmination at carbon 6 and the subsequent desuccinylation is

L,L-diaminopimelic acid. For the decarboxylase to function, the compound must be racemized to form *meso*-diaminopimelic acid.

3. In bacteria, and presumably in blue-green algae, lysine is formed by decarboxylation of diaminopimelic acid. In fungi and in higher animals, lysine is formed by a different route as seen in the α-ketoglutaric acid family (Fig. 10).

4. A series of reactions probably involves transfer of an active formaldehyde group from serine, followed by reduction.

5. Intramolecular rearrangement and reduction in α-aceto-α-hydroxybutyric acid take place in one step. The same enzyme catalyzes the analogous step in valine biosynthesis (Fig. 13).

6. Transamination from glutamic acid. The same transaminase functions for the keto acids of both isoleucine and valine.

Serine family. The serine family is composed of serine, glycine, cysteine, and tryptophan. The numbered items refer to **Fig. 12**.

1. Transamination from alanine. There is equal evidence for a second pathway in which dephosphorylation precedes transamination.

2. The terminal group of serine is transferred to tetrahydrofolic acid (THFA) to form *N*(10)-hydroxymethyl-THFA. In this form, it can be transferred at various levels of oxidation for biosynthesis of compounds methionine, purine, and thymine.

3. This reaction, inferred to occur in microorganisms, is yet to be directly demonstrated. In animal tissues, serine receives the sulfhydryl group by transsulfuration from homocysteine, which is formed in animal tissues from dietary methionine.

4. See the aromatic family for considerations of this reaction.

Pyruvic acid family. This family is composed of valine, leucine, and alanine. The numbered items refer to **Fig. 13**.

1. Intramolecular rearrangement and reduction take place in one step.

2. Generally by transamination from glutamate, although cases of direct reductive amination with ammonia have been cited.

3. A single enzyme, dihydroxy acid dehydrase, catalyzes the dehydration of both the isoleucine and valine dihydroxy acid precursors.

4. The complex series of reactions involved in the formation of α-ketoisocaproic acid is exactly analogous to the steps in the citric acid cycle leading to formation of α-ketoglutarate, with the keto acid taking the place of oxaloacetate.

5. Transamination from glutamate. The valine transaminase also functions in isoleucine biosynthesis.

Edward A. Adelberg; Paul T. Magee

Histidine biosynthesis. The pathway of histidine biosynthesis shown in **Fig. 14** is known to occur in the mold *Neurospora* and in coliform bacteria. It should be noted that the imidazole ring of histidine is formed by the pathway and is not derived from the five-membered ring of adenosine triphosphate (ATP).

David W. E. Smith

Bibliography. G. C. Barrett et al. (eds.), *Chemistry and Biochemistry of the Amino Acids*, 1983; H. D. Jakubke and H. Jeshkeit, *Amino Acids, Peptides and Proteins*, 1978; A. Meister, *Biochemistry of the Amino Acids*, 2 vols., 2d ed., 1965.

Amino sugar

A sugar in which one or more nonglycosidic hydroxyl groups is replaced by an amino or substituted amino group. The most abundant example is D-glucosamine (2-amino-2-deoxy-D-glucose), whose structure is shown below in the α-pyranose ring form.

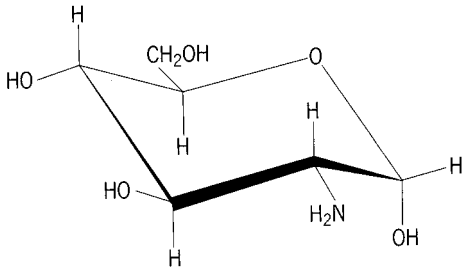

A linear polymer of *N*-acetyl-D-glucosamine is widely distributed as chitin, the exoskeletal material of arthropods. The glycoproteins of higher animals, which are components of the proteoglycans of cartilage and skin, consist of polysaccharides that are generally sulfated and have *N*-acetylated glucosamine or galactosamine alternating with a uronic acid.

Amino sugars are important constituents of glycoproteins and oligosaccharides involved in biological recognition. The 9-carbon amino sugar *N*-acetylneuraminic acid occurs widely. Biosynthetic incorporation of amino sugars commonly involves the nucleotide-sugar pathway. Amino sugars of the greatest structural diversity are found in microorganisms as constituents of cell walls, in antigenic carbohydrates produced at the cell surface, and as antibiotic substances secreted from the cell. Streptomycin is the first demonstrated example of numerous amino-sugar-containing antibiotics produced notably by Actinomycetes (bacteria). *See* GLYCOPROTEIN; OLIGOSACCHARIDE; POLYSACCHARIDE.

Derek Horton

para-Aminobenzoic acid

A compound also known as PABA, often considered to be a water-soluble vitamin, with the structure below. *p*-Aminobenzoic acid is widely distributed in

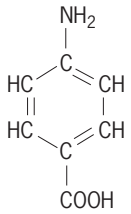

foods and has been isolated from liver, yeast, and other sources rich in vitamin B. There is doubt, however, that it is significant as a nutrient. In the early 1940s a considerable literature concerning the role of *p*-aminobenzoic acid in curing deficiency disease appeared, but it is now known that this effect was due in great part to other vitamins, particularly folic acid, about which little was then known. *p*-Aminobenzoic acid is a part of the folic acid molecule, and its presence in a folic acid–deficient diet results in increased intestinal synthesis of the folic acid.

p-Aminobenzoic acid antagonizes the bacteriostatic action of sulfonamides. Because of the similar chemical compositions of these substances, it is probable that the sulfonamides function at least in part by displacing *p*-aminobenzoic acid in bacterial enzyme systems. *p*-Aminobenzoic acid is an effective antirickettsial agent and has been used to treat typhus, scrub typhus, and spotted fever. There is no evidence that humans have a dietary requirement for this vitamin. The compound is also used as a sun-screen agent in suntan lotions. *See* FOLIC ACID; RICKETTSIOSES; VITAMIN. Stanley N. Gershoff

Ammeter

An electrical instrument for measuring electric current. Currents are usually either unidirectional and steady (direct current or dc) or alternating in direction at a relatively low frequency (alternating current or ac). A current that is unidirectional but regularly fluctuating is a superposition of dc and ac. Higher-frequency ac is often referred to as radio-frequency or RF current. At frequencies above about 10 MHz, where the wavelength of the signal becomes comparable with the dimensions of the measuring instrument, current measurements become inaccurate and finally meaningless, since the value obtained depends on the position where the measurement is made. In these circumstances, power measurements are usually used. *See* CURRENT MEASUREMENT; MICROWAVE MEASUREMENTS.

Magnetic fields, thermal (heating) effects, and chemical effects resulting from the passage of electric current may be employed to create measuring instruments. Current may also be measured in terms of the voltage that appears across a resistive shunt through which the current passes. This method has become the most common basis for ammeters, primarily because of the very wide range of current measurement that it makes possible, and more recently through its compatibility with digital techniques. *See* MULTIMETER; VOLTMETER.

The unit of current, the ampere, is the base unit on which rest the International System (SI) definitions of all the electrical units. The ampere itself is defined in terms of the mechanical forces that are produced by the magnetic interaction of two currents. Primary instruments in national standards laboratories measure currents in accordance with this definition and are at the head of a hierarchy of everyday calibrated current-measuring instruments. *See* CURRENT BALANCE; ELECTRICAL UNITS AND STANDARDS; WATT BALANCE.

Direct-current ammeters. The moving-coil, permanent-magnet (d'Arsonval) ammeter remains an important type. Their mechanical construction is identical with that of a d'Arsonval moving-coil instrument. *See* VOLTMETER.

These ammeters continue to be in large-scale production and are available in many versions, from small indicators to large instruments suitable for control panels. Generally they are of modest accuracy, no better than 1%. Digital instruments have taken over all measurements of greater precision because of the greater ease of reading their indications where high resolution is required.

By suitable choice of coil winding and strength of magnet, moving-coil meters of this basic design can give full deflection at currents in the range from 10 microamperes to about 10 milliamperes. Increased sensitivity can be obtained by the use of electronic amplifiers. Ammeters intended for higher currents use similar instruments, modified to require some 50 mV for full-scale deflection, in conjunction with shunts. *See* AMPLIFIER.

Low-frequency ac ammeters. Moving-iron instruments are widely used as ammeters for low-frequency ac applications and their mechanical construction is identical with a moving-iron repulsion voltmeter. *See* VOLTMETER.

Although the design will also operate with direct current, hysteresis in the magnetic material limits accuracy. Eddy currents limit the usefulness to low frequencies.

High-frequency ammeters. High-frequency currents are measured by the heating effect of the current passing through a physically small resistance element. One early design, the hot-wire ammeter, used the increase in length of a wire that was heated by the current. The increase was mechanically amplified in order to move the pointer. In modern instruments the temperature of the center of the wire is sensed by a thermocouple, the output of which is used to drive a moving-coil indicator. A wide range of full-scale sensitivities can be provided, from about 50 A with a heavy-gage heater in air to 1 mA by using a fine winding on an insulating bead around the thermocouple, working in a small evacuated glass bulb. *See* THERMOCOUPLE.

Digital ammeters. Although most forms of digital ammeter operate by balancing currents derived from the signal and a reference source, analog-to-digital converters are restricted to full-range currents in the region of 1 mA. Higher values require inconveniently large circuit currents; lower currents bring problems of accuracy because of leakage. Digital ammeters are therefore almost invariably based on a conventional digital voltmeter which measures the voltage drop caused by the current flowing through a shunt resistor. *See* VOLTMETER.

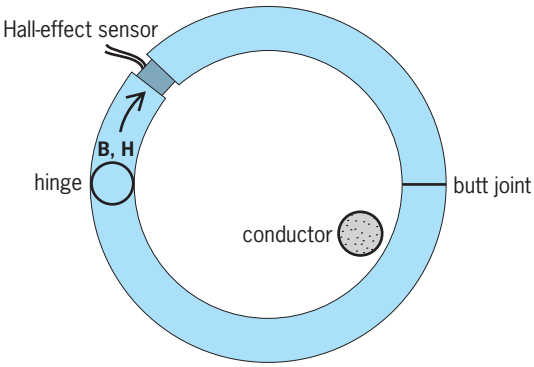

Magnetic circuit of a clip-on ammeter or current probe. B = magnetic flux density, H = magnetic field strength.

Alternating-current measurement. Alternating currents may be measured by the addition of an ac converter. This converter transforms the ac signal at the output of the scaling amplifier to an equivalent dc signal for processing by the analog-to-digital conversion circuits. Shunts and current transformers can be used for ranges up to tens of thousands of amperes; electronic amplifiers and electrometers can extend the current measurement capability to the femtoampere (10^{-15} A) region. More refined analog-to-digital converters can be used to provide improved resolution and accuracy. *See* ANALOG-TO-DIGITAL CONVERTER; TRANSFORMER.

Clip-on ammeters (clamp meters). These hand-held devices possess the exceedingly useful attribute of being able to measure a current in a particular conductor without interrupting the circuit to insert, for example, a current-measuring resistor. They accomplish this by clamping a hinged high-permeability magnetic core around the conductor, which thus becomes a single-turn primary winding of a current transformer. If alternating current is to be measured, a secondary winding around the core which is connected to an ammeter suffices to complete the instrument. If direct current is to be measured, a Hall-effect flux sensor is inserted transversely in a narrow gap in the core to measure the magnetic flux (see **illustration**). The flux density in the magnetic circuit resides perdominantly in the gap whose dominant and constant reluctance ensures that this flux, and therefore the voltage output of the Hall sensor, is proportional to the current. This direct-current instrument is best calibrated with known currents. *See* HALL EFFECT.

Both the alternating-current and direct-current instruments possess the convenient property that the measurement is, in principle, independent of the position of the current-carrying conductor within the core. This is because of Ampère's theorem that the line integral of flux around a closed magnetic path is independent of the position within the magnetic path of the current which is the source of the flux.

The accuracy of these instruments is usually only a few percent, but this is adequate for many applications. Connecting the secondary winding or Hall-sensor output to an oscilloscope produces the particularly useful fault-tracing tool called a current probe. It is often the current in one particular conductor of a network rather than the voltage differences between junction points that is important.

Sampling techniques. A different digital solution to the problem of measuring alternating or varying currents is the use of sampling. The signal, after suitable scaling and conversion to a voltage, is sampled repetitively. These samples are digitized, using a fast analog-to-digital converter. The resulting stream of numerical values is then processed mathematically to yield the information required, such as the root-mean-square (rms) value, mean value, and dc component. In order to accurately reflect the properties of the input waveform, the sampling rate needs to be at least twice the highest frequency present in the signal. *See* ELECTRICAL MEASUREMENTS; INFORMATION THEORY; ROOT-MEAN-SQUARE.

R. B. D. Knight; Bryan P. Kibble

Bibliography. H. M. Berlin and F. C. Getz, Jr., *Principles of Electronic Instrumentation and Measurement,* 1988; H. H. Chiang, *Electrical and Electronic Instrumentation,* 1984; S. Geczy, *Basic Electrical Measurements,* 1984; A. D. Helfrick and W. D. Cooper, *Modern Electronic Instrumentation and Measurement Techniques,* 1990; L. Jones, *Electronic Instruments and Measurements,* 1991; B. Kibble et al., *A Guide To Measuring Direct and Alternating Current and Voltage Below 1 MHz*, Institute of Measurement and Control, 2003; M. V. Reissland, *Electrical Measurements: Fundamentals, Concepts, Applications,* 1990.

Ammine

One of a group of complex compounds formed by the coordination of ammonia molecules with metal ions and, in a few instances, such as calcium, strontium, and barium, with metal atoms. Some typical examples of ammines include $[Co(NH_3)_6]Cl_2$ (rose), $[Cu(NH_3)_4]Cl_2$ (blue), $[Cr(NH_3)_6]Cl_3$ (yellow), $[Cr(NH_3)_4Cl_2]Cl$ (cis is violet, trans is green), $[Ni(NH_3)_6]Cl_2$ (blue), $[Pt(NH_3)_4]Cl_2{\cdot}H_2O$ (colorless), and $[Hg(NH_3)_2]Br_2$ (colorless). Although these ammines are formally analogous to many salt hydrates, the general characteristics of the group of ammines differ considerably from those of the hydrates. For example, hydrated Co(III) salts are strong oxidizing agents whereas Co(II) ammines are strong reducing agents. The ammines of principal interest are those of the transition metals and of the zinc family, but even here there is wide variation in stability or rate of decomposition. For example, iron ammines are unstable in aqueous solution; Cu(II) and Co(II) ammines exist in aqueous solution but are decomposed by aqueous acids; Co(III) and Pt(IV) ammines can be recrystallized from strong acids. Ammines are prepared by treating aqueous solutions of the metal salt with ammonia or, in some instances, by the action of dry gaseous or liquid ammonia on the anhydrous salt. These, and similar, differences have motivated much

theoretical study of the bonding in such compounds. *See* AMMONIA; CHEMICAL BONDING; COORDINATION CHEMISTRY.

Harry H. Sisler

Ammonia

The most familiar compound composed of the elements nitrogen and hydrogen, NH_3. It is formed as a result of the decomposition of most nitrogenous organic material, and its presence is indicated by its pungent and irritating order.

Uses. Because of the wide range of industrial and agricultural applications, ammonia is produced in tremendous quantities. Examples of its use are the production of nitric acid and ammonium salts, particularly the sulfate, nitrate, carbonate, and chloride, and the synthesis of hundreds of organic compounds including many drugs, plastics, and dyes. Its dilute aqueous solution finds use as a household cleansing agent. Anhydrous ammonia and ammonium salts are used as fertilizers, and anhydrous ammonia also serves as a refrigerant, because of its high heat of vaporization and relative ease of liquefaction. *See* FERTILIZER.

Molecular structure. The NH_3 molecule has a pyramidal structure of the type shown in the diagram below, in which the nitrogen atom has achieved a

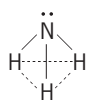

stable electronic configuration by forming three electron pair bonds with the three hydrogen atoms. The HNH bond angle in the pyramid is 106.75°, which is best explained as resulting from the use of sp^3 hybrid bonding orbitals by the nitrogen atom. This should yield tetrahedral bond angles with the result that one sp^3 orbital is occupied by the unshared pair of electrons. The repulsive effect of this unshared pair, which is concentrated relatively near to the nitrogen nucleus on the shared pairs of electrons forming the N-H bonds, produces a slight compression of the HNH bond angles, thus accounting for the fact that they are slightly less than tetrahedral (109.5°). The dipole moment of the ammonia molecule, 1.5 debyes, is a resultant of the combined polarities of the three N-H bonds and of the unshared electron pair in the highly directional sp^3 orbital. The pyramidal ammonia molecule turns inside out readily, and it oscillates between the two extreme positions at the precisely determined frequency of 2.387013×10^{10} Hz. This property has been used in the highly accurate time-measuring device known as the ammonia clock.

Physical characteristics. The physical properties of ammonia are analogous to those of water and hydrogen fluoride in that the physical constants are abnormal with respect to those of the binary hydrogen compounds of the other members of the respective periodic families. This is particularly true of the boiling point, freezing point, heat of fusion, heat of vaporization, and dielectric constant. These abnormalities may be related to the association of molecules through intermolecular hydrogen bonding. The principal physical constants for ammonia are summarized in **Table 1**. Ammonia is highly mobile in the liquid state and has a high thermal coefficient of expansion.

TABLE 1. Physical properties of ammonia

Property	Value
Melting point	−77.74°C (−107.9°F)
Boiling point	−33.35°C (−28.03°F)
$\Delta H_{fus.}$ at mp	5657 J mole^{-1}
$\Delta H_{vap.}$ at bp	23,350 J mole^{-1}
Critical temperature	132.4°C (270.3°F)
Critical pressure	112.5 atm (11.40 MPa)
Dielectric constant (−77.7°C)	25
Density (−70°C)	0.7253 g cm^{-3}
Density (−30°C)	0.6777 g cm^{-3}
$\Delta H^\circ_{form.}$ (25°C)	46.19 kJ mole^{-1}
$\Delta G^\circ_{form.}$ (25°C)	16.64 kJ mole^{-1}
S° (298.1 K) (exptl.)	192.5 J deg^{-1} mole^{-1}
C°_p	35.66 J deg^{-1} mole^{-1}
Viscosity (liquid, 25°C)	0.01350 Pa · s
Vapor pressure (−20°C)	1426.8 torr (190.22 kPa)
Vapor pressure (0°C)	3221.0 torr (429.43 kPa)
Vapor pressure (20°C)	6428.5 torr (857.06 kPa)
Solubility in H_2O	
(1 atm,* 0°C)	42.8% by wt
(1 atm, 20°C)	33.1% by wt
(1 atm, 40°C)	23.4% by wt

*1 atm = 101.325 kilopascals.

Chemical properties. Most of the chemical reactions of ammonia may be classified under three chief groups: (1) addition reactions, commonly referred to as ammonation; (2) substitution reactions, commonly referred to as ammonolysis; and (3) oxidation-reduction reactions.

Ammonation. Ammonation reactions include those in which ammonia molecules add to other molecules or ions either through the mechanism of covalent bond formation using the unshared pair of electrons on the nitrogen atom, through ion-dipole electrostatic interactions, or through hydrogen bonding. Most familiar of the ammonation reactions is the reaction with water, which may be represented schematically as reaction (1). The strong tendency of water

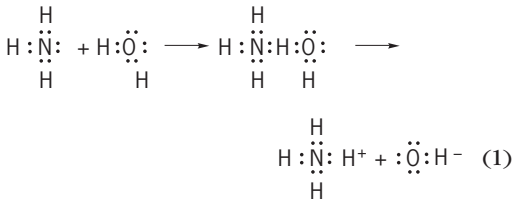

(1)

and ammonia to combine is evidenced by the very high solubility of ammonia in water (700 vol of ammonia gas in 1 vol of water at 20°C and 1 atm ammonia pressure). The ammonia hydrate (ammonium hydroxide) is a weak electrolyte in aqueous solution as is indicated by an ionization constant of 1.77×10^{-5} at 25°C. Phase diagrams for the NH_3-H_2O system indicate the existence of $NH_3 \cdot H_2O$ (mp −79.0°C

or −110.0°F) in the solid state at low temperatures. Under these conditions, the compound $2NH_3{\cdot}H_2O$ (mp −78.8°C or −109.8°F) also exists. Ammonia reacts readily with strong acids to form ammonium salts, reaction (2). Ammonium salts of weak acids in the solid state dissociate readily into ammonia and the free acid.

Included among ammonation reactions is the formation of complexes (called ammines) with many

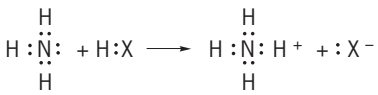

(2)

($:X^-$ = an anion)

metal ions, particularly transition-metal ions, for example, $Hg(NH_3)_2^+$, $Cr(NH_3)_6^{3+}$, $Zn(NH_3)_4^{2+}$, and $Co(NH_3)_6^{3+}$. *See* AMMINE.

Ammonation occurs with a variety of molecules capable of acting as electron acceptors (Lewis acids). The reaction of ammonia with such substances as sulfur trioxide, sulfur dioxide, silicon tetrafluoride, and boron trifluoride are typical and illustrated by reaction (3).

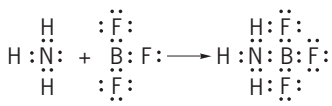

(3)

Ammonolysis. Ammonolytic reactions include reactions of ammonia in which an amide group (—NH_2), an imide group (═NH), or a nitride group (≡N) replaces one or more atoms or groups in the reacting molecule. Examples include reactions (4) to (8).

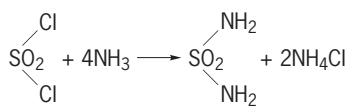

(4)

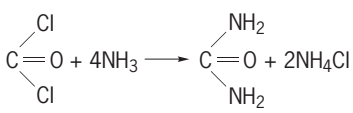

(5)

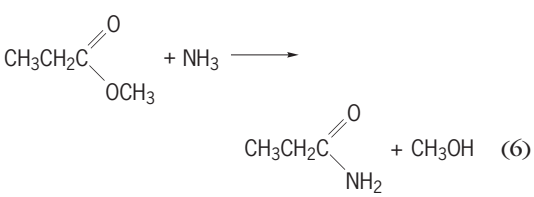

(6)

$$HgCl_2 + 2NH_3 \longrightarrow Hg(NH_2)Cl + NH_4Cl \qquad (7)$$

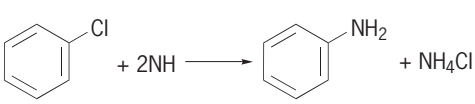

(8)

Oxidation-reduction. These reactions may be subdivided into those which involve a change in the oxidation state of the nitrogen atom and those in which elemental hydrogen is liberated. An example of the first group is the catalytic oxidation of ammonia in air to form nitric oxide, shown in reaction (9). In the

$$4NH_3 + 5O_2 \xrightarrow{Pt} 4NO + 6H_2O \qquad (9)$$

absence of a catalyst, ammonia burns in oxygen to yield nitrogen. This is shown by reaction (10). Another example, reaction (11) is the reduction with ammonia of hot metal oxides such as cupric oxide. Still another example is the reaction of ammonia with chlorine (12).

$$4NH_3 + 3O_2 \longrightarrow N_2 + 6H_2O \qquad (10)$$

$$3CuO + 2NH_3 \longrightarrow 3Cu + 3H_2O + N_2 \qquad (11)$$

$$2NH_3 + Cl_2 \longrightarrow NH_2Cl + NH_4Cl \qquad (12)$$

Oxidation-reduction reactions of ammonia of the second type are exemplified by reactions of active metals with ammonia, shown in reactions (13) and (14).

$$2Na + 2NH_3 \rightarrow 2NaNH_2 + H_2 \qquad (13)$$

$$3Mg + 2NH_3 \xrightarrow{\Delta} Mg_3N_2 + 3H_2 \qquad (14)$$

Liquid ammonia as a solvent. The physical and chemical properties of liquid ammonia make it appropriate for use as a solvent in certain types of chemical reactions. The solvent properties of liquid ammonia are, in many ways, qualitatively intermediate between those of water and of ethyl alcohol. This is particularly true with respect to dielectric constant; therefore, ammonia is generally superior to ethyl alcohol as a solvent for ionic substances but is inferior to water in this respect. On the other hand, ammonia is generally a better solvent for covalent substances than is water. The chemical properties of ammonia, for example, its ability to undergo ammonation, ammonolysis, and oxidation-reduction reactions, are roughly analogous to reactions of water (hydration, hydrolysis, and oxidation-reduction). Both water and liquid ammonia undergo autoionization; liquid ammonia, reaction (15), undergoes the process to a lesser extent, as its very low ion product constant indicates, reaction (16).

$$2NH_3 \rightleftharpoons NH_4^+ + NH_2^- \qquad (15)$$

$$[NH_4^+][NH_2^-] = 1.9 \times 10^{-33} \text{ at } -50°C \qquad (16)$$

Among the particularly interesting aspects of chemistry in liquid ammonia is the fact that metals of sufficiently low lattice energy, high cation solvation energy, and low ionization energy are reversibly soluble in liquid ammonia. The solutions obtained are highly colored—blue if dilute, bronze if concentrated. Metals which form such solutions include the alkali metals, the heavier alkaline earth metals, and the divalent lanthanides. Concentrations as high as 10 to 20 molal are obtainable for the alkali metals. Dilute solutions exhibit the phenomenon of strong electrolytic conduction, whereas the concentrated solutions act as metallic conductors. At very high

TABLE 2. Analogous compounds in the water and ammonia systems

Aquo	Ammono
$(H_3O)Cl$	$(NH_4)Cl$
$C(=O)(OH)_2$	$C(=NH)(NH_2)_2$
KOH	KNH_2
$Na_2[Zn(OH)_4]$	$Na_2[Zn(NH_2)_4]$
$Cu(H_2O)_4^{2+}$	$Cu(NH_3)_4^{2+}$
MgO	$MgNH$, Mg_3N_2
$CH_3C(=O)OH$	$CH_3C(=NH)NH_2$
C_2H_5OH	$C_2H_5NH_2$
$(CH_3)_2O$	$(CH_3)_2NH$, $(CH_3)_3N$
$Hg(OH)Cl$	$HgNH_2Cl$
$HOCl$	H_2NCl

dilutions magnetic measurements indicate the presence of 1 mole of unpaired electrons per mole of dissolved metal, but as the concentrations of the solutions are increased, their paramagnetism decreases. Much remains to be learned concerning the nature of these solutions, but presently available data may be interpreted in terms of the five solute species M, $M_{2(am)}$, $M^-_{(am)}$, $M^+_{(am)}$, and $e^-_{(am)}$ participating in equilibria (17)–(19). The ammoniated electron, $e_{(am)}$,

$$M \rightleftharpoons M^+_{(am)} + e^-_{(am)} \tag{17}$$

$$2M \rightleftharpoons M_{2(am)} \tag{18}$$

$$e^-_{(am)} + M \rightleftharpoons M^-_{(am)} \tag{19}$$

appears to consist of an electron trapped in a cavity (approximately 0.3 nanometer in diameter) of NH_3 molecules. These solutions are unstable but, in the absence of catalysts, decompose only very slowly to yield the metal amide plus hydrogen gas [reaction (20)]. Solutions of metals in liquid ammonia are excellent reducing agents and are particularly useful for the reduction of organic compounds which are miscible with liquid ammonia.

$$e^-_{(am)} + NH_3 \rightarrow {}^1/_2H_2 + NH_2^- \tag{20}$$

The usefulness of liquid ammonia as a solvent is based on the differences in the chemical properties of liquid ammonia and of other common solvents, notably water. Principal among these differences (compared with water) are (1) the lesser tendency of ammonia to release protons, (2) the greater electron donor tendency (or proton affinity) of ammonia, and (3) the stronger reducing character of ammonia may be used as a solvent for very strong bases (such as NH_2^- or $C_2H_5O^-$) which would undergo complete protolysis in aqueous solution. Liquid ammonia may also be used as a solvent for very strong reducing agents (such as solvated electrons) which would immediately displace hydrogen from water. Because of the second difference, liquid ammonia solutions do not provide very strong acids, since all strong acids are converted immediately to ammonium ion, NH_4^+, which is a much weaker acid than hydronium ion, H_3O^+, its counterpart in aqueous systems. In summary, it may be said that as a solvent for chemical reactions, liquid ammonia affords much stronger bases and stronger reducing agents but much weaker acids and weaker oxidizing agents than does water. By application of these differences, a number of interesting synthetic procedures may be carried out in liquid ammonia.

Ammonia system of compounds. Many of the familiar compounds which contain oxygen may be considered to be derived from water as the parent solvent. In an analogous way it is sometimes useful to consider many nitrogen-containing compounds to be derived from ammonia as parent solvent. These latter compounds are sometimes considered to constitute the nitrogen or ammonia system of compounds. This analogy is useful in understanding the chemistry of various nitrogen compounds (**Tables 2** and **3**).

TABLE 3. Analogous compounds in the water and ammonia systems

Aquo	Ammonia
$KOH + (H_3O)Cl \longrightarrow KCl + 2H_2O$	$KNH_2 + (NH_4)Cl \longrightarrow KCl + 2NH_3$
$Zn + 2H_3O^+ \longrightarrow Zn^{2+} + H_2 + 2H_2O$	$Zn + 2NH_4^+ \longrightarrow Zn^{2+} + H_2 + 2NH_3$
$CH_3C(=O)OC_2H_5 + H_2O \xrightarrow{H_3O+} CH_3C(=O)OH + C_2H_5OH$	$CH_3C(=NH)NHC_2H_5 + NH_3 \xrightarrow{NH_4^+} CH_3C(=NH)NH_2 + C_2H_5NH_2$
$C(=O)(OH)_2 + 2OH^- \longrightarrow C(=O)(O^-)_2 + 2H_2O$	$C(=NH)(NH_2)_2 + 2NH_2^- \longrightarrow C(=NH)(NH^-)_2 + 2NH_3$
$Zn(OH)_2 + 2OH^- \longrightarrow Zn(OH)_4^{2-}$	$Zn(NH_2)_2 + 2NH_2^- \longrightarrow Zn(NH_2)_4^{2-}$

Synthesis of ammonia. The Haber-Bosch synthesis is the major source of industrial ammonia. In a typical process, water gas (CO, H_2, CO_2) mixed with nitrogen is passed through a scrubber cooler to remove dust and undecomposed material. The CO_2 and CO are removed by a CO_2 purifier and ammoniacal cuprous solution, respectively. The remaining H_2 and N_2 gases are passed over a catalyst at high pressures (up to 1000 atm or 100 megapascals) and high temperatures (approx. 700°C or 1300°F) [reaction (21)].

$$3H_{2(g)} + N_{2(g)} \rightleftharpoons 2NH_{3(g)} \qquad (21)$$

The ammonia is separated by absorption in water. Processes used vary widely in the sources of N_2 and H_2, treatment of the catalysts, temperature, pressure, and methods of ammonia separation.

Other industrial sources of ammonia include its formation as a by-product of the destructive distillation of coal, and its synthesis through the cyanamide process, which is indicated by reactions (22) and (23).

$$CaC_2 + N_2 \xrightarrow{1000^\circ C} CaCN_2 + C \qquad (22)$$

$$CaCN_2 + 3H_2O\ (\text{steam}) \longrightarrow CaCO_3 + 2NH_3 \qquad (23)$$

In the laboratory, ammonia is usually formed by its displacement from ammonium salts (either dry or in solution) by strong bases, as indicated by reaction (24). Another source is the hydrolysis of metal nitrides, as indicated reaction (25).

$$NH_4^+ + OH^- \xrightarrow{\text{heat}} NH_3 + H_2O \qquad (24)$$

$$Mg_3N_2 + 6H_2O \longrightarrow 3Mg(OH)_2 + 2NH_3 \qquad (25)$$

See AMIDE; AMINE; HIGH-PRESSURE PROCESSES; HYDRAZINE; NITROGEN. Harry H. Sisler

Bibliography. J. R. Jennings, *Catalytic Ammonia Synthesis*: *Fundamentals and Practice*, 1991; H. H. Sisler, R. D. Dresdner, and W. T. Mooney, Jr., *Chemistry*: *A Systematic Approach*, 1980; S. Strelzoff, *Technology and Manufacture of Ammonia*, 1981, reprint 1988.

Ammonium salt

A product of a reaction between ammonia, NH_3, and various acids. The general reaction for formation is (1). Examples of ammonium salts are ammo-

$$NH_3 + HX \longrightarrow NH_2X \qquad (1)$$

nium chloride, NH_4Cl, ammonium nitrate, NH_4NO_3, and ammonium carbonate, $(NH_4)_2CO_3$. These compounds are addition products of ammonia and the acid. For this reason, their formulas are sometimes written as $[H(NH_3)]X$.

All ammonium salts decompose into ammonia and the acid when heated. Their stability, however, varies according to the nature of the acid. Salts of weak acids decompose at lower temperatures than do salts of strong acids. Ammonium chloride, the salt of the strong acid hydrogen chloride, HCl, decomposes at 320°C (608°F), whereas ammonium sulfide, $(NH_4)_2S$, the salt of the weak acid hydrogen sulfide, H_2S, decomposes at 32°C (90°F). If the salt is heated in a closed vessel, a definite pressure of ammonia is established in the presence of the solid salt. This pressure is determined solely by the temperature and, if the acid is nonvolatile, is called the dissociation pressure at that temperature. For a detailed discussion of such equilibra *see* CHEMICAL EQUILIBRIUM.

If anhydrous ammonia is added to many of the ammonium salts at very low temperatures, salts containing several molecules of ammonia are formed. Ammonium chloride, for example, can add three or six molecules of ammonia to form complex salts. The reaction for the formation of the tetra compound is (2). The number of such complexes which maybe

$$[H(NH_3)]Cl + 3(NH_3) \longrightarrow [H(NH_3)_4]Cl \qquad (2)$$

formed depends upon the nature of the acid radical. When warmed, they lose ammonia; all are unstable above 0°C (32°F).

Ammonium chloride is made by absorbing ammonia in hydrochloric acid. It crystallizes from the solution in feathery crystals of the regular crystal system. It is a colorless solid with a density of 1.52. This salt, sometimes called sal ammoniac, is used in galvanizing iron, in textile dyeing, and in manufacturing dry cell batteries.

Ammonium nitrate, NH_4NO_3, a colorless salt with a density of 1.73, is prepared from ammonia and nitric acid. The solid salt deliquesces, or absorbs water from moist air, thus appearing to melt. It is used as a source of nitrous oxide, N_2O, or laughing gas, and in the manufacture of explosives. A mixture of ammonium nitrate and trinitrotoluene is referred to as amatol.

Ammonium sulfate, $(NH_4)_2SO_4$, obtained from ammonia and sulfuric acid, is a colorless solid with a density of 1.77. It is prepared commercially by passing ammonia and carbon dioxide, CO_2, into a suspension of finely ground calcium sulfate, $CaSO_4$, as shown in reaction (3). Large quantities are also pro-

$$CaSO_4 + 2NH_3 + CO_2 + H_2O \longrightarrow CaCO_3 + (NH_4)_2SO_4 \qquad (3)$$

duced as a by-product of coke ovens and coal-gas works. The chief use of ammonium sulfate is as a fertilizer.

Ammonium carbonate, $(NH_4)_2CO_3$, may be prepared by bringing ammonia and carbon dioxide together in aqueous solution, shown in reaction (4). It

$$2NH_3 + CO_2 + H_2O \longrightarrow (NH_4)_2CO_3 \qquad (4)$$

is also obtained by heating a mixture of ammonium sulfate and a fine suspension of calcium carbonate, shown in reaction (5).

$$(NH_4)_2SO_4 + CaCO_3 \longrightarrow (NH_4)_2CO_3 + CaSO_4 \qquad (5)$$

Ammonium thiocyanate, NH_4SCN, a colorless solid with a density of 1.31, is prepared by the reaction of ammonia and carbon disulfide, CS_2, as in reaction (6). It is used as a protective agent in dyeing.

$$2NH_3 + CS_2 \longrightarrow NH_4SCN + H_2S \qquad (6)$$

Except for several complex species such as ammonium chloroplatinate, $(NH_4)_2PtCl_6$, ammonium salts are very soluble in water. In aqueous solutions, they ionize to produce the ammonium ion, NH_4^+, and an acid anion. Solutions of ammonium salts of strong or moderately strong acids are acidic as a result of hydrolysis of the ammonium ion. The reaction involving a molecule of water and producing a hydrogen ion, H^+, is shown as reaction (7).

$$NH_4^+ + H_2O \rightleftharpoons NH_3 + H_3O^+ \qquad (7)$$

When a strong base is added to a solution of an ammonium salt, ammonia is evolved. This is a test for the presence of NH_4^+ ion and occurs because the reaction of reaction (8) is driven toward the left.

$$NH_3 + H_2O \rightleftharpoons NH_4^+ + OH^- \qquad (8)$$

See AMMONIA; FERTILIZER; HYDROLYSIS.

Francis J. Johnston

Amnesia

A significant but relatively selective inability to remember. Amnesia can be characterized along two dimensions with respect to its onset: an inability to remember events that occurred after the onset of amnesia is referred to as anterograde amnesia, and a deficit in remembering events that occurred prior to the onset of amnesia is referred to as retrograde amnesia. Amnesia can be due to a variety of causes, or etiologies, and it can be classified according to whether the cause is primarily neurological or psychological in origin.

Neurological amnesia. Neurological amnesias are the result of brain dysfunction and can be transient or permanent. They are usually characterized by a severe anterograde amnesia and a relatively less severe retrograde amnesia.

Transient amnesias are temporary memory disturbances and can range in duration from hours to months, depending on the cause and severity. They can be caused by epilepsy, head injury, and electroconvulsive therapy (most frequently used for the treatment of depression). In cases of transient global amnesia, an extensive amnesia that is usually sudden in onset and resolves within a day, the etiology is still not known, although many believe that it is vascular in origin. *See* SEIZURE DISORDERS.

Temporal lobe damage. Permanent amnesia can occur following brain damage to the medial temporal lobe. This discovery was made in 1953 following an experimental bilateral temporal lobe resection in an individual suffering from epilepsy. Although this procedure was moderately successful in treating the epilepsy, the patient developed severe anterograde amnesia. Although this individual had an average intelligence quotient (I.Q.) and normal short-term memory, he had great difficulty remembering recent events in his life. For example, he did not know where he lived, what he ate during his last meal, what year it was or how old he was; and he found it very difficult to learn new words and people's names. A temporally graded retrograde amnesia was also exhibited, with more remote memories being better preserved. The length of his retrograde amnesia was somewhere between 2 and 11 years. Bilateral temporal lobectomies are now avoided because of the resulting amnesia.

On occasion it is necessary to do a unilateral temporal lobectomy, which results in material specific memory deficits. Left temporal lobectomies result in verbal memory deficits and right temporal lobectomies result in nonverbal, visual memory deficits. Amnesia resulting from impairment to the medial temporal lobe can also occur following anoxia, cerebrovascular accidents, head injury, and viral infections to the brain. The primary structures involved in the processing of memory within the medial temporal lobe are the hippocampus and the amygdala.

Wernicke-Korsakoff syndrome. Neurological amnesia can also occur following damage to the diencephalon. One of the most common causes of diencephalic amnesia is Wernicke-Korsakoff syndrome, a disorder caused by a thiamine deficiency, usually related to chronic alcoholism. It is characterized by confusion, ataxia, and an oculomotor disturbance. Following thiamine replacement therapy, these symptoms reverse, but the individual is left with an amnesia. The major characteristics in diencephalic amnesia are very similar to medial temporal lobe amnesia. Damage to the mediodorsal thalamus and the mamillary bodies is largely responsible for the amnesia. *See* ALCOHOLISM.

Preserved memory function. Amnesics have provided important insights into how the diencephalon and medial temporal lobe are involved in preserved memory function. Amnesics have been shown to demonstrate normal learning on tasks involving motor, perceptual, and cognitive skills. They generally show a normal pattern of performance on tasks measuring how previously presented information indirectly influences subsequent responses. For example, amnesics will show normal rates of decreasing response times for identifying previously shown stimuli. This process is referred to as priming. Other examples include normal tendencies to complete word stems (such as MOT___) and word fragments (such as _O_EL) with previously presented words (Motel) when asked to complete the words with the first word that comes to mind. *See* COGNITION.

Multiple memory systems. Studies of preserved memory functions in amnesics supported the view that there are multiple memory systems represented in the brain. Damage to the medial temporal lobe area or the diencephalon appears to disrupt memory requiring conscious recollection of previously experienced information or events, frequently referred

to as explicit memory; however, amnesics' ability to show evidence of memory when conscious recollection is not required (referred to as implicit memory) is generally intact. Furthermore, since the retrograde amnesia is limited to events closer to the onset of the amnesia, with more remote memories better preserved, it appears that the diencephalon and medial temporal lobe play a role in the initial formation of explicit memory, as well as for some period of time, perhaps even years, following formation of the memory.

Memory improvements. Attempts to improve amnesics' ability to remember have generally met with little success. The most common approach to retraining has involved practicing remembering. Since damage to the medial temporal lobes and diencephalon is relatively permanent, and regeneration of damaged areas in the brain is limited, there is little hope that memory practice alone will improve the impaired functions, at least in adults. More encouraging approaches depend on using preserved implicit memory function in ways that will allow the patient to learn new information. *See* BRAIN; NEUROBIOLOGY.

Functional amnesia. Memory impairment that is not associated with brain damage is referred to as functional amnesia. Functional amnesia can be classified according to whether the amnesia is nonpathological or pathological. Nonpathological functional amnesia is a normal memory loss for events occurring during infancy and early childhood, sleep, hypnosis, and anesthesia. Pathological functional amnesia is an abnormal memory loss found in cases of functional retrograde amnesia and multiple personality. In contrast to neurological amnesia, pathological functional amnesia is usually associated with more severe retrograde than anterograde amnesia. *See* ANESTHESIA; HYPNOSIS; SLEEP AND DREAMING.

Functional retrograde amnesia. Functional retrograde amnesia usually occurs in response to a severe emotional trauma. Initially, individuals typically enter a fugue state in which they may wander around for an indefinite period of time. This stage is characterized by loss of personal identity, loss of memory of their past, and a lack of awareness of their amnesia. This stage usually ends when the individual is challenged by questions about his or her identity or past. Individuals who do not fully recover from this stage will enter a period characterized by awareness of the amnesia. The amnesia may spontaneously remit in response to an external trigger, such as the appearance of a familiar person, an occurrence that reminds them of their identity or past, or as the result of hypnosis or sodium amytal therapy. Recovered patients typically remain amnestic for the events during the fugue state, but usually remember the events during the amnestic period following the fugue state. Various explanations of functional retrograde amnesia have been proposed, including a dissociation of traumatic information from the ego, repression of some threatening event or information by the ego, and a selective failure of episodic memory, memory characterized by events in one's life.

Multiple personality disorder. A common characteristic in multiple personality disorders is amnesia among different personalities. The amnesia may be unidirectional, in which one personality will be amnestic for another personality, but the latter personality will not be amnestic for the former; or bidirectional, in which two personalities will be amnestic for each other. The amnesia can be quite problematic for an individual with a multiple personality disorder when the present personality cannot account for a predominant action of another personality (for example, when one personality spends the money earned by another personality). Amnesia in individuals with multiple personality disorders has been explained in terms of a dissociation from the ego of other personalities. Additional explanations include difficulty in retrieving stored information associated with the different states related to different personalities, and difficulty in retrieving analogous to posthypnotic states. Just as in the neurological amnesias, there have been demonstrations of intact implicit memory in functional amnesia with relatively severe impairment of explicit memories. *See* MEMORY; SCHIZOPHRENIA.

Richard S. Lewis

Bibliography. A. J. Parkin and N. R. C. Leng, *Neuropsychology of the Amnesic Syndrome,* 1992; L. Squire (ed.), *Memory and Its Disorders,* 1991; L. R. Squire and N. Butters (eds.), *Neuropsychology of Memory,* 2d ed., 1992.

Amnion

A thin, cellular, extraembryonic membrane forming a closed sac which surrounds the embryo in all reptiles, birds, and mammals and is present only in these forms; hence the collective term amniotes is applied to these animals. The amnion contains a serous fluid in which the embryo is immersed. *See* AMNIOTA.

Typically, the amnion wall is a tough, transparent, nerve-free, and nonvascular membrane consisting of two layers of cells; an inner, single-cell-thick layer of ectodermal epithelium, continuous with that covering the body of the embryo as the outer layer of its skin, and an outer covering of mesodermal, connective, and specialized smooth muscular tissue, also continuous with the mesodermal germ layer of the embryo. Early after the formation of the amnion, waves of contraction of the muscles pass over the amniotic sac and produce a characteristic rocking of the embryo. *See* GERM LAYERS.

In reptiles, birds, and some mammals, the amnion arises by a process of folding over the embryo body of the extraembryonic ectoderm along with its underlying, closely applied mesoderm (collectively, the somatopleure). Head, tail, and lateral folds of this sheet of tissue meet and fuse over the back of the embryo. Only the inner limb of the fold forms the true amnion, the outer limb of the fold becoming part of another fetal membrane, the chorion. In other mammals, including humans, the amnion arises by a process of

cavitation in a mass of cells in which embryonic and extraembryonic cells become separated. The cavity forms above those cells destined to form the embryo body and eventually spreads over and around the embryo up to the region of the developing umbilical cord.

The major function of the amnion and its fluid is to protect the delicate embryo. Thus, developmental stages of terrestrial animals are provided with the same type of cushioning against mechanical shock as is provided by the water environment of aquatic forms. *See* FETAL MEMBRANE. Nelson T. Spratt, Jr.

Amniota

A collective term for the classes Reptilia (reptiles), Aves (birds), and Mammalia (mammals) of the subphylum Vertebrata. The remaining vertebrates, including the several classes of fishes and the amphibians, are grouped together as the Anamnia. Members of the Amniota are characterized by having a series of specialized protective extraembryonic membranes during development. Three of the membranes—amnion, chorion or serosa, and allantois—occur only in this group, but a fourth, the yolk sac, is sometimes present and is found in many anamniotes. The presence of the extraembryonic membranes makes it possible for the embryonic development of the amniotes to take place out of the water. In the most primitive forms the early stages of development take place inside a shell-covered egg that is deposited on land. This pattern is typical of most reptiles, all birds, and some mammals (**Figs. 1** and **2**). In these animals the amnion and chorion form fluid-filled sacs which protect the embryo from desiccation and shock. The allantois usually acts as a storage place for digestive and nitrogenous wastes and, in conjunction with the chorion, as a respiratory structure. In viviparous reptiles and mammals the chorion and allantois generally fuse and become more or less intimately associated with the uterine lining of the mother. Nutritive, excretory and respiratory exchanges take place across the chorioallantoic membrane between the allantoic circulation of the embryo and the uterine circulatory vessels of the mother. *See* ALLANTOIS; AMNION; ANAMNIA; CHORION; VERTEBRATA; YOLK SAC. Jay M. Savage

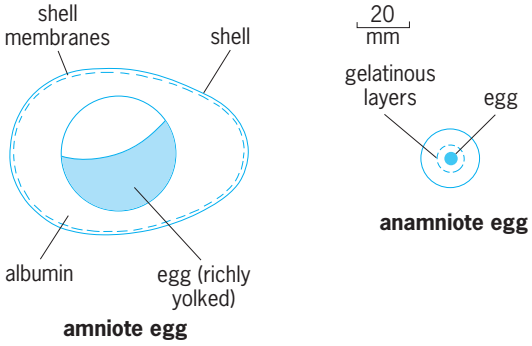

Fig. 1. Basic types of vertebrate eggs.

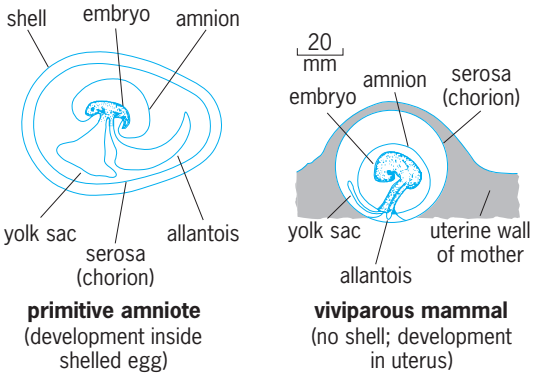

Fig. 2. Embryo developing in amniote eggs.

Amoebida

An order of Lobosia without protective coverings (tests). These protozoa range in size from about 4 micrometers to 0.08–0.12 in. (2–3 mm). Pellicles may be thin, as in *Amoeba proteus*, or thicker and less flexible, as in *Thecamoeba verrucosa* (**Fig. 1***a*). Pellicular folds may develop during locomotion, particularly in species with thick pellicles. Both flagellate and ameboid stages occur in certain soil amebas. In *Naegleria*, for example, the cycle includes an ameba, a flagellate, and a cyst. Most Amoebida have no flagellate stage.

Locomotion. The term ameboid movement is rather loose because locomotion of Amoebida varies somewhat from genus to genus. In some cases movement involves protoplasmic flow of the body as a whole, without typical pseudopodia. In *A. proteus* (Fig. 1*c* and *d*) there may be several ridged indeterminate pseudopodia, into one of which the organism appears to flow in locomotion. In other species (Fig. 1*b*) determinate pseudopodia never become large enough to direct locomotion. In some cases (Fig. 1*l*–*n*) the form of the pseudopodia may vary in a single species. Certain amebas commonly have a relatively inert posterior mass, the uroid, which may or may not be partially constricted, is sometimes covered with projections, and often contains food vacuoles. The uroid may be discarded by autotomy.

Locomotion involving protoplasmic flow depends upon sol-gel reversibility. According to the contractile-hydraulic theory, endoplasm of *A. proteus* is under constant pressure from the ectoplasm (plasmagel). The plasmagel contracts posteriorly to drive the endoplasm (plasmasol) forward to a point at which pressure from the plasmagel is relaxed momentarily. The result is a bulge, a developing pseudopodium (Fig. 1*c* and *m*). The plasmasol reaching the pseudopodial tip is diverted peripherally to make contact with, and become part of, the plasmasol at the posterior end of the body. The new plasmasol flows anteriorly and the process is repeated. In

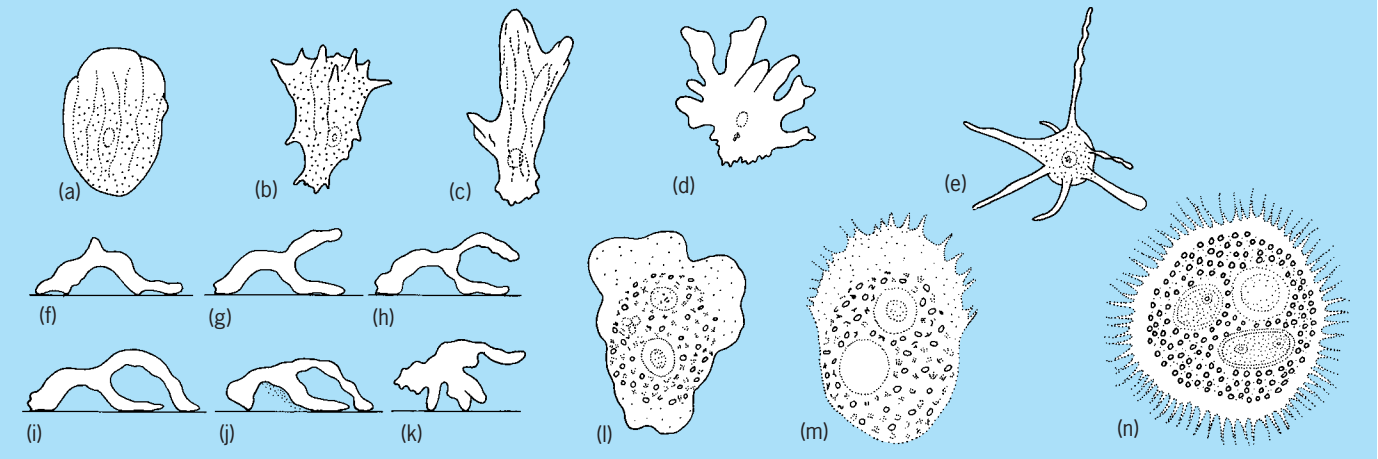

Fig. 1. Typical ameboid locomotion. (*a*) *Thecamoeba verucosa*, locomotion without formation of distinct pseudopodia. (*b*) *Mayorella bigemia*, formation of conical pseudopodia. (*c*) *Amoeba proteus*, formation of large pseudopodia. (*d*) *Amoeba dubia*, formation of a number of large pseudopodia. (*e*) *Astramoeba flagellipodia*, floating form with slender and sometimes spiral pseudopodia. (*f–k*) *Chaos* (*Pelomyxa*) *carolinensis*, locomotion of walking type, as seen in thriving cultures. (*l–n*) *Acanthamoeba castelanii*, specimens with different forms of pseudopodia, 12–30 μm. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)

Thecamoeba there may be a rolling type of locomotion. A peculiar kind of walking has been described in one of the giant amebas (Fig. 1*f–k*). *See* CELL MOTILITY.

Cellular inclusions. Amebas, normally phagotrophic, usually contain food vacuoles. Certain species contain crystals of apparently differing chemical nature. Other inclusions are globules of different sizes, mitochondria, and stored food reserves. In addition, bacteria or algae may occur in the cytoplasm, changing the color to a gray or green. Nuclei range in number from one to several hundred, as in *Chaos carolinensis*. The giant amebas are visible without a microscope. *See* CELL (BIOLOGY).

Endoparasitic species. Those species found in the digestive tract of invertebrates and vertebrates include relatively harmless species and a few pathogens, such as *Entamoeba histolytica* of humans and *E. invadens* of reptiles.

Amebiasis. Entamoeba histolytica causes amebiasis. In primary cases the amebas are localized in the colon. Cases range from mild amebiasis to acute amebic dysentery. The amebas (**Fig. 2***f–k*) invade the wall of the colon, causing ulceration in symptomatic infections. Complications may include perforations of the appendix or colon, adhesions involving the colon, or other visceral damage. In secondary amebiasis the amebas may invade the ileum or liver (causing hepatic abscess) or, rarely, the lungs, brain, spleen, lymph glands, urinary bladder, uterus, penis, vagina, and skin. The amebas may migrate to the ileum; the usual route to most other organs is the blood, but contamination is probably involved in invasion of the urogenital tract. Severe infections may be correlated with vitamin deficiencies or generally substandard diets. Intestinal amebiasis has been treated with various drugs, such as arsenicals (carbarsone), quinoline derivatives (diodoquin, vioform), alkaloid derivatives such as emetine, and antibiotics such as terramycin.

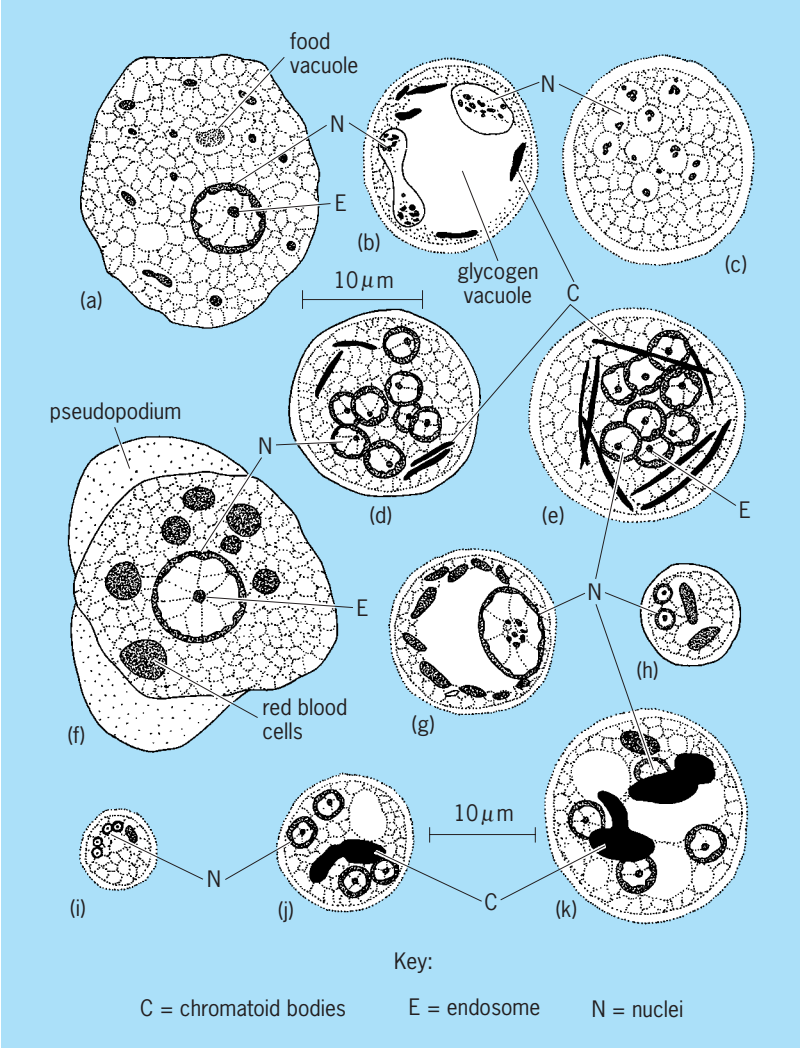

Fig. 2. Commensal and parasitic amebas. (*a–e*) *Entamoeba coli*: (*a*) rounded form; (*b*) cyst; (*c*)multinucleate cyst; (*d*) typical cyst from naturally infected monkey; (*e*) octunucleate cyst. (*f–k*) *E. histolytica*: (*f*) ameboid form; (*g*) cyst with glycogen vacuole; (*h*) binucleate cyst; (*i, j, k*) quadrinucleate cysts. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)

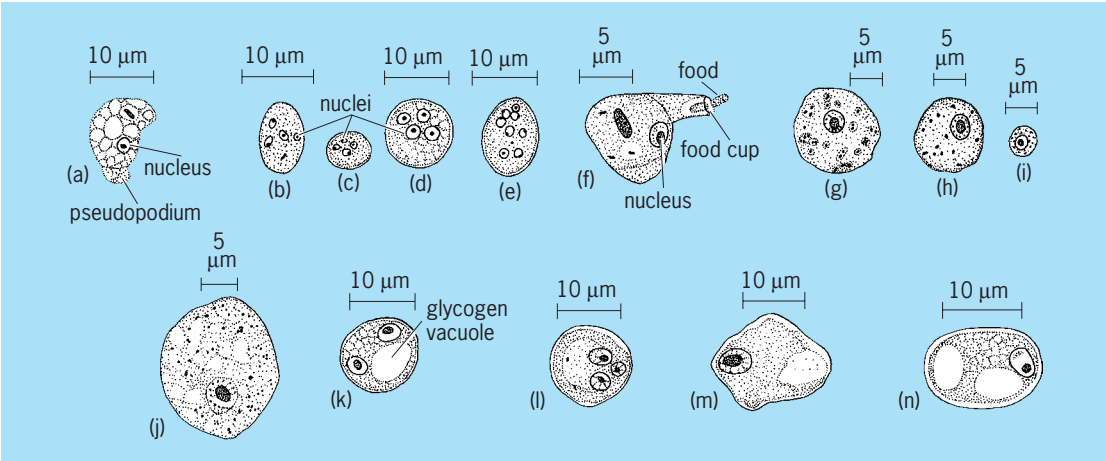

Fig. 3. Nonpathogenic amebas. (*a–e*) *Endolimax nana*: (*a*) ameboid stage; (*b*) cyst from monkey; (*c–e*) cyst from human. (*f–n*) *Iodamoeba bütschlii*: (*f*) ameboid stage; (*g,h*) ameboid stage, medium size; (*i*) ameboid stage, small size; (*j*) ameboid stage, large size; (*k*) binucleate cyst; (*l*) cyst showing three cysts; (*m*) uninucleate cyst; (*n*) uninucleate cyst from monkey. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)

Nonpathogenic species. *Entamoeba coli* (Fig. 2*a–e*), a similar ameba, does not invade human tissues. Also limited to the lumen of the colon are *Endolimax nana, Iodamoeba bütschlii* (**Fig. 3**), and *Dientamoeba fragilis*. These four are relatively harmless although sometimes associated with digestive disturbances. In laboratory identification, number and structure of nuclei, shape of chromatoid bodies (Fig. 2) if present, and size and shape of cysts are important criteria. Intestinal amebas are transferred usually as cysts which are passed in the feces. Voided cysts remain viable for several days if kept moist and shaded. *Dientamoeba fragilis*, for which cysts are unknown, is an apparent exception to usual methods of transfer. Cysts are distributed widely by an infected individual and a new host becomes infected by swallowing cysts. Uncooked vegetables from soil fertilized with human feces are a potential source of infection. Standard methods of water purification seem reasonably protective, but it is difficult to control spread of cysts by food handlers. *See* LOBOSIA; PROTOZOA; SARCODINA; SARCOMASTIGOPHORA.

Richard P. Hall

Amorphous solid

A rigid material whose structure lacks crystalline periodicity; that is, the pattern of its constituent atoms or molecules does not repeat periodically in three dimensions. In the present terminology amorphous and noncrystalline are synonymous. A solid is distinguished from its other amorphous counterparts (liquids and gases) by its viscosity: a material is considered solid (rigid) if its shear viscosity exceeds $10^{14.6}$ poise ($10^{13.6}$ pascal · second). *See* CRYSTAL; SOLID-STATE PHYSICS; VISCOSITY.

Preparation. Techniques commonly used to prepare amorphous solids include vapor deposition, electrodeposition, anodization, evaporation of a solvent (gel, glue), and chemical reaction (often oxidation) of a crystalline solid. None of these techniques involves the liquid state of the material. A distinctive class of amorphous solids consists of glasses, which are defined as amorphous solids obtained by cooling of the melt. Upon continued cooling below the crystalline melting point, a liquid either crystallizes with a discontinuous change in volume, viscosity, entropy, and internal energy, or (if the crystallization kinetics are slow enough and the quenching rate is fast enough) forms a glass with a continuous change in these properties. The glass transition temperature is defined as the temperature at which the fluid becomes solid (that is, the viscosity = $10^{14.6}$ poise = $10^{13.6}$ Pa · s) and is generally marked by a change in the thermal expansion coefficient and heat capacity. [Silicon dioxide (SiO_2) and germanium dioxide (GeO_2) are exceptions.] It is intuitively appealing to consider a glass to be both structurally and thermodynamically related to its liquid; such a connection is more tenuous for amorphous solids prepared by the other techniques. *See* GLASS.

Types of solids. Oxide glasses, generally the silicates, are the most familiar amorphous solids. However, as a state of matter, amorphous solids are much more widespread than just the oxide glasses. There are both organic (for example, polyethylene and some hard candies) and inorganic (for example, the silicates) amorphous solids. Examples of glass formers exist for each of the bonding types: covalent [As_2S_3], ionic [$KNO_3{-}Ca(NO_3)_2$], metallic [Pd_4Si], van der Waals [*o*-terphenyl], and hydrogen [$KHSO_4$]. Glasses can be prepared which span a broad range of physical properties. Dielectrics (for example, SiO_2) have very low electrical conductivity and are optically transparent, hard, and brittle. Semiconductors (for example, As_2SeTe_2) have intermediate electrical conductivities and are optically opaque and brittle. Metallic glasses (for example, Pd_4Si) have high electrical and thermal conductivities, have metallic luster, and are ductile and strong.

Uses. The obvious uses for amorphous solids are as window glass, container glass, and the glassy

polymers (plastics). Less widely recognized but nevertheless established technological uses include the dielectrics and protective coatings used in integrated circuits, and the active element in photocopying by xerography, which depends for its action upon photoconduction in an amorphous semiconductor. In optical communications a highly transparent dielectric glass in the form of a fiber is used as the transmission medium. In addition, metallic amorphous solids have been considered for uses that take advantage of their high strength, excellent corrosion resistance, extreme hardness and wear resistance, and unique magnetic properties. *See* OPTICAL COMMUNICATIONS.

Semiconductors. It is the changes in short-range order (on the scale of a localized electron), rather than the loss of long-range order alone, that have a profound effect on the properties of amorphous semiconductors. For example, the difference in resistivity between the crystalline and amorphous states for dielectrics and metals is always less than an order of magnitude and is generally less than a factor of 3. For semiconductors, however, resistivity changes at a given temperature of 10 orders of magnitude between the crystalline and amorphous states are not uncommon, and accompanying changes in optical properties can also be large.

Electronic structure. The model that has evolved for the electronic structure of an amorphous semiconductor is that the forbidden energy gap characteristic of the electronic states of a crystalline material is replaced in an amorphous semiconductor by a pseudogap. Within this pseudogap the density of states of the valence and conduction bands is sharply lower but tails off gradually and remains finite due to structural disorder (**Fig. 1**). The states in the tail region are localized; that is, their wave functions extend over small distances in contrast to the extended states that exist elsewhere in the energy spectrum. Because the localized states have low mobility (velocity per unit electric field), the extended states are separated by a mobility gap (Fig. 1) within which charge transport is markedly impeded. In each band, the energy at which the extended states meet the localized states is called the mobility edge. *See* BAND THEORY OF SOLIDS.

An ideal amorphous solid can be conceptually defined as having no unsatisfied bonds, a minimum of bond distortions (bond angles and lengths), and no internal surfaces associated with voids. Deviations from this ideality introduce localized states in the gap in addition to those in the band edge tails due to disorder alone. One important defect is called an unsatisfied, broken, or dangling bond. These dangling bonds create states deep in the gap which can act as recombination centers and markedly limit carrier lifetime and mobility. A large number of such states introduced, for example, during the deposition process will dominate the electrical properties.

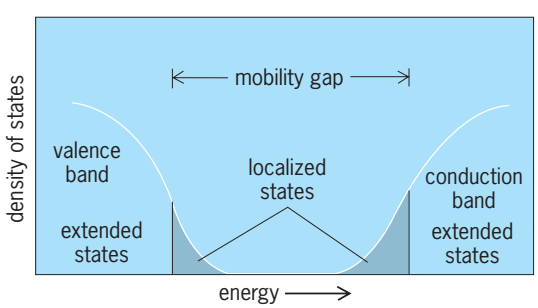

Fig. 1. Density of states versus energy for an amorphous semiconductor.

Charge transport can occur by two mechanisms. The first is conduction of mobile extended-state carriers (analogous to that which occurs in crystalline semiconductors), for which the conductivity is proportional to exp $(-E_g/2kT)$, where E_g is the gap width, T is the absolute temperature, and k is Boltzmann's constant. The second mechanism is hopping of the localized carriers, for which the conductivity is proportional to exp $[-(T_o/T)^{1/4}]$, where T_o is a constant (Mott's law). At low temperatures carriers hop from one localized trap to another, whereas at high temperatures they can be excited to the mobility edge.

Glassy chalcogenides. One class of amorphous semiconductors is the glassy chalcogenides, which contain one (or more) of the chalcogens sulfur, selenium, or tellurium as major constituents. These amorphous solids behave like intrinsic semiconductors, show no detectable unpaired spin states, and exhibit no doping effects. It is thought that essentially all atoms in these glasses assume a bonding configuration such that bonding requirements are satisfied; that is, the structure accommodates the coordination of any atom. These materials have application in switching and memory devices. *See* GLASS SWITCH.

Tetrahedrally bonded solids. Another group is the tetrahedrally bonded amorphous solids, such as amorphous silicon and germanium. These materials cannot be easily formed by quenching from the melt (that is, as glasses) but can be prepared by one of the deposition techniques mentioned above.

When amorphous silicon (or germanium) is prepared by evaporation, not all bonding requirements are satisfied, so a large number of dangling bonds are introduced into the material. These dangling bonds are easily detected by spin resonance or low-temperature magnetic susceptibility and create states deep in the gap which limit the transport properties. The number of dangling bonds can be reduced by a thermal anneal below the crystallization temperature, but the number cannot be reduced sufficiently to permit doping.

Amorphous silicon prepared by the decomposition of silane (SiH_4) in a plasma has been found to have a significantly lower density of defect states within the gap, and consequently the carrier lifetimes are expected to be longer. This material can be doped *p*- or *n*-type with boron or phosphorus (as examples) by the addition of B_2H_6 or PH_3 to the SiH_4 during deposition. This permits exploration of possible devices based on doping, which are analogous to devices based on doping of crystalline silicon.

One reason plasma-deposited silicon has a

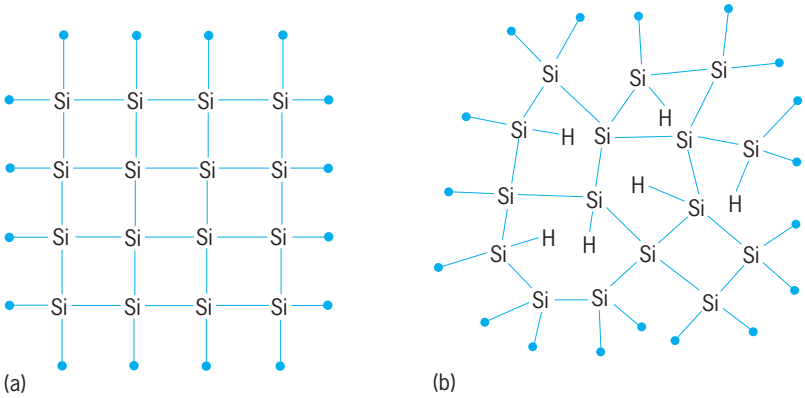

Fig. 2. Bonding of silicon. Bonds which continue the network are shown terminated by a dot. (*a*) Crystalline arrangement. (*b*) Amorphous structure with dangling (or unsatisfied) bonds terminated by hydrogen.

significantly lower density of defect states within the gap is that the process codeposits large amounts of hydrogen (typically 5–30% of the atoms, depending upon deposition conditions), and this hydrogen is very effective at terminating dangling bonds (**Fig. 2**). Other possible dangling-bond terminators (for example, fluorine) have been explored.

The ability to reduce the number of states deep in the gap and to dope amorphous silicon led directly to the development of an amorphous silicon photovoltaic solar cell. Solar conversion efficiencies of 10% have been achieved in the laboratory, leading to a development effort aimed at large-scale power applications. The appeal of amorphous silicon is that it holds promise for low-cost, easily fabricated, large-area cells.

Amorphous silicon solar cells have been constructed in heterojunction, *pin*-junction, and Schottky-barrier device configurations, and are widely used in consumer products such as calculators and watches. The optical properties of amorphous silicon provide a better match to the solar spectrum than do those for crystalline silicon, but the transport properties of the crystalline material are better. Experiments indicate that hole transport in the amorphous material is the limiting factor in the conversion efficiency. *See* SEMICONDUCTOR; SEMICONDUCTOR DIODE; SEMICONDUCTOR HETEROSTRUCTURES; SOLAR CELL; SOLAR ENERGY.

Brian G. Bagley

Bibliography. S. Elliott, *Physics of Amorphous Materials*, 2d ed., 1990; N. F. Mott, *Conduction in Non-Crystalline Materials*, 2d ed., 1993; K. Morigaki, *Physics of Amorphous Semiconductors*, 1994; R. Zallen, *The Physics of Amorphous Solids*, 1983.

Ampère's law

A law of electromagnetism which expresses the contribution of a current element of length dl to the magnetic induction (flux density) B at a point near the current. Ampère's law, sometimes called Laplace's law, was derived by A. M. Ampère after a series of experiments during 1820–1825. It plays a fundamental role in the International System (SI) definition of the ampere. *See* PHYSICAL MEASUREMENT.

Whenever an electric charge is in motion, there is a magnetic field associated with that motion. The flow of charges through a conductor sets up a magnetic field in the surrounding region. Any current may be considered to be broken up into infinitesimal elements of length dl, and each such element contributes to the magnetic induction at every point in the neighborhood. The contribution dB of the element is found to depend upon the current I, the length dl of the element, the distance r of the point P from the current element, and the angle θ between the current element and the line joining the element to the point P (see **illus.**). Ampère's law expresses the manner of the dependence by Eq. (1).

$$dB = k\frac{I\,dl\,\sin\theta}{r^2} \tag{1}$$

Choice of units. The proportionality factor k depends upon the units used in Eq. (1) and upon the properties of the medium surrounding the current. As in other equations expressing observed relationships, there is an arbitrary choice as to which of the units is to be defined from Ampère's law, or by a relationship derived from the law. If values are assigned to B, I, and l, the factor k can be found by experiment. If, however, an arbitrary value is assigned to k, Ampère's law or any equation derived from Ampère's law may be used to define the unit of current since units of B are otherwise defined. In SI, the latter choice is made, a value is assigned to k, and the ampere as a unit of current is defined from an equation derived from Ampère's law. When the current is in empty space, the factor k is assigned a value of 10^{-7} weber/A-m.

As in other equations associated with electric and magnetic fields, for example Coulomb's law, it is convenient to replace k by a new factor μ_0 related to k as in Eq. (2). This substitution removes the factor

$$\mu_0 = 4\pi k \tag{2}$$

4π from many derived equations in which it would otherwise appear. With this substitution Ampère's law becomes Eq. (3). The factor μ_0 is called the

$$dB = \frac{\mu_0}{4\pi}\frac{I\,dl\,\sin\theta}{r^2} \tag{3}$$

permeability of empty space: it has the value $4\pi \times 10^{-7}$ henry/m. *See* ELECTRICAL UNITS AND STANDARDS.

Right-hand rule. The direction of dB is always perpendicular to the plane determined by the line

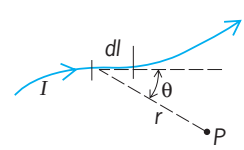

Graphic representation of Ampère's law.

tangent to the current element dl and the line joining dl to P. The sense of the lines is clockwise when looking in the direction of the current. The direction of each point may also be described in terms of a right-hand rule. If the current element is grasped by the right hand with the thumb pointing in the direction of the current, the fingers encircle the current in the direction of the magnetic induction. (If electron flow is used as the convention for current direction, the left hand is used in place of the right.)

Since the magnetic induction is everywhere perpendicular to the current element, it follows that the lines of induction or flux always form closed paths.

Field near a current. From Ampère's law, the field near a current may be calculated by finding the vector sum of the contributions of all the various elements that make up the current. This sum can be found (provided the integration can be carried out) by integrating over the whole length of the current, as in Eq. (4), where the limits are taken so that all

$$B = \int dB = \frac{\mu_0}{4\pi} \int \frac{I\, dl\, \sin\theta}{r^2} \tag{4}$$

current elements are included and the integral represents the vector sum.

The experimental test of the validity of Ampère's law is not direct since experiments are made not upon the flux density due to individual elements, but upon the flux density resulting from the current as a whole. Thus, the applications of Ampère's law are in the computation of the flux density for known geometrical arrangements of the current. For simple current paths, the summation is easily carried out, as for the flux density at the center of a single circular conductor of radius a. For this conductor, $r = a$, and r is always perpendicular to dl, so that $\sin\theta = 1$. Furthermore, all contributions are perpendicular to the plane of the coil, so that the vector sum is the arithmetic sum as in Eq. (5). The flux density anywhere

$$B = \frac{\mu_0}{4\pi} \int_0^{2\pi a} \frac{I\, dl}{a^2} = \frac{\mu_0}{4\pi} \frac{I}{a^2} \int_0^{2\pi a} dl = \frac{\mu_0 I}{2a} \tag{5}$$

on the axis of a flat coil, on the axis of a solenoid, or near a long straight conductor may be computed readily, and the result compared with experimental values. In general, the field of any current may be evaluated if dB can be expressed in Ampère's law, and the integration can be carried out. For the case of a long straight conductor *see* BIOT-SAVART LAW.

Ampère's law can be used as an alternative definition of magnetic induction (flux density). *See* MAGNETISM.

Kenneth V. Manning

Bibliography. B. I. Bleaney and B. Bleaney, *Electricity and Magnetism*, 3d ed., 1976, paper 1989; W. J. Duffin *Electricity and Magnetism*, 4th ed., 1990; D. Halliday, R. Resnick, and K. Krane, *Physics*, 5th ed., 2002; E. M. Purcell, *Electricity and Magnetism*, 2d ed., 1985; H. D. Young and R. A. Freedman, *Sears and Zemansky's University Physics*, 11th ed., 2003.

Amphetamine

A stimulating drug that affects the brain and the body in a variety of ways, also known by the trade name Benzedrine. Chemically, amphetamine is a racemic mixture of the *l* and *d* isomers of α-methyl-β-phenethylamine. The *l* isomer has more pronounced effects on the body, while the *d* isomer (commonly referred to as Dexedrine) has a greater effect on the brain. On the whole, the pharmacological effects of amphetamine are to produce an increase in blood pressure, a relaxation of bronchial smooth muscle, a constriction of the blood vessels supplying the skin and mucous membranes, and a variety of alterations in behavior.

Pharmacology. The mechanisms by which amphetamine produces its effects are not precisely defined. The effects on the body seem to be mediated predominantly through an increase in the activity of the neurons in the sympathetic nervous system via the transmitter norepinephrine. Amphetamine has the ability to release norepinephrine from nerve terminals. The consequence of release is that amphetamine has a spectrum of activity similar to the normal physiological effects of norepinephrine on the peripheral nervous system. For example, amphetamine and norepinephrine increase blood pressure by increasing the strength with which the muscles of the heart contract. Both drugs cause a constriction of blood vessels in the skin and mucous membranes, and they both cause the smooth muscles in the lung to relax.

Similarly, the major effects of amphetamine on the brain have been related to its ability to release norepinephrine in the hypothalamus, the reticular activating system, and the cerebral cortex. However, amphetamine may also release another neurotransmitter, dopamine, in the mesolimbic system and the cerebral cortex, and some of the effects of amphetamine on the brain are due to this other chemical.

The effects of amphetamine on the brain are generally stimulating. The drug causes an increase in motor activity, a decrease in appetite, an increase in arousal and ability to concentrate, an elevation of mood, insomnia, an enhanced sense of well-being, and hyperventilation. *See* CENTRAL NERVOUS SYSTEM; SYMPATHETIC NERVOUS SYSTEM.

Indications. The pharmacological effects noted above are the basis for the use of this drug in medicine. For example, the ability of amphetamine to contract blood vessels in the mucous membranes and to relax smooth muscles in the lung make it an efficacious nasal decongestant. Indeed, the original use of amphetamine was in Benzedrine inhalers for the purpose of clearing the nasal passages.

In a like manner, the pronounced excitatory effects of amphetamine on the brain have led to several other medical uses. For example, its ability to cause hyperventilation (by stimulation of the respiratory centers in the medulla) has resulted in its use as an analeptic. The arousing and insomnia-producing

effects (via stimulation of the reticular activating system and the cerebral cortex) have been exploited to treat narcolepsy, a disease characterized by an inability to stay awake. Furthermore, the ability of amphetamine to cause a loss of appetite (via the hypothalamus) has promoted its use in diet programs as an anorexic; however, tolerance to this effect limits its usefulness. Finally, the enhanced sense of well-being and mild euphoria that are seen after taking amphetamine have led to its use in certain forms of psychiatric depression.

One important use of amphetamine is in the treatment of hyperkinetic children. Hyperkinesis is a behavioral disorder in children that is characterized by increased motor activity, short attention span, lack of concentration, impulsiveness, emotional lability, and low tolerance for frustration. Consequently the child is difficult to control and disrupts discipline in the home and classroom. Amphetamine has been used successfully to treat this syndrome, although the paradox of how a stimulant can reduce hyperkinetic behavior is not understood.

Side effects. Just as the therapeutic uses of amphetamine are a consequence of its pharmacology, so are its side effects. Thus, the typical side effects on the body include dry mouth, heart rhythm alterations (palpitations and arrhythmias), hypertension, stomach cramps, and decreased urinary frequency. The central nervous system side effects include dizziness, dysphoria, headache, tremor, restlessness, insomnia, decreased appetite, increased aggressiveness, anxiety, and paranoid panic states. Extreme overdosage can result in convulsions, cerebral hemorrhaging, coma, and death.

Illegal uses. The stimulant and euphorogenic side effects of amphetamine have made this drug subject to widespread abuse. The patterns of this abuse, however, vary greatly. For example, in the early 1930s, when amphetamine was in Benzedrine inhalers, the drug was taken by dunking the inside paper liner of the inhaler in a glass of beer and then drinking the mixture. Today, most self-administration is with pills and capsules that are illegally diverted from legitimate sources. There also developed growing abuse of methamphetamine (a chemical derivative of amphetamine). This drug is taken in crystal form, either intranasally ("snorted") or intravenously ("mainlined") after being dissolved in water. Another problem concerns multiple drug use in a single individual (polydrug abuse). For example, amphetamine is often taken in "binges" which are terminated by taking large doses of depressants (such as secobarbital). A cycle then develops of alternating "uppers" and "downers." Amphetamine is also used as an adulterant of illegal cocaine; the mixture purportedly has a more pleasant and longer-lasting effect. Finally, amphetamine has been mixed with heroin, with the resulting concoction (a "speedball") supposedly having a more potent and euphorogenic effect. However, an implicit danger in taking illegal amphetamine, as with taking any illegal drug, is that the user has no idea as to purity or the source of the drug. Thus, the safety of the user is jeopardized.

The diversity of people taking amphetamine without proper medical supervision makes it difficult to characterize a typical user. Amphetamine abuse by students has been observed. There is also evidence that it is used by truck drivers to fend off fatigue and boredom so that they can drive longer trips. The drug is used by housewives to give them a "lift" and help them through their day. It is thought to be used by athletes to increase endurance and improve performance. It is also taken by diet-conscious people who wish to use amphetamine's anorexic effects to help lose weight. Amphetamine abuse by physicians and nurses has also been reported; they use the drug to stay alert during their long working hours.

More serious abuse occurs in some individuals who take the drug over long periods. As with other drugs of abuse (such as morphine, alcohol, or barbiturates), tolerance will occur after repeated dosing. Thus, whereas in the past one pill might be sufficient to induce the appropriate high, it becomes necessary to increase that amount in the future to achieve the same feeling. Over time, the users may become "psychologically dependent" on amphetamine, and feel that they cannot live without it. Subsequent drug usage and the acquisition of more drugs may then dominate and deleteriously affect their lives. *See* ALCOHOLISM; BARBITURATES; MORPHINE ALKALOIDS.

Chronic use of large amounts of amphetamine can have severe effects on personality. It has been shown that large doses of amphetamine can induce a behavioral state in humans that is nearly indistinguishable from paranoid schizophrenia, but can be reversed upon cessation of the drug. This drug-induced panic state is dangerous not only to the user (whose irrational acts may be self-destructive) but also to those in contact with the user. This is because, in these fearful, illogical, and agitated states, the user may become deluded that other persons have conspiratorial or hurtful intentions; this may lead to impulsive, "self-defensive" acts by the user. The potential for injury then becomes very high, and acute psychiatric intervention is critical.

The question of whether or not physical dependence occurs with amphetamine is unresolved. It is known that after long-term treatment with amphetamine, removal of the drug results in feelings of anxiety, stomach cramps, headache, and extreme fatigue. However, the mild severity of these symptoms makes it unclear if this is a true withdrawal response. Regardless, the other effects of long-term abuse of amphetamine (weight loss, insomnia, behavioral disorders) make unsupervised self-administration highly risky and unadvisable. *See* ADDICTIVE DISORDERS; NARCOTIC.

Richard E. Chipkin

Bibliography. L. S. Goodman and A. Gillman (eds.), *The Pharmacological Basis of Therapeutics*, 9th ed., 1996; L. E. Hollister, *Clinical Use of Psychotherapeutic Drugs*, 1973; R. M. Julien, *A Primer of Drug Action*, 7th ed., 1995; O. S. Ray, *Drugs, Society and Human Behavior*, 8th ed., 1998.

Amphibia

One of four classes (the others are Reptilia, Aves, and Mammalia) composing the superclass Tetrapoda of the subphylum Vertebrata. The living amphibians number over 4700 species and are classified in three orders: Salientia or Anura (frogs and toads, over 4200 species); Urodela or Caudata (salamanders, approximately 420 species); and Gymnophiona or Apoda (caecilians, about 165 species). The orders in the subclasses Labyrinthodontia and Lepospondyli are now extinct. A classification scheme for the Amphibia follows. See separate articles on each group listed.

Class Amphibia
- Subclass Labyrinthodontia
  - Order: Ichthyostegalia
    - Temnospondyli
    - Anthracosauria
- Subclass Lepospondyli
  - Order: Nectridea
    - Aistopoda
    - Microsauria
    - Lysorophia
- Subclass Lissamphibia
  - Order: Anura
    - Caudata
    - Apoda

General characteristics. A typical amphibian is characterized by a moist glandular skin, the possession of gills at some point in its life history, four limbs, and an egg lacking the embryonic membrane called the amnion. The closest evolutionary relatives of the amphibians are the fishes, from which they evolved, and the reptiles, to which they gave rise (**Fig. 1**). Present-day amphibians, however, are highly specialized animals, rather different from the primitive forms that probably arose from crossopterygian fishes and far removed from those that gave rise to the earliest reptiles. *See* AMNION; ANAMNIA.

Fish and amphibians. In general, modern amphibians as adults differ from fishes in lacking scales, breathing by means of the skin and lungs instead of gills, and having limbs in place of fins. There are many exceptions to these generalizations, however. Some fishes lack scales, and some members of one group of amphibians, the Apoda, have scales buried in the skin. Some salamanders retain the gills throughout life, and there are air-breathing lungfish. There are no amphibians with paired fins, but the caecilians lack limbs entirely; presumably these animals have evolved from ancestors that possessed typical tetrapod limbs. The significant differences that distinguish typical adult amphibians from fishes are those closely associated with adaptation to life on land. Respiration is by means of skin and lungs rather than

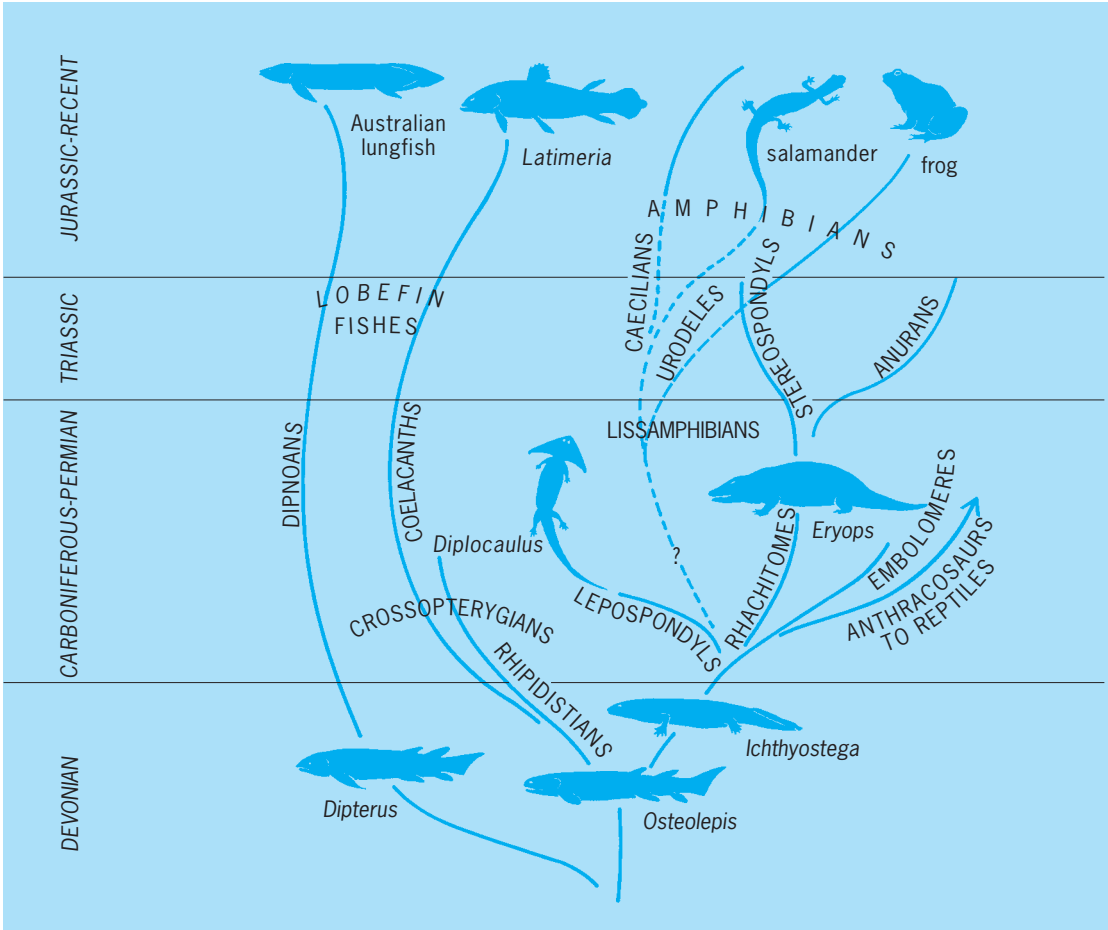

Fig. 1. Vertebrate phylogeny, including tetrapod ancestors, stem reptile groups, and amphibian groups. (*After E. H. Colbert, Evolution of the Vertebrates, Wiley, 1955*)

gills, and tetrapod limbs are for terrestrial locomotion. *See* OSTEICHTHYES; RESPIRATION.

Reptiles and amphibians. Reptiles usually have a dry scaly skin that greatly reduces water loss, and so are very different from the amphibians with their moist skin that permits much evaporation. Young (larval) amphibians have gills, but there is no comparable gill-breathing, larval stage in the life history of a reptile. An important difference between the two groups is the absence of the amniotic egg in the Amphibia and its presence in the Reptilia. The amniotic egg is far more complex than its anamniotic predecessor (represented today by the eggs of fish and modern amphibians). The amniotic egg has three layers: the chorion, the allantois, and the amnion. These layers are involved in protection, nutrient supply, and waste removal, all important functions when moving away from a dependence on the aquatic environment, a predominant condition in all fishes and amphibians today. Lacking these membranes, amphibian eggs must be laid in water or in very moist places. The shell of the reptile egg also makes it more able to resist desiccation, and the eggs can be laid in relatively dry places. The ability to resist water loss through the skin and the development of a land egg are major evolutionary steps for reptiles. *See* REPTILIA.

Ecology. The all-important factor in amphibian life is water. Most species must return to the water to breed, and all must have access to water (even if only in the form of rain or dew) or die of desiccation in a short time. An important consequence of this basic fact of physiology is that vast arid and semiarid areas of the Earth are inhabited by a relatively few specialized amphibians. The majority of amphibian species are found in moist, tropical regions.

Amphibians typically are absent from highly saline waters. A notable exception is the crab-eating frog (*Rana cancrivora*) of southeastern Asia, which can tolerate full-strength seawater. A high level of urea in this frog's blood (as in the blood of sharks and their allies) produces an osmotic pressure high enough to help withstand the otherwise dehydrating effect of salt water.

Physiology. Amphibians are poikilothermic animals; that is, the temperature of the body of an amphibian is largely regulated by the temperature of the surrounding environment. An amphibian in the water will have a temperature close to, if not the same as, that of the water; an amphibian out of the water will often be somewhat cooler than the air as a result of evaporative cooling.

An animal such as an amphibian that burns little of its food energy in keeping warm is able to survive on less food than a bird or mammal of similar size. This advantage is offset by the inability of amphibians to be active under cold conditions. Thus the far northern and southern parts of the world which support large populations of birds and mammals have few amphibian species.

Evolution. The amphibians mark a significant point in the evolution of the vertebrates, the transition from aquatic to terrestrial life. Although restricted to aquatic or at least moist habitats, modern amphibians are an important component of many ecosystems. In the South American rainforest, for example, frogs are a dominant component of the biome. There are as many as 80 or more species of anurans in a small geographic area alone. Also, in certain eastern deciduous forests of the United States, terrestrial salamanders have been calculated to be one of the highest-biomass components of the forest. *See* TETRAPODA; THERMOREGULATION.

W. Ben Cash; Richard G. Zweifel

History and classification. Traditionally, the living vertebrates have been divided into five main groups: fishes, amphibians, reptiles, birds, and mammals. Recent studies have shown, however, that only three of these groups (amphibians, birds, and mammals) are monophyletic, with each including all the descendants of a single common ancestor. Fishes are paraphyletic, as they include the ancestors of all other vertebrate groups. Reptiles are also paraphyletic, as they include the ancestors of birds and mammals. Fortunately, reptiles, birds, and mammals constitute a monophyletic group, the Amniota.

The principal feature which distinguishes living amphibians from amniotes is their mode of reproduction. All amniotes have extraembryonic membranes, including the amnion, which surround the developing embryo within a fluid-filled sac, making it independent of an external aquatic environment. Amphibians, however, still depend on water as the environment into which eggs are normally laid.

Living amphibians, sometimes called the Lissamphibia, are divided into three main groups which are anatomically distinct (Fig. 1). The Caudata (salamanders, including newts) are the least derived with fore- and hindlimbs of equal length and a distinct tail. They are largely aquatic and found mainly in the Northern Hemisphere. The Gymnophionia, apodans and caecilians, are rare, limbless burrowers restricted mainly to the tropics. The Anura, frogs and toads, have large hindlimbs which are used in swimming and hopping; they have well-developed hearing; and many have the ability to produce a characteristic croak. They are truly amphibious and have a worldwide distribution.

Despite these obvious anatomical differences, the three groups of living amphibians are united by the common possession of some unique features: (1) a specialized structure in the inner ear used in hearing, the papilla amphibiorum; (2) green rods in the retina; (3) specialized mucous glands in the skin; (4) specialized glands in the mouth; and (5) bicuspid pedicellate teeth. These features have not been observed in any other living vertebrate, and they are clear evidence that all three groups must have descended from a single common ancestor.

Fossils. The fossil record of the three groups of living amphibians is extensive, and the earliest member of each has been found in Mesozoic rocks. The earliest salamanders are from the Middle Jurassic of England and Kirghizia (Kyrgyzstan), the earliest gymnophionan is from the Lower Jurassic of Arizona;

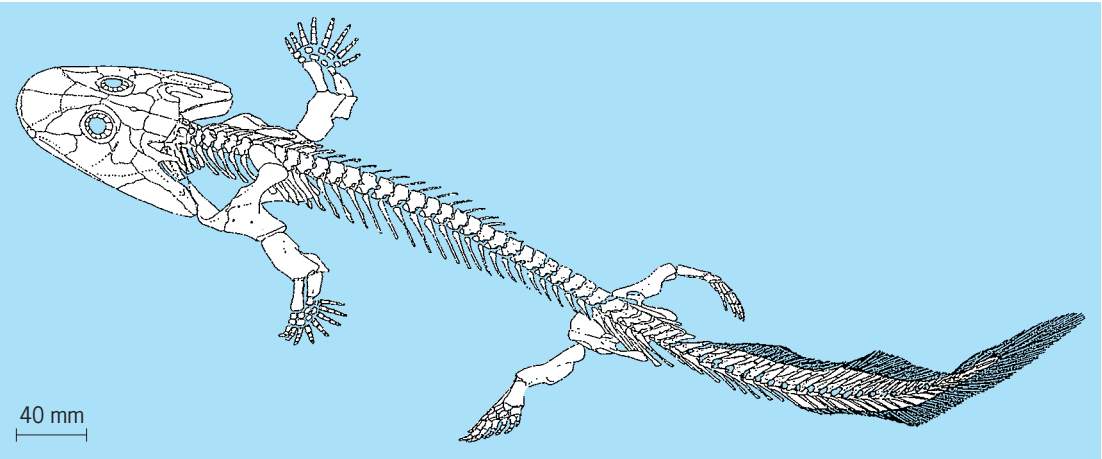

Fig. 2. Reconstruction of the skeleton of the Devonian amphibian *Acanthostega*. (*From M. I. Coates, The Devonian tetrapod Acanthostega gunnari Jarvik: Postcranial anatomy, basal tetrapod interrelationships and patterns of skeletal evolution, Trans. Roy. Soc. Edinburgh Earth Sci., 87:363–421, 1996*)

and the earliest anuran is from the Lower Jurassic of South America. However, no intermediary forms which link the three groups together have been found in the Mesozoic, and it is necessary to look in the Paleozoic, some 100 million years earlier, for the common ancestor of modern amphibians.

The earliest known amphibians have been found in the Upper Devonian rocks of Greenland. Animals such as *Acanthostega* (**Fig. 2**) retain a large number of fishlike features, including internal gills and a caudal fin supported by fin rays. However, they lack an anal fin, and the paired appendages bear digits and form distinct flexible limbs rather than fins. The earliest amphibians had as many as eight digits on each foot, which may have been used as flexible paddles to assist aquatic locomotion. The first amphibians with the characteristic pentadactyl limb, with five digits on each foot, are recorded in the early Carboniferous.

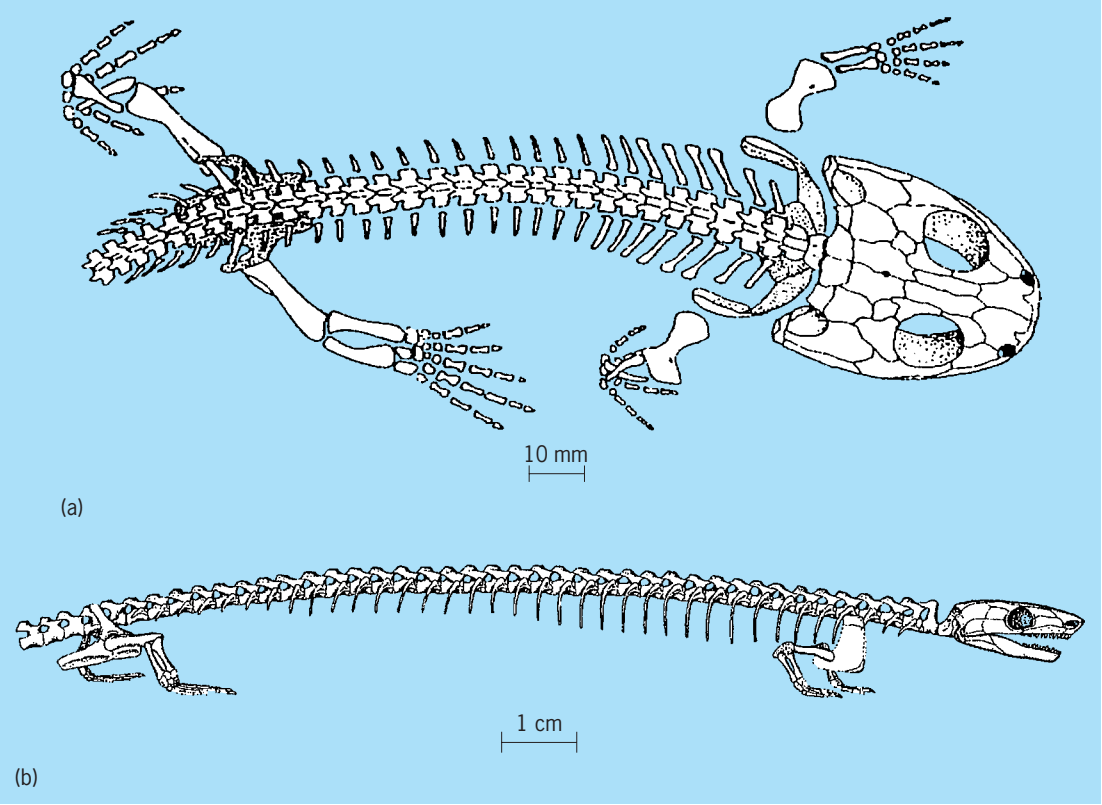

Fig. 3. Reconstructions of amphibian skeletons. (*a*) Carboniferous temnospondyl *Balanerpeton* (*from A. R. Milner and S. E. K. Sequeira, The temnospondyl amphibians from the Visean of East Kirkton, West Lothian, Scotland, Trans. Roy. Soc. Edinburgh Earth Sci., 84:331-362, 1994*). (*b*) Permian microsaur *Rhynchonkos* (*as Goniorhynchus, from R. L. Carroll and P. Gaskill, The Order Microsauria, Mem. Amer. Phil. Soc., 126:1–211, 1978*).

The development of the pentadactyl limb coincided with a rapid radiation of amphibians, and by the end of the Mississippian (325 million years ago) all the main groups had evolved, including those from which both the modern amphibians and the amniotes were to evolve.

Much of the evidence of this rapid radiation is based on single, isolated fossil specimens or is from fossil localities where species diversity is very low. In contrast, the East Kirkton Quarry, a site found recently in the Mississippian of Scotland, has an extensive fauna of terrestrial amphibians, and includes some of the earliest known members of groups of amphibians which were to dominate the later Carboniferous and Permian periods. Among these is the temnospondyl *Balanerpeton* (**Fig. 3**). This is the earliest known member of the group from which all modern amphibians may have evolved. Temnospondyls have a salamanderlike postcranium, with a four-digit manus (hand) and five-digit pes (foot), but a froglike skull with large vacuities in the palate and an otic notch which almost certainly supported a tympanum. Some temnospondyl species also have bicuspid pedicellate teeth. However, the gymnophionans are so different from temnospondyls that a close relationship has been disputed. Instead, a separate origin has been proposed among the microsaurs (Fig. 3). The earliest record of this group is from rocks similar in age to these at East Kirkton, in the Upper Mississippian of Illinois.

This dispute clearly influences the date of the most recent common ancestor of the three groups of modern amphibians. If all three descended from temnospondyls, their most recent common ancestor is possibly represented by the dissorophid temnospondyls of the late Carboniferous and early Permian. However, if the gymnophionans have a separate origin among the microsaurs, then their most recent common ancestor was a much earlier amphibian from the Mississippian, when the microsaurs and temnospondyls had not yet diverged into distinct groups.

Whichever hypothesis is accepted, it is clear that modern amphibians have a very long history extending back almost to the time of the origin and radiation of land vertebrates 340 million years ago. Their unique sensory biology and specialized glands must have evolved at that time and have remained unchanged to the present day. *See* REPTILIA.

T. R. Smithson

Bibliography. W. F. Blair et al., *Vertebrates of the United States*, 2d ed., 1968; J. F. Breen, *Encyclopedia of Reptiles and Amphibians*, 1974; R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1987; W. E. Duellman and L. Trueb, *Biology of Amphibians*, 1986; M. J. Lannoo (ed.), *Amphibian Declines: The Conservation Status of United States Species*, University of California Press, Berkeley, 2005; D. W. Linzey, *Vertebrate Biology*, McGraw-Hill, New York, 2001; A. R. Milner, The Paleozoic relatives of lissamphibians, *Herp. Monogr.*, 7:8–27, 1993; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; F. H. Pough et al., *Herpetology*, 3d ed., Prentice Hall, Upper Saddle River, NJ, 2004; W. D. I. Rolfe, E. N. K. Clarkson, and A. L. Panchen, Volcanism and early terrestrial biotas, *Trans. Roy. Soc. Edinburgh Earth Sci.*, 84:175–464, 1994; A. S. Romer, *Vertebrate Paleontology*, 3d ed., 1966; R. D. Semlitsch (ed.), *Amphibian Conservation*, Smithsonian Institution Press, Washington, DC, 2003; R. C. Stebbins and N. W. Cohen, *A Natural History of Amphibians*, Princeton University Press, 1995; G. R. Zug, L. J. Vitt, and J. P. Caldwell, *Herpetology*, Academic Press, San Diego, 2001 .

Amphibole

A group of common ferromagnesian silicate minerals that occur as major or minor constituents in a wide variety of rocks. The crystal structure of the amphiboles is very flexible and, as a result, the amphiboles show a larger range of chemical composition than any other group of minerals. The structural and chemical complexity of the amphiboles reveals considerable information on the geological processes that have affected the rocks in which they occur. *See* MINERAL; SILICATE MINERALS.

Chemistry. A general formula for amphiboles may be written as $A_{0-1}B_2C_5T_8O_{22}W_2$, where

$$\begin{aligned} A &= \text{Na, K, Ca} \\ B &= \text{Ca, Na, } Mn^{2+}, Fe^{2+}, \text{Li, Mg} \\ C &= \text{Mg, } Fe^{2+}, \text{Al, } Fe^{3+}, Ti^{4+}, \text{Mn, Li} \\ T &= \text{Si, Al, } Ti^{4+} \\ O &= \text{oxygen} \\ W &= \text{OH, F, } O^{2-}, \text{Cl} \end{aligned}$$

and the chemical species are written in order of their importance. Amphiboles are divided into four main groups, according to the type of chemical species in the B group:

Iron-magnesium-manganese amphiboles	$B = (Fe^{2+}, Mg, Mn^{2+}, Li)_2$
Calcic amphiboles	$B = Ca_2$
Sodic-calcic amphiboles	$B = NaCa$
Sodic amphiboles	$B = Na_2$

Iron-magnesium-manganese amphiboles. These amphiboles may be orthorhombic or monoclinic, and the names and formulas of the common amphiboles are as follows:

Orthorhombic amphiboles

Anthophyllite	$(Mg,Fe^{2+})_7Si_8O_{22}(OH)_2$
Gedrite	$Na_{0-1}(Mg, Fe^{2+},Al)_7(Si,Al)_8O_{22}(OH)_2$
Holmquistite	$Li_2(Mg,Fe^{2+})_3(Fe^{3+},Al)_2Si_8O_{22}(OH)_2$

Monoclinic amphiboles

Cummingtonite	$(Mg,Fe^{2+})_7Si_8O_{22}(OH)_2$
Grunerite	$(Fe^{2+},Mg)_7Si_8O_{22}(OH)_2$
Tirodite	$Mn_2^{2+}(Mg,Fe^{2+})_5Si_8O_{22}(OH)_2$
Dannemorite	$Mn_2^{2+}(Fe^{2+},Mg)_5Si_8O_{22}(OH)_2$

Calcic amphiboles. These amphiboles are monoclinic; the more important species are as follows:

Tremolite	$Ca_2Mg_5Si_8O_{22}(OH)_2$
Actinolite	$Ca_2(Mg,Fe^{2+})_5Si_8O_{22}(OH)_2$
Edenite	$NaCa_2Mg_5(Si_7Al)O_{22}(OH)_2$
Pargasite	$NaCa_2(Mg_4Al)(Si_6Al_2)O_{22}(OH)_2$
Hastingsite	$NaCa_2(Fe_4^{2+}Fe^{3+})(Si_6Al_2)O_{22}(OH)_2$
Tschermakite	$Ca_2(Mg_3Al_2)(Si_6Al_2)O_{22}(OH)_2$
Hornblende	$Na_{0-1}Ca_2(Mg,Fe^{2+},Fe^{3+},Al)_5(Si,Al)_8O_{22}(OH,F)_2$
Kaersutite	$NaCa_2(Mg_4Ti)(Si_6Al_2)O_{22}(O,OH)_2$
Cannilloite	$CaCa_2(Mg_4Al)(Si_5Al_3)O_{22}(OH)_2$

Sodic-calcic amphiboles. The sodic-calcic amphiboles are monoclinic; the more important species are as follows:

Richterite	$NaCaNaMg_5Si_8O_{22}(OH)_2$
Winchite	$CaNa(Mg_4Al)Si_8O_{22}(OH)_2$
Barroisite	$CaNa(Mg_3Al_2)(Si_7Al)O_{22}(OH)_2$
Katophorite	$NaCaNa(Mg_4Al)(Si_7Al)O_{22}(OH)_2$
Taramite	$NaCaNa(Fe_3^{2+}Al_2)(Si_6Al_2)O_{22}(OH)_2$

Sodic amphiboles. The sodic amphiboles are monoclinic; the most important species are as follows:

Glaucophane	$Na_2(Mg_3Al_2)Si_8O_{22}(OH)_2$
Riebeckite	$Na_2(Fe_2^{3+}Fe_2^{3+})Si_8O_{22}(OH)_2$
Arfvedsonite	$NaNa_2(Fe_4^{2+}Fe^{3+})Si_8O_{22}(OH)_2$
Ungarettiite	$NaNa_2(Mn_2^{2+}Mn_3^{3+})Si_8O_{22}O_2$
Obertiite	$NaNa_2(Mg_3Fe^{3+}Ti^{4+})Si_8O_{22}O_2$

Other amphiboles. Solid solution is extremely common in amphiboles and gives rise to a large number of additional distinct mineral species. These species are named by attaching prefixes to the names listed above, with the prefix indicating the substituent element differentiating the two minerals. For example, ferropargasite is related to pargasite by substitution of Fe^{2+} for Mg, and ferriwinchite is related to winchite by substitution of Fe^{3+} for Al:

Pargasite	$NaCa_2(Mg_4Al)(Si_6Al_2)O_{22}(OH)_2$
Ferropargasite	$NaCa_2(Fe_4^{2+}Al)(Si_6Al_2)O_{22}(OH)_2$
Winchite	$CaNa(Mg_4Al)Si_8O_{22}(OH)_2$
Ferriwinchite	$CaNa(Mg_4Fe^{3+})Si_8O_{22}(OH)_2$

In this way, a large number of distinct amphibole minerals can be named in a rational way, without generating a large number of unrelated names.

Structure. Amphiboles can have orthorhombic and monoclinic symmetries; these can be distinguished either by x-ray crystallography or by the optical properties of the mineral in polarized light. *See* X-RAY CRYSTALLOGRAPHY.

Monoclinic amphiboles are by far the most common. The characteristic feature of the amphibole structure is the chain of corner-sharing tetrahedrally coordinated groups, labeled T(1) and T(2) [**Fig. 1**].

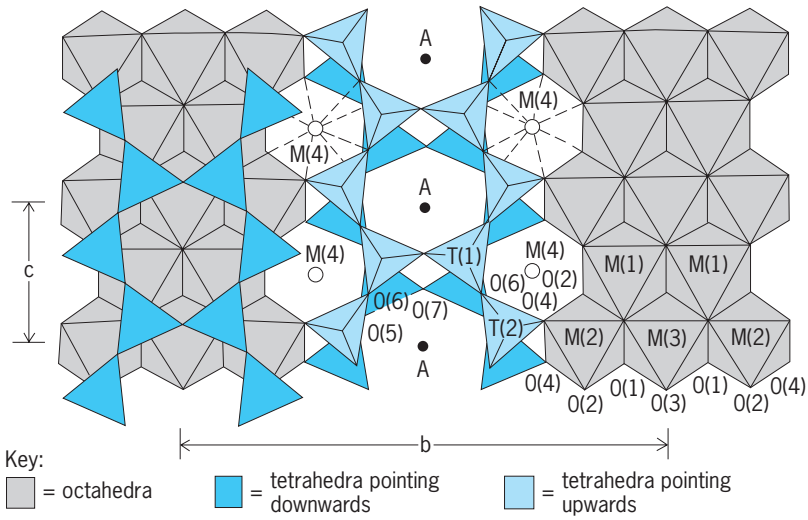

Fig. 1. Crystal structure of monoclinic (*C2/m*) amphibole viewed perpendicular to the tetrahedral double chains; the corner-sharing tetrahedral chains [T(1) and T(2)] and the edge-sharing octahedral strips [M(1) to M(3)] are shaded. O(1) to O(7) = distinct oxygen atoms. (*After F. C. Hawthorne, The crystal chemistry of the amphiboles, Can. Mineral., 21:173–480, 1983*)

These tetrahedra are occupied by the T cations (commonly Si and Al) of the formula, and their repeat distance along the z axis defines the unit-cell length, c, in this direction. The C cations of the formula (commonly Mg, Fe^{2+}, Al, and Fe^{3+}) occupy the M(1), M(2), and M(3) octahedra. These octahedra share edges to form strips that extend along the z axis, and link to the tetrahedral double chains in the x and y directions to form a continuous three-dimensional framework. The B cations occupy the M(4) site (Fig. 1) and are surrounded by eight anions; they provide additional linkage between the tetrahedral double chains and the octahedral strips, and are important in controlling the miscibility and stability behavior of the amphiboles. The A cations occupy the A site, which occurs in a large cavity between adjacent tetrahedral double chains (Fig. 1) and is surrounded by 12 oxygen atoms.

At the center of the octahedral strip is the O(3) site (Fig. 1) that is occupied by the W anions of the chemical formula. This site is bonded to the cations at the adjacent two M(1) sites and one M(3) site, and is usually occupied by monovalent anions (OH^-, F^-, and Cl^-). In the calcic amphibole kaersutite, it has long been recognized that O^{2-} (oxygen) can occur at the O(3) site, associated with the presence of high Fe^{3+} and Ti^{4+} that are characteristic of a highly oxidizing environment. In sodic amphiboles, it has been shown that major amounts of Ti^{4+} can couple to major amounts of O^{2-} at O(3), to the extent that the new anhydrous amphibole species obertiite [with O^{2-} at O(3)] has been recognized.

Inspection of the structure down the y axis (**Fig. 2**) shows that the amphibole structure consists of sheets of tetrahedra interleaved with octahedrally coordinated C-group cations. It is here that the difference between the orthorhombic and monoclinic structures becomes apparent. In the monoclinic structure (Fig. 2*a*), the tetrahedral layers all stack in

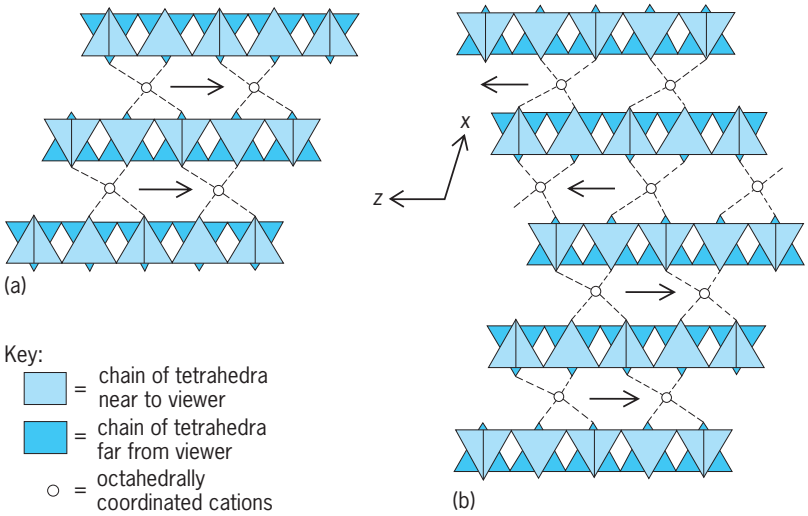

Fig. 2. Crystal structures of (*a*) monoclinic (*C2/m*) and (*b*) orthorhombic (*Pnma*) amphiboles, shown schematically as a stacking of layers. (*After F. C. Hawthorne, The crystal chemistry of the amphiboles, Can. Mineral., 21:173–480, 1983*)

the x direction with the same sense of displacement (along the z direction) relative to the underlying layer, and hence the x axis is inclined to the z axis, producing a monoclinic structure. In the orthorhombic structure (Fig. 2*b*), the displacement of the tetrahedral layers reverses every third layer, and hence the x axis is orthogonal to the z axis, producing an orthorhombic structure. The different stacking of layers in the orthorhombic structure results in a much smaller cavity around the M(4) site. Consequently, the M(4) site can accept only small B-group cations (Mn^{2+}, Fe^{2+}, Mg, and Li), and only iron-magnesium-manganese amphiboles can be orthorhombic. The greater flexibility of the monoclinic structure allows all B-group cations to occupy the M(4) site; as a result, all of the principal chemical groups of amphiboles can have the monoclinic structure. *See* CRYSTAL; CRYSTALLOGRAPHY; CRYSTAL STRUCTURE; SYMMETRY.

Immiscibility and exsolution. Amphiboles do not show the complete range of possible compositions suggested by the general chemical formula and common idealized compositions described above. In particular, there is not a continuous range of chemical composition between the four main amphibole groups: the iron-magnesium-manganese amphiboles, the calcic amphiboles, the sodic-calcic amphiboles, and the alkali amphiboles. This lack of so-called solid solution is a result of the structure not being able to accommodate two types of cations (positively charged atoms) of very different size (or charge) at the same set of sites in the structure of a single crystal. This effect is very prominent between the iron-magnesium-manganese amphiboles and the calcic amphiboles, in which the M(4) site (Fig. 1*a*) is occupied by the B-group cations (Fe^{2+}, Mg, Mn) and Ca, respectively. There is a wide range of chemical compositions common in simple calcic and iron-magnesium-manganese amphiboles (**Fig. 3**). For some chemical compositions (within the shaded areas of Fig. 3), a single amphibole crystallizes. For other chemical compositions (the unshaded areas), a single amphibole is not stable. Instead, two coexisting amphiboles form: a calcic amphibole and an iron-magnesium-manganese amphibole. Thus actinolite and cummingtonite-grunerite are immiscible.

The degree to which two amphiboles are immiscible often varies as a function of temperature and pressure; at high temperatures or pressures, miscibility is usually enhanced. The immiscible region, which is known as a miscibility gap, is narrow at high temperature but widens at lower temperature. When the amphibole composition is within the miscibility gap, a single amphibole is no longer stable. It is here that the process of exsolution (or unmixing) occurs. Actinolite exsolves cummingtonite-grunerite to form lamellae of cummingtonite within the actinolite. Coexisting amphiboles and exsolution textures are very informative about temperatures of crystallization and cooling history, particularly in metamorphic rocks. *See* PHASE EQUILIBRIUM; SOLID SOLUTION; SOLID-STATE CHEMISTRY.

Occurrence. Amphiboles are common minerals in many types of igneous rocks, and the composition of the amphibole reflects the silica content of the rock. The high melting temperature of many rocks often precludes crystallization of amphiboles from the melt. However, high-temperature minerals such as pyroxenes often react with hydrothermal fluids during cooling to produce secondary amphiboles. At higher temperatures such secondary amphiboles can form an integral part of the original rock, whereas at lower temperatures they often occur as veins infilling fractures formed during cooling. Calcic amphiboles, particularly pargasite, are characteristic of ultramafic and metabasaltic rocks and are usually quite rich in magnesium. Titanium-rich hornblendes and kaersutites occur in intermediate rocks and are often strongly oxidized—an unusual feature in amphiboles from any other environment. Acidic rocks, particularly granites, contain a wide range of amphiboles, from hastingsite to riebeckite and arfvedsonite; these are often iron-rich and can contain significant amounts of more unusual elements such as Li, Zn, and Mn. Understanding the variation in amphibole composition as a function of the progressive crystallization of a suite of rocks is important for understanding the petrological processes in an igneous environment.

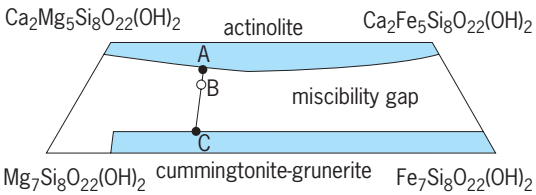

Fig. 3. Compositional variations in calcic and iron-magnesium-manganese amphiboles at low temperature. Shaded areas denote compositional regions within which single amphiboles are stable; compositions within the unshaded area (for example, composition B) cannot form a single amphibole but form two coexisting amphiboles (compositions A and C); composition A represents a calcic amphibole.

Amphiboles usually weather very easily and hence are not important in sedimentary rocks, although they can be significant components of soil. Iron-rich alkali amphiboles can form at essentially ambient conditions in the sedimentary environment, but this occurrence is rare. Amphiboles are common and important rock-forming minerals in many types of metamorphic rocks. They are particularly abundant in rocks of basaltic composition at most grades of metamorphism. In the greenschist facies, the first amphiboles to form are tremolite-actinolite at low to high pressures. With increasing grade, the amphibole composition becomes more complex, changing in the general direction of pargasite. The specific compositional changes are sensitive to differences in temperature and pressure, and are useful in following the detailed history of metamorphic rocks. Blueschists are characteristic of very high pressures and relatively low temperatures. At low grade, riebeckite and crossite form, giving the blueschists their characteristic color. With increasing temperature and pressure, the amphibole composition evolves toward glaucophane, which can coexist with actinolite in transitional facies. At very high pressures and temperatures, the miscibility gap between the alkali and calcic amphiboles disappears with the crystallization of sodic-calcic amphiboles such as winchite. Understanding these processes allows derivation of the tectonic history of a rock from the compositional variations recorded in its constituent amphiboles. *See* BASALT; BLUESCHIST; FACIES (GEOLOGY); GRANITE; IGNEOUS ROCK; METAMORPHIC ROCK; METAMORPHISM; SCHIST.

Industrial uses. Amphiboles are economically important as commercial asbestos minerals and as semiprecious gem materials. World asbestos production is dominated by the serpentine-group mineral chrysotile; but the amphibole minerals anthophyllite, cummingtonite-grunerite (amosite), actinolite, and riebeckite (crocidolite) are also important, particularly in Australia and South Africa. It is now well established that the mineralogical characteristics of asbestos have a major influence on its degree of carcinogenicity; amphibole asbestos seems to be far more harmful in this respect than chrysotile. *See* ASBESTOS.

Some amphiboles with attractive physical properties are marketed as semiprecious gem material. Most important is nephrite, a dense compact form of fibrous actinolite that is a principal variety of jade. Fibrous riebeckite is marketed as one of the less common varieties of tiger's eye. Iridescent gedrite and gem-quality pargasite are used as semiprecious gems in contemporary jewelry. *See* GEM; JADE; MINERALOGY.

Frank C. Hawthorne

Bibliography. W. A. Deer et al., *An Introduction to the Rock-Forming Minerals*, 2d ed., 1992; F. C. Hawthorne, The crystal chemistry of the amphiboles, *Can. Mineral.*, 21:173–480, 1983; F. C. Hawthorne et al., The role of Ti in hydrogen-deficient amphiboles, *Can. Mineral.*, 36:1253–1265, 1998; B. E. Leake et al., Nomenclature of amphiboles: Report of the Subcommittee on Amphiboles of the International Mineralogical Association Commission on New Minerals and Mineral Names, *Can. Mineral.*, 35:219–246, 1997; D. R. Veblen (ed.), *Reviews in Mineralogy*, vol. 9A: *Amphiboles and Other Hydrous Pyriboles: Mineralogy*, 1981; D. R. Veblen and P. H. Ribbe (eds.), *Reviews in Mineralogy*, vol. 9B: *Amphiboles: Petrology and Experimental Phase Relations*, 1982.

Amphibolite

A class of metamorphic rocks with one of the amphibole minerals as the dominant constituent. Most of the amphibolites are dark green to black crystalline rocks that occur as extensive layers widely distributed in mountain belts and deeply eroded shield areas of the continental crust. Amphibolite is the main country rock that has been intruded by the large granite masses found in most mountain ranges, with small and large masses of amphibolite present also as inclusions in granites. *See* AMPHIBOLE.

Amphibolites are the products of regional metamorphism and crustal deformation of older materials of appropriate composition. The deformation that accompanied the formation of the amphibolites is expressed by the segregation of the minerals into layers and the orientation of the needles and prisms of the amphibole crystals along a direction in the layer. The temperature range of formation is roughly 400–500°C (750–930°F) of the epidote-amphibolite and amphibolite metamorphic facies. The amphibolites are of remarkably simple mineralogies with most consisting of only common hornblende and plagioclase, although epidote, chlorite, diopside, garnet, biotite, magnetite, sphene, and quartz may be present as minor minerals.

The features of the original rock are obliterated; thus it is difficult and sometimes impossible to determine the premetamorphic rock. The bulk composition of the amphibolites could result by the addition of water and recrystallization of original intrusive masses of gabbro, diabase, or piles of basaltic lava flows and, if so, are referred to as ortho amphibolites. The bulk composition could result by the dehydration and recrystallization of lower-temperature-of-formation serpentine and chlorite schists, which also could represent altered original gabbro or basaltic lavas, but which also could be original clay-rich marine sediments (and then the rocks are referred to as para amphibolites). Compositional relations between the minor elements titanium, chromium, and nickel have been used to distinguish the ortho from para amphibolites in some occurrences.

Amphibolite layers commonly are associated with layers of various biotite schists, granite gneisses, quartzite, and marble masses that formed during the same metamorphic episode. Extensive folding, shearing, and plastic deformation such as pinch and swell structures of the amphibolite layers are characteristic features of most occurrences. Multiple episodes of deformation and recrystallization are

common in many localities, and imply a long and complex history of evolution of the rock mass.

The amphibolites are seldom sites of ore deposits, although mineralized quartz veins that crosscut the amphibolites occur in some localities.

The plate tectonic model of crustal evolution interprets the large masses of amphibolites as the alterations of basaltic crustal materials in subduction zones. *See* METAMORPHIC ROCKS; SCHIST.

George De Vore

Amphidiscosa

An order of the subclass Amphidiscophora in the class Hexactinellida. These sponges are distinguished from the order Hemidiscosa in that the birotulates are amphidiscs. Examples of this order are *Pheronema, Monorhaphis*, and *Hyalonema*. *See* HEMIDISCOSA.

The record of fossil Amphidiscosa is poor, but goes back to the Carboniferous *Uralonema*.

Willard D. Hartman; Robert M. Finks

Amphilinidea

An order of tapeworms of the subclass Cestodaria. These worms have a protrusible proboscis and frontal glands at the anterior end. No holdfast organ is evident. All members of the order inhabit the coelom of sturgeon and other fish. The only life history which is completely known is that of *Amphilina* (**illus.** *a*). The 10-hooked embryos leave the parental uterus through a pore, and if upon escaping into the water they are eaten by an amphipod crustacean, they undergo further development to the procercoid larva (illus. *b*). When the parasitized amphipod is eaten by a sturgeon, the larval worm enters the coelom of the fish and develops to sexual maturity.

Clark P. Read

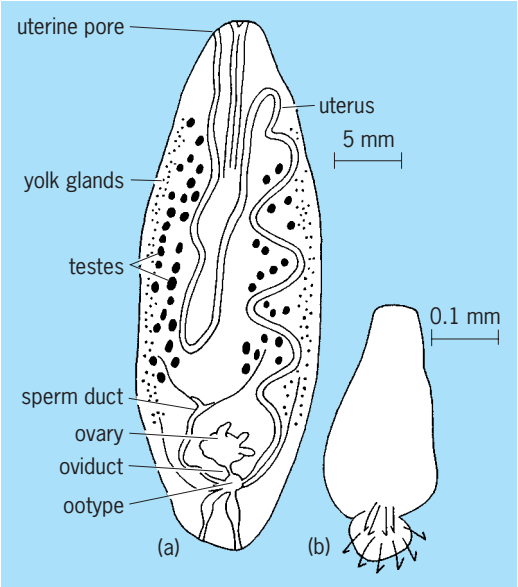

***Amphilina*. (*a*) Adult. (*b*) Procercoid larva.**

Amphionidacea

An order of the Eucarida comprising a single species, *Amphionides reynaydii* (see **illus.**). Because of the similarity of its early larval stages to those found among the caridean shrimps, for more than a century it was classified as an aberrant member of the Caridea. However, the fact that the pleuron of the second abdominal somite never overlaps that of the first in any stage immediately distinguishes this species from all true Caridea.

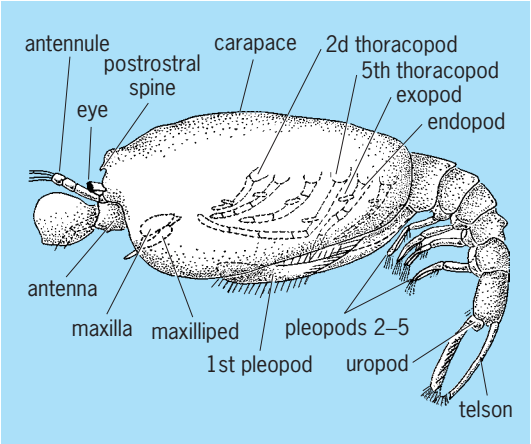

Amphionides reynaudii. (After P. A. McLaughlin, Comparative Morphology of Recent Crustacea, W. H. Freeman, 1980)

The cephalothorax is covered by a large but very thin carapace. Only the first pair of thoracic appendages is modified as maxillipeds. The following six pairs of limbs are biramous, with the outer rami short and virtually functionless; the inner rami are segmented and somewhat sticklike. There are no appendages on the last thoracic somite in females, but a four-segmented uniramous pair of appendages has been seen in juvenile males. The eyestalk has a large tubercle on the inner side, which has been postulated to be a photophore used for attracting prey. In addition to the single pair of maxillipeds, the pronounced development of the first pair of pleopods in adult females clearly sets *Amphionides* apart from other eucarids. When extended, these pleopods close off the ventral carapace region and appear to form a unique type of brood pouch beneath the thorax. Although no brooding females have ever been captured, the fact that the genital pores on the coxae of the sixth thoracopods open into this chamber gives support to the brood pouch concept. The very long fifth thoracopods could be used for cleaning and egg tending.

Amphionides may have up to 12 or 13 free-swimming larval stages followed by a metamorphic

molt to the juvenile form. It is these larval stages that are most frequently captured in plankton samples, and in the nineteenth century several of the larval stages were described as distinct species. Adult males have never been collected; however, sexual dimorphism in the last juvenile stage has been described. It is presumed that adult males are much more efficient swimmers, and thus able to avoid capture. It is also possible that their only function is one of reproduction, and that adult males are relatively short-lived.

Amphionides reynaudii has worldwide distribution, primarily in equatorial regions. Adult females have been collected at depths between 21,000 and 111,000 ft (700 and 3700 m), while most larvae occur in depths of only 90–300 ft (30–100 m). *See* EUCARIDA.

Patsy A. McLaughlin

Bibliography. P. Heegaard, *Larvae of Decapod Crustacea: The Amphionidae, Dana*, Rep. 77, pp. 1–67, 1969; D. I. Williamson, *Amphionides reynaudii* (H. Milne Edwards), representative of a proposed new order of eucaridan Malacostraca, *Crustaceana*, 25:35–50, 1973.

Amphipoda

An order of crustaceans in the subclass Malacostraca, which lack a carapace, bear unstalked eyes, and respire by thoracic branchiae, or gills. The abdomen usually bears three pairs of biramous swimmerets (pleopods), three pairs of rather rigid uropods, and a telson which may be lobed or entire. The body is usually flattened laterally, and the pereiopods (walking legs), unlike those of the Isopoda, are elongated so that walking is difficult. In contrast to the Isopoda, the maxillipeds lack epipodites. The sexes are separate, but reproductive and copulatory organs are very simple. The eggs are extruded by the female into a ventral brood pouch composed of setose lamellae attached to the medial bases of the legs. The young hatch as miniature adults, growing usually to a length of 0.12–0.48 in. (3–12 mm), and in exceptional cases to 5.6 in. (140 mm). *See* ISOPODA.

Amphipods are very abundant in the oceans, being represented by 3200 species. More than 600 other species occur in streams, lakes and subterranean waters and in terrestrial leaf molds and mosses. Many are excellent swimmers. Nonpelagic species of the suborders Gammaridea and Caprellidea live on aquatic bottoms, plants, and epifaunal growths. Predation by amphipods is occasional. Their mouthparts are well adapted for chewing (**Fig. 1**); they either eat aquatic plants, debris, and detritus or swallow mud containing food particles. A few are filter feeders; others feed by sucking animal tissues.

Four suborders are known, the Gammaridea, Hyperiidea, Caprellidea, and Ingolfiellidea.

Gammaridea. The pleon, or abdomen, is well developed and the maxilliped bears a palp (**Fig. 2***a*). This is the largest suborder, comprising 3200 species. They are principally marine, but the group contains

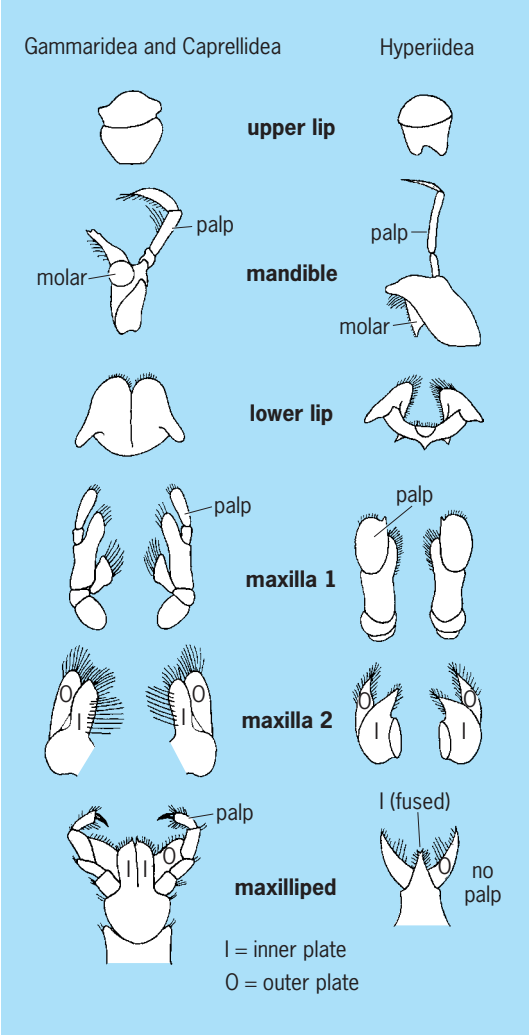

Fig. 1. Generalized mouthparts of Amphipoda. (*After G. O. Sars, An Account of the Crustacea of Norway, vol. 1, 1895*)

the only nonmarine forms, of which 500 are limnetic and 80 terrestrial. About 200 of the marine species are pelagic, the remainder being benthic or intertidal. Gammarideans have been discovered in the deepest oceanic trenches and at altitudes of 13,000 ft (4000 m) on tropical islands.

Hyperiidea. The abdomen is well developed (Fig. 2*b*), but the maxilliped lacks a palp. The suborder comprises about 300 species which are entirely marine and pelagic. Adaptations of pelagic life include suspensory mechanisms such as oil storage, broadened body surfaces, and elongated appendages. The eyes are conspicuous, because of the proliferation of ommatidia to cover the entire cephalic area (Fig. 2*b*) or to their absence in bathypelagic species. Some hyperiids spend much of their lives encased within medusae and salps.

Caprellidea. The abdomen is vestigial, and the maxilliped bears a palp. The body is very slender, resulting in the name "skeleton shrimp" (Fig. 2*c*). Caprellids are entirely marine. This group contains

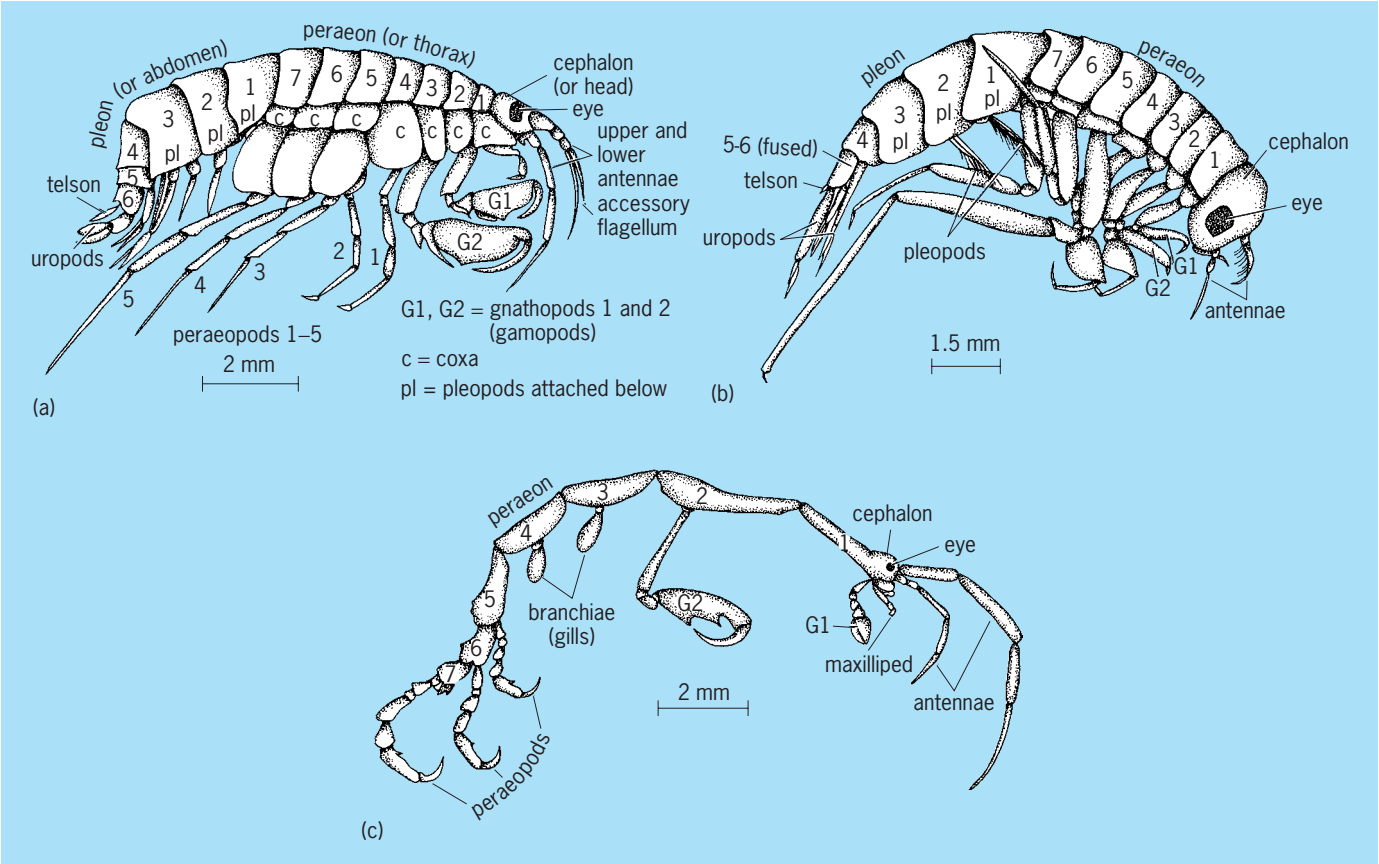

Fig. 2. Examples of the suborders of the Amphipoda. (*a*) Gammaridea, *Lilljeborgia pallida,* male; (*b*) Hyperiidea, *Parathemisto bispinosa*, female; (*c*) Caprellidea, *Caprella linearis*, male. (*After G. O. Sars, An Account of the Crustacea of Norway, vol. 1, 1895*)

200 species of skeleton shrimps and about 30 species of whale lice, or cyamids. The cyamids are dorsoventrally flattened animals, which are epibionts in external orifices of cetaceans (**Fig. 3**). The first thoracic segment bearing the first gnathopod is usually fused solidly to the head in caprellids, leaving only six free thoracic segments. Most species lack the first two pairs of pereiopods, while a few bear remnants of these legs. Branchial lamellae occur on two or three free segments, and female brood lamellae occupy the second and third free segments. Caprellids commonly occur on epifaunal growths in shallow water; some of them feed on hydrozoan coelenterates. A behavioral motion similar to that of the praying mantids (insects) has been described for them.

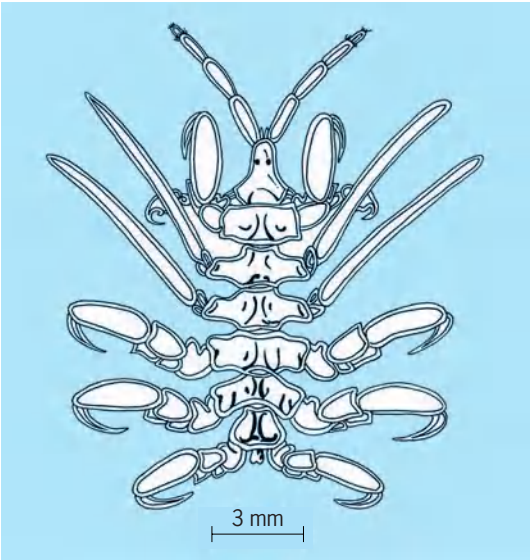

Fig. 3. *Cyanis boopis*, male; a whale louse (Caprellidea) which is modified as an epibiont on cetaceans. (*After G. O. Sars, An Account of the Crustacea of Norway, vol. 1, 1895*)

Ingolfiellidea. Both abdomen and maxilliped are well developed as in the Gammaridea, but the head often bears a separate ocular lobe lacking eyes. The pleopods are simple and leaflike, and the body is thin, resembling a tanaid. The dactyls of the gnathopods are composed of two articles. The suborder comprises a dozen species, in the abyssal sea, in the shallow sea, and in subterranean localities.

Economic importance. Marine species are important food for various stages of many commercial fishes. Hyperiids are the principal food of seals at certain seasons, and also of balaenoid whales at times. One gammaridean genus, *Chelura*, is a minor wood borer, associated with the isopod *Limnoria*.

A few fossil species are known in Tertiary amber deposits. *See* HYPERIIDEA; MALACOSTRACA; MARINE ECOLOGY.

J. Laurens Barnard

Amplifier

A device capable of increasing the magnitude of a physical quantity. This article discusses electronic amplifiers whose operation depends on transistors. Some amplifiers are magnetic, while others may take the form of rotating electrical machinery. Forms of nonelectrical amplifiers are hydraulic actuators and levers which are amplifiers of mechanical forces. *See* DIRECT-CURRENT MOTOR; HYDRAULIC ACTUATOR; LEVER.

The operation of an amplifier can be explained with a model (**Fig. 1**). A controlled voltage source of

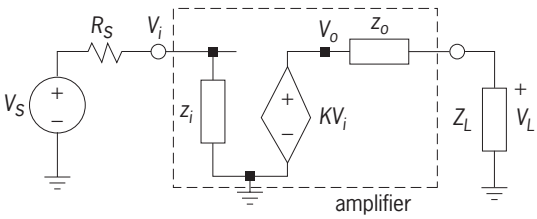

Fig. 1. Amplifier model with source, load, and input and output impedances.

gain K generates an output voltage $V_o = KV_i$ from an input voltage V_i. This input voltage is obtained from a source voltage V_S with source resistance R_S via voltage division with the amplifier's input impedance z_i. The load voltage V_L across the load impedance Z_L is obtained from V_o by voltage division with the amplifier's output impedance z_o. The input voltage and load voltage are given by Eqs. (1), where k_i

$$V_i = \frac{z_i}{z_i + R_S} V_S = k_i V_S$$
$$V_L = \frac{Z_L}{Z_L + z_o} V_o = k_o V_o \tag{1}$$

and k_o, respectively, express the effects of the amplifier loading the source and of the load impedance loading the amplifier. The impedances z_i and z_o mostly consist of a resistor in parallel with a capacitor; often they may be assumed to be purely resistive: $z_i = r_i$ and $z_o = r_o$. From Eq. (1), the amplifier's operation is given by Eq. (2). Thus, the amplification

$$V_L = k_i K k_o V_S = \frac{z_i}{z_i + R_S} K \frac{Z_L}{Z_L + z_o} V_S \tag{2}$$

is decreased from the ideal value K by the two load factors k_i and k_o. The reduction in gain is avoided if the two factors equal unity, that is, if $z_i = \infty$ and $z_o = 0$. Thus, in addition to the required gain K, a good amplifier has a very large input impedance z_i and a very small output impedance z_o. *See* GAIN.

Operational Amplifiers

The operational amplifier (op amp) is a commonly used general-purpose amplifier. It is implemented as an integrated circuit on a semiconductor chip, and functions as a voltage amplifier whose essential characteristics at low frequencies are very high voltage amplification, very high input resistance, and very low output resistance. *See* INTEGRATED CIRCUITS; VOLTAGE AMPLIFIER.

The operational amplifier has three signal terminals (**Fig. 2***a*): an inverting and a noninverting input and one output. There are also two terminals to apply dc power, which are normally not drawn in circuit diagrams. Further terminals may be present for other functions, such as offset control. Because the electronic circuitry inside the operational amplifier does not permit excursions of the output voltage to exceed the supply voltages of either polarity, both positive and negative supply voltages, $\pm V_B$, with respect to ground are required if both positive and negative excursions of output voltage with respect to ground are desired. All signal voltages are referenced to ground.

If the voltages at the inverting and noninverting input terminals are called V^- and V^+, respectively, the operational amplifier input voltage V_{in} equals $V^+ - V^-$, and the output voltage V_o is given by Eq. (3).

$$V_o = AV_{\text{in}} = A(V^+ - V^-) \tag{3}$$

Because the magnitude of the output voltage V_o cannot become larger than the power supply voltage V_B, and the operational amplifier gain A is very large, the magnitude of the voltage V_{in} across the input terminals of an op amp is very small, being limited as in Eq. (4). With $A = 10^5$ and, typically, $V_B = 12$

$$V_{\text{in}} = V^+ - V^- = \frac{V_o}{A} < \frac{V_B}{A} \tag{4}$$

V, it follows that $V_{\text{in}} = 12 \times 10^{-5}$ V $= 120\ \mu$V. This voltage is normally much smaller than any other signal in the circuit, and V_{in} is therefore approximated as zero. Expressed differently, because the gain of the operational amplifier is so large, even very small

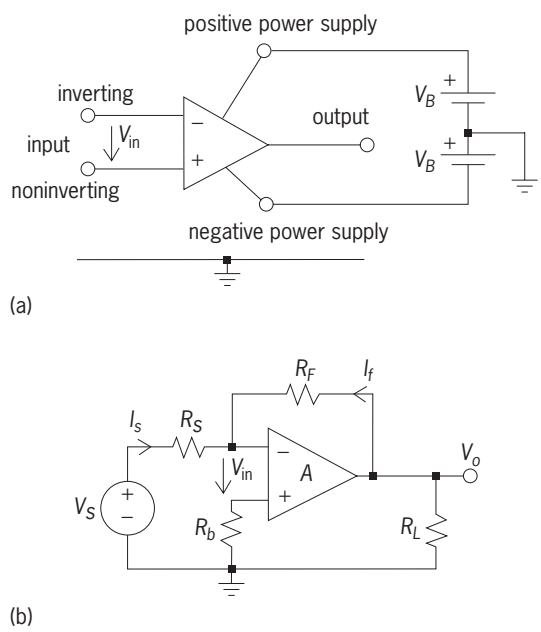

Fig. 2. Operational amplifier. (*a*) Circuit symbol, showing signal terminals and power supply terminals. (*b*) Simple operational amplifier circuit to amplify the signal voltage V_S by the gain $-R_F/R_S$, and apply it to load resistor R_L.

input voltages would be amplified to levels larger than the power supplies permit. Thus, to obtain practical amplifiers with gain less than $A = V_o/V_{in}$, operational amplifiers are always used with negative feedback (Fig. 2*b*). Feedback also reduces distortion and increases gain stability by the same factor as the reduction in voltage gain. *See* ANALOG COMPUTER; DISTORTION (ELECTRONIC CIRCUITS); FEEDBACK CIRCUIT; OPERATIONAL AMPLIFIER.

Transconductance Amplifiers

The transconductance amplifier has become widely used. In contrast to operational amplifiers, which convert an input voltage to an output voltage, transconductors are voltage-to-current converters described by the transconductance parameter g_m, which satisfies Eq. (5). Thus, the output current I_{out}

$$I_{out} = g_m V_{in} \tag{5}$$

is proportional to the input voltage V_{in}. Whereas operational amplifier circuits are mostly found with single-ended output terminals (Fig. 2*a*), transconductors have either single-ended (**Fig. 3***a*) or (mostly)

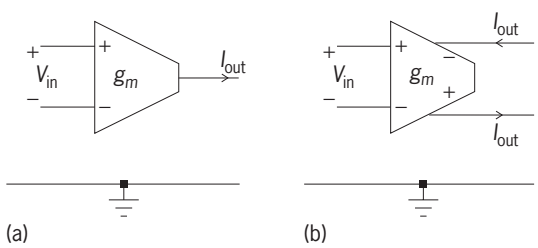

Fig. 3. Circuit symbols for transconductance amplifiers with (*a*) single-ended, output and (*b*) differential output.

differential outputs (Fig. 3*b*). As do operational amplifiers, ideal transconductors have an infinite input resistance, but they also have an infinite output resistance so that the output is an ideal current source (see **table**). One of the attractive properties of transconductance amplifiers is their wide bandwidth. Very simple transconductance circuits can be designed which maintain their nominal g_m values over bandwidths of several hundred megahertz (10^8 Hz), whereas operational amplifiers often have high gain only over a frequency range of less than 100 Hz, after which the gain falls off rapidly. Consequently, in high-frequency communications applications, circuits built with transconductance amplifiers

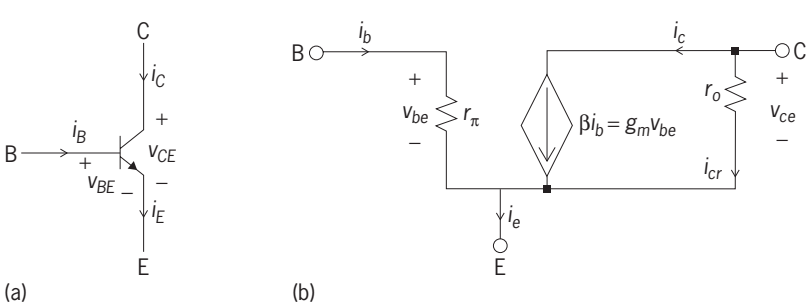

Fig. 4. An *npn* transistor. (*a*) Circuit symbol. (*b*) Small-signal model.

Transconductance amplifier parameters

Parameter	Symbol	Typical value
Transconductance	g_m	10–1000 μA/V
Input resistance	r_i	1,000,000 Ω
Output resistance	r_o	100,000 Ω

generally give much better performance than those with operational amplifiers.

Bipolar Junction Transistor Amplifiers

A bipolar junction transistor has three terminals named the base (*B*), emitter (*E*), and collector (*C*) (**Fig. 4***a*). It consists of two *pn* junctions, the emitter-base junction, and the collector-base junction. In normal amplifier operation, the emitter-base junction is forward biased and the collector-base junction is reverse biased; that is, for an *npn* transistor (Fig. 4*a*), the base voltage is more positive than the emitter voltage by $V_{BE} \approx 0.65$ V, and the collector is more positive than the base. These polarities are inverted in a *pnp* transistor.

Because it is necessary to distinguish between constant currents and voltages (dc or bias conditions), time-varying signals (ac), and the total quantities in the analysis of electronic circuits, the following standard notation is used: dc is denoted as uppercase sub uppercase, ac is lowercase sub lowercase, and the total quantity is lowercase sub uppercase. Thus, the total collector current i_C, consisting of the sum of the dc bias current I_C and the time-varying ac signal current i_c, would be written as Eq. (6). Other

$$i_C(t) = I_C + i_c(t) \tag{6}$$

quantities are expressed similarly; for example, $v_{BE} = V_{BE} + v_{be}$ for the base-emitter voltage.

Transistor operation. The operation of the bipolar transistor can be described by the simple equations (7) and (8), where V_T is the strongly

$$i_C \cong I_S[\exp(v_{BE}/V_T) - 1] \approx I_S \exp(v_{BE}/V_T) \tag{7}$$

$$i_B = \frac{i_C}{\beta} \qquad i_E = \frac{i_C}{\alpha} \tag{8}$$

temperature-dependent thermal voltage (≈ 0.026 V at room temperature), and the approximation in Eq. (7) is valid in the forward-biased region, where $v_{BE} \approx 0.65$ V is much larger than V_T. The quantity I_S is the reverse saturation current (the current that flows in the collector when $v_{BE} < 0$; a few nanoamperes), β is the base-to-collector current gain, and α is the emitter-to-collector current gain. It follows from the transistor configuration (Fig. 4*a*) that $i_E = i_C + i_B$, and this result, along with Eq. (8), leads to Eqs. (9).

$$\alpha = \frac{\beta}{\beta + 1} \qquad \beta = \frac{\alpha}{1 - \alpha} \tag{9}$$

Typical numbers are $\alpha \approx 0.99$ and $\beta \approx 100$. Thus, i_B is much smaller than both i_C and i_E, so that $i_C \approx i_E$. *See* TRANSISTOR.

Linear performance. To achieve linear amplifier performance for small ac signals, the highly nonlinear bipolar junction transistor is operated at a dc bias point (I_C, V_{CE}) by choosing a dc base current I_B as discussed below. The ac signals are then superimposed on the dc bias points. For small-signal operation, the bipolar junction transistor can be modeled by a simple linear circuit (Fig. 4*b*). To derive this circuit, the small-signal approximation, $v_{be} \ll V_T$, is used with Eq. (7). The small-signal current i_c is found to be linearly related to the small-signal voltage v_{be} by Eq. (10), where the transconductance parameter g_m, given by Eq. (11), is proportional to the dc bias

$$i_c = g_m v_{be} = \beta i_b \tag{10}$$

$$g_m = \frac{I_C}{V_T} \tag{11}$$

current I_C that is selected by the designer. The second equality in Eq. (10) follows from Eq. (8) Dividing both sides of Eq. (10) by $i_b g_m$ yields Eq. (12), where

$$r_\pi = \frac{v_{be}}{i_b} = \frac{\beta}{g_m} \tag{12}$$

r_π is the small-signal input resistance of the transistor. In some amplifier configurations the input resistance r_e, given by Eq. (13), is also of interest.

$$r_e = \frac{v_{be}}{i_e} = \frac{\alpha}{g_m} \approx \frac{1}{g_m} \tag{13}$$

An additional component in the small-signal circuit is the output resistor r_o. It cannot be obtained from Eqs. (7) and (8) but arises from the Early effect, which causes the collector current to increase with increasing collector-emitter voltage. Thus, a current i_{cr} is added to i_c such that Eq. (14) is satisfied.

$$r_o = \frac{v_{ce}}{i_{cr}} \tag{14}$$

With the small-signal transistor parameters identified, the gain and input and output resistances of bipolar junction transistor amplifiers can now be readily derived. First, however, it is necessary to deal with the problem of bias design, which will be explained on the case of a common-emitter amplifier. *See* BIAS (ELECTRONICS).

Common-emitter amplifier. Because bipolar junction transistor parameters are strong functions of temperature and β varies widely from device to device, amplifier performance must be made fairly independent of the transistor parameters and the bias circuitry must be designed such that the collector bias current I_C is constant (**Fig. 5***a*). The bias circuitry consists of four resistors, labeled R_1, R_2, R_E, and R_C, and a battery, which provides voltage V_{CC}. Two coupling capacitors, labeled C_C, conduct no dc current and serve to isolate the dc bias quantities from the load resistor R_L and the source v_S with source resistor R_S. The resistor R_E, between the emitter terminal and ground, provides feedback between the output and the input of the amplifier. This feedback causes the desired stabilization of the dc operating point as discussed below, but also reduces the gain. Thus, to achieve a stable operating point, R_E must be included in the dc circuit but removed from the ac circuit. This step is accomplished by a bypass capacitor C_B which shunts ac currents away from R_E and places the emitter essentially at ac ground. It makes the emitter terminal common to both input and output and leads to the name common-emitter amplifier.

The base bias current I_B is derived by replacing the bias circuit seen from the base by a Thévenin equivalent: The voltage V_B is obtained by voltage division as Eq. (15), and the Thévenin resistance at

$$V_B = \frac{R_2}{R_1 + R_2} V_{CC} \tag{15}$$

the base equals the parallel connection of R_1 and R_2, given by Eq. (16). For sufficiently large values of β,

$$R_B = \frac{R_1 R_2}{R_1 + R_2} \tag{16}$$

the bias point is given approximately by Eqs. (17) and (18). Thus, the bias point is almost independent

$$I_C = \frac{(V_B - V_{BE})}{R_E} \tag{17}$$

$$V_{CE} \approx V_{CC} - I_C(R_C + R_E) \tag{18}$$

of the transistor parameters such as β, and is essentially fixed by the choice of V_{CC} and of the resistors. *See* THÉVENIN'S THEOREM (ELECTRIC NETWORKS).

The operating point is the dc bias point on a graph of I_C versus V_{CE} (Fig. 5*b*). It is specified by the intersection of the straight load line, given by Eq. (18), and the transistor characteristics given by the manufacturer. The operating point is chosen approximately in the center of the linear range, at a suitable current I_C and at $V_{CE} \approx V_{CC}/2$. Placing the operating point at $V_{CC}/2$ permits the largest possible undistorted signal amplitude of approximately $V_{CC}/2$ at the output. The dc voltage across R_E, V_E, is chosen approximately as $V_E \approx 0.2\ V_{CC}$.

With the operating point set as prescribed, the transistor can be replaced in the common-emitter amplifier (Fig. 5*a*) by its small-signal model (Fig. 4*b*) to compute the gain of the amplifier and its input and output resistances. In the resulting small-signal circuit (Fig. 5*c*), all capacitor values are assumed to be infinite, that is, the capacitors are short circuits for ac signals. Furthermore, the power supply V_{CC} is ac signal ground.

The small-signal circuit shows that the input resistance seen by the source is R_B in parallel with r_π; that is, with Eqs. (11) and (12) and the chosen bias conditions, it is given by Eq. (19). (The resistor R_B

$$r_{\text{in}} = \frac{r_\pi R_B}{r_\pi + R_B} \approx r_\pi \tag{19}$$

is usually much greater than r_π.) Similarly, the output resistance seen by the load resistor is given by

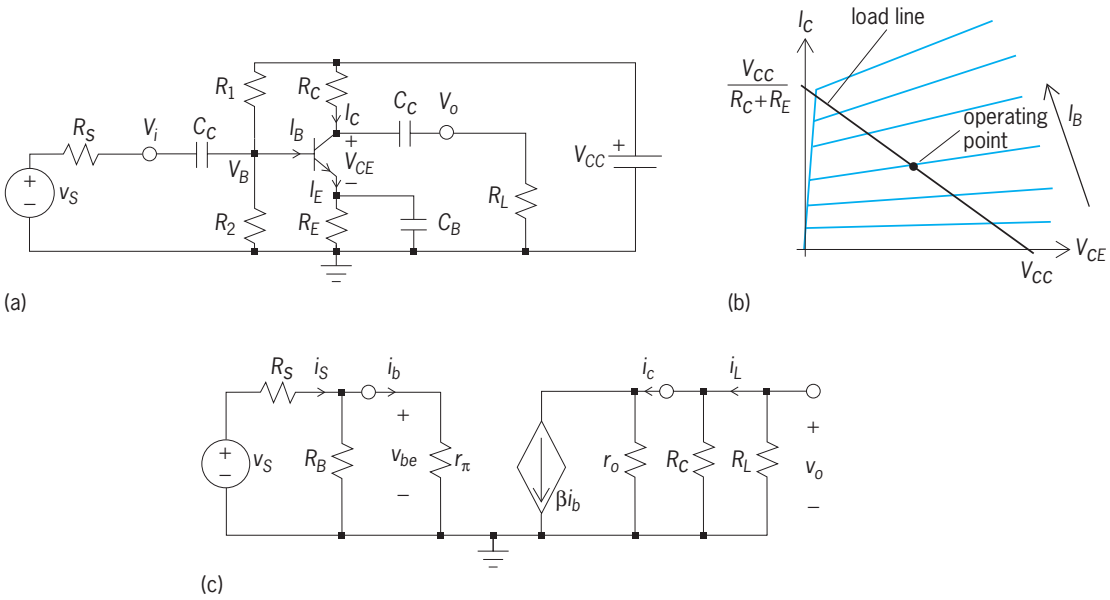

Fig. 5. Bipolar junction transistor common-emitter amplifier. (*a*) Circuit diagram. (*b*) Transistor dc characteristics and load line, showing placement of operating point. Arrow indicates increasing values of I_B. (*c*) Small-signal model of amplifier.

Eq. (20). (The resistor R_C is usually much smaller than r_o.)

$$r_{\text{out}} = \frac{r_o R_C}{r_o + R_C} \approx R_C \tag{20}$$

Finally, the voltage gain can be computed from the small-signal circuit, neglecting R_B and r_o, as Eq. (21),

$$A_v = \frac{v_o}{v_S} = -\frac{r_\pi}{R_S + r_\pi}\left(\frac{\beta}{r_\pi}R_C\right)\frac{R_L}{R_C + R_L} \tag{21}$$

which compares directly with Eq. (2). The quantity $(\beta/r_\pi)R_C = g_m R_C$ is the intrinsic gain of the unloaded amplifier, labeled K in Eq. (2). The amplifier's current gain is readily obtained from the small-signal circuit (Fig. 5*c*), neglecting r_o, as Eq. (22).

$$A_i = \frac{i_L}{i_S} \approx -\beta\frac{R_C}{R_C + R_L} \tag{22}$$

Thus, the common-emitter amplifier provides substantial voltage and current gains. It has low input resistance r_{in} and somewhat larger, but still small, output resistance r_{out}.

Common-base amplifiers. In a common-base amplifier (**Fig. 6***a*), the input voltage v_S with source resistor R_S is applied through a coupling capacitor C_C to the emitter-base junction which is forward-biased by a dc bias source $-V_{EE}$. Bias current flows from the power supply V_{CC} to $-V_{EE}$. The load R_L is connected through a second coupling capacitor C_C. In the small-signal model (Fig. 6*b*), the controlled current source is labeled $\alpha\ i_e$ because the emitter current i_e is the controlling input current into the bipolar junction transistor. The model uses the input resistance r_e defined in Eq. (13).

According to the small-signal model (Fig. 6*b*), the input resistance seen by the source to the right of R_S is given by Eq. (23). (The resistance r_e is usually much

$$r_{\text{in}} = \left(\frac{1}{R_E} + \frac{1}{r_e}\right)^{-1} \approx r_e \tag{23}$$

smaller than R_E, the resistor between the emitter and $-V_{EE}$.) The resistance r_{in} is bias dependent and by Eqs. (11) and (13) is typically quite small. Thus, the common-base amplifier has a very low input resistance. The output resistance r_{out} seen from the output voltage terminal v_o back into the circuit is found by setting $v_S = 0$, so that $i_e = 0$. Thus, this resistance

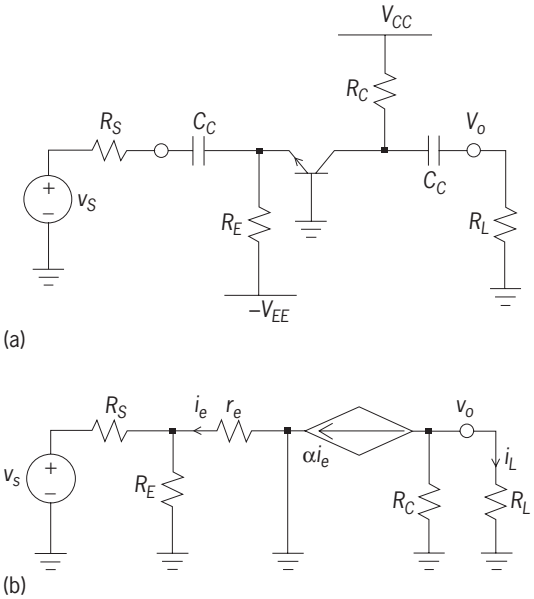

Fig. 6. Common-base amplifier. (*a*) Circuit diagram. (*b*) Small-signal model.

is given by Eq. (24), where R_C is the resistor between

$$r_{\text{out}} = R_C \tag{24}$$

the collector and V_{CC}.The voltage and current gains are derived from the small-signal model (Fig. 6*b*) as Eqs. (25) and (26).

$$A_v = \frac{V_o}{V_S} \approx \alpha \frac{1}{r_e + R_S} \frac{R_C R_L}{R_C + R_L} \tag{25}$$

$$A_i = \frac{i_L}{i_e} = \alpha \frac{R_C}{R_C + R_L} \tag{26}$$

The common-base amplifier is used as a current buffer which accepts an input current into a very small input resistance r_e and delivers that current (multiplied by $\alpha \approx 1$) with a much larger output resistance into the load. If R_C can be used as load resistor and R_L is absent, the current gain is almost equal to unity. Another advantage of the common-base amplifier is that its bandwidth is much wider than that of the common-emitter amplifier.

Common-collector amplifiers. The common-collector configuration, (**Fig. 7***a*), also called an emitter follower, is the third amplifier connection of a bipolar junction transistor. Here the collector is connected directly to the power supply, that is, to ac signal ground, so that it is common to input and output. The output is now taken from the emitter terminal. The input resistance r_{in} seen from the source with resistor R_S is computed from the small-signal model (Fig. 7*b*), neglecting r_o, as in Eq. (27), where R_B is the

$$r_{\text{in}} = R_B || R_{\text{in}} = R_B || [r_\pi + (\beta + 1)(R_E || R_L)] \tag{27}$$

resistor between the base and ground and R_E is the resistor between the emitter and a dc bias source $-V_{EE}$. (The notation || means "in parallel with.") Here, the resistance connected to the emitter is multiplied by $\beta + 1$ and adds to the resistance seen into the base of the amplifier, so that $R_{\text{in}} = v_b/i_b \approx (\beta + 1)(R_E || R_L)$ is very large. This is one of the advantages of the emitter-follower: it does not load the signal source. The output resistance r_{out} seen from the load R_L into the circuit, which is calculated from the small-signal model by setting $v_S = 0$, is approximated by Eq. (28). Thus, into the emitter is seen the total re-

$$r_{\text{out}} \approx \frac{r_\pi + R_S}{\beta + 1} \tag{28}$$

sistance of the base circuit divided by $\beta + 1$; that is, the output resistance is very small.

From the small-signal model, neglecting R_B, it follows that the voltage and current gains are given approximately by Eqs. (29) and (30).

$$A_v = \frac{v_o}{v_S} \approx 1 \tag{29}$$

$$A_i = \frac{v_o/R_L}{v_S/r_{\text{in}}} \approx \frac{r_{\text{in}}}{R_L} \tag{30}$$

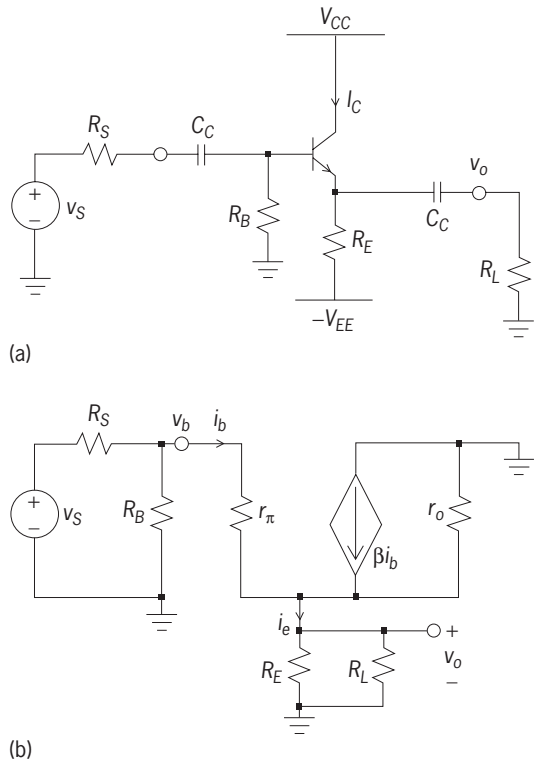

Fig. 7. Common-collector amplifier (emitter follower). (*a*) Circuit diagram. (*b*) Small-signal model.

These results show that the emitter follower is a very useful circuit; it is a buffer: without loading the source because of its large input resistance, it has a voltage gain of (somewhat less than) unity combined with a large current gain and a very small output resistance. *See* EMITTER FOLLOWER.

Field-Effect Transistor Amplifiers

In a field-effect transistor (FET) the currents are controlled by an applied electric field, generated by a voltage at a control electrode (the gate). There are two types of field-effect transistors, the junction field-effect transistor (JFET) and the insulated-gate or metal-oxide-semiconductor FET (MOSFET), each of which is available in two polarities (*n* channel and *p* channel). Further, there are depletion- and enhancement-mode MOSFETs, depending on how the conducting channel is established. The field-effect transistor differs from the bipolar junction transistor in the following characteristics: it is simpler to fabricate and occupies less space; in normal operation, the gate draws essentially no current, resulting in a very high input resistance; it exhibits no offset voltage at zero drain current; but, unfortunately, it generally has a significantly lower transconductance g_m at normal bias current levels. The operation of JFET and MOSFET amplifiers is similar in most respects. Differences arise from the fact that in a MOSFET the gate is insulated from the semiconducting channel by an oxide, whereas in a JFET this oxide insulator is replaced

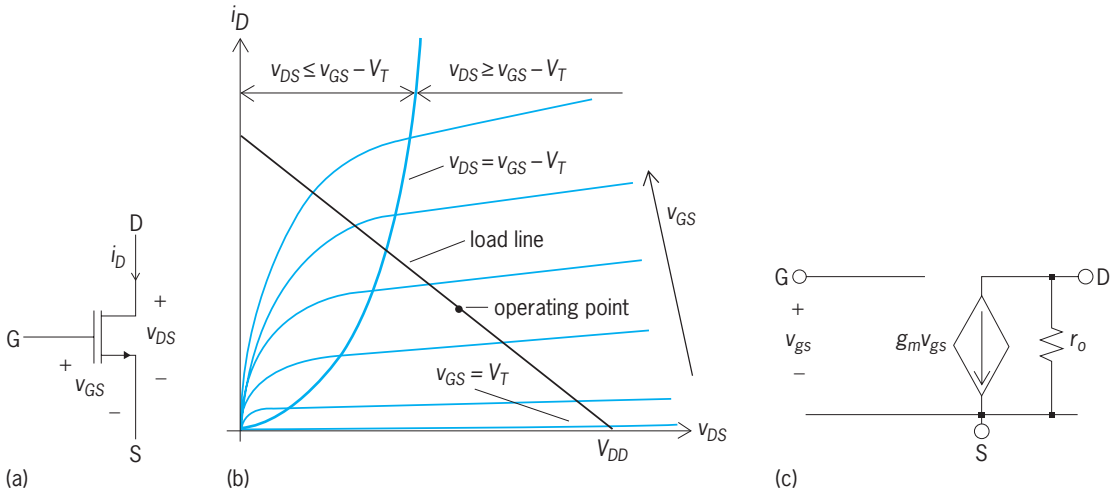

Fig. 8. An *n*-channel MOSFET. (*a*) Symbol. (*b*) Transistor dc characteristics. (*c*) Small-signal model.

by the depletion region of a reverse-biased *pn* junction. Only *n*-channel MOSFET amplifiers will be discussed.

MOSFET operation. A MOSFET has source (*S*), drain (*D*), and gate (*G*) terminals (**Fig. 8***a*). In an *n*-channel MOSFET, electrons flow from the source through the channel to the drain by applying a positive bias voltage V_{DS} to the drain. In an enhancement-type MOSFET, the channel is established by applying a positive voltage V_{GS} to the gate. The voltage V_{GS} must at least equal the threshold voltage V_T; increasing V_{GS} widens the channel and increases the current (Fig. 8*b*).

In the region $v_{DS} \ge v_{GS} - V_T$, the operation of the MOSFET is described by Eq. (31), where K is

$$i_D = K(v_{GS} - V_T)^2(1 + \lambda v_{DS}) \tag{31}$$

a parameter that depends on the electron mobility, the oxide capacitance per unit area, and the channel width and length; and λ is the channel-length modulation parameter that accounts for the increase in i_D with increasing v_{DS} (Fig. 8*b*). As in bipolar junction transistors, to achieve linear amplification from the nonlinear MOSFET it is necessary to establish an operating point, that is, a dc bias point on a graph of I_D versus V_{DS} (Fig. 8*b*), and superimpose the ac signals on the dc conditions. The bias voltage V_{DS} must lie in the range $V_{GS} - V_T < V_{DS} < V_{DD}$, where V_{DD} is the power supply voltage. The small-signal operation is then described by a small-signal model (Fig. 8*c*) with the output resistance given by Eq. (32) and the transconductance given by Eq. (33). The derivatives

$$r_o = \left[\frac{di_D}{dv_{DS}}\right]^{-1} = \frac{1}{\lambda K(V_{GS} - V_T)^2} \simeq \frac{1}{\lambda I_D} \tag{32}$$

$$g_m = \frac{di_D}{dv_{GS}} = 2K(V_{GS} - V_T) \tag{33}$$

are evaluated at the dc operating point. The gate is an open circuit so that no gate current flows.

MOSFET amplifiers. Analogous to bipolar junction transistor circuits, there are common-source, common-gate, and common-drain amplifiers. Biasing methods use the same principles as in bipolar junction transistors, but the situation is simpler here because no gate current flows. For example, the common-source amplifier (**Fig. 9***a*) can be constructed in a manner analogous to the bipolar junction transistor common-emitter circuit with bias circuitry consisting of resistors labeled R_1, R_2, R_B, and R_D, and battery V_{DD}. The resistors R_1 and R_2 are chosen large to keep the input resistance $R_G = R_1 || R_2$ of the amplifier high. The gain obtained from the

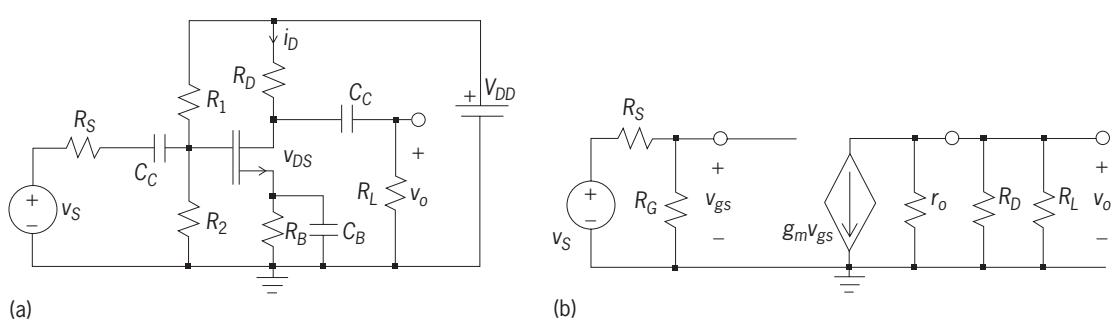

Fig. 9. Common-source amplifier. (*a*) Circuit diagram. (*b*) Small-signal model.

small-signal model (Fig. 9*b*) is given by Eq. (34),

$$A_V = \frac{v_o}{v_S} = -g_m \frac{R_G}{R_G + R_S} \frac{R_D R_L}{R_D + R_L} \qquad (34)$$

where R_S has been neglected compared to R_G, and also r_o. (Typical values for R_G and r_o are 5.4 MΩ and 50 kΩ, respectively.) This gain is much smaller than the one obtainable with a bipolar junction transistor amplifier. The output resistance equals $r_{\text{out}} = R_D \,||\, r_o \approx R_D$.

Common-gate and common-drain (source-follower) amplifiers can be constructed in a way analogous to bipolar junction transistor common-base and common-collector circuits. The source follower is found to be a good buffer circuit which does not load the signal source, provides a voltage gain of unity, and has a very low output resistance.

Differential Amplifiers

Single-ended (one input and one output) amplifiers have a number of disadvantages. They require large-valued bias resistors to achieve bias stability, which in turn restricts the dc gain. Also, they have low noise immunity, and are generally not useful for implementation in integrated circuits. The differential amplifier avoids many of these problems, but uses two or more active devices, which are either bipolar junction transistors or MOSFETs (**Fig. 10**). It is used as the input stage for almost all integrated amplifiers, including operational amplifiers.

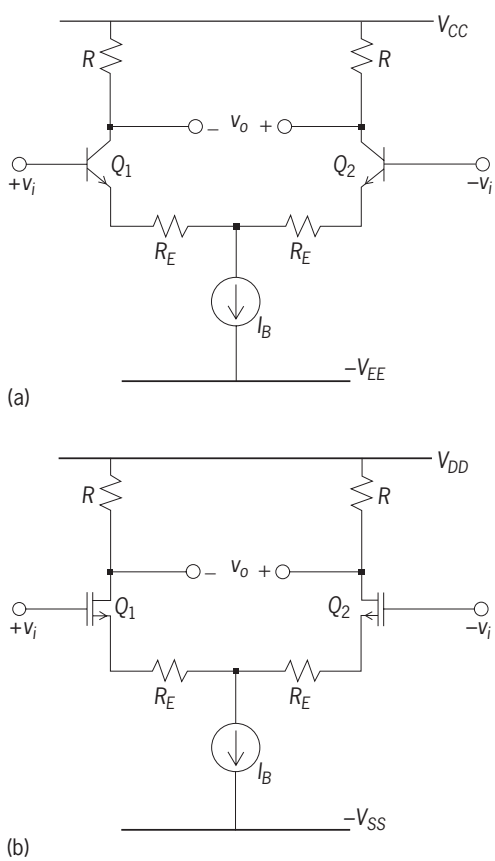

Fig. 10. Differential amplifiers. (*a*) Bipolar junction transistor amplifier. (*b*) MOSFET amplifier.

In the bipolar junction transistor stage (Fig. 10*a*), the differential gain is given by Eq. (35*a*), with the

$$A_{vd} = \frac{v_o}{v_d} = g_m(R||r_o) \approx g_m R \qquad (35a)$$

$$A_{vd} \approx \frac{g_m R}{1 + g_m R_E} \qquad (35b)$$

transconductance g_m defined in Eq. (11), where v_d is the differential voltage applied between the bases and R represents the load resistors. This simple differential transistor pair structure has the very small linear input range of only $|v_d| \approx V_T/2 \approx 13$ mV at room temperature. To increase the range, the so-called emitter-degeneration resistors R_E are included so that $v_d/2$ does not entirely appear across the base-emitter junctions of the bipolar junction transistors but also across R_E. The presence of R_E increases the linear range approximately to $|v_d| \approx R_E I_B/4$, but at the cost of a reduced gain given by Eq. (35*b*). For large R_E, the gain is essentially determined by a ratio of resistors, R/R_E. The differential input resistance between the bases of the amplifier can be derived from small-signal analysis as Eq. (36).

$$r_{d,\text{in}} = 2[r_\pi + (\beta + 1)R_E] \qquad (36)$$

The analysis of the MOSFET differential amplifier (Fig. 10*b*) is largely identical to that of the bipolar junction transistor circuit: The gain is given by Eq. (35) with g_m in Eq. (33).

The current source I_B can be realized in principle by a large resistor R_B; this approach, however, requires large power supply voltages because the voltage drop $R_B I_B$ across the bias "source" subtracts from the voltage available on the transistors. Large power supplies are normally not available in integrated circuits, and large-valued resistors may be difficult to implement. The solution is to establish a reference bias current, $I_{\text{ref}} = I_B$, possibly external to the integrated circuit chip, and use current mirrors to obtain bias currents throughout the integrated circuit. At the same time, current mirrors are used to avoid the load resistors R. *See* CURRENT SOURCES AND MIRRORS; DIFFERENTIAL AMPLIFIER.

Output Stages

Typically, amplifiers increase the power or signal levels from low-power sources, such as microphones, strain gages, magnetic disks, or antennas. After the small signals have been amplified, the amplifier's output stage must deliver the amplified signal efficiently, with minimal loss, and with no distortion to a load, such as a loudspeaker. Because the signal at the output stage is large, the small-signal approximations used earlier are normally not useful and the large-signal behavior must be investigated. The following discussion will concentrate on bipolar junction transistors, but it could equally well be based on MOSFETs; the circuit principles and concepts are the same.

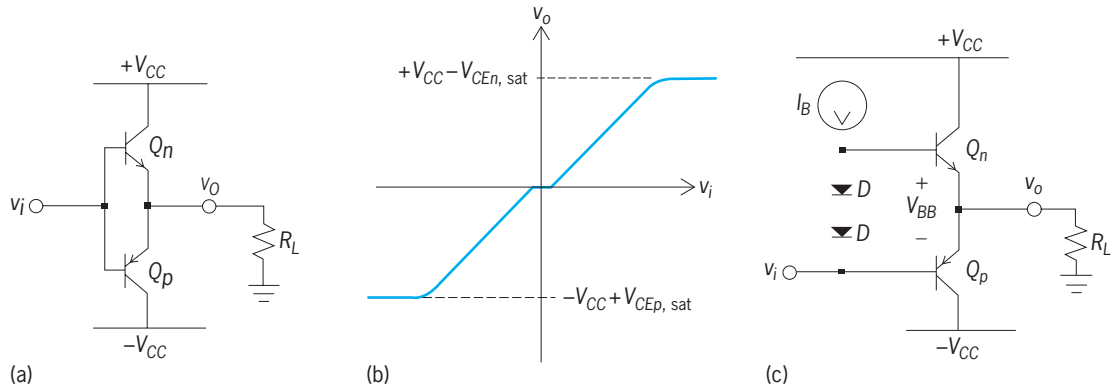

Fig. 11. Complementary bipolar junction transistor output stage. (*a*) Basic circuit, illustrating principle of operation. (*b*) Transfer curve. (*c*) Practical class AB stage.

To transfer power efficiently implies that the power dissipated in the output stage should be as small as possible. A serious problem arises when the power P dissipated in the transistors, given by Eq. (37), rises above a safe level specified by the

$$P = V_{CE}I_C \tag{37}$$

manufacturer, because an undue rise in the transistor's junction temperature will reduce the device's useful life or destroy it.

Buffer stage. An output stage should take a signal with minimal loading from an amplifier and deliver it with the least possible attenuation to the load. A buffer circuit will perform that job, such as the emitter follower (Fig. 7), where I_C must always flow and must be at least as large as the peak expected load current because I_C must be positive. This arrangement is referred to as class A operation. The average power dissipated in the transistor equals $V_{CC}I_C$, because the full power supply voltage V_{CC} appears across the device when the output voltage $v_o = 0$, that is, in the center of the signal range. The instantaneous power rises to 2 $V_{CC}I_C$ (assuming symmetrical supplies, $V_{CC} = V_{EE}$) when v_o reaches $-V_{CC}$. The power in the transistor depends on the value of the load resistor R_L; it can become infinite when $R_L = 0$, that is, when the load is short circuited, because then the collector current becomes arbitrarily large (flowing theoretically without resistance from V_{CC} through the transistor to ground) and the voltage V_{CE} equals V_{CC}. This dangerous situation must be avoided by clever circuit techniques that make an output stage short-circuit proof.

Since the power into a buffer is almost zero (because $i_b \approx 0$), the output power P_L must be drawn from the power supply and the conversion efficiency is a parameter of interest. It is defined by Eq. (38).

$$\eta = \frac{\text{load power}}{\text{supply power}} = \frac{P_L}{P_S} = \frac{0.5V_{o-\text{peak}}^2/R_L}{2V_{CC}I_C}$$

$$= \frac{1}{4}\frac{V_{o-\text{peak}}}{V_{CC}}\frac{V_{o-\text{peak}}}{R_L I_C} \le \frac{1}{4} \tag{38}$$

The factor $2V_{CC}$ appears in the denominator because (for $-V_{EE} = -V_{CC}$) the voltage across R_E is also equal to V_{CC} on average and R_E dissipates the same average power as the transistor. The last two factors in Eq. (38) are less than unity, so that the efficiency reaches only 25% at best. Much better performance can be obtained by using complementary devices.

Complementary (class AB) stage. This output stage has a complementary transistor pair, consisting of an *npn* transistor Q_n and a *pnp* transistor Q_p, connected between symmetrical power supplies $\pm V_{CC}$ (**Fig. 11***a*). When the input voltage $v_i = 0$, no current flows because neither of the transistors is forward biased ($V_{BE} = 0$), and therefore $v_o = 0$. If $v_i > V_{BEn} \approx 0.65$ V, Q_p is off, that is, does not conduct ($V_{EBp} < 0$), but Q_n conducts and current flows from $+V_{CC}$ through Q_n and the load resistance R_L to ground: $v_o > 0$. If $v_i < V_{EBp} \approx 0.65$ V, the situation is reversed: Q_n is off ($V_{BEn} < 0$), and Q_p conducts so that current flows from ground through R_L and Q_p to $-V_{CC}$: $v_o < 0$. Both transistors act as emitter followers, so that $v_o \approx v_i$, but only one at a time conducts and no dc current flows from $+V_{CC}$ to $-V_{CC}$. Apart from the small saturation voltage, $V_{CE,\text{sat}} \approx 0.2$ V, v_o can swing from $+V_{CC}$ to $-V_{CC}$ with a gain of approximately unity and very little distortion. This behavior is depicted in the graph of v_o versus v_i, called the transfer curve (Fig. 11*b*), which also shows the so-called crossover distortion at the origin. This distortion arises because v_i must at least equal the base-emitter voltages of the transistors before the emitter followers turn on and become active.

The problem is avoided by arranging the circuit such that a dc voltage V_{BB} equal to two forward-biased *pn*-junction voltages, that is, ≈ 1.3 V always appears between the bases of the transistors (Fig. 11*c*). Two diodes D (in practice, two diode-connected transistors) are forward biased by a current source I_B, which also supplies base currents to Q_n and Q_p. For ac signals, the forward-biased diodes are simply very small resistors. The operation is then as just described for the basic circuit (Fig. 11*a*), but the crossover distortion is avoided.

The power conversion efficiency for this circuit, neglecting the small power dissipated in the diodes, can be shown to equal $\eta = 0.785V_{o-\text{peak}}/V_{CC}$, that is, maximally 78.5%. By using the earlier results from the emitter follower, the gain is found to equal

approximately unity, and the input and output resistances are given by Eq. (39), where i_n and i_p are the

$$r_{\text{in}} \approx (\beta + 1)R_L \qquad r_{\text{out}} \approx r_{en} || r_{ep} = \frac{V_T}{i_n + i_p} \quad (39)$$

currents in the *npn* and *pnp* devices, respectively, and r_{en} and r_{ep} are their emitter resistances as defined in Eq. (13). Thus, the input resistance is very large and the output resistance very small, as required for a good follower or buffer circuit. *See* POWER AMPLIFIER.

Rolf Schaumann

Bibliography. P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3d ed., 1993; R. Gregorian and G. C. Temes, *Analog MOS Integrated Circuits for Signal Processing*, 1986; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Amplitude (wave motion)

The maximum magnitude (value without regard to sign) of the disturbance of a wave. The term "disturbance" refers to that property of a wave which perturbs or alters its surroundings. It may mean, for example, the displacement of mechanical waves, the pressure variations of a sound wave, or the electric or magnetic field of light waves. Sometimes in older texts the word "amplitude" is used for the disturbance itself; in that case, amplitude as meant there is called peak amplitude. This is no longer common usage. *See* LIGHT; SOUND.

As an example, consider one-dimensional traveling waves in a linear, homogeneous medium. The wave disturbance y is a function of both a space coordinate x and time t. Frequently the disturbance may be expressed as $y(x, t) = f(x \pm \upsilon t)$, where υ denotes the wave velocity. The plus or minus sign indicates the direction in which the wave moves, and the shape of the wave dictates the functional form symbolized by f. Then, the amplitude of the disturbance at some point x_0 is the maximum magnitude (that is, the maximum absolute value) achieved by f as time changes over the duration required for the wave to pass point x_0. A special case of this is the one-dimensional, simple harmonic wave $y(x, t) = A \sin [k(x \pm \upsilon t)]$, where k is a constant. The amplitude is A since the absolute maximum of the sine function is $+1$. The amplitude for such a wave is a constant. *See* HARMONIC MOTION; SINE WAVE.

If the medium which a wave disturbs dissipates the wave by some nonlinear behavior or other means, then the amplitude will, in general, depend upon position. *See* WAVE MOTION.

S. A. Williams

Amplitude modulation

The process or result of the process whereby the amplitude of a carrier wave is changed in accordance with a modulating wave. This broad definition includes applications using sinusoidal carriers, pulse carriers, or any other form of carrier, the amplitude factor of which changes in accordance with the modulating wave in any unique manner. *See* MODULATION.

Amplitude modulation (AM) is also defined in a more restrictive sense to mean modulation in which the amplitude factor of a sine-wave carrier is linearly proportional to the modulating wave. AM radio broadcasting is a familiar example. At the radio transmitter the modulating wave is the audio-frequency program signal to be communicated; the modulated wave that is broadcast is a radio-frequency, amplitude-modulated sinusoid. *See* RADIO BROADCASTING.

In AM the modulated wave is composed of the transmitted carrier, which conveys no information, plus the upper and lower sidebands, which (assuming the carrier frequency exceeds twice the top audio frequency) convey identical and therefore mutually redundant information. J. R. Carson in 1915 was the first to recognize that, under these conditions and assuming adequate knowledge of the carrier, either sideband alone would uniquely define the message. Apart from a scale factor, the spectrum of the upper sideband and lower sideband is the spectrum of the modulating wave displaced, respectively, without and with inversion by an amount equal to the carrier frequency.

For example, suppose the audio-frequency signal is a single-frequency tone, such as 1000 Hz, and the carrier frequency is 1,000,000 Hz; then the lower and upper sidebands will be a pair of single-frequency waves. The lower-sideband frequency will be 999,000 Hz, corresponding to the difference between the carrier and audio-signal frequencies; and the upper-sideband frequency will be 1,001,000 Hz, corresponding to the sum of the carrier and audio-signal frequencies. The amplitude of the signal appears in the amplitude of the sidebands. In practice, the modulating waveform will be more complex. A typical program signal might occupy a frequency band range of perhaps 100 Hz to 5000 Hz.

An important characteristic of AM, as illustrated by **Fig. 1**, is that, apart from a scale factor and constant term, either the upper or lower envelope of the modulated wave is an exact replica of the modulating wave, provided two conditions are satisfied: first, that the carrier frequency exceeds twice the highest speech frequency to be transmitted; and second, that the carrier is not overmodulated.

Single-sideband (SSB) modulation. This is a form of linear modulation that takes advantage of the fact that in an AM signal the carrier conveys no information and each of the two sidebands (that is, the upper sideband and the lower sideband) contains the same information. Hence it is not necessary to transmit both of them, and it is possible to achieve a 50% savings in bandwidth relative to normal AM or double-sideband (DSB) modulation.

The price for this bandwidth efficiency is increased complexity. The generation of SSB is typically accomplished either by filtering out the unwanted sideband of a DSB signal or by using a phase-shift modulator. This latter system involves use of a Hilbert

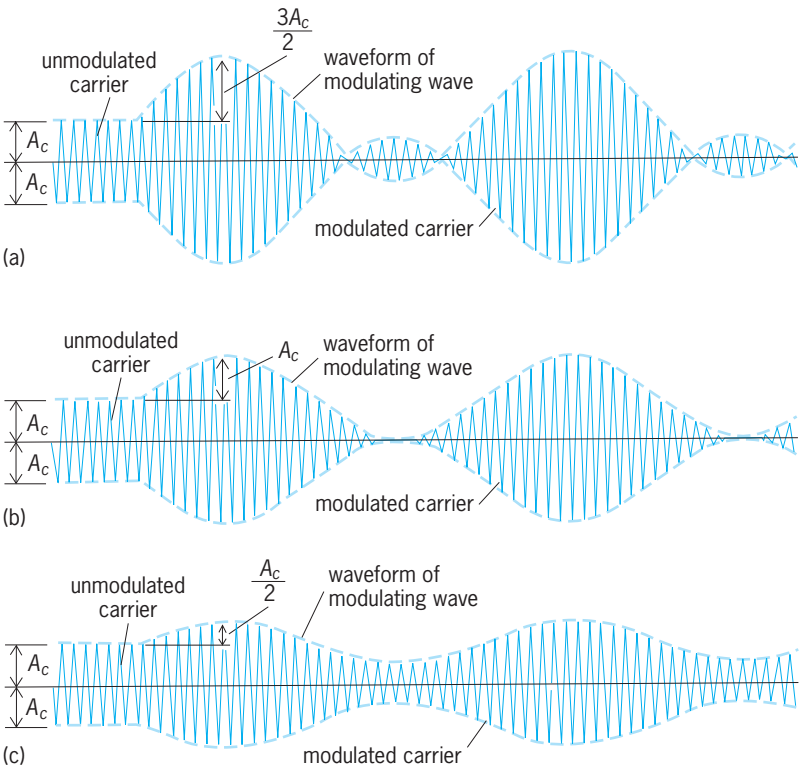

Fig. 1. Amplitude modulation of a sine-wave carrier by a sine-wave signal, with (*a*) 50% overmodulation, (*b*) 100% modulation, and (*c*) 50% modulation. (*After H. S. Black, Modulation Theory, Van Nostrand, 1953*)

transformer to shift the phase of the frequency components of the message by -90°. Specifically, if the message is denoted by $m(t)$, the transmitted waveform $s(t)$ is given by Eq. (1), where ω_0 is the angular

$$s(t) = m(t)\cos\omega_0 t \pm \hat{m}(t)\sin\omega_0 t \qquad (1)$$

frequency of the carrier and $\hat{m}(t)$ is the Hilbert transform of $m(t)$. If the plus sign is used in the above equation, the lower sideband is retained, whereas if the minus sign is used, the upper sideband is retained.

Either a coherent demodulator or a carrier reinsertion receiver can be used to demodulate an SSB signal. In either case, there are certain limits on phase and frequency errors that can be tolerated. Also of concern in SSB transmission are linearity requirements, since the envelope of an SSB signal can undergo significant variations.

SSB signaling is extensively used in voice transmission. Indeed, for applications such as telephone transmission over satellite links, one of the most common forms of modulation is the use of SSB subcarriers, which in turn frequency-modulate a carrier. The technique works as follows: Each voice message occupies roughly a 3-kHz bandwidth. The various voice messages are SSB-modulated onto individual subcarriers spaced 4 kHz apart. This composite baseband signal then frequency-modulates a carrier which is transmitted over the radio channel.

SSB modulation is also used for voice transmission in radio relay systems and coaxial cable systems. Indeed, this type of transmission has become so well accepted on a worldwide basis that standard frequency allocation plans have been set up by the Consultative Committee in International Telegraphy and Telephony (CCITT).

The choice of whether or not to use SSB depends upon the particular conditions under which the system is operating. SSB requires both less power and less bandwidth than does AM, but requires a more complex transmitter and receiver. Also, SSB systems have been shown to be less sensitive to the effects of selective fading than are both AM and DSB. *See* SINGLE SIDEBAND.

Vestigial-sideband (VSB) modulation. This is modulation whereby in effect the modulated wave to be transmitted is composed of one sideband plus a portion of the other adjoining the carrier. The carrier may or may not be transmitted. VSB is like SSB except in a restricted region around the carrier. The overall frequency response to the wanted sideband and to the vestigial sideband is so proportioned by networks that upon demodulation, preferably but not necessarily by a product demodulator, the original modulating wave will be recovered with adequate accuracy.

By thus transmitting a linearly distorted copy of both sidebands in a restricted frequency region above and below the carrier, the original modulating wave can now be permitted to contain significant components at extremely low frequencies, even approaching zero in the limit. By this means, at the cost of a modest increase in bandwidth occupancy, network requirements are greatly simplified. Furthermore, the low-frequency limitation and the inherent delay associated with SSB are avoided.

In standard television broadcasting in the continental United States the carrier is transmitted, envelope detection is used, and the vestigial sideband possesses a bandwidth one-sixth that of a full sideband. *See* TELEVISION.

Comparison of modulation spectra. In order to better understand the differences between AM, DSB, SSB, and VSB, it is helpful to consider the resulting spectrum of each waveform when the signal given by Eq. (2) is used as the modulation, where A and B

$$m(t) = A\cos\omega_1 t + B\cos\omega_2 t \qquad (2)$$

are constant amplitudes. **Figure 2***a* shows the single-sided spectrum of the modulation itself, where $f_1 = \omega_1/2\pi$ and $f_2 = \omega_2/2\pi$; and Fig. 2*b*–*e* shows the spectra of the AM, DSB, SSB, and VSB waveforms, respectively, when modulated by $m(t)$. In Fig. 2*b*, f_0 is the carrier frequency and a is the modulation index. In Fig. 2*e*, ϵ represents the fraction of the lower sideband at frequency $f_0 - f_1$ which is retained, that is, that is, it represents a vestige of the lower sideband of the DSB waveform of Fig. 2*c*.

Uses of AM in multiplexing. Multiplexing is the process of transmitting a number of independent messages over a common medium simultaneously. To multiplex, it is necessary to modulate. Two or more communicating channels sharing a common propagation path may be multiplexed by arranging them

along the frequency scale as in frequency division, by arranging them in time sequence as in time division, or by a process known as phase discrimination, in which there need be no separation of channels in either frequency or time.

Frequency-division multiplexing. When communication channels are multiplexed by frequency division, a different frequency band is allocated to each channel. Single-sideband carrier telephone systems are a good example. At the sending end, the spectrum of each channel input is translated by SSB to a different frequency band. For example, speech signals occupying a band of 300–3300 Hz might be translated to occupy a band range of 12,300–15,300 Hz corresponding to the upper sideband of a sinusoidal carrier, the frequency of which is 12,000 Hz. Another message channel might be transmitted as the upper sideband of a different carrier, the frequency of which might be 16,000 Hz.

At the receiving end, individual channels are separated by electric networks called filters, and each original message is recovered by demodulation. The modulated wave produced by SSB at the sending end becomes the modulating wave applied to the receiving demodulator.

When communication channels are multiplexed by frequency division, all channels may be busy simultaneously and continuously, but each uses only its allocated fraction of the total available frequency range.

Time-division multiplexing. When communication channels are multiplexed by time division, a number of messages is propagated over a common transmitting medium by allocating different time intervals in sequence for the transmission of each message. **Figure 3** depicts a particularly simple example of a two-channel, time-division system. Ordinarily, the number of channels to be multiplexed would be considerably greater. Transmitting and receiving switches must be synchronized; time is of the essence in this system, and the problem is to switch at the right time.

On the theoretical side there is a certain basic, fundamental question which must always be answered about any time-division system. The question is: At what rate must each communication channel be connected to its common transmitting path? Today it is known from the sampling principle that for successful communication each channel must be momentarily connected to the common path at a rate that is in excess of twice the highest message frequency conveyed by that channel. *See* PULSE MODULATION.

Viewed broadly, amplitude modulation of pulse carriers generates the desired amplitude and phase relationships essential to time-division multiplexing. In addition, whereas each communication channel may use the entire available frequency band for transmitting its message, it may transmit only during its allocated fraction of the total time.

Phase-discrimination multiplexing. This type of multiplexing, like SSB, saves bandwidth, may save signal power, and, like AM and VSB, has the important advantage of freely transmitting extremely low modulating frequencies. Furthermore, each communication channel may utilize all of the available frequency range all of the time.

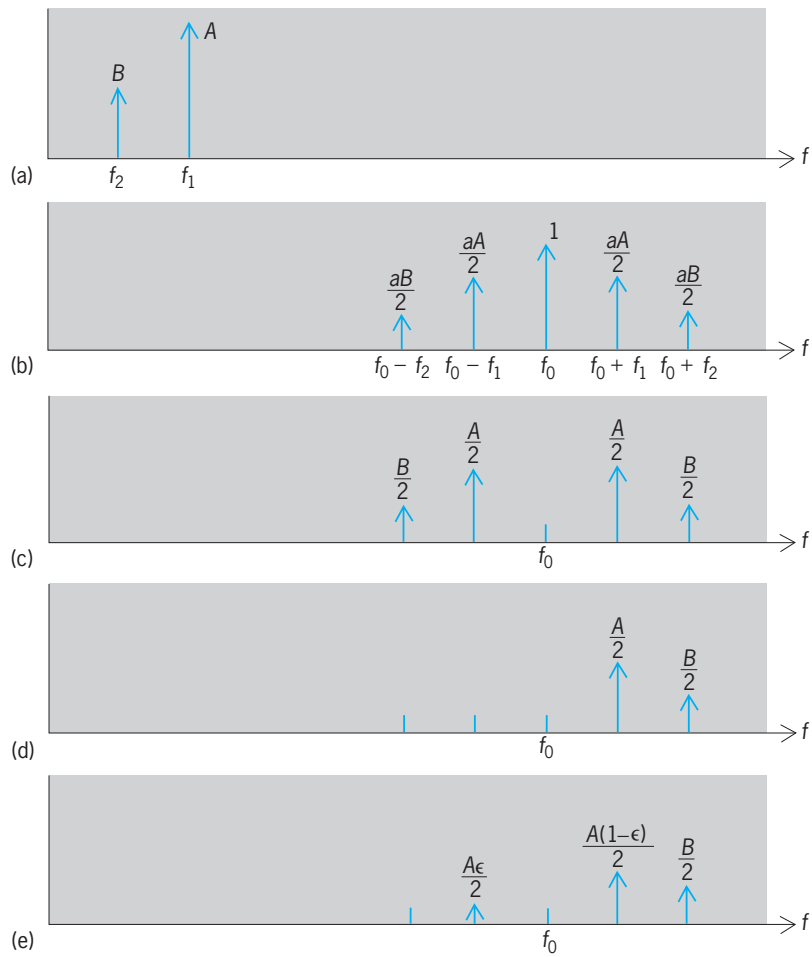

Fig. 2. Comparison of spectra of modulation waveforms. Signal amplitude is plotted as a function of frequency *f*. (*a*) Spectrum of modulation itself. (*b*) Normal amplitude modulation (AM). (*c*) Double-sideband (DSB). (*d*) Single-sideband (SSB). (*e*) Vestigial-sideband (VSB).

When n channels are multiplexed by phase discrimination, the modulating wave associated with each channel simultaneously amplitude-modulates $n/2$ carriers, with a different set of carrier phases provided for each channel. All sidebands are transmitted; $n/2$ carriers may or may not be transmitted. At the receiving end, with the aid of locally supplied carriers and an appropriate demodulation process, the n channels can be separated, assuming distortionless transmission, ideal modulators, and so on. Systems with odd numbers of channels can also be devised. Equality of sidebands and their exact phases account for the suppression of interchannel interference.

Day's system. This is a simple example of phase-discrimination multiplexing. Two sine-wave carriers of the same frequency but differing in phase by 90° are amplitude-modulated, each by a different message wave. The spectrum of each modulated sinusoid occupies the same frequency band. These modulated sinusoids are then added and propagated without distortion to a pair of demodulators. Quadrature carriers of correct frequency and phase are applied

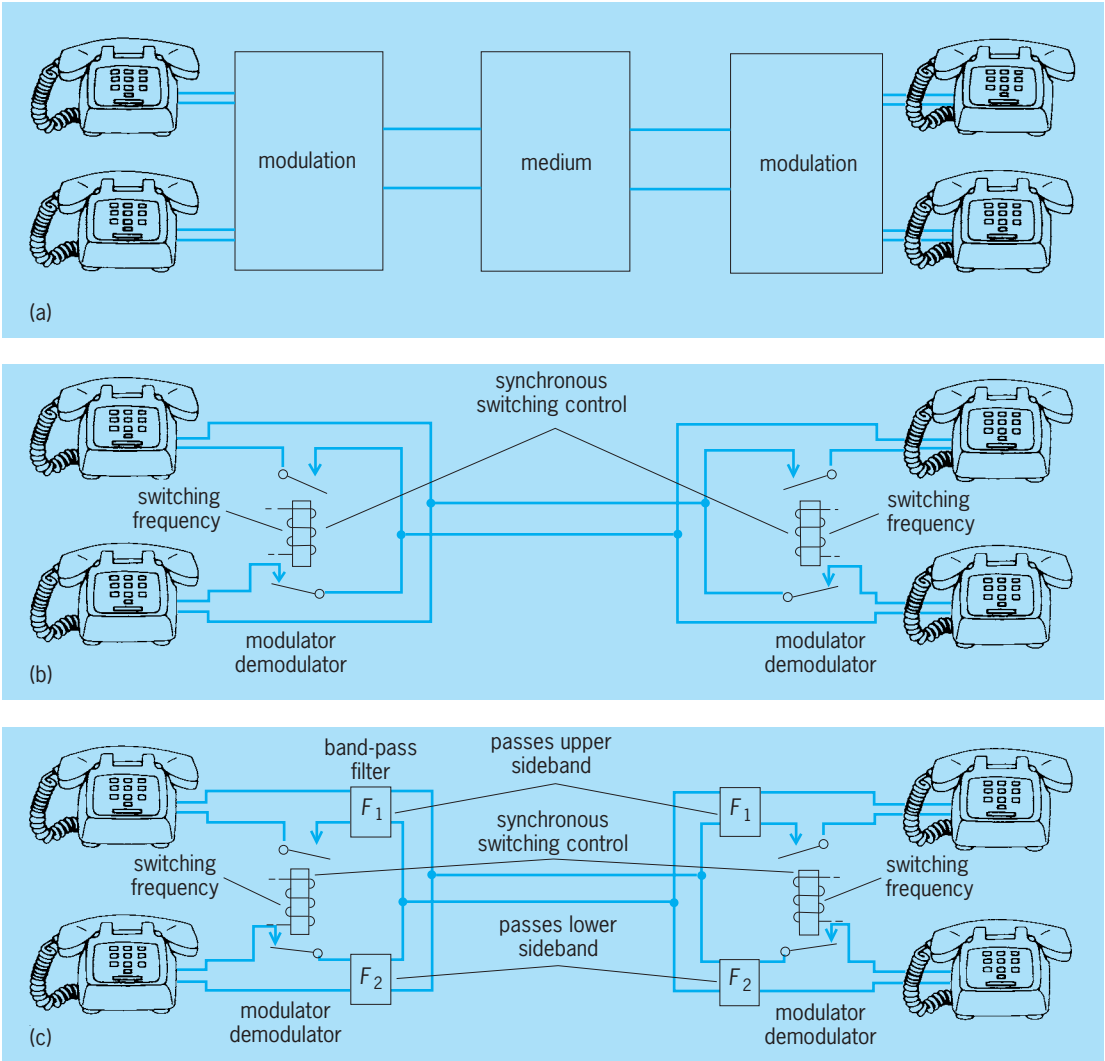

Fig. 3. Two-channel, time-division carrier system and two-channel, frequency-division carrier system. (*a*) General diagram of two-channel carrier system. (*b*) Amplitude modulation, time division. (*c*) Single-sideband modulation, frequency division. (*After H. S. Black, Modulation Theory, Van Nostrand, 1953*)

locally, one to each demodulator. Theoretically, a faithful copy of each message can be recovered.

Within the continental United States, for purposes of saving bandwidth, Day's system is used for multiplexing the two so-called color components associated with color television broadcasting. *See* MULTIPLEXING AND MULTIPLE ACCESS.

Modulator and demodulator. Many methods of modulating and demodulating are possible, and many kinds of modulators and demodulators are available for each method. *See* MODULATOR.

Fundamentally, since the sidebands of an amplitude-modulated sinusoid are generated by a multiplication of wave components which produces frequency components corresponding to the products, it is natural to envisage a product modulator having an output proportional to the product of two inputs: modulating wave and carrier. An ideal product modulator suppresses both modulating wave and carrier, transmitting only upper and lower sidebands. *See* AMPLITUDE MODULATOR.

At the receiving end, the original message may be recovered from either sideband or from both sidebands by a repetition of the original modulating process using a product modulator, commonly referred to as a product demodulator, followed by a low-pass filter. Perfect recovery requires a locally applied demodulator carrier of the correct frequency and phase. For example, a reduced carrier system creates its correct demodulator carrier supply by transmitting only enough carrier to control the frequency and phase of a strong, locally generated carrier at the receiver. *See* AMPLITUDE-MODULATION DETECTOR.

SSB systems commonly generate their modulator and demodulator carriers locally with independent oscillators. For the high-quality reproduction of music, the permissible frequency difference between modulator and demodulator carriers associated with a particular channel is limited to about 1–2 Hz. For monaural telephony, frequency differences approaching 10 Hz are permissible.

Unbalanced square-law demodulators, rectifier-type demodulators, and envelope detectors are often used to demodulate AM. However, even though

overmodulation (Fig. 1) is avoided, significant distortion may be introduced. In general, the amount of distortion introduced in this manner will depend upon the kind of demodulator or detector used, the amount of noise and distortion introduced prior to reception, and the percentage modulation. Harold S. Black; Laurence B. Milstein

Bibliography. L. W. Couch, *Digital and Analog Communication Systems*, 5th ed., 1996; M. Schwartz, *Information, Transmission, Modulation and Noise*, 4th ed., 1990; H. Stark, F. B. Tuteur, and J. B. Anderson, *Modern Electrical Communications*, 2d ed., 1988; F. G. Stremlers, *Introduction to Communication Systems*, 3d ed., 1990; H. Taub, *Principles of Communication Systems*, 2d ed., 1986; R. E. Ziemer and W. H. Tranter, *Principles of Communications*, 4th ed., 1994.

Amplitude-modulation detector

A device for recovering information from an amplitude-modulated (AM) electrical signal. Such a signal is received, usually at radio frequency, with information impressed in one of several forms. The carrier signal may be modulated by the information signal as double-sideband (DSB) suppressed-carrier (DSSC or DSBSC), double-sideband transmitted-carrier (DSTC or DSBTC), single-sideband suppressed-carrier (SSBSC or SSB), vestigial sideband (VSB), or quadrature-amplitude modulated (QAM).

The field of amplitude-modulation detector requirements splits by application, complexity, and cost into the two categories of synchronous and asynchronous detection. Analog implementation of nonlinear asynchronous detection, which is typically carried out with a diode circuit, is favored for consumer applications, AM-broadcast radio receivers, minimum-cost products, and less critical performance requirements. Synchronous detectors, in which the received signal is multiplied by a replica of the carrier signal, are implemented directly according to their mathematics and block diagrams, and the same general detector satisfies the detection requirements of all SSB, DSB, and VSB signals. Although synchronous detectors may operate in the analog domain by using integrated circuits, more commonly digital circuits are used because of cost, performance, reliability, and power advantages. A quadrature-amplitude-modulation detector is simply a two-channel version of a synchronous detector.

Synchronous detection. The idea behind synchronous detection is simple. There are two conceptual approaches: to reverse the modulation process (which is rather difficult), or to remodulate the signal from the passband (at or near the transmitter's carrier frequency) to the baseband (centered at dc or zero frequency). The remodulation approach is routine. Unfortunately, a nearly exact replica of the transmitter's carrier signal is needed at the receiver in order to synchronously demodulate a transmitted signal. Synchronous means that the carrier-signal reference in the receiver has the same frequency and phase as the carrier signal in the transmitter. There are three means available to obtain a carrier-signal reference. First, the carrier signal may actually be available via a second channel. There are no difficulties with this method because a perfect replica is in hand. Second, the carrier signal may be transmitted with the modulated signal. It then must be recovered by a circuit known as a phase-lock loop with potential phase and frequency errors. Third, the carrier signal may be synthesized by a local oscillator at the receiver, with great potential for errors. Unless otherwise stated, it will be assumed that a perfect replica of the carrier signal is available. *See* OSCILLATOR; PHASE-LOCKED LOOPS.

SSB, VSB, and DSB signals. Demodulation of an SSB-modulated (upper or lower), VSB-modulated, or DSB-modulated signal is as simple as multiplying it by the transmitter carrier-signal replica and low-pass filtering to remove unwanted high-frequency terms (sometimes called images) from the product in order to obtain the base-band signal (**Fig. 1***a*). Sometimes the multiplication is carried out in a single step, sometimes in two steps called conversion stages. If the type of transmitted-signal modulation is transmitted carrier (TC), the consequent dc component must be removed by a filter. If the information signal itself has a dc component, an amplitude-modulated TC method cannot be used to transmit the data because of the difficulty in separating the dc component of the signal from the carrier which is shifted to dc by the demodulation process. *See* ELECTRIC FILTER.

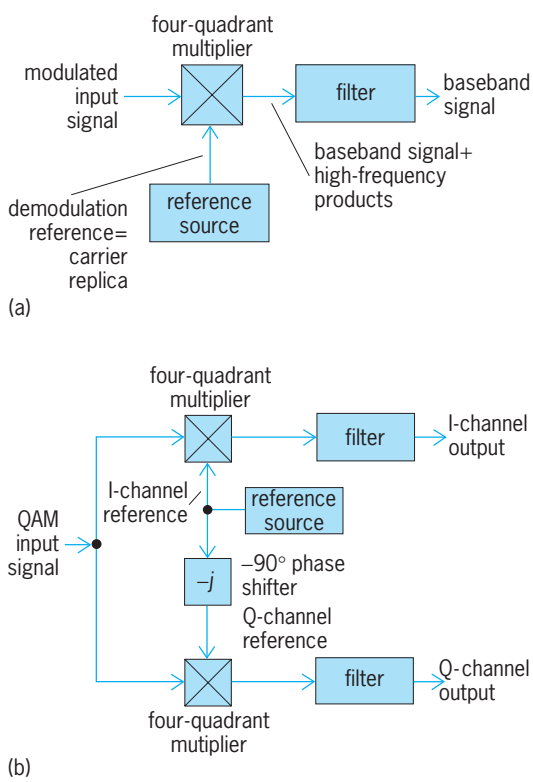

Fig. 1. Synchronous (coherent) detectors for (*a*) double-sideband (DSB), single-sideband (SSB), and vestigial-sideband (VSB) signals, and (*b*) for quadrature-amplitude-modulated (QAM) signals.

If the ith information component in a modulated signal is $A_i \cos(w_i t)$ and the modulating carrier is $\cos(w_c t)$, then the ith modulated element is given by Eq. (1). The ideal local oscillator signal is $y(t) =$

$$x(t) = A_i \cos(w_i t) \cos(w_c t) \tag{1}$$

$2 \cos(w_c t)$, so the demodulated signal is given by Eq. (2). The last term is removed by the filter, so that

$$\begin{aligned} z(t) = x(t)y(t) &= 2A_i \cos(w_i t) \cos^2(w_c t) \\ &= 2(A_i/2) \cos(w_i t)[1 + \cos(2w_c t)] \\ &= A_i \cos(w_i t) + [A_i \cos(w_i t) \cos(2w_c t)] \end{aligned} \tag{2}$$

the information signal is completely recovered. *See* TRIGONOMETRY.

The center frequency of the signal is often shifted down in two conversion stages. First an analog conversion by an oscillator translates the frequency of the signal to a value convenient for the analog-to-digital converter to operate; then the four-quadrant digital multiplier serves as a second converter and completes the frequency translation. The increasing performance and decreasing costs of analog-to-digital converters favor detectors in which analog conversion is eliminated and the analog-to-digital converter operates directly on the higher-frequency signal.

The importance of the accuracy of the reference in synchronous demodulation cannot be overemphasized. An average frequency error in the demodulating reference is intolerable because it will translate the modulated signals to the wrong frequencies. The cosine of the phase angle in the demodulating reference is the demodulator gain. A phase error can therefore seriously reduce the gain. A small frequency error integrates with time to become a phase error, causing objectionable periodic amplitude modulation of the received signal.

However, it is not necessary for the reference waveform to be a pure sinusoid, as long as it does have the correct periodicity and phase. An appeal to Fourier theory shows that this reference waveform can be represented by a sum of sinusoids harmonically related to the fundamental frequency of the reference. Multiplying the signal to be detected by this nonsinusoidal reference causes so-called images to appear after the multiplication process. These images, which exist at multiples of the fundamental frequency, can be removed by suitably filtering the multiplier output signal. *See* FOURIER SERIES AND TRANSFORMS; NONSINUSOIDAL WAVEFORM.

QAM signals. A synchronous detector of quadrature-amplitude-modulated signals consists of two synchronous demodulators operating in parallel (Fig. 1*b*). The two demodulating carrier-reference signals are separated in phase by precisely 90°. The accuracy of the phase angle of the carrier-reference signal limits the degree of separation (after demodulation) that can be maintained between the two QAM signals that have been transmitted on top of one another. If the ith transmitted signal element is given by Eq. (3), and the I and Q demodulation carrier references are given by Eqs. (4) and (5), then the filtered

$$x(t) = A_i \cos(w_i t) \cos(w_c t) + B_i \cos(w_i t) \sin(w_c t) \tag{3}$$

$$\begin{aligned} y_I(t) &= 2 \cos(w_c t + \Theta) \\ &= 2[\cos(w_c t) \cos(\Theta) - \sin(w_c t) \sin(\Theta)] \end{aligned} \tag{4}$$

$$\begin{aligned} y_Q(t) &= 2 \sin(w_c t + \Theta) \\ &= 2[\sin(w_c t) \cos(\Theta) + \cos(w_c t) \sin(\Theta)] \end{aligned} \tag{5}$$

(that is, the double-frequency products have been

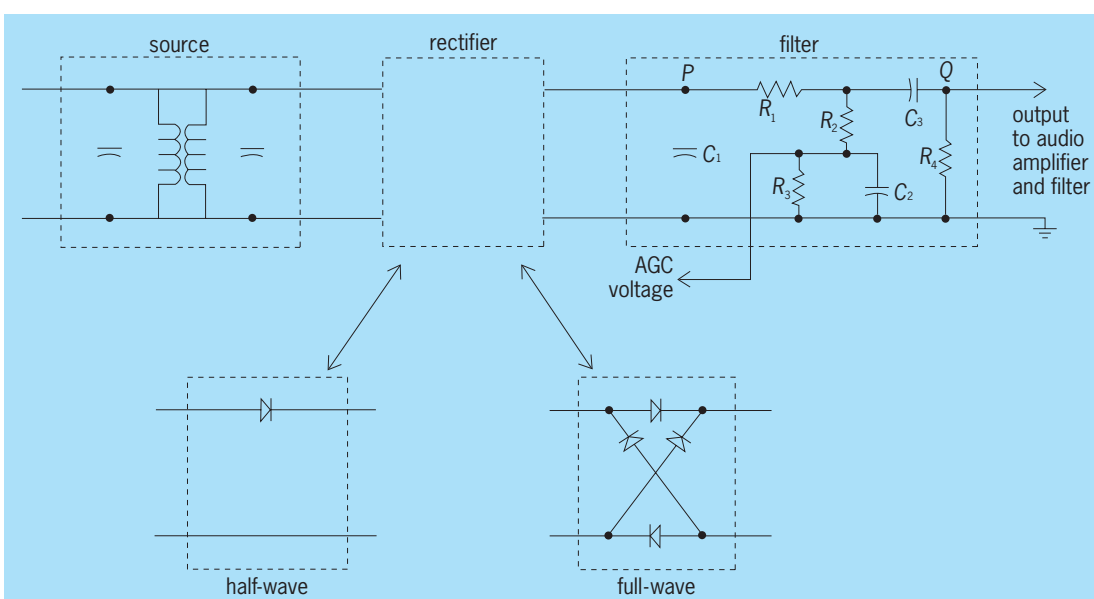

Fig. 2. Diode detector circuit. *P* = input point of filter; *Q* = output point of filter.

removed) demodulated outputs are given by Eqs. (6) and (7). Separation is complete for $\Theta = 0$.

$$z_I(t) = A_i \cos(w_i t)\cos(\Theta) - B_i \cos(w_i t)\sin(\Theta) \quad (6)$$

$$z_Q(t) = B_i \cos(w_i t)\cos(\Theta) + A_i \cos(w_i t)\sin(\Theta) \quad (7)$$

Asynchronous detection. Asynchronous detection applies to DSTC signals whose modulation index is less than 1, and is quite simple. First the received signal is full-wave rectified, then the result is low-pass filtered to eliminate the carrier frequency and its products, and finally the average value is removed (by dc blocking, that is, ac coupling). This result is identical to that which is obtained by synchronous detection with a reference that has been amplitude distorted to a square wave. The cheaper but less efficient half-wave rectifier can also be used, in which case the demodulation process is called envelope detection.

Diode detector. A diode detector circuit in a radio receiver has three stages (**Fig. 2**). The first stage is the signal source, which consists of a pair of tuned circuits that represent the last intermediate-frequency (i.f.) transformer which couples the signal energy out of the i.f.-amplifier stage into the detector. The second stage is a diode rectifier, which may be either full wave or half wave. Finally, the signal is passed through the third stage, a filter, to smooth high-frequency noise artifacts and remove the average value. *See* DIODE; INTERMEDIATE-FREQUENCY AMPLIFIER; RADIO RECEIVER; RECTIFIER.

The waveform shaping at the input point of the filter (Fig. 2) is determined by a capacitor C_1 in parallel with an equivalent resistance of R_1, the latter in series with a parallel combination of resistors, R_2 and R_4. The filter also has a capacitor C_3 between R_2 and R_4, and a parallel combination of a resistor R_3 and a capacitor C_2 in series with R_2. The reactances of both capacitors C_2 and C_3 are quite small at the information frequency, so the capacitors can be viewed as short circuits. Meanwhile, the C_3–R_4

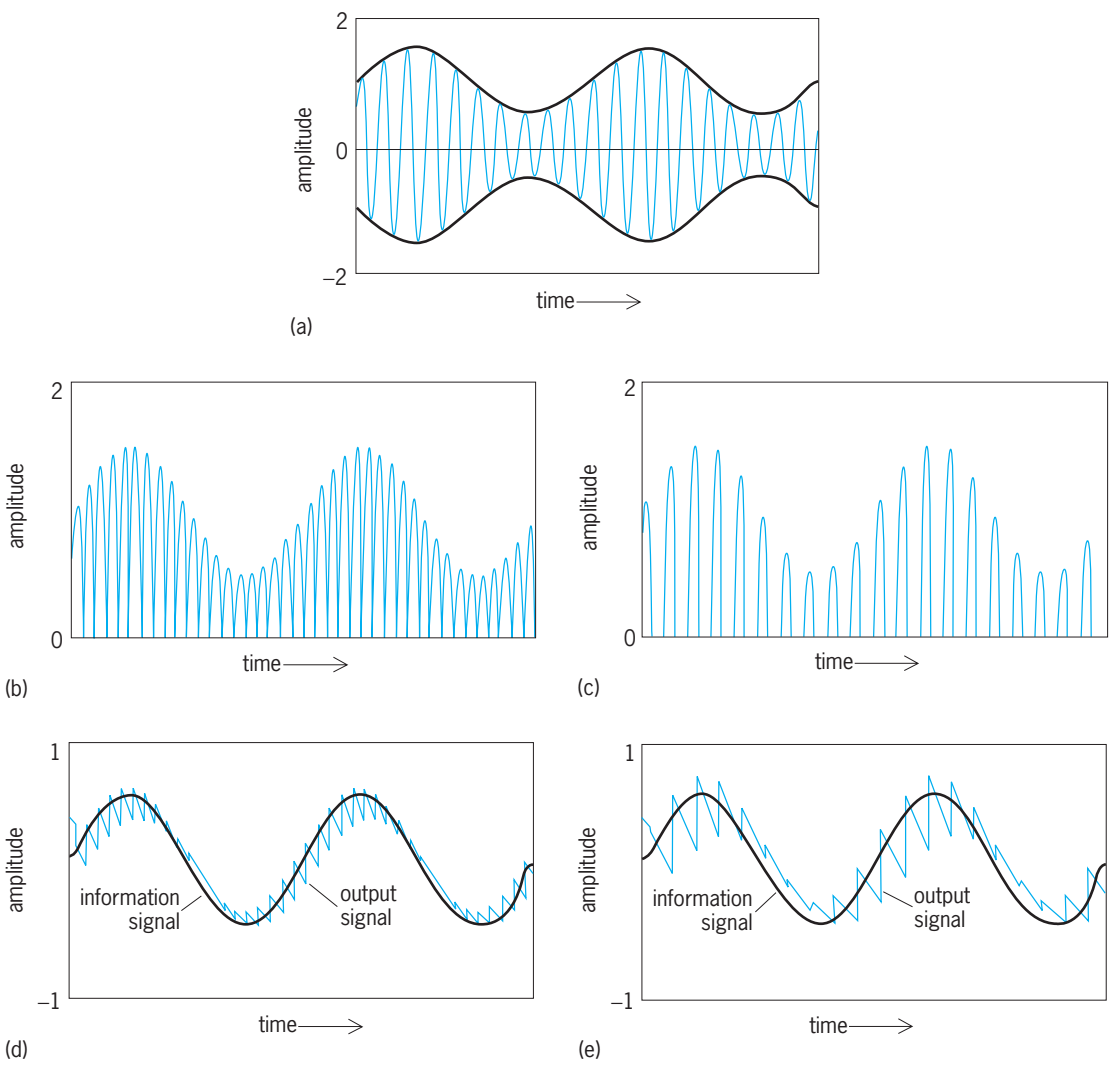

Fig. 3. Detection of a 50%-modulated double-sideband transmitted-carrier (DSTC) signal. An artificially low frequency ratio of 10:1 is used to illustrate raggedness of the output. (*a*) Received signal and modulating information-signal envelope. (*b*) Full-wave rectified received signal. (*c*) Half-wave rectified signal. (*d*) Information signal and output signal after full-wave rectifier and filter. (*e*) Information signal and output signal after half-wave rectifier and filter.

combination serves as a dc-blocking circuit to eliminate the constant or dc component at the output point of the filter. In order to bias the output point properly for the next amplifier stage, R_4 is replaced by a biasing resistor network in a real filter; R_4 is the single-resistor equivalent. *See* ELECTRICAL IMPEDANCE.

The strength of the signal arriving at the detector (**Fig. 3***a*) is proportional to the mean value of the signal at the input point of the filter and is sensed as the automatic-gain-control (AGC) voltage. Changes in this voltage level are used to adjust the amplification before the detector so that the signal strength at the detector input can remain relatively constant, although the signal strength at the receiver's antenna may fluctuate. Since capacitor C_2 shunts all signal energy and any surviving carrier energy to ground, the AGC voltage is roughly the average value at the input point of the filter scaled by $R_3/(R_1 + R_2 + R_3)$ because R_1 is much smaller than R_4. *See* AUTOMATIC GAIN CONTROL (AGC).

Waveforms. The received signal (Fig. 3*a*) is delivered from the i.f. stage to the diode rectifier. If the filter were simply a resistor, then the output from the diode rectifier would simply be the full-wave rectification (Fig. 3*b*) or half-wave rectification (Fig. 3*c*) of the received signal. (The presence of the capacitors in the filter provides the waveform shaping.) In both cases the modulating (information) signal shapes the envelopes of the rectified waveforms. In an amplitude-modulation radio, the output from the filter after full-wave rectification (Fig. 3*d*) would be the audio signal (plus some noise) that is delivered to the audio amplifier. Additional filtering can be provided as necessary to reduce the noise to an acceptable level. This amplified and filtered signal is finally delivered to an output device, such as a loudspeaker. The ragged waveform of the filter output contrasts with the smooth waveform of the information signal (Fig. 3*d*). The raggedness vanishes as the ratio of the i.f. frequency to the information frequency increases. While a synthetic example with a low ratio can be used to clearly show the effects within the demodulator (Fig. 3*d*), the amplitude of this raggedness noise decreases in almost direct proportion to the increase in the frequency ratios. The raggedness of the output signal from the filter after half-wave rectification (Fig. 3*e*) is much greater than that of the full-wave-rectifier case.

The actual worst-case ratio for standard-broadcast amplitude-modulation radio is 46.5:1. In this case the raggedness on the filtered outputs is reduced to a fuzz that can be seen in graphs of the waveforms but is well outside the frequency range of audio circuits, loudspeakers, and human hearing. *See* AMPLITUDE MODULATION; AMPLITUDE MODULATOR; MODULATION; WAVEFORM.

Stanley A. White

Bibliography. R. C. Dorf (ed.), *The Electrical Engineering Handbook*, 2d ed., 1997; M. S. Roden, *Analog and Digital Communication Systems*, 3d ed., 1991; M. Valkenburg (ed.), *Reference Data for Engineers: Radio, Electronics, Computer, and Communications*, 8th ed., 1996.

Amplitude-modulation radio

Radio communication employing amplitude modulation of a radio-frequency carrier wave as the means of conveying the desired intelligence. In amplitude modulation the amplitude of the carrier wave is made to vary in response to the fluctuations of a sound wave, television image, or other information to be conveyed. *See* AMPLITUDE MODULATION; RADIO.

Amplitude modulation (AM), the oldest and simplest form of modulation, is widely used for radio services. The most familiar of these is broadcasting; others include radiotelephony and radiotelegraphy, television picture transmission, and navigational aids. The essentials of these radio systems are discussed in this article.

Low frequency (long wave). European and Asian countries use frequencies in the range 150–255 kHz for some broadcast services. An advantage of these frequencies is stable and relatively low-attenuation wave propagation. When not limited by atmospheric noise, large areas may be served by one station. In the United States these frequencies are reserved for navigational systems and so are not available for broadcasting. *See* RADIO SPECTRUM ALLOCATION; RADIO-WAVE PROPAGATION.

Low-frequency (lf) broadcast antennas are omnidirective and radiate vertically polarized waves. Unless special means are used to reduce antenna selectivity, the highest program frequencies transmitted are substantially below 10,000 Hz.

Medium frequency. The frequencies in the range from 535 to 1705 kHz are reserved for AM (standard) broadcasting. In the Western Hemisphere this band is divided into channels at 10-kHz intervals, certain channels being known as clear, regional, and local, according to the licensed coverage and class of service. The local channels are occupied by stations, usually of 250-W output, servicing smaller localities. Many stations occupy a channel, but they are situated far enough apart to permit interference-free coverage in the local area. Fewer stations of higher power, but with greater distances between them, share the regional channels. A few clear channels are occupied by high-power stations (50,000-W maximum output in the United States). These stations may have exclusive use of a channel, or may share it with another distant station. *See* RADIO BROADCASTING.

Interference between co-channel regional stations and clear-channel stations is minimized by use of directive antennas, which suppress radiation toward other stations and direct it to main populated areas from a properly located station.

European medium-frequency (mf) broadcasting channels are assigned at 9-kHz intervals rather than the 10-kHz intervals used in the Western Hemisphere. This reduced spacing provides more channels within the mf band. The technique of directive antennas, which also provides more channels within a band by minimizing interference between stations, has not been used extensively in Europe.

Vertically polarized radiation is used at medium and low frequencies propagated over the Earth's

surface. There is also propagation of high-angle radiation via reflection from the ionosphere, a phenomenon that predominates at night but is relatively absent during daylight. This sky-wave propagation accounts for the familiar long-distance reception at night. At distances where the downcoming sky waves overlap the ground wave, fading and distortion of the signal occurs. Receivers in the groundwave zone get stable signals day and night. In the overlap zone, daylight reception may be stable but night reception disturbed by fading. In the sky-wave zone at night, assuming no interference from other stations, satisfactory service may be received over long distances. Reception in this zone depends on atmospheric noise, which varies seasonally, and the state of the ionosphere, which varies greatly from night to night depending upon astronomical conditions that affect the upper atmosphere.

Individual AM broadcast stations transmit program frequencies ranging from 30 to 10,000 Hz with excellent fidelity. To obtain suitable tuning selectivity between channels, AM broadcast receivers may reproduce less extensive program frequencies, according to make, cost, type, and condition of the receiver.

High frequency (shortwave). Small bands of frequencies between 3000 and 7500 kHz are used in tropical areas of high atmospheric noise for regional broadcasting. This takes advantage of the lower atmospheric noise at these frequencies and permits service under conditions where medium frequencies have only severely limited coverage. Wave propagation day and night is by sky wave. Groundwave coverage from such stations is usually negligible, since high-angle horizontally polarized radiation is used. Short-distance coverage by this mode is by ionospheric reflection of waves radiated almost vertically.

Long-distance international broadcasting uses high-power transmitters and directive (beam) antennas operating in the bands 5950–6200 kHz, 9500–9775 kHz, 11,700–11,975 kHz, 15,100–15,450 kHz, 17,700–17,900 kHz, 21,450–21,750 kHz, and 25,600–26,100 kHz. These bands are allocated throughout the world for this purpose, and the band used for any path depends upon ionospheric conditions which vary with direction, distance, hour, season, and the phase of the 11-year sunspot cycle. Typically, waves are propagated in low-angle beams by multiple reflections between ionosphere and Earth to cover transoceanic distances, and signals are often distorted during transmission. These bands are so crowded that a signal is seldom received without interference for any prolonged period. Reception from particular stations can be improved by the use of special directive receiving antennas. Propagation partially through the ionosphere also causes rotation of the transmission's plane of polarization due to the Faraday effect. *See* FARADAY EFFECT.

The technical performance of high-frequency (hf) broadcast transmission systems is usually to the same standards employed for mf broadcasting, although propagation and interference conditions seldom make this evident to a listener.

AM telephony and telegraphy. The first radiotelephony was by means of amplitude modulation, and its use has continued with increasing importance. Radiotelephony refers to two-way voice communication. Amplitude modulation and a modified form called single-sideband are used almost exclusively for radiotelephony on frequencies below 30 MHz. Above 30 MHz, frequency or phase modulation is used almost exclusively, a notable exception being 118–132 MHz, where amplitude modulation is used for all two-way very high frequency (vhf) radiotelephony in aviation operations.

The least expensive method known for communicating by telephony over distances longer than a few tens of miles is by using the high frequencies of 3–30 MHz. Furthermore, since radio is the only way to communicate with ships and aircraft, hf AM radiotelephony has remained essential to these operations, except for short distances that can be covered from land stations using the very high frequencies. Therefore, hf radiotelephony has become established for a great variety of pioneering and exploration enterprises and private, public, and government services needing telephone communication, fixed or mobile, over substantial distances where there are no other ways to telephone. Because AM techniques are simple and inexpensive, this form of modulation has predominated. The economic development of distant hinterland areas depends greatly on hf AM radio telephony to the outside world, although satellite transmission and reception has generally replaced hf AM radio telephony for all purposes except broadcasting.

Widespread use has led to serious crowding of the hf band. All governments are suspending the use of the hf band wherever it is technically and economically feasible to employ frequencies above 30 MHz using either direct transmission or radio repeater stations. The trend to single-sideband modulation also alleviates the pressure of congestion in the hf band.

Many of the radiotelephone systems in use operate two ways on one frequency in simplex fashion, that is, all stations on the frequency are normally in a receiving status. Transmission, by press-to-talk (manual) or voice-operated carrier (automatic) switching from reception to transmission, occurs only while the sender is talking. Many stations can thus occupy one frequency provided the volume of traffic by each of the stations is small. Two-way telephony must be strictly sequential. This system is not adapted for connection to a normal two-wire telephone.

Full duplex radiotelephony, for interconnection with wire telephone systems, is essential for most public correspondence. This requires two frequencies, each available full time in one direction. Even so, typical fading of signals during propagation requires that voice-operated antiregeneration devices be used to maintain circuit stability. Talking must be sequential between speakers as there can be no interrupting, but the system will interconnect with conventional business or home telephones.

Amplitude-modulated telegraphy consists of

interrupting a carrier wave in accordance with the Morse dot-dash code or codes used for the printing telegraph. Much of the radiotelegraphic traffic of the world uses AM telegraphy, although there has been extensive conversion to frequency-shift (frequency-modulation) telegraphy since 1944, the latter being better adapted to automatic teleprinting operations. Radiotelegraph operations have been refined, speeded, and mechanized, but, under adverse noise and fading conditions, AM manual telegraphy between experienced operators is still more reliable. Most aviation and marine radiotelegraphy uses AM manual methods. *See* TELEGRAPHY; TELEPHONE SERVICE.

Single-sideband (SSB) hf telephony. This is a modified form of amplitude modulation in which only one of the modulation sidebands is transmitted. In some systems the carrier is transmitted at a low level to act as a pilot frequency for the regeneration of a replacement carrier at the receiver. Where other means are available for this purpose, the carrier may be completely suppressed. Where intercommunication between AM and SSB systems is desired, the full carrier may be transmitted with one sideband.

Since 1954 the use of SSB has expanded rapidly in replacing common AM telephony for military and many nonpublic radio services. In time, SSB will gradually displace AM radiotelephony to reduce serious interference due to overcrowding of the radio spectrum. SSB transmission also is less affected by selective fading in propagation.

Multiplexing, both multiple-voice channels or teleprinter channels included with voice, is applied to SSB transmission. Teleprinting and data transmission by frequency multiplex using SSB radiotelephone equipment is increasing and displacing older methods of radiotelegraphy for fixed point-to-point government and common-carrier services. *See* SINGLE SIDEBAND.

Aviation and marine navigation aids. Amplitude-modulated radio has a dominant role in guidance and position location, especially in aviation, which is almost wholly under radio guidance. Radio is used in traffic control from point to point, in holding a position in a traffic pattern, and in approach and landing at airports. Distance measuring from a known point, radio position markers, runway localizers, and glidepath directors all use AM in some form, if only for coded identification of a facility.

Marine operations are not so dependent on radio facilities as are those of aviation, but almost every oceangoing ship makes use of direction finding, at least to determine its bearing from a radio station of known location. Special marine coastal beacon stations emit identified signals solely for direction-finding purposes. Certain navigational systems (Decca, loran) for aviation are also used by ships for continuous position indication and guidance. *See* ELECTRONIC NAVIGATION SYSTEMS.

Television broadcasting. Amplitude modulation is used for the broadcasting of the picture (video) portion of television. In England, France, and a few other places amplitude modulation is also used for the sound channel associated with the television picture, but frequency modulation is more commonly used for sound.

Countries of the Western Hemisphere, Japan, Philippines, Thailand, and Iran broadcast television video in an emission band of 4.25 MHz; the English video bandwidth is 3 MHz; the French system, 10 MHz. The rest of continental Europe (except Russia) use a bandwidth of 5.25 MHz. The carrier frequencies employed are between 40 and 216 MHz, and 470 to 890 MHz. A channel allocation includes the spectrum needed for both sound and picture. Japan and Australia also use 88–108 MHz for television broadcasting. Reception of such transmissions is often impaired by multipath (the same signal arriving out-of-phase at the receiver from different propagation paths due to reflections by large objects and terrain), causing effects such as ghosting.

The English and French systems employ positive video modulation, in which white corresponds to higher amplitudes of the modulation envelope and black corresponds to lower amplitudes. All other established video broadcasting uses negative modulation, working in the opposite sense. Synchronizing pulses in the negative system are at maximum carrier amplitude. The dynamic range from black to white in the picture varies from 75 to 25% of maximum amplitude.

The upper-frequency portion of one video sideband is suppressed by filters so that its remaining vestige, together with the other complete sideband, is transmitted. This is called vestigial sideband transmission and avoids unnecessary spectrum usage. *See* TELEVISION TRANSMITTER.

Edmund A. Laport;
Michael C. Rau

Stereophonic broadcasting. A number of systems for stereophonic AM radio broadcasting have existed since the late 1970s. These systems allow AM radio stations to transmit two channels of information in the same spectrum space where only one could exist previously. The result is similar to that of stereophonic FM, audio cassette, and other binaural entertainment media.

To transmit in AM stereo, a broadcast station employs a device known as an exciter to adapt its existing transmitter. The exciter has left- and right-channel audio inputs as well as summed audio- and radio-frequency outputs. Stereophonic audio program material is connected to the audio inputs, while the outputs attach to the AM transmitter. The summed audio (left channel plus right channel) information is used to amplitude-modulate the transmitter, to assure that compatibility with existing monophonic receivers is maintained. The radio-frequency exciter output, however, contains the stereophonic information to be transmitted. This output is connected in place of the transmitter's crystal oscillator, and although the signal present is at the exact carrier frequency of the radio station, it is also phase-modulated by the difference (left channel minus right channel) audio information. A subaudible and low-level identification tone is usually added to this difference audio information. This allows AM radio

receivers to identify stereophonic transmissions and activate the appropriate decoding circuitry, as well as a stereo indicator lamp commonly found on the radio front panel. *See* PHASE MODULATION; RADIO RECEIVER.

Stereophonic AM broadcasting methods all have their roots in a system known as quadrature multiplexing. This system allows the transmission of two channels of information on a single carrier frequency, but true quadrature transmissions are inherently incompatible with the majority of radios available to consumers. The differences between various systems employed to transmit stereophonic AM broadcasts all relate to the methods used to overcome this incompatibility. *See* AMPLITUDE MODULATION; RADIO; STEREOPHONIC RADIO TRANSMISSION.

Stanley Salek

Bibliography. National Association of Broadcasters, *NAB Guide to Advanced Television Service,* 2d ed., 1991; J. Whitaker (ed.), *NAB Engineering Handbook*, 9th ed., National Association of Broadcasters, 1999.

Amplitude modulator

A device for moving the frequency of an information signal, which is generally at baseband (such as an audio or instrumentation signal), to a higher frequency, by varying the amplitude of a mediating (carrier) signal. The motivation to modulate may be to shift the signal of interest from a frequency band (for example, the baseband, corresponding to zero frequency or dc) where electrical disturbances exist to another frequency band where the information signal will be subject to less electrical interference; to isolate the signal from shifts in the dc value, due to bias shifts with temperature or time of the characteristics of amplifiers, or other electronic circuits; or to prepare the information signal for transmission.

Familiar applications of amplitude modulators are standard-broadcast or amplitude-modulation (AM) radio; data modems (modem = modulator + demodulator); and remote sensing, where the information signal detected by a remote sensor is modulated by the remote signal-conditioning circuitry for transmission to an electrically quiet data-processing location. *See* AMPLITUDE-MODULATION RADIO; MODEM; REMOTE SENSING.

The primary divisions among amplitude modulators are double sideband (DSB); single sideband (SSB); vestigial sideband (VSB); and quadrature amplitude (QAM), where two DSB signals share the same frequency and time slots simultaneously. Each of these schemes can be additionally tagged as suppressed carrier (SC) or transmitted carrier (TC).

To amplitude-modulate an information signal is (in its simplest form) to multiply it by a second signal, known as the carrier signal (because it then carries the information). A real (as opposed to complex) information signal at baseband, or one whose spectrum is centered about zero frequency, has an amplitude spectrum which is symmetric (an even function; **Fig. 1***a*) and a phase spectrum which is asymmetric (an odd function; Fig. 1*b*) in frequency about zero. Because of this symmetry, the information in the upper or positive sideband (positive frequencies) replicates the information in the lower or negative sideband (negative frequencies). Balanced modulation or linear multiplication (that is, four-quadrant multiplication where each of the two inputs is free to take on positive or negative values without disturbing the validity of the product) of this information signal by a sinusoidal carrier of single frequency, f_c, moves the information signal from baseband and replicates it about the carrier frequencies f_c and $-f_c$ (Fig. 1*c*). *See* FOURIER SERIES AND TRANSFORMS.

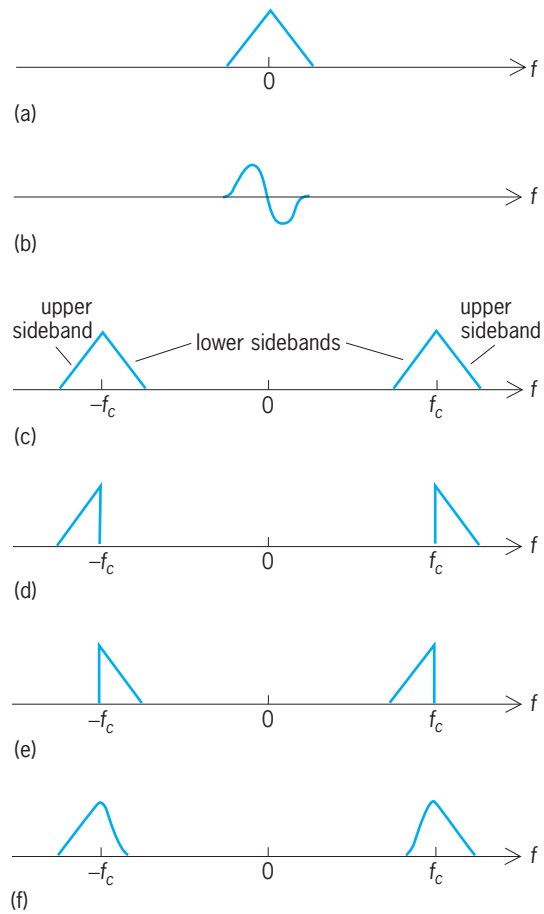

Fig. 1. Baseband spectra and spectra of outputs from suppressed-carrier modulators. Amplitudes and phases are shown as functions of frequency *f*. (*a*) Baseband amplitude spectrum. (*b*) Baseband phase spectrum. (*c*) Amplitude spectrum of output from double-sideband (DSB) modulator. (*d*) Amplitude spectrum of output from upper single-sideband (SSB) modulator. (*e*) Amplitude spectrum of output from lower single-sideband (SSB) modulator. (*f*) Amplitude spectrum of output from vestigial-sideband (VSB) modulator.

Suppressed-carrier modulators. Linear multiplication (described above) produces double-sideband suppressed-carrier (DSSC or DSBSC) modulation of the carrier by the information. One of the redundant sidebands of information may be eliminated by filtering (which can be difficult and costly to do adequately) or by phase cancellation (which is usually a much more reasonable process) to produce the more

efficient single-sideband suppressed-carrier (SSBSC or SSB) modulation (Fig. 1*d, e*). The process of only partially removing the redundant sideband (usually by deliberately imperfect filtering) and leaving only a vestige of it is called vestigial-sideband (VSB) modulation (Fig. 1*f*). *See* ELECTRIC FILTER.

The DSB and SSB generation mechanisms may be illustrated by considering one sinusoid, $A_i \cos(w_i t)$, of a weighted sum of sinusoids that may be assumed to constitute the information signal, $u(t)$. The ith information signal component can be multiplied by the carrier signal, $\cos(w_c t)$, to produce the DSSC product, $x(t)$, as in Eq. (1). This expression clearly shows the generation of the lower-sideband-frequency signal component at $w_c - w_i$ and the upper-sideband-frequency signal component at $w_c + w_i$. *See* TRIGONOMETRY.

$$x(t) = A_i \cos(w_i t) \cos(w_c t) = \frac{A_1}{2}\{\cos[(w_c - w_i)t] + \cos[(w_c + w_i)t]\} \quad (1)$$

A pair of phase shifters may now be introduced to operate on both the carrier and information signals to produce $\sin(w_c t)$ and $A_i \sin(w_i t)$, respectively. The product of this new signal pair is $y(t)$, given by Eq. (2), which exhibits sideband-frequency signal components at the same frequencies as $x(t)$. However, the sum of $x(t)$ and $y(t)$ is given by Eq. (3), which is the lower-sideband-frequency signal only, and the difference of $x(t)$ and $y(t)$ is given by Eq. (4), which is the upper-sideband-frequency signal only.

$$y(t) = A_i \sin(w_i t) \sin(w_c t) = \frac{A_1}{2}\{\cos[(w_c - w_i)t] - \cos[(w_c + w_i)t]\} \quad (2)$$

$$x(t) + y(t) = A_i \cos[(w_c - w_i)t] \quad (3)$$

$$x(t) - y(t) = A_i \cos[(w_c + w_i)t] \quad (4)$$

Implementation of this phase-shifting method to generate either an upper- or a lower-SSB signal requires a pair of linear modulators (**Fig. 2**), one to generate $x(t)$, the other to generate $y(t)$. The phase shifting of the information signal is generally performed by a wideband Hilbert transformer, a filter whose gain is $-j$ for positive frequencies and j for negative frequencies, where j is the square root of -1. The narrowband phase shifter for the carrier signal can be a simple second-order filter. However, such a filter may not be necessary because the carrier-frequency generator often provides both sine and cosine outputs simultaneously. Because the required transmission bandwidth is 50% less than that of DSB, SSB is an exemplary modulation system for bandwidth preservation. *See* SINGLE SIDEBAND.

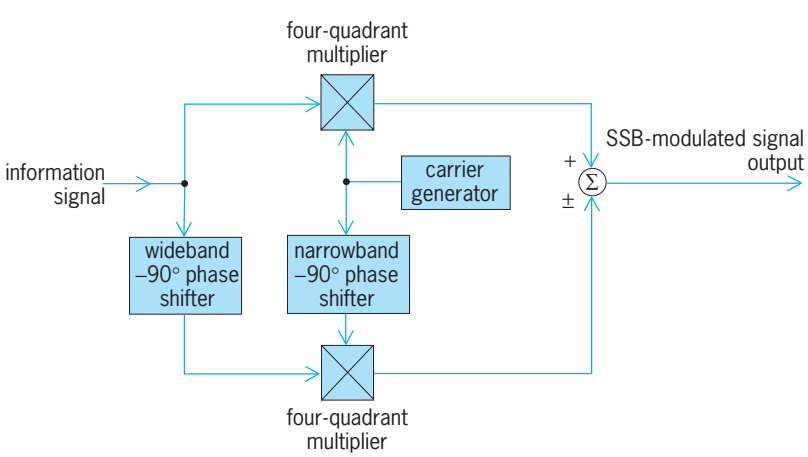

Fig. 2. Single-sideband (SSB) amplitude modulator.

Since the redundant sideband has been reduced (but not eliminated) in the generation of a VSB signal, its transmission efficiency is less than that of a SSB signal but greater than that of a DSB signal. The VSB signal can be generated in two steps: manufacture of a DSB signal, followed by the partial filtering out of a redundant sideband.

Carrier insertion. Receivers demand a carrier-signal reference to properly demodulate (detect) the transmitted signal. This reference may be provided in one of three ways: it may be transmitted with the modulated signal, for example, in double-sideband transmitted-carrier (DSTC or DSBTC) modulation, which is the method used for standard-broadcast AM radio; transmitted on a separate channel (often the case for instrumentation systems); or generated at the receiver.

Receivers for double-sideband signals can be very inexpensive if the carrier is transmitted with the information-bearing sidebands, as in the case of DSTC signals. In concept, the carrier may be inserted by simply adding it to the DSSC signal. For maximum transmission efficiency in generating the DSTC signal, the minimum amount of carrier should be transmitted. (Transmission efficiency may be defined as the ratio of total sideband power to the sum of sideband-plus-carrier power.) However, in order to avoid distortion in simple low-cost receivers, the amplitude of the inserted carrier must be at least equal to the peak amplitude, A, of the information signal. The ratio of the actual amplitude of the information signal to A is known as the modulation index, which should never exceed unity (**Fig. 3**). *See* AMPLITUDE-MODULATION DETECTOR; DISTORTION (ELECTRONIC CIRCUITS).

The transmitter can also be simplified by adding to the information signal a dc value equal to the peak magnitude of the information signal. The value of this modified information signal is now always positive. No longer is a four-quadrant multiplier required to modulate; rather, a far simpler two-quadrant multiplier generates the DSTC signal directly.

Mechanization. DSSC modulation, no matter how it is disguised, is just ordinary (linear) multiplication, or a reasonable approximation to that multiplication. Furthermore, as discussed above, a DSTC signal can be modeled as a DSSC signal with the carrier added. Any amplitude modulator is therefore simply some sort of embodiment of a linear multiplier

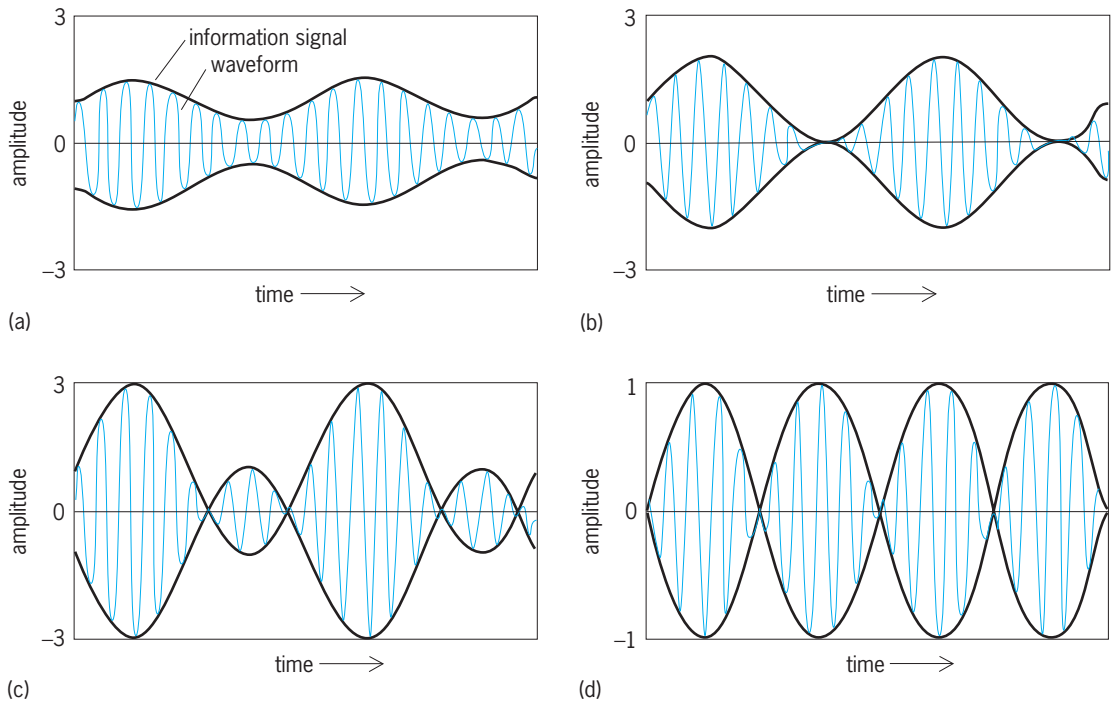

Fig. 3. Waveforms of a signal after various types of double-sideband modulation. Information signal envelopes are shown superimposed on the modulated waveforms. Double-sideband transmitted-carrier (DSTC) modulation with modulation indices of (*a*) 50%, (*b*) 100%, and (*c*) 200% are compared with (*d*) double-sideband suppressed-carrier (DSSC) modulation.

with or without a means to add the carrier.

High-level DSTC modulation is usually done by varying the power-supply voltage to the final high-power amplifier in the radio-frequency transmitter by summing the information signal with the dc output voltage of the power supply. Low-level DSTC modulation is carried out by integrated circuits that approximate the multiplications, followed by a linear amplifier. *See* AMPLIFIER; ELECTRONIC POWER SUPPLY.

Because of the falling costs and improving performance of digital devices, signals are now generally represented by digital data, that is, signal values which are sampled in time and encoded as numbers to represent their sampled values. Digital signal-processing (DSP) functions are carried out by some sequence of addition (or subtraction), multiplication, and delays. Efficient mechanization of each of these functions has been the subject of considerable effort. Digitally, waveform generation and linear multiplication are highly optimized processes. Advances in integrated-circuit fabrication techniques have so lowered the cost of digital circuits for modulation that the overwhelming majority of amplitude modulators in use are now digital. Custom application-specific integrated circuits (ASICs), general-purpose programmable DSP devices, and customizable arrays such as programmable logic arrays (PLAs) and field-programmable arrays (FPAs) are all in widespread use as amplitude modulators and demodulators. Digital modems as data transmission equipment dominate the production of amplitude modulators. *See* INTEGRATED CIRCUITS.

Quadrature amplitude modulation (QAM). It was observed above that the DSSC signal could be reduced to an SSB signal because of the redundant information in the second sideband. This redundancy can be exploited in a different way by quadrature modulation. In addition to the first information signal, $u(t)$, quadrature amplitude modulation acts on a second information signal, $v(t)$. The purpose of the process is to amplitude-modulate the two signals to the same frequency, w_c, with the goal of recovering each signal intact and uncontaminated by the other.

When a first DSSC signal, $u(t)\cos(w_c t)$, is demodulated with a first reference signal, $\cos(w_c t)$, the product signal given by Eq. (5) is obtained. After low-pass

$$u(t)[\cos(w_c t)]^2 = 0.5[1 + \cos(2w_c t)]u(t) \qquad (5)$$

filtering to eliminate the double-frequency term, $\cos(2w_c t)$, and scaling by 2 to offset the 0.5 gain, the first information signal, $u(t)$, is recovered. If the first received signal, $u(t)\cos(w_c t)$, had been demodulated with a second reference signal, $\sin(w_c t)$, the product signal given by Eq. (6) would have been ob-

$$u(t)\sin(w_c t)\cos(w_c t) = 0.5\sin(2w_c t)u(t) \qquad (6)$$

tained, and low-pass filtering this signal to eliminate the double-frequency term, $\sin(2w_c t)$, would leave nothing.

Similarly, when a second transmitted and received DSSC signal, $v(t)\sin(w_c t)$, is demodulated with the second reference signal, $\sin(w_c t)$, the product signal

given by Eq. (7) is obtained. After low-pass filter-

$$v(t)[\sin(w_c t)]^2 = 0.5[1 - \cos(2w_c t)]v(t) \qquad (7)$$

ing to eliminate the first double-frequency term, cos $(2w_c t)$, and scaling by 2 to offset the 0.5 gain, the second information signal, $v(t)$, is obtained. If the second received signal, $v(t)\sin(w_c t)$, had been demodulated with the first reference signal, $\cos(w_c t)$, the product signal given by Eq. (8) would have been

$$v(t)\sin(w_c t)\cos(w_c t) = 0.5\sin(2w_c t)v(t) \qquad (8)$$

obtained, and low-pass filtering this signal to eliminate the double-frequency term, $\sin(2w_c t)$, would again leave nothing.

The two carrier signals, $\sin(w_c t)$ and $\cos(w_c t)$, are 90° apart; they are also orthogonal, or are in quadrature. That is, the average value of their product is zero. An obvious deduction is that, if a third signal is transmitted that is the sum of the first and second transmitted signals of the discussion above, that is, $u(t)\cos(w_c t) + v(t)\sin(w_c t)$, and the received signal is applied to two parallel channels (named by convention I and Q), then it is possible to demodulate in the first (I) channel with the first reference signal, $\cos(w_c t)$, and in the second (Q) channel with the second reference signal, $\sin(w_c t)$, thereby recovering the first and second information signals, $u(t)$ and $v(t)$, respectively, from the I and Q channels.

The modulation method in the discussion above could be DSBTC if $u(t)$ and $v(t)$ were replaced by $u'(t) = K_1 + u(t)$ and $v'(t) = K_2 + v(t)$, where K_1 and K_2 are equal to or greater than the maximum amplitudes of $u(t)$ and $v(t)$, respectively, in order to keep the modulation indices positive.

In the above discussion, the first and second transmitted signals are both amplitude-modulated signals, but the modulating signals [$\cos(w_c t)$ and $\sin(w_c t)$] are in quadrature. The pair of information signals, $u(t)$ and $v(t)$, are therefore quadrature-amplitude-modulated (QAM) in the formation of the third transmitted signal. *See* AMPLITUDE MODULATION; MODULATION; MODULATOR.

Stanley A. White

Bibliography. R. C. Dorf (ed.), *The Electrical Engineering Handbook*, 2d ed., 1997; M. S. Roden, *Analog and Digital Communication Systems*, 3d ed., 1991; M. Valkenburg (ed.), *Reference Data for Engineers: Radio, Electronics, Computer, and Communications*, 8th ed., 1996.

Amylase

An enzyme which breaks down (hydrolyzes) starch, the reserve carbohydrate in plants, and glycogen, the reserve carbohydrate in animals, into reducing fermentable sugars, mainly maltose, and reducing nonfermentable or slowly fermentable dextrins. Amylases are classified as saccharifying (β-amylase) and as dextrinizing (α-amylases). The α- and β-amylases are specific for the α- and β-glucosidic bonds which connect the monosaccharide units into large aggregates, the polysaccharides. The α-amylases are found in all types of organs and tissues, whereas β-amylase is found almost exclusively in higher plants. *See* CARBOHYDRATE; ENZYME; GLYCOGEN.

Animals. In animals the highest concentrations of amylase are found in the saliva and in the pancreas. Salivary amylase is also known as ptyalin and is found in humans, the ape, pig, guinea pig, squirrel, mouse, and rat. Pig pancreas is rich in amylase, whereas cattle, sheep, and dog pancreases have lower concentrations.

Starch is one of the most important constituents of human food. The food prepared by the mouth for swallowing (the bolus) is converted by the gastric juices into chyme (a semiliquid paste). Chyme is passed through the pylorus into the duodenum, where intestinal digestion occurs. Part of this digestion is caused by pancreatic amylase, which, like ptyalin, hydrolyzes starch to maltose. A maltase also found in pancreatic juice hydrolyzes maltose to glucose. Glucose is picked up by the bloodstream for use in the tissues for respiration and for conversion to glycogen in the liver for storage.

Plants. Starch is broken down during the germination of seeds (rich in starch) by associated plant enzymes into sugars. These constitute the chief energy source in the early development of the plant. β-Amylase occurs abundantly in seeds and cereals such as malt. It also is found in yeasts, molds, and bacteria.

Industry. Amylase is also used in industry. It is used (1) in brewing and fermentation industries for the conversion of starch to fermentable sugars, (2) in the textile industry for designing textiles, (3) in the laundry industry in a mixture with protease and lipase to launder clothes, (4) in the paper industry for sizing, and (5) in the food industry for preparation of sweet syrups, to increase diastase content of flour, for modification of food for infants, and for the removal of starch in jelly production. The amylase enzyme used in industry comes from many sources: fungi, malt, bacteria, and the pancreas gland of cattle. *See* MALT BEVERAGE.

Daniel N. Lapedes

Amyloidosis

A disorder characterized by the accumulation of an unusual extracellular fibrous protein (amyloid) in the connective tissue of the body. The deposition of amyloid may be widespread, involving major organs and leading to serious clinical consequences, or it may be very limited with little effect on health.

The term amyloid originated with R. Virchow (1854), who mistakenly believed that this substance was akin to starch. Amyloid protein has now been found to have the following unique electron microscopic, x-ray diffraction, and biochemical characteristics. It appears as a homogeneous, eosinophilic extracellular substance when stained with hematoxylin and eosin. When it is stained with Congo red, green birefringence patterns are visible in the polarizing microscope. With the electron microscope, it appears in the form of fine, nonbranching fibrils 7–10 nanometers in diameter, and x-ray diffraction

studies show a cross beta pattern. Amino acid sequence studies demonstrate the presence of unusual proteins (AA, AL transthyretin, beta-2 microglobin, beta/A4 protein, and others).

Amyloidosis has been classified clinically as: (1) primary amyloidosis, with no evidence for preexisting or coexisting disease; (2) amyloidosis associated with multiple myeloma; (3) secondary amyloidosis, associated with chronic infections (such as osteomyelitis, tuberculosis, leprosy), chronic inflammatory disease (such as rheumatoid arthritis, ankylosing spondylitis, regional enteritis), or neoplasms (such as medullary carcinoma of the thyroid); (4) heredofamilial amyloidosis, associated with familial Mediterranean fever and a variety of heritable neuropathic, renal, cardiovascular, and other syndromes; (5) local amyloidosis, with local, often tumorlike, deposits in isolated organs without evidence of systemic involvement; (6) amyloidosis associated with aging in the heart or brain; (7) amyloid associated with hormonal disorders; and (8) amyloid of chronic hemodialysis.

Immunochemical findings. There is an immunochemical classification based on definition of the various types of proteins in amyloidosis: (1) amyloid associated with tissue deposition of protein AA, a tissue protein occurring in secondary amyloidosis and the amyloid of familial Mediterranean fever; and (2) amyloid associated with tissue deposition of protein AL, a protein consisting primarily of the amino-terminal variable segment of the light chain of homogeneous immunoglobulins, or in a few instances the whole light chain, which occurs in primary amyloid and amyloid associated with multiple myeloma. Transthyretin, the major protein of autosomal dominant inherited neuropathic amyloid, is also found in the amyloid lesions in the aged heart. Other chemical types also exist; for example, beta protein is found in the amyloid of Alzheimer's disease, IAPP (amylin) in that of adult onset (type 2) diabetes, and beta-2 microglobulin in the amyloid of chronic hemodialysis.

Antisera to protein AA react with an antigenically related component, serum AA (SAA), that is found in normal human serum, but has a higher molecular weight than AA and is regarded as the precursor of AA. Serum AA has several distinct characteristics: (1) It behaves as an acute phase reactant and is elevated in concentration in infection and inflammation. (2) It migrates as an apolipoprotein in the serum. (3) It is produced in the liver. In the heredofamilial amyloidosis, a variety of single point mutations have been found as the cause of the disorder; that is, the prototype is transthyretin with the substitution of methionine for valine at position 30.

A second component of amyloid, P-component (plasma component or pentagonal unit), with different ultrastructure, x-ray diffraction pattern, and chemical characteristics, has been isolated from amyloid and shown to be identical with a circulating human alpha globulin present in serum in minute amounts. It has certain similarities to C-reactive protein (CRP, an acute-phase protein in humans). However, it does not behave as an acute-phase protein in humans, and has biophysical and immunologic distinctions from CRP.

Clinical features. The different types of amyloid overlap considerably with regard to organ distribution. Since any organ may be involved, the clinical manifestations of amyloidosis are extremely varied. Infiltration of and about peripheral nerves, skin, tongue, joints, and heart is most frequent in the primary forms, while in secondary amyloidosis the liver, spleen, and kidney are the major sites of deposits.

Carpal tunnel syndrome (median neuropathy), resulting from local deposition of amyloid, is a frequent finding in certain types of primary amyloidosis and in the amyloidosis associated with myeloma and some heredofamilial forms. Some individuals with primary amyloidosis develop a coagulation disorder in which there is a deficiency of factor X (a procoagulant in plasma). This condition has been traced to the binding of these proteins to amyloid fibrils.

Amyloid in the heredofamilial (involving multiple siblings of successive generations) forms is responsible for peripheral neuropathy in certain families in Portugal, Japan, and Sweden; upper limb neuropathy in certain families of German or Swiss ancestry in the United States; cranial nerve neuropathy in a group of Finnish ancestry; and kidney disease and cardiovascular disease in others. All of the above conditions are inherited as autosomal dominant traits except for the amyloid associated with familial Mediterranean fever, which is inherited in an autosomal recessive manner.

Clinical features in the diagnosis of primary amyloidosis include unexplained proteinuria, peripheral neuropathy, enlargement of the tongue, enlargement of the heart, intestinal malabsorption, bilateral or familial carpal tunnel syndrome, or low blood pressure when standing upright. Symptoms frequently exist for several years before correct diagnosis. When primary amyloidosis is suspected, evidence of an underlying disease is investigated to rule out possible inflammatory disorders or malignant tumors. In cases of chronic infectious or inflammatory diseases, secondary amyloidosis is suspected if there is unexplained proteinuria, or enlargement of the liver or spleen.

The diagnosis is established by means of microscopic examination of appropriate biopsy specimens. Subcutanous abdominal fat or tissue from the rectum have proved to be particularly useful, and in selected cases, skin, gingival, kidney, and liver biopsy specimens are valuable. Since a biopsy specimen may contain only a small quantity of amyloid, it is important to look for the characteristic green birefringence on polarization microscopy after Congo red staining.

Treatment. There is no specific treatment for amyloidosis, but supportive treatment is very useful. Some individuals with secondary renal amyloidosis have improved after the cure of a predisposing disease, such as active tuberculosis or chronic osteomyelitis. Kidney transplants have been carried out successfully in a few cases of renal failure. Colchicine

has been shown to relieve the acute attacks of familial Mediterranean fever and may decrease the formation of amyloid. Combinations of melphalan, prednisone, and colchicine are used in primary (AL) amyloidosis, and liver transplantation has been introduced in heredofamilial amyloid to remove the site of synthesis of the mutant protein. Alan S. Cohen

Bibliography. A. S. Cohen, Amyloidosis, *Bull. Rheumatic Dis.*, 40:1–12, 1991; A. S. Cohen, Amyloidosis, *N. Engl. J. Med.*, 277:522–530, 1967; J. R. Harris (ed.), *Electron Microscopy of Proteins,* vol. 3, 1982; J. B. Natvig et al., *Amyloid and Amyloidosis*, 1990.

Anaerobic infection

An infection caused by anaerobic bacteria (organisms that are intolerant of oxygen). Most such infections are mixed, involving more than one anaerobe and often aerobic or facultative bacteria as well.

Anaerobes are prevalent throughout the body as indigenous flora, and virtually all anaerobic infections arise endogenously, the principal exception being *Clostridium difficile* colitis. Factors predisposing to anaerobic infection include those disrupting mucosal or other surfaces (trauma, surgery, and malignancy or other disease), those lowering redox potential (impaired blood supply, tissue necrosis, and growth of nonanaerobic bacteria), drugs inactive against anaerobes (such as aminoglycosides), and virulence factors produced by the anaerobes (toxins, capsules, and collagenase, hyaluronidase, and other enzymes).

Anaerobic gram-negative bacilli (*Bacteroides, Prevotella, Porphyromonas, Fusobacterium*) and anaerobic gram-positive cocci (*Peptostreptococcus*) are the most common anaerobic pathogens. *Clostridium* (spore formers) may cause serious infection. The prime pathogen among gram-positive nonsporulating anaerobic bacilli is *Actinomyces*.

Infections commonly involving anaerobes and the incidence of anaerobes in such infections are noted in the **table**. In terms of frequency of occurrence, the oral and dental pleuropulmonary, intraabdominal, obstetric-gynecologic, and skin and soft tissue infections are most important. Foul odor is definitive evidence of involvement of anaerobes in an

Infections commonly involving anaerobic bacteria

Infections	Incidence, %	Proportion of cultures positive for anaerobes yielding only anaerobes
Bacteremia	20	4/5
Central nervous system		
Brain abscess	89	1/2 to 2/3
Extradural or subdural empyema	10	
Head and neck		
Chronic sinusitis	52	4/5*
Chronic otitis media	33 to 56	0 to 1/10
Neck space infections	100	3/4
Wound infection following head and neck surgery	95	0
Dental, oral, facial		
Orofacial, of dental origin	94	4/10
Bite wounds	47	1/34
Thoracic		
Aspiration pneumonia	62 to 100	1/3 to 1/2†
Lung abscess	85 to 93	1/2 to 3/4
Empyema (nonsurgical)	62 to 76	1/3 to 1/2
Abdominal		
Intraabdominal infection (general)	81 to 94	1/10 to 1/3
Appendicitis with peritonitis	96	1/100
Liver abscess	52	1/3
Other intraabdominal infection (postsurgery)	93	1/6
Biliary tract	41 to 45	2/117 to 0
Obstetric-gynecologic		
Miscellaneous types	71 to 100	1/3
Pelvic abscess	88	1/2
Vulvovaginal abscess	75	1/4
Vaginal cuff abscess	98	1/30
Septic abortion, sepsis	63 to 67	
Pelvic inflammatory disease	25 to 48	1/14 to 1/7
Soft tissue and miscellaneous		
Nonclostridial crepitant cellulitis	75	1/12
Pilonidal sinus infection	73+	
Diabetic foot ulcers	95	1/20
Soft-tissue abscesses	60	1/4
Cutaneous abscesses	62	1/5
Decubitus ulcers with bacteremia	63	
Osteomyelitis	40	1/10

*23/28 cultures (82%) yielding heavy growth of one or more organisms had only anaerobes present.
†Aspiration pneumonia occurring in the community, rather than in the hospital, involves anaerobes to the exclusion of aerobic or facultative forms two-thirds of the time.
SOURCE: S. M. Finegold, Antimicrobial therapy of anaerobic infections: A status report, *Hosp. Pract.,* 14:17–81, October 1979.

infection, but absence of such odor does not exclude the possibility. Other important clues to anaerobic infection, though not specific, are location near a mucosal surface; bite infection; tissue necrosis; gas formation; prior aminoglycoside, quinolone, or trimethoprim/sulfamethoxazole therapy; "sulfur granules" (actinomycosis); certain distinctive clinical presentations (such as gas gangrene); septic thrombophlebitis; unique morphology of certain anaerobic organisms on smear; and failure to grow organisms aerobically from specimens. To document anaerobic infection properly, specimens for culture must be obtained so as to exclude normal flora and must be transported under anaerobic conditions. *See* GANGRENE; GAS.

Therapy includes surgery (debridement, drainage) and antimicrobial agents. *Bacteroides fragilis* is the most resistant anaerobe; only chloramphenicol, carbapenems, beta lactam/beta lactamase inhibitor combinations, and metronidazole are essentially active against all strains. Other useful drugs include clindamycin, cefoxitin, ceftizoxime, and cefotetan. *See* ANTIBIOTIC; INFECTION. Sydney M. Finegold

Bibliography. S. M. Finegold and W. L. George (eds.), *Anaerobic Infections in Humans*, 1989; P. N. Levett (ed.), *Anaerobic Microbiology: A Practical Approach*, 1992; T. Willis and L. Phillips, *Anaerobic Infections: Clinical and Laboratory Practice*, 1991.

Analcime

A mineral with a framework structure in which all the aluminosilicate tetrahedral vertices are linked, thus allying it to the feldspars, feldspathoids, and zeolites. The structure is cubic, space group Ia3d, $\alpha =$ 1.371 mm. Its formula is $Na(H_2O)[AlSi_2O_6]$; in this sense it is a tectosilicate.

The analcime structure type includes several other mineral species. These include high-temperature leucite, β-$K[AlSi_2O_6]$; pollucite, $Cs(H_2O)$-$[AlSi_2O_6]$; and wairakite, $Ca(H_2O)[AlSi_2O_6]_2$. Temperature-dependent order-disorder relationships between AlO_4 and SiO_4 tetrahedra lead to modifications of lower structural symmetry, birefringent optical properties, and complex composite twinning about a single crystal nucleus. Crystals are most often trapezohedra (see **illus.**); rarely the mineral is massive granular. Hardness is 5–5.5 on Mohs scale; specific gravity is 2.27.

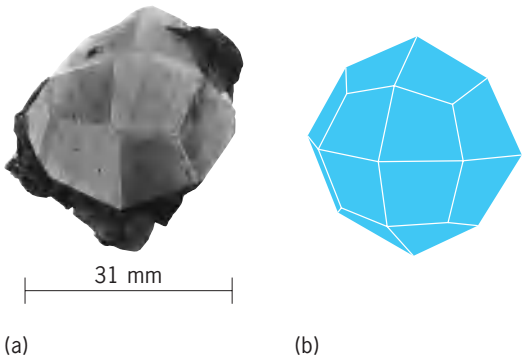

Analcime, (*a*) Specimen from Keweenaw County, Michigan (*American Museum of Natural History*). (*b*) Trapezohedral crystal typical of analcime (*after C. Klein, Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*).

Analcime most frequently occurs as a low-temperature mineral in vesicular cavities in basalts, where it is associated with zeolites (particularly natolite), datolite, prehnite, and calcite. Small grains are frequent constituents of sedimentary rocks and muds in oceanic basins associated with volcanic sources. Noted localities for large crystals include the Bay of Fundy region, Nova Scotia; the Watchung basalts of New Jersey; and the Keweenaw Peninsula in northern Michigan. Analcime also occurs in certain hydrothermal sulfide deposits as a fissure mineral. *See* FELDSPAR; FELDSPATHOID; SILICATE MINERALS; ZEOLITE. Paul B. Moore

Analgesic

Any of a group of drugs of diverse chemical structure and physiological effects which are commonly used for the relief of pain. To qualify as an analgesic a drug must selectively reduce or abolish pain without causing impairment of consciousness, mental confusion, incoordination or paralysis, or other derangements of the nervous system.

Narcotic alkaloids. The oldest and best-known analgesics are opium, a drug obtained by extracting the juice of the poppy seed, and its most active alkaloid, morphine. Morphine and related drugs reduce or block the activation of pain neurons in the gray matter of the spinal cord, and at receptor sites in the brainstem and thalamus. The mechanisms by which this inhibition is effected are complex. Morphine does not act directly on the pain-mediating neurons. Instead, it involves an elaborate pain-modulating system in the posterior horns of the spinal cord, the dorsal-ventral reticular formation of the medulla, the periaqueductal gray matter, and the hypothalamus and thalamus. Separate neurons in these regions have opioid receptors on their surface, and the strength of this affinity between drug and neuron is proportional to the analgesic potency of the opiate (or opioid). All parts of this neuronal system acting locally or by way of descending connections inhibit pain transmission cells. In these same regions there are also neurons that elaborate a naturally occurring (endogenous) group of opioid peptides whose function is to modulate or suppress pain. They are called enkephalins, and the most potent of them is beta-endorphin. Since they are abundant not only in these sites but also in the gut, in sympathetic neurons, and in chromafin cells of the adrenal glands, they are believed to account for the natural inhibition of pain by strong emotion (stress), and expectancy of pain relief by nonanalgesic drugs (placebo effect). *See* ENDORPHINS; MORPHINE ALKALOIDS; OPIATES.

In addition to their use as analgesic drugs, opiates have other biological effects such as sedation,

pupillary constriction, suppression of cough reflex, respiratory depression, reduction of intestinal motility, impairment of segmental flexor reflexes, and decrease in body temperature. This functional diversity is attributed to the activation of other inhibitory systems of neurons. *See* SEDATIVE.

While morphine is the most powerful medical analgesic substance, there are many other naturally occurring alkaloids derived from opium. The best known of these is codeine.

Common to all opiates is the attribute that if they are taken for weeks or months the recipient will need larger doses to obtain the same analgesic and sedative effects. This response is called tolerance. If the drug is stopped, disagreeable withdrawal or abstinence effects are experienced within hours to days. There is severe pain, sweating, salivation, hyperventilation, restlessness, and confusion. These abstinence symptoms, which are marks of habituation, pressure the addicted person to take extreme measures to obtain the narcotic in order to avoid the symptoms. *See* ADDICTIVE DISORDERS; ALKALOID; NARCOTIC.

Synthetics. Because of the strong addictive properties of morphine and related compounds, chemists have synthesized other drugs of similar chemical structure, in the hope of securing analgesia without addiction. This effort has been only partially successful. Methadone, a drug that has been given to addicts as a substitute for morphine, is an effective analgesic when taken orally and is less addictive than morphine. Meperidine (Demerol) is a strong synthetic analgesic but definitely addictive. Other synthetic analgesics are oxycodone (Percodan), levorphanol (levodromoran), propoxyphene (Darvon), and pentazocine (Talwin). The last two of this series cause little or no addiction but, unfortunately, are not strong analgesics. Another synthetic drug, Naloxone, blocks the analgesic effect of all opiate agonists and precipitates withdrawal symptoms in addicted individuals.

Salicylates. Another class of analgesic drugs, which are nonnarcotic (nonaddictive), are the salicylates, the most familiar being acetysalicylic acid (aspirin), and salicylatelike drugs such as phenylbutazone (Butazolidine), indomethacin (Indocin), acetaminophen, and phenacetin. These drugs are most effective in relieving skeletal pain due to inflammation (such as arthritis). Their analgesic properties, which are not nearly as strong as those of morphine and the synthetic opioids, are due to their action on both the peripheral and central nervous system. Peripherally they block the synthesis of prostaglandins in the inflamed tissues and thereby prevent the sensitization of pain receptors. Centrally they act in some obscure way on the hypothalamus. These drugs also have many other effects, such as reducing fever (antipyrexia) and preventing platelet agglutination. They are the most commonly used of all analgesic medications and are often combined with caffeine or a barbiturate sedative under a variety of trade names and sold for the relief of headache, backache, and so forth. *See* ASPIRIN; EICOSANOIDS; NERVOUS SYSTEM (VERTEBRATE); PAIN.

Raymond D. Adams

Analog computer

A computer or computational device in which the problem variables are represented as continuous, varying physical quantities. An analog computer implements a model of the system being studied. The physical form of the analog may be functionally similar to that of the system, but more often the analogy is based solely upon the mathematical equivalence of the interdependence of the computer variables and the variables in the physical system. *See* SIMULATION.

Types. An analog computer is classified either in accordance with its use (general- or specific-purpose) or based on its construction (hydraulic, mechanical, or electronic). General-purpose implies programmability and adaptability to different applications or the ability to solve many kinds of problems. Most electronic analog computers were general-purpose systems, either real-time analog computers in which the results were obtained without any significant time-scale changes, or high-speed repetitive operation computers. The latter utilized time compression typically in the range 1000:1 to 10,000:1, completing a computation or simulation in a few milliseconds and enabling the display of results as flicker-free images.

Special-purpose analog computers contain fixed programs permitting few or no adjustments. They are generally built into physical systems where they serve typically as a subsystem simulator, function controller or results analyzer. For example, the pneumatic computer shown in **Fig. 1** used air bellows and flapper nozzles to generate accurate multiplication, division, squaring, or square-root functions of input signals, encoding data as air pressures.

Use. Large electronic analog computer systems with many hundreds of operational amplifiers were widely used from the early 1960s to the mid-1980s. They solved extremely complex and extensive sets of differential equations (mathematical models) such as six-degree-of freedom space flights, exothermal chemical reaction kinetics, control systems for food processing plants, and the human immunosuppressive system. During the 1970s, they were often paired with a digital computer in hybrid systems

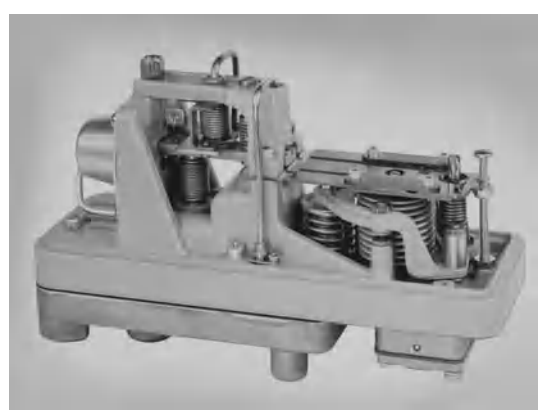

Fig. 1. Special-purpose pneumatic analog computer, model Foxboro 556. (*Foxboro Co.*)

Fig. 2. Analog-hybrid laboratory equiped with a HYSHARE 600 system. This system can include up to six analog computer consoles and one or more digital processors for multiuser, multitask applications. (*Electronic Associates, Inc.*)

(**Fig. 2**). The digital computer then primarily controlled the setting up and running of the analog computers, collecting and checking results and often providing logic and initial values to allow extensive searching through vast multidimensional problem domains. In some instances, hybrid systems also entailed automatic patching (interconnecting) of analog computing units.

Smaller analog computers, with fewer than 100 operational amplifiers, were extensively used for education and training purposes. Models of a theoretical nature, expressed by sets of differential equations, or practical engineering simulations based on empirical data and heuristic models often lent themselves to study by analog computers. The ease by which the user could interact with the modeled system was a major feature, since learning is greatly enhanced when the user safely can explore, experiment with, and even misuse the modeled system.

Digital equivalents. Since the 1970s, digital computer programs have been developed which essentially duplicate the functionality of the analog computer. Modern simulation languages, such as ACSL, GASP, GPSS, SLAM, and Simscript, have replaced electronic analog computers. They provide nearly the same highly interactive and parallel solution capabilities of electronic analog computers, but without the technical shortcomings of electronics: accuracy inherently limited to 0.01%, effective bandwidths of 1 MHz, and cumbersome and time-consuming programming. Simulation languages also avoid the large purchase investments and the continual maintenance dependencies of complex electronic systems.

Digital multiprocessor analog system. Another type of analog computer is the digital multiprocessor analog system (**Fig. 3**), in which the relatively slow speeds of sequential digital increment calculations have been radically boosted through parallel processing. In this type of analog computer it is possible to retain the programming convenience and data storage of the digital computer while approximating the speed, interaction potential, and parallel computations of the traditional electronic analogs.

The digital multiprocessor analog computer typically utilizes several specially designed high-speed processors for the numerical integration functions, the data (or variable) memory distributions, the arithmetic functions, and the decision (logic and control) functions. All variables remain as fixed or floating-point digital data, accessible at all times for computational and operational needs.

The digital multiprocessor analog computer achieves an overall problem-solving efficiency comparable to the very best continuous electronic analog computers, at a substantially lower price. An example of such a computer, the model AD10 (Fig. 3*a*), can solve large, complex, and multivariate problems at very high speeds and with the advantages of all-digital hardware. Its computation system (Fig. 3*b*) is based on five parallel processors working in a special computer architecture designed for high-speed operation. The various elements are interconnected by means of a data bus (MULTIBUS). A highly interleaved data memory of up to 10^6 words serves the data and program storage functions. The five processors working in parallel are: the control processor (COP), which controls all operations; the arithmetic processor (ARP), which runs the numerical

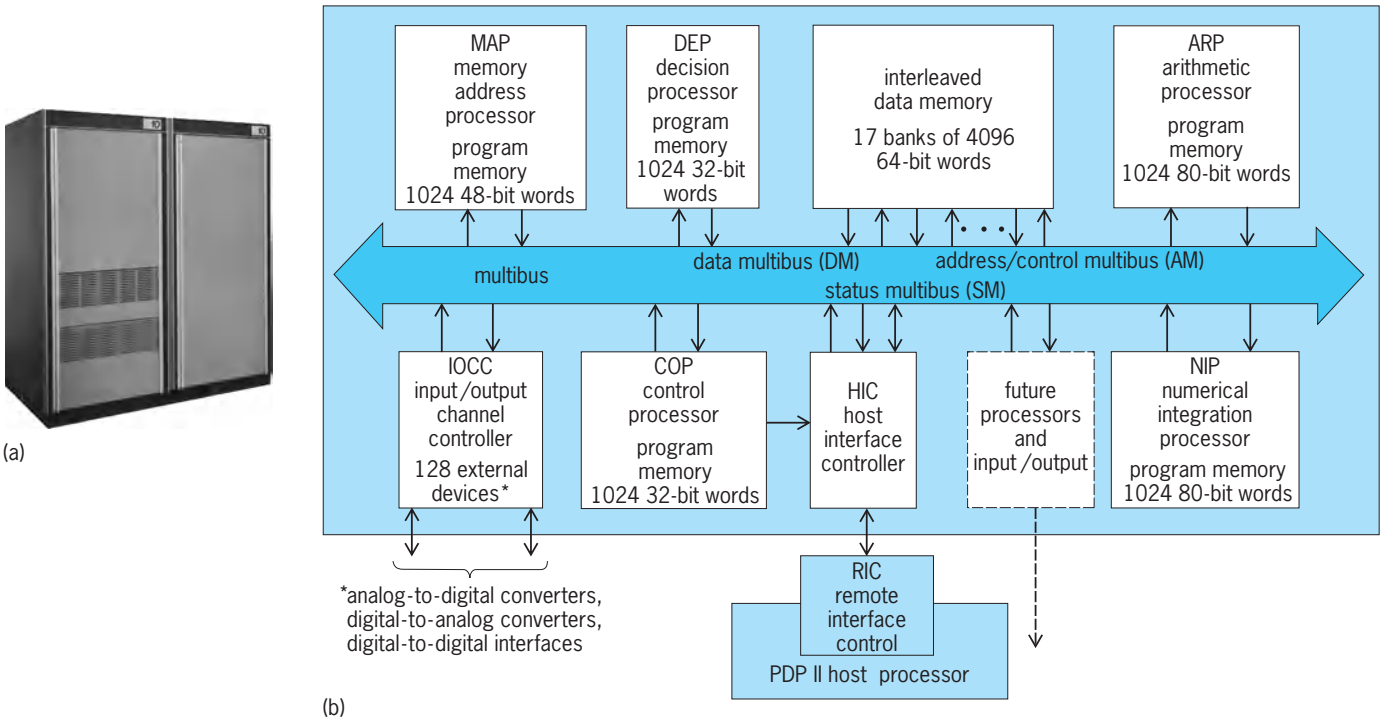

Fig. 3. Multiprocessor analog computer. (*a*) Exterior. (*b*) Organization. (*Applied Dynamics International*)

calculations; the decision processor (DEP), which executes the logic parts of the program; the memory address processor (MAP), which makes sure that all data are fetched and stored efficiently; and the numerical integration processor (NIP), which carries out the integration functions that are crucially important in an analog computer.

Description. The typical electronic general-purpose analog computer consists of a console containing a collection of operational amplifiers; computing elements, such as summing networks, integrator networks, attenuators, multipliers, and function generators; logic and interface units; control circuits; power supplies; a patch bay; and various meters and display devices. The patch bay is arranged to bring input and output terminals of all programmable devices to one location, where they can be conveniently interconnected by various patch cords and plugs to meet the requirements of a given problem. Prewired problem boards can be exchanged at the patch bay in a few seconds and new coefficients set up typically in less than a half hour. Extensive automatic electronic patching systems have been developed to permit fast setup, as well as remote and time-shared operation.

The analog computer basically represents an instrumentation of calculus, in that it is designed to solve ordinary differential equations. This capability lends itself to the implementation of simulated models of dynamic systems. The computer operates by generating voltages that behave like the physical or mathematical variables in the system under study. Each variable is represented as a continuously varying (or steady) voltage signal at the output of a programmed computational unit. Specific to the analog computer is the fact that individual circuits are used for each feature or equation being represented, so that all variables are generated simultaneously. Thus the analog computer is a parallel computer in which the configuration of the computational units allows direct interactions of the computed variables at all times during the solution of a problem.

Unique features. The unique features of the analog computer which are of value in science and technology are as follows:

1. Within the useful frequency bandwidth of the computational units and components, all programmed computations take place in parallel and are for practical purposes instantaneous. That is, there is no finite execution time associated with each mathematical operation, as is encountered with digital computer methods.

2. The dynamics of an analog model can be programmed for time expansion (slower than real system time), synchronous time (real time), or time compression (faster than real time).

3. The computer has a flexible addressing system so that almost every computed variable can be measured, viewed with vivid display instruments, and recorded at will.

4. One control mode of the computer, usually called HOLD, can freeze the dynamic response of a model to allow detailed examination of interrelationships at that instant in time, and then, after such study, the computing can be made to resume as though no stop had occurred.

5. By means of patch cords, plugs, switches, and adjustment knobs the analog computer program or model can be directly manipulated, especially during dynamic computation, and the resultant changes

in responses observed and interpreted.

6. Because of the fast dynamic response of the analog computer, it is easy to implement implicit computation through the use of problem feedback. (This important and powerful mathematical quality of the analog computer is discussed more fully below.)

7. The computer can be used for on-line model building; that is, a computer model can be constructed in a step-by-step fashion directly at the console by interconnecting computational units on the basis of one-for-one analog representation of the real system elements. Then, by adjusting signal gains and attenuation parameters, dynamic behavior can be generated that corresponds to the desired response or is recognizable as that of the real system. This method allows a skillful person to create models when no rigorous mathematical equations for a system exist.

8. For those applications to which it is well suited, the analog computer operates at relatively low cost, thus affording the analyst ample opportunity to investigate, develop, and experiment within a broad range of parameters and functions.

Components. Manipulations of the signals (voltages) in the analog computer are based upon the fundamental electrical properties associated with circuit components. The interrelation of voltages for representing mathematical functions is derived by combining currents at circuit nodes or junctions. *See* KIRCHHOFF'S LAWS OF ELECTRIC CIRCUITS; OHM'S LAW.

The simplest arrangement of components for executing addition would be to impress the voltages to be added across individual resistors (**Fig. 4***a*). The resistors would then be joined (Fig. 4*b*) to allow the currents to combine and to develop the output voltage across a final resistor. Use of this simple configuration of elements for computation is impractical, because of the undesirable interaction between inputs. A change in one input signal (voltage) causes a change in the current that flows through the input resistor; this changes the voltage at the input resistor junction, and the change secondarily causes a different current to flow in the other input resistors. The situation gets more interactive when another computing circuit is attached so that part of the summing current flows away from the summing junction. This interaction effect also prevents exact computing. If, in some way, each voltage to be summed could be made independent of the other voltages connected to the summing junction, and if the required current fed to other circuits could be obtained without loading the summing junction, then precise computation would be possible. *See* DIRECT-CURRENT CIRCUIT THEORY.

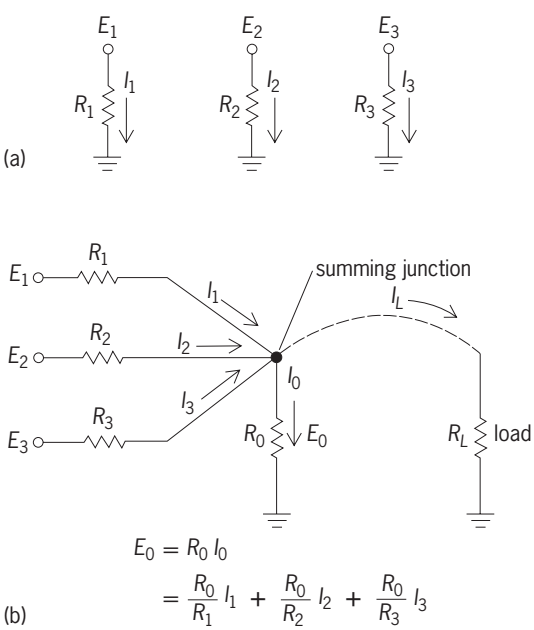

Fig. 4. Addition of electric currents by using a passive network of resistors. (*a*) Individual voltages and resistors. (*b*) Voltages and corresponding currents summed into a common resistor.

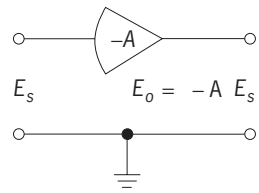

Fig. 5. Symbol for a high-gain direct-current amplifier, with one inverting input referenced to ground.

The electronic analog computer satisfies these needs by using high-gain (high-amplification) dc operational amplifiers. A symbol to represent a dc direct-coupled amplifier is shown in **Fig. 5**. According to convention, the rounded side represents the input to the amplifier, and the pointed end represents the amplifier output. A common reference level or ground exists between the amplifier input and output, and all voltages are measured with respect to it. The ground reference line is understood and is usually omitted from the symbol. The signal input network (consisting of summing resistors) connects to the inverting (or negative) amplifier input terminal. The noninverting (or positive) amplifier input terminal is normally connected to the reference ground. Generally the inverting input is called the summing junction (SJ) of the amplifier. Internal design of the amplifier is such that, if the signal at the summing junction is positive with respect to ground, the amplifier output voltage is negative with respect to ground. The amplifier has an open-loop voltage gain of $-A$; therefore an input voltage of E_s results in an output voltage of $-AE_s$. Gain of a commercial computing amplifier is typically 10^8; thus, an input voltage of less than 1 μV can produce several volts at the amplifier output. *See* AMPLIFIER; DIRECT-COUPLED AMPLIFIER.

Because the operational amplifier thus inverts the signal (changes its sign or polarity), it lends itself to the use of negative feedback, whereby a portion of the output signal is returned to the input. This arrangement has the effect of lowering the net gain of the amplifier signal channel and of improving

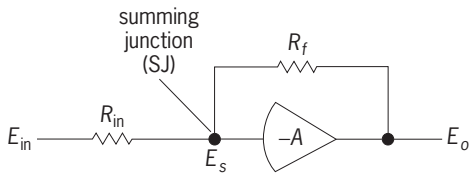

Fig. 6. High-gain amplifier which has been made into an operational amplifier through the inclusion of an input resistor and a feedback resistor tied together at the amplifier's summing junction (SJ).

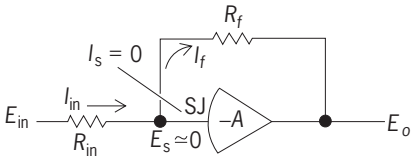

$$I_{in} = I_f \qquad (1)$$

$$\frac{E_{in}}{R_{in}} = -\frac{E_o}{R_f} \qquad (2)$$

$$\boxed{E_o = -\frac{R_f}{R_{in}} E_{in}} \qquad (3)$$

Fig. 7. Summing junction currents into an operational amplifier create the fixed gain function, determined by resistor values, as indicated in the box.

overall signal-to-noise ratio and increasing computational accuracy.

Circuit operation (**Fig. 6**) can be viewed in the following manner. A dc voltage E_{in} applied to input resistor R_{in} produces a summing junction voltage E_s. The voltage is amplified and appears at the amplifier output as voltage E_o (equal to $-AE_s$, where A is voltage gain of the amplifier). Part of output voltage E_o returns through feedback resistor R_f to the summing junction. Because the returned or feedback voltage is of opposite (negative) polarity to the initial voltage at the summing junction, it tends to reduce the magnitude of E_s, resulting in an overall input-output relationship that may be expressed as Eq. (1). In fact,

$$\frac{E_o}{E_{in}} = \frac{-A}{A+1} = \frac{-1}{1+1/A} \simeq -1 \quad A > 10^8 \qquad (1)$$

the summing junction voltage E_s is so small that it is considered to be at practically zero, a condition called virtual ground.

To illustrate how the operational amplifier serves the needs of the computing network, consider the currents that flow and the corresponding algebraic expressions (**Fig. 7**). The operational amplifier is designed to have high input impedance (high resistance to the flow of current into or out of its input terminal); consequently the amplifier input current I_s can then be considered to be practically zero. The resulting current equation states that input current I_{in} is equal to feedback current I_f. Since the amplifier has a very high gain, the summing junction voltage is virtually zero. Voltage drop across R_{in} is thus equal to E_{in}; voltage drop across R_f is E_o. The equation in the box is the fundamental relationship for the amplifier. As long as the amplifier has such a high gain and requires a negligible current from the summing junction, the amplifier input and output voltages are related by the ratio of the two resistors and are thus not affected by the actual electronic construction of the amplifier. If several input resistors are connected to the same summing junction and voltages are applied to them (**Fig. 8**), then because the summing junction remains at practically zero potential, none of the inputs will interfere with the other inputs. Thus all inputs will exert independent and additive effects on the output.

Because amplifier gain has a negative sign, output voltage equals the negative sum of the input voltages, weighted according to the values of the individual resistors in relation to the feedback resistor, as shown in the box in Fig. 8.

When a computing circuit is connected to the amplifier output, a demand for load current is introduced. Such a required output load current must be supplied by the amplifier without a measurable change in its output voltage. That is, the operational amplifier must act as a voltage controller, supplying whatever current is required within limits, while maintaining the net voltage gain established by the mathematical ratio of its input and feedback elements. The operational amplifier and network is a linear network because once the input-feedback ratios are adjusted, signal channel gains remain constant (a straight-line function) during computation.

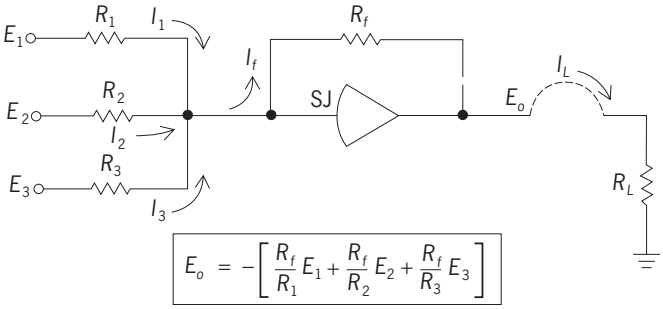

Fig. 8. Operational amplifier which has been made into a summer through the use of several input resistors connected to the summing junction. Output is equal to the inverted weighted sum of the inputs.

Programming. To solve a problem using an analog computer, the problem solver goes through a procedure of general analysis, data preparation, analog circuit development, and patchboard programming. He or she may also make test runs of subprograms to examine partial-system dynamic responses before eventually running the full program to derive specific and final answers. The problem-solving procedure typically involves eight major steps, listed below:

1. The problem under study is described with a set of mathematical equations or, when that is not possible, the system configuration and the interrelations of component influences are defined in block-diagram form, with each block described in terms of black-box input-output relationships.

2. Where necessary, the description of the system (equations or system block diagram) is rearranged

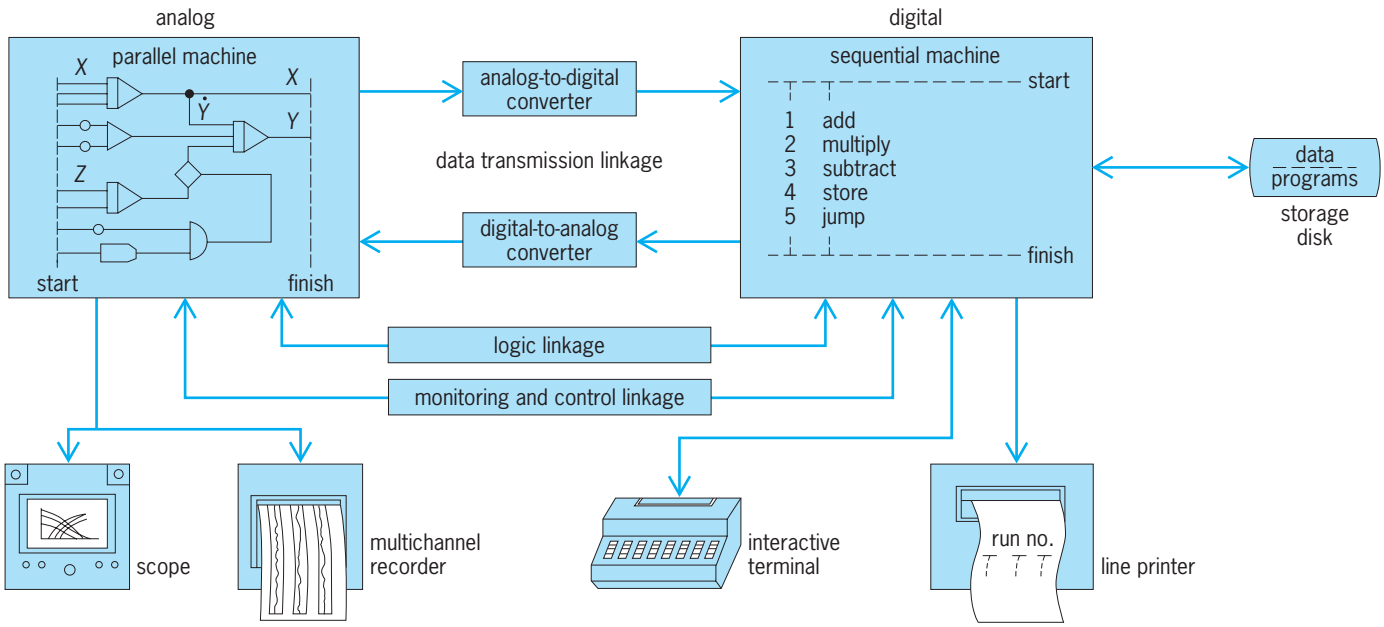

Fig. 9. Block diagram representation of a hybrid analog-to-digital computer.

in a form that may better suit the capabilities of the computer, that is, avoiding duplications or excessive numbers of computational units, or avoiding algebraic (nonintegrational) loops.

3. The assembled information is used to sketch out an analog circuit diagram which shows in detail how the computer could be programmed to handle the problem and achieve the objectives of the study.

4. System variables and parameters are then scaled to fall within the operational ranges of the computer. This may require revisions of the analog circuit diagram and choice of computational units.

5. The finalized circuit arrangement is patched on the computer problem board.

6. Numerical values are set up on the attenuators, the initial conditions of the entire system model established, and test values checked.

7. The computer is run to solve the equations or simulate the black boxes so that the resultant values or system responses can be obtained. This gives the initial answers and the "feel" for the system.

8. Multiple runs are made to check the responses for specific sets of parameters and to explore the influences of problem (system) changes, as well as the behavior which results when the system configuration is driven with different forcing functions.

Hybrid computers. The accuracy of the calculations on a digital computer can often be increased through double precision techniques and more precise algorithms, but at the expense of extended solution time, due to the computer's serial nature of operation. Also, the more computational steps there are to be done, the longer the digital computer will take to do them. On the other hand, the basic solution speed is very rapid on the analog computer because of its parallel nature, but increasing problem complexity demands larger computer size. Thus, for the analog computer the time remains the same regardless of the complexity of the problem, but the size of the computer required grows with the problem.

Interaction between the user and the computer during the course of any calculation, with the ability to vary parameters during computer runs, is a highly desirable and insight-generating part of computer usage. This hands-on interaction with the computed responses is simple to achieve with analog computers. For digital computers, interaction usually takes place through a computer keyboard terminal, between runs, or in an on-line stop-go mode. An often-utilized system combines the speed and interaction possibilities of an analog computer with the accuracy and programming flexibility of a digital computer. This combination is specifically designed into the hybrid computer.

A modern analog-hybrid console (Fig. 2) contains the typical analog components plus a second patchboard area to include programmable, parallel logic circuits using high-speed gates for the functions of AND, NAND, OR, and NOR, as well as flip-flops, registers, and counters. The mode switches in the integrators are interfaced with the digital computer to permit fast iterations of dynamic runs under digital computer control. Data flow in many ways and formats between the analog computer with its fast, parallel circuits and the digital computer with its sequential, logic-controlled program (**Fig. 9**). Special high-speed analog-to-digital and digital-to-analog converters translate between the continuous signal representations of variables in the analog domain and the numerical representations of the digital computer. Control and logic signals are more directly compatible and require only level and timing compatibility. *See* ANALOG-TO-DIGITAL CONVERTER; BOOLEAN ALGEBRA; DIGITAL-TO-ANALOG CONVERTER.

Programming of hybrid models is more complex than described above, requiring the user to consider

the parallel action of the analog computer interlaced with the step-by-step computations progression in the digital computer. For example, in simulating a space mission, the capsule control dynamics will typically be handled on the analog computer in continuous form, but interfaced with the digital computer, where the navigational trajectory is calculated. *See* COMPUTER.

Per A. Holst

Bibliography. R. M. Howe, *Design Fundamentals of Analog Computer Components*, 1961; G. A. Korn, *Electronic Analog and Hybrid Computers*, 1971; D. M. MacKay et al., *Analog Computing at Ultra-High Speed*, 1962; R. Tomovic, *Repetitive Analog Computers*, 1958.

Analog states

States in neighboring nuclear isobars that have the same total angular momentum, parity, and isotopic spin. They also have nearly identical nuclear structure wave functions except for the transformation of one or more neutrons into an equivalent number of protons, which occupy the same single-particle states as the neutrons. Analog states (or isobaric analog states, IAS) have been observed throughout the periodic table, indicating that isotopic spin is a good quantum number. *See* ANGULAR MOMENTUM; I-SPIN; ISOBAR (NUCLEAR PHYSICS); NUCLEAR STRUCTURE; PARITY (QUANTUM MECHANICS).

Since the nucleon–nucleon interaction has been found to be approximately charge-independent, it is possible to consider protons and neutrons as representing different charge states of a single particle, that is, a nucleon. Thus, a level (commonly referred to as a parent state) in a nucleus with Z protons and N neutrons can be expected to have an analog in the neighboring isobar with $Z + 1$ protons and $N - 1$ neutrons (and the same total number of nucleons, $A = Z + N$), where the protons and neutrons occupy the same orbits as those in the parent state. The energy difference between the parent and analog states predominantly arises from the increased contribution from the electrostatic Coulomb interaction to the total energy arising from the extra proton in the analog state. From this amount must be subtracted the neutron-proton mass difference of 0.782 MeV (energies are given on the atomic mass scale). The agreement between such calculated energies of analog states and their measured values is in general fairly precise but not exact. The reason is that small additional factors influence the level energies, such as electromagnetic effects, a small charge-dependent nuclear interaction, isospin mixing, and nuclear structure effects.

Isotopic spin. Since the nucleon has spin $s = {}^1/_2$, it obeys Fermi statistics. Charge independence of nuclear forces then implies that systems of two nucleons that can be described by the same wave functions that are antisymmetric in both space and spin would have identical energies. There are three such combinations of two-nucleon systems that can satisfy these conditions, namely, neutron-neutron, proton-proton, and neutron-proton. The isotopic spin (also called isobaric spin or isospin) quantum number $t = {}^1/_2$ has been introduced to distinguish between the two states of the nucleon, and provides a means to further identify these three states. The neutron is associated with a projection of t along the direction of quantization, that is, the z axis, with $t_z = {}^1/_2$, and the proton with $t_z = -{}^1/_2$. The total z-projection of isospin, T_z, is given by the sum of the t_z of the individual nucleons. Hence, $T_z = 1, 0, -1$ for the neutron-neutron, neutron-proton, and proton-proton systems, respectively, and these represent an isobaric triplet with each system having total isospin $T = 1$. Only one two-nucleon system exists with a wave function that is symmetric in space and spin, that is, the neutron-proton combination for which $T = T_Z = 0$. For a nucleus with N neutrons and Z protons, $T_z = {}^1/_2\,(N - Z)$ and, as with angular momentum, the value of isospin T must be equal to or greater than its z-projection, T_z. Analog states that differ solely in the interchange of either one neutron or proton (neighboring isobars) have values of T_z that differ by ± 1. Those that differ by the interchange of two neutrons or protons have T_z that differ by ± 2, and so forth. *See* EXCLUSION PRINCIPLE; FERMI-DIRAC STATISTICS.

Nuclei with N = Z. Nuclei with $N = Z$ (equal numbers of neutrons and protons) have $T_z = 0$. Hence, there can be no analog for states in such nuclei that also have total isospin $T = 0$. This prohibition follows because, as noted above, T_z changes by ± 1 between neighboring isobars which must then have a total isospin T equal to or greater than 1. As a result, there is no analog of the ground state for these nuclei. However, excited states in $N = Z$ nuclei for which T is greater than zero can have analogs.

Nuclei with N > Z. Nuclei with $N > Z$ (neutron excess) have $T_z = {}^1/_2(N - Z)$, which is equal to or greater than ${}^1/_2$. Therefore, all states in these nuclei have $T \geq {}^1/_2$, and in principle, they can thus have analogs. For heavy nuclei where the Coulomb energy becomes quite large, a question arose as to whether this large energy might destroy the isobaric symmetry. However, this question was resolved in 1961 with the observation of analog states in heavy nuclei by experiments that utilized the (p, n) reaction. *See* NUCLEAR REACTION.

Coulomb displacement energy. The energy required to transform a neutron into a proton in a nucleus with Z protons and N neutrons is called the Coulomb displacement energy, ΔE_c. This energy can be calculated in terms of the energies of the parent and analog states from Eq. (1), where M_{Z+1} is the energy of the

$$\Delta E_c = M_{Z+1} - M_Z + \Delta_{np} \qquad (1)$$

analog state, M_Z the energy of the parent state, and $\Delta_{np} = 0.782$ MeV (as noted above) is the neutron-proton mass difference (all energies are given on the atomic mass scale). The difference in the analog energies, that is, $M_{Z+1} - M_Z$, can be expressed in terms of the Q-value for beta decay between the ground states of the isobars, and the energy of excitation of the analog state in the $Z + 1$ nucleus. From analyses of the abundant data on analog states

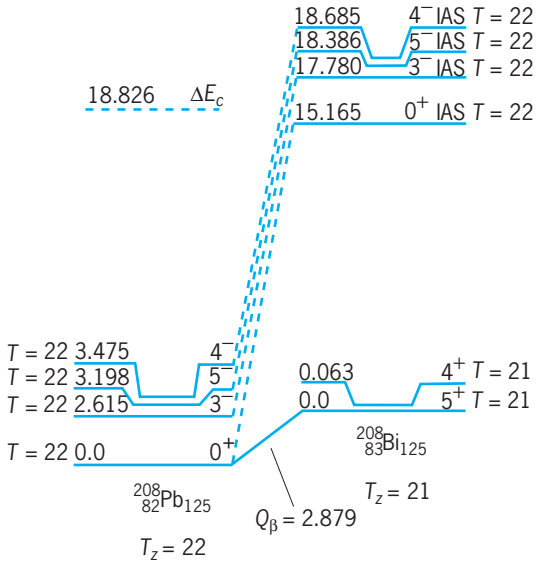

Partial level diagram of the ^{208}Pb and ^{208}Bi isobar system. Broken lines connect parent–analog state pairs. Energies of excited states are given relative to the ground states, respectively, in megaelectronvolts. The isospin, *T*, and the angular momentum and parity, J^{π}, are shown for each level.

throughout the periodic table, it is found that the Coulomb displacement energy can be represented to a high degree of accuracy by Eq. (2), where $\bar{Z} =$

$$\Delta E_c = 1.412(\bar{Z}/A^{1/3}) - 0.861 \text{ (MeV)} \qquad (2)$$

$Z + {}^1/_2$. The Coulomb displacement energy represents the minimum energy of excitation in the parent nucleus for which states with $T = T_Z + 1$ can exist.

Mass 208 system. A partial level diagram for the ^{208}Pb and ^{208}Bi isobar system demonstrates the energy relationships discussed above (see **illus.**). In such a diagram, only a few of the many observed levels are shown. Below an excitation energy of $\Delta E_c =$ 18.826 MeV in ^{208}Pb, all levels have $T = T_z = 22$; the Coulomb displacement energy ΔE_c is the minimum excitation at which $T = 23$ states can occur in ^{208}Pb. The energy difference between the ^{208}Pb and ^{208}Bi ground states is $Q_\beta = 2.879$ MeV. The *z*-component of isospin for ^{208}Bi is $T_z = 21$, which is also the total isospin, *T*, for all levels up to an excitation energy of 15.165 MeV. At this energy, a level (actually a resonance) is observed in ^{208}Bi which possesses characteristics that are similar to those of the ground state of ^{208}Pb, and is identified as the analog. The energy of excitation of the analog state lies in the nuclear continuum of the ^{208}Bi mass system, that is, well above the energy at which ^{208}Bi is unstable to particle decay. As a result, the analog state is accessible only by means of some nuclear reaction. It was first identified in low-energy studies of proton scattering on ^{207}Pb, and later observed in the (*p, n*) reaction on ^{208}Pb. Schematically, in the (*p, n*) reaction an energetic proton impinges upon a ^{208}Pb nucleus and knocks out one of the excess neutrons that occupy the single-particle orbits that lie above those occupied by the protons, and the incident proton is captured into an equivalent orbit. This event is known as a charge-exchange reaction (here, a neutron is transformed into a proton), and is related to the inverse reaction that occurs in positron beta decay. *See* RADIOACTIVITY.

Fermi strength. Analog states are connected by either pure Fermi or mixed Fermi and Gamow-Teller transitions. The selection rules for Fermi transitions are $\Delta J = 0$, $\Delta\pi = 0$, $\Delta T = 0$, and $\Delta T_z = \pm 1$, where *J* represents the total angular momentum and π represents the parity. Those for Gamow-Teller transitions are $\Delta J = 0$ or ± 1 (but transitions between analogs with $J = 0$ are forbidden), $\Delta\pi = 0$, $\Delta T =$ 0 or ± 1, and $\Delta T_z = \pm 1$. Both types of transitions require $\Delta L = 0$, where *L* represents the total orbital angular momentum. *See* SELECTION RULES (PHYSICS).

Of special interest in the study of nuclear structure is the distribution of strength for various types of transitions based upon the ground state. For Fermi transitions, all of the strength is concentrated in the transition between the analog states. The total strength is proportional to the isospin matrix element between the analog states, which is given by Eq. (3), where $|TT_z\rangle$ represents the isospin part of the

$$|\langle TT_z - 1\,|T_-|\,TT_z\rangle|^2 = (T + T_z)(T - T_z + 1) \qquad (3)$$

parent wave function, and T_- is an operator which changes a neutron into a proton. Since $T = T_z$, the total Fermi strength is proportional to $(N - Z)$, which is just equal to the number of excess neutrons. This outcome reflects the fact that each of the excess neutrons occupies an orbit (with specific space and spin descriptions) that lies at an energy above those occupied by the protons, and thus there is equal probability for each excess neutron to be transformed into a proton and to participate in the Fermi transition.

Since the strength is well defined in practice, Fermi transitions can be used to obtain information pertaining to the reaction mechanisms that effect them. For beta decay, this approach usually involves determination of the vector coupling constant, G_V, as well as fine details about the nuclear wave functions. In charge-exchange studies, knowledge is sought pertaining to details of the effective nucleon–nucleus interaction and, to some extent, details of the overlap between the spatial wave functions of the parent and analog states.

Other applications. The study of analog states provides important information used to test nuclear theories. For example, the double charge-exchange reactions (π^+, π^-) [where π^+ and π^- represent a pion with positive and negative charge, respectively] have been used to identify double isobaric analog states, that is, analogs in isobars removed by 2 charge units with $T_z = T - 2$. Such data have been useful for testing various formulas for predicting the relative masses of isobaric multiplets. Single and double charge-exchange reactions utilizing incident pions have also been used to investigate giant resonances built upon analog states. The single-particle structure of parent states can be studied by

observing the particle decay of the analog state when the decay resides in the nuclear continuum. Measurements of the widths of analog states provide information pertaining to their fragmentation, for example, their mixing with states of the same spin and parity but with total isospin lower by one unit, that is, equal to $T - 1$. *See* GIANT NUCLEAR RESONANCES.

Daniel J. Horen

Bibliography. M. S. Antony, J. Britz, and A. Pape, Coulomb displacement energies between analog levels for $44 \leq A \leq 239$, *Atom. Nucl. Data Tables*, 40:9–56, 1988; M. S. Antony et al., Isobaric mass equation for $A = 1$–45 and systematics of Coulomb displacement energies, *Atom. Nucl. Data Tables*, 33:447–478, 1985; A. Bohr and B. R. Mottelson, *Nuclear Structure*, vol. 1, 1969, vol. 2, 1975; D. H. Wilkinson (ed.), *Isospin in Nuclear Physics*, 1969.

Analog-to-digital converter

A device for converting the information contained in the value or magnitude of some characteristics of an input signal, compared to a standard or reference, to information in the form of discrete states of a signal, usually with numerical values assigned to the various combinations of discrete states of the signal.

Analog-to-digital (A/D) converters are used to translate analog information, such as audio signals or measurements of physical variables (for example, temperature, force, or shaft rotation), into a form suitable for digital handling, which might involve any of these operations: (1) processing by a computer or by logic circuits, including arithmetical operations, comparison, sorting, ordering, and code conversion, (2) storage until ready for further handling, (3) display in numerical or graphical form, and (4) transmission.

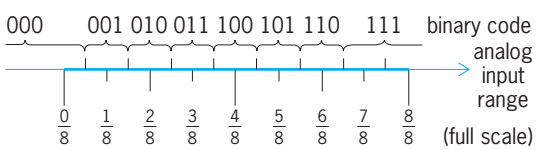

Fig. 1. A three-bit binary representation of a range of input signals.

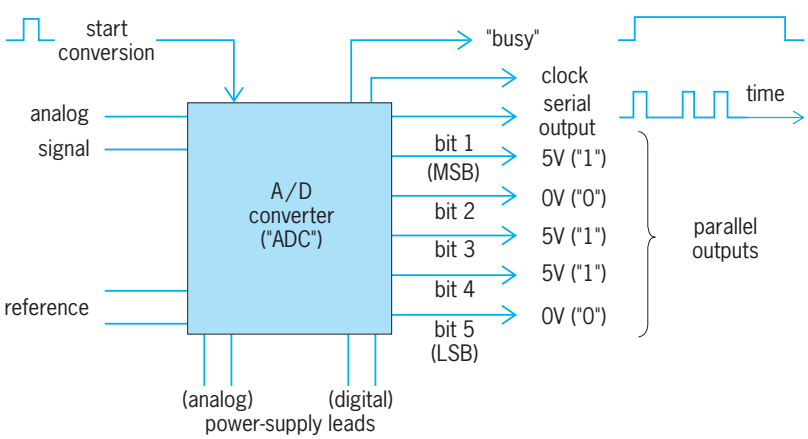

Fig. 2. Basic diagram of an analog-to-digital converter, showing parallel and serial (return-to-zero) output formats for code 10110.

If a wide-range analog signal can be sampled and converted, with sufficient frequency, to an appropriate number of two-level digits, or bits, to represent each sample or snapshot, the digital representation of the signal can be transmitted through a noisy medium without essential degradation of the fine structure of the original signal. *See* COMPUTER GRAPHICS; DATA COMMUNICATIONS; DIGITAL COMPUTER.

Concepts and structure. Conversion involves quantizing and encoding. Quantizing means partitioning the analog signal range into a number of discrete quanta and determining to which quantum the input signal belongs. Encoding means assigning a unique digital code to each quantum and determining the code that corresponds to the input signal. The most common system is binary, in which there are 2^n quanta (where n is some whole number), numbered consecutively; the code is a set of n physical two-valued levels or bits (1 or 0) corresponding to the binary number associated with the signal quantum. [The term binary is used here in two senses: a binary (radix-2) numbering system, and the binary (two-valued) decisions forming each bit.] *See* BIT.

Figure 1 shows a typical three-bit binary representation of a range of input signals, partitioned into eight (that is, 2^3) quanta. For example, a signal in the vicinity of 3/8 full scale (between 5/16 and 7/16) will be coded 011 (binary 3). *See* NUMBERING SYSTEMS.

Conceptually, the conversion can be made to take place in any kind of medium: electrical, mechanical, fluid, optical, and so on (for example, shaft-rotation-to-optical); but by far the most commonly employed form of A/D converters comprises those devices that convert electrical voltages or currents to coded sets of binary electrical levels (for example, +5 V or 0 V) in simultaneous (parallel) or pulse-train (serial) form, as shown in **Fig. 2**. Serial outputs are not always made available in the form of binary numbers.

The converter depicted in Fig. 2 converts the analog input to a five-digit "word." If the coding is binary, the first digit (most significant bit, abbreviated MSB) has a weight of 1/2 full scale, the second 1/4 full scale, and so on, down to the nth digit (least-significant bit, abbreviated LSB), which has a weight of 2^{-n} of full scale (1/32 in this example). Thus, for the output word shown, 10110, the analog input must be given approximately by the equation below.

$$\frac{16}{32} + \frac{0}{32} + \frac{4}{32} + \frac{2}{32} + \frac{0}{32} = \frac{22}{32} = \frac{11}{16}\text{FS (full scale)}$$

The number of bits, n, characterizes the resolution of a converter. The **table** translates bits into other conventional measures of resolution in a binary system.

Figure 2 also shows a commonly used configuration of connections to an A/D converter: the analog signal and reference inputs; the parallel and serial

Binary resolution equivalents*

Bit	2^{-n}	$1/2^n$ (fraction)	dB (decibels)	$1/2^n$ (decimal)	%	Parts per million
FS†	2^0	1	0	1.0	100	1,000,000
MSB‡	2^{-1}	1/2	−6	0.5	50	500,000
2	2^{-2}	1/4	−12	0.25	25	250,000
3	2^{-3}	1/8	−18.1	0.125	12.5	125,000
4	2^{-4}	1/16	−24.1	0.0625	6.2	62,500
5	2^{-5}	1/32	−30.1	0.03125	3.1	31,250
6	2^{-6}	1/64	−36.1	0.015625	1.6	15,625
7	2^{-7}	1/128	−42.1	0.007812	0.8	7,812
8	2^{-8}	1/256	−48.2	0.003906	0.4	3,906
9	2^{-9}	1/512	−54.2	0.001953	0.2	1,953
10	2^{-10}	1/1024	−60.2	0.0009766	0.1	977
11	2^{-11}	1/2048	−66.2	0.00048828	0.05	488
12	2^{-12}	1/4096	−72.2	0.00024414	0.024	244
13	2^{-13}	1/8192	−78.3	0.00012207	0.012	122
14	2^{-14}	1/16,384	−84.3	0.000061035	0.006	61
15	2^{-15}	1/32,768	−90.3	0.0000305176	0.003	31
16	2^{-16}	1/65,536	−96.3	0.0000152588	0.0015	15
17	2^{-17}	1/131,072	−102.3	0.00000762939	0.0008	7.6
18	2^{-18}	1/262,144	−108.4	0.000003814697	0.0004	3.8
19	2^{-19}	1/524,288	−114.4	0.000001907349	0.0002	1.9
20	2^{-20}	1/1,048,576	−120.4	0.0000009536743	0.0001	0.95

* From D. H. Sheingold (ed.), *Analog-Digital Conversion Handbook,* 3d ed., Analog Devices, Inc., 1986.
† Full scale.
‡ Most significant bit.

digital outputs; the leads from the power supply, which provides the required energy for operation; and two control leads—a start-conversion input and a status-indicating output (busy) that indicates when a conversion is in progress. Serial-output converters often furnish a clock output pulse to indicate the time at which a bit is available. The reference voltage or current is often developed within the converter.

Second in importance to the binary code and its many variations is the binary-coded decimal (BCD), which is used when the encoded material is to be displayed or printed in numerical form. In BCD, each digit of a radix-10 number is represented by a four-digit binary subgroup. For example, the BCD code for 379 is 0011 0111 1001. The output of the A/D converter used in digital panel meters is usually BCD.

Techniques. There are many techniques used for A/D conversion, ranging from simple voltage-level comparators to sophisticated closed-loop systems, depending on the input level, output format, control features, and the desired speed, resolution, and accuracy. The three most popular techniques, used individually or in combination, are direct or flash conversion; integrating conversion, exemplified by dual slope; and feedback conversion, exemplified by the method of successive approximations.

Comparator. The most elementary A/D converter is the comparator, a 1-bit (2-value) converter. It has two inputs, the signal and the reference: whenever the signal is greater than the reference, the output is 1; when less, the output is 0. A digitally controlled latch may be included to lock in the value of the comparison at a definite time.

Direct conversion. The fastest, and conceptually simplest, approach to n-bit conversion is via the direct or flash converter. It quantizes by using $2^n - 1$ comparators to compare the input signal simultaneously to a set of reference signals representing all amplitude quanta. Then it employs a logical scheme called a priority encoder to determine the parallel digital output corresponding to a given set of comparator output states.

Figure 3 is a simplified block diagram of a 2-bit (4-level) flash converter. The resistor string acts as a

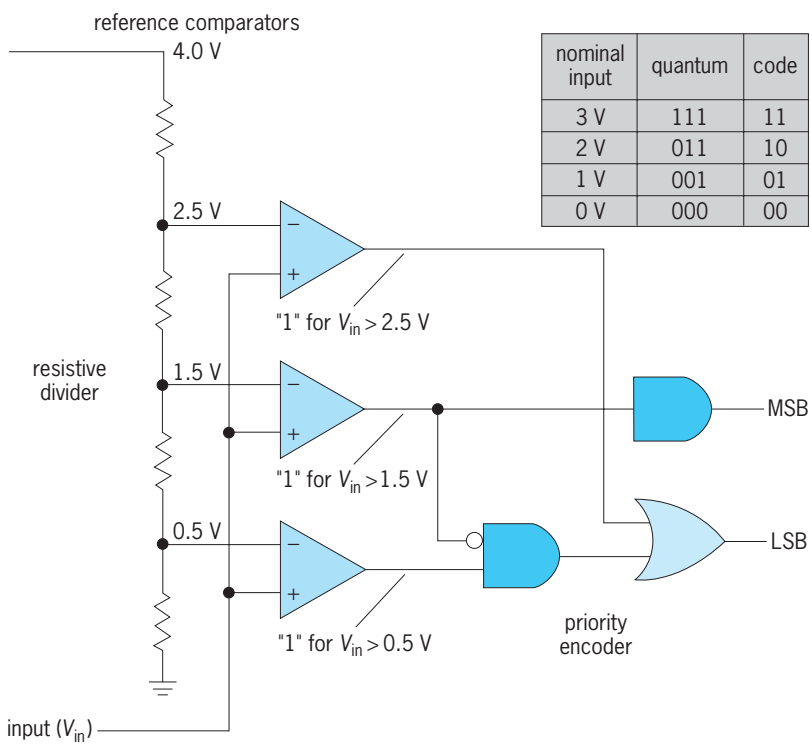

nominal input	quantum	code
3 V	111	11
2 V	011	10
1 V	001	01
0 V	000	00

Fig. 3. Diagram of a 2-bit flash converter.

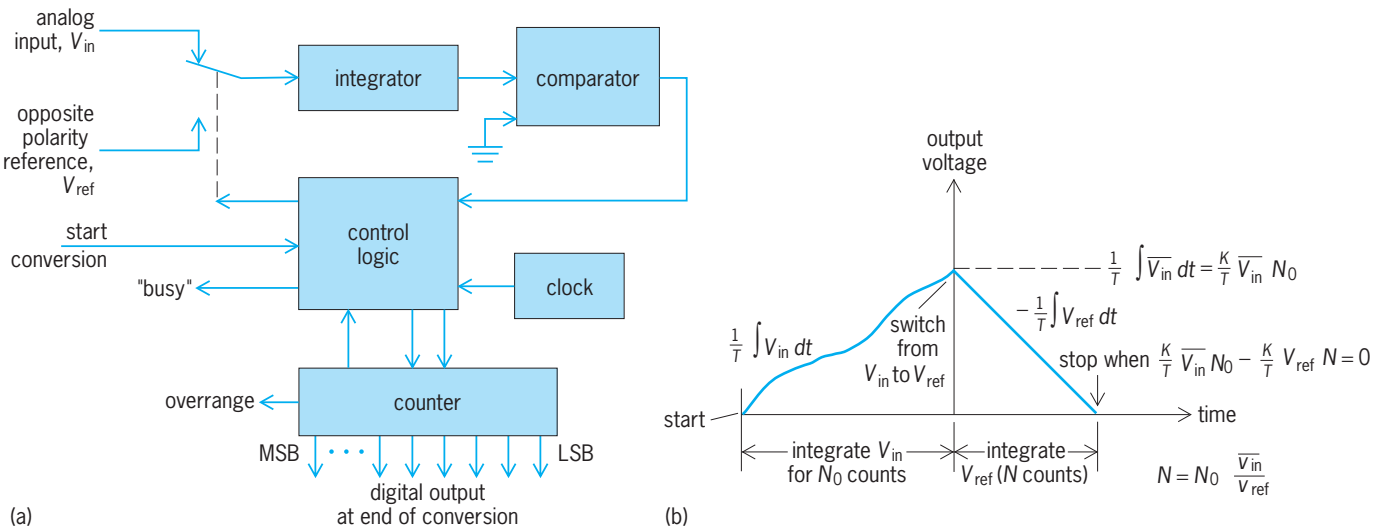

Fig. 4. Example of a dual-slope conversion. (*a*) Block diagram of converter. (*b*) Integrator output. Here, *k* is a constant, and *T* is the *RC* (resistor-capacitor) time constant of the integrator.

voltage divider to provide reference voltages for the three (that is, $2^2 - 1$) comparators. When the common input is below the lowest reference level, all of the comparators have 0 output. When the input is greater than the reference for the lowest comparator, its output switches to 1; when the next comparator's reference is exceeded, its output too switches to 1; and when the third comparator's reference level is exceeded, all three comparators are at 1. Thus, the length of the series of 1's identifies the input voltage's quantum.

The quantum must now be encoded as a 2-bit digital word at the output. In the priority-encoder scheme shown in Fig. 3, the output is a binary word, and the correspondence between quantum and code is given in the table.

Practical flash converters require lengthy resistor strings, large numbers of comparators, and rather complicated encoders. For example, a 10-bit flash converter calls for 1023 comparators and requires encoding 1024 levels into a 10-bit word. Nevertheless, in integrated-circuit form, all of the circuity exists on a single monolithic silicon chip.

Integrating conversion. Unlike flash converters, which measure the input instant by instant, integrating converters use analog integrating circuits in a variety of circuit architectures to convert average values of the input into trains of pulses, which are then encoded by digital counters or processors. Integrating converters have high resolution and low noise sensitivity, and the train of pulses can be transmitted as a high-level noise-insensitive signal for encoding

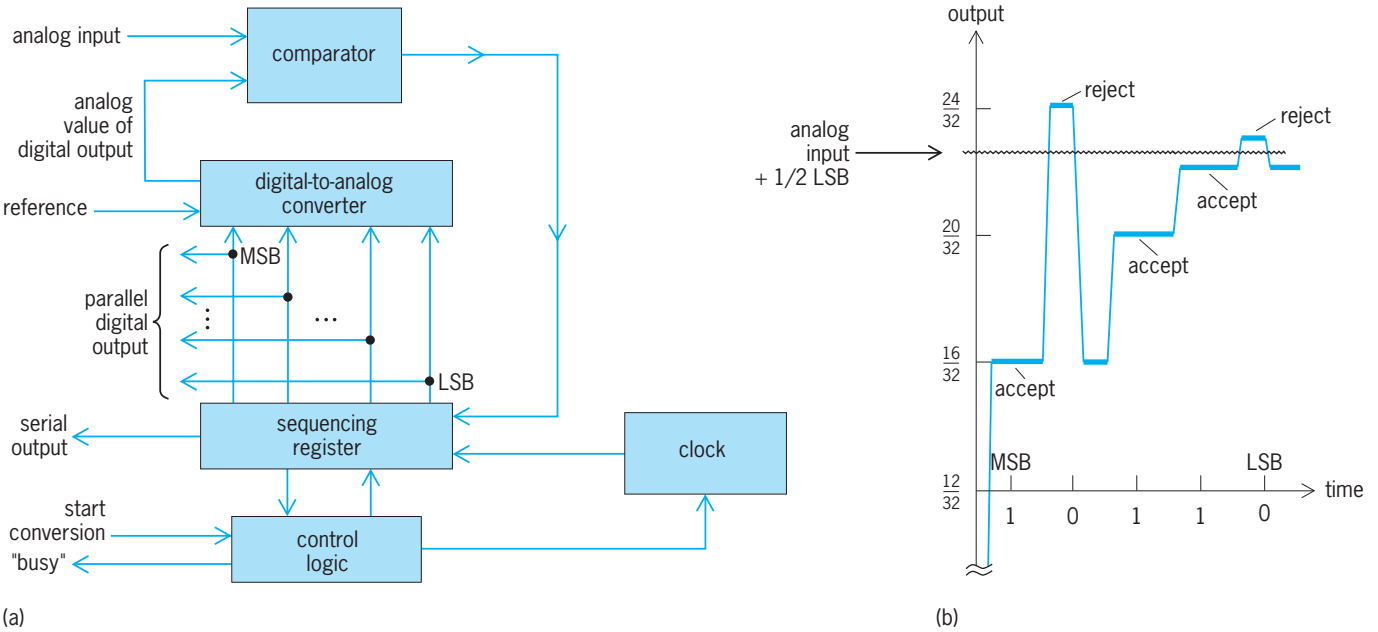

Fig. 5. Successive-approximations conversion. (*a*) Block diagram of converter. (*b*) Digital-to-analog converter output for the example in Fig. 2.

at a remote location. The class includes voltage-to-frequency converters, delta-modulation types, and dual-slope converters.

The dual-slope converter, discussed here as an example of the class, has long been popular in digital voltmeters; it can perform complete conversions with high accuracy at low speeds, generally a few conversions per second. **Figure 4***a* is a simplified block diagram of a dual-slope converter. The input is integrated for a period of time determined by a clock-pulse generator and counter (Fig. 4*b*). The final value of the signal integral becomes the initial condition for integration of the reference in the opposite sense, while the clock output is counted. When the net integral is zero, the count stops. Since the integral "up" of the input over a fixed time (N_0 counts) is equal to the integral "down" of the fixed reference, the ratio of the number of counts of the variable period to that of the fixed period is equal to the ratio of the average value of the signal to the reference.

Feedback conversion. Some A/D converter types employ an *n*-bit digital-to-analog converter to compare an analog version of the digital output with the analog input; the decision (higher or lower) causes a change in the digital output to bring its value closer to the input value. The change may be made by a counter that increases or decreases the digital value in small steps, or by the educated guessing of a successive-approximations converter.

Successive-approximations conversion combines superior speed and accuracy; it is the principal technique used for data-acquisition and computer-interface systems. **Figure 5***a* is a simplified block diagram of a successive-approximations converter. In a manner analogous to the operation of an apothecary's scale with a set of binary weights, the input is weighed against a set of successively smaller fractions of the reference, produced by a digital-to-analog (D/A) converter that reflects the number in the output register. *See* DIGITAL-TO-ANALOG CONVERTER.

First, the MSB is tried (1/2 full scale). If the signal is less than the MSB, the MSB code is returned to zero; if the signal is equal to or greater than the MSB, the MSB code is latched in the output register (Fig. 5*b*). The second bit is tried (1/4 full scale). If the signal is less than 1/4 or 3/4, depending on the previous choice, bit 2 is set to zero; if the signal is equal to or greater than 1/4 or 3/4, bit 2 is retained in the output register. The third bit is tried (1/8 full scale). If the signal is less than 1/8, 3/8, 5/8, or 7/8, depending on previous choices, bit 2 is set to zero; otherwise, it is accepted. The trial continues until the contribution of the least-significant bit has been weighed and either accepted or rejected. The conversion is then complete. The digital code latched in the output register is the digital equivalent of the analog input signal.

Physical electronics. The earliest A/D converters were large rack-panel chassis-type modules using vacuum tubes, requiring about 1.4 ft^3 (0.04 m^3) of space and many watts of power. Since then, they have become smaller in size and cost, evolving through circuit-board, encapsulated-module, and hybrid construction, with improved speed and resolution. Single-chip 12-bit A/D converters with the ability to interface with microprocessors are now available in small integrated-circuit packages. Integrated-circuit A/D converters, with 8-bit and better resolution and conversion rates of hundreds of megahertz, are also commercially available. *See* MICROPROCESSOR.

Daniel H. Sheingold

Bibliography. Analog Devices, Inc. Engineering Staff, *Analog-Digital Conversion Handbook*, ed. by D. H. Sheingold, 3d ed., 1986; M. Demler, *High-Speed Analog to Digital Conversion*, 1991; D. F. Hoeschele, Jr., *Analog-to-Digital and Digital-to-Analog Conversion Techniques*, 2d ed., 1994; R. Van de Plassche, *Integrated Analog-to-Digital and Digital-to-Analog Converters*, 1994.

Analysis of variance

Total variation in experimental data is partitioned into components assignable to specific sources by the analysis of variance. This statistical technique is applicable to data for which (1) effects of sources are additive, (2) uncontrolled or unexplained experimental variations (which are grouped as experimental errors) are independent of other sources of variation, (3) variance of experimental errors is homogeneous, and (4) experimental errors follow a normal distribution. When data depart from these assumptions, one must exercise extreme care in interpreting the results of an analysis of variance. Statistical tests indicate the contribution of the components to the observed variation.

In an illustrative experiment, t methods of treatment are under study, and n samples are measured for each treatment for a total of nt samples. Measurement X_{ij} of the ith sample that received the jth treatment records an overall effect μ, an effect β_j produced by the jth treatment, and an effect ϵ_{ij} produced by experimental error. The three effects are additive, so that Eq. (1) holds, where $i = 1, \ldots, n$;

$$X_{ij} = \mu + \beta_j + \epsilon_{ij} \tag{1}$$

and $j = 1, \ldots, t$. The statistical problem is to test for the existence of these effects.

The analysis of variance in this example is presented in the **table**. Entries in the sum of squares column represent that part of the total variation that is attributable to each source. Total sum of squares Q is the sum over all squared deviations of observations X_j from the grand mean $\overline{X}$, Eq. (2). Similarly,

$$\overline{X} = (\Sigma_i X_{ij})/nt \tag{2}$$

within treatments, sum of squares E is the sum over all squared deviations of observations X_{ij} within a treatment from the mean $\overline{X}_j$ of that treatment, Eq. (3).

$$\overline{X}_j = (\Sigma_i X_{ij})/n \tag{3}$$

Also, between treatments, sum of squares T is n times

Illustrative example of the analysis of variance

Source of variation	Sum of squares	Degrees of freedom	Mean square
Between treatments	$T = n\Sigma_j(\overline{X}_j - \overline{X})^2$	$t - 1$	$T = T/(t-1)$
Within treatments	$E = \Sigma_{i,j}(X_{ij} - \overline{X}_j)^2$	$t(n-1)$	$E' = E/t(n-1)$
Total	$Q = \Sigma_{i,j}(X_{ij} - \overline{X})^2$	$nt - 1$	

the sum over all treatments of the squared deviations of treatment means $\overline{X}_j$ from a grand mean $\overline{X}$ as defined by Eqs. (2) and (3). The sum of squares is generally computed more easily from the equivalent formulas (4)–(6).

$$Q = \Sigma_{i,j}X_{ij}^2 - \frac{(\Sigma_{i,j}X_{ij})^2}{nt} \tag{4}$$

$$T = \Sigma_j\frac{(\Sigma_i X_{ij})^2}{n} - \frac{(\Sigma_{i,j}X_{ij})^2}{nt} \tag{5}$$

$$E = Q - T \tag{6}$$

The entries under degrees of freedom represent the number of independent comparisions upon which the sum of squares for the source of variation is based. In every case the linear restriction imposed by the relationship of the particular mean to the observations results in the loss of one degree of freedom. Therefore the number of degrees of freedom is always one less than the number of deviations used to compute the sum of squares.

The mean squares in the analysis of variance are obtained by dividing the sum of squares by the corresponding degrees of freedom. The within-treatments mean square is an estimate of σ^2, the variance of the error term ϵ_{ij} in the additive model. It represents the random or unexplained variation in the data. The between-treatments mean square is an estimate of $\sigma^2 + \sigma_\beta^2$, where σ_β^2 is the variance of the treatment effects β_j.

If the treatment means differ substantially, the β_j effects estimated by $(\overline{X}_j - \overline{X})$ will differ correspondingly and will have a large variance σ_β^2. If on the other hand the means do not differ, the treatment effects β_j would be zero and σ_β^2 would be zero. In this case the treatment mean square would be equal to the error mean square and both would be independent estimates of σ^2. By comparing the ratio T'/E' of between treatment mean square T' to within-treatment mean square E' with unity, the variation due to treatments is compared with the variation due to random or unexplained factors. If this ratio, called the F ratio, is close to unity, there is no evidence of a treatment effect. However, if ratio T'/E' is substantially greater than unity there may be a significant treatment effect.

To compare the mean squares objectively, one uses the F test of significance in which the statistical hypothesis is that $\sigma_\beta^2 = 0$. Under this hypothesis it can be concluded that the treatment effects are significantly different from zero at the significance level α if the calculated F ratio is greater than the value of F at the α point on the F distribution with $t - 1$ and $t(n - 1)$ degrees of freedom. *See* BIOMETRICS; QUALITY CONTROL; STATISTICS. Robert L. Brickley

Bibliography. D. R. Anderson, D. F. Sweeny, and T. A. Williams, *Introduction to Statistics*: *Concepts and Applications*, 3d ed., 1993; S. Jarrell, *Basic Statistics*, 1994; N. A. Weiss, *Introductory Statistics*, 5th ed., 1999.

Analytic geometry

A branch of mathematics in which algebra is applied to the study of geometry. Because algebraic methods were first systematically applied to geometry in 1637 by the French philosopher-mathematician René Descartes, the subject is also called cartesian geometry. The basis for an algebraic treatment of geometry is provided by the existence of a one-to-one correspondence between the elements, "points" of a directed line g, and the elements, "numbers," that form the set of all real numbers. Such a correspondence establishes a coordinate system on g, and the number corresponding to a point of g is called its coordinate. The point O of g with coordinate zero is the origin of the coordinate system. A coordinate system on g is cartesian provided that for each point P of g, its coordinate is the directed distance $\overline{OP}$. Then all points of g on one side of O have positive coordinates (forming the positive half of g) and all points on the other side have negative coordinates. The point with coordinate 1 is called the unit point. Since the relation $\overline{OP} + \overline{PQ} = \overline{OQ}$ is clearly valid for each two points P, Q, of directed line g, then $\overline{PQ} = \overline{OQ} - \overline{OP} = q - p$, where p and q are the coordinates of P and Q, respectively. Those points of g between P and Q, together with P, Q, form a line segment. In analytic geometry it is convenient to direct segments, writing PQ or QP accordingly as the segment is directed from P to Q or from Q to P, respectively. To find the coordinate of the point P that divides the segment P_1P_2 in a given ratio r, put $\overline{P_1P}/\overline{P_2P} = r$. Then $(x - x_1)/(x - x_2) = r$, where x_1, x_2, x are the coordinates of P_1, P_2, P, respectively, and solving for x gives $x = (x_1 - rx_2)/(1 - r)$. Clearly r is negative for each point between P_1, P_2 and is positive for each point of g external to the segment. The midpoint of the segment divides it in the ratio -1, and hence its coordinate $x = (x_1 + x_2)/2$. *See* MATHEMATICS.

The plane. Choose any two intersecting lines g_1, g_2 of the plane, with cartesian coordinate systems selected on each so that the intersection point has coordinate O in each system. To each point P of the

plane an ordered pair of numbers (x, y) is attached as coordinates, where x is the coordinate of the point of intersection of g_1 with the line through P parallel to g_2, and y is the coordinate of the point of intersection of g_2 with the line through P parallel to g_1. If P lies on g, its coordinates are $(x, 0)$, where x is the coordinate of P in the coordinate system on g_1. Similarly, each point of g_2 has coordinates $(0, y)$. The origin O has coordinates (0, 0). The lines g_1 and g_2 are called the x axis and y axis, respectively. It is usually convenient to take the same scale on each axis; that is, the segments joining the unit points on the two axes to the origin are congruent. The notation $P(x, y)$ denotes a point P with coordinates (x, y). If ω denotes the angle made by the positive halves of the two axes, and d the distance of points $P_1(x_1, y_1)$, $P_2(x_2, y_2)$, application of the law of cosines yields Eq. (1).

$$d = \left[(x_1 - x_2)^2 + (y_1 - y_2)^2 + 2(x_1 - x_2)(y_1 - y_2)\cos\omega\right]^{1/2} \quad (1)$$

Since $\cos 90^\circ = 0$, this important formula is simplified by taking the axes mutually perpendicular. Though it is occasionally useful to employ oblique axes (that is, $\omega \neq 90^\circ$), the simplifications resulting from a rectangular cartesian coordinate system make it the usual choice. Such a cartesian coordinate system is assumed in what follows. Thus, the distance d of $P_1(x_1, y_1)$, $P_2(x_2, y_2)$ is given by Eq. (2).

$$d = [(x_1 - x_2)^2 + (y_1 - y_2)^2]^{1/2} \quad (2)$$

Let θ denote the smaller of the two angles that a line g makes with the positive half of the x axis, measured in the (positive) direction of rotation (that is, from the positive half of the x axis to the positive half of the y axis). Angle θ is called the slope angle of g, and the number $\lambda = \tan\theta$, the slope of g, plays an important role in plane analytic geometry. Slope is not defined for any line perpendicular to the x axis. If $P_1(x_1, y_1)$, $P_2(x_2, y_2)$ are two distinct points of line g with slope angle $\theta \neq 90^\circ$, and Q denotes the intersection of the line through P_1 perpendicular to the y axis with the line through P_2 perpendicular to the x axis, then Eq. (3) holds.

$$\lambda = \tan\theta = \frac{\overline{QP_2}}{\overline{P_1Q}} = \frac{y_2 - y_1}{x_2 - x_1} \quad (3)$$

Loci and equations. The correspondence between the geometric entity "point" and the arithmetic entity "pair of real numbers," upon which plane analytic geometry is based, results in associating with each geometric locus one or more equations that are satisfied by the coordinates of all those (and only those) points forming the locus (equations of the locus), and in associating with each system of equations in the variables x, y the figure (graph of the equations) whose points are determined by the pairs of numbers satisfying the equations. Thus the algebraic method of studying geometry is balanced by a geometric interpretation of algebra. A central problem in analytic geometry is that of finding equations of certain important figures among curves and surfaces. *See* CURVE FITTING.

Equations of lines. A line may be determined by data of various kinds, each yielding a different form for its equation. Thus an equation for a line g through $P_1(x_1, y_1)$ perpendicular to the x axis is evidently $x = x_1$. If g goes through $P_1(x_1, y_1)$, $P_2(x_2, y_2)$ and is not perpendicular to the x axis, let $P(x, y)$ be the coordinates of any other point of g. Then $(y - y_1)/(x - x_1) = \lambda = (y_2 - y_1)/(x_2 - x_1)$, whereas if (x, y) are the coordinates of a point not on g, $(y - y_1)/(x - x_1) \neq \lambda$. Hence an equation for g is shown as Eq. (4), from which follow Eqs. (5) and (6).

$$\frac{y - y_1}{x - x_1} = \frac{y_2 - y_1}{x_2 - x_1} \quad (4)$$

$$y - y_1 = \lambda(x - x_1) \quad (5)$$

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} \quad (6)$$

The two-point form is given by Eq. (4), the point-slope form by Eq. (5), and the symmetric form by Eq. (6). The validity of the determinant form,

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0$$

as well as the slope-intercept form, $y = \lambda x + b$, and the intercept for $x/a + y/b = 1$, $a \cdot b \neq 0$, also follow readily from Eq. (1). *See* DETERMINANT.

A line g is determined by its distance p from the origin O ($p \geqq 0$) and the angle β that the perpendicular to g from O makes with the x axis. (The perpendicular is directed from O to g in case g does not go through O, and so as to make $\beta < 180^\circ$ in the contrary case.) The equation of g in terms of p and β (the so-called normal form) is $x\cos\beta + y\sin\beta - p = 0$. The directed distance from g to any point $P(x_0, y_0)$ is $x_0\cos\beta + y_0\sin\beta - p$. It follows that the equations of the bisectors of the angles formed by two lines

$$x\cos\beta_i + y\sin\beta_i - p_i = 0 \quad (i = 1, 2)$$

are

$$x\cos\beta_1 + y\sin\beta_1 - p_1 = \pm(x\cos\beta_2 + y\sin\beta_2 - p_2)$$

The general form of an equation of a line is $Ax + By + C = 0$, where A, B are not both zero. The normal form is obtained from the general form upon dividing the left-hand member by $\pm\sqrt{A^2 + B^2}$, choosing the sign of the radical opposite to that of C in case $C \neq 0$, the same as that of B in case $C = 0$ and $B \neq 0$, and the same as that of A in case $B = 0$ and $C = 0$. If $B = 0$, the general form reduces to x = constant, a line perpendicular to the x axis, and if $B \neq 0$, one obtains $y = -(A/B)x + C/A$, a line with slope $-A/B$ and y intercept C/A. Thus to each line there corresponds a linear equation in x and y, and with each such equation is associated a line.

Angle between two lines. The angle ϕ ($0^\circ < \phi < 180^\circ$) from a line g_1 to another intersecting line g_2 is that through which g_1 must be rotated (about the point of intersection) to coincide with g_2. If θ_1, θ_2 are the slope angles of g_1, g_2, respectively, then

$$\phi = \theta_2 - \theta_1$$

and

$$\tan\phi = \frac{\tan\theta_2 - \tan\theta_1}{1 + \tan\theta_1 \tan\theta_2} = \frac{\lambda_2 - \lambda_1}{1 + \lambda_1 \cdot \lambda_2}$$

provided $\theta_1 \neq 90^\circ \neq \theta_2$. Consequently, g_1 and g_2 are mutually perpendicular if and only if $1 + \lambda_1 \cdot \lambda_2 = 0$. The formula for $\tan\phi$ holds in case g_1, g_2 are parallel. Then $\phi = 0^\circ$ and $\lambda_1 = \lambda_2$. It follows that two distinct lines $A_i x + B_i y + C_i = 0$, with $i = 1, 2$, are mutually perpendicular if and only if $A_1A_2 + B_1B_2 = 0$, and parallel provided $A_1B_2 - A_2B_1 = 0$.

Area of a triangle. Let $P_i(x_i, y_i)$, with $i = 1, 2, 3$, be vertices of a triangle whose area is denoted by A. Then $A = \frac{1}{2}d \cdot D$, where $d = [(x_1 - x_2)^2 + (y_1 - y_2)^2]^{1/2}$ and D is the distance of P_3 from the line joining P_1, P_2. Substituting the coordinates (x_3, y_3) of P_3 for (x, y) in the normal form of the equation of that line gives

$$D = \frac{\pm\left[(y_2 - y_1)x_3 - (x_2 - x_1)y_3 - x_1y_2 + x_2y_1\right]}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}}$$

and hence

$$A = \pm(^1/_2)(x_1y_2 + x_2y_3 + x_3y_1 - x_3y_2 - x_2y_1 - x_1y_3)$$

or

$$A = \pm(1/2) \cdot \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

The positive sign holds provided the vertices P_1, P_2, P_3 are in counterclockwise order, and the negative sign in the contrary case.

Linear combinations. If $u_1 = 0$, $u_2 = 0$ are equations of lines through a point P, for every choice of constants c_1, c_2 (except $c_1 = c_2 = 0$) the linear combination $c_1u_1 + c_2u_2 = 0$ is an equation of a line through P, and every line through P has an equation of that form. It follows that three lines $u_i = 0$, with $i = 1, 2, 3$, are concurrent provided there exist constants c_1, c_2, c_3 (not all zero) such that the linear combination $c_1u_1 + c_2u_2 + c_3u_3 = 0$ for every pair of numbers (x, y). Putting $u_i = A_ix + B_iy + C_i$, with $i = 1, 2, 3$, then c_1, c_2, c_3 are nontrivial solutions of the system of equations $c_1A_1 + c_2A_2 + c_3A_3 = 0$, $c_1B_1 + c_2B_2 + c_3B_3 = 0$, $c_1C_1 + c_2C_2 + c_3C_3 = 0$. Hence the lines $u_i = 0$, with $i = 1, 2, 3$, are concurrent if and only if

$$\begin{vmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix} = 0$$

Circle. By use of the formula for the distance of two points and the definition of a circle, an equation for a circle with center $C(x_0, y_0)$ and radius r ($r \geqq 0$) is found to be $(x - x_0)^2 + (y - y_0)^2 = r^2$. If $r = 0$, the only (real) point of the locus is the center (x_0, y_0) and the circle is a point circle. The above equation is called the standard form. Expansion yields $x^2 + y^2 - 2x_0x - 2y_0y - x_0^2 + y_0^2 - r^2 = 0$, a quadratic in x, y with equal, nonzero, coefficients of x^2 and y^2, and with the product term xy lacking. Conversely, the locus of each such equation $A(x^2 + y^2) + 2Dx + 2Ey + F = 0$, $A \neq 0$, is a (real) circle with center $(-D/A, -E/A)$ and radius $r = (1/A) \cdot [D^2 + E^2 - AF]^{1/2}$ provided $D^2 + E^2 - AF \geqq 0$, for by "completing the square" the above equation may be put in the standard form. Let $P(x_1, y_1)$ be on the circle $(x - x_0)^2 + (y - y_0)^2 = r^2$. An equation for the tangent to the circle at P is easily seen to be $(y_1 - y_0)(y - y_1) + (x_1 - x_0)(x - x_1) = 0$. Adding $(x_1 - x_0)^2 + (y_1 - y_0)^2$ to the left side of this equation, and its equal r^2 [since $P(x_1, y_1)$ lies on the circle] to the right side, yields $(x_1 - x_0)(x - x_0) + (y_1 - y_0)(y - y_0) = r^2$ as an equation of the tangent. The formal process by which this equation may be obtained from the standard form of the equation of a circle is called polarization. If $P(x_1, y_1)$ is not on the circle, the line that is the locus of that equation is called the polar line of P with respect to the circle. If P is outside the circle, its polar line joins the contact points of the two tangents from P to the circle. *See* CIRCLE.

Polarization of the general equation of a circle gives $A(x_1x + y_1y) + D(x + x_1) + E(y + y_1) + F = 0$, as an equation of the tangent at (x_1, y_1) if that point is on the circle, and of the polar line of (x_1, y_1) otherwise. The tangential distance t from (x_1, y_1) to the circle $(x - x_0)^2 + (y - y_0)^2 = r^2$ is given by $t^2 = (x_1 - x_0)^2 + (y_1 - y_0)^2 - r^2$. Hence $(x - x_0)^2 + (y - y_0)^2 - r_0^2 = (x - x_1)^2 + (y - y_1)^2 - r_1^2$ is an equation of the locus of points with equal tangential distances from the two circles with centers (x_0, y_0), (x_1, y_1) and radii r_0, r_1, respectively. Since the equation is linear, it represents a line which evidently contains the common chord of the circles in case they intersect in two distinct points. If $u_i = 0$, with $i = 1, 2, 3$, denotes equations of three circles that intersect pairwise in two distinct points, then the three common chords are concurrent, since the equation $u_1 - u_2 = 0$ is a linear combination of the equations $u_2 - u_3 = 0$, $u_3 - u_1 = 0$. An equation of the circle through three noncollinear points $P_i(x_i, y_i)$, with $i = 1, 2, 3$, may be written

$$\begin{vmatrix} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{vmatrix}$$

Conic sections. Much of plane analytic geometry deals with a class of curves which (from the way in which they were first studied) are known as conic sections or conics. A conic is the locus of a point P that moves so that its distance from a fixed point F (the focus) is in a constant positive ratio ϵ (the

eccentricity) to its distance from a fixed line (the directrix) not through F. Let $(c, 0)$ be the coordinates of F, $c > 0$, and take the y axis as the directrix. Then $P(x, y)$ satisfies the equation $[(x - c)^2 + y^2]^{1/2} = \epsilon x$; that is, $(1 - \epsilon^2)x^2 + y^2 - 2cx + c^2 = 0$, and it is easily seen that each point whose coordinates satisfy this equation is on the conic. Hence each conic is represented by a second-degree equation in the cartesian coordinates (x, y). A conic is called a parabola, ellipse, or hyperbola accordingly as $\epsilon = 1$, $\epsilon < 1$, $\epsilon > 1$, respectively. *See* CONIC SECTION.

Parabola. When $\epsilon = 1$, the equation for the conic obtained above becomes $y^2 - 2cx + c^2 = 0$ or $y^2 = 2c(x - c/2)$. The curve has an axis of symmetry (the x axis when the focus and directrix are chosen as above). The point $(c/2, 0)$ at which the axis intersects the curve is the vertex. Putting $x = c$ in the equation, it is seen that the chord through the focus that is perpendicular to the axis (the latus rectum) has length $2c$. All quadratics, $x = Ay^2 + By + C$, $y = Ax^2 + Bx + C$, $A \neq 0$, are equations of parabolas (with axes perpendicular to the y axis, or the x axis, respectively). All other parabolas have equations $Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$, with discriminant

$$\begin{vmatrix} A & B & D \\ B & C & E \\ D & E & F \end{vmatrix} \neq 0 \quad \text{and} \quad AC - B^2 = 0$$

(A quadratic with nonvanishing discriminant is called irreducible.) A simple standard form of an equation of a parabola is obtained by taking $(c/2, 0)$ for focus $(c > 0)$ and the line $x = -c/2$ for directrix, resulting in $y^2 = 2cx$. Polarizing gives $y_1y = c(x + x_1)$ for an equation of the tangent at the point $P(x_1, y_1)$ on the parabola. This tangent cuts the axis at the point $Q(-x_1, 0)$ whose distance from the focus is $x_1 + c/2$. But this is the distance from the directrix to P, and consequently the triangle FPQ is isosceles with $\measuredangle FQP = \measuredangle FPQ$. It follows that the line joining F to any point P of the parabola makes the same angle with the tangent at P as does the line through P that is parallel to the axis of the parabola. Thus each light ray emanating from the focus is reflected parallel to the axis. *See* PARABOLA.

Ellipse. When $\epsilon < 1$, the equation of the general conic given above becomes

$$\left(\frac{x - c}{1 - \epsilon^2}\right)^2 + \frac{y^2}{1 - \epsilon^2} = \frac{\epsilon^2 c^2}{(1 - \epsilon^2)^2} \quad (\epsilon < 1)$$

This is an equation of the ellipse with focus $(c, 0)$ and directrix the y axis. Putting $X = x - c/(1 - \epsilon^2)$, $Y = y$, the standard form $X^2/A^2 + Y^2/B^2 = 1$ of the equation is obtained, where $A = \epsilon c/(1 - \epsilon^2) > 0$, $B = \epsilon c/(1 - \epsilon^2)^{1/2} > 0$. This is equivalent to referring the ellipse to a new set of coordinate axes obtained by translating the origin to the point $[c/(1 - \epsilon^2), 0]$. The coordinates of the focus F in the XY coordinate system are $[-c\epsilon^2/(1 - \epsilon^2), 0]$ and the equation of the directrix is $X = -c/(1 - \epsilon^2)$. Since the standard form shows the curve to be symmetric with respect to each of the new axes and the origin, it is clear that $F'[c\epsilon^2/(1 - \epsilon^2), 0]$ and $X = c/(1 - \epsilon^2)$ may be taken as focus and directrix. Putting $C = c\epsilon^2/(1 - \epsilon^2)$, then $C^2 = A^2 - B^2$, and $X^2/A^2 + Y^2/B^2 = 1$ is seen to be the locus of points $P(X, Y)$, the sum of whose distances from $(C, 0)$ and $(-C, 0)$ is $2A$. This relation is frequently used to define an ellipse. The chord containing the foci is the major axis; its length is $2A$. The chord of length $2B$ through the midpoint of the foci, perpendicular to the major axis, is the minor axis. The tangent to $X^2/A^2 + Y^2/B^2 = 1$ at $P(X_1, Y_1)$ on the ellipse is $X_1X/A^2 + Y_1Y/B^2 = 1$. Lines PF, PF' make equal angles with the tangent, so sound or light that emanates from one focus is reflected to the other. An irreducible quadratic $Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$ is an equation of an ellipse provided that $AC - B^2 > 0$. *See* ELLIPSE.

Hyperbola. When $\epsilon > 1$, the equation for the conic is written $[x + c/(\epsilon^2 - 1)]^2 - y^2/(\epsilon^2 - 1) = \epsilon^2c^2/(1 - \epsilon^2)^2$, $(\epsilon > 1)$. Putting $X = x + c/(\epsilon^2 - 1)$, $Y = y$ gives the standard form of a hyperbola $X^2/A^2 - Y^2/B^2 = 1$, where $A = \epsilon c/(\epsilon^2 - 1) > 0$, $B = \epsilon c/(\epsilon^2 - 1)^{1/2} > 0$. It is clear that the curve consists of two branches and is symmetric to the axes and the origin. Putting $C = c\epsilon^2(\epsilon^2 - 1)$, then both $F(C, 0)$, $F'(-C, 0)$ are foci of the hyperbola, with respective directrices $X = c/(\epsilon^2 - 1)$, $X = -c/(\epsilon^2 - 1)$, and $A^2 + B^2 = C^2$. Then $X^2/A^2 - Y^2/B^2 = 1$ is seen to be an equation of the locus of points $P(X, Y)$ such that $PF - PF' = 2A$. The lines $X/A \pm Y/B = 0$ are asymptotes of the hyperbola. As a point P traverses either branch of the hyperbola, its distances from an asymptote approach zero. An irreducible quadratic $Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$ is an equation of an hyperbola provided $AC - B^2 < 0$. *See* HYPERBOLA.

Three-dimensional space. Let cartesian coordinate systems be established on each of three pairwise mutually perpendicular lines of three-space that intersect in O, the common origin of the systems. Suppose equal scales and call the lines the x axis, y axis, and z axis. To each point P of space an ordered triple (x, y, z) of real numbers is attached as rectangular cartesian coordinates, where x is the coordinate of the foot of the perpendicular from P to the x axis, and y and z are similarly defined. Thus every point of space has unique coordinates, and each ordered triple of real numbers is the coordinates of a point of space. If d denotes the distance of two points $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$, then $d = [(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2]^{1/2}$. Let g denote any directed line and g' the line through O parallel to g and directed in the same sense. If α, β, γ denote the angles that g' makes with the x, y, and z axes, respectively (the direction angles of g), then $\cos\alpha$, $\cos\beta$, $\cos\gamma$ are the direction cosines of g. They satisfy the relation $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$, and any three numbers λ, μ, ν such that $\lambda^2 + \mu^2 + \nu^2 = 1$ are the direction cosines of a (directed) line. Three numbers a, b, c proportional to the direction cosines λ, μ, ν of a directed line g are direction numbers of g. Clearly direction numbers of parallel lines are proportional. If $P_1(x_1, y_1, x_1)$, $P_2(x_2, y_2, z_2)$ are distinct points of g, then $x_2 - x_1$, $y_2 - y_1$, $z_2 - z_1$ are direction numbers of g. It follows that $P(x, y, z)$ is on g if and only if

$x - x_1 = t(x_2 - x_1)$, $y - y_1 = t(y_2 - y_1)$, $z - z_1 = t(z_2 - z_1)$, $-\infty > t > \infty$. These are parametric equations of g (t is the parameter), from which the symmetric equations $(x - x_1)/(x_2 - x_1) = (y - y_1)/(y_2 - y_1) = (z - z_1)/(z_2 - z_1)$ follow at once. The direction cosines of g, directed from P_1 to P_2, are $\cos\alpha = (x_2 - x_1)/d$, $\cos\beta = (y_2 - y_1)/d$, $\cos\gamma = (z_2 - z_1)/d$, where $d = [(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2]^{1/2}$.

If α_1, β_1, γ_1, and α_2, β_2, γ_2 are direction angles of directed lines g_1, g_2, respectively, and θ denotes the angle between them, then $\cos\theta = \cos\alpha_1\cos\alpha_2 + \cos\beta_1\cos\beta_2 + \cos\gamma_1\cos\gamma_2$. Hence g_1, g_2 are mutually perpendicular if and only if $a_1a_2 + b_1b_2 + c_1c_2 = 0$, where a_1, b_1, c_1 and a_2, b_2, c_2 are direction numbers of g_1, g_2, respectively. Let the plane π go through $P_0(x_0, y_0, z_0)$ and be perpendicular to a line g with direction numbers a, b, c. Then $P(x, y, z)$ is in π if and only if $P = P_0$ or the line joining it to P_0 is at right angles to a line through P_0 that is parallel to g; that is, if and only if $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$. Hence to each plane corresponds a linear equation. Conversely, if P_0 (x_0, y_0, z_0) satisfies the linear equation $Ax + By + Cz + D = 0$, with A, B, C not all zero, then $A(x - x_0) + B(y - y_0) + C(z - z_0) = 0$ is an equation of a plane through P_0, perpendicular to a line with direction numbers A, B, C, and so $Ax + By + Cz + D = 0$, with $A, B, C \neq 0, 0, 0$, is an equation of a plane. Clearly $x = 0$ is an equation for the plane determined by the y and z axes; equations of the other two coordinate planes are $y = 0$ and $z = 0$. If a directed perpendicular g from O to a plane meets the plane at P (g is directed from O to P in case the plane is not through O) the plane has equation $x\cos\alpha + y\cos\beta + z\cos\gamma = p$, where α, β, γ are the direction angles of g and $p = OP$. This is the normal form of equation of a plane. The general form is reduced to it upon dividing by $\pm[A^2 + B^2 + C^2]^{1/2}$. The distance from $Ax + By + Cz + D + 0$ to $P(x_0, y_0)$ is $(Ax_0 + By_0 + Cz_0 + D)/\pm(A^2 + B^2 + C^2)^{1/2}$, where the sign is selected opposite that of D. (In case $D = 0$, other conventions are used.) Two planes $A_ix + B_iy + C_iz = 0$, with $i = 1, 2$, are parallel in case the number triples A_1, B_1, C_1 and A_2, B_2, C_2 are proportional and are mutually perpendicular provided $A_1A_2 + B_1B_2 + C_1C_2 = 0$ (since A_1, B_1, C_1 and A_2, B_2, C_2 are direction numbers of lines that are perpendicular to the respective planes). If $P_i(x_i, y_i, z_i)$, with $i = 1, 2, 3$, is not collinear, the plane determined has the equation

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0$$

Three planes $A_ix + B_iy + C_iz + D_i = 0$, with $i = 1, 2, 3$, intersect in one point if and only if

$$\begin{vmatrix} A_1 & B_2 & C_1 \\ A_2 & B_2 & C_2 \\ A_3 & B_3 & C_3 \end{vmatrix} = 0$$

The locus of points whose coordinates satisfy an equation $f(x, y, z) = 0$ is a surface. Curves may be thought of as intersections of two surfaces, and as such the coordinates of their points satisfy two equations $f(x, y, z) = 0$, $g(x, y, z) = 0$. Thus a line, considered as the intersection of two planes, is given by the two (simultaneous) equations $A_ix + B_iy + C_iz = D_i = 0$, with $i = 1, 2$. That line has direction numbers

$$\begin{vmatrix} B_1 & C_1 \\ B_2 & C_2 \end{vmatrix} \qquad \begin{vmatrix} C_1 & A_1 \\ C_2 & A_2 \end{vmatrix} \qquad \begin{vmatrix} A_1 & B_2 \\ A_2 & B_2 \end{vmatrix}$$

Curves are also represented parametrically by equations $x = f(t)$, $y = g(t)$, $z = h(t)$, the parameter t varying in an interval (a, b), finite or infinite. Parametric equations for the line have already been given. Additional examples are $x = r\cos t$, $y = r\sin t$, $z = 0$, with $0 \le t \le 2\pi$, parametric equations of the circle in the xy plane, with center at O and radius r; $x = a \cdot \cos t$, $y = a \cdot \sin t$, $z = kt$, with $k \neq 0$, $-\infty < t < \infty$, parametric equations of a circular helix—the curve of the thread of a machine (untapered) screw.

Special surfaces. It follows from the definition of a sphere and the formula for distance of two points that $(x - a)^2 + (y - b)^2 + (z - c)^2 = r^2$ is an equation for the sphere with center (a, b, c) and radius r, and by completing the squares of the x, y, and z terms in the equation $x^2 + y^2 + z^2 + 2Dx + 2Ey + 2Fz + G = 0$, it is seen that the locus of such an equation is a sphere with positive or zero radius, or there is no (real) locus.

Any equation in just two of the three coordinates is an equation of a cylinder whose elements are parallel to the axis of the missing variable. Thus the locus in 3-space of $x^2 + y^2 = r^2$ is a (right) circular cylinder whose elements are parallel to the z axis and which intersects the xy plane in the circle $x^2 + y^2 = r^2$, $z = 0$.

Any equation $f(x, y, z) = 0$, with $f(x, y, z)$ homogeneous in x, y, z (for example, $4xy - xz + yz = 0$, $x^3 - xy^2 + z^3 = 0$) has a cone with vertex O as locus.

A surface of revolution is obtained by rotating a plane curve C about a line g of its plane. If $f(x, y) = 0$, $z = 0$ are equations of C, and g is the x axis, the resulting surface of revolution has the equation $f(x, \sqrt{y^2 + z^2}) = 0$. Thus the surface generated by revolving the circle $x_2 + (y - b)^2 = a^2$, $z = 0$, about the x axis (the torus or anchor ring, if $b > a$) has the equation $x^2 + (\sqrt{y^2 + z^2} - b)^2 = a^2$.

A quadric surface is the locus of points whose coordinates satisfy an equation of the form $Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Jz + K = 0$, where at least one coefficient of a second-degree term is not zero. Some surfaces obtained by rotating conics about a line belong to this class, for example, spheres, prolate and oblate spheroids (given by rotating an ellipse about its major and minor axes, respectively), hyperboloids and paraboloids resulting from rotations of hyperbolas and parabolas about their axes of symmetry, and right circular cones and cylinders. Cylinders with conics for directrix curves are also members. Apart from degenerate cases, such as two planes, the remaining quadric surfaces and the standard forms of their equations are (1) ellipsoid,

$x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$; (2) hyperboloid of one sheet, $x^2/a^2 + y^2/b^2 - z^2/c^2 = 1$; (3) hyperboloid of two sheets, $x^2a^2 - y^2/b^2 - z^2/c^2 = 1$; (4) elliptic paraboloid, $x^2/a^2 + y^2/b^2 = 2z$; and (5) hyperbolic paraboloid, $x^2/a^2 - y^2/b^2 = 2z$. Hyperboloids of one sheet and hyperbolic paraboloids are ruled surfaces; that is, each contains an infinity of straight lines, called generators. In fact, each of those surfaces contains two sets of generators. *See* QUADRIC SURFACE.

n-Dimensions. Let cartesian coordinate systems with equal scales be established on each of n pairwise mutually perpendicular lines intersecting in the common origin O, and label the lines $OX_1, OX_2, \ldots, OX_n$. To each point P of n space an ordered n-tuple $(x_1, x_2, \ldots, x_n)$ of numbers is attached as coordinates, where x_i is the coordinate of the foot of the perpendicular from P to OX_i, with $i = 1, 2, \ldots, n$. Two points $P(x_1, x_2, \ldots, x_n)$, $Q(y_1, y_2, \ldots, y_n)$ have distance $d = [(x_1 - y_2)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2]^{1/2}$. Direction angles $x_1, x_2, \ldots, x_n$ of a directed line are defined as in three-dimensional analytic geometry, and the direction cosines satisfy the relation $\cos^2 \alpha_1 + \cos^2 \alpha_2 + \cdots + \cos^2 \alpha_n = 1$. Direction cosines of the line $g(P, Q)$, directed from P to Q, are $(y_1 - x_1)/d, (y_2 - x_2)/d, \ldots, (y_n - x_n)/d$, and numbers proportional to them (for example, the numerators) are direction numbers of $g(P, Q)$. Hence $(X_1, X_2, \ldots, X_n)$ are coordinates of a point on $g(P, Q)$ if and only if $(X_1 - x_1)/(y_1 - x_1) = (X_2 - x_2)/(y_2 - x_2) = \cdots = (X_n - x_n)/(y_n - x_n)$. These are symmetric equations for the line determined by P, Q. Denoting the common value of the quotients by t gives the n parametric equations $X_i = x_i + t(y_i - x_i)$, with $i = 1, 2, \ldots, n$, $-\infty < t < \infty$. Let g be a line through $C(c_1, c_2, \ldots, c_n)$ with direction numbers $a_1, a_2, \ldots, a_n$. Point C, together with all points X such that the line $g(C, X)$ is perpendicular to g, defines an $(n - 1)$-dimensional subspace (hyperplane). Its equation is $a_1(X_1 - c_1) + a_2(X_2 - c_2) + \cdots + a_n(X_n - c_n) = 0$, since $X_i - c_i$, with $i = 1, 2, \ldots, n$, are direction numbers of $g(C, X)$ and the equation is the condition that g and $g(C, X)$ be mutually perpendicular. It is readily seen that the locus of every linear equation $A_1X_1 + A_2X_2 + \cdots + A_nX_n + K = 0$ is a hyperplane perpendicular to a line with direction numbers $A_1, A_2, \ldots, A_n$. It may be put in normal form by dividing by $\pm(A_1^2 + A_2^2 + \cdots + A_n^2)^{1/2}$, and the distance from it to a point $C(c_1, c_2, \ldots, c_n)$ if found by substituting the coordinates of C for $X_1, X_2, \ldots, X_n$. Subspaces of dimension k $(1 \le k < n)$ are given by systems of $n - k$ linear equations. An equation of the hyperplane determined by n points is readily expressed in determinant form, as well as the condition that $n + 1$ points be on a hyperplane [substitute the coordinates of the $(n + 1)$-st point in the first row of the determinant equation of the hyperplane]. Discussion of loci of higher order lies outside of the scope of this article.

This brief sketch of analytic geometry has dealt only with that coordinate system most frequently used. For other coordinate systems (polar), as well as a discussion of transformation of coordinates, *See* ALGEBRA; COORDINATE SYSTEMS; DIFFERENTIAL GEOMETRY; EUCLIDEAN GEOMETRY; PROJECTIVE GEOMETRY; TRIGONOMETRY.

Leonard M. Blumenthal

Bibliography. G. Fuller, *Analytic Geometry*, 7th ed., 1993; D. F. Riddle, *Analytic Geometry*, 6th ed., 1996.

Analytic hierarchy

A framework for solving a problem. The analytic hierarchy process is a systematic procedure for representing the elements of any problem. It organizes the basic rationality by breaking down a problem into its smaller constituents and then calls for only simple pairwise comparison judgments, to develop priorities in each level.

The analytic hierarchy process provides a comprehensive framework to cope with intuitive, rational, and irrational factors in making judgments at the same time. It is a method of integrating perceptions and purposes into an overall synthesis. The analytic hierarchy process does not require that judgments be consistent or even transitive. The degree of consistency (or inconsistency) of the judgment is revealed at the end of the analytic hierarchy process.

Human reasoning. People generally provide subjective judgments that are based on feelings and intuition rather than on well-worked-out logical reasoning. Also, when they reason together, people tend to influence each other's thinking. Individual judgments are altered slightly to accommodate the group's logic and the group's interests. However, people have very short memories, and if asked afterward to support the group judgments, they instinctively go back to their individual judgments. Repetition is needed to effect deep-rooted changes in people's judgment.

People also find it difficult to justify their judgments logically and to explicate how strong these judgments are. As a result, people make great compromises in their thinking to accommodate ideas and judgments. In groups, there is a willingness to compromise. If truth is to be an objective reality, reality must be very fuzzy because the search for truth often ends in compromise. What is regarded as truth may often be essentially a social product obtained through interaction rather than by pure deduction.

Logical understanding does not seem to permeate judgment instantaneously. It apparently needs time to be assimilated. However, even when logical understanding has been assimilated, people still offer judgment in a spontaneous emotional way without elaborate explanation. Even when there is time, explanations tend to be fragmentary, disconnected, and mostly without an underlying clear logical foundation.

Outline of the process. People making comparisons use their feelings and judgment. Both vary in intensity. To distinguish among different intensities, the scale of absolute numbers in **Table 1** is useful.

The analytic hierarchy process can be decomposed into the following steps. Particular steps may be emphasized more in some situations than in

TABLE 1. Scale of relative importance

Intensity of relative importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
3	Slight importance of one over another	Experience and judgment slightly favor one activity over another
5	Essential or strong importance	Experience and judgment strongly favor one activity over another
7	Demonstrated importance	An activity is strongly favored and its dominance is demonstrated in practice
9	Absolute importance	The evidence favoring one activity over another is of the highest possible order of affirmation
2, 4, 6, 8	Intermediate values between the two adjacent judgments	When compromise is needed
Reciprocals of above nonzero numbers	If an activity has one of the above numbers assigned to it when compared with second activity, the second activity has the reciprocal value when compared to the first	

others. Also as noted, interaction is generally useful for stimulation and for representing different points of view.

1. Define the problem and determine what knowledge is sought.

2. Structure the hierarchy from the top (the objectives from a broad perspective) through the intermediate levels (criteria on which subsequent levels depend) to the lowest level (which usually is a list of the alternatives).

3. Construct a set of pairwise comparison matrices for each of the lower levels, one matrix for each element in the level immediately above. An element in the higher level is said to be a governing element for those in the lower level since it contributes to it or affects it. In a complete simple hierarchy, every element in the lower level affects every element in the upper level. The elements in the lower level are then compared to each other, based on their effect on the governing element above. This yields a square matrix of judgments. The pairwise comparisons are done in terms of which element dominates the other. These judgments are then expressed as integers according to the judgment values in Table 1. If element A dominates element B, then the whole number integer is entered in row A, column B, and the reciprocal (fraction) is entered in row B, column A. Of course, if element B dominates element A, the reverse occurs. The whole number is then placed in the B,A position with the reciprocal automatically being assigned to the A,B position. If the elements being compared are equal, a one is assigned to both positions. The numbers used express an absolute rather than an ordinal relation.

4. There are $n(n - 1)/2$ judgments required to develop the set of matrices in step 3 (taking into account the fact that reciprocals are automatically assigned in each pairwise comparison), where n is the number of elements in the lower level.

5. Having collected all the pairwise comparison data and entered the reciprocals together with n unit entries down the main diagonal (an element is equal to itself, so a "one" is assigned to the diagonal positions), the eigenvalue problem $Aw = \lambda_{\max} w$ is solved and consistency is tested, using the departure of $\lambda_{\max}$ from n (see below).

6. Steps 3, 4, and 5 are performed for all levels and clusters in the hierarchy.

7. Hierarchial composition is now used to weigh the eigenvectors by the weights of the criteria, and the sum is taken over all weighted eigenvector entries corresponding to those in the lower level of the hierarchy.

8. The consistency ratio of the entire hierarchy is found by multiplying each consistency index by the priority of the corresponding criterion and adding them together. The result is then divided by the same type of expression, using the random consistency index corresponding to the dimensions of each matrix weighted by the priorities as before. The consistency ratio should be about 10% or less to be acceptable. If not, the quality of the judgments should be improved, perhaps by revising the manner in which questions are asked in making the pairwise comparisons. If this should fail to improve consistency, it is likely that the problem should be more accurately structured; that is, similar elements should be grouped under more meaningful criteria. A return to step 2 would be required, although only the problematic parts of the hierarchy may need revision.

If the exact answer in the form of hard numbers was actually available, it would be possible to normalize these numbers, form their ratios as described above, and solve the problem. This would result in getting the same numbers back, as should be expected. On the other hand, if firm numbers were not available, their ratios could be estimated to solve the problem.

Example of the process. In the following example the analytic hierarchy process is used to assist a young family (a father, a mother, and a child) of specified income to buy a new car, say, either model A, B, or C. The choice will be determined through four important criteria.

TABLE 2. Comparison matrix comparing criteria for buying a car

Decision to buy a car	Price	Running cost	Comfort	Status	Priorities
Price	1	3	7	8	.582
Running cost	1/3	1	5	5	.279
Comfort	1/7	1/5	1	3	.090
Status	1/8	1/5	1/3	1	.050

$\lambda_{max} = 4.198$
C.I. = .066
C.R. = .073

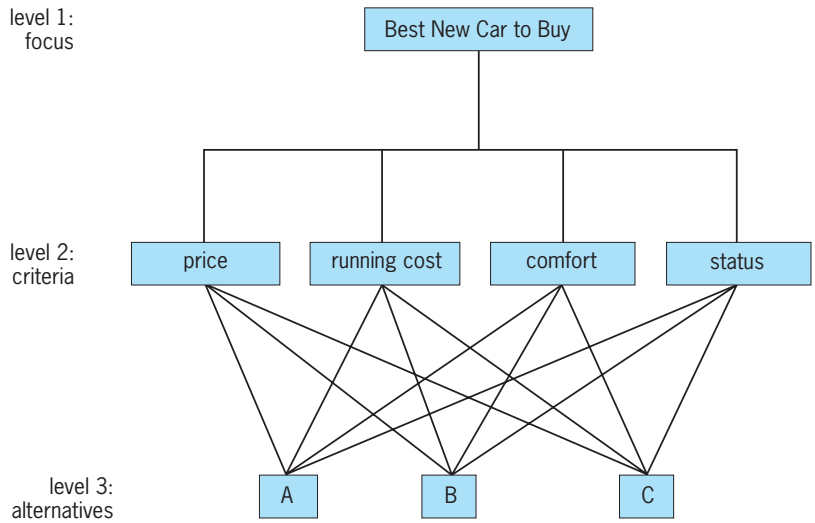

Analytic hierarchy used to assist a family in buying a new car.

The hierarchy of such a decision often takes the form shown in the **illustration**. In this hierarchy, level 1 is the single overall objective: Best New Car to Buy. On level 2 are the criteria which are perceived to compose what is meant by Best New Car, such as Price and Running Cost (operating and maintenance). On level 3 are the various alternative cars from which the family will choose.

This downward decomposition format can easily be used on a wide class of problems. In addition, a slight further modification to incorporate feedback loops will cover an even wider range.

The questions asked when comparing each criterion are of the following kind: Of the two alternatives being compared, which is considered more important by the family buying a car and how much more important is it? The comparison matrix of **Table 2** is then formed. Since Price and Running Cost have the highest priorities, the other factors are discarded and only these two are used in continuing the process. Care must be taken in doing this. The original priorities are then normalized by dividing each of their sum to obtain the new relative priorities in Eqs. (1).

$$\begin{aligned} .67 &= \frac{.58}{.58 + .28} \\ .33 &= \frac{.28}{.58 + .28} \end{aligned} \tag{1}$$

The process is then repeated for the third level, where each car is compared with respect to the two high-priority factors from the second level, as shown in **Table 3**.

The two priority columns are then recorded as in **Table 4**. All entries of the first column are then multiplied by .67, the priority of Price, and those of the second column by .33, the priority of Running Cost, and added. This gives the third column. Thus Car B was selected for its efficient operation even though its initial price is considerably higher than Car A. A car dealer in the income neighborhood of this family may find it profitable to stock up on the cars in the respective proportions.

One of the most powerful contributions that the analytic hierarchy process makes is to test out the degree of inconsistency or incompatibility of new ideas or new policies adopted with older, more familiar, better tried successful methods. For example, in doing the above problem the participants were not sure whether the judgments for Price over Running Cost should be 7, 5, or 3. Each one was tried separately, and it was found that 3 yielded the highest consistency. Those who voted for 3 won that argument.

Priorities and consistency. An easy way to get a good approximation of the priorities is to use the geometric mean. This is done by multiplying the elements in each row and taking their nth root, where n is the number of elements. Then, normalize the column of numbers thus obtained by dividing each entry by the sum of all entries. Alternatively, normalize the elements in each column of the matrix and then average each row.

The consistency index can also be determined by hand calculations. Add the numbers in each column of the judgment matrix, multiply the first sum by the first priority, the second by the second, and so on, and add. For the first matrix the column sums (1.60, 4.40, 13.33, 17) are obtained, and multiplying by (.582, .279, .090, .050) gives 4.20. This number is

TABLE 3. Comparison matrices comparing alternative cars

Price	Car A	Car B	Car C	Priorities	Running cost	Car A	Car B	Car C	Priorities
Car A	1	2	3	.540	Car A	1	1/5	1/2	.106
Car B	1/2	1	2	.297	Car B	5	1	7	.745
Car C	1/3	1/2	1	.163	Car C	2	1/7	1	.150
				$\lambda_{max} = 3.009$ C.I. = .005 C.R. = .008					$\lambda_{max} = 3.119$ C.I. = .059 C.R. = .103

TABLE 4. Composition of priorities

	Price (priority .67)	Running cost (priority .33)	Composite priority of cars
Car A	.540	.106	.396
Car B	.297	.745	.445
Car C	.163	.150	.159

denoted by λ_{max}. The consistency index is given by Eq. (2).

$$\text{C.I.} = \frac{\lambda_{max} - n}{n - 1} \tag{2}$$

The consistency is now checked by taking the ratio (C.R.) of C.I. with the appropriate one of the set of numbers in **Table 5** to see if it is about 10% or less (20% may be tolerated in some cases but not more). Otherwise the problem must be studied again and judgments revised. The consistency of the hierarchy in the above example, as given by Eq. (3), is .06, which is good.

$$\frac{.066 \times 1 + .005 \times .67 + .059 \times .33}{.900 \times 1 + .580 \times .67 + .580 \times .33} = \frac{.089}{1.48} = .06 \tag{3}$$

Judgment formation. When several people participate, judgments are often debated. Sometimes the group accepts a geometric average of their combined judgments. If there is strong disagreement, the different opinions can each be taken and used to obtain answers. Those which subsequently display the highest consistency within the group are the ones usually retained.

The analytic hierarchy process incorporates equally both tangible factors, which require hard measurements, and such intangible factors as comfort, which require judgment. Eventually one finds that so-called hard numbers have no meaning in themselves apart from their utilitarian interpretations. In the above example, buying a $10,000 car is more than twice as "painful" as buying a $5000 car.

The interdependence of criteria, such as Comfort and Price, has to be considered carefully since there may be some perceived overlap. For example, higher price buys more comfort, but it also buys other desirable attributes. Judging the relative importance of such things as price and comfort, therefore, must be done as independently as possible with avoidance of overlaps.

Validation by physical laws. Using the scale 1–9 has been justified and demonstrated by many examples. However, the following simple optics illustration, carried out with small children, shows that perceptions, judgments, and these numbers lead to results which can be validated by laws of physics. In this example, four identical chairs were placed at distances of 9, 15, 21, and 28 yards (1 yd = 0.9144 m) from a floodlight. The children stood by the light, looked at the line of chairs, and compared the first with the second, the first with the third and then with the fourth, and so on for the second, third, and fourth chairs. Each time the children said how much brighter one chair was, compared to the other.

Their judgments were entered in the matrix of **Table 6** to record the relative brightness of the chairs. The reciprocals were used in the transpose position.

TABLE 6. Comparison matrix comparing perceived brightnesses of chairs

	Chair 1	Chair 2	Chair 3	Chair 4	Brightness ratios
Chair 1	1	5	6	7	0.61
Chair 2	1/5	1	4	6	0.24
Chair 3	1/6	1/4	1	4	0.10
Chair 4	1/7	1/6	1/4	1	0.05

λ_{max} = 4.39
C.I. = 0.13
C.R. = 0.14

The inverse-square law of optics is now used to test these judgments. Since the distances are 9, 15, 21, and 28 yd, these numbers are squared and their reciprocals calculated. This gives .0123, .0044, .0023, and .0013 respectively. Normalization of these values gives .61, .22, .11, and .06, which are very close to the brightness ratios obtained in the test using the analytic hierarchy process.

Structuring a hierarchy. There are no rules for structuring a hierarchy. However, a typical analytic hierarchy for allocating resources—either by measuring costs or by measuring benefits—will often be stratified roughly as follows. The top level will include the overall objectives of the organization or system. Benefit-cost criteria may appear in the next level. A subordinate level may further clarify these criteria in the context of the particular problem by itemizing specific tasks which are to be accomplished at some level of performance. This is followed by the alternatives being evaluated.

Capabilities. Designing an analytic hierarchy—like the structuring of a problem by any other method—is more art than science. It necessitates substantial knowledge of the system in question. A very strong aspect of the analytic hierarchy process is that the knowledgable individuals who supply judgments for the pairwise comparisons usually also play a prominent role in specifying the hierarchy.

Although a hierarchy to be used in resource allocation will tend to have the vertical stratification

TABLE 5. Random consistencies

n:	1	2	3	4	5	6	7	8	9	10
Random consistency:	0	0	.58	.90	1.12	1.24	1.32	1.41	1.45	1.49

indicated above, it can also be much more general. The only restriction is that an element on a higher level must serve as a governing element for at least one element (which can be the element itself) on the immediately lower level. The hierarchy need not be complete; that is, an element at an upper level need not function as a criterion for all the elements in the lower level. It can be partitioned into nearly disjoint subhierarchies sharing only a common topmost element. Thus, for instance, the activities of separate divisions of an organization can be structured separately. As suggested above, the analyst can insert and delete levels and elements as necessary to clarify the task or to sharpen a focus on one or more areas of the system.

The analytic hierarchy process has already been successfully applied in a variety of fields, including planning the allocation of energy to industries; designing a transport system for the Sudan; planning the future of a corporation and measuring the impact of environmental factors on its development; designing future scenarios for higher education in the United States; selecting candidates and winners in elections; setting priorities for the top scientific institute in a developing country; solving a faculty promotion and tenure problem; and predicting oil prices. *See* DECISION THEORY; SYSTEMS ENGINEERING.

Thomas L. Saaty

Bibliography. B. I. Golden (ed.), *The Analytic Hierarchy Process*, 1989; T. L. Saaty, *The Analytic Hierarchy Process*, 1980; T. L. Saaty, *Decision Making for Leaders*, 1982, reprint 1990; T. L. Saaty and L. Vargas, *The Logic of Priorities*, 1981, reprint 1991.

Analytical chemistry

The science of chemical characterization and measurement. Qualitative analysis is concerned with the description of chemical composition in terms of elements, compounds, or structural units, whereas quantitative analysis is concerned with the measurement of amount.

Scope. Originally, chemical analysis conformed closely to its literal meaning and consisted of the separation of a sample into its components, which were weighed. Later, such gravimetric techniques were supplemented by volumetric or, more appropriately, titrimetric analysis, which consists of measurement of the amount of a standardized solution of a reagent that reacts with a measured portion of the sample to reach an end point—a process known as titration. The end point is usually detected by means of an indicator, that is, a substance that changes color when a given amount of reagent has been added. These forms of chemical analysis are called classical or wet methods to distinguish them from the instrumental methods, which have greatly enlarged the scope of analytical chemistry. *See* GRAVIMETRIC ANALYSIS; TITRATION.

Instrumental methods, which might better be termed physicochemical methods, have been developed to the degree that they make up the vast bulk of both qualitative and quantitative analysis. They include measurements of purely physical characteristics, as well as the use of physical measurements, to follow the course of chemical reactions.

Over the years, the scope of analytical chemistry has been enlarged from its original meaning, which was limited to the determination of chemical composition in terms of the relative amounts of elements or compounds in a sample. The discipline has been expanded to involve, for example, the spatial distribution of elements or compounds in a sample, the distinction between different crystalline forms of a given element or compound, the distinction between different chemical forms (such as the oxidation state of an element), the distinction between a component on the surface or in the interior of a particle, and the detection of single atoms or molecules of a substance. To permit these more detailed questions to be answered, as well as to improve the speed, accuracy, sensitivity, and selectivity of traditional analysis, a large variety of physical measurements are used. These methods are based on spectroscopic, electrochemical, chromatographic, chemical, and nuclear principles.

Modern analysis has also placed significant demands on sampling techniques. It has become necessary, for example, to handle very small liquid samples [in the nanoliter (10^{-9} liter) range or less] as part of the analysis of complex mixtures such as biological fluids, and to simultaneously determine many different components. The sample may be a solid that must be converted through vaporization into a form suitable for analysis.

Spectroscopy. One important area is spectroscopy, which includes the measurement of emission, absorption, reflection, and scattering phenomena resulting from interaction of a sample with gamma rays and x-rays at the high-energy end of the spectrum and with the less energetic ultraviolet, visible, infrared, and microwave radiation. *See* SPECTROSCOPY.

Molecular spectroscopy. Lower-energy forms of excitation such as ultraviolet, visible, or infrared radiation are used in molecular spectroscopy. Ultraviolet radiation and visible radiation, which are reflective of the electronic structure of molecules, are used extensively for quantitative analysis. The radiation absorbed by the sample is measured. It is also possible to measure the radiation emitted (emission, fluorescence). The absorption of infrared radiation is controlled by the properties of bonds between atoms, and accordingly it is most widely used for structure identification and determination. It is less widely used for quantitative analysis except for gases such as carbon monoxide (CO) and hydrocarbons. X-rays are used through emission of characteristic radiation, absorption, or diffraction. In the last case, characteristic diffraction patterns reveal information about specific structural entities, such as a particular crystalline form. Extended x-ray absorption fine structure (EXAFS) is based on the use of x-rays from a synchrotron source to reveal structural details such as interatomic distances. *See* EMISSION SPECTROCHEMICAL ANALYSIS; EXTENDED X-RAY ABSORPTION

FINE STRUCTURE (EXAFS); INFRARED SPECTROSCOPY; X-RAY FLUORESCENCE ANALYSIS.

Mass spectrometry. Mass spectrometry is an important and increasingly applied method of analysis, especially for organic and biological samples. Among the applications are the analysis of more than 70 elements (spark-source mass spectrometry), surface analysis (secondary ion mass spectrometry and ion-probe mass spectrometry), and the determination of the structure of organic molecules and of proteins and peptides (high-resolution mass spectrometry, Fourier-transform mass spectrometry). *See* MASS SPECTROMETRY; SECONDARY ION MASS SPECTROMETRY (SIMS).

Nuclear magnetic resonance. This technique measures the magnetic environment around individual atoms and provides one of the most important means for deducing the structure of a molecule. Atoms possessing nuclear spin are probed by monitoring the interaction between their nuclear spin and an applied external magnetic field. For large molecules, these interactions are complex, and a variety of nuclear excitation techniques have been developed that permit establishment of the connectivity between the various atoms in a molecule. Since the technique is nondestructive, it can be used to monitor living systems. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

Electron spectroscopy. Several forms of spectroscopy are especially useful for surface analysis. The scanning electron microscope (SEM) involves a finely collimated electron beam that sweeps across the surface of a sample to produce an image. At the same time, the surface atoms are excited to emit characteristic x-rays, making it possible to obtain an image of the surface along with its spatially resolved elemental composition. The resolution of this technique (electron microprobe) is in the micrometer (10^{-4} cm) range. Images with a resolution of nanometers (10^{-9} cm) have been obtained by using the techniques of atomic force microscopy (AFM) and scanning tunneling microscopy (STM), and this resolution corresponds to the dimensions of individual atoms. A significant advantage of the latter two techniques is that a high vacuum is not required, so samples can be analyzed at atmospheric pressure. *See* ELECTRON-PROBE MICROANALYSIS; ELECTRON SPECTROSCOPY; SCANNING ELECTRON MICROSCOPE; SCANNING TUNNELING MICROSCOPE.

There are a number of other important surface analysis techniques involving the interaction of radiation with surface atoms. Low-energy electron diffraction (LEED) involves the observation of the diffraction pattern of a reflected electron beam to give structural information about the crystallographic orientation of the atoms on the surface. Several techniques are used to determine composition and structural information by measuring the energies of electrons emitted from surface atoms. These include ultraviolet photoelectron spectroscopy (UPS) and x-ray photoelectron spectroscopy (also called electron spectroscopy for chemical analysis, or ESCA), which differ in the energy of the photon beam striking the surface. Other techniques include Auger electron spectroscopy (AES) and electron energy loss spectroscopy (EELS). *See* AUGER EFFECT; ELECTRON DIFFRACTION; SURFACE AND INTERFACIAL CHEMISTRY.

Radiochemistry. The radiation emitted by a radionuclide can be determined accurately, forming the basis of a number of important analytical techniques. In some cases, these nuclides are incorporated directly into a molecule, rendering it radioactive [carbon-14 (^{14}C), hydrogen-3 (^{3}H), iodine-125 (^{125}I), phosphorus-32 (^{32}P)]. In other cases, the radioactivity is induced in the sample by bombardment with neutrons or photons (neutron or photon activation analysis). The radionuclides are identified by their characteristic radiation. For example, gamma-ray spectrometry is used to measure the characteristic radiation from nuclides emitting gamma rays. Many nuclides can be analyzed simultaneously by this method. *See* ACTIVATION ANALYSIS; GAMMA-RAY DETECTORS; RADIOACTIVITY.

Electrochemical analysis. There are a number of widely applied techniques that are based mainly on four electrochemical phenomena: (1) potentiometry, involving the relationship between the measurement of an electrical potential between two electrodes, one of which responds selectively to the concentration (activity) of a species in solution; (2) amperometry, involving the measurement of a current proportional to the concentration of a species in solution at a given applied potential between two electrodes; (3) coulometry, based on the relationship between the charge passed in an electrochemical cell and the amount of a species generated (Faraday's law); and (4) conductimetry, involving the relationship between the impedance (resistance) of a solution and the concentration of ionic species present. *See* ELECTROCHEMISTRY.

Potentiometry. This is undoubtedly the most widely applied electrochemical technique, since it includes a variety of ion-selective electrodes, the most important of which is the glass electrode used to measure pH. Other important ion-selective electrodes measure ions of sodium, potassium, calcium, sulfide, chloride, and fluoride. When the electrodes are used in conjunction with gas-permeable membranes, gases such as ammonia, carbon dioxide, and hydrogen sulfide can be measured. *See* ION-SELECTIVE MEMBRANES AND ELECTRODES; pH; POLAROGRAPHIC ANALYSIS.

Amperometry. This method depends on the detection of species that can undergo electrolysis; many substances such as oxygen are detected in this way. If the detection electrode is covered with an enzyme layer, a variety of substances in complex mixtures can be determined by taking advantage of the selectivity of the enzymatic reaction. An electroactive product of the enzymatic reaction is detected, and this configuration is called an enzyme electrode. These determinations are typically performed at a constant applied potential. A number of important techniques are based on measuring the current resulting from time-dependent variation of the potential. These techniques include cyclic

voltammetry, chronocoulometry, pulse techniques (such as pulsed amperometric detection), and chronoamperometry. They are used to study electrochemical reactions at electrodes, to study properties of electrode surfaces, and to analyze (for example, to detect an effluent from a chromatographic column or to determine a trace metal in solution).

Other phenomena. Coulometry is used in quantitatively generating analytical reagents and in connection with a system capable of detecting an end point. Conductimetry is a nonselective technique that is used to measure the overall ionic content of an electrically conducting solution. *See* COULOMETER.

Herbert A. Laitinen; George S. Wilson

Trace analysis. Trace analysis involves the determination of those elemental constituents of a sample that make up approximately 0.01% by weight of the sample or less. There is no sharp boundary between nontrace and trace constituents. The lower limit of detected concentration is set by the sensitivity of the available analytical methods and, in general, is pushed downward with progress in analytical techniques. A large number of different physical and chemical techniques have been developed for the measurement of the elemental composition at the microgram-per-gram and nanogram-per-gram level, thereby constituting the field of trace analysis.

Methods. All trace analytical methods can be divided into three component steps: sampling, chemical or physical pretreatment, and measurement. Depending upon the type of information desired in an analysis and the requirements of sensitivity, precision, and other performance figures of merit, an appropriate measurement technique is selected. Most frequently, this will be an instrumental approach. Therefore, it is essential to know the capabilities and limitations of the various methods of instrumental measurement available for trace determinations. Once they are known, appropriate steps can be taken in the sampling and pretreatment steps to provide sufficient amounts of the microconstituents that are free of interferences and in the appropriate form for the final measurement. In a number of methods the pretreatment step may be omitted, and in others the sampling and measurement occur simultaneously. In spite of possible deviations, these steps are interrelated and require different degrees of emphasis, depending upon the individual analytical situation. In many cases, the analyst uses a particular physical or physicochemical method in which manifestations of energy provide the basis of measurement. These methods are indirect in the sense that the emission or absorption of radiation or transformation of energy must be related in some way to the mass or concentration of the species that are being determined. The establishment of these relations almost invariably requires calibration, with the use of standards of known content of the constituent in question. As such, many of the available techniques do not provide absolute results.

Trace analysis ranges from the more classical chemical methods of colorimetric and absorption spectrophotometric analysis to modern instrumental approaches. The wide diversity of methods is apparent from the number of approaches and the attendant classes and subclasses in the **table**. Within any one of these categories, there are techniques that are more specialized.

Of the various criteria used in the selection of an

Some trace analysis methods and limits of detection

Method	Absolute, g	Limits of detection: concentration, ppm
Chromatography		
Thin-layer	10^{-5} to 10^{-3}	
Gas-liquid		10 to 10^{6}
Liquid-liquid		10^{-3} to 10^{0}
Electrochemical		
Coulometry	10^{-9} to 10^{0}	
Ion selective electrode		10^{-2} to 10^{2}
Polarography		
Conventional		10^{0} to 10^{3}
Modern		10^{-3} to 10^{3}
Laser probe microanalysis		10^{2} to 10^{4}
Nuclear		
Neutron activation		10^{-3} to 10^{-1}
Electron probe		10^{2} to 10^{3}
Ion probe		10^{-1} to 10^{1}
Mass spectrometry		
Isotope dilution		10^{-5} to 10^{6}
Spark source		10^{-3} to 10^{1}
Electrical plasma*		10^{-5} to 10^{0}
Organic microanalysis	$>10^{-5}$	
Optical-absorption		
Atomic		
Flame		10^{-3} to 10^{1}
Nonflame	10^{-15} to 10^{-9}	
Molecular		
UV-visible		10^{-3} to 10^{2}
Infrared		10^{3} to 10^{6}
Microwave		10^{0} to 10^{3}
Optical-emission		
AC spark		10^{1} to 10^{3}
DC arc		10^{-2} to 10^{2}
Electrical plasma†		10^{-4} to 10^{2}
Optical-fluorescence		
Atomic		
Flame		10^{-3} to 10^{2}
Nonflame	10^{-15} to 10^{-9}	
Molecular		10^{-3} to 10^{-1}
Optical-phosphorescence		10^{-3} to 10^{2}
Optical-Raman		10^{0} to 10^{5}
Spectrometric-resonance		
Nuclear magnetic		10^{1} to 10^{5}
Electron spin	10^{-9} to 10^{-6}	
Thermal analysis	10^{-5} to 10^{-4}	
Wet chemistry		
Gravimetry	10^{-3} to 10^{-2}	
Titrimetry		10^{-2} to 10^{4}
X-ray spectrometry		
Auger		10^{3} to 10^{5}
ESCA		10^{3} to 10^{5}
Fluorescence		10^{-1} to 10^{2}
Mössbauer		10^{0} to 10^{3}
Photoelectron		10^{0} to 10^{3}
Immunoassay	10^{-2} to 10^{-9}	10^{-3} to 10^{0}

*Inductively coupled plasma or microwave-induced plasma.
†Inductively coupled plasma or direct-current plasma.

appropriate trace analytical method, sensitivity, accuracy, precision, and selectivity are of prime importance. Other important considerations, such as scope, sampling and standards requirements, cost of equipment, and time of analyses, are of great practical significance.

Sensitivity and detection limits. For all analysis regimes, the measure x of some physical parameter is related to the concentration c of the analyte in a certain sample. The sensitivity S is the slope of the analytical calibration curve, which is the plot of the measure x versus the concentration of the analyte in a series of standards having known analyte concentrations. The term sensitivity must not be used to indicate either the limit of detection or the concentration required to give a certain signal.

The standard deviation σ is given by Eq. (1), where

$$\sigma = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}} \tag{1}$$

x_i is the individual measure, and $\bar{x}$ is the mean of n measurements. If n is less than about 10, σ is replaced by the term s. The relative standard deviation (RSD) is given by Eqs. (2).

$$\text{RSD} = \frac{\sigma}{\bar{x}} \quad \text{or} \quad \text{RSD} = \frac{s}{\bar{x}} \tag{2}$$

Precision is defined as the random uncertainty for the measure x of the corresponding uncertainty in the estimate of the concentration. It is often expressed as the relative standard deviation.

Accuracy is a measure of the agreement between the estimated concentration and the true value. Accuracy can never be better than the precision of the analytical procedure. Bias (systematic errors) in the calibration procedure always causes the estimated uncertainties to disagree with the true value by an amount equal to the bias.

The limit of detection c_L is the smallest concentration that can be detected with a reasonable certainty. This concentration is given by Eq. (3), where

$$c_L = \frac{x_L - \bar{x}_b}{S} \tag{3}$$

x_L is the limiting detectable measure and x_b is the average blank measure. The value for x_L is given by Eq. (4), where k is a factor determining a margin of

$$x_L = x_b + ks_b \tag{4}$$

confidence and s_b is the standard deviation of a limited number of measures of the blank. Generally, k is taken as 3; thus, the value of c_L can be determined by Eqs. (5). This means that any concentration re-

$$c_L = \frac{3s_b}{S} \quad \text{or} \quad c_L = \frac{3N_b}{S} \tag{5}$$

sulting in a signal three times the background standard deviation is considered just detectable. To evaluate x_b and s_b, at least 20 measurements should be used. If the major sources of variation are electrical noises, s_b can be replaced with N_b, the background noise level. If $3s_b$ is chosen, the confidence level is 99.86% for a purely Gaussian distribution of errors. At low concentrations, broader and asymmetric distributions are likely, so $3s_b$ corresponds to a practical confidence level of about 90%. It is noted that at the limit of detection the relative standard deviation is about 0.5; that is, there is equal confidence in deciding whether the analyte is detected or not.

As a result of widely varying pathways of development of many analytical techniques, there is a lack of consistency in the definition or specification of detection limits in the analytical literature. Consequently, it is difficult to make critical comparisons between methods for a particular element. However, since a particular element may be determined by a number of different techniques, depending upon the matrix in which it is being sought, a summary of the experimental values that have been published is pertinent. *See* ACTIVATION ANALYSIS; ANALYTICAL CHEMISTRY; CHROMATOGRAPHY; ELECTROCHEMICAL TECHNIQUES; GAS CHROMATOGRAPHY; SPECTROSCOPY.

Andrew T. Zander

Separation techniques. This collection of techniques includes various forms of chromatography and electrophoresis. They are based on the separation of a mixture of species in a sample due to differential migration. Two forces act in opposition: a stationary phase acts to retard a migrating species, while the mobile phase tends to promote migration. The mobile phase may be liquid (liquid chromatography) or gaseous (gas chromatography), while the stationary phase may be a solid or a solid covered with a thin film of liquid. The stationary phase is typically packed in a column through which the mobile phase is pumped. High-performance liquid chromatography (HPLC) has become especially important for the separation of complex mixtures of nonvolatile materials, because the use of high pressure emphasizes the differences in the rate of diffusion among the species to be separated. Separations may often be accomplished in a matter of several minutes. The stationary phase can preferentially interact with the migrating species according to charge, size, and hydrophobicity, or in some cases because of the special affinity that a species has for the stationary phase (affinity chromatography). The stationary phase can also be a thin layer of solid support deposited on a plate (thin-layer chromatography). *See* CHROMATOGRAPHY; GAS CHROMATOGRAPHY; LIQUID CHROMATOGRAPHY.

Alternatively, the driving force for separation will be the migration of charged species in an electric field (electrophoresis). The stationary phase may be a gel on a plate or in a tube, or a solution maintained in a capillary through which the analytes move. The important techniques in this area are capillary electrophoresis, isotachophoresis, and isoelectric focusing. *See* ELECTROPHORESIS.

In order to detect the various species after they are separated, most of the techniques described above have been employed. In some cases, it is desir-

able to have detection methods that are specific for the species of interest. In other cases, such specificity is not necessary since the species have already been separated and a detector with a general response can be used.

Chemical methods. Chemical and biological reactions form the basis of many analytical methods, besides the classical forms of gravimetric and volumetric analysis. From biology have come a number of widely used techniques: immunoassays, enzyme-catalyzed reactions, gene probes, and bioluminescence. These methods are often based on the rate of a reaction, which is proportional to the concentration of the analyte of interest. *See* BIOLUMINESCENCE; IMMUNOASSAY; KINETIC METHODS OF ANALYSIS.

Thermal methods. Thermal methods are based on the heating of a sample over a range of temperatures. This approach may result in absorption of heat by the sample or in evolution of heat due to physical or chemical changes. Thermogravimetry involves the measurement of mass; differential thermal analysis involves a detection of chemical or physical processes through a measurement of the difference in temperature between a sample and a stable reference material; differential thermal calorimetry evaluates the heat evolved in such processes. A variety of calorimetric techniques are used to measure the extent of reactions that are otherwise difficult to evaluate. *See* CALORIMETRY.

Sample handling and processing methods. As the demands increase for more rapid and reliable analysis at lower cost, automation of sampling and sample handling, measurement, and data analysis have become extremely important. Applications may involve monitoring a manufacturing process stream, air quality, or critical diagnostic analyses of a patient undergoing surgery. Fortunately, the advent of low-cost but powerful computers has greatly aided development in this area.

Computers. The digital computer has had a major impact on modern analytical chemistry. The microprocessor has been incorporated as an integral part of a variety of analytical instruments to permit detailed programming of their operation. More complex programming and automation are possible through the personal computer, which is being used not only to operate a single instrument but also to tie various operations together. These systems are increasingly linked to networks, which provide for storage of and access to data. *See* DIGITAL COMPUTER; MICROCOMPUTER; MICROPROCESSOR.

Chemometrics. Chemometrics refers to the use of statistical and other mathematical approaches to data handling. An important application is the optimization of an instrumental procedure by processing the initial instrument output rapidly and feeding the information back in order to control the settings of the instrument, thereby achieving optimal accuracy and sensitivity from the instrument. Still more complex computer operations, such as pattern recognition, are designed to draw conclusions through comparisons of arrays of data. *See* CHEMOMETRICS.

Automation. Analytical methods can be automated by several approaches. Clinical methods are often carried out by means of an automated analyzer. The sample is introduced into a flowing stream (unsegmented flow) or separated from the next sample by an air bubble (segmented flow), and a sequence of analytical operations is performed on the sample, culminating in a final measurement step based on an instrumental reading. The instrument is calibrated by running a standard sample through the system. A computer readout compares the analytical result with the normal range expected for the analysis. Other physical principles are also used in automated analytical systems. In robotics, a computer-controlled robot carries out the routine steps of chemical methods, including sample preparation and cleanup, weighing, dissolution, adjustment of conditions, reagent manipulation, and instrument reading. This capability increases sample throughput and reduces errors caused in many cases by operator boredom and inattention due to the repetitive nature of the manipulations. Such analyses can be performed very rapidly with small volumes (nanoliters) of solution. *See* COMBINATORIAL CHEMISTRY; ROBOTICS.

Instrumentation. As there is a constant need for improved instrumentation, the development of new instrumentation, rather than simple application of old techniques, is an area of active interest in analytical chemistry research. *See* CHEMISTRY; INSTRUMENT SCIENCE.

George S. Wilson

Bibliography. D. C. Harris, *Quantitative Chemical Analysis*, 6th ed., 2002; A. G. Howard and P. J. Statham, *Inorganic Trace Analysis: Philosophy and Practice*, 1994; J. R. Lawrence (ed.), *Trace Analysis*, vol. 1, 1981, vol. 2, 1982; J. F. Rubinson and K. A. Rubinson, *Contemporary Chemical Analysis*, 1998; D. A. Skoog et al., *Principles of Instrumental Analysis*, 5th ed., 1998.

Anamnia

Those vertebrate animals, sometimes called Anamniota, which lack an amnion in development. The amnion is a protective embryonic envelope that encloses the embryo and its surrounding liquid, the amniotic fluid, during fetal life. An amnion is present in mammals, birds, and reptiles, but is absent in fishes and amphibians. In early classifications the vertebrates were commonly separated on this basis, and the expressions Amniota and Anamniota are still useful in grouping higher and lower vertebrates. It should be recognized that these terms represent grades of development, however, and do not carry the connotation of established classificatory ranks. Anamnia, then, is a group name that includes the Recent classes Agnatha, Chondrichthyes, Osteichthyes, and Amphibia and, by presumption, the class Placodermi, which is known only from fossils. *See* AMNION; AMNIOTA; AMPHIBIA; PISCES (ZOOLOGY).

Reeve M. Bailey

Anaphylaxis

A generalized or localized tissue reaction occurring within minutes of an antigen-antibody reaction. Similar reactions elicited by nonimmunologic mechanisms are termed anaphylactoid reactions. The term anaphylaxis was introduced by C. R. Richet to describe acute adverse reactions which followed reinjection of an eel toxin into dogs. Instead of becoming immune (nonreacting) to the toxin, the animals developed severe and often fatal congestion and hemorrhage of the gastrointestinal tract, lungs, pleura, and endocardium.

It was initially thought that anaphylaxis in humans was associated with precipitating antibody (known today to be of the IgG class of immunoglobulins), and that it differed from the common allergic manifestations of hay fever due to "reaginic" antibody. The reaginic antibody has now been identified as IgE. *See* ALLERGY; IMMUNOGLOBULIN.

Signs and symptoms. In animals undergoing anaphylaxis, one organ system is usually affected more than others (thus the term shock-organ). In the guinea pig, the manifestations of anaphylaxis are respiratory, with cough, wheezing, and gasping respirations caused by severe bronchoconstriction. In the dog and the rat, the manifestations are gastrointestinal with hemorrhage. Severe pulmonary vasoconstriction leading to heart failure and death prevail in the rabbit. In humans, the clinical manifestations of anaphylaxis include reactions of the skin with itching, erythema, and urticaria; the upper respiratory tract with edema of the larynx; the lower respiratory tract with dyspnea, wheezing, and cough; the gastrointestinal tract with abdominal cramps, nausea, vomiting, and diarrhea; and the cardiovascular system with hypotension and shock. Individuals undergoing anaphylactic reactions may develop any one, a combination, or all of the signs and symptoms.

Anaphylaxis may be fatal within minutes, or may occur days or weeks after the reaction, if the organs sustained considerable damage during the hypotensive phase. Tissue changes in acute anaphylaxis are minimal and include edema and a mild eosinophilic infiltration. Myocardial ischemia and ischemic changes of the renal tubules and the brain may be seen after prolonged hypotension. Macroscopic changes are variable, and in fatal cases usually involve the respiratory system, where pulmonary edema, pulmonary infiltrates, and increased bronchial secretions are noted.

Mechanism. Anaphylaxis in humans is most often the result of the interaction of specific IgE antibody fixed to mast cells and antigen. Two molecules of IgE are bridged by the antigen, which may be a complex protein or chemical (hapten) bound to protein. The antigen-antibody interaction leads to increased cell-membrane permeability, with influx of calcium and release of either preformed or newly formed pharmacologic mediators from the granules. Preformed mediators include histamine and eosinophilic or neutrophilic chemotactic factors. Newly formed molecules include leukotrienes or slow-reacting substance of anaphylaxis and prostaglandins. The mediator action induces bronchoconstriction, vasodilation, cellular infiltration, and increased mucus production. Degranulation of mast cells and basophils is augmented when intracellular cyclic guanine monophosphate (cGMP) is increased; the reaction is inhibited in the presence of increasing intracellular concentrations of cyclic adenine monophosphate (cAMP). *See* EICOSANOIDS.

Another mechanism for induction of anaphylaxis in humans occurs when antigen binds to preformed IgG antibody and complement components interact with the antigen-antibody complex. The early components of the complement system bind to the antibody molecule, leading to activation of other complement components. During the activation, components known as anaphylatoxins (C3a and C5a) are released which may directly cause bronchoconstriction with respiratory impairment, and vasodilation with hypotension or shock. Activation of the complement system by agents like iodinated radiocontrast media, or degranulation of mast cells by drugs like opiates, will elicit similar manifestations without an antibody-antigen interaction. *See* COMPLEMENT.

Anaphylaxis due to IgE mechanisms has been associated with foreign proteins such as horse antitoxins, insulin, adrenocorticotropic hormone (ACTH), protamine, and chymopapain injected into herniated discs; drugs such as penicillin and its derivatives; foods such as shellfish, nuts, and eggs; and venom of stinging insects. Anaphylaxis mediated by IgG is seen in blood-transfusion reactions and following the use of cryoprecipitate, plasma, or immunoglobulin therapy. Direct mast-cell degranulation with release of mediators occurs with opiates, radiocontrast media, and curare.

In certain individuals, mast cell degranulation occurs after exercise and following exposure to cold. Individuals with systemic mastocytosis spontaneously release mediators from the increased number of tissue mast cells, which may cause an anaphylactic reaction. When no cause for the anaphylaxis can be found, the reaction is termed idiopathic.

Prevention and treatment. After the identification of the inciting agent for the anaphylactic reaction, prevention is the best mode of therapy. Individuals should not be given antibiotics to which they are sensitive, and foods which cause reactions should be avoided. Individuals sensitive to insect-sting venom should take preventive measures to avoid stings and should use insect repellents. Individuals who react to exercise should avoid such activities; when cold is the cause of anaphylaxis, warm clothes should be worn, and submerging in the cold water of swimming pools should be avoided. Immunotherapy with insect venom and desensitization with certain drugs are effective prophylactic measures.

Individuals with recurrent episodes of anaphylaxis, when the etiological cause is unknown and preventive measures are impractical, should be provided with epinephrine in a form that can be self-administered whenever symptoms occur.

Administration of corticosteroids may also be helpful. Such prophylaxis may be life-saving. *See* EPINEPHRINE.

The treatment of anaphylaxis is aimed at reducing the effect of the chemical mediators on the end organs and preventing further mediator release. The drug of choice for this is epinephrine given subcutaneously in repeated doses. Additionally, a clear airway and appropriate oxygenation must be maintained; hypotension should be treated, as should any cardiac arrhythmia. Aminophyllin, a rapid-onset bronchodilator, and corticosteroids, which alter prostaglandin and leukotriene metabolism but have an onset of action in several hours, are second-line indications to be used in patients who do not respond adequately to epinephrine. Cromolyn sodium, known to prevent mast-cell degranulation, has been effective in preventing bronchoconstriction, but has no place in the treatment of the acute anaphylactic reaction. Such effective therapeutic measures must be carried out as soon as possible because of the potentially fatal nature of anaphylaxis. *See* HYPERSENSITIVITY; SHOCK SYNDROME.

Shlomo Bar-Sela; Jordan N. Fink

Bibliography. E. Bacal, R. Patterson, and C. R. Zeiss, Evaluation of severe (anaphylactic) reactions, *Clin. Allergy*, 8:295, 1978; E. Middleton, C. E. Reed, and E. F. Ellis (eds.), *Allergy: Principles and Practices*, 4th ed., 1993; R. Patterson (ed.), *Allergic Diseases: Diagnosis and Management*, 5th ed., 1997; P. L. Smith et al., Physiologic manifestations of human anaphylaxis, *J. Clin. Invest.*, 66:1072–1080, 1980.

Anaplasmosis

A disease of mammals caused by a specialized group of gram-negative bacteria in the order Rickettsiales, family Anaplasmataceae, genus *Anaplasma*. *Anaplasma* is an obligate intracellular parasite that infects erythrocytes (red blood cells) of cattle, sheep, goats, and wild ruminants in much of the tropical and subtropical world. A newly reclassified species (*Anaplasma phagocytophilum*) that infects the white blood cells of mammals was recently added to this genus. However, the most important species and the one traditionally associated with the disease anaplasmosis is *A. marginale*. Anaplasmosis is one of the most important diseases of cattle, causing anemia and sometimes death, resulting in significant economic losses worldwide.

Occurrence and transmission. *Anaplasma* resides in erythrocytes within a vacuole formed from the host cell membrane. The rickettsiae remain within the vacuole and multiply by binary fission, producing numerous rickettsiae. The vacuole—which can be identified in stained blood films as a round, densely staining body approximately 1 micrometer in diameter—is referred to as an inclusion body. *Anaplasma marginale* contains both deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), with a total genome size of 1200–1260 kilobase pairs (kbp). Molecular studies based on the analysis of ribosomal RNA relate *Anaplasma* closely to *Ehrlichia* spp.

Transmission of *Anaplasma* occurs biologically by multiple species of ticks. In the United States *Dermacentor* spp ticks are the primary vectors, but some 20 species act as vectors worldwide. Rickettsiae are ingested with a blood meal and undergo complex development beginning in the tick gut cells. *Anaplasma* is transmitted from the salivary glands while ticks feed on the host. Infections acquired by one tick stage may be transmitted after molting to the next stage (transtadial). Ticks may remain infected indefinitely (without feeding) and serve as a reservoir for infection of susceptible cattle. Mechanical transmission occurs when the mouthparts of biting flies (such as horse flies and stable flies) become contaminated while feeding on infected cattle and then quickly move to uninfected animals. There is no development of rickettsiae in these mechanical vectors. Contaminated needles and instruments for dehorning, castration, and tagging may also transmit the organism throughout a herd. Isolates of *Anaplasma* from different geographic areas differ from each other in biology, morphology, and antigenic characteristics.

Epidemiology. Once cattle are infected with *Anaplasma*, they remain persistently infected for life and may serve as reservoirs of infection for other cattle or tick vectors. During this time the parasitemia (presence of parasites in the blood) fluctuates and at times may be undetectable. The spread of anaplasmosis may occur when these carrier cattle are moved to nonendemic areas. Ticks have been shown to transmit new infections after feeding on animals with undetectable parasitemias. Carrier animals may also have relapse infections, producing new infections in susceptible herds.

Five major surface proteins of *Anaplasma* have been identified, and some have been found to be encoded by polymorphic multigene families. This variability results in distinct geographic strains of *Anaplasma* which are antigenically different and complicate vaccine strategies.

Clinical disease. Once an animal is infected it takes approximately 21 days before organisms can be identified in blood smears; but animals may not exhibit clinical disease for as long as 60 days postinfection. In acute anaplasmosis, cattle exhibit depression, loss of appetite, increased temperature, labored breathing, dehydration, jaundice, and a decrease in milk production. Erythrocyte count, packed cell volume, and hemoglobin values decrease significantly, and death may result from severe anemia. Abortions may also occur following infection. These clinical manifestations of anemia are associated with the rapid removal of infected and uninfected erythrocytes by macrophages (phagocytic cells of the immune system). Recovered animals gradually regain condition but remain chronically infected and are subject to periodic relapses. They may also serve as a source of infection for susceptible animals. The disease is most severe in older animals not previously exposed; calves up to 1 year old become infected but usually

have a mild or subclinical (showing no symptoms) reaction.

Diagnosis. Parasites can be identified as a round inclusion body at the periphery of red blood cells in a Giemsa-stained blood film. Infection can be confirmed by identifying specific antibodies in serological tests such as the competitive enzyme-linked immunosorbent assay (C-ELISA). Molecular amplification techniques such as the polymerase chain reaction (PCR) are also used, particularly when parasite density is very low or undetectable.

Immunity. Cattle that recover from clinical anaplasmosis are protected from subsequent exposure to the same geographic isolate of *Anaplasma*. Calves have a natural immunity to clinical disease if exposed within their first year. Protection is attributed to a complex combination of antibody and cell-mediated responses, with initial responses directed against the outer-membrane proteins of *Anaplasma*. Animals develop high levels of antibody during acute and recovery phases of anaplasmosis, but passive transfer of the antiserum (the serum component of blood that contains antibodies specific to one or more antigens) does not convey protection to naive animals (which have not been previously exposed to the disease). Splenectomy of recovered animals results in recrudescence (return of symptoms) of the rickettsemia (presence of rickettsiae in the blood), suggesting a role for cellular immunity in protection. Cell-mediated mechanisms of protection may involve macrophage activation and specific T-lymphocyte responsiveness. Specific antibodies may coat infected erythrocytes, enhancing their susceptibility to phagocytosis by activated macrophages. The ability of the rickettsiae to persist within the host at low levels despite a strong immune response is thought to be the result of antigenic variants created during cyclical rickettsemia. *See* ANTIBODY; ANTIGEN; CELLULAR IMMUNOLOGY; IMMUNITY.

Control and treatment. The clinical manifestations of anaplasmosis can be halted or prevented if tetracycline antibiotics are administered early. The drug inhibits rickettsial protein synthesis but does not kill the organism. Tetracyclines have been added to feeds to prevent clinical symptoms, but this does not prevent infection. Controlling fly and tick populations by chemical spraying or dipping have also been used to reduce the spread of disease.

Vaccination with killed *A. marginale* from erythrocytes has been shown to provide some protection against the acute disease but it does not prevent infection. Blood-derived vaccines are no longer used in most of the United States. The related and less virulent species, *A. centrale*, is used as a live-blood vaccine in some countries and provides some protection against *A. marginale*, but it can revert to a virulent form. Recombinant surface proteins of *A. marginale* have been evaluated as candidate vaccines and may be included in next-generation vaccines. In areas where anaplasmosis is endemic and control measures are not applied, most calves become infected, resulting in endemic stability with minimal clinical disease but high infection rates. *See* ANTIBIOTIC; VACCINATION.

Economic impact. Anaplasmosis continues to be a major disease of cattle in the tropical and subtropical world. Economic losses result from antibiotic control and vaccine costs, vector control, and production losses (such as decreased weight gain) from infected and recovered cattle and deaths. It is the most important tick-borne disease of cattle in the United States. *See* RICKETTSIOSES. Edmour F. Blouin

Bibliography. A. F. Barbet, Recent developments in the molecular biology of anaplasmosis, *Vet. Parasitol.*, 57:43–49, 1995; E. F. Blouin et al., Applications of a cell culture system for studying the interaction of *Anaplasma marginale* with tick cells, *Animal Health Res. Rev.*, 3(2):57–68, 2002; J. de la Fuente et al., Molecular phylogeny and biogeography of North American isolates of *Anaplasma marginale* (Rickettsiaceae: Ehrlichiae), *Vet. Parasitol.*, 97:65–76, 2001; D. M. French et al., Expression of *Anaplasma marginale* major surface protein 2 variants during persistent cyclic rickettsemia, *Inf. Immun.*, 66(3):1200–1207, 1998; K. M. Kocan, Adaptations of the tick-borne pathogen, *Anaplasma marginale*, for survival in ticks and cattle, *Exp. Appl. Acarol.*, 28:9–25, 2002; G. H. Palmer and T. F. McElwain, Molecular basis for vaccine development against anaplasmosis and babesiosis, *Vet. Parasitol.*, 57:233–253, 1995.

Anapsida

Formerly, a group of reptiles that was recognized in rank-based classifications. Amniotes were once divided into subclasses, based on the patterns of temporal fenestrae (openings in the skull roof posterior to the openings for the eyes). Subclass Synapsida included taxa that exhibit a single pair of temporal fenestrae (synapsid condition). Subclass Diapsida included taxa that feature two pairs of temporal fenestrae (diapsid condition) and those that are descended from ancestors that had them. The absence of temporal fenestrae is known as the anapsid condition. Mesosaurids, pareiasaurids, and captorhinids are examples of fossil reptiles that lack temporal fenestrae, whereas turtles are the only living reptiles that lack them; these reptiles were placed together in the subclass Anapsida. Certain Paleozoic reptiles (for example, millerettids) that exhibited synapsid-like temporal fenestrae but clearly did not belong in Synapsida (or Diapsida) were placed in Anapsida. *See* AMNIOTA; CAPTORHINIDA; CHELONIA; DIAPSIDA; MESOSAURIA; REPTILIA; SYNAPSIDA.

Studies of early amniotes that use cladistic methodology, in which derived features are used to group taxa, revealed that the absence of temporal fenestrae is a primitive feature. Moreover, amniote taxa that had been placed in the subclass Anapsida in rank-based classifications were found to form a paraphyletic (artificial) group at the base of Reptilia in cladograms. Phylogenetic systematists refuse to recognize artificial groups with formal taxon names;

but with the advent of phylogenetic nomenclature, attempts were made to conserve traditional group names by linking them with extant representatives, in order to adapt them as "clade names." Thus, Anapsida was linked to a clade that included turtles. In the earliest cladistic studies of amniotes, turtles and captorhinids formed a clade, which was called Anapsida. Subsequent studies placed turtles in a clade that contained pareiasaurs, mesosaurs, and their close relatives, and the name Anapsida was transferred to that clade. That act generated confusion because captorhinids, which are regarded by paleontologists as exemplary fossil reptiles with anapsid skulls, were no longer to be regarded as anapsid reptiles. *See* ANIMAL SYSTEMATICS; PHYLOGENY; TAXONOMY.

Recent discoveries have led taxonomists to advocate the abandonment of Anapsida as a taxonomic entity. Paleontological research indicates that synapsid-like temporal fenestrae arose at least four times in the clade of pareiasaurs, mesosaurs, and their close relatives. Morphological and molecular studies now place turtles in Diapsida, a corollary of which is that the anapsid condition of turtles can be interpreted as a derived feature of turtles. If turtles are truly nested within Diapsida, the name Anapsida can no longer be used as a clade name (because the phylogenetic definition for Anapsida used Diapsida as a reference taxon).

Whereas it is no longer appropriate to use the term anapsid in a taxonomic context, it can still be used in a morphological sense to describe the skull of an amniote that lacks temporal fenestrae. Thus, the anapsid condition was common among Paleozoic amniotes, where it represented the retention of a primitive feature in those taxa. Conversely, the absence of temporal fenestrae in turtles is a derived feature of these reptiles (if the diapsid origin of turtles is correct). Sean Modesto

Bibliography. M. J. Benton, *Vertebrate Palaeontology*, 3d ed., Blackwell, Oxford, 2004; R. L. Carroll, *Vertebrate Paleontology and Evolution*, W. H. Freeman, New York, 1988; A. S. Romer, *Vertebrate Paleontology*, 3d ed., University of Chicago Press, 1966.

Anaspida

A group of fossil jawless fishes, roughly similar in appearance to extant lampreys, known from Early Silurian–Early Devonian deposits (about 410–430 million years ago) that represent marine, coastal environments. Anaspid fossils are known predominantly from the Northern Hemisphere.

Most anaspids, such as *Rhyncholepis*, are small, rarely exceeding 15 cm (6 in.), with a rather slender and laterally compressed body. A slanting row of 7–15 branchial (gill) openings is present between the head region and the trunk. Triradiate postbranchial spines at the ventral end of the branchial row characterize the group. The head of most anaspids is covered by a series of smaller plates, in contrast to the massive headshield seen in most other contemporary jawless fishes (such as Osteostraci, Heterostraci), and the trunk bears rod-shaped scales arranged in a chevronlike pattern in five variously inclined overlapping rows. The relatively fragile nature of this dermal skeleton means that well-preserved articulated fossils are rare but isolated scales and plates are fairly common. The anaspids are also characterized by a median dorsal row of peculiar, often hook-shaped spines/scales. The tail is strongly hypocercal (the muscular lobe containing the notochord turns downward). Some articulated specimens have traces of paired ventrolateral fins, which vary in length from genus to genus (see **illustration**).

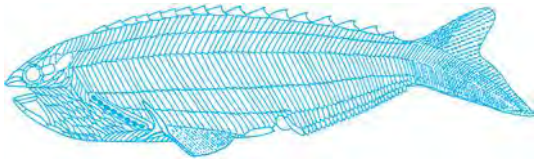

Reconstruction of *Rhyncholepis parvula* from the Silurian of Norway. (*Modified from A. Ritchie, 1980; reproduction by permission of Royal Society of Edinburgh, Trans. Roy. Soc. Edinburgh: Earth Sci., 92:263–323, 2002*)

Due to their close external similarities to living lampreys, anaspids have traditionally been regarded as close relatives or ancestors of the former. However, current views on the relationships of early vertebrates suggest that anaspids, together with other fossil jawless vertebrates, are more closely related to the jawed vertebrates (gnathostomes) than to lampreys. Their exact position, though, on the family tree is still a matter of considerable debate, mainly due to a lack of information regarding their internal anatomy. *See* ACANTHODII; HETEROSTRACI; JAWLESS VERTEBRATES; OSTEOSTRACI. Henning Blom

Bibliography. H. Blom, T. Märss, and G. C. Miller, Silurian and earliest Devonian Birkeniid anaspids from the Northern Hemisphere, *Trans. Roy. Soc. Edinburgh: Earth Sci.*, 92:263–323, 2002; P. Janvier, *Early Vertebrates*, Oxford Monogr. Geol. Geophys. 33, Oxford University Press, 1996.

Anaspidacea

An order of the crustacean superorder Syncarida, the most primitive of the Eumalacostraca. The families Anaspididae and Koonungidae have living representatives, while the Clarkecarididae are true fossils. The families Gampsonychidae and Palaeocarididae, formerly included in this order, now are assigned to a new order, Palaeocaridacea. The Anaspidacea arose during the Paleozoic Era, and the recent species are now restricted as living fossils to freshwater habitats in Tasmania and South Australia (not New Zealand). They exhibit a great adaptability to special habitats. The first thoracic somite is incorporated with cephalic tagmata.

Anaspididae. *Anaspides tasmaniae* has a body which is similar to that of the amphipods. It is about 2 in. (5 cm) long and is found in mountain lakes and cold rivers at elevations of 2000–4000 ft

(600–1200 m) on Mount Wellington. *Paranaspides lacustris*, from the same region, is 1.2 in. (3 cm) long and resembles *Mysis*. *Micrasides calmani* is the smallest species and is found associated with *Sphagnum*. Anaspidites antiquus is from the Triassic of Australia. *See* AMPHIPODA.

Koonungidae. *Koonunga cursor* is 0.8 in. (2 cm) long and is found in the mud bottom of temporary springs near Melbourne. The eyes are sessile and the first thoracic limb is modified for digging. *See* SYNCARIDA.

Hans Jakobi

Bibliography. R. D. Barnes, *Invertebrate Zoology*, 1968; H. K. Brooks, On the fossil Anaspidacea, with a revision of the classification of the Syncarida, *Crustaceana*, 4:239–242, 1962; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; R. Siewing, Uber die Verwandtschaftsbeziehungen der Anaspidaceen, *Zool. Anz. Suppl.*, 18:242-252, 1955; G. Smith, On the Anaspidacea, living and fossil, *Quart. J. Microsc. Sci.*, 53:489-578, 1909.

Anatomy, regional

The detailed study of the anatomy of a part or region of the body of an animal, most commonly applied to regional human anatomy. This is in contrast, but supplementary, to the study of organ systems, such as the cardiovascular, where all the structures pertaining to the system are studied in their continuity. To obtain the maximum information about the anatomy of an animal, both methods are used, as well as histological, embryological, and functional studies.

The various regions of the body have particular interest for certain medical and other scientific specialists. For example, the obstetrician has a detailed knowedge of the perineal region, and the ophthalmologist of the eye. Regional anatomy is of great importance to the surgeon, and certain texts concerning it are labeled surgical anatomy.

Closely related to regional anatomy is topographic anatomy, in which bony and soft tissue landmarks on the surface of the body are used to indicate the known location of deeper structures. An example is the point halfway between the umbilicus and the ventral prominence of the right hipbone; this marks the approximate location of the vermiform appendix and is known as McBurney's point.

There are many methods of dividing the body into regions for study, and one such means of classification is shown in the **illustration**. This system includes only externally visible areas; other systems would include special internal regions as well.

The head, trunk, and extremities are the principal regions. Subdivision of these can be carried out, as illustrated, so that many distinct areas or especially vital regions are indicated. There is no end to the dividing and subdividing a specialist may do to make his task eventually less difficult.

Regional anatomy is the natural and historical approach to the study of human and animal forms. The ancient scientists did not know much about the systems of the body except through conjecture. They approached the dissection of animals and cadavers in a manner intended to display the structures of some part, such as limb. The integration of such information led to the development of systematic anatomy.

Thomas S. Parsons

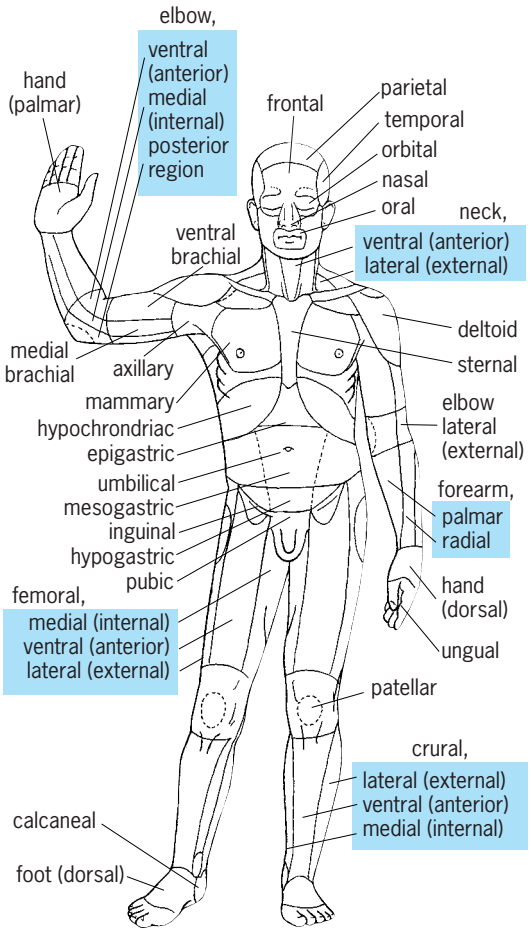

Some of the major regions of the human body seen in ventral view. (*After B. Anson, ed., Morris' Human Anatomy, 12th ed., McGraw-Hill, 1966*)

Andalusite

A nesosilicate mineral, composition Al_2SiO_5, crystallizing in the orthorhombic system. It occurs commonly in large, nearly square prismatic crystals. The variety chiastolite has inclusions of dark-colored carbonaceous material arranged in a regular manner. When these crystals are cut at right angles to the *c* axis, the inclusions form a cruciform pattern (**Fig. 1**). There is poor prismatic cleavage; the luster is vitreous and the color red, reddish-brown, olive-green, or bluish. Transparent crystals may show strong dichroism, appearing red in one direction and green in another in transmitted light. The specific gravity is 3.1–3.2; hardness is 7.5 on Mohs scale, but may be less on the surface because of alteration.

Occurrence and use. Andalusite was first described in Andalusia, Spain, and was named after this locality. It is found abundantly in the White Mountains near

Fig. 1. Andalusite, variety chiastolite. Prismatic crystal specimens from Worchester County, Massachusetts. (*American Museum of Natural History specimens*)

Laws, California, where for many years it was mined for manufacture of spark plugs and other highly refractive porcelain. Chiastolite, in crystals largely altered to mica, is found in Lancaster and Sterling, Massachusetts. Water-worn pebbles of gem quality are found at Minas Gerais, Brazil. *See* SILICATE MINERALS.

Cornelius S. Hurlbut, Jr.

Aluminum silicate phase relations. The three polymorphs of Al_2SiO_5 are andalusite (Fig. 1), kyanite, and sillimanite. These minerals occur in medium- and high-grade metamorphic rocks, and andalusite and sillimanite occur rarely in granites. Laboratory determination of the pressure-temperature stability fields of these minerals has been difficult and remains a subject of controversy, but the phase diagram shown in **Fig. 2** is generally accepted. It is consistent with thermochemical data and natural occurrence of the minerals. The univariant boundaries of the three stability fields meet at an invariant point located at 3.8 kilobars (380 megapascals) and 930°F (500°C).

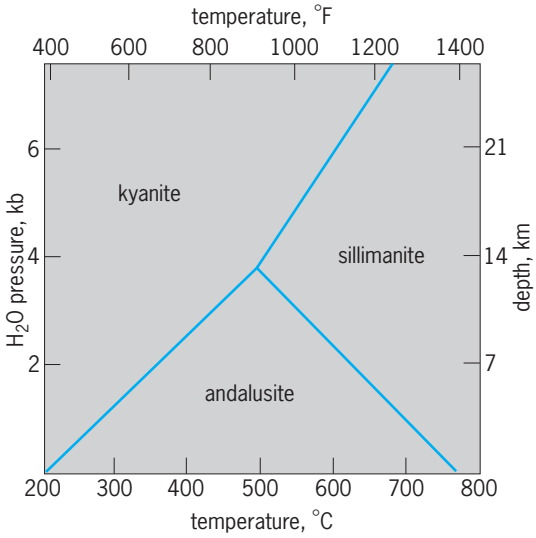

Fig. 2. Phase diagram for the polymorphs of aluminum silicate (Al_2SiO_5). 1 km = 0.6 mi; 1 kb = 10^2 MPa. (*After M. J. Holdaway, Stability of andalusite and the aluminum silicate phase diagram, Amer. J. Sci., 271:97–131, 1971*)

Grade (degree) of metamorphism is commonly a function of increasing temperature at some nearly constant or slowly changing pressure. Temperature variation is a reflection of changing temperature gradients (or heat sources) in the Earth, and pressure is an indication of thickness of overlying rock at the time of metamorphism. In pelitic (former shale) schists the grade or degree of metamorphism may be expressed by index minerals, such as the aluminum silicates. In regional metamorphic terranes the sequence of kyanite followed by sillimanite with increasing grade or temperature corresponds to pressures above 3.8 kbar (380 MPa) or depths greater than 8 mi (13 km) in the Earth's crust. In other regional metamorphic terranes the sequence andalusite to sillimanite indicates shallower levels in the crust (Fig. 2). For the New England region of the United States, the sequence is kyanite to sillimanite in Vermont, western New Hampshire,western Massachusetts, and Connecticut, while in Maine, eastern New Hampshire, eastern Massachusetts, and Rhode Island the sequence is andalusite to sillimanite, all of which formed during Early Devonian metamorphism. This indicates that central and western New England were metamorphosed at greater pressure (deeper in the crust) than northern and eastern New England. Thus, natural occurrence of aluminum silicates combined with knowledge of their phase relationships is a useful aid in deducing extent of vertical movement and erosion occurring in the Earth's crust after metamorphism. *See* KYANITE; METAMORPHISM; SILLIMANITE.

M. J. Holdaway

Bibliography. W. A. Deer, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, vol. 1a: *Orthosilicates* 2d ed., 1997; M. J. Holdaway, Stability of andalusite and the aluminum silicate phase diagram, *Amer. J. Sci.*, 271:97–131, 1971; D. M. Kerrick, The Al_2SiO_5 Polymorphs, *Reviews in Mineralogy*, vol. 22, 1990; C. Klein, *Manual of Mineralogy*, 21st ed., 1993; E-an Zen, W. S. White, and J. B. Hadley (eds.), *Studies of Appalachian Geology: Northern and Maritime*, 1968.

Andesine

A plagioclase feldspar with composition $Ab_{70}An_{30}$ to $Ab_{50}An_{50}$ (Ab = $NaAlSi_3O_8$; An = $CaAl_2Si_2O_8$). Andesine occurs primarily in igneous rocks, often in a glassy matrix as small, chemically zoned, lathlike crystals known as microlites. The rock types may be called andesinites (if dominantly feldspar), andesites (see **illus.**), andesitic basalts (or olivine-bearing andesites, as in Hawaiian lava flows), or pyroxene-, hornblende- or biotite-andesites (depending on the second most dominant mineral—all are volcanic). *See* ANDESITE.

The symmetry of andesine is triclinic, hardness on the Mohs scale 6, specific gravity 2.69, melting point ~1210°C (2210°F). If quenched at very high

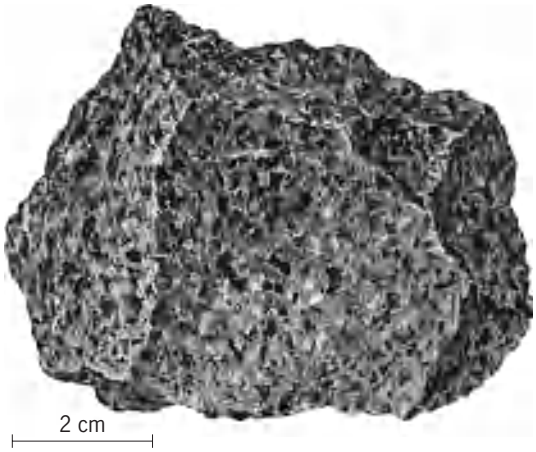

Andesine grains with biotite in a specimen from Zoutpansberg, Transvaal, South Africa. Andesine is rarely found except as grains in igneous rocks.

temperatures, andesine has an albitelike structure, with aluminum (Al) and Silicon (Si) essentially disordered in the structural framework of the crystals. But, in the course of cooling, most natural andesines develop an Al-Si ordered structure with unusual, strain-induced, periodic features detectable only by single-crystal x-ray diffraction or by transmission electron microscopy. These are called e-plagioclases, based on the presence of non-Bragg diffraction maxima (e-reflections) in x-ray or electron patterns. *See* ALBITE; CRYSTAL STRUCTURE.

Calcic andesines and labradorites ($Ab_{55}An_{45}$–$Ab_{40}An_{60}$) may exsolve into two distinctly intergrown lamellar phases whose regularity of stacking produces beautiful interference colors like those in the feathers of a peacock. Polished specimens of this material are called spectrolite in the gem trade, and at some localities (notably eastern Finland) crystals up to 10 in. (25 cm) are mined by hand. Smaller crystals are made into cabochons for jewelry. They may be abundant enough in the host rock to be valued as a decorative stone. Hematite (Fe_2O_3) is present in rare oligoclases, and some andesines as micrometer-scale thin flakes, oriented parallel to certain structurally defined planes, producing brilliant reddish-gold reflections. These semiprecious stones are called aventurine or sunstone. *See* FELDSPAR; GEM; HEMATITE; IGNEOUS ROCKS; LABRADORITE.

Paul H. Ribbe

Andesite

A typical volcanic rock erupted from a volcano associated with convergent plate boundaries. The process of subduction, which defines convergent plate boundaries, pushes oceanic lithosphere beneath either oceanic lithosphere or continental lithosphere. Andesites are the principal rocks forming the volcanoes of the "ring of fire," the arcuate chains of volcanoes which rim the Pacific Ocean basin. The Marianas and Izu-Bonin islands, the islands of Japan, the Aleutian Islands, the Cascades Range of the northwest United States, the Andes mountain chain of South America, and the Taupo Volcanic Zone of New Zealand (**Fig. 1**) are andesitic. *See* LITHOSPHERE; PLATE TECTONICS; VOLCANO.

Andesite volcanoes have been responsible for some of the most destructive, globally significant eruptions of historic and modern times. Examples include the 1815 eruption of Tambora and the 1883 eruption of Krakatau in Indonesia, the 1902 eruption of Mount Pelee in the West Indies, the 1980 eruption of Mount St. Helens in the United States, the 1991 eruption of Pinatubo in the Philippines, and the 1995–1998 eruption of the Soufriere Hills in the West Indies. The activity of most active andesite volcanoes is monitored continuously, but loss of life can still occur. For example, mud flows (lahars) from Nevado del Ruiz volcano in Colombia caused around 25,000 deaths in November 1985.

Andesites are mostly dark-colored vesicular volcanic rocks which are typically porphyritic (containing larger crystals set in a fine groundmass). Phenocrysts (the larger crystals) comprise plagioclase; calcium-rich, calcium-poor pyroxene; and iron-titanium oxides set in a fine-grained, frequently glassy, groundmass. Some andesites contain phenocrysts of olivine, and some contain amphibole and biotite; these latter rocks generally contain more potassium. The porphyritic nature of andesites is derived from a complicated history of magmatic crystallization and evolution as the melts rise toward the surface from deep in the Earth. Phenocryst minerals commonly are strongly zoned and show evidence for disequilibrium during growth, consistent with an origin involving crystal fractionation and mixing processes. Andesites are readily classified in terms of their silicon dioxide (SiO_2) content, between 53 and 63 wt %, and potassium oxide (K_2O) content at a given SiO_2 content (**Fig. 2***a*; **table**). They can also be readily discriminated on a total alkali versus SiO_2 diagram (the TAS diagram; Fig. 2*b*). Most andesite volcanoes erupt lavas and tephras (volcanic ash) which range in composition from basaltic andesite to dacite. Eruptions are often explosive, reflecting the relatively high water and gas content of the magmas. During an eruption, a column may

Fig. 1. Mount Ruapehu (2797 m; 9140 ft) in the North Island, New Zealand, erupting July 1996. This active andesite volcano is at the southern end of the Taupo Volcanic Zone continental margin arc system. Tephra (volcanic ash) is shown falling from the eruptive column to blanket the landscape to the north (left) of the volcano.

Chemical analyses (expressed in percent) of typical andesites from a continental margin arc (Mount Ruapehu, New Zealand) and an oceanic island arc (Raoul Island, Kermadec Islands, Southwest Pacific)

Component	Ruapehu	Raoul	Average continental crust
	Percent		
SiO_2	59.44	57.27	59.1
TiO_2	0.67	1.23	0.70
Al_2O_3	16.92	14.10	15.8
Fe_2O_3	6.37	13.31	6.6†
MnO	0.10	0.23	0.11
MgO	4.03	2.93	4.4
CaO	6.79	7.52	6.4
Na_2O	3.38	2.69	3.2
K_2O	1.70	0.62	1.9
P_2O_5	0.12	0.19	0.2
LOI*	0.27	−0.47	
Total	99.79	99.62	98.41
	Trace elements, ppm		
Sc	16.7	41	22
V	152	294	131
Cr	51	4	119
Ni	25	4	51
Cu	49	102	24
Zn	68	122	73
Ga	17	13	16
Rb	59	8.5	58
Sr	274	188	325
Y	19	37	20
Zr	130	75	123
Nb	4.54	0.82	12
Cs	5.13	0.35	2.6
Ba	339	202	390
La	13.54	4.4	18
Ce	29.24	13.0	42
Pr	3.8		5.0
Nd	14.92	11.7	20
Sm	3.49	3.84	3.9
Eu	1.04	1.35	1.2
Gd	3.87	4.61	3.6
Tb	0.57	0.93	0.56
Dy	3.29	5.41	3.5
Ho	0.69		0.76
Er	1.87	3.15	2.2
Tm	0.30		
Yb	2.03	3.95	2.0
Lu	0.31	0.62	0.33
Hf	4.26	2.37	3.7
Ta	0.81	0.066	1.1
Pb	14.53	2.7	12.6
Th	5.47	0.563	5.6
U	1.63	0.233	1.4

* Loss on ignition (≈volatile content).
† All Fe as Fe_2O_3.

rise tens of kilometers above the volcano, leading to stratospheric dispersion of tephra and volcanic gases and generation of pyroclastic flows (for example, Pinatubo). Pyroclastic flows are a particular feature of andesite-type volcanism and are among the most dangerous of volcanic hazards. Indeed, it was pyroclastic flows from Vesuvius which overwhelmed Pompeii in A.D. 79; and at Mount Pelee on the island of Martinique in 1902 the pyroclastic flow engulfed the town of St. Pierre, killing around 28,000 people. *See* BASALT; LAVA; PYROCLASTIC ROCKS.

Less commonly, rocks of andesite and basaltic-andesite composition are associated with sites of intraplate volcanism, such as Iceland, Galápagos, and Hawaii, and are unrelated to subduction or orogenic processes. These rocks have basaltic-andesite to andesite composition (that is, in the composition range of 53–63% SiO_2) and have been called icelandite and hawaiite. They have formed in response to fractional crystallization of basaltic magma.

It is generally recognized that most andesites cannot be direct partial melts of peridotitic mantle. As such, andesites are evolved magmas and the end product of a plethora of processes, including crystal

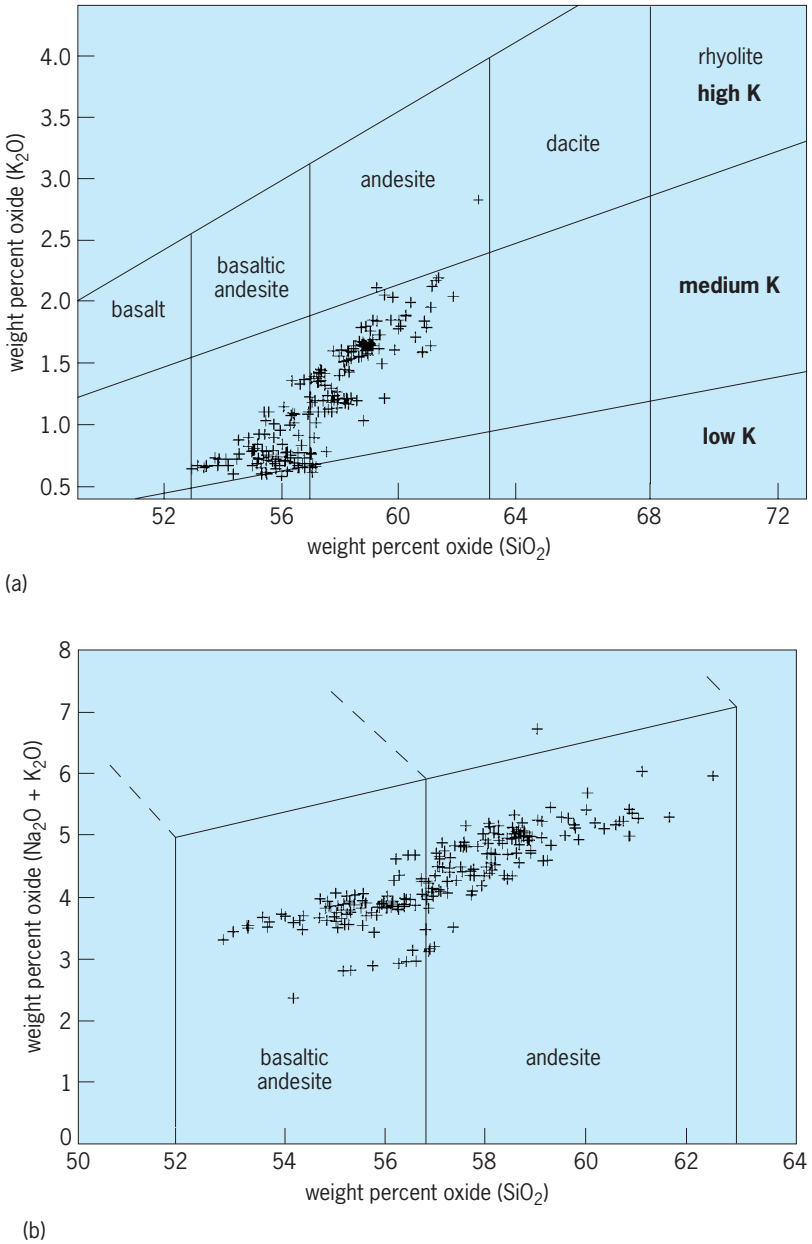

Fig. 2. Igneous rock and andesite classification. (*a*) Igneous rocks, in terms of SiO_2 versus K_2O. Note that a continuum exists between basalt, basaltic andesite, andesite, dacite, and rhyolite, and that it is possible to subdivide igneous rocks into low-, medium-, and high-K varieties. The data shown are andesites from Mount Ruapehu, which classify in this diagram as medium-K basaltic andesites and andesites. (*b*) Andesitic rocks, in terms of SiO_2 versus ($Na_2O + K_2O$), showing a part of the total alkali versus silica (TAS) diagram. (After *M. J. Le Bas et al., A chemical classification of volcanic rocks based on the total alkali-silica diagram, J. Petrol., 27:745–750, 1986*)

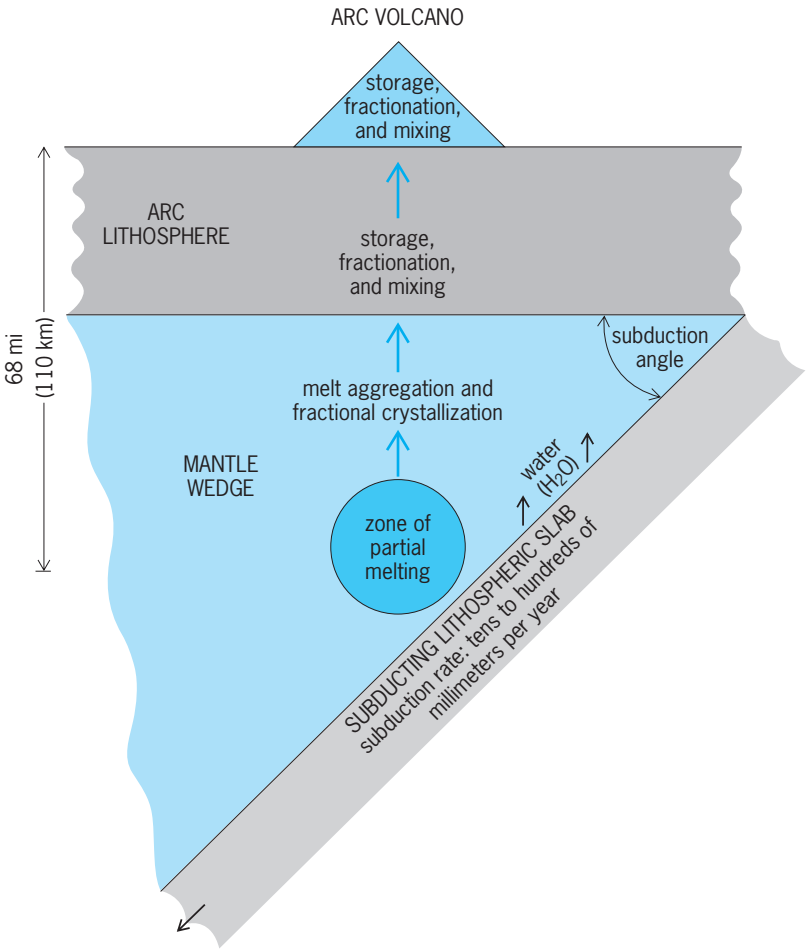

Fig. 3. Schematic cross section through a typical subduction system. The major components of a convergent plate boundary are the arc lithosphere, the subducting lithospheric slab, and the mantle wedge. Note that arc volcanoes overlie the subducting slab by around 110 km (68 mi), that the volume of the mantle wedge is determined by arc lithosphere thickness and slab dip, and that the mantle wedge is the principal site of magma production. After magmas form by partial melting, they can evolve further by fractional crystallization, mixing, and storage in reservoirs over a range of depths.

fractionation, magma mixing, mingling, assimilation, and storage, which have acted on primary basaltic magmas produced by partial melting in the mantle wedge. *See* MAGMA.

In more detail, the process of subduction recycles oceanic lithosphere back into the deep Earth; and as such, it complements the process of sea-floor spreading which generates new basalt crust and oceanic lithosphere at the mid-ocean ridges. The zone along which subduction takes place is demarcated at the Earth's surface by a deep oceanic trench and within the Earth by an increase of earthquake intensity along an inclined plane, the Wadati-Benioff Zone. This zone derives from the fact that cold, brittle material (the subducting oceanic lithosphere) is pushing into warmer upper mantle, and the earthquakes reflect the contrasting physical properties between the two mediums. The subduction process takes place at rates varying from a few tens to a few hundreds of millimeters per year and at angles of around 20° (shallow subduction) to nearly 90° (steep subduction). Typically, andesite volcanoes are located about 110 km (68 mi) above the surface of the subducting plate (**Fig. 3**). Between the surface of the subducting plate, which geologists refer to as the subducting slab, and the overlying plate is the mantle wedge, where basaltic melts form by the process of partial melting. Partial melting involves the generation of melt by dissolution of minerals, leaving a residual, more refractory, mineral assemblage. The melts aggregate and then move toward the surface by virtue of their lower density and viscosity. The melting occurs because the subducting slab, which may have spent many millions of years on the ocean floor as oceanic crust, contains minerals with seawater locked into their structure. When the minerals are subjected to higher pressure and temperature as subduction proceeds, chemical reactions in the minerals in the slab lead to loss of water and other volatile constituents which diffuse upward into the mantle wedge, together with an array of readily soluble elements such as potassium, rubidium, cesium, barium, and uranium. The mantle wedge becomes relatively enriched in these elements. The introduction of water is especially significant as it depresses the melting temperature and facilitates partial melting among constituent mineral grains. Importantly, the mantle wedge is recognized as the source of most arc-related magmas. On rare occasions, where subducting oceanic lithosphere is young, is still hot, and is subducting rapidly, the slab itself may experience partial melting, generating a unique suite of high-silica, low-potassium rocks of andesite composition, called adakites (after Adak Island in the Aleutian island arc chain). Still another unique suite of andesitic rocks, called boninites, which have high-magnesium and typically low-titanium contents, lacking plagioclase phenocrysts but with magnesium-rich, calcium-poor pyroxene phenocrysts, are associated with some suites of island arc rocks (for example, the Bonin island arc, which is the type area), where water-saturated melting of depleted mantle peridotite produced partial melts of high-silica, highmagnesium content. *See* SUBDUCTION ZONES.

The bulk composition of typical andesites is similar to that of continental crust (see table), contrasting with oceanic crust which is basaltic. Over geological time, subduction and the resulting andesite volcanism has proved a means of generating new continental crust, possibly at continental margin arc systems.

John Gamble

Bibliography. P. W. Francis, *Volcanoes: A Planetary Perspective*, Oxford University Press, 1993; J. A. Gamble et al., A fifty year perspective of magmatic evolution on Ruapehu Volcano, New Zealand: Verification of open system behaviour in an arc volcano, *Earth Planet. Sci. Lett.*, 170:301–314, 1999; J. B. Gill, *Orogenic Andesites and Plate Tectonics*, Springer-Verlag, Berlin, 1981; M. J. Le Bas et al., A chemical classification of volcanic rocks based on the total alkali-silica diagram, *J. Petrol.*, 27:745–750, 1986; A. R. McBirney, *Igneous Petrology*, Jones and Bartlett, Boston, 1993; R. L. Rudnick, Making continental crust, *Nature*, 378:571–578, 1995; R. S. Thorpe (ed.), *Andesites: Orogenic Andesites and Related Rocks*, Wiley, Chichester, 1982;

M. Wilson, *Igneous Petrogenesis: A Global Tectonic Approach*, Unwin Hyman, London, 1989.

Andreaeopsida

A class of the plant division Bryophyta, containing plants commonly called granite mosses. The class consists of one order and two families, the Andreaeobryaceae and the Andreaeaceae. The large tapered foot and short, massive seta of the Andreaeobryaceae show possible linkage to the Bryopsida, but the differences seem more fundamental and more significant than the similarities.

The small, brittle plants grow as perennials on rocks in montane regions in dark red, red-brown, or blackish tufts. The stems are erect, simple or forked, and grow from a single apical cell with three cutting faces; they consist of a homogeneous tissue of thick-walled cells. The rhizoids are multicellular and filamentous or platelike. The leaves are spirally arranged, of various shapes and with or without a midrib. The leaf cells are thick-walled and often papillose. Paraphyses are present in both male and female inflorescences.

The sex organs are superficial in origin. The antheridia are long-ellipsoid and stalked. The archegonia are flask-shaped. The terminal sporophytes consist of a small foot and an ellipsoid capsule elevated on an extension of the gametophytic axis (see **illus**.), the pseudopodium (or, in *Andreaeobryum*, they consist of a well-developed foot, a short seta and no pseudopodium). The capsule dehisces by 4–8 vertical slits. The slender columella is surrounded (and overarched) by sporogenous tissue of endothecial origin. The wall consists of solid tissue. Stomata, peristomes, and elaters are lacking. The spores begin division within the spore wall before dehiscence of the capsule. The protonema is a branched filament producing numerous buds; erect filaments may produce at their tips leaflike plates. The calyptra is small and mitrate or, in *Andreaeobryum*, larger and campanulate-mitrate. *See* ANTHOCEROTOPSIDA; BRYOPHYTA; BRYOPSIDA; HEPATICOPSIDA; SPHAGNOPSIDA.

Howard Crum

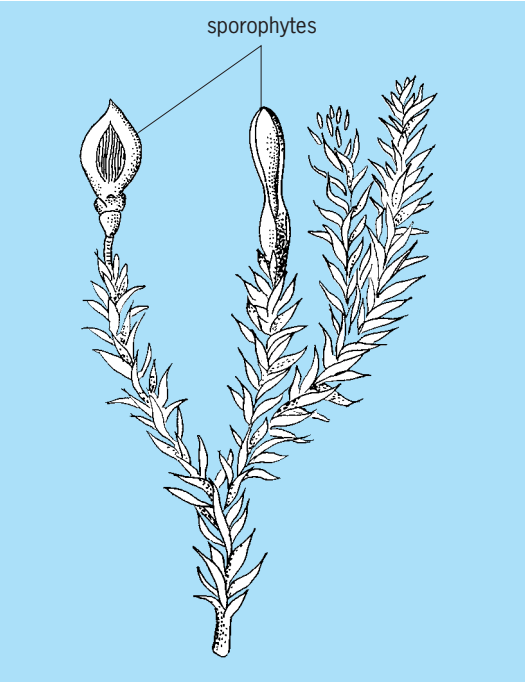

Granite moss (*Andreaea ruperstris*), showing gametophyte with sporophytes. (*After G. M. Smith, Cryptogamic Botany, vol. 2, 2d ed., McGraw-Hill, 1955*)

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; W. Schultze-Motel, Monographie der Laubmoosgattung *Andreaea*, I: Die costaten Arten, *Willdenowia*, 6:25–110, 1970; W. C. Steere and B. M. Murray, *Andreaeobryum macrosporum*, a new genus and species of Musci from northern Alaska and Canada, *Phytologia*, 33:407–410, 1976.

Androgens

A class of steroid hormones produced in the testes and adrenal cortex which act to regulate masculine secondary sexual characteristics. The major naturally occurring androgens, such as testosterone, are C-19 steroids (that is, they all contain 19 carbon atoms) with a hydroxyl or ketone group at the C-3 and C-17 positions (see testosterone in **Fig. 1**).

Reproductive functions. Androgens play several important roles related to male reproductive functions. These steroids are responsible for the development of the entire reproductive tract, including the epididymis, seminal vesicles, and vas deferens, in the developing male embryo as well as the external genitalia, including the penis. At the onset of puberty, androgen secretion increases dramatically; once testosterone levels reach a plateau, they remain fairly constant until a man reaches the age of 70–80, after which they slowly decline. During puberty, these androgenic hormones are essential for development of the masculine secondary sex characteristics, which include a deepening of the voice, growth of a beard and coarse body hair, growth of the penis, and increased development of muscles. They are also required for development of the male sexual drive, or libido; initiation and maintenance of spermatogenesis, or sperm production; and the growth and function of accessory sex glands, including the prostate gland. *See* PROSTATE GLAND; SPERMATOGENESIS.

Although the interstitial cells (Leydig cells) in the male testes secrete testosterone, other less potent androgens, including dehydroepiandrosterone (DHEA) and androstenedione, are secreted by the cortex of the adrenal glands in both males and females. In females these adrenal androgens modulate the physiological function of reproductive and sexual organs including the ovaries, vagina, clitoris, and mammary glands. They are also responsible for growth of hair in the armpits and appear to be required for sexual arousal of the female. Overproduction of these adrenal androgens in a female can cause

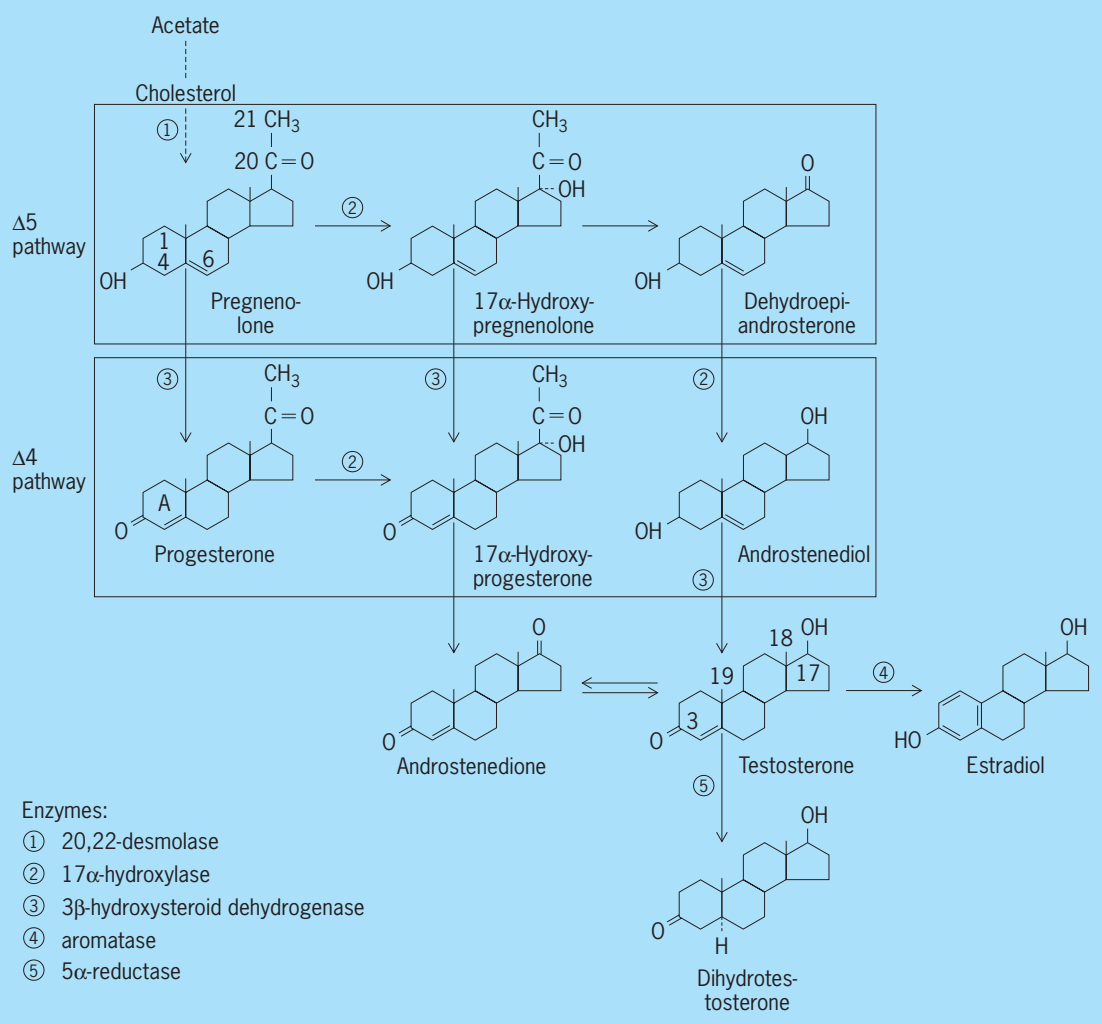

Fig. 1. Major pathways of testicular androgen and estrogen synthesis.

masculinization, which includes the growth of coarse body hair. Although the synthesis and secretion of adrenal androgens in both males and females is stimulated by adrenocorticotropic hormone, testosterone secretion by the Leydig cells of the male testes is specifically stimulated by luteinizing hormone. Both adrenocorticotropic hormone and luteinizing hormone are tropic hormones secreted by the anterior pituitary gland. *See* ADENOHYPOPHYSIS HORMONE; ADRENAL CORTEX; OVARY; REPRODUCTIVE SYSTEM; TESTIS.

Nonreproductive functions. Androgens also produce nonreproductive effects that include increased secretory activity of the sebaceous glands in a boy's skin at the time of puberty and fusion of the epiphyseal growth plate in the long bones, preventing an adult male from continuing to grow taller. DHEA and its sulfated ester, DHEA(S), have been reported to exert a number of potential health benefits that range from antiaging and anticancer effects to enhanced weight loss and stimulation of the immune system. However, the use of synthetic androgens, often referred to as anabolic steroids, to enhance muscle mass and athletic ability can have deleterious consequences. In a teenage boy, these undesirable side effects include not only acne and breast development but also a shutting down of bone growth. In adult males, overuse of these anabolic steroids can result in: infertility (decreased sperm production); liver damage; acceleration of the growth of prostate tumors; and the development of liver and brain tumors.

Biosynthesis. Like all steroid hormones, testosterone is synthesized from cholesterol, which in turn is either synthesized from acetate (Fig. 1) or derived from serum lipoproteins. The rate-limiting step for testosterone synthesis is the side-chain cleavage of cholesterol to form pregnenolone. The desmolase enzyme, located within the mitochondria of Leydig cells, catalyzes this key reaction. Once pregnenolone is generated, it is converted to testosterone via two different pathways. The Δ5 pathway is the major pathway in humans; in this pathway side-chain cleavage (carbons 20 and 21) and reduction of the 17-keto group occur before oxidation of the first ring (A ring) of the steroid nucleus. In the Δ4 pathway this sequence is required. The enzymes that catalyze these steroid modifications are identical for both pathways

and are located in the smooth endoplasmic reticulum of testosterone-secreting cells. Once the synthesized testosterone has been released into the bloodstream, it circulates throughout the body bound to proteins that are synthesized in the liver. Most of the secreted testosterone binds with high affinity to a serum transport protein known as sex steroid-binding globulin. Roughly 40% of the circulating testosterone is bound with low affinity to serum albumin. Although only 1–3% of the circulating testosterone is unbound, it is this free form of testosterone that is biologically active. *See* CHOLESTEROL; STEROID.

Further reduction of testosterone to generate the more potent androgen dihydrotestosterone (DHT) occurs in some androgen-responsive tissues, including cells in the prostate gland, that express the 5α-reductase enzyme (Fig. 1). This potent androgen binds to intracellular androgen receptors more tightly, or with higher affinity, than testosterone itself. Once either testosterone or dihydrotestosterone has bound to these receptor proteins located within target cells, these hormone-receptor complexes interact with other molecules (coactivators and corepressors) and bind directly to specific nuclear deoxyribonucleic acid (DNA) sequences. This nuclear binding subsequently regulates the expression or transcription of androgen-responsive genes. *See* GENE.

Conversion to estradiol. Testosterone can also be converted to estradiol via an enzyme that catalyzes the aromatization (introduction of double bonds) of the A ring of the steroid (Fig. 1). This reaction occurs primarily within the testicular Sertoli cells when they are stimulated by another anterior pituitary tropic hormone, follicle stimulating hormone (FSH). Similar amounts of estradiol can also be generated from testosterone within adipose tissue and the central nervous system. Although estradiol is a well-characterized female reproductive hormone, normal concentrations are also required for spermatogenesis, as well as regulation of bone density in males. Overproduction of estradiol in obese men can lead to feminization, including enlargement of the breasts. *See* ESTROGEN.

Catabolism. The metabolic breakdown, or catabolism, of testosterone takes place fairly rapidly in the liver. The major metabolites that are produced are 17-ketosteroids, generated from either androstenedione or dihydrotestosterone (**Fig. 2**). The major end product of this metabolic pathway is androsterone, which is excreted in the urine in the free form of conjugated with glucuronate or sulfate. DHEA, the predominant androgen secreted by the adrenal glands, is the major source of these excreted 17-ketosteroids. Other metabolites of testosterone that are found in small amounts in urine include

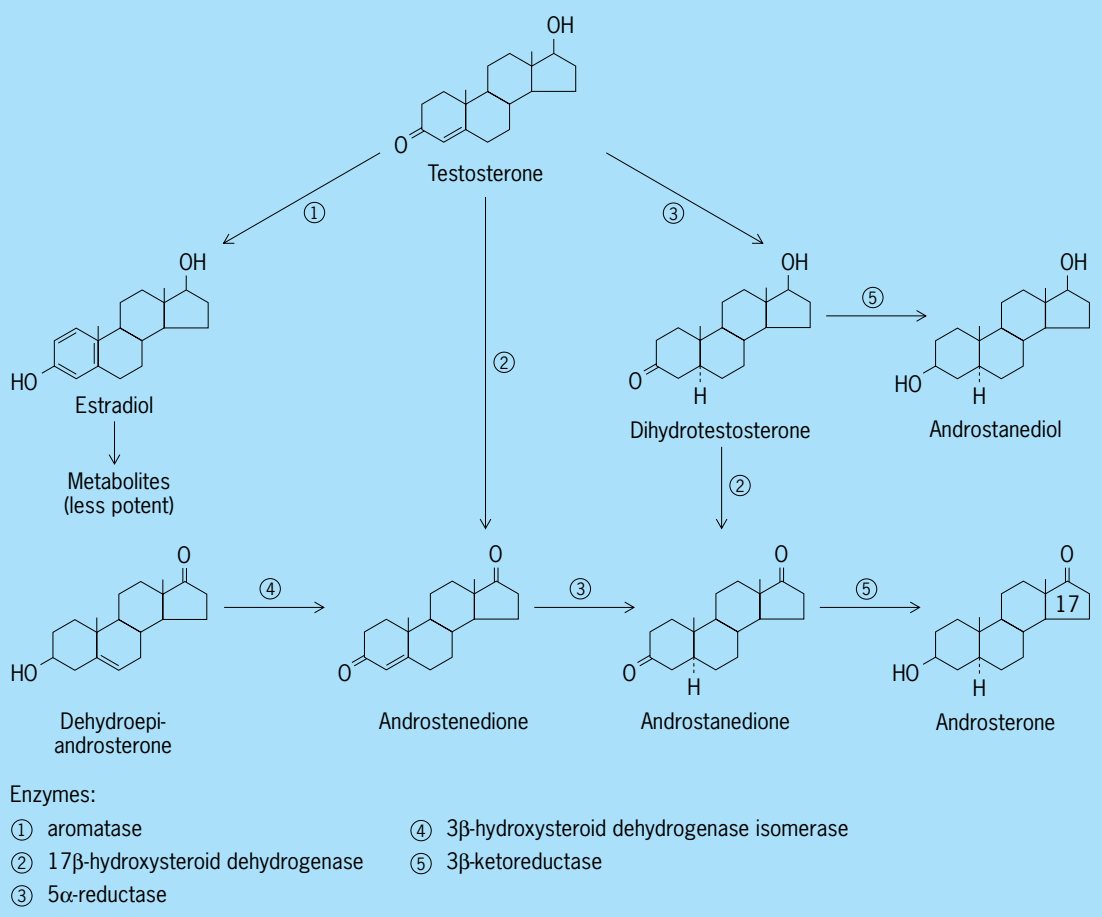

Fig. 2. Catabolism of androgens to generate active metabolites or inactive 17-ketosteroids.

androstanediol, which is formed by the reduction of the 3-keto group of dihydrotestosterone, and estrogen metabolites (estrone and estriol) which are formed after testosterone has been converted to estradiol. Thomas J. Schmidt

Bibliography. C. J. Bagatelli and W. J. Bremner, *Androgens in Health and Disease*, Humana Press, Totowa, NJ, 2003; R. M. Berne et al., *Physiology*, 5th ed., Mosby, St. Louis, 2003; C. Chang, *Androgens and Androgen Receptor: Mechanisms, Functions, and Clinical Applications*, Kluwer Academic Publishers, Boston, 2002.

Andromeda

A prominently located constellation in the northern sky (see **illustration**), named for the daughter of Cassiopeia in Greek mythology: When Cassiopeia bragged that her daughter Andromeda was more beautiful than Poseidon's daughters, the Nereids, Poseidon created Cetus, the sea monster. (In some versions of the myth, Cassiopeia boasted of her own beauty.) The situation required Andromeda's sacrifice. However, Perseus saved Andromeda by showing Cetus the head of Medusa, turning Cetus to stone. *See* CASSIOPEIA; PERSEUS.

The Great Galaxy in Andromeda (M31) is the nearest spiral galaxy to the Earth, and along with the galaxy M33 in Triangulum is the farthest object that can be seen with the unaided eye, since the glow that comes from its core can be seen even though it is 2.4 million light-years away. Only photographs show the galaxy's extent. *See* ANDROMEDA GALAXY.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION. Jay M. Pasachoff

Andromeda Galaxy

The giant spiral galaxy nearest to Earth. The Andromeda Galaxy is seen in the constellation of Andromeda, which is visible from Earth in the autumn skies. The galaxy is also known as M31 and NGC 224, based on its entries in two widely used catalogs. *See* ASTRONOMICAL CATALOG; CONSTELLATION; MESSIER CATALOG.

The Andromeda Galaxy can be seen without a telescope, one of a very few external galaxies that are sufficiently close and bright. Visually through a telescope, it appears as a large, faint, nearly featureless object. But photography, especially with modern electronic detectors, shows that it is an immense system of stars and other material, arranged in a nearly edge-on disk and large, spherical, faint halo (**Fig. 1**). It has been mapped in great detail at many wavelengths, including radio, infrared, visual, ultraviolet, and x-ray.

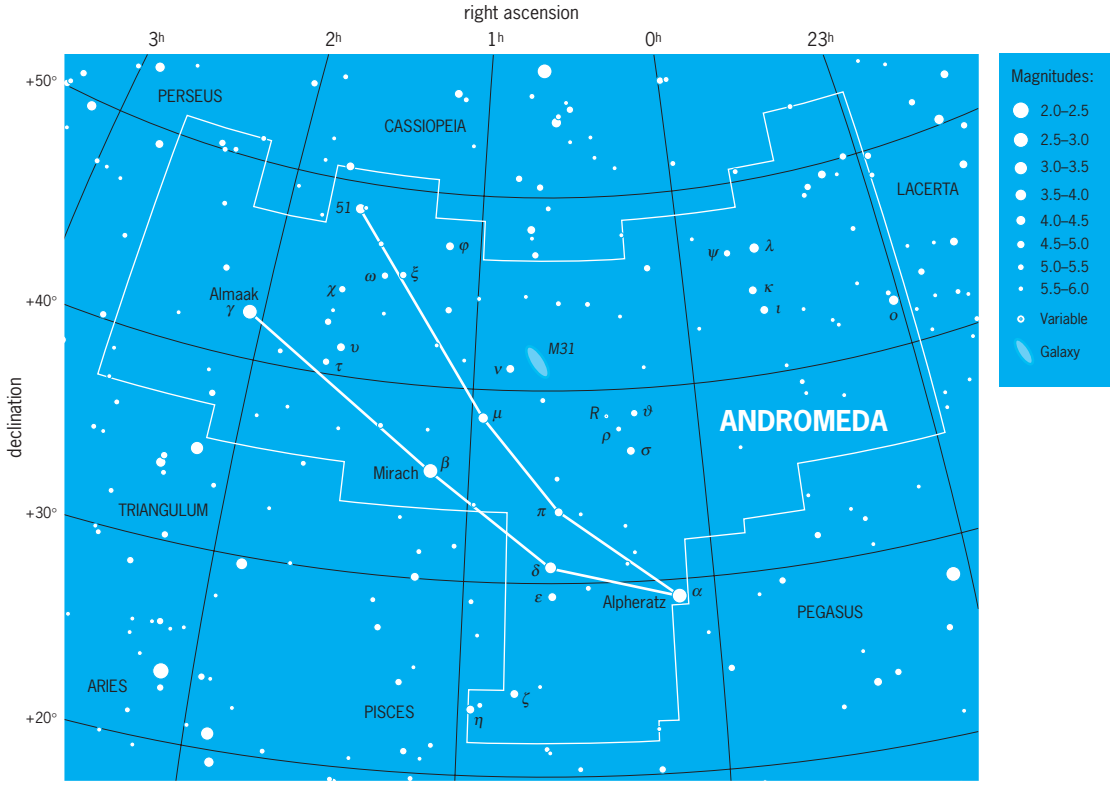

Modern boundaries of the constellation Andromeda, daughter of Cassiopeia. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

Fig. 1. Photographic image of the Andromeda Galaxy obtained with the 1.2-m Schmidt telescope of the Palomar Observatory. (*Courtesy of Paul Hodge*)

The Andromeda Galaxy is one of 36 galaxies that are clustered together in the Local Group. It and the Milky Way Galaxy are the largest members and are two of only three spiral galaxy members. Its distance can be measured in a number of ways, most accurately by the use of the Cepheid variable period-luminosity law. The distance is found to be 2,500,000 light-years (1.5×10^{19} mi or 2.4×10^{19} km). *See* CEPHEIDS; LOCAL GROUP.

Size. The main body of the Andromeda Galaxy is about 3° across in the sky, corresponding to about 135,000 light-years (8×10^{17} mi or 1.3×10^{18} km). However, the hydrogen gas in the galaxy's disk can be detected to almost twice that size, and the dark-matter halo, the extent of which is only roughly measured, is even larger. (Dark matter is material of an unknown nature that can be detected only by its gravitational influence on visible matter. It appears to be the major constituent of galaxies and galaxy clusters.)

Contents. There are a few hundred billion stars (a few times 10^{11}) in the Andromeda Galaxy, most of them apparently quite similar in their properties to the stars in the Milky Way Galaxy. The Andromeda Galaxy also is similar in containing hundreds of star clusters, both globular clusters (**Fig. 2**) and smaller, young clusters, many of which have been studied in detail by the *Hubble Space Telescope*. *See* HUBBLE SPACE TELESCOPE; STAR; STAR CLUSTERS.

Between the stars and clusters there is an interstellar medium of gas, mostly atomic hydrogen, detected by radio telescopes. The gas is very tenuous except in certain areas where it is dense enough to condense into new stars. Star-forming regions are made up of both gas and dust, the gas in these places being largely molecular hydrogen, as well as other molecules. The inner parts of the Andromeda Galaxy are relatively free of atomic hydrogen, but contain much cool molecular gas. Currently star formation is concentrated at intermediate distances from the center, where the spiral arms are most conspicuous. The total mass of gas is a few billion times the mass of the Sun. *See* INTERSTELLAR MATTER; RADIO ASTRONOMY.

There is also dust located throughout the disk of the galaxy, visible both at optical and at infrared wavelengths. The northwest side of Andromeda is heavily obscured by this dust, as seen in telescopic images. The total mass of dust in the galaxy is about 1/100 the mass of gas.

Early results from space-borne x-ray telescopes searching for x-ray sources in the Andromeda Galaxy turned up a few dozen sources, mostly located in the inner parts. Searches with the *Chandra X-ray Observatory* have pinpointed hundreds of x-ray sources, many of which correspond in position with globular clusters. Most of them are neutron stars like those found in Milky Way clusters. *See* CHANDRA X-RAY OBSERVATORY; NEUTRON STAR; X-RAY ASTRONOMY.

Structure. The nucleus of the Andromeda Galaxy is unusual. As discovered by the *Hubble Space Telescope*, it is a double nucleus, the two components being too close together to have been resolved by ground-based telescopes. The two revolve around each other, and one contains a black hole of several million solar masses. *See* BLACK HOLE.

Surrounding the nucleus is an ellipsoidal bulge, made up of old and intermediate-age stars.

Fig. 2. Globular star cluster G1 in the Andromeda Galaxy, imaged by the *Hubble Space Telescope*. (*NASA*)

Enveloping the whole disk and extending even farther out than the bulk of the disk stars is a nearly spherical halo, primarily containing thinly spaced old stars. However, unlike the halo of the Milky Way Galaxy, that of the Andromeda Galaxy is not mainly made up of ancient stars with low heavy-element abundances. The origin and evolution of the halo differs in some unexplained way from those of the Milky Way Galaxy. Paul Hodge

Bibliography. P. Hodge, *The Andromeda Galaxy*, Kluwer Academic, 1992; P. Hodge, *Atlas of the Andromeda Galaxy*, University of Washington Press, 1981; T. R. Lauer et al., Planetary camera observations of the double nucleus of M31, *Astron. J.*, 106: 1436–1447, 1993; B. F. Williams, Recent star formation history of the M31 disk, *Astron. J.*, 126:1312–1325, 2003.

Anechoic chamber

A room whose boundaries absorb effectively all the waves incident on them, thereby providing free-field conditions. A free field is a field whose boundaries exert negligible effect on the incident waves. In practice, it is a field in which the effects of the boundaries are negligible over the frequency range of interest. Acoustic chambers and radio-frequency and microwave chambers will be discussed.

Acoustic chambers. An acoustic anechoic chamber is a room in which essentially an acoustic free field exists. It is sometimes referred to as a free-field or dead room. The word anechoic is derived from the Greek, meaning "without echo." However, a room that is free from echoes (in the usual sense of the word) is not necessarily an anechoic room: in addition, there must be no significant reflections of sound waves from the boundaries of the room.

Free-field conditions can be approximated when the absorption by the boundaries of the room approaches 100%. To reduce sound reflected by the boundaries to a minimum, the absorption coefficient must be very high and the surface areas of the boundaries should be large.

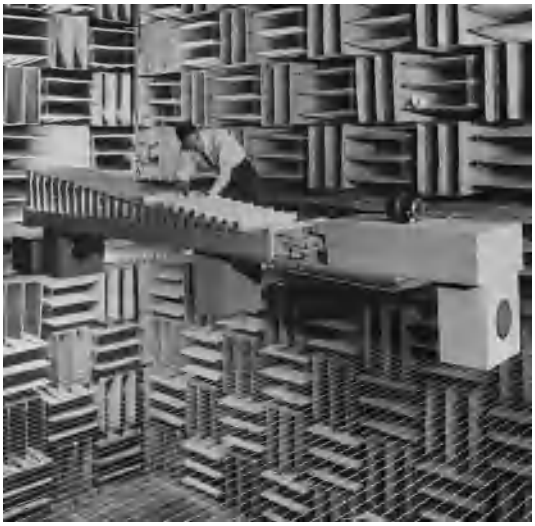

Interior of an acoustic anechoic chamber. The room's special construction eliminates 99.9% of reflected sound.

The sound absorptive material usually installed in such rooms consists of glass fibers held together with a suitable binder. In order to achieve large surface area, a wall construction (see **illus.**) is used that includes wedges of sound absorptive material, the base of which is usually between 8 × 8 in. (20 × 20 cm) and 8 × 24 in. (20 × 60 cm), and the length of which is usually 3 to 5 ft (0.9 to 1.5 m). These wedges resemble stalagmites and stalactites and absorb about 99% of incident sound energy over most of the audio-frequency ranges. A horizontal net of thin steel cables just above floor wedges permits walking in the room. *See* REFLECTION OF SOUND; SOUND ABSORPTION. Cyril M. Harris

Radio-frequency and microwave chambers. The radio-frequency or microwave anechoic chamber is a shielded (screened) enclosure in which internal reflections of electromagnetic waves are reduced to an absolute minimum, thus providing a zone that is free from radio noise, broadcast transmissions, and other extraneous signals, and that simulates free-space conditions. Its main use is to provide a controlled and well-defined environment for the testing of electronic equipment.

Shielded or screened enclosures may vary in size from small metallic boxes that are intended for the protection of circuit elements to large rooms capable of housing vehicles or aircraft. The purpose of such enclosures is to isolate sources of electromagnetic energy from electronic, radio, or other sensitive equipment which might suffer degradation of performance as a result of interaction with transmitters, radio-frequency heating equipment, and so forth. The provision of a radio-frequency quiet zone is an essential feature of many applications in electromagnetic tests and investigations, for example, the determination of antenna characteristics and the measurement of low-level emissions during electromagnetic compatibility assessments. *See* ELECTRICAL SHIELDING; ELECTROMAGNETIC COMPATIBILITY; MICROWAVE FREE-FIELD STANDARDS.

The mechanism by which shielded enclosures function is a combination of absorption as the electromagnetic wave progresses through the material of the shield and reflection of the wave at the interface. The principle of reciprocity applies, and reflection occurs at the internal as well as at the external interface. In most practical shielded enclosures any radio-frequency signals generated internally are reflected and the bulk of the energy remains within the enclosure. At certain frequencies where the dimensions of the enclosure are comparable with half a wavelength, reinforcement of reflected waves take place and the enclosure becomes a resonant cavity. The internal conditions are thus somewhat different from free space, and it is the attempt to reproduce free-space or open-field conditions within a shielded enclosure which has led to the development of resistive material to line the enclosure walls to absorb rather than reflect the

radio-frequency energy and hence create an anechoic chamber. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; CAVITY RESONATOR; RECIPROCITY PRINCIPLE; REFLECTION OF ELECTROMAGNETIC RADIATION.

Theoretical investigations indicate that reflection-free conditions can be established at discrete frequencies by provision of sheets of resistive material that have a surface impedance of 120π (377) ohms per square, that is, the free-space or plane-wave impedance, and are spaced at one-quarter wavelength from the reflecting surfaces. Extension of this principle to achieve broadband coverage and hence an anechoic enclosure capable of use over a wide frequency range led to the development of various types of material having the appropriate resistive, lossy characteristics. Such material must have good basic dielectric properties, with a relative permittivity near unity to minimize reflection at the interface, and to incorporate increasing resistive loss through the material. In practice, foam plastics are ideal for the purpose, and these can be readily loaded with carbon, ferrite, or other conducting particles on a graded or layer basis to provide the increasing loss. Furthermore, the use of cones or pyramids of the material assists considerably in the transition from minimal reflection at the apex to full absorption at the base where the resistive loading is at a maximum. The lining of the walls and ceiling of a shielded enclosure with such materials provides for absorption of any internal signals and therefore anechoic or minimal reflection conditions simulating propagation in free space. *See* DIELECTRIC MATERIALS; ELECTROMAGNETIC RADIATION; PERMITTIVITY.

In the microwave band, that is, at frequencies greater than 1 GHz, the wavelength is less than 1 ft (30 cm), and hence a lining thickness of only 1 or 2 in. (2.5 or 5 cm) provides for a reflectivity better than −40 dB or less than 1%. Enclosures that are used for test and measurement purposes have walls and ceilings with conical or pyramidal linings up to 7 ft (2 m) thick and are capable of simulating free-space conditions with reflections less than 10% over a wide frequency range extending down to 50 MHz. Special precautions are taken with the installation of lossy material on the floor, because the latter generally needs to be mechanically strong in addition to having desirable radio-frequency characteristics. *See* RADAR-ABSORBING MATERIALS. G. A. Jackson

Bibliography. P. A. Chatterton and M. A. Houlden, *EMC: Electromagnetic Theory in Practical Design*, 1992; C. M. Harris, *Handbook of Acoustical Measurements and Noise Control*, 3d ed., 1991.

Anemia

A reduction in the total quantity of hemoglobin or of red blood cells (erythrocytes) in the circulation. Because it generally is impractical to measure the total quantity, measures of concentration are used instead. Hemoglobin is contained in red blood cells, which are suspended in plasma, the liquid component of blood. Therefore, concentration is affected not only by quantities of hemoglobin and red blood cells but also by plasma volume. Thus, the apparent anemia found in many women in the third trimester of pregnancy is not really anemia at all: the red cell mass is actually increased, but the plasma volume is expanded even more. In other words, hemodilution is present. Conversely, in dehydration and other circumstances of hemoconcentration, the plasma volume is reduced, thereby tending to mask anemia. *See* BLOOD; HEMOGLOBIN.

The three measures of concentration most often employed are the hemoglobin, the red blood cell count, and the volume of packed red cells. Hemoglobin is determined photocolorimetrically. The red blood cell count is determined by automated optical-electronic devices; until recently it was determined by microscopic observation of a counting chamber—a less accurate method. The volume of packed red cells, often called the hematocrit after the instrument originally used in its measurement, is a measure of the proportion of volume that red cells occupy in a specimen of whole blood, and is determined by centrifugal methods, or by computation from light scatter or displacement volume of individual cells during automated enumeration procedures. When total red cell mass must be known, it is measured by removing some of the patient's red cells and tagging them with the radioisotope chromium-51, reinjecting them, and then computing the red cell mass from the resulting dilution of the tagged red cells by the unlabeled red cells in the individual's bloodstream. For diagnostic and therapeutic purposes, anemia is categorized according to the average volume of the individual red cells (mean corpuscular volume) and the concentration of hemoglobin within the red cells (mean corpuscular hemoglobin concentration).

General considerations. In a group of healthy individuals, the values for hemoglobin, red cell count, and hematocrit approximate a "normal" distribution. Values that are less than 2.5 standard deviations below the mean are indicative of anemia if other clinical factors do not indicate a condition of hemodilution. The mean values are greater for adult males than for adult females, and greater for adults than children. Lower atmospheric oxygen tension at higher altitudes results in higher mean values for hemoglobin, red cells, and hematocrit in healthy individuals living under these conditions; hence, anemia would be defined at a higher value.

Red blood cells have a finite life-span, averaging 120 days in the circulation in healthy persons, but the span may be shortened in certain diseases. Every day, about 45 billion aged red cells are removed from each liter of blood, or a total of 2×10^{11} red cells per day in the adult of average size. In the state of health, the rate of production of new red blood cells (erythropoiesis) equals the rate of removal of senescent red blood cells. Anemia occurs if the rate of erythropoiesis is reduced below normal. It also occurs if hemorrhage or destruction (hemolysis) of red blood cells within the body increases the rate of loss of

erythrocytes and the rate of erythropoiesis does not increase enough to compensate.

Mechanisms. Red blood cell production may be affected by several different mechanisms. Erythropoietin, a growth factor produced by healthy kidneys, stimulates the bone marrow to produce more erythroblasts and accelerates their maturation. An intact bone marrow can respond by increasing production of red blood cells by about sevenfold after a few days or weeks. The proliferative response of the bone marrow to anemia may be defective or absent if the production of erythropoietin is diminished or absent, as occurs in chronic renal disease and some endocrine disorders. Erythropoietin produced synthetically has become available on prescription to overcome that deficiency and thereby resolve that specific type of anemia. *See* KIDNEY.

Ineffective erythropoiesis results when, in the intact, stimulated bone marrow, red blood cell precursors either fail to mature, or die in the bone marrow prior to their delivery to the circulation as erythrocytes.

Aplastic anemia occurs when the bone marrow stem cells, that give rise to precursors of erythroblasts, are markedly diminished in number or may respond inadequately to erythropoietin. This condition occurs when the bone marrow is adversely affected by certain chemicals or autoantibodies; is injured by irradiation; atrophies, or is replaced by fat; is replaced by fibrous (scar) tissue; or is infiltrated by cancer cells. Drugs and other chemicals that have been associated with aplastic anemia include those whose effects are dose-related and are an extension of their intended purpose, for example, some of the drugs used in cancer chemotherapy. Other drugs, notably chloramphenicol and benzol, seem to operate in an idiosyncratic fashion; many other agents have been implicated, though each in fewer cases.

Defective synthesis of deoxyribonucleic acid (DNA) and abnormal nuclear maturation result from malabsorption of vitamin B_{12} (as in pernicious anemia) or dietary deficiency of folic acid or its malabsorption (sprue). Methotrexate, fluorouracil, hydroxyurea, mercaptopurine, and other drugs used in chemotherapy affect purine and pyrimidine metabolism and can produce the same lesion. All of these anemias are generally characterized by large red blood cells, which are called macrocytes, and a specific morphological abnormality of the nuclear chromatin of erythroblasts which characterizes them as megaloblasts, and the condition as megaloblastic anemia. *See* VITAMIN B_{12}.

Defective synthesis of hemoglobin impairs cytoplasmic maturation. The majority of cases are due to deficiency of body stores of iron and to abnormal release of iron from reticuloendothelial stores. The former occurs in iron-deficiency anemia, and the latter in the anemia of chronic inflammatory diseases. Qualitative abnormalities of globin synthesis occur in the hereditary hemoglobinopathies (as in sickle-cell anemia, hemoglobin-C disease), whereas reductions in the rate of synthesis of either the α- or β-globin chains, in the absence of qualitative abnormalities, occurs in the thalassemia syndromes. The hemoglobinopathies and thalassemia may be inherited independently of one another or simultaneously. Heterozygotes have little if any anemia, whereas homozygotes and double heterozygotes have moderate or marked anemia. Rarely, disorders affect the porphyrin moiety of the heme prosthetic group in globin (as in lead intoxication and hereditary erythropoietic porphyria).

Acute blood loss reduces the total blood volume and produces symptoms of weakness, dizziness, thirst, faintness, and shock, in that order, according to increasing magnitude of blood loss. The anemia which results is not detectable by measures of concentration until hemodilution occurs over subsequent days, or more rapidly if replacement fluids are given intravenously. The proliferative response of the healthy marrow will correct the anemia in 2–6 weeks, depending upon the size of the deficit and provided there are sufficient body stores of iron, folic acid, and vitamin B_{12}, required for hemoglobin synthesis. Chronic blood loss results in iron-deficiency anemia.

Hemolysis, the accelerated destruction of red blood cells, also induces a proliferative response from the marrow. However, it differs from hemorrhage because red blood cells are lost without plasma, and it thus diminishes the measures of concentration at the outset. Disorders external to the red blood cell (such as autoantibodies, chemicals including drugs, splenomegaly, irradiation, and thermal injury) are almost always acquired, and may destroy otherwise healthy cells of the host, as well as transfused, normal red blood cells. Even trauma to the red blood cells in the circulation (such as from artificial heart valves, abnormal capillaries, sometimes marching or running on hard surfaces) has caused hemolytic anemia. Disorders intrinsic to the red blood cell are almost always hereditary, and may affect the membrane, or envelope itself (as in hereditary spherocytosis and elliptocytosis), or the glycolytic enzymes of the cytosol (as in hereditary deficiency of pyruvate kinase or glucose-6-phosphate dehydrogenase), or be the consequence of anomalies of globin synthesis. When an external factor interacts with the intrinsic defect (such as the spleen: hereditary spherocytosis; a redox drug or chemical: abnormality of glucose-6-phosphate dehydrogenase), the anemia is abated when the external factor is removed, even though the hereditary factor remains. In hemolytic anemia the iron is recovered and used again for hemoglobin synthesis, except when there is intravascular hemolysis, in which case it may be lost in the urine.

Diagnosis and treatment. Pallor, weakness, and fatigue are common to all anemias. They may not be noticed until anemia is advanced, if it is of gradual onset and there has been time for cardiovascular and biochemical adaptation. Faint jaundice in the sclerae is a feature of hemolytic anemia, whereas in anemia due to lack of vitamin B_{12}, glossitis and neuropathy occur and may be noted. The morphological study of the blood is often the most simple

guide. Macrocytic anemias are often found to be megaloblastic and due to deficiency of vitamin B_{12} or folic acid. The latter contention can be proved by measuring the quantity of those substances in the plasma. Administration of one of those vitamins will only cure individuals in whom its specific deficiency is established. Microcytic anemias are often also hypochromic. Hypochromia is a reduction in the concentration of hemoglobin within individual red blood cells. It is detected by microscopic examination of the thin blood film after Wright's stain, and by finding the ratio of hemoglobin concentration/hematocrit to be less than 0.31.

Microcytic-hypochromic anemias are most often due to iron deficiency. The administration of iron will resolve iron-deficiency anemia. Since the body normally conserves iron and has no excretory route for it, iron deficiency implies a pathological blood loss which must be explained. The administration of iron or iron-containing over-the-counter medications without investigation of the underlying cause may result in the failure to detect an asymptomatic neoplasm of the gastrointentinal tract which is causing occult blood loss until it has progressed to the incurable state. In the absence of proved deficiency, the administration of iron, vitamin B_{12}, and folic acid are not of value to the anemic individual. Further, the administration of iron, especially by injection, may be harmful. Prednisone and other adrenal corticosteroids are helpful in hemolytic anemias associated with autoantibodies.

Hereditary disorders are generally not amenable to therapy, except for those hemolytic diseases which may benefit after splenectomy. In all other cases, the treatment of the anemia is achieved by treating the underlying disease, such as hypothyroidism, rheumatoid arthritis, or leukemia. Blood transfusions are reserved for acute blood loss when symptoms of hypovolemia and shock are present, or in chronic anemia if there are signs of inadequate cardiovascular or pulmonary compensation and an underlying cause cannot be found or treated. Since blood transfusions entail a risk of immune reactions, hepatitis, acquired immune deficiency syndrome, and iron overload, they should be avoided when other alternatives for treatment are possible. *See* CLINICAL PATHOLOGY; HEMATOLOGIC DISORDERS. Arthur Haut

Bibliography. G. R. Lee et al., *Clinical Hematology*, 8th ed., 1981; W. J. Williams et al. (eds.), *Hematology*, 3d ed., 1983.

Anesthesia

Loss of sensation with or without loss of consciousness. There are several ways of producing anesthesia, with the choice dependent on the type of surgery and the medical condition and preference of the patient. Each person responds differently to a given anesthetic, and anesthetic techniques and drugs often have marked effects on bodily functions, especially those of the cardiovascular and respiratory systems. Therefore, these systems are monitored closely during anesthetic administration, with measurements such as heart sounds, blood pressure, heart action (electrocardiogram), temperature, and oxygenation taken using a variety of sophisticated devices.

General anesthesia. A state of complete insensitivity or unconsciousness is produced when anesthetic gases are inhaled; adjuvant drugs are often given intravenously. Anesthetics commonly used by inhalation are halogenated-hydrocarbon liquids (isoflurane, enflurane, and halothane), which are vaporized and administered in oxygen, usually with nitrous oxide. Adjuvant drugs include thiopental (to produce rapid loss of consciousness for induction of anesthesia), narcotic opioids (for supplemental analgesia, that is, insensitivity to pain), muscle relaxants, and sedatives. Although the mechanism of general anesthesia is unknown, the anesthetics act on the upper reticular formation of neurons in the thalamus and midbrain (neuronal structures necessary for activating the cerebral cortex and maintaining an active, attentive state).

Regional anesthesia. Analgesia, without loss of consciousness, results from injecting a solution of local anesthetic drug either into the cerebrospinal fluid surrounding the spinal cord (spinal anesthesia) or into the epidural space surrounding the cerebrospinal fluid (epidural anesthesia), usually in the lower back (lumbar) area. Epidural injection is also performed at the base of the spine (caudal anesthesia). The duration and extent of analgesia depend on the nature and amount of local anesthetic injected, among other factors. The local anesthetic acts by blocking the conduction of nerve impulses. Narcotic opioids are injected postoperatively into either the epidural space or cerebrospinal fluid for pain relief.

Local anesthesia. Analgesia can be localized to a small area, for example, the forearm, by injecting a local anesthetic solution around nerves supplying the area (the bracheal plexus in the upper arm and chest supplies the forearm). Large peripheral nerves may also be blocked individually by using this method. The duration of analgesia is determined by the choice and amount of local anesthetic.

Injection of a dilute solution of local anesthetic into the skin and superficial tissues provides very localized analgesia for minor surgery.

Acupuncture. Acupuncture is an ancient procedure, once used only in China but now practiced in the United States and elsewhere. It involves inserting needles into specific points around the body, as determined from historical charts, and manipulation of the needles; sometimes electric current is applied. Weak analgesia results through alteration of pain perception. Although sometimes helpful for chronic pain, acupuncture generally has not been found satisfactory for surgical anesthesia. *See* CENTRAL NERVOUS SYSTEM; PAIN. Fredrick K. Orkin

Bibliography. M. J. Cousins and P. O. Bridenbaugh (eds.), *Neural Blockade in Clinical Anesthesia and Management of Pain*, 3d ed., 1998; R. D. Dripps, J. E. Eckenhoff, and L. D. Vandam (eds.), *Introduction to Anesthesia: The Principles of Safe Practice*,

7th ed., 1988; R. D. Miller (ed.), *Anesthesia*, 4th ed., 1994; R. K. Stoelting, *Pharmacology and Physiology in Anesthetic Practice*, 3d ed., 1998.

Aneurysm

A localized, abnormal arterial dilation usually caused by a weakening of the vessel wall. Aneurysms are commonly found in the abdominal aorta, intracranial arteries, and thoracic aorta; however, they may also involve the femoral, popliteal, splenic, carotid, or renal arteries. Aneurysms vary in size from less than 1 in. (2–3 cm) to more than 4 in. (10 cm). They are usually classified as true or false: true aneurysms involve all layers of the artery (inner endothelium or intima, middle muscular layer or media, and outer connective tissue or adventitia); false aneurysms do not involve all layers. With the exception of intracranial aneurysms, atherosclerotic vessel disease is generally the cause of aneurysms; other causes include syphilis, trauma, cystic medial necrosis, bacterial infections, and arteritis.

Complications. The major clinical sign of all aneurysms (except intracranial) is a large pulsatile mass; other manifestations depend on the location of the aneurysm. Aneurysms of the thoracic aorta can produce symptoms of compression or erosion as well as dissection. In dissection, the blood creates a channel between the intima and media of the aorta, which may result in blockage of major branches of the aorta, a condition that is often fatal. Debris in aneurysms may also embolize; that is, it may break free and block distal arteries. The major complication of an aneurysm is rupture, the possibility of which is directly related to its size. The mortality rate for a ruptured abdominal aortic aneurysm is 60–70%.

Intracranial aneurysms, often called berry aneurysms, are congenital weaknesses in the intracranial arteries. These aneurysms are saclike outpocketings of the vessel and vary in size from 0.2 in. (0.4 cm) to greater than 0.6 in. (1.5 cm). The major manifestations are usually related to bleeding and can vary from a severe headache, to neurologic impairment, to death. In many instances, a small initial hemorrhage (a "herald bleed") is often followed by a second bleeding episode, which is much more severe and often fatal. *See* HEMORRHAGE.

Diagnosis and treatment. The best method of diagnosis is ultrasound or a computer x-ray analysis (computed tomography scan). The recommended treatment for all aneurysms is surgery. In patients whose aneurysms are due to atherosclerosis, associated diseases can increase the risk of surgery. Abdominal and thoracic aneurysms are generally treated with resection of the aneurysm and replacement with a graft made of synthetic material (see **illus.**). Blood flow must be established to the branches of the aorta. Intracranial aneurysms are also treated surgically; however, surgery involves excluding the aneurysm from the cerebral circulation by placing a metallic clip across the neck of the aneurysm sac. A ruptured aneurysm or dissecting aneurysm is a true emergency and requires immediate surgical intervention. *See* ARTERIOSCLEROSIS; CARDIOVASCULAR SYSTEM; CIRCULATION DISORDERS; HYPERTENSION; MEDICAL IMAGING.

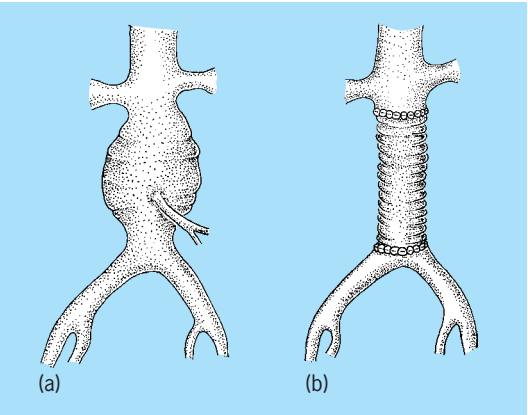

Surgical repair of an abdominal aortic aneurysm. (*a*) Before resection. (*b*) Tubular-type graft anastomosis. *(After M. D. Kerstein, P. V. Moulder, and W. R. Webb, eds., Aneurysms, Williams and Wilkins, 1983*)

Donald L. Akers; Morris D. Kerstein

Bibliography. C. Arbost, M. Allary, and N. Economou, Resection of an aneurysm of the abdominal aorta: Reestablishment of the continuity by a preserved human arterial graft which results after five months, *Arch. Surg.,* 64:405, 1952; M. D. Kerstein, P. V. Moulder, and W. R. Webb (eds.), *Aneurysms*, 1983; S. I. Schwartz and G. T. Shires (eds.), *Principles of Surgery*, 7th ed., 1998.

Angiogenesis

The origin and development of blood vessels. Blood vessels are composed of two basic cell types, vascular endothelial cells and peri-endothelial cells (including vascular smooth muscle cells and elongated contractile cells called pericytes, both of which support the underlying endothelial cells). The inner epithelial lining of all blood vessels, adjacent to the lumen, is a single layer of endothelial cells (**Fig. 1**). In larger blood vessels, such as arteries and veins, the inner endothelial lining, called the tunica intima, is surrounded by a medial layer, the tunica media, composed of multiple layers of vascular smooth muscle cells embedded in elastin-rich extracellular matrix. The tunica media layer is surrounded by an extracellular matrix-rich layer called the tunica adventitia. In contrast, capillary walls consist of only a single layer of endothelial cells, sometimes surrounded by pericytes.

Arteries and veins are the two fundamental types of blood vessels, carrying blood away from and toward the heart, respectively. Although both have the basic structure noted above, higher-pressure arteries generally have thicker medial smooth muscle–containing layers, whereas larger veins have thinner, more elastic walls and valves to prevent backflow. Recent studies of the process of blood vessel formation, or angiogenesis, have shown that the endothelial cells contributing to arteries and veins have

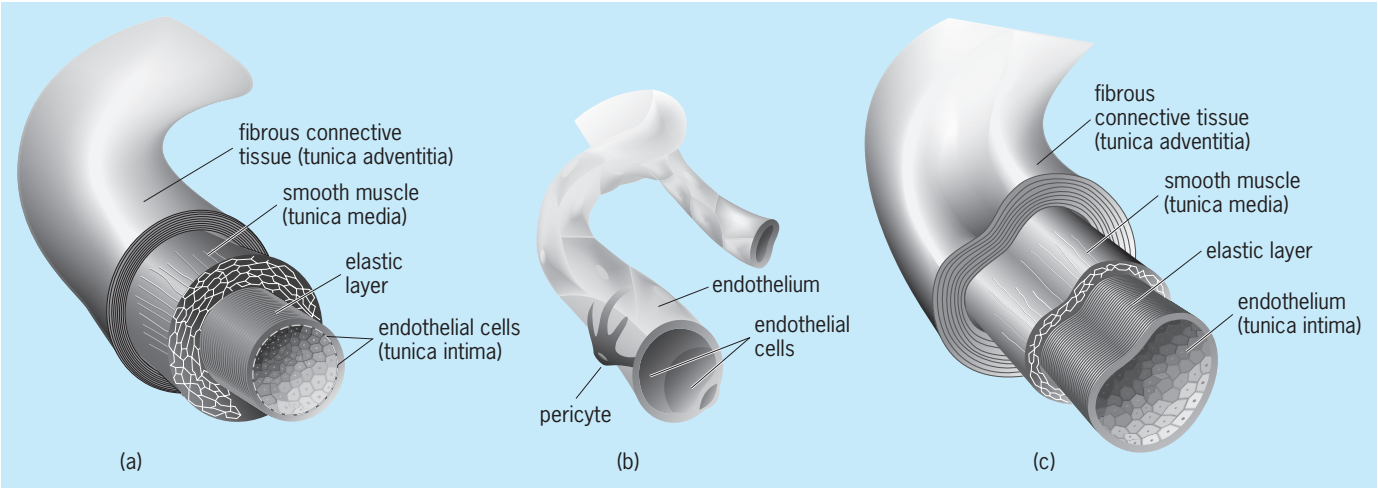

Fig. 1. Structure of blood vessels. (*a*) Arteries and (*c*) veins have the same tissue layers. (*b*) Capillaries are composed of only a single layer of endothelial cells, sometimes surrounded by elongated contractile cells known as pericytes (not to scale). (*Adapted from P. H. Raven and G. B. Johnson, Biology, 6th ed., McGraw-Hill, New York, 2002*)

distinct molecular and functional identities that are established early during development.

Blood vessel formation. Blood vessels are found only within the chordates, but within the vertebrates the anatomical architecture of the major blood vessels of the circulatory system is surprisingly well conserved, as highlighted by recent detailed characterization of the anatomy of the developing zebrafish vasculature. The first major vessels to emerge during embryonic development form by coalescence of individual mesodermal endothelial progenitor cells, or angioblasts, into cords of attached cells which then form open vascular cells, or lumenize. This process, called vasculogenesis, was thought to be restricted to early embryonic development; however, recent evidence has shown that vasculogenesis also occurs later in development and even postnatally. Most later developmental and postnatal blood vessel formation, however, occurs by sprouting and elongation of new vessels from preexisting vessels or remodeling of preexisting vessels, collectively known as angiogenesis.

Vessels generally form first as unlined endothelial tubes but generally become rapidly associated with supporting pericyte or vascular smooth muscle cells. A variety of studies have shown that the acquisition of these supporting cells is critical for proper morphogenesis, stability, and survival of nascent blood vessels. Although most larger blood vessels form during development and are relatively stable thereafter, significant growth and remodeling of blood vessels continues throughout adult life, for example during uterine cycling or in conjunction with gain or loss of fat or muscle mass.

Postnatal vessel growth also occurs in pathologic contexts, notably cancer. Tumors recruit new blood vessels to provide themselves with access to the host circulation and obtain oxygen and nutrients. Acquiring a vascular blood supply is essential for the progression and survival of a tumor. New tumor vessels form primarily through endothelial proliferation and growth from preexisting, adjacent host blood vessels. These vessels are usually morphologically abnormal, poorly structured, and deficient in vascular smooth muscle cells. Some tumor vascular spaces appear to lack even an endothelial lining and are surrounded only by tumor-derived cells.

Molecular regulators. Recently, interest in factors that might promote or inhibit blood vessel growth and the application of molecular methods to vascular biology and the study of vascular development have led to the discovery of many new genes encoding proteins involved in blood vessel growth and assembly. The vascular endothelial growth factor (VEGF) family of ligands and their tyrosine kinase vascular endothelial growth factor receptors (VEGFRs) expressed on the surface of endothelial cells play critical roles in the differentiation of blood vessels. VEGF-A plays a particularly central role—it is essential for the survival, proliferation, migration, and arterial differentiation of endothelial cells. Targeted disruption or "knockout" of just one of the two copies of the gene encoding this protein in mice causes (heterozygous) embryos to die early in development with a dramatic reduction in endothelial cell number. VEGFR2 is a key endothelial receptor transducing VEGF-A signals, and loss of the gene encoding this protein is also lethal (although only in homozygotes with both functional copies of the gene knocked out).

Other related VEGF ligands and VEGFR receptors have also been uncovered, and these also play important roles in modulating the development of blood vessels. VEGF-C signaling through the VEGFR3 receptor plays an essential role in the formation of lymphatic vessels. Another set of ligands, the angiopoietins, play a critical role in vascular endothelial–vascular smooth muscle cell interaction and blood vessel remodeling. Angiopoietin-1 binds to the endothelially expressed Tie-2 receptor to promote the acquisition of vascular smooth muscle cells and stabilization and maturation of nascent blood vessels. Loss of either angiopoietin-1 or Tie-2 function is lethal

during embryogenesis; mice with knockouts of the genes encoding these proteins die early in development with highly enlarged, poorly remodeled blood vessels deficient in vascular smooth muscle cells. Many additional genes have been identified that also play important roles in the specification, differentiation, remodeling, and maturation of blood vessels. It has become clear from the myriad of molecular studies of blood vessel formation conducted to date that regulation of vessel assembly is a highly complex and exquisitely regulated process that we have only just begun to understand.

Clinical importance. Much of the recent explosion of scientific interest in the mechanisms regulating growth and assembly of blood vessels has been driven by potential antiangiogenic (inhibiting vessel development) or proangiogenic (enhancing vessel development) therapeutic applications.

Antiangiogenic therapy. The idea that growth of tumors can be inhibited by targeting the blood vessels that supply them, rather than the tumors themselves (**Fig. 2**), grew out of observations that tumors acquire blood vessels, and that their ability to promote local angiogenesis correlates with progression from relative dormancy to rapid growth and aggressiveness (the term "angiogenic switch" has been coined to describe this transition). A variety of tumor-secreted factors mediate the angiogenic switch, including the factors described above that regulate normal vessel growth and development, notably the critical vascular regulator VEGF. A worldwide search is now on for antiangiogenic compounds that can halt or reverse the growth of tumor blood vessels and thereby stop tumor growth or even cause tumor regression. One focus of attention has been on endogenously produced antiangiogenic factors, many of them normal products of the in-vivo processing of extracellular matrix or plasma proteins. More recently, some of the prominent molecular players described above (particularly VEGF) have been directly targeted for antiangiogenic cancer therapy using various means, including chemical inhibitors, interfering fragments of receptors or ligands, or interfering antibodies. Although human trials of antiangiogenic therapies are mostly in early stages, animal studies and preliminary human trials have yielded encouraging findings, including a lack of general toxicity (unlike chemotherapeutic agents traditionally used for attacking tumor cells directly, for example) and efficacy against many tumors in animal models. Fortuitously, it appears that tumor vessels are differentially sensitive to antiangiogenic therapies compared with normal host blood vessels, probably because the morphologically abnormal tumor vessels deficient in pericyte/smooth muscle cells are less stable than normal host vessels.

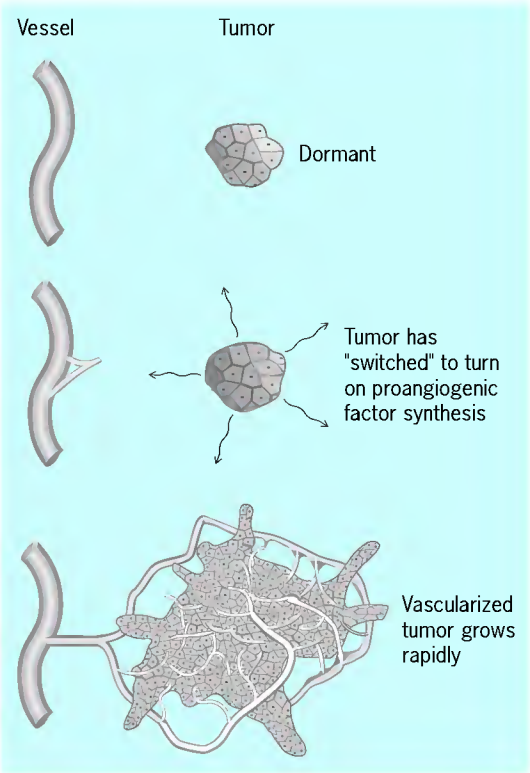

Fig. 2. Tumor angiogenesis and antiangiogenic therapy.

Proangiogenic therapy. Treatment of limb or cardiovascular ischemia (insufficient blood supply) using proangiogenic therapy is also an exciting new possibility, promoting regrowth of new vessels into tissues that because of disease and/or injury have become deficient in blood supply. To date, attempts at proangiogenic gene therapy using the genes encoding VEGF and FGFs (fibroblast growth factors) have been mixed.

Animal developmental models. One difficulty in developing effective therapies for either antiangiogenic treatment of cancer or proangiogenic treatment of ischemia is that although we have now identified many important molecular regulators of blood vessel growth and maintenance, we still have a relatively rudimentary understanding of how all of these regulators function together in the complex orchestration of blood vessel morphogenesis. A great deal of the insight into the functional roles of the key vascular signaling molecules as well as the receptors and intracellular pathways that transduce these signals has come from studies of blood vessel formation during development in animal model organisms. Targeted disruption and overexpression experiments in mice have provided definitive assays for the in-vivo functions of vascular genes during development. In most cases these genes have later been shown to play analogous functional roles in pathologic or nonpathologic angiogenesis in adults (neoangiogenesis). Studies in developing avians, *Xenopus* (frogs), and zebrafish have also begun to provide important new insights into vessel specification, differentiation, and assembly during early development.

Zebrafish. The zebrafish shows particular promise as a new model organism for understanding how blood vessels are assembled in vivo. Zebrafish are tropical fish native to Southeast Asia that have a number of advantageous attributes for developmental studies, particularly studies of vascular development.

Zebrafish adults are small; fully grown animals are only about an inch long, so large numbers of fish can be housed in a small space. Zebrafish have a relatively short generation time, and adult females lay hundreds of eggs every few weeks, so large numbers of progeny can be obtained for genetic or experimental studies. The eggs are fertilized and develop externally to the mother; thus zebrafish embryos and larvae are readily accessible for noninvasive observation or experimental manipulation at all stages of their development. Early development is very rapid: Blood circulation begins within 24 hours after fertilization, larvae hatch by approximately 2.5 days, and larvae are swimming and feeding by 5–6 days.

Two additional features of zebrafish make them particularly useful for studying vascular development. First, developing zebrafish are very small—a 2-day postfertilization (dpf) embryo is just 2 mm long. The embryos are so small, in fact, that the cells and tissues of the zebrafish receive enough oxygen by passive diffusion to survive and develop in a reasonably normal fashion for the first 3 to 4 days, even in the complete absence of blood circulation. This makes it fairly straightforward to assess the cardiovascular specificity of genetic or experimental defects that affect the circulation. Second, zebrafish embryos and early larvae are virtually transparent. The embryos of zebrafish (and many other teleosts) are telolecithal; that is, yolk is contained in a single large cell separate from the embryo proper. The absence of obscuring yolk proteins gives embryos and larvae a high degree of optical clarity. Genetic variants deficient in pigment cells or pigment formation are even more transparent. This remarkable transparency is probably the most valuable feature of the fish for studying blood vessels in vivo, and has been exploited in a number of different ways.

Confocal microangiography, a technique that permits high-resolution three-dimensional visualization of fluorescent structures, has been used to study functioning blood vessels throughout the zebrafish at every stage of embryonic or early larval development. Transgenic zebrafish expressing green fluorescent proteins in blood vessels allow visualization of vascular endothelial cells and their angioblast progenitors prior to initiation of circulation through a vessel, and even before vessel assembly. A complete atlas of the anatomy of the developing vasculature throughout early development has been prepared using confocal microangiography, the first such atlas to be compiled for any vertebrate. Together these and other tools make it possible to perform in-vivo genetic and experimental dissection of the mechanisms of blood vessel formation in the zebrafish. This facility has already been used to dissect the nature of molecular signals guiding arterial-venous differentiation of endothelial cells during development, and promises to make the fish an important resource for understanding the cues and signals that guide the patterning of growing blood vessels. *See* ARTERY; BLOOD VESSELS; CANCER (MEDICINE); CAPILLARY (ANATOMY); CIRCULATORY SYSTEM; DEVELOPMENTAL BIOLOGY; TUMOR; VEIN. Brant M. Weinstein

Bibliography. E. M. Conway, D. Collen, and P. Carmeliet, Molecular mechanisms of blood vessel growth, *Cardiovasc. Res.*, 49:507–521, 2001; N. Ferrara, VEGF and the quest for tumour angiogenesis factors, *Nat. Rev. Cancer*, 2:795–803, 2002; J. M. Isner, Myocardial gene therapy, *Nature*, 415:234–239, 2002; S. Isogai, M. Horiguchi, and B. M. Weinstein, The vascular anatomy of the developing zebrafish: An atlas of embryonic and early larval development, *Dev. Biol.*, 230:278–301, 2001; R. Kerbel and J. Folkman, Clinical translation of angiogenesis inhibitors, *Nat. Rev. Cancer*, 2:727–739, 2002; N. D. Lawson and B. M. Weinstein, Arteries and veins: Making a difference with zebrafish, *Nat. Rev. Genet.*, 3:674–682, 2002; B. Weinstein, Vascular cell biology in vivo: A new piscine paradigm?, *Trends Cell Biol.*, 12:439–445, 2002.

Angle modulation

Modulation in which the instantaneous angle varies in proportion to the modulating waveform. When used in an analog manner, it is referred to as either frequency modulation (FM) or phase modulation (PM), depending upon whether the instantaneous frequency or instantaneous phase varies with the modulation. When used in a digital modulation format, it is referred to as either frequency-shift keying (FSK) or phase-shift keying (PSK). Both FSK and PSK can be used with either a binary or an M-ary alphabet, where M is an integer greater than 2. Indeed, the most common forms of PSK correspond to an input alphabet size of either two or four; while for FSK alphabets of two, four, or eight, symbols are typically employed.

The spectrum of an angle-modulated signal for an arbitrary modulating waveform (that is, an arbitrary message) is difficult to determine. However, if that modulating waveform is a sine wave of frequency f_m, modulating a carrier of frequency f_c, then it is straightforward to show that the modulated signal has a discrete frequency spectrum with components at $f_c, f_c \pm f_m, f_c \pm 2f_m, f_c \pm 3f_m, \ldots$ (**illus.** *a*). The amplitude of the nth harmonic of the spectrum, at frequency $f_c + nf_m$, is given by the magnitude of $J_n(\beta)$, a Bessel function of order n and argument β. This quantity β is known as the modulation index. For phase modulation, β is the maximum deviation of the signal phase from the carrier phase, in radians; while for frequency modulation, it is the ratio $\Delta f/f_m$, where Δf is the maximum deviation of the signal frequency from the carrier frequency. *See* MODULATION.

For a PSK waveform modulated by a message that is modeled as a sequence of random binary data, the power spectral density is a continuous function of frequency, f, which has a peak centered at the carrier frequency, f_c (illus. *b*). It has the form $A \sin^2 x/x^2$, where A is a constant and x is proportional to $f-f_c$. If, instead of phase-shift-modulating a carrier, the message frequency-shift modulates the signal between the carrier frequencies f_{c1} and f_{c2}, the power spectral

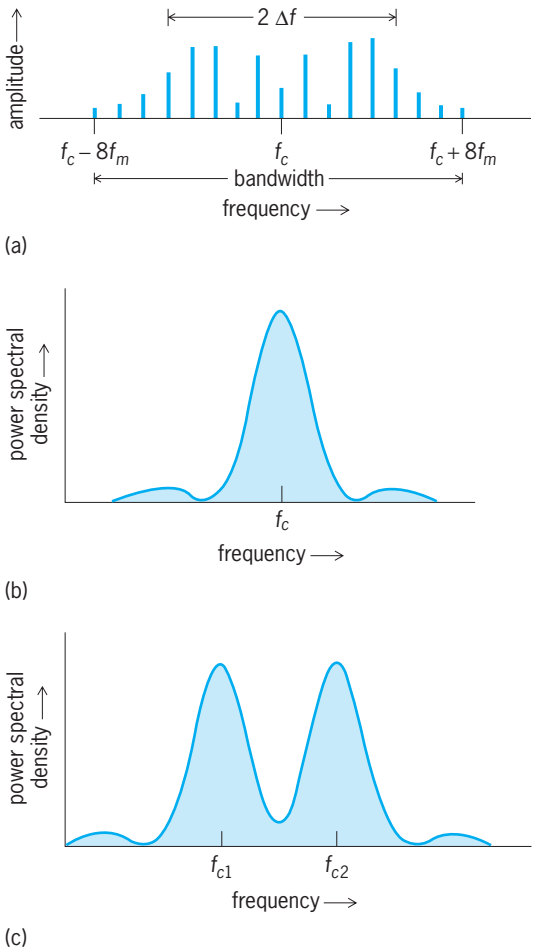

Spectral diagrams of angle-modulated signals. (*a*) Amplitude-frequency spectrum of a signal that is frequency-modulated or phase-modulated by a sine wave, with modulation index $\beta = 5$ (*after M. Schwartz, Information Transmission, Modulation, and Noise, 4th ed., McGraw-Hill, 1990*). (*b*) Power spectral densities of PSK signal and (*c*) FSK signal (*after B. P. Lathi, Modern Digital Communications Systems, 3d ed., Oxford University Press, 1998*).

density has peaks at both these frequencies (illus. *c*). *See* FREQUENCY MODULATION; MODULATION; PHASE MODULATION. Laurence B. Milstein

Bibliography. M. Schwartz, *Information Transmission, Modulation and Noise*, 4th ed., 1990; H. Taub and D. L. Schilling, *Principles of Communication Systems*, 2d ed., 1986; R. E. Ziemer and W. H. Tranter, *Principles of Communications*, 4d ed., 1994.

Anglesite

A mineral with the chemical composition $PbSO_4$. Anglesite occurs in white or gray, orthorhombic, tabular or prismatic crystals or compact masses (see **illus.**). It is a common secondary mineral, usually formed by the oxidation of galena. Fracture is conchoidal and luster is adamantine. Hardness is 2.5–3 on Mohs scale and specific gravity is 6.38. Anglesite fuses readily in a candle flame. It is soluble with difficulty in nitric acid. The mineral does not

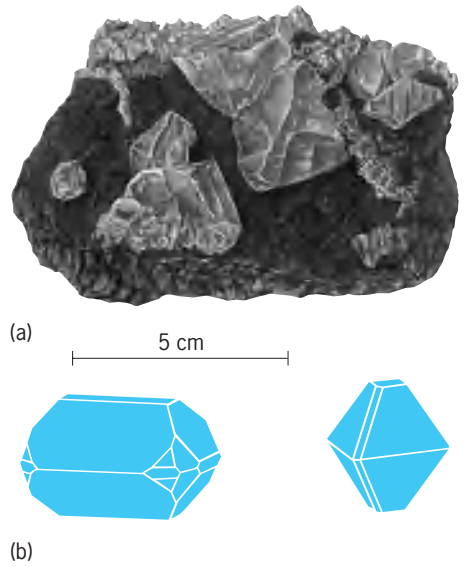

Anglesite. (*a*) Crystals on galena from Phoenixville, Pennsylvania. (*Bryn Mawr College specimen*). (*b*) Crystal habits (*after L. G. Berry and B. Mason, Mineralogy, W. H. Freeman, 1959*)

occur in large enough quantity to be mined as an ore of lead, and is therefore of no particular commercial value. Fine exceptional crystals of anglesite have been found throughout the world. In the United States good crystals of anglesite have been found at the Wheatley Mine, Phoenixville, Chester County, Pennsylvania, and in the Coeur d'Alene district of Shoshone County, Idaho. Edward C. T. Chao

Anguilliformes

The true eels, a large order of actinopterygian fishes, also known as the Apodes. The Anguilliformes are related to Saccopharyngiformes (sack pharynx fishes), Elopiformes (tarpons), and Notacanthiformes (spiny eels and halosaurs); all have a leptocephalous larval stage in development. The chief characteristics of the Anguilliformes include an elongate body with numerous vertebrae; a pectoral girdle which, when present, is free from the head; absence of a pelvic girdle and pelvic fins in Recent adults; dorsal and anal fins confluent with the caudal fin (caudal fin absent in some ophichthids); loss of skeletal parts, especially about the head (for example, the orbitosphenoid, posttemporal, symplectic, posttemporal bones, and gular plate are absent); mesocoracoid and postcleithra absent; swim bladder present, usually physostomous (that is, with a connection to the esophagus); restricted gill openings; no fin spines; scales usually absent or, if present, cycloid and embedded in the skin; and 6 to 49 branchiostegal rays. *See* ACTINOPTERYGII; EEL; ELOPIFORMES; NOTACANTHOIDEI; SACCOPHARYNGIFORMES; SWIM BLADDER.

The Anguilliformes comprise three suborders, 15 families, 141 genera, and about 791 species. The largest family of each suborder is treated below.

Suborder Anguilloidei. Frontal bones are divided (sutured), and scales present or absent. There are three families, five genera, and about 29 species.

The largest family is the Anguillidae (freshwater eels). Unlike most eels of the world, anguillids have small scales embedded in the skin; however, scales do not develop until an individual is 2 or 3 years old. They are further distinguished by having pectoral fins, and a caudal fin continuous with the dorsal and anal fins; branchiostegals do not overlap; the anterior nostril is tubular, near the tip of the snout; the posterior nostril is round and in front of the eye. One genus and about 16 species occur worldwide, except in mainland Asia, the eastern Pacific, and most of South America. Adult anguillids live in freshwater streams or in brackish-water estuaries and, upon reaching maturity, migrate to sea to spawn.

American eel (*Anguilla rostrata*). (*After G. B. Goode, Fishery Industries of the U.S., 1884*)

In the case of the common American eel (*Anguilla rostrata*) [see **illustration**], the life cycle encompasses several morphological and physiological stages: egg > leptocephalus > glass eel > elver > yellow eel > silver eel > death after spawning. The act of spawning has never been observed. The leptocephalus is a planktonic larva; the glass eel is the first stage of the metamorphosed leptocephalus and lacks pigment; the elver is completely metamorphosed and pigmented; the yellow eel is the nonmigratory juvenile; and the mature silver eel is ready for seaward migration. The leptocephali drift toward the coast of the Americas, a journey requiring about one year, at which time they attain 60–80 mm (2.4–3.2 in.) in length and begin to metamorphose. During metamorphosis, larvae shrink, leaving the glass eels shorter than their precursors. Female eels ascend Atlantic and Gulf coastal rivers, often far into the heartland of North America. There they remain 5–40 years (depending on latitude) before returning to sea. In the meantime, males remain on or near the coast. Males are generally smaller than females and develop larger eyes just before migration.

It remains a mystery as to the exact place of spawning. Adult American eels have never been captured beyond the continental shelf, and neither spawning adults nor their eggs have been observed. The only empirical evidence of a possible spawning site is in 2000 m (6560 ft) of water of a geologically old, deep basin of the Bahamas where two adults were photographed.

Suborder Muraenoidei. Frontals are divided (sutured), and scales absent. There are three families, 24 genera, and about 207 species.

The largest family is the Muraenidae (moray eels). Members of the family, consisting of 15 genera and about 185 species, lack pectoral fins; the gill opening is restricted to a small oval aperture; the teeth are long and fanglike; posterior nostrils are above and in front of the eyes; and some species of *Gymnothorax* are ciguatoxic (ciguatoxin is a poisonous substance accumulated up the food chain in the flesh and viscera of some fish). Maximum length is about 3 m (9.8 ft). Moray eels occur in the tropical and temperate seas of the world.

Suborder Congroidei. Frontals are fused scales rarely present. There are nine families, 112 genera, and about 555 species.

The largest family is the Ophichthidae (snake eels and worm eels). Ophichthidae is a family of marine eels, none of which regularly occurs in freshwater. The family differs from other eels in the unusual form of the branchiostegal rays, which are long, slender, and numerous, do not attach basally to the hyoid bone, and broadly overlap those of the other side, leaving the branchiostegal region inflated or slightly bulbous. It is the most diverse and speciose family of eels, consisting of about 250 known species in 55 genera. Ophichthids occur in coastal marine habitats of tropical and warm temperate waters of the world.

Herbert Boschung

Bibliography. E. B. Böhlke (ed.), *Fishes of the Western North Atlantic*, Part 9, vol. 1: *Orders Anguilliformes and Saccopharyngiformes* (pp. 1–655), vol. 2: *Leptocephali* (pp. 657–1055), Sears Foundation for Marine Research Memoir (Yale University), New Haven, 1989; J. E. McCosker and R. H. Rosenblatt, A revision of the Eastern Pacific snake-eel genus *Ophichthus* (Anguilliformes: Ophichthidae) with the description of six new species, *Proc. Calif. Acad. Sci.*, 50(19):397–432, 1998; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Angular correlations

An experimental technique that involves measuring the manner in which the likelihood of occurrence (or intensity or cross section) of a particular decay or collision process depends on the directions of two or more radiations associated with the process. Traditionally, these radiations are emissions from the decay or collision process. However, a variant on this technique in which the angular correlations are between an incident and emitted beam of radiation has been widely used; this variant is known as angular distributions.

The fundamental reason for performing such measurements, rather than just scrutinizing a single radiation in a particular direction or measuring the total intensity for a process, is that the angular correlation or angular distribution measurement provides much more information on both the decay or collision process and on the structure and properties of the emitter of the radiation. The technique is used to study a variety of decay and collision processes in atomic and molecular physics, condensed-matter

(solid-state) and surface physics, and nuclear and particle physics.

Detection of multiple emissions. The principal use of this technique in nuclear physics has been to determine the angular momentum, or spin, and parity of excited nuclear states which are radioactive, that is, decay spontaneously, by measuring in coincidence the radiation in specific directions from two successive transitions in the radioactive cascade. The measurements are generally of coincidences between gamma rays, but coincidences between gamma rays and electrons (beta particles) are also used. The form of the angular correlation, the measured intensity as a function of the angle between the two radiations, gives the information about the intermediate excited state in the cascade. *See* GAMMA RAYS; NUCLEAR SPECTRA; NUCLEAR STRUCTURE; RADIOACTIVITY.

In atomic and molecular collisions as well as in nuclear and particle collisions, this technique is employed as a means of completely specifying the dynamics of the collision, with the added proviso that the energies of the emitted radiations are also to be measured. Wide use has been made of angular correlations in the impact ionization of atoms by electrons where the directions of both the scattered electron and the ejected electron are measured. This application is known as (e,2e) spectroscopy. These studies, particularly just above the ionization threshold energy, have revealed the details of the law governing threshold ionization by charged particles, known as the Wannier law. Furthermore, these measurements reveal considerable information about the quantum-mechanical wave function of the system after the ionization process has taken place, that is, the final state. This type of final state, two electrons and one positive ion, is an example of a three-body continuum Coulomb state. *See* ATOMIC STRUCTURE AND SPECTRA; SCATTERING EXPERIMENTS (ATOMS AND MOLECULES).

Angular distributions. This technique is much simpler experimentally than is the general angular correlations technique, since coincidence measurements are not required. Consequently, it can be employed in situations where the probability (or cross section) for the process is so small that a coincidence experiment would not lead to enough counts to make the result meaningful. This technique allows the extraction of the multipolarity of the incident radiation, provided the target orientation is known. The widest application of this technique has been in atomic, molecular, and condensed-matter investigations, and both emitted electrons and electromagnetic radiation (x-rays, ultraviolet, and visible) have been studied. The most important incident particles are electromagnetic radiation, generated principally by synchrotron radiation. Other incident particles include electrons, protons, and heavy ions and atoms. *See* MULTIPOLE RADIATION; SYNCHROTRON RADIATION.

Photoelectron spectroscopy. In photoionization, the study of the angular distribution of photoelectrons emitted from a target as a result of ionization by electromagnetic radiation (photons), the technique is used in several different ways. The atomic and molecular targets are generally randomly oriented. If unpolarized photons are used, the photoelectron direction with respect to the photon beam direction is measured; for linearly polarized photons, the photoelectron direction with respect to the polarization is measured. This technique is used primarily at photon energies low enough (below about 500 electronvolts, corresponding to photon wavelengths longer than about 2.5 nanometers) that they constitute electric dipole radiation, to an excellent approximation. The general form of the angular distribution is determined by a parameter β, known as the asymmetry parameter, that contains dynamical information about both the target and the photoionization process. The dependence of β on photon energy is extremely helpful in understanding the details of the photoionization cross section, particularly in energy regions where resonances or minima occur. In addition, since β depends on the quantum-mechanical phases of each of the possible final states, phase information, which is unavailable from total cross section measurements, can be obtained. *See* ELECTRON SPECTROSCOPY; PHOTOIONIZATION; POLARIZED LIGHT.

For oriented targets, that is, target atoms or molecules whose magnetic moments, spins, or angular momenta are all aligned, the general form of the angular distribution is more complex and depends upon both the degree of alignment and the state of polarization of the incident radiation. This is often the case for studies of surfaces and crystalline solids where the crystal structure aligns the constituent atoms and molecules; then the technique is generally known as angle-resolved photoelectron spectroscopy. In surface studies, this technique can often aid in the understanding of the details of how the atoms and molecules are attached to the surface along with the geometry of their orientation with respect to the surface. *See* ANGULAR MOMENTUM; SPIN (QUANTUM MECHANICS); SURFACE PHYSICS.

Decay of excited states. When an excited state decays, energy is given off in the form of some kind of radiation. Measuring the angular distribution of the radiation provides certain information about the excited state. In particular, if the emission is isotropic, that is, independent of direction, the excited state was also isotropic. An angular distribution that does depend upon direction (anisotropic) implies that the excited state too had a preferred direction. Thus, the technique provides information about the process that created that excited state. The technique is used in nuclear physics with observation of emitted gamma rays, particularly in connection with radioactive decay. It is also used extensively in atomic and molecular physics. When electromagnetic radiation emitted from atoms or molecules is observed, the technique is called fluorescence spectroscopy, and with emitted electrons it is known as Auger spectroscopy. *See* AUGER EFFECT; X-RAY FLUORESCENCE ANALYSIS.

Steven T. Manson

Bibliography. J. Boyle and M. Pindola (eds.), *Many-Body Interactions in Atomic Physics,* 1995; P. C. Deshmukh, V. Radojevič, and S. T. Manson, Photoionization of the outer shells of radon and radium, *Phys. Rev. A,* 45:6339–6348, 1992.

Angular momentum

In classical physics, the moment of linear momentum about an axis. A point particle with mass m and velocity $\mathbf{v}$ has linear momentum $\mathbf{p} = m\mathbf{v}$. Let $\mathbf{r}$ be an instantaneous position vector that locates the particle from an origin on a specified axis. The angular momentum $\mathbf{L}$ can be written as the vector cross-product in Eq. (1) The **illustration** shows the geometrical meaning of this equation.

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \tag{1}$$

See CALCULUS OF VECTORS; MOMENTUM.

The time rate of change of the angular momentum is equal to the torque $\mathbf{N}$. A rigid body satisfies two independent equations of motion (the dynamical equations) given by Eqs. (2) and (3), where d/dt

$$\frac{d}{dt}\mathbf{p} = \mathbf{F} \tag{2}$$

$$\frac{d}{dt}\mathbf{L} = \mathbf{N} \tag{3}$$

denotes the rate of change, the derivative with respect to time t. Only Eq. (1) is required for a point particle. Equation (2) indicates that a rigid body acts as a point particle located at its center of mass. The motion of the center of mass depends upon the net force $\mathbf{F}$, which is the vector sum of all applied forces. Equation (3) gives the angular motion about the center of mass. The case of statics occurs when the net force and net torque both vanish. *See* KINETICS (CLASSICAL MECHANICS); STATICS; TORQUE.

Vectors and pseudovectors. Given the definition of angular momentum in Eq. (1), although it has three components, it is not a vector. The parity transformation π transforms vectors such as $\mathbf{r}$ and $\mathbf{p}$ according to Eqs. (4). Thus, $\pi \cdot \mathbf{L} \rightarrow +\mathbf{L}$, so that

$$\mathbf{r}' = \pi \cdot \mathbf{r} = -\mathbf{r} \tag{4a}$$

$$\mathbf{p}' = \pi \cdot \mathbf{p} = -\mathbf{p} \tag{4b}$$

$\mathbf{L}$ is called a pseudovector or axial vector. Mathematically, these quantities are called tensors of odd relative weight. Examples of pseudovectors include angular momentum, torque, and magnetic fields. Two or more pseudovectors are combined by adding their components, so that the total angular momentum is the sum of the angular momenta of the individual particles.

Among the fundamental particles, the photon field is a massless vector field and the rho meson is a pseudovector field. *See* ELEMENTARY PARTICLE; PARITY (QUANTUM MECHANICS).

Rigid-body motion. To study the nonrelativistic mechanics of a rigid body made up of N particles, it is convenient to write the instantaneous velocity $\mathbf{v}_\alpha$ of the α-th particle in terms of angular velocity $\boldsymbol{\omega}$ as $\mathbf{v}_\alpha = \boldsymbol{\omega} \times \mathbf{r}_\alpha$. For this system, Eq. (1) becomes Eq. (5),

$$\mathbf{L} = \sum_{\alpha=1}^{N} m_\alpha \mathbf{r}_\alpha \times (\omega \times \mathbf{r}_\alpha) = \mathbf{I} \cdot \omega \tag{5}$$

where $\mathbf{I}$ is the moment of inertia tensor. The moment of inertia tensor can be written in matrix form as Eq. (6). Representative components of this tensor are given by Eqs. (7). Since $\mathbf{I}$ is symmetric, it can al-

$$I_{ij} = \sum_{\alpha=1}^{N} m_\alpha [r_\alpha^2 \delta_{ij} - (r_\alpha)_i (r_\alpha)_j] \tag{6}$$

$$I_{11} = \sum_{\alpha=1}^{N} m_\alpha (y_\alpha^2 + z_\alpha^2) \tag{7a}$$

$$I_{12} = -\sum_{\alpha}^{N} m_\alpha x_\alpha y_\alpha \tag{7b}$$

ways be diagonalized and expressed in principal axis (or symmetry axis) form, $\mathbf{I}_{PA} = \text{diag}\,(I_1, I_2, I_3)$, and in this form it is constant in time: $\dot{\mathbf{I}}_{PA} = 0$. The Euler angle coordinates (α, β, γ) are useful because they can be used to rotate a rigid body into its principal axis frame. *See* EULER ANGLES; MOMENT OF INERTIA; ROTATIONAL MOTION.

When all of the forces that act on a system are conservative (that is, $\nabla \times \mathbf{F} \equiv \mathbf{O}$), a potential energy function $V(\mathbf{r})$ exists for which $\mathbf{F} = -\nabla V(\mathbf{r})$. Consider a rigid body of mass M and inertia tensor $\mathbf{I}$ subject to two conservative forces $\mathbf{F}_i = -\nabla_R V_1(\mathbf{R})$ and $\mathbf{F}_2 = -\nabla_{\alpha 0} V_2(\alpha_0)$, where $\mathbf{R}$ is the position vector of the center of mass, and $\alpha_0 = (\alpha, \beta, \gamma)$ are moment-arm-weighted Euler angles. The quantities ∇_R and $\nabla_{\alpha 0}$ are gradient operators. The equations of motion (2) and (3) become Eqs. (8) and (9). Using $\mathbf{I}_{PA}$ and writing out Eq. (9) in component form in the rotating principal-axis frame yields Eqs. (10).

$$M\mathbf{R} = \mathbf{F}_1 \tag{8}$$

$$\mathbf{I} \cdot \dot{\omega} = \mathbf{N}_2 = \mathbf{r} \times \mathbf{F}_2 \tag{9}$$

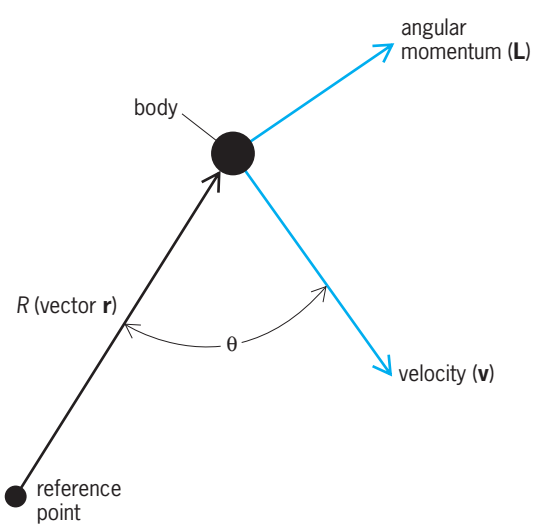

Geometrical definition of angular momentum. If we choose a reference point and draw a line *R* from it to a moving body, then the line *R* and the velocity v of the body of mass *m* define a plane. The angular momentum is a vector that lies perpendicular to this plane. If the momentum of the body is p = *m*v and the line *R* is a vector r, then the magnitude of the angular momentum L is *pr* sin θ.

$$I_1\dot{\omega}_x - (I_2 - I_3)\omega_y\omega_z = N_x \qquad (10a)$$

$$I_2\dot{\omega}_y - (I_3 - I_1)\omega_x\omega_z = N_y \qquad (10b)$$

$$I_3\dot{\omega}_z + (I_1 - I_2)\omega_x\omega_y = N_z \qquad (10c)$$

See ENERGY; RIGID-BODY DYNAMICS.

Conservation. A symmetry is a transformation that leaves a physical system unchanged. A physical quantity is called invariant under a transformation if it remains the same after being transformed. For example, the solutions to Eqs. (2) and (3), or equivalently Eqs. (8) and (9), are invariant under change of the coordinate origin or orientation of the **i**, **j**, and **k** axes. The freedom to choose any orientation of coordinate axis is called rotational invariance, because one choice of axes can be rotated into another.

In physics, the rotational invariance follows from the isotropy and homogeneity of space that has been experimentally established to high accuracy. The diagonalization transformation $\mathbf{I} \rightarrow \mathbf{I}_{PA}$ is carried out by rotating the axes, and thus is a consequence of rotational invariance. The principal axes of a rigid body are called the symmetry axes because their mechanical description is simplest there. (Rotational invariance says that the mechanics can be solved in any orientation of the axes but not that an arbitrary orientation is simplest.)

The study of symmetry shows that one of the deepest relations in physics is that between dynamics and conservation. A physical quantity B is conserved if Eq. (11) is satisfied; that is, B is constant in time

$$\frac{dB}{dt} = 0 \qquad (11)$$

although it may vary in space. Noether's theorem states that if a physical system is invariant under a continuous symmetry, a conservation law exists, provided that the observable in question decreases rapidly enough at infinity. Thus, when the force is zero everywhere (the system is invariant under translation in space), the linear momentum is conserved. If the torque is zero everywhere (the system is invariant under rotation), the angular momentum is conserved. If the system is invariant under translations in time, the total energy is conserved. The converse to Noether's theorem is false, as conserved quantities do not imply continuous symmetries. For example, the parity transformation π of Eqs. (4) can be conserved but is not continuous.

The set of all three-dimensional rotations form a group called SO(3): S for special (det = +1), O for orthogonal (length-preserving), and 3 for the space dimension. The set of 2×2 complex, unitary (length-preserving) special transformations is called SU(2). There is a 2:1 homorphism of SU(2) $\rightarrow$ SO(3), and SU(2) is called the covering group of SO(3).

Angular momentum is important in the evolution of celestial objects. The shapes of these bodies and collections of these bodies, spiral galaxies for instance, follow from the space-time variation of their torques. There is a puzzling problem along these lines: If the universe can rotate, it is not clear why it rotates so slowly; that is, why the anisotropics are so small. A proposed solution is based on the hypothesis of an inflationary epoch in the very early universe, about 10^{-35} s after the big bang. *See* BIG BANG THEORY; CONSERVATION OF ENERGY; CONSERVATION OF MOMENTUM; COSMOLOGY; SYMMETRY LAWS (PHYSICS); UNIVERSE.

Quantum angular momentum. Quantum mechanics has a richer and more complicated structure than classical physics. Because of this, the relationship between symmetry and conservation is even more useful. Whereas in classical physics the observables and states coincide, in quantum mechanics the states of the system correspond to probability amplitudes ψ and the observables **A** to self-adjoint operators (hermitian operators with suitable boundary conditions). According to the Heisenberg uncertainty principle, all observables cannot be simultaneously known, rather only those that commute. Two commuting self-adjoint operators can be diagonalized by a single unitary transformation. The complete set of commuting observables specifies the maximum amount of sharp information possible about a quantal system. The time dependence of a state, represented in **x**-space, where **x** is the position coordinate, as $\psi(\mathbf{x}, t)$, is given by the Schrödinger equation (12), where $\hat{H}$ is the system hamiltonian

$$\hat{H}\psi(\mathbf{x}, t) = i\hbar\left(\frac{\partial\psi}{\partial t}\right)(\mathbf{x}, t) \qquad (12)$$

(energy) operator and $\hbar = 1.05 \times 10^{-34}$ joule-second and is Planck's constant divided by 2π. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Rotational invariance under SO(3) transformations in quantum mechanics implies the angular momentum commutation relations for a single particle given in vector form by Eq. (13). The z component of this equation can be written as Eq. (14), and the x

$$\mathbf{J} \times \mathbf{J} = i\hbar\,\mathbf{J} \qquad (13)$$

$$[J_x, J_y] = i\hbar\,J_z \qquad (14)$$

and y components are expressed in similar fashion, where $[A, B]$ is the commutator of A and B, defined by Eq. (15). The quantum-mechanical momentum

$$[A, B] = AB - BA \qquad (15)$$

J is the most general observable satisfying Eq. (13). The orbital angular momentum $\mathbf{L} = i\hbar\ \mathbf{r} \times \nabla$, where ∇ is the gradient operator, is a special case of a quantum-angular momentum. The total angular momentum $\mathbf{J} = J_x\mathbf{i} + J_y\mathbf{j} + J_z\mathbf{k}$ has squared length $J^2 = J_x{}^2 + J_y{}^2 + J_z{}^2$. It is straightforward to use the operator identity $[A^2,B] = A[A,B] + [A,B]A$ together with Eq. (13) to show that Eq. (16) is satisfied for

$$[J^2, J_i] = 0 \qquad (16)$$

$i = x, y, z$. An inspection of Eqs. (13) and (16) shows that the angular momentum complete set of commuting observables is $\{J^2, J_i\}$, where any single component J_i can be simultaneously diagonal with J^2. Conventionally J_z is chosen as the diagonal component, and J^2 is called the Casimir operator of SO(3).

The two nonself-adjoint ladder operators $J_\pm = J_x \pm J_y$ are useful for determining the properties of angular momentum states.

Let $\psi(J, M, r)$ be a simultaneous eigenfunction of the operators J^2 and J_z with eigenvalues λ and μ, so that Eqs. (17) and (18) are satisfied. The variable r in

$$J^2\psi(J, M, r) = \lambda\psi(J, M, r) \tag{17}$$

$$J_z\psi(J, M, r) = \mu\psi(J, M, r) \tag{18}$$

ψ stands for all other observables which commute with J^2 and J_z. Repeated use of the angular momentum commutation relations of the operators J_+ and J^2 proves that $\lambda = J(J+1)\hbar^2$ and that $\mu = M\hbar$, where $M = -J, -J+1, \ldots, J-1, J$ for each value of J. Further, $J = n/2$, where n is a positive integer, so that $0, {}^1/_2, {}^3/_2, 1, {}^5/_2, \ldots$ are the allowed values of total angular momentum. The eigenfunctions for half-odd-integer values of total angular momentum are called spinors. A value S of spin, together with S_z, labels intrinsic spin states. Spinors are new features of rotational invariance. They have the surprising property of changing sign under an SO(3) rotation through 2π radians, $\mathbf{R}(2\pi)$; that is, Eq. (19) is satisfied. The states

$$\mathbf{R}(2\pi)\psi(J, M, r) = -\psi(J, M, r) \tag{19}$$

$\psi(J, M, r)$ can be spinor-valued since measured probability densities depend only upon $|\psi(J, M, r)|^2 dV$ (where dV is a volume element), whereas the observables must be tensors to have sensible classical limits. The doublet Zeeman splitting of an unpaired electron in an external magnetic field is an indication that the electron has this strange internal structure. Thus, $\mathbf{J} = \mathbf{L} + \mathbf{S}$ is the total angular momentum. *See* SPIN (QUANTUM MECHANICS).

Quantum addition of angular momenta. The discussion above gives the properties of a single angular momentum. Two angular momenta, $\mathbf{J}_1$ and $\mathbf{J}_2$, are called independent if all components of one commute with all components of the other. From the preceding discussion, the sets of eigenfunctions $\psi(J_1, M_1, r)$ and $\psi(J_2, M_2, r)$ can be determined. Let $\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2$. Then the eigenfunction $\psi(J, M, r)$ can be expressed as a linear combination of the $\psi(J_1, M_1, R)$ and $\psi(J_2, M_2, r)$ by Eq. (20), where

$$\psi(J, M, r) = \sum \mathrm{C}(J_1J_2J, M_1M_2)\psi(J_1, M_1, r) \cdot \psi(J_2, M_2, r) \tag{20}$$

$\mathrm{C}(J_1J_2J, M_1M_2)$ is the Wigner coefficient (often erroneously called the Clebsch-Gordon coefficient). The summation is over values of M_1 and M_2 such that $M = M_1 + M_2$, and J must have one of the set of values $J = J_1 + J_2, J_1 + J_2 - 1$.

The Wigner coefficients play another important role in the Wigner-Eckart theorem. Let $\mathbf{T}(JM)$ be a family of tensor operators with angular momentum JM, and consider the quantum matrix elements $\langle j'm'|\mathbf{T}(JM)|jm\rangle$. These matrix elements represent transition probability amplitudes from quantum state jm to state jm' caused by $\mathbf{T}(JM)$. The Wigner-Eckart theorem states that these matrix elements can be expressed by Eq. (21), where $\langle j'|\mathbf{T}(J)|j\rangle$ is the

$$\langle j'm'|\mathbf{T}(JM)|jm\rangle = \mathrm{C}(jJj', mM)\langle j'|\mathbf{T}(J)|j\rangle \tag{21}$$

reduced matrix element and is independent of mMm'. The angular momentum conservation is contained in the Wigner coefficient; the details of the particular operator $\mathbf{T}$ are contained only in the reduced matrix element.

The operator $\mathbf{T}$ in atomic spectroscopy is a perturbation hamiltonian from an incident electromagnetic wave that drives atomic transitions. Thus, the Wigner-Eckart theorem, through the Wigner coefficients, determines which transitions vanish, that is, the selection rules governing the process. Much of atomic spectroscopy follows from the rotational symmetry, or the quantum theory of angular momentum. *See* ATOMIC STRUCTURE AND SPECTRA; SELECTION RULES (PHYSICS). Brian De Facio; John L. Safko

Bibliography. L. C. Biedenharn and J. D. Louck, *Angular Momentum in Quantum Physics*, 1981; E. U. Condon and H. Odabasi, *Atomic Structure*, 1980; E. P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra*, 1959.

Anharmonic oscillator

An oscillator that does not obey Hooke's law. This law is an idealized expression that assumes that a system displaced from equilibrium responds with a restoring force whose magnitude is proportional to the displacement. The use of Hooke's law results in a linear equation of motion that fails to describe many properties of the real world. Nature demonstrates two fundamentally different forms of nonlinearity, which may be called elastic anharmonicity and damping anharmonicity. To understand their difference, and the nature of a harmonic oscillator, it is necessary to understand potential functions. *See* HARMONIC MOTION; HARMONIC OSCILLATOR; HOOKE'S LAW; NONLINEAR PHYSICS.

Atomic interaction. In the classical formulation, where the simplest case is one-dimensional, the x-variable in the potential function $V(x)$ represents displacement away from the stable position of equilibrium that was established between competing forces of attraction and repulsion. For example, in the case of two atoms constrained against rotation, the potential describing their interaction can be approximated by the asymmetric potential curve in **Fig. 1***a*. It is not symmetric because of the Pauli exclusion principle, involving the overlap of electronic orbitals. Thus, it is more difficult to push the atoms together than it is to pull them apart. Therefore, the potential rises more quickly for compression than it does for extension. *See* EXCLUSION PRINCIPLE; INTERMOLECULAR FORCES; MOLECULAR STRUCTURE AND SPECTRA.

Although the potential of Fig. 1 was generated by means of an empirical formula, based on classical physics, $V(x)$ is foundational to the Hamiltonian with which the Schrödinger equation is derived. The Schrödinger equation, in turn, is used to calculate the probability of observing various states of position or

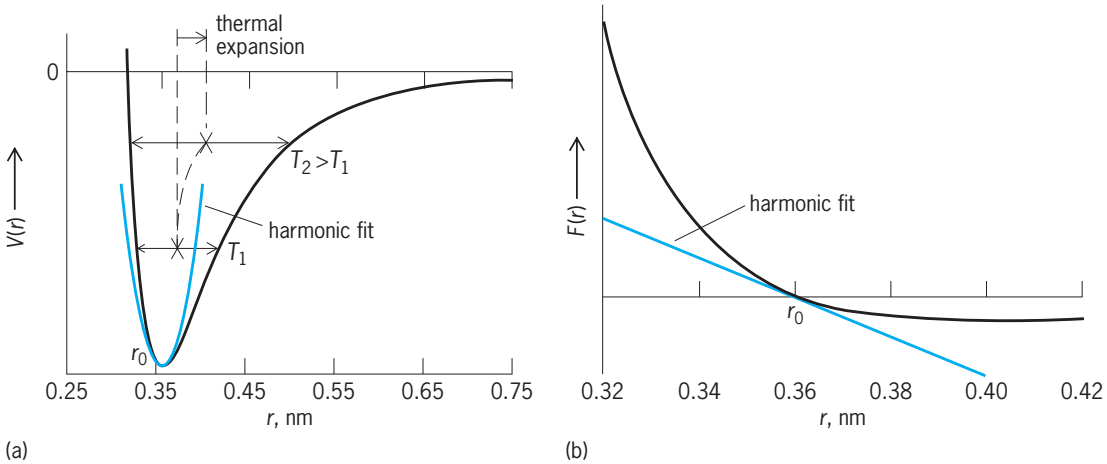

Fig. 1. Interaction between atoms. (*a*) Potential function of the interaction, *V*, plotted against *r*, the separation distance of the atomic nuclei. *V*(*r*) is plotted on a linear scale in units of negative joules. (*b*) Force, *F*(*r*), that is associated with the potential and results from its negative derivative [$F(r) = -dV/dr$] when the atoms are displaced from the equilibrium position r_0. *F*(*r*) is plotted on a linear scale in units of newtons.

momentum. Shown superposed on the anharmonic potential of Fig. 1 is a parabola that represents the harmonic approximation. An appreciation for its limitations can be realized by considering two different oscillatory motions, whose turning points are illustrated by the pair of horizontal lines with arrows on each end. These correspond to two different temperatures, T_1 and T_2, and the amplitude of oscillation is greater at the higher temperature, T_2. The shift toward larger time-averaged separation of the atoms as the temperature increases is the basis for thermal expansion. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; SCHRÖDINGER'S WAVE EQUATION; THERMAL EXPANSION.

There are numerous other cases in which the harmonic potential is known to be inadequate. To cite one example, its use precludes the proper description of the heat capacity of solids. *See* SPECIFIC HEAT OF SOLIDS.

Simple pendulum. Elastic nonlinearity at large amplitudes of the motion is exhibited also by many macroscopic-scale oscillators, such as the pendulum. **Figure 2** shows the potential for a rigid, planar simple pendulum, in which the bob mass is able to go past vertical, that is, past the unstable equilibrium at the angle $\theta = \pi$ radians. This is not possible for a mass on a string, since the string can support only tension and not compression. *See* PENDULUM.

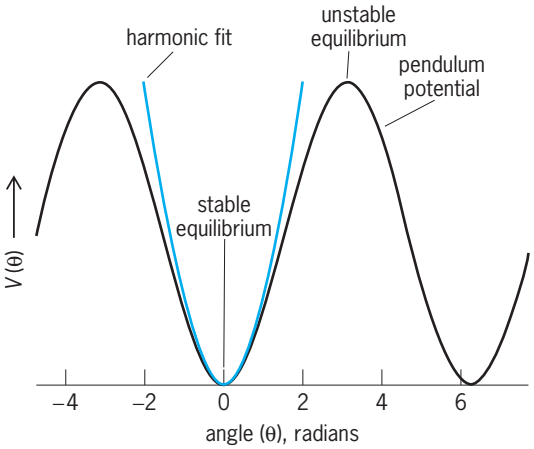

Fig. 2. Potential function for the rigid, planar simple pendulum.

The pendulum has become a macroscale archetype of chaos. For chaotic motion, the energy with which the damped oscillator is driven must be large enough to allow winding modes, in which the bob is allowed to move "over the top." Otherwise, it is a system demonstrating that nonlinearity is a necessary but not sufficient condition for chaos. At low levels, but not so small as to introduce complications due to damping nonlinearities, the motion of the pendulum is nearly harmonic and thus isochronous, meaning that the period is independent of amplitude. At high levels, the motion can be periodic and yet highly nonlinear. Highly nonlinear cases involve subharmonic response, in which it takes longer for the pendulum to repeat its motion than is required for the drive. Such modes cannot be explained by a linear equation of motion. *See* CHAOS.

Highly nonlinear but nonchaotic pendulum motion is a realm of amplitude jumps, as drive of the instrument is "tuned" back and forth in frequency across "resonance." Unlike the harmonic oscillator, when the pendulum is driven at large levels there is not a well-defined resonance frequency. If the drive frequency is below the pendulum's small-motion resonance frequency, small phase perturbations introduced by an external disturbance can seriously disrupt the steady-state motion. The pendulum can transition from a high-level, long-period mode to a low-level, short-period mode, even though the driving force was not altered. Such a change is suggestive of an analogy in physiology, ventricular fibrillation. *See* HEART DISORDERS.

Damping anharmonicity and internal friction. Anharmonicity manifests itself in different ways at different scales. At both the atomic scale and the macroscopic scale, the assumption of a smooth potential function is oftentimes reasonable. At the mesoscale,

this is no longer true, as defect organization takes place via the exceedingly complex interactions responsible for friction. When nature fills the vacuum it abhors, it rarely does so with perfection. At finite temperatures, for example, solids contain lattice sites devoid of an atom. These vacancies are called Schottky defects. Materials are never as strong as theoretical estimates of strength, based on the assumption of perfect lattices, because of defect structures such as dislocations. *See* CRYSTAL DEFECTS.

When material is strained in a cyclic manner, stress/strain hysteresis is observed, and it has been thoroughly studied by engineers. The anharmonic nature of this hysteresis has been described in terms of its influence on oscillator damping as a mechanism of internal friction. It differs in type from elastic anharmonicity, discussed above, and is illustrated by the graph of **Fig. 3**. *See* HYSTERESIS.

This form of anharmonicity has been the source of challenges to those who seek to study gravitational waves. Metastabilities, illustrated by the metastable equilibria in the enlargement in Fig. 3, are present in the springs used to isolate the mirrors of the LIGO (Laser Interferometer Gravitational-Wave Observatory) instrument from earth noise. Their presence disallows the occurrence of an absolute equilibrium position. A primary problem with trying to explain friction from first principles is that the fine structure responsible for these metastabilities is not static. Rather, because they are both temperature- and stress-dependent, defect structures evolve in time. *See* GRAVITATIONAL RADIATION; LIGO (LASER INTERFEROMETER GRAVITATIONAL-WAVE OBSERVATORY).

Damping anharmonicity is of a totally different character than viscous damping, and only the latter case is well known to physics. Because damping anharmonicity is incompatible with the Hooke's-law description of a spring, the description of energy loss must be modified (**Fig. 4**). *See* DAMPING.

In the internal friction case (Fig. 4*a*), it is seen that the minimum of the potential shifts back and forth as the spring is cycled in strain during free decay. This is due to material creep responsible for hysteresis, and it involves alterations in the configuration of defect structures. The amount of a given shift is a function of the amplitude of motion in the half cycle that just preceded the shift. For common exponential damping, the shift is proportional to amplitude. An energy-based description of exponential damping is provided by the equation below for the friction force shown in **Fig. 5**, which involves the function

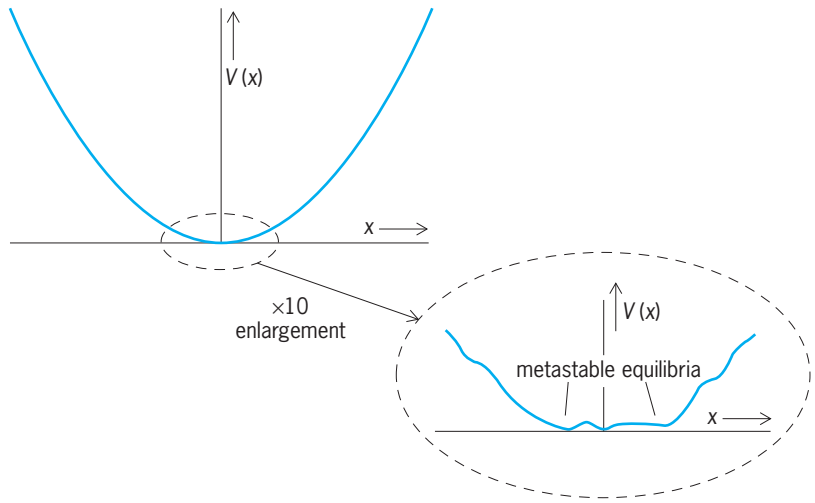

Fig. 3. Anharmonic potential responsible for internal friction damping. Only at low levels (as noted by the ×10 enlargement) does the nonlinear nature of the potential become observable.

$$\text{Friction force} = \text{(per unit mass)} \begin{Bmatrix} \dfrac{\omega}{Q} v & \text{for viscous damping} \\ \dfrac{\pi\omega}{4Q}\sqrt{\omega^2 x^2 + v^2}\,\mathrm{sgn}(v) & \text{for hysteretic damping} \end{Bmatrix}$$

$\mathrm{sgn}(v)$. [When $v > 0$, $\mathrm{sgn}(v) = 1$, and when $v < 0$, $\mathrm{sgn}(v) = -1$.] Here, v is the velocity, ω is the angular frequency, and Q is the quality factor. *See* CREEP (MATERIALS); Q (ELECTRICITY).

Unlike viscous damping which is linear, hysteretic damping is nonlinear. Since the term under the

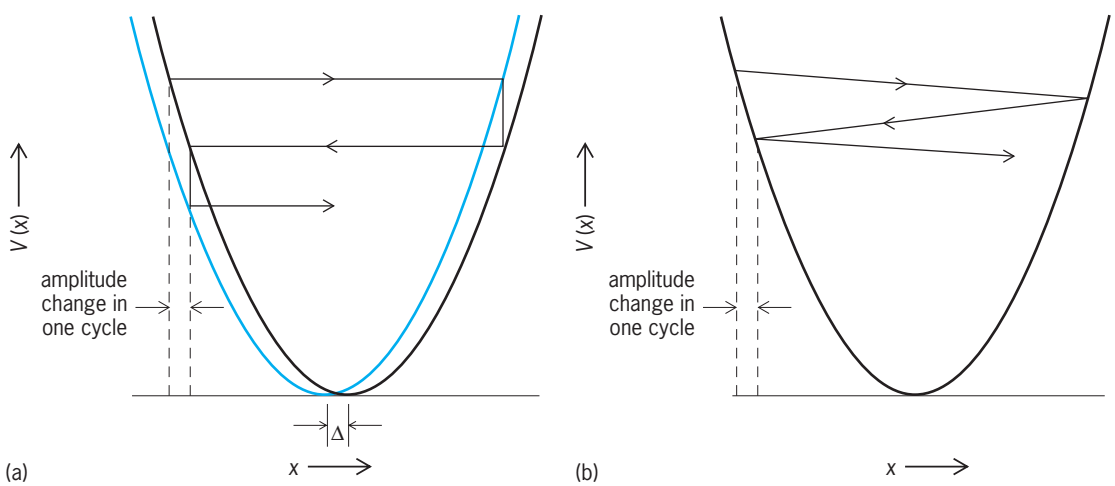

Fig. 4. Comparison of nonlinear and linear oscillator damping. (*a*) Nonlinear damping due to internal friction. (*b*) Linear damping due to an external force.

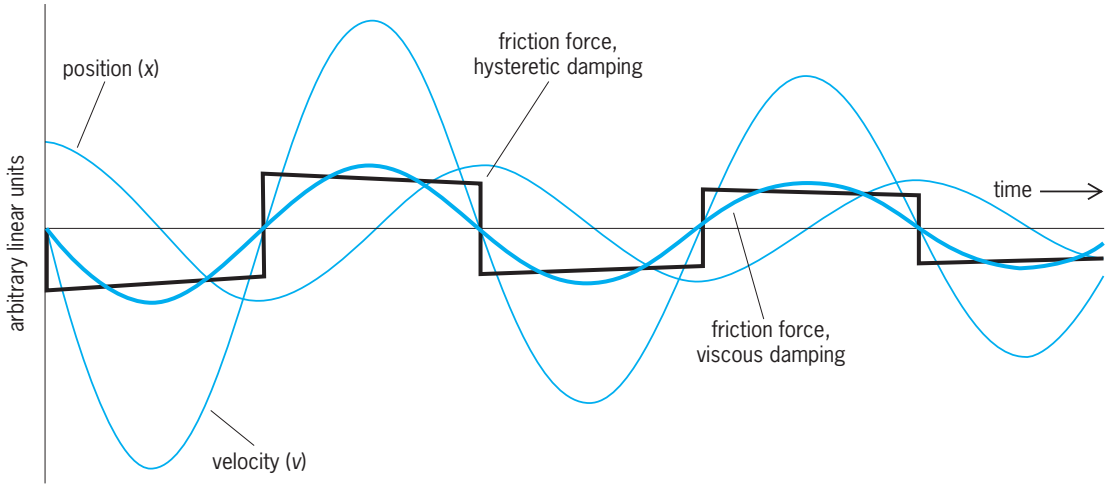

Fig. 5. Comparison of (1) viscous (linear) damping that derives from an external-to-the-oscillator friction force and (2) hysteretic (nonlinear) damping of the internal friction type. Both oscillators are in free decay with quality factor $Q = 10$ and angular frequency $\omega = 3.07$ Hz.

square root in the equation above is twice the energy of oscillation per unit mass, the velocity is seen to be important only by way of its algebraic sign. The two forms are similar only in that they yield the same quality factor Q for the exponential free decay. If one considers a Fourier series to describe the nonlinear friction force, it can be shown that the fundamental component of the series is the only term responsible for the energy loss per cycle that determines the Q. For this reason, many have failed to recognize the ubiquitous nature of nonlinear damping. The two forms differ greatly with regard to frequency dependence of Q. Because $\omega/Q = 2\beta$, where β is the coefficient commonly referred to as the damping constant, it is seen that Q must be proportional to frequency in the case of viscous damping. When monitored at low frequency and low energy, the vast majority of mechanical oscillators exhibit a Q dependence that is quadratic in the frequency, which is consistent with the nonlinear damping model presented. *See* FOURIER SERIES AND TRANSFORMS.

The properties of internal friction responsible for $Q \propto \omega^2$ were first discovered early in the twentieth century, and the investigators claimed that internal friction in solids appeared to be universal. Experiments since the 1990s support this belief, which implies a closer connection than previously recognized to the well-known case of surface friction described by C. A. Coulomb. In an additional early study by L. Portevin and F. Le Chatelier, which has been also mostly overlooked by physics, it was found that strain of alloys under stress is not smooth but jerky. The discontinuous changes observed are reminiscent of the Barkhausen effect, also discovered early in the twentieth century. *See* BARKHAUSEN EFFECT; FRICTION.

This jerky behavior, related to stick-slip features of friction, implies that atomic-scale quantum mechanics is not best suited to the description of internal friction. The electronvolt is infinitesimal compared to the quantum changes observed. There is evidence for an energy of mesoscale quantization, given by mc^2/α, where $m = 9.1 \times 10^{-31}$ kg is the electron mass, $\alpha = 1/137$ is the fine-structure constant, and $c = 3 \times 10^8$ m/s is the speed of light. This Compton scale of energy, having the mesoscale value of 1.1×10^{-11} J, is thought to regulate, through self-organizing metastabilities, the processes "still unknown from first principles" responsible for internal friction. It is possible that one of the last frontiers of physics actually lies in one of the most accessible parts of our world. Randall D. Peters

Bibliography. G. L. Baker and J. P. Gollub, *Chaotic Dynamics, an Introduction*, 2d ed., Cambridge University Press, 1996; R. Desalvo et al., Study of quality factor and hysteresis with the state-of-the-art passive seismic isolation system for Gravitational Wave Interferometric Detectors, *Nucl. Instrum. Meth. Phys. Res., Sec. A*, 538:526–537, 2005; C. Kittel, *Introduction to Solid State Physics*, 8th ed., Wiley, 2005; M. R. Matthews, C. F. Gauld, and A. Stinner (eds.), *The Pendulum: Scientific, Historical and Educational Perspectives*, Springer, 2005; R. D. Peters, Friction at the mesoscale, *Contemp. Phys.*, 45:475–490, 2004; R. D. Peters, Resonance response of a moderately driven rigid planar pendulum, *Amer. J. Phys.*, 64:170–173, 1996.

Anhydrite

A mineral with the chemical composition $CaSO_4$. Anhydrite occurs commonly in white and grayish granular masses, rarely in large, orthorhombic crystals (see **illus.**). Fracture is uneven and luster is pearly to vitreous. Hardness is 3–3.5 on Mohs scale and specific gravity is 2.98. It fuses readily to a white enamel. It is soluble in acids and slightly soluble in water. Anhydrite is an important rock-forming mineral and occurs in association with gypsum, limestone, dolomite, and salt beds. It is deposited directly by evaporation of seawater of high salinity at or above 108°F (42°C). Anhydrite can be produced artificially by dehydration of gypsum at about 390°F (200°C). Under natural conditions anhydrite

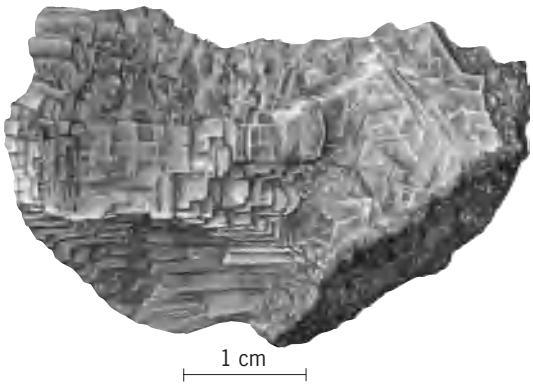

Anhydrite from Montanzas, Cuba. (*Specimen from Department of Geology, Bryn Mawr College*)

hydrates slowly, but readily, to gypsum. It is not used as widely as gypsum. Anhydrite is of worldwide distribution. Large deposits occur in the Carlsbad district, Eddy County, New Mexico, and in salt-dome areas in Texas and Louisiana. *See* GYPSUM; SALINE EVAPORITES. Edward C. T. Chao

Animal

Any living organism which possesses certain characteristics that distinguish it from plants is a member of the animal kingdom. There is no single criterion that can be used to distinguish all animals from all plants. Animals usually lack chlorophyll and the ability to manufacture foods from raw materials available in the soil, water, and atmosphere. Animal cells are usually delimited by a flexible plasma or cell membrane rather than a cell wall composed either of cellulose or chitin, as are the cells of most plants. Animals generally are limited in their growth and most have the ability to move in their environment at some stage in their life history, whereas plants are usually not restricted in their growth and the majority are stationary.

The presence or lack of chlorophyll in an organism does not determine its affinity to the plant or animal kingdom. Among the protozoa, the class Phytamastigophora includes animals, such as the euglenids, which have chromatophores containing chlorophyll. These organisms are considered to be animals by zoologists and plants by phycologists. The vestige of a feeding apparatus in these protozoa indicates that they have descended from forms that ingested food particles, that is, animals. Higher parasitic plants and the large plant group Fungi also lack chlorophyll. Another borderline group is the slime molds: the Mycetozoa of zoologists and the Myxomycophyta of the botanists. These organisms exhibit both plant and animal characteristics during their life history. Movement is not a characteristic restricted to the animal kingdom; many of the thallophytes such as *Oscillatoria*, numerous bacteria, and colonial chlorophytes are motile.

The grouping of living organisms into kingdoms is still unsettled. Previously biologists used only two groups, the Animalia and the Plantae, based on the large, familiar organisms then known to them. The discovery of microorganisms, such as bacteria and viruses, introduced difficulties for this simple scheme. Today biologists recognize up to five kingdoms, but with considerable differences of opinion. Most, however, place the one-celled animals and plants, sometimes along with algae and certain other groups, into the Protista. Other kingdoms are the Monera for the bacteria and blue-green algae, and the Fungi for the slime molds and true fungi. These schemes for recognizing additional kingdoms have the practical advantage of eliminating the difficulties of delimiting and describing the kingdoms of multicellular animals and plants. *See* ANIMAL KINGDOM; PLANT; PLANT KINGDOM. Walter J. Bock

Animal communication

A discipline within the field of animal behavior that focuses upon the reception and use of signals. Animal communication could well include all of animal behavior, since a liberal definition of the term signal could include all stimuli perceived by an animal. However, most research in animal communication deals only with those cases in which a signal, defined as a structured stimulus generated by one member of a species, is subsequently used by and influences the behavior of another member of the same species in a predictable way (intraspecific communication). In this context, communication occurs in virtually all animal species, if only as a means by which a member of one sex finds its partner.

The field of animal communication includes an analysis of the physical characteristics of those signals believed to be responsible in any given case of information transfer. A large part of this interest is due to technological improvements in signal detection (for example, the use of tape and video recorders and gas chromatographs), coupled with analysis of the signals obtained with such devices.

Communication modes. Information transmission between two individuals can pass in four channels: acoustic, visual, chemical, and electrical. An individual animal may require information from two or more channels simultaneously before responding appropriately to reception of a signal. Furthermore, a stimulus may evoke a response under one circumstance but be ignored in a different context.

Acoustic. Sound signals have characteristics that make them particularly suitable for communication, and virtually all animal groups have some forms which communicate by means of sound. Sound can travel relatively long distances in air or water, and obstacles between the source and the recipient interfere little with an animal's ability to locate the source.

Sounds are essentially instantaneous and can be altered in important ways. Both amplitude and frequency modulation can be found in sounds emitted by animals; in some species sound signals have discrete patterns due to frequency and timing of utterances. Since a wide variety of sound signals are

possible, each species can have a unique set of signals in its repertoire. Animals can travel directly toward a sound source, despite variable wind speed and direction, the onset of darkness, or simultaneous signaling by thousands of other animals of various species.

Sound signals are produced and received primarily during sexual attraction, including mating and competition. They may also be important in adult–young interactions, in the coordination of movements of a group, in alarm and distress calls, and in intraspecific signaling during foraging behavior. *See* REPRODUCTIVE BEHAVIOR.

The human hearing range (for example, of frequency and intensity) is somewhat limited, but with new technology sounds outside the range of human perception can be tape recorded. Such research reveals that many species of animals, particularly insects, apparently can communicate at frequencies much higher or lower than the human hearing range; a few species, perhaps elephants, communicate at frequencies below the human range. *See* HEARING (HUMAN); PHONORECEPTION.

Visual. Visual signaling between animals can be an obvious component of communication, as is the case for fireflies. Travel time is essentially instantaneous; however, visual signals have very different characteristics from those of sound. Sound requires a medium for transmission (such as air, water, or a substrate), but light travels unimpeded unless blocked by obstacles. Besides the normal range of human vision (visible light), visual signals, that is, the entire electromagnetic spectrum, include additional frequencies in the infrared (heat) and ultraviolet ranges.

The quality of light that is often considered is color, but other characteristics are important in visual communication. Alterations of brightness, pattern, and timing also provide versatility in signal composition. Two other features of vision deserve mention. First, an animal equipped with binocular vision can easily locate the precise source of a signal. Second, color patterns in animals (for example, butterflies) can serve as relatively permanent signals.

Even with all its favorable characteristics, the visual channel suffers from the important limitation that all visual signals must be line of sight. Information transfer is therefore largely restricted to the daytime (except for animals such as fireflies) and to rather close-range situations because of obstacles normally present in the environment.

Intraspecific visual signaling, as with acoustical and chemical signaling, appears to occur primarily during mate attraction. The color dimorphism of birds, the interesting patterns of butterfly wings, the posturing of some fish, and firefly flashing are examples. Among fireflies, for example, flying males flash at regular intervals in specific patterns as they cruise along. Stationary females of the same species, if receptive, flash at a set interval of time after the male, enabling the male to travel directly to them and not to a member of another species.

Some parent–young interactions involve visual signaling. A young bird in the nest may open its mouth when it sees the underside of its parent's beak. Other examples are the synchronized behavior observed in schooling fish and flocking birds.

As with sounds, the human range of perception of visual signals is somewhat limited. Some other species, such as honeybees, see well into the ultraviolet spectrum, and a few forms, such as some beetle species, can perceive objects in the infrared range.

Chemical. Chemical signals, like visual and sound signals, can travel long distances, but with an important distinction. Distant transmission of chemical signals requires a movement of air or water. Therefore, an animal cannot perceive an odor from a distance; it can only perceive molecules brought to it by a current of air or water. Animals do not hunt for an odor source by moving other than upwind or upcurrent in water, because chemical signals do not travel in still air or water since diffusion is far too slow.

The fact that chemical signals comprise molecules means that they have special attributes important in information transmission, attributes not shared by the other channels. Although acoustical or visual signals are essentially instantaneous, chemical signals have a time lag. A single molecule might stimulate a flying insect to head upwind, but if no other similar molecules are perceived, the insect might well abandon the search or drop downwind until it perceives other molecules of the same type.

Chemical signals at times have to be of an appropriate concentration if they are to be effective. A chemical normally considered to be an attractant can serve as a repellent if it is too strong. Chemical signals may persist for a while, and time must pass before the concentration drops below the threshold level for reception by a searching animal. Since molecules of different sizes and shapes have varying degrees of persistence in the environment, the chemical channel is often involved in territorial marking, odor trail formation, and mate attraction to a fixed location. This channel is particularly suitable where acoustical or visual signals might betray the location of a signaler to a potential predator.

The array of molecular structure is essentially limitless, permitting a species-specific nature for chemical signals. Unfortunately, that specificity can make interception and analysis of chemical signals a difficult matter for research.

Pheromones are chemical signals that are produced by an animal and are exuded through glandular openings (or otherwise) and influence the behavior of other members of the same species. If pheromones are incorporated into a recipient's body (by ingestion or absorption), they may chemically alter the behavior of such an individual for a considerable period of time. *See* PHEROMONE.

In social insects, long-term regulation of the behavior of individuals within a colony appears to be mainly, if not exclusively, due to chemical inhibition or enhancement of the endocrine system of individuals by pheromones received from other members of the colony. Queen honeybees or queen termites produce one or more chemical substances that inhibit

other females (workers) from becoming functional queens. *See* SOCIAL INSECTS.

All animal species seem to have a somewhat limited range of types of molecules that they can perceive. A substance with a particular odor or taste for humans may not even be perceived by a member of another species. The converse is also true: another species may well be able to smell or taste a compound that humans do not notice. *See* CHEMICAL ECOLOGY; CHEMORECEPTION.

Electrical. Some electric fish and electric eels live in murky water and have electric generating organs that are really modified muscle bundles. Communication by electric signaling is rapid; signals can travel throughout the medium (even murky water), and rather complex signals can be generated, permitting species-specific communication during sexual attraction. However, the electrical mode is apparently restricted to those species that have electric generating organs. *See* ELECTRIC ORGAN (BIOLOGY).

Context. Any of the above types of signals can be effective under the right circumstance but ineffective under other circumstances. For example, when recruited honeybees are searching for a potential hive site found by experienced foragers (scout bees), they are attracted by the chemical components of an exudate from a gland at the upper side and tip end of the abdomen. Similarly, when a honeybee swarm is moving from the location of its parent colony to a new site, an odor from the gland orients bees during transit and stimulates settling once the new site has been reached. However, recruited bees searching for flowers exploited by foragers from their colony apparently ignore those same chemicals.

Similarly, male moths of some species are first attracted to the odor of the host plant of their future offspring (plant-feeding caterpillars) and are then attracted to the pheromones exuded by females perched on that host plant. In some cases, if males are not first exposed to the odor of the larval host plant, they may not respond to pheromones exuded by females of their species. *See* PLANT.

Methodology. Animal communication is one of the most difficult areas of study in science for several reasons. First, experiments must be designed and executed in such a manner that extraneous cues (artifacts) are eliminated as potential causes of the observed results. Second, once supportive evidence has been obtained, each hypothesis must be tested. In animal communication studies, adequate tests often rely upon direct evidence—that is, evidence obtained by artificially generating the signal presumed responsible for a given behavioral act, providing that signal to a receptive animal, and actually evoking a specific behavioral act in a predictable manner.

A classic example of such a test entails holding a wooden dowel with a red dot on its underside over a gull chick, eliciting the same behavior in the chick as when a parent gull holds its beak over the nest. Whereas this example is quite simple and can be repeated with little equipment, acoustic, visual, chemical, and electric signals can be quite complex, requiring information derived from physics, chemistry, and mathematics and employing electronics and computers.

Nonhuman language. The fact that human beings have a language indicates that language is possible in an animal species, and this possibility has been investigated repeatedly in several nonhuman species. Dog, cat, and horse owners, for example, may have a strong bias on this point; those who work with personable animals, such as chimpanzees and dolphins, serve as another example.

Experiments conducted by Karl von Frisch in the 1940s yielded results suggestive of a language among honeybees in the form of movements by scout bees through which the location of food sources is communicated to other bees. Although the hypothesis drew wide interest and support because of its emotional appeal, experiments in the 1960s, coupled with a careful analysis of the design and controls of the original series of experiments, indicated that Frisch's own earlier (1937) odor-search hypothesis more than adequately accounted for the observation.

Continued attempts to demonstrate language use in various species (bees, pigeons, chimpanzees, dolphins), and experiments using a computer-simulated bee, have not been able to satisfy critics of the design or interpretation of the experiments. It is possible that human beings may well be the only animal species capable of using symbolic language. *See* ETHOLOGY; PSYCHOLINGUISTICS. Adrian M. Wenner

Bibliography. T. Lewis (ed.), *Insect Communication*, 1984; T. A. Sebeok (ed.), *How Animals Communicate*, 1977; A. M. Wenner, *Concept-centered vs. Organism-centered Research*, 1989; A. M. Wenner, D. Meade, and L. J. Friesen, Recruitment, search behavior and flight ranges of honeybees, *Amer. Zool.*, 31:768–782, 1991; A. M. Wenner and P. H. Wells, *Anatomy of a Controversy: The Question of a "Language" among Bees*, 1990.

Animal evolution

The theory that modern animals are the modified descendants of animals that formerly existed and that these earlier forms descended from still earlier and different organisms.

Animals are multicellular organisms that feed by ingestion of other organisms or their products, being unable to derive energy through photosynthesis or chemosynthesis. Animals are currently classed into about 30 to 35 phyla, each of which has evolved a distinctive body plan or architecture; representatives of the major phyla are depicted in **Fig. 1**.

Invertebrates

All phyla began as invertebrates, but lineages of the phylum Chordata developed the internal skeletal armature, with spinal column, which was exploited in numerous fish groups and which eventually gave rise to terrestrial vertebrates. The number of phyla is uncertain partly because most of the branching patterns and the ancestral body plans from which

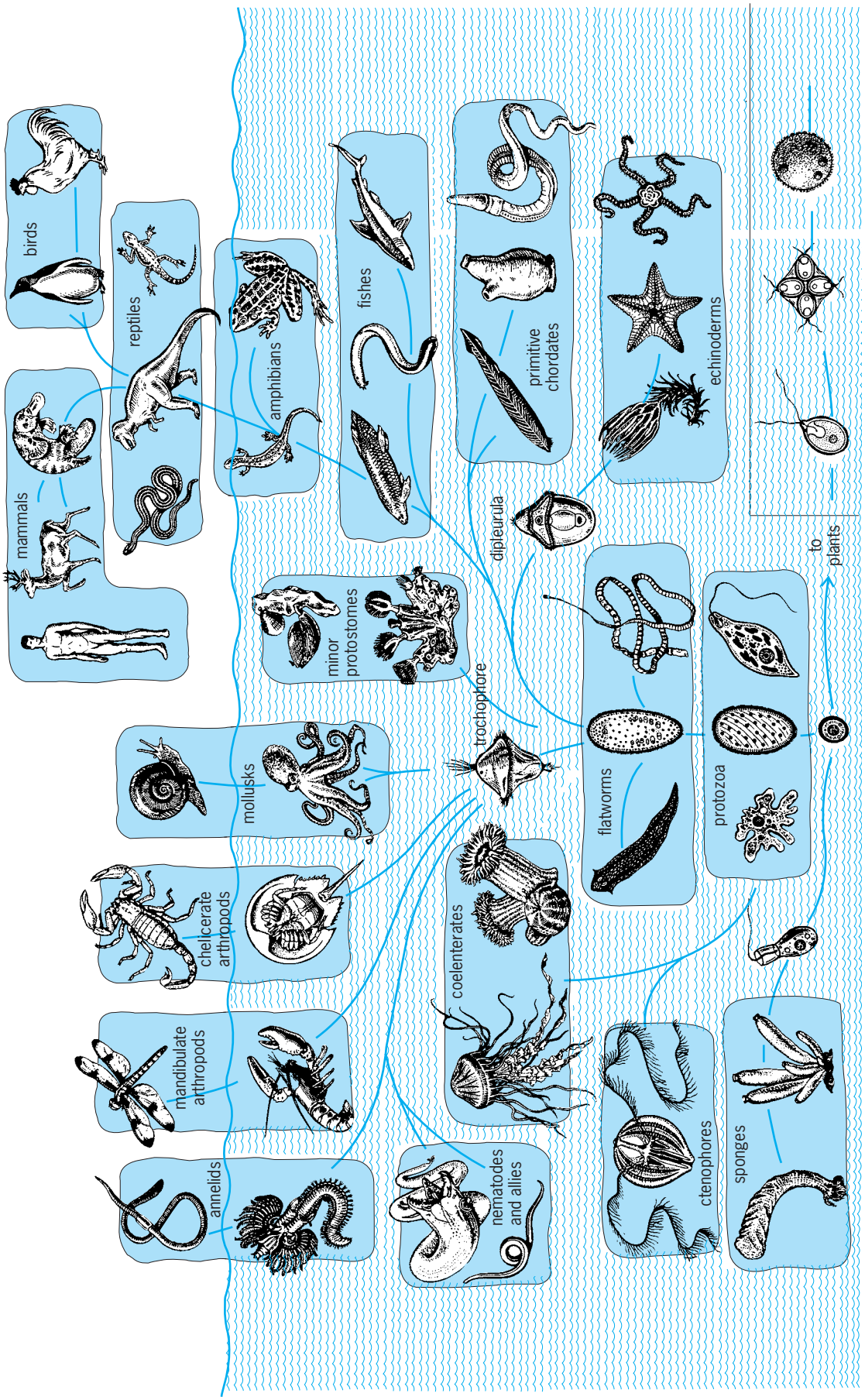

Fig. 1. Evolution of animal groups, showing hypothetical relationships. (*After G. B. Moment, General Zoology, 2d ed., Houghton Mifflin, 1967*)

putative phyla have arisen are not yet known. For example, arthropods (including crustaceans and insects) may have all diversified from a common ancestor that was a primitive arthropod, in which case they may be grouped into a single phylum; or several arthropod groups may have evolved independently from nonarthropod ancestors, in which case each such group must be considered a separate phylum. So far as known, all animal phyla began in the sea. *See* ANIMAL; ANIMAL KINGDOM.

Several lines of evidence bear significantly upon the ancestry of these animals and of their major subdivisions. Classically, animals were grouped according to the similarities of their adult body plans, and while the body plans have provided a basis for the recognition of most of the phyla and other distinctive animal groups, they have not provided definitive evidence of the interrelations of the phyla. A second approach is to compare patterns of early development, which should be more conservative than adult features; such comparisons have permitted the grouping of some phyla into likely alliances. Hypotheses as to the body plans of phylum ancestors have been erected from these morphologic data. It has not proven possible to corroborate any hypotheses, however, for the earliest members of the phyla (and of invertebrate classes) appear abruptly in the fossil record, and their ancestors cannot be traced. Furthermore, numbers of extinct phyla or other major animal groups also appear in the fossil record, adding branches to the tree of life which must be reckoned with, but with no definitive indication of either their origins or branching patterns. Finally, genetic changes that have accumulated after branching events separated the major groups should have left a pattern of variation in the structure of genes that may be reconstructed to yield a phylogeny. If genetic changes have occurred in a reasonably regular manner, then the amount of divergence in the structure of genes and gene products, such as protein and ribonucleic acid (RNA) molecules, should be proportional to the time since branching occurred. Some genes change so fast that their pattern of divergence is useful only for recent branching events, while others are so conservative that they are used to study divergences that occurred billions of years ago. These methods are promising when used in conjunction with other evidence.

Major animal groups. The simplest living phyla may have no direct relationships to other animals (**Fig. 2**), but indicate the versatility of multicellular design. Porifera (sponges) contain several differentiated cell types that form body walls, but their architecture is limited to folding these walls into chambers that are sometimes highly convoluted. Coelenterata (=Cnidaria; sea anemones, corals, and jellyfish) have two cellular body layers, separated by a gelatinous layer; the inner tissue forms a digestive cavity. *See* CNIDARIA; PORIFERA.

Flatworms (Platyhelminthes) and their allies have solid bodies, with both external (ectoderm) and internal (endoderm) tissue layers separated by a third cellular layer (mesoderm); there is a gut lined

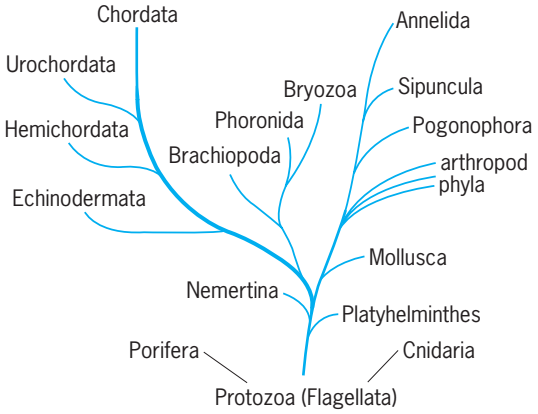

Fig. 2. Phylogenetic tree depicting the possible relations among the phyla according to evidence from all sources. The branching pattern is constrained by models of development and of body-plan evolution and by molecular data. Branch length is not to scale.

with endoderm. All phylogenetic evidence suggests that a flatwormlike form gave rise to the remaining phyla, which possess body cavities (other than a gut) in which reproductive and other organs can be sequestered. In about 15 of these phyla this cavity (the coelom) is lined by mesoderm and usually functions as a hydrostatic skeleton. *See* PLATYHELMINTHES.

The coelomates constitute all of the familiar, larger-bodied animals and are commonly grouped into two main divisions (Fig. 2). In one group, the body was originally segmented, as in annelid worms or arthropods, though the segmentation may be reduced or lost in some branches. In the other group, the coelom is divided into two or three regions; these include the Enchinodermata (starfish and sea urchins), Tunicata (sea squirts), and Chordata (including humans). Additionally, there are coelomate phyla that cannot yet be assigned to either of these groups; some, such as the Brachiopoda, Phoronida, and Bryozoa, display an enigmatic mixture of traits that otherwise characterize one or the other of the main groups. In about 10 phyla (not shown in Fig. 2) the body cavity is not a true coelom but lies between endoderm and mesoderm and is termed a pseudocoel. These pseudocoelomates are soft-bodied, small to minute, and commonly parasitic; some of them have never been found as fossils. *See* ANNELIDA; BRACHIOPODA; BRYOZOA; CHORDATA; ECHINODERMATA; PHORONIDA; TUNICATA.

Origins of animals. Some features of the cells of primitive animals resemble those of the single-celled Protozoa, especially the flagellates, which have long been believed to be animal ancestors. Molecular phylogenies have supported this idea and also suggest that the phylum Coelenterata arose separately from all other phyla that have been studied by this technique. Thus animals may have evolved at least twice from organisms that are not themselves animals, and represent a grade of evolution and not a single branch (clade) of the tree of life. Sponges have also been suspected of an independent origin, and it is possible

that some of the extinct fossil phyla arose independently or branched from sponges or cnidarians. *See* PROTOZOA.

The earliest undoubted animal fossils (the Ediacaran fauna) are soft-bodied, and first appear in marine sediments nearly 650 million years (m.y.) old. This fauna lasted about 50 m.y. and consisted chiefly of cnidarians or cnidarian-grade forms, though it contains a few enigmatic fossils that may represent groups that gave rise to more advanced phyla. Then, nearly 570 m.y. ago, just before and during earliest Cambrian time, a diversification of body architecture began that produced most of the living phyla as well as many extinct groups. The body plans of some of these groups involved mineralized skeletons which, as these are more easily preserved than soft tissues, created for the first time an extensive fossil record. The earliest mineralized skeletons are chiefly minute, the small shelly fauna of earliest Cambrian age; and while some of these fossils are primitive members of living phyla, many—perhaps 10 to 20 kinds of them—are distinctive and quite possibly represent phyla or at least major branches of phyla that soon became extinct, geologically speaking. The soft-bodied groups were also markedly diversified, though their record is so spotty that their history cannot be traced in detail. A single, exceptionally preserved soft-bodied fauna from the Burgess Shale of British Columbia that is about 530 m.y. old contains not only living soft-bodied worm phyla, but extinct groups (perhaps a dozen) that cannot be placed in living phyla and do not seem to be ancestral to them. Among these is *Wiwaxia,* a novel form which bears spinelike elements similar to some of the small shelly fossils of Early Cambrian age, thus providing evidence that the small shelly forms did indeed include novel types. Clearly, an early radiation of animals produced a vast array of novel body plans in an evolutionary episode that has never been surpassed. *See* BURGESS SHALE; CAMBRIAN; EDIACARAN BIOTA; FOSSIL.

History of major groups. Following the early phase of rampant diversification and of some concurrent extinction of phyla and their major branches, the subsequent history of the durably skeletonized groups can be followed in a general way in the marine fossil record (**Fig. 3**). The composition of the fauna changed continually, but three major associations can be seen: one dominated by the arthropodlike trilobites during the early Paleozoic (Fig. 3*a*), one dominated by articulate brachiopods and crinoids (Echinodermata) in the remaining Paleozoic (Fig. 3*b*), and one dominated by gastropod (snail) and bivalve (clam) mollusks during the

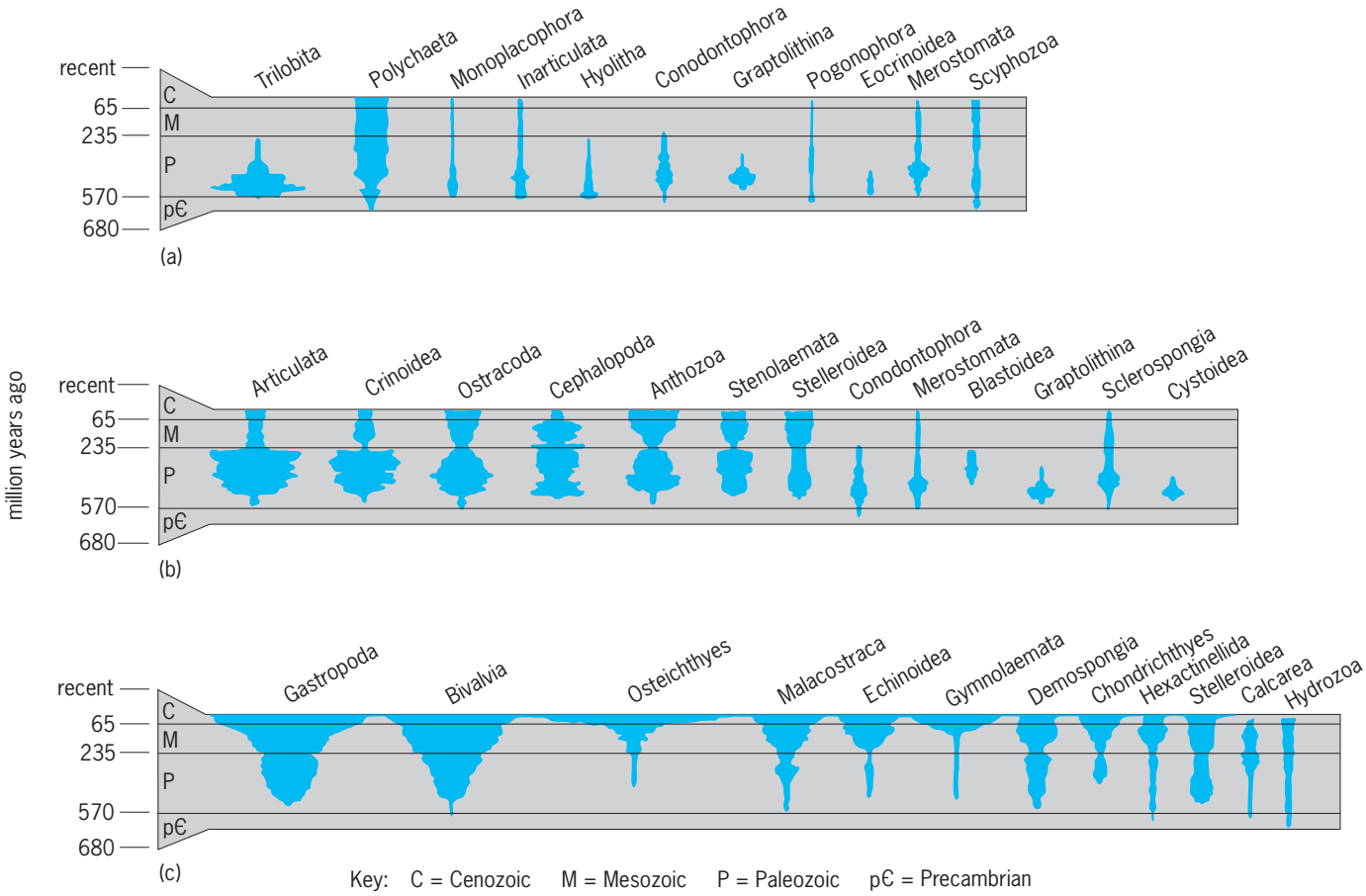

Fig. 3. Marine fossil record of major groups during the last 680 million years. The width of the spindles is proportional to the number of families present. The shifting of dominance from (*a*) trilobites and other forms during the early Paleozoic to (*b*) articulate brachiopods, crinoids, and others in the late Paleozoic and then to (*c*) gastropod and bivalve mollusks can be seen. (*After J. J. Sepkoski, Jr., in J. W. Valentine, ed., Phanerozoic Diversity Patterns: Profiles in Macroevolution, Princeton University Press, 1985*)

Mesozoic and Cenozoic (Fig. 3*c*). The mass extinction at the close of the Paleozoic that caused the contractions in so many groups (Fig. 3) may have extirpated over 90% of marine species and led to a reorganization of marine community structure and composition into a modern mode. Resistance to this and other extinctions seems to have been a major factor in the rise of successive groups to dominance. On land, the fossil record is too spotty to trace in detail the invasions of the terrestrial environment, which must have been under way by the mid-Paleozoic. Annelids, arthropods, and mollusks are the more important invertebrate groups that made the transition to land. The outstanding feature of terrestrial fauna is the importance of the insects, which appeared in the late Paleozoic and later radiated to produce the several million living species, surpassing all other life forms combined in this respect. *See* ANNELIDA; ARTHROPODA; CENOZOIC; INSECTA; MESOZOIC; MOLLUSCA; PALEOZOIC.

James W. Valentine

Chordate Origins

The phylum Chordata consists largely of animals with a backbone, the Vertebrata, including humans (**Fig. 4**). The group, however, includes some primitive nonvertebrates, the protochordates: lancelets, tunicates, acorn worms, pterobranchs, and possibly the extinct graptolites and conodonts. The interrelationships of these forms are not well understood. With the exception of the colonial graptolites, they are soft-bodied and have only a very limited fossil record. They suggest possible links to the Echinodermata in developmental, biochemical, and morphological features. In addition, some early Paleozoic fossils, the carpoids, have been classified alternatively as chordates and as echinoderms by various students, again suggesting a link. In spite of these various leads, the origin of the chordates remains basically unclear. *See* VERTEBRATA.

Chordates are characterized by a hollow, dorsal, axial nerve chord, a ventral heart, a system of slits in the pharynx that serves variously the functions of feeding and respiration, a postanal swimming tail, and a notochord that is an elongate supporting structure lying immediately below the nerve chord. The protochordates were segmented, although sessile forms such as the tunicates show this only in the swimming, larval phase. Even the free-swimming forms were not truly active, and they used the gill apparatus primarily for feeding on small particles

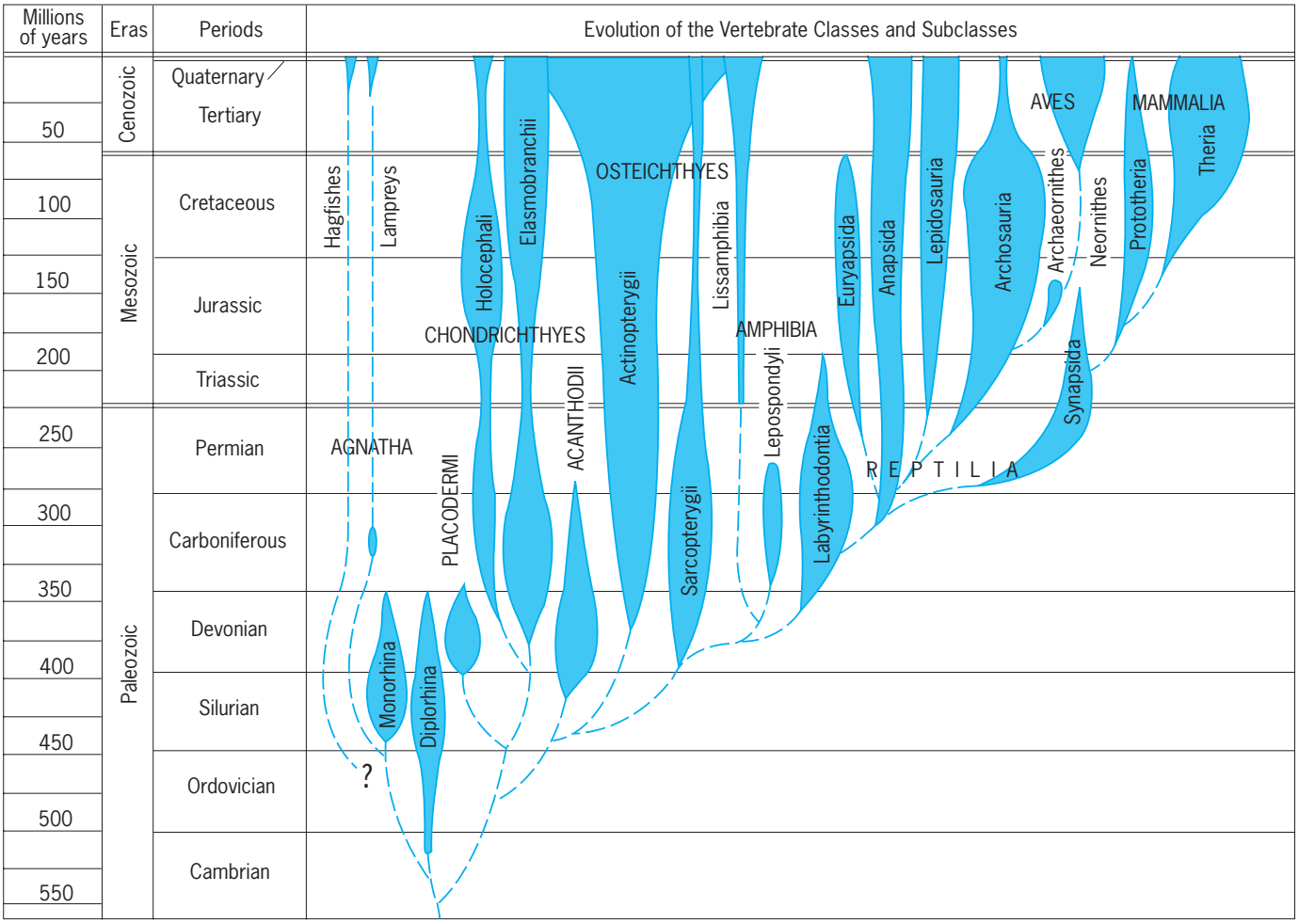

Fig. 4. Vertebrate evolution in relation to geologic time scale. Classes are shown in capital letters; subclasses are shown in lowercase letters, except that "Hagfishes" and "Lampreys" are common names for orders. (*After M. Hildebrand, Analysis of Vertebrate Structure, Wiley, 1982*)

trapped on mucous streams generated in the pharynx. There are two hypotheses concerning the nature of the protochordates. One holds that all protochordates were originally sessile, with, however, active larvae, and that the vertebrates arose from the larval phase through increased development of the segmented muscular trunk and associated nervous system. The alternative hypothesis proposes that the sessile condition seen in many protochordates is secondary.

The first vertebrates were fishlike animals in which the pharyngeal slits formed a series of pouches that functioned as respiratory gills. An anterior specialized mouth permitted ingestion of food items large in comparison with those of the filter-feeding protochordates. Vertebrates are first known from bone fragments found in rocks of Cambrian age, but more complete remains have come from the Middle Ordovician. Innovations, related to greater musculoskeletal activity, included the origin of a supporting skeleton of cartilage and bone, a larger brain, and three pairs of cranial sense organs (nose, eyes, and ears). At first the osseous skeleton served as protective scales in the skin, as a supplement to the notochord, and as a casing around the brain. In later vertebrates the adult notochord is largely or wholly replaced by bone, which encloses the nerve chord to form a true backbone. All vertebrates have a heart which pumps blood through capillaries, where exchanges of gases with the external media take place. The blood contains hemoglobin in special cells which carry oxygen and carbon dioxide. During the exchanges the oxygen content of the cells is increased and that of carbon dioxide decreased. In most fishes the blood passes from the heart to the gills and thence to the brain and other parts of the body. In most tetrapods, and in some fishes, blood passes to the lungs, is returned to the heart after oxygenation, and is then pumped to the various parts of the body.

Fish evolution. The jawless fish, known as Agnatha, had a sucking-rasping mouth apparatus rather than true jaws. They enjoyed great success from the Late Cambrian until the end of the Devonian. Most were heavily armored, although a few naked forms are known. They were weak swimmers and lived mostly on the bottom. The modern parasitic lampreys and deep-sea scavenging hagfish are the only surviving descendants of these early fish radiations. *See* DEVONIAN; JAWLESS VERTEBRATES.

In the Middle to Late Silurian arose a new type of vertebrate, the Gnathostomata, characterized by true jaws and teeth. They constitute the great majority of fishes and all tetrapod vertebrates. The jaws are modified elements of the front parts of the gill apparatus, and the teeth are modified bony scales from the skin of the mouth. With the development of jaws, a whole new set of ecological opportunities was open to the vertebrates. Along with this, new swimming patterns appeared, made possible by the origin of paired fins, forerunners of which occur in some agnathans. *See* GNATHOSTOMATA; SILURIAN.

Four groups of fishes quickly diversified (Fig. 2). Of these, the Placodermi and Acanthodii are extinct. The Placodermi were heavily armored fishes, the dominant marine carnivores of the Silurian and Devonian, rivaled only by the large eurypterid arthropods. The Acanthodii were filter-feeders mostly of small size. They are possibly related to the dominant groups of modern fishes, the largely cartilaginous Chondrichthyes (including sharks, rays, and chimaeras) and the Osteichthyes (the higher bony fishes). These also arose in the Late Silurian but diversified later. *See* ACANTHODII; CHONDRICHTHYES; OSTEICHTHYES; PLACODERMI.

Conquest of land. The first land vertebrates, the Amphibia, appeared in the Late Devonian and were derived from an early group of osteichthyans called lobe-finned fishes, of which two kinds survive today, the Dipnoi or lungfishes, and the crossopterygian coelacanth *Latimeria*. They were lung-breathing fishes that lived in shallow marine waters and in swamps and marshes. The first amphibians fed and reproduced in or near the water. True land vertebrates, Reptilia, with a modified (amniote) egg that could survive on land, probably arose in the Mississippian. *See* AMNIOTA; AMPHIBIA; DIPNOI; MISSISSIPPIAN.

Reptile radiations. By the Middle Pennsylvanian a massive radiation of reptiles was in process. In it can be traced the line leading to mammals as well as the lineages of the great reptiles that dominated the Mesozoic Era. The most prominent reptiles belong in the Diapsida: dinosaurs, lizards and snakes, and pterosaurs (flying reptiles). The birds, Aves, which diverged from the dinosaur radiation in the Late Triassic or Early Jurassic, are considered to be feathered dinosaurs, and thus members of the Diapsida, whereas older authorities prefer to treat them as a separate class. In addition, there were several Mesozoic radiations of marine reptiles such as ichthyosaurs and plesiosaurs. Turtles (Chelonia) first appeared in the Triassic and have been highly successful ever since. *See* AVES; DINOSAURIA; JURASSIC; PENNSYLVANIAN; REPTILIA.

Mammalian origins. The line leading to mammals can be traced to primitive Pennsylvanian reptiles, Synapsida, which diversified and spread worldwide during the Permian and Triassic. The first true mammals, based on characteristics of jaw, tooth, and ear structure, arose in the Late Triassic. Derived mammals, marsupials (Metatheria) and placentals (Eutheria), are known from the Late Cretaceous, but mammalian radiations began only in the early Cenozoic. By the end of the Eocene, all the major lines of modern mammals had become established. Mammals are easily separated into distinct groups (orders), but their relationships are not easy to discover from fossil records because of the explosion of mammalian evolution in the early Cenozoic. Molecular analyses (blood proteins, deoxyribonucleic acid, ribonucleic acid) of living mammals show that the most primitive group of placentals is the edentates (sloths, armadillos, and anteaters). An early large radiation included the rodents, primates (including monkeys, apes, and

humans), and bats, possibly all closely related to the insectivores and carnivores. The newest radiations of mammals are of elephants and sea cows, while the whales are related to the artiodactyls (cattle, camels).

Because of biogeographic isolation, marsupials came to flourish in Australia and South America, whereas the placentals diversified widely in Eurasia and North America. *See* CENOZOIC; CRETACEOUS; EOCENE; MAMMALIA; PERMIAN; SYNAPSIDA; TRIASSIC.

Primates and humans. Among the early and primitive lines of placental mammals were the Paleocene and Eocene plesiadapoids from which the primates arose. They were mostly nocturnal, insect- and fruit-eating animals with forwardly directed eyes and a locomotor system well developed for an arboreal life. The living primates consist of prosimians; tarsiers; lemurs; monkeys and anthropoids; apes; and humans and near relatives, the hominids. During the Oligocene, monkeys diverged into Old and New World lines. From the former, the apes arose in the late Oligocene. Known as the Pongidae, they are represented today by gibbons, orangutans, chimpanzees, and gorillas. *See* APES; OLIGOCENE.

The Hominidae, which include the human subspecies *Homo sapiens sapiens*, diverged from the apes about 5 million years ago. The closest living relatives among the apes are chimpanzees and gorillas. The earliest known hominids consist of fossils of the genus *Australopithecus*, excavated from rocks formed about 3.75 million years ago in eastern Africa. Several distinct lines of *Australopithecus*, under various generic names, coexisted for a long period of time; and also, during the later part of their existence, with *H. habilis*, which lived from about 2 million to 1.75 million years ago. This early species of *Homo* was notable for the relatively large brain, and probably included the first tool user. It was superseded by *H. erectus*, a larger animal with a greatly increased relative brain size. About 1 million years ago, *H. erectus* spread from Africa to eastern and southern Asia. It too has been given several generic names, of which *Pithecanthropus* and *Sinanthropus* from eastern Asia are the most familiar. *See* AUSTRALOPITHECINE; FOSSIL HUMANS.

Skeletal remains anatomically similar to those of *H. sapiens* have been found in northern Africa in beds formed about 300,000 years ago. Truly modern forms appeared some 100,000 years ago. They penetrated Europe some 30,000 to 40,000 years ago and spread throughout the Earth, occupying the Old World first, penetrating to Australia, and about 12,000 to 15,000 years ago entering North America and spreading rapidly to South America. *See* PALEONTOLOGY.

Keith S. Thomson

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988; K. G. Field et al., Molecular phylogeny of the animal kingdom, *Science*, 239: 748–752, 1988; M. F. Glaessner, *The Dawn of Animal Life*, 1984; A. S. Romer, *Vertebrate Paleontology*, 3d ed., 1966; J. W. Valentine (ed.), *Phanerozoic Diversity Patterns: Profiles in Macroevolution*, 1985; H. B. Whittington, *The Burgess Shale*, 1985; J. Z. Young, *Life of the Vertebrates*, 3d ed., 1991.

Animal feeds

Substances suitable for the nutrition of animals. This article discusses the main category—livestock and poultry feed— then briefly describes the secondary area—pet foods.

Livestock and Poultry Feed

In formulating livestock and poultry diets, animal nutritionists select from a variety of bulk feed ingredients including forages, feed grains, protein concentrates, by-product feeds, and vitamin-mineral supplements. These are formulated into complete rations in modern feed manufacturing plants (see **illus.**) to provide a balance of energy, protein, vitamins, and minerals to meet the nutrient requirements of various classes of livestock and poultry. Drugs and other additives may also be needed to meet specific needs. Today most feed manufacturers use high-speed electronic computers and various linear programming techniques to formulate diets. The overall goal is to prepare the lowest-cost ration that will provide economical livestock and poultry products for the consumer while maximizing returns to the producer.

Forages. In a broad sense, forages consist of grasses and legumes fed to ruminants (cattle, sheep, goats) and horses. Because of symbiotic microbes (bacteria and protozoa) in the digestive tract of these animals, they are uniquely equipped to utilize nutrients in forages; thus they do not compete with humans for the same food source. *See* FORAGE CROPS.

Grazing. Pasture and rangeland areas are covered with forage that is harvested by the grazing animal. Pasture provides about one-third of the nutrients consumed by dairy cattle and about half of those used by beef cattle and sheep. Cattle and sheep on rangeland obtain more than half their feed from forage. Grazing quality forage has the advantages of providing more digestible protein per acre than most other feeds, eliminating harvesting and feeding costs, and reducing soil erosion that results from the action of water and wind. While a portion of the forage is returned to the land as manure, appreciable forage may be lost due to trampling and selection by the grazing animals. To preclude such loss, forages are often mechanically harvested and fed as hay, silage, or dehydrated material. Although this practice minimizes loss and permits forage to be harvested at the optimum stage of maturity and highest quality, additional labor and energy are required to operate the harvesting, processing, and feeding equipment. *See* GRASS CROPS.

Hay and silage. In the United States, cool-season grasses, such as tall fescue and orchard grass, grow most rapidly during the spring and fall when rainfall is adequate and temperatures are moderate. Conversely, warm-season species, such as bahia grass, bermuda grass, and digit grass, grow best during the summer. Furthermore, because of variable climatic conditions, forage species do not grow at a uniform rate throughout the growing season. Because of this uneven distribution in forage production, most farms produce more forage than can be consumed by the

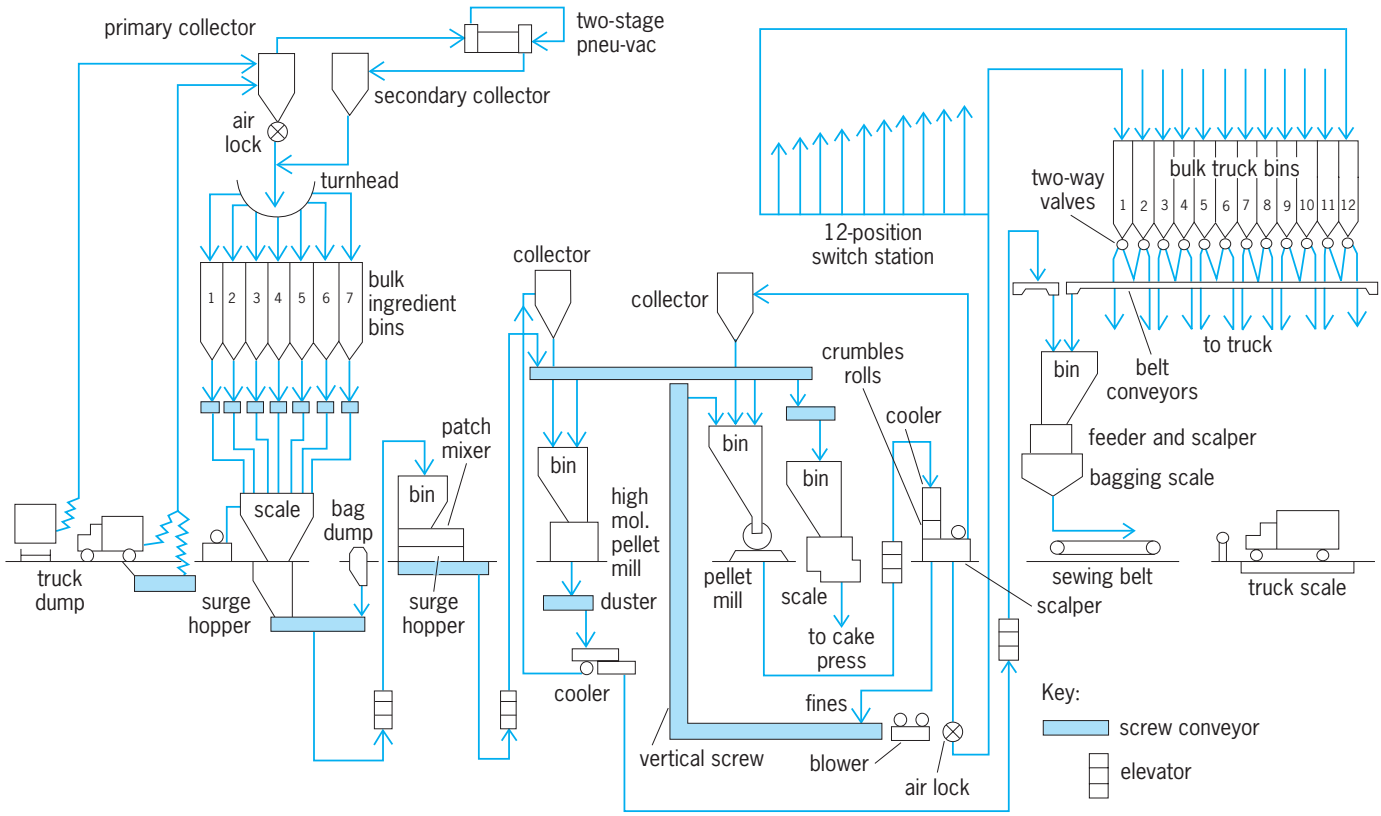

Flow diagram of formula feed plant. (*Sprout Waldron and Co., Inc.*)

livestock. This surplus forage is conserved as hay or silage and either fed or sold at later periods.

Hay is dried forage which has been cut and sun-cured in the field to about 10–20% moisture and then baled or otherwise packaged. It is commonly fed to ruminants and horses when grazing is inadequate or when the animals are confined. In the United States, alfalfa is the major hay crop, but other legumes, as well as annual and perennial grasses, are also extensively grown. The type of hay produced in a region depends primarily on those species which are best adapted to that area. For example, in the Northeast, timothy and alfalfa are important hay crops, while in the humid Southeast, warm-season perennial grasses such as Coastal bermuda grass and Pensacola bahia grass predominate.

Silage is fermented forage which is produced by the action of acetic and lactic acid–producing anaerobic bacteria in closed structures called silos. Although almost any forage crop can be preserved as silage, corn ranks first in the United States because good-quality, chopped, whole corn plant ensiles well and produces more digestible nutrients per acre than any other forage. Other silage consists of sorghum, grass, legumes, small grains, and miscellaneous crops. *See* CLOVER; LEGUME FORAGES.

Dehydration. Preserving forage by dehydration is relatively new compared to hay-making or ensiling. Although this industry is based largely on alfalfa, other crops, including Coastal bermuda grass, Johnson grass, and whole corn plant are also dehydrated.

In this process, freshly harvested or field-wilted forage is chopped and fed directly into the combustion gases (1600–2000°F or 870–1090°C) inside rotating horizontal drum dryers. These may be single- or triple-pass and range 8–12 ft (2.4–3.7 m) in diameter and 24–65 ft (7–20 m) in length. The evaporative capacity ranges 3–30 tons (2.7–27 metric tons) of water per hour. During drying, moisture diffuses to the surface of the forage particles as fast as it is evaporated. This rapid evaporation of surface moisture keeps the plant material cool enough to prevent burning. After remaining in the drum for about 2–10 min, the dehydrated forage is separated from the moisture-laden exit gases (250–350°F or 120–180°C) by means of a cyclone separator or other type of separator. Next, the forage is ground in a hammer-mill and pelleted for ease of storage, transportation, and feeding. To control dust, about 1.0% animal fat or vegetable oil may be added prior to grinding.

Mechanically dehydrated forages such as alfalfa and Coastal bermuda grass usually do not compete with hay or silage as roughages for ruminants. These products are high in protein, minerals, carotene (provitamin A), tocopherol (vitamin E), vitamin K, xanthophylls (substances that provide yellow color to poultry skin and egg yolks), and unidentified growth and reproduction factors. They are used primarily as supplements in swine and poultry rations and pet foods. The importance of carotene, xanthophylls, and vitamin E is such that producers of pelleted, dehydrated alfalfa or Coastal bermuda grass usually store the product under inert gas or add an

antioxidant (ethoxyquin) to minimize oxidation and ensure a high-quality product.

Feed grains. Feed grains are any of several grains commonly used as livestock and poultry feeds; they include corn, barley, oats, and grain sorghum (milo). Although wheat is second only to corn as a cereal grain in the United States, wheat is not usually fed to livestock. Most of the wheat crop is used for the manufacture of flour and other human foods. However, when properly used, wheat is a satisfactory feed for all classes of livestock. Accordingly, when the price of feed grains is high and surplus wheat is available at low price, considerable amounts may be fed to animals. *See* BARLEY; CORN; OATS; SORGHUM; WHEAT.

The purpose of processing feed grains is to improve their palatability, digestibility, and overall feed efficiency. For many years, stockmen sought to do this by grinding, crushing, rolling, and soaking feed grains, but since the late 1950s several new processing techniques have been used. One such technique is to ensile high-moisture (32%) ground ear corn. This has been shown to result in 10–15% greater utilization by cattle compared to regular ground ear corn. Additional techniques for improving grains by processing include reconstitution (moistening), flaking, roasting, micronizing, popping, and exploding.

Reconstitution. The process in which water is added to grain to bring its moisture content up to 20–30% is known as reconstitution. Usually grain so treated is held in limited-oxygen storage for a minimum of 3 weeks. Generally, cattle fed reconstituted corn or milo gained 8–15% more weight than those fed regular grain; and the feed efficiency of the reconstituted grain is increased by 7–11%.

Flaking. In a modification of steam rolling, the grain is cooked or steamed at approximately 200°F (93°C) at atmospheric pressure for 15 to 30 min, or at higher temperature and pressure, approximately 300°F (150°C) and 50 lb/in.^{2} (345 kilopascals) for 1 to 2 min, and then is rolled into flakes $^1/_{\sqrt{32}}$ in. (0.8 mm) thick. Because this process increases the moisture content to about 15–20%, the rolled flakes are immediately dried to 15% moisture. Flaking was the first modern process whereby the rate of gain and feed efficiency of milo were markedly increased. For beef cattle, gain and feed efficiency increased up to about 11% and 15%, respectively.

Roasting. In roasting corn, the grain is heated to 300°F (150°C). This results in expansion of the kernel, as evidenced by the fact that the roasted material weighs 39 lb/ft^3 (625 kg/m^3) compared to 45 lb/ft^3 (720 kg/m^3) for the raw corn material. Moisture content also is decreased about 2–4%. For fattening cattle, roasted corn improves weight gain by about 11% and has a feed efficiency about 14% greater than that of unroasted corn.

Micronizing. Another method used in the feed industry involves heating grain to 300°F (150°C) with microwaves. This micronized grain contains about 7% moisture. It is then rolled to form a uniform, free-flowing product. Prior to feeding, water is usually added to adjust the moisture content to 10%. Cattle fed micronized grain sorghum (milo) compare favorably with those fed steam-flaked grain sorghum in rate of gain and feed efficiency. However, the cost of processing favors the micronizing technique over the steam-flaking method.

Popping. This method involves rapidly heating grain containing about 15–20% moisture to 300–310°F (150–155°C). At this temperature, the internal moisture of the grain is volatilized and explodes the kernel. The exploded product is then dry-rolled to form a uniform product and reduce storage space. All grains can be processed by this method; however, it appears to be particularly effective in processing milo. Feeding trials conducted in Texas revealed that cattle consumed less of the popped milo than the regular grain. Although this resulted in lower average daily gains, feed efficiency was increased 17%.

Explosion puffing. In another method, explosion puffing, live steam is injected into high-tensile-strength steel bottles which hold about 200 lb (90 kg) of grain. About 20 s after the pressure reaches 250 lb/in.^{2} (1.7 megapascals) a valve is opened. Rapid release of the pressurized moisture within the grain explosively puffs the kernels. In practice, tandem vessels are alternately filled and pressurized to achieve a continuous operation. The resulting product resembles puffed breakfast cereal. Cattle fed puffed milo have exhibited feed intake, gain, and feed efficiency similar to that of cattle fed flaked grain.

Protein supplements. Protein supplements are feeds which contain more than 20% protein or protein equivalent. They are primarily added to rations containing feed grains or forages which are low in protein. Addition of protein supplements depends upon the animal's requirement for protein and the amount and quality of grain and forage in the ration.

Plant sources. High-protein meals can be made from soybean, cottonseed, sunflower seed, safflower seed, rapeseed, peanut, linseed, and coconut (copra). These meals are by-products of the oilseed-crushing industry; they vary in feeding value (that is, protein quantity and quality) depending upon the amount of hull or seedcoat remaining in the meal, conditions of time and temperature during extraction, and constituent amino acids of the protein. For example, although soybean meal and sunflower meal contain about the same amount of protein (45.8% versus 46.8%, respectively), they differ in amino acid content. Compared to sunflower meal, soybean meal is high in lysine (6.6 versus 3.8 g/16 g nitrogen) and low in methionine (1.1 versus 2.1 g/16 g nitrogen). In the United States, soybean meal is the most widely used protein supplement.

Quality forage and other leafy plants contain considerable protein (12–30%). The amino acid content of this protein compares favorably to that of soybean meal, fishmeal, and other high-quality proteins. Considerable research has been conducted in the United States, England, and Pakistan to investigate the preparation of leaf protein concentrates from forages. In the United States, the Pro-Xan process was developed for concentrating protein from alfalfa. The final product constitutes 4.5% of the starting material and

contains 40–50% protein and 600–900 mg/kg xanthophyll. Feeding trials with chicks have indicated that the product is a good source of protein and xanthophyll. If the cost of processing can be lowered, leaf protein concentrate may contribute significantly as an alternate protein source in the production of animal protein—meat, milk, and eggs—for humans.

Animal sources. Protein supplements of animal origin primarily consist of inedible products derived from meat packing or rendering plants, marine sources, poultry processing plants, and dairies. In general, such products are high in protein and of excellent quality, being well balanced in amino acids and providing vitamins and minerals. Feather meal, a by-product of the poultry processing industry, has also been used as a protein source. Although this product contains about 85% protein, it is poorly digested and must be hydrolyzed for good utilization. Animal-derived protein supplements are subject to autoxidation and rancidity because of their high fat content, and are often a source of bacterial (particularly *Salmonella*) contamination of poultry and meat products used for human food. It is therefore important that proper sanitation be practiced during processing and handling of these protein supplements to prevent bacterial contamination of human food.

Single-cell sources. There has been considerable interest in the production of protein from single-cell organisms such as yeasts, bacteria, and algae. A wide variety of carbon sources can be used as substrate for the growth of these organisms, including petroleum, methane, alcohol, starch, molasses, cellulose, and paper pulp waste liquor. Potential yields are quite high. One thousand pounds (450 kg) of single-cell organisms might produce 50 tons (45 metric tons) of protein per day. Although production of single-cell protein as a feedstuff for livestock appears promising, problems relating to palatability, digestibility, nucleic acid content, toxins, protein quality, and economics must be solved before this becomes a reality.

Microbial synthesis. Cattle and other ruminants are able to derive a portion of their dietary protein from nonprotein nitrogen via microbial protein synthesis in the rumen. Amino acids, amides, ammonium salts, and other nonprotein nitrogen sources provide nitrogen from which symbiotic rumen bacteria synthesize microbial protein. This microbial protein is then digested by the ruminant animal in the abomasum and gastrointestinal tract. Providing that adequate energy and other nutrients are available, about one-fourth to one-third of the protein requirements of ruminants can be met by the feeding of nonprotein nitrogen. If this amount is exceeded and the animals are not managed properly, toxicity and death may result. Commercially available products include urea and the ammoniated products of molasses, citrus pulp, beet pulp, cottonseed meal, and rice hulls. Both solid (range cubes and pellets) and liquid (lick tanks) products containing urea, molasses, trace minerals, vitamin A, and so forth, are available. Slow-release urea products also are marketed. With concern about world food supply and competition between humans and animals for available protein, the feeding of nonprotein nitrogen to ruminants is increasingly important. *See* PROTEIN.

Vitamins. Vitamins are complex organic compounds which are required in exceedingly small amounts by one or more animal species. Most function as cofactors in enzyme systems involving the regulation of metabolism and transformation of energy. Certain vitamins (such as vitamin K) are apparently needed by only a few species; others (such as vitamin A and thiamine) are required by all animals. Some vitamins (such as vitamin A) must be present in the diet; others are synthesized in the tissues of the animals (such as vitamin C) or by bacteria in the digestive tract (such as B-complex vitamins). The absence in the diet of one or more of the required vitamins may prevent the animal from growing or reproducing, or may cause deficiency diseases. If the deficiency is severe, the animals may die. Farm animals fed diets containing high-quality green forage usually receive enough of most required vitamins. Because monogastric animals (poultry, swine) thrive when small amounts of leaf meal are added to their diets, alfalfa and other forages are usually described as containing unidentified growth factors.

High temperatures and oxidation destroy certain vitamins; therefore, care must be taken in the processing and storage of feedstuffs to protect the vitamins they contain. Certain feedstuffs also may contain antivitamins, which prevent the proper assimilation and functioning of vitamins. To ensure that rations contain sufficient vitamins to meet dietary requirements, manufacturers often fortify feed formulations with one or more vitamins. *See* VITAMIN.

Minerals. Animals, like plants, require minerals to grow, and if deficiencies occur, livestock become unthrifty, lose weight, or exhibit other deficiency symptoms. Minerals serve many vital functions. The skeletons of vertebrates are composed chiefly of calcium and phosphorus. Not only must livestock receive adequate calcium and phosphorus, but they must also receive these minerals in the proper proportion, because a great excess of one or the other may be detrimental. Also, if the proper calcium-phosphorus ratio is provided, less vitamin D is needed. Minerals also are necessary constituents of soft tissues and body fluids, and they function in maintaining the proper acid-base balance of body tissues. In addition to macroelements (for instance, calcium, phosphorus, sodium, chlorine, and iron), trace elements (zinc, cobalt, manganese, and selenium, among others) are required for proper nutrition. Trace elements primarily function as cofactors in metabolic reactions. Feedstuffs produced in particular areas may be deficient in or have an excess of certain minerals. For example, in the Great Lakes and Northwest regions of the United States, soils are deficient in iodine. Also, soils containing more than 0.5 ppm (part per million) of selenium are potentially dangerous for production of livestock feed. The form and availability of certain minerals also are an important consideration in formulating livestock diets. Phytin phosphorus,

for example, is not a good source of phosphorus for poultry or swine; however, it appears to be satisfactory for ruminants. To ensure an adequate supply of minerals, feed manufacturers often add mineral supplements. Farm animals may also be permitted free access to a mineral mixture. If this is allowed, care should be taken, particularly with trace minerals, to see that the animals do not consume too much, or toxicity may result. Like vitamins, use of mineral supplements adds to production costs. Therefore, the need for such supplements should be given careful consideration.

By-products. This group of feeds includes a number of by-products formed in the processing of livestock, plants, and other materials. To be acceptable as livestock feeds, such products should be palatable, provide digestible nutrients or roughage, and present no health hazard. Use of by-products as feed not only lowers the cost of animal production but helps to reduce the competition between livestock and humans for available grain and protein supplies. Another advantage is that industry derives income from products that otherwise might be wasted and disposed of at a cost. By-product feeds include inedible products and residues from animal, poultry, and marine processing plants, fruit and vegetable processing plants (citrus pulp, tomato pomace), the brewing and distilling industry (brewers' and distillers' grains), dairy processing plants (whey, cheesemeal), wood and paper mills (wood molasses, treated wood scraps), as well as miscellaneous fibrous materials such as cottonseed gin trash, cottonseed, soybean, and peanut hulls. Research has been conducted on the feeding of animal wastes (manure and litter) to livestock. There is also interest in the treatment of crop residues (cornstalks, straw, and so forth) with alkali (sodium hydroxide) to improve their digestibility for ruminants. Although certain by-products (such as peanut hulls) are available only in a limited area, others (such as brewers' grains) are more widely available.

Feed additives and implants. These may be defined as nonnutritive substances which enhance the utilization of a feed or productive performance of the animal. It is estimated that in the United States 75% of the finishing cattle, finishing lambs, and growing-finishing pigs receive feed additives or implants. Use of these products and guidelines for administration and withdrawal of approved products are regulated and prescribed by the U.S. Food and Drug Administration (FDA). Also, feed-control officials in individual states may regulate use of these substances for livestock marketed intrastate. Prior to receiving approval for marketing a product, the manufacturer must prove its safety and efficacy through a rigorous testing program. Since the early 1950s, antibiotics have been used to stimulate growth and promote feed efficiency in ruminants, swine, and poultry. However, there is concern that continuous feeding of antibiotics to livestock may increase the resistance of certain strains of bacteria and that such resistance may be transferred to bacteria associated with human diseases. It is therefore expected that the use of such antibiotics as penicillin and the tetracyclines in feeds may come under increased scrutiny.

In 1954 the hormone diethylstilbesterol (DES) was approved for use in cattle finishing rations. Two years later, implants for steers were approved. In 1973 the use of DES was banned when it was proved to be carcinogenic and present in trace amounts in the meat of animals. However, in 1974 a Federal Court of Appeals overruled this ban. On November 1, 1979, the 1974 decision was reversed, and DES may no longer be used as a feed additive or implant for cattle and sheep.

Another hormone, melengestrol acetate (MGA), is marketed for feeding to heifers being finished for slaughter. This compound, a synthetic progestin, affects metabolic rate and suppresses estrus, thereby increasing feed efficiency of feedlot heifers.

Other hormones or hormonelike compounds which are marketed for implantation in beef cattle to improve growth and feed efficiency are Synovex and Ralgro. Synovex-S contains progesterone and estradiol benzoate, and is used for steers; Synovex-H contains testosterone and estradiol benzoate, and is used for heifers. Although both hormones are FDA-approved, they are being evaluated to determine the advisability of withdrawing approval of these products in consideration of human safety. Ralgro is the trade name for zeranol, a substance produced by the mold *Giberella zeae*. Although not a hormone itself, this compound is thought to influence the production and release of certain hormones in the body. It is approved for implantation in beef cattle.

In fermenting feeds, rumen microorganisms produce carbon dioxide and methane in addition to fatty acids. Methane accounts for about 10% of the total energy of feedstuffs; hence its production reduces the energy efficiency of feeds. Therefore, compounds which may inhibit methane production in the rumen are being investigated. Dichlorovinyldimethyl phosphate, hemiacetal derivatives of chloral and starch, and halogenated analogs of methane have shown promise.

A compound which has proved effective in altering rumen metabolism is monensin, a fermentation product of *Streptomyces cinnamonensis*. The proprietary product, Rumensin, has been shown to significantly improve feed efficiency in feedlot cattle. This compound acts by altering the proportion of fatty acids produced in the rumen; more propionic acid (and less methane) is produced than acetic and butyric acids. Because conversion of feed to propionic acid is energetically more efficient than its conversion to the other acids, overall production efficiency increases. *See* NUTRITION.

International feed nomenclature. Because feed names have evolved through usage, the name of a particular feedstuff has not always been descriptive enough to fully identify and characterize it. For example, the term Linseed meal does not convey whether the product is obtained by mechanical or solvent extraction or whether it contains 33%, 35%, or 37% crude protein. To overcome this deficiency, a

system was developed that makes it possible to identify the contents and other characteristics of a feed from its name. This system, known as the National Research Council (NRC) Feed Nomenclature System, or International Feed Nomenclature System, gives to each feed a specific name that may consist of up to nine terms: scientific name, origin (or parent material), species (variety or kind), part actually eaten, process(es) and treatment(s) to which the parent material or the part eaten was subject before being fed to the animal, stage of maturity (applicable only to forages), cutting or crop, grade (quality designations and guarantees), and classification (according to nutritional characteristics). In addition, each feed is grouped into one of eight feed classes, each of which is designated by a number in parenthesis: (1) dry forage or dry roughage; (2) pasture, range plant, and forages fed green; (3) silages; (4) energy feeds; (5) protein supplements; (6) minerals; (7) vitamins; (8) additives; then each feed is given an international reference number. The classification number, for example, (1), is the last term in the international feed name and the first digit of the feed reference number. An example of an international feed name for dehydrated alfalfa pellets is: alfalfa, aerial part, dehydrated ground pelleted, early bloom, cut 2, (1). The international reference number for this feed would be 1-07-733. This six-digit reference number is the numerical name of the particular feedstuff and may be used when formulating rations using linear programming and computers. Only the first digit is related to a specific feed characteristic. The remaining digits are used only to number the feed. This international system of naming feeds is being widely adopted for use in scientific writing and journal articles. International reference numbers particularly are being used to identify feeds in tables and computer formulations. Donald Burdick

Pet Foods

Pet foods are usually nutritionally complete products which adequately support growth and life, even if fed exclusively.

Pet foods are made from proper mixtures of meat, fish, fowl, bone, cereal, milk, minerals, and vitamins. The minerals and vitamins are added in quantities necessary to augment the natural foodstuffs. They may be prepared in various finished forms, canned, frozen, or dried.

The canned products are sterilized at high temperatures to permit storage without refrigeration. The frozen variety often is limited to raw meat or seafood. Dry items are stable at room temperature, though usually not as long as canned items.

Those pet foods which include meat contain a fairly large proportion of highly nutritious organ meats as well as other meat products.

The manufacture of the canned variety is fairly standardized in that the grinding, blending, filling, and sterilizing procedures are well known. The dry variety, particularly dry dog food, can be manufactured in a number of different ways, ranging from pelletizing of dry ingredients to hot-air and vacuum drying of moistened, expanded pastes and extrusions. Little or no preparation is needed for the frozen products since these are generally raw.

A significant innovation is intermediate-moisture pet food that keeps without canning or refrigeration. Careful control of amount of "active water" in the product avoids spoilage.

Aseptic canning of certain pet foods appears entirely possible. In this procedure the product is sterilized outside the can and then filled into a sterilized container, whereupon it requires no further cooking and is stable at room temperature for indefinite periods. Clarence K. Wiesman

Bibliography. P. R. Cheeke, *Applied Animal Nutrition*: *Feeds and Feeding*, 2d ed., 1998; D. C. Church, *Livestock Feeds and Feeding*, 4th ed., 1997; A. Cullison and R. S. Lowrey, *Feeds and Feeding*, 4th ed., 1986; R. M. Ensminger, *The Stockman's Handbook*, 7th ed., 1992; M. E. Ensminger, J. E. Oldfield, and W. W. Heinemann, *Feeds and Nutrition*, 2d ed., 1998; D. Natz (ed.), *Feed Additive Compendium*, 1977; W. G. Pond et al., *Basic Animal Nutrition and Feeding*, 1995.

Animal growth

The increase in size or weight of an animal. A number of processes are involved in growth. These include hyperplasia, or increase in cell number due to cell proliferation (cell division or mitosis) and recruitment from stem cell populations; hypertrophy, or increase in the size of cells; and differentiation (precursor cells achieving mature functioning—often this process cannot be reversed). Schematic diagrams showing muscle and adipose development are included below. *See* ANIMAL MORPHOGENESIS; CELL DIFFERENTIATION; CELL DIVISION; CELL SENESCENCE AND DEATH; EMBRYONIC DIFFERENTIATION; MITOSIS.

Measurement of growth. The indices of growth include height (as for humans and horses), body weight (agricultural animals), dry weight, length (rodents, lizards, fish), total body DNA or nitrogen retention (used in agricultural animals as an indicator of muscle growth), or rates of whole-body protein synthesis and/or accretion:

Protein accretion (net protein synthesis)
= Protein synthesis − Protein degradation

Mathematical models for growth. Growth appears to follow an S or sigmoidal curve. There are multiple equations that mathematically describe growth. These include the logistic model and the Gompertz equation

$$W = \int A \cdot e^{-b \cdot e-kt}$$

$$dW/dt = k \cdot W \cdot (A/W)$$

The Gompertz equation describes growth of an animal within a specific species very well from embryonic stage to mature adult size. Here W is the weight at any given age (time t following the beginning of

development), A is the maximal weight (mature or adult weight), k is the growth rate or constant, and b is a second constant.

At significant periods of growth, a linear function can easily be applied for growth.

Telomeres. Telomeres are multiple TTAGGG tandem repeats at the end of chromosomes in at least higher vertebrates. The telomeres stablize the chromosomal ends and protect the genes in the subtelomeric regions. Telomeres shorten with progressive cell cycles (either in the animal or in tissue culture), leading to the view that shortened telomeres are related to senescence or aging. Addition of telomeric repeats can be accomplished by the enzyme telomerase, an RNA-dependent DNA polymerase. High levels of telomerase activity are observed in tissues with a high rate of cell proliferation, for example, malignant tissues and fetal tissues. Given the number of cell divisions, it is not surprising that cloned animals resulting from nuclear transfer from somatic cells have reduced telemetric length, with corresponding effects on growth rate. *See* CHROMOSOME.

Embryonic and fetal development. Development begins following fertilization of the ovum (egg) by the spermatozoon. There is cell division to form first a two-cell stage, then four-cell, then eight-cell stage. Cell divisions continue but become desynchronized.

Prior to implantation, in mammals, the part of the conceptus that will develop into the animal (as opposed to the membranes and placenta) is called the embryo. At the end of embryonic development, the major organs are normally formed and "true" bone formation begins (replacing cartilage). The fetus develops in the uterus until birth. Considerable growth and development occurs with organs achieving the capability of full functioning.

In mammals, the placenta plays an important role influencing fetal growth. The availability of critical nutrients and oxygen to the growing fetus is related to placental size and functioning. In addition, the placenta can produce hormones that influence fetal growth. Other critically important hormones affecting fetal growth are insulin and insulinlike growth factor-II (IGF-II). In birds, the yolky egg provides the nutrients for the developing embryo. In amphibians, there is a free-living larval (for example, "tadpole") stage that metamorphoses into the adult form as stimulated by the hormone thyroxine. *See* EGG (FOWL); HORMONE; PLACENTATION.

Birth and postnatal growth and development. At birth, all animals have their organs formed. However, animals can be born (mammals) or hatched (birds) either with their systems well functioning (precocial development) or with the systems not completely functional (altricial development). For instance, horses can walk very soon after birth, where as cats and rats have offspring that are blind at birth and rats have hairless offspring.

Following birth, the vast majority of growth (increase in weight) occurs. There is not complete synchrony of growth with different organ systems. First, the rate of skeletal growth is maximal, then comes maximal rate of muscular growth, followed by maximal rate of adipose growth. Growth in most mammals ceases shortly after puberty (as discussed below). At puberty, there is an increase in sex steroid hormones in the circulation. The hormones induce a "growth spurt" in many species, including humans. Another result is the closure of the growth plate and the cessation of growth (discussed below).

Muscle growth. Muscle development and growth involves the massive cell division (proliferation) of the myogenic precursor cells. These cells migrate to specific sites in the body, for instance, the developing limbs. They are attracted to the sites by gradients of growth factors and are undergoing some differentiation. The precursor cells then form myoblast cells, which are the last stage of cells that can undergo division. The myoblasts differentiate into myocytes. Multiple myocytes then aggregate to form multinucleated myotubes, which start to produce the contractile proteins and then fibers of muscle. Postnatally, muscle growth is accomplished by increasing the size of the muscle fibers. More nuclei are added by the recruitment of satellite cells. Muscle growth and development may be summarized schematically:

Mesodermal somites
↓
Myogenic precursor cells (→ Satellite cells)
↓
Proliferating, differentiating, and migrating
(embryonic and fetal life)
↓
Myoblast cells
↓
Proliferating and differentiating myocyte cells
(capable of synthesizing many muscle proteins)
↓ Differentiating
Myotubes
(multinucleated cells that are assembling
muscle contractile fibers)
↓
Myofibers (multinucleated)
↓
Muscles growing in size due to hypertrophy
as well as recruitment of satellite cells

See MUSCLE.

Bone and cartilage growth. Longitudinal growth or lengthening of the long bones of the skeleton occurs at the epiphyseal or growth plate. Growth of this cartilage (including proliferation of cartilage cells or chondrocytes) occurs under the influence of the hormone insulinlike growth factor (IGF-I) (discussed below). At puberty or sexual maturation in most higher vertebrates, there is closing or ossification of epiphyseal plates, and longitudinal growth of bones ceases. This occurs under the influence of sex steroid hormones. *See* BONE; CARTILAGE.

Adipose tissue growth. Adipose tissue provides energy storage as triglyceride. This is critically important to animals at a time of nutritional deprivation or during hibernation or when energy consumption is not meeting their needs (for example, during pregnancy and lactation). Unlike other organs, adipose

tissue can continue to increase in size after sexual maturation. *See* ADIPOSE TISSUE; LIPID; TRIGLYCERIDE.

Adipose tissue grows by proliferation and migration of adipose precursor cells. These differentiate into preadipocytes, which are the last cells in the adiocyte lineage capable of division. They may divide multiple times to form adipocytes. Adipocytes do not divide but produce the key enzymes necessary for triglyceride or fat formation and they fill with triglyceride. Adipose tissue growth and development may be summarized schematically:

Adipose precursor cells
↓
Proliferating, differentiating, and migrating
(embryonic and fetal life)
↓
Preadipocyte cells
↓
Proliferating and differentiating adipose cells
(capable of synthesizing adipose proteins
and triglyceride/fat)
↓
Mature fat-filled adipose cells
(growing by increasing in size)

Growth hormone. Multiple hormones influence growth. A prime example is growth hormone (GH), produced by specific cells, the somatotropes, found in the anterior pituitary gland. The development of the somatotropes is induced by pit-1. Release of GH is controlled by peptides released from specialized nerve terminals in the hypothalamus. These peptides include GH releasing hormone (GHRH) which stimulates GH release and somatostatin (SRIF) which inhibits GH release. Growth hormone acts by binding to GH receptors (GHR) on the cell membrane of cells in the liver and some other tissues. In turn, GH increases the release of IGF-I. The IGF-I binds reversibly to IGF-binding proteins (IGF-BPs) in the circulation. Unbound IGF-I stimulates skeletal growth:

Hypothalamus (in brain)
↓ ↓
$GHRH^{+}$ $SRIF^{-}$
↓ ↓
Anterior pituitary gland (somatotropes)
↓
Growth hormone (GH)
↓
Liver and other tissues (GHR)
↓ ↓
IGF-I IGF-BPs
↓ ↓
Growing tissues, such as cartilage

Lack of sufficient production of GH and/or IGF-I is associated with dwarfism in, for example, humans, cattle, chickens, dogs, and mice. Excess GH production causes giantism and, in adult animals, acromegaly.

Other hormones influencing growth. Normal levels of thyroid hormones are required for growth. Glucocorticoid hormones, produced by the adrenal cortex [cortisol in humans and domestic mammals, and corticosterone in rodents and poultry and other birds], inhibit growth.

Manipulation of growth rate. Some chemical agents (related to hormones) are used to stimulate growth in animals. These include the use of GH in children with abnormally low growth rates (and likely low adult heights) due to hypopituitary dwarfism or low height for age. Muscle growth is stimulated in many but not all higher vertebrates by androgens. In commercial cattle, growth can be enhanced by the administration of estrogens, alone or with androgens. This treatment is approved for commercial use in some countries. There are beta-adrenergic agonists (for instance, Ractopamine) that enhance growth, particularly muscle growth. This treatment is used commercially in pigs in some countries.

Factors influencing growth rate. Growth rates within a species are influenced considerably by the environment and by the animal's genetics.

Growth may be reduced by lack of availability of nutrients, the presence of toxicants in the environment, environmental stressors (for example, high or low temperatures), and disease. Diets deficient in critical nutrients, for example, energy components, proteins (and even specific amino acids such as lysine, methionine, or tryptophan), vitamins, and minerals, will result in reduced growth rates. When full nutrition is restored, compensatory or catch-up growth occurs. Not only is the growth rate restored but it rebounds to a higher level, above that normally seen. One mechanism by which stressors reduce growth is via increased release of the glucocorticoid hormones, cortisol and corticosterone. Other mechanisms of growth reduction include suppression of IGF-I levels, for instance, by disease; and toxicants that influence the thyroid gland (goitrogens). *See* NUTRITION; PROTEIN.

Genetics. Genetics profoundly influences growth rate. There is marked sexual dimorphism in growth rates, with males of most species growing faster and larger than females. Different breeds of animals have been produced with markedly different growth rates and mature body sizes. Perhaps the best example is seen in the widely different breeds of dogs and even the disparate-sized poodles such as miniature and toy poodles. Selective breeding can greatly influence growth rate. An example is seen in chickens (meat or broiler type), where since 1945 there has been increased growth rates by more than fivefold with consequent improvements in efficiency and reduced cost to the consumer. More recently, geneticists working for the livestock industry have employed quantitative trait loci (QTL) following genome scans or specific point mutations (for instance, single-nucleotide polymorphorisms, or SNPs) in candidate genes to identify superior parents or grandparents, and then the geneticists have used these in select breeding programs. *See* BREEDING (ANIMAL); GENETICS.

Gene imprinting. Fetal growth rates are influenced by imprinting of genes. This is the result of methylation of DNA. The pattern of imprinting varies with the parent of origin of the genes. An example

is the imprinting of the IGF-II gene in eutherian mammals. There are distinct differences in imprinting sites, depending on maternal or paternal origin of the gene. There is also imprinting of the IGF-II receptor gene (responsible for degradation of IGF-II); this is found in most mammals but not in the primate lineage. Colin G. Scanes

Bibliography. S. Brody, *Bioenergetics and Growth*, Waverley Press, Baltimore, 1945; M. F. Rothschild, Porcine genomics delivers new tools and results: This little piggy did more than just go to market, *Genet. Res.*, 83:1–6, 2004; C. G. Scanes, *Biology of Growth of Domestic Animals*, Iowa State Press/Blackwell, 2003.

Animal kingdom

One of five kingdoms of organisms: Animalia, Plantae, Fungi, Protoctista (or, in part, in some classifications Protozoa), and Procaryotae (bacteria). Animals are eukaryotic multicellular organisms that take food into their bodies and that develop from blastula embryos. More than 30 million species of organisms have been described in all five kingdoms; the animal kingdom includes over 1 million described species. Animal species are organized into phyla (see **illus.**) that are defined according to comparative patterns of development, body structures, behavior, biochemical pathways, modes of nutrition, and ancestry. *See* ANIMAL SYSTEMATICS.

Form and habitat. Animals take diverse forms. The largest animals are great blue whales, which grow up to 100 ft (30 m) in length; the smallest are marine loriciferans, measuring as little as 300 micrometers. Thus, the largest animals are about 100,000 times as large as the smallest.

Some one-celled organisms are classified as members of the Protoctista kingdom. Most protoctists are microorganisms, but some are larger than animals. For example, rotifers, nematodes, and loriciferans are smaller than protoctists such as ciliates.

Most animal phyla are aquatic; that is, they are found in shallow seas, ponds, puddles, and streams. Some members of two phyla are called terrestrial because they live on the land surface. These phyla are Chordata, including humans and other mammals, reptiles, amphibians, fish, and birds; and Arthropoda, including moths, beetles, and millipedes. Animals that fly are referred to as aerial; insects (phylum Arthropoda), birds, bats, and flying dinosaurs are included in this category. Some animals inhabit a moist, underground environment but are not strictly terrestrial; annelid worms, nematodes, and other soil-dwelling animals are adapted to life in the thin film of water surrounding soil particles.

Diversity. There are approximately 1.2 million animal species, but new species and phyla are

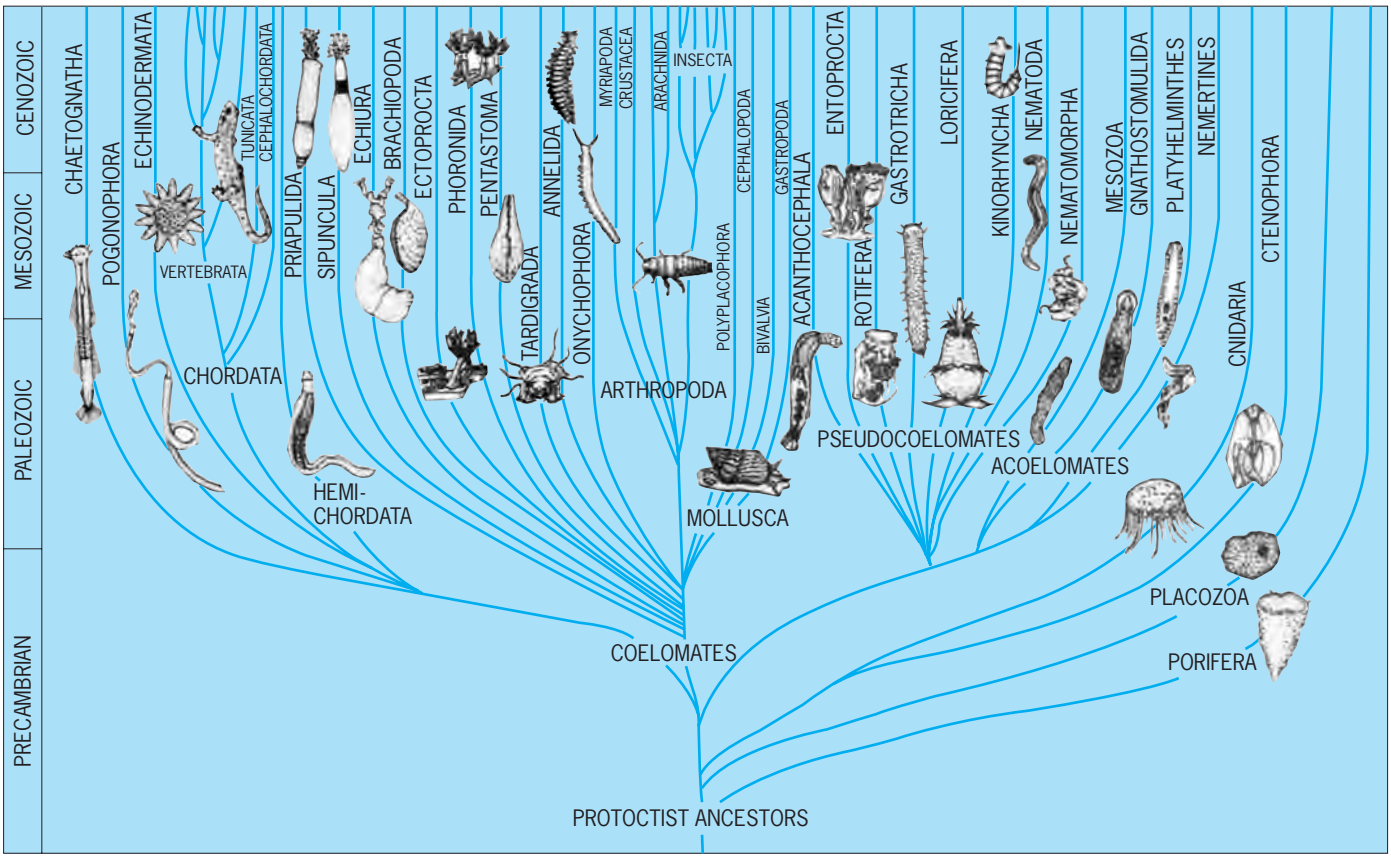

Animal phylogeny or family tree, diagramming the relationships among phyla of living animals. The ancestors of all present-day animals probably were soft-bodied marine forms, such as the placozoa. (*After L. Margulis and K. V. Schwartz, "The Animal Kingdom" Teachers' Guide, Wards Natural Science, Rochester, New York, 1987*)

constantly being identified. Of the known animal species, some 75% are arthropods. One phylum, Vestimentifera, includes giant tube worms that measure 5 ft (1.5 m) long. These animals of the abyss live in hot, submarine gardens some 1.5 mi (2.5 km) below the ocean surface. Another phylum, Loricifera, includes spiny marine animals that live in offshore sediments. Much remains to be learned about loriciferan food and development because these animals stick tenaciously to shelly gravel. Unfortunately, animals along with their habitats are being obliterated thousands of times faster than any of the earlier mass extinctions recorded in the geological record. Biologists estimate that about 99.9% of all animal species that ever lived are now extinct. *See* EXTINCTION (BIOLOGY).

Nutrition and development. Unlike plants, seaweeds, and some bacteria which photosynthesize (autotrophs), animals are fed by other sources (that is, they are heterotrophs). Animals obtain food by ingestive or absorptive nutrition. Most animals take food into their bodies (ingestive), but some aquatic animals absorb dissolved nutrients through their cell membranes either from symbiotic partners or from the surrounding water. The energy source is organic compounds.

Animals may reproduce sexually or asexually or by both methods. Sexual reproduction involves two different reproductive cells, a female egg cell and a smaller male sperm cell. At fertilization these two cells join and produce a zygote that is destined to become a new organism. The animal zygote divides to form a blastula embryo that is a hollow, liquid-filled ball of cells. Blastulas are found only in the animal kingdom. Animal species also vary broadly in their mode of embryonic development. Embryos of some species (human, for example) develop directly into smaller versions of the adult, but in other species the embryos develop first into a larva, an immature stage preceding adulthood. For example, frog embryos develop into a tadpole, which is a larval stage, and then the tadpole dramatically reorganizes through the process of metamorphosis into the adult form. Adult animals produce sex cells that carry on the continuous life cycle of the species. Different patterns of development are one line of evidence utilized to deduce ancestry of a phylum.

In contrast to sexual reproduction, asexual reproduction is by a single individual without fertilization of egg by sperm. New individuals are produced asexually by various processes, such as budding by sponges, detaching of body parts by annelids, parthenogenesis by rotifers, and pseudogamy by ticks. Asexual reproduction occurs in about half of the animal phyla. *See* BLASTULATION; EMBRYOGENESIS; REPRODUCTION (ANIMAL).

Multicellularity. Animals are multicellular, but since all five kingdoms include certain multicellular species multicellularity alone does not define an animal. Animals, like plants, fungi, and protoctists, but unlike bacteria, are also eukaryotic. Eukaryotes have membrane-bound nuclei; more than one chromosome; and organelles (that is, "little organs" within a cell, such as mitochondria and a nucleus). In contrast to plant cells, animal cells lack cell walls (which contain cellulose). Moreover, animal cells intercommunicate and support one another by means of elaborate junctions. In addition, the cells of animals form tissues which specialize in a wide array of functions; examples of such tissues are muscle, skin, and blood. The only exception is animals of phyla Porifera (sponges) and Placozoa, which have poorly defined tissues. *See* CELL (BIOLOGY); EUKARYOTAE; PLANT CELL; TISSUE.

Invertebrates and vertebrates. Traditionally, animals have been grouped into invertebrates (without backbones) and vertebrates (with backbones). Vertebrates include mammals, amphibians, reptiles, birds, fish, all members of the phylum Chordata. However, a few chordates, such as tunicates, are invertebrates. Members of all other animal phyla, more than 98% of all animal species, are invertebrates. Although invertebrates lack backbones, they achieve physical support by structures ranging from delicate glass spicules, to tough rings and rods, to hydrostatic pressure. The number of animal species within a single invertebrate phylum varies considerably: Placozoa comprises a single species, but the great phylum Arthropoda comprises more than 800,000 species. If tropical species were better described, the arthropods might include as many as 10 million living insect species. *See* AMPHIBIA; AVES; CHONDRICHTHYES; CHORDATA; MAMMALIA; OSTEICHTHYES; REPTILIA.

Animal ancestry. Microscopic evidence, fossils, geographic distribution, function (physiology), behavior, embryonic patterns, and form (morphology) are lines of evidence from which contemporary classifications are drawn, both within the animal kingdom and in linking it to other kingdoms. The common ancestors of living animals probably evolved over 700 million years ago. Most biologists agree that animals evolved from choanomastigotes, which are microscopic protozoa ancestral to the sponges. Examination of ribosomal ribonucleic acid sequences indicate that animals and fungi share an evolutionary history: their common ancestor may well be a single-celled organism similar to choanomastigotes. *See* ANIMAL EVOLUTION; PHYLOGENY; PLANT KINGDOM.

Karlene V. Schwartz

Bibliography. R. M. Alexander, *The Chordates*, 1975; M. F. Glaessner, *The Dawn of Animal Life*, 1984; E. D. Hanson, *Origin and Early Evolution of Animals*, 1977; R. Lewin, *Thread of Life*, 1991; L. Margulis and K. V. Schwartz, *Five Kingdoms: An Illustrated Guide to the Phyla of Life on Earth*, 3d ed., 1998; S. P. Parker, *Synopsis and Classification of Living Organisms*, 2 vols., 1982; E. E. Ruppert and R. D. Barnes, *Invertebrate Zoology*, 6th ed., 1994.

Animal reproduction

The formation of new individuals, which may occur by asexual or sexual methods. In the asexual methods, which occur mainly among the lower animals, the offspring are derived from a single individual.

Sexual methods are general throughout the animal kingdom, with offspring ordinarily derived from the paired union of special cells, the gametes, from two individuals. Basic to all processes of reproduction is the origin of the new individual from one or more living cells of the parent or parents. There has been no acceptable demonstration of the origin of a new living organism from other material than preexisting living cells. *See* PREBIOTIC ORGANIC SYNTHESIS.

Asexual reproduction. Asexual processes of reproduction include binary fission, multiple fission, fragmentation, budding, and polyembryony. Among the protozoans and lower metazoans, these are common methods of reproduction. However, the last-mentioned process can occur in mammals, including humans.

Binary fission involves an equal, or nearly equal, longitudinal or transverse splitting of the body of the parent into two parts, each of which grows to parental size and form. This method of reproduction occurs regularly among protozoans, in which it is essentially the process of cell division, with complete separation of the daughter cells. To a limited extent, binary fission may be observed among metazoans, such as sea anemones, as longitudinal fission and among planarians as transverse fission.

Multiple fission, schizogony, or sporulation produces several new individuals from a single parent. It is common among the Sporozoa, such as the malarial parasite, which form cystlike structures containing many cells, each of which gives rise to a new individual. The cyst cells arise from a series of divisions of the nucleus, followed by cytoplasmic divisions of the original cell. *See* SPOROZOA.

Fragmentation is a form of fission occurring in some metazoans, especially the Platyhelminthes, or flatworms; the Nemertinea, or ribbon worms; and the Annelida, or segmented worms. The parent worm breaks up into a number of parts, each of which regenerates missing structures to form a whole organism. It occurs also in certain starfish, as *Linckia*, in which single arms may pinch off and regenerate a complete animal.

Budding is a form of asexual reproduction in which the new individual arises from a relatively small mass of cells that initially forms a growth or bud on the parental body. The bud may assume parental form either before separation from the body of the parent as in external budding, or afterward, as in internal budding. External budding is common among sponges, coelenterates (**Fig. 1**), bryozoans, flatworms, and tunicates. Among certain of the coelenterates, such as the colonial hydroid *Obelia*, buds give rise to medusae, or jellyfish, rather than to the parental-type polyp. The medusae represent the sexual generation. They are free-swimming and of separate sexes, producing eggs and sperm, respectively. Upon fertilization of the eggs the asexual, polyp-type individual develops. Another example of asexual individuals budding sexual individuals is found in the cestodes, or tapeworms. Here, the head or scolex, by which the animal is attached to the host tissue, produces

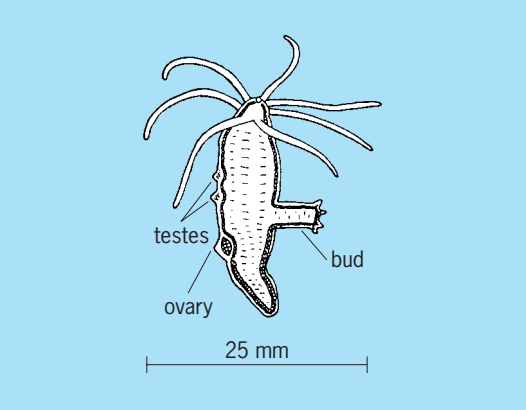

Fig. 1. Budding in *Hydra*.

a series of segments, termed proglottids, each of which is a sexual individual. This phenomenon is also known as strobilization. In some species of sponges, coelenterates, bryozoans, and tunicates, budding may occur without separation of the buds and thus lead to the formation of an organic colony, accompanied in some cases by specialization of parts for particular functions, such as in the Portuguese man-of-war, *Physalia*.

Internal budding occurs among fresh-water sponges and bryozoans. In the sponges the internal buds, termed gemmules, consist of groups of primitive cells surrounded by a dense capsule formed by the body wall. If the parent animal dies as a result of desiccation or low temperature, the cells of the gemmules can later be released and form new sponges. In the bryozoans the similarly functioning buds are known as statoblasts.

Polyembryony is a form of asexual reproduction, occurring at an early developmental stage of a sexually produced embryo, in which two or more offspring are derived from a single egg. Examples are found scattered throughout the animal kingdom, including humans; in humans it is represented by identical twins, triplets, or quadruplets. In some flatworms, polyembryony is illustrated by the rediae, each of which in turn produces many young tadpole-like cercariae. A striking example of polyembryony is found among the insects in the hymenopteran *Litomastix*, which is a parasite on the egg of the moth *Plusia*. The embryo of this wasp subdivides so extensively that about 1500 individuals are formed.

In mammals, polyembryony is a regular feature of the development of armadillos. The four offspring produced at a single birth in the nine-banded armadillo (*Dasypus novemcinctus*) are identical quadruplets. The quadrupling process occurs during the late blastocyst stage of development. In humans, twins occur in about 1.1% of births, triplets in 0.012%, and quadruplets in 0.00014%. About one-third of the twins are identical and have thus arisen by polyembryony, probably occurring in a late blastocyst stage at the onset of gastrulation.

Sexual reproduction. Sexual reproduction in animals assumes various forms which may be classified under conjugation, autogamy, fertilization

(syngamy), and parthenogenesis. Basically, the various processes all involve the occurrence of certain special nuclear changes, termed meiotic divisions, preliminary to the production of the new individual. *See* GAMETOGENESIS; MEIOSIS.

Conjugation occurs principally among the ciliate protozoans, such as *Paramecium*, and involves a temporary union of two individuals during which each is "fertilized" by a micronucleus from the other. In this process the macronucleus of each conjugant breaks down, and the micronucleus undergoes two meiotic divisions to form four nuclei, of which three degenerate. The fourth nucleus divides again, and each conjugant transfers one of these micronuclei to its partner, where it fuses with the stationary micronucleus. The conjugants then separate and the fusion-micronucleus in each divides three times. Of the eight nuclei, four form macronuclei and three degenerate. The exconjugants then divide twice, along with mitotic division of the micronuclei, so that each of the four cells obtains one macro- and one micronucleus. The parental state is thus restored, and further reproduction occurs by simple fission. The ability to conjugate again is not attained until after a large number of divisions. Conjugation does not ordinarily occur within clones, which are organisms derived by mitotic division from a single individual. Strains that are capable of conjugating with one another are designated mating types, rather than males or females, because there are no correlated morphological distinctions between them, and moreover a particular species may have several mating types that are interfertile and thus represent more than two sexes.

In autogamy the nuclear changes described for conjugation take place, but since there is no mating, there is no transfer of micronuclei. Instead, the prospective migratory micronucleus reunites with the stationary one. The process may be considered related to parthenogenesis.

Fertilization, or syngamy, comprises a series of events in which two cells, the gametes, fuse and their nuclei, which had previously undergone meiotic divisions, fuse. In metazoans, the gametes are of two morphologically distinct types: spermatozoa, or microgametes, and eggs, also called ova or macrogametes. These types are produced by male and female animals, respectively, but in some cases both may be produced by a single, hermaphroditic individual. The nucleus of the spermatozoon has half the number of chromosomes characteristic of the ordinary (somatic) cells of the animal. The nucleus of the ripe egg in some animals, for instance, coelenterates and echinoderms, also has attained this haploid condition, but in most species of animals it is at an early stage of the meiotic divisions when ready for fertilization. In the latter situation, the meiotic divisions of the egg, characterized by formation of small, nonfunctional cells termed polar bodies, are completed after the sperm enters, whereupon the haploid egg nucleus fuses with the haploid sperm nucleus. Fertilization thus produces a zygote with the diploid chromosome number typical of the somatic cells of the species (23 pairs in humans) and this is maintained during the ensuing cell divisions (**Fig. 2**). *See* FERTILIZATION.

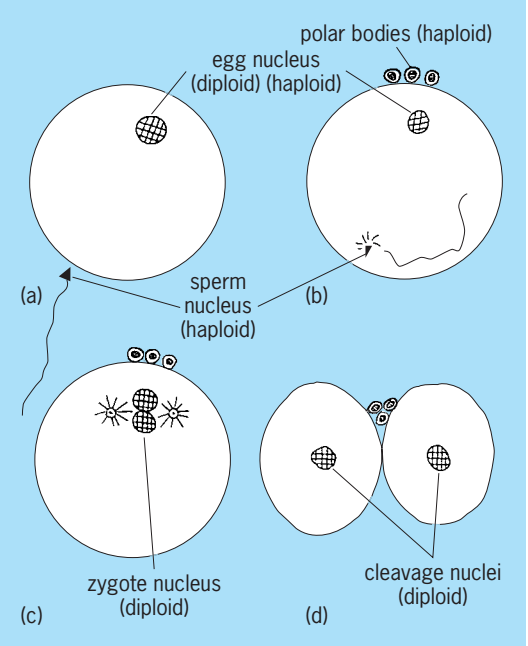

Fig. 2. Stages (*a–d*) of fertilization in animals. In different species the sperm enters at one or another of the stages between *a* and *b*, during which the two meiotic divisions of the egg occur.

In many protozoans, such as the flagellate *Polytoma*, the fusing gametes are not visibly different from one another nor from the ordinary organism. They are termed isogametes. In others they may differ (anisogametes), usually in size as in *Chlamydomonas braunii*; in some cases (*C. coccifera*), one gamete is motile and the other not.

Some metazoans are able to produce functional gametes while still in a larval condition. Reproduction of this type, termed neoteny, is exhibited in amphibians by the axolotl, and also in various insects. The Larvaceae among the tunicates are considered to represent persistent larvae reproducing neotenically.

Hermaphroditism is rare among the vertebrates but is found among most groups of invertebrates. Despite the production of eggs and sperm by the same individual, most hermaphroditic animals reproduce by cross-fertilization. In some cases, this is due to differences in time of ripening of the eggs and sperm. In others, with internal fertilization, crossing is assured by differences in location of sperm and egg ducts. In ascidians there is a physiological block to the union of the spermatozoon with an egg from the same individual. The latter situation illustrates that the interaction of egg and sperm in fertilization is not only generally species-specific but may also be individual-specific. *See* HERMAPHRODITISM.

Parthenogenesis is the development of the egg without fertilization by a spermatozoon. It is listed as a form of sexual reproduction because it involves development from a gamete. Rotifers, crustaceans,

and insects are the principal groups in which it occurs naturally. It has also been induced (artificial parthenogenesis) in species from all the major phyla by various kinds of chemical or physical treatment of the unfertilized egg. Even in mammals, several adult rabbits have reportedly been thus produced. *See* DEVELOPMENTAL BIOLOGY.

The honeybee provides a classic example of natural parthenogenesis. The males, or drones, all develop from unfertilized eggs and are haploid. The females, or workers and queen, arise from fertilized eggs and are diploid. Since the queen is inseminated only once, during the nuptial flight, she stores the sperm during her egg-laying life, which is 5 years or more, and can evidently permit or prevent them from fertilizing the eggs that are laid.

In certain animals periods of parthenogenesis may alternate with fertilization. Thus aphids produce parthenogenetic females during part of the year, but as winter approaches males appear, whereupon fertilization ensues. The fertilized eggs hatch out into females the following spring.

Among certain of the Lepidoptera, like the gypsy moth, *Lymantria*, parthenogenesis occurs at times when there appears to be a scarcity of males, and the unfertilized eggs can form males or females. There is also sporadic occurrence of parthenogenesis in chickens and turkeys. Mostly the embryos are abnormal and die at an early stage, but in the turkeys many develop quite far, and a few have been reared to adults.

Parthenogenetic development may also occur in eggs that are produced during a larval stage, as in the gallfly, *Miastor*. This condition is termed pedogenesis and is the parallel of neoteny. *See* ESTRUS; OOGENESIS; OVUM; SPERM CELL; SPERMATOGENESIS.

Albert Tyler; Howard L. Hamilton

Animal symmetry

Animal symmetry relates the organization of parts in animal bodies to the geometrical design that each type suggests. The term asymmetrical applies to most sponges and some protozoa because the body lacks definite form or geometry, as it cannot be subdivided into like portions by one or more planes. Spherical symmetry is exhibited by some protozoa, such as the Heliozoia and Radiolaria. The body is spherical with its parts concentrically around, or radiating from, a central point. Radial symmetry is exemplified by the echinoderms and most coelenterates. The body is structurally a cylinder, tall or short, having a central axis named the longitudinal, anteroposterior, or oral-aboral axis (see **illus.**). Any plane through this axis divides the animal into like halves. Often several planes, from the axis outward, can divide the body into a number of like portions, or antimeres, five in most echinoderms. Ctenophores and many sea anemones and corals possess biradial symmetry, basically radial but with some parts arranged on one plane through the central axis. Most animals have bilateral, or two-sided, symmetry, in which a median or sagittal plane divides the body into equivalent right and left halves, each a mirror image of the other.

Tracy I. Storer

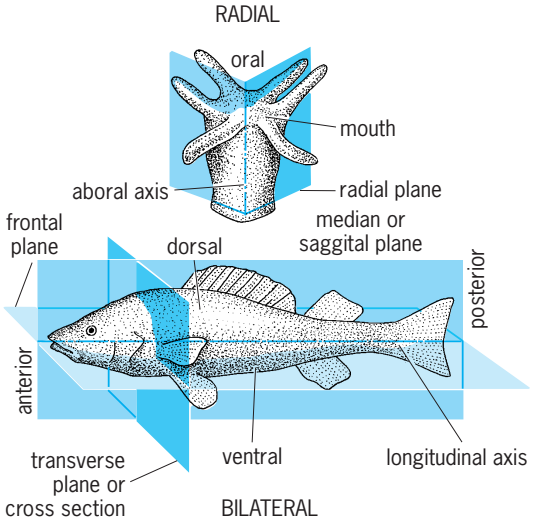

Types of symmetry and the axes, planes, and regions in animal bodies. (*After T. I. Storer et al., eds., General Zoology, 6th ed., McGraw-Hill, 1979*)

Animal virus

A small infectious agent that is unable to replicate outside a living animal cell. Unlike other intracellular obligatory parasites (for example, chlamydiae and rickettsiae), they contain only one kind of nucleic acid, either deoxyribonucleic acid (DNA) or ribonucleic acid (RNA), but not both. They do not replicate by binary fission. Instead, they divert the host cell's metabolism into synthesizing viral building blocks, which then self-assemble into new virus particles that are released into the environment. During the process of this synthesis, viruses utilize

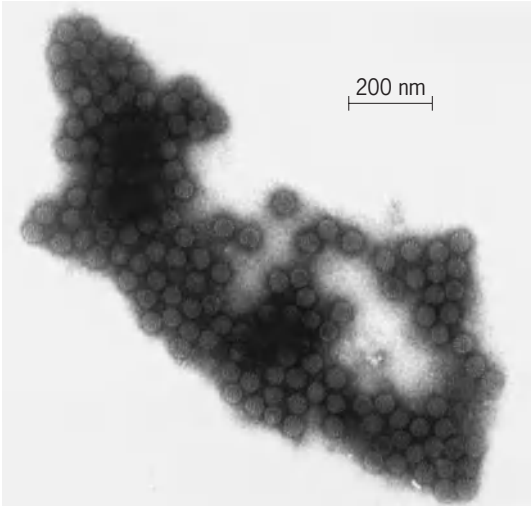

Fig. 1. Electron micrograph of negatively stained, purified bovine papilloma virions with icosahedral symmetry. (*Courtesy of David Hill and John P. Sundberg*)

cellular metabolic energy, many cellular enzymes, and organelles which they themselves are unable to produce. For this reason they are incapable of sustaining an independent synthesis of their own components. Animal viruses are not susceptible to the action of antibiotics. The extracellular virus particle is called a virion, while the name virus is reserved for various phases of the intracellular development. *See* DEOXYRIBONUCLEIC ACID (DNA); RIBONUCLEIC ACID (RNA).

Morphology. Virions are small, 20–300 nanometers in diameter, and pass through filters which retain most bacteria. Due to this property, viruses were called filterable agents of disease. However, large virions (for example, vaccinia, which is 300 nm in diameter) exceed in size some of the smaller bacteria.

The major structural components of the virion are proteins and nucleic acid, but some virions also possess a lipid-containing membranous envelope. The protein molecules are arranged in a symmetrical shell, the capsid, around the DNA or RNA. The shell and the nucleic acid constitute the nucleocapsid.

In electron micrographs of low resolution, virions appear to possess two basic shapes: spherical and cylindrical. High-resolution electron microscopy and x-ray diffraction studies of crystallized virions reveal that the "spherical" viruses are in fact polyhedral in their morphology, while the "cylindrical" virions display helical symmetry. The polyhedron most commonly encountered in virion structures is the icosahedron, in which the protein molecules are arranged on the surface of 20 equilateral triangles. Based on these morphological features, viruses are classified as helical or icosahedral (**Fig. 1**). Certain groups of viruses do not exhibit any discernible features of symmetry and are classified as complex virions (**Fig. 2**). Further distinction is made between virions containing RNA or DNA as their genomes and between those with naked or enveloped nucleocapsids. These assignments are shown in the **table**, in which the major groups of animal viruses are listed.

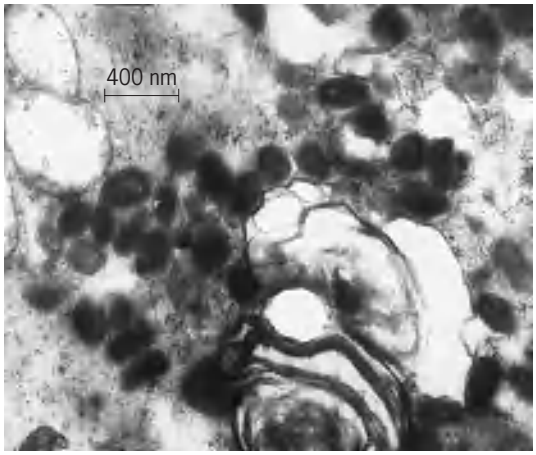

Fig. 2. Electron micrograph of a thin section of squirrel cells afflicted with squirrel fibroma. The brick-shaped pox virions contain a characteristic dumbbell-shaped nucleocapsid. (*Courtesy of David Hill and John P. Sundberg*)

Viral nucleic acid. The outer protein shell of the virion furnishes protection to the most important component, the viral genome, shielding it from destructive enzymes (ribonucleases or

Classification of animal viruses

Family*	Prototype	Nucleic acid	Structure
Poxviridae	Vaccinia virus	Double-stranded DNA	Complex, enveloped
Parvoviridae	Adeno-associated	Single-stranded DNA	Icosahedron, naked
Reoviridae	Reovirus	Double-stranded RNA, fragmented	Icosahedron, naked
Rhabdoviridae	Vesicular stomatitis virus rabies	Single-stranded RNA negative strand	Bullet-shaped, helical nucleocapsid, enveloped
Herpesviridae	Herpes simplex	Double-stranded DNA	Icosahedron, enveloped
Adenoviridae	Adenovirus	Double-stranded DNA	Icosahedron, naked
Papovaviridae	Simian virus 40	Double-stranded DNA, circular	Icosahedron, naked
Retroviridae	Rous sarcoma	Single-stranded RNA	Complex, enveloped
Paramyxoviridae	Newcastle disease	Single-stranded RNA negative strand	Helical nucleocapsid, enveloped
Orthomyxoviridae	Influenza	Single-stranded RNA negative strand, fragmented	Helical nucleocapsid, enveloped
Togaviridae	Alpha: Sindbis Flavi: Yellow fever	Single-stranded RNA	Icosahedron, enveloped
Coronaviridae	Avian infectious bronchitis	Single-stranded RNA	Complex, enveloped
Arenaviridae	Lymphocytic choriomeningitis	Single-stranded RNA negative strand	Complex, enveloped
Picornaviridae	Polio	Single-stranded RNA	Icosahedron, naked
Bunyaviridae	Bunyamwera	Single-stranded RNA negative strand, fragmented	Complex, enveloped

*After F. Fenner, The classification and nomenclature of viruses: Summary of results of meetings of the International Committee on Taxonomy & Viruses in Madrid, September 1975, *Virology,* 71:371, 1976.

deoxyribonucleases). The viral genome carries information which specifies all viral structural and functional components required for the initiation and establishment of the infectious cycle and for the generation of new virions. This information is expressed in the alphabet of the genetic code (sequence of nucleotides) and may be contained in a double-stranded or single-stranded (parvoviruses) DNA, and double-stranded (reoviruses) or single-stranded RNA. The viral DNA may be linear or circular, and the viral RNA may be a single long chain or a number of shorter chains (fragmented genomes), each of which contains different genetic information. Furthermore, some RNA viruses have the genetic information expressed as a complementary nucleotide sequence. These are classified as negative-strand RNA viruses. Finally, the RNA tumor viruses have an intracellular DNA phase, during which the genetic information contained in the virion RNA is transcribed into a DNA and integrated into the host cell's genome. The discovery of this process came as a surprise, since it was believed that the flow of genetic information was unidirectional from DNA to RNA to protein and could not take place in the opposite direction, from RNA to DNA. The transcription of RNA to DNA was termed reverse transcription, and the RNA tumor viruses are sometimes referred to as retroviruses. A classification based on the relationship between viral genomes and their messenger RNA (mRNA) is shown in **Fig. 3**. *See* GENETIC CODE; TUMOR VIRUSES.

When introduced into a susceptible cell by either chemical or mechanical means, the naked viral nucleic acid is in most cases itself infectious. Two exceptions are the negative-strand RNA viruses and the RNA tumor viruses. In these cases the RNA has to be first transcribed and reverse-transcribed, respectively, into the proper form of genetic information before the infectious process can take place. This task is carried out by means of an enzyme which is contained in the protein shell of the virion nucleocapsid.

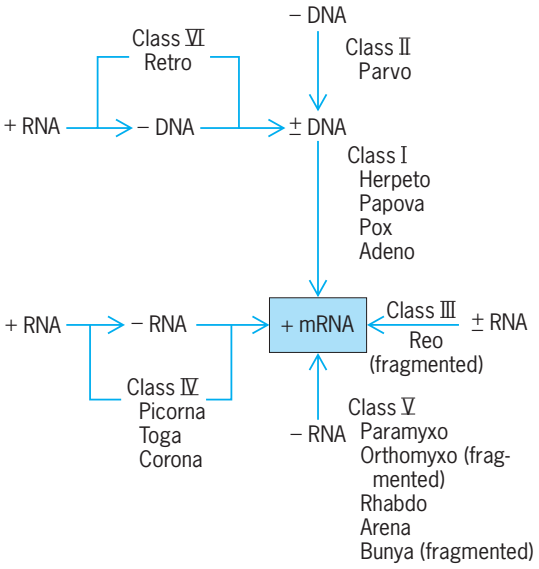

Fig. 3. Classification based on the relationship between the viral genome and its mRNA. (*After D. Baltimore, Expression of animal virus genomes, Bact. Rev., 35:235–241, 1971*)

The whole nucleocapsid is therefore required for infectivity.

Assay. Advances in virus research are closely linked with the utilization of new laboratory techniques. The introduction of tissue cultures eliminated the necessity to use experimental animals and revolutionized animal virus research. The application of plaque assays, first developed in studies of bacterial viruses, was of similar importance. In this assay, a monolayer of cells growing on solid support (usually in a petri dish) is infected with a dilute solution of virions and overlaid with a soft nutrient agar layer. The high viscosity of the agar layer prevents newly released virions from diffusing farther than the cells adjacent to the ones originally infected. After a sufficiently large number of surrounding cells are killed by newly released virions (after 1 day to 3 weeks, depending on the virus), a clear plaque becomes visible. Since each plaque was initiated by a single virus particle, the method allows a quantitative determination of the number of viable virions and offers a means of virus purification (the virus population isolated from a single plaque is relatively homogeneous). Plaque assays are frequently used in diagnosis and in the clinical isolation of viruses. Other assays are based on the ability of some viruses to hemagglutinate (aggregate) red blood cells. The highest dilution of a virus-containing extract which is still capable of hemagglutinating (end-point dilution) is used as a measure of virion concentration. Titrations with specific antibodies are similarly applied in the characterization and quantitation of viruses. *See* TISSUE CULTURE.

Infectious cycle. Viral infection is composed of several steps: adsorption, penetration, uncoating and eclipse, and maturation and release.

Adsorption. Adsorption takes place on specific receptors in the membrane of an animal cell. The presence or absence of these receptors determines the tissue or species susceptibility to infection by a virus. Enveloped viruses exhibit surface spikes which are involved in adsorption; however, most animal viruses do not possess obvious attachment structures.

Penetration. Penetration takes place through invagination and ingestion of the virion by the cell membrane (phagocytosis or viropexis). Enveloped viruses often also enter the cell by a process of fusion of the virion and cell membranes. Penetration is followed by uncoating of the nucleic acid, or in some cases by uncoating of the nucleocapsid. At this stage, the identity of the virion has disappeared, and viral infectivity cannot be recovered from disrupted cells (special methods are required to utilize the infectious nucleic acid). *See* PHAGOCYTOSIS.

Eclipse. The absence of infectious particles in cell extracts is characteristic of the eclipse period. During the eclipse the biochemical processes of the cell are manipulated to synthesize viral proteins and nucleic acids. The survival of viruses depends on this subversive ability. In the process of evolution they have developed an extraordinary efficiency and a remarkable repertoire of strategies. The eclipse period in infections with DNA viruses starts with the

transcription of the genetic information in the nucleus of the cell (except poxviruses), processing into mRNAs, and their translation into proteins (in the cytoplasm). This process is divided into early and late transcription. In the absence of newly synthesized viral components, the immediate early events must be entirely catalyzed by cellular enzymes. The early proteins, however, are virus-encoded functional proteins which will participate in the synthesis of viral DNA and of intermediate and late viral proteins, as well as in the shutoff of various cellular functions which might be detrimental to viral synthesis. The event that distinguishes early from late mRNA transcription and translation is the onset of viral DNA synthesis. The major late products are the structural proteins of the nucleocapsid. Almost as soon as these proteins are synthesized, they assemble with newly synthesized DNA molecules into virion nucleocapsids. The appearance of these particles signals the end of the eclipse period.

The events of the eclipse period in infections with RNA viruses are similar, except that they take place in the cytoplasm (influenza virus excepted), and a division into early and late transcription cannot be made. In the case of positive-strand RNA viruses, the viral RNA is itself the mRNA. In infections with negative-strand RNA viruses, the virion RNA in the nucleocapsid is first transcribed into positive mRNAs. Intracellular nucleocapsids are present throughout the entire infectious cycle, and the eclipse period cannot be defined in the classical sense. RNA tumor viruses reverse-transcribe their RNA into DNA, which enters the cell nucleus and becomes integrated into the cellular DNA. All viral mRNAs and genomic RNAs are generated by transcription of the integrated DNA (Fig. 3).

Maturation and release. The event characteristic of the maturation step is virion assembly and release. In many cases the protein shell is assembled first (procapsid) and the nucleic acid is inserted into it. During this insertion, processing of some shell proteins by cleavage takes place and is accompanied by a modification of the structure to accommodate the nucleic acid. Unenveloped viruses which mature in the cytoplasm (for example, poliovirus) often exit the cell rapidly by a reverse-phagocytosis process, even before the breakdown of the cell. In some cases, however, a large number of virus particles may accumulate inside the cell in crystalline arrays called inclusion bodies (**Fig. 4**). Viruses that mature in the nucleus are usually released slowly, and the damage to the cell is extensive. Enveloped viruses exit the cell by a process of budding. Viral envelope proteins (glycoproteins) become inserted at various sites into the cell membrane, where they also interact with matrix proteins and with nucleocapsids. The cellular membrane then curves around the complex and forms a bud which detaches from the rest of the cell. **Figure 5** shows virus particles budding from the cell surface of a mouse cell infected with murine leukemia virus.

Effect of viral infections. Two extreme types of effects are identified: lytic infections, which cause cell death by a variety of mechanisms with cell lysis as the most common outcome, and persistent infections, accompanied either by no apparent change in the host cell or by some interference with normal growth control, as in transformation of normal to cancer cells. The degenerative phenomena in tissue cultures during a lytic infection are called cytopathic or cytolytic effects. In animals, extensive destruction of tissue may accompany an infection by a lytic virus. *See* LYTIC INFECTION.

As a defense to certain conditions of infection, animal cells generate a group of substances called interferons which, by a complex mechanism, inhibit replication of viruses. They are specific to the cell species from which they were derived but not to the virus which elicited their generation. (Mouse interferon will protect mouse but not human cells from any viral infection.)

Genetics. Like other microorganisms, viruses mutate either spontaneously or following exposure to

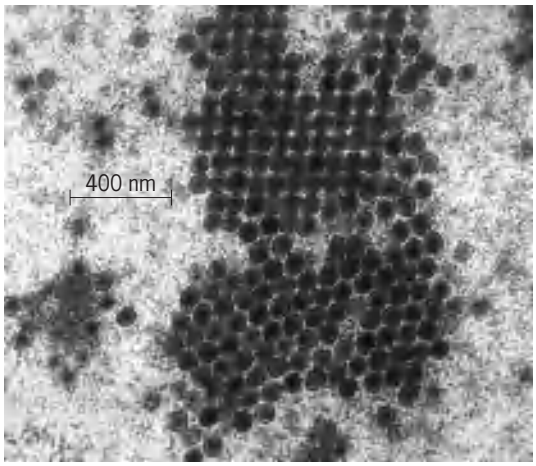

Fig. 4. Electron micrograph of thin section of chicken cells containing crystalline inclusion bodies formed by avian reovirus during infection. (*Courtesy of David Hill and John P. Sundberg*)

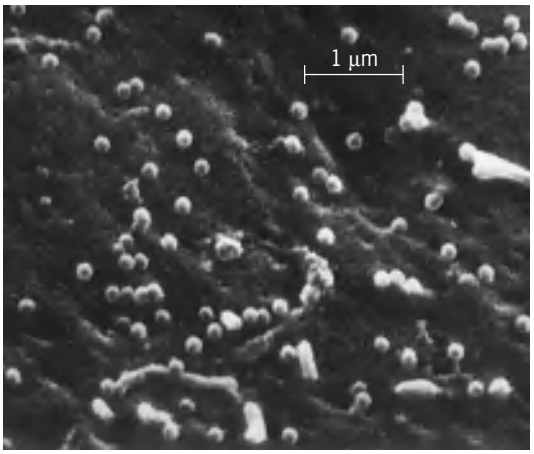

Fig. 5. Scanning electron micrograph of the surface of a mouse cell infected with murine leukemia virus. A large number of virus particles are shown in the process of budding. (*Courtesy of R. MacLeod*)

mutagenic agents. With the introduction of cell tissue cultures and with the development of plaque purification procedures, isolation and characterization of well-defined mutants became possible. In the past the most useful mutants in genetic studies of animal viruses were temperature-sensitive *(ts)* and host-range *(hr)* mutants. The *ts* mutants grew efficiently at lower than normal but not at normal temperatures (permissive and nonpermissive temperatures, respectively). The mutation caused a change of an amino acid in a viral protein which resulted in a biologically inactive configuration at the higher temperatures. Studies of the impaired proteins and their respective genes helped to elucidate the molecular mechanisms of various events in the infectious cycle. The *hr* mutants, which grew normally in certain host cells but not in others, were useful in studies of virus–host interactions. This approach to viral genetics was dependent on the isolation of the proper mutants, which was often fortuitous. With the advent of modern DNA technology of restricting, cloning, and sequencing, it became possible to generate site-specific mutants. Rather than depending on an accidental change of a nucleotide by a mutagen, the investigator could chemically alter a nucleotide at any selected site. In this manner, mutants were generated to answer very specific questions about viral gene expression and virus–host interactions. Since reverse transcription can be carried out in the test tube, these techniques could, in principle, be applied also to RNA viruses as long as their RNAs are infectious. DNA cloning, which permits the preparation of relatively large quantities of viral genomes, is making it possible to study viruses which were inaccessible to investigation either because of low yields obtained in living organisms or because of the lack of suitable tissue cultures. *See* GENETIC ENGINEERING; MUTATION; RECOMBINATION (GENETICS); RESTRICTION ENZYME.

This modern approach to animal virus genetics also rapidly advanced understanding of the complex molecular biology of the animal cell. Many structural and regulatory features of animal virus genes are similar to those of the animal cell, yet are easier to study because of the smaller size of viral genomes. *See* GENETICS; MOLECULAR BIOLOGY.

Pathology. Virus infections spread in several ways: through aerosols and dust, by direct contact with carriers or their excretions, and by bites or stings of animal and insect vectors. At the point of entry, infected cells undergo viremia. From there, the virus becomes disseminated by secretions. It is carried through the lymphatic system and bloodstream to other target organs, where secondary viremias occur (except in localized infections like warts). In most cases viral infections are of short duration and great severity. However, persistent infections are not uncommon (herpes, adeno, various paramyxoviruses like measles). *See* VIRUS INFECTION, LATENT, PERSISTENT, SLOW.

The afflicted organism mounts a variety of defenses, the most important of which is the immune response. Circulating antibodies against viral proteins are generated. Those interacting with virion surface proteins neutralize the infectious potential of the virus. Although the antibodies are specific against the virus which has elicited them, they will cross-react with closely related virus strains. The specificity of neutralizing antibodies obtained from experimentally injected animals is utilized for diagnostic purposes or in quantitative assays. In addition to the circulating antibodies, cell-mediated immune responses also take place. The most important of these is the production of cytotoxic thymus-derived lymphocytes, found in the lymph nodes, spleen, and blood. They destroy all cells which harbor in their membrane viral glycoproteins, regardless of whether these cells are actively involved in virus synthesis or acquired the viral proteins by membrane fusion with inactive virions or cell debris. The cell-mediated immunity has been demonstrated to be more important to the process of recovery than circulating antibodies. In spite of their beneficial role, immune responses often seriously contribute to the pathology of the disease. Circulating antigen–antibody complexes can lodge in organs and cause inflammation; cell-mediated responses have been known to produce severe shock syndromes in patients with a history of previous exposures to the virus. *See* ANTIBODY; AUTOIMMUNITY; CELLULAR IMMUNOLOGY; IMMUNITY; VIRUS INTERFERENCE.

Control. Viruses are resistant to the antibiotics commonly used against bacterial infections. The use of chemotherapeutic agents with antiviral activity is plagued by their toxicity to the animal host. However, the application of vaccines has been successful in the control of several viruses. The vaccines elicit immune responses and provide sometimes lifelong protection. Two types of vaccines have been applied: inactivated virus and live attentuated virus. Various inactivation procedures are available. For example, in preparations of polio vaccines (Salk), a combination of formalin and heat was utilized. Protection by attenuated live virus vaccines dates back as early as 1776, when Edward Jenner used infections with harmless cowpox to achieve protection from the devastating effects of smallpox. In modern times, an attentuated laboratory strain of smallpox has been applied so successfully that the disease is considered to be eradicated. A small probability of back mutations of the attenuated virus to a virulent strain makes applications of live vaccines somewhat riskier. On the other hand, protection is longer-lasting and, by virtue of spread to nonvaccinated individuals, more beneficial to the population group (herd effect). *See* CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; POLIOMYELITIS; VACCINATION; VIRUS CHEMOPROPHYLAXIS.

In order to achieve full protection, it is important that the vaccine contain all the distinct antigenic types of the virus. Development of monoclonal antibodies led to a better characterization of these types in naturally occurring viruses. This information will undoubtedly lead to better vaccines. Moreover, monoclonal antibodies have aided investigations into the molecular structure of viral antigenic groups and brightened future prospects for synthetic vaccines.

See MONOCLONAL ANTIBODIES; VIRUS; VIRUS CLASSIFICATION. M. E. Reichmann

Bibliography. R. Dulbecco and H. S. Ginsberg, *Virology*, 1980; H. Fraenkel-Conrat and P. C. Kimball, *Virology*, 1982; H. Fraenkel-Conrat and R. R. Wagner (eds.), *Comprehensive Virology*, 1974–1983; R. A. Lerner, Synthetic vaccines, *Sci. Amer.*, pp. 66–74, February 1983; S. E. Luria et al., *General Virology*, 3d ed., 1978; D. Nathans, Restriction endonucleases, Simian Virus 40 and the new genetics, *Science*, 206:903–909, 1979; K. M. Smith and M. A. Lauffer (eds.), *Advances in Virus Research*, 1958–1983.

Anise

One of the earliest aromatics mentioned in literature. The plant, *Pimpinella anisum* (Apiaceae), is an annual herb about 2 ft (0.6 m) tall and a native of the Mediterranean region. It is cultivated extensively in Europe, Asia Minor, India, and parts of South America. The small fruits are used for flavoring cakes, curries, pastry, and candy. The distilled oil is used in medicine, soaps, perfumery, and cosmetics. *See* APIALES; SPICE AND FLAVORING. Perry D. Strausbaugh; Earl L. Core

Ankerite

The carbonate mineral $Ca(Fe,Mg)(CO_3)_2$, also commonly containing some manganese. The mineral has hexagonal (rhombohedral) symmetry (see **illus.**) and has the cation-ordered structure of dolomite. The name is applied only to those species in which at least 20% of the magnesium positions are occupied by iron or manganese; species containing less iron are termed ferroan dolomites. The pure compound, $CaFe(CO_3)_2$, has never been found in nature and has never been synthesized as an ordered compound. *See* DOLOMITE.

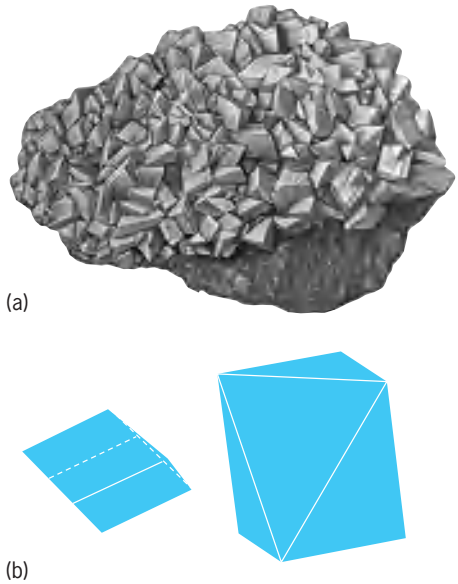

Ankerite. (*a*) Specimen of the mineral (*specimen from Department of Geology, Bryn Mawr College*). (*b*) Crystal habits (*after C. Klein, Manual of Mineralogy, 21st ed., John Wiley & Sons, Inc., 1993*).

Ankerite typically occurs in sedimentary rocks as a result of low-temperature metasomatism. It is commonly white to light brown, its specific gravity is about 3, and its hardness is about 4 on Mohs scale. *See* CARBONATE MINERALS; METASOMATISM. Alan M. Gaines

Bibliography. L. L. Y. Chang, R. A. Howie and J. Zussman, *Rock-Forming Minerals*, vol. 58: *Non-silicates: Sulphates, Carbonates, Phosphates, Halides*, 1995; W. A. Deer, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, vol. 1a: *Orthosilicates* 2d ed., 1997.

Annelida

The phylum comprising the multisegmented, invertebrate wormlike animals, of which the most numerous are the marine bristle worms and the most familiar the terrestrial earthworms. The Annelida (meaning little annuli or rings) include the Polychaeta (meaning many setae); the earthworms and fresh-water worms, or Oligochaeta (meaning few setae); the marine and fresh-water Hirudinea or leeches; and two other marine classes having affinities with the Polychaeta: the Archiannelida (meaning primitive annelids) are small heteromorphic marine worms, and the Myzostomaria (meaning sucker mouths) are parasites of crinoid echinoderms. *See* HIRUDINEA; MYZOSTOMIDA; OLIGOCHAETA; POLYCHAETA.

General Characteristics

These five groups share few common characters and little resemblance except that most have a wormlike body. Typically they are bilaterally symmetrical, lack a skeleton, and have a short to long linear body divided into rings or segments, which are separated from one another by transverse walls or septa. The mouth is an anteroventral or anterior vent at the forward end of the alimentary tract, and the anus posterodorsal or posterior at the hind end of the gut.

Metamerism. The linear series of segments, or metameres, from anterior to posterior ends constitute the annelid body. These segments may be similar throughout, resulting in an annulated cylinder, as in earthworms and *Lumbrineris*. More frequently the successive segments are dissimilar, resulting in regions modified for particular functions. Each segment may be simple or uniannular, corresponding to a metamere, or it may be divided or multiannulate. The total number of segments varies from five to several hundred, according to kind. Segments may have lateral fleshy outgrowths called parapodia (meaning side feet), armed with special secreted bristles or rods, called setae and acicula; they provide protection and aid in locomotion. Setae are lacking in Hirudinea and some polychaetes. The body is covered by a thin to thick epithelium which is never shed.

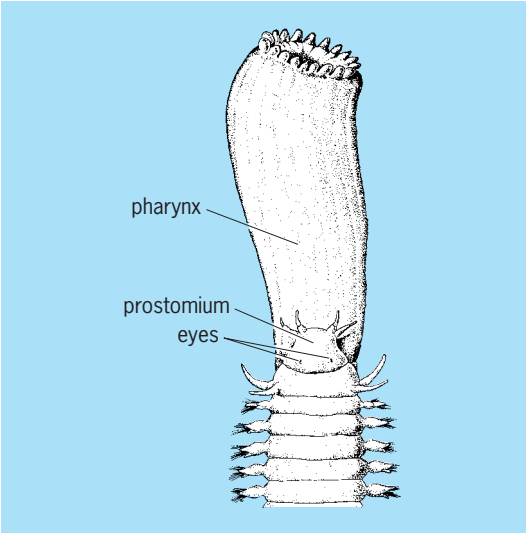

Fig. 1. ***Eteone*** **(Phyllodocidae), everted pharynx, simple prostomial eyes, in dorsal view. (*After O. Hartman, The littoral marine annelids of the Gulf of Mexico, Publ. Inst. Mar. Sci., 2:7–124, 1951*)**

Sense organs. Annelids have sense organs of many kinds. Most conspicuous are those on the anterior, or head end, and on parapodia. Eyes which function as photoreceptors may be variously developed, simple (**Fig. 1**) to complex, small to large, absent to numerous, and obscure to bizarre. Structurally they are simple light-absorbing pigment spots, usually paired on the head, but also on body segments or anal end. In some sabellids they are on the tentacular crown and may number in the hundreds, or they are on the pygidium (**Fig. 2**) and aid in backward movement.

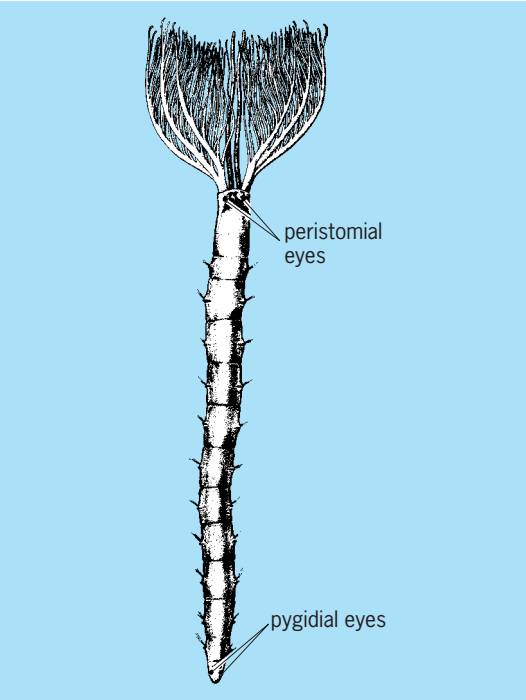

Fig. 2. ***Fabricia*** **(Sabellidae), peristomial and pygidial eyes, in ventral view. (*After O. Hartman, Fabricinae feather-duster polychaetous annelids in the Pacific, Pac. Sci., 5:379–391, 1951*)**

In some polychaetes they are highly developed, with lenses and retina (**Fig. 3**), and may be at the end of an erect stalk. Swarming stages of benthic annelids may acquire large modified eyes which are an attractant at sexual maturity. Some pelagic polychaetes have large eyes which make up most of the anterior end (**Fig. 4**). *See* EYE (INVERTEBRATE).

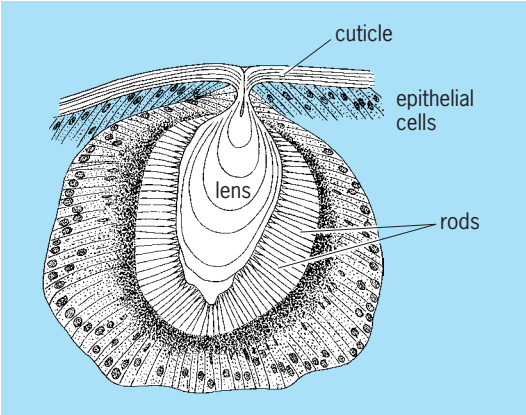

Fig. 3. Cross section of eye of *Eunice* (Eunicidae). (*After O. Pflugfelder, Ueber den feineren Bau der Augen freilebender Polychaeten, Z. wiss. Zool., 142:540–586, 1932*)

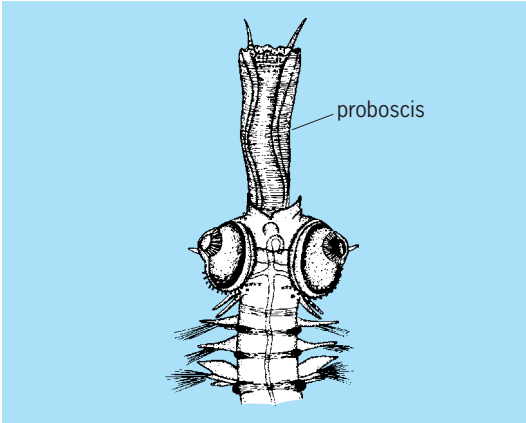

Fig. 4. Anterior end of *Torrea* (Alciopidae), large compound eyes and everted proboscis, in dorsal view. (*After E. Claparède, Les Annelides Chetopodes du Golfe de Naples, Mém. Soc. Phys. Genève, 20:1–225, 1870*)

Feelers or tactoreceptors are frequently on the prostomium as antennae; on anterior segments as long filiform or thick fleshy structures called tentacles; and on parapodia as cirri, papillae, scales, or tactile hairs. Cilia in bands or clusters or in grooves occur in specific patterns. The most conspicuous receptors are the large caruncles (nuchal organs) of amphinomid polychaetes, as simple (**Fig. 5**) or complex (**Fig. 6**) fleshy, folded, paired organs surrounding the cephalic structures. The nuchal organs of some syllids and phyllodocids, as well as the ciliated ridges of errantiate and sedentary polychaetes, illustrate the diversity and importance of these structures. Proprioreceptors are sense organs which respond to stimuli originating within the animal, such

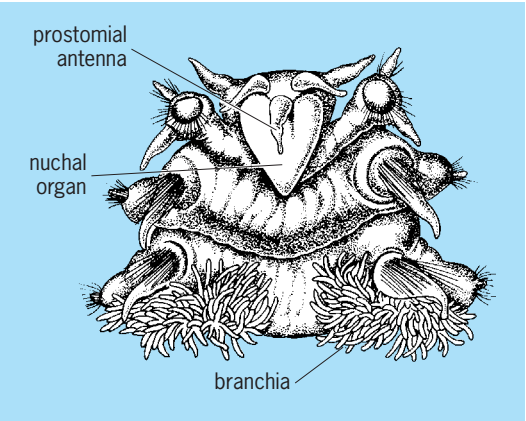

Fig. 5. Simple nuchal organ or caruncle of *Amphinome* (Amphinomidae), in dorsal view. (*After O. Hartman, The littoral marine annelids of the Gulf of Mexico, Publ. Inst. Mar. Sci., 2:7–124, 1951*)

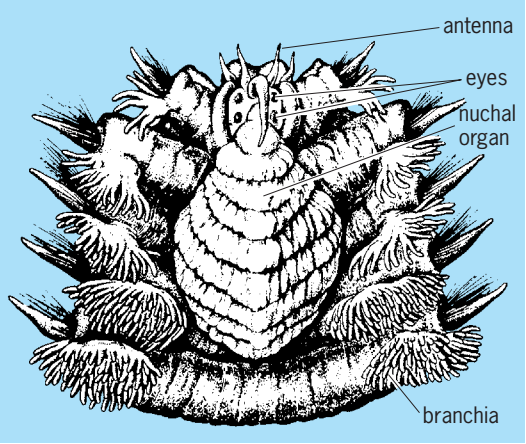

Fig. 6. Anterior end of *Hermodice* (Amphinomidae) with complex caruncle, in dorsal view. (*After O. Hartman, The littoral marine annelids of the Gulf of Mexico, Publ. Inst. Mar. Sci., 2:7–124, 1951*)

as hunger. Trophic reactions to sense food and repelling organs to deter enemies are well developed in many annelids.

Nuchal organs are ciliated, eversible or stationary pouches or grooves located in the anterior or neck region. They relay messages of the immediate environment of the individual. When partly or wholly enclosed as pouches, they may contain small grains either of secreted or foreign origin to form statocysts; they function in righting movements of the body. Setae are part of the sensory system for they detect changes in the environment through their basal attachments. Shallow-water species often have short, strong, resistant setae, whereas abyssal species have long, slender, simple setae. Each organ is unique and well adapted to its role in the development, growth, protection, and reproduction of the species involved.

Chromatophores are cell clusters which change their shape and size to conform to the shadows of the animal's background, responding to changes in light intensities, and therefore are generally protective. They are well developed in translucent pelagic larvae which exist at the surface of the sea; they screen damaging intensities of light from delicate tissues. Oligochaetes and hirudineans, which generally lack eyes or special light receptors, are sensitive to light changes through the surface epithelium. *See* CHROMATOPHORE.

Luminescence. Light production is known in many annelids. Some oligochaetes shed luminous mucus through the mouth or anus. Many polychaetes have special photocytes which luminesce. Some *Chaetopterus* colonies shed bright mucus from the palpi and parapodia. *Polycirrus* among the terebellids is highly luminescent. Pelagic spawning stages of *Odontosyllis* (a syllid) and *Tharyx* (cirratulid) flash bright light when disturbed. Some scale worms (polynoids) have light-producing cells on the undersides of elytra. Flashes are rhythmic and may continue for a minute or more. Luminescence may lure other individuals of the species or aid escape. The posterior, broken-off end continues to light up long after the anterior end has escaped. *See* BIOLUMINESCENCE.

Digestive system. The alimentary tract is a straight or sinuous tube, consisting of mouth, pharynx, esophagus, stomach, gut, and pygidium or anus. Its several parts may be further differentiated and named for structure and function. The mouth may be a simple anterior (oligochaetes) or anteroventral (many polychaetes) pore provided with highly complex organs or accessory parts. The upper lip is formed by the prostomium, and the lower lip by the ventrum of the first segment. Accessory organs include the short to long grooved palpi (**Fig. 7**) of many polychaetes, which direct food to the mouth or also select and propel nutrients along ciliated tracts. The building organs and cementing glands of some tubicolous annelids are associated with the mouth; they select inert particles for shape and size and attach them in specific patterns to a basic secreted mucoid membrane, resulting in tubes which are highly characteristic (**Fig. 8**). In sand-cementing or reef-building colonial species, the beach front may be covered with cemented bands of tubes highly

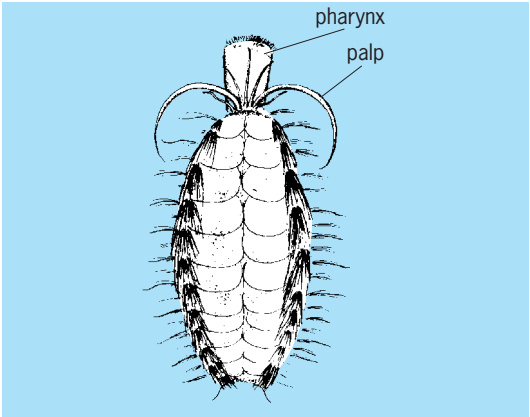

Fig. 7. *Laetmonice* (Aphroditidae), everted pharynx, long palpi, in dorsal view. (*After W. McIntosh, A monograph of the British annelids, Roy. Soc. London, 1(pt. 2):215–442, 1900*)

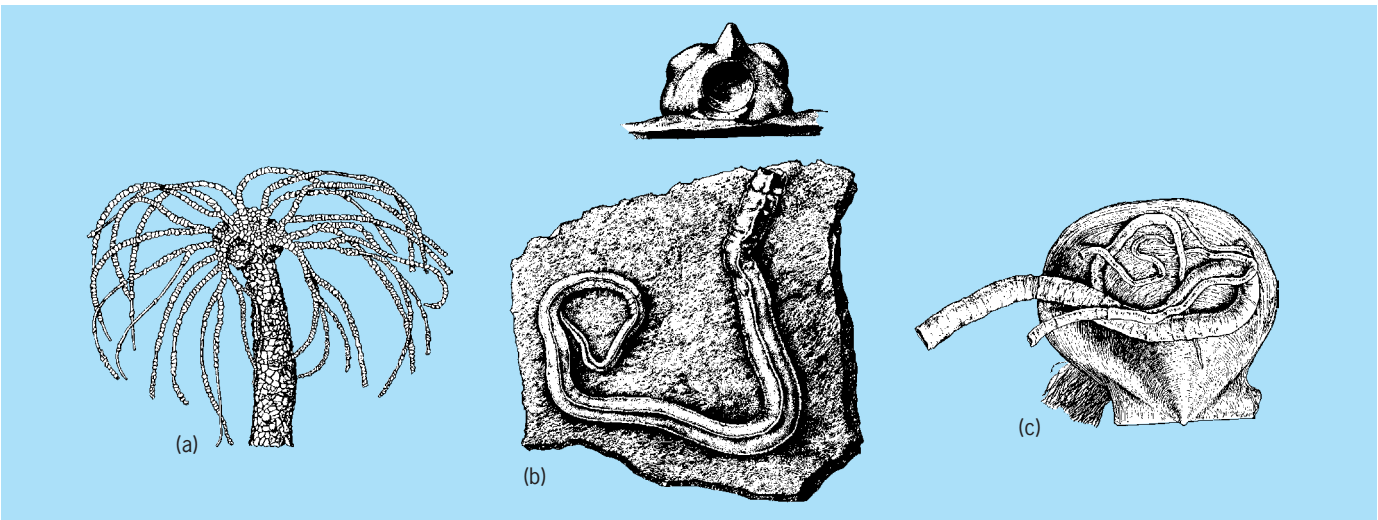

Fig. 8. Tubes of sedentary polychaetes. (*a*) Distal end of tube of *Lanice* (Terebellidae) (*after O. Hartman, Polychaetous annelids collected by the USNS Eltanin and Staten Island cruises, chiefly from Antarctic seas, Allan Hancock Monogr. Mar. Biol., 2:1–387, 1967*). (*b*) Tube of *Vermiliopsis* (Serpulidae), entire and in frontal view. (*c*) Several tubes of *Protula* (Serpulidae) on deep-water pelecypod (*after O. Hartman, Endemism in the North Pacific Ocean, with emphasis on the distribution of marine annelids, and descriptions of new or little known species, Essays in the Natural Sciences in Honor of Captain Allan Hancock, University of Southern California, Los Angeles, 1955*).

resistant to wave action. Colonial serpulid worms secrete calcium carbonate from special collar cells, formed into highly organized tubes with the aid of oral organs.

The mouth is followed by the buccal cavity which may be modified as a proboscis or saclike eversible pouch. In nonselective, mud-eating annelids the pouch may be voluminous when everted, its surface papillated or branched so as to increase its sweeping surface as it brushes over the surface muds; the adherent materials are swept into the oral cavity and swallowed.

The buccal cavity may be followed by a short to long muscularized eversible pharynx (**Fig. 9**) which captures and breaks up or compresses food particles. Its inner walls may be fortified by papillae (**Fig. 10**) or hard gnaths or jaws (**Fig. 11**) called para-, micro-, or macrognaths, maxillae, mandibles, chevrons, trepan, or other descriptive names. A short esophagus leads to the muscular stomach or digestive region. Lateral ceca or pouches may be present, along esophageal and stomach portions, to increase the amount of surface for secretion and digestion, especially in short-bodied worms. Peristaltic (clasping and compressing) movements are rhythmic and result in the food bolus being digested, and wastes separated and pushed into the gut. In nonselective-feeding annelids the gut may be distended with great amounts of inert materials; in selective feeders there are few remains but those of living animals. The proctodaeum, or region preceding the anus, expels the wastes as fecal pellets of characteristic form.

The proctodaeum with pygidium may be a simple, tapering end with posterior pore or may be highly modified as a plaque. In some pelagic swarming nereids the pygidium of the male acquires a circlet of papillae through which sperm cells are shed into the sea. Defecation, or disposal of wastes from the gut, is usually posteriorly. Tubicolous annelids which lie head down may reverse their position for this function; others unable to turn may have an anal ciliated groove directing fecal materials forward and dorsal, away from the mouth.

Nervous system. The nervous system consists of a dorsal, bilaterally symmetrical, ganglionic mass or brain within or behind the preoral region. The brain is connected to the ventral cord through the circumesophageal connectives which extend about the

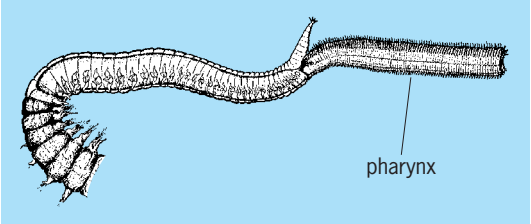

Fig. 9. Anterior end of *Glycinde* (Goniadidae), pharynx partly everted, in right lateral view. (*After O. Hartman, Goniadidae, Glyceridae, Nephtyidae, Allan Hancock Pacific Expedition, 15(1):1–181, 1950*)

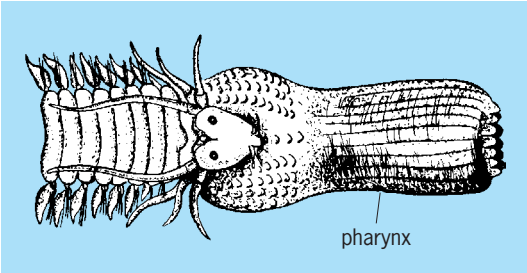

Fig. 10. Anterior end of *Anaitides* (Phyllodocidae) with pharynx fully everted, in dorsal view. (*After E. Rioja, Datos para el conocimiento de la fauna de Anélidos poliquetos del Cantábrico, Trab. Mus. Nac. Madrid, Zool. 37, pp. 1–99, 1918*)

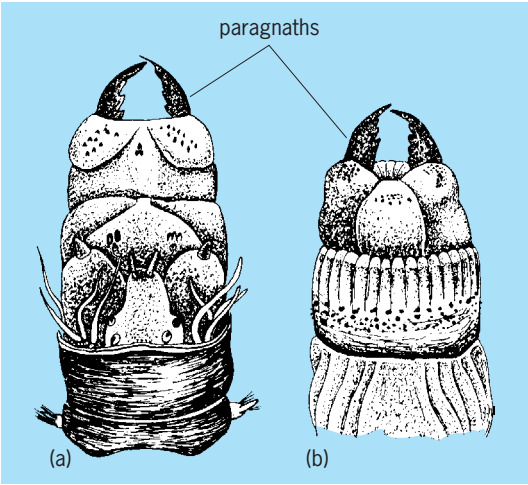

Fig. 11. Anterior end of *Nereis* (Nereidae) with everted pharynx showing distal jaws and paragnaths, (*a*) dorsal and (*b*) ventral views. (*After E. Rioja, Datos para el conocimiento de la fauna de Anelidos poliquetos del Cantábrico, Trab. Mus. Nac. Madrid, Zool. 37, pp. 1–99, 1918*)

oral cavity (**Fig. 12**). The ventral cord may be single or paired, nearly smooth or nodular, or ganglionated according to the segmental pattern of the body. The brain sends out lateral branches to the eyes, palpi, antennae, or other structures; the ventral cord has lateral branches to all fleshy parts which receive stimuli. A giant axon is an enlarged part of the ventral cord, present in many long, muscular, or actively moving oligochaetes and polychaetes. It permits rapid transfer of stimuli and muscular response, resulting in abrupt response. Messages are relayed directly from the source of stimulation to the muscles, without reaching the brain. These giant fibers were first discovered in oligochaetes in 1861, and have since been found in many other annelids. They vary in size and complexity, but are always larger than normal neurons to which they are connected. Their cells arise in the circumesophageal ganglion and extend throughout the length of the worm, or they arise much farther back and may extend longitudinally or crosswise. Each giant fiber may represent a single nerve cell or fusion of several nerve cells. Their diameters are largest in *Myxicola* (Sabellidae) and much smaller in most other annelids. Their occurrence has been reported from widely related families.

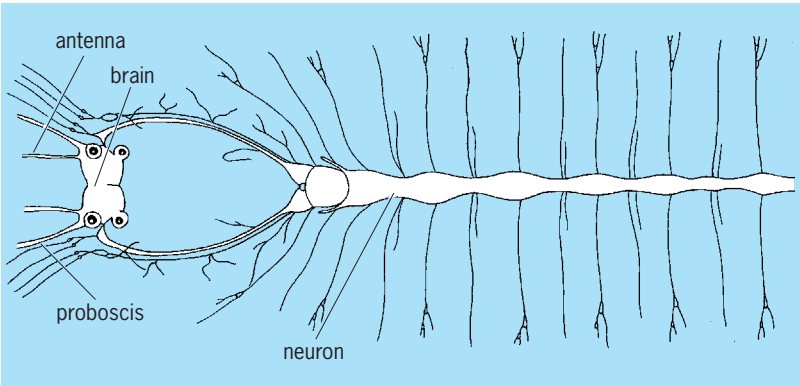

Fig. 12. Diagram of nervous system of *Nereis* showing some of the nerve branches. (*After S. Maxwell, Beiträ zur Gehirnphysiologie der Anneliden, Pfluegers Arch., Berlin, 67:263–297, 1897*)

Neurosecretion is a process of special cells located in the brain. Typically, nerve cells are polarized, consisting of an axon or head and a slender fiber. Secretory cells lack a fiber and secrete hormones. In some annelids they affect the time and rate of metamorphosis at maturity. In nereids the hormones inhibit the processes of metamorphosis, until the sexual cells have arrived at their full stage of development. The two processes, sexual development and transformation to an epitoke, are separate and each under its own control. *See* NEUROSECRETION.

Circulatory system. The circulatory system consists of dorsal and ventral longitudinal, median vessels located above and below the alimentary tract. Lateral branches extend to all parts of the body. Pulsating or propelling contractile portions, sometimes called hearts, are in the anterior dorsal vessel or also at intervals in the ventral vessels. In oligochaetes these are segmental vessels of varying number connecting the dorsal and ventral vessels and surrounding some portion of the alimentary tract.

The contained blood is red through the presence of a hemoglobinlike substance, which may be in corpuscles or granular in the fluid, or it is green, through the presence of chlorocruorin. These colors when diluted are yellow or colorless, as in many small annelids. *See* RESPIRATORY PIGMENTS (INVERTEBRATE).

The purpose of the blood is to transfer nutrients to all parts of the body and to exchange carbon dioxide for oxygen, either through the skin or through special organs called branchiae (Fig. 5). They are fingerlike fleshy prolongations of the epithelium, either simple, branched, or spiraled structures, at anterior or also posterior ends of the body. Each branchia contains a vascular loop through which the blood flows in one direction. The oxygen requirements of annelids may be very low, illustrated in worms existing in stagnant pools and in deep oceanic basins. Irrigation in the tube, burrow, or environment may be achieved through U-shaped structures which open to the surface or through undulations of the body or parapodia, to ensure a flow of water and nutrients.

Some annelids lack a closed circulatory system so that blood and coelomic fluids mix freely, resulting in a hemocoel; it may be partial or complete. Many annelids have a special organ called a cardiac or heart body surrounding the pulsating vessel and sometimes visible as a thick brown or red body of spongy tissue; its function is to dispose of circulatory wastes. The circulatory and coelomic fluids aid in maintaining turgidity of the annelid body, and with musculature control they act as a kind of skeleton.

Excretory system. The excretory system functions to dispose of wastes in the coelom, resulting from metabolism. The special organs for the function are called nephridia. In their simplest form they are protonephridia, consisting of a strand of cells connecting the coelom to the body wall, usually a pair to a segment. The inner end terminates in a flagellum which propels wastes outward; the distal end is a

pore at the surface of the body called a nephridiopore. Primitive nephridia occur in larval, small, or primitive annelids. Clusters of simple nephridia, solenocytes, basally joined to a tube, are present in many errantiate polychaetes. More complex organs, or metanephridia, have a ciliated nephrostome or funnel opening into the coelom and continued to the surface as a complex organ. They function to transport wastes and at sexual maturity may serve as gonoducts to release gonadial products. Complex nephridia are present in all oligochaetes and many polychaetes.

Nephridia may be few in number and conspicuous, or limited to anterior segments as in many sedentary families, or may be segmental, a pair to a segment, or dispersed and irregular. Some coelomic wastes, which may result through phagocytosis, or cell breakdown, cannot be eliminated through nephridia and may accumulate to form conspicuous dark blobs in posterior segments.

Muscular system. The muscular system consists of an outer circular and an inner longitudinal system of muscles, each varying in extent and density according to species. In addition, an oblique series, between outer and inner layers, is well developed in annelids performing complex lateral movements. Long, very active burrowers or crawlers have an extensive musculature, whereas short, sluggish forms may have diminished musculature development. The circular muscles form a sheath beneath the epithelium; they effect an increase or decrease in body diameter, and are best developed in tubicolous and burrowing forms. Longitudinal muscles are often in paired thin to thick bands in dorsal and ventral positions; they effect a shortening or lengthening of the body, and are most conspicuous in errantiate species. The oblique muscles occur in incomplete bands, between circular and oblique layers, connecting parts of the parapodia to the body wall and to other muscular layers.

Movements are achieved mainly by coordinated muscular contractions and expansions of the laterally projecting parapodia or setae, resulting in an undulating or meandering movement. Swimming species may move from side to side or by successive forward darts and stops. Some annelids with reduced parapodia and a long proboscis use the latter in progression by extension and withdrawal of the eversible part of the alimentary tract. Some annelids secrete mucus to move over irregular surfaces.

Regeneration. The ability to replace lost parts is highly developed in annelids. Most frequent is the replacement of tail, parapodia, and setae. The anterior end may be replaced provided the break is postpharyngeal. The torn end is first covered over with scar tissue, then differentiated into epithelial cells and all other tissues characteristic of the whole animal. The tendency is to restore the number of segments peculiar to the species. Least differentiated species regenerate more completely than highly modified ones. Younger individuals repair more easily or quickly than older ones. Fragmentation may result in as many new individuals as there are pieces.

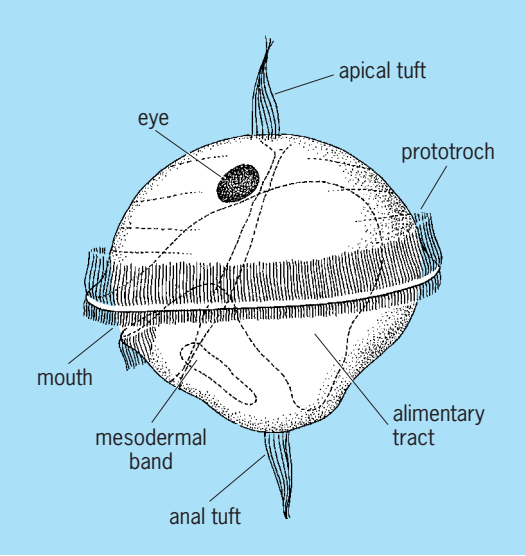

Fig. 13. Lateral view of larval trochophore with parts indicated and much enlarged.

The presence of a part of the ventral nerve cord is necessary to develop a new head. *See* REGENERATIVE BIOLOGY.

Reproduction. Depending on the species, reproduction may be sexual, asexual, or both. Sexual reproduction may be dioecious, in which male and female are similar, rarely dissimilar. Individuals may be hermaphroditic, both male and female, but with cross fertilization. Some annelids are protandric hermaphrodites, in which the sexual stages alternate. Epigamy refers to the transformation of mature individuals from a benthic to a planktonic form; changes occur in parapodial lobes and the alimentary tract may be absorbed. Fully metamorphosed individuals leave their tubes, swarm to the surface of the sea, sometimes in rhythmic or lunar patterns, deposit sperm and ova, and then disintegrate. In the palolo (Eunicidae) only posterior ends are released, and head ends retreat to regenerate new tails. The ova which are deposited at the surface of the sea develop into ciliated spheres called trochophores (**Fig. 13**), which develop into segmented metatrochs (**Fig. 14**) and then into larvae that begin to resemble adults (**Fig. 15**) into which they will metamorphose.

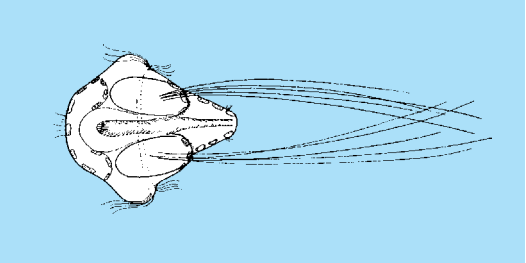

Fig. 14. Metatroch of planktonic sabellarian larva, showing early segmentation and long swimming setae. (*After O. Hartman, Polychaetous annelids, Paraonidae, Magelonidae, Longosomidae, Ctenodrilidae and Sabellariidae, Allan Hancock Pacific Expedition, 10(3):311–389, 1944*)

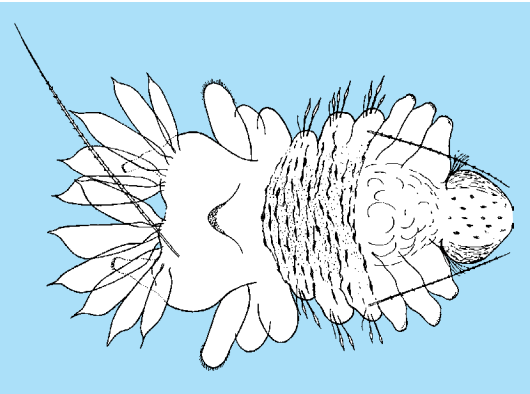

Fig. 15. Transitional stage of metatroch larva of sabellarian larva, with larval setae nearly replaced. (*After O. Hartman, Polychaetous annelids, Paraonidae, Magelonidae, Longosomidae, Ctenodrilidae and Sabellariidae, Allan Hancock Pacific Expedition, 10(3):311–389, 1944*)

At the appropriate time they sink to the bottom and assume a benthic development. *See* PROTANDRY.

Olga Hartman

Fossils

Fossil annelids, or segmented worms, are mostly soft-bodied, yet sufficiently common to indicate that this large and varied group of invertebrates has been abundant for more than 500 million years.

Origins. The bulk of the annelidan fossil record is represented by the polychaetes. The earliest definite occurrences are from the Lower Cambrian of Greenland (Sirius Passet fauna). Possible precursors to the polychaetes exist in the form of wiwaxiids and the more primitive halkieriids. Both are sclerite-bearing, and in the wiwaxiids the microstructure is identical to annelidan chaetae; in these animals the sclerites form a protective coat, but it is plausible to derive the notochaetae and neurochaetae from, respectively, an imbricated dorsal array and fanlike extensions of leg-like parapodia. The coexistence of polychaetes and halkieriids in the Sirius Passet fauna indicates an earlier divergence, but a number of Ediacaran fossils previously assigned to the annelids, such as *Dickinsonia*, probably are not closely related.

Soft-bodied polychaetes. The soft-bodied record from the Cambrian onward is sparse, but polychaetes are known from a number of exceptionally preserved fossil biotas. A diverse assemblage occurs in the Middle Cambrian Burgess Shale of British Columbia. Not only do these polychaetes (**Fig. 16**) show a very wide morphological range, but with one possible exception it has not proved feasible to accommodate any of the six species in a known fossil or living family. *See* BURGESS SHALE.

Other landmarks in the polychaete fossil record are a series of deposits from the Carboniferous that include the Francis Creek Shale (Mazon Creek ironstone nodules; **Fig. 17**) of Illinois, the Bear Gulch Limestone of Montana, and the Granton shrimp bed of Edinburgh, Scotland. The first two units have yielded a wide variety of polychaetes, many with jaws. Although only one polychaete occurs at Granton, it is significant in apparently representing the earliest tomopterid. In younger deposits, soft-bodied polychaetes are known from horizons in the Triassic (Voltzia Sandstone of eastern France, also deposits in Italy and Madagascar), Jurassic (Osteno beds of Italy, Solnhofen Limestone of Germany), Cretaceous (Sahel Alma fish beds of Lebanon), and Eocene (Monte Bolca of Italy).

Scolecodonts. Polychaete jaws, or scolecodonts (**Fig. 18**), first appear in the lower Ordovician but seem to be conspicuously more abundant in the Paleozoic than in either the Mesozoic or Cenozoic. Many are attributed to the eunicids, the jaws of which seem to have a high preservation potential in comparison to those of many other families. Such variation is due to differences in original scolecodont mineralogy, which are still imperfectly understood. However, it seems clear that composition varies taxonomically, with different families having scolecodonts either composed of organic material (scleroproteins) or having this organic matrix mineralized to varying degrees by either aragonite or calcite. In Recent polychaetes, for example, eunicids and onuphids have calcite mineralization and lumbrinerids have aragonite, while arabellids and glycerids are

Fig. 16. *Canadia spinosa* from the Middle Cambrian Burgess Shale of British Columbia. Note the pair of anterior tentacles, broad dorsal chaetae extending across much of the trunk, and ventral chaetae extending from either margin.

unmineralized. It also appears that the composition of scolecodonts has changed during time. Paleozoic scolecodonts were predominantly organic and strongly sclerotized, whereas calcareous mineralization became widespread in the Mesozoic. This mineralogical shift may explain the scarcity of post-Paleozoic scolecodonts, given that aragonite is relatively unstable, and most scolecodonts are extracted in the laboratory by acid digestion. Work on preservation potential, including survival during digestion by predators and resistance to decay in the sediment, is improving understanding, but the correlation between expected survival according to composition and field observation is by no means exact.

Other fossil records. Apart from scolecodonts, the hard-part record of polychaetes consists of tubes, especially the calcareous serpulids and related spirorbids. Spirorbids first appear in the Ordovician, and both groups are a frequent component of many Mesozoic and Cenozoic assemblages. They may occur as encrusters on reefs, hard grounds, and shell debris, or as free-living on sandy substrates. Numerous groups of polychaetes inhabit tubes that are variously reinforced so as to give a wide spectrum of fossilization potential. Most important in the fossil record are the terebellids, sabellariids, and pectinariids, all of which construct an agglutinated tube bound together by mucous secretions. This may be of mud, sand grains, or other selected particles such as foraminiferal tests, bivalve shells, echinoderm ossicles, and fish scales. Assignment to particular groups largely depends on comparisons with living relatives. However, a remarkable specimen from the Miocene of New Zealand shows a soft-bodied sabellid within its tube. In the Paleozoic a variety of tubicolous fossils have been attributed to polychaetes, but many of these are of questionable status.

It has also been customary to attribute many burrows and other trace fossils to polychaete activities, but in few cases are such assignments secure because a variety of vermiform phyla are capable of making effectively indistinguishable traces, but trace fossils (*Walcottia*) from the Upper Ordovician of Cincinnati are so similar to polychaetes that an assignment to this group seems secure. Characteristic borings into hard substrates, such as by spionid polychaetes, are also diagnostic. A unique specimen from the Middle Devonian of New York has a spionidlike polychaete still preserved in its boring. *See* TRACE FOSSILS.

Oligochaetes and leeches. The fossil record of oligochaetes is lamentable, in part because their predominantly terrestrial and fresh-water habitats have a relatively poor rock record. Nevertheless, the reproductive cocoons characteristic of the clitellates (oligochaetes and leeches) have been recognized as far back as the Triassic and may be more common as fossils than is generally realized. In addition, there is a record of oligochaete trace fossils in the form of burrows and associated fecal pellets from ancient soils (paleosols). The most convincing of these come from Cenozoic paleosols, especially in the western interior of the United States. However, the oldest records appear to date back to the Triassic.

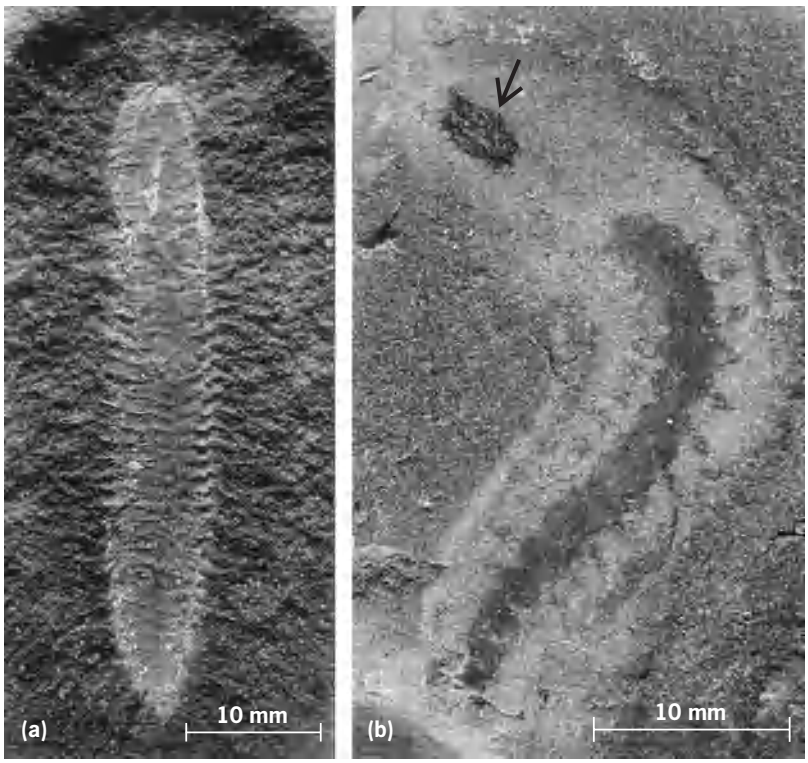

Fig. 17. Polychaetes from the Francis Creek Shale (Mazon Creek ironstone nodules). (*a*) *Astreptoscolex anasillosus*, a nephytid; note the marginal chaetae and cuticular creases. (*b*) *Esconites zelus*, a glycerid; the body is comparatively poorly preserved, but the anterior scolecodont apparatus (arrow) is in place.

Fig. 18. Scolecodont element (*Kettnerites martinssoni*) from the Silurian of Gotland, Sweden. (*Courtesy of Claes Bergman, Kristianstad, Sweden*)

The leeches are best known for their blood-sucking abilities, but they also include a number of predators. Apart from the cocoons, they are poorly represented in the fossil record. A tentative record from the Lower Silurian of Wisconsin is based on the putative terminal sucker of an annulated soft-bodied worm. Much younger worms from the Upper Jurassic Solnhofen Limestone are also claimed, on slender evidence, to represent leeches.

Echiurans. Molecular data indicate that the echiurans, long thought to be related to the Annelida, may nest within the Polychaeta. The fossil record, however, is restricted to examples from the upper Carboniferous (Mazon Creek) and, more speculatively, various traces attributed to the activity of the characteristic proboscis.

Pogonophorans. Pogonophorans were once regarded as a separate phylum, but evidence (including molecular) now indicates that they nest within the annelids. Soft-bodied remains are not known, but certain tubicolous fossils have been assigned with varying degrees of confidence to the pogonophorans. Some are associated with hydrothermal deposits, ancient analogues of the midocean ridge "black smokers." *See* FOSSIL. S. Conway Morris

Bibliography. D. E. G. Briggs and E. N. K. Clarkson, The first tomopterid, a polychaete from the Carboniferous of Scotland, *Lethaia*, 20:257–262, 1987; G. K. Colbath, Jaw mineralogy in eucinean polychaetes (Annelida), *Micropaleontology*, 32:186–189, 1986; S. Conway Morris, Middle Cambrian polychaetes from the Burgess Shale of British Columbia, *Phil. Trans. Roy. Soc. London*, B285:227–274, 1979; S. Conway Morris and J. S. Peel, Articulated halkieriids from the Lower Cambrian of North Greenland and their role in early protostome evolution, *Phil. Trans. Roy. Soc. London*, B347:305–358, 1995; D. Jones and I. Thomson, Echiura from the Pennsylvanian Essex fauna of northern Illinois, *Lethaia*, 10:317–325, 1977; S. B. Manum, M. N. Bose, and R. T. Sawyer, Clitellate cocoons in freshwater deposits since the Triassic, *Zool. Scripta*, 20:347–366, 1991; N. Kotake, Deep-sea echiurans: Possible producers of *Zoophycos*, *Lethaia*, 25:311–316, 1992; I. Thompson, Errant polychaetes (Annelida) from the Pennsylvanian Essex fauna of northern Illinois, *Palaeontographica*, Abteilung A, 163:159–199, 1979.

Anomopoda

An order of fresh-water branchiopod crustaceans, formerly included in the Cladocera. Exceptionally these organisms may be as much as 6 mm (0.24 in.) in length, but often are less than 1 mm (0.04 in.). So-called water fleas of the genus *Daphnia* are the most familiar anomopods. *Daphnia* and its close relatives swim freely in open water; however, many anomopods are benthic, predominantly crawling species.

The short trunk is protected by a functionally bivalved carapace that covers the five or six pairs of trunk limbs and the terminal postabdomen that bears two claws. The trunk limbs are extremely diverse in structure, both among themselves and between species. They collect food, usually particulate, by scraping or filtration or both, and are used by many benthic species for crawling. The postabdomen is often used for pushing. Swimming is by means of antennae. The short antennules (first antennae) are sensory. There is a single median sessile eye, derived by fusion of two eyes.

Most species feed on minute particles, but *Pseudochydorus* is a scavenger, and *Anchistropus* has the remarkable habit of living parasitically on *Hydra*, against whose stinging cells it is protected. The mandibles are of the rolling, crushing type.

Reproduction is mainly by parthenogenesis, eggs and young being carried dorsally in a brood pouch beneath the carapace. Some races of some species of *Daphnia* are obligate parthenogens. Cyclical parthenogenesis is common in temperate zone species; summer parthenogenesis alternates with sexual reproduction in autumn. Fertilized eggs are cast off, protected by a modified portion of the carapace, which is often called an ephippium, and the eggs can overwinter or resist drought. Worldwide in distribution, anomopods are among the most abundant and successful of all fresh-water invertebrates. *See* BRANCHIOPODA. Geoffrey Fryer

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Anopla

A class of the phylum Rhynchocoela which is divided into the orders Palaeonemertini and Heteronemertini. The simple tubular proboscis lacks stylets and resembles the body wall in structure. The mouth is posterior to the brain. The nervous system lies either immediately below the epidermis or among the musculature of the body wall. The vascular system is well developed.

Anoplan rhynchocoelans are common on rocky shores, living beneath stones or interstitially in shell debris or coarser sands. They are carnivorous, preying almost entirely on annelids, but will also take carrion if this is not too decomposed. Acid secretions kill or immobilize prey in the foregut, and digestion occurs rapidly in the intestine by extra- and intracellular processes. *See* ENOPLA; HETERONEMERTINI; PALAEONEMERTINI; NEMERTEA. J. B. Jennings

Anoplura

A small group of insects usually considered to constitute an order. They are commonly known as the sucking lice. All are parasites living in the covering hair of mammals. About 250 species are now known, and these comprise probably about half of the species in the world, but the group is known in detail to few persons.

Morphology. These lice are distinguishable from the Mallophaga by their mouthparts and manner of feeding. The mouthparts consist of three very slender stylets which form a tube. When at rest, they are retracted in a pocket which lies just behind the mouth. The antennae are usually five-segmented, rarely three-segmented. The thoracic segments are always very closely fused and lack wings. The claw is one-segmented, and on at least one pair of legs this claw is enlarged and can be folded into a process from the tibia to grasp a hair of the host. In some species two and in others all of the legs are thus modified. The ovipositor is very much reduced. It consists of but little more than two flaps which are able to close around a hair. Eyes are commonly lacking, and in those species in which they do occur, they are reduced to a pair of simple lenses or two light-receptive spots. All the species are quite small;

the largest scarcely exceeds 0.2 in. (5 mm) in length and the smallest does not attain 0.04 in. (1 mm).

Parasitic activity. The eggs are attached to the hairs of the host by a glue which issues from the base of the ovipositor and envelops a hair and one end of the egg. The young escape from the egg through a lid or operculum at the free end. In general, they are much like the adult and immediately begin feeding upon the blood of the host. Feeding is accomplished by thrusting forth the tube formed by the stylets, piercing the skin, and sucking blood by a pump in the throat of the insect. Very few species have been studied closely enough to determine their life-span, but in the human louse this is about a month.

Disease transmission. This habit of sucking blood gives the Anoplura a special importance. They take up any disease-producing organisms in the host's blood and may transfer these organisms to another individual. Thus, they transfer the organisms which cause two of the most important diseases of humans, epidemic typhus and relapsing fever. A third disease, trench fever, which gained great prominence during World War I, has also been of very great importance. The transmission of these diseases depends entirely upon certain habits of the insects, which are coincident with certain habits of their hosts. *See* RELAPSING FEVER.

Host specificity. Indeed, these lice have become so closely adapted to life upon a particular kind of mammal, both physiologically and physically, that they can scarcely move about except when they are surrounded by its hair or body covering, and they can subsist only on its blood for any length of time. The transfer of a louse, from one host to another, can occur only when the hosts are in close bodily contact, as in the nest or at the time of mating. In the case of humans, the lice may become detached and enter the clothing or the bedding. In times of social disturbance, such as war or when people are crowded together as formerly occurred on ships, or in jails or slums, the opportunities for an exchange of these parasites are enormously increased. Under these conditions, in certain parts of the world louse-borne diseases may become epidemic and produce tremendous mortality. The close restriction of each species of lice to its special host is sufficient to account for the fact that diseases are not transferred from one species of mammal to another. This has made it possible to prevent the epidemic spread of human diseases transmitted by lice by employing methods of cleanliness and destroying the lice. One epidemic disease of rodents that is carried by lice is known from India, and there are in all probability others.

Human species. The degree to which lice are restricted to one kind of mammal or to a few closely related kinds may be indicated in the following discussion. There are two species of lice which occur as common ectoparasites upon humans. One of them is *Pediculus humanus*, the head and body louse which transmits the diseases mentioned above (**Fig. 1**). Apparently it sometimes occurs upon gibbons in zoos and has been recorded from certain New World monkeys. Other quite distinct species of the genus *Pediculus* are known from the chimpanzee and from New World monkeys. The other species occurring on humans is *Phthirius pubis*, known as the crab louse, and it is restricted as far as is known to humans, although a second species has been described from the gorilla (**Fig. 2**). It transmits no known disease. The lice of the Old World monkeys are referred to another genus, *Pedicinus*. *See* INSECTA.

Gordon F. Ferris; Dwight M. DeLong

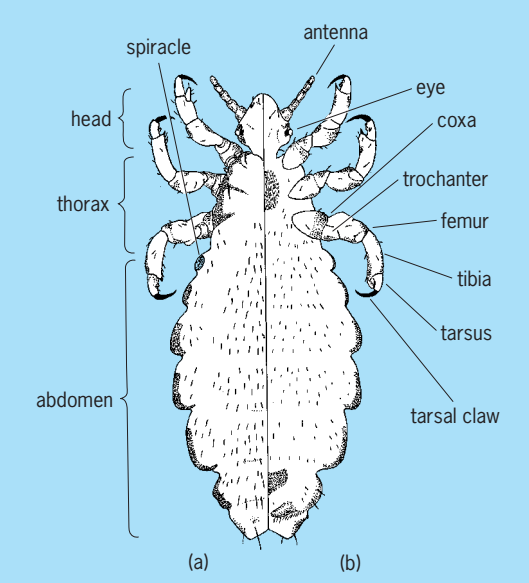

Fig. 1. The louse *Pediculus humanus*, female, shown in (*a*) dorsal and (*b*) ventral views.

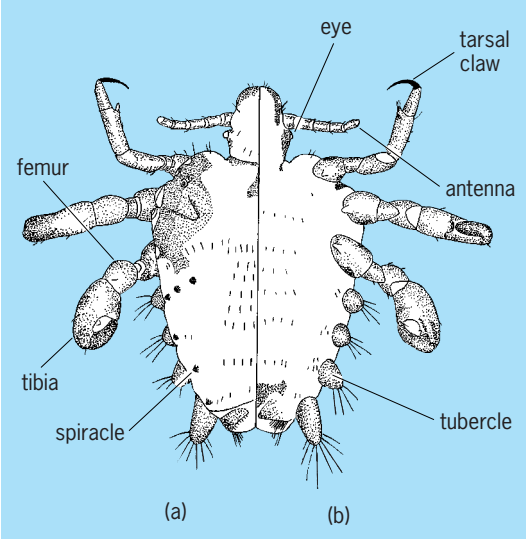

Fig. 2. The louse *Phthirius pubis*, female, shown in (*a*) dorsal and (*b*) ventral views.

Anorexia nervosa

A psychiatric disorder in which a dramatic reduction in caloric intake consequent to excessive dieting leads to significant bodily, physiological,

biochemical, emotional, psychological, and behavioral disturbances.

Clinical presentation. Anorexia nervosa is typically an illness of adolescent females: 90% of all cases begin in girls who are between 12 and 20 years of age. Nevertheless, this disorder can also occur in males, in prepubertal girls, and in women well into their third decade. Moreover, if the illness becomes chronic, it can persist into mid-life and beyond. Estimates of the incidence of anorexia nervosa are limited but indicate that about 1–5% of adolescent girls living in industrialized cultures are affected, and that the rate is rising.

Anorexia nervosa literally means "nervous loss of appetite" but appears to have little to do with such. Rather, there is usually a conscious decision made by a teen-age girl, most commonly around the ages of 14 or 18 years, to embark upon a diet. This decision often occurs in the context of some stressful life change situation. In addition, there are usually two personality characteristics: a morbid fear of losing control over one's weight (the majority of girls who develop anorexia nervosa usually are slightly overweight prior to the illness' onset) and a distorted body image (that is, the perception of oneself as larger than one is in reality).

The amount of weight lost can vary considerably. The usual criterion for making a diagnosis of anorexia nervosa is a weight change of at least 25% from premorbid weight (to at least 15% below ideal weight in persons who were overweight at the onset). But this figure should be viewed as only a rule of thumb; some persons may be affected by the disorder and do not reach the 25% weight loss. In the most extreme cases, weight losses of up to 50% and more can occur.

In addition to these core disturbances, there is an array of associated symptoms that characterize most persons with anorexia nervosa. Amenorrhea, or absence of menstruation, occurs in 95–100% of anorectic women; about one-third of the cases of amenorrhea actually precede the onset of substantial weight loss. Increased physical activity, such as excessive running, swimming, or calisthenics, is noticeable in a majority of cases; early in the course of the illness it is associated with a sense of well-being and is goal-directed, but later it becomes more obsessional and frenetic. Insomnia, whether secondary to the starvation or the emotional turmoil characteristic of the disorder, is a frequently reported symptom; in some cases, it can lead to an abuse of sedatives or tranquilizers. Use of emetics, cathartics, and diuretics is not uncommon in women who suffer from anorexia nervosa for at least several months. These practices are employed in an attempt to minimize the absorption of food, to rid the body of weight associated with body fluids, and to symbolically cleanse the self of "alien" substances. Difficulty in recognizing satiation is a common complaint, and if the illness becomes chronic, sensations of hunger may also become difficult to discriminate. Obsessional thinking and depression are not unusual in anorectic women; whether these symptoms precede the illness is not always clear, but they become increasingly apparent with chronicity.

The course and prognosis of anorexia nervosa is highly variable. While a high recovery rate, perhaps above 67%, is found in those persons whose illness begins acutely in their early teens and who quickly receive treatment, the outlook is considerably bleaker in those persons who develop the disorder later, who do not receive early treatment, and who develop bulimia. For all individuals with anorexia nervosa, the recovery rate is probably around 40%, with another 30% showing moderate improvement and the remaining 30% running a debilitated, chronic course. Included in the last group is the 5–10% of those who actually die from such complications of the disorder as cardiovascular collapse, infection, and suicide.

Bulimia. Perhaps as many as 40–50% of anorectic patients whose illness persists for more than 1–2 years will develop the additional eating disturbance known as bulimia. Persons whose anorexia starts relatively late, for example, in their twenties, may be particularly vulnerable to becoming bulimic. Literally meaning "oxhunger," bulimia refers to binge eating or compulsive overeating wherein thousands of calories are consumed in a relatively brief period of time (for example, 2 h). The episodes usually occur in the evening in private and may be no more frequent than a few times per month. However, they can occur several times per day; the average rate is around several times per week. The binge characteristically involves carbohydrates—the very type of food that the anorectic so rigidly avoids the rest of the time—and will usually culminate in self-induced vomiting. There is usually enormous shame and self-loathing associated with bulimia; whereas dieting is equated with control and strength, binging has the opposite implication. Patients characteristically wake up the next morning with even greater resolve to diet, initiating a fast–feast cycle.

The precise nature of the relationship of bulimia to anorexia nervosa remains unclear. Not all anorectics become bulimic and not all bulimics were anorectic. But dieting is common as a precursor to bulimia, and high premorbid weight and chronicity of weight loss seem to predispose the anorectic to developing bulimia. There is some indication that anorectic persons who develop bulimia are also more prone to depression (and often have a family history of depression), impulsive behavior (for example, shoplifting), and more outgoing but chaotic social relationships than those persons who remain "pure restricting" anorectics.

A major concern about bulimia relates not to the overeating per se but to the subsequent vomiting. The chronic loss of gastric juices can lead to a depletion of the body's potassium, which in turn can seriously affect cardiac function (to the point of death). In general, the onset of bulimia is believed to signify a poor prognosis.

Psychological aspects. Although there is a broad range of symptoms, personality styles, precipitants, and outcomes that characterize anorexia nervosa, there is no simple explanation of its origins. In one

widely accepted conception, the illness is viewed as a desperate struggle by the vulnerable female adolescent to establish a sense of identity separate from that of her domineering, overbearing, controlling, and intrusive mother. By virtue of the mother's insensitivity to the cues of her child, such as hunger as an infant, the future anorectic has not received clear-cut confirmation of her internal experiences. This, coupled with the mother's overintrusiveness and domination, leaves the developing child with serious deficits in her sense of self and in her ability to identify and discriminate feelings and bodily sensations. Her body image is disturbed, and she feels a potentially paralyzing sense of ineffectiveness. In the context of the typical stresses of puberty and adolescence, for example, bodily changes of portending adult responsibility and sexuality, separation from mother and home, the vicissitudes of relations with peers, social and sexual encounters with the opposite sex, and the demands of school or work—the preanorectic, who is already sensing high expectations from her achievement-oriented family, turns toward control of her body weight as the only area left in her life in which she can achieve fully independent and effective control.

There is considerable emphasis on viewing the family system as the matrix in which the illness develops and for which the illness must have significance. Four characteristics of the families of anorectics are distinguishable: excessive mutual enmeshment, mutual overprotectiveness, rigidity in relation to internal problems and the external world, and an inability to achieve resolution of conflict. Anorexia is viewed as both a response to the lack of "living space" that the adolescent experiences and as a defense against the threats to the stability of the family system that the girl's normal development implies. Development of anorexia is thus a function of both individual and family.

It should be noted, however, that these formulations suffer from a common problem. They are based on assessments of anorectic patients—and their families—after the illness has become established. Thus, these theories cannot very well differentiate among predisposing, precipitating, and sustaining factors. Moreover, many of the problems and conflicts emphasized in anorectics—such as those concerning separation and independence, appearance, adult heterosexual relations—are no different from those of normally developing adolescents. And the more pathological features described in these individuals and their families—for example, feelings of emptiness, precarious sense of self, family enmeshment and rigidity—have been reported in other pathological but nonanorectic psychiatric conditions such as schizophrenia, borderline personality, and psychosomatic disorders. *See* SCHIZOPHRENIA; STRESS (PSYCHOLOGY).

Biological aspects. There are numerous physical, physiological, and biochemical changes that reflect primarily, although not solely, the ravages of starvation. In addition to the general bodily emaciation they manifest, anorectic individuals show brittle nails and thinning hair. Cold extremities, a slow pulse, a small heart, and a hypometabolic state, in compensation for the starvation, are also characteristic. In advanced cases, edema may occur. In anorectic women who also binge and vomit, tooth decay and enlargement of the salivary glands are common. A mild-to-moderate anemia and a diminished white blood cell count develop with progressive malnutrition. Diabetes insipidus, consequent to disturbed hypothalamic function, can also occur in advanced cases, producing large amounts of dilute urine and constant thirst. In addition, abnormalities in glucose tolerance and blood electrolytes have been observed. Particularly in women who vomit, the blood potassium level can be significantly low, possibly producing abnormalities on the electrocardiogram. A small percentage of anorectic subjects show vague abnormalities in the electroencephalogram. Interestingly, these abnormalities of electroencephalogram are not always clearly correlated with the extent of emaciation. *See* MALNUTRITION; PITUITARY GLAND DISORDERS.

Because of the prominence of amenorrhea in its symptomatology, there has been a long-standing interest in the endocrinology of anorexia nervosa. A number of reliable hormonal abnormalities have been documented. Most prominent among these are diminished hypothalamic–pituitary–gonadal axis function and elevated hypothalamic–pituitary–adrenal axis function. While urinary and plasma gonadotropin levels are low and their circadian and monthly cycles are deficient, plasma and urinary levels of cortisol are high and usually will not be suppressed by dexamethasone (a synthetic adrenal steroid) administration. These abnormalities point to hypothalamic dysfunction, presumably consequent to weight loss and starvation. Possibly they also reflect functional changes secondary to affective disturbances (for example, depression), behavioral changes (for example, excessive exercise, altered sleep patterns), or some preexisting biological vulnerability. Reversal of these endocrine aberrations usually occurs with clinical improvement, although considerable time may be required for full normalization. *See* ENDOCRINE MECHANISMS.

Treatment. As with virtually all psychiatric conditions for which the etiology is unknown and where no single empirically effective treatment exists, the therapeutic approaches to anorexia nervosa are diverse and reflect the different disciplines, training biases, and experiences of their proponents.

Nevertheless, there are some general considerations in the management of this illness. First, there are probably mild cases that resolve without any intervention whatsoever. Second, despite this, anorexia nervosa is a potentially pernicious illness and, once established, can be strikingly resistant to all treatments; hence, early intervention is called for once the diagnosis is made. Third, the approach to treatment must be tailored to the stage of the illness.

In the earliest cases and youngest children, the pediatrician's involvement in a supportive, gently inquiring, and educational fashion may be sufficient.

If no improvement is seen within a few weeks, the child may be referred to a psychiatrist, but the pediatrician should remain involved. Individual, insight-oriented psychotherapy directed toward increasing confidence in identifying and accepting bodily feelings, understanding the sources of one's low self-esteem and poor sense of self, and "trying out one's wings" in the context of a reliable, supportive, respectful, and inquiring, but nonintrusive, relationship has generally been the essence of treatment for nonhospitalized individuals, particularly before chronicity has set in. Individual psychotherapy still remains a critical part of any approach to treatment, but the recovery rate can be increased, perhaps beyond 80%, by the inclusion of regular family therapy as part of the treatment approach. This implies the willingness of all the family members to participate and requires the individual to be young enough to still be involved with her family system. Hospitalization must be seriously considered whenever weight loss becomes substantial, exercising or food-related rituals begin to interfere with day-to-day functioning, or the individual begins to appear physically weak or emotionally hopeless. *See* PSYCHOTHERAPY.

In the hospital, the first priority of treatment is directed toward correcting the biological abnormalities created by the extreme dieting (and, when present, the vomiting). If the individual appears unable or unwilling to resume adequate caloric intake, despite firm but supportive nursing and concomitant individual and family psychotherapy, more extreme measures may have to be instituted, including behavior modification or intravenous hyperalimentation.

The behavior modification approach assumes that the food avoidance, whatever its original determinants, has become a habit that can best be treated by appropriate positive and negative reinforcement. Thus, a schedule is mapped out which progressively "punishes" weight loss with confinement to bed, loss of visitors, and so on, while "rewarding" weight gain with off-ward privileges, access to exercising, and such.

While intravenous fluids can correct electrolyte imbalance, they cannot provide sufficient calories and nutrition to achieve weight gain. However, the availability now of total parenteral nutrition, that is, hyperalimentation with concentrated nutrients via a peripheral or a central large intravenous catheter, has provided a technique for reversing the severe emaciation caused by advanced anorexia nervosa.

There is some evidence that, particularly in anorectics who are also characterized by depression and bulimia, antidepressant or possibly anticonvulsant medication may be helpful in damping down the binging and thereby gradually normalizing eating behavior in general. Neuroleptics (such as chlorpromazine) can also be helpful in the acute hospital management, particularly in the anxious or excessively exercising patient, but their beneficial effects are probably only indirect, that is, secondary to general tranquilization. *See* PSYCHOPHARMACOLOGY; TRANQUILIZER.

Anorexia nervosa can become a chronic illness for many of its victims. Improvement in the hospital does not mean that care need not be provided after discharge. Moreover, while a certain percentage of individuals do make a complete recovery, the real therapeutic challenge is in working with the chronic anorectic victim over many years in an attempt to minimize the morbidity imposed by the emaciation, medical complications, depression, and social isolation which often occur. *See* AFFECTIVE DISORDERS.

Jack L. Katz

Bibliography. H. Bruch, *Eating Disorders: Obesity, Anorexia Nervosa, and the Person Within*, 1979; H. Bruch, *The Golden Cage: The Enigma of Anorexia Nervosa*, 1978; P. E. Garfinkel and D. M. Garner, *Anorexia Nervosa: A Multidimensional Perspective*, 1982; S. Minuchin, B. Rosman, and L. Baker, *Psychosomatic Families: Anorexia Nervosa in Context*, 1978; J. A. Sours, *Starving to Death in a Sea of Objects: The Anorexia Nervosa Syndrome*, 1980.

Anorthite

The calcium-rich plagioclase feldspar with composition Ab_0An_{100} to $Ab_{10}An_{90}$ (Ab = $NaAlSi_3O_8$; An = $CaAl_2Si_2O_8$), occurring in olivine-rich igneous rocks and rare volcanic ejecta (for example, at Mount Vesuvius, Italy, and Miyakejima, Japan). Hardness is 6 on Mohs scale, specific gravity 2.76, melting point 1550°C (2822°F). The crystal structure of Ab_0An_{100} (triclinic space group $P\bar{1}$) consists of an infinite three-dimensional array of corner-sharing $[AlO_4]$ and $[SiO_4]$ tetrahedra, alternately linked together in a framework of $[Al_2Si_2O_8]_\infty{}^{2-}$ composition in which charge-balancing calcium (Ca^{2+}) cations occupy four distinct, irregular cavities. The crystal structure of anorthite inverts to higher symmetry ($I\bar{1}$) with each of the following: (1) heating above 242°C (468°F), (2) adding ~10 mole % Ab component (substituting $Na^+ + Si^{4+} \leftrightarrow Ca^{2+} + Al^{3+}$), (3) increasing hydrostatic pressure above ~2.6 gigapascals (26,000 atmospheres), and (4) effectively disordering the aluminum (Al) and silicon (Si) atoms at very high temperatures. Each inversion is evidenced by certain types of structurally out-of-step (antiphase) domains that can be imaged by transmission electron microscopy. Natural anorthite has no commercial uses, but the synthetic material [$CaO \cdot Al_2O_3 \cdot 2SiO_2$ (known as CAS2)] is important in the ceramic industry and in certain composite materials with high-temperature applications. *See* CRYSTAL STRUCTURE; FELDSPAR; IGNEOUS ROCKS.

Paul H. Ribbe

Anorthoclase

The name usually given to alkali feldspars which have a chemical composition ranging from $Or_{40}Ab_{60}$ to $Or_{10}Ab_{90}\pm$ up to approximately 20 mole % An (Or, Ab, An = $KAlSi_3O_8$, $NaAlSi_3O_8$, $CaAl_2Si_2O_8$) and which deviate in one way or another from monoclinic symmetry tending toward triclinic symmetry. When found in nature, they usually do not consist

of a single phase but are composed of two or more kinds of K- and Na-rich domains mostly of submicroscopic size. In addition, they are frequently polysynthetically twinned after either or both of the albite and pericline laws. It appears that they originally grew as the monoclinic monalbite phase inverting and unmixing in the course of cooling during geological times. They are typically found in lavas or high-temperature rocks. *See* FELDSPAR; IGNEOUS ROCKS.

Fritz H. Laves

Anorthosite

A rock composed of 90 vol % or more of plagioclase feldspar. Strictly, the rock is composed entirely of crystals discernible with the eye, but some finely crystalline examples from the Moon have been called anorthosite or anorthositic breccia. Two principal types of anorthosite are based on field occurrence: layers in stratified complexes of igneous rock, and large massifs of rock up to 12,000 mi^2 (30,000 km^2) in area. Scientists have been fascinated with anorthosites because they are spectacular rocks (dark varieties are quarried and polished for ornamental use); valuable deposits of iron and titanium ore are associated with anorthosites; and the massif anorthosites appear to have been produced during a unique episode of ancient Earth history (about 1–2 billion years ago).

Definition, occurrence, and structure. Pure anorthosite has less than 10% of dark minerals—generally some combination of pyroxene, olivine, and oxides of iron and titanium; amphibole and biotite are rare, as are the light minerals apatite, zircon, scapolite, and calcite. Rocks with less than 90% but more than 78% of plagioclase are modified anorthosites (such as gabbroic anorthosite), and rocks with 78–65% of plagioclase are anorthositic (such as anorthositic gabbro). *See* GABBRO.

The structure, texture, and mineralogy vary with type of occurrence. One type of occurrence is as layers (up to 10 ft or 3 m thick) interstratified with layers rich in pyroxene or olivine. The bulk composition of the layered rock samples containing anorthosite layers is gabbroic. The Stillwater igneous complex in Montana is an example where the lower part of the complex is rich in olivine, pyroxene, and chromite, and the upper part is rich in plagioclase. The second type of occurrence is the massif type. Commonly, the massifs are domical in shape and weakly layered. There appear to be two kinds of massif anorthosites: irregular massifs typified by the Lake St. John body in Quebec, and domical massifs similar to the Adirondack (New York) massif made classic through the studies of A. Buddington. The irregular massifs are older than the domical massifs and commonly contain olivine. The domical massifs are associated with silicic rocks of uncertain origin. *See* MASSIF.

Possibly there is a third group of anorthosite occurrences: extremely ancient bodies of layered rock in which the layers of anorthosite contain calcium-rich plagioclase and the adjacent layers are rich in chromite and amphibole in addition to pyroxene. There are only a few examples of these apparently igneous complexes, in Greenland, southern Africa, and India. However, they appear to be terrestrial counterparts of lunar anorthosites. Recrystallization and deformation have pervasively affected these rocks, and it is difficult to discover their original attributes.

Layered anorthosites. Zones of anorthosite in layered complexes are composed mostly of wellformed, moderately zoned crystals of plagioclase (labradorite to bytownite) up to several millimeters long. The tabular crystals of plagioclase commonly are aligned subparallel to the layering. Irregular crystals of pyroxene, olivine, and, rarely, magnetite are interstitial to the plagioclase crystals. These features are best explained by the hypothesis of accumulation of plagioclase crystals onto the floor or roof of the reservoir of melt, where gravity and perhaps motion of melt caused them to lie with their centers of gravity in the lowest position (like pennies spilled on the floor). After accumulation, the intergranular melt gradually crystallized as overgrowths of plagioclase and interstitial pyroxene. The separation of accumulated plagioclase crystals in anorthosite layers from pyroxene or olivine in pyroxenite and peridotite layers is a puzzle to scientists, for it is expected that plagioclase and pyroxene crystallize together in sub-equal proportions from basaltic magmas, not first one mineral, then the other and back again as suggested by the layering.

Compositional zoning of plagioclase within layers and other structural and textural features suggest that plagioclase floated in certain intrusions. This is consistent with calculated densities of some postulated liquids, if the liquids are poor in water. Evolution of the magma body may be dynamically complex at the stage of plagioclase flotation and anorthosite formation, particularly if new batches of magma arriving from below are less dense than the old differentiated magma. Certain unusual rocks, including the platinum-rich Merensky reef horizon of the Bushveld intrusion located near the transition between mafic layers and anorthositic layers, may be related to such dynamical complexities. *See* MAGMA.

Massif anorthosites. In massifs the plagioclase has an intermediate composition, mostly labradorite. The plagioclase crystals are poorly formed and practically unzoned and may be up to around 10 ft (3 m) long, but crystals about a centimeter long constitute the bulk of most massifs. Rare dark varieties have moderately well-formed plagioclase crystals (see **illus.**) filled with tiny specks of dark minerals. In many specimens the large labradorite crystals display a play of colors (blue and green are common) in a particular orientation, an attractive quality related to very slow cooling and sought after in ornamental stone. Lighter varieties have smaller crystals of plagioclase lacking inclusions of dark minerals, which instead occur as millimeter-sized interstitial grains. (Thus the various shades of anorthosite colors result from variations in the size or granularity of the pigmenting dark minerals.) In transitional varieties large dark crystals of plagioclase are set in a fine-grained

Outcrop of massif anorthosite on the shore of Pipmuacan Reservoir, Quebec. Exceptionally well-preserved texture shows well-formed relict plagioclase crystals (now recrystallized) outlined by irregular black pyroxene grains. Two dark cores of twinned plagioclase are shown (circles).

sugarlike white matrix. Many of the plagioclase grains are bent and broken. The associated dark minerals are pyroxene, olivine, magnetite, and ilmenite. Deposits of almost pure magnetite occur and have been used for iron ore, although most deposits have too much titanium to be highly desired for iron ore. Other deposits are rich in olivine, spinel, and apatite as well as magnetite. The massifs are surrounded by gneissic rocks, and the original age relationships between anorthosite and gneiss are commonly uncertain. The textural and mineralogical features of massifs are indicative mainly of extensive recrystallization and deformation at high temperature. The dark varieties have textures similar to those of layered anorthosites and presumably have a similar igneous origin, but this cannot be determined with certainty.

Many anorthosite massifs formed about 1.1–1.7 billion years ago. Possibly they mark a unique episode in the history of the differentiation of the Earth. The restricted age of massif anorthosite still awaits a convincing explanation.

Lunar anorthosites. By comparison with terrestrial occurrences, most lunar anorthosites are very fine grained, although one rock has crystals up to a centimeter long. Much of the fine grain size results from comminution by meteorite impact, and some of it probably results from rapid crystallization of impact melts. The unbrecciated lunar anorthosites have well-formed crystals of calcium-rich plagioclase (anorthite) intergrown with traces of olivine, spinel, pyroxene, ilmenite, metallic iron, and a few other exotic minerals. Associated rocks and minor minerals suggest that lunar anorthosites are of the layered type formed by accumulation of crystals of plagioclase from basaltic melts. Evidently, lunar plagioclases floated in iron-rich lunar basaltic melts, and most scientists agree that rocks rich in plagioclase and including anorthosite compose a major part of the skin of the Moon. *See* ANDESINE; LABRADORITE; MOON. Alfred T. Anderson, Jr.

Role in growth of continental crust. Large occurrences of Archean anorthosite are formed principally in back-arc basin or oceanic settings. Although intrusions of Archean anorthosite into continental crust are known, they seem limited in size, and lack of exposure further hampers evaluation of their contribution to continental growth. Based upon what is known of these rocks, it may be concluded that their contributions to Archean continents were vastly less than that of gray tonalitic gneiss and komatiite-basalt associates of the greenstone belts. *See* ARCHEAN.

In contrast to the Archean, Proterozoic anorthosites were emplaced into continental crust in large massifs usually localized within belts such as the Grenville Province of Canada or the Rogaland-Ergusund region of southern Norway. In these belts anorthositic massifs may account for as much as 20% of the local crust. Most of these bodies were emplaced during the time period 1650–1000 million years ago (Ma) and appear to be related to the slow, anorogenic break-up of a mid-Proterozoic supercontinent. Within the Grenville Province the anorthosite massifs appear to span the 500-million-year interval from 1650 to 1150 Ma, and it is probable that the interval consists of anorthositic pulses at 1650, 1450, and 1150 Ma. These rocks account for a minimum surface area of close to 39,000 mi^2 (100,000 km^2), or about 13% of the crustal area of the Grenville Province. If this same concentration is assumed downward throughout the continental crust, a maximum value for anorthositic additions to continental crust may be computed. For a crust 24 mi (40 km) thick this results in 1×10^6 mi^3 (4×10^6 km^3) of anorthositic addition to the regional crust over 500 million years, or an average rate of crustal addition of 0.2 mi^3/year (0.8 km^3/year). This is about 150–250 times less than the 0.3 mi^3/year (1.2 km^3/year) rate estimated for present-day arc accretion or the 0.2–0.4 mi^3/year (1–2 km^3/year) estimates for crustal accretion during the interval 1900–1700 Ma. The belt of anorthosite massifs in the Grenville Province is about 900 mi (1500 km) in length. This gives an accretion rate per mile (kilometer) of magmatic belt of about 2 $mi^3/10^6$ years (5 $km^3/10^6$ years) as compared with about 12 $mi^3/10^6$ years (30 $km^3/10^6$ years) per mile of magmatic belt for modern arcs. If the anorthosite massifs are continued downward for only 3 mi (5 km) into the crust, their rate of crustal addition becomes about 0.2 $mi^3/10^6$ years (1 $km^3/10^6$ years) per mile of magmatic belt. This is a small fraction of other crustal growth mechanisms, but it is far from insignificant. It may be concluded that within the mid-Proterozoic the emplacement of anorogenic, massif anorthosites constituted a small but significant means of continental crustal growth. *See* CONTINENTS, EVOLUTION OF; IGNEOUS ROCKS; METAMORPHISM; PROTEROZOIC. James McLelland

Bibliography. L. D. Ashwal, *Anorthosites, Minerals and Rocks*, vol. 21, Springer-Verlag, 1993; R. B. Hargraves (ed.), *Physics of Magmatic Processes*, 1980; Y. Isachsen (ed.), *Origin of Anorthosite and Related Rocks*, N.Y. State Mus. Sci. Serv. Mem. 18, 1969; A. R. McBirney, *Igneous Petrology*, 2d ed., 1992; J. McLelland, Crustal growth associated

with anorogenic, mid-Proterozoic anorthosite massifs of northeastern North America, *Tectonophysics*, 161:331–334, 1989; B. Mason and W. G. Melson, *The Lunar Rocks*, 1970; A. Reymer and G. Schubert, Phanerozoic addition sites to the continental crust and crustal growth, *Tectonics*, 3:63–77, 1984.

Anostraca

An order of branchiopod crustaceans, known as fairy shrimps and, in some cases, brine shrimps (Artemia). These organisms range up to about 4 in. (100 mm) in length, but usually are much smaller. The trunk consists of 19 to 27 segments, of which usually the first 11, sometimes the first 17 or 19, bear limbs, plus a telson bearing flattened furcal rami (see **illus**.). The trunk limbs are foliaceous, and all are of the same basic type but differ a little among themselves; each limb is differentiated into a series of endites and is usually filtratory. By beating in metachronal rhythm they propel the animal forward and draw food particles toward it. Anostracans usually swim with their ventral surface uppermost.

There is no carapace. The eyes are pedunculate. The antennae are large and modified as claspers in the male. The large mandibles are of the rolling, crushing type. The other mouthparts are very small.

Most anostracans subsist on minute particles that they sieve from the water with their trunk limbs and pass forward to the mouthparts. A few large species become carnivores as they grow, seizing small crustaceans, including in some cases smaller species of fairy shrimps.

Reproduction is usually bisexual, but some brine shrimps are parthenogenetic. Females pass eggs into a ventral brood pouch and then shed them. Eggs are drought-resistant. They hatch as nauplii which develop by a series of molts, gradually adding segments posteriorly.

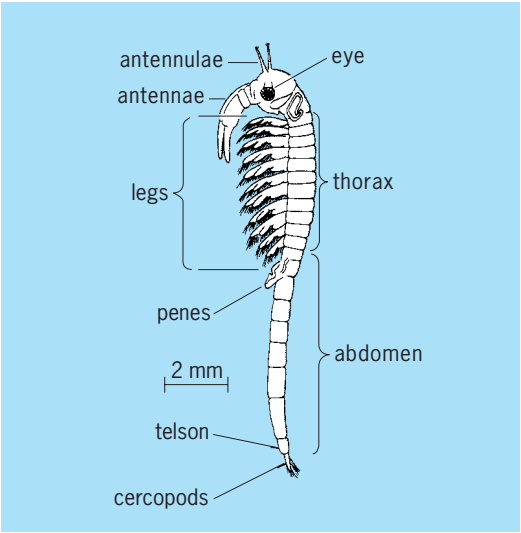

***Branchinecta paludosa*, male, small specimen, lateral aspect.**

Fairy shrimps frequent fresh water, usually temporary pools, in all regions of the world, but are also found in predator-free waters in Arctic and Antarctic regions. Brine shrimps also escape predation by living in hypersaline habitats such as salt lakes and salt pans to which they have become physiologically adapted. *See* BRANCHIOPODA. Geoffrey Fryer

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Anoxic zones

Oxygen-depleted regions in marine environments. The dynamic steady state between oxygen supply and consumption determines the oxygen concentration. In regions where the rate of consumption equals or exceeds the rate of supply, seawater becomes devoid of oxygen and thus anoxic. In the open ocean, the only large regions that approach anoxic conditions are in the equatorial Pacific between 50 and 1000 m (165 and 3300 ft) depth and in the northern Arabian Sea and the Bay of Bengal in the Indian Ocean between 100 and 1000 m (330 and 3300 ft) depth. The Pacific oxygen-depleted region consists of vast tongues extending from Central America and Peru nearly to the middle of the ocean in some places. In parts of this zone, oxygen concentrations become very low, 15 micromoles per liter (atmospheric saturation is 200–300 μmol/L). Pore waters of marine sediments are sometimes anoxic a short distance below the sediment–water interface. The degree of oxygen consumption in sediment pore waters depends upon the amount of organic matter reaching the sediments and the rate of bioturbation (mixing of the surface sediment by benthic animals). In shallow regions (continental shelf and slope), pore waters are anoxic immediately below the sediment-water interface; in relatively rapid sedimentation-rate areas of the deep sea, the pore waters are usually anoxic within a few centimeters of the interface; and in pore waters of slowly accumulating deep-sea sediments, oxygen may never become totally depleted. *See* MARINE SEDIMENTS.

Restricted basins (areas where water becomes temporarily trapped) are often either permanently or intermittently anoxic. Here, the anoxia can be a consequence of natural conditions or anthropogenic causes such as eutrophication, or a combination of them. Classic examples are the Baltic Sea, Black Sea, Gulf of Mexico, Carioca Trench off the coast of Venezuela, and fiords on the Norwegian and British Columbia coasts. Lakes that receive large amounts of nutrients or organic matter (either from natural or human-produced sources) are often anoxic during the period of summer stratification. *See* BALTIC SEA; BLACK SEA; EUTROPHICATION; FIORD; GULF OF MEXICO.

The chemistry of many elements dissolved in seawater (particularly the trace elements) is vastly changed by the presence or absence of oxygen. Since large areas of the ocean water mass are in contact with oxygen-depleted pore waters, the potential

exists for anoxic conditions to have a marked effect on the chemistry of the sea. *See* SEAWATER.

Reduction-oxidation (redox) reactions. Organic matter is formed directly or indirectly through photosynthesis, which is driven by the energy of the Sun. Carbon dioxide, CO_2, the stable form of combined carbon, is here converted to energy-rich organic carbon compounds. The required electrons for this reduction are provided by water (H_2O) through the release of elemental oxygen. Organic carbon is a very unstable form of carbon and reacts with various oxidants, yielding back the energy required for its photosynthesis. The most common of these reactions is the oxidation of organic carbon. The electrons are transferred back from carbon to oxygen, yielding CO_2 and H_2O as final products. This reaction fuels animal life on land and in the sea. In the absence of oxygen, organic matter reacts with various other oxidants (that is, electron acceptors) to form CO_2 and to gain energy. *See* PHOTOSYNTHESIS.

The major redox reactions that occur in natural systems are given below. They are listed in the order of decreasing gain of free energy (2 and 3 are interchangeable). For simplicity, CH_2O is used to represent organic carbon.

1. Oxygen reduction:

$$O_2 + CH_2O \longrightarrow CO_2 + H_2O$$

2. Nitrate reduction (denitrification):

$$4NO_3^- + 5CH_2O + 4H^+ \longrightarrow 5CO_2 + 2N_2 + 7H_2O$$

3. Manganese(IV) reduction:

$$2MnO_2 + CH_2O + 4H^+ \longrightarrow CO_2 + 2Mn^{2+} + 3H_2O$$

4. Iron(III) reduction:

$$2Fe_2O_3 + CH_2O + 8H^+ \longrightarrow CO_2 + 4Fe^{2+} + 5H_2O$$

5. Sulfate reduction:

$$SO_4^{2-} + 2CH_2O + 2H^+ \longrightarrow 2CO_2 + H_2S + 2H_2O$$

6. Fermentation:

$$2CH_2O \longrightarrow CH_4 + CO_2$$

The reaction with oxygen is the most energy-yielding and continues until O_2 is nearly entirely consumed. Nitrate, NO_3^-, and manganese, Mn(IV), reduction follow the consumption of oxygen. NO_3^- generally exists in low concentrations, $<100\ \mu$mol/L, in seawater depleted in oxygen, but Mn(IV) is insoluble and exists mainly as a solid manganese oxide, MnO_2. The importance of the latter redox reaction thus depends upon the proximity of organic carbon and MnO_2 and, as a result, the reaction will be more important in sediments than in water. The same is true for the iron–organic matter reaction, because Fe(III) is insoluble in oxic waters and exists as an oxide or hydroxide (Fe_2O_3 or FeOOH). The sulfate reaction is very important in seawater because of its high concentration (28,000 μmol/L). In a system closed from the atmosphere, sulfate represents the largest oxidant reservoir in seawater (without sediments). In anoxic parts of the water column, such as in the Baltic Sea or Black Sea or as shown in **Fig. 1***a* for a fiord in British Columbia, the concentrations of hydrogen sulfide, H_2S, increase soon after oxygen and nitrate are depleted. In anoxic water or sediment

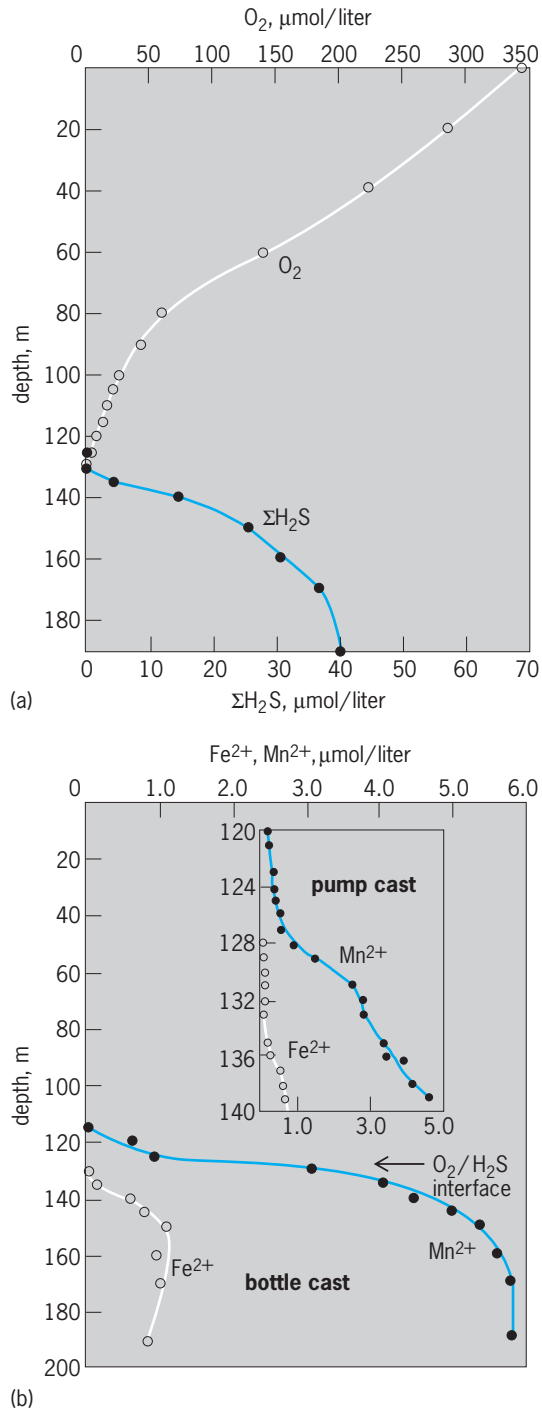

Fig. 1. Water column profiles from Saanich Inlet, an intermittent anoxic fiord in British Columbia, Canada. (*a*) Oxygen and hydrogen sulfide profiles showing the interface at 130 m (430 ft). (*b*) Dissolved iron and manganese at the same location. Pump cast indicates a more closely spaced profile at the location. 1 m = 3.3 ft. (*After S. Emerson, R. Cranston, and P. Liss, Redox species in a reducing fiord: Equilibrium and kinetic considerations, Deep Sea Res., 26A:859-878, 1979*)

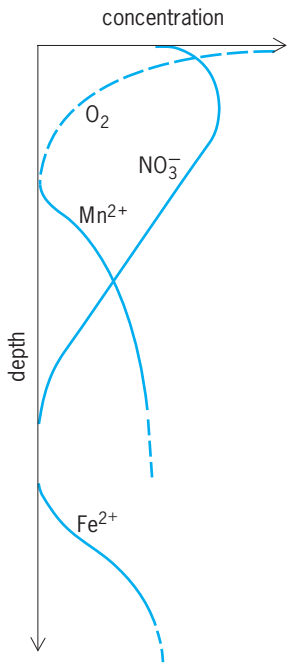

Fig. 2. Schematic representation of trends in pore water profiles. Depth and concentration are arbitrary units but represent concentrations measured or inferred (in the case of O_2) in equatorial Atlantic pore waters. (*After P. Froelich et al., Early oxidation of organic matter in pelagic sediments of the eastern equatorial Atlantic: Suboxic diagenesis, Geochim. Cosmochim. Acta, 43:1075-1090, 1979*)

samples, H_2S often can easily be smelled. After sulfate is consumed, organic matter fermentation results in the formation of CO_2 and methane, CH_4.

Chemical analysis of anoxic basins and marine-sediment pore waters reveals that the sequence of reactions shown above generally holds in nature. **Figure 2** is a schematic summary of the depth distribution of oxidants and oxidation products observed in pore waters of the equatorial Atlantic Ocean. The depth interval represents the top half meter of sediment. In these pore waters, oxygen consumption is followed by nitrate depletion and Mn^{2+} and Fe^{2+} production. In shallow-water marine sediments, rich in organic matter, the above sequence is confined to the top few centimeters, and sulfate reduction followed by methane formation dominates the pore water chemistry.

The fate of nutrients and transition metals in these systems is largely dependent upon the degradation of organic matter (nutrients and many transition metals are incorporated into organic matter at the sea surface and released during degradation); manganese and iron remobilization (both iron and manganese oxides are known to be excellent absorbers of transition metals); and sulfate reduction (some metals form very insoluble sulfides). Thus, redox processes are important to the control of the chemistry of oxidants, oxidation products, and transition metals. *See* HYDROSPHERE; OXIDATION-REDUCTION.

Reaction kinetics. Nearly all organic matter degradation reactions, which occur in natural systems, are mediated by bacteria. The above reactions are energetically favorable; however, activation energies are high enough that the rates are very slow at earth surface temperatures without bacterial catalysis. The rates of reaction are dependent upon the concentrations of organic carbon, oxidant concentration, and the number of bacteria present. In many cases, the last of these is large enough that it is not limiting, and in some cases both number of bacteria and oxidant concentration are abundant (as in some marine sediments), and the organic matter degradation reaction can be considered pseudo first-order.

The environmental oxidation rates of dissolved reduced species (such as Mn^{2+}, Fe^{2+}, NH_4^+, and H_2S), when they enter oxygenated water, are generally not well known. A case study, however, is the oxidation rate of manganese. Figure 1*a* shows the O_2 and H_2S distributions in Saanich Inlet, an intermittently anoxic fiord in British Columbia. The dissolved iron and manganese profiles are plotted in Fig. 1*b*. The difference in the rates of iron and manganese oxidation is revealed by the disparity in distance that they extend into the oxic region. The rate of iron oxidation is much faster than that of Mn^{2+}. The rate of manganese oxidation has been estimated in Saanich Inlet from the profile in Fig. 1*b* and information about water mixing at the O_2/H_2S interface. The mean lifetime for Mn^{2+} with respect to oxidation in the water, which contains O_2, is on the order of a few days. Although this is much slower than the Fe^{2+} oxidation rate, which is probably a few minutes, it is many orders of magnitude faster than rates of manganese oxidation based on abiotic laboratory experiments. Studies of the particulate matter at the O_2/H_2S interface in the water column of Saanich Inlet indicate that it is rich in bacteria, some of which have been identified as manganese oxidizers. This finding supports the hypothesis that the Mn^{2+} oxidation rate is bacterially catalyzed in natural systems. *See* CHEMICAL DYNAMICS.

Summary. The reasons that anoxic regions exist in the marine environment and the sequence of reactions, which occur during organic matter degradation, are reasonably well understood. To a first approximation, thermodynamic (energetic) predictions about the major redox reactions have been confirmed in pore waters of marine sediments. However, we have only begun to understand the kinetic aspects of organic matter degradation and oxidation of reduced species as they enter oxygenated waters. Reaction rates caused by bacterial catalysts in the environment have not been studied to the extent that quantitative generalizations can be made. This is one of the major challenges facing marine geochemists in the understanding of anoxic processes. *See* SEAWATER FERTILITY. Steven R. Emerson; Helmuth Thomas

Bibliography. F. M. M. Morel and J. C. Hering, *Principles and Applications of Aquatic Chemistry*, 1993; J. F. Pankow, *Aquatic Chemistry Concepts*, 1991; J. L. Sarmiento and N. Gruber, *Ocean Biogeochemical Dynamics*, 2006; W. Stumm (ed.), *Aquatic Surface Chemistry: Chemical Processes at the Particle–Water Interface*, 1987; W. Stumm and J. J. Morgan, *Aquatic Chemistry*, 3d ed., 1996.

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

2 ANS-BIN

www.MHEST.com

A

Anseriformes

An order of birds comprising two families, the screamers (Anhimidae) of South America and the worldwide waterfowl (Anatidae). They are closely related to the Galliformes (fowls), with the screamers being a rather intermediate group. The waterfowl and the fowls are closely related within the Aves and are often placed within a distinct superorder, the Galloanserae. The large, flightless diatrymids (giant ground birds) of the early Tertiary appear to be specialized offshoots of the Anseriformes. *See* GALLIFORMES; GASTORNITHIFORMES.

Classification. The order Anseriformes is divided into the suborder Anhimae, containing the single family Anhimidae (screamers; 3 species), and the suborder Anseres, including only the family Anatidae (ducks, geese, and swans; 147 species). The waterfowl (Anatidae) are further subdivided into seven subfamilies: the primitive Anseranatinae (magpie goose of Australia; this genus is sometimes placed in a separate family), Dendrocygninae (tropical tree ducks), Anserinae (swans and geese), Tadorninae (shelducks of the Old World and South America), Anatinae (true ducks), Merginae (mainly Northern Hemisphere sea ducks and mergansers), and Oxyurinae (stiff-tailed ducks).

Screamers. The South American screamers (*Anhima*) are turkey-sized, fowllike aquatic birds with a short, heavy, chickenlike beak. The legs are of medium length and heavy, with four toes having only a basal web. They fly slowly but soar well. Screamers live in marshes, walk on mats of floating vegetation, and sometimes swim. They feed on vegetable matter, are gregarious, and nest in solitary pairs. The nest is a pile of rushes in reed beds. The clutch is three to six eggs. Both parents incubate and care for the chicks, which are downy and leave the nest right after hatching.

Waterfowl. The waterfowl vary in size from the tiny pygmy geese (*Nettapus*) to large swans (*Cygnus*). They occur worldwide in fresh and marine (coastal) waters. All species have strong legs and feet with webbed toes and a flattened bill with comblike lamellae or teeth. The plumage is waterproof and varies from pure white to multihued to all black; females usually have a brown, cryptic plumage. The tongue is large and fleshy and serves in filter-feeding. Waterfowl feed on both plants and animals, obtained by filtering (many true ducks), grazing (swans and geese), and diving (sea ducks, mergansers, stiff-tailed ducks, and some true ducks). Most diving ducks feed on mollusks and water plants. In mergansers (*Mergus*), the lamellae of the narrowed bill are modified into toothlike structures suitable for catching fish. All species swim well, and most are strong fliers. A few, such as the Hawaiian goose (*Branta sandvicensis*), are almost completely terrestrial. Two of the three species of the South American steamer ducks (*Tachyeres*) are flightless, as were several larger fossil species.

Most species are gregarious except at breeding; the Australian black swan (*Cygnus atratus*) is an exception as it breeds in large colonies. Waterfowl are monogamous with a strong pair bond (some mate for life) and elaborate courtship. Courtship and pair formation is usually completed on the wintering ground, with the pair returning to the birth area of the female. The nest is usually solitary and can be placed on the ground, in burrows, in holes in trees, on cliffs, and in marshes; it is usually lined with down. The clutch is from 2 to 16 eggs, incubated by both sexes or in most species by the female alone. The young leave the nest right after hatching and are cared for by both sexes (for example, swans and geese) or by the female alone (most ducks). The males molt into an eclipse plumage similar to the female plumage and then into the breeding plumage. All flight feathers are molted simultaneously, leaving the bird flightless for several weeks.

Many high-latitude species migrate, wintering as large mixed flocks in more southern fresh waters

Trumpeter swan (*Cygnus buccinator*). (*Photo by Gerald and Buff Corsi; © 2000 California Academy of Sciences*)

or along the marine coast. Steller's eider (*Polysticta stelleri*) and spectacled eiders (*Somateria fischeri*) winter in large flocks in open waters within the frozen Bering Sea.

The magpie goose (*Anseranas semipalmata*) of Australia is primitive among the anatids, and is a good intermediate form between the screamers and the advanced waterfowl. Its feet are only partly webbed, and it molts its flight feathers gradually instead of simultaneously. However, it possesses the typical broadened, somewhat-flattened bill, with a thin covering of skin found in the waterfowl.

Fossil record. The screamers (Anhimidae; 3 species) do not have a fossil record. Waterfowl are first found as fossils in the early Oligocene, but not commonly until the Miocene, which is surprising because the heavy bones of aquatic anatids would be expected to fossilize readily if these birds were present. Apparently, the radiation of the anatids into the modern subfamilies and genera only took place in the Miocene, as fossils known from the early Miocene are difficult to place into modern taxa. The much older long-legged and wading *Presbyornis* (Eocene) appears to be a primitive anseriform and may be a connecting link between the Charadriiformes and the Anseriformes. *See* CHARADRIIFORMES.

Economic importance. The waterfowl are of immense economic importance. The mallard (*Anas platyrhynchos;* the common domesticated duck), muscovy duck (*Cairina mosebata*), gray-lag goose (*Anser anser;* the common domesticated goose), and swan goose (*Anser cygnoides*) have been domesticated since ancient times for their flesh, eggs, and feathers. A special breed of domesticated duck is used in southeastern Asia for ridding the rice fields of animal pests. Species of waterfowl, especially the mute swan (*Cygnus olor*), the black swan (*Cygnus atratus*), and the mandarin duck (*Aix galericulata*), are commonly kept as ornamental birds on ponds in parks and estates. The down of the common eider (*Somateria mollissima*) is collected in great quantities for quilting. Waterfowl of many species are among the most popular game birds. Some species of waterfowl have become a pest in urban areas, notably the Canada goose (*Branta canadensis*) in the eastern United States, the black swan in New Zealand, and the mallard, which has been introduced into areas in the Southern Hemisphere where it has become common and has threatened some other members of the genus *Anas* by interbreeding and competition.

Conservation and extinction. Many species of waterfowl have declined in numbers from overshooting and in more recent years from loss of habitat, both breeding areas and wintering grounds. Several species are now extinct, one example being the Labrador duck (*Camptorhynchus labradorius*) of the east coast of North America. Recovery and spread of the trumpeter swan (*Cygnus buccinator*) is an outstanding example of saving a species of waterfowl close to extinction (see **illustration**). The Hawaiian goose (*Branta sandvicensis*) has been reintroduced successfully into its native islands from stocks bred in captivity. Most species of waterfowl are carefully protected and managed under strict conservation laws.

Waterfowl are of intense interest to ornithologists, aviculturists, zoo keepers, conservationists, and game managers. They are one of the most popular groups of birds kept by zoos and aviculturists, with most species breeding readily in captivity. *See* AVES.

Walter J. Bock

Bibliography. J. Delacour et al., *The Waterfowl of the World*, 4 vols., 1954–1964; E. A. Johnsgard, *Ducks, Geese, and Swans of the World*, 1978; J. Kear, *Ducks, Geese, and Swans*, Oxford University Press, 2005.

Antarctic Circle

The Antarctic Circle is an imaginary line that delimits the northern boundary of Antarctica. It is a distinctive parallel of latitude at approximately 66°30′ south. Thus it is located about 4590 mi (7345 km) south of the Equator and about 1630 mi (2620 km) north of the south geographic pole.

All of Earth's surface south of the Antarctic Circle experiences one or more days when the Sun remains above the horizon for at least 24 h. The Sun is at its most southerly position on or about December 21 (slightly variable from year to year). This date is known as the summer solstice in the Southern Hemisphere and as the winter solstice in the Northern Hemisphere. At this time, because Earth is tilted on its axis, the circle of illumination reaches 23.50° to the far side of the South Pole and stops short 23.50° to the near side of the North Pole.

The longest period of continuous sunshine at the Antarctic Circle is 24 h, and the highest altitude of the noon Sun is 47° above the horizon at the time of the summer solstice. The long days preceding and following the solstice allow a season of about 5 months of almost continuous daylight.

Six months after the summer solstice, the winter solstice (Southern Hemisphere terminology) occurs on or about June 21 (slightly variable from year to year). On this date the Sun remains below the horizon for 24 h everywhere south of the Antarctic Circle; thus the circle of illumination reaches 23.50° to the far side of the North Pole and stops short 23.50° to the near side of the South Pole. *See* ARCTIC OCEAN.

Tom L. McKnight

Bibliography. A. H. Strahler and A. Strahler, *Introducing Physical Geography*, 4th ed., 2005; H. Veregin (ed.), *Rand McNally Goode's World Atlas*, 21st ed., 2004.

Antarctic Ocean

The Antarctic Ocean, sometimes called the Southern Ocean, is the watery belt surrounding Antarctica. It includes the great polar embayments of the Weddell Sea and Ross Sea, and the deep circumpolar belt of ocean between 50 and 60°S and the southern fringes of the warmer oceans to the north. Its northern boundary is often taken as 30°S (**Fig. 1**). The Antarctic is a cold ocean, covered by sea ice during the winter from Antarctica's coast northward to approximately 60°S.

The remoteness of the Antarctic Ocean severely hampers the ability to observe its full character. The sparse data collected and the more recent addition of data obtained from satellite-borne sensors have led to an appreciation of the unique role that this ocean plays in the Earth's ocean and climate. Between 50 and 60°S there is the greatest of all ocean currents, the Antarctic Circumpolar Current sweeping seawater from west to east, blending waters of the Pacific, Atlantic, and Indian oceans. Observed within this current is the sinking of cool (approximately 4°C; 39.2°F), low-salinity waters to depths of near 1 km (0.6 mi), which then spreads along the base of the warm upper ocean waters or thermocline of more hospitable ocean environments. The cold polar atmosphere spreading northward from Antarctica removes great amount of heat from the ocean, heat which is carried to the sea surface from ocean depths, brought into the Antarctic Ocean from warmer parts of the ocean. At some sites along the margins of Antarctica, there is rapid descent of cold (near the freezing point of seawater, −1.9°C; 28.6°F) dense water, within thin convective plumes. This water reaches the sea floor, where it spreads northward, chilling the lower 2 km (1.2 mi) of the global ocean, even well north of the Equator.

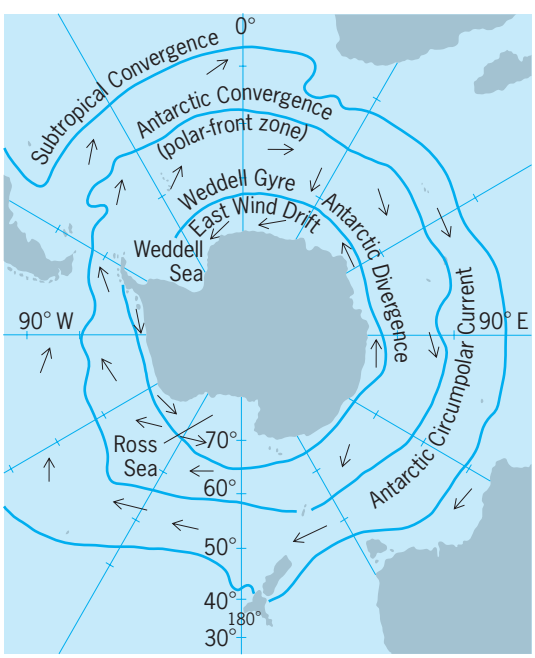

Fig. 1. Direction of the surface circulation and major surface boundaries of the Antarctic Ocean.

Oceanographic investigations. There has been much exploration of the Antarctic Ocean since the sixteenth century. Curiosity, exploration for rich unclaimed lands, and the need for new ship routes were the impetuses. There was the search for Terra Australis Incognita, the proposed great southern continent. Sightings of many of the larger islands in the southwestern Pacific Ocean were first considered to be this long-sought land, but circumnavigation proved that these were islands and, with the exception of a few of the smaller ones, not economically worth exploiting. In the 1770s, under the command of Captain James Cook, the first scientific voyages were carried out in the Antarctic Ocean and added much to the knowledge of the Southern Hemisphere's geography and climate. From the size of the icebergs observed and a latitudinal variation of the pack ice field, Cook surmised a frozen continent not symmetrical with the geographic South Pole. The nineteenth century witnessed further explorations of the Antarctic Ocean and the delineation of the coastlines of Antarctica. This series of explorations was sparked mainly by the growing seal and whaling industries.

In the twentieth century, further scientific studies have been carried out by the German ship *Deutschland* in the Weddell Sea and by the circumpolar expeditions of the English vessels *Discovery II* and *William Scoresby*. The United States has contributed much through the Navy Deep Freeze operations and the 1962–1972 circumpolar study of the Antarctic Ocean by the National Science Foundation–sponsored ship, USNS *Eltanin*. In the period 1975–1980, an intensive study of the Drake Passage was the focus of a program called International Southern Ocean Studies (ISOS). In October-November 1981, the first modern expedition well into the seasonal sea ice cover was carried out along the Greenwich Meridian in a joint U.S.-Soviet oceanographic program.

Currents. The major flow is the Antarctic Circumpolar Current, or West Wind Drift (Fig. 1). Along the Antarctic coast is the westward-flowing East Wind Drift. The strongest currents are in the vicinity of the polar front zone and restricted passages such as the Drake Passage, and over deep breaks in the meridionally oriented submarine ridge systems (**Fig. 2**).

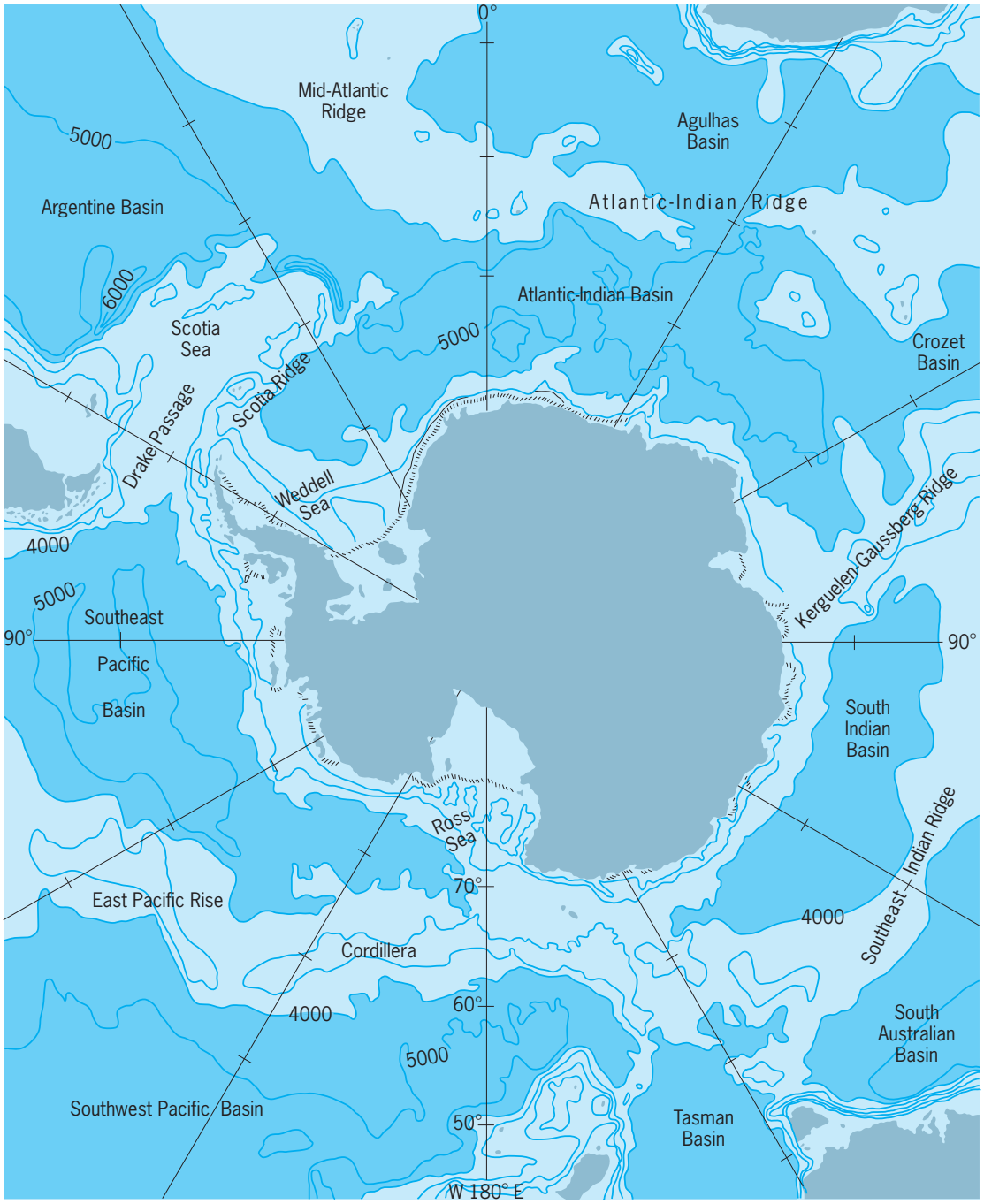

Fig. 2. Bottom topography of the Antarctic Ocean. The depths shown are in meters. 1 m = 3.3 ft.

International Southern Ocean Studies. The structure and volume transport of the Antarctic Circumpolar Current are central to the understanding of the general circulation of the Antarctic Ocean. Estimates of the total volume transport through the Drake Passage based on observations prior to 1970 varied from zero to 2.4×10^8 m^3/s (8.5×10^9 ft^3/s). To resolve this controversy and to study the dynamics and structure of the Antarctic Circumpolar Current, the International Southern Ocean Studies program was organized. Major emphasis was placed on monitoring the volume transport, determining the thermohaline and chemical properties of the waters passing through the Drake Passage, and investigating the nature and magnitude of meridional exchanges of heat and salt. As part of the 6-year program, 65 current meter moorings were maintained, and more than 600 hydrographic stations were taken during some 15 oceanographic cruises.

Transport through Drake Passage. The volume transport of an oceanic current can be decomposed into a baroclinic transport and a barotropic transport. The baroclinic transport arises from gradients in the horizontal and vertical density or mass field. The barotropic transport is due to a uniform pressure field and gives rise to a velocity which does not vary with depth.

The density of seawater is a function of pressure, temperature, and salinity and can be obtained from routine hydrographic measurements. Thus, from a series of hydrographic stations it is possible to determine the mass field and calculate the baroclinic transport. To determine the barotropic velocity and the total velocity profile, it is necessary to obtain the total velocity at least at one depth (with, for example, a current meter). By comparing this velocity with the baroclinic velocity, it is possible to determine the barotropic component and the total transport. The baroclinic transport through the Drake Passage between the free surface and the depth at 3000 decibars or 30 megapascals (approximately 3 km or 1.8 mi) is relatively constant (9.5×10^7 m^3/s or 3.4×10^9 ft^3/s $\pm 15\%$).

Though the magnitude of the current is small, there is little attenuation of flow with depth, which results in the great volume transport of the Antarctic Circumpolar Current. ISOS data show the Circumpolar Current carries about 1.25×10^8 m^3/s (4.4×10^9 ft^3/s) of ocean water through the Drake Passage. About 1×10^8 m^3/s is due to the baroclinic pressure field, the rest being barotropic. Variations from the mean of 8×10^7 m^3/s occur, closely correlated to variations in the wind field, which drives the ocean circulation.

Antarctic Circumpolar Current. Constant-density surfaces in the Drake Passage do not rise uniformly toward the south, but shoal in a series of distinct zones of large horizontal density gradients. These regions are vertically aligned and appear as a series of steplike structures, approximately 50 km (30 mi) wide (**Fig. 3**). Associated with these regions of large horizontal density gradient are high-velocity cores where the Antarctic Circumpolar Current attains surface velocities of 40 cm/s (16 in./s).

In the upper waters, these horizontal density discontinuities are associated with water-mass boundaries and in the deep waters with steplike rises of the core of Circumpolar Deep Waters. The polar front zone, the transition region between Antarctic and Subantarctic Surface Waters, is confined to the region between the two central density steps and is approximately 200 km (120 mi) wide. Small-scale mixing processes cause the Antarctic and Subantarctic

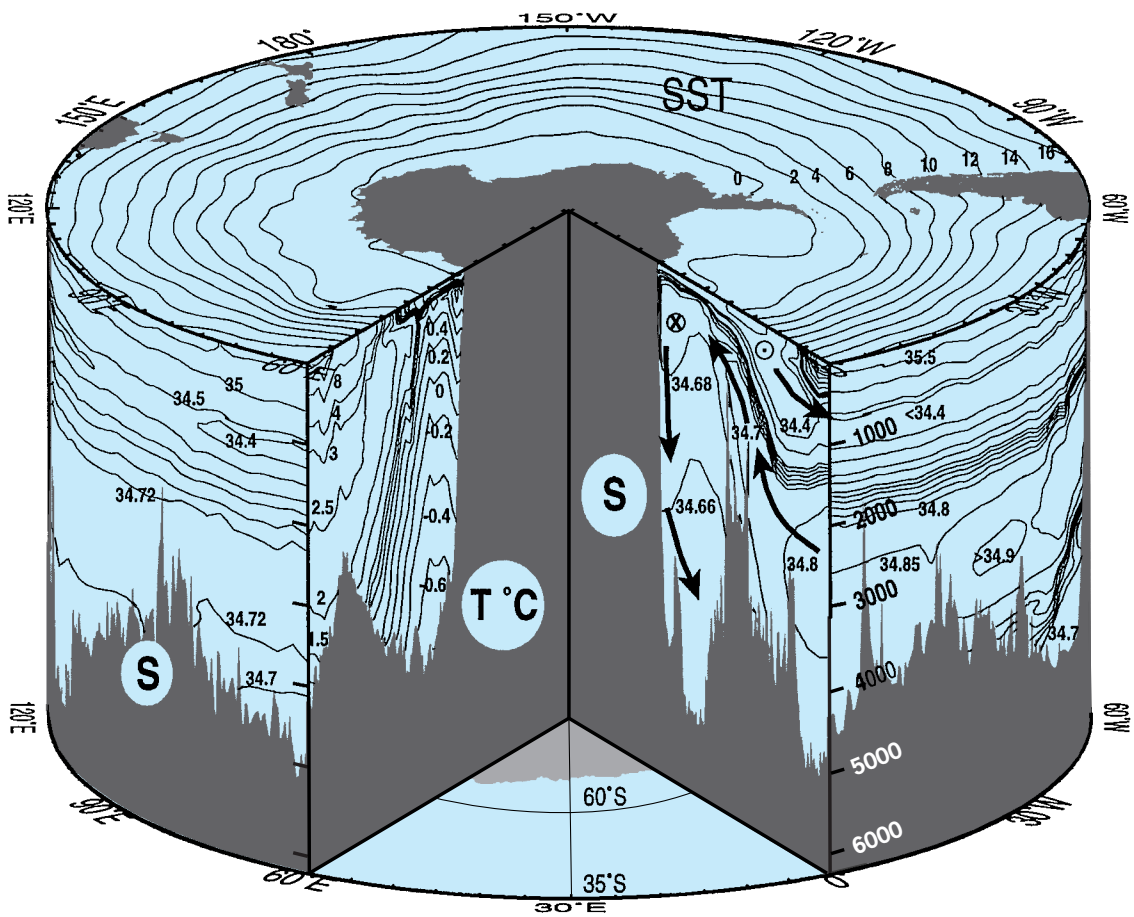

Fig. 3. Antarctic Ocean temperature and salinity. Isotherms of the annual average sea surface temperature (SST in °C) are shown on the plane of the sea surface. Winter SST adjacent to Antarctica is at the freezing point of seawater, −1.9°C. Sea surface temperature increases with distance from Antarctica. The core of the eastward flowing Antarctic Circumpolar Current and associated polar front occurs near the 4°C isotherm. The right-hand plane of the slice shows the salinity (*S*) values as a function of latitude (the 35°S and 60°S latitudinal circles are shown along the floor of the figure) and depth (shown in thousands, in meters), values in approximately parts per thousand of dissolved sea salt (‰). These data are derived from the oceanographic observations along the Greenwich Meridian shown at the floor of the figure. (*After A. Gordon, The Southern Ocean, J. Mar. Educ., 15*(3), *1999*)

-S

Surface Waters to interleave with vertical scales of order 100 m (330 ft) and less. However, these processes are relatively ineffective at transporting heat and salt across the zone.

Meanders, rings, and meridional heat and salt fluxes. The meridional transfer of heat and salt across the polar front zone is apparently accomplished by large-scale processes associated with instabilities of the Antarctic Circumpolar Current. Although the transport of the Antarctic Circumpolar Current is large and primarily toward the east through the Drake Passage, there are large disturbances called waves or meanders which give rise to north-south velocities. The meanders involve the entire water column, and have 250-km (160-mi) wavelengths and eastward phase velocities of 10–15 cm/s (4–6 in./s). These meanders are attributed to baroclinic and barotropic instabilities. Because the meanders transport colder waters northward and warm waters southward, they decrease the potential energy (baroclinic instability) associated with the inclined density surfaces and give rise to a poleward heat flux. The calculated poleward heat flux for the Drake Passage is larger than the required circumpolar averaged heat flux needed to balance heat lost by the Antarctic Ocean to the atmosphere.

Occasionally, the waves or meanders in the Antarctic Circumpolar Current become unstable, growing in time until the meander pinches off, a ring is formed, and the current and front are reestablished with a primarily zonal orientation. (Rings are commonly observed near other major oceanic current systems, such as the Gulf Stream and Kuroshio.) When a northward meander pinches off, colder Antarctic waters become enclosed by Subantarctic waters, giving rise to a cold-core clockwise-rotating cyclonic ring. Similarly, a southward meander can pinch off to form a warm-core ring (**Fig. 4**). In the Drake Passage the existence of such rings is well documented, having been photographed from satellites, studied from ships, and sensed with current meters. In other regions of the Antarctic Circumpolar Current, the existence of rings has also been inferred from hydrographic and current meter measurements.

The rings observed in the Antarctic Circumpolar

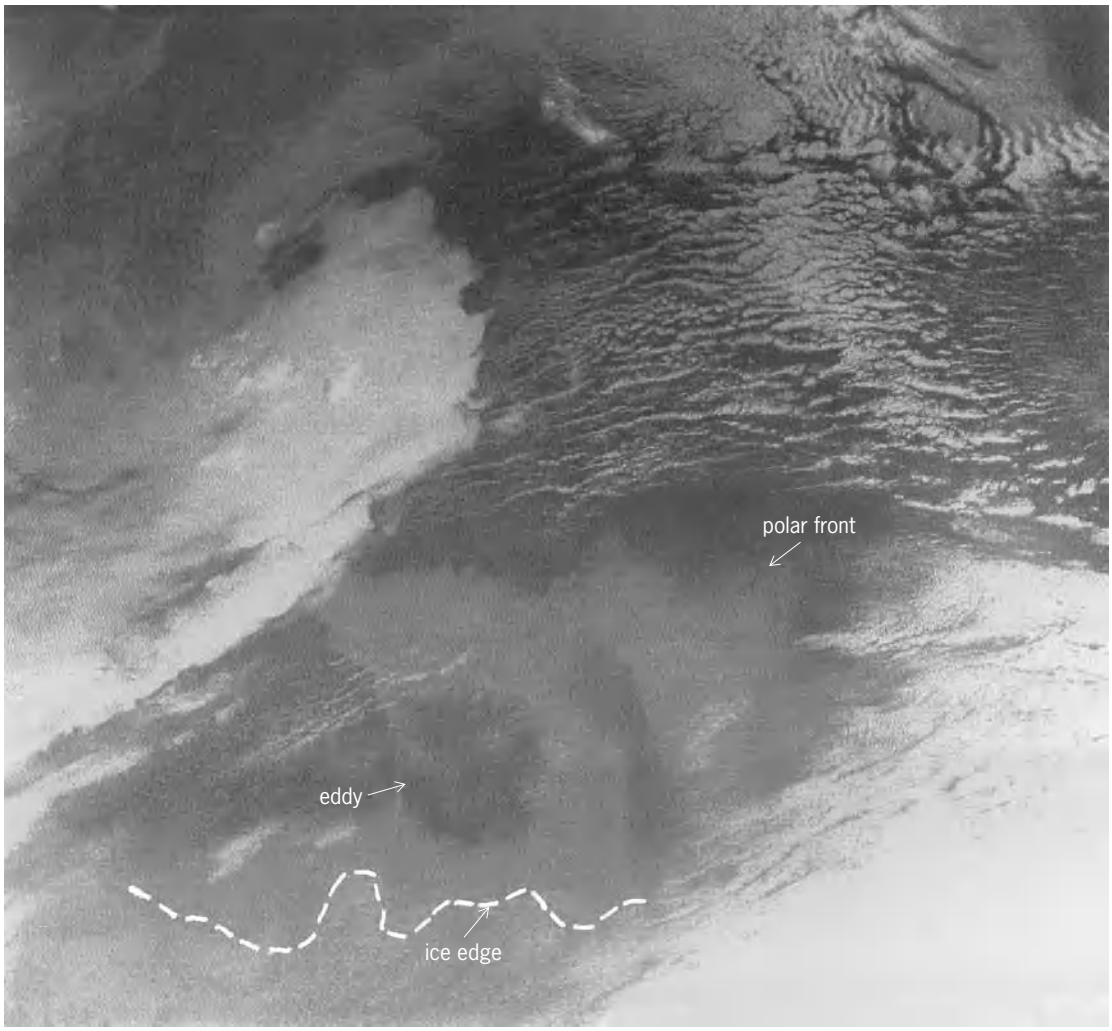

Fig. 4. Thermal infrared image from a polar orbiting satellite over the Drake Passage. The polar front, a warm–core ring (eddy), and the edge of the pack ice are indicated. (*From R. Legeckis, Oceanic polar front in the Drake Passage: Satellite observations during 1976, Deep–Sea Res., 34:701-704, 1977*)

-s

Current are typically 3 km (1.8 mi) deep and 80–120 km (50–74 mi) in diameter, involving a volume of water equal to about one-tenth the water in all the world's fresh-water lakes. The average temperature difference between the waters inside and outside a cold-core ring is 0.25°C (0.45°F). Thus the northward ejection of a cold-core ring represents a significant heat transfer. The formation of 200 such rings per year in the Antarctic Circumpolar Current would account for one-half of the heat lost by the ocean to the atmosphere south of the polar front.

Between the Antarctic Circumpolar Current and Antarctica, the baroclinic velocity is very small. Characteristic velocities of less than 5 cm/s (2 in./s) are typical. However, the deep-reaching nature of these currents yields significant volume transports. A dominant circulation feature is the clockwise-flowing Weddell Gyre, which transfers (7.6×10^7 m^3/s; 2.7×10^9 ft^3/s) water between the Antarctic Circumpolar Current and the coastal East Wind Drift. The northward-flowing limb of the Weddell Gyre, just east of the Antarctic Peninsula, meets the Pacific water carried into the Atlantic via Drake Passage, forming a strong frontal zone over the Scotia Ridge (Fig. 2), called the Weddell-Scotia Confluence.

A smaller clockwise-flowing gyre exists within the Ross Sea, and possibly east of the Kerguelen-Gaussberg Ridge (Fig. 2). These circulation cells, with the Weddell Gyre, exchange water between the circumpolar belt and the Antarctic coastal region. *See* OCEAN CIRCULATION.

The nonzonal character of the flow is associated with irregularities in the bottom topography. On approaching a ridge, the current accelerates and turns northward. It becomes more diffuse and turns southward when approaching a basin.

Water masses. The seawater comprising the Antarctic Ocean can be conveniently divided into several water masses. The criteria used for this division are based on the temperature and salinity (T/S relation) of the water. Each water mass will fall in a different region of the diagram (**Fig. 5**). In schematic form, Fig. 3 shows the positions of these water masses and the average meridional circulation of the Antarctic Ocean. Within each of the water masses is an associated core layer. This layer contains the most undiluted water characterizing the water mass. It is observed as an extreme (maximum or minimum) in temperature, salinity, or dissolved-oxygen concentration.

As observed in the T/S diagram, the coldest, freshest water is the Antarctic Surface Water. The cooling results from the loss of heat to the atmosphere, and the low salinity is due to the excess of precipitation over evaporation. The heat and salt needed to maintain this water at a constant temperature and salinity over a period of years are supplied by the wind-induced upwelling at a rate of 4×10^7 m^3/s (1.4×10^9 ft^3/s) of the southward-flowing deep water discussed above. This water mass is called the Circumpolar Deep Water.

The Antarctic Surface Water is bounded to the north by thick, homogeneous Subantarctic Surface Water, which is warmer and slightly more saline than the Antarctic Surface Water. A polar front zone separates these water masses (Fig. 1).

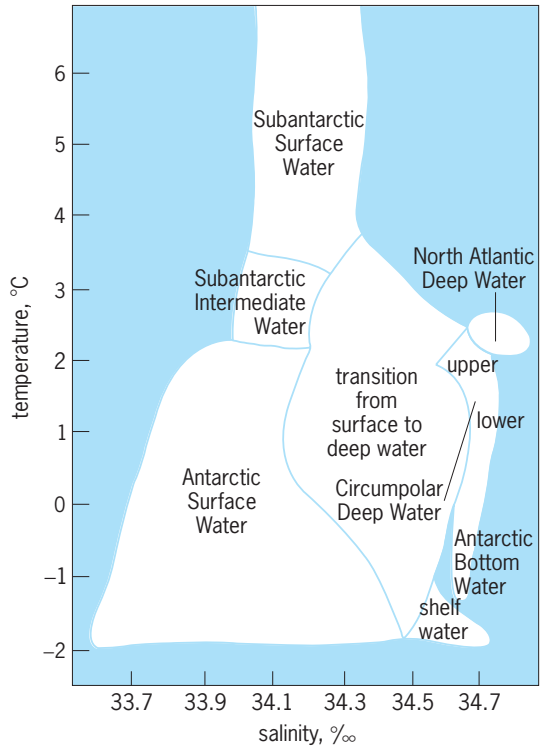

Fig. 5. Generalized diagram of the several water masses of the Antarctic Ocean; division of the water masses is based on the temperature–salinity relationship. °F = (°C × 1.8) + 32°.

The polar front zone is characterized by a large surface-temperature gradient. Its position varies in a series of meanders or waves within a 3–4° latitude band. It is also called the Antarctic Convergence, because there is occasionally evidence of a convergence of the two surface waters. During convergence, the mixture of both surface waters sinks and flows northward as Subantarctic Intermediate Water. However, later work suggests that this does not always occur; at times, signs of divergences are found, or signs of neither convergence nor divergence but rather a simple meeting of two water masses. The direct contribution of winter-cooled Subantarctic Surface Water to Subantarctic Intermediate Water, with a significant component from the Antarctic Surface Water, may also occur.

In contrast to the Arctic, the Antarctic Ocean is not landlocked, but then it is not exactly fully open to the lower warmer latitudes, as the Antarctic Circumpolar Current surrounding Antarctica effectively blocks the passage of warmer surface waters from lower latitudes into high southern latitudes. Antarctica's glacial state of the last tens of millions of years is related to the establishment of the deep circumpolar belt and the development of the Antarctic Circumpolar Current. Even at depth the Antarctic Circumpolar Current tends to hem in southern polar waters, but there are a few deep cracks in the ridge system

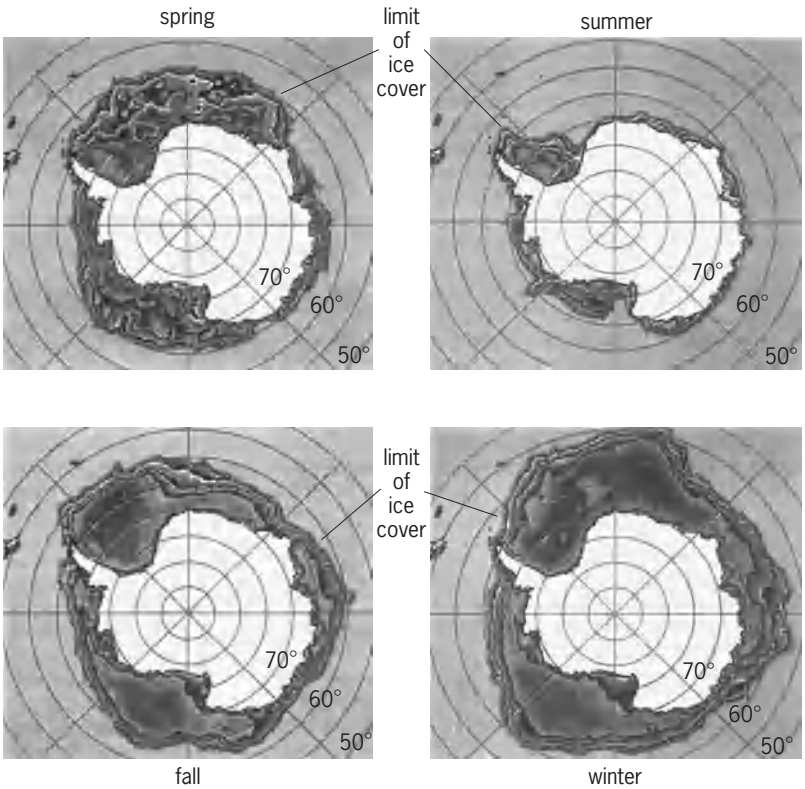

Fig. 6. Ice cover limits in the Antarctic Ocean at various seasons.

that surrounds Antarctica, allowing for meridional exchange of deep and bottom water.

Along its 21,000 km (13,020 mi) path the Antarctic Circumpolar Current transports 130×10^6 m^3 (4.6×10^9 ft^3) of seawater each second toward the east, making it the strongest of the ocean's currents. The polar front within the center of the Antarctic Circumpolar Current separates the very cold polar surface waters to the south from the somewhat warmer waters to the north. The Antarctic Circumpolar Current path and associated polar front position are aligned with the position of the maximum westerly winds, but guided to a significant degree by bottom topography, following along the northern flank of a series submarine ridges. The Antarctic Circumpolar Current and polar front reach south of 60°S, near New Zealand and near 48°S in the South Atlantic. As the ocean surface temperature pattern responds to the circulation pattern and the atmosphere responds to the sea surface temperature, there is the surprising result that the bottom topography influences the position of the maximum westerly winds. *See* SEAWATER.

Deep relatively saline water (>34.7‰) spreads poleward and upwells toward the sea surface (Fig. 3). It is balanced by northward flow of lower-salinity waters (<34.4‰) near 1 km (0.6 mi) [Antarctic Intermediate Water] and by sinking of slightly lower-salinity water along the continental slope of Antarctica. This process (saltier water in, fresher water out) removes the slight excess of regional precipitation from the Southern Ocean. The temperature is based on data collected at the same points as used for the salinity section. Shallowing of isotherms is evident as the deep water rises up toward the sea surface. There it is cooled and sinks, flooding the bottom layers with water of less than 0°C (32°F). This cold bottom water spreads well into the global ocean (Antarctic Bottom Water). The low salinity (Antarctic Intermediate Water) is shown as less than 34.45 band near 1 km (0.6 mi). More saline deep water spreading southward is seen near the 4 km (2.5 mi) depth.

At many sites along the continental margin of Antarctica, notably within the Weddell Sea (the extreme southern Atlantic Ocean), deep-reaching convective plumes of very cold surface water descend over the continental slope into the deep ocean. This Antarctic Bottom Water, high in dissolved oxygen, ventilates the bottom-layer adjacent ocean and eventually cools the lower 2 km (1.2 mi) of the global ocean. The transport of bottom water of less than −0.7°C (30.7°F) emanating from the Weddell Sea is estimated to vary from around 2 to 4×10^6 m^3/s (92 to 184×10^6 ft^3/s). Perhaps an equal amount descends at other sites around Antarctica.

Sea ice. Naturally the extreme cold of the polar regions causes an extensive ice field to form over the southern regions of the Antarctic Ocean. The extent of the ice is seasonal (**Fig. 6**) in that during the October-to-March period the area decreases, and it increases during the remaining months. The seasonal difference in the volume of sea ice is estimated as 2.3×10^{19} grams (8.1×10^{17} oz). Satellite photographs reveal that the sea ice field is not uniform, but has many large polynyas. The sea ice plays an important role in the heat balance since it reflects much more solar radiation (and therefore heat) into space than would be the case for a water surface. The polynyas would therefore be of special interest in radiation and heat-balance studies. In addition to the ice formed at sea, the ice calving at the coast of Antarctica introduces icebergs into the ocean at a rate of approximately 1×10^{18} g/year (3.5×10^{12} oz/year). *See* HEAT BALANCE, TERRESTRIAL ATMOSPHERIC; ICEBERG; SEA ICE.

In winter, a thin veneer of frozen seawater or sea ice stretches from Antarctica northward, reaching half the distance to the Antarctic Circumpolar Current. Antarctic Ocean sea ice is an important part of the global climate system, and changes in its behavior have important feedbacks to the global climate. Scientists recognize that accurate predictions of future climate requires that the ocean and atmosphere forces that govern Antarctic Ocean sea ice be understood, and that they be properly simulated in numerical global climate models.

South of the Antarctic Circumpolar Current, there is a significant transfer of the warm deep water into the surface layer of the ocean. This heat limits the thickness of the winter sea ice cover to between 0.5 and 1 m (1.6 and 3.3 ft), in sharp contrast to Arctic where upwelling of relatively warm water is small, and the sea ice is greater than 1 m (3.3 ft) thick. Upwelling of the warm deep water contributes on average 40 W/m^2 into the atmosphere during the winter

months, attaining values well above 100 W/m^2 during storm periods. As the heat is transferred to the atmosphere, cold surface water is produced. It is this water that sinks along the Antarctic margins.

Glacial (fresh-water) ice and the ocean meet along the shores of Antarctica. This occurs not only at the northern face of the ice sheet but also at hundreds of meters depth along the bases of floating ice shelves. Ocean-glacial ice interaction is believed to be a major factor in controlling Antarctica's glacial ice mass balance and stability.

During the austral winters of 1974 to 1976, near the Greenwich Meridian and 66°S in the vicinity of a seamount called Maud Rise, the ice displayed strange behavior, which has not been repeated since. A large region normally covered by sea ice in winter remained ice-free, though it was surrounded by sea ice. This remarkable climate anomaly is referred to as the Weddell Polynya.

Normally the waters over the deep regions of the Weddell Sea remain stratified during the winter, with the thick relatively warm (>0°C; 32°F) saline deep water, separated from cold (near freezing, −1.85°C; 28.7°F), fresher, thin surface layer, only slightly less dense than that of the deep water. Upwelling of the warm deep water provides heat to the surface layer and atmosphere. As long as there is enough fresh water, including that stored within the sea ice, within the surface layer to dilute and make more buoyant the cooled deep water, this configuration remains in place. The immense storage of heat within deep water is sufficient to easily remove the winter sea ice cover if it were allowed to more freely and reach the sea surface. Should upwelling of deep water somehow overwhelm the surface-layer fresh-water storage, the system would abruptly flip into a convective state. This is exactly what happened during the Weddell Polynya event. Ocean cooling to nearly 3 km (1.8 mi) was observed, indicative of an average heat flux into the winter atmosphere over the polynya of 135 W/m^2.

Sediments. The composition of the sediments are influenced by two factors: the high biological productivity in the nutrient-rich euphotic zone, and the input of glacial debris carried seaward by icebergs. The siliceous ooze (diatoms and radiolaria) is found roughly between the polar front zone and the limit of the pack ice. To the north, the siliceous ooze is replaced by calcareous ooze and red clay in deeper basins. Within the pack ice field, sediments are of glacial marine type (**Fig. 7**).

The differences in concentration of diatom and radiolaria species with depth in the ooze indicate past minor climatic changes of the Antarctic region. A major transition in sediment types is found to have occurred 1.6 × 10^6 years ago when the sediments were red clay, indicating lower productivity. This suggests that the present Antarctic circulation has existed for at least the past 1.6 × 10^6 years. *See* MARINE SEDIMENTS.

Biology. Antarctic Ocean marine mammal populations of whales and seals were the targets of much exploitation in the early twentieth century. Many

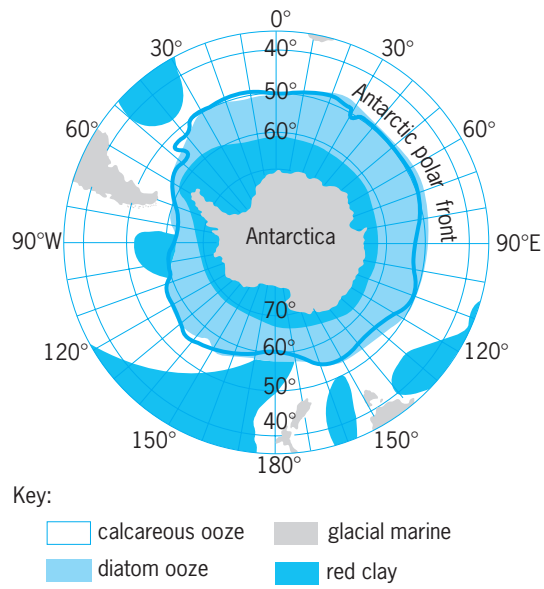

Fig. 7. Generalized distribution of the four sediment types which are found around Antarctica. (*After J. D. Hays, Quaternary sediments of the Antarctic Ocean, Prog. Oceanogr., 4:117–131, 1967*)

species were nearly depleted. The marine mammals are sustained by high productivity of plankton within the Antarctic Ocean, particularly associated with frontal zones and the diverse and complex ecosystem supported at the northern fringes of the seasonal sea ice. The estimated initial stock of 884,000 metric tons (975,000 tons) of baleen whales has diminished to 307,000 metric tons (338,000 tons), mostly among the larger (more commercial) whales.

There has been renewed interest in the living marine resources of the Antarctic Ocean which were once "whale food," particularly in the inch-long shrimplike Antarctic krill, *Euphausia superba*. Krill form large schools or swarms with densities of up to 60,000 individuals per cubic meter (1700 individuals per cubic foot) of ocean, extending from ten to hundreds of meters with perhaps giant schools of over 1 km (0.6 mi) in horizontal dimension. Significant krill populations exist near the boundaries of the Weddell Gyre, north of the Ross Sea, and over the continental margins of Antarctica at a number of sites. Estimates of the total annual sustainable yield of krill (9 × 10^6 metric tons) is nearly double the total world fish and shellfish catch.

Many nations have exploited this source of protein in the Antarctic Ocean. Because of the remoteness of the Antarctic Ocean and the seasonal sea ice cover, the full life cycle, feeding habits, and dynamics of the swarming activity of krill are not known; management of this living resource requires such knowledge. An international program called Biological Investigations of Marine Antarctic Systems and Stocks (BIOMASS) was developed in the early 1980s to investigate krill and the population dynamics, food chain, and behavior of a number of elements in Antarctic Ocean living resources, notably whales, seals, birds, fish, and squid. *See* MARINE FISHERIES.

Arnold L. Gordon

Bibliography. D. T. Georgi and J. M. Toole, The Antarctic Circumpolar Current and the oceanic heat and freshwater budgets, *J. Mar. Res.*, 40(supp.):183–917, 1982; A. L. Gordon, Polar oceanography, *Rev. Geophys. Space Phys.*, 12(5):1124–1131, 1983; A. L. Gordon, Weddell Deep Water variability, *J. Mar. Res.*, 40(supp.):199–217, 1982; A. L. Gordon and T. N. Baker, *Southern Ocean Atlas: Objective Contouring and Grid Point Data Set*, 1982; A. L. Gordon and E. M. Molinelli, *Southern Ocean Atlas: Thermohaline-Chemical Distributions and the Atlas Data Set*, 1982; W. D. Nowlin, Jr., and M. Clifford, The kinematic and thermohaline zonation of the Antarctic Circumpolar Current at Drake Passage, *J. Mar. Res.*, 40(supp.):481–507, 1982; D. Walton, *Antarctic Science*, 1987.

Antarctica

Antarctica is the coldest, windiest, and driest continent. The lowest temperature ever measured on Earth, −89.2°C (−128.5°F), was recorded at the Russian Antarctic station of Vostok in July 1983. Katabatic (cold, gravitational) winds with velocities up to 50 km/h (30 mi/h) sweep down to the coast and occasionally turn into blizzards with 150 km/h (nearly 100 mi/h) wind velocities. Antarctica's interior is a cold desert averaging annually only a few centimeters of water-equivalent precipitation, while the coastal areas average 30 cm (12 in.).

Antarctica's area is about 14 million square kilometers (5.4 million square miles), which is larger than the contiguous 48 United States and Mexico together (**Fig. 1**). Antarctica ranks fifth in size among the continents, and is followed by Europe and Australia. About 98% of it is buried under a thick ice sheet, which in places is 4 km (13,000 ft) thick, making it the highest continent, with an average elevation of over 2 km (6500 ft).

Most of Antarctica is covered by ice, and some mountains rise more than 3 km (almost 10,000 ft) above the ice sheet. The largest range is the Transantarctic Mountains (**Fig. 2**) separating East from West Antarctica. The highest peak is Mount Vinson, 5140 m (16,850 ft), in the Ellsworth Mountains. Other ranges, such as the Gamburtsev Mountains in East Antarctica, are completely buried, but isolated peaks called nunataks frequently thrust through the ice around the coast.

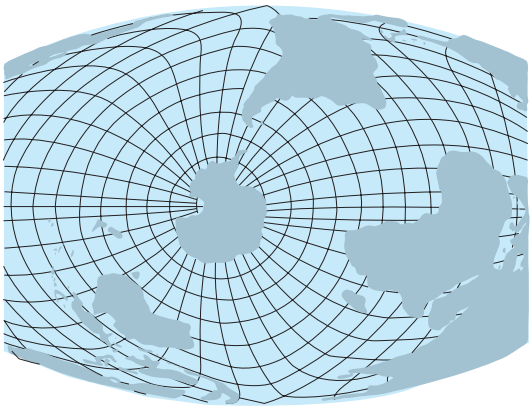

Fig. 1. Equal-area projection showing Antarctica's size in relation to that of the nearest continents.

The Antarctic continent is surrounded by a large belt of sea ice. At its maximum annual extent in September, sea ice covers an area of over 18 million square kilometers (7 million square miles); however, it melts during the summer season to a minimum extent of about 2–3 million square kilometers (1 million square miles). The ice reaches to 52°S latitude in some areas and has an average thickness of about 1.5 m (5 ft). In summer, it can generally be found only close to the coast and in the Weddell Sea.

The waters of the Southern Ocean are considered Antarctic waters south of the Antarctic Convergence. This convergence marks a sharp boundary between cold Antarctic water and warmer water to the north, and it lies roughly between 50 and 60°S latitude. Inside or close to the Antarctic Convergence lie the subantarctic islands. South of the Antarctic Convergence there are numerous icebergs; however, icebergs have also been sighted as far north as 45°S latitude. *See* ANTARCTIC OCEAN; ICEBERG.

Discovery. It is unclear who first sighted the Antarctic continent. The English navigator James Cook crossed the Antarctic Circle for the first time in his two great voyages between 1772 and 1776 and circumnavigated Antarctica, but he failed to sight the continent. Thaddeus von Bellingshausen, sent by the Russian czar to be the first to circumnavigate Antarctica since Cook, sighted land at 70°S latitude, 2°W longitude, in 1820 but did not clearly recognize the ice field he saw as land. A few days later, the Englishman Edward Bransfield sighted the Antarctic Peninsula, but his log books are now lost, making it difficult to substantiate any claims of discovery. The American sealer Nathanial Palmer sighted land almost a year earlier in the same region that Bransfield saw, but his log entries are vague.

Ice sheet. The Antarctic ice sheet is the largest remnant of previous ice age glaciations. It has probably been in place for the last 20 million years and perhaps up to 50 million years. It is the largest reservoir of fresh water on Earth, with a volume of about 25 million cubic kilometers (6 million cubic miles). If melted, Antarctic ice would raise the level of the world's oceans by about 60 m (200 ft). The ice load has depressed the continent by about 600 m (2000 ft) on average, and if it were removed, Antarctica would appear as an archipelago of scattered, mountainous islands. Glaciers flow out from this ice sheet and feed into floating ice shelves along 30% of the Antarctic coastline. The two biggest ice shelves are the Ross and Filchner-Ronne. These shelves may calve off numerous large tabular icebergs, with thicknesses of several hundred meters, towering as high as 70–80 m (250 ft) above the sea surface. The largest berg ever sighted in Antarctic waters was 335 km (208 mi) long and 97 km (60 mi) wide, about the size of Belgium. *See* GLACIOLOGY.

Fig. 2. Aerial view of the Transarctic Mountains. (*U.S. Navy photograph*)

Life on land. Year-round life on land in Antarctica is sparse and primitive. North of the Antarctic Peninsula, a complete cover of vegetation, including moss carpets and only two species of native vascular plants, may occur in some places. For the rest of Antarctica, only lichen, patches of algae in melting snow, and occasional microorganisms occur. In summer, however, numerous migrating birds nest and breed in rocks and cliffs on the continental margins, to disappear north again at the beginning of winter. South of the Antarctic Convergence, 43 species of flying birds breed annually. They include petrels, skuas, terns, cormorants, and gulls. Several species of land birds occur on the subantarctic islands. The largest and best-known of the Antarctic petrels are the albatrosses, which breed in tussock grass on islands north of the pack ice. With a wing span of 3 m (10 ft), they roam freely over the westerly wind belt of the Southern Ocean. *See* PROCELLARIIFORMES.

Ocean life. In contrast to the generally sparse and primitive lifeforms on the Antarctic continent, the Southern Ocean teems with life. Nutrient- and oxygen-rich, it is one of the most productive oceans and supports a rich and diverse food web. The base of this food web is phytoplankton (tiny sea plants), which support a large crop of zooplankton (small sea animals), about half of which are the 3–6-cm-long (1–2-in.) shrimplike krill. Krill are the primary food of baleen whales and crabeater seals, and they are also eaten in substantial quantities by penguins, flying birds, and fish. *See* PHYTOPLANKTON.

Seven species of whales frequent Antarctic waters, including the blue, fin, humpback, and right whales, all of which are protected by international convention, and the sei, sperm, and minke whales. There are also six species of seals: crabeater, Weddell, elephant, leopard, Ross, and fur seals. Some species were severely diminished in numbers by sealing but are now recovering. There are also 21 species of penguins scattered throughout the Southern Hemisphere. The most species and the largest colonies of penguins, some of them containing several hundred thousand birds, occur on the subantarctic islands. Only two species, Emperor and Adélie penguins, are found on the Antarctic continent. *See* CETACEA; SPHENISCIFORMES.

Aurora. Modern science in Antarctica encompasses many areas of research. A fascinating phenomenon studied in the polar regions is the aurora. Auroras (northern or southern lights) are caused by subatomic particles (electrons and protons) from the Sun colliding with molecules in the upper atmosphere and producing light. The typical form of an aurora is curtain- and draperylike, moving and rippling rapidly across the sky. Bands, arcs, rays, and spirals occur as auroras form. The color of the aurora is determined by the atoms and molecules present, violet for nitrogen and red and green for atomic oxygen, and the intensity is determined by the energy of the subatomic particles that strike them. Auroras interfere with communication systems, may cause power blackouts, and can perhaps even affect the climate. *See* AURORA.

Ozone hole. A discovery was made in 1983 that ozone (O_3), a gas that occurs in small quantities in the atmosphere, has been steadily decreasing over Antarctica. The reduction is greatest in October, when the polar upper atmosphere is coldest. Temporary ozone losses as high as 50% have been recorded, forming a hole of low ozone concentrations over Antarctica that is as large as the United States. In recent years, ozone concentrations have been lower,

and the size of the hole has been larger and the hole has lasted longer. The explanations for this ozone hole are complex but heavily implicate chlorofluorocarbons (CFCs), the industrial materials used as coolants in refrigerators and air conditioners and for making plastic foams. Ozone, although present in very small quantities, effectively absorbs the ultraviolet radiation in sunlight that causes sunburn and skin cancer. Ultraviolet radiation is also reported to cause cataracts and to weaken the immune system.

Ice core records. Some of the most spectacular scientific results from Antarctica in recent years have come from deep ice cores, particularly from the Russian station Vostok (**Fig. 3**), where ice core records date back about 400,000 years. The method of reconstructing past climates is through analyses of isotope ratios, since the ratio of oxygen-18 to oxygen-16 in the ice depends on the air temperature at the time when the ice was deposited as snow on the surface. Another interesting result comes from the analyses of the trace gases contained in the tiny air bubbles in the ice. As the snow turns into ice, the air is trapped and can give an indication of the chemical composition of the atmosphere at that time. Pre–industrial revolution levels of carbon dioxide (CO_2), for example, have been found to be much lower (275 parts per million) than at present (360 ppm). *See* CLIMATE HISTORY; PALEOCLIMATOLOGY.

Gondwanaland. Up until about 140 million years ago, East Antarctica formed the central block of a giant supercontinent called Gondwanaland. When this continent broke up, the continents of Antarctica, Australia, Africa, and South America, as well as the Indian subcontinent, drifted to their present positions. In reconstructing Gondwanaland, the edges of the continental shelves are matched, and rocks and fossils on either side are compared. There are many examples supporting the idea that these continents were once joined together. One crucial line of evidence arises from the discovery of about 280-million-year-old tillites—consolidated ground moraines from a former widespread and broadly synchronous glaciation of the southern continents. *See* CONTINENTS, EVOLUTION OF.

International Geophysical Year. International scientific cooperation flourished during the International Geophysical Year (IGY) which took place from June 1957 to December 1958. It was the first major worldwide scientific effort that involved Antarctica. About 50 stations were maintained by 12 nations in Antarctica, with the United States establishing a station at the South Pole and the Soviet Union establishing

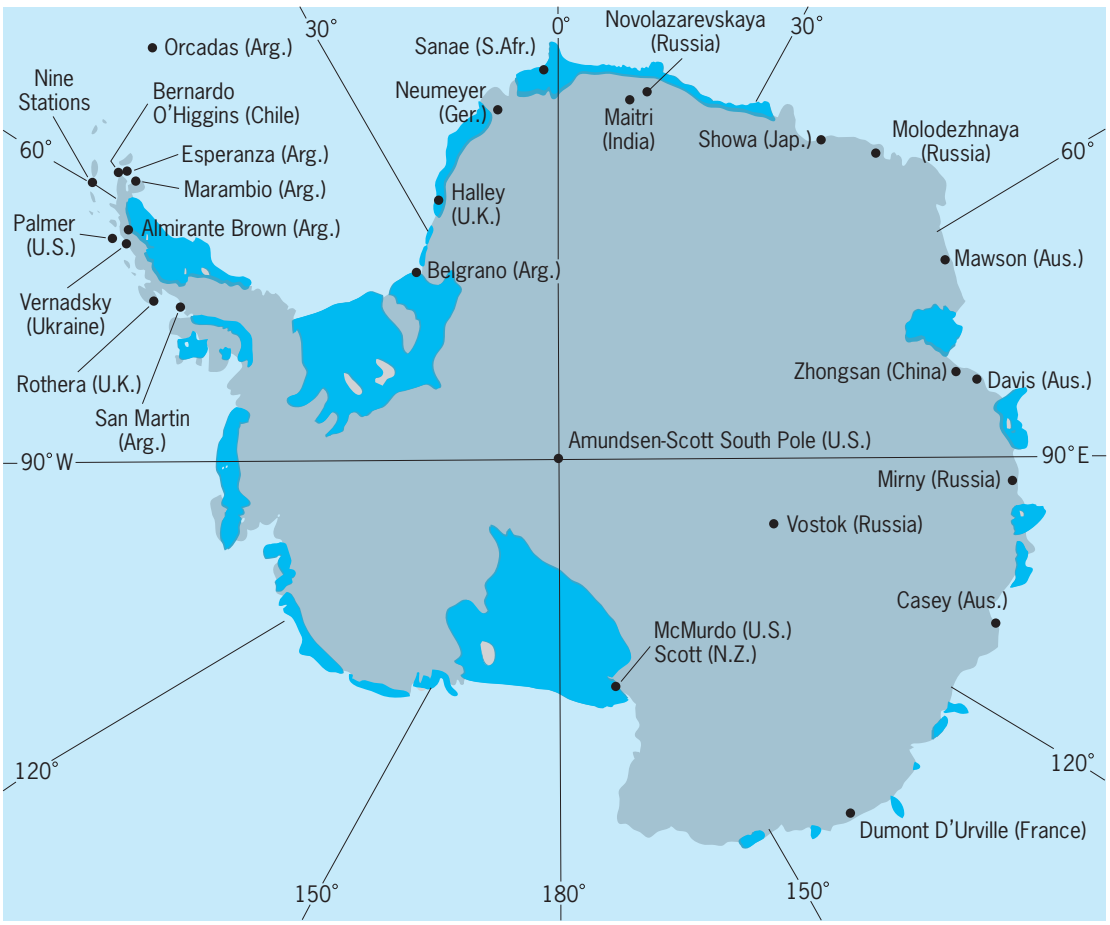

Fig. 3. Stations occupied during winter 2002. Gray areas show the continental ice sheet; color areas are ice shelves and glacier tongues. There are nine stations on King George Island: Jubany (Argentina), Commandante Ferraz (Brazil), Capitan Arturo Prat (Chile), Presidente Eduardo Frei (Chile), Escudero (Chile), Great Wall (China), King Sejong (Korea), Arctowski (Poland), and Bellinghausen (Russia).

-s

one near the geomagnetic pole. Extensive traverses were conducted with large oversnow vehicles, and the Commonwealth Transantarctic Expedition completed the first crossing of Antarctica. The IGY activities in Antarctica made a significant contribution to the study of weather and climate, the upper atmosphere, and the antarctic ice sheet and underlying bedrock. The cooperation between scientists of all the participating nations produced a lasting impact.

Antarctic Treaty. The United States convened a conference in 1958 to discuss the future of Antarctica, and a treaty was signed on December 1, 1959, by the nations that had been active in Antarctica during the IGY: Argentina, Australia, Belgium, Chile, France, Japan, New Zealand, Norway, South Africa, the Soviet Union, the United Kingdom, and the United States. The treaty temporarily freezes all territorial claims and sets aside Antarctica for peaceful purposes only. In the meantime, 32 other nations have joined the initial 12 in signing the treaty. The overriding achievement of the Antarctic Treaty has been to bring peace and political stability to a major world region. Consultative meetings have established standards for environmental protection, sharing of research results, tourism, and protection of special areas. The Antarctic Treaty has an indefinite duration and continues to provide the framework for cooperation and harmony in Antarctica. Gunter Weller

Bibliography. D. J. Drewry, R. M. Laws, and J. A. Pyle (eds.), *Antarctica and Environmental Change*, Royal Society, Clarendon Press, Oxford, 1993; R. Fifield, *International Research in the Antarctic*, Oxford University Press, New York, 1987; D. W. H. Walton (ed.), *Antarctic Science*, Cambridge University Press, New York, 1987.

Antares

α Scorpii, a cool supergiant star of spectral type M1Ib, whose red color stands out in the midsummer sky. With an effective temperature of approximately 6000°F (3600 K), Antares resembles Betelgeuse, the brightest of the red supergiants, and would fill the solar system beyond the orbit of Mars if it replaced the Sun. Antares has a distance of 185 parsecs (600 light-years) from the Sun, and its angular diameter of about 0.045 arc-second has been measured by interferometric and lunar occultation methods. Red supergiants of this type originate as stars with mass at least 20 times that of the Sun. Such stars quickly evolve through successive stages of thermonuclear fusion of heavier and heavier elements in their cores with lighter elements undergoing fusion in concentric shells surrounding the hot core. Eventually, the core builds up with iron, the most tightly bound of all atomic nuclei. The fusion of iron nuclei absorbs rather than releases energy, and the supergiant star implodes in a type II supernova explosion. This fate is likely for Antares in less than a million years. Prior to this dramatic event, Antares will have shed up to 50% of its mass through a stellar wind of material blown away from the star into the surrounding interstellar medium. *See* BETELGEUSE; SCORPIUS; SPECTRAL TYPE; STELLAR EVOLUTION; SUPERGIANT STAR; SUPERNOVA.

Antares is a member of an association of young and primarily hot stars, and is gravitationally bound in a binary star system with a hot blue star of spectral type B3V. The two stars orbit each other with a period of about 900 years, from which their masses can be determined to be about 15 and 7 times that of the Sun for the red and blue components respectively. The interaction of the strong wind of matter from the supergiant with the radiation from the less massive, hot companion produces an unusual nebulosity surrounding the hot star, contributing emission lines to that star's spectrum. Antares is also a source of radio-wavelength emission arising separately from each component of the binary. *See* BINARY STAR; NEBULA; STAR. Harold A. McAlister

Bibliography. E. E. Bloemhof and R. M. Danen, Direct measurement of the inner radius of the dust shell around the cool supergiant star α Scorpii, *Astrophys. J.*, 440: L93–L96, 1995; A. Frankoi, D. Morrison, and S. C. Wolff, *Voyages Through the Universe*, 3d ed., Brooks/Cole, 2004; J. B. Kaler, *The Hundred Greatest Stars*, Copernicus Books, 2002; J. M. Pasachoff and A. Filippenko, *Astronomy in the New Millennium*, 3d ed., Brooks/Cole, 2007; G. L. Verschuur, Journey into the Galaxy, *Astronomy*, 21(1):32–39, January 1993.

Anteater

A name associated with several animals in four different orders of living mammals (see **table**). They are so named because they are insectivorous, having a diet of ants and termites which they detect mainly by smell. These animals provide a good example of the evolutionary concept of convergence, in which similar adaptations are made by organisms of different groups. The animal most frequently associated with this name is the giant anteater (*Myrmecophaga tridactyla*), a ground-dwelling member of the family Myrmecophagidae in the order Xenarthra (see **illustration**). Three arboreal species—the northern tamandua (*Tamandua mexicana*), southern tamandua (*T. tetradactyla*), and the silky, or pygmy, anteater (*Cyclopes didactylus*)—are also classified in this family. They are found in Central and South America from Mexico to northern Argentina. Although they share membership in Xenarthra with sloths and armadillos, anteaters are the only toothless members of the order. Tamanduas and silky anteaters do possess either distinct tooth sockets or vestiges, which indicate that the teeth have been lost in the course of phylogenetic development. All members of the order Xenarthra differ from all other mammals by having xenarthrous vertebrae, lumbar vertebrae that have additional articulations (xenarthrales) to provide reinforcement for digging.

Morphology and diet. Anteaters range from 155 mm, or 6 in. (silky anteater) to about 1200 mm or 4 ft (giant anteater) in head-plus-body length. The giant anteater has a dense, coarse, stiff gray coat with

Classification of animals commonly referred to as anteaters

Scientific name	Common name	Range
Order Monotremata		
Family Tachyglossidae		
Tachyglossus aculeatus	Short-nosed echidna or spiny anteater	Australia, Tasmania, New Guinea
Zaglossus bruijni	Long-nosed echidna	New Guinea
Order Dasyuromorphia		
Family Myrmecobidae		
Myrmecobus fasciatus	Numbat or banded anteater	Australia, New South Wales
Order Pholidota		
Family Manidae		
Manis pentadactyla	Chinese pangolin	Southeast Asia
M. crassicaudata	Indian pangolin	Southeast Asia
M. javanica	Malayan pangolin	Southeast Asia
M. tricuspis	Tree pangolin	Senegal, Kenya, Zambia
M. gigantea	Scaly anteater or giant pangolin	Senegal, Uganda, Angola
M. temmincki	Cape pangolin	Chad, Sudan, South Africa
M. tetradactyla	Long-tailed tree pangolin	Senegal, Uganda, Angola
Order Xenarthra		
Family Myrmecophagidae		
Myrmecophaga tridactyla	Giant anteater	Central and South America
Tamandua mexicana	Northern tamandua or lesser anteater	Mexico to Peru
T. tetradactyla	Southern tamandua or lesser anteater	South America
Cyclopes didactylus	Silky anteater	Central and South America

black and white shoulder stripes. Adults weigh 22–39 kg (48–86 lb), although animals in zoos may weigh as much as 60 kg (132 lb). The dense, short fur of tamanduas is tan to dark brown with patches of black or reddish brown on the trunk. Completely black or dark brown individuals are found in some regions of South America. They weigh 3.2–7 kg (7–15.5 lb). The silky anteater has soft, silky gray to yellowish-orange fur with a darker middorsal stripe. It weighs 375–410 g (13.2–14.4 oz).

Anteaters have elongated, tapered snouts and tubular mouths. The long, narrow, rounded tongues are longer than their heads. The tongue is covered in minute, posteriorly directed spines (filiform papillae) and is wet with saliva when the animal is feeding. It is not sticky as is sometimes reported. The retractor muscle of the tongue is highly efficient, with motion pictures showing 160 tongue strokes a minute by a giant anteater at a termite mound. Approximately 35,000 ants and/or termites are consumed daily. Their keen sense of smell allows anteaters to be selective in choosing which anthills or termite mounds are suitable for clawing open for food. Neither anthills nor termite mounds are destroyed during feeding, as was once believed. Rather, the anteater takes only a portion of the population, thus conserving its source of food. The stomach does not secrete hydrochloric acid as in most mammals, but depends on the formic acid of the ingested ants to assist with digestion.

Giant anteater (*Myrmecophaga tridactyla*). (*Photo © 2001 by John White*)

Tamanduas may occasionally eat small fruits, while silky anteaters supplement their diet with bees and wasps. The ears are short and rounded and the eyes are small. The muscular, prehensile tail and long claws (up to 400 mm or 16 in.) of the arboreal tamanduas and silky anteater are used to grip branches. The strong foreclaws in all four species are also used to open ant and termite nests and as defensive weapons. When walking, only the outer edge of the hand, with claws retracted, makes contact with the ground and only half the sole of the hind foot makes contact. The four toes of the hind foot of the silky anteater are joined by a callused sole whose surface broadens on the inside to make an elastic rounded cushion for grasping. The genus name *Cyclopes* (Greek *kyklos* = "circle"; Latin *pes* = "foot") alludes to this feature. When agitated, posterior glands of tamanduas produce a musky secretion that has a very disagreeable odor.

Ecology and distribution. Giant anteaters are diurnal and prefer grassland, swamps, and lowland tropical forest. They swim well, even crossing wide rivers. A study of giant anteaters in Brazil showed a

population density of 1.2 individuals per square kilometer. Each giant anteater lives a solitary existence during most of the year. Contact between the sexes occurs only at the time of estrus. Adults maintain their own territories, which may overlap considerably. They sleep in a shady trench or depression on the ground. Giant anteaters have the lowest recorded body temperature for a placental mammal, 32.7°C (90.9°F). The species is classified as "vulnerable" by the International Conservation Union (IUCN) and is on Appendix 2 of the CITES.

Tamanduas are also diurnal and inhabit savannas, thorn scrub, and wet and dry forests, whereas silky anteaters inhabit tropical forests. The silky anteater is the most arboreal of the four species and will not voluntarily descend to the ground. This squirrel-sized, nocturnal species spends the daylight hours usually curled up in a ball in the fork or hollow of a kapok tree.

Reproduction and development. Giant anteaters mate in the fall and give birth to a single precocial young in the spring. Gestation ranges from about 150 to 190 days. A newborn weighs approximately 1.2 kg (2.5 lb). Immediately following its birth, the baby will climb onto its mother's back and be carried wherever she goes. Mothers may carry their young until they are half-grown, at which time they are almost as big as the mother. The young nurse for about six months and may remain with their mother for as long as two years. Little is known about reproduction in tamanduas and silky anteaters. Tamanduas usally give birth to a single young in the spring following a gestation period of 130–190 days. The young is carried on the mother's back and remains with its mother for about one year. It is known that both silky anteater parents care for their young by regurgitating crushed insects to feed the baby. *See* MONOTREMATA; PHOLIDOTA.

Donald W. Linzey

Bibliography. D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Antelope

The name given to a group of African and Asian hoofed ruminants of the order Artiodactyla. The 91 species in 31 genera are classified in five subfamilies (Bovinae, Cephalophinae, Hippotraginae, Antilopinae, and Caprinae) in the family Bovidae. This group includes the bongos and kudus (*Tragelaphus*), elands (*Taurotragus*), duikers (*Cephalophus* and *Sylvicapra*), waterbucks (*Kobus*), reedbucks (*Redunca*), roan and sable antelopes (*Hippotragus*), oryx (*Oryx*), wildebeests or gnus (*Connochaetes*), impalas (*Aepyceros*), blackbucks (*Antilope*), gerenuks (*Litocranios*), gazelles (*Gazella*), and springbucks (*Antidorcas*). One genus (*Pseudonovibos*) was described only on the basis of a series of 14 horns from Cambodia and Viet Nam; the living animal has never been seen by scientists.

Eland bull (*Taurotragus derbianus*).

Morphology. Antelopes vary in size from the royal antelope (*Neotragus pygmaeus*), which has a head and body length of about 500 mm (19–20 in.), a shoulder height of 250–305 mm (10–12 in.), and a weight of 2–3 kg (4–6.5 lb), to the giant eland (*Taurotragus derbianus*; see **illustration**) which may have a head and body length of 1800–3450 mm (70–134 in.), a shoulder height of 1000–1800 mm (39–70 in.), and a weight of 400–1000 kg (875–2200 lb). The royal antelope is the second smallest hoofed mammal in the world, after some individuals of Asiatic chevrotains or mouse deer (*Tragulus javanicus*) that inhabit southeast Asia. *See* CHEVROTAIN.

All adult male antelopes and most adult females possess unbranched horns composed of a bony core attached to the frontal bones of the skull and covered by a hard sheath of horny material. With the exception of the four-horned antelope (*Tetracerus*) of India and Nepal, all individuals possess a single pair of horns. The length and configuration of the horns varies considerably among the various species of antelopes. The dental formula is I 0/3, C 0/1, PM 3/3, M 3/3 × 2 for a total of 32 teeth. The lower incisors generally project forward and serve as cutting teeth, while the upper lip and tongue serve the function of upper incisors. These mammals feed by twisting grass, stems, and leaves around the tongue and cutting the vegetation off with the lower incisors. All ruminants have a four-compartment stomach and chew the cud. When food is initially swallowed, it goes into the largest of the four compartments (rumen) where it is acted upon by bacteria. It is later regurgitated while the animal is resting, chewed, and swallowed a second time for thorough digestion. The ulna and fibula are reduced, and the main foot bones are fused into a cannon bone. On each foot the third and fourth toes are well developed, whereas the second and fifth toes are reduced or absent. *See* DENTITION.

Reproduction and development. Reproduction may be seasonable in some species, whereas in others mating may occur throughout the year. Gestation periods range from 168 to 277 days. There is normally a single offspring, but twins are frequent in some species of gazelles. Sexual maturity is attained

between 6 (dwarf antelope) and 48 (eland) months of age. Longevity may be over 23 years (eland).

Ecology. Antelopes are grazers and browsers. Depending on the species, antelopes may inhabit grasslands, brushy areas, wooded savannahs, open forests, and, in the case of the klipspringer (*Oreotragus*) and beira (*Dorcatragus*), even rocky hills, mountains, and cliffs. Most are gregarious, although some (*Oreotragus, Ourebia, Raphicerus, Neotragus, Madoqua*, and *Dorcatragus*) are usually solitary and territorial.

Excessive hunting, excessive overgrazing by domestic livestock, agricultural development, and other habitat modifications have adversely affected most populations of antelopes. The International Conservation Union (IUCN) has designated some subspecies and/or populations of most genera as critically endangered, endangered, threatened, conservation dependent, or vulnerable. *See* ARTIODACTYLA; MAMMALIA.

Donald W. Linzey

Bibliography. D. Macdonald, *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Antenna (electromagnetism)

The device that couples the transmitter or receiver network of a radio system to space. Radio waves are used to transmit signals from a source through space. The information is received at a destination which in some cases, such as radar, can be located at the transmitting source. Thus, antennas are used for both transmission and reception. *See* RADAR.

Size and configuration. To be highly efficient, an antenna must have dimensions that are comparable with the wavelength of the radiation of interest. At long wavelengths such as the part of the spectrum used in broadcasting (a frequency of 1 MHz corresponds to a free-space wavelength λ of 300 m), the requirement on size poses severe structural problems, and it is consequently necessary to use structures that are portions of a wavelength in size (such as 0.1 λ or 0.25 λ). Such antennas can be described as being little more than quasielectrostatic probes protruding from the Earth's surface. The characteristics of antenna properties can be expressed in terms of the ratio of the dimensions of the antenna to the wavelength of operation. Therefore, in describing antennas, the dimensions are given in wavelengths.

In order to control the spread of the energy, it is possible to combine antennas into arrays. As the wavelength gets shorter, it is possible to increase the size of the antenna relative to the wavelength; proportionately larger arrays are also possible, and techniques that are familiar in acoustics and optics can be employed (**Fig. 1**). For example, horns can be constructed with apertures that are large compared with the wavelength. The horn can be designed to make a gradual transition from the transmission line, usually in this case a single-conductor waveguide, to free space. The result is broadband impedance characteristics as well as directivity in the distribution of energy in space. Another technique is to use an elemental antenna such as a horn or dipole together with a reflector or lens. The elemental antenna is essentially a point source, and the elementary design problem is the optical one of taking the rays from a point source and converting them into a beam of parallel rays. Thus a radio searchlight is constructed by using a paraboloidal reflector or a lens. A very large scale structure of this basic form used as a receiving antenna (together with suitably designed receivers) serves as a radio telescope. Antennas used for communicating with space vehicles or satellites are generally large (compared to wavelength) structures as well. *See* RADIO TELESCOPE; SPACE COMMUNICATIONS; TRANSMISSION LINES; WAVEGUIDE.

Types of problems. Antenna problems are frequently formulated in one of the following forms. (1) Given a specific radiating system, it is required to find its directional characteristics, polarization of fields, gain, input impedance, and other important properties which need to be known in order that the communication system may be designed. (2) Given a directional characteristic, for example, a certain beam-width, polarization, and beam orientation, it is required to find or select a radiating system that will produce it. Certain other constraints such as impedance behavior and physical size will have to be met. This antenna problem may be classed as a synthesis problem.

Elemental dipole antenna. Radiation properties of many antennas may be obtained by assuming that the current distribution on the antenna consists of a superposition of current elements of appropriate magnitude and, where necessary, of appropriate phase. With knowledge of the fields produced by a current

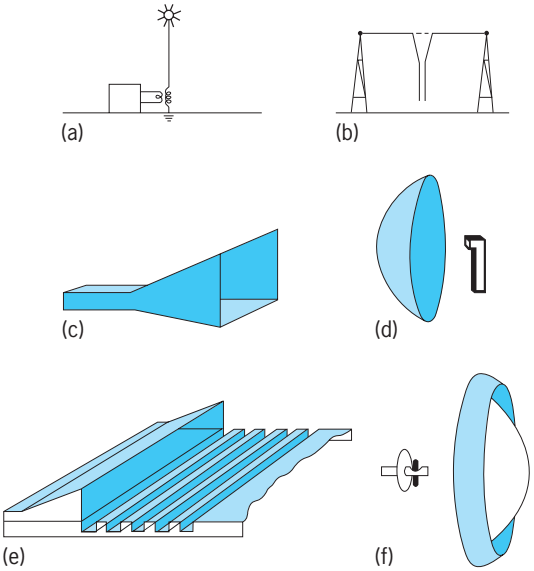

Fig. 1. Various types of antennas. (*a*) Top-loaded vertical mast; (*b*) center-fed horizontal antenna; (*c*) horn radiator; (*d*) paraboloidal reflector with a horn feed; (*e*) corrugated-surface wave system for end-fire radiation; (*f*) zoned dielectric lens with a dipole-disk feed. *(After D. J. Angelakos and T. E. Everhart, Microwave Communications, Krieger, 1983)*

-s

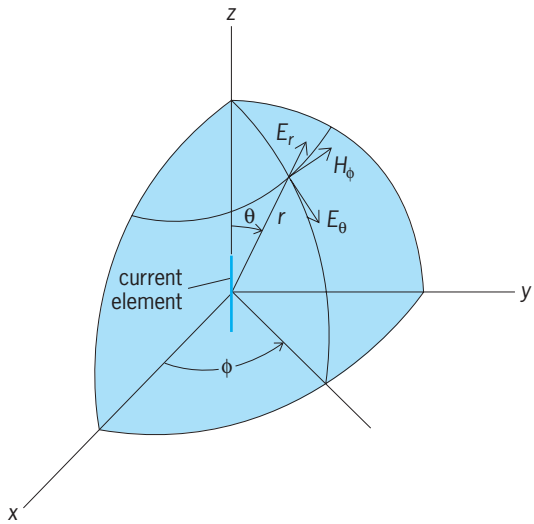

Fig. 2. Spherical coordinate system for an electric dipole.

element, by superposition it is then possible to obtain the composite fields of the antenna. Care must be taken to add the fields with proper consideration for their vector and phasor properties. An equivalent current source may also be defined for aperture-type antennas. For these reasons and because additional characteristics of antennas may be illustrated by the simple current element, the elemental dipole is a good starting point in the study of antennas.

For a differential current element of cosinusoidal time variation oriented along the z axis, the electric (E) and magnetic (H) field components, in a spherical coordinate system (**Fig. 2**), are given by Eqs. (1)–(3), where the differential current element

$$E_r = 60\beta^2 Idz\left[\frac{\cos(\omega t-\beta r)}{(\beta r)^2} + \frac{\sin(\omega t-\beta r)}{(\beta r)^3}\right]\cos\theta \quad (1)$$

$$E_\theta = 30\beta^2 Idz\left[\frac{-\sin(\omega t-\beta r)}{\beta r} + \frac{\cos(\omega t-\beta r)}{(\beta r)^2} + \frac{\sin(\omega t-\beta r)}{(\beta r)^3}\right]\sin\theta \quad (2)$$

$$H_\phi = \frac{\beta^2}{4\pi} Idz\left[\frac{-\sin(\omega t-\beta r)}{\beta r} + \frac{\cos(\omega t-\beta r)}{(\beta r)^2}\right]\sin\theta \quad (3)$$

is a current $I\cos(\omega t)$ in a very short piece of wire of length dz, r is the radial distance to the observation point, θ and ϕ are the angular coordinates, $\beta = 2\pi/\lambda$, $\omega = 2\pi f$, and λ and f are, respectively, the wavelength and frequency of the radio wave. *See* COORDINATE SYSTEMS; ELECTROMAGNETIC RADIATION; ELECTROMAGNETIC WAVE TRANSMISSION.

The product $\beta r = 2\pi r/\lambda$ represents the distance in terms of wavelengths. Hence for large values of r/λ ($\beta r \gg 1$) only the $1/(\beta r)$ terms are retained. The far-field components of significance are then given by Eqs. (4) and (5). The first part of the field expres-

$$E_\theta = \frac{-60\pi Idz}{\lambda}\left[\frac{\sin\theta}{r}\right]\sin(\omega t-\beta r) \quad (4)$$

$$H_\phi = -\frac{Idz}{2\lambda}\left[\frac{\sin\theta}{r}\right]\sin(\omega t-\beta r) = \frac{E_\theta}{120\pi} \quad (5)$$

sions shows that the magnitude of the intensity depends on the ratio of the differential current element to the wavelength. The second part, in square brackets, is the variation of field intensity with angle and distance from the element. It represents the radiation pattern of the elemental antenna. The last term contains the wave phenomenon, a radio wave traveling from the current element outward in a radial direction (**Fig. 3**).

For a small magnetic dipole (loop), the far-field components are given by Eqs. (6) and (7), where dA

$$E_\phi = 120\pi^2\frac{IdA}{\lambda^2}\left[\frac{\sin\theta}{r}\right]\cos(\omega t-\beta r) \quad (6)$$

$$H_\theta = -\pi IdA\left[\frac{\sin\theta}{r}\right]\cos(\omega t-\beta r) = \frac{-E_\phi}{120\pi} \quad (7)$$

is the differential area of a loop of current I. The radiation pattern of the magnetic dipole is the same as that

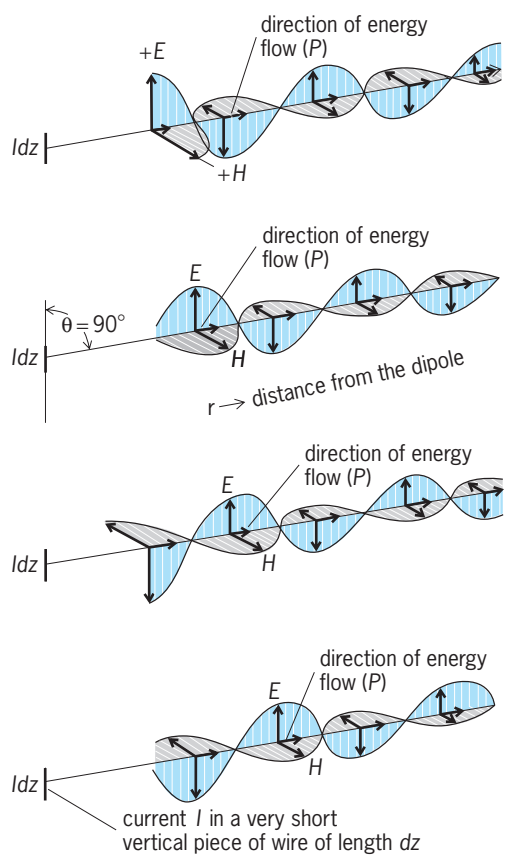

Fig. 3. Radiation field of an elemental dipole at various instants in a cycle, for a position in the plane that perpendicularly bisects the dipole, and at a radical distance *r* that is large compared to the wavelength.

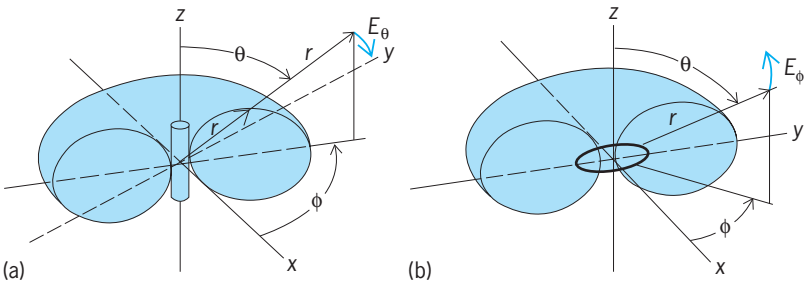

Fig. 4. Far-field patterns. (*a*) Small electric dipole. (*b*) Small magnetic dipole.

of the electric dipole, except that the electric field and magnetic field polarizations are interchanged.

Pattern shape. A small electric or magnetic dipole radiates no energy along its axis, the contour of constant energy being a toroid (**Fig. 4**). The most basic requirements of an antenna usually involve this contour in space, called the radiation pattern. The purpose of a transmitting antenna is to direct power into a specified region, whereas the purpose of a receiving antenna is to accept signals from a specified direction. In the case of a vehicle, such as an automobile with a car radio, the receiving antenna needs a nondirectional pattern so that it can accept signals from variously located stations, and from any one station, as the automobile moves. The antenna of a broadcast station may be directional; for example, a station in a coastal city would have an antenna that concentrated most of the power over the populated land. The antenna for transmission to or from a communication satellite should have a narrow radiation pattern directed toward the satellite for efficient operation, preferably radiating essentially zero power in other directions to avoid interference. Each special application of radio waves has its unique requirements for radiation patterns. *See* DIRECTIVITY; RADIO BROADCASTING.

For highly directional beam-type patterns, the concept of directivity is useful in the description of pattern shape. A simple approximation is based on beam area B at the half-power beam width of the pattern, that is, at the angles to the sides of the central beam axis where the radiated power is half the power radiated along the beam axis. Beam area b is defined as the solid angle through which all power would be radiated if the power were constant over the complete solid angle and at maximum level. For these hypothetical conditions, directivity D is defined as the surface of the sphere (4π radians) divided by beam area B in radians, as shown in Eq. (8). If the

$$D = \frac{4\pi}{B} \tag{8}$$

beam is described in terms of its two widths, θ and ϕ (in degrees), then $D = 41{,}253/(\theta\phi)$. These approximations neglect power radiated from the secondary lobes and power that is lost in the antenna conductors. Keeping the secondary lobes small is important to reduce interference in communications systems having adjacent cofrequency transmitters.

Total radiated power P_r from an antenna can be computed by integrating the energy flow outward through an imaginary sphere surrounding the antenna system. For a hypothetical spherical radiation pattern surface which has a radius of r, the radiated power P_r is given by Eq. (9), where E_s is the field

$$P_r = \frac{4\pi r^2 E_s^2}{120\pi} \tag{9}$$

strength of the radiated electric field at the surface of the imaginary sphere, and 120π is the characteristic resistance of free space. If the radiated power is 1 kW, the field strength E_s at the surface of the imaginary sphere 1 mi (1.609 km) in radius is 107.6 mV/m. This uniform spherical radiator has no directivity and hence no gain with a lossless antenna system. It is referred to as an isotropic radiator and is used as the standard for comparison of the gain of directional antennas. This antenna gain can be defined as the ratio of maximum field strength from the antenna of interest relative to the field strength that would be radiated from a lossless isotropic antenna fed with the same input power. In relation to directivity D, gain G can be expressed as in Eq. (10), where η is antenna efficiency.

$$G = \eta D \tag{10}$$

Efficiency. Antenna efficiency η is defined as the radiated power divided by the input power to the antenna. Input per P_t supplies antenna system losses P_l and the radiated power P_r, as shown by Eq. (11).

$$\eta = \frac{P_r}{P_l + P_r} \tag{11}$$

Losses in an antenna system take place in the resistance of the transmission line to the antenna, the coupling network between line and antenna, antenna conductors, insulators, and the ground system. For electrically small antennas, even small-loss resistances can significantly reduce the efficiency. In the design of transmitting antennas, the heating effect of losses must also be considered together with losses due to corona. *See* CORONA DISCHARGE.

Polarization. The plane of the electric field depends on the direction in which the current flows on the antenna (Fig. 3). The electric field is in a plane orthogonal to the axis of a magnetic dipole (Fig. 4). This dependence of the plane of the radiated electromagnetic wave on the orientation and type of antenna is termed polarization. A receiving antenna requires the same polarization as the wave that it is to intercept. By combining fields from electric and magnetic dipoles that have a common center, the radiated field can be elliptically polarized; by control of the contribution from each dipole, any ellipticity from plane polarization to circular polarization can be produced. Some antennas of the conical horn type lend themselves to symmetrical excitation of wave guides. Circular polarization can be easily detected with appropriate feed arrangements such as probes placed orthogonal to each other within the waveguide. The orientation and sense of the circular

polarization components can be ascertained and the accompanying signals separated. Similar cross-polarized (simultaneous vertical and horizontal) signals of the same or different frequencies can be detected. Thus, for a single transmitting or receiving station it is possible to almost double the useful capability. Such cross-polarization systems are used extensively in both terrestrial and satellite microwave facilities. *See* POLARIZATION OF WAVES.

Impedance. The input impedance Z_{in} of an antenna is the ratio of the voltage to current at the terminals connecting the transmission line and transmitter or receiver to the antenna. The impedance can be real for an antenna tuned at one frequency but generally would have a reactive part at another frequency. The impedance can be determined through theoretical and other methods, some of which involve numerical computations. For simple antennas, approximate input or self-impedance can be computed from a knowledge of radiation patterns and approximate current distribution. For more complicated radiating structures, input impedance is usually measured experimentally. *See* ELECTRICAL IMPEDANCE.

The concept of antenna impedance may be illustrated by the example of the simplest wire conductor antenna. The current distribution and charge distribution on the wire antenna can be approximated to a fair degree of accuracy by the distributions on an open-ended transmission line for which the current is zero at its end and the charge accumulation is a maximum. For a quarter-wavelength section of line, at the input, the current is a maximum and the charge is a minimum. Since the electric field is proportional to the charge and represents the voltage, the ratio of voltage to current represents a low impedance. As the wires are separated into a V antenna, the antenna impedance still remains low since the current and voltage at the input have not changed. When the configuration is fully expanded into the half-wave antenna, the input impedance is low but not negligible since the radiation of energy must be accounted for. The power carried by the emitted radiation is I^2R_r, where R_r is called the radiation resistance. For this half-wave antenna case, R_r has a value of about 73 ohms with a small reactive part. For the electrically small antenna (elemental dipole), the radiation resistance is given by Eq. (12). For $dz/\lambda = 0.05$,

$$R_r = 80\pi^2(dz/\lambda)^2 \tag{12}$$

R_r is about 2 ohms. The input impedance for this small antenna has a very large capacitive reactance which when tuned out with a corresponding inductor yields a resistive input of 2 ohms plus a resistance due to the imperfectly conducting antenna and inductor wires. This is not a very efficient antenna: its efficiency is low, and its low input resistance is difficult to match to the transmission line.

As a component of a receiving circuit, the antenna may be represented by Thévenin's or Norton's equivalence theorems. In the former the representation is a simple series circuit containing impedance, and also containing voltage induced from the incoming wave. Useful power is calculated in terms of the resulting current in the resistance of the load circuit. *See* NETWORK THEORY; THÉVENIN'S THEOREM (ELECTRIC NETWORKS).

If a current element of length dz is placed in a field E parallel to it, the root-mean-square voltage induced is E_s. If the element is tuned, the available power P that it captures is shown by Eq. (13). The average power per square meter P_{av} in the plane wave is shown in Eq. (14), which indicates that the effective

$$P = \frac{E^2\lambda^2}{320\pi^2} \tag{13}$$

$$P_{av} = \frac{E^2}{120\pi} \tag{14}$$

area of the elemental dipole is $3\lambda^2/(8\pi)$.

If the reactive component of antenna impedance is appreciable, techniques similar to those used in other transmission circuits are required to obtain maximum power; that is, the impedance of the connected circuit must be conjugate to that of the antenna. Reflections caused by a mismatched antenna can cause trouble at a distant point in receiving circuitry. The solution may be to obtain a better match or to employ a nonreciprocal device which discriminates against reflected waves. *See* IMPEDANCE MATCHING.

When a driven antenna is in proximity to others, the radiation pattern of the antenna can be altered so as to yield the desired fields. The presence of the other elements can modify the input impedance of the antenna, called driving-point impedance. The combination of the elements and driven antenna constitute an array, as discussed below.

Because of the interaction between the elements of an array, a distinction is made between a driven element to which a transmission line directly connects and a parasitic element that couples into the array only through its mutual electromagnetic coupling to the driven elements. Any conductor, such as a guy wire, in the vicinity of a primary or driven antenna element may act as a parasitic element and affect antenna impedance and radiation pattern (**Fig. 5***a*). The driving-point impedance and the radiation pattern are controlled by the tuning and spacing of parasitic elements relative to driven elements. If a parasitic element reduces radiation in its direction from a driven element, it is called a reflector; if it enhances radiation in its direction, it is called a director. *See* YAGI-UDA ANTENNA.

Array antennas. An array of antennas is an arrangement of several individual antennas so spaced and phased that their individual contributions add in the preferred direction and cancel in other directions. One practical objective is to increase the signal-to-noise ratio in the desired direction. Another objective may be to protect the service area of other radio stations, such as broadcast stations. *See* SIGNAL-TO-NOISE RATIO.

The simplest array consists of two antennas. Using phasor notation, the azimuth pattern is given by

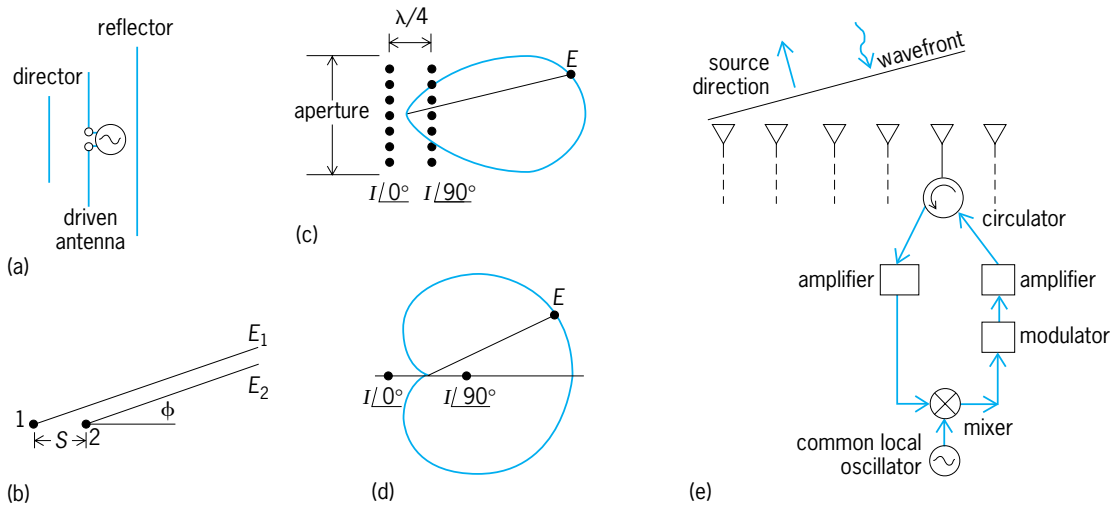

Fig. 5. Antenna arrays. (*a*) Simple three-element antenna. (*b*) Two-antenna array. (*c*) Broadside box array. (*d*) Unidirectional (cardioid pattern) array. (*e*) Retrodictive array. Each element has the same electronics circuit. (*Part e after R. E. Collin, Antennas and Radiowave Propagation, McGraw-Hill 1985*)

Eq. (15), where E is field strength for unit distance at

$$E = E_1 \underline{/0^\circ} + E_2 \underline{/2\pi(S/\lambda)\cos\ \phi + \psi_2} \qquad (15)$$

azimuth angle ϕ, E_1 and E_2 are nondirectional values of the field strength of the individual antennas, S is the spacing from antenna 1 to antenna 2, and ψ_2 is the electrical phasing of antenna 2 with respect to antenna 1 (Fig. 5*b*). *See* ALTERNATING-CURRENT CIRCUIT THEORY.

The four variables that the simple two-antenna array provides make possible a wide variety of radiation patterns, from nearly uniform radiation in azimuth to a concentration of most of the energy into one hemisphere, or from energy in two or more equal lobes to radiation into symmetrical but unequal lobes.

In a directional antenna a feeder system is required for the proper division of power and control of the phase of the radiated fields. The feeder system is also called upon to match impedances between the transmitter of transmission line and the antenna.

For further control over the radiation pattern a preferred arrangement is the broadside box array. In this array, antennas are placed in a line perpendicular to the bidirectional beam. Individual antenna currents are identical in magnitude and phase. The array can be made unidirectional by placing an identical array 90° to the rear and holding its phase at 90° (Fig. 5*c*). The directivity of such a box array increases with the length or aperture of the array.

Another popular arrangement is the in-line array, in which the current in each antenna is equal but the phase varies progressively to give an end-fire unidirectional pattern (Fig. 5*d*). The end-fire array, for a given number of elements, does not have as much gain as the broadside box array.

If antennas are stacked vertically in line to form a collinear array with the currents of each antenna in phase, a symmetrical pattern results in the azimuth plane. This type of array is used for commercial broadcasting to increase the nondirectional ground coverage for a given transmitter power.

One useful antenna system consists of many, usually identical, antenna elements that are connected through separate networks capable of introducing signal phase changes so that a progressive phase shift along the array results in a corresponding shift in direction of the radiation beam. Of course, to obtain the broadside array property mentioned above, the phase-shift difference is maintained at zero or a multiple of 360°. For the end-fire array, maximum radiation is desired along one end of the array and a minimum, or zero, in the back radiation. The progressive phase shift ζ (a linearly increasing phase delay) depends on the uniform spacing d of the elements. For an array with N elements, $\zeta = \pi/N$ and $d = \lambda/(2N)$, where λ is the wavelength.

Further use of array concepts has enabled improvements in communications. By introducing a network for each antenna element (Fig. 5*e*), it is possible to receive a signal from a source direction and to return a signal in the direction of the source. The returned signal can be modulated or amplified or have its frequency changed. Such an array is called a retrodirective array. Basically, the array seeks out the incoming signal and returns one of useful characteristics, such as that which is needed for the communication between a moving vehicle and a stationary or slowly moving source.

The incoming pilot signal excites the individual elements with a progressive phase delay; the element closest to the source receives the incoming wavefront in advance of the other array elements, with the element farthest from the source experiencing the largest phase delay. In order that a transmitted signal (with information) be sent in the direction of the source, it is necessary that the phase delays be converted into phase advances. That is to say, the element farthest from the source must have the largest phase advance, and the other elements of the array must have progressively lesser phase advances. The

overall array thus transmits toward the source direction.

The phase reversal (from a phase delay to phase advance) is accomplished at each array element through the use of a circulator (separating the incoming signal from the outgoing), a common local oscillator of twice the frequency of communication, a mixer, a separation of the resulting frequencies, and a transmission of the responding signal with information sent to the source.

Antenna arrays can also be designed to adapt their radiation or receiving patterns through the use of proper signal processing of the energy received by each element of the array. It is possible to enhance the signal received from a desired direction or to reject an unwanted signal or interference from another direction. The signal processing involves phase comparison and phase-locked loops to obtain the desired results. Such antenna arrays are called adaptive arrays. *See* PHASE-LOCKED LOOPS.

It is possible to use more than two antennas in an array to achieve a desired pattern shape. If a symmetrical pattern is satisfactory, the antennas can be in line. If the arrays are not in line, the protection directions can be specified by the nulls of the individual pairs of antennas used in the array. When necessary, the number of pairs of antennas can be increased to provide further control of the pattern shape.

Bandwidth. The bandwidth of an antenna may be limited by pattern shape, polarization characteristics, and impedance performance. Bandwidth is critically dependent on the value of Q; hence the larger the amount of stored reactive energy relative to radiated resistive energy, the less will be the bandwidth. *See* Q (ELECTRICITY).

For efficient operation, a low standing-wave ratio at the antenna input terminals is necessary over the operating frequency range. Usually the input impedance will vary in both resistance and reactance over the desired frequency interval. It is, therefore, of interest to determine the bandwidth that can be achieved for a given standing-wave ratio. For a particular antenna the bandwidth that can be obtained by compensating networks has a theoretical limit; therefore, if a wider band is required, it becomes necessary to select antenna types with the inherent characteristic of wider bandwidth, such as the frequency-independent antennas discussed below.

Electrically small antennas. Antennas whose mechanical dimensions are short compared to their operating wavelengths are usually characterized by low radiation resistance and large reactance. This combination results in a high Q and consequently a narrow bandwidth. Current distribution on a short conductor is sinusoidal with zero current at the free end, but because the conductor is so short electrically, typically less than 30° of a sine wave, current distribution will be essentially linear. By end loading to give a constant current distribution, the radiation resistance is increased four times, thus greatly improving the efficiency but not noticeably altering the pattern.

An end-fed monopole antenna is common at very low frequencies for long-range communication, commercial broadcasting, and mobile use. The small loop antenna is used extensively for direction finding and navigation. Where height is a limiting factor at higher frequencies, the monopole height can be reduced by forming the conductor into a helical whip or by top loading.

Slot antenna. Where no height is permitted, a slot is used. Just as the magnetic loop is related to the electric dipole, so a slot in a conductive surface is related to a conductive wire in space. The conductive surface is usually the outside surface of a waveguide within which the radio energy travels. A slot in the waveguide allows the energy to radiate. Typically the slot is narrow and a half-wavelength long. The configuration of the electric field radiated from the slot is the same as the magnetic field from a wire of like dimensions; thus one is the dual of the other. The slot may be fed by a transmission line connected across its narrow dimension, by a resonant cavity behind it, or (as mentioned above) from a waveguide. Because it is flush with a metallic surface, a slot antenna is advantageous in aircraft.

Broadband antenna. The impedances of very thin wire antennas such as half-wave antenna dipoles and quarter-wave monopoles are highly sensitive to frequency, although the radiation pattern of these antennas is insensitive to frequencies off resonance. The impedance can have large capacitive reactances for frequencies lower than resonance and inductive reactances for frequencies above resonance. This variation can be minimized by choosing dipoles that are thick, by minimizing reflections at the feed points by using conical conductor, or by combinations of these choices.

Nonresonant antennas. Long-wire antennas, or traveling-wave antennas, are usually one or more wavelengths long and are untuned or nonresonant.

Horizontal single-conductor type. The radiation pattern of a long conductor in free space depends upon its length in wavelengths. A 0.5-λ conductor will radiate broadside, but as it is made longer the pattern splits, the major lobe coming closer to, but never reaching, the direction of the conductor. The pattern is a figure of revolution of its own cross section around the conductor as an axis.

When a horizontal unterminated line with standing electric current waves of uniform amplitude is placed above perfect ground, it has a radiation pattern symmetrical about the center with its major lobes closest to the line in both directions. In the practical case with constant phase velocity and exponential variation in amplitude, the pattern lobes will be unsymmetrical with larger lobes toward the open end. The total number of lobes is the same as the number of 0.5-λ lengths of the conductor. If the far end of the line is terminated, the antenna radiates only one major lobe in the direction the wave travels down the line to the termination.

Arrays of long-conductor type. Because a single long-conductor terminated antenna has rather high side lobes and a major lobe at an angle to the axis of the conductor, two long terminated conductors can

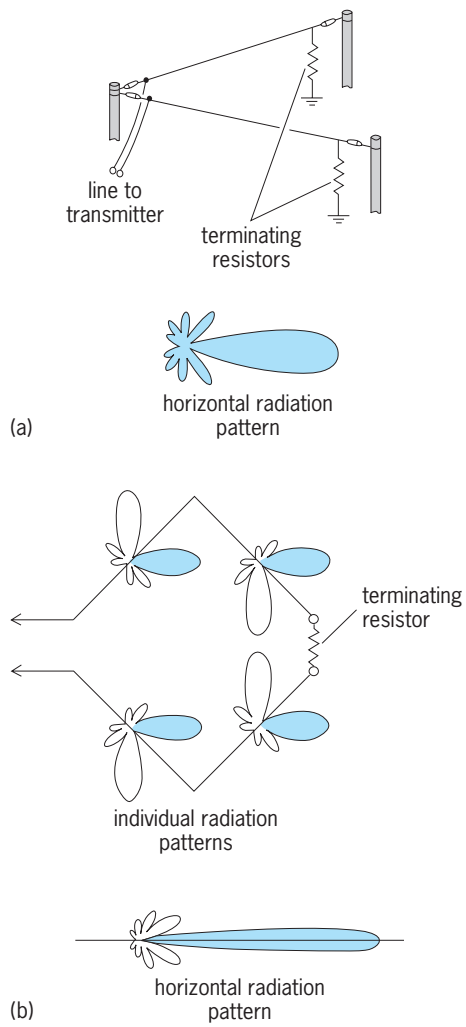

Fig. 6. Arrays of long-conductor antennas. (*a*) Horizontal terminated V. (*b*) Horizontal terminated rhombic.

be combined into a horizontal V array (**Fig. 6***a*). The angle between the V array is determined by the length of the conductors and is usually chosen so that the major lobes will add in phase.

The commonest type of long-conductor array is the rhombic antenna (Fig. 6*b*). Major lobes of four legs add in phase to form the resultant major lobe, while major lobes at right angles tend to cancel, leaving only smaller side lobes. Because of its simplicity, low cost, and wide bandwidth, the rhombic array is widely used for both transmission and reception where the side lobes can be tolerated and the real estate dimensions permit.

Dielectric type. A dielectric material, formed into either a solid rod or a hollow cylinder, is a waveguide. The wavelength of a wave entirely inside a large, solid dielectric rod is less than the free-space wavelength; rather, the wavelength is this free-space length divided by the square root of the dielectric constant of the rod. If the rod diameter is large compared to the wavelength inside it, most of the energy in the wave travels inside the dielectric. Even if the rod diameter is somewhat less than the wavelength inside the dielectric, the wave travels along the rod, with some of the energy inside the dielectric and some outside. However, if the diameter of the rod is reduced below a half-wavelength, the velocity of the wave increases to that of free space and the wave continues beyond the end of the rod into free space. The major lobe is then in the direction of the rod. *See* DIELECTRIC MATERIALS.

Frequency-independent antennas. Antennas that consist of lengths of wires are sensitive to changes in frequency as the wire lengths become equal to a half-wavelength, a full wavelength, and so forth. An infinitely long biconical antenna would have characteristics independent of frequency since its dimension is specified by angle only, but, of course, if such an antenna is truncated its length is also specified, making it frequency-dependent.

There are two principal approaches to constructing frequency-independent antennas. The first is to shape the antenna so that it can be specified entirely by angles; hence when dimensions are expressed in wavelengths, they are the same at every frequency. Planar and conical equiangular spiral antennas adhere to this principle. The energy radiates as the wave progresses along an equiangular spiral antenna, so that beyond a dimension of circumference of the spiral equal to the wavelength the antenna structure can be terminated. This criterion determines the lowest frequency of practical application. The preciseness of the input portion of the spiral determines the highest frequency of operation. If r_1 is the radial distance to the spiral corresponding to a wavelength λ_1, such that Eq. (16) is valid, then at a new

$$r_1 = r_0 e^{a(\phi_1 - \phi_0)} \tag{16}$$

wavelength λ_2 this radial distance should be given by Eq. (17) if the antenna characteristics are to be

$$r_2 = (\lambda_2/\lambda_1) r_1 \tag{17}$$

frequency-independent. This yields Eq. (18). Here a,

$$r_2 = r_0 e^{a(\phi_2 - \phi_0)} \tag{18}$$

r_0, and ϕ_0 represent the spiral parameters, and ϕ_1 and ϕ_2 are the angles at which the radial distance is r_1 and r_2 respectively (**Fig. 7***a*).

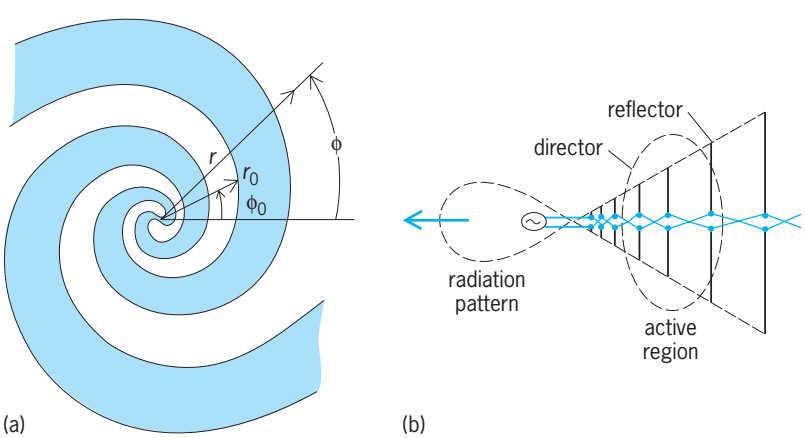

Fig. 7. Frequency-independent antennas. (*a*) Equiangular spiral (*after D. J. Angelakos and T. E. Everhart, Microwave Communications, Krieger, 1983*). (*b*) Log-periodic structure.

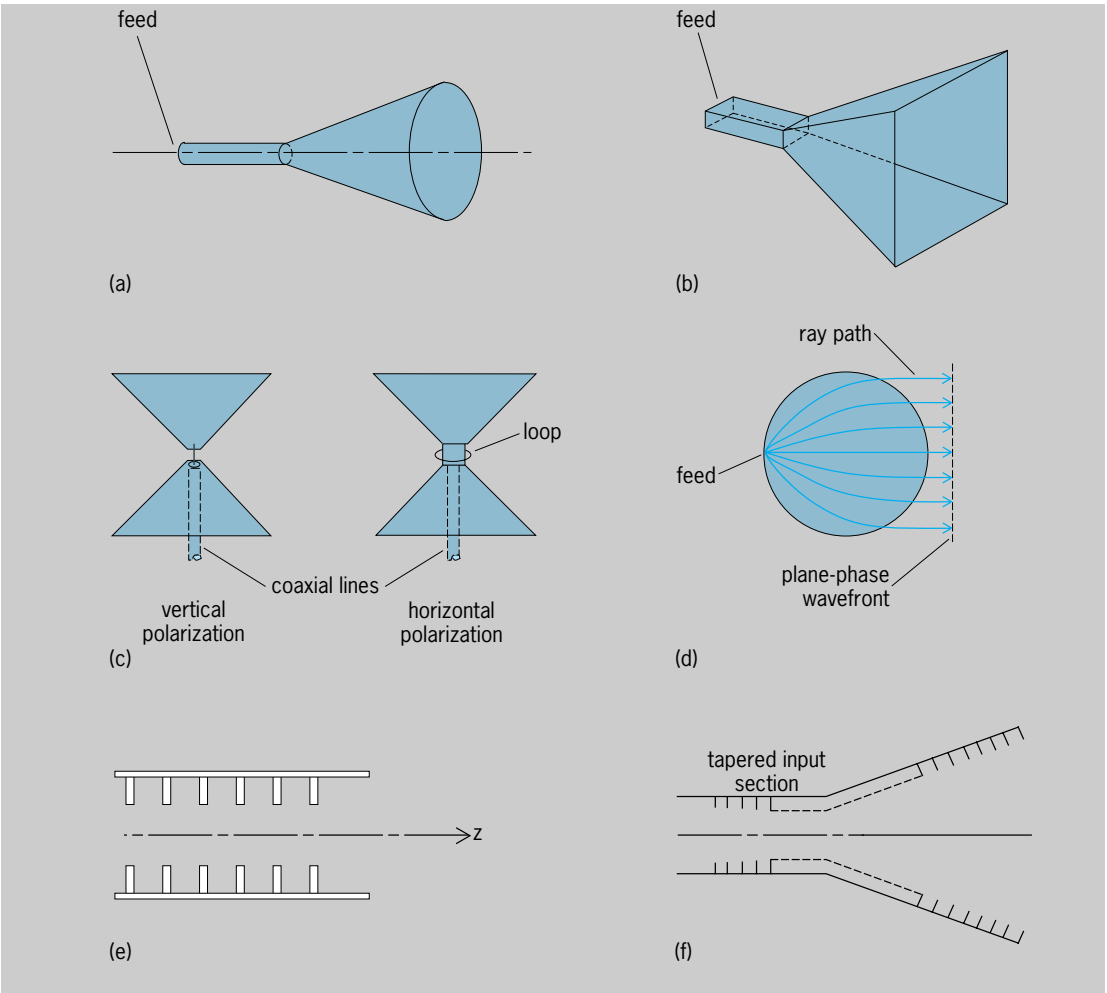

Fig. 8. Direct-aperture antennas. (*a*) Conical horn. (*b*) Pyramidal horn. (*c*) Biconical horns, electrically fed for vertical polarization or magnetically fed for horizontal polarization. (*d*) Luneberg lens with dielectric sphere between feed and plane-phase wavefront. (*e*) Corrugated circular waveguide. (*f*) Corrugated conical horn. (*Parts e and f after R. E. Collin, Antennas and Radiowave Propagation, McGraw-Hill, 1985*)

The second approach depends upon complementary shapes. According to this principle, which is used in constructing log-periodic antennas, before the structure shape changes very much, when measured in wavelengths, the structure repeats itself. This periodicity in the exponent of the structural dimensions yields a class of antennas which is frequency-independent in the sense that as the frequency of the input signal is changed, another set of elements comes into resonance, with the same electrical dimensions for each frequency, within the limits set by the fineness of the element spacing (Fig. 7*b*).

By combining periodicity and angle concepts, antenna structures of very large bandwidths become feasible.

Direct-aperture type. When they are to be used at short wavelengths, antennas can be built as horns, mirrors, or lenses. Such antennas use conductors and dielectrics as surfaces or solids in contrast to the antennas described thus far, in which the conductors were used basically as discrete lines. *See* MICROWAVE OPTICS.

Horn radiators. The directivity of a horn increases with the size of the mouth of the horn; however, tolerances must be held close if the variations in magnitude and phase distribution are to be low enough to maintain high gain over a broad band. Conical and pyramidal horns are commonly used for high-gain beams (**Fig. 8***a* and *b*). Biconical horns (Fig. 8*c*) have an omnidirectional pattern with either vertical or horizontal polarization.

Luneberg lens. If in a dielectric sphere the index of refraction varies with distance from the center of the sphere, a plane wave falling on the surface of one hemisphere is focused at the center of the opposite hemisphere. Such an arrangement is called a Luneberg lens (Fig. 8*d*). By reciprocity, energy is fed into the lens at the focal point for transmitting; energy is removed from the lens at this point for receiving. Only the feed point need be moved around the lens to steer the pencil beam.

For high frequencies a horizontal conducting grid structure serves as a Luneberg lens antenna. Because it will focus rays from any azimuth direction, several high-frequency beams can be generated

in different azimuth directions at the same time.

Corrugated conical horns. A very special antenna system that is used either as a feed of a reflector type or as a direct radiator is the corrugated conical horn antenna (Fig. 8*e* and *f*). For the ordinary metallic-wall waveguide and horn, the metal walls present boundary conditions for the electric field that are different from those for the magnetic field. By replacing the metal walls by pseudowall-nonwall in the corrugated structure, the boundary conditions become impartial to the type of field: electric or magnetic. The hybrid wall appears as both a short circuit and an open circuit. The latter is obtained by depressions of length comparable to a quarter wavelength at some middle frequency. This horn has a high degree of circular symmetry in its radiation and receiving patterns.

Reflector type. By using reflectors it is possible to achieve high gain, modify patterns, and eliminate backward radiation. A low-gain dipole, a slot, or a horn, called the primary aperture, radiates toward a larger reflector called the secondary aperture. The large reflector further shapes the radiated wave to produce the desired pattern. *See* REFLECTION OF ELECTROMAGNETIC RADIATION.

Plane sheet reflectors. One common plane reflector antenna consists of a dipole parallel to a flat conducting sheet (**Fig. 9***a*). The perfectly conducting screen creates an image of the physical structure (mirror image) but with the currents of the image flowing in the opposite direction. The forward pattern is the same as that of a two-element dipole array. If the screen is large enough, it prevents all back radiation. Usually the screen is limited in size, and some radiation exists to the rear. Gain depends on spacing and can exceed 6 dB above a free-space dipole.

Corner reflectors. If two flat conducting sheets intersect along the y axis at an angle α (Fig. 9*b*), an effective directional antenna results when a dipole is placed a distance S from the corner. Gain depends on corner angle α, dipole spacing S, and antenna losses; gain can exceed 12 dB for $\alpha = 45^\circ$. *See* CORNER REFLECTOR ANTENNA.

Parabolic reflectors. If an isotropic source is placed at the focus F of a parabola, the radiated wave is reflected from the parabolic surface as a plane wave at the aperture plane (Fig. 9*c*). The physical size of the parabola determines the size of the aperture plane.

If the parabolic reflector is a surface of translation (Fig. 9*d*), an in-phase line source serves as the primary aperture, such as the two dipoles in this illustration. Beam width in the vertical, z direction is controlled by the parabola aperture, while the azimuth beam width, in the y direction, is controlled by the length of the cylindrical parabola.

If the parabolic reflector is a surface of rotation, it converts a spherical wave from an isotropic source at its focus to a uniform plane wave at its aperture. The presence of the primary antenna in the path of the reflected wave has two disadvantages. (1) The reflected wave modifies the input impedance of the primary radiator. (2) The primary radiator obstructs the reflected wave over the central portion of the aperture (Fig. 9*e*).

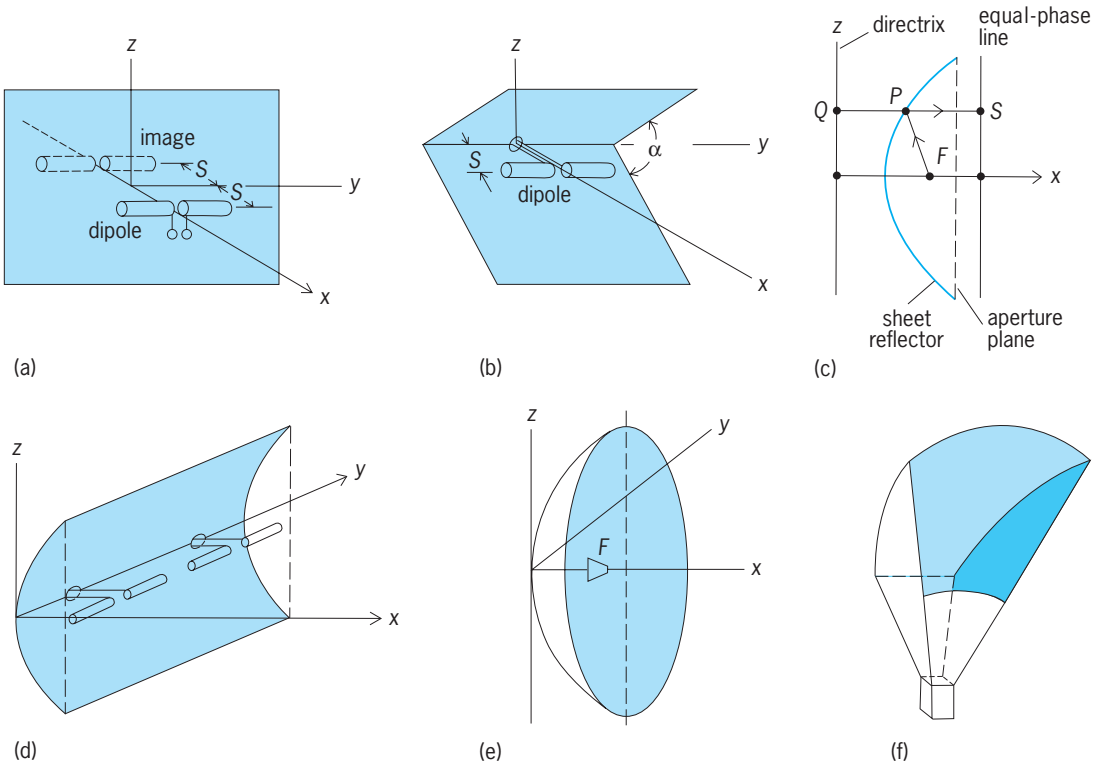

Fig. 9. Reflector antennas. (*a*) Plane sheet reflector. (*b*) Corner reflector. (*c*) Parabolic reflector, where *FP* = *QP* and *FP* + *PS* is a constant for all points *P* on the reflector. (*d*) Cylindrical parabolic reflector. (*e*) Paraboloidal reflector. (*f*) Horn paraboloid reflector.

-S

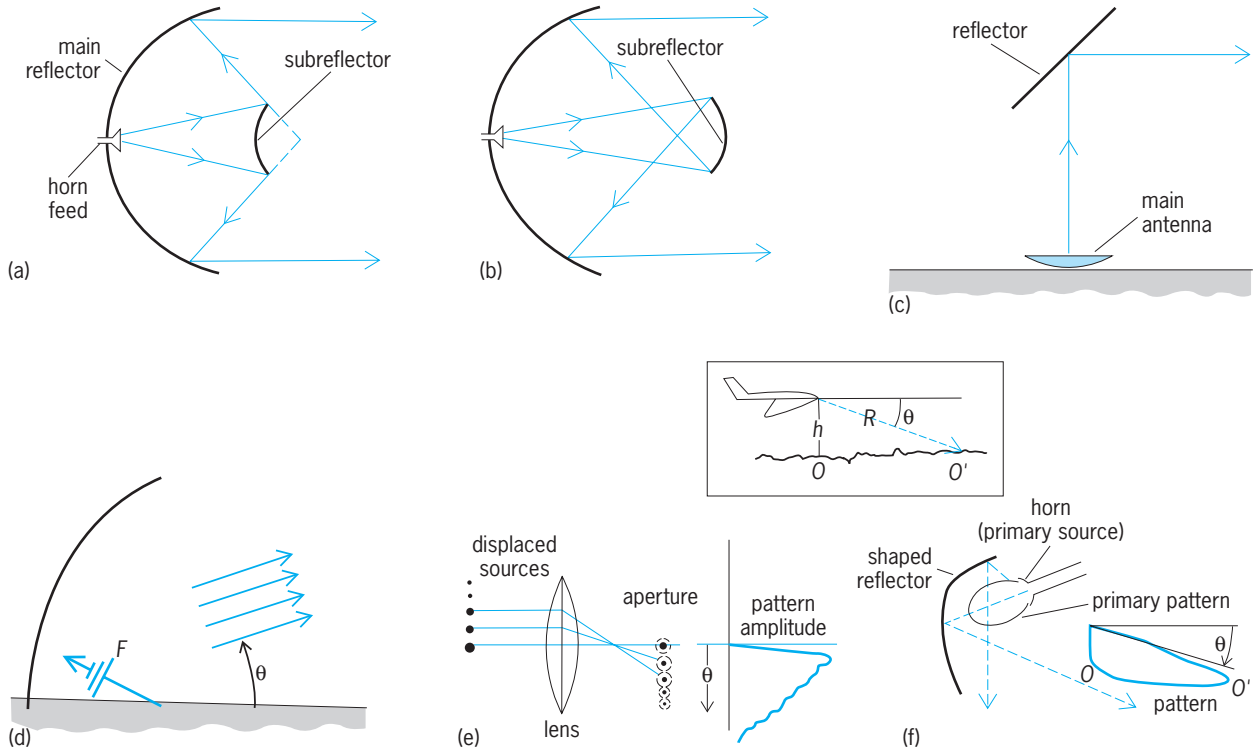

Fig. 10. Two-element systems. (*a*) Cassegrain. (*b*) Gregorian. (*c*) Periscope. (*d*) Spherical steerable system. Feed is located at focus *F*. θ = elevation angle. (*e*) Distributed sources. (*f*) Shaped reflector. Inset to parts *e* and *f* shows geometry of aircraft radar application. Power emitted in the direction of *O′* must be greater than that emitted from *O* to compensate for weaker signals reflected from *O′* because range *R* to *O′* is greater than distance *h* to *O*. (*Part e and f after D. J. Angelakos and T. E. Everhart, Microwave Communications, Krieger, 1983*)

To avoid both of these disadvantages, only a portion of the paraboloid can be used; the focus, where the primary aperture must be placed, is then to one side of the secondary aperture. For example, a primary-aperture horn placed at one side can illuminate the parabolic reflector (Fig. 9*f*). This type of antenna is commonly used in point-to-point microwave systems, because it is broadband and has a very low noise level. Because of its high front-to-back ratio, the horn paraboloid reflector is used at the ground station for satellite communication.

Two-reflector antennas. A beam can be formed in a limited space by a two-reflector system. The commonest two-reflector antenna, the Cassegrain system, consists of a large paraboloidal reflector. It is illuminated by a hyperbolic reflector, which in turn is illuminated by the primary feed (**Fig. 10***a*). The Gregorian system (Fig. 10*b*) is similar, except that an elliptical reflector replaces the hyperbolic reflector and consequently must be placed farther out than the focal point of the paraboloid. This system is less compact than the Cassegrain system, and less frequently used. A variation of the Gregorian involves a shaping of the primary reflector to further modify the pattern; it is called the Dragonian.

In microwave relay systems, one common practice is to place the paraboloid at ground level (Fig. 10*c*). The feed system is readily accessible, and long lengths of transmission line are not required. A flat sheet at the top of a tall tower redirects the beam over obstacles to the next antenna.

By using a portion of a hemisphere (Fig. 10*d*) illuminated by a movable feed, the spherical reflecting surface causes the wave to be reflected off the ground to form a beam at a specified elevation angle. With this arrangement it is possible to steer the beam both in elevation and in azimuth. This system is used in the Arecibo instrument, a 1000-ft (300-m) radio telescope antenna.

A pattern of desired shape can alternatively be obtained by displacing the primary source (Fig. 10*e*) or by shaping the reflector (Fig. 10*f*). An example of the application of such patterns is an aircraft radar system.

Low-profile antennas. A series of antennas are useful in situations which require a low profile. The slot antennas mentioned above constitute a large portion of this group. In essence, replacing a wire (metal) by a slot (space), which is a complement of the wire, yields radiation characteristics that are basically the same as those of the wire antenna except that the electric and magnetic fields are interchanged. The impedance Z_d of the driven antenna, in the simplest case of a dipole antenna, is related to that of the equivalent driven slot antenna Z_s by Eq. (19). A

$$Z_s Z_d = \frac{(120\pi)^2}{4} \tag{19}$$

folded dipole antenna has an impedance $Z_{fd} = 4Z_d$ that is related to the impedance Z_{fs} of the equivalent

slot antenna by Eq. (20).

$$Z_{fs}Z_{fd} = \frac{(120\pi)^2}{4} = Z_{fs}(4Z_d) \qquad (20)$$

Because flush-mounted antennas present a low profile and consequently low wind resistance, slot-type antennas have had considerable use in aircraft, space-launching rockets, missiles, and satellites. They have good radiation properties and are capable of being energized so as to take advantage of all the properties of arrays, such as scanning, being adaptive, and being retrodirective. These characteristics are obtained without physical motion of the antenna structures. Huge slot antenna arrays are commonly found on superstructures of aircraft carriers and other naval ships, and slot antennas are designed as integral parts of the structure of aircraft, such as the tail or wing.

Patch antenna. This antenna consists of a thin metallic film which is attached to a dielectric substrate mounted on a metallic base. Depending on its use, the patch can be of different shapes and can be driven in various fashions. Driven at one end, the radiated electric field at this end has a polarization that is in phase with the radiated electric field at the farther end of the patch antenna.

MMIC antennas. Planar antennas are designed as integral parts of monolithic microwave integrated circuits (MMICs). Coupling between can be effected through the use of planar (flush-mounted) antennas fabricated directly on the microelectronics chips (integrated circuits). This arrangement eliminates the need for coaxial lines, which at these microwave frequencies exhibit considerable losses. As is the case with other planar antennas, it is possible to design circuitry so as to obtain many, if not all, the properties of arrays mentioned above. The elements of these arrays can take on the form of slot antennas or patch antennas (of course with suitable modification for use on the MMICs). *See* MICROWAVE.

Diogenes J. Angelakos

Bibliography. American Radio Relay League, *ARRL Antenna Book*, 18th ed., 1997; D. J. Angelakos and T. E. Everhart, *Microwave Communications*, 1968, reprint 1983; C. A. Balanis, *Antenna Theory: Analysis and Design* 2d ed., 1996; R. E. Collin, *Antennas and Radiowave Propagation*, 1985; R. C. Johnson, *Antenna Engineering Handbook*, 3d ed., 1993; J. D. Kraus, *Antennas*, 2d ed., 1988; W. L. Pritchard et al., *Satellite Communication Systems Engineering*, 2d ed., 1992; W. L. Stutzman and G. A. Thiele, *Antenna Theory and Design* 2d ed., 1997.

Anthocerotopsida

A small class of the plant division Bryophyta, made up of plants commonly called hornworts. There is a single order with six genera.

Gametophyte structure. The gametophytes are flat thalli, often forming rosettes of uniform thickness (see **illus.**) or, in *Dendroceros*, with a thickened midrib and crisped unistratose wings. Ventral scales are absent. Rhizoids are unicellular, with thin, smooth walls. The thallus of undifferentiated tissue is sometimes provided with cavities containing mucilage (and often occupied by *Nostoc* colonies) and ventral pores leading to them. The cells are thin-walled and lack oil bodies; at least those at the surface have one, two, or more large chloroplasts with a central pyrenoid. Slime papillae (or mucilage hairs) are lacking. The stalked antheridia develop from internal cells and occur singly or in groups in cavities beneath the upper surface of the thallus and become exposed by the rupture of overlying tissue. The archegonia, also endogenous, are embedded in the dorsal tissue of the thallus. Paraphyses are lacking. The sporophyte is surrounded at the base by a tubular involucre (or completely surrounded in *Notothylas*). A calyptra is lacking.

Sporophyte structure. The sporophyte consists of a massive foot and an erect, long-cylindric, green capsule (see illus.), though in *Notothylas* the capsule is horizontal and spindle-shaped. The capsule, indeterminate in growth owing to a basal meristem, dehisces from the apex into two valves which are usually twisted when dry. The wall consists of severa l cell layers in a more or less solid tissue having one, two, or more chloroplasts per cell. Stomata with two guard cells are usually present. The spore sac, derived from the amphithecium, surrounds and overarches the slender columella (lacking in *Notothylas*). Spore maturation proceeds from the apex downward. The tetrahedral spores are mingled with pseudelaters of one to several cells, with or without spiral bands. A protonema is lacking. The haploid chromosome number is 5 or 6. *See* REPRODUCTION (PLANT).

The elongate capsule dehiscing by two valves and its indeterminate growth from a basal meristematic

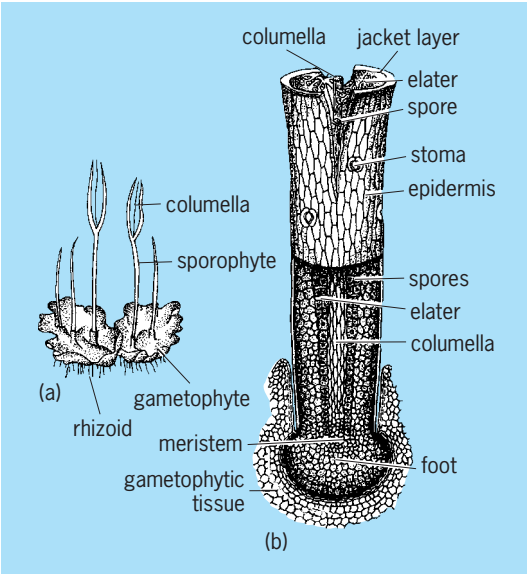

Hornwort *Anthoceros*. (*a*) Thalloid gametophyte with long hornlike sporophytes. (*b*) Elongated sporophyte with basal foot in gametophytic tissue. (*After H. J. Fuller and O. Tippo, College Botany, rev. ed., Holt, 1954*)

tissue are unique, as are the large chloroplasts with a central pyrenoid and the endogenous origin of sex organs. (The chloroplasts are sometimes single in the gametophyte and paired in the sporophyte.) The spore mother cells undergo meiosis directly, but the elater mother cells usually undergo several mitotic divisions before differentiation as pseudelaters of an unreduced chromosome number. As a result, the pseudelaters may be more numerous than the spores. (In the Hepaticopsida, by contrast, the spore mother cells undergo meiosis directly or undergo several mitotic divisions before meiosis, whereas the elater mother cells mature directly into diploid elaters.) *See* ANDREAEOPSIDA; BRYOPHYTA; BRYOPSIDA; HEPATICOPSIDA; SPHAGNOPSIDA. Howard Crum

Bibliography. M. Fulford, Contemporary thought in plant morphology: Hepaticae and Anthocerotae, *Phytomorph*, 14:103–119, 1964; G. G. Häel de Menéndez, Estudio de las Anthocerotales y Marchantiales de la Argentina, *Opera Lilloana*, 7:1–297, 1962; J. Proskauer, Studies on the morphology of *Anthoceros*, pt. I., *Ann. Bot.*, n.s., 12:237–265, 1948, pt. II, 12:427–439, 1948; J. Proskauer, Studies on Anthocerotales, pt. III: The genera *Anthoceros* and *Phaeoceros, Bull. Torrey Bot. Club*, 78:331–349, 1951, pt. IV, 80:65–75, 1953; J. Proskauer, Studies on Anthocerotales, pt. V: The genera Anthoceros and Phaeoceros, *Phytomorph*, 7:113–135, 1957, pt VI: On spiral thickenings in the columella and its bearing on phylogeny, *Phytomorph*, 10:1–19, 1960.

Anthophyllite

A magnesium-rich orthorhombic amphibole with perfect {210} cleavage and a color which varies from white to various shades of green and brown. It is a comparatively rare metamorphic mineral which occurs as slender prismatic needles, in fibrous masses, and sometimes in asbestiform masses. Anthophyllite may occur together with calcite, magnesite, dolomite, quartz, tremolite, talc, or enstatite in metacarbonate rocks; with plagioclase, quartz, orthopyroxene, garnet, staurolite, chlorite, or spinel in cordierite-anthophyllite rocks; and with quartz and hematite in metamorphosed iron formations and with talc, olivine, chlorite, or spinel in metamorphosed ultrabasic rocks. Anthophyllite is distinguished from other amphiboles by optical examination or by x-ray diffraction, and from other minerals by its two cleavage directions at approximately 126° and 54°.

Anthophyllite has the general formula below, with

$$(Mg, Fe^{2+})_{7-x}Al_x(Al_xSi_{8-x}O_{22})(OH, F, Cl)_2$$

$x < 1.0$. For aluminum-poor varieties, up to about 40% of the Mg may be replaced by Fe^{2+}; higher iron contents result in the formation of the monoclinic amphibole cummingtonite. Increasing the aluminum content in anthophyllite beyond $x = 1.0$ results in the formation of the orthohombic amphibole

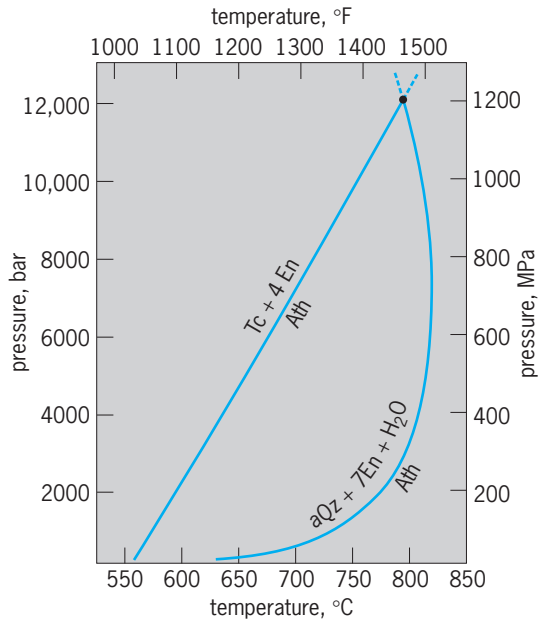

Stability field of anthophyllite in pressure-temperature space calculated with the program GEO-CALC. Broken lines extend from the invariant point at which the equilibrium terminates. Ath = anthophyllite, En = enstatite, Fo = forsterite, *a*Qz = quartz (*a* is variable coefficient), Tc = talc.

gedrite; aluminous anthrophyllite can accommodate more Fe^{2+} than Al-poor varieties.

The stability field of anthophyllite in pressure–temperature space (see **illus.**) is bounded at low temperatures by the equilibrium 4 enstatite + talc = anthophyllite and at high temperatures by the equilibrium anthophyllite = 7 enstatite + quartz + H_2O. *See* AMPHIBOLE; BIOPYRIBOLE; CUMMINGTONITE.

J. V. Chernosky

Bibliography. W. A. Deer, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, vol. 2B: *Double-Chain Silicates*, 2d ed., 1997.

Anthozoa

A class of phylum Cnidaria (formerly known as Coelenterata) whose members are marine, live exclusively as polyps, and occur as solitary individuals or in clonal or colonial groups. Most adult anthozoans are attached to a firm object or the sea bottom but some burrow into soft sediments, and rare ones are planktonic. As is typical of cnidarians, sexual reproduction commonly gives rise to a planktonic larva, the planula, which metamorphoses to a polyp, usually when it settles onto the substratum. *See* CNIDARIA.

Defining characteristics. Anthozoans are distinguished from members of other classes of Cnidaria by the absence of a medusoid phase in the life cycle as well as by certain characteristic features of the polyp: (1) hollow tentacles, into which the coelenteron (gastrovascular cavity) extends; (2) some to many cells in the mesoglea; (3) gametes of endodermal origin; (4) an actinopharynx (also called a stomodeum), a throatlike tube that extends from

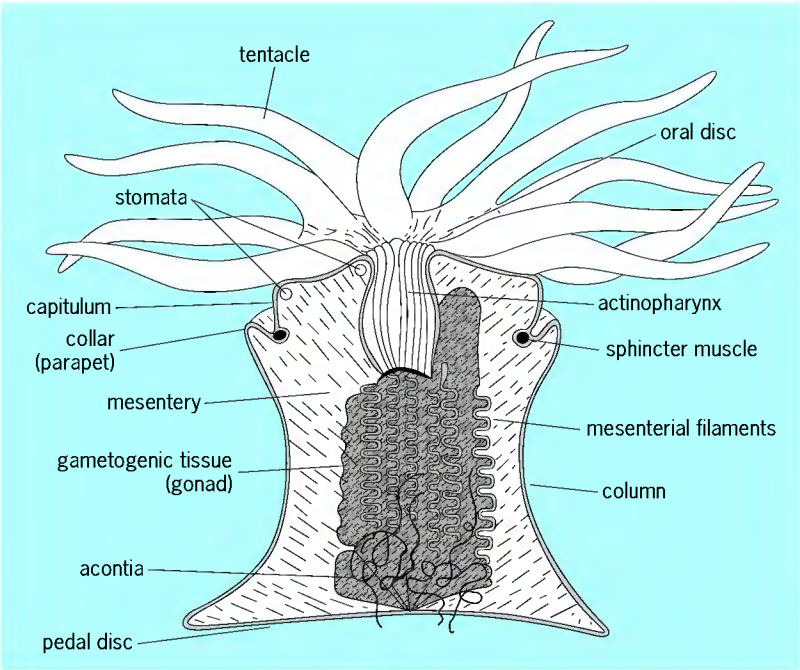

Fig. 1. Longitudinal section through a sea anemone. (*Reprinted with permission from Interactive Glossary of Actinarian (Sea Anemone) Anatomical Terms*, http://darwin.nhm.ku.edu/inverts/InterGlossary/basicintro.htm)

the mouth partly into the coelenteron (**Fig. 1**); and (5) mesenteries, which are radially arrayed sheets of tissue extending from the body wall that partition the coelenteron into wedge-shaped compartments. Some or all mesenteries, termed "complete," extend from the body wall to the actinopharynx, whereas "incomplete" ones do not reach the actinopharynx (**Fig. 2**). A polyp of the subclass Octocorallia has eight mesenteries and eight tentacles; each tentacle has small side branches, pinnately arrayed. A polyp of the subclass Hexacorallia typically has roughly a multiple of six tentacles and mesenteries, although some have octamerous, decamerous, or, rarely, irregular symmetry.

Running the length of the actinopharynx, which is more or less ovoid in cross section, are one or more grooves called siphonoglyphs Fig. 2, which are smoother and have longer cilia than the rest of the actinopharynx. Octocorals and some hexacorals have one siphonoglyph; many hexacorals have two (or rarely more), located at diametrically opposite ends of the actinopharynx. The arrangement of retractor muscles on mesenteries that connect to the siphonoglyphs may differ from that of the other mesenteries. The presence of siphonoglyph(s) and these anatomically distinctive mesenteries provide a bilaterality to animals that belong to a group commonly considered radially symmetrical. Thus anthozoans are described as biradial, with elements of both radial and bilateral symmetry.

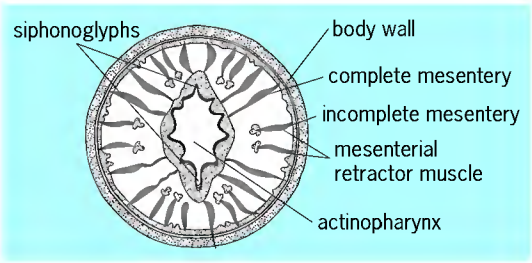

Fig. 2. Cross section (at the level of the actinopharynx) through the column of a sea anemone. (*Adapted from L. H. Hyman, The Invertebrates: Protozoa Through Ctenophora, vol. 1, McGraw-Hill, 1940*)

Classification. Class Anthozoa includes soft, stony, and black corals, sea pens, sea fans and whips, and sea anemones sensu lato (that is, in the broad sense—including all its subordinate taxa). The current classification of Anthozoa is listed below. Two groups of corals that lived in the Paleozoic Era, Rugosa and Tabulata, may be included in Anthozoa; but without knowing their internal anatomy, their placement and how modern anthozoans are related to them are uncertain.

Class Anthozoa
- Subclass Octocorallia (Alcyonaria)
  - Order: Alcyonacea [soft corals]
    - Helioporacea [blue corals]
    - Pennatulacea [sea pens]
- Subclass Hexacorallia (Zoantharia)
  - Order: Actiniaria [sea anemones in the strict sense]
    - Antipatharia [black corals]
    - Ceriantharia [tube anemones]
    - Corallimorpharia [corallike anemones]
    - Scleractinia [stony corals]
    - Zoanthidea [mat anemones]

The three groups of octocorals share so many features that there is no doubt they constitute a clade which, it has been proposed, merits class status. Some classification schemes put Ceriantharia in its own subclass because it differs so much in morphology from other hexacorals; although molecular sequence data support the distinctness of this group, whether it belongs with hexacorals is still debated. Previously, cerianthids were thought to be most closely allied to antipatharians, based on a resemblance between antipatharian polyps and cerianthid larvae; but differences between the two groups are so great that this relationship is no longer considered tenable. Actiniarians, scleractinians, and corallimorpharians share many morphological features, including details of mesentery and tentacle arrangement. The position of the Zoanthidea is obscure; these animals have some features reminiscent of rugosans, but because nothing is known about the soft tissue morphology of the latter, and the former do not form the skeletons, which are all that is known of rugosans, it is impossible to determine if this resemblance is phylogenetically significant. *See* PHYLOGENY.

Morphology. As is typical of cnidarian polyps, at the distal end of the cylindrical column is the oral disc (Fig. 1). The oval or slitlike mouth at its center, which opens into the actinopharynx, is surrounded by the tentacles, each of which arises from the space

between two mesenteries, so its cavity is part of the coelenteron. The eight tentacles of an octocoral are arrayed in a single cycle, as are the six in most antipatharians; those of a cerianthid are arrayed in two cycles, one immediately around the mouth and one at the margin of the oral disc; in other hexacorals, the tentacles may cover all or only part of the oral disc, and be arrayed radially or in cycles. In colonial anthozoans, the proximal (basal) end of the body is connected to other polyps or to a common mass of tissue from which other polyps also arise. In solitary anthozoans, the base may form a pedal disc (Fig. 1), by which the animal attaches to something solid, the base may be more or less pointed or bulbous and used for burrowing into and anchoring in soft sediments.

The coelenteron of an anthozoan is filled with seawater, which, in combination with the animal's musculature, serves as a hydrostatic skeleton. Water is driven into the coelenteron largely through the siphonoglyphs (Fig. 2) by the action of the cilia that distinguish these structures. The column has ectodermal longitudinal and endodermal circular muscles by means of which the animal shortens and narrows, respectively; the tentacle muscles are similarly arrayed. Each mesentery typically has longitudinal muscles along one side and transverse ones along the other.

The skeleton of a cnidarian, if present, functions only to support and protect the animal—it does not serve movement. Actiniarians, ceriantharians, and corallimorpharians do not form skeletons; zoanthideans of many species incorporate debris, such as sand grains and sponge spicules, into the mesoglea of the column wall, which provides a sort of adventitious skeleton. Mineralic skeletons, produced by all octocorals and scleractinians, are made of calcium carbonate in the crystal form aragonite. The scleractinian calcareous skeleton consists of a cup around each polyp, with radiating septa that parallel the mesenteries; in a colony, cups of adjacent polyps may share walls, or the cups may be separate.

Most octocorals form sclerites (sometimes termed spicules) that are commonly shaped like needles, discs, or rods and are embedded in the mesoglea (gelatinous layer between the ectoderm and endoderm) of living tissue. The sclerites of a helioporacean colony are fused into a solid skeleton that appears similar to scleractinian corals except that it is blue; the sclerites of the organ-pipe coral (order Alcyonacea) are fused into long red parallel tubes linked by transverse platforms. The color of these skeletons is imparted by iron salts, which scleractinian skeletons lack and so are white. A colony of antipatharians and some alcyonaceans, such as the sea fans and sea whips, is supported by an internal organic skeleton that may be calcified to some degree; the colony of some pennatulaceans has an internal stiffening rod of organic material.

Physiology. All cnidarians are carnivorous, and most capture large prey with the tentacles, which are richly supplied with stinging cells called nematocysts and, in hexacorals, with spirocysts; in dimorphic (or trimorphic) species of octocorals, polyps of only one type feed. The tentacles deliver the prey to the mouth and it is swallowed (through the actinopharynx). Small planktonic prey enter the coelenteron, with water pumped in to support the hydrostatic skeleton. Filaments borne along the free edges of many or all mesenteries, which consist largely of gland cells and nematocysts, enwrap prey items to effect digestion. Undigested food is egested through the single body opening, the mouth. Many shallow-water anthozoans, particularly tropical ones, contain zooxanthellae, microscopic photosynthetic organisms that have a symbiotic relationship with marine organisms of several kinds, and live within endodermal cells of anthozoans. Some of the carbon fixed by photosynthesis of the zooxanthellae is leaked to the anthozoan host; some of the oxygen generated in the photosynthetic process is presumably also available to the host. Reciprocally, nitrogenous wastes and carbon dioxide from the animal may be used by the photosymbionts. Like all cnidarians, anthozoans have no specialized structures for respiration and excretion, which are effected through the entire body surface.

Although sedentary, some anthozoans (especially actiniarians and pennatulaceans) are capable of limited locomotion as adults. The nerve net of an anthozoan has no aggregations of nerve cells or specialized nervous structures. Otherwise, however, the nerves function essentially as they do in other animals. Members of a colony may be poorly or well integrated neurologically. Some pennatulaceans and zoanthids are bioluminescent. *See* BIOLUMINESCENCE.

Reproduction and development. Anthozoans of all species can undergo sexual reproduction (although not all polyps of a colony may be able to do so). Some species of anthozoans are gonochoric (have separate sexes) whereas some are hermaphroditic; hermaphroditism can be serial (in which the animal changes from one sex to the other during its lifetime) or simultaneous (in which the animal produces eggs and sperm at the same time), and some hermaprodites are self-fertile. Gametes are formed in endodermal cells of the mesenteries and sink into the mesenterial mesoglea to mature; when ripe, the gametes break through the mesoglea and endoderm surrounding them, to be freed into the coelenteron. Typically they are then released (through the actinopharynx and mouth) into the surrounding water, where fertilization and larval development occur. The zygote develops into either a ciliated swimming larva, the planula, or directly into a young polyp. In some species, only sperm are released into the water, whence they are taken into a female and the eggs are fertilized; some or all of larval development occurs in the female's coelenteron or on the surface of her body.

Many, but not all, anthozoan species can undergo asexual reproduction by a variety of processes, including budding (a term that is not synonymous with asexual reproduction), fission, and laceration. The products of asexual reproduction are genetically identical to one another, thus constituting a clone;

if the clonemates remain physically attached to one another, the multipolyp result is termed a colony. Asexual reproduction requires regenerative powers to reconstitute the missing body parts; members of species that do not undergo asexual reproduction typically have poorer regenerative powers than those that do, and some anthozoans, while able to heal, cannot regenerate at all. *See* REGENERATION (BIOLOGY).

Ecology. Anthozoans occur in all oceans of the world at all depths, including trenches. They can be ecologically very important, and they dominate the substratum in some habitats. Since most anthozoans require full or nearly complete salinity, and are intolerant of sedimentation and desiccation, they are poorly represented in estuaries and turbid environments, and high in the intertidal zone. Many groups are most diverse in the shallow tropics, but all groups (including scleractinian corals) occur in cold and deep water. [Deep-water scleractinians, however, do not live deeper than 5–6 km (16,404 ft–19,685 ft), probably due to the decrease in concentration of calcium carbonate in seawater with depth.]

Because species of coral that build shallow-water reefs are adversely affected by low temperature, coral reefs are confined to tropical and subtropical regions. Nearly all species of these reef-forming corals contain zooxanthellae, so tropical coral reefs are also confined places that can be reached by sufficient sunlight for photosynthesis [thus not at depths much greater than about 100 m (330 ft) or in caves]. Deep-water coral reefs differ from shallow tropical reefs: they are not latitudinally confined; the diversity of species is much lower; octocorals are relatively more important in forming them; and the scleractinians have less robust skeletons than their tropical reef-forming counterparts (which seems to be related to their lack of zooxanthellae). *See* REEF.

Daphne G. Fautin

Fossil record and evolution. Fossil anthozoans are known from the late Precambrian (during the Ediacaran geologic period, about 600 million years ago) and from rocks of every later geologic period. Fossilized soft polyps are extremely rare, but several groups of anthozoans with a mineralic external skeleton were well preserved and have good (and even excellent) records. The Paleozoic Rugosa and Tabulata are common in rocks of Middle Ordovician to Permian age, and the Mesozoic-Cenozoic Scleractinia are common from the Middle Triassic to the present. Rugosans and tabulates were important reef builders in the Middle Silurian to early Late Devonian, and scleractinians were important reef builders from the Jurassic through the Cretaceous, and again from the Miocene to the present. Fossil Octocorallia are represented by impressions from the Ediacaran period, and although sclerites are recognized sporadically through time, the fossil record is not good enough to trace the evolutionary pattern within this group or its relationship to the Hexacorallia. *See* EDIACARAN BIOTA; GEOLOGIC TIME SCALE; RUGOSA; TABULATA.

The relationship among the classes of Cnidaria had been uncertain until the discovery that the mitochondrial deoxyribonucleic acid (DNA) of anthozoans is circular, as it is in all non-cnidarians, whereas that of cubozoans, hydrozoans, scyphozoans, and staurozoans is linear. Thus the non-anthozoans, now commonly referred to as the Medusozoa, form a clade that is derived (evolutionarily advanced), in this feature relative to Anthozoa, so Anthozoa is the sister group to all other cnidarians. *See* ANIMAL EVOLUTION; CUBOZOA; HYDROZOA.

William A. Oliver, Jr., Daphne G. Fautin

Bibliography. D. F. Dunn, Cnidaria, pp. 669–706 in *Synopsis and Classification of Living Organisms*, vol. 1, S. P. Parker (editor-in-chief), McGraw-Hill, New York, 1982; L. H. Hyman, *The Invertebrates: Protozoa Through Ctenophora*, vol. 1, McGraw-Hill, New York, 1940; J. W. Wells and D. Hill, Anthozoa—general features, pp. F161–F165 in *Treatise on Invertebrate Paleontology*, vol. F: *Coelenterata*, Geological Society of America and University of Kansas Press, Lawrence.

Anthracosauria

An order of Paleozoic tetrapods that arose in the Early Carboniferous (Mississippian) (345 million years ago) and lasted until the Early Triassic (about 245 mya). The term anthracosaurs (meaning "coal reptiles") encompasses a range of genera that may not all be close relatives; rather, they form a stem group (members of the lineage leading to amniotes, but not amniotes themselves) possibly related to amniotes (reptiles, birds, and mammals). They share a number of features, such as the pattern of skull roof bones and vertebral construction, that distinguish them from temnospondyls, a larger group of Paleozoic tetrapods that are more closely related to modern amphibians. Anthracosaurs also share a number of features that may be primitive for tetrapods, including a "skull table" unit that is only loosely attached to the cheek, and a palate in which bones almost meet along the midline.

Anthracosaurs include both terrestrial and aquatic forms, with a subgroup, the embolomeres, being large, long-bodied, crocodilelike fish eaters. Embolomeres include genera such as *Pholiderpeton* and *Anthracosaurus* (**illus.** *a*) from the Carboniferous of Great Britain and *Archeria* from the Early Permian (about 290 mya) of the United States. Embolomeres are characterized by having two ring- or horseshoe-shaped parts forming the central portion (centrum of their vertebrae), and a prong or "horn" on the tabular bone, at the back of the skull (illus. *a*).

The earliest anthracosaurs, *Silvanerpeton* (illus. *b*) and *Eldeceeon*, are from Scotland and are Viséan in age (about 340 million years old). *Eldeceeon* was terrestrial, but *Silvanerpeton* was probably partially aquatic. Another mid-Carboniferous (about 320 mya) form, *Eoherpeton*, was also probably terrestrial, though it is known only from a few incomplete specimens. *Gephyrostegus*, from the Late Carboniferous (about 300 mya) of eastern Europe, is a terrestrial form which at one time was thought to be a close ancestor of reptiles, though its position among the anthracosaurs is

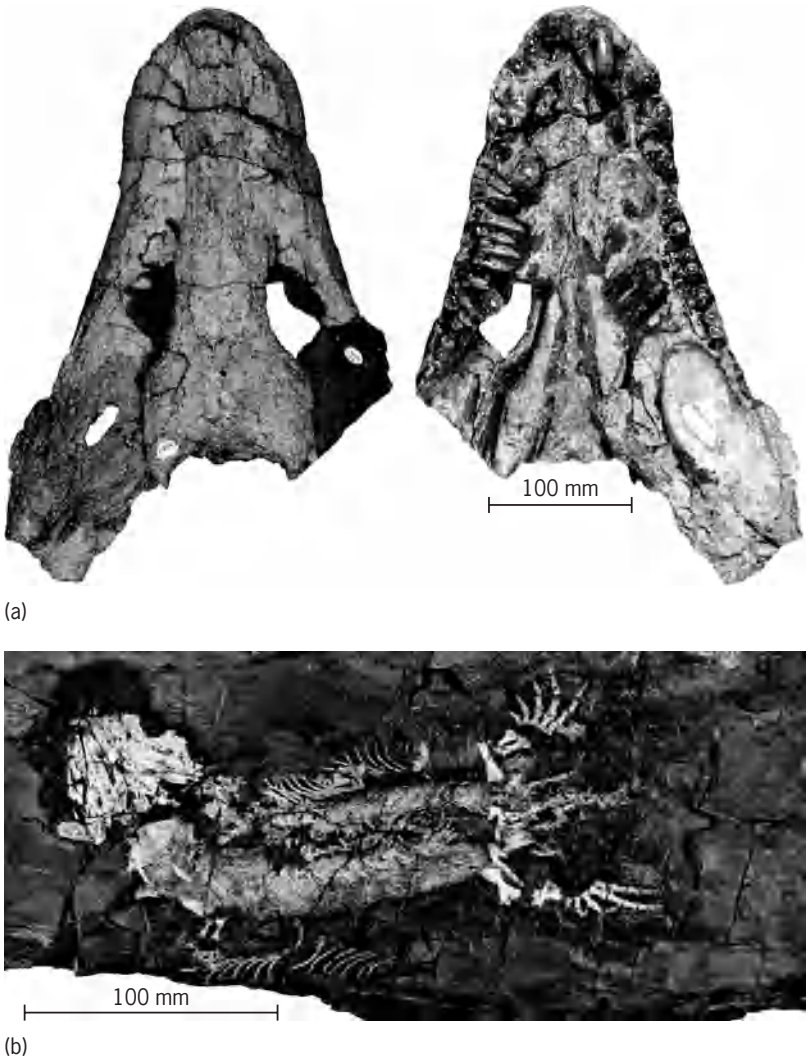

Anthracosaur features. (*a*) *Anthracosaurus russelli*, one of the largest and fiercest anthracosaurs, from Great Britain. The left view shows the skull roof, with the central "skull table" bearing a short tabular horn; the right view shows the palate with enormous teeth. (*b*) *Silvanerpeton miripedes*, one of the earliest anthracosaurs, from Scotland. The head is flattened, and the body has been "kippered (slit open)," with the ribs separated from the vertebral column. The dermal gastralia (scalelike structures) formed a protective armour (seen well here) down the middle of the animal.

disputed. The most recent anthracosaurs (last in their line) have been found in Russia, where some embolomere-like forms have reverted to terrestrial habits. *See* AMNIOTA; AMPHIBIA; TEMNOSPONDYLI; TETRAPODA. Jennifer A. Clack

Bibliography. M. J. Benton, *Vertebrate Palaeontology*, 3d ed., Blackwell, Oxford, 2005; R. L. Carroll, *Vertebrate Paleontology and Evolution*, W. H. Freeman, New York, 1987; J. A. Clack, *Gaining Ground: The Origin and Early Evolution of Tetrapods*, Indiana University Press, Bloomington, 2002; T. R. Smithson, Anthracosaurs, in *Amphibian Biology*, vol. 4, pp. 1053–1063, ed. by H. Heatwole and R. L. Carroll, Surrey Beatty, Chipping Norton, NSW, 2000.

Anthrax

An acute, infectious worldwide zoonotic disease caused by the spore-forming bacterium *Bacillus anthracis* to which most animals, especially grazing herbivores, are susceptible. In natural conditions, human infections (predominantly cutaneous) usually result from contact with infected animals or contaminated animal products, such as hides or wool. The awareness of the use of *B. anthracis* as a bioterrorist weapon has considerably increased. Anthrax is endemic as a zoonosis in many areas of Africa, Asia, and the Americas, where spores can lie dormant in the soil for many years and commonly affect grazing animals such as sheep, cattle, and goats. *See* ZOONOSES.

Bacteriology. *Bacillus anthracis*, the causative agent of anthrax, derives its name from the Greek word for coal, *anthraxis*, because of the black skin lesions it causes. *Bacillus anthracis* is the only obligate pathogen within the genus *Bacillus* and is a large gram-positive, aerobic, spore-bearing bacillus, 1–1.5 × 3–10 micrometers in size. It is nonmotile and nonhemolytic (does not destroy red blood cells) on horse or sheep's blood agar, grows at 37°C, and forms typical white colonies, which are characteristically tacky on teasing with a loop. The oval spores

(0.75 × 1 μm) resist drying, cold, heat, and disinfectants and can survive in soil for decades.

Spores do not form in host tissues, unless the infected body fluids are exposed to air that contains a lower level of carbon dioxide. Provided no antibiotics have been prescribed, *B. anthracis* is readily isolated from clinical specimens such as skin or blood cultures. In environmental samples, because of the presence of other nonpathogenic *Bacillus* spp., a selective medium such as polymixin lysozyme ethylenediaminetetraacetic acid (EDTA) thallus acetate agar is recommended for isolating specimens.

R. C. Spencer

Pathogenesis. *Bacillus anthracis* possesses three virulence factors: lethal toxin, edema toxin, and a poly-D-glutamic acid capsule. Lethal toxin is composed of two proteins, lethal factor (LF, 82 kilodaltons, a metalloprotease) and protective antigen (PA, 63 kDA), complexed at a ratio of seven protective antigens per lethal factor. The protective antigen is produced by the anthrax bacillus at a molecular weight of 83 kDA, but must be cleaved by either serum or target cell surface proteases to 63 kDa before it complexes with lethal factor to form lethal toxin. The edema toxin is composed of edema factor (EF, 89 kDa, adenylate cyclase) and protective antigen, and it is believed to complex in a manner similar to that seen for lethal toxin. Protective antigen plays a central role in that it is required for transport of lethal factor and edema factor into host target cells. Macrophages appear to be the primary host target cells for lethal toxin, whereas neutrophils appear to be the target cells for edema toxin in addition to other cells involved in edema formation. The third virulence factor is the capsule, which inhibits phagocytosis through its negatively charged poly-D-glutamic acid composition.

All three toxin components are encoded by a plasmid, pXO1, whereas the genes required for capsule synthesis (cap B, cap C, and cap A) are encoded by the smaller pXO2 plasmid. The pXO2 plasmid also encodes a depolymerase enzyme, which cleaves the capsule to form small glutamic acid peptides that also appear to play a role in virulence. Strains lacking either or both plasmids are avirulent, such as the veterinary vaccine Sterne strain, which lacks the pXO2 plasmid.

John W. Ezzell

Infection in humans occurs after introduction of the spore through a break in the skin (cutaneous anthrax) or entry through mucosa (gastrointestinal anthrax). After ingestion by macrophages at the site of entry, germination of the vegetative form then occurs, followed by extracellular multiplication, together with the production of capsule and toxins. In inhalation anthrax, spores of 1–2 μm in diameter are inhaled and deposited in the alveolar spaces from where they are transported by the lymphatics to the mediastinal lymph nodes, where they cause mediastinal lymphadenopathy and hemorrhage (that is, enlarged mediastinal lymph glands and bleeding within them). Vegetative bacteria then spread through blood and lymph where, if unchecked, they cause septicemia (blood poisoning). In turn, large amounts of exotoxins are produced, which are responsible for the overall symptoms and death.

Clinical features. The symptoms of anthrax infection in humans are dependent on the route through which the infection is contracted:

Cutaneous anthrax. Cutaneous anthrax is by far the most common form, accounting for over 95% of naturally occurring cases. It is caused by direct contact with spores, especially via open wounds. Following an incubation period of 1–3 days, localized itching precedes the appearance of a papular lesion that turns vesicular; 2–6 days later the lesion develops a black eschar (dry crust), giving anthrax its name.

Inhalation/pulmonary anthrax. Inhalation/pulmonary anthrax follows infection via the lungs, where the spores can easily cross into the blood at the level of the alveoli. Initially, infection resembles the common cold and is characterized by nonspecific symptoms including fever, nonproductive (dry) cough, malaise, and fatigue. Chest x-ray at this stage often shows mediastinal widening (due to mediastinal lymphadenopathy) and pleural effusions (abnormal accumulation of fluid in the pleural space, the area between the membranes lining the lungs and the chest cavity). Two to five days after initial symptoms there is an abrupt onset of marked pyrexia (fever) and severe respiratory depression often accompanied by dyspnea (labored breathing), cyanosis (bluish tinge due to lack of oxygen), and stridor (peculiar, harsh, vibrating sound produced during respiration). Untreated, patients will go into shock and die within 24–36 h.

Gastrointestinal anthrax. Gastrointestinal anthrax is rare and usually follows consumption of contaminated meat. It is characterized by acute inflammation of the gastrointestinal tract. Initial signs of nausea, anorexia, vomiting, and pyrexia are followed by abdominal pain, hematemesis (vomiting of blood), and severe diarrhea, which is often bloody. Bacteremia (presence of bacteria in the blood) may develop after 2–3 days, which if untreated is usually fatal.

Person-to-person transmission. Person-to-person spread of anthrax is very rare, and contacts need not be treated unless they have been exposed to the same source of infection. Although direct exposure to exudates (excretions) from cutaneous anthrax lesions may result in secondary cutaneous infection, there have been no known cases of person-to-person transmission of inhalation disease.

Treatment. Antibiotic therapy is the recommended treatment for anthrax. Route of administration and type of medication vary depending on the type of disease, whether diagnosis has been confirmed, and organism sensitivity.

Inhalation and gastrointestinal anthrax. When the diagnosis of inhalation or gastrointestinal anthrax is suspected but not confirmed, it may be necessary to start empirical treatment to protect against the possibility of anthrax. However, in these circumstances, it will also be necessary to treat concurrently for other causes of acute respiratory illness. Recommended initial treatment for inhalational and gastrointestinal anthrax is with intravenous ciprofloxacin (with

a change to oral delivery when appropriate) for 60 days. Treatment can be changed to penicillin if the organism is found to be sensitive.

Cutaneous anthrax. Treatment for cutaneous anthrax should be initiated with oral ciprofloxacin 500 mg twice daily for 7 days. This can be changed to penicillin if the organism is found to be sensitive. If there is suspicion of deliberate release, treatment may need to be continued for up to 60 days to protect against inhalation anthrax, which may have been acquired concurrently.

Prophylactic treatment. In the event of a known exposure to anthrax spores, antibiotic prophylaxis should be initiated as soon as possible with oral ciprofloxacin (a fluoroquinolone) or with doxycycline (a tetracycline) if fluoroquinolones are unavailable or contraindicated. Prophylaxis should continue until *B. anthracis* exposure has been excluded. If exposure is confirmed, prophylaxis should continue for 60 days. During this period, no special precautions are required for exposed persons; however, they should receive an anthrax information sheet and be instructed to seek medical attention immediately in the event of any suspicious symptoms.

Pediatric use of fluoroquinolones and tetracyclines can be associated with adverse effects that must be weighed against the risk of developing a serious disease. If *B. anthracis* exposure is confirmed, the organism must be tested for penicillin susceptibility. If susceptible, exposed children may be treated with oral amoxycillin. *See* ANTIBIOTIC.

Vaccination. Vaccination is the most cost-effective form of mass protection. Although the first anthrax animal vaccine was developed by Pasteur in 1881, human vaccines did not emerge until the middle of the twentieth century.

Animal vaccines. Pasteur's initial vaccine had problems with declining potency and variability in immunogenicity (ability to induce an immune response), which led to the search for a more stable and effective vaccine. Sterne developed an attenuated spore vaccine, based on an avirulent nonencapsulated strain derived from a subculture from a case of bovine anthrax. This has proved to be extremely safe and effective, requiring little modifications. Although extremely effective, repeated vaccinations are required for long-term protection, because a single dose provides immunity for only about one year.

Human vaccines. Vaccination of humans with live spores, either by scarification or subcutaneous injection, was developed in the former Soviet Union and China. In Russia a strain analogous in its derivation to the Sterne strain is used.

The vaccine developed in the United States is a filtrate of a bovine strain of *B. anthracis* V770-NPI-R. It was licensed in 1972 for administration to those in at-risk occupations. It is not licensed for use in children or pregnant women, and at present is not available for civilian use. The vaccine is licensed to be given in a six-dose series over an 18-month period. The current vaccine from the United Kingdom consists of a filtrate of an aerobic supernatant from the nonencapsulated Sterne strain of *B. anthracis*. The UK vaccine was introduced for workers in at-risk occupations in 1965, and licensed for general human use in 1979. At present, on empirical grounds, boosters are administered 6 months after the initial series of three doses (at 0, 3, and 6 weeks), and annually thereafter.

Future vaccines. Considerable effort has been expended in the development of next-generation vaccines to meet current licensing criteria, but with improved safety and efficacy performance characteristics. The favorites are subunit vaccines that contain whole-length (83,000 kDa) recombinant protective antigen. *See* VACCINATION. R. C. Spencer

Bibliography. L. Baillie and T. D. Read, *Bacillus anthracis*: A bug with attitude, *Curr. Opin. Microbiol.*, 4:78–81, 2001; T. C. Dixon et al., Anthrax, *N. Engl. J. Med.*, 341:815–826, 1999; J. Guilleman, *Anthrax: The Investigation of a Deadly Outbreak*, University of California Press, Berkeley, 1999; R. C. Spencer, *Bacillus anthracis*, *J. Clin. Pathol.*, 56:182–187, 2003; M. N. Swartz, Recognition and management of anthrax—an update, *N. Engl. J. Med.*, 345:1607–1610, 2001; P. C. B. Turnbull, Definitive identification of *Bacillus anthracis*—a review, *J. Appl. Microbiol.*, 2:237–240, 1999.

Anthropology

The observation, measurement, and explanation of human variability in time and space. This includes both biological variability and the study of cultural, or learned, behavior among contemporary human societies. These studies are closely allied with the fields of archeology and linguistics. Studies range from rigorously scientific approaches, such as research into the physiology, demography, and ecology of hunter-gatherers, to more humanistic research on topics such as symbolism and ritual behavior. *See* ARCHEOLOGY; PHYSICAL ANTHROPOLOGY.

Anthropology lacks a unified theory comparable to neo-Darwinian evolution in the biological sciences and is characterized, instead, by a wide variety of subfields that analyze and integrate studies of human behavior in different ways. Social-cultural anthropology examines the various ways in which learned techniques, values, and beliefs are transmitted from one generation to the next and acted upon in different situations. Most studies stress the historical development and internal structure and workings of particular cultural traditions, and anthropologists have amassed detailed bodies of documentation on different human societies, ranging from small, nomadic hunter-gatherer groups like the Eskimo or Australian aboriginals to complex, sedentary societies like India, Iran, or Japan. Such ethnographic studies are carried out within a relativistic framework which assumes that one cannot evaluate the behavior and attitudes of one society by applying the standards of another. This attitude has been a useful corrective to earlier attempts to judge non-Western behavior by Western standards. Given this view, social-cultural

anthropologists strive, by various means, to achieve a reliable understanding of each cultural system in its own terms, with as little distortion as possible due to ethnocentrism.

Significant, too, within social-cultural anthropology are cross-cultural studies that seek to identify essential structural or behavioral properties of human society. While anthropologists like Leslie White have proposed general laws of cultural evolution, others like Claude Lévi-Strauss have argued for universal principles of human cognition by examining how people in different societies classify natural and social phenomena by means of logical systems based upon the principle of binary opposition. Modern scholars have sought to identify universal patterns of symbolic behavior and belief, and there are other social-cultural anthropologists actively testing these kinds of propositions in particular cases.

Increasingly, too, social-cultural anthropologists have applied their training and skills to issues of contemporary importance such as economic development in third world countries, public policies affecting ethnic minorities, and changes arising from contact between different societies (especially Western and non-Western ones). Sometimes referred to as applied anthropology, such studies are often made in situations where conflicting social values or expectations may arise. A cogent example is a series of studies into the social consequences arising from technical and economic changes due to the so-called Green Revolution, involving the introduction of new and more productive strains of rice into traditional rice-growing societies in southern Asia. This nonacademic area of social-cultural anthropology, where application is emphasized over basic research, assumed a larger role in social planning and policymaking of all kinds, ranging from the actions of individual corporations (such as Western oil companies operating in traditional Moslem societies in the Middle East) to national governments.

Cultural linguistics is closely allied with both the goals and methods of social-cultural anthropology, especially with respect to the way in which linguists strive for a reliable understanding of how each different language works according to its own sound system (phonology) and grammatical structure. Linguists have produced a growing body of valuable sources, including dictionaries, grammars, and texts of particular languages and comparative analyses, that enable them to infer historical relationships between different languages that share common ancestry. Such controlled, empirical studies, free of ethnocentric bias, also provide finely tuned views of verbal and behavioral interaction in the context of particular situations, such as non-Western artistic performances or rituals. *See* PSYCHOLINGUISTICS.

There has also been a developing tendency in anthropology toward integration of different subfields. For example, ethnoscience is a subject in which anthropologists apply approaches derived from linguistics to understand the grammatical structure and manipulation of cognitive perceptions by people in different societies of such things as color, weather, and biotic environment. Studies range from traditional Philippine farmers' plant taxonomies and agricultural practices to long-distance sailing and navigation by Micronesian islanders in outrigger canoes. Another growing subfield is ethnoarcheology, in which observations of material behavior (especially discard) in contemporary societies are used to interpret the archeological remains of prehistoric cultures. Modern anthropology is characterized by its breadth and diversity of approaches to the study of variability in human behavior. Richard A. Gould

Bibliography. M. Douglas, *Natural Symbols,* 1972; M. Harris, *The Rise of Anthropological Theory,* 1968; D. Levinson and Martin J. Malone, *Toward Explaining Human Culture: Critical Review of the Findings of Worldwide Cross-Cultural Research,* 1980; C. Lévi-Strauss, *The Savage Mind,* 1967; E. F. Moran, *Human Adaptability,* 1979; J. E. Pfeiffer, *The Emergence of Man,* 1978; E. B. Ross (ed.), *Beyond the Myths of Culture: Essays in Cultural Materialism,* 1980.

Anthropometry

The systematic quantitative representation of the human body. Anthropometric techniques are used to measure the absolute and relative variability in size and shape of the human body.

Instrumentation. Depending on the objective, anthropometric techniques involve instrumentation that may include weighing scale, anthropometer, skinfold calipers, body volume tanks, and bioelectrical impedance analyzers that are used for measurements of body size and body composition. Similarly, radiographic instruments and x-ray scanners such as dual-energy-ray absorption meters and ultrasound densitometers are used for quantifying cortical bone density, bone mass, subcutaneous fat density, and lean body mass.

Quality control. Anthropometry follows a rigorous set of guidelines that include standardization of the measurement techniques, uniform landmarks, and establishing conditions of the measurements. Despite its simplicity, it requires a great deal of quality control to ensure minimal error of measurements. Such an objective is usually obtained through meticulous training in the technique that results in replicable measurements.

Applications. In biological anthropology and human paleontology, anthropometry is the technique of choice for quantifying variability and relationship of fossils and extant populations. Anthropometric measurements of the head, face, and long bones are also used in analyzing fossil taxa (using measurements from radiographs). The mathematical approaches are used in describing the size and proportions of various fossil hominids. For example, through measurements of long bones it is known that the earliest human ancestors such as *Australopithecus afarensis* were about 3.5–5 ft (1.05–1.5 m) tall. In the same vein, it has been established that *Homo erectus* and Neandertals were

as tall as modern-day humans, contrary to previous assumptions. *See* ANIMAL GROWTH; BIOMETRICS.

Anthropometry is the most universally applicable, inexpensive, and noninvasive method available to assess the nutritional history throughout life. It has been used to assess and predict the health, nutritional history, welfare, performance, and survival of nonindustrialized and industrialized societies. For example, since fat is the main form of energy storage, and body muscle is composed largely of protein, anthropometric measurements of body composition provide indirect estimates of energy and protein reserves of the body. These reserves are sensitive to changes in dietary intake. Reserves can be depleted during chronic malnutrition, resulting in muscle wasting, while during overnutrition reserves can grow, resulting in obesity. *See* NUTRITION.

To achieve such objectives, anthropologists and population biologists have developed various references that can be used as base lines for expressing the absolute and relative deviation from the average. Techniques of data analysis include the expression of individual values in the form of Z scores (the individual value minus the reference mean for the age and sex, divided by the corresponding standard deviation). This approach permits the investigator to express the measurements in terms of Z score units from the mean. Another approach involves expressing individual values in the form of percentiles placement. For this purpose, the investigator needs to compare the individual value to the percentile ranges given in the anthropometric standard. Thus, an individual measurement may be expressed as being either close to the 50th percentile or above or below the 95th or 5th percentile.

Anthropometry is also essential to the field of forensics, specifically forensics anthropology, which is concerned with the relationship between medicine and the law. Forensic anthropologists (so-called bone detectives) make extensive use of anthropometry in human identification, whether for isolated cadavers, commingled remains, victims of mass disasters, or genocide victims. As applied to metacarpal and phalangeal (reterning to bones of the hand) patterns, anthropometry helps in the diagnosis of congenital malformation syndromes (dysmorphogenesis states), even in fetuses. Anthropometric measurements of the head and face are also extensively employed in orthodontic diagnosis, in treatment planning, and following orthodontic treatment. Measurements made from cephalometric radiographs also serve in the identification of syndromes. Further extensions of cephalometry measurement of the living human head), in three dimensions (cartesian anthropometry), are used in sculpting head forms not just for equipment design but for use in the plastic and reconstructive surgery of accident victims. Forensic anthropologists are involved in the identification of human remains using techniques that include direct skeletal measurements, allowing the identification of the age at death, sex, and cause of death, and deduction of the life history of the deceased. For example, identification of healed bone fractures, age, and sex of Neandertals has permitted anthropologists to infer the life-span and hunting techniques used. Analysis of the skeletal remains revealed that these fossil humans did not live beyond the age of 50 years and practiced manual (hand-to-hand) hunting of their game, resulting in fractures similar to those suffered by rodeo cowboys. *See* NEANDERTALS.

The use of anthropometry for the design of clothing, equipment of all sorts, and interiors is relatively new in the history of humankind. For example, through anthropometric techniques to establish human dimensions, gas masks, oxygen masks, dust masks, and respirators as well as military helmets were designed. Anthropometry has enabled the production of clothes that accurately match the various sizes of the population. The dimensional standards used for sports equipment and military applications have been designed to match the distribution of size in the population as whole. Hence, anthropometry plays an essential role in the armed forces, by providing information about the necessary dimensions for military equipment of all sorts to accommodate individuals in the military service. Similarly, the equipment used in space demands knowledge of human dimensions. Anthropometry has evolved from the manual to automated approach, whereby measurements are entered directly into the computer for data processing.

Anthropometry plays an essential role in all areas of human endeavor concerned with the relative and absolute quantification of the human body. It serves a profound function in all aspects of science and human activity. *See* ANTHROPOLOGY; PHYSICAL ANTHROPOLOGY.

A. Roberto Frisancho

Bibliography. W. M. Bass, *Human Osteology: A Laboratory and Field Manual of the Human Skeleton*, 4th ed., 1995; B. Bogin, *Pattern of Human Growth*, 1988; C. Bouchard, R. Malina, and L. Perusse, *Genetics of Fitness and Physical Performance*, 1997; P. B. Eveleth and J. M. Tanner, *Worldwide Variation in Human Growth*, 1976; A. R. Frisancho, *Anthropometric Standards for the Assessment of Growth and Nutritional Status*, 1990; S. Garn, Physical anthropology: Methods, *Med. Phys.*, 3:25–30, 1960; T. Lohman, A. Roche, and R. Martorell (eds.), *Anthropometric Standardization Reference Manual*, 1988; F. Spencer (ed.), *History of Physical Anthropology*, vol. 1, 1997.

Antibiotic

The original definition of an antibiotic was a chemical substance that is produced by a microorganism and, in dilute solutions, can inhibit the growth of, and even destroy, other microorganisms. This definition has been expanded to include similar inhibitory substances that are produced by plants, marine organisms, and total- or semi-synthetic procedures.

History. The first to apply antibiotic therapy, over 2500 years ago, were the Chinese. They were aware

of the therapeutic properties of moldy soybean curd applied to infections, and used this material as standard treatment. L. Pasteur and J. F. Joubert recognized the clinical potential of antibiosis in 1877. They observed that large quantities of anthrax bacilli could be given to an animal without adverse effects, provided that ordinary bacteria were given at the same time. The modern era of antimicrobial therapy was launched in the mid-1930s with the discovery of prontosil; sulfamethoxazole, a derivative of this synthetic compound, is still in clinical use (see **table**). Penicillin, discovered by A. Fleming in 1928, was not developed until the early 1940s because of the low potency and high instability of the crude material. The potency and therapeutic potential of penicillin was recognized by H. Florey and E. Chain. World War II provided the impetus for the Anglo-American research effort that led to the large-scale production of penicillin and its development as a systemic therapeutic agent. In the mid-1940s, cephalosporin C was discovered by G. Brotzu and streptomycin by S. Waksman and his associates. The latter discovery was the result of a well-planned, systematic, and intensive search for antibiotics in soil actinomycetes and fungi. Subsequent observations on the mode of action of bacitracin, penicillin, polymyxin, gentamycin, chloramphenicol, tetracycline, and amphotericin revealed the importance and therapeutic potential of antibiotics and added momentum to the search for new ones. Since the 1940s, thousands of antibiotics have been isolated and identified; some have been found to be of value in the treatment of infectious disease. They differ markedly in physicochemical and pharmacological properties, antimicrobial spectra, and mechanisms of action.

Biosynthesis. Penicillin is produced by strains of the fungus *Penicillium notatum* and *P. chrysogenum*. Most of the other antibiotics in clinical use are produced by actinomycetes, particularly streptomycetes (natural antibiotics) [see table]. Other antibiotics, such as nalidixic acid and clotrimazole, are produced by chemical synthesis (synthetic antibiotics). Based on structure, the major antibiotic classes are the β-lactams (penicillins and cephalosporins), aminoglycosides, macrolides, tetracyclines, quinolones, rifamycins, polyenes, azoles, glycopeptides, and polypeptides.

Manufacture. The key step in the production of natural antibiotics is a fermentation process. Strains of microorganisms, selected by elaborate screening procedures from randomly isolated pure cultures, are inoculated into sterile nutrient medium in large vats and incubated for varying periods of time. Both cell growth and antibiotic production, which are influenced by culture conditions such as composition of the medium, temperature, aeration, and pH, are monitored during this period.

Different strains of a single microbial species may differ greatly in the amounts of antibiotics they produce. Strain selection is thus the most powerful tool in effecting major improvements in antibiotic yield. In addition, variations in culturing conditions often markedly affect the amount of antibiotic that is produced by a given strain. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; FERMENTATION.

The production of an antibiotic in quantity by fermentation and its subsequent recovery from the culture medium usually require 5–8 days. The process can be shortened sometimes by establishing the optimal conditions for growth and antibiotic production. Paper, thin-layer, and high-pressure liquid chromatography, techniques designed to separate minute quantities of different substances, allow rapid detection of the antibiotic compounds. Fluorescence, ultraviolet, and infrared spectroscopy are often used to monitor antibiotic distribution in various fractions, obtained by such procedures as solvent extraction and column or high-pressure liquid chromatography. *See* CHROMATOGRAPHY; SPECTROSCOPY.

Chemical modifications of antibiotics produced by fermentation processes have led to semisynthetic ones with improved antimicrobial activity or pharmacological properties. For example, chemically modified penicillins, such as methicillin, are effective against bacteria resistant to the parent compound, penicillin G. Other penicillins, such as ampicillin, are active against a wider variety of microorganisms than the parent compound. Still others are better absorbed by the body or have greater chemical stability, more favorable tissue distribution and excretion, or reduced toxicity.

In the United States, the Federal Pure Food and Drug Act of 1906, as amended in 1938 and 1962, requires that all drugs for human use must be approved by the Food and Drug Administration (FDA) for purity, efficacy, and safety. Before an antibiotic is tested in humans, extensive microbiological, pharmacological, and toxicological studies are needed to support claims for efficacy and safety. The data are compiled and submitted to the FDA as part of an application for an Investigational New Drug. If the FDA standards are met, the antibiotic is approved for human clinical trials. The trials study the safety and pharmacokinetic profile of the new drug in healthy volunteers (phase I trials) and its efficacy in patients with different infectious diseases (phase II and III trials). Data from these trials are then submitted to the FDA to support a New Drug Application. If approved by the FDA, the antibiotic can be marketed for human use. Other countries have their own regulatory agencies, with most European nations bound by the European Economic Community's rules and regulations. *See* PHARMACEUTICALS TESTING.

Antimicrobial activity. In general, microorganisms are divided into bacteria, fungi, viruses, and protozoa. All four groups can cause infectious diseases in animals and humans, though the majority of infections are caused by bacteria. Most antibiotics are active against bacteria. Although for the proper treatment of serious infections cultures and antibiotic sensitivities are required, antibiotic therapy is often empiric, with etiology being inferred from the clinical features of a disease.

Bacteria. Bacteria are divided by a staining reaction into the gram positive and the gram negative; each

group comprises a wide variety of different species. Gram-positive bacteria may be subdivided by another staining reaction into acid-fast and non-acid-fast bacteria. *See* STAIN (MICROBIOLOGY).

Staphylococci, pneumococci, and streptococci are the more common gram-positive organisms, while enterobacteria, *Pseudomonas*, and *Hemophilus* are the most common gram negative. Tubercle bacilli are the most important acid-fast organisms. Certain antibiotics, such as erythromycin and vancomycin, are effective only against gram-positive bacteria. Others, such as cephalosporins, quinolones, tetracyclines, and chloramphenicol, are effective against both gram-positive and gram-negative bacteria and are referred to as broad-spectrum antibiotics. *See* BACTERIA; MEDICAL BACTERIOLOGY.

Fungi. Pathogenic fungi may be divided on the basis of their pathogenicity into true pathogens, such as *Histoplasma*, and opportunistic pathogens, such as *Candida*, *Aspergillus*, and *Cryptococcus*. The opportunistic occur mainly in debilitated and immunocompromised patients. Although hundreds of antibiotics active against fungi have been isolated, clinically useful ones are amphotericin B, nystatin, griseofulvin and the azole antifungals. As broad-spectrum antifungals, amphotericin is active against systemic infections and nystatin against local infections. Griseofulvin is active against fungi that cause skin diseases. Azoles, which constitute perhaps the largest group of antifungals, are broad-spectrum and synthetic. *See* FUNGI; MEDICAL MYCOLOGY; OPPORTUNISTIC INFECTIONS.

Viruses. Viruses may be divided on the basis of nucleic acid content (RNA and DNA), morphology, and composition of the virus shell (enveloped and nonenveloped viruses). With some viruses that cause mild infections, such as the common-cold viruses (rhinoviruses), treatment is symptomatic. With others, such as the polio, smallpox (now eradicated), and hepatitis B viruses, the only way to prevent disease is by vaccination. With still other viruses (see table), antibiotics, mostly synthetic, are the appropriate treatment. Clinically useful antibiotics are ribavirin, acyclovir, and zidovudine, which are active against, respectively, respiratory, herpes, and human immunodeficiency viruses. *See* ANIMAL VIRUS; VACCINATION.

Protozoa. Protozoa may be divided, on the basis of the site of infection, into intestinal (*Entamoeba histolytica, Giardia lamblia*), urogenital (*Trichomonas vaginalis*), blood (*Plasmodium falciparum*), and tissue (*Pneumocystis carinii*, which was reclassified as a fungus). Protozoan diseases such as malaria, trypanosomiasis, and amebiasis are particularly common in the tropics, in populations living under poor housing and sanitary conditions. In the developed countries, *P. carinii* is the most important opportunistic pathogen, being associated almost exclusively with acquired immune deficiency syndrome (AIDS). Antibiotics active against protozoa include metronidazole, trimethoprim-sulfamethoxazole, and quinine. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); MEDICAL PARASITOLOGY; PROTOZOA.

Antitumor activity. The pioneering observation in 1952 on the antitumor activity of actinomycin sparked an intensive search for antitumor antibiotics in plants and microorganisms. Among the antibiotics used clinically against certain forms of cancer are daunorubicin, doxorubicin, mitomycin C, and bleomycin. Although antitumor antibiotics are significantly toxic to the patient, the lack of nontoxic therapeutic alternatives has required their continued use. *See* CANCER (MEDICINE).

Mechanism of action. Antibiotics active against bacteria are bacteriostatic or bacteriocidal; that is, they either inhibit growth of susceptible organisms or destroy them. Bactericidal activity may be necessary for the eradication of infections where the host immune system is impaired. On the basis of their mechanism of action, antibiotics are classified as (1) those that affect bacterial cell-wall biosynthesis, causing loss of viability and often cell lysis (penicillins and cephalosporins, bacitracin, cycloserine, vancomycin); (2) those that act directly on the cell membrane, affecting its barrier function and leading to leakage of intracellular components (polymyxin); (3) those that interfere with protein biosynthesis (chloramphenicol, tetracyclines, erythromycin, spectinomycin, streptomycin, gentamycin); (4) those that affect nucleic acid biosynthesis (rifampicin, novobiocin, quinolones); and (5) those that block specific steps in intermediary metabolism (sulfonamides, trimethoprim). Some antibiotics, such as sulfonamides, penicillins, cephalosporins, and quinolones, are specific inhibitors of bacterial enzymes. *See* ENZYME; SULFONAMIDE.

Antibiotics active against fungi are fungistatic (ketoconazole, fluconazole) or fungicidal (amphotericin, nystatin). Their mechanisms of action include (1) interaction with the cell membrane, leading to leakage of cytoplasmic components (amphotericin, nystatin); (2) interference with the synthesis of membrane components (ketoconazole, fluconazole); (3) interference with nucleic acid synthesis (5-fluorocytosine); and (4) interference with microtubule assembly (griseofulvin). *See* FUNGISTAT AND FUNGICIDE.

Pharmacological properties. For an antibiotic to be effective, it must first reach the target site of action on or in the microbial cell. It must also reach the body site at which the infective microorganism resides in sufficient concentration, and remain there long enough to exert its effect. The concentration in the body must remain below that which is toxic to the human cells. The effectiveness of an antibiotic also depends on the severity of the infection and the immune system of the body, being significantly reduced when the immune system is impaired. Complete killing or lysis of the microorganism may be required to achieve a successful outcome. *See* IMMUNITY.

Antibiotics may be given by injection, orally, or topically. When given orally, they must be absorbed

Spectrum of activity and clinical utility of antibiotics and other antimicrobial agents

Infecting microorganism:			*Staphylococcus aureus*	*Staphylococcus epidermidis*	*Streptococcus pneumoniae*	*Streptococcus pyogenes*	*Streptococcus viridans*	*Enterococcus faecalis*	*Corynebacterium diphtheriae*	*Clostridium tetani*	*Clostridium difficile*	*Bacillus anthracis*	*Mycobacterium tuberculosis*	*Escherichia coli*	*Klebsiella pneumoniae*	*Salmonella* species	*Shigella* species
Antibiotic			Disease produced*														
Generic name	Trade name	Produced by	Bacteremia Endocarditis Osteomyelitis Pneumonia Skin infections Toxic shock syndrome	Bacteremia Endocarditis Urinary tract infections	Bronchitis Meningitis Otitis media Pneumonia	Pharyngeal infections Rheumatic fever Scarlet fever Septic sore throat Tonsillitis	Bacteremia Endocarditis Wound infections	Bacteremia Endocarditis Urinary tract infections	Diphtheria	Tetanus	Antibiotic-associated colitis	Anthrax	Tuberculosis	Abdominal infections Pneumonia Traveler's diarrhea Urinary tract infections	Pneumonia Rhinitis	Bacteremia Gastroenteritis Typhoid fever	Dysentery
Acyclovir	Zovirax	Chemical synthesis															
Amantadine	Symmetrel	Chemical synthesis															
Amikacin	Amikin	*Streptomyces* sp. and chemical modification	⊕	⊕				⊕					+	⊕	⊕	+	+
p-Aminosalicylic acid	PAS	Chemical synthesis											⊕				
Amoxicillin	Amoxil, Polymox, Trimox, etc	*Penicillium chrysogenum* and chemical modification	+	+	⊕	⊕	⊕	⊕				+		⊕		⊕	
Amphotericin B	Fungizone	*Streptococcus nodosus*															
Ampicillin	Polycillin, Principen, etc.	*Penicillium chrysogenum* and chemical modification	+	+	⊕	⊕	⊕	⊕		+	+	+		⊕	+	⊕	⊕
Azidothymidine	AZT, Retrovir	Chemical synthesis															
Azithromycin	Zithromax	*Streptomyces* sp. and chemical modification	⊕	⊕	⊕	⊕	⊕	+	⊕	+	+	⊕					
Azlocillin	Azlin, Securopen	*Penicillium chrysogenum* and chemical modification	+	+	⊕	⊕	⊕	⊕		+	+	+		⊕	+	⊕	⊕
Aztreonam	Azactam	Chemical synthesis												⊕	⊕	+	+
Bacitracin	Bacitracin	*Bacillus subtilis* *Bacillus licheniformis*	+	+	+	+	+	+	+	+	+	+					
Cefactor	Ceclor	*Cephalosporium* sp. and chemical modification	⊕	⊕	⊕	⊕			+					⊕	⊕	+	+
Cefadroxil	Duricef	*Cephalosporium* sp. and chemical modification															⊕
Cefamandole	Mandol	*Cephalosporium* sp. and chemical modification															⊕
Cefoperazone	Cefobid	*Cephalosporium* sp. and chemical modification	⊕	⊕	⊕	⊕	⊕	⊕	+							+	
Cefotaxime	Claforan	*Cephalosporium* sp. and chemical modification	⊕	⊕	⊕	⊕			+					⊕	⊕	+	+
Cefoxitin	Mefoxin	*Streptomyces* sp. and chemical modification	⊕	⊕	⊕	⊕	⊕		+					⊕	⊕	+	+
Ceftazidime	Fortaz, Tazicef, Tazidime	*Cephalosporium* sp. and chemical modification	⊕	+	⊕	+								⊕	⊕	+	+
Ceftriaxone	Rocephin	*Cephalosporium* sp. and chemical modification	⊕	⊕	⊕	⊕	+		+					⊕	⊕	⊕	⊕
Cephalexin	Keflex	*Cephalosporium* sp. and chemical modification	⊕	⊕	⊕	⊕	+		+					+	+	+	+
Cephalothin	Keflin	*Cephalosporium* sp. and chemical modification	⊕	⊕	⊕	⊕	⊕		+	+				⊕	⊕	⊕	⊕
Chloramphenicol	Chloromycetin	*Streptomyces venezuelae*	+	+	⊕	⊕	+	⊕	+	⊕	⊕	+		+	⊕	⊕	⊕
Ciprofloxacin	Cipro	Chemical synthesis	⊕	⊕	⊕	⊕	+	⊕	+	+	+			⊕	⊕	+	⊕
Clarithromycin	Biaxin	*Streptomyces* sp. and chemical modification	⊕	⊕	⊕	⊕	⊕	+	⊕	+	+	⊕					
Clindamycin	Cleocin	*Streptomyces lincolnensis* and chemical modification	⊕	⊕	⊕	⊕	+		+	⊕	⊕						
Clotrimazole	Lotrimin, Mycosporin, etc.	Chemical synthesis															
Cyclophosphamide	Cytoxan, Neosar	Chemical synthesis															
Cycloserine	Seromycin	*Streptomyces orchidaceus*											⊕				
Dideoxyadenosine		Chemical synthesis															⊕
Dideoxycytidine	Hivid	Chemical synthesis															
Dideoxyinosine		Chemical synthesis															
Doxycycline	Vibramycin	*Streptomyces* sp. and chemical modification	+	+	+	+	+	+		+	⊕	⊕	+	⊕	⊕	⊕	⊕
Econazole	Spectazole	Chemical synthesis															
Erythromycin	Erythrocin, Ilotycin, etc.	*Streptomyces erythreus*	⊕	⊕	⊕	⊕	⊕	+	⊕	+	+	⊕					
Ethambutol	Myambutol	Chemical synthesis											⊕				
Fluconazole		Chemical synthesis															

*+ signifies test-tube activity. ⊕ signifies clinical usefulness.

				⊕									+	⊕	+	+	+	+	+	+	+			+		+												Brucellosis	Disease produced*	*Brucella* species
				+									⊕	⊕	+		⊕	⊕	⊕	⊕	⊕			⊕		⊕	⊕			⊕		⊕		⊕			Urinary tract infections		*Proteus* species
				+									⊕	+			+	⊕		⊕	⊕					⊕								⊕			Bacteremia Pneumonia Urinary tract infections Wound and burn infect.		*Pseudomonas* species
				⊕									+	⊕												+								+			Plague		*Pasteurella* species
		+		⊕								+	⊕	⊕	⊕	+	⊕	⊕	⊕	⊕	⊕			⊕	+	⊕	⊕	+		⊕		⊕		+			Bacteremia Bronchitis Meningitis Otitis media		*Haemophilus influenzae*
		⊕		⊕								⊕																⊕									Whooping cough		*Bordetella pertussis*
		⊕		⊕								⊕	+	⊕	+	+	⊕	+	⊕	⊕	+			+	+	⊕	⊕	⊕		⊕		⊕					Gonorrhea Meningitis Otitis media		*Neisseria* species
				⊕										⊕																							Cholera		*Vibrio* species
		⊕		⊕										+	+	+	⊕		+	⊕	+			+				⊕									Syphilis, yaws		*Treponema* species
																	⊕			⊕																	Lyme disease		*Borellia* species
				+							⊕		+	⊕			+	⊕	⊕	⊕	⊕																Abdominal infections Bacteremia Mixed wound infections		*Bacteroides* species
		⊕		⊕								⊕																⊕									Legionnaire's disease (pneumonia)		*Legionella pneumophila*
		⊕		⊕								⊕		⊕														⊕									Lymphogranuloma venereum Nongonococcal urethritis Pneumonia Psittacosis Trachoma		*Chlamydia* species
		+	⊕	⊕								+		⊕														+									Rickettsial pox Spotted fever Typhus		*Rickettsia* species
+			⊕							⊕																					⊕						Candidiasis (topical systemic)		*Candida* species
+			+							+																					⊕						Aspergillosis		*Aspergillus* species
+			+							+																					⊕						Histoplasmosis		*Histoplasma capsulatum*
⊕										+																					⊕						Meningitis Systemic infections		*Cryptococcus neoformans*
																																					Pneumonia		*Pneumocystis carinii*
																																					Dermatophytosis		*Trichophyton* and *Epidermophyton* spp.
																																					Amebiasis		*Entamoeba histolytica*
																																					Trichomoniasis		*Trichomonas vaginalis*
									⊕																												Mononucleosis Burkitt's lymphoma		Epstein-Barr virus
																																				⊕	Cold sores Encephalitis Genital herpes Keratoconjunctivitis		Herpes simplex virus
																																				⊕	Chicken pox Herpes zoster		Varicella zoster virus
																																					Hepatitis		Hepatitis (A, B, C) virus
																																					Congenital defects Hepatitis Systemic infection		Cytomegalovirus
																																					Genital warts		Papilloma virus
																																			⊕		Influenza		Influenza virus
																																					Bronchitis Pneumonia		Respiratory syncytial virus
					⊕	⊕	⊕														⊕	⊕						⊕									AIDS AIDS-related complex		Human immuno-deficiency virus

Spectrum of activity and clinical utility of antibiotics and other antimicrobial agents (cont.)

Infecting microorganism:			*Staphylococcus aureus*	*Staphylococcus epidermidis*	*Streptococcus pneumoniae*	*Streptococcus pyogenes*	*Streptococcus viridans*	*Enterococcus faecalis*	*Corynebacterium diphtheriae*	*Clostridium tetani*	*Clostridium difficile*	*Bacillus anthracis*	*Mycobacterium tuberculosis*	*Escherichia coli*	*Klebsiella pneumoniae*	*Salmonella* species	*Shigella* species
Antibiotic			Disease produced*														
Generic name	Trade name	Produced by	Bacteremia Endocarditis Osteomyelitis Pneumonia Skin infections Toxic shock syndrome	Bacteremia Endocarditis Urinary tract infections	Bronchitis Meningitis Otitis media Pneumonia	Pharyngeal infections Rheumatic fever Scarlet fever Septic sore throat Tonsillitis	Bacteremia Endocarditis Wound infections	Bacteremia Endocarditis Urinary tract infections	Diphtheria	Tetanus	Antibiotic-associated colitis	Anthrax	Tuberculosis	Abdominal infections Pneumonia Traveler's diarrhea Urinary tract infections	Pneumonia Rhinitis	Bacteremia Gastroenteritis Typhoid fever	Dysentery
5-Fluorocytosine	Ancobon	Chemical synthesis															
Fusidic acid	Fucidine	*Fucidium coccineum*	⊕	⊕	+	+	+										
Gentamycin	Garamycin	*Micromonospora purpurea*	⊕	⊕				⊕					+	⊕	⊕	+	+
Griseofulvin	Fulvicin, Grifulvin	*Penicillium griseofulvum*															
Imipenem		*Streptomyces cattleya* and chemical modification	⊕	⊕	⊕	⊕	⊕	⊕		⊕	⊕			⊕	⊕	⊕	⊕
Isoniazid	Cotinazin, Rimifon, Nydrazid	Chemical synthesis											⊕				
Itraconazole	Sporanox	Chemical synthesis															
Kanamycin	Kantrex	*Streptomyces kanamyceticus*	+	+										⊕	⊕	+	+
Ketoconazole	Nizoral	Chemical synthesis															
Methicillin	Azapen, Staphcillin	*Penicillium chrysogenum* and chemical modification	⊕	⊕	+	+										+	
Metronidazole	Flagyl, Protostat	Chemical synthesis								⊕	⊕						
Mezlocillin	Baypen, Mezlin	*Penicillium chrysogenum* and chemical modification	+	+	⊕	⊕	⊕	⊕		+	+	+		⊕	+	⊕	⊕
Miconazole	Monistat	Chemical synthesis															
Nafcillin	Nafcil, Unipen	*Penicillium chrysogenum* and chemical modification	⊕	⊕	+	+	+										
Netilmicin	Netromycin	*Micromonospora inyoesis* and chemical modification	+	+									+	⊕	⊕	+	+
Nitrofurantoin	Furadantin	Chemical synthesis	+	+				⊕	+					⊕	⊕		
Norfloxacin	Noroxin	Chemical synthesis	⊕	⊕				+						⊕	⊕	+	+
Novobiocin	Albamycin	*Streptomyces niveus* *Streptomyces spheroides*	+	+	+	+		+									
Nystatin	Mycostatin	*Streptomyces noursei*															
Oxacillin	Bactocill, Prostephen	*Penicillium chrysogenum* and chemical modification	⊕	⊕	+	+	+										
Penicillin G		*Penicillium notatum* *Penicillium chrysogenum*	+	+	⊕	⊕	⊕	+	⊕	⊕	⊕	⊕					
Piperacillin	Isipen, Pipracil	*Penicillium chrysogenum* and chemical modification	+	+	⊕	⊕	⊕	⊕		+	+	+		⊕	+	⊕	⊕
Polymyxin B	Aerosporin	*Bacillus polymyxa*												⊕	⊕	+	+
Ribavirin	Virazole	Chemical synthesis															
Rifampicin	Rifadin, Rifamate, Rimactane	*Nocardia mediterranea* and chemical modification	⊕	⊕	+	+	+	+	+	+	+		⊕				
Rimantadine		Chemical synthesis															
Spectinomycin	Trobicin	*Streptomyces spectabilis*															
Streptomycin		*Streptomyces* sp.	+	+			+	+	+			+	⊕	+	+	+	+
Sulfamethoxazole	Gantanol	Chemical synthesis	⊕	⊕	+	+								⊕	⊕	⊕	⊕
Teicoplanin		*Actinoplanes teichomyceticus*	⊕	⊕	+	+		⊕									
Tetracycline	Achromycin, Sumycin, etc.	*Streptomyces* sp.	+	+	+	+	+	+		+	⊕	⊕	+	⊕	⊕	⊕	⊕
Ticarcillin	Aerugipen, Monapen, Ticar	*Penicillium chrysogenum* and chemical modification	+		⊕	⊕	⊕	⊕		+	+	+		⊕	⊕	⊕	⊕
Tobramycin	Nebcin, Tobracin	*Streptomyces tenebrarius*	⊕	⊕				⊕					+			+	+
Trimethoprim	Proloprim, Trimpex	Chemical synthesis	⊕	⊕	⊕	⊕	⊕	+						⊕	⊕		
Vancomycin	Vancocin	*Streptomyces orientalis*	⊕	⊕	⊕	⊕	⊕	⊕		⊕	⊕	⊕					
Vidarabine	Ara A	Chemical synthesis															
Zidovudine	Retrovir, AZT	Chemical synthesis															

*+ signifies test-tube activity. ⊕ signifies clinical usefulness.

-S

						⊕		+	⊕					+						+		+							+			+					Brucellosis	Disease produced*	*Brucella* species
			⊕	⊕	⊕	+		⊕	+					+	⊕					⊕	⊕	⊕			+				+			⊕		⊕			Urinary tract infections		*Proteus* species
				⊕	⊕	+			+					⊕						⊕		⊕							⊕			⊕		⊕			Bacteremia Pneumonia Urinary tract infections Wound and burn infect.		*Pseudomonas* species
			+	+		⊕		+	⊕					+								+							+			⊕		+			Plague		*Pasteurella* species
			⊕	+	⊕	⊕		⊕	+					+	⊕							+			⊕							⊕		+			Bacteremia Bronchitis Meningitis Otitis media		*Haemophilus influenzae*
						⊕																															Whooping cough		*Bordetella pertussis*
			+		⊕	⊕		+	+	⊕		+		+	⊕	⊕				+		+			⊕							⊕					Gonorrhea Meningitis Otitis media		*Neisseria* species
						⊕		+	+					+																		+					Cholera		*Vibrio* species
						⊕			+					+																		+					Syphilis, yaws		*Treponema* species
																⊕																					Lyme disease		*Borellia* species
						+								+												⊕						⊕					Abdominal infections Bacteremia Mixed wound infections		*Bacteroides* species
			⊕			⊕		⊕				⊕																									Legionnaire's disease (pneumonia)		*Legionella pneumophila*
						⊕		⊕				⊕																									Lymphogranuloma venereum Nongonococcal urethritis Pneumonia Psittacosis Trachoma		*Chlamydia* species
						⊕																															Rickettsial pox Spotted fever Typhus		*Rickettsia* species
																		⊕						⊕				⊕		⊕						⊕	Candidiasis (topical systemic)		*Candida* species
																		+						+				+		+						+	Aspergillosis		*Aspergillus* species
																		+						+				⊕		⊕						+	Histoplasmosis		*Histoplasma capsulatum*
																		+						+				+		+						⊕	Meningitis Systemic infections		*Cryptococcus neoformans*
			⊕					⊕																													Pneumonia		*Pneumocystis carinii*
																																		⊕			Dermatophytosis		*Trichophyton* and *Epidermophyton* spp.
																										⊕											Amebiasis		*Entamoeba histolytica*
																										⊕											Trichomoniasis		*Trichomonas vaginalis*
																																					Mononucleosis Burkitt's lymphoma		Epstein-Barr virus
	⊕																																				Cold sores Encephalitis Genital herpes Keratoconjunctivitis		Herpes simplex virus
																																					Chicken pox Herpes zoster		Varicella zoster virus
																																					Hepatitis		Hepatitis (A, B, C virus)
																																					Congenital defects Hepatitis Systemic infection		Cytomegalovirus
																																					Genital warts		Papilloma virus
											⊕																										Influenza		Influenza virus
													⊕																								Bronchitis Pneumonia		Respiratory syncytial virus
⊕							⊕	⊕												⊕	⊕						⊕										AIDS AIDS-related complex		Human immuno-deficiency virus

into the body and transported by the blood and extracellular fluids to the site of the infecting organisms. When they are administered topically, such absorption is rarely possible, and the antibiotics then exert their effect only against those organisms present at the site of application.

Microbial resistance. The therapeutic value of every antibiotic class is gradually eroded by the microbial resistance that invariably follows broad clinical use. For example, resistance to penicillins was recognized soon after the introduction of this class, and use eventually shifted to cephalosporins. Early resistance to macrolides (erythromycin) and chloramphenicol severely limited their use. The emergence of antibiotic resistance is simply natural selection in action. The driving force is antibiotic pressure. In turn, antibiotic resistance is the driving force for the discovery of new antibiotics.

Some bacteria are naturally resistant to certain antibiotics (inherent resistance). For example, enterobacteria are naturally resistant to erythromycin and fusidic acid by virtue of their outer membrane. Clinical resistance is commonly due to the emergence of resistant organisms following antibiotic treatment (acquired resistance). This emergence, in turn, is due to selection of resistant mutants of the infective species (endogenous resistance) or, usually, to transfer of resistance genes from other, naturally resistant species (exogenous resistance). A major challenge in antimicrobial chemotherapy is the horizontal spread of resistance genes and resistant strains, mostly in the hospital but also in the community. The consequences are increased patient morbidity and mortality, reduced drug options, and more expensive and toxic antibiotics.

Detection. Antibiotic resistance is initially detected by susceptibility testing, which provides the resistance phenotype of a given pathogen and has practical implications for patient treament. Rapid detection of resistance and pathogen identification are critical for the rational use of antibiotics and implementation of infection control measures. In the absence of such information, treament is empiric, usually involving broad-spectrum agents, which exacerbates resistance development. Inadequate infection control measures encourage dissemination of resistant strains. The molecular features of resistance are assessed by combinations of biochemical techniques (such as function assays) and molecular biology techniques such as DNA restriction analysis, DNA probes, and polymerase chain reaction (PCR).

Genetics. The genetics of antibiotic resistance are best understood in bacteria. Resistance can result from gene mutations, transfer of R-plasmids between strains or species, and movement of genetic elements (transposons, integrons, gene cassettes) between plasmids or chromosomes in the same organism or to a new organism. For example, R-plasmids were responsible for the spread of sulfonamide resistance to *Shigella* in the 1950s, and penicillin resistance from enterobacteria to *Neisseria* and *Haemophilus* in the 1970s. Plasmids are extrachromosomal circular pieces of DNA that replicate independently of chromosomal DNA and may be present in many copies in each cell. Transposons are smaller pieces of DNA that occasionally move from one site on the chromosome to another site on the chromosome or plasmid. They can replicate only as part of a chromosome or plasmid. Integrons may be part of a plasmid or transposon, while gene cassettes may be part of an integron or may exist free in a circularized (but nonreplicating) form. Gene cassettes and integrons can contain one or more resistance genes. Some resistance genes may have originated from antibiotic-producing organisms or other fortuitously resistant species, or may have evolved from genes coding for normal, mechanistically related cell functions. *See* BACTERIAL GENETICS; PLASMID; TRANSPOSONS.

Mechanisms. The development of resistance varies with the microorganism and the antibiotic. The mechanisms of resistance may be classified as follows: (1) decreased uptake, as with the aminoglycosides, or increased efflux, as with tetracyclines; (2) increased destruction of the drug, as with penicillins, aminoglycosides, and chloramphenicol; (3) altered target site, as with sulfonamides, trimethoprim, streptomycin, erythromycin, and rifampicin; and (4) increased concentration of a metabolite antagonizing the drug action, as with sulfonamides. *See* DRUG RESISTANCE.

Impact on disease. The effects of antibiotics on disease in humans, animals, and plants are discussed below.

Humans. Since not all infectious diseases are reportable to health officials within a community, it is difficult to determine the true incidence of many such illnesses. It is estimated, however, that the average duration of many infectious diseases and the severity of certain others have decreased significantly since the introduction of antibiotic therapy. The dramatic drop in mortality rates for such dreaded diseases as meningitis, tuberculosis, and septicemia offers striking evidence of the effectiveness of these agents. Bacterial pneumonia, bacterial endocarditis, typhoid fever, and certain sexually transmitted diseases are also amenable to treatment with antibiotics. So are infections that often follow viral or neoplastic diseases, even though the original illness may not respond to antibiotic therapy.

Credit for better health, longer life, and diminished mortality rates must be given to a variety of factors. Improved sanitation and housing, immunization programs, and better nutrition have complemented advances in antibiotic therapy to achieve substantial control of infectious diseases. Thus, the vast majority of deaths in developed countries today result from degenerative diseases rather than infections. However, worldwide, infectious diseases remain the leading cause of death. *See* EPIDEMIOLOGY.

Animals. Antibiotics in small amounts are widely used as feed supplements to stimulate growth of livestock and poultry. They probably act by inhibiting organisms responsible for low-grade infections and by reducing intestinal epithelial inflammation. Many experts believe that this use of antibiotics contributes to the emergence of antibiotic-resistant bacteria that

could eventually pose a public health problem. They advocate that those antibiotics used in human therapy should not be used for animal growth promotion. Such concerns may eventually affect the practice of promoting animal growth with antibiotics. *See* ANIMAL FEEDS.

In cattle, sheep, and swine, antibiotics are effective against such economically important diseases as bacterial diarrhea, pneumonia, leptospirosis, foot rot, mastitis, and infections of the reproductive and urinary tracts. The use of antibiotics in dogs and cats closely resembles their use in human medical practice. Administration of the antibiotic may be by oral, parenteral, or topical means.

In fish farms, antibiotics are effective against such economically important diseases as bacterial kidney disease in salmon. They are usually added to the food or applied to the fish by bathing. The incidence of infections in fish, and animals in general, may be reduced by the use of disease-resistant stock, better hygiene, and better diet. *See* AQUACULTURE.

Plants. Although effective against many microorganisms causing disease in plants, antibiotics are not widely used to control crop and plant diseases. Some of the limiting factors are instability of the antibiotic under field conditions, the possibility of harmful residues, and expense. Nevertheless, antibiotic control of some crop pathogens is being practiced, as is true of the rice blast in Japan, for example. *See* PLANT PATHOLOGY. Nafsika H. Georgorapadakou

Chemistry of major groups. Most antibiotics are produced by fermentation of microorganisms or semisynthetic approaches that involve chemical derivatization of the naturally occurring antibiotics. Some antibiotics have been isolated from terrestrial plants and marine organisms. Over 13,000 fermentation-derived antibiotics have been reported. Many thousands of synthetic derivatives have been prepared from these compounds in efforts to improve their biological properties. However, relatively few of the natural or semisynthetic compounds have reached commercial status since most of these compounds are not active in an animal model, are too toxic, or have very limited efficacy.

Microorganisms that produce antibiotics include actinomyces, algae, bacteria, and fungi. During the fermentation of pure cultures of these organisms, the antibiotics are frequently produced as complex mixtures of many closely related components. It is not uncommon to have complexes of 10 or more components; however, there are usually one or two major components, with the others present at very low levels. At one time the very minor components were usually not detected, but modern analytical instrumentation such as high-performance liquid chromatography and liquid chromatography/mass spectrometry has greatly facilitated the identification of these compounds. On a number of occasions, more than one apparently unrelated antibiotic complex is produced by the same microorganism.

Generally, the chemical structures of these compounds are complex; however, some are very simple with only a few carbon atoms. Total synthesis of even the relatively simple antibiotics can be challenging because of the multiplicity of sterogenic centers and functionality that these compounds frequently possess. The smallest clinically used antibiotic is fosfomycin (**1**), which contains only three carbon

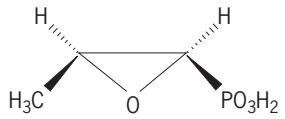

(**1**)

atoms. This compound is produced by fermentation of strains of *Streptomyces*. Most of the fermentation-derived antibiotics have molecular weights in the range of 300–800.

The major antibacterial antibiotics in terms of structure class are (in order of commercial importance) cephalosporins, penicillins, quinolones (synthetic), macrolides, aminoglycosides, and tetracyclines. Compared to antifungal, antiviral, antituberculin, and anticancer antibiotics, the antibacterial antibiotics are the most widely used. All of these antibiotics are fermentation products or are derived from fermentation products, except the quinolones, which are synthetic.

The first clinically used compound in the quinolone family was nalidixic acid, which is orally effective and has activity against gram-negative bacteria. Chemical modifications of this basic structure have led to compounds such as norfloxacin (**2**),

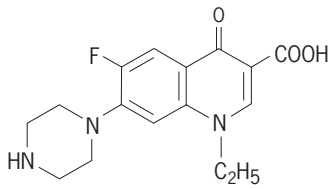

(**2**)

which has an improved antibacterial spectrum and significantly less susceptibility to resistance. By the original definition, these compounds are not antibiotics since they have not been produced by microorganisms, but current usage includes them.

There have been many systems to classify the fermentation-derived antibiotics based on structure, mechanism of action, biosynthesis, producing organism, or other features. Although chemical classification is usually the most informative, it is complicated by the complex nature and diversity of chemical structures of these compounds. The following classification of some of the major groups is based on chemical structure, is generally consistent with the more widely used chemical schemes, and covers most of the commercially important antibiotics.

Aminoglycosides. This group of antibiotics is characterized by compounds containing amino sugars and deoxystreptamine or streptamine. Representative members are gentamicin [structure (**3**) is gentamycin

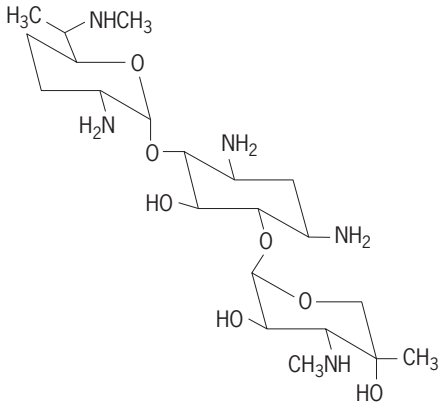

(3)

C_1] and neomycin; they are used primarily for gram-negative bacterial infections treated by injection and for topical applications, respectively. These antibiotics destroy bacteria by interfering with their protein biosynthesis. All of the members of this family of antibiotics are extremely water-soluble, polybasic compounds, reflecting the highly polar units that make up these molecules. *See* AMINO SUGAR; POLAR MOLECULE.

Ansamycins. Antibiotics in this family are large-ring structures that contain benzenoid or naphthalenoid aromatic units in which nonadjacent positions are bridged by an aliphatic chain to form the macrocyclic ring. An amide bond is always found at one of the aliphatic-aromatic junctions. Rifampin (**4**), a

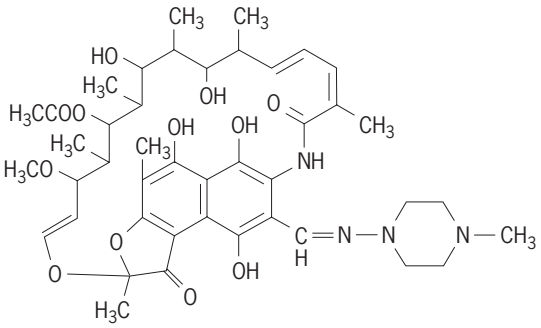

(4)

semisynthetically derived compound, is a member of this family. It kills bacteria by inhibiting their ribonucleic acid (RNA) polymerase. Clinically, it is used against infections by gram-positive bacteria and mycobacteria. *See* MACROCYCLIC COMPOUND.

Anthracyclines. Antibiotics in this family are characterized by having a tetrahydrotetracenequinone chromophore linked to one or more sugars by glycosidic bonds. Doxorubicin (**5**), also known as adriamycin, is a member of this family and is one of the most important clinical antitumor agents. Anthracyclines are antibiotics that appear to destroy bacteria by intercalating with their deoxyribonucleic acid (DNA). Although most anthracyclines are very active against gram-positive bacteria, they are quite toxic to mammals and are not used as antibacterial

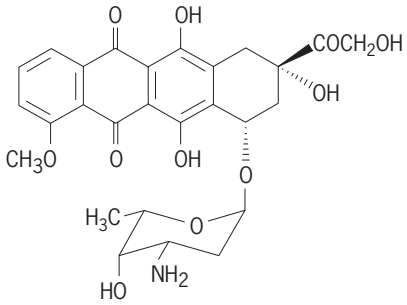

(5)

antibiotics. The clinically used antibacterial antibiotics have much greater safety margins. Doxorubicin, however, has selective activity against certain types of tumors and, with carefully controlled dosing to minimize toxic effects against normal cells, this antibiotic has been very effective for cancer chemotherapy.

Chloramphenicol. Many naturally occurring antibiotics belong to families that are relatively large, with 20 or more closely related structures. In contrast, there are only a few compounds that are closely related to chloramphenicol (**6**). The structure of chlo-

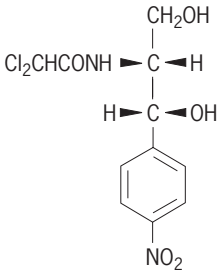

(6)

ramphenicol is relatively simple: it is a nitrobenzene derivative of dichloroacetic acid. This antibiotic has a broad spectrum of antibacterial activity and kills bacteria by the inhibition of protein biosynthesis. Although chloramphenicol is very effective clinically, its use has been limited because of the potential toxic side effects. *See* NITROAROMATIC COMPOUND.

Enediynes. This is a unique family of antitumor antibiotics in which all members contain a cyclic array of two acetylenic units and a double bond. Calicheamicin and esperamicin are representative members of this family. The aglycone portion [structure (**7**), where R1 and R2 represent sugars and aro-

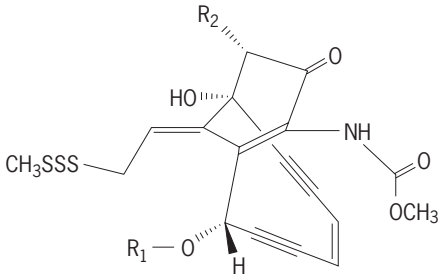

(7)

matic groups] of these antibiotics has been called a

warhead, since it can be triggered to generate diradicals which destructively interact with cellular DNA to cause single- and double-strand cleavage and subsequent cell death. This damaging effect to DNA is probably responsible for the antibacterial and antitumor activities of these compounds. They are the most potent antitumor agents that have been discovered and are approximately 1000 times more potent than adriamycin (**5**).

Glycopeptides. Antibiotics within this family are relatively large molecules with molecular weights frequently in the range of 1400–2000. They all contain peptide units, with a number of aromatic amino acids that have apparently formed cyclic structures by phenolic oxidative coupling. Attached to this aglycone by glycosidic linkages are usually two or more sugar moieties. Vancomycin and avoparcin are glycopeptide antibiotics that are used commercially for medical and agricultural applications. These compounds inhibit bacterial cell-wall biosynthesis by interaction of the antibiotic aglycones with bacterial cell-wall peptides terminating in acyl-D-ala-D-ala. The term acyl-D-ala-D-ala represents the peptides in the bacterial cell wall that end in acyl-D-alanyl-D-alanine. The acyl term indicates linkage to the peptide through the carboxyl group of the bacterial peptide. Peptides ending in this D-ala-D-ala sequence bind to glycopeptide antibiotics. Vancomycin is effective clinically against gram-positive bacteria, especially staphylococci that are resistant to other antibiotics.

β-Lactams. For chemical classification the β-lactam antibiotics are within the general class of peptide antibiotics. All antibiotics in this family contain a β-lactam ring, and within the family are several substructure groups. The more important groups include the penicillins, cephalosporins, carbapenems, and monobactams as represented by ampicillin (**8**), cephalothin (**9**), imipenem (**10**), and aztreonam (**11**), respectively. Ampicillin and cephalothin

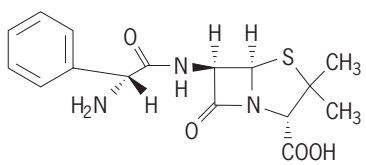

(**8**)

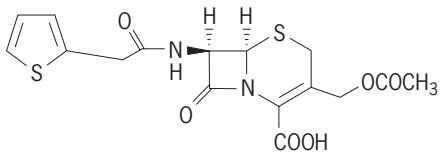

(**9**)

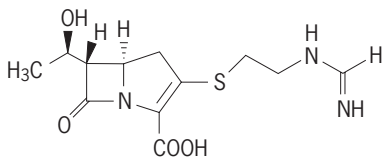

(**10**)

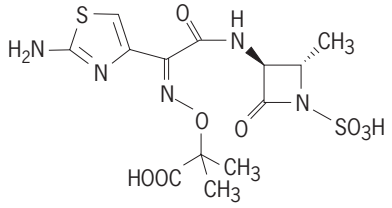

(**11**)

are produced semisynthetically, whereas imipenem and aztreonam are derived from total syntheses. Although the latter two compounds are synthetic, the key compounds leading to their discovery were fermentation-derived antibiotics. The β-lactam family is commercially the most important class of antibiotics used for bacterial infections. The β-lactams are relatively nontoxic, since they selectively inhibit bacterial cell-wall biosynthesis. The development of resistance by bacteria, frequently from acquisition of new β-lactamases that inactivate the antibiotics, has spurred the synthesis of thousands of new β-lactam derivatives in efforts to find compounds with improved biological properties. Another important group of β-lactams comprises the β-lactamase inhibitors such as clavulanic acid and tazobactam. These compounds have poor antibiotic activity but are excellent inhibitors of the enzymes that inactivate commercially important β-lactam antibiotics. Consequently, they are used in combination with some of the commercially important β-lactam antibiotics to improve the spectrum of activity. *See* DRUG RESISTANCE; LACTAM.

Lincomycin/celesticetin. Antibiotics within this small family are carbohydrate-type compounds containing a thiosugar (a sugar that has one or more of its oxygen atoms replaced with sulfur) linked by an amide bond to a substituted proline unit. Clindamycin (**12**)

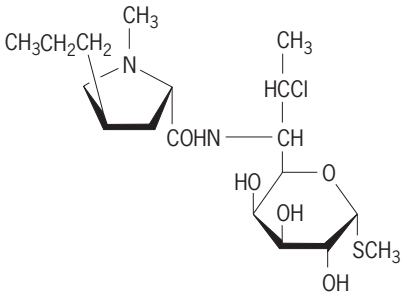

(**12**)

is clinically one of the most useful compounds in this family and is derived by chemical modification of lincomycin. Antibiotics in this family are effective against gram-positive aerobic and anaerobic bacteria. They destroy bacteria by inhibiting their protein biosynthesis.

Macrolides. Macrolides constitute a broad category of antibiotics that are chemically defined as macrocyclic lactones. Within this broad category are a number of subfamilies. The two more prominent subfamilies are the antibacterial antibiotics related to erythromycin and the antifungal polyene antibiotics related to amphotericin B. The latter group are

known as polyenes. *See* LACTONE.

The general class of antibacterial macrolides related to erythromycin have macrocyclic lactone rings that are 12, 14, or 16 membered. Erythromycin (**13**)

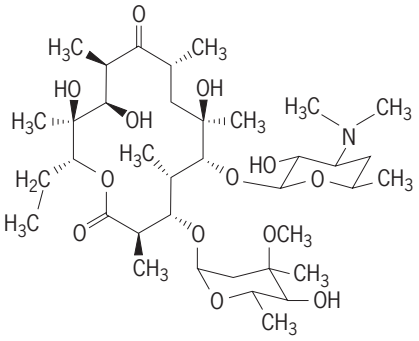

(**13**)

has a 14-membered ring and is usually referred to as a basic macrolide, since an amino sugar is attached to its macrocyclic lactone through a glycosidic linkage and gives basic properties to the overall molecule. Other related macrolides such as neutramycin have neutral sugars. Both the neutral and basic macrolides of this type have activity against gram-positive bacteria and inhibit bacterial protein synthesis. Basic macrolides are important commercial antibiotics used for medical (for example, erythromycin) and agricultural (for example, tylosin) applications. Erythromycin is the drug of choice for treatment of legionnaire's disease. *See* LEGIONNAIRES' DISEASE.

Nucleosides. This antibiotic family is relatively large, with over 200 known structures. The nucleoside antibiotics are analogs of the essential purine and pyrimidine nucleosides that are subunits of RNA and DNA. The differences between the essential nucleosides and the antibiotics are usually in the chirality

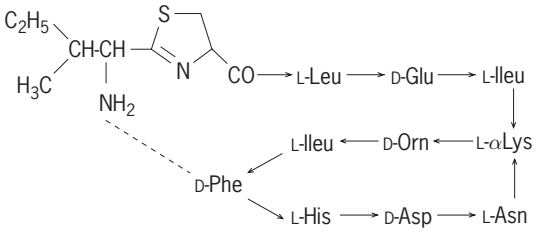

(**14**)

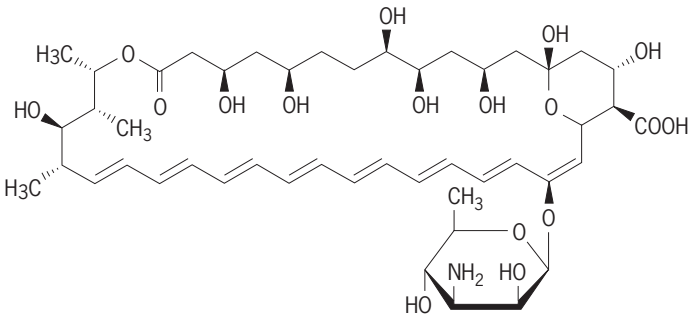

(**15**)

and type of sugar, substitution on the sugar, the substitution pattern of the base and its oxidation state, and the type of bonding between the sugar and base units. The similarity in chemical structure to numerous critical metabolic products has resulted in nucleoside antibiotics having a number of modes of action as shown by the following examples: (1) puromycin inhibits bacterial protein synthesis by acting as an analog of aminoacyl-tRNA, which takes part in ribosomal peptide bond formation; and (2) the polyoxins interfere with fungal cell-wall assembly by inhibition of chitin synthetase. Many nucleoside antibiotics show antitumor or antiviral activity. Ara-A is used topically for treatment of herpes viral infections. *See* NUCLEOTIDE.

Peptides. Structurally the peptide antibiotics represent a very large and extremely diverse class. Examples of the families within this classification are the linear oligopeptides, cyclic oligopeptides, diketopiperazines, depsipeptides (subunits linked by amide and ester bonds), and β-lactams. The peptide antibiotics contain unusual amino acids residues such as β- or γ-amino acids, amino acids with the D-configuration, or amino acids totally unrelated to the common amino acids. Peptide antibiotics have been used for medical and agricultural applications. Bacitracin [structure (**14**) is bacitracin A] is used as a topical antibiotic in human therapy and as a growth promoter in animal feeds. Like many of the peptide antibiotics, it has a cyclic structure, contains several common amino acids in the D-configuration, and has an unusual amino acid containing a dihydrothiazole unit. *See* AMINO ACIDS.

Polyenes. The antifungal polyenes are macrolides with 20- to 44-membered rings containing three to seven conjugated double bonds. They are classified as trienes, tetraenes, pentaenes, hexaenes, or heptaenes, depending on the length of the polyene chromophore. Many of the polyenes contain a glycosidically linked amino sugar, mycosamine or perosamine. Some of the heptaenes have an aromatic moiety, *p*-aminoactophenone or its *N*-methyl analog, on a side chain to the carbon involved in lactone formation. *Streptomyces* are the organisms usually found to produce these compounds. The polyenes usually have broad-spectrum antifungal activity, which results from interaction of these compounds with membrane sterols to change the fungal membrane permeability. Although the polyenes have excellent antifungal activity, only amphotericin B (**15**) is used parenterally for life-threatening antifungal infections such as candidiasis, aspergillosis, histoplasmosis, and coccidioidomycosis since the other polyenes are too toxic. Amphotericin B also has a small safety margin, but it is used because there are no better antifungal antibiotics for these applications.

Polyethers. Most of the polyether antibiotics are characterized by a linear series of tetrahydrofuran and tetrahydropyran moieties that frequently involve spiroketal systems (structures that possess a fusion of two rings, usually five or six membered). These compounds terminate in a carboxylic acid group.

Some polyether antibiotics have a sugar unit linked through a glycosidic bond to one of the tetrahydrofuran or tetrahydropyran rings. This sugar is almost always 4-*O*-methylamicetose. The large majority of these compounds are produced by *Streptomyces* species and to some extent the rarer actinomyces.

Polyethers selectively bind and transport certain cations through cellular membranes. Because of this unusual property, these compounds are referred to as ionophores. Each antibiotic has its own ion specificity. These antibiotics are active against gram-positive bacteria, mycobacteria, fungi, yeast, and certain protozoa. Their ion transport capabilities also makes them quite toxic for parenteral use in mammals. However, some of these compounds are commercially very useful as anticoccidial agents that are administered orally in feed to poultry. Monensin and maduramicin are very effective for this application. Maduramicin is one of the few polyether antibiotics containing a sugar different from 4-*O*-methylamicetose. *See* ION TRANSPORT; IONOPHORE.

Tetracyclines. The naturally occurring tetracyclines are a relatively small family of about 25 antibiotics that are produced mainly by *Streptomyces* species. These antibiotics contain a polyhydronaphthacene nucleus. None of the tetracyclines have a quinone functionality, which is common to numerous other tetracyclic antibiotics such as doxorubicin (**5**). Instead, the tetracyclines are characterized by having an array of hydroxyl and keto groups on one side of the molecule and a relatively fixed substitution pattern on the A-ring, which contains a carboxamide group ($CONH_2$) in the 2-position, as represented by tetracycline (**16**). The mechanism of action for most

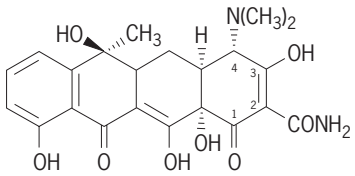

(**16**)

tetracyclinelike antibiotics is inhibition of bacterial protein synthesis. Although this family of naturally occurring antibiotics is relatively small, a number of the compounds have become commercial products for both medical and agricultural applications. The tetracyclines have a broad antibacterial spectrum, with activity against both gram-positive and gram-negative bacteria; and they are effective against *Rickettsia, Mycoplasma*, and *Spirochaetes*. They are orally active and exhibit relatively low toxicity. Chemical modifications have resulted in compounds such as minocycline with improved biological properties. *See* ANTIMICROBIAL AGENTS; CHEMOTHERAPY.

Donald B. Borders

Bibliography. B. Austin, *Marine Microbiology*, 1988; L. E. Bryan, *Microbial Resistance to Drugs*, 1988; B. W. Bycroft (ed.), *Dictionary of Antibiotics and Related Substances*, 1988; J. E. Conte and S. L. Barriere, *Manual of Antibiotics and Infectious Diseases*, 1988; W. Crueger and A. Crueger, *Biotechnology: A Textbook of Industrial Microbiology*, 2d ed., 1989; E. F. Gale et al., *The Molecular Basis of Antibiotic Action*, 2d ed., 1981; A. Kucers, N. Mck. Bennit, and R. J. Kemp, *The Use of Antibiotics*, 4th ed., 1987; B. Pratt and R. Fekety, *The Antimicrobial Drugs*, 1986; R. Reiner, *Antibiotics: An Introduction*, 1982; H. Umezawa (ed.), *Index of Antibiotics from Actinomycetes*, vol. 1, 1967, vol. 2, 1978; M. J. Weinstein and G. H. Wagman (eds.), *Antibiotics: Isolation, Separation and Purification*, 1978; M. Woodbine, *Antibiotics and Antibiosis in Agriculture*, 1977.

Antibody

A member of the immunoglobulin class of proteins that is produced and secreted by the B cells (or B lymphocytes) of the immune system in response to antigens (foreign substances). Antibodies are found in the tissue fluids and mucous membranes and are essential for protection against and recovery from infection. *See* ANTIGEN; IMMUNITY; IMMUNOGLOBULIN.

Historical perspective. In the 1890s, E. von Behring and S. Kitasato showed for the first time that a toxin-neutralizing activity appeared in the blood plasma of patients or animals that had survived tetanus or diphtheria. Furthermore, plasma from these animals could be transferred to other animals, protecting them. Although the chemical nature of the protective agent, antibody, was unknown, use of antitoxic plasma or serum to treat these and other infections soon became widespread. (Plasma is the liquid, acellular part of blood; serum is the liquid that remains after blood has clotted.) In fact, this was the only effective treatment for bacterial infections until the advent of antibiotics.

In a variety of tests, antibodies were shown to agglutinate bacteria or, in fresh serum, to lyse them; to precipitate in combination with antigen; to cause the skin wheals typical of allergy; and to make bacteria more readily ingested (opsonized) by phagocytic cells. Fierce debates raged about whether each of these activities was the property of different classes of antibody, or simply reflected the nature of the test. Once the structure of antibodies was unraveled, it became clear that several different classes existed.

The earliest experiments showed that antibodies were specific for the antigens that elicited them. They were heterogeneous (varying in charge and size) and migrated at different rates in an electrophoretic field, although generally in the gamma globulin region. Some antibodies were much larger than others, so they were named macroglobulin. When structural studies showed that all antibodies were closely related, the term "immunoglobulin" was adopted, and gamma globulin became immunoglobulin G (IgG), macroglobulin became IgM, and so on. *See* ELECTROPHORESIS; GLOBULIN.

Structure. The five classes of antibodies (IgM, IgD, IgG, IgA, and IgE) are structurally similar, consisting of a basic unit of two heavy (H) and two light (L) polypeptide chains. The H chains are linked to

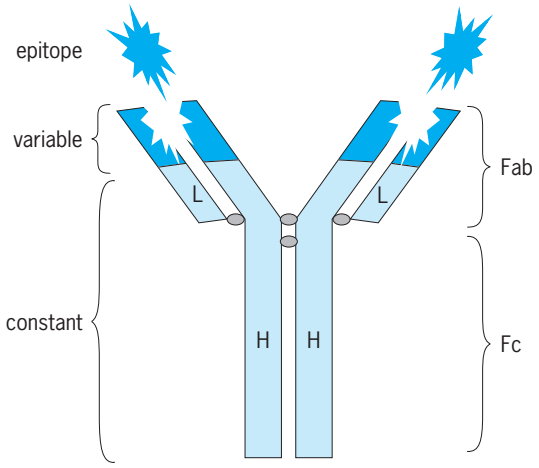

Fig. 1. Typical antibody molecule. There are two H and two L chains, held together by disulfide bonds, and the N-terminal variable domains are shown in dark color. The molecule is bilaterally symmetrical so two identical epitopes are bound. The antigen-binding regions (Fab) and the constant region (Fc) are indicated.

each other by one or more disulfide bonds, and each L chain is similarly linked to an H chain. In any one antibody molecule, the H and L chains are the same, so the molecule is said to have axial symmetry (**Fig. 1**).

X-ray crystallography and molecular sequence analyses show that each chain is composed of immunoglobulin folds, that is, two or more compact domains containing beta-pleated sheets that are held together by disulfide bonds. They probably all derive from an ancient ancestral domain and have a molecular weight of about 10,000 daltons. L chains have two domains, and H chains have four or five. When many antibodies were sequenced, it was found that the N- (or amino-) terminal domains of both H and L chains vary from antibody to antibody, while the other domains are more or less constant within a single class (**Fig. 2**). The variable domains (V_L and V_H) are where antigenic determinants (or epitopes) bind, based on complementarity of surface charge and hydrophobicity. (In this regard, antigen-antibody binding is similar to that between enzymes and substrates.) Between the first and second constant domains of H chains lies a flexible "hinge" peptide region (Fig. 2) that is readily cleaved by enzymes to produce two antigen-binding branches, called Fab, and a common fragment, known as Fc. The Fab regions confer specificity, whereas the Fc is responsible for the function of the particular class of antibody.

Because antibodies are bilaterally symmetrical, each can bind two epitopes (that is, they are divalent). Natural antigens usually have many epitopes, so they can be linked together by antibodies to form a large antigen-antibody complex, or immune complex; this is how, for example, bacteria are agglutinated by antibodies. Complexes between soluble antigens and antibodies become insoluble, which explains the phenomenon of precipitation. Immune complexes that form in the blood may get trapped in the basement membranes that underlie capillaries, resulting in harmful inflammation.

Antibody classes. Each of the five classes, or isotypes, of immunoglobulins probably arose over evolutionary time by gene duplication. This is often the case when a gene product confers a strong survival advantage. (With two copies, one copy of the gene can mutate without risk to the survival of the individual, and so evolve into a new useful gene; then there is even more pressure to duplicate, as every new variant may confer further survival advantages.) Each of the classes has unique structural and biological properties, while retaining overall similarity and function in immune responses. The specific properties of each class depend on the structure of the constant domains of its H chain; they all share the same pool of L chains. The very same L chains, and V_H domains, can be found in each of the five classes. In fact, a B cell that is making, for example, IgM against a certain epitope may be found a few days later to be making IgG or IgE of exactly the same specificity.

IgD. IgD consists of two L chains and two delta (δ) H chains. Although IgD is present in the blood, its major role is in the activation of B lymphocytes. Before exposure to antigen, a mature B cell expresses surface IgD (sIgD) and surface IgM (sIgM) of the same specificity. Antigen binding by both sIgD and sIgM activates the B cell.

IgM. IgM is the largest of the antibody molecules. Secreted IgM is composed of five basic units [composed of two L chains and two mu (μ) H chains] connected by disulfide binds to a J (joining) chain. Although it can bind 10 epitopes, in practice it binds only one or two. IgM is present in the blood (at a concentration of 100 mg/dL), but its large size hinders its movement from the blood to the tissue spaces. In addition, IgM is not capable of opsonization, as phagocytic cells do not have receptors for its Fc end.

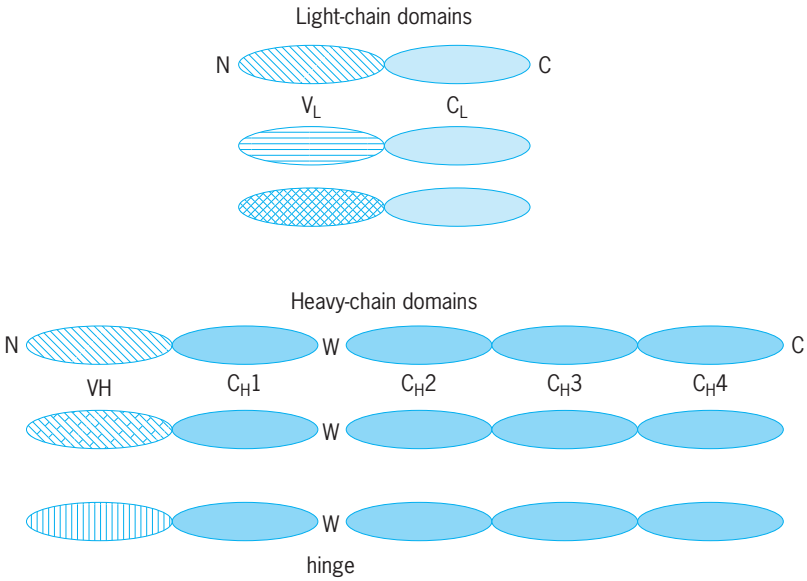

Fig. 2. Domain structure of three typical light and heavy chains shows their similarities and differences. The amino- and carboxy-terminals of the chains are marked N and C, respectively. The flexible hinge region is also shown. At the N terminals the first domain has a variable sequence that differs from antibody to antibody; it is to these domains that antigens bind. The C terminal domains are constant within any one class of immunoglobulin.

Its main function in immune responses is to activate complement, a group of interacting proteins that are important mediators of inflammation and can attract and assist white blood cells in capturing bacteria and directly lyse bacterial cells. *See* COMPLEMENT.

IgG. IgG is a typical 2L–2H basic unit, with gamma (γ) H chains. It is the most abundant immunoglobulin class in blood (1000 mg/dL). It activates complement, but it is much less efficient than IgM. Since its Fc end fits receptors on most phagocytic white blood cells, it is capable of opsonization. IgG is the only isotype that can cross the placenta from human mother to fetus, providing vital antenatal and neonatal protection. However, if a mother is negative for the red blood cell antigen Rh(D) and her fetus is positive for the antigen, red cells from the fetus can cross the placenta and immunize the mother. The mother's IgG antibodies then cross the placenta in the other direction and destroy the fetal red cells, causing a condition known as hemolytic disease of the newborn. It is prevented by administering IgG against Rh(D) to the mother at 28 weeks' gestation and again at delivery. *See* RH INCOMPATIBILITY.

IgA. IgA is present in secretory fluids such as saliva, bronchial fluid, and milk. It is found in high concentrations in the colostrum (the first milk produced by the mother after giving birth) and continues to protect a baby's oral, nasal, and pulmonary mucosa as long as it nurses. IgA is produced by B cells located under the mucosa of secretory epithelia and released as a dimer of two basic units [of two L chains and two alpha (α) H chains] joined at their Fc ends by a J chain. The antibody is bound by receptors on the basal aspect of the mucosal cells and transported to the luminal side, where it is released into the secretions along with the receptor, known as secretory component. The secretory component seems to protect IgA from digestion by enzymes secreted by humans and many microorganisms.

IgE. IgE, a basic unit molecule with two L and two epsilon (ϵ) H chains. IgE antibodies are present in blood in very small amounts (micrograms per deciliter), but they are potent and can bind strongly to the membrane of tissue mast cells and blood basophils (the major effector components in an immediate allergic reaction).When antigen cross-links adjacent molecules of IgE, it causes mast cells and basophils to release their contents of many inflammatory mediators, mainly histamine. Thus, IgE is the antibody responsible for allergies and asthma. The mast cells triggered by IgE attract specialized phagocytes called eosinophils, which contain defense molecules that allow them, uniquely, to kill parasites. *See* ALLERGY.

Antibody production. Antibodies are produced by B lymphocytes and are the product of gene recombination events that lead to a unique V_L and V_H domain in each cell. Initially, this gene cassette (the set of gene segments encoding the H and L variable regions) is further recombined with the genes encoding the constant domains of the H chains of IgM and IgD. A primary ribonucleic acid (RNA) transcript is then made, and it is processed to messenger RNA (mRNA) for the mature H chains of IgM and IgD. These then combine with L chains as they are synthesized. Thus, the mRNAs that are translated into the H and L polypeptide chains are produced by various recombinations of gene segments that encode the variable and constant domains.

The resulting immunoglobulin is inserted into the plasma membrane, where it serves the B lymphocyte as an antigen receptor. When the cognate antigen binds to its corresponding sIgD and sIgM on the B cell and other environmental conditions are met, the B cell matures into an antibody-secreting plasma cell, which can release a million antibodies per minute. These secreted antibodies are identical to the ones that served as B-cell receptors.

IgM and IgD are always synthesized first and inserted into the B-cell membrane as receptors. If the B cell never encounters antigen, this is all that happens. If it does meet antigen, it will begin to secrete IgM (IgD is only there as a receptor). If so instructed by helper T cells, B cells will switch to making IgG, IgA, or IgE by recombining the V domain of their H chain gene with the corresponding new constant (or C) domain gene. Thus, the class of the secreted antibody changes, but the specificity for antigen does not. *See* CELLULAR IMMUNOLOGY; GENE; IMMUNOLOGICAL ONTOGENY.

Induction. After exposure to a new antigen, for example a vaccine, sensitive tests may detect IgM antibody in the blood in 2 or 3 days. For most antigens, the IgM concentration will peak in a week or two, and then decline as the IgM is replaced by IgG, the levels of which will decline in weeks to months. A second (booster) injection of antigen induces an IgG response which is quicker, rises more steeply, and is sustained longer. *See* VACCINATION.

Polyclonal and monoclonal antibodies. Each B cell (and each antibody) is specific for a single epitope, but most proteins have several epitopes (which can be 10 to 20 amino acids long), and organisms such as bacteria have hundreds. When many different B-cell clones (identical B cells) produce different antibodies against various antigenic determinants, or epitopes, the antibodies and the response are said to be polyclonal. A monoclonal antibody is produced by only one of these B-cell clones; it is not usually seen in nature but is produced by isolation of single B cells and their progeny. Newer methods are available to produce monoclonal antibodies, which are valuable in research, diagnosis, and therapy. *See* ANTIGEN-ANTIBODY REACTION; MONOCLONAL ANTIBODIES.

J. John Cohen

Bibliography. D. R. Davies and H. Metzger, *Structural Basis of Antibody Function*, in *Annual Review of Immunology*, 1:87–115, 1983; P. Holliger and P. J. Hudson, Engineered antibody fragments and the rise of single domains, *Nat. Biotech.*, 23:1126–1136, 2005; C. A. Janeway et al., *Immunobiology*, 6th ed., Garland Publishing, 2004; R. A. Manz et al., Maintenance of serum antibody levels, *Annu. Rev. Immunol.*, 23:367–386, 2005; W. E. Paul (ed.), *Fundamental Immunology*, 5th ed., Lippincott Williams & Wilkins, 2003.

Anticline

A fold in layered rocks in which the strata are inclined down and away from the axes. The simplest anticlines (see **illus.**) are symmetrical, but in more highly deformed regions they may be asymmetrical, overturned, or recumbent. Most anticlines are elongate with axes that plunge toward the extremities of the fold, but some have no distinct trend; the latter are called domes. Generally, the stratigraphically older rocks are found toward the center of curvature of an anticline, but in more complex structures these simple relations need not hold. Under such circumstances, it is sometimes convenient to recognize two types of anticlines. Stratigraphic anticlines are those folds, regardless of their observed forms, that are inferred from stratigraphic information to have been anticlines originally. Structural anticlines are those that have forms of anticlines, regardless of their original form. *See* FOLD AND FOLD SYSTEMS; SYNCLINE.

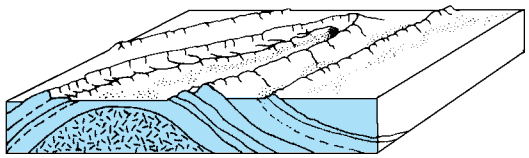

Diagram relating anticlinal structural to topography.

Philip H. Osberg

Antiferromagnetism

A property possessed by some metals, alloys, and salts of transition elements in which the atomic magnetic moments, at sufficiently low temperatures, form an ordered array which alternates or spirals so as to give no net total moment in zero applied magnetic field. **Figure 1** shows the simple antiparallel arrangement of manganese moments at temperatures below 72 K in the unit cell of manganese fluoride (MnF_2). The most direct way of detecting such arrangements is by means of neutron diffraction. *See* NEUTRON DIFFRACTION.

Néel temperature. This is the transition temperature (L. Néel, 1932) below which the spontaneous antiparallel magnetic ordering takes place. A plot of the magnetic susceptibility of a typical antiferromagnetic powder sample versus temperature is shown in **Fig. 2**. Below the Néel point, which is characterized by the sharp kink in the susceptibility, the spontaneous ordering opposes the normal tendency of the magnetic moments to align parallel to the applied field. Above the Néel point, the substance is paramagnetic, and the susceptibility χ obeys the Curie-Weiss law, as in Eq. (1), with a negative paramagnetic Curie-Weiss temperature $-\theta$. The Néel temperature is similar to the Curie temperature in ferromagnetism. *See* CURIE TEMPERATURE; CURIE-WEISS LAW; MAGNETIC SUSCEPTIBILITY.

$$\chi = C/(T + \theta) \tag{1}$$

The cooperative transition that characterizes antiferromagnetism is thought to result from an interaction energy U of the form given in Eq. (2), where

$$U = -2\Sigma J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j \tag{2}$$

$\mathbf{S}_i$ and $\mathbf{S}_j$ are the spin angular momentum vectors associated with the magnetic moments of neighbor atoms i and j, and J_{ij} is an interaction constant which probably arises from the superexchange coupling discussed later, although formally Eq. (2) is identical to the Heisenberg exchange energy. If all J_{ij} are positive, the lowest energy is achieved with all $\mathbf{S}_i$ and $\mathbf{S}_j$ parallel, that is, coupled ferromagnetically. Negative J_{ij} between nearest-neighbor pairs (i, j) may lead to simple antiparallel arrays, as in Fig. 1; if the distant neighbors also have sizable negative J_{ij}, a spiral array may have lowest total energy. *See* FERROMAGNETISM; HELIMAGNETISM.

A simple lattice like MnF_2 (Fig. 1) can be divided into sublattice 1, containing all corner atoms, and sublattice 2, containing all body-centered atoms. Nearest-neighbor interactions connect atoms on different sublattices. On the average, the interaction may be replaced by a single antiparallel coupling between the total magnetizations M_1 and M_2 of the two sublattices. Each sublattice acts as if it

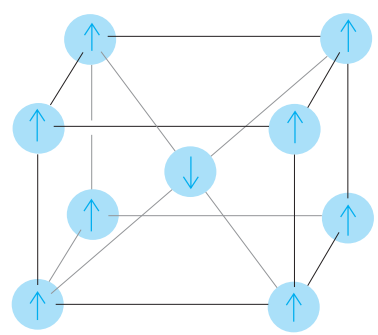

Fig. 1. Antiferromagnetism in manganese fluoride. Only manganese atoms are shown.

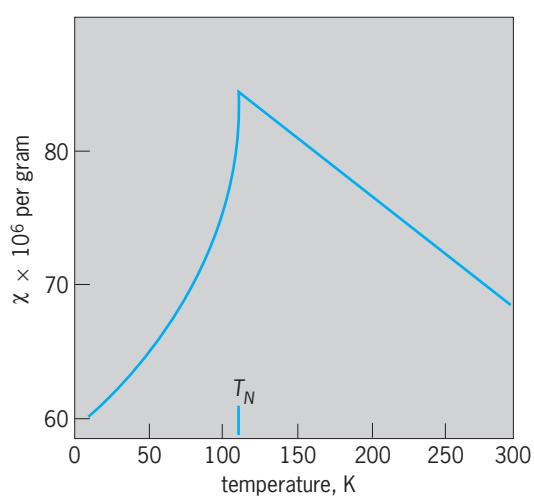

Fig. 2. Magnetic susceptibility of powdered manganese oxide. T_N = Néel temperature. (*After H. Bizette, C. F. Squire, and B. Tsai, 1938*)

were in a large internal magnetic field (Weiss field) proportional to the negative magnetization of the other sublattice. This elementary approach was first given by Néel in 1932 and is analogous to the Weiss molecular field theory of ferromagnetism. More exact treatments of Eq. (2) have been made, but the basic features of antiferromagnetism appear in this simple model.

At high temperatures, the sublattice magnetizations obey the Curie law, as in Eqs. (3) and (4), where

$$M_1 = (C'/T)(H_0 - \lambda M_2) \tag{3}$$

$$M_2 = (C'/T)(H_0 - \lambda M_1) \tag{4}$$

C' is the Curie constant for a sublattice, H_0 is an applied external field, and $-\lambda$ is the proportionality constant of the internal Weiss field. From Eqs. (3) and (4), Eq. (5) derived, is derived, which fits the Curie-Weiss law, Eq. (1), with the values shown in Eqs. (6).

$$\chi = (M_1 + M_2)/H_0$$
$$= 2(C'/T)[1 + \lambda(C'/T)] \tag{5}$$
$$C = 2C' \qquad \theta = C'\lambda \tag{6}$$

The condition that M_1 and M_2 can have finite values in the absence of H_0 (condition of spontaneous sublattice magnetization) is that the determinant of the coefficients of M_1 and M_2 in Eqs. (3) and (4) vanishes, which is satisfied at a temperature given by Eq. (7). This is the Néel temperature. Equations (3)

$$T_N = C'\lambda \tag{7}$$

and (4) hold only for $T \geq T_N$; the Curie law takes a complicated form for $T < T_N$. In this latter region, the sublattice magnetization varies with temperature essentially in the same manner as does the magnetization of ferromagnetism.

The preceding theory predicts $\theta/T_N = 1$; the experimental values (see **table**) range from 0.7 to 5. P. W. Anderson ascribed this disagreement to the oversimplified two-sublattice model. Anderson's multisublattice theory not only accounts for $\theta/T_N > 1$, but also predicts a variety of magnetic ordering arrangements, many of which have been confirmed by neutron diffraction. For example, the arrangement in MnO is shown in **Fig. 3**; the magnetic moments are all parallel in alternating planes.

Some representative antiferromagnets

Substance	Crystal type	Néel temp. (T_N), K	Curie-Weiss temp. (θ), K
MnF_2	Rutile	67	80
MnO	NaCl	122	610
FeO	NaCl	198	507
$KMnF_3$	Perovskite	88	158
$CuCl_2 \cdot 2H_2O$	Orthorhombic	4.3	4.5
CrSb	NiAs	723	550
Cr_2O_3	Al_2O_3	307	485
$ZnFe_2O_4$	Spinel	9	
EuTe	NaCl	7.8	6
MnTe	Hexagonal close-packed	403	690
La_2CuO_4	Orthorhombic	250	
$YBa_2Cu_3O_6$	Orthorhombic	500	

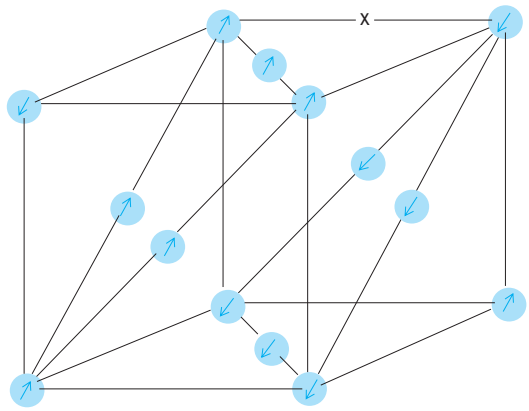

Fig. 3. Antiferromagnetism in manganese oxide. Only manganese atoms are shown.

Superexchange. This is an effective coupling between magnetic spins which is indirectly routed via nonmagnetic atoms in salts and probably via conduction electrons in metals. Consider the oxygen atom at the position labeled X in Fig. 3. The three dumbbell-shaped electronic wave functions of the oxygen will each overlap a pair of manganese atoms (**Fig. 4**). Along any one of these dumbbells, the ground state is $Mn^{2+}O^{2-}Mn^{2+}$, and the overlap mixes in the excited states $Mn^{+}O^{-}Mn^{2+}$ and $Mn^{2+}O^{-}Mn^{+}$, in which an electron "hops" from oxygen to manganese. The electron hops more easily if its magnetic moment is antiparallel to the manganese magnetic moment. Detailed consideration shows that there is an indirect tendency, from this mechanism, for the magnetic moments of the two manganese ions to be anti-parallel; this can be expressed by an energy of the form $-J_{ij}\,\mathbf{S}_i \cdot \mathbf{S}_j$, with negative J_{ij}. This coupling aligns the moments of second-neighbor manganese ions in an antiparallel array, as in Fig. 3. First neighbors are coupled by "right-angled" superexchange from π-like bonding. This is probably comparable to second-neighbor coupling in MnO but does not affect ordering, primarily because it is geometrically impossible for all first neighbors to be antiparallel to one another.

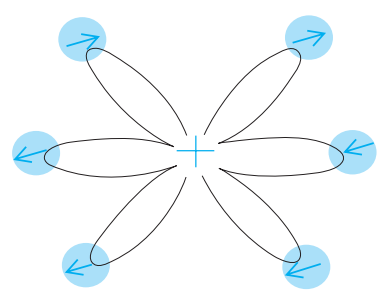

Fig. 4. Superexchange mechanism, in which coupling between magnetic spins of manganese atoms is indirectly routed through oxygen atom in center.

In metals the conduction electrons may play the "hopping" role ascribed above to O^{2-} electrons, or the antiferromagnetism may be related to periodic magnetic order in the electron energy bands.

Magnetic anisotropy. The magnetic moments are known to have preferred directions; these are shown in Figs. 1 and 3. In MnO, it is not known exactly in which direction the moments point, except that it is some direction in the (111) planes. Anisotropic effects come from magnetic dipole forces (predominant in MnF_2 and in MnO) and also from spin-orbit coupling combined with superexchange. Some nearly antiparallel arrays, such as Fe_2O_3, show a slight bending (canting) and exhibit weak ferromagnetism. The anisotropy affects the susceptibility of powder samples and is of extreme importance in antiferromagnetic resonance. *See* MAGNETIC RESONANCE; MAGNETISM. E. Abrahams; E. Keffer; M. Brian Maple

Bibliography. N. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; D. Craik, *Magnetism*, 1995; D. C. Jiles, *Introduction to Magnetism and Magnetic Materials*, 1991; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996.

Antifreeze (biology)

Glycoprotein or protein molecules synthesized by polar and north temperate fishes to enable them to survive in freezing seawater. Similar antifreezes are found in some insects, but relatively little is known about their structure and function.

Freezing-point depression. In a marine fish, the amount of salt and other small molecules in the blood depresses its freezing point to about 30°F (−0.8°C). In the winter, the polar oceans and the nearshore water of north temperate oceans are at the freezing point of seawater, 28.6°F (−1.9°C). In the absence of ice, many fishes survive by supercooling, a thermodynamic state of equilibrium in which a solution (the body fluids of the fish in this case) can be in liquid state, in the absence of ice nuclei, at a temperature lower than the equilibrium freezing point. However, polar waters are often laden with ice that can enter the fish by ingestion of seawater. Propagation of ice in the body fluids or tissues of the fish always leads to freezing damage and death. To avoid freezing, many fishes have evolved biological antifreezes that further lower the freezing point of their body fluids to 28°F (−2.2°C), which is 0.6°F (0.3°C) below the freezing point of seawater.

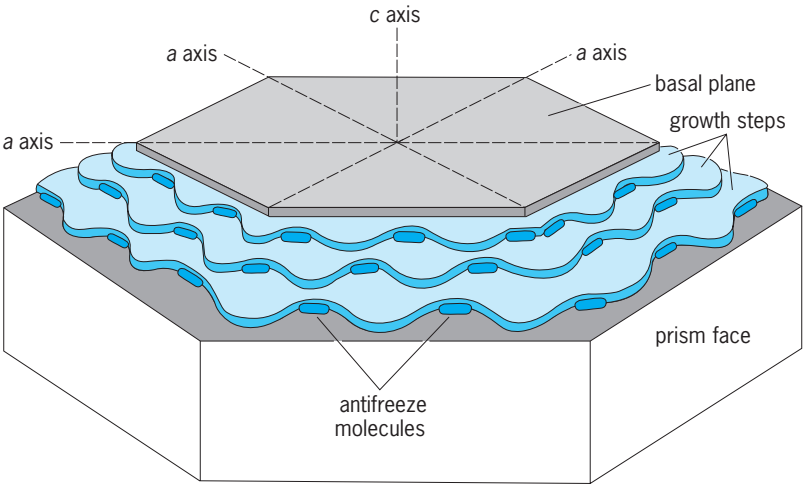

Model of adsorption inhibition as a mechanism of noncolligative lowering of the freezing point of water.

Freezing-point depression of water by a solute is one of the colligative properties and is solely a function of the number of dissolved particles. A 1-molal solution of sodium chloride depresses the freezing point of water by about 6.5°F (3.6°C). The freezing-point depression by biological antifreezes is 200 to 300 times greater than that predicted by their molal concentration, and is therefore a noncolligative process. However, antifreezes show the expected colligative effect on the melting point. This separation of freezing and melting points, sometimes referred to as thermal hysteresis, is a unique property of antifreezes not found in any other class of molecules.

Types. Two types of antifreeze have been isolated; they are either peptides or glycopeptides, the latter containing carbohydrate moieties. Antifreeze glycopeptides were discovered in Antarctic fishes belonging to the family Nototheniidae. They are a series of at least eight different-sized polymers, antifreeze glycopeptides 1 to 8, composed of different numbers of the basic repeating tripeptide unit, alanyl-alanyl-threonine with the disaccharide galactose-*N*-acetylgalactosamine linked to the threonine residue by an α-glycosidic linkage. The molecular masses range from 2600 to 34,000 daltons. The smaller antifreeze glycopeptides (6, 7, and 8) differ slightly from the larger ones in that a few alanines are replaced by prolines. Antifreeze glycopeptides of identical or very similar structure have subsequently been isolated from several northern cods belonging to the family Gadidae. *See* PEPTIDE.

Antifreeze peptides are relatively smaller molecules, mostly of several thousand daltons, and unlike the structurally conserved antifreeze glycopeptides, they have diverse composition and structure. They have been isolated from a number of northern fishes and two Antarctic species. Antifreeze peptides of the winter flounder and Alaskan plaice (flat fishes) and of several sculpins (cottids) have high alanine content. Their primary structures consist of stretches of nonpolar alanines separated by small clusters of polar residues, and their secondary structure is largely an alpha helix. Antifreeze peptides of sea raven, another cottid, however, are rich in cystine and possess beta structure. A third class of antifreeze peptides has been isolated from members of the family Zoarcidae (eel pouts), including two Antarctic, one Arctic, and one Atlantic ocean species. Unlike all the other antifreezes, this type has no biased distribution of amino acids and appears to lack defined secondary structure.

Mechanism of effect. Adsorption of antifreeze molecules to the surface of ice crystals, leading to inhibition of the crystals' growth, has been proposed to explain the noncolligative depression of the freezing

point. Such adsorption has been demonstrated, but precisely how it is achieved is not entirely clear. For the antifreeze glycopeptides and the winter flounder antifreeze peptide, a lattice match between polar side chains and oxygen atoms in the ice crystal allowing hydrogen bonding may account for adsorption. The helical secondary structure of flounder antifreeze peptide places the side chains of the polar residues on one side of the helix, and those of the nonpolar alanine on the other. The spacing between many of the polar side chains is 0.45 nanometer and matches that between the oxygen atoms in the ice crystal parallel to the *a* axes. In the antifreeze glycopeptides, many of the hydroxyl groups of the disaccharide moieties are also spaced 0.45 nm apart.

An ideal ice crystal is a hexagon characterized by a *c* axis normal to three *a* axes (see **illus**.). The preferred direction of growth is along the *a* axes, presumably by water molecules joining the basal plane at the growth steps. Antifreeze molecules are thought to adsorb along the growth steps and force the ice to grow in between them, leading to the formation of many small, highly curved growth fronts. These curved fronts result in a large surface area-to-volume ratio, which raises the surface free energy. In order for freezing to continue, energy must be removed by further reducing the temperature; in other words, the freezing point is lowered. When the temperature is lowered below the freezing point of an antifreeze solution in the presence of a small seed ice crystal, ice growth occurs extremely rapidly along the *c* axis, the nonpreferred axis of growth, in the form of ice needles or spicules. Antifreeze molecules therefore not only inhibit ice crystal growth in between melting point and freezing point but also drastically change the crystal habit of ice growth. *See* COLD HARDINESS (PLANT); CRYPTOBIOSIS.

Arthur L. DeVries

Bibliography. A. L. DeVries, Biological antifreeze agents in cold-water fishes, *Comp. Biochem. Physiol.*, 73A:627–640, 1982; A. L. Devries, Role of glycopeptides and peptides in inhibition of crystallization of water in polar fishes, *Phil. Trans. Roy. Soc. London*, B304:575–588, 1984; A. L. DeVries, The role of antifreeze glycopeptides and peptides in the freezing avoidance of Antarctic fishes, *Comp. Biochem. Physiol.*, 90B:611–621, 1988; J. T. Eastman and A. L. DeVries, Antarctic fishes, *Sci. Amer.*, 254(11):106–114, 1986; J. A. Raymond and A. L. DeVries, Adsorption inhibition as a mechanism of freezing resistance in polar fishes, *Proc. Nat. Acad. Sci. USA*, 74:2589–2593, 1977; J. D. Schrag et al., Primary and secondary structure of antifreeze peptides from Arctic and Antarctic zoarcid fishes, *Biochim. Biophys. Acta*, 915:357–370, 1987.

Antifreeze mixture

A chemical substance that, when added to a liquid such as water, reduces the freezing point of the mixture. Antifreezes are used in a wide variety of applications, the most common being automotive cooling systems. Antifreeze liquids are also used in refrigeration systems (as a secondary coolant), heating and air-conditioning systems, ice skating rinks, and solar energy units, and as deicing agents for runways and aircraft. *See* ENGINE COOLING; REFRIGERATION.

Properties of a desirable antifreeze include the ability to depress the freezing point of the liquid (typically water), good solubility, high boiling point (to provide boil-over protection), chemical compatibility with materials of construction, good heat transfer properties, appropriate viscosity, and low cost. Freezing-point depression is a colligative property and follows Raoult's law, that is, the reduction in freezing point is proportional to the number of molecules of solute per unit of solvent. For example, 46 grams (1 mole) of methanol will depress the freezing point of 1000 grams of water to the same extent as will 342 grams (1 mole) of sucrose (1.86°C or 3.35°F). *See* SOLUTION.

Chemicals that have been used as antifreezes include glycols, glycerol, brines (such as calcium chloride), and alcohols. Ethylene glycol is the most common antifreeze used in automotive cooling systems because of the outstanding freezing-point depression effect, boil-over protection, heat transfer characteristics, high flash point, and low vapor pressure. Propylene glycol, diethylene glycol, and methanol have also been used to a limited extent. Propylene glycol has a lower oral toxicity to mammals than does ethylene glycol, but both glycols are readily biodegraded and are essentially nontoxic to aquatic life. *See* ETHYLENE GLYCOL.

Commercial automotive antifreezes contain corrosion inhibitors to protect the various types of metals in the cooling system. Different formulations are available for different materials of construction and type of service, but typical inhibitors for light-duty automotive use (passenger cars) include phosphates, borates, molybdates, and nitrates for steel and iron protection and thiazoles for copper and brass protection. To protect against corrosion of heat-rejecting aluminum surfaces, special silicate-based corrosion inhibitor packages are used. Antifoam agents and dyes are also frequently added. Heavy-duty applications, such as in diesel trucks, require a low-silicate formulation (because aluminum use in the engine is low, and potential silicate gel formation is avoided) and use supplemental cooling additives to protect against cavitation corrosion of the cylinder liners. Long-life coolants are specially formulated with carboxylic acid salt inhibitors designed to be nondepleting with use. Some locations around the world require phosphate-free inhibitor packages and use carboxylic acid salt, benzoate, and nitrate inhibitor packages. *See* INHIBITOR (CHEMISTRY).

Antifreeze may become unsuitable for use because of depletion of inhibitors, the presence of corrosion products or corrosive ions, or degradation of the glycol. At that point the fluid is replaced and the old fluid is discarded. Although the glycols are readily biodegraded, used antifreeze may pick up heavy metals from the cooling system (such as lead from the solder) so recycling of used coolant may be preferred

for environmental and economic reasons. Recycling can be done a number of ways, but distillation of used antifreeze to remove water and reclaim the glycol is the most common.

Glycol antifreeze solutions are often used in aircraft deicing. These contain additional components for corrosion protection and wetting. Aircraft antiicing fluids also contain a polymeric thickening agent to increase the fluid viscosity, which allows the fluid to adhere to the aircraft surface and provide protection against freezing for a limited period of time. These thickeners impart non-newtonian behavior so that at high shear, such as when the aircraft is taking off, the fluid viscosity is reduced and the fluid flows off the aircraft. Deicing compounds are also used on airport runways to break the ice-surface bond. Chemicals used in this application include glycols, urea, and potassium acetate. *See* NON-NEWTONIAN FLUID.

Kathleen F. George

Bibliography. *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed. vol. 3, 1992; I. L. Rozenfeld, *Corrosion Inhibitors*, 1981; Society of Automotive Engineers, *Engine Coolant, Cooling Systems, Materials and Components*, SP-960, 1993.

Antifriction bearing

A machine element that permits free motion between moving and fixed parts. Antifrictional bearings are essential to mechanized equipment; they hold or guide moving machine parts and minimize friction and wear. Friction wastefully consumes energy, and wear changes dimensions until a machine becomes useless.

Simple bearings. In its simplest form a bearing consists of a cylindrical shaft, called a journal, and a mating hole, serving as the bearing proper. Ancient bearings were made of such materials as wood, stone, leather, or bone, and later of metal. It soon became apparent for this type of bearing that a lubricant would reduce both friction and wear and prolong the useful life of the bearing. Convenient lubricants were those of animal, vegetable, or marine origin such as mutton tallow, lard, goose grease, fish oils, castor oils, and cottonseed oil. Egyptian chariot wheels show evidence of the use of mutton tallow for bearing lubrication.

The use of mineral oils dates back principally to the discovery of Drake's well at Titusville, Pennsylvania, in 1859. Petroleum oils and greases are now generally used for lubricants, sometimes containing soap and solid lubricants such as graphite or molybdenum disulfide, talc, and similar substances.

Materials. The greatest single advance in the development of improved bearing materials took place in 1839, when I. Babbitt obtained a United States patent for a bearing metal with a special alloy. This alloy, largely tin, contained small amounts of antimony, copper, and lead. This and similar materials have made excellent bearings. They have a silvery appearance and are generally described as white metals or as Babbitt metals. For many decades they have served as the measure of excellence against which other bearing materials have been compared. *See* ALLOY.

Wooden bearings are still used, however, for limited applications in light-duty machinery and are frequently made of hard maple which has been impregnated with a neutral oil. Wooden bearings made of lignum vitae, the hardest and densest of all woods, are still used. Lignum vitae is native only to the Caribbean area. This wood has a density of approximately 80 lb/ft^3 (1.3 g/cm^3) and has a resin content of some 30% by volume; thus it is remarkably self-lubricating. The grain is closely interwoven, giving the material high resistance to wear and compression and making it difficult to split. Applications are found in chemical processing and food industries where lignum vitae wood can successfully resist the destructive action of mild acids, alkalies, oils, bleaching compounds, liquid phosphorus, and many food, drug, and cosmetic compounds.

About 1930 a number of significant developments began to occur in the field of bearing metals. Some of the most successful heavy-duty bearing metals are now made of several distinct compositions combined in one bearing. This approach is based on the widely accepted theory of friction, which is that the best possible bearing material would be one which is fairly hard and resistant but which has an overlay of a soft metal that is easily deformed. **Figure 1** shows a number of combinations of bearing materials made up as indicated of combinations of hard and soft materials. **Figure 2** shows actual bearings in which graphite, carbon, plastic, and rubber have been incorporated into a number of designs illustrating some of the material combinations that are presently available.

Rubber has proved to be a surprisingly good bearing material, especially under circumstances in which abrasives may be present in the lubricant. Rubber bearings have found wide application in the stern-tube bearings of ships, on dredge cutter heads, on a number of centrifugal pumps, and for shafting bearings on deep-well pumps. The rubber used is a tough resilient compound similar in texture to that in an automobile tire. These bearings are especially effective with water lubrication, which serves as both coolant and lubricant.

Cast iron is one of the oldest bearing materials. Iron bearings were used in ancient India and China. With the advent of more complicated machinery during the industrial revolution, cast iron became a popular bearing material. It is still used where the duty is relatively light.

Porous metal bearings are frequently used when plain metal bearings are impractical because of lack of space or inaccessibility for lubrication. These bearings have voids of 16–36% of the volume of the bearing. These voids are filled with a lubricant by a vacuum technique. During operation they supply a limited amount of lubricant to the sliding surface between the journal and the bearing. In general, these

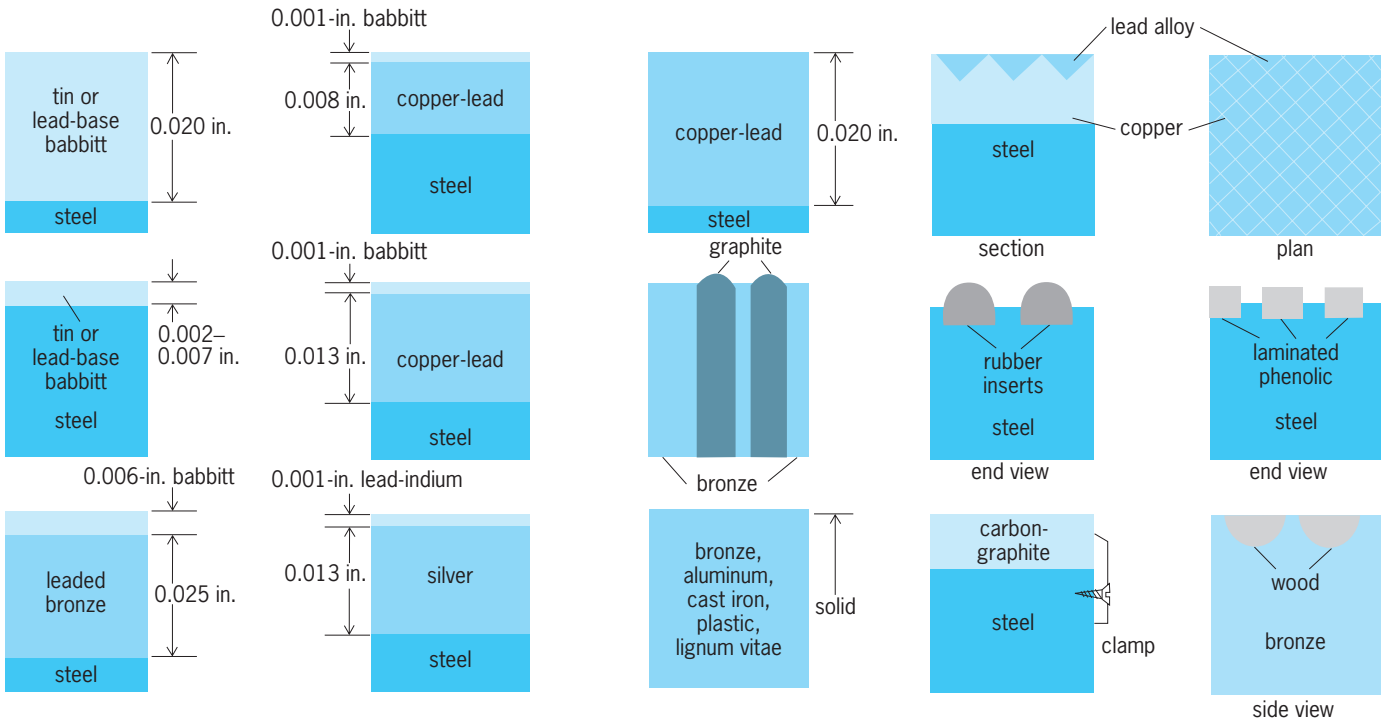

Fig. 1. Schematic of a number of combinations of bearing materials made up of hard and soft materials. 1 in. = 2.54 cm. (*After J. J. O'Connor, ed., Power's Handbook on Bearings and Lubrication, McGraw-Hill, 1951*)

bearings are satisfactory for light loads and moderate speeds.

In some areas recent research has shown that, surprisingly enough, very hard materials when rubbed together provide satisfactory bearing characteristics for unusual applications. Materials such as Stellite, Carboloy, Colmonoy, Hastelloy, and Alundum are used. Because of their hardness, these bearings must be extremely smooth and the geometry must be precise for there is little possibility that these materials will conform to misalignment through the process of wear.

Lubricants. Petroleum oils and greases have been fortified by chemical additives so that they are effective in reducing wear of highly stressed machine elements such as gears and cams. The additives include lead naphthenate, chlorine, sulfur, phosphorus, or similar materials. In general, compounds containing these elements are used as additives to form—through reaction with the metal surfaces—chlorides, sulfides, and phosphides which have relatively low shear strength and protect the surface from wear and abrasion.

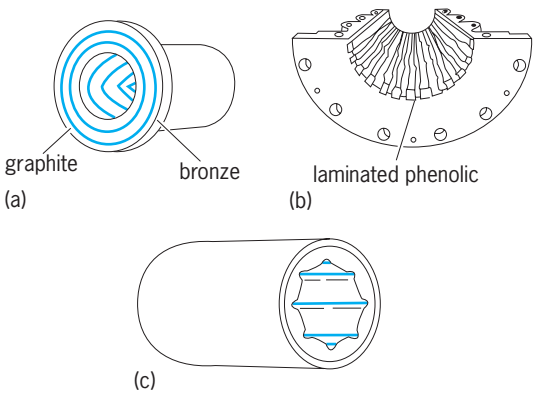

Fig. 2. Bearings with (*a*) graphite, (*b*) wood, plastic, and nylon (*after J. J. O'Connor, ed., Power's Handbook on Bearings and Lubrication, McGraw-Hill, 1951*); (*c*) rubber (*Lucian Q. Moffitt, Inc.*).

The method of supplying the lubricant and the quantity of lubricant which is fed to the bearing by the supplying device will often be the greatest factor in establishing performance characteristics of the bearing. For example, if no lubricant is present, the journal and bearing will rub against each other in the dry state. Both friction and wear will be relatively high. The coefficient of friction of a steel shaft rubbing in a bronze bearing, for example, may be about 0.3 for the dry state. If lubricant is present even in small quantities, the surfaces become contaminated by this material whether it is an oil or a fat, and depending upon its chemical composition the coefficient of friction may be reduced to about 0.1. Now if an abundance of lubricant is fed to the bearing so that there is an excess flowing out of the bearing, it is possible to develop a self-generating pressure film in the clearance space as indicated in **Fig. 3**. These pressures can be sufficient to sustain a considerable load and to keep the rubbing surfaces of the bearing separated. This is the type of bearing that is found on the crank shaft of a typical automobile engine. Unit pressures on these bearings reach and exceed 2500 psi (1.7×10^7 pascals) and with little attention will last almost indefinitely, provided that the oil is kept clean and free from abrasive particles

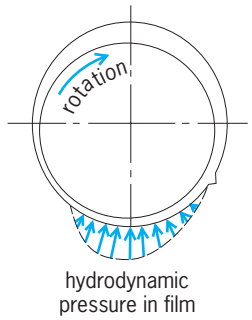

Fig. 3. Hydrodynamic fluid-film pressures in a journal bearing. (*After W. Staniar, ed., Plant Engineering Handbook, 2d ed., McGraw-Hill, 1959*)

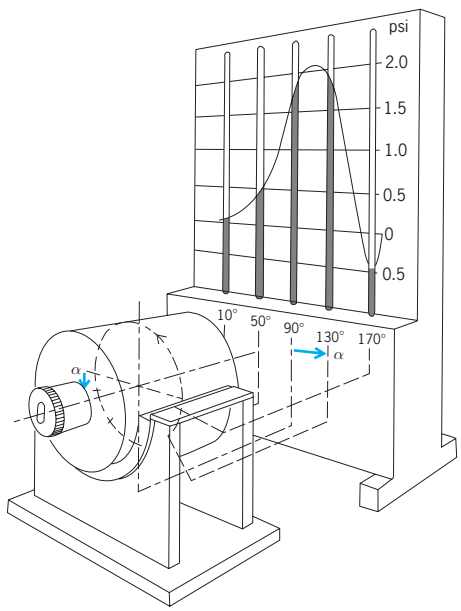

Fig. 4. Schematic of a test device for determining pressure in a journal bearing. 1 psi = 6.9 kPa. (*After H. Drescher, Air lubrication bearings, Eng. Dig., 15:103–107, 1954*)

and that the bearing materials themselves do not deteriorate from fatigue, erosion, or corrosion.

Figure 4 shows a schematic of a simple test device which indicates the pressure developed in the converging clearance space of the journal bearing. The position shown where the center of the journal is eccentric to the center of the bearing is that position which the journal naturally assumes when loaded. If the supply of lubricant is insufficient to fill the clearance space completely or if the load and speed of operation are not favorable to the generation of a complete fluid film, the film will be incomplete and there will be areas within the bearing which do not have the benefit of a fluid film to keep the rubbing surfaces apart. These areas will be only lightly contaminated by lubricant.

The types of oiling devices that usually result in insufficient feed to generate a complete fluid film are, for example, oil cans, drop-feed oilers, waste-packed bearings, and wick and felt feeders.

Oiling schemes that provide an abundance of lubrication are oil rings, bath lubrication, and forced-feed circulating supply systems. The coefficient of friction for a bearing with a complete fluid film may be as low as 0.001. **Figure 5** shows typical devices which do not usually supply a lubricant in sufficient quantity to permit the generation of a complete fluid film. **Figure 6** shows devices in which the flow rate is generally sufficient to permit a fluid film to form. Figure 6*a*, *b*, and *d* shows some typical forms of fluid-film journal bearings. The **table** shows current design practice for a number of bearings in terms of mean bearing pressure. This is the pressure applied to the bearing by the external load and is based on the projected area of the bearing. *See* ENGINE LUBRICATION; JEWEL BEARING; LUBRICANT.

Fluid-film hydrodynamic types. If the bearing surfaces can be kept separated, the lubricant no longer needs an oiliness agent, such as the fats, oils, and greases described above. As a consequence, many extreme applications are presently found in which fluid-film bearings operate with lubricants consisting of water, highly corrosive acids, molten metals, gasoline, steam, liquid refrigerants, mercury, gases, and so on. The self-generation of pressure in such a bearing takes place no matter what lubricant is used, but the maximum pressure that is generated depends upon the viscosity of the lubricant. Thus, for example, the maximum load-carrying capacity of a gas-lubricated bearing is much lower than that of a liquid-lubricated bearing. The ratio of capacities is in direct proportion to the viscosity.

Current practice in mean bearing pressure

Type of bearing	Permissible pressure, psi of projected area*
Diesel engines, main bearings	800–1500
Crankpin	1000–2000
Wristpin	1800–2000
Electric motor bearings	100–200
Marine diesel engines, main bearings	400–600
Crankpin	1000–1400
Marine line-shaft bearings	25–35
Steam engines, main bearings	150–500
Crankpin	800–1500
Crosshead pin	1000–1800
Flywheel bearings	200–250
Marine steam engine, main bearings	275–500
Crankpin	400–600
Steam turbines and reduction gears	100–220
Automotive gasoline engines, main bearings	500–1000
Crankpin	1500–2500
Air compressors, main bearings	120–240
Crankpin	240–400
Crosshead pin	400–800
Aircraft engine crankpin	700–2000
Centrifugal pumps	80–100
Generators, low or medium speed	90–140
Roll-neck bearings	1500–2500
Locomotive crankpins	1500–1900
Railway-car axle bearings	300–350
Miscellaneous ordinary bearings	80–150
Light line shaft	15–25
Heavy line shaft	100–150

*1 psi = 6.9 kPa.

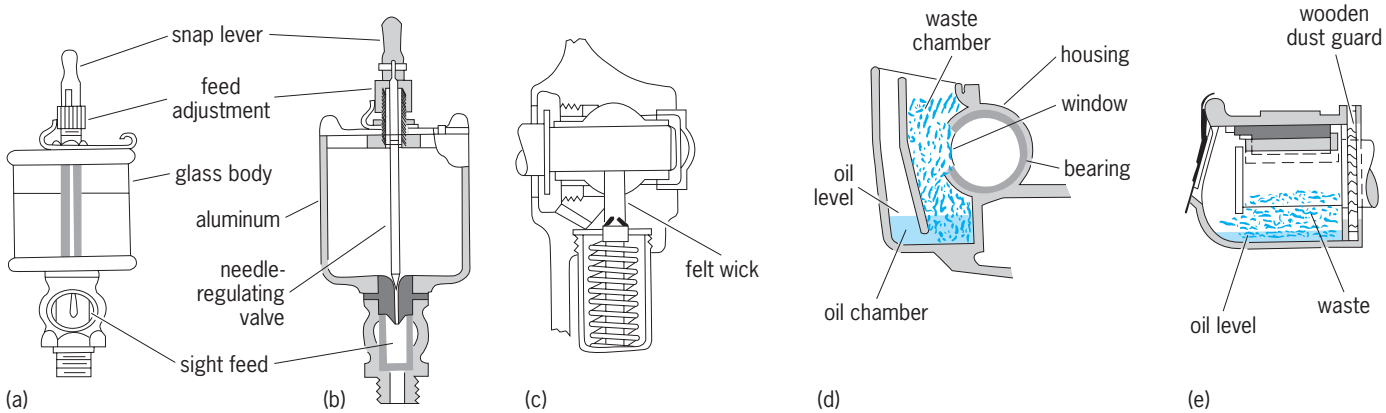

Fig. 5. Schematic showing some of the typical lubrication devices. (*a, b*) Drop-feed oilers. (*c*) Felt-wick oiler for small electric motors. (*d*) Waste-packed armature bearing. (*e*) Waste-packed railroad bearing. (*After W. Staniar, ed., Plant Engineering Handbook, 2d ed., McGraw-Hill, 1959*)

Considerable research has been directed toward the operation of machinery at extremes of temperature. On the low end of the scale this may mean −400°F (−240°C). On the upper end of the scale expectation is that some devices may be called upon to function at 2000–3000°F (1100–1600°C). Gas is the only presently known lubricant that could possibly be used for such extremes of temperature. Because the viscosity of gas is so low, the friction generated in the bearing is correspondingly of a very low order. Thus gas-lubricated machines can be operated at extremely high speeds because there is no serious problem in keeping the bearings cool. A rotor system has been operated on gas-lubricated bearings up to 433,000 revolutions per minute.

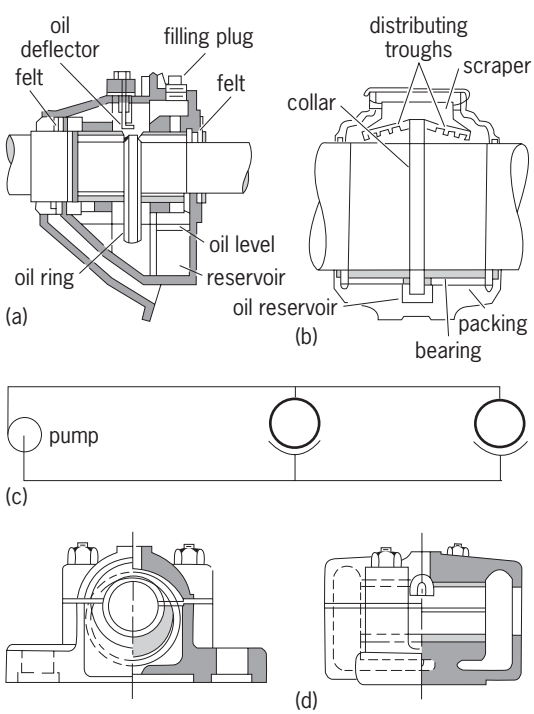

Fig. 6. Lubrication devices. (*a*) Ring-oiled motor bearings. (*b*) Collar-oiled bearing (*after W. Staniar, ed., Plant Engineering Handbook, 2d ed., McGraw-Hill, 1959*). (*c*) Circulating system for oiling bearings (*after J. J. O'Connor, ed., Power's Handbook on Bearings and Lubrication, McGraw-Hill, 1951*). (*d*) Rigid ring-oiling pillow block (*after T. Baumeister, ed., Standard Handbook for Mechanical Engineers, 8th ed., McGraw-Hill, 1978*).

The self-generating pressure principle is applied equally as well to thrust bearings as it is to journal bearings. One of the first commercial applications of the self-generating type of bearing was in the tilting-pad type of thrust bearing. One tilting pad is shown schematically in **Fig. 7***a*. There is certainly no question regarding the great value of the tilting-pad thrust bearing. It excels in low friction and in reliability. A model of a typical commercial thrust bearing is shown in Fig. 7*b*. The thrust bearing is made up of many tilting pads located in a circular position. One of the largest is on a hydraulic turbine at the Grand Coulee Dam. There, a bearing 96 in. (244 cm) in diameter carries a load of 2.15×10^6 lb (976,000 kg) with a coefficient of friction of about 0.0009. Large marine thrust bearings are of this type and transfer the entire thrust of the propeller to the hull of the ship.

Fluid-film hydrostatic types. Sleeve bearings of the self-generating pressure type, after being brought up to speed, operate with a high degree of efficiency and reliability. However, when the rotational speed of the journal is too low to maintain a complete fluid film, or when starting, stopping, or reversing, the oil film is ruptured, friction increases, and wear of the bearing accelerates. This condition can be eliminated by introducing high-pressure oil to the area between the bottom of the journal and the bearing itself, as shown schematically in **Fig. 8**. If the pressure and quantity of flow are in the correct proportions, the shaft will be raised and supported by an oil film whether it is rotating or not. Friction drag may drop to one-tenth of its original value or even less, and in certain kinds of heavy rotational equipment in which available torque is low, this may mean the difference between starting and not starting. This type of lubrication is called hydrostatic lubrication and, as applied to a journal bearing in the manner indicated, it is called an oil lift. Synchronous condensers need the oil lift when the unit is of large size. Rolling-mill and foil-mill bearings may be equipped with an oil lift

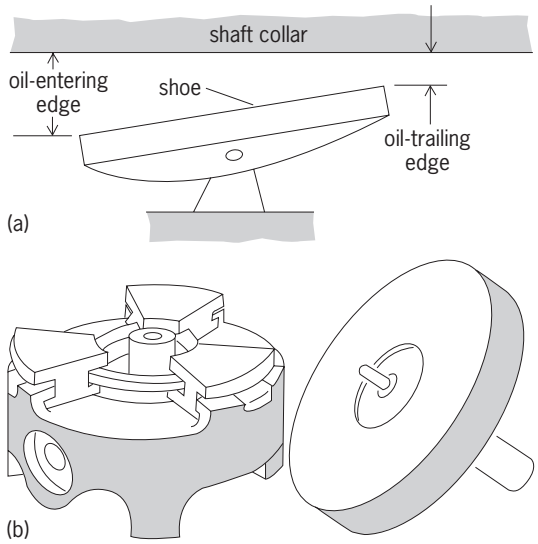

Fig. 7. Tilting-pad-type bearing. (*a*) Schematic of tilting pad (*after W. Staniar, ed., Plant Engineering Handbook, 2d ed., McGraw-Hill, 1959*). (*b*) Thrust bearing (*after D. D. Fuller, Theory and Practice of Lubrication for Engineers, Wiley, 1956*).

to reduce starting friction when the mills are under load. Occasionally, hydrostatic lifts are used continuously on bearings that are too severely overloaded to maintain a normal hydrodynamic or self-pressurizing oil film.

Hydrostatic lubrication in the form of a step bearing has been used on various machines to carry thrust. Such lubrication can carry thrust whether the shaft is rotating or not and can maintain complete separation of the bearing surfaces. **Figure 9** shows a schematic representation of such a step-thrust bearing. High-speed machines such as ultracentrifuges have used this principle by employing air as the lubricant at speeds upward of 100,000 rpm.

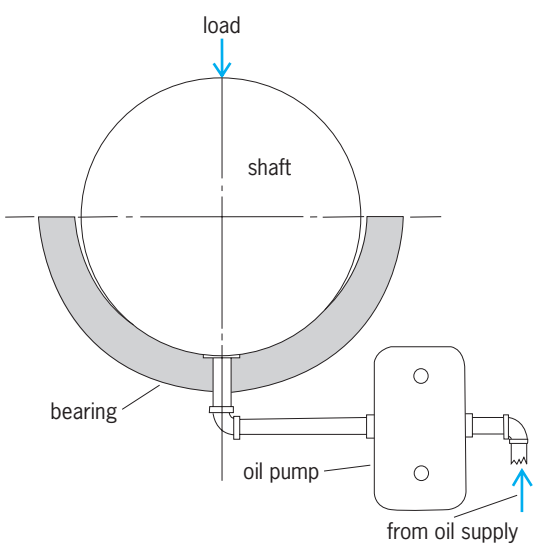

Fig. 8. Fluid-film hydrostatic bearing. Hydrostatic oil lift can reduce starting friction drag to less than one-tenth of usual starting drag. (*After W. Staniar, ed., Plant Engineering Handbook, 2d ed., McGraw-Hill, 1959*)

Large structures have been floated successfully on hydrostatic-type bearings. For example, the Hale 200-in. (5-m) telescope on Palomar Mountain weighs about 1×10^6 lb (500,000 kg); yet the coefficient of friction for the entire supporting system, because of the hydrostatic-type bearing, is less than 0.000004. The power required is extremely small and a $^1/_{12}$-hp clock motor rotates the telescope while observations are being made. Hydrostatic bearings are currently being applied to large radio telescopes and radar antennas, some of which must sustain forces of 5×10^6 lb (2×10^6 kg) or more, considering wind loads as well as dead weight. One such unit constructed at Green Bank, West Virginia, by the Associated Universities has a parabolic disk or antenna 140 ft (42 m) in diameter.

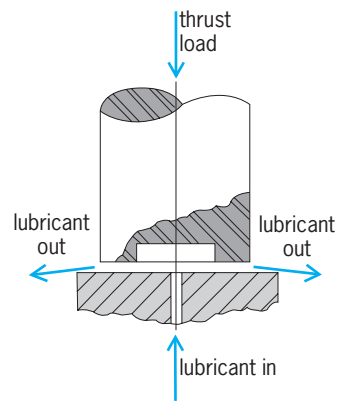

Fig. 9. Step bearing. (*After W. Staniar, ed., Plant Engineering Handbook, 2d ed., McGraw-Hill, 1959*)

Rolling-element types. Everyday experiences demonstrate that rolling resistance is much less than sliding resistance. The wheelbarrow, the two-wheeled baggage truck, and similar devices are striking examples of the reduction in friction by the use of the wheel. Heavy crates and similar objects are easily moved by introducing rollers under the leading edge of the load while the crate or object is pushed along. Egyptian engineers building the pyramids transported huge blocks of stone from the quarry to the building site by means of rollers. This principle is used in the rolling-element bearing which has found wide use.

The first major application of these bearings was to the bicycle, the use of which reached its peak just before the year 1900. In the development of the automobile, ball and roller bearings were found to be ideal for many applications, and today they are widely used in almost every kind of machinery.

Structure. These bearings are characterized by balls or cylinders confined between outer and inner rings. The balls or rollers are usually spaced uniformly by a cage or separator. The rolling elements are the most important because they transmit the loads from the moving parts of the machine to the stationary supports. Balls are uniformly spherical, but the rollers may be straight cylinders, or they may be barrel- or cone-shaped or of other forms, depending upon

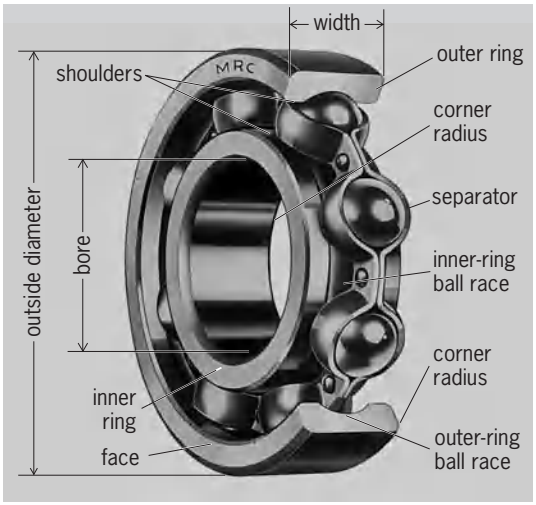

Fig. 10. Deep-groove ball bearing. (*Marlin-Rockwell*)

the purpose of the design. The rings, called the races, supply smooth, hard, accurate surfaces for the balls or rollers to roll on. Some types of ball and roller bearings are made without separators. In other types there is only the inner or outer ring, and the rollers operate directly upon a suitably hardened and ground shaft or housing. **Figure 10** shows a typical deep-grooved ball bearing, with the parts that are generally used.

These bearings may be classified by function into three groups: radial, thrust, and angular-contact bearings. Radial bearings are designed principally to carry a load in a direction perpendicular to the axis of rotation. However, some radial bearings, such as the deep-grooved bearings shown in Fig. 10, are also capable of carrying a thrust load, that is, a load parallel to the axis of rotation and tending to push the shaft in the axial direction. Some bearings, however, are designed to carry only thrust loads. Angular-contact bearings are especially designed and manufactured to carry heavy thrust loads and also radial loads.

Life. A unique feature of rolling-element bearings is that their useful life is not determined by wear but by fatigue of the operating surfaces under the repeated stresses of normal use. Fatigue failure, which occurs as a progressive flaking or pitting of the surfaces of the races and rolling elements, is accepted as the basic reason for the termination of the useful life of such a bearing.

Because the load on a bearing determines the severity of the stress in the surfaces of the races and the rolling elements, it follows that if the load is increased, the life of the bearing will be decreased, and conversely if the load is decreased, the life of the bearing will be increased. This is usually expressed by the relationship that the life of the bearing is inversely proportional to the load cubed. Thus, doubling the load will reduce the life of the bearing by a factor of 8.

The life of a bearing under a given load will therefore be a certain number of revolutions. If this number of revolutions is used up at a relatively high rate, the life of the bearing will be correspondingly short. If the total number of revolutions is used up at a low rate, the life of the bearing is correspondingly longer; that is, the life is inversely proportional to the speed. Life expectancy is a prediction based on results obtained from tests of a large number of identical bearings under identical loads and speeds. Individual bearings may deviate from this figure on a statistical basis, but manufacturers have followed the law of averages in establishing their ratings. For example, some manufacturers specify load and speed ratings for their bearings based on 3000 h of operation. The bearing manufacturer who uses 3000 h as a life-expectancy figure assumes that at least 90% of all bearings will last 3000 h under the specified conditions of load and speed. Based on statistical averages, however, this means that 10% of the bearings will fail before reaching the 3000-h life expectancy, 50% of the bearings will attain five times the design life, and a few bearings may reach 20–30 times the design life of 3000 h.

Characteristics. The various types of rolling-contact bearings can be identified in terms of their broad general characteristics (**Fig. 11**). A separable or a magneto-type bearing is useful where disassembly is frequent (Fig. 11*a*). The outer race is pressed firmly in the removable housing; the inner race may be pressed against a shoulder on the shaft. A deep-grooved bearing with a filling slot (Fig. 11*b*) allows more balls to be used in the bearing than are shown in Fig. 10 and will thus carry heavier radial loads. Because of the filling slot, however, it should be used only for light thrust loads. If the thrust is in one direction, the bearing should be mounted with the slot away from the thrust direction. The double-row radial bearing with deep grooves handles heavier radial and thrust loads than a single-roll bearing of the same dimensions. Internal self-aligning double-roll bearings (Fig. 11*d*) may be used for heavy radial loads where self-alignment is required. The self-aligning feature should not be used to correct poor design or assembly because excessive misalignment will harm the bearing. Thrust loads should be light because a thrust load will be sustained by only one row of balls. Figure 11*e* shows an external self-aligning bearing which requires a larger outside diameter, but has the advantage of being able to carry thrust in either direction, as well as providing a self-aligning feature. Angular contact bearings (Fig. 11*f*) provide for a maximum thrust and modest radial loads. They may be mounted back to back as duplex bearings and carry thrust in either direction. To minimize axial movement of such a bearing and the shaft that it restrains, these bearings may be preloaded to take up any possible slack or clearance. A ball bushing is shown in Fig. 11*g*. This is used for translational motion or motion of a shaft in its axial direction. Any such motion should be at a very low speed and with light radial loads. Roller bearings with short straight rollers are shown in Fig. 11*h*. They permit free movement in the axial direction when the guide lips are on either the outer or inner race. Needle bearings (Fig. 11*i*)

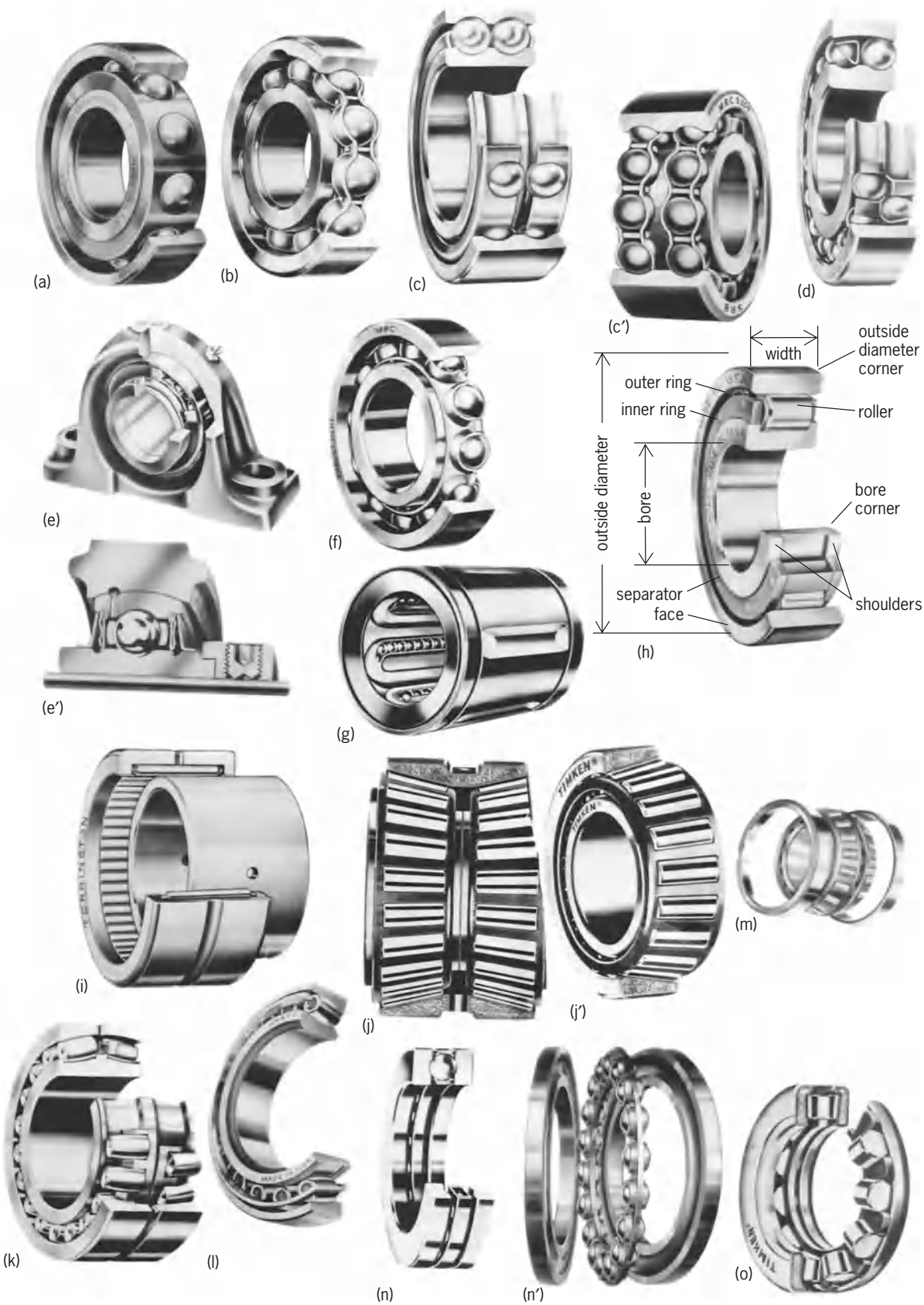

Fig. 11. Views of the various ball and roller bearings. (*a*) Separable angle-contact ball bearing (*New Departure*). (*b*) Deep-groove ball bearing with filling slot (*New Departure*). (*c*) Double-row deep-groove ball bearings (*SKF; Marlin-Rockwell*). (*d*) Internal self-aligning bearing (*SKF*). (*e*) Deep-groove ball bearing with spherical outer ring. (*f*) Standard angular-contact ball bearing (*Marlin-Rockwell*). (*g*) Ball bushing that permits unlimited travel, linear motion (*Thomson Industries*). (*h*) Roller bearing showing nomenclature (*Hyatt*). (*i*) Needle roller bearing (*Torrington*). (*j*) Tapered roller bearings, two-row and single-row (*Timken*). (*k*) Radial spherical roller bearings (*Torrington*). (*l*) Barrel roller bearing (*Hyatt*). (*m*) Concavex roller bearing (*Shafer*). (*n*) Single-direction ball-bearing thrust (*SKF; Marlin-Rockwell*). (*o*) Cylindrical roller thrust bearing (*Timken*).

-s

have rollers whose length is at least four times the diameter. They may be furnished with a retainer but have their maximum load capacity when there is no retainer and the clearance space is completely filled with rollers. They are with or without an inner race and are most useful where space must be saved. If the shaft is used as an inner race, it should be hardened.

The tapered roller bearings shown in Fig. 11*j* permit heavy radial and thrust loads. These may be procured as double tapered rollers with two inner races and one outer or two outer races and one inner. The race cones intersect at the axis of the bearing. Figure 11*k* shows a self-aligning bearing with two rows of short barrel rollers. The spherical outer race is one piece, or it may be in two parts for preloading. The thrust is taken on the spherical surfaces of the center flange of an inner race. Figure 11*l* shows a barrel roller with spherical outer race. Figure 11*m* is a self-aligning bearing with hourglass rollers. This bearing is also built as a two-row bearing with a one-piece inner race.

The ball thrust bearing shown in Fig. 11*n* is used only for carrying thrust loads acting in the direction of the axis of the shaft. It is used for low-speed applications, while other bearings must support the radial load. Straight roller thrust bearings are made of a series of short rollers to minimize twisting or skewing. The tapered roller thrust bearing shown in Fig. 11*o* eliminates the twisting that may take place with the straight rollers but cause a thrust load between the ends of the rollers and the shoulder on the race.

Mountings. Many types of mountings are available for rolling-element bearings, and their selection will depend upon the variety of service conditions encountered. In the preferred method of mounting, the inner and outer races are held securely in place to prevent creeping or spinning of the races in the housing or on the shaft. In rigid mountings, provision must be made for expansion either by building slip clearances into one mount or by using a straight roller bearing on the free end of the shaft. A bearing can be mounted on a shaft with a snap ring on the outer race which locates the bearing against a shoulder. Bearings may also be mounted on shafts by means of an eccentric ring cut on the side of the extended inner race. A set screw can be used to hold the bearings in place. *See* FRICTION; WEAR.

Dudley D. Fuller.

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; J. E. Shigley and C. R. Mischke, *Bearings and Lubricaton*: *A Mechanical Designers' Handbook*, 1990.

Antigen

Any substance that causes the immune system to produce specific antibodies or T cells against it. An antigen may be a foreign substance from the environment (such as chemicals, bacterial or viral proteins, or pollen) or formed within the host's own body. Reactions of antigens with antibodies or T cells serve as a defense against microorganisms and other foreign bodies, but can be detrimental if the immune response is mounted against the "self," as in autoimmune disorders. Antigen-antibody complexes are used in laboratory tests for detecting the presence of either antigen or antibody to determine a subject's previous exposure to pathogens. *See* ANTIBODY; ANTIGEN-ANTIBODY REACTION.

Classes of antigens. Antigens can be classified into three distinct groups: immunogens, allergens, and tolerogens.

Immunogens. An immunogen is any substance that provokes an immune response. It is usually a protein or polysaccharide, and its ability to stimulate the immune response depends on many factors, including its foreignness, size, chemical composition, and conformation. Peptide sequences with high segmental mobility are frequently antigenic. Because an immunogen is rather large, it usually contains a large number of antigenic determinants, or epitopes (sites recognized by antibodies). Immunizing an animal with a protein results in the formation of numerous antibody molecules with different specificities, the number of different antibodies depending on the number of antigenic determinants. Since the introduction of technology for producing monoclonal antibodies, it has become possible to produce in the laboratory antibodies specific to virtually any antigen. *See* MONOCLONAL ANTIBODIES.

Immunogens may be introduced into an animal by ingestion, inhalation, contact with skin, or by injection into the bloodstream, skin, peritoneum, or other tissue.

Allergens. Allergens are substances that cause an allergic reaction by preferentially driving the immune system to an immunoglobulin E (IgE) response. Allergens can be ingested, inhaled, injected, or come into contact with skin. Allergens include dust mites, dander, peanuts, shellfish, and many other common environmental agents. *See* ALLERGY; IMMUNOGLOBULIN.

Tolerogens. Tolerogens are antigens that cause a specific immune non-responsiveness due to their molecular form. They are thought to provide only partial signals to the T or B cells, which respond by shutting down immune functions. Tolerogens are usually "self," that is, produced by the host's own body, which explains why, in general, we do not have an immune response to proteins normally made in the body. When this system does not work it is termed autoimmunity. *See* AUTOIMMUNITY; CELLULAR IMMUNITY.

Antibody response. Since antigens are complex molecules, the antibody response to them may also be complex. A mixture of antibodies may be produced, each corresponding to only a portion (epitope) of the original antigen surface (antigenic mosaic). If the antigenic entity is not a single molecular species but, as in the case of whole microorganisms, a mixture of separate antigens, the complexity is multiplied accordingly. A small part of an antigen may stimulate an antibody capable of recognizing not only that antigen but related antigens; this is known as a cross-reaction. The most famous example of a cross-reaction came from Edward Jenner's

discovery that milkmaids infected with cowpox were protected from smallpox.

Sources of antigens. Bacteria, viruses, protozoans, and other microorganisms are important sources of antigens. These may be proteins or polysaccharides derived from the outer surfaces of the cell (capsular antigens), from the cell interior (somatic or O antigens), or from the flagella (flagellar or H antigens). Other antigens are either secreted by the cell or are released during cell death and disruption; these include many enzymes and toxins, of which diphtheria, tetanus, and botulinum toxins are important examples. The presence of antibody to one of these constituent antigens in human or animal sera is presumptive evidence of past or present contact with specific microorganisms, and this finds application in clinical diagnosis and epidemiological surveys. *See* BOTULISM; DIPHTHERIA; TETANUS; TOXIN.

Use in vaccines. In medicine, the ability of antigen to stimulate the production of antibodies or T cells is the basis of vaccines. In traditional antigen-based vaccines, the "target" antigens of a microorganism are produced and deliberately administered to induce an antibody response. There are four types of antigen-based vaccines: live, purified, recombinant, and synthetic. Live vaccines are made from viruses that have been attenuated (made nonpathogenic) by long culture in cells or animals until the dangerous genes have mutated. They are among our most effective vaccines. Purified antigen vaccines are those that are composed of molecules purified directly from a pathogen. Because these vaccines use a limited number of molecules from the parent pathogen, there is no danger of pathogen replication. Recombinant antigen vaccines are created by insertion of deoxyribonucleic acid (DNA) encoding an antigenic protein into bacteria, yeast, or viruses that infect cultured mammalian cells, or by direct transfection of the DNA into the mammalian cells. The organisms or cells will then produce the protein, which can be further purified for use as a vaccine. Synthetic antigen vaccines are those that use peptide antigens synthesized by automated machines. While the technology exists, this method of making vaccines is not yet in use. *See* VACCINATION.

Vaccines have been made from protein antigens derived from cancer cells. These aim to stimulate the patient's immune system to attack the cancer. Scientists have recently worked out the genetic codes of many of these cancer cell proteins, so they can make them in the laboratory in large quantities.

Margaret J. Polley; Zoë Cohen

Bibliography. J. Chinen and W. Shearer, Basic and clinical immunology, *J. Allergy Clin. Immunol.*, 116(2):411–418, 2005; T. Doan, R. Melzold, and C. Waltenbaugh, *Concise Medical Immunology*, Lippincott Williams & Wilkins, 2005; C. A. Janeway, Jr. et al., *Immunobiology*, Garland Publishing, 6th ed., 2004; W. E. Paul (ed.), *Fundamental Immunology*, Lippincott Williams & Wilkins, 5th ed., 2003; M. Tosi, Innate immune responses to infection, *J. Allergy Clin. Immunol.*, 116(2):241–249, 2005.

Antigen-antibody reaction

A reaction that occurs when an antigen combines with a corresponding antibody to produce an immune complex. A substance that induces the immune system to form a corresponding antibody is called an immunogen. All immunogens are also antigens because they react with corresponding antibodies; however, an antigen may not be able to induce the formation of an antibody and therefore may not be an immunogen. For instance, lipids and all low-molecular-weight substances are not immunogenic. However, many such substances, termed haptens, can be attached to immunogens, called carriers, and the complex then acts as a new immunogen capable of eliciting antibody to the attached hapten. *See* ANTIBODY; ANTIGEN.

Antibody-antigen specificity. An antibody, which belongs to a group of proteins in blood plasma called immunoglobulins, consists of two functionally and structurally distinct portions. The major portion is structurally identical for all antibodies of its immunoglobulin class and takes no part in the reaction with antigen. The other part varies markedly between antibodies and is responsible for binding antigen; in fact, its variability confers on it specificity, that is, the ability to react with one specific antigen among many. *See* IMMUNOGLOBULIN.

A molecule of antibody has two identical binding sites for one antigen or more, depending on its class. Each site is quite small and can bind only a comparably small portion of the surface of the antigen, which is termed an epitope. The specificity of an antibody for an antigen depends entirely upon the possession of the appropriate epitope by an antigen. The binding site on the antibody and the epitope on the antigen are complementary regions on the surface of the respective molecules which interlock in the antigen-antibody reaction. If a similar epitope occurs on a molecule other than the original immunogen, the antibody will bind to that molecule in a cross-reaction. The intensity with which an antibody binds to the antigen depends on the exactitude of the fit between the respective binding site and epitope, as well as some inherent characteristics of the reacting molecules and factors in the environment. The epitope must be continuous spatially, but not structurally: in other words, if the molecule of the antigen consists of several chains, like several pieces of string that have been dropped on top of one another, then an epitope may be formed by adjacent regions on two different chains, as well as by adjacent regions on the same chain. If the epitope is now modified either chemically (for example, by altering the hapten) or physically (for example, by causing the chains to separate), then its fit in the binding site will be altered or abolished, and the antigen will react with the antibody either less strongly or not at all.

Nature of immune complex. The immune complex formed in the reaction consists of closely apposed, but still discrete, molecules of antigen and antibody. Therefore, the immune complex can dissociate

into the original molecules. The proportion of the dissociated, individual molecules of antigen and antibody to those of the immune complex clearly depends on the intensity of the binding. These proportions can be measured in a standardized procedure, so that the concentration of antigen [Ag], antibody [Ab], and the immune complex [AgAb] becomes known. A fraction is then calculated and called either the dissociation constant ($K_d = [Ag] \times [Ab]/[AgAb]$) or the association constant ($K_a = [AgAb]/[Ag] \times [Ab]$). The magnitude of either of these constants can be used subsequently to assess the intensity of the antigen-antibody reaction, for example, to select the most strongly binding of several antibodies for use in immunoassay; to compare the binding of an antibody to an immunogen with that to the corresponding hapten; or to detect modification in the epitope. *See* IMMUNOASSAY.

Only one epitope of its kind generally occurs on each molecule of antigen, other than that which consists of multiple, identical units, though many epitopes of different configuration are possible. Particles, however, either natural ones such as cells or suitably treated artificial ones made of, for example, latex or glass, typically carry multiple identical epitopes, as well as nonidentical ones, because their surfaces contain many molecules of the same antigen. An antibody molecule, bearing two or more binding sites, can become attached to a corresponding number of antigen molecules. If the antibody is specific for only one epitope, as is the case with monoclonal antibody, then it can link with one molecule of antigen at each binding site, but the antigen cannot link with yet another antibody, because it has no more identical epitopes. In a solution of single molecules of antigen, therefore, monoclonal antibody can form at most complexes no larger than those composed of one antibody molecule and the number of antigen molecules corresponding to that of the number of binding sites. By contrast, polyclonal antibody to an antigen consists of antibodies to a variety of epitopes on that antigen, and it can, therefore, link together a very large number of molecules. *See* MONOCLONAL ANTIBODIES.

Immune complexes comprising many molecules eventually reach sufficient size to scatter light, at which point they can be detected by nephelometry or turbidimetry; if their growth continues, they become visible as precipitates, which can also be assayed by such methods as immunodiffusion. Since particles typically carry many molecules of antigen, they can be, in principle, aggregated by monoclonal antibody as well as polyclonal antibody, though the latter is much more effective; and the reaction can be detected by inspection, for example the agglutination of bacteria or red cells, though much more sensitive methods, such as particle counting, are now available. Antigen-antibody reactions can also be detected at very low concentration of reactants through special techniques such as immunofluorescence and radioimmunoassay. *See* IMMUNOASSAY; IMMUNOFLUORESCENCE; IMMUNONEPHELOMETRY; RADIOIMMUNOASSAY.

The reaction between antigen and antibody is followed by a structural change in the remainder of the antibody molecule. The change results in the appearance of previously hidden regions of the molecule. Antibodies to these hidden regions are known as rheumatoid factors owing to their frequently high concentration in the blood of individuals suffering from rheumatoid arthritis. Some of these hidden regions have specific functions, however, such as binding complement. Fixation of complement by immune complexes has been used to detect and measure antigen-antibody reactions. *See* COMPLEMENT.

Factors affecting reaction. The antigen-antibody reaction is affected by various factors, including pH, temperature, and concentration, as well as the type of other dissolved substances. Acidic or markedly alkaline solutions, excessively high or low salt concentration, and the presence of chaotropic substances all prevent the formation of immune complexes or cause the dissociation of preformed ones. Substances that reduce charge on particles, allowing them to aggregate more easily, or reduce free water in solution, for example, polymeric hydrophilic materials such as poly(ethylene glycol), assist the formation of immune complexes.

Uses. The chief use of antigen-antibody reactions has been in the determination of blood groups for transfusion, serological ascertainment of exposure to infectious agents, and development of immunoassays for the quantification of various substances. *See* BLOOD GROUPS; IMMUNOLOGY; SEROLOGY.

Alexander Baumgarten

Bibliography. B. Cinader and R. G. Miller (eds.), *Progress in Immunology VI*, 1986; E. L. Cooper, *General Immunology*, 1982; E. A. Kabat, *Structural Concepts in Immunology and Immunochemistry*, 2d ed., 1976; N. R. Rose, H. Friedman, and J. L. Fahey, *Manual of Clinical Laboratory Immunology*, 3d ed., 1986.

Antihistamine

A type of drug that inhibits the combination of histamine with histamine receptors. These drugs are termed either H-1 or H-2 receptor antagonists depending on which type of histamine receptor is involved. H-1 receptor antagonists are used largely for treating allergies, and H-2 receptor antagonists are used to treat peptic ulcer disease and related conditions.

H-1 receptor antagonists. H-1 receptor antagonists have been available for many years and include agents such as diphenhydramine, chlorpheniramine, doxylamine, cyclizine, tripelennamine, promethazine, and terfenadine. Their primary therapeutic use is to antagonize the effects of histamine that is released from cells by antigen-antibody reactions; they can thus inhibit histamine-induced effects, such as bronchoconstriction, skin reactions, for example, wheals and itching, and nasal inflammation. These drugs, therefore, are quite effective in reducing

allergy signs and symptoms, especially if they are administered before contact with the relevant antigen (for example, ragweed or other pollens); however they are not effective in treating asthma. *See* ALLERGY; ASTHMA.

Other uses of H-1 receptor antagonists may involve actions having little or nothing to do with their ability to antagonize histamine. For example, some (diphenhydramine, cyclizine, meclizine) are effective as antimotion-sickness drugs, and some (diphenhydramine, doxylamine, pyrilamine) are important components of over-the-counter sleep aids. Although preparations that are marketed as being useful in treating the common cold often contain H-1 receptor antagonists, these drugs have no specific actions against the cold virus.

H-1 receptor antagonists have low toxicity. The chief adverse effect is sedation. However, this effect varies widely, both among the drugs and from individual to individual; in young children excitement may be seen. Another common set of effects caused by many of these drugs, including dry mouth, blurred vision, and urinary retention, can be ascribed to their anticholinergic actions. Agents such as terfenadine and loratadine have weaker sedative and anticholinergic actions than do diphenhydramine and doxylamine. Overdoses of H-1 receptor antagonists may be associated with excitement or depression, and although there is no pharmacologic antidote for these drugs, good supportive care should be adequate in managing cases of poisoning. *See* SEDATIVE.

H-2 receptor antagonists. H-2 receptor antagonists are much newer, but they rapidly became some of the most frequently prescribed drugs. Histamine stimulates gastric acid secretion by combining with H-2 receptors. By preventing this combination, H-2 antagonists can reduce acid secretion in the stomach, an effect that makes these drugs useful in managing various conditions.

The medical treatment of peptic ulcer disease, a common disorder in the Western world, involves giving antacids and other medication to reduce gastric acidity. The effectiveness of H-2 antagonists in reducing gastric acid has made these drugs a primary treatment for peptic ulcer. In patients with active ulcers, these drugs reduce symptoms and promote healing of the ulcers. They also prevent recurrence of ulcers. All of the H-2 antagonists (cimetidine, ranitidine, famotidine, nizatidine) appear equally effective in reducing gastric acidity. *See* ULCER.

Other conditions in which H-2 antagonists are used to lower gastric acidity include reflux esophagitis, stress ulcers, and hypersecretory states such as the Zollinger-Ellison syndrome, in which tumor cells secrete large amounts of the hormone gastrin, which stimulates gastric acid secretion. In these conditions, administration of H-2 antagonists reduces symptoms and promotes healing.

The toxicity of H-2 antagonists is quite low, and adverse effects are reported by only 1–2% of patients. The most common side effects are gastrointestinal upsets, including nausea, vomiting, and diarrhea.

One effect associated only with cimetidine is the ability to inhibit the metabolism of other drugs. Because actions of drugs taken concurrently with cimetidine may thus be enhanced or prolonged, dosages of those other drugs may have to be reduced in order to prevent toxicity. *See* HISTAMINE. Alan Burkhalter

Bibliography. B. Ulnas et al. (eds.), *Histamine and Histamine Antagonists*, 1991.

Antimatter

Matter which is made up of antiparticles. At the most fundamental level every type of elementary particle has its anti-counterpart, its antiparticle. The existence of antiparticles was implied by the relativistic wave equation derived in 1928 by P. A. M. Dirac in his successful attempt to reconcile quantum mechanics and special relativity. The antiparticle of the electron (the positron) was first observed in cosmic rays by C. D. Anderson in 1932, while that of the proton (the antiproton) was produced in the laboratory and observed by E. Segré, O. Chamberlain, and their colleagues in 1955. *See* ELECTRON; ELEMENTARY PARTICLE; POSITRON; PROTON; QUANTUM MECHANICS; RELATIVITY.

Properties of antiparticles. The mass, intrinsic angular momentum (spin), and lifetime (in the case of unstable particles) of antiparticles and their particles are equal, while their electromagnetic properties, that is, charge and magnetic moment, are equal in magnitude but opposite in sign. Some neutrally charged particles such as the photon and π° meson are their own antiparticles. Certain other abstract properties such as baryon number (protons and neutrons are baryons and have baryon number $+1$) and lepton number (electrons and muons are leptons and have lepton number $+1$) are reversed in sign between particles and antiparticles. *See* ANGULAR MOMENTUM; BARYON; LEPTON.

The quantum-mechanical operation of turning particles into their corresponding antiparticles is termed charge conjugation (C), that of reversing the handedness of particles is parity conjugation (P), and that of reversing the direction of time is time reversal (T). A fundamental theorem, the *CPT* theorem, states that correct theories of particle physics must be invariant under the simultaneous operation of C, P, and T. Simply put, the description of physics in a universe of antiparticles with opposite handedness where time runs backward must be the same as the description of the universe. One consequence of the *CPT* theorem is that the above-mentioned properties of antiparticles (mass, intrinsic angular momentum, lifetime, and the magnitudes of charge and magnetic moment) must be identical to those properties of the corresponding particles. This has been experimentally verified to a high precision in many instances; for example, the magnitude of the magnetic moments of positrons and electrons are equal to within 1 part in 10^{12}, the charge-to-mass ratios of antiprotons and protons are the same to within 1 part in 10^9, and the masses of the K^0 meson and its antiparticle

are equal to 1 part in 10^{18}. *See* CPT THEOREM; PARITY (QUANTUM MECHANICS); SYMMETRY LAWS (PHYSICS); TIME REVERSAL INVARIANCE.

It is presently thought that leptons (electrons, muons, tau leptons, and their associated neutrinos) are fundamental particles. Baryons and mesons are composite particles, being made of yet more fundamental particles, quarks. As with all elementary particles, for every type of quark there is a corresponding antiquark. Baryons are composed of three quarks, and mesons of quark-antiquark pairs. Antibaryons are then made up of three antiquarks, while antimesons are made of the conjugate antiquarkquark pair of the associated meson. *See* BARYON; MESON; NEUTRINO; QUARKS.

Annihilation and production. When a particle and its antiparticle are brought together, they can annihilate into electromagnetic energy or other particles and their antiparticles in such a way that all memory of the nature of the initial particle and antiparticle is lost. Only the total energy (including the rest-mass energy, $E = mc^2$, where m is the rest mass and c is the speed of light) and total angular momentum remain. The annihilation of an electron and positron releases about 10^6 electronvolts (1 MeV) of electromagnetic energy (1.6×10^{-13} J). This can be compared with the energy of 4 eV released in a typical chemical reaction, or 200 MeV of a typical nuclear fission reaction. Annihilation of a proton and antiproton can release 1880 MeV of electromagnetic energy. *See* ELECTRON-VOLT; JOULE'S LAW.

In the reverse process, antiparticles can be produced in particle collisions with matter if the colliding particles possess sufficient energy to create the required mass. For example, a photon with sufficient energy which interacts with a nucleus can produce an electron-positron pair. The creation of antibaryons or antileptons, for example, in collisions of protons with matter at high-energy accelerators, requires enough energy to create not only the antiparticle but also a corresponding particle such that total baryon or lepton number and charge can be conserved in the process. *See* ELECTRON-POSITRON PAIR PRODUCTION.

Since mesons do not possess baryon or lepton number, only charge, energy, and angular momentum need be conserved in their production. Thus, a process such as a collision of a proton with a proton can produce a single neutral pi meson, for example, $p = p \rightarrow p = p = \pi^0$. Other quantum numbers, such as strangeness and charm, must be conserved if production of mesons possessing these quantum numbers is to proceed through strong or electromagnetic interactions. In these cases a particle with the negative values of the particular quantum number must also be produced. Such a process is termed associated production. *See* CHARM; QUANTUM NUMBERS; STRANGE PARTICLES.

Matter-antimatter oscillations. Isolated neutral particles, notably K^0 and B^0 mesons, can spontaneously transform into their antiparticles via the weak interaction. These quantum-mechanical phenomena are termed K–$\bar{K}$ or B–$\bar{B}$ mixing, respectively. Mixing can lead to particle-antiparticle oscillations wherein a K^0 can become its antiparticle, a $\bar{K}^0$, and later oscillate back to a K^0. It was through this phenomenon that observation of *CP* violation first occurred. That observation, coupled to the *CPT* theorem, implies that physics is not exactly symmetric under time reversal, for example, that the probability of a K^0 becoming a $\bar{K}^0$ is not exactly the same as that in the reverse process. *See* WEAK NUCLEAR INTERACTIONS.

The possibility of neutron-antineutron ($n\bar{n}$) oscillations also exists if baryon number is not strictly conserved. While there are theories which allow, and even require, baryon number to be violated, sensitive searches for $n\bar{n}$ oscillations have found none, and place lower limits on oscillation times of $\tau_{n\bar{n}} \geq 5 \times 10^7$ s.

Antimatter in cosmology. Experimental observations, both ground- and balloon-based, indicate that the number of cosmic ray antiprotons is less than 1/10,000 that of protons. This number is consistent with the antibaryon production that would be expected from collisions of cosmic protons with the Earth's atmosphere, and is consistent with the lack of appreciable antimatter in the Milky Way Galaxy. Clouds of positrons have been observed at the center of the Galaxy. Their origin is unclear, but it is suspected that they arise from massive star formation near a large black hole, explosion of massive stars, or other such phenomena, rather than from a primordial source. *See* COSMIC RAYS.

Attempts to find antimatter beyond the Milky Way involve searches for gamma radiation resulting from matter-antimatter annihilation in the intergalactic gas that exists between galactic clusters. The null results of these searches suggests that at least the local cluster of galaxies consists mostly of matter.

If matter dominates everywhere in the universe, a question arises as to how this came to be. In the standard model of cosmology, the big bang model, the initial condition of the universe was that the baryon number was zero; that is, there was no preference of matter over antimatter. The current theory of how the matter-antimatter asymmetry evolved requires three ingredients: interactions in which baryon number is violated, time reversal (or *CP*) violation, and a lack of thermodynamic equilibrium. The last requirement was satisfied during the first few microseconds after the big bang. Time reversal violation has been observed in the laboratory in K^0 decays, albeit perhaps not of sufficient size to explain the observed baryon-antibaryon asymmetry. But the first ingredient, baryon number violation, has not yet been observed in spite of sensitive searches. Thus, the origin of the dominance of matter over antimatter remains an outstanding mystery of particle and cosmological physics. *See* BIG BANG THEORY; COSMOLOGY; THERMODYNAMIC PROCESSES.

Antiprotons in accelerators. Antiprotons are stable against decay, just as are their proton counterparts. Thus, after their production in high-energy proton-matter collisions they can be cooled, collected, and stored for an unlimited amount of time. The only condition on their continued existence is that they

not come in contact with matter, where they would annihilate. Stored antiprotons have been accelerated to high energies and brought into head-on collision with similarly high-energy protons at particle accelerators at both CERN, near Geneva, Switzerland, and Fermilab, near Batavia, Illinois. Experiments with proton-antiproton collisions led to the first observation of *W* and *Z* bosons, the mediators of the weak force, and to the discovery of the top quark, the heaviest known elementary particle with a mass of about 185 proton masses. *See* INTERMEDIATE VECTOR BOSON; PARTICLE ACCELERATOR.

Antihydrogen atoms. Antihydrogen atoms, composed of an antiproton and a positron, have been created in high-energy particle physics experiments. This was accomplished by passing a stream of high-energy antiprotons through a gas jet of xenon atoms at CERN and protons at Fermilab. Occasionally a photon was created in the interaction between particles in the jet and the antiprotons. Sometimes this photon had sufficient energy to produce a subsequent electron-positron pair, and the positron became attached to the antiproton. Only a few atoms were created by this method, and they had high velocities in the laboratory.

A method of forming antihydrogen atoms with low velocities in the laboratory is to cool, decelerate, and trap antiprotons, and bring similarly cooled positrons into their proximity. Some of the positrons will be captured by the antiprotons, forming thousands of antihydrogen atoms. *See* PARTICLE TRAP.

Antiparticle nuclei. Antiparticle nuclei are difficult to form in the laboratory because their constituent antiprotons and antineutrons must be created in high-energy collisions of projectile and target particles. While several antiparticles can be created in these collisions, the probability that they will have small enough relative velocities to form themselves into nuclei without first annihilating in the rest of the debris of the collision is small. The greater the number of constituents, that is, the larger the nuclei, the smaller is the probability. Nonetheless, nuclei as large as antihelium have been observed in high-energy proton collisions with matter. Because of the difficulty of forming heavy antinuclei from particle collisions, the observation of such nuclei in cosmic rays would imply that their origin was from stars completely composed of antimatter.

Because of the short duration of balloon flights, balloon-based experiments that search for antimatter above the Earth's atmosphere lack the ability to systematically study the nature of that antimatter which is found. These experiments have never found heavier antimatter particles than antiprotons. Experiments with longer duration of observation and more sophisticated experimental instrumentation, for example, mounted on the International Space Station, can have greater sensitivity to antimatter in cosmic rays. Observation of antihelium would be indicative of primordial antimatter that survived since the big bang, and that of heavier antinuclei such as anticarbon would imply the existence of antimatter stars. *See* SPACE STATION.

Michael E. Zeller

Bibliography. Committee on Elementary Particle Physics, *Elementary Particle Physics: Revealing the Secrets of Energy and Matter*, National Research Council, National Academy Press, 1998; G. Gabrielse, Extremely cold antiprotons, *Sci. Amer.*, 267(6):78–89, December 1992; D. Perkins, *Introduction to High Energy Physics*, 3d ed., Addison-Wesley, Menlo Park, CA, 1987; *Readings from Scientific American: Particles and Fields*, Scientific American, New York, 1980; C. Schwartz, *A Tour of the Subatomic Zoo: A Guide to Particle Physics*, American Institute of Physics, New York, 1992; G. Tarlé and S. Swordy, Cosmic antimatter, *Sci. Amer.*, 278(4):36–44, April 1998; S. Weinberg, *The First Three Minutes*, 2d ed., Basic Books, New York, 1993.

Antimicrobial agents

Chemical compounds biosynthetically or synthetically produced which either destroy or usefully suppress the growth or metabolism of a variety of microscopic or submicroscopic forms of life. On the basis of their primary activity, they are more specifically called antibacterial, antifungal, antiprotozoal, antiparasitic, or antiviral agents. Antibacterials which destroy are bactericides or germicides; those which merely suppress growth are bacteriostatic agents. *See* FUNGISTAT AND FUNGICIDE.

Of the thousands of antimicrobial agents, only a small number are safe chemotherapeutic agents, effective in controlling infectious diseases in plants, animals, and humans. A much larger number are used in almost every phase of human activity: in agriculture, food preservation, and water, skin, and air disinfection. A compilation of some common uses for antimicrobials is shown in **Table 1**.

Almost 5000 years before L. Pasteur (1870) clearly enunciated the germ theory of disease, plant drugs were described by Shen Nung in *Pen Tsao* (*The Great Herbal*). The Ebers Papyrus about 1500 B.C. prescribed the use of onions and other plants for the cure of septic wounds. Through the Middle Ages the search for effective antimicrobials from plants continued, providing a major stimulus for the development of systematic botany. The herbals of this time gradually replaced superstition and incantations with precise descriptions of both the plant and the proper method of its preparation for therapeutic use. Three contemporary drugs, emetine, quinine, and chaulmoogra oil, were used as early as the seventeenth century as crude extracts of the ipecacuanha plant, cinchona bark, and *Taraktogenos kurzii*, respectively. During the next two centuries the herbal and flora evolved into the modern pharmacopoeia. A survey by E. M. Osborn (1943) of 2300 species of plants revealed that 440 species produced antimicrobial agents.

Pasteur, R. Koch, P. Ehrlich, and Lord Lister founded modern microbiology during the last three decades of the nineteenth century. Pasteur and J. Joubert (1877) discovered the phenomenon of microbial antagonism. Pyocyanase was purified from culture filtrates of *Pseudomonas aeruginosa* by

TABLE 1. Common antimicrobial agents and their uses

Use	Agents
Chemotherapeutics (animals and humans)	
Antibacterials	Sulfonamides, isoniazid, *p*-aminosalicylic acid, penicillin, streptomycin, tetracyclines, chloramphenicol, erythromycin, novobiocin, neomycin, bacitracin, polymyxin
Antiparasitics (humans)	Emetine, quinine
Antiparasitics (animal)	Hygromycin, phenothiazine, piperazine
Antifungals	Griseofulvin, nystatin
Chemotherapeutics (plants)	Captan (*N*-trichlorothio-tetrahydrophthalimide), maneb (manganese ethylene bisdithiocarbamate), thiram (tetramethylthiuram disulfide)
Skin disinfectants	Alcohols, iodine, mercurials, silver compounds, quaternary ammonium compounds, neomycin
Water disinfectants	Chlorine, sodium hypochlorite
Air disinfectants	Propylene glycol, lactic acid, glycolic acid, levulinic acid
Gaseous disinfectants	Ethylene oxide, β-propiolactone, formaldehyde
Clothing disinfectants	Neomycin
Animal-growth stimulants	Penicillin, streptomycin, bacitracin, tetracyclines, hygromycin
Food preservatives	Sodium benzoate, tetracycline

R. Emmerich and O. Low (1899) and became the first microbial antagonist to be extensively investigated as a chemotherapeutic agent. Although its action was thought to combine antimicrobial and immunizing activities, the discovery, isolation, purification, pharmacology, and clinical evaluation of pyocyanase established a pattern of investigation employed with the sulfonamides, penicillin, and all subsequently discovered antimicrobials.

Lister employed phenols and cresols as surface germicides to convert the hospitals of the late nineteenth century from pesthouses into clean, safe quarters for surgery and childbirth with a reduced hazard from cross infection. His practice is followed today not only in hospitals but in homes, food establishments, research laboratories, and wherever disinfection is essential to safe techniques. *See* ANTISEPTIC.

At the turn of the century Ehrlich, from his observations on the specific staining reactions of dyes upon certain tissue cells, conceived of highly specific dyelike molecules, "magic bullets" whose germicidal action would specifically destroy microbes without killing cells of the host. An extension of Ehrlich's approach by G. Domagk led to the discovery in 1935 of sulfonamido-crysoidin (Prontosil), the first synthetic antimicrobial of broad clinical usefulness. J. Tréfouël and coworkers quickly established that sulfanilamide was the active moiety of the Prontosil molecule and sulfanilamide was soon modified into a series of useful derivatives, comprising the sulfa drugs, which vary in their usefulness and types of side effects. *See* CHEMOTHERAPY.

Erwin Smith (1900) established that plant diseases were as common as those of animals. Antimicrobial therapy, however, has not been equally effective. The synthetic captans and manebs are most widely used. Antibiotics, although broadly tested, have not been so useful in the control of plant diseases.

The most important antimicrobial discovery of all time, that of the chemotherapeutic value of penicillin, was made after hope of developing its clinical usefulness had been abandoned. As E. Chain (1954) recalls this dramatic period:

"Work on penicillin was begun about one year before the war (1938); it was thus not stimulated, as is so often stated, by wartime demands for new and effective methods of treating infected wounds. It is perhaps important to emphasize that the decision to reinvestigate the properties of penicillin was not motivated by the hope of developing the clinical application of an antibacterial substance which had been shown to have great practical possibilities, but for some incomprehensible reasons had been overlooked for nine years. . . . The investigation was undertaken essentially as a biochemical problem, with the aim of establishing the chemical nature of a natural, very unstable, antibacterial product active against the staphylococcus."

The subsequent establishment of penicillin as a nontoxic drug with 200 times the activity of sulfanilamide opened the flood gates of antibiotic research. In the next 20 years, more than a score of new and useful microbially produced antimicrobials entered into daily use. *See* ANTIBIOTIC.

New synthetic antimicrobials are found today, as in Ehrlich's day, by synthesis of a wide variety of compounds, followed by broad screening against many microorganisms. Biosynthetic antimicrobials, although first found in bacteria, fungi, and plants, are now being discovered primarily in actinomycetes.

S. Waksman (1943) was the first to recognize the

TABLE 2. Functional groups of some antimicrobial agents

Structural feature	Agents
Amino sugar	Streptomycin, neomycin, kanamycin
Polyene	Amphotericin B, nystatin, trichomycin
Polypeptide	Polymyxins, circulin, colistin, thiactin
Lactam	Penicillin
Diazo	Azaserine, DON
Coumarin	Novobiocin
Macrolide	Erythromycin, filipin, carbomycin
Phenazine	Griseolutein, echinomycin
Tetracyclic	Chlortetracycline, oxytetracycline, tetracycline, netropsin, grisein
Pyrrole	Fradicin, prodigiosin
Quinone	Xanthomycin A, cyanomycin, chartreusin
Thio-	Actithiazic acid, aureothricin, celesticetin
Alkaloid	Griseomycin
Nucleoside	Psicofuranine, puromycin, nebularine
Nitro-	Chloramphenicol, azomycin
Acetylenic	Mycomycin, nemotins
Spirane	Griseofulvins, helvolic acid
Tropolone	Puberulic acid, lactaroviolin

prolific antibiotic productivity of the actinomycetes. His rapid discovery of actinomycin, streptothricin, streptomycin, and neomycin inspired the intensive study of this group. Approximately three-quarters of the known antibiotics and the same fraction of all useful antibiotics are produced by actinomycetes, which live almost exclusively in the soil. Despite their habitat, a striking feature of actinomycetes is that to date it has been impossible to demonstrate the production of any antibiotics in the soil under natural growth conditions.

Antimicrobial agents contain various functional groups (**Table 2**). No particular structural type seems to favor antimicrobial activity. The search for correlation of structure with biological activity goes on, but no rules have yet appeared with which to forecast activity from contemplated structural changes. On the contrary, minor modifications may lead to unexpected loss of activity. George M. Savage

Bibliography. D. Gottlieb and P. D. Shaw (eds.), *Antibiotics*, 2 vols., 1967; S. Hammond, *Antibiotics and Antimicrobial Action*, 1978; B. M. Kagan, *Antimicrobial Theory*, 3d ed., 1980.

Antimony

The element is dimorphic, existing as a yellow, metastable form composed of Sb_4 molecules, as in antimony vapor and the structural unit in yellow antimony; and a gray, metallic form, which crystallizes with a layered rhombohedral structure. Antimony

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

differs from normal metals in having a lower electrical conductivity as a solid than as a liquid (as does its congener, bismuth). Metallic antimony is quite brittle, bluish-white with a typical metallic luster, but a flaky appearance. Although stable in air at normal temperatures, it burns brilliantly when heated, with the formation of a white smoke of Sb_2O_3. Vaporization of the metal gives molecules of Sb_4O_6, which break down to Sb_2O_3 above the transition temperature. Some of the more significant properties of atomic and bulk antimony are given in the **table**.

Antimony occurs in nature mainly as Sb_2S_3 (stibnite, antimonite); Sb_2O_3 (valentinite) occurs as a decomposition product of stibnite. Antimony is commonly found in ores of copper, silver, and lead. The

Some atomic and bulk properties of antimony

Property	Value
Atomic weight	121.75
Electron configuration (S^4 ground state)	$[Kr]4d^{10}5s^25p^3$
Covalent radius	141 pm
Ionic radius (Sb^{3+})	90 pm
Metallic radius	159 pm
Ionization energies, 1st–6th in kJ mole^{-1}	834, 1592, 2450, 4255, 5400, 10,420
Electrode potential, Sb^{3+}/Sb	0.21 V
Electronegativity (Allred-Rochow)	1.82
Oxidation numbers	−3, 0, +3, +5
Specific gravity	6.691
Melting point	630.8°C (1167.4°F)
Boiling point	1753.01°C (3187.42°F)
Electrical resistivity	39.0 μohm cm (273 K)
Toxicity level	0.5 mg·m^{-3} of air

metal antimonides NiSb (breithaupite), NiSbS (ullmannite), and Ag_2Sb (dicrasite) also are found naturally; there are numerous thioantimonates such as Ag_3SbS_3 (pyrargyrite).

Antimony is produced either by roasting the sulfide with iron, or by roasting the sulfide and reducing the sublimate of Sb_4O_6 thus produced with carbon; high-purity antimony is produced by electrolytic refining.

Commercial-grade antimony is used in many alloys (1–20%), especially lead alloys, which are much harder and mechanically stronger than pure lead; batteries, cable sheathing, antifriction bearings, and type metal consume almost half of all the antimony produced. The valuable property of Sn-Sb-Pb alloys, that they expand on cooling from the melt, thus enabling the production of sharp castings, makes them especially useful as type metal. John L. T. Waugh

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook of Chemistry and Physics*, 85th ed., CRC Press, 2004; D. F. Shriver and P. W. Atkins, *Inorganic Chemistry*, 3d ed., 1999.

Antioxidant

A substance that, when present at a lower concentration than that of the oxidizable substrate, significantly inhibits or delays oxidative processes, while being itself oxidized. In primary antioxidants, such as polyphenols, this antioxidative activity is implemented by the donation of an electron or hydrogen atom to a radical derivative, and in secondary antioxidants by the removal of an oxidative catalyst and the consequent prevention of the initiation of oxidation.

Antioxidants have diverse applications. They are used to prevent degradation in polymers, weakening in rubber and plastics, autoxidation and gum formation in gasoline, and discoloration of synthetic and natural pigments. They are used in foods, beverages, and cosmetic products to inhibit deterioration and spoilage. Interest is increasing in the application of antioxidants to medicine relating to human diseases

attributed to oxidative stress. This article aims to summarize the mechanisms by which antioxidants prevent lipid autoxidation; to present some products formed as a result of oxidation, some methods available for their detection, and the structure of some synthetic and natural antioxidants; and to emphasize the role of antioxidants in biological systems, including diseases associated with their absence.

Lipid autoxidation. The autoxidation process is shown in reactions (1)–(3). Lipids, mainly those con-

$$\text{RH} + \text{initiator (L)} \longrightarrow \text{R}\cdot + \text{LH} \quad (1)$$

$$\text{R}\cdot + \text{O}_2 \longrightarrow \text{ROO}\cdot \quad (2)$$

$$\text{ROO}\cdot + \text{RH} \longrightarrow \text{ROOH} + \text{R}\cdot \quad (3)$$

taining unsaturated fatty acids, such as linoleic acid [RH in reaction (1)], can undergo autoxidation via a free-radical chain reaction, which is unlikely to take place with atmospheric oxygen (ground state) alone. A catalyst (L) is required, such as light, heat, heavy-metal ions (copper or iron), or specific enzymes present in the biological system [reaction (1)]. The catalyst allows a lipid radical to be formed (alkyl radical R·) on a carbon atom next to the double bond of the unsaturated fatty acid. This radical is very unstable and reacts with oxygen [reaction (2)] to form a peroxyl radical (ROO·), which in turn can react with an additional lipid molecule to form a hydroperoxide [ROOH in reaction (3)] plus a new alkyl radical, and hence to start a chain reaction. Reactions (2) and (3), the propagation steps, continue unless a decay reaction takes place (a termination step), which involves the combination of two radicals to form stable products. *See* AUTOXIDATION; CATALYSIS; CHAIN REACTION (CHEMISTRY).

When lipid autoxidation occurs in food, it can cause deterioration, rancidity, bad odor, spoilage, reduction in nutritional value, and possibly the formation of toxic by-products. Oxidation stress in a lipid membrane in a biological system can alter its structure, affect its fluidity, and change its function. A number of assays are available for the detection of lipid autoxidation, most of them based on the detection of intermediate products, such as a lipid hydroperoxide (ROOH), or conjugated double bonds (each double bond separated by one single bond), formed during the oxidation of the unsaturated fatty acid. Other methods of detection are direct measurement of the intermediate free radicals formed (electron spin resonance methods) or the measurement of fatty acid degradation products, such as ketones and aldehydes (thiobarbituric assay). *See* FREE RADICAL.

In oxidative stress the antioxidant can eliminate potential initiators of oxidation and thus prevent reaction (1). It can also stop the process by donating an electron and reducing one of the radicals in reaction (2) or (3), thus halting the propagation steps. A primary antioxidant can be effective if it is able to donate an electron (or hydrogen atom) rapidly to a lipid radical and itself become more stable then the original radical. The ease of electron donation depends on the molecular structure of the antioxidant, which dictates the stability of the new radical; and it may be predicted by measuring the oxidation potential of the molecule, by the use of computer software, or by calculating quantum parameters of the molecule. Many naturally occurring polyphenols, such as flavonoids, anthocyanins, and saponins, which can be found in wine, fruit, grain, vegetables, and almost all herbs and spices, are effective antioxidants that operate by this mechanism. This type of antioxidant action is shown in reactions (4) and (5), with AH given as a general

$$\text{AH} + \text{ROO}\cdot \longrightarrow \text{A}\cdot + \text{ROOH} \quad (4)$$

$$\text{ROO}\cdot + \text{InH} \rightleftarrows \text{ROOH} + \text{In}\cdot \quad (5)$$

antioxidant and α-tocopherol (vitamin E) as an example. *See* VITAMIN E.

There are a number of different mechanisms by which antioxidants can prevent reaction (1) from taking place in biological and nonbiological systems. These mechanisms can be absorbing by ultraviolet light, scavenging oxygen, chelating transition metals, or inhibiting enzymes involved in the formation of reactive oxygen species, for example, NADPH oxidase and xanthine oxidase (reducing molecular oxygen to superoxide and hydrogen peroxide), dopamine-β-hydroxylase, and lipoxygenases. The common principle of action in the above examples is the removal of the component acting as the catalyst that initiates and stimulates the free-radical chain reaction. The structures of some antioxidants, synthetic and natural, are shown below. *See* ENZYME.

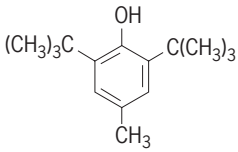

Butylated hydroxy toluene

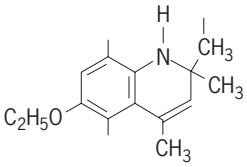

Ethoxyquin

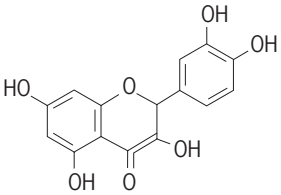

Quercetin

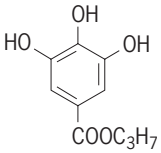

Propyl gallate

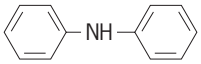

Diphenylamine

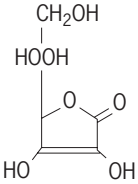

Ascorbic acid (vitamin C)

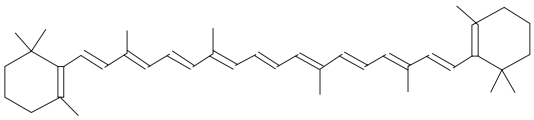
β-Carotene

Among antioxidants, the synthetic compounds butylated hydroxyanisole (BHA), propyl gallate, ethoxyquin, and diphenylamine are commonly used as food additives. Quercetin belongs to a large natural group of antioxidants, the flavonoid family, with more than 6000 known members, many acting through both mechanisms described above. Ascorbic acid is an important water-soluble plasma antioxidant; it and the tocopherols, the main lipid soluble antioxidants, represent the antioxidants in biological systems. β-Carotene belongs to the carotenoid family, which includes lycopene, the red pigment in tomatoes; the family is known to be very effective in reacting with singlet oxygen ($^1O_2$), a highly energetic species of molecular oxygen. *See* ASCORBIC ACID; CAROTENOID; FLAVONOID; FOOD PRESERVATION.

Role in human body. Antioxidants play an important role in the human body. Each year an increasing number of diseases are attributed to oxidative stress, involving radicals and oxidants. In any biological system, the balance between the formation of reactive oxygen species and their removal is vital. New ones are formed regularly, either as a result of normal functions of the organs or of oxidative stress. Thus, superoxide ($O_2^{\bullet -}$), hydrogen peroxide (H_2O_2), and hydroxyl radical (HO) are mutagenic compounds and also by-products of normal metabolic pathways in the organs. Superoxide, which is the most important source of initiating radicals in vivo, is produced or secreted regularly from such normal cellular tissue as the mitochondria or from macrophage cells. To maintain balance, the system protects itself from the toxicity of reactive oxygen species in different ways, including the formation of antioxidants of various structures and functions. Human plasma contains a number of different enzymatic antioxidants, such as catalase, superoxide dismutase, glutathione peroxidase, reductase, and transferase, which degrade reactive oxygen species and thus protect circulating lipids, proteins, and cells from oxidative damage. Nonenzymatic protein antioxidants found in plasma, such as transferrin and albumin, are able to chelate transition metals and eliminate their catalytic effect. Other protective antioxidants are the low-molecular-weight compounds, such as vitamin C, glutathione, bilirubin, uric acid (all water-soluble), vitamin E, β-carotene, lycopene, ubiquinol-10 (all lipid-soluble), and polyphenols, found in plasma and extracellular and intracellular fluids, and in lipoproteins and membranes.

When the oxidation/reduction balance is altered toward higher oxidative stress, with more reactive oxygen species and less protecting elements, cell damage, tissue injury, and specific diseases result. Among the disorders and effects on organs associated with the proliferation of free radicals and the violation of the oxidation/reduction balance are the oxidation of lipids and proteins, with changes in their structure and function; and damage of deoxyribonucleic acid (DNA) and its contribution to the onset of cancer, cardiovascular diseases, reperfusion injury, cataracts, neurological disorders, and lung diseases. The aging process is also thought to result from the constant exposure of the organs to reactive oxygen species, with concomitant and continuous infringement of this balance, and with cumulative damage resulting from a gradual decrease in repair capacity and an increase in degenerative changes in the organs. *See* OXIDATION-REDUCTION; OXYGEN TOXICITY; SUPEROXIDE CHEMISTRY. Jacob Vaya; Lester Packer

Bibliography. S. Baskin and H. Salem (eds.), *Oxidants, Antioxidants, and Free Radicals*, Taylor & Francis, 1997; E. Cadenas and L. Packer (eds.), *Handbook of Antioxidants (Oxidative Stress and Disease)*, 2d ed., Marcel Dekker, New York, 2001; B. Frei (ed.), *Natural Antioxidants in Human Health and Disease*, Academic Press, New York, 1994; B. Halliwell and J. M. C. Gutteridge, *Free Radicals in Biology and Medicine*, Oxford University Press, 1999; J. Pokorny, *Antioxidants in Foods: Practical Applications*, CRC Press, 2001.

Antipatharia

An order of the cnidarian subclass Hexacorallia. These animals are the black or horny corals which live in rather deep tropical and subtropical waters and usually form regular or irregularly branching

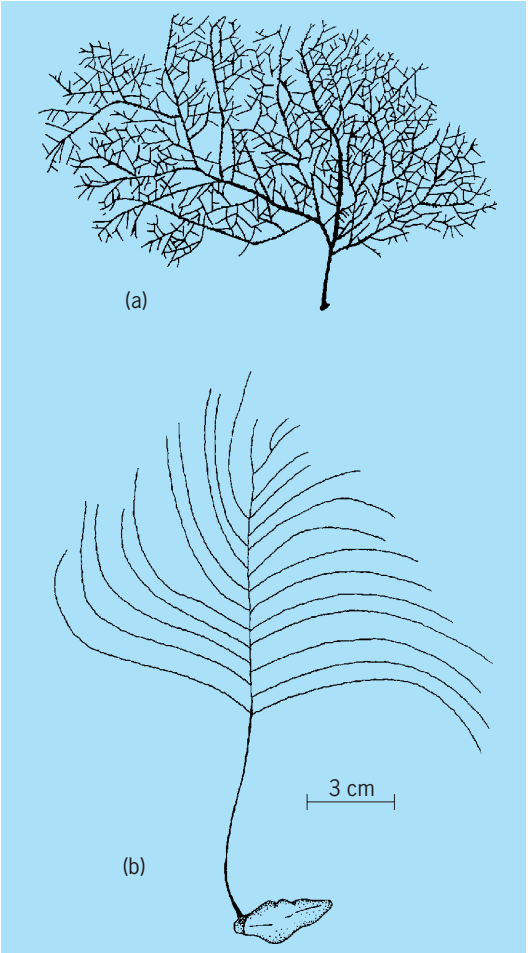

Fig. 1. Antipatharians. (*a*) *Antipathes rhipidion*. (*b*) *Bathypathes alternata*.

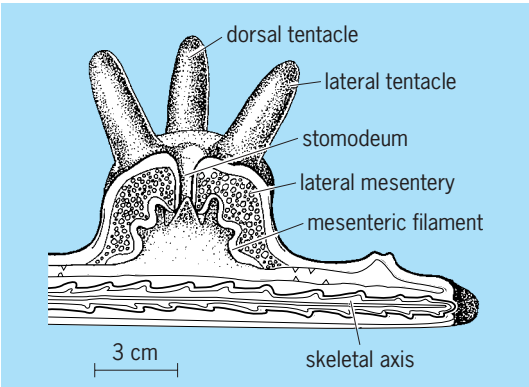

Fig. 2. Antipatharian polyp cut in vertical section.

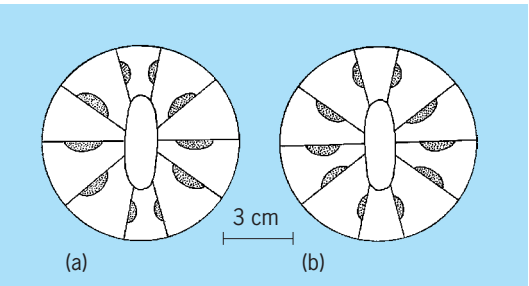

Fig. 3. Mesenteries of (*a*) *Stichopathes ceylonensis* and (*b*) *Antipathes longibrachiata*.

plantlike colonies, often 7 to 10 ft (2 to 3 m) in height, with thorny, solid lamellar, horny axial skeletons (**Fig. 1**). *Stichopathes* forms an unbranching wirelike colony.

The polyp or zooid (**Fig. 2**) has six unbranched, nonretractile tentacles with a warty surface due to the presence of nematocysts. Six primary, complete, bilaterally arranged mesenteries (**Fig. 3**) occur, of which only two lateral ones bear filaments and gonads. *Dendrobranchia*, however, has eight retractile pinnate tentacles. Both edges of the stomodeum appear to be differentiated into siphonoglyphs morphologically, but not histologically. Adjacent zooids are united by a coenenchyme, but their gastrovascular cavities have no connection. The musculature is the most weakly developed in the anthozoans.

The polyps are dioecious. Schizopathidae are dimorphic; the gastrozooid has a mouth and two tentacles, while the gonozooid, the only fertile polyp, lacks a mouth. *See* CNIDARIA; HEXACORALLIA.

Kenji Atoda

Antiresonance

The condition for which the impedance of a given electric, acoustic, or dynamic system is very high, approaching infinity. In an electric circuit consisting of a capacitor and a coil in parallel, antiresonance occurs when the alternating-current line voltage and the resultant current are in phase. Under these conditions the line current is very small because of the high impedance of the parallel circuit at antiresonance. The branch currents are almost equal in magnitude and opposite in phase.

The principle of antiresonance is used in wave traps, which are sometimes inserted in series with antennas of radio receivers to block the flow of alternating current at the frequency of an interfering station, while allowing other frequencies to pass. Wave traps are used in electric power substations to separate circuits for power-line carrier communications. *See* ANTENNA (ELECTROMAGNETISM); RESONANCE (ALTERNATING-CURRENT CIRCUITS).

John Markus

Antiseptic

A drug used to destroy or prevent the growth of infectious microorganisms on or in the human or animal body, that is, on living tissue. The legal definition is stated in the Federal Food, Drug, and Cosmetic Act, Chap. II, Sect. 201, Para. (o) as follows: "The representation of a drug, in its labeling, as an antiseptic shall be considered to be a representation that it is a germicide, except in the case of a drug purporting to be, or represented as, an antiseptic for inhibitory use as a wet dressing, ointment, dusting powder, or such other use as involves prolonged contact with the body." This means that antiseptics will render microorganisms innocuous, by either killing them or preventing their growth, according to the character of the preparation or the method of application.

Antiseptics have been in use for well over a century, first empirically in the prevention of puerperal sepsis and then more specifically in antiseptic surgery. During this period many chemical substances have been employed as antiseptics, such as certain halogens (iodine), mercurial compounds, essential oils, silver compounds, alcohols, bisphenols, and quaternary ammonium compounds. *See* HALOGEN ELEMENTS.

Iodine. This is the most important of the halogens used as an antiseptic. Although iodine solutions were used as an external application in wounds in 1839 and were admitted to the U.S. Pharmocopeia (USP) in 1840, they were not widely used in surgery until after 1860. Tincture of iodine (iodine in an alcohol solution) has been employed widely as a preoperative antiseptic and in first aid. Tincture of iodine is germicidal by laboratory test in 0.02% concentration but 2.0% solutions are usually employed in surgery and first aid. The official solutions are iodine tincture, USPXV, which consists of 2.0% iodine and 2.4% sodium iodide in alcohol; iodine tincture, strong, NFX (National Formulary X), 7% iodine and 5.0% potassium iodide in alcohol; and iodine solution, NFX, 2.0% iodine and 2.4% sodium iodide in water. The tincture is a better antiseptic than the water solution because it penetrates rapidly, is highly germicidal, and is nonspecific in antibacterial activity. Also, the tincture is not appreciably counteracted by organic matter and has a low surface tension, a low pH (which further enhances its germicidal activity) and, in usable solutions, a high degree of antimicrobial effectiveness.

Mercurial compounds. Compounds of mercury were used to prevent infection before the germ theory of disease was established, and were widely used in surgery following the work of Robert Koch on mercury bichloride in 1881. Although Koch appeared to prove that this mercurial could kill spores, the test was such that the results were actually the effect of bacteriostatic, not germicidal or sporicidal, action. In spite of this erroneous evaluation, mercury bichloride became the antiseptic of choice in place of the phenol solution used by Joseph Lister; it was employed by the professions and laypeople for many years, and still has some applications. Because of their high toxicity and severe caustic action, such inorganic mercurials as mercuric chloride, mercuric oxycyanide, and potassium mercuric iodide have been largely replaced by certain organic mercury compounds. Organic mercurial compounds are far less toxic and are nonirritating in concentrated solutions. They are highly bacteriostatic, and in concentrated solutions germicidal as well. They are also nonspecific in antimicrobial activity. The organic mercurials as well as the inorganic salts are readily counteracted by organic matter and have a high surface tension in water solutions. However, tinctures of organic mercurials (that is, solutions in alcohol) are more germicidal and, because of their lower surface tensions, are able to penetrate into tissue crevices. Tinctures of organic mercurials are effective in practice as a result of both germicidal and bacteriostatic activity. The most important organic mercurials are phenylmercuric nitrate and acetate; merbromin NF (mercurochrome), the sodium salt of dibromohydroxymercurifluorescein; thimerosal NF (merthiolate), the sodium salt of ethyl mercury thiosalicylic acid; nitromersol (metaphen), the anhydride of 4-nitro-3-hydroxymercuri-*o*-cresol; and mercresin, an alcohol-acetone solution containing a mixture of secondary amyltricresols and orthohydroxyphenylmercuric chloride.

Essential oils. Essential oils have been defined as odoriferous oily substances obtained from such natural sources as plants by steam distillation. Essential oils in alcoholic solutions also were early employed in place of the carbolic acid solution of Lister, and because of the toxic and corrosive action of mercury bichloride, they also replaced this compound. Alcoholic solution of essential oils was first developed in 1881 and was admitted as liquor antisepticus to the U.S. Pharmacopeia in 1900 and to the National Formulary IV in 1916. It was still included in the National Formulary X. While thymol is the most active ingredient in this formulation, the other essential oils included act in a synergistic manner; that is, the total effect is greater than the additive effect of the individual components. Although these essential oils are present in low concentration, the final combination is highly germicidal, in fact, equal in germicidal activity to the 2.5% solution of carbolic acid so successfully used by Lister and others in antiseptic surgery. This antiseptic is also nonspecific in its germ-killing property and kills all varieties of infectious microorganisms, with the exception of spores. Its low surface tension facilitates spreading and penetration, and its low pH enhances germicidal activity. The antiseptic is not counteracted by organic matter, is stable under all conditions of storage, and is nontoxic, noncaustic, and safe to use on all body tissues. For these reasons, the alcoholic solutions of essential oils as represented by liquor antisepticus have proved effective in a wide variety of clinical applications and in first aid.

Silver compounds. These compounds have been employed as antiseptics since 1884 and have been widely used for a variety of purposes. Because of the bland nature of most of these compounds, they have been successfully used in the eyes, nose, throat, urethral tract, and other organs. The most widely used silver compounds are silver nitrate, USP, ammoniacal silver nitrate solution, NF, silver picrate, NNR (New and Nonofficial Remedies), and certain colloidal silver preparations such as strong protein silver, NF, and mild silver protein, NF (Argyrol). These silver compounds are effective germicides of low tissue toxicity and are not counteracted by organic matter.

Alcohols. Such compounds as ethyl alcohol and isopropyl alcohol have been used as antiseptics since 1894 and are still widely employed. They are germicidal rather than bacteriostatic and are effective against the vegetative forms of bacteria and virus, but do not kill spores. Ethyl alcohol in 62.5–70% solution is most commonly used, being widely employed for disinfecting the skin before hypodermic injections and other skin punctures. It is rapidly germicidal, nonspecific in germicidal activity, nontoxic, nonirritating, and not counteracted by organic matter. Isopropyl alcohol is equal, if not superior, to ethyl alcohol and is widely used for degerming the skin and for disinfecting oral thermometers. Alcohols are also widely used in other antiseptic preparations, in which they serve to lower the surface tension and to promote spreading and penetration.

Bisphenols. Compounds such as dichlorophene and tetrachlorophene are essentially bacteriostatic agents and are weaker as germicides. The antibacterial activity of bisphenols was first demonstrated in 1906, and much new information has been developed since 1927. They have proved quite effective as skindegerming agents, when used in soaps and other detergents, and as mildew-preventing formulations. The halogenated form, such as dichlorophene, tetrachlorophene, hexachlorophene, and bithionol, is most commonly employed. When used repeatedly on the skin, as in soaps and detergents, bisphenols have a tendency to remain for long periods, thus reducing skin bacteria to a significant degree. For this purpose they are especially useful in preoperative hand washing. Bisphenols are counteracted to a variable degree by organic matter.

Quaternary ammonium compounds. These organic pentavalent nitrogen compounds have high germicidal activity. Although they are more properly classified as surface-active disinfectants, some of them are employed in certain antiseptic formulations especially suited for use on the skin and for mucous surfaces. They are counteracted to a certain extent

by organic matter, especially by blood serum. Nontoxic and nonirritating, they may be used in place of alcohol after preoperative scrub-up.

Comparison of antiseptics. It is difficult, if not impossible, to evaluate the relative merits of the various classes of antiseptics under all conditions of use. However, such an effort has been made with respect to reduction of skin bacteria in the serial-basin test. The results substantiate the efficiency of iodine solutions. *See* BIOASSAY; FUNGISTAT AND FUNGICIDE.

George F. Reddish; Francis Clark

Bibliography. American Pharmaceutical Association, *National Formulary*, 12th ed., 1965; *Pharmocopeia of the United States*, 20th ed., 1980; G. F. Reddish (ed.), *Antiseptics*, *Disinfectants*, *Fungicides*, *and Chemical and Physical Sterilization*, 2d ed., 1957.

Antisubmarine warfare

All measures required to combat enemy submarines, including strategy, operational employment of forces, tactics, and a wide variety of equipment and weapons to find and destroy submarines and to neutralize their weapons.

Goals. The key physical facts of antisubmarine warfare (ASW) are that submerged submarines are, effectively, invisible to surface observers and that sound is the only form of energy that travels any substantial distance underwater. All forms of antisubmarine warfare are based on strategies which cause the submarines to surrender either their invisibility (or stealth) or their mobility. Generally, since the submarine can be destroyed once it is located, the effort in antisubmarine warfare goes primarily into finding it.

Whereas in the past submarines were important only as antiship weapons, some can now deliver strategic attacks. Immobilization has little effect on a strategic submarine capable of firing missiles with intercontinental range; it is an open question as to what extent antisubmarine warfare measures can be effective against such craft. Because the relative invulnerability of the strategic submarines is widely considered to be an important element of strategic stability, it is also debatable to what extent antisubmarine warfare measures directed against them are desirable. *See* BALLISTIC MISSILE.

Revised strategic emphasis. With the end of the Cold War, the emphasis has changed, at least for the major sea powers, from maintaining sea control to expeditionary operations in littoral areas, which are often relatively shallow. In the past, sea control might have been considered statistical; that is, it would have been acceptable to trade numbers of merchant ships for the elimination of an enemy submarine threat. Now the war is likely to be concentrated in an enemy's home waters, and all of the ships of the expeditionary force are quite valuable. Losses are far less acceptable. Diesel submarines, moreover, may be able to lie in ambush on the seafloor. On the other hand, likely enemies, at least minor powers, probably will not have the resources to find targets in the open ocean and cue submarines to attack them. This was much the situation the Royal Navy faced in the Falklands in 1982, when one Argentine submarine became a major and very difficult threat.

To the extent that they are intended to attack surface ships, submarines are armed with torpedoes and cruise missiles. A submarine can generally attack with homing torpedoes up to a range of about 6–12 mi (10–20 km); missiles may extend this range to 60 mi (100 km), which is about the limit of submarine-borne sensors. Attacks from greater ranges require the assistance of external sensors, and in that case the most effective antisubmarine warfare measure may be an attack on the external system. Effective antitorpedo weapons do not yet exist (although decoys can be used to negate a homing torpedo), but there are reasonably effective antimissile weapons. These considerations define the goals of tactical antisubmarine warfare: to find and destroy any submarine armed with torpedoes and approaching to within 6–12 mi (10–20 km). Ideally, any submarine armed with missiles must be dealt with beyond 60–90 mi (100–150 km), but failures to intercept torpedo-armed submarines are more likely to be devastating.

A central strategic question, then, is whether to destroy a submarine approaching its target or to destroy it in the open ocean. The first option implies a tradeoff between antisubmarine measures and weapons designed to neutralize submarine weapons. An alternative question might be whether to neutralize the submarine's offensive strategy or to destroy the submarine. If there are many potential targets, all more or less equivalent, it may be possible to destroy virtually all the submarines before their weapons can be used effectively. On the other hand, if there are a few but very important potential targets, such as carriers and major amphibious ships, then even a brief period of submarine success may be fatal to the overall war effort.

A second key question is the area over which the antisubmarine war is to be fought. During the Cold War, the area was the world ocean, and the West could take advantage of the fact that Soviet submarines had to pass through choke points on their way to patrol areas and also on the way back to rearm. Forces could be concentrated at the choke points. Had the war turned hot, anywhere a submarine was found, it could be attacked. In a modern expeditionary war, on the other hand, the combat zone will be quite restricted. The only choke point may be the entrance to a harbor, and it may be impossible to place forces there. Rules of engagement may drastically limit the circumstances under which submarines can be attacked. It may even be that more friendly than enemy submarines are in the area, because friendly submarines may be vital for roles such as reconnaissance and bombardment using cruise missiles. Thus classification of detected submarines may be much more important than it was in the past.

Sonar. Since invisibility is the key element of submarine operations, and since this is violated only by sound, sound sensor (that is, sonar) performance

shapes the tactics of antisubmarine warfare. Sonars may be active, that is, the listening device may impose a signature on the target, or passive, in which case the device may listen for noise from the target. There are two limits on active sonar: water can transmit only a limited amount of energy, and active sensing will tend to alert the target submarine.

The distance that sound travels through the ocean depends on its frequency: under given conditions, the lower the frequency, the farther a sound signal will travel. Thus, very low frequency sound (typically on the order of 100 Hz or less) could be transmitted across hundreds, or even thousands, of miles, whereas very high frequency sound (hundreds of kilohertz) may travel only 1 mi (1.6 km) or less. On the other hand, for a given transducer, beam definition is inversely related to sound wavelength. Therefore, even though a very low frequency signal may be received many thousands of miles from its source, the receiver may have only a vague indication of the direction of origin. *See* UNDERWATER SOUND.

During World War II, ship-borne sonars operated at very high frequency (typically about 25 kHz), but range was limited to about 1–2 mi (2–3 km), depending on local conditions. Because wavelengths were short, even very small sonar could produce a relatively narrow beam (15° wide in typical United States sets) which gave adequate definition to control the fire of short-range weapons, such as depth charges. After World War II, sonar frequency was reduced to improve range. Ultimately, large surface ships and submarines were equipped with sonars of about 3 kHz, capable in some cases of detecting targets at a distance of as much as 30 mi (50 km); but at such ranges, target position is relatively ill-defined. Passive sonars can perform even better, but only at much lower frequencies.

The best low-frequency active sonar is the size of small boats (**Fig. 1**), which is probably about the largest size that would be acceptable within a conventional hull. Due to ship impacts, especially of lower frequency and thus ever larger sonars, there is increasing interest in towed devices. Typically the pinger and the linear-array receivers are separately towed. To hold down pinger dimensions, very wide transmitted beams are accepted, the array forming sufficiently narrow receiving beams.

Fig. 1. SQS-26 low-frequency hull sonar, the largest U.S. hull sonar (the version onboard current surface ships, SQS-53, uses much the same sonar dome). Nearby workers give idea of scale. (*U.S. Navy*)

During the Cold War, Western navies concentrated on passive rather than active sonars for two reasons. First, using passive sonar denied a submarine information as to the whereabouts of those hunting it and, often, as to how close they were to attacking. Second, if the submarine can be detected passively, passive range often seems to be greater than active range. In addition, different types of submarines can sometimes be distinguished by the sounds they produce (sound can also indicate what the submarine is doing). The caveat is important because under some circumstances diesel submarines are extremely difficult to detect passively. Thus active sonars have become more important as attention has shifted from nuclear submarines, which must make some sound associated with their machinery, to diesel submarines, which can lie silently on the bottom with electric batteries for essential life support for a while.

Unfortunately, the virtues of low-frequency active sonars can reverse in shallow water. Sound bounces off both the bottom and the surface. The farther the sound propagates, the more it bounces and the more potential confusion. Yet long range seems essential. The main solution to date has been to use more complex waveforms which can register the small differences in path length associated with the bounces. Growing computer power makes such waveforms more usable. There have also been proposals to use low-frequency dipping sonar, whose beams would be projected more or less horizontally, grazing bottom and surface at such shallow angles that there would be very little reflection. No such sonar currently exists.

The most important current shipboard low-frequency (hence long-range) passive sonars are towed arrays. They are lines of passive transducers trailed far astern of a ship at the optimum listening depth. Although the array can detect distant targets (perhaps positioned hundreds of miles away), it cannot provide range and bearing data that are sufficient for fire control. Therefore, an airplane or helicopter is generally used to search the area defined by the sensor quickly enough to detect the submarine before it leaves the area. Passive sonar technology is also used by fixed listening systems (like the Cold War SOSUS and the current ADS) and by sonobuoys launched by aircraft (and occasionally by surface ships).

Sonar performance also depends on local water conditions and on the depth of the sonar in the water. Variable-depth sonars, which can be lowered to exploit water conditions, are towed below and behind the ship (**Fig. 2**). Submarines (**Fig. 3**) can also dive to reach favorable sonar conditions and can carry towed arrays that function as passive variable-depth sonars. The ridge running up the hull of the submarine in Fig. 3 houses such an array, which is streamed far behind the submarine, where the water is relatively undisturbed and hence better for listening, by using cable housed in the submarine's ballast tanks. *See* SONAR.

Alternatives to sonar. Sonar operates in an extremely difficult environment. For decades there have been attempts to find an alternative, particularly

one which would permit very rapid search from the air or from space. The only current widely used nonacoustic sensor is the magnetic anomaly detector, which typically has a range of 1 km (0.6 mi) or less. It is generally used by aircraft as a last-minute confirmation of submarine location just before a torpedo is dropped. Magnetic detectors have also been proposed as seafloor sensors in relatively shallow water, the idea being that a submarine crossing a line of such detectors would reveal itself. A series of lines would indicate submarine course and speed.

During the Cold War the Soviet Union experimented very actively with both infrared and wake sensors. Many of their submarines had the latter either on their hulls or atop their sails. Apparently these devices sought to detect the wake of a submarine through which the Soviet submarine might pass, so that the Soviet craft could trail another submarine. Detection was based on expected differences between water in the wake and outside it. Fairly wide deployment of such sensors suggests that the Soviets had enjoyed some success, but it is not clear just how much. Since the Cold War, U.S. and British submarines have been fitted with environmental sensors which may be used either to detect submarine wakes or to distinguish water conditions in which wakes would be difficult to detect (to frustrate the Russian sensors, details of which presumably became known at the end of the Cold War). Despite claims, it seems very unlikely that an effective airborne or space-borne submarine detector has yet been perfected.

Classification of objects. A major area of effort in antisubmarine warfare has been to distinguish ever fainter (more distant) submarine noises from the surrounding random ocean noise, so that attacks are not made against nonexistent targets. Ideally it would also be possible to distinguish one kind of submarine from another (for example, to avoid accidental attacks on friendly craft). The ability to distinguish the sounds of real versus nonreal objects and friendly versus enemy submarines is generally termed classification. *See* ACOUSTIC SIGNAL PROCESSING; UNDERWATER SOUND.

Classification is often based on the sound spectrum of a detected submarine. The spectrum has two components: a series of lines (narrow-band signals) associated with rotating machinery and propeller(s), and broadband sound associated with flow over the submarine's hull and also with liquid flows inside. Much of nuclear submarine design in recent years has been associated with reducing narrow-band sounds, because the nuclear submarine cannot turn off their sources (such as turbo-generators and pumps). Moreover, a signal processor can exploit the constancy of the narrow-band lines to bring them up out of noise; they can be tracked and integrated over time. The main narrow-band signals consistently produced by a well-designed diesel submarine are associated with propeller blade rate, caused by interaction between the blades and the asymmetrical flow due to the sail and particularly the tail fins. Current sickle-shaped propellers are designed to reduce blade rate. Note that the identification of nuclear submarines has generally been associated with the narrow-band spectrum. It is difficult to identify a quiet diesel submarine. However, it is hoped that by using the details of the flow noise to visualize the shape of the submarine, identification can be accomplished. That may be possible using a different mathematical basis for analysis, perhaps wavelets rather than Fourier analysis.

Fig. 2. Variable-depth sonar in streamlined housing (fish). (*U.S. Navy*)

Fig. 3. Nuclear attack submarine USS *San Francisco*. The ridge running up the hull houses a towed array. (*U.S. Navy*)

Weapons. The main weapons used in antisubmarine warfare are homing torpedoes (using active or passive sonars) and depth bombs, the latter generally fused to explode at a preset depth. Most depth bombs are unguided, but in recent years the Russians have marketed a guided homing version, and some current Russian homing torpedoes behave like guided depth bombs whose propulsion is turned on only after they detect a submarine. During the

Fig. 4. Mark 46 lightweight homing torpedo aboard a helicopter. (*U.S. Navy*)

Cold War, Western and Soviet navies deployed nuclear depth bombs, but they are no longer in service. In effect the lethal radius of such a weapon could be equated with the effective search radius of a homing torpedo, and some delivery systems (such as the U.S. ASROC) had alternative nuclear depth bomb and torpedo payloads. Because sonar range may exceed the range at which a torpedo can find a submarine target, weapons are often delivered by air (several navies use ship-launched rocket, such as the U.S. ASROC to deliver torpedoes). The U.S. Mark 46 lightweight homing torpedo (**Fig. 4**) is used by many navies: as a ship-launched torpedo, as a main airplane- and helicopter-delivered attack weapon, and as the warhead of the ship-launched ASROC. The current U.S. Mk 54 lightweight torpedo combines Mk 46 propulsion with a much more sophisticated seeker and computer. A goal of ASW and torpedo research and development is increasing precision and smaller torpedoes, enabling both larger effective magazine capacity on ships and utility on developing smaller, crewless ASW vehicles. *See* ACOUSTIC TORPEDO.

Most navies have discarded conventional depth charges because a fast submarine can evade them as they sink slowly through the water. However, a few navies retain rocket-thrown depth charges because they are effective in shallow water and against submarines lying on the bottom (a homing torpedo generally needs the Doppler effect of a moving target to distinguish it from the bottom). Homing torpedoes have thus become very nearly universal antisubmarine weapons. The rule of thumb is that a torpedo needs a 50% speed margin over a submarine. Thus, a 30-knot (55-km/h) torpedo might be effective against nonnuclear submarines whose maximum speed is about 20 knots (35 km/h), but 30-knot (55-km/h) nuclear submarines require the use of 45-knot (80-km/h) torpedoes. The faster the torpedo, the more difficult it is to distinguish the echoes of a real submarine from its own flow noise. *See* NAVAL ARMAMENT.

Target characteristics. Submarines fall into two broad classes: nuclear and conventional. Nuclear submarines can operate submerged at high speed essentially indefinitely. Although nonnuclear craft may have maximum submerged speeds only slightly less than those of nuclear submarines, they can sustain those speeds for only less than an hour because they depend upon batteries with limited energy content. (Their uninterrupted underwater endurance, at very limited speed, may be as much as a week.) The batteries are recharged by air-breathing diesel engines, and recharging thus requires the submarine to draw air from the surface. A breathing pipe, the snorkel, permits the diesels to run while the submarine is operating at a shallow depth. For several years some submarines have been built with an additional air-independent power plant, which they can use to loiter for weeks at very low speed, conserving battery power for an attack and subsequent evasion. This is designed to drastically reduce snorkeling, but comes at some cost of adding complication into a craft with a very small crew.

The type of machinery used to generate power relates directly to the noise made by the submarine. The reactor of a nuclear submarine is always running, and produces noise even when it is running very slowly. A diesel-electric (conventional) submarine is noisy when snorkeling and possibly at high speed, but the electric motor is inherently fairly quiet. However, a nuclear submarine can be designed to reduce noise significantly to the point where the craft makes no more noise than a diesel. In this case the submarine would have the enormous advantage of mobility.

In water less than the maximum submarine depth (say, less than 1000 ft or 300 m deep), a diesel submarine under pursuit can hide on the bottom, almost indistinguishable from other bottom features and largely immune to attack. Nuclear submarines cannot generally do so, because their steam power plants require a constant flow of water through their condensers, and they could therefore be disabled by ingesting mud. Diesel submarines are increasingly the threat of concern due to emphasis on littoral operations (**Fig. 5**).

Operations. Most antisubmarine warfare operations fall into three distinct categories: long-range detection (which may be followed by attacks, often by aircraft), interception by submarines, and counterattack by escorts.

Long-range detection. If a submarine can be detected at very long range, it can be attacked by an airplane cued by the detection system. Until the end of the Cold War, much of U.S. antisubmarine strategy was based on the ability of fixed undersea arrays (parts of the SOSUS sound surveillance system) and strategic ship-towed arrays to detect Russian submarines at distances of hundreds or even thousands of miles in the open Atlantic and Pacific oceans. NATO navies acquired maritime patrol aircraft, such as P-3C Orions, specifically to intercept the submarines thus detected, using sonobuoys to locate them precisely enough to attack. *See* SONOBUOY.

In the aftermath of the Cold War, naval operations are likely to occur in areas not covered by the fixed systems, and possibly not covered by a limited number of slow towed-array ships. A current research trend is toward deployable submarine detection systems, both long-range (like SOSUS) and very short

range (for use in large numbers in shallow water).

Interception. Because it is almost invisible, a submarine can lie near enemy submarine bases in hopes of intercepting emerging enemy ships. In the absence of effective long-range detectors, this sort of submarine blockade becomes an important tactic in possible post–Cold War confrontations.

Escort. If the submarines are likely to concentrate on ships as targets and if they cannot be located efficiently at long ranges, antisubmarine ships can be concentrated in a convoy around the potential targets. Convoys do not actively protect the escorted ships against attack. Rather, they act as a deterrent, since any submarines that attack must surrender their invisibility, inviting counterattacks from escort ships. Even with crude submarine detectors, it may be possible to track back the submarine torpedoes to find the submarines. Convoy escorts were responsible for the bulk of all German submarines destroyed at sea in World War II.

Convoy tactics have another effect. An individual submarine cannot detect and track potential targets at a very great range. Typically, submarine campaigns require some form of assistance, for example, by scouting aircraft. If targets are distributed randomly across the ocean, submarines may hunt them efficiently. If the targets are bunched in convoys, the submarines at sea require considerable assistance, and that assistance in turn requires communication with some central command. In World War II, this communication by the German U-boat force provided the Allies with the means of finding some submarines (for example, those awaiting fresh fuel and other supplies) and also of predicting the positions of submarine concentrations.

Convoy tactics were largely dropped during the 1950s and 1960s because it was felt that a single nuclear weapon could destroy an entire convoy. That threat has diminished because of nuclear deterrence and the possibility of limited and regional conflicts, but it is no longer clear that enough escorts can be provided in an emergency.

Escort is likely to be far more important in limited conflicts, not least because rules of engagement will almost certainly prohibit attacks on submarines that do not directly threaten surface ships. However, in the case of a limited conflict, as in the Falklands, the benefits to be gained by sinking or even damaging a single important ship, such as a carrier, may impel a submarine commander to chance counterattack from an escort.

It therefore becomes important to develop direct counters to a submarine's weapons, to be used to back up escort forces. Many ships tow noisemakers designed to decoy acoustic homing torpedoes, but such decoys do not address the threat of wake-following torpedoes.

Networked sensor system. Another possibility is to change the concept of underwater warfare by employing large numbers of sensors with inherently short ranges. Taken together, the sensors can create an underwater tactical picture. In theory, that is not too different from the Cold War use of very long range sensors. The difference is in the precision and response speed being aimed at. Typical Cold War practice was to use the long-range sensors to cue platforms, such as ships or maritime patrol aircraft, which in turn would use their own tactical sensors to locate a submarine well enough to become the target of a homing torpedo. The process was inherently drawn out, and it required large numbers of tactical platforms. In a littoral area, a submarine once out of port might be quite close to its targets, so time would be short. Probably there would be relatively few platforms in position to relocate and attack it. A possible future development involves large numbers of inexpensive short-range sensors, strewn over a littoral seafloor, sharing their information via underwater data links or buoyed radios. The resulting tactical picture would be quite precise, the submarine being located well enough that a very short range homing weapon would be able to attack it (the new concept is being developed in tandem with a very small homing torpedo).

Fig. 5. Swedish diesel submarine *Gotland* in the United States on loan from Sweden to help train U.S. naval crews against the type of submarine they are likely to encounter in littoral waters. During the Cold War, the U.S. Navy concentrated mainly on the blue-water threat presented by Soviet nuclear submarines, which used very different tactics and had very different capabilities. The light-colored plating covers the submarine's passive sonar, and torpedo tubes are visible below it. The lengthwise strake along the lower hull is a low-frequency passive array. (*U.S. Navy*)

Platforms for ASW. There are four main platforms for antisubmarine warfare: surface ships equipped with sonars; airplanes equipped with sonobuoys (which they drop into the water); helicopters carrying both sonobuoys and dipping sonars; and submarines. A fifth category, unmanned underwater vehicles (UUVs), is becoming significant.

Surface ships. These are most effective for an escort or convoy strategy. Depending on their sonar range, they may fire torpedoes or rocket-propelled depth bombs directly over the side, or they may be able to employ missiles carrying torpedoes or nuclear depth bombs. Generally, surface ships equipped with long-range, low-frequency sonars, either active or passive, support aircraft for target reacquisition and attack. The aircraft may be shipboard (like U.S. LAMPS [light airborne multi-purpose system] helicopters) or they may be land-based, either fixed-wing

Fig. 6. Missile destroyer *Forrest Sherman*. With the demise of the Soviet Union, specialized antisubmarine ships no longer seem worthwhile. *Forrest Sherman* is a multirole destroyer designed to deal with air and underwater threats and to bombard land targets. Vertical launchers, one of which is set into the deck forward of the bridge, can fire vertically launched ASROC antisubmarine missiles as well as Standard antiaircraft missiles and Tomahawk land-attack missiles. The radome atop the mast receives the data link from the LAMPS helicopter, which is used both to deal with submarines and to attack small surface craft. The anchor is located in the bow, instead of in the traditional position on the side, to clear the big sonar dome below water. (*U.S. Navy*)

or helicopters (**Fig. 6**). *See* NAVAL SURFACE SHIP.

Aircraft. Aircraft equipped with sonobuoys are primarily a means of reacquiring submarines, not of locating them. That is because the aircraft carry only a limited number of buoys, which they must expend in order to search, and which cannot be recovered. Helicopters can lower recoverable (dipping) sonars into the water. *See* HELICOPTER.

Submarines. These are not really useful for large-scale search because their sensor range is quite limited. They can be used as escorts because they often enjoy better sonar conditions than the surface ships; however, they then have the problem of communication with surface ships. A lack of communication complicates the classification problem: antisubmarine warfare is frustrating, and surface craft or aircraft are sometimes tempted to attack any submarine they detect, without knowing whether or not it belongs to the enemy. *See* SUBMARINE.

Crewless underwater vehicles. The use of UUVs is currently of interest mainly as sensor or sensor-distribution platforms. In littoral areas, for example, they or their surface equivalents may be the best means of creating the sensor net which is intended to detect and track submarines. However, in future it is quite possible to imagine armed UUVs which might lie in wait at a harbor entrance, attacking emerging enemy submarines either under command (via an acoustic data link, which is now practicable thanks to increased computer power) or autonomously, based on rules of engagement in the vehicle's computer. A stealthy UUV might also be used to tag a submarine, even one in harbor, with a transponder which would make later detection relatively easy. Yet another possibility, given a sufficient energy source, would be for the UUV to trail the submarine, either cueing antisubmarine forces or itself attacking on command. *See* UNDERWATER VEHICLE. Norman Friedman

Bibliography. A. D. Baker III and J. L. Labayle-Couhat, *Combat Fleets of the World: Their Ships, Aircraft, and Armament*, published every 2 years (current U.S. editor is Eric Wertheim); D. R. Frieden, *Principles of Naval Weapons Systems*, 1985; N. Friedman, *Naval Institute Guide to World Naval Weapon Systems*, various editions; N. Friedman, *The Postwar Naval Revolution*, 1987; N. Friedman, *U.S. Destroyers since 1945: An Illustrated Design History*, 2d ed., 2004; N. Friedman, *Seapower as Strategy*, 2001; N. Friedman, *U.S. Submarines since 1945: An Illustrated Design History*, 1994; W. Hackmann, *Seek and Strike*, 1984; R. Harding (ed.), *The Royal Navy 1930–2000: Innovation and Defence*, 2004; J. R. Hill, *Anti-Submarine Warfare*, 2d ed., 1989.

Antitoxin

An antibody that will combine with and generally neutralize a particular toxin. When the manifestations of a disease are caused primarily by a microbial toxin, the corresponding antitoxin, if available in time, may have a pronounced prophylactic or curative effect. Apart from this, the other properties of an antitoxin are those of the antibody family (IgG, IgA, IgM) to which it belongs. *See* ANTIBODY; IMMUNOGLOBULIN.

Antitoxins have been developed for nearly all microbial toxins. Diphtheria, tetanus, botulinus, gas gangrene, and scarlatinal toxins are important examples. Antitoxins may be formed in humans as a result of the disease or the carrier state, or following vaccination with toxoids, and these may confer active immunity. The status of this can be evaluated through skin tests, for example, the Schick test for diphtheria antitoxin and the Dick test for scarlatinal antitoxin of scarlet fever, or by titration of the serum antitoxin level. Animals, especially horses, may also be immunized, and their antitoxic sera, usually refined, employed in the passive immunization of humans. *See* BOTULISM; DIPHTHERIA; GANGRENE; IMMUNITY; TETANUS; TOXIN-ANTITOXIN REACTION; VACCINATION.

Antitoxin standardization is accomplished by comparing the abilities of standard and unknown antitoxins to neutralize the toxic or serologic activities of a reagent toxin, as in the Romer skin-neutralization test in rabbits or the Ramon flocculation test. Internationally recognized standard reference antitoxins

are available from various governmental laboratories, which also define procedures for their use. *See* NEUTRALIZATION REACTION (IMMUNOLOGY).

Henry P. Treffers

Bibliography. G. S. Wilson and A. A. Miles (eds.), *Topley and Wilson's Principles of Bacteriology and Immunity*, 2 vols., 6th ed., 1975.

Anura

One of the three living orders of class Amphibia, encompassing the frogs and toads. Anurans differ most obviously from salamanders (order Caudata) in lacking a tail. Usually, frogs and toads have long hindlimbs adapted for the hopping locomotion characteristic of members of the order. There can be no confusion of anurans with the limbless caecilians, members of the third order of amphibians, Apoda.

Morphology. Anurans are short-bodied animals with a large mouth and protruding eyes. The externally visible part of the ear, absent in some forms, is the round, smooth tympanum situated on the side of the head behind the eye. There are five digits on the hindfeet and four on the front. Teeth may be present on the upper jaw and the vomerine bones of the roof of the mouth, and are found on the lower jaw of only one species. Often teeth are totally lacking, as in toads of the genera *Bufo* and *Rhinophrynus*.

The short vertebral column consists of six to ten vertebrae, usually nine, and the elongate coccyx. The sacral vertebra precedes the coccyx and bears more or less enlarged lateral processes with which the pelvic girdle articulates (**Fig. 1**). A characteristic feature of anurans is the fusion of the bones in the lower arm and lower leg, so that a single bone, the radioulna in the arm and the tibiofibula in the leg, occupies the position of two in most other tetrapods (**Fig. 2**).

Distribution. Over 4200 species of anurans are known, so these animals are far more diversified than the salamanders (with less than 420 species) or caecilians (with about 165 species). Only the extreme

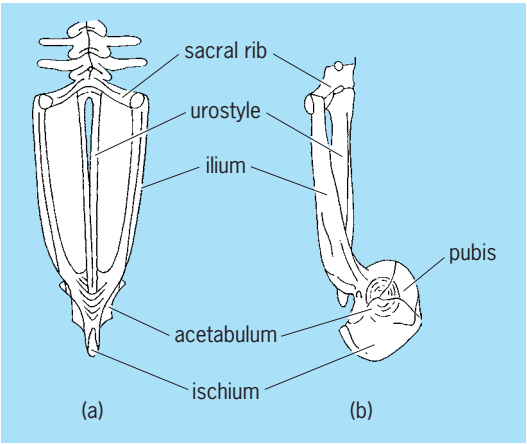

Fig. 1. Pelvic girdle of the frog in (*a*) dorsal view and (*b*) lateral view. (*After W. Montagna, Comparative Anatomy, Wiley, 1959*)

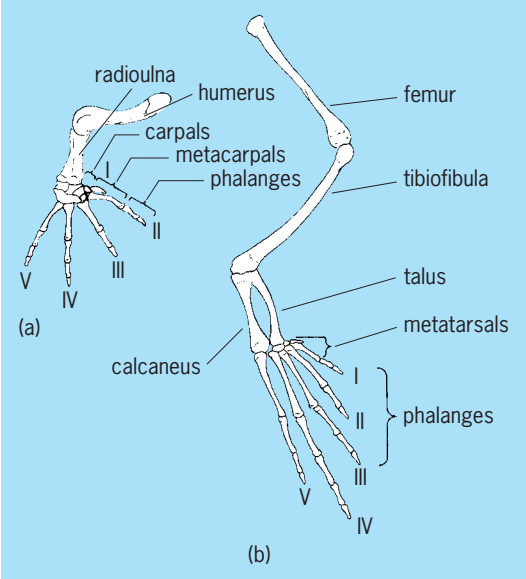

Fig. 2. Limbs of frog showing characteristic fusion of bones in (*a*) anterior limb and (*b*) posterior limb. (*After W. Montagna, Comparative Anatomy, Wiley, 1959*)

frozen polar regions and remote oceanic islands lack native anurans; 80% of the species live in the tropics. The concentration of species in tropical regions contrasts with the distribution of salamanders, mostly found in more temperate areas.

Vocalization. The one character of anurans that comes to the attention of most people, including many who may never see one, is the voice. Unlike salamanders, which are mute or nearly so, most anurans have voices and use them in a variety of ways. In the breeding season, great numbers of males may congregate in favorable sites and call, each species giving its own characteristic vocalization. Because no two species breeding at the same time and place have identical calls, it is assumed that the call is important in aiding individuals to find the proper mate (species recognition). In some species, it appears that the female is active in selecting the mate and may be responding to the mating call, but the call may not act in exactly the same way in other species. The mating call is given with the mouth closed. Air is shunted back and forth between the lungs and the mouth, so individuals can call even when submerged. Many species possess one or two vocal sacs, which are expansible pockets of skin beneath the chin or behind the jaws. The sacs (**Fig. 3**), which may be inflated to a volume as great as that of the animal itself, serve as resonators.

Other noises made by anurans include the so-called fright scream, given with the mouth open, and the warning chirp, which evidently serves as a sex recognition signal when one male contacts another. Some calls evidently serve as territorial signals. Efforts to monitor anuran populations often take advantage of the calling behavior of frogs and toads. *See* ANIMAL COMMUNICATION.

Reproduction and development. In general, male anurans grasp the female about the body with the

Fig. 3. Toad of the genus *Bufo* giving mating call with vocal sac expanded. (*American Museum of Natural History photograph*)

forelegs, a procedure called amplexus, and fertilizes the eggs externally as they are extruded in the water. The number of eggs may be quite large (up to 20,000 in the bullfrog or 25,000 in a common toad) or may be as few as one in some species. The eggs are each surrounded by concentric coats of jelly and may be deposited, according to the habit of the species, singly, in groups of various sizes and shapes, or in strings.

The larvae (tadpoles) are at first limbless and have external gills and a muscular tail with dorsal and ventral fins (**Fig. 4**). At hatching, there is no mouth opening present, but one soon forms that develops a horny beak and several rows of labial teeth not at all like the true teeth of the adult animal. Shortly after the tadpole hatches, the gills become enclosed within chambers and are no longer visible externally. Except for the gradual development of the hindlimbs, no additional external changes take place as the tadpole grows until the time for metamorphosis. The anterior limbs, which have been forming hidden in the gill chambers, break through the covering skin as metamorphosis begins. The tail dwindles in size as it is absorbed, while the mouth assumes its adult shape. Many other changes are taking place internally, including shortening of the intestine and its adaptation to the carnivorous diet of the adult frog.

The pattern of breeding and development outlined above (referred to as the complex life cycle) is widespread among anurans and is undoubtedly the primitive one. However, many modifications of this pattern have evolved. Many species lay eggs in moist places on land. Each egg is provided with sufficient yolk to allow the embryo to pass an abbreviated larval stage within the egg and emerge as a transformed small frog. The female marsupial frog *Gastrotheca* of South America carries the eggs in a pouch on the back, from which tadpoles or fully formed frogs emerge, according to the species. In the Surinam toad *Pipa*, also of South America, the eggs undergo development while situated in pits in the back of the mother. The male of Darwin's frog (*Rhinoderma darwini*), another South American species, has the remarkable habit of carrying the eggs and larvae in his vocal sac until metamorphosis occurs. The most highly specialized breeding habit among the Anura is seen in an African genus, *Nectophrynoides*, which is ovoviviparous; that is, the young develop within the maternal oviduct.

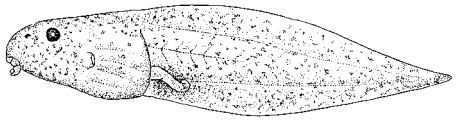

Fig. 4. Tadpole, or larval, stage of the frog *Rana pipiens*. (*After W. F. Blair et al., Vertebrates of the United States, 2d ed., McGraw-Hill, 1968*)

Nutrition. All adult anurans are carnivorous. Diet seems to depend largely upon the size of the individual, and the capacious mouth of some species permits astonishing feats of swallowing. A large bullfrog, for example, may snap up low-flying bats, ducklings, snakes, and turtles. Insects and other invertebrates form the bulk of the diet of most anurans. The tongue, moistened by a sticky secretion from the intermaxillary gland in the roof of the mouth, is used to catch smaller prey, while larger items of food may bring the front limbs into play. When swallowing, an anuran will usually depress its eyeballs into its head to aid in forcing the food down the pharynx.

In contrast to adult anurans, most tadpoles are vegetarian and feed on algae. A few are largely carnivorous or sometimes cannibalistic, and even vegetarian species will scavenge for dead animal matter. Striking feeding specializations occur, such as the funnel mouth of certain tadpoles that skim food from the surface of the water. Some are suspension feeders, like the eastern narrowmouth toad (*Gastrophryne carolinensis*), filtering organisms and organic particles out of the water column.

Ecology. The habitats of anurans are as various as the places where freshwater accumulates. Lakes and streams provide year-round habitat for many species, while others may be found in the aquatic habitat only during the breeding season. For instance, temporary aquatic habitats of the Atlantic Coastal Plain provide critical habitat for a striking diversity of anurans. The cycles of filling and drying preclude large predators from the habitat and provide ideal places for developing anurans. Any permanent source of water in the desert is likely to support a population of one or more species, and when rainstorms occur, the air around a temporary pool may be filled with mating calls for a few nights, while the animals take advantage of the water for breeding. Sometimes, the pool goes dry before the tadpoles metamorphose, and the adult frogs retreat underground to await another rain. Moist tropical regions provide an abundance of habitats little known to temperate regions, such as the air plants (bromeliads) that hold water and so provide a moist home and breeding site for anurans that may never leave the trees.

Economics. Anurans are used by humans in two important ways, as food and as laboratory animals.

Many thousands of frog legs are consumed annually, and the demand in the United States is sufficiently great that the domestic supply is supplemented by imports from overseas. Thousands more of frogs are used each year as laboratory animals, both as specimens for dissection and study in zoology classes, and as experimental animals for research on a variety of zoological and medical topics. Perhaps a more important service of anurans results from their ecological position as consumers of insects.

Classification. Although the majority of anurans fall into fairly well-defined familial categories, the arrangement of the families into subordinal groups by different authorities is not consistent, and there is controversy about the relationships of some smaller groups. The arrangement adopted here is tentative and undoubtedly will need to be changed as additional studies are made of the anatomy of anurans and are correlated with discoveries in the fossil record.

Archaeobatrachia. This suborder of the Anura is distinguished from others in having vertebrae that are concave on both anterior and posterior surfaces. Ribs are present in the adult animals, a feature confined to this group and the suborders Opisthocoela and Aglossa. Only four living species belong to the simple family Ascaphidae: three of the genus *Liopelma*, the only frogs native to New Zealand, and *Ascaphus truei* of the western United States and adjacent Canada. *Ascaphus* is called the tailed frog because of the taillike extension of the male cloaca that serves as an intromittent organ, thus permitting internal fertilization (**Fig. 5**). Small streams in forested regions are the habitat of *Ascaphus*; the flattened tadpole, with reduced tail fin and suckerlike mouth, is adapted to life in swiftly running waters.

Mesobatrachia. Some members of this suborder have trunk vertebrae that are convex anteriorly and concave posteriorly, and the adults typically have free ribs. The typical opisthocoelous family is the Discoglossidae with three genera, *Bombina*, *Alytes*, and *Discoglossus*, in Europe, North Africa, and the Orient, and a fourth, *Barbourula*, endemic to a single island in the Philippines. There are only nine species of discoglossids. Also included within the Opisthocoela at times, but of obscure relationship, is the monotypic Mexican family Rhynophrynidae.

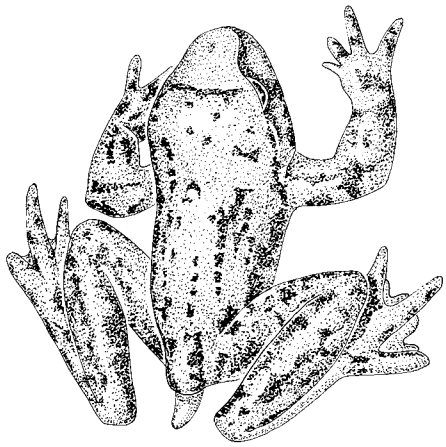

Fig. 5. Male tailed frog (*Ascaphus truei*) with a cloacal appendage, unique to the species, between the legs.

The Pipidae are also included within the Mesobatrachia. As the subordinal name suggests, a definitive tongue is lacking. Free ribs are present in the tadpole, but they fuse to the vertebrae in the adult. The vertebrae are opisthocoelous. This family of four genera and 30+ species is found in South America and Africa. A South American genus, *Pipa*, is noted for its odd breeding habits. Amid peculiar gyrations, the eggs are fertilized and deposited on the female's back, where they become implanted. Protected there, they develop in individual pouches, and eventually the miniature toads rupture the skin and emerge. An African genus is *Xenopus*, the clawed frogs, widely used in human pregnancy tests. Pelobatidae and Pelodytidae are also families assigned to this suborder. The species of the North American genus *Scaphiopus* and the European genus *Pelobates*, which is also found in northwestern Africa, are burrowing forms commonly known as spadefoot toads because of the sharp-edged digging tubercle on the hindfoot.

Neobatrachia. The suborder Neobatrachia contains most of the diversity of the order Anura and is characterized by a procoelous vertebral column, in which each vertebra is concave anteriorly and convex posteriorly, and with a free coccyx articulating by a double condyle. This suborder commonly includes 19 families, including Leptodactylidae, Bufonidae, Hylidae, Centrolenidae, Ranidae, Rhacophoridae, and Microhylidae.

Leptodactylidae. The leptodactylid frogs are most abundant in number of genera and species in the American tropics and Australia, with a very few species found as far north as the southwestern United States and New Guinea. A single genus in South Africa possibly belongs to this family. There are about 60 genera of leptodactylids with more than 600 species; about 300 species are in the genus *Eleutherodactylus* alone. Evolutionary lines within the family have undergone extensive adaptive radiation, so that one or another species lives in almost every fashion known to frogs. Species with similar habits are often similar in appearance too, so that leptodactylid frogs are often almost indistinguishable externally from a variety of species of other families.

Bufonidae. Members of the large genus *Bufo*, the true toads, are native to virtually every place in the world where anurans live, except the Australian region. The other four genera of the Bufonidae are found in Africa and the Malay region. These genera are of little numerical importance, but one of them, *Nectophrynoides* of eastern Africa, includes the only ovoviviparous anurans. All bufonids lack teeth, and the true toads of the genus *Bufo* are for the most part rather warty, short-legged, terrestrial animals that enter the water only during the breeding season.

Hylidae. The tree frogs are one of the larger amphibian families, with over 700 species known. The majority of these, about 350 species, belong to the genus *Hyla*, which is found in both Eastern and Western hemispheres. Many of the Hylidae are adapted

Fig. 6. North American tree frogs: green tree frog, *Hyla cinerea* (*left*) and gray tree frog, *H. versicolor*. (*American Museum of Natural History photograph*)

to arboreal life in having expanded digital disks that facilitate climbing (**Fig. 6**), but some have adopted other modes of existence and lead a terrestrial, aquatic, or even burrowing life.

The paradox frog of South America, genus *Pseudis*, and its relatives are sometimes placed within the Hylidae but may be accorded their own family, Pseudidae. The paradox frog is so called because the adult frog is smaller than the tadpole.

Centrolenidae. In the American tropics is found a small group of three genera and about 20 species of small, translucent, arboreal frogs grouped together in the family Centrolenidae. A peculiar characteristic of some species (but not restricted to this family) is that the bones are green.

Ranidae. The Ranidae number about 500 species, of which approximately 170 belong to the genus *Rana* (**Fig. 7**). This genus is fairly well represented wherever anurans live, except in South America and Australia. In each of these continents, there is only a single species of *Rana*, and the ecological position of the ranids appears to be filled by leptodactylids. *Rana* includes the abundant typical frogs, as opposed to toads, of North America and Europe. In Africa and Asia, in addition to *Rana*, there are more than 30 other genera of diverse habits and appearance that belong to the Ranidae. The poison frogs of Central and South America are small, brilliantly colored frogs whose skin secretions serve to poison the arrows of Indians (**Fig. 8**). These frogs were formerly grouped in the Procoela, but studies indicate that they may more correctly be considered as a subfamily, Dendrobatinae, of the family Ranidae.

Fig. 7. Northern leopard frog (*Rana pipiens*), the most widespread species in North America, found from Hudson Bay in Canada south to Panama and from the Atlantic almost to the Pacific. (*American Museum of Natural History photograph*)

Fig. 8. Panamanian poison frog (*Dendrobates auratus*). (*American Museum of Natural History photograph*)

Rhacophoridae. Externally, hylids and rhacophorids may resemble one another very closely, but the characteristics of the vertebral column and pectoral girdle indicate their true relationships. The geographic distribution of the two families is largely mutually exclusive. Where hylids are abundant in the New World, Australia, and New Guinea, there are no rhacophorids. Over 400 species of Rhacophoridae are credited to the fauna of Africa, where only a single *Hyla* is known. About five species of *Hyla* occur in Asia among over 100 rhacophorids. The center of rhacophorid diversity is Africa, where about 20 genera are found. Only two genera, and these are doubtfully distinct, occur in Asia.

Microhylidae. The frogs of the family Microhylidae are found in three regions of the world: southern Africa and Madagascar; Sri Lanka (Ceylon), India, and southern China to New Guinea and the extreme northern part of Australia; and South America northward into the southern United States. Many of the microhylids are chunky little frogs with a pointed head and tiny mouth, and lead a terrestrial or burrowing life. In some regions, notably in New Guinea, extensive adaptive radiation has produced a variety of ecological types paralleling more typical members of other families. The specialization of microhylids is such that there are more genera recognized (58) than of ranids and rhacophorids together, though in

number of species (about 230) the microhylids are inferior to either of the other groups. Investigators question the relationship of the Ranidae and Microhylidae. The two families may not be so closely related as inclusion in the same suborder would indicate; their skeletal similarities may be convergent rather than truly indicative of close phylogenetic relationship. Richard G. Zweifel; W. Ben Cash

Bibliography. W. E. Duellman and L. Trueb, *Biology of Amphibians*, 1986; M. J. Lannoo (ed.), *Amphibian Declines: The Conservation Status of United States Species*, University of California Press, Berkeley, 2005; D. W. Linzey, *Vertebrate Biology*, McGraw-Hill, New York, 2001; R. W. McDiarmid and R. Altig (eds.), *Tadpoles: The Biology of Anuran Larvae*, University of Chicago Press, 1999; F. H. Pough et al., *Herpetology*, 3d ed., Prentice Hall, Upper Saddle River, NJ, 2004; R. D. Semlitsch (ed.), *Amphibian Conservation*, Smithsonian Institution Press, Washington, DC, 2003; R. C. Stebbins and N. W. Cohen, *A Natural History of Amphibians*, Princeton University Press, 1995; G. R. Zug, J. V. Laurie, and J. P. Caldwell, *Herpetology*, Academic Press, San Diego, 2001.

Anxiety disorders

A group of distinct psychiatric disorders (including generalized anxiety disorder, panic disorder, obsessive-compulsive disorder, social anxiety disorder, and posttraumatic stress disorder) that are characterized by marked distress and impairment of functioning. Fear is an adaptive emotion that is seen in many animal species, while anxiety can be viewed as the subjective component of fear and, therefore, a normal emotion selected for by evolutionary processes. Anxiety disorders are among the most prevalent, disabling, and costly of psychiatric disorders. They are frequently underrecognized and undertreated. Fortunately, there is growing understanding of the mechanisms underlying these disorders, and a number of effective treatments are available. *See* NEUROTIC DISORDERS.

Generalized anxiety disorder. Generalized anxiety disorder (GAD) is characterized by excessive worry, tension, and anxiety. Accompanying physical symptoms include muscle tension, restlessness, fatigability, and sleep disturbances. Generalized anxiety disorder occurs in around 4–6% of the population and is the anxiety disorder most frequently encountered by primary care physicians, to whom sufferers may go to seek help for the physical symptoms of the disorder. This disorder is more common in women, and frequently begins in the teenage years with increasing prevalence into the fourth decade. Research on the impairment of functioning found in people with generalized anxiety disorder relative to sufferers with other medical disorders indicates that generalized anxiety disorder is one of the most disabling of all medical disorders.

Etiology and pathophysiology. Animal studies of fear and clinical studies of people with generalized anxiety disorder suggest that similar brain circuits are involved in both cases. For example, the amygdala (a small structure located in the front [anterior] part of the temporal lobes) plays an important role in fear and anxiety. Numerous complex connections to a range of other brain areas allow it to coordinate cognitive, emotional, and physiological responses to fear and anxiety. Thus in the "fight or flight" response, the organism makes cognitive-affective decisions about how to respond to the perceived danger and has a range of somatic (for example, increased heart rate and respiration rate) and endocrine (for example, release of "stress hormones") responses that act together to increase the likelihood of avoiding the danger.

Within the fear and anxiety circuits in the brain, various neurochemical systems [for example, amino acid neurotransmitter systems, such as glutamate and gamma-aminobutyric acid (GABA), and monoaminergic neurotransmitter systems, such as serotonin and noradrenaline] are responsible for mediating the communication between the functionally connected regions. Medications acting on these systems, such as benzodiazepines, selective serotonin reuptake inhibitors (SSRIs), and noradrenergic/serotonergic reuptake inhibitors (NSRIs), are effective in treating generalized anxiety disorder. *See* NEUROBIOLOGY; NORADRENERGIC SYSTEM; SEROTONIN.

Treatment. Both specific medications and particular kinds of psychotherapy have proven effective in the treatment of generalized anxiety disorder. Although benzodiazepines have often been used to treat generalized anxiety disorder, selective serotonin reuptake inhibitors (SSRIs) and noradrenaline serotonin reuptake inhibitors (NSRIs) are currently viewed as first-line options because of their favorable safety profile. Cognitive-behavioral psychotherapy focuses on using behavioral techniques and changing underlying thought patterns; it is the best studied of the psychotherapies in generalized anxiety disorder, with growing evidence for its efficacy. *See* PSYCHOPHARMACOLOGY; PSYCHOTHERAPY.

Panic disorder. Panic disorder (PD) is characterized by repeated, sudden panic attacks, which are accompanied by a range of physical symptoms, including respiratory (shortness of breath), cardiovascular (fast heart rate), gastrointestinal (nausea), and occulovestibular (dizziness) symptoms. Agoraphobia frequently coexists with panic disorder and involves anxiety about and avoidance of situations that provoke a panic attack. The prevalence of panic disorder is approximately 2% in the general population; it is more common in women and is often complicated by depression. Panic attacks mimic a number of important medical conditions that need to be carefully excluded at initial assessment. *See* PHOBIA.

Etiology and pathophysiology. The same brain circuits and neurotransmitters implicated in fear and generalized anxiety disorder are also likely to play a role in panic disorder. "Challenge" studies that activate particular neurochemical systems have been used to demonstrate the involvement of GABA, serotonin, and noradrenaline in panic disorder. This is

complemented by evidence of treatment efficacy of medications that modulate these systems. Carbon dioxide provokes panic attacks, an observation that has led to the hypothesis that panic disorder involves a false suffocation alarm. As predicted by this theory, panic attacks are more common when pulmonary carbon dioxide is increased (for example, in lung disease) and is less common when pulmonary carbon dioxide is decreased (for example, after pregnancy).

As in generalized anxiety disorder, the amygdala appears to play an important role in panic. Damage to this area may be associated with loss of fear responses. Conversely, in conditions with increased activation of the amygdala, such as temporal lobe epilepsy, there may be an increased predisposition to panic. Brain imaging has also suggested a role for the parahippocampal region, which appears to be involved in remembering contextual aspects of fear (that is, where the panic attack occurred). *See* BRAIN.

Treatment. Treatment of panic disorder, as other anxiety disorders, involves the use of specific medications and psychotherapies. The first-line choice of medication is typically an SSRI or NSRI. Benzodiazepines are effective alone or in combination with SSRIs, but their use as the only medication is generally avoided due to the potential for dependence and withdrawal. Cognitive-behavioral therapy that addresses avoidance behavior and irrational dysfunctional beliefs is also effective.

Obsessive-compulsive disorder. Obsessive-compulsive disorder (OCD) is characterized by obsessions (unwanted, persistent, distressing thoughts) and compulsions (repetitive acts to relieve anxiety caused by obsessions). The disorder occurs in 2–3% of the population and often begins in childhood or adolescence. Interestingly, obsessive-compulsive disorder is also seen in the context of certain infections, brain injury, and pregnancy or the puerperium (6-week period after delivery). Unfortunately, since people are often embarrassed by their symptoms and unaware of available treatments, it is common for diagnosis and appropriate treatment to be delayed for many years. *See* OBSESSIVE-COMPULSIVE DISORDER.

Etiology and pathophysiology. A range of evidence now implicates a brain circuit (loop) between the frontal cortex, basal ganglia, and thalamus in mediating obsessive-compulsive disorder. Lesions to the basal ganglia (for example, in certain neurological disorders) may lead to obsessions and compulsions. Brain-imaging studies have demonstrated increased brain activity in the basal ganglia in obsessive-compulsive disorder, and show that both medication and specific psychotherapies are able to normalize this activity. In certain very treatment-resistant cases of obsessive-compulsive disorder, neurosurgical procedures that effect minor interruptions in communication between the implicated cortico-basal brain regions, can bring some relief of symptoms.

Key neurotransmitters in this circuit include the dopamine and serotonin neurotransmitter system. Genetic studies have suggested a role for aspects of the dopamine system in obsessive-compulsive disorder. Medications that increase dopamine in the basal ganglia increase repetitive movements in animal models of obsessive-compulsive disorder, and can increase tics in people with obsessive-compulsive disorder. Furthermore, medications that lower dopamine transmission appear to be helpful as part of the treatment of obsessive-compulsive disorder. A range of data also points to the importance of the serotonin system. This includes the selective response of obsessive-compulsive disorder to serotonergic and not to noradrenergic antidepressants.

Treatment. Countering misperceptions in the community is crucial to treating many of the anxiety disorders. A wider understanding of obsessive-compulsive disorder as a treatable medical disorder would encourage people to seek treatment sooner. SSRIs are current first-line treatments for obsessive-compulsive disorder, with dopamine blockers added in those who do not respond to these agents. Behavioral therapy focuses on exposure and response prevention, while cognitive strategies address the distortions in beliefs that underlie the perpetuation of symptoms.

Social anxiety disorder. Social anxiety disorder (SAD) is characterized by persistent fears of embarrassment, scrutiny, or humiliation. People with social anxiety disorder may avoid social situations (such as talking in small groups and dating) and performance situations (eating or speaking in front of others), resulting in marked disability. For some, symptoms are confined to one or more performance situations, while for others symptoms may be generalized to include most social and performance situations. Generalized social anxiety disorder is usually more severe, and sufferers are more likely to have a family history of social anxiety disorder.

Social anxiety disorder is particularly common, with prevalence figures in some studies upwards of 10%. While women are more likely to have social anxiety disorder, men with the disorder may be more likely to come for treatment. Onset is typically in the teenage years, with symptoms continuing for many years. Social anxiety disorder is often complicated by depression, and people with it may self-medicate their symptoms with alcohol, leading to alcohol dependence. *See* AFFECTIVE DISORDERS; ALCOHOLISM.

Etiology and pathophysiology. There is a growing search for the structures and chemicals that are involved in thoughts and feelings related to social interaction and status. Animal studies have, for example, found that serotonin and dopamine receptors differ in dominant and subordinate primates. People with social anxiety disorder who are exposed to facial expressions that invoke fear of scrutiny will selectively increase brain activity in the amygdala (previously discussed in relation to fear in panic and generalized anxiety disorder). In addition, brain-imaging studies have found that effective treatment with medication and psychotherapy normalizes activity in the amygdala and the closely related hippocampal region in social anxiety disorder.

Treatment. Many people with symptoms of social anxiety disorder view these as immutable aspects

of their personality. Education of the community through advocacy groups can be helpful in changing this perception. When social anxiety disorder is identified in the clinic, it is important to begin treatment by addressing such misperceptions. SSRIs, NSRIs, and cognitive-behavioral therapy are all effective in the treatment of social anxiety disorder. Monoamine oxidase inhibitors and benzodiazepines are also known to be effective treatments for the disorder, but they have a number of disadvantages. *See* MONOAMINE OXIDASE.

Posttraumatic stress disorder. By definition, posttraumatic stress disorder (PTSD) is an abnormal response to severe trauma. It is characterized by distinct clusters of symptoms: reexperiencing of the event (for example, in flashbacks, dreams, or distress in response to reminders of the trauma), avoidance (of reminders of the trauma), numbing of responsiveness to the environment, and increased arousal (for example, insomnia, irritability, and being easily startled). Although exposure to severe trauma occurs in more than 70% of the population, posttraumatic stress disorder has a lifetime prevalence of 7–9% in the general population. Risk factors for developing posttraumatic stress disorder following exposure to severe trauma include female gender, previous history of psychiatric illness, trauma severity, and absence of social support after the trauma. *See* POSTTRAUMATIC STRESS DISORDER.

Etiology and pathophysiology. Brain-imaging studies have suggested that in posttraumatic stress disorder frontal areas (for example, the anterior cingulate cortex) may fail to effectively dampen the "danger alarm" of the amygdala (known to mediate cognitive, affective, and physiological responses to intense fear). Whereas stress responses ordinarily recover after exposure to trauma, in posttraumatic stress disorder they persist. Furthermore, brain-imaging studies have suggested that hippocampal volume may be reduced in posttraumatic stress disorder, perhaps accounting for the memory deficits that are characteristic of this disorder.

One possibility is that stress hormones are responsible for the decreased hippocampal volume in posttraumatic stress disorder. Indeed, there is growing evidence that functioning of the hypothalamic-pituitary-adrenal hormonal axis in posttraumatic stress disorder is disrupted. Other systems, such as serotonin and noradrenaline, may however also be involved. *See* STRESS (PSYCHOLOGY).

Treatment. Given that posttraumatic stress disorder may occur in the context of interpersonal trauma, establishing a relationship with a trusted therapist is important. As with many of the anxiety disorders, both SSRIs and cognitive-behavioral therapy are effective in decreasing posttraumatic stress disorder symptoms. Behavioral techniques (using different forms of exposure in the safety of the consultation room) or cognitive retraining (addressing irrational thoughts on the trauma and its consequences) can both be helpful. Paul D. Carey; Dan J. Stein

Bibliography. E. Hollander and D. Simeon, *Concise Guide to Anxiety Disorders*, American Psychiatric Publishing, Washington, 2003; Social anxiety disorder; *Acta Psychiat. Scand.*, vol. 108, suppl. 417, 2003; D. J. Stein, *Clinical Manual of Anxiety Disorders*, American Psychiatric Publishing, Arlington, 2004; D. J. Stein and E. Hollander (eds.), *Textbook of Anxiety Disorders*, American Psychiatric Publishing, Washington, DC, 2002; Update on posttraumatic stress disorder, *J. Clin. Psychiat.*, vol. 65, suppl. 1, 2004.

Anyons

Particles obeying unconventional forms of quantum statistics. For many years, only two forms of quantum statistics were believed possible, Bose-Einstein and Fermi-Dirac, but it is now realized that a continuum of possibilities exists.

In quantum mechanics, the probability that a given state will occur is calculated by adding the amplitudes for all processes leading to that state, and then squaring this total amplitude. Thus, in the case of two indistinguishable particles A and B, the amplitude for the process that leads to A arriving at point x while B arrives at point y must be added to the amplitude for the process that leads to A arriving at y while B arrives at x—the so-called exchange process—because the final states cannot be distinguished. Actually, this simple recipe is appropriate only for particles obeying Bose-Einstein statistics (bosons). The recipe for particles obeying Fermi-Dirac statistics (fermions) is to subtract the amplitude for the exchange process. *See* BOSE-EINSTEIN STATISTICS; FERMI-DIRAC STATISTICS.

The primary definition of anyons posits other possible recipes for adding exchange processes. Anyons of type θ are defined by the rule that the amplitudes for processes where A winds around B n times in a counterclockwise direction are added with a factor $e^{2in\theta}$. In this definition, half-integer values of n are allowed; half a winding is an exchange. The values $\theta = 0$ and $\theta = \pi$ correspond to the previous definitions of bosons and fermions, respectively.

The more general possibilities for quantum statistics, corresponding to other possible values of the angle θ, are defined in a mathematically consistent way only for particles whose motion is restricted to two space dimensions. However, many of the most interesting materials in modern condensed-matter physics are effectively two-dimensional, including microelectronic circuits and the copper-oxide layers fundamental to high-temperature superconductivity. The question as to whether the elementary excitations in these systems, the quasiparticles, are anyons can be settled only by a deep investigation of specific cases. The quasiparticles in fractional quantized Hall states are known to be anyons. *See* HALL EFFECT; INTEGRATED CIRCUITS; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; QUANTUM STATISTICS; SUPERCONDUCTIVITY. Frank Wilczek

Bibliography. F. Wilczek, Anyons, *Sci. Amer.*, 264(5):58–64, May 1991; F. Wilczek (ed.), *Fractional Statistics and Anyon Superconductivity*, 1990.

Aorta

The main vessel of the systemic arterial circulation arising from the left ventricle of the heart; it is divided into three parts for convenience only. The first portion, the ascending aorta, passes upward under the pulmonary artery; the coronary arteries arise at the base of the ascending aorta behind the aortic valves. The second part, or aortic arch, curves over the hilum of the left lung, giving off the innominate, left carotid, and left subclavian arteries, which supply the neck, head, and forelimbs. The third portion, or descending aorta, continues downward in the thorax on the left side of the vertebral column to the diaphragm, giving off small arteries to the bronchi, esophagus, and other adjacent tissues. Below the diaphragm this vessel, known as the abdominal aorta, descends to the level of the fourth lumbar vertabra where it bifurcates into the two common iliac arteries supplying the hindlimbs.

In the abdomen the major branches of the aorta include the single celiac, superior mesenteric and inferior mesenteric, and the paired renal and internal spermatic (male) or ovarian (female) arteries. In addition, many small branches go to other organs and to the body wall.

The aorta is a large elastic artery. Its endothelial lining is continuous with that of the heart and its arterial branches. Around this lining is a thin tunica intima, consisting of longitudinal elastic fibers, followed by a heavy layer of circularly arranged tunica media, which has a predominant proportion of elastic fibers arranged in dense layers between muscle tissue and collagenous fibers. The outer tunica adventitia consists of loose connective tissue which contains fat, nerves, and nutrient blood vessels called vasa vasorum. This arrangement of layers, especially the elastic fibers, enables the aorta to expand and recoil in response to each pulsation of blood through it. Walter Bock

Aortic body

One of the cellular aggregations traditionally known as chemoreceptors which are generally found adjacent to the aortic arch and pulmonary artery and which often are adjacent to the orgin of the subclavian arteries (see **illus.**). The term is best restricted to those groups of cells that have topographic, morphological, and functional similarities. However, many aggregations of epitheloid cells have been described which lie adjacent to the aorta in the thorax, in the root of the neck, and in the abdomen, and any of these could be included, with some topographic descriptive justification, in the group called aortic bodies.

The comparative physiology and anatomy of aortic bodies are not well documented because these organs are microscopic and, therefore, not readily accessible and identifiable. Their functions are not well studied, even in mammals. However, studies have shown that the morphological substrate of these bodies appears to be similar in clusters of two types of epitheloid cells, one of which is innervated, surrounded by a number of large capillaries.

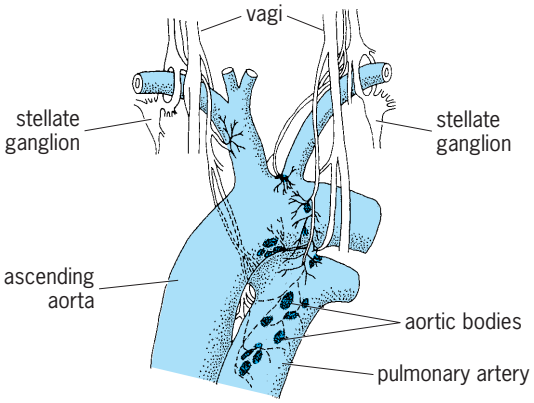

Oval-shaped aortic bodies are found lying on the anterior and posterior surfaces of the aortic arch and pulmonary artery.

There are two types of parenchymal cells in aortic bodies. Chief or glomus (type I) cells are usually polyhedral in shape, have fingerlike processes, and contain membrane-bound electron-dense granules that contain neuropetides and biogenic amines, usually dopamine and norepinephrine. These parenchymal cells are connected by gap junctions. Sustentacular (type II) cells are supporting elements which closely surround several chief cells. These cells are not innervated and do not contain vesicles, which are characteristic of chief cells. Sustentacular cells resemble the supporting Schwann cells associated with chromaffin tissues and have many ultrastructural features in common with such tissues. In addition to functional differences, the principal dissimilarities between the two types of cells are the size of the cytoplasmic granules and the nature of the contained amines, the nature of the innervation of the chief cells, and the distinct structure of capillary blood vessels that form a characteristic feature of the chemoreceptors.

The presence of catecholamines has been demonstrated by high-performance liquid chromatography. Acetylcholine and neuropeptides, which are known to be neurotransmitters in other tissues, have also been found in chief cells. Receptors for these neurochemicals have been localized in aortic body cells and nerve endings. However, the neurotransmitter which is specific for arterial O_2 and CO_2 chemoreception in the aortic bodies has not been identified. It is likely that several neurotransmitters are released in response to a single stimulus and the sensory responses are dependent on the simultaneous effects of these transmitters.

Aortic bodies are innervated by afferent nerve fibers derived from the vagus nerve and efferent nerve fibers of the sympathetic nervous system. Both types of fibers terminate in synaptic-type expansions on chief cells. The larger (afferent) endings of the vagal fibers contain a variable number of clear synaptic vesicles and mitochondria; the endings of sympathetic nerve fibers contain synaptic-type

dense vesicles that are typical of postganglionic noradrenergic nerve endings. However, all chief cells are not innervated. There are no loose nerve endings between the cells, and the sustentacular cells are not innervated. The distribution and innervation of aortic bodies by the left and right aortic nerves are not symmetrical.

The aortic bodies in adults receive arterial blood supplies from small and slender blood vessels that originate from either the coronary artery or the aortic arch. In the fetal stage, oxygen-poor blood of the pulmonary arterial system supplies these chemoreceptor tissues as well. The extensive capillary network is surrounded by a continuous fenestrated endothelium.

The aortic bodies resemble carotid bodies in that they respond to systemic hypoxia eliciting chemosensory discharge. They are poorly responsive to carbon dioxide partial pressure or to hydrogen ion concentration, although these stimuli augment the response to hypoxia. The impulses that travel up the aortic nerves increase in hypoxia; under conditions of high oxygen tensions the reverse happens. The mechanism by which this effect is achieved is still unknown, but it appears that chief cells are the first chemoreceptive sensors. These cells influence the central nervous system by generating impulses through the release of neurotransmitters. The chemosensory fibers of the aortic body show a profound excitatory response to hypotension, and this response is synergistically augmented by hypoxia. Conditions such as anemia and carboxyhemoglobinemia are potent stimuli for the aortic chemoreceptors. However, the response is ultimately due to a fall in the partial pressure of oxygen in the cells. Accordingly, aortic bodies appear to have less of an oxygen supply than the carotid bodies and hence are stimulated by any factor that tends to compromise oxygen delivery to them. These chemoreceptors are also transiently stimulated by the metabolic inhibitors of oxidative phosphorylation. *See* CAROTID BODY.

As a part of the reflex arc, the aortic bodies subserve the control function of both the respiratory and the cardiovascular systems. The respiratory function of the aortic body appears to be relatively weak compared to that of carotid bodies. The effects on the cardiovascular reflex are expected to be significant owing to the fact that blood pressure effects on the chemosensory discharge are vigorous. These effects would have to be mediated by the autonomic nervous system. However, the physiological function of the aortic bodies is still under investigation.

Tumors may arise that involve either chief cells or chromaffin cells (rather than sustentacular cells). The tumor may contain high concentrations of catecholamines, usually norepinephrine. *See* CHEMORECEPTION.

Sukhamay Lahiri

Bibliography. H. Acker and R. G. O'Regan (eds.), *Physiology of the Peripheral Arterial Chemoreceptors,* 1983; N. S. Cherniack and J. G. Widdicombe, *Handbook of Physiology: The Respiratory System,* vol. 2, pp. 247–362, 1986.

Apatite

The most abundant and widespread of the phosphate minerals, crystallizing in the hexagonal system, space group $P6_3/m$. The apatite structure type includes no less than 10 mineral species and has the general formula $X_5(YO_4)_3Z$, where X is usually Ca^{2+} or Pb^{2+} or As^{5+}, and Z is F^-, Cl^-, or $(OH)^-$. The apatite series takes X = Ca, whereas the pyromorphite series includes those members with X = Pb. Three end members form a complete solid-solution series involving the halide and hydroxyl anions and these are fluorapatite, $Ca_5(PO_4)_3F$; chlorapatite, $Ca_5(PO_4)_3Cl$; and hydroxyapatite, $Ca_5(PO_4)_3(OH)$. Thus, the general series can be written $Ca_5(PO_4)_3(F,Cl,OH)$, the fluoride member being the most frequent and often simply called apatite.

The apatite series is further complicated with other substitutions. Carbonate-apatites involve the coupled substitution $(PO_4)^{3-} \rightleftharpoons (H_4O_4),(CO_3)^{2-}$; $Ca^{2+} \rightleftharpoons H_2O$; but the exact nature of this substitution is not yet known. The fluoride- and chloride-bearing carbonate-apatites are called francolite and dahllite, respectively. They are of considerable biological interest since they occur in human calculi, bones, and teeth. Dental enamels are composed in part of dahllite, and the efficacy of fluoride-bearing toothpaste partly concerns apatite crystal chemistry. Other less frequent substitutions occur in nature, involving $(AsO_4)^{3-} \rightleftharpoons (PO_4)^{3-}$ and the arsenate end members are known as svabite, $Ca_5(AsO_4)_3F$, and hedyphane, $Ca_5(AsO_4)_3Cl$. More limited substitutions can include many other ions, such as Mn^{2+} for Ca^{2+} (mangan-apatite), and Si^{4+} and S^{6+} for P^{5+}.

Apatites have been synthesized hydrothermally by the hydrolysis of monetite, $CaH(PO_4)$, or by direct fusion of the components, such as $Ca_3(PO_4)_2 + CaF_2$, which yields fluorapatite. Other apatite structures have been synthesized, usually by fusion of the components, involving the following elements and ions: Ba, Pb, Mg, Ni, Sr, Zn, Cr (3+ and 6+), Al, Fe, Na, Ce, Y, O^{2-}, $(OH)^-$, F^-, Cl^-, $(CO_3)^{2-}$, and $(SO_4)^{2-}$.

Diagnostic features. The apatite isomorphous series of minerals occur as grains, short to long hexagonal prisms terminated by pyramids, dipyramids, and the basal pinacoid (see **illus.**). The minerals are transparent to opaque, and can be asparagus-green (asparagus stone), grayish green, greenish yellow, gray, brown, brownish red, and more rarely violet, pink, or colorless. Mangan-apatite is usually blue or blue-gray to blue-green. Some apatites fluoresce in ultraviolet radiation, in particular the mangan-apatite variety. The cleavage is poor parallel to the base and fracture uneven. Apatites are brittle, with hardness 5 on Mohs scale, and specific gravity 3.1–3.2; they are also botryoidal, fibrous, and earthy.

Occurrence. Apatite occurs in nearly every rock type as an accessory mineral. In pegmatites it occurs as a late-stage fluid segregate, sometimes in large crystals and commonly containing Mn^{2+}. It often crystallizes in regional and contact metamorphic rocks, especially in limestone and associated with chondrodite and phlogopite. It is very common in basic

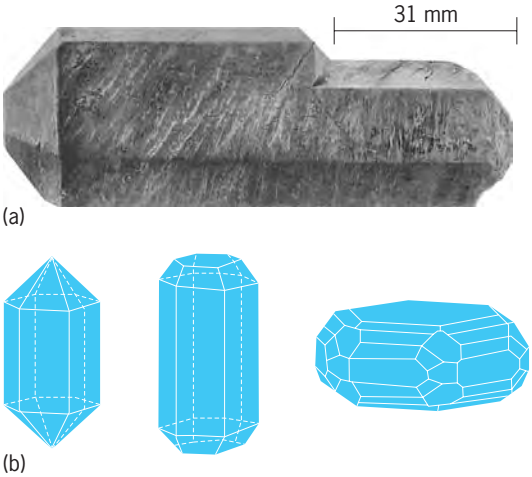

Apatite. (*a*) Specimen for Eganville, Ontario, Canada (*American Museum of Natural History specimens*). (*b*) Typical crystal habits (*after C. Klein, Manual of Mineralogy, 21st ed., John Wiley & Sons, 1993*).

to ultrabasic rocks; enormous masses occur associated with nepheline-syenites in the Kola Peninsula, Russia, and constitute valuable ores which also contain rare-earth elements. Apatite is a common accessory with magnetite ores associated with norites and anorthosites and is detrimental because it renders a high phosphorus content to the ores. Apatite occurs rarely in meteorites along with other phosphates.

Large beds of oolitic, pulverulent, and compact fine-grained carbonate-apatites occur as phosphate rock, phosphorites, or collophanes. They have arisen by the precipitation of small concretions formed by organisms and by the action of phosphatic water on bone materials or corals. Extensive deposits of this kind occur in the United States in Montana and Florida and in North Africa. The material is mined for fertilizer and for the manufacture of elemental phosphorus. *See* FERTILIZER; PHOSPHATE MINERALS; PHOSPHORUS; PYROMORPHITE. Paul B. Moore

Apes

The group of primates most closely related to humans. They include the African great apes, the gorillas and chimpanzees; the Asian great apes, the orangutans; and the lesser apes from Asia, the gibbons. The apes can be distinguished from the rest of the primates by a number of anatomical and behavioral traits, which indicate their common origin; thus they are considered a monophyletic group and are called the Hominoidea.

Apes are distinguished from other primates through such obvious features as the absence of a tail and presence of an appendix. They share a number of specializations (synapomorphies) of the skeleton, which are useful as diagnostic characters, particularly when it comes to distinguishing fossil apes, because bones and teeth are the most readily preserved parts in the fossil record. The distal end of the humerus is especially useful, both because it is one of the most robust body parts and therefore readily preserved, and because it is diagnostic of the ape condition, with a large trochlea (ulnar forearm articulation) and a well-developed trochlea ridge. The wrist is also modified for mobility of the joint. The lumbar vertebrae are distinctive in having robust pedicles, caudally inclined spinous processes, and dorsally oriented transverse processes arising from the pedicle. There are few synapomorphies of the skull, which in general retains the primitive primate condition except in individual species, but two shared specializations are the deep arched palate and relatively small incisive foramina. The teeth also are generally primitive, except in the broad, low-crowned incisors and enlarged molars. *See* DENTITION; SKELETAL SYSTEM.

Each group of apes differs from this basic ape pattern in varying degrees.

Gibbons. The gibbons retain many of the primitive ape characteristics. (Note that the specialized, or derived, conditions that characterize the apes as a group must subsequently be considered primitive when differences within the group are considered.) They have also developed a number of unique characters that are different from any other ape: They have very elongated arms (relative to body weight) with many modifications of the musculature for a brachiating form of locomotion (swinging by the arms). Their legs are also lengthened, so that they are quite efficient bipeds. They have adopted a monogamous family system, which is unusual in primates, and one of the outcomes is role sharing between males and females and lack of size variation between the sexes. They have also developed a complex system of vocal communication related both to their close social bonds and to their thick tropical forest environment, which makes visual communication difficult. At present there are 13 species of gibbon (*Hylobates*) divided into three subgenera [Hylobates (**Fig. 1**), Nomascus, Symphalangus] occupying most parts of Southeast Asia where primary forest still remains. *See* SOCIAL MAMMALS.

Orangutan. The sister group to the gibbons is the great ape and human group, which is readily distinguished by the shared presence of enlarged premolars in all its members. Within this group can be distinguished the orangutans and the African apes. The orangutans have a great many specializations that support this separation, although because of its common heritage with the chimpanzee and gorilla, and the great increase in body size of all three, the orangutan has a superficial similarity which has led in the past to all being grouped in the same family, the Pongidae. The differences separating the orangutans from other apes are in part biochemical (for example, the structure of their blood proteins) and in part morphological. The deep face, enlarged premaxilla, narrow distance between the eyes, massive zygomatic bones, smooth floor of nose, and enlarged central incisors are characters unique to the orangutan, and in all of these traits the African apes and humans retain the primitive ape condition.

The orangutan is today confined to the tropical rainforests of Borneo and Sumatra, where it exists

Fig. 1. *Hylobates lar* (white-handed gibbon). (*Photo by Gerald and Buff Corsi, © 2004 California Academy of Sciences*)

Fig. 2. *Pongo pygmaeus* (orangutan). (*Photo by H. Vannoy Davis, © 2001 California Academy of Sciences*)

as two variable species. Protein differences between them are well above the subspecies level, and there are differences in the chromosomes and in mitochondrial deoxyribonucleic acid (mtDNA) cleavage sites, and it would appear that the orangutan is a good example of recent speciation. The species from Borneo is known as *Pongo pygmaeus* (**Fig. 2**), and the one from Sumatra, as *P. abelii*. Both are largely arboreal despite their great size, which ranges in body weight from 88 to 308 lb (40 to 140 kg). They lead solitary or small-group existences, leading to the massive sexual-size variation indicated by body weight variation. Little is known of the social structure.

African apes. The other part of the great ape group consists of four species of African apes, two each of gorilla and chimpanzee. They are distinguished from the orangutans (and other primates) by specializations of the wrist and frontal sinus and the development of massive brow ridges (all of which they also share with humans), and by a further series of unique modifications of the hand that are related to its unusual form of locomotion, called knuckle walking. Their legs are reduced in length (relative to body weight) so that their arms appear long in proportion to the rest of their body. The effects of this are further increased both by elongation of the hand and by the fact that when they walk quadrupedally on the ground they support their weight on the middle phalanges (finger bones) of the hand. This raises the body to a semiupright position even when the animals are walking on all fours. The orangutan, which parallels the leg reduction and therefore the proportionally longer arms, does not use its knuckles for weight support in this way but supports itself on its clenched fists.

The common chimpanzee (*Pan troglodytes*; **Fig. 3**) inhabits much of the forested region of tropical Africa, extending beyond the forest into more seasonal wooded habitats as well. The pygmy chimpanzee, or bonobo (*Pan paniscus*), is confined to the southern loop of the Congo River, where it inhabits mainly swamp forest. The two gorilla species are also confined to tropical Africa, but the genus is divided into the eastern gorilla (*Gorilla beringei*) and the western gorilla (*G. gorilla*; **Fig. 4**). The gorillas are the largest of the primates, with body weights ranging from 165 to 396 lb (75 to 180 kg), while the chimpanzee at 88 to 110 lb (40 to 50 kg) is much

Fig. 3. *Pan troglodytes* (chimpanzee). (*Photo by Gerald and Buff Corsi, © 2002 California Academy of Sciences*)

Fig. 4. ***Gorilla gorilla gorilla*. (*Photo by Gerald and Buff Corsi, © 2002 California Academy of Sciences*)**

smaller. Gorilla social groups consist of a dominant male with several females and immature males and females, while chimpanzees live in fluctuating multimale or multifemale groups.

The relationships between the species of ape are shown in **Fig. 5**, with the gibbons at the top, divided into three subgenera of the genus *Hylobates*; orangutans (*Pongo pygmaeus* and *P. abelii*); the African great apes and humans at the bottom; and some fossil species that will be described below. Notice that the orangutans are put into a separate group (or clade) from the other great apes (the chimpanzees and gorillas) because the latter have been shown to be more closely related to humans in evolutionary terms than they are to the orangutan. This signifies that the idea of the "great apes" encompassing all three is not a valid taxonomic concept, and the use of a single family, the Pongidae, to include them all is not correct. The great apes, however, have many superficial similarities to each other, mostly related to their large size, so that they may be said to be similar in grade, but in terms of clade (ancestral-descendant relationship) they are different. This could be recognized by restricting the use of the family Hominidae, but other classifications are also possible based on this set of relationships.

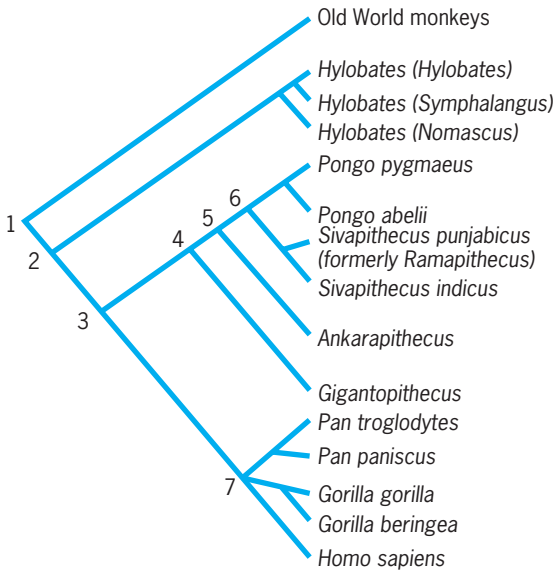

Fig. 5. Cladogram showing the relationships of fossil and living apes. The relationships are based on shared morphological characters (synapomorphies) as described in the text, and there is no time element inherent in the cladogram. The timing of divergences between groups can be inferred from the oldest data for which fossils are known: the date at which the Old World monkey split with the apes must have been in excess of 20 million years; the divergence of the orangutan must have occurred before 11 million years ago; and the divergence of humans from the African apes must have taken place before 4 million years ago. Nodes 1–7 are explained in the text.

Fossil apes. The fossil history of the apes can be traced back about 22 million years to the early Miocene. Claims have been made that the Oligocene primate *Aegyptopithecus zeuxii* from the Egyptian Fayum can be considered an ape, which could take ape history back to around 35 million years ago, but recent evidence from the postcranial bones, especially the distal end of a humerus, and from the skull, shows that *Aegyptopithecus* has none of the characters described earlier as being diagnostic of the living apes. Thus there is no reason for including them as apes, however great their superficial similarity to apes based on shared primitive characters. *Aegyptopithecus* was a small, rather heavily built arboreal primate that could be viewed as being physiologically similar to the ancestral monkeys and apes (at node 1 in the illustration).

Even by the early Miocene, for which there is a wealth of fossil evidence from many sites in Kenya and Uganda, it is far from certain that apes have been definitely identified. *Proconsul* is the best known of these primates, together with four other genera and a total of 10 species; like *Aegyptopithecus* it mainly shares primitive characters with living apes. The evidence for its being an ape comes from its postcranial morphology, which shows many similarities to ape morphology in the shoulder, elbow, and foot. These similarities have now been demonstrated for at least two of the species of *Proconsul*, and they appear to settle the question of the hominoid affinities of this fossil genus. This means that two things can be said about ape evolution: first, the apes must have differentiated from the other primates by the time *Proconsul* lived 22 million years ago; and second, the initial modifications that characterized them as apes were postcranial. This is interesting, because it will have been seen from the descriptions of the living apes that their greatest specializations, both as a group and as individuals, are postcranial ones, with all of the living apes having adopted quite extraordinary methods of locomotion: gibbons, brachiating; orangutans, fist walking; chimpanzees and gorillas, knuckle walking; and humans, bipedal walking. Perhaps the trend to this extraordinary diversity of locomotion started with the beginnings of ape evolution.

Sivapithecus. By the middle Miocene, three groups of apes had differentiated which have apparent affinities with groups of living apes. The best-documented group is *Sivapithecus* from India, Pakistan, Greece, and Turkey. These are small to medium-size apes that lived some 8–13 million years ago, and they share with the orangutan many of the specializations of the face mentioned earlier as being diagnostic of this species. The synapomorphies include the narrow distance between the eyes, concave face, smooth floor of nose, massive zygomatic bones, and enlarged central incisors (relative to the lateral incisors). One or all of these characters are known to be present in *S. indicus* and *S.* (formerly *Ramapithecus*) *punjabicus* from India and Pakistan and *Ankarapithecus meteai* from Turkey. Skulls of *S. indicus* from Pakistan and *A. meteai* from Turkey show these characters clearly, but similar fossils from Greece, which some authorities have linked with the *Sivapithecus* group, appear to lack these orangutan synapomorphies. These species are also characterized by thick enamel on their molar teeth, a feature present in orangutans and also in humans, which has led in the past to some or all of these fossils being described as humanlike or even as being actual human ancestors.

Mention can also be made of *Gigantopithecus*, which appears to share many of the characters of this node, but there is no evidence for any of the body parts characteristic of the orangutan clade and so it is not possible to place it exactly. *Gigantopithecus* is a giant ape known from India and China, and like *Sivapithecus* it has thick enameled teeth and a massively built mandibular body. Unlike the other fossils, however, it also has molariform premolars and canines and greatly reduced incisors; the most that can be said of its relationships at present is that it probably fits in the illustration between nodes 3, 4, and 6.

Dryopithecus. The second group of middle Miocene ape consists of at least three species from Europe. It includes the classic *Dryopithecus fontani*, a species comparable in size with *S. punjabicus*, and it is known mainly from France, Germany and Spain, where it has been described under numerous synonyms. There is also evidence from Hungary and Spain for several other species of *Dryopithecus* considerably smaller in size. All have the enlarged premolars, mentioned earlier to be diagnostic of the great ape and human group; and for most of the characters described above as uniquely shared by the orangutan and *Sivapithecus* species, the specimens of *Dryopithecus* from Hungary and Spain retain the primitive condition; that is, they lack the orangutan synapomorphies. They lack the thickened enamel present in the sivapithecines, but they appear to have dental enamel slightly thicker than was present in earlier apes like *Proconsul*. They have postcranial adaptations for suspensory locomotion similar to that seen in the living great apes, particularly in their limb proportions and lengthened hands, so that this group of fossil apes may now be placed near the origin of the great ape and human clade, which is at or near node 3 in Fig. 5.

A group of late Miocene apes from Greece (*Graecopithecus freybergi*, or "*Ouranopithecus macedoniensis*") have similarities in skull form to *Dryopithecus*, but what little is known of their postcranial skeleton suggests that they retain the primitive condition for hominoids, and they have thickened enamel on their teeth. Thus, their present taxonomic position is uncertain, and they are not shown in Fig. 5.

Kenyapithecus. The third group of middle Miocene fossil ape is an African group represented by the genera *Kenyapithecus* and *Afropithecus* from Kenya and *Heliopithecus* from Saudi Arabia. There is some uncertainty about the status of this group and of its phylogenetic affinities. *Kenyapithecus* shows some similarities to later specimens of *Sivapithecus*, and like this Asian genus the molars have thickened enamel, but on the other two genera the enamel is only moderately thickened. All of them, however, share massive premolar enlargement, and it is this feature that groups them. In other respects, especially their postcrania, they appear to retain ancestral characters, so that they also fall on or near node 3 in Fig. 5. One species of this genus is also known from outside Africa, in Turkey, and the date of this site (Pasalar ~15 million years ago) provides the earliest evidence of the emigration of apes from Africa to Europe and Asia. There is no fossil evidence for the later stages of African ape and human evolution (node 7 or higher in Fig. 5). *See* FOSSIL HUMANS; MONKEY.

Peter Andrews

Bibliography. C. Groves, *Primate Taxonomy*, 2001; W. Hartwig, *The Primate Fossil Record*, 2002; J. R. Napier and P. H. Napier, *A Handbook of Living Primates*, 1967; D. R. Swindler (ed.), *Comparative Primate Biology*, vol. 1, 1986; F. S. Szalay and E. Delson, *Evolutionary History of the Primates*, 1979.

Aphasia

Impairment in the use of spoken or written language caused by injury to the brain, which cannot be accounted for by paralysis or incoordination of the articulatory organs, impairment of hearing or vision, impaired level of consciousness, or impaired motivation to communicate. The language zone in the brain includes the portion of the frontal, temporal, and parietal lobes surrounding the sylvian fissure and structures deep to these areas. In right-handed persons, with few exceptions, only injury in the left cerebral hemisphere produces aphasia. Lateralization of language function is variable in left-handers, and they are at greater risk for becoming aphasic from a lesion in either hemisphere. *See* HEMISPHERIC LATERALITY.

Distinctive recurring patterns of deficit are associated with particular lesion sites within the language zone. These patterns may entail selective impairment of articulation, ability to retrieve concept names, or syntactic organization. Other dissociations affect principally the auditory comprehension of speech, the repetition of speech, or the recognition of written words. The erroneous production of

unintended words in speech (paraphasia), oral reading (paralexia), or writing (paragraphia) is a feature of some forms of aphasia.

In general, injury to the speech zone anterior to the rolandic fissure (Broca's area) results in nonfluent aphasia, with effortful articulation, loss of syntax, but relatively well-preserved auditory comprehension. Postrolandic lesions typically result in fluent aphasia. Facility of articulation, grammatical organization, and rate of speech are well preserved, while the comprehension of language and word choice are most affected. Within the group of fluent aphasias three major subtypes may be distinguished. Wernicke's aphasia, associated with injury to the posterior portion of the first temporal convolution, is characterized by severely impaired auditory comprehension, and fluent, but markedly paraphasic speech, sometimes at faster than normal rate. Anomic aphasia, typically caused by injury at the temporoparietal junction, leaves the patient unaffected in comprehension, ease of articulation, or grammatical organization, but unable to recall the principal nouns and verbs for the message. Conduction aphasia features selectively impaired repetition, with relatively fluent speech production and good comprehension. This form of aphasia is often associated with damage in the white matter underlying the parietal operculum. In the transcortical aphasias both repetition and the ability to recite overlearned material are remarkably preserved, while communicative speech is virtually impossible. Anterior (transcortical motor) and posterior (transcortical sensory) forms are recognized. *See* BRAIN.

While the above syndromes often occur in clearcut form, mixed forms of aphasia, caused by multiple lesions or lesions spanning anterior and posterior portions of the speech zone, are quite common, and massive destruction of the entire language area results in a global aphasia. Further, individual variations in behavioral manifestations of similar lesions have set limits on the strict assignment of function to structures within the language area.

Preadolescent children suffering aphasia after unilateral injury usually recover rapidly, presumably by virtue of the capacity of the right cerebral hemisphere early in life to acquire the language functions originally mediated by the left hemisphere. Capacity for recovery of function decreases during later adolescence and young adulthood.

Complete recovery in adults after a severe injury is much less common, and severe aphasia may persist unchanged for the duration of the person's life. Many individuals are aided by remedial language training, while others continue severely impaired. *See* MEMORY.

Harold Goodglass

Apiales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Rosidae of the class Magnoliopsida (dicotyledons). The order (also known as the Umbellales) comprises two families, the Araliaceae, with about 700 species, and the Apiaceae, with about 3000. They are herbs or woody plants with mostly compound or conspicuously lobed or dissected leaves (see **illus.**), well-developed schizogenous secretory canals, separate petals, and an inferior ovary that has only one ovule in each locule. The nodes are generally multilacunar, and the pollen is trinucleate. Ginseng (*Panax*) and English ivy (*Hedera helix*) are well-known members of the Araliaceae.

Cow parsnip (*Heracleum lanathum*), family Umbelliferae, order Umbellales. (*Photograph by A. W. Ambler, from National Audubon Society*)

The Apiaceae are mostly aromatic herbs, most numerous in temperate regions. The flowers consistently have an ovary of two carpels, ripening to form a dry fruit that splits into halves, each containing a single seed. Some common garden vegetables and spice plants, including carrot (*Daucus carota*), parsnip (*Pastinaca sativa*), celery (*Apium graveolens*), parsley (*Petroselinum crispum*), caraway (*Carum carvi*), and dill (*Anethum graveolens*), belong to the Apiaceae, as do also such notorious poisonous plants as the poison hemlock (*Conium*) and water hemlock (*Circuta*). *See* ANISE; CARROT; CELERY; FENNEL; GINSENG; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PARSLEY; PARSNIP; PLANT KINGDOM; ROSIDAE.

Arthur Cronquist

Apical dominance

Correlative inhibition of the growth of lateral (axillary) shoots exerted by the growing apical bud of the plant stem. Partial or complete apical dominance can also occur in the initiation of lateral roots and in the orientation of lateral organs such as branches, leaves, and stolons.

Origin of lateral meristems. In the apical meristem, cell division occurs in the young forming leaves and

along the flanks of the apical bud. As the number of cells increases, they elongate, pushing the apical meristem upward and leaving a small portion of the dividing cells behind the axil of each of the laterally forming leaves. This meristem, called the lateral meristem, may remain a small group of cells with little organization or may develop into an axillary bud possessing short internodes, unexpanded leaves, and an apical meristem.

In roots, branching does not directly involve the apical meristem. Lateral roots do not form from organized buds but originate in the layer of cells between the endodermis and root vascular system (pericycle). Cell division within the pericycle results in meristematic tissue which will form a lateral root. *See* APICAL MERISTEM; LATERAL MERISTEM; ROOT (BOTANY).

Lateral plant growth. The degree of apical dominance over the lateral buds varies with the plant species. Some plants, such as pea and sunflower, exhibit strong apical dominance, causing the formation of single branchless shoots. Other plants, such as tomato, have weak apical dominance, allowing axillary bud growth and a bushy growth habit.

Apical dominance can be broken by several factors, including apical bud removal (decapitation), horizontal positioning of the plant's main axis (gravistimulation), shoot inversion, low light intensity, or short-day photoperiods. In some situations, apical dominance is weakened as the plant becomes older. For example, in sycamore and ash, strong apical dominance controls the early growth of the plant, which results in a single leader shoot. In later years, weak apical dominance causes a branching habit. In herbaceous plants, apical dominance is often weakened at the end of the growing season, especially when the terminal bud forms a terminal flower. After apical dominance is broken, the released lateral shoots have a hierarchy of dominance of their own. Frequently, the uppermost shoot becomes dominant and inhibits the growth of the lower shoots. *See* PHOTOPERIODISM; PLANT MOVEMENTS.

Plant organs other than the main shoot are under the control of apical dominance. In roots, strong apical dominance causes tap-root growth, whereas weak apical dominance results in a fibrous root system. Also, increased lateral root initiation in the pericycle occurs in detipped roots. Leaves and branches which grow at characteristic angles to the main axis of the stem will grow more upright after removal of the apical bud of the main stem. The growth form of some organs is also controlled by the presence of the apical bud. Rhizomes (underground axillary shoots lacking chlorophyll and having rudimentary leaves) of potato are one example of such control. Rhizomes will grow upright, forming green, leafy shoots if the apical bud and all the aboveground axillary buds are removed.

Lower plants, such as mosses and ferns, as well as fungi and algae, also exhibit apical dominance. In one example, the removal of the fruiting body of the fungus, *Phycomyces*, results in the formation of lateral branches which are normally absent.

Nutritional factors. One hypothesis for the mechanism of apical dominance is based on the competition of nutrients between centers of growth. According to this hypothesis, nutrients are preferentially transported to the apical bud, causing deficiencies in the axillary buds. Decapitation may shift nutrient transport from the sink (apical bud) and allow for accelerated transport to the lateral buds. Nutrient deficiencies have been found to inhibit bud growth, and nutrient supplementation can overcome bud inhibition in some cases. However, the concept that apical dominance is controlled solely by nutritional factors is unlikely. Frequently, there has been failure to relieve bud inhibition by direct application of nutrients, and the nutrient levels of inhibited buds are not always lower than the levels in growing buds. *See* PLANT MINERAL NUTRITION.

Hormonal factors. A second hypothesis for the mechanism of apical dominance proposes that a plant hormone produced in the apical bud acts as the correlative signal.

Auxin. The plant growth hormone, auxin, is synthesized in the young leaves of the apical bud and is transported down the stem. Addition of auxin to decapitated shoots restores apical dominance, thus replacing the function of the apical bud in its inhibition of lateral bud growth. Therefore, auxin is proposed to be the primary signal involved in maintaining apical dominance in stems. Accordingly, if auxin directly inhibits lateral bud growth, then high levels of auxin are also expected in the inhibited buds. However, contrary to this theory, lateral buds have been found to be low in auxin. Auxin-directed nutrient transport has been proposed to account for inhibition of lateral bud growth by auxin since some nutrients accumulate in areas of high auxin levels. Thus, tissues low in auxin, such as axillary buds, may be starved for nutrients. Auxin may also control the transport of other hormones essential for lateral bud growth. However, it is interesting to note that auxin levels actually increase in growing lateral buds. Thus, auxin may be required for shoot growth after buds are released from apical dominance.

In roots, auxin promotes the initiation of lateral roots. Therefore, auxin may not act as the correlative inhibitor in roots as it is proposed to be in shoots. However, little research has been conducted on the hormonal control of apical dominance in roots. *See* AUXIN.

Cytokinins. Cytokinins are well known as cell-division factors and promote lateral bud growth in plants under strong apical dominance. Cell division is blocked in inhibited axillary buds. In the early stages of cytokinin-promoted bud growth, cell division is stimulated within 1 h, accompanied by increases in nucleic acids. Doubtless, outgrowth of suppressed lateral buds involves activation of certain genes, resulting in changes in the rates of synthesis of key regulatory proteins. However, investigations of hormone-induced gene products involved in apical dominance have not been completed.

It is postulated that the auxin cytokinin ratio is the important factor in determining whether lateral

buds grow or remain inhibited. There is evidence that auxin inhibits cytokinin transport from the roots (primary site of cytokinin biosynthesis) to the lateral bud. Supporting this observation is the fact that levels of cytokinin at the bud site are low when the buds are suppressed and increase greatly when the suppression is broken by decapitation or gravistimulation.

Cytokinin application directly to potato rhizomes results in their conversion to leafy, upright shoots. In decapitated and disbudded plants, the presence of roots is necessary for this rhizome-shoot conversion. Thus, lack of cytokinin from the roots may be responsible for the maintenance of the rhizome growth form.

In roots, cytokinin strongly inhibits root initiation in detipped roots. It has been proposed that cytokinin is the primary inhibiting agent from the root apex controlling apical dominance over lateral rooting. *See* CYTOKININS.

Other hormones. Although auxin and cytokinin are the major hormones regulating apical dominance, there is increasing evidence that other hormones (gibberellin, abscisic acid, and ethylene) contribute to this process. Gibberellin application enhances auxin's effect in maintaining apical dominance in decapitated shoots and has been found to promote auxin accumulation in the area of application. Thus, plant tissues high in gibberellin, such as the apical bud, may accumulate a high concentration of auxin. However, gibberellin also stimulates cytokinin-promoted lateral bud growth. Thus, it appears that gibberellin may contribute to the elongation of already growing buds. In rhizome growth of potatoes, addition of gibberellin and auxin together will cause aerial axillary buds to form rhizomes, whereas either hormone alone will not yield that effect. *See* GIBBERELLIN.

The growth inhibitor, abscisic acid, occurs in high levels in inhibited lateral buds under strong apical dominance in some plant species. The direct addition of abscisic acid to lateral buds will counteract the growth-promoting effect of cytokinin application or decapitation of the main stem apex. It has been proposed that a high level of auxin in the stem may trigger abscisic acid synthesis in the lateral buds. Abscisic acid also inhibits lateral rooting when applied to detipped roots. However, the role of abscisic acid in apical dominance of roots and shoots is still unclear. *See* ABSCISIC ACID.

The gaseous plant hormone, ethylene, inhibits lateral bud growth promoted by cytokinin or decapitation when it is supplied continuously. Ethylene is produced at especially high levels at the plant nodes where the lateral buds are situated. High levels of auxin in plant tissues promote ethylene biosynthesis in these nodes, and it is postulated that auxin levels in the stem stimulate ethylene production, thus suppressing lateral bud growth. Accordingly, decapitation should cause a lowering of ethylene production at the nodes, reducing bud inhibition. However, no lowering of ethylene is observed in the nodes of decapitated shoots. Therefore, an inhibitory role for ethylene in apical dominance seems unlikely. In contrast to an inhibitory effect, a short burst of ethylene promotes lateral bud elongation. Thus, under certain conditions, ethylene may contribute to bud growth after release from apical dominance. Still another hypothesis suggests that ethylene produced in inverted or gravistimulated shoots retards main shoot elongation and allows for the subsequent axillary bud outgrowth. *See* ETHYLENE; PLANT GROWTH; PLANT HORMONES.

Marcia Harrison

Bibliography. L. J. Audus and F. T. Addicott (eds.), *Abscisic Acid*, 1983; M. A. Harrison and P. B. Kaufman, Estimates of free and bound indole-3-acetic acid and zeatin levels in relation to regulation of apical dominance and tiller release in oat shoots, *J. Plant Growth Regul.*, 2:215–223, 1983; M. A. Harrison and P. B. Kaufman, The role of hormone transport and metabolism in apical dominance in oat, *Bot. Gaz.*, 145:293–297, 1984; I. D. J. Phillips, Apical dominance, *Annu. Rev. Plant Physiol.*, 26:341–367, 1975; T. K. Prasad and M. G. Cline, The role of gravity in apical dominance: Effects of clinostating on shoot inversion-induced ethylene producion, shoot elongation and lateral bud growth, *Plant Physiol.*, 83:505–509, 1987; H. E. Street and H. Opik (eds.), *The Physiology of Flowering Plants: Their Growth and Development*, 3d ed., 1992; P. F. Wareing and I. D. J. Phillips, *Growth and Differentiation in Plants*, 3d ed., 1981.

Apical meristem

Permanently embryonic tissue involved in cell division at the apices of roots and stems, and forming dynamic regions of growth. These apical meristems, usually consisting of small, densely cytoplasmic cells, become established during embryo development. Thereafter they divide, producing the primary plant body of root and shoot. Below the apical meristems, tissue differentiation begins; the protoderm gives rise to the epidermal system, the procambium to the primary vascular system, and the ground meristem to the pith and cortex (**Figs. 1** and **2**). Plant apical meristems have been the object of experiments on development similar to those carried out on animal embryos.

Roots. Root apical meristem is covered by a root cap, a region of parenchymatous, often starch-containing cells which has a protective function and is responsible for perceiving gravitational changes. In many monocotyledons, including grasses, the root cap has a separate initiating region, the calyptrogen. In many vascular cryptogams the root apex has a prominent tetrahedral apical cell (**Fig. 3**). In flowering plants, root apices have multicellular apical meristems. Root tips have been shown to possess a central region, usually hemispherical, which consists of cells which rarely divide or synthesize deoxyribonucleic acid (DNA), and have less ribonucleic acid (RNA) and protein than adjacent cells; this region is known as the quiescent center (**Fig. 4**). The cells which divide and give rise to root tissues

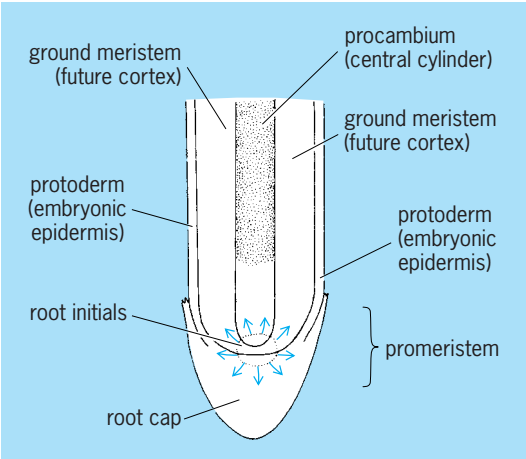

Fig. 1. Diagram of a root apical meristem. Cortex and central cylinder have separate initials; epidermis and root cap have a common origin.

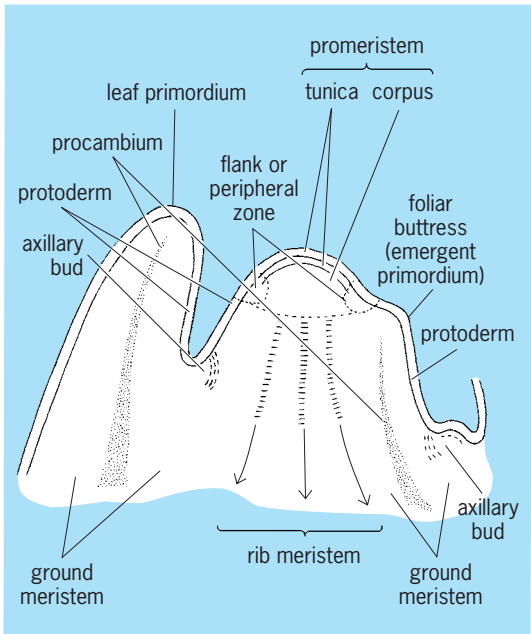

Fig. 2. Diagram of a hypothetical shoot apical meristem with a two-layered tunica.

lie around the periphery of this region, and are sometimes described as constituting a proximal and distal meristem (Fig. 4). Cells in the quiescent center are regarded as cells that are mitotically young and genetically sound; they can renew the initial cells from time to time. Surgically isolated quiescent centers of corn roots grown in sterile nutrient medium could form whole roots.

Single apical cells of cryptogams were considered to become quiescent during development and to represent a quiescent center, but work on the aquatic fern *Azolla* demonstrated conclusively that the apical cell functions as an initial, dividing repeatedly in a very regular fashion to give rise to merophytes, structural units whose origin and subsequent differentiation can be accurately followed.

Lateral roots. Primordia of lateral roots have an endogenous (internal) origin in deep-seated tissues of the parent root, usually the pericycle, at some distance from the root apex. In cryptogams, the endodermis may contribute to lateral root formation. In some species, lateral roots arise in greater profusion, close to the apical meristem and to one another (**Fig. 5***a*, *b*); they are then usually formed according to a more or less regular pattern. Adventitious roots may be formed on stems or leaves, usually after excision from the plant; their meristems originate by dedifferentiation of mature stem of leaf tissues. This occurs in propagation by cuttings. *See* ENDODERMIS; PERICYCLE.

Shoots. Shoot apices vary greatly in size and shape. The diameter can vary from about 50 micrometers in *Zea* to 0.14 in. (3.5 mm) in *Cycas revoluta*; the shape may be elongated and conical, dome-shaped, flat, or even slightly concave. The distance from the center of the apex to the youngest leaf primordium

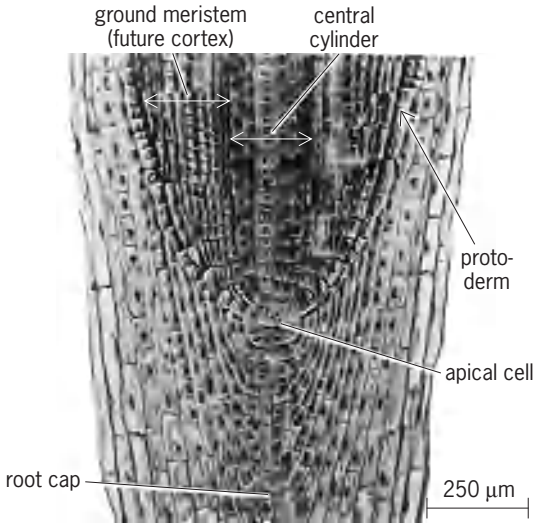

Fig. 3. Median longitudinal section of the root apical meristem of *Equisetum*.

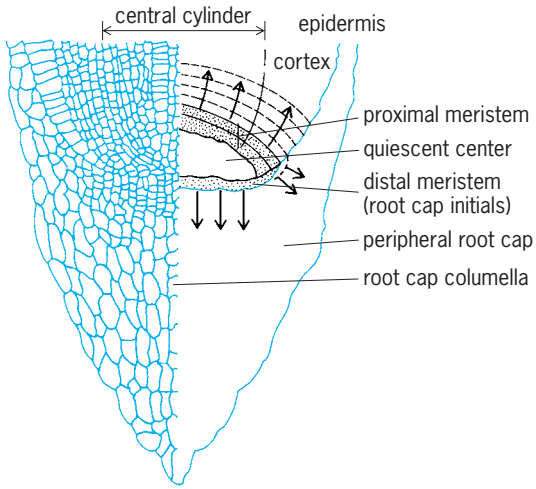

Fig. 4. Longitudinal section of corn root tip, showing the quiescent center with the proximal and distal meristems at its periphery. (*After J. G. Torrey and L. J. Feldman, The organization and function of the root apex, Amer. Sci., 65:334–344, 1977*)

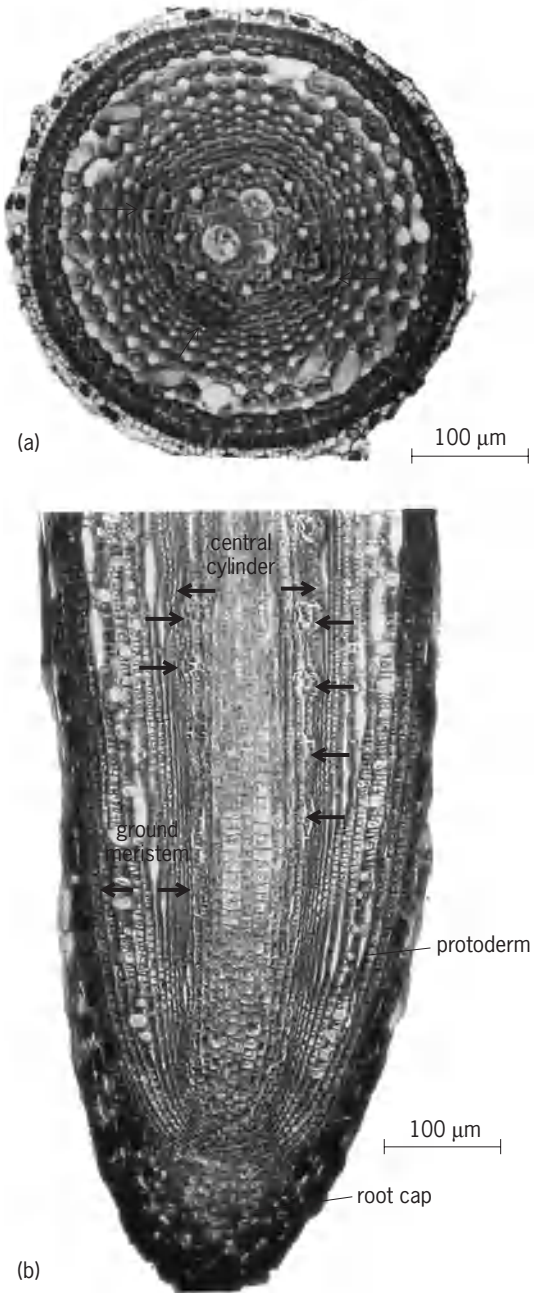

Fig. 5. Root tip of *Pontederia cordata*: (*a*) cross section and (*b*) longitudinal section showing the endogenous origin of lateral root primordia (arrows), and the various tissue regions of the root. (*From W. A. Charlton, Distribution of lateral roots and pattern of lateral initiation in Pontederia cordata L., Bot. Gaz., 136:225–235, 1975*)

also varies considerably. Apices increase in size during the development of a single plant; for example, the apical meristem of flax *(Linum usitatissimum)* increases in area 20-fold from the seedling up to the time of flowering. Apices may also change in size during the time (plastochron) between the formation of one leaf primordium and the next.

A single apical cell is present in shoot apices of bryophytes and vascular cryptogams (**Fig. 6**); however, surrounding cells are also mitotically active, and these plants have multicellular apical meristems. In flowering plants, the outer layer or layers of cells (tunica) may divide predominantly by walls at right angles to the surface; the inner tissue (corpus), in less regular planes (**Figs. 7** and **8**). Regions of the apical meristem may react differently to certain stains, reflecting so-called cytohistological zonation. One type of zonation may be superimposed upon another.

Cells in the central terminal region of the vegetative shoot apex divide less actively than those on the flanks or at the periphery, where leaf primordia are formed. Various surgical experiments, involving incision of the apical meristem, have shown that new apices can be regenerated from portions of the flank. Excised apical meristems, devoid of leaf primordia, can be successfully grown on agar nutrient medium, in the presence of auxin, and will eventually yield new plants.

Leaf primordia. Leaf primordia are formed on the flanks of the shoot apex, in an arrangement or phyllotaxy characteristic of the species. Usually, periclinal divisions (walls parallel to the surface) occur in

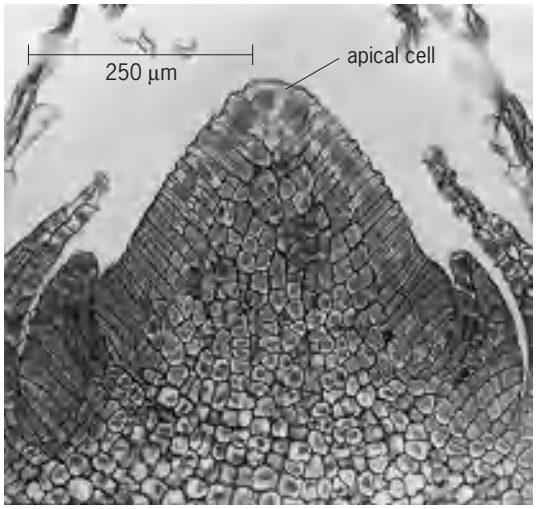

Fig. 6. Median longitudinal section of the shoot apical meristem of *Equisetum*, showing apical cell.

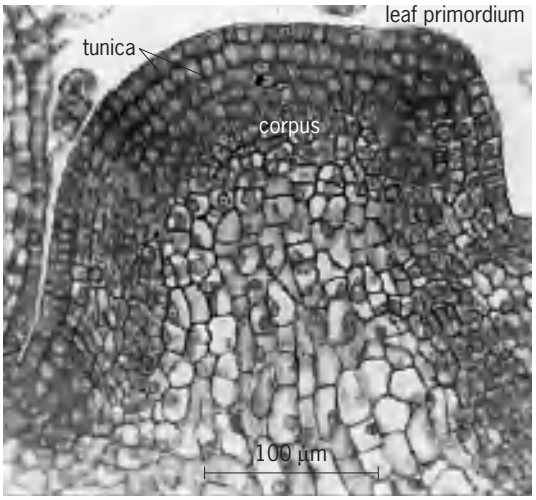

Fig. 7. Longitudinal section of the shoot apical meristem of potato, *Solanum tuberosum*, showing a two-layered tunica and a young leaf primordium on the right.

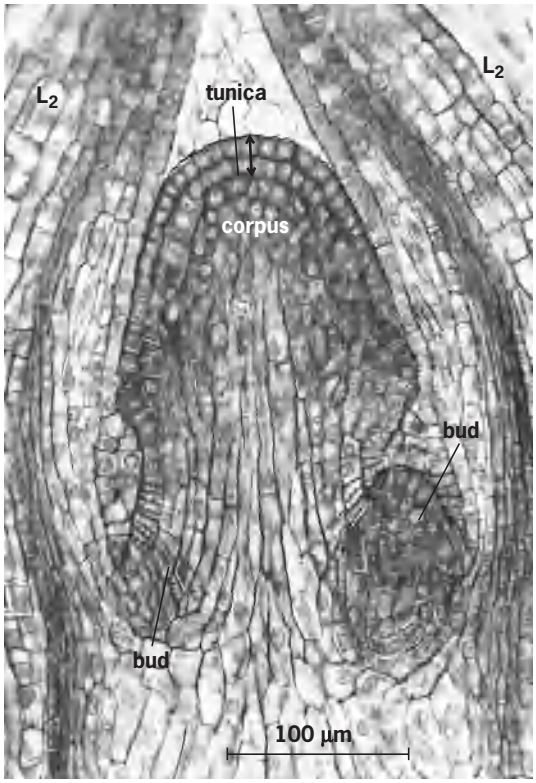

Fig. 8. Longitudinal section of the shoot apical meristem of *Alternanthera philoxeroides*, showing a two-layered tunica. Unequal axillary buds are present in the axils of the second-youngest pair of leaves, L_2. (*From E. G. Cutter, Morphogenesis and developmental potentialities of unequal buds, Phytomorphology, 17:437–445 1967*)

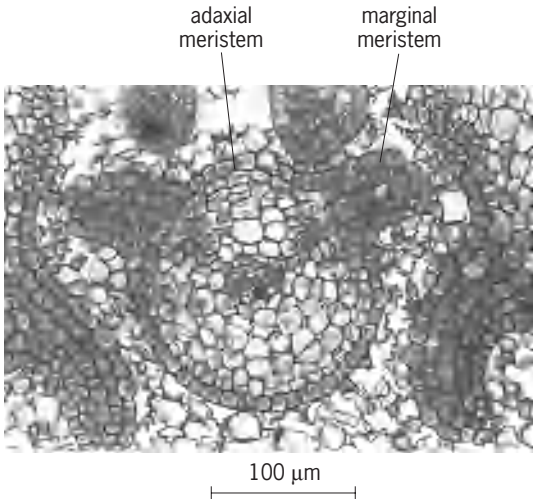

Fig. 9. Transverse section of a young leaf primordium of *Alternanthera philoxeroides*, showing marginal and adaxial meristems. (*From E. G. Cutter, Plant Anatomy, pt. 2: Organs, Edward Arnold Ltd., London, 1971*)

the second and perhaps deeper layers of cells, in localized positions (Fig. 7). If the leaf is sheathing when mature (as in grasses), mitotic activity spreads around the circumference of the apex. The leaf primordium grows apically at first, but in angiosperms apical growth is short-lived and is followed by intercalary growth; in fern leaf primordia, apical growth is prolonged. Marginal meristems divide to form the leaf blade or lamina (**Fig. 9**); an adaxial meristem may contribute to the thickness of the leaf. *See* LEAF.

Lateral buds. In most angiosperms, in addition to leaf primordia, the apical meristem, or its derivatives, gives rise to bud primordia. These usually occupy axillary positions (that is, above the junction of the leaf with the stem), and are exogenous (that is, originate superficially), sometimes developing from pockets of meristematic tissue separated from the apical meristem by more mature, differentiated tissue (Fig. 8). In some plants, buds occurring in the axils of a pair of opposite leaves may be of unequal size and growth potentiality (Fig. 8). Sometimes groups of buds are present. A few plants do not branch, and no bud primordia are formed. Branching may also occur by equal division, or dichotomy, of the shoot apex; this is common in cryptogams, and has now been reported in a number of angiosperms. Bud primordia are usually inhibited to a greater or lesser degree by the parent shoot apex.

Adventitious buds are formed on root cuttings (as in raspberry) or leaf cuttings (as in *Begonia*), and

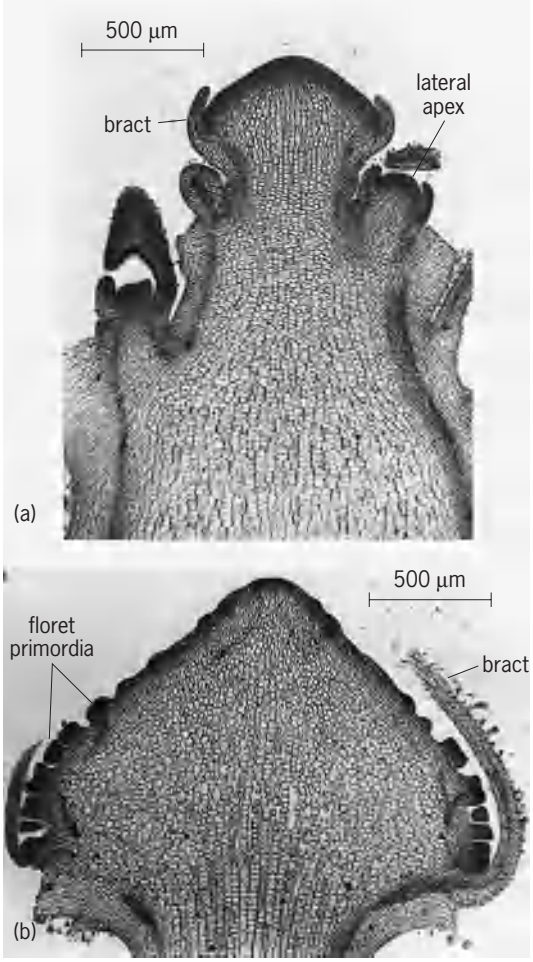

Fig. 10. Longitudinal sections of inflorescence meristems of *Calendula*. (*a*) Dome-shaped inflorescence apex with bract primordia at the periphery; a vegetative or transitional lateral apex is present on the right; note the difference in size. (*b*) Older inflorescence apex with floret primordia, at same magnification. Note increase in both width and height. (*From E. G. Cutter, Plant Anatomy pt. 2: Organs, Edward Arnold Ltd., London, 1971*)

they may be exogenous or endogenous in origin. *See* BUD.

Primary thickening meristems. These meristems bring about growth in thickness of many monocotyledons, and a few dicotyledons which have little secondary growth. They consist of files of elongated cells (somewhat resembling cambium in appearance) on the "shoulders" of the shoot tip, usually below the leaf primordia; by their division, they add to the width of the shoot tip and sometimes raise the periphery of the apical region so that the apex itself lies in a depression.

Floral meristems. Under appropriate environmental stimuli (such as daylength), vegetative shoot apices may be induced to develop as apices of inflorescences or individual flowers. This involves a change from indeterminate to determinate growth. Indeed, the transition to flowering involves many changes, for example in size, shape, rate of mitosis, or amount of RNA. Inflorescence apices, such as those of the garden marigold, *Calendula*, may increase markedly in size, become more steeply domed, and finally form floret primordia (**Fig. 10***a*, *b*). These dramatic changes may occur in a matter of hours. Floral apices give rise to a sequence of primordia of floral organs (sepals, petals, stamens, carpels), many of which originate in a manner similar to leaf primordia. *See* FLOWER; INFLORESCENCE. Elizabeth G. Cutter

Bibliography. P. W. Barlow, Towards an understanding of the behaviour of root meristems, *J. Theoret. Biol.*, 57:433–451, 1976; E. M. Gifford, Jr., and G. E. Corson, Jr., The shoot apex in seed plants, *Bot. Rev.,*37:143–229, 1971; J. D. Mauseth, *Plant Anatomy*, 1988; J. G. Torrey and L. J. Feldman, The organization and function of the root apex, *Amer. Sci.*, 65:334–344, 1977.

Aplacophora

A class of vermiform marine mollusks covered by a cuticle invested with innumerable calcareous spicules; also known as solenogasters. There are 277 named species. Most species are less than 0.4 in. (10 mm) in length, but a few attain 12 in. (300 mm). The head with a mouth is poorly differentiated; there is a small posterior mantle cavity, or cloaca. Despite their specialized shape (see **illus.**), Aplacophora retain a primitive molluscan condition in the radula, integument, gonad-pericardium, muscles, and nervous system. There are two distinct subclasses, the creeping neomenioids (Neomeniomorpha) with a narrow foot within a ventral groove, and the burrowing chaetoderms (Chaetodermomorpha) without either a groove or foot, and with a cuticular shield around the mouth. Aplacophora are continental-shelf and deep-sea forms found from subtidal depths to hadal depths (30,000 ft or 9000 m), and in some rare localities may be dominant organisms. *See* CHAETODERMOMORPHA; MOLLUSCA; NEOMENIOMORPHA. Amelie H. Scheltema

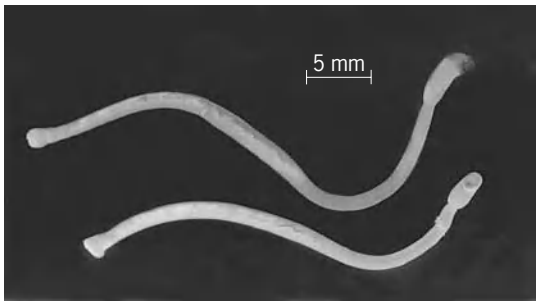

Living *Chaetoderma nitidulum*. (*Courtesy of R. Robertson*)

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; A. H. Scheltema, Ancestors and descendents: Relationships of the Aplacophora and Polyplacophora, *Amer. Malacol. Bull.*, 6:57–68, 1988; A. H. Scheltema, Position of the class Aplacophora in the phylum Mollusca, *Malacologia*, 17(1):99–109, 1978; E. R. Trueman and M. R. Clarke (eds.), *The Mollusca*, vol. 10: *Evolution*, 1985.

Aplite

A fine-grained, sugary-textured rock, generally of granitic composition; also any body composed of such rock. *See* GRANITE.

The light-colored rock consists chiefly of quartz, microcline, or orthoclase perthite and sodic plagioclase, with small amounts of muscovite, biotite, or hornblende and trace amounts of tourmaline, garnet, fluorite, and topaz. Much quartz and potash feldspar may be micrographically intergrown in cuneiform fashion.

Aplites may form dikes, veins, or stringers, generally not more than a few feet thick, with sharp or gradational walls. Some show banding parallel to their margins. Aplites usually occur within bodies of granite and more rarely in the country rock surrounding granite. They are commonly associated with pegmatites and may cut or be cut by pegmatites. Aplite and pegmatite may be gradational or interlayered, or one may occur as patches within the other. *See* PEGMATITE.

Aplites form in different ways. Some represent granite or pegmatite which recrystallized along fractures and zones of shearing. Others are of metasomatic (replacement) origin. Many form from residual solutions derived from crystallizing granitic magma (rock melt). If these fluids retain their volatiles, pegmatites may form. If the volatiles escape, a more viscous fluid may be created, and a fine-grained (aplitic) texture may be developed. *See* IGNEOUS ROCKS; MAGMA; METASOMATISM. Carleton A. Chapman

Apoda

The smallest order of class Amphibia. The order Apoda is sometimes called Gymnophiona and is known by the common name "caecilian." Caecilians

Mexican burrowing caecilian (*Dermophis mexicanus*). (*Photo © 2003 J. A. Campbell*)

are wormlike, legless amphibians with indistinct or even hidden eyes. A series of annular grooves (annuli) are usually present along the length of the body, heightening the resemblance to an earthworm (see **illustration**). Most caecilians lead a burrowing existence, though members of one genus, *Typhlonectes*, are aquatic. Some species have the eyes hidden beneath the bones of the skull and probably are blind, but others at least are able to distinguish movement. A unique feature of some caecilians among modern Amphibia is the presence of scales buried in the skin.

There are more than 165 species (6 families, 33 genera) of caecilians confined to tropical regions of both the Eastern and Western hemispheres. Due to the generally cryptic nature of caecilians, the true species richness may be much greater. Many species are less than 1 ft (30 cm) in length, but three species of the genus *Caecilia* grow to over 3 ft (90 cm). The breeding habits of caecilians are poorly known. Caecilians are internal fertilizers, and some species lay eggs, while others bring forth their young alive. The embryos of the species that bear living young are nourished in the later part of their embryonic development by "uterine milk," which is secreted by the mother. In some of the species that lay eggs, there is an aquatic larval stage. Caecilians are carnivorous, but little is known of their food habits. Captive specimens have fed on earthworms, and in the natural state caecilians have been known to eat lizards.

Fossil caecilians are rare. The earliest known were found in Arizona and date from the Early Jurassic period (206 million years before present). Among the significant characteristics of this fossil series was the existence of girdles that supported limbs. It was once proposed that caecilians share a closer relationship to salamanders (sister group) than to frogs, but recent molecular phylogenetic data support a more distant relationship of caecilians to other amphibians. New fossil representatives and continued molecular data will shed increasingly clear light on the phylogeny of Order Apoda. *See* AMPHIBIA; ANURA; URODELA. Richard G. Zweifel; W. Ben Cash

Bibliography. D. W. Linzey, *Vertebrate Biology*, McGraw-Hill, New York, 2001; F. H. Pough et al., *Herpetology*, 3d ed., Prentice Hall, Upper Saddle River, NJ, 2004; E. H. Taylor, *The Caecilians of the World: A Taxonomic Review*, 1968; R. Zardoya and A. Meyer, On the origin of and phylogenetic relationships among living amphibians, *Proc. Nat. Acad. Sci.*, 98(13):7380–7383, 2001; G. R. Zug, L. J. Vitt, and J. P. Caldwell, *Herpetology*, Academic Press, San Diego, 2001.

Apodiformes

An order of birds consisting of two dissimilar groups, the swifts (Apodi) and the hummingbirds (Trochili). These birds have been placed together because of several anatomical specializations of the wings and feathers, although an alternative classification based on morphological and molecular characters places the swifts and hummingbirds in two separate orders. Swifts and hummingbirds have strong wings and are excellent fliers but have small, weak feet and cannot walk. Some of these features may be convergences because of the rapid but different flight of these birds. The two groups share a unique crossover structure of a superficial neck muscle which cannot be related to their ways of life. The relationship of the swifts and hummingbirds to other birds is obscure.

Classification. The order Apodiformes is divided into the suborder Apodi, containing the families Aegialornithidae (fossils), Hemiprocnidae (crested swifts; 4 species), and Apodidae (true swifts; 83 species), and the suborder Trochili, containing the single family Trochilidae (hummingbirds; 341 species). The Aegialornithidae are primitive swifts from the Early Eocene of Europe. Otherwise the fossil record of swifts and hummingbirds is very poor.

Characteristics. Swifts are fast-flying, aerial birds with dull, hard plumage; long, curved, pointed wings; and a short, broad, weak bill with a wide gape, adapted to catching insects in flight. They rest by clinging to cliffs, hollow trees, and other vertical surfaces. Their nest is composed of sticks and other plant material glued together and to vertical surfaces, with the extreme condition being a nest built completely of their gluelike mucus. The clutch varies from one to six eggs, and the young are cared for by both sexes until they fly from the nest. Swifts are found worldwide except at high latitudes. True swifts (Apodidae) never perch crosswise on branches, but crested swifts (Hemiprocnidae), found in tropical Asia to New Guinea, are able to perch on branches.

The hummingbirds (Trochilidae) are small, brightly colored, nectar-feeding birds found only in the Western Hemisphere, with most species living south of the United States–Mexican border. The bill is slender and varies in length and shape, correlated closely with the shape of the flowers utilized by each species. They have a rapid wing beat and flight, and are able to hover in front of a flower while feeding or even fly backward (see **illustration**). Hummingbirds are attracted to the color red, and flowers specialized on hummingbirds for cross-pollination are

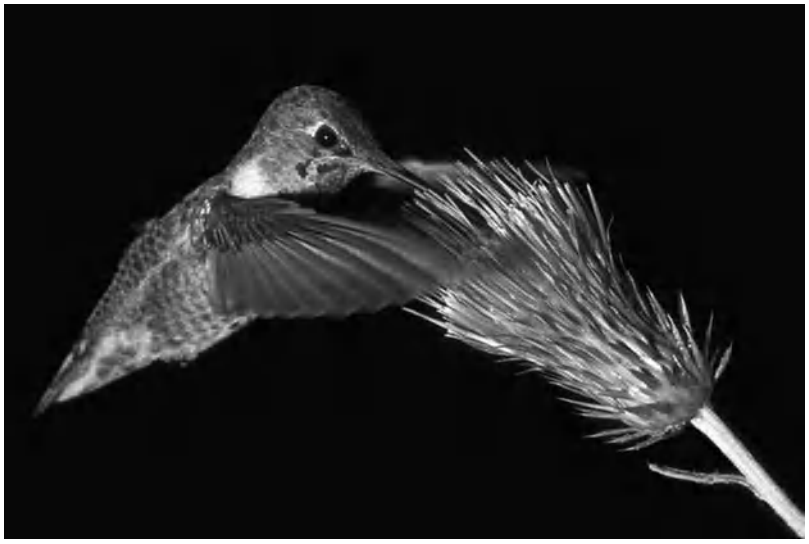

Anna's hummingbird (*Calypte anna*). (*Photo by Dr. Lloyd Glenn Ingles; © 1999 California Academy of Sciences*)

red. Hummingbirds feed on nectar and insects, and can aggressively defend a feeding territory. They are among the smallest birds, including the bee hummingbird (*Lisugsa belenae*) of Cuba, the smallest of all birds. The diminutive size of hummingbirds results in large loss of heat because of their small body mass compared to surface area. If they are in poor physiological condition, hummingbirds may hibernate overnight to save energy. Moreover, high-altitude species place their nests under an overhanging tree branch or in a cave to reduce radiational cooling during the night. The brilliant colors of the males serve both for courtship, including species recognition, and for defense of their feeding territories. Females of a few species also possess bright colors when they defend feeding territories. Males have a courtship flight and sometimes a "song," but are not involved in nest-building, incubation, or feeding the young. The deep cup nest is built on a branch and consists of plant down and spider web. The clutch is usually two eggs, and the young are cared for by the female until they can fly from the nest.

Economic importance. A few species of cave swiftlets (such as *Collocalia brevirostris*) are economically important in providing nests that serve as the base for "birds-nest soup." Prior to the early decades of the twentieth century, hummingbirds were collected extensively to use as decoration on women's hats, a practice, long over, that severely affected the status of many species. Hummingbirds are important in ecotourism and in backyard bird watching, particularly in the western United States, where many people set out hummingbird feeders. *See* AVES.

Walter J. Bock

Bibliography. P. Chantler, *A Guide to the Swifts and Treeswifts of the World*, 2d ed., 1999; K.-L. Schuchmann, Family Trochilidae (Hummingbirds), pp. 468–535, in J. Del Hoyo et al. (eds.), *Handbook of the Birds of the World*, vol. 5, Lynx Edicions, 1999.

Apophyllite

A hydrous calcium potassium silicate containing fluorine. The composition is variable but approximates to $KFCa_4(Si_2O_5)_4 \cdot 8H_2O$. It resembles the zeolites, with which it is sometimes classified, but differs from most zeolites in having no aluminum. It exfoliates (swells) when heated, losing water, and is named from this characteristic; the water can be reabsorbed. The mineral decomposes in hydrochloric acid, with separation of silica. It is essentially white, with a vitreous luster, but may show shades of green, yellow, or red. The symmetry is tetragonal and the crystal structure contains sheets of linked SiO_4 groups, and this accounts for the perfect basal cleavage of the mineral (see **illus.**). It occurs as a

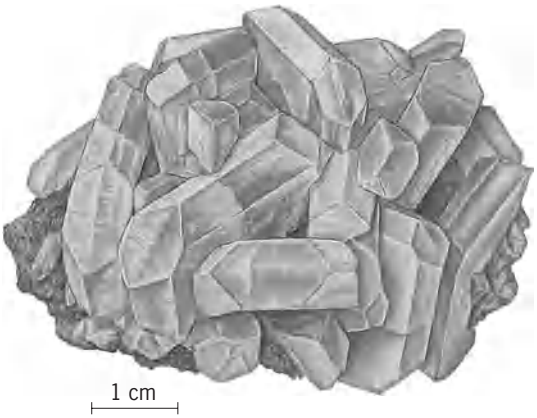

Apophyllite crystals with basalt from French Creek, Pennsylvania. (*From Department of Geology, Bryn Mawr College*)

secondary mineral in cavities in basic igneous rocks, commonly in association with zeolites. The specific gravity of apophyllite is about 2.3–2.4, the hardness is 4.5–5 on Mohs scale, the mean refractive index is about 1.535, and the birefringence is 0.002. *See* SILICATE MINERALS; ZEOLITE.

George W. Brindley

Apoptosis

Cell death triggered by extracellular signals or genetically determined events and carried out by physiological processes within the cell. It is therefore sometimes called "programmed cell death" as opposed to uncontrolled cell death following acute injury. Apoptosis is characterized by a systematic breakdown of cellular structure, including cytoplasmic shrinkage, breakdown of chromatin and structural proteins of the cell nucleus, and blistering (blebbing) of the cell membrane.

Cell death is a fundamental concept in biology and medicine, and it is characteristic of many diseases. Cell death can take place through the mechanisms of apoptosis or necrosis. While necrotic cell death has been known for a long time, not until the 1970s

was apoptotic cell death clearly recognized. Apoptosis requires energy, and it can be regulated by the cell itself. Necrosis is passive cell death, and it is usually caused by an outside factor, such as loss of blood supply. In necrosis, a large group of cells die simultaneously in the same area, and the process is often associated with inflammation and damage to the surrounding tissues. In apoptosis, individual cells die, and the fragmented cell is removed by surrounding healthy housekeeping cells, the macrophages, making the process neat; no harm is done to the surrounding tissues. *See* CELL (BIOLOGY); CELL NUCLEUS; INFLAMMATION; NUCLEOSOME.

Characteristics. Apoptotic cell death generally takes 12–24 h. It is divided into three phases: initiation, effector, and degradation. In the initiation phase, the apoptotic stimulus is introduced to the cell; depending on the strength and the nature of the stimulus, this phase may take several hours. In the effector phase, the apoptotic stimulus is modulated in the cell; this phase is still reversible (the cell can be salvaged). In the degradation phase, the cell is inevitably doomed to die; degradation of cellular proteins and fragmentation of nucleic acids in the cell (see **illustration**) characterize this phase, which normally takes 1–2 h.

Through apoptosis the organism can regulate the number of cells in a given tissue or destroy any unwanted or damaged cells without harming adjacent tissues. Apoptosis is already present at the embryonic stage, where it participates in the formation of different organs and tissues. Human digits, for instance, are formed because cells between the digit anlages die through apoptosis during embryonic development. Thus, apoptosis is the sculptor, which by deleting unnecessary cells gives rise to the appearance of organisms.

Apoptosis participates in the development of the human immune system. The immune system consists of white cells which attack and destroy foreign elements introduced into the human body. Early in development, the defense system is nonselective and consists of white blood cells which are not able to distinguish foreign molecules from self-derived molecules. When the defense system matures, white cells that cannot make that distinction are destroyed in the thymus through apoptosis. In adult tissues, a restricted loss of such a self-tolerance may sometimes take place (that is, harmful, self-directed white cells develop and fail to undergo apoptosis); the individual then suffers from an autoimmune disease. In such a disease, these white blood cells attack some of the host's own tissue components and induce apoptosis. Apoptosis thus functions to protect the host by participating in the development of the immune system, but it may also contribute to the development of an autoimmune disease.

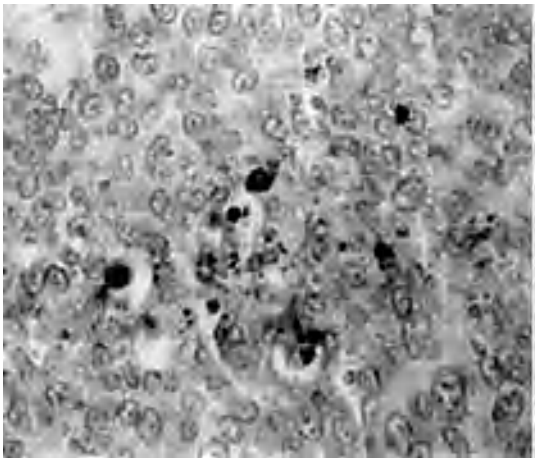

Several apoptotic cells and fragmented apoptotic bodies in cancer cells (center area of the figure) located by the TUNEL method, which detects apoptotically fragmented nucleic acids in tissue sections.

Activation and inhibition. Apoptosis can be triggered by several factors, including irradiation, heat, several chemical drugs, and toxic compounds. Some of these factors may cause genetic alterations. If such alterations occur, the cells are often capable of correcting such damage, but if this fails, an apoptotic cell death program is stimulated. Apoptosis can also be induced through stimulation or inhibition of certain types of cell surface receptor molecules. A receptor serves as a sensor of the outside milieu of the cell. Hormone-dependent cells, such as breast or prostate cells, have hormonal receptors on their surface; the well-being of these cells depends on a constant stimulation of the receptors. If the receptors are blocked, the cells suffer an apoptotic death. For example, when receptors fail to respond to certain growth-promoting molecules (such as interleukins), apoptosis is initiated. Some cells, such as white blood cells, harbor ligand molecules, which are able to stimulate receptors that induce apoptosis. Inside the cell, two families of proteins, the bcl-2's and the caspases, regulate apoptosis. The bcl-2 family proteins operate in the effector phase of apoptosis and may either promote or inhibit apoptosis. A balance between individual bcl-2 family proteins determines whether the cell will live or die. Caspases are enzymes which operate at the degradation phase of apoptosis. Their activation leads to degradation of cellular proteins and fragmentation of cellular nucleic acids. Recently, other groups of proteins involved in apoptosis have been discovered. Inhibitor of apoptosis proteins (IAPs) are capable of inhibiting caspases. There are also proteins, Fas-associated death-domain-like interleukin-1 (IL-1) beta converting enzyme inhibitory proteins (FLIPs), which inhibit receptor-mediated apoptosis early in the signaling cascade. *See* TUMOR SUPPRESSOR GENE.

Pathology. Apoptosis takes part in the development of several diseases. These include viral and other infections, autoimmune diseases, many neurodegenerative diseases such as Alzheimer's and Parkinson's diseases, and several metabolic and endocrinologic diseases. In these diseases, apoptosis is usually increased, leading to a loss of cells. In Alzheimer's disease, loss of nerve cells in the brain leads to loss of brain matter and deterioration of the intellectual function. Alzheimer's disease is characterized by a gradual accumulation of an aberrant

protein in the neuronal cells called amyloid. Genes which produce this protein also influence apoptosis, and it has been suggested that dysregulation of apoptosis plays a primary role in causing the neuronal loss seen in Alzheimer's disease.

Apoptosis is also operating in vascular diseases such as ischemic heart disease. This disease is characterized by formation of plaques in the vessel walls, eventually leading to obstruction of the blood flow. A great amount of apoptosis has been shown in cells within such plaques. It has been suggested that apoptosis may contribute to the softening of plaques through deletion of cells capable of synthesizing hard collagenous tissue; this softening makes plaques more prone to rupture. A plaque rupture precipitates the formation of an obstructive clot in the blood vessel which leads to an acute heart infarct. When this happens, a part of the heart muscle nourished by the obstructed vessel is destroyed by necrosis. However, some cells at the marginal zone of the infarcted heart muscle undergo apoptosis. By different treatment modalities at an early stage of an acute heart infarct, one may diminish the size of the damaged tissue and save cells from death because at an early stage of apoptosis cells can still be revived and brought back to normal. Apoptosis also plays a role in the further development of ischemic heart disease. As a consequence of heart infarct, some people develop heart failure; one reason is that the remaining heart cells, which are under continuous strain, slowly undergo apoptosis. This apoptosis gradually leads to a decrease in the number of functioning muscle cells in the wall of the heart and development of a failing heart. Decreased apoptosis may also cause disease. For example, some types of lymphomas arise due to deranged apoptotic machinery and overproduction of a protein, which lengthens cell survival. In many other types of cancer, apoptosis is increased. The apoptotic rate in different cancer types may vary, but usually about 1–2% of cancer cells show apoptosis. Cancers arise, in part, because even though cell death in cancer tissues is high, the rate of cell division is higher. One method to treat cancer would be to increase the proportion of apoptotic cells. In fact, treatments directed at increasing apoptosis are already in use. Both irradiation of tumors and cancer chemotherapy lead to apoptosis of cancer cells, and thus the tumors grow more slowly. In hormone-dependent tumors (such as breast cancer), cancer cells are deprived of hormonal stimulus by antiestrogenic treatment, resulting in their apoptosis. Unfortunately, cancer cell populations are heterogeneous, and such treatment destroys only a part of them. The remaining cells are usually resistant to the treatment and must be dealt with in some other way. More selective treatment modalities, including transfection of viral vectors expressing apoptotic genes, are under development. It is plausible that through increased knowledge of apoptosis a cure for some types of cancer will become possible in the near future. In addition, researchers will be able to develop better treatments for other diseases where apoptosis plays a role. *See* ALZHEIMER'S DISEASE; CANCER (MEDICINE); CELL (BIOLOGY); DEVELOPMENTAL BIOLOGY; GENETICS; TUMOR.

Ylermi Soini

Bibliography. J. M. Adams and S. Cory, The Bcl-2 protein family: Arbiters of cell survival, *Science*, 281:1322–1326, 1998; B. Alberts et al., *Molecular Biology of the Cell*, 4th ed., Garland Publishing, 2002; A. Haunstetter and S. Izumo, Apoptosis: Basic mechanisms and implications for cardiovascular disease, *Circ. Res.*, 82:1111–1129, 1998; J. Jenkins and E. P. Reddy, Reviews: Apoptosis, *Oncogene*, 17:3203–3399, 1998; J. F. Kerr, A. H. Wyllie, and A. R. Currie, Apoptosis: A basic biological phenomenon with wide-ranging implications in tissue kinetics. *Brit. J. Cancer*, 26(4):239–257, 1972; D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 4th ed., Worth, 2004; N. A. Thornberry and Y. Lazebnik, Caspases: Enemies within, *Science*, 281:1312–1316, 1998.

Apostomatida

A little-studied group of ciliates comprising an order of the Holotrichia. The majority occur as commensals on marine crustaceans. This association is generally more intimate and complex than that of the chonotrichs found on related hosts. Their life histories may become exceedingly complicated, and they appear to bear a direct relationship to the molting cycles of their hosts. Apostomes are particularly characterized by the presence of a unique rosette in the vicinity of an inconspicuous mouth opening and the possession of only a small number of ciliary rows wound around the body in a spiral fashion. *Foettingeria* (see **illus.**) is an example commonly encountered. *See* CILIOPHORA; HOLOTRICHIA; PROTOZOA.

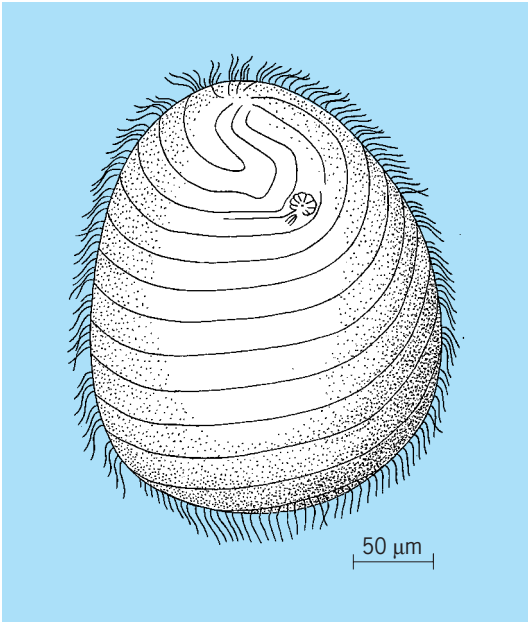

***Foettingeria*, an example of an apostomatid.**

John O. Corliss

Appendicitis

An inflammation of the vermiform appendix. Acute appendicitis is the most common cause of emergency abdominal surgery, occurring in 5–6% of the population of the United States. Although about 80% of the individuals are between 5 and 35 years of age, the disease does occur in the very young as well as in the elderly. In the latter groups, the clinical picture is atypical, and frequently results in a significant delay in diagnosis with a higher incidence of complications. Acute appendicitis is rare in children less than 1 year of age. *See* APPENDIX (ANATOMY).

Cause. Acute appendicitis develops when the lumen of the appendix becomes obstructed, usually by fecal material, a foreign body, or hyperplasia of lymphatic tissue that is normally present in the wall of the appendix. The obstructed appendix becomes distended because of continued secretion of mucus by the lining cells. Distention eventually interferes with the blood supply of the appendix, and reduced blood supply leads to the development of areas of cell death with eventual gangrene and perforation. There is also a proliferation of bacteria within the obstructed appendix, a factor that contributes to infection as gangrene develops.

Typically, acute appendicitis progresses from obstruction of the lumen and distention of the appendix to spread of the inflammation beyond the appendix. Initially, there is localized peritonitis confined to the area of the appendix. If unrecognized and untreated, this may progress to an inflammatory mass or abscess, or to perforation of the appendix with resultant diffuse peritonitis, generalized toxic reaction, and even death. Each year, as many as 10,000 deaths in the United States are attributed to appendicitis and its complications, although it is a much rarer cause of significant morbidity and mortality since the introduction of antibiotics. *See* PERITONITIS.

The usual progression of symptoms includes pain in the region around the navel; loss of appetite, nausea, and occasionally vomiting; localization of the pain to the right lower quadrant of the abdomen; and mild fever. Although the pain typically is localized in the right lower quadrant of the abdomen, there are variations because the appendix may be located in a number of other positions within the abdominal cavity. Fever is a fairly late sign, with mild elevation the rule; a high fever increases the suspicion of perforation or of some other inflammatory process.

Diagnosis and symptoms. The diagnosis of appendicitis is generally made by history and physical examination, although laboratory and radiologic studies may be helpful in differentiating appendicitis from other conditions. The condition must be considered in any case of abdominal pain, especially involving the lower abdomen. Appendicitis is suspected when there are signs of peritoneal inflammation in the right lower quadrant of the abdomen, specifically tenderness on examination by touch and diminished bowel sounds. Pain may also accompany examination of the rectum or the pelvis. With progression of the disease beyond 12–18 h, the incidence of perforation along with secondary complications of infection and a generalized toxic reaction increases significantly. To avoid this crisis, early surgery is essential.

Treatment. The treatment of acute appendicitis is prompt surgical removal of the inflamed appendix. Prior to surgery, the patient may be given intravenous fluids to correct dehydration and electrolyte imbalances. The use of antibiotics before surgery to decrease wound infection is often recommended. Antibiotics are continued after surgery in cases where the inflammation has extended beyond the appendix. Delay in removal of the appendix increases the chance of perforation. Death following appendectomy for acute appendicitis is less than 0.5% overall, but does increase to about 15% in the elderly. This increase is related to delay in diagnosis, associated cardiopulmonary complications, and the frequent presence of unrelated major diseases. The incidence of complications in appendicitis is 10–15%, but increases to 30–40% when perforation is present. The most common complication of appendectomy for acute appendicitis is infection. To reduce the risk of such complications, prompt diagnosis and surgery are necessary. *See* ANTIBIOTIC; GASTROINTESTINAL TRACT DISORDERS. Anthony L. Imbembo

Bibliography. D. Law, R. Law, and B. Eiseman, The continuing challenge of acute and perforated appendicitis, *Amer. J. Surg.*, 131:533, 1976; F. R. Lewis et al., Appendicitis: A critical review of diagnosis and treatment in 1000 cases, *Arch. Surg.*, 110:677, 1975.

Appendicularia (Larvacea)

A class of marine, planktonic animals in the subphylum Tunicata. This class is characterized by the persistence of a tail, notochord, gill slits, and dorsal nerve cord throughout life and by a unique feeding structure, the "house." Appendicularians are free-swimming, solitary animals that are believed to have evolved through neoteny from the tadpole larvae of a bottom-dwelling, ascidianlike, ancestral tunicate. They resemble a bent tadpole with a flat, muscular tail and a trunk containing all major organs. Included are a complex digestive system, reproductive organs, two ciliated openings of the gill slits (the spiracles) leading to a mucus-lined pharynx, a mucus-producing gland called the endostyle, and a simple circulatory system with a single, muscular heart (**Fig. 1**). *See* NEOTENY.

All appendicularians, except the species *Oikopleura dioica*, are hermaphroditic; that is, each individual bears the reproductive organs of both sexes. While sperm may be liberated from the testes via tiny ducts to the outside, release of eggs from the ovary occurs only when the trunk splits open, resulting in the death of the animal. Each individual produces from about 50 to several hundred eggs, which are fertilized externally and develop into free-swimming tadpole larvae that undergo metamorphosis

without settling. Life-spans range from one to several weeks.

The epidermal cells on the trunks of all appendicularians secrete a gelatinous tunic made of mucopolysaccharides. This tunic is then expanded into a nonliving, balloonlike structure, the house, used in feeding. Of the three families of appendicularians, the Oikopleuridae have the most complex houses. The animal resides within the house and forces water, by means of the muscular beating of its tail, through two incurrent filters on the external surface of the house (**Fig. 2***a*). These filters contain a fibrous mesh which removes large particles, particularly large diatoms and other algae. Water laden with smaller particles then passes through a complex internal filter, where small phytoplankton and bacteria become further concentrated. The ciliated spiracles on the animal's trunk draw these food particles in a slurry up a hollow mucous tube to the animal's mouth.

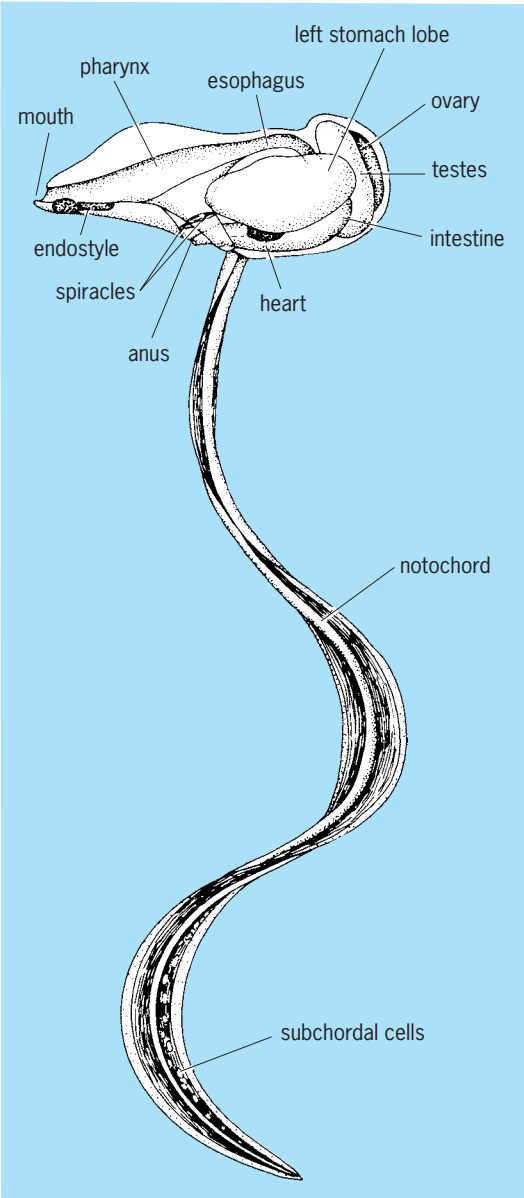

Fig. 1. Anatomy of *Oikopleura albicans*, a typical appendicularian. (*After A. Alldredge, Appendicularians, Sci. Amer., 235:94–102, 1976*)

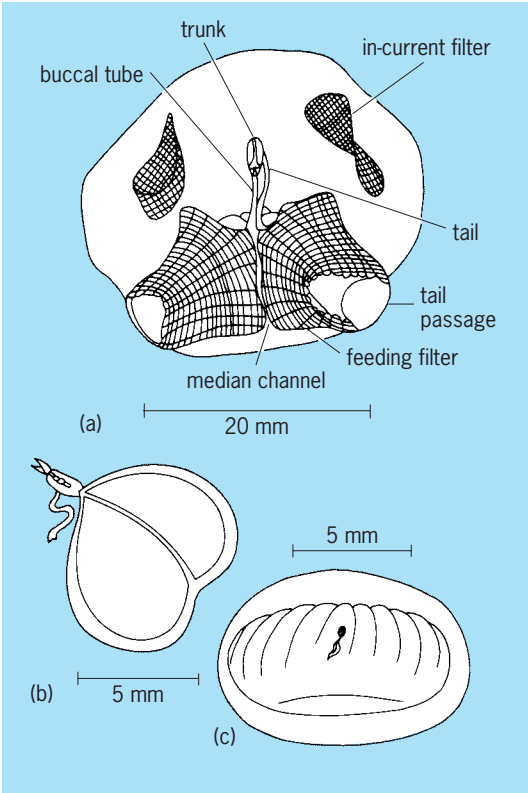

Fig. 2. Houses of the three appendicularian families: (*a*) Oikopleuridae, (*b*) Fritillaridae, and (*c*) Kowalevskiidae.

The house becomes readily clogged with the larger nonfood particles, and appendicularians may abandon the old house and expand a new one as often as six times a day. The house built by the family Kowalevskiidae is a much simpler umbrellalike structure lacking complex filters (Fig. 2*c*). Members of the family Fritillaridae deploy the house as a large, collapsible bubble in front of the mouth (Fig. 2*b*).

Appendicularians feed primarily on small particles from 0.1 to 30 micrometers in diameter. Larger individuals may filter up to several thousand milliliters of seawater per day. As one of the few metazoan groups capable of capturing bacteria and tiny phytoplankton, appendicularians are important marine grazers which may significantly reduce phytoplankton populations in the ocean. About 13 genera and 70 species of appendicularians are known. They are found in all oceans of the world and are a common component of plankton samples, particularly in coastal waters. The most abundant genera are *Oikopleura* and *Fritillaria*. *See* CHORDATA; TUNICATA.

Alice L. Alldredge

Bibliography. A. L. Alldredge, Appendicularians, *Sci. Amer.*, 235:94–102, 1976; A. L. Alldredge and L. P. Madin, Pelagic tunicates: Unique herbivores in the marine plankton, *Bioscience*, 32:655–663, 1982; R. D. Barnes, *Invertebrate Zoology* 6th ed., 1994.

-s

Appendix (anatomy)

A narrow, elongated tube closed at one end, extending from the cecum, a blind pocket off the first part of the large intestine (see **illus.**). It is found in only a few mammals.The size and detailed structure of the appendix vary markedly, depending on the species and the age of the individual. Most reptiles, birds, and mammals have either a single or a paired cecum at the anterior end of the large intestine, but it is quite rare that this cecum has a thinner projection or true appendix.

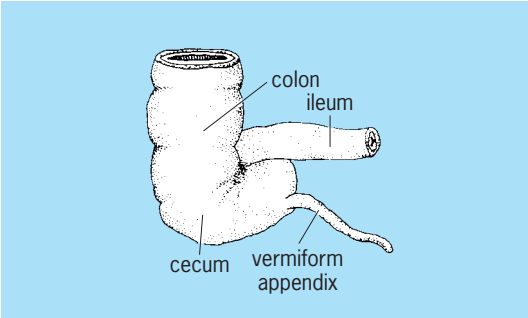

Junction of human ileum and colon, showing cecum and vermiform appendix. (*From C. K. Weichert, Anatomy of the Chordates, 4th ed., McGraw-Hill, 1970*)

In humans the appendix is about 3 in. (7.5 cm) long, but it varies greatly in both size and its specific location in the lower right quarter of the abdomen. Rabbits and their relatives have a very large cecum with an elongate and rather thick appendix, which may reach a length of several inches. A few rodents, such as some of the Old World porcupines, also possess an appendix. The pyloric ceca at the anterior end of the intestine of many fishes are sometimes called appendixes, but they are quite different structures.

The inner lining (mucosa) of the appendix is continuous with the remainder of the digestive lining. In young humans the glandular epithelium is surrounded by clusters of lymphoid follicles which tend to regress with age. The next layer is a smooth muscle coat which, in turn, is surrounded by a connective tissue coat, the outermost part being covered with visceral peritoneum (serosa).

The exact function of the human appendix is unknown, and it is considered to be a remnant of a portion of the digestive tract which was once more functional and is now in the process of evolutionary regression. The appendixes presumably function in digestion in forms with larger ones. *See* APPENDICITIS.

Thomas S. Parsons

Apple

Apples (genus *Malus*) belong to the family Rosaceae (see **illus.**), which includes many other deciduous fruits such as pears, peaches, and cherries. There are about 30 *Malus* species in the North Temperate Zone. The fruits of most species are edible. Although selections of the Asiatic species *M. prunifolia* have been cultivated for their fruits for centuries in China and Japan, they have now been largely replaced by varieties of the "European" cultivated apple. *See* FRUIT; ROSALES.

More apples are consumed than any other temperate-zone tree fruit. Apples are eaten fresh, processed into jellies or preserves, cooked in pies and pastries, or made into sauces. Apple juice is drunk fresh, or at various stages of fermentation as cider, applejack, or brandy. Apple cider vinegar is popular for use in salads and in many processed foods.

Origin and breeding. The "European" cultivated apple is now thought to have been derived principally from *M. pumila*, a Eurasian species which occurs naturally from the Balkans eastward to the Tien Shan of central Asia. In the wild, some forms of *M. pumila* approach present cultivars in size and quality. Another Asian species, *M. sylvestris*, whose range extends into western Europe, grows side by side with *M. pumila* and hybridizes with it in the Caucasus Mountains. Thus *M. sylvestris* probably also had some genetic input into the cultivated apple.

Wild apples, mainly the edible *M. baccata*, grow so thickly east of Lake Baikal in Siberia that the region is called Yablonovy Khrebet ("Apple Mountains").

The success of the Delicious and Golden Delicious cultivars may be laid to the demand for better-quality fresh fruit. Both originated on farms as chance seedlings near the end of the nineteenth century. In spite of the dominance of such "chance" cultivars, cultivars such as Cortland and Idared that were produced by scientific breeding have begun to achieve prominence. Apple crosses are made by first removing the stamens from the flowers of the female parent and then applying previously gathered pollen of the selected male parent to the receptive pistil. After fruits form and ripen, seed is removed and germinated.

Most of the apple breeding programs under way earlier this century have ceased, but the few remaining ones are now introducing some exciting new cultivars. Many of the new cultivars are of

Apple varieties: (*a*) Priscilla and (*b*) Silverspur. (*Bountiful Ridge Nurseries*)

excellent quality and in addition are resistant to the most damaging diseases, such as scab, rust, and fire blight. These cultivars promise to revolutionize apply growing in the future. *See* BREEDING (PLANT).

Distribution. The apple is probably the most widely distributed fruit crop in the world, although it ranks behind grapes, bananas, and oranges in total production. There are substantial apple-growing areas on all temperate-zone continents; the United States, Italy, France, and Germany are leading producers.

Apples can be grown as far north as 60° latitude in the maritime climate of northwestern Europe. In North America, apple culture does not extend much above 50° north latitude. Away from the coasts, the buffering effects of the Great Lakes on temperature extremes facilitate heavy apple production in New York, Michigan, and Ontario. Hardier cultivars are continuing to be developed for use in colder regions. Apples can be grown in the tropics at higher elevations where sufficient chilling to break dormancy is available. Cultivars with lower chilling needs are being developed.

The principal apple-growing regions in North America are, in order of importance: the Northwest (Washington, Oregon, Idaho, and British Columbia), the Northeast (New York, New Jersey, New England, Ontario, Quebec, and Nova Scotia), the Cumberland-Shenandoah area (Pennsylvania, Maryland, Virginia, West Virginia, and North Carolina), Michigan, California, the Ohio Basin (Ohio, Indiana, and Illinois), and Colorado.

Planting systems and rootstocks. Historically apples have been grown in America as large trees on seedling rootstocks, widely spaced, with a density of 25–40 per acre (62–100 per hectare). This system has many disadvantages: the trees take many years to reach full production, they are difficult to prune and harvest because ladders are needed, and the fruit is often poorly colored and small in size.

Dwarfing rootstocks have been used by European growers for centuries to produce smaller trees that bear earlier, and produce better-quality fruit. Earlier in this century the various clonal rootstocks were cataloged into size categories (the M. series) according to the numbering system of the East Malling Research Station, England. New clonal rootstocks (the MM. series) have been developed at East Malling by breeding. There are now rootstock-breeding programs under way in several European countries and at the New York State Agricultural Experiment Station at Geneva.

In America, most apple trees are planted on clonal rootstocks which produce semidwarf trees at a density of 100–200/acre (250–500/ha). There are also some plantings of full-dwarf trees (400–800/acre or 1000–2000/ha), sometimes grown as espaliers on a trellis. These plantings are much more productive on an area basis than the older ones. Although the acreage of apple orchards has declined drastically in this century, the number of trees and the production of apples have increased. *See* PLANT PROPAGATION.

Orchardists in many areas are using sod culture with herbicide treatments in the row to prevent competition from weeds for moisture and nutrients. Nutrients are applied by broadcasting granular fertilizers or by spraying solutions on the soil near the trunk of trees. In special situations, urea in solution may be applied to foliage.

Apples are usually irrigated in western North American growing areas. The use of trickle irrigation rather than flood, furrow, or sprinkler irrigation has increased. In trickle irrigation, small amounts of water are applied over a long period of time and as needed by the trees. Fertilizer can be applied by trickle irrigation.

In certain areas, particularly in western North America, spring frosts can damage blossoms. To reduce the danger of damage, several techniques are available. Orchards can be heated with smudge pots which burn waste oil, or with gas heaters, or by burning "bricks" of petroleum. The higher prices of petroleum-based fuel and more strict pollution control regulations are making this method less popular. Large fans ("wind machines") can be used in orchards to break inversion layers by mixing warm air posed above the trees with the cold air near the ground. When trees are sprinkled in the late dormant period, evaporative cooling results in delay of bloom beyond the frost danger period. Sprinkling can also be used during periods of freezing temperatures to prevent frost injury directly. The latent heat of fusion liberated when the sprinkled water freezes provides sufficient heat to prevent the blossoms from freezing.

Pruning. To maintain the correct balance between wood, vegetative growth, and fruiting spurs, apple trees must be pruned regularly. On most cultivars, blossoms, and later fruits, are formed on spurs produced in previous years. Pruning seeks to maximize spur formation on a strong tree structure into which is permitted the penetration of adequate light for coloring and ripening. Traditionally, pruning has been carried out in the dormant season, but in recent years a system of pruning during the summer has been developed that requires less work and gives better results. Manual, hydraulic, or pneumatic pruning tools are used. Pruners gain access to trees by climbing, with the aid of ladders and self-propelled mechanical platforms.

Pollination. Most apple cultivars are self-incompatible and require pollen from a second cultivar for fertilization and, therefore, fruit formation to occur. The pollen is transferred from tree to tree by insects, usually bees. Fruit growers often plant two cultivars that bloom at the same time in close proximity, so that each will pollinate the other effectively. Triploid cultivars, such as Idared, must be planted in close proximity with two other cultivars since the triploid's pollen is infertile. In many areas natural populations of bees are insufficient to provide adequate pollination, and hives are placed in the orchard during bloom. *See* POLLINATION.

Pest control. Apples are attacked by viruses, bacteria, mycoplasmas, fungi, nematodes, mites, insects, and mammals. In most apple-growing regions precautions must be taken against all of these pests,

though to different degrees. In practice, damage caused by fungi and insects is most important. Much protection can be afforded by growing cultivars that are resistant to one or more of the pests prevalent in the area. But a limited number of cultivars resistant to only a few pests are now available. Organisms that are spread solely in propagating material, like some viruses, can be controlled by maintaining virus-free nurseries. Some pests can be reduced by cultural practices that interfere with their life cycle. Thus, close mowing limits weed growth, and removing dropped fruits, which harbor pests, may reduce subsequent damage from those pests.

However, good pest control on the present commercial apple cultivars usually requires the use of chemical pesticides. Chemicals for use on apples are stringently tested for efficacy and for safety to orchard workers, consumers, and the environment. Several compounds previously used have been discontinued for safety reasons. Modern pesticides, properly used, provide excellent control of most pests without damaging the trees and fruit. However, many pests have developed tolerance to some of the best pesticides, rendering them ineffective. In an effort to forestall the development of resistance, several different chemicals are used; pests are unlikely to develop tolerance to all of them. Since the development, testing, and registration of new pesticides have become such a lengthy and expensive process, the future of chemical control is uncertain. *See* PESTICIDE.

Three approaches that have become more important for pest control are integrated pest management (IPM), the use of resistant cultivars, and the development of biological controls. In integrated pest management all available relevant methods of control are integrated in such a way that each method supplements and does not interfere with any other. Thus chemicals would be used only when pests are predicted to reach sufficient levels to result in economic losses. Cultural practices would be used that result in reduced damage from pests by encouraging the development of more resistant plants and by favoring the action of predators and parasites that are active against apple pests.

Considerable progress has been achieved on biologically based controls for apple pests. Biological control in the strict sense involves the use of predators, parasites, or pathogens to reduce pest populations and activity, thus causing a reduction in damage by the pest species of interest. Generally, biological controls are highly specific to the target organism and have little effect on other organisms in the vicinity. Once established, biological controls are likely to be long-lasting because of the basic compatibility of the agent with the pest. *See* INSECT CONTROL, BIOLOGICAL.

In many areas, the control of the European red mite is accomplished through the activity of predacious mites. Populations of the predacious mites often increase in orchards to levels that result in suppression of the apple-damaging red mite. Highly specific insecticides that do not affect predacious mites have been developed for use in orchards. Occasionally, however, certain insecticides must be used to control other apple-damaging insects, and these also affect the predacious mites, thus disrupting the biological control process of the red mites. *See* INSECTICIDE.

Great advances have been made in understanding insect pheromones, chemical substances that are produced in minute quantities by female insects and that males can sense over great distances. Several pheromones have been identified and synthesized chemically. Synthetic pheromones are used as attractants in traps to determine insect populations. If large amounts of pheromones are released in orchards, sexual activity of insects is disrupted and reproduction is reduced. *See* PHEROMONE.

As far as diseases of apple are concerned, only a limited number of biological control systems are in use. The crown gall disease caused by a bacterium is being controlled in some nurseries by treating nursery stock with suspensions of a nonpathogenic bacterium that produces an antibiotic that inhibits the pathogenic bacterium. *See* CROWN GALL.

Absolute control of a pest can be achieved by growing a cultivar that is resistant to the pest. Even partial resistance may be sufficient in areas and years when pest pressure is low. The amount of pesticide needed for pest control on partially resistant cultivars is less than that required on more susceptible cultivars.

There are several projects in the United States and other countries to breed new apple cultivars that are more resistant to pests, particularly diseases. Some scab-resistant cultivars, such as Liberty, Nova Easygro, and Priscilla, are already available. Some of these cultivars are also resistant to other diseases such as rusts, fire blight, and powdery mildew. *See* PLANT PATHOLOGY.

Harvesting. Traditionally, apples have been harvested manually by pickers standing on ladders. Two developments are changing this practice: (1) The use of mechanical tree shakers with catching frames and conveyor belts that detach and move fruit to bulk bins or trucks without individual handling. Apples so harvested usually are destined for processing. (2) The use of dwarf trees whose fruit can be picked from the ground by inexperienced personnel. *See* AGRICULTURAL MACHINERY.

Marketing. Apples may be sold fresh immediately, or after a period of storage, or they may be processed into less perishable products such as canned, frozen, or dried slices or chunks for baking, applesauce, apple juice or cider, and vinegar.

Fresh apples may be sold locally at roadside stands or at farmers' markets, or sold in wholesale quantities to supermarkets. Often, large quantities of fruit are traded by cooperatives or independent buyers for sale in distant markets. Thousands of tons of apples are shipped annually from the state of Washington to the Midwest and the East Coast. Apples can be stored for long periods at low temperatures under controlled atmosphere (CA storage), so that fresh fruits are now available year-round.

There is a large international trade in apples, particularly in Europe between the large producing countries of France and Italy and the net-importing countries Germany, Britain, and Scandinavia, and from the United States to Canada. Apples grown in the Southern Hemisphere (Australia, New Zealand, South Africa, and Argentina) are shipped in large quantities to western Europe during the northern winter. *See* FRUIT, TREE.

Herb S. Aldwinckle; Steven V. Beer

Apraxia

An impairment in the performance of voluntary actions despite intact motor power and coordination, sensation and perception, and comprehension. The apraxic person knows the act to be carried out, and has the requisite sensory-motor capacities; yet performance is defective. The abnormality is highlighted when the act must be performed on demand and out of context. The individual may perform normally in such activities as hammering or tooth brushing performed with the object in hand, especially in a natural situation, but will often fail when required to pantomime such acts.

Defects in performance vary from total inability to initiate the action, to incorrect serial ordering of elements, to partial approximations. A common apraxic behavior is the use of a body part as an object. Pantomiming the act of brushing the teeth, for example, a person may run the index finger across the teeth as though it were a toothbrush, while in normal performance, the hand assumes the posture of holding and moving the brush.

Apraxia is usually observed in both upper extremities. When it occurs unilaterally, it is usually the left arm and hand that are affected. This has been explained by assuming that the left cerebral hemisphere is specialized in the organization of voluntary movements, just as it is in language. The left hand is under the immediate control of the right hemisphere, but for skilled voluntary actions, the right hemisphere is dependent on information transmitted from the dominant left hemisphere over the corpus callosum. Callosal lesions produce apraxia of the left hand, because the right hemisphere is incapable of organizing the plan of movement independently. With an appropriately placed left-hemisphere lesion, a bilateral apraxia will result. *See* HEMISPHERIC LATERALITY.

When the left-hemisphere lesion also destroys the primary motor zone, the right arm is paralyzed and the apraxia is masked. The observable apraxia on the left side is referred to as sympathetic apraxia. This is seen in many individuals with right hemiplegia (unilateral paralysis of the body) and Broca's aphasia. Another apraxia often coupled with Broca's aphasia is nonspeech oral apraxia (or buccofacial apraxia). Individuals with this disorder can be observed to struggle to perform such simple acts as protruding the tongue or licking the lips on command or imitation, even though these movements are executed easily as part of the act of eating. *See* APHASIA.

There are several disorders that are controversial with regard to their interpretation as forms of apraxia. The nonfluent speech pattern of Broca's aphasia, often riddled with speech-sound errors, is considered as apraxia of speech by some authorities, while others view it as an integral part of the linguistic deficit of the aphasia. In dressing apraxia and in some types of constructional apraxia, the defect appears to be perceptually based. Limb-kinetic apraxia is widely interpreted today as a mild spastic paresis, while ideational apraxia, commonly associated with dementia, is likely due to conceptual confusion rather than to a disturbance of motor organization. *See* AGNOSIA.

Gerald J. Canter

Bibliography. G. R. Hammond (ed.), *Phenomenology of Will and Action*, 1967; K. M. Heilman and E. Valenstein (eds.), *Clinical Neuropsychology* 3d ed., 1993; P. J. Vinken and G. W. Bruyn (eds.), *Cerebral Control of Speech and Limb Movements*, 1991; R. W. Wertz et al., *Apraxia of Speech in Adults: The Disorder and Its Management*, 1991.

Apricot

The stone fruit *Prunus armeniaca*, thought to be native to China, and then distributed throughout Asia, Europe, and eventually to North and South America and Oceania. The species is genetically diverse and can grow in a wide range of climates depending upon the cultivar. Such diversity occurs in North America, where apricots are produced near Penticton, British Columbia, Canada, in northern New York, as far south as southern California, and even near Puebla, Mexico. Most commercial production in the world is limited to areas where temperatures do not fall below −10 to −20°F (−23 to −29°C) for extended periods; however, certain cultivars can tolerate even severer conditions. Many apricot cultivars can tolerate high summer temperatures in excess of 105°F (40°C). Some cultivars develop an internal browning of the flesh if high temperatures persist with fruit on trees. Apricots tend to bloom earlier than other stone fruit and are sensitive to frost. Frostfree areas are generally preferred.

Horticulture. Flowers from most commercial cultivars are self-fertile, but examples of self-infertility are found in commercial cultivars. In those self-incompatible cultivars, cross-pollination with another cultivar is required. Honeybees are normally used to facilitate cross-pollination. The Japanese apricot (*P. mume*) is a favorite ornamental and fruit tree in China and Japan. In Japan, for example, the *mume* is used in juice and processed and sold as a consumable fruit product. There are small orchards of *mume* in California; it has also been used in hybridization with apricot.

Fruits for commerce are generally yellow to orange in skin color and flesh (**Fig. 1**). The skin surface is generally smooth, with a thin suture line. New cultivars have very smooth skin, with very intense orange skin, sometimes with a red blush and deep orange flesh color. Apricots can also be red, or white with varying shades of yellow or orange; however,

Fig. 1. Castlebrite apricot branch with a fruit cluster.

these are normally not found in commercial outlets. Apricot size ranges from small (about 25–30 g per fruit) to large (100–130 g). The fruit can be consumed fresh, dried, frozen, or processed as canned product, as juice, or as baby food (pureed).

Apricots thrive best in well-drained soils. They tolerate sandy, gravel, and clay soils, but grow best in deep alluvial, sandy loams. Fruit with increased color and taste are probably produced on soils which have high amounts of sand and gravel. Apricots are adapted to soils where the pH is 6.2–7.2. Soils with neutral pH help to minimize micronutrient deficiencies. Apricots require only a modest amount of nitrogen compared with many other fruit trees. Typically 50–100 pounds of actual nitrogen per acre per year are applied to commercial apricot orchards. Too much nitrogen can delay fruit maturity, reduce fruit firmness, and generally reduce fruit quality. Fruits from high-nitrogen fertilized soils tend to be more prone to disease and produce excessive shoot growth with ample water. Additions of potassium, zinc, manganese, and boron are normally required on a regular basis. In some regions, phosphorus- and sulfur-containing fertilizers are applied. Calcium sprays appear to increase fruit firmness.

Production and propagation. Of the total commercial production of apricots in North America, California produces 90–95%, followed by Washington and Utah. Trees are propagated by budding on hardwood cuttings of Marianna 2624 (*P. cerasifera* × *P. musoniana?*) or Myrobalan 29C (*P. cerasifera*) seedlings of apricot, peach (Lovely, Nemaguard, Nemared), and Myrobalan. Peach rootstocks are the most commonly used. Plum roots are used in sites that are poorly drained. Citation (peach × plum) is a vegetatively propagated rootstock which has gained use in commercial apricot plantings. The rootstock is thought to control tree size and improve fruit quality. It is sensitive to crown gall disease. Trees are set from 24 ft (7.3 m) apart, to 8 ft (2.5 m) between rows in the orchard. Orchard undergrowth is either clean-cultivated (free of undergrowth) or controlled by chemical herbicides along tree rows, and the remainder (row middles) are mowed. Planted cover crops or natural vegetation is mowed in middles. Trees are pruned and trained to an open-center framework by summer and dormant pruning. *See* PLANT PROPAGATION.

Harvest. Apricots are harvested mostly by hand, but some machine harvesting is done for processed fruit. Fresh fruit is picked and packed in the orchard or is graded and sorted in sheds specifically designed for fruit packing. The fruit is cooled and shipped in volume fill or tray packs to markets throughout the world. Processed fruit is generally harvested into large bins and then loaded onto trucks and carried to canneries, freezers, or facilities that can process the fruit into juice. Apricots are a very nutritious fruit high in vitamin A and antioxidants lycopene and beta carotene. They are also a good source of vitamin C and soluble fiber.

Diseases. There are several fungal diseases of apricot that require annual control. The brown rot fungus destroys blossoms and twigs (*Monilinia laxa*) in the spring and frequently causes fruit decay (*M. fructicola* and *laxa*) in the summer (**Fig. 2**). The disease can be devastating and can cause reduced yield. Sprays are applied from early in the bloom through full bloom to reduce blossom and twig blight. Typically one to three sprays are applied, and chemicals such as Iprodione and Vinclozolin (Dicarboximides), or Tebuconazole, Propiconazole, or Fenbuconazole (DMI-triazoles) are effective. New fungicides to control brown rot disease are continually under development. Protection of the fruit is achieved by sprays applied from 1 to 3 weeks

Fig. 2. Spurs and shoots infected with brown rot fungus.

prior to harvest with the aforementioned chemicals. A disease that infects during the flowering and small fruit stage is known as jacket rot, caused by *Botrytis cinerea*. Symptoms occur on small green fruit when remnants of flower parts are still attached to the fruit. The disease causes a brown discoloration on the fruit under the petal remnants. The young fruit may wither and fall within weeks. Control can be achieved with fungicides. Powdery mildew (*Sphaerotheca pannosa* and *Podosphaera tridactyla*), a spring fungal disease, appears as web-like white growth on fruit, leaves, and stems. Fruit may not be packed because of associated lesions on the surface. The fungus *Wilsonomyces carpophilius*, causing shot-hole disease, kills dormant buds, spots fruit and leaves, and, in severe cases, induces defoliation. The fungus survives within infected buds and on twigs. It is controlled by spraying a fungicide at early dormancy (fall) and during blossoming. *See* FUNGISTAT AND FUNGICIDE.

A disease destructive to the framework of trees is caused by *Eutypa lata* (gummosis, cytosporina, limb dieback). This fungus invades fresh pruning wounds when rainfall occurs 2–6 weeks after pruning, causing infections that kill branches and sometimes the entire tree. The greatest incidence of disease occurs during periods when rainfall is highest (fall and winter). Control is based largely on removal and destruction of infected branches, pruning during June, July, August, and September, or no pruning (California). No other treatments have been consistently commercially feasible and reliable.

A lethal root rot of apricot is caused by the fungus *Armillaria mellea*. It can be partially controlled by the use of Marianna 2624 rootstock. Another root and crown rot is caused by *Phytophthora* spp. Both *Armillaria* and *Phytophthora* diseases can kill trees. Site selection to avoid poorly drained soils aids in control of *Phytophthora*. Verticillium wilt (*Verticillium dahliae*) is a soil-borne fungal disease that causes branches and limbs on trees to collapse and die. Symptoms typically appear in young trees less than 6 years old. Careful water management, and tree removal followed by soil fumigation with methyl bromide can aid in the control of *Armillaria*, *Phytophthora*, and *Verticillium* spp.

Bacterial canker caused by *Pseudomonas syringae* is the most important bacterial disease. The disease is found typically in young trees less than 7 years old and is associated with trees growing in stressful conditions, such as cold, high rainfall, poor nutrition, or nematode infestation. Twigs, branches, and sometimes entire trees are killed. Disease severity usually is greatest in trees on sandy soil and in those on apricot and plum rootstocks. Reducing tree stress and budding the cultivar high on the rootstock seem to reduce the incidence of disease. Fumigation with methyl bromide may aid in controlling nematode populations. Bactericidal sprays are relatively ineffective.

A few virus and viruslike diseases of apricot are known. The most common, the plum pox virus (sharka) and chlorotic leaf spot, are found throughout producing regions of Europe. Most of these are controlled by the use of disease-free planting stock. *See* PLANT PATHOLOGY.

Insects. The most common insect pests of apricot are peach twig borer (*Anarsia lineatella*), worms (*Archips argyrospila*, *Orthosia hibisci*, and others), earwigs (*Forficula auricularia*), mites (*Tetranychus* spp.), and aphids. Integrated pest management programs have been developed to control most pests of apricots.

Stephen M. Southwick

Bibliography. N. F. Childers, *Modern Fruit Science*, 10th ed., 1995; J. M. Ogawa and H. English, *Diseases of Temperate Zone Tree Fruit and Nut Crops*, 1991; J. M. Ogawa and S. M. Southwick, Introduction: Apricot, in J. M. Ogawa et al. (eds.), *The Compendium of Stone Fruit Diseases*, American Phytopathological Society, 1995; A. Rosati, T. M. DeJong, and S. M. Southwick, Comparison of leaf mineral content, carbon assimilation and stem water potential of two apricot (*Prunus armeniaca*) cultivars grafted on Citation and Marianna 2624 rootstocks, *Acta Hort.*, 451:263–267, 1997; S. M. Southwick and G. Stokes, Apricots, the fruit and its importance, in *The Encyclopedia of Food Science, Food Technology, and Nutrition*, Academic Press, London, 1991; S. M. Southwick and K. G. Weis, Selecting and propagating rootstocks to produce apricots, *HortTechnology*, 8(2):164–170, 1998; S. M. Southwick, J. T. Yeager, and K. G. Weis, Use of gibberellins on Patterson apricot (*Prunus armeniaca*) to reduce hand thinning and improve fruit size and firmness: Effects over three seasons, *J. Hort. Sci.*, 72(4):645–652, 1997; S. M. Southwick, J. T. Yeager, and H. Zhou, Flowering and fruiting in Patterson apricot (*Prunus armeniaca*) in response to postharvest application of gibberellic acid, *Scientia Horticulturae*, 60:267–277, 1995; G. Stokes et al., Apricots and apricot processing, in Y. H. Hui (ed.), *The Encyclopedia of Food Science and Technology*, Wiley, Cutten, CA, 1989.

Apsides

In astronomy, the two points in an elliptical orbit that are closest to, and farthest from, the primary body about which the secondary revolves. In the orbit of a planet or comet about the Sun, the apsides are, respectively, perihelion and aphelion. In the orbit of the Moon, the apsides are called perigee and apogee, while in the orbit of a satellite of Jupiter, these points are referred to as perijove and apojove. The major axis of an elliptic orbit is referred to as the line of apsides. *See* CELESTIAL MECHANICS; ORBITAL MOTION.

Raynor L. Duncombe

Aqua regia

A mixture of one part by volume of concentrated nitric acid and three parts of concentrated hydrochloric acid. Aqua regia was so named by the alchemists because of its ability to dissolve platinum and gold. Either acid alone will not dissolve these noble metals. Although chlorine gas and nitrosyl chloride

are formed as indicated in reaction (1), the oxidiz-

$$HNO_3 + 3HCl \rightarrow Cl_2 + NOCl + 2H_2O \quad (1)$$

ing properties of aqua regia are not believed to be increased. Instead, the metal is easier to oxidize because of the high concentration of chloride ions which form a stable complex ion as illustrated with gold in reaction (2).

$$Au + 4H^+ + NO_3^- + 4Cl^- \rightarrow AuCl_4^- + NO + 2H_2O \quad (2)$$

See CHLORINE; GOLD; NITRIC ACID; PLATINUM.

E. Eugene Weaver

Aquaculture

The cultivation of fresh-water and marine species (the latter type is often referred to as mariculture). The concept of aquaculture is not new, having begun in China about 4000 years ago, and the practice of fish culture there has been handed down through the generations. China probably has the highest concentration of aquaculture operations in a single country today. Yet the science of aquaculture, as a whole, is far behind that of its terrestrial counterpart, agriculture. Aquaculture, though still partly art, is being transformed into a modern, multidisciplinary technology.

Aquacultural ventures occur worldwide. China grows macroalgae (seaweeds) and carp. Japan cultures a wide range of marine organisms, including yellowtail, sea bream, salmonids, tuna, penaeid shrimp, oysters, scallops, abalone, and algae. Russia concentrates on the culture of fish such as sturgeon, salmon, and carp. North America grows catfish, trout, salmon, oysters, and penaeid shrimp. Europe cultures flatfish, trout, oysters, mussels, and eels. Presently, plant aquaculture is almost exclusively restricted to Japan, China, and Korea, where the national diets include substantial amounts of macroalgae.

Aquaculture accounts for approximately 10% of the world's production of fish and shellfish, compared with an estimated 7% contribution in 1967. Roughly 65–70% of the world's aquaculture production consists of fresh-water, brackish-water, and marine fish, about 30–35% is divided equally between seaweeds and mollusks, and 0.3% crustaceans. In Japan, aquaculture supplies more than 98% of the yellowtail consumed, and in Israel about half of the finfish consumed are cultured.

Extensive to intensive culture. The worldwide practice of aquaculture runs the gamut from low-technology extensive methods to highly intensive systems. At one extreme, extensive aquaculture can be little more than contained stock replenishment, using natural bodies of water such as coastal embayments, where few if any alteration of the environment are made. Such culture usually requires a low degree of management and low investment and operating costs; it generally results in low yield per unit area. At the other extreme, intensive aquaculture, animals are grown in systems such as tanks and raceways, where the support parameters are carefully controlled and dependence on the natural environment is minimal. Such systems require a high degree of management and usually involve substantial investment and operating costs, resulting in high yields per unit area. Obviously, many aquacultural operations fall somewhere between the two extremes of extensive and intensive, using aspects of both according to the site, species, and financial constraints.

Fig. 1. Typical earthen pond as used for catfish culture.

In the Mediterranean area, culturists successfully use extensive methods to grow fish such as mullet, usually in polyculture with other species. Mullet are contained in large estuarine lagoons requiring the voluntary entrance of juveniles from the wild. These systems rely on naturally occurring feeds, and water is exchanged through arrangements of canals and sluices. Typically, the fish are harvested after 2–3 years, with relatively low production levels (88–176 1b/acre or 100–200 kg/hectare). Additional species that enter the lagoons are considered a welcome secondary crop.

An example of aquaculture further up the scale of intensity is catfish farming as normally practiced in warmer regions of the Western Hemisphere. In many areas, such as Mississippi, catfish are grown in fresh-water ponds that are often as large as 25 acres (10 ha) [**Fig. 1**]. These systems are stocked seasonally, and the animals are fed dried, milled feeds supplemented by organisms that occur naturally in the ponds. Catfish grown in this manner usually reach a marketable size of 0.7 kg within 2 years, with yields as high as 2640 lb/acre (3000 kg/ha). This yield, though far higher than that of extensive mullet culture, is restricted by the grower's lack of environmental control. For example, in densely stocked pond systems, oxygen depletion is a continual concern, and it is not uncommon for a farmer to experience massive fish deaths as a result of low oxygen levels. Aeration devices are commonly used in an attempt to avoid such disasters (**Fig. 2**).

To increase yields per unit area and limit environmental dependence, culturists such as catfish farmers are beginning to rear fish in raceway and tank systems similar to the techniques that have been

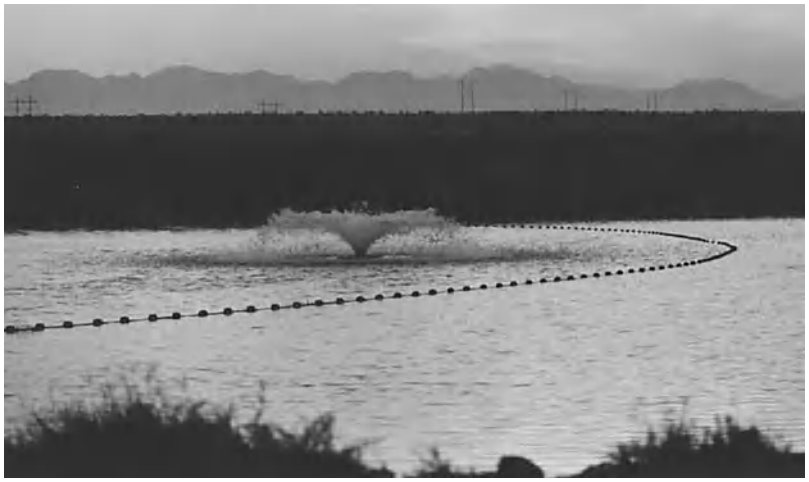

Fig. 2. Aerator operating in catfish pond.

used for years to grow trout. An example of such an operation is one in which catfish are reared in tanks ranging in size from 485 to 1450 ft^3 (45 to 135 m^3) [**Fig. 3**]. In this operation, geothermal water sources are used for year-round maintenance of optimal water temperature, and oxygen levels are carefully controlled. Such systems, when compared with pond culture, are capable of a thousandfold higher yields, providing harvestable fish of 2 lb (0.9 kg) within 12 months.

Both of the above examples of catfish-rearing techniques rely on classical brood stock and hatchery systems where mature breeding fish are maintained in ponds. Breeding pairs segregate; the females lay their egg masses in containers provided (usually milk cans); and the eggs are fertilized by the males. These fertilized egg masses are retrieved and hatched in troughs continually aerated by paddles (**Fig. 4**). Hatched larvae are maintained in small tanks or raceways and given artificial feeds until they reach a stocking size of 2.6–5.2 in. (7–15 cm).

Another outstanding example of intensive production is that of penaeid shrimp culture (the common marine prawn), as pioneered by the National Marine Fisheries Service Laboratory in Galveston, Texas, and the Universities of Arizona and Sonora. These systems consist of shallow raceways housed in greenhouselike domes made of inflated polyethylene (**Fig. 5**). The raceways may have either a closed, recirculating water source or a flow-through water system (**Fig. 6**). Penaeid species are reared on artificial feeds in high density (3.3 kg/m^2 or 0.68 $1b/ft^2$) from the postlarval stage. In such systems, environmental parameters are carefully maintained, and marketable animals can be produced within 4–6 months. The life cycle of the penaeid shrimp recently has been closed in captivity, and hatchery technology is available for the production of seed stock.

A unique combination of highly intensive and extensive aquaculture occurs in ocean ranching, as commonly employed with anadromous fish (which return from the ocean to rivers at the time of spawning). The two most notable examples are the ranching of salmon and sturgeon. In both instances, highly sophisticated hatchery systems are used to rear young fish, which are then released to forage and grow in their natural environment. The animals are harvested upon return to their native rivers.

Fig. 3. Circular raceways which are used for intensive culture of catfish. This operation employs geothermal heat to ensure growth on a year-round basis. (*Calaqua, Inc., Paso Robles, California*)

Fig. 4. Mechanized paddle troughs used in catfish aquaculture to provide aeration during egg incubation.

Industrialized aquaculture. While extremely extensive aquacultural operations are continuing, as they have been for generations, in the lesser-developed areas of the world, the more industrialized nations—where land and water resources are limited—have turned increasingly to intensive aquaculture. Intensive aquaculture brings with it high energy costs, necessitating the design of energy-efficient systems. As this trend continues, aquaculture will shift more to a year-round, mass-production industry using the least amount of land and water possible.

With this change to high technology and dense culturing, considerable knowledge and manipulation of the life cycles and requirements of each species are necessary. Specifically, industrialized aquaculture has mandated the development of reproductive control, hatchery technology, feeds technology, disease control, and systems engineering.

Improvements in each of these aspects are occurring and are critical to the future of intensive aquaculture. Reproductive control of a species is particularly important to maintain a steady supply of seed stock and to undertake genetic selection. Such control eliminates reliance on unpredictable wild stocks and is the first step toward the domestication of species for a production-oriented industry. In some fish species (notably trout and carp), reproduction has been controlled for a considerable time; in others (penaeid shrimp and lobster, for example), control has been achieved recently and is not yet routine. For a number of species (such as milkfish and eel), culturists still depend on wild stocks for replenishment.

Reproduction in captivity paves the way for geneticists to improve culture stocks through selective breeding designed to increase important traits such as growth rate, resistance to disease, tolerance to crowding, and reproductive performance. In terrestrial animal husbandry, genetic selection has substantially improved production, and the same is expected for aquaculture. *See* BREEDING (ANIMAL).

Unfortunately, many intensive aquacultural operations experience high mortalities during the animal's embryonic and larval development, hence the importance of creating efficient hatchery techniques. Many aquatic species exhibit a complicated series of larval stages, each requiring distinct environmental conditions and feeding regimes. As more biological information becomes available, hatchery systems are being carefully designed to allow high survival rates at each of these developmental levels.

In the current state of intensive aquaculture, feeds technology has been achieved only for a few species. When animals are grown in controlled systems, they are unable to forage for natural feeds, and the culturists must supply all required nutrients. To do this, however, necessitates knowledge of the animal's specific nutrient requirements and translation of that knowledge into inexpensive, formulated feeds that remain stable in the aqueous environment and are attractive to the particular animal. The lack of feed availability results in the widespread use of expensive, natural foods. *See* ANIMAL FEEDS.

Fig. 5. Prototype inflated building housing raceways for intensive culture of marine shrimp. (*National Marine Fisheries Service Laboratory, Galveston, Texas*)

Fig. 6. Inside view of inflated building in Fig. 5 showing two long raceways used for shrimp production. (*National Marine Fisheries Service Laboratory, Galveston, Texas*)

With the trend toward more mechanized aquaculture, disease control has become increasingly important. In densely packed culture conditions, pathogens considered relatively minor in the wild or in extensive systems are capable of wiping out entire stocks. Modern aquaculture, therefore, has mandated disease-oriented research and the development of therapeutics and prophylactic techniques.

As manipulation and knowledge of the life cycles of aquatic species increases, so does the sophistication of the systems needed to support such control. The current emphasis in aquatic systems engineering is to design practical culture facilities that are appropriate to each species (including all developmental stages), are energy-efficient, provide adequate food delivery, and are suitable for the routine maintenance necessary for high water quality.

Regardless of the type of system used, aquacultural products are marketed as are fisheries products, except for some advantages. For one, fisheries products often must be transported on boats and may experience spoilage; whereas cultured products, which are land-based, can be delivered fresh to the various nearby markets. Also, intensively cultured products through genetic selection can result in a more desirable food than those caught in the wild, with uniform size and improved taste resulting from controlled feeding and rearing in pollution-free water.

Fisheries products traditionally have been the main source of animal protein in certain population centers of the world. In fact, approximately 13% of the animal protein consumed worldwide is provided by fisheries products, and the percentage is continually increasing. Coupled with the dramatic rise in world population, this growing demand for fisheries products has resulted in heavy pressure on the natural resources, which are rapidly approaching the maximum level of sustained harvest. Without the development of aquatic farming techniques, the natural fisheries would not be able to keep up with demand.

Aquaculture is following the lead of the development of agriculture, beginning with hunting, moving through stock containment and low-technology farming, and finally arriving at high-density, controlled culture. As the methodology improves, the modern techniques of aquaculture—and of agriculture—will have to be transferred to the lesser-industrialized areas of the world in an effort to offset the predicted food shortages in the coming years. *See* AGRICULTURE; MARINE FISHERIES.

Wallis H. Clark Jr.; Ann B. McGuire

Bibliography. J. E. Bardach, J. H. Ryther, and W. O. McLarney, *Aquaculture: The Farming and Husbandry of Freshwater and Marine Organisms,* 1972; Food and Agriculture Organization, *Proceedings of the FAO Technical Conference on Aquaculture, Kyoto, Japan, 1976*, FAO Fish. Rep. 188; P. Korringa, *Farming Marine Fishes and Shrimps: A Multidisciplinary Treatise,* 1976; P. R. Limburg, *Farming the Waters,* 1981; R. R. Stickney, *Principles of Warmwater Aquaculture,* 1979.

Aquarius

The Water Bearer, in Greek mythology, is a zodiacal constellation, meaning that the path of the Sun and planets passes through it. (The Sun actually passes

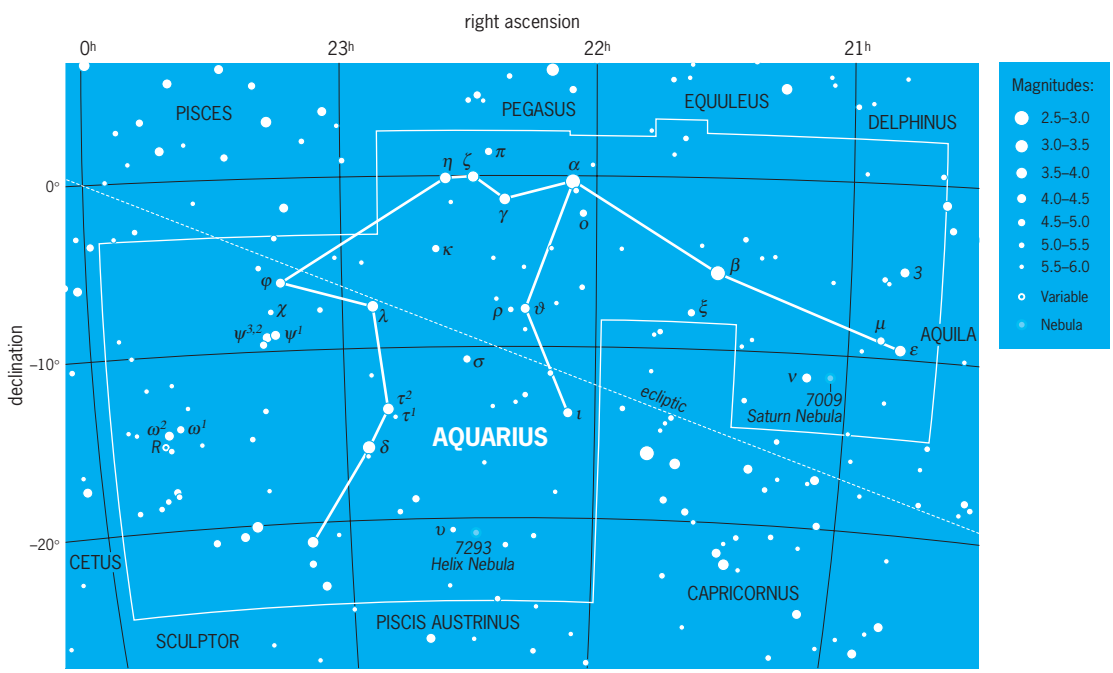

Modern boundaries of the constellation Aquarius, the Water Bearer. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

through 13 of the modern constellations each year, not the traditional 12.) It is almost entirely south of the celestial equator; none of its stars is especially bright (see **illustration**). Early constellation drawings show Aquarius as a man or boy pouring water from a bucket. The stream of water is oriented so as to flow into Piscis Austrinus, the constellation of the Southern Fish, pointing at the bright star Fomalhaut, which can be taken to correspond to the mouth of the fish. The constellation may have gotten its name from the Sun's position in it during the often-rainy month of February, which often led to floods in Babylonia and Egypt. The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION; ZODIAC.

The constellation boasts of two beautiful planetary nebulae, the Helix Nebula and the Saturn Nebula. The former subtends an angle in the sky larger than that of the Moon, although it is too faint to see without a telescope. *See* PLANETARY NEBULA.

Jay M. Pasachoff

Aquifer

A subsurface zone that yields economically important amounts of water to wells. The term is synonymous with water-bearing formation. An aquifer may be porous rock, unconsolidated gravel, fractured rock, or cavernous limestone. Economically important amounts of water may vary from less than a gallon per minute for cattle water in the desert to thousands of gallons per minute for industrial, irrigation, or municipal use.

Among the most productive are the sand and gravel formations of the Atlantic and Gulf Coastal plains of the southeastern United States. These layers extend for hundreds of miles and may be several hundred feet thick. Also highly productive are deposits of sand and gravel washed out from the continental glaciers in the northern United States; the outwash gravel deposits from the western mountain ranges; certain cavernous limestones such as the Edwards limestone of Texas and the Ocala limestone of Florida, Georgia, and South Carolina; and some of the volcanic rocks of the Snake River Plain in Idaho and the Columbia Plateau.

Aquifers are important reservoirs storing large amounts of water relatively free from evaporation loss or pollution. If the annual withdrawal from an aquifer regularly exceeds the replenishment from rainfall or seepage from streams, the water stored in the aquifer will be depleted. This "mining" of ground water results in increased pumping costs and sometimes pollution from seawater or adjacent saline aquifers. Lowering the piezometric pressure in an unconsolidated artesian aquifer by overpumping may cause the aquifer and confining layers of silt or clay to be compressed under the weight of the overlying material. The resulting subsidence of the ground surface may cause structural damage to buildings, altered drainage paths, increased flooding, damage to wells, and other problems. Subsidence of 10 to 15 ft (3.0 to 4.6 m) has occurred in Mexico City and parts of the San Joaquin Valley of California. Careful management of aquifers is important to maintain their utility as a water source. *See* ARTESIAN SYSTEMS; GROUND-WATER HYDROLOGY.

Ray K. Linsley

Arachnida

The largest class of the subphylum Chelicerata in the phylum Arthropoda. Most familiar of the included orders are the spiders, scorpions, harvestmen, and mites and ticks. Arachnids are mainly terrestrial and may be the oldest of the Recent terrestrial animals; scorpions are known from the Silurian (over 4×10^8 years ago).

Characteristics. The arachnid body is divided into a cephalothorax (prosoma) and an abdomen (opisthosoma). The cephalothorax has six pairs of appendages: the chelicerae (jaws), a pair of pedipalps, and four pairs of walking legs. There are no antennae. The abdomen may be either segmented or unsegmented and usually lacks appendages, or the appendages may be modified into specialized structures, for example, the spinnerets of spiders.

As in other arthropods there is an exoskeleton, the outside layer of which (the epicuticle) consists mainly of lipids; it is water-repellent and prevents water loss. Respiration is by means of book lungs or tracheae or both. Book lungs are paired invaginations of the ventral abdominal wall; their structure indicates that they are derived from the gills of aquatic animals. They consist of thin leaves that are filled with hemolymph (blood) and hang into a cavity. Air enters this cavity through a slit and flows between the blood-filled leaves; oxygen is taken up by hemocyanin in the blood and transported. Some arachnids have book lungs alone, while in others one or both pairs of book lungs may be modified into tracheae (internal tubes similar to those of insects), which open to the outside through small holes (spiracles) and carry air directly to tissues.

Arachnids have an open circulatory system. The heart lies in the anterior dorsal part of the abdomen. Blood enters the heart from the pericardial chamber through small openings (ostia). The blood is pumped into the prosoma through an anterior artery, and posteriorly through a posterior artery. It flows through vessels and chambers and around the book lungs and then into the pericardial cavity and back to the heart.

Digestion takes place outside the body. The captured prey is held by the chelicerae while digestive enzymes are secreted over it; then the broth is sucked up.

Excretory organs may be either thin-walled coxal glands that open to the outside on the basal segment (coxa) of each leg, or Malpighian tubules which enter the midgut. Arachnids excrete nitrogenous waste as guanine.

The nervous system is concentrated in the cephalothorax except in the primitive scorpions,

which have ganglia in each segment. The brain encircles the esophagus; the protocerebrum lies above the esophagus, the remainder of the brain below. In most orders, the lower brain includes the ganglia of the appendages, ancestrally and in scorpions located in the thorax and abdomen.

Sense organs include simple eyes (ocelli), which consist of a cornea and lens that are continuous with the cuticle, a retinal layer with photoreceptors behind, and a vitreous body in between. Some eyes have a reflective layer (tapetum), which may improve vision in dim light. Hollow hairs located at the tips of palps and legs may be olfactory. Fine hairs (trichobothria) on the appendages are sensitive to air currents and vibrations. Membrane-covered pits in the exoskeleton occur all over the body and appendages, often in groups; called slit sense organs, they respond to tension in the exoskeleton and to some vibrations.

The genital opening and gonads are on the underside at the anterior of the abdomen. Males of many species perform an elaborate courtship. In many arachnids the male deposits a package of sperm (spermatophore). The female is attracted to the spermatophore or may be guided to it by the male, and takes it into her gonopore. Females produce yolk-rich eggs and may provide some care to their young.

The arachnids are predominantly predacious. The main exceptions are among the mites, where herbivores and parasites of plants and animals are common.

Classification. The Arachnida comprise more than 10 Recent orders. Scorpiones, the scorpions, have a segmented abdomen and a postabdomen bearing a sting; they are found worldwide. Palpigradi and Schizomida are minute arachnids inhabiting litter and soil. Uropygi, the tailed whip scorpions or vinegarones, are large (up to 6 cm or 2.5 in. in body length) with large pedipalps and a posterior flagellum; they are found mostly in warm climates. Amblypygi, the tailless whip scorpions, are flat, tropical animals with whiplike first legs; they capture insects with strong, spiny pedipalps. Araneae, the spiders, includes 34,000 known species worldwide; they are most abundant in the tropics. Spiders have a spherical spinneret-bearing abdomen attached to the cephalothorax by a thin stalk or pedicel. Solifugae (or Solpugida), the sun spiders, are quick-moving arachnids with large jaws; they inhabit deserts. Pseudoscorpionida, the pseudoscorpions, resemble scorpions but lack the post abdomen; smaller than 8 mm (0.3 in.), they respire with tracheae and are common in leaf litter worldwide. Opiliones, the harvestmen or daddy long- legs, have an eye on each side of a central anterior tubercle, and many have odor-producing glands. Males transmit sperm with a penis; females have a ovipositor. Members are found worldwide. Ricinulei, a small group living in tropical leaf litter or in caves, resemble ticks and have their jaws covered by a hood. Acari, the mites, form the largest group and are considered to comprise several orders by some workers. They are important because many plant and animal parasites are included, some of which (such as ticks) may transmit diseases. A few mites are found in fresh water and oceans. *See* AMBLYPYGI; ARANEAE; ARTHROPODA; CHELICERATA; OPILIONES; PALPIGRADI; PSEUDOSCORPIONIDA; RICINULEI; SCHIZOMIDA; SCORPIONES; SOLIFUGAE; UROPYGI. H. W. Levi

Araeolaimida

An order of nematodes in which the amphids are simple spirals that appear as elongate loops, shepherd's crooks, question marks, or circular forms. The cephalic sensilla are often separated into three circlets: the first two are papilliform or the second coniform, and the third is usually setiform; rarely are the second and third whorls combined. Body annulation is simple. The stoma is anteriorly funnel shaped and posteriorly tubular; rarely is it armed. Usually the esophagus ends in a bulb that may be valved. In all but a few taxa the females have paired gonads. Male preanal supplements are generally tubular, rarely papilloid.

There are three araeolaimid superfamilies: Araeolaimoidea, Axonolaimoidea, and Plectoidea. The distinguishing characteristics of the Araeolaimoidea are in the amphids (sensory receptors), stoma, and esophagus. The amphids are in the form of simple spirals, elongate loops, or hooks. Although araeolaimoids are chiefly found in the marine environment, many species have been collected from fresh water and soil. The amphids in the Axonolaimoidea are generally prominent features of the anterior end, visible as a single-turn loop of a wide sausage shape. Species have been collected only from marine and brackish waters. Their feeding habits are unknown. All known species occur in marine or brackish-water environments.

Plectoidea is a superfamily of small free-living nematodes, found mainly in terrestrial habitats, frequently in moss; some are fresh-water, and a few are marine. Those that inhabit moss cushions can withstand lengthy desiccation. For most, the feeding habit is unconfirmed; where known, they are microbivorous. Many are easily raised on agar cultures that support bacteria. *See* NEMATA. Armand R. Maggenti

Araeoscelida

An order of Paleozoic diapsid reptiles including the families Petrolacosauridae and Araeoscelidae. Members of these families resemble primitive modern lizards, such as the green iguana, in size and most body proportions, but are distinguished by their elongate necks and distal limb elements (see **illus.**). *Petrolacosaurus*, from the Upper Pennsylvanian of Kansas, is the earliest known diapsid. The skull shows well-developed upper and lateral temporal openings and a suborbital fenestra that are characteristic of the diapsids.

Araeoscelis, from the lower Permian of Texas, lacks the lateral temporal opening, but the retention

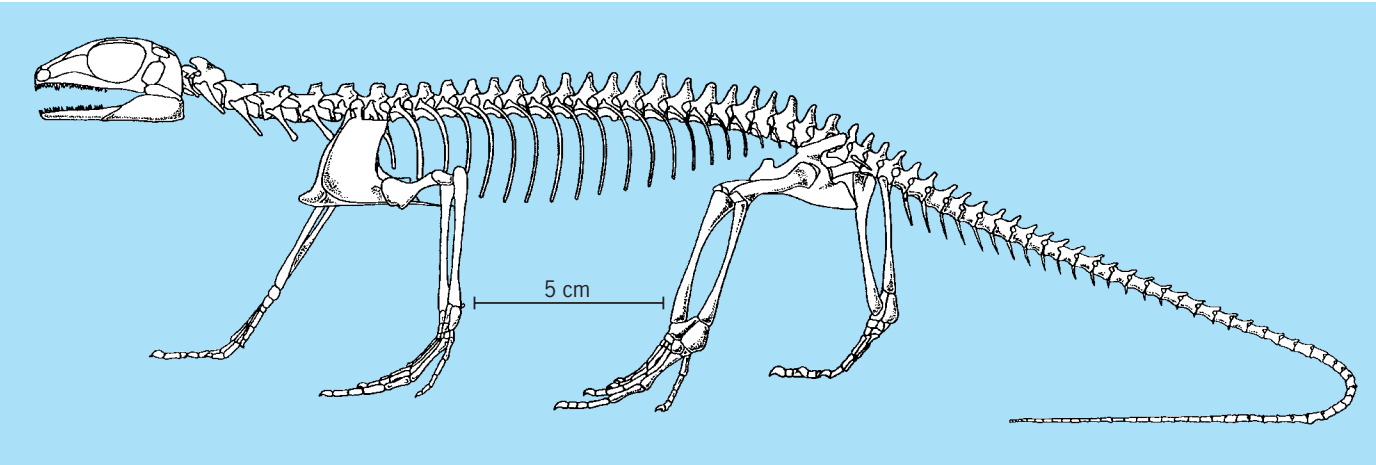

Reconstruction of *Petrolacosaurus kansensis*. (*After R. R. Reisz, A Diapsid Reptile from the Pennsylvanian of Kansas, Spec. Pub. Mus. Nat. Hist., Univ. Kans. 7, 1981*)

of a triradiate jugal suggests that the squamosal has secondarily expanded to fill this area. The cheek may have been re-elaborated to provide additional support for the expanded jaw musculature associated with the blunt teeth in this genus.

The following derived features distinguish members of the Araeoscelida from other early diapsids: six to nine elongate neck vertebrae; a radius as long as the humerus and a tibia as long as the femur; expanded neural arches; posterior cervical and anterior dorsal neural spines with mammillary processes; a coracoid process; and enlarged lateral and distal pubic tubercles. These characters distinguish the Araeoscelida as a monophyletic group that does not include the ancestors of any of the other major diapsid groups. Primitive features of the skeleton of araeoscelids demonstrate that diapsids have a sister group relationship with the protorothyrids, the most primitive known amniotes.

In contrast to later diapsids, members of the Araeoscelida show no evidence of an impedance-matching middle ear. The stapes is massive, the quadrate is not emarginated for support of a tympanum, and there is no retroarticular process. *See* REPTILIA.

Robert Carroll

Bibliography. R. R. Reisz, *A Diapsid Reptile from the Pennsylvanian of Kansas*, Spec. Pub. Mus. Nat. Hist., Univ. Kans. 7, 1981; R. R. Reisz, D. S. Berman, and D. Scott, The anatomy and relationships of the Lower Permian reptile *Araeoscelis, J. Vert. Paleontol.*, 1984.

Aragonite

One of three naturally occurring mineral forms of calcium carbonate ($CaCO_3$). The other forms (or polymorphs) are the abundant mineral calcite and the relatively rare mineral vaterite. Still other forms of calcium carbonate are known, but only as products of laboratory experiments. The name aragonite comes from Aragon, a province in Spain where especially fine specimens occur. *See* CALCITE; CARBONATE MINERALS.

Mineralogy. Aragonite has an orthorhombic crystal structure in which layers of calcium (Ca) atoms alternate with layers of offset carbonate (CO_3) groups. A common crystallographic feature of aragonite is twinning, in which regions of crystal are misoriented as though they were mirror images of each other. This can give rise to a pseudohexagonal symmetry which is readily identified in large crystals (see **illus.**). Aragonite crystals are usually colorless or white if seen individually; however, aggregates of small crystals may exhibit different colors. Most aragonites are nearly pure calcium carbonate; however, small amounts of strontium (Sr) and less commonly barium (Ba) and lead (Pb) may be present as impurities. *See* TWINNING (CRYSTALLOGRAPHY).

Stability. At the low temperatures and pressures found near the Earth's surface, aragonite is metastable and should invert spontaneously to calcite, which is stable at these conditions. This, in part, explains why calcite is far more abundant than aragonite. However, at low temperatures the transformation of aragonite to calcite effectively occurs only in the presence of water, and aragonite may persist for long periods of geologic time if isolated from water. Increased temperature also promotes the transformation to calcite. Despite being metastable, aragonite rather than calcite is sometimes the favored precipitate from certain solutions, such as seawater, in which magnesium (Mg) ions inhibit precipitation of calcite.

Aragonite becomes more stable than calcite only at very high pressures—approximately 3500 bars (350 megapascals) at 77°F (25°C) and 7000 bars (700 MPa) at 570°F (300°C). Since it is the high-pressure polymorph, aragonite has a higher density than calcite (2.94 versus 2.71 g/cm^3 at 77°F or 25°C). This is achieved by packing nine oxygen atoms (from six CO_3 groups) around each calcium, whereas only six oxygens surround each calcium in calcite.

Occurrence. Aragonite occurs most abundantly as the hard skeletal material of certain fresh-water and

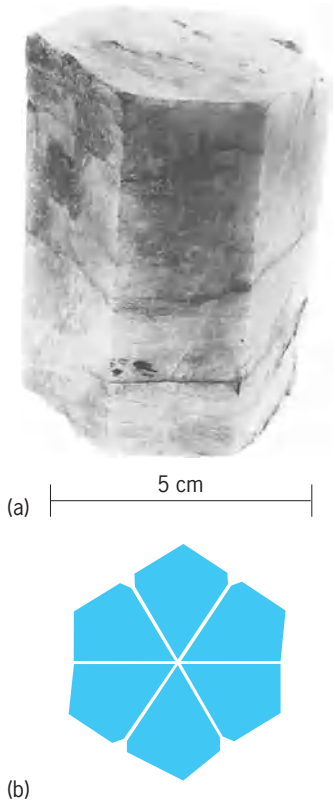

Aragonite. (*a*) Pseudohexagonally twinned specimen from Girgenti, Sicily (*American Museum of Natural History specimens*). (*b*) Arrangement of pseudohexagon twins (*after C. Klein, Dana's Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*).

marine invertebrate organisms, including pelecypods, gastropods, and some corals. The accumulated debris from these skeletal remains can be thick and extensive, usually at the shallow sea floor, and with time may transform into limestone. Most limestones, however, contain calcite and little or no aragonite. The transformation of the aragonite to calcite is an important step in forming limestone and proceeds by the dissolution of aragonite followed by the precipitation of calcite in the presence of water. This process may take more than 100,000 years. *See* LIMESTONE.

Aragonite also forms as a primary mineral during metamorphism at high pressures and moderate temperatures (330–570°F or 150–300°C). Its presence along with other high-pressure minerals in blueschist facies metamorphic rocks is regarded as important evidence for pressures in excess of 5000–7000 bars (500–700 MPa), corresponding to depths of greater than 10–15 mi (16–24 km) in Earth's crust. Examples of aragonite-bearing blueshists are found in the Franciscan Complex of the California Coast Ranges. *See* BLUESCHIST.

Other occurrences of aragonite include cave deposits (often in unusual shapes) and weathering products of calcium-rich rocks. Richard J. Reeder

Bibliography. L. L. Y. Chang, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, vol. 5B: *Non-Silicates: Sulphates, Carbonates, Phosphates, Halides*, 1995; C. Klein, *Dana's Manual of Mineralogy*, 21st ed., 1993; R. J. Reeder (ed.), *Carbonates: Mineralogy and Chemistry*, vol. 11 of *Reviews in Mineralogy*, Mineralogical Society of America, 1983.

Arales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Arecidae of the class Liliopsida (monocotyledons). It consists of two families, the Araceae, with only 1800 species, and the Lemnaceae, with only about 30.

The Araceae are herbs (seldom woody climbers) with ordinary roots, stems, and leaves (often broad and net-veined) and with the vessels confined to the roots. They have numerous tiny flowers grouped in a small to very large spadix. They are commonest in forested tropical and subtropical regions. *Anthurium* (elephant ear), *Arisaema* (jack-in-the-pulpit), *Dieffenbachia* (dumb cane), *Monstera*, and *Philodendron* are some well-known members of the Araceae.

The Lemnaceae, or duckweeds, are small, free-floating, thalloid aquatics that are generally conceded to be derived from the Araceae. Their flowers, seldom seen, are much reduced and form a miniature spadix. *Pistia* (water lettuce), a free-floating aquatic (but not thalloid) aroid is seen as pointing the way toward *Spirodela*, the least reduced genus of the Lemnaceae. *See* ARECIDAE; LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM. Arthur Cronquist

Araneae

A natural order of the class Arachnida, also called Araneida, commonly known as the spiders. These animals are widespread over most of the land areas of the world and are well adapted to many different habitats. They are known to be one of the oldest of all groups of arthropods, and their remains are known from the Devonian and Carboniferous geological deposits. Through successive geological periods spiders have become adapted to insects as their chief source of food. On the other hand, certain insects consume the eggs of spiders, others parasitize the eggs, and still others capture adults and place them in their nests for food for their young.

Morphology. Spiders have but two subdivisions of the body, the cephalothorax and the abdomen, joined by a slender pedicel (**Fig. 1**). All parts of the body are covered by chitinous plates which often extend into curious outgrowths, such as spines, horns, and tubercles. Only simple paired eyes, ocelli, are present with the number varying from eight, the most common number, to none in a few species inhabiting lightless habitats.

The first of six pairs of appendages are termed the chelicerae or jaws (**Fig. 2*a***), each of which consists of a stout basal segment and a slender distal fang. A poison gland is usually associated with the basal segment, and the related duct opens near the distal

end of the fang. A few spiders have lost their poison glands through retrogressive evolution. The second pair of appendages is the six-segmented pedipalps, simple organs of touch and manipulation in the female (Fig. 2*b*), but curiously and often complexly modified in males for use in copulation. The four pairs of thoracic legs consist of seven segments each, with considerable variation throughout the order (Fig. 2*c*). The spination of these appendages together with their relative lengths, associated sense organs, terminal claws, and other structures is all-important in classification and in adaptation.

Just in front of the genital opening in most females there is a more or less specific and often elaborately formed plate, the epigynum. This organ is also of importance in classification and in the reproductive activities of the female. The openings to the book lungs usually occur anterior to the epigynum. In most of the true spiders internal breathing tubules also occur with ventral openings posterior to the genital apertures. Distinctive paired ventral spinnerets occur near the posterior end of the abdomen. These vary in number from four to eight, eight being the most primitive condition. In certain families a sieve-like plate lies immediately anterior to the foremost spinnerets. From this plate, the cribellum, a special kind of banded silk is extruded and used in conjunction with a comb on the fourth legs. The spinnerets and cribellum are directly associated with several types of abdominal glands and responsible for the production of the different types of silk characteristic of these animals.

Silk. Silk produced by spiders is a scleroprotein which is fine, light, elastic, and strong. At present it is used industrially only in the making of cross hairs in optical instruments. This use is diminishing as metal filaments and etched glass come into more common usage. In addition to the attractive orb webs, these animals also construct sheet webs, funnel webs, tube webs, and reticular webs. The spider's reliance upon silk extends its use to the making of egg cocoons, sperm webs by males, molting sheets, gossamer threads for ballooning, attachment disks, lining for burrows, hinges for trap doors, binding for captives, retreats, and drag lines.

Classification. No general agreement exists at present among araneologists concerning the classification and exact arrangement of the families of spiders. A. Petrunkevitch (1939) recognized 5 suborders: Liphistiomorphae containing only 2 families with a primitively segmented abdomen, and restricted to regions in the Eastern Hemisphere; Mygalomorphae with 8 families; Hypochilomorphae with a single relict family native to the southern Appalachian region; Dipneumonomorphae with 48 families; and finally the Apneumonomorphae with 3 families, which lack the book lungs and represent the most highly modified members of the order. There is now a tendency to increase the number of recognized families.

Mygalomorphae. This suborder includes the trap-door spiders, purse-web spiders, and the tarantulas of the Western Hemisphere. These tarantulas are entirely different from those of southern Europe, which are wolf spiders, Lycosidae. American tarantulas reach the largest size of any known spiders. Those from the Amazon region may attain a body length of more than 3 in. (75 mm) with a leg span of 10 in. (250 mm). About 30 species living within the confines of the United States have been recognized, most of them in the Southwest. Perhaps the true trap-door spiders are those of most interest to the general public. These creatures have a rake on their chelicerae

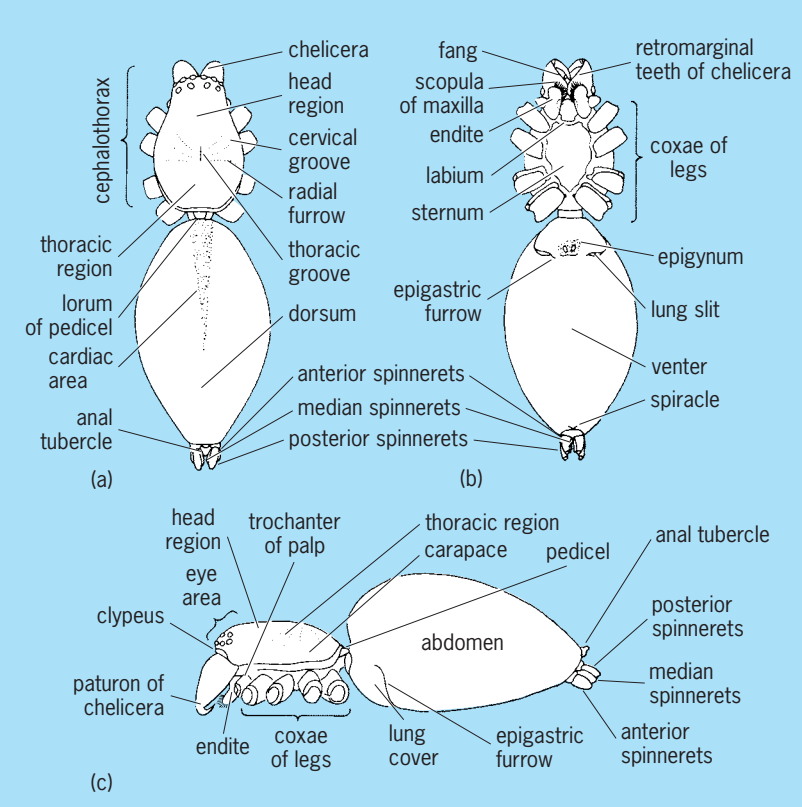

Fig. 1. Morphology of a spider. (*a*) Dorsal aspect. (*b*) Ventral aspect. (*c*) Lateral aspect. (*After B. J. Kaston, How to Know the Spiders, William C. Brown, 1953*)

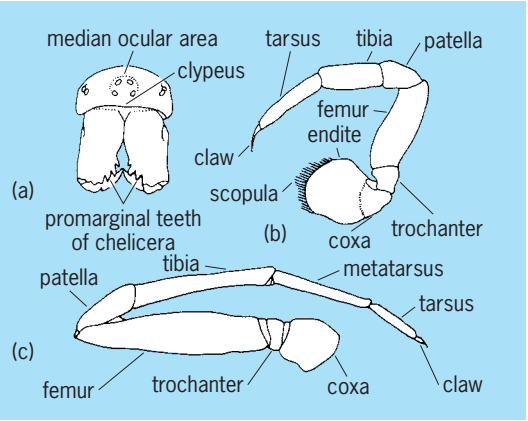

Fig. 2. Appendages of spider. (*a*) Front view of face showing the chelicerae. (*b*) Pedipalp of female. (*c*) Leg. (*After B. J. Kaston, How to Know the Spiders, William C. Brown, 1953*)

with which they perform their digging operations. Several species living in the southern and western United States are remarkable in that their burrows are waterproofed, lined with silk, and finally capped with a cleverly concealing lid.

Dipneumonomorphae. These are the spiders most commonly seen in the United States. The common grass spiders, members of the genus *Agelenopsis* and closely related genera, build their silken sheets in great profusion in grassy areas. The sheet terminates in a funnel in which the spider sits in readiness for a quick dash to obtain an insect caught in the silken platform. Several families are commonly named hunting spiders because they do not usually build webs but actively hunt their prey. The genera *Lycosa, Pardosa*, and *Pirata* are typical of the family Lycosidae. These actively pursue their prey over the ground, sometimes across bodies of water, and even dive beneath the surface for a time. Some species build retreats or tunnels. Some resemble the true trap-door spiders in their ability to build hinged doors over the mouths of their burrows, and others construct cleverly arranged turrets at the entrance to their burrows. Females typically construct a carefully built silken bag, fill it with eggs, and carry it attached to their spinnerets. When the young are ready to emerge from the sac, they are liberated or escape and climb upon the mother's body to be carried by her until they are capable of caring for themselves. The crab spiders have gained this name because of their crablike lateral movements. Some of these animals have the remarkable ability of changing their color to agree with that of their background. The jumping spiders, Salticidae, are probably the most colorful of any family group, especially in the tropics where they abound. They have keen vision and rapid movements, and are active and highly efficient hunters. In stalking insects they exhibit many interesting aspects of behavior. The courtship dances of males have been studied and carefully described. Males are often provided with conspicuous plumes and other ornaments for display during these activities. A dragline is always spun out behind each of these spiders as it goes jumping or walking about. They do not spin webs, but are skillful in constructing retreats, where they may retire at night or for the laying of eggs. Some of the most successful ant mimics belong in this family.

Much notoriety has been given to the several species of the genus *Latrodectus* belonging to the Theridiidae or comb-footed spiders. In the United States the females are commonly known as black widows because of their color and the popular belief that the female always eats the male immediately after mating. Their bad reputation has arisen from the general belief that they are very aggressive and readily attack human beings with little or no provocation. There is general agreement among araneologists, however, that the spiders are in reality timid and unaggressive. Their biting is a fear reaction, and a bite is not inflicted unless the animal is strongly stimulated in some way. The venom is a potent neurotoxin, and care should be taken to avoid the animals. This is especially important, of course, where children are concerned. Fatalities are rare and competent medical care will usually carry the bitten person through to complete recovery. *See* ARACHNIDA.

Arthur M. Chickering

Bibliography. W. S. Bristowe, *The World of Spiders*. 1958; J. H. Comstock, *The Spider Book*, 1965; W. J. Gertsch, *American Spiders*, 1949; B. J. Kaston, *Spiders of Connecticut*, Conn. State Geol. Nat. Hist. Surv. Bull. 70, 1948; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; A Petrunkevitch, *Catalogue of American Spiders*, pt. 1, *Trans. Conn. Acad. Arts Sci.*, vol. 33, 1939.

Arboretum

An area set aside for the cultivation of trees and shrubs for educational and scientific purposes. An arboretum differs from a botanical garden in emphasizing woody plants, whereas a botanical garden includes investigation of the growth and development of herbaceous plants, as well as trees and shrubs.

The largest of the arboretums in the United States is the Arnold Arboretum of Harvard University, founded in 1872, as the result of a bequest of James Arnold. More than 6500 kinds of woody plants are grown there. The United States National Arboretum, near Washington, D.C., was founded in 1927. Here hundreds of native and imported woody shrubs and trees are grown experimentally. Studies are made of shade and light tolerance, water requirements, temperature range, environmental influence, reproduction, and other botanical problems. *See* BOTANICAL GARDENS.

Earl L. Core

Arboriculture

A branch of horticulture concerned with the selection, planting, and care of woody perennial plants. Knowing the potential form and size of plants is essential to effective landscape planning as well as to the care needed for plants. Arborists are concerned primarily with trees since they become large, are long-lived, and dominate landscapes both visually and functionally.

Plants can provide privacy, define space, and progressively reveal vistas; they can be used to reduce glare, direct traffic, reduce soil erosion, filter air, and attenuate noise; and they can be positioned so as to modify the intensity and direction of wind. They also influence the microclimate by means of evaporative cooling and interception of the Sun's rays, as well as by reflection and reradiation. Certain plants, however, can cause human irritations with their pollen, leaf pubescence, toxic sap, and strong fragrances from flowers and fruit. Additionally, trees can be dangerous and costly: for example, branches can fall, and roots can clog sewers and break paving.

Plant selection. A plant's growth habit and its size at maturity are important characteristics in the selection process. It is important to select plants with roots that are free from kinks or circling. It is also desirable that trees be able to stand without support and have small branches along the trunk to provide protection and nourishment during establishment. Whether the plant is to be used for shade, screening, privacy, accent, or protection will determine what kind is selected. Sturdy root systems protect plants from being uprooted during storms, and if the roots are tolerant of poor soil the plants will remain healthy and can resist pests more readily. Leaves, flowers, and fruit are another consideration; not only are they visually important, but they also can have a considerable impact on maintenance.

Most plants grow best in deep, loamy soils, although some are able to withstand unfavorable soil conditions such as poor drainage, strata of different textures, extreme acidity or alkalinity, or chemical toxicities. The lowest recorded temperatures in a given region will determine which plants can survive in that region. Long-lived plants in particular should be selected to withstand the lowest temperatures expected.

Species native to an area may not perform well, however, particularly around buildings where microclimate and soil can differ considerably from preconstruction conditions.

Planting and care. Unless the soil has been compacted, the planting hole should only be deep enough to take the roots. The soil around the roots should be firmed and watered to settle it and remove air pockets. Unless the plant is overgrown, the only pruning needed after planting is some thinning to begin tree structure development, to shape shrubs, or to select vine canes to trellis. Little pruning is necessary for central-leader trees, conifers, and some hardwoods to grow strong and be well shaped. Species that become round-headed, however, may need considerable pruning to ensure the desired height of branching and a strong branch structure.

The less a branch or tree is pruned, the larger it will become. Therefore, only large branches that are too low or will compete or interfere with more desirable branches should be removed. Permanent branches, particularly of large-growing trees, should be at least one-third smaller in diameter than the trunk where they arise and be vertically spaced at least 18 in. (45 cm) apart along the trunk.

A tree will be more open and better retain its natural form if branches are removed completely (thinned) in contrast to being headed or stubbed back (see **illus.**). Heading concentrates subsequent growth just below the pruning cut and results in dense foliage with weakly attached branches. In removing a branch the final cut should be just to the outside of the branch bark ridge in the crotch and the collar below. Such a cut minimizes the size of the wound and the possibility that the trunk will decay. Seldom is it advisable to paint pruning wounds.

Fertilization of young plants is necessary to achieve rapid growth; mature plants, however, may need little or no added nutrients. Nitrogen is almost universally deficient in soils, and it usually is the only element to which trees and large shrubs will respond. Nitrogen fertilizers are water-soluble and can be applied to the soil or lawn surface and then watered in. In alkaline soils, the availability of iron or manganese may be so low for certain kinds of plants that they exhibit the typical pale leaves with narrow (due to iron deficiency) or wide (due to manganese deficiency) darker green extending bands along the veins. Increasing soil acidity or applying chelated nutrients usually ameliorates these problems. *See* FERTILIZER.

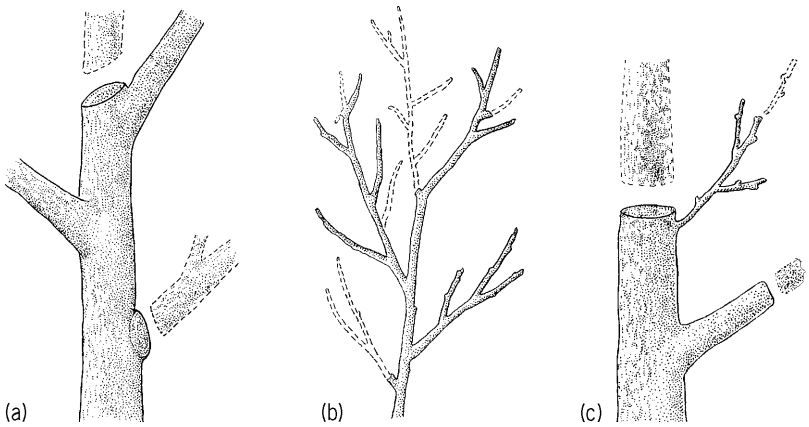

Methods of plant pruning. (*a*) Thinning (removing a branch at its origin) of a mature tree. (*b*) Thinning a young tree. (*c*) Heading (pruning to a small lateral branch or to a stub) of a mature tree. (*After R. W. Harris, Arboriculture: Care of Trees, Shrubs, and Vines in the Landscape, Prentice-Hall, 1983*)

Irrigation can ensure establishment of young plants, the attractive appearance of foliage, and even survival. Many mature plants can endure long periods without rain or irrigation, if a large volume of moist soil is present and the plants have extensive root systems. If the water supply is limited, it is important for the soil to be fully moist at the start of the growing season. A few heavy irrigations are more efficient than frequent light irrigations; more plants do poorly from too much water than not enough. *See* PLANT-WATER RELATIONS.

Mulch, that is, material placed on the soil surface, can control weeds, protect the soil from compaction and erosion, conserve moisture, moderate soil temperatures, provide an all-weather surface for walking, and allow plants to root in the most fertile and well-aerated surface soil. A wide range of organic and inorganic or synthetic materials can be employed. Mulch should be kept 2 in. (5 cm) away from the trunks of plants to reduce trunk disease and rodent damage. *See* LANDSCAPE ARCHITECTURE.

Richard W. Harris

Bibliography. B. Ferguson, *All about Trees*, 1982; C. W. Harris and N. T. Dines (eds.), *Time-Saver Standards for Landscape Architecture* 2d ed., 1997; R. W. Harris, *Arboriculture: Care of Trees, Shrubs, and Vines in the Landscape*, 1983; J. Hudak, *Trees for Every Purpose*, 1980; A. L. Shigo, *A New Tree Biology*, 1936; B. B. Wilson, *The Growing Tree*, 1984.

Arborvitae

A plant, sometimes called the tree of life, belonging to the genus *Thuja* of the order Pinales. It is characterized by flattened branchlets with two types of scalelike leaves. At the edges of the branchlets the leaves may be keeled or rounded; on the upper and lower surfaces they are flat, and often have resin glands. The cones, about $^1/_2$ in. (1.2 cm) long, have the scales attached to a central axis (see **illus.**). *See* PINALES; RESIN; SECRETORY STRUCTURES (PLANT).

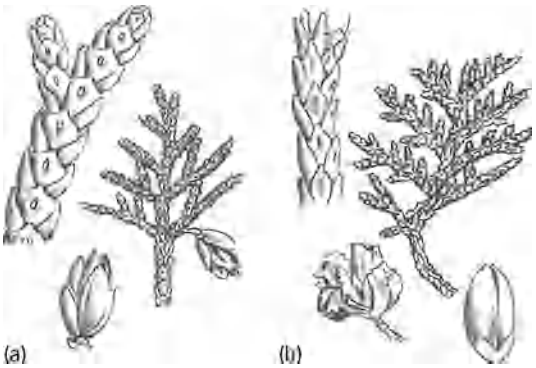

Examples of arborvitae. (*a*) Eastern arborvitae (*Thuja occidentalis*). (*b*) Oriental arborvitae (*T. orientalis*).

The tree is valued both for its wood and as an ornamental. *Thuja occidentalis*, of the eastern United States, grows to a height of 60 ft (18 m). Known as the northern white cedar, it occurs in moist or swampy soil from Nova Scotia to Manitoba and in adjacent areas of the United States, and extends south in the Appalachians to North Carolina and Tennessee. The wood is light, soft, and fragrant. It resists deterioration from moisture, and is used for construction, cabinetmaking, and making barrels and shingles. A large number of arborvitae seedlings are sold annually as ornamentals.

Other important species include the giant arborvitae (*T. plicata*), which grows to 180 ft (54 m) in the Pacific Northwest and is highly prized for its resistance to decay; oriental arborvitae (*T. orientalis*); and Japanese arborvitae (*T. standishii*). Among the horticultural forms are the dwarf, pendulous, and juvenile varieties. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Arboviral encephalitides

A number of diseases, such as St. Louis, Japanese B, and equine encephalitis, which are caused by arthropod-borne viruses (arboviruses). In their most severe human forms, the diseases invade the central nervous system and produce brain damage, with mental confusion, convulsions, and coma; death or serious aftereffects are frequent in severe cases. Inapparent infections are common (see **table**).

Arboviruses. The arbovirus "group" comprises more than 250 different viruses, many of them differing fundamentally from each other except in their ecological property of being transmitted through the bite of an arthropod. A large number of arboviruses of antigenic groups A and B are placed in the family Togaviridae, in two genera, alphavirus (serological group A) and flavivirus (serological group B). Still other arboviruses, related structurally and antigenically to one another but unrelated to Togaviridae, are included in the family Bunyaviridae, consisting chiefly of the numerous members of the Bunyamwera supergroup—a large assemblage of arboviruses in several antigenic groups which are cross-linked by subtle interrelationships between individual members. The nucleic acid genomes of all arboviruses studied thus far have been found to be RNA; most arboviruses are inactivated by lipid solvents (ether or sodium deoxycholate) and are unstable at room temperature. *See* VIRUS CLASSIFICATION.

Arboviruses will grow in many laboratory animals and in embryonated eggs and tissue cultures, often inapparently; they are almost universally pathogenic for newborn mice. Of the groups that have been established on the basis of antigenic relationships, most of those causing encephalitis belong to group A or B or the California complex. Animals inoculated repeatedly with one virus may develop serum antibodies against other group members. Depending upon the closeness of relationship, heterologous responses range from low and transient up to responses nearly equaling the homologous response. In some instances only hemagglutination-inhibiting antibodies cross, whereas in others complement-fixing and neutralizing antibody crossings occur. Within the arbovirus groups are several important members which, although neuropathogenic in the mouse and antigenically related to those causing encephalitis in humans, invade the human central nervous system only atypically.

Members of serological group A include western equine encephalitis (WEE), eastern equine encephalitis (EEE), and Venezuelan equine encephalitis (VEE) viruses; and Mayaro, Semliki Forest, Chikungunya, and Sindbis viruses, which have nonencephalitic syndromes. Group A viruses are chiefly mosquito-borne. Serological group B viruses include Japanese B, St. Louis, and Murray Valley encephalitis viruses (mosquito-borne), and the viruses of the Russian tick-borne complex, some of which produce encephalitis (Russian spring-summer), whereas others cause hemorrhagic fevers (Omsk, Kyasanur Forest) or other syndromes, such as louping ill. Also in group B are the nonneurotropic viruses of West Nile fever, yellow fever, dengue, and other diseases. *See* LOUPING ILL; YELLOW FEVER.

Pathogenesis. Illness begins 4–21 days after the bite of the infected vector, with sudden onset of fever, headache, chills, nausea, and generalized pains; marked drowsiness and neck rigidity may ensue. Signs and symptoms are often diphasic, the second phase being more severe and involving mental confusion, speech difficulties, convulsions, and coma. Mortality case rates vary in different diseases and epidemics. Mortality in Japanese B encephalitis in older age groups has been reported to be as high as

Summary of six major human arbovirus infections which occur in the United States

Disease	Exposure	Distribution	Vectors	Infection:case ratio (age incidence)	Sequelae	Mortality rate, %
Western equine encephalitis	Rural	Pacific Mountain West Central Southwest	*Culex tarsalis*	50:1 (under 5) 1000:1 (over 15)	+	2–3
Eastern equine encephalitis	Rural	Atlantic Southern coastal	*Aedes sollicitans* *A. vexans*	10:1 (infants) 50:1 (middle-aged) 20:1 (elderly)	+	50–70
St. Louis encephalitis	Urban-rural	Widespread	*C. pipiens* *C. quinquefasciatus* *C. tarsalis* *C. nigrapalpus*	>400:1 (young) 64:1 (elderly)	±	5–10
Venezuelan equine encephalitis	Rural	South (also South America, Central America)	*Aedes* *Psorophora* *Culex*	Unknown ratio	Unknown	0.5%
California encephalitis	Rural	North Central Atlantic South	(*Aedes* sp.?)	Unknown ratio (most cases under 20)	±	Fatalities rare
Colorado tick fever	Rural	Pacific Mountain	*Dermacentor andersoni*	Unknown ratio (all ages affected)	Rare	Fatalities rare

80%. Serious aftereffects such as mental and sensory defects, epilepsy, and paralysis occur. Abortive cases may produce the early symptoms or resemble aseptic meningitis. California encephalitis is seldom fatal, and prognosis excellent, but convalescence may be prolonged.

Laboratory diagnosis. Testing for hemagglutination-inhibiting (HI) serum antibodies may be complicated by antigenic cross-reactions, but can establish the arboviral group responsible. Specific identification depends on the detection of variations in the HI, complement-fixing, or neutralizing antibody titers between appropriately timed acute and convalescent specimens. Virus isolation from the blood or spinal fluid is usually not successful.

Epidemiology. All arboviruses causing encephalitis require an infected arthropod vector for transmission. The usual natural hosts are mammals or birds, with humans accidentally infected in the cycle. Ticks may serve not only as vectors but also as a natural reservoir, because the virus may pass transovarially to offspring; in other instances birds or mammals are the probable reservoir. In highly endemic regions many persons have antibodies, chiefly from inapparent infections. Each virus is limited to definite geographic areas. The population's immunity to its local virus may cross-immunize against closely related viruses. For example, endemic West Nile fever may partially protect its areas against infections by other flaviviruses, some of which can cause encephalitis. Another hypothesis suggests a common ancestor, different descendants having adapted to various regions, each with its special vector and reservoir.

Treatment. There is no proved specific treatment. In animals, hyperimmune serum given early may prevent death. Killed virus vaccines have been used in animals and in persons occupationally subjected to high risk. A live, attenuated vaccine against Japanese B encephalitis virus, developed in Japan, has been used experimentally with some success, not only in pigs to reduce amplification of the virus in this important vertebrate reservoir but also in limited trials in humans. In general, however, control of these diseases continues to be chiefly dependent upon elimination of the arthropod vector. *See* ANIMAL VIRUS; VIRUS.

Joseph L. Melnick

Bibliography. E. Jawetz, J. L. Melnick, and E. A. Adelberg, *Review of Medical Microbiology*, 12th ed., 1976.

Arc discharge

A type of electrical conduction in gases characterized by high current density and low potential drop. The electric arc was discovered by Humphry Davy in 1808, when he connected a piece of carbon to each side of an electric battery, touched the two pieces of carbon together, then drew them slightly apart. The result is a dazzling steam of ionized air, or plasma, at a temperature of 6000°C (10,800°F), the surface temperature of the sun. A typical arc runs at a voltage drop of 100 V with a current drain of 10 A. The arc has negative resistance-the voltage drop decreases as the current increase-so a stabilizing resistor or inductor in series is required to maintain it. The high-temperature gas rises like a hot-air balloon while it remains anchored to the current-feeding electrodes at its ends. It thereby acquires an upward-curving shape, which accounts for its being called an arc.

Applications. There are many applications of such an intensely hot object. The brilliant arc and the

incandescent carbon adjacent to it form the standard light source for movie theater projectors. The electronic flashgun in a camera uses an intense pulsed arc in xenon gas, simulating sunlight. Since no solid-state material can withstand this temperature for long, the arc is used industrially for welding steel and other metals. Alternatively, it can be used for cutting metal very rapidly. Electric arcs form automatically when the contacts in electrical switches in power networks are opened, and much effort goes into controlling and extinguishing them. Lightning is an example of a naturally occurring electric arc. *See* ARC HEATING; ARC LAMP; ARC WELDING; CIRCUIT BREAKER; LIGHTNING; OPTICAL PROJECTION SYSTEMS; STROBOSCOPIC PHOTOGRAPHY; WELDING AND CUTTING OF MATERIALS.

Temperature. The arc has been pushed to extremely high temperatures in the search for thermonuclear fusion, the record temperature being 4×10^5°C in a long pulse in helium. The arc temperature appears to be limited by the energy lost in intense radiation from the interface region between the intensely hot, fully ionized plasma core and the surrounding cooler, un-ionized gas.

Structure. The structure of an electric arc has several components. The most prominent feature is the brilliant column of partially ionized gas, or plasma, referred to as the arc body. The arc body in general has a relatively low voltage drop along its length. The percentage of ionization can be calculated from Saha's equation, shown below. Here, n_0 and n_i are

$$\log_{10}\left(\frac{n_i^2}{n_0}\right) = -5040\frac{v_i}{T} + \frac{3}{2}\log_{10}(T) + 15.385$$

the neutral and ionic densities (cm^{-3}), respectively, v_i the ionization potential in volts, and T the absolute temperature in kelvins.

The electrons exciting the arc are emitted from an intensely hot spot on the negative electrode, or cathode. This spot is heated by positive ion bombardment from the plasma column. The electrons exit the arc from a second brilliant spot on the positive electrode, or anode. Ion bombardment also occurs here. The arc column automatically becomes positive with respect to both electrodes in order to contain the hot, rapidly moving, negative electrons, and this positive potential ejects ions. Both electrodes are separated from the plasma column by a thin transition region, the Debye sheath. The voltage drop between the cathode and the arc body is called the cathode fall, and is in general a major source of energy to the arc body. The voltage increase between the arc body and the anode is called the anode rise. The details of the structure of the electric arc are still not fully understood.

Anode spots. If the current to the anode is not sufficient to maintain the discharge, an interesting phenomenon can frequently be observed. Spherical luminous plasmas that are called anode spots or fireballs form on the anode (see **illus.**). Their formation and size depend on the ionization potential, the background pressure, and the possible impurities in the host gas. There plasmas usually are found at an imperfection in a smooth surface where the electric field is larger. Once formed, they are very stable and can exist for hours.

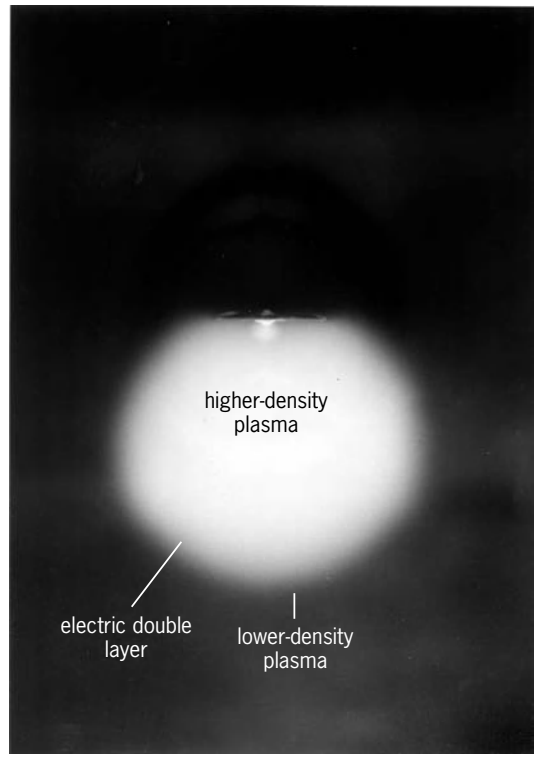

Anode soot or fireball in argon. (*From B. Song, N. D'Angelo, and R. L. Merlino, On anode spots, double layers and plasma contractors, J. Phys. D: Appl. Phys., 24:1789–1795, 1991*)

Experiments have shown that anode spots have an electric potential that is slightly greater than the ambient plasma in which they reside. The indication is that some of the higher-mobility electrons have escaped, leaving the anode spot with a slight excess positive charge. Under certain conditions, more than one anode spot can form. If a circular anode contains an edge in the shape of a sharp lip, the additional anode spots will be located as far from each other as possible on the circular lip. Two anode spots will be at the opposite ends of the diameter, three will be at the corners of a square, and so forth. The mutual repulsion of anode spots and their search for a stable equilibrium position result from the fact that each one has approximately the same amount of net positive charge enclosed within it.

Simulation of natural phenomena. Some experiments have been directed toward simulating natural phenomena in the laboratory. If a voltage pulse is applied to the anode, the high-density anode spot explodes into the lower-density ambient plasma. Voltage signals detected by small movable Langmuir probes within the plasma can be displayed by using an oscilloscope that is triggered by the voltage pulse. Hence, the spatial and temporal characteristics of the expansion of a higher-density plasma into a lower-density plasma can be monitored and investigated. The stable transition region from the higher-density

spherical plasma to the ambient lower-density spherical plasma to the ambient lower-density plasma is similar to phenomena found in space. The term double layer is applied to the two adjacent laminar charge-density regions that are found within the transition region. The signs of the two charged regions are opposite, with the positively charged region being adjacent to the higher-potential side of the double layer. *See* ELECTRICAL CONDUCTION IN GASES.

Igor Alexeff; Karl E. Lonngren

Arc heating

The heating of matter by an electric arc. The matter may be solid, liquid, or gaseous. When the heating is direct, the material to be heated is one electrode; for indirect heating, the heat is transferred from the arc by conduction, convection, or radiation. *See* ARC DISCHARGE.

At atmospheric pressure, the arc behaves much like a resistor operating at temperatures of the order of thousands of kelvins. The energy source is extremely concentrated and can reach many millions of watts per cubic meter. Almost all materials can be melted quickly under these conditions, and chemical reactions can be carried out under oxidizing, neutral, or reducing conditions.

Arc heating is applied in, but is not limited to, electric arc furnaces (both open and submerged arc), plasma torches (including cutting, spraying, and gas heating), arc welders, arc machining, and arc boring. By far the greatest application in installed capacity is in electric arc furnaces. *See* ARC WELDING; ELECTRIC FURNACE; WELDING AND CUTTING OF MATERIALS.

Direct-arc furnaces. In a direct-arc furnace, the arc strikes directly between the graphite electrodes and the charge being melted. These furnaces are used in steelmaking, foundries, ferroalloy production, and some nonferrous metallurgical applications. Although an extremely large number of furnace types are available, they are all essentially the same. They consist of a containment vessel with a refractory lining, a removable roof for charging, electrodes to supply the energy for melting and reaction, openings and a mechanism for pouring the product, a power supply, and controls. The required accessory components include water-cooling circuits, gas cleaning and extraction equipment, cranes for charging the furnace, and ladles to remove the product. Because the electrodes are consumed by volatilization and reaction, a mechanism must be provided to feed them continuously through the electrode holders.

Most direct-arc furnaces operate in a batch mode whereby they are charged with raw materials (scrap steel, ores, and fluxes) which are then melted down to a two-phase liquid of slag and metal. The metal is refined by the addition of further raw materials, and finally the slag and metal products are removed from the furnace by pouring.

In a typical direct-arc furnace for steelmaking (**Fig. 1**), three electrodes are used in three-phase alternating-current operation. Furnace powers may vary from a few megavolt-amperes to over 100 MVA, and furnace inside diameters may range from less than 3 ft (1 m) to more than 20 ft (7 m). Generally, the furnace shell is cylindrical and the bottom is hemispherical; in larger furnaces the bottom is mounted on rockers for pouring. Because productivity is almost as important as energy economy in modern steelmaking furnaces, large portions of the furnace are water-cooled to reduce refractory erosion and minimize shutdown time for maintenance. *See* ELECTROMETALLURGY; FURNACE CONSTRUCTION; STEEL MANUFACTURE.

Refractory lining. The furnace lining functions to contain the liquid product within the furnace and to reduce heat losses. It must be carefully chosen for chemical and thermal compatibility with the specific process being used and must have sufficient strength and resilience to withstand the rigors of charging by heavy loads of scrap steel. In basic steelmaking, the bottom usually consists of a subhearth of magnesia bricks topped by a monolithic layer of rammed high-magnesia refractory which forms the hearth. The side walls are usually made of key-shaped magnesia bricks or magnesia—carbon bricks. The slag line is particularly susceptible to erosion, and therefore a wide variety of different materials are used. Since the water-cooled side walls are eventually coated with a layer of slag, only a thin layer of refractory is required for startup. The roof of the furnace, which generally is also water-cooled, is commonly made of 70% alumina brick. *See* REFRACTORY.

Electrical supply. Furnace transformers for direct-arc furnaces normally accept input voltages up to 69 kV, and so a step-down transformer is normally needed

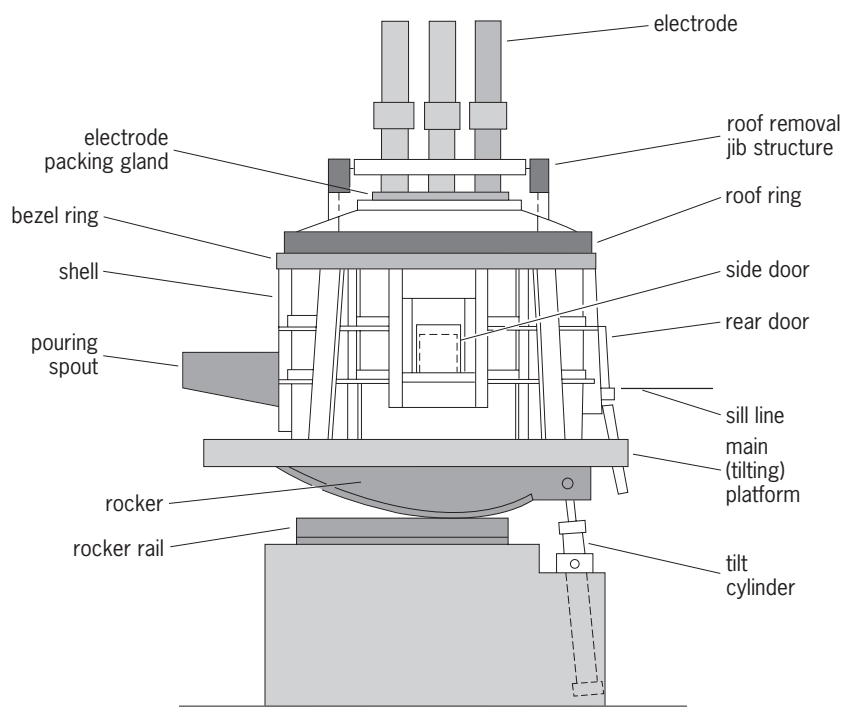

Fig. 1. Schematic representation of an electric arc furnace used in steelmaking. (*After AIME, Iron and Steel Society, Electric Furnace Steelmaking, ed. by C. R. Taylor, L. G. Kuhn publisher, 1985*)

in a substation if the utility voltage is more than this value. The primary side of the power supply consists of a disconnect switch, an oil circuit breaker (to protect equipment on the primary side of the furnace transformer), the furnace-operating vacuum switch, and the primary transformer windings with tap changers. The operating switch must be particularly rugged since it may have to interrupt furnace power up to 60 times per day. The secondary furnace circuit consists of the secondary windings on the transformer; flexible cables, which allow tilting of the furnace and movement of the roof; tubular bus-bars, which carry current to the electrodes; and electrodes. The transformer and most current-carrying lines are heavily water-cooled. The correct layout of the cables and bus tubes to the electrodes is very important to stable furnace operation so that the mutual coupling between the phases can be equalized and the currents between the phases balanced.

Instrumentation, additional reactors (which may have to be added to smaller furnace power supplies to limit the arc current during short circuit), and devices for power factor correction and harmonic suppression are also part of the electrical supply. The power levels required at different stages of the steelmaking cycle vary considerably, being low for startup, high for melt-down, and medium during refining. This power is controlled by adjusting the voltage taps on the primary of the furnace transformer to give the desired operating voltage and by continuously moving the electrodes up and down to adjust the current to the desired level. The latter operation is automated in modern furnaces by means of computer-controlled electric motors or hydraulic systems. Modern furnace transformers have at least six voltage taps; since the transformer can be changed from delta to star connection, this gives a choice of 12 voltage settings. Tap settings may be changed on load or off load; off-load changers are more common. *See* TRANSFORMER; WIRING.

Submerged-arc furnaces. The arcs are submerged below the solid feed and sometimes below the molten product. Submerged-arc furnaces differ from those used in steelmaking in that raw materials are fed continuously around the electrodes and the product and slag are tapped off intermittently. The furnace vessel is usually stationary. Submerged-arc furnaces are often used for carbothermic reductions (for example, to make ferroalloys), and the gases formed by the reduction reaction percolate up through the charge, preheating and sometimes prereducing it. Because of this, the energy efficiency of this type of furnace is high. The passage of the exhaust gas through the burden also filters it and thus reduces air-pollution control costs.

Plasma torches and devices. Although carbon arcs are plasmas, common usage of the term plasma torch suggests the injection of gas into or around the arc. This gas may be inert, neutral, oxidizing, or reducing, depending on the application and the electrodes used. Plasma torches are available at powers ranging from a few kilowatts to over 10 MW; usually they use

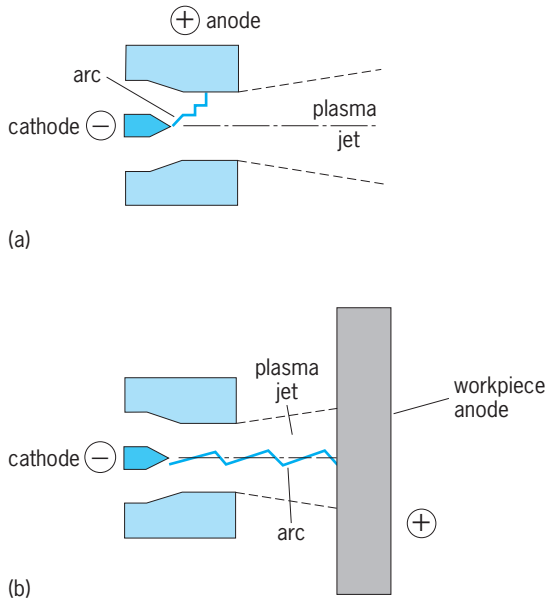

Fig. 2. Diagrams of small-scale plasma torches. (*a*) Nontransferred-arc torch. (*b*) Transferred-arc torch.

direct-current electricity and water-cooled metallic electrodes.

Small scale. Small-scale (20–150-kW) plasma torches are used in spraying and metal-cutting operations as well as in laboratory-scale research. Plasma spraying uses a plasma jet formed by striking an arc between a thoriated tungsten cathode and a cylindrical copper anode (**Fig. 2***a*). This type of plasma torch is called a nontransferred-arc torch since the arc remains within the torch and is not transferred externally. Particles of the material to be sprayed are injected radially into this jet where they are entrained, melted, and projected onto the substrate to be coated. Metallic or ceramic powders may be used to coat existing parts or to build up new parts on a mandrel. Since the plasma gas and surrounding atmosphere must be carefully controlled to give the desired coating properties, the entire process is usually highly automated and computer-controlled. Plasma spraying is widely used in the aerospace industry and in the manufacture of metallurgical and chemical process equipment. *See* WELDING AND CUTTING OF MATERIALS.

Plasma cutting is done in the transferred-arc mode where the arc is struck between the cathode and the workpiece which forms the anode (Fig. 2*b*). A large fraction of the arc power is liberated at the anode, thus melting the metal which is blown away by a high-velocity plasma jet. In both spraying and cutting operations, the plasma gas must be nonoxidizing to avoid rapid erosion of the tungsten cathode; oxidizing gases may be added downstream of the cathode in some applications.

Large scale. Large-scale plasma torches may also be operated in the transferred- or nontransferred-arc mode. In these torches, the internal electrodes are usually water-cooled cylinders. For nontransferred-arc operation, plasma gas is introduced tangentially

between the two electrodes and the arc is struck across the gap. The arc length may be increased, to increase the plasma voltage, by inserting a number of electrically insulated water-cooled cylindrical spacers between the cathode and anode. Additional gas may also be injected through gaps between these spacers. The advantage of increasing arc voltage is to minimize the current required for a given arc power; this reduces power supply cost, resistance losses, and electrode erosion. Erosion of the electrodes is further reduced by the proper choice of electrode material, electrode cooling, and a combination of aerodynamic and magnetic rotation of the arc to reduce its residence time at any location.

Other applications. The nontransferred-arc operation is used for the heating of large quantities of gas which may then be used as a heating fluid or reactant. Torches of this type are used for the production of acetylene from various hydrocarbons, for the production of titanium dioxide pigment, and in a number of metallurgical operations, including ironmaking and the treatment of metallurgical dusts. An application with commercial potential is the superheating of blast furnace tuyere gases.

Transferred-arc operation is used for the heating of electrically conducting solids or liquids in which the product forms one electrode. High efficiencies can be obtained even at relatively low gas flow rates since only the internal torch electrode need be cooled. The main applications for this type of plasma torch are the smelting and remelting of ferroalloys, the melting of specialty steels and alloys, and ladle heating.

Arc remelting. In arc remelting applications, material which has been melted by conventional means is remelted for consolidation or purification. The process is used for highly reactive metals such as zirconium, titanium, and superalloys. A great variety of geometries and operating conditions are used, inlcuding vacuum-arc remelting (VAR), vacuum-arc double-electrode remelting (VADER), plasma-arc remelting (PAR), and electroslag remelting (ESR). In all cases, the product is collected into a water-cooled mold, and sometimes an ingot is continuously withdrawn. In most cases, the material to be remelted forms one electrode while the product forms the second. The differences between the systems are essentially the arc atmosphere (vacuum, plasma gas, or slag) and the geometry of electrodes used. Sometimes, particulate material is fed, and plasma torches are used as electrodes. In vacuum-arc double-electrode remelting, a vacuum arc is struck between two electrodes made of consolidated raw material, and the product drips into a water-cooled crucible below them. Because the material in the crucible is not directly heated, better control over the product macrostructure is possible.

Direct-current carbon arc furnaces. A direct-current arc is more stable than its alternating-current counterpart, and can, therefore, be operated at lower current and higher voltage by increasing the arc length. This reduces both the electrode diameter and the electrode consumption compared to alternating-current operation at similar powers. Tests have also shown that injecting gas through a hole drilled through the center of the electrode further increases stability and reduces wear. Powdered ore and reductants may be injected with this gas, reducing the need for agglomerating the arc furnace feed.

In most cases, direct-current carbon arc furnaces have one carbon electrode, with the product forming the second electrode. The current is usually removed from the furnace through a bottom constructed of electrically conducting material. Several direct-current plasma furnaces with powers ranging from 1 to 45 MW are in operation. R. J. Munz

Bibliography. American Institute of Mining, Metallurgical, and Petroleum Engineers, Iron and Steel Society, *Electric Furnace Steelmaking*, ed. by C. R. Taylor, 1985; AIME, Iron and Steel Society, *Plasma Technology in Metallurgical Processing*, ed. by J. Feinman, 1987; AIME, *Proceedings of Electric Furnace Steel Conferences*, annually; A. Choudhury, *Vacuum Metallurgy*, ASM International, Materials Information Society, 1990; A. G. E. Robiette, *Electric Smelting Practice*, 1973.

Arc lamp

A type of electric-discharge lamp in which an electric current flows between electrodes through a gas or vapor. In most arc lamps the light results from the luminescence of the gas; however, in the carbon arc lamp a major portion of the light is produced by the incandescence of one or both electrodes. The color of the arc depends upon the electrode material and the surrounding atmosphere. Most lamps have a negative resistance characteristic so that the resistance decreases after the arc has been struck. Therefore some form of current-limiting device is required in the electric circuit. For other electric-discharge lamps. *See* VAPOR LAMP.

The carbon arc lamp was the first practical commercial electric lighting device, but the use of arc lamps is limited. In many of its previous functions, the carbon arc lamp has been superseded by the high-intensity mercury-vapor lamp. Arc lamps are now used to obtain high brightness from a concentrated light source, where large amounts of radiant energy are needed and where spectral distribution is an advantage. Typical uses of arc lamps are in projectors, searchlights, photography, therapeutics, and microscope lighting, and for special lighting in research. *See* MERCURY-VAPOR LAMP.

Carbon arc lamp. The electrodes of this lamp are pure carbon. The lamp is either open, enclosed, or an intensified arc with incandescence at the electrodes and some light from the luminescence of the arc. The open-arc form of the carbon arc is obsolete. It is unstable on a constant voltage supply, although on a constant current system the operating characteristics are good. In the enclosed type, a restricted supply of air slows the electrode consumption and secures a life of approximately 100 h. The intensified type used a small electrode, giving a higher-intensity output that is more white in color and more steady

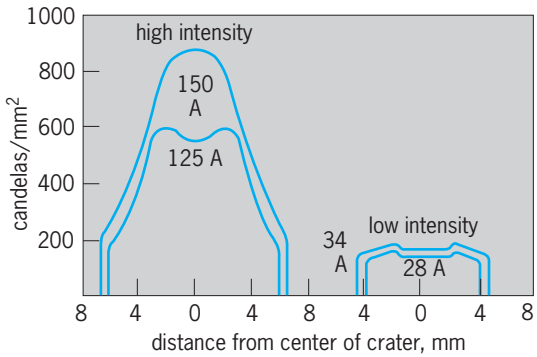

Crater brightness distribution in forward direction for typical low- and high-intensity carbons. 1 mm = 0.04 in.

than either the open or enclosed types. The electrode life is approximately 70 h. In the high-intensity arc the current may reach values of 125–150 A with positive volt-ampere characteristics, and the lamp may operate directly from the line without a ballast. The **illustration** shows two basic differences in brightness characteristics of the low-intensity and high-intensity arc lamps. The brightness of the high-intensity arc is higher and depends markedly on current. *See* LUMINOUS INTENSITY.

Although carbon arc lamps are built for alternating-current operation, direct current produces a steadier and more satisfactory operation. To start the arc, the carbons are brought together for a short period and then separated to form an arc between the electrodes. Some forms of the lamps introduce a third electrode between the main electrodes for a brief period of time to start the lamp. As the carbon burns away, the arc gap must be adjusted for efficient operation. This may be done by manual adjustment; however, an automatic feed mechanism is more efficient and satisfactory.

Flame arc lamp. The flame arc lamp radiates light from the arc instead of from the electrode. The carbon is impregnated with chemicals, which are more volatile than the carbon and, when driven into the arc, become luminous. The chemicals commonly used are calcium, barium, titanium, and strontium, which produce their characteristic color of light as chemical flames and radiate efficiently in their specific wavelengths in the visible spectrum. Some flames use several chemicals, some for light and others for high arc temperature and steady arc operation. A variety of other chemicals may be introduced to produce radiation outside the visible spectrum for specific commercial or therapeutic purposes.

The flame arc is available in short-burning, medium-electrode-life type and long-burning type. With enclosure and automatic adjusting mechanism, and with consideration of the proper chemical composition, the electrode life of the long-burning type may be more than 100 h.

Metallic electrode arc lamp. In this type of lamp, light is produced by luminescent vapor introduced into the arc by conduction from the cathode. The positive electrode is solid copper, and the negative electrode is formed of magnetic iron oxide with titanium as the light-producing element and other chemicals to control steadiness and vaporization. These lamps are limited to direct-current use. A regulating mechanism adjusts the arc to a constant length.

Xenon short-arc lamp. This special type of lamp operates with the arc enclosed in a fused quartz bulb filled with xenon gas at a pressure of several atmospheres. The light emitted is continuous throughout the visible portion of the spectrum and approximates daylight in color. Lamps are available in sizes ranging from 30 to 30,000 watts electrical input with efficacies as high as 50 lumens per watt. The light is primarily emitted by the gas arc rather than from the hot electrodes. Large-wattage lamps require liquid cooling of the electrodes. Special power sources are required to operate these lamps. These lamps are used where it is important to have a source of very bright white (6000 K or 10,300°F color temperature) light radiated from a very small (0.4 in. or 10 mm or less) light source. Because the bulbs have internal operating gas pressures above 10 atm (150 lb/in.^{2} or 1.0 megapascal), the lamps must be in some form of protective enclosure. Protection is also needed from the large amount of ultraviolet radiation these lamps produce.

G. R. Peirce

Bibliography. D. G. Fink and M. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; J. E. Kaufman and J. F. Christensen (eds.), *IES Lighting Handbook: Applications*, rev. ed., 1987.

Arc welding

A welding process utilizing the concentrated heat of an electric arc to join metal by fusion of the parent metal and the addition of metal to the joint usually provided by a consumable electrode. This discussion of the arc-welding process is divided into five general subjects: kinds of welding current, methods of welding, arc-welding processes, types of arc welders, and classification of electrodes. **Figure 1** presents a fundamental arc-welding circuit, and **Fig. 2** shows the elements of the weld at the arc. For the metallurgical aspects of welding *See* WELDING AND CUTTING OF MATERIALS.

Welding current. Electric current for the welding arc may be either direct or alternating, depending upon the material to be welded and the characteristics of the electrode used. The current source may

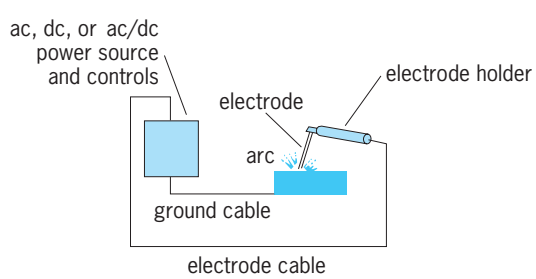

Fig. 1. Fundamental connections of electric equipment and workpiece for arc welding.

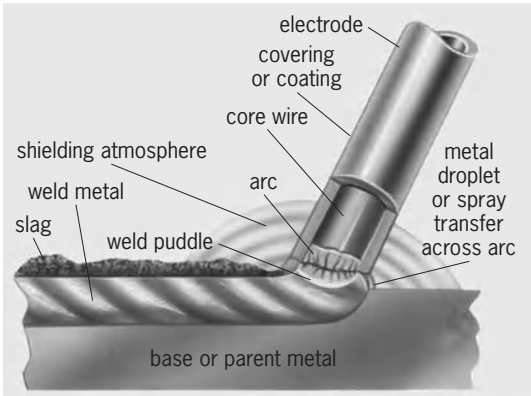

Fig. 2. Metallic welding arc.

be a rotating generator, rectifier, or transformer and must have transient and static volt-ampere characteristics designed for arc stability and weld performance.

On direct current the distribution of heat in the arc generally produces a concentration of heat at the positive terminal. The work generally has the larger mass and requires more heat for proper fusion; therefore the work is made positive and the electrode negative. This condition is known as straight polarity. When certain covered electrodes are used, the electrode is made positive and this condition is referred to as reverse polarity. With alternating current there is no polarity because of the current reversal each half cycle.

Methods. There are three basic welding methods: manual, semiautomatic, and automatic. Each has its benefits and limitations.

Manual welding. This is the oldest method, and though its proportion of the total welding market diminishes yearly, it is still the most common. Here an operator takes an electrode, clamped in a hand-held electrode holder, and manually guides the electrode along the joint as the weld is made. Usually the electrode is consumable; as the tip is consumed, the operator manually adjusts the position of the electrode to maintain a constant arc length.

Semiautomatic welding. This method is fast becoming the most popular welding method. The electrode is usually a long length of small-diameter bare wire, usually in coil form, which the welding operator manually positions and advances along the weld joint. The consumable electrode is normally motor-driven at a preselected speed through the nozzle of a hand-held welding gun or torch.

Automatic welding. This method is very similar to semiautomatic welding, except that the electrode is automatically positioned and advanced along the prescribed weld joint. Either the work may advance below the welding head or the mechanized head may move along the weld joint.

Welding processes. There are, in addition to the three basic welding methods, many welding processes which may be common to one or more of these methods. A few of the more common are described below.

Carbon-electrode arc welding. This process has been superseded, to a great extent, by other welding processes, but is still in limited use for welding ferrous and nonferrous metals. Normally, the arc is held between the carbon electrode and the work. The carbon arc serves as a source of intense heat and simply fuses the base materials together, or filler metal may be added from a separate source. A modification in which the arc is formed between two carbon electrodes is known as twin carbon arc welding, where the work does not form a part of the electrical circuit and only the heat of the arc is played against the work. The carbon electrodes are slowly consumed, and adjustments must be made while welding to maintain a constant arc length.

Shielded metal arc welding. This is the most widely used arc-welding process. A coated stick electrode, usually 14 in. (35 cm) long, is consumed during the welding operation, and therefore provides its own filler metal. The electrode coating burns in the intense heat of the arc and forms a blanket of gas and slag that completely shields the arc and weld puddle from the atmosphere. Its use is generally confined to the manual welding method.

Submerged-melt arc welding. In this process a consumable bare metal wire is used as the electrode, and a granular fusible flux over the work completely submerges the arc. This process is particularly adapted to welding heavy work in the flat position. High-quality welds are produced at greater speed with this method because as much as five times greater current density is used. Automatic or semiautomatic wire feed and control equipment is normally used for this process (**Fig. 3**). A modification of this process uses twin arcs (lead and trail) supplied from separate power sources but phased to reduce magnetic interaction of the arc stream.

Tungsten–inert gas welding. This process, often referred to as TIG welding, utilizes a virtually nonconsumable electrode made of tungsten (**Fig. 4**). Impurities, such as thorium, are often purposely added to the tungsten electrode to improve its emissivity for direct-current welding. The necessary arc shielding is provided by a continuous stream of chemically inert

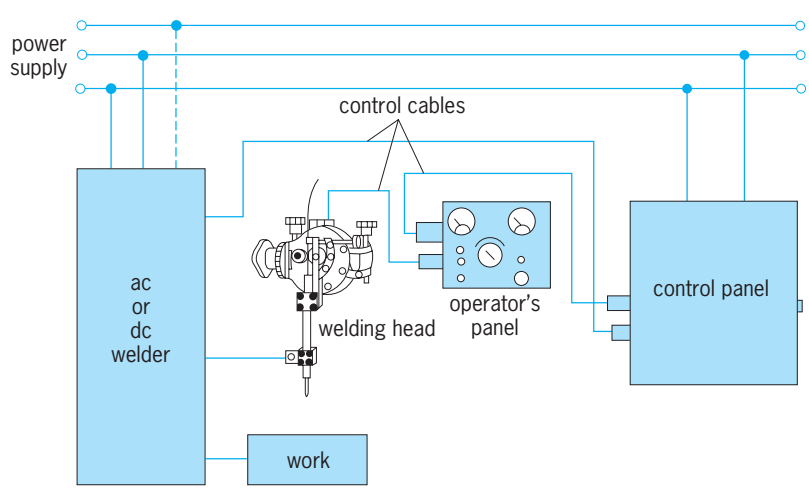

Fig. 3. Schematic diagram of automatic wire feed and control for arc welding.

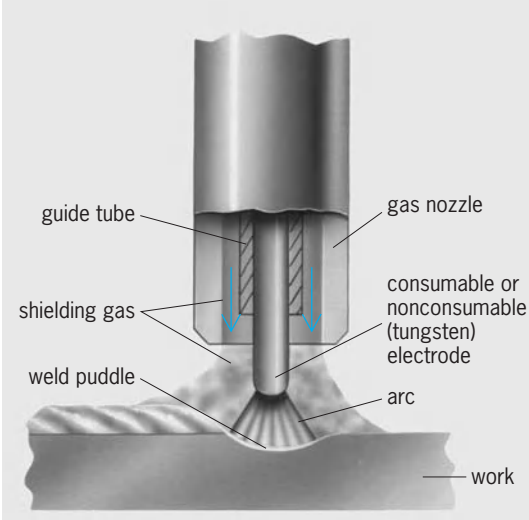

Fig. 4. Inert-gas-shielded metal welding arc.

gas, such as argon, helium, or argon-helium mixtures, which flows axially along the tungsten electrode that is mounted in a special welding torch. This process is used most often when welding aluminum and some of the more exotic space-age materials. When filler metal is desired, a separate filler rod is fed into the arc stream either manually or mechanically. Since no flux is required, the weld joint is clean and free of voids.

Metal–inert gas welding. This welding process, often referred to as MIG welding, saw its greatest growth in the 1960s. It is similar to the TIG welding process, except that a consumable metal electrode, usually wire in spool form, replaces the nonconsumable tungsten electrode. This process is adaptable to either the semiautomatic or the automatic method. In addition to the inert gases, carbon dioxide has become increasingly common as a shielding means. In MIG welding, the bare, but sometimes coated, wire is fed into the arc at current densities as high as 10 times those used for conventional shielded metal arc welding. Automatic wire feed and controls are used to drive the wire at speeds as high as 1000 in./min (2500 cm/min). This extremely versatile process produces high-quality welds at high speeds and with little or no spatter, thus eliminating cleaning.

Space does not allow describing all of the other welding processes, but a few of those most commonly referred to are electroslag welding, flux-cored welding, vapor-shielding welding, and plasma-arc welding.

Arc welders. An arc welder is an electric power generator, or conversion unit, for supplying electric current and voltage of the proper characteristics to a welding arc. Welders may be classified by the type of current supplied to the arc, alternating or direct. Most arc welders are designed for single operator use. Multiple-operator dc arc welding has limited application and is used primarily where operators are concentrated in a welding area.

Alternating-current welders. These are generally of the static type employing a two-winding transformer and a variable inductive reactor, which is connected in series with the welding circuit. The reactor may be combined with the transformer to provide adjustable leakage reactance between windings (**Fig. 5**). The transformer isolates the welding circuit from the primary line, and steps down the line voltage to an open-circuit voltage of 80 V or less for manual welding and 100 V or less for automatic welding.

The ac arc is extinguished each half cycle as the current wave passes through zero. The reactor, by reason of its stored inductive energy, provides the voltage required for reignition of the arc at each current reversal.

Several types of reactor construction are employed, including a movable core with variable air gap, a movable coil to vary leakage between windings, and a saturable reactor with dc control in combination with a tap changer.

Direct-current welders. These may be of the rotating dc generator or static rectifier type, with either constant-current or constant-voltage characteristics.

The dc generator type is usually driven by a directly coupled induction motor or internal combustion engine. The conventional design employs a generator with a combination shunt-field rheostat control and a differentially compounded series field tapped to give an adjustable welding current with drooping volt-ampere characteristic.

Other schemes of current adjustment are used, most of which employ the principle of field control. These include excitation from a third brush utilizing the effect of armature reaction or a movable core to vary the leakage between poles of the generator. Because of the extreme and rapid fluctuation of the arc voltage during welding, the welding generator must be designed for rapid response to load changes.

The rectifier-type arc welder is a combination of transformer, variable reactor, and rectifier. The transformer and reactor designs are similar to those employed in ac welders, except that dc welders are generally used on a three-phase supply (**Fig. 6**). Some single-phase welders of this type are used, particularly where a dual ac/dc output is required. The rectifiers are the selenium dry plate or silicon diode full-wave bridge type.

Both the rotating generator and rectifier types are also available with constant-voltage characteristics. On these units the output voltage is adjustable, and the welders are designed for either flat or adjustable droop volt-ampere characteristics. In addition, a direct-current inductor is often included in

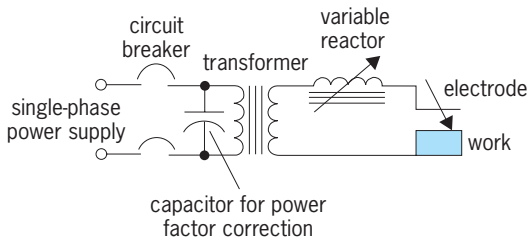

Fig. 5. Schematic diagram of a typical alternating-current arc welder.

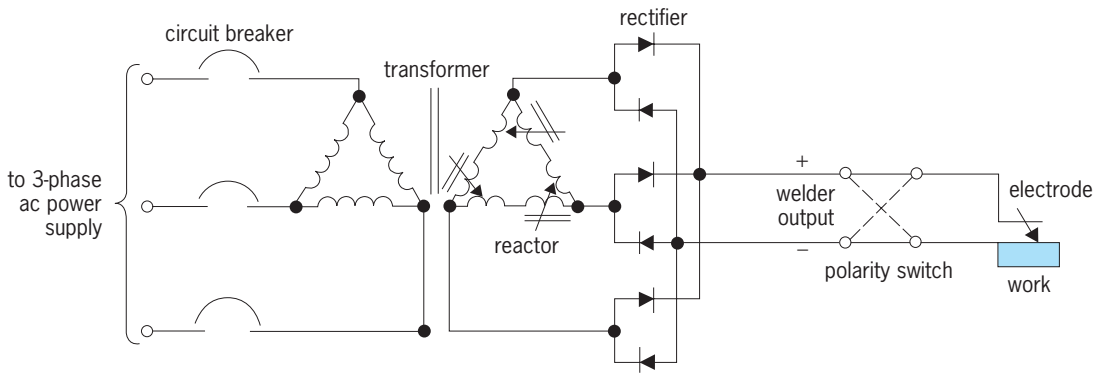

Fig. 6. Schematic diagram of a typical direct-current arc welder.

the welder output to tailor the welder's response to the process requirements. The constant-voltage type of welder is most often used with the MIG welding process where the welding current is adjusted by varying the speed of the electrode wire feed. An addition to the continually expanding line of constant-potential type of welders is a direct-current arc welder with a pulsing output superimposed on a relatively smooth background characteristic. This welder provides the power source requirements imposed by a new MIG process, the applications of which have not yet been fully explored.

Multiple-operator dc welding. This method employs one or more power units with approximately 80 load volts and relatively flat voltage regulation. The units operate in parallel to supply a number of resistor outlets, one for each operator. The resistor units provide ballast impedance and a means for adjusting the welding current. This system effects appartus economy by utilizing the load diversity of multiple welding arcs.

Electrodes. An arc-welding electrode is a wire or rod of current conducting material, with or without a covering or coating, forming one terminal of the electric arc used in arc welding. Electrodes may be classified as nonconsumable and consumable. Nonconsumable electrodes (carbon, tungsten) are consumed at a very low rate and add no filler metal to the weld. Consumable electrodes melt with the heat of the arc, and the metal is transferred across the arc to form part of the weld metal.

Consumable electrodes may be in the form of rods 9–18 in. (22.5–45 cm) long or continuous wire. Practically all welding electrodes for manual welding are coated or covered. This covering (usually extruded on the core wire) serves a number of important functions. The gas and slag shielding produced by the covering protects the molten metal from the atmosphere during the transfer across the arc and after deposit.

The materials in the electrode covering are intended to perform one or more of the following: (1) purify the weld metal; (2) control melting rate, penetration, bead shape, and cooling rate; (3) add alloying elements to the weld metal; (4) stabilize the arc; and (5) provide insulating coating to reduce shock hazard.

The covering of some electrodes includes powdered iron to increase the speed of welding by stepping up the rate of metal deposit.

Welding electrodes have been classified according to tensile strength of the deposited weld metal, type of covering, welding position, and type of current. The electrodes are identifiable by electrode classification numbers under specifications prepared jointly by the American Welding Society and the American Society for Testing and Materials. *See* RESISTANCE WELDING.

Emil F. Steinert

Bibliography. American Welding Society, *Arc Welding Safety and Health*, 1982; R. Bakish and S. S. White, *Handbook of Electron Beam Welding*, 1964; A. L. Phillips (ed.), *Welding Handbook*, 4th ed., 1957.

Arcellinida

An order of Lobosia. The shell (test) of these protozoa has a single, well-defined aperture through which slender fingerlike pseudopodia (lobopodia) can be extended. The test often has an internal chitinous layer known as a glycocalyx and an external layer composed of secreted siliceous elements, organic hollow platelets, or collected sand grains and other matter cemented together by an adhesive secretion. The secreted platelets are synthesized within membranous vesicles in the cytoplasm, where they first appear as dense granules. These granules are transported to the surface of the cell, where they are deposited on the outer surface of the cell membrane and mature into the alveolate, closely assembled platelets.

Most arcellinidans are uninucleate, but *Arcella* is binucleate. During asexual reproduction certain species prefabricate their tests before fission to some extent. During growth, the organism may assemble materials by secretion or ingestion in preparation for new test construction. At fission, when a daughter cell is produced, such materials pass through the aperture into the extruded protoplasm, where they are moved to the surface and cemented to the outer wall of the new test. The old test is kept by the sister organism. Cyst walls may be formed inside the test; the aperture is often plugged with debris

before encystment. Food consists of bacteria and smaller protistans. The order includes *Arcella, Centropyxis, Cochliopodium, Difflugia*, and many other genera. *See* LOBOSIA; PROTOZOA; SARCODINA; SARCOMASTIGOPHORA.

O. Roger Anderson

Arch

A structure, usually curved, that when subjected to vertical loads causes its two end supports to develop reactions with inwardly directed horizontal components. The commonest uses for an arch are as a bridge, supporting a roadway, railroad track, or footpath, and as part of a building, where it provides a large open space unobstructed by columns. Arches are usually built of steel, reinforced concrete, or timber.

The designations of the various parts of an arch are given in **Fig. 1**. The skewback is the abutment of pier surface upon which the arch rests. Because the arch springs from its skewback, the intersection of the arch and the skewback is called the springing line. The upper surface of the arch is the extrados, the inner surface the intrados. The line passing through the center of gravity of each section is the arch axis. The crown is the highest section.

The advantage of the arch over a simple truss or beam of the same span is the reduction in the positive moment by the negative moment resulting from the horizontal thrust at the supports. The arch rib is subjected to large axial forces and small bending moments. To minimize the moments, the center line of the rib should closely approximate the funicular polygon for dead load, plus perhaps some portion of the live load. For a uniformly distributed load the funicular polygon is a parabola.

The principal dimensions of the center line of the arch are span and rise. These may be dictated by conditions at the site, if the structure is a bridge, or by architectural requirements, if the arch is to form part of a building. A rise of from one-fifth to one-third of the span may prove economical.

On the basis of structural behavior, arches are classifed as fixed (hingeless), single-hinged, two-hinged, or three-hinged (**Fig. 2**). An arch is considered to be fixed when rotation is prevented at its supports. Reinforced concrete ribs are almost always fixed. For long-span steel structures, only fixed solid-rib arches

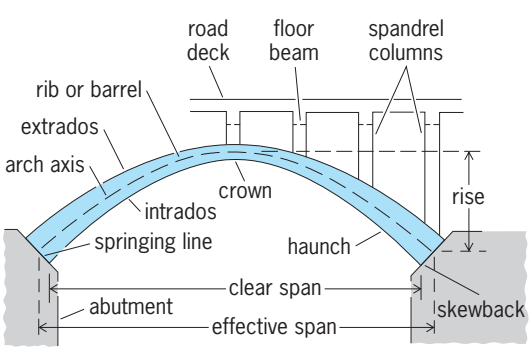

Fig. 1. An open-spandrel, concrete, fixed-arch bridge.

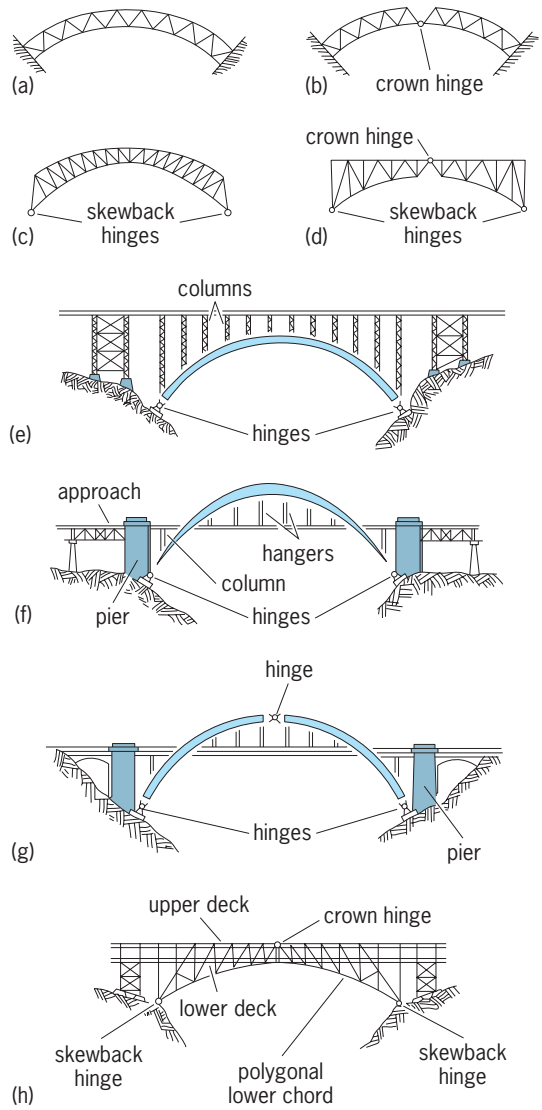

Fig. 2. Various types of bridge arches. (*a*) Fixed. (*b*) Single-hinged. (*c*) Two-hinged. (*d*) Three-hinged. (*e*) Parallel curve rib arch. (*f*) Half-through, two-hinged, crescent-rib arch. (*g*) Parallel curve, three-hinged rib arch. (*h*) Double-deck spandrel-braced arch. (*After G. A. Hool and W. S. Kinne, Movable and Long-span Steel Bridges, 2d ed., McGraw-Hill, 1943*)

are used. Because of its greater stiffness the fixed arch is better suited for long spans than hinged arches.

A hinge introduced into the arch rib produces an axis of zero moment forcing the thrust line (funicular polygon) to pass through the hinge. An arch hinged at the crown is a single-hinged arch; it is rarely used because it has no distinct advantages. The two-hinged arch has hinges at each skewback. Because foundations for two-hinged arches receive thrust forces only, the abutments are more easily planned. A hinge introduced at the crown of a two-hinged arch is statically determinate, and no stresses result from temperature, shrinkage, or rib-shortening effects. *See* STRUCTURAL ANALYSIS.

Because the horizontal component of an arch reaction is large, excellent foundation conditions must exist at the site. Abutments constructed on a steeply sloping gorge of strong, sound rock transmit the

Fig. 3. Fort Pitt highway bridge, a double-deck tied-arch bridge across Monongahela River at Pittsburgh, Pennsylvania. Unusual are a truss instead of girders as ties, and box sections as arch ribs. (*Lead Industries Ass.*)

thrust directly to the bedrock. Abutments on rock can usually be designed to resist the moment developed at the skewback of a fixed-end arch. Foundations on earth or piles may rotate and relieve part of the assumed restraint. A slight yielding of foundations usually causes no serious harm to three-hinged arches. If suitable foundation conditions are not available, the horizontal reaction component may be provided by a tie between skewbacks. An arch so constructed is classified as a tied arch (**Fig. 3**). The tied arch is usually more expensive than the two-hinged type.

Concrete arches. Concrete is relatively weak in tension and shear but strong in compression and is therefore ideal for arch construction. In Fig. 1, the arch proper consists of two or more solid ribs or a single solid barrel whose width is equal to that of the bridge. The solid-rib arch supports its roadway by means of a system of beams and spandrel columns. The rib type is particularly desirable for skew crossings because each rib may be arranged to act independently and is designed as a right arch. Ribs are usually interconnected by cross struts. Hollow box sections have been used in place of solid ribs to form a hollow-rib arch.

The roadway in the barrel-type arch is supported on an earth fill, which is confined by spandrel walls, extending from the barrel to the road deck. The barrel type, frequently called a spandrel-filled arch, is generally less economical of material than the rib type. Where for architectural reasons a solid wall effect is desired for a low, flat arch, the spandrel-filled design is often adopted. Placing solid curtain walls over the spandrel openings of an open-spandrel arch achieves the same effect.

Because the rib or barrel is subject to compression and some bending, it is designed as an eccentrically loaded column, with reinforcement placed near both the intrados and extrados. Steel percentages should be kept small.

Precast reinforced concrete arches of the three-hinged type have been used in buildings for spans up to 160 ft (49 m). *See* CONCRETE.

Steel arches. Steel arches are solid-rib or braced-rib arches. Solid-rib arches usually have two hinges but may be hingeless. A parallel curved rib (with a constant depth throughout its length) is the most commonly used form of plate girder or solid webbed rib.

The braced-rib arch has a system of diagonal bracing replacing the solid web of the solid-rib arch. The world's longest arch spans are both two-hinged arches of the braced-rib type, the Sidney Harbor Bridge in Australia and the Bayonne Bridge at Bayonne, New Jersey (**Fig. 4**), which have spans of 1650 and 1652 ft (502.9 and 503.5 m), respectively. Both are of the half-through type. Other classifications according to the method by which the roadway is supported by the arch proper are through arches and deck arches. Through and half-through arches are usually of the rib type.

The spandrel-braced arch is essentially a deck truss with a curved lower chord, the truss being capable of developing horizontal thrust at each support. This type of arch is generally constructed with two or three hinges because of the difficulty of adequately anchoring the skewbacks.

Wood arches. Wood arches may be of the solid-rib or braced-rib type. Solid-rib arches are of laminated construction and can be shaped to almost any

Fig. 4. Bayonne Bridge across the Kill Van Kull in New Jersey. (*Port of New York Authority*)

required form. Arches are usually built up of nominal 1- or 2-in. (2.5- or 5-cm) material because bending of individual laminations is more readily accomplished. For wide members a lamination may consist of two or more pieces placed side by side to form the required width. All end-to-end joints are scarfed and staggered. Although the cross section is usually rectangular, I-beam sections have been used. Because of ease in fabrication and erection, most solid-rib arches are of the three-hinged type. This type has been used for spans of more than 200 ft (60 m).

The lamella arch has been widely used to provide wide clear spans for gymnasiums and auditoriums. The wood lamella arch is more widely used than its counterpart in steel. The steel lamella roof for the civic auditorium in Corpus Christi, Texas, has a clear span of 224 ft (68 m). The characteristic diamond pattern of lamella construction provides a unique and pleasing appearance. Basically, lamella construction consists of a series of intersecting skewed arches made up of relatively short straight members. Two members are bolted, riveted, or welded to a third piece at its center. These structures are erected from scaffolds because no supporting action develops until a large section is in place. Construction starts at both sills, moving up until the sides meet in the center.

The horizontal thrust of the lamella roof must be resisted by steel tie rods or wood ties at wall height or by buttresses. The thrust component developing in the longitudinal direction of the building may be resisted by ties extending the length of the building or by the roof decking. *See* BRIDGE; ROOF CONSTRUCTION; TRUSS. Charles N. Gaylord

Masonry skeletons. Masonry skeletons in the form of arches, vaults, and domes have supported bridges and buildings since ancient times. An arch, dated 1400 B.C., was discovered among Babylonian ruins. It was built of brick with clay mortar joints. Although other civilizations understood the arch concept, the Romans significantly advanced the art of masonry skeleton construction using stone and brick. Extant Roman masonry skeletons include the Pont du Gard near Nimes, France, built about 18 B.C.; the Ponte di Traiano, Alcantara, Spain, built about A.D. 105; and the Hagia Sophia in Istanbul, Turkey, built about A.D. 535. *See* BRICK; CLAY; MASONRY; STONE AND STONE PRODUCTS.

The Pont du Gard comprises three tiers of stone arches that support two roadways and an aqueduct at a height of 50 m (160 ft) above the Gard River. The longest clear arch span of the lower tiers is 25 m (81 ft), and the spans of the top story arches are 5 m (16 ft). The cut stones were laid dry, except for those forming the water channel, which were set in mortar.

The Ponte di Traiano comprises six successive stone arches of up to 30 m (100 ft) span supporting the roadway at about 62 m (200 ft) above the Tagus River. One arch was destroyed on two occasions by military forces but has been restored.

The Hagia Sophia is an assemblage of partial domes surrounding the main dome of 31 m (102 ft) clear diameter. The domes are mostly brick covered by stucco on the exterior and mosaic on the interior. The Hagia Sophia was damaged during a severe earthquake in A.D. 558 but was repaired by 563.

As the virgin forests of Europe were exhausted, masonry construction gained importance for reasons of economy, architecture, and permanence. Canal, and later railroad network, construction during the eighteenth and nineteenth centuries necessitated many bridges. For example, large-scale masonry arch construction was started in the United Kingdom about 1750 and continued through the railway construction period, ending about 1910. In the United Kingdom there are some 40,000 highway and 33,000 railway masonry arch spans, most of brick. Many have

been used continuously and subjected to much heavier loading than when originally put into service, and have received only minimal maintenance. Large masonry arch bridges were also built on the Continent. Examples are the stone arch bridge completed in 1903 at Plauen, Germany, and the stone arch viaduct for the Rhaetian Railway near Filisur, Switzerland. The stone arch bridge has a clear span of 91 m (295 ft) with a low rise of one-fifth the span, and is built of local slate. The railway viaduct comprises segmental arches on very tall piers and was completed in 1904.

In the eastern and midwestern United States, virgin timber was not exhausted until after inexpensive steel and reinforced concrete became available; hence the need for labor-intensive masonry arch bridges did not develop except for early railroad viaducts. Two examples of stone arch railroad viaducts are the Carrollton Viaduct at Baltimore, Maryland, and the Starrucca Viaduct at Lanesboro, Pennsylvnia. The Carrollton Viaduct was built in 1829 of dressed granite with a 25-m (80-ft) main span and a smaller side span. The Starrucca Viaduct, completed in 1848, has 17 segmental arch spans of 16 m (51 ft) on slender piers about 34 m (110 ft) high. Only the outer voussoirs were dressed. *See* TRESTLE.

The masonry arch can provide structure and beauty, is fireproof, requires comparatively little maintenance, and has a high tolerance for foundation settlement and movement due to other environmental factors. Most arches are curved, but many hectares (acres) of floor in highrise office and public buildings are supported by hollow-tile jack (flat) arches (**Fig. 5**). The arch shape selected for a particular structure depends on architectural considerations, but must be such that the statical requirements for the anticipated loading are satisfied.

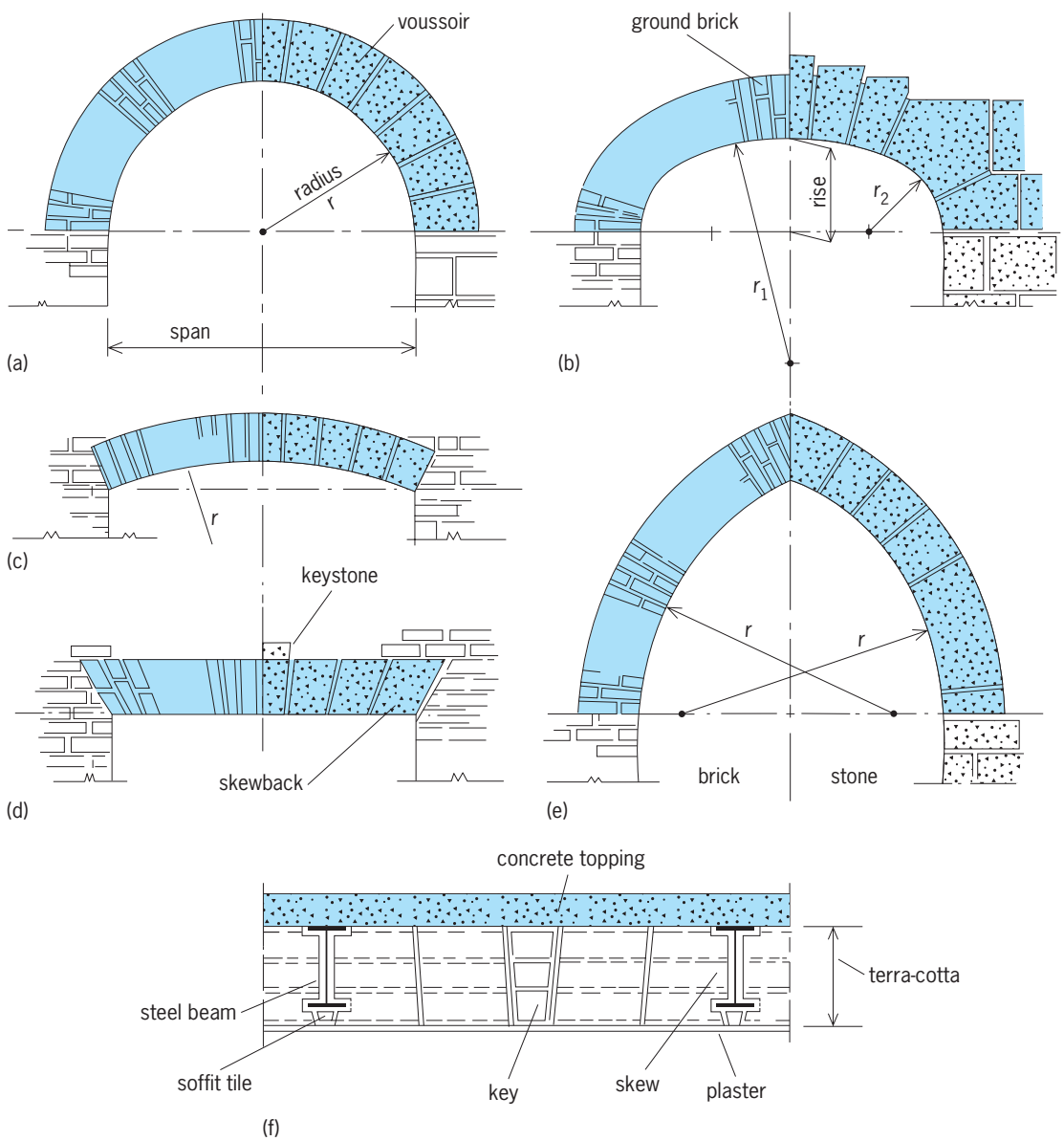

Fig. 5. Common forms of masonry arches. (*a*) Roman (semicircular). (*b*) Elliptical. (*c*) Segmental. (*d*) Jack or flat. (*e*) Gothic. (*f*) Hollow-tile jack. (*After C. G. Ramsay and H. R. Sleeper, Architectural Graphic Standards, 9th ed., John Wiley, 1994*)

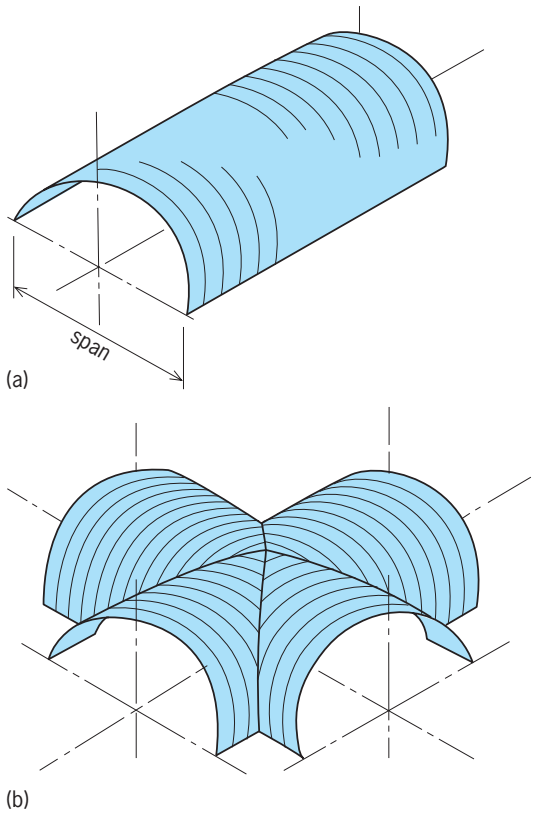

Fig. 6. Types of vaults. (*a*) Individual or barrel. (*b*) Quadripartite. (*After J. Heyman, The Stone Skeleton, Cambridge University Press, 1995*)

If a curved arch is wide (dimension normal to span), the arch is referred to as a barrel arch or vault. The vault cross section may have several different shapes (Fig. 5). Contiguous vaults may be individual, may intersect, or may cross (**Fig. 6**). A four-part vault is termed quadripartite. Contiguous quadripartite vaults that are supported at the corners by columns are masonry skeletons of large cathedrals. For example, the cathedral of Notre-Dame in Paris, France, was built as a sexpartite vault about A.D. 1200.

Stone for masonry skeletons is cut from three classes of rock; igneous (granite, traprock), metamorphic (gneiss, slate, quartzite), and sedimentary (limestone, sandstone). The physical properties vary widely within each class, and the properties of the stone actually used should be determined by test. The minimum compressive strength can range from 3000 pounds per square inch (21 megapascals) for sandstone to 15,000 psi (103 MPa) for granite. The corresponding upper limits are 20,000 psi (138 MPa) and 30,000 psi (207 MPa). *See* IGNEOUS ROCKS; METAMORPHIC ROCKS; SEDIMENTARY ROCKS.

The primary requirements for brick as a structural material are compressive strength and weathering resistance. Building brick (common brick) is now graded as Severe-, Moderate-, and No-Weathering (SW, MW, and NW) by the American Society for Testing and Materials (ASTM). The minimum acceptable compressive strength for a single brick loaded over the full bedding area varies from 2500 psi (17 MPa) for SW down to 1250 psi (9 MPa) for NW. An index of weathering durability is water absorption. The maximum acceptable absorption of brick submerged 24 h in cold water is 16% by weight for SW and unlimited for NW.

Hollow clay tiles (terra-cotta) are made by extruding a clay mixture through a die, cutting pieces from the extrusion, and then heating them to a high temperature (burning). Terra-cotta for floor arches was made semiporous in order to improve fire resistance. When the arches were loaded in the direction of the cells, the compressive strength varied from 2000 to 12,000 psi (14 to 83 MPa). Water absorption was not considered an issue for floor arches.

Masonry skeletons depend on a general state of compression for stability. Old brick or stone masonry cannot reliably transmit tensile stress, even when laid with mortar joints. With time the tensile bond between the mortar and masonry is lost because of environmental factors. Steel reinforcment can be placed in the joints between units to improve the tensile capacity of the masonry. However, the long-term viability of such construction has yet to be determined. The compressive stresses in the skeleton are usually low compared to the crushing strength of the masonry, but are sufficient to develop friction and prevent slip of adjacent units.

Until the middle of the eighteenth century the geometrical proportions of masonry skeletons were based on complex rules (closely guarded by their possessors), applied by an architect-engineer (master builder) who had served a long apprenticeship. These rules usually lead to skeletons which are in a compressive state due to gravity loads and wind and which have survived for centuries. Analytical and experimental work in the latter part of the eighteenth century led to a limit-state approach to design, but most construction still followed the old rules.

At the beginning of the nineteenth century, C. L. M. Navier advocated elastic structural analysis which offered engineers a design tool based on the working-stress philosophy. The funicular polygon of the actions on the skeleton could now be computed from statics for any arch, assuming elastic behavior and geometrical boundary conditions at the springings. This force polygon (pressure line, thrust line) must lie within the kern of the arch cross sections throughout, for all combinations of actions, if the voussoirs are not to separate at the outer or inner curves of the arch (extrados or intrados). For a solid rectangular arch cross section, the eccentricity of the pressure line at the kern limit is $e = d/6$ (**Fig. 7***a*). Because the joints between the voussoirs will not separate until $e = d/2$ (assuming no local crushing at the joints), the arch is said to have a geometrical factor of safety (GFS) = 3 for $e = d/6$. Masonry arches are usually proportioned to have GFS > 2.

Arches are often designed as if fixed at the springings. In reality, no arch is fully fixed-ended, and a slight change in the geometrical boundary conditions influences the location of the pressure line. The abutments of masonry arches displace an unknown amount in time, and the masonry is subjected

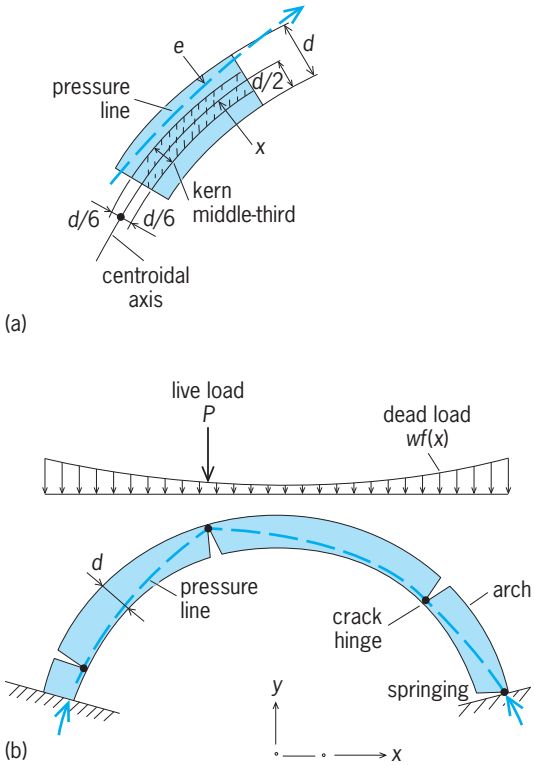

Fig. 7. Pressure line of an arch. (*a*) Notation for arches. (*b*) Arch collapse mechanism for a masonry arch subjected to a point live load.

to volume and material property changes due to the environment. Because of these effects, the validity of the conventional elastic analysis has been questioned, especially regarding long-term behavior, with the return to the limit state approach, especially for assessment.

Pipard and associates developed a limit analysis (about 1935–1962) for curved arches using a mechanism approach and assuming that the masonry has (1) no tensile strength, (2) infinite compressive strength, (3) infinite elastic modulus, and (4) no slip between adjacent voussoirs. J. Heyman related this limit-state approach to the plasticity theorems.

Analyses have been developed which are intended to predict curved masonry arch behavior prior to collapse. In 1990 a method was proposed for "fixed arches" based on the development of three crack hinges until the collapse load is reached when the fourth crack hinge develops. Procedures are available for analyzing inelastic masonry arches utilizing nonlinear finite element techniques.

Charles Birnstiel

Bibliography. J. Henrych, *The Dynamics of Arches and Frames*, 1981; J. Heyman, *The Stone Skeleton: Structural Engineering of Masonry Architecture*, Cambridge University Press, Cambridge, U.K., 1995; F. Leonhardt, *Brucken-Bridges*, MIT Press, Cambridge, MA, 1984; C. G. Ramsey and H. R. Sleeper, *Architectural Graphic Standards*, John Wiley, New York, 1994; W. Schueller, *Horizontal-Span Building Structures*, John Wiley, New York, 1983.

Archaea

A group of diverse microscopic prokaryotic organisms that constitute one of the three domains of life. Comparison of the sequences of small subunit ribosomal ribonucleic acid (rRNA) from a large variety of living organisms indicates that they may be divided into three major groups: Archaea, Bacteria, and Eukaryota (**Fig. 1**). *See* BACTERIA; EUKARYOTA.

Archaea are often found in some of the most extreme environments on this planet; specifically, anaerobic, hypersaline, extremely cold, or extremely high temperature environments. They comprise about one-third of the prokaryotic biomass in coastal Antarctic water, providing food for many marine animals. A few are symbionts in the digestive system of animals; however, none are known to cause disease in humans or animals.

General Characteristics

The Archaea are quite distinct from bacteria and eukaryotes, and within the domain there is great diversity.

Morphology. Archaea may be spherical, rod-shaped, spiral, plate-shaped, or pleomorphic (variable in shape) [**Fig. 2**]. Some are single cells, whereas others form filaments. They range in diameter from 0.1 micrometer to more than 15 μm, and some filaments grow longer than 200 μm. Although there is considerable variation in archaeal cell wall structure, all members lack peptidoglycan in their cell walls. They stain either gram-positive or gram-negative. The gram-positive archaea have a cell wall with a single thick homogeneous layer resembling that in gram-positive bacteria and thus stain gram-positive. The gram-negative archaea lack the outer membrane and peptidoglycan network and thus stain gram-negative.

Membrane biochemistry. Archaeal membrane lipids are branched-chain hydrocarbons attached to glycerol by ether links, rather than fatty acids connected by ester links, as in bacteria and eukaryotes. The ether-linked lipids confer greater structural rigidity and chemical stability, enabling these organisms to live in a more extreme environment. The membranes of extreme thermophiles (organisms that thrive at high temperatures) consist almost completely of monolayers of tetraethers, which gives these archaea thermal stability.

Physiology. Many archaea are extremophils, requiring extreme conditions of temperature, pH, and salinity to survive. Some are mesophiles, and grow optimally at intermediate temperatures ranging from 68 to 113°F (20 to 45°C); others are hyperthermophiles that can live and grow in environments above 212°F (100°C). They can be aerobic (require oxygen for growth), facultatively anaerobic (grow equally well with or without oxygen), or strict anaerobes (grow in the absence of oxygen). Some archaea obtain energy from inorganic compounds (chemolithoautotrophs), whereas others derive their nutrients from organic substances.

Reproduction is via budding, fragmentation, binary fission, or unknown mechanisms.

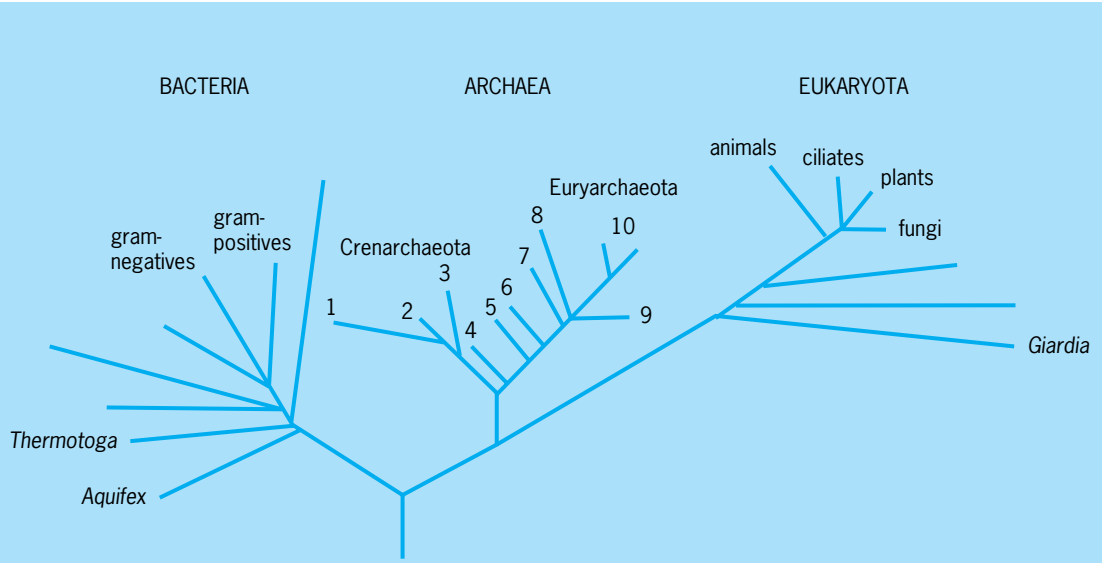

Fig. 1. Phylogenetic relationships among the three domains of living organisms: Bacteria, Archaea, and Eukaryota. The phylogenetic positions of representative archaea are indicated by numbers: (1) *Sulfolobus;* (2) *Pyrodictium;* (3) *Thermoproteus;* (4) *Methanopyrus;* (5) *Thermococcus;* (6) *Methanococcus;* (7) *Methanobacterium;* (8) *Methanosarcina;* (9) *Archaeoglobus;* (10) halophiles. (*After O. Kandler, Where next with the archaebacteria?, Biochem. Soc. Symp., 58:195–207, 1992*)

Metabolism. Although many aspects of archaeal metabolism appear similar to other organisms, the archaea differ with respect to glucose catabolism, pathways for CO_2 fixation, and the ability of some to synthesize methane. Archaea use a number of unique coenzymes in their metabolism, some of which are associated with methanogenesis (the production of methane: other coenzymes are unknown in the bacteria and eukaryotes). The enzymes involved in archaeal protein and cell wall synthesis are sufficiently different from those used by bacteria and are not inhibited by antibiotics such as penicillin, tetracycline, and streptomycin, which are effective against most bacteria.

A few archaea ferment sugars via an unusual metabolic pathway. Most bacteria and eukaryotes catabolize glucose by the Embden-Meyerhof-Parnas (EMP) pathway, which involves phosphorylated intermediates and yields energy-rich end products. Archaea use a nonphosphorylated pathway termed the pyroglycolytic pathway, which generates less energy for each sugar molecule catabolized than the EMP pathway. However, most of the EMP enzymes are present in archaea, since they synthesize sugars by reversing the reactions of the EMP pathway. It is unknown why glycolysis evolved differently in bacteria and archaea.

Very little is known about biosynthetic pathways in the archaea. Preliminary data suggest the synthetic pathways for amino acids, purines, and pyrimidines are similar to those in other organisms. *See* BACTERIAL PHYSIOLOGY AND METABOLISM.

Genetics and molecular biology. Although archaea and bacteria share some genetic similarities, there are important differences. For example, the initial amino acid in their polypeptide chains, coded by the AUG start codon, is methionine (as it is in eukaryotes), whereas bacteria use *N*-formyl methionine. Archaea are also distinguished from both bacteria and eukaryotes by their transfer ribonucleic acid (tRNA), ribosomes, elongation factors, and RNA polymerases. *See* GENETIC CODE; MOLECULAR BIOLOGY; RIBONUCLEIC ACID (RNA); RIBOSOMES.

Taxonomy

Based on rRNA data, the archaea are divided into two phyla: the Euryarchaeota and Crenarchaeota. The euryarchaeotes occupy many different ecological niches and have a wide variety of metabolic patterns. The crenarchaeotes are thought to resemble the ancestor of the archaea, and almost all of the well-characterized species are thermophiles or hyperthermophiles. *See* CLASSIFICATION, BIOLOGICAL.

Crenarchaeota. Most of the crenarchaeotes are extremely thermophilic, and many are acidophiles (organisms that have a preference for acidic environments) and sulfur-dependent. Almost all are strict anaerobes. The sulfur may be used as an electron acceptor in anaerobic respiration or as an electron source by lithotrophs (organisms that use inorganic substrates as electron donors in energy metabolism). They grow in geothermally heated waters or soils that contain elemental sulfur. Examples include the sulfur-rich hot springs in Yellowstone National Park (the orange color of these springs is due to the carotenoid pigments of these archaea), and the waters surrounding areas of submarine volcanic activity. Genera in Crenarchaeota include *Thermoproteus*, *Sulfolobus*, and *Pyrodictium*. *See* CAROTENOID; HYDROTHERMAL VENT; SULFUR.

Euryarchaeota. This is the largest phylum of archaea with many diverse classes, orders, and families. Euryarchaeota comprises five major groups that

differ with respect to their physiology and ecology: methanogens, halophiles, thermoplasms, extremely thermophilic S^0 metabolizers, and sulfate-reducing archaea.

Methanogens. The methanogens constitute the largest group of archaea and are responsible for the production of all of the biologically produced methane in the atmosphere (Fig. 2). They grow in the absence of oxygen (strict anaerobes) and consume carbon dioxide and hydrogen or simple organic compounds (such as acetic acid, formic acid, alcohols, and methylaminines), from which they produce methane or methane and CO_2. They are the last organism on the food chain in the conversion of organic carbon to water-insoluble gas, most of which escapes into the atmosphere. They inhabit the digestive tract of mammals (ruminants, such as cattle, deer, elk, sheep, goats, and giraffes), sewage sludge digesters, swamps (where they produce marsh gas), hot springs, and sediments in freshwater and marine environments, and even live within anaerobic protozoa. A few methanogens can grow at very high temperatures and live alongside the extreme thermophilic S^0 metabolizers in marine hydrothermal vents. It is possible that methanogenic archaea were among the earliest organisms, as they seem well adapted to living under conditions similar to those presumed to have existed on the young Earth. *See* METHANE; METHANOGENESIS (BACTERIA).

Halophiles. The halophilic archaea are characterized by their dependence on high concentrations of salt (sodium chloride) from their environment. They carry out aerobic respiration and are chemoheterotrophic, relying on complex nutrients (typically proteins and amino acids) for growth. Halophiles are found in the Great Salt Lake in Utah, the Dead Sea between Israel and Jordan, alkaline salt lakes of Africa, marine salterns (solar evaporation ponds used to concentrate salt for use in seasoning and for the production of fertilizers), and on salt-preserved fish and animal hides. Most strains require salt concentrations near the saturation point of sodium chloride (about 36%) to maintain the integrity of their cell walls.

The most studied halophile is *Halobacterium salinarium*. This species is best known for a unique type of photometabolism that is used to derive energy from sunlight when the oxygen content of the environment declines. It uses the rhodopsin pigments (bacteriorhodopsins) as photoreceptors to produce adenosine triphosphate (ATP). *Halobacterium* also uses this ATP to rotate its flagella to propel itself to

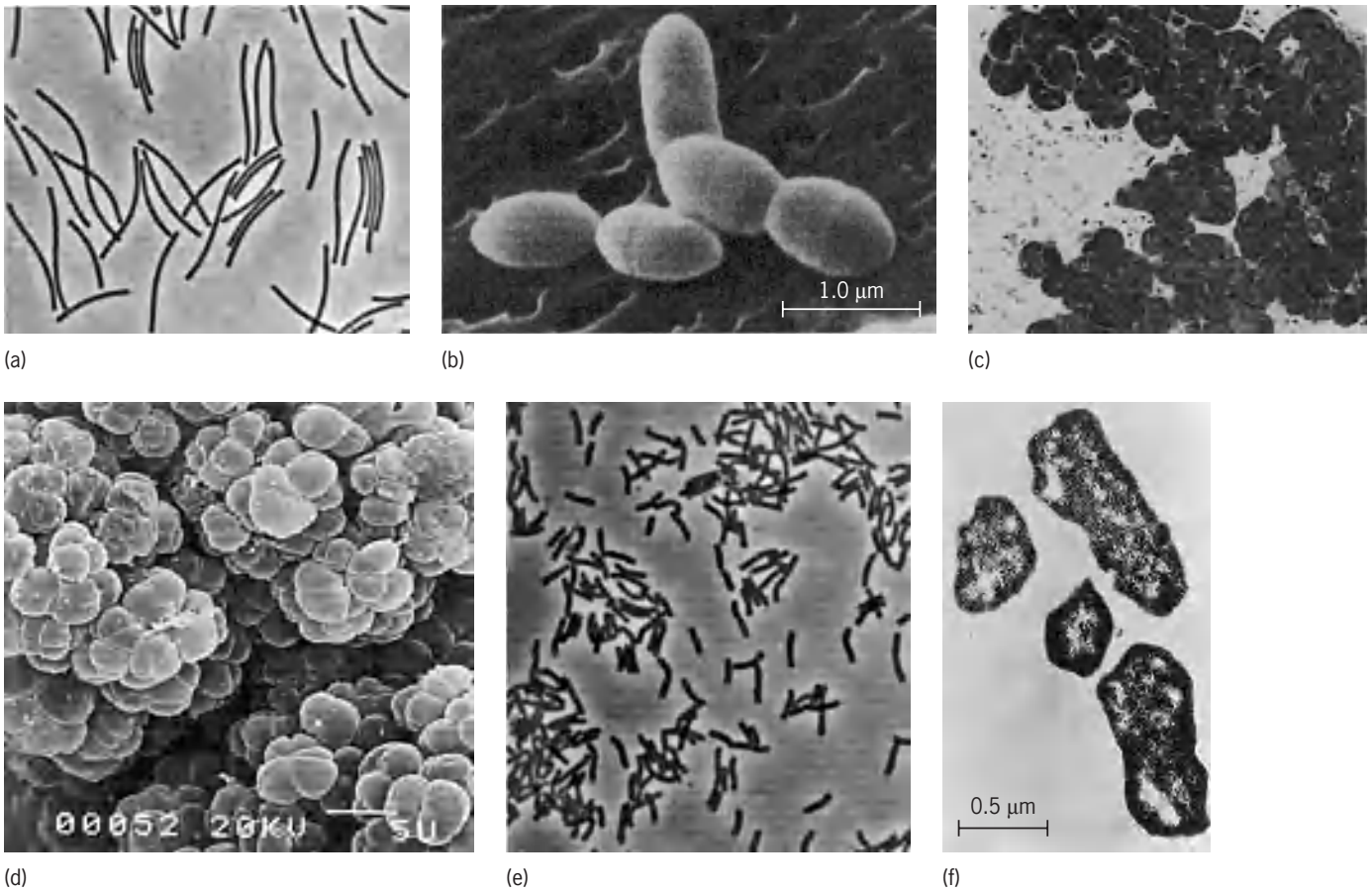

Fig. 2. Selected methanogens. (*a*) *Methanosppirillium hungatei;* phase contrast (×2000). (*b*) *Methanobrevibacter smithii.* (*c*) *Methanosarcina barkeri* from sewage digester; transmission electron microscope (×6000). (*d*) *Methanosarcina mazei;* scanning electron microscope. Scale bar = 5 μm. (*e*) *Methanobacterium bryantii;* phase contrast (×2000). (*f*) *Methanogenium marisnigri;* electron micrograph (×45,000). (*Reprinted with permission from L. Preslott, J. Harley, and D. Klein, Microbiology, 5th ed., McGraw-Hill, 2001*)

the proper water depth for maximum light absorption. *See* HALOPHILISM (MICROBIOLOGY).

Thermoplasms. The thermoplasms lack cell walls and are thermoacidophilic, requiring high temperatures and extremely acidic conditions to grow; their deoxyribonucleic acid (DNA), RNA, cytoplasmic membranes, and proteins do not function at temperatures below 113°F (45°C). These archaea grow in coal mine refuse piles that contain large amounts of iron pyrite (FeS), which is oxidized to sulfuric acid. As a result, these piles become very hot and acidic. *See* PYRITE.

Extremely thermophilic S^0 metabolizers. These archaea are strictly anaerobic and can reduce sulfur to sulfide. They are motile by flagella and have optimum growth temperatures around 190 to 212°F (88 to 100°C).

Sulfate-reducing archaea. These archaea are characterized not only by their unique ability to reduce sulfate but also by their possession of the methanogen coenzymes F_{420} and methanopterin. Archaeal sulfate reducers can extract electrons from a variety of electron donors (such as H_2, lactate, glucose) and reduce sulfate, sulfite, or thiosulfate to sulfide. (Elemental sulfur is not used as an acceptor.) Many species are found in marine hydrothermal vents. John P. Harley

Bibliography. T. Brock, Life at high temperatures, *Science*, 203:132–138, 1985; R. Charleobois, *Archaea*: Whose sister lineage?, in *Organization of the Prokaryotic Genome*, ed. by R. L. Charleobois, pp. 63–76. ASM Press, Washington, DC, 1999; C. R. Woese, Archaebacteria, *Sci. Amer.*, 244(6):98–122, 1981; C. R. Woese et al., Towards a natural system of organisms: Proposal for the domain Archaea, Bacteria, and Eukarya, *Proc. Nat. Acad. Sci. USA*, 87:4576–4579, 1990.

Archaeocyatha

An extinct group of mainly Lower Cambrian marine sponges which, although lacking spicules, possessed an intricate, highly porous skeleton of calcite. It was probably a monophyletic group; that is, all representatives were derived from a single ancestor. The position of the Archaeocyatha within the Porifera is uncertain, but they were probably most closely related to the class Demospongiae. Their fossil record is well known, as archaeocyaths represent the first large skeletal animals to have been associated with reefs; they were widespread in the shallow, warm waters that surrounded the many continents that occupied tropical latitudes during the Cambrian.

Morphology. Archaeocyath sponges displayed a great variety of sizes and growth forms. They ranged from a few millimeters to over 500 mm in length and width, and included open cups, convoluted plates, and complex branching forms (**Fig. 1**). Cup-shaped and branching forms 20–100 mm in height were, however, by far the most common. A typical cup-shaped archaeocyath skeleton is composed of a double-walled, inverted cone; the area between the outer and inner walls is known as the intervallum (**Fig. 2**). The intervallum may bear radially arranged, vertical plates known as septa, and other skeletal structures such as horizontal tabulae and dissepiments, or radial rods. The inner wall encloses a central cavity, which is open to the exterior at the top of the cup.

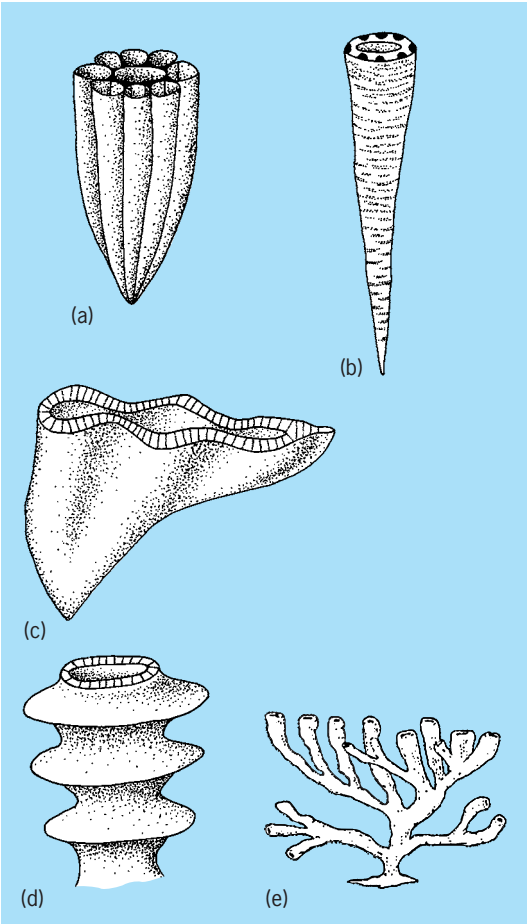

Fig. 1. Some growth forms in the Archaeocyatha: (*a*) ribbed; (*b*) tubular conical; (*c*) asymmetric, conical; (*d*) annulated; (*e*) branching.

The calcareous skeleton of archaeocyaths was densely microcrystalline, and may have been composed of high-magnesium calcite. This microstructure is remarkably similar to that found in the living calcified sponge *Vaceletia*, and so it has been inferred to have formed in a similar fashion via calcification of a series of stalked collagenous templates.

Functional morphology. The archaeocyath soft tissue was probably restricted to the intervallum, with a thin veneer of membrane around the inner and outer walls. Since archaeocyaths were suspension feeders, water almost certainly entered the animal via the numerous small pores in the outer wall, was filtered by the choanocyte chambers housed in the intervallum, and exited via the porous inner wall from the opening at the top of the central cavity. The outer skeletal wall probably acted as a sieve to exclude large, potentially clogging particles from the soft tissue.

Flume experiments have shown that the presence of septa enhances passive water flow through a model of an archaeocyath cup. Moreover, septa lacking pores appear to minimize leakage of flow from the porous outer wall. This observation might

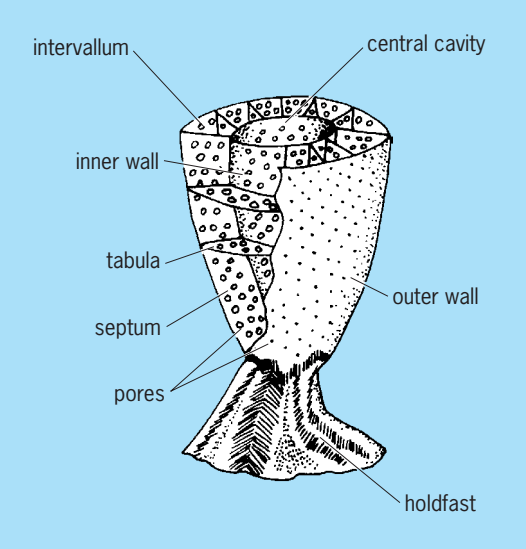

Fig. 2. Archaeocyath skeletal structure.

explain the trend found in some archaeocyaths which show a reduction in septal pores through the lifetime of an individual. As an individual increases in height, it encounters increasingly higher current velocities, necessitating greater control of water flow.

Classification. Archaeocyatha has been traditionally subdivided into two subclasses, the Regulars and Irregulars, according to differences in the early development of the skeleton. It has now been demonstrated, however, that these skeletal characters are a function of differences in soft-tissue distribution, which are independent of systematic placing. Regulars (orders Monocyathida, Ajacicyathida, Coscniocyathida, and Tabulacyathida) generally show no tabulae or dissepiments, and they are inferred to have had intervalla that were completely filled with soft tissue. By contrast, the skeletons of Irregulars (orders Archaeocyathida and Kazachstanicyathida) are believed to have borne soft tissue in their upper parts only, as they possessed abundant structures such as tabulae which served to section off abandoned areas of the skeleton as the soft tissue migrated upward.

History. Archaeocyaths appeared about 535 million years ago (Ma) on the Siberian platform with a considerable standing diversity of species, possibly indicating some preskeletal history. The group migrated from this center of origin to become globally widespread in low latitudes by about 530 Ma, and rapidly diversified to reach an acme of over 170 genera by approximately 525 Ma. However, between approximately 523 and 520 Ma, the group suffered extinction (probably due to loss of habitat) and diversity plummeted; only two species are known from later in the Cambrian.

Ecology. Archaeocyaths were a common element in Lower Cambrian reefs, where they intergrew with a varied community of calcimicrobes (possible calcified cyanobacteria). Most were attached to calcimicrobes by a holdfast or epitheca (Fig. 2), although some large solitary Regular individuals may have been rooted in lime mud. Many had buttresses which aided in stabilization and binding, as well as inhibiting the growth of competitors. Reworked archaeocyath skeletal debris was also a common component of interreef sediments.

Archaeocyaths were both upright open-surface frame builders and common crypt dwellers; indeed, up to half of the species in some communities were cryptic, living within small cavities within the reef (**Fig. 3**). Representatives from all six archaeocyathan orders were present in crypts, but in markedly different proportions. A far greater proportion of archaeocyathids, monocyathids, and coscinocyathids were represented in cryptic communities than were ajacicyathids and tabulacyathids. Ajacicyathids were a very minor component of the cryptos, even though they were the most diverse order in open surface communities.

In addition to differences in systematic distribution, cryptic and open-surface archaeocyaths display distinct morphological differences. First, all cryptic archaeocyaths have porous septa. This suggests that such forms may have been adapted hydrodynamically to low-turbulence conditions, as might be expected in crypts. Second, while cryptic individuals were often small cup-shaped and juvenile individuals, branching forms dominated open surfaces.

Some have suggested that archaeocyaths possessed photosymbionts by analogy with modern reef corals. However, the inferred absence of large areas of external soft tissue that could have been exposed to light, and the preferred cryptic niche of many representatives, does not support this hypothesis.

Evolutionary trends. Archaeocyaths show a strongly progressive trend through their history of an increasing proportion of branching forms.

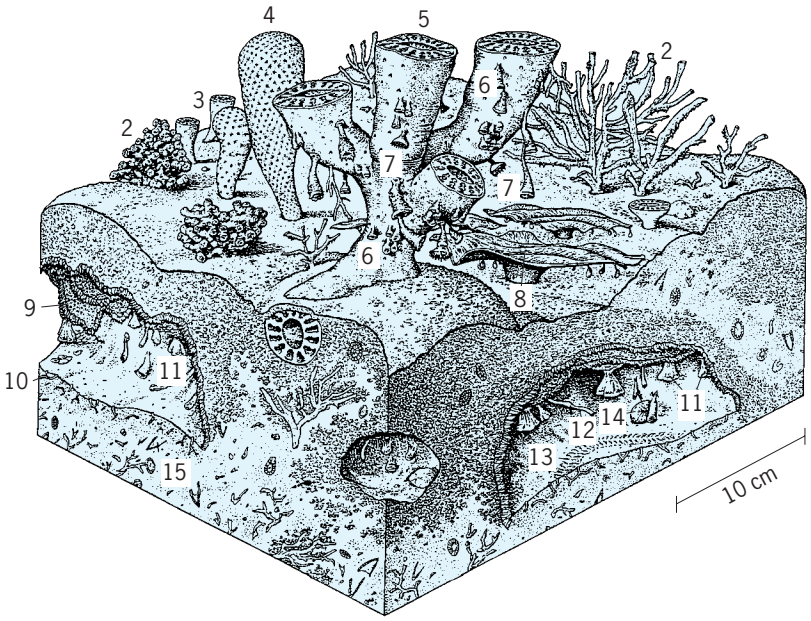

Fig. 3. Reconstruction of a Lower Cambrian archaeocyath reef community: (1) calcimicrobe; (2) branching archaeocyath sponges; (3) solitary cup-shaped archaeocyath sponges; (4) chancelloriid (problematicum); (5) radiocyaths (problematicum); (6) small archaeocyath sponges; (7) "coralomorphs"; (8) archaeocyath sponge; (9) fibrous cement; (10) microburrows (traces of a deposit feeder); (11) cryptic archaeocyaths and coralomorphs; (12) cribricyaths (problematicum); (13) trilobite trackway; (14) cement; (15) sediment with skeletal debris. (*Copyright, John Sibbick*)

Branching morphologies were derived from the more primitive solitary, cup-shaped state and were acquired independently in several lineages. Few possessed a branching habit in the early stages of archaeocyath history, but by the end of the Lower Cambrian about three-quarters of genera possessed such growth forms. This suggests that branching archaeocyaths were better competitors for space in the Lower Cambrian reef habitat. *See* PARAZOA; PORIFERA. Rachel Wood

Bibliography. F. Debrenne and A. Yu. Zhuravlev, Irregular Archaeocyaths: Morphology, ontogeny, systematics, biostratigraphy, palaeoecology, *Cahiers de Paléontologie*, 1992; M. Savarese, Functional analysis of archaeocyathan skeletal morphology and its paleobiological implications, *Paleobiology*, 18:464–480, 1992; R. Wood, A. Yu. Zhuravlev, and F. Debrenne, Functional biology and ecology of Archaeocyatha, *Palaios*, 7:131–156, 1992; A. Yu. Zhuravlev and R. A. Wood, Lower Cambrian reefal cryptic communities, *Palaeontology*, 38:443–740, 1995.

Archaeognatha

An order of ancestrally wingless insects, commonly known as bristletails, in the subclass Apterygota; about 350 species are known in the two families, Machilidae and Meinertellidae. All have slender, spindle-shaped bodies of about 0.4–0.8 in. (1–2 cm) in length, tapered tailward to a long, multisegmented median caudal filament and paired cerci. The relatively thin integument is covered with pigmented scales after the first instar. The antennae are long and filiform, without intrinsic musculature in the flagellum, as in all true insects; and large, contiguous compound eyes are accompanied by a trio of ocelli. The mandibles are exposed and attached to the cranium by a single joint (condyle), and the maxillary palpi are extended, with seven segments, while the labial palpi have only three segments. Each leg has a large coxa, two-segmented trochanter, femur, tibia, and a two- or three-segmented tarsus terminating in a two-clawed pretarsus. The posterior pair or two pairs of coxae and most of the ventral abdominal coxites bear styles, which may represent modified legs. Outwardly projecting vesicles that are apparently water absorbing also occur on most coxites. Internally, the gut is relatively simple, with small crop, digestive ceca, and a dozen or more Malpighian tubules. The nervous system is primitive, with three thoracic and eight abdominal ganglia and twin connectives. Tracheae are well developed.

In mating, the male deposits sperm on a thread or in a spermatophore that is picked up by the female, but in at least one species sperm is placed directly on the female ovipositor. Immature forms are similar to adults in body form, though the first instar may lack cerci. Sexual maturity is attained after eight or more molts.

Bristletails live in a variety of habitats worldwide, often on stones or tree trunks where they find the algae and lichens that appear to be their principal sources of food. Most species are primarily nocturnal feeders. *See* INSECTA. William L. Brown, Jr.

Bibliography. D. J. Borror, C. A. Triplehorn, and N. F. Johnson, *An Introduction to the Study of Insects*, 6th ed., 1997; Commonwealth Scientific and Industrial Research Organization Staff, *The Insects of Australia*, 1991.

Archaeopteridales

An order of extinct, free-sporing plants (cryptogams) that lived during the Late Devonian through Early Mississippian. The order includes the genera

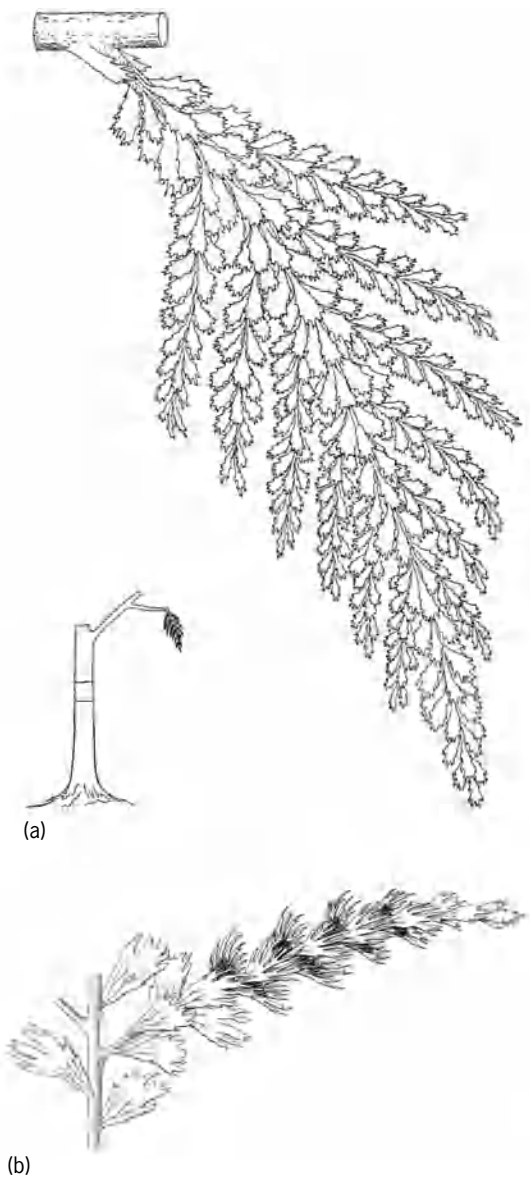

***Archaeopteris macilenta*. (*a*) Vegetative lateral branch system bearing simple leaves (*from C. B. Beck, On the anatomy and morphology of lateral branch systems of Archaeopteris, Amer. J. Bot., 58:758–784, 1971*). (*b*) Strobilus showing adaxial sporangia (*from C. B. Beck, Archaeopteris and its role in vascular plant evolution, in K. J. Niklas, ed., Paleobotany, Paleoecology, and Evolution, 1:193–230, Praeger, 1981*).**

Archaeopteris which has been extensively investigated, *Actinoxylon*, *Actinopodium*, *Svalbardia, Eddya*, and *Siderella*.

Archaeopteris had a worldwide distribution on what today are Northern Hemisphere continents, as well as on Australia. The genus is characterized by an arborescent habit and large, determinate, deciduous, frondlike lateral branch systems that bear simple, oval to fan-shaped leaves, with dichotomous venation, which vary in different species from highly dissected to nearly entire. The main axes of lateral branch systems have radially symmetrical internal anatomy and helically arranged leaves; ultimate axes are borne suboppositely to alternately in one plane (see **illus.**). Anatomy and leaf arrangement of ultimate vegetative axes reflect a nonplanate, bilateral symmetry. Fertile leaves (sporophylls) bear microsporophylls and megasporophylls on their adaxial (upper) surfaces. Radial symmetry of fertile ultimate axes (strobili) is expressed both internally and externally in the region of sporophylls.

In large stems, a cylinder of discrete vascular bundles encloses a broad pith. Leaf and branch traces diverge from the vascular bundles along radial planes. Extensive secondary wood and lesser amounts of secondary phloem are produced by a bifacial vascular cambium. The wood is compact tissue comprising tracheids and rays of variable height and width in different species. The circular-bordered pits in the radial walls of the tracheids are arranged in groups of several to 30 or more aligned in bands at right angles to the long axes of the tracheids, a characteristic also found in isolated wood and stem fragments called *Callixylon*. *See* PHLOEM.

Archaeopteridales belongs to the class Progymnospermopsida. The progymnosperms are thought by some botanists to be ancestral to seed plants. *See* PLANT KINGDOM. Charles B. Beck

Bibliography. C. B. Beck, (ed.), *Origin and Evolution of Gymnosperms*, 1988; S. V. Meyer, *Fundamentals of Paleobotany*, 1987; W. N. Stewart and G. W. Rothwell, *Paleobotany and the Evolution of Plants*, 2d ed., 1993.

Archaeopteryx

A genus of fossil birds in the extinct order Archaeopterygiformes, characterized by flight feathers like those of modern birds. *Archaeopteryx* is generally accepted as the oldest known fossil bird. It represents a unique snapshot of evolution; most of its skeletal characters are shared with small carnivorous dinosaurs, yet it has fully feathered modern wings. *Archaeopteryx* is effectively a flying dinosaur—an ancient bird about the size of a modern magpie. *See* AVES; DINOSAUR.

Fossil record. *Archaeopteryx* is known from one feather impression and seven skeletons, discovered over the course of 150 years from the Altmuhl Valley in Bavaria, Germany. All have been found in lithographic limestone of uppermost Jurassic (Tithonian) age, approximately 147 million years old, and were excavated from several locations that lie within a 15-mile radius of the town of Solnhofen. The Solnhofen Lithographic Limestone preserves exceptionally fine detail, including the impressions of feathers; the animal carcases sank into fine plastic sediment on a shallow lagoon bed and were buried in anaerobic conditions. The original specimen, the impression of a single feather preserved as part and counterpart, was discovered in 1861 and named *Archaeopteryx lithographica* by H. von Meyer. A few months later, a nearly complete skeleton was discovered and subsequently sold to the Natural History Museum in London. A second, complete specimen was discovered in 1877 and eventually acquired by the Berlin Natural History Museum. It was named as a different species, *Archaeopteryx siemensii*, by W. Dames in 1897 and as a separate genus *Archaeornis* by B. Petronievics in 1925, although it is widely regarded as a smaller individual of *A. lithographica*.

Five additional well-documented skeletons of *Archaeopteryx* have been discovered since the 1950s. Not all are complete, but all show some evidence of feather impressions. The identity of three skeletons were not recognized at first. A fragment collected in 1855 showing feather impressions, claws, and a knee joint was named as a pterosaur (a flying reptile) by von Meyer in 1859; its correct identity was not realized until 1970. Two specimens had been known for many years and misidentified as examples of the small theropod dinosaur, *Compsognathus*. The largest (sixth) specimen has been described as a separate genus and species, *Wellnhoferia grandis*, on the basis of size and distinctive manual (hand) and pedal (foot) morphology. The most recent discovery in 1993 was named *A. bavarica* by P. Wellnhofer based principally on its longer limb proportions. Eighth and ninth specimens are also known to exist, but they have not been described. *See* FOSSIL.

Morphology. The most striking feature of *Archaeopteryx* is the feather imprints in the surrounding fine-grained lithographic limestone, which are preserved most completely in the London and Berlin specimens. The structure of the primary flight feathers in the wing is exactly like that seen in modern flying birds, with the midrib (rhachis) offset to give a short leading-edge vane and a longer trailing-edge vane. Furthermore, the number and arrangement of primary and secondary flight feathers is identical to that in modern birds. A pair of symmetrical feathers was attached to each tail vertebra. The skull has two temporal openings (as is characteristic of diapsid reptiles), with a large braincase. The jaws were lined with slightly curved, conical, unserrated teeth that were separated from each other by interdental plates that closely resemble those of small theropod dinosaurs.

Many features of the postcranial skeleton are shared with maniraptorian (small advanced theropod) dinosaurs. Significant features include a long scapula and coracoid that are almost equal in length. The lateral orientation of the shoulder socket is intermediate between the dorsolateral position in

modern birds and the ventrolateral position in maniraptorans. The furcula (wishbone), formed from the fusion of the clavicles, is a simple boomerang shape, similar to that in maniraptorans and unlike the long, springy deeply-V-shaped one present in modern birds. There was no ossified sternum (keel); an element originally interpreted as a keel in *A. bavarica* has been reidentified recently as part of a coracoid. The forearm skeleton is long and the wrist includes a semilunate carpal. This element allowed much greater mobility of the hand, which is regarded as a prerequisite of the flight stroke. The hand has three long fingers, ending in hooked claws with long horny sheaths. The hind limb is slender with a long tibia, three walking toes, and a short reversed perching toe. *See* FEATHER.

Biology. The lifestyle of *Archaeopteryx* has been much debated, particularly the issue of whether it was basically terrestrial or arboreal. It may have foraged on the ground; the dentition suggests feeding on a range of small animals—arthropods and perhaps small vertebrates. *Archaeopteryx* had some climbing ability based on the mobility of the hands and curvature of the manual claws. The pedal claws are less curved and reflect a compromise between climbing and terrestrial locomotion.

Archaeopteryx had some flight capability since it possessed flight feathers and wings. Its brain and sensory systems were also flight ready. Computed tomography and three-dimensional reconstruction of the London specimen has shown that the brain volume of *Archaeopteryx* is within in the range for modern birds. It possessed a birdlike brain pattern, with large optic lobes (reflecting the importance of sight), large cerebellar lobes for flight control, and large floccular lobes concerned with balance. The inner ear was also birdlike, with large semicircular canals for balance and spatial orientation.

All the evidence suggests that *Archaeopteryx* was capable of limited, slow, stable powered flight. It lacked anatomical improvements seen in modern birds for maneuverable flight and ground takeoff. A combination of tree climbing and launching from a perch to fly to a safe height may have been a predator escape strategy and a stimulus for the evolution of bird flight.

Phylogeny. It is generally accepted that birds evolved from small feathered maniraptoran dinosaurs. *Archaeopteryx* shares many derived skeletal characters with maniraptorans, particularly the dromaeosaurids, the family to which it is most closely related. The discovery of several lineages of small feathered dinosaurs, including dromaeosaurs, from an exceptional preservation site of Lower Cretaceous age in Liaoning Province, China, strengthens the case for a dinosaurian origin of birds. These discoveries suggest that feathers evolved initially to provide an insulatory covering and were much later elaborated and co-opted for flight. A small minority supports the hypothesis that birds evolved from basal archosaurs. However, there is much testable evidence to support a maniraptoran origin. *See* ANIMAL EVOLUTION.

Angela C. Milner

Bibliography. P. J. Currie et al., *Feathered Dragons*, Indiana University Press, Bloomington, 2004; P. Dominguez Alonso et al., The avian nature of the brain and inner ear of *Archaeopteryx, Nature*, 430:666–669, 2004; A. Elzanoski, Archaeopterygidae (Upper Jurassic of Germany), pp. 129–159, in L. M. Chiappe and L. M. Witmer (eds.), *Mesozoic Birds: Above the Heads of the Dinosaurs*, University of California Press, Berkeley, 2002; A. C. Milner, *Dino-Birds*, Natural History Museum, London, 2002; K. Padian and L. Chiappe, The origin and early evolution of birds, *Sci. Amer.*, pp. 28–33, February 1998.

Archaic ungulate

The most diverse group of early Tertiary mammals are the archaic ungulates (condylarths). Although closely related, they are not a monophyletic group but, along with various extinct and extant mammals, form the taxon Ungulata (**Fig. 1**). They are ancestral to as many as 7 of 18 living orders of mammals: Artiodactyla, Cetacea, Hyracoidea, Perissodactyla, Proboscidea, Sirenia, and possibly Tubulidentata. Along with the Late Cretaceous zhelestids, best known from Asia, Ungulata forms the taxon Ungulatomorpha. Fossil (and molecular) evidence suggests that Ungulatomorpha separated from other placentals 85 million years ago.

Early ungulatomorphs had lower-crowned, more squared molars compared to contemporary placentals, which have a more slicing dentition, indicating a trend toward omnivory and herbivory. Although "ungulate" implies hooves, most archaic ungulates had at best rudimentary hooves or even claws.

Arctocyonidae is the sister taxon to other ungulates, with the possible exception of artiodactyls. Arctocyonids retained more primitive character states in their dentition, such as more teeth and relatively little molarization of premolars. Smaller arctocyonids were like the coati-sized *Chriacus*, a scansorial mammal probably equally adept in trees or on the ground. It had powerful limb musculature, mobile joints, flexible plantigrade pentadactyl feet with claws, a slightly divergent hallux (first digit of the hindfoot), and a long and robust tail for grasping or taking hold of something. Wolf-sized arctocyonids such as *Arctocyon* were terrestrial with an extremely long powerful tail, short stout limbs, a plantigrade manus (forefoot) and pes (foot), and an opposable hallux. The dentitions of all arctocyonids were low-crowned with rounded cusps, and moderately to well developed canines. Arctocyonids are thought to have had a bearlike diet of fruits, nuts, seeds, insects, and small animal prey. Arctocyonids reached their diversity peak in the middle to late Paleocene, disappearing in the early Eocene. They are known from North America and Europe.

Members of Hyopsodontidae were generally smaller animals. The ubiquitous, rat-sized early Eocene *Hyopsodus* was quite slender and possessed shortened limbs. The hyopsodontid dentition was

similar to that of arctocyonids but showed the beginnings of molarization of premolars and a hint of crest development on molars. Hyopsodontidae ranges from early Paleocene through late Eocene, reaching its peak in late Paleocene. They are best known from North America and Europe, but with Asian and north African records.

Mioclaenidae was quite different in dental architecture. Mioclaenids had simplified biting surfaces and some premolar inflation, suggesting consumption of tougher, fibrous plants. They seem in general to have been slightly larger (up to hare size) than hyopsodontids, but similar to smaller arctocyonids. The mioclaenids are diverse in the middle Paleocene but appeared in the early Paleocene and disappeared in the late Paleocene. They occurred in North and South America and Europe.

Didolodontidae, from the middle Paleocene through Oligocene of South America, was dentally similar to Mioclaenidae. The size of didolodontids was similar to or slightly larger than that of mioclaenids. Didolodontidae gave rise to the totally extinct South American Meridiungulata, comprising Litopterna, Astrapotheria, Notoungulata, Pyrotheria, and Xenungulata.

Premolar inflation was greater in Periptychidae than Mioclaenidae. In *Periptychus*, the quite bulbous premolars were grooved. Some molar cusps were accentuated, while crests developed both on molars and on premolars. This suggests that tough, fibrous plant material was eaten by the sheep-sized periptychids. Some smaller, squirrel-sized periptychids showed similar specializations but, by virtue of size, probably ate higher-energy diets, which may have included fruits, nuts, or insects. Postcranially, periptychids are best known from the squirrel-sized *Gillisonchus* and the sheep-sized *Ectoconus*. In size and general body proportions, but not in head, tail, or pedal structure, *Ectoconus* resembled the aardvark, *Orycteropus*. *Periptychus* and *Ectoconus* were stoutly built with no reduction of the distal ulna or the fibula and no elongation of limbs, and had a pentadactylous manus and pes. At least *Ectoconus* among periptychids possessed a subplantigrade manus and pes that was short and wide with considerable padding below, wide spreading digits, and small tapirlike hooves (**Fig. 2**). Periptychidae was taxonomically diverse in the early Paleocene of North America but rapidly declined, disappearing in late Paleocene.

Triisodontines, Hapalodectidae, and Mesonychidae, along with Cetacea, constitute the Cete. Artiodactyla might be the sister taxon to Cete. In the earliest triisodonts, the suggestion of later dental trends toward carnivory can be seen. In cetans such as *Synoplotherium* and *Harpagolestes*, excessive wear on molars and incisors, massiveness of some of the cusps, and large canines point to carrion feeding. *Dissacus*, *Mesonyx*, and *Hapalodectes* lacked these specializations but probably were carnivorous, as was the slender-toothed *Hapalodectes*. These animals ranged from the raccoon-sized *Microclaenodon* and coyote-sized *Dissacus* to the wolf-sized *Mesonyx* and bear-sized *Harpagolestes* and *Pachyaena*. Based on *Mesonyx*, the skeleton of cetans developed more cursorial adaptations. Cetan skeletons tended to be more gracile, and in some such as *Mesonyx* the manus and pes were digitigrade (walking on toes) with closely aligned metapodials. Some taxa such as *Dissacus* retained the primitive plantigrade posture and hooved unguals, rather than reverting to claws as in some arctocyonids. The triisodonts were moderately diverse in the early and mid Paleocene. In North America and Asia (and a lesser extent Europe), another radiation occurred during the Eocene, before cetans disappeared.

Fig. 1. Cladogram relating archaic ungulates to major clades of extant and extinct ungulates.

Phenacodontidae was taxonomically the least diverse of the archaic ungulates. Both dentally and postcranially, phenacodontids showed several trends away from arctocyonids. In *Phenacodus*, whose teeth were quite low-crowned with distinctive cusps, there was a trend toward lophs and crests on both upper and lower molars. Phenacodontids tended toward molarization of the premolars more than other archaic ungulates. The molarization of the premolars and the development of lophs and selenes (crescents) reached its extreme in *Meniscotherium*. Postcranially, the sheep-sized *Phenacodus* is one of

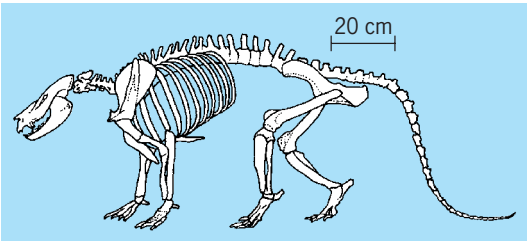

Fig. 2. Skeleton of *Ectoconus*, an early Paleocene phenacodont condylarth.

the best-known taxa. Its limbs were somewhat elongated, but flexibility was still similar to arctocyonids except possibly in the carpus and tarsus. Digits I and V were reduced. Phenacodontidae appeared in the late Paleocene and disappeared in the middle Eocene. They were from North America and Europe. Phenacodontidae is the probable sister taxon of Paenungulata, which among extant mammals includes Perissodactyla, Hyracoidea, Proboscidea, and Sirenia. J. David Archibald

Bibliography. M. J. Benton, *Vertebrate Paleontology*, Chapman & Hall, London, 1997; R. L. Carroll, *Vertebrate Paleontology and Evolution*, W. H. Freeman, New York, 1988.

Archean

A period of geologic time from about 3.8 to 2.5 billion years ago (Ga). During the Archean Eon a large percentage of the Earth's continental crust formed, plate

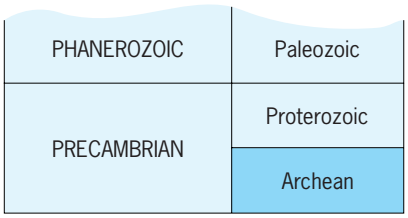

tectonics began, very warm climates and oceans existed, and life appeared on Earth in the form of unicellular organisms.

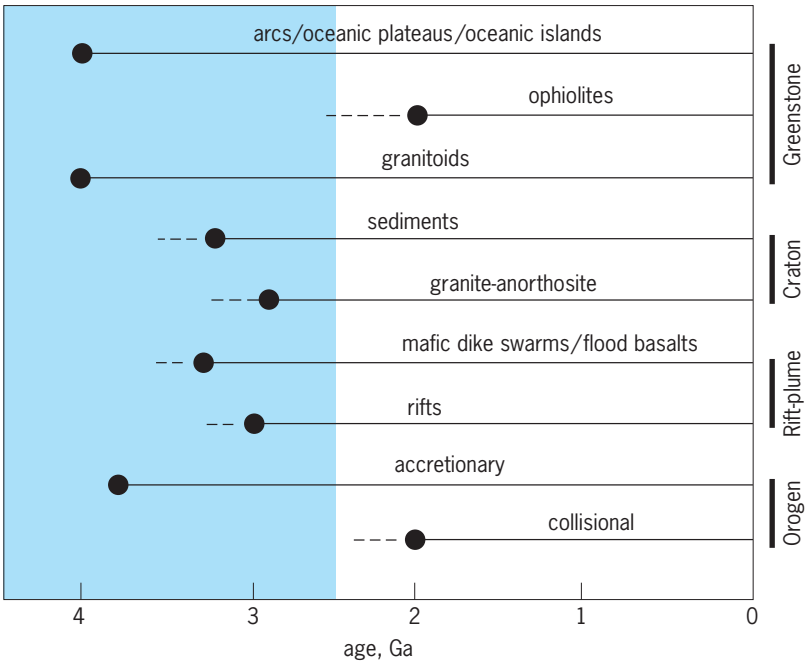

Fig. 1. Distribution of tectonic settings with geologic time. Note that most modern tectonic settings extend back into the Archean (shaded region). Ophiolites are remnants of oceanic crust.

Tectonic regimes. Plate tectonics can be tracked with time by using rock assemblages that are characteristic of certain tectonic settings. Although at one time greenstones (rock successions in which submarine basalts and related sediments dominate) were thought to be an Archean phenomenon, it is now clear that they have formed throughout geologic time. It is also clear that all greenstones do not represent the same tectonic setting: some reflect island arcs, whereas others reflect oceanic plateaus, volcanic islands, or oceanic crust. Except for ophiolite greenstones (remnants of oceanic crust), greenstones are recognized throughout the geologic record from the oldest known rocks at 4 Ga to the present (**Fig. 1**). The oldest rocks indicative of craton and passive-margin environments are sediments about 3 Ga, suggesting that cratons developed by this time, very early during the Archean, although they may not have been large in area. Giant dike swarms and flood basalts, which are probably derived from mantle plumes, are present in southern Africa and western Australia by 2.7 Ga, with the oldest known dike swarms dating to over 3 Ga (Fig. 1). The oldest continental rift assemblages, comprising basalts and immature sediments, occur in southern Africa at about 3 Ga, and the oldest accretionary orogens (linear belts of deformation resulting from plate collisions) are the Acasta gneisses (4 Ga) and the Amitsoq gneisses (3.9 Ga) in northwest Canada and southwest Greenland, respectively. Although the oldest well-documented collisional orogens are Paleoproterozoic (2.5–1.6 Ga), it is likely that late Archean collisional orogens with reworked older crust occur on some continents. *See* CRATON; GNEISS; PLUTON.

The occurrence of rock assemblages typical of arcs, oceanic plateaus, and oceanic islands and the presence of accretionary orogens in the very earliest vestiges of the geologic record at 4–3.5 Ga strongly supports some sort of plate tectonics operating on the Earth by this time. By 3 Ga, cratons, passive margins, and continental rifts were also widespread. Although plate tectonics appears to have occurred since 4 Ga, there are geochemical differences between Archean and younger rocks that indicate that Archean tectonic regimes must have differed in some respects from modern ones. The degree that Archean plate tectonics differed from modern plate tectonics is unknown; however, these differences are important in terms of the evolution of the Earth. *See* PLATE TECTONICS.

Earth's oldest rocks. The oldest rocks occur as small, highly deformed terranes tectonically incorporated within Archean crustal provinces (**Fig. 2**). Although the oldest known igneous rocks on Earth are the 4 Ga Acasta gneisses of northwest Canada, the oldest minerals are detrital zircons (zircons in sediments) from the 3 Ga Mount Narryer quartzites in western Australia. These zircons have uranium-lead (U-Pb) ages, determined by ion probe, ranging from about 4.3 to 3.5 Ga, and are important in that they indicate the presence of granitic sources for the sediments, some of which were 4.3 Ga. These sources may have been remnants of continental crust,

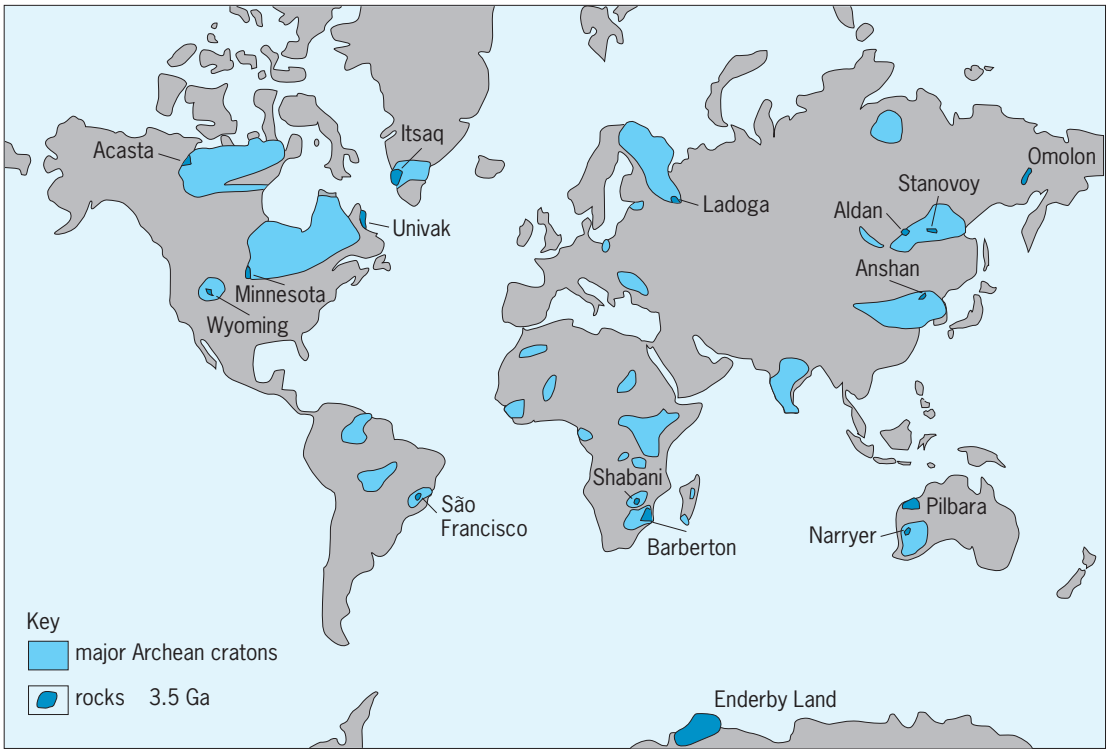

Fig. 2. Distribution of the Earth's oldest rocks.

although the volume of this crust may have been small.

The oldest isotopically dated rocks on Earth are the Acasta gneisses, which are a heterogeneous assemblage of highly deformed granitic rocks, tectonically interleaved with mafic and ultramafic rocks, and metasediments. Uranium-lead zircon ages from the granitic components of these gneisses range from 4 to 3.6 Ga, and thus it would appear that this early crustal segment evolved over about 400 million years and developed a full range in composition of igneous rocks. The chemical compositions of Acasta mafic rocks are very much like less deformed Archean greenstones representing various oceanic tectonic settings. *See* DATING METHODS; ROCK AGE DETERMINATION.

The largest and best-preserved fragment of early Archean continental crust is the Itsaq Gneiss Complex in southwest Greenland. In this area, three terranes, each with its own tectonic and magmatic history, collided about 2.7 Ga, forming the continental nucleus of Greenland. Although any single terrane records less than 500 million years of precollisional history, collectively the terranes record over 1 billion years of history before their amalgamation. The most extensively studied greenstone in southwest Greenland is the Isua sequence, composed of basalts and komatiites interbedded with banded iron formation, intrusive sheets of granite, and volcanic sediments.

The Pilbara craton in western Australia also comprises terranes that accreted between 3.46 and 3.2 Ga. Within this craton are thick successions of submarine volcanics, volcanic sediments, and chemical sediments (chert and banded iron formation). Some of these rest unconformably on an older greenstone complex with a U-Pb zircon age of about 3.5 Ga, an important observation indicating that land emerged above sea level by 3.46 Ga. The Barberton greenstone in southern Africa is one of the most studied early Archean greenstones. Together with coeval granites, the Barberton succession formed between 3.55 and 3.2 Ga and includes four tectonically juxtaposed terranes, with similar stratigraphic successions in each terrane. Each succession begins with submarine basalts and komatiites (high-temperature, magnesium-rich lavas) that may be remnants of oceanic plateaus. Overlying these rocks is a suite of felsic to basaltic submarine volcanics, fine-grained volcanic sediments and cherts, possibly representing an oceanic island arc.

Mineral deposits. The Archean is known for its reserves of iron, copper, zinc, nickel, and gold. Some of the world's largest copper-zinc deposits occur as massive sulfide beds associated with submarine volcanics in Archean greenstones in Canada and western Australia. These deposits are typically lenticular in shape with more than 60% massive sulfides of zinc and iron at the top, with a copper-rich base. They are typically associated with hydrothermal cherts and underlain by large alteration zones. These deposits are formed at the same time as surrounding volcanic rocks by the leaching of metals from the volcanics by hydrothermal fluids and deposition on the sea floor. *See* CHERT.

Nickel sulfides are associated with komatiite flows in Archean greenstones, and they are particularly important in western Australia. Like the copper-zinc

deposits, they are layered deposits of finely disseminated ores within submarine komatiite flows. The komatiites appear to be lava channels following submarine topography. Typical statiform nickel ores consist of a thin rich sulfide layer (>80% sulfides) overlain by a layer of disseminated sulfides. The massive Archean nickel-sulfide ores appear to be of magmatic origin. Voluminous submarine komatiite flows assimilated sulfide-rich sediments during eruption of the sea floor, and immiscible nickel-sulfide droplets settled from the komatiitic lavas, producing the massive layered nickel ore zones at the base. *See* LAVA.

Banded iron formations are common in Archean greenstones worldwide, and many of these deposits are important sources for iron ore. Banded iron formations are sedimentary rocks formed by chemical precipitation of iron oxides in oxygen-depleted seawater. The delicate banding, which can be followed over great distances in these sediments, suggests an origin by cyclic sedimentation in an arc-related basin, or in some instances, like the banded iron formations in the Hamersley Basin in western Australia, in passive margin or intracratonic basins. *See* BANDED IRON FORMATION.

The Archean is a principal producer of the world's gold. The lode gold deposits typical of Archean greenstones vary in their geologic setting, ore minerals, and styles of mineralization. Silver is typically associated with the gold, with varying amounts of such elements as arsenic, antimony, and tungsten. Most of the Archean gold is associated with altered country rocks in which carbon dioxide (CO_2) and potassium have been introduced with salt-poor, gold-bearing fluids. Almost all Archean deposits are controlled by shear zones or faults, and mineralization occurred during the late stages of deformation or after deformation. Studies have shown that many of the deposits formed at shallow depths (<5 km or 3 mi) from hydrothermal fluids with a wide range of temperatures (200–500°C; 390–930°F).

Plate tectonics. Because the Earth has steadily cooled with time, there is reason to suspect that plate tectonics may have been different in the Archean. For instance, a hotter mantle in the Archean would have produced more melt at ocean ridges and hence a thicker oceanic crust, perhaps as thick as 20 km (12 mi). Since oceanic crust is less dense than the mantle, Archean oceanic plates would have been more buoyant and thus more difficult to subduct. A hotter mantle would also convect faster, probably causing plates to move faster, and hence oceanic plates would have less time to cool and become gravitationally unstable. Both of these factors would tend to cause plates to resist subduction in the Archean.

Because of the difficulty of subduction, Archean plates may not have been moving fast enough to remove the excess Archean heat. Although it seems that mantle plume activity was probably greater in the Archean than afterward, plumes cannot be a substitute for heat loss by plate tectonics, since they bring heat into the mantle from the core. In effect, plates cool the mantle while plumes cool the core. What is required is a mechanism in the lithosphere that promotes heat loss at the surface. Three possible explanations for excess Archean heat loss are: (1) The Archean ocean ridge system was unusually long. (2) The inversion of basalt to eclogite (a high-density mafic rock) in thick Archean oceanic crust may have increased the density of the lithosphere sufficiently for it to subduct. (3) Spreading rates of Archean ocean ridges may have been faster than at present.

Although plate tectonics undoubtedly operated in the Archean, it must have differed in some fundamental aspects from modern plate tectonics to effectively remove the excess heat from the Earth's upper boundary layer. Perhaps the most important factor leading to evolutionary changes is the cooling of Earth. If fragments of oceanic plateaus constitute a significant proportion of Archean greenstones, as seems likely, Archean mantle plume activity may have been more widespread than at present. Perhaps the first continental nuclei were oceanic plateaus (**Fig. 3**). Partial melting of the thickened roots of these plateaus would produce granitic magmas, changing the plateaus into continental nuclei. Shallow subduction around the perimeters of the plateaus would result in both lateral continental growth and thickening of the lithosphere, and collision of oceanic plateaus and island arcs around the edges would contribute to continental growth. *See* CONTINENTS, EVOLUTION OF.

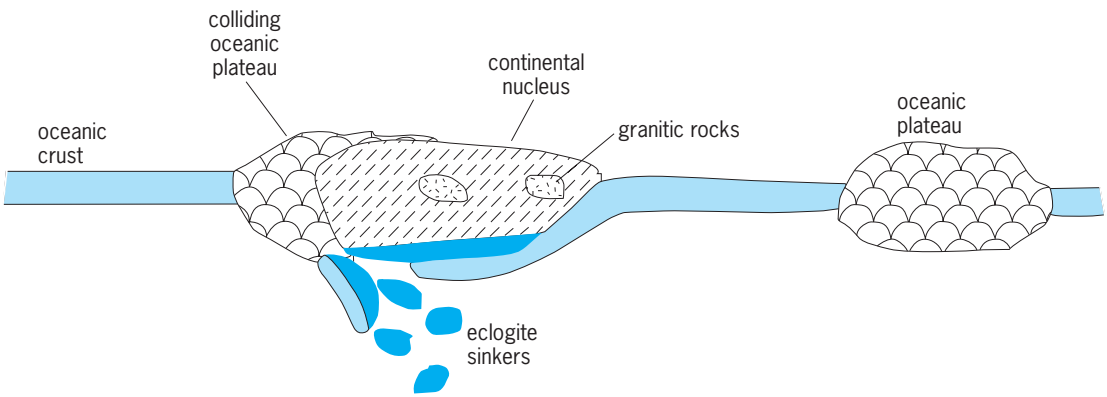

Fig. 3. Model of Archean plate tectonics. Eclogite is a heavy mafic rock that may have pulled buoyant subducting plates into the mantle.

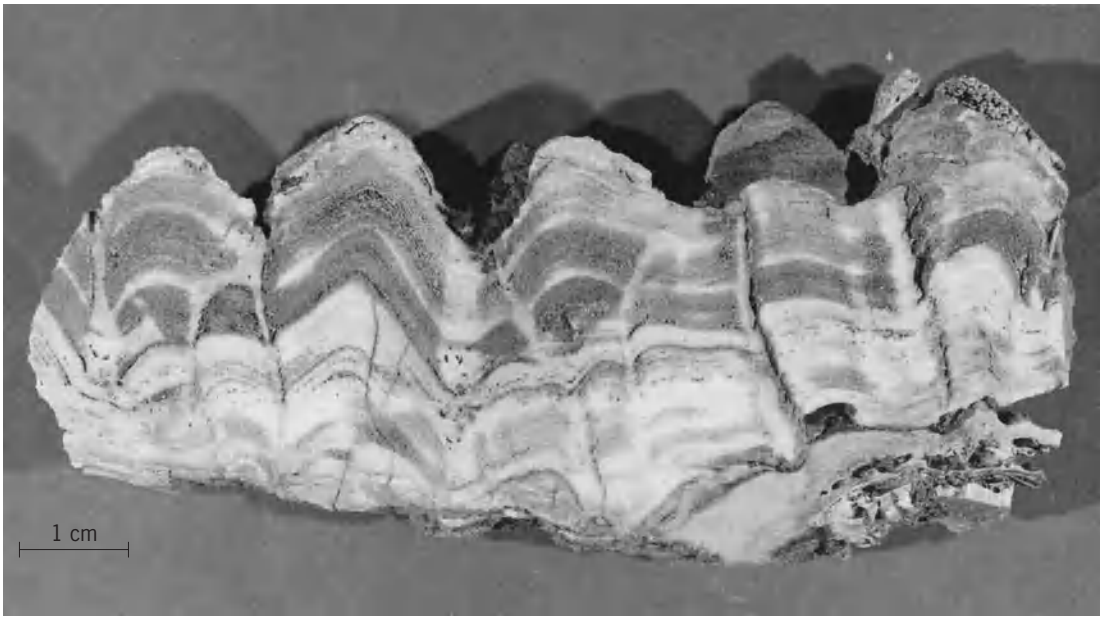

Fig. 4. Early Archean stromatolites from the Barberton greenstone in southern Africa. The laminae were probably deposited by cyanobacteria, and are among the oldest known fossils. (*Courtesy of Don Lowe*)

Atmosphere and oceans. The Earth's first atmosphere was probably composed chiefly of gases such as helium and hydrogen inherited from the solar nebula from which the solar system formed, as well as from the asteroidlike bodies that collided to form Earth. As Earth heated up from core formation, it released gases and formed a secondary atmosphere composed chiefly of CO_2, methane (CH_4), nitrogen (N_2), and water (H_2O). In support of this view, the surviving rock record includes carbonates that reflect a carbon dioxide–rich atmosphere; and also one or more greenhouse gases (carbon dioxide, methane) must have been present to prevent the surface of the Earth from freezing over. The Earth's surface could not have been frozen, because there are sedimentary rocks that were deposited in seawater by 4 Ga. If carbon dioxide and methane were very abundant in the Archean atmosphere, as seems probable, Archean climates must have been very warm due to greenhouse warming. In fact, oxygen isotopes from marine sediments such as chert indicate hot oceans, perhaps reaching 80°C (176°F). The Archean atmosphere and oceans contained very little free oxygen, since Archean sediments contain reduced detrital minerals such as uraninite (UO_2) that could not have survived in an oxidizing system. Also, ancient soil horizons preserved in Archean rocks contain iron chiefly in the reduced state (Fe^{+2}), reflecting low oxygen contents in the atmosphere. *See* ATMOSPHERE; GREENHOUSE EFFECT.

Changes in atmosphere composition in the late Archean and Proterozoic occurred in response to declining oceanic volcanism and hydrothermal input into the oceans and atmosphere after 2.4 Ga. Less carbon dioxide and methane entered the near-surface reservoirs. Also, a supercontinent formed at the end of the Archean, increasing the area of land available for weathering. The weathering extracted carbon dioxide from the atmosphere, causing a decrease in the amount of atmospheric carbon dioxide, which cooled the atmosphere and led to the first worldwide glaciations about 2.3 Ga. As this supercontinent broke up about 2.1 Ga, organic matter was rapidly buried in numerous small basins around the continental fragments, leading to a substantial increase in oxygen in the atmosphere-ocean system.

Life. There are three lines of evidence for life in the Archean: (1) fossil stromatolites, which are laminated structures deposited by microorganisms; (2) fossils of cells or cellular tissue; and (3) carbonaceous matter identifiable from its carbon isotopic composition as a product of biologic activity. Some of the oldest fossil stromatolites occur in the 3.5 Ga Barberton greenstone in southern Africa and in the 3.5 Ga Pilbara greenstone in western Australia. These have wavy or domical shapes and occur in thin chert beds (**Fig. 4**). Threadlike filamentous microfossils of bacteria also occur in some of these cherts. In the Kromberg Formation (3.4 Ga) in southern Africa, the filamentous fossils occur in association with wavy flat-laminated stromatolites composed in part of amorphous carbon. Some of the best-preserved fossil bacteria occur in the Apex chert in western Australia, where cell structure is preserved. A biologic origin for these structures is established by their morphology, characterized by well-defined, barrel- or disc-shaped cells with rounded or conical terminal cells (**Fig. 5**). The similarity of these forms to unbranched, uniseriate prokaryotic cells of cyanobacteria, both living and fossil, is striking. *See* FOSSIL; STROMATOLITE.

It is well known that carbonaceous matter of biologic origin is enriched in the light isotope carbon-12 (^{12}C), compared to carbon-13 (^{13}C). The oldest carbon isotopic evidence for terrestrial life comes from the 3.7-Ga Isua greenstone in western

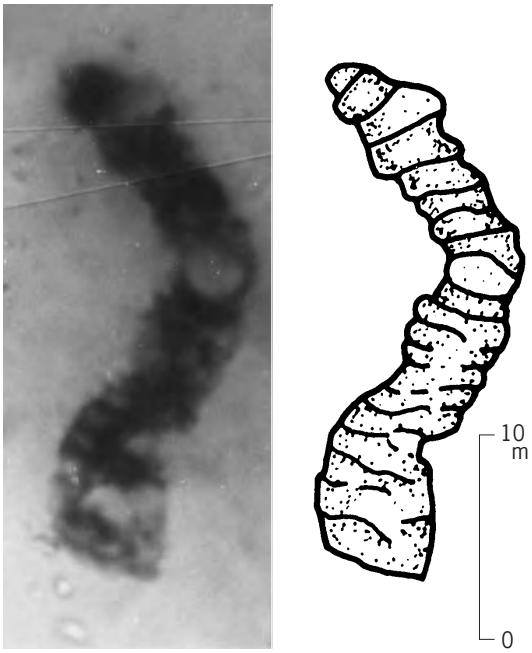

Fig. 5. Typical microfossil bacteria from the 3.46-Ga Apex chert in western Australia, shown in a photograph and interpretive drawing. (*Courtesy of J. W. Schopf*)

Greenland from carbonaceous samples taken from deformed and metamorphosed sedimentary rocks, probably deep-sea sediments. Microscopic carbonaceous globules analyzed from these rocks suggest that these globules came from planktonic, photoautotrophic microorganisms. When the fossil and carbon isotopic evidence is considered collectively, investigators conclude that autotrophic forms of primitive life were well established on the Earth by 3.5 Ga.

Catastrophic mantle plume event. An increase in the production rate of continental crust at 2.7 Ga is commonly attributed to a rapid change from layered to whole-mantle convection in the Earth. The result was a catastrophic mantle plume event, a short period (<50 million years) during which the lithosphere was bombarded by numerous mantle plumes. Evidence supporting this mantle plume event includes a large peak in the frequency of komatiites, oceanic plateau basalts, flood basalts, giant dike swarms, and large layered intrusions. The formation of perhaps the first supercontinent followed the mantle plume event and peak crustal production by 10–50 million years, as indicated by U/Pb zircon chronology of Archean rocks. A peak in gold production between 2700 and 2650 Ma also correlates with supercontinent formation.

A catastrophic mantle plume event at 2.7 Ga would also be expected to affect the ocean/atmosphere/biosphere systems and leave a permanent imprint in the geologic record. For example, (1) low sea level at 2.7 Ga may reflect direct hits of mantle plumes beneath the only two large cratons, Kaapvaal and Pilbara; (2) shales of this age derived from intensely weathered sources may record global warming caused by increased input of greenhouse gases (carbon dioxide and methane) into the atmosphere; (3) a corresponding peak in black shale deposition may reflect increased input of nutrients into the oceans and increased anoxia related to methane input; (4) a strong peak in banded iron formation deposition at 2.7 Ga may be related to enhanced hydrothermal spring activity pumping more iron into the oceans; (5) an increase in the number of occurrences of stromatolites per unit volume of juvenile continental crust at 2.7 Ga may reflect an increase in biomass related to enhanced nutrient/carbon dioxide levels caused by a mantle plume event; (6) a well-defined decrease in the carbon-12/carbon-13 ratio of kerogens at 2.7 Ga could reflect enhanced activity of methanotropic bacteria due to greater input of methane into the oceans. *See* KEROGEN.

Kent C. Condie

Bibliography. S. Bengtson (ed.), *Early Life on Earth*, 1994; K. C. Condie (ed.), *Archean Crustal Evolution*, 1994; K. C. Condie, *Plate Tectonics and Crustal Evolution*, 4th ed., 1997; E. G. Nisbet, *The Young Earth*, 1987.

Archeoastronomy

The interdisciplinary study that attempts to determine how much astronomy prehistoric people knew and how it influenced their lives. It involves multiple disciplines: astronomy to chart the heavens, archeology to probe the cultural context, engineering to survey sites, and ethnology to provide clues to the cultural past. Archeoastronomy has prompted valuable insights into the astronomy of the past, even to revolutionizing some models of prehistoric cultures. *See* ASTRONOMY.

It has been suggested that archeoastronomy and its loose family of disciplines should be subsumed under a broader study, cultural astronomy. The reason to do cultural astronomy is that the sky can perform a special role in the scheme of cultural systems. The sky then serves as a cultural resource of many uses.

The cultural context is the key to understanding the findings of archeoastronomers. Finding astronomical orientations at sites is easy; interpreting these as intentional alignments is hard. It is necessary to consider what their purpose might be (to keep a seasonal calendar? to regulate sacred time? both or neither?). The great danger is the imposition of modern astronomy and culture upon an alien culture of the past.

Naked-Eye Astronomy

The kinds of observations of key celestial cycles that can be made without a telescope are reviewed in this section. Only the Sun and the Moon will be considered.

Sun. Most people are aware that the height of the Sun in the sky at noon changes with the seasons, with the greatest height in summer and the lowest in winter. There is, however, less familiarity with the Sun's seasonal motion along the horizon. On the day of the summer solstice (around June 22), for

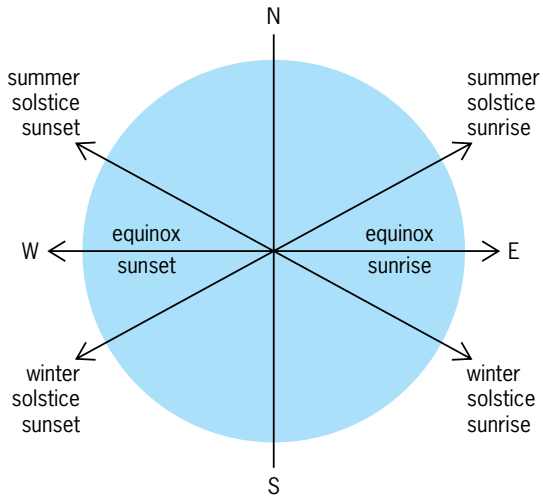

Fig. 1. Angular swing of the Sun along the horizon at rising and setting from solstice to solstice. The angle is for a latitude of 36°, that of the North American Southwest.

example, the Sun rises the farthest north of east that it will get for the year (**Fig. 1**). On the equinoxes (March 21 and September 23), it rises due east. And on the winter solstice (December 22), it reaches its farthest point south of east. The same occurs, mirror-reflected, at sunset. (This description refers to the Northern Hemisphere midlatitudes; in the Southern Hemisphere, summer and winter are reversed by date.) *See* EQUINOX; SEASONS; SOLSTICE.

Thus, from summer to winter, the sunrise point moves to the south; from winter to summer, to the north. The rate at which the sunrise point moves from day to day varies during the year. At the solstices, the sunrise points do not noticeably move for a few days. The Sun appears to "stand still" (which is the meaning of the word solstice). In contrast, at the equinoxes, the sunrise points move at their fastest rate, by almost the Sun's own diameter in a day at midlatitudes.

This seasonal voyage of the Sun along the horizon differs with latitude with respect to the size of the solstice-to-solstice swing along the horizon. At 36°, the latitude of the North American Southwest, the swing amounts to 60°, one-sixth of the total horizon circle (Fig. 1). Farther north, the arc is greater; at the latitude of Stonehenge (about 51°), it is about 80°.

The solstitial horizon points are symmetrical, and so they are also reversible. If an observer faces points that align to the summer solstice sunrise, then turning around and sighting along the points in the opposite direction reveals very closely the winter solstice sunset point (with a flat horizon).

Moon. The Moon's most obvious change is that of its phases. The month of phases (synodic month), the time from one phase of the Moon to the repetition of that phase, averages 29.5 days. *See* PHASE (ASTRONOMY).

Suppose the point of moonrise is observed for a month. It would be seen that the moonrise point varies from a point farthest south to one farthest north during the month. In other words, the moonrise motion mimics the sunrise motion but occurs about 12 times as fast. Depending on when the observations are made, the moonrise arc may be larger than, the same as, or smaller than the sunrise arc. This difference results because the Moon's path in the sky with respect to the stars is not the same as the Sun's, but is inclined at about 5°, crossing the Sun's path at two points. Thus, the Moon can appear as much as 5° below the Sun's path, 5° above it, or right on it. *See* MOON.

Complication. The matter is complicated in that the two points where the Sun's and Moon's paths cross (the nodes) move with respect to the stars, taking 18.6 years to circle the sky once. The result is that when the Moon's path reaches its highest point above the Sun's, the Moon's horizon swing is greater than the Sun's. When the two line up, the swings are the same. When the Moon's path falls below the Sun's, the total arc is less.

How much greater or less can be considerable. At a latitude of 36°, the Moon moves through a maximum arc of 70° and a minimum arc of 45° during the 18.6-year cycle (**Fig. 2**). In analogy to the Sun standing still at the solstices, the two extremes of the Moon's positions are also called standstills: major standstill for the maximum angle and minor standstill for the minimum, with 9.3 years between. Again in analogy to the Sun, these angular changes are more pronounced at more northern or southern latitudes.

When the Moon moves from minor to major standstill, its monthly arc stays within that of the Sun for about half the time. Then, an alignment that works for the Moon will work for the Sun at some time during the year. Hence, it may be that an orientation is not specifically for the Moon. However, during the time when the Moon's swing lies outside the angle of the Sun, an orientation can apply to the Moon only, and not at all to the Sun. (Again, these alignments are reversible.)

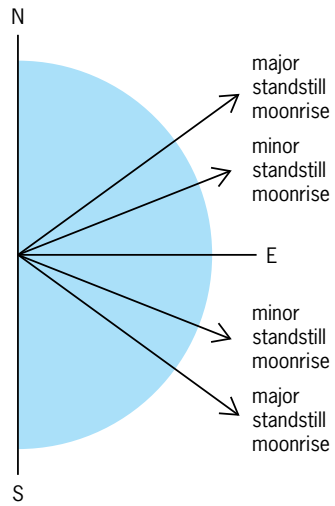

Fig. 2. Monthly angular swing of the Moon along the horizon at moonrise, for times of maximum angle (major standstill) and minimum angle (minor standstill). The angles are shown for a latitude of 36°.

Horizon-marking system. It can now be seen how a simple horizon-marking system is set up. First, a location with a clear view of the horizon must be found, with at least a few prominent features over the angular range of the sunrise (or sunset). Then it is necessary to return to this spot daily and note the rising positions of the Sun throughout the year at significant times: the solstices, the equinoxes, and, perhaps, important times to plant crops. Thus, a basic solar calendar is established. Since the Sun's positions at various dates along the arc remain fixed for a long time, once established the solar calendar will be good for many years.

A solar-horizon calendar keeps reliable track of the seasonal year. To subdivide that interval, people usually rely upon a lunar calendar that follows the phases of the Moon. The essential astronomical problem is that a seasonal cycle and a lunar-phase cycle are incommensurate. The solution can involve complicated, long-term strategies, such as that used in the Gregorian calendar, or more casual, practical ones, such as counting a "short" month of a few days at the end of each year. *See* CALENDAR.

The Old World: Stonehenge and Other Megalithic Sites

In a direct sense, Stonehenge created the interest in archeoastronomy, and the site exemplifies the problems and potential of the archeoastronomical enterprise.

Stonehenge is popularly known for the massive upright stones that form a central horseshoe and circle (some 82 ft or 25 m in diameter) in the center of the site (**Fig. 3**). Such large stones are commonly called megaliths in Great Britain; this term has come to be applied to all sites where stones, even fairly small ones, are arranged in some pattern. The horseshoe opens out on the main axis of Stonehenge, called the Avenue. Some 260 ft (80 m) from the center, within but not in the center of the Avenue, sits the tilted Heel Stone. This main axis of Stonehenge aligns roughly with the summer solstice sunrise. In fact, at summer solstice the Sun rises somewhat to the left of the Heel Stone as seen by an observer near the center of the structure.

Astronomical alignments at Stonehenge. The modern controversy concerning the astronomical use of Stonehenge developed in the 1960s. G. Hawkins, an astronomer, searched for astronomical alignments to the Sun, Moon, stars, and planets for the main features of the site. He found them for the Sun and the Moon, including moonrise and moonset during major and minor standstills. Later, he proposed that the site could even have been used to anticipate times of possible eclipses.

Radiocarbon dates indicate that Stonehenge was built over a span from 3100 to 1000 B.C. in three separate stages. The muddle over the astronomical use of Stonehenge comes, in large part, from the fact that it is a mosaic of structures, most likely built by different people, perhaps with a continuity of design. The great stones were erected between 2000 and 1500 B.C.; they attract the most attention now, but other aspects are also important.

It is the earliest parts of Stonehenge, constructed between 3100 and 2100 B.C., that have the most astronomical promise (**Fig. 4**). These comprise the outer earthwork ring and ditch (about 330 ft or 100 m in diameter and 7 ft or 2 m high) broken only in the direction of the Heel Stone; a ring of 56 holes (the Aubrey Holes) that were dug and then quickly filled with chalk; an array of postholes near the opening to the Heel Stone; and the four Station Stones that lie along the circle of the Aubrey Holes.

Lunar and solar observing can be done with these elements. The four Station Stones form a fairly good

Fig. 3. Inner great trilithons of Stonehenge. (*Courtesy of O. Gingerich*)

rectangle. From its center, the summer solstice Sun rises along the opening to the Heel Stone. The short sides of the rectangle are parallel to this line, so they point to the summer solstice sunrise and winter solstice sunset. The long sides of the rectangle and its diagonals line up the moonrises and moonsets at the major and minor standstills. (These are reversible alignments.)

Hawkins also contended that the inner megaliths of the horseshoe sighting outward through the ring around them also aligned to important settings and risings of the Sun and Moon when viewed from inside the monument.

Evaluation of claims. Using positions near the inner megaliths as observing stations fails, whether for the Sun or the Moon. The gaps between the megaliths are wide, and depending on where the observer stands, the views can cover wide angles on the horizon. As noted above, even the central alignment—observing the summer solstice sunrise over the Heel Stone—does not work now, nor did it ever operate in the past 5000 years. Finally, it is not clear where the observer is supposed to stand among the massive stones.

J. North flipped the situation around so that the astronomical arguments gain considerable strength. In this interpretation, the observer stands outside the monument and uses it as the foresight. For example, an observer standing next to the Heel Stone would see standing stones and their lintels block out the glare of the western horizon at sunset. (It is again important that the extremes of the Sun and Moon are reversible alignments.) Only narrow sight lines are visible; one reveals the winter solstice setting Sun. As the Sun descends, the stones make a dark mass, pierced by sunlight through a central slit near the base of the stones. Also from the Heel Stone, the Moon's minor southern standstill appears through a central slit (not the same one as the winter solstice Sun).

The focus on the winter solstice sunset inspires a new view of the Avenue, which comes into Stonehenge from the summer solstice sunrise direction. As people in a precession around sunset of the winter solstice paced up the Avenue, they would behold first the top of the monument, and more would slide into view as they approached it. If timed right, they would reach the Heel Stone at just the moment of the final gleam of the winter sun.

The alignment of the sides of the inner rectangle holds up, in part because of its fairly large size (112 by 260 ft or 34 by 79 m). At Stonehenge's latitude, the solar and lunar extreme lines naturally cross almost at right angles; these astronomical symmetries seem to have caught the attention of the builders of this earlier part of the structure. The Station Stones fall closely around a circle with a radius and center roughly that defined by the Aubrey Stones. Using the edges of the Station Stones as backsights and the bank as an artificial horizon, the spectacle of the northern extreme moonset works out well. All told, the older parts of Stonehenge make a reasonable solar and lunar observatory.

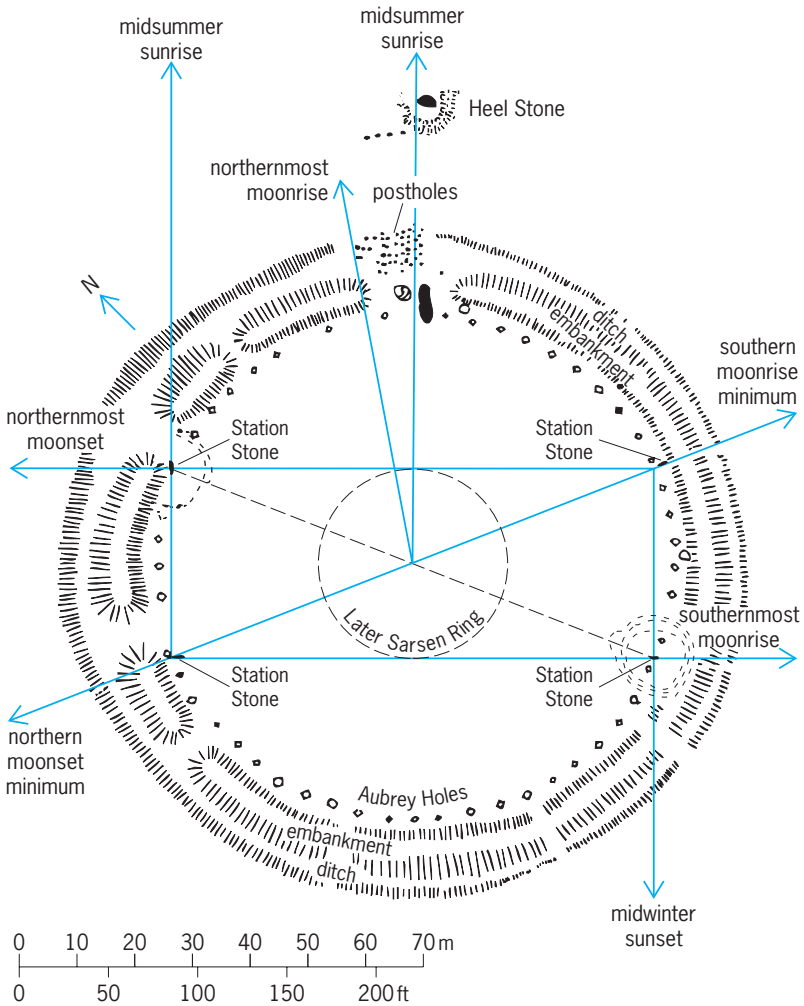

Fig. 4. Diagram of the major features of Stonehenge that may have been used astronomically. (*After O. Gingerich, in K. Brecher and M. Feirtag, eds., Astronomy of the Ancients, MIT Press, 1979*)

Other megalithic sites. The basic problem in all of this is that even if the astronomy works, the cultural context is unclear. Horizon watching can be used simply to tell the time of year or more forcefully to set a ritual calendar. One hope of guessing about the importance of astronomy in megalithic societies is to examine many sites to try to integrate astronomical and cultural evidence.

Two large burial sites, constructed about 3000–2500 B.C., also display winter solstice orientations. At Newgrange in Ireland, a long passage directs sunlight to reach the center of the tomb every morning for about a week around the winter solstice. At Maes Howe in the Scottish Orkney Islands, the setting Sun's rays illuminate the central chamber around the time of the winter solstice.

Hundreds of prehistoric sites in Great Britain and France were carefully surveyed by A. Thom, an engineer. He first found indications of alignments for the solstices and equinoxes, then for the lunar standstills. He promoted the idea that megalithic astronomers made extremely accurate observations of the Moon (using very distant foresights, tens of kilometers long) so as to pick out very small,

long-term variations of the Moon's motions.

The precise lunar observations have been questioned, and different analyses lead to the conclusion that the Moon was observed, but not with the precision inferred by Thom. From the view of cultural necessity, it is unclear how such precision would benefit megalithic people in terms of simple survival value. However, more so than Hawkins's efforts, Thom's work forced archeologists to account for the astronomy in megalithic cultures.

Lacking rich cultural evidence, British investigations turn to the statistical evidence of a large number of sites, selected in an unbiased way, from ones of a similar cultural background. Such work has revealed (mostly for sites in Scotland) strong hints of rough (within a degree or so) orientations to the position of the Moon at major and minor standstills. Hence, general trends show up, even if embedded in a background of noise (that is, orientations that have no relation to astronomy).

Skywatching in the New World

Compared to the Old World, the New World archeoastronomer has the advantage of the survival of remnants of the cultures from pre-Columbian times. Even the great destruction wielded by the Spanish in Mesoamerica, especially their burning of Mayan books that contained much astronomy, could not wipe out completely the astronomy inherent in that culture.

In the Southwest, the Spanish encountered adobe villages, which they called pueblos, of the native peoples who had lived in them at least 1000 years. Many of these pueblos disappeared in historic times (from 1540 onward); those that survived are the cultural connection to the prehistoric people called the Anasazi, who occupied a vast area in the Southwest, centered on the Four Corners area (where New Mexico, Arizona, Utah, and Colorado now meet). Here stand ruins deserted from A.D. 1000 to 1400, stone and adobe constructions that provide some insight into the life of the Anasazi.

Pueblo astronomy. The Hopi (in Arizona) and Zuñi (in New Mexico) pueblos provide the best clues to the past because these villages were touched only lightly by the Spanish. Ethnographers gathered cultural information here around 1900, before the severe pressures on the part of Anglos occurred. It is inferred that the Hopi and Zuñi are cultural descendants of the Anasazi (although it is not known from which specific Anasazi sites). Thus, these pueblos preserve a remarkable cultural connection to prehistory.

Among the Hopi and Zuñi, astronomy played a central role in the agricultural and ceremonial life. The seasonal cycle of the Sun set the ritual calendar and determined the times of specific crop plantings and harvestings. The dry Southwest demands an observant farming, for raising crops is a marginal activity; in the past, failed crops could mean death.

The observing was invested in a religious office, usually called the Sun Priest. He watched daily from a special spot within the pueblo or not far outside it and carefully observed sunrise (occasionally sunset) relative to the horizon features. From past experience, he knew the points that marked the summer and winter solstices and the times to plant crops.

A crucial aspect of the Sun Priest's work was the ability to forecast ceremonial dates; he did so by making anticipatory observations about 2 weeks ahead of time. The rising (or setting) points of the Sun showed a daily change that could be reliably discerned against a horizon profile. By counting down a certain number of days, the Sun Priest could announce ahead of time the day for the ritual, allowing the people of the pueblo enough time to prepare for the ceremony. These forecasting procedures enabled the historic pueblo Sun Priest to predict the actual dates of the solstices with good accuracy, mostly within 1 day of the actual astronomical dates, as given in the historic records of the celebratory dates. The proper choice of ceremonial dates was the major responsibility of the Sun Priest, and it is likely that a prehistoric Sun Priest had the same responsibilities.

Along with horizon features, the Zuñi Sun Watcher, called Pekwin, used a natural pillar to chart the seasons. When the shadow cast by the pillar lined up in a special fashion, Pekwin knew that the summer solstice would soon occur. Also, within the pueblo, special windows and portholes allowed sunlight to hit special plates or markings on the walls at significant times of the year. Thus light and shadows, along with horizon features, made up the basis of the puebloan solar astronomy. The puebloan ancestors are believed to have done much the same as an adaptive survival strategy in a similar cultural context and environment.

Anasazi solar astronomy. Around A.D. 1000, the Anasazi prospered in the San Juan Basin and other regions of the Colorado Plateau. They built community houses that were four or five stories high, contained hundreds of rooms, and many large and small kivas—round, underground rooms used for ritual purposes.

Chaco Canyon in northwestern New Mexico grew to be a center of Anasazi culture. By 1130, eight large villages were located within 9 mi (15 km) of the central canyon. An extensive road system within the San Juan Basin, perhaps a trade network, connected Chaco to many outlying villages. Perhaps a few thousand people lived here in the large and small villages. These Chacoans faced climatic conditions similar to those of today; like the historic pueblos, they probably also had Sun Priests and seasonal solar calendars (and also lunar ones).

Three problems must be kept in mind when evaluating sites in Chaco Canyon. Many Chacoan buildings have been reconstructed, perhaps in ways quite different from the original way, so building orientations, unless based on general alignments of the original foundations, must be taken with caution. Second, few locations within the canyon offer good horizon profiles for a seasonal watch. Finally, it is not known whether each large village had its own, independent Sun Priest (as is true in the historic pueblos) or whether a central religious figure had the authority and responsibility for sun watching for the

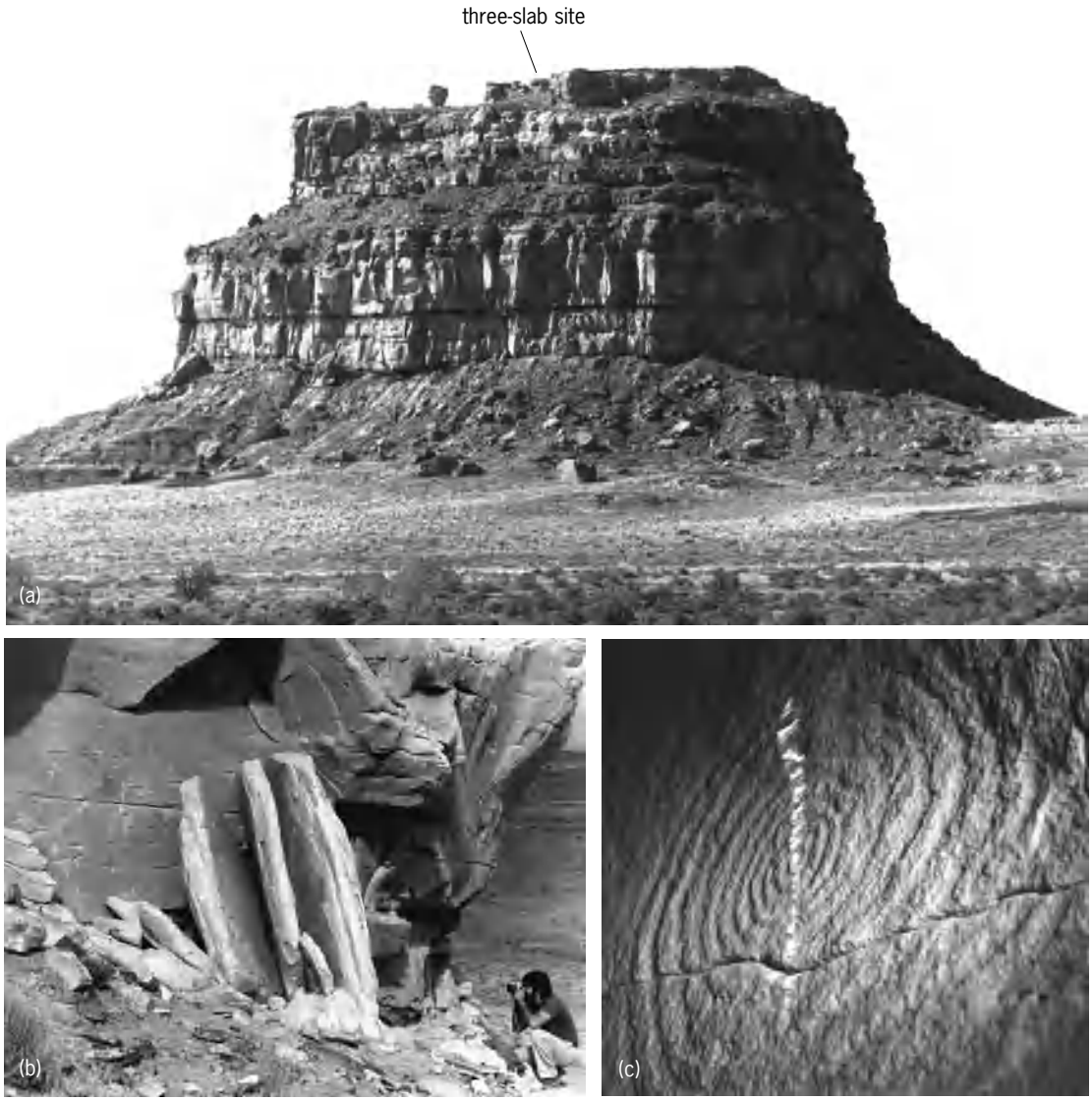

Fig. 5. Fajada Butte in Chaco Canyon. (*a*) View of entire butte (*photo by M. Zeilik*). (*b*) Rock slabs near the top of the butte. They rest against the rock surface on which two spirals are pecked. Their upper edges cause sunlight in late morning to play upon the rock face and around the spirals (*photo by R. Elston*). (*c*) Summer solstice sunlight cutting through the large spiral at about 11:13 a.m. in 1980. Due to a shift of the middle slab, the current view is somewhat different (*photo by W. Wampler*).

all the villages (including those at the outlier sites networked by the road system).

Fajada Butte. This butte (**Fig. 5***a*) thrusts upward at the eastern end of the canyon. Within 30 ft (10 m) of the summit, three rock slabs lie against the butte's southeast face (Fig. 5*b*). The slabs are a few meters long with about 4-in. (10-cm) gaps between them. They shield the rock face on which they rest from the Sun except at times before local solar noon. Then the edges of the slabs allow sunlight to strike the rock face, on which are carved two spirals: a large one (almost 1.5 ft or 0.5 m wide) right behind the slabs, and a smaller one above and to the left of the larger.

The spirals mark the Sun's yearly cycle by light patterns visible late in the morning. On the summer solstice, a shaft of light materializes above the large spiral at about 11 a.m. In about 20 min, it descends and slices through the heart of the spiral design (Fig. 5*c*). On the winter solstice, two shafts of light appear on the outside of the large spiral and pass through its outer edges at about 10 a.m. At both equinoxes, two shafts appear, one shorter and to the left of the other. The shorter shaft cuts through the center of the small spiral, while the larger one drops through one side of the large spiral.

It is possible that the play of light and shadow from a natural rock fall was noted by the Anasazi, and that they then made the spirals on the rock face. Certainly the equinoxes and solstices have distinctive patterns of light and shadow. However, it is unlikely that these patterns were used to forecast, say, the solstices with the 1-day precision achieved by the historic pueblo Sun Priests, since the horizontal motion of the main shaft of light is much too small on an average daily basis to predict the solstices any better than within a week or so. The site might well have served as a sun shrine, a place to which religious officials journeyed at important times of the year to place offerings to

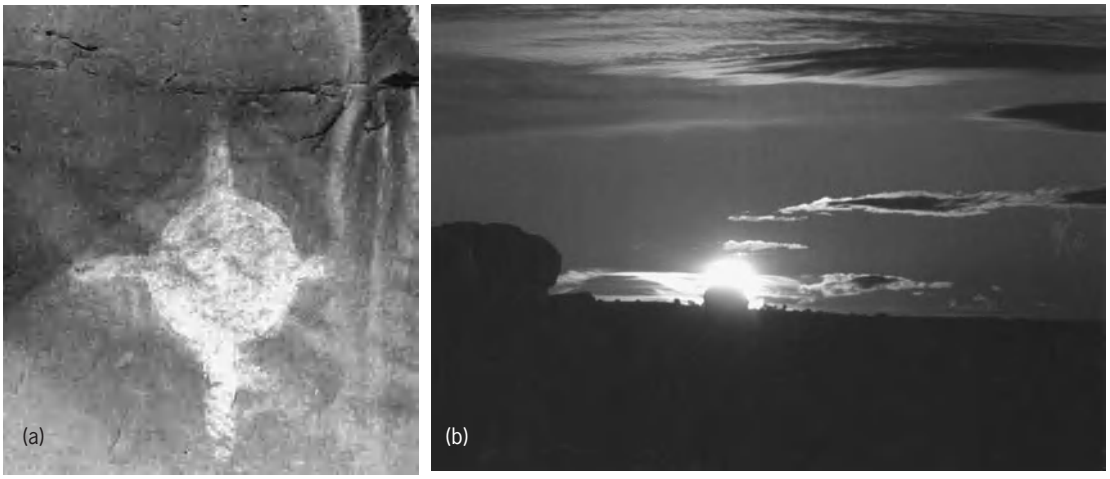

Fig. 6. Wijiji site in Chaco Canyon. (*a*) Painted white sun symbol above the ledge near the ruin. The emblem resembles that of the Zia pueblo sun symbol and a sun symbol on a war shield from Jemez pueblo. (*b*) Winter solstice sunrise observed from the ledge at a position a few meters north of the sun symbol. This sunrise view is the same from the position at the sun symbol 16 days before the winter solstice. The angular width of the rock pillar is a little smaller than the angular diameter of the Sun. (*Photos by M. Zeilik*)

the Sun. In the historic pueblos, sun shrines can be natural rock formations located some distance from the pueblo; a priest would visit them before sunrise.

Unfortunately, it may never be possible to validate these interpretations about the site. The middle rock slab moved in the mid-1980s, and the pattern of the light shafts has changed considerably. Hence, what is now seen is not what was seen before, and may not have been the pattern viewed by the Anasazi if erosion has shifted the slabs in the past 1000 years.

Wijiji. This is a pueblo built about A.D. 1100, late in Chaco's history. About 0.6 km (1 mi) to the east of it, a large rincon (valley) opens up in the mesa. On the northwest side of the rincon runs a narrow ledge, which can be reached by climbing a prehistoric staircase. Here on the wall is painted a large four-pointed symbol (**Fig. 6***a*) that resembles the Zia Pueblo sun sign.

North of the symbol, three boulders rest on the ledge; the largest has a double spiral carved in its surface. The design and technique are clearly Anasazi. Eastward from the ledge, a large rock pillar, across the rincon, rises above the horizon. From a spot a few meters south of the boulder with the double spiral, the winter solstice Sun rises behind the pillar (Fig. 6*b*). Because the angular width of the pillar against the sky is somewhat smaller than the angular diameter of the Sun, the winter solstice sunrise behind the pillar is a clear event when viewed from the location near the boulders. If an observer moves a few meters (yards) to either side, the shift moves the Sun to either side of the pillar. Near the sun symbol the Sun rises behind the pillar 16 days prior to the winter solstice, making possible an anticipatory observation that works extremely well to forecast the day of the winter solstice.

The area around the ledge contains much rock art, some of it Navajo, and it has been argued that the white "sun" symbol is Navajo in origin and that the site was used by Navajo (perhaps in the late seventeenth century) for sun watching. That may be the case, but there are certainly Anasazi relics here, too. Thus the site may have been first used by the Anasazi and then adopted by the Navajo for ceremonial purposes.

Anasazi lunar astronomy. There is no firm evidence so far of prehistoric monthly lunar calendars among the Anasazi. However, occasional claims have been made that the Anasazi noted the 18.6-year standstill cycle of the Moon. The most convincing case for attention to this interval has been made for the Chimney Rock Archaeological Area in Colorado, where the natural rock pillars act as a natural foresight when viewed from the site of Chimney Rock Pueblo, a Chacoan outlier, which was occupied from about 1075 to 1175. The orientation allows an anticipation and forecasting of the major lunar standstills. For about 2.5 years prior to the standstill, the Moon rises between the pillars for 1 or 2 days per month. Starting after the summer solstice, the Moon appears as a waxing crescent. Finally, near the winter solstice, the full moon stands between the pillars at around sunset. Hence, the priest in charge of the moon watching could forecast the date of the standstill and also the full moon nearest the winter solstice, an important conjunction among the historic pueblos.

Regional ceremonialism. Many models have been proposed for the development of Chaco Canyon as the regional center of the San Juan Basin. The "economic redistribution" model is fairly plausible. Here the roads serve to direct people to Chaco to trade goods, which they take back to their home pueblos.

Astronomical scheduling may have been used to tell people when to commence their travels. In a decentralized social system, each outlying pueblo might have its own Sun Priest tracking time locally. He would anticipate important ceremonial dates, such as the winter solstice, so that the people of the pueblo could prepare for a pilgrimage to Chaco Canyon. Other Sun Priests would do the same so that the arrivals at Chaco would be synchronized by the sunwatching.

The concept of regional ceremonialism scheduled by astronomy may serve as a synthesizing concept for other prehistoric sites. For instance, there is increasing evidence that Stonehenge played a central role in a network of ritual sites with the winter solstice ceremony as a cosmic organizer in sacred space and time.

Other Places and Cultures

Every prehistoric culture appears to have developed its own astronomy. The traditional navigators of Oceania needed to memorize guide stars and to employ them as the bearing markers for island-to-island travel over thousands of kilometers of water. The Carib people of northern South American developed a calendar that relied on the positions of stars relative to each other and to the Sun at times of rising and setting. A bone from the shores of Lake Edward, Zaire and Uganda, may have markings of a lunar calendar, tallied at a time over 8000 years ago, perhaps used to forecast marine activity or the weather; a focus on the Moon continues in Africa today. Chinese astronomical records cut into bones and shells may have begun as early as the twelfth century B.C., well before the Babylonians incised their earliest records on clay.

By the time the Spanish invaded Mesoamerica and South America, the use of astronomy went well beyond complex systems of cycles and calendars. For instance, in some Mayan cities knowledge of the cycles of Venus timed the onset of warfare. In others, key political events incorporated the summer solstice and perhaps conjunctions of Jupiter and Saturn. Ancient Inca city planning and politics embodied astronomy, such as the ceque system of radial lines from the Temple of the Sun in the valley of Cuzco. These lines mark the directions to sacred places as well as to specific astronomical phenomena. The ceque system had a calendric manifestation in knotted cords that tallied the days of the agricultural year.

Michael Zeilik

Bibliography. A. F. Aveni, *Ancient Astronomers*, St. Remy Press/Smithsonian Books, 1993; J. M. Malville and C. Putnam, *Prehistoric Astronomy in the Southwest*, rev. ed., Johnson Books, 1993; J. North, *Stonehenge: A New Interpretation of Prehistoric Man and the Cosmos*, Free Press, 1996; C. L. N. Ruggles and N. Saunders (eds.), *Astronomies and Cultures*, University of Colorado Press, 1993; R. A. Williamson and C. R. Farrer, *Earth & Sky: Visions of the Cosmos in Native American Folklore*, University of New Mexico Press, 1992.

Archeological chemistry

The application of chemical techniques to the study of archeological finds, natural or anthropogenic, in order to ascertain their composition or, in some cases, their age. Traditional chemical analysis uses wet methods, in which a sample is brought into solution and its components are assayed by precipitation or titration. These methods were applied to ancient coins as early as the late eighteenth centuary. The obvious need to minimize damage to an irreplaceable object spurred the development of microchemical techniques. Modern analysis relies on instrumental methods that require only very small samples or are entirely nondestructive. Although these methods rely on physical phenomena rather than chemical transformation, all procedures that are capable of the qualitative and quantitative determination of the atomic or molecular composition of the object under study are usually included under the broad heading of archeological chemistry.

Methods. Various analytical methods are utilized in archeological chemistry, as described below. It should be noted that these methods of analysis are not competing but complementary. The choice of method depends on the nature of the object, on the elements to be determined, and on the accuracy required.

Optical emission spectography. The oldest instrumental method of analysis is optical emission spectrography, in which thermally excited atoms emit light at wavelengths corresponding to energy differences between electronic orbitals. A set of wavelengths identifies a given atom; the intensity of the emitted light, measured photographically or by means of a photomultiplier tube, is proportional to the concentration of that atom. Optical emission spectrography has been largely replaced by more sensitive methods covering a wider range of elements.

Atomic absorption spectroscopy. In this method the atoms in a sample flame absorb part of the energy of a reference light source containing the same atom. Accuracies range from $\pm 1\%$ for major components to $\pm 15\%$ for trace elements in the parts-per-million range.

Inductively coupled plasma. In this modern version of emission spectroscopy, an inductively coupled plasma is the means of excitation, attaining temperatures in excess of 8,000–10,000°C (14,400–18,000°F).

Neutron activation analysis. This method involves the bombardment of the sample with slow neutrons in a nuclear reactor. Some atoms are converted to radioactive isotopes whose subsequent decay produces gamma rays that are detected by a semiconductor counter coupled with a multichannel pulse analyzer to yield both qualitative and quantitative data. Small objects can be irradiated in their entirety, obviating sample removal and making the method nondestructive. Variations include instrumental neutron activation analysis (INNA), proton-induced gamma-ray emission (PIGE), and prompt gamma neutron activation analysis (PGNAA), in which the gamma rays of short-lived radioisotopes are analyzed while the sample absorbs neutrons.

X-ray fluorescence spectrometry. The sample, which may be the entire object, is exposed to x-rays that excite electrons in the lower shells to higher energy levels. When these electrons drop back to lower levels, secondary or fluorescent x-rays are emitted and analyzed for energy and intensity, thus providing identification of the elements present as well as their

concentrations. X-ray fluorescence offers accuracy of ±2–5% of concentrations from 10 ppm to 100%, but since x-rays penetrate only to depths of 20–200 micrometers, only the surface of an artifact will be analyzed.

Electron microprobe analysis and proton-induced x-ray emission. The related techniques of electron microprobe analysis, which uses high-energy electrons for excitation, and proton-induced x-ray emission (PIXE), which uses protons, permit focusing the incident beam on areas of 1 μm^2 and 1 mm^2, respectively. This not only makes it possible to analyze small regions of heterogeneous surfaces, like those of painted pottery, but also allows scans to be made across the surface of a cross section to yield compositions as a function of depth, although only at the expense of cutting through the object. Both methods can be applied to bulk analysis by grinding up a sample and converting it to a glass bead or pressed pellet.

Auger electron spectroscopy. This method is similar to the electron microprobe technique, but measures electrons ejected by the absorption of x-rays in the matrix.

X-ray photoelectron spectroscopy. This method also ejects electrons by irradiation with x-rays and has the advantage of identifying the valence state of atoms. Photoelectron spectroscopy bridges the gap between elemental analysis and the determination of the valence state of ions and of the identification of molecular species, as does Mössbauer spectroscopy, which specifically reveals the oxidation state of iron. *See* ACTIVATION ANALYSIS; ATOMIC SPECTROMETRY; AUGER EFFECT; MÖSSBAUER EFFECT; PROTON-INDUCED X-RAY EMISSION (PIXE); RADIOISOTOPE; X-RAY FLUORESCENCE ANALYSIS.

Organic archeometry. Organic materials constitute only a small portion of archeological finds, but since they include such basic necessities as food, drink, and clothing, they have the potential of revealing much about past life. Because they consist of covalently bound, complex, and sensitive molecules, their study requires special methods of analysis. Organic archeometry is the newest and most rapidly expanding field of archeological chemistry. Organic dyes have long been determined qualitatively and quantitatively by absorption spectroscopy in the visible and ultraviolet ranges. The extension into the infrared range allows not only the identification of organic materials by visual or computer-aided comparison of infrared spectra ("fingerprinting") but also some structural interpretation. Since organic residues typically consist of mixtures of dozens or even hundreds of individual compounds, the progress of organic archeometry has crucially depended on the development of chromatographic separation procedures. These include column chromatography, paper and thin-layer chromatography, gas chromatography, with or without prior pyrolysis, and liquid chromatography.

All of these techniques not only separate mixtures into individual components but permit their identification if the rate at which they travel through the chromatographic substrate, the retention time, can be matched to those of authentic reference compounds. The most powerful techniques combine a separation step with a spectroscopic identification method.

The detection of components in liquid chromatography is commonly made by ultraviolet spectroscopy, but it can now also be coupled with mass spectrometry. Gas chromatography has been coupled with infrared spectroscopy and, most powerfully, with mass spectrometry. In mass spectroscopy, individual components are converted to usually positively charged ions by electron impact or by chemical ionization. These ions are separated by electrical and magnetic fields and recorded separately. The resulting mass spectrum can identify even very complex molecules with virtual certainty. This has made possible the identification of food remains in vessels, house floors, and refuse middens.

An important adjunct to the study of prehistoric diets is the determination of stable isotopes of carbon and nitrogen by mass spectrometry, which has been applied to collagen extracted from skeletal remains. A recent advance is a single instrument that first separates organic components by gas chromatography, then converts them singly into carbon dioxide by combustion, and lastly determines the isotopic composition of each component by mass spectrometry.

Another method that is gaining use in organic archeometry is nuclear magnetic resonance spectrometry (NMR), which detects a limited number of atomic nuclei, among them ordinary hydrogen, carbon-13, nitrogen-15, fluorine-19, and phosphorus-31, by their simultaneous interaction with an external magnetic field and a radio-frequency field. The sample requirements are too high to apply this technique to the small amounts of compounds emerging from a capillary gas chromatograph (although a preparative gas chromatograph would make this possible), and the archeological applications of both proton magnetic resonance (PMR) and carbon magnetic resonance (CMR) have been to unseparated mixtures, including fossil resins and tars. *See* GAS CHROMATOGRAPHY; INFRARED SPECTROSCOPY; MASS SPECTROMETRY; NUCLEAR MAGNETIC RESONANCE (NMR); SPECTROSCOPY.

Identification. The determination of the chemical composition of an archeological find is not an end in itself, but provides the archeologist with factual evidence not otherwise obtainable and touching on many aspects of early human life. The changing elemental composition of coins detects progressive debasement and reveals economic history and fiscal policy. The metals added to copper to make bronze and brass outline the history and spread of technology in general and of metallurgy in particular. The foodstuffs consumed are indicators of the advent and progress of agriculture and animal husbandry. Together, all these paint a picture of prehistoric social, cultural, and economic stratification. *See* PREHISTORIC TECHNOLOGY.

Provenience analysis. The chemical analysis of archeological artifacts has firmly established that the

composition of an object offers clues to its geographic origin, which may be far from the excavation site, as had been evident in some cases from purely typological considerations. This provides evidence of trade and exchange in commodities and raw materials. The elemental or molecular makeup of natural products differs from one deposit to another. Matching the composition of artifacts to that of all possible sources reveals extensive trade over impressive distances. For obsidian, a natural volcanic glass, this has been done by elemental analysis; for amber, a fossil resin, by infrared and carbon resonance spectrometry. The composition of manufactured objects reflects their source less directly because the transformation of clay into ceramics or of ore into metal may cause changes in composition. In addition, there is the problem that deposits of clay or ore may have been exhausted in antiquity and therefore are no longer available for comparison. Provenience analysis of metals is especially difficult because the reuse of broken implements from various sources, possibly together with new ore, will lead to the superimposition of a number of compositional indicators. Nevertheless, the use of sophisticated statistical procedures, notably cluster analysis, has made it possible to distinguish local ceramic and metal products from imports and, in many cases, to identify the sources of the latter.

Chemical dating. While the most widely used methods for dating archeological material—radioactive decay, thermoluminescence, and archeomagnetism—deal with physical processes, three depend on the progress of conventional chemical reactions. (1) Amino acid dating uses the rate of racemization of optically active organic molecules. (2) Hydration dating measures the thickness of the weathering layer produced by the action of water on natural and artificial glass, including obsidian and flint. (3) NFU dating of bone relies on the loss of nitrogen (N) from the organic collagen component and on the uptake of fluorine (F) and uranium (U) by the inorganic hydroxyapatite component. Like all nonnuclear chemical reactions, these changes are a function not only of time but also of temperature, of acidity and, in the case of fluorine and uranium uptake, of the concentrations of these elements in the surrounding soil. Chemical methods cannot produce absolute dates unless these other variables are known or can be estimated reasonably closely. They are, however, useful in establishing relative ages of finds within a single site in which the depositional characteristics are likely to have been uniform. *See* AMINO ACID DATING; ARCHEOLOGICAL CHRONOLOGY; CHEMICAL MICROSCOPY; DATING METHODS; PALEOMAGNETISM; RACEMIZATION; RADIOCARBON DATING; THERMOLUMINESCENCE. Curt W. Beck

Bibliography. J. B. Lambert, *Traces of the Past: Unraveling the Secrets of Archaeology through Chemistry*, 1997; A. M. Pollard (ed.), *New Developments in Archaeological Science*, 1992; A. M. Pollard and C. Heron, *Archaeological Chemistry*, 1996; S. U. Wisseman and W. S. Williams (eds.), *Ancient Technologies and Archaeological Materials*, 1994.

Archeological chronology

The establishment of the temporal sequence of prehistoric events. During the nineteenth and early twentieth centuries archeologists devised a number of methods to construct a relative chronology for cultural remains. These included stratigraphy (correlation of layered sediments and rocks), seriation (ordering of materials based on changes in style), and cross-dating (using temporally diagnostic ceramic and stone tool types). Beginning in the early twentieth century, dating methods developed in the natural sciences began to be applied to archeological remains to provide absolute (that is, calendar) dates. These are sometimes called chronometric dating methods, and the more commonly used ones are described here. *See* ARCHEOLOGY; DATING METHODS.

Radioactive decay methods. The decay of a radioactive material is monitored by its half-life, the amount of time it takes for half of a given amount to decay. Because half-lives for any given radioactive isotope are nearly constant, this information can be used to measure time. *See* HALF-LIFE.

Radiocarbon dating. Radiocarbon dating is the most commonly used chronometric method in archeology. The radiocarbon isotope carbon-14 is produced in the atmosphere by cosmic-ray bombardment of nitrogen-14. Carbon-14 is quickly spread throughout the biosphere beginning with photosynthesis and dissolution in the ocean and other water bodies, so that plants and animals have roughly the same concentration of carbon-14 as the atmosphere. When a plant or animal dies, it ceases to interact with the carbon cycle, and the amount of carbon-14 diminishes according to its half-life of about 5730 years. By measuring the current decay rate (conventionally by counting the number of beta particles detected over a given time) or the ratio of carbon-14 to stable carbon isotopes (by accelerator mass spectrometry, AMS), the time since the organism died can be derived. Any organic remains found at archeological sites (such as bones, shells, and preserved plant materials) can therefore potentially be dated, and AMS allows dating of very small samples. Radiocarbon dates are given in terms of radiocarbon years, which must be converted to calendar years, because the rate of cosmic-ray bombardment has varied through time. Correlations have been established between carbon-14 and tree-ring dating (dendrochronology), which are used for calibration back to about 10,000 years; correlations based on uranium series dating on corals are used for calibration back to about 30,000 years. The upper limit of radiocarbon dating is about 50,000 years. *See* BIOGEOCHEMISTRY; RADIOCARBON DATING.

Uranium-series dating. Uranium-series dating involves the use of uranium-238 (and, less commonly, uranium-235), which has a very long half-life and decays through a succession of "daughter" radioisotopes, each with a much shorter half-life. The decay series ends with lead-206, which is stable. Over a long enough time, the radioactivity of the daughters

becomes the same as that of the parent uranium, a situation called secular equilibrium. However, uranium can be soluble in water, whereas most of the daughter radioisotopes cannot. Water-precipitated materials (such as stalagmites, stalactites, and travertines) or organisms such as coral, which derive carbonate material from the water, initially contain only uranium without the daughters. But after deposition, decay of the uranium leads to the buildup of the daughters until equilibrium is reestablished. This provides a clock, extending back about 500,000 years, to measure when precipitation occurred. Due to this long dating range and because much of the early archeological record is contained in limestone caves and rock shelters, uranium-series dating has been very important in dating the Paleolithic period. *See* LEAD ISOTOPES (GEOCHEMISTRY); RADIOISOTOPE (GEOCHEMISTRY).

Potassium-argon dating. Potassium-argon dating is used to date volcanic events. The radioactive isotope potassium-40, which decays to argon-40, is commonly found in volcanic rocks. During a volcanic eruption, the argon gas escapes. After cooling, the decay of potassium-40 leads to a new buildup of argon in the rocks. The ratio of potassium-40 to argon-40 determines the time since the eruption. Traditionally, potassium and argon were measured separately, but a more recent technique is to bombard the sample with neutrons, which converts potassium-40 to argon-39. Then both argon-40 and argon-39 can be measured simultaneously by mass spectrometry. This technique, known as the Ar-Ar method, provides more accurate and precise results. Although potassium-argon dating is restricted to dating volcanic events, it has been very important in dating early human evolution in East Africa, where cultural materials are often separated by volcanic ash layers. *See* GEOLOGIC TIME SCALE; ROCK AGE DETERMINATION.

Radiation-exposure methods. In some materials natural radioactivity produces free electrons that can be subsequently trapped in crystalline defects. The trapped charge accumulates at a rate governed by the natural radioactive dose rate, which is usually fairly constant. This accumulation, calibrated by laboratory irradiation as an equivalent dose, is proportional to the time since the traps were last emptied. Dividing by the dose rate yields an age.

Luminescence dating. Luminescence dating uses heat (thermoluminescence, TL) or light (optically stimulated luminescence, OSL) to empty the electron traps, resulting in the emission of light, or luminescence. The intensity of the luminescence signal is proportional to the time since the sample was last heated or exposed to light. It thus provides a way to date pottery, stone tools that have been accidentally or deliberately heated, or buried sediments. Quartz and feldspar are the two most common minerals dated, and their ubiquity at archeological sites, either in pottery or stone tools or in archeological sediments, gives luminescence dating wide applicability. Recent research on OSL, in particular, has greatly increased both the accuracy and precision of the method, and dates of up to 100,000 years and more are routinely obtained. OSL dating of single grains of quartz or feldspar is the newest development, opening up the possibility of using luminescence to study site formation processes. *See* THERMOLUMINESCENCE.

Electron spin resonance. Electron spin resonance measures the trapped electron concentration directly by inducing a change in the quantum spin of the electrons via exposure to microwave radiation in the presence of a magnetic field. This method has been commonly used to date mammalian tooth enamel (to the time of its formation), as well as shells, coral, and quartz (to time of last heating or exposure to light). It is mainly used for dating older materials, and has an upper age limit of about 2 million years, giving it a prominent role in understanding the chronology of early human evolution. *See* SPIN LABEL.

Climatic variation methods. Variations in climatic parameters such as temperature and precipitation are often recorded in biological structures and sediment accumulations. Use of pollen compositions, lake sediments, and loess deposits to track temporal climatic trends has a long history.

Dendrochronology. Dendrochronology uses climatic-induced variations in annual tree rings to produce single-year resolution dating. Each ring consists of an inner band of light-colored wood and an outer band of dark-colored wood, reflecting different parts of the growing season. The width and density of the rings in some trees is strongly influenced by temperature and precipitation, and the resulting pattern of changing thicknesses and densities is reproduced by all trees of the same species in a given region. This allows cross-referencing of the patterns from living trees to ancient wood remains until a master sequence is developed against which wood from archeological sites can be referenced and thus dated. The advantages of high resolution are countered by the disadvantage of limited applicability. Although dendrochronology has been applied all over the world, the most developed sequences are from conifers in the American southwest and oaks in northern Europe, where master sequences in both places extend more than 10,000 years. *See* DENDROCHRONOLOGY.

Oxygen isotope ratio dating. Oxygen isotope ratios recorded in deep-sea and ice cores also provide high-resolution dating. Oxygen has three isotopes, the lightest and most common being oxygen-16 and the heaviest being oxygen-18. Vapor (and thus rain and snow) above the ocean has a higher ratio of ^{16}O to ^{18}O than the ocean. During glacial periods, the lighter snow gets locked up in glaciers, so that the ocean is isotopically heavier than during nonglacial periods. These ratios are preserved in the tiny shells of foraminifera, which accumulate on the ocean floor. Cores of ocean sediments thus contain a record of past climate. Similar records are available from ice cores (with the ratios calculated on the ice itself) from polar regions. The oxygen isotope records provide only relative dating based on stratigraphy. Attempts to anchor the sequence in absolute time have

most commonly used paleomagnetism, carbon-14, and uranium-series dating. Precise dates for the isotope sequence have also been obtained by alignment with astronomical data, on the justification that ice-age cycles are caused by minor periodic variations in the Earth's orbit. Application to archeological dating requires correlating the ocean data with shellfish isotope ratios or with terrestrial climatic records, which can be difficult.

Weathering methods. Archaeologists have put considerable effort into trying to date materials by their degree of weathering. Thickness of patinas (thin, colored films) on rocks, degradation of carbon, uptake of fluorine in bones, and the extent of lichen growth on rocks have all be studied.

Obsidian hydration dating. Obsidian hydration dating is the most commonly used weathering dating method, because obsidian (volcanic rock that has cooled too rapidly to crystallize) was often used to make stone tools in prehistoric times (see **illustration**). Water diffuses into the surface of freshly fractured obsidian and forms a hydration rind. The rind thickness increases with time, but it is also a function of temperature, composition, burial depth, and other factors, all of which have to be controlled for accurate dating. Experts in the field disagree on how the rind should be measured and on the proper equation for describing the diffusion process. Most success has been achieved by comparing rind thicknesses with known-age obsidians from nearby sites or by just dating on a relative scale. *See* HYDRATION; OBSIDIAN.

Amino acid racemization. Amino acid racemization refers to the conversion of the chemical structure of amino acids, the building blocks of proteins, from "left-hand" arrangements to "right-hand" arrangements after an organism dies. The conversion rate toward an equilibrium level, at which there are equal numbers of both arrangements, is a function of time but also depends on climatic variables, particularly temperature. Early applications to bone were not successful, but recent work with chemically stable egg shells from large flightless birds, such as ostriches and emus, has shown more promise. These egg shells are common in archeological sites in southern Africa and Australia, so most work is now being done in those two areas. *See* AMINO ACIDS; RACEMIZATION.

Archaeomagnetism. In a class of its own, archaeomagnetism is based on temporal changes in the location of the Earth's magnetic pole and in the intensity of the Earth's magnetic field. When a sample is heated to a high temperature, magnetic dipoles within iron-bearing inclusions align themselves with the Earth's magnetic field. When the sample cools, the alignments are frozen in that position. Dating requires the construction of a master curve on which known past positions and intensities are plotted. The alignment recorded from an archeological sample, such as a hearth or a kiln, can then be matched with the master curve and thus dated. Dating is limited to areas for which such curves have been constructed and rarely extends further back than 10,000 years. The Earth's magnetic poles have also reversed from north to south and back several times through Earth's history, and these reversals have been recorded and dated (by potassium-argon dating) to allow coarse resolution dating on a geological scale. *See* PALEOMAGNETISM.

James K. Feathers

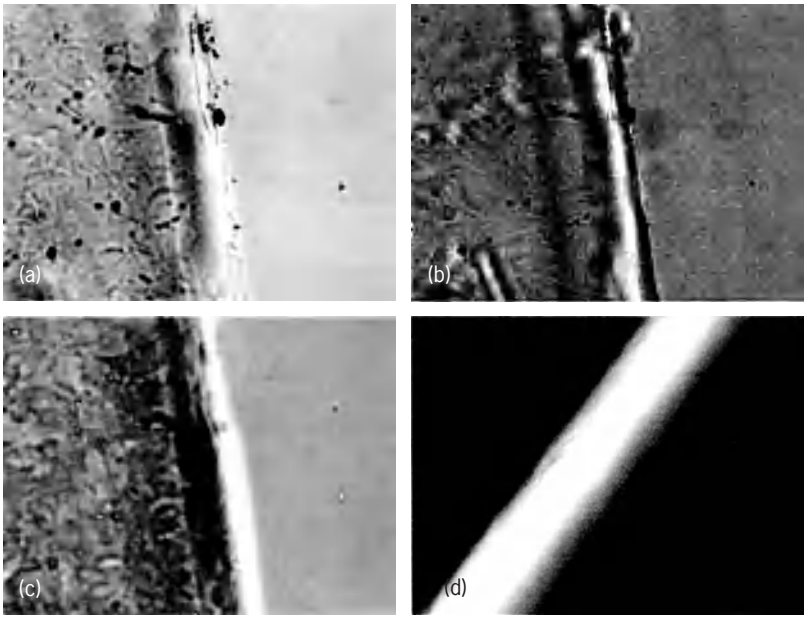

Photomicrographs of a hydration rim of a 150,000-year-old obsidian tool from Ethiopia taken using different types of imaging. Each image is 0.125 mm wide. (*a*) Using conventional light at optimum focus, the rim appears bounded by bright lines, demarcating it from the obsidian to the left and the mounting medium to the right. The bright lines are probably functions of varied refraction indices among the unhydrated obsidian, the hydration rim, and the mounting medium. (*b*) Reduction of the light enhances the boundary between the rim and the mounting medium. (*c*) Bringing the microscope out of focus moves both lines to the right, which is expected if the lines are refraction artifacts. (*d*) In cross-polarized light, the boundaries are shown to be gradational rather than sharp, in contrast to earlier assumptions. (*Reprinted from L. M. Anovitz et al., The failure of obsidian hydration dating: Sources, implications and new directions, J. Archeolog. Sci., 26:735–752, ©1999, with permission from Elsevier*)

Bibliography. M. J. Aitken, *An Introduction to Optical Dating*, Oxford University Press, Oxford, 1998; M. G. L. Baillie, *A Slice through Time*, B. T. Batsford, London, 1995; S. Bowman, *Radiocarbon Dating*, British Museum Publications, London, 1990; M. J. O'Brien and R. L. Lyman, *Seriation, Stratigraphy and Index Fossils: The Backbone of Archaeological Dating*, Kluwer Academic, New York, 1990; R. E. Taylor and M. J. Aitken (eds.), *Chronometric Dating in Archaeology*, Plenum Press, New York, 1997; G. A. Wagner, *Age Determination of Young Rocks and Artifacts: Physical and Chemical Clocks in Quaternary Geology and Archaeology* (translated from German), Springer, Berlin, 1998.

Archeology

The scientific study of past material culture. The initial objective of archeology is to construct cultural chronologies, attempting to order past material culture into meaningful temporal segments. The intermediate objective is to breathe life into these chronologies by reconstructing past lifeways. The ultimate objective of contemporary archeology is

to determine the cultural processes that underlie human behavior, both past and present.

Thus, archeology is both scientific and humanistic. Throughout many parts of the world, archeology is considered a subdiscipline of anthropology, focusing on the anthropology of extinct cultures. In other parts of the world, archeology is regarded as an extension of history, attempting to write a prehistory of people who lack an adequate history of their own.

The material culture of the past is of infinite variety. It is the scientific study of this evidence which differentiates the contemporary archeologist from the nineteenth-century antiquarian. So broad is the task that there is no such thing as any single archeological method, although over the past century archeologists have evolved what can be termed an overall archeological approach.

Methods. By constant confirmation, the archeologist often attempts to establish synchronism with what has already been established historically. In 1945, for instance, Mortimer Wheeler found at Pondicherry pottery shards of well-dated Roman types in close association with certain East Indian wares. Subsequently, he was able to establish a date for the unknown Indian ceramics on the basis of the historically dated Roman ceramics.

Similarly, scarabs exported and traded from the Nile Valley into Palestine, the Sudan, and further afield have been dated with reference to the Egyptian king lists, providing at least approximate dates hundreds of miles away from the place of manufacture. Even when dates in years are less readily available, some kind of relative synchronization may be attempted, such as when similar bronze axes or sword hilts were discovered with locally produced objects in Sweden, Britain, and Spain, or when daggers of Mycenaean type were found in Wiltshire, England.

Types. Archeologists use a number of types in order to categorize similar artifacts. Most common is the temporal type, a principle similar to the index fossil concept used by the geologist. A temporal type can be any kind of archeological artifact or feature, but ideally it is some object of common use in which the form is subject to change, due to either the whim of fashion or technological improvement. One example is the simple flint arrowhead with side barbs and central tang. It is typical of the British Bronze Age and was not in fashion earlier or later. Ceramic types have been established by archeologists working around the world, and a thoroughly tested ceramic chronology is invaluable as a temporal ordering device, no matter where the archeologist is working. *See* INDEX FOSSIL.

The nature of the artifact employed as a temporal type is irrelevant, and its use may not even be known. In the future, for instance, the metal springs of modern clothespins remaining after the wooden parts have decayed may serve, for the archeologist, as a means to define periods and cultural associations, even though the original use may be lost in the past.

Archeologists also establish other types. Functional types attempt to group artifacts on the basis of known or presumed functions. Function may be established through experimentation, through analogy to existing primitive groups, or through microscopic analysis of damage to the tool itself. Archeologists also establish technological types, divisions which reflect the mode of manufacture. Technological types are particularly helpful when studying stone-tool manufacture.

Cultures. The concept of culture is used in two different ways by contemporary archeologists. When dealing with cultural chronologies, the archeologist most commonly uses a modal or shared view of culture. It is this normative collection of shared ideas which cause artifacts to change in systematic ways through time, and temporal types can be established on the basis of this shared culture. When attempting to reconstruct lifeways, however, the archeologist can no longer rely on the shared aspects of culture. When transcending temporal associations, contemporary archeologists tend to view culture systematically, as people's extrasomatic (that is, learned) method of dealing with the social and cultural environment. In this sense, one does not share a cultural system—one participates in it. Participation is controlled by cultural roles, often expressed through intricate sets of status relationships.

As worldwide cultural sequences become better known, archeologists are able to abandon the modal concept of culture in favor of the more satisfactory, if more complex, systemic view of cultural phenomena.

Law of superposition. The law of superposition was formulated initially by Nicolaus Steno (1638–1687). Steno's law, simply stated, says that in any pile of sedimentary rocks that have not been disturbed by folding or overturning, the strata on the bottom were deposited first. The law of superposition thus holds that, all else being equal, older deposits will tend to be buried beneath younger ones. Thomas Jefferson is generally credited with being the first archeologist to have applied systematic principles of stratigraphy to archeology.

Mere stratigraphic equivalence, however, does not necessarily indicate contemporaneity, as there can be misleading mixtures of successive occupational debris on one surface. In sandy areas, successive occupations may have been covered by layers of blown sand. Where the covering layers are intact, no better sealing and separation of occupational surfaces can be found. If, however, wind activity is reversed, and sand is removed rather than deposited, two or more distinct layers of human occupation may become falsely associated. Archeologists must therefore study the processes of cultural deposition in order to recognize the difference between intact and disturbed strata.

Excavation. Contemporary excavation must be conducted with a plan, a firm research design that attempts to provide answers to definite questions. Archeology is one of the few sciences which destroys its own data in the process of generating them. Archeologists must therefore be extremely careful to make the appropriate observations at the time of excavation. If problems are not in mind, delicate

archeological associations might be overlooked. Once a site is excavated, only the data from that excavation are available; the site itself has actually been destroyed.

Archeological technique attempts to record and clarify such relationships. The chronological relationship, for instance, between two structures lying side by side may be of interest; one structure can be established as later if it overlies the other. On a larger scale, one may wish to ascertain the duration, successive extents, economies, and purposes of occupation of a given site or area.

Contemporary archeologists are faced with the unprecedented destruction of sites because of deep plowing, quarrying, construction, and leveling for parking lots and airports. In Europe such excavation has been termed rescue archeology; in the New World the unwieldy term cultural resource management has been applied to the effort to preserve, or at least mitigate, the impact upon the survival of the archeological site.

Relation to other disciplines. The task of deciphering meaning from past material culture is so complex that the archeologist is often required to borrow from allied disciplines in the physical and natural sciences. Only a few can be mentioned.

Geology. Geological studies provide information on the natural barriers and obvious routes and the source of ores which people use. Geomorphology provides information regarding changes in sea level and the extent of glacial ice sheets and loess deposits. It serves as a basis for assigning relative dates to archeological discoveries, particularly of Paleolithic stone implements. *See* GLACIAL EPOCH; PLEISTOCENE; STRAND LINE; VARVE.

Climatology and paleobotany. The analysis of ancient plant pollen and spores—palynology—and information regarding former climates has recently become one of archeology's best methods for examining prehistoric ecological adaptations. Dates also have been determined by studying annual growth in trees. The identification of cereal grains in the surface of ancient pottery has helped in piecing together the story of the beginnings of agriculture. *See* DENDROCHRONOLOGY; PALYNOLOGY; POSTGLACIAL VEGETATION AND CLIMATE.

Paleontology. The study of fossil remains helps the archeologist to identify finds of bone, antler, and ivory and to reconstruct people's natural environment in successive periods. Paleontological studies also shed light upon the history of the domestication of animals and the development of humans' own bodily structure. *See* FOSSIL HUMANS; PALEONTOLOGY.

Land snails often have restricted habitats; a study of the shells of these creatures taken, for example, from different levels in the material which has accumulated at different times in a ditch made by humans may show that when the ditch was first excavated there were large trees growing in the immediate neighborhood. Later, pottery may have found its way into the half-filled ditch and, with it, the shells of snails typical of open grassland, therefore indicating that by the time the pottery was being made, the area had been cleared of trees.

Along with the pottery and grassland snail shells, there may be found the wing cases and other remains of beetles. The restricted habitat of beetles serves the same purpose as that of snails as far as environmental studies are concerned. There have been cases where the remains of a particular kind of dung beetle have shown that cattle were present at the time a certain deposit was made. In one case, the finding of the wings of certain ants showed that a mound had been built up at a particular time of the year, since these ants are known to fly only during their mating season.

Mineralogy. The mineralogist has been called upon to compare and identify the sources of metal ores and materials used in the manufacture of implements and to name the origin of stones. Without this information archeologists would not have known that the blue stones of Stonehenge were quarried in South Wales, nor would archeologists have found the sites in North Wales and Westmoreland where stone axes discovered in southern England were made.

Physics. The science of physics has provided the archeologist with a mean of dating finds in years, for it has made available the study of the rates of decay of radioactive substance. The study of paleomagnetism and the orientation of magnetic particles in strongly heated structures and pottery may eventually provide a means of dating archeological finds. By using the proton magnetometer, which measures minute differences in magnetic fields, the archeologist is often able to confirm the suggested existence of filled-in pits and ditches of which no surface evidence appears.

Other techniques include the use of air photographs to map known sites and to discover new ones; the use of special underground photographic apparatus, such as photographic drills, to penetrate into the earth and to record the interior of buried chambers; and the use of electrical instruments which record differences in electrical potential and thus detect irregularities below the surface. *See* AERIAL PHOTOGRAPH; RADIOCARBON DATING.

Chemistry. The study of chemistry provides knowledge of natural processes in the soil and helps in the identification of invisible evidences of occupation, such as decayed layers of buried turf. Certain chemical techniques also make possible the preservation and reconstruction of archeological finds for study and exhibition. *See* AMINO ACID DATING; ARCHEOLOGICAL CHEMISTRY.

Synthesis. The archeologist must have some understanding of all these sciences to extract from sites and materials every possible piece of information which may lead to a better understanding of prehistory. One requirement, however, is necessary for the archeologist to make any contribution to humanity's inheritance of knowledge: the archeologist must be able to record and publish every minor fact for the benefit of colleagues and successors, because the writing of prehistory requires the synthesis of all archeological discovery and interpretation. *See* ANTHROPOLOGY; PHYSICAL ANTHROPOLOGY.

David Hurst Thomas

Archiacanthocephala

An order (or class, according to some classifications) of the phylum Acanthocephala. The adult worms are parasitic in terrestrial vertebrates. The body wall and lemnisci of species in this order have numerous amitotically fragmented nuclei or a few ameboid giant nuclei. The main trunks of the lacunar system are dorsal and ventral, or dorsal. Typically, there are eight separate cement glands in the male. Two ligament sacs are present in the female, one dorsal, the other ventral. They are persistent and united with the openings of the uterine bell. The eggs are elliptical and have a thick shell. The proboscis receptacle has a conspicuous ventral cleft or is a closed sac with two concentric muscle layers. The proboscis hooks occur in long rows or spiral rows, and the trunk lacks spines. The cystacanth occurs in grubs, roaches, and grasshoppers. Some common archiacanthocephalans are *Oncicola canis*, *Moniliformis moniliformis*, and *Macracanthorhynchus hirudinaceus*.

Oncicola canis is a short plump acanthocephalan, primarily parasitic in dogs and other Canidae. It occurs also in cats. The body of the adult, 0.2–0.5 in. (6–14 mm) long, is short and heavy with irregular cross furrows. The globular proboscis has six spiral rows of six hooks each. Male organs occupy more than one-half the length of the body. The arthropod intermediate host is unknown. Cystacanths have been found in armadillos and in the esophageal walls of turkeys, indicating a transport host in the life cycle.

Moniliformis moniliformis is an elongate acanthocephalan which is parasitic in house rats. The females are 4–12 in. (10–30 cm) long, whereas the males measure 2–5 in. (6–13 cm). The body of both sexes exhibits conspicuous pseudosegmentation except on the extremities. The proboscis is cylindrical with 12–15 rows of 10 or 11 hooks each. Male organs are confined to the posterior half of the body. Eight pyriform cement glands are present and the testes are long and elliptical. The egg is elliptical with a clear, thick outer shell. Cockroaches (*Periplaneta americana*) serve as the intermediate host; however, in Europe a beetle, *Blaps mucronata*, is the intermediate host. The cystacanth develops broad lateral flanges. Occasionally infections have been found in humans.

Macracanthorhynchus hirudinaceus is the giant thorn-headed worm of hogs and is probably the best known of all acanthocephalans because of its cosmopolitan distribution (see **illus.**) Females measure 10–24 in. (25–60 cm) in length and the males 2–4 in. (5–10 cm). The worms are pinkish with a transversely wrinkled body which tapers from a rather broad anterior end to a slender posterior end. The proboscis is globular with six spiral rows of six hooks each. The testes are elongate and located in the middle of the body or anterior to it. The eight cement glands are elliptical in shape, arranged in four consecutive pairs in the posterior portion of the body. Eggs are elliptical in shape with a heavy, dark-brown outer shell which is irregularly grooved and ridged, giving it a sculptured appearance. At least 25 species of scarabaeid beetle larvae have been reported as intermediate hosts. The cystacanth is cylindrical in shape. In addition to their occurrence in the domestic pig, adults have been reported to occur in squirrels, chipmunks, moles, and occasionally humans. This acanthocephalan is of considerable economic importance to the hog-raising industry.

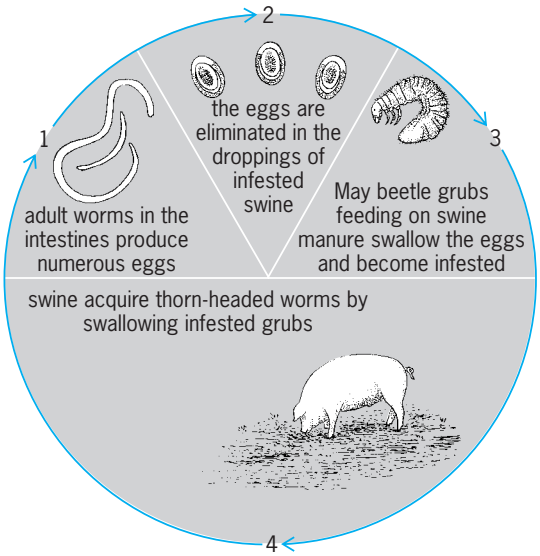

Life cycle of the giant thorn-headed worm of hogs.

Macracanthorhynchus ingens is an acanthocephalan found in raccoons (*Procyon lotor lotor*) which is similar in appearance to *M. hirudinaceus*. Female worms are 7–12 in. (18–30 cm) in length, and males 5–6 in. (13–15 cm). Mature eggs are elliptical in shape, with a heavy, dark-brown outer shell which is irregularly sculptured. Larvae of scarabaeid beetles serve as intermediate hosts. Two species of *Phyllophaga* and one species of *Ligyrus* have been successfully infected. Experimental work has shown that frogs may serve as transport hosts. *Macracanthorhynchus ingens* has been reported from a child who had a history of eating "bugs." Nine worms were removed from the child after the administration of an anthelminthic. *See* ACANTHOCEPHALA; COLEOPTERA; ORTHOPTERA.

Donald V. Moore

Archidiidae

A subclass of the plant class Bryopsida (mosses). The Archidiidae consists of a single genus, *Archidium*, with 26 species, occurring in ephemeral habitats, especially in wet, grassy places.

The small gametophytes consist of erect, simple or forked stems, and elongate, singly costate leaves in numerous rows. The capsules, usually terminal, are globose, immersed, and irregularly rupturing. The capsule wall consists of a single layer of cells; stomata, peristome, and columella are lacking. The spores are few, developing 1–44 spore mother cells scattered in the outermost layer of the endothecium, and are large (50–310 micrometers) and thick-walled.

The calyptra is scarcely differentiated. The haploid chromosome numbers are 13 and 26.

The subclass is unique in the scattered origin of spores in a single layer of the endothecium and in the absence of a quadrant stage in the early ontogeny of the capsule. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA; BUXBAUMIIDAE; DAWSONIIDAE; POLYTRICHIDAE; TETRAPHIDIDAE. Howard Crum

Bibliography. J. A. Snider, A revision of the genus *Archidium* (Musci), *J. Hattori Bot. Lab.*, 39:105–201, 1975; J. A. Snider, Sporophyte development in the genus *Archidium* (Musci), *J. Hattori Bot. Lab.*, 39:85–104, 1975.

Archigregarinida

An order of the protozoan subclass Gregarinia, class Telosporea, subphylum Sporozoa. All gregarines are parasites of the digestive tract and body cavity of invertebrates or lower chordates; their large, mature trophozoites (vegetative stages) live outside the host's cells. The Archigregarinida are primitive gregarines, and live in marine worms (annelids and lower chordates—enteropneustids, sipunculids, and ascidians). Their life cycle includes sexual and asexual phases and involves three periods of schizogony (multiple fission).

There are only 28 named species in 5 genera. The most important genus is *Selenidium,* which has 24 species. Its members occur in the intestine of marine polychaete annelids, sipunculids, enteropneustids, and ascidians.

The hosts are infected by eating oocysts containing sporozoites. The sporozoites emerge, enter host intestinal epithelial cells, become ellipsoidal, grow into large schizonts, and produce many merozoites or schizozoites. These enter the intestinal lumen, become attached to intestinal cells by their anterior end, and grow into gamonts. The gamonts become free in the lumen of the intestine, and two of them (sporadins) become attached to each other in an association known as syzygy. They secrete a cyst wall around themselves, and the resultant gametocysts or gamontocysts are passed into the seawater with the feces. Within the gamontocyst, one of the gamonts produces a large number of female gametes by multiple fission, while the other gamont produces an equal number of male gametes; the two types of gamete are about the same size, but the male gametes have a flagellum. Fertilization takes place, and each zygote forms an oocyst (often called sporocyst or spore) within the gamontocyst. Then by a third multiple fission, eight sporozoites are formed in each oocyst. *See* GREGARINIA; PROTOZOA; SPOROZOA; TELOSPOREA. Norman D. Levine

Archimedes' principle

The principle that there is an upward fluid force on a body submerged (or floating) in a stationary fluid that is equal to the weight of the fluid displaced by the body. This concept, perhaps the oldest stated principle in fluid mechanics, was first put forth by Archimedes in the third century B.C.

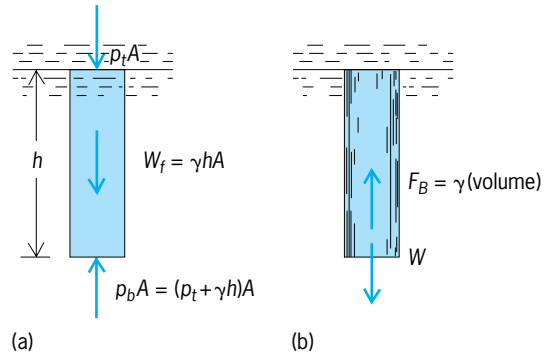

Fig. 1. Demonstration of Archimedes' principle. (*a*) Forces on a water column; p_t = pressure at top of column, p_b = pressure at bottom of column, h = height of column, A = area of column, W_f = weight of fluid column, γ = specific weight of fluid. (*b*) When the water column is replaced with a solid object, hydrostatic forces on this object are the same as on the original water column; F_B = buoyant force, W = weight of object.

In a static fluid, the weight of the fluid causes an increase in pressure with depth. Thus, at the surface of the fluid, the pressure is atmospheric pressure ($p_0 = 14.7$ lb/in.$^2 = 101$ kilonewtons/m^2), while at a depth h the pressure has a larger value of p_1, given by Eq. (1), where γ is the specific weight of the fluid

$$p_1 = p_0 + \gamma h \tag{1}$$

(weight/volume). The difference in pressure force between the bottom and the top of a water column (**Fig. 1***a*) is therefore given by Eq. (2), where h and

$$(p_b - p_t)A = \gamma hA \tag{2}$$

A are the height and area of the column, and p_b and p_t are the pressures at the bottom and top of the column. When the water is in equilibrium, this difference is precisely equal to the weight W_f of the water within the column, given by Eq. (3). If the

$$W_f = \gamma(\text{volume}) = \gamma hA \tag{3}$$

water column were replaced with a solid object (Fig. 1*b*), the pressure forces on the object would be the same as on the original water column. That is, the net hydrostatic force on the object, termed the buoyant force F_B, would be equal to the weight of the water displaced (which is the statement of Archimedes' principle). The same concept holds for a body of arbitrary shape, which can be thought of a consisting of many small vertical columns fastened together. Archimedes' principle is valid for submerged or floating bodies in liquids or gases. *See* BUOYANCY; SPECIFIC GRAVITY.

The fact that a helium-filled balloon rises through air is explained by any of the following equivalent statements: (1) the balloon's weight is less than its buoyant force; (2) the weight of the displaced air is greater than the balloon's weight; or (3) the average air pressure on the lower surface of the balloon is

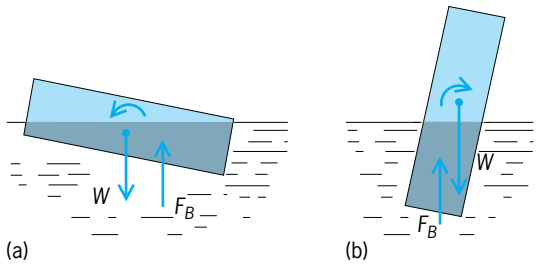

Fig. 2. Stability of a floating body depends on locations of the buoyant force F_B, which acts through the centroid of the displaced fluid volume, and the weight W, which acts through the center of gravity. Curved arrows indicate moments that result from combination of F_B and W. (*a*) Body in a stable position. (*b*) Body in an unstable position.

sufficiently greater than that on its upper surface.

A neutrally buoyant submarine displaces a volume of water precisely equal to its weight. The average specific weight of the submarine (the steel hull, power plant, occupants, enclosed air, and so forth) equals the specific weight of seawater; the buoyant force equals the total weight. In contrast, if a scale indicates a person's weight as 150 lb (667 newtons), the actual weight is approximately 150 + 0.18 = 150.18 lb (667.2 + 0.8 = 668.0 N) because of the approximate 0.18-lb (0.8-N) buoyant force of the air (specific weight of air times the person's volume) pushing upward on the person. However, since a person is much more dense than air, the buoyant force in this situation is usually negligible.

The location of the buoyant force (which acts through the centroid of the displaced fluid volume) and the location of the center of gravity of the body determine the stability of submerged or floating bodies. Specifically, a body's stability depends on what happens when the body is given a slight angular displacement (as when a ship rolls because of wave motion). If the moment (torque) caused by the buoyant force and the equal-but-opposite weight tends to rotate the body back to its horizontal equilibrium position, the body is stable (**Fig. 2***a*). However, the body is unstable if the moment causes the body to rotate further from its equilibrium position (Fig. 2*b*) *See* HYDROSTATICS. Bruce R. Munson

Bibliography. E. N. Gilbert, How things float, *Amer. Math. Month.*, 98(3):201–216, 1991; B. R. Munson, D. F. Young, and T. H. Okiishi, *Fundamentals of Fluid Mechanics*, 3d ed., 1997.

Architectural acoustics

The science of sound as it pertains to buildings. There are three major branches of architectural acoustics. (1) Room acoustics involves the design of the interior of buildings to project properly diffused sound at appropriate levels and with appropriate esthetic qualities for music and adequate intelligibility for speech. Room acoustics is an essential component of the design of theaters, concert halls, lecture rooms, classrooms, and churches, among other building types. (2) Noise control or noise management involves the reduction and control of noise between a potentially disturbing sound source and a listener. The walls, floors, ceilings, windows, and doors in buildings reduce sound energy as it travels through them. Disturbing noise sources may be equipment and people within the building or intruding noise from external sources of sound such as amplified music, airplanes, or highways. Noise from building services, including heating, ventilating, and air-conditioning systems, lighting systems, and electrical power systems, must also be controlled so that it does not disturb people using buildings. (3) The design of sound reinforcement and enhancement systems uses electronic equipment to improve the quality of sounds heard in rooms.

Properties of building materials. When a sound wave moving through the air strikes a building surface, three things can happen to it. A portion of the sound energy is absorbed within the wall, and converted into another form of energy such as heat; a portion of the energy is reflected back into the room; and a portion is transmitted through the wall to the other side (**Fig. 1**).

Sound absorption. The sound absorption coefficient α is the portion of energy that is absorbed compared to the total energy that strikes the surface. At a given frequency, the absorption coefficient is a function of angle of incidence. In architectural acoustics it is convenient to use an average absorption coefficient which represents an average absorption over all angles of incidence. The average absorption coefficient is assumed to depend only on the physical characteristics of the material and not on the sound field (**Table 1**). The absorption of a surface can be expressed in sabins (or metric sabins), units indicating the absorption of 1 ft^2 (or 1 m^2) of a totally absorbing surface, that is, a surface having an absorption coefficient of unity (**Table 2**). *See* SOUND ABSORPTION.

Sound is absorbed by any mechanism that converts incident sound waves into other forms of energy and ultimately to heat. All materials used in buildings absorb some sound. Generally, materials that are heavy and massive such as brick and concrete will not absorb much sound energy, and instead reflect sound back into the room. Materials that are light and porous will absorb significant amounts of

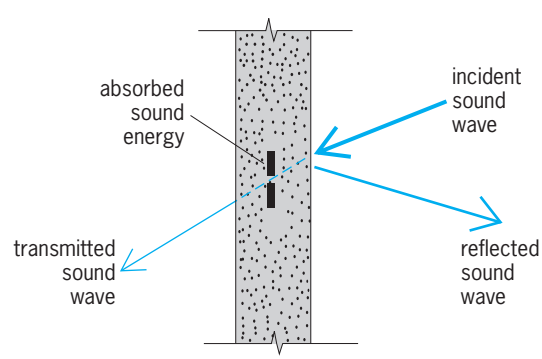

Fig. 1. Sound wave striking a building surface. Absorption of a portion of the sound energy in the wall, reflection of a portion back into the room, and transmission of a portion through the wall are shown. (*Siebein Associates*)

TABLE 1. Sound absorption coefficients of general building materials*

Material	Values of α					
	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz
Brick, unglazed	.03	.03	.03	.04	.05	.07
Carpet, heavy						
On concrete	.02	.06	.14	.37	.60	.65
On 40-oz/yd^2 (1.36 kg/m^2) hair-felt or foam rubber	.08	.24	.15	.69	.71	.73
Concrete block						
Coarse	.36	.44	.31	.29	.39	.25
Painted	.10	.05	.06	.07	.09	.08
Fabrics						
Light velour, 10 oz/yd^2 (0.34 kg/m^2), hung straight, in contact with wall	.03	.04	.11	.17	.24	.35
Heavy velour, 18 oz/yd^2 (0.61 kg/m^2), draped to half area	.14	.35	.55	.72	.70	.65
Floorings						
Concrete or terrazzo	.01	.01	.015	.02	.02	.02
Linoleum, asphalt, rubber, or cork tile on concrete	.02	.03	.03	.03	.03	.02
Wood	.15	.11	.10	.07	.06	.07
Glass						
Large panes of heavy plate glass	.18	.06	.04	.03	.02	.02
Ordinary window glass	.35	.25	.18	.12	.07	.04
Gypsum board 1/2 in. (12.7 mm) nailed to 2 × 4's 16 in. (406 mm) on center	.29	.10	.05	.04	.07	.09
Marble or glazed tile	.01	.01	.01	.01	.02	.02
Plaster, gypsum or lime						
Smooth finish on tile or brick	.01	.01	.02	.03	.04	.05
Rough finish on lath	.14	.10	.06	.05	.04	.03
Plywood paneling, 3/8 in. (9.5 mm) thick	.28	.22	.17	.09	.10	.11

*Data from Acoustical Materials Association.

sound energy depending upon their porosity, thickness, density, and mounting. The sound waves are converted into heat energy by propagating through the interstices of the material and by vibration of the small fibers of the material. Another important mechanism of absorption occurs when sound waves force a panel into motion; the resulting flexural vibration converts a fraction of the incident sound energy into heat. Materials especially designed to have relatively high absorption coefficients are called acoustical materials, acoustical wall panels, and acoustical ceiling tiles.

The noise reduction coefficient (NRC) is an average of the absorption coefficients of a material in the 250-, 500-, 1000-, and 2000-Hz octave bands. The noise reduction coefficient is often used to make initial comparisons of the effectiveness of absorbent materials used in offices, atriums, corridors, and other noncritical spaces.

Sound reflection. When a sound wave is larger than the dimension of the surface it strikes, it will reflect from the surface at an angle of reflection that is equal to the angle of incidence. Diffusion or scattering occurs when sound waves strike a surface with dimensions that are the same order of magnitude as their wavelengths. In that case, the scattering of sound waves from the irregular surface randomizes the directions of the reflected sound waves. Splayed wall panels, convex curves, coffers, and specially designed diffusing panels effectively diffuse or scatter sounds (**Fig. 2**).

Sound transmission. The sound transmission coefficient τ is the portion of sound energy that is transmitted through the wall compared to the total energy that strikes the surface. The sound transmission loss in decibels is equal to 10 log $(1/\tau)$. It measures

TABLE 2. Approximate sound absorption of occupied and unoccupied auditorium chairs

Type	Absorption, sabins per chair*					
	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz
Upholstered chair, occupied, depending on type of chair and spacing	3.2–4.4	3.5–5.4	4.2–6.5	4.8–7.0	5.0–7.0	4.8–7.0
Upholstered chair, unoccupied, depending on type of chair and spacing	1.8–3.5	2.6–5.0	3.6–5.0	3.7–5.3	3.7–6.1	4.3–5.3
Wood pews, occupied, depending on spacing	3.1	3.4	4.1	4.7	4.8	4.7
Metal or wood chair, unoccupied	0.15	0.2	0.25	0.40	0.45	

*1 sabin = 0.093 metric sabin.

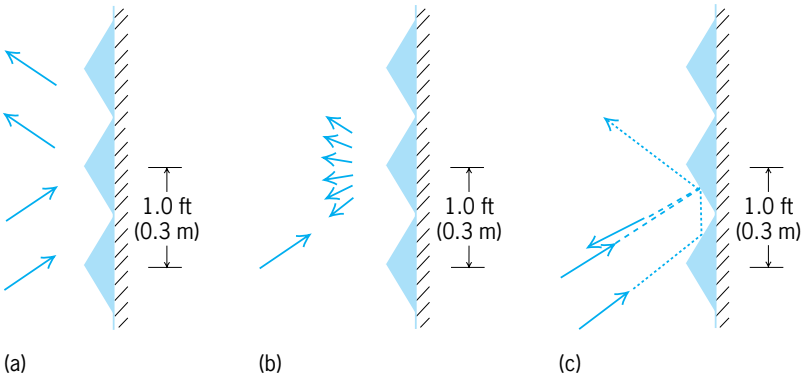

Fig. 2. Reflection of sound from a plane surface having splays with centers 1 ft (0.3 m) apart. Wavelength of sound is (*a*) much larger than the splays (frequency 100 Hz), (*b*) similar in size to that of the splays (frequency 1000 Hz), and (*c*) much smaller than the splays (frequency 100,000 Hz). (*After C. M. Harris, ed., Handbook of Acoustical Measurements and Noise Control, 3d ed., McGraw-Hill, 1991*)

how much sound is lost or how much the sound is reduced as it is transmitted through a construction assembly (an assemblage of materials that are used to construct a wall, roof, ceiling, or other surface in a building).

Frequency dependence. Sound reflection, absorption, and transmission all vary with the frequency or wavelength of sound. Therefore, it is essential to know the absorption and transmission coefficients over the entire frequency range of audible sounds. These quantities are usually measured in laboratory tests in 1/3 octave bands from 100 Hz or lower to 4000 Hz and higher.

Reverberation. Reverberation is the persistence of sound in a room after the source has stopped due to multiple reflections from the walls and ceiling of the room. Around 1900, Wallace Clement Sabine conducted pioneering research in reverberation and sound quality in rooms. He found that the reverberation time, RT, the amount of time it takes for a sound to decay to the point of inaudibility in a room after the source has stopped, was related to the volume of the room in cubic meters, V, and the absorbing power, $\Sigma_i S_i \alpha_i$, of the interior finish materials in the room by Eq. (1). Here, the absorbing power is the

$$\mathrm{RT} = \frac{0.161V}{\Sigma_i S_i \alpha_i} \tag{1}$$

sum, over the room surfaces, labeled i, of the surface area in square meters, S_i, multiplied by the respective absorption coefficient, α_i. The corresponding formula in U.S. Customary units is Eq. (2), where V'

$$\mathrm{RT} = \frac{0.049V'}{\Sigma_i S_i' \alpha_i} \tag{2}$$

and S_i' are in cubic feet and square feet respectively. *See* REVERBERATION.

Room acoustics. One essential component of room acoustics is an understanding of psychoacoustics and the qualitative evaluation of sounds heard by people in rooms. Psychoacoustics is the study of the psychology of sounds. It includes studies conducted in laboratories and in actual listening rooms of how people react to the level, frequency content, direction, and arrival time of sounds. These studies have established a set of relationships among the acoustical qualities that have been found to be important in the perception of sound, the room surfaces that contribute to these qualities, and the physical components of the sound field in a room that contribute to these properties (**Table 3**). *See* PSYCHOACOUSTICS.

TABLE 3. Summary of acoustical qualities in rooms and the architectural features of rooms that contribute to them

Acoustic quality	Architectural feature that creates acoustic quality	Physical acoustic event that creates acoustic quality
Envelopment and source width	Narrow rooms up to 70–80 ft (21–24 m) across Multiple tiers of narrow balconies	Early sound reflections arriving at the listener from the side (up to 80 ms after the direct sound)
Clarity	Sound reflecting ceiling	Sound reflections (usually from overhead) that arrive shortly after the direct sound
Reverberance	Large room volume Sound-reflecting materials	Prolonging of sound in the room after the source has stopped due to multiple reflections from the room surfaces
Loudness	Limited size of room Provision of multiple sound reflections to seats	Sound reflections from the ceiling and walls arriving shortly after the direct sound
Intimacy	Reflecting surfaces close to listeners	Arrival of the first sound reflection from a building surface shortly after the direct sound
Warmth	Heavy massive building materials	Strength of reflected low-frequency sound
Brilliance	Reflecting surfaces for high-frequency sounds	Strength of high-frequency sound
Spaciousness	Surface texture and sound-diffusing materials	Late sound energy arriving from the sides (after 80–100 ms)
Localization of sound	Clear sight and sound line between listener and source	Strength of direct sound relative to subsequent reflections
Ensemble	Overhead and side-wall sound reflecting surfaces near the performing area	Sound reflections that allow the musicians across the stage to hear each other

These relationships are based upon the concept of the impulse response of the room. The use of this concept has been an essential component in increasing the understanding of room acoustics and improving the acoustical design of rooms. The impulse response consists of the response of a room to an impulsive sound such as a single note of music or a single syllable of speech. It has four components. The direct sound is the sound that travels directly through the air by the shortest path from the source to the listener. Early reflected sounds are sounds that reflect off room surfaces and reach the ears of listeners within 50–80 milliseconds after the direct sound. These sounds will increase the apparent loudness and clarity of sounds when they come from overhead surfaces (the ceiling); increase the apparent width of the sound source and contribute to a sense of acoustic spaciousness when they come from the sides; and create a sense of acoustic intimacy if they arrive very shortly after the direct sound (less than 20 to 40 ms) regardless of direction. If the sounds contain low-frequency or bass energy, a sense of warmth will be achieved. If they contain higher-frequency or treble sounds, a sense of brilliance will be achieved. The reverberant tail is the sound energy that has reflected off multiple surfaces before it reaches the ears of listeners, gradually decreasing in amplitude or loudness and arriving at progressively longer time periods after the direct sound. It contributes to a sense of reverberance or fullness of musical sounds if it lasts as long as 2.0 to 2.3 s in a concert hall. If the reverberant sound energy arrives at the ears of listeners from the sides and the sounds that arrive at the left and right ears of listeners are configured somewhat differently, a sense of spaciousness will be enhanced. The fourth component of the impulse response is the background noise. This is the level of ambient noise in a room from building services, people moving about, equipment in the room, and intruding sounds from outside the room (**Fig. 3**).

Sound absorbent materials are strategically located in rooms to absorb unwanted sounds. They are used on the ceilings and upper wall surfaces of atriums, gymnasiums, cafeterias, and large open offices to reduce reverberant sounds that reflect off these surfaces, creating excessive noise. They are also used on the rear walls of lecture rooms and auditoriums to reduce the level of potential echoes from these surfaces. An echo is a sound reflection that arrives at a listener's ear 70 ms or more after the direct sound and is loud enough to be heard as a distinct acoustic event or a second image of the original sound. Absorbent materials are also used on the surfaces of large concave shapes such as domes. A concave shape will focus much of the reflected sound energy at its geometric center, creating an acoustic hot spot or a location where sounds are heard as much louder than at other locations. Absorbent materials are also used to vary the reflected and reverberant energy levels in rooms that are used for several acoustic functions such as music and theater. Sound absorbent banners or curtains may be lowered in a concert hall to cover wall surfaces when the room is used for theatrical performances and less reverberation is desired. *See* ECHO.

Several important design concepts are used to provide good listening conditions in rooms for speech and music. First is to provide good access to the direct sound for all people in the room. This usually involves raising the source of sound on an elevated stage, altar, or podium at the front of the room and sloping the floor surface to elevate the ears of people above the heads of those seated in front of them. The width and depth of the room should also be limited so that the natural direct sound can project from the speaker or instruments at the front of the room to the listeners. Second is to limit the background noise level in the room so that people can hear the sound they want to hear above the level of the ambient sound. Third is to limit the reverberation time in

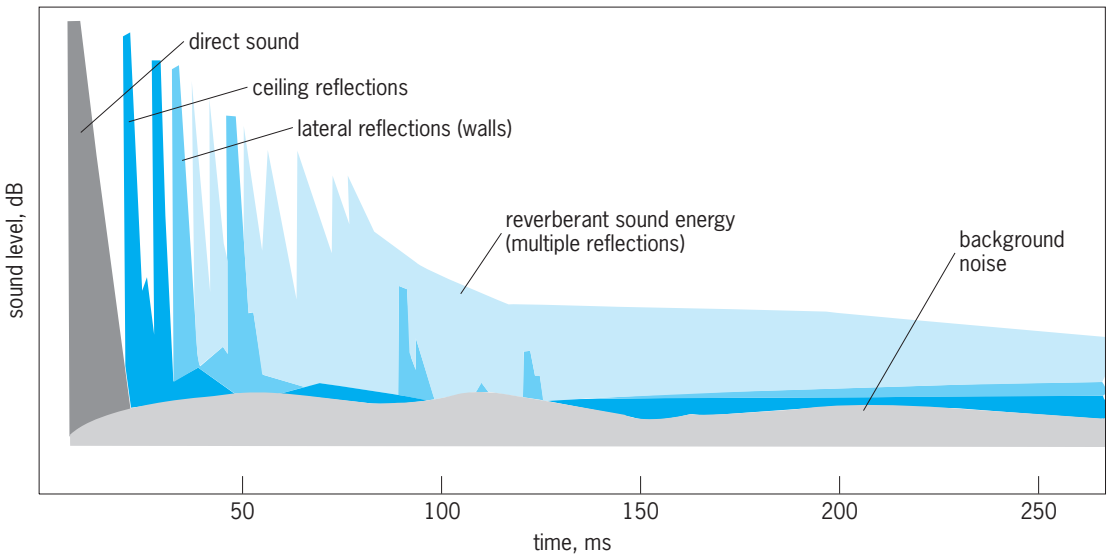

Fig. 3. Impulse response, showing four basic parts: the direct sound, early sound reflections from the ceiling and walls, the reverberant tail, and the background or ambient noise in the room. (*Siebein Associates*)

TABLE 4. Appropriate reverberation times for various rooms

Room	Reverberation time at middle frequencies (500–1000 Hz), s
Class room	0.3–0.6
Lecture hall	0.6–0.9
Music rehearsal room	0.6–1.1
Small theater	0.8–1.1
Multipurpose auditorium	1.2–2.0
Church	1.0–2.0*
Concert hall	1.8–2.3

*Depending upon size and nature of music program.

the room so that sounds are heard clearly and fully, while providing enough reverberant sound energy that sounds are heard as "full" and "live." If there is too much reverberation in a room, the persistence of an initial syllable will cover up or mask the one that follows it, making it difficult to understand what is being said (**Table 4**).

The ceilings and walls of rooms are designed to direct early reflected sounds to listeners' ears. The wall and ceiling panels in the front of the room are shaped to direct reflected sounds over the entire audience seating area. The panels in the center of the room are designed to cast reflections over the rear half of the seating area. The panels at the rear of the room cast reflections over the last rows of seating. This approach to room acoustics design uses reflected sounds to increase the apparent level of sounds toward the rear of the room where the direct sound is weakest. This helps to create more uniform sound levels throughout the room. When these panels are free-form floating panels suspended below the upper ceiling of a room, they are often called acoustic clouds (**Fig. 4**).

The Eugene McDermott Concert Hall in Dallas, TX, provides examples of several acoustic design concepts (Fig. 4). The canopy over the orchestra platform provides strong early reflections to the audience for clarity, intimacy, and envelopment. The room is surrounded by large reverberation chambers with doors that can open and close. This allows sound to enter the chambers, reflect off the walls, and reenter the room as reverberant energy, providing a rich, full sound for orchestral concerts and organ. The room has a series of narrow balconies that surround the main seating area, creating an intimate space. The undersides and faces of the balconies provide surfaces to cast sound reflections from the sides of the room to the audience for a spacious and enveloping sound. The room is built of heavy concrete with wood laminated to its surface. These materials reflect all frequencies of sound, creating a sense of wamth (reflection of low-frequency sounds) and brilliance (reflection of high-frequency sounds).

Noise control. Acoustical planning concepts for buildings include placing noisy activities away from activities that require relative quiet and locating noise-sensitive activities away from major sources of noise. Buffer spaces such as corridors or storage spaces are often used to separate two rooms that require acoustical privacy such as music rehearsal

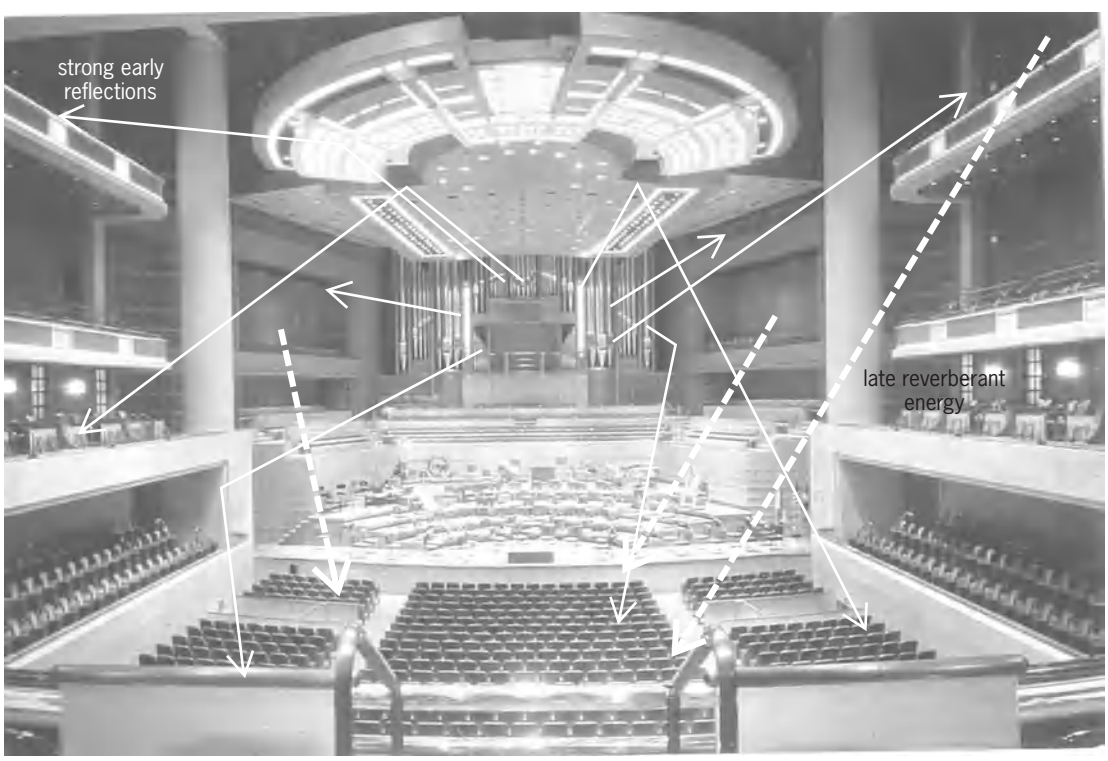

Fig. 4. Interior of the Eugene McDermott Concert Hall in the Morton H. Meyerson Symphony Center in Dallas, TX. Russell Johnson of Artec Consultants, Inc., acoustical consultants, and I.M. Pei of Pei, Cobb, Freed and Partners, architects, collaborated on the design of the room. Paths of strong early reflections and late reverberant energy are shown.

TABLE 5. STC ratings for various construction assemblies and the magnitude of perceived sounds heard by people listening on the other side of the assembly*

STC rating	Typical construction assembly	Typical application	Perception of intruding sounds on opposite side
35	0.5-in. (1.25-cm) gypsum board on both sides of metal studs	Office walls	Normal speech heard somewhat
45	0.5-in. (1.25-cm) gypsum board on both sides of metal studs with glass fiber insulation in stud space	Conference room walls	Normal speech heard faintly if at all Loud speech heard faintly
50	8-in. (20-cm) concrete masonry unit sealed on both sides	Classroom walls; mechanical rooms with air-handling units	Loud speech barely heard Music reduced in volume Low-frequency mechanical noise reduced
50	Two layers of 0.5-in. (1.25-cm) gypsum board on one side of metal studs, one layer of 0.5-in. (1.25-cm) gypsum board on the other side with glass fiber insulation in stud space	Classroom walls	Loud speech barely heard Music reduced in volume
55	Two layers of 0.5-in. (1.25-cm) gypsum board on both sides of metal studs with glass fiber insulation in stud space	Music practice room walls in schools	Loud speech generally not heard Music sounds, except loud bass, faintly heard

*Used with permission; © G. W. Siebein, *Tutorial on Architectural Acoustics*, 1992.

rooms in a school. Intruding noises from the exterior or from adjoining rooms can be reduced by using walls, ceilings, windows, and doors with appropriate transmission losses. The performance of a construction assembly used to separate rooms is often compared using a sound transmission class (STC) rating. This is a weighted average of the transmission loss values for the assembly in the 125- to 4000-Hz octave bands (**Table 5**).

Airborne sound is transmitted through a partition by forcing the partition into vibration. The vibrating partition becomes a secondary source of sound and radiates sound to the room on the other side. For a homogeneous wall, the transmission loss is related to its mass. The transmission loss increases by 3–6 dB per octave band over many frequencies. The transmission loss value of a partition is usually poor at frequencies corresponding to strongly excited natural frequencies of the partition and at certain critical frequencies where the phases of incident sound waves coincide with the phases of vibration of the wall. The transmission loss also increases by 4–6 dB in a given frequency band when the mass of the wall is doubled. A compound or double wall assembly can be used to reach a relatively high transmission loss with low mass per unit wall area. The separation between the two leaves or surfaces of the wall must be maintained as completely as possible for this to occur.

It is essential to control noise from building services. The location of air-conditioning plants on a site should be chosen so as to reduce propagation of noise to neighbors. Mechanical rooms in buildings that house air handling units, pumps, and other equipment should be located away from noise-sensitive rooms. Noise control treatments in the air-conditioning system include providing vibration isolators for equipment; providing flexible connections between ducts, conduits, and pipes to equipment; designing air ducts to operate with air velocities that will not create turbulent flow noise; and installing silencers or attenuators in the ducts to reduce noise produced by fans from traveling through the duct work. *See* MECHANICAL VIBRATION; VIBRATION ISOLATION.

Noise buildup within a room can be reduced by applying absorbent materials to the interior surfaces. The amount of noise that is reduced by adding absorbent material is estimated by the quantity 10 log (a_2/a_1). Here, a_2 is the amount of absorption in the room after absorbent materials have been added to it, and a_1 is the original amount of absorbtion in the room. The amount of absorption in a room is equal to the sum of the surface areas of each of the room surfaces multiplied by their respective sound absorption coefficients. *See* ACOUSTIC NOISE.

Sound reinforcement. Sound reinforcement systems, electronic enhancement systems, and sound amplification systems are used in many buildings. A sound reinforcement system amplifies the natural acoustic sounds in a room that is too large for people to hear with just "natural" room acoustics. This type of system reinforces the natural sounds that come from the room, increasing their apparent loudness with a series of loudspeakers.

In an electronic enhancement system, loudspeakers act as virtual room surfaces to create the perception that sounds are reflected from these surfaces at the proper times and with the proper loudness. These systems usually have a network of loudspeakers located throughout a room and connected to a microprocessor. The microprocessor can delay the signals to arrive at times corresponding to reflected sounds from the virtual room surfaces. It can also add reverberation and other special acoustic effects to create a virtual acoustic space. A sound amplification system makes all sounds played in a space louder. It is usually not designed to supplement the

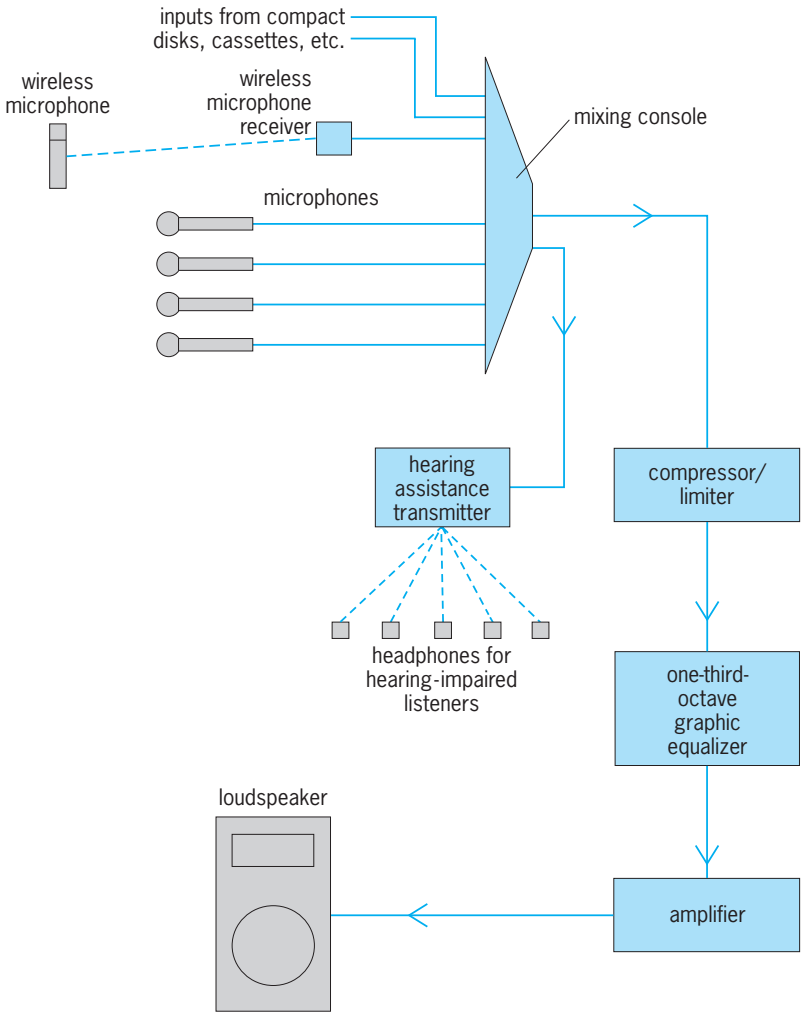

Fig. 5. Simple sound-reinforcement system.

natural room acoustics or to provide subtle virtual room effects to the amplified sounds (**Fig. 5**).

Several basic types of loudspeaker arrangements are used in rooms. When the ceiling height of a room is 20 ft (6 m) or more, as in a church or theater, a central cluster system is often used. This consists of a group of loudspeakers located high above the sound source, which projects sound down toward the audience seating. In a room with lower ceilings, such as a convention center or hotel ball room, a distributed loudspeaker system is used. It consists of a network of loudspeakers mounted in the ceiling and spaced at even increments.

In a long, horizontal room, the sound from the loudspeaker will reach the ears of listeners before the direct sound from the person speaking because the sound waves travel in air at 1128 ft/s (344 m/s) and the sound moving through the wires as an electronic signal travels many times faster. A digital delay is added to the sound arriving from the loudspeaker so that it will arrive after the direct sound from the person speaking, maintaining its correct location in the impulse response as a reinforcing "electronic reflection." *See* SOUND; SOUND-REINFORCEMENT SYSTEM.

Gary W. Siebein

Bibliography. L. L. Beranek, *Concert and Opera Halls: How They Sound*, 1998; W. C. Cavanaugh and J. Wilkes, *Architectural Acoustics: Principles and Practice*, 1998; M. D. Egan, *Architectural Acoustics*, 1989; C. M. Harris (ed.), *Handbook of Acoustical Measurements and Noise Control*, 3d ed., 1991; M. Mehta, J. Johnson, and J. Rocafort, *Architectural Acoustics: Principles and Design*, 1999; C. M. Salter and Associates, Inc., *Acoustics: Architecture, Engineering, the Environment*, 1998.

Architectural engineering

A discipline that deals with the technological aspects of buildings, including the properties and behavior of building materials and components, foundation design, structural analysis and design, environmental system analysis and design, construction management, and building operation. Environmental systems, which may account for 45–70% of a building's cost, include heating, ventilating and air conditioning, illumination, building power systems, plumbing and piping, storm drainage, building communications, acoustics, vertical and horizontal transportation, fire protection, alternate energy sources, heat recovery, and energy conservation. In addition, to help protect the public from unnecessary risk, architectural engineers must be familiar with the various building codes, plumbing, electrical and mechanical codes, and the Life Safety Code. The latter code is similar to a building code and is designed to require planning and construction techniques in buildings which will minimize possible hazards to the occupants. *See* ELECTRICAL CODES; FIRE TECHNOLOGY.

Architectural engineering differs from other engineering disciplines in two important aspects. Most engineers work with other engineers, while most architectural engineers work or consult with architects. Furthermore, an architectural engineer not only must be fully qualified in engineering, but must also be thoroughly versed in all architectural considerations involved in design and construction. An architectural engineer designing a structural or environmental system is expected to be familiar not only with that system, but also with the multitude of architectural considerations which may affect its design, installation, and operation.

Although retaining more or less conventional forms, buildings have become increasingly complex during the last century. The need for a professional engineer to solve inherent technological problems was first recognized by the University of Illinois, which established an architectural engineering curriculum in the 1890s. There are now a number of such programs throughout the United States. All of these programs meet the same requirements as other engineering disciplines in mathematics, basic science, engineering science, humanities, and engineering design. All are accredited by the Accreditation Board for Engineering and Technology (ABET). However, there is considerable diversity in the architectural engineering curricula offered by the various

institutions. Some programs are of a traditional 4-year duration. These programs meet the engineering requirements at the expense of architectural content. Other programs are 5 years in length. In general, they contain more architecturally related instruction as well as greater breadth in the environmental technologies.

A majority of the graduates of architectural engineering curricula become registered engineers. However, a sizable number elect to continue their education and become registered as architects as well. Career opportunities include partnerships with architects, consultants to architects, facilities engineers, and construction managers. Opportunities are also available as technical representatives for equipment manufacturers. *See* BUILDINGS; ENGINEERING.

Thomas S. Dean

Archosauria

A diversified group of diapsid reptiles, comprising dinosaurs (and their descendants, birds), crocodylians, pterosaurs (flying reptiles), and various forms closely related to these groups. In the past, the primitive archosaurs from the Triassic Period were collectively referred to as Thecodontia, but this is an artificial grouping that is no longer accepted by specialists. *See* CROCODYLIA; ORNITHISCHIA; PTEROSAURIA; SAURISCHIA.

Archosaurian reptiles are readily distinguished by the presence of one or more openings (antorbital fenestrae) on each side of the skull between the nostril and eye socket, and an opening in the lower jaw (mandibular fenestra) [see **illustration**]. Among extant vertebrates, the two surviving groups of archosaurs, crocodylians and birds, differ from other reptiles and from mammals in many anatomical and behavioral characteristics. Both have a muscular foregut, or gizzard, containing stones for processing food, a heart with completely divided ventricles, and a pulmonary artery with three semilunar valves. Both crocodylians and birds use calls to mark territories and attract potential mates.

Early archosaurs employed quadrupedal locomotion, but the hindlimbs were already longer than the forelimbs. Later forms were typically bipedal, but many groups independently reverted to quadrupedal stance and gait. Most archosaurs were carnivorous, but herbivory evolved independently in a number of lineages, mainly among dinosaurs.

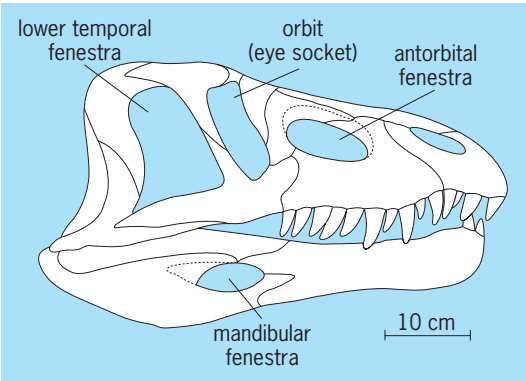

Skull of *Erythrosuchus africanus*, a basal archosaur from the early Middle Triassic of South Africa.

The earliest archosaurian fossils are known from the Late Permian. Archosaurian reptiles underwent a major diversification during the Triassic Period, when they became the dominant terrestrial vertebrates. Birds (with about 9000 known species) are the most diverse group of amniote vertebrates on land today, but crocodylians are represented by only about two dozen species.

The most primitive archosaurs, mainly from the Early Triassic and often grouped together as the Proterosuchia, already had the diagnostic openings in the skull but retained a sprawling limb posture with the elbows and knees sticking out sideways. Among more advanced archosaurs, the limbs moved partially or completely beneath the body, perhaps permitting more rapid locomotion. Paleontologists now distinguish two major evolutionary lineages among these advanced forms, the Crurotarsi and the Ornithodira. The Crurotarsi comprise crocodylians (Crocodylia) and various closely related forms (Crocodylomorpha), the long-snouted, superficially crocodilelike phytosaurs, the heavily armored aetosaurs, the facultatively bipedal ornithosuchids, and various large predators from the Middle and Late Triassic, commonly grouped together as Rauisuchia. The Ornithodira include dinosaurs (including birds), pterosaurs, and a few still poorly known forms probably related to the ancestry of these two groups.

Crurotarsans and ornithodirans are easily distinguished from each other by the structure of their ankle joints. In the former, the two principal ankle bones, the astragalus and the calcaneum, rotate against each other by means of a peg-and-socket joint (crurotarsal condition). In the latter, a simple hinge joint is developed between the astragalus and calcaneum (which are firmly attached to the bones of the lower leg) and the remainder of the ankle and foot (mesotarsal condition). Most ornithodirans, especially dinosaurs, show substantial modifications of the pelvic girdle and hindlimbs, which indicate erect gait. The posture of the hindlimbs in pterosaurs has been the subject of much debate. Recent discoveries of trackways suggest that they moved on all fours when walking on the ground, rather like bats. *See* AVES; DINOSAUR; REPTILIA.

Hans-Dieter Sues

Bibliography. P. J. Currie and K. Padian (eds.), *Encyclopedia of Dinosaurs*, San Diego, 1997; P. C. Sereno, Basal archosaurs: Phylogenetic relationships and functional implications, *Mem. Soc. Vert. Paleontol.*, 2:1–53, 1991; P. C. Sereno, The evolution of dinosaurs, *Science*, 284:2137–2147, 1999; D. B. Weishampel, P. Dodson, and H. Osmólska (eds.), *The Dinosauria*, 2d ed., University of California Press, Berkeley, 2004; P. Wellnhofer, *The Illustrated Encyclopedia of Pterosaurs*, Crescent Books, New York, 1991.

Arctic and subarctic islands

Defined primarily by climatic rather than latitudinal criteria, arctic islands are those in the Northern Hemisphere where the mean temperature of the warmest month does not exceed 50°F (10°C) and that of the coldest is not above 32°F (0°C). Subarctic islands are those in the Northern Hemisphere where the mean temperature of the warmest month is over 50°F (10°C) for less than 4 months and that of the coldest is less than 32°F (0°C).

Such islands generally are in high latitudes. Distribution of land and sea masses, ocean currents, and atmospheric circulation greatly modify the effect of latitude so that it is often misleading to use location relative to the Arctic Circle as a significant criterion of arctic or subarctic. The largest proportion by area of the islands lies in the Western Hemisphere, primarily in Greenland and in the Canadian Arctic Archipelago.

Diversity of land surfaces. Physiographically, the islands include all the varied major landforms found elsewhere in the world, from rugged mountains over 8000 ft (2500 m) high, through plateaus and hills, to level plains only recently emerged from the sea. All have been glaciated except Sakhalin and some of the islands in the Bering Sea sector. R. F. Flint reported that of the 5,800,000 mi^2 (15,000,000 km^2), or 10%, of land area of the world still ice-covered, over 5,000,000 mi^2 (13,000,000 km^2) is in Antarctica, and over 700,000 mi^2 (1,812,000 km^2) lies in the arctic islands with over 600,000 mi^2 (1,554,000 km^2) of this in Greenland. Removal of the weight of ice sheets and the resultant crustal rebound has exposed prominent marine beaches and wave-cut cliffs on many of the islands. These now commonly occur at elevations of over 300 ft (90 m) above sea level. *See* ANTARCTICA.

Islands of the arctic and the subarctic regions.

Climate. The general climatic pattern of these islands is set by their location relative to the two semipermanent centers of low pressure over the Aleutian Islands and over Iceland (the Aleutian Low and the Icelandic Low). Especially during the winters these low-pressure centers affect the areas from Kamchatka to southeastern Alaska and from Newfoundland to Novaya Zemlya, respectively. The intervening areas, including the islands of the northwestern Canadian Arctic Archipelago and those of the central and eastern Russian Arctic, are much less subject to cyclonic storms; rather they are dominated by stable, dry air masses often linked across the polar basin from either continent. During the summer this pattern decreases in intensity. Most of the precipitation is cyclonic in origin. Because they are marine areas, the islands receive more precipitation than they otherwise would, yet even so this is very light for most of the arctic islands removed from the zone of cyclonic activity. Also, because they are marine areas, the islands, regions of low temperatures by definition, are not regions of extreme low temperatures. Much lower temperatures are reported from continental land areas farther south than from the most northern arctic islands. In general, the larger the island and the closer its proximity to a continental landmass, the higher are the summer temperatures and the lower its winter temperature. *See* POLAR METEOROLOGY.

Vegetation and soils. The climatic differences between arctic and subarctic islands are reflected in their natural vegetation. The arctic islands are treeless. Natural vegetation consists of the tundra—mosses, sedges, lichens, grasses, and creeping shrubs. Luxuriance and continuity of ground cover vary with such factors as moisture, insolation, and soil nutrient conditions. Bare ground is often exposed, and in places plant growth may be lacking completely except for a few rock-encrusting lichens. In such places the ground surface may consist of frost-shattered rock fragments, tidal mud flats, boulder-strewn fell fields, or snow patches and ice. Permafrost (permanently frozen ground) occurs throughout the Arctic (and in parts of the subarctic) and is reflected in impeded drainage and patterned ground. *See* PERMAFROST.

The natural vegetation of subarctic islands characteristically is the boreal forest or taiga, composed predominantly of conifers such as spruce, fir, pine, and larch with deciduous trees such as birch, aspen, and willow; the latter are especially common in regrowth of clearings in the forest. Impeded drainage because of permafrost or glaciation gives rise to numerous ponds and muskeg areas. A transitional type of vegetation, the forest-tundra, is recognized

Size and elevation of larger arctic and subarctic islands*

Name	Area		Highest point (elevation)		
	mi^2	km^2	Name	ft	m
Aleutian Islands					
Unimak Island	15,500	40,100	Shishaldin Volcano	9,978	3,040
Unalaska Island	10,800	28,000	Makushin Volcano	6,680	2,030
St. Lawrence Island	18,200	47,100	Kookoligit Mountains	2,207	670
Nunivak Island	16,000	41,400	Roberts Mountain	1,675	510
Kodiak Island	37,400	96,900	Grayback Mountain	3,317	1,000
Canadian Arctic					
Archipelago	500,000	1,295,000			
Baffin Island	196,000	507,000	Penny Ice Cap	6,750	2,060
Ellesmere Island	76,000	197,000	Barbeau Peak	8,600	2,600
Victoria	84,000	217,000	Shaler Mountains	2,100	655
Banks	27,000	70,000	Durham Heights	2,400	732
Devon	21,000	55,000	Treuter Mountains	6,190	1,920
Axel Heiberg	17,000	43,000	Name not available	8,400	2,560
Melville	16,000	42,000	Raglan Range	3,500	1,070
Southampton	16,000	42,000	Porsild Mountains	1,750	530
Prince of Wales	13,000	33,000	Name not available	500	150
Newfoundland	42,734	109,000	Gros Morn	2,644	806
Greenland	840,000	2,176,000	Mountain Forel	11,286	3,500
Iceland	39,961	102,000	Oraefajökull	6,955	2,120
Svalbard (archipelago)	24,100	62,000			
Vest-Spitsbergen	15,250	39,000	Newtontopp	5,630	1,720
Franz Josef Land (archipelago)	7,000	18,000	Name not available	3,000	910
Novaya Zemlya (archipelago)	36,000	93,000			
Severny Island	21,000	54,000	Name not available	3,510	1,070
Yughny Island	15,000	39,000			
Severnaya Zemlya (archipelago)	14,000	36,000	Name not available	1,500	450
New Siberian Islands	12,000	31,000	Name not available	1,033	315
Wrangel Island	2,000	5,000	Peak Berry	3,510	1,070
Sakhalin Island	27,000	70,000	Nevelskoi Mountain	6,600	2,010
Kurile Islands	6,000	16,000	Alaid Volcano	7,674	2,030

*Approximate only in some cases because of incomplete mapping.

on some subarctic islands in sectors where smaller trees are widely spaced and abundant mosses cover the ground.

The typical soils of the subarctic islands are podzols—the surface soil grayish-white beneath the raw humus layer and highly acidic in nature. The tundra soils of the arctic islands really consist only of a dark-brown peaty surface layer over poorly defined thin horizons, and much of the ground cannot properly be termed soil. *See* SOIL.

Character of major islands. Within this general description, individual islands vary considerably (see **illus.** and **table**) A brief summary of some of the larger islands and archipelagos in the Western Hemisphere is given in the following sections.

Aleutian Islands. Extending southwest for more than 1000 mi (1600 km), from the Alaska Peninsula at 163°W to Attu Island at 175°W, the Aleutians separate the Bering Sea to the north from the Pacific Ocean to the south. Their rugged, mountainous surface consists of the drowned continuation of the Alaska and Aleutian ranges. Several active volcanoes are included among them. Most of the islands were glaciated by local mountain glaciers, a few of which persist in reduced size on some of the eastern islands (Unimak, Unalaska, and Umnak). There are about 150 islands and innumerable reefs. They are grouped from east to west as the Fox, Andreanof, Rat, and Near islands. Unimak, the easternmost island, is the largest, about 65 mi (105 km) long and 25 mi (40 km) wide.

The Aleutian Islands experience extremely variable weather because of the frequent cyclonic storms, complicated by the mountainous terrain. An associated gusty wind is known as the williwaw. Swift sea currents run among the islands and there is much fog. At Dutch Harbor (Unalaska Island) the mean temperatures are January, 32°F (0°C), and July, 51°F (11°C); while at Atka they are 33°F (1°C) and 50°F (10°C). Mean annual precipitation at the two stations is 56.7 and 70.2 in. (1440 and 1783 mm), respectively. The islands are treeless but generally support a luxuriant growth of grass, willow, and alder.

Canadian Arctic Archipelago. All those islands lying north of the continental mainland and west of Greenland to 141°W are included in this group. In rough outline the archipelago resembles a triangle, from a rather irregular base at about 61°N in the east and 67°N in the west, to its apex at the northernmost tip of Ellesmere Island in latitude 83°30′N. Those lying north of 74°N are known as the Queen Elizabeth Islands.

The easternmost islands of the archipelago are mountainous, whereas those to the west and northwest are plateaulike or plains. The mountains in Baffin Island, eastern Devon Island, and southeastern

Ellesmere Island are composed of Precambrian rock and average 5000–7000 ft (1520–2130 m) elevation. Continuing the mountain line northward through Ellesmere Island and Axel Heiberg Island, summit elevation exceeds 8000 ft (2500 m), partly reflecting a different fold axis, the Innuitian. In contrast the more southerly of the Queen Elizabeth Islands are more typically plateaulike, even though they contain more of the Innuitian fold structure; here the maximum elevation is about 3000 ft (920 m) in Melville Island but is commonly less than 1000 ft (305 m). The northwestern part of the Queen Elizabeth Islands is a recently emerged coastal plain, in which salt domes occur at places. Most of the islands in the western part of the southern archipelago are low and comparatively flat.

Most of the islands were glaciated, although there still is not sufficient information to establish conditions in the northwestern islands. Relic ice caps remain upon Ellesmere, Axel Heiberg, Devon, Bylot, and Baffin islands, because of the higher elevations and greater precipitation.

The Archipelago has long, cold winters. For 3–4 months mean monthly temperatures range from −20 to −30°F (−30 to −34°C) over most of the area. February is usually the coldest month; temperatures of −35 to −40°F (−37 to −40°C) are commonly recorded then. Rarely do temperatures drop below −50°F (−46°C), although a minimum of −63°F (−53°C) has been recorded at three stations. The southeastern and eastern parts of the archipelago have milder winters, because of the proximity to open water in Davis Strait and the passage of cyclonic storms. During the cool summer, temperatures are more uniform over the archipelago, with the July mean only from 40 to 50°F (4 to 10°C). Maximum temperatures exceed 60°F (16°C) at most but not all stations.

Precipitation is light. Most of the archipelago receives less than 10 in. (254 mm) a year, with only three stations (in the southeast) recording more than 15 in. (381 mm). One station (Eureka) reported less than 2 in. (50 mm) a year over a 2-year period. Snow may fall in any month, but over most of the archipelago rain accounts for about half the precipitation. Snowfall averages from 12.5 to 100 in. (320 to 2540 mm) in the southeast. Natural vegetation is of the tundra type.

Sea ice closes archipelago waters to shipping for 9–10 months in the year. Breakup usually occurs between late June and mid-July with freezeup between mid-September and late October, depending on location. The waters around the northwestern islands are rarely clear of ice. Icebergs are confined to the eastern parts, chiefly in the waters separating the archipelago from Greenland.

Newfoundland. Generally a plateau of rolling surface, this subarctic island is tilted west to east from the Long Range Mountains (over 2000 ft or 610 m) to the Avalon Peninsula (700 ft or 213 m). The northeast-southwest grain of the island, the result of folding and faulting, is reflected in the coastal configuration and physiography. The island was completely glaciated and most of it shows the results of intense ice scour. The indented seacoasts are commonly cliffed and rugged, with marked fiords in the high sectors of the Northern Peninsula. The only significant lowland area is a coastal plain in the west.

From the standpoint of climate and natural vegetation the island is predominantly subarctic, yet it includes aspects of the arctic and of the more continental regions. The moderating marine influence on climate is lessened by the island's east (that is, lee) location relative to the continental landmass, and by the cold Labrador Current, which flows south along the east coast and swings around to affect the south and parts of the west coasts.

Winters are cold, particularly in the interior (January mean, 15 to 20°F or −9 to −7°C), but are somewhat milder on the coast, especially in the southeast (St. John's January mean, 24°F or −4°C). Conversely, the coasts remain cooler in the summer (about 55°F or 13°C in July), whereas the western lowlands exceed 60°F or 16°C in mean daily July temperatures. The entire island has abundant precipitation, well distributed through the year, with heaviest falls (over 55 in. or 1400 mm) in the south. Snowfall is abundant (over 100 in. or 2540 mm) everywhere except along the south coasts. Sea ice seals off all coasts at its maximum extent, except the south. Fog is fairly frequent both on the coast and inland. It is most common in summer and on the southeast coasts. The juxtaposition of the cold Labrador Current and the warm Gulf Stream in the offshore ocean is largely responsible.

Boreal forest covers less than half of the island, with the best stands in the western lowlands and the north central valleys. Poor drainage, resulting from recent glaciation, and elevation restrict its extent. Where the altitude exceeds 1200 ft (365 m), the forest gives way to barrens—extensive areas of tundra—in the west and southwest.

Greenland. The world's largest island, Greenland extends over 1600 mi (2575 km) from Cape Farewell (59°46′N) at the south to Cape Morris Jesup (83°39′N) at the north, the latter being the nearest land to the North Pole. Its greatest width (77°N) is just under 700 mi (1130 km). Five-sixths of the surface is buried beneath the largest remaining land ice sheet in the Northern Hemisphere, and numerous smaller ice caps and glaciers occur as separate bodies around its margin. In profile the ice sheet is similar to a flat shield rising gently to form three broad domes, the highest of which exceeds 10,000 ft (3050 m). A maximum thickness of 7000 ft (2130 m) of ice has been estimated, but more work has to be done on this as well as on the suggestion that the underlying surface may consist of several separate islands rather than just one. Occasional peaks (nunatak) project through the ice sheet near its margin. Tongues of ice from the main glacier descend to the sea at many points. Most of the icebergs in the North Atlantic originate from such glaciers in southwest Greenland.

The ice-free margins are widest in the southwest and on the northeast coasts, although access to the

latter is impeded by the arctic pack ice (storis) in the southward-moving East Greenland Current. All the land has been intensely glaciated, except the northernmost (Peary Land), which probably received insufficient precipitation. The margins include alpine mountains, plateaus (particularly in areas of basalt bedrock in the central areas on both east and west coasts), and lowlands. The skaergaard, a swarm of low islands and reefs, is very prominent along the southwest coast. The shore is much indented by bays and fiords.

The full significance of the ice cap in the climatic pattern is still not known. Temperatures there usually range from 27 to −49°F (−3 to −45°C) through the year. One associated phenomenon is the outward movement of strong katabatic winds from the ice margins which often produce a foehn or chinook effect on the valleys through which they are channeled. Climatic conditions in the ice-free margins are extremely variable, and the local complex topography has a great influence; for example, the inner parts of the fiords usually are warmer in summer than the outer, and colder in winter. Mean winter temperatures range rather uniformly on the west coast, from a February mean of 18°F (−8°C) at Ivigtut in the southwest to −20°F (−29°C) at Smith Sound in the northwest. During the summer there is much less contrast, with July mean temperatures of 50°F (10°C) in the south and 35–40°F (2–4°C) in the north. Precipitation, mainly as snow, decreases rapidly from 46 in. (1170 mm) at Ivigtut in the south to less than 10 in. (254 mm) north of 69°N. Notable climatic change has occurred in Greenland and its adjacent seas within historical times. Fog is common through the summers, especially near the broken sea ice. The island becomes ice-locked in winter except for part of the southwest coast. Natural vegetation of the ice-free areas is essentially tundra. Plant growth reaches its maximum development in the inner parts of the fiords, particularly in southwest Greenland. Five varieties of arboreal growth occur in the latter area, and in the Julianehaab district some birches grow to the height of 20 ft (6 m). Copses up to 7 ft (2 m) high occur in favored areas as far north as Disko Bay, but the grasses, mosses, and stunted growth of the true tundra are more typical and commonly make up the only vegetation over much of the island. *See* ARCTIC OCEAN; ASIA; EUROPE; NORTH AMERICA.

William C. Wonders

Bibliography. V. D. Aleksandrova, *The Arctic and Antarctic*, 1980; T. Armstrong et al., *The Circumpolar North*, 1978; J. D. Ives and R. G. Barry (eds.), *Arctic and Alpine Environments*, 1974; W. C. Wonders (ed.), *The Arctic Circle*, 1976.

Arctic Circle

The parallel of latitude approximately 66°30′ north of the Equator, or 23.5° from the North Pole. Named for the northern constellation Bear, the Arctic Circle has the same angular distance from the Equator as the inclination of the Earth's axis from the plane of the ecliptic. Thus, when the Earth in its orbit is at the Northern Hemisphere summer solstice, June 21, and the North Pole is tilted 23.5° toward the Sun, the Sun's rays extend beyond the pole 23.5° to the Arctic Circle, giving that parallel 24 h of sunlight. On this same date the Sun's rays at noon will just reach the horizon at the Antarctic Circle, 66°30′ south. The highest altitude of the noon Sun at the Arctic Circle is on June 21, when it is 47° above the horizon.

Within the Arctic Circle the Sun remains above the horizon continuously for 24 h at the longest period. However, with twilight considered, it remains daylight or twilight continuously for about 5 months. Twilight can be considered to last until the Sun drops 18° below the horizon. *See* MATHEMATICAL GEOGRAPHY; SOLSTICE.

Van H. English

Bibliography. A. H. Strahler and A. Strahler, *Introducing Physical Geography*, 4th ed., 2005; H. Veregin (ed.), *Rand McNally Goode's World Atlas*, 21st ed., 2004.

Arctic Ocean

The Arctic Ocean is an enclosed ocean, connected to the Pacific Ocean through Bering Strait and Bering Sea, and to the Atlantic Ocean through Fram Strait and the Barents, Greenland, and Norwegian seas (**Fig. 1**). The Arctic Ocean is the central region including the marginal seas whose water characteristics are both arctic in nature and quite distinct from those of the Atlantic and Pacific oceans. It includes the area between Bering Strait on the Pacific side and Fram Strait on the Atlantic side together with the Canadian Archipelago and seas along the continental shelf, and generally includes the Barents Sea. Usually it does not include the Norwegian, Greenland, or Bering seas. *See* ATLANTIC OCEAN; BERING SEA; PACIFIC OCEAN.

Expeditions. At the end of the nineteenth century, F. Nansen noted that driftwood caught ice formed near Siberia and exited the Arctic Ocean into the North Atlantic Ocean. In 1893, in an attempt to learn about the oceanography of the Arctic Ocean and to reach the North Pole, Nansen allowed his vessel, *Fram*, to be frozen into the ice pack north of Siberia. *Fram* drifted across three major basins of the Arctic Ocean, though it failed to drift far enough north for Nansen to claim the North Pole. The expedition gave the first real insight into the nature of the Arctic Ocean, obtaining a vast amount of knowledge from an ocean region where virtually none had existed before. In intervening years, almost all oceanography was done from isolated ice camps. Not until 1987 did a surface research vessel, FS *Polarstern*, exceed the farthest north point in the *Fram* drift path, and then by only a few miles. On September 7, 1991, IB *Oden* and FS *Polarstern* became the first nonnuclear-powered ships to reach the North Pole. The expedition proved the feasibility of carrying out standard modern oceanographic measurements, common in other oceans, throughout most of the Arctic Basin. Since then, several expeditions have

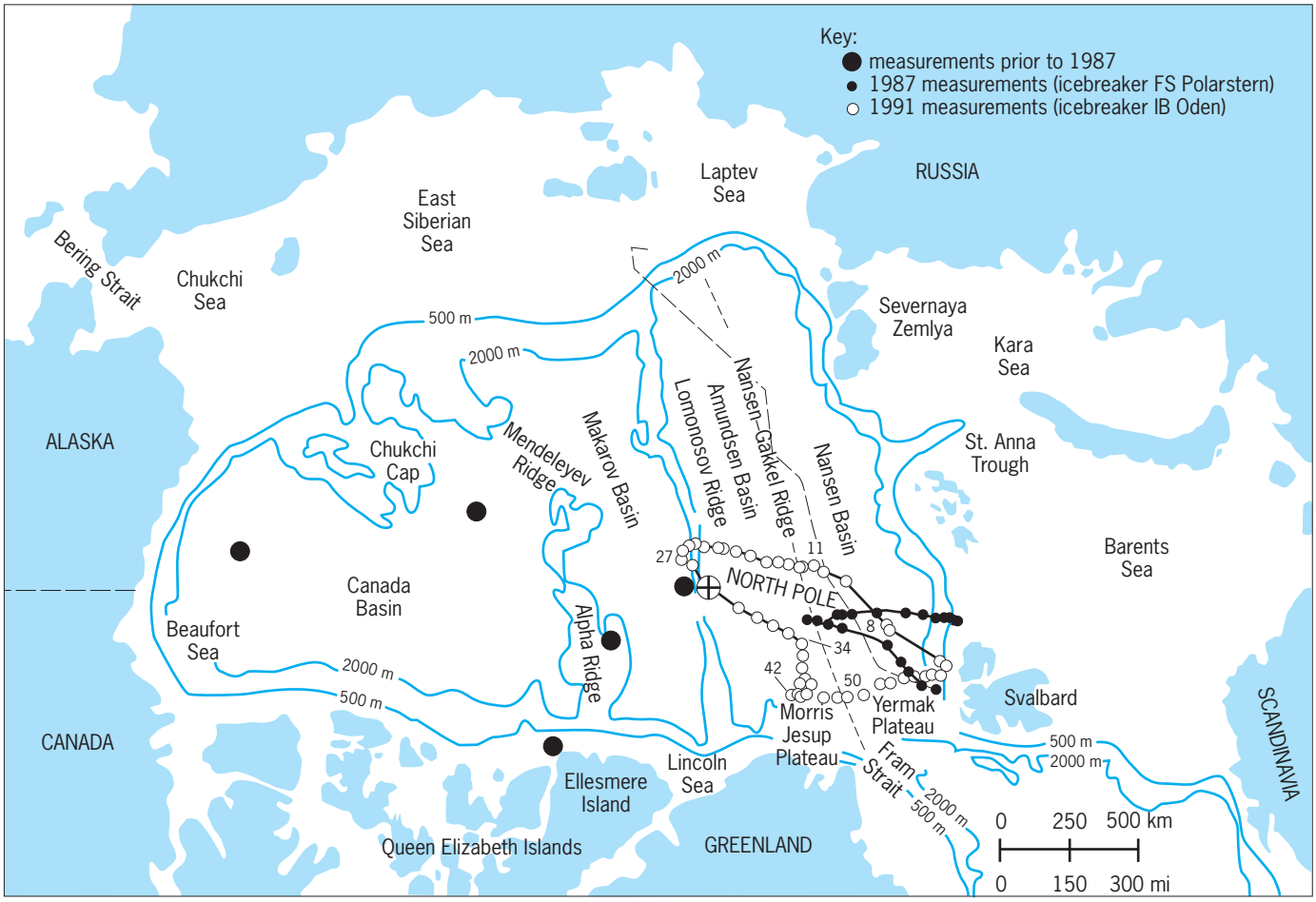

Fig. 1. Arctic Ocean basins.

traversed many of the key regions of the Arctic Ocean (**Fig. 2**).

General features. The Arctic Ocean, including its marginal seas and the Canadian Archipelago, has a total area of approximately 9.4×10^6 km^2 (3.6×10^6 mi^2), with the continental shelves, mostly on the Eurasian side, constituting about one-third of this area. The continental shelves, cut by several canyons, are typically 800 km (480 mi) wide on the Eurasian side and about 100 km (62 mi) wide on the North American side. The central Arctic Ocean Basin is separated by the subsurface Lomonosov Ridge, which extends from North America to the Eurasian continent almost through the North Pole, into the Eurasian Basin on the Atlantic side and the Canadian Basin on the Pacific side. Each of these two basins is subdivided into two basins by the less prominent subsurface Nansen-Gakkel and Alpha-Mendeleyev ridges, the Nansen and Amundsen basins constituting the Eurasian Basin, and the Makarov and Canada basins constituting the Canadian Basin. The deepest parts of the Arctic Ocean, which lie on either side of the Lomonosov Ridge, are more than 4000 m (13,000 ft). The depth at the North Pole is 4270 m (14,000 ft). *See* BASIN; CONTINENTAL MARGIN; MARINE GEOLOGY; SUBMARINE CANYON.

For many months of the year, daylight is weak or absent. Air temperatures over the Arctic Ocean range from a little above freezing in summer to an average of −35°C (−31°F) in January. Temperatures as low as −50°C (−58°F) have been recorded.

Biological productivity is low in most regions but is high in some parts of the marginal seas (Barents and Bering seas), where commercial fisheries exist. *See* MARINE FISHERIES.

Much of the Arctic Ocean is ice-covered throughout the year, typically with ice 2–3 m (7–10 ft) thick. In the central Arctic Ocean, ice is ever present, moving continually and mostly driven by average wind patterns (**Fig. 3**). As a result of the ice motion, large open cracks, or leads, appear and disappear as the open water freezes. As in regions where the ice is under pressure, large pressure ridges, up to 20 m (66 ft), can be formed. These pressure ridges, which can be a major barrier to icebreakers, may help mix the near-surface waters as the ice moves. Large parts of some of the marginal seas (Chukchi, Laptev, and Kara seas) and some of the Beaufort Sea become mostly ice-free for part of the summer. Much of the Eurasian coastline becomes navigable, if not completely ice-free, in the summer. It represents a major shipping route for communities in northern Russia.

Fig. 2. Icebreaker expedition stations during 1987–1998 show that a substantial region of the Canadian Basin remains to be explored.

Seasonal production of ice over the large continental shelves strongly influences the structure and properties of the water column in the central regions.

Bering Strait is shallow (45 m; 150 ft), restricting the flow of water from the Pacific Ocean into the Arctic Ocean. Water from the Atlantic Ocean flows into the Arctic Ocean through Fram Strait and the Barents Sea. Deeper water exchanges through Fram Strait, 2500 m (8200 ft) deep, though communication with the main part of the North Atlantic is restricted to the upper 800 m (2600 ft) by the Greenland-Scotland Ridge farther to the south. Water flows out of the Arctic Ocean through Fram Strait and the Canadian Archipelago. There is a large fresh-water flow into the Arctic Ocean. About 90% of river input comes from Siberian rivers and about 10% from North

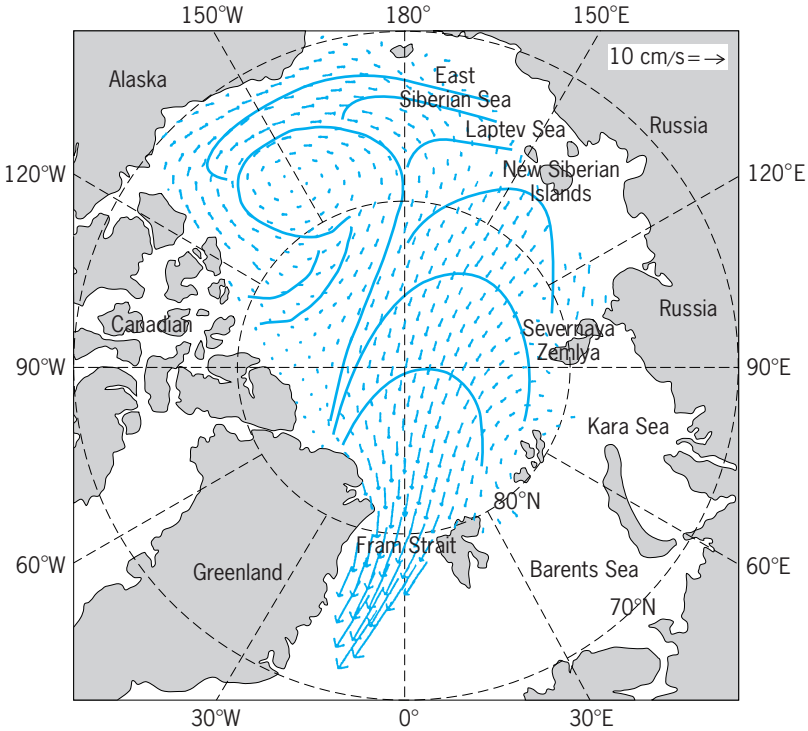

Fig. 3. Mean drift of ice in the Arctic Ocean. The data were obtained from ice station drifts and satellite tracking buoys placed on the ice. (*After S. Pfirman et al., Reconstructing the origin and trajectory of drifting Arctic sea ice, J. Geophys. Res., 102:12575–12586, 1997*)

American rivers, primarily the Mackenzie River. An unassessed amount comes from other runoff land drainage.

Water masses and circulation. Up until the early 1980s, conventional oceanography with research ships was not feasible because of severe conditions and extensive ice cover. While the former Soviet Union had an ambitious oceanographic program collecting data from ice stations and using long-range helicopters, only a limited amount of these data were available to western scientists. Icebreaker expeditions have added vastly more oceanographic data (Fig. 2), and the release of Russian historical ice station data is under way. These new data have given a much more complete picture of the Arctic Ocean and processes operating in it.

General properties. The Arctic Ocean has well-defined layers throughout most of its volume: the Polar Mixed Layer; the halocline, water with a high salinity and density gradient; the Atlantic Layer; intermediate-depth water sometimes called Upper Polar Deep Water; and deep water (see **table**).

Ice formation over the large continental shelves is responsible for many of the characteristics of the different water masses. When ice forms, brine is rejected from the ice and mixes with underlying water, producing dense (saline) water. This water flows off the shelves at many locations, forming plumes that sink through and entrain the surrounding water. The plumes influence the characteristics of all of the water masses below the Polar Mixed Layer. *See* SEA ICE; SEAWATER.

Water mass identification and circulation patterns. Salinity and temperature, as well as several natural chemical constituents (for example, nutrients, oxygen), are used to characterize water masses, to determine their sources, and to trace their circulation patterns. Anthropogenic materials (for example, chlorofluorocarbons, tritium) that were introduced into the ocean during the twentieth century are used to provide information on residence or exchange times of water masses for waters that are renewed in about 150 years or less. Naturally produced carbon-14 dates older water. *See* RADIOCARBON DATING.

The near-freezing Polar Mixed Layer lies over the entire central deep basin. Its salinity varies from relatively fresh values near Bering Strait to quite high values near Fram Strait. The salinity reflects the relative amounts of the constituent source waters (Atlantic, Pacific, river runoff, and sea ice meltwater). The mean circulation of Pacific origin waters in the Polar Mixed Layer as indicated from nutrient distributions enters through Bering Strait and exits through the Canadian Archipelago and Fram Strait (**Fig. 4**). Fresh water mainly from Siberian rivers enters this flow and exits with it. Atlantic origin water in the Polar Mixed Layer flows along the Eurasian coast toward the Canadian Basin, where it becomes overlain with fresher water from the Pacific Ocean and from rivers.

Characteristics of water masses in the Arctic Ocean*

Water mass	Depth range, m	Salinity†	Temperature	Residence time, years
Polar Mixed Layer	0–50	30–33	Near −1.8°C	5–10
Halocline				
Upper Halocline Water	50–130	Near 33.1	Near −1.6°C	5–15
Lower Halocline Water	20–200	Near 34.2	Near 1.2°C	5–15
Atlantic Layer	200–600	34.8–34.95	0–3°C	~25
Upper Polar Deep Water	600–1700	Near 34.9	0 to −0.5°C	20–40
Deep Water				
Eurasian Basin	1700–4400	34.92	−0.94°C	50–150
Canadian Basin	1700–4000	34.94	−0.5°C	~700

*Properties vary with depth and over the Arctic Basin. Values are typical of the respective water masses.
†Salinity values are measured in terms of the relative conductivity of seawater compared to a standard.

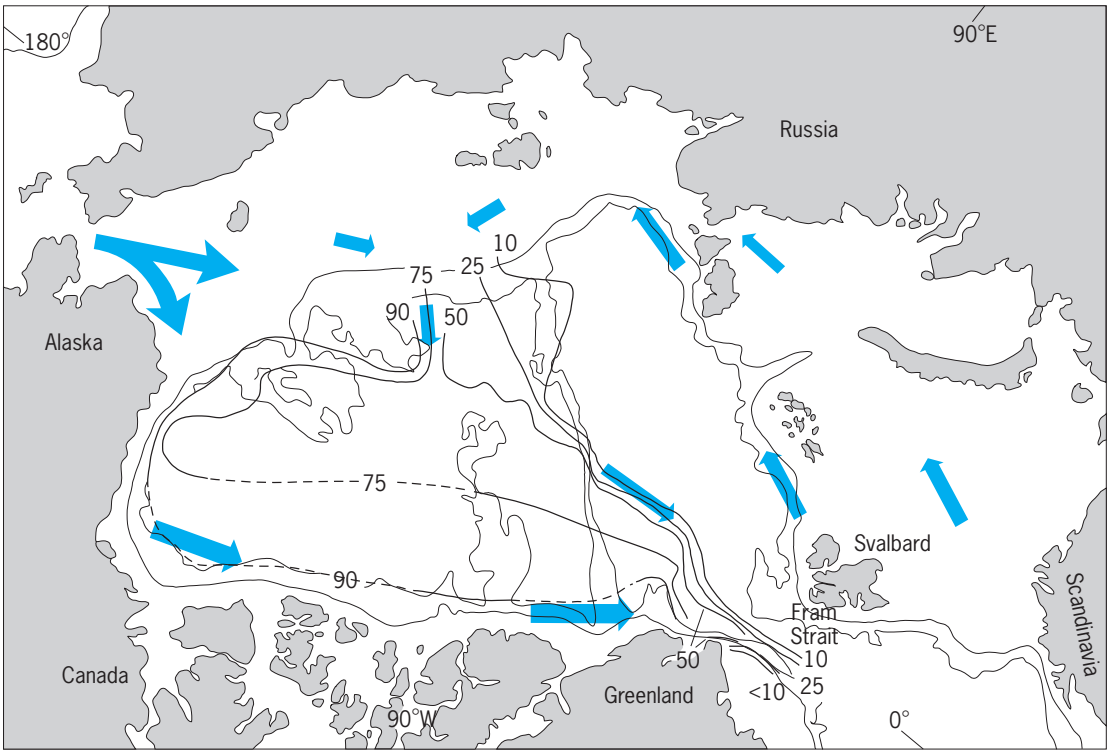

Fig. 4. Surface water (top 30 m; 100 ft) circulation determined from natural chemical tracers. (*After E. P. Jones et al., Distribution of Atlantic and Pacific waters in the upper Arctic Ocean: Implications for circulation, Geophys. Res. Lett., 25:765–768, 1998*)

Two distinct water masses make up the halocline. In the Canadian Basin, the shallower part of the halocline is characterized by high nutrient concentrations. This part of the halocline is mostly of Pacific origin but is modified probably in the Chukchi-East Siberian Sea region. It is confined largely to the Canadian Basin and flows out through the Canadian Archipelago. A narrow stream is seen flowing by the Morris Jesup Plateau and exiting through Fram Strait. Throughout both the Eurasian and Canadian basins is deeper halocline water. In the Eurasian Basin, it is found immediately under the Polar Mixed Layer; in the Canadian Basin, it lies under the less dense water originating mostly in the Pacific Ocean. This deeper halocline water is formed from water of Atlantic origin modified by melting sea ice as the Atlantic water enters though Fram Strait. Both parts of the halocline receive some fresh water from rivers and from sea ice meltwater. The circulation pattern of the whole halocline seems to generally follow that of the Atlantic Layer (see below), though it is not as definitively traced.

Warmer water, with core temperatures between 0.5 and about 2.5°C (33 and 36.5°F), has historically been denoted the Atlantic Layer. This water is warmest as it enters through Fram Strait near Svalbard. As it circulates around the Arctic basins, it cools to about 0.5°C (33°F) before returning to and exiting through Fram Strait (Fig. 4). Mixing with surface water that would otherwise allow heat exchange with the atmosphere is strongly inhibited by the density gradient of the halocline. The cooling of the Atlantic Layer is largely a result of entrainment of cold water from shelf plumes, which in turn warm and transport heat to the deeper layers as they sink.

The Upper Polar Deep Water extends from beneath the Atlantic Layer to the sill depth of the Lomonosov Ridge, about 1700 m (5575 ft) deep. It is characterized by increasing salinity and decreasing temperature with depth, both likely also influenced by shelf plumes. Its circulation is similar to that of the Atlantic Layer (**Fig. 5**).

The Eurasian Basin Deep Water is slightly colder and fresher than the Canadian Basin Deep Water. The differences are small, but have been hard to explain because the major source of the Canadian Basin Water was presumed to be cold water from the Eurasian Basin overflowing the Lomonosov Ridge. The higher temperature especially was a puzzle. An explanation again lies in the shelf plumes. Since the continental shelves of the Canadian Basin are generally shallower than those of the Eurasian Basin, shelf plumes dense enough to reach deep waters in the Canadian Basin would have to pass through the warm Atlantic Layer, entraining warmer water and carrying with them both heat and salt to the deep water. In the Eurasian Basin, the shelves are generally deeper. Shelf plumes reaching the deep waters in the Eurasian Basin bypass the Atlantic Layer by flowing under it in deep canyons such as the St. Anna Trough. Very little can be said about the circulation of the deep waters except that it is slow.

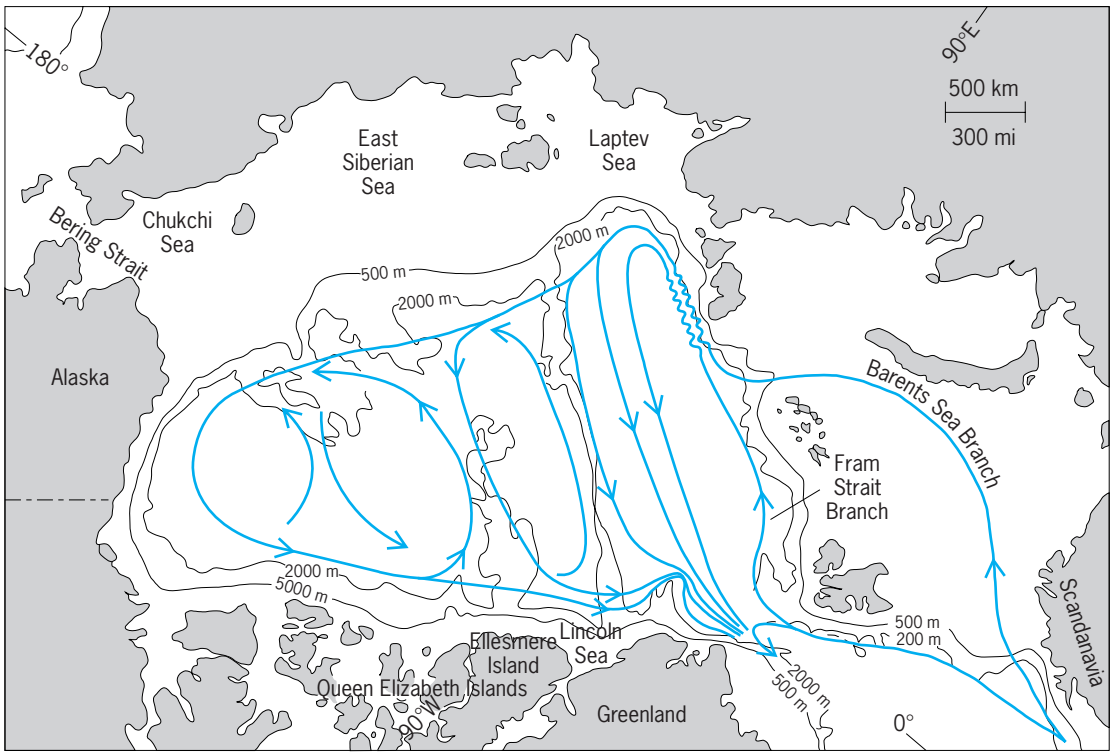

Fig. 5. Mean circulation of the Atlantic Layer and Upper Polar Deep Water. (*After B. Rudels et al., On the intermediate depth waters of the Arctic Ocean, in O. M. Johannessen, R. D. Muench, and J. E. Overland, eds., The Polar Oceans and Their Role in Shaping the Global Environment: The Nansen Centennial Volume, American Geophysical Union, Washington, DC, 1994*)

Climate change affecting the ocean. Up until the 1980s, the concept of the Arctic Ocean was of a quiet, steady-state ocean not interacting much with the global ocean. At the end of the twentieth century, a very different picture emerged. Changes include warming of the Atlantic Layer and a retreat and diminishing of the cold halocline. There has also been a major shift in the boundary between Atlantic and Pacific waters in the surface Polar Mixed Layer and halocline waters. The new observations, together with evidence of changing circulation patterns in the near-surface waters, show that the Arctic Ocean is not a quiescent environment remaining in a quasi-steady state, but exhibits as large a variability as other oceans.

In 1990, temperatures in the Atlantic Layer in the Eurasian Basin were observed to be about 1°C (2°F) warmer than those reported earlier. In 1993, warmer Atlantic Layer water (1.4°C; 34.5°F) was observed in the Canadian Basin over the Mendeleyev Ridge, far past the historical boundary of warmer part of the Atlantic Layer near the Lomonosov Ridge. At the North Pole, where the temperature maximum was historically near 0.8°C (33.4°F), the Atlantic Layer has become warmer and thicker, moving higher in the water column. By contrast, in the Canadian Basin north of Alaska, the temperature, salinity, and density were almost identical in 1985 and 1997, and are consistent with past findings. The type of change observed elsewhere had not occurred in this region.

Climate models predict that global warming as a result of the buildup of greenhouse gases (carbon dioxide, methane, and others) in the atmosphere will be enhanced in polar regions. A significant finding regarding the warming in the Atlantic Layer is that it can be related to warming in West Spitsbergen current off South Cape (Svalbard) and warming of the source waters in the Norwegian Sea. The explanation for this seems to lie in a shifting weather pattern over the North Atlantic Ocean rather than the warming predicted by climate models. *See* CLIMATE MODELING; GREENHOUSE EFFECT.

In addition to warming, two major changes have occurred in Arctic Ocean waters. There has been a diminishment of the density gradient in the halocline, and there has been a shift in the boundary between waters of Atlantic and of Pacific origin in the Polar Mixed Layer and halocline. The halocline is the barrier to heat transport from the relatively warm Atlantic Layer to the surface, where the ice is. If the halocline were to disappear or be much diminished, there is more than enough heat to melt all of the Arctic Ocean ice. Observations in the Eurasian Basin show an increase in near-surface salinity, with the halocline deepening and disappearing. The barrier depends, initially at least, on where fresh water from rivers goes. Higher salinity near surface water suggests that there has been a change in where the river runoff flows as it enters the Arctic Ocean.

Whether the observed changes are due to a large, but normal, fluctuation or are part of a trend is not yet known. Higher temperatures in the Atlantic Layer

or Norwegian Sea would not, by themselves, lead to a shift in the boundary between the Atlantic and Pacific domains or to the changes in the halocline. Most striking, the Polar Mixed Layer in the interior of the Eurasian Basin is more saline than historical data (34.3 relative conductivity), and it is deeper (140 m; 460 ft) and reaches down to the thermocline at the top of the Atlantic Layer. In the near-surface water in the Eurasian and Makarov basins, there is less fresh water. This points to the river runoff into the Arctic Ocean being redistributed or decreased.

Global climate change. There is the question of whether changes in the ocean-atmosphere system of the Arctic will have global change implications. Two ways in which the Arctic Ocean could influence the global oceans and climate change are through the transport of fresh water and through changes to the global ocean "conveyor belt." Fresh-water transport is basic to global atmospheric and oceanic climate systems. The Arctic Ocean is a major pathway for the return flow of fresh water from the North Pacific Ocean, where precipitation exceeds evaporation, to the North Atlantic Ocean, where evaporation is greater than precipitation. Fresh water is returned to the Atlantic Ocean via rivers flowing into the Arctic Ocean and from the North Pacific via the relatively fresh flow into the Arctic Ocean through Bering Strait.

The global ocean conveyor belt is a highly schematic depiction of how the ocean operates in a climate context. The conveyor belt depends on processes that make surface water dense enough to sink. In polar regions, this happens in winter when surface waters cool and brine is released during ice formation. The newly formed deep water flows toward the Equator, with upwelling in tropical regions, where it becomes warm. Of the heat transported from tropical to northern regions, the atmosphere accounts for about one-half and the northward flowing warm water carries the other half. Most portrayals show the northern "terminus" of the conveyor belt to be in the Nordic seas. Recent work suggests, however, that more than 30% of the water contributing to the deep circulation from the Arctic region could be the result of processes occurring within the Arctic Ocean. *See* OCEAN CIRCULATION.

A change in the thermohaline circulation brought about by changes in fresh-water budgets or in ice formation in the Arctic Ocean could have a significant impact on this heat transport. On a longer time scale, such changes in the past are thought to have brought on ice ages, for example. In view of the fact that variability in the Arctic has long seemed to have some special involvement in global change, there is still the question of its role in global climate change. E. P. Jones

Bibliography. W. S. Broecker, Chaotic climate, *Sci. Amer.*, 273:5, November 1995; S. G. Gorshkov, *World Ocean Atlas*, vol. 3, Pergamon Press, Oxford, 1983; O. M. Johannessen, R. D. Muench, and J. E. Overland (eds.), *The Polar Oceans and Their Role in Shaping the Global Environment: The Nansen Centennial Volume*, American Geophysical Union, Washington, DC, 1994; F. Nansen, *Farthest North*, vols. 1 and 2, George Newnes, London, 1898.

Arcturus

The brightest star in the northern sky, apparent magnitude −0.05, also known as α Boötis. It is a yellow giant star of spectral type K1.5, one of the nearest giants to the Earth at a distance of 11.25 parsecs (2.16×10^{14} mi or 3.47×10^{14} km). Unlike the Sun, which is currently converting hydrogen into helium in its core, Arcturus has already exhausted its central hydrogen and has evolved away from the main sequence. It is approximately 25 times larger in diameter than the Sun, and more than 100 times more luminous. Its effective temperature is estimated to be 7700°F (4300 K). *See* BOÖTES; SPECTRAL TYPE; STELLAR EVOLUTION.

Arcturus has a large space motion relative to the Sun, and it is a member of the high-velocity group of stars known as Population II, associated with the halo of the Milky Way Galaxy. The component of its space velocity in the direction of galactic rotation is −75 mi/s (−120 km/s) relative to the local standard of rest, implying it moves around the galactic center much more slowly than the Sun. Its galactocentric orbit is highly eccentric, while the Sun's is essentially circular. The pattern of chemical abundances is also quite typical of Population II; heavy elements such as iron and nickel are about three times more deficient compared to hydrogen than they are in the Sun. while light elements such as oxygen, magnesium, and silicon are enhanced in about the same degree as in the halo stars. Water vapor has been detected in the photosphere of the star. *See* STELLAR POPULATION.

Precise measurements have revealed small variations in the radial velocity of Arcturus with a period close to 2 days and an amplitude of about 660 ft/s (200 m/s), which are believed to be due to pulsation. Variations with much longer periods of 1 or 2 years have also been reported, and could be related to convective motions in the star's photosphere or the presence of surface features moving across the disk of the star as it rotates around its axis. Similar variations have also been found in other giant stars. Brightness variations at the level of a few percent have been seen in Arcturus, and are presumably also related to pulsation. *See* STAR. David W. Latham

Area

The superficial contents of a geometrical figure of two dimensions. The area of any rectangle or square is the product of two adjacent sides, one of which may be called the base and the other the altitude. In general, any line segment that partially bounds a plane geometric figure may be called a base if its line does not separate the figure, and a perpendicular

Area formulas

Figure	Formula
Triangle	$hb/2$, where h = altitude, b = base; $\sqrt{s(s-a)(s-b)(s-c)}$, where $s = (1/2)(a+b+c)$, and a, b, and c are sides of the triangle
	ab, where a and b are adjacent sides
Square	a^2, where a = side
Parallelogram	$ab \sin \theta$, where a and b are adjacent sides, and θ is the angle between the sides
Trapezoid	$(1/2)(a+b)h$, where a and b are the parallel sides, and h is the altitude
Quadrilateral	$(1/2)ab \sin \theta$, where a and b are the diagonals, and θ is the angle between them
Regular polygon	$(1/4)nl^2 \cot 180^\circ/n$, where n is the number of sides, each of length l
Circle	πr^2, where r = radius
Ellipse	πab, where a and b are semiaxes
Sphere	$4\pi r^2$, where r = radius
Spherical triangle	$(A+B+C-\pi)r^2$, where A, B, and C are angles (radians), and r is the radius

drawn to the base line from one of its points at greatest distance may be called the altitude. The area of a parallelogram is equal to the product of its base times its altitude. The area of a triangle is one-half the product of its base times its altitude. The area of a trapezoid is equal to one-half the product of the sum of its parallel sides (bases) times its altitude. Some area formulas are given in the **table**. *See* EUCLIDEAN GEOMETRY. J. Sutherland Frame

Arecales

An order of flowering plants, division Magnoliophyta (Angiospermae), of the subclass Arecidae in the class Liliopsida (monocotyledons). The name of the subclass is derived from the ordinal name. The order consists of the single family Arecaceae (Palmae), the palms, with more than 200 genera and nearly 3000 species, largely confined to tropical and subtropical regions. The order Arecales has also been called Palmales or Principes.

Most palms are trees with an unbranched trunk and a terminal crown of large leaves. Less common types are scandent, with scattered leaves, or more or less acaulescent, with the leaves arising from the ground. The leaf has a blade, petiole, and sheath; the sheath is open or closed, but in any case it fully encircles the stem at the base. The blade as a whole is usually pinnately compound, less often palmately compound or merely lobed. It is pinnately or less often palmately veined, but the individual pinnae or segments have more or less parallel veins and are plicate between the veins. The plicate structure of all palm leaves relates to a complex ontogeny shared only by some of the Cyclanthaceae (Cyclanthales), in which new tissue continues to develop along the folds after the flanking tissues have matured. *See* ARECIDAE; CARNAUBA WAX; COCONUT; CYCLANTHALES; DATE; LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM; VEGETABLE IVORY. Arthur Cronquist

Arecidae

A subclass of the class Liliopsida (monocotyledons) of the division Magnoliophyta (Angiospermae), the flowering plants, consisting of four orders (Arecales, Cyclanthales, Pandanales, and Arales), five families, and nearly 6000 species. Except for the highly reduced family Lemnaceae (Arales), they have an inflorescence of usually numerous, small flowers, generally subtended by a prominent spathe and often aggregated into a spadix. Except in the Araceae (Arales), the endosperm seldom contains much starch.

More than 80% of the species have broad, petiolate leaves that do not have the typical parallel venation commonly associated with monocotyledons, and more than half of the species are arborescent (likewise an unusual character among the monocotyledons). *See* ARALES; ARECALES; CYCLANTHALES; LILIOPSIDA; MAGNOLIOPHYTA; PANDANALES; PLANT KINGDOM. Arthur Cronquist; T. M. Barkley

Arenaceous rocks

The arenaceous rocks (arenites) include all those clastic sedimentary rocks whose particle sizes range from 2 to 0.06 mm, or if silt is included, to 0.004 mm. Some arenites are composed primarily of carbonate particles, in which case they are called calcarenites and grouped with the limestones. Some oolitic iron ores and glauconite beds are properly classified as arenites. But the vast majority of arenites are commonly called sandstones, and the two words are almost synonymous. *See* CALCARENITE; GRAYWACKE; OOLITE; SANDSTONE; SEDIMENTARY ROCKS. Raymond Siever

Bibliography. H. Blatt and R. Tracy, *Petrology: Ingenious, Sedimentary, and Metamorphic*, 1995; M. E. Tucker, *Sedimetary Petrology: An Introduction to the Origin of Sedimentary Rocks*, 3d ed., 2001.

Argentiniformes

An order of teleost fishes in the superorder Protacanthopterygii. These fishes, collectively called the argentines or herring smelts, were elevated from Osmeriformes (which are treated herein as a superfamily of the Salmoniformes) to ordinal rank.

Argentiniforms are identifiable by the following combination of characters: body variable, elongate and cylindrical in some to short and compressed

Argentiniformes features

Argentinoidei (families Argentinidae, Opisthoproctidae, and Microstomatidae)	Alepocephaloidei (families Platytrocidae, Bathylaconidae, and Alepocephalidae)
Eyes tubular or not tubular	Eyes not tubular
Adipose fin usually present	Adipose fin absent
Dorsal fin inserted near center of body	Dorsal fin inserted well back on body
Maxillary and premaxillary, when present, toothless	Upper jaw usually with teeth
Mouth usually small	Mouth usually large
Swim bladder present or absent	Swim bladder absent
Eggs small, development gradual	Eggs large, development direct
12 genera, 61 species	38 genera, 103 species

in others; jaw mechanism and dentition greatly reduced; premaxillary absent in some species, if present lacks teeth; maxillary usually toothless; eyes variously directed, some species with tubular eyes; some species with specialized light organs, which may be associated with tubular eyes; no serial photophores; adipose fin present or absent; caudal fin forked; swim bladder, when present, physoclistous (not attached to gut). In addition, the argentines have a crumenal organ (paired branchial pouches in which food particles are trapped by large interlocking toothed gill rakers).

Two distinct groups (suborders in this classification), Argentinoidei and Alepocephaloidei, are identified by the characters listed in the **table**.

Argentinoidei. This suborder consists of the families Argentinidae, Opisthoproctidae, and Microstomatidae.

The Argentinidae (argentines or herring smelts) is found in the Atlantic, Indian, and Pacific oceans, and consists of two genera, with about 19 species.

The Opisthoproctidae (barreleyes or spookfishes) is found in tropical to temperate Atlantic, Indian, and Pacific oceans, and consists of six genera with about 10 species.

The Microstomatidae (includes Bathylagidae, deep-sea smelts) is found in all seas, mostly ranging from subarctic to Antarctic waters, and consists of four genera, with about 32 species. *Bathylagus*, the largest genus, contains about 15 species.

Alepocephaloidei. This suborder consists of the families Platytrocidae, Bathylaconidae, and Alepocephalidae.

The Platytrocidae (searsides) is found in all oceans, but is absent from the Mediterranean Sea. It consists of 13 genera, with 37 species.

The Bathylaconidae (no common name) is found in circumtropical waters, and consists of two genera, with three species.

The Alepocephalidae (slickheads) is found in all oceans, and consists of 22 genera, and no less than 60 species.

Opisthoproctidae specializations. Some of the most bizarre fishes in all the seas of the world are the family Opisthoproctidae. As a family, most of the species have tubular eyes; pectoral fins on the lower side of the body; pelvic fins posteriorly placed; and parietal bones that do not meet on the midline of the skull. Most lack a swim bladder. Most species are long and subcylindrical, but one genus, *Opisthoproctus*, has a foreshortened, deep and compressed body and sole (an elongate forward projection on the underside of the abdomen that may act as a reflector for the light organ inside the body near the anus) covered with large deciduous cycloid scales. The tubular eyes are directed dorsally and topped with a gelatinous transparent tissue. The interorbital bones are reduced to a mere sliver, exposing the brain through the transparent roof of the skull; the premaxillary is absent; and the maxillary is reduced to a thin scalelike bone which is easily lost, leaving the toothless dentary to occupy almost half the length of the snout. Some of these bathypelagic fishes have never been seen alive and are known from only a few specimens. *See* AULOPIFORMES; SALMONIFORMES. Herbert Boschung

Bibliography. D. P. Begle, Monophyly and relationships of the argentinoid fishes, *Copeia*, 1992(2):350–366, 1992; D. M. Cohen, Argentinoidea, pp. 1–70 in *FWNA Fishes of the Western North Atlantic*, pt. 4, 1964; D. M. Cohen, Argentinidae, pp. 215–216 in M. M. Smith and P. C. Heemstra (eds.), *Smiths' Sea Fishes*, Springer-Verlag, Berlin, 1986; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, 2006; J. R. Paxton and D. M. Cohen, *Argentinidae*, p. 1884 in K. E. Carpenter and V. H. Niem (eds.), *FAO Species Identification Guide for Fishery Purposes: The Living Marine Resources of the WCP*, vol. 3: *Batoid Fishes, Chimaeras and Bony Fishes, Part 1 (Elopidae to Linophrynidae)*, FAO, Rome, 1999; M. Schneider, Argentinidae: Argentinas, pp. 858–859 in W. Fischer et al. (eds.), *Guia FAO para Identification de Especies para lo Fines de la Pesca*, Pacifico Centro-Oriental, 3 vols., FAO, Rome, 1995.

Argillaceous rocks

Clastic sedimentary rocks containing silt- or clay-sized particles that are less than 0.0625 mm and/or clay minerals. The argillaceous rocks (lutites) include shales, argillites, siltstones, and mudstones. They are

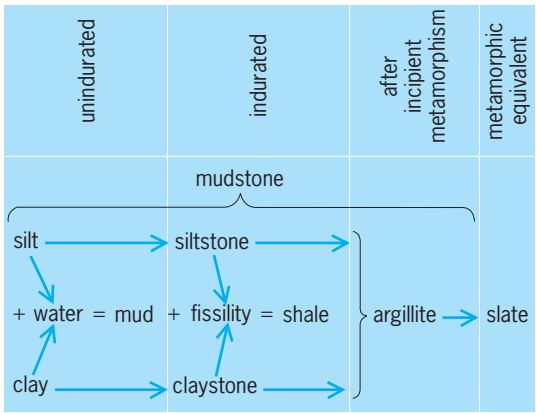

Classification of fine-grained mechanical sediments. (*After F. J. Pettijohn, Sedimentary Rocks, 2d ed., Harper & Row, 1957*)

the most abundant sedimentary rock type, varying according to different estimates from 44 to 56% of the total sedimentary rock column. Claystone is hardened or cemented (indurated) clay, which consists dominantly of fine material of which at least a major proportion is clay mineral (hydrous aluminum silicates). Shale is a laminated or fissile claystone or siltstone, in general more consolidated than claystone. Mudstone is a claystone that is blocky and massive. The term argillite is used for rocks which are more indurated than claystone or shale but not metamorphosed to slate. All these argillaceous rocks are consolidated equivalents of muds, oozes, silts, and clays (see **illus.**). Loess is a fine-grained, unconsolidated, windblown deposit. The term shale has been used by many authors generically to denote all of these types of rock. *See* BENTONITE; CLAY; CLAY MINERALS; LOESS; SEDIMENTARY ROCKS; SHALE. Raymond Siever

Bibliography. H. Blatt and R. Tracy, *Petrology: Ingenious, Sedimentary, and Metamorphic*, 1995; M. E. Tucker, *Sedimetary Petrology: An Introduction to the Origin of Sedimentary Rocks*, 3d ed., 2001.

Argon

A chemical element, Ar, atomic number 18, and atomic weight 39.948. Argon is the third member of group 0 in the periodic table. The gaseous elements in this group are called the noble, inert, or rare gases, although argon is not actually rare. The Earth's atmosphere is the only natural argon source; however, traces of this gas are found in minerals and meteorites. Argon constitutes 0.934% by volume of the Earth's atmosphere. Of this argon, 99.6% is the argon-40 isotope; the remainder is argon-36 and argon-38. There is good evidence that all the argon-40 in the air was produced by the radioactive decay of the radioisotope potassium-40. *See* INERT GASES; PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Argon is colorless, odorless, and tasteless. The element is a gas under ordinary conditions, but it can be liquefied and solidified readily. Some properties of the gas are listed in the **table**. Argon does not form any chemical compounds in the ordinary sense of the word, although it does form some weakly bonded clathrate compounds with water, hydroquinone, and phenol. There is one atom in each molecule of gaseous argon.

Properties of argon

Property	Value
Atomic number	18
Atomic weight (atmospheric argon)	39.948
Melting point (triple point), °C	−189.4
Boiling point at 1 atm pressure, °C	−185.9
Gas density at 0°C and 1 atm (101.325 kPa) pressure, g/liter	1.7840
Liquid density at normal boiling point, g/ml	1.3998
Solubility in water at 20°C, ml argon (STP) per 1000 g water at 1 atm (101.325 kPa) partial pressure of argon	33.6

The oldest large-scale use for argon is in filling electric light bulbs. Welding and cutting metal consumes the largest amount of argon. Metallurgical processing constitutes the most rapidly growing application. Argon and argon-krypton mixtures are used, along with a little mercury vapor, to fill fluorescent lamps. Argon mixed with a little neon is used to fill luminous electric-discharge tubes employed in advertising signs (similar to neon signs) when a blue or green color is desired instead of the red color of neon. Argon is also used in gas-filled thyratrons, Geiger-Müller radiation counters, ionization chambers which measure cosmic radiation, and electron tubes of various kinds. Argon atmospheres are used in dry boxes during manipulation of very reactive chemicals in the laboratory and in sealed-package shipments of such materials.

Most argon is produced in air-separation plants. Air is liquefied and subjected to fractional distillation. Because the boiling point of argon is between that of nitrogen and oxygen, an argon-rich mixture can be taken from a tray near the center of the upper distillation column. The argon-rich mixture is further distilled and then warmed and catalytically burned with hydrogen to remove oxygen. A final distillation removes hydrogen and nitrogen, yielding a very high-purity argon containing only a few parts per million of impurities. A. W. Francis

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; M. Ozima and F. A. Podosek, *Noble Gas Geochemistry*, 2001.

Arhynchobdellae

An order of the class Hirudinea (leeches) which do not have an eversible proboscis, but frequently have three jaws armed with sharp teeth. The blood of these annelids contains hemoglobin. They may be divided into the Gnathobdellae, with jaws, and the Pharyngobdellae, without jaws. *See* HEMOGLOBIN.

Gnathobdellae have bodies which are oval in cross section and have a conspicuous posterior sucker. The anterior sucker does not project beyond the

sides of the body but forms a deep cup on the underside of the head in which the jaws can work to make their incision in the host. This group contains most of the important bloodsucking, leech parasites of humans and other warm-blooded animals. Some have often been used by medical practitioners such as *Macrobdella decora* in North America, *Hirudo medicinalis* in Europe and northern Asia, and *Hirudinaria granulosa* in southern and eastern Asia. The land leeches, such as *Haemadipsa*, are members of this group. They occur in great numbers on vegetation in swamp and jungle areas and attach themselves to passing warm-blooded animals. They make traveling on foot in these areas exceedingly unpleasant, and death from loss of blood or secondary infection is not uncommon. *Haemopis* is an example of a gnathobdellid leech which has lost its bloodsucking habits (**illus.** *a*). Its food consists of small invertebrates such as earthworms, which it ingests whole. The teeth are blunt.

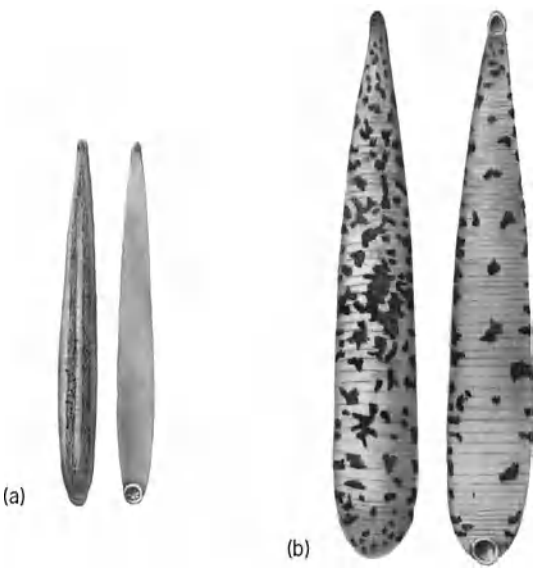

Arhynchobdellae. (*a*) *Haemopis grandis*. (*b*) *Erpobdella punctata*.

Pharyngobdellae are specialized for carnivorous diets and in many cases have completely lost the jaws. They have a strong muscular pharynx which extends nearly half the length of the body. *Erpobdella* (illus. *b*) is common in lakes and streams in the Northern Hemisphere, while *Trocheta* tends to leave the water and forage in moist soil. *See* HIRUDINEA; RHYNCHOBDELLAE.

K. H. Mann

Aries

The Ram, a northern zodiacal constellation, meaning that the path of the Sun and planets passes through it (see **illustration**). In Greek mythology, the golden fleece of this ram was pursued by Jason and the Argonauts, whose ship is memorialized in four current-day constellations in the southern sky. In Johannes Bayer's 1603 star atlas, the first three lettered stars are located in the ram's horns. *See* ZODIAC.

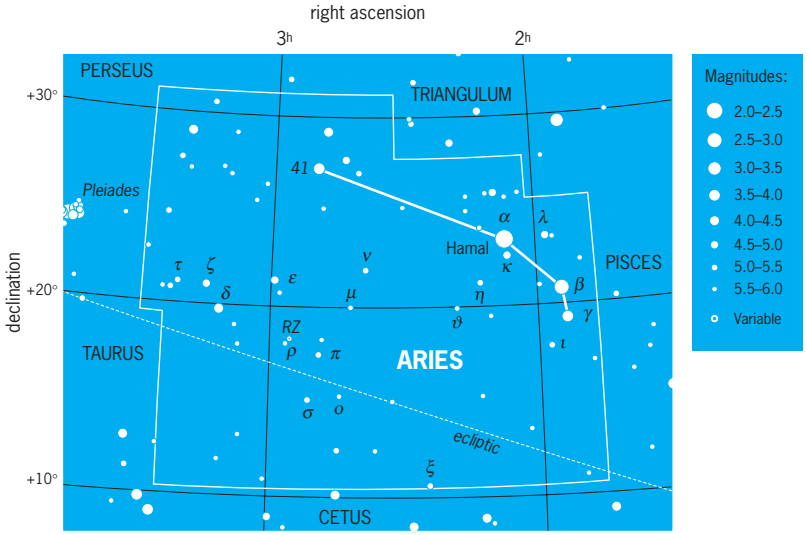

Modern boundaries of the constellation Aries, the Ram. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

When the constellations were organized about 2000 years ago, the point where the ecliptic and the celestial equator cross, the point of the spring equinox, was in Aries, so that point is known as the first point of Aries. (Because of the precession of the equinoxes, that point has now moved into the constellation Pisces.) *See* EQUINOX; PRECESSION OF EQUINOXES.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Aristolochiales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Magnoliidae of the class Magnoliopsida (dicotyledons). It contains only the family Aristolochiaceae, with 8 or 10 genera and about 600 species, most of them in tropical and subtropical regions. Within its subclass the order is marked by the presence of ethereal oil cells, by its uniaperturate or nonaperturate pollen, and especially by its strongly perigynous to epigynous flowers, usually with united carpels, that typically lack petals and have the sepals joined into a highly irregular, corolloid calyx. Many of the species are climbing vines. *Aristolochia* (birthwort or Dutchman's pipe) and *Asarum* are well-known genera of the order.

Several parasitic, nonchlorophyllous families that have been included by many authors in the Aristolochiales are here treated as forming a separate order, Rafflesiales in the subclass Rosidae. *See* MAGNOLIIDAE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; RAFFLESIALES.

Arthur Cronquist

Arithmetic

A branch of mathematics dealing with numbers, operations on numbers, and computation. Arithmetic is useful in solving many practical problems, such as buying, selling, budgets, sports statistics, and measurement. The usual numbers of arithmetic are whole numbers, fractions, decimals, and percents. Beyond the numbers of arithmetic are negative numbers, rational numbers, and irrational numbers. The rational and irrational numbers together constitute the real numbers.

Whole numbers. The whole numbers include the infinite sequence of counting numbers—one, two, three, four, five, . . . —and the number zero. For numbers to ten, a single symbol is used, and for larger numbers a combination of symbols.

Numbers to ten. Numbers to ten are designated with a single digit; the digits are 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. Counting by ones shows "how many." To extend this counting concept and to prepare for operations on the numbers, it is helpful to recognize the numbers as made up of parts, for example to see five on dice as 4 and 1, as 3 and 2, or as 2 and 2 and 1.

Numbers ten and greater. Numbers ten and greater are expressed by using a combination of the ten digits, with the place of the digit indicating the value of the digit. This place-value system, named the Hindu-Arabic numeration system, is now used around the world. *See* NUMBERING SYSTEMS.

In a multidigit numeral, the value of each place from right to left is a successive power of ten and the total values for all places are combined or added. For example, the meaning of 1,234 is given in (1).

$$1 \text{ thousand} + 2 \text{ hundreds} + 3 \text{ tens} + 4 \quad \text{or}$$

$$1 \times 10^3 + 2 \times 10^2 + 3 \times 10 + 4 \qquad (1)$$

Each three places denotes a period. The first period on the right is the ones or units period, but this name is usually omitted. The next six period names are thousand, million, billion, trillion, quadrillion, and quintillion. Numbers within each period are read as if they were in the ones period and then the period name is given. For example, 123,456,789,721,042 is read "one hundred twenty-three trillion four hundred fifty-six billion seven hundred eighty-nine million seven hundred twenty-one thousand forty-two."

Operations. The basic operations are addition (+) and multiplication (× or ·), with subtraction (−) and division (÷) defined, respectively, by using addition and multiplication.

If two numbers, a and b, are combined or added, the result is a number, c, called the sum. In the example, $4 + 6 = 10$, 4 and 6 are addends and 10 is the sum. The whole amount, 10, is the result of combining two parts, 4 and 6.

If a given number, n, sets of objects with the same number in each set, r, are combined, then multiplication of n and r is the total number of objects. In $3 \times 4 = 12$, 3 and 4 are factors and 12 is the product (**Fig. 1**).

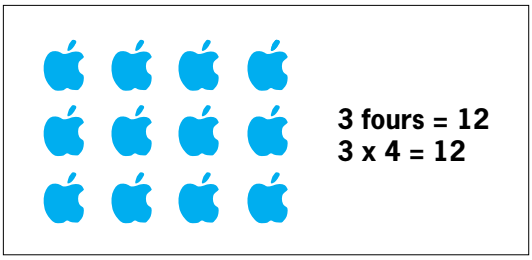

Fig. 1. One meaning of multiplication.

The multiplication 3×4 can also be thought of as the number of members in a rectangular array with 3 rows and with 4 objects in each row. Another use is to find the number of ways that 3 objects can be paired with 4 objects, for example, matching 3 pairs of trousers and 4 shirts. Yet another use of multiplication is to find the area of a rectangular region, and in turn the area of other kinds of regions. *See* AREA.

Subtraction is the inverse operation to addition, finding an addend when a sum and one addend are known. If there are 18 children on the playground and 10 are boys, then the number of girls is $18 - 10$ or 8, illustrating that the whole minus a part leaves the other part. Subtraction is used also to find the difference, for example, to see how many more are in one group of 18 children than in a group of 10 children. Subtraction is checked with addition; $18 - 10 = 8$ because $8 + 10 = 18$.

Division is the inverse operation to multiplication, that is, finding a factor when a product and a factor are known. In $12 \div 3 = 4$, 12 is called the dividend, 3 is called the divisor, and 4 is the quotient. There are two applications of division, in finding the number of threes in 12, or in finding the number in each set when 12 is separated into 3 equal groups. In division such as $13 \div 4$, the quotient of 3 and a remainder of 1 is usually written as 3r1. With fractions, the quotient can be shown as a factor, $3^1/_4$.

Basic facts. Any of the four operations with single-digit numbers is a basic fact. There are 100 basic facts each for addition, subtraction, and multiplication, but only 90 for division because 0 cannot be a divisor. Thinking strategies are useful in learning facts (see **table**).

Computation. The algorithms, or rules, for computation rely on basic facts, place value, and properties of the operations. In addition and subtraction, the total values of each of the places are added or subtracted. The sum of $345 + 67 + 162$ is found by combining the amount in the ones places, the amount in the tens places, the amount in the hundreds places, and then combining these three totals. A vertical format, as in (2), makes it easy to combine like values. Mathematical reasons for combining this way are the associative

$$\begin{array}{r} 345 \\ 67 \\ +162 \\ \hline 574 \end{array} \qquad (2)$$

Strategies for learning basic facts of arithmetic

Operation	Strategy	Example
+	Count on, for 6 + 3	Six, seven, eight, nine
	Doubles, for 7 + 8	Seven plus seven = 14; so 7 + 8 is one more, or 15
	Use 10, for 9 + 6	Nine and one make 10; 10 plus 5 = 15
−	Add on, for 12 − 9	What added to make 9 makes 12?
	Count up, for 12 − 9	Nine, ten, eleven, twelve—keeping track of the counts
×	Doubles, for 8 × 7	4 sevens = 28; so 8 sevens = 28 + 28 or 56
	Use known facts, for 7 × 7	5 sevens = 35; 2 sevens = 14; so 7 sevens = 35 + 14 or 49
	Relating to 10, for 9 × 6	10 sixes = 60; so 9 sixes = 60 − 6 or 54
÷	Use multiplication, for 72 ÷ 9	How many nines make 72? What times nine = 72?
	Count up, for 20 ÷ 5	Five, ten, fifteen, twenty; so four—keeping track of the counts

property, $a + (b + c) = a + (b + c)$, and the commutative property, $a + b = b + a$. *See* ALGEBRA.

The algorithm for multiplication depends on the distributive property for multiplication over addition, $a \times (b + c) = a \times b + a \times c$. For example, $7 \times (60 + 8) = 7 \times 60 + 7 \times 8$. The algorithm (3*a*) illustrates this property and shows why the 5 is "carried" to the tens place and why 47 is moved one place to the left in the standard algorithm (3*b*).

$$\begin{array}{r} 68 \\ \underline{\times 7} \\ 7 \times 8 \rightarrow 56 \\ 7 \times 60 \rightarrow \underline{420} \\ 7 \times 68 \rightarrow 476 \end{array} \quad (3a) \qquad \begin{array}{r} {\scriptstyle 5} \\ 68 \\ \underline{\times 7} \\ 476 \end{array} \quad (3b)$$

The written algorithms for addition, subtraction, and multiplication are done right to left, but division is done left to right. For example, if six people share the \$750 cost of a bus rental, the quotient for $6\overline{)750}$ is found by dividing hundreds, then the 15 remaining tens, and then the 30 remaining ones, for a quotient of 125. Each person needs to pay \$125.

Most of the arithmetic that people use everyday involves either mental computation or estimation.

Mental computation means finding the exact answer in one's head without paper and pencil or calculators. For all operations, mental computation is most often done left to right. For 68 + 76, one might think "60 + 70 = 130; 130 + 8 = 138; 138 + 6 = 144."

Estimation is done when only an approximate answer is needed. For example, an estimate for the area of a lawn that measures 325 ft by 62 ft might be 300 × 60 or 18,000 ft^2. A slightly better estimate is 300 × 70 or 21,000 ft^2. There are many different ways to estimate, depending on the need in the practical problem. Beyond estimation as a practical skill, it is needed to check calculations on a calculator.

Fractions and decimals. Understanding fractions and decimals, as well as operations on these numbers, is essential for practical uses and for long-term memory.

Meanings of fractions. The initial and most basic idea is that a fraction shows "part of a whole." In the fraction $^3/_4$, read "three fourths," the 4 shows the number of equal-size pieces in each whole unit as well as the size of one piece, "fourth". The 3 shows the number of equal-size pieces being taken or considered. The top number, 3, "numbers" the parts and is called the numerator. The bottom number, 4, "names" the parts and is called the denominator. For a set of objects such as a dozen eggs, the egg carton is considered the whole and individual eggs the parts. The fraction $^3/_{12}$ might show the part of a carton with brown eggs if 3 of the 12 eggs are brown and the rest are white.

Fraction meaning is extended to include division of a whole number by a counting number. For example, $1 \div 4$ could mean one whole pie divided among four people. Since each person would get $^1/_4$ of a pie, $1 \div 4 = {}^1/_4$. In a similar manner, $3 \div 4 = {}^3/_4$ (**Fig. 2**).

Fractions that represent the same part of a unit are equivalent fractions. A ruler with customary units shows equivalents such $^1/_4 = {}^2/_8 = {}^4/_{16} = {}^8/_{32}$. The algorithm for generating equivalent fractions is: multiply or divide the numerator and denominator by the same nonzero number. For example, the fraction $^{15}/_{100}$ can be expressed as $^3/_{20}$ by dividing 15 and 100 by 5.

A mixed number shows a combination of wholes and parts. For example, $3^4/_5$ means 3 wholes and $^4/_5$ of another whole. The fraction does not occupy a place, as each symbol does for whole numbers. It is sometimes useful to find mixed number and fraction equivalents. For $3^4/_5$, one whole can be expressed as $^5/_5$, so 3 wholes equals $^{15}/_5$; $^{15}/_5$ and $^4/_5$ more make $^{19}/_5$. Hence, $3^4/_5 = {}^{19}/_5$.

Operations on fractions. The meaning for the operations on fractions depends on the meaning of the

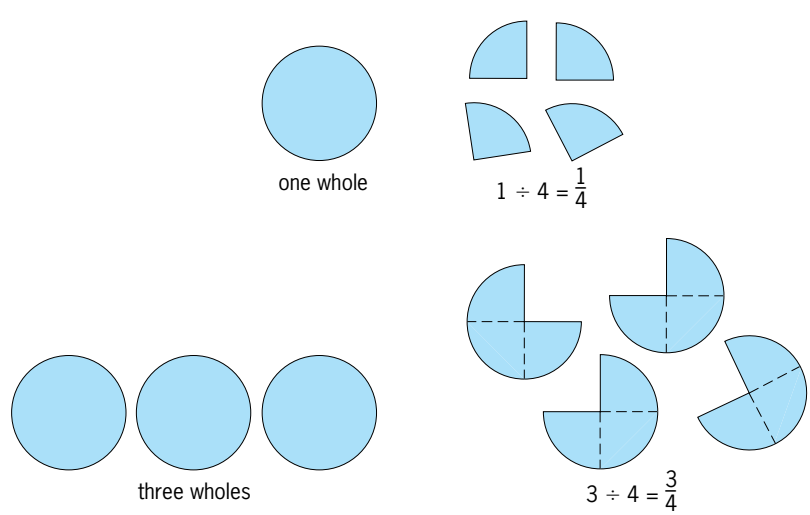

Fig. 2. Division of counting numbers as fractions.

numbers. A key idea for addition and subtraction is that equal-size parts are separated or combined. For example, each fraction in $^3/_8 + {}^2/_8$ shows eights, so the parts can be combined: 3 eighths + 2 eighths makes 5 eighths. The general algorithm is given by Eq. (4).

$$\frac{a}{b} + \frac{c}{b} = \frac{a+c}{b}, \quad b \neq 0 \tag{4}$$

To add or subtract fractions with different-size pieces, both fractions are expressed as equivalent fractions with a common denominator, such as in Eq. (5). One relatively easy way to find the common

$$\frac{3}{4} - \frac{1}{6} = \frac{9}{12} - \frac{2}{12} = \frac{7}{12} \tag{5}$$

denominator for 6 and 4 is to take successive multiples of the 6—that is, 6, 12, 18,. . . —until a number is found for which 4 is a factor. The first multiple, 6, does not have 4 as a factor. Since the next multiple, 12, has 4 as a factor, 12 is a common denominator, in fact the least common denominator, for 6 and 4. The least common denominator for three or more fractions is the least multiple of all of the three or more denominators.

One way to think of multiplication of fractions is to think of finding a fractional part of another number. For example, $^1/_2$ of $^6/_{11}$ is one-half of six equal pieces, each of which is one-eleventh. Finding "$^1/_2$ of the amount" is multiplication by $^1/_2$, as in Eq. (6). The

$$\frac{1}{2} \text{ of } \frac{6}{11} = \frac{1}{2} \times \frac{6}{11} = \frac{3}{11} \tag{6}$$

usual algorithm, multiplying numerators and multiplying denominators, as in Eq. (7), gives Eq. (8). One

$$\frac{a}{b} \times \frac{c}{d} = \frac{a \times c}{b \times d}, \quad b \neq 0, \quad d \neq 0 \tag{7}$$

$$\frac{1}{2} \times \frac{6}{11} = \frac{6}{22} = \frac{3}{11} \tag{8}$$

way to multiply mixed numbers is to express each as a fraction, such as in Eq. (9).

$$2\frac{1}{4} \times 3\frac{1}{2} = \frac{9}{4} \times \frac{7}{2} = \frac{63}{8} \tag{9}$$

One use of division with fractions is to find the number of parts contained in a number. For example, $5 \div {}^1/_8$ could mean to find how many $^1/_8$ pieces of a candy bar are in 5 candy bars. Since there are eight eighths in one whole, there are 5×8 or 40 in 5 wholes. Hence, $5 \div {}^1/_8 = 5 \times 8$. The general rule is: to divide by c/d, multiply by its reciprocal, d/c, as in Eq. (10), where none of the denominators b, c, or d

$$\frac{a}{b} \div \frac{c}{d} = \frac{a}{b} \times \frac{d}{c} \tag{10}$$

is 0. To divide mixed numbers, each mixed number is first expressed as a fraction.

Fig. 3. Decimals with symmetry around the ones place.

Meaning of decimals. To show decimals less than one, the place value system for whole numbers is extended to the right of the ones place. The value of each place to the right is a successive power of $^1/_{10}$. The decimal point is needed to designate the ones place because the place on the right is no longer the ones place. The decimal point also separates the whole number from the decimal part. The names and values of the places are symmetric with respect to the ones place (**Fig. 3**). The 3. (decimal point included) is the center, surrounded by tens and tenths, as well as hundreds and hundredths. The meaning can be shown as $100 + 20 + 3 + {}^4/_{10} + {}^5/_{100}$. The decimal part, .45, can be thought of as 45 hundredths because $^4/_{10} = {}^{40}/_{100}$, and ${}^{40}/_{100} + {}^5/_{100} = {}^{45}/_{100}$.

Decimal operations. As with whole numbers, addition and subtraction with decimals is done by combining or separating equal-size parts. For example, to add 62.4 + 5.62, hundredths are first combined, then tenths, then ones, then tens for a total of 68.02.

One way to think about multiplication of decimals is to think of multiplying fractions, such as in Eq. (11).

$$0.2 \times 0.3 = \frac{2}{10} \times \frac{3}{10} = \frac{6}{100} = 0.06 \tag{11}$$

Counting the total number of places to the right of the decimal points in the two factors shows the power of $^1/_{10}$ and consequently the number of places to the right of the decimal point in the product.

Division with decimals is usually done by solving an equivalent problem with the divisor a whole number. For example, $0.6\overline{)4.62}$ is changed to an equivalent problem, $6\overline{)46.2}$, by multiplying divisor and dividend by 10.

Percent is another way to express fractions and decimals that show hundredths. For example, 7 hundredths can be expressed as $^7/_{100}$, as 0.07, or as 7%. All three expressions show the same part of a whole. *See* PERCENT.

Joseph N. Payne

Bibliography. E. F. Krause, *Mathematics for Elementary Teachers*, 2d ed., 1991; National Council of Teachers of Mathematics, *Mathematics for the Young Child*, 1990; J. A. Van de Walle, *Elementary School Mathematics*, 2d ed., 1994.

Arkose

Arenite (rock composed of sand-size fragments) that contains a high proportion of feldspar in addition to quartz and other detrital minerals. Arkose is also known as feldspathic sandstone. Although there is

no universal agreement, many geologists consider a minimum of 25% feldspar a requisite for calling sandstone an arkose. Other geologists accept a lower value.

Composition. Arkoses may contain a high proportion of other nonquartz detritus, such as igneous and metamorphic rock fragments, micas, amphiboles, and pyroxenes. Frequently the accessory heavy mineral suite consists of a variety of species. Though the arkoses are rarely as well sorted as orthoquartzites, they may be moderately well sorted. The grains are angular or poorly rounded. Clay matrix is generally subordinate in arkoses, but there may be as much as 10–15%. If there is that much, the rocks have textural similarity to feldspathic graywackes, though the mineralogy may be appreciably different in the latter. The clay is dominantly kaolinite with smaller proportions of micaceous and montmorillonitic clay. Presumably much of the kaolinite has come from the weathering of feldspar. Conglomeratic zones are common in many arkoses. Feldspathic sandstones that contain less than 25% of feldspar have been termed subarkoses. They are in general similar to the more highly feldspathic true arkoses.

Structure. Sedimentary structures of arkoses are similar in kind to those of the orthoquartzites. Cross-bedding, the major feature, may be displayed on a huge scale, some cross-bedded units being many feet thick. Bedding is crude and many times not distinguishable; many beds are thick and massive. Ripple marks may be present but are not common. Some arkoses, such as those of the Newark Series along the Atlantic Coast of the United States, contain mud cracks, frost crystal impressions, and footprints of small dinosaurs.

Occurrence. Arkoses are associated with a variety of clastic rocks, dominantly conglomerates, and reddish-colored shales. Arkoses also are found with basic lava flows. The formations occur as thick, wedge-shaped deposits, the thick end of the wedge being in close proximity to the source area and sometimes separated from it by normal faults of large displacement. Other arkoses are relatively thin formations that overlie granitic terrain. These formations grade laterally into other kinds of sandstone away from the area of underlying granite. Most arkoses are found in geosynclinal areas, but the thin, reworked, granite-wash arkoses can be found on stable continental platforms. *See* GEOSYNCLINE.

Origin. The granite-wash arkoses appear to have formed as the result of a transgression of the sea over a land area underlain by granite. The fragmented granite in the soil and mantle rock is incorporated in the basal sediment. In some areas the original granite is changed so slightly the arkose is called recomposed granite and may be almost indistinguishable from the original granite.

The origin of the arkoses is best understood in terms of the abundance of feldspar. Feldspar is unstable both chemically and mechanically as compared with quartz and, given sufficient rigors of chemical weathering at the source and abrasion during transportation, will disappear. The lack of appreciable chemical weathering at the source (which allows the contribution of much feldspar to the sediment) may be due to high topographic relief or to climatic extremes.

Since high relief and climatic extremes generally are associated with orogenic movements, arkoses are usually interpreted as sediments that result from tectonically active regions. Rift valleys formed by the divergence of two continental plates at a spreading center are now favored as the tectonic environment for arkose formation. This explains too the association with basaltic lavas. The abundant iron oxides present in most arkoses, the mud cracks, and the fanglomerates point to a mostly terrestrial mode of deposition. The thin granite washes are marine and represent different conditions. *See* FELDSPAR; GRAYWACKE; SANDSTONE; SEDIMENTARY ROCKS.

Raymond Siever

Armadillo

A heavily armored mammal in the family Dasypodidae, order Cingulata (previously Edentates or Xenarthra). Armadillos are characterized by bony dermal plates, which help to protect their bodies. Being mammals, they do have hair, but it is usually sparse and often quite inconspicuous. The anterior and posterior parts of the body are often heavily armored, whereas the midportion of the body is often encircled or partially encircled by rings. These rings are used taxonomically in describing and identifying many of the species of armadillos. The skin of the ventral part of the body is soft. The toes are clawed and used to burrow and dig for food. The hindlimbs always have five toes with claws; the forelimbs may have three, four, or five heavy claws used for digging. The snout is long, and the cylindrical viscous tongue is used in capturing food. The teeth are simple pegs with no roots and no enamel. Most species have about 14 to 18 teeth, but the giant armadillo, *Priodontes maximus*, has about 80 to 100 small teeth. It reaches a length of approximately 10 feet (3 m), whereas the smallest species, the fairy armadillo, *Chlamyphorus truncatus*, is about 5–6 in. (12–15 cm) long. When disturbed, many species roll into a ball or wedge themselves into a burrow opening, with the dermal plates helping to protect them.

Armadillos occur only in the New World. They range from the southern United States through most of South America. There are 21 species of armadillos in nine genera, but only one, the nine-banded armadillo, *Dasypus novemcinctus*, occurs in North America (see **illustration**). It is the best-known species. It is adaptable and has extended its range rapidly in the United States: it is now found from Texas and Oklahoma eastward to Florida, Georgia, and southern South Carolina. It feeds on a great variety of invertebrates, including beetles, termites, great numbers of ants, and many other insects such as centipedes, millipedes, and spiders. It also feeds on small vertebrates, including lizards and their eggs,

Nine-banded armadillo (*Dasypus novemcinctus*). (*Photo © 2005 Pam Y. Burtt*)

and on some plant foods, such as various fruits, tubers, and subterranean fungi. This species uses scent a great deal in hunting. It places its nose close to the ground and holds its breath (up to 6 min) while digging so as to avoid inhaling dust. It is of public health importance because it is a reservoir host for *Trypanosoma cruzi*, the causative agent of Chagas' disease.

The nine-banded armadillo has a very unusual life cycle. Four young are born in a den or chamber at the end of the burrow. The young are always of the same sex and are identical quadruplets, all arising from the division of a single egg. The multiple-birth condition, or polyembryony, is unique among mammals. The young are well developed at birth and are weaned at about 8 weeks of age.

The genus *Dasypus*, the long-nosed armadillos, includes six other species in addition to the nine-banded armadillo. The genera *Priodontes* and *Chlamyphorus* contain one species each. Species in other genera including only one species are *Calyptophractus retusus*, the greater fairy armadillo; *Euphractus sexcinctus*, the six-banded armadillo; and *Zaedyus pichiy*, the pichi. Other genera are *Cabassous*, the naked-tailed armadillos, with four species; *Chaetophractus*, the hairy armadillos, with three species; and *Tolypeutes*, the three-banded armadillos, with two species. *See* EDENTATA; TRYPANOSOMIASIS.

John O. Whitaker, Jr.

Bibliography. A. L. Gardner, Order Cingulata, in *Mammals of the World*, ed. by D. Wilson and D. A. Reeder, in press, 2005; G. J. Galbreath, Armadillo, pp. 71–79 in *Wild Mammals of North America*, ed. by J. A. Chapman and G. A. Feldhamer, Johns Hopkins University Press, Baltimore, 1982; G. G. Montgomery, *The Evolution and Ecology of Armadillos, Sloths and V. vermilinguas*, Smithsonian Institution Press, Washington, DC, 1985; R. M. Nowak, *Walker's Mammals of the World*, 5th ed., vol. 1, Johns Hopkins University Press, Baltimore, 1991; J. O. Whitaker, Jr., and W. J. Hamilton, Jr., *Mammals of the Eastern United States*, Cornell University Press, Ithaca, NY, 1998; D. E. Wilson and D. M. Reeder, *Mammal Species of the World*, 3d ed., Johns Hopkins University Press, Baltimore, 2005.

Armature

That part of an electric rotating machine which includes the main current-carrying winding. The armature winding is the winding in which the electromotive force (emf) produced by magnetic flux rotation is induced. In electric motors this emf is known as the counterelectromotive force.

On machines with commutators, the armature is normally the rotating member (**Fig. 1**). On most ac machines, the armature (**Fig. 2**) is the stationary member and is called the stator. The core of the armature is generally constructed of steel or soft iron to provide a good magnetic path, and is usually laminated to reduce eddy currents. The armature windings are placed in slots on the surface of the core. On machines with commutators, the armature winding is connected to the commutator bars. On ac machines with stationary armatures, the armature winding is connected directly to the line. *See* COMMUTATOR; CORE LOSS; ELECTRIC ROTATING MACHINERY; WINDINGS IN ELECTRIC MACHINERY.

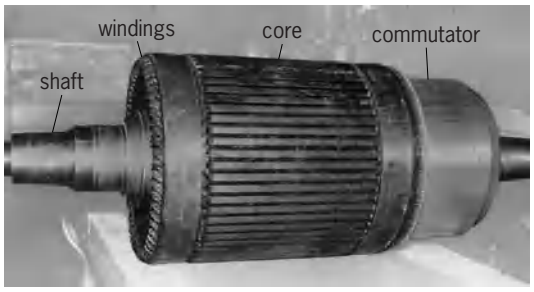

Fig. 1. A rotor armature of of a direct-current generator or motor. (*General Electric*)

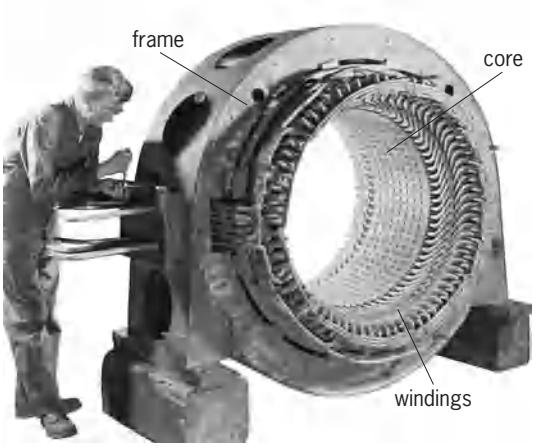

Fig. 2. A stator armature of an ac induction motor. (*Allis Chalmers*)

Arthur R. Eckels

Bibliography. S. A. Nasar (ed.), *Handbook of Electric Machines*, McGraw-Hill, 1987; S. A. Nasar and L. E. Unnewehr, *Electromechanics and Electric Machines*, 2d ed., Wiley, New York, 1983; National

Electrical Manufacturer's Association, Pub. no. MG1-1972, *Motors and Generators*, New York, 1972; M. S. Sarma, *Electric Machines: Steady State Theory and Dynamic Performance*, 2d ed., Brooks-Cole, Monterey, 1994.

Armature reaction

The effects of the magnetomotive force (mmf) of the armature on the air-gap field of direct-current (dc) and synchronous alternating-current (ac) machines. Since armature current varies directly with the electrical or mechanical load on the machine, armature-reaction effects are load-dependent. In dc machines, the armature reaction causes a distorted flux-density distribution in the air gap. **Figure 1** shows the magnetic field produced in the air gap of a two-pole dc motor or generator by (*a*) the field mmf, (*b*) the armature mmf,and (*c*) both armature and field mmf acting together. **Figure 2** shows the distorted flux distribution in the air gap, curves *a, b*, and *c* corresponding to the above three cases. Due to the saturation of the armature teeth, the flux density is decreased by a greater amount under pole tip *e* than it is increased under pole tip *f*, and therefore the armature reaction produces a demagnetizing effect, and the generated voltage or countervoltage will be reduced when the armature is loaded. In a generator this degrades the voltage regulation. In a motor it tends to increase the speed and may cause instability.

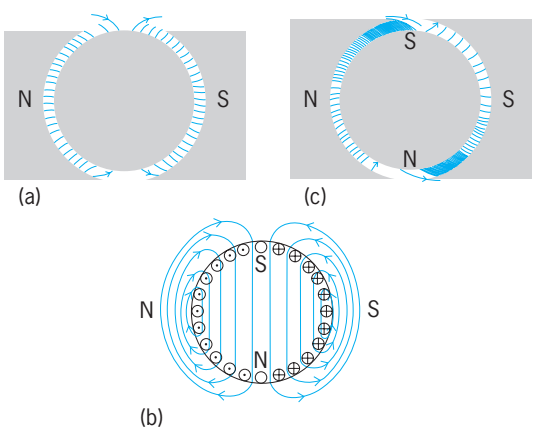

Fig. 1. Magnetic fields in air gap of two-pole machine. (*a*) Main field. (*b*) Armature field. (*c*) Load conditions.

For dc machines subject to heavy overloads, rapidly reversing loads, or operation with a weak field, the resultant flux-distribution distortion by excessive armature reaction will cause nonuniform distribution of voltage between commutator segments, and may result in flashover between commutator segments. A pole-face (or compensating) winding, embedded in slots in the pole face and excited by armature current, is provided to neutralize the armature mmf under the pole faces.

In polyphase synchronous machines, balanced polyphase load currents flowing in the armature winding produce a revolving mmf. This mmf rotates

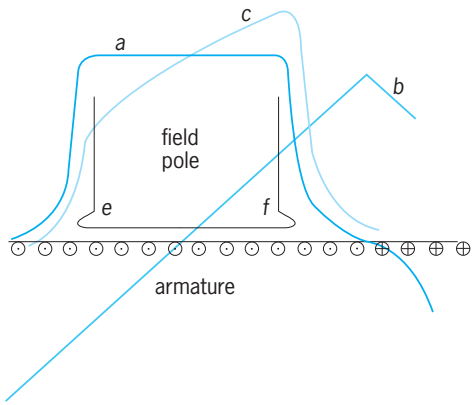

Fig. 2. Distorted flux distribution in air gap.

at exactly the speed of the rotating field poles. Hence the mmf of armature reaction reacts with the mmf of the field winding on the rotor poles to produce a displaced resultant magnetic field. The angle of the armature reaction mmf with respect to the pole axis depends upon the power factor and kilovolt-amperes of the load.

In cylindrical-rotor machines, the position and magnitude of the peak air-gap flux density are determined by the resultant mmf field, and armature reaction in such machines therefore causes a change in the amplitude and phase of the generated voltage without substantially altering its waveform. In salient-pole machines, the armature reaction tends to concentrate the flux at one side of the pole face (**Fig. 3**), and some voltage waveform distortion results.

In analyzing the effects of armature reaction in salient-pole machines, the armature mmf is often resolved into two components. That component which acts along the pole axis is known as direct-axis armature reaction and has a direct magnetizing or demagnetizing effect. That component located at right angles to the pole axis is known as quadrature-axis

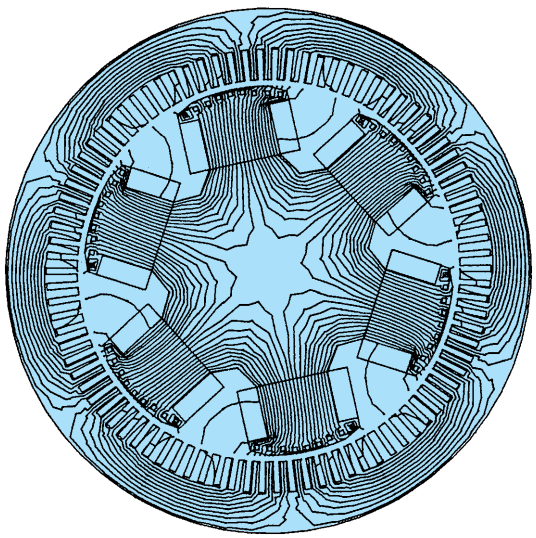

Fig. 3. Computer plot showing skewing of pole flux effected by armature reaction in a salient-pole synchronous machine. (*E. G. Strangas, Michigan State University*)

armature reaction. It is analogous to the armature reaction effect in modern dc machines. *See* ARMATURE; ELECTRIC ROTATING MACHINERY; MAGNETOMOTIVE FORCE.

George McPherson, Jr.

Army armament

The weapons, equipment, and supplies that permit lethal or nonlethal devices to strike their targets. Armaments can be classified as field artillery, individual and crew-served weapons (infantry), armor and antiarmor, antiaircraft, helicopter armaments, and mines and countermines.

Supplies must be available to all army units to support armaments. For example, a battery of six howitzers can expend 1.5 tons (1360 kg) of projectiles, propellant, and fuzes during 1 min of firing. A logistics system which includes aircraft, ships, and a wide variety of ground vehicles and material-handling equipment (such as forklifts and cranes) is designed to maintain all forces with supplies. General-purpose ground vehicles used for supply include 0.5- and 0.75-ton (450-and 680-kg) trucks and 6-ton (5400-kg) trailers. Specialized cargo-carrying vehicles are used to accompany and supply ammunition to self-propelled artillery howitzers. For example, supporting the 155-mm self-propelled howitzer fleet is the M992A2 field artillery ammunition supply vehicle (FAASV), which has a loaded weight of 58,500 lb (26,500 kg) and a built-in automated conveyor to deliver fuzed artillery ammunition.

Field Artillery

Artillery can deliver highly lethal warheads to ranges well beyond the reach of infantry weapons. Artillery includes rockets, missiles, and self-propelled and towed howitzers. Artillery fire support is designed to meet adversarial forces by attacks on personnel and on medium and hard targets such as troop carriers and tanks. This mission is conducted at extended ranges with high-explosive projectiles, projectiles with shaped-charge submunitions, and the Copperhead guided projectile. Artillery can also channel, delay, and destroy oncoming forces by delivery of antipersonnel and antitank mines. Future goals include the use of smart munitions; increased range, lethality, and firing rates; new propulsion techniques; cartridge course correction in flight; and systems which permit much greater battlefield versatility.

Cannons (howitizers). There are two calibers of cannons in the U.S. Army's inventory, the 105-mm and the 155-mm. The army-specific name for these cannons are howitzers, and they are the primary armaments of field artillery units. They can be either towed or self-propelled systems.

105-mm howitzers. These are the M102 weapon and the M119 towed weapon. They are intended for rapid deployment of light infantry and air assault divisions. The M102 weighs approximately 5242 lb (2380 kg), and has a range of 7.0 mi (11.3 km) with the M1 high explosive projectile and 9.2 mi (14.8 km) with the rocket-assist M548 projectile. The M119A1 howitzer is derived from the British L119 light gun and is designated to replace M102 howitzers. The weapon weighs approximately 4270 lb (1816 kg) and fires the standard 105-mm artillery cartridges. Its range is 8.9 mi (14.3 km) with the M760 high-explosive projectile and 12.1 mi (19.5 km) through rocket assist.

155-mm howitzers. These consist primarily of the self-propelled, fully tracked M109 series weapon and the towed M198 howitzers.

The M109 series howitzer is the main fire support artillery system of the armored and mechanized units. The most advanced version is the M109A6 Paladin (**Fig. 1***a*), which is capable of semiautonomous operations for more rapid gun positioning and firing. Paladin has both a chassis lined with high-strength synthetic fiber material (aramid) and a pressurized compartment to ensure survivability of the crew against ballistic and nuclear, biological, and chemical threats. The combat-loaded weapon weighs 63,500 lb (28,900 kg). Its M284 cannon has a maximum range of 11 mi (18 km) with the M107 high

Fig. 1. 155-mm howitzers. (*a*) M109A6 Paladin self-propelled howitzer (*BMY Combat Systems*). (*b*) M198 towed howitzer (*U.S. Army*).

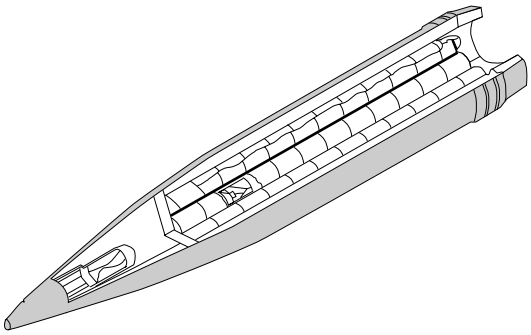

Fig. 2. M483A1 155-mm cargo-carrying projectile. (*U.S. Army*)

explosive and 19 mi (30 km) with the M549A1 rocket-assist projectile. The M483A1 projectile (**Fig. 2**), which was used effectively against Iraqi forces in the Persian Gulf War, contains 88 dual-purpose, anti-light-armor/antipersonnel M42/M46 submunitions and can be fired at a range of 10.9 mi (17.5 km). The M864 projectile contains 72 submunitions and a baseburner motor which extends the range to 17 mi (27 km). The 155-mm howitzer also fires the M712 shaped-charge Copperhead guided projectile to a range of 10 mi (16 km). The M712 uses semiactive laser guidance for the terminal flight and homing on target.

Crusader is a 155-mm self-propelled artillery system being designed to replace the M109A6 Paladin system in 2007. The Crusader will offer the most advanced technologies in automation, electronics, and software, digitization, mobility, and lethality. Through use of autonomous operations, Crusader will have unprecedented positioning and firing rate. Its XM297 cannon has a maximum range of 25 mi (40 km) with the M549A1 projectile and nearly 31 mi (50 km) with the developmental XM982 projectile.

The towed M198, air-transportable weapon (Fig. 1*b*) replaced the M114 series. It weighs 15,000 lb (6810 kg). Its M199 cannon has a maximum range of 19 mi (30 km) with the M549A1 rocket-assist projectile and 14 mi (22 km) with the M107 high-explosive projectile.

The lightweight 155-mm towed howitzer (XM777) system (**Fig. 3**) replaces the M198. The XM777 through the use of high-strength titanium maintains strength but at a significantly reduced weight (less than 9000 lb or 4100 kg). It is designed to meet or exceed the capabilities of the M198. Because it is lightweight, it has improved mobility and terrain coverage. The XM777 is also the first towed artillery system to use digitization for improved lethality and battlefield dominance.

Rockets. The multiple launch rocket system (MLRS; **Fig. 4**) is a free-flight artillery rocket weapon used primarily as a counterfire weapon with a range of up to 25 mi (40 km). It consists of a 12-rocket launcher, pod-mounted on a tracked vehicle. The warhead contains 644 of M77 dual-purpose, shaped-charge, antipersonnel, antimateriel bomblets similar to those used in the M483A1 projectile. The MLRS was used with great effectiveness against Iraqi forces in the Persian Gulf War.

The 2.75-in. rocket systems are a family of "fire and forget" rockets used primarily for air-to-ground applications by fixed and rotary wing aircraft. The rockets are fired from either a 7- or 19-tube launcher for antipersonnel, antimaterial, and antiarmor purposes. Screening smoke and both visible and infrared illumination are also available for supportive missions.

Individual and Crew-Served Weapons (Infantry)

Infantry armament is composed largely of line-of-sight, direct-fire weapons and indirect-fire mortar weapons used by nonmechanized light infantry (60-mm and 81-mm mortars) and mechanized infantry and armor divisions (120-mm mortars)

Small-caliber weapons. Many small-caliber weapons have been introduced as replacements or additions to the inventory.

M9 and M10 pistols. These 9-mm pistols will replace the M1911A1 .45-caliber pistol and various .38-caliber revolvers. The pistols are semiautomatic,

Fig. 3. XM777 lightweight 155-mm towed howitzer. (*U.S. Army*)

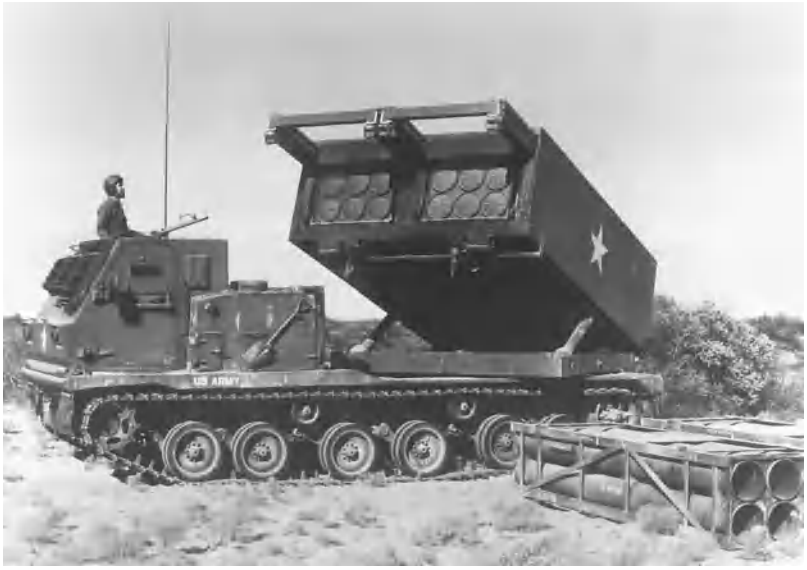

Fig. 4. Multiple-launch rocket system (MLRS). (*Vought Corp.*)

Fig. 5. M16A2 rifle. (*U.S. Army*)

magazine-fed, recoil-operated, double-action handguns with a 15-round magazine. They are chambered to accept M882 ball ammunition and the standard 9-mm NATO ammunition. The M9 and M10 weigh approximately 2.52 lb (1.145 kg).

M16A2 rifle. The rifle (**Fig. 5**) is an improved version of the 5.56-mm-caliber M16A1 rifle. It includes a muzzle brake compensator to reduce muzzle jump, a burst control to limit automatic fire to three rounds per trigger pull, and a barrel designed to accommodate NATO standard 5.56-mm ammunition. The rifle has a maximum effective range of 2900 ft (880 m) against area targets. It fires M855 ball and M196 tracer ammunition.

M4 carbine. The operating characteristics of the M4 5.56-mm carbine are similar to those of the M16A2 rifle; the same ammunition is used. However, the carbine weighs 7.2 lb (3.5 kg) with a loaded 30-round magazine versus 8.8 lb (4 kg) for the M16A2. The M4 is also almost 25 cm (10 in.) shorter. The maximum effective range against area targets is 2600 ft (800 m).

M24 sniper weapon system. This replacement for the current M21 sniper rifle is a bolt-action, 7.62-mm-caliber rifle with a detachable 10-power optical sight and a five-round magazine containing M118 special ball cartridges. The maximum effective range is 2600 ft (800 m).

M203 grenade launcher. This is a 40-mm, pump-action, single-shot weapon that attaches under the barrel of the M16 series rifles. It weighs 3.6 lb (1.6 kg) loaded, has a maximum range of 1150 ft (350 m), and fires practice, buckshot, dual-purpose, pyrotechnic, and spotting ammunition.

Mark 19 Mod 3 grenade machine gun. This weapon is an advanced version of the Mark 19 gun. It is an air-cooled, automatic machine gun that fires M430 40-mm dual-purpose antipersonnel, anti-light-materiel cartridges at a cyclic rate of 325–375 rounds per minute to a range of 1.4 mi (2.2 km) from a tripod- or vehicle-mounted position.

M60 7.62-mm machine gun. This is a general-purpose widely used machine gun mounted on a bipod (**Fig. 6**), on a tripod, or in a vehicle. It weighs 23 lb (10.4 kg) empty and fires 100 rounds per minute at a sustained rate to a range of 2.3 mi (3.7 km). An M60E3 version has been designed to be lighter and more versatile.

M249 squad automatic weapon. This is a one-person portable machine gun that fires the 5.56-mm M855 ball and M856 tracer ammunition. It is used to replace the M16 for sustained and accurate fire to approximately 3300 ft (1000 m) at a rate of 800 rounds per minute.

M231 5.56-mm firing port machine gun. This is an air-cooled, gas-operated, magazine-fed system mounted in the Bradley fighting vehicle (discussed below). It weighs 9.5 lb (4.3 kg) with a 30-round magazine, has a sustained rate of fire of 50–60 rounds per minute, and an effective range of 1000 ft (300 m).

M242 25-mm chain gun. This is the main armament system for the Bradley. It is externally powered by a 1.5-hp (1.1-kW) motor, has a dual feed, and can fire single shots and multiple shots at up to 500 rounds per minute of high-explosive M792 and armor-piercing M791 and M919 ammunition. It has an effective range of 6500–10,000 ft (2000–3000 m).

M230 30-mm Hughes chain gun. This is a major turret-mounted weapon of the AH-64 Apache helicopter (discussed below) used in air-to-air and air-to-ground roles. It weighs 118 lb (53.6 kg) and fires the M789 high-explosive, dual-purpose antitank and antipersonnel rounds at a rate of 625 rounds per minute.

Next-generation weapons. Improvements in hit probability, combat load, and target acquisition of small-caliber weapons over those of the current systems are being pursued. The long-term objective is a family of small arms that may feature composite materials, bursting ammunition, laser fire control, night-vision devices, and microelectronics. The Objective Individual Combat Weapon and the Objective Crew-Served Weapon are slated to be the next-generation small-caliber systems.

Mortars. The U.S. Army uses three mortar weapon systems: the 60-mm M224 lightweight company mortar, the 81-mm improved M252 mortar, and the 120-mm M120 mortar. The 107-mm M30 mortar has been replaced by the 120-mm M120 mortar (towed and ground mounted) and the 120-mm M121 mortar (carrier mounted). Each type of mortar fires a family of projectiles, including high-explosive, smoke, illuminating, and full-range and short-range practice cartridges. The main direction of mortar development has been to increase weapon and cartridge performance (through improvements in lethality, smoke obscuration, white-light and infrared-light illumination, range, reliability, and rate of fire and reductions in weight and unit cost; to improve fuze versatility through use of a multioption fuze; and to use waterproof combustible cases for propulsion charges. There has been significant improvement in mortar fire control through the use of the M23 Mortar Ballistic Computer, and initiatives are under way to enable mortars to be active participants in the digital

Fig. 6. M60 7.62-mm machine gun. (*U.S. Army*)

Fig. 7. M224 60-mm lightweight company mortar system. (*U.S. Army*)

battlefield via the XM30 Improved Mortar Ballistic Computer and the XM35 Digital Mortar Fire Control System. A program is also under way to make mortar cartridges safer for the user through insensitive munition design.

60-mm M224 mortar. This weapon is a lightweight, smoothbore system which can be drop- or trigger-fired (**Fig. 7**). It fires the M720 or M888 high-explosive cartridges with the M734 multioption fuze and the M935 point detonation fuze respectively to a range of 2.2 mi (3.5 km). The M224 also uses the M721 illuminating cartridge, the M722 white phosphorus marking cartridge, and the M767 60-mm short-range practice cartridge. The M767 infrared illuminating cartridge, the M720A1 high-explosive cartridge, the M768 high-explosive cartridge, and the M769 full-range practice cartridge are in development.

81-mm M252 mortar. This is the basic United Kingdom L16 smoothbore mortar with a U.S.-developed blast-attenuating device and an improved M177 bipod mount. The weapon weighs 93.25 lb (42.3 kg). It fires the M821 and M821A1 high-explosive cartridges with the M734 multioption fuze to a range of 3.5 mi (5.7 km). The M853A1 illumination, the M819 red phosphorus smoke, the M889 and M889A1 high-explosive cartridges with the M935 point detonation fuze, the M879 full-range practice cartridge, and the M880 short-range practice cartridge are standard rounds for the M252 system. The XM816 infrared illuminating cartridge is in development.

120-mm M120 and M121 mortars. This mortar weapon has replaced the 107-mm M30 mortar. It is an improved version of the Soltam K-60 Tampella mortar and is fielded in two configurations. The M120 towed version is transported on a two-wheel carriage. The M121 carrier-mounted version is mounted on the M1064 mortar carrier. The M121 weapon can be fired from the carrier or removed and fired from the ground by using an auxiliary baseplate and mount. The towed mortar weighs 318 lb (144 kg) in firing position and 642 lb (291 kg) in towed position. It fires a full family of ammunition within the range from 560 ft to 4.5 mi (170 m to 7.2 km). An improved high-energy cartridge (M933/M934), an illuminating cartridge (M930), and a smoke cartridge (M929) have been developed for use with this weapon. The 81-mm M303 insert allows 120-mm crews to live-fire-train with the older 81-mm M300 series ammunition and the 81-mm M880 short-range practice cartridge with a significant reduction in training costs. Cartridges in development include the XM983 infrared illuminating cartridges. The XM984 extended-range DPICM submunition round, and the Precision-Guided Mortar Munition (PGMM).

Shoulder-launched munitions. The primary shoulder-launched infantry munitions consist of the AT4 recoilless rifle and the Shoulder-Launched Multipurpose Assault Weapon-Disposable (SMAW-D) Bunker Defeat Munition.

M136 AT4 recoilless rifle. Infantry forces use this weapon primarily to defeat light-armored targets. It is issued as a complete system containing an 84-mm projectile with a shaped-charge warhead and a disposable composite launcher. The system weighs 14.8 lb (6.7 kg), is 39 in. (1.0 m) long, and has an effective range of between 800 and 1600 ft (250 and 500 m). It replaces the M72 LAW (light antitank weapon). The system has been employed in several conflicts, including the Persian Gulf War.

XM141 SMAW-D bunker defeat munition. Infantry forces use the SMAW-D (**Fig. 8**) to defeat bunkers, walls, and very light armor targets. It is issued as a complete system and includes a dual-mode warhead and disposable composite launcher. The system weighs 15.7 lb (7.1 kg) and is 32 in. (0.8 m) long. The system has an effective range of 800–1600 ft (250–500 m).

Fig. 8. XM141 SMAW-D bunker defeat munition. (*U.S. Army*)

Armor and Antiarmor

The primary armored weapon systems consist of the M1 and M1A1 main battle Abrams tanks; the more lightly armored Bradley M2 infantry and M3 cavalry fighting vehicles; and the M113 family of armored personnel carriers. Major thrusts in armament development have resulted in conversion of the rifled 105-mm cannon in the M1 tank to the 120-mm smoothbore cannon in the M1A1 tank, and the utilization of tungsten and depleted-uranium long-rod penetrators for maximum antiarmor effectiveness. Increased protection against penetration by kinetic-energy and shaped-charge ammunition has been provided by the use of laminate Chobham-type armor and reactive armor.

Main battle tanks. The main purpose of the Abrams series of main battle tanks is to provide mobility, firepower, and shock effect to overwhelm and destroy enemy forces on the battlefield. The Abrams powerful main gun, special armor, and 1500-hp (1.1-MW) engine make it particularly suitable for both attacking and defending against enemy armor forces. The M1 tank (**Fig. 9**) is equipped with a 105-mm M68 main gun, a M219/M240 coaxial weapon which fires 7.62-mm ammunition, another M240 7.62-mm loader's machine gun, and a commander's M85 .50-caliber machine gun. The M1A1 tank's main gun is the 120-mm M256 smoothbore cannon; small-caliber weapons carried are identical to those on the M1 tank. M1A1 features include increased armor protection, suspension improvements, and a protection system for nuclear, biological, and chemical attacks. These afford the tank and the soldiers greater survivability in a variety of contaminated environments. Antitank ammunition available for use in the 105-mm cannon includes the M735, M774, and M833 kinetic-energy, heavy-metal, armor-piercing, fin-stabilized, discarding-sabot projectiles; and the M456 series of high-explosive, shaped-charge projectiles. Projectiles developed for use in the 120-mm cannon include the M829A2 high-velocity kinetic-energy round and the M830A2 shaped-charge round, which has high antiarmor effectiveness and has an added proximity switch selection that provides the tank with a self-defense capability against helicopters. The propellant for the 120-mm gun is contained in combustible cartridge cases; metal cartridge cases are used in the 105-mm weapons. Both cannons fire kinetic-energy and shaped-charge projectiles.

Fig. 9. M1 Abrams battle tank. (*U.S. Army*)

Fig. 10. M2 Bradley infantry fighting vehicle with M242 25-mm turret-mounted chain gun. (*U.S. Army*)

Infantry fighting vehicles. The Bradley M2 infantry fighting vehicle (**Fig. 10**) transports and supports infantry soldiers on the battlefield, while the M3 cavalry fighting is used to conduct reconnaissance and surveillance operations. Both vehicles have identical armament, which consists of the 25-mm M242 turret-mounted chain gun, a 7.62-mm M240C coaxial machine gun, and six M231 firing port weapons. The vehicles also carry a dual-tube TOW missile launcher (discussed below). The M3 variant carries more TOW missiles than does the M2 variant. This space is used on the M2 to carry more infantry soldiers.

M113 armored personnel carrier. This versatile vehicle (**Fig. 11**) exists in many variations to accomplish different missions. Normal armament is the .50-caliber machine gun. The vehicle may also serve as the carrier for 81-mm, 107-mm, or 120-mm mortar weapons, the M163 20-mm Vulcan air-defense gun, TOW launchers, and the Mark 19 Mod 3 machine gun.

Appliqué armor. These appliqués consist of packages that may contain a series of plates and an insensitive explosive which reacts with incoming shaped-charge jets to greatly reduce their armor penetration effectiveness. The armor tiles are flat metal boxes of various dimensions that are mounted on the front, rear, top, and sides of the vehicle. Appliqué armor has been designed for the M2A2 and M3A2 Bradley fighting vehicles and for the M60A3 tanks.

TOW. The tube-launched, optically tracked, wire-guided missile (TOW; **Fig. 12**) is a heavy antiarmor

assault weapon that can be fired from a ground tripod, from helicopters, and from armored and unarmored vehicles. The missile is 6 in. (15 cm) in diameter and contains a variety of shaped-charge warheads. It has a range of up to approximately 2.5 mi (4.0 km). Following initial launch to a safe distance from the gunner, the rocket motor accelerates the missile to the tactical velocity. The initial TOW system was fielded in 1970, and since then has continually been upgraded.

Air Defense

There are a number of air defense weapons. These include the Avenger Air Defense System, the handheld Stinger rocket, and the Patriot system.

Stinger. This is a human-portable, shoulder-fired, two-stage, infrared homing missile, designed to attack and destroy at short range any low-altitude, fixed- and rotary-wing aircraft with a high-explosive fragmentation warhead. The Stinger is to be replaced by the Linebacker Air Defense System, which is in production and being fielded in the Armor and Mechanized Divisions (as of 2000).

Avenger Air Defense System. This air defense system is mounted on a Humvee (high-mobility multipurpose wheeled vehicles; HMMWV). It is lightweight and transportable and designed to counter enemy cruise missiles, crewless aerial vehicles, and low-flying fixed-wing or rotary aircraft.

Patriot. This system fires MIM-104 surface-to-air missiles (**Fig. 13**) at attacking aircraft and missiles at ranges up to 37.3 mi (62.3 km), providing medium- and high-altitude defense. It was used in the Persian Gulf War to defend against Iraqi-launched SCUD missiles. The system consists of up to eight M901 mobile launchers, each of which contains four ready-to-fire missiles. The launcher is supported by major units including a power plant, an MPQ-53 phased-array radar, a communications center, and the crewed control station. The radar searches, detects the incoming target, and provides the electronic illumination of the target, which signals the target's location to the intercepting missile's semiactive radar homing unit. The MIM-104 missile has a 200-lb (90-kg) high-explosive warhead initiated by a radar proximity fuze.

Helicopter Armaments

The major combat aircraft used in support of ground operations are the AH-64 Apache and the AH-1 Cobra attack helicopters. Both helicopters are highly mobile and are capable of destroying moving armored columns and other point and area targets on the modern battlefield.

The AH-64 Apache is armed with the Hellfire laser-guided antitank missile system, 2.75-in. rockets, and the M230 30-mm turret-mounted chain gun. The M230 gun is used in air-to-air and air-to-ground roles.

The AH-1 Cobra helicopter is armed with point and area weapon systems: the TOW missiles, 2.75-in. rockets, and the 20-mm cannon.

The RAH-66 Comanche (formerly the LHX, light helicopter experimental) is under development to perform both reconnaissance and attack missions against airborne and ground targets. Each helicopter is equipped with three-barrel 20-mm cannon and can carry up to 14 Hellfire or 28 Stinger missiles, or 56 2.75-in. rockets. *See* HELICOPTER.

Fig. 11. M113 armored personnel carrier. (*U.S. Army*)

Fig. 12. TOW 2 (tube-launched, optically tracked, wire-guided) missile (tripod-mounted). (*U.S. Army*)

Mines

The U.S. Army places great emphasis on the use of mines in modern warfare. A family of mines and modes of delivery have been developed, which include scatterable antipersonnel and antitank mines (family of scatterable mines; FASCAM) that have high utility because they can be emplaced by artillery, aircraft, and ground vehicles at close-in and extended ranges. They are autonomous after deployment, are effective over a time period, and self-destruct to permit occupation of the area by friendly forces.

Artillery-delivered mines. The two 155-mm artillery projectiles which carry mines are the antipersonnel M692/M731 ADAM (area denial artillery munition) and the antivehicular M718/M741 RAAM (remote antiarmor mine) rounds. The ADAM contains 36 wedge-shaped mines, and the RAAM has 9 mines. The ADAM

Fig. 13. Patriot system firing an MIM-104 surface-to-air missile. (*Raytheon*)

Fig. 14. Mine dispenser of GEMSS (ground-emplaced mine scattering system), mounted on flatbed chassis. (*U.S. Army*)

Fig. 15. Volcano modular mine-dispensing system (helicopter-mounted). (*U.S. Army*)

mine is tripwire-fired and produces a pop-up fragmenting warhead; a magnetic sensor, which detects the vehicle signature, activates the RAAM mine. The 155-mm howitzer can deploy both mine systems at ranges up to 11 mi (18 km).

Ground-emplaced scatterable mines. The vehicle-mounted GEMSS (ground-emplaced mine scattering system; **Fig. 14**) and the FLIPPER mine system emplace the magnetic-influenced antiarmor M75 and the tripline-activated antipersonnel M74 mines. They are used to rapidly establish a defensive position for friendly forces. The GEMSS consists of the M128 mine dispenser mounted on the M794 flat-bed chassis. The dispenser contains two independently operated magazines holding 800 mines and a mine launcher which can be programmed to control the density and width of the minefield and to specify either a long or a short self-destruct time. The FLIPPER is used for slower, low-volume operation. It consists of the manually loaded M138 mine dispenser which can be mounted on tracked or wheeled vehicles.

Modular pack mine system (MOPMS). This system is a four-person portable M131 mine dispenser which contains 17 antitank M78 and 4 antipersonnel M79 mines. The dispenser can be activated to disperse mines by the M71 remote control unit at a range of up to 3300 ft (1000 m). The remote control unit can also reset the mine self-destruct time or can command-destruct the minefield.

Wide-area munition (WAM). This planned mine system is designed to fire a stand-off, top-attack, antiarmor, explosively formed penetrator warhead. The WAM has a lethal range up to 330 ft (100 m), and will also have countermeasure resistance. Initial deployment will be by troop emplacement with subsequent variants being capable of delivery by helicopter and missile.

Air-launched mines. The GATOR mine cluster weapon system consists of the CBU78B/CBU89B dispenser which carries BLU91B antiarmor and BLU92B antipersonnel mines. The system is deliverable by high-speed Air Force and Navy tactical aircraft. GATOR mines are also used in the VOLCANO dispensing system. Self-destruct times are field-selectable by a switch on the dispenser. The minefield density can be controlled by the number of dispensers used and dispenser settings as well as the speed and altitude of the plane.

VOLCANO is a versatile modular mine system for rapidly deploying GATOR mines from vehicles on the ground and from UH-60A helicopters (**Fig. 15**). The system comprises a dispenser control unit and one to four launcher racks, which snap into place on mounting-beam fixtures. Each rack contains up to 40 expendable M87 mine canisters, which in turn contain five antiarmor BLU91B and one antipersonnel BLU92B GATOR mines. As many as 960 mines can be dispensed from UH-60A helicopters and M817 dump trucks. Jamie Ruffing; Gary Pacella; James Wejsa

Bibliography. M. A. Blake (ed.), *Army*, October 1999; B. Davis et al., *TACOM-ARDEC Brochure*, 1997; C. F. Foss (ed.), *Jane's Armour and Artillery 1998–1999*, 19th ed., 1999; C. F. Foss and T. J.

Gander (eds.), *Jane's Military Vehicles and Logistics 1998–99*, 19th ed., 1999; *Weapon Systems 1999*, U.S. Government Printing Office, Washington, DC, 1999.

Aromatic hydrocarbon

A hydrocarbon with a chemistry similar to that of benzene. Aromatic hydrocarbons are either benzenoid or nonbenzenoid. Benzenoid aromatic hydrocarbons contain one or more benzene rings and are by far the more common and the more important commercially. Nonbenzenoid aromatic hydrocarbons have carbon rings that are either smaller or larger than the six-membered benzene ring. Their importance arises mainly from a theoretical interest in understanding those structural features that impart the property of aromaticity.

Benzenoid aromatic hydrocarbons are also called arenes. Benzene itself is the prototypical arene. The properties associated with aromaticity have little to do with aroma, although the aromatic hydrocarbons were first studied in connection with naturally occurring fragrances. Instead, these compounds possess special stability; take part in certain types of reactions; and exhibit persistence of the structural integrity of aromatic rings during chemical reactions, while groups attached to those rings are chemically altered or manipulated.

Benzene. With the molecular formula C_6H_6, benzene is highly unsaturated; it has fewer hydrogens than are required to accommodate the four valences of each carbon atom. Under appropriate conditions, benzene (**1**) can be hydrogenated to cyclohexane (**2**, C_6H_{12}), as shown in reaction (1). This result is

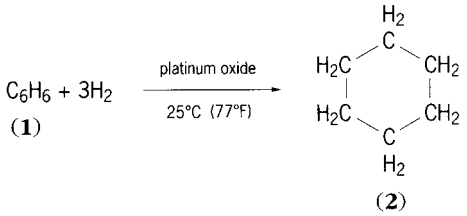

(1)

most easily explained if the six carbon atoms of benzene are linked in a ring, with three double bonds, alternating with single bonds as in structures (**3**) or (**4**)—a structure suggested as early as 1865 by

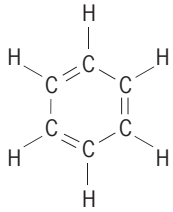

(**3**)

(**4**)

F. A. Kekulé. Quantitative study of the hydrogenation, however, showed that this structure cannot be entirely correct. For example, hydrogenation of cyclohexene (**5**) to cyclohexane is exothermic and liberates 28.6 kcal/mol of heat [reaction (2)]. Hydrogenation of 1,3-cyclohexadiene (**6**), with two double bonds, liberates approximately twice as much heat [reaction (3)]. Therefore it would be expected that

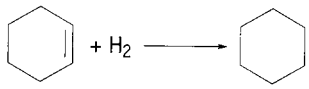

(**5**) (**2**) (2)

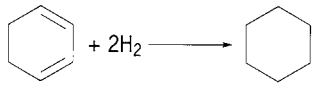

(**6**) (3)

benzene, if it had the 1,3,5-cyclohexatriene structure (**3**) or (**4**), would liberate about three times as much heat, or 84–86 kcal/mol. Experimentally, however, reaction (1) is exothermic by only 49.8 kcal/mol, meaning that addition of the first mol of hydrogen to benzene is actually endothermic by about 5 kcal/mol [reaction (4)]. Benzene is more stable than the

(**4**) (**6**) (4)

1,3,5-cyclohexatriene structure would suggest, and its stabilization energy (also called delocalization or resonance energy) is approximately 36 kcal/mol (the energy difference between the theoretical heat of hydrogenation of 1,3,5-cyclohexatriene and the experimental heat of hydrogenation of benzene, 86 − 50 = 36). Other methods for calculating the resonance energy of benzene give a similar result. *See* BENZENE.

Resonance. The double bonds in the benzene structure can be arranged in two ways, (**4**) and (**4′**).

(**4**) (**4′**)

Benzene is a resonance hybrid of these two Kekulé structures; the double-headed arrow is used to signify that the benzene structure is neither (**4**) nor (**4′**), but a single structure that is a hybrid of the two. That is, the bonds between adjacent carbon atoms are neither double nor single, but of some intermediate or hybrid type.

Physical measurements bear this out. For example, x-ray determinations show that benzene is a planar, regular hexagon. Its carbon-carbon bond lengths are all identical and 0.139 nanometer longer than a normal double bond (0.134 nm) but shorter than a single bond (0.154 nm) in ethane. All the bond angles (C—C—C and C—C—H) are 120°. *See* BOND ANGLE AND DISTANCE; RESONANCE (MOLECULAR STRUCTURE).

Molecular orbital view. Each carbon atom in benzene is connected to three atoms, two adjacent carbon atoms and a hydrogen atom. These three bonds lie in a single plane and use three of the carbon's four valence electrons. The fourth valence electron of each carbon is located in a *p* orbital, extending perpendicularly above and below the plane of the other three bonds. These electrons, one from each carbon atom and called π electrons, form three molecular orbitals located above and below, but parallel to, the plane of the ring.

The symbol of a hexagon with an inscribed circle (**7**) is often used to express the delocalized nature

(**7**)

of the π electrons in benzene and other arenes. There is physical evidence that the π electrons circulate around the ring carbons, as implied by this formula. For example, in the nuclear magnetic resonance (NMR) spectra of arenes, the chemical shifts of arene hydrogen atoms (protons) are characteristically at lower magnetic fields than those of protons attached to carbon-carbon double bonds. This difference is due to an induced magnetic field caused by circulation of the π electrons in the molecular orbitals above and below the arene ring plane. Indeed, this chemical shift difference, due to a diamagnetic ring current, is sometimes used as evidence for aromaticity in nonbenzenoid aromatic hydrocarbons. *See* DELOCALIZATION; MOLECULAR ORBITAL THEORY.

Reactions of arenes. Although benzene and other arenes react with hydrogen by addition [as in reaction (1)], they react with most other reagents by substitution, as in the formation of bromobenzene [(**8**); reaction (5)], since addition of bromine would lead

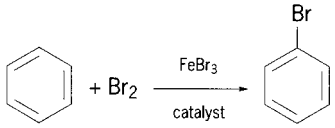

(5)

(**8**)

to a dibromocyclohexadiene in which the aromatic stabilization would be lost.

Stability of benzene ring. Whereas most unsaturated hydrocarbons are easily oxidized by, for example, potassium permanganate at room temperature, benzene is untouched by this reagent even at 100°C (212°F). The stability of the benzene ring toward oxidation is strikingly illustrated with toluene (**9**) in which the aromatic stabilization would be lost.); methylbenzene, where it is the methyl (CH_3) side chain and not the arene ring that is oxidized to yield structure (**10**), as in reaction (6). Not only is the arene ring unscathed, but the methyl group attached to the ring is more susceptible toward oxidation than would normally be the case. The methyl groups of ethane ($H_3C—CH_3$), for example, are not oxidized under similar conditions.

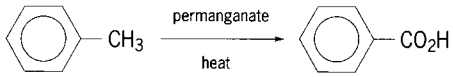

(6)

(**9**) (**10**)

Benzene can be oxidized at high temperature to the polyester intermediate maleic anhydride (**11**), as in reaction (7), which is carried out industrially on a large scale.

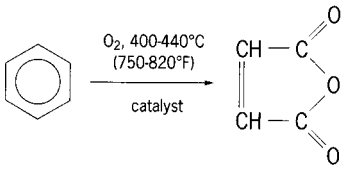

(7)

(**11**)

Arene rings and substituents. Arene rings stabilize reactive intermediates with a positive charge (cation), negative charge (anion), or odd electron (radical) on the atom attached to the ring, as in the benzyl cation (**12**), anion (**13**), or radical (**14**).

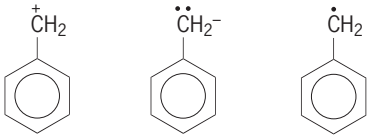

(**12**) (**13**) (**14**)

The reason is that the adjacent +, −, or · can be delocalized into the arene ring, as shown by the curved arrows for structure (**12**) in notation (8).

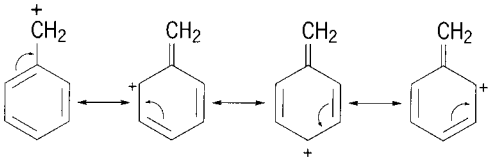

(8)

For this reason, the methyl group in toluene but not ethane is easily oxidized.

Arenes with one benzenoid ring. Besides benzene itself, several alkylbenzenes are commercially important and produced on a large scale—millions of pounds annually. Production is commonly by the cyclodehydrogenation of alkanes at high temperatures over metallic catalysts such as platinum.

Benzene, toluene, and the xylenes are added to unleaded gasoline to raise the octane number. These arenes are also essential to the petrochemical industry. Products derived from them include polyesters, polyurethanes, polystyrene, and synthetic rubber; alkylbenzenesulfonate detergents; phenol and acetone; pharmaceuticals, flavors, and perfumes; plasticizers; and many others. *See* PETROCHEMICAL.

Arenes with rather specialized structures can be produced by the cyclotrimerization of acetylenes. Transition-metal catalysts are used. Monoalkylacetylenes (**15**) may give mixtures of 1,3,5- and

1,2,4-trisubstituted benzenes[reaction (9)]. With

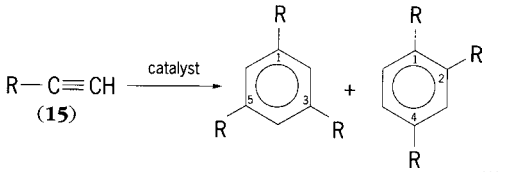

(9)

appropriate substituents, diacetylenes and monoacetylenes can be co-"trimerized" to give specifically substituted benzenes (**16**) in high yields, as in reaction (10), where R = $(CH_3)_3Si$ and n = 0–2.

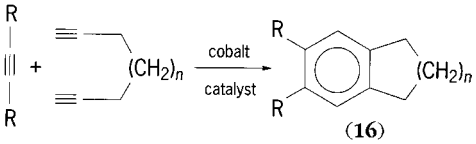

(10)

Arenes with linked benzene rings. Benzene rings can be linked, either directly as with biphenyl (**17**), terphenyls (**18**) and others, or with one or more

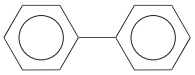

(**17**)

(**18**)

nonaromatic carbons between, as with diphenylmethane (**19**) and triphenylmethane (**20**). *Para*-polyphenylene (**21**) is an important thermally stable

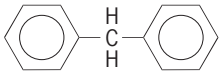

(**19**)

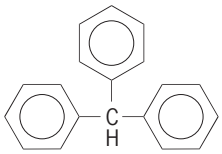

(**20**)

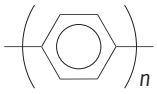

(**21**)

polymer. Triphenylmethane provides the carbon framework for a major class of dyes and food colorings.

The two arene rings in biphenyl are not coplanar but tilted about 45° with respect to each other. Rotation around the bond that joins the two rings is facile, but if large groups are placed in the *ortho*-positions of each ring, rotation becomes difficult and leads to separable mirror-image isomers (**22a**) and (**22b**).

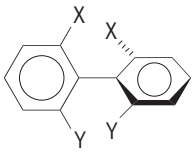

(**22a**)

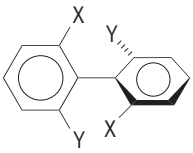

(**22b**)

Such molecules are chiral and differ from one another only in a left- and right-handed way. *See* STEREOCHEMISTRY.

Arene rings can be linked through three-dimensional frameworks. An interesting example is triptycene (**23**), which has three arene rings held rigidly at angles of 120°. The framework is useful for studying interactions of groups attached to the rings and held at fixed angles and distances from one another. Triptycene is the first of a general class known as iptycenes, another of which is pentiptycene (**24**), with five fixed arene rings.

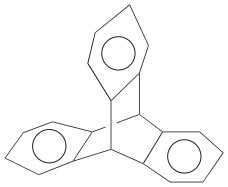

(**23**)

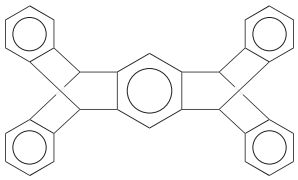

(**24**)

Fused rings. Arenes with fused rings are also known as polynuclear aromatic hydrocarbons. Rings are said to be fused when they share two carbon atoms. The simplest example is naphthalene, a colorless crystalline compound found in coal tar, best known as a moth repellent, whose structure can be represented by the various forms shown in notation (11). Though planar, not all the carbon-carbon

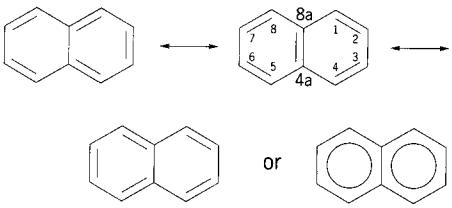

(11)

bond lengths in naphthalene are identical. The C_1–C_2 and C_{4a}–C_{8a} bonds are shorter (0.137 and 0.139 nm) than the C_2–C_3 and C_1–C_{8a} bonds (0.140 and 0.142 nm), suggesting that the resonance contributor shown with the numbering and having double bonds at the short bond distances contributes more to the resonance hybrid than do the other two contributors. The NMR spectrum of napthalene indicates a diamagnetic ring current, although the resonance energy is somewhat less than twice that of benzene.

Additional arene rings can be fused. For example, anthracene, tetracene, and pentacene are linearly fused, while phenanthrene, triphenylene, and pyrene are angularly fused (**Fig. 1**). In general, angular fusion results in more stable systems than linear

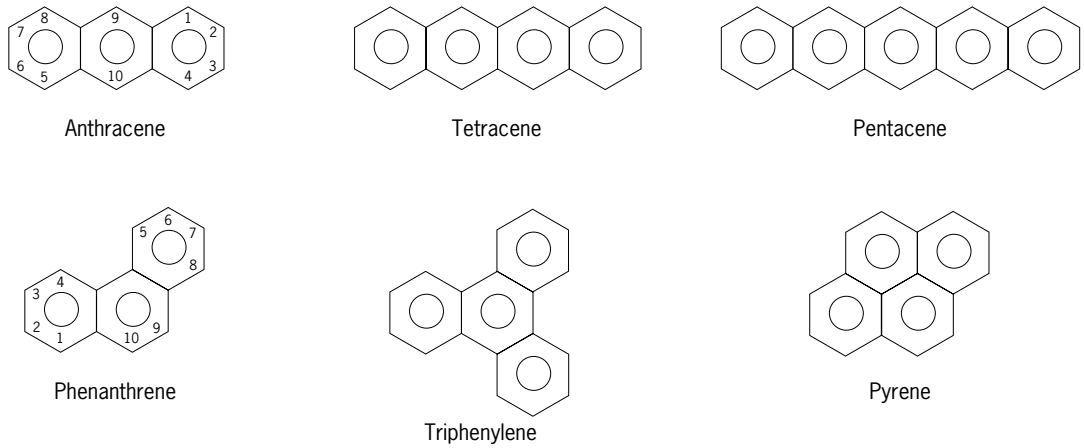

Fig. 1. Structures of some fused-ring (polynuclear) aromatic hydrocarbons.

fusion. Phenanthrene, for example, is about 6 kcal/mol more stable than its linear isomer anthracene. Stability falls off sharply in the linearly fused series, and compounds with more than seven such rings are unknown.

Anthracene and phenanthrene undergo addition reactions at the 9,10-positions [reactions (12) and (13)]. The products, each with two fully aro-

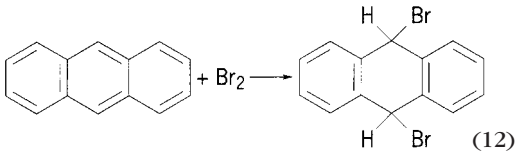

(12)

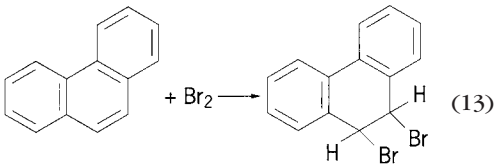

(13)

matic benzene rings, retain more than two-thirds of the original resonance energy. Even so, these bromine adducts readily eliminate hydrogen bromide (HBr) on heating to give the fully aromatic substitution products, 9-bromoanthracene or 9-bromophenanthrene.

Some polynuclear aromatic hydrocarbons are carcinogenic. Examples include benz[a]pyrene, 7,12-dimethylbenz[a]anthracene, and cholanthrene. Such compounds are environmental pollutants, being formed during the combustion of fuel oil, in the burning of cigarettes, in forest fires, and in roasting red meats. In the body, these hydrocarbons are enzymatically oxidized to related structures that react with deoxyribonucleic acid (DNA) bases, causing mutations that lead to cancer.

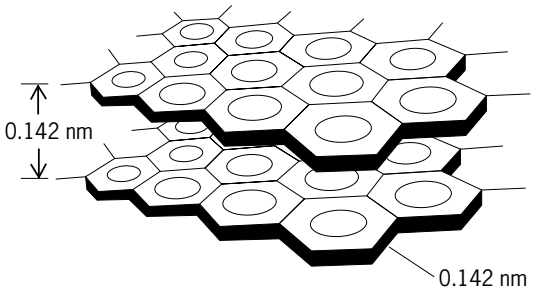

Fig. 2. Structure of graphite.

Increasing the number of ring fusions results in higher carbon:hydrogen (C/H) ratios. For example, the C/H ratio for benzene is 1.0; for naphthalene, 1.25; for anthracene, 1.4; and for pyrene, 1.6. If fusions are extended infinitely in two dimensions, sheets of carbon atoms are obtained that are arranged in hexagons, the known structure of graphite (**Fig. 2**). There are no formal bonds between the graphite layers. *See* GRAPHITE; POLYNUCLEAR HYDROCARBON.

Bent and deformed benzene rings. Although arene rings are normally flat, they are more flexible than was originally thought. They can be bent substantially without losing their aromaticity.

Coronene (**25**) is a planar arene present in coal tar.

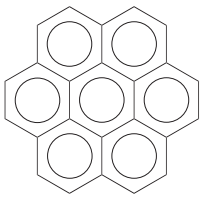

(**25**)

If two of the peripheral rings are disconnected and each completed, a planar geometry is no longer possible, and the molecule, called hexahelicene (**26**), assumes a spiral or helical arrangement of the

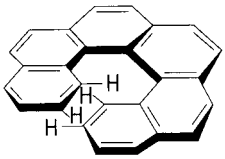

(**26**)

rings. Each arene ring is distorted a bit to accommodate the strain. The molecule can exist in two mirror-image forms, depending on whether the helix is right or left handed. Helicenes with additional rings (7–14) in the spiral have been prepared.

Fusion of small rings to an arene ring, as in structures (**27**)–(**31**), distorts the structure and imparts

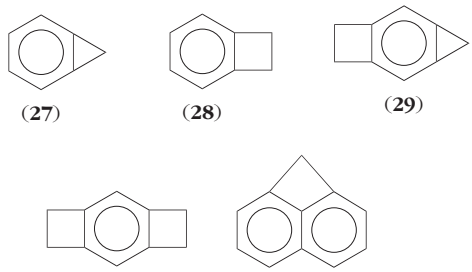

(**27**) (**28**) (**29**) (**30**) (**31**)

unusual reactivity to the arene ring. Benzocyclobutene (**28**), for example, undergoes thermal ring-opening to the highly reactive tetraene [(**32**); reaction (14)]; release of strain compensates for the loss of aromatic resonance energy. Despite the strain,

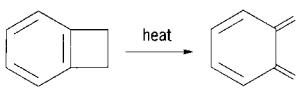

(**28**) (**32**) (14)

fusion of more than one small ring to a single arene ring is possible [structures (**29**) and (**30**)]. Spectra of such compounds show that aromaticity is retained (for example, NMR spectra show the usual chemical shifts due to π-electron ring currents).

Heavy adjacent substitution distorts arene rings. In octamethylnaphthalene (**33**), for example, the

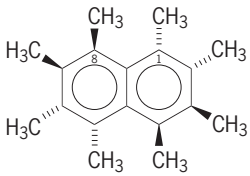

(**33**)

methyl groups and the carbon atoms to which they are attached are bent from planarity because of steric requirements of the substituents. The molecule is like a two-bladed propeller, and the angle between the C_1 and C_8 methyl substituents is about 40°. Yet spectral properties show that aromaticity is preserved.

Cyclophanes are molecules in which an arene ring is part of a larger ring, as in *para*-cyclophanes (**34**) and *meta*-cyclophanes (**35**). If n or m is large the

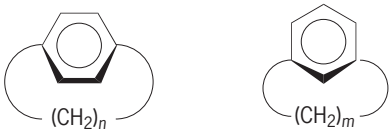

(**34**) (**35**)

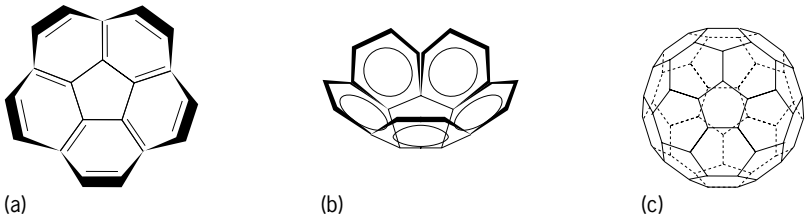

Fig. 3. Bent benzenoid structures. (*a*) Corannulene, showing the five arene rings fused to a central five-membered ring. (*b*) Corannulene, showing the cup shape. (*c*) Buckminsterfullerene, a spherical allotrope of carbon.

arene ring remains planar, but if small ($n = 6$, $m = 5$) the ring will distort, and if even smaller the molecule becomes impossible to construct. Two or more arene rings may be incorporated in a single cyclophane. In [2.2]paracyclophane (**36**), the "end" carbons are

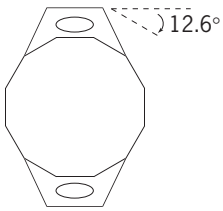

(**36**)

bent from planarity by 12.6°, yet the arene rings undergo most reactions (for example, substitution) characteristic of benzenoid compounds. Analogs of compound (**36**) have been synthesized in which from three to all six of the arene ring positions are bridged. *See* MACROCYCLIC COMPOUND.

Corannulene (**Fig. 3***a*) has five arene rings fused to the central five-membered ring. The molecule, which cannot be planar without deforming the hexagons, is cup shaped (Fig. 3*b*). This structural unit is present in the spherical allotrope of carbon, C_{60} (Fig. 3*c*), known as buckministerfullerene; it has no hydrogens and is not a hydrocarbon. Nevertheless, it is aromatic, since the bonding resembles that of benzene. That is, each carbon atom is covalently attached to three other carbons, and the fourth valence electron is in a p orbital that is perpendicular to the surface of the sphere. C_{60} has a π-electron cloud spread over the outside and inside surfaces of the sphere, just as benzene has a π-electron cloud above and below the plane of the hexagon. Though it has its own distinctive chemistry, C_{60} displays special stability associated with aromaticity. *See* FULLERENE.

Nonbenzenoid aromatic hydrocarbons. Given the special aromatic properties associated with the formal arrangement of alternating double and single bonds inscribed in a hexagon, it was natural to explore other ring sizes with similar arrangements. Cyclobutadiene (**37**) was exceptionally difficult to synthesize, and it is extremely reactive. Though it has been trapped in a rigid argon matrix at −265°C (−445°F), it dimerizes [reaction (15)] even at

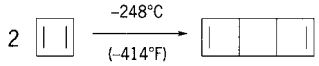

(**37**) (15)

−248°C (−414°F) when the matrix softens. X-ray studies of more stable tetra-substituted analogs of cyclobutadiene show rectangular rather than square geometry, with alternate single and double bonds; that is, the π electrons are not delocalized as they are in benzene.

Even if one of the double bonds is incorporated in a benzene ring, as in (**38**), the compound resists isolation and rapidly dimerizes. Biphenylene (**39**),

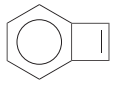

(**38**)

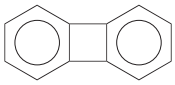

(**39**)

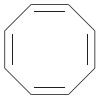

(**40a**)

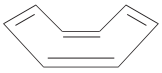

(**40b**)

however, with both of cyclobutadiene's double bonds as part of a benzene ring, is a stable white crystalline compound.

Cyclooctatetraene (**40a**) has the properties of a cycloalkene with four double bonds. The structure is nonplanar and tub shaped (**40b**), and the π electrons are localized at the double bonds.

Hückel rule. Originally it was thought that cyclobutadiene was not aromatic and unstable because of the strain associated with the four-membered ring, and that cyclooctatetraene was not aromatic because it was nonplanar. But the difference between these molecules and benzene turned out to be more fundamental. From molecular orbital theory, E. Hückel derived the rule that planar, cyclic conjugated (alternate single and double bonds) systems with $4n + 2$ π electrons (n is an integer, 0, 1, 2...) will be aromatic and have substantial resonance energy, whereas those with $4n$ such electrons will not; indeed, it was later shown that $4n$ systems are often destabilized, hence antiaromatic.

Benzene is a $4n + 2$ system ($n = 1$) and aromatic; cyclobutadiene is a $4n$ ($n = 1$) system and antiaromatic. Cyclooctatetraene would be a $4n$ system ($n = 2$) if it were planar, but since such a system would be antiaromatic, there is no driving force for the molecule to become planar. As striking confirmation of these ideas, pentalene (**41**), a planar analog of cyclooctatetraene, is exceptionally reactive, unstable, and antiaromatic (a $4n$ system, $n = 2$), whereas the purple hydrocarbon azulene (**42**; a $4n = 2$ system, $n = 2$, and an isomer of naphthalene) is stable and undergoes substitution reactions analogous to those of benzenoid arenes.

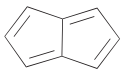

(**41**)

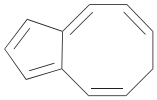

(**42**)

Annulenes. Monocyclic conjugated systems $(CH)_n$, where n is twice the number of double bonds, are called annulenes (benzene, where $n = 6$, is a [6]annulene). According to the Hückel rule, annulenes with 6, 10, 14, 18. . . π electrons, if planar, should be aromatic. In [10]annulene (**43**) with two trans double bonds, the inside hydrogen atoms must occupy the same space. Replacement of these hydrogens by a —CH_2— group as in (**44**) yields a "bridged"

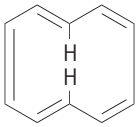

(**43**)

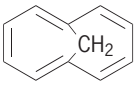

(**44**)

annulene with a nearly planar 10-π-electron system. In accord with the Hückel theory, the compound is stable and shows aromatic properties (diamagnetic ring current, substitution reactions).

The [14]annulene (**45**) and [18]annulene (**46**)

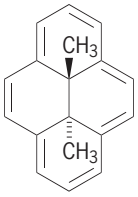

(**45**)

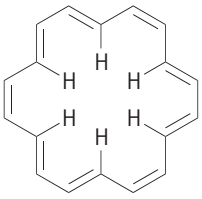

(**46**)

are among the many $4n + 2$ aromatic annulenes that have been synthesized. Despite the numerous nonplanar conformations accessible to such molecules, they tend to be planar. The 12 "outside" protons in (**46**) are strongly deshielded (NMR chemical shift 9.28) as in benzene, but the 6 "inside" protons are strongly shielded (NMR chemical shift −2.99). This result is entirely in accord with a strong ring current around the periphery of the molecule.

Annulenes with $4n$ π electrons have also been synthesized ($n = 3, 4, 5\ldots$), and in general they are less stable than the $4n + 2$ systems. The proton NMR spectra of the [12] through [24] annulenes alternate, being diamagnetic (aromatic) for the $4n + 2$ members and paramagnetic (nonaromatic) for the $4n$ members, consistent with the Hückel rule.

Kekulene, a cycloarene with 12 arene rings, provides an interesting comparison of benzenoid and nonbenzenoid arenes. Two of its possible π-electron arrangements are shown in (**47a**) and (**47b**). The former has six benzenoid rings connected by single and double bonds. The latter has no benzenoid rings but two $4n + 2$ annulene rings, the "inner" [18] ($n = 4$) and the "outer" [30] ($n = 7$). X-ray and NMR data all support structure (**47a**) as the best representation.

Aromatic ions. One of the most striking confirmations of Hückel theory came with its prediction of

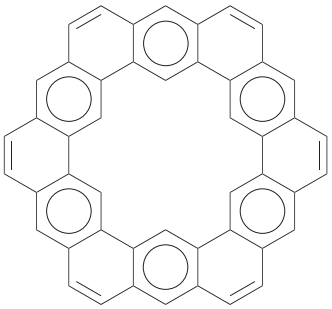

(**47a**)

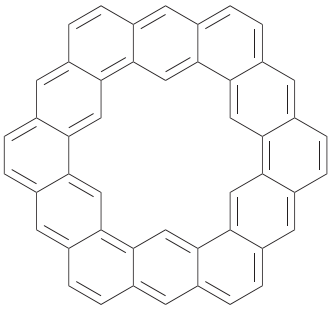

(**47b**)

aromatic ions; the concept of aromaticity is not restricted to neutral molecules. For example, the cyclopropenyl cation (**48**), cyclopentadienyl anion (**49**), and tropylium cation (**50**) follow the Hückel $4n + 2$

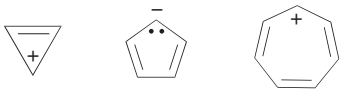

(**48**) (**49**) (**50**)

rule, with $n = 0$, 1, and 1 respectively. Each of these ions shows exceptional stability, and in each the ionic charge is delocalized over all ring atoms. A particularly striking example is seen with cyclooctatetraene. Though nonplanar and not aromatic, the hydrocarbon (**40a**) becomes planar when reduced to the dianion (**51**) which, being a Hückel system

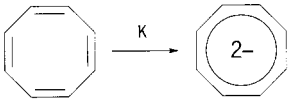

(16)

(**40a**) (**51**)

($n = 2$), is aromatic, as in reaction (16) [K = potassium].

Harold Hart

Bibliography. A. T. Balaban, M. Banciu, and V. Ciorba, *Annulenes, Benzo-, Hetero-, Homo-derivatives and their Valence Isomers*, vols.1–3, 1987; D. H. R. Barton and W. Ollis (eds.), *Comprehensive Organic Chemistry*, vol. 1, ed. by J. F. Stoddard, 1979; J. R. Dias, *Handbook of Polycyclic Hydrocarbons*, 1987; F. Diederich, *Cyclophanes*, 1991; P. J. Garratt, *Aromaticity*, 1986; M. Hornby and J. M. Peach, *Foundations of Organic Chemistry*, 1993; D. Lloyd, *Nonbenzenoid Conjugated Carbocyclic Compounds*, 1984.

Aromatization

The conversion of any nonaromatic hydrocarbon structures, especially those found in petroleum, to aromatic hydrocarbons. There are numerous routes and means to accomplish this transformation, the simplest and most important of which are direct dehydrogenation of naphthenes to aromatics, reaction (1); dehydroisomerization of naphthenes to aromatics, reaction (2); dehydrocyclization of aliphatics to aromatics, reaction (3); and high-temperature condensation of hydrocarbons to aromatics, reaction (4).

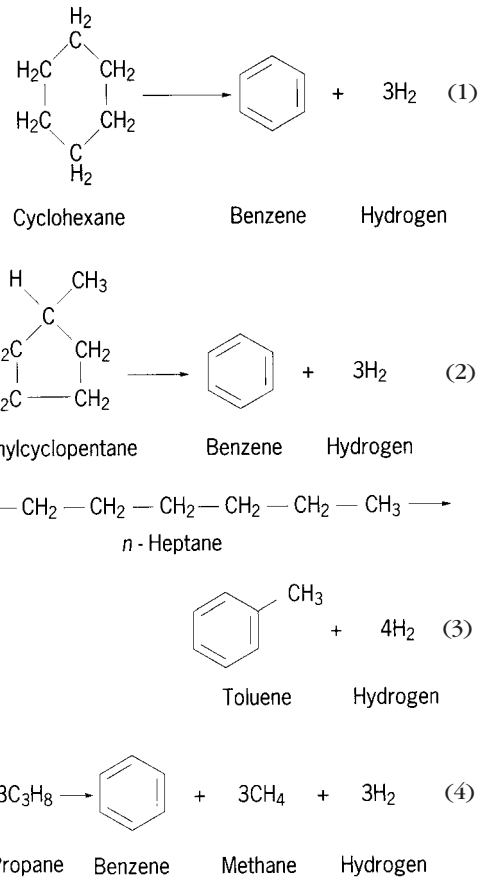

Reaction (1) was performed on a huge scale by the petroleum industry during World War II for the production of toluene for TNT, starting from methylcyclohexane and utilizing molybdenum oxide on porous activated alumina or mixed tungsten-nickel sulfides as catalysts. At that time, reaction (3), employing chromia on alumina catalysts, was investigated extensively but was found impracticable commercially.

Subsequently, reforming of naphthas with catalysts comprising small amounts of platinum on an acidified alumina support provided a means of aromatization that rapidly displaced the earlier processes, since, it accomplishes reactions (1), (2), and (3) readily and simultaneously. It is a major process for benzene, toluene, and other aromatics from petroleum sources.

Reaction (4) merely illustrates one type of reaction

that may occur in the high-temperature (600–800°C or 1100–1500°F) thermal cracking of petroleum fractions. Small hydrocarbons from the cracking of heavier components, as well as any originally present, may condense to aromatics, usually with considerable splitting to methane as well as hydrogen. Thermal cracking processes emphasizing this type of reaction have achieved only limited use in the petroleum industry. *See* PETROLEUM PROCESSING AND REFINING.

Bernard S. Greensfelder; Mott Souders

Bibliography. W. A. Gruse and D. R. Stevens, *Chemical Technology of Petroleum*, 3d ed., 1960; *Kirk-Othmer Encyclopedia of Chemical Technology*, 3d ed., vol. 1, 1978; E. I. Shaheen, *Catalytic Processing in Petroleum Refining*, 1983; J. G. Speight, *Chemistry and Technology of Petroleum: An Introduction*, 1982.

Arsenic

A chemical element, symbol As, atomic number 33. Arsenic is found widely distributed in nature (approximately 5×10^{-4} of the Earth's crust). It is one of the 22 known elements composed of only one stable nuclide, $^{75}_{33}$As; the atomic weight is 74.92158. There are 17 other radioactive arsenic nuclides known.

1												13	14	15	16	17	18
1 H	2																2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

There are three polymorphic modifications of arsenic. The yellow cubic α-form is made by condensing the vapor at very low temperatures. The black β-polymorph is isostructural with black phosphorus. Both these modifications revert to the stable γ-form, gray or metallic, rhombohedral arsenic, on heating or exposure to light. The metallic form is a moderately good thermal and electric conductor and is brittle, easily fractured, and of low ductility. Some of the important atomic and bulk properties of arsenic are given in the **table**.

Arsenic is found native as the mineral scherbenkobalt, but generally occurs among surface rocks combined with sulfur or metals such as Mn, Fe, Co, Ni, Ag, or Sn. The principal arsenic mineral is FeAsS (arsenopyrite, mispickel); other metal arsenide ores are $FeAs_2$ (löllingite), NiAs (nicolite), CoAsS (cobalt glance), NiAsS (gersdorffite), and $CoAs_2$ (smaltite). Naturally occurring arsenates and thioarsenates are common, and most sulfide ores contain arsenic. As_4S_4 (realgar) and As_4S_6 (orpiment) are the most important sulfur-containing minerals. The oxide, arsenolite, As_4O_6 is found as the product of the weathering of other arsenical minerals, and is also recovered from flue dusts collected during the extraction of Ni, Cu, and Sn from their ores; it also results when the arsenides of Fe, Co, or Ni are roasted in air or oxygen. The element may be obtained by roasting FeAsS or $FeAs_2$ in the absence of air or by reduction of As_4O_6 carbon, when As_4 may be sublimed away.

Some atomic and bulk properties of arsenic

Property	Value
Electron configuration (S^4 ground state)	[Ar] $3d^{10}4s^24p^3$
Covalent radius	121 pm
Ionic radius (As^{3+})	69 pm
Metallic radius	139 pm
Ionization energies 1st–6th in kJ mole^{-1}	947, 1798, 2734, 4834, 6040, 12300
Electrode potential, As^{3-}/As	0.25 V
Electronegativity (Allred-Rochow)	2.20
Oxidation numbers	−3, 0, +3, +5
Specific gravity (α, β, γ)	2.026, 4.7, 5.727
Melting point (γ)	817°C (1503°F) (3.6 MPa)
Boiling point (γ)	616°C (1141°F) (sublimes)
Electrical resistivity	33.3 $\mu\Omega$ cm (273 K)
Toxicity level	0.5 mg · m^{-3} of air

Elemental arsenic has few uses. It is one of the few minerals available in 99.9999+% purity, which is largely used in the laser material GaAs and as a doping agent in the manufacture of various solid-state devices. Arsenic oxide is used in glass manufacture. The arsenic sulfides are used as pigments and in pyrotechnics. Dihydrogen arsenate is used in medicine, as are several other arsenic compounds. Most of the medicinal uses of arsenic compounds depend on their toxic nature. *See* ANTIMONY; PHOSPHORUS.

John L. T. Waugh

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. F. Shriver and P. W. Atkins, *Inorganic Chemistry*, 3d ed., 1999.

Art conservation chemistry

The application of chemistry to the technical examination, authentication, and preservation of cultural property. Chemists working in museums engage in a broad range of investigations, most frequently studying the chemical composition and structure of artifacts, their corrosion products, and the materials used in their repair, restoration, and conservation. The effects of the museum environment, including air pollutants, fluctuations in temperature and relative humidity, biological activity, and ultraviolet and visible illumination, represent a second major area of research. A third area of interest is the evaluation of the effectiveness, safety, and long-term stability of materials and techniques for the conservation of works of art. Though analytical techniques appear to dominate, many other areas of chemistry, biology, physics, and engineering, including polymer

chemistry, kinetic studies, imaging methodologies, biodegradation studies, dating methods, computer modeling, metallography, and corrosion engineering, play active roles in conservation science.

Examination and analysis. Methods of examination may be divided into two classes: those that provide an image of the entire object (holistic examination; **Fig. 1**) or a section of it; and those that provide an analysis at a point on the object, with or without sampling. Nondestructive methods, not requiring sampling, are always preferable. However, modern methods of analysis can be employed on such minute samples that they are in effect nondestructive. In some cases, samples must be taken for methods that are in principle nondestructive because the object is too large to fit into a sample chamber. The ability to analyze minute samples introduces the serious concern that the sample may not be representative of the composition of the artifact but may be an inclusion or contaminant introduced by the experimentalist. With specimens from painted surfaces, great care must be taken to identify areas of restoration.

A further concern arises from the differing depths from which signals originate. On a metal surface, ion scattering spectrometry (ISS) would see the initial fraction-of-a-nanometer, predominantly adsorbed species and contaminants. Secondary ion mass spectrometry (SIMS) would begin to penetrate the oxidized area; Auger electron spectrometry (AES) would examine the bulk of the oxidized layer; and x-ray-induced photoelectron spectrometry (XPS) would give data on the bulk sample some 10 nm below the specimen surface. In the analysis of silver artifacts by neutron activation, significant differences were observed in results for samples taken by drilling when compared with streak samples obtained by rubbing a quartz slide across the surface of the artifact. *See* ACTIVATION ANALYSIS; AUGER EFFECT; NONDESTRUCTIVE EVALUATION; SECONDARY ION MASS SPECTROMETRY (SIMS); SURFACE PHYSICS; X-RAY FLUORESCENCE ANALYSIS.

The most commonly employed holistic method is x-ray radiography, where variations in the density and average atomic number of the sample attenuate an x-ray beam, leaving a negative image on film. Ordinary medical x-ray equipment is routinely used with paintings and small three-dimensional objects, while high-energy 300-kV x-ray radiography is used for large stone and metal sculpture. Watermarks are detected in prints and drawings by β-radiography with a poly(methyl methacrylate) plate doped with carbon-14. Computerized tomography (CAT) scans and nuclear magnetic resonance (NMR) imaging have been applied to Egyptian mummies both to assess their condition and to conduct physical anthropological studies. *See* COMPUTERIZED TOMOGRAPHY; NUCLEAR MAGNETIC RESONANCE (NMR); RADIOGRAPHY.

Other methods, such as ultraviolet and infrared reflectance and fluorescence, are used to show areas of compositional difference indicating restoration or variation in the pigments used by the artist. Infrared reflectography has been most successfully employed in the study of preparatory drawings under the paint layers in Netherlandish paintings and has been introduced to the study of underdrawings in medieval manuscript illumination. *See* INFRARED IMAGING DEVICES; INFRARED RADIATION; LUMINESCENCE ANALYSIS; ULTRAVIOLET RADIATION.

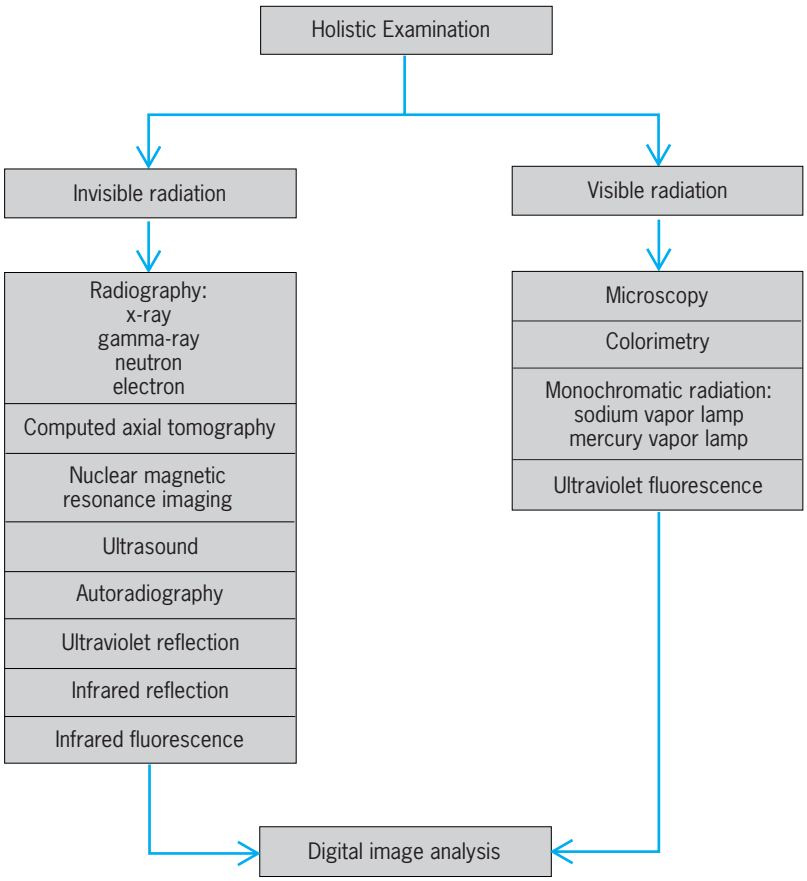

Fig. 1. Methods for holistic investigation of art objects. (*After F. Mairinger and M. Schreiner, New methods of chemical analysis: A tool for the conservator, in N. S. Bromelle and G. Thomson, eds., Science and Technology in the Service of Conservation, International Institute for Conservation, 1982*)

In the examination of paintings, small samples are taken under the binocular microscope, embedded in transparent resin, and polished to produce a cross section for microscopic examination. This permits a study of the artist's painting technique and shows how several layers may have been built up to achieve a desired effect. The same specimens may then be subjected to electron beam microprobe examination. The microprobe in a sequence of point analyses permits the creation of a holistic picture of the elemental distribution through the several layers of the paint specimen. A series of elemental scans can be combined by image-processing software to give a color-coded mapping of the pigments distributed throughout the specimen. *See* CHEMICAL MICROSCOPY; IMAGE PROCESSING; MICROSCOPE.

Conservation studies of composition and technique embrace the entire spectrum of modern chemical analysis (**Fig. 2**).

Authentication. The separation of the fake from the authentic is a small but often spectacular aspect

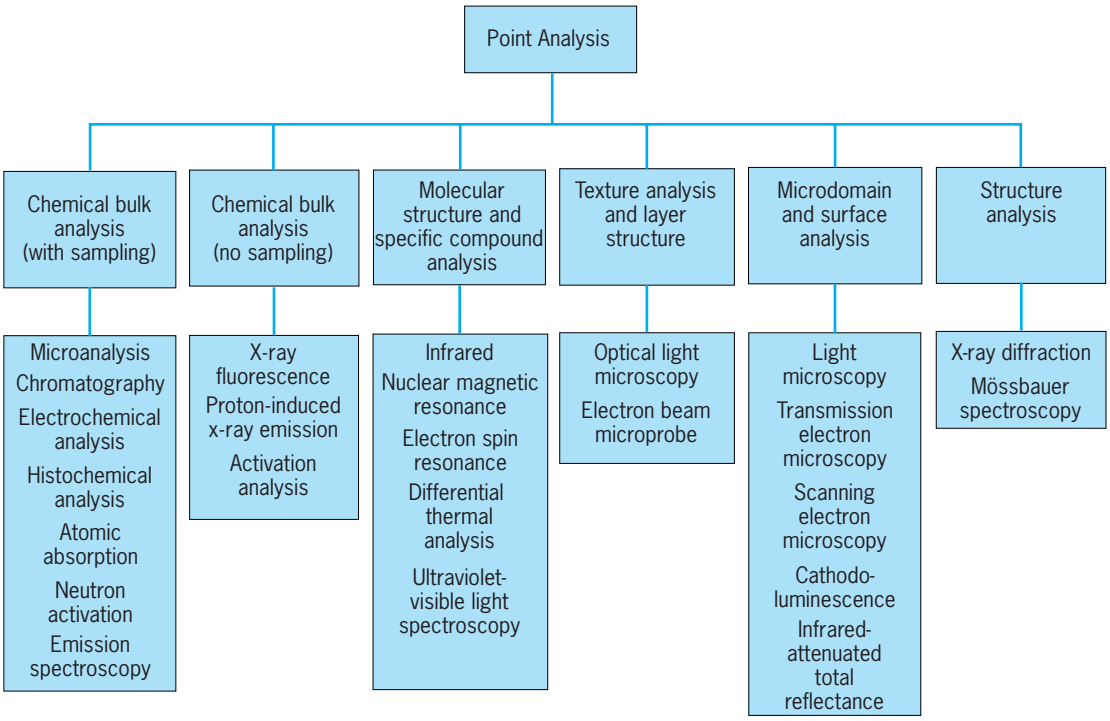

Fig. 2. Methods for point analysis applied to the examination of works of art. (*After F. Mairinger and M. Schreiner, New methods of chemical analysis: A tool for the conservator, in N. S. Bromelle and G. Thomson, eds., Science and Technology in the Service of Conservation, International Institute for Conservation, 1982*)

of the technical examination of artifacts. In some cases, direct age determination (dendrochronology for panel paintings, fission track dating for uranium glass, radiocarbon dating for organic materials, thermoluminescence dating for ceramics) is possible. The development of direct counting methods for radiocarbon dating promises to expand the application of this technique as the minimum sample size has been reduced from grams to milligrams. In 1988, the Shroud of Turin was declared a medieval forgery based on radiocarbon dates obtained by three independent laboratories. *See* ARCHEOLOGICAL CHRONOLOGY; DATING METHODS; DENDROCHRONOLOGY; FISSION TRACK DATING; RADIOCARBON DATING.

More commonly, the issue of authenticity turns upon anachronisms in composition or technique when the artifact in question is compared to accepted artifacts of the period. Thus, the greater part of the work in the conservation laboratory concerns the building of databases of analyses of composition, trace-element distributions, and studies of technique. For example, the analysis of 1000 Sassanian silver coins shows gold impurity levels (**Fig. 3**) in modern fake coins to be distinct from those of authentic specimens. In some cases, the silver used was so pure (less than 0.001% gold) that it could only have been refined in the twentieth century.

Anachronisms in the pigments used have unmasked many forgers, including Hans Van Meergeren (1889–1947), who forged and sold a series of paintings in the style of Vermeer to major Dutch collections. The presence of one of these paintings in the collection of the Nazi leader Hermann Goering led to a notorious trial that relied heavily on technical analysis, including pigment identification, x-ray radiography, and microchemical demonstration that the paint medium was a twentieth-century phenol-formaldehyde resin, selected to imitate the reduced solubility of aged drying oil media. In 1972, pyrolysis

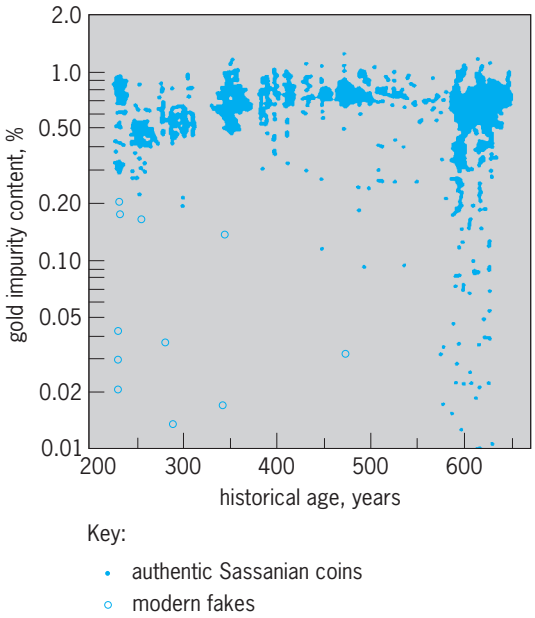

Fig. 3. Gold impurity levels versus historical age for Sassanian (Near East, A.D. 224–641) coins. Data points representing modern fakes indicate date of the originals that the fake coins intended to imitate. (*After S. J. Fleming, Authenticity in Art: The Scientific Deduction of Forgery, Crane Russak Co., 1976*)

gas chromatography was used to identify the specific resin used in the most famous of these faked Vermeers, *Christ at Emmaus*, sold to the Museum Boymans in Rotterdam. *See* PIGMENT (MATERIAL).

A demonstration of the limits of technical examination is provided by a comprehensive study of the painting *Deborah Kip, Wife of Sir Balthasar Gerbier, and Her Children* (*The Gerbier Family*) by Peter Paul Rubens, acquired by the National Gallery of Art, Washington, D.C., in 1971. Among the techniques employed were x-ray radiography, optical microscopy, emission spectroscopy, and electron beam microprobe study of the pigments and paint layer structure. Transmission and scanning electron microscopy were applied to the chalk white in the ground layer of the painting in order to identify coccoliths, calcareous remains of unicellular marine algae. Lead isotope mass spectrometry was used to develop a database of the distribution of lead isotope ratios in lead-white samples from all paintings examined: all paintings examined painted before 1800; Dutch, Flemish, and Belgian paintings painted before 1800; and different portions of *The Gerbier Family* (**Fig. 4**). The ever-narrowing distribution of lead isotope ratios reflecting ever-narrowing sources of supply and the pigments observed were highly consistent throughout the painting. Yet, despite the wealth of technical data, it could not be stated with certainty that Rubens had painted the entire picture, or that the six portions of canvas that had been sewn together to produce this large painting were painted at the same time by the same hand. *See* ELECTRON MICROSCOPE; EMISSION SPECTROCHEMICAL ANALYSIS; LEAD ISOTOPES (GEOCHEMISTRY); MASS SPECTROMETRY.

Pigment analysis by optical light microscopy, x-ray diffraction, and ion microprobe of ultramicro samples from the Vinland Map owned by Yale University identified significant amounts of anatase, a modern titanium dioxide pigment first available in the early twentieth century. The map, thought to be a fifteenth-century copy, was purchased as the earliest map of the New World providing evidence that Leif Ericsson had visited North America some 500 years before Columbus. Researchers using proton-induced x-ray emission (PIXE) have attempted to challenge these pigment findings, claiming that their elemental analyses show only small amounts of titanium in the map. These results are less persuasive than those of the earlier analysis, since in that work the microscopic pigment particles were positively identified while in the proton-induced x-ray emission PIXE analyses an average elemental analysis over a much wider area was obtained. *See* PROTON-INDUCED X-RAY EMISSION (PIXE).

Substantial collections of notable occurrences of pigments in paintings have been developed, permitting estimates of the probability of finding a specific pigment in a painting of a given period. Typically, in paintings from 1300 to 1800 there is almost 100% certainty that lead white [$2PbCO_3 \cdot Pb(OH)_2$] will be detected, while in paintings of around 1900 estimates of the probability of finding certain pigments are as

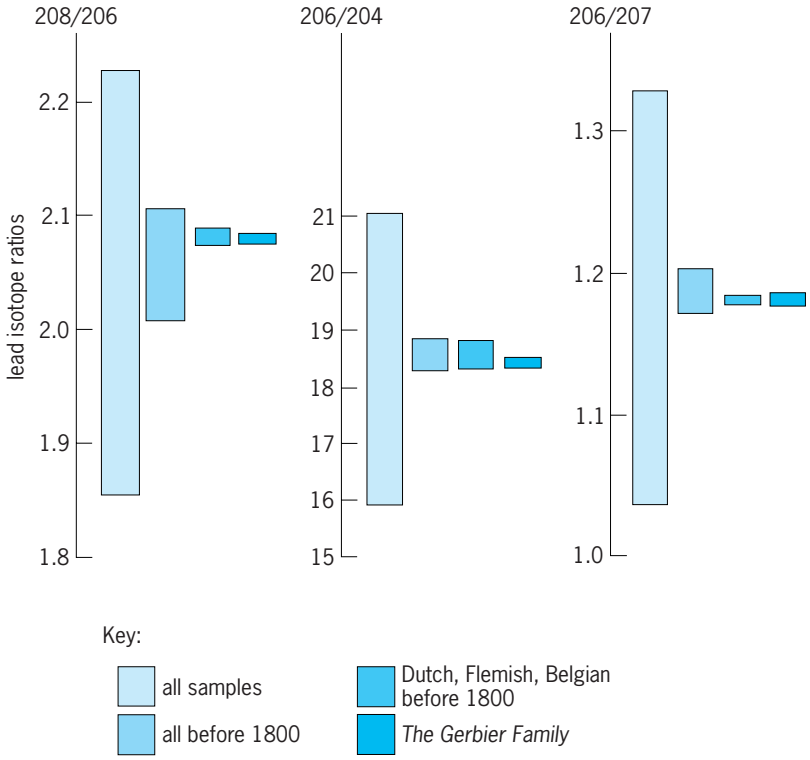

Fig. 4. Distribution of lead isotope ratios in lead-white pigment samples taken from paintings. (*After B. Keisch and R. C. Callahan, Rubens' The Gerbier Family: Investigation by Lead Isotope Mass Spectrometry, National Gallery Studies in the History of Art, pp. 75–78, 1973*)

follows: lead white, 90%; zinc white, 50%; and barium sulfate, as an extender in white pigments, 50%. The data show that several pigments of the same hue will frequently appear in a single painting.

Other studies have used thin-layer chromatography to follow the introduction of natural dyes such as cochineal from the New World in Flemish tapestries of the fifteenth and sixteenth centuries. Cochineal, derived from scale insects, was discovered in Mexico but was extensively imported to Europe shortly after, since its properties were superior to those of kermes, the principal insect dye used in Europe up to that time. Similar studies on yellow dyes taken from textiles of the period 1500–1850 demonstrated that 80% were weld (dyers' weed), which was used in both European textiles and Anatolian carpets. *See* ARCHEOLOGICAL CHEMISTRY; DYE; TEXTILE CHEMISTRY.

Museum environment. Many artifacts are sensitive to destructive agents in the museum atmosphere. Rapid changes in relative humidity will cause dimensional changes in wood furniture, polychrome sculpture, and panel paintings, leading to cracking and splitting of the wood with loss of painted surface decoration. High relative humidity can lead to mold growth and foxing on books and prints, while low relative humidity will cause photographic prints and films to become brittle. Chemists have developed treated silica gel to stabilize the relative humidity where sensitive artifacts are displayed in sealed cases. *See* HUMIDITY; HUMIDITY CONTROL.

Fig. 5. Lead bullet with lead formate corrosion product formed by reaction with formic acid released from display case materials. (*Smithsonian Institution*)

Oxidation of iron objects, tarnishing of silver plate, and the development of corrosion products on lead artifacts by the action of formic and acetic acids emitted by wooden display cases have regularly been observed in museums. Lead formate was identified by infrared spectroscopy as the corrosion product (**Fig. 5**) found on a lead bullet in a plywood-backed display case. Some artifacts are made of unstable materials and decompose if left untreated. For example, cellulose nitrate film has, through spontaneous combustion, caused major fires in film repositories. Books and works of art on paper supports made of acid paper rapidly degrade by acid hydrolysis. Methods for treating acid paper with an alkaline reserve have been developed, including the Library of Congress diethyl zinc vapor phase process for books and a process using a solution of methyl magnesium carbonate to deposit magnesium carbonate in acid paper. *See* PAPER.

The common air pollutants sulfur dioxide (SO_2) and ozone (O_3) have been monitored at elevated levels in museums, libraries, and archives. These pollutants cause the degradation of leather, spotting of photographic prints, and fading of dyes and pigments. Chemical methods of analysis are used to identify degradation products and to study the kinetics of degradation mechanisms. Specialists in air-pollution monitoring use analytical instrumentation to measure ambient pollution levels in museums. Some of the museums remove SO_2 and O_3 from the air by using filter systems containing activated carbon or aluminum oxide impregnated with potassium permanganate in their heating, ventilation, and air conditioning (HVAC) systems.

With a few highly important artifacts, inert atmospheres have been used. Thus, the Declaration of Independence, the Constitution, and the Bill of Rights are displayed, in the National Archives of the United States, in a helium atmosphere within a metal frame bonded to glass. To prevent contamination and oxidation, some Moon rocks are kept in inert atmospheres.

Conservation treatment. The most frequently employed techniques of the museum conservator involve chemicals and chemical reactions. Adhesives are used to mend broken pots, glass, metal, and so forth; benzotriazole to stabilize bronzes suffering from so-called bronze disease; bleaches such as hydrogen peroxide to remove stains and foxing from prints and drawings; diethyl zinc or methyl magnesium carbonate to deacidify paper; ethylene oxide and sulfuryl fluoride to fumigate infested textiles to arrest biodeterioration; and solvents to remove selectively discolored varnishes or overpainting from paintings on canvas. Friable materials such as waterlogged wood must be freeze-dried; or, if too large to fit into a vacuum chamber, they must be consolidated with resins or poly(ethylene glycol). Consolidants are also used with friable stone, leather, archeological ivory, and any other material that has been severely degraded. These processes are developed by collaboration of the scientist and the conservator to assure that the techniques are effective, safe, and, wherever possible, reversible.

Part of the evaluation process for conservation materials and techniques involves the examination of long-term stability. Two major approaches are

Fig. 6. Folding endurance tester used to evaluate the increase in strength of paper that is treated with the impregnating agent poly(vinyl alcohol).

used: light exposure and accelerated thermal aging. With the former, the natural or treated specimen is subjected to intense illumination for extended periods, commonly 2000 h. With thermal aging, the system under evaluation is heated for 1–50 days at temperatures of 122–212°F (50–100°C). In both cases, the aged samples are then studied by observing changes in physical properties [tensile strength, folding endurance (**Fig. 6**), color, molecular weight, compressive strength] and chemical properties [solubility, pH, change in composition]. Many materials which have ideal properties in commercial applications are unacceptable in conservation treatments, because they cannot be expected to remain stable for the extended periods required in museum practice. Much of the work of museum conservators involves removing materials used in the past which have discolored or become unstable or brittle. Mending tape on prints and drawings, shellac varnishes and pottery mends, and linseed oil consolidation are a few examples of poor materials used in attempts to repair works of art. Often such toxic solvents as dimethyl formamide and tetrahydrofuran must be used to undo the damage caused by materials. *See* ANALYTICAL CHEMISTRY. Norbert S. Baer

Bibliography. N. S. Bromelle and P. Smith (eds.), *Adhesives and Consolidants*, 1984; N. S. Bromelle and G. Thomson (eds.), *Science and Technology in the Service of Conservation*, 1982; R. L. Feller (ed.), *Artists' Pigments: A Handbook of Their History and Characteristics*, 1994; S. J. Fleming, *Authenticity in Art: The Scientific Detection of Forgery*, 1975; J. S. Mills, *The Organic Chemistry of Museum Objects*, 2d ed., 1999; G. Thompson, *The Museum Environment*, 1986.

Arteriosclerosis

The name given a group of degenerative diseases of arteries characterized by thickening and hardening of their walls. The group includes three types of lesions: (1) atherosclerosis involves the aorta and its major branches; (2) medial sclerosis involves the muscular arteries of the legs; and (3) arteriolosclerosis involves the small branches of the arterial tree, called the arterioles.

Atherosclerosis. Atherosclerosis is by far the most common and important form of arteriosclerosis, and the two terms are often used interchangeably. Some degree of atherosclerosis is found at autopsy in persons of all ages, even children, but it increases in severity and extent with advancing years. One study on 22-year-old American soldiers in the Korean War revealed the presence of atherosclerotic lesions in the coronary vessels in 77% of them. It is more common in males than in females, especially females before the age of menopause. After menopause the sex incidence and severity are about equal.

Incidence. Atherosclerosis has global distribution and occurs in virtually epidemic proportions in the industrialized nations of Western civilization. There is an extremely high prevalence and severity of the lesions in the Northern European countries, the United States, Australia, and New Zealand. There is a relatively low occurrence in economically deprived areas such as India, Africa, and South America. Factors associated with the high incidence and severity of atheromas (anatomic lesions of atherosclerosis) include a high total caloric intake, high fat intake, sedentary living, aggressive personality, emotional stress, and cigarette smoking. Hypertension (high blood pressure) does not induce arteriosclerosis, but it augments its development and accelerates the progress of the disease if it is present. The excessive incidence of myocardial infarction (heart attack) in cigarette smokers has been clearly documented. High blood lipid levels especially of cholesterol and triglycerides are also associated with higher incidence. The strong supposition thus arose that environmental and nutritional factors are of prime etiologic importance in the development of the disease, although the specific cause has not been identified. Although virtually ubiquitous, atherosclerosis fortunately does not cause serious disease symptoms in most individuals. *See* HEART DISORDERS; HYPERTENSION.

Pathology. Atheromas begin within the lining layer of the aorta (intima) or its branches and subsequently extend into the middle layer (media) of these vessels. The basic lesion has been shown to be a focal overgrowth of the smooth muscle cells of these layers. There is evidence that the initial lesion is a monoclonal overgrowth of smooth muscle cells, in some respects resembling a benign smooth muscle tumor such as a leiomyoma. This monoclonal overgrowth of smooth muscle cells, forming the initial atherosclerotic plaque, represents a mutation; its most likely cause, therefore, is a chemical mutagen from the environment or viruses. The cells subsequently degenerate, producing an accumulation of lipids in their cytoplasm, then necrosis, and finally calcification and scarification. As these lesions enlarge in the intima, the overlying endothelial cells may become disrupted, and fibrin clots are deposited on the surface. The lumen of the vessel is diminished in diameter (stenosis) by both the atheromatous plaque and the overlying clot. Subsequently, the fibrin clot may be incorporated into the fibrous scar in the intima. The occurrence of successive layers of clots and enlarging plaque produces narrowing of the lumen. When this process happens in large vessels such as the thoracic or abdominal aorta, it may not produce occlusion. However, if the process involves all layers of the arterial wall, it may weaken the wall, and an aneurysm may result. This may rupture, and the person may die from hemorrhage into the body cavity. *See* ANEURYSM.

When the atherosclerotic process occurs in the smaller branches of the aorta, a variety of consequences may ensue. Complete occlusion of the lumen by clot or hemorrhage into the plaque may occur. If the process takes place in the vessels supplying blood to the heart muscle (coronary arteries), the blood supply may be restricted or stopped completely, producing a myocardial infarction. If the

process involves the arteries supplying the brain, a similar occlusion may produce an infarction of the brain. Sometimes hemorrhage into the infarction may ensue. This is called a cerebrovascular accident (stroke). If arteries leading to the legs, arms, or internal organs are occluded, gangrene may result. Because of these sequelae, atherosclerosis assumes awesome importance as the major cause of death in the United States. *See* HEMORRHAGE.

Causes. Although the vast majority of cases of atherosclerosis appear to be principally caused by environmental factors, there are some specific genetic defects associated with the genesis of the process. Individuals with primary diabetes mellitus develop severe arteriosclerosis at an earlier age than nondiabetic individuals do. Familial hypercholesterolemia, homocystinuria, and hypothyroidism are other examples of metabolic defects associated with arteriosclerosis. *See* DIABETES; THYROID GLAND DISORDERS.

Medial sclerosis. Medial sclerosis is an uncommon type of arteriosclerotic lesion of little clinical significance. It affects the arteries of the arms, legs, and genital tracts of both sexes equally. It rarely occurs under the age of 50, and its cause is obscure. The disorder is characterized by ringlike calcifications within the media (middle layer) of affected vessels. The endothelial lining of the vessel is not altered, and the size of the lumen not impinged upon. Therefore, occlusion of the vessel seldom occurs.

Arteriolosclerosis. Arteriolosclerosis is an increased generalized thickening of the walls of arterioles related to hypertension. The change is often most prominent in the kidneys, although other internal organs may be similarly affected. Individuals with diabetes mellitus have an increased incidence of the lesion. *See* CIRCULATION DISORDERS.

N. Karle Mottet

Bibliography. J. D. Wilson et al. (eds.), *Harrison's Principles of Internal Medicine*, 14th ed., 1998; R. W. Wissler (ed.), *Arteriosclerosis Plagues: Advances in Imaging for Sequential Quantitative Evaluation*, 1992.

Artesian systems

Ground-water conditions formed by water-bearing rocks (aquifers) in which the water is confined above and below by impermeable beds. These systems are named after the province of Artois in France, where artesian wells were first observed.

Because the water in the intake area of an artesian system is higher than the top of the aquifer in its artesian portion, the water is under sufficient head to cause it to rise in a well above the top of the aquifer. Many of the systems have sufficient head to cause the water to overflow at the surface, at least where the land surface is relatively low. Flowing artesian wells were extremely important during the early days of the development of ground water from drilled wells, because there was no need for pumping. Their importance has diminished with the decline of head that has occurred in many artesian systems and with the development of efficient pumps and cheap power with which to operate the pumps. When they were first tapped, many artesian aquifers contained water that was under sufficient pressure to rise 100 ft (30 m) or more above the land surface. Besides furnishing water supplies, many of the wells were used to generate electric power. With the increasing development of the artesian aquifers through the drilling of additional wells, the head in most of them has decreased and it is now from a few feet to several hundred feet below the land surface in many areas of former artesian flow. A majority of artesian wells are now equipped with pumps. *See* AQUIFER.

Perhaps the best-known artesian aquifer in the United States is the Dakota sandstone, of Cretaceous age, which underlies most of North Dakota, South Dakota, and Nebraska, much of Kansas, and parts of Minnesota and Iowa at depths ranging from 0 to 2000 ft (0 to 600 m). The water is highly mineralized, as a general rule, but during the latter part of the nineteenth century, when these areas were being settled, the Dakota sandstone provided a valuable source of water supply under high pressure. Few wells in this aquifer flow more than a trickle of water today. The St. Peter sandstone and deeper-lying sandstones of early lower Paleozoic age, which underlie parts of Minnesota, Wisconsin, Iowa, Illinois, and Indiana, form another well-known artesian system. Formerly, wells on low ground flowed abundantly, but now wells have to be pumped throughout most of the area. Some of the water is highly mineralized, but in many places it is of good quality. In New Mexico, in the Roswell artesian basin, cavernous limestone of Permian age provides water to irrigate thousands of acres of cotton and other farm crops. Although the head has been steadily declining, many wells still have large flows and others yield copious supplies by pumping. Among the most productive artesian systems are the Cretaceous and Tertiary aquifers of the Atlantic and Gulf Coastal plains. These provide large quantities of water for irrigation and industrial use and supply large cities, such as Savannah, Georgia; Memphis, Tennessee; and Houston and San Antonio, Texas. Numerous artesian basins are found in intermontane valleys of the West. Some of the best known are in the Central Valley, California, where confined aquifers provide water to irrigate millions of acres of farmland, and the San Luis Valley, Colorado. Numerous other lesser artesian systems are found in all parts of the United States. *See* GROUND-WATER HYDROLOGY.

Albert N. Sayre; Ray K. Linsley

Arthritis

A group of diseases affecting joints or their component tissues. Several types of arthritis are recognized, and these can be divided into groups by their clinical course and pathologic appearance.

Types. There are four basic types of arthritis: inflammatory arthritis, degenerative joint disease, nonarticular rheumatism, and miscellaneous arthritis.

Inflammatory arthritis. This type is characterized by inflammation of tissues associated with joints. Connective tissue diseases, crystal deposition diseases, infectious arthritis, and spondyloarthropathies are examples of inflammatory arthritis.

Connective tissue diseases are a group of acute and chronic diseases characterized by involvement of joints, connective tissue, serosal membranes, and small blood vessels. These diseases are divided into acquired disorders (for example, rheumatoid arthritis, systemic lupus erythematosus, scleroderma, polymyositis, vasculitis) and rare hereditary diseases (for example, Ehlers-Danlos syndrome).

Rheumatoid arthritis is the most common variety of inflammatory arthritis. It occurs in younger and middle-aged persons and is characterized by noninfectious inflammation of the synovium (joint-lining membrane) frequently associated with extraarticular manifestations other than in the joints. The etiology is unknown, but genetic, immunologic, infectious, and psychologic disturbances have all been suggested. The systemic disease follows a variable but slowly progressive course, marked by spontaneous flares and remissions. Symmetrical involvement of the smaller joints of the hands and feet are often the initial clinical signs of disease, with wrists and other large joints affected later. Destruction of joints is common with advancing disease, resulting in a high incidence of disability.

Other connective tissue diseases, such as systemic lupus erythematosus, scleroderma, and vasculitis, frequently include joint symptoms. Other manifestations—involving soft-tissue structures such as kidneys, muscle, heart, and lungs—are the primary source of morbidity and mortality in these conditions. *See* CONNECTIVE TISSUE DISEASE.

There are three groups of crystal deposition disease classified according to type of crystal involvement: gout (monosodium urate), pseudogout (calcium pyrophosphate), and calcific tendonitis (hydroxyapatite). Gout, a disorder of uric acid breakdown, has long been associated with a violent arthritis, particularly of the first toe. Deposits of monosodium urate crystals in the affected joints and irregular episodes of symptoms are typical features. *See* GOUT; URIC ACID.

Infectious arthritis is an inflammatory joint disease caused by the invasion of the synovial joint by living microorganisms such as gonorrheal, streptococcal, and staphylococcal bacteria. Such arthritis usually results from a generalized infection but may appear following local spread or after trauma. The outcome varies considerably with the person's age and immune status, the adequacy of the treatment, and the nature of the organism. The most common type of infectious arthritis occurs in disseminated gonococcemia, but the gonococcal organisms are frequently difficult to isolate from the involved joints. Rheumatic fever, a sequel to streptococcal infection, tends to include a migratory arthritis that may involve several joints in succession; permanent damage to the joints is minimal. Tuberculous arthritis is less prevalent now than before the advent of adequate therapy for tuberculosis. *See* GONORRHEA; RHEUMATIC FEVER; STAPHYLOCOCCUS; STREPTOCOCCUS; TUBERCULOSIS.

The spondyloarthropathies are types of inflammatory arthritis characterized by involvement of the axial (central) skeleton (for example, the spine rather than the limbs). Ankylosing spondylitis (sacroiliitis) and Reiter's syndrome (clinical triad of arthritis, conjunctivitis, and urethritis) are examples of the spondyloarthropathies. All are hereditary to some extent through association with the histocompatibility system.

Degenerative joint disease (osteoarthritis). This is a ubiquitous joint disease characterized pathologically by deterioration of cartilage lining the joints and new bone formation beneath the cartilage. The disease is very common in older persons and is thought to be inherent in the aging process. In addition, various other factors are involved, including joint injury and the presence of other forms of arthritis. Degenerative joint disease is marked by a progressive stiffness, loss of function, and destruction of the larger, weight-bearing joints of the body. With advancing age, the continued slow damage causes increasing disability. *See* AGING.

Nonarticular rheumatism. This group of diseases, also called soft-tissue rheumatisms, includes tendonitis, bursitis, tenosynovitis, and fibrositis. Calcific tendonitis and nonseptic bursitis usually follow local trauma, muscle strain, or excessive exercise but can result from deposition of hydroxyapatite crystals (crystal deposition disease). Septic bursitis may also occur from local invasion by microorganisms. Fibrositis is a complex of musculoskeletal pains that frequently involves the neck, shoulder girdle, and extremities and is associated with multiple trigger points. The etiology is unclear, but the disorder may relate to psychobiologic or sleep disturbances or muscular and soft-tissue abnormalities. *See* BURSITIS; RHEUMATISM.

Miscellaneous arthritis. Systemic diseases of other or unknown etiology may produce arthritis or joint destruction. There are neurologic, blood, and endocrine examples of these unusual rheumatic diseases.

Treatment. Disability can often result from arthritis but can be curtailed by general health maintenance, rest, and rehabilitation. Occupational and physical therapies can be helpful. Diagnosis of the particular type of arthritis is extremely important in choosing drug therapy. Rheumatoid arthritis can be treated with nonsteroidal anti-inflammatory drugs such as aspirin. Disease-modifying antirheumatic drugs such as gold, antimalarial drugs, and immunosuppressive drugs are also frequently used. Osteoarthritis is treated with anti-inflammatory and analgesic drugs. Cortisone compounds are also used to treat arthritis, but the dosage and duration of treatment

must be carefully monitored because of significant side effects. Surgical treatment includes arthroscopic surgery and joint replacement.

Arthritis affects an estimated 11 million persons in the United States. The annual cost of treatment and care and the losses in time, money, and productivity are incalculable. With an increasing survival time in the population, arthritis constitutes one of the greatest medical, social, and economic problems. *See* AUTOIMMUNITY; JOINT DISORDERS. Robert Searles

Bibliography. R. G. Lahita (ed.), *Systemic Lupus Erythematosus*, 3d ed., 1998; D. J. McCarty, *Arthritis and Allied Conditions*: *A Textbook of Rheumatology*, 13th ed., 1996; A. Soren, *Arthritis and Related Affectations*: *Clinic, Pathology, and Treatment*, 1993.

Arthropoda

The largest phylum in the animal kingdom, including the well-known insects, spiders, ticks, and crustaceans, as well as many smaller, less well-known groups, and an abundance of strange forms known only as fossils. Arthropods make up about 95% of all species of animals. The estimated number of known species exceeds 1 million; of this number, most are insects. No one really knows the exact number of arthropod species. Some authorities estimate that it may range as high as 10 million.

Arthropods vary in size from the microscopic mites to the giant decapod crustaceans, such as the king crab with an appendage span of 5 ft (1.5 m) or more. The adult arthropod typically has a body composed of a series of ringlike segments as well as a pair of many-jointed limbs on each segment, movable on each other by means of muscles. However, some parasites, such as the pentastomids or the rhizocephalans, as adults have virtually no sign of segmentation. The integument of arthropods secretes a multilayered cuticle containing chitin. Often this cuticle is sclerotized by the deposit of hardening proteins. To accommodate growth, this exoskeleton must be shed periodically in a process called molting, or ecdysis. Young stages may be quite different from the adults, and some parasitic species differ very radically in body form from their nearest relatives. These characteristics, taken together, distinguish the arthropods from all other animals.

Classification

At one time, it was thought that the arthropods consisted of several phyla, that is, that they had a multiplicity of origins from different ancestors (polyphyly). However, recent investigations have successfully demonstrated that Arthropoda is a coherent group that arose from a single ancestor in the distant past (monophyly). Nevertheless, there are as many ways of classifying the arthropods as there are arthropod researchers. This is especially true when attempting to incorporate the wide array of extinct arthropods seen in the fossil record. The taxonomic system given here is generally accepted and places all the living groups in a logical scheme along with only a few of the fossil forms. The asterisk indicates extinct forms. The classification is modified from R. C. Brusca and G. Brusca, *The Invertebrates*, 2d ed., 2003.

Generalized Higher Classification of Phylum Arthropoda*

Subphylum Fuxianhuiida*
- Trilobitomorpha* (including Trilobita)
- Cheliceriformes

Subphylum Cheliceriformes
- Class Chelicerata
  - Subclass Merostomata
    - Arachnida
- Class Pycnogonida

Subphylum Myriapoda
- Class Diplopoda
  - Chilopoda
  - Pauropoda
  - Symphyla

Subphylum Hexapoda
- Class Entognatha (including Parainsecta)
  - Insecta

Subphylum Crustacea
- Class Remipedia
  - Cephalocarida
  - Branchiopoda
  - Malacostraca
  - Subclass Phyllocarida (Leptostraca)
    - Hoplocarida
    - Eumalacostraca
- Class Maxillopoda
  - Subclass Thecostraca (including Cirripedia)
    - Tantulocarida
    - Ostracoda
    - Copepoda
    - Branchiura (including Pentastomida)
    - Mystacocarida

The phylogenetic relationships between these subphyla are still not certain. The hexapods are believed by many to be closely related to the crustaceans, but as to which group of crustaceans is still strongly argued. However, depending on whether or not one includes fossils in the analysis, hexapods can emerge as more closely related to the myriapods (a traditional affiliation). Conversely, recent molecular evidence seems to indicate that the myriapods by themselves might be closely allied with the cheliceriforms, but strong arguments also ally them with other mandibulate groups such as the hexapods and crustaceans. Trilobitomorphs, long thought to bear some relationship to the cheliceriforms, recently have been suggested to have common origins with mandibulates. The phylogenetic relationships within each of these subphyla are only beginning to be settled. A massive, worldwide effort (called "Assembling the Tree of Life") is under way to utilize information from a variety of sources [comparative anatomy,

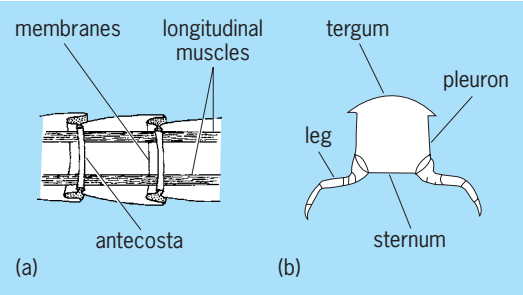

Fig. 1. Arthropod segmentation. (*a*) Segmental plates. (*b*) Cross section of a body segment. (*After R. E. Snodgrass, Principles of Insect Morphology, McGraw-Hill, 1935*)

embryology, developmental genetics, molecular deoxyribonucleic acid (DNA) sequences] to elucidate kinship connections.

Metamerism

Body segmentation, or metamerism, is one of the most fundamental characteristics of the arthropods (**Fig. 1***a*). However, it is a feature shared with several other phyla, such as the annelid worms, kinorhynch worms, and even chordates. The link with annelid worms has been formalized into a superphylum group called Articulata. In this respect, metamerism would be viewed as the key character. However, as striking as this feature is, other aspects of the biology suggest that the link between annelids and arthropods is not so close. The molecular sequencing of deoxyribonucleic acid (DNA) has advanced an alternative view in which arthropods are united into a superphylum called Ecdysozoa, which also includes kinorhynchs, priapulids, nematodes, onychophorans, and tardigrades. In this respect, molting would be viewed as the key character. The ecdysozoans are contrasted to another superphylum, called Lophotrochozoa, in which the annelids, mollusks, and some other phyla are united.

The arthropod cuticle may be sclerotized continuously around the segments. More typically, however, it forms discrete segmental plates, or sclerites. A dorsal plate of a segment is called a tergum or notum; a ventral plate is a sternum; lateral plates are pleura (Fig. 1*b*). The consecutive tergal and sternal plates, unless secondarily united or fused, are connected by infolded arthrodial membranes, and are thus movable on each other by longitudinal muscles attached on anterior and posterior marginal ridges on the underside of the plates. Since nearly all the body and limb muscles are attached to either integumental sclerites or to skeletal elements infolded from the surface sclerites called endophragms, there is effectively no limit to the development of skeletomuscular mechanisms. Furthermore, the body segments, associated with specialized sets of appendages, tend to become consolidated or united into segment groups, or tagmata, forming differentiated body regions, such as the head, thorax, and abdomen.

The segmentally arranged limbs of all modern arthropods develop in the embryo from small lateroventral outgrowths of the body segments that lengthen and become jointed. One phylum, the onychophorans, appears to present a model for an ancestor to arthropods, whereby rudimentary limbs could have given rise to more elaborate appendages. Whether this ancestor was a crawling or swimming form is still widely debated. The fossil record provides examples of both walking and swimming forms that could serve equally well as models for an arthropod ancestor. *See* ONYCHOPHORA; ANNELIDA.

The first segmental limbs may have been all leglike appendages (as they were in the trilobites), which served for walking across the bottom of the sea. On the other hand, some authorities have suggested that the earliest limbs were flagellar (very long with many small segments) in form with the ancestral scuttling along the bottom or swimming. In their later evolution, however, arthropod legs became modified in structure for many other purposes, such as feeding, grasping, swimming, respiration, silk spinning, egg laying, and sperm transfer. Thus, modern arthropods possess legs that are highly modified and more specialized than those seen in any other group of animals. *See* METAMERES.

Development

A compelling set of arguments for the unity of Arthropoda arises from the relatively new field of evolutionary developmental genetics, or "evo-devo." The exploration of how genes control various aspects of development in both animals and plants has yielded new information relative to assessing the origin of anatomical structures and systems. For example, a series of maternal genes set the polarity of the animal egg even before development begins. Other genes in the zygote quickly specify numbers and polarity of individual segments by actually segregating distinct anterior and posterior halves of segments in arthropods. A critical gene in this regard, known as *engrailed*, specifies the anterior portions of what in the embryo are called the parasegments, the first manifestations during ontogeny of any metamerism. As development progresses, the two halves of the parasegment uncouple from each other and associate with the tissues of the adjacent parasegment. As a result, the *engrailed* gene comes to be expressed in the posterior portion of the adult segment, a pattern seen in all arthropods.

The *Hox* genes, another family of genes responsible for setting the anterior-posterior polarity in the animal body, have become specialized in the arthropods into a family of genes that also help specify the identity of segments within body regions, or tagmata. The genes *labial*, *proboscipedia*, *Deformed*, and *Sex combs reduced* play a role in determining the identity of segments in the arthropod head (**Fig. 2**). The genes *Antennapedia*, *Ultrabithorax*, and *abdominal A* are expressed in various ways in the thorax. *Abdominal B* and *caudal* are genes that are expressed in the tail end of the body, for example, in the pleon of a crayfish (Fig. 2). The *Hox* genes are located on

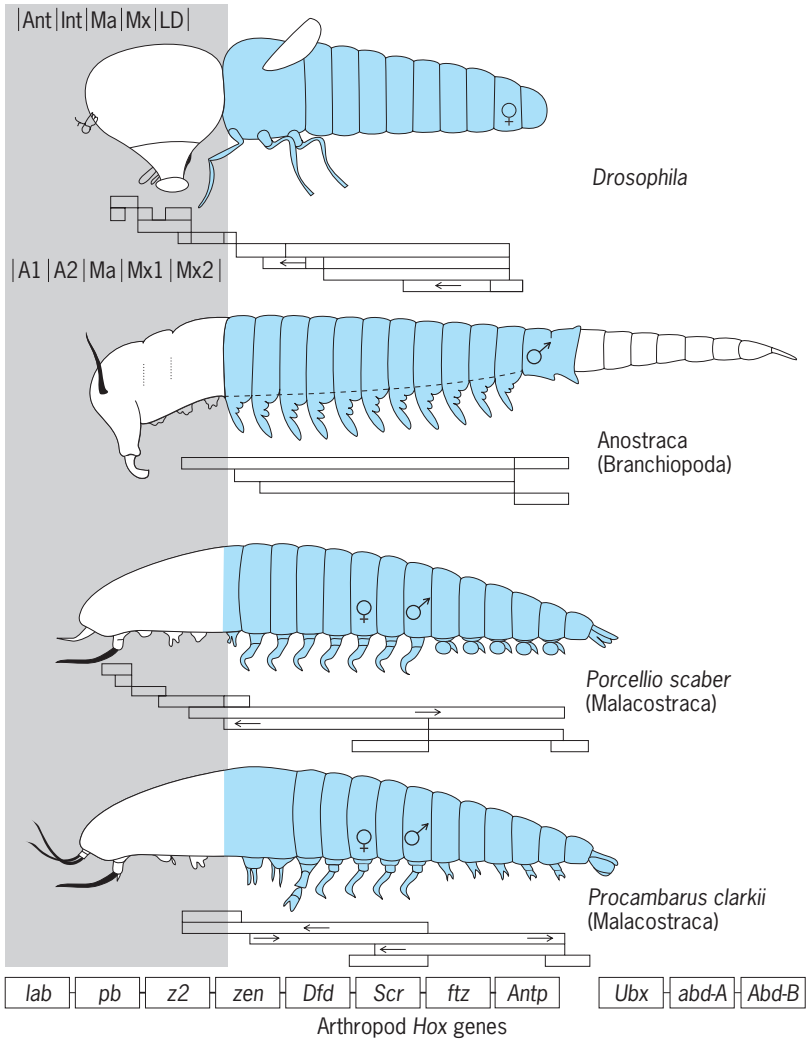

Fig. 2. Areas of expression of the various *Hox* genes in selected mandibulate arthropods. *lab* = labial; *pb* = proboscipedea; *z2* = zerknuellt; *zen* = zen; *Dfd* = deformed; *Scr* = sex combs reduced; *ftz* = fushi terazu; *Antp* = Antennapedia; *Ubx* = Ultra biothorax; *abd-A* = abdominal-A; *AbdB* = abdominal-B.

the chromosome in the order in which they are read and expressed from anterior to posterior. Variations occur, but the underlying pattern is identical across all arthropod groups.

In addition, the genes that control specification of the limbs are employed in a distinctive way characteristic of all the arthropods examined so far. Two systems operate. The basalmost parts of the limb, the coxopodites, are controlled by a gene called *extradenticle*. This system is quite separate from the complex array of genetic interactions that specifies various aspects of the more distal portions of the arthropod limb. Here, the *engrailed* gene triggers a gene series called the *hedgehog* cascade, wherein the genes engaged in a specific sequence specify various quadrants of the developing limb bud, culminating in the tip of the appendage where a gene called *distalless* is expressed.

No matter how different the body plan of an adult arthropod may be, the same genes are present in all arthropods and are expressed in the same areas of the body. While most of these genes are also found in a wide array of animal phyla, it is the distinctive combination outlined above that is unique to arthropods. In other words, all arthropods share a common pattern of developmental control and body region specification. *See* DEVELOPMENTAL GENETICS.

Internal Organs and Physiology

The arthropods, from lobsters to the tiniest gnat, have all the internal organs essential to any complex animal (**Fig. 3**).

Gut. A gut extends either straight or coiled from the anterior, ventrally directed mouth to the terminal anus. There are three regions in the gut. The anterior stomodeal region (foregut) is formed by an ingrowth of the outer germ layer of the embryo, the ectoderm. This stomodeum is often elaborately developed with teeth and plates of hardened chitin, and setose sieves and filters. The middle part of the gut, or mesenteron, is formed from the endoderm and contains the digestive diverticula as well as a region for absorption. This middle region can be quite short in some arthropods. The posterior gut region is the proctodeum (hindgut), also formed from ingrowths of ectoderm.

Nervous system. The nervous system includes a brain and a subesophageal ganglion in the head, united by connectives around the stomodeum, and the system generally displays paired ventral nerve cords with segmental, interconnected ganglia. Some of the successive ganglia, however, may be condensed into composite ganglionic masses. Peripherally directed nerves proceed from the ganglia. Internal proprioceptors and surface sense organs of numerous kinds (chiefly tactile, olfactory, and optic) are present. The sensory range can be quite remarkable: Insects can see into the ultraviolet range; mantis shrimp have color perception that far exceeds the abilities of humans; and spiders and scorpions "see" by vibratory excitement of minute hairs on their legs.

Heart. The dorsal heart is often tubular, but it also can be short and compact. Arteries can distribute blood in a number of different ways. The heart can pump either directly through arteries to the various tissues and organs (a more or less closed system) or from the anterior end of the tube directly into the body cavity (an open system). Conversely, blood can reenter the heart through openings (valves) along its sides from the body cavity, or return to the heart by means of veins from the body tissues.

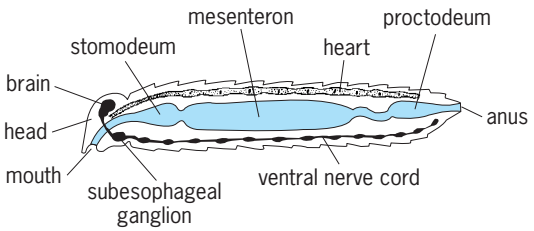

Fig. 3. Diagram of position of principal internal organs in a typical arthropod.

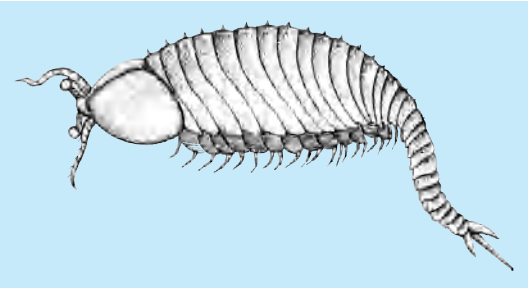

Fig. 4. The Cambrian proto-arthropod, *Fuxianhuia protensa*, from China.\vskip8pt

Respiration and excretion. Aquatic arthropods breathe by means of gills. Most terrestrial species have either flat air pouches, or tubular tracheae opening from the outside surface; some have both. A few small, soft-bodied forms respire directly through the thin cuticle. Excretory organs open either at the bases of some of the appendages or into the midgut.

Reproduction. Most arthropods have separate sexes, but some are hermaphrodites. Parthenogenesis is known but can be said to be common only in certain groups of insects. The genital openings differ in position in different groups and are a key component in defining the distinct body plans found in the phylum. The gonopores usually (but not always) are on the same body segment in the two sexes.

Fuxianhuiida

This is an extinct group of arthropods known from the Cambrian more than 550 million years ago. The currently recognized forms had rather large and long bodies (**Fig. 4**). However, the most interesting body part is the head. It appears to have only two segments: one that contained the eyes and one with an antennalike limb between which was the mouth. What is perhaps significant is that in modern arthropods these same two segments form the anteriormost part of the head and are the only segments that do not have any *Hox* gene expression. The mouth was anteroventral but had no special limbs to assist in capturing or rendering prey, so it is possible the fuxianhuiids merely plowed through bottom sediments and ate the mud.

Trilobitomorpha

The trilobitomorphs are extinct and are among the oldest known arthropods. Trilobita (or trilobites) were the dominant group, yet they are not the actual ancestors of any modern arthropods. *See* TRILOBITA.

A typical trilobite has a flattened oval shape with an elevation, the rachis, running lengthwise along the middle of the back, giving it a three-part appearance from which the trilobites get their name (**Fig. 5**). More important, however, is the crosswise division of the body into a head, a thorax, and a pygidium. The dorsal aspect of the head is not overtly segmented, but indentations of the median elevation suggest a primitive metamerism. On its upper surface, laterally, it bears a pair of compound eyes. The antennae extend forward on the undersurface, and a posteriorly directed mouth is located just posterior to a ventral sternite between the antennae called the hypostome. The posterior half of the ventral head bears four pairs of limbs. The thorax is completely segmented, and each metamere bears a pair of limbs similar to those on the head. The minimal number of thorax segments is eight, but there are often many more than that. The pygidium is a rounded plate at the posterior end of the body showing evidence of former segmentation.

The trilobite ancestors were probably fully segmented primitive arthropods. All of the postoral head and trunk limbs were alike and consisted of a long series of seven-segmented legs. The similarity of these limbs suggests that the diversified appendages of other arthropods were once jointed legs used for locomotion. Furthermore, each trilobite leg (**Fig. 6**) has a fringed or setose branch, perhaps a gill, arising laterally from the basal leg segment. Many trilobites have medially directed, spiny gnathobases on the basalmost limb segment, and these structures probably served for grasping food and passing it forward to the mouth, similar to what occurs in the living horseshoe crabs. The position of the reproductive

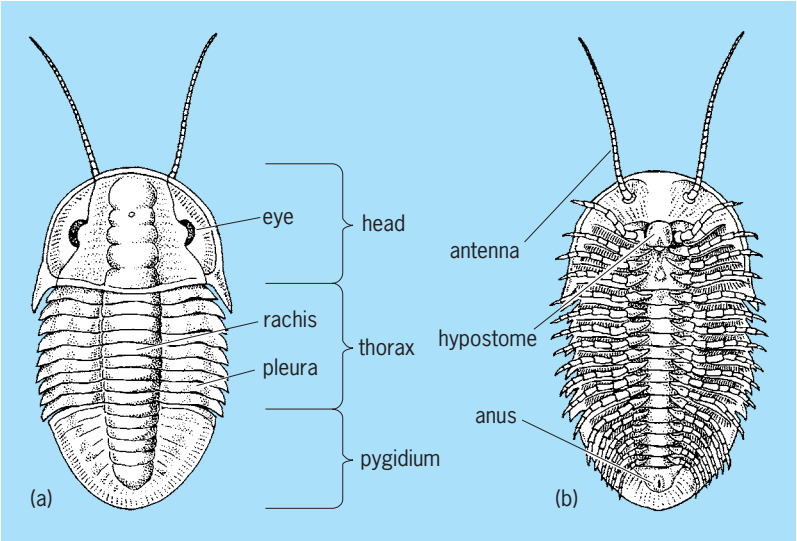

Fig. 5. Generalized trilobite. (*a*) Dorsal aspect. (*b*) Ventral aspect. (*After R. E. Snodgrass, Textbook of Arthropod Anatomy, Cornell University Press, 1952*)

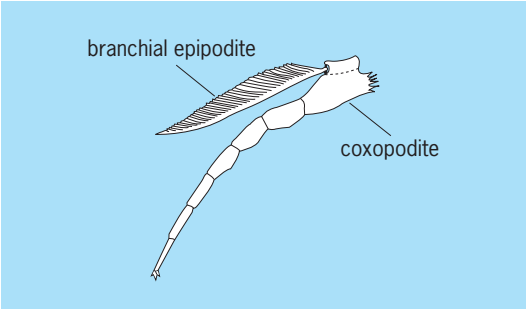

Fig. 6. Generalized trilobite leg.

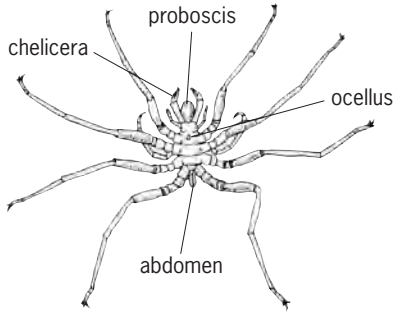

Fig. 7. ***Nymphon hirtipes*, a pycnogonid, an inhabitant of ocean shores. (*After R. E. Snodgrass, A Textbook of Arthropod Anatomy, Cornell University Press, 1952*)**

openings is not known. The anus is at the end of the pygidium.

Cheliceriformes

The cheliceriformes get their name from the fact that their first appendages are a pair of small pincers, known as chelicerae. They have no antennae, but it is now clear from developmental genetic studies of *Hox* genes that the chelicerae are on the same segment that bears the antennae of other arthropods. Cheliceriforms can be viewed as arthropods that had a distinctive *Antennapedia* mutation early in their history, leading to the formation of chelicerae rather than antennae.

Two classes are recognized: Pycnogonida and Chelicerata.

Pycnogonida. The pycnogonids are generally small, long-legged creatures that live mostly along the shores of the ocean, or in the deep sea. Their slender legs give them the superficial appearance of spiders. They differ, however, in that the abdomen in living forms is an insignificant lobe projecting from between the posteriormost legs (**Fig. 7**), and the mouth is situated on a large, anteriorly directed proboscis. Pycnogonids feed particularly on the tentacles of sea anemones, which they reduce to a pulp and swallow in liquid form. Although most pycnogonids have eight walking legs, some deepwater forms in Antarctica have 10 and 12 legs. Pycnogonids have a set of gonopores at the base of each walking limb. *See* PYCNOGONIDA.

Chelicerata. The body of the chelicerata is divided into two parts: a prosoma (or cephalothorax) and an opisthosoma (or abdomen). The prosoma carries the chelicerae, a second pair of appendages (termed pedipalps), and usually four pairs of walking legs. Appendages are usually absent on the opisthosoma, but when present they are never leglike. The genital aperture in each sex is always on the eighth postoral segment. The chelicerates include the subclasses Merostomata (which includes the orders Xiphosurida and Eurypterida) and Arachnida. *See* CHELICERATA.

Merostomata. The subclass Merostomata includes the order Xiphosurida (horseshoe crabs), which are not crabs at all (**Fig. 8**), and Eurypterida (extinct sea scorpions). As in the trilobites, the basal segments are very spiny and directed medially. These spiny lobes work against each other to "chew up" food before it is passed forward toward the mouth. The xiphosurids feed on solid food. All the other modern chelicerates are liquid feeders.

The eurypterids are extinct aquatic merostomes that lived in the Paleozoic Era. A typical eurypterid has an elongate body divided into an unsegmented prosoma or cephalothorax carrying six pair of appendages (**Fig. 9**) and a segmented opisthosoma or abdomen without appendages. Among the prosomal appendages, the first pair are the chelicerae, while the second pair are the homologs of the pedipalps of the arachnids. In some species, the opisthosoma ends in a tail bearing a large terminal spine and so resemble scorpions, such that they are

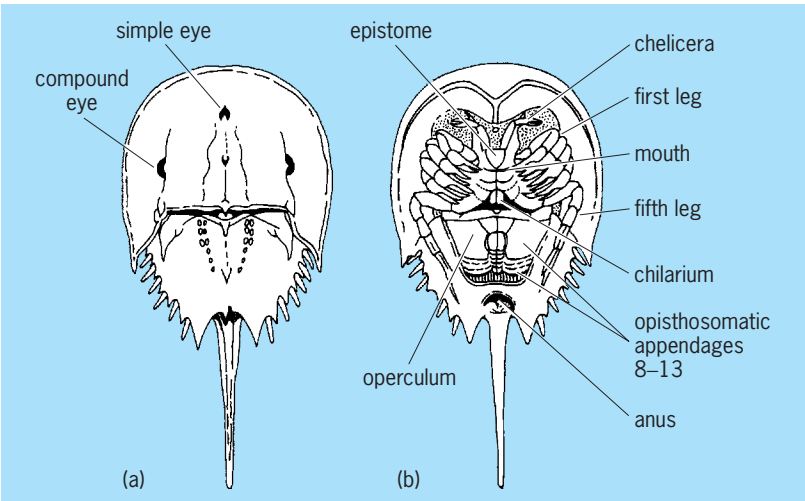

Fig. 8. Horseshoe crab (*Limulus*). (*a*) Dorsal aspect. (*b*) Ventral aspect. (*After R. E. Snodgrass. A Textbook of Arthropod Anatomy, Cornell University Press, 1952*)

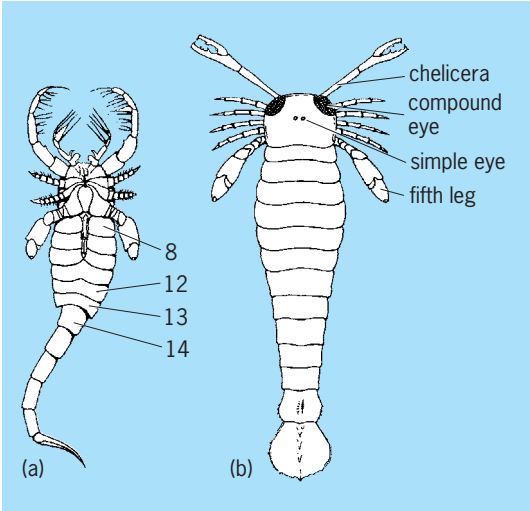

Fig. 9. Eurypterids. (*a*) *Mixopterus kiaeri*. (*b*) *Pterygotus buffaloensis*. (*After R. E. Snodgrass, A Textbook of Arthropod Anatomy, Cornell University Press, 1952*)

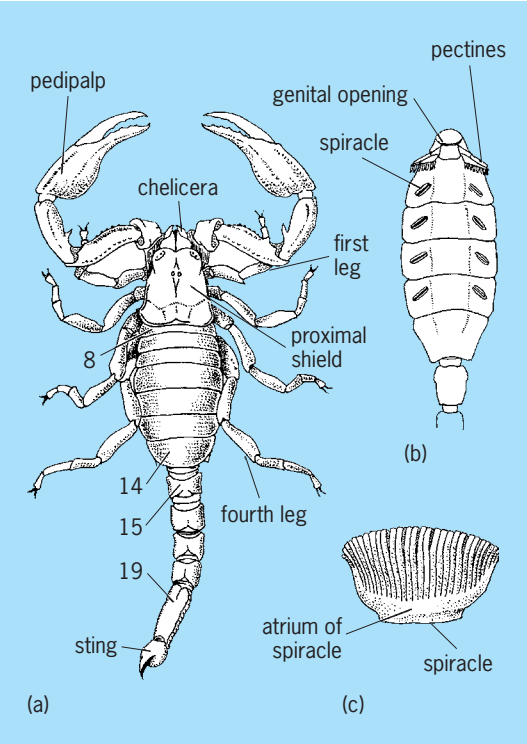

Fig. 10. Scorpion (*Chactas vanbenedeni*). (*a*) Entire animal, dorsal aspect (*after R. E. Snodgrass, A Textbook of Arthropod Anatomy, Cornell University Press, 1952*). (*b*) Anterior abdomen, ventral aspect. (*c*) Book lung, diagrammatic.

presumed to be the ancestors of modern scorpions (Fig. 9*a*). Although most of the eurypterids were 6–15 in. (15–38 cm) long, some attained a length of 6–8 ft (1.8–2.4 m). A scorpion of such a size would indeed be a formidable animal. *See* EURYPTERIDA; MEROSTOMATA; XIPHOSURIDA.

Arachnida. The familiar subclass Arachnida includes the modern scorpions (**Fig. 10**) as well as the spiders (**Fig. 11**), daddy longlegs, ticks (**Fig. 12**), mites, and others, besides a large number of fossil species. The first leg posterior to the chelicerae is modified as a grasping appendage, or pedipalp. The prey is pierced (or torn open by the chelicerae), digestive juices are discharged into it from the stomach, and the predigested tissue is then ingested by the sucking pharynx. These corrosive digestive secretions are why the bites of spiders are often so painful and difficult to heal.

Arachnida includes the orders Palpigradi, Schizomida, Scorpiones (scorpions), Uropygi (tailed whip scorpions), Amblypygi (tailless whip scorpions), Araneae (spiders), Solfugae (sun spiders), Pseudoscorpionida (pseudoscorpions), Opiliones (harvestmen), Ricinulei, and Acari (ticks and mites). *See* ARACHNIDA.

Mandibulata

The mandibulate arthropods include the myriapods (centipedes and millipedes), hexapods (insects), and crustaceans. There is great diversity within the group, but they may be readily distinguished from the chelicerates by the presence of antennae and the possession of mandibles, which are typically a pair of biting and chewing jaws, but these limbs may be drawn out into cutting blades or piercing stylets. The antennae arise from the front of the head and are never leglike in form. The crustaceans have a pair of second antennae that arise behind the mouth and belong to the series of segmental limbs. In place of a second pair of antennae, the myriapods and hexapods lack a distinct appendage on that segment, but appear to modify limb rudiments as an upper lip, or labrum. The labrum forms an anterior lip or shield over the mouth. The mandibles posterior to the second antennae are the second postoral appendages. These have evolved from a pair of primitive legs by the development of teeth or masticatory surfaces on the basal segments and the complete elimination, or reduction to a small palp, of the rest of the limb. Posterior to the mandibles are two accessory feeding

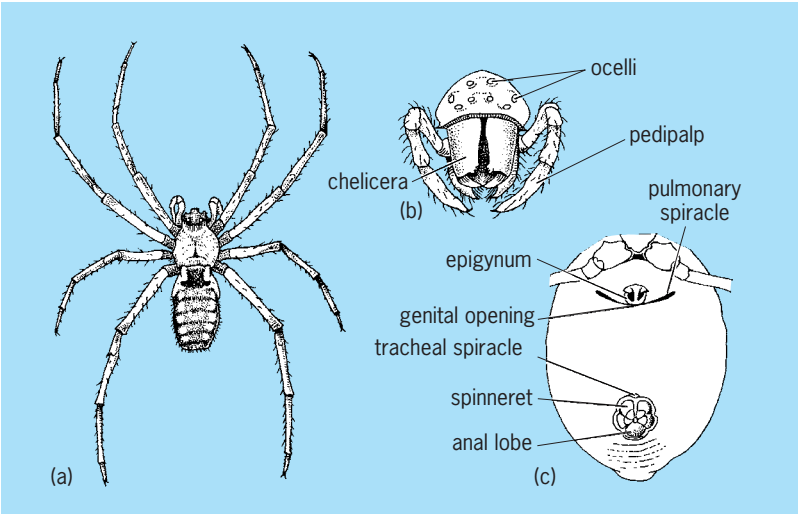

Fig. 11. Orb-weaving spider. (*a*) Entire animal, dorsal aspect. (*b*) Anterior view of head. (*c*) Undersurface of abdomen.

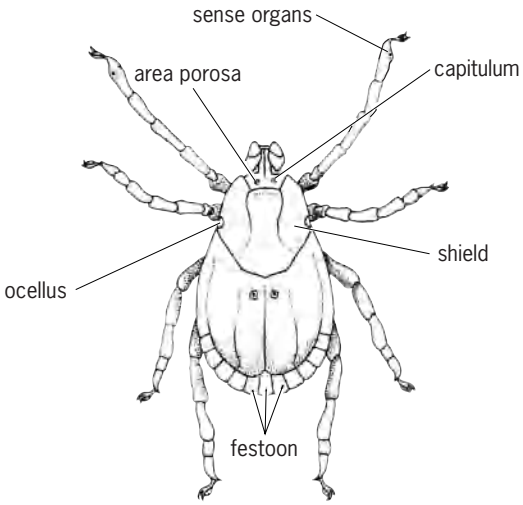

Fig. 12. Tick (*Dermacentor*). (*After R. E. Snodgrass, A Textbook of Arthropod Anatomy, Cornell University Press, 1952*)

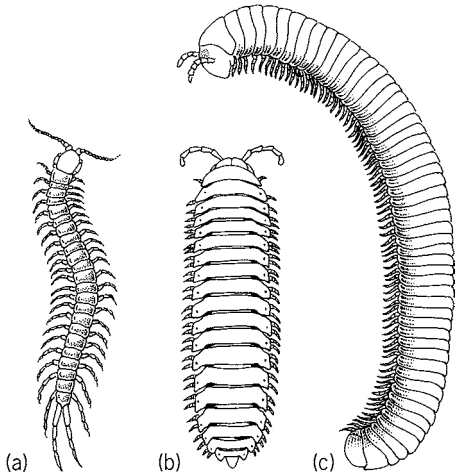

Fig. 13. Chilopoda and Diplopoda. (*a*) Chilopod, or centipede (*Scolopocryptops*; formerly *Otocryptops*). (*b*) Polydesmoid diplopod. (*c*) Julid diplopod. (*After R. E. Snodgrass, A Textbook Anatomy, Cornell University Press, 1952*)

limbs of various forms known as the first and second maxillae. The trunk limbs, which vary in number in the several mandibulate groups, are used mostly for walking or swimming, but some are commonly specialized for other purposes. The trunk may consist of only a head and body, or it may be divided into head, thorax, and "tail region." This tail can be either a limbless abdomen or a limb bearing pleon. *See* MANDIBULATA.

Myriapoda. The myriapods include the common centipedes and millipedes, and some other groups of a similar nature. There are four rather distinct living classes: Chilopoda, Diplopoda, Pauropoda, and Symphyla. Most of them develop after hatching by anamorphosis, in which the young hatch with only five to seven body segments and three pairs of legs. Further development takes place in a subterminal zone of growth, in which new segments and legs are generated and added to those already present. Two groups of the chilopods, however, develop after hatching by epimorphosis. Their development (in terms of segment number) is completed within the egg; they increase only in size after hatching. The adult myriapod consists of only two parts: a head and a segmented body.

Chilopoda. The chilopods are often called centipedes (**Fig. 13***a*). Though they include four groups of species, these groups differ mainly in the number of body segments, which varies from 15 to 100 or more. In some species, the number of tergal plates is less than the number of segments. Except for the house centipede that prefers human dwellings, they all live in similar places in the ground, under rocks or logs, or beneath loose bark. The appendages of the first body segment are developed as poison fangs that are directed forward beneath the head; a venom gland opens on the tip of each fang. The other appendages, except those of the reduced last two segments, are legs. Centipedes move quickly, as one would expect of a rapacious obligate carnivore.

Eyes, when present, are in lateral clusters that may look like compound eyes but do not have a compound-eye structure. They detect prey not so much by sight, however, as by tactile detection of vibrations. The reproductive openings are at the posterior end of the body. *See* CHILOPODA.

Diplopoda. The diplopods, or millipedes, differ from all other arthropods in that each body segment, except for the first three or five, bears two pairs of legs (Fig. 13*b* and *c*)—thus their name, Diplopoda. The body is either flattened somewhat or cylindrical and is ideally suited to bulldoze through leaf litter and decaying woody material that they consume for food. The mouthparts consist of a pair of mandibles with a large flat lobe, known as the gnathochilarium, forming a lower lip. They respire by means of tracheae arising from small ventral plates that carry the legs. The reproductive openings of each sex are for the most part on the third body segment. However, in the males of the common diplopods, one or both pairs of legs on the seventh segment (called the gonopods) are modified to receive the sperm from the genital outlets on the third segment and transfer it to the sperm receptacles on the third segment of the female. One group of diplopods have the gonopods at the rear end of the body, but these posterior gonopods are used for holding the female, whereas the male uses his mandibles to transfer the sperm. *See* DIPLOPODA.

Pauropoda. These are minute creatures, having generally not more than 12 body segments and 9 pairs of legs (**Fig. 14**). They are distinguished by their branched antennae. The mouthparts consist of a pair of simple undivided mandibles and a gnathochilarium resembling that of the diplopods. The genital ducts open on the ventral plate of the third body segment. *See* PAUROPODA.

Symphyla. The symphylans look like small, white, soft-bodied centipedes with long antennae, 12 pairs of legs, and two tapering unjointed appendages at the rear end of the body (**Fig. 15**). A characteristic feature of the symphylans is the presence of

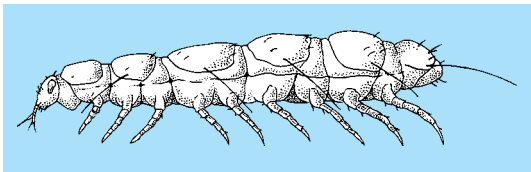

Fig. 14. Pauropod (*Pauropus silvaticus*). (*After R. E. Snodgrass, A Textbook of Arthropod Anatomy, Cornell University Press, 1952*)

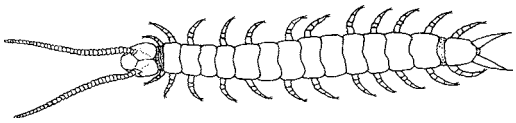

Fig. 15. Symphylan, *Scutigerella immaculata*, with long antennae, 12 pairs of legs, and two unjointed appendages. (*After R. E. Snodgrass, A Textbook of Arthropod Anatomy, Cornell University Press, 1952*)

small styli on the undersurface of the body medial to the legs. They live mostly in damp places under stones, in rotting logs, or in the ground, but one species, known as the garden centipede, sometimes becomes a destructive pest of garden crops and greenhouse plants. The mandibles resemble those of the diplopods in being composed of two parts, of which the distal ones are the functional jaws. The head, however, bears two pairs of maxillae, and those of the second pair, as in the insects, are united in a labiumlike organ. The reproductive openings are ventral on the fourth body segments. *See* SYMPHYLA.

Hexapoda. Insects and their allies may or may not have wings, but are easily recognized by the division of the body into head, thorax, and abdomen and by the presence of only three pairs of legs carried by the thorax (**Fig. 16***a*). The head always includes the segments of the mandibles and two pairs of maxillae. Most adult insects have both compound and simple eyes. There is only one pair of antennae. The second maxillae are united in a single organ, the labium. The mouthparts undergo endless modifications in different orders by which they are adapted to biting, chewing, piercing, and sucking, according to the nature of the food. Respiration is by means of tracheae. The reproductive openings in each sex are near the end of the abdomen.

Entognatha. This class of hexapods lacks wings and is characterized by having the jaws partially enclosed within the head capsule. There are three groups; none of them is large, either in species numbers or in body size (Fig. 16*b*). The entognaths are generally only a few millimeters in size and among the cryptofauna that inhabits leaf litter and rooting vegetation.

Insecta. There are both wingless and winged forms in this class. The wingless insects include those such as silverfish and bristletails (Fig. 16*c*). The winged insects, or Pterygota, include some groups that have lost their wings. The pterygotes not only make up the vast majority of the Insecta but also represent the dominant form of animal life in terms of existing species, with probably 900,000 described forms (as a conservative estimate). The reason for their abundance and diversity rests on their ability to disperse by flight, small size, very specialized feeding habits and ecological requirements, and the niche segregation of larvae from adults. *See* INSECTA.

Crustacea. The familiar crustaceans are the aquatic shrimps, crayfish, lobsters, crabs, and the terrestrial sow bugs. The oceans and fresh waters, however, swarm with lesser members of the clan, which constitute the basic food for most of the larger animals. *See* CRUSTACEA.

While crustaceans share many characters with other arthropod groups, it is the crustacean head (cephalon) that is distinctive. The cephalon consists of five, usually fused segments. Each cephalic somite has a pair of appendages, which include the antennules and antennae (also referred to as first and second antennae), mandibles, maxillules, and maxillae (the last two also referred to as first and second

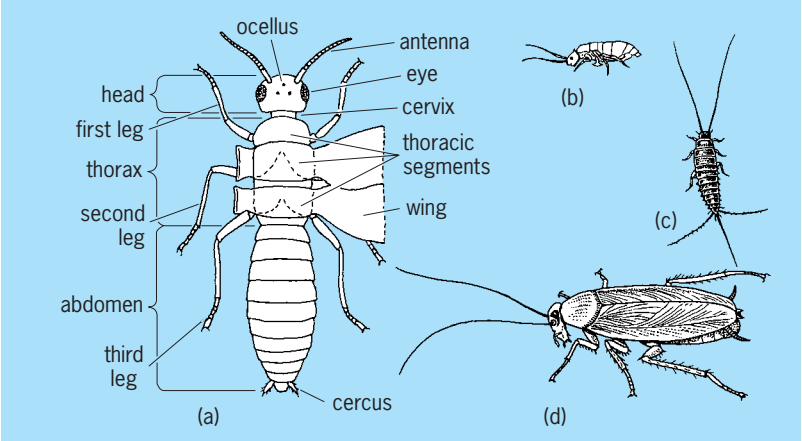

Fig. 16. Insecta. (*a*) Diagram of insect structure (*after R. E. Snodgrass, Principles of Insect Morphology, McGraw-Hill, 1935*). (*b*) Collembolan. (*c*) Thysanuran. (*d*) Pterygote, cockroach, *Periplaneta* (*parts c and d after R. E. Snodgrass, A Textbook of Arthropod Anatomy, Cornell University Press, 1952*).

maxillae). It is the first antenna that is the homologue of the antennae of other arthropods. One problem with studying the phylogenetic relationships of crustaceans is that while the body is typically divided into three distinct regions—cephalon, thorax (trunk), and abdomen (pleon)—the gonopores can be located in a number of places. In other groups of arthropods, the location of the gonopores is fixed for each group, and this variability in crustaceans suggests that the subphylum Crustacea is actually composed of several disparate groups. The crustaceans can be subdivided into five classes, though there is much doubt whether all of these are true monophyletic groups, that is, with a single shared origin: Cephalocarida, Remipedia, Branchiopoda, Maxillopoda, and Malacostraca.

Cephalocarida and Remipedia. The cephalocarids are a group of very small (2.0–3.6 mm) bottom-dwelling crustaceans, which at one time were regarded as the most primitive of crustaceans. The structure and function of the serially homologous appendages was the focal point of the cephalocarid theory, which postulated that evolution in the Crustacea came about through specialization, simplification, and reduction of cephalocarid-like appendages (**Fig. 17***a*). However, the 1981 discovery by cave divers of the Remipedia (Fig. 17*b*) has challenged this hypothesis. The lack of trunk regionalization in remipedes and their simple, biramous appendages have caused some carcinologists to propose that the Remipedia more correctly reflect the crustacean ancestral type. *See* CEPHALOCARIDA; REMIPEDIA.

Branchiopoda. The branchiopods for many years have been divided into four major groups. Although the Anostraca (fairy shrimps; **Fig. 18***a*) and Notostraca (tadpole shrimps; Fig. 18*b*) have been retained as monophyletic taxa, the names Conchostraca (clam shrimps) and Cladocera (water fleas) have been replaced. The former has been split into the orders Laevicaudata and Spinicaudata, which both have bivalved carapaces. The cladocerans, while still a monophyletic group, are subdivided into

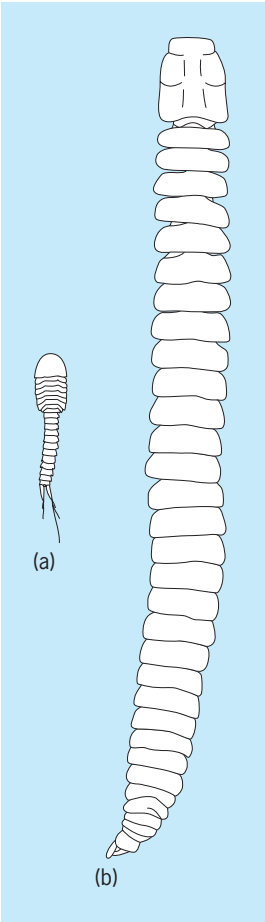

Fig. 17. Two primitive living crustaceans suggested as representing something close to an ancestral type. (*a*) Cephalocarid. (*b*) Remipede.

the orders Ctenopoda, Anomopoda, Onychopoda, and Haplopoda (Fig. 18*c–f*). *See* BRANCHIOPODA.

Maxillopoda. The class Maxillopoda has not received universal acceptance among carcinologists. One of the principal threads of continuity seen throughout the group is the ways in which the maxillae and anterior thoracic appendages are used in food gathering and locomotion. However, there are probably at least two quite separate groups included in the maxillopodans. One group, the largest by far, generally conform to a body plan with five segments in the head, seven segments in the thorax, and four segments in the abdomen. This would include the copepods, thecostracans, and perhaps some of the ostracodes. Maxillopoda contains the subclasses Thecostraca, Tantulocarida, Ostracoda, Copepoda, Branchiura, and Mystacocarida. *See* MAXILLOPODA.

Copepoda (**Fig. 19**) include some 8000 to 10,000 species; they are by far the largest taxon assigned to the Maxillopoda. Copepods are found in virtually all aquatic habitats and a few moist terrestrial habitats, but a large number are parasitic and infest a vast variety of vertebrate and invertebrate hosts. *See* COPEPODA.

The second largest maxillopodan subclass is Thecostraca (**Fig. 20**), which includes the stalked and sessile barnacles (order Thoracica; Fig. 20*b–d*). The cirripedes also include two parasitic orders, Rhizocephala and Ascothoracica (collectively known as the cirripedes). The rhizocephalans are noted for their "parasitic castration" of crabs, but they have been found to infest a variety of crustacean hosts. As adults, they completely lack any similarity to other crustaceans, being merely a bag of gonads attached to a root system that permeates the host body. The ascothoracicans (Fig. 20*a*) parasitize coelenterates and echinoderms. *See* CIRRIPEDIA.

The Ostracoda is one of the largest of crustacean taxa. It is divided into two subgroups: the Myodocopa, in which the antennae are strongly developed for swimming, and the Podocopa, in which the antennae are specialized for walking. These are generally minute forms that occur in just about every aquatic environment. *See* MYODOCOPA; OSTRACODA; PODOCOPA.

Tantulocarida is a small subclass comprising ectoparasites of other crustaceans and are less than 0.3 mm in body length. However, they appear to be related as a sister group to the thecostracans. *See* TANTULOCARIDA.

The remaining maxillopodan subclasses, Mystacocarida and Branchiura, represent relatively few species. Both groups have gonopores on the fourth segment of the trunk. All mystacocarids are small, free-living meiofaunal animals with a preference for wave-washed, sandy, intertidal beaches; they have a small limbless abdomen extending posteriorly from the gonopores. The branchiurans, or fish lice, are ectoparasites of the gills of fish and occasionally amphibians. The branchiurans now contain another group, Pentastomida, which was long thought to be a separate phylum. However, the ultrastructure of their sperm and DNA sequence analysis indicates they are closely related to the fish lice. The pentastomids are all parasites of the respiratory systems of vertebrates. However, the ultrastructure of the sperm as well as DNA molecular sequence data clearly ally the pentastomids with the argulid branchiurans. *See* BRANCHIURA; MYSTACOCARIDA; PENTASTOMIDA.

Malacostraca. When most people think "crustacean," it is probable that it is a malacostracan that they have in mind, such as a shrimp, lobster, crayfish, or crab. The Malacostraca constitute a clearly definable, monophyletic group. However, it is very diverse and is subdivided into three subclasses: Phyllocarida, Hoplocarida, and Eumalacostraca, the last with the superorders Syncarida, Peracarida, and Eucarida. The most distinctive characters of the malacostracans include eight thoracic segments, each with a pair of appendages (pereiopods), and a pleon with six pairs of appendages. The first five pleopods are biramous limbs used in swimming, while for the hoplocaridans and eumalacostracans the sixth pair (uropods), together with the telson, form the tail fan. The female genital openings are invariably on the sixth thoracic segment, while male genital openings are on the eighth. Females may be provided with sperm receptacles, and males often have the first two

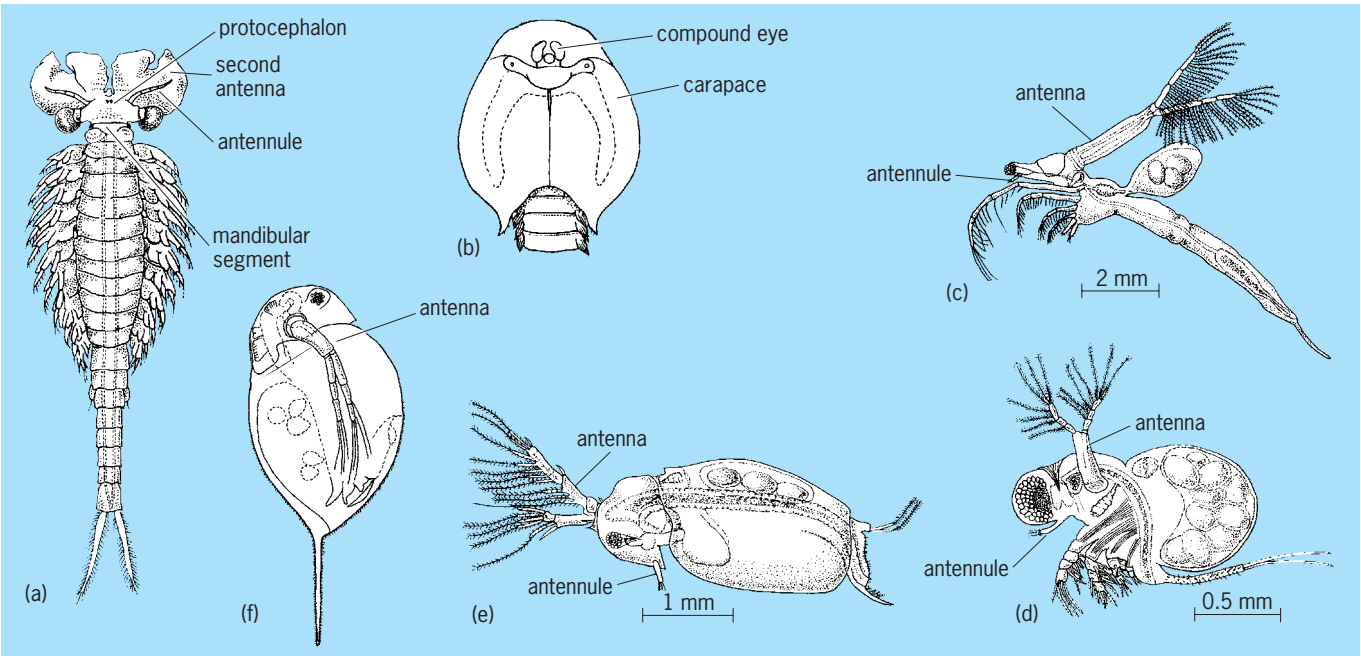

Fig. 18. Examples of branchiopods. (*a*) Anostraca (*Branchipus*). (*b*) Notostraca (*Triops*). (*c*) Hoplopoda (*Leptodora*). (*d*) Onychopoda (*Polyphemus*). (*e*) Ctenopoda (*Sida*). (*f*) Anomopoda (*Daphnia*). (*Parts c–f after W. T. Edmondson, ed., Freshwater Biology, John Wiley and Sons, 1966*)

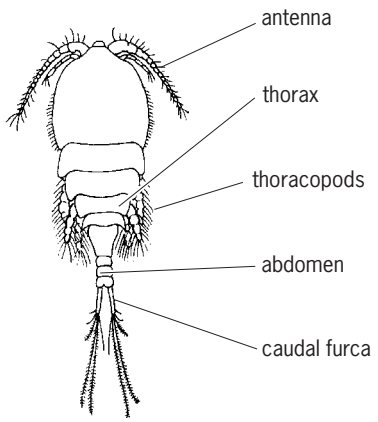

Fig. 19. *Cyclops*, a copepod.

pairs of pleopods modified to serve in sperm transfer (gonopods). *See* MALACOSTRACA.

The most primitive of the Malacostraca are the Phyllocarida, the only group of higher crustaceans to have the carapace modified into a bivalve shell. Extant representatives of the Hoplocarida are the highly specialized and predatory stomatopods, or mantis shrimp. Within the Eumalacostraca, the simplest body plan is exhibited by the syncarids, who lack a carapace (**Fig. 21**), while most other forms generally possess a carapace to some degree (**Fig. 22**). These latter include the eucarids, that is, the lobsters, crabs, shrimp and their allies, and the peracarids, an extremely diverse group among which are some of the few crustaceans adapted to true terrestrial conditions. *See* EUCARIDA; EUMALACOSTRACA; HOPLOCARIDA; PERACARIDA; PHYLLOCARIDA; SYNCARIDA.

Fossils

In recent years, studies of the arthropod fossil record have made significant contributions to our understanding of the origin and early evolution of the subphyla.

Cheliceriformes. The cheliceriform arthropods have a fascinating record. For example, while living and many fossil Pycnogonida (or sea spiders) lack an abdomen, a few Paleozoic fossils actually possess a long tubular abdomen. Furthermore, within the fossil record of Chelicerata we can note the presence of a variety of the eurypterids and scorpions, horseshoe crabs, and true spiders and their allies.

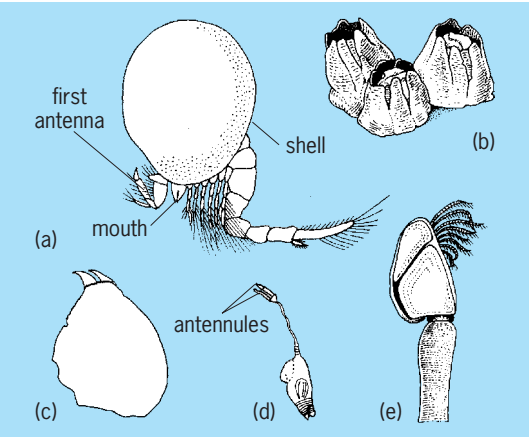

Fig. 20. Some types of Cirripedia. (*a*) Ascothoracica (*Synagoga*). (*b*) Thoracica (*Megabalanus*). (*c*) Acrothoracica (*Lithoglyptes*), female and (*d*) male. (*e*) Thoracica (*Lepas*). (*Parts c and d after J. T. Tomlinson and W. A. Newman, Lithoglyptes spinatus, a burrowing barnacle from Jamaica, Proc. U.S. Nat. Mus., 112:517–526, 1960*)

Eurypterids, with lengths of up to 6 ft (2 m), were the largest known arthropods. In addition, there are many fossils of the class Chelicerata that have no living analogs; these include the chasmataspids and some peculiar proto-chelicerates.

Myriapoda and Insecta. Although the insects and myriapods have more species than any other animal group, their fossil record is relatively poor. The earliest insects are Devonian wingless forms, but most insect fossils are little more than wings. The Late Paleozoic insect fauna was composed of so-called archaic types that—with a few notable exceptions such as cockroaches (Fig. 16*d*) and dragonflies—went extinct.

Crustacea. Most crustacean classes are small to microscopic in size and thus have a poor to modest fossil record. However, the fossil record of the crustaceans has greatly improved in the past two decades. The malacostracans (crabs, lobsters, shrimp, and their allies) are known from abundant fossils from the early Paleozoic to the present. Many of the malacostracan groups that were common in the late Paleozoic either went extinct in the Permian, or are today restricted to geographic or ecologic refuges, such as the southern continents, which once formed a single supercontinent called Gondwanaland, or groundwater habitats. However, the most noteworthy fossil arthropods are microscopic forms from the Cambrian of Sweden, in which minute details are preserved and which lend themselves to rather detailed stories as to the origin and early evolution of the group.

Other forms. Finally, several other "near-arthropods" have interesting fossil records, especially in the Cambrian. The Burgess Shale and similar deposits in China and Greenland contain peculiar animals that were once classified among the trilobitomorphs, but that researchers now realize may not even be arthropods *sensu stricto* (in the strict sense). These animals, such as the unique creature *Anomalocaris*, have wormlike bodies but possess one or more arthropod-like structures such as compound eyes or jointed limbs. The Onychophora, generally conceded to represent proto-arthropods, have a very limited fossil record; only one species is known from the Pennsylvanian Mazon Creek concretion faunas of Illinois, but a host of truly strange forms are known from the Cambrian. Finally, there are several fossil groups that are definitely arthropods, but that do not easily fit into the four major groups cited earlier. These include unusual creatures such as the euthycarcinoids and the thylacocephalans; much more infromation about them is needed before their relationships to other arthropods can be judged.

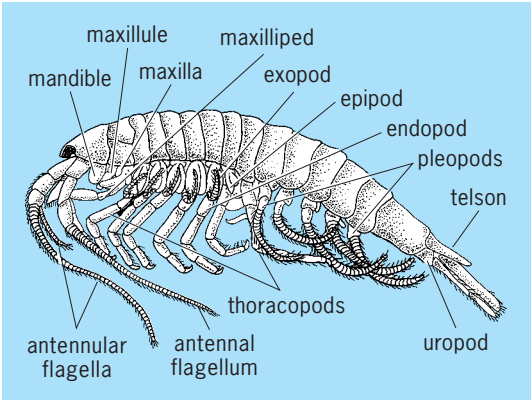

Fig. 21. Primitive malacostracan, *Anaspides tasmaniae* (Anaspidacea). (*After P. A. McLaughlin, Comparative Morphology of Recent Crustacea, W. H. Freeman, 1980*)

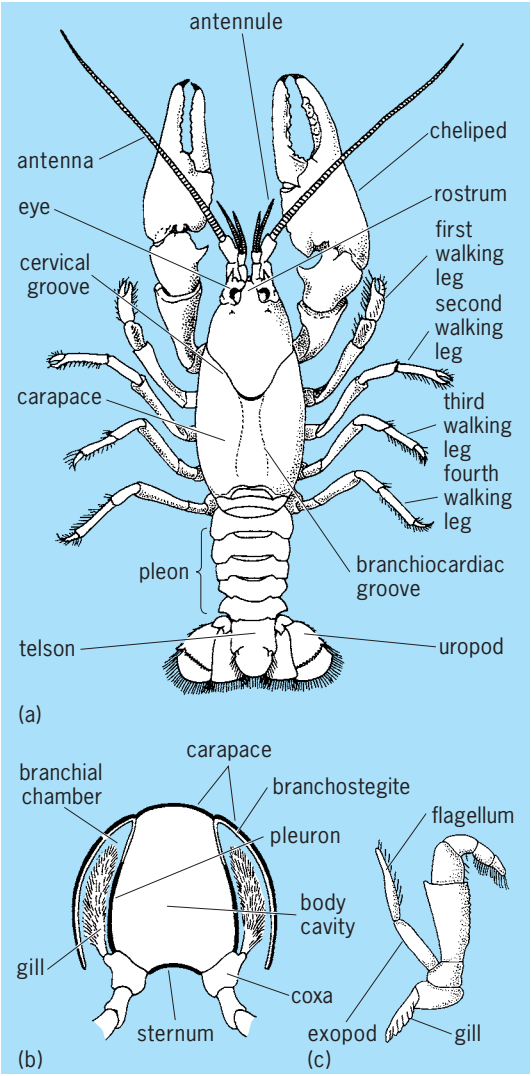

Fig. 22. Decapoda. (*a*) *Cambarus*, a fresh-water crayfish. (*b*) Cross section of thoracic segment of a decapod. (*c*) Third maxilliped. (*After R. E. Snodgrass, A Textbook of Arthropod Anatomy, Cornell University Press, 1952*)

Frederick R. Schram; Jerome C. Rozen, Jr.; Patsy A. McLaughlin

Bibliography. R. C. Brusca and G. Brusca, *The Invertebrates*, 2d ed., 2003; J. Cracraft and M. J. Donaghue (eds.), *Assembling the Tree of Life*, 2004; R. A. Fortey, *Trilobite! Eyewitness to Evolution*, 2000; R. A. Fortey and R. H. Thomas (eds.), *Arthropod Relationships*, 1998; S. Koenemann and R. A. Jenner (eds.), *Crustacea and Arthropod Relationships*, Taylor and Francis, 2005; F. R. Schram, *Crustacea*, 1986.

Artichoke

Cynara scolymus, a herbaceous perennial plant, in the family Compositae; also called globe artichoke. Its origin is in the Mediterranean region. Artichoke requires a mild winter and cool summer with fog and little bright sunshine. It is a delicacy in Europe, Africa, and North and South America. Artichoke is also a medicinal plant; it is rich in the cynarin and orthophenol constituents. Of the total acres planted worldwide, approximately 56%, 14%, and 12% are planted in Italy, France, and Spain, respectively.

The marketable portion of the plant, the so-called bud, is actually the immature flower head, made up of numerous closely overlaid bracts or scales (see **illus.**). The edible portion consists of the tender bases of the bracts, the young flowers, and the receptacle or fleshy base upon which the flowers are borne. The bud can be various shapes, from round to oblong to flat, and the color can be light green to dark green, often with purple or red.

Green Globe artichoke. (*Burpee Seeds*)

Artichoke is grown readily from seed, but the seedlings are highly variable and plants tend to throw back to the wild; hence, vegetative propagation by means of stumps (rootstocks), offshoots, or ovoli is recommended. The latter two are believed to produce earlier crops. The plant grows to a height of 3 or 4 ft (0.9 or 1.2 m), and sends up seasonal shoots from a permanent crown—as many as 12 or more in plants 4–5 years old. Each shoot forms a cluster of large basal or rosette leaves, from the center of which the stem grows. Buds are produced terminally on this elongated stem and on the lateral branches. If the buds are not removed, they develop into purple-centered thistlelike flowers and viable seeds are produced.

In the United States, artichokes are grown in the Pacific Coast area between south San Francisco and Los Angeles, mainly Monterey County, on about 10,000 acres (4050 hectares). The production per acre fluctuates according to weather conditions.

Aly M. Ibrahim

Articulata (Echinodermata)

The only surviving subclass of the Crinoidea, differentiated during Triassic times. The calyx is dicyclic, but considerable reduction of the infrabasals and basals may occur. The uniserial arms bear pinnules and usually branch, and the arm retains its movable articulation with the radial plate, despite the incorporation of the lower brachial ossicles into the calyx. Extant stalked forms with nodal rings of cirri (*Metacrinus*) are included in the order Isocrinida (see **illus.**). They do not tolerate turbulant waters and live at depths below current action, although they do inhabit shallow water when the conditions

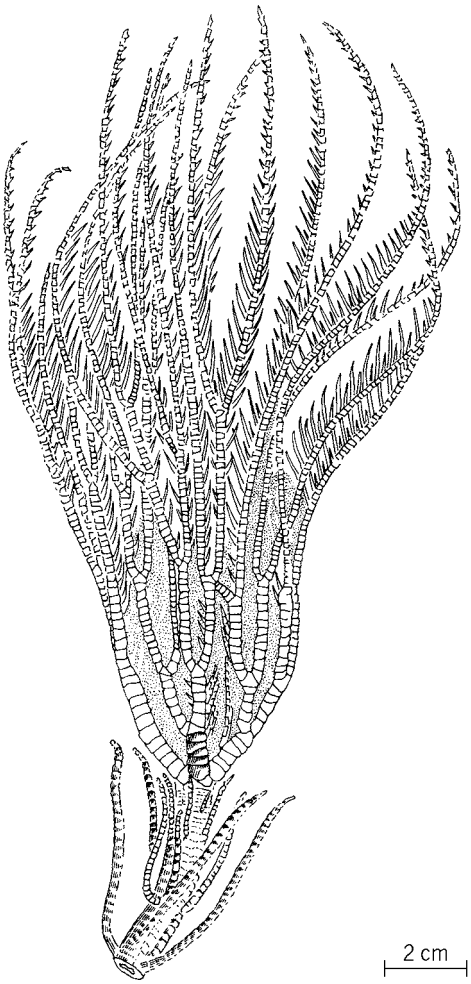

***Metacrinus cyaneus*, a living, articulate stalked crinoid from Bass Strait, at 200 fathoms (1200 ft or 365 m).**

are suitable. These forms may trail the stem during temporary free-swimming phases, using the cirri for reattachment. The feather stars, of the order Comatulida, discard the stem when young, and thereafter remain free, either as swimming animals or as creeping benthic forms. They prefer shallow, clear water, rich in nutrients, and therefore abound on tropical coasts and in polar seas rather than in temperate waters. Four other orders have been defined; one of them includes extinct floating forms. *See* CRINOIDEA; ECHINODERMATA. Howard B. Fell

Bibliography. A. H. Clark, *Monograph of the Existing Crinoids*, U.S. Nat. Mus. Bull. 82, 1915–1967.

Artificial intelligence

The subfield of computer science concerned with understanding the nature of intelligence and constructing computer systems capable of intelligent behavior. Many activities, such as perception, learning, planning, natural language, and understanding, involve intelligent behavior and include an immense diversity of phenomena. Fields such as psychology, linguistics, and philosophy share scientific concern for these phenomena. *See* COGNITION; INTELLIGENCE; INFORMATION PROCESSING (PSYCHOLOGY); LINGUISTICS.

Artificial intelligence is primarily concerned with representations of knowledge and heuristic methods of reasoning using common assumptions and rules of thumb. Two examples of problems studied in artificial intelligence are (1) planning how a robot, or person, might assemble a complicated device or move from one place to another and (2) diagnosing the nature of a person's disease or of a machine's malfunction from the observable manifestations of the problem.

Approach. Approaches of artificial intelligence research can be experimental or theoretical. As in other experimental fields, investigators build devices (in this case, computer programs) to carry out their investigations. Programs are created to explore ideas about how intelligent action might be attained, as well as to test hypotheses about concepts or mechanisms involved in intelligent behavior to confirm theories.

Much of the inspiration for the organization and details of artificial intelligence programs comes from observing and studying the behavior of people in different situations.

Foundations. The four foundations of artificial intelligence are representation, search, reasoning, and learning. These four foundations are typical ingredients of any artificial intelligent system and are addressed to some degree in most artificial intelligence research.

Representation. A computer system must have an internal representation of the problem or related knowledge. Examples of representation include the symbolic description of a room for a moving robot and a description of a person with a disease for a classification program. A representation also includes all the knowledge, including basic programs for testing and measuring a structure in question, plus all the programs for transforming the structure into another one in ways appropriate to the task. The representation of a problem typically implies an architecture on which it can be manipulated. Two examples of architectures are the programming language used for the representation (often high-level languages, like LISP) and the database used to store and query knowledge. Changing the representation used for a task can make an immense difference, potentially turning a problem from impossible to trivial. *See* DATABASE MANAGEMENT SYSTEM; PROGRAMMING LANGUAGES.

Search. An important aspect of most artificial intelligence systems is that they use a problem-solving methodology that involves search. Artificial intelligence has gradually built up a stock of general search methods that can be used. In the worst case, there is no information to guide the search and the search is blind. In heuristic search, extra information is used to guide the search. *See* GENETIC ALGORITHMS.

Search methods include blind breadth-first and depth-first (as used in game-tree search); generate-and-test (a sequence of candidates is generated, each being tested for a solution); hill climbing (a measure of progress is used to guide each step); means-ends analysis (the difference between the desired situation and the present one is used to select the next step); impasse resolution (the inability to take the desired next step leads to a subgoal of making the step feasible); planning by abstraction (the task is simplified, solved, and the solution used as a guide); and matching (the present situation is represented as a schema to be mapped into the desired situation by putting the two in correspondence).

Reasoning. Given the representation and knowledge of a problem, many forms of reasoning exist to turn the knowledge into a solution. Examples of reasoning include expert systems (typically rule-based systems developed using human expertise to identify the rules of the problem); case-based reasoning (a database of previous problems and solutions are searched for the closest match to the current problem); deductive and inductive reasoning (problem knowledge is used to deduce a set of possible solutions, or used to build a hypothesis that best explains the existing knowledge and possibly the current problem); and resolution in logic systems (tools used to determine if the current hypothesis or problem is true or not given the facts we know about the universe in which the problem exists). *See* LOGIC; PROBLEM SOLVING (PSYCHOLOGY).

Learning. Most artificial intelligent systems use the ability to adapt, or learn, based on the history, or knowledge, of the system. Learning takes the form of updating knowledge, adjusting the search, reconfiguring the representation, and augmenting the reasoning. Some of the most common learning methods used in artificial intelligence systems are statistical learning (using the number of the different types of historical events to bias future action or to develop inductive hypotheses, typically assuming that events

follow some known distribution of occurrence); neural networks (the updating of weights in a network, where networks are first trained on existing knowledge and then applied to the problem, outputting a value that can be interpreted as a solution); and reinforcement learning (the actions taken by an artificial intelligent system are rewarded or penalized based on their ability to solve the problem more or less accurately). *See* NEURAL NETWORK.

Examples. The following are examples of research in artificial intelligence.

Games. Games are an important application area in artificial intelligence. The initial efforts were in writing general problem-solving programs that would learn how to play games. After discovering that learning was not sufficient to produce game-playing programs that could compete against moderately good human players, the emphasis changed to writing sophisticated programs that were excellent at playing a specific game.

A good example of this approach is chess programs. In 1997, IBM's Deep Blue chess program beat Garry Kasparov, the world champion. A typical chess program uses a representation of a chessboard to store the current positions of the chess pieces, plus representations of standard chess openings, favorable chess positions, and possibly significant games. The program also stores representations of the possible chess moves. A chess program will often use a search function to find a good move. A chess program will take the current board and try all the program's possible moves (a half-step) and then all the possible responses that the opponent could make (the second half-step). The sequence of two such half-steps is a called a ply. A simple chess-playing program might investigate two or three plys ahead. Since the number of such positions in a chess game will be huge, the program may have an evaluation function (heuristic) for the "goodness" of a possible set of positions. This function will be used to decide which possible moves to investigate. The evaluation function may be used to eliminate (or prune) low-value (to the computer) positions and to expand the high-value positions to more depth.

Recently, efforts have been made to refocus the game-playing research back to general game players which accept the game rules as input and can play a variety of games, instead of programs that are carefully designed to play just one game.

Natural language processing. Extracting information from natural language input is an important capability, since so much human knowledge is stored in natural language (for example, textbooks) and much interaction (for example, commands and advice) occurs as spoken language. While an easy task for humans, understanding natural language is a very difficult task for computers. One important part of a natural language processor is the representation of the syntax and semantics of the language. Computer science has learned how to deal with formal languages effectively. However, natural languages inherently involve ambiguity. Approaches to natural language processing often try to build databases of previous examples and meanings of natural language, from which statistically driven hypotheses can be generated and applied to new examples. A good example of natural language processing is found in audio data-entry systems, such as those used by telephone-based help systems. Using a restricted set of expected word utterances at a given time, the system compares the acoustical signatures of the users spoken responses to those it expects to hear. By repeating the understood response to the user and asking for a confirmation, the system can update its database of acoustical signatures for that response. In doing so, the system is able to learn from specific users and all users in general. *See* NATURAL LANGUAGE PROCESSING.

Perception. Perception is the extraction, from a sensory input, of an internal representation of the external world suitable for intelligent processing. Though there are many types of sensory signals, the richest sensory input is vision. A common task in vision-perception programs is object recognition. The perception program has a representation of the objects to be identified. These representations of objects can often be sets of features that distinguish each object from other objects. Features can be edges, surfaces, or shapes. The object-recognition program then tries to match the features in the input with the objects and to identify which objects are present. However, this task is very difficult. Most objects in the real world are three-dimensional. From different angles, an object can look very different. Human vision is good at compensating for the different angles; computer programs are not. *See* COMPUTER VISION; PERCEPTION.

A different approach to object recognition is the use of artificial neural nets (networks). Each object to be identified is handled by a neural net. The neural net is trained to recognize its object. With sufficient training, the set of neural nets will be able to identify which of the set of objects appears in the scene. This approach is very useful in situations such as the military application of recognizing enemy vehicles in surveillance images.

Autonomous mobile robotics. Popular literature has many examples of autonomous robots that make human life easier by doing those tasks that are repetitious, dangerous, or unpleasant. Although there have been some successes, such as rovers on Mars and industrial robots that help in manufacturing tasks, there are many more possibilities in which autonomous robots would be very useful. Mobile robots are devices that have means of locomotion (wheels), sensors (sonars, lasers, and cameras), communications (wireless), and often actuators (arms or grippers).

A typical autonomous robot would have a representation of its world, possibly a map that shows the locations of walls, doors, obstacles, and other features of interest. Localization, the task of determining the robot's exact position in such a representation, is a major problem in robotics. Probabilistic methods have been used very successfully in matching sensor readings to the representation to determine the most

likely position of the robot. An autonomous robot will have a goal to be achieved, such as exploring a disaster area looking for victims. A major task of the program will be planning the actions for achieving the goal. This will include path planning to search the whole area, sensing and moving around obstacles, and identifying possible victims. *See* ROBOTICS.

Expert systems. Even when no firm scientific base exists to make a decision, extensive experience enables people to accomplish many tasks successfully. A class of artificial intelligence programs, called expert systems, attempt to exhibit equivalent performance by acquiring and incorporating the same knowledge that human experts have. Many attempts to apply artificial intelligence to medicine, government, and other socially significant tasks take the form of expert systems. Even though the emphasis is on knowledge, all the standard ingredients are present, including the representation of the task, representation of the knowledge, and reasoning or search methods. *See* EXPERT SYSTEMS.

A typical example of an expert-system task is a medical diagnosis program, which would have "if-then" rules that a doctor would use. The program might prompt for answers to questions very similar to what a doctor would ask. With a sufficient number of correct if-then rules, the program might be able to determine a correct diagnosis. One advantage of the expert-systems program is that it can use many more if-then rules than a human could keep in mind. Very rare diseases or conditions could be covered with the rules.

Obtaining and codifying implicit expert knowledge in if-then rules is a critical aspect of developing artificial intelligence expert programs. That entire bodies of knowledge can be sets of active if-then rules, even though they did not yet form a coherent scientific theory, has been an important discovery. An important variation of the expert-systems program is the inclusion of fuzzy logic, which allows uncertainty to be handled explicitly. *See* FUZZY SETS AND SYSTEMS.

In careful tests, a number of expert systems have shown performance at levels of quality equivalent to or better than average practicing professionals on the restricted domains over which they operate. Nearly all large corporations, and many smaller ones, use expert systems. A common application is to provide technical assistance to persons who answer customers' trouble calls. Computer companies use expert systems to assist in configuring components from a parts catalog into a complete system that matches a customer's specifications, a kind of application that has been replicated in other industries by tailoring assembled products to customers' needs. Troubleshooting and diagnostic expert systems programs are also commonplace.

Expert systems have sparked important insights in reasoning under uncertainty, causal reasoning, reasoning about knowledge, and acceptance of computer systems in the workplace. They illustrate that there is no hard separation between pure and applied artificial intelligence, since finding what is required for intelligent action in a complex applied area makes a significant contribution to basic knowledge.

Scope and implications. Research in artificial intelligence explores the full range of intellectual tasks. In addition to the subject areas mentioned, significant work has been done on puzzles and reasoning tasks, induction and concept identification, symbolic mathematics, theorem-proving in formal logic, natural language understanding and generation, vision, robotics, chemistry, biology, engineering analysis, computer-assisted instruction, and computer-program synthesis and verification, to name only the most prominent. As in any developing technology, there occurs both a reinforcement of the basic ideas (about representation, search, reasoning and learning) and the discovery of new mechanisms that both extend existing ideas and reveal limits.

Artificial intelligence has close ties to several surrounding fields. As part of computer science, it plays the role of expanding the intellectual sophistication of the tasks to which computers can be applied. Various subfields, once viewed as part of artificial intelligence, have become autonomous fields, most notably symbolic mathematics and program verification. Work on vision and speech overlaps with the field of pattern recognition in electrical engineering in the concern with efficient signal processing.

Numerous applications of artificial intelligence have been deployed that demonstrate savings in money or time to business, industrial, military, scientific, and health organizations. The success in expert systems, robotics, and perception has precipitated substantial commercial activity, including the formation of many small venture firms. As computers become smaller and less expensive, more intelligence is built into automobiles, appliances, and other machines, as well as computer software, in everyday use. *See* COMPUTER; CONTROL SYSTEMS; DIGITAL COMPUTER; INTELLIGENT MACHINE; SOFTWARE.

Steven M. Gustafson; David A. Gustafson

Bibliography. T. Mitchell, *Machine Learning*, McGraw-Hill, New York, 1997; S. Russell and P. Norvig, *Introduction to Artificial Intelligence*, Prentice Hall, New Jersey, 2003; S. C. Shapiro (ed.), *Encyclopedia of AI*, Wiley, New York, 2000.

Artificially layered structures

Manufactured, reproducibly layered structures having layer thicknesses approaching interatomic distances. Modern thin-film techniques are at a stage at which it is possible to fabricate these structures, also known as artificial crystals or superlattices, opening up the possibility of engineering new desirable properties into materials. In addition, a variety of problems in solid-state physics can be studied which are otherwise inaccessible. The various possibilities include the application of negative pressure, that is, stretching of the crystalline lattice; the study of dimensional crossover, that is, the transition from a situation in which the layers are isolated and two-dimensional in character to a situation in which the

layers couple together to form a three-dimensional material; the study of collective behavior, that is, properties that depend on the cooperative behavior of the whole superlattice; and the effect and physics of multiple interfaces and surfaces. These structures serve as model systems and as a testing ground for theoretical models and for other naturally occurring materials that have similar structures. For instance, ceramic superconductors consist of a variable number of conducting copper oxide (CuO_2) layers intercalated by various other oxide layers, and therefore artificially layered structures may be used to study predictions of the behavior of suitably manufactured materials of this class. A variety of applications have also been proposed or discovered. Of course, one of the most exciting prospects is the discovery of new, as yet unpredicted phenomena. For a discussion of semiconductor superlattices *see* CRYSTAL STRUCTURE; SEMICONDUCTOR HETEROSTRUCTURES

Preparation. Artificially layered structures have been prepared since the beginning of the century, using mostly chemical methods for deposition. With the advent of sophisticated high-vacuum preparation techniques, there has been a significant increase in activities in this field since the early 1970s. The preparation techniques can be conveniently classified into two groups: evaporation and sputtering. In an evaporation system, two or more sources of particles (which end up being part of the final structure) are aimed at a heated substrate where the artificially layered structure is grown. The rates are precisely controlled by using rate monitors, and the various beams are chopped by using shutters in various configurations. The ultrahigh-vacuum (UHV) version has been commonly designated as a molecular beam epitaxy (MBE) apparatus. The sputtering method relies on bombarding targets of the proper materials with an inert gas, such as argon, thus producing the beams of the various elements. Commonly, in this case, the substrate is held against a rotating, heated table and so moved from one beam to the next. *See* CRYSTAL GROWTH; MOLECULAR BEAMS; SPUTTERING.

Structure. Once the artificially layered structure is prepared, it is necessary to characterize whether the layer structure is stable at the growth temperature. This is of considerable importance, since the interdiffusion of the constituents in many cases eliminates the layered growth. A direct image of the layers in gallium arsenide–aluminum arsenide superlattices, using electron micrography, is shown in **Fig. 1**. One of the most successful methods of characterizing layered growth has been x-ray diffraction. In this method the intensity of x-rays elastically reflected from a sample is measured as a function of incidence angle. The layered growth is indicated by the existence of many superlattice peaks, which are due to the Bragg reflection by the superlattice planes. The spacing of these peaks is related to the thickness of the layers, whereas the amplitude is related to the admixture of one constituent into the other, the interfacial roughness, and the amount of strain present. *See* ELECTRON MICROSCOPE; X-RAY DIFFRACTION.

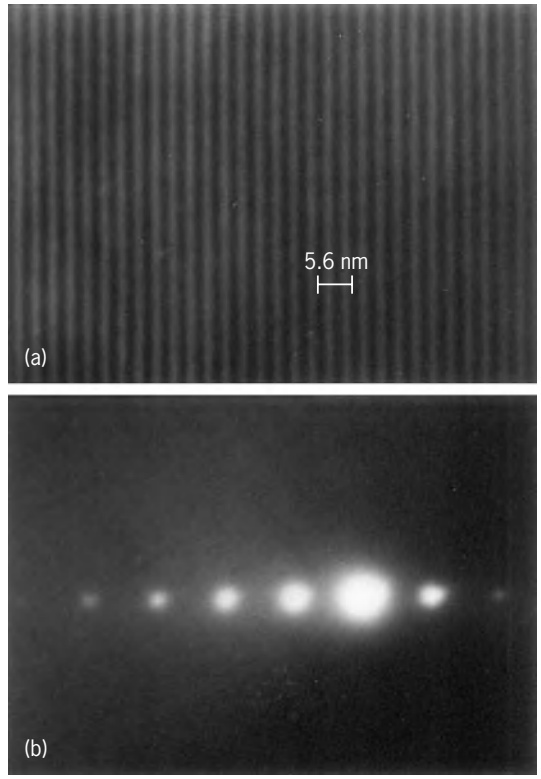

Fig. 1. Ordered superlattice consisting of alternating layers of gallium arsenide (GaAs) and aluminum arsenide (AlAs), each four atomic layers thick, deposited by molecular beam epitaxy. (*a*) Transmission electron micrograph of the superlattice. (*b*) Electron diffraction pattern of the superlattice. New diffraction spots around the central diffraction peak are the result of Bragg scattering from the long periodicity of the superlattice. (*From P. M. Petroff et al., Crystal growth kinetics in* $(GaAs)_n$-$(AlAs)_m$ *superlattices deposited by molecular beam epitaxy, J. Cryst. Growth, 44:5–13, 1978*)

Physical properties. The normal-state properties of non-lattice-matched metallic superlattices show anomalies at small thicknesses of the order of 2.0 nanometers periodicity (that is, repeat distance; **Fig. 2**). The lattice expands slightly (approximately 2%) in the direction perpendicular to the layers, the shear elastic constant decreases markedly (approximately 35%), and the temperature coefficient of resistivity changes sign, as in a metal–nonmetal transition.

Artificially layered structures that are superconducting exhibit dimensional crossover, as is shown (**Fig. 3**) beautifully by the temperature dependence of the superconducting critical field. When the superconducting layers are thin and well separated, the critical field shows a characteristic square-root-like behavior, as is expected for two-dimensional superconductors. If the superconducting layers are close together, the layers are strongly coupled and show a linear temperature dependence typical of a three-dimensional material. At intermediate separation the behavior can change from three-dimensional at high temperatures to two-dimensional at low temperatures, a phenomenon known as dimensional crossover. Since the high-temperature ceramic superconductors are also layered, many of their properties

are similar to those of artificially layered superconductors. *See* SUPERCONDUCTIVITY.

Because of a short-range coupling mechanism that is not yet understood, the ferromagnetic layers in ferromagnetic-normal superlattices order with antiparallel (antiferromagnetic) alignment. As a consequence, the resistivity can exhibit changes as large as 150% with applied field (giant magnetoresistance). This amount is much larger than the ordinarily observed magnetoresistance of 2–3% in magnetic materials. In some cases a variety of magnetic coupling mechanisms have been observed across nonmagnetic layers, giving rise to new magnetic structures (in dysprosium-yttrium superlattices, for instance) and collective spin waves in molybdenum-nickel superlattices. *See* MAGNETORESISTANCE; MAGNON.

Applications. Artificially layered structures are especially useful for the construction of mirrors for soft x-rays since there are no suitable, naturally occurring crystals for this purpose. This application was one of the main motivations for research on artificially layered structures. Mirrors and polarizers for neutrons have also been manufactured and are currently used in the United States and Europe.

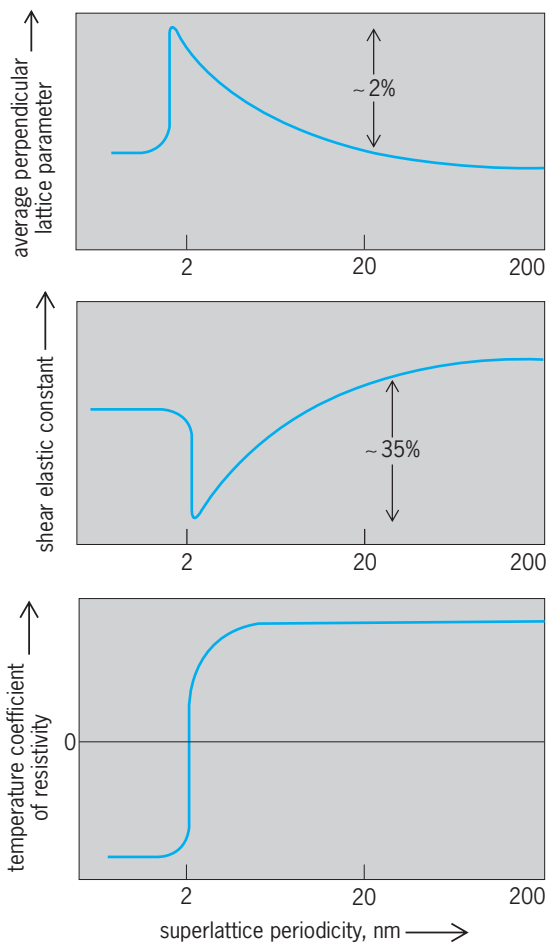

Fig. 2. Average perpendicular lattice parameter, shear elastic constant, and temperature coefficient of resistivity versus superlattice periodicity. (*After M. R. Khan et al., Structural, elastic and transport anomalies in molybdenum-nickel superlattices, Phys. Rev., B27:7186–7193, 1983*)

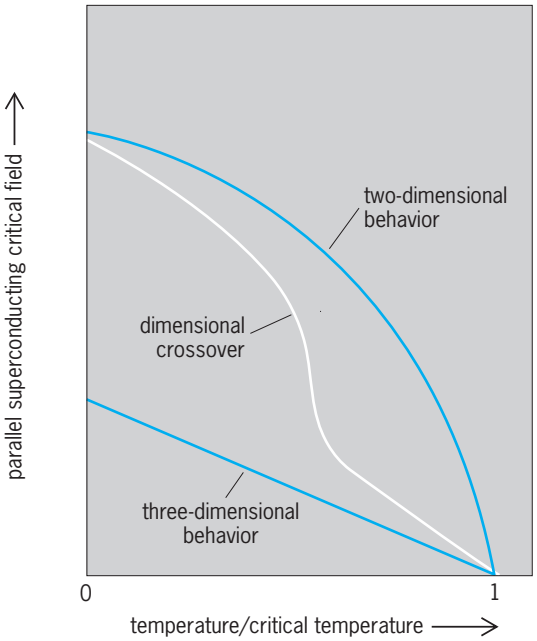

Fig. 3. Parallel superconducting critical field versus temperature at various layer thicknesses. (*After S. T. Ruggiero, T. W. Barbee, and M. R. Beasley, Superconductivity in quasi-two-dimensional layered deposits, Phys. Rev.Lett., 45:1299–1302, 1980*)

Superlattices with zero temperature coefficient of resistivity are useful as resistor material, and high-critical-field-magnet tapes using superconducting-insulator superlattices have been proposed. Magnetic superlattices that exhibit giant magnetoresistance are incorporated into magnetoresistive recording heads. *See* ELECTRICAL INSULATION; NEUTRON OPTICS; X-RAYS. Ivan K. Schuller

Bibliography. M. Bartusiak, Architects in the laboratory, *Discover*, 2:60–61, 1981; P. Dhez and C. Weisbuch (eds.), *Physics, Fabrication and Applications of Multilayered Structures*, 1988; J. D. Dow and I. K. Schuller (eds.), *Interfaces, Superlattices and Thin Films,* 1987; S. Maekawa et al. (eds.), *Metallic Multilayers*, 1993; I. K. Schuller, E. E. Fullerton, and Y. Bruynseraede, Quantitative x-ray diffraction from superlattices, *Mater. Res. Soc. Bull.*, 12(12): 33–38, 1992.

Artiodactyla

An order comprising the even-toed ungulates (hoofed mammals). There are two main radiations: the predominantly omnivorous Bunodontia, including suoids (such as pigs, peccaries, and hippos); and the more herbivorous Selenodontia, including camels and ruminants (such as deer, giraffe, cattle, sheep, and antelope). Artiodactyla contains about 213 living species, making it the fifth most speciose order of mammals (exceeded primarily by rodents and bats, and slightly by carnivores and primates). First known from the early Eocene, artiodactyls have proliferated during the last 55 million years to reach

great diversity (especially among the family Bovidae). Their radiation is often contrasted with that of the odd-toed ungulates, or Perissodactyla (horses, rhinos, and tapirs). Artiodactyls are also important for human economy and agriculture, comprising most of the domestic animals, providing milk, wool, and most of the meat supply. *See* PERISSODACTYLA.

Morphology . Artiodactyls are defined by a unique morphology of the ankle joint, possessing a "double-pulley" astragalus. In the general mammalian condition this ankle bone has a single pulleylike articular surface on its dorsal surface (for articulation with the tibia, or shin bone), whereas in artiodactyls it also has a second pulleylike articulation on its ventral surface, forming an articulation with the lower row of ankle bones (the cuboid and the navicular) [**Fig. 1**]. This morphological feature is frequently considered to be a key innovation of artiodactyls, but its precise functional significance is poorly understood.

Artiodactyls also have a paraxonic foot structure, where the axis of limb support passes between the third and fourth digits. The first toe is usually completely lost. The foot posture is usually unguligrade

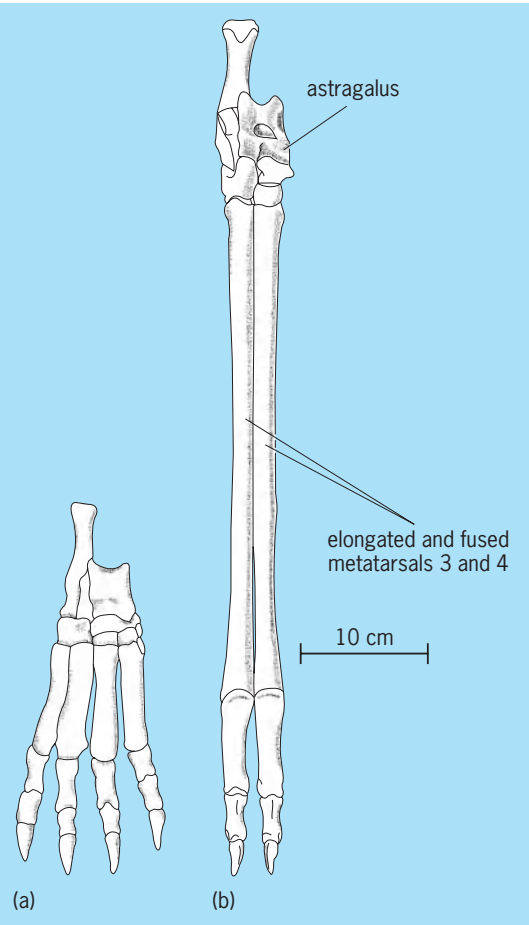

Fig. 1. Representative left hindfeet of artiodactyls, showing the double-pulley astragalus. (*a*) Primitive condition, as in the Oligocene oreodont *Agriochoerus* (although the clawed condition in this animal is a secondary one). (*b*) Derived condition, as in the Miocene camelid *Oxydactylus*. (*After A. S. Romer, Vertebrate Paleontology, 3d ed., University of Chicago Press, 1966*)

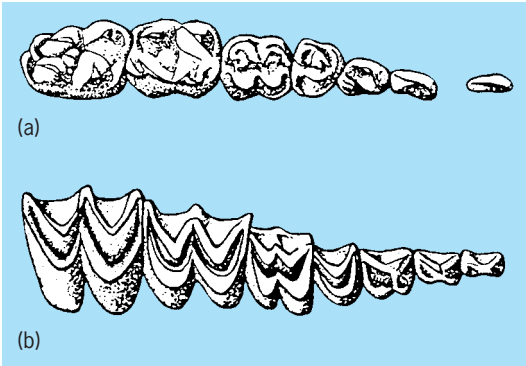

Fig. 2. Representative right upper cheek teeth of artiodactyls. (*a*) Primitive, bunodont condition in *Chleuastochoerus*, a Pliocene suid. (*b*) Derived, selenodont condition in *Merycochoerus*, a Miocene oreodont. (*After A. S. Romer, Vertebrate Paleontology, 3d ed., University of Chicago Press, 1966*)

(standing on the most distal phalanx, like a ballerina on point). Some primitive extinct artiodactyls retain a more generalized, digitigrade stance (standing on the digits, like a person on tiptoe). Digitigrady is also secondarily adopted in derived camelids, where the hooves are replaced by a padlike foot. In primitive artiodactyls and in suoids, the foot is four-toed, and the metapodials (hand and foot bones) are short and unfused as in the generalized mammalian condition (Fig. 1*a*). In more derived artiodactyls, digits 2 and 5 have been reduced or lost entirely, and metapodials 3 and 4 are lengthened and fused to form a "cannon bone" (Fig. 1*b*). It is this fusion of the metapodials with digits 3 and 4 free that gives ruminants their cloven-hoofed appearance (in contrast with the single and solid hoof of horses).

In the primitive artiodactyl condition, retained in most living suoids, the cheek teeth have a bunodont crown morphology (**Fig. 2***a*). This is the generalized morphology of low, rounded cusps seen in many omnivorous mammals, including humans. In more derived artiodactyls, which have a herbivorous rather than an omnivorous diet, the cheek teeth are selenodont: the crown pattern is for the individual cusps to be run together into crescentic lophs, one pair on the buccal (outer) side of the each tooth and another pair on the lingual (inner) side (Fig. 2*b*). These lophs provide the teeth with high-standing ridges, better adapted to shred fibrous food than the bunodont teeth, which are adapted to pulp fruits and roots. The primitive condition is for these teeth to be brachydont, or low-crowned, but in some selenodont artiodactyls (especially in grazers) the teeth are hypsodont (high-crowned), which renders them more durable for a highly abrasive diet. *See* DENTITION; TOOTH.

The skulls of suoids and selenodonts also reflect their differences in diet (**Fig. 3**). In ruminants (and also convergently camelids and protoceratids, an extinct related family) the upper incisors and canine are reduced or absent. However, an enlarged upper canine may be retained for display in some hornless males, as seen in living mouse deer and musk

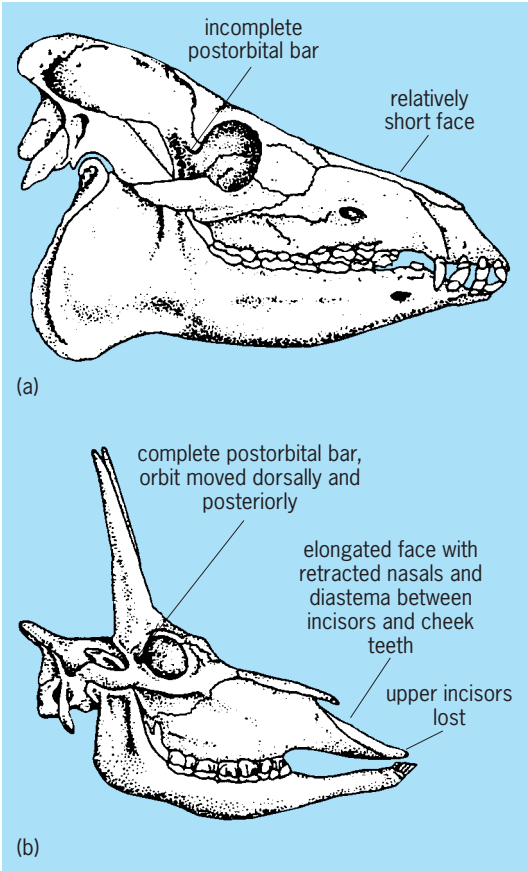

Fig. 3. Representative artiodactyl skulls. (*a*) Suoid condition, in the Oligocene anthracothere *Bothriodon*. (*b*) Selenodont condition, in the Miocene giraffe *Samotherium*. (*After A. S. Romer, Vertebrate Paleontology, 3d ed., University of Chicago Press, 1966*)

deer (**Fig. 4***a*). The lower canine has become incisiform in shape, and these lower incisors now meet against a horny pad in the upper jaw. Horns or other types of cranial appendages (such as antlers and ossicones) are a typical feature of ruminants, although they also evolved in protoceratids, and a single median horn on the forehead is known in the extinct pig *Kubanochoerus*. Different types of cranial appendages are not homologous in their structure and development, and most likely they evolved independently in the different families (Fig. 4). Horns are usually known only in the males, and are used for intraspecific display and combat. Horns have evolved in the females in species which form year-round associations of mixed sex herds, such as cattle, some grazing antelope (such as the wildebeest), giraffe, and reindeer. *See* ANIMAL COMMUNICATION; TERRITORIALITY.

Selenodont artiodactyls, like other herbivorous mammals, require a gut fermentation chamber to house microorganisms to ferment the cellulose of plant cell walls. Selenodonts have a forestomach site of fermentation, situated before the true stomach. This site comprises three chambers in ruminants (the rumen, omasum, and abomasum), but only two in camels (where a distinct reticulum is absent). Selenodonts differ from other forestomach-fermenting mammals in their remastication of the food, or chewing of the cud. Selenodonts retain their food for a long time and extract the maximal amount of nutrients out of the cellulose. This gives them an advantage where food is of limited quantity, but they are not able to bulk-process food of low quality, in contrast with less efficient hindgut fermenters such as horses.

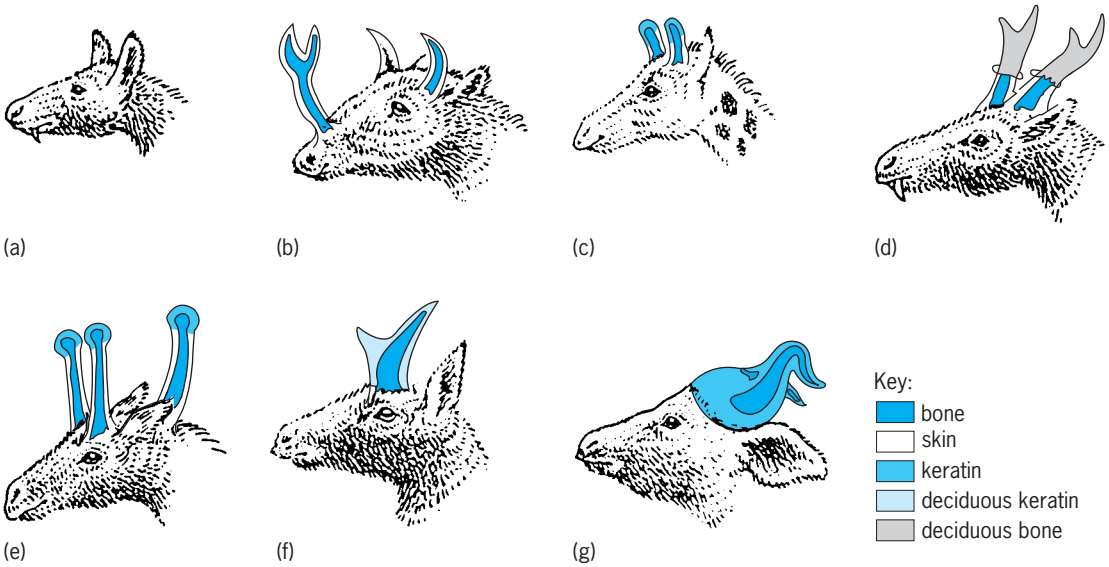

Fig. 4. Heads of male selenodont artiodactyls, showing different types of cranial appendages. Relative thickness of skin or keratin cover is not to scale. (*a*) Tragulid, showing primitive hornless condition with enlarged upper canines. (*b*) Protoceratid, an extinct tylopod convergent with ruminants in horn acquisition, with a forked nasal horn in addition to postorbial ones. (*c*) Giraffid (giraffe), with ossicones. (*d*) Cervid (muntjac deer), with antlers. (*e*) Dromomerycid, extinct cervid-related group with an occipital horn in addition to postorbital ones. (*f*) Antilocaprid (pronghorn), with deciduous keratinous horn sheath. (*g*) Bovid (cow), with true horns. (*After C. M. Janis, Evolution of horns and related structures in hoofed mammals, Discovery, 19(2):8–17, Yale Peabody Museum of Natural History, 1986*)

-s

Classification, evolution, and distribution. Despite their host of unique (apomorphic) features, in terms of branching patterns artiodactyls represent an early divergence from the main stem of ungulate evolution. The earliest artiodactyls, rabbit-sized animals found in the Eocene of North America and Europe, are commonly known as dichobunids. By the late middle Eocene the earliest members of the two modern radiations, the Bunodontia and the Selenodontia, had arisen from among different groups of dichobunids. The three main types of living artiodactyls—suoids, camelids, and ruminants—can trace their roots back to this Eocene divergence (**Fig. 5**). There has been much debate as to whether whales (order Cetacea) should be included among the Artiodactyla. Whales are clearly the sister taxon to artiodactyls among modern mammals, but some molecular studies imply that they are closely allied with the hippos. Morphologists and paleontologists would generally prefer to keep whales separate from artiodactyls. *See* CETACEA.

Suborder Suiformes. All living members of the Bunodontia are contained within the Suiformes (suoids) comprising the families Suidae (pigs), Tayassuidae (peccaries), and Hippopotamidae (hippos). Suoids are primarily omnivorous and live in forest, woodland, or brushland. Some suoids, such as the warthog (Suidae) and the hippo, live in grasslands and eat more fibrous herbage (such as grass). Such suoids have evolved lophed teeth and some type of forestomach fermentation convergently with selenodont artiodactyls, although they do not chew the cud. Suoids are primarily an Old World radiation, except for the New World peccaries. Fossil suoids include the Anthracotheriidae, piglike forms with lophed teeth known from the early Tertiary of North America and until the Pliocene of the Old World. Anthracotheres may have been ancestral to hippos. Other extinct bunodonts (possibly grouped within the Suiformes) include the European small to medium-size Cebochoeridae and Choeropotamidae, and the Entelodontidae of North America and Eurasia, which were a radiation of giant hogs, some as large as bison, that may have been carrion scavengers. *See* HIPPOPOTAMUS.

Suborder Tylopoda. The only living selenodont artiodactyls that are not members of the suborder Ruminantia are the camelids. The suborder Tylopoda is an assemblage of nonruminant selenodont families that may not be closely related to each other. Today, the family Camelidae is known from the large, desert-dwelling camels of the Old World and the antelope-size llamas (including alpacas and vicunas) of the mountains and pampas of South America. Present-day camelids are mixed feeders, living in herds in open habitats. However, until around 2 million years ago, camelids were an exclusively North American group. They first appeared in the late Eocene, and during the later Cenozoic they were a highly diverse radiation, paralleling the African ruminants in their morphological diversity (although none appeared to be true grazers). Their diversity declined in the Pliocene, and they became extinct in North America around 10,000 years ago at the end of the Pleistocene. *See* ALPACA; CAMEL; LLAMA.

Extinct tylopods can be divided into three main groups: (1) small to medium-size European forms of the early Tertiary, including the sheeplike Anoplotheriidae, the rabbitlike Cainotheriidae, and the llamalike Xiphodontidae; (2) some North American forms probably closely related to camelids, the rabbitlike Oromerycidae and the deerlike Protoceratidae, whose later members sported horns both over the eyes and on the nose (Fig. 4*b*); (3) and the North American oreodonts (families Agriochoeridae and Merycoidodontidae), the most common North American fossils of the Oligocene, that resembled a cross between a pig, a sheep, and a hyrax.

Suborder Ruminantia. Ruminants can be divided into the more primitive Tragulina (traguloids), an assemblage of small, hornless forms, and the more derived Pecora, which usually comprise larger, horned forms. Traguloids are first known from the late middle Eocene of Eurasia and North America, and are a mostly early Tertiary radiation. They include the families Amphimerycidae, Hypertragulidae, and Leptomerycidae. Traguloids usually have relatively short legs with unfused metapodials, and the males have saberlike upper canines. Surviving traguloids are the members of the family Tragulidae, including the mouse deer (*Tragulus*) of southeast Asia and the chevrotain (*Hyemoschus*) of central Africa. Tragulids are solitary, tropical-forest-dwelling animals that eat mainly nonfibrous leaves and fruit and do not appear to chew the cud. *See* CHEVROTAIN.

The earliest known pecorans were the Eurasian Gelocidae, first appearing in the late Eocene. These animals were rather like traguloids in general appearance but had more derived features of the limbs, skull, and dentition. The living musk deer (family Moschidae) has a similar appearance: small-size, with the males possessing large upper canines rather than horns. Modern moschids are known only from Asia, but are also recorded from the Miocene of North America (the blastomerycines) and the Oligocene of Europe. The modern, horned pecoran families date from the early Miocene. Pecorans also have more derived limb, with longer, fused metapodials; the reduction or loss of digits 2 and 5; and a reduced ulna and fibula. These limbs are better adapted for fast, sustained locomotion.

The family Giraffidae is an exclusively Old World group, known today only from Africa, but also from Asia until the Recent. Giraffids have ossicones, simple skin-covered cranial appendages (Fig. 4*c*). There are two living species of giraffids, both browsers (although some extinct giraffids had a more mixed-feeding diet): the forest-dwelling, solitary okapi (*Okapia*) and the savanna-dwelling, herd-forming giraffe (*Giraffa*). There was also a Plio-Pleistocene radiation of short-necked mooselike giraffids called sivatheres. *See* GIRAFFE.

The family Cervidae is a primarily Eurasian group, although cervids first reached North America in the Pliocene (5 million years ago), and later invaded South America to produce an extensive radiation

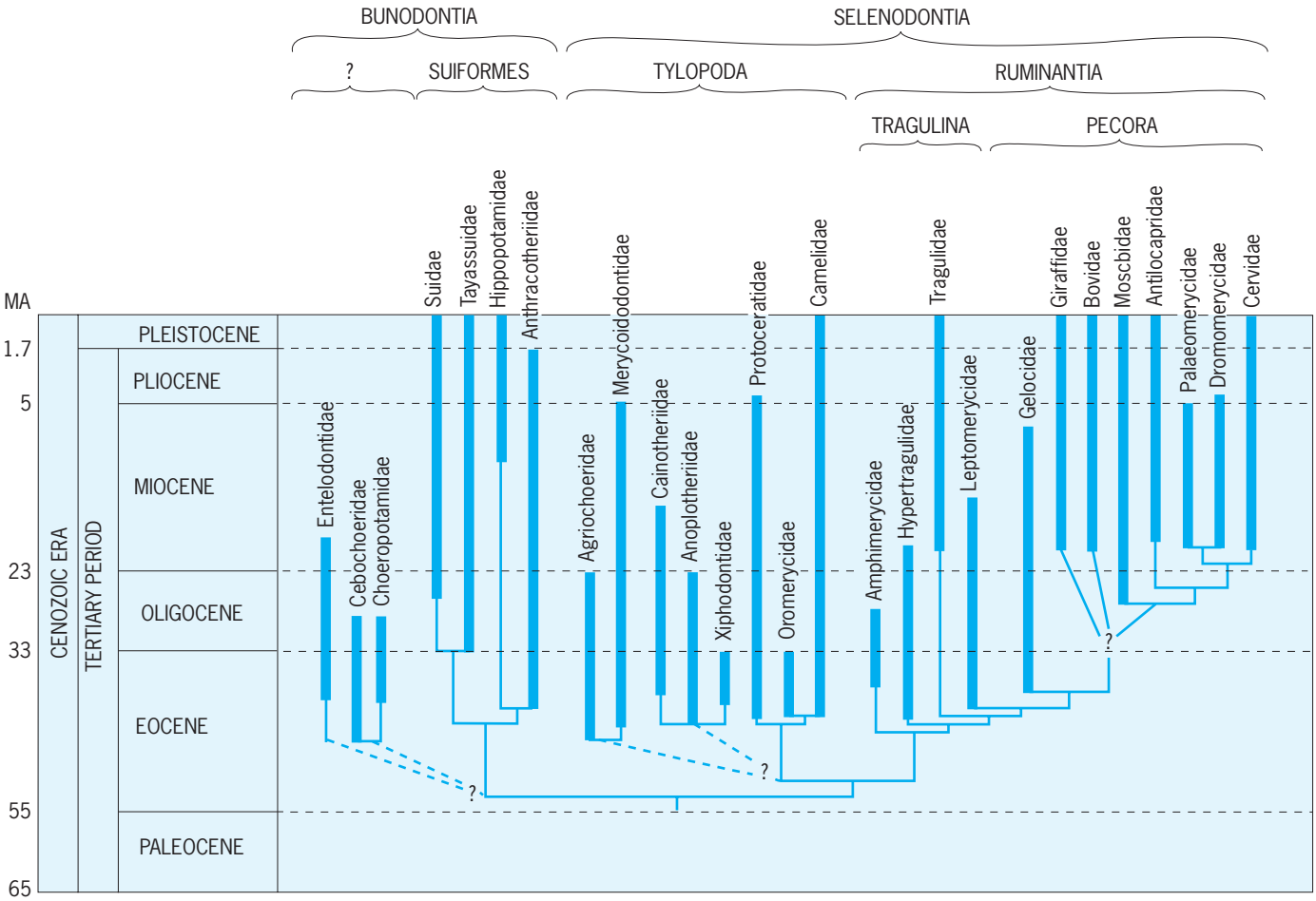

Fig. 5. Consensus of the interrelationships of the families of artiodactyls, including all living families and selected extinct ones (the basal Eocene radiation of dichobunids is not shown). Ma = million years ago.

of New World tropical and temperate forms. Most cervids possess antlers, a branched bony cranial appendage that is shed and regrown annually. Cervids that lack antlers, such as the Chinese water deer (*Hydropotes*), or that have only short antlers, such as the muntjac (*Muntiacus*), may have an enlarged upper canine in the males (Fig. 4*d*). Most cervids are browsers or mixed feeders, living in woodland or brushland, and forming small groups rather than permanent large herds (the reindeer or caribou is the one exception, and this deer is also the only one where antlers are present in the female). Extinct related groups, which were also apparently woodland browsers, include the Eurasian Palaeomerycidae and the North American Dromomerycidae. Some members of these groups sported a third horn on the occipital region as well as horns over the eyes (Fig. 4*e*). *See* DEER; MOOSE; REINDEER.

The family Antilocapridae, surviving only as the pronghorn "antelope," is an exclusively North American group. Antilocaprids are highly convergent on the antelope in anatomy and behavior, and all appear to have been savanna- or plains-dwelling mixed feeders, but they are probably more closely related to cervids than to bovids. The pronghorn combines a permanent bony horn core with a deciduous sheath that is shed annually (Fig. 4*f*). Extinct antilocaprids probably did not shed the sheath, or had horns that resembled nondeciduous antlers made from naked bone. *See* PRONGHORN.

The family Bovidae is today the most diverse of the artiodactyl families, containing cattle, sheep, goats, and antelope. Bovids are a primarily Old World radiation: they did not reach North America until the Pleistocene, and never reached South America (except as forms introduced by humans). Bovids vary greatly in their behavior and ecology, ranging from tragulid-like solitary tropical-forest selective browsers (such as duikers), to deerlike tropical-woodland browsers (such as the bushbuck and bongo), to savanna-dwelling herd-forming grazers (such as the wildebeest and buffalo), to cold or temperate-dwelling mixed feeders (such as sheep, goats, and musk-ox). Bovids are characterized by true horns, unbranched cranial appendages that consist of a bony horn core covered by a nondeciduous keratin sheath (Fig. 4*g*). The great radiation and diversification of bovids is a primarily recent (Plio-Pleistocene) phenomenon in association with the spread of African savannas. *See* ANTELOPE; BISON; BUFFALO; MUSK-OX. Christine Janis

Bibliography. G. A. Feldhamer, B. C. Thompson, and J. A. Chapman (eds.), *Wild Mammals of North*

America: Biology, Management, and Conservation, 2d ed., Johns Hopkins University Press, 2003; *Grzimek's Encyclopedia of Mammals*, vol. 5, McGraw-Hill, 1990; D. MacDonald, *The Encyclopedia of Mammals*, Andromeda Oxford Limited, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Asbestos

Any of six naturally occurring minerals characterized by being extremely fibrous (asbestiform), incombustible, and having high tensile strength. Historically they were utilized in commerce for fire protection; for fiber reinforcing material in tiles, plastics, and cements; for friction materials; and for thousands of other uses. Currently the vast majority of asbestos used worldwide is chrysotile type which is used for asbestos-cement production, while use continues to be strong only in friction products, coating and compounds, and roofing products. Because of great concern over the health effects of asbestos, many countries have promulgated strict regulations or bans on its use. The six minerals designated as asbestos also occur in a nonfibrous form. In addition, there are many other minerals which occasionally occur naturally with a highly fibrous morphology similar to asbestos. These ambiguities generated a controversy over whether the mineralogical definition of asbestos fibers should be applied when doing health hazard evaluations for legal and regulatory purposes. The U.S. Occupational Safety and Health Administration (OSHA) has reviewed this matter and concluded that the medical and scientific information available does not support an expansion of the definition of asbestos to include minerals not mineralogically defined as asbestos in the regulations.

Minerals. The six naturally occurring minerals exploited commercially for their desirable physical properties, which are in part derived from their asbestiform habit, are chrysotile asbestos—a member of the serpentine mineral group; and anthophyllite asbestos, grunerite asbestos (known historically by the commercial name amosite), riebeckite asbestos (known historically by the commercial name crocidolite), tremolite asbestos, and actinolite asbestos—all members of the amphibole mineral group. Populations of these mineral fibers, however processed, can be demonstrated to be asbestos if the length varies independently of the diameter.

Chrysotile asbestos. Chrysotile has the ideal chemical formula $Mg_3[Si_2O_5](OH)_4$ [Mg = magnesium, Si = silicon, O = oxygen, and OH = hydroxy]. Small amounts of aluminum (Al), iron (Fe), manganese (Mn), calcium (Ca), potassium (K), and sodium (Na) may appear in the bulk chemical analyses of this mineral (see **table**). The chrysotile fibers consist of long hollow tubes, and each fiber is approximately 25 nanometers in diameter, with lengths varying from well under 1 micrometer to well over 10 cm. The fibers are characteristically curly and ropelike. Chrysotile is identified by a combination of quantitative elemental composition; distinctive shape, size, and hollow-tube morphology of the fibers (**Fig. 1**); and characteristic x-ray or electron diffraction pattern. *See* SERPENTINITE.

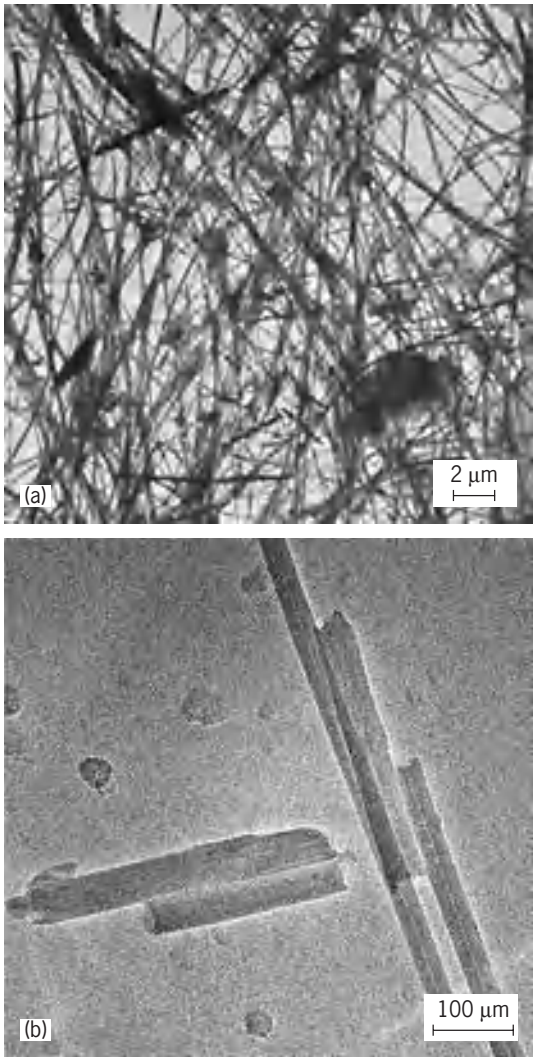

Fig. 1. Transmission electron photomicrographs of chrysotile asbestos from the Bazhenovskye Deposit in the Russian Federation. (*a*) At 1000 × magnification. (*b*) At 40,000 × magnification, where the hollow-tube morphology is visible.

Amphibole asbestos. The five regulated forms of amphibole asbestos and their ideal chemical formulas are grunerite asbestos, $(Fe^{2+},Mg)_7[Si_8O_{22}](OH)_2$; riebeckite asbestos, $Na_2Fe_2^{3+}(Fe^{2+},Mg)_3[Si_8O_{22}](OH)_2$; anthophyllite asbestos, $Mg_7[Si_8O_{22}](OH)_2$; tremolite asbestos, $Ca_2Mg_5[Si_8O_{22}](OH)_2$; and actinolite asbestos, $Ca_2(Mg,Fe^{2+})_5[Si_8O_{22}](OH)_2$. A considerable amount of chemical substitution takes place in these asbestos minerals, as indicated in representative analyses presented in the table. The amphibole asbestos minerals are characterized by their various elemental compositions and distinctive diffraction patterns, and generally vary in width

Anhydrous wt % of the major elements of representative specimens of the six regulated asbestos minerals

Element	Amosite (Penge, RSA*)	Anthophyllite (Paakila, Finland)	Chrysotile (Coalinga,† CA)	Crocidolite (Kuruman, RSA*)	Actinolite (Wrightwood, CA)	Tremolite (Korea)
Si	24.3	28.8	22.3	24.4	27.3	27.8
Fe	31.1	5.2	2.7	28.3	4.9	0.7
Mn	0.5	—	—	—	—	—
Mg	4.0	19.2	29.2	2.1	14.5	15.1
Ca	—	—	—	—	8.0	10.5
Na	—	—	—	4.3	—	—
O	40.1	46.9	45.7	40.8	45.3	45.9

*Republic of South Africa.
†Coalinga is part of Diablo Mountain Range.

from 0.1 to over 1 μm and in length from 1 μm to over 10 cm (**Fig. 2**). *See* AMPHIBOLE.

Geology. The major deposits of commercial chrysotile asbestos are found in massive serpentinite bodies formed by alteration of magnesium- and iron-rich igneous rocks such as dunite, peridotite, and iherzolite. The world's largest deposits of this type are found in Québec Province in Canada, in the Ural Mountains of the Russia Federation, in the Troodos Mountains of Cyprus, and in the Diablo Range of

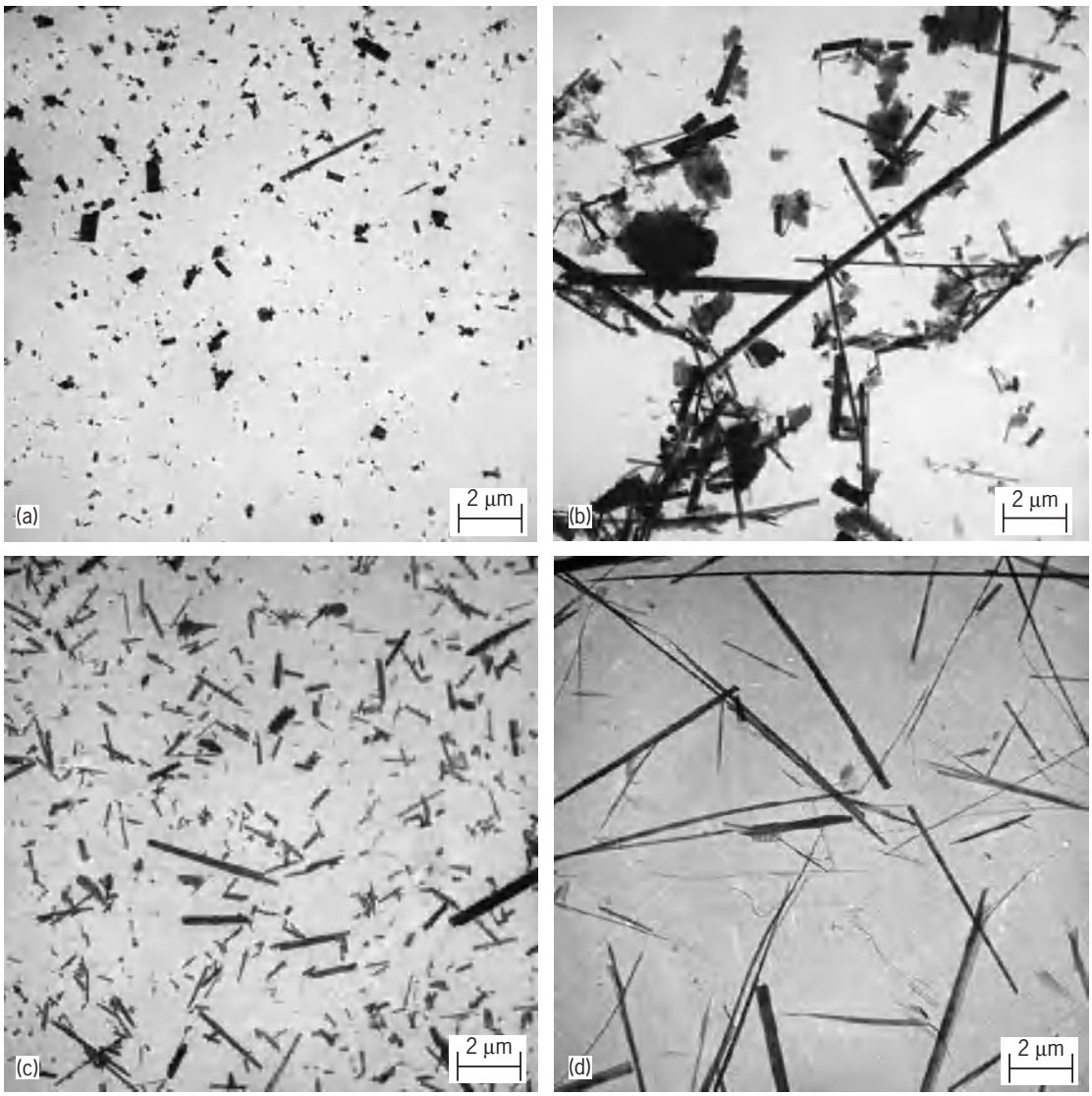

Fig. 2. Transmission electron photomicrographs of (*a*) amosite asbestos, Penge, Republic of South Africa; (*b*) anthophyllite asbestos, Paakila, Finland; (*c*) crocidolite asbestos, Koegas, Republic of South Africa; and (*d*) tremolite-actinolite asbestos, Jamestown, CA. Note that tremolite and actinolite are part of a solid solution series and differ only in iron content.

California. The asbestos in these serpentinite deposits is usually mined in large open pits (**Fig. 3**). Minor amounts of fibrous and nonfibrous anthophyllite, tremolite, and talc are also found in serpentinites. Serpentinized limestones also contain minor deposits of chrysotile asbestos. All the important commercial deposits of grunerite and riebeckite asbestos occur in Precambrian banded iron formations located in three areas—in the Cape and Transvaal provinces of the Republic of South Africa and in Western Australia. Anthophyllite asbestos, found within serpentinized periodotites of the Karelian Mountains, East Finland, was mined in modest commercial quantities from 1918 to 1975. Asbestos fibers crystallize from aqueous fluids that penetrate the open cracks, fissures, faults, and slip planes within tectonically stressed host rock. The fibers that form oriented perpendicular to the walls of the open spaces are known as cross fibers, while the fibers that form oriented parallel to a slip or fault plane are known as slip fibers. *See* DUNITE; PERIDOTITE; SERPENTINITE.

Fig. 3. Bazhenovskoye chrysotile asbestos deposit located in the Urals Region of the Russian Federation. The deposit was discovered by Alexey Ladizhensky in 1889.

Sources. Asbestos, mainly the chrysotile variety, is mined throughout the world. The two most important sources of chrysotile asbestos are the Urals Region of the Russian Federation and Québec Province in Canada. Lesser amounts of amosite and crocidolite asbestos have been mined in the Republic of South Africa, and anthophyllite in East Finland. Commercial production of tremolite and actinolite asbestos has been small and sporadic. The countries of the former Soviet Union are the major producers of chrysotile asbestos, followed by Canada, China, Brazil, Zimbabwe, and South Africa. The only producing asbestos mine in the United States is located within the New Idria serpentinite at the south end of the Diablo Range in California; ore containing 50–70% short-fiber chrysotile asbestos is excavated from an open pit and then trucked to King City, CA, for wet processing.

Uses. At present more than 99% of the asbestos used in the world is chrysotile, although small regional amphibole asbestos operations are thought to continue; however, the amphibole asbestos minerals essentially have left international commerce. Chrysotile asbestos continues to be used for many types of products because of its chemical and thermal stability, high tensile strength, flexibility, low electrical conductivity, and large surface area. Past uses of asbestos, such as sprayed-on insulation, where the fibers may become easily airborne, have been generally abandoned. Chrysotile asbestos is used predominantly for the construction industry in the form of cement sheets, coatings, pipes, and roofing products. Additional important uses are for reinforcing plastics and tiles, for friction materials, and for packings and gaskets.

Health effects. Historically, unaware of the dangers, workers in the asbestos trades and in mining and milling industries were excessively exposed to needlessly high levels of asbestos, which caused a very high incidence of morbidity and mortality. The three principal diseases associated with exposure to the asbestos minerals are lung cancer; mesothelioma, a rare cancer of the pleural and peritoneal membranes that enclose the chest and abdominal cavities; and asbestosis, a nonmalignant disease characterized by a diffuse interstitial fibrosis of the lung, which causes the lung tissue to become stiff and exchange oxygen poorly. Excessive exposure to all the asbestos fiber types is associated with asbestosis and increased risk of lung cancer. Mesothelioma, a rare tumor accounting for approximately 1 in 10,000 deaths in the general population, can be dramatically increased by exposure to amosite, crocidolite, or tremolite asbestos. These last two fiber types are strongly associated with an increased incidence of nonoccupational mesothelioma and therefore are thought to present a risk at rather low exposures. As the current commercial use of amphibole asbestos is very limited, the risk of environmental mesothelioma associated with naturally occurring asbestos outcroppings, particularly tremolite asbestos, has become of concern. The ingestion of asbestos has not been proven to cause disease in either humans or animals. Workers employed in the friction materials, cement, and roofing materials industries who were exposed only to chrysotile asbestos, generally at levels below two fibers per cubic centimeter, do not appear to have an excess of either malignant or nonmalignant disease. *See* MUTAGENS AND CARCINOGENS; ONCOLOGY; RESPIRATORY SYSTEM DISORDERS.

Malcolm Ross; Robert P. Nolan

Bibliography. *Asbestos in Public and Commercial Buildings*, Health Effects Institute, 1991; B. Levadie (ed.), *Definitions for Asbestos and Other Health Related Silicates*, American Society for Testing and Materials, STP 834, 1984; G. D. Guthrie and B. T. Mossman (eds.), *Health Effects of Mineral Dusts*, Mineralogical Society of America, 1993; R. P. Nolan et al. (eds.), *Health Effects of Chrysotile-Asbestos: Contribution of Science to Risk Management Decisions*, Can. Mineralog. Spec. Pub., 2000; H. C. W. Skinner, M. Ross, and C. Frondel, *Asbestos and Other Fibrous Materials*, 1989; R. L. Virta, *Minerals Yearbook*, vol. 1: *Metals and Minerals* (*Asbestos*), U.S. Bureau of Mines, 1989.

Ascaridida

An order of nematodes in which the oral opening is generally surrounded by three or six labia; in some taxa labia are absent, but the cephalic sensilla are always evident. Usually there are eight cephalic or labial sensilla; the submedians may be fused and then only four sensilla are seen. The stoma varies from being completely reduced to spacious or globose. The esophagus varies from club shaped to nearly cylindrical, never rhabditoid. There may be posterior esophageal or anterior intestinal ceca. The collecting tubules of the excretory system may extend posteriorly and anteriorly. Males generally have two spicules; however, in some taxa there may be none or only one. The gubernaculum may also be present or absent. Though females generally have two ovaries, multiple ovaries do occur. The number of uteri is also variable: two, three, four, or six. Phasmids are sometimes large and pocketlike. Reportedly, the larvae lack a stomatal hook or barb.

The order probably comprises seven superfamilies: Ascaridoidea, Seuratoidea, Camallanoidea, Dracunculoidea, Subuluroidea, Dioctophymatoidea, and Muspiceoidea (*incertae sedis*).

Ascaridoidea. This superfamily of about 65 genera comprises large parasitic roundworms, whose adult stages usually occur in the stomach or small intestine of terrestrial and aquatic mammals, birds, reptiles, and fishes (see **table**); the parasitic larval stages of many species occur, either temporarily or indefinitely, in other parts of the host's body.

General morphology. These roundworms are mostly medium to large and thick bodied, with three large lips, interlabia, and other cuticular specializations of the head. Eight incompletely fused submedian cephalic papillae are located on the external circle, and a lateroventral pair is present. The buccal capsule is usually more or less rudimentary. The esophagus is more or less cylindroid, is sometimes wholly muscular, and sometimes terminates posteriorly in a glandular ventriculus with or without one or more appendices. The intestine sometimes has one or more anterior ceca. Females are oviparous, sometimes with more than two uteri; the eggs are generally more or less spherical and thick shelled. Males have two spicules, but rarely caudal alae or a precloacal sucker.

Life cycles. Many species have a direct life cycle. Others, mainly species with marine mammals, birds, and fishes as definitive hosts, require an intermediate host, such as a fish, amphibian, insect, crustacean, or small mammal. Distribution is cosmopolitan, and these worms are the causative agents of ascariasis, or ascaridosis, occurring in many agricultural animals (and occasionally in humans). Infestation is typically characterized by pulmonary damage and distress initially, and digestive disturbances later. Damage may also occur during larval migration to other parts of the body, including the liver and brain.

Common ascarids. Ascaris lumbricoides, the large roundworm of humans, has a life cycle similar to that of *A. suum*. Infestations are very common in tropical and temperate regions where sanitation is poor or lacking. These worms often are as big as an ordinary pencil.

Ascaris suum, the pig ascarid, usually is considered distinct from *A. lumbricoides*, on certain biological criteria. The eggs of this worm pass with feces onto the ground, where they develop until each contains an infective larva, capable of hatching in the gut of another pig. The liberated larvae penetrate into the intestinal wall, enter the bloodstream, break out of capillaries into the air spaces of the lungs (causing "thumps," or ascaroid pneumonia), and via the trachea reenter the gut where they develop into adults. In a mouse, the same larval migration occurs, but the rest of the cycle is not completed.

Lung damage from the larval invasion may lead to a fatal pneumonia in young pigs. Reportedly, hog cholera virus becomes pathogenic under this pulmonary stress. The virus is innocuously present in nematodes of the genus *Metastrongylus* that commonly infest the lungs of pigs. Intestinal infestations are debilitating and cause economic loss.

Heavy infestations of *Ascaridia galli*, the large intestinal roundworm of chickens, can cause intestinal blockage and death; lesser infestations stunt growth and interfere with good bone development. The life cycle resembles that of *Ascaris suum*, but with limited migration of the larvae into and out of the intestinal wall. Turkeys also are subject to infestation.

Toxacara canis, one of the common ascarids of dogs, is very injurious to young puppies and can be lethal in heavy infestations. Its life cycle is like that of *Ascaris suum*, but infection also occurs prenatally, and by ingestion of rats with encapsulated *T. canis* larvae in the tissues.

The exposure of humans to ascarids of lower animals, mainly those associated with dogs and cats, can result in a very serious condition known as visceral larva migrans. As with normal human *Ascaris*, migration to ectopic foci often results in serious and sometimes fatal conditions. In human tissues and organs such as the liver, lungs, brain, eye, viscera, or muscles, the larvae become encapsulated by granulomatous tissue. Prognosis varies, but in the majority of cases recovery is complete. However, blindness, paralysis, and death are not uncommon.

Dracunculoidea. This superfamily of parasitic nematodes comprises obligate tissue parasites of fishes, reptiles, and mammals. All known species require an intermediate host in order to complete their life cycle, and that host is always a water flea (*Cyclops*). Most often the stoma is reduced and vestibular. In mature females, the vulva is equatorial and atrophied; egg and larval deposition often results from the rupturing of the female. The most widely known example is *Dracunculus medinensis*, the guinea worm. Knowledge of this worm and its lesion extends back to antiquity.

Ingestion of water containing infective *Cyclops* is the only known source of infection. The encysted nematode larvae are released from *Cyclops* by the digestive juices of the duodenum. Then the larvae burrow through the intestinal wall, and upon reaching

Some Ascaridoidea of domesticated mammals and birds in the United States*

Name	Host†	Remarks
Toxocara cati	Cat	Occasional parasite of humans; *Toxocara* larvae can cause human visceral larva migrans; cats can become infected by eating mice with the larvae in their tissues
Toxascaris leonina	Cat, dog	Occasional parasite of humans; normally develops without passage of larvae through the lungs
Neoascaris vitulorum	Cattle	Not generally distributed; infestations are injurious to calves; infection occurs prenatally, perhaps exclusively
Parascaris equorum	Horse	Larval migration of tracheal type damages liver and lungs; heavy infestations are sometimes fatal to foals
Ascaridia columbae	Pigeon	Heavy infestations can cause anemia, droopiness, and diarrhea
Ascaridia dissimilis	Turkey	Exact information on pathogenicity is not available

*All have a direct life cycle.
†Some of the species also occur in other hosts.

the loose connective tissue, usually retroperitoneal in position, they develop to adulthood in 8 months to 1 year.

The gravid females, 28–48 in. (70–120 cm) long, migrate from the site of development to the subcutaneous tissues. At that time the ovaries are postfunctional, and the uteri are highly coiled, distended, and filled with masses of rhabditoid larvae. The females make their way to the surface of the skin, and a papule is formed. The papule continues to develop and a blister forms, usually on the lower extremities. When the blister comes in contact with fresh water, the uterus bursts through the anterior part of the nematode's body, and the worm also bursts, releasing cloudlike swarms of motile larvae. These larvae are then filtered from the water by *Cyclops* and subsequently ingested. Development to the infective stage within *Cyclops* takes about 10 days.

The formation of the blister and subsequent rupturing of the female produce a profound allergic reaction. This reaction results from the release of large amounts of toxic by-products from the worm. Upon discharge of larvae in fresh water, much of the allergic reaction abates. The reactions and systemic prodromes include erythema, urticarial rash, pruritus, vomiting, diarrhea, and giddiness. Septicemia, suppurating cysts, and chronic abscesses are not uncommonly associated with these infections. The worms can be removed surgically, or in the native manner of winding upon a stick. Chemotherapy is also available.

Control in endemic areas includes keeping infected persons from wading or bathing in water used for drinking purposes, and the education to avoid drinking suspect water. *See* NEMATA.

Armand R. Maggenti

Bibliography. R. Muller, *Dracunculus* and dracunculiasis, in *Advances in Parasitology*, vol. 9, 1971; F. A. Neva and H. W. Brown, *Basic Clinical Parasitology*, 6th ed., 1996.

Ascidiacea

A class of Tunicata which occurs as solitary zooids or, by a process of asexual budding, develops into colonies. Zooids vary in length from a few millimeters to 10 in. (25 cm). Individuals or colonies are invested by a protective covering, the tunic or test, made of polysaccharide material structurally close to cellulose. Beneath the test is the body wall or mantle. Each zooid has two apertures: inhalant (oral) and exhalant (atrial; see **illus.**). Water currents, created by cilia on the margins of stigmata in the pharyngeal wall, draw water into the branchial sac, where it is filtered and passed out through the exhalant aperture. The entrance to the branchial sac is guarded by tentacles which prevent large particles from entering. Filtration takes place on a mucous sheet secreted by the endostyle; the sheet is passed across the internal face of the branchial sac by ciliary or muscular action and is then rolled into a cord by the dorsal lamina (or by languets). Digestive enzymes are secreted into the stomach, and a pyloric gland, of unknown function, enters at the junction of stomach and intestine. Gonads are hermaphroditic, and may be situated in the loop of the intestine or in the mantle wall.

The tadpole larva does not feed and settles and metamorphoses within a few hours; the tail tissues are resorbed at this time into the trunk and provide

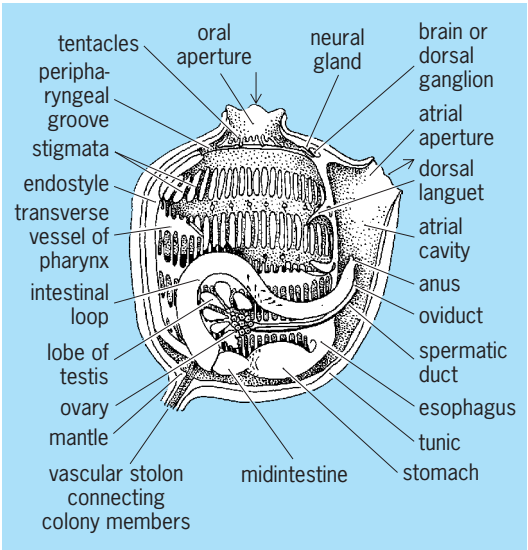

Left side of a zooid of the colonial ascidian *Perophora*. The tunic, mantle, and anterior wall of the pharynx have been removed from the left. The two arrows indicate direction of water intake and expulsion.

a nutrient reserve for further development of the adult. Asexual reproduction giving rise to colonies may occur by budding from stolons, by budding from tissues of the mantle wall, or by dedifferentiation of zooids, which then constrict into a number of buds.

Common examples of solitary ascidians are *Ciona*, *Molgula*, and *Ascidia*; examples of colonial forms are *Botryllus*, *Perophora*, and *Amaroucium*. Three orders of ascidian are recognized, Aplousobranchia, Phlebobranchia, and Stolidobranchia, differing from one another by characters of the brancial sac and position of the gonads. Ascidians occur throughout the seas of the world and at all depths, including the abyssal region. Most species feed on minute particulate matter, but a few are carnivores and engulf small zooplankters. *See* TUNICATA. Ivan Goodbody

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; F. S. Russell and C. M. Yonge (eds.), *Advances in Marine Biology*, vol. 9, 1971, vol. 12, 1974; W. G. Van Name, The North and South American ascidians, *Bull. Amer. Mus. Nat. Hist.*, vol. 84, 1945.

Ascomycota

A phylum in the kingdom Fungi, representing the largest of the major groups of fungi, and distinguished by the presence of the ascus, a specialized saclike cell in which fusion of nuclei and reduction division occur. The resulting nuclei are delimited by membranes through free-cell formation and a wall to form ascospores inside the ascus. In most ascomycetes, each ascus contains eight ascospores, but the number may vary from one to several hundred.

In the simplest ascomycetes (yeasts), the vegetative body (thallus) is unicellular; however, in the majority of ascomycetes, the thallus is more complex and consists of a tubular, threadlike hypha with cross walls which grows in or on the substrate. These hyphae eventually form structures called ascomata (ascocarps), on or in which the asci are formed. In addition to their sexual reproduction, most ascomycetes reproduce asexually by means of conidia. Traditionally, the structure of the ascoma and ascus has served as the basis for subdividing the Ascomycota into five classes: Hemiascomycetes, Plectomycetes, Pyrenomycetes, Discomycetes, Loculoascomycetes. The recent introduction of molecular data, however, is changing concepts of the relationships of different groups of ascomycetes and will eventually lead to a much-revised classification scheme. *See* DISCOMYCETES; HYMENOMYCETES; LOCULOASCOMYCETES; PLECTOMYCETES.

As a group, the ascomycetes occur throughout the world in all types of habitats and on both living and dead substrates. An estimated 33,000 species of ascomycetes are arranged in about 3300 genera, with new species being described regularly. Ecologically ascomycetes function as primary decomposers of plant materials, but they also are important as plant and human pathogens; in baking, brewing, and winemaking; in enzyme and acid production; and as sources of antibiotics and other drugs. *See* EUMYCOTA; FUNGI; PLANT PATHOLOGY; YEAST. Richard T. Hanlin

Bibliography. C. J. Alexopoulos, C. W. Mims, and M. Blackwell, *Introductory Mycology*, 4th ed., John Wiley, New York, 1996; J. Breitenbach and F. Kränzlin, *Fungi of Switzerland*, vol. 1: *Ascomycetes*, Edition Mycologia, Lucerne, 1984; R. W. G. Dennis, *British Ascomycetes*, rev. ed., J. Cramer, Vaduz, 1981; R. T. Hanlin, *Illustrated Genera of Ascomycetes*, vols. I and II, APS Press, St. Paul, 1990, 1998.

Ascorbic acid

A vitamin also known as vitamin C. It is a white crystalline compound, highly soluble in water, a stronger reducing agent than the hexose sugars, which it resembles chemically. Vitamin C deficiency in humans has been known for centuries as scurvy. The compound has the structural formula shown below.

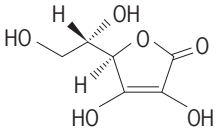

L-Ascorbic acid

The stability of ascorbic acid decreases with increases in temperature and pH. This destruction by oxidation is a serious problem in that a considerable quantity of the vitamin C content of foods is lost during processing, storage, and preparation. Biological methods for estimating ascorbic acid are rarely used. The vitamin is determined chemically by making use of its reducing properties.

While vitamin C is widespread in plant materials, it is found sparingly in animal tissues. Of all the animals studied, only the guinea pig, the red vented bulbul bird, the fruit-eating bat, and the primates, including humans, require a dietary source of vitamin C. The other species studied are capable of synthesizing the vitamin in such tissues as liver and kidneys. Some drugs, particularly the terpenelike cyclic ketones, stimulate the production of ascorbic acid by rat tissues.

Vitamin C–deficient animals suffer from defects in their mesenchymal tissues. Their ability to manufacture collagen, dentine, and osteoid, the intercellular cement substances, is impaired. This may be related to a role of ascorbic acid in the formation of hydroxyproline, an amino acid found in structural proteins, particularly collagen. People with scurvy lose weight and are easily fatigued. Their bones are fragile, and their joints sore and swollen. Their gums are swollen and bloody, and in advanced stages their teeth fall out. They also develop internal and subcutaneous hemorrhages.

The biochemical role of ascorbic acid is obscure. It seems reasonable to expect that it functions metabolically in oxidation-reduction systems, since it has been shown that ascorbic acid and dehydroascorbic

acid are readily interconverted in plant and animal tissues. It may also act as an antioxidant, protecting hydrogen carriers from destructive oxidation. Ascorbic acid has a role in tyrosine metabolism. It also appears to function in the conversion of folic to folinic acid. *See* FOLIC ACID.

There has been great difficulty in establishing the human requirements for vitamin C. Usually, vitamin requirements are based on data obtained from dietary surveys accompanied by blood or urine analyses and often by saturation tests. There is evidence that vitamin C may play roles in stress reactions, in infectious disease, or in wound healing. Therefore, many nutritionists believe that the human intake of ascorbic acid should be many times more than that level which produces deficiency symptoms.

The recommended dietary allowances of the Food and Nutrition Board of the National Research Council are 30 mg per day for 1–3-month infants, 80 mg per day for growing boys and girls, and 100 mg per day for pregnant and lactating women. These values represent an intake which tends to maintain tissue and plasma concentrations in a range similar to that of other well-nourished species of animals. The Accessory Food Factors Subcommittee of the British Medical Research Council does not believe that a dietary intake of more than 30 mg per day has beneficial effects. *See* VITAMIN.

In most processes for the large-scale synthesis of ascorbic acid, technical dextrose is hydrogenated to sorbitol. Biological oxidation converts sorbitol to sorbose. The next step is diacetone sorbose, which is oxidized to diacetone gulosonic acid. Diacetone gulosonic acid is hydrolyzed to gulosonic acid. The methyl ester of the latter is reacted with sodium methylate to form the sodium salt of ascorbic acid. Stanley N. Gershoff; Werner A. Lindenmaier

Bibliography. S. N. Gershoff and W. A. Lindenmaier, Ascorbic acid, in AccessScience@McGraw-Hill, http://www.accesscience.com, DOI 10.1036/1097-8542.054400.

Ascothoracica

An order of the subclass Cirripedia. Members are ecto- or endoparasites of coelenterates and echinoderms. The body is enclosed in a voluminous saclike mantle up to 0.8 in. (20 mm) long, whereas the body is only about one-fourth of this length (see **illus.**). Ascothoracica are not attached permanently to the host by the antennular region, but the antennules may be modified as a clasping organ. The mouthparts are also modified as a clasping organ and for piercing and sucking. Up to six pairs of thoracic appendages are present, reduced in number and development in the endoparasites. Unlike all other Cirripedia, Ascothoracica retain a more or less segmented abdomen in the adult. There are no cement glands. Diverticula of the alimentary canal extend into the mantle.

The sexes are separate, males being smaller than females. Larval development takes place mainly in the parental mantle. The nauplii have no frontal horns. There are said to be two cypris stages, the first being passed in the mantle, the second free-swimming and most probably the dispersal and infective stage.

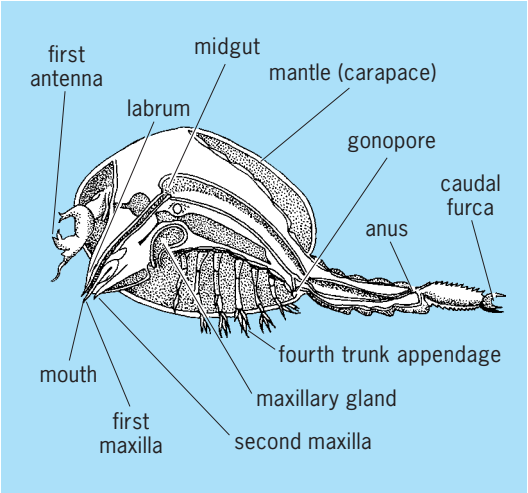

***Ascothorax ophioctenis*, a parasite in the bursae of brittle stars. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., Saunders, 1968*)**

The virtually motile existence of individuals, the possession of an abdomen, and discrepancies in the larval cycle have prompted some writers to separate the Ascothoracica from the Cirripedia. The presence of a cypris stage indicates a relationship nearer to the Cirripedia than to any other subclass of the Crustacea. *See* CIRRIPEDIA. H. G. Stubbings

Ash

A genus, *Fraxinus*, of deciduous trees of the olive family Oleaceae, order Scrophulariales, which have opposite, pinnate leaflets, except in one species, *F. anomala*, which has only a single leaflet. There are about 65 species in the Northern Hemisphere. This tree occurs in America south to Mexico, in Asia south to Java, and in Europe. *See* SCROPHULARIALES.

The white ash (*F. americana*), of the eastern United States, is the most important species, and attains a height of 120 ft (36 m). This species has stalked leaflets, rusty-colored winter buds, and an erect trunk that is valuable for lumber. The wood is light, strong, and flexible, and is used for oars, baseball bats, furniture, motor vehicle parts, boxes, baskets, and crates. The black ash (*F. nigra*; **illus.** *a*) grows 75 ft (23 m) in height in wet soils in the northeastern United States and Canada and has sessile leaflets and friable outer bark. The wood of black ash is used for the same purposes as that of white ash. The red ash (*F. pennsylvania*), also in the eastern United States and adjacent Canada, grows to 60 ft (18 m) and has pubescent (hairy) twigs and leaf stalks. The uses of the wood of this species are also similar to those of white ash. Some species of ash are ornamental trees, such as the flowering ash (*F. ornus*) with gray winter buds and white flowers, and the

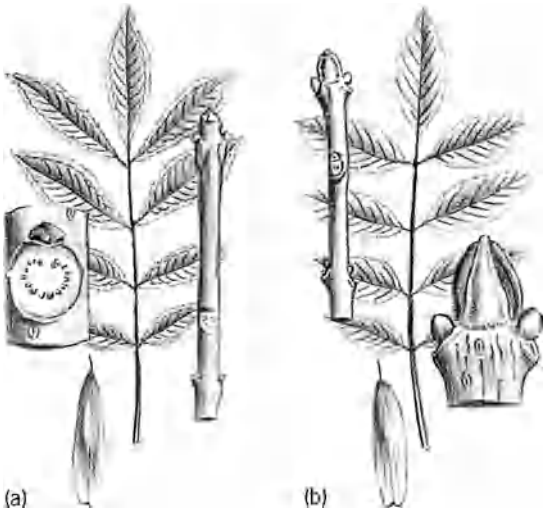

Two species of ash. (*a*) Black ash (*Fraxinus nigra*). (*b*) European ash (*F. excelsior*).

European ash (*F. excelsior*; illus. *b*) with black buds and sessile leaflets. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Asia

The largest of the world's continents. With its peninsular extension, commonly called the continent of Europe, it is the major portion of the broad east-west extent of the Northern Hemisphere landmasses. In many ways Asia is more a cultural concept than a physical entity. There is no logical physical separation between Asia and Europe, and even Africa is separated from Asia merely by the width of the Suez Canal. For convenience, however, the Eurasian landmass is considered to be divided by the Ural Mountains into Europe in the west and Asia in the east. Thus restricted Asia has an area of about 17,700,000 mi^2 (45,800,000 km^2), about one-third of the land area of the Earth. In the north, mainland Siberia reaches past the 77th parallel. Southward, India and Sri Lanka (Ceylon) reach nearer than 10° north of the Equator, while the Indonesian islands extend more than 10° south of the Equator. The continental heart of Asia is more than 2000 mi (3200 km) from the nearest ocean. These vast distances have great significance in the climates and vegetation of Asia and thus in the erosional and depositional patterns of the land (**Fig. 1**). *See* CONTINENT; EUROPE.

As the world's largest continent, Asia has been shaped by the processes of cratonization, geosynclinal development, and associated orogenesis. By accretion several major stable structural units were formed, and these collectively form the basement for much of the continent (Fig. 1). The three northern cratons of Asia—the Russian, the Siberian or Anggaran, and the Sino-Korean—were evidently stabilized as early as 1800 or 1700 million years ago (Ma), whereas the more southerly ones—the Chang Jiang, the Tarim, and the Arabian—appear to have stabilized and been consolidated into the Chinese protoplatform much later (around 700–500 Ma). The Indian craton, though formed early, has been in collision with the Eurasian platform since about 40 Ma.

Geosynclinal development occurred in the central Asian region on continental margins. As a consequence, sedimentary deposits now blanket much of the basement, and these sediments have been involved in a series of orogenic episodes. In the early Paleozoic (500 Ma) there were several areas of activity—Kunlun, Qinling, Beishan, and Tianshan—and the closing of this system merged the Tarim platform with the Eurasian platform. In the Mesozoic (200 Ma), disintegration and geosynclinal development occurred on the southern flank (the Tethys-Himalayan Domain, in what is now central-northern Tibet), and in the Marginal Pacific Tectonic Domain, covering present-day eastern China, Korea, eastern Russia, and the adjacent marine margins. From these events arose the major structural alignments which persist to the present, running west-east throughout central Asia, and southwest-northeast through the Pacific margins of the continent. The collision of the Indian craton with the Eurasian landmass added the predominantly west-east alignment of the Himalayan system, as well as forcing the Tibetan plateau uplift, and the reactivation of preexisting west-east mountain structures in central Asia.

It has been postulated that the Marginal Pacific Tectonic Domain of Asia may have been in the Paleozoic, an Atlantic-type margin (tensional, stable, and relatively quiescent) but that in the Mesozoic it was converted to a Pacific type, with strong compression, and a highly active Benioff zone situated to the east of the island (and quasi-island) arcs—Kamchatka, Japan, Taiwan, and the Philippines. It has also been postulated that southeast Asia was accreted as a series of relatively small fragments detached from the Australian landmass and propelled northwestward until they became part of the Asian landmass.

Tectonic and geosynclinal processes continue at the present time. Volcanism and seismic activity are widespread in the island arcs and indeed on parts of the Asian mainland affected by geosynclinal activity. In the marine basins and the lowlands behind the Pacific Rim arcs, sedimentation is proceeding actively. The convergence of India with Tibet sustains the mountain ranges of the Himalayas, with elevations of over 25,000 ft (7600 m) as well as the 16,400-ft (5000-m) plateau to the north. In the same manner the Arabian plate in collision with Iran sustains elevations in excess of 13,000 ft (4000 m). *See* OROGENY; PLATE TECTONICS.

Mountain chains. Asia's great mountain chains are related to the tectonic structures. The Pamir knot marks the northern penetration of the Indian craton, while the ranges encircling Assam in northeast India represent its northeasterly wing. Elevations in the vicinity of the Pamirs reach 24,600 ft (7500 m); ranges then radiate in several directions. The Hindu Kush range curves southwestward into Afghanistan.

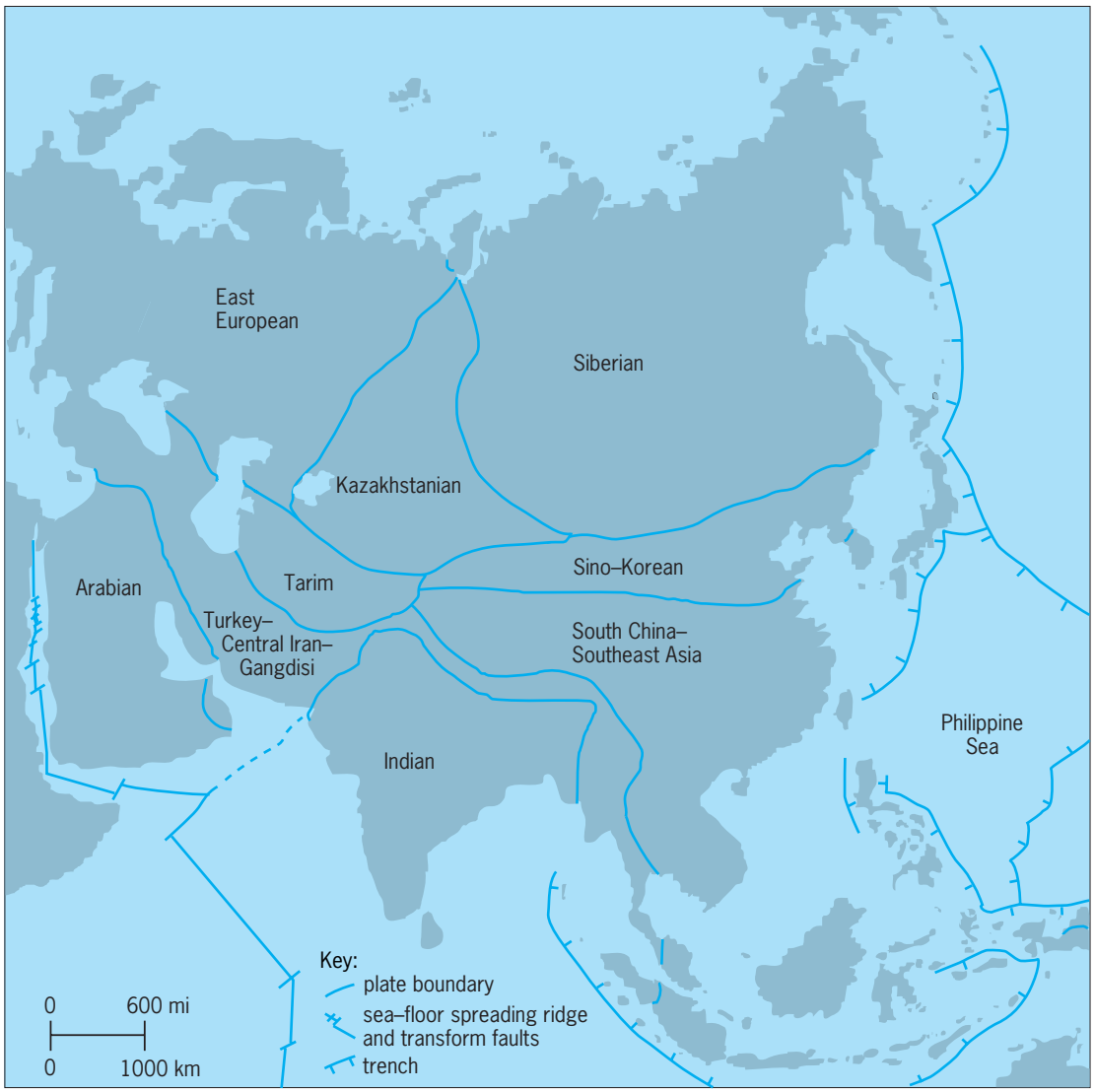

Fig. 1. Major cratons that have been involved in the Eurasian tectonic plate. (*After The Research Institute of Geology, Chinese Academy of Geological Sciences, Scheme showing the subdivisions of paleo and present plates in Asia, Cartographic Publishing House, Shingxa, 1982*)

The mountain trendline continues, after a jog northwestward, in the Elburz of northern Iran and thence in the Armenian highlands and the Caucasus, each with elevations reaching 18,000 ft (5500 m), decreasing thereafter to the Pontus and Taurus ranges of northern and southern Turkey. In western and southern Iran are the massive Zagros and Makran ranges.

Southeastward from the Pamir knot run the three most imposing mountain chains on Earth: the Karakorum, which continues the line of the Hindu Kush eastward in an arc convex to the north; the Himalaya in an arc convex to the south; and the shorter Trans-Himalaya, north of the Himalaya, with higher average elevations but peaks of lesser height. In all of these, the average elevations exceed 4 mi (6400 m), with several scores of peaks reaching a height in excess of 25,000 ft (7600 m) above sea level. Everest, 29,141 ft (8882 m), and Kinchinjunga, 28,146 ft (8579 m), lie in the Himalaya, while the peak designated as K2, 28,250 ft (8611 m), rises in the Karakorum. The few passes range between 14,000 and 15,000 ft (4300 and 4600 m), although the Karakorum Pass lies at 18,317 ft (5583 m).

In eastern Tibet the Himalaya and Nyanqen-tanglha curve sharply around the northeast extremity of the Indian craton, the Himalaya being cut by the great gorge of the Brahmaputra River. From the bend zone, great ridges divided by deep gorges run south to form the Burma-China frontiers and the mountain backbones of the Malay peninsula and Vietnam. The Nan Ling system of southern China diverges eastward to divide the Chang Jiang from the Xi drainage.

From the western Himalaya, the 11,000-ft (3400-m) Sulaiman Range runs south and, together with the Kirthar Range, divides West Pakistan from Afghanistan.

Branching off from the Karakorum south of Kashgar, the Kunlun mountains run eastward across China. In the west they have been elevated along

with the uplifted Tibetan plateau, and reach elevations of 21,000 ft (6100 m). In the east, lacking the elevated Tibetan base, they nonetheless reach altitudes of 12,000 ft (3700 m) in the Qin Ling and lesser Dabie mountains. In eastern China they reach almost to the Pacific coast. Together with the northeastward arc of the Altyn Tagh and the Nan Shan branching from it, the Kunlun forms the northern wall of the Tibetan plateau. Near the eastern end of the Kunlun proper lie the A'nyemaqen Mountains, with peaks up to 25,000 ft (7600 m).

Trending east and west is a series of great ranges to its north, with mutual connections in the west. These include the Altai-Sayan, the Tannu Ola, and the Hentyin which form natural boundaries for Outer Mongolia. They continue the systems of young mountains crossing central Asia; farther northeast, they extend further in the Stanovo Mountains of eastern Siberia.

Hill lands and plateaus. Sedimentation over the ancient craton surfaces as well as in the intercraton zones has in some instances been followed by deformation, erosion, and the formation of hill regions. In other cases, where deformation has been minor, plateau surfaces have resulted. Most of southern China, which lies upon the Chang Jiang craton, consists of hill areas with local relief of up to 1640 ft (500 m), though absolute elevations above sea level are considerable. The elevated western and eastern edges of the southern portions of the Indian craton are marked by the Western and Eastern Ghats, and in the south the Nilgiri Hills rise to elevations of 8000 ft (2400 m). The Cambodian craton in its eastern regions is buried under younger geological materials that make up the Laotian and Vietnamese hill zones; and in the south it is overlain by the Cardamom Mountains.

The Tibetan plateau is maintained at its overall elevation of around 16,000 ft (4900 m) by the constant and continuing motion of the colliding Indian plate. Though much of the surface of the Tibetan plateau is flat, mountain ranges traverse it, and deep river gorges have been cut into it.

Other plateau regions represent areas in which sedimentary materials laid down on the craton surfaces have been left relatively undisturbed. In eastern Asia lies the subhumid loess plateau of northern China, the humid Yunnan and Guizhou plateaus of southern China, and the Kaima plateau of northern Korea. Much of the Siberian craton is covered by the Siberian plateau. In southern Asia most of peninsular India is a plateau block on the Indian craton, tilted from west to east, and covered in its northeasterly portions by recent volcanic lava. Saudi Arabia is likewise a tilted plateau block, with sand and other materials overlying the basement of the Arabian plate.

In southeastern Asia the Shan plateau is an undulating surface, deeply incised by the Salween river in the east. In contrast, the Korat plateau in eastern Thailand represents sedimentary rock formations very little disturbed since they were deposited on the craton surface in Mesozoic times.

Lowland plains. The most significant topographic units of Asia are the great alluvial plains and river deltas; in some of these live the vast majority of the world's population. The gross drainage pattern of Asia is radial; the rivers flow from the highlands in the heart of the continent and run outward in all directions. Only in the south, east, and north sectors of the continent do the rivers reach the sea. Flowing into the peripheral seas of the Pacific are such mighty rivers as the Mekong, the Xi, the Chang Jiang, the Huai, the Yellow, and the Amur, each building large, heavily populated plains and, with the exception of the Amur, densely settled deltas. The Yellow Plain (North China Plain) with some 125,000 mi^2 (324,000 km^2) of area, and the Chang Jiang Plain, with about 75,000 mi^2 (194,000 km^2), are among the most extensive alluvial plains of the Earth. In the shallow South China, East China, and Yellow seas, the deltas of the Mekong, Xi, Chang Jiang, Huai, and Yellow rivers are pushing steadily seaward. *See* DELTA; FLOODPLAIN.

North of the Amur, all the great Asian rivers drain across Siberia into the Arctic, with plains supporting only boreal forests and tundra. Westward, the radial rivers end in internal drainage basins and saltwater bodies such as Lake Balkash and the Aral Sea. These include the Ili, flowing into Lake Balkash from Chinese Xinjiang, and the Amu and Syr Darya, running into the Aral Sea. On the frontier of Iran and Afghanistan lies the seasonal Lake Helmand with the Helmand River emptying into it. The long Tarim River flows in the desert basin of southern Xinjiang until it dries up in its easterly course or empties into shifting salt lakes such as Lop Nur. *See* BASIN; PLAINS.

From the glaciers of eastern Tibet and the Himalaya run meltwaters to which are added the heavy summer monsoon rains of southern and southeastern Asia. These waters collect in such great streams as the Indus and Ganges-Brahmaputra of India, and the Irriwaddy and Salween of Burma, all emptying into the Indian Ocean, and all except the Salween have constructed great delta plains which support some of the densest populations of southern and southeastern Asia. India's east coasts have such great rivers as the Kistna, Godavari, and Cauveri; on the west coast are the Narbada and the Tapti. Most streams of the Indian peninsula are dry in winter.

Most rivers in southwestern Asia also are intermittent seasonal streams. The only large rivers here are the Tigris-Euphrates streams, whose sediments have formed the Mesopotamian plain extending inland from the head of the Persian Gulf. Their waters, however, are derived neither from the Arabian deserts nor the dry Mesopotamian region, but from the uplands of Armenia and southern Turkey. Smaller streams flow out of northern and western Turkey into the Black, Aegean, and Mediterranean seas. In Palestine the small Jordan flows down a rift valley to the Dead Sea, 1292 ft (394 m) below sea level. The only other such low-lying surface of Asia is the Turfan depression of Xinjiang, 928 ft (283 m) below sea level. *See* MEDITERRANEAN SEA; RIFT VALLEY; RIVER.

Islands. Important sectors of Asia, containing some 400,000,000 people, are completely insular. The most important are the Japanese, Philippine, and

Indonesian islands and Taiwan. Almost all of Asia's islands lie in great volcanic arcs bounding large seas off the continent's Pacific coast. At least 160 active volcanoes are found here and in Kamchatka. With such tectonic activity accompanied by periodic fractures, uplifts, and downfaultings, it is not surprising that the geological structures and topographic features of many of these islands are exceedingly complex and that the soils are generally immature. It also is to be expected that level alluvial land must be scarce, for rivers are short and, although their swift drop brings large loads of sediments, many of the islands rise from such great oceanic depths that delta extension seaward is slow. *See* OCEANIC ISLANDS; VOLCANO.

Few islands lie along the Asiatic coasts of the Indian Ocean, although the Sunda chain of Indonesia has perhaps more of a claim to Indian Ocean frontage than to Pacific frontage. The 400-mi (640-km) Andaman and Nicobar chains in the Bay of Bengal form genetic connections between the Arakan Yoma of western Burma and the mountains of northern Sumatra. Sri Lanka, however, is the only significant island in the northern part of the Indian Ocean west of Sumatra. In the Persian Gulf off the north coast of Arabia lies the small island Bahrein, which is rich in oil. *See* INDIAN OCEAN.

Few islands lie off the alluviated coastlands of northern Siberia. Some moderately large ones are included in the barren and rocky Severnaya Zemlya group, the New Siberian Islands, and Wrangel Island. The Commander Islands and Karaginski Island lie in the Bering Sea only a short distance from the Aleutians.

Climates. Five major climatic types may be distinguished in the Asian region: (1) the monsoonal system of eastern Asia, (2) the monsoonal system of southern Asia, (3) the equatorial regions of southeastern Asia and their extension into the Southern Hemisphere as they are influenced by the Australian monsoon, (4) the winter rainfall areas of southwestern Asia, and (5) the cyclonic and convectional storm systems of central and northern Asia.

Fundamental to understanding the climates of Asia are the vastness of the unbroken landmass and the long latitudinal stretch from the polar realm to south of the Equator. These are responsible for the great temperature and humidity extremes that occur. The greatest ranges of temperatures in the world have been recorded in interior Asia. Continentality, therefore, is the outstanding feature of climates of interior Asia. In coastal and insular areas of east Asia, however, winds moving over the warm, northward-flowing Japan Current and the western Pacific waters moderate the coastland and island climates. *See* MARITIME METEOROLOGY.

Monsoonal circulations. The term monsoon derives from an Arabic word referring to seasonal winds affecting southern and southeastern Asia that are predominantly onshore, warm, and wet during the summer and offshore, cool, and dry during the winter. While both southern Asia and eastern Asia have monsoonal regimes that are related to the large-scale western Pacific and Eurasian circulation systems, they differ somewhat in their characteristics. In each instance, however, the summer monsoon represents a vast penetration of tropical influences and humidity into more northerly latitudes.

The winter in central Asia is characterized by extremely low temperatures, associated with surface divergence. Outflow of air at ground level is blocked to the south by the high relief of the Tibetan plateau and the Himalayan mountain system. Bursts of cold air break out periodically across central Russia, northern China, and Korea, and reach as far as Japan, where they deposit large quantities of snow derived from moisture captured in passage across the Sea of Japan. Over China and the western Pacific this air makes its way gradually south, eventually passing as the northeast monsoon into southeastern and southern Asia. It is, however, a dry period in much of southern and southeastern Asia.

In summer the Asian landmass, and especially the Tibetan plateau, becomes a heat source for the atmosphere. In response, surface-level circulation patterns are established. One pushes a surge of humid warm air across the Indian Ocean from the southwest and into the Indian subcontinent, extending itself steadily during May and June all the way up to the northwestern part of Pakistan, where its progress is blocked by the mountain ranges. Some of this air passes across northern southeastern Asia to join with Pacific maritime air, and advances with a series of sudden steps into central and northern China, Korea, and the southern two-thirds of Japan. The initial onset of these monsoonal surges is accompanied by heavy rains, which are followed by intermittent bursts of rainfall as long as the dynamics of this southwest flow are sustained. By September both of these northward surges of energy have diminished, and gradually the monsoon conditions retreat to the south.

While southern and eastern Asia typify the classic monsoon pattern, parts of northern southeast Asia share in this seasonal reversal. Burma, Thailand, and Indochina receive the bulk of their precipitation from the summer monsoon. Their cool seasons (in the Northern Hemisphere winter) are dry, except for some coastal locations in Vietnam, but the three remain sheltered from the extreme cold that is associated with the outflow of Siberian air in regions farther to the north. *See* MONSOON METEOROLOGY.

Equatorial regions of southeastern Asia. Across southeastern Asia, and especially in parts of Malaysia and Indonesia, equatorial conditions prevail. At no time of the year is this region without substantial rainfall and high humidity. However, in moving eastward across Indonesia it becomes apparent that there is an effect derived from the "monsoonal" regime of the Australian landmass. Less pronounced than the monsoon of the Northern Hemisphere, that of Australia does nonetheless show similar features, with high rainfall and humidity in the "high sun" months (November through February) and cooler drier conditions in the "low sun" months (May through August).

Southwestern Asia. Southwestern Asia extends the desert zone of northern Africa eastward, with low

precipitation and extremely high summer temperatures. During the winter months cyclonic storm systems track through the Mediterranean trough and occasionally reach as far east as the Persian Gulf area. *See* DESERT.

Central Asia. Continentality is a predominant feature of the climates of this region. There are few sources of accesible moisture to provide humidity and precipitation. As is the case with southwestern Asia, midlatitude cyclonic storms penetrate this area, particularly in the summer months, to bring small amounts of precipitation. Winters are cold and dry.

Typhoons. Parts of southern, southeastern, and eastern Asia lie in the path of typhoons. Typhoons develop from initial disturbances or "waves" in the tropical easterly surface flows (flows sometimes referred to as the trade winds), usually at some distance north or south of the Equator. Such waves are not uncommon, but a portion of them, especially in the later summer months, develop into full-scale typhoons, which travel westward and then gradually curve away from the Equator in a northward or southward direction, depending on the hemisphere. Massive energy release comes from the high humidity of the tropical maritime air entrained in them, and the typhoons strike with ferocious winds and torrential downpours of rain. On making a landfall they usually lose their energy source and gradually die out. Those in the Bay of Bengal may cause extensive damage and bring heavy rains to the east coast of India. In the Pacific, the destructive typhoons may cross the northern Philippines and hit parts of the continental coast, usually between Amoy and central Vietnam. Many of them, however, follow a path gradually curving northward, past or across Taiwan, Okinawa, and the southern part of Japan as well as southern coastal Korea. These typhoons spread wind, wave, and flood destruction but also bring heavy beneficial rainfall over areas 100–300 mi (160–480 km) in diameter along their paths. *See* TROPICAL METEOROLOGY.

General rainfall patterns. The driest portions of Asia include the vast areas of southern Mongolia, Xinjiang, central Asia, and southwestern Asia. Except for small, favored mountain areas, most of this region from the Gobi to the Red Sea gets less than 10 in. (25 cm) of precipitation per year. With the exception of southern Arabia, which is subtropical desert, these are midlatitude desert and dry steppe regions. Favored with higher rainfall are the Yemen Mountains and the coastal mountains of Turkey, together with Lebanon, Syria, and northern Israel. The highlands of Armenia and the Elburz of Iran are favored also with more abundant rainfall, which ranges from 25 to 50 in. (64 to 127 cm) or more per year.

The northeastern Siberian mountains and the arctic coastal lands also receive meager rainfall, less than 8 in. (20 cm), but are not dry because evaporation is low and the water table is high. Most of Siberia has permafrost below a few feet of surface soil, so that rainwater does not filter far down. Between the arid belt of central Asia and the northeast Siberian low-precipitation zone, the annual rainfall ranges between 10 and 18 in. (25 and 45 cm).

In eastern Asia the precipitation increases in a southeasterly direction from interior Asia to the coast. The annual maximum seldom exceeds 80 in. (203 cm) in the wetter southeast coastal regions, whereas this drops to less than 30 in. (76 cm) in the North China Plain and less than 15 in. (38 cm) at the Great Wall. In some mountainous parts of Japan and Taiwan, the yearly average may be more than 100 in. (254 cm).

In the Indian subcontinent rainfall is heaviest along the western plateau fringe and in East Bengal, where it averages over 100 in. (254 cm) per year. The interior of the peninsula is relatively dry. Northwestern India and Pakistan share the drought of southwestern Asia. With the exception of the extreme north, Sri Lanka generally has abundant rainfall.

With the exception of some small areas in Assam, southeastern Asia has the heaviest rainfall of the entire Asiatic region. The mainland mountains facing the southwest summer monsoon crossing the Bay of Bengal, and parts of the Vietnamese and Laotian cordilleras facing the humidified northeast winter monsoons of eastern Asia regularly get average rainfalls of 120–150 in. (305–380 cm) or even more. Equally heavy rainfalls occur in the southwestern half of Sumatra, southwestern Java, the northwestern half of Borneo, and the Pacific fringe of the Philippine Islands. In general, and with only a few minor exceptions, southeastern Asia has no areas that are subject to severe drought.

General soil patterns. This large continent presents a variety of zonal soils and accumulations of soil material, particularly alluvium.

Podzolic soils (spodosols). In cold to cool regions of sufficient soil moisture, such as most of Siberia, soils are acid, podzolized, coarse-textured, and relatively sterile. Soils here generally are overlain with a semicarbonized coniferous leaf layer in the forest belt. The tundra ground layer may be bare rock, coarse sediment, or sterile sands. Sphagnum peat bogs occupy depressions. Elsewhere a peaty mat of little-decayed organic matter forms a "soil." This supports only tundra mosses and lichens and some scrub bushes. The greater part of Siberia, however, is in taiga, or boreal coniferous forests, scrublike toward the tundra belt but taller in the south.

Brown, coarse, acidic, and somewhat podzolized soils are characteristic of most forested areas in temperate regions of Asia because of the acidity of the leaf matter and the leaching of finer soil particles from the top soil horizon. In the higher mountains of southern and middle Asia, podzolic soils associated with coniferous forest belts are common. With warmer latitudes and lower altitudes, deciduous broad-leafed forests replace the coniferous forests, and the more rapid decomposition of leaf and litter on the forest floor produces a richer surface soil.

Calcareous soils. As ultisol-oxisol continua, the calcareous soils contrast sharply with the humid podzolic soils of the northlands and high-altitude slope lands of Asia. They include the arid and subhumid calcareous, chestnut-colored soils of the southern

Siberian, central Asian, and southwest Asian steppelands, and of the drier agricultural areas of China and India. Meager precipitation in these areas provides too little groundwater to support deep-rooted plants such as trees. Shallow-rooted grasses and herbs form a ground cover, which in time, if left uncultivated, becomes a thick layer of dead plant matter. Because of the low humidity, leaching of either organic matter or soluble minerals is much reduced. In fact, capillary action tends to bring dissolved minerals back to, or near, the surface after rains have ceased. The soils have a high percentage of soluble calcium and nitrogenous matter and are therefore of great agricultural richness. These are the soils of the west Siberian wheatlands and the steppes surrounding the central Asian deserts and such areas as southern Inner Mongolia, China's loess plateau, and Turkey's Anatolia. With sufficient water, either through rainfall or irrigation, they are exceptionally productive of cereals and other crops.

Toward the desert margins of these soils the humus thins with the thinning grass stands and the soil turns to light chestnut, but mineral plant foods are abundant. In the bottoms of poorly drained basins here, alkaline accumulations may inhibit vegetation growth. By contrast, where rainfall is somewhat more abundant, as in the west Siberian plain, the central Manchurian plains, and the Amur-Ussuri lowlands, high stands of thick grass may lead to deep black earths or chernozems.

Loess soils. The loess soils (various orders of cumulic soils) are a specialized type resulting from the accumulation of dust blown from desert regions where strong outblowing winds occur. The most significant and widespread deposits of this type in Asia are found in China's northwest loess plateau. Depths of fine-textured soils with a high calcium content may reach 100–300 ft (30–90 m). Incorporated with the loess are decayed grass roots and other organic matter. Vertical root channels with a cementation of soluble mineral salts produce a vertical structure so that columnar erosional cliffs or remnants are often found. This columnar pattern of calcification permits the Chinese in this region to tunnel into cliffs and carve out underground rooms for homes with comparative ease. Heavy rainfall results in sheet and gulley erosion, however, and much of the area has developed badland characteristics. Nevertheless, the soil is rich for agriculture. Grasses tend to be the natural vegetation where cultivation is not carried on.

Humid lateritic and leached soils. As ultisol-oxisol continua, these humid tropical and subtropical soils involve different processes in formation and have differing characteristics from the podzolized soils of the boreal and temperate forests. They also are acid and are formed under forest conditions, generally under broad-leafed forests, either deciduous monsoon forests or evergreen rainforests. Bacterial decay of the forest litter is so rapid under the warm, humid conditions and rainfall so heavy during part or all of the year that soluble minerals and comminuted organic matter are either leached away rapidly to deep soil horizons or washed away to rivers and the sea. The chemical weathering and solution of silica and the lack of carbonization of leaf matter to form carbonic acid leave iron and aluminum oxides in surface layers as red and yellow earths and clays. In extreme and long-term development, the subsurface red earths change to true laterite.

Most of the humid slope lands of Japan, southern China, southeastern Asia, and southern Asia have iron-rich red earths. In humid, mist-shrouded upland parts of southern China such as Guizhou Province, aluminum-rich yellow earths predominate. The fertility of these slope lands is quickly lost when stripped of trees and used for agriculture. Erosion then cuts the hillsides into ravines, while heavy sediment loads are carried by the numerous streams to the plains and the sea.

In the humid tropical and subtropical forest realms the fertility of the soil also differs considerably according to the nature of the parent materials. The older soils overlying the ancient rock complexes of the stable platforms such as are found in Borneo, southeastern Sumatra, large parts of southeastern Asia, and southern peninsular India are rather infertile compared with younger soil materials based upon more recent volcanic lavas and ash (andosols). The areas with numerous volcanoes and basic lava and ash have relatively fertile soils and support dense populations. Java, southwestern Sumatra, Sulawesi, and the Japanese islands demonstrate this relationship. Borneo, although large, is without volcanoes, and its infertile soils support a population only one-thirtieth as large as that of smaller Java. In peninsular India the most fertile soils are associated with the basaltic lava of the Deccan that produces the black waxy regur.

Alluvial soils. As entisol-inceptisol continua, the youngest, or least maturely developed, soils of Asia are the great areas of alluvium, which are intensively cultivated wherever water and temperature conditions are suitable. Their fertility in part derives from frequent inundations by floods which bring new layers of sediment containing organic debris and minerals in solution and suspension. Where irrigation water is available or can be brought to these alluvia, the soil is utilized for wet-rice cultivation in most parts of Asia except Russia (**Fig. 2**). *See* SOIL.

Major vegetation patterns. Asia's vegetation belts and zones follow, in general, the climatic patterns from desert lands through tropical to Arctic margins.

Tundra zone. A belt of tundra comprises the islands in the Arctic Ocean as well as a strip of territory on the Siberian lowlands. Its north-south extent is perhaps greatest in the northeast. In addition, it extends south into elevated portions of the hills and mountains. The frozen subsoil permits the growth of little more than mosses, lichens, dwarfed trees, and scrub. Because the flow of the great rivers here is northward, the mouths often remain blocked with ice even after their upper reaches have already melted. The flow downstream thus floods and spreads over immense regions of the tundra zone, making large areas impassable on foot. *See* PERMAFROST; TUNDRA.

Fig. 2. Alluvial bottomlands, long heavily cultivated for rice, in the hill country of Zhejiang Province, China, have been subject to damaging sand and gravel deposits washed from unprotected clearings for tung trees and sweet potatoes. (*From G. B. Cressy, Asia's Lands and Peoples, 3d ed., McGraw-Hill, 1963*)

The bush tundra comprises willow and dwarf birch, which hug the ground. Farther southward, fir, birch, and larch trees in stunted form make their appearance. All tundra vegetation is perennial. Tundra also is characteristic of the alpine heights of mountains of interior Asia.

Taiga (boreal coniferous forest). One of the largest expanses of forest left in the world is the Siberian taiga, a dominantly coniferous forest of larches, spruce, fir, and pines, with such deciduous trees as birch and aspen occurring intermixed with the conifers or taking over as a secondary growth in burnt-over areas. The width of this belt in Siberia is more than 1000 mi (1600 km) and it stretches about 4000 mi (6400 km) from the Sea of Okhotsk to the Urals (**Fig. 3**). *See* TAIGA.

West of Lake Baikal, the taiga reaches to the Alta-Sayan, and with little transition, changes to the steppelands of central Asia.

Midlatitude mixed forests. Various admixtures of coniferous and deciduous trees compose this vegetation. Large areas in eastern Asia either have or had this kind of forest prior to clearance. It occurs in Hokkaido and northern Honshu in the Japanese archipelago, and over much of the Korean peninsula. Much of the far eastern parts of Russia is also under this vegetation type, as in the Amur-Ussuri basin, and in the mountainous regions that surround the plains of Manchuria. From the plains themselves forests have long since been cleared.

Mixed midlatitude deciduous and coniferous forest areas of a similar type occupy the hill lands surrounding the Yellow Plain, as well as the Qin Ling Mountains and on into the hill country of the Chang Jiang river valley. This type of forest also may at one time have covered the North China (Yellow) Plain, but the plain has been under cultivation so long that only a guess can be made as to the original vegetation cover. It needs to be pointed out, however, that little commercially usable forest remains in the accessible areas of China and Korea whereas Japan has splendid forests of this so-called temperate type.

Mixed mountain and highland systems. A similar kind of forest makes up an element of the mixed mountain and highland forest systems. These systems occur in the Iranian uplands, in the Altai highlands, in the Pamir-Tian Shan highlands, in the Hindu Kush highlands, and in the Himalayan highlands at lower elevations. Similar forests also surround the Sichuan basin in southwestern China, and spread onto some of the adjacent hill and plateau regions.

Central and east Asian steppelands. From the forests of the west Siberian plain southward, an increasingly dry steppeland is encountered. It extends for 400–500 mi (640–800 km) in a belt about 1000 mi (1600 km) long between the Urals and the Altai-Sayan and associated uplands. The northern half of this belt with its higher annual precipitation of 12–16 in. (30–40 cm) [comprising the Ob dry plain and the northern part of the Kazakh upland running southeastward from it] is the agricultural heart of the plain. The southern part gradually changes to desert steppe and then to desert along about the 50th parallel north. A well-defined zone of dry mountain steppeland occurs in the Tian Shan flanks at elevations of 5500–6000 ft (1680–1830 m), below which desert prevails.

Eastward of Lake Baikal a broadened steppe zone occupies the Trans-Baikal region extending southward to the Gobi Desert of southern Mongolia and eastward to the Great Xingan Mountains, where the zone, about 200 mi (320 km) wide, runs southward in Inner Mongolia. The steppe zone in Inner Mongolia widens with the increasing moisture south of the Great Wall to include most of China's loess

Fig. 3. Winter view of boreal forest (taiga) in Siberia. Hunting and trapping continue to be characteristic activities in great portions of this forest. (*Sovfoto*)

plateau. Grasses also form the natural vegetation of the Manchurian plain, with tall grass in the eastern portion thinning out to short-grass steppe in the Xingan Mountain flanks. The Gobi Desert is flanked by steppelands to its north, east, and south, as well as by mountain steppe zones in the eastern Altai and eastern Tian Shan. From the heart of the Gobi both northward and southeastward, the grasses become richer and taller with the increasing average annual precipitation, and they are at their best in northern Mongolia. In southern Mongolia the richest grasslands have been preempted by Chinese agriculturalists for cultivation.

The typical grasses of the central and east Asian steppelands are feather grass, fescues, and koeleria. In moister parts of the steppe, dicotyledons such as sickle alfalfa, milk vetches, and other legumes are prominent herbs. A grazing economy occupies much of the steppelands, but in the Russian and Chinese spheres the richer grasslands have been put to the plow. *See* GRASSLAND ECOSYSTEM.

Mixed evergreen forests. This tropical and subtropical forest type appears to be limited in Asia mostly to interior southern China and to Japan from the Kwanto Plain southward. South of the Chang Jiang valley it extends from the coast at Shanghai to the gorge lands of eastern Tibet. It includes most of the Chang Jiang drainage and the western part of the Xi drainage, as well as the Yunnan plateau. The coast south of Shanghai in an ever widening zone inland, however, is warm and moist, finally becoming part of the subtropical rainforest.

In Asia the characteristic trees of the mixed forest include broad-leafed evergreen trees such as banyans and camphor, and coniferous trees such as pines, cedars, and cypresses, as well as varieties of bamboo. On the karst limestone hills of Guizhou and Guangxi the tree growth is mainly deciduous, but pines predominate in the leached soils of other hill areas. Horsetailed pines are very prominent. In Japan this forest zone is characterized by live oak, laurel, and other hardwoods such as camphor, intermixed with stands of Japanese cedar, pine, and bamboos. These forests occupy most of the areas of red and yellow earths in China and Japan. The Japanese mountains are well covered with forest, but large parts are in secondary growth of genya or fuel wood. In China most of the original forest has long been cut down and the hills are largely barren of large trees. Much of the cut-over slope land is covered with coarse grasses, ferns, and various scrub, although a beginning has been made in reforestation.

Tropical and subtropical rainforest. The rainforest type is restricted to warm or hot regions of southern and southeastern Asia which get ample rainfall the year round or get so much rain during a large part of the year that a high ground-water table is maintained during the short dry season. The subtropical sectors are found along the southeastern China coast, in Taiwan, and in northern Burma; they merge with the tropical rainforest farther south, where rainfall and temperature increase. Most of southeastern Asia lies in this forest zone, but there are large areas which have such a long dry winter season that they do not support evergreen rainforests. These exceptions are in monsoon forest or savanna grassland described in subsequent sections on Asia's vegetation. In southern Asia the rainforest zones include Assam and the west slopes of the Chindwin, Chin, and Arakan hills, the southern half of Sri Lanka, and the Western Ghats and associated lowlands of southwestern India. *See* RAINFOREST.

The variety of trees and plants in the rainforest is enormous; 11,000 species of flowering plants have been recorded from Borneo alone. Trees are seldom in pure stands, making lumbering a selective and costly process. Camphor and sweetgum, bamboo and palm, citrus, and banyan are common trees in the subtropical coastal areas of China. In the southern tropical rainforest, the coastal forests in areas of tidal mudflats and brackish ground water are formed by varieties of mangrove. Since these are not so dependent upon rainfall, they are found also in coastal areas having seasonal droughts, for example, in eastern India. Mangrove is the characteristic forest in all the muddy tidal soils of the tropics. Next inland and in brackish water, the nipa palm may form dense stands in river estuaries. Sandy coastal strips may have *Casuarina*, *Pandanus*, *Barringtonia*, and dense plantations of coconut palms.

Dipterocarps occupy a prominent place in the lowland rainforests of the interior, because of both their height and the large amount of space they cover. They often form the top story of the forest, with their crowns towering up to 200 ft (60 m) or more. Below an altitude of 2600 ft (800 m) in Malaya they cover 50% of the wooded area, in the Philippines 75%, and in Indonesia as a whole about 50% of the forest stands. Camphor and species of *Shorea* and *Kompassia* are among the giant trees of this forest in Indonesia. Other prominent trees of this forest are gumwoods, mahogany, and ebony. Many banyan-type trees send down aerial roots from their lower branches to take root upon reaching the ground, forming a forest of smaller boles from a single tree. Typical features of the taller trees of this forest are great flaring buttresses that anchor them to the ground. Below the larger trees, a closed canopy up to 60 ft (18 m) high may be formed by secondary trees. Tangles of lianas, which in the case of the rattan palm may reach 500–600 ft (150–180 m) in length, and the abundance of epiphytes and parasitic plants growing on other trees give the impression of an utter confusion of plant life.

Clearing of the tropical rainforest and subsequent abandonment results in a succession of ferns, tropical grasses, and scrub which may be transitional to ultimate reforestation. Often, however, coarse, tough, tall savanna grasses get established; these are more permanent and stifle other growths. Burning may destroy their surface parts, but underground runners quickly send up new sprouts. Forest clearance for agriculture, for settlement, and for timber harvesting has gone on rapidly in southeastern Asia since the end of World War II, and the effective tropical forest area has been reduced by well over half. In

addition, three decades of warfare in Indochina resulted in heavy damage to forest ecosystems.

Monsoon tropical deciduous forest. In the tropical parts of Asia, which have a moderately high rainfall but a long dry season (usually in the low-sun period or winter), broad-leafed trees drop their leaves during the season of drought. Most of India is in this vegetation zone, with the exception of the dry northwest and the south-central peninsular areas, the rainforest zone of the Western Ghats and southwestern coast, and the Himalaya mountain zone. In southeastern Asia this zone is found in central Burma, the Shan plateau, northwest Thailand, and in a small part of Assam. It also is found in southern Vietnam and southern Laos together with adjoining areas of Cambodia. Similar monsoon forests occur in the Lesser Sundas, east Java, and Sulawesi.

The monsoon deciduous forest consists mostly of mixed species, but sometimes a single species becomes dominant as a result of selection from frequent burnings. Teak, which has especially attracted commercial attention, is one of these. Its seeds are hard-cased and fire-resistant; even its saplings have a high degree of fire resistance. Sometimes pure stands of teak develop.

Perhaps even more widespread in monsoon forests are bamboos, which frequently grow in pure stands over wide areas. Where burnt over during the dry season, their root clumps quickly send up new shoots during the rainy season. In the Lesser Sundas, Sulawesi, and the Philippines, areas of eucalyptus forests resembling teak forests in origin have developed as migrants from Australia. *See* BAMBOO.

The monsoon dry forest, an open woodland of very few species, is associated with the poor soils and low rainfall—generally not much more than 40 in. (100 cm)—found in large areas of Burma, Thailand, and Vietnam. In densely populated regions such as India, cultivation has made large parts of the monsoon forest zone treeless.

Hot deserts. Much of the lower Indus valley in Pakistan, along with extensive regions in southwestern Asia—in Arabia and in the lands around the Persian Gulf—are areas of hot desert. Vegetation is sparse and strongly drought-resistant. Except where irrigated, these deserts are lightly populated, although land-use systems involving grazing have placed considerable stress on the land.

Cold deserts and semidesert regions. Large areas of central Asia are desert, but unlike the deserts in southwestern Asia they experience cold winters. Plant life is very restricted, and there are large areas of bare rock. Shifting sand dunes are also common. The Gobi and Taklimakan are at lower elevations. The Tibetan and Iranian plateaus are likewise desert, and because of altitude even show an Arctic character in places.

In the Gobi, precipitation may be in the form of showers or protracted drizzles, but the tropical desert areas generally receive their meager rainfall in torrential downpours on rare occasions. Sudden floods rush through the dry washes known as wadis in Arabia but soon drain away or sink into the sands.

After such rains numerous herbs may spring to life and flower, while the bunch grass here and there may become green for a short season. Xerophytic shrubs such as wormwood and saxaul and varieties of sage are common in the desert. Stunted willows occur where fresh ground water is near the surface. In the wetter climatic fringes, a thin grass cover appears.

The preceding descriptions serve to demonstrate that even in broad characterizations the Earth's largest landmass exhibits the most varied of physical characteristics.

Herold J. Wiens; Jerome D. Fellmann; W. D. McTaggart

Bibliography. C. Hutchison, *Geological Evolution of Southeast Asia*, 1989; R. Jishun et al., *Geotechtonic Evolution of China*, Science Press, Beijing, and Springer-Verlag, Berlin, 1987; A. A. Meyerhoff et al., *China: Stratigraphy, Paleogeography, and Tectonics*, Kluwer, 1991; V. Pokshishevkii, *Geography of the Soviet Union*, 1974; K. Takahashi and H. Arakawa (eds.), *Climates of Southern and Western Asia*, 1981; T. R. Tregear, *An Economic Geography of China*, 1970; G. Trewartha, *Japan: A Geography*, rev. ed., 1965; U.S. Department of the Interior, Board on Geographic Names, *Gazetteer of the People's Republic of China: Pinyin to Wade-Giles, Wade-Giles to Pinyin*, 1979; R. K. Verma, *Geodynamics of the Indian Peninsula and the Indian Plate Margin*, 1991; S. Zhao, *Physical Geography of China*, 1986.

Asparagales

A large, widespread order of petaloid monocotyledons consisting of 29 families and about 26,000 species. Asparagales are clearly circumscribed in deoxyribonucleic acid (DNA) sequence analyses but are difficult to define morphologically, and separation from Liliales has proved particularly problematic. Phytomelan, a dark seedcoat pigment present in most families of Asparagales but not found in other plants, is an obvious characteric of this order, although seeds of the largest family, Orchidaceae, and a few other taxa lack phytomelan. Most Asparagales are herbaceous perennials, but there are some vines including some species of asparagus (Asparagaceae), and woody taxa including aloes (Asphodelaceae). The order contains many horticultural taxa, including members of Orchidaceae (orchids), Amaryllidaceae (daffodils, belladonna lilies, and others), Iridaceae (irises, gladiolus, freesias), Convallariaceae (lily of the valley, Solomon's seal), and Hemerocallidaceae (daylilies), in addition to food crops including onions, leeks, and garlic (Alliaceae). *See* FLOWER; LILIIDAE; LILIOPSIDA; MAGNOLIOPHYTA; ORCHID; ORCHIDALES; PLANT KINGDOM.

Michael F. Fay; Mark W. Chase

Asparagus

A dioecious perennial monocot (*Asparagus officinalis*) of Mediterranean origin belonging to the plant order Liliales. Asparagus is grown for its young shoots or spears, which are canned, frozen, or cooked fresh

as a vegetable. These aerial stems arise from rhizomes (underground stems). The rhizomes and the fleshy and fibrous roots constitute the massive underground part of the plant. *See* LILIALES.

Propagation. Asparagus is propagated by seed with 1-year-old crowns transplanted to the field and spaced 8-18 in. (20–45 am) apart in 4–8 ft (1–2 m) rows. Although male female plants, separating crowns on the basis of sex has not been economical. Mary Washington is the principal variety cultivated; several new varieties developed by the University of California are widely planted. Blanched or white asparagus is grown by ridging soil over the rows and cutting the spears beneath the soil surface. Chemical weed control is commonly weed.

Harvesting. The length of the annual harvest season varies with age of the crowns and with climatic conditions. Generally, spears are cut for 8–10 weeks each spring after the crowns are 3–4 years old. In area with longer growing seasons, such as California, harvesting begins earlier and continues 10–12 weeks. Commercial plantings are often harvested for 12–16 years. In most areas special knives are used for harvesting; however, spears for canning and freezing are sometimes snapped off by hand above the ground level. Several types of mechanical harvesters are now in commercial use.

Commercial production is limited to areas where crowns will have a dormant period of 3–5 months each year. Dormancy in the northern states is induced by low temperatures and in California by withholding irrigation. California, New Jersey, and Washington are important asparagus-producing states.

H. John Carew

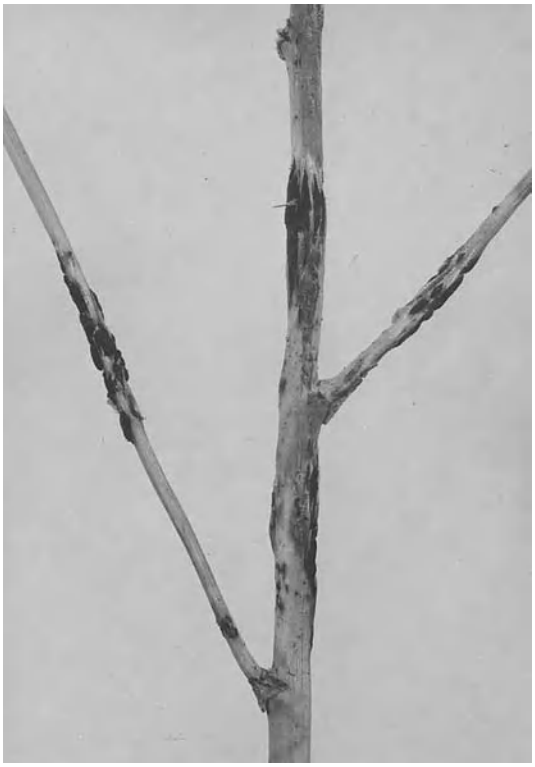

Asparagus rust. (*Courtesy of Arden F. Sherf*)

Diseases. Asparagus yields can be reduced by two fungal diseases, rust and *Fusarium*-induced wilt and root rot complex.

Asparagus rust (see **illus**.) is caused by a long-cycle autoecious rust fungus, *Puccinia asparagi*. The uredinial stage of the fungus is the most damaging and appears in the summer on stalks that grow after the harvest season. Uredinial lesions appear as oval reddish-brown pustules. Severe infection for 2 or 3 consecutive years reduces the vigor of the crowns and results in lower yields. Growers are advised to plant resistant varieties. Applications of fungicidal sprays in summer and fall may give partial control.

The Fusarium wilt and root rot complex is caused by two species of *Fusarium*. The fungi soil-borne, and infections occur in the roots or crowns. In addition, airborne spores can infect through insect injuries or other lesions near the base of the stalks. The fungi are also seed-borne. Affected plants become wilted, stunted, and yellowish-brown in color. Red streaks occur in the crowns, and the roots rot. In an established planting, individual diseased plants may be scattered among healthy ones, but eventually entire fields may be destroyed.

Control practices for Fusarium wilt and root rot include seed treatment, the planting of disease-free crowns, and planting in fields where asparagus has not previously been grown. *See* PLANT PATHOLOGY.

Thomas H. Barksdale

Aspartame

A white, crystalline compound, 1-aspartyl-1-phenylalanine methyl ester (APM), with formula (**1**). It is slightly soluble in water. Its sweetening

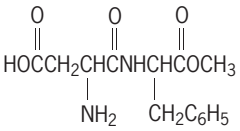

(**1**)

properties were discovered accidentally in 1965 when the compound, a dipeptide, was produced as an intermediate in the synthesis of the C-terminal tetrapeptide of gastrin. Aspartame is the L,L-diastereoisomer; the three other possible diastereoisomers are not sweet. The taste of aspartame would not have been predictable based on its component amino acids, aspartic acid and phenylalanine.

The sweetness of aspartame relative to sucrose is a function of the latter's concentration, and is also dependent upon the presence of other flavors and materials. In a number of applications, such as chewing gum and various fruit-flavored products, aspartame favorably extends and enhances the flavor perception, and it shows synergy with other sweeteners. The sweetness perception may also last longer with aspartame than with sucrose or other sweeteners. *See* SUCROSE; TASTE.

Aspartame is metabolized to its component amino acids, which are further metabolized by the usual

metabolic pathways. Under certain conditions of heat and pH in aqueous solution, aspartame is transformed into its diketopiperazine derivative, 3,6-dioxo-5-benzyl-2-piperazineacetic acid (**2**), which is

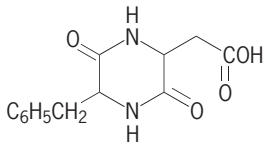

(**2**)

tasteless. This property limits the use of aspartame when it is exposed to high temperatures, such as in baking. The stability of aspartame in aqueous solution is pH-dependent; it is most stable at a pH of approximately 4. The rate of conversion (its half-life is 262 days at 77°F or 25°C) is sufficiently slow under the conditions of normal use that aspartame has found an increasing number of applications in various food products, and is particularly successful in soft drinks. The safety of aspartame has been established by studies in animals and human beings. Aspartame has been approved in many countries for uses in both dry and wet applications. *See* FOOD ENGINEERING.

Daniel L. Azarnoff

Asphalt and asphaltite

Varieties of naturally occurring bitumen. Asphalt is also produced as a petroleum by-product. Both substances are black and largely soluble in carbon disulfide. Asphalts are of variable consistency, ranging from a highly viscous fluid to a solid, whereas asphaltites are all solid. Asphalts fuse readily, but asphaltites fuse only with difficulty. Asphalts may, moreover, occur with or without appreciable percentages of mineral matter, but asphaltites usually have little or no associated mineral matter. *See* BITUMEN.

Natural occurrence. Many asphalts occur as viscous impregnations in sandstones, siltstones, and limestones. Most such deposits are thought to be petroleum reservoirs from which volatile constituents have been stripped by exposure of the rock.

The asphaltites (gilsonite, graphamite, and glance pitch) were probably derived from a saline lacustrine sapropel and owe their variable properties to differences in environment of deposition. These substances occur on a large scale in the Uinta Basin of northeastern Utah, where they are derived from upper Eocene Green River sediments, most of which are oil shales high in carbonate content. *See* OIL SHALE; SAPROPEL.

Asphalt. Relatively pure asphalt occurs in Kern, San Luis Obispo, and Santa Barbara counties, California. The asphalt, or tar seep, may contain up to 30% mineral matter and occasionally, as in the La Brea tar pits, has remains of insects or animals that became entrapped in the viscous material. Many of these occurrences are associated with the Miocene Monterey shale, which is thought to have been a source bed for the local crude oil. Occurrences of asphalt are also known in Kentucky and Oklahoma. Although asphalt seeps have long been known in France, Greece, Russia, Cuba, and other countries, the best known and largest are those of Venezuela and Trinidad.

Venezuelan asphalt has been recovered from the Bermudez Pitch Lake covering over 1000 acres (400 hectares) and averaging 5 ft (1.5 m) in depth. The lake is thought to be fed by asphalt springs. The asphalt at the edges of the lake becomes hard enough to walk on. After treatment to remove water and volatile constituents, the asphalt fuses at 130–140°F (54–60°C), is 92–97% soluble in carbon disulfide, and contains about 83% carbon, 11% hydrogen, 6% sulfur, and 1% nitrogen.

The Trinidad Asphalt Lake covers 115 acres (47 ha) and is 135 ft (41 m) deep at its center. The asphalt is softest at the center, where it is probably fed by underground springs, and hardest at the edges. Even at the center, however, the material is hard enough to support a person and can be broken into blocks. Gas is evolved from the asphalt, and rainwater tends to collect in depressions forming an emulsion of asphalt and water containing gas, sand, and clay. The crude material is 39% soluble in carbon disulfide, contains 27% mineral matter, and loses 29% of its weight as water and gas on heating to 100°F (38°C). After the refining process which drives off gas and water, the product fuses at about 190°F (88°C), is 56% soluble in carbon disulfide, contains 38% mineral matter, and consists of 80–82% carbon, 10–11% hydrogen, 6–8% sulfur, and less than 1% nitrogen, all on an ashfree basis. The Trinidad lake contains about 10,000,000–15,000,000 tons (9,000,000–14,000,000 metric tons) of asphalt.

Sandstones impregnated with asphalt occur in Oklahoma, Kentucky, Arkansas, Alabama, Utah, California, and other states. Many of these deposits have been developed as a source of paving material. Similar deposits also occur in Canada, South America, Europe, Asia, and Africa. One of the most extensive occurrences is that of Asphalt Ridge near Vernal, Utah, where asphalt saturates sandstones of the Uinta formation of Eocene age and of the Mesaverde formation of Upper Cretaceous age. This deposit crops out over a distance of 11.5 mi (18.5 km) and the sandstone contains 8–15% asphalt. It has been estimated that over 1,000,000,000 tons (900,000,000 metric tons) of high-grade material is readily available. The asphaltic sandstone is scooped from the deposit and used directly in road construction. If necessary, it is mixed with sand to attain the desired consistency.

Most asphalts are of marine origin and consist of the high-molecular-weight compounds normally present in petroleum residues. Asphalts and

TABLE 1. Major physical differences of asphaltite groups

Group	Specific gravity at 77°F (25°C)	Softening point, °F (°C)
Gilsonite	1.03–1.10	230–350 (110–177)
Glance pitch	1.10–1.15	230–350 (110–177)
Grahamite	1.15–1.20	350–600 (177–316)

TABLE 2. Asphalts and their uses

Asphalt type and % of production	Manufacturing process	Properties	Uses
Straight-run, 70–75%	Distillation or solvent precipitation	Nearly viscous flow	Roads, airport runways, hydraulic works
Air-blown, 25–30%	Reacting with air at 400–600°F (204–316°C)	Resilient; viscosity less susceptible to temperature change than straight-run	Roofing, pipe coating, paints, underbody coatings, paper laminates
Cracked, less than 5%	Heating to 800–1000°F (427–538°C)	Nearly viscous flow; viscosity more susceptible to temperature change than that of straight-run asphalt	Insulation board saturant, dust laying

asphaltites often contain unusually high percentages of vanadium.

Asphaltites. These substances are divided into three groups: gilsonite, glance pitch, and grahamite. The major physical differences in these substances are in specific gravity and softening point, as shown in **Table 1**. All three substances are nearly completely soluble in carbon disulfide. *See* IMPSONITE; WURTZILITE.

Differentiation of the asphaltites into three groups is based only on physical properties and not on a genetic basis. For this reason, similarly categorized substances may have somewhat different origins and variable compositions.

Glance pitch occurs on Barbados, and material from this deposit has been marketed as manjak. Other veins of glance pitch, some of which contain up to 27% mineral matter and up to 7.4% sulfur, also occur in Haiti, Cuba, Mexico, Argentina, Colombia, Chile, the Baltic states, and the Near East. Glance pitch has been used to make lacquers.

Grahamite occurs in West Virginia, Texas, Oklahoma, and Colorado. It is also known in Mexico, Cuba, Trinidad, Argentina, and Peru. The Peruvian grahamite is particularly rich in vanadium, and some vanadium minerals are associated with it. In general, most deposits are relatively small and are no longer of commercial interest. Irving A. Breger

Petroleum by-product. Asphalt is derived from petroleum in commercial quantities by removal of volatile components. It is an inexpensive construction material used primarily as a cementing and waterproofing agent. Over 27,000,000 tons (24,000,000 metric tons) of asphalt is used in the United States annually, of which more than 98% is derived from petroleum. *See* PETROLEUM PRODUCTS.

Asphalt is composed of hydrocarbons and heterocyclic compounds containing nitrogen, sulfur, and oxygen; its components vary in molecular weight from about 400 to 5000. It is thermoplastic and viscoelastic; at high temperatures or over long loading times it behaves as a viscous fluid, at low temperatures or short loading times as an elastic body.

The three distinct types of asphalt made from petroleum residues and their uses are described in **Table 2**.

In the construction of pavement surface for major roads and airport runways, hot asphalt is mixed with hot graded-stone aggregate. The mixture is spread on a dense, compacted stone base and rolled while still hot to give a smooth surface.

Roads having only light traffic are often given a thin, inexpensive wearing surface by spraying fluid asphalt on the road base and covering it immediately with stone. The fluid asphalt may be hot paving asphalt or a liquid asphalt. Liquid asphalts are produced by blending asphalt with various petroleum distillates or by emulsifying hot asphalt with water containing a small amount of soap. The liquid asphalts are fluid at ambient temperatures but harden as the solvent or water evaporates. *See* PAVEMENT.

Air-blown asphalt is used mainly for roofs. Hot asphalt may be mopped on the roof and covered with decorative gravel, or prefabricated asphalt shingles may be nailed onto the roof. *See* ROOF CONSTRUCTION.

Asphalt is also used in hydraulic works to line canals and reservoirs, to face dams and dikes, and to bind together the rocks in breakwaters.

Thomas K. Miles

Aspidogastrea

A group of entoparasites considered to be a subclass or order of the Trematoda. They have strongly developed ventral holdfasts whose alveolations are taxonomically important. Two families, Aspidogastridae and Stichocotylidae, are recognized. Aspidogastridae, which are the commonest, occur in various cavities of mollusks and in digestive tracts of fishes and turtles. Development is usually direct, involving nonciliated juveniles and one host. Fishes and turtles probably acquire infestations of adult worms by ingesting parasitized mollusks. Elongate Stichocotylidae have a single row of alveoli and occur in the digestive tracts of skates. Juvenile *Stichocotyle nephrops* encyst in lobsters and develop to maturity after ingestion by skates, thus approaching the digeneid life cycle. Little is known of the physiology of aspidogastreids, but they appear less host-specific than other trematodes. *See* DIGENEA; TREMATODA.

William J. Hargis, Jr.

Aspidorhynchiformes

A small order of specialized holostean fishes which are first recorded from Middle Jurassic deposits of Europe. Later records indicate that they were a very successful group in terms of geographic range, and

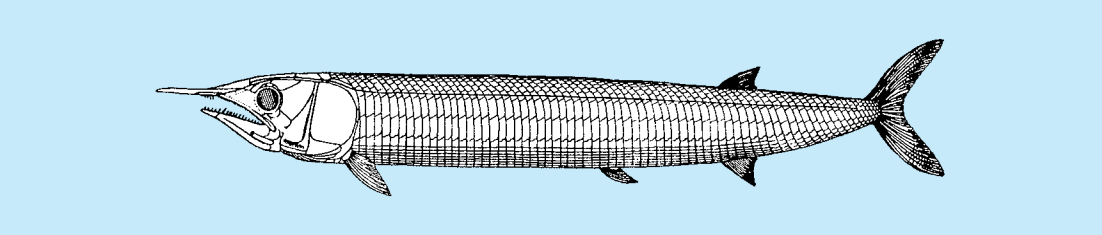

***Aspidorhynchus acutirostris*, Upper Jurassic, Bavaria; a typical fish of the order Aspidorhynchiformes.**

probably had a worldwide distribution in the warm seas of the Cretaceous Period. The order contains one family, Aspidorhynchidae, and two genera, *Aspidorhynchus* and *Belonostomus*. These fishes, some of which reached a length of over 3 ft (0.9 m), are characterized by a ganoid scale covering with much deepened scales along the flank, by an elongate fusiform body and head with long slender snout, and by an externally symmetrical tail. All the fins are small, and fringing fulcra are reduced or absent. The dorsal and anal fins are positioned opposite one another far back on the body, and the pelvic fins are inserted closer to the anal than to the pectorals (see **illus**.). The vertebral column shows a fair amount of ossification with thin ringlike or half centra. *See* HOLOSTEI.

Noteworthy features of the skull are the small supraoccipital in the neurocranium, a large preopercle, a single series of plates behind the eye which cover the cheek region, and a mandible bearing closely set sharp teeth, with a single median (predentary) bone in front of the symphysis.

Aspidorhynchiforms are an isolated group and their relationships are not clear. They do not appear to have arisen from the earlier but superficially similar chondrosteans of the family Saurichthyidae because of substantial differences in the opercular series between the two groups. A relationship with the Pholidophoriformes has been suggested by some authors because of several skull characters that are shared. *See* CHONDROSTEI; PHOLIDOPHORIFORMES.

In the body form, fin position, and elongated snout, the aspidorhynchiforms resemble some of the living teleostean Exocoetoidei (needlefishes and sauries). It seems likely that these two widely separated and unrelated groups of fishes shared a similar mode of life, being predaceous open-water forms that utilized their long snouts and strong swimming ability in capturing prey. *See* OSTEICHTHYES; TELEOSTEI.

Ted M. Cavender

Bibliography. R. M. Alexander, *Chordates*, 2d ed., 1981; D. V. Obruchev, *Fundamentals of Paleontology*, 1967; A. S. Romer, *Vertebrate Paleontology*, 3d ed., 1966.

Aspirin

The acetyl ester of salicylic acid, also known as 2-(acetyloxy)-benzoic acid and acetylsalicylic acid. Aspirin is prepared by the acetylation of salicylic acid with acetic anhydride. Its structural formula is below.

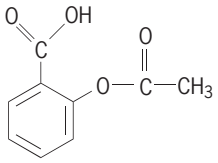

Pain relief by the bark of the willow tree has been known since antiquity. In principle, therefore, aspirin is a natural remedy. Various techniques to extract the active principle, salicin, from the bark and to prepare salicylic acid were developed during the first half of the nineteenth century. During the second half, salicylates were used to preserve food, to control pain and fever, and to relieve gout. Aspirin was introduced into practice around 1899, mainly through the efforts of chemist Felix Hoffmann.

Measurable amounts of the acetylated compound are absorbed from the stomach, but most of it is absorbed (after hydrolysis) from the small intestine; none is absorbed from the large intestine. Once absorbed, aspirin is deacetylated; the peak salicylate level occurs within 1 or 2 h after absorption. Salicylic acid binds to plasma protein. Aspirin is excreted as a glucuronide, salicyluric acid, or gentisic acid. Many individual differences exist for the absorption, tolerance, and therapeutic usefulness of aspirin.

Aspirin is effective as an analgesic, antipyretic, and antiinflammatory drug. It prevents the aggregation of platelets, and there is some evidence that it can prevent stroke. Aspirin, if tolerated, is the preferred drug for the treatment of rheumatoid arthritis, and it has been used in the treatment of osteoarthritis. Aspirin lowers fever, probably by acting on the hypothalamus. Salicylates inhibit aldose reductase in the lens; it has been suggested that they might retard the development of cataracts. Aspirin might encourage the development of Reye's syndrome, an acute encephalopathy which occurs in children who recover from viral disease, but this cause-and-effect relationship remains to be confirmed. *See* ANALGESIC; ARTHRITIS.

Arachidonic acid is a cyclic fatty acid which is the raw material for prostaglandins and leukotrienes. Transformation of arachidonic acid into prostaglandins requires cyclooxygenase; transformation into leukotrienes requires lipoxygenase. Aspirin is known to inactivate cyclooxygenase by

acetylation, but has no effect on lipoxygenase. It has been suggested that the inhibition of cyclooxygenase by aspirin offers lipoxygenase a more-than-normal access to arachidonic acid; and thus could induce an exaggerated production of leukotrienes. Since some leukotrienes are potent bronchoconstrictors, they have been thought to explain the violent reactions to aspirin in aspirin-sensitive patients, but some research has cast doubt on this hypothesis.

Intolerance to aspirin is not uncommon. It tends to develop in middle age and involve the skin or the respiratory tract, or both. In the skin, it causes urticaria. In the respiratory tract, nasal polyps precede the development of aspirin sensitivity. Intermittent (but occasionally progressive) bronchial asthma might occur. Death rarely ensues because people rapidly become aware of their intolerance.

The clinical symptoms of aspirin sensitivity are almost identical with those encountered in clinical allergy and anaphylaxis. Consequently, reactions caused by aspirin have been thought to represent an aspirin allergy, that is, an antibody-mediated sequence, although aspirin-sensitive individuals do not have an increased titer of immunoglobulin E (IgE). Intolerance to aspirin may be the result of a specific disease which alters receptors. Aspirin-sensitive individuals also react to nonsteroidal antiinflammatory drugs, and, occasionally, to the food color FD&C yellow no. 5, which is a pyrazolone derivative. *See* ALLERGY; ANAPHYLAXIS; IMMUNOGLOBULIN; PAIN.

Max Samter

Bibliography. L. Lasagna and F. G. McMahon (eds.), New perspectives on aspirin therapy, *Amer. J. Med.,* June 1983; M. Samter, Intolerance to aspirin, *Hosp. Pract.,* 8:85–90, 1973; M. Samter (ed.), *Immunological Diseases,* 1978; M. Samter and R. F. Beers, Jr., Intolerance to aspirin: Clinical studies and consideration of its pathogenesis, *Ann. Int. Med.,* 68:975–983, 1968.

Assembly machines

Machines that take discrete components as come into an assembly department and bring them together so as to produce a configuration of some practical value. Such machines differ from packaging machinery in two ways: in assembly machinery, components must be inserted in specific sequence and spatial attitude; and they must often be tested functionally as part of the assembly process.

Assembly machinery was originally conceived for situations where volume or hazard of production, parts size, or availability of labor made manual assembly impractical from an operational or economic viewpoint. The early applications of assembly machinery and the majority of modern applications are found in the automotive industry, consumer products, manufacturing, or hazardous assembly.

Although automatic assembly machinery competes with inexpensive manual labor in emerging countries or depressed areas in mature industrial countries, manufacturers have different attitudes toward justification of assembly machinery because of single world economy, rising costs in developing countries, and consumer attitudes toward product quality. Assembly machinery is often justified on its increased quality, reduced in-process inventory, and rapid response to customer specification, not on labor savings alone.

Types. Assembly machinery can be classified in several ways, including work path (rotary, carrousel, or linear); index–dwell ratios (continuous motion, intermittent motion, or power and free); actuation (mechanical, fluid power, or electronically programmable); work nest configuration (pallet or walking beam); and design (special or standard modular). The engineers who specify assembly machinery will usually select from one or more of these categories based on product size and weight, volume of production, product life cycle, future and present flexibility needs, human resources, and return on investment. Choice among these types of machines is restricted by choice of builders, industry backlog, and importance of prior operating experience.

Work path. Rotary dial machines have a number of pallets fastened to a rotating dial or ring. Transfer devices are usually mounted to the machine base, while feeders are placed outside the periphery of the dial. Advantages are low initial cost, high cyclic speed, low inertia, and good index accuracy (the ability of a machine to present the work-holding fixtures or pallets to a specific operating station). Disadvantages include limited number of work stations, lack of future flexibility, congested work stations, and poor access for maintenance. Rotary machines usually have high operating costs, primarily due to intensive use of pneumatic components and the high downtime required for maintenance.

Carrousel machines are usually configured in race track shape. Like rotary machines, they usually have fixed index–dwell ratios. Work-holding nests (the machine elements used to hold all of the parts being assembled) are secured to one another (a configuration known as precision link) or fastened to a chain. Where high index accuracy is required, they are fastened to stressed steel bands. Modern carrousel machines usually have a high level of mechanically actuated and synchronized work stations, driven from one or more lateral main camshafts. Highly modular, they can be reconfigured readily to meet product modifications or even product changes. Advantages include a larger number of work stations and good maintenance characteristics. Disadvantages include high capital cost, limited line-balancing ability and, with the exception of stressed band machinery, modest index accuracy. They have good to excellent operating costs. Carrousel machines may have pallets in cantilever form around the periphery of the machine or may be built in an over-under configuration, a system in which the fixtures pass the various functional stations on the machine and then go over a drum or wheel for return to start a new cycle.

Linear machines (see **illus.**) have an open-loop configuration in which pallets transport the product being assembled along one or more linear paths,

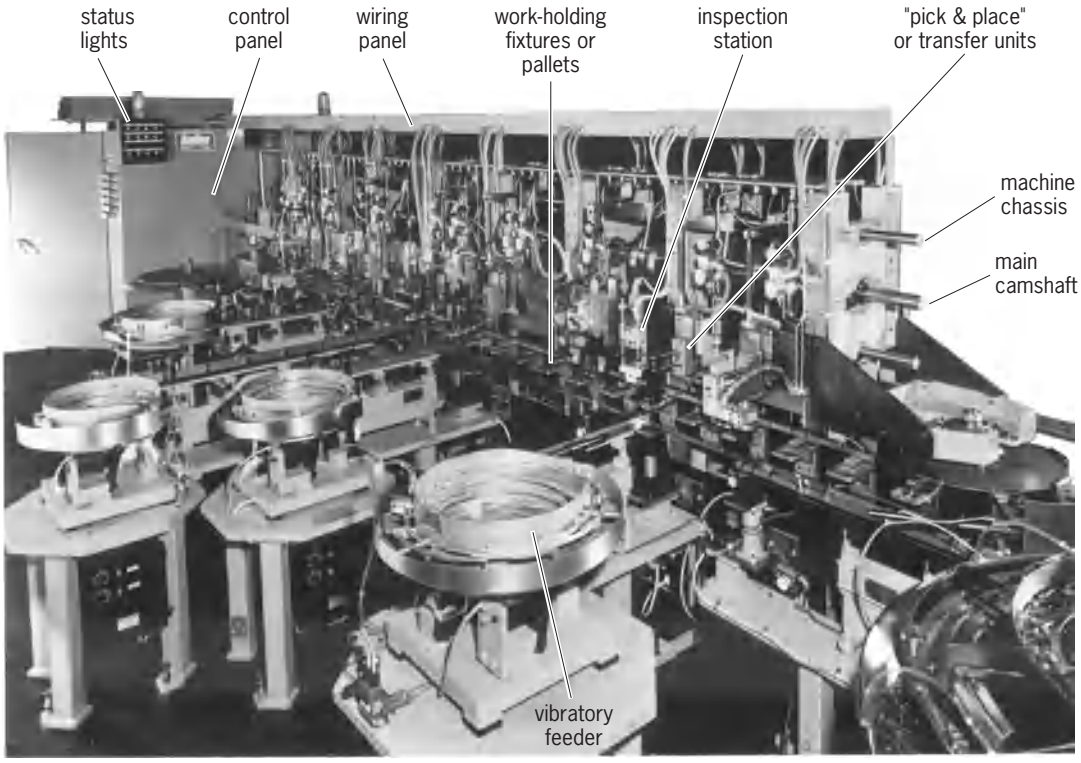

Linear synchronous assembly machine.

rather than in a closed circuit such as rotary or carrousel machines. Most linear machines are power and free systems in which each individual station is activated by the presence of the work-holding pallet before it.

Index–dwell ratios. Motion of a pallet or nest is called index, and lack of motion is called dwell. Together they form one machine cycle. In most machines, the index portion of the cycle is used by transfer devices to acquire components from feeding mechanisms. During the dwell portion of the cycle, parts are inserted, and joining operations and inspections are performed. In general, the dwell should exceed the index portion of the cycle. The ratio of dwell to index is often 1:1 at high cyclic rates, but the ratio will be much higher at low cyclic rates.

In any form of automatic assembly, production rates will be controlled by the single longest operation. If new parts can be added to the assembly without the fixture being stopped, the machine can be operated in a continuous motion. Usually such equipment is suitable for assemblies in which component parts do not need radial orientation, for example, deodorant dispensers and simple aerosol valves. Continuous-motion machines are capable of extremely high rates of production, but their application range is very restricted.

Power and free machines are sophisticated conveyor systems in which work pallets are transferred from work station to work station by free-flow conveyors, walking beam systems, or motorized pallets. These transfer machines are specifically indicated where line-balancing requirements are severe and fixture/workpiece weight is high (for example, above 11 lb or 5 kg). Primary disadvantages include high procurement costs and high engineering costs. Advantages include high index–dwell ratios (since independent stations operate only when a fixture pallet is present) and good ability to interface people among automatic sequences; line balancing can be achieved by parallel paths or redundant stations.

Actuation. Initiation of movement of machinery can be achieved by pneumatic or hydraulic fluid power, cams or electric driving motors, solenoids, or other rotary or linear actuators. Fluid power is the least costly to build, but it has the highest operating cost and often has high noise levels. Such machines usually have sequential controls. Cam-operated machines are quiet and highly reliable, and over prolonged periods they have low operating costs. They require more engineering and have a higher purchase price. Electrically actuated machines are common in Europe and Asia. They have very complex control systems including significant interlocking systems. These machines are often specified because they are perceived to have a high level of flexibility.

Machine design. There are significant differences in philosophy of machine design between European, Asian, and North American builders and between in-house and commercial builders. In-house builders (more typically found in Europe and Asia) tend to consider each assembly machine project a unique task, while commercial builders consider assembly a set of generic tasks capable of being done by generic modular stations. Commercial builders claim greater operational up time because of the reliability of proven modules, and greater flexibility for product modification because of the generic station design.

Tasks. Assembly machines perform several tasks: parts feeding (including orientation, separation, and transfer); parts joining (for example, welding, riveting, and soldering); parts inspection (condition, parameters, presence, and position); functional testing (for example, capacitance, torque, and pressure decay); marking (date coding, model number, and operational characteristics); and ejection (in controlled or uncontrolled positions). Many of the tasks of assembly machinery such as joining, functional testing, and product marking use commercial units identical to those used on manual lines. Commercial screwdrivers, orbital riveters, ultrasonic welders, vision systems, leak testing, printers, laser markers, and other commercial units can be adapted to automatic assembly machinery with little or no modification.

The unique task of automatic assembly machinery is that of parts feeding, which consists of accepting the component parts as they come to the assembly area and taking one of these components, separating that component from other parts, changing its spatial orientation to a usable insertion attitude, and transferring it from the orientation device into the work-holding fixture or into a partially completed assembly.

Some parts can be fabricated directly on the assembly line. This is true of paper or film components such as gaskets, diaphragms, and labels. Material is brought to the machine in reels and is fabricated directly on the machine. Some fragile parts are often preformed and rewound in reels or bandolier devices for cutoff or separation at the assembly machine. Transistors, resistors, and capacitors are often fed this way, although many are fed from special tubes. Many fragile mechanical parts or subassemblies usually stored in X-Y coordinate trays can also be stored more efficiently in plastic chainlike bandoliers.

Magazines. These are often used to hold flat parts or to store fragile subassemblies, and take many shapes. Certain parts that are not easily damaged can be fed by pushing one part from the stack with a technique referred to as coin changing. Parts such as motor or transformer laminations are usually fed in this mode. Other more fragile parts are fed from elevator-type magazines in which one part is lifted from the top of the stack without lateral motion relative to the other parts in the stack.

All magazine feeding systems are dependent on several technical and economic conditions. Of the economic questions, the most critical is proximity to the fabrication operations. Magazine racks, tubes, and containers are expensive, and the quantity required is proportional to the distance from fabrication to assembly. Technically, magazines are the most desirable option when the component can be ejected from a primary operation directly to the assembly machine magazine storage device without any secondary operation. Parts requiring secondary operations or manual loading of magazine tubes do not enjoy the same savings as those in which magazine loading is done at the primary fabrication operations.

Most parts coming to the assembly floor do not come in bandoliers or in convenient magazine carriers. Fragile parts are usually stored in X-Y coordinate trays or boxes requiring some level of coordinate programming (including robots) for parts retrieval. This storage is very expensive and adds greatly to machine development costs. Larger parts such as engine blocks, transmission housings, and motor frames are usually stored between fabricating, surface finishing, and assembly in overhead storage conveyors. Retrieval of such parts is often manual but is increasingly done by robots.

Parts feeders. Most of the parts coming to the assembly floor arrive in unoriented bulk condition. While ideally such parts should be captured at fabrication and carried in an oriented condition, economic reality usually makes this impractical. Many parts require a variety of secondary operations. Vast differences between fabrication cycle rates and assembly rates often dictate multiple fabricating machines, molds, or dies. Components often come to the assembly floor from outside vendors quite remote from the assembly area. When such parts are durable and the pipeline of supply is extensive, bulk unoriented shipment of components is the normal mode of delivery.

To be able to feed such parts, a variety of feeders have been developed over the years. Basically such parts feeders, operating by mechanical, vibratory, or centrifugal means, move a large number of unoriented parts and discharge them onto selective devices in feeder bowls or discharge tracks. Those parts with usable attitudes are captured and passed along to the discharge point of the feeder; those without are returned to the storage area of the feeder to try again.

Most early feeders used mechanical rotors or oscillating blades to move the parts. In some specialized applications, centrifugal forces move the parts. In most modern feeders, vibratory action is used to produce the motion of components. Vibratory feeders consist of bowls, trays, or tracks mounted on a series of springs at an angle to the plane of the bowl, tray, or track. Electromagnets are used to pull down the feeders, compressing the springs until the magnet is turned off, freeing the spring, and causing the parts container to move upward. This repeated cycling causes the components to be thrown upward and forward in either a circular or a linear unidirectional path. The parts, if properly oriented, proceed through a series of capturing and retaining devices until they are fed to the discharge point. The parts to be discharged are moved sequentially into an escapement device. Escapements are used to separate the part that is to be transferred and inserted from subsequent parts in feeder tracks. Escapements must accurately locate parts for pickup and transfer.

Unique aspects. Unlike fabricating machines, assembly machinery must handle discrete parts with significant variations in dimensional tolerance. All assembly machinery must be debugged prior to operation. This debugging is a normal component of assembly machine development, ensuring that the

machine is capable of accepting a given spectrum of variations in component parts.

Assembly machines are usually one-of-a-kind systems, and they need extensive sensor networks to ensure that the machine functions are working properly. Few products are manufactured in such volumes as to permit full development of process reliability. The prototype nature of most assembly machinery requires that machine controls contain a large element of process verification not usually found in other industrial machinery. *See* MANUFACTURING ENGINEERING; PRODUCTION METHODS. Frank Riley

Bibliography. P. Niland, *Management Problems in the Acquisition of Special Automatic Equipment*, 1961; F. J. Riley, *Assembly Automation: A Management Handbook* 2d ed., 1996.

Astatine

A chemical element, At, atomic number 85. Astatine is the heaviest of the halogen groups, filling the place immediately below iodine in group 17 of the periodic table. Astatine is a highly unstable element

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

existing only in short-lived radioactive forms. About 25 isotopes have been prepared by nuclear reactions of artificial transmutation. The longest-lived of these is ^{210}At, which decays with a half-life of only 8.3 h. It is unlikely that a stable or long-lived form will be found in nature or prepared artificially. The most important isotope, used for tracer studies, is ^{211}At. Astatine exists in nature in uranium minerals, but only in the form of trace amounts of short-lived isotopes, continuously replenished by the slow decay of uranium, The total amount of astatine in the Earth's crust is less than 1 oz (28 g).

In aqueous solution, astatine resembles iodine except for differences attributable to the fact that astatine solutions are of necessity extremely dilute. Like the halogen iodine, when astatine exists as a free element in solution, it is extracted by benzene. The element in solution is reduced by agents such as sulfur dioxide and is oxidized by bromine. It is more electropositive than the other halogens. It has oxidation states with coprecipitation characteristics similar to those of the iodide ion, free iodine, and the iodate ion. Powerful oxidizing agents produce an astatate ion, but not a perastatate ion. The free state is most readily obtained and is characterized by high volatility and high extractability into organic solvents. *See* HALOGEN ELEMENTS. Earl K. Hyde

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. F. Shriver and P. W. Atkins, *Inorganic Chemistry*, 3d ed., 1999.

Asterales

An order of flowering plants, division Magnoliophyta (Angiospermae), which gives its name to the subclass Asteridae in the class Magnoliopsida (dicotyledons). The Asterales have often been included in the order Campanulales, but they are perhaps more closely allied to the Rubiales and Dipsacales. The order consists of only the single very large family Asteraceae (Compositae), with about 20,000 species, occurring in nearly all parts of the world but most abundant and conspicuous in areas which are not densely forested. *See* CAMPANULALES; DIPSACALES; RUBIALES.

The Asterales are marked by their inferior ovary, single basal ovule, specialized pollen presentation mechanism, and pseudanthial, centripetally flowering heads which often have specialized marginal flowers with a strap-shaped corolla resembling the petal of an ordinary flower. The anthers are introrsely dehiscent and connate (or at least connivent) into a tube around the style, which grows up through the anther tube and pushes out the pollen. A more or

New England aster (*Aster novae-angliae*), a characteristic member of the order Asterales. The numerous apparent petals around margin of flower head are corollas of individual flowers, called ray flowers in contrast to disk flowers which occupy the center of head. (*Courtesy of Alvin E. Staffen, National Audubon Society*)

less similar pollen presentation mechanism occurs in many of the Campanulales and some of the Rubiales, in the same subclass.

Most members of the order are herbaceous, but some, such as the sagebrush (*Artemisia tridentata*), are shrubs, and a few tropical species are trees. Many well-known garden ornamentals, such as aster (see **illus.**), chrysanthemum, dahlia, daisy, sunflower (*Helianthus*), and zinnia, belong to the Asterales. A few garden vegetables, for example, lettuce (*Lactuca*) and artichoke (*Cynara*), and some common weeds, such as dandelion (*Taraxacum*), thistle (*Cirsium*), and ragweed (*Ambrosia*), also belong to the order. *See* ARTICHOKE; ASTERIDAE; LETTUCE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; ORNAMENTAL PLANTS; PLANT KINGDOM; SUNFLOWER.

Arthur Cronquist; T. M. Barkley

Asteridae

A large subclass of the class Magnoliopsida (dicotyledons) of the division Magnoliophyta (Angiospermae), the flowering plants, consisting of 11 orders, 49 families, and more than 60,000 species. The Asteridae are mostly sympetalous with unitegmic, tenuinucellate ovules and with the stamens usually as many as, or fewer than, the corolla lobes and alternate with them. Most of them have two carpels, but a few have as many as five or even more carpels, and a few others are pseudomonomerous. The largest orders of the group are the Asterales (about 20,000 species), Scrophulariales (about 11,000 species), Lamiales (about 7800 species), and Rubiales (about 6500 species). Other orders are the Gentianales, Plantaginales, Solanales, Callitrichales, Campanulales, Calycerales, and Dipsacales. See individual articles on each order. *See* MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Asteroid

One of the many thousands of small planets (minor planets) revolving around the Sun, mainly between the orbits of Mars and Jupiter. The presence of a gap in J. A. Bode's empirical law of planetary spacings motivated a search for the missing planet. G. Piazzi discovered Ceres on January 1, 1801, and three other small planets were discovered in the next few years. Visual and photographic searches for additional asteroids have continued to the present day, augmented by electronic detection technology. Newly discovered ones are assigned a catalog number and name (such as 433 Eros) only after they are observed often enough to compute an accurate orbit. The Minor Planet Center in Cambridge, Massachusetts, maintains a file of 21 million measurements of asteroid positions. The Institute for Theoretical Astronomy in St. Petersburg, Russia, publishes an annual ephemeris of predicted asteroid positions for the over 73,000 cataloged asteroids. *See* PLANET.

Orbits. The majority of asteroids have semimajor axes (mean distances to the Sun; symbolized a) between 2.2 and 3.2 astronomical units (1 AU = distance from Earth to the Sun = 1.496×10^8 km = 9.3×10^7 mi). However, numerous small asteroids orbit between Venus and Mars, and two large groups, the Trojan asteroids, orbit at Jupiter's distance from the Sun. *See* TROJAN ASTEROIDS.

In 1992, the first of the trans-Neptunian "asteroids" was discovered. Called Kuiper Belt Objects (KBOs) and Scattered Disk Objects, nearly 800 had been found by early 2004. They represent a population of bodies much more numerous than the main-belt or Trojan asteroids, but are more properly thought of as comets. They originally accreted from the primordial solar nebula (the disk of gas and dust from which the planets formed) beyond the orbit of Neptune, but never were incorporated into an outer planet. As a few of these objects slowly leak away and enter the inner solar system closer to the Sun, their ices sublime and they develop the heads and tails characteristic of comets. There are also a modest number of minor planets orbiting the Sun in temporary orbits beyond Jupiter but well inside the Kuiper Belt; they are termed Centaurs. Although there may be no sharp distinctions in physical properties between asteroids and outer solar system small bodies, and some old short-period comets may be counted among the near-Earth asteroids (NEAs), this article deals with classical asteroids that originated in the inner parts of the solar system: the Trojans, the main-belt asteroids, and the NEAs, which approach the Sun to within 1.3 AU. *See* COMET; KUIPER BELT.

Most asteroid orbits are more elliptical and inclined to the plane of the ecliptic than the orbits of major planets. Eccentricities (e) average about 0.15, and inclinations (i) about 10°; occasionally they exceed 0.5 and 30°, approaching the characteristics of short-period comet orbits. Near-Earth asteroids are of three different orbital types: Amors orbit beyond Earth's orbit and do not cross it; Apollos cross Earth's orbit but generally do not intersect it; and Atens orbit inside Earth's orbit.

Asteroids are not uniformly distributed in a, e, and i. Vacant lanes in the main asteroid belt, known as Kirkwood gaps (**Fig. 1**), occur at distances where the periods of revolution would be a simple fraction (such as $^1/_3$ or $^2/_5$) of Jupiter's period of 11.86 years. Asteroids in these gaps may have been preferentially removed when Jupiter's powerful gravity sent them into wild, chaotic orbits; in these orbits, they may have collided with the Sun, the Earth, or another planet. Most asteroids originally beyond the $^1/_2$ resonance were directly ejected from such orbits by nearby Jupiter early in the history of the solar system, except those grouped near the stable $^2/_3$ (Hilda group) and $^3/_4$ (Thule) resonances. Asteroid fragments continue to be ejected from the $^1/_3$ gap and from other resonant orbits. These chaotic processes, combined with forces due to asymmetric solar heating, are believed to be responsible for bringing most of the Earth-approaching asteroids and smaller meteorites close to the Earth. *See* CHAOS.

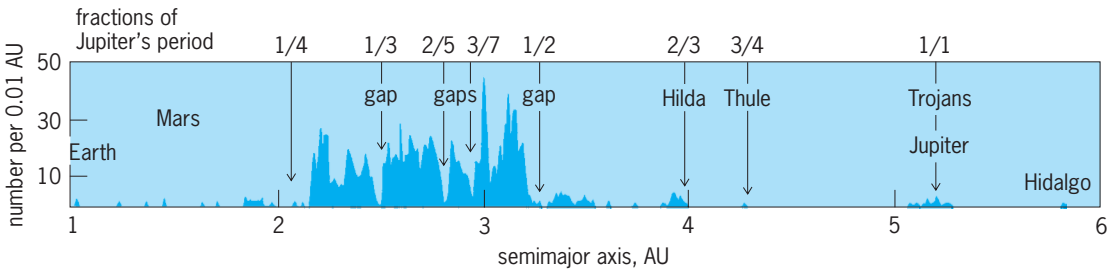

Fig. 1. Distribution of the brighter asteroids, with distance from the Sun, between the Earth and Jupiter. Major fractions of asteroid orbital periods to Jupiter's period are shown. The arrows point to associated clusters or Kirkwood gaps.

Clusterings of asteroids with similar *a, e*, and *i* are known as Hirayama families, named for the Japanese astronomer K. Hirayama who discovered some of the major ones in 1918. A few dozen families can be recognized. Some of the larger families are composed of asteroids that have the same colors, and are presumably made of the same minerals. Such a family probably consists of fragments of a larger, precursor asteroid that was broken apart by a catastrophic collision with another asteroid or comet. Some of these families are immersed in huge, torus-shaped dust belts, discovered by the *Infrared Astronomical Satellite* (*IRAS*). The colliding asteroids are gradually grinding each other down to dust, which then spirals into the Sun or is blown away by light pressure from the Sun. *See* ORBITAL MOTION.

Shapes, spins, and satellites. When the brightnesses of most asteroids are measured, they vary in a periodic manner, occasionally by more than one magnitude (a factor of 2.5), but more commonly 0.3 magnitude or less. The light curves usually are double-peaked (**Fig. 2**), characteristic of an irregularly shaped body spinning in space rather than the more complex curves that would result from albedo differences (spots) on a spherical body. Light curves have been measured for many hundreds of asteroids.

The spin periods are typically a few hours to about a day, but range from about 2 h to roughly a month. Asteroids smaller than about 500 ft (150 m) across spin much more rapidly, indicating that they are solid rocks. The fact that larger asteroids never spin faster than 2 h indicates that they are weak bodies, perhaps rubble piles held loosely together by gravity.

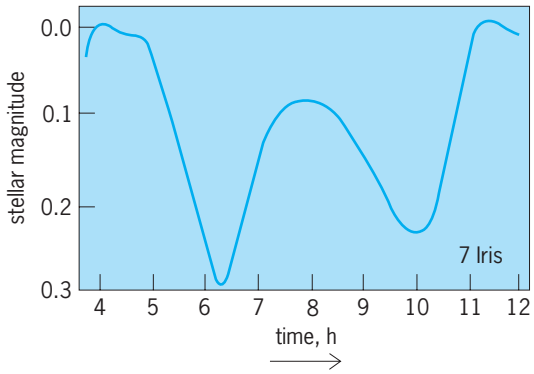

Fig. 2. Variation with time in the brightness of 7 Iris due to its irregular shape and 7.1-h rotation. (*After G. Kuiper*)

If they were to spin faster, the centripetal acceleration at their equator would exceed their weak gravity and they would tend to fly apart. A few large asteroids spin very slowly; they are believed to have been slowed by asymmetric solar heating and reradiation. Collisions among asteroids are not so important in controlling the spins of medium-sized asteroids as was once thought. A few asteroids have variable periods or they wobble and tumble chaotically; some of them may have been struck recently by another asteroid.

The very largest asteroids, especially those composed of weak materials, cannot maintain highly irregular shapes against the compression of their own gravitational fields, and are roughly spherical. Some of the larger, rapidly spinning asteroids have large light-curve amplitudes. They may be weak "rubble piles," previously fractured and partly disrupted by collisions and immediately reaccumulated into a pile of boulders. Some small asteroids, notably 4179 Toutatis, which was studied intensively by radar when it passed close to the Earth in 1992, have been found to have dumbbell shapes. They may actually be two bodies, lightly resting against each other.

In August 1993, the *Galileo* spacecraft flew past the main-belt asteroid 243 Ida and discovered that it possesses a mile-wide satellite, which was subsequently named Dactyl. Five years later, a second asteroidal satellite was discovered around 45 Eugenia with a telescope in Hawaii, using new adaptive optics technology. Since then (through early 2004), 13 more main-belt asteroids and one Trojan have been found by adaptive optics either to be a double object (binary) or to possess a satellite. Also, 10 near-Earth asteroids have been observed by radar to be binary or have a satellite, and another 10 near-Earth asteroids (and two more main-belt asteroids) show unusual light curves that strongly suggest a binary configuration. Only a few percent of main-belt asteroids have substantial satellites, which were probably produced by catastrophic collisions or impact cratering events. The fraction is much higher, maybe 20%, for near-Earth asteroids, most of which may result from tidal disruption during close passes to the Earth or Venus. *See* ADAPTIVE OPTICS; BINARY ASTEROID.

Sizes and densities. Improvements in radar technology make it possible to image small asteroids that pass close to the Earth almost as well as by spacecraft flybys. With the exception of these observations, no asteroid shows a disk large enough to measure

Noteworthy asteroids

Size rank	Number and name	Spectral type	Diameter, mi (km)*	Spin period, h	Orbital elements *a*, AU	*e*	*i*, degrees
1	1 Ceres	G (C-like)	590 (949)	9.1	2.77	0.08	10.6
2	2 Pallas	B (C-like)	331 (533)	7.8	2.77	0.23	34.8
3	4 Vesta	Achondrite	329 (529)	5.3	2.36	0.09	7.1
4	10 Hygiea	C	253 (407)	27.6	3.14	0.12	3.8
5	511 Davida	C	200 (322)	5.1	3.18	0.17	15.9
6	704 Interamnia	F (C-like)	197 (317)	8.7	3.06	0.15	17.3
7	52 Europa	C	183 (295)	5.6	3.09	0.11	7.5
8	87 Sylvia	P	172 (277)	5.2	3.48	0.09	10.9
9	65 Cybele	P	167 (269)	4.0	3.43	0.11	3.6
10	15 Eunomia	S	161 (259)	6.1	2.64	0.19	11.8
11	16 Psyche	M	155 (249)	4.2	2.92	0.14	3.1
12	31 Euphrosyne	B (C-like)	154 (248)	5.5	3.15	0.23	26.3
13	451 Patienta	B (C-like)	153 (247)	9.7	3.06	0.07	15.2
14	3 Juno	S	150 (242)	7.2	2.67	0.25	13.0
15	324 Bamberga	C	149 (240)	29.4	2.68	0.34	11.2
16	13 Egeria	G (C-like)	139 (224)	7.0	2.58	0.09	16.5
17	624 Hektor	D	186 × 93 (300 × 150)	6.9	5.15	0.03	18.3
18	532 Herculina	S	137 (220)	9.4	2.77	0.17	16.3
19	107 Camilla	C	137 (220)	4.8	3.49	0.07	10.0
20	423 Diotima	P?	135 (217)	4.8	3.07	0.03	11.2
21	121 Hermione	C	135 (217)	5.6	3.46	0.14	7.6
22	45 Eugenia	F (C-like)	133 (215)	5.7	2.72	0.08	6.6
23	19 Fortuna	C	130 (210)	7.4	2.44	0.16	1.6
24	24 Themis	C	129 (207)	8.4	3.13	0.13	0.8
25	7 Iris	S	127 (204)	7.1	2.39	0.23	5.5
26	6 Hebe	S	126 (202)	7.3	2.42	0.20	14.8
27	702 Alauda	C	126 (202)	8.4	3.19	0.03	20.5
28	88 Thisbe	C	124 (200)	6.0	2.77	0.16	5.2
			Other interesting asteroids				
	41 Daphne	C	116 (187)	6.0	2.77	0.27	15.8
	44 Nysa	Aubrite?	42 (68)	6.4	2.42	0.15	3.7
	165 Loreley	C	99 (160)	7.2	3.14	0.07	11.2
	216 Kleopatra	M	87 (140)	5.4	2.79	0.25	13.2
	243 Ida	S	18.6 × 7.9 × 5.8 (29.9 × 12.7 × 9.3)	4.6	2.86	0.04	1.1
	250 Bettina	M	53 (86)	5.1	3.14	0.14	12.9
	349 Dembowska	Achondrite?	90 (145)	4.7	2.93	0.09	8.3
	433 Eros	Chon.	21.4 × 7.0 × 7.0 (34.4 × 11.2 × 11.2)	5.3	1.46	0.22	10.8
	747 Winchester	P	116 (186)	9.4	3.00	0.34	18.2
	951 Gaspra	S	11.3 × 6.5 × 5.5 (18.2 × 10.5 × 8.9)	7.0	2.21	0.17	4.1
	1566 Icarus	Chon.?	1 (2)	2.3	1.08	0.83	22.9
	1620 Geographos	S	1.3 (2.1)	5.2	1.24	0.33	13.3
	4179 Toutatis	S	~2 (4)	130	2.51	0.64	0.5

*Diameters are accurate to around 10%; size rankings may vary.

accurately by using Earth-based telescopes. For more distant asteroids the chief technique used to measure asteroid diameters is radiometry, which compares the brightness of reflected visible sunlight from an asteroid with the brightness of the asteroid's emitted thermal radiation in the infrared. In the mid-1980s, *IRAS* measured the diameters and albedos of over 2000 asteroids. Radiometric diameters have been checked by timings of occasional disappearances of stars when asteroids chance to pass between the Earth and a star. *See* ALBEDO; INFRARED ASTRONOMY; OCCULTATION.

There are about 30 asteroids larger than 124 mi (200 km) in diameter (see **table**); about 75% of them are soot black (geometric albedos of 3–5%). Asteroids are much more numerous at smaller sizes, generally following a size distribution characteristic of fragmentation processes, as would be expected if the asteroids were smashing into each other. Indeed, there are so many asteroids confined in the volume of the asteroid belt that collisions sufficient to fragment medium-sized asteroids occur every few billion years, and much more often for smaller ones. Thus all asteroids have been extensively battered, and most of the smaller ones are collisional fragments.

A few of the largest asteroids are massive enough to detectably change the orbits of other asteroids and even the orbit of Mars, from which the masses of the large asteroids can be determined. The total mass of all asteroids is less than 5% that of the Moon. For asteroids with measured masses and diameters (volumes), their bulk densities can be calculated. The resulting densities of Ceres, Pallas, and Vesta are roughly 2.1, 2.7, and 3.5 times that of water, respectively. Vesta's density is similar to that of ordinary rocks, but Ceres is as underdense as carbonaceous chondritic meteorites, which contain a large fraction of volatiles. Perturbations on the orbit of Mars, combined with estimates of densities from the orbital periods of some asteroidal satellites, suggest that most C-type

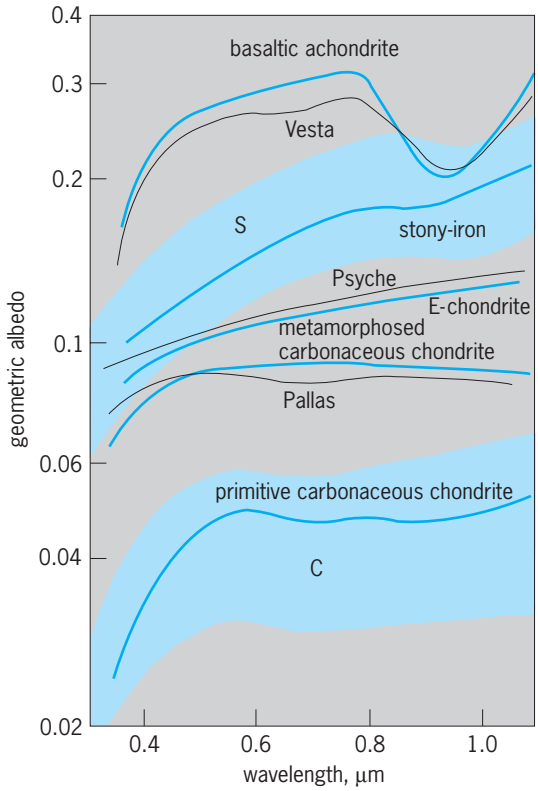

Fig. 3. Visible and near-infrared reflection spectra of the major S and C compositional classes of asteroids, plus several unusual ones. Laboratory spectra for some similar meteorites are shown with colored lines.

asteroids have densities less than 1.5, even more porous than carbonaceous chondrites. Whether this is due to inherently lightweight or porous materials or due to large voids in a rubble-pile structure, or a combination, is not yet known.

Surface compositions. Spectra of sunlight reflected from asteroids have shapes, including absorption bands, characteristic of different rock-forming minerals (**Fig. 3**). Combined with the albedo data from radiometry, the spectral colors of surfaces of thousands of asteroids show that more than three-quarters of them have very low albedos and are composed of carbon-rich material (often with hydrated, or water-rich, minerals). The black asteroids located in the middle and outer parts of the belt (called C type) resemble carbonaceous meteorites, which are believed to be among the most primitive materials in the solar system, little altered since the planets were forming. The black asteroids near the outer edge of the main belt, and most of the Hildas, have a reddish tinge and are not represented by known meteorites on the Earth; they are called P types, and may be even richer in organic components. Still farther out, many of the Trojans are even redder and more mysterious; they are termed D types.

Closer to the inner edge of the belt, most asteroids are so-called S types, characterized by moderately high albedos and by absorption bands due to the common silicate minerals pyroxene and olivine. They also contain considerable metal, and probably are akin to either the stony-iron meteorites or the ordinary chondritic meteorites. The spectral data alone favor stony-iron compositions, but such compositions would leave no other bodies remaining as sources for the ordinary chondrites, which constitute approximately 85% of all meteorite falls. It appears that the spectral properties of the surfaces of asteroids are altered by a process called space weathering. Micrometeorite impacts and other effects alter the colors of the surface regoliths (fragmental soils), changing the spectral signature of the meteoritic bed rock. Only asteroids too small to retain appreciable regoliths—less than about 1 km in diameter—look sensibly like ordinary chondrite meteorites, but it may be presumed that many of the larger S-types are also ordinary chondrites. *See* METEORITE; REGOLITH.

The general progression of asteroid compositions, from S types in the inner belt, to C's, to P's, and then to D's at Jupiter's distance, is thought to reflect the variation with distance from the Sun in the composition of the original nebular dust from which the planets were formed. *See* COSMOCHEMISTRY; SOLAR SYSTEM.

There are also some less common asteroid compositions, which are nevertheless interesting because they imply that some of the asteroids were somehow heated to the point that they melted. The asteroid Vesta is apparently mostly covered with lava flows, made of the same minerals as the basaltic achondritic meteorites. When a planetary body is heated to the point that volcanic lavas flow onto its surface, metal sinks to the center, forming a core. A number of asteroids, notably 16 Psyche, have so-called M-type spectra and radar echoes suggesting they are made of solid metal. Astronomers believe that Psyche's parent body was melted into a body that looked like Vesta does today, with a basaltic surface, an olivine-rich mantle, and a metallic core. Then collisions fragmented and stripped away the rocky crust and mantle, exposing the core.

There are questions about this picture, however. If there have been enough powerful collisions to strip some asteroids down to their metallic cores, it is strange that Vesta's crust has been so well preserved. In fact, Vesta is in the middle of a family of small, basaltic asteroids, apparently pieces of its crust blasted away by a cratering impact. Pictures of Vesta taken by the Hubble Space Telescope have revealed the huge crater created by this impact. But other asteroids should have suffered collisional fragmentation intermediate between Vesta and Psyche, and the fate of the pieces of olivine mantles of such asteroids remains a mystery. Indeed, only a few small, so-called A-type asteroids have been found that show the expected spectral signature of the supposedly abundant olivine.

Surface conditions and geology. Subsolar surface temperatures of main-belt asteroids mostly range from −80 to +10°F (210 to 260 K). The low gravity of asteroids prevents them from retaining any atmospheric gases, so their surfaces are exposed directly to interplanetary space where they are

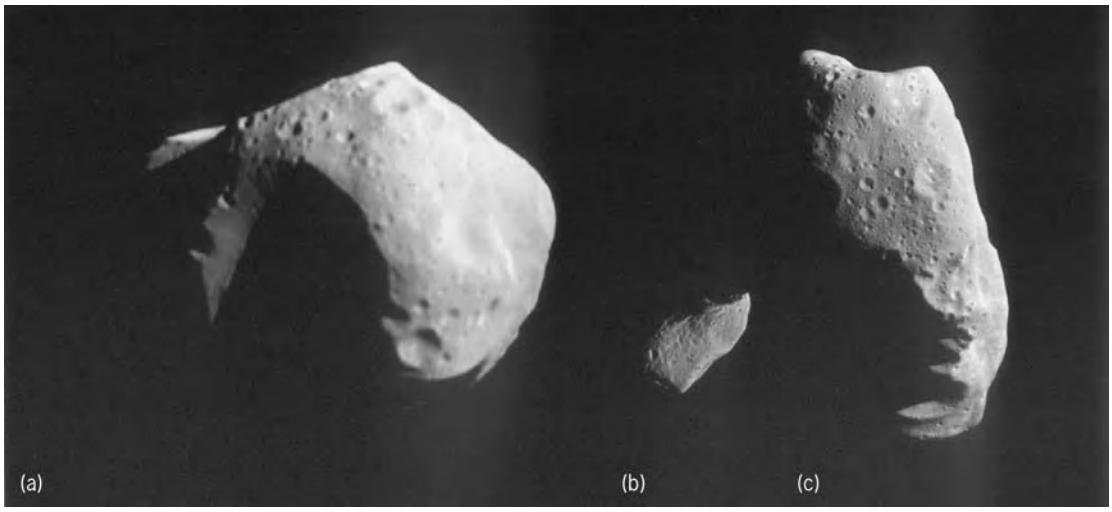

Fig. 4. Main-belt asteroids imaged by spacecraft. (*a*) Mathilde imaged by *Near Earth Asteroid Rendezvous* (*NEAR*) spacecraft on June 27, 1997. (*b*) Gaspra imaged by *Galileo* spacecraft on October 29, 1991. (*c*) Ida imaged by *Galileo* on August 28, 1993. All three objects are presented at the same scale, but relative brightnesses have been altered. Mathilde is actually much darker than the other two asteroids and has more large craters. (*NASA*)

impinged by the solar wind and micrometeorite bombardment. Polarimetric measurements suggests that asteroid surfaces are dusty, and the larger asteroids are expected to retain a fraction of the material ejected from cratering impacts, which accumulates into thick deposits of regolith.

Two S-type main-belt asteroids have been studied close-up by the *Galileo* spacecraft, which flew past them on the way toward its 1995 rendezvous with Jupiter. The encounter of 951 Gaspra on October 29, 1991, revealed an irregularly shaped body, 11.3 × 6.5 × 5.5 mi (18.2 × 10.5 × 8.9 km) in size (**Fig. 4**). Excellent data were obtained from the encounter on August 28, 1993, with the somewhat larger asteroid 243 Ida, a member of the populous Koronis family (Fig. 4). Ida's surface is covered by many more craters, especially moderate- and large-sized ones, than is Gaspra's, implying that Ida's creation as a fragment of the Koronis parent body may have happened billions of years ago. *Galileo*'s magnetometer observed magnetic anomalies as the spacecraft flew near both Gaspra and Ida, revealing unexpected interactions of the asteroids with the solar wind, possibly due to remanent magnetization of the asteroids.

A very different asteroid is the dark, 32-mi (52-km) main-belt object 253 Mathilde (Fig. 4). It is pockmarked by giant craters, five of which have diameters similar to or larger than the radius of Mathilde itself. They may result from crushing of Mathilde's porous material, rather than explosive excavation by impacting projectiles.

Earth-approaching asteroids. Apollo, Amor, and Aten asteroids are of special interest, particularly because they stand a chance of striking the Earth. Indeed, Meteor Crater (Arizona), and other craters on the Earth and the Moon, testify to the potential for collisions with near-Earth asteroids. Many scientists believe that just such a collision 6.5×10^7 years ago rendered most species of life, including the dinosaurs, extinct. A huge, eroded crater of that age on the Yucatán Peninsula in Mexico must have been caused by the impact of an asteroid or comet about 10 mi (16 km) in diameter.

Impacts by near-Earth asteroids are not confined to the past. In 1908, a small asteroid, perhaps 160 ft (50 m) across, exploded over the Tunguska region of Siberia with energy equivalent to 15 megatons of TNT. It has been estimated that the impact of a mile-wide asteroid could loft enough particles into the stratosphere to dim the Sun, disrupt agriculture, and cause worldwide famine. Such an impact has a 1-in-500,000 chance of occurring each year. Only about one-half of such potentially threatening objects have been discovered so far.

These small asteroids have not always been in Earth-approaching orbits. After perhaps 10 million years, most of the current crop will have struck the Earth, the Moon, the Sun, or one of the other inner planets, or will have been ejected from the solar system. Most are probably fragments of main-belt asteroids, traveling in chaotic orbits, just like their smaller cousins, the meteorites. (In fact, some meteorites are probably chips broken off Earth approachers.) A few near-Earth asteroids are dead comet nuclei, which have lost their ices after thousands of years spent close to the Sun in the inner solar system.

One of the largest near-Earth asteroids, 433 Eros, was the first to be visited by a spacecraft. *NEAR-Shoemaker* orbited around Eros for a year, before landing on its surface, providing the most detailed information ever obtained about the chemistry, mineralogy, and geology of an asteroid. Eros has the composition of ordinary chondrites. Seen up close (**Fig. 5**), the asteroid's surface lacks small craters but is covered with boulders. In places, there are mysterious, flat, dry "ponds".

Origin and evolution. Current cosmogonical models for the origin of planets involve accretion from

Fig. 5. Views of the two opposite sides of the asteroid 433 Eros, imaged from the *NEAR-Shoemaker* spacecraft. (*NASA*)

myriads of asteroidlike planetesimals. It is likely that asteroids are a remnant of the planetesimals that failed to accrete into a planet between Mars and Jupiter. Perhaps shifting positions of the young planets, including massive, nearby Jupiter, increased the relative velocities of asteroids to the present value of 3 mi/s (5 km/s) so that asteroids fragment rather than accrete when they meet each other. Instead of forming a planet, the asteroids have been smashing each other to bits.

Evidently some asteroids of primitive, nonvolatile solar composition were heated within the first few hundred million years after the origin of the solar system, perhaps by the solar wind or extinct radionuclides, and they melted. While the unmelted, weak, C-type asteroids may have been depleted by a large factor by collisions, most of the strong stony-iron cores of the melted proto-asteroids have survived; perhaps they are among the M- and S-type asteroids observed today. The asteroids still collide and fragment. The fragments drift due to solar forces and enter resonances, which ultimately send some of them into the inner solar system, where they produce craters or fall as meteorites.

This picture of the evolution of the asteroids is incomplete and tentative. As new data are collected, research continues on alternative interpretations.

Clark R. Chapman

Bibliography. J. K. Beatty, C. C. Peters, and A. Chaikin (eds.), *The New Solar System*, 4th ed., Sky Publishing Corp. and Cambridge University Press, 1999; W. F. Bottke, Jr., et al. (eds.), *Asteroids III*, University of Arizona Press, 2002; H. Y. McSween, Jr., *Meteorites and Their Parent Planets*, 2d ed., Cambridge University Press, 1999; J. L. Remo (ed.), Near-Earth Objects: The United Nations Conference, *Ann. N.Y. Acad. Sci.*, vol. 822, 1997.

Asteroidea

A subclass of the class Stelleroidea in the subphylum Asterozoa; asteroids are known as starfish (**Fig. 1**). The arms are not sharply demarcated from the rest of the body. The ambulacral ossicles never fuse to form vertebrae. The tube feet are locomotor organs, usually suctorial, emerging from an open ambulacral groove. The dominant growth gradients are such as to cause the skeletal ossicles to lie in longitudinal rows known as series (for example, adambulacral series, ventrolateral series, and inferomarginal series).

For many years the Asteroidea and the Ophiuroidea were accorded separate class status, but in 1951 W. K. Spencer showed that they have a common origin. Most differences between these groups disappear when fossil forms are considered. Thus both groups are now ranked as subclasses of Asterozoa. The 1700 or so known fossil and recent species of Asteroidea are now grouped into six orders: Platyasterida, Paxillosida, Valvatida, Forcipulatida, Notomyotida, and Trichasteropsida. Asteroids range in size from about 0.4 to 40 in. (10 mm to 1 m) across. Many starfish are brightly colored and attractive animals, but some are dowdy and cryptic. Their conjugated carotenoid pigments fade on preservation. Like most echinoderms, starfish have a lifespan of about 5 years. *See* OPHIUROIDEA.

Economic importance. Several species of starfish are regarded as pests, particularly in oyster, mussel, and scallop fisheries. Beyond this they are of little significance, with the notable exception of the coral predator *Acanthaster planci*, which occurs in the Indo-Pacific. When present in large numbers, it may kill extensive areas of reef-building corals. However, large aggregations of this species are occasional phenomena, and coral killed by it is capable of regenerating more quickly than was hitherto supposed.

Ecology. Starfish inhabit all types of bottom throughout the world's seas and oceans. Some burrow in sand and mud, and others live on rocks and

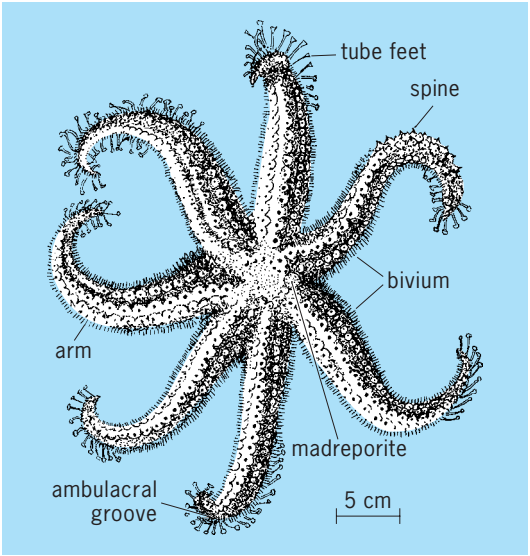

Fig. 1. A representative asteroid, *Astrostole scabra*.

coral reefs, where they are at their most diverse. *Calvasterias* may live at the surface, attached to floating weeds. Most families show well-defined bathymetrical preferences, but there are exceptions, for instance, the offshore fauna of New Zealand. There are 12 families which occur at depths greater than 2 mi (3 km), and several genera extend below 4 mi (6.5 km).

Parasitism and commensalism. Asteroids are parasitized by a variety of organisms. Small sea snails of the genus *Stylifer* bore into their tissues. A polychaete worm *Acholoë astericola* lives in the ambulacral groove of *Astropecten*; a cirripede castrates *Cascinasterias* by destroying the gonad; and a slender fish, *Fierasfer*, inhabits the body cavity of *Culcita*. This appears to enter the mouth and then bores through the stomach wall. A remarkable insect-asteroid association has been reported from Australia, where a marine caddis fly lays its eggs in the tissues of *Patiriella*.

Anatomy. The outer surface is coated with a thin layer of ciliated epithelium. Below this lies connective and muscle tissue which forms the body wall and in which the skeletal ossicles are embedded, forming the test. In some species the skeletal component is very considerable, making them heavily armored and relatively rigid; in others it is less, so they appear lighter and more flexible. Some of the ossicles, notably the spines, protrude to the exterior (**Fig. 2**). The arms can be slowly moved by the body wall musculature. Although 5 arms are common, up to 12 occur in a number of genera, and more than 12 may occur, as in *Acanthaster*.

Pedicellariae occur in some starfish, but these have evolved independently from echinoid pedicellariae. They are small seizing organs used to grip intruders and, according to one report, even to catch food. They probably assist in protecting the delicate epithelium from predators and sediments. Pedicellariae are formed from two or four modified spines which are snapped together by adductor muscles like minute tongs. Their shape and type is important in classification because different orders of starfish have evolved different forms of pedicellariae. In

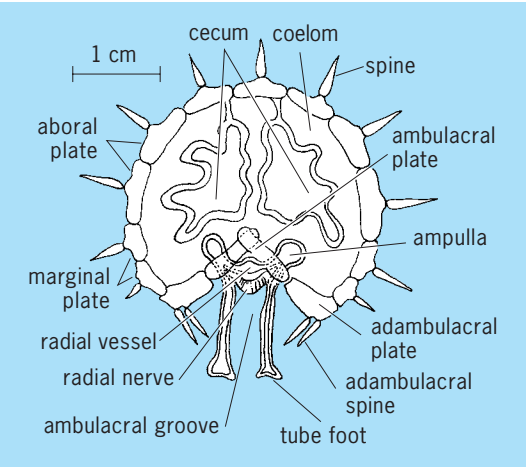

Fig. 2. Transverse section through arm of asteroid.

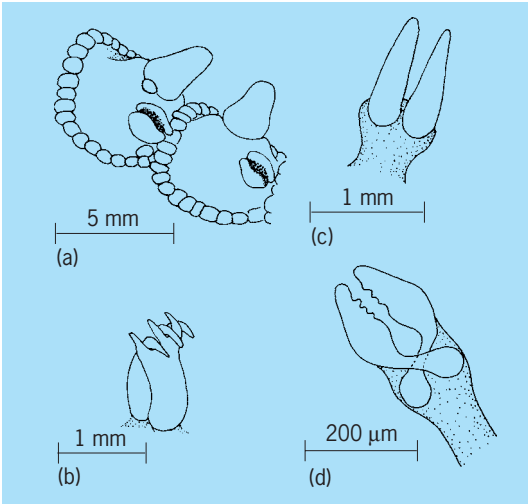

Fig. 3. Pedicellariae of asteroids. (*a*) Valvate type on marginal plates of *Hippasteria trojana*. (*b*) Felipedal type in *Cosmasterias dyscrita*. (*c*) Straight type. (*d*) Crossed type, as in Forcipulatida.

many species they are sessile, that is, attached directly to the test. They may be simple and spine-like, recessed like minute clamshells, or felipedal like miniature cat's paws. In the Forcipulatida they are stalked and may be crossed or straight. Their activities are controlled by their own nerves and the basiepithelial plexus of the test (**Fig. 3**).

The well-developed water-vascular system follows the pattern for the phylum, with the following features. With the exception of some Paleozoic genera, the madreporite is always on the upper (aboral) side. There is usually only one madreporite situated interradially, but *Allostichaster* has 2 to 4 and *Acanthaster* up to 16 or more. The tube feet may be peglike and suckerless or columnar and suckered. The former occur in burrowing species and in species which inhabit soft deposits, where they are presumably used like oars or stilts. The latter are used for climbing over obstacles and are better at adhering to hard substrates like rocks. Suckered tube feet can be used to pull open the bivalve prey.

The nervous system also follows the pattern for the phylum. The radial nerves retain a superficial position just below the epithelium lining the ambulacral groove. The basiepithelial nerve plexus envelops most of the body and coordinates the activities of peripheral organs such as pedicellariae, papulae, and paxillae. Starfish cannot see, but they clearly detect changes in light intensity such as those caused by shadows. They have photosensitive eyespots located at the tip of each arm. The terminal tube feet of each ambulacrum appear to be especially sensitive to water-borne chemicals which may be emitted by prey species, and they are thus used in food detection.

Feeding habits. Starfish are voracious feeders. The prey is detected by water-borne odors or by direct contact as the result of random movement. Small food, such as amphipods and young mollusks, may be engulfed whole. *Acanthaster* everts part of its

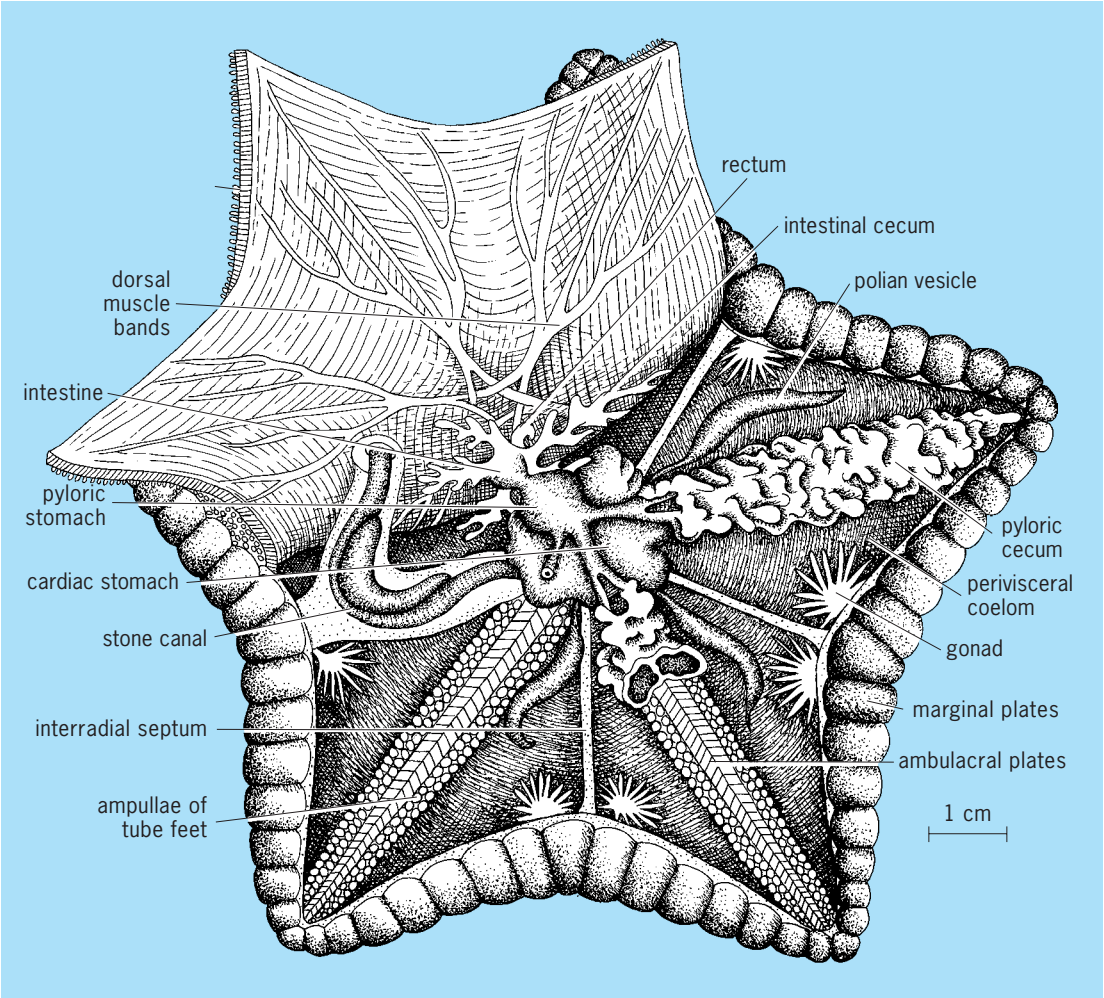

Fig. 4. Anatomy of a starfish, *Asterodon robustus*, as viewed from aboral surface.

stomach out of the mouth and wraps it over coral polyps, which are thus digested outside the predator's body. Some of the Forcipulatida, for example, *Asterias*, can open large clam and other bivalve shells by applying a continuous force which overcomes the molluscan adductor muscles. The starfish uses both tube-feet muscles and body wall muscles. The mechanical stresses applied by a starfish are sufficient to create the shell gap necessary for the insinuation of the folds of the everted starfish stomach. Toxic secretions are not used to paralyze the prey.

The mouth is in the middle of the lower (oral) surface of the disk and leads through a short esophagus to the cardiac stomach (**Fig. 4**). The number of lobes of this organ correspond to the arm number of the starfish. It is these lobes that can be everted through the mouth to carry out extracorporeal digestion. Above the cardiac stomach lies the pyloric stomach, again with as many branches as there are arms. Within each arm, the branch forks to give two blind digestive ceca. In most species, a short intestine leads upward to the rectum and anus, generally situated near the middle of the upper (aboral) surface. In starfish which lack an anus, the feces are extruded through the mouth.

Respiration and water balance are aided by small contractile outgrowths of the test epithelium called papulae. Coelomic fluid circulates on their inside and seawater on their outside, so that exchange can occur across their then membranous surface.

Reproduction. Asteroids are generally mature and able to reproduce at 1 year. They normally continue to grow for about 4 years, growth taking place when food is abundant. The sexes are usually separate, but in a few genera, such as *Fromia*, hermaphrodites occur, the sex changing with age. The sex of adults is not discernible externally. The gonads lie in the interradii, sometimes in pairs, and sometimes extending into the arms. Each has a separate duct to the exterior, and sperm and eggs are shed into the sea, where external fertilization occurs. If a larva develops, it is a bipinnaria or derived type. Many species lack a larval stage. Some brood the eggs, and hatching may take place in the stomach, the mouth, or a special marsupium. Some species reproduce by transverse asexual fission. Regeneration of lost or damaged parts is a characteristic of most genera. In *Linckia* this is most marked, and a whole new individual can be regenerated from a fragment of an arm. *See* ECHINODERMATA.

Andrew C. Campbell

Bibliography. *Asteroidea of the North Pacific and Adjacent Waters*, U.S. Nat. Mus. Bull. 76, pts. 1–3, 1911–1930; M.E. Downey, Starfishes from the Caribbean and the Gulf of Mexico, *Smithson. Contrib. Zool.*, vol. 126, 1973; H. B. Fell, Phylogeny of sea stars, *Phil. Trans. Roy. Soc. London*, ser. B, 246:381–485, 1963; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; C. Schuchert, *Revision of Paleozoic Stelleroidea*, U.S. Nat. Mus. Bull. 88, 1915.

Asteroxylales

A small extinct order of the class Lycopsida that bridges the evolutionary gap between the primitive Zosterophyllopsida and relatively advanced Lycopsida; hence the asteroxylaleans are often termed prelycopsids. The best-known asteroxylalean species are of Early Devonian age, although similar forms survived to the Late Devonian. All possessed at least some hydrophytic features; together with a paucity of strengthening tissues, this characteristic suggests that they largely relied on hydrostatic pressure for structural support. *See* ZOSTEROPHYLLOPSIDA.

With the exception of *Asteroxylon*, asteroxylalean fossils are not well preserved. The asteroxylaleans are regarded as the most primitive lycopsids, but they can also be considered as the most advanced zosterophyllopsids; some authorities prefer to disperse the prelycopsids among other taxa.

The pivotal plant in this order is *A. mackiei*, which originated from the remarkable biotic communities petrified in Lower Devonian volcanigenic cherts at Rhynie, Scotland (see **illus.**). This species had naked horizontal rhizomes that produced primitive roots and aerial branches up to 20 in. (50 cm) high and 0.4 (1 cm) in diameter. The radially symmetrical actinostele (a star-shaped protostele) is typical of primitive lycopsids rather than zosterophyllopsids. Vascular traces extended toward, but did not enter, the densely packed clasping tissue outgrowths termed protomicrophylls.

The sporangia were homosporous, kidney shaped, and distributed singly and randomly among the leaves on the more distal axes. Like the zosterophyllopsids, *Asteroxylon* bore sporangia on stalks with a vascular supply that was independent of, and more extensive than, that of the surrounding enations. These reproductive structures are generally regarded as homologous with vegetative lateral branches.

Other relevant but less well understood Devonian plants are *Drepanophycus* and *Kaulangiophyton*. They were geographically widespread in the Lower and Middle Devonian, when they probably formed extensive clonal mats of low diversity in water-rich habitats. Although typically larger than *Asteroxylon*, they had the same basic growth architecture. However, the unvascularized enations were sparser and smaller, and the scattered large sporangia were egg shaped. The anatomy of these plants is not well known, and the position of the sporangia is ambiguous. For both genera, it is uncertain whether the sporangia occupied separate stalks or were positioned on the upper surfaces of sporophylls, like true lycopsids. *See* LYCOPHYTA; LYCOPSIDA.

Richard M. Bateman; William A. DiMichele

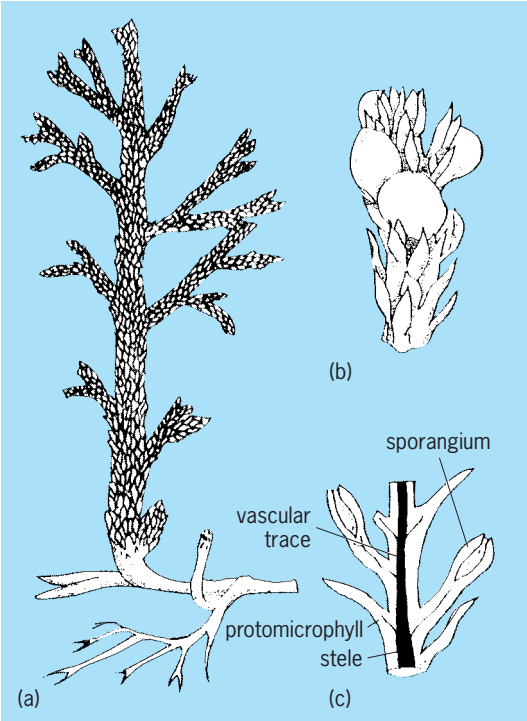

The most completely known asteroxylalean, *Asteroxylon mackiei*, from the Lower Devonian Rhynie Chert of Scotland. (*a*) Reconstruction of vegetative parts of the plant. (*b*) Exterior of the fertile region (*after W. G. Chaloner and P. MacDonald, Plants Invade the Land, HMSO, 1980*). (*c*) Longitudinal section of the fertile region, contrasting the partial vascularization of the protomicrophylls with the complete vascularization of the larger sporangia (*after A. G. Lyon, The probable fertile region of Asteroxylon mackiei, Nature, 203:1082–1083, 1964*).

Bibliography. P. G. Gensel and H. N. Andrews, *Plant Life in the Devonian*, 1984; W. H. Stewart and G. W. Rothwell, *Paleobotany and the Evolution of Plants*, 2d ed., 1993.

Asthenosphere

A layer of variable thickness within the upper mantle of the Earth that is mechanically weaker (less rigid) than both the overlying lithosphere and the underlying mesosphere (**Fig. 1**). The depth of transition from lithosphere to asthenosphere can be as little as 20 km (12 mi) in oceanic settings and as great as 250 km (155 mi) in continental settings. In both cases, the asthenosphere may extend in depth to over 300 km (186 mi). The presence of a mechanically weak layer in the upper mantle can be attributed to the temperature and pressure conditions being such that the ratio between the temperature of mantle material and its melting point (solidus) is a maximum. *See* EARTH INTERIOR; LITHOSPHERE.

The existence of the asthenosphere has primarily been inferred from geophysical measurements,

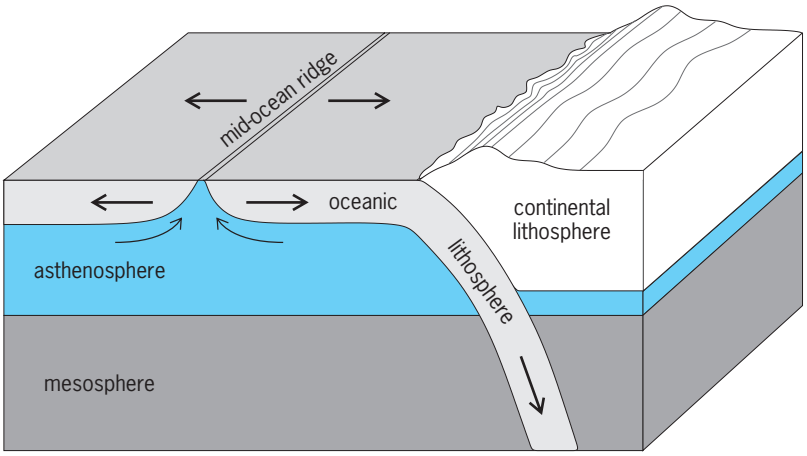

Fig. 1. Location of the asthenosphere in the upper mantle (not to scale). The thickness of the asthenosphere varies laterally due to the dynamics of plate tectonics.

including seismic, magnetotelluric, glacial rebound, and lithospheric stress distribution. The seismic low-velocity zone (LVZ) is a well-known feature of the upper mantle that is often directly correlated with the low-rigidity asthenosphere. Electromagnetic observations indicate the presence of a high-conductivity zone (HCZ) beneath the lithosphere that is consistent with partially molten upper-mantle material. The melting of large ice sheets at the end of the last ice age has significantly changed the surface load distribution throughout much of the Earth; consequent rebound rates can be related to the viscosity of the underlying mantle. Glacial rebound models indicate the presence of a low-viscosity layer beneath the lithosphere that is associated with the asthenosphere. In addition, the forces and motions associated with plate tectonics require the presence of convecting mantle material beneath the lithosphere. The predominant view is that the asthenosphere is unlikely to completely decouple the lithosphere from the mesosphere with its own flow regime; rather, the lithosphere, asthenosphere, and mesosphere are all part of a single convecting system. *See* EARTH, CONVECTION IN; ISOSTASY; PLATE TECTONICS.

Although the asthenosphere has a strict definition as a mechanically weak layer constituting the top part of the sublithospheric mantle, its meaning has broadened in recent times to such an extent that many geophysicists avoid using the term altogether. For example, many now appear to equate the asthenosphere with the LVZ, HCZ, or even the entire upper mantle. In addition, geochemists have given chemical connotations to the asthenosphere, and often describe it as a homogeneous, convecting, and depleted layer beneath the lithosphere. However, from a physical point of view, the asthenosphere could be enriched or depleted, homogeneous or inhomogeneous, and yet still adhere to its original definition.

Mechanical properties. The asthenosphere is thought to be largely composed of the mineral olivine (~60%), with lesser proportions of garnet, spinel, and pyroxenes. In the presence of stresses applied at geological time scales, mantle rocks in the asthenosphere will flow as a result of diffusion creep, a mechanism which involves the thermally activated migration of crystal defects due to stress. The ability of a material to flow can be quantified by its viscosity, which is a measure of a fluid's resistance to flow. In the lithosphere, viscosities range from 10^{21} pascal seconds at its base to infinite at shallower depths as the rheology (that is, the deformation and flow characteristics of the material) enters a brittle-elastic regime. In the asthenosphere, viscosities can be as low as 10^{18}–10^{19} Pa·s, but will increase to 10^{21} Pa · s or more in the underlying mesosphere. It is probable that some parts of the asthenosphere are molten and have extremely low viscosities, but the volume and lateral extent of such material is poorly understood. The viscosity of mantle rock and its mechanical strength depends strongly on the ratio between its temperature and melting temperature (solidus). Above the asthenosphere, this ratio decreases rapidly, resulting in a change from a ductile regime to a brittle-elastic regime. Below the asthenosphere, the mantle solidus increases with depth at a greater rate than the mantle temperature, which increases the mechanical strength of the material. The temperature in the asthenosphere is generally thought to be between 1300 and 1500°C (2372 and 2732°F), but caution is required in attributing a temperature range to the asthenosphere. For example, the mechanical strength of upper-mantle rocks can be greatly reduced by hydration, with no change in temperature. *See* GEODYNAMICS; RHEOLOGY; VISCOSITY.

Seismic properties. The measurement and interpretation of seismic waves within the Earth provides some of the strongest evidence for the existence of the asthenosphere. The speed of seismic shear (S) waves can be significantly reduced in the

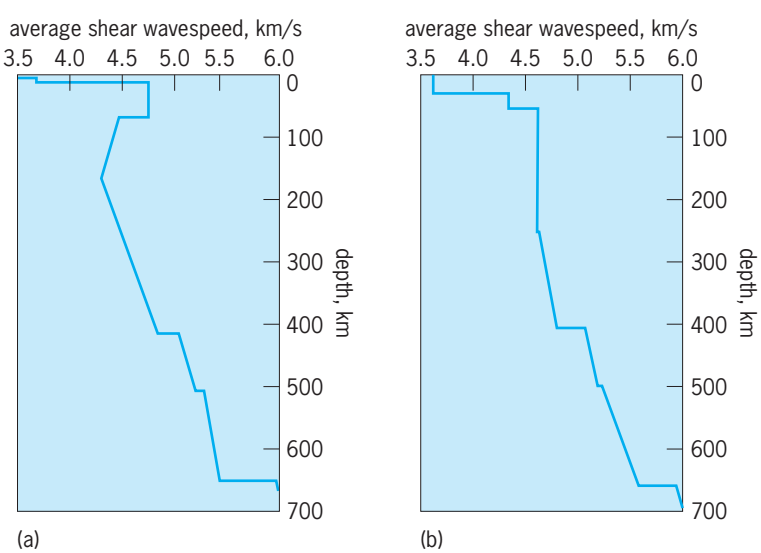

Fig. 2. Variation of seismic shear wave speed with depth in (*a*) suboceanic and (*b*) subcontinental settings. The LVZ is present beneath oceanic lithosphere but absent beneath stable continents.

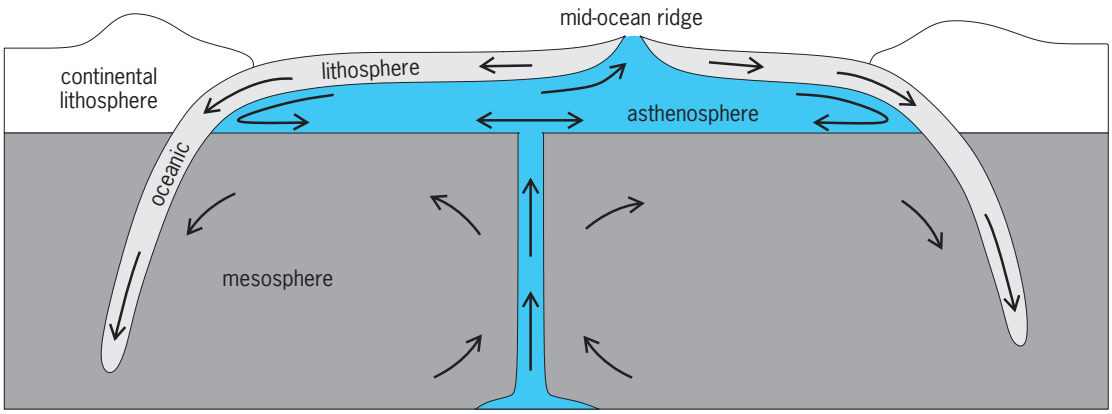

Fig. 3. Proposed mantle convection paradigm (not to scale). In this scenario, the asthenosphere accommodates large lateral flow from hotspots toward ridges and possibly a partial return flow of asthenospheric material from subduction zones. Plate motions and deep mantle flow couple directly at subduction zones and beneath the roots of continental shields.

asthenosphere due to the presence of low-viscosity material. Regional and global studies of the Earth consistently reveal the presence of the LVZ as a pronounced drop in shear-wave speed relative to the lithosphere and mesosphere. However, it appears that the LVZ is largely confined to the suboceanic lithosphere, with little evidence of slower wave speeds beneath inactive continents (**Fig. 2**). Seismic waves (especially shear waves) are also strongly attenuated in regions of relatively low viscosity. Seismic models of mantle attenuation indicate high attenuation beneath continents and low attenuation beneath oceans, a result that correlates well with studies of seismic-wave speeds. *See* SEISMOLOGY.

Another indicator of a weak zone beneath the lithosphere comes from seismic-wave observations showing pronounced changes in the fast orientation of seismic wavespeed (anisotropy) with depth in the upper mantle, which may be related to a change in the preferred alignment of olivine crystals due to variations in the stress conditions (magnitude and direction) over time. In the lithosphere, anisotropy is caused by past deformation that has been "frozen" in, while anisotropy in the asthenosphere is related to contemporary plate motion.

The relationship between the various seismic properties discussed above and mechanical strength as measured by viscosity are not unique. For example, although most of the LVZ is probably due to a change in the physical property of the mantle (that is, reduced viscosity), changes in composition can also affect seismic-wave speeds. Therefore, terms such as LVZ or "high attenuation zone" should not be used synonymously with asthenosphere.

Electromagnetic properties. Electromagnetic studies of the asthenosphere show it to be partially molten beneath mid-ocean ridges with high conductivities (0.1 S/m). Beneath stable continental shields, the conductivity is much less (0.02 S/m), which suggests the absence of an electrical asthenosphere. In regions of active continental tectonics, the depth to the top of the high-conductivity zone (HCZ) can range from 70 to 200 km (43 to 124 mi) in a little over 100 km (62 mi) laterally. Recent studies suggest that the HCZ may also be electrically anisotropic due to the contribution of hydrogen diffusivity to the conductivity of olivine. As in the seismic case, one should not directly associate the HCZ with a mechanically weak asthenosphere, since the HCZ is characterized by the presence of melt and/or volatiles and may be distinct from the LVZ and the rheological asthenosphere. *See* MID-OCEANIC RIDGE; ROCK, ELECTRICAL PROPERTIES OF.

The usefulness of electromagnetic observations in the upper mantle for inferring physical properties has improved considerably in the past decade or so, with major developments in laboratory studies of mantle mineral conductivities. Future work in this area is likely to greatly improve our understanding of the asthenosphere.

Paradigms. Several early models of plate tectonics decouple the motion of the lithosphere from convection in the mesosphere and instead suggest that flow in the mechanically weak asthenosphere could entirely account for the motion of tectonic plates. Subsequent models, which are still popular today, require the lithosphere, asthenosphere, and mesosphere to be part of a unified convective system. An alternative paradigm has been proposed in which the asthenosphere is a layer that exists predominantly beneath the oceanic lithosphere and is fed entirely by upwelling plumes from the lower mantle. The mass balance of the system is preserved via accretion to the overlying lithosphere and subduction (**Fig. 3**). The asthenosphere flow model is one of several paradigms that seek to explain the precise role of the upper mantle in plate tectonics, and helps illustrate the difficulty in trying to define the structure and dynamics of the Earth's interior in the absence of direct measurement. Nick Rawlinson

Bibliography. D. L. Anderson, Lithosphere, asthenosphere and perisphere, *Rev. Geophys.*, 33:125–149, 1995; J. B. Gaherty, M. Kato, and T. H.

Jordan, Seismological structure of the upper mantle: A regional comparison of seismic layering, *Phys. Earth Planet. Int.*, 110:21–41, 1999; G. Heinson, Electromagnetic studies of the lithosphere and asthenosphere, *Surveys Geophys.*, 20:229–255, 1999; I. Jackson (ed.), *The Earth's Mantle*, Cambridge University Press, 1998; J. P. Morgan, W. J. Morgan, and Y.-S. Zhang, Observational hints for a plume-fed, suboceanic asthenosphere and its role in mantle convection, *J. Geophys. Res.*, 100:12,753–12,767, 1995; J. Plomerová, D. Kouba, and V. Babuška, Mapping the lithosphere-asthenosphere boundary through changes in surface-wave anisotropy, *Tectonophysics*, 358:175–185, 2002.

Asthma

An allergic inflammatory disease of the airways, involving mast cells, eosinophils, macrophages, lymphocytes, and neutrophils. Such inflammatory changes are associated with widespread airflow obstruction, which is variable and improves (reverses) spontaneously or with appropriate therapy. Inflammation progresses to increased airway irritability (hyperresponsiveness) induced by the inhalation of allergens, cold air, and occupational factors. Although bronchospasm can be induced immediately after exposure to a specific allergen in an appropriately sensitized recipient, it is the late allergic response that most resembles the inflammatory reaction occurring in asthma. Central to this reaction is the release from mast cells, eosinophils, and lymphocytes of chemical mediators such as histamine, leukotrienes (potent bronchoconstricting agents), and various cytokines which perpetuate the response. Potent neurohumoral agents derived from neural pathways contribute further to the bronchospasm. *See* CYTOKINE; HYPERSENSITIVITY.

Diagnosis. Wheezing, nocturnal breathlessness, coughing, and chest tightness often relieved by expectoration are highly suggestive of asthma. Episodes of breathlessness which result from exposure to an irritant (such as cold air) or an allergen (such as dust mites) following exercise or a viral infection and which are reversed spontaneously or with therapy are diagnostic of asthma. A physical examination may not reveal anything. However, eczema and edema in the folds of the nasal chambers are suggestive of a hereditary allergy, the major predictor of asthma. Nasal polyps may indicate aspirin intolerance in an asthma sufferer. Objective measures of airflow obstruction which improved spontaneously or with therapy are also central to establishing an asthma diagnosis.

Atopy, the genetic predisposition for developing an immunoglobulin-E (IgE) mediated (allergic) response to inhaled environmental allergens, is the strongest predisposing factor for developing asthma. The demonstration of an allergen-specific IgE by the presence of specific IgE antibodies using skin tests or serum analysis allows for easier identification. Utilizing the suspected sensitizing allergen, visual examination by fiber-optic bronchoscopy, or even sputum analysis may provide greater insight into the pathogenesis of asthma. *See* IMMUNOGLOBULIN.

Asthma may be classified, therefore, according to severity, etiology, or pattern of airflow obstruction. It is helpful to differentiate those factors that induce inflammation from those that incite acute bronchospasm in susceptible individuals. The association of an elevated serum IgE and the occurrence of asthma in all age groups, including those who are not atopic, makes antigenic stimulation causal in all instances of asthma. The severity of asthma can best be defined in terms of peak-flow monitoring (monitoring the severity of the allergy). Such evaluations (see **table**) as mild, moderate, and severe are useful in applying therapy in a stepwise manner contingent on severity.

Peak-flow meter assessment provides severity classification and permits early detection of deteriorating condition, often prior to the onset of an asthma attack. Factors associated with a higher risk of developing a severe asthmatic condition and possible fatality include a history of acute life-threatening attacks, hospitalization for asthma within the previous year, a history of respiratory failure, recent reductions in corticosteroid (anti-inflammatory) therapy, and failure to follow a recommended therapy. Asthma sufferers with low incomes who receive inadequate medical care and who are inner-city residents are at especially high risk for fatal asthma.

Therapy. Successful management of asthma requires education of the sick individual coupled with the development of a partnership with an asthma management health-care team; assessing and monitoring the severity of asthma, with utilization of objective parameters of assessment (for example, the peak-flow meter, a device that measures the amount of air that enters and leaves the lungs); environmental management to avoid asthma triggers; and establishment of a drug regimen that controls asthma, as well as a written plan to prevent the condition from becoming worse. Adequate management of asthma should control the symptoms, prevent asthma attacks, return and maintain pulmonary function as close to normal as possible, maintain normal activity levels including exercise, avoid adverse side effects from the drugs, reduce and prevent irreversible airway changes, and prevent mortality.

The treatment of asthma requires that the individual and the family participate with the health-care provider in care. Individuals with moderate-to-severe or difficult-to-manage mild asthma should be trained to measure their morning pulmonary status. Monitoring evaluates effectiveness of the treatment, provides warning of early deterioration in asthma status, and provides critical assessment regarding treatment changes. Measuring the amount of air that enters and leaves the lungs (peak flow rate) for all initial and follow-up examinations provides the most objective assessment, because peak-flow determination is effort-dependent and is not suitable in all instances of asthma. The peak flow rate permits assessment of severity, airway hyperresponsiveness,

Classification of asthma severity

Asthma severity	Clinical features before treatment	Lung function*	Regular medication usually required to maintain control
Mild	Intermittent, brief symptoms less than 1 or 2 times per week Nocturnal asthma symptoms less than twice a month Asymptomatic between exacerbations	PEF >80% predicted at baseline PEF variability <20%	Intermittent inhaled short-acting beta-agonist (taken as needed) only
Moderate	Exacerbations more than 1 or 2 times per week Nocturnal asthma symptoms more than twice a month Symptoms requiring inhaled beta-agonist almost daily	PEF 60–80% predicted at baseline PEF variability 20–30% PEF normal after bronchodilator	Daily inhaled anti-inflammatory agent Possibly a daily long-acting acting bronchodilator, especially for nocturnal symptoms
Severe	Frequent exacerbations Continuous symptoms Frequent nocturnal asthma symptoms Physical activities limited by asthma Hospitalization for asthma in previous year Previous life-threatening exacerbation	PEF <60% predicted at baseline PEF variability 20–30% PEF normal after bronchodilator	Daily inhaled anti-inflammatory agent at high doses Daily long-acting bronchodilator, especially for nocturnal symptoms Frequent use of systemic corticosteroids

*PEF = peak expiratory flow.

response to therapy, symptomatic deterioration, and identification of asthma triggers. Peak flow rate improvement, following bronchodilator therapy, is suggestive of asthma. However, in occasional cases of asthma where there is severe obstruction of the airway, lack of a response to bronchodilators may occur until effective anti-inflammatory therapy has restored the response.

Environmental management with avoidance of indoor and outdoor asthma triggers is central to the successful management of asthma. Tight buildings with inadequate natural ventilation often lead to indoor pollutants as well as dust mites. Such environmental allergens are a major factor in the increase of airway hyperresponsiveness. A reduction in allergens reduces airway hyperresponsiveness and leads to asthma improvement. However, it may take 6 months for complete removal of allergens to be effective. Tobacco smoke and wood smoke always accentuate asthma and must be avoided. Also, drugs such as beta-blocking agents, angiotensin-converting enzyme inhibitors, and aspirin and nonsteroidal anti-inflammatory agents should be avoided in susceptible individuals.

Asthma management requires a step-by-step introduction of medications, dependent upon the standard peak flow rate zone system applicable to all cases of asthma. Medications include bronchodilators (which act as relievers) and anti-inflammatory agents (which act as preventers). Anti-inflammatory agents reduce airway inflammation, whereas bronchodilators relax smooth muscles, thus leading to airway dilatation. Higher drug concentrations can be delivered by aerosol with an inhaler or by dry powder, and there will be little systemic effect. The use of a metered dose inhaler should be repeatedly reviewed. Spacers (a direct drug delivery system) improve delivery of the drug to the airways and reduce deposition of the drug both on the mucous membranes in the mouth and in systemic adsorption.

Bronchodilators are airway smooth-muscle relaxants that enhance microciliary clearance, reduce vascular permeability, and appear to modulate most mediator release. Short-acting beta-agonists are the most predictable and potent bronchodilating agents available and are the drug of choice in reversing acute exacerbation. If beta-agonist therapy is required more than three times per week, therapy with anti-inflammatory agents should begin. If there is no prompt, sustained effect from using beta-agonists, immediate medical attention is required as well as administration of systemic steroids.

Theophylline in sustained-release form has a prolonged bronchodilator effect which is useful in reducing nocturnal asthma. Occasional toxicity requires periodic blood monitoring.

Anti-inflammatory agents are fundamental to the pharmacotherapy of asthma. Corticosteroids are the most potent and effective agents, but at higher concentrations bone metabolism and adrenal function may be inhibited. Such agents inhibit the cytokine effect as well as the synthesis of leukotriene and prostaglandin. Nonsteroid inhaled anti-inflammatory agents inhibit both the immediate and late allergic reactions.

A stepwise approach to pharmacologic therapy is based on the severity of the asthma condition and the current therapy. The number and frequency of drug therapies is increased with increasing severity. Once control is sustained for several weeks or months, a reduction, or step-down in drug therapies can be carefully developed to identify the minimum required for controlling asthma. *See* ALLERGY; RESPIRATORY SYSTEM DISORDERS.

Albert L. Sheffer

Bibliography. J. Brownell and T. B. Casale, Anti-IgE therapy, *Immunol. Allergy Clin. N. Amer.*, 24(4):551–568, 2004; National Asthma Education and Prevention Program Expert Panel Report, NIH Pub. no. 97–4051, 2002; National Asthma Education and Prevention Program Expert Panel Report: Managing asthma during pregnancy: Recommendations for pharmacologic treatment—2004 update, *J. Allergy Clin. Immunol.*, 115(1):34–46, January 2005 [Erratum, *J. Allergy Clin. Immunol.*, 115(3):477, March 2005]; National Heart, Lung, and Blood Institute, *International Consensus Report on Diagnosis and Treatment of Asthma*, 1992; U.S. Department of Health and Human Services, National Heart, Lung, and Blood Institute, *Guidelines for the Diagnosis and Management of Asthma*, 1991.

Astomatida

An order of protozoans, subclass Holotrichia, in which all species are mouthless. All species are parasitic in other animals, typically oligochaete annelids. Many astomatids possess an elaborate holdfast organelle, of value in attachment to the cells or tissues of the host's alimentary tract. Some authorities have considered the species of this rather large and ill-defined group to represent nothing more than degenerate, astomatous members of certain other groups, such as thigmotrichs, apostomes, and possible hymenostomes. Still, there seems to be value in recognizing the evolutionary distinctiveness of the group as a whole. *Anoplophyra* is a typical example. *See* CILIOPHORA; HOLOTRICHIA; OLIGOCHAETA; PROTOZOA.

John O. Corliss

Astrapotheria

A relatively small group of extinct South American ungulates, ranging from the late Paleocene to the late Miocene. They are customarily divided into two suborders: the late Paleocene–Eocene Trigonostylopoidea, and the early Eocene–late Miocene Astrapotheroidea. Eocene members of the order, such as *Albertogaudrya*, tended to be large animals and exhibited the enlargement of the canines characteristic of later astrapotheres.

The order, as recognized today, was probably derived independently from the ancestral condylarths, which were very likely present in South America at the end of the Cretaceous or early Paleocene. One of the most spectacular and advanced members of the order was *Astrapotherium* (see **illus.**). This animal, known from the late Oligocene to the late Miocene, averaged 9–10 ft (2.7–3 m) in length, although some other forms grew even larger. The anterior part of the skull was striking with the huge, persistently growing, curving canines. Although the upper incisors were lost and the premaxillae were drastically reduced, the lower incisors were present along with large lower canines. The cheek teeth were composed of relatively small premolars and of very large molars. The retracted, chopped-off appearance of the snout region strongly suggests that this animal had a moderately large trunk. It is likely that, as in living cows and deer, the well-developed lower incisors cropped against a horny pad above. All this, however, leaves the function of the large canines unexplained. The front legs were somewhat more strongly constructed than the hind ones, and this makes the habits of this animal a puzzle. *See* ARCHAIC UNGULATE; MAMMALIA.

Frederick S. Szalay

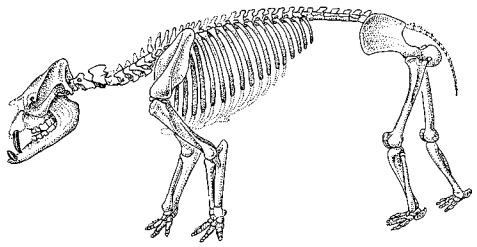

Astrapotherium magnum **skeleton. Genus is known from late Oligocene to late Miocene.**

Astrobiology

The scientific discipline that studies the origin, evolution, distribution, and future of life in the universe. It is an interdisciplinary science integrating contributions from biology, geology, astronomy, paleontology, and planetary science, among others. The first use of the term appears to have been in 1941 when it was defined as the subject of life in the universe other than on Earth. The U.S. National Aeronautics and Space Administration (NASA) adopted the term in 1996 and expanded the meaning to include the origin and history of life on Earth, and the future of life in the universe. In 1998, the NASA Astrobiology Institute was established, solidifying this new discipline within the science community.

Astrobiology is perhaps best understood by starting with one of its most tangible areas of study, the multiplicity of living organisms that surround us here on Earth. It is natural to ask how such diversity evolved, and when during the Earth's early history did life itself begin? To address these questions one must understand the composition of prebiotic chemicals which were present during this period, as well as the physical and environmental conditions of the early Earth in which these chemicals interacted. Investigating these areas will naturally lead one to pose broader questions, "How and when during life's origin and earliest evolution were these critical prebiotic materials formed or delivered to the Earth?" To understand the origin and evolution of life on the early Earth, one must also look to other examples of planet formation both within our solar system and as part of other solar systems throughout our galaxy. What are the factors that make a planet such as the Earth habitable? Are habitable planets common or quite rare in the universe? If life does exist on such habitable planets, how can it be detected? Such is the thread that ties all of the cosmos together, both

inanimate and living, into the new discipline that is astrobiology.

Life on early Earth. The first record of life on Earth has been observed in the form of cyanobacteria-like microfossils believed to have been laid down approximately 3.5 billion years ago during the Archaean (3.8 to 2.5 billion years ago). This stage of life on Earth appears to have already evolved to the level of complex interacting microbial communities, and seems to have done so in a relatively brief period of time. There is compelling evidence indicating that until about 3.8 billion years ago the Earth was subject to an intense meteorite bombardment of debris left over from the formation of the solar system, and the intensity of these impacts would have likely sterilized the planet—destroying any nascent life. The geologic record suggests, therefore, that only about 300 million years was required for the first life to appear and then to evolve to a high level of complexity. This initial life was microbial, which continued as the only form of life for the next billion years of the Earth's history. Much of the current research in astrobiology is involved with understanding the mechanisms of adaptation and evolution which shaped this microbial life on Earth and tracing (mostly) microbial life back to it earliest forms. Comparing the genetic information carried within each organism, in the form of the nucleic acids RNA (ribonucleic acid) and DNA (deoxyribonucleic acid), astrobiologists have been able to resolve the "tree of life" to great detail, and have begun to understand the nature and metabolism of the earliest microbes. This biological information, together with geological evidence of the physical Earth during this early period of its history, sets the conditions for understanding the origin of life on Earth. *See* ARCHEAN; BIOSPHERE; CYANOBACTERIA; HADEAN.

Chemical origin of life. While a rigorous definition of life has proved elusive, there is general agreement among astrobiologists that three conditions must be present for life to exist: liquid water, organic compounds as a source of nutrients, and a biologically useful source of energy. There are a variety of potential sources of energy that astrobiologists are examining, including radiation, photochemical products, minerals, and reduced gases. Life seems quite opportunistic in this regard; it has been found in nearly every possible habitat on Earth, including the polar regions, beneath the subsurface, in desert environments, and even associated with the extreme conditions of hydrothermal vent systems. Liquid water is the one factor that is present in all of these environments, and to the best of our knowledge it appears to be a required factor, one which may limit the distribution of life elsewhere in the universe.

Given that these conditions were met on early Earth, how then did life begin? To begin to address this, researchers have acknowledged that all life on Earth shares a common set of biochemical machinery to orchestrate the processes of life.

These processes involve the use of DNA to store genetic information, a universal genetic code enabling a DNA sequence to specify a protein, protein enzymes to act as chemical catalysts which drive the cells complex chemical processes, and additional protein synthesis machinery based on RNA. Identifying the evolutionary origin of DNA and its associated proteins has proven very difficult, since each requires the other for its own synthesis. Researchers have been faced with the dilemma of which came first? One possible scenario which has been proposed that could bypass this circularity is the concept of an entirely "RNA world" that might have been the basis for the first forms of life on Earth. RNA molecules can replicate and carry information (like DNA), but they can also act as catalysts (ribozymes) much as proteins do. These abilities suggest that RNA could have originally provided both the genetic code and the catalytic function, and these roles could subsequently have been taken over by DNA and proteins. This stage of life's early history, however, continues to be controversial as well as an area of intense study. *See* DEOXYRIBONUCLEIC ACID (DNA); ENZYME; NUCLEIC ACID; PROTEIN; RIBONUCLEIC ACID (RNA); RIBOZYME.

Although the concept of an RNA world has reached some level of acceptance, it is unlikely that RNA itself, with the associated complexity of its four nucleotide bases and a ribose phosphate backbone, represents the earliest prebiotic molecule. Research on nucleic acids with hexoses (six-carbon monosaccharide sugars), for example, suggests that a wide variety of related informational macromolecules are possible. Nonsugar alternatives have also been proposed, such as peptide nucleic acid (PNA). *See* MONOSACCHARIDE; PEPTIDE.

In any case, it is clear that the prebiotic chemistry which led to life at some point must have been segregated from the changing surrounding environment to achieve the stability necessary for continued evolution. Among present-day organisms this segregation is achieved by means of a membrane composed of phospholipids and proteins. A fundamental area of research related to prebiotic chemistry, therefore, concerns the origin of the first cell membrane, which made possible a regulated vesicle for the chemistry and metabolism of life. *See* PREBIOTIC ORGANIC SYNTHESIS.

Source of life's building blocks. To fully understand the biochemical mechanisms leading to the origin of life it is necessary for astrobiologists to search back further in time, and identify what building blocks, in the form of the earliest organic compounds, were present on the early Earth. In characterizing this early material, researchers must address the related question, "How were these critical compounds delivered to our planet?" A significant fraction of this precursor material is believed to have been acquired as a late-accreting layer from impacts of C-type asteroids and comets during the period of heavy bombardment of the inner solar system 4.5 to 3.8 billion years ago. In addition to simple volatile molecules such as water and short-chain hydrocarbons, these C-type asteroids and comets are also rich in complex organic chemicals. *See* ASTEROID; COMET; COSMOCHEMISTRY; METEORITE.

An important part of the puzzle, the idea of a

cometary source for prebiotic organics on Earth, in some respects simply moves the question out into the cosmos, leading astrobiologists to ask, "What is the ultimate source and distribution of organic matter in the universe?" One place they are looking at are the dense interstellar clouds, the birth sites of stars and star systems which include planets. Interstellar ices are composed primarily of water, but they have also been shown to contain some ammonia, carbon monoxide, carbon dioxide, and the simplest alcohol, methanol. It has been known for some time that when similar icy solids are exposed to conditions of ultraviolet irradiation, more complex chemicals can be produced than those originally present in the ice, and astrobiologists speculate that some of these chemicals might have played an important role in the chemistry of the early Earth and in the prebiotic chemistry of planets in general. *See* INTERSTELLAR MATTER; MOLECULAR CLOUD.

Life in extreme environments. Life on Earth has been found in nearly every environment in which water is present, including extremes of temperature, oxygen, pH, salinity, radiation, light intensity, and pressure. The study of life in these extreme environments is one of the most active areas of research in astrobiology, focusing on the environmental limits at which life can be found, the types of microorganisms which can be found under these conditions, how they are able to adapt, and where they occur in the "tree of life." Viable organisms have been discovered frozen within permafrost over geological time scales as well as associated with hot springs and hydrothermal fluids venting from Earth's subsurface crust onto the deep seafloor at temperatures near 120°C (248°F). Other organisms, such as the primitive unicellular red alga *Cyanidium caldarium* or the acidiphilic archaebacterium *Sulpholobus acidocaldarius*, can tolerate acidic conditions down to a pH of 1–2. Halophiles such as *Halobacterium* and *Halococcus*, which can tolerate high salinity, have been recorded at salt concentrations as high as 1.5–4 M (moles per liter of solution). This aspect of astrobiological research has greatly expanded our view of the versatility of life. The results of this work have broadened the scope of possible environments within our solar system which could harbor life as well as formed the basis for searches for particular sites, for example on Mars or Europa, which should be the focus of future space missions. *See* ARCHAEBACTERIA; HALOPHILISM (MICROBIOLOGY); MARINE MICROBIOLOGY.

Biosignatures. The chemistry of life is inherently different from abiotic chemistry, and all organisms leave evidence of this in the environments with which they interact. Physical properties of an environment indicative of previous life are termed biosignatures, and their characterization and use for detecting possible life elsewhere in the universe is a major component of astrobiological research. Biosignatures are generally of two types, geologic or astronomical. Those associated with geological samples include the distribution of organic compounds, the chirality of the molecules (whether they are left or right handed), the isotopic composition of biological elements, biomineralization, and even the presence of microfossils. Astronomical biosignatures are also being investigated now that it is possible to characterize the chemical composition of planets and other extraterrestrial bodies both within and outside our solar system. Such biosignatures are often associated with atmospheric composition but can also reflect the chemical make-up of the surface as well. Spectral biosignatures of the Earth, for example, would indicate the presence of liquid water, chlorophyll on the surface, and abundant oxygen in the atmosphere. More generally, the simultaneous presence of both strongly reduced gases and oxidized gases which are not in chemical equilibrium (such as oxygen and methane in the Earth's atmosphere) is considered a reliable biosignature for many types of planetary atmospheres. *See* ASTRONOMICAL SPECTROSCOPY; MOLECULAR ISOMERISM.

Habitability of other planets. One of the fundamental questions in astrobiology asks, "Where among what is likely to be a vast number of planets and moons are conditions suitable for life as we know it on Earth?" In other words, "What parts of the universe are habitable?" Recalling the three ingredients upon which life depends—liquid water, organic molecules (or carbon), and a suitable energy source—leads to the conclusion that life is a planetary phenomenon. Initially, the habitable zone of planets around stars was thought to encompass planets whose orbits are close enough to their sun for solar energy to drive the chemistry of life, but not so close as to boil off water or break down the organic molecules on which life depends. Within our own solar system this region would encompass perhaps Venus and Mars in addition to the Earth. More recently, however, astrobiologists have come to realize that the habitable zone may be much larger than originally conceived, since other localized sources of heat and energy could be present at great distances from the Sun. The strong gravitational pull caused by large planets, for example, may produce enough energy to sufficiently heat the cores of moons orbiting planets at great distances from their star. This situation is believed to be occurring between Jupiter and its moon Europa, where evidence suggests that a large, salty ocean is present beneath the surface. *See* JUPITER; SATELLITE ASTRONOMY.

The likelihood of planets or other bodies being present in the habitable zone of solar systems other than our own is dependent on, among other factors, the frequency with which planets form around stars in general. After centuries of speculation, scientists have confirmed just since 1995 that there are indeed planets orbiting other stars, and more than 160 extrasolar planets had already been discovered by 2005. While almost all of these extrasolar planets seem to be gas giants, like Jupiter, it is likely that Earth-like worlds also orbit other stars, and instrumentation with sufficient precision to detect a world as small as Earth should be available by 2014. *See* EXTRASOLAR PLANETS.

Edward M. Goolish

Bibliography. L. J. LaFleur, Astrobiology, *Astronomical Society of the Pacific Leaflet Series*, Leaflet no. 143, pp. 333–340, 1941.

Astrometry

That part of astronomy dealing with the position, distance, and motion of celestial objects, including solar system objects, stars, radio sources, and galaxies.

New astrometry. Astrometry has changed since the 1980s. Very long baseline interferometry (VLBI) observations of extragalactic radio sources at milliarcsecond accuracies have provided the basis for a fixed, epoch-independent reference frame. The *Hipparcos* astrometric satellite has provided an optical catalog at milliarcsecond accuracy. Charge-coupled devices (CCDs) have replaced photographic plates and visual observing. The old, single-star position, observing programs have been replaced by programs for observing large numbers of stars and detecting near-Earth objects and extrasolar planets. Large telescopes with CCD mosaics are being used for astrometric surveys, and space missions are planned. The purpose of astrometry is no longer just to determine star positions; rather it is also to obtain positions, motions, and distances in support of astrophysics and the discovery of new objects.

Reference systems. In 1994 the International Astronomical Union (IAU) adopted the International Celestial Reference Frame (ICRF), based on about 400 extragalactic, distant, radio sources, which have no apparent proper motions, as the fundamental reference frame. The individual sources may be subject to changes in the source structure, but the reference frame should be quasi-inertial. This kinematically based system replaced the dynamical system based on the solar system motions. The ICRF replaced the previous fundamental catalogs, such as the FK5, which were based on positions and proper motions of bright stars. *See* CELESTIAL REFERENCE SYSTEM.

The *Hipparcos Star Catalog*, from observations by that astrometric satellite, provides an accurate optical catalog based on the ICRF. Other star catalogs, such as the Tycho and USNO CCD Astrographic Catalog (UCAC), provide a denser coverage of the sky and can reach fainter magnitudes, but with reduced accuracies. The positions and proper motions of the stars provide a two-dimensional map of the sky for a given time. To provide the three-dimensional aspect, the parallaxes (which give the distances to the stars) and the radial velocities are necessary. *See* PARALLAX (ASTRONOMY).

In addition to a reference frame, the International Celestial Reference System (ICRS) includes the definition of time, algorithms (such as the precession-nutation model), constants, and the relationship between the celestial and terrestrial reference frames. *See* ASTRONOMICAL COORDINATE SYSTEMS; ATOMIC TIME; DYNAMICAL TIME; EARTH ROTATION AND ORBITAL MOTION; NUTATION (ASTRONOMY AND MECHANICS); PRECESSION OF EQUINOXES; TIME.

Positional catalogs. Astrometric data are provided in the form of catalogs. One kind is an observational catalog, which can be the publication, or a data set, of individual observations. The observations can also be reduced to a common instrumental system with the individual observations of a given star combined to form a single position for each star. Such observational catalogs that determine their own reference frames are known as absolute catalogs.

Observations can be made at different wavelengths, such as radio, optical, and infrared. The transformations between the ICRF, dynamical reference frames, and optical, infrared, x-ray, and other catalogs must be determined from observational data. *See* ASTRONOMICAL CATALOGS.

Observing programs. Observations can be divided into those involving large-angle and small-angle measurements. Large-angle measurements determine the difference in position between objects over large angular distances in the sky. Historically, such observations were made by transit circles, which record, at the time of transit of an object, both its right ascension and its declination. *See* ASTRONOMICAL TRANSIT INSTRUMENT.

Interferometers, observing at radio, optical, or infrared wavelengths, combine the reception of the emission from a source at two separate detectors. By measuring the time difference between the two detections, a very accurate measurement of the angle to the source can be determined. Thus, at radio wavelengths, VLBI uses antennas located as far apart as possible to maximize the baseline between the two instruments and to develop the most accurate observations. The data are recorded on tapes and shipped to a correlator for analysis. At optical wavelengths, the signals must be combined immediately, and the differences in the atmospheric effects must be minimized. However, since optical wavelengths are so much shorter than radio wavelengths, milliarcsecond accuracies can be achieved over baselines of tens of meters, while for radio-frequency measures, distances of thousands of kilometers are required. *See* RADIO ASTRONOMY.

The *Hipparcos Astrometric Satellite* used a technique for observing pairs of stars separated by approximately 60° to form a catalog of stars located throughout the sky. The *Hipparcos Catalog* provides the positions and proper motions for stars, measured over a 3-year period in 1990–1993, at accuracies of approximately 1 milliarcsecond at the observational epoch and degrading in accuracy at 1 milliarcsecond per year, because of the accuracy limits of the proper motions.

Small-angle measurements are made by individual-exposure techniques, for example, measuring the positions of all the objects on a CCD exposure. They provide accurate relative positions of the observed objects. They can also provide, by means of multiple observations, the parallaxes and motions of the stars with respect to reference stars in the field. CCDs, which can be mounted in mosaic cameras, offer the advantages of a very linear and efficient response to detected photons, and an instantaneous digital readout of the field. *See* ASTRONOMICAL IMAGING; CHARGE-COUPLED DEVICES.

Overlapping exposures and global solution techniques can be used to develop the small-angle measurements into large-angle solutions. Also,

combining many exposures can provide surveys of the skies. *See* ASTRONOMICAL ATLASES.

For accurate positional work, multiple stars must be identified, since it is necessary to take into account the motion of the center of light of a multiple star as the components move with respect to each other, or to observe the individual moving components. Speckle interferometers take very rapid exposures (approximately 30 per second) to freeze atmospheric effects. These short exposures can then be added together to measure the separation, relative position, and magnitude of pairs of stars that could not otherwise be observed with such accuracy through the atmosphere. The atmosphere is the primary limitation on ground-based astrometric accuracy. *See* BINARY STAR; SPECKLE; TWINKLING STARS.

P. K. Seidelmann

Bibliography. E. Hoeg and P. K. Seidelmann (eds.), *Astronomical and Astrophysical Objectives of Submilliarcsecond Astrometry*, IAU Symp. 166, 1995; G. H. Kaplan, *The IAU Resolutions on Astronomical Reference Systems, Time Scales, and Earth Rotation Models*, USNO Circ. no. 179, 2005; J. Kovalevsky and P. K. Seidelmann, *Fundamentals of Astrometry*, 2004; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, 1992; P. K. Seidelmann and A. K. B. Monet (eds.), *Astrometry in the Age of the Next Generation of Large Telescopes*, 2005.

Astronautical engineering

The engineering aspects of flight and navigation in space. The latter is also known as astronautics. The terminology thus parallels aeronautical engineering and aeronautics as applied to aviation. Astronautical engineering deals with vehicles, instruments, and other equipment used in space, but not with the sociological or economic aspects of space flight, except as they influence the equipment.

Space. The boundary between the atmosphere and space is a matter of debate. Few contend that, for practical purposes, it lies much more than 100 mi (160 km) above the surface of the Earth; some would put the boundary lower. For purposes of record, the Fédération Aéronautique Internationale has established the line of demarcation between the atmosphere and outer space at 62 mi (100 km) above the surface of the Earth. Over 99% of the Earth's atmosphere lies within 20 mi (32 km) of its surface, but the air thins slowly above that height, merging imperceptibly with outer space, which even in its farthest reaches may contain a molecule or two per cubic inch. With the boundary of space so indeterminate, a more precise definition of a space mission is on the basis of velocity, the space mission being at or above the velocity required to circle the Earth completely as a satellite.

Because there is no sharp line of demarcation between the atmosphere and space, some overlap exists between aeronautical and astronautical engineering. Nevertheless, astronautical engineering is unique in many aspects; there are far more differences than similarities in the two disciplines. *See* AERONAUTICAL ENGINEERING.

Aircraft and spacecraft contrasts. There is a lack of parallelism between current astronautic and aeronautic vehicle terminology. An aircraft is a self-contained vehicle, having within its structure essentially all the equipment required to transport its payload from one place to another. A spacecraft, in the more restricted sense, is the container for the payload. Sometimes the word is used to denote the container and payload. Most spacecraft, to date, have had either very limited propulsion or none at all.

Since enormous speeds are the hallmark of all astronautic missions, unpowered spacecraft require a booster, or launch vehicle, usually a rocket many times as large as the spacecraft. The weight of the spacecraft, in fact, seldom exceeds 5% of the total launch vehicle weight.

The distinction between spacecraft and launch vehicle has been further clouded by the NASA space shuttle, which uses as its orbital stage a winged airplane. This so-called orbiter is capable of launching spacecraft in the more traditional sense, but also of remaining in orbit for several days and performing useful missions during its orbital stay. The shuttle also carries a crew. Its wings and other airplanelike features are used only in the descent and landing. *See* SPACE SHUTTLE.

Aircraft structures are designed to house the payload and to obtain lift and control by interaction with the atmosphere. Spacecraft structures must properly house the payload in both the atmospheric and space environments. Space vehicles are usually expended on one flight; aircraft are designed for many years of service. Spacecraft are lifted by the thrust of their rocket engines, steered by directing the thrust, and usually employ no wings or fins. Where aerodynamic effects are encountered, such as in leaving and reentering the atmosphere, the speeds involved usually differ from those of airplanes by an order of magnitude. *See* SPACECRAFT STRUCTURE.

Airplanes are most frequently propelled by air-breathing engines, which use the oxygen of the air for combustion. They produce their thrust by changing the momentum of the air. Spacecraft use rocket engines exclusively, although some proposals have been made to use air-breathing engines for the initial phases of a spaceship's ascent. Rockets contain within themselves both the energy (that is, the fuel and oxidizer) and the mass whose momentum is changed for propulsion. *See* ROCKET PROPULSION.

Airplanes are usually guided by a human pilot, spacecraft by automatic or remote controls. The precision required in space navigation makes it probable that spacecraft will always depend heavily on automatic control, although people may serve to monitor and adjust the equipment. *See* SPACE NAVIGATION AND GUIDANCE.

Electric power for auxiliary purposes is easily generated in aircraft. The propulsion engines for airplanes must operate continuously during flight to

overcome the ever present aerodynamic drag. Electric generators, driven by the main engines, supply the required power. In contrast, space vehicles coast freely during practically all of their useful lives. Solar cells and nuclear reactors are two obvious means of supplying electric power. Fuel cells using hydrogen and oxygen have also been used for shorter-duration applications. *See* SPACECRAFT PROPULSION.

Cost and reliability. It is extremely expensive to put a pound of payload into Earth orbit. Thus designers have been justified in going to great lengths to convert a pound of structure into a pound of payload. Great improvement appears possible in this respect; only the cost of the propellant seems to be irreducible.

In view of the high cost of space operations, it is especially important that space vehicles operate long enough to successfully fulfill their missions. A severe reliability requirement is thus imposed upon vehicles and equipment intended for missions, such as journeys to the planets, which may require up to a year or more to accomplish.

The requirements for high reliability and low weight add tremendously to the cost of the payloads themselves, to the extent that their cost approaches that of the launch vehicle. In the Apollo lunar missions the cost of the last three stages, which included the lunar module and the service module, came very close to that of the first three stages comprising the Saturn V launch vehicle, even though the weight of the payload was only about 1.5% that of the launch vehicle. In their eagerness to minimize the cost of the launch vehicle, the designers may have overreached themselves. Obviously, when both expendable launch vehicles and expendable spacecraft are used, the objective should be minimum combined cost. This minimum does not necessarily occur when the costs of launch vehicles and spacecraft are equal.

All space missions through 1975 used expendable launch vehicles. The reusable space shuttle was expected to reduce the costs of Earth-to-orbit transportation. However, preliminary indications are that the shuttle will not prove nearly so economical as originally expected, largely because it is only partially reusable, and because the recovery provisions are heavy and complex.

Designing for minimum cost is of extreme importance in astronautical engineering because of the very high unit costs involved and because frequently only one approach is selected. The wrong approach could therefore waste billions of dollars and still not be recognized as an inferior one.

Because astronautical engineering is a transportation science, it is convenient to discuss it in terms of vehicle, environment, and interaction between the two.

Escape velocity. Gravity is a dominating influence in the design of space launch vehicles. To achieve even a low orbit, a vehicle must expend an amount of effort equivalent to climbing out of a well some 4000 mi (6400 km) deep. The task is roughly equivalent to driving a car from San Diego, California, to Bangor, Maine, while dragging a cement block as heavy as another automobile, with no gas stations en route. As a result, a space vehicle as it sits on a launching pad consists mostly of propellant.

The pull of the Earth's gravity on a body extends indefinitely far into space and varies inversely with the square of the distance to the center of the body. More generally, gravitational attraction exists between any two bodies. Isaac Newton expressed the quantitative relation as $F = g(m_1m_2/r^2)$, where g is a universal constant, m_1 and m_2 are the masses of the bodies in question, and r is the distance between their centers of mass.

Despite the fact that the pull of gravity extends to infinity, it is nonetheless possible to escape permanently from the Earth's gravity in the sense of never being drawn back to the ground. The key is speed. Circular velocity is the minimum at which a space vehicle can remain permanently above the Earth. At low altitudes, this velocity is about 25,000 ft/s (7.9 km/s). *See* SATELLITE (SPACECRAFT).

As the speed is increased above the circular velocity, the path of a vehicle becomes a larger circle or an elongated ellipse. When the speed reaches 37,000 ft/s, or about 7 mi/s (11.2 km/s), the path becomes a parabola and the vehicle will travel along one of the legs to infinity without further propulsion. *See* ESCAPE VELOCITY.

Propulsion requirements. These velocities are tremendous by any previous standard. To reach them, a vehicle must carry the corresponding amount of energy in the form of propellant.

Even with the most energetic propellants and the lightest structures, it has not yet been possible to reach orbital velocity with a single rocket. To overcome this seemingly insurmountable obstacle, one rocket is carried as the payload of a larger one. When the larger burns out, the second is ignited and adds its velocity to that of the first. This is known as the step-rocket or staging technique. *See* ROCKET STAGING.

For lunar and planetary missions, lightweight vehicles, powerful propellants, and many stages are used. The lunar orbit rendezvous method required a total of six stages to take the Apollo astronauts to the Moon and back. The launch involves the full thrust of the first and second stages and the partial burning of the third stage of a Saturn V. The third stage is reignited to achieve escape velocity. The fourth, the service module (SM), uses some of its propellant to enter an orbit around the Moon. The fifth stage, the lunar module (LM), descends to the surface of the Moon, and the sixth returns to lunar orbit to rendezvous with the service module, at which time the astronauts and payload of the lunar ascent module are transferred to the command module. The service module then propels the command module out of lunar orbit and on a highly elliptical return to Earth. The command module accomplishes braking in the Earth's atmosphere and finally lands by parachute in the ocean.

Three different propellant combinations are used for primary propulsion. The first stage, which comprises 78% of the liftoff weight, uses dense, cheap

propellants—liquid oxygen and kerosine. The second and third stages use liquid oxygen and liquid hydrogen. Hydrogen gives high performance but has a very low density and boiling point. The service module and the lunar module, which must coast in space for several days, use noncryogenic nitrogen tetroxide and mixed hydrazines.

The main engines of the space shuttle have pushed chemical propulsion technology almost to its limit. Hydrogen and oxygen are the propellants, and the combustion pressures have been raised to some 3000 lb/in.^{2} (20.7 megapascals). Very little, if any, further reduction in fuel consumption can be had through additional increases in pressure. Many people, in fact, question whether the small gains achieved in going higher than the pressures used in the Saturn V upper stages have justified the increased expense, reduced reliability, and increased maintenance.

Future vehicles may use nuclear power. Fission reactors were used to heat hydrogen gas in the Nuclear Engine Rocket Vehicle Application (NERVA) Program, and propellant consumption amounting to a third less than that achieved by the best chemical combination was demonstrated. Costs, political pressure, and the presence of intense harmful radiation have thus far prevented any practical application of nuclear power for rocket propulsion. Controlled nuclear fusion represents a longer-term, potentially very attractive possibility.

Guidance and control. Although propulsion is the key to space flight, other elements are essential and present numerous new problems. One such element is guidance and control. For the ascent phase of space vehicle flight, guidance systems similar to those used for ballistic missiles are employed.

Another control requirement of many types of space vehicles is that of maintaining the desired vehicle attitude over long periods of time. Displacement gyroscopes, even excellent ones with very low drift rates, cannot provide an accurate reference for days or weeks. Such devices must be corrected frequently by an external reference.

At least two such references are available: sources of electromagnetic radiation, and the gravitational gradient. The first might be used by such devices as a Sun seeker, a star tracker, or a horizon scanner. All of these detect sources of emitted or reflected radiation. Of special interest is the horizon scanner, which indicates the direction of the vertical with respect to the Earth. This direction is vital for orbital missions that relate to occurrences on the ground such as meteorological observation.

In the vicinity of the Earth (or any large celestial body) the difference in the pull of gravity between points on the craft having different distances from the Earth can be usefully employed.

Reaction wheels or other devices capable of storing angular momentum may be used to provide the torque to effect or maintain a given orientation. Such devices are very efficient, both from a weight and an energy standpoint, where disturbing torques on the spacecraft are small, random, and long continued. At the opposite end of the torque spectrum, torques that are large and uncompensating, rocket engines are the most suitable. For many applications reaction wheels are useful to absorb the small random torques, and rockets handle torques and total momenta which exceed the capacity of the wheel system.

Auxiliary power. Vehicle and payload equipment require electric power. This power, generally speaking, must be provided in rather generous quantities, and, in many cases, for long periods of time. The energy required can be carried along, or it can be supplied by the Sun. For small amounts of energy, chemical sources, such as batteries or chemically fueled generators, may be used. A great deal more energy can be obtained from a nuclear reactor. Energy also comes continuously from the Sun but at a fairly low density at Earth's distance.

Nuclear reactors no larger than a basketball appear possible. Large power outputs are possible, but at the expense of considerable weight. An additional, stringent limitation to the use of nuclear reactors is the radiation that they give off. For human crew members or sensitive payloads, heavy shielding is required.

Where small amounts of power are needed, but for long periods of time, solar energy presents an ideal source. Solar batteries are particularly attractive, because they have no moving parts and thus offer good dependability. Present solar cells operate at 8–10% efficiency with some improvement in prospect. In space near the Earth, each square foot of surface normal to the Sun's rays receives about 150 W of solar energy. Of this, perhaps 10 W can be delivered to the payload. Various schemes for obtaining the required intercept area have been proposed. If power requirements are low, the solar cells are mounted on the spacecraft. If the entire external area of the craft does not provide sufficient area, special panels are provided that unfold once a space environment is reached. *See* SOLAR CELL; SPACE POWER SYSTEMS.

Communications equipment. The communications equipment comprises an essential item of nearly all space vehicles. This equipment is designed for light weight, low power consumption, and, usually, long life. It may be designed especially to operate in a strong radiation environment. The equipment transmits data generated within the space vehicle, receives commands from the Earth, or emits signals to permit the vehicle to be tracked by a ground station. *See* SPACE COMMUNICATIONS.

Payload. Nearly all space vehicles have the items previously mentioned: propulsion, guidance, power supply, communications. In addition, there will be specialized payload items, depending on the function the vehicle is to perform. These may include scientific instruments, telescopes, communications relay and storage equipment, or human beings. There must also be a space frame to house all of these and to provide the required environment in the vacuum of space.

Ground communication and support. Although a large percentage of the problems and most of the

romance of space flight are associated with the vehicles, it would be a mistake to assume that these constitute even a major fraction of the total operating system. Indeed, the cost of overcoming the Earth's gravity is so great that any portion of the total operation which can be performed on the ground should be done there.

The supporting ground equipment consists of the preparation and launching equipment, and the tracking, communications, and payload-oriented equipment for turning the received data into usable form. For missions which involve return of space vehicles or booster rockets, recovery equipment may also be required.

Launch and checkout equipment closely resembles (and sometimes is identical with) that used for large ballistic missiles. Ground-to-space communications equipment is generally distinguished by the use of large, high-gain antennas, coupled with extremely sensitive receivers. These receive the low-power signals from the spacecraft. Precise radio and optical tracking systems, together with electronic computers, permit rapid and accurate determination and prediction of the position of space ships. *See* LAUNCH COMPLEX; SPACECRAFT GROUND INSTRUMENTATION.

Reentry. In their interaction with the terrestrial and atmospheric environment during reentry, space vehicles resemble ballistic missiles. However, although ballistic reentry techniques have been proved successful, the use of winged vehicles also has certain attractive aspects. There is a basic difference in these two methods in respect to the way atmospheric heat is handled. The ballistic approach absorbs the heat in the reentry body or rejects it back to the air by mass transfer. The winged vehicle dissipates the heat by radiation. Considerable research has been done on compromise reentry vehicles, such as the lifting body approach. *See* ATMOSPHERIC ENTRY.

The orbital stage of the shuttle is a winged craft designed to land like an airplane. It will utilize a combination of techniques to overcome the reentry heating problems: lift, temperature-resistant materials, and local ablative cooling.

In addition to the problem of reentry into the Earth's atmosphere, some missions require entry into the atmospheres of other planets, notably Mars, Venus, and Jupiter. These atmospheres are strikingly different from that of the Earth in composition, density, and density gradient. The same general techniques will probably apply, but optimal designs will differ quite sharply from Earth atmosphere reentry systems. *See* HYPERSONIC FLIGHT.

Environment. Astronautical engineering must contend with the unique environment of space outside the Earth's atmosphere. Here it is necessary for the engineer not only to learn and apply new information, but to dismiss many old concepts and rule-of-thumb procedures.

Although gravity is present in space, whenever a vehicle is coasting unpropelled, the shell and everything in it are acted on equally by gravity and therefore appear weightless. Loose internal objects do not "fall" relative to fixed items. Fluids do not flow naturally, but must be confined and extruded. Liquids exposed to the vacuum of space evaporate or freeze. External transfer of heat takes place only by radiation.

Metals exposed to the ultraviolet rays of the Sun emit electrons. Small particles of cosmic dust strike external surfaces at extremely high velocities and gradually erode them. Cosmic radiation creates a spectrum of secondary radiation that may reach levels damaging to equipment or personnel.

Starting rotating equipment within the vehicle may set the latter into counterrotation so the momentum may be conserved. In dealing with this environment the modern engineer must think twice to avoid being misled by earthbound experience. *See* SPACE FLIGHT. Robert C. Truax

Bibliography. V. S. Avduevskii and G. Uspenskii, *Scientific and Economy-Oriented Space Systems*, 1988; L. Friedman, *Star-Sailing: Solar Sails and Interstellar Travel*, 1988; N. Johnson and D. McKnight, *Artificial Space Debris,* 1987; V. L. Pisacane and R.C. Moore (eds.), *Space Systems*, 1994; O. Zarrouati, *Space Trajectories*, 1987.

Astronautics

The application of scientific principles and engineering techniques to flight in space. Astronautics deals with space vehicles in the sense that aeronautics deals with aircraft. A distinguishing feature between astronautics and aeronautics is the extent to which the vehicles are influenced by the Earth's atmosphere. *See* AERONAUTICS; ASTRONAUTICAL ENGINEERING.

The subject matter of astronautics is flight at altitudes where a vehicle overcomes gravitational attraction and controls its course by reactive propulsion. Aeronautics concerns flight in regions where a vehicle resists gravitational attraction and controls its course by aeromechanical forces. The distinction is convenient but not clear cut. Rockets, by their reaction, assist airplanes to take off. Space vehicles, such as the shuttle, may glide back to Earth. *See* INTERPLANETARY PROPULSION; ROCKET PROPULSION.

During the twentieth century, aeronautics progressed from a human ambition to a commercial and military reality. Astronautics looks forward to similar progress during the twenty-first century. *See* INERTIAL GUIDANCE SYSTEM; NAVIGATION; RELIABILITY, AVAILABILITY, AND MAINTAINABILITY; SPACE; SPACE BIOLOGY.

Thousands of successful space missions have been flown. Crewed Earth orbital missions included two practice missions under Project Apollo and three Skylab missions. The former Soviet Union completed many crewed Earth orbital missions with its Soyuz and Salyut spacecraft. *See* SPACE FLIGHT; SPACE STATION.

A total of nine crewed lunar voyages have been made, seven of them involving landing. Two included considerable surface travel with a small,

jeep-sized vehicle, the lunar rover. All were accomplished without fatalities. One launch pad accident resulted in the death of three Apollo astronauts, V. I. Grissom, E. H. White, and R. R. Chaffee. Four Soviet cosmonauts died during reentry from Earth orbital missions; there was one fatality in 1967 and three in a 1971 mission.

The Apollo lunar landing missions involved complex feats of navigation and control, as well as many stages of propulsion. The basic flight plan included Earth orbit, transfer from Earth to lunar orbit, descent to the surface of the Moon, return to lunar orbit, and return to Earth. The lunar descent was made by the two-stage lunar module (LM). The service module (SM), crewed by one of the three astronauts, was left in orbit. After the surface exploration, the ascent stage of the lunar module rejoined the service module in lunar orbit. The service module provided propulsion to return the command module to the Earth. The command module alone made the final reentry with the three astronauts.

Equally sophisticated feats of astronautics were performed by some of the crewless missions. There were literally hundreds of relatively simple Earth orbital missions by both the United States and the Soviet Union. They included military reconnaissance satellites, communications satellites, and satellites for weather observation, recording radiation intensities, and biological investigations. *See* COMMUNICATIONS SATELLITE; MILITARY SATELLITES; SATELLITE NAVIGATION SYSTEMS; SCIENTIFIC AND APPLICATIONS SATELLITES.

The United States alone has launched successful crewless probes to every planet but Pluto. It is now possible to manipulate many of these probes by radio command from Earth—to correct their courses, start, stop, and aim their cameras, and control a great variety of other on-board equipment. *See* SPACE NAVIGATION AND GUIDANCE; SPACE PROBE.

Two notable accomplishments were the soft landing of crewless spacecraft on Mars and swingbys of Jupiter and Saturn by the Voyager probes. In both cases, spectacular photographs were relayed back to Earth, and a great deal of other data were obtained. The photographs of the moons of Jupiter were particularly breathtaking.

After termination of the Apollo program the only crewed spaceflights were orbital missions carried out by the Soviet Union. These missions set new records for human endurance in space and established the feasibility of voyages lasting at least a year.

The crewed space program of the United States resumed with completion of the space shuttle, a partially reusable craft having a winged orbital stage. A four-shuttle fleet made numerous flights with great technological success. In January 1986, on the twenty-fifth shuttle mission, however, the shuttle *Challenger* exploded shortly after liftoff, killing all seven astronauts on board. The accident was attributed to a leak in a joint in one of the solid-propellant booster rockets. The accident set the United States space program back at least 2 years. Further flights were suspended until the causes could be ascertained and appropriate remedies applied. Shuttle flights resumed in September 1988. *See* SPACE SHUTTLE.

In July 1988, Russia launched a crewless probe to the planet Mars. That country has been accumulating thousands of crew hours in space, presumably in preparation for a crewed mission to Mars. The United States has preferred to concentrate on construction of a rather large crewed space station. China has become the third spacefaring nation with a brief orbital crewed mission in 2003 and a successful five-day mission in 2005.

There has been an increased interest in commercial activities in space worldwide. The European space community has a strictly commercial launch activity in French Guiana, where Ariane vehicles are regularly launched with commercial payloads. *See* SPACE TECHNOLOGY. Robert C. Truax

Bibliography. H. S. F. Cooper, *Before Lift-off*, 1987; J. D. Rummel, V. M. Ivanov, and J. Rummel (eds.), *Space and Its Exploration*, 1993; H. L. Shipman, *Humans in Space: Twenty-First Century Frontiers*, 1989.

Astronomical atlases

Books, CD-ROMs, or computer-accessible sets showing stars, constellations, or other astronomical phenomena and their locations in the sky. Astronomical atlases provide two-dimensional views of the distributions of objects.

The first sets of astronomical sky positions were mere lists of stars visible to the unaided eye and were not drawn to make an atlas. Hipparchus's star catalog (circa 127 B.C.) divided the stars into brightness classes, from which the current magnitude scale descends. Hipparchus's work is known mostly from Ptolemy's *Almagest*, written about A.D. 150 but available from translations about 1000 years after that. Ptolemy's catalog included 1028 stars. It is said that Ptolemy plotted his stars onto a globe; no trace of that globe is known, though other star globes may descend from it. *See* MAGNITUDE (ASTRONOMY).

Early star maps. Comets were pictured in the *Nuremberg Chronicle* (1493), but the same woodcuts were used multiple times, often in different orientations. So a question arises as to when representations of the sky become scientifically accurate.

Albrecht Dürer, in Nuremberg, made woodcuts dated 1515 that show the Ptolemaic stars and constellations, one map for each hemisphere. The first real star atlas was *De le stelle fisse* by Alessandro Piccolomini (1540), published in Venice. Piccolomini showed the different constellations, though crudely. He showed all but one of Ptolemy's constellations, and the Southern Cross as well. No constellation figures accompany the star charts. He showed four levels of brightness and marked several stars in each constellation with letters. His book was written in

Italian, a spoken language, rather than scientific and scholarly Latin. *See* CONSTELLATION.

Copernicus's historic Sun-centered diagram of 1543 was schematic, and diagramed only a circle, the outermost, for the "sphere of fixed and immobile stars." When Thomas Digges, in 1576, added an appendix to the book *Prognostication Everlasting* (1556) of his father, Leonard Digges, he provided the first discussion and first diagram of the Copernican system in English. His diagram shows stars extending outward at varying distances, something he discusses in the text. But the stars are not accurately located.

The star positions could be plotted with increasing accuracy as measuring accuracy improved. Tycho Brahe (1572) discussed his supernova, and Johannes Kepler (1604) discussed his own supernova with star diagrams included. *See* SUPERNOVA.

Golden age of star atlases. The first of the major beautiful star atlases was the *Uranometria* of Johannes Bayer, published in Augsburg in 1603. Bayer's 51 charts (**Fig. 1**), engraved by Alexander Mair, show the individual constellations, and additional northern and southern hemispheres show the overall arrangement of constellations. Bayer's atlas has grids to mark the star positions.

Bayer's lettering scheme is still used today for the stars that were so lettered. Bayer usually assigned the Greek letters in approximate order of brightness in a given constellation, with similar brightnesses assigned in order of position, and with special objects like the Big Dipper assigned Greek letters in order around the asterism. Bayer also used some lowercase Roman letters when he ran out of Greek letters for a given constellation. He used Tycho's observations for northern hemisphere objects and observations of Pieter Dirckszoon Keyzer, a Dutch navigator, for southern hemisphere objects.

Johannes Hevelius used his own observations to make an even more handsome catalog, *Firmamentum Sobiescianum sive Uranographia Joh. Hevelii*, published in Gdansk in 1687 (**Fig. 2**). He used Edmond Halley's observations of southern stars. Hevelius drew the constellations as they would be seen from outside the celestial sphere looking in (God's view), so they appear backward from the way they were in Bayer's charts or, indeed, to human eyes. He introduced 11 new constellations, 7 of which survive. Hevelius's constellation figures are sharply and incisively drawn, basically one to a page.

John Flamsteed, the first Astronomer Royal, made his own set of detailed observations of the northern stars and used them to create his *Atlas Coelestis*, published under the supervision of his friends in London in 1729 after his death. Most pages show a few constellations together. The French globe maker Fortin revised the book in 1776, dividing Hydra into two parts, and the now 27-plate atlas appeared in several French editions.

Johann Doppelmayr, a Nuremberg professor, published his *Atlas Coelestis* in 1742. It contains a variety of illustrations of astronomical events and

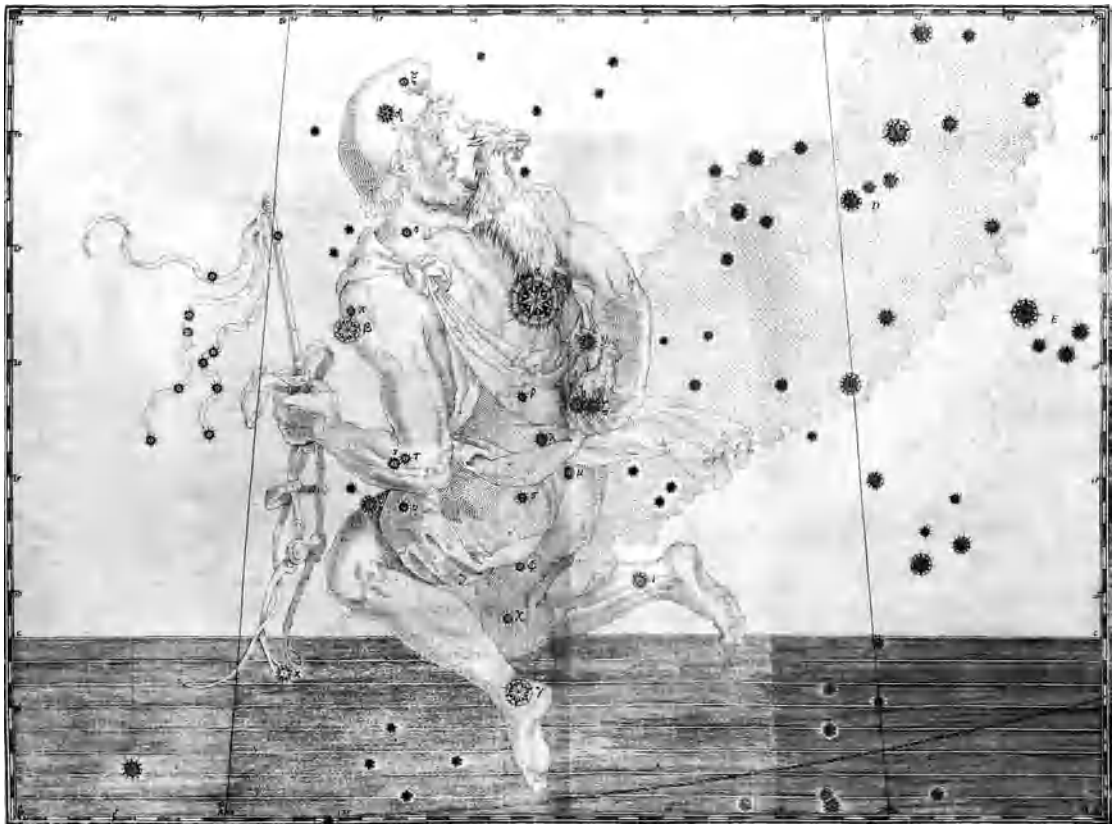

Fig. 1. Auriga, from Johann Bayer's star atlas, *Uranometria* (1603). (*Jay M. Pasachoff collection*)

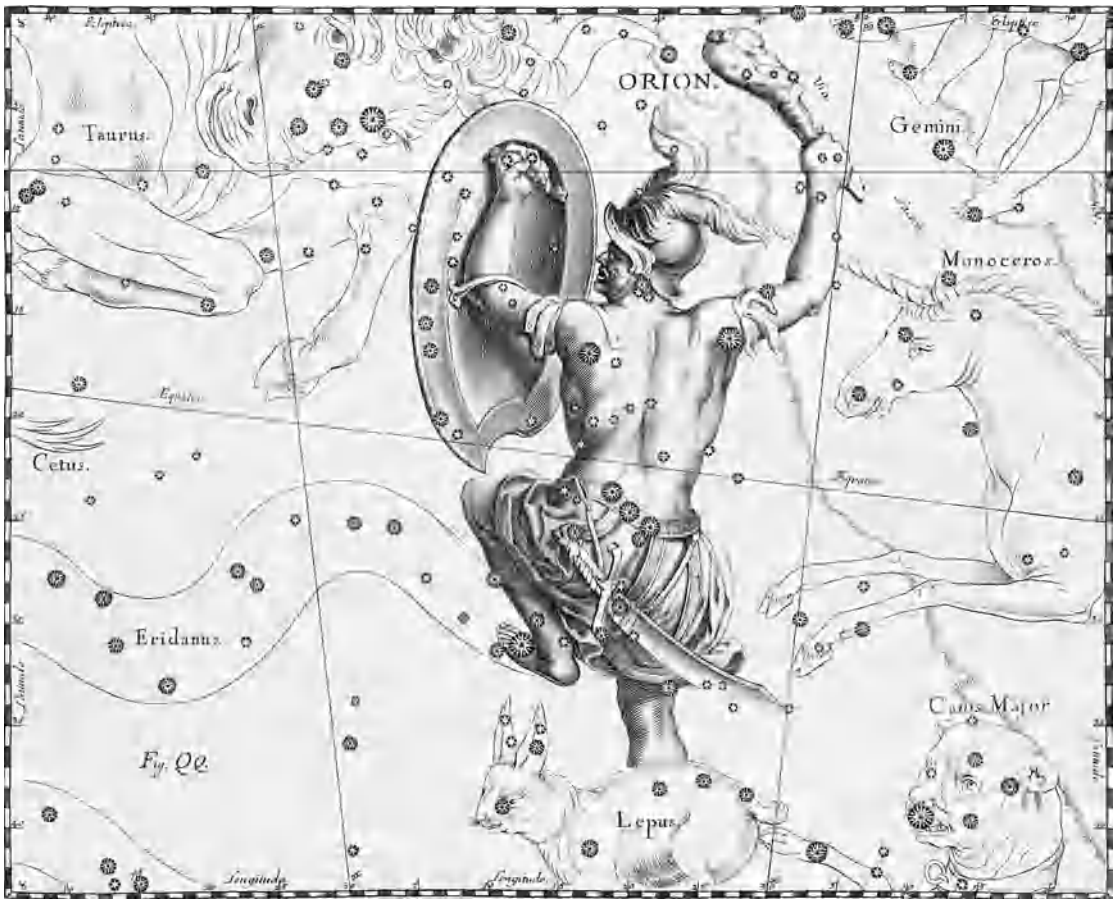

Fig. 2. Orion from Johannes Hevelius's star atlas (1687). The constellation is drawn as it would be seen from outside the celestial sphere and is therefore backward from the way it appears from Earth. (*Jay M. Pasachoff collection*)

configurations, including eclipses and transits, and its star atlas is a minor part of the book, with many constellations shown together.

John Bevis, discoverer of the Crab Nebula in 1731, produced a star atlas based on Bayer's (as is apparent from comparing the constellation figures), using also the star catalogs of Halley and Hevelius. But his apparent partner, the instrument maker John Neale, went bankrupt from the cost of the plates, and the atlas, *Uranographia Britannica*, was never published. Some two dozen sets of plates, mostly in the form of bound atlases, are known (**Fig. 3**). Some of them were sold at auction in 1785 and then to the public as *Atlas celeste* in 1786. In his plates, Bevis depicted not only the main constellations on a page but also the constellations around it in a lighter line strength. Using Hevelius's observations, he wound up with 79 constellations. His atlas had over 3550 stars.

The era of great star atlases ended with the overwhelming *Uranographia* of Johan Elert Bode, published in Berlin in 1801. The book contains 18 maps showing 99 constellations (**Fig. 4**) plus maps of the northern and southern hemispheres. This huge atlas, unwieldy to handle, contained over 17,000 stars, over 10 times more than Bayer's. It also included nonstellar objects that had been cataloged by William Herschel.

Nineteenth-century progress. The cataloging and mapping of nonstellar objects by Charles Messier in 1770 and, more extensively, by William Herschel, subsequent to his discovery of Uranus, led to a shift in emphasis. Herschel's catalog of stars led to the *General Catalogue* and eventually to the *New General Catalogue* and its two additional *Index Catalogues*, which provide many of the notations still in use today. *See* ASTRONOMICAL CATALOGS; MESSIER CATALOG.

Crude popular atlases appeared, such as that by Elijah Burritt in the 1830s in the United States. F. W. A. Argelander's *Atlas des nordlichen gestirnten Himmels*, published in Bonn in 1863, shows over 300,000 stars. The B.D. numbers from the accompanying star catalog, the *Bonner Durchmusterung*, are still in common use. A supplement of 1887 added stars below the equator; it was revised in 1951. The third edition of the northern atlas was not ready until 1954.

Carte du Ciel. In the 1880s, an international consortium set up the *Carte du Ciel*, a project to use 17 identical refracting telescopes around the world to chart the entire sky photographically. It is now held that this project, by tying up European telescopes for decades on slow-going and fruitless work, led to the assumption of the lead in astrophysics by United States astronomers as the twentieth century arrived.

Fig. 3. Taurus and surrounding constellations from John Bevis's unpublished star atlas. (*Jay M. Pasachoff collection*)

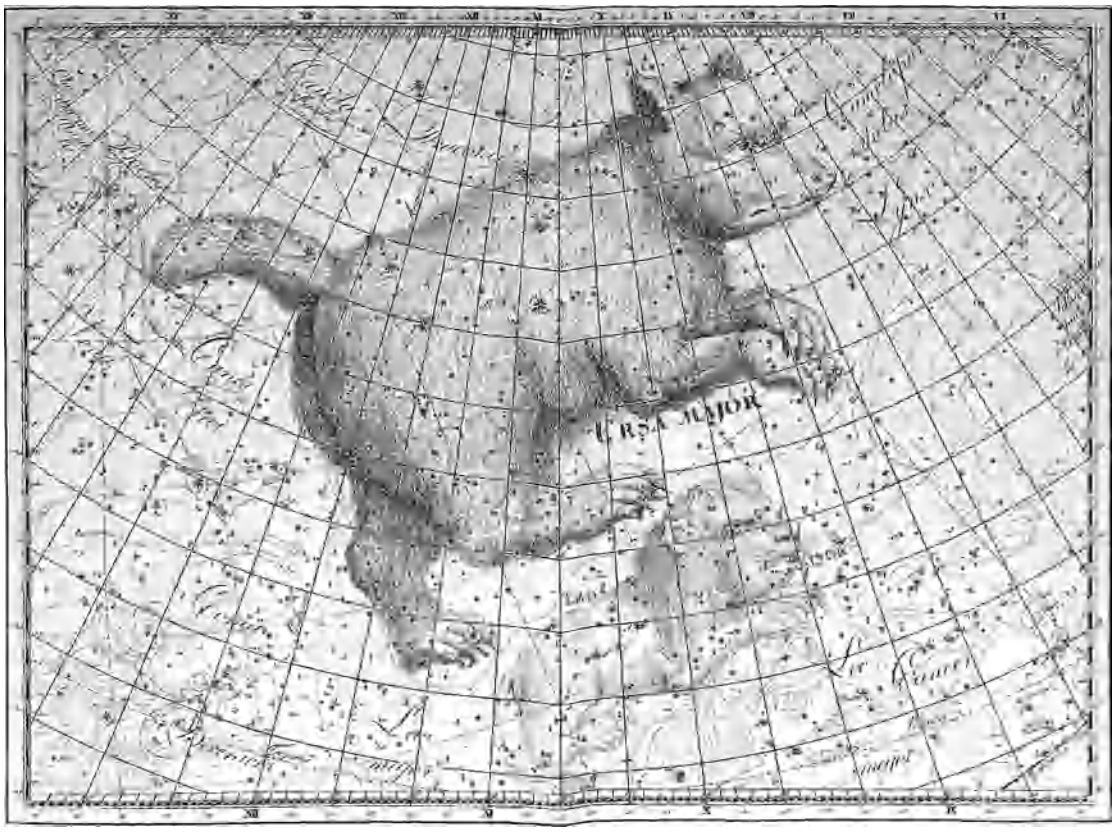

Fig. 4. Ursa Major from Johan Elert Bode's star atlas, *Uranographia* (1801). (*Jay M. Pasachoff collection*)

Photographic atlases. In 1927, E. E. Barnard produced *A Photographic Atlas of Selected Regions of the Milky Way*, edited by E. B. Frost and M. R. Calvert after Barnard's death. Volume 1 contains 51 prints, and volume 2 provides accompanying charts with identifications.

The invention in the 1930s of the Schmidt camera, capable of making wide-field photographs by using fourth-order optics, allowed the whole sky to be mapped. The National Geographic Society–Palomar Observatory Sky Survey began in 1949 to take plates using the 48-in. (1.2-m) Schmidt camera at the Palomar Observatory in California. By 1956, the whole sky above a certain southern latitude, -33°, was photographed in two colors, showing stars down to about 20th magnitude, about a million times fainter than stars that can be seen with the unaided eye. The glass plates, over 900 pairs, were printed onto transparencies and onto paper. The Ohio State University Radio Observatory provided transparent overlays showing and identifying stars, nebulae, galaxies, and other objects.

A correcting plate that was good in different colors was later installed on the Palomar Schmidt camera, which was renamed the Oschin Schmidt camera. A second-generation survey added an infrared color and used finer-grained plates, winding up with 894 sets of three plates.

Both Palomar Observatory Sky Surveys have been digitized. They are now available online, and the original Palomar Observatory Sky Survey can be purchased inexpensively on CD-ROMs.

The United Kingdom Schmidt Telescope at the Anglo-Australian Observatory in Siding Spring, Australia, and a Schmidt telescope at the European Southern Observatory's site at La Silla in Chile were used to photograph the southern sky in two visible colors and one infrared color, completing the sky coverage begun at Palomar. *See* SCHMIDT CAMERA.

Popular atlases. Many atlases showing reasonable numbers of stars for amateur observers appeared in the twentieth century. The most popular included the *Skalnate Pleso Atlas* of Antonin Becvar, compiled at Tatranska Lomnica, now in Slovakia. *Norton's Star Atlas*, with a few dozen plates showing the dome of the sky, has been through many editions.

Wil Tirion of the Netherlands is a leading celestial cartographer whose *Sky Atlas 2000.0* led atlases into the 21st century. He has since changed his method of drawing from handwork to computers and has provided atlases for many sources in a variety of projections. *Sky Atlas 2000.0*'s second edition, by Tirion and Roger W. Sinnott, was published in 1998.

In his 1963 *Peterson Field Guide to the Stars and Planets*, Donald H. Menzel used photographic star charts. Starting in 1983, a reduced-scale version of Tirion's atlas started appearing in the *Peterson Field Guide to the Stars and Planets*, then under the names of Menzel and Jay M. Pasachoff, revised in 1992 under the names of Pasachoff and Menzel and in 2000 by Pasachoff.

Contemporary amateur astronomers and others rely on *Sky Catalogue 2000.0* by Alan Hirshfeld, R. W. Sinnott, and François Ochsenbein (1991, 1985) and *NGC 2000.0*, edited by Sinnott. Sinnott and Michael A. C. Perryman did the *Millennium Star Atlas* (1997), using the accurate positions measured by the *Hipparcos* satellite of the European Space Agency. Tirion, Barry Rappaport, and George Lovi published *Uranometria 2000.0*, covering the whole sky, in 1986 with stars to magnitude 9.5 and an essay on historic atlases. Tirion, Rappaport, and Will Remaklus produced *Uranometria 2000.0 Deep Sky Atlas*, with volume 1 dealing with the northern hemisphere and volume 2 with the southern hemisphere. The volumes show over 280,000 stars on 220 star charts down to magnitude 9.75, plus over 25,000 galaxies, and use *Hipparcos* data. *See* PARALLAX (ASTRONOMY).

Galaxies. The deep-sky atlases mentioned above chart the apparent two-dimensional positions of galaxies. In his *Hubble Atlas of Galaxies*, (1961, 1984), Allan Sandage illustrated the various galaxy types but not their positions. *The Carnegie Atlas of Galaxies* (1994) by Sandage and John Bedke extended the work. Hans Vehrenberg, a distinguished amateur, covered the whole sky photographically in his *Atlas of Deep Sky Splendors*, showing the orientation of the objects.

A true mapping of space through the three-dimensional positions of galaxy clusters was published by Brent Tully of the University of Hawaii. In the early twenty-first century, two projects have mapped hundreds of thousands of galaxies and extended three-dimensional knowledge to greater distances than ever before on a uniform scale. They are the 2dF project carried out at Siding Spring, Australia, and the Sloan Digital Sky Survey. They are available online and in various technical publications. *See* GALAXY, EXTERNAL.

Solar-system atlases. Galileo made seven wash drawings of the lunar surface in 1609, and his *Sidereus Nuncius* published seven other engravings of the lunar surface in 1610. Hevelius's *Selenographia*, published in Gdansk in 1647, mapped the lunar surface in detail. A full set of lunar photographs from ground-based observations is the *Photographic Lunar Atlas* were produced by Gerard P. Kuiper in 1960. The National Geographic Society produced an airbrushed set of lunar maps; a reduced-scale version is available in Pasachoff's *Peterson Field Guide to the Stars and Planets*. In 1999, Michael Light produced *Full Moon*, in which original NASA *Apollo* photographs of the lunar surface were reprocessed. *See* MOON.

Ron Greeley and R. Batson have produced the *NASA Atlas of the Solar System* (1997) and a *Compact NASA Atlas of the Solar System* (2001). Gridded charts of various solar-system objects are produced on a regular basis by the U.S. Geological Survey's Flagstaff office of Astrogeology.

Twenty-first-century atlases. Atlases are now available online. The *Hubble Space Telescope's Guide Star Catalogue* provides positions for over 16 million objects, and can be called up on a computer

in atlas form to make charts. *See* SPACE TELESCOPE, HUBBLE.

Programs are available for home computers that allow users to chart any given area of the sky on any scale, including deep-sky objects and solar-system objects. Capabilities continue to expand.

Jay M. Pasachoff

Bibliography. R. Greeley and R. Batson, *Compact NASA Atlas of the Solar System*, Cambridge University Press, 2001; M. Light, *Full Moon*, Knopf, 1999; J. M. Pasachoff, *Peterson Field Guide to the Stars and Planets*, 4th ed., Houghton Mifflin, 2000; A. Sandage and J. Bedke, *The Carnegie Atlas of Galaxies*, 2 vols., Carnegie Institution of Washington, 1994; R. W. Sinnott and Michael A. C. Perryman, *Millennium Star Atlas*, Sky Publishing Corp., 1997; W. Tirion, B. Rappaport, and W. Remaklus, *Uranometria 2000.0 Deep Sky Atlas*, 2 vols., Willmann-Bell, 2001; W. Tirion and R. W. Sinnott, *Sky Atlas 2000.0*, 2d ed., Sky Publishing Corp., 1998.

Astronomical catalogs

Lists or enumerations of astronomical data relevant to astronomy, navigation, geodesy, and space science applications. Astronomical catalogs vary a great deal in form and content depending upon the type of the data and the objects to which they are referring. The essential data may be positions and motions of the celestial objects, magnitudes, spectra, radial velocities of stars, or energy fluxes. The celestial objects may be stars, galaxies and other galactic and extragalactic objects, or solar system bodies, which are generally ordered by increasing right ascensions and identified by a catalog number. Catalogs are either derived directly from observations or are compiled from different sources.

Access to the catalogs. Astronomical data centers have been established by international exchange agreements to develop and provide, in machine-readable form and on-line, catalogs that were previously published in book form. The Astronomical Data Center (ADC) at the National Aeronautics and Space Administration's Goddard Space Flight Center discontinued its services in 2002. The Centre de Données Astronomiques de Strasbourg (CDS), in France, currently provides the most complete library of published astronomical catalogs available on line with standard documentation and format and a relational database management system. This service is also available in Canada, the United States, Japan, India, the United Kingdom, and China. Other data centers provide access to data and catalogs related to specific projects or astronomical objects. Such data centers are now developed in the general frame of the Virtual Observatory, by which astronomers will have World Wide Web access to any kind of data in common standardized formats. *See* WORLD WIDE WEB.

Reference systems and fundamental catalogs. The positions and motions listed for celestial objects should refer to a primary reference system which is realized by a reference frame. A considerable change was the transition from a reference system realized by stellar positions and proper motions to the International Celestial Reference System (ICRS), based on observed extragalactic radio sources. The ICRS and the International Celestial Reference Frame (ICRF) that realizes the ICRS have been adopted by the International Astronomical Union starting from 1998.

The ICRS is centered at the solar system barycenter and is nonrotating with respect to the ensemble of distant extragalactic objects. It has no intrinsic orientation but was aligned close to the mean equator and dynamical equinox of J2000.0 (the date January 1, 2000, at 12 hours Terrestrial Time) for continuity with previous fundamental reference systems. *See* ASTRONOMICAL COORDINATE SYSTEMS; DYNAMICAL TIME; EQUATOR; EQUINOX.

The ICRF is the catalog of adopted positions (right ascensions and declinations) of 608 extragalactic radio sources, mostly quasars, observed with very long baseline interferometry (VLBI), which is currently the most accurate realization of the ICRS. Of these objects, 212 are defining sources, with position uncertainties of the order of 0.5 milliarcsecond (mas), that establish the orientation of the ICRS axes. Extensions of the ICRF were released in 1999 and 2004 with 59 and 109 additional sources, respectively. *See* ASTRONOMICAL COORDINATE SYSTEMS; QUASAR; RADIO TELESCOPE.

The ICRS is realized at optical wavelengths by a subset of about 100,000 stars in the *Hipparcos Catalogue*. This is referred to as the *Hipparcos* Celestial Reference Frame (HCRF) which realizes the *Hipparcos Catalog* Reference System (HCRS). *See* ASTROMETRY; CELESTIAL REFERENCE SYSTEM.

Pre-Hipparcos catalogs of stellar positions. The *Fifth Fundamental Catalogue* (*FK5*), published in 1988 by the Astronomisches Rechen-Institut in Heidelberg, Germany, was the International Astronomical Union (IAU) realization of the celestial reference system up to 1998. It contains the positions at J2000.0 of 1535 stars, mostly brighter than ninth magnitude, and an additional 3117 stars, down to a magnitude of about 9.5. *See* MAGNITUDE (ASTRONOMY).

Moderately bright stars (seventh to ninth magnitude), selected on the basis of one star per square degree of the sky, have been related to the *FK5* by meridian circle observations in order to form a system of sufficient density to serve as position references for photographic observations. Typical of catalogs of such reference stars is the *International Reference System* (*IRS*) catalog (1991), which was the world standard reference system in the pre-*Hipparcos* era. *See* ASTRONOMICAL TRANSIT INSTRUMENT.

A photographic survey of the northern celestial hemisphere resulted in the *AGK3* (1975), the third in a series of catalogs published by the German Astronomical Society. A revised version of this catalog has been published by the Astronomisches Rechen-Institut as the *Position and Proper Motion* (*PPM*) *Catalog* (1991) of 181,731 stars brighter than the 9th magnitude, north of -2.5° declination. A

similar photographic catalog for the southern celestial hemisphere is the *Second Cape Photographic Catalog* (*CPC2*) [1993] of 250,000 stars to a limiting magnitude of 10.5.

An important photographic catalog covering the entire sky to a limiting magnitude of 13 was the result of an international undertaking involving 19 observatories, with each assigned zones of declination and observing with nearly identical telescopes. The catalog, known as the *Carte du Ciel* (*CdC*) or *Astrographic Catalogue* (*AC*), finally completed for all zones in 1964, provides star positions in the form of rectangular coordinates, as measured on the plates. A reduction of the *AC* was made at U.S. Naval Observatory (USNO) in Washington, D.C. using its *Astrographic Catalogue Reference Stars* (*ACRS*) with positions and proper motions of about 320,000 stars.

The *Smithsonian Astrophysical Observatory Star Catalogue* (*SAOC*) was compiled from selected visual and photographic catalogs, initially for the determination of artificial satellite positions. It contains positions and proper motions of 259,000 stars over the entire sky. It was published in book form (1966) and updated in machine-readable form with a reduction to the *FK5* system.

The Guide Star Catalog (*GSC*), constructed for the pointing of the *Hubble Space Telescope*, provides positions and magnitudes for nearly 19 million objects in the 6th to 15th magnitude range, of which more than 15 million are classified as stars. The *USNO-A1.0 Catalog* contains *GSC*-based positions of 488,006,860 sources which were detected by the U.S. Naval Observatory's Flagstaff Station Precision Measuring Machine (PMM) in the framework of several photographic sky surveys. *See* HUBBLE SPACE TELESCOPE.

Hipparcos and Tycho Catalogues. The *Hipparcos* and *Tycho Catalogues* (1997) have both resulted from instruments on board the European Space Agency's *Hipparcos* satellite, operational from late 1989 to 1993. They include a large quantity of astrometric and photometric data, as well as annexes featuring variability and double- and multiple-star data. *See* PARALLAX (ASTRONOMY).

The *Hipparcos Catalogue* of 118,218 stars is complete to visual magnitude $V = 7.3$ and includes stars as faint as visual magnitude 12. It was linked to the ICRF through radio stars. Accuracies are 1 to 3 mas for positions at epoch 1991.25 and 1 to 2 mas/year for proper motions. *Hipparcos* also contains broadband visual photometric data including variability information in the 1–2 millimagnitude range for the brighter stars.

The *Tycho Catalogue* (also named *Tycho-1 Catalog*) of about 1 million stars is an extension of the *Hipparcos Catalogue* that is more complete at magnitudes 8–11 but less accurate by a factor of 30. It contains two-color photometry for most of its stars.

Post-Hipparcos star catalogs. New versions of star catalogs have been formed by referring them to the HCRS, such as the *AC 2000* and the *USNO-A2.0* catalog. The *USNO B1.0 Catalog* (2003) expands this series in providing 1,042,618,261 stars or galaxies and in the inclusion of proper motions. It is expected to be complete down to magnitude 21 with accuracies of 200 mas for positions at J2000, 0.3 mag for brightnesses in up to five colors, and 85% for distinguishing stars from nonstellar objects.

Several catalogs have aimed to combine Tycho observations with century-old positions, such as the *Tycho Reference Catalogue* (*TRC*), the *ACT Reference Catalog*, and the *All-Sky Compiled Catalogue* (*ASCC-2.5*). They have been superseded by the *Tycho-2 Catalog* of 2,539,913 brightest stars, which combines a sophisticated reduction of the Tycho stars and 144 ground-based star catalogs. Accuracies are from 10 to 100 mas for positions and from 1 to 3 mas/year for proper motions.

The *USNO CCD Astrographic Catalog* (*UCAC*) has been undertaken for global sky coverage down to 16th magnitude at red wavelengths. Preliminary releases of this catalog are *UCAC1* in the southern hemisphere and *UCAC2* covering 86% of the sky. The latter is a compiled catalog of 48 million stars which are primarily in the 8.0–16.0 magnitude range, including all catalogs used for *Tycho-2*, plus the recent epoch ground-based observations of the *UCAC* project. Accuracies are 20 to 70 mas for positions and 1 to 7 mas/year for proper motions.

A revision of the *Guide Star Catalog, GSC II*, is intended to provide positions, proper motions, magnitudes, and classifications for all objects derived from the Digitized Sky Survey. A preliminary version (*GSC2.2*) includes 435,457,355 objects with magnitude limits 18.5 and 19.5 in the two relevant photographic wavelengths, respectively.

Catalogs of nonstellar objects. J. L. E. Dreyer's *New General Catalogue of Nebulae and Starclusters* (1888) contains 7840 objects originally classified as nonstellar, and was supplemented by two Index Catalogs in with an additional 5386 objects. The NGC and IC numbers assigned in these catalogs remain the most commonly used designations. A *Master List of Nonstellar Astronomical Objects* (1980) provides approximately 185,000 listings from 270 catalogs, with multiple listings of objects appearing in several catalogs. *See* GALAXY, EXTERNAL; NEBULA; STAR CLUSTERS.

With the expansion of astronomical research beyond the visible part of the electromagnetic spectrum to the ultraviolet, x-ray, and gamma-ray regions at short wavelengths and to the infrared and radio regions at long wavelengths, and with modern very large sky surveys, the number of nonstellar objects has increased dramatically. An *Index of Extragalactic Radio Source Catalogs* (1977) provides precise references to the large number of individual catalogs. In this category, the third version of the *Reference Catalog of Bright Galaxies* (1991) contains 23,022 objects, the *Cambridge Observatory Discovery Catalogs of Quasars*, named *3C* and *4C* (1959 and 1965), are reference catalogs; and the *Catalog of Quasars and Active Galactic Nuclei*, (11th edition, 2003) contains 48,921 entries with position, redshift, photometry, and flux densities. *See* RADIO ASTRONOMY; SLOAN DIGITAL SKY SURVEY.

Catalogs and databases of several million extragalactic objects with various facilities (including cross-identification) are available at the NASA/IPAC Extragalactic Database (NED) and the HYPERLEDA database maintained at Lyon Observatory (France), with mirror sites in several countries.

The number of known infrared sources has rapidly increased, especially since the launch in 1982 of the *Infrared Astronomical Observatory Satellite* (*IRAS*), whose survey has provided the *IRAS Faint Source Catalog* and the *IRAS Catalogue of Point Sources*, which contains some 250,000 infrared point sources. Astrometric positions of stellar sources from the *IRAS Point Source Catalog* have been published by the U.S. Naval Observatory in the *Catalog of Positions of Infrared Stellar Sources* (2001), containing 37,700 entries with *Hipparcos* and *Tycho-2* astrometry and photometry.

The Two Micron All Sky Survey (2MASS) project has imaged the entire sky in three near-infrared bands, producing a point-source catalog (2000, 2003) of 470 million objects, mostly stars, and an extended-source catalog of 1.6 million objects, with accuracy in position of about 70 mas. The Deep Astronomical Survey of the Southern Sky (DENIS), in two near-infrared bands and one optical, was conducted by a European consortium and produced (2003) a catalog of 195,204,157 point sources.

Ultraviolet astronomical observations of stellar and nonstellar objects have been obtained from rocket vehicles, spacecraft, and satellites. Catalogs of these photometric and spectroscopic observations are available at CDS. *See* ULTRAVIOLET ASTRONOMY.

Early discoveries of x-ray sources were from rocket observations, but a rapid increase in known sources occurred after the launch of the *UHURU* satellite in 1970 and other satellites which followed. Catalogs and databases of these observations are available from NASA High Energy Astrophysics Science Archive Center (HEASARC) or CDS. The *ROSAT All-Sky Bright Source* and *ROSAT All-Sky Faint Source Catalogs* are among the most frequently used catalogs in this category. *See* CHANDRA X-RAY OBSERVATORY.

Gamma rays have been detected from quasars, galaxies, and highly evolved stars. Catalogs of observations from spacecraft and satellites, such as the *International Gamma-Ray Astrophysics Laboratory* (*INTEGRAL*), are available from HEASARC. *See* GAMMA-RAY ASTRONOMY.

Catalogs of astrophysical data. Among the numerous catalogs in this category is the monumental *Henry Draper Catalog* (*HD*) of spectral classification, which with its extension includes data for 275,000 stars, with an updated machine-readable version (1989). The introduction of luminosity classes in spectral classification, such as the prevailing Morgan-Keenan-Kellman (MK) system (1943), led to classification of some 33,000 stars in this system, compiled into a catalog by the Mount Stromlo Observatory in Australia (1981). This catalog also contains the magnitudes and colors of stars on the U, B, V photometric system. The Asiago Database on Photometric Systems (ADPS) project (2000) aims to list and investigate all existing photometric systems, and provides a link to a regularly updated *General Catalogue of Photometric Data* with an associated master index. *See* ASTRONOMICAL SPECTROSCOPY; SPECTRAL TYPE.

Catalogs of stellar radial velocities are listed in two bibliographies produced by the Kitt Peak National Observatory, covering data upto 1970, and the *General Catalog of Mean Radial Velocities* (2000) has been compiled with data from 1970 to 1990.

Catalogs of special stellar objects. There are many catalogs in this category. Among the most frequently used are the *Combined General Catalog of Variable Stars* (*GCVS*); the *Washington Double Star Catalog* (*WDS*) of the U.S. Naval observatory (2000), which is the official database for about 100,000 double and multiple stars for the IAU; the *Bright Star Catalog* of all stars brighter than magnitude 6.5; and the *Yale General Catalog of Trigonometric Stellar Parallaxes* (4th edition, 1994), containing more than 15,000 entries with a precision of 4 mas. *See* BINARY STAR; STAR; VARIABLE STAR.

Nicole Capitaine

Bibliography. K. Aa. Strand, *Basic Astronomical Data*, 1963; H. Eichhorn, *Astronomy of Star Positions*, Ungar, 1974; K. J. Johnston et al. (eds.), *Towards Models and Constants for Sub-Microarcsecond Astrometry*, International Astronomical Union Colloquium 180, 2000 ; J. Kovalevsky and P. K. Seidelmann, *Fundamentals of Astrometry*, Cambridge University Press, 2004; H. G. Walter and O. J. Sovers, *Astrometry of Fundamental Catalogs: The Evolution from Optical to Radio Reference Frames*, Springer, 2000.

Astronomical coordinate systems

Schemes for locating astronomical objects in space. To an observer on the Earth's surface, the stars of the night sky appear to be placed upon a spherical shell of infinite radius with the observer at the center. Celestial objects appear to move or to maintain constant locations with respect to the stars, and at any given time their position on this imaginary sphere, called the celestial sphere, can be specified by two angles, called celestial coordinates, whose values depend upon what coordinate system is used. *See* CELESTIAL SPHERE.

Each coordinate system is defined by a fundamental plane and a principal axis. For example, on the Earth's surface, longitude and latitude coordinates are used to determine positions. In this system, which is analogous to astronomical coordinate systems, the fundamental plane is that of the Earth's Equator, and the principal axis is defined by a line running from the Earth's center to a point on the Equator at the longitude of Greenwich, England. Longitude is measured east or west of Greenwich. Angular positions along the fundamental plane, like longitude, can be expressed in terms of either degrees or hours of time. One hour equals 15°. Angular measurements north or south of the fundamental plane, like

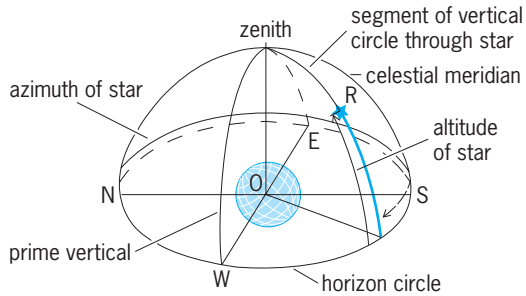

Fig. 1. Horizon system of astronomical coordinates.

latitude, are measured from 0 to 90° at the poles. In astronomical coordinate systems, the degrees north are denoted with a plus sign (+), and degrees south are with a minus sign (−). *See* EQUATOR; LATITUDE AND LONGITUDE.

Horizon system. The boundary between the hemisphere of the sky that is visible and the hemisphere which is hidden from view by the Earth is called the horizon (**Fig. 1**). The observer is located at the center of the system (*O*), the pole directly overhead is termed the zenith, and the opposite pole, the nadir. These pole directions are aligned with a plumb line, which is determined by the observer's local gravity. The fundamental horizon plane is 90° from the poles, and for astronomical applications the principal axis is most often taken to pass through the north point. Great circles that pass through the zenith and nadir are termed vertical circles; the one passing through the east and west points is termed the prime vertical, and that passing through the north and south points is called the celestial meridian. *See* HORIZON; ZENITH.

The longitudinal coordinate of a celestial object is termed its azimuth and is most often measured eastward from the north point to the object's vertical circle; and the latitudinal coordinate, termed its altitude, is measured along the object's vertical circle, north or south from the horizon plane to the object. For example, the position of the star *R* in Fig. 1 would have an approximate azimuth and altitude of 210° and +60°. The object's zenith distance, measured along the object's vertical circle, is the angle measured from the zenith point to the object.

Because the horizon coordinate system is fixed with respect to the observer on the Earth's surface, it must rotate with the Earth. Hence the azimuth and altitude of a celestial object are constantly changing with time. For many astronomical applications, such as the listing of stellar and planetary positions, a nonrotating, coordinate system is required. The most common of these nonrotating, or inertial, coordinate systems is the equatorial system.

Equatorial system. The fundamental plane of the equatorial coordinate system can be visualized by imagining that the Earth's equatorial plane is extended to intersect the celestial sphere. An alternate fundamental plane, the ecliptic plane, is the extension of the Earth's mean orbital plane onto the celestial sphere (**Fig. 2**). These planes intersect at two points, called equinoxes, with the angle between them ϵ being termed the obliquity of the ecliptic. This angle is about 23.4°. *See* ECLIPTIC; EQUINOX.

Due to the Earth's motion about the Sun, observers on Earth see an apparent motion of the Sun along the ecliptic plane. The point where the Sun's annual apparent motion takes it northward across the equatorial plane is called the vernal equinox ♈, and the line between the Earth's center and this point defines the principal axis for both the equatorial and ecliptic coordinate systems. The apparent passage of the Sun through the vernal equinox, on about March 21, marks the beginning of spring in the Northern Hemisphere. Because of disturbing effects of the Sun and Moon on the Earth's figure, the Earth's rotation axis precesses, causing the celestial pole to describe an approximate circular motion about the ecliptic pole once every 26,000 years. This causes the location of the vernal equinox to drift westward along the ecliptic about 50 arc-seconds each year. Hence for an inertial coordinate system, where the principal axis is not moving, an epoch must be specified at which time the coordinate system is held fixed. In practice, the beginning of the year 2000 is most often used as an epoch. *See* EARTH ROTATION AND ORBITAL MOTION; PRECESSION OF EQUINOXES.

The north and south celestial poles represent the extension of the Earth's North and South poles onto the celestial sphere. For a celestial object (for example, object *X* in Fig. 2), the longitudinal coordinate is termed the right ascension α and is measured eastward along the celestial equator from the vernal equinox ♈ to the great circle passing through the object and the north and south celestial poles. The latitudinal coordinate, called the declination δ, is then measured along the object's great circle, north or south from the Equator to the object.

Ecliptic system. The ecliptic coordinate system is often used when representing the orbital motions of the planets, asteroids, and comets. The fundamental plane is that of the ecliptic, and as in the equatorial system, the principal axis is the line extending from the Earth's center to the vernal equinox. The position of a celestial object is defined, in the ecliptic system, by the ecliptic longitude and latitude. The longitude

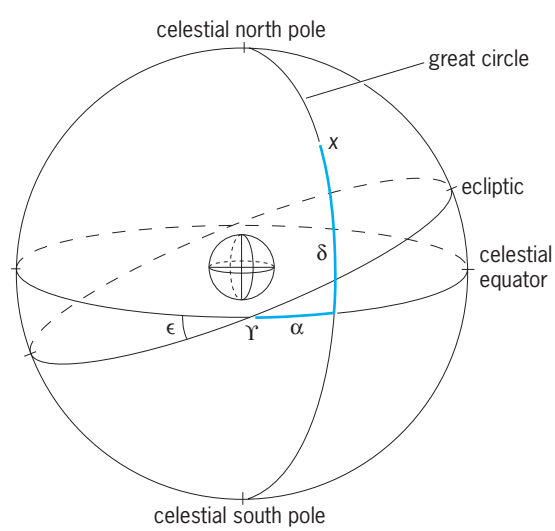

Fig. 2. Equatorial system of astronomical coordinates.

is measured eastward along the ecliptic plane from the vernal equinox to the great circle that passes through the object as well as the north and south ecliptic poles. The ecliptic latitude is then measured along this great circle either north or south from the ecliptic plane to the object.

Galactic system. Astronomers working with stars and other objects within the Milky Way Galaxy often find it convenient to use the galactic disk as the fundamental plane of their coordinate system, and the line extending from the galactic center to the Sun's location as the principal axis. With the Sun as the origin, galactic longitude is measured from the principal axis along the galactic equator eastward to the great circle that passes through the object and the north and south galactic poles. Galactic latitude is measured along the object's great circle, either north or south from the galactic plane to the object. In equatorial coordinates, the position of the north galactic pole, at the 2000 epoch, is approximately $\alpha = 12.9$ h (192.9°) and $\delta = +27.1°$. *See* MILKY WAY GALAXY.

Donald K. Yeomans

Bibliography. S. Bell and C. Hohenkerk, Correspondence: Controversy in *The Astronomical Almanac 2006, The Observatory*, 125:271–272, 2005; P. Duffett-Smith, *Practical Astronomy with Your Calculator*, 3d ed., Cambridge University Press, 1988; C. A. Murray, *Vectorial Astronomy*, Adam Hilger, 1983; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, University Science Books, 1992; L. G. Taff, *Computational Spherical Astronomy*, 1981, rev. ed., Krieger, 1991.

Astronomical imaging

The production of a permanent two-dimensional record of a scene of interest to astronomy. In direct imaging, the scene is a section of the sky optically imaged with a telescope. In other types of imaging, the scene is the output of an instrument, often a spectroscope. Although astronomical images can be produced using light from any part of the electromagnetic spectrum, this article mainly concerns images made in the optical and near-infrared regions (wavelengths of 320–1000 nanometers). The overriding concerns in astronomical imaging are resolution, wavelength sensitivity, and above all, detective quantum efficiency.

Astronomers can use three techniques to construct an image. The first, the subject of this article, employs a two-dimensional, multielement panoramic detector such as a photographic plate or the retina of the human eye to detect the elements of the scene. The other two techniques build an image over time, either by scanning the scene with a single-element detector or by interferometry (interpreting the changing interference pattern of light waves detected by multiple telescopes as the spacing of the telescopes changes). *See* INTERFEROMETRY.

Most panoramic detectors belong to a class called photon detectors, which make use of an individual photon's ability to alter the quantum-mechanical state of one or more electrons in the detector. The other detector classes, thermal detectors and wave detectors, are difficult to fabricate as panoramic devices.

Arrays. At optical wavelengths as well as in the very near infrared, astronomers usually select the charge-coupled device (CCD) as the best available panoramic detector. CCDs are arrays—they consist of a grid of spatially discrete photon detectors, rather than the almost continuous distribution of sensitive emulsion found on a photographic plate. *See* CHARGE-COUPLED DEVICE.

At other wavelengths, astronomers also use arrays of photon detectors. At shorter wavelengths (x-rays and ultraviolet light), these are sometimes modified CCDs, but at wavelengths longer than 1.0 micrometers (infrared to submillimeter) the arrays are other devices. The following discussion concentrates on CCDs, but most of the general principles discussed apply to imaging with any detector array. *See* INFRARED ASTRONOMY.

MOS capacitor. Each detecting element in a CCD array is a metal-oxide-semiconductor (MOS) capacitor, a good example of a photon detector. The basic structure of this very simple solid-state device is a three-layer sandwich (**Fig. 1**). As illustrated, the left-hand layer is a block of *p*-type silicon semiconductor whose left-hand face is connected to electrical ground. The middle of the sandwich is a thin layer of insulator (usually silicon dioxide). The right-hand layer is a thin coating of metal (usually highly doped polysilicon), which is held at a positive voltage. *See* SEMICONDUCTOR; TRANSISTOR.

The conductivity of the *p*-type silicon is due to positively charged holes (the majority carriers), which are repelled toward ground by the positive voltage of the metal layer, producing a depletion region devoid of charge carriers. Illuminated through the transparent electrode, the device is sensitive to light because a photon absorbed in the depletion region will ionize a silicon atom, creating an electron-hole pair. The electron will move to the right, toward the positive electrode (the potential well), but cannot cross the nonconducting oxide layer. For every photon absorbed by the silicon, another electron is stored at the semiconductor-oxide boundary. At the end of an exposure, the charge on the capacitor is read out and converted to a digital signal.

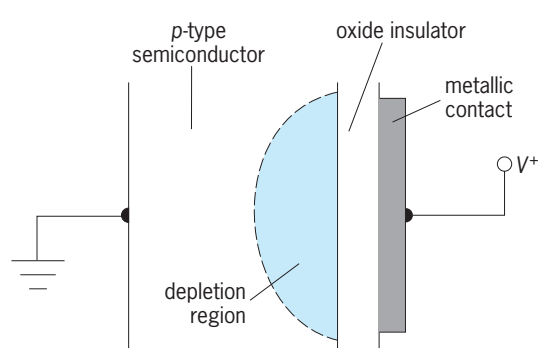

Fig. 1. Cross-sectional view of the physical structure of an MOS capacitor. Positive voltage (usually a few volts) applied to the metal layer creates a depletion region in the semiconductor.

Characterization of Array Detectors

From roughly 1880 to 1980, photography served as the primary method of astronomical imaging. What motivated professional astronomers to move to CCDs and other arrays? More generally, what makes one detector better than another? All the following characteristics are important.

Efficiency. Astronomers construct telescopes to gather as many photons as possible, and it seems perverse if a detector does not use a large fraction of these expensive photons to produce its signal. The quantum efficiency (QE) is a common measure of detector efficiency. It is usually defined as the fraction of photons incident on the detector that actually contribute to the signal. In a perfect detector, every incident photon would be absorbed in a fashion that contributes equally to the signal, and the detector would have a QE of 1.00.

Noise. Although efficiency in a detector is important, the real value of any measurement depends crucially on its uncertainty. The uncertainty in the output signal produced by a detector is often called the noise, and the signal-to-noise ratio (SNR) is a widely used indicator of the quality of a measurement. However, a perfect detector does not produce a signal with zero noise and infinite signal-to-noise ratio. *See* ELECTRICAL NOISE; SIGNAL-TO-NOISE RATIO.

There is an uncertainly inherent in measuring the strength of any incident stream of photons. For a photon-counting device, this uncertainty arises from the Poisson statistics of random photon arrivals, and is given by Eq. (1), where N is the number of photons

$$\sigma = \sqrt{N} \tag{1}$$

actually counted and σ is the uncertainty in N. A perfect detector, with QE = 1, counts all incident photons and will therefore produce a signal-to-noise ratio given by Eq. (4).

$$\mathrm{SNR}_{\mathrm{perfect}} = \frac{N_{\mathrm{out}}}{\sigma_{\mathrm{out}}} = \frac{N_{\mathrm{in}}}{\sigma_{\mathrm{in}}} = \sqrt{N_{\mathrm{out}}} = \sqrt{N_{\mathrm{in}}} \tag{2}$$

Real detectors will have a smaller signal-to-noise ratio, because they either count fewer photons or add output noise. The detective quantum efficiency (DQE) describes this departure of a real detector from perfection and, better than the QE alone, measures the quality of a detector in a particular situation. For a perfect detector, DQE = QE = 1. For any detector, DQE $\leq$ QE.

In detecting faint sources, a typical photographic emulsion, hypersensitized to increase its QE, might have QE = 0.04 and, because of noise due to emulsion granularity, DQE = 0.005. Modern astronomical CCDs degrade the information in the input hardly at all, and their DQEs are close to their QEs—in the range 0.20–0.95. Thus, both efficiency and noise considerations strongly select the CCD over photography. *See* PHOTOGRAPHY.

Spectral response and discrimination. The QE of a detector is generally a function of the wavelength of the input photons. Some excellent detectors are useless at some wavelengths. For example, silicon devices, including CCDs, cannot respond to photons with wavelengths greater than 1.1 μm. One can imagine an ideal detector that measures both the intensity and the wavelength of each incoming photon. Designs of such detectors have been proposed and might well replace CCDs if proven equivalent in DQE and other properties.

In professional astronomy, color images (which illustrate rough spectral distributions) are almost always assembled indirectly by combining three separate monochrome frames with different spectral responses.

Linearity. In an ideal detector, the output signal is directly proportional to the input illumination. Departures from strict linearity are common in modern infrared-sensitive devices and are a serious problem in photography. CCDs are highly linear devices except at very high levels of input, where individual array elements will saturate. Further increases in input will not move the output signal above the saturation level. *See* SATURATION.

Response time. The minimum time interval over which the detector can measure changes in illumination is an important parameter. Readout procedures for large astronomical CCDs, for example, can limit their response time to several seconds or more, and measuring rapid brightness changes is an application where astronomers replace CCDs with specialized devices.

Size, pixel number, and array resolution. Clearly, the physical size of each detector of an array will determine how closely spaced its elements, or pixels (for picture elements) can be. Large arrays are more easily manufactured for some types of detectors (such as MOS capacitors) than for others (such as bolometers and wave detectors). There is an obvious advantage in field-of-view for detectors with a large number of pixels. For any array recording telescopic images, it is best if the pixel size and telescope resolution match. Astronomical imaging is improved if the resolving power of the telescope is improved, either by placing it in space or by employing adaptive optics. *See* ADAPTIVE OPTICS.

At the beginning of the CCD era, photographic plates had a clear advantage in pixel number: For a very moderate cost, a photographic plate had a very large area (tens of centimeters on a side), and thus, in effect, contained up to 10^9 pixels (1000 megapixels). Mosaics of CCD arrays, although quite expensive, are beginning to approach the pixel count of medium-sized photographic plates.

Data handling. An important characteristic of modern solid-state detectors is the ease with which their output can be stored as a digital image accessible by a computer. This is critically important in the usual situation in which an astronomer wishes to analyze the image quantitatively. It is nearly essential for images transmitted by space-based telescopes.

General Operation and Properties of CCDs

Researchers at Bell Laboratories produced the first CCDs in the early 1970s as a kind of computer memory. A few astronomers quickly recognized the

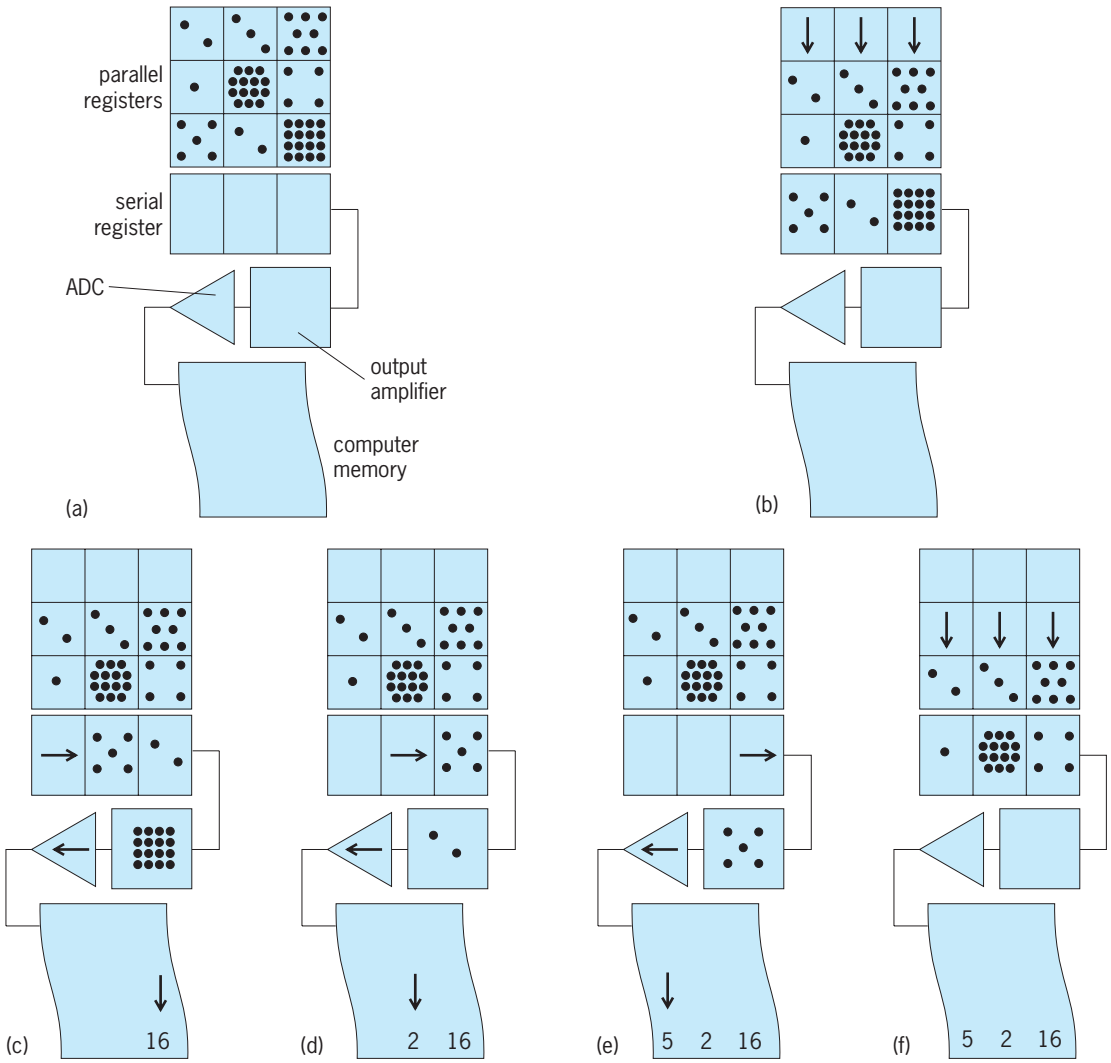

Fig. 2. CCD components and readout. (*a*) Accumulated photoelectrons in a 3 × 3 array of capacitors. Reading the array: (*b*) The bottom row shifts into the serial register; all remaining rows shift down by one pixel in the parallel register. (*c*–*e*) Serial register reads one column at a time. (*f*) Step *b* is repeated: all the rows shift down one pixel.

device's potential as a panoramic light detector, and by 1976 had recorded the first CCD images of celestial objects. The CCD array has become a standard component in military, industrial, and consumer applications that include scanners, copiers, and still and video cameras. Such widespread applicability has supported research and development costs. The resulting rapid evolution of the scientific CCD, initially spurred by the design for detectors used in the *Galileo* mission to Jupiter, has revolutionized the practice of optical observational astronomy.

To understand the basic ideas behind the array operation, one may imagine a matrix of MOS capacitors placed behind a shutter in the focal plane of a telescope. Photons strike the pixels of the array, and each accumulates photoelectrons at a rate that is proportional to the intensity of light incident on the pixel. At the end of the exposure, the array contains an electronic record of the image.

The defining attribute of the CCD is the method whereby the electron pattern stored in the array is converted into a useful form (numbers in a computer). One may consider, for example, an array of detector pixels three columns wide and three rows tall (**Fig. 2***a*). Each column of pixels is called a parallel register. There is one additional row of capacitors (not exposed to light), called the serial register.

Reading the array first requires a shift of pixel content down the columns of the parallel registers by one pixel (Fig. 2*b*), that is, "coupling" the charge, providing the name "charge-coupled device." Electrons originally stored in row 3, for example, shift to the serial register.

A second operation now reads the newly filled serial register by shifting its contents to the right by one pixel (Fig. 2*c*). The electrons in the rightmost pixel shift into a new structure: an output amplifier that ultimately converts the charge to a voltage, and an analog-to-digital converter (ADC) that converts the voltage into the number stored in computer memory. The CCD continues this shift-and-read of the serial register, one pixel at a time (Fig. 2*d* and 2*e*), until the entire serial register has been read. The

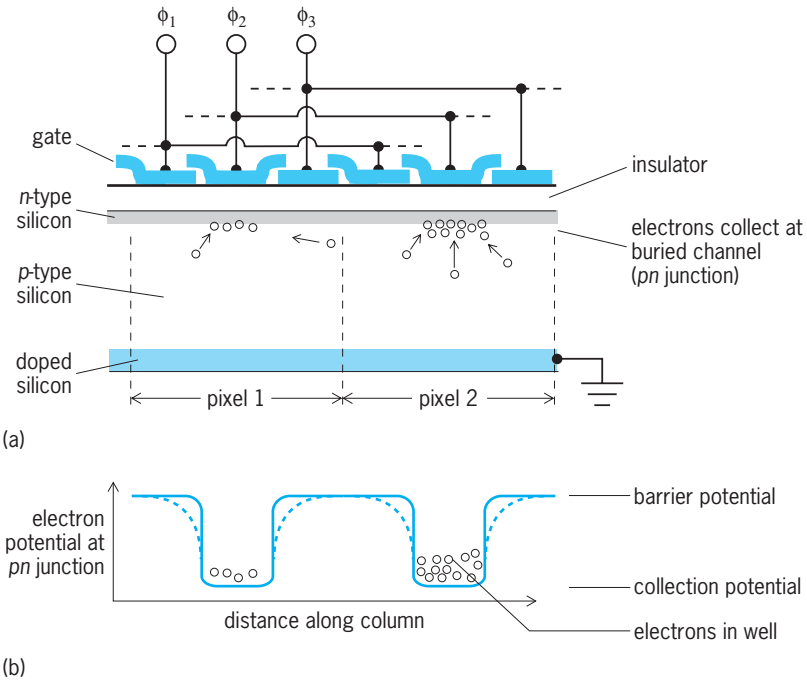

Fig. 3. Gate structure in a three-phase buried-channel CCD. (*a*) Two pixels, in cross section. (*b*) Electron potential *pn* junction. Collection and barrier potentials on the gates isolate the pixels from each other during an exposure. Overlapping gates produce a gradient in the barrier region (broken curve in lower figure) that enhances collection.

device then repeats parallel shifts (Fig. 2*f*) [followed by serial shifts and reads] until the entire array has been read to memory. *See* ANALOG-TO-DIGITAL CONVERTER.

To persuade the electrons stored in one capacitor to move to the neighboring capacitor, the single electrode of a pixel is usually replaced with two or more electrodes whose voltages are manipulated. In a three-phase CCD, for example, each pixel contains three electrodes or gates (**Fig. 3**).

Readout time and read noise. To maximize the DQE, the amplifier and analog-to-digital converter of a scientific CCD introduce very little noise, usually less than a few electrons per pixel. The noise these circuits adds depends on how quickly they do their job—the faster, the noisier. A large, slow-scan CCD for astronomy may require many tens of seconds to read a single image. This is an important difference between astronomical CCDs and commercial devices like digital cameras and webcams.

Lengthy readout times increase response time and waste telescope time. To speed read time, some chips use multiple amplifiers on a single array. Each of four amplifiers, for example, can simultaneously read one-quarter of an array (**Fig. 4*a***).

An additional strategy is to build a mosaic of several very closely spaced but electrically independent CCDs, an approach that also circumvents the practical difficulties in manufacturing monolithic CCDs larger than 10–20 megapixels. A relatively simple combination of offset multiple exposures will fill in those parts of an image masked by the gaps between mosaicked CCDs. Kitt Peak National Observatory, for example, placed the eight-element Mosaic Imager in service in 1998 (Fig. 4*b*). As of 2005, several large observatories support mosaic arrays of up to 40 CCDs with total sizes of 200–400 megapixels.

Dark current, cooling, and vacuum enclosures. Thermal agitation can generate electron-hole pairs in the depletion zone of a MOS capacitor. The resulting steady flow of thermal electrons into the CCD potential wells is called dark current. At room temperature, the dark current can easily fill the potential wells and saturate a scientific CCD in a few seconds. Dark current adds some number of electrons to whatever number of photoelectrons are produced in a pixel and also introduces an associated uncertainty (noise). This uncertainty can never be removed, so the signal-to-noise ratio is always degraded by dark current.

Astronomers almost always operate CCD cameras at very low temperatures to reduce dark current. Cameras on spacecraft, if properly shaded from sunlight, can operate at low temperatures without other provisions for cooling. On the ground, a common cooling method places the CCD in thermal contact with a bath of liquid nitrogen (LN_2), a chemically inert substance that boils at 77 K (−196°C or −321°F). For liquid nitrogen–cooled CCDs, a typical operating temperature is around −100°C (−148°F). This technique requires sealing the detector in a small vacuum chamber as well as a carefully designed system for handling liquid nitrogen.

At the cost of somewhat higher dark current, relatively inexpensive and easy-to-use solid-state thermoelectric (Peltier junction) coolers can maintain a detector in the −20 to −50°C (−4 to −58°F) range, where the dark current of specially engineered CCDs (the MPP-CCD) is acceptably low for many astronomical applications. *See* PELTIER EFFECT; THERMOELECTRICITY.

Some non-CCD arrays used in the infrared require cooling with expensive liquid helium, which boils at 4.2 K (−269°C or −452°F). A relatively complicated apparatus, the closed-cycle helium refrigerator, can cool detectors to the 10 to 60 K (−263 to −213°C or −442 to −352°F) range without consuming large amounts of liquid helium.

Charge-transfer efficiency. In a CCD readout, the fraction of electrons in a pixel that are successfully

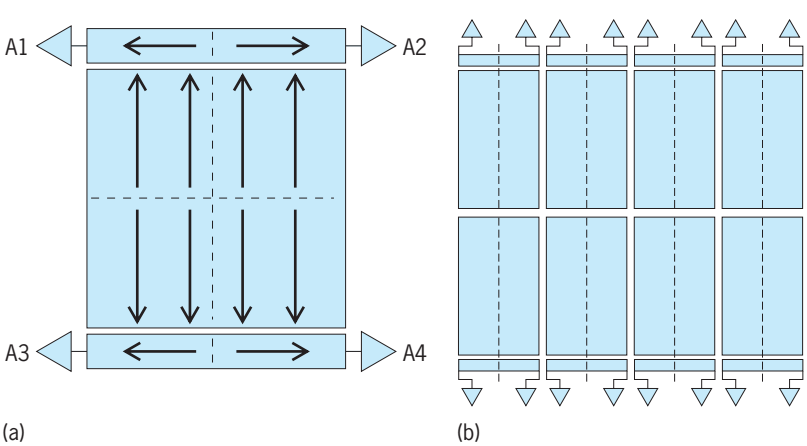

Fig. 4. Large CCD strategies. (*a*) Large monolithic detector with multiple serial registers and amplifiers (four, in this case). Read time is reduced by a factor equal to the number of amplifiers, and total charge transfer efficiency (CTE) is improved. (*b*) Mosaic of eight arrays butted to form a single large-area detector.

moved to the next pixel is called the charge-transfer efficiency. One can avoid errors in the image only if the charge-transfer efficiency is very close to one. Multi-megapixel arrays require charge-transfer efficiency values approaching 0.999999 (known as "six nines"). To avoid decreases in charge-transfer efficiency, astronomical CCDs cannot operate at temperatures below about −100°C (−148°F) and must be designed so that all electrons collect in a region safely removed from the semiconductor–insulator interface. Manufacturers can produce this so-called buried channel by adding a thin layer of *n*-type semiconductor between the *p*-type material and the insulator (Fig. 3).

Surface issues. If a CCD is illuminated from the front side, the insulator and (especially) the gates will absorb and reflect photons at all wavelengths, but particularly in the blue and ultraviolet.

Backthinning. QE improves at all wavelengths, but especially at short wavelengths, if the semiconductor layer is made very thin, so that the buried channel can be illuminated from the back. The delicate technology for fabricating the very fragile backthinned CCD is costly to set up and difficult to perfect, so a thin, back-illuminated CCD is substantially more expensive than the thick, front-illuminated version of the same device.

Frontside options. Several strategies can improve the short-wavelength QE of a front-side illuminated CCD. Gate materials like indium tin oxide are more transparent than polysilicon, but harder to fabricate. Some designs leave gaps between the electrodes, and direct incoming light through them with microlenses. Another strategy coats the front surface with a thin layer of phosphor. A typical phosphor might glow with green light when excited by incoming ultraviolet light, thus producing a detector response to the ultraviolet. Except for cost, all these strategies have so far proven inferior to backthinning.

Antireflection coatings. An appropriate antireflection coating can minimize the reflection losses from any surface. The coating is most effective for light whose wavelength is four times its thickness, so a CCD designer must choose the coating with the intended use of the detector in mind. Usually CCDs are coated either to enhance the short-wavelength response or to minimize reflectivity near the device's QE maximum.

Processing Array Data

No matter what source is being investigated, certain calibration steps, called preprocessing, are necessary to remove the instrumental signature from array data. Further processing is then carried out to improve image quality.

Preprocessing. The calibration includes subtractive corrections and steps to improve linearity and flat-field response.

Subtractive corrections. If a detector is not exposed to a signal from the telescope but simply sits in the dark for a time, it can nevertheless produce a positive signal. This results from two effects: the signal bias, or zero level, which is always present to ensure amplifier linearity, and the accumulated thermal dark current, which grows with exposure time. The calibration step here is to take several exposures with the shutter closed, compute the bias and dark rate, and subtract the appropriate levels from every image produced by the array.

Linearity. A nonlinear detector requires a special calibration step to convert its response into one directly proportional to the input illumination.

Flat-field response. Identical light signals generally do not produce identical responses in every pixel of a detector array. This defect can arise because of structural quantum-efficiency differences intrinsic to the array or because of imperfections in the optical system. The calibration procedure is to observe images of a perfectly uniform source (for example, a projector screen or the twilight sky) and then divide every data frame by this flat field image.

Processing. After preprocessing, subsequent treatment of images depends on the information being sought. Background removal is often important. Removal of a smooth background (due to sky glow or light pollution) is usually a simple subtractive operation. Cosmic rays strike pixels randomly and generate electrons in them, and these are most easily removed if multiple exposures of the same scene can be combined (astronomical light sources repeat positions, but cosmic rays do not). Combining multiple exposures can also fill in detector defects like insensitive pixels or the gaps in a mosaic. Since the SNR generally increases with the number of electrons or photons counted, combining multiple exposures increases data quality.

Frederick R. Chromey

Bibliography. R. Berry and J. Burnell, *The Handbook of Astronomical Image Processing*, Willmann-Bell, Richmond, 2000; S. B. Howell, *Handbook of CCD Astronomy*, Cambridge University Press, 2000; C. Kitchin, *Telescopes and Techniques*, 2d ed., Springer, London, 2003; I. McLean, *Electronic Imaging in Astronomy*, Wiley, 1997.

Astronomical photography

The application of the photographic process to astronomy, including monochrome photography and color photography.

Monochrome Photography

Monochrome photography was one of the premiere tools of astronomical research during most of the twentieth century. It offered two major advantages: the integration of signal (through time exposure) allowed the accumulation of photons from very faint objects which could not otherwise be seen; and the storage of information in an efficient and permanent form allowed protracted, in-depth study of astronomical objects away from the telescope. Monochrome astronomical photography became a standard technique that was applied to direct imaging, spectroscopy, photometry, polarimetry, and astrometry. *See* ASTROMETRY; ASTRONOMICAL SPECTROSCOPY; PHOTOMETRY; POLARIMETRY.

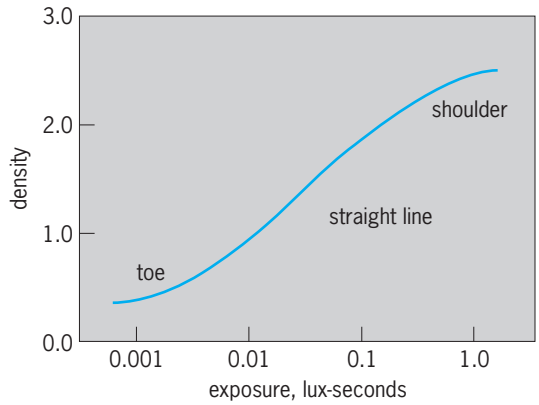

Fig. 1. Typical characteristic curve describing the relationship between density and exposure for a hypothetical monochrome emulsion.

Nonetheless, because astronomers were continually searching for increased sensitivity and wavelength coverage, the aggressive attempts to improve photographic emulsions had, by the 1970s, begun to reach the limits of what was reasonably attainable. At the same time a new type of electronic imager, the charge-coupled device (CCD), developed for military applications, began to become available to the scientific community. Space-based imaging applications of the charge-coupled device, particularly the development of the TI 800 × 800-px chip for use in the *Galileo* spacecraft, stimulated expanded use of this mode of imaging at ground-based observatories. By the end of the twentieth century, charge-coupled devices had become so pervasive in astronomical observatories that it was clear that monochrome astronomical photography was, in many respects, a technique of the past. Rarely used now by professional astronomers, the techniques of high-quality monochrome astrophotography are primarily preserved by amateur astronomers. Some advances are still being achieved, primarily through computer-processing of digitized photographic images. *See* CHARGE-COUPLED DEVICES.

Performance. Although the acquisition of astronomical photographs has declined dramatically, large photographic archives will continue to serve as important resources. An understanding of the properties of photographic emulsions will remain necessary to fully exploit these archives.

Characteristic curve. One important property of a photographic emulsion is represented by the characteristic curve, also known as the Hurter-Driffield (H-D) curve (**Fig. 1**). This curve is variously plotted as the transmittance, opacity, or density of the processed emulsion versus the exposure, defined to be the product of light intensity and exposure time. Unlike the now-favored electronic detectors, which exhibit a linear analog to the characteristic curve, the H-D curve of an emulsion has a distinct nonlinearity, which complicates the use of photographic data. One result is that on the toe and shoulder of the curve, that is, for both very faint and very bright objects, the monochrome photograph provides little photometric distinction.

Nonetheless, the characteristic curve is very useful in selecting the most suitable emulsion for a particular application, and in optimizing the exposure time and processing necessary for making the most useful image of the specific objects of interest. The straight-line portion of the characteristic curve defines the range of exposures that may most reliably be used for determination of brightnesses of stars. The slope of the curve gives an index defining the level of contrast of the photograph.

Wavelength sensitivity. Photographic emulsions are also characterized by their sensitivity as a function of wavelength (**Fig. 2**). O emulsions are blue sensitive, J are useful into the green, E and F have very broad responses into the red, and N is usable into the near-infrared. The most important professionally used astronomical emulsions, which are still available, are the IIIa-J, IIIa-F, I-N and IV-N (progressively more sensitive to the red and near-infrared portions of the electromagnetic spectrum). These plates are produced in large sizes and were most recently used professionally in all-sky-survey imaging. The general wavelength sensitivity of a photographic emulsion can be matched to the specific scientific problem at hand; fine tuning of the spectral window that is imaged is accomplished by use of optical filters in conjunction with the emulsion. The characteristic curves of these emulsions vary widely. For example, near-infrared emulsions have much lower sensitivity (and hence longer required exposure) than more blue-sensitive emulsions.

Image structure. The image-structure characteristics of emulsions are critical to the application of the photograph. The most important of these characteristics is the granularity, or graininess, of the photograph. The size of the clumps of silver halide grains which form the image in the emulsion have a strong effect on the resolving power of the photograph, that is, the fineness of the detail which can be measured. Granularity is indicated by the prefix in the emulsion designator, types 103, II, III, IV, V, and 649 being progressively finer grained. In general, lower granularity is obtained at the expense of longer exposure times to achieve similar densities in the processed photographs.

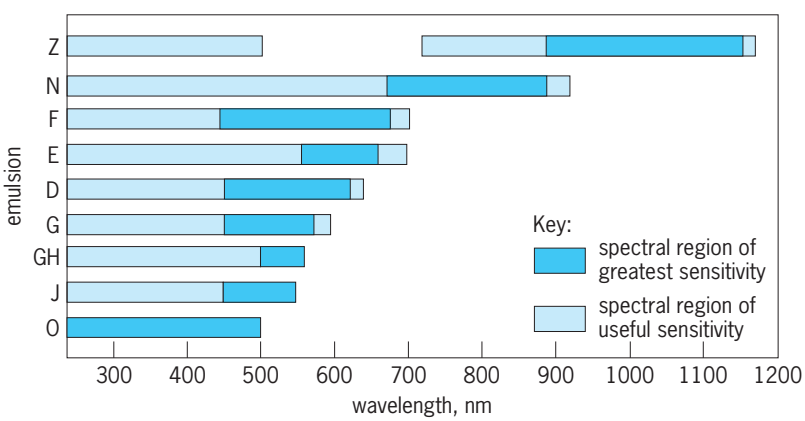

Fig. 2. Spectral sensitization for a variety of monochrome astronomical emulsions. (*After Eastman Kodak Co., Scientific Imaging with Kodak Films and Plates, Publ. P-315, 1987*)

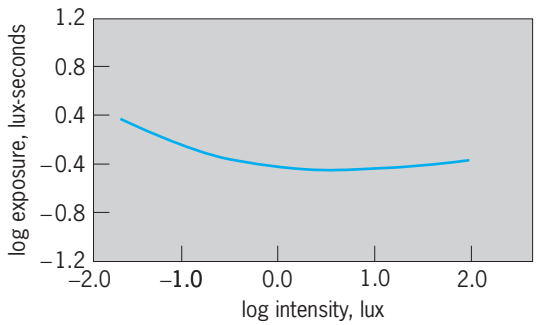

Fig. 3. Result of reciprocity failure for a hypothetical monochrome emulsion. The graph shows the amount of exposure required to produce a constant density at different levels of illumination for a single photographic emulsion type. Loss of sensitivity at very high or very low levels of illumination is reflected in the greater exposures required at these extreme levels.

Reciprocity failure. In the low-light-level environment of astronomical imaging, exposure times can be quite long, from minutes to hours. In theory, the law of reciprocity for chemical reactions implies that the total exposure is equal to the intensity times the exposure time. In practice, for most photographic emulsions, this law fails for long exposure times (**Fig. 3**). Reciprocity failure, whereby increased exposure time does not produce a commensurate increase in total exposure, has been partially mitigated by special processing of certain astronomical emulsions. These emulsions are distinguished by the letter a in the type designator, for example, IIa-O. Two primary factors in the reciprocity characteristics of photographic emulsions are the temperature of the emulsion during exposure, and the gaseous environment in the camera as the emulsion is exposed.

Hypersensitization. Several techniques have been used to improve the performance of astronomical emulsions. Generally described as hypersensitization, they include cooling the emulsion during exposure (to mitigate reciprocity failure), baking the plate prior to exposure, removing oxygen and water from the emulsion prior to use, treating the plates with hydrogen gas prior to exposure, and preflashing the emulsion with a low level of uniform light. In addition, there are processing steps which can improve the performance of the emulsion, such as the extension of development times beyond what is normally specified.

Applications. The great majority of monochrome astrophotography is now done by amateur astronomers, though much of it is quite professional in nature and can provide interesting images with resolutions and image scales not presently attainable with charge-coupled-device images. Targets for this type of imaging are frequently large, extended objects, such as nearby galaxies, galactic star fields, and various types of extended nebulae, as well as comets with extended tails. Large, systematic photographic programs have disappeared from astronomy, and the remaining photography is done primarily on an ad-hoc basis.

Photographic emulsions are now most commonly used by amateur astronomers and, for this application, are found primarily on flexible film substrates, and only rarely on glass plates. The emulsion types which are most popular for astronomical imaging are Kodak Technical Pan Film (Type 2415), and Kodak T-MAX 400 Professional Film. These emulsions are widely used in a standard 35-mm format, although they are also available in larger-size sheets on very stable Estar substrates. Even among amateur astronomers, the process of gas-hypering film to increase photosensitivity is fairly common.

Computers are widely used to provide enhancements to monochrome photographic images. Films are typically scanned at high resolution to provide digital image files for subsequent processing. Software tools of many types are used, ranging from commercial applications packages designed for photoprocessing for desktop publishing to special-purpose astronomical packages. Multiple photographic images of the same field can be added in order to increase signal-to-noise ratio and dynamic range, and images are frequently contrast-enhanced, or digitally filtered to improve sharpness. Very precisely controlled color images are often produced from monochrome photography using multiple images of a given field obtained through glass band-pass filters, and digital image processing techniques are used to produce very large, high-resolution mosaics of astronomical images by precisely abutting adjacent images.

Both existing and new photographic archives will be given additional power as techniques of digitization of large databases improve, and as digital image processing technology continues to evolve. Many important photographic archives have already been digitized and made available in electronic formats, and it is expected that many more will be digitized in coming years. *See* IMAGE PROCESSING. Eric R. Craine

Color Photography

The language of astronomy abounds with references to color, and the concept of color is implicit in many of the measurements that astronomers make. Color index, for example, is a quantity related to the temperature of a star, while the redshift of a galaxy is used to indicate its recessional velocity. More directly, stars may be described as red giants or white dwarfs or even blue stragglers. *See* BLUE STRAGGLER STAR; COLOR INDEX; REDSHIFT; STAR.

These names reflect the underlying importance of color in astrophysics and cosmology, and though the colors involved are subtle and difficult to distinguish by the eye in its dark-adapted state, special photographic techniques can be used to display them. A realistic representation of the true colors of celestial bodies can reveal new relationships in familiar objects and add an important third dimension to the morphology and brightness information of the more usual monochrome representations.

Color films. As discussed above, special photographic materials are necessary to accommodate the unusual requirements of photography in astronomy. Not only is the amount of incoming radiation to be

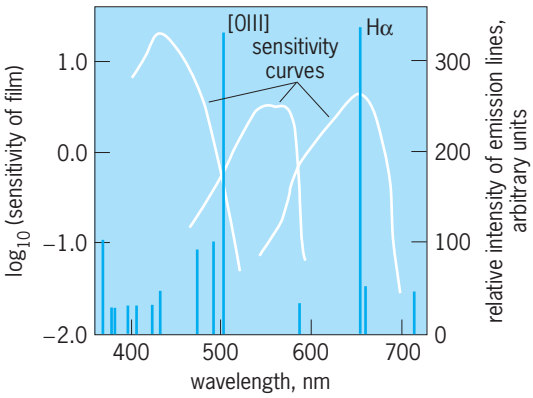

Fig. 4. Comparison of the spectral sensitivity curves of a color film with the emission-line spectrum of the Orion Nebula.

detected extremely small, but it is accompanied by unwanted light from the night sky (the airglow). The materials must therefore combine extreme sensitivity at long exposures with high contrast. The ability to detect faint objects is ultimately more dependent on the contrast and resolution of the photographic material than on the light grasp of the telescope or available observing time. On the other hand, color films are designed for general use at levels of illumination where high contrast and low-light-level efficiency are unimportant. In addition, these films are intended primarily to reproduce the broadband colors of everyday life, and for this the rather uneven spectral response of their individual layers is unimportant (**Fig. 4**). Unfortunately, gaseous nebulae emit most of their visible radiation in the form of monochromatic emission lines from the ionized elements present. In the visible region, the strong green forbidden lines of doubly ionized oxygen [O III] near 500 nanometers and the rich red of a hydrogen recombination line Hα at 656 nm often predominate. The green line falls at a minimum in the blue-green sensitivity of a typical color film and is not well recorded; however, the red line coincides with the peak sensitivity of the red-sensitive layer. Thus, color films always show gaseous nebulae as red, largely irrespective of the contribution from the green oxygen line, whereas yellow (red + green) would be a more realistic representation in many cases. *See* NEBULA.

A further problem is the effect of long exposures on the relative sensitivity of the three layers, which are differently affected by low-intensity reciprocity failure. Changes in both sensitivity and contrast of the layers are found, and exposures which are long enough to be astronomically useful often produce severe color-balance distortion. However, in the absence of materials prepared specifically for color astrophotography, attempts have been made to minimize the problems of ordinary color films and exploit their advantages of low cost, ease of use, and ready availability. Both reversal and negative films have been used. Most workers appear to favor the higher contrast and brighter colors of reversal films, but the greater latitude and color-correction possibilities of negative materials are considerable advantages, especially since exposure times are difficult to assess. The long-exposure properties of both types of film can be improved by user-applied preexposure treatments.

Cooled cameras. As discussed above, low-intensity reciprocity failure of both color and monochrome films is reduced if the long exposure is made at a low temperature. Most experiments have been made with cameras designed for fairly small formats and cooled to about −103°F (−75°C) with solid carbon dioxide. Care must be taken to avoid the condensation of water vapor on the film during exposure and to ensure that the cooling is uniform across the frame. This becomes increasingly difficult with the large formats used in professional astronomy, and is impossible in some types of telescopes. Exposure at low temperature reduces the effects of low-intensity reciprocity failure on both the speed and color balance of all types of films.

Hypersensitization of color film. Some of the techniques which are used for spectroscopic plates may also be applied to color films. Baking both in nitrogen and in forming gas, a 2–4% hydrogen-in-nitrogen mixture, is useful. Forming gas is available in special kits specifically for hypersensitizing small quantities of films for astrophotography. Films are baked for several hours in a flow of the gas at 150°F (65°C) just prior to exposure and then (preferably) exposed in a nitrogen atmosphere. Substantial long-exposure speed gains are reported. However, the process affects each of the three sensitive layers differently, and some shift in color balance may be experienced.

These user-applied processes reduce some of the disadvantages of color films for astrophotography, and push development may be used in addition to the above, to increase both speed and contrast. However, the basic problem of uneven spectral response remains. As a result, color films can reproduce only realistic colors of the brighter, continuous-spectrum objects, such as planets, stars, and galaxies. Faint objects and emission nebulae are not well recorded.

Indirect color photography. An alternative approach is to use the oldest system of color photography, the three-color separation technique. This process derives directly from James Clerk Maxwell's 1861 demonstration of additive color imaging. In this system, three exposures are made with combinations of photographic emulsions and filters chosen to record the red, green, and blue parts of the spectrum on separate plates or film. In professional astronomy, these passbands are achieved with Eastman Kodak's fine-grain IIIa-J and IIIa-F (blue-green and red-sensitive) emulsions with glass filters to record the appropriate broadband colors. Kodak's Technical Pan film also works well after hypersensitization, but for the faintest objects under truly dark skies exposure times can be 90 min per color, even with an *f*/2.5 telescope such as the United Kingdom Schmidt. *See* SCHMIDT CAMERA.

Filter-emulsion passbands are chosen to ensure adequate overlap between adjacent colors so that hues intermediate between red, green, and blue are well

recorded. With care in selection of these parameters, coverage of the visible region is much more uniform than is possible with conventional color film. This overlap ensures that narrow emission lines across the visible spectrum are well recorded, which is not the case with color film (Fig. 4).

The same principles are used with digital detectors such as charge-coupled devices, which are much more sensitive than hypersensitized silver-based photographic materials. As yet, they lack the essentially unlimited area and small pixels that give conventional materials their "photographic" characteristics that translate into a distinctive esthetic quality. Charge-coupled-device images are usually combined into three-color images with a computer, using commercially available software, and similar methods can be used to combine digitized versions of photographic red, green, and blue exposures.

Recombining color separations. Two methods are possible to recover the color information in three-color separations. The subtractive process involves imagewise combinations of yellow, magenta, and cyan dyes or pigments and is now rarely practiced outside professional printing applications, though it was once popular as the dye transfer process used for the finest color reproduction. It was also the basis of the Technicolor process. Dyes or pigments used subtractively also form the the color systems used in all commercial films and printing processes. *See* PRINTING.

Much more flexible in the astronomical application is the additive process, which allows several levels of image manipulation before the monochromes are combined. Additive color photography involves mixing colored light, rather than colored compounds, and its most common manifestation is the color television or computer screen image, which a magnifier shows to be made up of blue, green, and red dots or strips. When used photographically, monochrome positive copies are made by contact copying the three original separation negatives onto a suitable film material. *See* ELECTRONIC DISPLAY; PICTURE TUBE.

At this stage, a wide range of image-manipulation techniques can be applied to enhance small or faint features and to adjust the contrast of the original images. The positives are subsequently enlarged sequentially in register onto a positive-working color material (such as Cibachrome or Ektachrome) or onto a negative stock for subsequent printing by normal substractive methods. The positives are enlarged through the equivalent of the taking filters, the red positive through a red filter, and the green through a green filter, and so forth. This process is easier to use and more flexible than any of the subtractive systems and can yield color prints of both scientific and esthetic merit.

Color balance. The eye is a poor discriminator of color at low light levels, and it is not possible to check the color balance of astronomical photographs against the original scene or against color memory to verify color fidelity. Color pictures produced on the 154-in. (4-m) Anglo-Australian telescope are balanced by comparison with a neutral gray step-wedge image projected onto the plates during the telescope exposure. The light source of the projector is filtered to a color temperature of 5500 K (9400°F; equivalent to sunlight). Thus, sunlike stars appear white on pictures from this observatory, and stars hotter or cooler than the Sun appear blue and yellow respectively. Likewise, the balance of color in an emission-line nebula is adjusted with reference to the same gray step wedge on the assumption that the spectrum has been recorded equivalently in all three passbands. *See* PHOTOGRAPHY. David F. Malin

Bibliography. Eastman Kodak Co., *Scientific Imaging with Kodak Films and Plates*, Publ. P-315, 1987; T. Hallas and D. Hallas, Astrophotography with a twist, *Sky Telesc.*, 96(5):130–137, November 1998; I. McLean, *Electronic Imaging in Astronomy*, Wiley, 1997; D. F. Malin, An Integrated system of astronomical photography, Parts II and III, in Eastman Kodak, *TechBits Mag.*, Issue 1, pp. 1–10, and Issue 3, pp. 3–11, 1990; D. F. Malin, *A View of the Universe*, 1993; D. F. Malin and P. G. Murdin, *The Colours of the Stars*, 1984; G. Walker, *Astronomical Observations*, Cambridge University Press, 1987.

Astronomical spectroscopy

The use of spectroscopy (the analysis of light as a function of wavelength) as a tool for obtaining observational data on the chemical compositions, physical conditions, and radial velocities of astronomical objects. Astronomical applications of optical spectroscopy from ground-based observatories cover the electromagnetic spectrum from the near-ultraviolet (wavelengths around 0.3 micrometer) through the visible (0.4–0.7 μm) and into the near-infrared (2 μm). Space-based observatories, such as the Hubble Space Telescope, the *Far Ultraviolet Spectroscopic Explorer*, and the *Spitzer Space Telescope*, extend spectroscopic observations from the far-ultraviolet (0.1 μm) to the far-infrared (200 μm). Work at shorter wavelengths (x-ray and gamma-ray spectroscopy) and longer wavelengths (submillimeter and radio wavelengths) requires techniques other than those discussed here. *See* GAMMA-RAY ASTRONOMY; HUBBLE SPACE TELESCOPE; INFRARED ASTRONOMY; RADIO ASTRONOMY; SPITZER SPACE TELESCOPE; SUBMILLIMETER ASTRONOMY; ULTRAVIOLET ASTRONOMY; X-RAY ASTRONOMY.

Usually a spectrograph is fitted to a reflecting telescope, which serves as a light collector. The image of the celestial body being studied is focused on the spectrograph slit, which limits the region under study (thus improving the spectral resolution) and reducing the contribution by the night sky. The diverging light beam then passes from the slit to a collimator (either a lens or mirror). This produces parallel light, which is then dispersed by a diffraction grating or prism. The dispersed light enters a camera, which focuses the spectrum onto a detector, either a charge-coupled device (CCD) in the case of an optical spectrograph, or an electronic array

sensitive to infrared light. *See* CHARGE-COUPLED DEVICES; DIFFRACTION GRATING; SPECTROGRAPH; SPECTROSCOPY; TELESCOPE.

Resolving power. The resolving power of a spectrograph is usually denoted by R, and is equal to the wavelength of light, λ, divided by the spectral resolution, $\Delta\lambda$; that is $R = \lambda/\Delta\lambda$. The size of the spectral resolution element is set by the width of the slit, the grating, the focal length of the camera, and the detector resolution. Typically $R \approx 200$ for low-dispersion survey work, $R \approx 5000$ for "classical" spectroscopy, and $R \approx 50{,}000$–$100{,}000$ for the highest-resolution work. *See* RESOLVING POWER (OPTICS).

Multiobject spectroscopy. It is often desirable to obtain spectroscopy of many of the objects within a telescope's field of view in a single exposure. A variety of methods are available to accomplish such surveys.

Slitless spectroscopy. It is possible to take spectra of all of the brighter objects within the field of view by not using a spectrograph at all, but by combining a low-dispersing element directly with the telescope. For instance, an objective prism may be placed in front of the telescope, which is often a Schmidt camera. Slitless spectroscopy has been used for large stellar surveys such as the Henry Draper Catalogue and, more recently, the Case Michigan Survey. Alternatively, a grism or grens (a hybrid between a prism and a grating) may be placed in the converging beam near the focal plane. Such spectroscopy is usually of low resolving power ($R \approx 100$–500) and is limited to the objects brighter than the sky background, as slitless spectroscopy has no way of attenuating the amount of sky seen by the detector. However, the use of large-format arrays of charge-coupled devices which can minimize the effects of the sky background opens new possibilities in searches for faint objects. *See* ASTRONOMICAL CATALOGS; SCHMIDT CAMERA.

Slitlet masks. In this technique, a picture is usually taken of a region containing several astronomical objects of interest; the exact locations of these objects are determined, and small slits (slitlets) are then milled in the corresponding locations in a metal plate. This plate is substituted for the slit in a conventional spectrograph. Care must be taken that the spectra do not overlap, which greatly limits the number of objects that can be observed at any one time to 10–100. Such systems have the advantage of good subtraction of the night sky, and very high efficiency. This technique has enjoyed a revival, with multiobject slitlet masks in use on the world's largest telescopes [for example, the DEIMOS spectrograph on the Keck 10-m (386-in.) telescope, GMOS spectrographs on the Gemini 8-m (315-in.) telescopes, and the IMACS spectrograph on the 6.5-m (256-in.) Magellan (clay) telescope].

Fiber-fed spectroscopy. Rather than milling slitlets in a plate, holes may be drilled, which are then plugged with optical fibers. (Such an arrangement is often referred to as a plugboard.) The light is then transported via the fibers to a spectrograph mounted on an optical bench in a laboratorylike environment adjacent to the telescope. Alternatively, robotics may be used to position fibers in the focal plane; the fibers are then anchored to a metal plate via magnets. At the spectrograph, the fibers are arrayed in a line and act as the spectrograph slit. Although significant light losses may occur in this scheme (typically factors of 2–3), hundreds of objects can be observed simultaneously, leading to very effective use of the telescope. The faintest object that can be usefully observed is dependent upon how well the relative transmissions of the various fibers can be calibrated, since some fibers will be sampling the night-sky spectra alone, while others will contain the spectra of the night sky plus a faint astronomical object of interest. *See* FIBER-OPTICS IMAGING; OPTICAL FIBERS.

Echelle spectroscopy. Normal spectrographs employ diffraction gratings that are intended to be used in low orders ($n = 1$, 2, or 3), with colored glass filters used to prevent overlap of adjacent orders. Echelle spectrographs differ from conventional systems in that they employ gratings intended to be used in very high orders ($n > 10$), resulting in very high resolving power. Normally these orders would fall on top of one another, rendering the data useless. An echelle uses a second dispersal element, usually another grating but sometimes a prism, at right angles to the first, in order to separate the successive spectral strips from each other. A large range of wavelengths can be obtained in the format of nearly parallel segments, well suited for charge-coupled devices.

Coudé spectroscopy. In the past, high resolving power was achieved primarily by using a large, stationary spectrograph located off the polar axis of an equatorial telescope. Light would be brought through the hollow axis of the telescope by a series of flat mirrors (the coudé train). Coudé spectrographs are very stable, as they are located off the telescope, and the optics and gratings can be quite large as they do not need to be attached to the telescope. Echelles and fiber-fed bench-mounted spectrographs have largely replaced the coudé spectrograph; most modern telescopes are now built with altitude-azimuth mounts rather than equatorial mounts, leading to the further decline of coudé systems.

Integral field spectroscopy. In this application, a close-knit bundle of optical fibers is placed in the focal plane and is used to observe an extended astronomical object, such as a gaseous nebula or a galaxy. The light is transmitted via the fibers to a bench-mounted spectrograph. Although the fibers are in a linear array at the spectrograph, their locations in the focal plane are known, and sophisticated data reduction techniques allow the astronomer to reconstruct a spectral "image" of the object.

Infrared spectroscopy. Infrared spectroscopy is sensitive to light at wavelengths longer than about 1 μm through about 200 μm. At ground-based telescopes, observations are possible in the near-infrared (1–5 μm) and in the mid-infrared (5–30 μm). Far-infrared (30–200 μm) spectroscopy is limited to space-based observatories (such as the *Infrared*

Astronomical Satellite, Infrared Space Observatory, and *Spitzer Space Telescope*) because these wavelengths are absorbed by the Earth's atmosphere.

Infrared spectroscopy is a necessary tool for the study of objects embedded in dusty regions, for example the nuclei of certain galaxies, dust-enshrouded massive stars, and regions of star formation in molecular clouds local to the Milky Way Galaxy. Because dust very effectively scatters shorter-wavelength light, there are some objects which are visible only in infrared light. Furthermore, the Earth's atmosphere is more stable in the infrared, as the distorting twinkling of starlight is less at infrared wavelengths, resulting in sharper images of extended objects and binary star systems, useful for angularly resolved spectroscopy of such objects. Most molecules produce rich absorption spectra in the infrared. Cool objects like very young stellar embryos recently formed in molecular cloud cores, very low mass stars, brown dwarfs, and giant planets not only are brighter in infrared light but also manifest more molecular lines than atomic lines and thus are more suitably studied in the infrared regime. *See* MOLECULAR CLOUD; TWINKLING STARS.

Infrared spectroscopy is a relatively young field; early studies using Fourier transform spectrometers (discussed below) began in the 1970s but were limited to the most prominent targets, principally the Sun and the brightest stars. The development of sensitive infrared array spectrometers in the late 1980s and early 1990s paved the way for a plethora of discoveries, from the ultraluminous infrared galaxies, to the largest known redshifts in the universe, to the existence and nature of the brown dwarfs, failed stars intermediate in mass between stars and giant planets. Current generations of infrared spectrometers in use and under construction incorporate traditional features of visible light spectrometers, such as echelle gratings, integral field units, and multiobject capability. *See* BROWN DWARF; GALAXY, EXTERNAL; INFRARED SPECTROSCOPY; REDSHIFT.

Fourier transform spectroscopy. Fourier transform spectroscopy, used particularly in the near-infrared, employs a concept entirely different from the spectrographs described above. Instead of being dispersed in a spectrograph, the light of a wide band of wavelengths is passed through a Michelson interferometer with variable spacing of its two apertures. The resulting interferogram, which is an electronic record of the interference signal produced by the interferometer as the separation of the apertures is varied, is converted into a record of intensity versus wavelength by a computer, and is of extremely high spectral resolution. *See* INTERFEROMETRY.

Spectrophotometry. Often it is useful to know the actual flux of a source in physical units arriving at the Earth as a function of wavelength. In order to achieve this, care must be taken to account for atmospheric light losses. The data can then be calibrated by observing spectrophotometric standard stars; these are stars whose spectral energy distribution is known via comparison to laboratory sources. Such calibration removes the wavelength dependence in the response of the telescope, spectrograph, and detector, and hence allows a direct comparison of the spectrum with model calculations, providing information complementary to that determined from analysis of spectral lines.

Spectropolarimetry. Measurements of polarization in continuum and line spectra often enable the recognition of nonthermal processes (such as synchrotron radiation) in the source. In many instances, they make it possible to assess the importance of light scattering by electrons or small solid grains. *See* POLARIMETRY.

Precision radial velocities. Measurements of the wavelength displacement, due to the Doppler effect, between the spectral lines in a celestial object and a laboratory source give the line-of-sight velocity of the source with respect to the observer. With conventional comparison sources (discharge tubes) and standard spectrographs, the precision of radial velocities is generally of order 1–3 km/s (0.6–1.8 mi/s). However, improvements have pushed this to 5 m/s^{-1} (20 ft/s) or better by using fiber-fed bench-mounted spectrographs, or by using absorption cells of gas (such as iodine) through which the starlight passes to provide an excellent internal reference system. The use of precision radial velocities, combined with dedicated observational programs, has resulted in the detection of planets around nearby stars, evidenced by small, periodic radial velocity variations in their parent stars. *See* DOPPLER EFFECT; PLANET.

Applications. The application of astronomical spectroscopy extends from solar system objects (the Sun, planets, and comets) to Milky Way objects (stars, including binary stars, ordinary novae, and cataclysmic variables; and gaseous nebulae, such as supernova remnants, H II regions, and planetary nebulae) and to distant galaxies and quasars.

Emission-line spectroscopy. Rarefied gas seen against a cooler background will produce an emission-line spectrum. In gaseous nebulae, such as H II regions and planetary nebulae, gas is ionized by the extreme ultraviolet emission of hot stars. As the electrons recombine with atomic nuclei, emission lines are

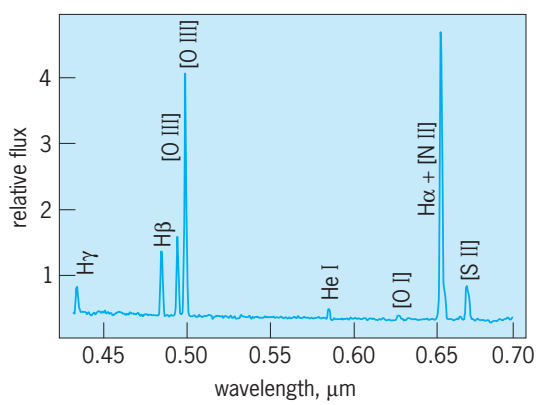

Fig. 1. Spectrum of the largest H II region in the irregular galaxy NGC 1156. The continuum comes from the underlying stellar population. Forbidden lines of neutral and doubly ionized oxygen (O I), (O III), singly ionized nitrogen (N II), and singly ionized sulfur (S II) are visible.

produced. Spectroscopy of the gas reveals physical conditions (density, electron temperatures) as well as chemical composition. The strongest optical nebular lines are the hydrogen Balmer lines and forbidden lines of doubly ionized oxygen ("forbidden" in that these lines would not be seen at higher densities). Forbidden lines of other elements are also usually present, in addition to neutral helium lines (**Fig. 1**). *See* NEBULA; PLANETARY NEBULA.

Some stars also show emission lines in their spectra, indicative of hot gas seen in extended envelopes or disks around these stars. Examples include Be stars (B-type stars whose spectra show emission lines) and Wolf-Rayet stars. *See* WOLF-RAYET STAR.

Absorption-line spectroscopy. More commonly, gas is seen projected against a hotter source, as is the case in a stellar atmosphere or in gas found in interstellar space seen against stars or galaxies. In this situation an absorption-line spectrum is produced.

The Sun was the first astronomical object whose spectrum was analyzed; Joseph Fraunhofer discovered hundreds of dark bands superposed on an otherwise continuous range of colors in 1815. Subsequent analysis of the dark-line Fraunhofer spectrum allowed astronomers to deduce the solar chemical composition, the temperature, density values, and their variation with depth in the solar photosphere. Later discovery of the splitting of spectral lines due to the Zeeman effect demonstrated the existence of magnetic fields on the Sun. The attenuated outer envelopes of the Sun, the chromosphere and corona, require special instrumentation, as they are very faint compared to the bright surface of the Sun. *See* CHROMOSPHERE; SOLAR CORONA; SOLAR MAGNETIC FIELD; SUN; ZEEMAN EFFECT.

Differences in the spectral appearance of stars are due primarily to differences in their effective temperature (the temperature at the surface), and only secondarily to surface gravity (pressure). Differences in chemical compositions are usually a tertiary effect, except for stars with significant chemical peculiarities. The spectra of most stars can be classified in a two-parameter spectral class and luminosity system. Strengths and appearance of different spectra lines supply classification criteria. Classification indices are usually temperature-sensitive lines; luminosity criteria are usually density-sensitive lines. Stars of high luminosity have lower atmospheric densities. The spectral sequence is described by letters O, B, A, F, G, K, M, and L, ranging from hot, about 50,000 K (90,000°F), to cool, 2000 K (3500°F), with decimal subdivisions. Luminosity classes range from I (very luminous) to V (main-sequence dwarf stars). The Sun's spectral class is G2 V, which means it is closer to G0 than to K0 and is a dwarf star. *See* HERTZSPRUNG-RUSSELL DIAGRAM; SPECTRAL TYPE; STAR.

The properties of stellar spectra may be illustrated by the spectrum of the star Sk-69° 22 (**Fig. 2***a*), one of the hottest, most luminous, and massive stars known, with a spectral type of O3 I. This star is located in the Large Magellanic Cloud, the nearest galaxy to the Milky Way Galaxy. Absorption lines of hydrogen (H) and singly ionized helium (He II) dominate. The lack of neutral helium (He I), and the presence of four-times ionized nitrogen (N V), require that the star's temperature be at least 50,000 K (90,000°F). The presence of emission lines of doubly and triply ionized nitrogen and triply ionized silicon (N III, N IV, Si IV), and the very strong, broad emission line of He II at 0.4686 μm, indicate that this star is of extremely high luminosity; these emission lines are formed in the expanding outer envelope of the star.

Certain stars, particularly white dwarfs, have intense magnetic fields whose strength can be measured by the splitting of spectral lines similar to that done for the Sun, although without the possibility of spatial resolution over the disk of the star. In fact, understanding of the importance of magnetic fields in astrophysics derives largely from spectroscopic data. *See* STELLAR MAGNETIC FIELD; WHITE DWARF STAR.

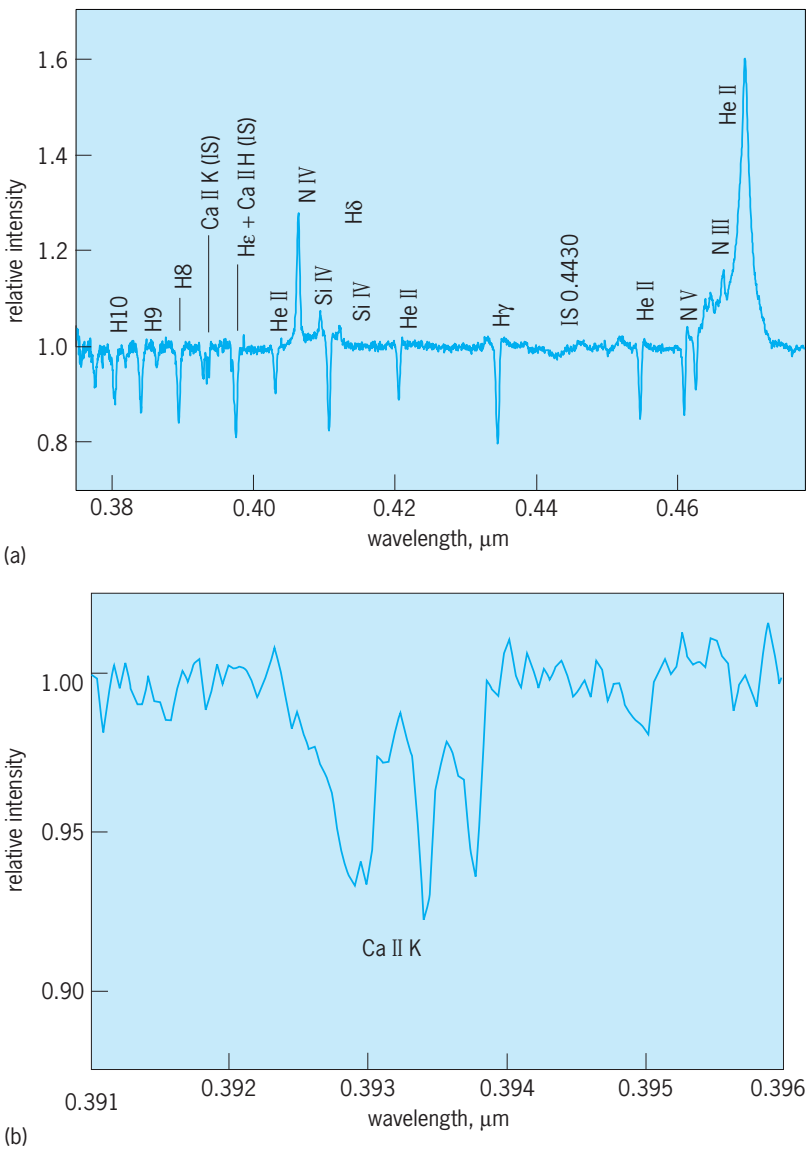

Fig. 2. Spectra of the star Sk-69° 22, one of the hottest, most luminous, and massive stars known, in the Large Magellanic Cloud. (*a*) Part of the optical spectrum, 0.375–0.478 μm. IS indicates absorption lines from interstellar gas. (*b*) Detailed spectrum of the K line of singly ionized calcium (Ca II), 0.391–0.396 μm.

The presence of matter in interstellar space manifests itself in sharp, strong absorption lines. Despite the extremely low density of gas found between stars, the distances to some stars are so great that the effective column depth can be appreciable. Among the strongest interstellar lines in the optical are the H and K resonance lines of singly ionized calcium (Ca II). Broad, diffuse interstellar bands are also seen, the strongest of which in the optical occurs at 0.443 μm. Both of these can be seen in the spectrum of Sk-69° 22, and detailed inspection of its Ca II K line (Fig. 2*b*) shows multiple components, with radial velocities corresponding to gas in both the Milky Way Galaxy and the Large Magellanic Cloud. Extremely luminous, distant quasars make it possible to probe the conditions of the interstellar medium when the universe was quite young, since their light passes through the interstellar media of a number of intervening galaxies. *See* INTERSTELLAR MATTER; QUASAR.

Spectroscopy can be obtained of the resolved, luminous members of the very nearest galaxies; for all but this handful, the spectrum of a galaxy is a composite of its typically 10^{12} stellar members. Such integrated spectroscopy can be analyzed by population synthesis, in which a library of stars is mixed in appropriate proportions until something similar to the observed spectrum is seen. Such studies can place useful limits on the relative numbers of various types of stars, but whether the mix is a unique match to the integrated spectrum is often open to question without additional information.

Reflection spectroscopy. The spectra of solar system objects (other than the Sun) are primarily reflected sunlight. Spectra of planets with atmospheres also show lines and bands of molecules which make it possible to assess atmospheric temperatures and densities in favorable circumstances. For instance, the presence of carbon dioxide (CO_2) in the atmosphere of Venus was established in 1932 from spectroscopic observations. The spectra of comets show a reflection spectrum of the Sun upon which are superposed bright bands of CN, molecular carbon, and other molecules. They also show intense fluorescence effects. *See* COMET; VENUS.

Continuous energy distributions. In addition to the spectral lines, the energy distribution of the continuum—the light between the spectral lines—often provides information about the physical conditions which is highly complementary to the information that can be gleaned from the spectral lines themselves. Measurement of the intrinsic continuum shape requires use of spectrophotometry, and the ability to correct for any reddening due to interstellar material. The continuous energy distribution will sometimes reveal limitations in stellar atmosphere models that would otherwise cause analysis of the spectral lines to yield misleading results. In addition, the continuous energy distribution, particularly over a large wavelength interval, provides information that is necessary for population synthesis to produce unique results.

Philip Massey; Lisa Prato; Lawrence H. Aller

Bibliography. L. H. Aller, *The Atmospheres of the Sun and Stars*, 2d ed., Ronald Press, New York, 1963; P. M. Gray (ed.), *Fiber Optics in Astronomy II*, Astronomical Society of the Pacific, San Francisco, 1993; C. R. Kitchin, *Optical Astronomical Spectroscopy*, Institute of Physics Publishing, Bristol, 1995; A. D. Thackeray, *Astronomical Spectroscopy*, Macmillan, New York, 1961.

Astronomical transit instrument

A telescope adapted to the observation of the passage, or transit, of an astronomical object across the meridian of the observer. The astronomical transit instrument is the classic instrument of positional astronomy, which is the study of the positions and motions of astronomical objects. (The specific categories of astronomy concerned with these investigations are astrometry and celestial mechanics.) The chief variants of the classic design include the vertical circle, the horizontal transit circle, the broken or prism transit, the photographic zenith tube and, most commonly, the meridian or transit circle. *See* ASTROMETRY; CELESTIAL MECHANICS.

The astronomical transit instrument was first developed by the Danish astronomer Ole Roemer in 1689. The modern transit instrument has an objective (lens) with a diameter of 6–10 in. (15–25 cm) and a focal length of 72–90 in. (180–230 cm). The instrument consists of a telescope mounted on a single fixed horizontal axis of rotation. The horizontal axis has a central hollow cube (or sometimes a sphere) and two conical semiaxes ending in cylindrical pivots. The objective and imaging halves of the telescope are also fastened to the cube of the instrument, perpendicular to the horizontal axis. Rotation of the instrument in its bearings, or wyes, permits the optical axis to sweep only in the plane of the meridian. Because it is constrained to a single plane, it can be solidly mounted on massive piers, creating a stability not present in other telescope designs. An accurate clock is the essential ancillary scale by which the transits of the astronomical objects are observed.

The astronomical transit instrument takes advantage of a special case of the astronomical triangle, which is composed of arcs of great circles on the celestial sphere. Its vertices are, respectively, the north celestial pole, the zenith point of the observer, and the celestial object under observation. The angle at the north celestial pole represents the hour angle of the object; hence, when the object is on the meridian, the hour angle is zero, and the triangle degenerates to a single arc, a segment of the meridian. At that instant the local sidereal time equals the right ascension of the celestial object. *See* ASTRONOMICAL COORDINATE SYSTEMS; CELESTIAL SPHERE; SIDEREAL TIME.

Meridian or transit circle. For many years, the major observatories of the world had astronomical transit instruments called meridian or transit circles. These instruments are similar to the transit instrument described above except that they are also capable of

measuring the distance of the object along the meridian to obtain its declination. The addition of a large, accurately calibrated circle attached to the horizontal axis allows large angles to be measured, and serves as the scale for measuring declination. The circle originally was read with a micrometer and later photographed to be measured with a special measuring engine. Most circles are now read with some form of charge-coupled-device (CCD) camera. An accurate (atomic) clock is used as the scale for determining the right ascension of the object. *See* ATOMIC CLOCK.

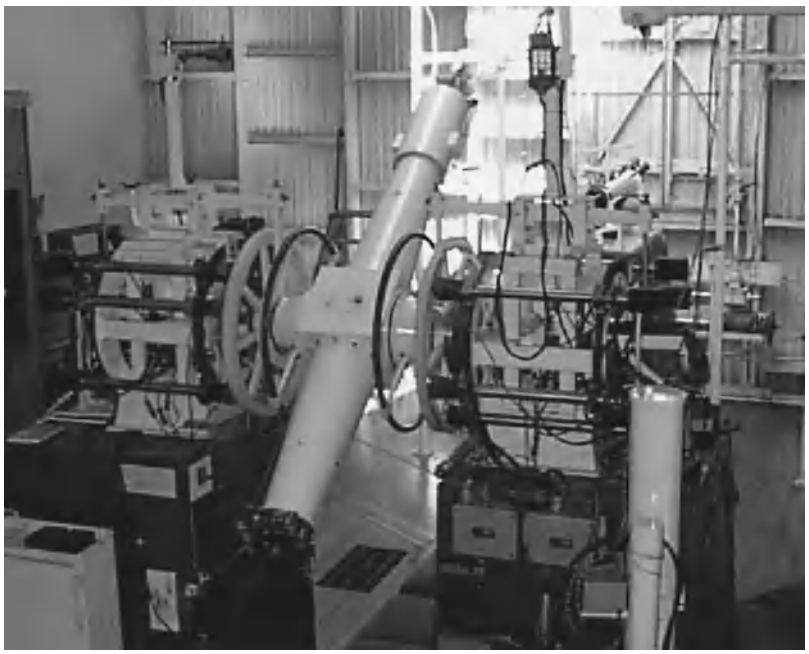

Fig. 1. Six-inch (15-cm) transit circle, U.S. Naval Observatory. (*Official U. S. Naval Observatory photograph*)

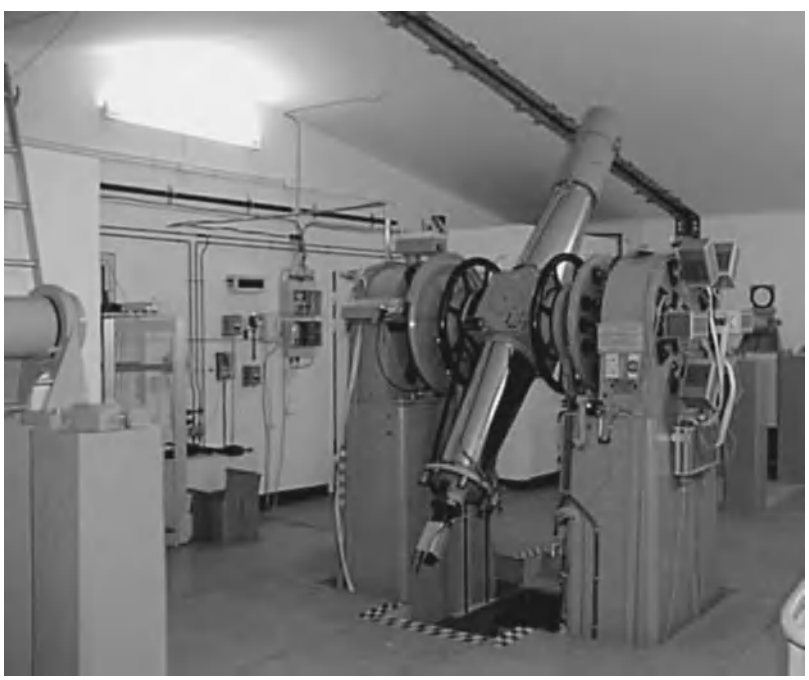

Fig. 2. Carlsberg Meridian Circle, La Palma, Canary Islands. (*University of Copenhagen*)

Traditionally, a very high precision micrometer at the eyepiece end of the instrument contained a movable wire, or pair of wires, and a stationary grid of vertical wires in the x and y directions for use in registering the transit and the declination. The micrometer and eyepiece have now been replaced with electronic detection systems, most commonly charge-coupled devices. While such systems have improved the accuracy and efficiency of the observations, they cannot be used far from the Equator due to the curved paths that the stars follow across the charge-coupled devices, and it is difficult to observe objects that are not point sources, such as the Sun and planets. *See* CHARGE-COUPLED DEVICES.

The 6-in. (15-cm) transit circle of the U.S. Naval Observatory was designed and constructed at the end of the nineteenth century, but it was improved continuously with the latest technological developments (**Fig. 1**). It was the first transit circle to have an electronic circle and the first to have the micrometer data read directly into an electronic computer. An atomic clock provided the computer with the signals to synchronize the right ascension readings from the micrometer. The divided glass circle was scanned with six charge-coupled-device cameras. The probable error of a single set of circle readings was 0.07″. However, the declination of a star may contain uncertainties of 0.25–0.50″ due to errors from other sources, such as atmospheric refraction, mechanical flexure of the instrument, and residual errors in the divided circle. The probable error of a single observation of the right ascension data of an equatorial star was about 0.012 s. When combined with other measurements to form a catalog, the positions have a probable error of around 0.05 second of arc. The introductions to the fundamental catalogs produced with this instrument contain complete descriptions of the equipment, methods, and results.

The 6-in. transit circle was used to observe not only the brighter stars (as faint as ninth magnitude) but also the Sun, Moon, planets, and several of the brighter asteroids. After 100 years of observations, this instrument was taken out of service.

Corrections. It is extremely difficult to adjust the instrument to the point of perfection, where a central charge-coupled-device column will trace the true meridian as the instrument is rotated on its pivots; therefore, corrections must be determined and applied to the observational data. The three principal instrumental errors that require correction are azimuth, collimation, and level. The azimuth correction is the horizontal angle between the axis of rotation and the true east-west direction and is obtained by observing circumpolar stars. The collimation correction is the angle between the line from the optical center of the telescope objective to the center of the charge-coupled device and the plane perpendicular to the horizontal axis of rotation, and is measured using small horizontal telscopes called collimators. The level correction is the angle that the axis of rotation makes with the plane of the horizon. In addition,

a nadir value is obtained that is the zero point of the declination scale. The level and nadir values are obtained by observing a reflection of the focal plane in a basin of mercury.

These instrumental errors vary hourly, daily, and seasonally and, in many cases, are reproducible over these time periods. The main contributors to these errors are the variation in temperature within the pavilion and around the opening in the roof and air strata close to the building caused by wind patterns. Using computers, the values of the temperature, pressure, and humidity can be sampled, and many of these errors can be modeled. The instrument is carefully designed with modern optics and metallurgy to attempt to minimize these effects.

Another error, the clock correction, represents the error between the true sidereal time that the star should transit the local meridian and the time of transit recorded by the local sidereal clock. At one time the clocks had error rates that were significant enough to warrant a correction; however, today they have sufficient accuracy that this correction is considered as a constant for each night's work. This correction is evaluated through observations on bright stars whose positions are already well known.

Applications. The classical astronomical transit instrument had three interrelated uses: (1) from a known position on the Earth, observations of transits of stars led to the determination of their right ascension and declination with respect to the astronomical coordinate system; (2) with a knowledge of the positions of the stars observed and of Greenwich time, the longitude of the observer could be computed from observations of the time of transit of a star; and (3) the determination of corrections to the clock were made by the observation of stars of known position with an instrument situated at a known longitude. As the accuracy of clocks improved (and, specifically, with the development of atomic clocks), the third item fell out of use. To produce a fundamental star catalog, it was necessary to observe solar system objects at the same time as the stars. *See* LATITUDE AND LONGITUDE.

In 1997 the European Space Agency published the results of observations made with the *Hipparcos* satellite. The accuracy of the 100,000 star positions far exceeds anything obtainable with transit circles. In addition, modern astrographs with charge-coupled-device cameras can measure far more star positions than the transit circle and with greater accuracy. The development of mathematical techniques involving overlapping charge-coupled-device frames globally interconnected using powerful computers extends the narrow astrograph field to wide angles. Nevertheless, transit circles still carry on specialized observing programs. The Carlsburg Meridian Telescope at La Palma in the Canary Islands is an example of a modern transit circle used to provide additional accurate star positions to improve the proper motions of previously measured stars and for the measurement of solar system bodies, especially asteroids (**Fig. 2**). It is completely automatic, being operated remotely via the Internet. Use of a charge-coupled device as the detector generally limits these instruments to areas near the Equator. *See* ASTEROID.

A transit circle with collimators and azimuth marks is capable of determining positions of stars over wide angles in a fundamental reference system determined by making concurrent observations of solar system bodies. With the International Celestial Reference Frame (ICRF), based on radio observations of quasars and realized in the optical wavelengths by the *Hipparcos Catalogue*, replacing the fundamental reference stars, as realized by the FK4 and FK5, the transit circle is no longer the determining instrument for the stellar reference frame, and is mostly used now for small, specialized programs. *See* ASTRONOMICAL CATALOGS; CELESTIAL REFERENCE SYSTEMS.

F. Stephen Gauss

Bibliography. H. Eichhorn, *Astronomy of Star Positions*, Dover, 1974; R. Green, *Spherical Astronomy*, Cambridge University Press, 1985; H. C. King, *The History of the Telescope*, 1955, reprint, Dover, 2003; G. P. Kuiper and B. M. Middlehurst (eds.), *Telescopes: Stars and Stellar Systems*, vol. 1, University of Chicago Press, 1960; T. Page and L. W. Page, *Telescopes*, Macmillan, 1966; P. van de Kamp, *Principles of Astrometry*, W. H. Freeman, 1967.

Astronomical unit

The basic unit of length in the solar system. The astronomical unit (AU) is also used to a limited extent for interstellar distances through the definition of the parsec (1 pc = 206,265 AU). It is nearly equal to the mean distance a between the center of mass of the Sun and the center of mass of the Earth-Moon system ($a = 1.00\,000\,23$ AU), and for that reason it is often convenient to think of it as the mean distance between the Sun and Earth. *See* PARSEC.

The most accurate determination of the length of the astronomical unit in physical units, such as meters, is obtained from phase-modulated continuous-wave (CW) radio signals beamed to other planets. The round-trip travel times of the signals are determined by cross-correlating the returned signal from the planet with the transmitted signal, and as a result, planetary distances are measured directly.

Prior to the inception of planetary radar technology around 1961, determinations of the astronomical unit relied on parallax effects in the positions of planets and asteroids as viewed from distant points on the surface of the Earth, or on spectroscopic Doppler shifts in the light of stars near the plane of the ecliptic. These techniques were limited in accuracy to about 43,000 mi (70,000 km). On the other hand, early radar determinations were accurate to 300 mi (500 km), and by accumulating several years of planetary distance measurements, it was possible to refine this accuracy to 0.6 mi (1 km), the limit of passive radar bounce technology. Continuous-wave signals returned from the National Aeronautics and Space Administration (NASA) landers on the surface of Mars

resulted in a determination of the astronomical unit to an accuracy of 20 m (65 ft). *See* DOPPLER EFFECT; EARTH ROTATION AND ORBITAL MOTION; PARALLAX (ASTRONOMY).

With an adopted value of 299,792,458 m/s (186,282.397051 mi/s) for the speed of electromagnetic propagation in vacuum, the value of the astronomical unit from NASA's Viking and Pathfinder missions is 149,597,870,692 m (92,955,807.268 mi).

John D. Anderson

Bibliography. D. W. Kurtz and G. E. Bromage (eds.), *Transits of Venus: New Views of the Solar System and Galaxy*, IAU Colloquium 196, 2005; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, 1992.

Astronomy

The study of the universe and the objects in it through scientific investigation. Since much of contemporary astronomy uses the laws and methods of physics, the terms "astronomy" and "astrophysics" are usually used interchangeably. However, modern astronomy also uses techniques from many other scientific disciplines, including chemistry, geology, and biology, for which the terms astrochemistry, planetary science, and astrobiology are increasingly used.

The goal of astronomical research is to understand celestial objects and the nature and evolution of the universe by using whatever techniques are appropriate. Astronomers' breadth of knowledge of different types of astronomical objects often gives insights into understanding any given object or physical process.

Scope. Since the advent of spectroscopy in the late nineteenth century, allowing the remote sensing and analysis of the composition and motion of distant objects, the distinction between astronomy and astrophysics has largely evaporated. The view that astronomers study objects in the sky and astrophysicists seek to explain them has been superseded. Some branches of astronomy that are not strictly included in astrophysics involve the study of the motions of the planets and satellites as well as of spacecraft launched from Earth (celestial mechanics) and the measurement of positions and motions of stars (astrometry). However, the extreme astrometric accuracy of the *Hipparcos* space mission in the 1990s gave data that were quickly applied to fundamental distance determinations that are used in assessing the expansion of the universe and other cosmological parameters. So even the traditional fields of astronomy are now so intertwined with astrophysics that no strict distinction is possible. *See* ASTROMETRY; CELESTIAL MECHANICS.

The use of geological knowledge and methods in analyzing close-up observations from spacecraft of planets and their satellites and of comets and asteroids closely links the disciplines of astronomy and planetary science. Indeed, the discovery of planets around distant stars holds for even closer relations in the future. Methods of studying molecules in interstellar clouds involve chemical knowledge. Planetary science and astrochemistry come together with astronomy in the search for life outside the solar system, part of the search for extraterrestrial intelligence (SETI). The National Aeronautics and Space Administration (NASA), the United States space agency, has placed a priority on astrobiology, including the investigation of Mars and bringing samples back to Earth. *See* ASTEROID; ASTROBIOLOGY; COMET; EXTRATERRESTRIAL INTELLIGENCE; INTERSTELLAR MATTER; PLANET; SATELLITE (ASTRONOMY); SOLAR SYSTEM.

Technologies. Astronomers often lead in employing new technologies, pushing them to the limit in exploring extremely faint signals in various parts of the electromagnetic spectrum. Nearly all astronomical research is now heavily dependent on computers, which control telescopes, help analyze data, make theoretical calculations, display results, provide Internet links for e-mail and the World Wide Web, and allow efficiencies through word processing, image processing, and statistical analysis programs. Even the shapes of telescope mirrors of some of the newest and largest instruments are controlled through computer feedback loops. *See* DIGITAL COMPUTER; IMAGE PROCESSING.

Astronomical imagery is now dominated by light-sensitive silicon chips known as charge-coupled devices (CCDs). In these devices, electric charge results when a pixel is struck by light, and this charge is then transferred to adjacent pixels in the process of being read out. CCDs are approximately 100 times more sensitive than film, and are used at all professional observatories as well as by increasing numbers of students and amateur astronomers. Thus an image or spectrum that formerly took over an hour to record on film can now be recorded in about a minute, with the added benefit over film that the law of reciprocity holds for the data; that is, the strength of the electronic signal produced is directly and linearly proportional to the intensity of the incoming signal. CCDs are sensitive throughout the optical and near-infrared portions of the electromagnetic spectrum. Other electronic devices are used in other parts of the spectrum to record incoming signals. *See* ASTRONOMICAL IMAGING; CHARGE-COUPLED DEVICES.

Fiber optics are used for a variety of astronomical purposes, including the taking of hundreds of galaxy images simultaneously from the field of view of a telescope and bringing the light to a spectrograph that can produce simultaneous spectra of all the objects. This multiplexing effect can increase the efficiency of sky surveys many times. *See* FIBER-OPTICS IMAGING.

The technology of adaptive optics, in which the shape of a mirror is changed slightly at a high rate (often faster than 1 Hz) to compensate for the blurring of astronomical images caused by the Earth's atmosphere, and for slower active optics, was developed largely in secret for military purposes. It has now been released to the public and is being increasingly pursued to eliminate the twinkling of stars. *See* ADAPTIVE OPTICS.

Optical telescopes. The opening of the 5-m (200-in.) Hale telescope at the Palomar Observatory on Palomar Mountain, California, in 1948 marked the beginning of a great period of development in optical astronomy. The light-gathering power of this telescope allowed cosmological study that extended most of the way to the beginning of time in the universe. It was joined in the task by several 4-m-class (160-in.) telescopes and by one less successful larger telescope. Apparently, the maximum size of such traditional telescopes had been reached.

In the 1990s, new techniques of telescope making allowed the completion of several telescopes in the 10-m (400-in.) class, twice the diameter and thus four times the collecting area of the Hale telescope. The first to be completed were the twin telescopes at the W. M. Keck Observatory on Mauna Kea in Hawaii. At this high site (about 4200 m or 13,800 ft in altitude) with excellent steady air (giving steady images known as good seeing), low amounts of water vapor overhead leading to good infrared transparency, and dark and usually clear skies are located several of the largest telescopes in the world. In 1999 the 8-m (320-in.) Subaru ("Pleiades") telescope of the Japanese National Observatory and the 8-m Gemini North telescope of an international consortium (with 50% U.S. participation) were established there. In the Southern Hemisphere, in Chile, better placed to observe various southern objects including the center of the Milky Way Galaxy, the four 8-m telescopes of the European Southern Observatory's Very Large Telescope (VLT) began operation. These telescopes, often operated independently, are also operated together as interferometers, as are the Keck telescopes, giving increased resolution. These techniques are rapidly improving. The Gemini South 8-m telescope also operates in Chile. An even larger telescope, though with limited pointing ability in the sky, is operated in Texas for spectroscopic use, and a similar spectroscopic telescope has been built in South Africa.

Still larger telescopes are in the conceptual design stage, though without definitive funding or construction plans. They include the 30-m (100-ft) California Extremely Large Telescope (CELT) and the Overwhelmingly Large Telescope (OWL).

The existing large telescopes have proven useful in taking spectra of the optical counterparts of gamma-ray bursts, proving that they are very far away; and in analyzing the distances to faraway galaxies and in measuring the redshifts of their spectra, leading to the current cosmological models of the expansion of the universe and the conclusion that the rate of expansion is accelerating. *See* COSMOLOGY; GAMMA-RAY BURSTS; HUBBLE CONSTANT; TELESCOPE.

Telescopes in space. The 1990s saw the thorough use of the vantage points of space for astronomical observation, exemplified by NASA's series of Great Observatories. In 1991 the *Compton Gamma-Ray Observatory* was launched, and in the following years mapped about one gamma-ray burst per day in addition to many other objects and events. The *Hubble Space Telescope* was launched in 1990 to study the ultraviolet and visible parts of the spectrum. Its repair in 1993, with secondary mirrors compensating for a focusing problem with the main mirror, brought it to full working order, and a 1996 upgrade included an improved two-dimensional spectrograph and infrared capability. The *Chandra X-Ray Observatory*, launched in 1999, provides high-resolution x-ray images, and is the same size and scope as Hubble. It studies various types of celestial objects and processes, such as black holes of stellar and galactic sizes. The *Spitzer Space Telescope*, launched in 2003, formerly the *Space Infrared Telescope Facility*, is the fourth of this series of Great Observatories. The European Space Agency's XMM-Newton x-ray telescope and INTEGRAL gamma-ray telescope are also important, large-scale spacecraft. *See* ASTROPHYSICS, HIGH-ENERGY; BLACK HOLE; GAMMA-RAY ASTRONOMY; HUBBLE SPACE TELESCOPE; INFRARED ASTRONOMY; X-RAY ASTRONOMY; X-RAY TELESCOPE.

Smaller spacecraft have also made valuable contributions. The *Cosmic Background Explorer* (*COBE*) in the early 1990s mapped the cosmic background radiation and provided important clues to the origin of the universe and its large-scale structure. It has been superseded by NASA's *Wilkinson Microwave Anisotropy Probe*. NASA's Explorer series continues with a variety of mid-sized spacecraft. The European Space Agency's *Solar and Heliospheric Observatory* (*SOHO*) and *Infrared Space Observatory* (*ISO*) have also sent back valuable observations. *See* COSMIC BACKGROUND RADIATION; SUN.

Telescopes across the spectrum. The atmosphere blocks most of the electromagnetic spectrum from reaching the Earth's surface, leaving windows of transparency mostly in the optical and radio parts of the spectrum. Radio astronomers have made the most of their window of transparency with such telescopes as the 100-m (328-ft) fully steerable telescope outside Bonn, Germany; the 330-m (1083-ft) Arecibo dish in Puerto Rico, in which the secondary mirror can be moved over a spherical bowl covered with mesh to give some limited tracking ability; and the precise, 100-m × 110-m (300-ft × 360-ft) Robert C. Byrd Green Bank Telescope in West Virginia, opened in 2000, for example, to allow studies of interstellar molecules. Radio astronomers pioneered the use of interferometry, taking advantage of the long wavelengths at which they observe. The Very Large Array (VLA) of radio telescopes of the National Radio Astronomy Observatory extends about 25 km (15 mi) over a plain in New Mexico, synthesizing the aperture to provide the resolution (though not the collecting area) of a single telescope of that size. The Very Long Baseline Array uses a dedicated set of telescopes on various sites across the globe to synthesize a telescope the size of the Earth. In India, the Giant Metrewave Radio Telescope near Pune has 30 antennas each 45 m (150 ft) across; larger than the VLA, it is more sensitive at long wavelengths but cannot image at the shorter wavelengths at which the VLA often observes. *See* RADIO ASTRONOMY; RADIO TELESCOPE.

The ozone layer and other constituents of the atmosphere block the shortest wavelengths from penetrating to the Earth's surface, so observations of gamma rays, x-rays, and most of the ultraviolet region require telescopes in space. Similarly, water vapor blocks most of the infrared, allowing only a few windows of transparency, and observations from balloons and, to go higher, from spacecraft are necessary to cover the whole infrared spectrum. *See* OZONE; ULTRAVIOLET ASTRONOMY.

Spectroscopy. Much of astronomy involves breaking down the incoming celestial radiation into its component wavelengths, a process known as spectroscopy. The brighter the object, the finer the spectroscopic resolution in wavelength that can be obtained. Spectroscopic studies can reveal the temperature of an object, the identity and proportions of its chemical elements, and the velocities of its constituents toward and away from the Earth. Light from the Sun and other objects is sometimes polarized, and studies of such polarization can give information about the magnetic fields present or about scattering processes. *See* ASTRONOMICAL SPECTROSCOPY; POLARIMETRY.

Non-electromagnetic-radiation telescopes. Though most of what astronomers study is electromagnetic radiation in its various forms, some particles do arrive at the Earth. The expansive definition of a telescope includes anything used in astronomy to observe the sky. Several neutrino telescopes have been used to detect neutrinos from the Sun and, in one instance, from a supernova. The original neutrino telescope was a large tank of cleaning fluid based in a mine far underground. Subsequent neutrino telescopes are tanks of gallium and even larger tanks of purified water, in which interactions with neutrinos have detectable consequences. These neutrino telescopes have pinpointed the "solar neutrino problem" as arising from problems in understanding the fundamental physics of elementary particles rather than from uncertainties or errors in the astronomical measurements. *See* NEUTRINO; NEUTRINO ASTRONOMY; SOLAR NEUTRINOS.

Cosmic rays are particles from the Sun and from more distant objects in space. These primary cosmic rays interact with particles in the Earth's atmosphere to make secondary cosmic rays, and the pace of observation of these secondary cosmic rays as well as the few primary cosmic rays that reach the Earth is increasing. *See* COSMIC RAYS.

Gravitational waves are a consequence of Einstein's general theory of relativity, and their existence was verified indirectly by careful study of the period of a binary pulsar. A pair of interferometers was built on Earth to attempt direct detection of such gravitational waves, which should result from such distant events as the merger of two neutron stars. This Laser Interferometer Gravitational-wave Observatory (LIGO) is in operation in Louisiana and in Washington but is in constant modification to improve its sensitivity, which is currently not at the level necessary to expect discoveries. Similar facilities are in operation and in construction at other locations around the world. *See* GRAVITATIONAL RADIATION.

Theory and computation. Theoretical calculations of the nature of astronomical objects or processes are known as theoretical astrophysics. Theoretical astrophysicists analyze the evolution of stars, analyze the solar wind of expanding plasma around the Sun, trace the explosion phase of supernovae, calculate the distribution of sizes of structures formed in the early eons of the universe, or analyze the formation of the light elements in the first minutes after the big bang. Cosmological calculations use Einstein's general theory of relativity. The discovery of the dark energy that makes up about 70% the content of the universe has led to a wide-ranging set of new cosmological ideas. *See* BIG BANG THEORY; PLASMA (PHYSICS); RELATIVITY; SOLAR WIND; STAR; STELLAR EVOLUTION; SUPERNOVA.

The availability of supercomputers, powerful and fast computers capable of handling large amounts of data, has led to three-dimensional simulations of, for example, the effects of varying amounts of cold dark matter (such as undiscovered types of subatomic particles) and hot dark matter (such as neutrinos) in the formation of large-scale structure in the early universe. Such dark matter, making up approximately 30% the contents of the universe, apparently exceeds the amount of matter detectable through more traditional methods of observation. From a variety of observations, including those of distant supernovae and those of the cosmic background radiation from the *Wilkinson Microwave Anisotropy Probe*, it has been found that the bulk of the contents of the universe, about 70%, are of dark energy. This dark energy may be in the form of the cosmological constant, energy transformable into mass in an amount calculable with Einstein's equation. *See* SIMULATION; SUPERCOMPUTER; UNIVERSE.

Models of the oscillations detectable on the Sun's surface through long-time-series observations made on the ground from the Global Oscillation Network Group (GONG) or from space with instruments on the *Solar and Heliospheric Observatory* (*SOHO*) are used to improve understanding of the solar interior, a process known as helioseismology. *See* HELIOSEISMOLOGY.

Experimental astrophysics. Laboratory astrophysics involves the measurement of basic parameters that are used in calculations of physical or chemical processes relevant to astronomy. Though the Sun and stars as well as interstellar space are excellent laboratories for studying processes at high temperature or in vacuums unobtainable on Earth, sometimes fundamental parameters such as cross sections of atomic and molecular collisional excitation and ionization, as well as nuclear parameters, can be measured in laboratories on Earth. *See* ATOMIC STRUCTURE AND SPECTRA; MOLECULAR STRUCTURE AND SPECTRA.

Historical astronomy. Studying astronomy's history provides data for comparison with the present. For example, the paths of solar eclipses from hundreds or thousands of years ago can be analyzed to study

the rate of rotation of the Earth and the relation between astronomical timekeeping and timekeeping by atomic methods, leading to an improved understanding of processes inside the Earth. *See* ARCHEOASTRONOMY; DAY; EARTH ROTATION AND ORBITAL MOTION; ECLIPSE; TIME; YEAR. Jay M. Pasachoff

Bibliography. G. R. Burbidge et al. (eds.), *Annual Review of Astronomy and Astrophysics*, Annual Reviews; M. Cassé (transl. by S. Lyle), *Stellar Alchemy: The Celestial Origin of Atoms*, Cambridge University Press, 2003; K. Croswell, *The Universe at Midnight*, Free Press, 2001; M. Harwit, *Astrophysical Concepts*, 3d ed., Springer, 1998; M. A. Hoskin, *The Cambridge Concise History of Astronomy*, Cambridge University Press, 1999; D. Leverington, *A History of Astronomy: From 1890 to the Present*, Springer, 1995; J. M. Pasachoff, *Astronomy: From the Earth to the Universe*, 6th ed., Brooks/Cole, 2002; J. M. Pasachoff and A. Filippenko, *The Cosmos: Astronomy in the New Millennium*, 2d ed., Brooks/Cole, 2004, 3d ed., 2007.

Astrophysics, high-energy

The study of the universe as revealed by high-energy, invisible forms of light: x-rays and gamma rays. These radiations are produced in the cosmos when gas is heated to millions of degrees Kelvin or electrons have been accelerated to near the speed of light by violent and extreme conditions. Exploding stars, neutron stars, black holes, and galaxy clusters, the most massive objects in the universe, are among the objects studied.

Instrumentation. The high energies of x-rays and gamma rays have two important consequences for astronomical research. First, these forms of light are absorbed by the atmosphere, so telescopes to detect them must be placed on spacecraft above the atmosphere. Second, the telescopes must be constructed differently. Gamma rays have such high energy that they cannot be focused by traditional techniques, although indirect methods can give a rough estimate of their direction. *See* SATELLITE ASTRONOMY.

X-rays will reflect off mirrors, but only if they strike at grazing angles, like a stone skipping across a pond. For this reason, x-ray mirrors have to be carefully shaped and aligned nearly parallel to the incoming x-rays. These barrel-shaped mirrors are nested one inside the other to increase the collection area, and therefore the sensitivity, of the telescope.

The *Chandra X-ray Observatory*, launched by the National Aeronautics and Space Administration (NASA) in July 1999, is the premier focusing x-ray telescope. It is an assembly of four pairs of mirrors. Chandra's mirrors are the smoothest mirrors ever constructed. The largest of the mirrors is almost 4 feet (1.2 m) in diameter and 3 ft (0.9 m) long. *See* X-RAY TELESCOPE.

The European Space Agency's *XMM*, a powerful telescope launched in December 1999, has 58 mirrors. These mirrors are not as smooth as *Chandra*'s mirrors, so *XMM* cannot make images of the same crispness, but it can detect fainter sources and measure the energies of x-rays very accurately.

NASA's *Rossi X-ray Timing Explorer* (*RXTE*), launched in December 1995, has the ability to study changes in the intensity of x-rays produced in the violent environment around neutron stars and black holes on time scales ranging from microseconds to months. Two other NASA missions involving international cooperation, the *High Energy Transient Explorer* (*HETE-2*), launched in 2001, and the *Swift* satellite, launched in 2004, are dedicated to the exploration of short-lived bursts of x-rays and gamma rays.

NASA's *Compton Gamma-Ray Observatory*, which inaugurated a new era in gamma-ray astronomy with its unprecedented sensitivity and coverage of a large range of gamma-ray energies, was deorbited and reentered the Earth's atmosphere on June 4, 2000. The *Compton Observatory* has been succeeded by the European Space Agency's *International Gamma-Ray Astrophysics Laboratory* (*INTEGRAL*), launched in October 2002, an observatory with a complement of gamma-ray, x-ray, and optical telescopes. Several new telescopes are designed to observe flashes of visible light created high in Earth's atmosphere when extremely high energy (teraelectronvolt or TeV) gamma rays hit the upper atmosphere. For example, the *High Energy Spectroscopic System* (*HESS*), an array of four 12-m (39-ft) mirrors on a mountain in central Namibia, has made a rough image of TeV gamma rays from the remnant of a supernova explosion. *See* CERENKOV RADIATION; GAMMA-RAY ASTRONOMY.

Supernovae. When a massive star (ten or more times as massive as the Sun) has used up the nuclear fuel that makes it shine, the pressure drops in the central core of the star. Gravity crushes the matter in the core to higher and higher densities. Temperatures soar to billions of degrees Kelvin. The intense heat generated in the collapse produces a cataclysmic rebound that sends high-speed debris flying outward at speeds in excess of 5000 mi/s (8000 km/s). A thermonuclear shock wave races through the now expanding stellar debris, fusing lighter elements into heavier ones and producing a brilliant visual outburst with the brightness of several hundred million suns.

A massive star explodes about once every 50 years in the Milky Way Galaxy. The shell of matter thrown off by the supernova creates a magnetized bubble of multimillion-degree gas and high-energy particles called a supernova remnant. The hot gas expands and produces x-rays for thousands of years (**Fig. 1**). Gamma rays from radioactive elements have also been detected from supernova remnants by gamma-ray telescopes such as those that were on the *Compton Gamma-Ray Observatory*.

The study of remnants of exploded stars, or supernovae, is essential for understanding the origin of life on Earth. The cloud of gas and dust that collapsed to form the Sun, Earth, and other planets was composed mostly of hydrogen and helium, with a small amount of heavier elements such as carbon, nitrogen, oxygen, and iron. The only place where these and other

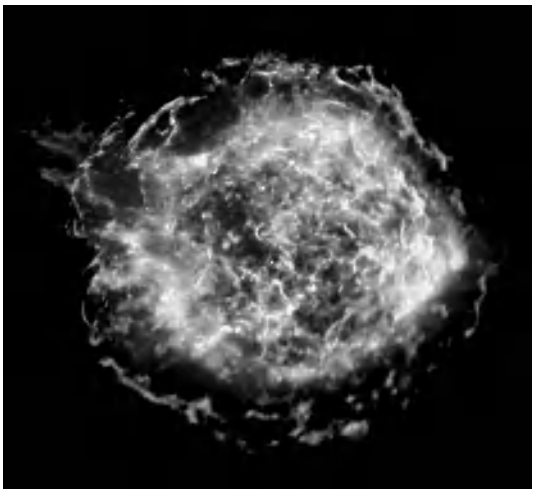

Fig. 1. ***Chandra X-ray Observatory*** **image of Cassiopeia A (Cas A), the 320-year-old remnant of a massive star that exploded. The image shows an expanding shell of hot gas produced by the explosion. This gaseous shell is about 10 light-years in diameter and has a temperature of about 5×10^7 Kelvin. (*NASA/Chandra X-ray Observatory Center/Goddard Space Flight Center/U. Hwang et al.*)**

heavy elements necessary for life are made is deep in the interior of a massive star. There they remain until a supernova explosion, spreads them throughout space. *See* NUCLEOSYNTHESIS; SUPERNOVA.

Neutron stars. When a massive star explodes, most of it is flung into space, but the core of the star is compressed to form a rapidly rotating dense ball of neutrons that is about 12 mi (20 km) in diameter. The collapse and rapid rotation of the neutron star cause it to become highly magnetized. A magnetized, rapidly rotating neutron star can produce electric voltages of 10^{16} V.

Neutron star gravity, which is more than 10^{11} times stronger than gravity on Earth, is overwhelmed by the electric field, and particles are pulled off the neutron star and accelerated to speeds near the speed of light. An intense shower of electrons and antimatter electrons, or positrons, is produced by these particles. The pulsed emission from the Crab Nebula, observed at all wavelengths from radio through gamma rays, is thought to be caused by this process (**Fig. 2**). *See* CRAB NEBULA; PULSAR.

As particles stream out from the pulsar and spiral around magnetic field lines, they produce a distinctive kind of radiation known as synchrotron radiation. The Crab Nebula's bell-shaped appearance in the x-ray image is due to synchrotron radiation from a huge magnetized bubble of high-energy electrons that is several light-years in diameter. Dozens of these so-called pulsar wind nebulae have been discovered by the *Chandra X-ray Observatory*. *See* SYNCHROTRON RADIATION.

The rotation-powered activity of neutron stars such as the Crab pulsar can last only a few thousand years. However, if the neutron star has a nearby companion star, its x-ray intensity may increase again in a million years or so. When the companion star enters the red giant stage of its life, it will increase greatly in size and gas will flow from the giant star onto the neutron star. The gas will be heated to tens of millions of degrees Kelvin as it falls onto the surface of the neutron star, and will glow brightly in x-rays.

The Milky Way Galaxy contains several hundred of these neutron star x-ray binaries. Depending on the details of the rate at which the matter falls onto the neutron star, and how the magnetic field of the neutron star guides the inflow, the star will be observed to pulse, flicker, or flare up violently in x-rays. X-ray binaries provide a unique opportunity to study neutron stars. A similar process allows the study of even stranger objects, black holes. *See* BINARY STAR; NEUTRON STAR.

Black holes and quasars. When some very massive stars collapse, they will form black holes. A black hole does not have a surface in the usual sense of the word. There is simply a region in space around a black hole beyond which nothing can be seen, because nothing can escape from inside this region. This region is called the event horizon.

Anything that passes beyond the event horizon is doomed to be crushed as it descends ever deeper into the gravitational well of the black hole. Neither visible light, nor x-rays, nor any other form of electromagnetic radiation given off by the particle can escape.

A black hole cannot be seen directly. One way to find one is by observing the energy released by matter that is falling toward the black hole. As gas and dust particles swirl toward a black hole, they speed up and form a flattened disk. Friction caused by collisions between the particles heats them to

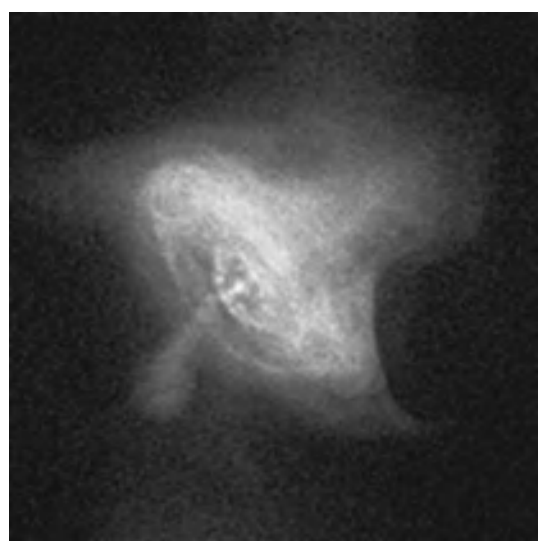

Fig. 2. ***Chandra X-ray Observatory*** **image of the Crab Nebula, a supernova remnant and pulsar in the constellation Taurus. The image shows the central pulsar, a rapidly spinning neutron star, or pulsar that emits pulses of radiation 30 times a second, surrounded by tilted rings of high-energy particles that appear to have been flung outward over a distance of more than a light-year from the pulsar. (*NASA/Chandra X-ray Observatory Center/Smithsonian Astrophysical Observatory*)**

extreme temperatures. Just before the particles pass beyond the event horizon, they produce x-rays and gamma rays as their temperatures approach 10^8 Kelvin.

One of the best methods for finding a black hole has been to study x-ray binary systems. Although most of these systems consist of a normal star in a close orbit around a neutron star, there are about 20 cases in which the observations indicate that the mass of the invisible companion star is so great—more than three times the mass of the Sun—that it must be a black hole.

Neutron stars or black holes may be the explanation for one of the most important unsolved problems in astrophysics, gamma-ray bursts. As determined by the BATSE detector that was on the *Compton Gamma-Ray Observatory* and the Italian-Dutch satellite *Beppo-Sax*, these mysterious outbursts are observed to occur about once per day. In about a minute, gamma-ray bursts release as much energy as the Sun will give off at all wavelengths in its lifetime of 10^{10} years. There appear to be at least two types of gamma-ray bursts. Short bursts are likely produced by the merger of two neutron stars, or a neutron star and a black hole. Long bursts represent the explosions of extremely massive stars, greater than 50 times the mass of the Sun. According to the theory, a massive black hole forms in the center of the star just before the explosion. As matter in the center of the star pours into the black hole, a titanic explosion occurs, ejecting matter outward at nearly the speed of light. *See* GAMMA-RAY BURSTS.

Black holes grow when matter falls into them. A black hole in the center of a galaxy where stars are densely packed may grow to the mass of 10^9 suns. Energy released from large clouds of gas as they fall into these supermassive black holes can be stupendous. This is the accepted explanation for quasars, sources in which the power output at the center of a galaxy can be a thousand times greater than an entire galaxy of 10^{11} stars. *See* QUASAR.

One of the most intriguing features of supermassive black holes is that they do not suck up all the matter that falls within their sphere of influence. Some of the matter falls inexorably toward the black hole, and some explodes away from the black hole in high-energy jets that move at near the speed of light (**Fig. 3**). These jets produce radio, optical, x-ray, and gamma radiation. The matter swirling around the black hole must somehow be producing enormous electric and magnetic fields that accelerate electrons to extremely high energies. Exactly how this happens is unknown and is a major focus of research. *See* BLACK HOLE.

Galaxy clusters and dark matter. More than half of all galaxies in the universe are members of groups of galaxies or larger collections of galaxies, called clusters. X-ray observations have shown that most clusters of galaxies are filled with vast clouds of multimillion-degree gas. The mass of this gas, which was heated when it collapsed from a much larger size, is greater than all the stars in all the galaxies in a cluster of a thousand galaxies. Galaxy clusters are the largest and most massive gravitationally bound objects in the universe.

The x-ray-producing hot gas found in a typical cluster of galaxies presents a great mystery. Over time this extremely hot gas should escape the cluster, since the galaxies and gas do not provide enough gravity to hold it in. Yet the gas remains in clusters of all ages. Scientists have concluded that some unobserved form of matter, called dark matter, is providing the gravity needed to hold the hot gas in the cluster. An enormous amount of dark matter is needed—about three to ten times as much matter as that observed in the gas and galaxies. This means that most of the matter in the universe may be dark matter.

The candidate that best reproduces the observations is called cold dark matter—hypothetical subatomic particles that produce no light and can at present be detected only through gravity. Detailed measurements of the size and temperature of the hot gas clouds in galaxy clusters with x-ray telescopes could help solve the dark matter mystery. *See* DARK MATTER; GALAXY, EXTERNAL.

While an explanation for the source of dark matter is still lacking, an effect that is even more enigmatic has been discovered. Astronomers have observed that the visible light from type 1a supernovae, which act as standard candles, is fainter than expected in distant galaxies. The best explanation is that they are more distant than originally thought, which implies that the expansion of the universe must be accelerating. *Chandra*'s measurements of the dark-matter content of clusters of galaxies have verified this astounding result by an independent method.

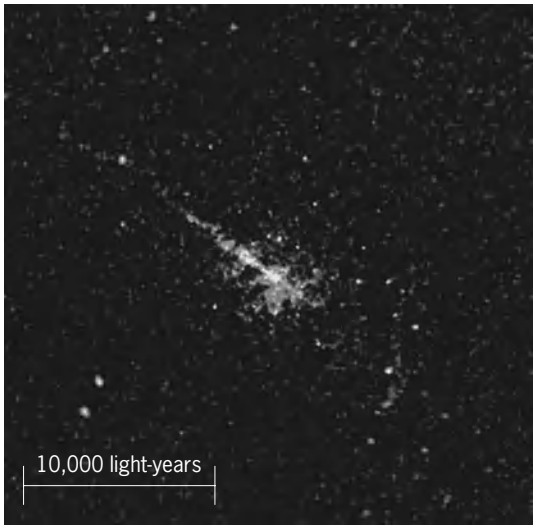

Fig. 3. ***Chandra X-ray Observatory*** **image of NGC 5128, a radio galaxy in the constellation Centaurus, 10^7 light-years from Earth. The image shows a bright source in the nucleus of the galaxy, which is thought to be due to a supermassive black hole. The jet extending to the upper left, far outside the galaxy, is caused by explosive activity around the black hole. (*NASA/Chandra X-ray Observatory Center/Smithsonian Astrophysical Observatory*)**

Cosmic acceleration can be explained if the space between galaxies is filled with a mysterious dark energy that has the property that, as the universe expands, more dark energy is created. The existence of dark energy requires either a modification of Einstein's theory of general relativity or a major revision of some other area of fundamental physics. Assuming that dark energy is responsible for the acceleration, combining the *Chandra* results with observations of type 1a supernovae and the cosmic microwave background radiation indicates that dark energy makes up about 75% of the energy density of the universe, dark matter about 21%, and visible matter about 4%. *See* COSMIC BACKGROUND RADIATION; COSMOLOGY; DARK ENERGY; RELATIVITY; UNIVERSE; X-RAY ASTRONOMY.

Wallace Tucker

Bibliography. M. Begelman and M. Rees, *Gravity's Fatal Attraction*, Scientific American Library, 1996; D. Berry, *Smithsonian Intimate Guide to the Cosmos*, Smithsonian/Madison Press, 2004; P. Charles and F. Seward, *Exploring the X-ray Universe*, Cambridge University Press, 1995; J. Pasachoff, *A Field Guide to the Stars and Planets*, 4th ed., Houghton Mifflin, 2002; G. Schilling, *Evolving Cosmos*, Cambridge University Press, 2004; E. Schlegel, *The Restless Universe*, Oxford University Press, 2002; J. Silk, *On the Shores of the Unknown: A Short History of the Universe*, Cambridge University Press, 2005; K. Thorne, *Black Holes and Time Warps*, W. W. Norton, 1994; W. Tucker and K. Tucker, *Revealing the Universe: The Making of the Chandra X-ray Observatory*, Harvard University Press, 2001; K. Weaver, *The Violent Universe: Joyrides through the X-ray Cosmos*, Johns Hopkins University Press, 2005.

Asymmetric synthesis

A chemical reaction or series of reactions leading to the predominant or exclusive formation of a single enantiomer, that is, one of an isomeric pair of chemical compounds whose molecules are nonsuperimposable mirror images. Enantiomers are said to be chiral. Many organic compounds of interest for pharmaceuticals, flavorings, and pesticides are chiral. A specific target of an asymmetric synthesis is a product that contains only the desired enantiomer rather than a mixture of enantiomers of chiral molecules. *See* MOLECULAR ISOMERISM.

One of the most common structural features of many chiral molecules is the presence of a chiral center, a carbon atom in a molecule that has four different atoms or groups attached to it. For example, two enantiomers, (**1**) and (**2**), of naproxen exist and

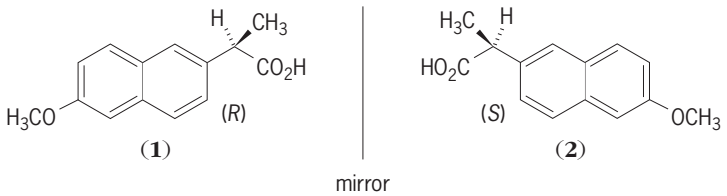

have the same constituents, but owing to the relative orientations of their substituents in space, these structures are not superimposable. In depicting three-dimensional structures, drawings use a wedge-shaped bond to the methyl group to indicate that it is above the plane of the naphthyl group and a dashed-line bond to indicate that it is below the plane. There are rules of nomenclature that unambiguously label the three-dimensional aspects of a center by a chirality descriptor, such that (**1**) is (*R*)-naproxen and (**2**) is (*S*)-naproxen. If a mirror plane is present in a molecule, it is not chiral and its mirror image will be identical with the original.

Although a chiral center is the most common source of chirality, one may also have chirality associated with an axis that can give rise to enantiomers. 2,2′-Bis(diphenylphosphino)-1,1′-binaphthyl (BINAP) is a common ligand used in transition-metal catalysis for synthesis of one enantiomer of a product, and BINAP can be obtained as the pure (*R*) or (*S*) enantiomer, (**3**) and (**4**).

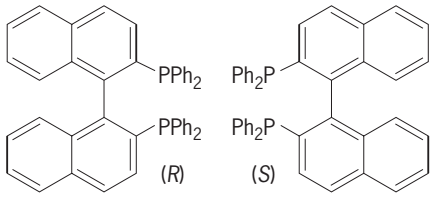

(**3**) (**4**)

See STEREOCHEMISTRY.

Chirality in molecules is an important factor in determining their biological activity. One of the enantiomers of a chiral substance is often more effective than the other for applications in pharmacology, agriculture (pesticides or herbicides), or food chemistry (flavoring agents). Most complex molecules in a living organism are chiral and therefore interact with enantiomers of a given substance in different ways. This is true with regard to perception of odors where the enantiomers of carvone (**5** and **6**) have

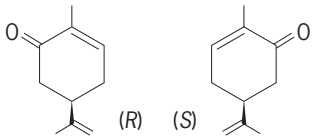

(**5**) (**6**)

distinctively different fragrances: the (*R*) enantiomer smells like spearmint and the (*S*) enantiomer like caraway seeds. Pharmacological effects can also be pronounced; for example, (2*S*, 3*R*)-propoxyphene (**7**) is an analgesic, whereas (2*R*, 3*S*)-propoxyphene (**8**) is a cough suppressant. In other cases, one enantiomer can be effective and the other toxic. In the past, drugs were often sold as racemates (an equal mixture of both enantiomers) when the least effective enantiomer was not a risk or was converted to the active enantiomer in the body. *See* RACEMIZATION.

There has been a movement toward preparation of pure enantiomers for pharmaceuticals owing to the

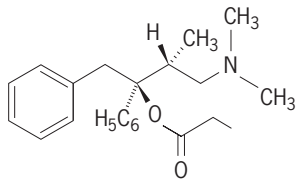

(**7**)

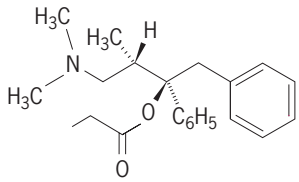

(**8**)

potential side effects of the ineffective enantiomer, which, if not converted to the active enantiomer, at least would present a metabolic burden on an individual. Another advantage of providing a single enantiomer drug is that, as viewed by the patent office, the pure enantiomer is a different compound than the racemate. Hence, the racemic version of omeprazole, a $6 billion-a-year drug, is marketed in the enantiomerically pure form as esomeprazole (**9**). The "es"

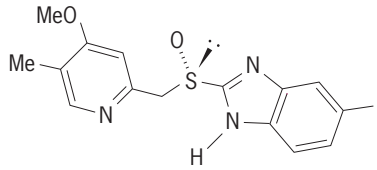

(**9**)

corresponds to the (*S*) configuration of the sulfoxide. This marketing tactic for preserving patent rights is known as a chiral switch.

Distinguishing enantiomers. The chirality descriptors (*R*) and (*S*) provide a method of labeling the three-dimensional aspects of a chiral center, and for a molecule with one center would describe the absolute configuration of that molecule. Although these enantiomeric molecules may behave differently in a biological system or with other chiral molecules, almost all of their physical properties are identical. One exception is their interaction with polarized light. Plane-polarized light passing through a solution containing a single enantiomer will rotate the plane of polarization clockwise (dextrorotatory or +) or counterclockwise (levorotatory or −), for a given wavelength of light. This provides an experimental parameter by which a particular enantiomer can be identified. The relationship between the direction of rotation of light and the absolute configuration is complex, and this physical property cannot often be used reliably to determine absolute configuration. For compounds containing heavy atoms, x-ray scattering can provide a method for determining absolute configuration in a crystal structure, and this provides a method for correlating observed optical rotation with configuration. When both are known, the nomenclature often signifies the physical property, as well as the absolute configuration, for example, (*R*)-(−)-carvone (**5**). By knowing the degree of rotation of a standard concentration (1 g/L) of a pure enantiomer, one can calculate a specific rotation, $[\alpha]_D$. Hence the amounts of each enantiomer in a sample that is not enantiomerically pure can often be determined by measuring its $[\alpha]_D$. The $[\alpha]_D$ of (**5**) is -61°, whereas the $[\alpha]_D$ of (**6**) is $+61^\circ$, and an equal mixture of both, a racemate, would have an $[\alpha]_D$ of 0°. *See* OPTICAL ACTIVITY.

A sample that contains an unequal amount of enantiomers is said to be chiral nonracemic, and the enantiomeric excess (ee) is given by Eq. (1), where *R* and

$$\text{ee} = \frac{|R - S|}{R + S} \times 100 \qquad (1)$$

S correspond to the concentrations of the respective enantiomers.

This definition in terms of percentages of enantiomers is analogous to the effect on $[\alpha]_D$ of the presence of one enantiomer rotating light in the opposite direction to the other in mixtures. The term ee is now preferred over the older term of optical purity as the rotation may not be proportional to concentration owing to effects such as dimerization. Modern methods, such as gas chromatography and nuclear magnetic resonance (NMR) spectroscopy, use interactions with other chiral molecules to determine the percentages of enantiomers.

In the absence of the influence of nonracemic substances, symmetrically constituted reagents would be expected to yield racemic mixtures. The aim of an asymmetric synthesis is to selectively prepare enantiomerically pure products through the use of a nonracemic material, particularly avoiding the necessity of carrying out physical separations.

Strategies. If one prepares a racemate, the enantiomers can be separated in a number of ways.

Diastereomer separation. A classic method involves reaction of the mixture of enantiomers with an enantiomerically pure reagent and then separating the diastereomeric products. Diastereomers are no longer related as mirror images and thus have different chemical and physical properties. For example, an enantiomerically pure acid (**10**) would react with a racemic chiral alcohol, (**11**) and (**12**), to yield diastereomeric esters [reaction (2)].

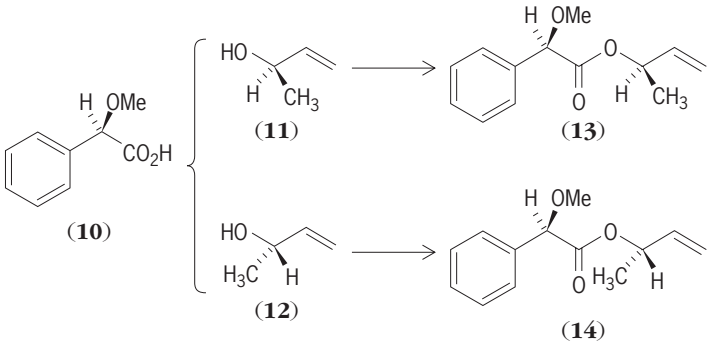

(**10**) (**11**) (**12**) (**13**) (**14**) (2)

Since diastereomers (**13**) and (**14**) are formed, they can be distinguished by physical methods such

as NMR or retention times on a gas chromatograph. This approach would provide an analysis of the enantiomeric purity of a nonracemic alcohol without resorting to the use of optical rotation.

In many cases, diastereomers can be separated by various physical methods, such as crystallization, and subsequent removal of the original chiral fragment would provide a method for isolation or resolution of a pure isomer. Thus, hydrolysis of pure (**13**) would yield (**11**) of 100% ee.

This type of separation of enantiomers does not fit the aim of an asymmetric synthesis, where physical separations are supposed to be avoided in principle.

Chiral reagent. If a chiral reagent, substituent in the molecule, or other entity influences the formation of new chiral center, one has the essence of an asymmetric synthesis. A common approach is the use of a chiral reagent. An example is the asymmetric reduction of a deuterated aldehyde with alpine-borane (**15**) which allows recovery of α-pinene (**17**) from which the reagent was derived and provides the chiral alcohol in ~100% ee [reaction (3)]. Note that (**18**) would be nearly impossible to resolve from a

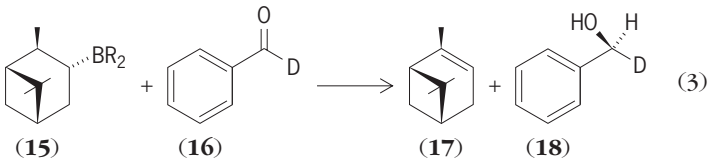

racemate by a conventional resolution via diastereomers.

Another reliable strategy is the temporary incorporation of an enantiomerically pure chiral auxiliary to direct the stereochemistry of a reaction and subsequent cleavage of it from the product. For example, if one wishes to enantioselectively substitute an α-position on a carboxylic acid, derivatization with an amino alcohol with a directing group can provide the necessary stereocontrol of the reaction [reaction (4)].

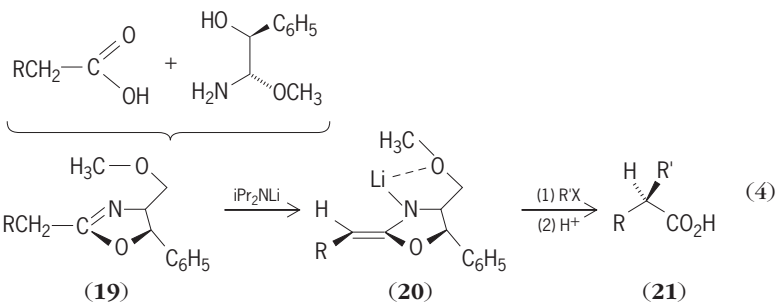

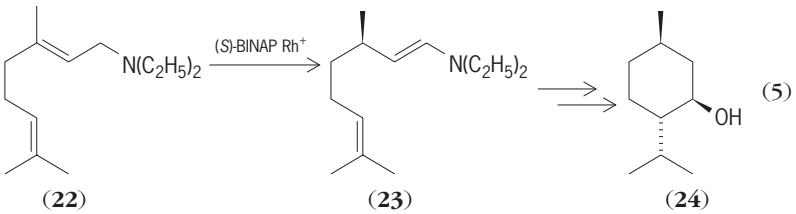

Chiral catalyst. These strategies require a 1:1 use of chiral reagent or auxiliary per chiral molecule produced. Ideally, one would prefer to have the chirality multiplied in some way that large numbers of enantiopure product molecules could be formed for each enantiopure agent that was used. Asymmetric catalysis provides a way to do this. An example is the Takasago process that is used commercially for the preparation of (−)-menthol (**24**). Cationic rhodium complexes prepared from BINAP are very efficient catalysts for the enantioselective isomerization of allylic amines to enamines [reaction (5)].

Thus, the chirality within the catalyst is transferred in the isomerization of the achiral geranylamine (**22**) to the enamine (**23**). A small amount of catalyst can isomerize a large number of molecules and create chiral product in high enantiomeric yield, and this provides a method of multiplying the chirality that exists in the catalyst. Asymmetric catalysis is an important component of modern approaches to asymmetric synthesis. *See* CATALYSIS.

Chiral amplification. Intuitively one would expect that an enantiomerically pure catalyst would be required to yield products in high enantiomeric purity; however, this is not always the case. The rate of addition of diethylzinc to benzaldehyde is dramatically enhanced by the addition of an amino alcohol. If the amino alcohol is chiral and nonracemic, an enantioselective addition can take place. When enantiopure DAIB [3-exo-(dimethylamino)isoborneol] (**26**) is used, a product with 98% ee is obtained [reaction (6)].

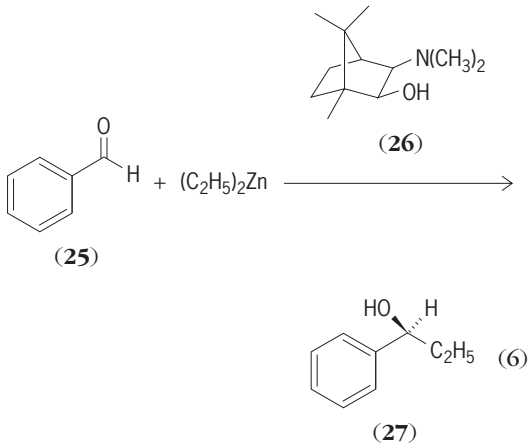

A fascinating feature of this reaction is that when DAIB has an enantiomeric purity of only 21% ee, the product (**27**) is obtained in 96% ee. The origin of this effect can be found in the relative stability of dimers containing either two molecules of the same enantiomer of DAIB or different enantiomers [reaction (7)].

Monomers (**29**) and (**31**) are considered to be the active species. In this case, the homo dimers (**28**) and (**30**) are highly dissociated, whereas the hetero dimer (**32**) remains mostly intact. Thus, virtually all of monomer (**31**) is tied up in the unreactive hetero dimer, and this leaves the excess (**29**) as the predominant species effecting the reaction.

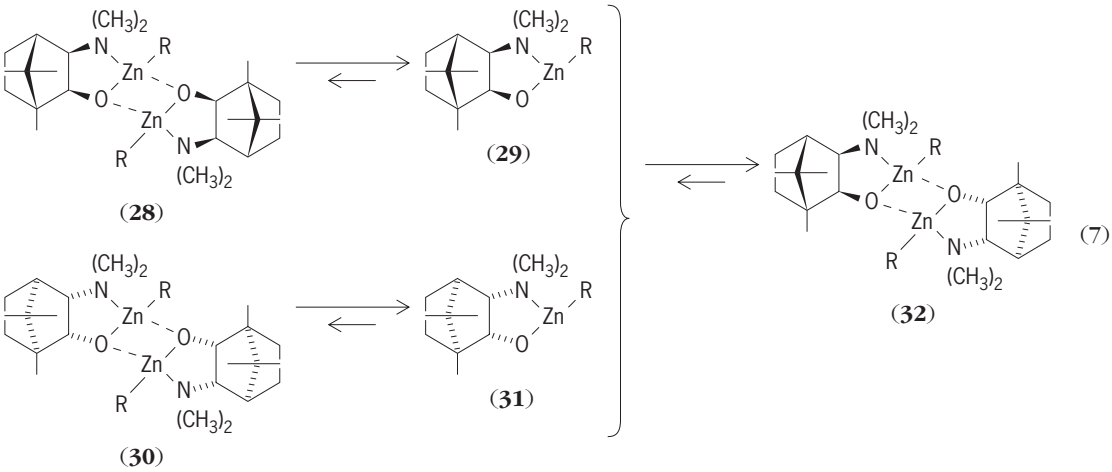

Kinetic resolution. Another approach to obtaining enantiomers in high ee from a racemate involves the enantioselective conversion of one of the enantiomers into a different compound. This can be achieved with enzyme hydrolysis of esters, but one of the most commonly used methods for deracemization employs selective epoxidation of allylic alcohols using tartrate esters [reaction (8)].

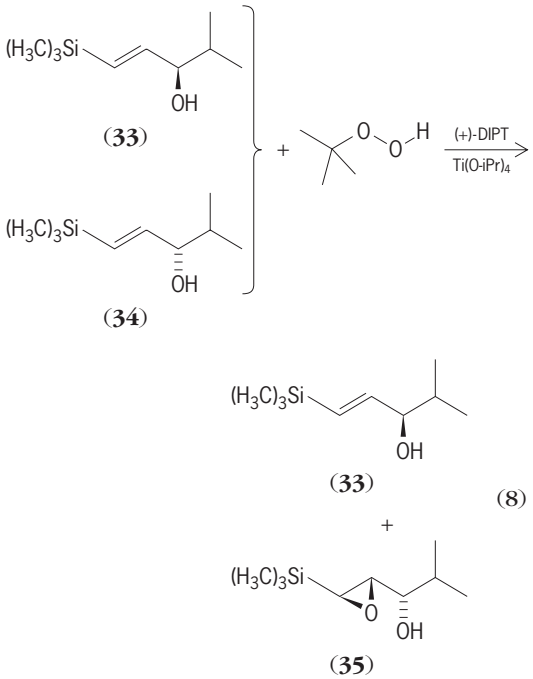

This approach not only provides an enantioselective synthesis of an epoxide (**35**) from (**34**) in 99% ee, but also leaves (**33**) behind in >99% ee. A disadvantage of this approach is that half of the starting material remains unreacted if the goal was to obtain pure epoxide.

Dynamic kinetic resolution. The problem associated with 50% conversion of a racemic starting material can be overcome if a method for racemization of the starting material is available. Under certain conditions, racemization of the substrate may occur on a time scale that is fast compared to the reaction. This can be observed in ketone hydrogenations [reaction (9)]. In this case, (**36**) hydrogenates more rapidly than (**37**) with a ruthenium catalyst using the (*R*)- BINAP ligand.

Treating palladium with (**39**) provides a system that effectively racemizes the intermediates formed with diene monoepoxides, but also catalyzes enantioselective attack by pthalimide (**40**).

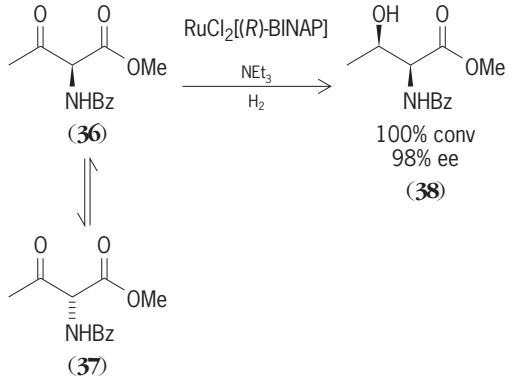

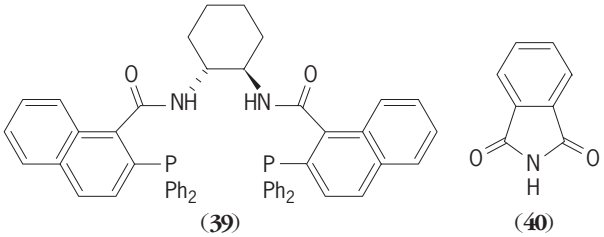

This procedure has been used by B. M. Trost to convert the racemic epoxide (**41**)/(**42**) to a precursor of the drug (*S*)-vigabatrin (**44**) [reaction (10)]. Since this process does not involve prior

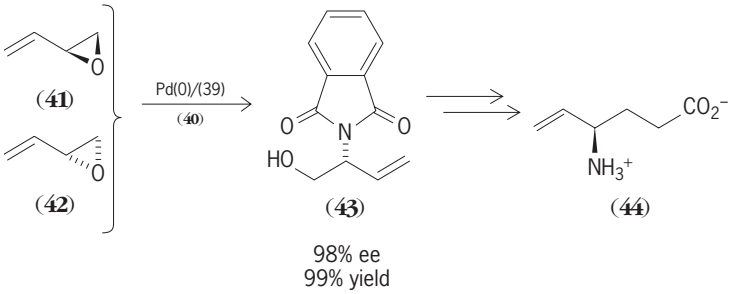

racemization of the substrate, it has been designated as a dynamic kinetic asymmetric transformation.

J. W. Faller

Bibliography. C. Bolm and J. A. Gladysz, Enantioselective catalysis, *Chem. Rev.*, 103:2761–3400, 2003; E. Eliel, S. Wilen, and L. Mander, *Stereochemistry of Organic Compounds*, Wiley, New York, 1994; R. E. Gawley and J. Aube, *Principles of Asymmetric Synthesis*, Elsevier, Oxford, 1996; R. Noyori, *Asymmetric Catalysis in Organic Synthesis*, Wiley, New York, 1994; J. Seyden-Penne, *Chiral Auxiliaries and Ligands in Asymmetric Synthesis*, Wiley, New York, 1995.

Asymptote

A line that is a limit of lines tangent to a curve as the contact points of those tangents approach infinity along the curve. Thus, an asymptote of a curve is an ordinary line (that is, not the "line at infinity") that is tangent to a curve at the points in which the curve intersects the line at infinity (see **illus.**). Analytically,

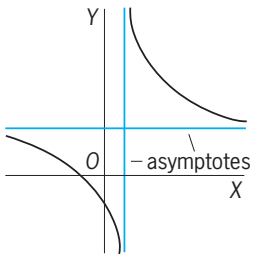

Asymptotes of a hyperbola.

a line $y = mx + b$ is an algebraic curve $f(x, y) = 0$, provided the constants m and b are such that the coefficients of the two terms of highest degree of $f(x, mx + b)$ are zero. For example, to find m and b so that $y = mx + b$ is an asymptote of $x^3 + y^3 - 3xy = 0$, elimination of y yields

$$(1 + m^3)x^3 + 3x^2(bm^2 - m) + 3x(mb^2 - b) + b^3 = 0$$

Equating the two leading coefficients to zero gives $m = b = 1$. Thus $y + x + 1 = 0$ is an asymptote. *See* ANALYTIC GEOMETRY; HYPERBOLA.

Leonard M. Blumenthal

Ateleopodiformes

An order of actinopterygian (ray-finned) fishes consisting of only one family, Ateleopodidae, known as jellynose fishes. Ateleopodiformes along with the order Stomiiformes constitute a teleost superorder called the Stenopterygii. Jellynose fishes are very distinctive in having a large head and bulbous snout; a long anal fin that unites with a much reduced caudal fin; a dorsal fin placed forward with 3 to 13 rays; a pelvic fin in a jugular position with a single ray in adults; and a largely cartilaginous skeleton (see **illustration**). Individuals may reach a length of 2 m (6.6 ft) at a depth range of 140–1281 m (called the bathydemersal environment). The family, consisting of 4 genera and 12 species, is represented in the eastern Atlantic, Caribbean Sea, eastern Pacific off Central America, and the Indo-West Pacific. *See* ACTINOPTERYGII; TELEOSTEI.

Herbert Boschung

Bibliography. J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994; M. M. Smith and P. C. Heemstra (eds.), *Smith's Sea Fishes*, Macmillian South Africa, Johannesburg, 1986.

***Itimaia plicatellus*, a jellynose fish found in deep water in the east-central Pacific Ocean. (*Photo © John E. Randall*)**

Athalamida

An order of Granuloreticulosia in which the naked amebas form branched, threadlike interconnected pseudopodia (reticulopodia). Species are

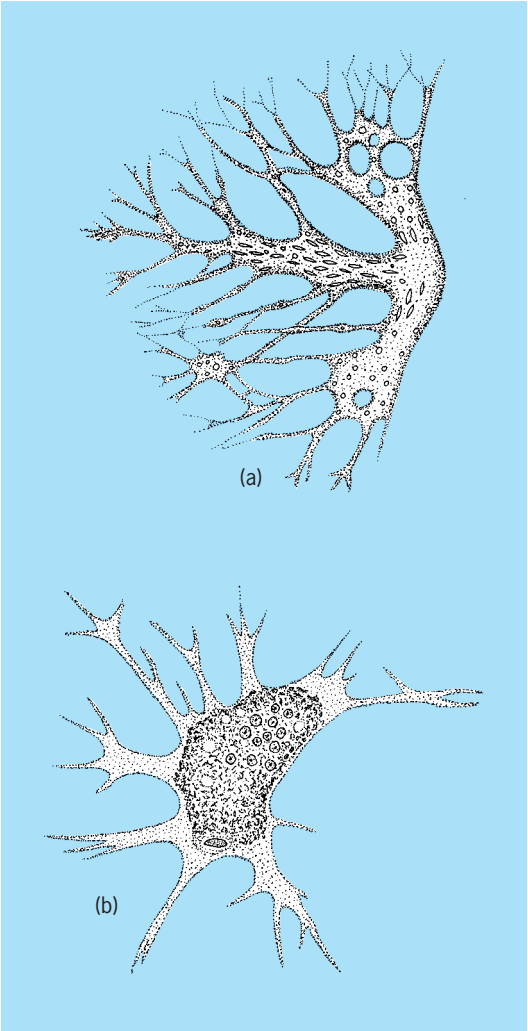

Athalamids. (*a*) *Biomyxa vagans*. (*b*) *Arachnula impatiens*. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)

known from fresh, salt, and brackish water. General characteristics are difficult to select, and the order may lack the taxonomic stability to survive accumulation of more nearly adequate information. Heterogeneity may extend even to a genus; described species of *Biomyza* differ appreciably in morphology and exhibit, for example, uninucleate and multinucleate condition in different species (although these may represent young and mature stages in life cycles). The granular cytoplasm in *Biomyza* (**illus.** *a*) shows no clear separation into ectoplasm and endoplasm. In addition to *Biomyza*, the genera *Arachnula* (illus. *b*), *Gymnophrys*, and *Pontomyxa* have been assigned to this order. *See* GRANULORETICULOSIA; PROTOZOA; RHIZOPODEA; SARCODINA; SARCOMASTIGOPHORA.

Richard P. Hall

Atheriniformes

An order of actinopterygian (ray-finned) fishes that comprise, with the Beloniformes and Cyprinodontiformes, the series Artherinomorpha. Atherinimorphs are more or less intermediate between typical soft-rayed and spiny-rayed fishes, but are not thought to be the ancestry of the latter. The order Atheriniformes, consisting of about 285 species in six families, are relatively small fishes, usually with an elongate body; opercular and preopercular margins lacking spines or serrations; lateral line absent or nearly so; usually two dorsal fins, with the first, if present, having weak spines; an anal fin larger than the soft dorsal fin and usually preceded by one spine; and usually abdominal pelvic fins, but subthoracic or absent in some (see family Phallostethidae below).

Habitats and distribution. Atheriniformes are represented by marine, fresh-water, and brackish water species. The marine species occupy continental and island shores of the tropical and temperate seas of the world. The fresh-water and brackish water species are found in northern and eastern Australia, New Guinea, the Philippines, eastern Indonesia, Southeast Asia, Malaya Peninsula, Madagascar, Africa, and North and South America.

The families Atherinidae and Atherionidae (Old World silversides) and the similar Atherinopsidae (New World silversides) share, in addition to the characters of the order, a slender, usually moderately compressed body; small terminal mouth; two widely separated dorsal fins; pectoral fins high on body and usually abdominal; scales large, usually cycloid, sometimes ctenoid; no lateral line; and broad silvery lateral band (black in preserved specimens). Atherinopsidae is the principal family of the western hemisphere. Fresh waters of Mexico have a rich silverside fauna: 39 species, primarily in two genera, *Atherinella* and *Chirostoma*. By contrast, all of the United States has only three species, which are limited to fresh water. The most widespread is the brook silverside (*Labidesthes sicculus*), ranging from the St. Lawrence River system south throughout much of the Mississippi Basin to the Gulf States. All in all, the family Atherinopsidae is represented in North America (including all of Mexico) by 11 genera and 56 species of which only 15 are limited to the marine environment, 10 from the Pacific, and 5 from the Atlantic. Two species of Old World silversides (Atherinidae) extend their range to the western North Atlantic, the hardhead silverside and reef silverside.

Another large family of atheriniforms is the Melanotaeniidae (rainbowfishes and blue eyes) consisting of 68 species from the fresh waters of northern and eastern Australia, New Guinea, and parts of eastern Indonesia. The body of the rainbow fishes is compressed; the dorsal fins are narrowly separated; the lateral line system is absent or weakly developed; the entire length of innermost pelvic fin ray is attached to the abdomen by a membrane; and, unlike most atheriniforms, they are colorful. The blue eyes are morphologically similar but differ in some features unique in the order; for example, the infraorbital series is reduced to a single bone (the lacrimal), and the mesethmoid is absent.

Reproduction and development. Reproduction in the brook silverside, as with some other atheriniforms, includes internal fertilization but without gross modification of fins and other skeletal structures to form clasping and copulatory organs; however, such organs are highly developed in the Phallostethidae, a family of 19 species of small fishes in Southeast Asia from the Philippines to Thailand. In this case, a unique copulatory structure of males called the pariapium is used to hold the female during mating and aids in transferring sperm packets. Derived from the pelvic skeleton, the pariapium is a complicated muscular and bony organ in the throat region that contains ducts from the urogenital system, as well as the terminal part of the intestine, all of which have migrated forward to below the pectoral fins in both sexes. The pelvic skeleton is absent in females.

Although fertilization is internal in both the brook silverside and the phallostethids, development is oviparous, with the female laying fertilized eggs. This reproductive strategy does not have all the advantages of viviparous development, as seen in some of the atheriniforms' near relatives in the orders Beloniformes and Cyprinodontiformes; however, fertilized eggs may be held and deposited at a time conducive to successful development, thus enhancing the reproductive potential of the species, especially one like the brook silverside, whose life cycle includes only one reproductive season. *See* ACTINOPTERYGII; CYPRINODONTIFORMES; TELEOSTEI.

Herbert Boschung

Bibliography. B. S. Dyer, Atherinopsidae (Neotropical silversides), pp. 515–525 in R. E. Reis et al. (eds.), *Checklist of the Freshwater Fishes of South and Central America*, EDIPUCRS, Porto Alegre, Brazil, 2003; B. S. Dyer and B. Chernoff, Phylogenetic relationships among atheriniform fishes (Teleostei: Atherinomorpha), *Zool. J. Linn. Soc.*, 117:1–69, 1996; R. J. Lavenberg and M. Chernoff, Atherinidae: Pejerreyes, pp. 889–901 in W. Fischer et al. (eds.), *Guia FAO para Identification de Especies para lo*

Fines de la Pesca, Pacifico Centro-Oriental, 3 vols., FAO, Rome, 1995; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994; L. R. Parenti, A phylogenetic revision of the pallostethid fishes (Atherinomorpha, Phallostethidae), *Proc. Calif. Acad. Sci.*, 46(11):243–377, 1989; L. R. Parenti, Relationships of atherinomorph fishes (Teleostei), *Bull. Mar. Sci.*, 52(1):170–196, 1993; D. E. Rosen, The relationships and taxonomic position of the halfbeaks, killifishes, silversides, and their relatives, *Bull. Amer. Mus. Nat. Hist.*, vol. 127, no. 5, 1964.

Atlantic Ocean

The large body of seawater separating the continents of North and South America in the west from Europe and Africa in the east and extending south from the Arctic Ocean to the continent of Antarctica. The Atlantic is the second-largest ocean water body, and in area it covers nearly one-fifth of the Earth's surface. It receives the fresh-water runoff from a continental drainage area approximately four times larger than that draining into either the Pacific or Indian oceans. The two major divisions, North and South Atlantic oceans, have the Equator as the common boundary. The North Atlantic, because of projecting land areas and island arcs, has numerous subdivisions. These include three large mediterranean-type seas, the Mediterranean Sea, the Gulf of Mexico plus Caribbean Sea, and the Arctic Ocean; two small mediterranean-type seas, the Baltic Sea and Hudson Bay; and four marginal seas, the North Sea, English Channel, Irish Sea, and Gulf of St. Lawrence. *See* ARCTIC OCEAN; BALTIC SEA; EQUATOR; GULF OF MEXICO.

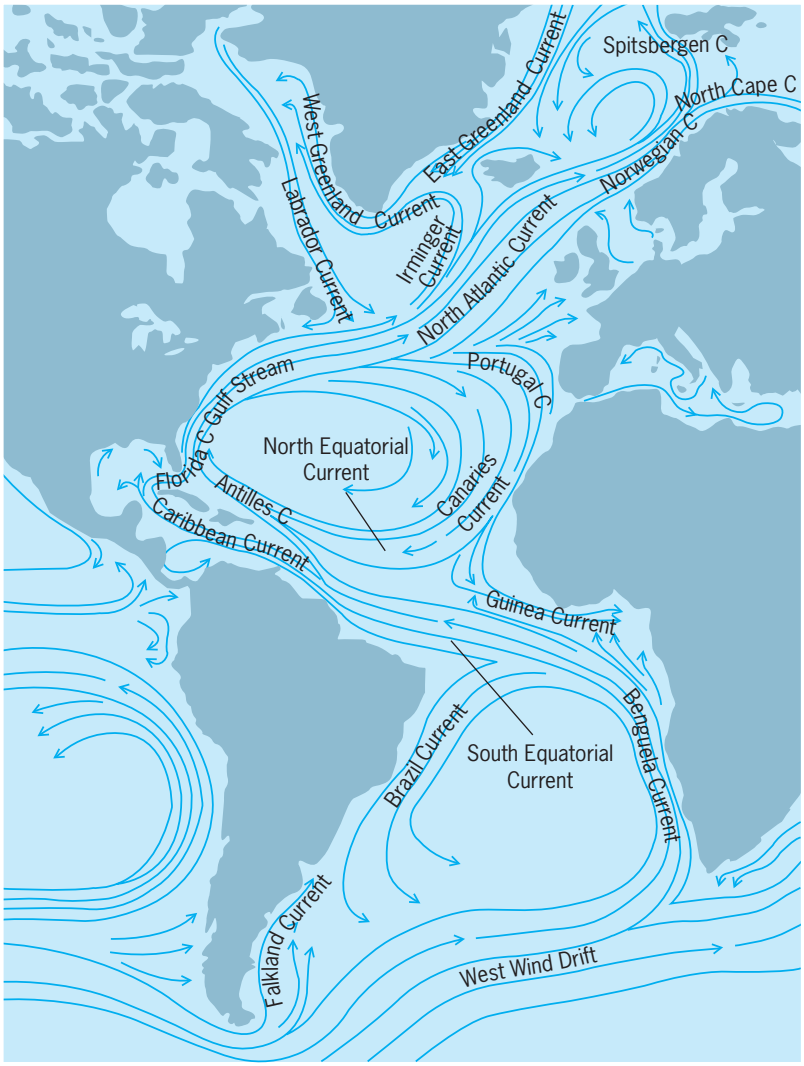

Fig. 1. Surface currents of the Atlantic Ocean. (*After J. Bartholomew, Advanced Atlas of Modern Geography, 3d ed., McGraw-Hill, 1987*)

Surface currents. Under the direction of M. F. Maury, the U.S. Naval Oceanographic Office began compiling wind and current measurements in the mid-1800s from logs of U.S. Navy ships as well as domestic and foreign merchant ships. The data were used to develop the first pilot charts and sailing directions. These historical ship drifts and pilot charts have provided most of what is known about the large-scale surface velocity patterns in the ocean. Modern measurements consisting of satellite-tracked surface drifters, moored current meters, and shipboard profiles have filled in many details of surface currents, including some of their variations in time. *See* INSTRUMENTED BUOYS.

Surface currents in the Atlantic Ocean are largely driven by the prevailing surface winds and flow in much the same direction as the winds (**Fig. 1**). Both the winds and currents circulate in large anticyclonic gyres centered at midlatitudes in the subtropical North and South Atlantic oceans. The currents tend to be much swifter on the western side of the gyres (Gulf Stream, Brazil Current) than on the east, because of the shape of the Earth and its rotation (Coriolis acceleration). A net northward flux of upper-layer water moves through the Atlantic, crossing the Equator near the western boundary. A compensatory southward flow of colder, denser water occurs at depth. The thermohaline circulation, which results in a net flux of heat from the South Atlantic to the North Atlantic, is driven by water mass density differences caused by geographical variations of temperature and salinity. *See* SEAWATER.

North Atlantic. The greater part of the water transported by the North Equatorial Current enters the Caribbean as the Caribbean Current and exits the Gulf of Mexico as the Florida Current, which continues northward along the North American coast as the Gulf Stream. Roughly half of the water flowing in the Caribbean Current and Florida Current enters from the North Equatorial Current; the other half originates in the South Atlantic and flows northward in the South Equatorial, North Brazil, and Guyana currents. *See* GULF STREAM.

South of the Grand Banks, the Gulf Stream divides into several branches. The North Atlantic Current flows across the Mid-Atlantic Ridge and forms several offshoots of relatively warm, saline waters that continue flowing in a northeastward direction. One of these, the Irminger Current, reaches Iceland.

Another flows across the Greenland-Scotland Ridge between the Faeroes and the Shetland Islands and thence along the Norwegian coast as the Norwegian Current. Part of this flow continues into the Barents Sea as the North Cape Current, and part of it reaches the Arctic Ocean as the West Spitsbergen Current. The transport of cold water southward in the East Greenland and Labrador currents partially compensates for the transport of relatively warmer water northward into the Arctic. The Irminger, West Greenland, and Labrador currents circulate around a cyclonic gyre centered in the Labrador Basin. Other branches of the North Atlantic Current flow southward as the Portugal and Canaries currents. These join the North Equatorial Current to complete the circuit of the North Atlantic.

Between the North and South Equatorial currents is the North Equatorial Countercurrent, which flows eastward across the Atlantic between 5°N and 10°N and into the Guinea Current. The Countercurrent is strongest during the second half of the year and disappears in the western Atlantic during the first half. The seasonal variation of the equatorial currents is largely due to the seasonal migration of the trade winds north and south across the Equator. Another eastward countercurrent that stretches across the Atlantic, the Equatorial Undercurrent, is located just below the surface layer centered on the Equator. This current is generated by the trade winds piling up water in the western Atlantic. *See* MARITIME METEOROLOGY; TROPICAL METEOROLOGY.

South Atlantic. The currents in the South Atlantic are, in many respects, the counterparts of those in the North Atlantic, for example the Brazil Current and the Gulf Stream, the Benguela and Canaries currents, and the Falkland and the Labrador currents. The Circumpolar Current or West Wind Drift flows eastward through the South Atlantic. Between the Circumpolar Current and Antarctica is located the Weddell Sea cyclonic gyre.

Ocean eddies. The Gulf Stream and other ocean currents are variable in position and time. Much of this variability is due to ocean eddies that are typically a few hundred kilometers in extent (**Fig. 2**). The eddies appear to be generated by instabilities of the swiftest currents and have velocities similar to them. Eddies are thought to be important to the dynamics of ocean circulation, and can transport water properties over long distances and times. Some of the longer-lived eddies survive several years, and change position by a thousand kilometers. In charts which show the general circulation, the eddies have been ignored (Fig. 1).

Surface temperature, salinity, and density. The sea surface temperature pattern is determined by geographical and seasonal variations of heating and cooling plus the advection of water by ocean currents and mixing (**Fig. 3**). Surface water is warmest in the equatorial band, where temperatures of 82°F (28°C) are found. Warm tropical water is carried northward in the Gulf Stream, where during winter large amounts of heat are released to the atmosphere. The northward heat flux in the Atlantic by ocean currents is thought to be about equal to that carried by the atmosphere.

Surface salinity is determined by geographical and seasonal variations in precipitation and evaporation and also advection and mixing. Highest salinities (greater than 37.0‰) are found in the subtropical gyres, where evaporation exceeds precipitation. Surface salinity is low (less than 35‰) in the doldrums, where heavy precipitation occurs. Salinity values below 30‰ occur where currents, such as the East Greenland, West Greenland, and Labrador, transport melting ice. Offshore of the major rivers such as the Amazon, which enters the Atlantic near the Equator, the surface salinity can be reduced considerably below usual oceanic values.

The warm and salty surface water of the subtropical gyres is less dense than the much colder and fresher water observed in the far north and south. The densest surface water in the open Atlantic is found during late winter in the Labrador Basin and Weddell Sea, sites of deep- and bottom-water formation.

Deep circulation. Subsurface water masses of the Atlantic originate from (1) water forced down from the surface in regions of convergence, (2) wintertime cooling and convection, and (3) inflows from other oceans and seas. The surface water in certain areas in the far north and south takes on a particularly high density in winter under the influence of climatic conditions. These water masses sink to a depth where their density is similar to that of the surrounding water, and then spread out at that level. At the same time the newly formed water mixes with the surrounding waters, modifying its characteristics. Overflows into the open North Atlantic from the Norwegian and Greenland seas and the Mediterranean Sea descend and mix with surrounding water until they, too, reach neutral buoyancy and spread out from their source regions as distinct layers. In addition, water from the Pacific enters the South Atlantic in the Circumpolar Current, and water from the Indian Ocean enters the South Atlantic around the southern tip of Africa as filaments and eddies of the Agulhas Current. In this way a multistoried stratification arises.

Of the Atlantic water masses, the densest is the very cold Weddell Sea Water, which is formed near the coast of Antarctica and extends northward along the sea floor to around 10°S (**Fig. 4**). Above this is Circumpolar Water, which enters the South Atlantic through the Drake Passage and which is split into lower and upper parts by higher-salinity water flowing southward from the North Atlantic at depths of 6600–9900 ft (2000–3000 m). The lower Circumpolar Water crosses the Equator in the west and flows northward in the western North Atlantic as far as 40°N. It also enters the eastern Atlantic through the Romanche Fracture Zone near the Equator. In the north, cold dense water flows into the North Atlantic from the Norwegian and Greenland seas over the ridge connecting Greenland and Scotland to form the abyssal water of the northern North Atlantic. Above this layer Labrador Sea Water is

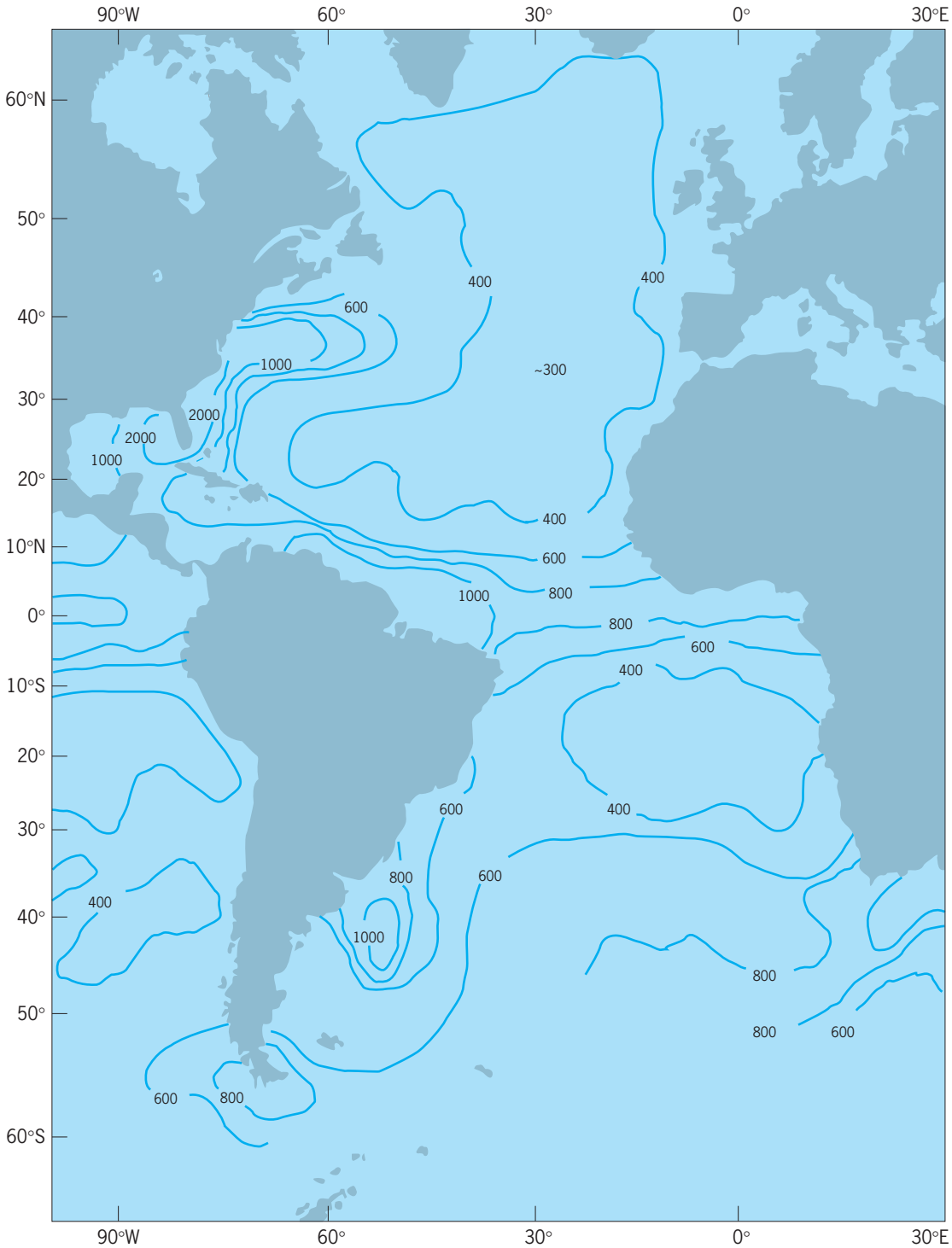

Fig. 2. Eddy kinetic energy per unit mass (cm^2/s^2) for the Atlantic Ocean based on historical ship drifts. Contour lines represent cm^2/s^2. 1 cm^2 = 0.155 $in.^2$ (*After K. Wyrtki, L. Magaard, and J. Hager, Eddy energy in the oceans, J. Geophys. Res., 81: 2641–2646, 1976*)

formed by deep convection in winter to roughly 5000 ft (1500 m).

Both Labrador Sea Water and the Norwegian Greenland Overflow Water are more saline at their sources than water originating in the South Atlantic, and their salinities are increased somewhat through mixing with the salty water of the Mediterranean Sea, which flows out of the Strait of Gibraltar. Mediterranean Water forms a huge lens of salty water which spreads westward across the Atlantic at depths of 3300–9900 ft (1000–3000 m) between 20° and 40°N with highest values in the eastern Atlantic near 3300 ft (1000 m). The combination of the Overflow Water, Labrador Sea Water, and Mediterranean Water forms North Atlantic Deep Water. This deep water mass flows southward along the western boundary

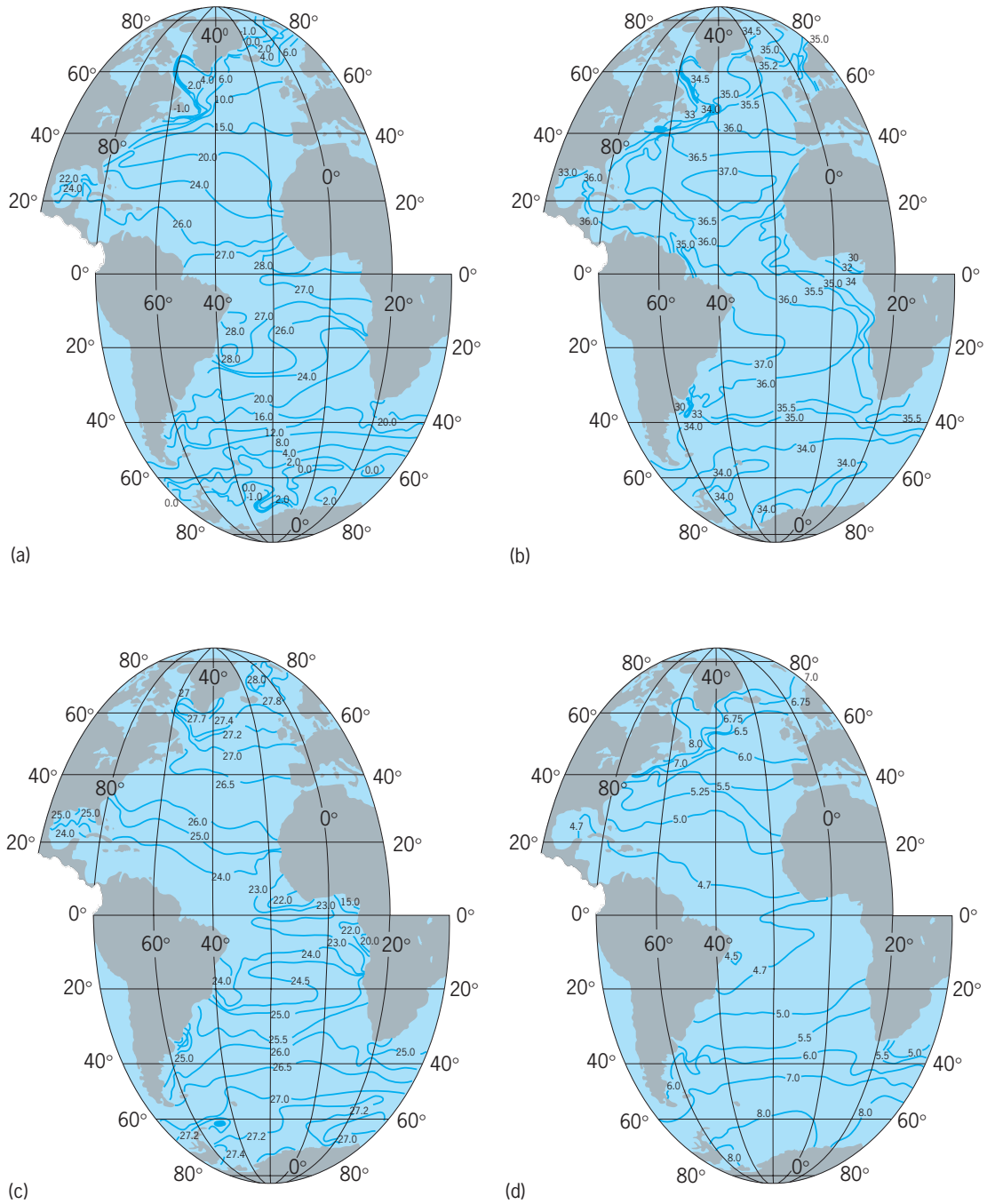

Fig. 3. Sea surface patterns. (*a*) Temperature; contour lines represent °C. (*b*) Salinity; contour lines represent a ratio based on relative conductivity. (*c*) Density; contour lines represent (kg/m^3) − 1000. (*d*) Oxygen; contour lines represent cm^3/liter. (*After J. L. Reid, Prog. Oceanog., 33:1–92, 1994*)

of the Atlantic, crosses the Equator, and extends into the Circumpolar Current, splitting it into upper and lower parts.

Above the deep water in the South Atlantic is the upper Circumpolar Water, and above this is the Intermediate Water, which is low in salinity and high in oxygen. The Intermediate Water spreads northward into the North Atlantic at depths of around 2600 ft (800 m) to at least 20°N, and even farther in the Gulf Stream. A warm upper layer overlies the Intermediate Water between the oceanic polar fronts. Temperatures in these waters are greater than 46°F (8°C).

Compared with that of the Indian and Pacific oceans, the deep circulation in the Atlantic is very vigorous, a result of deep water formation and strong thermohaline flows. The fastest speeds in the deep ocean reaching 50 cm/s (around 1 knot) are (1) underneath the major western boundary currents such as the Gulf Stream and Brazil Current; (2) along the western boundary, where the deep and abyssal meridional flows are concentrated; and (3) at the

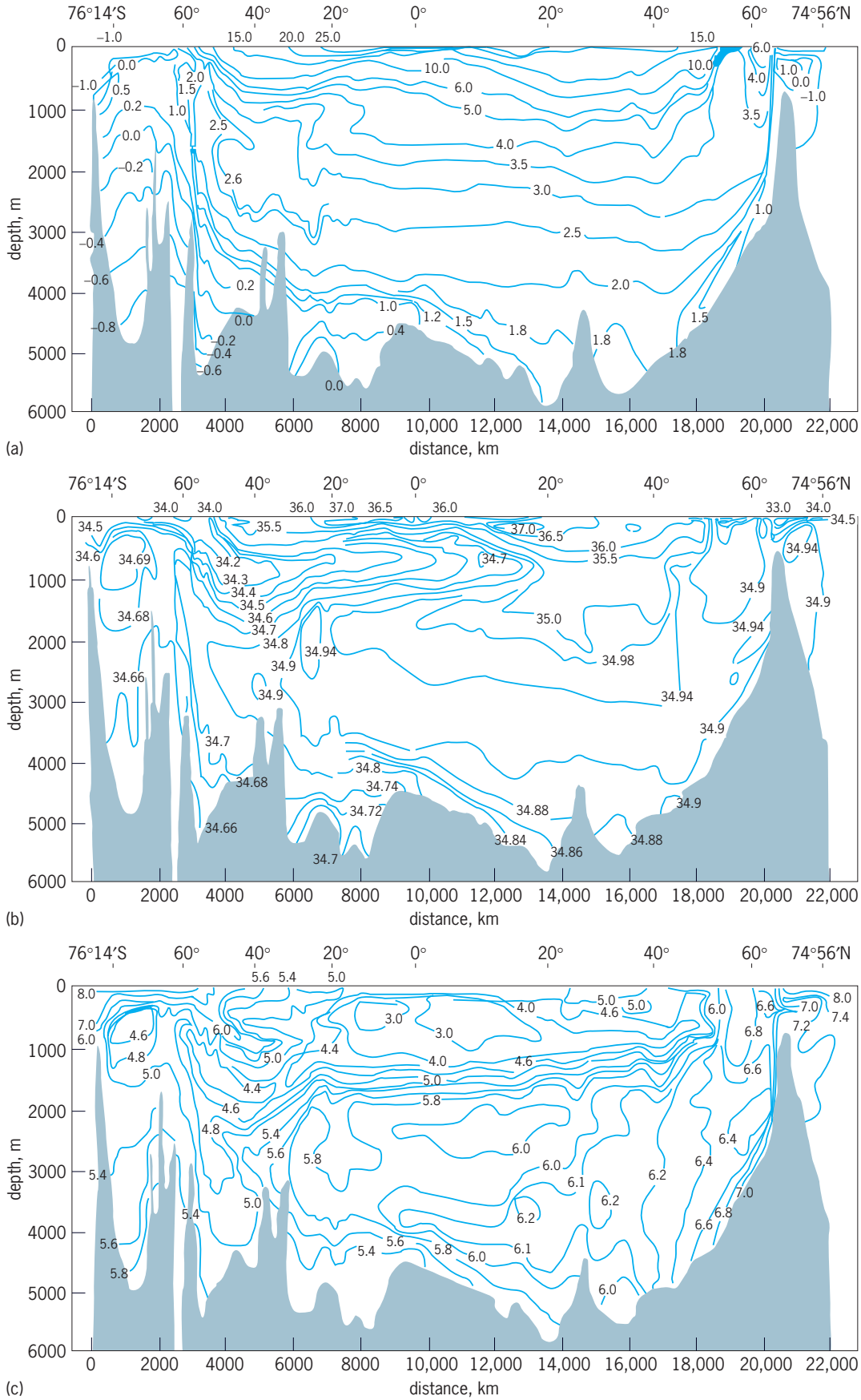

Fig. 4. North-south vertical sections of (*a*) potential temperature (°C) (*b*) salinity; and (*c*) oxygen (ml/liter) extending from the Weddell Sea at 76°S northward through the deeper part of the western basins and the Denmark Strait into the Norwegian-Greenland Sea at 75°N (*after J. L. Reid, On the total geostrophic circulation of the North Atlantic Ocean: Flow patterns, tracers, and transports, Prog. Oceanog., 33:1–92, 1994*).

sites of overflows, where dense water cascades down to its depth of neutral buoyancy much like an undersea waterfall. A particularly vigorous and complicated circulation scheme exists where the southward-flowing deep western boundary current encounters the deep extension of the Gulf Stream. Deep western boundary currents appear to have local energetic recirculations flowing counter to and offshore of the boundary currents. *See* OCEAN CIRCULATION. Philip L. Richardson

Bottom topography. Work on the transoceanic telegraph cables begun in 1860 stimulated development of deep-sea soundings. Since 1922 use of the echo sounder has made it possible to obtain a great number of depth soundings. The mean depth of the Atlantic Ocean is 12,690 ft (3868 m), and its volume is 76,000,000 mi^3 (318,000,000 km^3). *See* ECHO SOUNDER; MARINE GEOLOGY.

Broad shelves with depths less than 660 ft (200 m) are found in the region of the North Sea and the British Isles, on the Grand Banks of Newfoundland, and off the coasts of northeastern South America and Patagonia. The Mid-Atlantic Ridge, which extends from the Arctic Ocean to 55°S, is less than 9900 ft (3000 m) beneath the surface and is characterized by a pronounced relief. It separates the east and west Atlantic troughs, both of which have relatively uniform relief. The east and west troughs are connected in the vicinity of the Equator by the Romanche Deep, the only deep submarine passage through the Mid-Atlantic Ridge, with a depth of 25,500 ft (7728 m). *See* MID-OCEANIC RIDGE.

Three marked east-west ridges—the Greenland-Scotland Ridge in the North Atlantic and the Walvis and Rio Grande ridges in the South Atlantic—and several less conspicuous east-west rises separate the two Atlantic troughs into a series of basins including the West European, Canary, and Angola in the eastern Atlantic and the North American, Brazilian, and Argentine basins in the western Atlantic. Only isolated seamounts (such as the Great Meteor, Altair, and Atlantis seamounts) rise from the floor of the deep basins. Greatest depths occur in the narrow trenches along island arcs: 30,240 ft (9219 m) in the Puerto Rico Trench and 27,110 ft (8264 m) in the South Sandwich Trench. *See* DEEP-SEA TRENCH; SEAMOUNT AND GUYOT.

Islands. The islands of Jan Mayen, Iceland, St. Paul, Ascension, St. Helena, Tristan da Cunha, Gough, and Bouvet are parts of the Mid-Atlantic Ridge and are of purely volcanic origin. Other islands of volcanic origin but which lie outside the Mid-Atlantic Ridge are the Faeroes, Madeira, Fernando Poo, Príncipe, São Tomé, Annobón, Fernando Noronha, Trinidad, and the South Sandwich Islands. The Azores, Canary Islands, Cape Verde Islands, and Lesser Antilles are of predominantly volcanic origin. The Bermudas are the northernmost coral reefs. They rise from an old submarine volcanic cone. All the other islands in the Atlantic Ocean are continental in character, such as Spitsbergen, Bear Islands, the British Isles, Greater Antilles, Falkland Islands, and South Georgia. *See* ARCTIC AND SUBARCTIC ISLANDS; OCEANIC ISLANDS; WEST INDIES. Gunter O. Dietrich

Bibliography. G. Bearman, *Ocean Circulation*, 1989; K. O. Emery and E. Uchupi, *The Geology of the Atlantic Ocean*, 1984; S. G. Gorshkov, *World Ocean Atlas*, vol. 2: *Atlantic and Indian Ocean*, 1979; J. L. Reid, On the total geostrophic circulation of the South Atlantic Ocean: Flow patterns, tracers, and transports. *Prog. Oceanog.*, 23:149–244, 1989; J. L. Reid, On the total geostrophic circulation of the North Atlantic Ocean: Flow patterns, tracers, and transports, *Prog. Oceanog.*, 33:1–92, 1994; W. J. Schmitz, Jr., and M. S. McCartney, On the North Atlantic circulation, *Rev. Geophys.*, 31(1):29–49, 1993; B. A. Warren and C. Wunsch (eds.), *Evolution of Physical Oceanography: Scientific Surveys in Honor of Henry Stommel*, 1981.

Atmosphere

A gaseous layer that envelops the Earth and most other planets in the solar system. Earth, Venus, Mars, Jupiter, Saturn, Uranus, Neptune, and Titan (Saturn's largest satellite) are all known to possess substantial atmospheres that are held by the force of gravity. The structure and properties of the various atmospheres are determined by the interplay of physical and chemical processes. Structural features of Earth's atmosphere detailed below can often be identified in the atmospheres of other planetary bodies. *See* PLANETARY PHYSICS.

Composition. The composition of the Earth's atmosphere is primarily nitrogen (N_2), oxygen (O_2), and argon (Ar) [see **table**]. The concentration of water vapor (H_2O) is highly variable, especially near the surface, where volume fractions can vary from nearly 0% to as high as 4% in the tropics. There are many minor constituents or trace gases, such as neon (Ne), helium (He), krypton (Kr), and xenon (Xe), that are inert, and active species, such as carbon dioxide (CO_2), methane (CH_4), hydrogen (H_2), nitrous oxide (NO), carbon monoxide (CO), ozone (O_3), and sulfur dioxide (SO_2), that play an important role in radiative and biological processes.

In addition to the gaseous component, the atmosphere suspends many solid and liquid particles. Aerosols are particulates usually less than 5 micrometers in diameter that are created by gas-to-particle reactions or are lifted from the surface by the wind. A portion of these aerosols can become centers of condensation or deposition in the growth of water and ice clouds. Cloud droplets and ice crystals are made primarily of water with some trace amounts of particles and dissolved gases. Their diameters range from a few micrometers to about 100 μm. Water or ice particles larger than about 100 μm begin to fall because of gravity and may result in precipitation at the surface. *See* AEROSOL; CLOUD PHYSICS; PRECIPITATION (METEROLOGY).

One of the remarkable properties of the Earth's atmosphere is the large amount of molecular oxygen in the presence of gases, such as nitrogen, methane, water vapor, hydrogen, and others, that are capable of being oxidized. The atmosphere is in a highly oxidizing state that is far from chemical equilibrium.

Composition of the atmosphere*

Molecule	Fraction volume near surface	Vertical distribution
Major constituents		
N_2	7.8084×10^{-1}	Mixed in homosphere; photochemical dissociation high in thermosphere
O_2	2.0946×10^{-1}	Mixed in homosphere; photochemically dissociated in thermosphere, with some dissociation in mesosphere and stratosphere
Ar	9.34×10^{-3}	Mixed in homosphere with diffusive separation increasing above
Important radiative constituents		
CO_2	3.5×10^{-4}	Mixed in homosphere; photochemical dissociation in thermosphere
H_2O	Highly variable	Forms clouds in troposphere; trace amounts in stratosphere; photochemical dissociation above mesosphere
O_3	Variable	Small amounts, 10^{-8}, in troposphere; important layer, 10^{-6} to 10^{-5}, in stratosphere; dissociated above
Other constituents		
Ne	1.82×10^{-5}	
He	5.24×10^{-6}	Mixed in homosphere with diffusive separation increasing above
Kr	1.14×10^{-6}	
CH_4	1.15×10^{-6}	Mixed in troposphere; dissociated in upper stratosphere and above
H_2	5×10^{-7}	Mixed in homosphere; product of H_2O photochemical reactions in lower thermosphere, and dissociated above
NO	$\sim 10^{-8}$	Photochemically produced in stratosphere and mesosphere

*Other gases, for example, CO, N_2O, NO_2, and many by-products of atmospheric pollution also exist in small amounts.

This is in sharp contrast to the atmospheres of Venus and Mars, the planets closest to the Earth, which are composed almost entirely of the more oxidized state, carbon dioxide. The chemical disequilibrium on the Earth is maintained by a continuous source of reactive gases derived from biological processes. Life plays a vital role in maintaining the present atmospheric composition. *See* ATMOSPHERIC CHEMISTRY; MARS; VENUS.

Vertical structure. The total mass of the Earth's atmosphere is about 5.8×10^{15} tons (5.3×10^{15} metric tons). The vertical distribution of gaseous mass is maintained by a balance between the downward force of gravity and the upward pressure-gradient force. The balance is known as the hydrostatic balance or the barometric law. Hence, the declining atmospheric pressure that is measured while ascending in the atmosphere is a result of gravity. The globally averaged pressure at mean sea level is 1013.25 millibars (101,325 pascals).

Below about 60 mi (100 km) in altitude, the atmosphere's composition of major constituents is very uniform. This region is known as the homosphere to distinguish it from the heterosphere above 60 mi (100 km), where the relative amounts of the major constituents change with height. In the homosphere, there are sufficient atmospheric motions and a short enough molecular free path to maintain uniformity in composition. Above the boundary between the homosphere and the heterosphere, known as the homopause or turbopause, the mean free path of the individual molecules becomes long enough that gravity is able to partially separate the lighter molecules from the heavier ones. The mean free path is the average distance that a particle will travel before encountering a collision. Hence, the average molecular weight of the heterosphere decreases with height as the lighter atoms dominate the composition.

Radiative transfer. The vertical structure of the atmosphere (**Fig. 1**) is in large part determined by the transfer properties of the solar and terrestrial radiation streams. The energy of the smallest unit of radiation, the photon, is directly proportional to its frequency. The type of interaction that occurs between photons and the atmosphere depends on the energy of the photons. *See* PHOTON.

The most energetic of the photons are x-rays and extreme ultraviolet radiation of the electromagnetic spectrum, which are capable of dissociating and ionizing the gaseous molecules. The less energetic near-ultraviolet photons are able to excite molecules and atoms into higher electronic levels. As a result, most of the ultraviolet and x-ray radiation is attenuated by the upper atmosphere. A cloudless atmosphere is relatively transparent to visible light, where most of the solar energy resides. At the opposite end of the spectrum, toward the lower frequencies of radiation, is the infrared which is capable of inducing various vibrational and rotational motions in triatomic and polyatomic molecules.

In order to maintain an energy balance, the Earth must emit to space about the same amount of radiation as it absorbs from the Sun (**Fig. 2**). The terrestrial radiation occurs in the infrared part of the spectrum and is strongly affected by water vapor, clouds, carbon dioxide, and ozone and other trace gases. The ability of these gases to absorb and emit in the infrared allows them to effectively trap some of the outgoing radiation that is emitted by the surface, creating the so-called greenhouse effect which is responsible for making the average global surface temperature about 59°F (15°C). The Earth's carbon dioxide concentration is increasing as a result of fossil-fuel burning and deforestation. Since carbon dioxide absorbs and emits infrared radiation, the greenhouse effect is being enhanced, and this is responsible for much of the currently observed global warming which may become a major environmental problem in this century. *See* GREENHOUSE EFFECT; INSOLATION; TERRESTRIAL RADIATION.

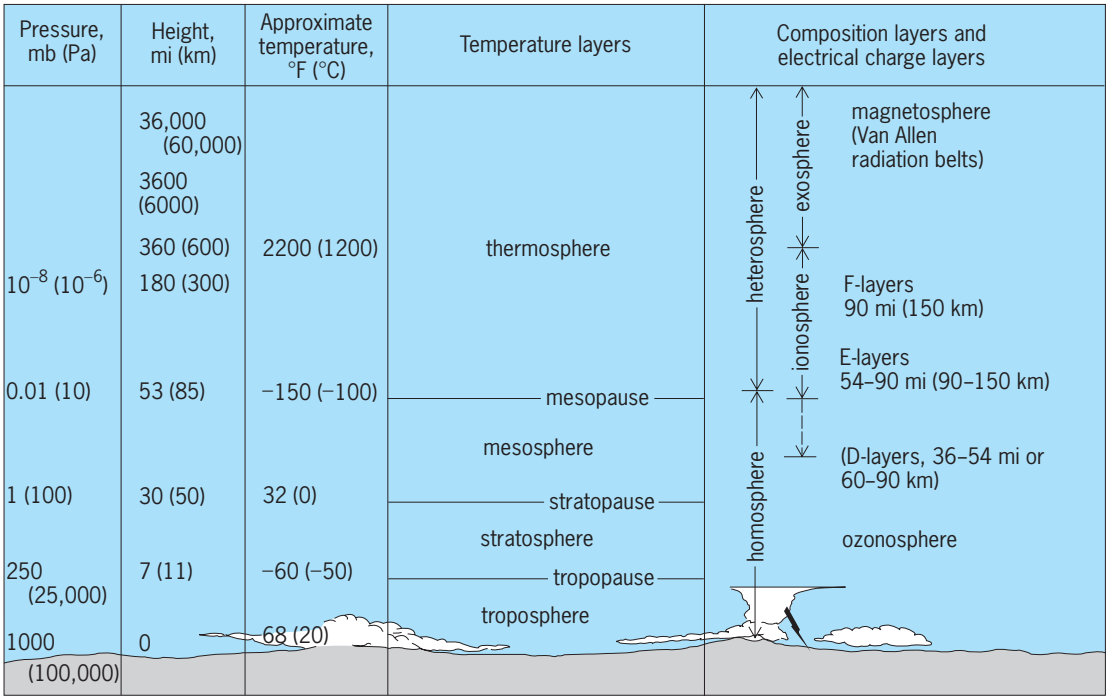

Fig. 1. Layers of the atmosphere, shown in terms of pressure, height, temperature, and compositional properties. (*After M. Neiburger, J. G. Edinger, and W. D. Bonner, Understanding Our Atmospheric Environment, W. H. Freeman, 1973*)

Troposphere. The atmospheric layer that extends from the surface to about 7 mi (11 km) is called the troposphere. The tropopause, which is the top of the troposphere, has an average altitude that varies from about 11 mi (18 km) near the Equator to about 5 mi (8 km) near the Poles. The actual tropopause height varies considerably on time scales from a few days to an entire year. *See* TROPOPAUSE; TROPOSPHERE.

The troposphere contains about 80% of the atmospheric mass and exhibits most of the day-to-day weather fluctuations that are observed from the ground. Temperatures generally decrease with increasing altitude at an average lapse rate of about 17°F/mi (6°C/km), although this rate varies considerably, depending on time and location (**Fig. 3**).

On average, about half of the total solar radiative flux incident at the top of the atmosphere is absorbed by the Earth's surface. This raises the surface temperature sufficiently to induce atmospheric circulations in the troposphere that redistribute the heat

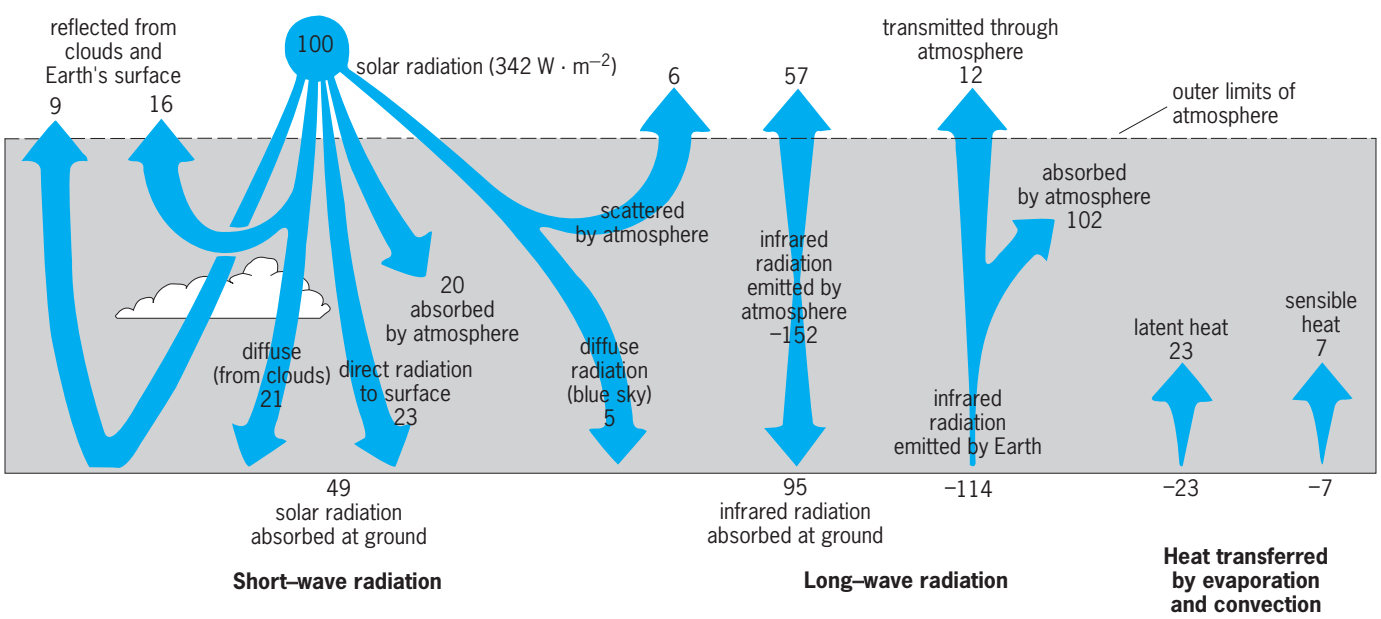

Fig. 2. Annually and globally averaged radiation energy budget of the Earth's atmosphere expressed as percentages of the total incoming solar radiation; the numerical values sum to a net balance for the surface, atmosphere, and space separately. (*After J. T. Kiehl and K. E. Trenberth, Earth's annual global mean energy budget, Bull. Amer. Meteorol. Soc., 78:197, 1997*)

-s

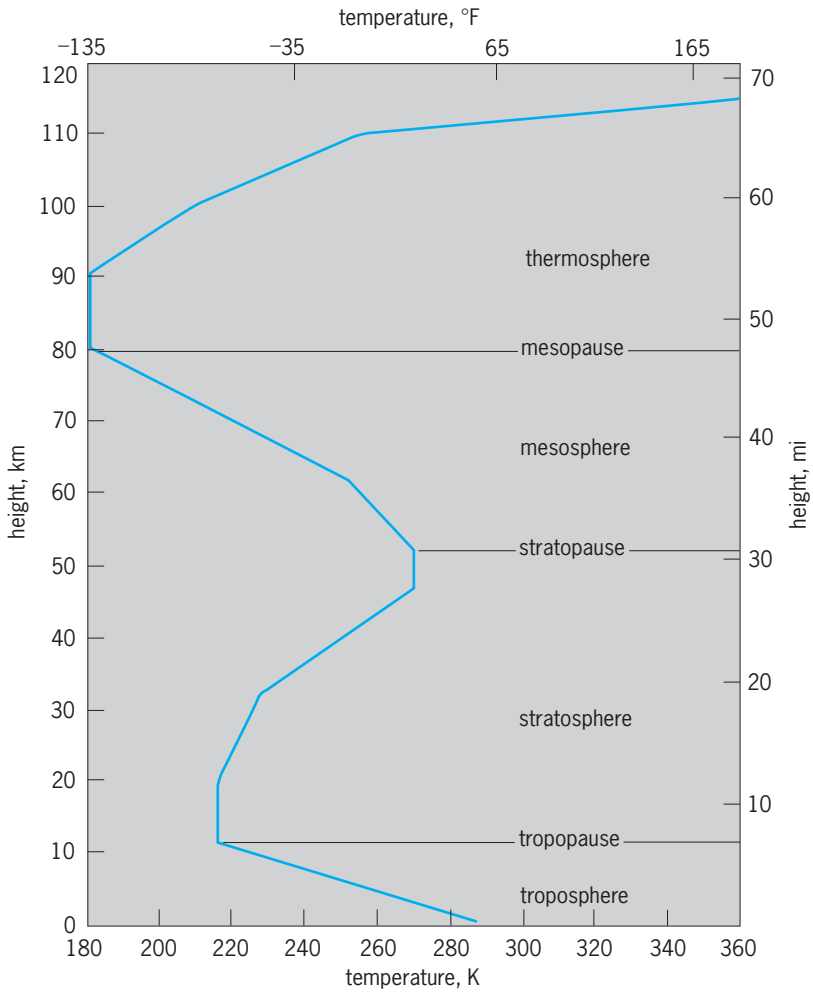

Fig. 3. Temperature of the United States standard atmosphere plotted as a function of height and showing the various thermal layers. The standard atmosphere is intended to be a global and annual representative average. (*After M. Neiburger, J. G. Edinger, and W. D. Bonner, Understanding Our Atmospheric Environment, W. H. Freeman, 1973*)

both vertically and horizontally throughout the troposphere. As an air parcel that is heated from contact with the surface is forced to rise, it becomes surrounded by lower pressure. The air expands and must cool in order to conserve total energy, resulting in the observed decrease of temperature with height.

In order to transfer heat from regions with a radiational surplus, such as the tropics, to regions with a radiational deficit, such as the high latitudes, the atmosphere goes into motion and produces the general circulation relative to the Earth's surface. The large-scale wind patterns are dominated by a balance between the horizontal pressure-gradient force and the Coriolis force. The horizontal pressure gradients form when uneven solar heating creates horizontal temperature gradients. The Coriolis force is a fictitious force that arises because the rotating Earth is an accelerating body. The Coriolis force acts at right angles to the wind and is proportional to the wind speed. Winds that start moving from high to low pressure become deflected to the right in the Northern Hemisphere and to the left in the Southern Hemisphere, so that eventually the winds blow parallel to lines of constant pressure. This is called geostrophic balance and is responsible for the westerlies and the jet streams that are found in the mid and high latitudes and the easterlies in the low latitudes or tropics. *See* ATMOSPHERIC GENERAL CIRCULATION; CORIOLIS ACCELERATION; GEOSTROPHIC WIND; JET STREAM; WIND.

Superimposed on the background westerly and easterly winds are numerous types of disturbances or waves that give the weather its complex character and dominate the transport of energy and momentum in the troposphere. Examples of such disturbances include (1) the Hadley cell, a thermally direct average circulation that extends from latitude 30°S to 30°N and transports heat and momentum away from the Equator; (2) the barotropic disturbance, a wave instability that forms in the presence of a very strong horizontal wind shear; (3) the baroclinic disturbance, a wave instability that grows in the presence of a strong north-south temperature gradient and is responsible for the extratropical cyclones which we observe as low-pressure systems with rain or snow in the midlatitudes; (4) tropical cyclones such as hurricanes and typhoons, disturbances that convert the potential energy of the moist and warm sea surface air into kinetic energy; and (5) thunderstorms, deep convective disturbances that rapidly overturn localized regions of the atmosphere in which relatively cold, dry air overlays warm, moist air. *See* ATMOSPHERIC ELECTRICITY; DYNAMIC INSTABILITY; HURRICANE; THUNDERSTORM.

The winds not only transport heat in a form that can be felt directly, namely sensible heat, but also transport latent heat, which is a form of heat that is carried by the water vapor. Latent refers to the fact that the heat which is absorbed when water is evaporated into the atmosphere is released only when the water condenses to form clouds and precipitation in the atmosphere. The global hydrological cycle, which includes the clouds and precipitation, is a major contributor to the global redistribution of heat. *See* HYDROMETEOROLOGY.

The amount of water vapor present in a parcel of air can be measured by the partial vapor pressure that is exerted by those water molecules. Since the maximum water vapor pressure (saturation vapor pressure) decreases exponentially with decreasing temperature, most of the water vapor condenses out to form clouds as a moist air parcel ascends through the troposphere. It is sufficiently cold at the tropopause (−110 to −40°F or −80 to −40°C) that very little water is able to escape above it, an effect known as the cold trap. As a result of the cold trap, the air is very dry and usually cloud-free above the tropopause.

Stratosphere. This is the atmospheric layer that extends from the tropopause up to the stratopause at about 30 mi (50 km) above the surface. It is characterized by a nearly isothermal layer in the first 6 mi (10 km) overlaid by a layer in which the temperature increases with height to a maximum of about 32°F (0°C) at the stratopause. *See* STRATOSPHERE.

The reversal in the temperature lapse rate is a result of direct absorption of solar radiation, mainly by ozone and oxygen at the ultraviolet frequencies. Ozone (O_3), is a poisonous and reactive gas. Most of the ultraviolet light is attenuated in the stratosphere by ozone and oxygen; hence the surface of the Earth is protected from receiving doses that are harmful to life. In spite of the great importance of stratospheric ozone, it is merely a trace gas with peak concentrations of about 10–15 parts per million by volume at an altitude of about 20 mi (33 km) in the tropics. Ozone concentrations are also influenced by stratospheric winds and hence are highly variable in space and time. Ozone photochemistry is highly complex and involves chemical cycles from the hydrogen, nitrogen, and chlorine groups. It is the chlorine chemistry that has received particular attention in light of the declining ozone concentrations over the Antarctic during spring in the Southern Hemisphere, creating the so-called ozone hole. The chlorofluorocarbons emitted by certain industrial activities have contributed to this decline. *See* PHOTOCHEMISTRY.

The reversal of the temperature lapse rate makes the stratosphere vertically stable. This stability limits the amount of vertical mixing and results in molecular residence times of many months to years. Another consequence of a stable stratosphere is that it acts as a lid on the troposphere, confining the strong vertical overturning and the surface-based weather phenomena of the troposphere. *See* WEATHER.

Occasionally a strong volcano injects large amounts of aerosols, commonly sulfate particles, into the stratosphere that may reside there for several years before they are removed by slow descent. While in the stratosphere, the sulfate particles can perturb the radiative balance by increasing the stratospheric temperature and cooling the troposphere. Significant events of this type were the El Chichón eruption in 1982 and the Pinatubo eruption in 1991. *See* VOLCANO.

Mesosphere. This is the atmospheric layer extending from the stratopause up to the mesopause at an altitude of about 53 mi (85 km). The mesosphere is characterized by temperatures decreasing with height at a rate of about 12°F/mi (4°C/km). Although the mesosphere has less vertical stability than the stratosphere, it is still more stable than the troposphere and does not experience rapid overturning. The coldest temperatures of the entire atmosphere are encountered at the mesopause, with values as low as −150°F (−100°C).

The temperature lapse rate found in the mesosphere is a result of the gradual weakening with height of the direct absorption of solar radiation by ozone. The radiative infrared cooling to space by the carbon dioxide molecules is responsible for the low temperatures near the mesopause. On rare occasions, the temperatures may be low enough for noctilucent clouds to form during the high-latitude summer. *See* MESOSPHERE; NOCTILUCENT CLOUDS.

Thermosphere. This part of the atmosphere is found above the mesopause. The thermosphere is characterized by rising temperatures with height up to an altitude of about 190 mi (300 km) and then is nearly isothermal above that. Although there is no clear upper limit to the thermosphere, it is convenient to consider it extending several thousand kilometers. Embedded within the thermosphere is the ionosphere, comprising those atmospheric layers in which the ionized molecules and atoms are dominant.

Molecular species dominate in the lower thermosphere, while atomic species are dominant above 190 mi (300 km). The distribution of the constituents is controlled by diffusive equilibrium in which the concentration of each constituent decreases exponentially with height according to its molecular weight. Hence the concentration of the heavier constituents such as nitrogen, oxygen, and carbon dioxide will decrease with height faster than the lighter constituents such as helium and hydrogen. At an altitude of 560 mi (900 km), helium becomes the dominant constituent, while hydrogen dominates above 1900 mi (3000 km).

Diffusion becomes the dominant mode of transport in the thermosphere, because the low air densities allow the mean free path of the molecules and atoms to become very large, about 3 ft at 60 mi (1 m at 100 km), and it increases rapidly with height. The thermosphere is exposed to the complete spectrum of radiation that is emitted by the Sun. This includes the extreme ultraviolet (wavelengths less than about 103 nanometers) and some x-ray radiation that is capable of dissociating and ionizing molecules. Although the amount of energy contained in these high-frequency (short-wavelength) radiation bands is small compared to the total solar flux in the visible band, they still dominate the radiative heating in the thermosphere. *See* SOLAR RADIATION.

Of the total absorbed solar energy in the thermosphere, about 35% goes into heating the neutral particles, raising the kinetic temperature, and about 45% is reradiated out of the thermosphere as ultraviolet airglow. Airglow occurs when molecules and atoms reemit temporarily absorbed single photons of solar radiation. The remainder of the incoming absorbed solar radiation (about 20%) is stored as chemical energy of oxygen atoms formed when an oxygen molecule is dissociated. This energy is released as heat when oxygen atoms recombine below 60 mi (90 km) in altitude. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; AIRGLOW.

In order to maintain thermal balance, the thermosphere must lose the heat it receives from the absorption of solar radiation. The emission of thermal, infrared radiation is weak, and so balance is achieved by thermal conduction of heat from the high temperatures of the upper thermosphere to the coldest temperatures at the mesopause. At the mesopause, there is sufficient carbon dioxide and ozone pressure to allow heat to be emitted radiatively to space. The reason that thermal conduction is such a good heat-transfer mechanism in the thermosphere, and not in the lower atmospheric layers, is the large mean free path of the atoms and molecules at these heights. *See* HEAT RADIATION.

The upper thermosphere, above 190 mi (300 km), is a very active region both thermally and dynamically; it exhibits large responses to changes in the solar radiative forcing and auroral particles. Temperatures can vary from about 600°F (300°C) at night to 2200°F (1200°C) during the day. A global circulation occurs in the thermosphere, with winds between 110 and 450 mi/h (50 and 200 m/s) blowing from the day to night side. This prevents the diurnal temperature change from becoming even larger. Also, auroral storms can induce a high- to low-latitude circulation in the thermosphere. *See* AURORA.

Ionosphere. This can be defined operationally as that part of the atmosphere that is sufficiently ionized to affect the propagation of radio waves. In the ionosphere, the dominant negative ion is the electron, and the main positive ions include O^+, NO^+, and O_2^+.

The ionosphere is classified into four subregions. The D region extends from 40 to 60 mi (60 to 90 km) and contains complex ionic chemistry; most of the ionization is caused by ultraviolet ionization of NO and by galactic cosmic rays. This region is responsible for the daytime absorption of radio waves, which prevents distant propagation of certain frequencies. The E region extends from 60 to 90 mi (90 to 150 km) and is caused primarily by the x-rays from the Sun. The F1 region from 90 to 125 mi (150 to 200 km) is caused by the extreme ultraviolet radiation from the Sun and disappears at night. The F2 region includes all the ionized particles above 125 mi (200 km), with the peak ion concentrations occurring near 190 mi (300 km). *See* COSMIC RAYS; IONOSPHERE.

Exosphere. The term exosphere is used to refer to the atmosphere above 300 mi (500 km), where the probability of interatomic collisions is so low that some of the atoms traveling upward with sufficient velocity can escape the Earth's gravitational field. The dominant escaping atom is hydrogen since it is the lightest constituent. Calculations of the thermal escape of hydrogen (also known as the Jeans escape) yield a value of about 3×10^8 atoms $\cdot$ $cm^{-2} \cdot s^{-1}$. This is a very small amount since at this rate less than 0.5% of the oceans would disappear over the current age of the Earth.

The main source of the escaping hydrogen is water vapor that becomes dissociated by ultraviolet radiation above the tropopause. The importance of the cold water trap at the tropopause is evident, as this limits the availability of water for dissociation and provides a limit on the hydrogen escape. It is speculated that Venus may have lost all its water because such a cold trap did not limit the hydrogen escape.

Magnetosphere. This is the region surrounding the Earth where the movement of ionized gases is dominated by the geomagnetic field. The lower boundary of the magnetosphere, which occurs at an altitude of nearly 75 mi (120 km), can be roughly defined as the height where there are enough neutral atoms that the ion-neutral particle collisions dominate the ion motion. The dynamics of the magnetosphere is dictated in part by its interaction with the plasma of ionized gases that blows away from the Sun, the solar wind. The solar wind interacts with the Earth's magnetic field and severely deforms it, producing a magnetosphere around the Earth. It extends about 40,000 mi (60,000 km) toward the Sun, but extends beyond the orbit of the Moon away from the Sun. *See* MAGNETOSPHERE; SOLAR WIND; VAN ALLEN RADIATION.

Vertical energy transport. The division of the Earth's atmosphere into layers is based primarily on the thermal and chemical properties of each layer. Even though physical properties are found that distinguish the layers from each other, it is important to realize that interactions of mass and energy do occur between the layers.

The most obvious energy transfer is the outgoing terrestrial infrared radiation that escapes from the Earth. However, atmospheric dynamicists have also been studying the upward transfer of energy and momentum by the vertical propagation of waves. As a wave propagates upward, it encounters lower air densities, and the wave amplitude grows to conserve the total energy flux. Eventually the wave amplitude becomes so large that the wave is said to break, not unlike ocean waves breaking on the shoreline, and it deposits its energy and momentum at that height. The dissipation that occurs when a wave breaks is achieved by molecular conduction, viscosity, and ion drag.

It is believed that wave breaking may play an important role in the heating and circulation patterns of certain upper atmospheric regions. The two main wave types are atmospheric tides and gravity waves. Atmospheric tides are initiated in the stratosphere when ozone is heated, and can produce a vertically propagating wave. Gravity waves are oscillations produced by the stable buoyancy force; they also grow in amplitude as they move upward. *See* AERONOMY; ATMOSPHERIC TIDES; METEOROLOGY; UPPER-ATMOSPHERE DYNAMICS.

Glen B. Lesins

Bibliography. R. M. Goody, *Principles of Atmospheric Physics and Chemistry*, 1995; J. T. Houghton, *The Physics of Atmospheres*, 3d ed., 2002; J. M. Wallace and P. V. Hobbs, *Atmospheric Science: An Introductory Survey*, 1977; R. P. Wayne, *Chemistry of Atmospheres*, 3d ed., 2000.

Atmosphere, evolution of

Variation with time of the chemical composition and total weight of the Earth's atmosphere. The atmosphere is a most tenuous envelope; its mass is less than one-millionth that of the solid Earth; its density even at sea level is less than one-thousandth that of rocks, and virtually all of the atmosphere is below a height only one-hundredth of an earth radius above the surface of the Earth. But the atmosphere is taken so much for granted that one tends to be surprised at the thought that it has a history, that its chemical composition and total weight have varied through time. On reflection, it would, however, be odd to find that the atmosphere has not changed during the long years of the Earth's existence, and that its weight

and composition have not responded to the complicated series of events that have left such clear marks on the Earth's crust.

The study of atmospheric evolution is difficult, because there seems to be no way of obtaining reliable samples of the atmosphere that are more than about 400,000 years old. But the quest is not altogether hopeless. The origin of life, the continued existence of animals since at least 600 million years ago (Ma), the marks which interaction of the atmosphere with surface rocks have left on ancient sediments, and the nature of volcanic gases through geologic time all give clues about the chemical evolution of the atmosphere. While it is not possible to solve all the puzzles of the history of the Earth's atmosphere with these clues, the broad pattern of atmospheric evolution is emerging; and a more complete understanding is within reach.

It may be wise to start by defining the problem rather precisely. Today the atmosphere contains a small number of major components and a very large number of minor and trace components. Each component exerts a pressure, which is essentially constant for some components at sea level and which is variable in time and space for other components (**Tables 1** and **2**). The problem of the chemical evolution of the Earth's atmosphere can be considered solved when an accurate plot can be made of the variation with time of the pressure of these components and others that may have existed in the atmosphere in the past. This is obviously a large undertaking, but it is made somewhat less so by the mutual exclusion of certain gases as major components of the atmosphere. At present, for instance, gases such as ammonia (NH_3), methane (CH_4), and hydrogen (H_2) can exist in the atmosphere only as trace components, because they are unstable in the presence of large quantities of oxygen. Conversely, oxygen could not have been a major component of an atmosphere in which ammonia, methane, and hydrogen were abundant. *See* ATMOSPHERIC CHEMISTRY.

TABLE 1. Essentially nonvariable constituents of today's air*

Constituent	Content, %
N_2	78.084
O_2	20.946
CO_2	0.036
A	0.934
Ne	18.18×10^{-6}
He^4	5.24×10^{-6}
He^3	6.55×10^{-12}
Kr	1.14×10^{-6}
Xe	0.087×10^{-6}
H_2	0.5×10^{-6}
N_2O	0.5×10^{-6}

*The classification of H_2 and N_2O as nonvariable constituents is uncertain.

TABLE 2. Variable constituents of today's dry air

Constituent	Content
O_3	0 to 0.07 ppm (summer) 0 to 0.02 ppm (winter)
SO_2	0 to 1 ppm
NO_2	0 to 0.02 ppm
CH_4	1 to 2 ppm
CH_2O	Uncertain
I_2	0 to 10^{-10} g · cm^{-3}
NaCl	Order of 10^{-10} g · cm^{-3}
NH_3	0 to trace
CO	0 to trace (0.8 cm atm)

Origin of free oxygen. Free oxygen is somewhat of an anomaly on the Earth. Most rocks that are more than a few feet below the Earth's surface are out of equilibrium with free oxygen and are oxidized in contact with the atmosphere. This is seen in the development of red hydrous ferric oxide minerals in soil zones above many rock types in temperate and tropical areas and in the presence of the red to reddish-brown sediments that are so common in the western United States.

The origin of oxygen in the Earth's atmosphere has been a source of continuing controversy for many decades. Of the two theories that have been dominant, the first proposes that atmospheric oxygen has been produced through geologic time by green plant photosynthesis, during which carbon is effectively separated from oxygen in carbon dioxide. In a very rough manner this can be written as reaction (1).

$$6CO_2 + 6H_2O \xrightarrow{\text{photo-synthesis}} \underset{\text{Carbohydrate}}{C_6H_{12}O_6} + 6O_2 \qquad (1)$$

This reaction runs in the opposite direction during the decay of plant material and the breathing of animals. It can be shown that nearly all of the oxygen produced by photosynthesis is consumed by respiration, but the small amount that is not lost in this manner could easily account for the present, rather large quantity of atmospheric oxygen.

The second theory proposes that ultraviolet light from the Sun decomposes water molecules in the upper atmosphere. Most of the water molecules recombine, but there is a finite possibility that some hydrogen atoms will escape from the Earth's atmosphere before recombination has taken place. Oxygen atoms, being 16 times as heavy as hydrogen atoms, escape much more slowly or not at all. The decomposition of water vapor followed by hydrogen escape is therefore a distinctly plausible manner of generating free atmospheric oxygen. The escape rate of hydrogen from the atmosphere is almost certainly rapid. The critical factor in determining escape rates is the temperature of the upper atmosphere, and this is now well known from rocket measurements. But a strongly limiting factor for oxygen production by this mechanism is the formation of ozone (O_3) in the stratosphere. Ozone absorbs ultraviolet light very readily, and forms a screen that prevents ultraviolet light from reaching the lower levels of the atmosphere, where water vapor is abundant. Most of the hydrogen that manages to reach the upper atmosphere today is a constituent not of water vapor but of biologically produced methane, and the rate of hydrogen escape from the Earth's atmosphere is controlled largely by the rate of upward transport of

this compound. Nevertheless, the hydrogen loss rate is currently a very small fraction of the oxygen production rate by photosynthesis. The oxygen content of the present atmosphere is therefore only slightly influenced by hydrogen loss from the upper atmosphere, and depends almost exclusively on the operation of the feedback system that links the rate of oxygen production during photosynthesis to the rate of oxygen use by weathering and the decay of organic matter. *See* PHOTOSYNTHESIS.

Biologic evidence. The proposition that the present high concentration of oxygen in the atmosphere is largely due to oxygen production during photosynthesis implies that oxygen was much less abundant prior to the existence of photosynthetic organisms. This view is in harmony with the requirement that free oxygen was absent from the atmosphere during the development of life. *See* PREBIOTIC ORGANIC SYNTHESIS.

It is likely that life evolved very early in Earth history. The oldest known unmetamorphosed sedimentary rocks, those found near North Pole, Australia, are about 3.5 billion years old and contain microscopic bits of carbon that are probably biologic in origin. They also have stromatolitic structures, which may be the work of cyanobacteria. There is little doubt, then, that life began more than 3.5 billion years ago (Ga), and there is reasonably compelling evidence that life began more than 3.8 Ga. It is not yet clear when organisms that were abl e to produce free oxygen developed, but evidence from the isotopic composition of carbon in sedimentary rocks suggests that this occurred before 2.7 Ga. *See* STROMATOLITE.

The biologic evidence for oxygen at levels approaching those of the present is still scant until the close of the Precambrian Era, some 543 Ma. The sequence of animals that developed during the latest part of the Precambrian Era and the beginning of the Phanerozoic Era strikingly parallels the zoning of animals in contemporary settings of progressively greater oxygen content. This correspondence suggests, but does not prove, that evolutionary events were related to changes in the level of atmospheric oxygen. *See* PRECAMBRIAN.

It has been proposed that the invasion of the land by plants and animals about 400 Ma occurred when the oxygen pressure had risen to about one-tenth of its present value. At this oxygen pressure the intensity of ultraviolet radiation at the Earth's surface would have been so low that it no longer presented a health hazard. Although this is no more than an interesting speculation, the persistence of animals requiring oxygen in rather large quantities indicates that oxygen levels in the atmosphere have probably never been lower than two-tenths of the present value during the past 400 million years. There may have been times when the oxygen pressure was greater than at present.

Most, if not all, coal deposits contain charcoal. This fact implies that during the past 350 Ma the atmosphere has contained more than about 15% oxygen, the minimum oxygen level required to sustain fires. Many coal deposits also contain the remains of very large trees. If the oxygen content of the atmosphere had been more than 30–35%, forest fires would probably have been so frequent and so intense that trees would not have been able to grow to such great heights. The evidence from coal deposits therefore indicates that the oxygen content of the atmosphere has been between 15 and 35% during the past 350 Ma.

Evidence from sediments. Today oxidation of rocks at the Earth-atmosphere interface is pervasive. If oxygen was essentially absent from the atmosphere in times past, one might expect to see relatively less oxidation in the minerals of ancient sediments, and to see the formation of new minerals in these sediments, which are less oxidized than their modern counterparts. Iron, manganese, uranium, and sulfur are among the elements which today respond most readily to oxidation at the Earth–atmosphere interface. Uraninite (UO_2), for instance, reacts rapidly with atmospheric oxygen to form a variety of higher oxides and hydrous oxides. A concerted search for uraninite in black sands during World War II was unsuccessful. And yet, uraninite, which has apparently survived weathering and transport, occurs as ores in the sedimentary rocks of the Dominion Reef series and the Witwatersrand series of South Africa (**Figs. 1** and **2**), as well as at Blind River, Canada, and at Serra

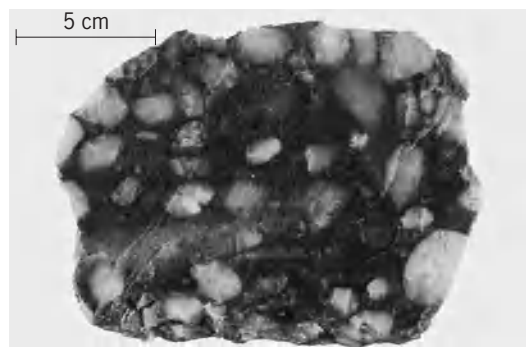

Fig. 1. Polished slab of gold ore, Modder Deep, South Africa. The white pebbles are quartz, SiO_2, and the major metallic constituent is pyrite, FeS_2. Most of the gold ore also contains uraninite, UO_2, which is apparently detrital.

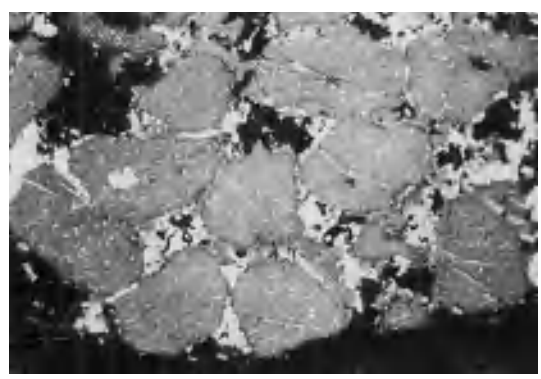

Fig. 2. Photomicrograph (×280) of a polished section of uranium ore, West Rand Cons Mine, South Africa. Gray grains are uraninite, UO_2, which are probably detrital. White material in cracks and between grains of uraninite is brannerite, an oxide of uranium, titanium, and calcium. (*From Ramdohr, Abhandl. Deut. Akad. Wiss. Berlin, 1958*)

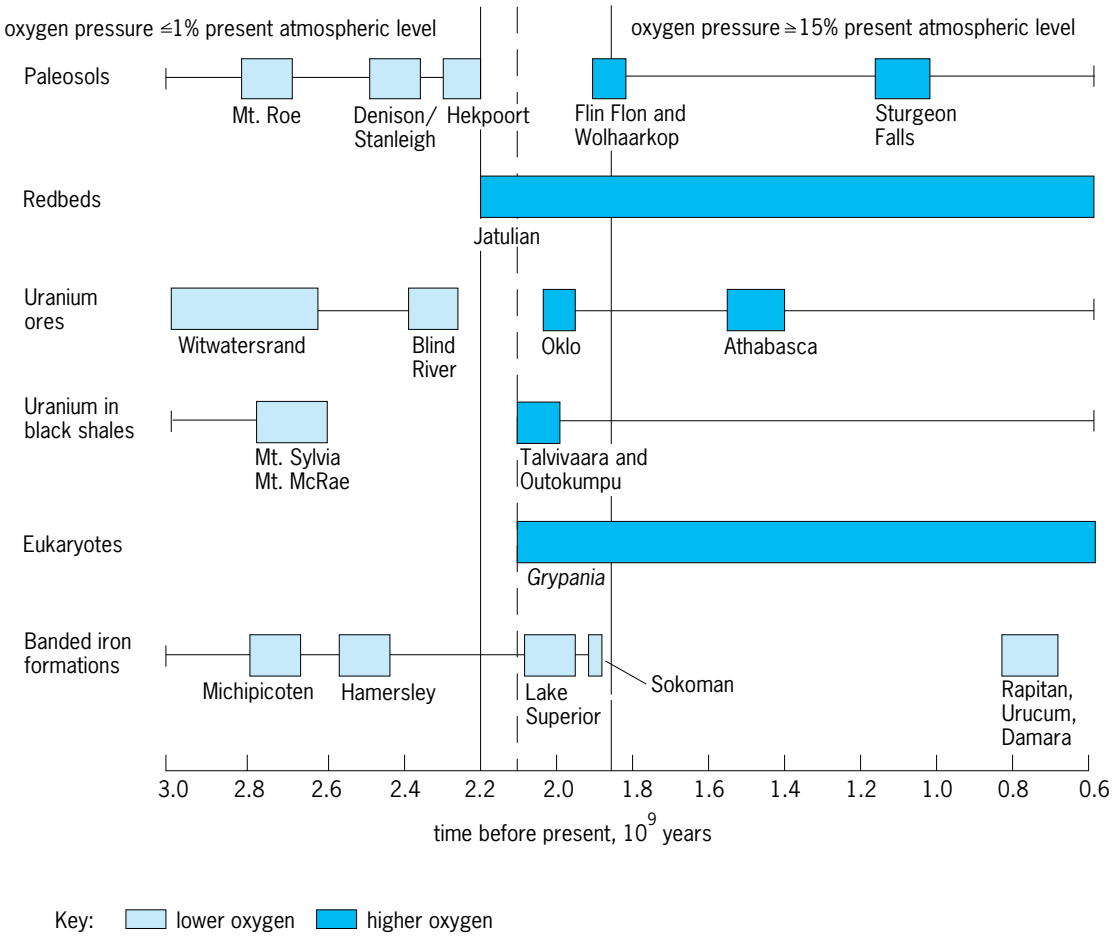

Fig. 3. Summary of data bearing on the O_2 content of the Precambrian atmosphere. (*After H. D. Holland, Early Proterozoic atmospheric change, in S. Bengston, ed., Early Life on Earth, Columbia University Press, 1994*; *also R. Rye and H. D. Holland, Amer. J. Sci., 298:621–672, 1998*)

de Jacobina, Brazil. The sediments in these areas are all more than 2.3 Ga. Nearly all geologists who have closely studied the ore deposits have agreed that much of the uraninite is a residue of weathering. This view has been corroborated by measurements of the age of uraninite grains in the deposits. They appear to be older than the sediments in which they occur. If the grains are indeed detrital, the rate of oxidation of uraninite must have been much slower prior to 2.3 Ga than it is today. Careful studies of the rate of oxidation of uraninite have shown that the oxygen pressure must have been less than about 4×10^{-3} atm (4×10^2 pascals) to permit the survival of uraninite during weathering, transport, and deposition of these ores.

Between 2.3 and 2.0 Ga the nature of uranium ore deposits changed from detrital to hydrothermal. The formation of hydrothermal uranium deposits requires a significant amount of atmospheric oxygen. A very significant rise in the oxygen level between 2.3 and 2.0 Ga is therefore indicated. Australian sandstones deposited between 2.75 and 3.25 Ga also contain other detrital minerals that are easily oxidized and that are absent from younger sandstones; this confirms the evidence from uranium minerals for a low-O_2 atmosphere prior to 2.7 Ga. Major differences between the oxidation state of iron in soils formed prior to 2.3 Ga and in those formed more recently than 1.9 Ga (**Fig. 3**) point to the same conclusion, as does the first appearance of highly oxidized redbeds about 2.2 Ga. A major positive excursion in the isotopic composition of carbon in carbonate rocks deposited between 2.25 and 2.06 Ga also suggests that oxygen levels climbed dramatically during this time interval (**Fig. 4**). In defining the changes in the isotopic composition, delta (δ) terms are used. The delta term for carbon is defined in Eq. (2), showing the

$$\delta^{13}C = \left[\frac{\left(\frac{^{13}C}{^{12}C}\right)_{sample} - \left(\frac{^{13}C}{^{12}C}\right)_{standard}}{\left(\frac{^{13}C}{^{12}C}\right)_{standard}} \right] \times 1000‰ \quad (2)$$

ratio of ^{13}C to ^{12}C in a sample and in a standard, which by convention is a Peedee belemnite. For sulfur, the delta term is defined in Eq. (3).

$$\delta^{34}S = \left[\frac{\left(\frac{^{34}S}{^{32}S}\right)_{sample} - \left(\frac{^{34}S}{^{32}S}\right)_{standard}}{\left(\frac{^{34}S}{^{32}S}\right)_{standard}} \right] \times 1000‰ \quad (3)$$

Evidence from gaseous emissions. A third line of evidence is available to tell something about the oxidation state of the atmosphere in a nonbiotic state. W. W. Rubey showed that many of the volatile

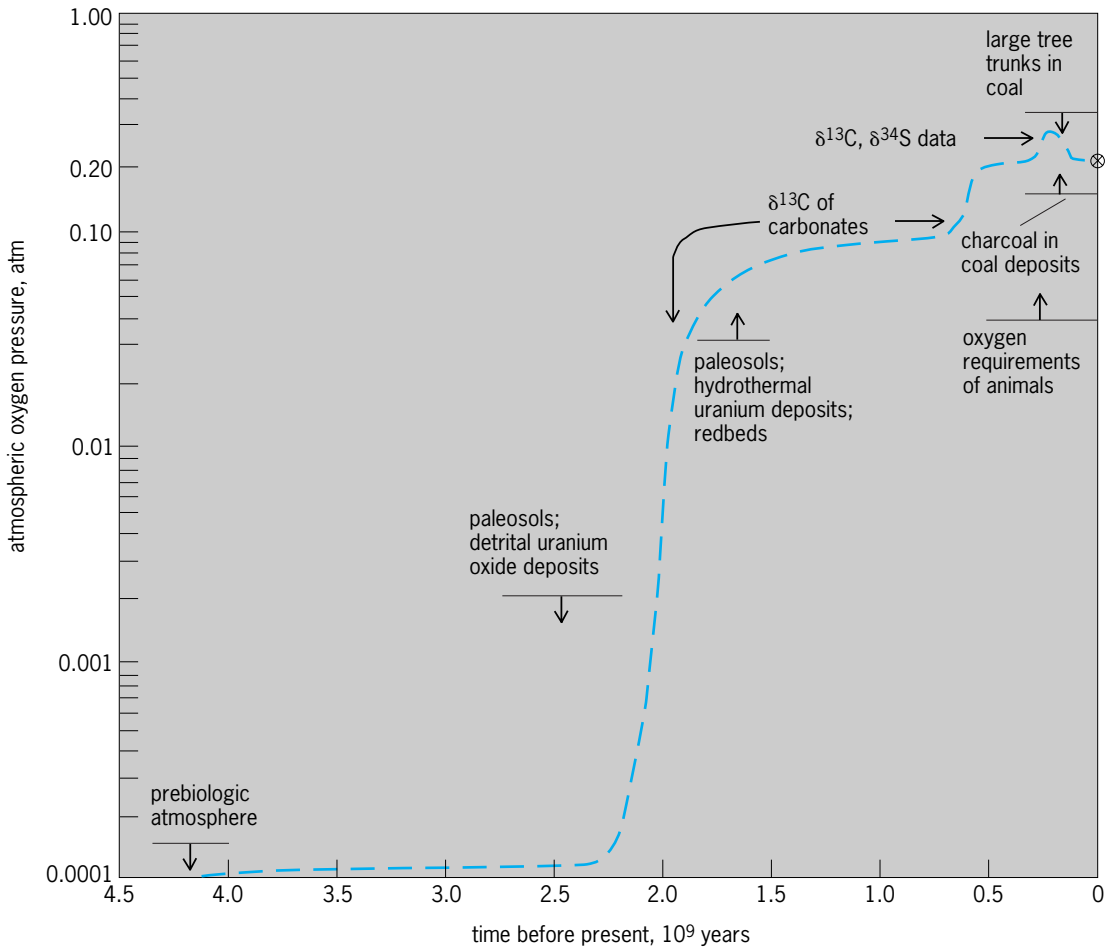

Fig. 4. Summary of the evidence for the evolution of atmospheric oxygen since the Earth was formed. 1 atm = 10^5 Pa.

constituents of the atmosphere and oceans cannot have been derived from the weathering of igneous rocks. He developed a very plausible argument for the concept that these "excess volatiles" have boiled out of the interior of the Earth during the course of its long history. Volcanoes give eloquent evidence for such boiling out today (**Fig. 5**), and it seems likely that a portion of the discharge of at least some hot springs has a deep source. In a nonbiotic state the chemistry of the atmosphere would be controlled largely by the chemistry of such emanations and by their interaction with surface rocks. Thus, if something were known about the oxidation of volcanic gases in the past, it might be possible to predict, at least within broad limits, the oxidation state of an atmosphere unaffected by biologic processes.

Fig. 5. Eruption of Mount Vesuvius in 1943. The gases emitted by volcanoes throughout geologic time have played a major role in the chemical evolution of the atmosphere of the Earth.

At present, volcanic gases consist mainly of water vapor, carbon dioxide, sulfur dioxide, hydrogen, carbon monoxide, and nitrogen (**Table 3**). Free oxygen is almost completely absent: The oxygen pressure

TABLE 3. Composition of typical Hawaiian volcanic gases

Gas	Vol %
H_2O	79.3
CO_2	11.6
SO_2	6.5
N_2	1.3
H_2	0.6
CO	0.4
Cl_2	0.05
Ar	0.04

in these gases as they emerge is about 10^{-7} atm (10^{-2} Pa). If a gas mixture is defined as being neutral from an oxidation-reduction point of view when it contains neither free hydrogen nor free oxygen, then these gases are slightly on the reduced side, since they contain a small amount of free hydrogen (H_2) and carbon monoxide (CO). Today these reduced gases react rapidly with atmospheric oxygen. In the absence of atmospheric oxygen, a variety of different reactions might take place. All of these tend to produce an essentially neutral atmosphere which would be similar to the present atmosphere, with the difference that oxygen would be virtually absent.

The oxidation state of volcanic gases today is controlled in large part by the oxidation state of the associated lavas. This in turn is reflected in the ratio of ferrous iron (Fe^{2+}) to ferric iron (Fe^{3+}) in lavas, a ratio which tends to be preserved on cooling. Studies of this ratio in basalts have shown that the Fe^{2+}/Fe^{3+} ratio has probably not changed significantly during at least the past 2 Ga. It is fairly certain, then, that the oxidation state of volcanic gases has not varied greatly during the second half of Earth history. There are good theoretical reasons for believing that this same state of affairs prevailed at least back to 3 Ga, but it is possible that shortly after its birth the Earth vented volcanic gases which were more highly reducing. The cause for this difference would have been the presence of metallic iron in the upper part of the Earth's mantle. Today this iron is largely concentrated in the Earth's core, far below the part of the mantle where lavas are generated. *See* EARTH INTERIOR.

Rise of oxygen pressure. It may be well to summarize the evidence regarding the history of the pressure of oxygen in the Earth's atmosphere before considering the pressure of the other constituents through geologic time. The largest area of ignorance surrounds the events prior to the accumulation of the earliest known sediments, some 3.8 Ga. It is possible that the atmosphere was quite reducing as a consequence of the injection of highly reduced volcanic gases into the atmosphere during the first 100 Ma of Earth history.

Shortly after 3 Ga the atmosphere probably contained less than 4×10^{-3} atm (4×10^2 Pa) oxygen. Green plant photosynthesis was already under way, but the rate of burial of organic carbon was insufficient to support more than a small oxygen pressure in the atmosphere. This condition probably prevailed until about 2.2 Ga. The first widespread redbed sequences were deposited at about that time, and no uranium deposits of the Witwatersrand–Blind River type younger than 2.3 Ga have been discovered. Both observations are consistent with a major increase in atmospheric oxygen between 2.3 and 2.0 Ga.

Another rise may well have taken place at the end of the Precambrian and may have been important for the development of animal life. Finally, the colonization of the land about 400 Ma may have caused a further rise in the oxygen pressure. Since that time the oxygen pressure has probably not fluctuated greatly (Fig. 4).

Rare gases other than helium. Of the atmospheric constituents other than oxygen, the rare gases, nitrogen, and carbon dioxide are probably the most interesting. Studies have shown that the Earth is very depleted in rare gases in comparison with the Sun. Even xenon, a heavy element which certainly cannot escape from the Earth's atmosphere today at anything but a geologically insignificant rate, appears to be only about one-millionth as abundant in comparison with silicon as in the Sun. From these studies it was concluded that the Earth was essentially devoid of an atmosphere after it had reached its present size and gravitational field, although it is possible that an original gaseous envelope containing the missing rare gases was swept away by strong magnetic fields which penetrated throughout the solar system, or during the violent Hayashi phase early in the history of the Sun, or during an impact with a Mars-size object. Whichever view turns out to be correct, the Earth seems to have been left nearly devoid of an atmospheric blanket at some time not long after the end of the accumulation phase.

Rubey's argument for the degassing of the Earth's interior is strongly supported by the rather anomalous abundance and isotopic composition of atmospheric argon. This gas is much more abundant than the other rare gases (Table 1), and consists almost entirely of the argon isotope of mass 40, which in the normal course of nuclear events should not be much more abundant than the argon isotopes of mass 36 and mass 38. The anomaly is readily explained by the degassing of argon-40 from the Earth after its production from the radioactive decay of potassium-40. Although the potassium content of the Earth is not well known, there is little doubt that the decay of the available potassium could account for the present abundance of atmospheric argon-40. It seems likely that degassing was most intense during the early history of the Earth, but the discovery of excess helium-3 in the Pacific Ocean indicates that the release of primordial gases is continuing from mid-ocean ridges. The atmospheric pressure of neon, argon, krypton, and xenon has probably increased gradually during the course of Earth history. *See* ARGON; INERT GASES; RADIOISOTOPE.

Other gases. The abundance of helium in the atmosphere is anomalously low. The quantity of helium which has been generated in the solid Earth by the decay of uranium and thorium series nuclides and has entered the atmosphere with other gases is much larger than the quantity of helium now present in the atmosphere. Helium has clearly escaped from the atmosphere into interplanetary space. *See* HELIUM.

Nitrogen is a good deal more reactive than the rare gases. It is cycled through living organisms at a geologically rapid rate but seems to reside mainly in the atmosphere. From a thermodynamic point of view, nitrate should be quite abundant in the oceans today. Its concentration in seawater is kept to very low levels by organisms for whom nitrate is an important and frequently limiting nutrient. *See* NITROGEN.

Ammonia is thermodynamically unstable in the presence of free oxygen; a major fraction of the nitrogen that is now in the atmosphere would be converted to ammonia only if the hydrogen pressure were in excess of 10^{-3} atm. Such a hydrogen pressure could have existed only early in the Earth's history.

The value of the hydrogen pressure, and even the very existence of a highly reducing early atmosphere, is presently very uncertain. R. Brett pointed out that the relatively large nickel content of olivine in the upper mantle is inconsistent with the presence of large amounts of metallic iron in the upper mantle early in Earth history. If this is correct, volcanic gases may never have been much more reducing than at present.

Methane could have been a major constituent of an early reducing atmosphere. It is likely that its pressure was already less than 10^{-3} atm (10^2 Pa) during the deposition of the sediments of the Isua area in West Greenland 3.8 Ga. Methanogens, that is, organisms that generate methane, may have been a significant part of the early biosphere, and may have been responsible for the presence of a significant amount of methane in the early atmosphere. However, methane must have been a very minor component ever since atmospheric oxygen became more than a trace constituent of the atmosphere. The very small concentration of CH_4 since then has almost certainly been controlled by the balance between its rate of biologic production and the rate of its photochemical decomposition in the atmosphere. *See* METHANE.

It seems likely that the carbon dioxide pressure has been reasonably close to the value at which hydrous magnesium aluminum silicate minerals are in equilibrium with kaolinite, quartz , calcite, and dolomite [reaction (4)]. All of these minerals are abundant in

$$\underset{\text{Chamosite}}{Mg_5Al_2Si_3O_{10}(OH)_8} + \underset{\text{Calcite}}{5CaCO_3} + 5CO_2 \rightarrow$$

$$\underset{\text{Dolomite}}{5CaMg(CO_3)_2} + \underset{\text{Kaolinite}}{Al_2Si_2O_5(OH)_4} + \underset{\text{Quartz}}{SiO_2} + 2H_2O \qquad (4)$$

sediments and sedimentary rocks, and it is likely that such reactions have helped to maintain the CO_2 pressure within two orders of magnitude of its present value for much of geologic time. There has been considerable debate concerning the CO_2 content of the atmosphere during the Archean, that is, more than 2.5 Ga. Some geochemists have proposed that the CO_2 pressure was in excess of 10^{-1} atm (10^4 Pa) and possibly as high as 10 atm (10^6 Pa). However, evidence from the composition of paleosols indicates that the CO_2 pressure has probably been less than 3×10^{-2} atm (3×10^3 Pa) since 2.75 Ga.

The analyses of small air bubbles trapped in the ice sheets of Greenland and Antarctica have had a profound impact on thinking about atmospheric evolution and the origin of ice ages. The CO_2 content of the atmosphere has varied by about 50% during the past 160,000 years (**Fig. 6**). Levels of CO_2 during the last

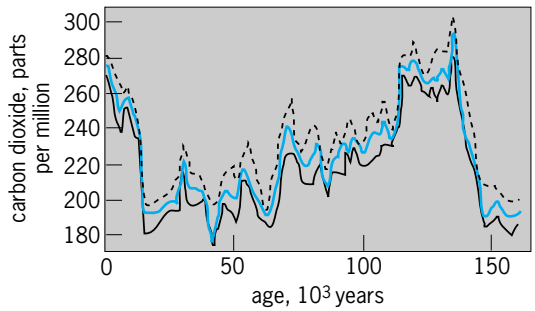

Fig. 6. Carbon dioxide concentration in air bubbles in the Vostok ice core, Vostok Station, Antarctica. (*After C. Lorius, Antarctic ice core: CO_2 and climatic change over the last climatic cycle, EOS, 69(26):681, 683–684, June 28, 1988*)

interglacial period around 125,000 years ago were about 270 parts per million (ppm) [2.7×10^{-4} atm or 27 Pa]. The CO_2 content in the atmosphere dropped gradually to about 200 ppm during the last ice age but rebounded rapidly to about 270 ppm at the end of the last ice age about 10,000 years ago. Thereafter it varied little until the beginning of the industrial revolution. Since then fossil fuel burning and, to a lesser extent, deforestation have pushed the CO_2 content of the atmosphere up to about 370 ppm (3.70×10^{-4} atm or 37 Pa). The CO_2 pressure is apt to rise to roughly twice that value by the end of the twenty-first century, unless economics and the fear of adverse climate changes produce a major shift away from the burning of fossil fuels as the dominant source of the world's energy.

Today CO_2 is second in abundance to water in gases issuing from volcanoes. Yet the ratio of CO_2 to water in the atmosphere plus the oceans is minuscule. Carbon dioxide has clearly been scavenged very thoroughly from the atmosphere-hydrosphere system. Its resting place is easily discovered in the elemental carbon of sedimentary rocks and in the carbonate rocks that have been deposited throughout the entire range of Earth history that is accessible through the study of sedimentary rocks (**Fig. 7**). The two dominant carbonate minerals, calcite ($CaCO_3$) and dolomite [$CaMg(CO_3)_2$], are the major components of limestones and dolomites. The removal of CO_2 from the atmosphere, its reaction with surface rocks, and its ultimate burial are processes that are chemically, mineralogically, and biologically complex.

Three stages of evolution. The above discussion seems to lead quite naturally to a threefold division of atmospheric history (**Table 4**). During the first stage, very shortly after the accretion of the Earth, the atmosphere may have been quite reducing. Reduced volatiles issuing from volcanoes probably would have given rise to an atmosphere consisting of methane with minor quantities of hydrogen, nitrogen, and ammonia. After metallic iron was removed from the upper mantle, the oxidation state of volcanic gases probably approached its present value, methane was replaced by carbon dioxide, and ammonia was converted to nitrogen. In the atmosphere during this second stage, nitrogen was its dominant component,

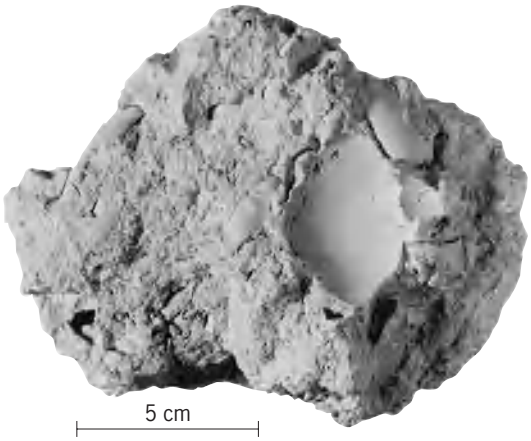

Fig. 7. Sample of Tertiary limestone from Florida. Nearly all carbon dioxide injected into the atmosphere has been removed in limestones and dolomites. During the past 6 × 10^4 years much removal has taken place through the action of organisms with carbonate shells. Prior to that time, removal by inorganic precipitation and by stromatolites was probably dominant.

and carbon dioxide and argon were its most important minor constituents.

The third stage opened when the rate of oxygen production became sufficiently great that oxygen became more than a trace component of the atmosphere. The transition from stage 2 to stage 3 probably occurred between 2.3 and 2.0 Ga. Since the opening of stage 3, the oxygen pressure has climbed to its present value by a path which was probably simple but could have been complex in detail. During this period the nitrogen, neon, argon, krypton, and xenon pressures have also gradually climbed to their present value, while the helium pressure remained reasonably constant, suspended between the rate of input and the rate of escape. The CO_2 pressure has probably been less than 100 times and more than half the present value during stage 3. There may have been a general decrease toward its present value, but fluctuations around a generally decreasing trend have probably been significant.

TABLE 4. Summary of data on the probable chemical composition of the atmosphere during stages 1, 2, and 3

Components	Stage 1	Stage 2	Stage 3
Major components:			
$P > 10^{-2}$ atm (10^3 Pa)	CH_4 H_2 (?)	N_2	N_2 O_2
Minor components:			
10^{-2} atm $> P >$ 10^{-4} atm (10 Pa)	H_2 (?) H_2O N_2 H_2S NH_3 Ar	H_2O CO_2 A O_2 (?)	Ar H_2O CO_2
Trace components:			
10^{-4} atm $> P >$ 10^{-6} atm (10^{-1} Pa)	He	Ne He CH_4 NH_3 (?) SO_2 (?) H_2S (?)	Ne He CH_4 Kr

Society is actively modifying the composition of the atmosphere. The major atmospheric constituents are essentially immune to human impacts, but the CO_2 content of the atmosphere has increased significantly, as has that of methane, and there have been major changes in the conce ntration of many trace gases. Some of these gases are entirely new to the atmosphere. Their concentration is extremely small, but the effects of some, like the chlorofluorocarbons, have been disproportionately large.

Heinrich D. Holland

Bibliography. S. Bengtson (ed.), *Early Life on Earth*, 1994; H. D. Holland, *The Chemical Evolution of the Atmosphere and Oceans*, 1984; F. S. Rowland and I. S. A. Isaksen (eds.), *The Changing Atmosphere*, 1988; S. H. Schneider and P. J. Boston (eds.), *Scientists on Gaia*, 1991; J. W. Schopf (ed.), *Earth's Earliest Biosphere*, 1992; J. W. Schopf and C. Klein (eds.), *The Proterozoic Biosphere*, 1992.

Atmospheric acoustics

The science of sound in the atmosphere. Atmospheric acoustics is concerned with sound outdoors rather than indoors. The atmosphere has a structure that varies in both space and time, and these variations have significant effects on a propagating sound wave. In addition, when sound propagates close to the ground, the type of ground surface has a strong effect.

Development. Atmospheric acoustics originated with the rebirth of science in the seventeenth century. Early work focused on measuring the speed of sound. In 1636, Marin Mersenne measured this speed by timing the interval between the flash and sound of a gun blast. He obtained the inaccurate value of 448 m/s (1470 ft/s). A contemporary, Pierre Gassendi, noted that the speed of sound was independent of its intensity, since the speed was the same whether the sound was made by a large cannon or a musket. The next century saw the first precise measurements of the speed of sound. In 1738 measurements under the direction of the Academy of Paris yielded a value of 332 m/s (1089 ft/s) when corrected to 0°C (32°F). At about the same time, G. L. Bianconi showed that the speed of sound definitely increased with increasing air temperature.

Sound travels relative to air with the speed given by Eq. (1), where T is absolute temperature in kelvins

$$c = (\gamma R_0 T/M)^{1/2} \qquad (1)$$

(temperature in degrees Celsius + 273.15), R_0 is the universal gas constant, γ is the specific heat ratio, and M is the average molecular weight. For dry air of normal composition, $M = 29.0$ and Eq. (1) gives $c = 331$ m/s (1086 ft/s) at 0°C (32°F). Thus, the value found in 1738 was within 0.3% of the best modern value.

The sustained development of atmospheric acoustics did not begin until the last half of the nineteenth century. The effects of atmospheric profiles on the

refraction of sound received considerable attention. In the early twentieth century, there were significant advances in understanding molecular absorption processes in the atmosphere. Starting in the 1960s, concern over increasing environmental noise led to increased understanding of the effects of the ground on sound propagation outdoors. At about the same time, the invention of the echosonde opened a new area of research on the acoustic sounding of the atmosphere.

Atmospheric sound attenuation. As sound propagates in the atmosphere, several interacting mechanism attenuate and change the spectral or temporal characteristics of the sound received at a distance from the source. The attenuation means that sound propagating through the atmosphere decreases in level with increasing distance between source and receiver. The total attenuation, in decibels, can be approximated as the sum of three nominally independent terms, as given in Eq. (2), where A_{div} is the

$$A_{total} = A_{div} + A_{air} + A_{env} \qquad (2)$$

attenuation due to geometrical divergence, A_{air} is the attenuation due to air absorption, and A_{env} is the attenuation due to all other effects and includes the effects of the ground, refraction by a nonhomogeneous atmosphere, and scattering effects due to turbulence.

Sound energy spreads out as it propagates away from its source due to geometrical divergence. At distances that are large compared with the effective size of the sound source, the sound level decreases at the rate of 6 dB for every doubling of distance. The phenomenon of geometrical divergence, and the corresponding decrease in sound level with increasing distance from the source, is the same for all acoustic frequencies. In contrast, the attenuation due to the other two terms in Eq. (2) depends on frequency and therefore changes the spectral characteristics of the sound.

Air absorption. Dissipation of acoustic energy in the atmosphere is caused by viscosity, thermal conduction, and molecular relaxation. The last arises because fluctuations in apparent molecular vibrational temperatures lag in phase the fluctuations in translational temperatures. The vibrational temperatures of significance are those characterizing the relative populations of oxygen (O_2) and nitrogen (N_2) molecules. Since collisions with water molecules are much more likely to induce vibrational state changes than are collisions with other oxygen and nitrogen molecules, the sound attenuation varies markedly with absolute humidity. *See* CONDUCTION (HEAT); MOLECULAR STRUCTURE AND SPECTRA; VISCOSITY.

The total attenuation due to air absorption increases rapidly with frequency (**Fig. 1**). For example, at a pressure of 1 atm, temperature of 20°C (68°F), and relative humidity of 70%, the total attenuation is about 1 dB for every 100 m (328 ft) at 2 kHz, but is close to 100 dB for every 100 m at 20 kHz. For this reason, applications in atmospheric acoustics are restricted to sound frequencies below a few thousand hertz if the propagation distance exceeds a few hundred meters. *See* SOUND ABSORPTION.

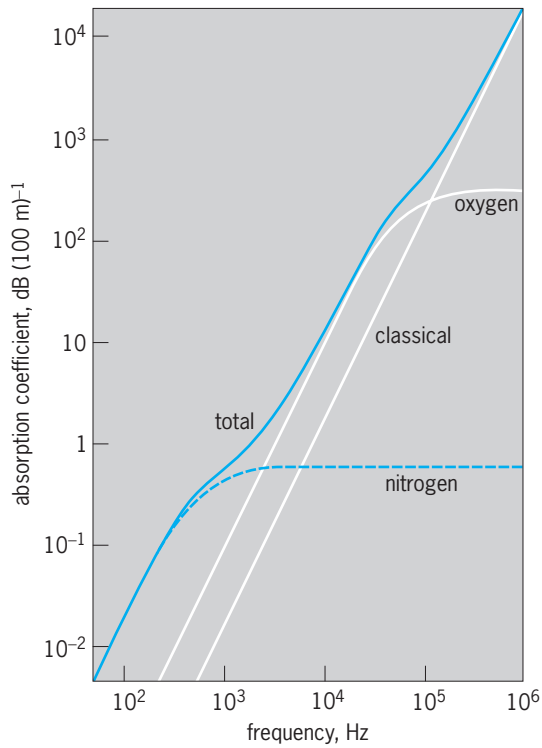

Fig. 1. Frequency dependence of attenuation coefficient in units of decibals per 100 m (dB/328 ft), temperature of 20°C (68°F), and a relative humidity of 70%. Total attenuation from air absorption and contributions from classical absorption (that due to viscosity and thermal conduction) and molecular absorption of oxygen and nitrogen are shown. (*After J. E. Piercy, T. F. W. Embleton, and L. C. Sutherland, Review of noise propagation in the atmosphere, J. Acous. Soc. Amer., 61:1403–1418, 1977*)

Effects of the ground. When the sound source and receiver are above a large flat ground surface in a homogeneous atmosphere, sound reaches the receiver via two paths (**Fig. 2**). There is the direct path from source to receiver and the path reflected from the ground surface. Most naturally occurring ground surfaces are porous to some degree, and their acoustical property can be represented by an acoustic impedance. The acoustic impedance of the ground is in turn associated with a reflection

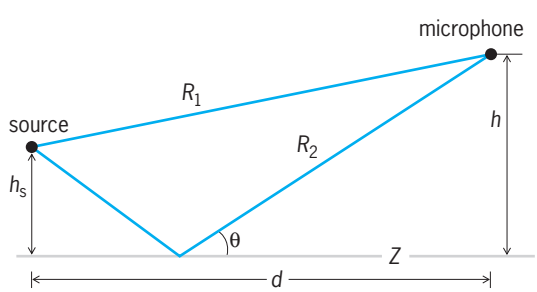

Fig. 2. Diagram of reflection of sound from a flat ground surface with impedance Z, showing direct (R_1) and reflected (R_2) paths from source to receiver, source height h_s, microphone height h, distance of propagation d, and angle of incidence or grazing angle θ.

coefficient that is typically less than unity. In simple terms, the sound field reflected from the ground surface suffers a reduction in amplitude and a phase change.

If the incident waves are plane, the reflection coefficient R_p on a plane ground surface is given in its simplest from by Eq. (3), where θ is the angle of

$$R_p = \frac{\sin\theta - (\rho c/Z)}{\sin\theta + (\rho c/Z)} \tag{3}$$

incidence or grazing angle and $\rho c/Z$ is the ratio of the characteristic impedance of the air (ρc, where ρ is the air density) to the acoustic impedance (Z) of the ground surface. At grazing incidence (θ approaches 0), R_P always approaches -1. Therefore, in the case of plane waves, the reflected sound cancels the direct sound at grazing incidence. *See* ACOUSTIC IMPEDANCE.

Real sound sources outdoors can be considered point sources at distances that are large compared with the effective size of the source. Wavefronts are no longer plane but spread spherically away from the source. A good approximation for the total sound pressure from a point source above a plane ground surface is given by Eq. (4), where p_d is the sound

$$p_t = p_d + R_p p_r + p_g \tag{4}$$

field along the direct path, p_r is the sound field along the reflected path, and p_g is a correction term that allows for the curvature of the incident sound field and, under certain conditions, the possible existence of a surface wave.

When the source and receiver are both relatively near the ground and are a large distance apart, the direct and reflected fields become nearly equal and the grazing angle tends to zero. The direct and reflected sound fields then cancel each other because R_P approaches -1, and any sound reaching the receiver is given by the third term, p_g. The consequences of these phenomena are illustrated by the spectra in **Fig. 3**. The sound pressure levels are calculated for a source height of 1.8 m (6 ft), a microphone height of 1.5 m (5 ft), and distances between 125 m (410 ft) and 4 km (2.4 mi). The levels are corrected for geometrical spreading and molecular absorption, and therefore show the effects of the ground only. The large dips around 500 Hz are not true interference dips due to comb filtering from path-length differences, but are the result of the third term in Eq. (3). The position of the dip does not change with source-receiver separation. Therefore, spatial averaging does not remove the decreased sound pressure levels at these frequencies. This has important consequences for environmental noise since 500 Hz is a common dominant frequency of many noise sources (ground transportation noise, for example). Naturally occurring ground provides a significant amount of natural shielding in the case of propagation at grazing incidence.

Refraction of sound. Straight ray paths are rarely achieved outdoors. In the atmosphere, both the wind and temperature vary with height above the ground. The velocity of sound relative to the ground is a function of wind velocity and temperature; hence it also varies with height, causing sound waves to propagate along curved paths. There are two distinct cases.

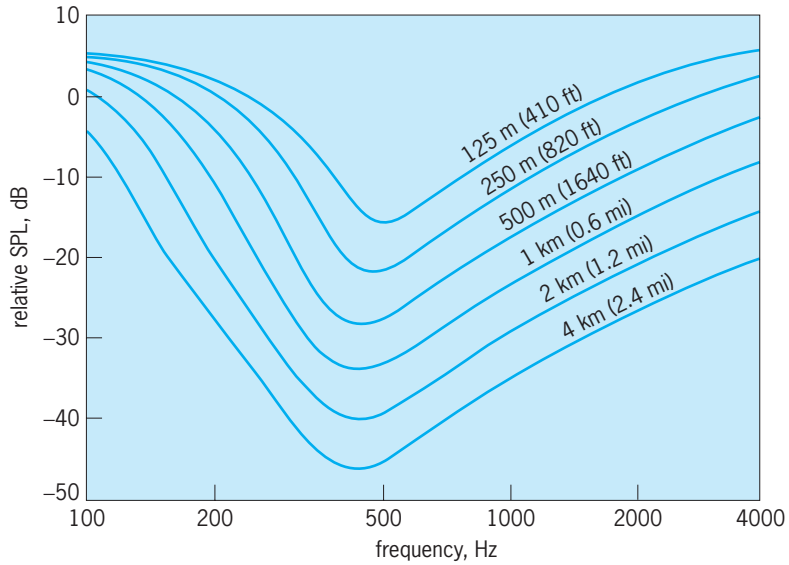

Fig. 3. Relative sound pressure levels (SPL) for propagation from a source over mown grass for source height of 1.8 m (6 ft), microphone height of 1.5 m (5 ft), and distances of propagation indicated.

The speed of the wind decreases with decreasing height above the ground because of drag on the moving air at the surface. Therefore, the speed of sound relative to the ground increases with height during downwind propagation, and ray paths curve downward (**Fig. 4***a*). For propagation upwind, the sound speed decreases with height, and ray paths curve upward (Fig. 4*b*). In the case of upward refraction, a shadow boundary forms near the ground beyond which no direct sound can penetrate. Some acoustic energy penetrates into a shadow zone via creeping waves that propagate along the ground and that continually shed diffracted rays into the shadow zones. The dominant feature of shadow-zone reception is the marked decrease in a sound's higher-frequency content. The presence of shadow zones explains

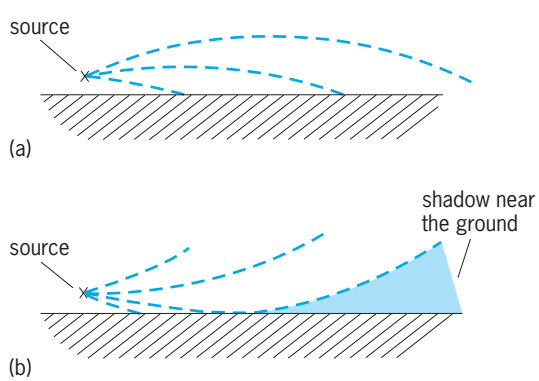

Fig. 4. Curved ray paths. (*a*) Refraction downward, during temperature inversion or downwind propagation. (*b*) Refraction upward, during temperature lapse or upwind propagation.

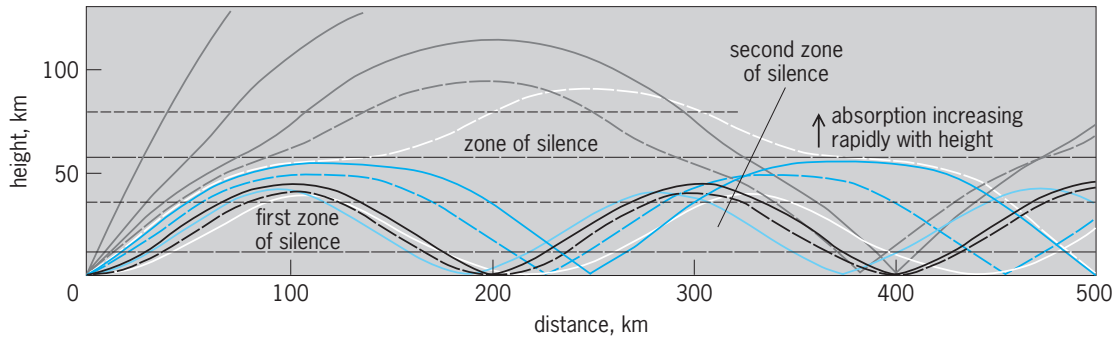

Fig. 5. Representative long-range ray paths of sound radiating from a source in the atmosphere and traveling from west to east in the Northern Hemisphere in winter. 1 km = 0.6 mi. (*After T. F. Malone, ed., Compendium of Meteorology, American Meteorological Society, 1951*)

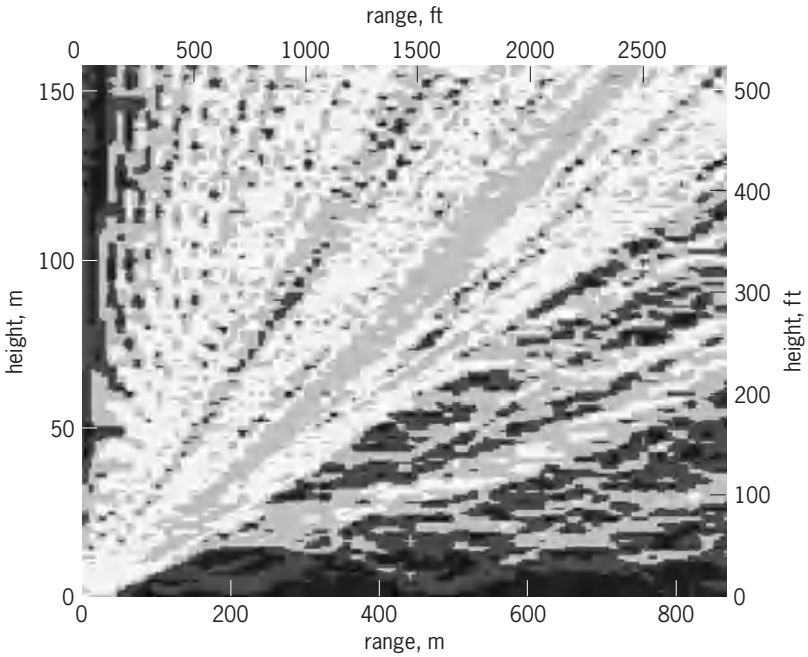

Fig. 6. Contour plot of sound levels in the presence of upward refraction in a turbulent atmosphere. The source is very close to the ground at zero range. The white areas are regions of higher sound levels.

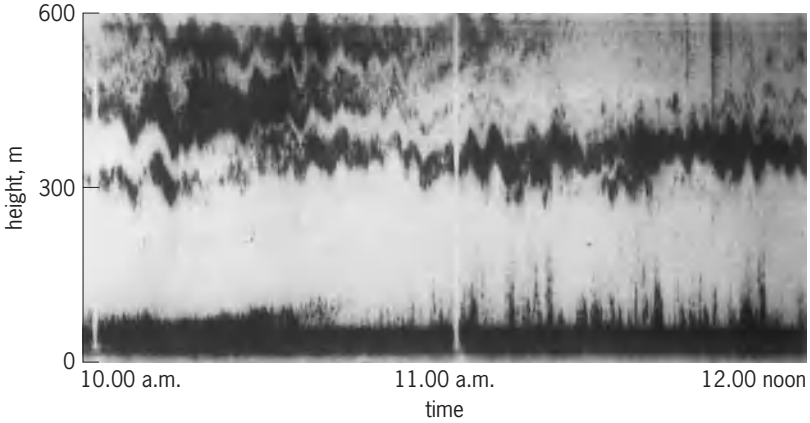

Fig. 7. Facsimile record of acoustic echo sounding of atmosphere at the South Pole, February 9, 1975. 300 m = 984 ft, 600 m = 1969 ft. Dark regions above 300 m altitude correspond to temperature inhomogeneities associated with the passage of convective plumes over the site. (*After E. H. Brown and F. F. Hall, Jr., Advances in atmospheric acoustics, Rev. Geophys. Space Phys., 16:47–110, 1978*)

why sound is generally less audible upwind of a source.

Refraction by temperature profiles is analogous. During the day, solar radiation heats the Earth's surface, resulting in warmer air near the ground. This condition is called a temperature lapse and is most pronounced on sunny days. A temperature lapse is the common daytime condition during most of the year, and also causes ray paths to curve upward. After sunset there is often radiation cooling of the ground, which produces cooler air near the surface. In summer under clear skies, such temperature inversions begin to form about 2 hours after sunset. Within the temperature inversion, the temperature increases with height, and ray paths curve downward.

The effects of refraction by temperature and wind are additive and produce rather complex sound speed profiles in the atmosphere (**Fig. 5**). Rays can be trapped in a sound channel where the effective speed of sound has a minimum. Audible sound is often received at distances of a few hundred kilometers from large explosions, even though the sound may be inaudible at closer distances. *See* REFRACTION OF WAVES.

Effects of turbulence. The atmosphere is also an unsteady medium with random variations in temperature, wind velocity, pressure, and density. In practice, only the temperature and wind velocity variations significantly contribute to the effects of turbulence on a sound wave. Turbulence in the atmosphere causes the effective sound speed to fluctuate from point to point, so a nominally smooth wave front develops ripples. One result is that the direction of a received ray may fluctuate with time in random manner. Consequently, the amplitude and phase of the sound at a distant point will fluctuate with time. The acoustical fluctuations are in some respects analogous to more familiar optical phenomena such as the twinkling of light from a star. The acoustical fluctuations are clearly audible in the noise from a large aircraft flying overhead. *See* TWINKLING STARS.

Turbulence in the atmosphere also scatters sound from its original direction. A collimated beam will spread in width so that the amplitude along the axis will appear to undergo an additional

attenuation. Another consequence of sound scattering by turbulence is the degradation of refractive acoustic shadow zones (**Fig. 6**). *See* TURBULENT FLOW.

Echosonde. Atmospheric acoustics has found extensive application in the study of meteorological disturbances in the lower atmosphere. The echosonde, invented by L.G. McAllister in the late 1960s, is the forerunner and also the prototype of current acoustic sounding instrumentation. A transducer pointing upward transmits a short acoustic pulse and then receives weak echoes over a relatively long period of time. Echoes received at time Δt after transmission have been reflected from inhomogeneities at height $c\Delta t/2$.

A facsimile display (**Fig. 7**) is produced by a side-by-side superposition of the echo histories for a long succession of such sounding experiments; this is interpreted as a picture of the time evolution of the height profile of the atmosphere's ability to backscatter acoustic waves. Darker regions on the display imply that, at the corresponding times and heights, the backscattering was strong. Meteorological interpretations of such records requires experience and an understanding of fundamental atmospheric processes. Refined systems in conjunction with the theory of sound propagation through a turbulent atmosphere enable a quantitative determination of temperature and wind-velocity structures above the ground and hence a characterization of the turbulence. *See* SOUND. Gilles A. Daigle

Bibliography. M. J. Crocker (ed.), *Encyclopedia of Acoustics*, Wiley-Interscience, 1997; H. H. Hubbard (ed.), *Aeroacoustics of Flight Vehicles: Theory and Practice*, Acoustical Society of America, 1994; A. D. Pierce, *Acoustics: Introduction to Its Physical Principles and Applications*, 1981.

Atmospheric chemistry

A scientific discipline concerned with the chemical composition of the Earth's atmosphere. Topics include the emission, transport, and deposition of atmospheric chemical species; the rates and mechanisms of chemical reactions taking place in the atmosphere; and the effects of atmospheric species on human health, the biosphere, and climate.

Atmosphere. Over 99.9% of the total mass of the atmosphere is contained in the lowest 50 km (30 mi) above the Earth's surface. This region is divided into the troposphere and the stratosphere. The troposphere extends from the surface to 8–18 km (5–11 mi) altitude, depending on latitude and season, and is characterized by a rapid decrease of temperature with altitude due to solar heating of the surface; it contains 85% of total atmospheric mass. The stratosphere extends from the top of the troposphere (the tropopause) to 50 km (30 mi; the stratopause), and is characterized by an increase of temperature with altitude due to absorption of solar ultraviolet radiation by ozone (O_3). Buoyancy facilitates vertical motions in the troposphere but suppresses these motions in the stratosphere. As a result, vertical mixing of the troposphere takes only about 1 month, while vertical mixing of the stratosphere takes about 5 years. *See* STRATOSPHERE; TROPOSPHERE.

Global transport in the atmosphere is driven primarily by the thermal contrast between high and low latitudes and by the effect of the Coriolis force. East-west transport of species around the globe requires only a few weeks; north-south transport is slower. Exchange of air between the Northern and Southern hemispheres takes about 1 year. *See* ATMOSPHERIC GENERAL CIRCULATION.

A useful quantity in atmospheric chemistry is the atmospheric lifetime, defined as the mean time that a molecule resides in the atmosphere before it is removed by chemical reaction or deposition. The atmospheric lifetime measures the time scale on which changes in the production or loss rates of a species may be expected to translate into changes in the species concentration. The atmospheric lifetime can also be compared to the time scales for transport to infer the spatial variability of a species in the atmosphere; species with lifetimes longer than a decade tend to be uniformly mixed, while species with shorter lifetimes may have significant gradients reflecting the distributions of their sources and sinks.

Principal constituents. The principal constituents of dry air are nitrogen (N_2; 78% by volume), oxygen (O_2; 21%), and argon (Ar; 1%). The atmospheric concentrations of N_2 and Ar are largely determined by the total amounts of N and Ar released from the Earth's interior since the origin of the Earth. The atmospheric concentration of O_2 is regulated by a slow atmosphere-lithosphere cycle involving principally the conversion of O_2 to carbon dioxide (CO_2) by oxidation of organic carbon in sedimentary rocks (weathering), and the photosynthetic conversion of CO_2 to O_2 by marine organisms which precipitate to the bottom of the ocean to form new sediment. This cycle leads to an atmospheric lifetime for O_2 of about 4 million years. It is thought that O_2 has remained at near-present levels in the atmosphere for the past 100 million years, reflecting a balance between sediment formation and weathering. An additional fast cycle between O_2 and CO_2 takes place by oxidation of organic carbon in the biosphere, balanced by photosynthesis; but this cycle does not regulate atmospheric O_2 levels, because it is limited by the supply of organic carbon. The total organic carbon in the biosphere, soil, and ocean reservoirs could convert at most 1% of atmospheric O_2 to CO_2. *See* ATMOSPHERE, EVOLUTION OF; BIOSPHERE; LITHOSPHERE; PHOTOSYNTHESIS.

Water vapor concentrations in the atmosphere range from 3% by volume in wet tropical areas to a few parts per million by volume (ppmv) in the stratosphere. Water vapor, with a mean atmospheric lifetime of 10 days, is supplied to the troposphere by evaporation from the Earth's surface, and it is removed by precipitation. Because of this short lifetime, water vapor concentrations decrease rapidly with altitude, and little water vapor enters the stratosphere. Oxidation of methane represents a major

source of water vapor in the stratosphere, comparable to the source contributed by transport from the troposphere.

Carbon dioxide. The most abundant carbon species in the atmosphere is CO_2. It is produced by oxidation of organic carbon in the biosphere and in sediments. The atmospheric concentration of CO_2 has risen from 280 ppmv in preindustrial times (as determined from ice core measurements) to 355 ppmv in 1993. There is concern that this rise may cause significant warming of the Earth's surface because of the ability of CO_2 to absorb infrared radiation emitted by the Earth (that is, the greenhouse effect). *See* GREENHOUSE EFFECT.

There is a natural cycle of carbon through geochemical reservoirs (**Fig. 1**). The total amount of carbon present in the atmosphere is small compared to that present in the other geochemical reservoirs, and therefore it is controlled by exchange with these reservoirs. Equilibration of carbon between the atmosphere, biosphere, soil, and surface ocean reservoirs takes place on a time scale of decades. The carbon is eventually transported to the deep ocean, mainly by sinking of cold surface water at high latitudes. Carbon then resides in the deep ocean for approximately 700 years before upwelling to the surface. A small fraction of the carbon in the deep ocean is incorporated into sediments, where it has a lifetime of 400 million years against uplift to the surface and weathering.

Human activity over the past 100 years has disrupted the natural carbon cycle by increasing the fluxes of carbon from the sediments to the atmosphere (fossil fuel combustion) and from the biosphere to the atmosphere (land-use changes, in particular tropical deforestation). The rate of increase of atmospheric CO_2 during the 1980s was 1.8 ppmv/year, that is, 4.0×10^{12} kg (8.8×10^{12} lb) C/year. Estimated sources from fossil fuel combustion and tropical deforestation during the decade were 6.0×10^{12} kg (13×10^{12} lb) C/year and 1.6×10^{12} kg (3.5×10^{12} lb) C/year, respectively, and the estimated sink of CO_2 from net uptake by the ocean was 2.0×10^{12} kg (4.4×10^{12} lb) C/year. Mass balance requires an additional CO_2 sink of 1.6×10^{12} kg (3.5×10^{12} lb) C/year, which could possibly be contributed by accumulation of carbon in the biosphere and in soils at northern midlatitudes. The capacity of the ocean, biosphere, and soil reservoirs to take up additional carbon is a crucial issue for assessing the response of atmospheric CO_2 to changes in emissions. *See* CARBON DIOXIDE.

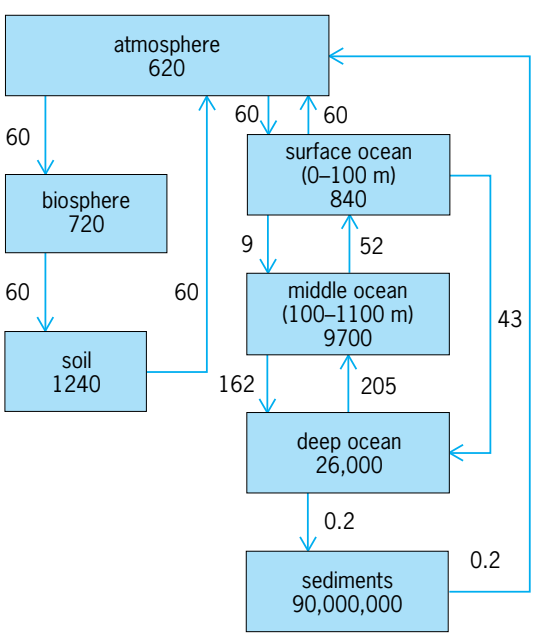

Fig. 1. Natural carbon (C) cycle. The total mass of carbon in each geochemical reservoir is in units of 10^{12} kg C; the carbon fluxes between reservoirs are in units of 10^{12} kg C/year. 1 kg = 2.2 lb. 1 m = 3.3 ft.

Other carbon species. Methane is the second most abundant carbon species in the atmosphere and an important greenhouse gas. It is emitted by anaerobic decay of biological carbon (for example, in wetlands, landfills, and stomachs of ruminants), by exploitation of natural gas and coal, and by combustion. It has a mean lifetime of 12 years against atmospheric oxidation by the hydroxyl (OH) radical, its principal sink. The concentration of methane in the atmosphere has risen from 0.8 ppmv in preindustrial times to 1.7 ppmv in 1993; this rise reflects an increase of sources over the past century, and could also reflect in part a decrease of OH concentrations. *See* METHANE.

Many hydrocarbons other than methane are emitted to the atmosphere from vegetation, soils, combustion, and industrial activities. The emission of isoprene [$H_2C{=}C(CH_3){-}CH{=}CH_2$] from deciduous vegetation is particularly significant. Nonmethane hydrocarbons have generally short lifetimes against oxidation by OH (a few hours for isoprene), so that their atmospheric concentrations are low. They are most important in atmospheric chemistry as sinks for OH and as precursors of tropospheric ozone, organic nitrates, and organic aerosols.

Carbon monoxide (CO) is emitted to the atmosphere by combustion, and it is also produced within the atmosphere by oxidation of methane and other hydrocarbons. It is removed from the atmosphere by oxidation by OH, with a mean lifetime of 2 months. Carbon monoxide is the principal sink of OH and hence plays a major role in regulating the oxidizing power of the atmosphere. Typical concentrations of CO in the troposphere range from about 50 parts per billion by volume (ppbv) in clean regions of the Southern Hemisphere to more than 200 ppbv over the polluted continents.

Nitrogen oxides. Nitrous oxide (N_2O) is of environmental importance as a greenhouse gas and as the stratospheric precursor for the radicals NO and NO_2. The principal sources of N_2O to the atmosphere are microbial processes in soils and the oceans; the main sinks are photolysis and oxidation in the stratosphere, resulting in an atmospheric lifetime for N_2O of about 130 years. The atmospheric concentration of N_2O has risen from 290 ppbv to 310 ppbv over the past century, a trend related at least in part to the growing use of nitrogen-based fertilizer.

The radical species NO and NO_2, referred to collectively as NO_x, play a major role in regulating the

concentration of ozone both in the stratosphere and in the troposphere. They are produced in the stratosphere by oxidation of N_2O, and are emitted to the troposphere by combustion, lightning, and microbial processes in soils. The lifetime of NO_x against oxidation to nitric acid is about 1 day; in the troposphere, nitric acid is removed principally by deposition, while in the stratosphere it cycles back to NO_x. Reaction of NO_x with organic radicals in hydrocarbon-rich regions of the troposphere produces organic nitrates, in particular peroxyacetylnitrate [$CH_3C(O)OONO_2$], which may be transported over long distances and eventually decompose to provide a diffuse reservoir source of NO_x. Tropospheric concentrations of NO_x range from over 1 ppbv in polluted regions with large combustion sources to 0.01–0.1 ppbv in remote regions.

The deposition of nitric acid is an important source of nitrogen nutrient to terrestrial ecosystems (in marine ecosystems, upwelling from the deep ocean dominates atmospheric deposition as a source of nitrogen). It has been argued that emission of NO_x from fossil fuel combustion could fertilize the biosphere and thus facilitate the removal of CO_2 from the atmosphere. In ecosystems where nitrogen is not a limiting nutrient, however, deposition of nitric acid causes acidification and a general decrease of vegetative productivity. *See* ECOSYSTEM; NITROGEN CYCLE; UPWELLING.

Tropospheric ozone and OH. About 90% of total atmospheric ozone resides in the stratosphere, where it is produced by photolysis of O_2. The ultraviolet photons ($\lambda < 240$ nm) needed to photolyze O_2 are totally absorbed by ozone and O_2 as solar radiation travels through the stratosphere. As a result, ozone concentrations in the troposphere (0.01–0.1 ppmv) are much lower than in the stratosphere (1–10 ppmv). *See* PHOTOLYSIS.

Tropospheric ozone plays a central role in atmospheric chemistry by providing the primary source of the strong oxidant OH. It is also an important greenhouse gas. In surface air, ozone is of great concern because of its toxicity to humans and vegetation. Ozone is supplied to the troposphere by slow transport from the stratosphere, and it is also produced within the troposphere by a chain reaction involving oxidation of CO and hydrocarbons by OH in the presence of NO_x. The chain is initiated by photolysis of ozone to produce singlet oxygen $O(^1D)$, as in reaction (1). The principal fate of $O(^1D)$ is stabilization to the ground-level triplet state $O(^3P)$ by collision with an inert molecule M [either N_2 or O_2; reaction (2)]. The $O(^3P)$ then reacts with O_2 to reform ozone [reaction (3)]. However, a small percentage of the $O(^1D)$ atoms reacts with water vapor, producing OH [reaction (4)].

$$O_3 + h\nu(\lambda < 320\,\text{nm}) \longrightarrow O_2 + O(^1D) \quad (1)$$

$$O(^1D) + M \longrightarrow O(^3P) + M \quad (2)$$

$$O(^3P) + O_2 + M \longrightarrow O_3 + M \quad (3)$$

$$O(^1D) + H_2O \longrightarrow 2OH \quad (4)$$

Once OH is produced, the chain propagates in sequence by reaction of hydrocarbons and CO with OH to yield peroxy radicals; reaction of the peroxy radicals with NO to produce NO_2 and recycle OH; and photolysis of NO_2 to produce O_3 while recycling NO. In the simplest case of the oxidation of CO, the propagation steps are as shown in reactions (5)–(8). The net reaction is given by (9). The chain is terminated by scavenging of radicals, for example reaction (10).

$$CO + OH \longrightarrow CO_2 + H \quad (5)$$

$$H + O_2 + M \longrightarrow HO_2 + M \quad (6)$$

$$HO_2 + NO \longrightarrow OH + NO_2 \quad (7)$$

$$NO_2 + h\nu(\lambda < 400\,\text{nm}) \longrightarrow NO + O(^3P) \quad (8)$$

$$O(^3P) + O_2 + M \longrightarrow O_3 + M \quad (3)$$

$$CO + 2O_2 \xrightarrow{h\nu} CO_2 + O_3 \quad (9)$$

$$HO_2 + HO_2 \longrightarrow H_2O_2 + O_2 \quad (10)$$

See CHAIN REACTION (CHEMISTRY); FREE RADICAL; TRIPLET STATE.

Ozone production by the above mechanism is particularly rapid in urban areas, where emissions of NO_x and of reactive hydrocarbons are high, resulting in unhealthful ozone levels. Regional-scale episodes of high ozone extending spatially over 10^5–10^6 km^2 (10^5 mi^2) are often observed in the United States during summer, and are thought to be primarily caused by interaction of NO_x emitted from fossil fuel combustion with isoprene emitted from vegetation. On a global scale, the above mechanism is estimated to represent a source of tropospheric ozone at least comparable in magnitude to input from the stratosphere, and perhaps much larger.

The strong oxidant OH produced by reactions (4) and (7) is the principal agent for removal of many environmentally important atmospheric species, including CO, methane, methyl bromide, hydrochlorofluorocarbons (HCFCs), NO_x, and sulfur species. Understanding the factors regulating OH concentrations is therefore a crucial issue in atmospheric chemistry. The lifetime of OH against oxidation of CO and hydrocarbons, its principal sinks, is only about 1 second; because of this short lifetime, OH concentrations are governed locally by a chemical steady state that is dependent both on solar ultraviolet radiation and on the concentrations of ozone, water vapor, NO_x, CO, and hydrocarbons. Typical daytime OH concentrations in the troposphere are of order 10^6 radicals cm^{-3}. Concentrations of OH may have declined over the past century because of anthropogenic emissions of CO and hydrocarbons, but this point is a matter of controversy. Any such decline would have been compensated at least in part by anthropogenic emissions of NO_x, resulting in increased ozone and corresponding OH production by reactions (4) and (7). There is concern that future

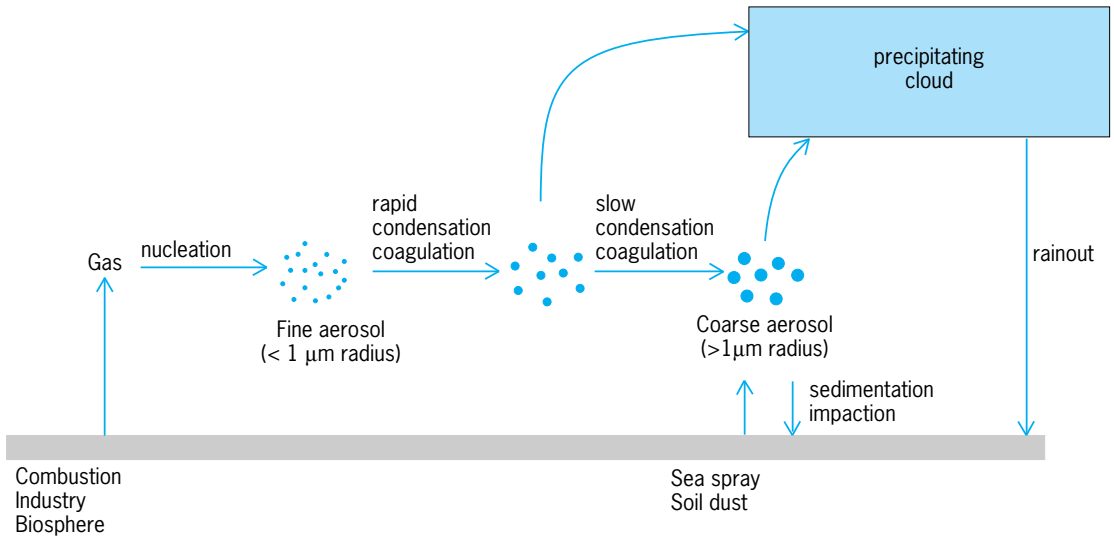

Fig. 2. Processes that determine aerosol concentrations.

changes in anthropogenic emissions of CO, hydrocarbons, and NO_x may have the potential for causing large changes in the concentration of OH and hence in the oxidizing power of the atmosphere. *See* ULTRAVIOLET RADIATION.

Atmospheric aerosols. The term aerosol refers to an ensemble of liquid and solid particles suspended in air. Aerosol number concentrations in the troposphere range from 10 to 10^6 particles cm^{-3}, with most of the aerosol mass present in particles with a radius of 0.1–10 micrometer. Aerosols are environmentally important in a number of ways. They cool the Earth by scattering solar radiation, limit visibility, serve as condensation nuclei for cloud formation, and provide media for certain chemical reactions that would otherwise not occur in the atmosphere. At the high concentrations sometimes present in urban air, aerosols may represent a significant public health risk. *See* ALBEDO.

There is a group of processes that determine aerosol concentrations (**Fig. 2**). Aerosols emitted mechanically to the atmosphere (for example, sea spray and soil dust) are called primary; they are relatively large (>1 μm radius) and deposit within a period of a few days by sedimentation or impaction on surfaces. Aerosols produced by condensation of gases (for example, sulfate and soot) are called secondary; they grow rapidly by condensation and coagulation to 0.1–1 μm radius, but further growth is slow, in part because the coagulation rate is limited by particle inertia. Secondary aerosols are removed from the atmosphere principally by rainout, with a mean lifetime in the troposphere of about 10 days. Aerosol lifetimes in the stratosphere are long, 1–2 years, because of the lack of precipitation. Volcanic aerosols injected into the stratosphere can thus effectively scatter solar radiation and have a sustained cooling effect on the Earth's climate.

Sulfuric acid produced in the atmosphere by oxidation of sulfur dioxide (SO_2) is a major component of the secondary aerosol and an important contributor to acid deposition. Sources of SO_2 to the atmosphere include emission from combustion, smelters, and volcanoes, and oxidation of oceanic dimethylsulfide [$(CH_3)_2S$] emitted by phytoplankton. It is estimated that about 75% of total sulfur emission to the atmosphere is anthropogenic. The increase of planetary albedo due to anthropogenic sulfate aerosol over the last century could possibly have caused a significant cooling of the Earth, offsetting the concurrent warming caused by rising concentrations of greenhouse gases. *See* AEROSOL; AIR POLLUTION.

Models. The general approach for computing the concentration c of a species in an atmospheric chemistry model is to solve the mass conservation equation for that species, including terms from transport, emission (E), deposition (D), chemical production (P), and chemical loss (L). The coordinate system of the model may be fixed with respect to the Earth (eulerian framework) or tied to moving parcels of air (lagrangian framework). In the eulerian framework, the mass conservation equation is written, for example, as Eq. (11), where N is the air density (both c

$$\frac{\partial c}{\partial t} = -\nabla \cdot (\mathbf{U}c) + \nabla \cdot \mathbf{K}N\nabla \left(\frac{c}{N}\right) + E - D + P - L \quad (11)$$

and N are in units of moles per unit volume). The first term on the right-hand-side represents transport by the mean circulation ($\mathbf{U}$ is the wind vector), and the second term is a simple parametrization for turbulent mixing derived by analogy with molecular diffusion ($\mathbf{K}$ is a turbulent diffusivity tensor). The P and L terms are usually functions of the concentrations of other species, so that closure may require additional mass conservation equations (one for each of the chemically interacting species). Models for atmospheric aerosols are further complicated by the need to account for a multiplicity of phases and for interactions between phases (for example, gas-particle mass

transfer and aerosol coagulation). *See* ATMOSPHERE; FLUID-FLOW PRINCIPLES. Daniel J. Jacob

Bibliography. B. J. Finlayson-Pitts and J. N. Pitts, Jr., *Chemistry of the Upper and Lower Atmosphere: Theory, Experiments, and Applications*, 1999; T. E. Graedel and P. J. Crutzen, *Atmospheric Change: An Earth's System Perspective*, 1993; J. Houghton, *Global Warming: The Complete Briefing*, 3d ed., 2004; D. Jacob, *Introduction to Atmospheric Chemistry*, 1999; J. H. Seinfeld and N. Pandis, *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*, 1997; P. Warneck, *Chemistry of the Natural Atmosphere*, 2d ed., 1999.

Atmospheric electricity

Electrical processes that take place in the lower atmosphere. This activity is ubiquitous and is of two kinds: the intense local electrification accompanying storms, and the much weaker fair-weather electrical activity occurring over the entire globe that is produced by the thousand or so electrified storms continuously in progress. The relative importance of the various mechanisms that cause storms to accumulate electrically charged particles is unknown, and the extent to which atmospheric electricity affects meteorological processes has yet to be determined. It is possible that atmospheric electricity may be important in forming precipitation and in providing energy to severe tornadoes. *See* PRECIPITATION (METEOROLOGY); TORNADO.

Disturbed-weather phenomena. The usual height of thunderstorms is about 6 mi (10 km); however, they can be as low as 2.5 mi (4 km) or as high as 13 mi (20 km). These electrified storms are characterized by heavy precipitation and strong updrafts and downdrafts with speeds as high as 65 mi/h (30 m/s).

As the result of various electrification processes, regions of charged water and ice particles accumulate in the cloud associated with the thunderstorm. In over 90% of such clouds, positive charge is in the upper part and negative charge is in the lower part, thus forming a positive, vertical dipole. When the electrified regions have dimensions of miles or kilometers and contain charges of tens of coulombs, electric fields of several hundred kilovolts per meter are produced, sufficient in intensity to cause dielectric breakdown in the form of lightning. A majority of these flashes, known as intracloud discharges, exchange charge between the upper positively charged and the lower negatively charged regions within the cloud. A minority, cloud-to-ground discharges, carry charge from the lower charged region, which is usually negative, to the ground. An average thunderstorm generates a charging current of roughly an ampere and potential differences of a hundred million volts that cause lightning flashes to occur about every 10 s, each transporting about 10 coulombs of charge.

Electrical conductivity. The atmosphere is an electrical conductor because it contains ions. These small electrically charged clusters of molecules are produced by radioactivity, solar radiation, and cosmic rays. Ions transport charge because they move under the influence of an electric field. The upper level of the atmosphere, above about 24 mi (40 km), known as the ionosphere or electrosphere, is a good conductor because of the presence of electrons and ions produced by solar and extraterrestrial ionizing radiation. In this region the air is so conductive that the time required for a charge to leak away, the electrical relaxation time, may be as little as 10^{-7} s. At lower levels, where the atmosphere is more dense and there is much less ionizing radiation, it is a much poorer conductor. At sea level, where ions are produced at the low rate of about 10 ion pairs $cm^{-3}\ s^{-1}$ by cosmic rays and terrestrial radioactivity, the atmosphere has very low conductivity. Here the electrical relaxation time can be 10^3 s or longer. If suspended cloud or aerosol particles are present, the conductivity can be further reduced, because many of the ions lose their mobility by becoming attached to these much more massive particles. In clouds, fog, or smoke the electrical relaxation time can sometimes exceed 10^4 s.

The air over thunderstorms is sufficiently conductive for a current to flow from the positive cloud top, carrying positive charge to the ionosphere. Measurements over the tops of thunderclouds show conduction currents with values ranging from one-tenth to several amperes. The frequency with which a cloud gives lightning is found to be roughly proportional to its conduction current, suggesting a relationship between this current and the electrification process that causes lightning.

The cloud is neither a sink nor a source of charge. Therefore, the current of positive charge flowing from the top of the cloud to the atmosphere is balanced by an equal transport of charge from the earth to the lower part of the cloud. This takes place through various processes, such as cloud-to-ground lightning, point discharge, falling precipitation, and conduction. Although the details are not clear, the transfer of charge within the cloud occurs as the result of intracloud lightning, the downward motion of negatively charged cloud and precipitation particles, and the upward motion of positively charged cloud particles.

The combined effect of the approximately one thousand thunderstorms in progress at any given time in the Earth's atmosphere provides an electric current of about a kiloampere that continuously deposits negative charge on the surface of the Earth and conveys an equal positive charge into the atmosphere by conduction through the ionosphere. *See* ATMOSPHERE; IONOSPHERE; THUNDERSTORM.

Fair-weather field. The currents from thunderstorms maintain a negative charge of about 1 million coulombs on the Earth and an equal net positive charge in the atmosphere, raising the ionosphere to a negative potential of about 200,000–300,000 V with respect to the Earth. Under the influence of this potential difference a vertical conduction current of about 1 microampere/km^2 carries positive charge to the Earth, giving rise to a fair-weather electric field which, because it is inversely proportional to

atmospheric conductivity, decreases exponentially with height. At sea level the field intensity is about 100 V/m and at 10 km 10 V/m, indicating 90% of the positive space charge in the atmosphere lies below 10 km (6 mi). The fair-weather field at the Earth's surface shows fluctuations over minutes or hours as a result of space charge carried by air motions and local variations in atmospheric conductivity. However, in undisturbed locations far at sea or over the polar regions, the field is less variable and a diurnal variation of about 20% can be recognized that is independent of position or local time. The variation is at a minimum a few hours after midnight and at a maximum in the late afternoon Greenwich Mean Time (GMT). This corresponds to the similar diurnal variation of thunderstorm electrical activity over the globe, which is at a minimum when the Sun is shining on the Pacific Ocean, and at a maximum when it is shining on the Earth's large landmasses.

Electricity as global index. Measuring lightning activity over large areas is accomplished by means of sferics detection, satellite observations, lightning detection networks, and evaluation of Schumann resonances. (Electromagnetic waves with frequencies of several to tens of hertz circle the Earth in the waveguide formed by the layer of air between the Earth and the ionosphere. Those waves with resonances of 8, 14, and 20 Hz are known as Schumann resonances.) The magnitude of the conduction currents flowing between thunderclouds and the upper atmosphere is closely related to the frequency of lightning. Therefore, measurements of the potential difference between the Earth and the ionosphere yield information concerning the rate that lightning is taking place over the entire globe. Because the frequency of lightning occurrence is a measure of convective activity, which increases rapidly with temperature, atmospheric electrical measurements are being used to detect and monitor global warming. *See* LIGHTNING; SFERICS.

Electrification mechanisms. Two processes give rise to the electrified cloud and precipitation particles that cause lightning. Such charged particles are produced within the cloud by charge-separation processes that take place when particles collide or break up. They are also introduced into the cloud when charge is brought to the cloud from its environment by conduction currents, convection, or lightning. To form the large regions of charged particles and the extensive electric fields that cause lightning, charged particles must be moved over large distances in opposition to electrical forces. This movement requires approximately 100 MW of power, which is supplied by falling, charged, precipitation particles and by convection that carries the charged particles in strong updrafts and downdrafts. It will not be possible to have a satisfactory understanding of the relative importance of these two charge transport mechanisms until far more is known about air motions within the cloud and details of the nature and the location of the various electrified cloud and precipitation particles. *See* CLOUD PHYSICS.

Effects of thunderstorm electricity. Electrical activity represents no more than a few percent of the total energy of a storm. It is probably for this reason that most meteorologists have assumed that atmospheric electricity plays a negligible role in weather phenomena. There is evidence, however, to suggest that electricity may be more important than formerly supposed. For example, a close correlation has been observed between the formation and dissipation of a very intense tornado and the lightning activity in the parent storm. Laboratory investigations of the effect of electric fields on the coalescence of water drops show that the electrical state of the cloud can influence the subsequent production of rain. Evidence that electricity may be important in natural clouds is afforded by radar observations that show heavy rain sometimes forms after less than a minute in the region of the cloud in which lightning occurs. Because the formation and fall of precipitation is known to produce a large effect on the dynamics of clouds and storm systems, thunderstorm electricity may, in an indirect way, produce significant effects on meteorological processes.

If the development of electric fields can affect the behavior of a cloud, it follows that variables that influence electrification will also influence the weather. Examples of such variables are extraterrestrial ionizing radiation, radioactivity, and aerosols from volcanoes or other natural and anthropogenic sources. Human activities can affect atmospheric electrical processes in several ways. Tall towers, skyscrapers, and airplanes can trigger lightning that otherwise would not have occurred. Techniques for research have been developed that initiate lightning under a thundercloud at a desired place and time by firing a rocket that trails a grounded wire. *See* WEATHER MODIFICATION.

Possible role of fair-weather electricity. There is evidence that fair-weather electricity, which is caused by thunderstorms, may in turn play a role in causing them to become electrified. Experiments show that a cloud growing from air artificially electrified with negative, instead of the usual natural, positive space charge develops into a thundercloud of inverted polarity. This finding suggests, in accord with various influence theories, that the electrification process in thunderclouds may be initiated, and its polarity determined, by the small fair-weather charges present in the atmosphere. Evidence that fair-weather atmospheric electrical variables may affect thunderstorm electrification was provided by observations of a thundercloud formed by a fire deliberately set for scientific studies (prescribed forest fire): the thundercloud grew from air that contained negative space charge, and it gave flashes that were exclusively of positive polarity.

Anomalies. With the advance of technology, observations have accumulated that reveal facets of atmospheric electricity that have yet to be explained. Pictures taken with sensitive video cameras from the ground, from airplanes, and from the space shuttle show various luminous phenomena that proceed out of the tops of active thunderclouds high into

the stratosphere or ionosphere. These rare phenomena, which have been seen but not photographed before, sometimes resemble a lightning channel and at other times an aurora or a glow discharge. Observations from lightning-detection networks show that, although most clouds produce almost exclusively negative flashes to ground, a small percentage of otherwise similar storms produce flashes of the opposite polarity. Space exploration discloses that intense electrical activity is present in the very different atmospheres of other planets. Lightning has now been reported on Venus, Jupiter, and Saturn. *See* PLANETARY PHYSICS; WEATHER. Bernard Vonnegut

Bibliography. J. A. Chalmers, *Atmospheric Electricity,* 1957; Geophysics Study Committee, Commission on Physical Sciences, Mathematics, and Resources, *The Earth's Electrical Environment,* 1986; C. Magono, *Thunderstorms,* 1980; R. Reiter, *Phenomena in Atmospheric and Environmental Electricity,* 1992.

Atmospheric entry

The motion of a body traveling from space through the atmosphere surrounding a planet. Entry bodies can be natural bodies, such as meteors or comets, or vehicles, such as ballistic missiles or the space shuttle orbiter. Entry begins when the body reaches the sensible atmosphere (defined as 400,000 ft or 122 km altitude for Earth).

The primary forces acting on an entry body are aerodynamic lift and drag, gravity, propulsion, and centrifugal acceleration. Of particular concern to the designer of entry vehicles is the control of the trajectory to minimize the effects of heating on the thermal protection system and aerodynamic loading on the vehicle structure. From Newton's law of motion, the forces on the vehicle determine the resulting trajectory as the body traverses the atmosphere. As shown in **Fig. 1**, the aerodynamic drag opposes the velocity vector. The aerodynamic lift is perpendicular to the velocity vector but not necessarily in the plane of the velocity and gravity vectors. The centrifugal force opposes the gravity vector. Any propulsive force is generally used to retard the velocity, prior to entry into the sensible atmosphere. *See* CENTRIFUGAL FORCE.

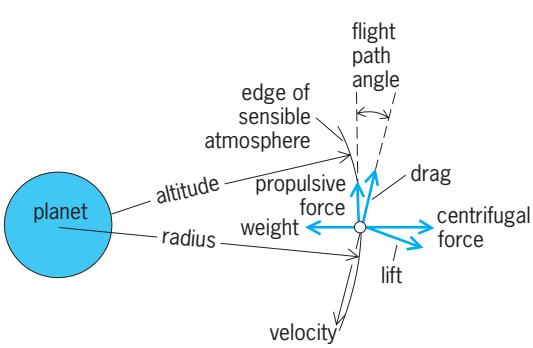

Fig. 1. Entry free-body diagram, showing forces on a vehicle entering the sensible atmosphere.

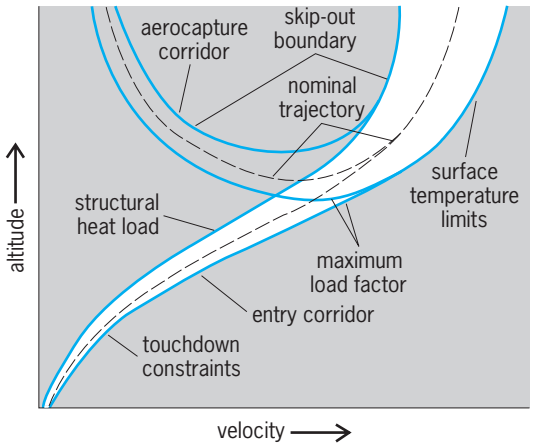

Fig. 2. Guidance corridor. Constraints on the trajectory, which determine the boundaries of the corridor, are indicated.

Aerocapture vehicles use the atmosphere to deplete energy prior to orbit capture in order to significantly reduce the amount of propellant required and thus reduce the mass requirements. Aerocapture is important for planetary exploration and will probably be used for the human exploration of the planets because of the large mass requirements of these missions.

Trajectory control. For controllable vehicles, the concept of trajectory control refers to the management of the kinetic and potential energy so as to maneuver from the initial entry conditions to the desired final conditions with due regard to system constraints and dispersions. One manner of accomplishing this is to establish a guidance corridor (**Fig. 2**). Initially, the flight path is steep enough to prevent skipping out of the atmosphere but shallow enough to keep the maximum temperature and structural load factor within limits. Later in flight, the total heat load into the structure, which increases with time of flight, becomes a constraint. Finally, touchdown or orbit capture conditions dictate orientation and velocity boundaries. The on-board guidance system attempts to follow the nominal trajectory, near the middle of the corridor, with the width of the corridor determined by dispersions. Dispersions are uncertainties, unknowns, errors, or unplanned changes in the guidance factors determining the control of the vehicle. Atmospheric dispersions on Earth are dominated by density deviations of up to 50% at higher altitudes and by surface winds associated with storms and weather fronts at lower altitudes. Other error sources include aerodynamic uncertainties and navigation uncertainties which are of particular concern for planetary aerocapture. *See* SPACE NAVIGATION AND GUIDANCE.

Aerodynamic characteristics. The aerodynamic forces arise from the effect of the airstream on the vehicle. The aerodynamic characteristics are primarily a function of the body shape, air speed, and orientation with respect to the velocity vector. The most dominant aerodynamic force is the vehicle drag, which provides the deceleration. High-drag

bodies are characterized by large, blunt reference profiles.

The lift force is perpendicular to the drag force, works perpendicular to the velocity vector, and is the primary force vector for trajectory control. Lift is modulated by changing the angle of attack or, more effectively, by rotating the lift vector about the velocity vector.

The ratio of lift to drag (L/D) determines the amount of trajectory control available. For the space shuttle this ratio is 1.1, while for the Apollo entry capsule this ratio was 0.3. The higher L/D of the shuttle gives larger ranging capability, more precise trajectory control, and point-landing capability. Depending on the navigation accuracy, aerocapture vehicles are being considered with L/D values of 0.3 to 1.5. *See* AERODYNAMIC FORCE.

Heating. Vehicles entering the atmosphere experience heat transferred from the hot air surrounding the spacecraft to the colder wall of the spacecraft. The transfer of heat or energy is accomplished by conduction, radiation, and convection. Conduction and radiation depend on a temperature difference between the fluid (air) and the vehicle wall, and convection is associated with the motion of the fluid. The temperature difference provides the potential for energy transfer or heat transfer.

A vehicle traveling at supersonic or hypersonic velocities deflects the air and forms a shock wave. The air between the detached bow shock and the vehicle is heated to very high temperatures by molecular collisions converting the kinetic energy to thermal energy. Approximately 97–98% of this energy flows past the vehicle into the free stream. The remaining 2% has to be managed by the thermal protection system on the spacecraft. The dissipation of heat depends on the vehicle's configuration and orientation to the free stream and the altitude and velocity at which the vehicle is traveling. *See* AEROTHERMODYNAMICS; HYPERSONIC FLIGHT; SHOCK WAVE; SUPERSONIC FLIGHT.

When the air flows along the surface, particles near the surface slow down through the action of viscous forces. At some distance from the surface, the velocity of the particles approaches the velocity of the undisturbed free stream. The region between the surface and some arbitrary percentage of the free-stream velocity (90–99%) is called the boundary layer, which is thin relative to the linear dimension of the vehicle. The flow within this boundary layer can be either laminar, which is smooth and orderly, or turbulent, which is irregular and characterized by eddies or fluctuating motion in the flow (**Fig. 3**). Turbulent flow has faster-moving particles and higher rates of heat transfer to the surface than laminar flow. Thus, it is desirable for entry vehicles to maintain laminar flow as long as possible to minimize the surface temperature. For aerocapture vehicles, flight near the upper corridor boundary helps to maintain laminar flow. *See* BOUNDARY LAYER FLOW; FLUID FLOW.

Heat protection systems. The advent of the space age with its progression from simple ballistic configurations, such as Mercury spacecraft, to the complex high-lift-to-drag orbiter of the Space Shuttle Transportation System has dictated development of new materials for use in the thermal protection system. Ablator materials, which were used on Mercury, Gemini, and Apollo spacecraft, accommodated the convective heating through absorption, vaporization, and the resultant char layer that reradiated the heat to the atmosphere. Ablators, however, are not reusable. *See* NOSE CONE.

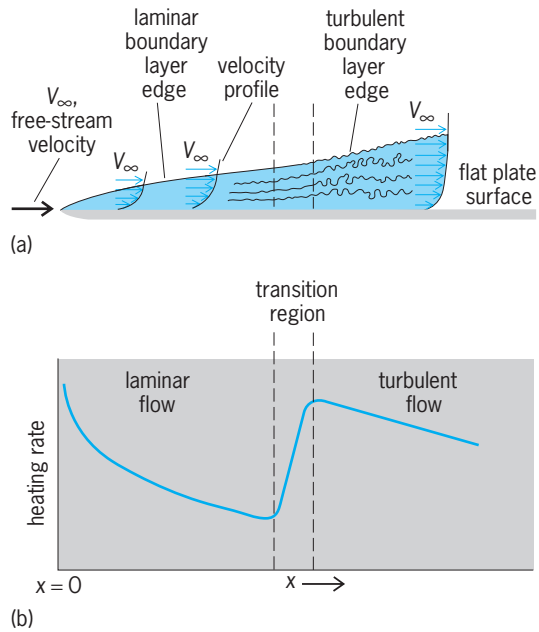

Fig. 3. Laminar and turbulent velocity profiles and heating-rate distribution on a flat plate. (*a*) Boundary layer development on a flat plate showing velocity profiles and streamlines in the boundary layer. (*b*) Surface heating rate as a function of distance *x* from flat-plate leading edge.

The reuse requirement for the shuttle orbiter necessitated development of new concepts of thermal protection. One concept is an external low-density insulator that can withstand high temperatures for multiple orbiter entries for at least 100 flights. The insulator is a tile fabricated from high-purity silica fibers reinforced with silica binder. Two temperature regimes utilize this insulator. One, designated HRSI (high-temperature reusable surface insulator), consists of 6 × 6 in. (15 × 15 cm) tiles applied to the lower surface, where temperatures vary from 1200 to 2300°F (922 to 1533 K). The other, designated LRSI (low-temperature reusable surface insulator), consists of 8 × 8 in. (20 × 20 cm) tiles on the side fuselage and upper wing, where temperatures vary from 700 to 1200°F (644 to 922 K). The materials are the same but have different coatings. These coatings are both waterproof borosilicate glass, but the HRSI coating contains black pigment for high emittance at high temperatures and the LRSI tiles have a white coating for low solar absorptance. *See* HIGH-TEMPERATURE MATERIALS.

Another insulator, FRSI (flexible reusable surface insulation), consists of 3 × 4 ft (0.9 × 1.2 m) sheets of Nomex felt and is bonded to areas on the upper

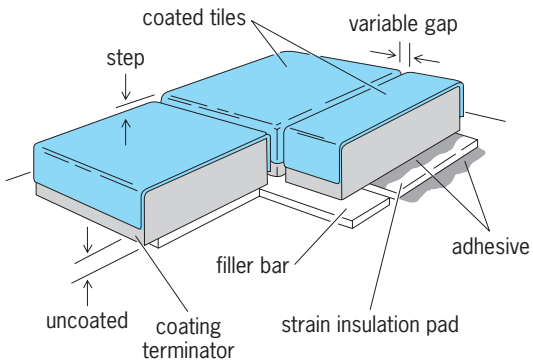

Fig. 4. Thermal-protection-system tile configuration on the shuttle orbiter. (*After R. L. Dotts, D. M. Curry, and D. J. Tillian, The shuttle orbiter thermal protection system materials, design and flight performance overview, 13th Intersociety Conference on Environmental Systems, San Francisco, July 11–15, 1983, SAE Tech. Pap. Ser., 831118, 1983*)

surface where temperatures are less than 700°F (644 K). Nomex felt is also used as the strain isolation pad, between the tiles and the structure. The tile configuration is shown in **Fig. 4**. The filler bar is also coated nomex felt, and the adhesive is room-temperature vulcanizing silicone.

The temperature on the wing leading edges and on the nose of the shuttle orbiter was predicted to exceed 2300°F (1533 K). A reinforced carbon-carbon system was developed to withstand up to 3000°F (1922 K). Reinforced carbon-carbon is a laminate of woven graphite cloth with a carbon binder and a silicon carbide coating which prevents oxidation. The locations of the various types of thermal protection on the surface of the orbiter are shown in **Fig. 5**. *See* COMPOSITE MATERIAL.

These insulating systems are installed as smooth surfaces so that tile-to-tile displacements and gap widths between the tiles are minimized. Stringent requirements of thermal protection system installation are imposed on the manufacturer to assure laminar boundary-layer flow during the significant heating portion of the entry trajectory.

Most of the heat protection systems that have been studied for planetary aerocapture vehicles are similar to the entry systems described above. However, the entry velocity for some of the return missions is too high for the use of tiles, and ablative heat shields will probably be required. *See* BALLISTIC MISSILE; MISSILE; SPACE PROBE; SPACE SHUTTLE; SPACECRAFT STRUCTURE. Richard L. Barton; Dorothy B. Lee; Joe D. Gamble

Bibliography. A. J. Chapman, *Fundamentals of Heat Transfer*, 1987; M. J. Griffin and J. R. French, *Space Vehicle Design*, 2d ed., 2004; W. Hankey, *Re-Entry Aerodynamics*, 1988; F. J. Regan and S. M. Anandakrishnan, *Dynamics of Atmospheric Entry*, 1993.

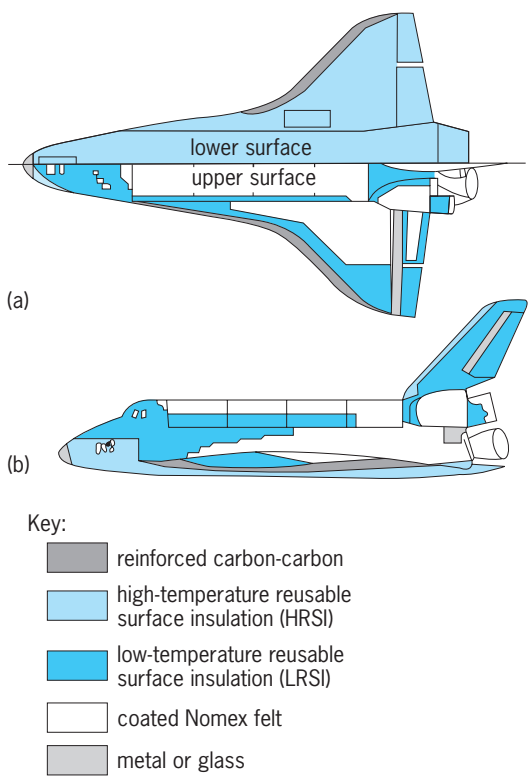

Fig. 5. Location of various types of thermal protection on the surface of the shuttle orbiter. (*a*) Top view of upper surface and bottom view of lower surface. (*b*) Side view. (*After R. L. Dotts, D. M. Curry, and D. J. Tillian, The shuttle orbiter thermal protection system materials design and flight performance overview, 13th Intersociety Conference on Environmental Systems, San Francisco, July 11–15, 1983, SAE Tech. Pap. Ser., 831118, 1983*)

Atmospheric general circulation

The statistical description of atmospheric motions over the Earth, their role in transporting energy, and the transformations among different forms of energy. Through their influence on the pressure distributions that drive the winds, spatial variations of heating and cooling generate air circulations, but these are continually dissipated by friction. While large day-to-day and seasonal changes occur, the mean circulation during a given season tends to be much the same from year to year. Thus, in the long run and for the global atmosphere as a whole, the generation of motions nearly balances the dissipation. The same is true of the long-term balance between solar radiation absorbed and infrared radiation emitted by the Earth–atmosphere system, as evidenced by its relatively constant temperature. Both air and ocean currents, which are mainly driven by the winds, transport heat. Hence the atmospheric and oceanic general circulations form cooperative systems. *See* MARITIME METEOROLOGY; OCEAN CIRCULATION.

Owing to the more direct incidence of solar radiation in low latitudes and to reflection from clouds, snow, and ice, which are more extensive at high latitudes, the solar radiation absorbed by the Earth–atmosphere system is about three times as great in the equatorial belt as at the poles, on the annual average. Infrared emission is, however, only about 20% greater at low than at high latitudes. Thus in low latitudes (between about 35°N and 35°S) the Earth–atmosphere system is, on the average, heated and in higher latitudes cooled by radiation. The Earth's

surface absorbs more radiative heat than it emits, whereas the reverse is true for the atmosphere. Therefore, heat must be transferred generally poleward and upward through processes other than radiation. At the Earth–atmosphere interface, this transfer occurs in the form of turbulent flux of sensible heat and through evapotranspiration (flux of latent heat). In the atmosphere the latent heat is released in connection with condensation of water vapor. *See* CLIMATOLOGY; HEAT BALANCE, TERRESTRIAL ATMOSPHERIC.

Considering the atmosphere alone, the heat gain by condensation and the heat transfer from the Earth's surface exceed the net radiative heat loss in low latitudes. The reverse is true in higher latitudes. The meridional transfer of energy, necessary to balance these heat gains and losses, is accomplished by air currents. These take the form of organized circulations, whose dominant features are notably different in the tropical belt (roughly the half of the Earth between latitudes 30°N and 30°S) and in extratropical latitudes. *See* METEOROLOGY; STORM.

Principal circulations. Characteristic circulations over the Northern Hemisphere are sketched in **Fig. 1**. In the upper troposphere, there are two principal jet-stream systems: the subtropical jet (STJ) near latitude 30°, and the polar-front jet (PFJ), with large-amplitude long waves and superimposed shorter waves associated with cyclone-scale disturbances. The long waves on the polar-front jet move slowly eastward, and the shorter waves move rapidly. At the Earth's surface, northeast and southeast trade winds of the two hemispheres meet at the intertropical convergence zone (ITCZ), in the vicinity of which extensive lines and large clusters of convective clouds are concentrated. Westward-moving waves and vortices form near the intertropical convergence zone and, in summer, within the trades. Heat released by condensation in convective clouds of the intertropical convergence zone, and the mass of air conveyed upward in them, drive meridional circulations (right of Fig. 1), whose upper-level poleward branches generate the subtropical jet stream at their poleward boundaries. *See* TROPICAL METEOROLOGY.

In extratropical latitudes, the circulation is dominated by cyclones and anticyclones. Cyclones develop mainly on the polar front, where the temperature contrast between polar and tropical air masses is concentrated, in association with upper-level waves on the polar-front jet stream. In winter, cold outbreaks of polar air from the east coasts of continents over the warmer oceans result in intense transfer of heat and water vapor into the atmosphere. Outbreaks penetrating the tropics also represent a sporadic exchange in which polar air becomes transformed into tropical air. Tropical airstreams,

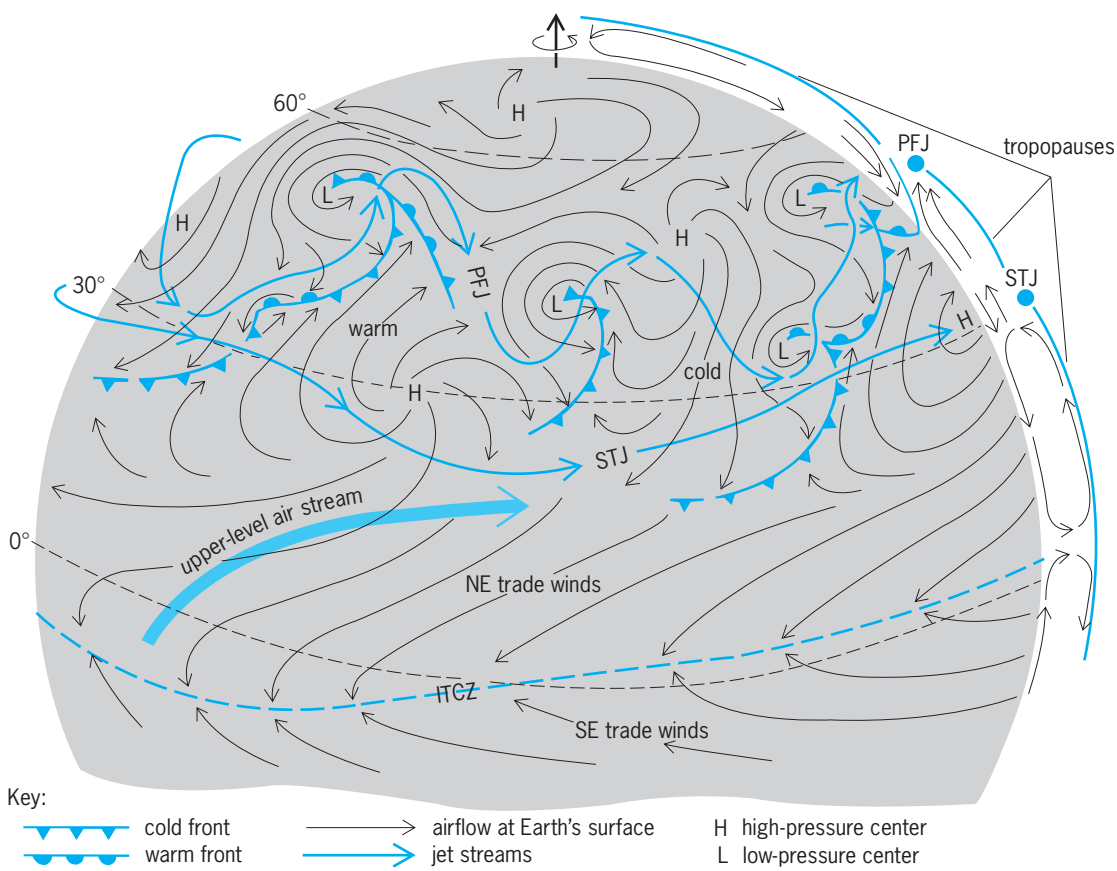

Fig. 1. Schematic circulations over the Northern Hemisphere in winter. The intertropical convergence zone (ITCZ) lies entirely north of the Equator in the summer. Eastward acceleration in the upper-level tropical airstream is due to Earth rotation and generates the subtropical jet stream (STJ). The vertical section (right) shows the dominant meridional circulation in the tropics and shows airstreams relative to the polar front in middle latitudes.

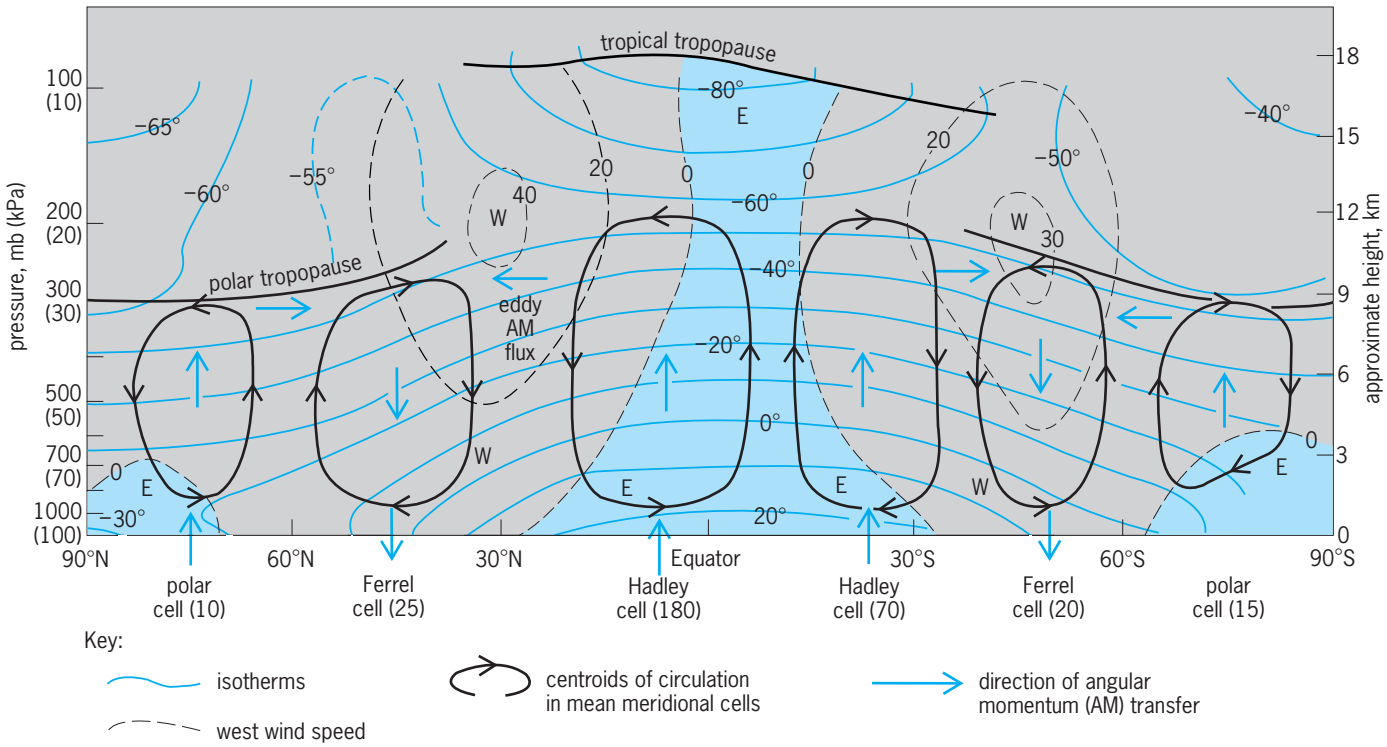

Fig. 2. Pole-to-pole section (450-fold vertical exaggeration) in January showing zonally averaged temperature in °C and wind speed in m/s; W and E denote westerlies and easterlies, respectively. Mean meridional cells are named and their intensities are given in terms of mass flow in megatons per second. °F = (°C × 1.8) + 32. 1 m/s = 2 knots. 1 km = 0.6 mi.

poleward on the west sides of the subtropical highs, then supply heat and water vapor to the extratropical disturbances. *See* CYCLONE; FRONT.

The characteristic flow in cyclones takes the form of slantwise descending motions on their west sides and ascent to their east in which extensive clouds and precipitation form. Heat that is released in condensation drives the ascending branch, and the descending branch consists of polar air that has been cooled by radiation in higher latitudes. When viewed relative to the meandering polar-front zone (right of Fig. 1), the combined sinking of cold air and ascent of warm air represents a conversion of potential energy into kinetic energy. This process maintains the polar jet stream. The branches of the circulation transfer heat both upward, to balance the radiative heat loss by the atmosphere, and poleward, to balance the radiative heat deficit in high latitudes.

Mean temperature and wind. A principal object of general circulation studies has been to explain the zonal mean (long-term average around latitude zones) structure of the atmosphere, shown for January in **Fig. 2**. The zonal (west-east) wind component is almost everywhere dominant and is in quasi-geostrophic balance with the mean meridional pressure gradient. The pressure gradient changes with height in accordance with the distribution of air density, which at a given pressure is inverse to temperature. Hence the distribution of zonal wind is related to that of temperature, as expressed by Eq. (1), where

$$\frac{\partial \overline{u}}{\partial z} = -\frac{g}{(2\Omega \sin \phi)\overline{T}} \frac{\partial \overline{T}}{\partial y} \qquad (1)$$

u denotes zonal wind component; z, height above sea level; g, acceleration of gravity; Ω, angular velocity of the Earth; ϕ, latitude; T, Kelvin temperature; and y, distance northward. Overbars denote values averaged over longitude and time. *See* GEOSTROPHIC WIND.

Only in the lowest kilometer or so, where surface friction disturbs the geostrophic balance, and in the vicinity of the Equator is the mean meridional (south-north) component comparable to the zonal wind. Because of the nature of the atmosphere as a shallow layer, the mean vertical wind component is weak. Whereas the magnitude of the mean zonal wind varies 100 mi/h (45 m/s), and the mean meridional wind up to 6.5 mi/h (3 m/s), the mean vertical wind nowhere exceeds 0.02 mi/h (1 cm/s). The vertical component cannot be observed directly, but can be calculated from the distribution of horizontal motions.

In the troposphere and lower stratosphere, the zonal circulation is similar in winter and summer, with easterlies in low latitudes and westerlies in higher latitudes, except in small regions of low-level easterlies around the poles. The strongest west winds are, in the winter hemispheres, observed near latitude 30° at about 7 mi (12 km). In summer, the west-wind maxima are weaker and located farther poleward. *See* STRATOSPHERE; TROPOSPHERE.

In the troposphere, the zonal wind increases upward according to Eq. (1) and with a general poleward decrease in temperature. In the lower stratosphere over most of the globe, this temperature gradient is reversed, and the wind decreases with

height. Above about 12 mi (20 km), separate wind systems exist, with prevailing easterly winds in summer and westerlies in winter that attain speeds up to 130–175 mi/h (60–80 m/s) near 40 mi (60 km) height in high latitudes. *See* JET STREAM.

The much weaker meridional circulation consists of six separate cells. Their general locations and nomenclatures are shown in Fig. 2, along with the approximate circulations in terms of mass flux. For each cell, only central streamlines are shown, but these represent flows that are several kilometers deep in the horizontal branches, while each vertical branch represents gentle ascending or descending motions over latitude belts some thousands of kilometers wide. The tropical Hadley cells, best developed in the winter hemisphere, are mainly responsible for maintaining the westerly winds as described below.

Angular momentum balance. The motion of the atmosphere plus the eastward speed of a point on the Earth's surface represent the total motion relative to the Earth's axis. The angular momentum (M) of a unit mass of air is given by Eq. (2), where a represents

$$M = (u + \Omega a \cos\phi) a \cos\phi \qquad (2)$$

the radius of the Earth and the other terms are as defined for Eq. (1). Considering the mean value in a zonal ring, this quantity is conserved unless the ring is subjected to a torque.

The surface easterlies in low latitudes and westerlies in middle latitudes (Fig. 2) exert the principal torques upon the atmosphere, due to frictional drags that are opposite to the direction of the surface winds.

Since the torques would tend to diminish the westerlies and easterlies, and this is not observed to occur over the long run, it follows that angular momentum has to be transferred from the belts of surface easterlies to the zones of westerlies. Calculations show that this meridional flux occurs mainly in the upper troposphere. Hence angular momentum has to be brought upward from the surface layer in low latitudes, transferred poleward, and ultimately brought down to the Earth in the belt of westerlies (Fig. 2).

The vertical transfers are mainly accomplished by the mean meridional circulations, involving the second term on the right-hand side of Eq. (2), which is much larger than the first. Considering the $\cos^2\phi$ factor in this term, there is a net upward transfer by the Hadley cells since their ascending branches are at lower latitudes than the descending branches. Similarly, the opposite Ferrel circulations bring angular momentum downward. With conservation of total angular momentum in the poleward upper branches of the Hadley cells, u in Eq. (2) increases at the expense of the Ω term. This process accounts for the generation of the strongest upper-level westerly winds in subtropical latitudes. From this source, eddies (waves in which poleward flow is associated with a stronger west-wind component than in equatorward flow) transfer angular momentum poleward to sustain the west winds at higher latitudes. *See* ANGULAR MOMENTUM.

Heat energy balance. Maintenance of an average temperature distribution such as in Fig. 2 depends upon a balance between the effects of heat sources and sinks (for example, radiation) and of heat transport by air motion. The sources and sinks are largely a function of latitude and elevation, such that meridional and vertical fluxes of heat energy are required. Two methods by which these fluxes are calculated from different kinds of observations give results that offer a check upon one another.

In the first method estimates are made of the rate of change of any property X per unit area of the Earth's surface due to sources and sinks. To maintain an unchanged condition, their integrated value, over the area north of a latitude ϕ, must equal the northward flux F_ϕ of the property across the latitude circle. This requirement is expressed by Eq. (3), in which t denotes time.

$$F_\phi = -2\pi a^2 \int_\phi^{90^\circ N} \frac{d\overline{X}}{dt} \cos\phi \, d\phi \qquad (3)$$

By employing this method, the energy balance may be calculated by substituting the quantities summarized in **Table 1** into Eq. (3). The listed heat sources comprise R_a, net (absorbed minus emitted) radiation of the atmosphere; R_e, net radiation at the Earth's surface; Q_s, flux of sensible heat from the surface to the atmosphere; LE, flux of latent heat from the surface (E denoting rate of evapotranspiration and L, heat of vaporization); and LP, release of latent heat in the atmosphere as estimated from the observed rate of precipitation P.

The second method is to compute the fluxes directly from aerological observations (made by balloon-borne radiosondes). If x denotes a given property per unit mass of air, the flux is given by Eq. (4), where v is the meridional wind component

$$F_\phi = \frac{2\pi a \cos\phi}{g} \int_0^{p_0} (\overline{x}\,\overline{v} + \overline{x'v'}) \, dp \qquad (4)$$

(positive northward). The integration, with pressure p as a vertical coordinate (related to height z as in Fig. 2), is extended from the bottom ($p = p_0$) to the top of the atmosphere ($p = 0$). Here $\overline{x}$ and $\overline{v}$ denote values averaged over time and longitude, and x' and v' are deviations from these mean values at a given pressure surface. A corresponding expression can be written for the vertical fluxes.

The forms of heat energy x are listed in **Table 2**. The atmospheric energy comprises the sum $c_pT + gz$, the two quantities being interchangeable during

TABLE 1. Sources of atmospheric properties

Different sources	$d\overline{X}/dt$
Atmospheric heat	$\overline{R}_a + \overline{Q}_s + L\overline{P}$
Latent heat	$L(\overline{E} - \overline{P})$
Heat of Earth's surface	$\overline{R}_e - \overline{Q}_s - L\overline{E}$
Heat of atmosphere and Earth	$\overline{R}_a + \overline{R}_e$

TABLE 2. Atmospheric properties used for flux computations

Property	x (per unit mass)*
Atmospheric energy	
Sensible heat	C_pT
Potential energy	gz
Total atmospheric energy	$C_pT + gz$
Latent heat	Lq

*Here c_p is specific heat of air; q is specific humidity, or mass of water vapor per unit mass of air.

vertical air movement; decompression of rising air diminishes its sensible heat content by an amount equal to the increase of potential energy associated with change of elevation. Almost everywhere $c_pT + gz$ increases upward, and Lq generally but not always decreases upward.

The first term of Eq. (4) represents the mean meridional circulations in Fig. 2, which dominate in the tropical belt. Water vapor conveyed equatorward by the trade winds is condensed in tall convective clouds near the intertropical convergence zone. As a result of latent heat release, the atmospheric energy ($c_pT + gz$) is augmented, so that the upper-level flow of a Hadley cell carries more heat poleward than does the equatorward low-level flow. The second term represents eddy fluxes, dominant in subtropical and higher latitudes, associated with cyclones and anticyclones (Fig. 1) whose poleward flows are both warmest and richest in water vapor.

Various components of the energy balance are summarized, by 30° belts of latitude, in **Fig. 3**. Of the net radiation absorbed at the Earth's surface over the whole globe, 81% is expended in evaporation. Correspondingly, 81% of the net radiative loss by the atmosphere is compensated by release of latent heat when water vapor condenses and falls out as rain, snow, or hail, and 19% by transfer of sensible heat from the Earth. In the tropical belt 30°N–30°S, the Earth–atmosphere system gains heat by radiation; the excess is exported to higher latitudes as atmospheric heat and latent heat, and by ocean currents. Considering the tropical-latitude and temperate-latitude belts of the two hemispheres, significant differences in the apportionments of evaporation and sensible heat transfer from the Earth to the atmosphere, and of the meridional transports of energy in the various forms, arise from the greater dominance of continents in the Northern Hemisphere. On the annual average, water-vapor-laden trade winds (the lower branches of Hadley cells in Fig. 2) converge at about 5°N, where the greatest precipitation is observed. Minima of precipitation occur in the belts near 30°N, where mean descending motions occur. Secondary maxima of precipitation, in latitudes 40–50°, are associated with frequent extratropical cyclones. *See* HYDROMETEOROLOGY.

The frequency and intensity of cyclones, and the contrasts between their cold and warm air masses, are much greater in winter than in summer in the Northern Hemisphere; these variations are much less pronounced in the oceanic Southern Hemisphere. Thus there is a fourfold greater poleward transport of sensible heat in middle latitudes of the Northern Hemisphere in winter than in summer, contrasted with only about a 30% seasonal variation in the Southern Hemisphere. In the tropics, large seasonal changes in the intensities of the Hadley circulation, together with a migration of the rain belt of the intertropical convergence zone (Fig. 1), are most pronounced in the monsoon regions of Asia-Australia and Africa. *See* MONSOON METEOROLOGY.

Between late spring and early autumn a given hemisphere receives solar radiation far in excess

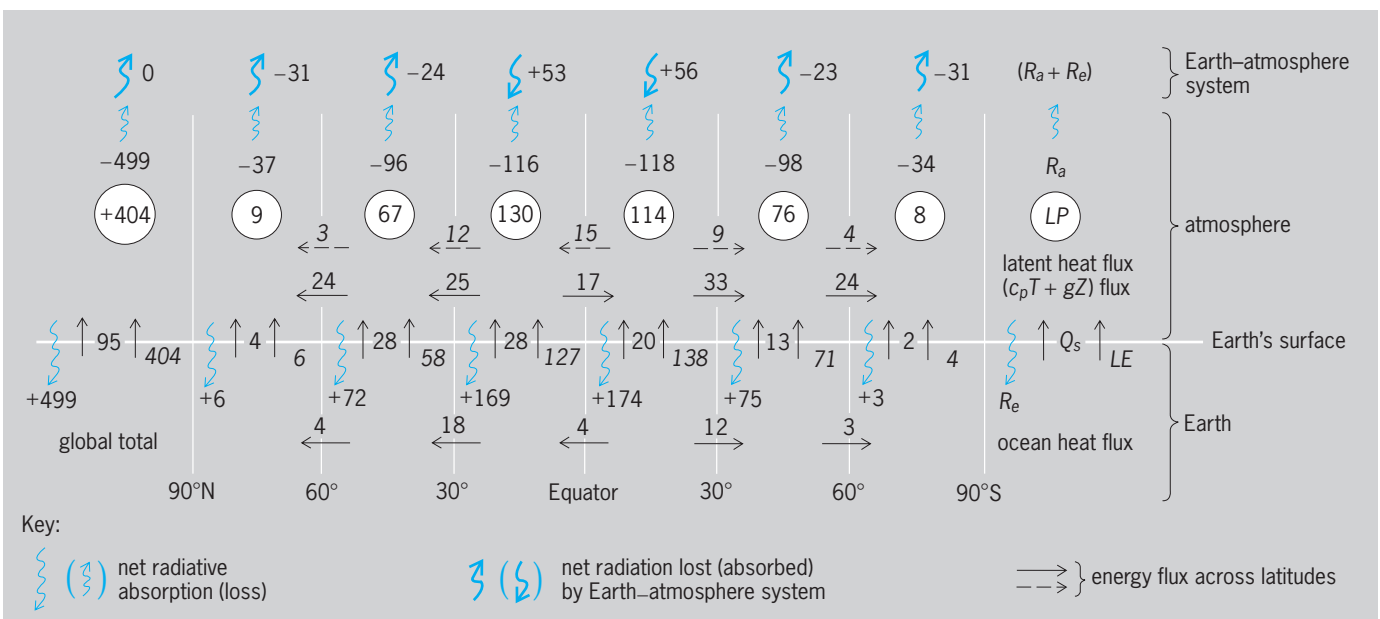

Fig. 3. Annual heat balance for the Earth as a whole and for each 30° latitude belt. Units are 10^{14} W. All values are keyed to the column at right. Italic numbers apply to water vapor (latent heat) flux.

of the amount it loses by infrared radiation. The excess heat is stored mainly in the oceans during the warm seasons and given up to the atmosphere as sensible and latent heat during the cooler seasons. Thus the oceans serve as an energy reservoir that tempers the seasonal changes of atmospheric temperature over them and over neighboring land areas invaded by marine air masses. *See* MARITIME METEOROLOGY. C. W. Newton

Bibliography. J. R. Holton, *An Introduction to Dynamic Meteorology,* 3d ed., 1992; J. P. Peixóto and A. H. Oort, Physics of climate, *Rev. Mod. Phys.*, 56:365–429, 1984; S. Petterssen, *Introduction to Meteorology*, 3d ed., 1969.

Atmospheric tides

Those oscillations in any or all atmospheric fields whose periods are integral fractions of either lunar or solar days. Oscillations with a period of a day are called diurnal, with a period of a half day semidiurnal, and with a period of a third of a day terdiurnal. A common nomenclature does not exist for smaller fractions of a day. The sum of all tidal variations is referred to as the daily variation. As a practical matter, the subject of atmospheric tides is generally restricted to oscillations on a global spatial scale (thus excluding sea breezes). Moreover, the bulk of attention is devoted to migrating tides, which are those tidal oscillations that depend only on local time.

Atmospheric tides tend to be rather small in the troposphere, although the tidal oscillations in rainfall are surprisingly large. Their importance stems from two primary factors: (1) Tidal oscillations tend to increase in amplitude with height and become major components of the total meteorology above about 50 km (30 mi). (2) The subject has played a prominent role in the intellectual history of meteorology, and it still provides a remarkable example of scientific methodology in an observational science. In connection with the second item, it should be noted that tides are unique among meteorological systems in that they have perfectly known periods and relatively well known sources of forcing (though the determination of these sources constitutes a major part of the history of the subject).

Observational description. The determination of an oscillation by means of data requires at least two measurements per period. Since most meteorological upper air soundings are taken only twice each day, such data can be used only to marginally determine diurnal oscillations. Occasionally, stations obtain soundings four times per day, which in turn permits determinations of semidiurnal oscillations. Rain gages assign rainfall to specific hours, and averages over many years allow the determination of the daily variation of rainfall. Surface pressure is monitored at a great many stations with effectively (from the point of view of tidal analyses) continuous time resolution. Therefore, surface pressure has traditionally been the field most thoroughly analyzed for tides.

To a fair degree of approximation, the distributions of solar semidiurnal (S_2) and diurnal (S_1) tides in surface pressure (P_s) are given by Eqs. (1) and (2), where

$$S_2(P_s) = a \sin^3\theta \sin(2t + 158^\circ) + bP_2(\theta)\sin(2t_u + 118^\circ) \quad (1)$$

$$S_1(P_s) = c \sin^3\theta \sin(t + 12^\circ) \quad (2)$$

$a = 1.16$ millibar (116 pascals), $b = 0.085$ millibar (8.5 Pa), $c = 0.593$ millibar (59.3 Pa), θ = colatitude, t = local time t_u = Greenwich (Universal) Time, and $P_2\theta \equiv {}^1/_2\cos^2\theta - 1)$.

The first term on the right-hand side of Eq. (1) represents a migrating semidiurnal tide with maximum surface pressure near 10 A.M. and 10 P.M. local time, while the second term represents a standing oscillation in Universal Time. The diurnal tide is primarily migrating with maximum pressure occurring near 5 A.M. local time. The lunar semidiurnal tide in surface pressure is similar in distribution to the migrating part of the solar semidiurnal tide but only about one-twentieth its strength; maximum lunar semidiurnal surface pressure typically occurs about 1 lunar hour and 13 lunar hours after lunar transit. Clearly, the solar semidiurnal tide dominates the surface pressure. The solar diurnal component is not only smaller but also far more irregular.

Rainfall is commonly observed to have a daily variation. The diurnal component, though often quite large, has a very irregular phase; on the other hand, the solar semidiurnal component is surprisingly uniform, amounting to about 10–20% of the mean daily rainfall in the tropics with maximum rainfall at about 4 A.M. and 4 P.M. local time. Maximum semidiurnal rainfall appears to occur somewhat later in middle latitudes. *See* PRECIPITATION (METEOROLOGY).

Data become more sparse when attempts are made to analyze tides above the surface. Analyses of radiosonde wind measurements have shown that solar semidiurnal oscillations in horizontal wind are primarily in the form of migrating tides. Diurnal oscillations, on the other hand, are significantly affected by regional, nonmigrating components up to at least 20 km (12 mi). Above this height, the diurnal oscillations also tend to be dominated by migrating components. There is a tendency for the diurnal oscillations to have more phase variation with height, especially at low latitudes. The **illustration** shows the approximate variation of tidal amplitudes with height. As a rough rule of thumb, oscillations in temperature tend to have magnitudes in kelvins comparable to the amplitudes in wind in meters per second. The illustration shows the general increase in amplitude with height. There is also no longer a clear dominance of the semidiurnal oscillations over the diurnal oscillations once the upper-level fields are considered. The amplitude increase with height renders the detection of tidal oscillations at greater altitudes somewhat easier since the tides are becoming a larger (and sometimes dominant) feature of the total fields. *See* OSCILLATION.

Meteorological rocket soundings have permitted extension of the knowledge about tides to about a 60-km (36-mi) altitude. Below an altitude of 50 km

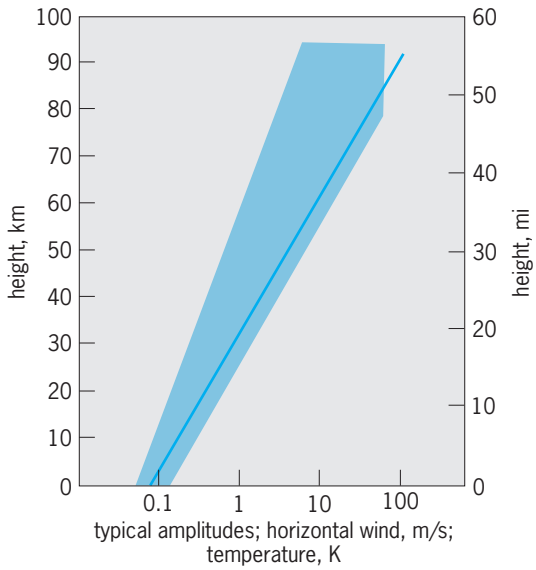

Variation of tidal amplitudes with height. The shaded area shows the general range of solar diurnal and semidiurnal amplitudes (on a logarithmic scale) as a function of height. The line corresponds to amplitude varying inversely with the square root of the basic pressure. In classical tidal theory, amplitudes scale by this factor, which leads to energy density being independent of height.

(30 mi), semidiurnal horizontal wind oscillations tend to be smaller than 1–2 m/s; however, diurnal oscillations are appreciably stronger, reaching 3–4 m/s near 40 km (24 mi) and rising to about 6–10 m/s above 50 km (30 mi). Above 50 km (30 mi) the semidiurnal oscillations also increase to about 6 m/s. Diurnal oscillations equatorward of 30° latitude tend to be associated with rapid vertical phase variations; poleward of 30° these phase variations tend to be absent. Semidiurnal oscillations tend to have no phase variations with height, except for a 180° phase shift someplace between 30 and 50 km (18 and 30 mi). Data above 60 km (36 mi) are obtained from larger rockets and satellite observations, as well as from a variety of radar techniques. Amplitudes are generally found to still increase with height up to about 80 or 90 km (48 or 54 mi). At 80 km (48 mi), amplitudes of as much as 40–50 m/s are found. At this height, diurnal oscillations tend to be dominant equatorward of 30°, while semidiurnal oscillations tend to be dominant poleward of 30°. The diurnal tidal contributions to the temperature come close to producing static instability, and may be an important source of turbulence in the upper atmosphere. Semidiurnal tides begin to show evidence of more phase variation with height. There tends to be a minimum in diurnal amplitudes near about 105 km (65 mi); within the thermosphere, tidal amplitudes are typically large, increasing asymptotically to constant values above 200 km (120 mi). These values are commonly in excess of 100 m/s and 100 K. *See* CONVECTIVE INSTABILITY; DYNAMIC INSTABILITY; METEOROLOGICAL ROCKET.

Historical considerations. The behavior of tides in the surface pressure was already well known by the nineteenth century. Ocean tides were known to be dominated by the lunar semidiurnal component. This was consistent with their being forced primarily by the tidal component of the gravitational potential. Here, lunar forcing is stronger than solar forcing (because the radius of the Earth is a much larger fraction of the lunar distance than the solar distance, and because gravitational forcing is primarily semidiurnal). The fact that atmospheric tides are primarily solar rather than lunar suggests a thermal origin. Lord Kelvin, however, noted that thermal forcing would be primarily diurnal rather than semidiurnal. It was thus puzzling that the surface pressure tide was primarily semidiurnal. He put forth the hypothesis that the atmosphere might be resonant at 12 h, and this might explain the preferential semidiurnal response.

The search for resonance was the primary focus of theoretical research on atmospheric tides through the 1950s. It was recognized that the resonance properties of the atmosphere were dependent on the thermal structure of the atmosphere. The high point in the search occurred in 1937, when C. L. Pekeris discovered that the thermal structure needed for solar semidiurnal resonance was almost identical to the independently inferred structure based on observations. The history of this inquiry had followed almost perfectly the paradigm for research in an observational science. Observations of tides had led to a hypothesis whose development led to independent implications (namely, the thermal structure of the atmosphere), which could be confirmed or refuted by additional measurements. Success, however, does not rigorously prove the original hypothesis; indeed, in the years following World War II, both observational and theoretical advances led to the rejection of the resonance hypothesis. Specific thermal sources of forcing (namely, insolation absorption by both water vapor and ozone) have proven adequate to account for observed tidal fields below about 100 km (60 mi) without any appeal to resonance. The agreement is by no means perfect, and it now appears that the latent heating associated with the daily variation in rainfall may constitute an additional important source of forcing.

Theory. The classical theory of atmospheric tides is based on a linearization of the equations of motion for a rotating spherical inviscid atmosphere about a state of rest. The resulting equations allow simple harmonic dependences on time and longitude, and are separable in their dependence on height and latitude. Latitude dependence is described by Laplace's tidal equation, which, for a given frequency and wave number, can be solved for an infinite number of eigenfunctions (known as Hough functions) and their associated eigenvalues (known as equivalent depths). Height dependence is described by the so-called vertical structure equation. This inhomogeneous equation includes (in either the equation or its boundary conditions) the gravitational and thermal forcing for each Hough mode, which is the complete solution for all physical fields (that is, horizontal and vertical wind, temperature, pressure, and density)

including their vertical structure associated with a given Hough function. The nature of the vertical response is determined by both the thermal structure of the atmosphere and the mode's equivalent depth. As a rule, the vertical structure equation yields only the trivial zero solution when there is no forcing. However, for a few equivalent depths it is sometimes possible to get nontrivial solutions in the absence of forcing; these are then said to be resonant equivalent depths. *See* EIGENFUNCTION.

The nature of the Hough functions and equivalent depths is significantly determined by the ratio of the oscillation frequency to the quantity 2Ω (the Coriolis parameter at the poles where Ω is the Earth's rotation rate). When this quantity is greater than 1, there are no latitudes where the frequency equals the local Coriolis parameter ($2\Omega \sin \phi$, where ϕ is the latitude). The Hough functions in these cases are closely related to spherical harmonics, and the equivalent depths are positive. Positive equivalent depths are generally associated with wavelike propagation in the vertical, though sufficiently large equivalent depths lead to vertical trapping (that is, exponential as opposed to wavelike behavior). The situation is markedly different when this ratio is less than 1. Then, two latitudes exist where the frequency equals the local Coriolis parameter. Beyond these latitudes the Coriolis parameter exceeds the frequency, and this, it turns out, inhibits vertical propagation. The way that this manifests itself for Laplace's tidal equation is that the eigensolutions divide into two classes: one with amplitudes concentrated poleward of these critical latitudes and having negative equivalent depths (which are always associated with trapping in the vertical), and the other with amplitudes concentrated equatorward of these critical latitudes and having positive equivalent depths. With reference to atmospheric tides, the critical latitudes for semidiurnal modes are essentially the poles, and all modes have positive equivalent depths. However, for diurnal modes the critical latitudes occur at 30°, and the eigenfunctions are divided into the two categories described above. As it turns out, the gravest diurnal mode, into which the bulk of the thermal forcing goes, has a negative equivalent depth. The so-called graveness of a mode is determined by the number of zeros that the associated Hough function has; the gravest mode is that mode for which the associated Hough function has no zeros except at the poles. Thus the effect of upper-level thermal forcing is unable to effectively penetrate to the surface. This is a primary reason for the dominance of the semidiurnal oscillation at the ground. The confinement of positive equivalent-depth diurnal modes to the region equatorward of 30° also accounts for the relative dominance of diurnal oscillations in this region in the neighborhood of 80 km (48 mi) altitude. *See* CORIOLIS ACCELERATION.

While the classical theory of atmospheric tides is adequate for many purposes, recent years have seen a substantial development of theory well beyond the classical theory to include the effects of mean winds, viscosity, and thermal conductivity. Very little progress has been made in explaining the tides in rainfall. *See* ATMOSPHERE; EARTH TIDES; TIDE.

Richard S. Lindzen

Bibliography. S. Chapman and R. S. Lindzen, *Atmospheric Tides*, 1970; Lord Kelvin (W. Thomson), On the thermodynamic acceleration of the earth's rotation, *Proc. Roy. Soc. Edinburgh*, 11:396–405, 1882; R. S. Lindzen, Atmospheric tides, *Annu. Rev. Earth Planet. Sci.*, 7:199–225, 1979; R. S. Lindzen, *Dynamics in Atmosphere Physics*, 1990; H. Volland, *Atmospheric Tidal and Planetary Waves*, 1988.

Atmospheric waves, upper synoptic

Horizontal wavelike oscillations in the pattern of wind flow aloft in the upper troposphere, usually with reference to the stronger portion of the westerly current in midlatitudes. The flow is anticyclonically curved in the vicinity of a ridge line in the wave pattern, and is cyclonically curved in the vicinity of a trough line. *See* TROPOSPHERE.

Any given hemispheric upper flow pattern may be represented by the superposition of sinusoidal waves of various lengths in the general westerly flow. Analysis of a typical pattern discloses the presence of prominent long waves, of which there are usually three or four around the hemisphere, and of distinctly evident short waves, of about half the length of the long waves.

Typically, each short-wave trough and ridge is associated with a particular cyclone and anticyclone, respectively, in the lower troposphere. The development and intensification of one of these circulations depends in a specific instance upon the details of this association, such as the relative positions and intensities of the upper trough and the low-level cyclone. These circulations produce the rapid day-to-day weather changes which are characteristic of the climate of the mid-latitudes.

The long waves aloft do not generally correspond to a single feature of the circulation pattern at low levels. They are relatively stable, slowly moving features which tend to guide the more rapid motion of the individual short waves and of their concomitant low-level cyclones and anticyclones. Thus, by virtue of their position and amplitude the long waves can exert an indirect influence on the character of the weather over a given region for a period of the order of weeks.

The motion, but not the amplification, of long waves and short waves can be predicted with considerable success by application of the principle of conservation of absolute vorticity to the flow pattern at midtropospheric elevations.

A blocking situation is one in which waves do not progress through the latitude belt of the westerlies. Blocking situations are frequently accompanied by extreme meteorological events; precipitation and cool temperatures persist near upper-level cyclones, and dry, warm weather persists near upper-level anticyclones. A blocking pattern usually consists of a ridge (anticyclone) over a trough (cyclone),

a high-amplitude ridge, or flow shaped like an uppercase Greek omega (Ω). Because of the preference for blocking off the west coasts of Europe and North America, it appears that topography must play an important role in blocking. *See* ATMOSPHERE; JET STREAM; STORM; VORTEX; WEATHER FORECASTING AND PREDICTION; WIND.

Frederick Sanders; Howard B. Bluestein

Bibliography. S. Colucci, An observational study of mid-tropospheric blocking during recent winter seasons, *The General Circulation: Theory, Modeling, and Observations,* Advanced Study Program of NCAR, Boulder, 397–405, 1978; R. Dole, The objective representation of blocking patterns, *The General Circulation: Theory, Modeling, and Observations,* Advanced Study Program of NCAR, Boulder, pp. 406–426, 1978.

Atoll

An annular coral reef, with or without small islets, that surrounds a lagoon without projecting land area.

Physiography. Most atolls are isolated reefs that rise from the deep sea. They vary considerably in size and shape. Small rings, some without islets, may be less than a mile in diameter, but many atolls have a diameter of 20 mi (32 km) or more and bear numerous islets. The largest atoll, Kwajalein in the Marshall Islands, has an irregular shape, is about 75 mi (120 km) long, 15–20 mi (24–36 km) wide, and covers 840 mi^2 (2180 km^2).

The reefs of the atoll ring are flat, pavementlike areas, large parts of which, particularly along the seaward margin, may be exposed at times of low tide (**Fig. 1**). The reefs vary in width from narrow ribbons to broad bulging areas more than a mile across. On windward reefs exposed to constant trade wind surf, the most prominent feature is commonly a low, cuestalike wave-resistant ridge whose steeper side faces the sea. The ridge, as much as 3 ft (1 m) above lowest tides, is composed of pink or rose-colored calcareous algae. To seaward of the ridge the gentle submarine slope is cut by a series of grooves. These may be slightly sinuous, in places branching as they extend seaward from the reef front. They are separated by flattened spurs or buttresses that appear to have developed mainly by coral and algal growth. The grooves and buttresses form the "toothed edge" that is so conspicuous when an atoll is viewed from the air. Beyond the toothed edge may be a shallow terrace, then a steep (35°) submarine slope that flattens progressively at deeper levels until it reaches the ocean floor. Some grooves of the toothed margin extend back through the marginal algal ridge for short distances as open surge channels; in places of exceptional development, deeper interconnecting channels may extend for several hundred feet under the reef flat, roofed over by algal growth and exposed here and there in open pools, or ending as a series of blowholes that spout with every wave. These structures taken together form a most effective baffle that robs the incoming waves of much of their destructive power, as they bring a constant supply of refreshing seawater with oxygen, food, and nutrient salts to wide expanses of reef. On leeward reefs, protected from the constant surf that characterizes windward reefs, the algal margin is less developed. Coral growth may be profuse to the reef edge which drops off sharply, in places vertically, for several hundred feet.

Inside the margin of the reef may be several zones that are rich in corals or other reef builders, but large parts of the reef flat are composed of cemented reef rock, bare or thinly veneered with loose sand or coarser reef debris. When the reef flats of Eniwetok Atoll were drilled, a solid plate of hard limestone was found below that extends to depths of 10–15 ft (3–4.5 m) and contains corals in growth position. Growth of organisms and deposition of sediments on the reef are controlled by the prevailing winds and currents and other ecologic factors. The sediments, including cemented rock, that make up the reef surface and its islets are composed entirely of carbonates secreted by shallow-water organisms.

The reefs of most atolls are broken, commonly to leeward, by one or more passes. Some of these may be as deep as the lagoon and give ready access to that body of water. Within atoll groups, lagoon depths show little relationship to size, except that very small lagoons are typically shallower than larger ones in the same group. Most atoll lagoons in the Caroline and Marshall islands range in depth from 100 to 300 ft (30 to 90 m), but the lagoon of Christmas Atoll, a very large atoll in the Line Islands, is only 23 ft (7 m) deep. Hundreds of coral knolls rise from the lagoon floor of many atolls. More than 2000 were charted in Eniwetok lagoon. Other atolls may show

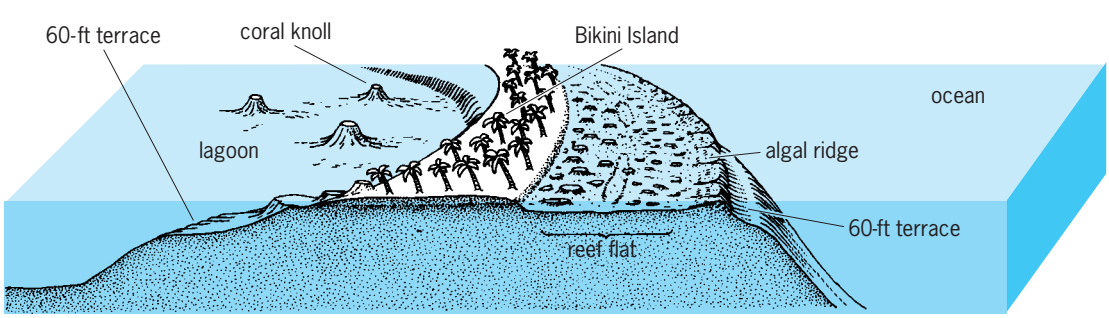

Fig. 1. Section showing features of rim of Bikini Atoll.

very few coral knolls, and some relatively shallow lagoons (Canton in the Phoenix Islands, Mataiva in the Tuamotus, Christmas Atoll) are divided into compartments by an intersecting network of linear coral ridges.

On many atolls islets are spaced at irregular intervals along the reef. Most of these are small and rarely rise as much as 20 ft (6 m) above the reef flat. They are composed of reef detritus (coral heads and sand) piled up by waves and winds. The sediments are unconsolidated except for layers of beach rock (cemented conglomerate and sandstone) at intertidal levels. Islets located on reefs that are subjected periodically to violent storms may exhibit one or a series of ridges or ramparts built of blocks of reef rock, coral heads, and finer material. These structures apparently can be built, moved, or destroyed by the waves of a single major storm. Low sand dunes may occur on the borders of islands inside sand beaches. Inland from the beaches and ramparts, sand and gravel flats occur and large areas are covered with coconut palms, breadfruit trees, and *Pandanus* (screw pines). More isolated islets may contain groves of native *Pisonia*, inside a beach ridge covered with *Tournefortia*, *Cordia*, and other native trees.

Many atolls exhibit terraces around their margins and inside their lagoons. Shallow drilling on Eniwetok islets has shown that the reef rock to a depth of 30 ft (9 m) or more below sea level was deposited beginning about 8000 years ago during the Holocene sea-level rise; the reef limestone immediately below that level and presumably that of the terraces was more than 130,000 years old and grew during the previous interglacial sea-level rise. *See* HOLOCENE.

Distribution. Existing atolls, like other types of coral reefs, require strong light and warm waters and are limited in the existing seas to tropical and near-tropical latitudes. A large percentage of the world's more than 400 atolls are contained in an area known as the former Darwin Rise that covers much of the central and western Pacific. Atolls are also numerous in parts of the Indian Ocean and a number are found, mostly on continental shelves, in the Caribbean area. *See* INDIAN OCEAN; PACIFIC OCEAN; TROPICAL METEOROLOGY.

The Darwin Rise also contains more than 150 flat-topped seamounts or guyots, which were atoll-like structures that existed at or nearly at sea level during Cretaceous time but are now drowned several thousand feet deep. Some, possibly many, atolls such as Bikini and Eniwetok have grown on a foundation of Cretaceous guyots during the Cenozoic. Similar atoll-like structures existed in shallow Paleozoic seas. These ancient atolls were built by a variety of organisms; the types of corals and algae that are primarily responsible for the present reefs were not then in existence. Many Paleozoic atolls became reservoirs of petroleum, and are being studied and mapped in detail by deep drilling. *See* PALEOZOIC; SEAMOUNT AND GUYOT.

Origin and development. Most atolls that rise from the deep sea have probably been built on submerged volcanoes. The reef caps of four such atolls (Eniwetok, Midway, Mururoa, and Fangataufa) have been drilled and the existence of a foundation of volcanic flows confirmed. In these atolls it was found that the volcanoes had been at least partly truncated by erosion prior to the initiation of reef growth. Many of the existing atolls probably originated as near-surface reefs in Tertiary time. The oldest limestones beneath Eniwetok are late Eocene, although the underlying basalt is Late Cretaceous (76.5 million years old), and Midway's oldest limestones are early Miocene. The bases of Mururoa and Fangataufa are late Miocene, and other atolls may prove to be appreciably younger. *See* EOCENE; MIOCENE; TERTIARY.

Wherever extended three-dimensional studies have been carried out on atolls—Funafuti in the Ellice Islands (Tuvalu), Kita-Daito-Jima, Bikini and Eniwetok in the Marshall Islands, Midway in Hawaii, or Mururoa and Fangataufa in the Tuamotus—they have proved the atolls to be complex structures; surface similarities may mask striking differences in underground structure. Wherever tested by the drill, the thicknesses of reef caps of atolls greatly exceed the depth limits of the reef builders, and these caps appear to have developed during long periods of subsidence.

The structure of Eniwetok Atoll in the Marshall Islands has been revealed by drilling (**Fig. 2**). The limestone cap, nearly a mile in thickness, rests on the truncated summit of a basaltic volcano that rises 2 mi (3 km) above the ocean floor. In one of the drillholes the lower limestones are late Eocene and formed as forereef deposits on an outer slope. Correlative limestones in a second hole were formed in shallow water. The 4000-ft (1200-m) section of limestone that forms the cap of Eniwetok records an overall subsidence of that amount over a period exceeding 40 million years. Subsidence of this magnitude even exceeds the amounts so brilliantly postulated by Charles Darwin. The sinking has not been steady or continuous but was interrupted by periods of relative emergence, as during periods of glaciation. The emergent stages are recorded in zones of

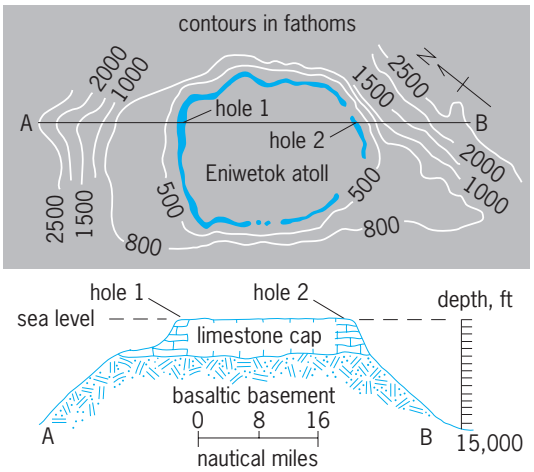

Fig. 2. Structure of Eniwetok Atoll in Marshall Islands. Lower limestones in hole 1 are late Eocene. Hole 2 reveals correlative limestones. 1 ft = 0.3 m; 1 naut mi = 1.85 km; 1 fathom = 1.83 m.

recrystallized and partially dolomitized sediments overlain by less altered limestone. Such sequences, termed solution unconformities, were formed when the top of the atoll stood above the sea for appreciable time. During such periods land snails lived on the island and land plants left a record of spores and pollen; Eniwetok and perhaps many other atolls functioned as stepping stones in the distribution of shallow-water marine life in the Pacific. *See* OCEANIC ISLANDS; REEF; SEQUENCE STRATIGRAPHY.

Harry S. Ladd; Joshua I. Tracey, Jr.

Bibliography. A. Guilcher, *Coral Reef Geomorphology*, 1988; N. M. Hill (ed.), *The Sea*, vol. 3, 1963, reprint 1981; H. S. Ladd et al., Drilling on Midway Atoll, Hawaii, *Science*, 156(3778):1088–1094, 1967; H. S. Ladd and S. O. Schlanger, *Drilling Operations on Eniwetok Atoll*, USGS Prof. Pap. 260-Y, 1960; H. J. Wiens, *Atoll Environment and Ecology*, 1962.

Atom

A constituent of matter consisting of z negatively charged electrons bound predominantly by the Coulomb force to a tiny, positively charged nucleus consisting of Z protons and $(A - Z)$ neutrons. Z is the atomic number, and A is the mass or nucleon number. The atomic mass unit is $u = 1.661 \times 10^{-27}$ kg. *See* ATOMIC MASS UNIT; ATOMIC NUMBER; COULOMB'S LAW; ELECTRON; MASS NUMBER; NEUTRON; PROTON.

Atoms are the smallest objects that have a chemical identity. Atoms are thus the basic building blocks of matter. Electrically neutral atoms ($z = Z$) with the range $Z = 1$ (hydrogen) to $Z = 92$ (uranium) make up the periodic table of the elements naturally occurring on Earth. (An element is the name of an atom having a given atomic number.) *See* ELEMENT (CHEMISTRY); PERIODIC TABLE.

Isotopes of a given element have different values of A but nearly identical chemical properties, which are fixed by the value of Z. Certain isotopes are not stable; they decay by various processes called radioactivity. Atoms with Z greater than 92 are all radioactive but may be synthesized, either naturally in stellar explosions or in the laboratory using accelerator techniques. *See* ISOTOPE; NUCLEAR STRUCTURE; RADIOACTIVITY; TRANSURANIUM ELEMENTS.

Atoms with $Z - z$ ranging from 1 to $Z - 1$ are called positive ions. Those having $z - Z = 1$ are called negative ions; none has been found with $z - Z$ greater than 1. *See* ION; NEGATIVE ION.

A Rydberg atom has one electron in a highly excited orbit. Its radius grows and binding energy decreases with the square and the inverse square, respectively, of the principal quantum number n. A neutral Rydberg atom with $n = 100$ has a huge diameter, approximately 10^{-6} m or 4×10^{-5} in. (near that of a bacterium). Its miniscule binding energy, approximately 1.36 millielectronvolts, explains its fragility in collisions with other particles. Occurring naturally in the interstellar medium, Rydberg atoms with n-values up to several hundred emit spectral lines that have been detected with radio telescopes. They have also been produced and studied in the laboratory. *See* RYDBERG ATOM.

An exotic atom has a different charged particle substituted for an electron, its nucleus, or both. For example, positronium consists of an electron and its antiparticle, a positron. *See* ATOMIC STRUCTURE AND SPECTRA; HADRONIC ATOM; MUONIUM; PIONIUM; POSITRONIUM.

Peter M. Koch

Atom cluster

Clusters are aggregates of atoms (or molecules) containing between three and a few thousand atoms that have properties intermediate between those of the isolated monomer (atom or molecule) and the bulk or solid-state material. The study of such species has been an increasingly active research field since about 1980. This activity is due to the fundamental interest in studying a completely new area that can bridge the gap between atomic and solid-state physics and also shows many analogies to nuclear physics. However, the research is also done for its potential technological interest in areas such as catalysis, photography, and epitaxy. A characteristic of clusters which is responsible for many of their interesting properties is the large number of atoms at the surface compared to those in the cluster interior. For many kinds of atomic clusters, all atoms are at the surface for sizes of up to 12 atoms. As the clusters grow further in size, the relative number of atoms at the surface scales as approximately $4N^{-1/3}$, where N is the total number of atoms. Even in a cluster as big as 10^5 atoms, almost 10% of the atoms are at the surface. Clusters can be placed in the following categories:

1. Microclusters have from 3 to 10–13 atoms. Concepts and methods of molecular physics are applicable.
2. Small clusters have from 10–13 to about 100 atoms. Many different geometrical isomers exist for a given cluster size with almost the same energies. Molecular concepts lose their applicability.
3. Large clusters have from 100 to 1000 atoms. A gradual transition is observed to the properties of the solid state.
4. Small particles or nanocrystals have at least 1000 atoms. These bodies display some of the properties of the solid state.

Geometry. Traditionally, solid-state physics describes a crystal as an infinitely extending periodic ordering of atoms with translational symmetry. Body-centered cubic (bcc), face-centered cubic (fcc), and hexagonal close packing (hcp) are the most common arrangements. These orderings are not normally found among clusters. The packing of the atoms in clusters can be investigated by electron diffraction on the cluster beam. The most favored geometry for rare-gas (neon, argon, and krypton) clusters of up to a few thousand atoms (**Fig. 1**) is icosahedral. The fivefold rotational axis of these structures, which first appears in clusters of seven atoms, is forbidden for the symmetry of the standard lattices of solid-state physics. As confirmed by computer simulation

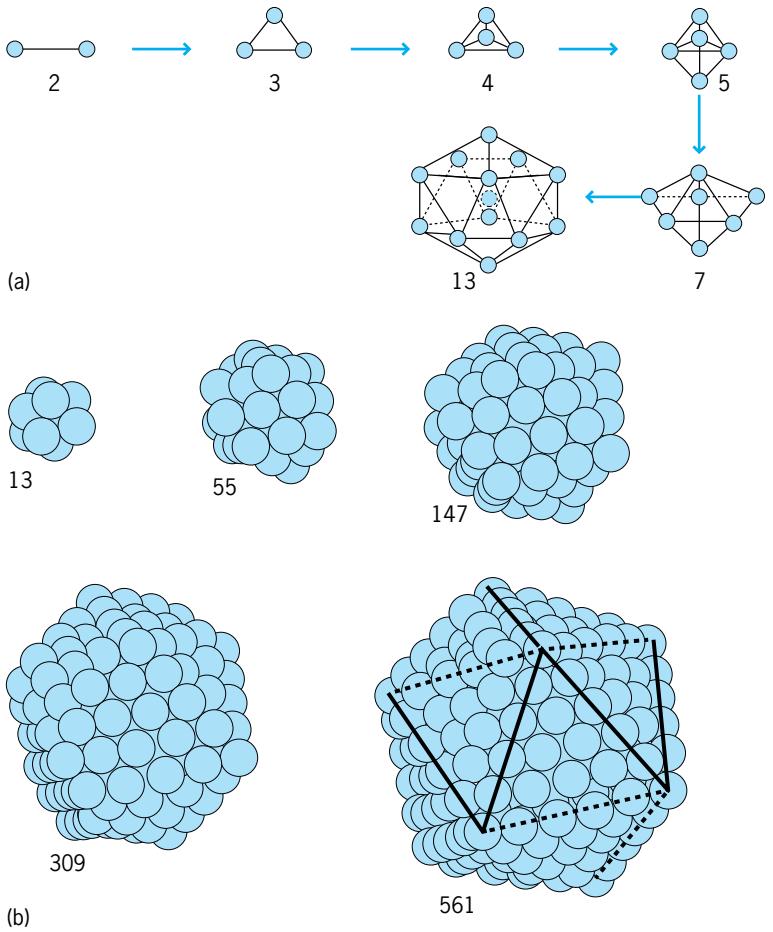

Fig. 1. Growth sequence of neutral rare-gas clusters. Number of atoms in cluster, *N*, is indicated for each cluster. (*a*) Microclusters. (*b*) Small and large clusters, all of which are icosahedra with fivefold symmetry. (*After H. Haberland, ed., Clusters of Atoms and Molecules, vol. 1, Springer-Verlag, 1994*)

methods, the icosahedral geometry allows the atoms to be aggregated into a more compact structure with a smaller potential energy than the structures allowed in solid-state physics, such as the face-centered cubic structure of bulk rare-gas solids. *See* CRYSTAL STRUCTURE.

However, the preferred cluster geometry depends critically on the bonding between the monomers in the clusters. For example, ionic clusters such as those of sodium chloride [$(NaCl)_N$] very rapidly assume the cubic form of the bulk crystal lattice, and for metallic clusters it is the electronic structure rather than the geometric structure which is most important. A summary of the different kinds of clusters, classified in terms of their chemical bonding, is given in the **table**. Mercury clusters (Hg_N) are particularly intriguing since they show van der Waals, covalent, or metallic bonding, depending on the cluster size. *See* CHEMICAL BONDING.

Production. A cluster apparatus typically consists of two separately pumped vacuum chambers which are connected by a small conical aperture, the skimmer. The central part of the cluster beam, formed in the first vacuum chamber, passes through the skimmer and is investigated in some way, ionized, and detected in a mass spectrometer in the second chamber.

There are two main types of sources for producing free cluster beams. In a gas-aggregation source, the atoms or molecules are vaporized into a cold, flowing rare-gas atmosphere. They cool by undergoing many collisions with the rare-gas atoms and aggregate into clusters. In a jet-expansion source, a gas is expanded under high pressure through a small hole into a vacuum. An extreme cooling of the rotational and vibrational degrees of freedom as well as the relative motion occurs during this adiabatic expansion, leading to cluster formation.

Classification of clusters according to their chemical bonding

Kind of cluster	Example	Average binding energy per atom, eV	Elements for which cluster type is found
Metallic clusters: half-filled band of delocalized electrons	(alkali metal)$_N$, Al_N, Cu_N, Fe_N, Pt_N, W_N, Hg_N with $N > 200$	0.5±3	Elements of the lower-left corner of the periodic table
Covalent clusters: directed bonding by electron pairs through *sp*-hybridization	C_N, Si_N, Hg_N, $80 \geq N \geq 30$	1±4 (Hg ≈ 0.5)	B, C, Si, Ge
Ionic clusters: bonding due to Coulomb force between ions	$(NaCl)_N$, $(CaBr_2)_N$	2±4	Metals from the left side of the periodic table with electronegative elements from the right side
Hydrogen-bonded clusters: strong dipole-dipole attraction	$(HF)_N$, $(H_2O)_N$	0.15±0.5	Molecules with closed electronic shells that contain H and strong electronegative elements (F, O, N)
Molecular clusters: like van der Waals with an additional weak covalent contribution	$(I_2)_N$, $(S_6)_N$, (organic molecules)$_N$	0.3±1	Organic molecules, some other closed-shell molecules
Van der Waals clusters: induced dipole interaction between atoms and molecules with closed electronic shells	(rare gas)$_N$, $(H_2)_N$, $(CO_2)_N$, Hg_N with $N < 10$	0.01±0.3	Rare gases, closed-shell atoms and molecules

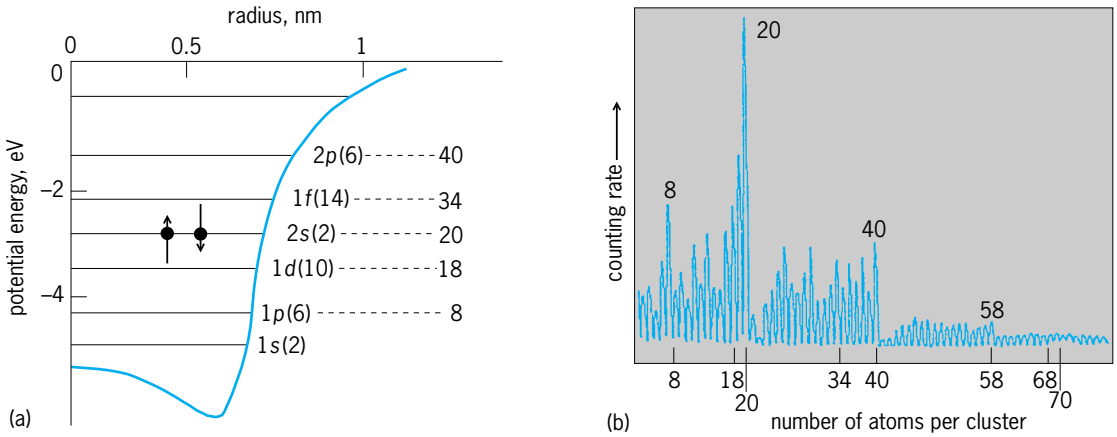

Fig. 2. Jellium model. (*a*) Potential energy of the electrons as a function of the cluster radius for a spherical sodium cluster with 20 atoms (Na_{20}). Electron energy levels are also shown, with their radial and angular-momentum quantum numbers and the degeneracies of the levels (in parentheses). The total number of electrons in a given shell is indicated on the right. Only the two electrons in the highest occupied orbital are indicated. (*b*) Mass spectrum of sodium (Na_N) clusters, showing high intensities (magic numbers) for clusters with completely filled electronic shells. (*After W. A. de Heer et al., Electronic shell structure and metal clusters, Sol. State Phys., 40:93–181, 1987*)

One crucial requirement for both these sources is to obtain a sufficient density of monomers to produce a beam with a high enough intensity of clusters for further investigation.The common way of doing this for materials with a high melting point is to use a laser-vaporization or ion-sputtering source. An intense ion or laser beam is directed onto the surface of the target material and removes material in the form of neutral or charged particles, which then cluster together. The number of clusters, their size, and their temperature can be controlled by combining this method with those of the gas-aggregation or jet-expansion source.

Properties. In most situations, the valence electrons of the atoms making up the clusters can be regarded as being delocalized, that is, not attached to any particular atom but with a certain probability of being found anywhere within the cluster. The simplest and most widely used model to describe the delocalized electrons in metallic clusters is that of a free-electron gas, known as the jellium model. The positive charge is regarded as being smeared out over the entire volume of the cluster while the valence electrons are free to move within this homogeneously distributed, positively charged background. The calculated potential for the electrons in a spherical jellium approximation typically looks like the example in **Fig. 2***a*. Here, the inner part of the effective potential resembles the bottom of a wine bottle. The electronic energy levels are grouped together to form shells. The jellium potential is very similar to the Woods-Saxon potential, used to describe the interaction between protons and neutrons in nuclear physics, and the same classification of energy levels is found, wherein the energy levels are characterized by the radial quantum number (1, 2, 3, . . .) and the angular-momentum quantum number (*s*, *p*, *d*, . . .). *See* NUCLEAR STRUCTURE.

The predicted shell closures can be seen in mass spectra of alkali-metal clusters. If the neutral clusters produced from a given cluster source are hot, they can fragment on leaving the source. The intensity of relatively stable clusters will thus grow at the expense of their less stable neighbors, producing so-called magic numbers in the mass distribution. The neutral mass distribution is then probed by ionizing the clusters with a photon energy that is just slightly higher than the ionization potential of the clusters to avoid further fragmentation after ionization. This operation was used to obtain the mass spectrum of sodium clusters in Fig. 2*b*, which gives a beautiful confirmation of the electronic shell model. *See* MAGIC NUMBERS.

The delocalized electrons can be excited by photon absorption or electron collisions into a collective motion with respect to the positively charged nuclei, which is known as a plasmon resonance. This is very analogous to the giant resonances, that is, the collective motion of neutrons and protons, found in atomic nuclei. The plasmon frequency for metal clusters depends on the cluster size and material but is generally in the visible range of the spectrum. It is the plasmon excitation of metal clusters or nanoparticles in glass that is responsible for the strong colors seen, for example, in medieval stained glass windows. The optical absorption of metal clusters is compared to that of atoms and the bulk metal in **Fig. 3**. *See* GIANT NUCLEAR RESONANCES; PLASMON.

Fullerenes. Fullerenes are clusters of carbon atoms, C_{2N}, with N greater than 11, which have unique hollow structures. They were discovered with a laser vaporization source combined with an adiabatic gas expansion. It is now possible to produce and isolate macroscopic amounts of some of these clusters, in particular C_{60} and C_{70}. The cluster C_{60} is especially interesting because of its highly symmetrical truncated icosahedral geometry (just like a soccer ball) and a range of fascinating properties, such as a quasi-three-dimensional aromatic chemistry and superconductivity when doped with alkali atoms. The availability of macroscopic amounts of a mass-selected, neutral atomic cluster with

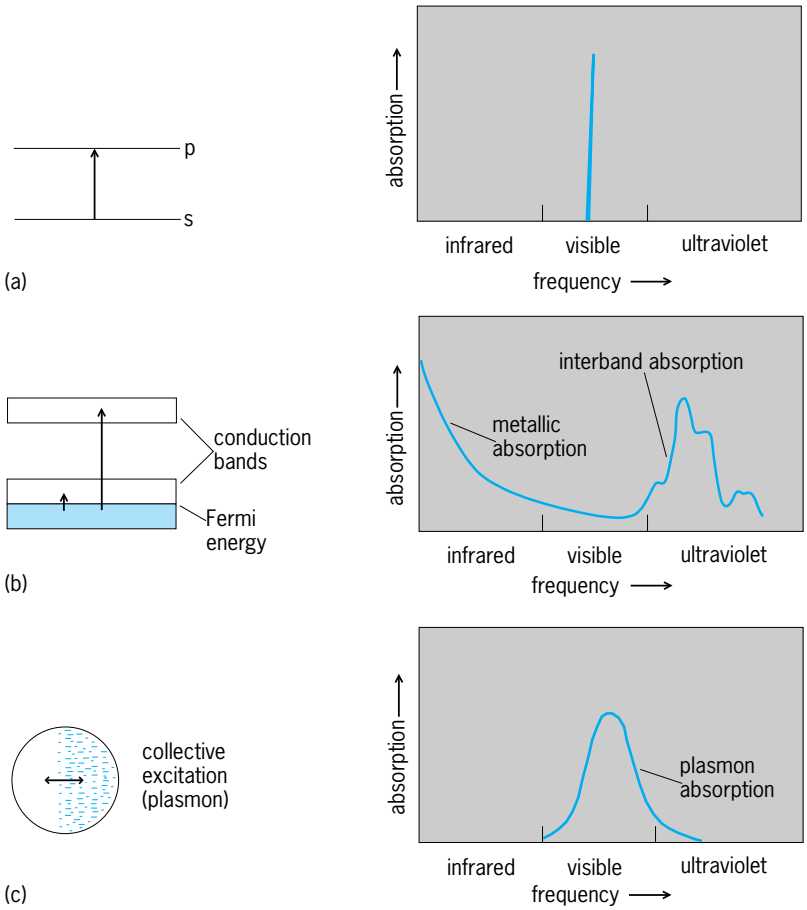

Fig. 3. Absorption of light by an alkali metal in various states of aggregation. Absorption spectra are shown on the right and schematic structures on the left. (*a*) Metal atom. (*b*) Metal crystal. (*c*) Metal cluster.

well-defined geometrical and electronic properties has opened up a new range of experiments and theoretical treatments to probe the dynamics of a relatively simple atomic cluster system with a large but finite number of degrees of freedom. *See* FULLERENE.

Probing with laser pulses. A recent development in the study of atomic clusters is the use of laser pulses with durations of less than 100 femtoseconds (10^{-13} s) to probe the time scale for energy coupling between the different degrees of freedom in the cluster in real time. This work, which is presently in its infancy, will provide very detailed information on the dynamics of such complex systems and will certainly provide a stringent test of theoretical models. *See* OPTICAL PULSES.

Clusters on surfaces. The overall emphasis in the field is gradually shifting from gas-phase studies to the study of clusters deposited on surfaces. The aim is to produce new materials with tailor-made electrical, optical, or catalytic properties. Cluster-beam epitaxy, where large metallic clusters are deposited at high kinetic energy (greater than 10 eV per atom) onto substrates, is already producing superior-quality mirror coatings for use in high-power laser applications.

Eleanor Campbell

Bibliography. M. S. Dresselhaus, G. Dresselhaus, and P. C. Eklund, *Science of Fullerenes and Carbon Nanotubes*, 1996; H. Haberland (ed.), *Clusters of Atoms and Molecules*, 2 vols., 1994; U. Kreibig and M. Vollmer, *Optical Properties of Metal Clusters*, 1995.

Atom economy

The maximum incorporation of the starting materials into the final product for any chemical reaction.

A broad imperative for synthetic chemists today is the need to develop innovative and sophisticated synthetic methods to keep pace with the continually changing variety of molecular targets. In 1991, Barry M. Trost presented a set of coherent guiding principles to evaluate the efficiency of a specific chemical process, which has subsequently altered the way many chemists design and plan their syntheses. He proposed that within synthetic efficiency there are two broad categories: selectivity and atom economy. Selectivity is defined as chemo-, regio-, and stereoselectivity, while atom economy seeks to maximize the incorporation of the starting materials into the final product. An additional corollary is that if maximum incorporation cannot be achieved, then ideally the side-product quantities should be minute and environmentally innocuous.

In the past, attaining the highest yield and product selectivity were the governing factors in chemical synthesis. As a result, multiple reagents often were used in stoichiometric quantities that were not incorporated into the target molecule, resulting in significant side products and waste. Today, the underlying standard is synthetic efficiency—the ability to maximize the incorporation of the starting materials into the final product and to minimize by-products.

The reaction yield and the atom-economy yield are calculated by different means. The reaction yield [Eq. (1)] is concerned only with the quantity of desired product isolated, relative to the theoretical quantity of the product, and does not provide information in terms of the efficiency of the transfer of molecular weight from the reactants into the desired product, which is the concern of atom economy. Atom economy takes into account all the reagents used and the unwanted side products, along with the desired product [Eq. (2)].

$$\text{Reaction yield} = \frac{\text{quantity of product isolated}}{\text{theoretical quantity of product}} \times 100\% \quad (1)$$

$$\text{Atom economy} = \frac{\text{molecular weight of products}}{\text{molecular weight of all products}} \times 100\% \quad (2)$$

Atom-uneconomic reactions. Examples of atom-uneconomic reactions are the Wittig reaction (formation of triphenylphosphine oxide), Grignard reaction, and substitution and elimination reactions (3). *See* GRIGNARD REACTION.

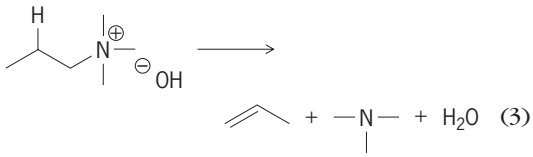

(3)

Atomic-economic reactions. Some atomic-economic reactions are the Diels-Alder reaction, rearrangements [reaction (4)], and concerted reactions. Not all

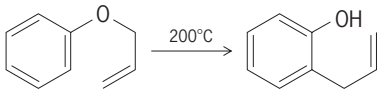

(4)

rearrangements are equal in that some require large amounts of acids to push the reaction forward (for example, the Beckmann rearrangement), and so those acids are considered to be ancillary reagents and factor into the economy of the reaction. *See* DIELS-ALDER REACTION.

Currently, many chemists are working on increasing the number of reactions in the "toolbox" by developing ideal and innovative methodologies that combine both selectivity and atom economy. Transition metals have proven to be excellent catalysts for the stereoselective transformation of many organic molecules, and since they are used catalytically, they are ideal for the development of atomic-economic methodologies. Reaction scheme (5) is an example of an transition-metal-catalyzed, atom-economic reaction.

The ruthenium catalyst complex is easily accessible, and its generality makes it the catalyst of choice. With this catalyst, the reactions are usually complete in 30 minutes at ambient temperature, and based upon the substituent there is regioselective control. In addition, this reaction is 100% atom-economic. *See* ASYMMETRIC SYNTHESIS; CATALYSIS; STEREOCHEMISTRY.

Synthesis of ibuprofen. A practical application of atom economy in industry is the synthesis of ibuprofen. The traditional synthesis of ibuprofen was developed in the 1960s. It was a six-step synthesis that used stoichiometric amounts of reagents and generated large quantities of waste by-products that needed further treatment [reaction scheme (6)]. There are numerous waste by-products, and the overall atom economy of the synthesis is 40%, or conversely, there is an overall waste of 60%.

Reaction scheme (7) shows a more atom-economic synthesis of ibuprofen, having an atom economy percentage of 77% (or 99% if the acetic acid generated in the first step is recovered).

For both reaction schemes (6) and (7), the starting materials and products are the same for the first step, with the only difference being the catalyst. In reaction scheme (6), $AlCl_3$ is not truly a catalyst but an auxiliary reagent that is needed in stoichiometric

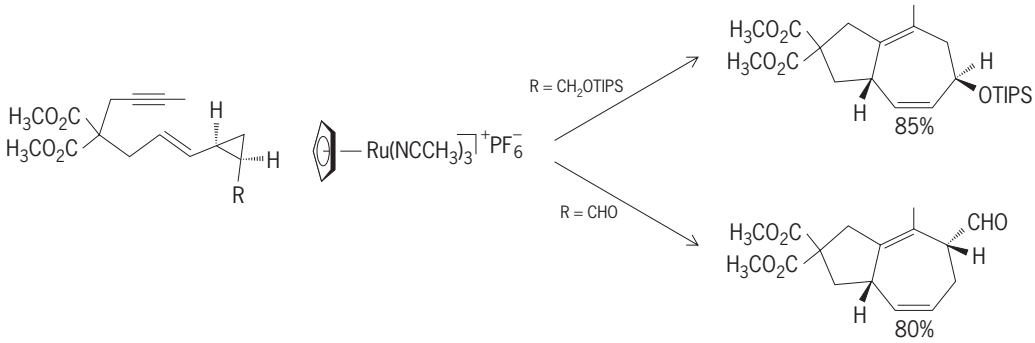

(5)

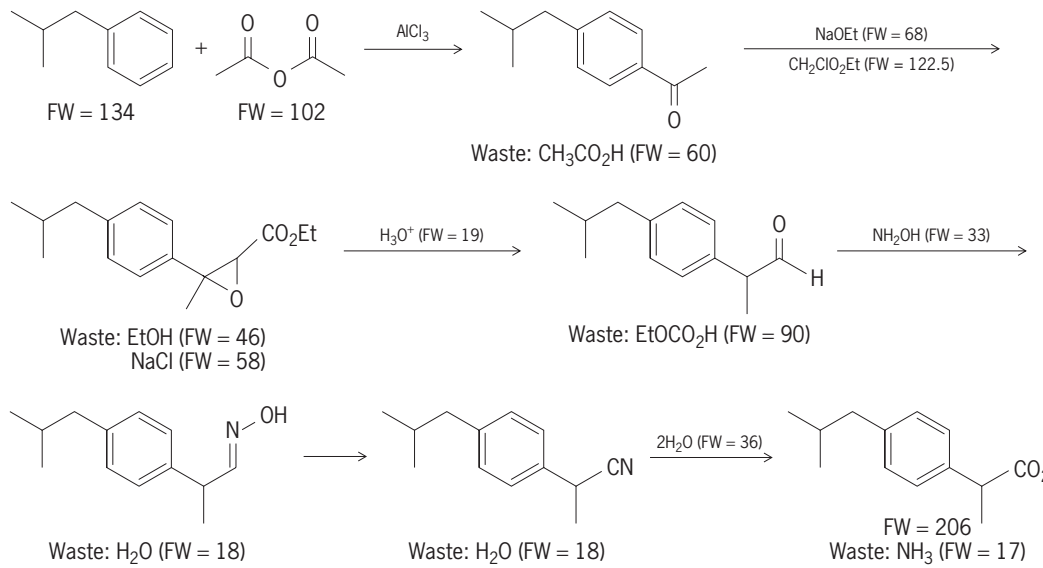

(6)

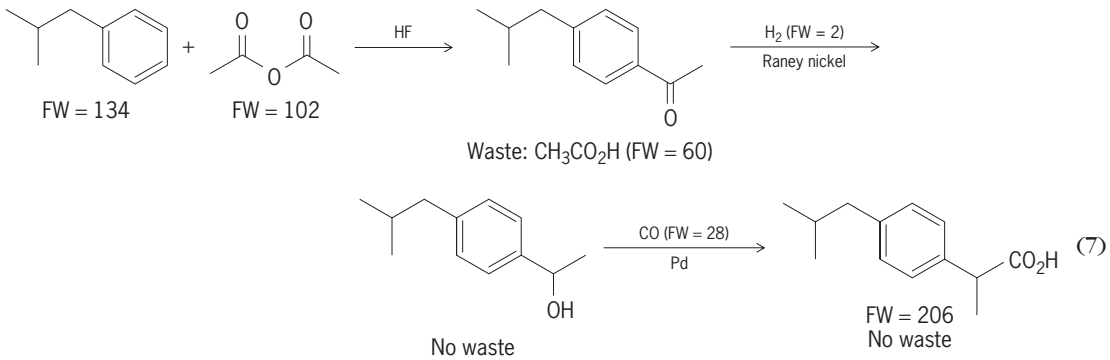

(7)

amounts and produces aluminum trichloride hydrate as a waste by-product. Hydrogen fluoride used in reaction shceme 7 is a catalyst, which is recovered and repeatedly reused. In addition, the Raney nickel and palladium catalysts are also recovered and reused. As a result, almost no waste is generated.

Since reaction scheme (7) requires three steps [relative to six for reaction scheme (6)], there is the additional elimination of waste by not having three extra steps. This results in the ability to produce larger quantities of ibuprofen in less time and with less capital expenditure. Therefore, besides having an atom-economic and environmentally friendly green synthesis, the product has a larger profit capability. *See* ORGANIC SYNTHESIS.

Charlene C. K. Keh; Chao-Jun Li

Bibliography. P. T. Anastas and J. C. Warner, *Green Chemistry: Theory and Practice*; Oxford, New York, 1998; B. M. Trost, The atom economy—a search for synthetic efficiency, *Science*, 254:1471–1477, 1991; B. M. Trost, On inventing reactions for atom economy, *Acc. Chem. Res.*, 35:695–705, 2002.

Atom laser

An atom laser is a device that generates an intense coherent beam of atoms through a stimulated process. It does for atoms what an optical laser does for light. The atom laser emits coherent matter waves, whereas the optical laser emits coherent electromagnetic waves. Coherence means, for instance, that atom laser beams can interfere with each other. *See* COHERENCE.

Properties. Laser light is created by stimulated emission of photons, a light amplification process. Similarly, an atom laser beam is created by stimulated amplification of matter waves. The conservation of the number of atoms is not in conflict with matter-wave amplification: The atom laser takes atoms out of a reservoir and transforms them into a coherent matter wave similar to the optical laser, which converts energy into coherent electromagnetic radiation (but, in contrast, the number of photons need not be conserved). *See* LASER.

The condition of high intensity means that there are many particles per mode or quantum state. A thermal atomic beam has a population per mode of only 10^{-12} compared to much greater than 1 for an atom laser. The realization of an atom laser therefore required the development of methods to greatly enhance the mode occupation. This enhancement is accomplished by cooling the reservoir of atoms to microkelvin temperatures or below. *See* LASER COOLING.

In the case of an ideal atom laser, the output beam should be monochromatic and directional, and have a well-defined phase and intensity. For atoms, being monochromatic means that their velocity spread is extremely small. Such beams propagate with minimum spreading, and can be focused by atom lenses to a small spot size. The minimum spreading and the minimum spot size are limited by Heisenberg's uncertainty relation in the same way as the propagation of a single-mode optical laser beam is diffraction limited. The analogy between light and matter waves is exploited in the field of atom optics. *See* ATOM OPTICS.

The different nature of atoms and photons implies different properties of light and atom beams. Unlike light, an atomic beam cannot travel far through air. It scatters off air molecules in less than a micrometer. Vacuum is thus required for all atom laser experiments. Also, slow atoms are strongly affected by gravity. Furthermore, a dense atom beam will show spreading in excess of the Heisenberg uncertainty limit because of the interactions between the atoms.

Elements. A laser requires a cavity (resonator), an active medium, and an output coupler (see **table**).

Cavity. Various analogs of laser cavities for atoms have been realized. The most important ones are

Analogies between an atom laser and the optical laser

Atom laser*	Optical laser
Atoms	Photons
Matter waves	Electromagnetic waves
Atom trap	Laser cavity
Atoms in the Bose condensate	Photons in the lasing mode
Thermal atoms	Gain medium
Evaporative cooling	Excitation of the gain medium
Stimulated scattering of atoms	Stimulated emission of photons
Critical temperature for Bose-Einstein condensation	Laser threshold

*Based on evaporative cooling.

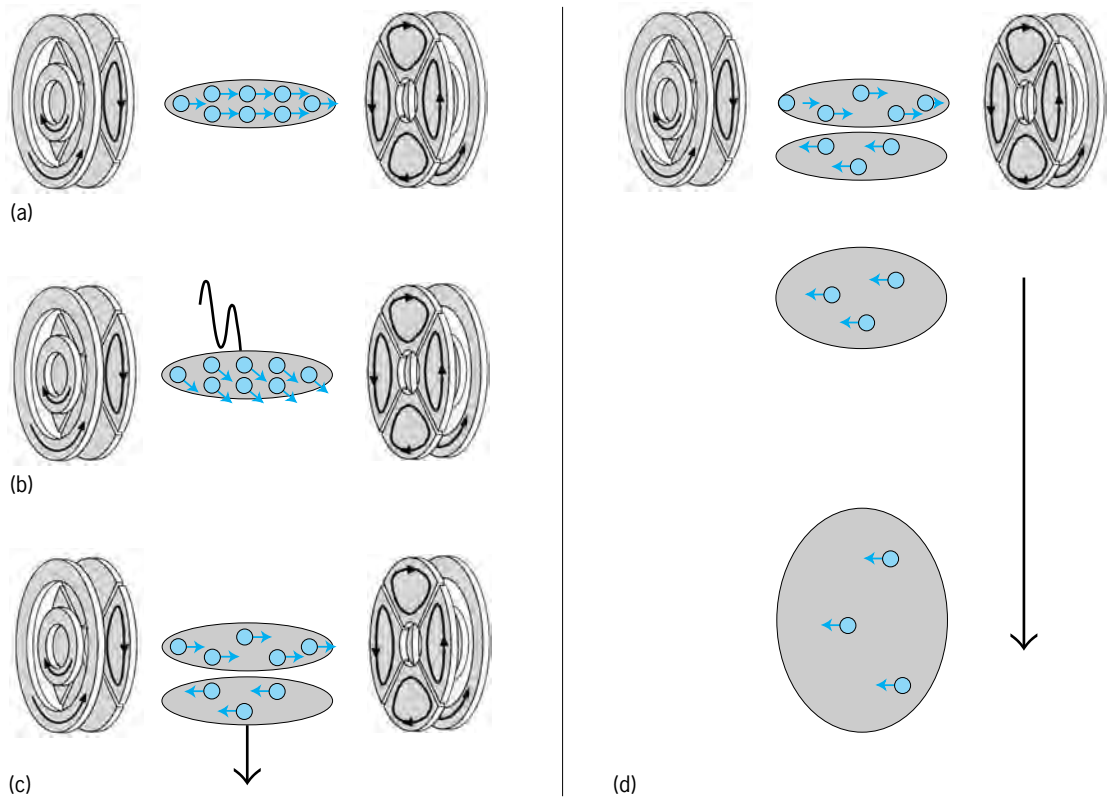

Fig. 1. Principle of the radio-frequency output coupler of an atom laser. (*a*) Bose-Einstein condensate confined in an atom trap. (*b*) Tilting of the spins of the atoms by a short radio-frequency pulse. (*c*) Splitting of the cloud into a trapped cloud and an out-coupled cloud. (*d*) Repetition of steps *a* and *c* creates a sequence of output pulses.

magnetic traps (which use the force of an inhomogeneous magnetic field on the atomic magnetic dipole moment) and optical dipole traps (which use the force exerted on atoms by focused laser beams). Confinement of atoms between two atom mirrors has been suggested and is analogous to a Fabry-Perot cavity for light. Even a single-mirror cavity is possible, where atoms perform multiple bounces off a mirror in the vertical direction and return because of gravity (an atomic "trampoline"). *See* INTERFEROMETRY; PARTICLE TRAP.

Active medium. The active medium is a reservoir of atoms which are transferred to one state of the confining potential, which is the analog of the lasing mode. The reservoir can be atoms confined in other quantum states of the atom cavity or an ultraslow atomic beam. The atoms are transferred to the lasing mode either by collisions or by optical pumping. The transfer of atoms is efficient only for an ultracold sample, which is prepared by laser cooling or evaporative cooling. This cooling ensures that the atoms in the reservoir occupy only a certain range of quantum states which can be efficiently coupled to the lasing mode.

Output coupler. The output coupler extracts atoms from the cavity, thus generating a pulsed or continuous beam of coherent atoms. A simple way to accomplish this step is to switch off the atom trap and release the atoms. This method is analogous to cavity dumping for an optical laser, and extracts all the stored atoms into a single pulse. A more controlled way to extract the atoms requires a coupling mechanism between confined quantum states and propagating modes.

Such a beam splitter for atoms can be realized by applying the Stern-Gerlach effect to atoms in a magnetic trap (**Fig. 1**). Initially, all the atoms have their electron spin parallel to the magnetic field, say spin up (Fig. 1*a*), and in this state they are confined in the trap. A short radio-frequency pulse rotates (tilts) the spin of the atoms by a variable angle (Fig. 1*b*). Quantum-mechanically, a tilted spin is a superposition of spin up and spin down. Since the spin-down component experiences a repulsive magnetic force, the cloud of atoms is split into a trapped cloud and an out-coupled cloud (Fig. 1*c*). By using a series of radio-frequency pulses, a sequence of coherent atom pulses can be formed (Fig. 1*d*). These pulses are accelerated downward by gravity and spread out. *See* QUANTUM MECHANICS.

Figure 2 shows such a sequence of coherent pulses. In this case, sodium atoms are coupled out from a magnetic trap by radio-frequency pulses every 5 ms. The atom pulses are observed by illuminating them with resonant laser light and imaging their shadows, which are caused by absorption of the light. Each pulse contains 10^5–10^6 sodium atoms.

Other output coupling schemes have been suggested, including optical transitions which eject atoms from the cavity because of the recoil of the absorbed photon, and tunneling through thin barriers of light.

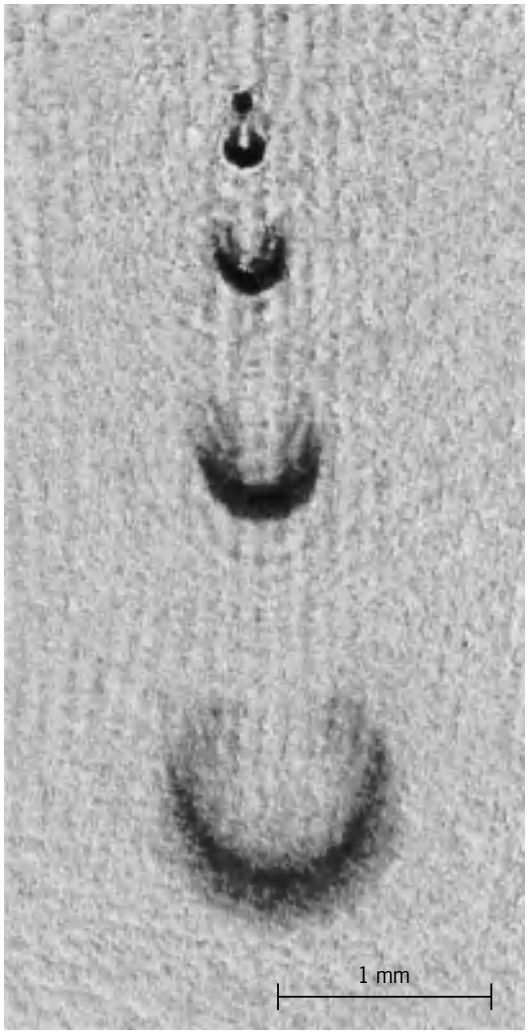

Fig. 2. Pulsed atom laser in operation, with pulses of coherent sodium atoms coupled out from a Bose-Einstein condensate that is confined in a magnetic trap.

Gain process. An atom laser is possible only for bosonic atoms. The accumulation of atoms in a single quantum state is based on Bose-Einstein statistics. Two different mechanisms have been discussed which may provide gain in an atom laser: elastic collisions and spontaneous emission of photons. *See* BOSE-EINSTEIN STATISTICS.

The case of elastic collisions is closely related to Bose-Einstein condensation. When a gas of bosonic particles is cooled down, it forms a Bose-Einstein condensate characterized by a macroscopic occupation of the ground state of the system. This process happens suddenly at the Bose-Einstein condensation transition temperature. The atoms in a Bose-Einstein condensate are coherent to first and higher order. An atom laser based on Bose-Einstein condensation operates in thermal equilibrium. Atom lasing is achieved simply by cooling down the gas. Below a certain temperature, nature maximizes entropy by generating a Bose-Einstein condensate. For photons, the situation is very different: Cooling a black-body cavity reduces its energy density in proportion to the fourth power of its absolute temperature (Stefan-Boltzmann law). Thus, at very low temperatures the cavity is empty. This is how entropy is maximized when the number of particles is not conserved. *See* BOSE-EINSTEIN STATISTICS.

It is instructive to look more closely at the stimulated amplification process that takes place when a Bose-Einstein condensate forms. In a normal gas, atoms scatter among a myriad of possible quantum states. But when the critical temperature for Bose-Einstein condensation is reached, they scatter predominantly into the lowest energy state of the system. This abrupt process is closely analogous to the threshold for operating an optical laser. The presence of a Bose-Einstein condensate causes stimulated scattering into the ground state. More precisely, the presence of a condensate with N_0 atoms enhances the probability that an atom will be scattered into the condensate by a factor of $N_0 + 1$.

In an atom laser, the analog to excitation of the active medium can be accomplished by evaporative cooling. The evaporation process creates a cloud which is not in thermal equilibrium and relaxes toward lower temperatures. This results in growth of the condensate. After equilibration, the gain process halts and the condensate fraction remains constant until further cooling is applied. In thermal equilibrium, there is still stimulated scattering of atoms into the condensate. However, this process is in dynamic equilibrium with collisions that knock atoms out of the condensate.

An atom laser was realized by extracting a beam of atoms from a Bose-Einstein condensate and explicitly demonstrating its coherence. The proof of the coherence was obtained by observing a high-contrast interference pattern when two Bose-Einstein condensates overlapped. The two condensates were created by cooling a gas of sodium atoms in a double-well potential. After the condensates were released from the trap, they accelerated downward, spread out ballistically, and eventually overlapped and interfered. The interference pattern could be directly photographed. It had a period of 15 micrometers, a gigantic length for matter waves. (Room-temperature atoms have a matter wavelength of 0.05 nm, 300,000 times smaller.) *See* INTERFERENCE OF WAVES.

An atom laser based on Bose-Einstein condensation is a special case of macroscopic occupation of a quantum state. In this case, the atoms accumulate in the ground state and are in thermal equilibrium. More generally, atom lasers can operate in higher-order modes and also as a driven system which is not in thermal equilibrium. (This is the situation in an optical laser.) The lasing mode is distinguished by preferential population of atoms or minimum loss. It has been suggested that this condition can be realized by optical pumping. In this case, atoms in the reservoir are optically excited. When they decay by spontaneous emission, they can reach final momentum states that differ from the initial momentum by the photon recoil. If one state within this range has a macroscopic population, then the rate of spontaneous emission into this final state is enhanced, and there is an amplification process similar to the one

described above for elastic collisions. The case of optically excited atoms shows very clearly the symmetry between the optical laser and the atom laser: The rate of emission to a final state, which is specified by a particular state of the atom inside the cavity and a particular mode of the photon field, is proportional to $(N + 1)(n + 1)$, where N is the number of atoms in this level of the cavity, and n is the number of photons in this mode. The first factor in this expression is the bosonic stimulation by atoms that is responsible for the amplification process in the atom laser, and the second describes the amplification process in the optical laser.

Potential applications. Although a basic atom laser has now been demonstrated, major improvements are necessary before it can be used for applications, especially in terms of increased output power and reduced overall complexity. The atom laser provides ultimate control over the position and motion of atoms at the quantum level, and might find use where such precise control is necessary, for example, for precision measurements of fundamental constants, tests of fundamental symmetries, atom optics (in particular, atom interferometry and atom holography), and precise deposition of atoms on surfaces. Since the matter wavelength of atoms can be extremely short (it decreases in inverse proportion to the atomic velocity), the ultimate limit to the spatial resolution is not the matter wavelength but the size of the atom. *See* FUNDAMENTAL CONSTANTS; NANOTECHNOLOGY; SYMMETRY LAWS (PHYSICS). Wolfgang Ketterle

Bibliography. M. R. Andrews et al., Observation of interference between two Bose condensates, *Science*, 275:637–641, 1997; B. Goss Levi, Bose condensates are coherent inside and outside an atom trap, *Phys. Today*, 50(3):17–18, March 1997; D. Kleppner, A beginner's guide to the atom laser, *Phys. Today*, 50(8):11–13, August 1997; H.-J. Miesner et al., Bosonic stimulation in the formation of Bose-Einstein condensate, *Science*, 279:1005–1007, 1998; G. Taubes, First atom laser shoots pulses of coherent matter, *Science*, 275:617–618, 1997.

Atom optics

The use of laser light and nanofabricated structures to manipulate the motion of atoms in the same manner that rudimentary optical elements control light. The term refers to both an outlook in which atoms in atomic beams are thought of and manipulated like photons in light beams, and a collection of demonstrated techniques for doing such manipulation (see **table**). Two types of atom optics elements have existed for some time: slits and holes used to collimate molecular beams (the analog of the pinhole camera), and focusing lenses for atoms and molecules (for example, hexapole magnets and quadrupole electrostatic lenses). However, in the 1980s the collection of optical elements for atoms expanded dramatically because of the use of near-resonant laser light and fabricated structures to make several types of mirrors as well as diffraction gratings. The diffraction gratings are particularly interesting because they exploit and demonstrate the (de Broglie) wave nature of atoms in a clear fashion. *See* LASER.

Diffraction gratings. Diffraction gratings for atoms have been made by using either a standing wave of light or a slotted membrane. The standing light wave makes a phase grating (that is, it advances or retards alternate sections of the incident wavefront but does not absorb any of the atom wave), so that the transmitted intensity is high. This approach requires the complexity of a single-mode laser, and introduces the complication that the light acts differently on the various hyperfine states of the atom. The slotted membrane, however, absorbs (or backscatters) atoms which strike the grating bars but does not significantly alter the phase of the transmitted atoms; it is therefore an amplitude grating. It works for any atom or molecule, regardless of internal quantum state, but with total transmission limited to about 40% by the opacity of the grating bars and requisite support structure.

The effect of a slotted membrane type of diffraction grating on atoms and molecules is demonstrated by an experiment in which a beam composed of both sodium atoms and a smaller number of sodium dimers (Na_2), moving at the same speed, is diffracted from a nanofabricated transmission grating (**Fig. 1**). Because the molecules have twice the mass of the atoms and are traveling at the same speed, they have twice the momentum, p. The de Broglie wave that describes their motion has a wavelength which is equal to Planck's constant divided by their momentum, and is therefore one-half of the atoms' de Broglie wavelength. As a result, the spacing of the peaks in the diffraction pattern is half as large. The diffraction pattern therefore has small peaks, attributed to the less populous molecules, in between adjacent larger ones.

If the atoms are removed from the beam by resonant laser light, leaving only the molecules, the smaller peaks are, indeed, unaffected, as is expected if they arise only from molecules in the combined beam. A similar diffraction experiment with a supersonic helium beam provided definitive evidence for the existence of the helium dimer, the most weakly bound molecule known. *See* INTERMOLECULAR FORCES.

The above diffraction pattern may be regarded as a demonstration that atoms and molecules behave

Comparison of optical elements for photons and atoms

Photons	Atoms
Lens	Electromagnetic fields (static, light) Zone plate
Mirror	Crystal surface Liquid helium Evanescent light
Phase grating	Standing light wave
Amplitude grating	Nanofabricated bars
Beam splitter	Only diffraction gratings (no achromatic elements)

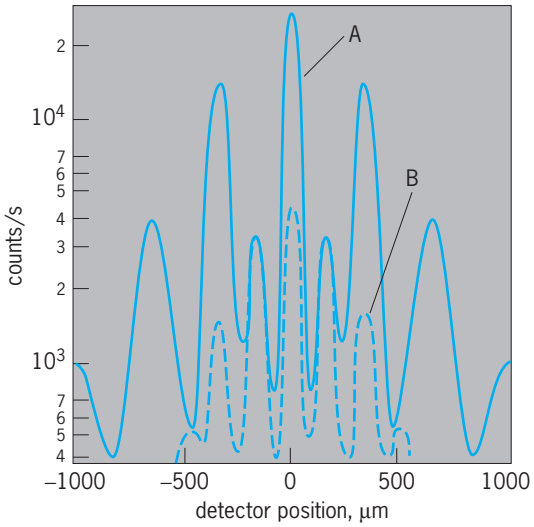

Fig. 1. Diffraction patterns from a transmission grating [period = 100 nm] (curve A) of a beam composed of both sodium atoms (Na) and sodium dimers (Na_2), and (curve B) of the same beam after the atoms have been removed with a deflecting laser beam, leaving only Na_2 molecules.

like waves. This demonstration differs fundamentally from similar demonstrations with neutrons and electrons in that atoms and molecules are composite particles; a sodium molecule is about 0.6 nanometer in diameter. In the experiment described above, both the de Broglie wavelength, 0.016 nm, and the coherence length, 0.12 nm, of these waves (the distance over which they remain in phase) were smaller than the size of the object which they describe. *See* DIFFRACTION GRATING.

In a diffraction grating, the constant spacing of the bars produces a constant angle of diffraction so that all the diffracted waves travel in parallel away from the grating. If the spacing is systematically decreased away from the center of the grating, the angle of diffraction is greater for rays that strike the grating farther from the center, and these rays intersect the undeflected rays, which pass through the center some distance downstream of the grating. For the Fresnel zone plate, in which the grating spacing varies inversely with distance from the center, all rays intersect at the same downstream point; that is, they converge to a focus. Such devices have been demonstrated for atoms in both a spherical variety that focuses initially parallel atoms to a point and a cylindrical type that focuses parallel atoms to a line. *See* DIFFRACTION.

Atom interferometers. Atom interferometers have been demonstrated through several different experimental routes, involving both microscopic fabricated structures and laser beams. These interferometers are the first examples of optical systems composed of the elements of atom optics like those discussed above. Atom interferometers, like optical interferometers, are well suited for application to a wide range of fundamental and applied scientific problems. Scientific experiments with atom interferometers divide naturally into three major categories: measurements of atomic and molecular properties, fundamental tests and demonstrations, and inertial effects.

Atomic polarizability measurements. An atomic property that has been precisely determined with an atom interferometer is the polarizability of sodium, a quantity that parametrizes its interaction with electric fields, low-frequency radiation, and nearby atoms. In the interferometer used to measure this, a diffraction grating was used to split the atom wave into two components whose centers were separated by 55 micrometers before a second grating diffracted them back toward each other so they recombined to form interference fringes at the surface of a third grating (**Fig. 2**). This approach allowed the insertion of a stretched metal foil between the atom waves in the two sides of the interferometer, so that a uniform electric could be applied to the atom wave on one side of the interferometer.

Because an atom is polarized by an electric field, this field acts as a slight depression in the potential energy, increasing the momentum of the atom as it passes through the electric field. The result is a shorter de Broglie wavelength and therefore an increased phase accumulation relative to the wave that passes on the side of the metal foil with no field. The measured phase shift increased quadratically with the applied field, permitting a precision determination of the polarizability of the ground state of sodium (**Fig. 3**). *See* DIPOLE MOMENT.

This experiment also illustrates the paradox that results from an attempt to imagine that the atom has a classical position even when no measurement is made. The polarizability depends on the size of the atom, and the phase shift is proportional to the product of the polarizability and the time spent in the field. The observed phase shift of the fringes is consistent with the entire atom being present on the side with the electric field for the entire time of passage through the interaction region. However, the observation of the interference fringes is possible only if the atom wave splits roughly evenly between the two sides of the interferometer; indeed, the fringes

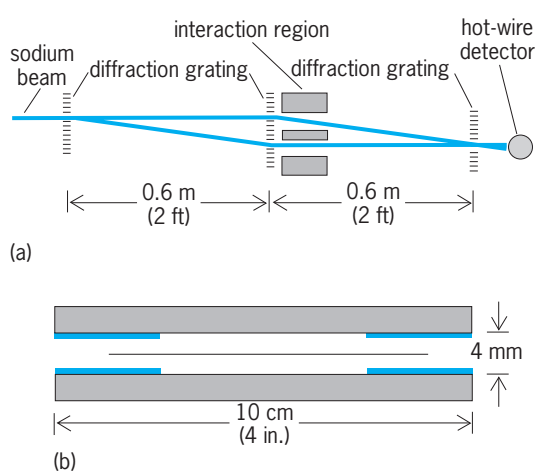

Fig. 2. Separated-beam atom interferometer. (*a*) Diagram of the apparatus. Diffraction gratings have a period of 200 nm. (*b*) Detail of the interaction region.

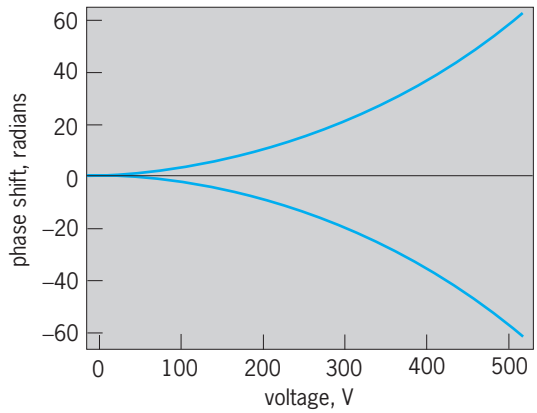

Fig. 3. Phase shift of the interference pattern of a separated-beam atom interferometer as a function of voltage applied to the left or right side of the interaction region. In either case, the magnitude of the phase shift shows a quadratic increase with voltage.

will disappear if a barrier prevents the atoms from traveling on either side of the interferometer. The quantum-mechanical description of the above situation is to add together the wave amplitudes of both possibilities (that is, the atom goes on one side and the atom goes on the other side) to predict the atom intensity observed at the detector. Along with this result comes the impossibility of localizing the atom, even in the imagination, without an intervening measurement to determine which side it traveled on.

Probing atomic interactions. Experiments have been conducted in which the atom wave on one side of the metal foil passes through a gas not present on the other side. The most useful perspective is that of conventional optics: The passage of a wave through a medium is described in terms of an index of refraction whose real part is proportional to the phase shift of the wave and whose imaginary part is proportional to the absorption of the wave. These factors shift the position and decrease the amplitude of the atom interference pattern, respectively. For sodium waves passing through helium and neon, the phase shift is found to be a much more sensitive probe of the interaction between sodium and the target gas than is the absorption. *See* ABSORPTION; REFRACTION OF WAVES; SCATTERING OF ELECTROMAGNETIC RADIATION.

Measurement of inertial effects. The sensitivity of atom interferometers to inertial effects results from the fact that there is no force on the atoms between the gratings, so that they travel along straight lines if viewed from an inertial (unaccelerating) coordinate system with respect to which the interferometer accelerates or rotates. When the atoms recombine to form interference fringes, they appear displaced because of the movement of the interferometer during the atoms' transit time. Since the de Broglie wavelength of the atoms is very small, atom interferometers are extremely sensitive to accelerations and rotations. The sensitivity of atom interferometers to rotation, which was demonstrated in the first interferometer to use light gratings, is of the order of 10^{10} greater than the phase shift in a light interferometer with the same geometry (a factor of $\sim 10^5$ from the longer transit time, and $\sim 10^5$ from the shortness of the de Broglie waves relative to light waves). Atom interferometer sensitivity to gravitation has been demonstrated at a level of 3 parts in 10^8, within an order of magnitude of the accuracy of the best absolute gravimeters. *See* ATOMIC STRUCTURE AND SPECTRA; FRAME OF REFERENCE; GRAVITY METER; INTERFERENCE OF WAVES; INTERFEROMETRY; MOLECULAR BEAMS; OPTICS; QUANTUM MECHANICS.

David E. Pritchard

Bibliography. C. S. Adams, M. Siegel, and J. Mlynek, Atom optics, *Phys. Rep.*, 240:143–210, 1994; F. Flam, Making waves with interfering atoms, *Science*, 252:921-922, 1991; R. Pool, Catching the atom wave, *Science*, 268:1129–1130, 1995; Special issue on atom optics and atom interferometry, *Appl. Phys. B*, vol. 54, 1992.

Atomic beams

Unidirectional streams of neutral atoms passing through a vacuum. These atoms are virtually free from the influence of neighboring atoms but may be subjected to electric and magnetic fields so that their properties may be studied. The technique of atomic beams is identical to that of molecular beams. For historical reasons the latter term is most generally used to describe the method as applied to either atoms or molecules.

The method of atomic beams yields extremely accurate spectroscopic data about the energy levels of atoms, and hence detailed information about the interaction of electrons in the atom with each other and with the atomic nucleus, as well as information about the interaction of all components of the atom with external fields. *See* MOLECULAR BEAMS.

Polykarp Kusch

Atomic bomb

A device for suddenly producing an explosive neutron chain reaction in a fissile material such as uranium-235 (^{235}U) or plutonium-239 (^{239}Pu). In a wider sense, any explosive device that derives its energy from nuclear reactions, including not only the foregoing fission weapon but also a fusion weapon, which gets its energy largely from fusion reactions of heavy hydrogen isotopes, and a fission-fusion weapon, which derives its energy from both fission and fusion. Because an atomic bomb derives its energy from nuclear reactions, it is properly called a nuclear explosive or nuclear weapon. *See* HYDROGEN BOMB; NUCLEAR FISSION; NUCLEAR FUSION; NUCLEAR REACTION; PLUTONIUM; URANIUM.

Materials. Of the two principal fissile materials, the cheaper but less potent ^{235}U is present in natural uranium usually in the proportion of 1 part to 139 parts of ^{238}U and is separated from it by various enrichment processes. Weapons-grade plutonium is

manufactured from ^{238}U in a special military production reactor that has enough excess neutrons for the reaction below.

$$^{238}U + n \rightarrow {}^{239}U \text{ (23-min half-life)} \rightarrow {}^{239}Np \text{ (2.3-day half-life)} \rightarrow {}^{239}Pu$$

See ISOTOPE SEPARATION; NUCLEAR REACTOR.

Weapon cores are made of very high fractions of fissile materials: highly enriched 93% uranium-235 or weapon-grade 94% plutonium-239. In most power reactors, fuel is only 3–4% uranium-235, and the plutonium-239 by-product is less than 80% isotopically pure. Although low-quality fissile materials, such as reactor-grade plutonium, can be used by expert weapon designers to make a fission explosive, such devices have not been placed in the arsenals of states known or suspected of making nuclear weapons.

Conceptual design. A fission bomb before ignition consists of a mass of fissile material and surrounding tamper—beryllium oxide or other reflector of neutrons intended ultimately to improve the neutron multiplication factor k—arranged in a geometry so favorable to neutron leakage that k is less than 1. These materials are suddenly compressed into a geometry where k substantially exceeds 1. This is done with chemical explosives that either implode a spherical subcritical mass of fission material (**illus.** *a*) or else drive two subcritical sections together in a gun-barrel type of arrangement (illus. *b*). At the same time, enough neutrons are artificially introduced to start an explosively divergent (expanding) chain reaction. *See* CHAIN REACTION (PHYSICS).

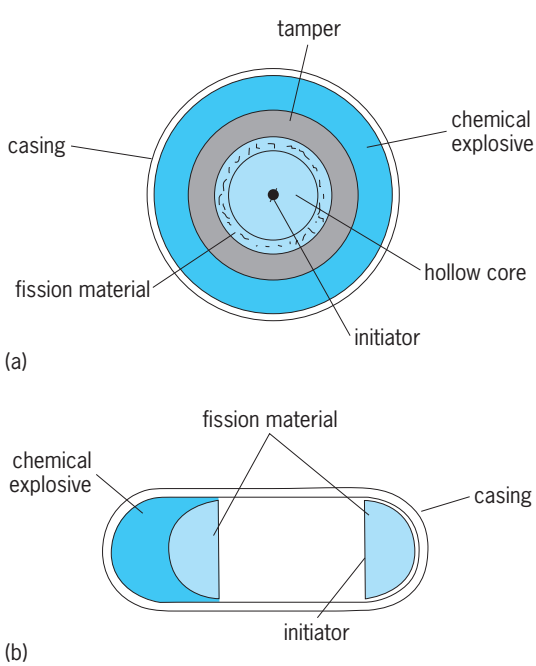

Conceptual diagrams illustrating the main features of (*a*) implosion and (*b*) gun-barrel type of fission weapons. (*After A. De Volpi et al., Born Secret: The H-Bomb, the "Progressive" Case and National Security, Pergamon, 1981*)

Fission-explosive devices intended for military application are highly sophisticated combinations of pure materials, precise design, and reliable electronics. The high-explosive lenses are shaped charges that generate and focus a compressive implosion wave. The fissile material or composite shell is levitated in a hollow core in order to allow the blast wave to impart maximum momentum to the imploding tamper-reflector. A chain-reaction initiator introduces at the optimal moment a burst of external neutrons. Either a forcibly mixed radioactive neutron source or a more useful miniature electronic neutron generator can be used for neutron injection. The latter allows controlled variability of explosive yields.

More neutrons are released in fission than are lost because of absorption and leakage, which means that a chain reaction can be maintained. With the right arrangement, the extra neutrons expelled will cause fissions in nearby fissile atoms, and the rate of fissioning grows exponentially. The rate of reaction ultimately tapers off. In a sufficiently efficient device, the reaction subsides as fission depletes the fuel (a total mass of 2.2 lb or 1 kg of heavy nuclei undergoes fission for each 17 kilotons or 7.1×10^{13} joules of energy release); but in a less efficient device, k is reduced below unity primarily through the explosive disassembly of the compact geometry. Temperatures are attained of the astrophysical order of 10^8 K, at which electrons have energies of the order 10 keV, and the energy peak of the blackbody (heat) radiation compares with that of the electronic kinetic energy. *See* HEAT RADIATION.

For optimum yield, the injection of neutrons must occur at precisely the right time as the fissile material approaches its critical configuration, within 10^{-3} s or less, depending on the type of weapon and its fissile material. In particular, if there are too many neutrons around too soon, the bomb may be triggered prematurely (predetonate) and "fizzle" (yield only a fraction of its potential). The implosion technique was developed to counter the predetonation tendency of plutonium, which has a high inherent neutron background. The poorer the quality of the plutonium, the worse the predetonation problem, and the more painstaking the weapon design becomes.

Yield. The explosive energy (yield) of a nuclear weapon is usually expressed in kilotons or megatons. A kiloton is the amount of energy liberated in the explosion of 1000 tons of TNT (10^{12} calories or 4.18×10^{12} J), and a megaton is a thousand times as large. The fission bombs that destroyed Hiroshima (gun-barrel type) and Nagasaki (implosion type) had estimated explosive yields of 13 and 22 kilotons, respectively. Fractional kiloton yields can be obtained (tactical nuclear weapons). Fission weapons have been tested up to approximately 500 kilotons, overlapping the yield of multistage fusion explosives (strategic nuclear weapons). Fission explosives are central to all nuclear warheads, including thermonuclear weapons, and some are known to be small enough to fit within a 127-mm (5-in.) artillery shell.

Acquisition of materials and components remain the major obstacle for state-supported production of military devices. Less knowledgeable individuals could conceivably assemble crude devices by using diverted materials, but highly lethal fission explosives, with yields of 1 kiloton or so, are extremely impractical with lower- (reactor-) grade materials. However, serious hazards could be generated through high-explosive dispersal of the radioactive materials. Meticulous control and safeguarding of separated fissile materials is essential, a process which is aided by the self-indicating radioactive nature of fissile materials.

Explosion effects. The nuclear explosive energy is communicated by mechanical shock and radiative transport to the surrounding water, earth, or air, ionizing it out to a radius which, in the case of explosions in air, is known as the fireball radius (150 yd or 140 m in about 1 s after a 20-kiloton nuclear explosion). Energy goes out from such a fireball into the surrounding relatively transparent air, in not very different orders of magnitude in the form of a shock wave and in the form of heat radiation that may continue for a number of seconds. *See* NUCLEAR EXPLOSION; RADIOACTIVE FALLOUT. A. De Volpi

Bibliography. A. C. Brown and C. B. McDonald, *The Secret History of the Atomic Bomb*, 1977; A. DeVolpi, Fissile materials and nuclear weapons proliferation, *Annu. Rev. Nuc. Particle Sci.*, 36:83–114, 1986.

Atomic clock

A device that uses an internal resonance frequency of atoms (or molecules) to measure the passage of time. The terms atomic clock and atomic frequency standard are often used interchangeably. A frequency standard generates pulses at regular intervals. A frequency standard can be made into a clock by the addition of an electronic counter, which records the number of pulses. *See* DIGITAL COUNTER.

Basic principles. Most methods of timekeeping rely on counting some periodic event, such as the rotation of the Earth, the motion of a pendulum in a grandfather clock, or the vibrations of a quartz crystal in a watch. An atomic clock relies on counting periodic events determined by the difference of two different energy states of an atom. According to quantum mechanics, the internal energy of an atom can assume only certain discrete values. A transition between two energy states with energies E_1 and E_2 may be accompanied by the absorption or emission of a photon (particle of electromagnetic radiation). The frequency ν of this radiation is given by the equation below, where h is Planck's constant. A basic

$$h\nu = |E_2 - E_1|$$

advantage of atomic clocks is that the frequency-determining elements, atoms of a particular isotope, are the same everywhere. Thus, atomic clocks constructed and operated independently will measure the same time interval, that is, the length of time between two events. In order for the two clocks to agree on the time, they must be synchronized at some earlier time. *See* ATOMIC STRUCTURE AND SPECTRA; ENERGY LEVEL (QUANTUM MECHANICS); QUANTUM MECHANICS.

An atomic frequency standard can be either active or passive. An active standard uses as a reference the electromagnetic radiation emitted by atoms as they decay from a higher energy state to a lower energy state. An example is a self-oscillating maser. A passive standard attempts to match the frequency of an electronic oscillator or laser to the resonant frequency of the atoms by means of a feedback circuit. The cesium atomic beam and the rubidium gas cell are examples of passive standards. Either kind of standard requires some kind of frequency synthesis to produce an output near a convenient frequency, such as 5 MHz, that is proportional to the atomic resonance frequency. *See* FEEDBACK CIRCUIT; LASER; MASER; OSCILLATOR.

Accuracy and stability. Two different gages of the quality of a clock are accuracy and stability. The accuracy of a frequency standard is defined in terms of the deviation of its frequency from an ideal standard. In practice, it might be defined in terms of the frequency differences measured between independently constructed and operated standards of the same type. Improving the accuracy depends on understanding and controlling all the parameters that might cause the frequency to shift. The stability of a frequency standard is defined in terms of the constancy of its average frequency from one interval of time to the next. For many frequency standards, the stability initially improves with increasing measurement time but eventually gets worse. That is, a more precise measurement of the frequency can be made by averaging together successive measurements, until some imperfection in the apparatus causes the frequency to change. The stability increases with increased Q (resonance frequency divided by the width of the resonance) and with increased measurement signal-to-noise ratio. *See* Q (ELECTRICITY); SIGNAL-TO-NOISE RATIO.

Common types. The three most commonly used types of atomic clock are the cesium atomic beam, the hydrogen maser, and the rubidium gas cell. The cesium clock has high accuracy and good long-term stability. The hydrogen maser has the best stability for periods of up to a few hours. The rubidium cell is the least expensive and most compact and also has good short-term stability.

Cesium atomic-beam clock. This clock (**Fig. 1**) uses a 9193-MHz transition between two hyperfine energy states of the cesium-133 atom. Both the atomic nucleus and the outermost electron have magnetic moments; that is, they are like small magnets, with a north and a south pole. The two hyperfine energy states differ in the relative orientations of these magnetic moments. The cesium atoms travel in a collimated beam through an evacuated region. Atoms in the different hyperfine states are deflected into different trajectories by a nonuniform magnetic field. Atoms in one of the two states are made to pass

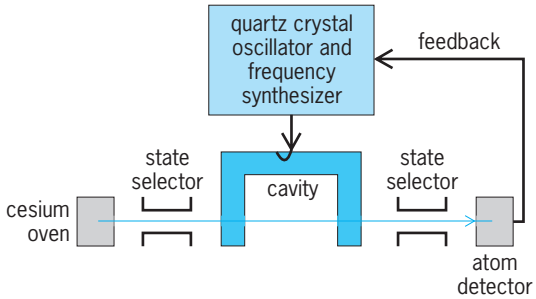

Fig. 1. Cesium atomic-beam clock. (*After G. Kamas, ed., Time and Frequency Users' Manual, NBS Tech. Note 695, 1977*)

through a microwave cavity, where they are exposed to radiation near their resonance frequency. The resonant radiation may cause the atom to make a transition from one state to the other; if that happens, the atom is deflected by a second, nonuniform magnetic field onto a detector. *See* ELECTRON SPIN; HYPERFINE STRUCTURE; MAGNETIC MOMENT; MOLECULAR BEAMS; NUCLEAR MOMENTS.

The Q of the resonance is over 10^8 for some laboratory standards and somewhat less for the smaller standards that are commercially available. Cesium atomic beams, including variants such as optically pumped atomic beams and atomic fountains, are the most accurate of all atomic clocks. The best models have an error of less than 1 part in 10^{15}, or about 1 s in 3×10^7 years. For this reason, cesium has become the basis of the international definition of the second: the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine states of the ground state of the cesium-133 atom. The cesium clock is especially well suited for applications such as timekeeping, where absolute accuracy without recalibration is necessary. Measurements from many cesium clocks throughout the world are averaged together to define an international time scale that is uniform to parts in 10^{14}, or about 1 microsecond in a year. *See* ATOMIC TIME; DYNAMICAL TIME; PHYSICAL MEASUREMENT.

Hydrogen maser. This instrument (**Fig. 2**) is based on the hyperfine transition of atomic hydrogen, which has a frequency of 1420 MHz. Atoms in the higher hyperfine energy state are selected by a focusing magnetic field, so that they enter an evacuated storage bulb inside a microwave cavity. The atoms bounce off the poly(tetrafluoroethylene)-coated walls for about 1 s before they are induced to make a transition to the lower hyperfine state by a process called stimulated emission. The stimulated emission from many atoms creates a self-sustaining microwave oscillation.

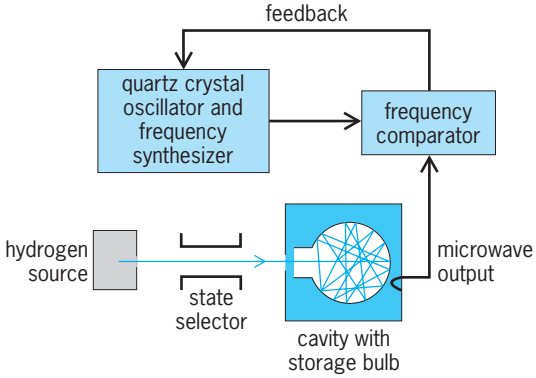

Fig. 2. Hydrogen maser. (*After G. Kamas, ed., Time and Frequency Users' Manual, NBS Tech. Note 695, 1977*)

The resonance Q is about 10^9. The best hydrogen masers have a stability of about 1 part in 10^{15} for averaging periods of 10^4 s. Over longer periods of time, the frequency drifts, primarily because of changes of the cavity tuning. Collisions with the walls cause the frequency to be shifted by about 1 part in 10^{11} relative to that of a free atom, but the magnitude of the shift varies from one device to another. This shift limits the accuracy of the hydrogen maser to about 1 part in 10^{12}.

The hydrogen maser can also be operated as a passive device, with improved long-term stability, due to the addition of automatic cavity tuning. The short-term stability is worse than that for an active maser.

Rubidium gas cell. This device (**Fig. 3**) is based on the 6835-MHz hyperfine transition of rubidium-87.

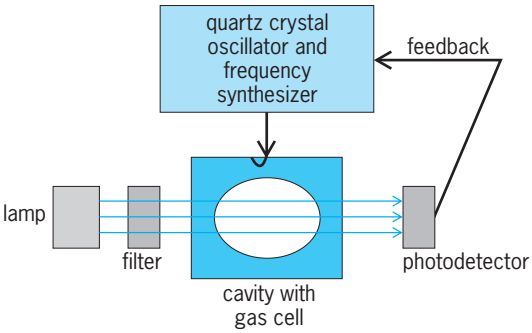

Fig. 3. Rubidium gas cell. (*After G. Kamas, ed., Time and Frequency Users' Manual, NBS Tech. Note 695, 1977*)

The rubidium atoms are contained in a glass cell together with a buffer gas, such as argon, that prevents them from migrating to the cell walls. A method called optical pumping is used to prepare the atoms in one hyperfine state. Filtered light from a rubidium resonance lamp is absorbed by atoms in one of the two hyperfine states, causing them to be excited to a higher state, from which they quickly decay to the other hyperfine state. If the atoms are then subjected to microwave radiation at the hyperfine transition frequency, they are induced to make transitions back to the other hyperfine state. They can then absorb light again from the lamp; this results in a detectable decrease in the light transmitted through the cell.

The Q is only about 10^7, but the short-term stability is quite good, reaching 1 part in 10^{13} for averaging times of 1 day. After longer periods, changes in the buffer gas pressure and the lamp cause the frequency to drift. The accuracy is not better than 1 part in 10^{10}. Rubidium standards are used in applications that do not require the accuracy of a cesium standard.

Laboratory types. Many other kinds of atomic clocks, such as thallium atomic beams and ammonia and rubidium masers, have been demonstrated in the laboratory. The first atomic clock, constructed at the National Bureau of Standards in 1949, was based on a 24-GHz transition in the ammonia molecule.

Some laboratories have tried to improve the cesium atomic-beam clock by replacing the magnetic state selection with laser optical pumping and fluorescence detection. Improved performance is expected because of increased signal-to-noise ratio and a more uniform magnetic field. One such standard, called NIST-7, was in operation at the U.S. National Institute of Standards and Technology and was formerly the primary frequency standard for the United States. Other laboratories have studied atomic-beam standards based on magnesium, calcium, or methane, which have frequencies higher than that of cesium.

Atomic frequency standards can also be based on optical transitions. One of the best-developed optical frequency standards is the 3.39-micrometer (88-THz) helium-neon laser, stabilized to a transition in the methane molecule. Research is being carried out on the use of laser-cooled, neutral atoms, such as calcium or strontium, as optical frequency standards.

Frequency synthesis chains have been built to link the optical frequency to radio frequencies. A new device for optical frequency synthesis, the femtosecond-laser frequency comb, has largely replaced the earlier, more cumbersome methods based on the frequency doubling and mixing of many separate lasers. In this device, a femtosecond mode-locked laser generates an optical spectrum made up of many equally spaced components. The frequency of a particular component is given by $f(m) = f_0 + mf_{rep}$, where m is an integer. The pulse repetition rate f_{rep} of the mode-locked laser can easily be measured by detecting the light with a photodiode, while the offset frequency f_0 can be determined by a method called self-referencing. The frequency of another laser can be determined from the beatnote with the nearest component of the spectrum. This provides a method to transfer the stability of an optical frequency to the frequency of the pulse repetition rate, which is in the radio-frequency region. This makes possible an optical clock, given a stable optical frequency.

Ion traps, which confine ions in a vacuum by electric and magnetic fields (**Fig. 4**), have been studied for use in both microwave and optical atomic clocks. They provide a benign environment for the ions while still allowing a long measurement time. Microwave clocks based on buffer-gas-cooled, optically pumped mercury-199 or ytterbium-171 ions have been built and show good stability. Values of Q as high as 1.5×10^{13} have been observed on the ytterbium-171 hyperfine transition. Other trapped ion standards make use of laser cooling to reduce frequency errors due to Doppler shifts. Laser cooling is a method by which resonant light pressure is used to damp the motion of atoms. A microwave clock based on laser-cooled mercury-199 ions has demonstrated an accuracy close to that of the best cesium clocks. It

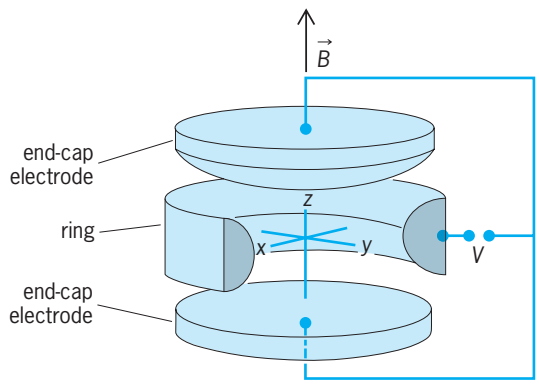

Fig. 4. Electrodes used to create the confining electric potential for a Penning ion trap or a Paul ion trap. An electric potential *V*, which is static for a Penning trap and oscillating for a Paul trap, is applied between the ring electrode and the end-cap electrodes. The Penning trap requires a uniform magnetic field *B*. (*After F. J. Rogers and H. E. Dewitt, eds., Strongly Coupled Plasma Physics, Plenum, 1987*)

is possible to observe extremely high values of Q on optical transitions of trapped ions. An optical transition has been observed in a single, trapped mercury ion with a Q of 1.6×10^{14}. Similar results have been obtained with other ions, such as strontium-88 and ytterbium-171. An optical frequency standard based on such an ion might be capable of an accuracy of 1 part in 10^{18}. *See* LASER COOLING; PARTICLE TRAP.

The atomic fountain (**Fig. 5**) is another method of using laser-cooled atoms in a frequency standard. In this method, neutral atoms, such as cesium, are laser cooled and state selected and then launched upward on a ballistic trajectory. The atoms pass through a microwave cavity, as in the atomic-beam method. They are moving slowly enough that they fall back down through the cavity a second time, after which they are detected by laser-induced fluorescence. The general principles are the same as in an atomic-beam standard, but greater precision is possible because the flight time of an atom through the apparatus is longer. Cesium atomic fountain clocks have now replaced cesium atomic beams at national measurement laboratories, such as the National Institute of Standards and Technology in Boulder, Colorado. The best ones have accuracies better than 1 part in 10^{15}.

The coherent population trapping (CPT) clock is a new type of microwave atomic clock, which is more interesting for its compact size than for its accuracy and stability. As in the rubidium cell clock, a hyperfine resonance in gas-phase atoms is excited. Rather than being excited directly by microwave radiation at the hyperfine transition frequency, the resonance is excited by two laser light fields whose frequency difference is equal to the hyperfine transition frequency. In practice, the two fields are generated by modulation of the frequency of a single laser. Such a clock can be made very compact, since there is no need for a microwave cavity to enhance a microwave field. CPT clocks based on the cesium hyperfine transition are currently being developed. In some of them, the cesium gas cell measures only a few millimeters on a side, and the entire device, exclusive of the power

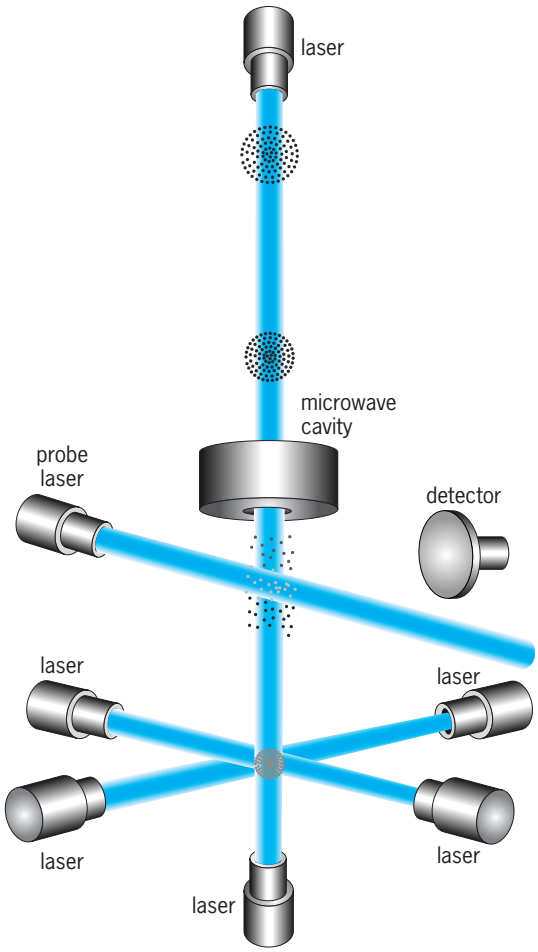

Fig. 5. Cesium atomic fountain clock. A sample of cesium atoms is trapped and laser-cooled at the intersection of six laser beams. The atoms are then launched upward, passing through the microwave cavity, then falling back through the cavity. After passing through the cavity the second time, the atoms are illuminated by the probe laser, and the induced fluorescence is detected. (*National Institute of Standards and Technology*)

supply, might fit into a volume of 1 cm^3.

Applications. Atomic clocks are used in applications for which less expensive alternatives, such as quartz oscillators, do not provide adequate performance. The use of atomic clocks in maintaining a uniform international time scale has already been mentioned; other applications are described below. *See* QUARTZ CLOCK.

Navigation. The Global Positioning System is a satellite-based system that enables a user with a suitable radio receiver to determine position within about 10 m (33 ft). The satellites send out accurately timed radio pulses, from which the user's receiver can calculate its location and time. The satellites and the ground stations, but not the users, need atomic clocks (usually cesium clocks). *See* SATELLITE NAVIGATION SYSTEMS.

Communications. Various digital communications systems require precise synchronization of transmitters and receivers in a network. Some systems use time-division multiplexing, in which many channels of information are sent over the same line by sequentially allotting a small time slot to each channel. Timing is very critical when there are several sources of information with their own clocks. The primary timing is provided by cesium clocks. *See* ELECTRICAL COMMUNICATIONS; MULTIPLEXING AND MULTIPLE ACCESS; PULSE MODULATION.

Radio astronomy. Very long baseline interferometry is a technique that allows two or more widely separated radio telescopes to achieve very high angular resolution by correlation of their signals. The system has the resolution that a single telescope would have if its aperture were equal to the distance between the telescopes. This can be thousands of miles. The stable timing needed to correlate the signals is provided by hydrogen masers. *See* RADIO ASTRONOMY; RADIO TELESCOPE.

Space exploration. Navigation of space probes by Doppler tracking requires very stable local oscillators, derived from atomic frequency standards. Doppler tracking relies on determining the velocity of the spacecraft by measuring the frequency shift of a signal after it has been echoed to the Earth by a transponder on the spacecraft. Stable local oscillators are also needed for studies of planetary atmospheres and rings by fluctuations of the radio signals transmitted through them. *See* SPACE NAVIGATION AND GUIDANCE.

Fundamental science. According to A. Einstein's special and general theories of relativity, a moving clock runs slower than a stationary one, and a clock on the Earth's surface runs slower than one far from the Earth. These predictions were verified to high accuracy by an experiment in which a hydrogen maser was launched in a rocket to an altitude of 10,000 km (6000 mi). *See* CLOCK PARADOX; RELATIVITY; TIME.

Wayne M. Itano

Bibliography. J. C. Bergquist, S. Jefferts, and D. J. Wineland, Time measurement at the Millennium, *Phys. Today*, 54(3):37–42, 2001; W. M. Itano and N. F. Ramsey, Accurate measurement of time, *Sci. Amer.*, 269(1):56–65, July 1993; J. Jespersen and J. Fitz-Randolph, *From Sundials to Atomic Clocks*, National Institute of Standards and Technology, Monogr. 155, U.S. Government Printing Office, Washington, D.C., 1999; T. Jones, *Spliting the Second: The Story of Atomic Time*, Institute of Physics Publishing, Bristol and Philadelphia, 2000; F. G. Major, *The Quantum Beat: The Physical Principles of Atomic Clocks*, Springer, New York, 1998; N. F. Ramsey, Precise measurement of time, *Amer. Sci.*, 76(1):42–49, 1988.

Atomic Fermi gas

A gas of atoms, generally ultracold, in which the individual atoms possess half-integer spin (that is, are fermionic). Since the 1980s, researchers have developed powerful new techniques for cooling atoms to ultralow temperatures. Among the most significant achievements made possible by these technical developments was the creation in 1995 of the long-sought Bose-Einstein condensate. In a

Bose-Einstein condensate, atoms are cooled to such a low temperature that they collect in the quantum-mechanical ground state of their confinement volume. Bose-Einstein condensation is possible only in the class of particles known as bosons. Particles of the other class, fermions, are forbidden to occupy the same quantum state and are thus prevented from directly condensing.

Progress in studying ultracold fermions has lagged behind that of bosons, not because of lack of interest, but because they are experimentally more difficult to cool. In fact, fermions are especially significant because the fundamental building blocks of matter, protons, neutrons, and electrons, are all fermions. Most of the field of condensed matter physics is concerned with the behavior of fermions, especially electrons. Of particular interest is the phenomenon of Cooper pairing, where fermions, such as the electrons in a superconductor, or the atoms in helium-3 (^{3}He), form correlated pairs when cooled to sufficiently low temperature. Recent experimental progress in cooling atomic Fermi gases has enabled the realization of Cooper pairing in the gas phase. Because of the inherent simplicity and tunability of the atomic interactions in an ultracold gas, this work may help clarify many troubling issues in condensed-matter physics, including the mechanisms at work in high- temperature superconductors.

Behavior of bosons and fermions. We have known since the work of Satyendranath Bose, Albert Einstein, Enrico Fermi, and Paul Dirac in the 1920s that there are two fundamental types of particles. Bosons are particles with integer spin angular momentum ($0, \hbar, 2\hbar, \ldots$) and fermions are those with half integer ($^1/_2\hbar, {}^3/_2\hbar, \ldots$). Since protons, neutrons, and electrons are each spin-$^1/_2$ particles, atoms made with an odd number of these constituents are themselves composite fermions. Examples are lithium-6 (^{6}Li) and potassium-40 (^{40}K), which happen to be the only stable fermionic isotopes of the alkali-metal elements. The differences between bosons and fermions are significant only when they become quantum-degenerate, that is, when the size of the atomic quantum-mechanical de Broglie wavelength, Λ, which increases as temperature decreases, is comparable to the average separation between atoms. When this condition is met, bosons undergo the Bose-Einstein condensation phase transition. Identical fermions, on the other hand, are forbidden by the Pauli exclusion principle from occupying the same quantum level (**Fig. 1**). As Cooper pairs are composite bosons, superconductivity and fermionic superfluidity are thought to be related to Bose-Einstein condensation of pairs. *See* BOSE-EINSTEIN CONDENSATION; EXCLUSION PRINCIPLE; QUANTUM MECHANICS; QUANTUM STATISTICS.

Cooling methods. The methods for cooling Fermi gases and for confining them in atom traps are similar to those used for bosons. The primary methods are laser cooling, atom trapping, and evaporative cooling. Laser cooling utilizes radiation pressure to slow atoms. Since temperature is related to the kinetic energy of the atoms, reducing their speed is equivalent

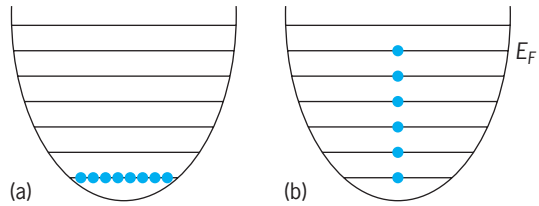

Fig. 1. Bosons and fermions at absolute temperature $T = 0$. Atoms in a magnetic trap feel a harmonic restoring potential, as indicated by the parabolic wells. The equally spaced lines in each well represent the quantized energy levels of the harmonic potential. (*a*) Bosons. At $T = 0$, they collectively occupy the ground level of the system. (*b*) Fermions. These particles must obey the Pauli exclusion principal, which forbids more than one identical fermion per level. At $T = 0$, they completely fill the energy-level ladder up to the Fermi energy E_F. The filling of energy levels in this manner is analogous to the way electrons fill the lowest available atomic orbitals in the periodic table of the elements.

to cooling. Laser cooling is typically limited to temperatures in the range of 100 millionths of a degree above absolute zero, or 100 microkelvin. Atom traps use laser radiation pressure or magnetic fields to confine atoms without walls. Since the atom temperatures are so low, isolation from physical boundaries is absolutely necessary to prevent heating.

Even lower temperatures, of the order of 100 nanokelvin (100 billionths of a degree above absolute zero), are needed to achieve quantum degeneracy in low-density atomic gases. These temperatures, as low as any produced by any method or found anywhere in the universe, are attainable using evaporative cooling. Evaporative cooling exploits the fact that there is always a distribution of energies in a gas, including some very energetic atoms in the "tail" of the distribution. By systematically removing only the hottest atoms from the trap, a large fraction of energy can be removed, while minimizing the number of atoms lost.

Effective evaporative cooling, however, requires that the atoms in the gas undergo elastic thermalizing collisions to continuously replenish the tail of hot atoms. The main experimental difficulty encountered with fermions is that the Pauli exclusion principle prevents identical fermions from interacting, so they are unable to thermalize. Several ways around this difficulty have been developed. First, one can make use of the different projections of the nuclear and atomic spin of the atom to make subsets of atoms that are distinguishable from other subsets. For example, one can make a mixture of ^{40}K atoms in two different nuclear spin states, so that the nuclear spins of some of the atoms point in a different direction than the others. Atoms in the same spin state will not interact, but two atoms in different spin states will. Unfortunately, this method cannot be universally used, because some spin states cannot be magnetically trapped, and because collisions between some spin states can result in the creation of atoms in a third state, with a release of excess energy that sends the atoms flying from the trap.

A more general way to cool fermions is to "sympathetically" cool them with bosons. The bosons are

evaporatively cooled as usual, while the fermions are cooled by interaction with the bosons. In this way, ^{6}Li has been cooled to temperatures as low as 20 nanokelvin using either lithium-7 (^{7}Li) or sodium-23 (^{23}Na) atoms as the "refrigerant." After the fermions have been cooled, the bosons can be selectively removed from the trap by driving transitions to untrapped quantum states using microwave or optical pulses. The gas is probed by shining a laser beam through the gas of atoms and imaging the shadow cast by the atoms onto a charge-coupled-device (CCD) camera. *See* LASER COOLING; PARTICLE TRAP.

Fermi pressure. In 1926, Fermi realized that the energy of a gas of fermions is finite at zero absolute temperature (Fig. 1), and consequently it exerts a pressure. This "Fermi pressure" is a purely quantum-mechanical effect, as the pressure p from the classical ideal gas law, $p = nkT$, vanishes as the temperature, T, approaches 0. (In this expression, n is the number density of the gas, and k is the Boltzmann constant.) The implications of this result were soon realized by S. Chandrasekhar: under certain conditions electron Fermi pressure could prevent the gravitational collapse of relatively cold, aging stars known as white dwarfs. This prediction was verified by observation, and the same stabilization method was later found to apply to neutron stars. The effect of Fermi pressure has also been demonstrated for a trapped gas of ^{6}Li atoms (**Fig. 2**). The figure shows a gas containing both bosons (^{7}Li) and fermions (^{6}Li), at three different temperatures. Although both types of atoms are mixed together in the trap at the same time, they can be individually imaged. The images are separated horizontally in the figure to show the difference in behavior between bosons and fermions. Figure 3*a* corresponds to the highest temperature, $T \sim T_F$, where T_F is the Fermi temperature (Fig. 1). At this relatively high temperature, there is little difference in the spatial size of the fermions compared with the bosons. However, as the temperature is reduced to $\sim 0.5\ T_F$ in Fig. 3*b*, and finally to $\sim 0.25\ T_F$ in Fig. 3*c*, it becomes clear that the fermions occupy more volume than do the bosons. This difference is a manifestation of Fermi pressure, albeit on a much different size and energy scale than for stars. *See* FERMI-DIRAC STATISTICS; NEUTRON STAR; WHITE DWARF STAR.

Cooper pairing in an atomic gas. One of the most remarkable achievements in physics in recent years was the realization of Cooper pairing in an ultracold atomic Fermi gas. The pairing of fermions is a phase transition, just as Bose-Einstein condensation is for bosons. In fact, these phase transitions are closely related because a pair of fermions is itself a composite boson. The realization of Cooper pairing required that the atoms be cooled to even lower temperatures than have been demonstrated thus far. Furthermore, the atoms must attract one another. Since identical ultracold fermions cannot interact, as was the case for evaporative cooling, the experimental schemes use a mixture of fermions in two different spin states. A powerful technique for producing strong, tunable interactions has been developed using a Feshbach resonance, in which a magnetic field tunes the energy of a bound state of two atoms, a diatomic molecule, to be equal to the energy of a pair of free atoms. The resulting collisional resonance can be used to create strong or weak interactions that are either repulsive or attractive depending on the value of the magnetic field. Because of the versatility it affords, this interaction "knob" is one of the most compelling aspects of quantum gas research. *See* RESONANCE (QUANTUM MECHANICS).

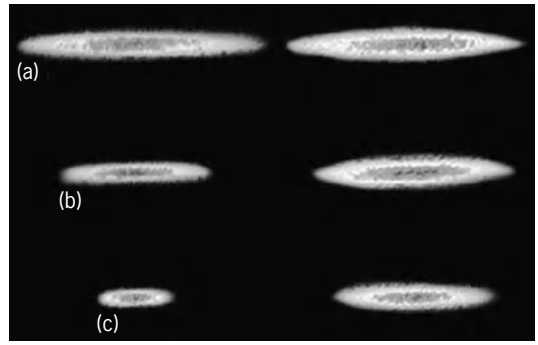

Fig. 2. Fermi pressure in a gas of trapped atoms. The images on the left are of the bosons (^{7}Li), while those on the right are of the fermions (^{6}Li) taken at progressively lower temperatures. Each pair of images corresponds to a particular temperature. The ^{6}Li and ^{7}Li atoms are simultaneously trapped in the same volume, but the images are separated for clarity. (*a*) Temperature $T \cong 1.5\ T_c \cong T_F$, where T_c is the critical temperature for Bose-Einstein condensation, and T_F is the Fermi temperature. (*b*) $T \cong 0.5\ T_F$. (*c*) $T \cong T_c \cong 0.25\ T_F$. (*From A. G. Truscott et al., Observation of Fermi pressure in a gas of trapped atoms, Science, 291:2570–2572, 2001*)

Traditional superconductors and superfluid ^{3}He form pairs in the weakly interacting limit, where the size of the Cooper pairs is much larger than the average distance between fermions. This situation is very well described by the BCS (Bardeen-Cooper-Schrieffer) theory of superconductivity. On the other hand, the interactions in high-temperature superconductors are believed to be much stronger,

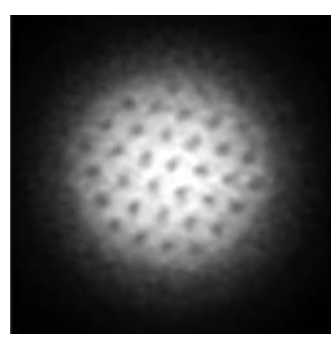

Fig. 3. Vortex array in a paired gas of ^{6}Li atoms. Vortices with quantized angular momentum form when the gas is stirred. They appear as dark regions in the figure because centrifugal forces push atoms away from the vortex core. Interactions between atoms cause the vortices to form a regular lattice array. Such an array of quantized vortices is very similar to those created in Bose-Einstein condensates of ultracold bosonic atoms and is a hallmark of superfluidity. (*From M. W. Zwierlein et al., Vortices and superfluidity in a strongly interacting Fermi gas, Nature, 435:1047–1051, 2005*)

and hence the pairs are spatially smaller. Feshbach resonances have enabled the investigation of pairing in atomic Fermi gases over a range of interaction strength, but the experiments thus far have explored mainly the strongly interacting regime where the pairs are almost molecular in character. In this regime, there is yet no complete theory of Cooper pairing, and so the experiments provide new insight into this physical realm. The evidence for Cooper pairing obtained in recent years includes direct measurements of the correlation between atom pairs, measurements of the heat capacity, and the observation of quantized vortices (**Fig. 3**), a hallmark of superfluidity and one of the consequences of Cooper pairing. *See* LIQUID HELIUM; QUANTIZED VORTICES; SUPERCONDUCTIVITY; SUPERFLUIDITY.

Randall Hulet

Bibliography. B. DeMarco and D. S. Jin, Onset of Fermi degeneracy in a trapped atomic gas, *Science*, 285:1703–1706, 1999; D. Jin, A Fermi gas of atoms, *Phys. World*, 15(4):27–32, April, 2002; H. J. Metcalf and P. van der Straten, *Laser Cooling and Trapping*, Springer, 1999; C. J. Pethick and H. Smith, *Bose-Einstein Condensation in Dilute Gases*, Cambridge, 2002; A. G. Truscott et al., Observation of Fermi pressure in a gas of trapped atoms, *Science*, 291:2570–2572, 2001; M. W. Zwierlein et al., Vortices and superfluidity in a strongly interacting Fermi gas, *Nature*, 435:1047–1051, 2005.

Atomic force microscopy

A form of microscopy in which a sharp tip is scanned over the surface of a sample while sensing the interaction force between the tip and the sample. Due to its ability to observe and manipulate surfaces in aqueous solution, atomic force microscopy (AFM) has revolutionized the way in which microscopists explore biological structures, from single molecules to living cells.

Principle and technique. The basic idea behind AFM is to use so-called near-field physical interactions (that is, repulsive or attractive forces) acting between a sharp tip and the specimen to generate three-dimensional images of the surface without dependence on an incident beam, as in optical and electron microscopies (**Fig. 1**). The sample is mounted on a piezoelectric scanner that allows high-resolution, three-dimensional positioning. The force interacting between tip and specimen is measured by the bending, or "deflection," of a soft cantilever spring, typically made of silicon or silicon nitride. The cantilever deflection is detected by a laser beam, usually focused on the free end of the cantilever, and reflected into a photodiode. *See* INTERMOLECULAR FORCES; PIEZOELECTRICITY; SURFACE PHYSICS.

Operational modes. AFM can be operated in several different modes: contact, tapping, noncontact, and force spectroscopy.

Contact mode. In contact mode, the tip and sample are placed in contact, and the tip is simply dragged across the surface, resulting in a topographic image of the surface. The scanning is usually done under feedback control, in which the sample is moved toward or away from the cantilever during the scan so that the bending of the cantilever, normal to the surface, remains constant in order to maintain constant force. The up and down motion of the sample is, therefore, a record of the sample topography.

Tapping mode. In tapping mode (also known as intermittent mode), the cantilever is oscillated near its resonant frequency using a piezoelectric actuator. Moving the oscillating tip until it lightly touches the surface reduces the oscillation amplitude. The reduction in oscillation amplitude now becomes the feedback control signal, which can be used to measure the surface topography. Due to its intermittent sample contact, tapping mode AFM can significantly reduce lateral forces during imaging, which may be advantageous for imaging soft biological samples.

Noncontact mode. AFM may also be operated in the noncontact mode, in which the tip is oscillated at a distance from the surface where attractive van der Waals forces dominate.

Force spectroscopy. Besides recording images, AFM can measure minute forces within or between biomolecules with piconewton (pN; 10^{-12} N) sensitivity. In this operating mode, known as force spectroscopy, the cantilever deflection is recorded as a function of the vertical displacement of the piezoelectric scanner; that is, deflection is recorded as the sample is pushed toward the tip and retracted. The resulting force versus separation distance curves provide a wealth of information on biophysical properties and biomolecular interactions and permit the manipulation of single biomolecules.

Nanoscale visualization of biological structures. Since the late 1980s, AFM has proved very useful in imaging the structure of biomolecules, lipid membranes, membrane proteins, and cells. Remarkably, images can be recorded in real-time under physiological conditions, with unprecedented signal-to-noise

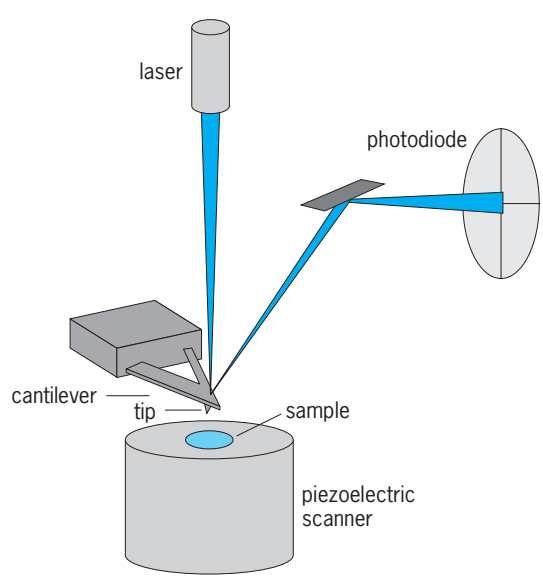

Fig. 1. Principle of AFM.

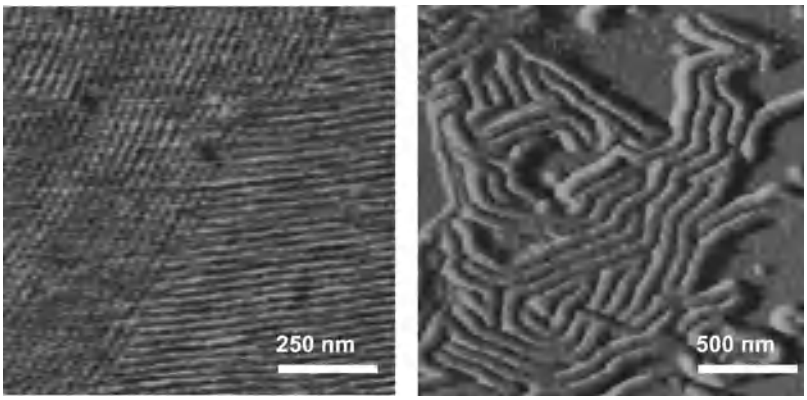

(a) (b)

Fig. 2. Application of AFM imaging to lipid membranes. (*a*) Nanostructures with well-defined bendings and a ~30 nm periodicity, formed at the surface of a mixed dipalmitoylphosphatidylcholine/surfactin bilayer supported on mica. (*b*) Another type of nanostructure with domains protruding ~5 nm above the surface of a dipalmitoylphosphatidylcholine/dioleoylphosphatidic acid bilayer imaged in the presence of the simian immunodeficiency virus fusion peptide.

ratio and resolution, thereby complementing the range of instruments available to investigate structural properties. The power of this approach for probing lipid bilayers is illustrated in **Fig. 2**. The AFM images, recorded directly in buffer solution, reveal that remarkable nanostructures are formed at the bilayer surface in the presence of lipopeptide or fusion peptide. This example illustrates how AFM is enabling biophysical researchers to investigate membrane structure, properties, and interactions (such as molecular packing, structural defects, nanodomains, interaction with external agents, fusion and adhesion events). These nanoscale studies are also expected to have an important impact on nanobiotechnology for the design of biosensor and biomimetic surfaces (synthetic surfaces that mimic biological surfaces). *See* BIOELECTRONICS; BIOSENSOR; NANOSTRUCTURE.

The stability and structural regularity of two-dimensional arrays of membrane proteins make them well suited for high-resolution imaging. In the past decade, progress in AFM instrumentation and methodology has allowed various membrane proteins to be explored with subnanometer resolution, such as the hexagonally packed intermediate (HPI) protein layer of *Deinococcus radiodurans*, the purple membrane of the archaean *Halobacterium*, and porin (protein channel) crystals of *Escherichia coli*. Remarkably, conformational changes of individual membrane proteins can be detected and related to function. For instance, voltage and pH-dependent conformational changes were observed in bacterial porins and correlated with the closure of the channel entrance, suggesting that this is a mechanism to protect the cells from drastic changes in the environment. *See* CELL MEMBRANES; CELL PERMEABILITY.

Another exciting feature of the instrument is the possibility to monitor dynamic processes using time-lapse imaging; that is, for the first time microscopists can observe single biomolecules at work, such as ribonucleic acid (RNA) polymerase moving along deoxyribonucleic acid (DNA) during the transcription process, and can follow dynamic events in living cells, such as cell growth and division. *See* DEOXYRIBONUCLEIC ACID (DNA); RIBONUCLEIC ACID (RNA).

There are several technical challenges to address in the future, including reducing tip–specimen interaction forces for high-resolution imaging of soft samples such as cells and improving time resolution to probe dynamic processes such as molecular conformational changes.

Detecting and manipulating single biomolecules. The piconewton force sensitivity of AFM force spectroscopy enables it to be used to manipulate single biomolecules and to measure their inter- and intramolecular interactions in relation to function. In particular, the AFM tip can be used to pull on flexible macromolecules (such as DNA, proteins, and polysaccharides) to learn about their elastic behavior as well as their conformational and folding/unfolding properties. While pioneering experiments were performed on pure molecules, such as the polysaccharide dextran, recent studies have also focused on membranes and cells. For instance, force curves recorded via AFM on the bacterial HPI layer showed multiple adhesion peaks that were attributed to the sequential unzipping of entire bacterial pores. Remarkably, the resulting molecular defects could be directly visualized in the manipulated specimen using high-resolution imaging. Another example of such a nanodissection experiment is the extraction of individual bacteriorhodopsins from purple membranes.

Measuring molecular recognition forces. AFM force spectroscopy is also well suited to measure molecular recognition forces between receptors and ligands, which play important roles in the life sciences and form the basis of many biotechnological applications. By attaching specific biomolecules on the AFM tips, researchers can measure the binding forces between complementary molecules, such as biotin–avidin, antibody–antigen, and DNA strands. The measured forces are usually in the range of 50–200 pN, depending on the nature of the molecules and the

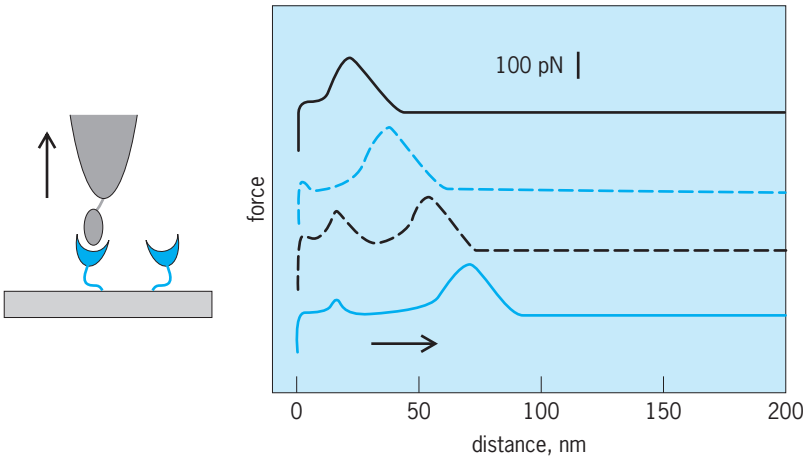

Fig. 3. Application of AFM force spectroscopy to the detection of single molecular recognition events. Force–distance curves were recorded between an AFM tip and a substrate functionalized with oligoglucose and concanavalin A, respectively. The ~100 pN unbinding forces reflect the interaction between single lectin and carbohydrate molecules.

pulling rate. An example of such an experiment is shown in **Fig. 3**. To measure lectin–carbohydrate interactions, a solid substrate was modified with the lectin concanavalin A and complementary oligoglucose molecules were attached to an AFM tip. Force–distance curves recorded between the modified tip and substrate showed unbinding forces of ~100 pN that were attributed to the specific interaction between individual lectin and carbohydrate molecules.

Besides providing fundamental insights into the properties and functions of biomolecules, these AFM-based analyses have important promise in medicine for the rapid, highly sensitive detection of biomolecules such as antigens and toxins. Moreover, biotips can also be used to detect specific molecules on cell surfaces and to map their distribution, opening the door to new applications in cell biology and microbiology for elucidating the molecular bases of cell-surface processes such as cell adhesion. *See* CELL (BIOLOGY); CELL ADHESION; MOLECULAR RECOGNITION.

Yves F. Dufrêne

Bibliography. G. Binnig, C. F. Quate, and C. Gerber, Atomic force microscope, *Phys. Rev. Lett.*, 56:930–933, 1986; R. J. Colton et al., *Procedures in Scanning Probe Microscopies*. Wiley, Chichester, 1998; B. P. Jena and J. K. H. Hörber, *Atomic Force Microscopy in Cell Biology: Methods in Cell Biology*, vol. 68, Academic Press, San Diego, 2002; V. J. Morris, A. R. Kirby, and A. P. Gunning, *Atomic Force Microscopy for Biologists*, Imperial College Press, London, 1999; Z. Shao et al., Biological atomic force microscopy: What is achieved and what is needed, *Adv. Phys.*, 45:1–86, 1996.

Atomic mass

The mass of an atom or molecule on a scale where the mass of a carbon-12 (^{12}C) atom is exactly 12.0 atomic mass units. The mass of any atom is approximately equal to the total number of its protons and neutrons multiplied by the atomic mass unit, $u = 1.6605387 \times 10^{-24}$ gram. (Electrons are much lighter, about 0.0005486 u.) No atom differs from this simple formula by more than 1%, and stable atoms heavier than helium all lie within 0.3%. *See* ATOMIC MASS UNIT.

This simplicity of nature led to the confirmation of the atomic hypothesis—the idea that all matter is composed of atoms, which are identical and chemically indivisible for each chemical element. In 1802, G. E. Fischer noticed that the weights of acids needed to neutralize various bases could be described systematically by assigning relative weights to each of the acids and bases. A few years later, John Dalton proposed an atomic theory in which elements were made up of atoms that combine in simple ways to form molecules.

In reality, nature is more complicated, and the great regularity of atomic masses more revealing. Two fundamental ideas about atomic structure come out of this regularity: that the atomic nucleus is composed of charged protons and uncharged neutrons, and that these particles have approximately equal mass. The number of protons in an atom is called its atomic number, and equals the number of electrons in the neutral atom. The electrons, in turn, determine the chemical properties of the atom. Adding a neutron or two does not change the chemistry (or the name) of an atom, but does give it an atomic mass which is 1 u larger for each added neutron. Such atoms are called isotopes of the element, and their existence was first revealed by careful study of radio-active elements. Most naturally occurring elements are mixtures of isotopes, although a single isotope frequently predominates. Since the proportion of the various isotopes is usually about the same everywhere on Earth, an average atomic mass of an element can be defined, and is called the atomic weight. Atomic weights are routinely used in chemistry in order to determine how much of one chemical will react with a given weight of another. *See* ATOMIC STRUCTURE AND SPECTRA; ISOTOPE; RELATIVE ATOMIC MASS.

In contrast to atomic weights, which can be defined only approximately, atomic masses are exact constants of nature. All atoms of a given isotope are truly identical; they cannot be distinguished by any method. This is known to be true because the quantum mechanics treats identical objects in special ways, and makes predictions that depend on this assumption. One such prediction, the exclusion principle, is the reason that the chemical behavior of atoms with different numbers of electrons is so different. Other consequences, such as the absence of certain features of the optical spectra of molecular gases, and the existence of Bose-Einstein condensates, have also been observed, confirming the truly indistinguishable nature of different atoms of the same isotope. *See* EXCLUSION PRINCIPLE; MOLECULAR STRUCTURE AND SPECTRA; QUANTUM MECHANICS; QUANTUM STATISTICS.

Measurement of atomic masses. In 1913, J. J. Thomson made the first true atomic mass measurements by deflecting a beam of neon ions (charged atoms) with a magnetic field. His resolution was just sufficient to separate neon-20 from neon-22. However, his basic method, measuring the curvature of a beam of ions (charged atoms) in a magnetic field, formed the basis of mass spectrometry for many years. F. W. Aston made great improvements to ion beam, mass spectrometers during the 1920s and 1930s. The common isotopes were identified, and improvements in resolution brought mass accuracies below a thousandth of a mass unit. This accuracy made it possible to verify the equivalence of mass and energy, expressed in the formula $E = mc^2$, where c is the speed of light. In this case, m is the mass difference between reactants and products, and E is the energy liberated in the associated nuclear reaction. For example, when two atoms of deuterium (mass 2.0141 u) fuse to form a helium atom (mass 4.0026 u), 0.0256 u is converted to energy. When multiplied by c^2, this is a tremendous amount of energy, several million times the chemical energy that would be released if the deuterium

reacted with oxygen to form heavy water. Thus a kilogram of nuclear fuel can liberate as much energy as several hundred tons of coal. Atomic mass measurements now provide measurements of nuclear reaction energies far more accurately than they can be predicted from nuclear theory. *See* MASS SPECTROSCOPE; NUCLEAR FUSION; NUCLEAR REACTION.

Single-ion mass measurements. Starting about 1990, a new method of atomic mass measurement was developed with accuracies of about 1 part in 10^{10}, a hundred or more times more accurate than had been possible before. The new technique involves isolating a single ion in a Penning trap, and measuring the frequency at which it orbits in a steady magnetic field produced by a superconducting magnet. The ion is set in motion in a tiny orbit, and its motion is detected with sensitive superconducting electronics. This cyclotron frequency is given by qB/m, where q is the charge of the ion, B is the magnetic field strength, and m is the ion mass. Comparing the cyclotron frequency of the unknown ion with the frequency of a carbon-containing atom gives its mass in u, since 1 u is defined as 1/12 the mass of a carbon-12 atom. *See* PARTICLE ACCELERATOR; PARTICLE TRAP.

The tremendous improvement in accuracy brought on by single-ion measurements not only has improved traditional applications of mass spectrometry such as the calibration of the gamma-ray spectrum, but also has opened the door to new applications in fundamental physics and metrology including evaluation of fundamental constants and setting a lower limit on the neutrino mass, and may ultimately lead to the replacement of the kilogram mass standard. *See* FUNDAMENTAL CONSTANTS; NEUTRINO.

Determination of α. Perhaps the most important dimensionless number in physics is the fine structure constant, $\alpha = 1/137.036$, which underlies much of atomic, nuclear, and elementary particle physics. In addition to determining the spacing of fine structure lines in atomic spectra, α sets the strength of various interactions, including the electromagnetic, and appears in numerous other important ratios. The square of the fine structure constant equals (with several small corrections) twice the ratio of the highest frequency of ultraviolet light emitted by a hydrogen atom, called the Rydberg, to the frequency of a photon whose energy equals the rest energy of an electron, mc^2. The hydrogen ultraviolet frequency has been measured to about 1 part in 10^{11}. The frequency associated with the electron mass is determined indirectly from Penning trap measurements of its atomic mass. The ratio of an atomic mass to its corresponding frequency equals $1000/c^2$ times the molar Planck constant, $N_A h$, where N_A is Avogadro's number and h is Planck's constant. Several measurements of $N_A h$ have been under taken, all requiring accurate atomic mass measurements. Together these may provide the best determination of α. *See* ELECTRON SPIN; FINE STRUCTURE (SPECTRAL LINES); FUNDAMENTAL INTERACTIONS; RYDBERG CONSTANT.

Redefinition of kilogram. The mass of an atom can now be measured much more accurately than the mass of the platinum-iridium cylinder that defines the kilogram. This artifact has been placed in a vault at the International Bureau of Weights and Measures in Sèvres, France. Adsorbed gases and damage from cleaning limit the accuracy of this standard to 1 part in 10^9 or worse. A difficult compromise is made between using the standard and preserving it, with the result that it was compared with other standards only three times in the twentieth century. These shortcomings make it very attractive to define the kilogram in terms of atomic masses, since atoms are widely available and all identical. This would be done by defining the Avogadro number N_A to be exactly $6.02214199 \times 10^{23}$, for example. This number of carbon atoms would then have a mass of 12 grams by definition. Such a definition can be useful only if standard masses can be made in which the number of atoms is known. This can most likely be accomplished using silicon, since the technology exists for making large ultrapure crystals of silicon and since the atomic mass of silicon-28 has been measured accurately. The spacing between atoms in silicon can be accurately measured using x-ray channeling and laser beams of known wavelength. An atomic definition of the kilogram would complete the process of making all the fundamental metrological standards (length, time, voltage, resistance, and so forth) depend on reproducible natural constants rather than the preservation and stability of any particular artifact. *See* AVOGADRO NUMBER; PHYSICAL MEASUREMENT.

Fred L. Palmer; David E. Pritchard

Bibliography. G. Audi et al., The Nubase evaluation of nuclear and decay properties, *Nucl. Phys.*, A624:1–122, 1997; F. DiFilippo et al., Accurate masses for fundamental metrology, *Phys. Rev. Lett.*, 73:1481–1484, 1994.

Atomic mass unit

A unit of mass equal to exactly $^1/_{12}$ of the mass of an atom of carbon-12 (^{12}C), the predominant isotope of carbon. The unit, also known as the dalton, is often abbreviated amu, and is designated by the symbol u. The relative atomic mass of a chemical element is the average mass of its atoms expressed in atomic mass units. *See* RELATIVE ATOMIC MASS.

Before 1961, two versions of the atomic mass unit were in use. The unit used by physicists was defined as $^1/_{16}$ of the mass of an atom of oxygen-16 (^{16}O), the predominant isotope of oxygen. The unit used by chemists was defined as $^1/_{16}$ of the average mass of the atoms in naturally occurring oxygen, a mixture of the isotopes ^{16}O, ^{17}O, and ^{18}O. In 1961, by international agreement, the standard based on ^{12}C superseded both these older units. It is related to them by: 1 amu (international) $\cong$ 1.000318 amu (physical) $\cong$ 1.000043 amu (chemical). *See* ATOMIC MASS.

Jonathan F. Weil

Atomic nucleus

The central region of an atom. Atoms are composed of negatively charged electrons, positively charged protons, and electrically neutral neutrons. The protons and neutrons (collectively known as nucleons) are located in a small central region known as the nucleus. The electrons move in orbits which are large in comparison with the dimensions of the nucleus itself. Protons and neutrons possess approximately equal masses, each roughly 1840 times that of an electron. The number of nucleons in a nucleus is given by the mass number A and the number of protons by the atomic number Z. Nuclear radii r are given approximately by $r = 1.2 \times 10^{-15}$ m $A^{1/3}$. *See* NUCLEAR STRUCTURE. Henry E. Duckworth

Atomic number

The number of elementary positive charges (protons) contained within the nucleus of an atom. It is denoted by the letter Z. Correspondingly, it is also the number of planetary electrons in the neutral atom.

Relation to x-ray wavelengths. The concept of atomic number emerged from the work of G. Moseley, done in 1913–1914. Using a crystal of potassium ferrocyanide in the Bragg x-ray spectrometer, he measured the wavelengths of the most energetic rays (K and L lines) produced by using the elements calcium to zinc as targets in an x-ray tube. The square root of the frequency, $\sqrt{\nu}$, of these x-rays increased by a constant amount in passing from one target to the next. These data, when extended, gave a linear plot of atomic number versus $\sqrt{\nu}$ for all elements studied, using 13 as the atomic number for aluminum and 79 for that of gold. Further, when extrapolated, this plot predicted the atomic number of hydrogen to be 1 and that of uranium to be 92. *See* X-RAY SPECTROMETRY.

Moseley's atomic numbers were quickly recognized as providing an accurate sequence of the elements, which the chemical atomic weights had sometimes failed to do. For example, the elemental pairs argon-potassium, cobalt-nickel, tellurium-iron, which had appeared in the periodic table in the reverse order of their atomic weights, were shown in their correct order by their atomic numbers. Additionally, the atomic number sequence indicated the positions of elements that had not yet been discovered. Two examples were $Z = 43$ (technetium) and $Z = 61$ (promethium); both elements are unstable and were not correctly identified until they had been produced in nuclear reactions.

Atomic and nuclear properties. The atomic number not only identifies the chemical properties of an element but facilitates the description of other aspects of atoms and nuclei. Thus, atoms with the same atomic number are isotopes and belong to the same element, while nuclear reactions may alter the atomic number: in alpha decay, $Z \rightarrow Z - 2$; in beta emission, $Z \rightarrow Z + 1$; and in positron emission or K capture, $Z \rightarrow Z - 1$. *See* ISOTOPE; RADIOACTIVITY.

When specifically written, the atomic number is placed as a subscript preceding the symbol of the element, while the mass number (A) precedes as a superscript, for example, ${}^{27}_{13}\mathrm{Al}$, ${}^{238}_{92}\mathrm{U}$. *See* ATOMIC STRUCTURE AND SPECTRA; ELEMENT (CHEMISTRY); MASS NUMBER. Henry E. Duckworth

Atomic physics

The study of the structure of the atom, its dynamical properties, including energy states, and its interactions with particles and fields. These are almost completely determined by the laws of quantum mechanics, with very refined corrections required by quantum electrodynamics. Despite the enormous complexity of most atomic systems, in which each electron interacts with both the nucleus and all the other orbiting electrons, the wavelike nature of particles, combined with the Pauli exclusion principle, results in an amazingly orderly array of atomic properties. These are systematized by the Mendeleev periodic table. In addition to their classification by chemical activity and atomic weight, the various elements of this table are characterized by a wide variety of observable properties. These include electron affinity, polarizability, angular momentum, multiple electric moments, and magnetism. *See* ATOMIC MASS; PERIODIC TABLE; QUANTUM ELECTRODYNAMICS; QUANTUM MECHANICS.

Each atomic element, normally found in its ground state (that is, with its electron configuration corresponding to the lowest state of total energy), can also exist in an infinite number of excited states. These are also ordered in accordance with relatively simple hierarchies determined by the laws of quantum mechanics. The most characteristic signature of these various excited states is the radiation emitted or absorbed when the atom undergoes a transition from one state to another. The systemization and classification of atomic energy levels (spectroscopy) has played a central role in developing an understanding of atomic structure. Control of populations of ensembles of excited atoms has led to the laser, which is itself now used to obtain even more refined information concerning atomic structure than has hitherto been possible. *See* LASER; SPECTROSCOPY.

Atomic radiation represents one of nature's most refined probes for the study of a wide range of natural phenomena, such as the effects of fluctuations in empty space, of the anomalous magnetic moment of the electron, and even subtler high-order corrections to atomic energy levels caused by quantum electrodynamics. The isolated atom is one of the most reliable systems known in nature in terms of the reproducibility of its energy levels, and is used as a primary frequency standard for time calibration. *See* ATOMIC CLOCK; FUNDAMENTAL CONSTANTS; PARITY (QUANTUM MECHANICS).

The problem of the mutual interaction of atoms or fragments (for example, electrons and ions) is still

more complicated than that of the isolated atom, but such interactions are important since they govern a wide variety of practical phenomena, including atmospheric physics, laser physics, plasma generation for controlled thermonuclear plasmas, materials research (including the influence of radiation on matter), and chemical reactions and molecular formation. This area is the domain of atomic collisions. *See* SCATTERING EXPERIMENTS (ATOMS AND MOLECULES).

Finally, the study of inner shell structure and interactions of heavier atoms at very high energies has an active field of research. This is the domain of high-energy atomic physics which has led to a better understanding of the structure and the dynamics of the inner, strongly bound atomic electrons. *See* ATOM; ATOMIC STRUCTURE AND SPECTRA; ELECTRON; NUCLEAR PHYSICS.

Benjamin Bederson

Atomic spectrometry

A branch of chemical analysis that seeks to determine the composition of a sample in terms of which chemical elements are present and their quantities or concentrations. Unlike other methods of elemental analysis, however, the sample is decomposed into its constituent atoms which are then probed spectroscopically.

In the ultimate atomic spectrometric procedure, a sample would be completely disintegrated into its constituent atoms. The atoms would then be sorted element by element and counted one by one. Advances in physics and chemistry have indicated that this atom-counting approach is feasible; however, the required equipment is complex and costly. Therefore, routine atomic spectrometric measurements are simpler but less elegant or sensitive.

In routine atomic spectrometry, a device called the atom source or atom cell is responsible for producing atoms from the sample; there are many different kinds of atom sources. After atomization of the sample, any of several techniques can determine which atoms are present and in what amounts, but the most common are atomic absorption, atomic emission, atomic fluorescence (the least used of these four alternatives), and mass spectrometry.

Most atomic spectrometric measurements (all those just mentioned except mass spectrometry) exploit the narrow-line spectra characteristic of gas-phase atoms. Because the atom source yields atomic species in the vapor phase, chemical bonds are disrupted, so valence electronic transitions are unperturbed by bonding effects. As a result, transitions among atomic energy levels yield narrow spectral lines, with spectral bandwidths commonly in the 1–5-picometer wavelength range. Moreover, because each atom possesses its unique set of energy levels, these narrow-band transitions can be measured individually, with little mutual interference. Thus, sodium, potassium, and scandium can all be monitored simultaneously and with minimal spectral influence on each other. This lack of spectral overlap remains one of the most attractive features of atomic spectrometry. *See* LINE SPECTRUM; SPECTRUM.

Atomic absorption spectrometry (AAS). In atomic absorption spectrometry, light from a primary source is directed through the atom cell, where a fraction of the light is absorbed by atoms from the sample. The amount of radiation that remains can then be monitored on the far side of the cell. The concentration of atoms in the path of the light beam can be determined by Beer's law, which can be expressed as the equation below, where P_0 is the light intensity

$$\log \frac{P_0}{P} = kC$$

incident on the atom cell, P is the amount of light which remains unabsorbed, C is the concentration of atoms in the cell, and k is the calibration constant, which is determined by means standard samples having known concentrations. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION.

The two most common kinds of atom cells employed in atomic absorption spectrometry are chemical flames and electrical furnaces. Chemical flames are usually simple to use, but furnaces offer higher sensitivity. The simplicity of the flame-based system results from the convenient means by which samples can be introduced. Ordinarily, a sample solution (prepared by dissolving the sample in suitable acids, solvents, or mixtures) is merely sprayed into the flame by means of a pneumatic nebulizer. To ensure that the greatest possible fraction of sample material reaching the flame is actually atomized, a spray chamber is interposed between the sprayer (nebulizer) and flame. In the spray chamber, large droplets in the aerosol settle out or are shattered, so only the finest mist reaches the flame. In the flame, the aerosol droplets are evaporated, the resulting solute (sample) particles are dried, and the dried particles are vaporized.

Unfortunately, even with this procedure the sample particles that reach the flame are seldom atomized completely. Moreover, the presence of particular sample concomitants can influence the degree of atomization of others. Such incomplete atomization often leads to effects known as matrix interferences. To overcome such interferences, so-called releasing agents are usually added to sample and standard (known-concentration) solutions so that the degree of atomization of both is as similar as possible.

Electrically heated carbon furnaces can reach temperatures comparable to those of chemical flames (2800 K or 4580°F). As a result, the furnaces also are effective but not 100% efficient at sample atomization. Thus matrix interferences are also common in furnace-based atomic absorption spectrometry.

The principal attraction of furnaces is that, unlike flames, they can generate sample atoms in a relatively still environment. Consequently, the atom cloud from even small quantities of a sample can be very concentrated, and sensitivity is very high for atomic absorption spectrometry. While the flame-based technique is capable of measuring sample concentrations down to 100 nanograms/ml (parts per

billion), furnace atomic absorption spectrometry offers detection limits on the order of 0.1 ng/ml, a thousandfold gain.

Another attribute of furnace atomization is the ability to handle microliter quantities of sample solution. While flames must usually be fed continuously with sample solutions at flow rates of 1–5 ml/min, furnaces operate satisfactorily with total sample volumes of 10–50 microliters. Because the total volume of required sample solution is small, furnaces can measure extremely low quantities of sample mass. Typical detection limits for many elements by furnace atomic absorption spectrometry are 10^{-13} g.

The most common primary light source employed in atomic absorption spectrometry is the hollow-cathode lamp. Conveniently, the hollow-cathode lamp emits an extremely narrow line spectrum of one, two, or three elements of interest. As a result, the atomic absorption spectrometry measurement is automatically tuned to the particular spectral lines of interest. This attribute simplifies the instrumentation and measurement procedure substantially. However, because an effective hollow-cathode lamp can be made to emit the narrow-line spectra of only one to three chemical elements, determining additional elements requires several lamps. Although several schemes for the automatic switching of lamps have been proposed and are commercially available, the inconvenience of lamp exchanging has kept atomic absorption spectrometry from becoming a truly simultaneous multielement technique. Instead, the method is usually employed for quantitative analysis of specific elements in a sample.

Although spectral interferences among elements are rare in atomic absorption spectrometry, a problem known as nonspecific absorption is fairly common. Regardless of whether a sample is atomized in a flame or furnace, it might generate molecular fragments or smoke. Understandably, furnaces cause greater problems in this regard than do flames. These molecular fragments and smoke can attenuate the light from the hollow-cathode lamp in a manner which is indistinguishable from atomic absorption. Consequently, the presence of nonspecific absorption artificially inflates the apparent concentration of the element in a sample. To overcome this problem, a number of background-correction methods are available on commercial instruments. They all operate on the principle that light scattering by smoke and absorption by molecular fragments both occur over a relatively broad spectral range, compared with atomic absorption. Comparing the broad-band absorption of an atomized sample with the narrow-band absorption of the light from a hollow-cathode lamp permits the contribution from atomic absorption spectrometry to be unraveled from that of nonspecific absorption.

Atomic emission spectrometry (AES). In this type of spectrometry, atomic species are measured by their emission spectra. For such spectra to be produced, the atoms must first be excited by thermal or nonthermal means. Therefore, the atom sources employed in atomic emission spectrometry are hotter or more energetic than those commonly used in atomic absorption spectrometry. Although several such sources are in common use, the dominant one is the inductively coupled plasma. From the simplest standpoint, the inductively coupled plasma is a flowing stream of hot, partially ionized (positively charged) argon. Power is coupled into the plasma by means of an induction coil. At a typical operating frequency of 27–40 MHz and an input power of 1–1.5 kW, the inductively coupled plasma achieves temperatures between 6000 and 8000 K (10,000 and 14,000°F). Such an elevated temperature dissociates sample components even more effectively than in a chemical flame, so the inductively coupled plasma does not suffer the serious matrix effects that are commonly encountered in flame atomic absorption spectrometry. Furthermore, atoms produced in the inductively coupled plasma are efficiently excited at the plasma's high temperature. Detection limits are accordingly low and are on the order of 0.1–1 ng/ml for most elements. Importantly, emission occurs from all elements simultaneously in the inductively coupled plasma, so that either rapid sequential or truly simultaneous multielement analysis is possible.

There are two common modes for observing emission spectra from an inductively coupled plasma. The less expensive and more flexible approach employs a so-called slew-scan spectrometer. With this instrument, spectral lines are accessed in rapid sequence, so that a number of chemical elements can be measured rapidly, one after the other. Moreover, because each viewed elemental spectral line can be scanned completely, it is possible to subtract spectral emission background independently for each element. Spectral interferences are therefore reduced somewhat.

The alternative approach is to view all spectral lines simultaneously, either with a number of individual photodetectors keyed to particular spectral lines or with a truly multichannel electronic detector driven by a computer. Although this alternative approach is somewhat less flexible than the slew-scan system, it enables samples to be analyzed more rapidly and permits transient atom signals (as from a furnace-based atomizer) to be recorded.

Because the inductively coupled plasma produces fewer severe interelement interferences, exhibits better detection limits, and permits analysis of a greater number of samples per unit time, it has assumed much of the workload that was formerly assigned to atomic absorption spectrometry. Nonetheless, inductively coupled plasma atomic emission is not without its shortcomings. For example, emission spectra from individual elements in the inductively coupled plasma are far more complex than are atomic absorption spectra. Also, the spectral background from the inductively coupled plasma consists not only of emission lines of argon (the plasma support gas) but also of a continuum (extremely broad-band) background in addition to emission bands from molecular fragments introduced with the sample or by entrainment of atmospheric gases. Therefore, spectral interferences in inductively coupled

plasma atomic emission are far more serious than are those in atomic absorption spectrometry.

Commercial instrument manufacturers have sought to overcome as many of these difficulties as possible. Most instruments using inductively coupled plasmas are computer-controlled, and many include programs with menus that permit the judicious selection of those spectral lines that are most likely to be free of spectral interference. Also, most instruments are equipped with spectral scanning features, whether they be of the slew-scan or direct-reading type. These features permit the background beneath each spectral line to be measured and the net atomic signal deduced.

Another limitation of inductively coupled plasma atomic emission measurements is sensitivity. Although the technique is capable of measuring lower concentrations than flame atomic absorption spectrometry, it does not yet possess the sensitivity that is achieved with furnace atomic absorption spectrometry. Moreover, except for specialized designs, most inductively coupled plasma units are not capable of measuring microliter quantities of sample solution. Finally, most instruments using inductively coupled plasmas are still somewhat more complex to operate than comparably priced units designed for atomic absorption spectromerty. *See* EMISSION SPECTROCHEMICAL ANALYSIS.

Mass spectrometry (MS). Elemental mass spectrometry has been practiced for many years in the form of spark-source mass spectrometry and, more recently, glow-discharge-lamp mass spectrometry. However, a hybrid technique that combines the inductively coupled plasma with a mass spectrometer (ICP-MS) has assumed a prominent place.

At the high temperatures present in an inductively coupled plasma, many atomic species occur in an ionic form. These ions can be readily extracted into a mass spectrometer via an interface that separates the atmospheric-pressure plasma and the mass spectrometer. In a typical configuration, a tapered, water-cooled metallic cone is inserted into the inductively coupled plasma. A small orifice (typically 0.04 in. or 1 mm in diameter) in the cone then allows the passage of ions from the plasma into a reduced-pressure zone held ordinarily at pressures near 1 torr (133 pascals). The ion beam generated in this differentially pumped region is then passed to one or more successively lower-pressure zones until it is extracted into a mass spectrometer, where ions are separated according to their mass. The mass spectrometer region is usually held at pressures near 10^{-5} torr (1.33×10^{-3} Pa).

In commercial instrumentation combining an inductively coupled plasma and mass spectrometry, a quadrupole mass spectrometer is employed. In this type of system, ions of different elements are separated sequentially according to mass. The separated ions then pass one after the other into a detector where they can be measured electronically. Differences in atomic mass make it possible to identify elements; the intensity of their signals then allows determination of their concentrations.

The advantages of the combination of inductively coupled plasma and mass spectrometry are substantial. The system is capable of some of the best detection limits in atomic spectrometry, typically 10^{-3} to 10^{-2} ng/ml for most elements. Also, virtually all elements in the periodic table can be determined during a single scan. The method is also capable of providing isotopic information, unavailable by any other atomic spectrometric method for such a broad range of elements. Finally, elemental mass spectra are relatively simple; there are very few naturally occurring isotopes for each element and also virtually no mono-isotopic elements that suffer severe spectral overlap with isotopes of other elements.

Use of this hybrid technique does present some disadvantages. A substantial number of diatomic and cluster ions are often observed in spectra produced by the combination of inductively coupled plasma and mass spectrometry. They originate from a combination of sample-borne species, from argon, from atmospheric gases, and from solvent vapor. These polyatomic ions often overlap with masses of desired elements. It has been found that the incidence of interfering ions can be reduced through use of aqueous or nitric acid solutions, but some problems remain.

From a practical standpoint, widespread use of the inductively coupled plasma combined with mass spectrometry is limited largely by the cost of instrumentation, which is between two and five times that of competing atomic absorption spectrometry and inductively coupled plasma atomic emission instruments. Also, few standard procedures have been developed, and a rather intensive effort would be needed for the development of methods for the determination of most new samples. Nonetheless, the extremely high sensitivity that is provided by the combination of inductively coupled plasma and mass spectrometry with its isotopic selectivity and relatively easily interpreted spectra suggests that it is likely to achieve increased prominence in the future. *See* MASS SPECTROMETRY.

Gary M. Hieftje

Bibliography. P. W. M. J. Boumans (ed.), *Inductively Coupled Plasma Emission Spectroscopy*, pt. I: *Methodology, Instrumentation, and Performance*, pt. II: *Applications and Fundamentals*, 1987; J. D. Ingle and S. R. Crouch, *Spectrochemical Analysis*, 1988; A. Montaser and D. W. Golightly, *Inductively Coupled Plasmas in Analytical Atomic Spectrometry*, 2d ed., 1992; D. G. Peters, J. M. Hayes, and G. M. Hieftje, *Chemical Separations and Measurements*, 1974; W. J. Price, *Spectrochemical Analysis by Atomic Absorption*, 1985.

Atomic structure and spectra

The idea that matter is subdivided into discrete building blocks called atoms, which are not divisible any further, dates back to the Greek philosopher Democritus. His teachings of the fifth century B.C. are commonly accepted as the earliest authenticated ones concerning what has come to be called

atomism by students of Greek philosophy. The weaving of the philosophical thread of atomism into the analytical fabric of physics began in the late eighteenth and the nineteenth centuries. Robert Boyle is generally credited with introducing the concept of chemical elements, the irreducible units which are now recognized as individual atoms of a given element. In the early nineteenth century John Dalton developed his atomic theory, which postulated that matter consists of indivisible atoms as the irreducible units of Boyle's elements, that each atom of a given element has identical attributes, that differences among elements are due to fundamental differences among their constituent atoms, that chemical reactions proceed by simple rearrangement of indestructible atoms, and that chemical compounds consist of molecules which are reasonably stable aggregates of such indestructible atoms. *See* CHEMISTRY.

Electromagnetic nature of atoms. The work of J. J. Thomson in 1897 clearly demonstrated that atoms are electromagnetically constituted and that from them can be extracted fundamental material units bearing electric charge that are now called electrons. These pointlike charges have a mass of 9.109×10^{-31} kg and a negative electric charge of magnitude 1.602×10^{-19} coulomb. The electrons of an atom account for a negligible fraction of its mass. By virtue of overall electrical neutrality of every atom, the mass must therefore reside in a compensating, positively charged atomic component of equal charge magnitude but vastly greater mass. *See* ELECTRON.

Thomson's work was followed by the demonstration by Ernest Rutherford in 1911 that nearly all the mass and all of the positive electric charge of an atom are concentrated in a small nuclear core approximately 10,000 times smaller in extent than an atomic diameter. Niels Bohr in 1913 and others carried out some remarkably successful attempts to build solar system models of atoms containing planetary pointlike electrons orbiting around a positive core through mutual electrical attraction (though only certain "quantized" orbits were "permitted"). These models were ultimately superseded by nonparticulate, matter-wave quantum theories of both electrons and atomic nuclei. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

The modern picture of condensed matter (such as solid crystals) consists of an aggregate of atoms or molecules which respond to each other's proximity through attractive electrical interactions at separation distances of the order of 1 atomic diameter (approximately 10^{-10} m) and repulsive electrical interactions at much smaller distances. These interactions are mediated by the electrons, which are in some sense shared and exchanged by all atoms of a particular sample, and serve as an interatomic glue that binds the mutually repulsive, heavy, positively charged atomic cores together. *See* SOLID-STATE PHYSICS.

Planetary atomic models. Fundamental to any planetary atomic model is a description of the forces which give rise to the attraction and repulsion of the constituents of an atom. Coulomb's law describes the fundamental interaction which, to a good approximation, is still the basis of modern theories of atomic structure and spectra: the force exerted by one charge on another is repulsive for charges of like sign and attractive for charges of unlike sign, is proportional to the magnitude of each electric charge, and diminishes as the square of the distance separating the charges. *See* COULOMB'S LAW.

Also fundamental to any planetary model is Thomson's discovery that electrically neutral atoms in some sense contain individual electrons whose charge is restricted to a unique quantized value. Moreover, Thomson's work suggested that nuclear charges are precisely equally but oppositely quantized to offset the sum of the constituent atomic electron charges. Boyle's elements, of which over a hundred have been discovered, may then be individually labeled by the number of quantized positive-charge units Z (the atomic number) residing within the atomic nucleus (unit charge $= +1.602 \times 10^{-19}$ coulomb). Each nucleus is surrounded by a complement of Z electrons (with charges of -1.602×10^{-19} C each) to produce overall charge neutrality. Molecules containing two, three, ..., atoms can then be thought of as binary, ternary, ..., planetary systems consisting of heavy central bodies of atomic numbers Z_1, Z_2, ..., sharing a supply of $Z_1 + Z_2 + \cdots$ electrons having the freedom to circulate throughout the aggregate and bond it together as would an electronic glue.

Atomic sizes. All atoms, whatever their atomic number Z, have roughly the same diameter (about 10^{-10} m), and molecular sizes as well as macroscopic solid sample sizes tend to be the sum of the aggregate atomic sizes. It is easy to qualitatively, though only partially, account for this circumstance by using Coulomb's law. The innermost electrons of atoms orbit at small radii because of intense electrical attractions prevailing at small distances. Because electrons are more than 2000 times lighter than most nuclei and therefore move less sluggishly, the rapid orbital motion of inner electrons tends to clothe the slow-moving positive nuclear charge in a negative-charge cloud, which viewed from outside the cloud masks this positive nuclear charge to an ever-increasing extent. Thus, as intermediate and outer electrons are added, each experiences a diminishing attraction. The Zth, and last, electron sees $+Z$ electronic charge units within the nucleus compensated by about $(Z - 1)$ electronic charges in a surrounding cloud, so that in first approximation the Zth electron orbits as it would about a bare proton having $Z = 1$, at about the same radius. Crudely speaking, an atom of any atomic number Z thus has a size similar to that of the hydrogen atom. When the molecular aggregates of atoms of concern to chemistry are formed, they generally do so through extensive sharing or exchanging of a small number (often just one) of the outermost electrons of each atom. Hence, the interatomic spacing in both molecules and solids tends to be of the order of one to a very few atomic diameters.

Chemical reactions. In a microscopic planetary atomic model, chemical reactions in which rearrangements result in new molecular aggregates of atoms can be viewed as electrical phenomena, mediated by the electrical interactions of outer electrons. The reactions proceed by virtue of changing positions and shapes of the orbits of the binding electrons, either through their internal electrical interactions or through their being electrically "bumped" from outside (as by collisions with nearby molecules).

Difficulties with the models. Before continuing with a detailed account of the successes of detailed planetary atomic models, it is advisable to anticipate some severe difficulties associated with them. Some of these difficulties played a prominent role in the development of quantum theory, and others present as yet unsolved and profound challenges to classical as well as modern quantum theories of physics.

A classical planetary theory fails to account for several atomic enigmas. First, it is known that electrons can be made to execute fast circular orbits in large accelerating machines, of which modern synchrotrons are excellent examples. As they are then in accelerated motion, they radiate light energy, as predicted by Maxwell's theory, at a frequency equal to that of their orbital circular motions. Thus, a classical planetary theory fails to explain how atoms can be stable and why atomic electrons do not continuously radiate light, falling into the nucleus as they do so. Such a theory also does not account for the observation that light emitted by atomic electrons appears to contain only quantized frequencies or wavelengths. Furthermore, a classical planetary theory would lead one to expect that all atomic electrons of an atom would orbit very close to their parent nuclei at very small radii, instead of distributing themselves in shells of increasing orbital radius and thickness, which seem able to accommodate only small, fixed numbers of electrons per shell.

Any theory of atomic structure must deal with the question of whether the atom is mostly empty space, as a planetary system picture suggests, or whether the entire atomic volume is filled with electronic charges in smeared out, cloudlike fashion. Such a theory must also be concerned with whether electrons and nuclei are pointlike, structureless, and indivisible, or whether they can be subdivided further into smaller constituents, in analogy to the way in which Thomson was able to extract electrons as constituents of atoms. Questions which still have not received definitive answers concern how much energy is needed to construct electrons and nuclei, what their radii are, and why they do not fly apart under the explosive repulsion of half their charge distribution for the other half. A more advanced theory than the classical planetary model is also required to determine whether electrons really interact with each other and with nuclei instantaneously, as Coulomb's law would have it, or whether there is a finite interaction delay time. Finally, a fundamental question, which can be answered only by quantum theory, is concerned with whether electrons and nuclei behave as pointlike objects, as light does when it gives rise to sharp shadows, or whether they behave as extended waves which can exhibit diffraction (bending) phenomena, as when light or water waves bend past sharp edges and partially penetrate regions of shadow in undulatory intensity patterns.

Scattering experiments. A key experimental technique in answering many important questions, such as those just posed, is that of scattering, in which small, high-speed probe projectiles are used to interact with localized regions of more extended objects from which they rebound, or scatter. A study of the number of scattered projectiles, and their distributions in speed and angle, gives much information about the structure of target systems, and the internal distributions, speeds, and concentrations of constituent bodies or substructures. *See* SCATTERING EXPERIMENTS (ATOMS AND MOLECULES).

The scattering of x-rays by solids, for example, was used by C. Barkla to indicate that the number of electrons in an atom is approximately half the atomic weight. The mechanism of the production of the secondary (scattered) x-rays is indicated in **Fig. 1***a*. X-rays are electromagnetic waves of wavelength considerably smaller than the size of the atom. If an x-ray sets an atomic electron into vibration, there follows the emission of a wave of lesser amplitude which can be observed at directions outside the incident beam. *See* X-RAYS.

Rutherford's experiments on the scattering of alpha particles represented an important step in understanding atomic structure. Alpha particles are helium ($Z = 2$) nuclei emitted at high velocities by some radioactive materials. A beam of these particles was directed at thin foils of different metals, and the relative numbers scattered at various angles were observed. While most of the particles passed through the foil with small deflections, such as the lower particle in Fig. 1*b*, a considerable number of very large deflections occurred. The upper particle in Fig. 1*b* has undergone such a deflection. The precise results could be explained only if the positive electric charge of the atom is concentrated in the very small volume of the nucleus, which also contains almost all of the atom's mass. The diameter of the nucleus, found to depend on the atomic mass, was about 10^{-14} m for heavy nuclei. The nuclear charge was found to be the atomic number Z times the magnitude of the electron's charge. *See* ALPHA PARTICLES.

The results of the scattering experiments therefore established the model of the atom as consisting of a small, massive, positively charged nucleus surrounded by a cloud of electrons to provide

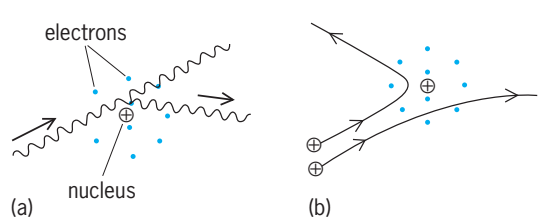

Fig. 1. Scattering by an atom (*a*) of x-rays, (*b*) of alpha particles.

electrical neutrality. As noted above, such an atom should quickly collapse. The first step toward an explanation as to why it did not, came from the Bohr picture of the atom.

Bohr Atom

The hydrogen atom is the simplest atom, and its spectrum (or pattern of light frequencies emitted) is also the simplest. The regularity of its spectrum had defied explanation until Bohr solved it with three postulates, these representing a model which is useful, but quite insufficient, for understanding the atom.

Postulate 1: The force that holds the electron to the nucleus is the Coulomb force between electrically charged bodies.

Postulate 2: Only certain stable, nonradiating orbits for the electron's motion are possible, those for which the angular momentum associated with the motion of an electron in its orbit is an integral multiple of $h/2\pi$ (Bohr's quantum condition on the orbital angular momentum). Each stable orbit represents a discrete energy state.

Postulate 3: Emission or absorption of light occurs when the electron makes a transition from one stable orbit to another, and the frequency ν of the light is such that the difference in the orbital energies equals $h\nu$ (A. Einstein's frequency condition for the photon, the quantum of light). The description of these apparently nearly instantaneous quantum jumps remains a subject of research, as discussed below.

Here the concept of angular momentum, a continuous measure of rotational motion in classical physics, has been asserted to have a discrete quantum behavior, so that its quantized size is related to Planck's constant h, a universal constant of nature. The orbital angular momentum of a point object of mass m and velocity v, in rotational motion about a central body, is defined as the product of the component of linear momentum mv (expressing the inertial motion of the body) tangential to the orbit times the distance to the center of rotation. *See* ANGULAR MOMENTUM.

Modern quantum mechanics has provided justification of Bohr's quantum condition on the orbital angular momentum. It has also shown that the concept of definite orbits cannot be retained except in the limiting case of very large orbits. In this limit, the frequency, intensity, and polarization can be accurately calculated by applying the classical laws of electrodynamics to the radiation from the orbiting electron. This fact illustrates Bohr's correspondence principle, according to which the quantum results must agree with the classical ones for large dimensions. The deviation from classical theory that occurs when the orbits are smaller than the limiting case is such that one may no longer picture an accurately defined orbit. Bohr's other hypotheses are still valid.

Quantization of hydrogen atom. According to Bohr's theory, the energies of the hydrogen atom are quantized (that is, can take on only certain discrete values). These energies can be calculated from the electron orbits permitted by the quantized orbital angular momentum. The orbit may be circular or elliptical, so only the circular orbit is considered here for simplicity. Let the electron, of mass m and electric charge $-e$, describe a circular orbit of radius r around a nucleus of charge $+e$ and of infinite mass. With the electron velocity v, the angular momentum is mvr, and the second postulatebecomes Eq. (1).

$$mvr = n\left(\frac{h}{2\pi}\right) \qquad (n = 1, 2, 3, \ldots) \qquad (1)$$

The integer n is called the principal quantum number. The centripetal force required to hold the electron in its orbit is the electrostatic force described by Coulomb's law, as shown in Eq. (2). Here ϵ_0 is

$$\frac{mv^2}{r} = \frac{e^2}{4\pi\epsilon_0 r^2} \qquad (2)$$

the permittivity of free space, a constant included in order to give the correct units to the statement of Coulomb's law in SI units. *See* QUANTUM NUMBERS.

The energy of an electron in an orbit consists of both kinetic and potential energies. For these circular orbits, the potential energy is twice as large as the kinetic energy and has a negative sign, where the potential energy is taken as zero when the electron and nucleus are at rest and separated by a very large distance. The total energy, which is the sum of kinetic and potential energies, is given in Eq. (3).

$$E = \frac{mv^2}{2} - mv^2 = -\frac{mv^2}{2} \qquad (3)$$

The negative sign means that the electron is bound to the nucleus and energy must be provided to separate them.

It is possible to eliminate v and r from these three equations. The result is that the possible energies of the nonradiating states of the atom are given by Eq. (4).

$$E = -\frac{me^4}{8\epsilon_0^2 h^2} \cdot \frac{1}{n^2} \qquad (4)$$

The same equation for the hydrogen atom's energy levels, except for some small but significant corrections, is obtained from the solution of the Schrödinger equation, as modified by W. Pauli, for the hydrogen atom. *See* SCHRÖDINGER'S WAVE EQUATION.

The frequencies of electromagnetic radiation or light emitted or absorbed in transitions are given by Eq. (5) where E' and E'' are the energies of the initial

$$\nu = \frac{E' - E''}{h} \qquad (5)$$

and final states of the atom. Spectroscopists usually express their measurements in wavelength λ or in wave number σ in order to obtain numbers of a convenient size. The frequency of a light wave can be thought of as the number of complete waves radiated per second of elapsed time. If each wave travels at a fixed velocity c (approximately 3×10^8 m/s in vacuum), then after t seconds, the distance traveled by the first wave in a train of waves is ct, the number

of waves in the train is νt, and hence the length of each must be $ct/(\nu t) = c/\nu = \lambda$. The wave number, defined as the reciprocal of the wavelength, therefore equals ν/c. The quantization of energy $h\nu$ and of angular momentum in units of $h/(2\pi)$ does not materially alter this picture. The wave number of a transition is shown in Eq. (6). If $T = -E/(hc)$, then Eq. (7) results. Here T is called the spectral term.

$$\sigma = \frac{\nu}{c} = \frac{E'}{hc} - \frac{E''}{hc} \tag{6}$$

$$\sigma = T'' - T' \tag{7}$$

The allowed terms for hydrogen, from Eq. (4), are given by Eq. (8). The quantity R is the important

$$T = \frac{me^4}{8\epsilon_0^2 ch^3} \cdot \frac{1}{n^2} = \frac{R}{n^2} \tag{8}$$

Rydberg constant. Its value, which has been measured to a remarkable and rapidly improving accuracy, is related to the values of other well-known atomic constants, as in Eq. (8). *See* RYDBERG CONSTANT.

The effect of finite nuclear mass must be considered, since the nucleus does not actually remain at rest at the center of the atom. Instead, the electron and nucleus revolve about their common center of mass. This effect can be accurately accounted for and requires a small change in the value of the effective mass m in Eq. (8). The mass effect was first detected by comparing the spectrum of hydrogen with that of singly ionized helium, which is like hydrogen in having a single electron orbiting the nucleus. For this isoelectronic case, the factor Z^2 must be included in the numerator of Eqs. (2) and (8) to account for the greater nuclear charge. The mass effect was used by H. Urey to discover deuterium, one of three hydrogen isotopes, by the shift of lines in its spectrum because of the very small change in its Rydberg constant. *See* ISOELECTRONIC SEQUENCE; ISOTOPE SHIFT.

Elliptical orbits. In addition to the circular orbits already described, elliptical ones are also consistent with the requirement that the angular momentum be quantized. A. Sommerfeld showed that for each value of n there is a family of n permitted elliptical orbits, all having the same major axis but with different eccentricities. **Figure 2***a* shows, for example, the Bohr-Sommerfeld orbits for $n = 3$. The orbits are labeled s, p, and d, indicating values of the azimuthal quantum number $l = 0$, 1, and 2. This number determines the shape of the orbit, since the ratio of the major to the minor axis is found to be $n/(l = 1)$. To a first approximation, the energies of all orbits of the same n are equal. In the case of the highly eccentric orbits, however, there is a slight lowering of the energy due to precession of the orbit (Fig. 2*b*). According to Einstein's theory of relativity, the mass increases somewhat in the inner part of the orbit, because of greater velocity. The velocity increase is greater as the eccentricity is greater, so the orbits of higher eccentricity have their energies lowered more. The quantity l is called the orbital angular momentum quantum number or the azimuthal quantum number. *See* RELATIVITY.

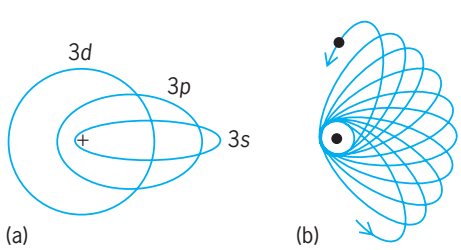

Fig. 2. Possible elliptical orbits, according to the Bohr-Sommerfeld theory. (*a*) The three permitted orbits for $n = 3$. (*b*) Precession of the 3*s* orbit caused by the relativistic variation of mass. (*After A. P. Arya, Fundamentals of Atomic Physics, Allyn and Bacon, 1971*)

A selection rule limits the possible changes of l that give rise to spectrum lines (transitions of fixed frequency or wavelength) of appreciable intensity. The rule is that l may increase or decrease only by one unit. This is usually written as $\Delta l = \pm 1$ for an allowed transition. Transitions for which selection rules are not satisfied are called forbidden; these tend to have quite low intensities. The quantum number n may change by any amount, although the probability of large changes in n is generally low. Selection rules for hydrogen may be derived from Bohr's correspondence principle. However, the selection rules, as well as the relation between n and l, arise much more naturally in quantum mechanics. *See* SELECTION RULES (PHYSICS).

Multielectron Atoms

In attempting to extend Bohr's model to atoms with more than one electron, it is logical to compare the experimentally observed terms of the alkali atoms, which contain only a single electron outside closed shells, with those of hydrogen. A definite similarity is found but with the striking difference that all terms with $l > 0$ are double. This fact was interpreted by S. A. Goudsmit and G. E. Uhlenbeck as due to the presence of an additional angular momentum of $^1/_2(h/2\pi)$ attributed to the electron spinning about its axis. The spin quantum number of the electron is $s = {}^1/_2$.

The relativistic quantum mechanics developed by P. A. M. Dirac provided the theoretical basis for this experimental observation. *See* ELECTRON SPIN.

Exclusion principle. Implicit in much of the following discussion is W. Pauli's exclusion principle, first enunciated in 1925, which when applied to atoms may be stated as follows: no more than one electron in a multielectron atom can possess precisely the same quantum numbers. In an independent, hydrogenic electron approximation to multielectron atoms, there are $2n^2$ possible independent choices of the principal (n), orbital (l), and magnetic (m_l, m_s) quantum numbers available for electrons belonging to a given n, and no more. Here m_l and m_s refer to the quantized projections of l and s along some chosen direction. The organization of atomic electrons into shells of increasing radius (the Bohr radius

scales as n^2) follows from this principle, answering the question as to why all electrons of a heavy atom do not collapse into the most tightly bound orbits. *See* EXCLUSION PRINCIPLE.

Examples are: helium ($Z = 2$), two $n = 1$ electrons; neon ($Z = 10$), two $n = 1$ electrons, eight $n = 2$ electrons; argon ($Z = 18$), two $n = 1$ electrons, eight $n = 2$ electrons, eight $n = 3$ electrons. Actually, in elements of Z greater than 18, the $n = 3$ shell could in principle accommodate 10 more electrons but, for detailed reasons of binding energy economy rather than fundamental symmetry, contains full $3s$ and $3p$ shells for a total of eight $n = 3$ electrons, but often a partially empty $3d$ shell. The most chemically active elements, the alkalies, are those with just one outer orbital electron in an n state, one unit above that of a completely full shell or subshell. *See* ELECTRON CONFIGURATION.

Spin-orbit coupling. This is the name given to the energy of interaction of the electron's spin with its orbital angular momentum. The origin of this energy is magnetic.

A charge in motion through either "pure" electric or "pure" magnetic fields, that is, through fields perceived as "pure" in a static laboratory, actually experiences a combination of electric and magnetic fields, if viewed in the frame of reference of a moving observer with respect to whom the charge is momentarily at rest. For example, moving charges are well known to be deflected by magnetic fields. But in the rest frame of such a charge, there is no motion, and any acceleration of a charge must be due to the presence of a pure electric field from the point of view of an observer analyzing the motion in that reference frame. *See* RELATIVISTIC ELECTRODYNAMICS.

A spinning electron can crudely be pictured as a spinning ball of charge, imitating a circulating electric current (though Dirac electron theory assumes no finite electron radius—classical pictures fail). This circulating current gives rise to a magnetic field distribution very similar to that of a small bar magnet, with north and south magnetic poles symmetrically distributed along the spin axis above and below the spin equator. This representative bar magnet can interact with external magnetic fields, one source of which is the magnetic field experienced by an electron in its rest frame, owing to its orbital motion through the electric field established by the central nucleus of an atom. In multielectron atoms, there can be additional, though generally weaker, interactions arising from the magnetic interactions of each electron with its neighbors, as all are moving with respect to each other and all have spin. The strength of the bar magnet equivalent to each electron spin, and its direction in space are characterized by a quantity called the magnetic moment, which also is quantized essentially because the spin itself is quantized. Studies of the effect of an external magnetic field on the states of atoms show that the magnetic moment associated with the electron spin is equal in magnitude to a unit called the Bohr magneton. *See* MAGNETISM.

The energy of the interaction between the electron's magnetic moment and the magnetic field generated by its orbital motion is usually a small correction to the spectral term, and depends on the angle between the magnetic moment and the magnetic field or, equivalently, between the spin angular momentum vector and the orbital angular momentum vector (a vector perpendicular to the orbital plane whose magnitude is the size of the orbital angular momentum). Since quantum theory requires that the quantum number j of the electron's total angular momentum shall take values differing by integers, while l is always an integer, there are only two possible orientations for s relative to l: s must be either parallel or antiparallel to l. (This statement is convenient but not quite accurate. Actually, orbital angular momentum is a vector quantity represented by the quantum number l and of magnitude $\sqrt{l(l=1)} \cdot (h/2\pi)$. There are similar relations for spin and total angular momentum. These statements all being true simultaneously, the spin vector cannot ever be exactly parallel to the orbital angular momentum vector. Only the quantum numbers themselves can be described as stated.) **Figure 3***a* shows the relative orientations of these two vectors and of their resultant j for a p electron (one for which $l = 1$). The corresponding spectral term designations are shown adjacent to the vector diagrams, labeled with the customary spectroscopic notation, to be explained later.

For the case of a single electron outside the nucleus, the Dirac theory gives Eq. (9) for the spin-

$$\Delta T = \frac{R\alpha^2 Z^4}{n^3} \times \frac{j(j+1) - l(l+1) - s(s+1)}{l(2l+1)(l+1)} \quad (9)$$

orbit correction to the spectral terms. Here $\alpha = e^2/(2\epsilon_0 hc) \cong 1/137$ is called the fine structure constant. The fine structure splitting predicted by Eq. (9)

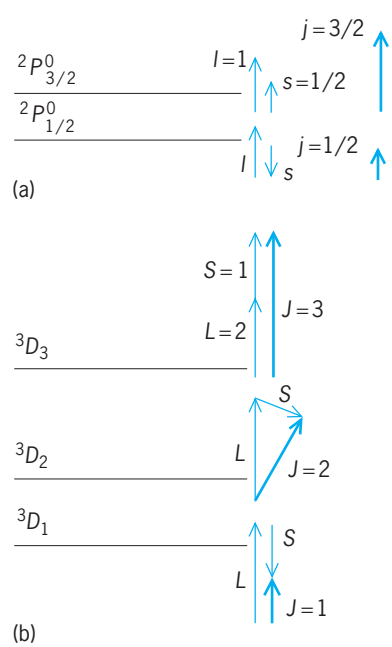

Fig. 3. Vector model for spectral terms arising from (*a*) a single *p* electron, and (*b*) two electrons, either two *p* electrons, or an *s* and a *d* electron.

is present in hydrogen, although its observation requires instruments of very high precision. A relativistic correction must be added.

In atoms having more than one electron, this fine structure becomes what is called the multiplet structure. The doublets in the alkali spectra, for example, are due to spin-orbit coupling; Eq. (9), with suitable modifications, can still be applied. These modifications may be attributed to penetration of the outer orbital electron within the closed shells of other electrons.

The various states of the atom are described by giving the quantum numbers n and l for each electron (the configuration) as well as the set of quantum numbers which depend on the manner in which the electrons interact with each other.

Coupling schemes. When more than one electron is present in the atom, there are various ways in which the spins and orbital angular momenta can interact. Each spin may couple to its own orbit, as in the one-electron case; other possibilities are orbit–other orbit, spin-spin, and so on. The most common interaction in the light atoms, called LS coupling or Russell-Saunders coupling, is described schematically in Eq. (10). This notation indicates that the l_i are

$$\{(l_1, l_2, l_3, \ldots)(s_1, s_2, s_3, \ldots)\} = \{L, S\} = J \qquad (10)$$

coupled strongly together to form a resultant L, representing the total orbital angular momentum. The s_i are coupled strongly together to form a resultant S, the total spin angular momentum. The weakest coupling is that between L and S to form J, the total angular momentum of the electron system of the atom in this state. Suppose, for example, it is necessary to calculate the terms arising from a p ($l = 1$) and a d ($l = 2$) electron. The quantum numbers L, S, and J are never negative, so the only possible values of L are 1, 2, and 3. States with these values of L are designated P, D, and F. This notation, according to which $L = 0, 1, 2, 3, 4, 5$, etc. correspond to S, P, D, F, G, H, etc. terms, survives from the early empirical designation of series of lines as sharp, principal, diffuse, and so on.

The spin $s = {}^1/_2$ for each electron, always, so the total spin $S = 0$ or $S = 1$ in this case. Consider only the term with $S = 1$ and $L = 2$. The coupling of L and S gives $J = 1$, 2, and 3, as shown in Fig. 3*b*. These three values of J correspond to a triplet of different energy levels which lie rather close together, since the LS interaction is relatively weak. It is convenient to use the notation ${}^3D$ for this multiplet of levels, the individual levels being indicated with J values as a subscript, as ${}^3D_1$, ${}^3D_2$, and ${}^3D_3$. The superscript has the value $2S + 1$, a quantity called the multiplicity. There will also be a term with $S = 0$ and $L = 2$ and the single value of $J = 2$. This is a singlet level, written as ${}^1D$ or ${}^1D_2$. The entire group of terms arising from a p and d electron includes ${}^1P$, ${}^3P$, ${}^1D$, ${}^3D$, ${}^1F$, ${}^3F$, with all the values of J possible for each term.

The ${}^2P$ state shown in Fig. 3*a* is derived from a single electron which, since $L = l = 1$ and $S = s = {}^1/_2$, has J values of ${}^1/_2$ and ${}^1/_2$, forming a doublet. If there are three or more electrons, the number of possible terms becomes very large, but they are easily derived by a similar procedure. The resulting multiplicities are shown in the **table**. If two or more electrons are equivalent, that is, have the same n and l, the number of resulting terms is greatly reduced, because of the requirement of the Pauli exclusion principle that no two electrons in an atom may have all their quantum numbers alike. Two equivalent p electrons, for example, give only ${}^1S$, ${}^3P$, and ${}^1D$ terms, instead of the six terms ${}^1S$, ${}^3S$, ${}^1P$, ${}^3P$, ${}^1D$, and ${}^3D$ possible if the electrons are nonequivalent. The exclusion principle applied to equivalent electrons explains the observation of filled shells in atoms, where $L = S = J = 0$.

Possible multiplicities with different numbers of electrons

Number of electrons:	1	2	3	4
Values of S:	${}^1/_2$	1, 0	${}^3/_2$, ${}^1/_2$	2, 1, 0
Multiplicities	Doublets	Singlets Triplets	Doublets Quartets	Singlets Triplets Quintets

Coupling of the LS type is generally applicable to the low-energy states of the lighter atoms. The next commonest type is called jj coupling, represented in Eq. (11). Each electron has its spin coupled to its

$$\{(l_1, s_1)(l_2, s_2)(l_3, s_3)\ldots\} = \{j_1, j_2, j_3, \ldots\} = J \qquad (11)$$

own orbital angular momentum to form a j_i for that electron. The various j_i are then more weakly coupled together to give J. This type of coupling is seldom strictly observed. In the heavier atoms it is common to find a condition intermediate between LS and jj coupling; then either the LS or jj notation may be used to describe the levels, because the number of levels for a given electron configuration is independent of the coupling scheme.

Spectrum of Hydrogen

Figure 4*a* shows the terms R/n^2 in the spectrum of hydrogen resulting from the simple Bohr theory. These terms are obtained by dividing the numerical value of the Rydberg constant for hydrogen (109,678 cm^{-1}) by n^2, that is, by 1, 4, 9, 16, etc. The equivalent energies in electronvolts may then be found by using the conversion factor 1 eV = 8066 cm^{-1}. These energies, in contrast to the term values, are usually measured from zero at the lowest state, and increase for successive levels. They draw closer together until they converge at 13.598 eV. This corresponds to the orbit with $n = \infty$ and complete removal of the electron, that is, ionization of the atom. Above this ionization potential is a continuum of states representing the nucleus plus a free electron possessing a variable amount of kinetic energy.

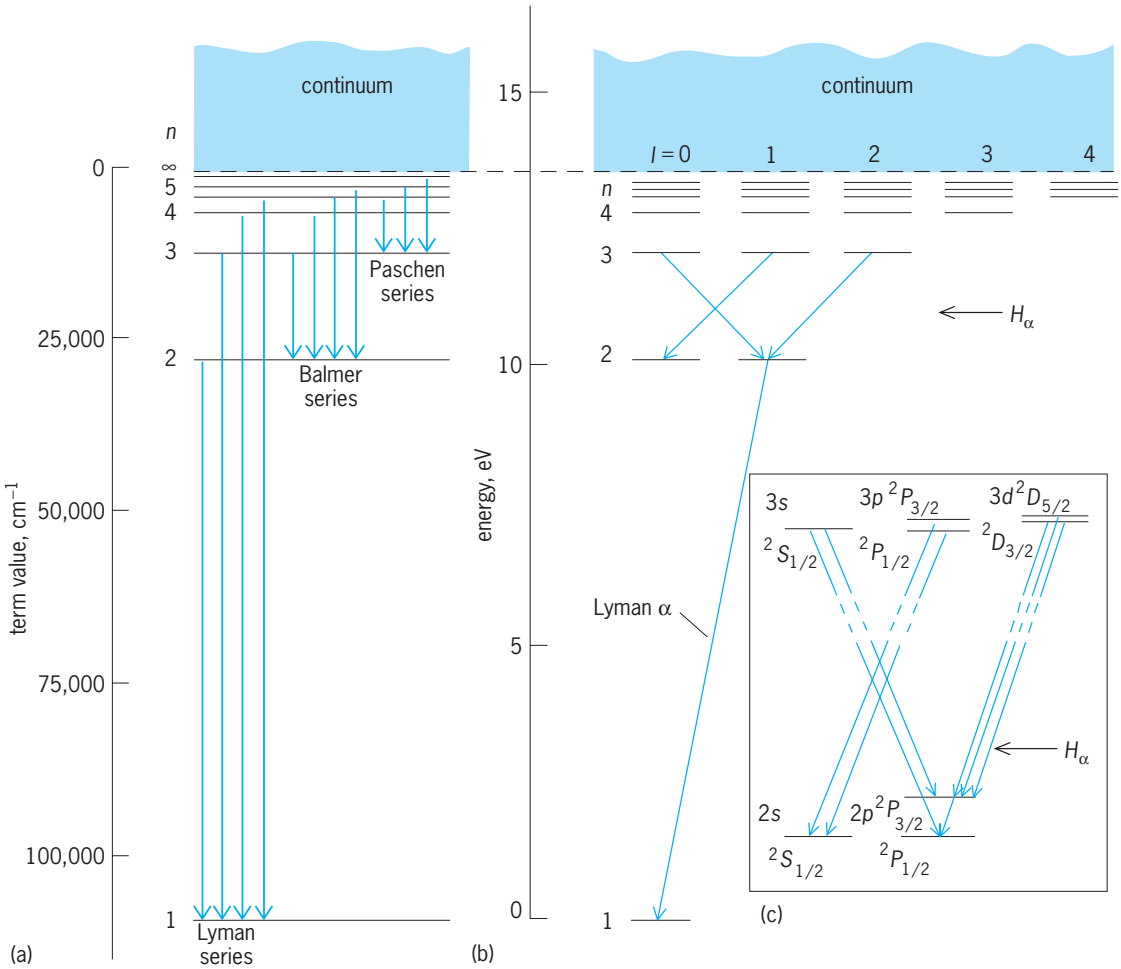

Fig. 4. Terms and transitions for the hydrogen atom. (*a*) Bohr theory. (*b*) Bohr-Sommerfeld theory. (*c*) Dirac theory.

The names of the hydrogen series are indicated in Fig. 4*a*. The spectral lines result from transitions between these various possible energy states. Each vertical arrow on the diagram connects two states and represents the emission of a particular line in the spectrum. The wave number of this line, according to Bohr's frequency condition, is equal to the difference in the term values for the lower and upper states and is therefore proportional to the length of the arrow. The wave numbers of all possible lines are given by Eq. (12), known as the Balmer formula,

$$\sigma = T'' - T' = R\left(\frac{1}{n''^2} - \frac{1}{n'^2}\right) \qquad (12)$$

where the double primes refer to the lower energy state (larger term value) and the single primes to the upper state. Any particular series is characterized by a constant value of n'' and variable n'. The only series with any lines in the visible region is the Balmer series, represented by Eq. (12) with $n'' = 2$. Its first line, that for $n' = 3$, is the bright red line at the wavelength 656.3 nanometers and is called H_α. Succeeding lines H_β, H_γ, and so on proceed toward the ultraviolet with decreasing spacing and intensity, eventually converging at 364.6 nm. Beyond this series limit there is a region of continuous spectrum. The other series, given by $n'' = 1, 3, 4$, etc., lie well outside the visible region. The Lyman series covers the wavelength range 121.5–91.2 nm in the vacuum ultraviolet, and the Paschen series 1875.1–820.6 nm in the infrared. Still other series lie at even longer wavelengths.

Since hydrogen is by far the most abundant element in the cosmos, its spectrum is extremely important from the astronomical standpoint. The Balmer series has been observed as far as H_{31} in the spectra of hot stars. The Lyman series appears as the strongest feature of the Sun's spectrum, as photographed by rockets and satellites orbiting above the Earth's atmosphere such as Skylab.

Nuclear Magnetism and Hyperfine Structure

Most atomic nuclei also possess spin, but rotate about 2000 times slower than electrons because their mass is on the order of 2000 or more times greater than that of electrons. Because of this, very weak nuclear magnetic fields, analogous to the electronic ones that produce fine structure in spectral lines, further split atomic energy levels. Consequently, spectral lines arising from them are split according to the relative orientations, and hence energies of interaction, of the nuclear magnetic moments with the electronic ones. The resulting pattern of energy

levels and corresponding spectral-line components is referred to as hyperfine structure. *See* NUCLEAR MOMENTS.

The fine structure and hyperfine structure of the hydrogen terms are particularly important astronomically. In particular, the observation of a line of 21-cm wavelength, arising from a transition between the two hyperfine components of the $n = 1$ term, gave birth to the science of radio astronomy. *See* ASTRONOMICAL SPECTROSCOPY; RADIO ASTRONOMY.

Investigations with tunable lasers. The enormous capabilities of tunable lasers have allowed observations which were impossible previously. For example, high-resolution saturation spectroscopy, utilizing a saturating beam and a probe beam from the same laser, has been used to measure the hyperfine structure of the sodium resonance lines (called the D_1 and D_2 lines). Each line was found to have components extending over about 0.017 cm^{-1}, while the separation of the D_1 and D_2 lines is 17 cm^{-1}. The smallest separation resolved was less than 0.001 cm^{-1}, which was far less than the Doppler width of the lines. *See* DOPPLER EFFECT; HYPERFINE STRUCTURE; LASER SPECTROSCOPY.

Isotope shift. Nuclear properties also affect atomic spectra through the isotope shift. This is the result of the difference in nuclear masses of two isotopes, which results in a slight change in the Rydberg constant. There is also sometimes a distortion of the nucleus, which can be detected by ultrahigh precision spectroscopy. *See* MOLECULAR BEAMS; PARTICLE TRAP.

Lamb Shift and Quantum Electrodynamics

The Bohr-Sommerfeld theory, which permitted elliptical orbits with small relativistic shifts in energy, yielded a fine structure for H_α that did not agree with experiment. The selection rule $\Delta l = \pm 1$ permits three closely spaced transitions. Actually, seven components have been observed within an interval of 0.5 cm^{-1}. According to the Dirac theory, the spin-orbit and the relativistic correction to the energy can be combined in the single formula (13), so that lev-

$$\Delta T = \frac{R\alpha^2 Z^4}{n^3} \cdot \left(\frac{2}{2j+1} - \frac{3}{4n} \right) \qquad (13)$$

els with the same n and j coincide exactly, as shown in Fig. 4*c*. The splittings of the levels are greatly exaggerated in the figure. Selection rules, described later, limit the transitions to $\Delta j = 0$ and ± 1, so there are just seven permitted transitions for the $H\alpha$ line. Unfortunately for the theory, two pairs of these transitions coincide, if states of the same j coincide, so only five lines of the seven observed would be accounted for.

The final solution of this discrepancy came with the experimental discovery by W. Lamb, Jr., that the $2s^2S$ level is shifted upward by 0.033 cm^{-1}. The discovery was made by extremely sensitive microwave measurements, and it has subsequently been shown to be in accord with the general principles of quantum electrodynamics. The Lamb shift is present to a greater or lesser degree in all the hydrogen levels, and also in those of helium, but is largest for the levels of smallest l and n. Accurate optical measurements on the wavelength of the Lyman α line have given conclusive evidence of a shift of the $1s^2S$ level.

The knowledge of this 1*s* level shift has improved with great rapidity. By 2000, the transition frequency (energy of the transition divided by Planck's constant h) separating the 1*s* and 2*s* states of atomic hydrogen had been measured with an uncertainty of only 2 parts in 10^{14}, representing the most accurate measurement up to that time of any frequency or wavelength in the visible or ultraviolet regions of the light spectrum. This high accuracy is possible primarily because (1) the 2*s* state in hydrogen is so long-lived that the natural width of the transition is only about 1 Hz; and (2) by probing a cold atomic hydrogen beam with two simultaneous laser photons, each with half the transition energy, from opposite directions, the Doppler broadening arising from the motion of the target atoms in the laser field is made to cancel to first order. As further discussed below, such Doppler spreads are usually a major source of smearing out such measurements.

Though an extended discussion of the highly abstract field theory of quantum electrodynamics is inappropriate here, a simple picture of the lowest-order contribution to the Lamb shift in hydrogen can be given. No physical system is thought ever to be completely at rest; even at temperatures near absolute zero (about −273°C or −460°F), there is always a zero-point oscillatory fluctuation in any observable physical quantity. Thus, electromagnetic fields have very low-level, zero-point fluctuations, even in vacuum, at absolute zero. Electrons respond to these fluctuations with a zero-point motion: in effect, even a point charge is smeared out over a small volume whose mean square radius, though tiny, is not exactly zero. The smearing out of the charge distribution changes the interaction energy of a charged particle with any additional external field in which it may reside, including that of an atomic nucleus. As *s* electrons spend more time near the nucleus, where the nuclear-field Coulomb electric field is most intense, than do, say, electrons having more extended orbits, the change in interaction energy is greater for *s* electrons than, say, *p* electrons, and is such as to raise the energy of *s* electrons above the exact equality with that for *p* electrons of the same n arising in Dirac's theory of the electron.

The higher the charge number Z of the atomic nucleus, the more intense is the electric field experienced by atomic electrons, and hence the Lamb shift of the energy level of each atomic electron is larger. Experiments pioneered in the 1960s and later extended showed that few-electron atoms analogous to ordinary hydrogen, helium, and so forth could be prepared from atoms of higher Z by ionizing fast beams of higher-Z elements by using the thin-foil electron stripping techniques described below, and that Lamb shifts in these analogs could be measured. By 1986 the method had been extended as far as heliumlike uranium (U^{90+}), for which a one-electron Lamb shift of 70.4 ± 8.3 eV was measured

compared with a theoretical value of 75.3 ± 0.4 eV. Once again the theory of quantum electrodynamics had exhibited its unchallenged status as the most exact and successful physical field theory ever created. *See* QUANTUM ELECTRODYNAMICS.

Doppler Spread

In most cases, a common problem called Doppler broadening of the spectral lines arises, which can cause overlapping of spectral lines and make analysis difficult. The broadening arises from motion of the emitted atom with respect to a spectrometer. Just as in the familiar case of an automobile horn sounding higher pitched when the auto is approaching an observer than when receding, so also is light blueshifted (to higher frequency) when approaching, and redshifted when receding from, some detection apparatus. The percentage shift in frequency or wavelength is approximately $(v/c) \cos \theta$, where v is the emitter speed, c the speed of light, and θ the angle between $\vec{v}$ and the line of sight of the observer. Because emitting atoms normally are formed with some spread in v and θ values, there will be a corresponding spread in observed frequencies or wavelengths. Several ingenious ways of isolating only those atoms nearly at rest with respect to spectrometric apparatus have been devised. The most powerful employ lasers and either involve the saturation spectroscopy mode mentioned above or use two laser photons which jointly drive a single atomic transition and are generated in lasers so arranged that the first-order Doppler shifts of the photons cancel each other.

Radiationless Transitions

It would be misleading to think that the most probable fate of excited atomic electrons consists of transitions to lower orbits, accompanied by photon emission. In fact, for at least the first third of the periodic table, the preferred decay mode of most excited atomic systems in most states of excitation and ionization is the electron emission process first observed by P. Auger in 1925 and named after him. For example, a singly charged neon ion lacking a 1*s* electron is more than 50 times as likely to decay by electron emission as by photon emission. In the process, an outer atomic electron descends to fill an inner vacancy, while another is ejected from the atom to conserve both total energy and momentum in the atom. The ejection usually arises because of the interelectron Coulomb repulsion. *See* AUGER EFFECT.

The vast preponderance of data concerning levels of excited atomic systems concerns optically allowed, single-electron, outermost-shell excitations in low states of ionization. Since much of the mass in nature is found in stars, and most of the elements therein rarely occupy such ionization-excitation states, it can be argued that presently available data provide a very unrepresentative description of the commonly occurring excited atomic systems in nature. When the mean lives of excited atomic systems are considered, the relative rarity of lifetime measurements on Auger electron-emitting states is even more striking. The experimentally inconvenient typical lifetime range (10^{-12} to 10^{-16} s) accounts for this lack.

Spectroscopy of Highly Ionized Atoms

Since the 1960s, an effective means of creating high ionization-excitation states, such as those that are common in stars and other high-temperature plasmas, has been provided by beam-foil spectroscopy. In this method, ions accelerated by a charged-particle accelerator pass through a thin carbon foil. The resulting beam of ions stripped of one or more electrons can be analyzed so the charge states and velocities are accurately known. Highly excited states are often produced by the interaction with the foil. The subsequent emission, detected with a spectrometer, allows the measurement of lifetimes of states which are difficult to produce in other sources. When, for example, electromagnetic atomic decays are observed, optical or x-ray spectrometers may be useful. Sometimes the intensity of the emitted radiation undergoes fluctuations called quantum beats, analogous to sound intensity heats of tuning forks of slightly different frequency, as can be seen in the photograph in **Fig. 5**. When, as is most frequently the case, Auger processes are dominant, less familiar but comparably effective electron spectrometers are preferred. *See* BEAM-FOIL SPECTROSCOPY.

Violent, high-velocity atomic collisions or electron-atom collisions are needed to reach the highest ionization-excitation states possible. Because highly ionized atoms are very important in stellar atmospheres, and in terrestrially produced highly ionized gases (plasmas) on which many potentially important schemes for thermonuclear fusion energy production are based, it is also important to overcome the Doppler spread problem in violent collisions. For sufficiently highly ionized, fast, heavy projectiles, lighter target atoms in a gaseous target

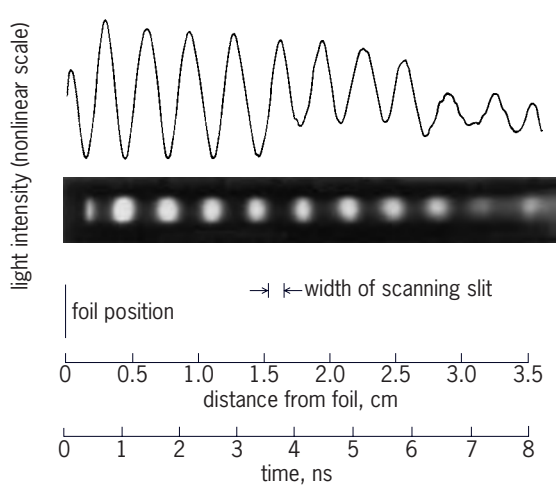

Fig. 5. Example of atomic process following excitation of a beam of ions in a thin solid target less than 1 μm in thickness. Quantum fluctuations in the light emitted by a 2-mm-diameter beam of such particles, traveling at a few percent of the speed of light, are shown as the atoms travel downstream from the target. (*After I. A. Sellin et al., Periodic intensity fluctuations of Balmer lines from single-foil excited fast hydrogen atoms, Phys. Rev., 184:56–63, 1969*)

can be ionized to very high states of ionization and excitation under single-collision conditions, while incurring relatively small recoil velocities. Sample spectra indicate that a highly charged projectile can remove as many as $Z - 1$ of the Z target electrons while exciting the last one remaining, all under single-collision-event conditions. Subsequent studies showed that the struck-atom recoil velocities are characterized by a velocity distribution no worse than a few times thermal, possibly permitting a new field of precision spectroscopy on highly ionized atoms to be developed.

In the 1980s and 1990s more direct methods of producing ions in high ionization-excitation states underwent rapid development. Ion sources have been developed that can produce beams of ions that have much greater intensities but suffer from higher kinetic energies (of the order of keV) than those which are readily available from ion sources based on the recoil ion production method (eV energies). With some exceptions, these ion sources usually produce lower ionization-excitation states than are found by using the very fastest ion beams or the beam-foil method; however, those methods are usually characterized by corresponding ion kinetic energies in the GeV range.

The two most prominent types of sources are electron cyclotron resonance (ECR) sources and cryogenic electron-beam ion sources (CRYEBIS). The former depend for their operation on ionization and excitation of ions by impact of electrons heated to high temperatures in an intense microwave discharge. The latter also exploit electron impact ionization and excitation, but accomplish this objective by first using electrostatic forces to trap ions in low ionization-excitation states inside intense electron beams, and then making use of additional electron impact excitation and ionization to achieve high ionization-excitation states. Periodic expulsion of ions from the trapping region after a suitable time delay is then used to form ion beams which are hence inherently pulsed in time, in contrast to beams from electron cyclotron resonance sources, which are normally extracted in a steady-state manner. *See* ION SOURCES.

In the 1990s a shorter, more compact version of the CRYEBIS, the electron beam ion trap (EBIT), underwent rapid development. Its operating principles are very similar, but it features close-up access to a short and very thin column of highly ionized ions excited by intense high-energy electron impact on atoms in the limited fields of view of specialized spectroscopic instruments mounted as close to the interaction region as possible. It is the favored instrument for increasingly precise x-ray spectroscopy on highly ionized atoms.

CRYEBIS sources are used as injectors for heavy-ion storage rings, and are designed to provide the much larger fluxes of high-charge-state ions needed to fulfill the potential of such rings. These machines recycle the supply of stored ions in the ring once every few microseconds through a given orbital point, enhancing their effective numbers, since each ion can be interrogated once per cycle by each of one or more local measuring instruments. Lifetime measurements can be carried out by monitoring decay times of the ions. The application of merged-electron-beam cooling in the ring (discussed further below) to greatly reduce the longitudinal and transverse velocities of stored ions is a further crucial attribute, permitting, for example, study of very low relative velocity interactions of merged electrons and ions inside the ring. *See* PARTICLE ACCELERATOR.

Relativistic Dirac Theory and Superheavy Elements

Dirac's theory of a single electron bound to a point nucleus implies a catastrophe for hydrogen atoms of high Z, if Z could be made on the order of $137 \cong 1/\alpha$ or larger [in which $\alpha = e^2/(2\epsilon_0 hc)$ is the fine-structure constant]. Dirac's expression for the electronic binding energy becomes unphysical at that point. Nuclear theorists have, however, contemplated the possible stable existence of superheavy elements, appreciably heavier than any observed heretofore, say, in the range $Z = 114$ to 126 or even larger. Though searches for such superheavy elements of either primordial or artificial origin have thus far proved unsuccessful, the advent of heavy-ion accelerators capable of bringing heavy particles to a distance of closest approach much smaller than the radius of the innermost shell in any atom, even a heavy one, raises the possibility of transient creation of a superheavy atom with a combined atomic number, $Z_{\text{comb}} = Z_1$ and Z_2, of about 180 or even greater (in which Z_1 and Z_2 are the individual atomic numbers of the colliding system).

The binding energy of a 1*s* electron in an atom increases rapidly with increasing atomic number. As already noted, in the usual linear version of the Dirac equation, a catastrophe occurs when the nuclear charge reaches $(1/\alpha) \cong 137$. When $Z_1 + Z_2$ is sufficiently large, the 1*s* binding energy can reach and exceed twice the rest energy of the electron $m_e c^2$, in which m_e is the electron mass and c is the speed of light. By varying the target or projectile, one can trace the path of the electron energies as they dive toward the negative-energy sea, or, in the case of the 1*s* electron, into the negative-energy sea, to give rise to bound states degenerate with the negative-energy continuum (**Fig. 6**). The concept of a negative-energy sea at $-m_e c^2$ was introduced by Dirac and used to predict the existence of a new particle, the positron, represented by a hole in the sea. The 1*s* level acquires a width because of the admixture of negative-energy continuum states, and the 3*s*-state width corresponds to a decaying state, provided there is nonzero final-state density. When a hole is present in the K shell, spontaneous positron production is predicted to take place. Production of an electron-positron pair is possible, since a final 1*s* ground state is available for the produced electron; the positron is predicted to escape with kinetic energy corresponding to overall energy balance. *See* ANTIMATTER; ELECTRON-POSITRON PAIR PRODUCTION.

Because of the effects of finite nuclear size in modifying the potential, $Z_1 + Z_2$ must reach some initial

value Z_{CR} greater than 137 at which the "splash" into the continuum is made. According to nuclear-model-dependent estimates, the splash would occur at about 170. Some types of nonlinear additions to the Dirac equation may remove the diving into the negative-energy continuum, and these could be tested by observation of positron production as a function of $Z_1 + Z_2$. Other "limiting field" forms of nonlinear electrodynamics lead to larger values of Z_{CR}, and since the positron escape width turns out to be approximately proportional to the square of the difference between the effective quasi-nuclear charge and Z_{CR}, these nonlinearities could be tested too. Intensive searches for spontaneous positron production have been undertaken but have not yet met with success.

Such spontaneous positron production, if ever found, would represent an effect due to the potential energy of interaction of a combined nucleus of superheavy charge $Z_1 + Z_2$ with Dirac's negative energy sea. However, experiments, carried out chiefly in the 1990s, have observed copious production of dynamically induced positrons from the reservoir of kinetic energy brought into collisions of heavy projectile ions with atoms at speeds very close to the speed of light.

The physical picture here is that of a naked ion (no electrons remaining) passing by a target atom at such close range and so fast that the disturbance due to the passage of the ion's electrical charge closely resembles a very energetic transverse light pulse, whose spectrum of photon energies is entirely sufficient to liberate positrons from Dirac's negative energy sea via electron-positron pair production, in which both an electron and a positron are born in the Coulomb field of the struck atom. In this picture, free electron-positron pairs are formed by collisional ionization of electrons initially lying in Dirac's negative electron continuum. Sometimes excitation of a negative-continuum electron into an unoccupied bound state of either the projectile or target particle manifests itself as atomic capture of the electron produced in such pair production, reducing the electrical charge of the ion by one unit, while simultaneously producing a free positron. Eventually, in either process, the positron inevitably collides with some other electron and annihilates, converting its Einstein rest energy and that of a struck electron back into photons. Rates for both the production of free pairs and pair-produced electron capture (**Fig. 7**) are large—of the scale for ordinary atomic as opposed to elementary-particle-interaction reaction rates—and increase with the energy of the collision. *See* POSITRON; QUASIATOM.

Uncertainty Principle and Natural Width

A principle of universal applicability in quantum physics, and of special importance for atomic physics, is W. Heisenberg's 1927 uncertainty principle. In its time-energy formulation, it may be regarded as holding that the maximum precision with which the energy of any physical system (or the corresponding frequency ν from $E = h\nu$) may be determined is limited by the time interval Δt available for a measurement of the energy or frequency in question, and in no case can ΔE be less than approximately $h/(2\pi\,\Delta t)$. The immediate consequence is that the energy of an excited state of an atom

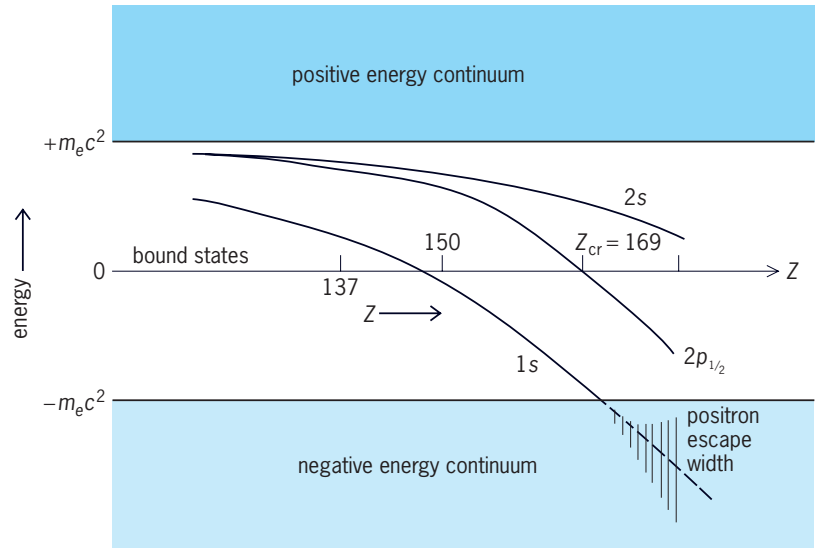

Fig. 6. Behavior of the binding energies of the 1s, $2p_{1/2}$, and 2s electrons with increasing atomic number Z. (*After J. H. Hamilton and I. A. Sellin, Heavy ions: Looking ahead, Phys. Today, 26(4):42–49, April 1973*)

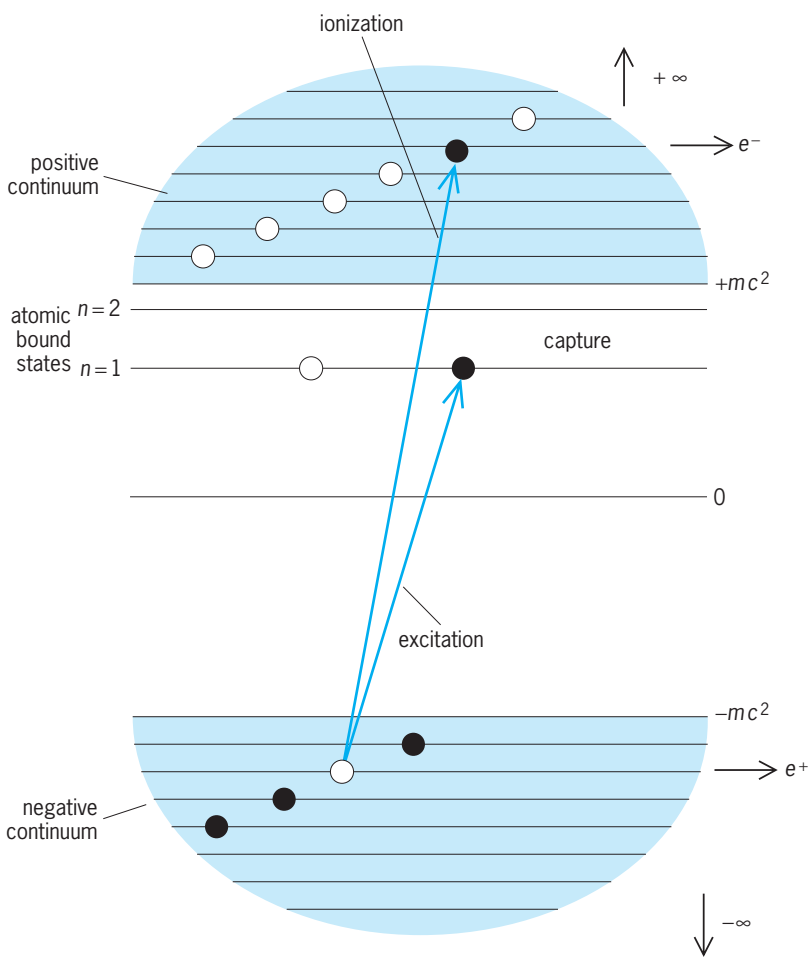

Fig. 7. Free positron (e^+) and free (ionization) or bound (capture) electron (e^-) production, arising from pair-production creation of a hole (positron) in Dirac's negative-electron-energy sea. The energy mc^2 refers to the einsteinian rest-mass energies of electrons and positrons, of 511,000 eV each. (*After Charles Vane*)

cannot be determined with arbitrary precision, as such states have a finite lifetime available for measurement before decay to a lower state by photon or electron emission. It follows that only the energies of particularly long-lived excited states can be determined with great precision. The energy uncertainty ΔE, or corresponding frequency interval $\Delta\nu$, is referred to as the natural width of the level. For excited atoms radiating visible light, typical values of $\Delta\nu$ are on the order of 10^9 Hz. (For a long time it was thought that the natural widths of levels participating in laser-cooling transitions set a lower limit to how low a temperature laser cooling can achieve. However, ingenious schemes have been developed that do not violate the uncertainty principle but skillfully evade it.) For inner electrons of heavy atoms, which give rise to x-ray emission, and also for Auger electron-emitting atoms, 10^{14} Hz is more typical. An energy interval $\Delta E = 1$ eV corresponds to about 2.4×10^{14} Hz. *See* UNCERTAINTY PRINCIPLE.

The time-energy uncertainty principle is perhaps best regarded as a manifestation of the basic wave behavior of electrons and other quantum objects. For example, piano tuners have been familiar with the time-frequency uncertainty principle for hundreds of years. Typically, they sound a vibrating tuning fork, or its modern electronic equivalent, of accepted standard frequency in unison with a piano note of the same nominal frequency, and listen for a beat note between the tuning fork and the struck tone. Beats, which are intensity maxima in the sound, occur at a frequency equal to the difference in frequency between the two sound sources. For example, for a fork frequency of 440 Hz and a string frequency of 443 Hz, three beat notes per second will be heard. The piano tuner strives to reduce the number of beats per unit time to zero, or nearly zero. To guarantee a frequency good to 440 ± 0.01 Hz would require waiting for about 100 s to be sure no beat note had occurred ($\Delta\nu\Delta t \gtrsim 1$). *See* BEAT.

Cooling and Stopping Atoms and Ions

Despite impressive progress in reducing Doppler shifts and Doppler spreads, these quantities remain factors that limit the highest obtainable spectroscopic resolutions. The 1980s and 1990s saw extremely rapid development of techniques for trapping neutral atoms and singly charged ions in a confined region of space, and then cooling them to much lower temperatures by the application of laser-light cooling techniques. Photons carry not only energy but also momentum; hence they can exert pressure on neutral atoms as well as charged ions. *See* LASER COOLING.

Schemes have been developed to exploit these light forces to confine neutral atoms in the absence of material walls, whereas various types of so-called bottle configurations of electromagnetic fields developed earlier remain the technique of choice for similarly confining ions. For neutral atoms and singly charged ions, whose energy levels (unlike those of most more highly charged ions) are accessible to tunable dye lasers, various ingenious methods have been invented to slow down and even nearly stop atoms and ions. More effective than cooling methods known from low-temperature physics, these methods often utilize the velocity-dependent light pressure from laser photons of nearly the same frequency as, but slightly less energetic than, the energy separation of two atomic energy levels to induce a transition between these levels. This procedure serves to cool atomic systems to submillikelvin temperatures. For example, an atom may absorb photons slightly below the precise resonance energy separation owing to the Doppler shift arising from its motion toward a laser source, and then reradiate fluorescent photons of slightly higher average energy in all directions. The atom thereby loses net directed momentum because of the redistribution in all directions and also loses net kinetic energy of motion because of the requirements of total energy conservation. The use of six laser beams in so-called optical molasses schemes permits atomic momentum (and hence kinetic energy) reduction in each of the six possible directions of motion. Other light forces on atoms, known as dipole gradient forces, have been used to deflect or focus beams of neutral atoms, exploiting the rapid spatial variation of electric fields inside laser beams. Depending on laser frequency, atoms are pulled toward or driven away from regions of high laser electric field because of the nature of the interaction of an inhomogeneous electric field with polarizable atoms.

This ingenious optical dipole trapping scheme was quickly improved upon by the invention of the magnetooptic trap. This device is larger and more robust. It combines optical forces provided by laser light with a weak magnetic field whose size goes through zero at the geometrical center of the trap and increases with distance from this center. The net result is a restoring force which confines sufficiently laser-cooled atoms near the center.

To universal surprise, initial measurements of the temperature of such trapped ions were only about 40×10^{-6} K above absolute zero. Until then, an anticipated value some six times greater had been regarded as a fundamental limit associated with the natural widths of the levels involved. Ingenious improvements have allowed cooling of ions to temperatures as low as 180×10^{-9} K.

For more highly ionized ions, annular storage rings are used in which radial confinement of fast ion beams (with speeds of approximately 10% or more of the speed of light) is provided by magnetic focusing, and longitudinal confinement automatically occurs because of the closed ionic orbits, similar to those in cyclotrons. Two cooling schemes are known to work on stored beams of charged particles, the so-called stochastic cooling method and the electron cooling method. In the former, deviations from mean stored particle energies are electronically detected, and electronic "kicks" that have been adjusted in time and direction are delivered to the stored particles to compensate these deviations. In electron cooling, which proves to be more effective for stored heavy ions of high charge, electron

beams prepared with a narrow velocity distribution are merged with the stored ion beams. When the average speeds of the electrons and the ions are matched, the Coulomb interaction between the relatively cold (low-velocity-spread) electrons and the highly charged ions efficiently transfers energy from the warmer ions, thereby reducing the temperature of the stored ions. Electron cooling occurs quickly, and has been demonstrated to cool beams of fast protons sufficiently to reduce the ratio of the spread in the longitudinal momentum to the total momentum to better than 1 part per million. Several heavy-ion storage rings are presently in operation.

Successful Explanations and Unresolved Problems

Many of the atomic enigmas referred to above have been explained through experimental discoveries and theoretical inspiration. Examples include Rutherford scattering, the Bohr model, the invention of the quantum theory, the Einstein quantum relation between energy and frequency $E = h\nu$, the Pauli exclusion and Heisenberg uncertainty principles, and many more. A few more are discussed below, but only in part, as the subject of atomic structure and spectra is by no means immune from the internal contradictions and defects in understanding that remain to be resolved.

In 1927 C. Davisson and L. Germer demonstrated that electrons have wave properties similar to that of light, which allows them to be diffracted, to exhibit interference phenomena, and to exhibit essentially all the properties with which waves are endowed. It turns out that the wavelengths of electrons in atoms are frequently comparable to the radius of the Bohr orbit in which they travel. Hence, a picture of a localized, point electron executing a circular or elliptical orbit is a poor and misleading one. Rather, the electrons are diffuse, cloudlike objects which surround nuclei, often in the form of standing waves, like those on a string. Crudely speaking, the average acceleration of an electron in such a pure standing wave, more commonly called a stationary state in the quantum theory, has no time-dependent value or definite initial phase. Since, as in Maxwell's theory, only accelerating charges radiate, atomic electrons in stationary states do not immediately collapse into the nucleus. Their freedom from this catastrophe can be interpreted as a proof that electrons may not be viewed as pointlike planetary objects orbiting atomic nuclei. *See* ELECTRON DIFFRACTION.

Atoms are by no means indivisible, as Thomson showed. Neither are their nuclei, whose constituents can be ejected in violent nucleus–nucleus collision events carried out with large accelerators. Nuclear radii and shapes can, for example, be measured by high-energy electron scattering experiments, in which the electrons serve as probes of small size. *See* NUCLEAR STRUCTURE.

However, a good theory of electron structure still is lacking, although small lower limits have been established on its radius. Zero-point fluctuations in the vacuum smear their physical locations in any event. There is still no generally accepted explanation for why electrons do not explode under the tremendous Coulomb repulsion forces in an object of small size. Estimates of the amount of energy required to "assemble" an electron are very large indeed. Electron structure is an unsolved mystery, but so is the structure of most of the other elementary objects in nature, such as protons, neutrons, neutrinos, quarks, and mesons. There are hundreds of such objects, many of which have electromagnetic properties, but some of which are endowed with other force fields as well. Beyond electromagnetism, there are the strong and weak forces in the realm of elementary particle physics, and there are gravitational forces. Electrons, and the atoms and molecules whose structure they determine, are only the most common and important objects in terrestrial life. *See* ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS.

Though the action-at-a-distance concept built into a Coulomb's-law description of the interactions of electrons with themselves and with nuclei is convenient and surprisingly accurate, even that model of internal atomic interactions has its limits. It turns out that electrons and nuclei do not interact instantaneously, but only after a delay comparable to the time needed for electromagnetic waves to travel between the two. Electrons and nuclei do not respond to where they are with respect to each other, but only to where they were at some small but measurable time in the past.

Quantum jumps. Einstein is credited with introducing the idea of discontinuous transitions in the time evolution of microscopic quantum states, such as are involved in atomic transitions. Bohr's postulate 3, discussed earlier, implicitly assumed that electrons changing states and radiating photons do so in discontinuously short time intervals—examples of what are commonly referred to as quantum jumps. E. Schrödinger, the formulator of the Schrödinger equation, which well describes many features of the quantum behavior of electrons in atoms as well as that of numerous other microscopic quantum systems, hoped that the necessity of assuming discontinuous quantum jumps could be avoided. He hoped that refined calculations would show a smooth change between the wave function solutions which satisfy the Schrödinger equation for individual stationary Bohr orbits when transitions between orbits occur. The presently accepted status of quantum theory nevertheless remains that of a purely statistical theory: the numerical coefficients of terms which when added together constitute an atomic wave function which satisfies the Schrödinger equation for an atom in a so-called general, superposition state, determine only the average probability of finding an atom in a stationary state described by one of the terms. Finding an atom in a particular state means making a state determination measurement, such as detecting a photon whose frequency labels the final stationary state of an electron after a transition has occurred.

If by a quantum-mechanical system is meant an ensemble of a large number of identically prepared atoms, then these average values may indeed evolve

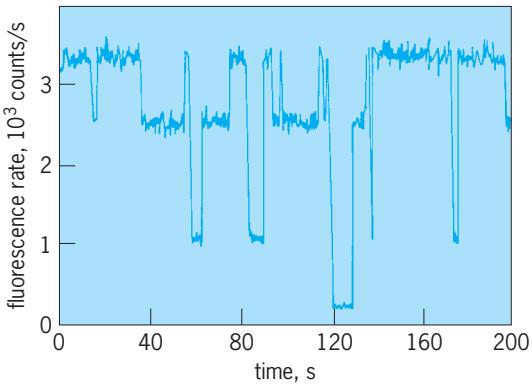

Fig. 8. Multiple quantum jumps occurring in the laser-excited 493-nm fluorescence of three Ba^+ ions. (*After W. Neuhauser and Th. Sauter, Observation of quantum jumps, Comments Atom. Mol. Phys., 21:83–95, 1988*)

continuously in time. However, if it is possible to prepare a very small number of isolated atoms in excited states in the same way again and again—in effect, to realize a time average, not an ensemble average over subsequent photon emissions—the question arises of whether discontinuous quantum jumps are truly observable.

Evidence that they are observable is seen in **Fig. 8**, which shows the intensity of light from large numbers of fluorescent photons which are emitted during rapid, repetitive laser excitation of a particular excited state in one or more of three Ba^+ ions contained in an electromagnetic ion trap. These ions most often relax back to the ionic ground states (states of no excitation) by the fluorescent photon emission whose intensity is observed. Occasionally, each of the ions relaxes instead to a state whose energy is intermediate, for which selection rules impede photon emission to the ground state—in effect, "shelving" the ion for a while in a state for which the laser is ineffective in repetitive reexcitation to the fluorescent state because its frequency does not match the frequency between the intermediate and upper levels. The apparently discontinuous switching to dark states is seen to reduce the fluorescent intensity to about two-thirds of the original level when one ion jumps to a dark state, to one-third of the level when two ions jump, and to nearly zero intensity when three ions jump. An additional remarkable feature of these observations is that the number of apparently simultaneous jumps of two or three ions at a time exceeds that expected on the basis of random coincidence of jumps by about two orders of magnitude. Cooperative interactions of even a very small number of atomic systems with the laser radiation field seem to have unexpectedly high probability.

Photon antibunching. Still other intimate connections between quantum jumps in atoms and the photons constituting the electromagnetic radiation field to which they give birth are illustrated by a phenomenon called photon antibunching. Fluorescent photon emission from a large ensemble of excited atoms in thermal equilibrium at some high temperature is said to be incoherent, meaning that individual atoms emit photons at random times; that is, the instant of photon emission of a given atom is not correlated with that of others in the ensemble. Under these circumstances the probability of a second photon being detected at some time interval Δt after detection of a first has a maximum at $\Delta t = 0$, a result called photon bunching since the photon emission tends to occur in bursts.

For one to a few ions stored in an electromagnetic trap, however, the smallest value of the probability can occur at $\Delta t = 0$, an effect called antibunching. That this result is plausible can be demonstrated by considering a trapped ion which returns to its ground state following fluorescent photon emission. If it is illuminated by a laser which reexcites it to the level from which fluorescence is again possible, the reexcitation is in general not instantaneous but occurs after some mean time delay determined by the laser intensity and the probability per unit time for resonant laser photon absorption, which is also a statistical process. The probability of two fluorescent photons being emitted in a time interval that is very short compared to that required on the average for reexcitation should therefore be very small, as the likelihood that the ion is still in its ground state should then be very high.

An illustration of antibunching for a single Mg^+ ion stored in an electromagnetic trap is shown in **Fig. 9**. The second-order correlation function (intensity correlation) $g^{(2)}$ of the $3^2S_{1/2} - 3^2P_{3/2}$ transition

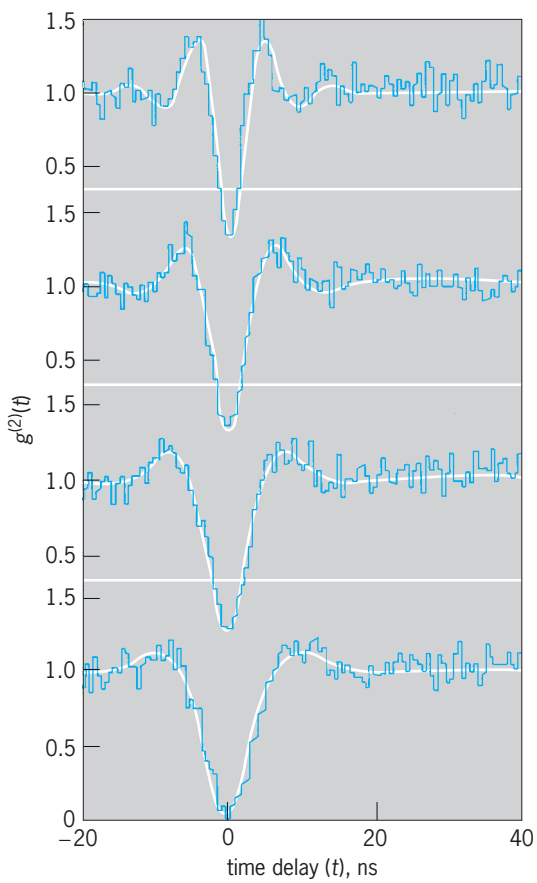

Fig. 9. Antibunching signal of a single Mg^+ ion for four different laser intensities. (*After F. Diedrich and H. Walther, Nonclassical radiation of a stored ion, Phys. Rev. Lett., 58:203–206, 1987*)

in Mg^+ is plotted versus the time delay t between two fluorescent photon emissions for four different laser intensities. The quantity $g^{(2)}(t)$ is proportional to the probability of a second photon being detected after a first emission with a time delay t. The large antibunching effect near zero time delay is strikingly demonstrated. Ivan A. Sellin

Bibliography. E. U. Condon and H. Odabasi, *Atomic Structure*, Cambridge University Press, 1980; R. D. Cowan, *The Theory of Atomic Structure and Spectra*, University of California Press, 1981; H. Friedrich, *Theoretical Atomic Physics*, 3d ed., Springer-Verlag, 2005; B. Levi, Work on atom trapping and cooling gets a warm reception in Stockholm, *Phys. Today*, 50(12):17–19, December 1997; B. Schwarzschild, Optical frequency measurement is getting a lot more precise, *Phys. Today*, 50(12):19–21, December 1997; L. Szasz, *The Electronic Structure of Atoms*, Wiley, 1992.

Atomic theory

The study of the structure and properties of atoms based on quantum mechanics and the Schrödinger equation. These tools make it possible, in principle, to predict most properties of atomic systems. A stationary state of an atom is governed by a time-independent wave function which depends on the position coordinates of all the particles within the atom. To obtain the wave function, the time-independent Schrödinger equation has to be solved. This is a second-order differential equation (1), where the first term in parenthesis is the so-

$$\begin{aligned} H\Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \ldots, \mathbf{r}_N) &= \left(\sum_{i=1}^{N} \frac{-\hbar^2}{2m_i}\nabla_i^2 + V(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \ldots, \mathbf{r}_N)\right) \\ &\quad \times \Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \ldots, \mathbf{r}_N) \\ &= E\Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \ldots, \mathbf{r}_N) \qquad (1) \end{aligned}$$

called kinetic energy operator, the second term is the potential energy operator, and E is the energy of the state described by Ψ. The potential energy term contains the Coulomb interaction between all the particles in the atom, and in this way they are all coupled to each other. *See* DIFFERENTIAL EQUATION; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; SCHRÖDINGER'S WAVE EQUATION.

A many-particle system where the behavior of each particle at every instant depends on the positions of all the other particles cannot be solved directly. This is not a problem restricted to quantum mechanics. A classical system where the same problem arises is a solar system with several planets. In classical mechanics as well as in quantum mechanics, such a system has to be treated by approximate methods. *See* CELESTIAL MECHANICS.

Independent particle model. As a first approximation, it is customary to simplify the interaction between the particles. In the independent particle model the electrons are assumed to move independently of each other in the average field generated by the nucleus and the other electrons. In this case the potential energy operator in Eq. (1) will be a sum over one-particle operators and Eq. (1) can be rewritten as Eq. (2). The simplest wave function, Ψ, which

$$\left(\sum_{i=1}^{N} \frac{-\hbar^2}{2m_i}\nabla_i^2 + u(\mathbf{r}_i)\right)\Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \ldots, \mathbf{r}_N) = E\Psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \ldots, \mathbf{r}_N) \qquad (2)$$

will satisfy Eq. (2) is a product of one-particle orbitals, $\Psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N = \varphi_1\varphi_2 \cdots \varphi_N$. To fulfill the Pauli exclusion principle, the total wave function must, however, be written in a form such that it will vanish if two particles are occupying the same quantum state. This is achieved with an antisymmetrized wave function, that is, a function which, if two electrons are interchanged, changes sign but in all other respects remains unaltered. The antisymmetrized product wave function is usually called a Slater determinant. *See* EXCLUSION PRINCIPLE.

Hartree-Fock method. In the late 1920s, only a few years after the discovery of the Schrödinger equation, D. Hartree showed that the wave function to a good approximation could be written as a product of orbitals, and also developed a method to calculate the orbitals. Important contributions to the method were also made by V. Fock and J. C. Slater (thus, the Hartree-Fock method).

For a stationary atomic state the energy should be stable with respect to small variations of the wave function. If the expectation value of the energy is minimized under the condition that the wave function should be a single Slater determinant, the expression for the Hartree-Fock potential is found. The Hartree-Fock model thus gives the lowest-energy ground state within the assumption that the electrons move independently of each other in an average field from the nucleus and the other electrons.

To simplify the problem even further, it is common to add the requirement that the Hartree-Fock potential should be spherically symmetric. This leads to the central-field model and the so-called restricted Hartree-Fock method. For a closed-shell system the requirement for the restricted Hartree-Fock method is automatically fulfilled.

The Hartree-Fock potential is a one-particle potential, but it depends on the orbitals of all the electrons. A common approach is to start with a reasonable approximation of the orbitals, calculate the Hartree-Fock potential from these, and solve the differential equation. This gives a new set of orbitals and the procedure can be repeated until convergence occurs. The solution is then said to be self-consistent.

The Hartree-Fock method gives a qualitative understanding of many atomic properties. Generally it is, for example, able to predict the configurations occupied in the ground states of the elements. Electron binding energies are also given with reasonable accuracy. For example, the ionization energy of neon is measured to be 21.5 eV, while the result with the Hartree-Fock method is 23 eV.

Electron correlation. Correlation is commonly defined as the difference between the full many-body problem, Eq. (1), and the Hartree-Fock model. More

specifically, the correlation energy is the difference between the experimental energy and the Hartree-Fock energy. There are several methods developed to account for electron correlation. Some important methods are the following.

In the configuration-interaction method, the total many-body wave function is expanded in a basis set, and the coefficients are determined by minimization of the expectation value of the energy. This method is widely used on both atoms and molecules. Basis sets consisting of analytic functions are often used. *See* QUANTUM CHEMISTRY.

The multiconfiguration Hartree-Fock method is a natural extension of the Hartree-Fock model. A number of configurations are chosen, and the mixing coefficients, as well as the radial part of the orbitals, are varied to minimize the expectation value of the energy.

In perturbation theory, the full hamiltonian operator, H, is divided up into two parts, $H = H_0 + V$, where H_0 can be, for example, the Hartree-Fock model (usually within the central field approximation), and then a perturbation expansion of V is carried out. The operator H_0 must be a reasonable approximation to H, otherwise the perturbation expansion may not converge. *See* PERTURBATION (QUANTUM MECHANICS).

All these methods require large-scale numerical calculations. The ability to perform accurate calculations increased dramatically during the 1980s and 1990s due to the development of fast computers.

One way to classify the different contributions to the correlation energy is to count the number of electrons involved in the different terms. The most important type of correlation is pair correlation. As an example, **Table 1** shows the importance of different correlation contributions to the ground state of the beryllium atom, a system with four electrons. The relative sizes of the contributions can be understood by comparison with a system of colliding classical particles. In such a system, collisions between two particles dominate and three-particle collisions are very rare. True four-particle collisions are also very rare, but two simultaneous two-particle collisions constitute a kind of effective four-particle effect which is more common.

Table 1 shows the contribution to the correlation energy only. The Hartree-Fock model accounts for 99% of the total binding energy of beryllium, and thus only around 1% of the binding energy is due to correlation. This success of the independent-particle model for the total energy is due to the large contributions from the inner electrons. These electrons interact primarily with the nucleus and only to a smaller extent with the other electrons, and thus are less influenced by correlation.

TABLE 1. Contributions to the total correlation energy for the ground state of beryllium

Number of particles involved	Contribution, %
Singles	0.7
Doubles	94.6
Triples	0.6
Quadruples	4.1

Strongly correlated systems. Although the Hartree-Fock model can qualitatively explain many atomic properties, there are systems and properties for which correlation is more important.

Many, but not all, atoms can bind an extra electron and form a negative ion. Since there is no long-range field outside a neutral atom, it is the detailed electron-electron interaction which makes binding possible. In general, the Hartree-Fock model makes quite poor quantitative predictions for the so-called electron affinity (the binding energy of the extra electron). Sometimes a calculation within the Hartree-Fock model will even predict that it is not possible to form a negative ion, but experiments as well as more refined correlated calculations show that it is indeed possible. Calcium is one such element; the electron affinity is around 0.02 eV, although no bound state is found in a Hartree-Fock calculation. *See* ELECTRON AFFINITY; NEGATIVE ION.

Excited states are often much more sensitive to electron correlation than the ground state. One well-known example comes from the beryllium atom. The ground state of beryllium can be denoted $(1s^2 2s^2)\ {}^1S_0$. In beryllium-like ions the first excited state of ${}^1S_0$ symmetry is the $(1s^2 2p^2)\ {}^1S_0$ state. In neutral beryllium, where the attractive interaction with the nucleus is relatively weaker, this state is not bound at all. The behavior can be explained only if electron correlation, which pushes it up above all the $(1s^2 2sns)\ {}^1S_0$ states (where n is the primary quantum number of the outer most electron), is accounted for. In fact, it is generally true that doubly excited states constitute strongly correlated systems. *See* ELECTRON CONFIGURATION.

In open-shell systems the electrons within the shell interact strongly with each other, and generally many states can be formed from one configuration. In an average level calculation the same Hartree-Fock orbitals are used for all these levels. This is a natural independent-particle starting point, but often large discrepancies are found when the result is compared with experiments. In the rare-earth elements (elements 57–70), the $4f$ shell is successively filled. For example, the Pr^{3+} ion is a system with two $4f$ electrons outside closed shells. Within this configuration it is possible to form 13 ${}^{2S+1}L_J$ states, and the independent-particle model does not even give their energies in the correct order. *See* RARE-EARTH ELEMENTS.

If the interest is not in calculating the total energy of a state but in understanding some other properties, effects beyond the central field model can be more important. Small changes in the wave function may have a large impact on some weak effect, although the total energy is insensitive. One example comes from the so-called hyperfine structure, that is, the interaction between the electronic and the nuclear magnetic moments. The most important contribution comes from the electronic wave function

inside the nucleus. In closed shells there are as many electrons with spin up as spin down, and the hyperfine interactions from these electrons cancel each other. If there are open outer shells, the electrons in these shells will interact slightly differently with the closed-shell spin-up and spin-down electrons. In this way the closed shell will be polarized; that is, the overlap with the nucleus will differ for a spin-up and a spin-down electron. Since inner electrons have a large overlap with the nucleus, even a small imbalance gives rise to a substantial effect. *See* HYPERFINE STRUCTURE.

Relativistic effects. The Schrödinger equation is a nonrelativistic wave equation. In heavy elements the kinetic energy of the electrons becomes very large. For example, the speed of the innermost electron in lead is around 40% of the speed of light. For heavy systems it is therefore not sufficient to use a nonrelativistic theory. Instead, the calculations are based on the relativistic counterpart to the Schrödinger equation, the Dirac equation. The original Dirac equation describes an electron moving in an external potential, provided, for example, by the nucleus. To describe many-electron systems, the interaction between the electrons has also to be specified. Nonrelativistically it is given by the Coulomb interaction. In Dirac theory the electronic spin is an integrated part of the equation, and this gives the electron a magnetic moment; that is, the electrons can also interact magnetically with each other. In addition, the interaction cannot be instantaneous, but has to propagate with the speed of light.

It is possible to construct a Hartree-Fock model based on the Dirac equation, where the electron-electron interaction is given by the Coulomb interaction, a magnetic contribution, and a term which corrects for the finite speed (retardation) with which the interaction propagates. This method is commonly used for heavy elements. Correlation can also be calculated in a relativistic framework with this expression for the electron-electron interaction. Here special problems arise, however, from the positron solutions (the positron is the electron's antiparticle) to the Dirac equation. It is important to separate electron and positron solutions in calculations to obtain reliable results. Practical solutions to this problem were found and implemented in calculations only during the 1980s. *See* ANTIMATTER; RELATIVISTIC QUANTUM THEORY.

Radiative corrections. With the measurement of the so-called Lamb shift in hydrogen in the 1940s, it became clear that there are additional effects not accounted for by the Dirac equation. These effects, which are called radiative corrections, were soon afterward shown to arise when the electromagnetic field was quantized within the theory of quantum electrodynamics. This theory, which combines special relativity (the Dirac equation) with the quantization of the electromagnetic field, has been very successful. It gives in principle a complete description of how to handle many-particle systems. In practice, however, a full quantum electrodynamics treatment of an atom becomes very cumbersome, and only a few calculations on two- and three-electron systems exist. The radiative corrections are, however, important for inner-shell electrons, and their effect on the interaction between an electron and a nucleus is well understood. For many-body systems, calculations of radiative effects are usually done within some independent-particle model, and the result is added to a correlated relativistic calculation based on the Dirac equation. *See* QUANTUM ELECTRODYNAMICS.

TABLE 2. Ionization energy of the innermost electron in lead

Type of contribution	Energy, eV
Relativistic Hartree-Fock model, Coulomb interaction*	88,498
Magnetic and retardation effects	−322
Relativistic correlation	4
Radiative contributions	−167

*A nonrelativistic calculation would account for only around 90% of this result.

As an example of a heavy system where both relativistic and radiative effects are important, **Table 2** shows the different contributions to the energy required to remove the innermost electron in lead. This quantity is not measured directly, but after the electron has been removed another electron will fall down into the hole left behind and an x-ray photon will be emitted. The least energetic photon is measured to have an energy of $74{,}970.01 \pm 0.17$ eV. The calculated value, which includes the ionization energy of the innermost electron, is close to that: 74,971 eV. *See* ATOMIC STRUCTURE AND SPECTRA.

Eva Lindroth

Bibliography. R. D. Cowan, *The Theory of Atomic Structure and Spectra*, Los Alamos Series in Basic and Applied Sciences, University of California, 1981; C. F. Froese Fischer, T. Brage, and P. Jönsson, *Computational Atomic Structure: An MCHF Approach*, IOP, 1997; L. Lindgren and J. Morrison, *Atomic Many-Body Theory*, Series on Atoms and Plasma, 2d ed., Springer-Verlag, Berlin, 1986.

Atomic time

Time that is based on physics, specifically on a unit of duration that corresponds to a defined number of periods of radiation of a specific atomic transition of a chosen isotope. The fundamental unit of atomic time is the International System (Système International or SI) second, which is defined as the duration of 9,192,631,770 periods of radiation corresponding to the transition between two hyperfine levels of the ground state of the cesium-133 atom. *See* ATOMIC CLOCK.

International Atomic Time. International Atomic Time (Temps Atomique International, or TAI) is determined by the Bureau International des Poids et Mesures (BIPM) in Sèvres, France from commercial atomic time standards and primary frequency standards in many countries. The time from the different

sources is compared by the Global Positioning System (GPS) and two-way satellite time transfer to determine the differences between the source time standards and a free running time scale at BIPM. The resulting TAI conforms as closely as possible to the definition of the SI second. TAI is steered to match the SI second; departures range $0.5–1.19 \times 10^{-14}$, with uncertainties of 0.2×10^{-14}.

TAI seeks to provide accuracy, long-term stability, and reliability. It was officially introduced in January 1972, but has been available since July 1955. TAI is the basis for a number of other time scales.

Terrestrial Time (TT). Terrestrial Time (TT) is the theoretical coordinate time whose mean rate is close to the mean rate of the proper time of an observer located on the geoid (the surface of constant gravity potential that extends the sea-level surface over the entire Earth). The unit of TT is the SI second. An accurate realization of TT is TT(TAI) = TAI + 32.184 seconds. *See* DYNAMICAL TIME.

Coordinated Universal Time (UTC). Since January 1, 1972, most, and now all, broadcast time services and civil times are based on the redefined Coordinated Universal Time (UTC), which is based on TAI and differs from TAI by an integer number of seconds. UT1 is based on the observationally determined rotational motion of the Earth, which conforms closely with the mean diurnal motion of the Sun. UTC is maintained within 0.9 s of UT1 by the introduction of 1-s steps (leap seconds) when necessary, usually at the end of December or June. The civil time around the world differs from UTC mostly by hour increments according to time zones. *See* EARTH ROTATION AND ORBITAL MOTION.

Satellite navigation. GPS and GLONASS (Global'naya Navigatsionnya Sputnikovaya Sistema) have atomic clocks on the satellites and use the accurate time and ephemerides of the satellites as a means of determining accurate positions on Earth. The satellite times are steered to match the rate of TAI and provide accurate sources of time. The GPS system is used for accurate time transfer and for the determination of the rotation and polar motion of the Earth. *See* SATELLITE NAVIGATION SYSTEMS.

Relativistic effects. The accuracy of atomic time requires that relativistic effects be taken into account. Hence, the definition of TAI is a coordinate time scale at what is known as the geocentric datum line, and having as its unit the SI second as obtained on the geoid in rotation. Then corrections to atomic clocks entering the formation of TAI must be made for gravitational potential and velocity and for the rotation of the Earth. Specifically, if there are two standard clocks at the same elevation on the Earth with respect to the geoid, the one closer to the Earth's equator will run slower than a clock closer to the pole, due to the greater rotational velocity near the equator. A standard clock at a lower elevation with respect to the geoid will run slower than a higher clock. *See* GEODESY; RELATIVITY.

Atomic clocks. There have been a large number of commercial cesium standards and a small number of primary cesium frequency standards available as the basis of atomic time. Since the 1990s, cesium fountains have been developed based on the principles of laser cooling of the cesium atoms. These cesium fountains (as of 2005, about five exist) have provided an order-of-magnitude improvement in accuracy. Optical frequency standards are now under investigation. These are based on optical atomic transitions, at optical rather than microwave frequencies, so they operate with a much smaller unit of time and can possibly achieve as much as five orders of magnitude improvement in accuracy. As of 2005, cesium atomic clocks achieve timing uncertainties of about 35 picoseconds per day, and optical atomic clocks may have uncertainties of 100 femtoseconds per day.

Due to the technology of atomic clocks, the SI second is the most accurately measured unit. Hence, the meter is now defined as the distance that light travels in 1/299,792,458 of a second in vacuum, rather than as a physical bar of metal. *See* LASER COOLING; LENGTH; PHYSICAL MEASUREMENT.

History. The civil time scales were defined based on the mean solar time until it was recognized that the rotation of the Earth was irregular. In the 1950s Ephemeris Time was introduced as a new time scale based on the apparent motion of the Sun. The ephemeris second was the new unit of time. However, it was difficult to determine Ephemeris Time from observations of the Sun, so observations of the Moon, which were more accurate, were used. Also, the definition of the ephemeris second was based on Simon Newcomb's theory of the motion of the Earth, which was fit to observations from the 1800s. So the ephemeris second matched the rate of rotation of the Earth from about 1850.

In the 1950s observations of the Moon were used to determine the atomic second with respect to the ephemeris second. The definition of the atomic second was given to agree with the ephemeris second. That is the definition of the SI second used today and the reason that atomic time does not match the current rate of rotation of the Earth.

P. Kenneth Seidelmann

Bibliography. S. A. Diddams et al. (2004), Standards of time and frequency at the outset of the 21st century, *Science*, 306:1318–1324, 2004; W. Markowitz et al., Frequency of cesium in terms of Ephemeris Time, *Phys. Rev. Lett.* 1:105–107, 1958; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, University Science Books, Mill Valley, California, 1992.

Atomization

The process whereby a bulk liquid is transformed into a multiplicity of small drops. This transformation, often called primary atomization, proceeds through the formation of disturbances on the surface of the bulk liquid, followed by their amplification due to energy and momentum transfer from the surrounding gas.

Spray formation processes are critical to the performance of a number of technologies and applications. These include combustion systems (gas turbine engines, internal combustion engines, incinerators, furnaces, rocket motors), agriculture (pesticide and fertilizer treatments), paints and coatings (furniture, automobiles), consumer products (cleaners, personal care products), fire suppression systems, spray cooling (materials processing, computer chip cooling), medicinals, (pharmaceutical), and spray drying (foods, drugs, materials processing). Current concerns include how to make smaller drops (especially for internal combustion engines), how to make larger drops (agricultural sprays), how to reduce the number of largest and smallest drops (paints and coatings, consumer products, medicinals, spray drying), how to distribute the liquid mass more uniformly throughout the spray, and how to increase the fraction of liquid that impacts a target (paints and coatings, spray cooling, fire suppression).

Spray devices (that is, atomizers) are often characterized by how disturbances form. The most general distinction is between systems where one or two fluids flow through the atomizer. The most common types of single-fluid atomizers (**Fig. 1**) are pressure

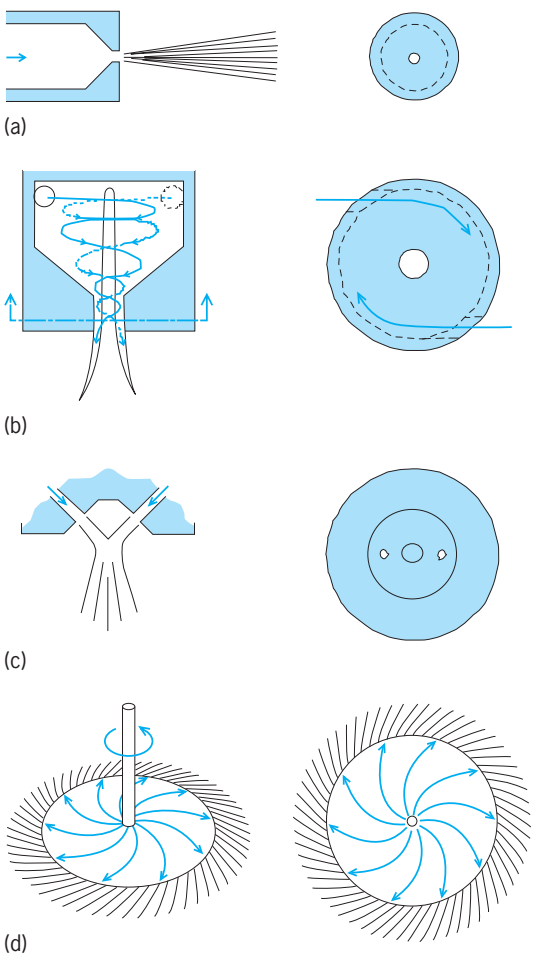

Fig. 1. Single-fluid atomizers. Diagrams on right are cross sections. (*a*) Pressure atomizer. (*b*) Pressure-swirl atomizer. (*c*) Impinging-jet atomizer. (*d*) Rotary atomizer.

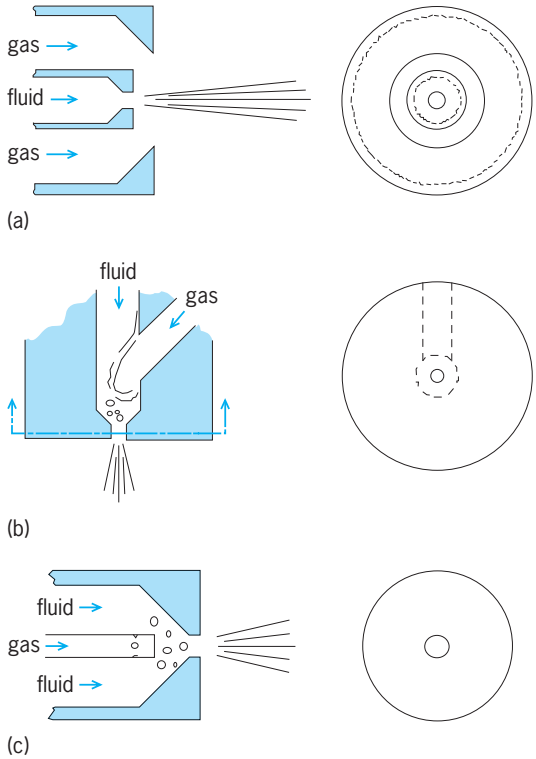

Fig. 2. Twin-fluid atomizers. Diagrams on right are cross sections. (*a*) External-mix atomizer. (*b*) Y-jet internal-mix atomizer. (*c*) Effervescent internal-mix atomizer.

(also called plain-orifice, hydraulic, or pneumatic), pressure-swirl, rotary, ultrasonic (sometimes termed whistle or acoustic), and electrostatic; other types include impinging jet, fan, and jet-swirl. Twin-fluid atomizers (**Fig. 2**) include internal-mix and external-mix versions, where these terms describe the location where atomizing fluid (almost always a gas) first contacts fluid to be sprayed (almost always a liquid).

While primary atomization is important, because of its role in determining mean drop size and the spectrum of drop sizes, subsequent processes also play key roles in spray behavior. They include further drop breakup (termed secondary atomization), drop transport to and impact on a target, drop evaporation (and perhaps combustion), plus drop collisions and coalescence. In addition, the spray interacts with its surroundings, being modified by the adjacent gas flow and modifying it in turn.

Basic processes. The spray formation process actually starts inside the atomizer itself. It is here that the fluid flow assumes its laminar or turbulent state. The latter is beneficial to forming small drops because precursors to fluid-mechanic instabilities are naturally present in turbulent flows. *See* LAMINAR FLOW; TURBULENT FLOW.

Primary atomization. Sheet or jet breakup resulting from fluid-mechanic instabilities is termed primary atomization. It depends on process operating conditions, fluid physical properties, and atomizer geometry.

Process operating conditions dictate the relative velocity between the fluid to be atomized and any surrounding gas. For single-fluid atomizers, fluid

supply pressure determines the velocity at which fluid exits the atomizer into the (assumed) quiescent environment. For twin-fluid atomizers, the keys are the fluid and atomizing gas velocities. In all cases, it has been established that a larger relative velocity will accelerate the growth of instabilities and lead to smaller drops.

The fluid physical properties that most strongly influence spray formation are viscosity and surface tension. Density usually plays a smaller role; an increase in density usually leads to a small decrease in drop size.

An increase in shear (or dynamic) viscosity always leads to an increase in droplet size because viscosity directly opposes instability growth. If viscosity is high (say, above 100 centipoise or 0.1 pascal second), some types of atomizers (including pressure and pressure-swirl) can no longer form sprays without the use of excessively high fluid supply pressures. *See* VISCOSITY.

The process is more complex if the fluid is non-newtonian, especially if viscoelastic. The elastic contribution is especially efficient at damping out instabilities, thereby inhibiting breakup. Non-newtonian behavior is prevalent in paints and coatings, pharmaceutical and medicinal sprays, consumer product sprays, and adhesives. A practical result is that single-fluid atomizers cannot form sprays of reasonable quality from viscoelastic fluids. *See* NON-NEWTONIAN FLUID.

An increase in surface tension can enhance or attenuate instability growth. For primary atomization, increasing surface tension leads to smaller drops at low relative velocity. High relative velocity causes an increase in surface tension to lead to larger drops. When spray formation is dominated by secondary atomization, increasing surface tension always leads to larger drops. *See* SURFACE TENSION.

Atomizer geometry affects spray formation through the initial configuration of the fluid to be sprayed. Fluid exiting the atomizer is often spread into thin sheets or filaments to enhance the breakup process since smaller drops result from thinner sheets and smaller-diameter jets. Sheets can be flat, as in fan sprays, or annular, as in pressure-swirl sprays and prefilming twin-fluid sprays (where the liquid is spread into a thin sheet prior to its contact with the atomizing air). *See* JET FLOW.

Secondary atomization. If the relative velocity between an already formed drop and its surrounding gas is sufficiently high, the drop may undergo breakup, that is, secondary atomization. Its importance is evaluated by the drop Weber number, a dimensionless group representing the competition between aerodynamic distortion of the drop and restorative surface tension, given by Eq. (1). Here ρ_{gas} is the mass

$$\mathrm{We} = \rho_{gas} V_{rel}^2 d/\sigma \qquad (1)$$

density of the surrounding gas, V_{rel} is relative velocity between the drop and the surrounding gas, d is drop diameter, and σ is surface tension. Values of the Weber number greater than 12 indicate that secondary atomization may be present. Higher values indicate that drops break up via the bag, boundary-layer stripping, and catastrophic modes (**Fig. 3**). *See* DIMENSIONLESS GROUPS.

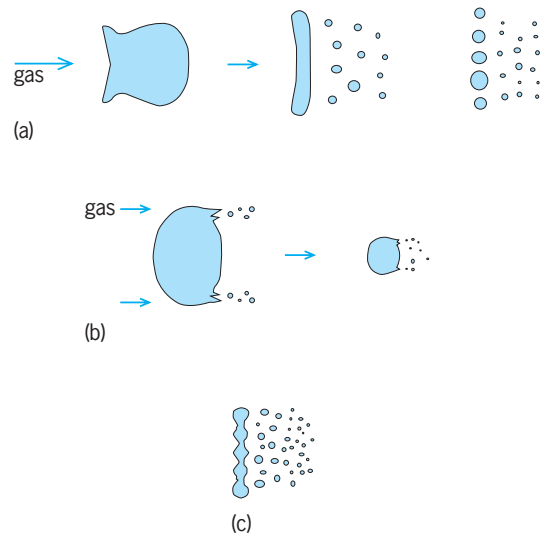

Fig. 3. Secondary atomization mechanisms. (*a*) Bag mode. (*b*) Boundary-layer stripping mode. (*c*) Catastrophic mode.

Drop transport. Droplet transport, or motion, is determined from the forces that action the drop using Newton's second law. Aerodynamic drag, F_{drag}, is usually the most important force. It is given by Eq. (2), where C_{drag} is the drag coefficient of the

$$F_{drag} = (1/2)\rho_{gas} C_{drag} (\pi/4) d^2 V_{rel}^2 \qquad (2)$$

drop. Since drops deform when moving, the drag coefficient depends on the drop's Reynolds number. Because droplet mass depends on diameter cubed while droplet drag force depends on diameter squared, large drops subject to aerodynamic drag retain their velocities relative to the surrounding gas much longer than smaller drops. *See* AERODYNAMIC FORCE; REYNOLDS NUMBER.

Additional forces of interest include gravity, when drops are moving slowly (for example, "drop-out" in consumer product sprays), and electrostatic, when significant electric fields or drop charging are present (automotive paint sprays, for instance). Electrostatic forces can be due to imposed external fields, or to fields resulting from the charged drops themselves. *See* ELECTROSTATICS.

Drop evaporation. Evaporation, along with primary and secondary atomization and coalescence, is a mechanism for changing drop size. Evaporation is dominated by energy transfer from the surrounding gas that goes into increasing the drop sensible and latent enthalpies. Both steady-state and transient (that is, drop heat-up) behavior can be important.

Analytical models are available to predict drop diameter as a function of time. The simplest assumes that all energy addition to the drop goes into latent enthalpy (that is, results directly in evaporation), the process is at steady-state, the drop is not affected by other drops, the process is spherically symmetric,

the interior temperature is uniform throughout the drop, the drop is stationary, and the gas-phase conditions are not changed by evaporation. The d^2-law, Eq. (3), results, where d_0^2 is the square of the droplet

$$d^2 = d_0^2 - \lambda_{evap} t \tag{3}$$

diameter when evaporation begins, t is the time since evaporation started, and λ_{evap} is the evaporation constant. More sophisticated models predict transient heat-up behavior and the influence of convective heat transfer when the droplet is in motion. Nonuniform temperatures and internal circulatory flows within the drop, as well as drop-to-drop interactions, have been considered. *See* ENTHALPY; EVAPORATION; PSYCHROMETRICS.

Drop collisions and coalescence. Drop collisions are rare in most parts of a spray. The exceptions are regions where the number of drops per unit volume is high, or where drops have widely differing velocities. Examples are immediately adjacent to the atomizer exit orifice, where the spray has not had time to spread, and near a target where drops are impinging on a surface and (perhaps) rebounding.

The importance of collisions can be assessed through a model that treats drops as rigid spheres. The probability of a collision is related to the square of the average diameter of the two drops, their relative velocity, and the number of drops per unit volume for both colliders.

Drop coalescence leads to an increase in drop diameter. However, not all droplet collisions result in coalescence. Colliding drops must have sufficient kinetic energy to overcome surface tension at impact, but not so much that they rebound. An increase in drop viscosity usually increases the likelihood of a collision resulting in coalescence.

Spray-surroundings interactions. A spray will interact with its surroundings as it moves away from the atomizer. This interaction changes spray cone angle, influences how far the spray penetrates into the surroundings, controls how much surrounding air is drawn into the spray, impacts mean drop velocity and the spread in droplet velocities, and varies the distribution of droplet mass throughout the spray cone.

There is a noticeable decrease in cone angle, or turning-in of the spray boundary. This results from the entrainment of surrounding gas which drives drops (especially the smallest ones) toward the spray centerline (**Fig. 4**).

The spray-surroundings interaction also limits how far a spray penetrates into the surrounding gas. Diesel engine, gas turbine, and consumer product sprays are applications where penetration is important. Droplet average velocity also depends on spray-surroundings interactions. It usually decreases with distance from the atomizer because the spray most often exits into a stagnant environment so that aerodynamic drag decelerates the droplets.

Spray-surroundings interactions also modify the distribution of fluid mass throughout the spray. For a hollow-cone spray (**Fig. 5**), such as that produced by

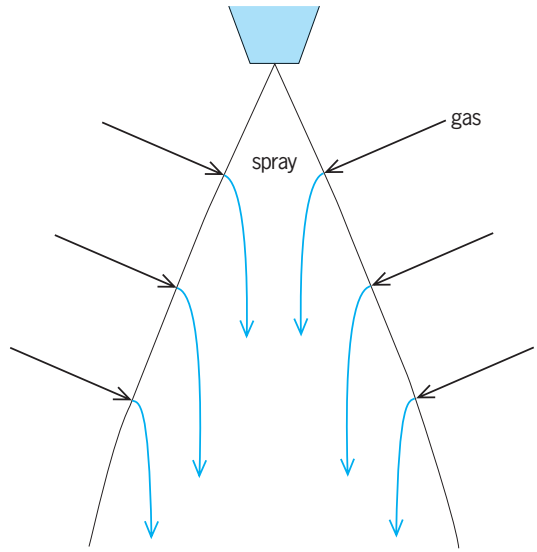

Fig. 4. Entrainment of surrounding air by a spray.

a pressure-swirl atomizer, entrainment helps to homogenize the fluid mass distribution by "filling in" the spray core. Furthermore, the size dependence of droplet motion means that larger drops are more likely to penetrate toward the edges of the spray while smaller drops are more likely to be dragged toward the interior of the spray by inbound gas, so that mean drop size increases with distance from the spray axis.

Single-fluid atomizers. Pressure and pressure-swirl atomizers require only a source of liquid at elevated pressure. Pressure atomizers (Fig. 1*a*) are found in diesel engines, aircraft engine afterburners, and some fire-suppression sprays. Their mean drop size performance is more sensitive to fluid viscosity than other types. In addition, fluid supply pressure variations can have such a large impact on atomizer performance that only a modest turn-down ratio (the ratio of the largest to the smallest liquid mass flow rates) is obtainable. Finally, pressure atomizers tend to produce sprays having narrow cone angles (half-angles of the order of 10°), so they are poor choices

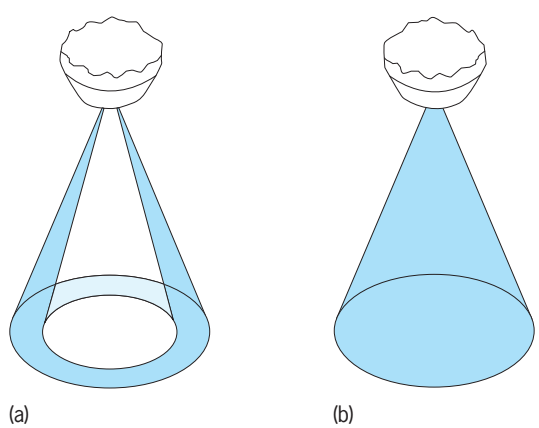

Fig. 5. Atomizer sprays. (*a*) Hollow-cone spray. (*b*) Solid-cone spray.

in applications where mixing the spray into its surroundings is important.

Pressure-swirl atomizers (Fig. 1*b*) produce sprays having larger, and controllable, cone angles. Values as high as 150° can be achieved. Modifications made to enhance the limited turn-down ratio characteristic of typical pressure-swirl atomizers have resulted in the dual-orifice (or duple), spill-return, and duplex designs. Pressure-swirl atomizers cannot produce reasonable quality sprays when the fluid viscosity climbs above about 100 cP (0.1 Pa · s).

Twin-fluid atomizers. The performance of twin-fluid atomizers tends to be less sensitive to fluid physical properties because they usually rely on secondary atomization. Secondary atomization depends somewhat on fluid surface tension, and to a lesser extent on viscosity and density.

Air-assist and air-blast atomizers are examples of external-mix units (Fig. 2*a*). Both can be used with highly viscous fluids. Their natural cone angles approach those of pressure atomizers, but can be extended using a mechanical deflection device called a pintle. Both types promote mixing between the fluid to be atomized and the surrounding gas, so they are desirable for combustion applications. An air-assist atomizer employs a smaller quantity of air than an air-blast unit, but at a higher velocity, and may require an external compressor for the atomizing gas supply.

Internal-mix designs include the Y-jet (Fig. 2*b*), effervescent (Fig. 2*c*), dissolved-gas, and flash atomizers. Effervescent atomizers can atomize highly viscous, and viscoelastic, fluids using a minimum amount of atomizing gas and large-diameter orifices. These energy-efficient devices form gas bubbles inside the atomizer body and often rely on primary atomization. Their natural cone angles are similar to those of pressure atomizers, but can be extended with mechanical deflection devices. Effervescent atomizers require an external supply of atomizing gas, but are relatively immune to clogging and plugging.

Dissolved-gas and flashing atomizers also form bubbles inside the atomizer body. They differ from effervescent types in their need for some combination of mass (dissolved gas) and energy transfer (flashing), in addition to bubble nucleation sites. This requirement can make them quite large. Moreover, it is necessary to dissolve gas into the fluid to be atomized, or mix in a high-vapor-pressure liquid, but an external compressor is unnecessary.

Measurement techniques. There are two principal methods for measuring drop sizes. Instruments based on forward light scattering (that is, diffraction) return information based on the number of drops per unit volume located along a line that crosses the spray. Instruments based on interference and refraction return information based on the rate of drops passing through a unit area located at a point in the spray. Diffraction instruments allow sprays to be characterized rapidly, but do not provide spatially precise information. Both techniques can interrogate transient sprays, and both provide the spray's mean drop size and the range of drop sizes produced. *See* DIFFRACTION; INTERFERENCE OF WAVES; REFRACTION OF WAVES; SCATTERING OF ELECTROMAGNETIC RADIATION.

Drop velocities are usually measured using interference-refraction instruments. The principles of operation are the same as those for laser Doppler velocimetry. Mean and root-mean-square (rms) velocity information is provided. *See* ANEMOMETER; FLOW MEASUREMENT.

Spray cone angles are usually measured via imaging. Charge-coupled-device (CCD) cameras, 35-mm cameras, movie cameras (both film and camcorders), and holography have been employed. Illumination sources include incandescent lamps, strobe lights, short-duration flashes, and arc lights. This same imaging apparatus can often be used to determine spray penetration.

Mass distribution throughout the spray (patternation) can be measured using either intrusive mechanical devices or nonintrusive optical devices. *See* AEROSOL; PARTICULATES. Paul E. Sojka

Bibliography. L. Bayvel and Z. Orzechowski, *Liquid Atomization*, Taylor and Francis, New York, 1993; A. H. Lefebvre, *Atomization and Sprays*, Hemisphere, New York, 1989; M. T. Lund, P. E. Sojka, and A. H. Lefebvre, Effervescent atomization at low mass flow rates, Part I: The influence of surface tension, *Atomiz. Sprays*, 3(1):77–89, 1993; S. D. Sovani, P. E. Sojka, and Y. R. Sivathanu, Prediction of drop size distributions from first principles: The influence of fluctuations in relative velocity and liquid physical properties, *Atomiz. Sprays*, 9(2):133–152, 1999.

Atropine

An alkaloid, $C_{17}H_{23}NO_3$, with the chemical structure below. The systematic chemical name is endo(±)-α-(hydroxymethyl)phenylacetic acid 8-methyl-8-azabicyclo[3.2.1]oct-3-yl ester, and in phamacy it is

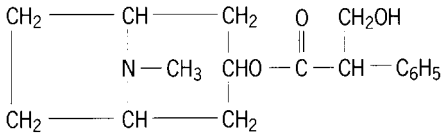

sometimes known as *dl*-hyoscyamine. It occurs in minute amounts in the leaves of *Atropa belladonna, A. betica, Datura stramonium, D. innoxia, and D. sanguinea* as well as many related plants. It is chiefly manufactured by racemization of *l*-hyoscyamine, which is isolated from the leaves and stems of the henbane, *Hyoscyamus niger*. It melts at 114–116°C (237–241°F) and is poorly soluble in water. The nitrate and sulfate are used in medicine instead of the free base.

Atropine is used clinically as a mydriatic (pupil dilator). Dilation is produced by paralyzing the iris and ciliary muscles. Atropine is also administered in small doses before general anesthesia to lessen oral and air-passage secretions. Its ability to reduce these secretions is also utilized in several

preparations commonly used for symptomatic relief of colds. *See* ALKALOID. Frank Wagner

Attention deficit hyperactivity disorder

A common psychiatric disorder of childhood characterized by attentional difficulties, impulsivity, and hyperactivity; known earlier as attention deficit disorder. A disorder or a syndrome comprises a group of signs and symptoms which tend to cluster together (symptoms refer to what the patient experiences; signs refer to what the physician observes). Other older names for this disorder include minimal brain dysfunction, minimal brain damage, hyperactivity, hyperkinesis, and hyperactive child syndrome. Over time, these names were modified due to their implications about etiology and core symptoms: minimal brain dysfunction seemed to imply that children with this disorder were brain-damaged, while hyperactivity and its synonyms named a feature seen in many but not all of these children. Recent studies indicate that attention deficit hyperactivity disorder tends to follow a chronic course, persisting in many cases not only into adolescence but into adulthood.

Clinical features. Some symptoms and signs may be present in all instances of the disorder, while others may occur in varying combinations or may not be present.

The three defining symptoms of attention deficit disorder are as follows: (1) *Attentional deficits.* The child is described as having a short attention span. Lacking "stick-to-itiveness," the child often fails to finish things he or she starts, does not seem to listen, and is easily distracted or disorganized. In more severe instances the child is unable to focus attention on anything, while in less severe cases attention can be focused on things of interest to the child. For example, the child may studiously absorb information about dinosaurs or tornadoes but ignore improper fractions. (2) *Impulsivity.* The child is often described as acting before thinking, shifting excessively and rapidly from one activity to another, or having difficulty waiting for a turn in games or group activities. (3) *Hyperactivity.* Many children with this disorder are hyperactive, and indeed, may have been noted to be so prior to birth. They are often described as always on the go or as driven by a motor. They may fidget, wiggle, move excessively, and have difficulty keeping still. This excessive activity is not noticeable when the children are playing. At such times, the child may be no more active than his or her peers; however, in the classroom or other quiet settings, the child cannot decrease his or her activity appropriately. Some affected children are active at a normal level or even sluggish (hence the inapplicability of the term "hyperactive" to a subset of individuals with attention deficit hyperactivity disorder); they have been studied less adequately than the hyperactive ones. On the basis of the predominating symptoms, children with attention deficit hyperactivity disorder are subcategorized as having hyperactive symptoms (hyperactive type), lacking hyperactivity (inattentive type), and having both inattention and hyperactivity or impulsivity (combined type).

Many children with attention deficit hyperactivity disorder frequently show an altered response to socialization. They are often described by their parents as obstinate, impervious, stubborn, or negativistic. With peers, many affected children are domineering or bullying, and thus may prefer to play with younger children. Another characteristic often seen in children with the disorder is emotional lability. Their moods change frequently and easily, sometimes spontaneously, and sometimes reactively so that the child is upset, sad, or angry.

Because of their behavioral difficulties, children with the disorder often have conflicts with parents, teachers, and peers. The child's behavior may engender friction with others. While this friction is sometimes used to explain the child's problems, more accurately it is a result of the child's disorder. Commonly, difficulties in discipline and inadequacies in schoolwork also lead to reproof and criticism. As a consequence, children with the disorder usually also have low self-esteem.

Associated disorders. Attention deficit hyperactivity disorder is frequently associated with other disorders, including disruptive behavior disorders, mood and anxiety disorders, and developmental disorders. The disruptive behavior disorders include oppositional defiant disorder and conduct disorder. Oppositional defiant disorder is characterized by stubbornness, defiance, and negativistic behavior. In conduct disorder, there is frequent recurrent violation of important rules at home or at school. There is often physical violence, lack of close relationships with others, and diminished or markedly reduced conscience. Specific learning disabilities represent a group of conditions (1) in which performance in psychometric testing in reading, spelling, mathematics, coordination, and other skills is below that which would be predicted on the basis of the child's intelligence, and (2) which is not explicable in terms of inadequate educational experience. In the case of reading and spelling, the disorder is often called dyslexia. When changes in mood are persistent and severe, the term "mood disorders" is applied; typical forms of mood disorders include depression and mania, sometimes known as bipolar disorder. Anxiety disorders refer to a group of conditions characterized by fearfulness and avoidance. Additional behavioral, developmental, and emotional difficulties are termed comorbidity. *See* AFFECTIVE DISORDERS.

Prevalence. It is estimated that 5–10% of children of elementary school age (roughly 6–19 years) manifest significant symptoms of attention deficit hyperactivity disorder. Moreover, it is estimated that the disorder may affect as many as 2–4% of adults. About twice as many boys as girls are affected with the disorder. The girls are much less likely than the boys to be aggressive and have serious behavioral difficulties, making the girls vulnerable to underidentification and undertreatment.

Causation. Formerly believed to be largely caused by brain damage, and more recently believed by some to be caused by food allergy, attention deficit hyperactivity disorder is now considered to be mainly hereditary. Supporting evidence is as follows: (1) The disorder runs in families. (2) In studies of adopted children with the disorder, psychiatric illnesses are found among biological parents rather than among adoptive parents. (3) Twin studies of the disorder document greater resemblance of monozygotic, or genetically identical, twins as compared to dizygotic twins, who share, on average, half of their varying genes. These twin studies suggest that attention deficit hyperactivity disorder is highly heritable. (4) Several investigators identified genes in the dopamine family as associated with this condition. Some researchers in biological psychiatry have suggested that many of the symptoms of the disorder are produced by an inherited metabolic abnormality in the brain. This abnormality is hypothesized to be underactivity in certain nerve tracts that is produced by deficient synthesis, release, or sensitivity to released neurotransmitters. (5) Not only do the parents (primarily the fathers) of affected children recount a history of the disorder in childhood, but there seems to be an increased frequency of other psychiatric disorders among relatives of children with the disorder. These other disorders include alcoholism, drug addiction, antisocial personality, and mood and anxiety disorders as well as learning and cognitive deficits.

Natural history. Formerly, it was believed that attention deficit hyperactivity disorder was outgrown during adolescence. Although some signs of the disorder such as excessive activity may diminish or disappear in some affected children, other signs such as attentional difficulties, impulsivity, and interpersonal problems may persist. It is now recognized that in about half of children with attention deficit hyperactivity disorder the disorder will persist into adulthood. Despite the fact that this disorder is common in adults, the lower rates of hyperactivity in adults may result in the condition being frequently overlooked.

Treatment. The treatment of the child or adult with this disorder involves three steps: evaluation, explanation of the problem to parents and child, and therapeutic intervention.

Evaluation requires a detailed history of the individual's psychological development and current functioning. When evaluating children, assessments commonly include information obtained from both parents and teachers and an interview with the child. Next, because the disorder is frequently associated with learning problems in school, it is desirable to obtain an individual intelligence test (such as the Wechsler Intelligence Scale for Children) as well as a test of academic achievement. Since attention deficit hyperactivity disorder is often associated with other psychiatric disorders, it is important to carefully evaluate the presence of these other conditions in the affected child, adolescent, or adult. If a diagnosis of attention deficit hyperactivity disorder is confirmed, the parents or family should be educated regarding the nature of the condition and associated conditions.

Medication and guidance are the mainstays of the treatment. Approximately 60–70% of children and adults manifest a therapeutic response to one of the major stimulant drugs, such as amphetamines and methylphenidate. A new generation of reformulated highly effective, once-daily, long-acting stimulants has emerged that has greatly facilitated the management of individuals with attention deficit hyperactivity disorder. Recently, a new nonstimulant medication (automoxetine) received FDA approval for the treatment of children and adults with attention deficit hyperactivity disorder. When effective, these medications increase attention, decrease impulsivity, and usually make the child more receptive to parental and educational requests and demands. Hyperactivity, when present, is usually diminished as well. These medications do not produce a "chemical straitjacket" but diminish the child's feeling of being driven. Although usually less effective (and, thus, second-line treatments), other medications can be helpful to individuals who cannot tolerate or do not respond to stimulants. The common mechanism of action for such medications is their impact upon the neurotransmitters dopamine and norepinephrine.

Because stimulants can be abused, concern about these medications has been expressed by the public. Yet, euphoria rarely occurs either in children or in adults with attention deficit hyperactivity disorder treated with adequate doses under medical supervision. Recent findings have emerged documenting that medication management of children with attention deficit hyperactivity disorder exerts a protective effect against the development of problems with drugs and alcohol in adolescence. Since medication is symptomatic rather than curative (that is, it only reduces the symptoms of the disorder), it must be used as long as the symptoms persist. Medication must be used in conjunction with concrete, specific, and flexible guidance techniques. Medication does not, however, improve specific developmental disorders (such as dyslexia) or coordination problems. Many children with attention deficit hyperactivity disorder alone or with associated disorders may require some form of remedial education since their behavioral problems have interfered with learning.

Joseph Biederman

Bibliography. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed., American Psychiatric Association, Washington, DC, 1994; J. Biederman et al., Gender differences in a sample of adults with attention deficit hyperactivity disorder, *Psych. Res.*, 53:13–29, 1994; J. Biederman, J. Newcorn, and S. Sprich, Comorbidity of attention deficit hyperactivity disorder with conduct, depressive, anxiety, and other disorders, *Amer. J. Psych.*, 148:564–577, 1991; S. V. Faraone et al., Attention deficit hyperactivity disorder in adults: An overview, *Biol. Psych.*, 48:9–20, 2000; S. V. Faraone and J. Biederman, The neurobiology of attention deficit hyperactivity disorder, in D. S. Charney, E. J. Nestler, and B. S. Bunney (eds.), *Neurobiology of Mental Illness*, pp. 788–801, Oxford,

New York, 1999; L. Goldman et al., Diagnosis and treatment of attention-deficit/hyperactivity disorder in children and adolescents, *JAMA*, 279:1100–1107, 1998; T. Spencer, J. Biederman, and T. Wilens, Pharmacotherapy of ADHD: A life span perspective, in *American Psychiatric Press Review of Psychiatry*, pp. 87–128, American Psychiatric Association, Washington, D.C., 1997; T. Wilens, Does the medicating ADHD increase or decrease the risk for later substance abuse?, *Rev. Brasil. Psychiat.*, 25:127–128, 2003; T. Wilens et al., Does stimulant therapy of attention deficit hyperactivity disorder beget later substance abuse? A meta-analytic review of the literature, *Pediatrics*, 111:179–185, 2003.

Attenuation

The reduction in level of a transmitted quantity as a function of a parameter, usually distance. It is applied mainly to acoustic or electromagnetic waves and is expressed as the ratio of power densities. Various mechanisms can give rise to attenuation. Among the most important are geometrical attenuation, absorption, and scattering.

For unconfined radiation from a point source in free space, the power density (watts per square meter) decreases in proportion to the square of the distance, for purely geometrical reasons. The power densities, I_1 and I_2, at distances r_1 and r_2 from the source, are related by Eq. (1).

$$I_2 = I_1 \left(\frac{r_1}{r_2}\right)^2 \tag{1}$$

See INVERSE-SQUARE LAW.

If the signal, in a parallel beam so that there is no geometrical attenuation, passes through a lossy medium, absorption reduces the power level, I, exponentially with distance, x, according to Eq. (2), where a is the attenuation coefficient.

$$I(x) = I(0)e^{-ax} \tag{2}$$

See ABSORPTION; ABSORPTION OF ELECTROMAGNETIC RADIATION; SOUND ABSORPTION.

Scattering is said to occur if the power is not absorbed in the medium but scattered from inhomogeneities. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

More complicated situations occur with guided waves, such as acoustic waves in pipes or electromagnetic waves in transmission lines or waveguides, where absorption may take place and irregularities may cause reflection of some power. *See* TRANSMISSION LINES; WAVEGUIDE.

In electric circuits, constituent elements are often described as attenuators when they reduce the level of signals passing through them. *See* ATTENUATION (ELECTRICITY).

Attenuation is usually measured in terms of the logarithm of the power ratio, the units being the neper or the decibel. *See* DECIBEL; NEPER. A. E. Bailey

Attenuation (electricity)

The exponential decrease with distance in the amplitude of an electrical signal traveling along a very long uniform transmission line, due to conductor and dielectric losses. If the peak voltage at the sending end of the transmission line is denoted by V_0, the peak voltage at a distance x from the sending end is given by Eq. (1), where α is the attenuation constant of the line.

$$V_x = V_0 e^{-\alpha x} \tag{1}$$

Characteristic impedance. Transmission-line theory shows that the input impedance of such a line has a purely resistive value Z_0 at high frequencies, where Z_0 is the characteristic impedance of the line. If the line is terminated at any point by a resistance of value Z_0, all of the power that reaches the termination will be dissipated and the line is said to be matched. If the line is terminated by any other impedance Z_t, part of the incident voltage will be reflected and the reflection coefficient Γ of the termination is defined as the ratio of the reflected voltage to the incident voltage. The value of Γ is shown by Eq. (2). Thus,

$$\Gamma = \frac{Z_t - Z_0}{Z_t + Z_0} \tag{2}$$

when $Z_t = 0$, $\Gamma = -1$; when $Z_t = 3Z_0$, $\Gamma = 0.5$; and when Z_t approaches infinity, Γ approaches $+1$. *See* ELECTRICAL IMPEDANCE; IMPEDANCE MATCHING; REFLECTION AND TRANSMISSION COEFFICIENTS.

Neper and decibel. Electronic engineers usually express power ratios in a logarithmic manner, using either nepers or decibels (dB) as in Eqs. (3) and (4),

$$\text{Power ratio in nepers} = {}^1/_2 \log_e \frac{P_1}{P_2} \tag{3}$$

$$\text{Power ratio in decibels} = 10 \log_{10} \frac{P_1}{P_2} \tag{4}$$

where P_1 and P_2 are the power levels to be compared. The advantages of using the decibel or neper, instead of just P_1/P_2, are that very large power ratios can be expressed by conveniently small numbers, and the overall gain or loss of a cascade of matched networks can be found by addition and subtraction instead of multiplication. *See* DECIBEL; NEPER.

The well-known relationship between napierian and common logarithms yields Eq. (5).

$$\text{Power ratio in decibels} = 8.686 \times \text{power ratio in nepers} \tag{5}$$

Insertion loss. When a generator with a reflection coefficient Γ_G is connected directly to a load with a reflection coefficient Γ_L, let the power dissipated in the load be denoted by P_3. Suppose now that a network, with two connectors called ports, is inserted between the generator and the load, and let this reduce the power dissipated in the load to P_4. Then,

the insertion loss of this two-port network is given in decibels by Eq. (6).

$$L = 10 \log_{10} \frac{P_3}{P_4} \tag{6}$$

Definition of attenuation. The attenuation of the two-port network is defined as the insertion loss when both the source and load are perfectly matched, that is, when $\Gamma_G = 0$ and $\Gamma_L = 0$. Insertion loss depends on the values of Γ_G and Γ_L as well as on the characteristics of the two-port network, whereas attenuation is a property only of the two-port network.

Returning now to the situation depicted by Eq. (1), the power entering the line is $(V_0)^2/2Z_0$, and the power at a distance x from the sending end is $(V_x)^2/2Z_0$; so, using Eq. (3), the attenuation in nepers over the distance x is seen to be given by Eq. (7).

$$A = {}^1/_2 \log_e \frac{(V_0)^2/2Z_0}{(V_x)^2/2Z_0} = {}^1/_2 \log_e(e^{2\alpha x}) = \alpha x \tag{7}$$

Thus, α gives the attenuation of the line in nepers per unit length; attenuation in decibels per unit length is seen from Eq. (5) to be 8.686α.

Types of attenuators. Attenuators find numerous applications, typical examples being: in a signal generator, to vary the amplitude of the output signal; and in the input line to a television receiver that is very close to a television transmitter, so that overloading can be avoided. *See* SIGNAL GENERATOR; TELEVISION RECEIVER.

Attenuators for the dc (steady voltage) to very high frequency (VHF) range (frequencies from 0 to 300 MHz) often contain resistors arranged in T or π configurations (**Fig. 1***a*). The resistor values are chosen so that the required attenuation is obtained without changing the matching conditions. When a single T section is required to give an attenuation of A_t decibels and be matched to input and output lines of characteristic impedance Z_0, the design equations for the resistors R_1 and R_2 are Eqs. (8) and (9), where k is given by Eq. (10). Thus, when $Z_0 = 50\ \Omega$ and an

$$R_1 = Z_0 \frac{1-k}{1+k} \tag{8}$$

$$R_2 = 2Z_0 \frac{k}{1-k^2} \tag{9}$$

$$k = 10^{-A_t/20} \tag{10}$$

attenuation of 10 dB is needed, $R_1 = 25.975\ \Omega$ and $R_2 = 35.136\ \Omega$.

Fixed coaxial attenuators can be designed to operate from dc to frequencies of more than 40 GHz. Various constructional techniques are used; for example, a thin resistive film can be deposited on the inner conductor, or a resistive film can be deposited on a substrate that is furnished with suitable launching, receiving, and earth electrodes.

Piston attenuators (sometimes called waveguide-beyond-cutoff attenuators) are used at both intermediate and microwave frequencies (Fig. 1*b*). The attenuation is varied by altering the separation between the two coils. The circular tube acts as a waveguide beyond cutoff, and the launching system is designed so that only one mode is excited in it. Assuming that the inner surface of the tube has perfect conductivity, an increase in the coil separation from x_1 to x_2 produces an attenuation change in decibels given by Eq. (11), where r is the radius of the cylinder, λ is

$$A_p = 8.686 \times 2\pi(x_2 - x_1)\left[\left(\frac{S_{\mathrm{nm}}}{2\pi r}\right)^2 - \frac{1}{\lambda^2}\right]^{1/2} \tag{11}$$

the free-space wavelength of the applied signal, and S_{nm} is a constant which depends upon the mode that is excited. The TE_{11} mode is usually employed, and in this case $S_{\mathrm{nm}} = 1.84118$. A piston attenuator has a very high dynamic range and a linear decibel scale. When great care is taken in the design and manufacture, an accuracy as high as 0.0002 dB per 10 dB can be achieved over a 120-dB range. *See* WAVEGUIDE.

A variable waveguide attenuator can be produced by moving a lossy vane either sideways across the waveguide or into the waveguide through a longitudinal slot.

The rotary vane attenuator is a very popular instrument. At the input end, there is a rectangular-to-circular waveguide taper containing a fixed lossy vane perpendicular to the incident electric vector. The central section contains a lossy vane diametrically across a circular waveguide that can be rotated, and the output section is a mirror image of the input section. When the central vane is at an angle θ relative to the two fixed end vanes, the total attenuation in decibels is given by Eq. (12), where A_0 is the residual attenuation when all three vanes lie in the same plane.

$$A_{\mathrm{rva}} = 40 \log_{10}(\sec\theta) + A_0 \tag{12}$$

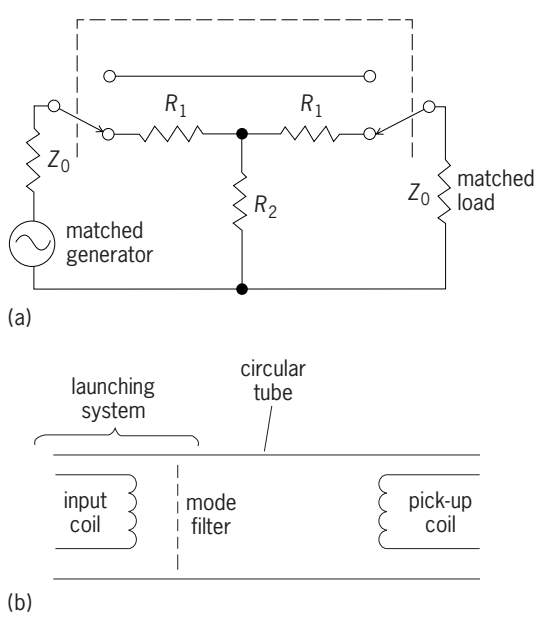

Fig. 1. Types of attenuators. (*a*) One-section T-type attenuator that can be switched in or out of a circuit between a matched generator and a matched load. (*b*) Piston attenuator.

Measurement of attenuation. Many different techniques for measuring attenuation have been devised. The power-ratio method is widely used. The simplest configuration requires only a stable well-matched filtered source and a well-matched low-drift power meter. Let the measured power be P_5 when the source is connected directly to the power meter and P_6 when the device under test is inserted between them. Then the attenuation of the device under test in decibels is given by Eq. (13). Insensitivity to

$$A_{\text{dut}} = 10 \log_{10} \frac{P_5}{P_6} \tag{13}$$

source power variations can be achieved by using two power meters (**Fig. 2**). A Kelvin-Varley voltage divider is adjusted to settings D_1 and D_2 that give nulls before and after inserting the device under test. Then the attenuation of the device under test is given by

$$A'_{\text{dut}} = 10 \log_{10} \frac{D_1}{D_2} \tag{14}$$

Eq. (14). If sufficient care is taken, this configuration will yield very accurate results over a 30-dB range.

Substitution methods of attenuation measurement are very popular, and the following techniques can be used: (1) radio-frequency (rf) substitution, for example, comparison with an accurate rotary vane attenuator; (2) intermediate-frequency substitution, for example, comparison with an i-f piston attenuator; (3) audio-frequency substitution, for example, comparison with an inductive voltage divider; and (4) dc substitution, for example, comparison with a Kelvin-Varley voltage divider. *See* INDUCTIVE VOLTAGE DIVIDER.

The device under test can be connected either in series or in parallel with the reference standard. For example, in a manually switched parallel i-f substitution system (**Fig. 3**), the device under test is inserted between matched isolators, and a continuous-wave signal at the required measurement frequency f is passed through it. This signal then has its frequency changed to, say, 30 MHz by a linear mixer that is driven by a local oscillator operating at a frequency of $f \pm 30$ MHz. A stable 30-MHz continuous-wave oscillator energizes the piston attenuator and, at each setting of the device under test, the piston is adjusted until the reading on the output voltmeter is the same in both positions of the switch. A system of this type has a dynamic range of about 100 dB.

When extremely high dynamic range (up to 160 dB) is needed, this type of receiver is inadequate. However, the necessary extra sensitivity can be obtained by employing a dual-channel phase-sensitive detection system, which can extract signals that are deeply buried in noise. In one configuration (**Fig. 4**), a synthesized signal source and a synthesized local oscillator are operated 10 kHz apart, so that, after mixing, 10-kHz signals are produced in both channels. Isolation amplifiers with very high reverse attenuation are used to eliminate mixer-to-mixer leakage. One channel provides a clean reference signal for a lock-in analyzer, and the measurement process

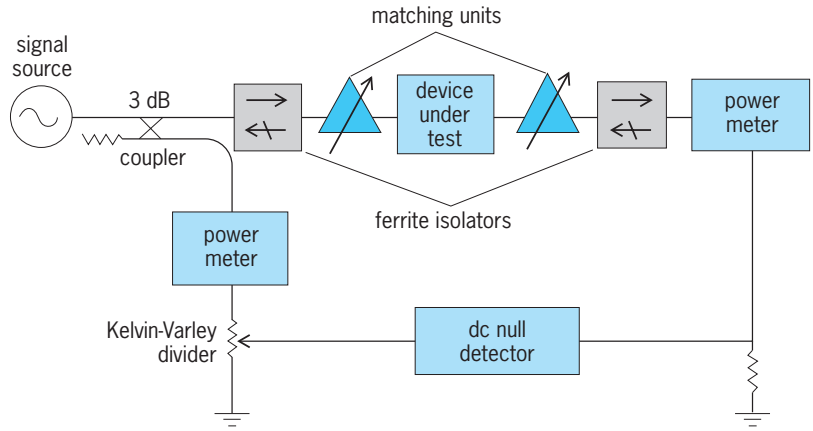

Fig. 2. Dual-channel power ratio system for attenuation measurement.

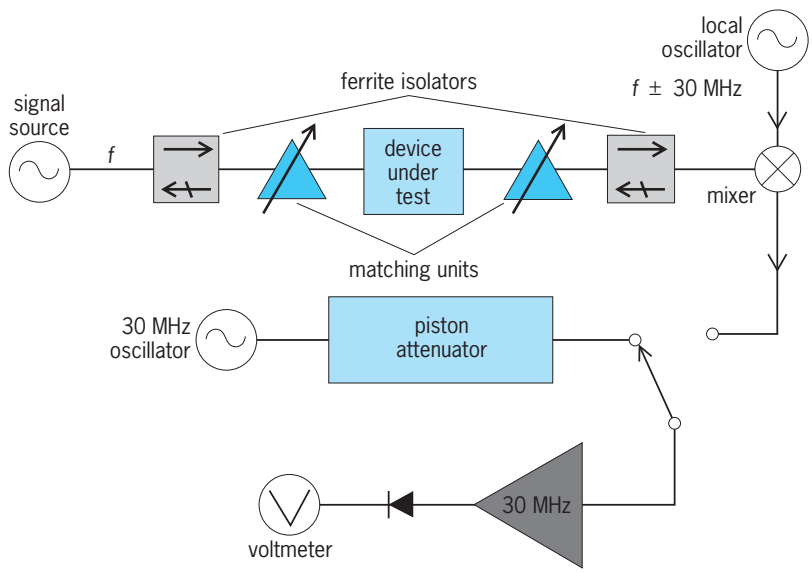

Fig. 3. Parallel i-f substitution system for attenuation measurement.

takes place in the other channel, which also feeds this analyzer. The lock-in analyzer contains in-phase and quadrature phase-sensitive detectors, whose outputs are $V \cos \phi$ and $V \sin \phi$, where V is the peak value of the 10-kHz signal emerging from the measurement channel, and ϕ is the phase difference between the signals in the two channels. These outputs are combined in quadrature to yield a direct voltage V_{out}, given by Eq. (15). Thus, V_{out} is directly related

$$V_{\text{out}} = \{(V \cos \phi)^2 + (V \sin \phi)^2\}^{1/2} = V \tag{15}$$

to the measurement signal V and is independent of the phase difference. Both a gage-block attenuator and an inductive voltage divider are placed in the measurement channel. When the device under test is inserted in this channel, appropriate amounts of attenuation are removed from both those devices to restrict the output voltage change to 20 dB (the maximum range over which the lock-in analyzer is highly linear). The attenuation in decibels through the device under test is given by Eq. (16), where $V_{\text{out},1}$ and

$$A_{\text{out}} = 20 \log_{10} \frac{V_{\text{out},1}}{V_{\text{out},2}} + A_{\text{gba}} + A_{\text{ivd}} \tag{16}$$

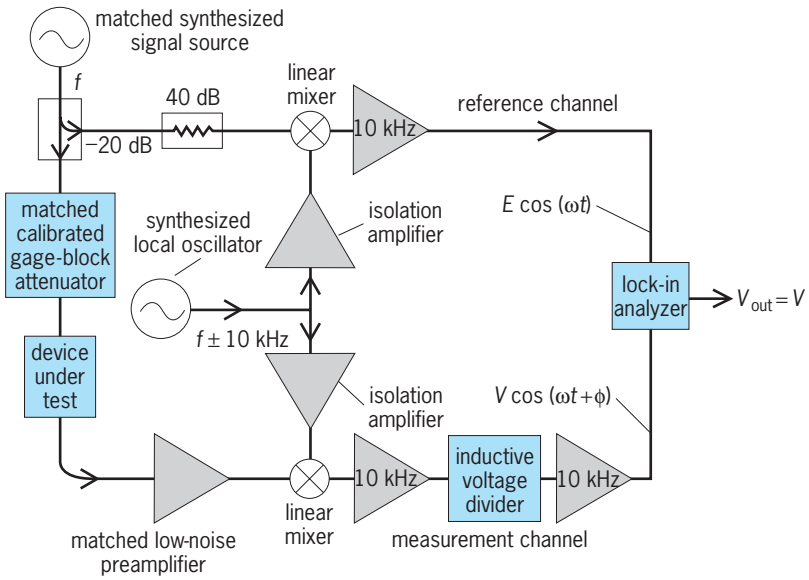

Fig. 4. Voltage-ratio-plus-gage-block system for attenuation measurements up to 160 dB. (*After G. J. Kilby, T. A. J. Smith, and F. L. Warner, The accurate measurement of high attenuation at radio frequencies, IEEE Trans. Instrum. Meas., 44:308–311, 1995*)

$V_{out,2}$ are the output voltages at the datum and calibration settings, respectively, and A_{gba} and A_{ivd} denote the attenuations in decibels removed from the gage-block attenuator and the inductive voltage divider when the device under test is inserted. The gage-block attenuator can be calibrated against the inductive voltage divider by using series substitution. To achieve accurate results up to 160 dB, careful attention must be paid to the elimination of leakage and circulating earth currents.

Low values of attenuation can be determined accurately by making reflection coefficient measurements on the device under test with a sliding short behind it. Several bridge techniques for measuring attenuation have been devised. Attenuation changes of less than 10^{-4} dB can be readily seen on a magic T bridge with a superheterodyne null detector. *See* MICROWAVE IMPEDANCE MEASUREMENT.

The attenuation in a waveguide can be found by making Q measurements on resonant sections of different lengths. *See* CAVITY RESONATOR; Q (ELECTRICITY).

The microwave reflection coefficient of a superconducting ring closed by a Josephson junction varies periodically as the current through an adjacent wire is increased. By using this effect (by connecting the device under test in series with the adjacent wire), discrete values of attenuation can be found very accurately by simply counting the number of zero crossovers and making use of a table of Bessel function roots. *See* JOSEPHSON EFFECT; SUPERCONDUCTIVITY.

When only moderate accuracy (on the order of ±0.5 dB) is required over a wide frequency range, a leveled swept source can be connected to the device under test, and the emerging signal can be fed to a diode detector that is followed by a logarithmic amplifier and oscilloscope. With this technique, resonances which can easily be missed using point-by-point methods are clearly revealed, and the effects of adjustments to the device under test can be seen immediately over a wide band. *See* AMPLITUDE-MODULATION DETECTOR.

Network analyzers became very popular after 1967. These instruments yield both the magnitude and phase angle of the transmission and reflection coefficients of the device under test over a wide frequency range. By using ingenious calibration and computer-correction techniques, high accuracy can be achieved. To obtain the phase information, a second channel is needed to provide a reference signal. In most network analyzers, mixers are employed and high-dynamic-range instruments of this type are now commercially available for use at frequencies up to 100 GHz. Considerable effort has been devoted since 1972 to six-port network analyzers in which the phase information is deduced solely from power measurements. *See* TRANSMISSION LINES.

Frank L. Warner

Bibliography. A. E. Bailey (ed.), *Microwave Measurements*, 2d ed., 1989; G. H. Bryant, *Principles of Microwave Measurements*, 1988; G. F. Engen, *Microwave Circuit Theory and Foundations of Microwave Metrology*, 1992; I. Kneppo, *Microwave Measurement by Comparison Methods*, 1988; S. R. Pennock and P. R. Shepherd, *Microwave Engineering with Wireless Applications*, Macmillan, London, 1998.

Audio amplifier

An electronic circuit for amplification of signals within or somewhat beyond the audio frequency range (generally regarded as 20 to 20,000 Hz). Audio amplifiers may function as voltage amplifiers (sometimes called preamplifiers), power amplifiers, or both. In the last case, they are often called integrated amplifiers. *See* POWER AMPLIFIER; VOLTAGE AMPLIFIER.

The function of integrated amplifiers (or of the combination of separate voltage amplifiers and power amplifiers used together) is to amplify a weak signal, such as from a microphone, phonograph pickup, tape player, radio tuner, or compact disc player, to a level capable of driving a loudspeaker or other type of transducer such as headphones at the desired sound level. Power amplifiers may have power ratings ranging from less than 1 W to several hundreds of watts. Stereo amplifiers consist of two identical, but electrically independent, amplifier circuits housed in a single chassis, often sharing a common power supply. Audio amplifiers are commonly constructed with solid-state devices (transistors and integrated circuits), although some amplifiers using vacuum tubes as the active, amplifying devices are still manufactured. *See* INTEGRATED CIRCUITS; TRANSISTOR; VACUUM TUBE.

Classification. Audio power amplifiers, or the power sections of integrated amplifiers, are classified in accordance with the manner in which the output stages conduct current. Class A amplifiers conduct a constant current regardless of whether an audio signal is applied to their input or not. They

are therefore relatively inefficient. Class AB amplifiers conduct a moderate amount of current even in the absence of an input signal, but the current increases as signal levels increase. Most modern audio power amplifiers for high-fidelity applications employ this type of circuitry. Class B circuits conduct minute amounts of current in the absence of an input signal and are therefore the most efficient in terms of power dissipation versus power delivered to the loudspeaker or other connected load. Other amplification classes have been developed by specific manufacturers, such as class G and class H, each of which attempts to improve efficiency beyond that attainable with either class AB or class B circuits.

Performance characteristics. The ideal amplifier delivers an output signal that, aside from its higher power level, is identical in relative spectral content to the input signal. Normally, various forms of distortion are generated by the amplifier, such as harmonic distortion (multiples of the desired signal frequency), intermodulation distortion (spurious sum or difference frequencies created when multiple tones are applied to the amplifier simultaneously, as in the case of music or speech amplification), and transient intermodulation distortion (caused by rapid fluctuations of the input signal level). All forms of distortion are measured as percentages of the desired signal amplitude. Generally, distortion levels of under 1% or 0.5% are considered to be low enough for high-fidelity applications. *See* DISTORTION (ELECTRONIC CIRCUITS); FIDELITY.

Other parameters used to define an amplifier's characteristics include frequency response and signal-to-noise ratio (S/N). The frequency response is the range of frequencies that the amplifier can handle, usually quoted with a tolerance in decibels, for example: "frequency response: 20 to 20,000 Hz, ±0.5 dB." Signal-to-noise ratio, also quoted in decibels, is indicative of the amount of residual noise generated by the amplifier itself, as compared with the desired output signal level. Signal-to-noise levels greater than 70 dB are generally considered to be acceptable, although some amplifiers offer much better values. *See* ELECTRICAL NOISE; RESPONSE; SIGNAL-TO-NOISE RATIO.

Speaker impedance. Although not related to quality of sound reproduction, other important amplifier specifications include a listing of suitable loudspeaker impedances that may be safely connected to the amplifier. Loudspeakers generally have impedances of 4 or 8 ohms. Most audio amplifiers can drive either of these impedances, often delivering higher power when the lower-impedance loudspeakers are connected. If the loudspeaker impedance dips to too low a value, however, the amplifier may not be able to deliver the higher current required and may either fail or cause its protection circuits to temporarily disable the amplifier. While many amplifiers can power more than one loudspeaker per channel, wiring multiple loudspeakers in parallel across the output terminals of an audio amplifier lowers the net impedance into which the amplifier must operate. Thus, if two 8-ohm loudspeakers are connected in parallel, the net impedance across the amplifier terminals becomes 4 ohms. *See* AMPLIFIER; ELECTRICAL IMPEDANCE; LOUDSPEAKER.

Leonard Feldman

Bibliography. R. Angus, L. Feldman, and N. Eisenberg, *The New World of Audio: A Music Lover's Guide*, 1982; G. Ballou, *Handbook for Sound Engineers:*, 2d ed., 1991; M. Clifford, *Modern Audio Technology*, 1992.

Audiometry

The quantitative assessment of individual hearing, both normal and defective. In typical tests, pure tones are presented through headphones, though some tests use speech instead of tones. In bone conduction tests, tones are presented through a vibrator placed on the head. Audiometric tests may serve various purposes, such as investigation of auditory fatigue under noise conditions, human engineering study of hearing aids and communication devices, screening of individuals with defective hearing, and diagnosis and treatment of defective hearing. In all of these situations, individual hearing is measured relative to defined standards of normal hearing (**Fig. 1**).

Audiometer. The pure-tone audiometer is the instrument used most widely in individual hearing measurement. It is composed of an oscillator, an amplifier, and an attenuator to control sound intensity. For speech tests of hearing, called articulation tests, word lists are reproduced on records or tape recorders. Measurements of detectability or

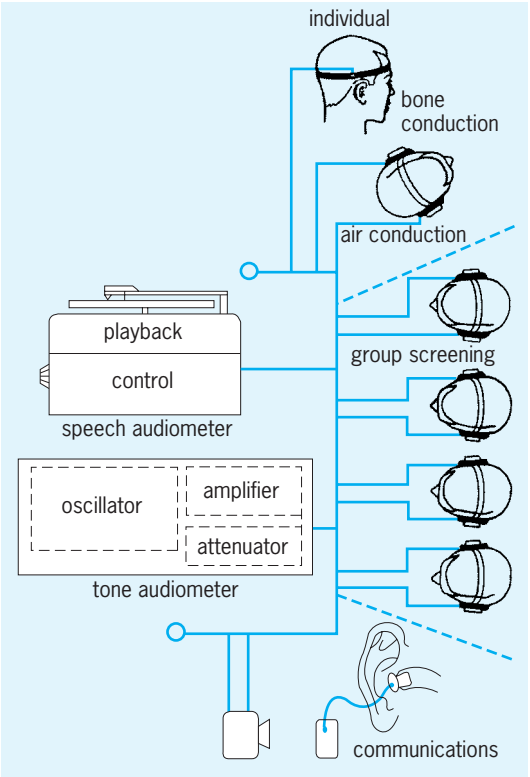

Fig. 1. Test equipment for tone and speech audiometry as applied to individual hearing measurement.

intelligibility can be made by adjusting the intensity of the test words. To make bone conduction tests, signals from the audiometer activate a vibrator located on the forehead or mastoid bone.

Reverberation and anechoic rooms. Scientific advance in audiometry demands careful control of all environmental sound. Two types of rooms especially constructed for research and measurement of hearing are shown in **Fig. 2**, the random diffusion, or reverberation, chamber and the anechoic room. In the reverberation chamber, sounds are randomly reflected from heavy nonparallel walls, floor, and ceiling surfaces. In the anechoic room, the fiber glass wedges absorb all but a small percentage of the sound. Such conditions make possible precise determination of the limits and variation of normal and defective hearing for all types and conditions of sound transmission.

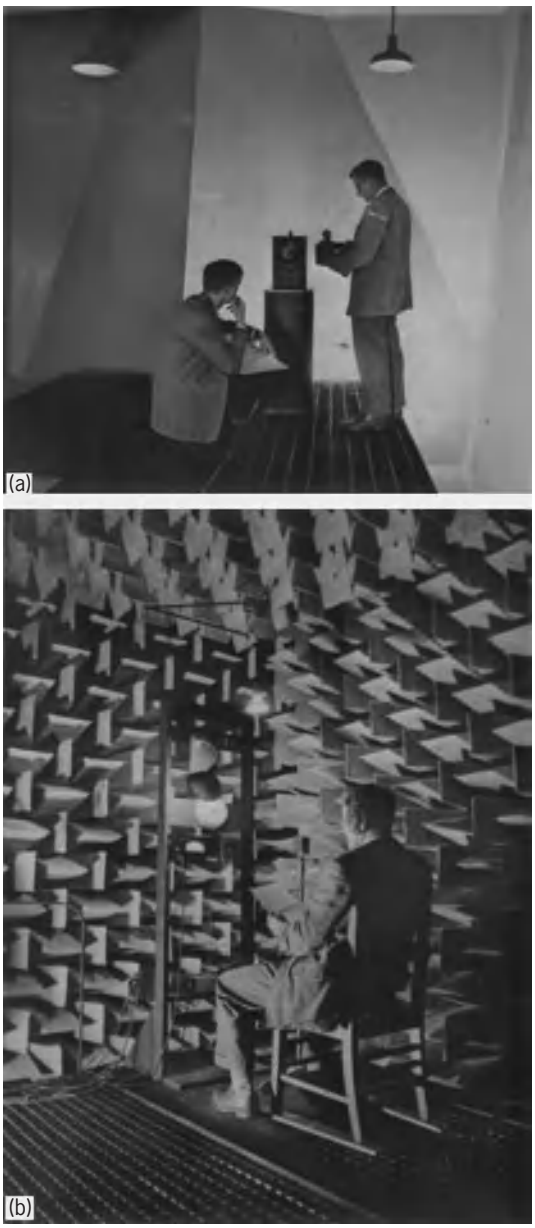

Fig. 2. Special environments for audiometry. (*a*) Random diffusion chamber for measuring hearing under controlled noise conditions. (*b*) Anechoic soundproofed room lined with fiber glass wedges. (*Courtesy of J. P. Egan*)

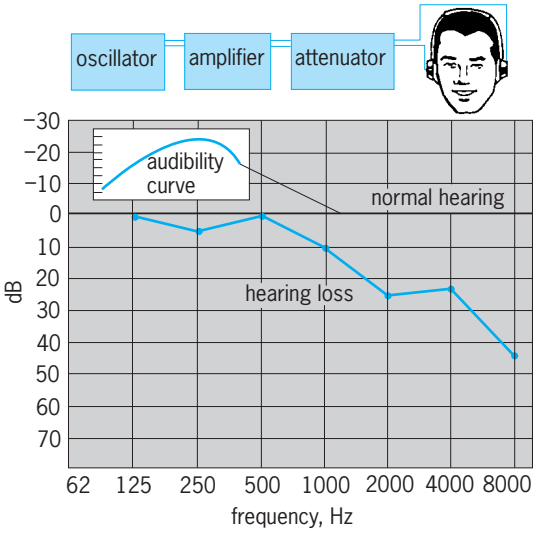

Fig. 3. Audiogram for determining pure-tone hearing loss at various frequency levels.

Audiogram. The measurement of hearing loss for pure tones is represented by the audiogram (**Fig. 3**). Sounds of different frequencies are presented separately to each ear of the individual, and the absolute threshold for each frequency is determined. The absolute threshold is the lowest intensity that can be detected by the individual who is being tested.

In clinical audiometry the status of hearing is expressed in terms of hearing loss at each of the different frequency levels. In the audiogram the normal audibility curve, representing absolute thresholds at all frequencies for the normal ear, is represented as a straight line of zero decibels. Amount of hearing loss is then designated as a decibel value below normal audibility, as defined by the American National Standards Institute (ANSI) or the International Organization for Standardization (ISO). The audiogram in Fig. 3 reveals a hearing loss for tones above 500 Hz. Automatic audiometers are now in use which enable one to plot an audiogram for oneself.

Articulation tests. Articulation tests are tests used to assess hearing and loss of hearing for speech. The threshold of intelligibility for speech is defined as the intensity level at which 50% of the words, nonsense syllables, or sentences used in the articulation test are correctly identified. The best of such articulation tests are phonetically balanced (PB) word tests, in which single-syllable words were selected so that the phonetic sounds naturally occurring in the English language would be proportionately represented in the test. The hearing loss for speech is determined by computing the difference in decibels between the individual intelligibility threshold and the normal threshold for that particular speech test. Discrimination loss for speech represents the difference between the maximum articulation score at a high intensity level (100 dB), expressed in percent of units identified, and a score of 100%. The measure of discrimination loss helps distinguish between conduction loss and nerve deafness.

Bone conduction tests. Bone conduction audiograms are compared with air conduction audiograms in order to analyze the nature of deafness. Loss in bone conduction hearing as well as air conduction hearing often indicates nerve deafness, as contrasted to middle-ear or conduction deafness. *See* EAR (VERTEBRATE).

Audiometric testing. Audiometric testing is adjusted to meet various needs of individuals. School audiometry makes possible rough determination of hearing loss in children, who can thereafter be tested with more precise techniques if some loss is found. Audiometry can be extended to assessment of hearing in very young children who cannot speak or follow directions. In such cases, certain reflexes, such as the psychogalvanic skin reflex, are conditioned to sound. Subsequently the presence or absence of the conditioned reflex can be used to indicate whether or not the child hears a particular sound. Young children can also be trained to press buttons to get a reward when sounds are present. Thereafter the occurrence of the learned response in relation to a test sound is used to determine the threshold of audibility.

Automatic audiometry. The main theoretical and technological issues of audiometry have focused on the design of psychometric and electrophysiological procedures, automation of testing, and application of computers to audiometer design.

Békésy audiometry. In 1974 G. von Békésy designed the first observer-controlled threshold-testing procedure. In Békésy audiometry, the observer indicates that he or she hears a sound by pressing a button. When the button is held down, the sound intensity is gradually reduced by a motor-driven attenuator until the observer releases the button, indicating that the sound is no longer heard. At this point, the sound intensity is gradually increased. As the observer makes these adjustments, the frequency of the test tone is advanced from 100 to 10,000 Hz and an audiogram is automatically charted. The audiogram in **Fig. 4** shows a marked hearing loss for tones above 3000 Hz.

Electrical response audiometry. Standard audiometric methods can be used to test children as young as 4 or even 3 years. The need to assess hearing in younger children and infants, and to be able to test some retarded and handicapped individuals, has prompted the development of automatic audiometric methods. Some of these methods measure reflexes and automatic reactions, such as the startle response and changes in heart rate.

More reliable and consistent than reflexes are various electrical measures of the auditory system's response to sounds. By placing electrodes at or near the eardrum, it is possible to measure the electrocochleogram, which is the auditory nerve's response to sounds. If electrodes are attached to the surface of the skull, it is possible to measure electrical potentials generated in the brainstem and cerebral cortex. Abnormalities in these electrical audiograms can indicate defective functioning in the ear or in the auditory nervous system.

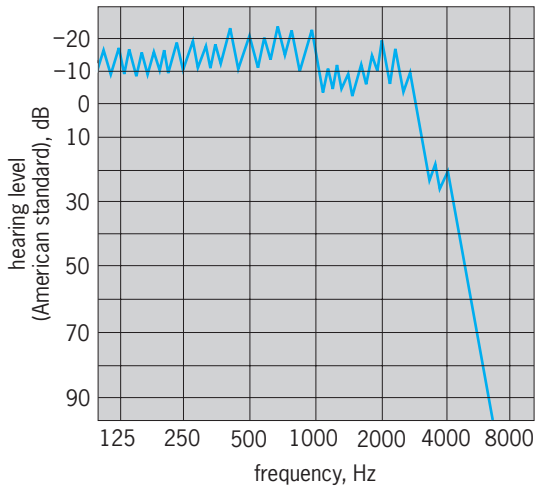

Fig. 4. Audiogram obtained by Békésy audiometric method of observer-controlled threshold testing.

Computer systems. Conventional data-processing instruments have been developed to program sequences of discrete sound stimuli that the observer must judge separately. These devices have been designed so that the observer should be limited to making only passive judgments, while the machine is programmed to do all the work. *See* COMPUTER.

Hybrid computer instruments have been constructed for the utilization of observer-controlled methods of audiometric measurement (**Fig. 5**). In

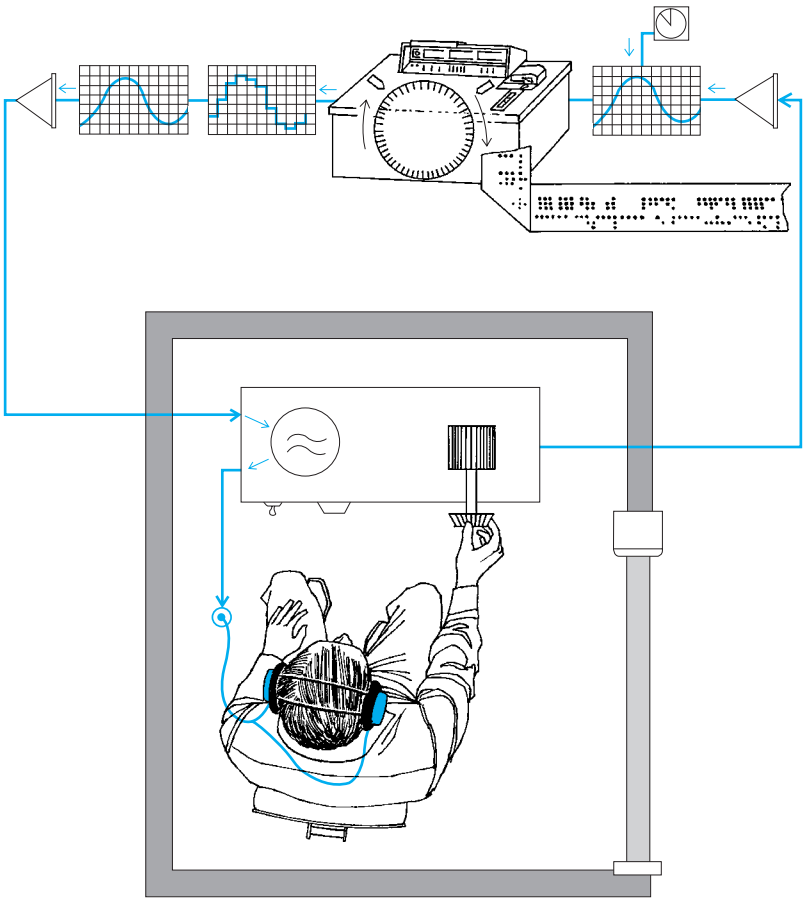

Fig. 5. Basic design of a hybrid computer system for dynamic audiometric testing.

these systems the subject uses a hand-controlled adjustment to control either intensity or frequency of a variable oscillator system. Dynamic programming is used to generate predetermined changes in the frequency and intensity of the tones heard by the subject. The computer program can require the subject to compensate or negate changes in sound, to track positively the loudness of a second sound, or to generate sounds to a point of zero loudness. For the future such methods will permit the measurement of hearing that can be associated directly with impedance measures of ear reactivity. *See* ELECTRODERMAL RESPONSE; HEARING (HUMAN); HEARING IMPAIRMENT; LOUDNESS. Lawrence E. Marks

Bibliography. H. A. Beagley (ed.), *Audiology and Audiological Medicine*, 1981; H. Kaplan et al., *Audiometric Interpretation: A Manual of Basic Audiometry*, 2d ed., 1993; F. N. Martin, *Introduction to Audiology*, 4th ed., 1990.

Auger effect

One of the two principal processes for the relaxation of an inner-shell electron vacancy in an excited or ionized atom. The Auger effect is a two-electron process in which an electron makes a discrete transition from a less bound shell to the vacant, but more tightly bound, electron shell. The energy gained in this process is transferred, via the electrostatic interaction, to another bound electron which then escapes from the atom. This outgoing electron is referred to as an Auger electron and is labeled by letters corresponding to the atomic shells involved in the process. For example, a KL_IL_{III} Auger electron corresponds to a process in which an L_I electron makes a transition to the K shell and the energy is transferred to an L_I electron (**illus**. *a*). By the conservation of energy, the Auger electron kinetic energy E is given by $E = E(K) - E(L_I) - E(L_{III})$ where $E(K,L)$ is the binding energy of the various electron shells. Since the energy levels of atoms are discrete and well understood, the Auger energy is a signature of the emitting atom. *See* ELECTRON CONFIGURATION; ENERGY LEVEL (QUANTUM MECHANICS).

The other principal process for the filling of an inner-shell hole is a radiative one in which the transition energy is carried off by a photon (illus. *b*).

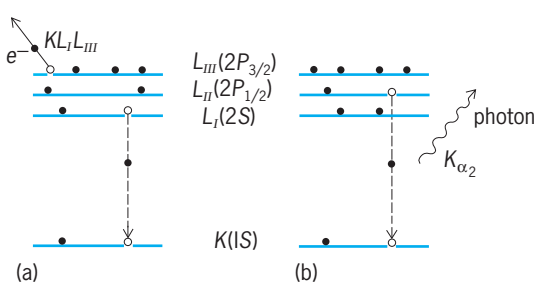

Two principal processes for the filling of an inner-shell electron vacancy. (*a*) Auger emission; a KL_IL_{III} Auger process in which an L_I electron fills the K-shell vacancy with the emission of a KL_IL_{III} Auger electron from the L_{III} shell. (*b*) Photon emission; a radiative process in which an L_{II} electron fills the K-shell vacancy with the emission of a $K_{\alpha 2}$ photon.

Inner-shell vacancies in elements with large atomic number correspond to large transition energies and usually decay by such radiative processes; vacancies in elements with low atomic number or outer-shell vacancies with low transition energies decay primarily by Auger processes. *See* ATOMIC STRUCTURE AND SPECTRA.

Auger electron spectroscopy is an important tool in the analysis of the near-surface region of solids. Inner-shell vacancies in the atoms of solids are created by incident radiation, usually an energetic electron beam, and the energy spectrum of the outgoing Auger electrons is then used to determine the types of atoms present, and hence the elemental composition of the material. In solids, the outgoing Auger electron has a high probability of undergoing an inelastic or energy loss event which changes the Auger electron energy. The probability of an inelastic process is characterized by a mean free path, a thickness in which approximately 63% of the Auger electrons undergo an energy changing event. The mean free path can vary from approximately 0.5 nanometer to 100 nm, depending on the electron energy and the solid; thus Auger electron spectroscopy is useful for measurements of the elemental composition of a material in its first 0.5–100 nm. Elemental composition as a function of depth is determined by using Auger electron spectroscopy combined with some erosion process which slowly removes the surface layers; the Auger spectrum measured at various stages of the erosion process yields a depth profile of the elemental composition of the material. *See* ELECTRON SPECTROSCOPY; SURFACE PHYSICS. Leonard C. Feldman

Bibliography. P. Auger, The compound photoelectric effect, *J. Phys. Radium*, 1925; G. Cubiotti, G. Mundio, and K. Wandett (eds.), *Auger Spectroscopy and Electronic Structure*, 1989.

Auger electron spectroscopy

Auger electron spectroscopy (AES) is a widely used technique that detects the elements in the first atomic layers of a solid surface. Although many elements can be detected, hydrogen usually cannot be observed. Excellent spatial resolution can be achieved. Auger electron spectroscopy is important in many areas of science and technology, such as catalysis, electronics, lubrication, and new materials, and also understanding chemical bonding in the surface region. Auger spectra can be observed with gas-phase species.

Basic principles. In Auger electron spectroscopy an electron beam, usually 2–20 kV in energy, impinges on a surface, and a core-level electron in an atom is ejected. An electron from a higher level falls into the vacant core level. Two deexcitation processes are possible: An x-ray can be emitted, or a third electron is ejected from the atom (**Fig. 1**). This electron is an Auger electron; the effect is named after its discoverer, Pierre Auger. Auger electrons for surface analysis can be created from other sources, such as, x-rays, ion beams, and positrons. K capture (radioactive decay) is another source of Auger

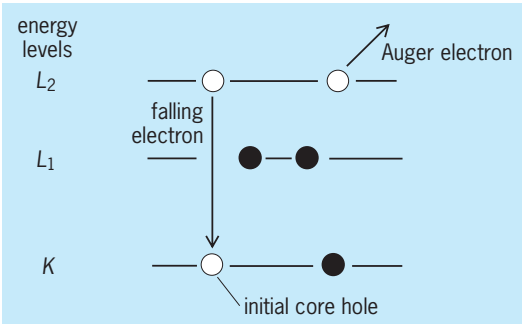

Fig. 1. Schematic diagram of an Auger electron being ejected from a carbon atom for a KL_2L_2 transition. Round forms are electrons.

electrons. The Auger yield (the ratio of Auger electrons to the number of core holes created) depends upon the element, the initial core level, and the excitation conditions. *See* AUGER EFFECT.

The surface sensitivity of Auger electron spectroscopy is based upon the fact that electrons with kinetic energy in the range of 30–2500 eV have an inelastic mean free path (IMFP) of about 0.5–3 nanometers for most materials. The inelastic mean free path is defined as the average distance that an electron will travel before it undergoes an inelastic collision with another electron or nucleus of the matrix constituent atoms. Very few Auger electrons created at depths greater than two to three times the IMFP will leave the bulk with their initial kinetic energy. Those electrons formed at greater depths undergo collisions with atoms and electrons in the bulk and lose part or all of the initial creation energy. In this energy range, all elements with an atomic number of 3 or more have detectable Auger electrons. Most elements can be detected to 1 atomic percent in the analysis volume, and relative atomic amounts usually can be determined.

X-ray notation is used to signify Auger transitions. For example, if a K (1s) electron in silicon is ejected and an L_3 ($2p_{3/2}$) electron falls into the K-level hole and another L_3 is ejected as the Auger electron, the transition is KL_3L_3. If a valence electron is involved, it often is denoted with a V. The energy relationship for Auger electrons is given by

$$E_{KE}(j,k,l) = E_{KE}(j) - E_{KE}(k) - E_{KE}(l) - U$$

where $E_{KE}(j,k,l)$ is the energy of the Auger electron, $E_{KE}(j)$ is the binding energy of the electron ejected during the initial ionization event, $E_{KE}(k)$ is the binding energy of the electron that falls into the empty level, $E_{KE}(l)$ is the binding energy of the ejected electron, and U is the hole-hole repulsion energy. U (about 1–10 eV) takes into account that there are two empty electron states in the excited atom. Handbooks containing the spectra of most elements are used for routine elemental identification. If a transition is within the same major elemental level, for example, $M_1M_{2,3}M_{2,3}$, it is called a Coster-Kronig transition. Chemical information sometimes can be obtained with Auger electron spectroscopy; for example, the $KL_{2,3}L_{2,3}$ silicon (Si) transition energies for Si and SiO_2 are 1616 and 1608 eV, respectively. The energy difference between L_2 and L_3 cannot be distinguished with Auger electron spectroscopy in this instance. Chemical information occasionally can be observed by line shape differences, for example, between carbides and graphite. With nonconducting samples an electron imbalance sometimes occurs on the surface that changes the observed transition energies. There are several ways to overcome this charging, but in extreme instances a spectrum cannot be observed.

A spectrum for copper is shown in **Fig. 2** in both integral and differential modes. In some instances, small peaks are easier to notice with differential spectra. Derivative spectra can be obtained instrumentally or mathematically from integral spectra. When an electron beam is used to produce Auger electrons, elemental maps can be made by analyzing only small energy regions over an area. This approach is called scanning Auger microscopy (SAM). These maps indicate spatial variations of elements over a selected area in the surface region. If the electron beam is moved in one direction, elemental line scans are produced. With the electron beam, secondary electrons are produced and images can be observed in the same way as with scanning electron microscopy. *See* ATOMIC STRUCTURE AND SPECTRA; ELECTRON MICROSCOPE.

Elemental distribution into the bulk can be obtained by sputtering, or "molecular sandblasting." This is done with a beam of gas ions usually 0.5–5 kV in energy. At selected intervals, spectra are recorded for chosen elements. This technique is used to study ion-implanted or layered materials and thin oxide layers. In some instances the depth resolution can

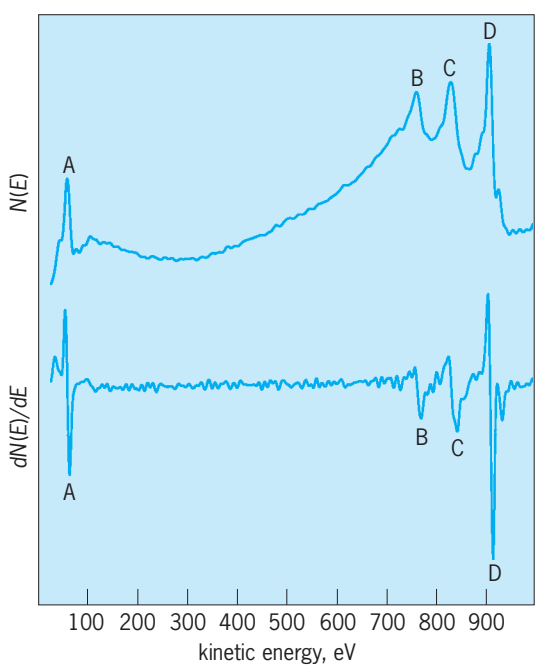

Fig. 2. Integral (upper) and differential (lower) mode spectra of sputter-cleaned copper with the most intense peaks identified: A, $M_{2,3}M_{4,5}M_{4,5}$; B, $L_3M_{4,5}M_{4,5}$; C, $L_3M_{2,3}M_{4,5}$; D, $L_3M_{2,3}M_{2,3}$.

be a few atomic layers with sample rotation during sputtering. Sputter profiles usually are not greater than a few micrometers.

Equipment. Auger electron spectroscopy spectra are obtained in a vacuum chamber at pressures below about 10^{-2} pascal (10^{-4} millibar), but most systems operate in the ultrahigh-vacuum (UHV) regime ($<10^{-6}$ Pa or 10^{-8} mb). These ultrahigh-vacuum chambers are constructed mostly with metal, and special handling of the materials in the system is required. In addition, specimens have to be handled with clean-room-type gloves to eliminate the transfer of oils and salts to the sample surface. Most systems have a sample introduction apparatus that permits the main chamber to remain at a low pressure when placing specimens in the analysis chamber. The samples are placed on a holder that permits their movement for analysis at specific locations. Many holders permit the sample temperature to be lowered to about 90 K (−183°C; −298°F) or raised to over 1000 K (727°C; 1340°F).

The electron beam is created with an electron gun that uses a heated filament of tungsten, a lanthanum boride (LaB_6) crystal, or a Schottky barrier filament (zirconium oxide, ZrO) coated on a single crystal face of tungsten. With the last source, the beam diameter can be below 20 nm. Current levels range from below 1 nanoamp to milliamps depending on the source. If x-rays are used to create Auger electrons, the ability to analyze small areas is lost, but chemical information can be obtained by measuring the energy difference between the x-ray photoelectrons and Auger electrons. The cylindrical mirror analyzer (CMA) or the concentric hemispherical analyzer (CHA) usually is employed to measure the energy of Auger electrons. The CMA has somewhat better throughput, but the CHA has much better energy resolution. Auger spectroscopy electrons are detected with channel or position-sensitive electron multipliers, and the signals are recorded with computers. Sputtering is done with an ion gun with current densities up to a few microamps per square centimeter. Usually an inert gas is employed at pressures of 10^{-2} Pa in the ionization region. Often, brief sputtering is used to remove atmospheric contaminants (usually containing carbon and oxygen) that are present with almost all materials that have been exposed to air.

Noel H. Turner

Bibliography. D. Briggs and M. P. Seah (eds.), *Practical Surface Analysis*, 2d ed., 1990; C. L. Hedberg (ed.), *Handbook of Auger Electron Spectroscopy: A Book of Reference Data for Identification and Interpretation in Auger Electron Spectroscopy*, 3d ed., 1995.

Augite

A group of monoclinic calcic pyroxenes which have the general chemical formula (Ca,Mg,Fe)(Mg,Fe)-Si_2O_6, in which calcium is the dominant cation in the first cation position. Monoclinic pyroxene with substantial iron or magnesium in place of calcium is called pigeonite, and has a different crystal structure from augite. Augite is generally considered to be a combination of the four end members, diopside ($CaMgSi_2O_6$), hedenbergite ($CaFe^{2+}Si_2O_6$), enstatite ($Mg_2Si_2O_6$), and ferrosilite ($Fe_2{}^{2+}Si_2O_6$), but it almost always has substantial aluminum and minor to substantial amounts of sodium, ferric iron, chromium, and titanium. These minor constituents enter as solid solutions of augite toward acmite ($NaFe^{3+}Si_2O_6$), jadeite ($NaAlSi_2O_6$), calcium Tschermak's molecule ($CaAl_2SiO_6$), ureyite ($NaCrSi_2O_6$), and $CaTiAl_2O_6$. The amount of substitution by any of these minor components depends upon the bulk composition of the rock as well as the conditions of formation. Four common varieties of augite are omphacite, which has about 50% substitution of jadeite; aegerine augite, which has substantial acmite substitution; fassaite, which has a considerable amount of calcium Tschermak's molecule in solution; and titanaugite, which is rich in titanium and has a characteristic purple color. *See* DIOPSIDE; ENSTATITE; PIGEONITE.

Augite in igneous and metamorphic rocks does not commonly occur as well-formed crystals; when it does, the crystals are dark green to black, short and prismatic. Augite crystallizes in a *C*-centered monoclinic cell, space group *C*2/*c*, and is commonly recognized by its prominent (110) pyroxene cleavages which form 87° angles. It displays typical pyroxene crystal chemistry—single chains of tetrahedral (SiO_4) groups alternating with strips containing two distinct types of octahedral sites accommodating Ca, Fe, Mg, Mn, Na, Cr, and Ti. The smaller Al cation can enter both tetrahedral and octahedral sites.

Augite occurs in both igneous and metamorphic rocks. It is nearly universal in basalts and gabbros, and occurs somewhat less frequently in less mafic igneous rocks. Alkali olivine basalts and other alkali-rich volcanic rocks contain augite which is commonly enriched in Na, Ti, Al, and Fe^{3+}. Magnesium-rich augite is a characteristic mineral in many ultramafic rocks and in rocks of the Earth's mantle. Augite and pigeonite are also rather common constituents of lunar basalts and basaltic meteorites. In metamorphic rocks, augite of a composition close to diopside occurs in marbles and in calc-silicate rocks. More compositionally complex augite occurs in high-grade metamorphosed amphibolites and granulites where it typically coexists with hornblende or orthopyroxene. The augite variety omphacite is found with magnesium-rich garnet in eclogite which is of basaltic composition, and apparently forms at very high pressures. *See* ECLOGITE; PYROXENE.

Robert J. Tracy

Bibliography. W. A. Deer et al., *Rock-Forming Minerals*, vol. 2: *Chain Silicates*, 1963; C. Klein, *Manual of Mineralogy*, 21st ed., 1993; J. J. Papike (ed.), *Pyroxenes and Amphiboles*, Mineral. Soc. Amer. Spec. Pap. 2, 1969.

Aulopiformes

An order of about 219 species of teleost fishes ranked in four suborders. The order is distinguished by certain skeletal specializations of the gill arches that

are unknown in other teleost fishes. They share characters with the order Myctophiformes (lanternfishes) to the extent that they are combined under the ordinal name Myctophiformes by some authors. In both orders the premaxilla excludes the maxilla from the gape of the mouth; the mouth is not protractile; an adipose fin is usually present; the caudal fin is forked; pelvic fins are usually abdominal; light-producing organs are often present; a swim bladder, if present, is physoclistous (not connected to the gut); and some families have hermaphroditic species with self-fertilization.

Morphological diversity. Descriptions of selected families follow to illustrate the diversity in morphology.

Giganturidae (telescope fishes). These fishes are mesopelagic in the Atlantic, Indian, and Pacific oceans). Telescope fishes are the most bizarre of the order. They have a large mouth, extending far behind the tubular eyes; a pectoral fin inserted above the gill opening; a forked caudal fin, with lower lobe greatly extended; and a greatly expandable stomach. The maximum length is 22 cm (9 in.). The family consists of one genus and two species.

Chlorophthalmidae (greeneyes). These fishes are benthic on the continental shelf and adjacent deep Atlantic, Indian, and Pacific oceans. They are moderately slender, rounded anteriorly, and compressed posteriorly; with large, lensless eyes that are directed dorsolaterally; a well-developed adipose fin; maxillary expanded post; a single supramaxilla. The tapetum of the eye is iridescent green; the top of the head is scaleless; and the caudal fin is forked, with lobes about equal length. The maximum length is about 25 cm (10 in.). The family consists of two genera and 20 species.

Synodontidae (lizardfishes). Members are benthic in the Atlantic, Indian, and Pacific oceans and, rarely, in brackish waters. The body is slender and cylindrical; head depressed; mouth large, with gape oblique and maxillary rudimentary; scales cycloid, and absent from top of head; adipose fin present, small to large; no luminous organs (see **illustration**). There are five genera and 55 species.

Paralepididae (barracudinas). These fishes are found mid-depth to near surface in all oceans from the Arctic to the Antarctic. Scales are either present or absent; the anal fin base is long; there are moderately long pointed jaws; eyes are normal with a circular iris; an adipose fin is present; squamation varies from a scaled head, body, and vertical fins to completely naked, except for embedded scales in the lateral line; the dorsal fin is posteriorly placed; pelvic fins are abdominal; the anal fin is long; and luminous organs may be present or absent. The maximum length is 1 m (3.3 ft). There are 12 genera and 56 species.

Alepisauridae (lancetfishes). Members are found in the Atlantic, Indian, and Pacific oceans. The body is slender and subcylindrical; the scales and light organs are absent; the dorsal fin is high (as in sailfishes), extending from just behind the head to near the adipose fin; the pectoral fins are ventral, the pelvic fins are abdominal; the mouth is large, with well-developed teeth, and long palatine teeth. The maximum length is no less than 2 m (6.6 ft). There is one genus and two species.

Orangemouth lizardfish (*Saurida flamma*). (*Photo © John E. Randall*)

Pelagic specializations. Ecologically there are two major groups of Aulopiforms, the benthic families and the pelagic to bathypelagic families. The latter group is confined to midwaters of deep oceans. These are predatory fishes ranging from only 12 cm (4.7 in.) in length to 2 m (6.6 ft). They have no swim bladder; no luminescent organs; a single row of small teeth on the premaxillaries, one or two rows of large teeth on the palatines (roof of mouth), and one to three rows on each dentary (mandible); and gillrakers in the form of spines or teeth. The skeleton of most is lightly ossified and the flesh may be flaccid. The lack of a swim bladder (hydrostatic organ) to maintain the fish's position in the water column is compensated to some extent by the light skeleton and the reduction or total loss of scales in many of the species. In total darkness species recognition, essential for reproduction in dioecious fishes, is usually enhanced by luminescent organs; however, this group of fishes lack the ability to produce light. For self-fertilizing hermaphroditic species, as some of these are, the lack of luminous organs is of little consequence. In this environment life is sparse. Many of these fishes are equipped with a large mouth, large knifelike teeth, and the ability to hover and make swift darts at their prey, assuring them success in obtaining a meal when the rare opportunity exists. In species lacking scales, the skin can stretch to accommodate very large prey. *See* OSTEICHTHYES.

Herbert Boschung

Bibliography. E. Bertelsen, G. Krefft, and N. B. Marshall, The fishes of the family Notosudidae, *Dana Rep.*, 86:1–114, 1976; K. E. Hartel and M. J. L. Stiassny, The identification of larval parasudis (Teleostei, Chlorophthalmidae): With notes on the anatomy and relationships of aulopiform fishes, *Breviora*, vol. 487, p. 23, 1986; R. K. Johnson, *Fishes of the Families Evermannellidae and Scopelarchidae: Systematics, Morphology, Interrelationships, and Zoogeography*, Fieldiana Zool. New Ser., vol. 12, p. 252, 1982; R. K. Johnson, Giganturidae: Development and relationships, pp. 199–201, in H. G. Moser et al. (eds.), *Ontogeny and Systematics of Fishes*, Spec. Publ. no. 1, Amer. Soc. Ichthy. Herp., 1984; R. K. Johnson and E. Bertelsen, The fishes of the family Giganturidae: Systematics, development, distribution, and aspects of biology, *Dana Rep.*,

91:1–45, 1991; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994; R. R. Rofen, Fishes of the Western North Atlantic: Families Paralepididae, Omosududae, Anotopteridae, Evermannellidae, and Scopelarchidae, *Sears Found Mar. Res., Mem.* (Yale University), 1(5):205–481, 1966.

Aurora

An optical manifestation of a large-scale electrical discharge process which surrounds the Earth. The discharge is powered by the so-called solar wind–magnetosphere generator. The Sun continuously blows out its upper atmosphere, the corona, with a supersonic speed. This fully ionized and magnetized gas flow interacts with the Earth's magnetic field, resulting in a comet-shaped cavity (the magnetosphere) carved around the Earth, while the lines of force of the Earth's magnetic field and of the solar wind magnetic field interconnect. Electric power of as much as 10^{12} W is generated as the solar wind blows across the interconnected field lines near the comet-shaped boundary. A part of the electric current (carried mainly by electrons) thus generated flows between the magnetospheric boundary and an annular, ring-shaped region of the polar upper atmosphere along the lines of force of the Earth's magnetic field. *See* MAGNETOSPHERE.

Fig. 1. Ring-shaped auroral glow observed from above the north polar region by the Dynamics *Explorer* satellite. (*L. A. Frank and J. D. Craven, University of Iowa*)

Fig. 2. Two curtain-shaped auroras stretching across the sky near Fairbanks, Alaska. (*Courtesy of Lee Snyder*)

As these electrons descend toward the Earth, they themselves develop an electrical potential drop of the order of a few kilovolts along the lines of force. As a result, the current-carrying electrons acquire energies of as much as a few kiloelectronvolts, sufficient to ionize and excite a few hundred atoms and molecules before they are stopped by the atmosphere at an altitude of about 60 mi (100 km).

Two ring-shaped glows, one in each hemisphere, are produced by upper atmospheric atoms and molecules which emit their own characteristic light after colliding with the current-carrying electrons. The most common light of the aurora (the greenish-white light) comes from excited oxygen atoms. Excited and ionized molecular nitrogen adds several band emissions. Imaging devices aboard satellites have successfully "photographed" both the northern and southern auroral rings (**Fig. 1**). Satellite-borne and rocket-borne detectors and radio probing devices, as well as advanced optical instrumentation, have increased considerably the understanding of the nature of the discharge electrons and the interaction between the upper atmosphere and auroral electrons. *See* SCIENTIFIC AND APPLICATIONS SATELLITES.

From a point on the ground, for example, Fairbanks, Alaska, only a small part of the ring-shaped glow can be observed. It is seen as a curtain-shaped glow, stretching from the eastern to the western horizon across the sky (**Fig. 2**). The bottom of the auroral curtain is sharply bounded and is located at about 60 mi (100 km) altitude. The upper boundary diffuses and extends to well above 180 mi (300 km). The auroral curtains undergo drastic changes from time to time, such as a sudden increase or rapid fluctuations in brightness, rapid poleward motions, or curling motions of various scales. When such activity occurs all along the ring-shaped region, it is called an auroral substorm. It lasts for a few hours and occurs a few times a day. The auroral discharge current during an auroral substorm can reach an intensity of several million amperes and generates intense magnetic disturbances, which cause polar magnetic substorms.

Magnetospheric physicists have determined the solar wind quantities which control the efficiency of the solar wind–magnetosphere generator and thus of auroral activity. They are the solar wind speed, and the intensity and the orientation (with respect

to that of the Earth's magnetic field) of the solar wind magnetic field. Therefore, it has become possible to forecast auroral activity about one hour in advance by monitoring the solar wind at a distance of 200 earth radii upstream of the solar wind. The aurora becomes active during geomagnetic storms which occur often about 40 h after an intense solar flare. This is because the efficiency of the solar wind–magnetosphere generator becomes high and variable when the solar wind becomes gusty after a solar flare. During a great magnetic storm, the auroral ring expands from its usual latitude of about 67 to 50° or a little less. It is on such an occasion when the aurora can be seen widely across the continental United States. *See* ATMOSPHERE; ATOMIC STRUCTURE AND SPECTRA; GEOMAGNETISM; IONOSPHERE; PLASMA (PHYSICS); SOLAR WIND.

S.-I. Akasofu

Australia

An island continent situated in the Southern Hemisphere and extending from 10 to 44°S and from 113 to 153°E. Australia's total area (2,941,526 mi^2 or 7,618,517 km^2 for the Australian mainland and 26,383 mi^2 or 68,332 km^2 for the island of Tasmania) is somewhat less than that of the United States. Bounded on the west by the Indian Ocean and on the east by the Pacific Ocean, Australia straddles the Tropic of Capricorn. Originally part of the ancient supercontinent of Gondwanaland, which began to disintegrate 200 million years ago, Australia broke away from what is now Antarctica about 70 million years ago and began its long drift northward to its present position. Australia's long-term isolation has contributed to the uniqueness of its flora and fauna. Australia possesses about 650 species of birds and 400 species of reptiles, a large proportion of which are endemic. Most interesting are its 255 species of mammals, in particular the marsupials (lacking placenta), which include the kangaroo, wallaby, and koala; and the monotremes (egg-laying), which include the platypus and spiny anteater. These animals appear to represent stages along the evolutionary path toward fully developed placental mammals. *See* CONTINENTAL DRIFT; TROPIC OF CAPRICORN.

Australia's global position astride the southern tropic provides a climate that for much of the continent is characterized by clear skies, high temperatures, and relatively low rainfall (50% of its area has a median annual rainfall of less than 12 in. or 300 mm). The major cause of Australian aridity is the global subsidence of air at a latitude of about 30°S, warming as it sinks to produce belts of high pressure, clear skies, and low rainfall. Equatorward of the high-pressure belt, easterly trade winds blow back toward the Equator, while poleward generally westerly winds blow toward the higher-latitude low-pressure systems. While aridity dominates central Australia, coastal climates range from the seasonally wet monsoon tropics in the north to relatively moist temperate climates in the south (**Fig. 1**). *See* MONSOON METEOROLOGY; TROPICAL METEOROLOGY.

Australia is generally of remarkably low elevation and moderate relief. Seventy-five percent of the landmass lies between 600 and 1500 ft (180 and 450 m) in the form of a high plateau. A cross section from east to west shows first a narrow belt of coastal plain, then the steep escarpments of the eastern face of the Great Dividing Range, stretching 1200 mi (1900 km) from the north of Queensland to the south of Victoria (**Fig. 2***a*). The descent on the western slope of the Dividing Range is gradual until elevation in the inland basins is often below sea level. The land surface then rises gradually again across the great plateau until the low ranges of Western Australia fringing the plateau are reached, and beyond these lies another coastal plain. With the exception of the Gulf of Carpentaria and Cape York peninsula in the north and the Great Australian Bight in the south, there are few striking features in the configuration of the coast (Fig. 2*a*). Australia may conveniently be divided into three great physiographic regions, revealing characteristic geologic structure and major landform regions (Fig. 2*b*).

East Australian Highlands. A narrow plain extends north and south along the eastern coast, seldom more than 60–70 mi (100–110 km) in width, and occasionally only a few miles wide, but broader in the north than in the south. Flanking the plain are the series of ranges and tablelands making up the Great Dividing Range. In the south, one branch sweeps westward through Victoria, whereas the main branch continues due south, interrupted by the waters of Bass Strait, and terminates in Tasmania. Elevations are low, and the highest peak, Mount Kosciusko, is 7328 ft (2234 m). A few other peaks rise above 500 ft (1500 m) [**Table 1**]. This is a region of ancient mountains and of old and hard igneous and metamorphic rocks (dating mainly from the Paleozoic Era, 235–570 million years ago) that were raised in relatively recent geological times as a series of plateaus and low mountain blocks. The uplift introduced the rich deposits of gold, silver, lead, and copper that were a major stimulus to the early development of eastern Australia. The soils of the region are primarily shallow, except for much of Cape York and the area around Sydney. Active dissection by numerous short rivers has produced a broken surface of deep valleys and gorges on the eastern slopes, but few of these give access to the interior, for most terminate abruptly in rocky cliffs or turn and run parallel to the coast. Others run back to high plateaus or steep-sided, hazy-blue ranges. *See* PALEOZOIC.

TABLE 1. Australian mountains and summit elevations in feet (meters)

Australian Mainland	
Kosciusko (New South Wales)	7328 (2234)
Townsend (New South Wales)	7266 (2215)
Bogong (Victoria)	6508 (1984)
Feathertop (Victoria)	6306 (1922)
Bartle Frere (Queensland)	5438 (1658)
Tasmania	
Legge's Peak	5160 (1573)
Cradle Mountain	5069 (1545)

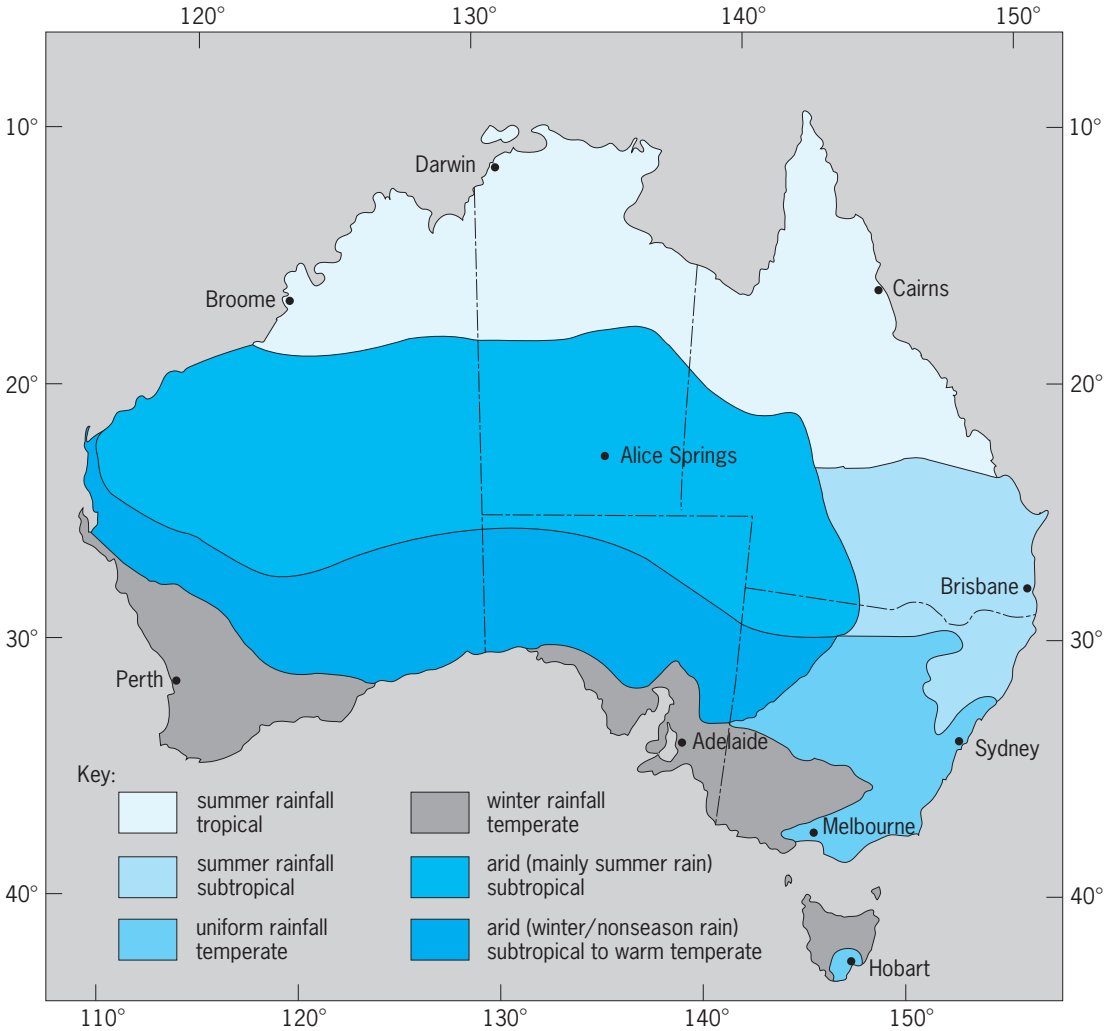

Fig. 1. Climate zones of Australia.

Located off the northeast coast and covering an area of 100,000 mi^2 (260,000 km^2) is the Great Barrier Reef, the largest living structure in the world. Extending 1300 mi (2000 km) from near the coast of Papua New Guinea to the Tropic of Capricorn, the reef actually comprises a chain of 2500 individual reef structures, many accumulating coral sand and debris to form small islands. Ribbon, or linear, reefs line the outer perimeter of the Great Barrier Reef at the continental shelf, at a distance ranging from 10 to 150 mi (16 to 240 km) offshore. There is concern about degradation of the central two-thirds of the reef as a result of the impact of tourism, agricultural runoff to the inshore waters, and infestations of the crown-of-thorns starfish which feed on growing coral. Partly in response to concerns about the health of the reef, it was declared a Marine Park in 1975 and was listed a UNESCO World Heritage Area in 1981. *See* REEF.

Climatically the region ranges from a tropical, primarily summer rainfall regime in the north (sometimes affected by the northwest monsoon) to a warm temperate, mainly uniform rainfall regime in the south (Fig. 1). This is the best-watered region in Australia, with annual rainfalls measuring generally above 25 in. (600 mm) and exceeding 250 in. (6000 mm) in some of the mountainous areas of the northeast. The generally sufficient rainfall is reflected in the extensive areas of gray-green *Eucalyptus* open forest and woodland (**Fig. 3**). Small pockets of tropical and temperate rainforest (closed forest) of more diverse floristic type occur along the east coast and in western Tasmania. Some of the river systems are of considerable size (**Table 2**). Large hydroelectric power plants have been developed in the East Australian Highlands and Tasmania. The elaborate Snowy River project, with a generating capacity of 3000 MW of electricity, also diverts water in tunnels under the mountains to the Murrumbidgee and Murray rivers, thus augmenting the supply of irrigation water in the drier interior of New South Wales and Victoria.

On the flanks of the East Australian Highlands are Australia's principal coal deposits, which are a major source of the nation's energy. Extensive black coal deposits were laid down in the Sydney and Bowen basins during the Permian Period (300–235 million years ago); and massive brown coal deposits of

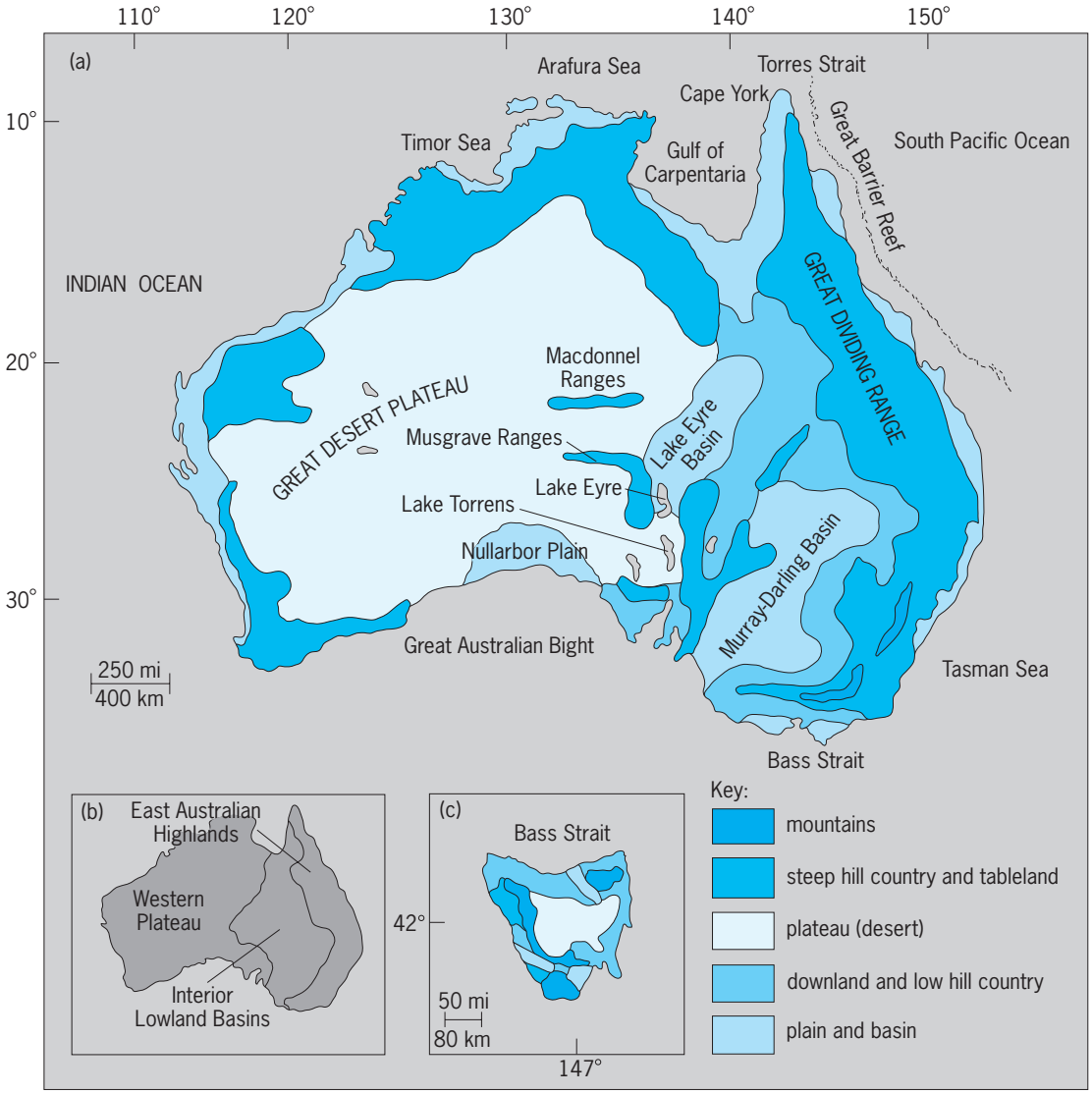

Fig. 2. Land surface characteristics of mainland Australia and Tasmania. (*a*) Landforms of Australia. (*b*) Sketch map of predominant geologic structure and major landform regions of Australia. (*c*) Landforms of Tasmania.

TABLE 2. Principal Australian rivers and lengths in miles (kilometers)

Australian Mainland	
Darling (Queensland and New South Wales)	1760 (2830)
Murray (Victoria, South Australia, and New South Wales)	1600 (2600)
Murrumbidgee (New South Wales)	1050 (1690)
Lachlan (New South Wales)	850 (1400)
Flinders (Queensland)	520 (840)
Diamantina (Queensland)	468 (753)
Murchison (Western Australia)	440 (710)
Burdekin (Queensland)	440 (710)
Bulloo (Queensland)	370 (600)
Mitchell (Queensland)	350 (560)
Tasmania	
South Esk	120 (190)
North Esk	45 (72)
Tamar (from confluence of North and South Esk)	40 (60)
Huon	105 (169)
Derwent	107 (172)
Gordon	90 (150)

Tertiary age (65–2 million years ago) lie in a great depression in the LaTrobe Valley of Victoria, south of the Great Divide and close to the coast of Victoria. Many important petroliferous basins lie adjacent to the major coal fields. The fields at Surat (Roma), flanking the divide in Queensland, and off the coast of Victoria in Bass Strait are among Australia's most productive deposits of petroleum and natural gas. *See* PERMIAN; PETROLEUM GEOLOGY; TERTIARY.

Tasmania. This small, relatively mountainous island lies 150 mi (240 km) southeast of the Australian mainland across Bass Strait. It is structurally similar to the East Australian Highlands and is a logical extension of that region. The dominant feature is the central plateau, falling from a general level of 3500 ft (1100 m) in the northwest toward the southeast. Its lake-studded surface is drained by the Derwent River (Table 2). To the west and south the plateau rises to 4000–5000 ft (1200–1500 m) in a range of rocky mountains, but to the north and east it is much lower

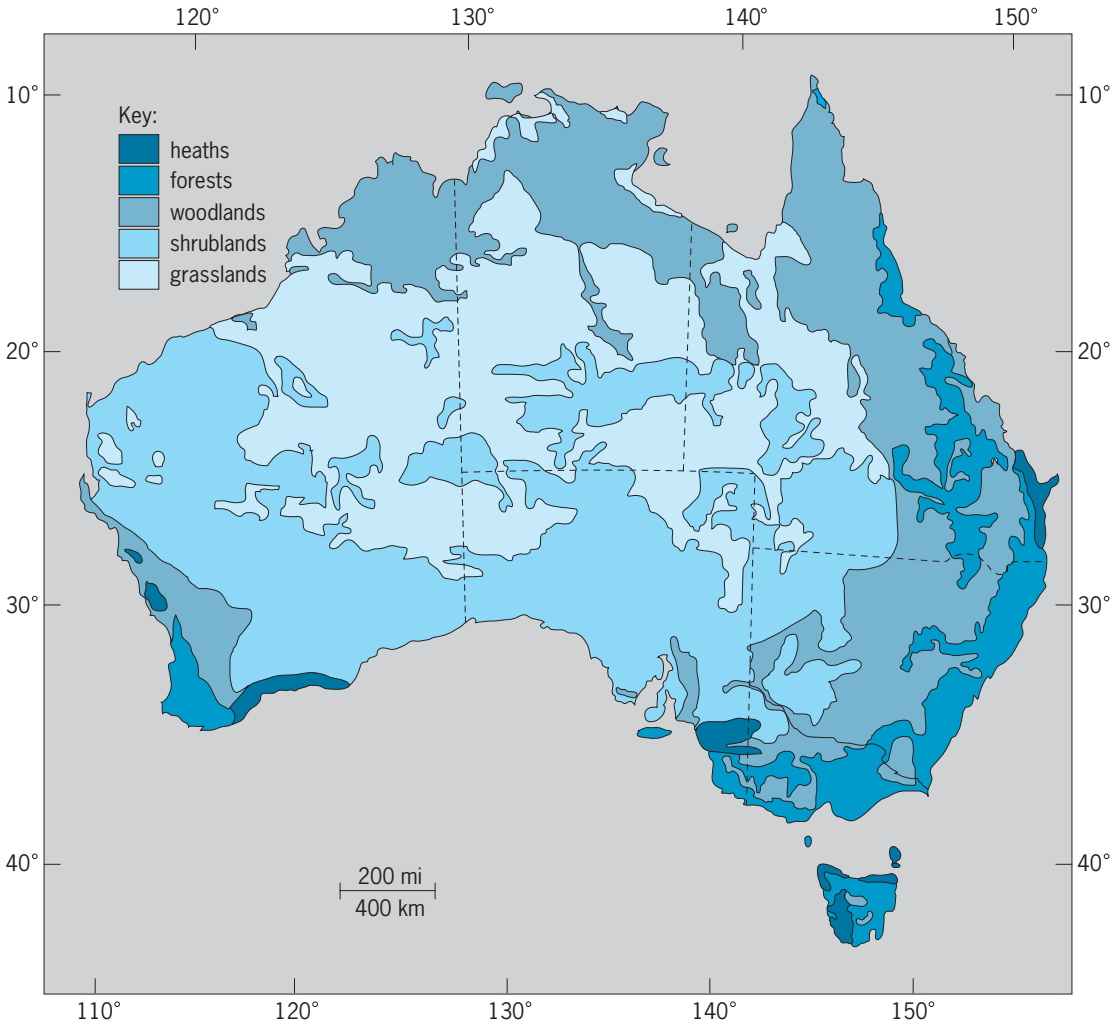

Fig. 3. Basic vegetation distributions in Australia. (*After R. M. Moore ed., Australian Grasslands, Australian National University Press, 1970*)

and terminates in bold scarps about 1000 ft (300 m) high which overlook the Tamar-Esk river lowlands. On the northeastern side of this valley lowland, an isolated block of rocky highlands includes the highest peak (5160 ft or 1570 m) in Tasmania (Table 1). In the southwest, the Gordon and Huon river valleys separate the mountains fringing the plateau from another outlying series of low mountains rising about 3500 ft (1100 m). The Tasmanian climate is temperate, with dominantly winter rainfall. A dense *Eucalyptus* forest and woodland covers most of the island except along the wetter west coast (rainfall exceeds 120 in. or 3000 mm in places), where dense beech forest predominates. Some coastal regions have extensive areas of heath vegetation. The rivers have short, rapid courses with relatively little seasonal variation in flow. Significant and exploitable mineral deposits of copper, lead, zinc, and tin occur, especially in northwest Tasmania.

Interior Lowland Basins. The inland comprises a region of mainly Mesozoic and Cenozoic (235 million years ago to the present) sedimentary rocks, surficial deposits, and deeply weathered mantle, occupying one-third of the continent between the western slope of the East Australian Highlands and the inner eastern margin of the ancient shield which forms the Western Plateau (Fig. 2*a* and *b*). Little land is over 500 ft (150 m), and some (in the region of Lake Eyre) is below sea level. *See* BASIN; CENOZOIC; MESOZOIC; SEDIMENTARY ROCKS.

Low rises separate three extensive drainage systems: One drains north to the Gulf of Carpentaria; the Lake Eyre Basin is one of inland drainage; and the third, the Murray-Darling Basin, is drained to the south. Except where the low Flinders, Barrier, and Mount Lofty ranges sharply break the surface in the south, the lowland is a dry and monotonous sandy-gravelly plain, crossed more often by broad, incised, dry river beds with occasional oxbow lakes (billabongs) than by rivers of running water. The rivers of the Murray-Darling Basin, draining the western slopes of the Great Dividing Range, have a marked seasonal variation in flow but never dry up in the lower reaches. South Australia's shallow lakes are more often dry expanses of encrusted white salt than bodies of water—the result of low rainfall and high

TABLE 3. Principal Australian lakes and areas in square miles (square kilometers)

Australian Mainland	
Eyre (South Australia)*	2970 (7690)
Torrens (South Australia)*	2230 (5780)
Mackay (Northern Territory and Western Australia)*	1990 (5100)
Gairdner (South Australia)*	1840 (2960)
Frome (South Australia)*	930 (2400)
Argyle (Western Australia)	430 (1100)
Corangamite (Victoria)	90 (230)
Tyrrell (Victoria)*	67 (170)
George (New South Wales)	60 (160)
Menindee (New South Wales)	60 (160)
Nerran (New South Wales)	55 (140)
Tasmania	
Great Lake	44 (110)
Sorell	19 (49)
St. Clair	15 (39)
Arthur	14 (36)
Echo	12 (31)

*Shallow salt pans, normally or seasonally dry.

evaporation (**Table 3**). In most parts of the region, water from deep artesian wells is available. *See* ARTESIAN SYSTEMS.

The climate of much of the region is arid, with annual rainfalls as low as 4 in. (100 mm) in the area around Lake Eyre. However, the coastal margins of the Gulf of Carpentaria exhibit a summer monsoonal climate, and in the far south of the region a temperate, moderate winter rainfall regime is present (Fig. 1).

Despite generally low rainfall, very little of the region is completely devoid of vegetation. In the south of the area, pockets of mallee heathland (dominated by low-growing *Eucalyptus*) predominate in the sand country of northwest Victoria and southeast South Australia. Much of the remainder of the Interior Lowland is covered in *Acacia* shrubland and grassland (Fig. 3), with patches of *Eucalyptus* savanna-woodland appearing in the north Gulf region and along the west coast of Cape York. *See* SAVANNA.

Minerals of importance in this region lie mainly along its margins, with significant deposits of gold, copper, lead, and zinc at the southeastern edge of the Barkly Tableland in the area of Mount Isa, and in the southern Barrier Range around Broken Hill. Extensive bauxite deposits occur on the west coast of Cape York, from where the raw material is transported to alumina refineries and aluminum smelters in Australia and overseas. Very significant natural gas fields dating from the Permian (300–235 million years ago) occur in the Cooper Basin in the Lake Eyre region. *See* BAUXITE; NATURAL GAS.

Western Plateau. This largest area occupies almost three-fifths of the continent; it is a great shield of ancient rocks, many of which date from the Archean Era (3600–2500 million years ago), standing 750–1500 ft (230–460 m) high. Sedimentary and low-grade metamorphic rocks dominate, but there are significant outcrops of high-grade metamorphic and igneous rocks. Some of the earliest evidence of life on Earth, microbe-sized fossils dating back 3465 million years, comes from the northwest of the Western Plateau. Much of the plateau is buried in deeply weathered desert sand, and only a few ridges of ancient mountains (such as the Macdonnel and Musgrave ranges) break the monotony of the surface. The sand is arranged in ridges as high as 50–60 ft (15–20 m), running remarkably parallel for great distances northwest to southeast. The ridges are more permanent features of the arid landscape than are the migratory dunes of other deserts. Local ephemeral streams form a net of inland drainage to scattered salt-lake basins. The limestone Nullarbor Plain, skirting the Great Australian Bight, is both treeless and riverless. Only in the southwestern corner of the continent (a temperate Mediterranean-like climate) and along the northwestern coast (a tropical monsoon climate) is rainfall sufficient to support a sclerophyll forest of *Eucalyptus* and a monsoon *Eucalyptus* woodland, respectively, each giving way inland to a succession of savanna, desert *Acacia* scrub, and dry grassland. Significant areas of heath occur in coastal regions of the southwest. In the north, coastal rivers are of considerable size, but they change from flooded torrents after seasonal summer rains to a succession of water holes in the winter dry season. In the same area, Lake Argyle, an engineered lake formed on the Ord River as part of an ambitious irrigation scheme, is Australia's largest fresh-water lake (Table 3). *See* ARCHEAN.

It is in this western region that many of the great postwar discoveries of minerals have been made that have helped transform the Australian economy. Chief among these is one of the world's greatest series of deposits of high-grade iron ore. To exploit them, settlements and networks of communications have been established in the hitherto unoccupied arid northwest of Western Australia, in the vicinity of the Hamersly and Opthalmia ranges; and ports have been constructed to facilitate the export of the ore to Japan and to transport ore to processing plants elsewhere in Australia. Significant gold deposits are located throughout the southern Western Plateau, and major oil and gas fields are located off the northwest coast in the Carnarvon Basin. Smaller gas and oil fields have been located west of Darwin, and in the Amadeus Basin in the south of the Northern Territory. Since the development of the Argyle deposit in the Kimberley region of the northern Western Plateau, Australia has become the world's largest diamond producer. *See* DIAMOND; ORE AND MINERAL DEPOSITS.

Nigel J. Tapper

Bibliography. Australian Surveying and Land Information Group, *Atlas of Australian Resources*, vol. 5: *Geology and Minerals*, 1988, vol. 6: *Vegetation*, 1990; D. N. Jeans (ed.), *Australia: A Geography*, vol. 1: *The Natural Environment*, 2d ed., 1986; D. Johnson, *The Geology of Australia*, 2005; State of the Environment Advisory Council, *Australia State of the Environment*, 1996; A. P. Sturman and N. J. Tapper, *Weather and Climate of Australia and New Zealand*, 1996.

Australopithecine

Any of the seven species that belong to the family Hominidae (comprising humans and their closest relatives). These species are not attributable to the genus *Homo*, but belong to at least three genera that existed between about 4.4 million years ago (Ma) and 1.2 Ma during the Pliocene and early Pleistocene epochs. All seven species are known only from Africa. Although some workers regard all of them as belonging to one genus, *Australopithecus*, it is clear that a second genus, *Ardipithecus*, should be recognized for the earliest known hominid fossils, and that the three "robust" species belong to a third genus, *Paranthropus*.

Fig. 1. Lateral view of the Taung skull of *Australopithecus africanus*, the type specimen of *Australopithecus* and the first early hominid specimen to be discovered in Africa. (*Courtesy of F. E. Grine*)

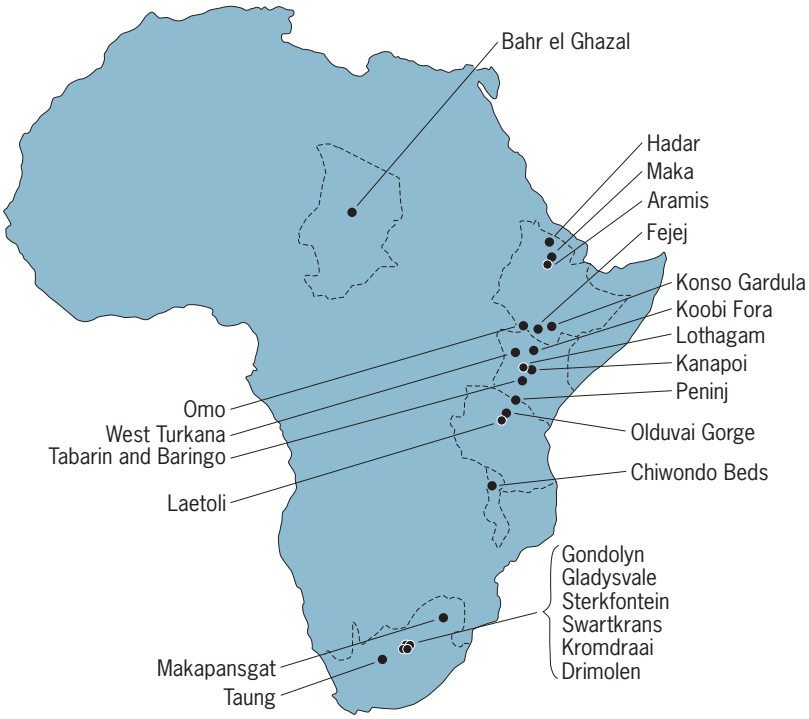

Fig. 2. Principal sites at which fossils of *Australopithecus* have been discovered in Chad, Ethiopia, Kenya, Tanzania, and South Africa.

Historical context. The name australopithecine comes from the taxon *Australopithecus* ("southern ape") *africanus*, which was coined by Raymond Dart for a fossil skull discovered in 1924 in Taung, South Africa (**Fig. 1**). The Taung skull had several distinctly hominid, or humanlike, features, but the claim that *A. africanus* was a human forebear was disputed by many of the leading paleoanthropologists of that time. The hominid status of *A. africanus* became widely accepted more than a decade later, largely because of work on *Australopithecus* fossils from Sterkfontein. Specimens of *A. africanus* are known also from the sites of Makapansgat and Gladysvale in southern Africa (**Fig. 2**). Faunal comparison with radiometrically dated sites in eastern Africa indicates that this species existed between about 3.0 and 2.3 Ma.

The name *Paranthropus* ("beside human") was coined in 1938 when fossils from Kromdraai, South Africa, were attributed to the taxon *P. robustus*. Fossils of this species are also known from the nearby sites of Swartkrans and Drimolen. These bones are dated, also by faunal comparisons with radiometrically dated sites in eastern Africa, to between about 1.8 and 1.5 Ma.

In 1959, a cranium with very large cheek teeth was discovered at Olduvai Gorge, Tanzania (Fig. 2). It represents a separate species of *Paranthropus*, *P. boisei*. Numerous *P. boisei* fossils have been discovered in Tanzania at Olduvai Gorge and Lake Natron, in southern Ethiopia along the Omo River (Shungura Formation), and especially in Kenya on the eastern side of Lake Turkana (Koobi Fora Formation). Fossils of *P. boisei* are dated to between 2.3 and 1.2 Ma.

In 1935, the fossil site of Laetoli was discovered south of Olduvai Gorge, but it was only during the 1970s that most of the hominid jaws and teeth were found there among volcanic ash layers that preserve spectacular hominid footprint trails. Between 1973 and 1977, numerous specimens, including the partial skeleton known as Lucy, were recovered from the Hadar Formation sediments in Ethiopia. It was proposed that the Laetoli and Hadar fossils belonged to a primitive species of *Australopithecus* named *A. afarensis*. *Australopithecus afarensis* fossils are known also from Maka in the Middle Awash of Ethiopia, and from Koobi Fora, Kenya. Fossils that possibly belong to *A. afarensis* (although their specific attribution is open to question) are known from the Ethiopian sites of Belohdelie and Fejej. *Australopithecus afarensis* is dated radiometrically to between 3.7 and 2.9 Ma; it may extend back to 4.0 Ma if the specimens from Belohdelie and Fejej belong to it.

In 1993, the anterior part of a mandible and an upper premolar fragment were discovered in the Bahr el Ghazal region of Chad (Fig. 2). The associated fauna includes several taxa found also in the Hadar Formation of Ethiopia, indicating an age of some 3.0

to 3.4 Ma. It has been suggested that the Chadian fossils represent a separate species, *A. bahrelghazali.* However, the Chad specimens do not afford a definitive diagnostic differentiation from *A. afarensis*; thus, as a conservative approach, the Bahr el Ghazal fossils are considered here as being tentatively attributable to *A. afarensis*.

In 1985, the discovery of a nearly complete skull in the Nachukui Formation on the western side of Lake Turkana, Kenya, led to the recognition of a third "robust" australopithecine species, *P. aethiopicus*. A handful of fossils attributable to this species are known also from southern Ethiopia (Shungura Formation). This species is known from about 2.7 to about 2.3 Ma.

Fossils found in the Aramis region of Ethiopia in 1992 and 1993 were recognized as belonging to a separate taxon, *Ardipithecus* ("ground ape") *ramidus*. It is the most primitive hominid species known. In some morphological features, it is more similar to apes than to other, later hominids. Most of the Aramis bones date to 4.4 Ma. Three fragmentary fossils from Kenya may be attributable to *Ar. ramidus*. One is from the site of Lothagam; it is likely between 5.0 and 5.5 million years old. The other two are from Tabarin and Baringo. Volcanic rocks beneath the Tabarin jaw are 5.1 million years old, and the fauna provides an upper age of 4.2 million years for the hominid. The fauna from Baringo also predates 4.2 Ma.

In 1995, the species *A. anamensis* was named to accommodate fossils that were found at the Kenyan sites of Kanapoi and Allia Bay. These specimens date to between about 3.9 and 4.2 Ma.

Species. Thus, at least seven australopithecine species, belonging to at least three genera, can be recognized in the Pliocene and early Pleistocene of Africa:

Ardipithecus ramidus	(?5.0–4.4 Ma)
Australopithecus anamensis	(3.9–4.2 Ma)
Australopithecus afarensis	(?4.0–2.9 Ma)
Australopithecus africanus	(3.0–2.3 Ma)
Paranthropus aethiopicus	(2.8–2.3 Ma)
Paranthropus boisei	(2.3–1.2 Ma)
Paranthropus robustus	(1.8–1.5 Ma)

All are defined on the basis of craniodental morphology.

Ardipithecus ramidus. This is the most primitive hominid species known. It has large canine teeth coupled with comparatively small molars. The anterior lower deciduous premolar is chimpanzeelike, and the anterior permanent premolar is asymmetrical, dominated by a single large cusp. It is unique among all hominids in having thin tooth enamel. The temporal bone has a very shallow glenoid fossa (jaw joint) without a distinct articular eminence; the external auditory meatus is small, the tympanic bone is tubular, and its lateral edge is positioned far laterally. The temporal bone is extensively pneumatized (filled with air cells). The upper limb bones display some features in common with modern humans, as well as characteristics found in living apes. The latter include a large styloid process at the end of the radius, a strong elbow-stabilizing mechanism of the radius and humerus, and strong muscle ridges. These features almost certainly indicate that *Ar. ramidus* was an accomplished climber.

The Aramis fauna is dominated by primates, and arboreal colobine monkeys are the most common of these. The mammalian remains and the presence of fossil seeds suggest that *Ar. ramidus* occupied a wooded habitat.

Australopithecus anamensis. This species has several primitive features in common with *Ardipithecus*, but it also exhibits some derived traits in common with taxa such as *A. afarensis*. It has a few unique characteristics. Thus, the anterior lower permanent premolar is virtually single-cusped and asymmetrical, and the canine crowns and especially the roots are very large (at least in presumptive males). The temporal bone is also extensively pneumatized, the glenoid fossa is shallow without a distinct articular eminence, the tympanic bone is tubular, and the external auditory meatus is small. The palate is shallow anteriorly. The postcanine tooth rows are nearly parallel to one another and close together. The mandibular symphysis, which has a marked postero-inferior slope, extends as far back as the first molar tooth. Its molar crowns have sloping sides, but its tooth enamel appears to be similar in thickness to that of other *Australopithecus* species.

The tibia has features that clearly indicate that this species was a biped. It is a large bone; its owner is estimated to have weighed 47–55 kg (103–120 lb). A wrist bone (the hamate) has a long hook, suggestive of large, powerful hand flexors running through the carpal tunnel. The radius and distal humerus have several apelike characters that are associated with vertical climbing. These bones suggest that *A. anamensis* was a terrestrial biped that was also heavily engaged in tree climbing. Current reconstructions of the paleoecology at Kanapoi and Allia Bay are consistent with this hypothesis. The fauna suggests a woodland-bushland habitat along a large river (at Kanapoi), and a gallery forest associated with a large river that formed a local delta into a large lake (Lake Lonyumun).

Australopithecus afarensis. This species is characterized primarily by a suite of primitive craniodental features, including a strongly protruding jaw (prognathism), a flat (unflexed) cranial base, and a relatively flat glenoid fossa without a distinct articular eminence. It has a tubular tympanic bone, although the external auditory meatus is large. The parietal mastoid angle is strongly flared, and there is close approximation of the cranial points lambda and inion at the back of the skull. The palate is anteriorly shallow, the mediolaterally and supero-inferiorly convex nasoalveolar clivus is demarcated from the floor of the nose by a horizontal sill, and the upper lateral incisor roots are situated to the side of the lateral walls of the nasal aperture. The canines are relatively large, and the anterior lower premolar tends to be asymmetrical and dominated by a single large cusp. The teeth are relatively large in comparison to body

size. Endocranial capacity ranges about 310–500 cm^3 (19–30 $in.^3$). In addition, *A. afarensis* has a high incidence of an occipital-marginal sinus by which blood is drained from the brain.

There is a high degree of (presumed sexual) dimorphism in body size. Estimates range 30.5–80.5 kg (67–177 lb), which is similar to body size ranges for gorillas and orangutans.

In 1978–1979, volcanic-ash layers bearing hominid footprint trails were discovered at Laetoli. These trails and the bones of the postcranial skeleton indicate that *A. afarensis* was a terrestrial biped, but its gait was not exactly the same as that of modern humans. It was not able to employ a full striding gait. Furthermore, a host of postcranial features, such as strongly curved finger bones, a superiorly oriented shoulder joint, a relatively long upper limb, and a relatively long forearm, indicate that *A. afarensis* was well adapted to arboreal climbing. This species almost certainly spent a considerable amount of time in an arboreal milieu, perhaps feeding and sleeping. The knee joints from Hadar suggest that the smaller (presumptive female) and larger (presumptive male) individuals varied in their reliance on arboreality, such as shown by orangutans.

Paleoenvironmental reconstructions indicate diverse habitats, from well-watered and wooded, to forested environments with edaphic grasslands (at Hadar), to more woodland-bushland conditions (at Laetoli). *Australopithecus afarensis* probably had a fairly broad range of locomotor abilities and dietary proclivities.

Australopithecus africanus. This is the type species of the genus *Australopithecus*. Its cranium is more globular and less pneumatized than that of *A. afarensis*. In comparison to *A. afarensis*, this species has a more steeply inclined forehead, a greater separation of the cranial points lambda and inion at the back of the skull, and a high glabella that is separated from the nasion. The nasoalveolar clivus is prognathic but flattened, the palate is anteriorly deep (shelved), and the maxillary lateral incisor roots are situated medial to the lateral margins of the nasal aperture. The glenoid fossa is deep, and there is a distinct articular eminence. The tympanic plate is vertically disposed, and the external auditory meatus is large. The lower anterior premolar is bicuspid. Endocranial capacity estimates for eight specimens range 428–515 cm^3 (26–31 $in.^3$), although the latter value is likely underestimated. In contrast to *A. afarensis*, *A. africanus* has a high incidence of the transverse sinus route for blood drainage from the brain.

Dental and skeletal dimensions indicate that sexual dimorphism in *A. africanus* was also of considerable magnitude. Body weight estimates range 33.5–67.5 kg (67–149 lb) for presumptive females and males. Postcranial elements are generally similar to those of *A. afarensis*, suggesting that *A. africanus* retained adaptations for tree climbing along with the adaptations for bipedal terrestrial locomotion. Some features, such as foot bone structure indicating a greater range of motion of the big toe, limb proportions indicating a relatively long upper limb, and the structure of the proximal end of the tibia, suggest that the skeleton of *A. africanus* may be somewhat more primitive (chimpanzeelike) than that of *A. afarensis*.

Paleoenvironmental reconstructions suggest that *A. africanus* inhabited well-watered environments with notable bush and tree cover. The microscopic details of tooth wear indicate that the diet of *A. africanus* consisted principally of fruits and leaves.

Paranthropus robustus. The three species in the genus *Paranthropus* are characterized by cranial and dental features indicative of powerful chewing. Thus, the *Paranthropus* face is flattened and somewhat "dished," with the cheeks located anterior to the margins of the nasal aperture. The facial skeleton is hafted high onto the neurocranium, which results in a low forehead with a concave frontal trigone. The prominent glabella is situated below the level of the supraorbital margin and is nearly coincident with the nasion. The nasoalveolar clivus is flattened and grades imperceptibly into the floor of the nose. The palate is extremely thick. The cheek bones are very robust, and males possess a sagittal crest from which the large temporalis muscles originate. In addition, the petrous portion of the temporal bone is coronally inclined, and the bulbous mastoid region of the temporal bone is laterally inflated. The premolars are fully bicuspid, and they and the molars are greatly enlarged. Tooth enamel is exceptionally thick.

Paranthropus robustus is the type species of the genus *Paranthropus*. It is further characterized by the presence of a triangular depression on the front of the cheek bone (maxillary trigone), a palate that is anteriorly shallow, and a deep glenoid cavity with a distinct articular eminence. The tympanic bone is a vertically deep plate, and the external auditory meatus is very large. An endocranial capacity estimate of some 530 cm^3 (32 $in.^3$) has been recorded for the single reliable specimen. The upper canines are in the same coronal plane as the incisors, and the anterior teeth not only are absolutely smaller than those of *Ardipithecus* and *Australopithecus* but are especially diminutive in relation to the sizes of the premolars and molars.

Analysis of the postcranial skeleton of *P. robustus* is complicated by the presence of specimens of the genus *Homo* in the same deposits at Swartkrans. Nevertheless, it appears that *Paranthropus robustus* bones share some features with those of *A. afarensis* and *Australopithecus africanus*. These features include a relatively small femoral head with a long and anteroposteriorly flattened femoral neck. The morphology of the proximal radius is like that seen in earlier australopithecine species and living great apes, suggestive of arboreal capabilities. Foot bones indicate bipedal terrestrial locomotion, with evidence of a more humanlike foot than is present in either *A. afarensis* or *A. africanus*. It has also been suggested that the hand of *P. robustus* may have been more similar to that of modern humans, indicating that *P. robustus* had the manual capability to manufacture stone tools. Body size estimates for *P. robustus* range 42.5–65.5 kg (93.5–144 lb); these values suggest that *P. robustus* may not have been larger,

on average, than either *A. afarensis* or *A. africanus*. Size (presumed sexual) dimorphism in cranial, dental, and postcranial elements of *P. robustus* appears to be somewhat less prominent than in the other two species of *Paranthropus* (*P. boisei* and *P. aethiopicus*) or in species of *Australopithecus* (such as *A. afarensis* and *A. africanus*). However, it is unclear whether this apparent pattern reflects a specific reduction of body size dimorphism in *P. robustus*, or whether it is a taphonomic artefact that reflects the preferred prey size of the predators (such as leopards) that were primarily responsible for the accumulation of *P. robustus* remains.

Paleoenvironmental reconstructions suggest an open habitat with a nearby river that probably supported a woodland or gallery forest, and also more open, wooded grassland or savanna conditions during the accumulation of *P. robustus* remains. Tooth wear suggests that the diet of *P. robustus* was composed of harder items than were chewed by *A. africanus*. Studies of the carbon isotope signatures of *P. robustus* teeth indicate that this species had an overall reliance on C3-based foods (that is, trees, shrubs, forbs, and tubers). Overall, the cranial and dental anatomy of *P. robustus* indicates a primary adaptation to the mastication of foods that required the application of powerful chewing forces.

Paranthropus boisei. Most of the morphological features that characterize this species are also possessed by *P. robustus*, although the dental and bony characteristics that are associated with mastication appear to be even more exaggerated in *P. boisei*. Thus, *P. boisei* differs from *P. robustus* in having a deeper maxilla, an anteriorly deep (shelved) palate, the absence of a maxillary trigone, zygomatics that tend to be laterally bowed with a visorlike configuration of the cheek, and a heart-shaped foramen magnum with a straight or posteriorly convex anterior margin. The cheek teeth of *P. boisei* tend to be larger than those of *P. robustus*; indeed, *P. boisei* premolars and molars are the largest of any hominid taxon. The canines and incisors, however, are diminutive. Endocranial capacity estimates range about 500–530 cm^3 (30–32 $in.^3$) with an average of approximately 515 cm^3 (31 $in.^3$). Thus, as in *P. robustus*, the brains of *P. boisei* specimens tend to be somewhat larger than those of most *Australopithecus* specimens. *Paranthropus boisei* has a high incidence of an occipital-marginal sinus by which blood is drained from the brain.

Although cranial and dental remains of *P. boisei* are quite abundant, there are comparatively few postcranial bones that can be attributed to this species with certainty. The bones that are reasonably attributed to *P. boisei* reveal a picture of a biped that retained adaptations to arboreal locomotion (for example, it had relatively long arms, such as displayed by *Australopithecus anamensis*, *A. afarensis*, and *A. africanus*, and an apelike proximal radius morphology that suggests a long lever arm of the biceps muscle as is found in *Australopithecus*). The ankle bones (talus and calcaneus) are generally humanlike, and the femur has a relatively small head coupled with a long and anteroposteriorly flattened neck.

Cranial and postcranial elements indicate considerable sexual dimorphism. Body weight estimates from more complete bones range 52–60 kg (114–132 lb), although other, less complete postcranial elements indicate a greater range, from less than 30 kg (66 lb) to greater than 100 kg (220 lb).

Paleoenvironmental reconstructions reveal a variety of potential habitats, including scrub woodlands with extensive wetlands, open woodlands with edaphic grasslands (in which the soil has been periodically inundated with water, for example, a bank overflow), riparian woodlands, and grassland or shrublands, with a possible preference for well-watered sites such as riverine gallery forests. Overall, it appears that *P. boisei* preferred fairly open habitats (woodland to scrub woodland) associated with water and edaphic grasslands. Wear on *P. boisei* cheek teeth suggests the mastication of some hard items, although they may have played less of a role than in the dietary regimen of *P. robustus*. The diet of *P. boisei* may have required the prolonged chewing of tough or fibrous vegetable foods, perhaps with relatively little nutritional value.

Paranthropus aethiopicus. This species is known from comparatively few fossils (one fairly complete cranium, two incomplete mandibles, and a handful of isolated teeth). Like *P. robustus* but unlike *P. boisei*, this species possesses a maxillary trigone and an anteriorly shallow palate. Like *P. boisei* specimens, *P. aethiopicus* has a heart-shaped foramen magnum with a straight anterior margin. The premolars and molars of *P. aethiopicus* are similar in size to those of *P. boisei*; thus, they tend to be larger than most *P. robustus* cheek teeth.

The cranium of *P. aethiopicus* differs from, and is more primitive than, those of *P. boisei* and *P. robustus*, in that it possesses a long, unflexed base, marked alveolar prognathism, a shallow glenoid fossa without a distinct articular eminence, and a strongly flared parietal mastoid angle. In addition, the anterior teeth of *P. aethiopicus* may have been larger than those of the other two *Paranthropus* species. The endocranial capacity of the single known (male) cranium is 410 cm^3 (25 $in.^3$), which is noticeably smaller than estimates for *P. boisei* and *P. robustus* crania. In addition, *P. aethiopicus* appears to have employed the transverse sinus rather than the occipital-marginal sinus to drain blood from the brain.

Cranial remains reveal considerable size (possibly sexual) dimorphism in *P. aethiopicus*. Postcranial bones that may belong to *P. aethiopicus* include a large ulna that is notable for its considerable length and substantial dorsoventral curvature, features also found in apes and associated with arboreal climbing. It seems safe to assume that the postcranium of *P. aethiopicus* will exhibit those characters that are possessed in common by other early hominid species, including considerable differentiation in size, features related to bipedal terrestrial locomotion, and other characteristics associated with arboreal climbing.

Evolutionary relationships. Paleoanthropologists disagree over the assignment of early hominid fossils to different genera and species. Such arguments

over what is known as Alpha Taxonomy are to be expected, as different workers view the fossil record from different philosophical perspectives. Such differences also account for disagreements over the phylogenetic relationships of these species, including the issue of which (if any) is most closely related to the human genus, *Homo*. Every phylogenetic hypothesis that has been put forward since the 1950s has been either falsified outright or at least substantially altered by ongoing research and new discoveries.

At present, no scientifically rigorous phylogenetic analysis has been undertaken that incorporates all seven of the species discussed above. The most comprehensive study of "australopithecine" evolutionary relationships to date is that by David Strait and colleagues (1997). That study does not include *Ar. ramidus* or *A. anamensis*. Nevertheless, it is evident from the descriptive account of *Ar. ramidus* that it very likely resembles the stem hominid taxon in its morphology (**Fig. 3**). *Australopithecus anamensis* shares some evolved (derived) traits with later species such as *A. afarensis*, but retains some primitive features that are displayed also by *Ar. ramidus*. Thus, *A. anamensis* most likely evolved from a species that was at least morphologically similar to *Ar. ramidus* in some respects. For the moment, *Ar. ramidus* represents the best candidate for the ancestor of *A. anamensis*. *Australopithecus anamensis*, in turn, possesses some unique features that make it unlikely to be the immediate ancestor of *A. afarensis*, but it is probable that *A. afarensis* evolved from a species that had a strong morphological resemblance to *A. anamensis*.

Australopithecus afarensis likely gave rise to a lineage that provided the ancestry of both *A. africanus* and another lineage that included the common ancestor of the genera *Paranthropus* and *Homo*.

Although *A. africanus* shares a number of derived morphological characters with species that are part of the *Paranthropus* and *Homo* lineage, it is not considered to be directly ancestral to that line because it exhibits derived morphology in several characters (for example, glenoid fossa depth and anterior palatal depth) that are more primitive in both *A. afarensis* and *P. aethiopicus*. It is perhaps more likely that the derived traits which *A. africanus* shares with some species of *Paranthropus* and *Homo* were evolved in parallel. Other workers, however, have argued cogently that *A. africanus* constitutes a reasonable morphotype for the last common ancestor of the lineage that leads to *Homo* and *Paranthropus*. In this case, the primitive features displayed by *P. aethiopicus* would represent evolutionary reversals. At present, then, it is safest to conclude that the phyletic position of *A. africanus* remains ambiguous.

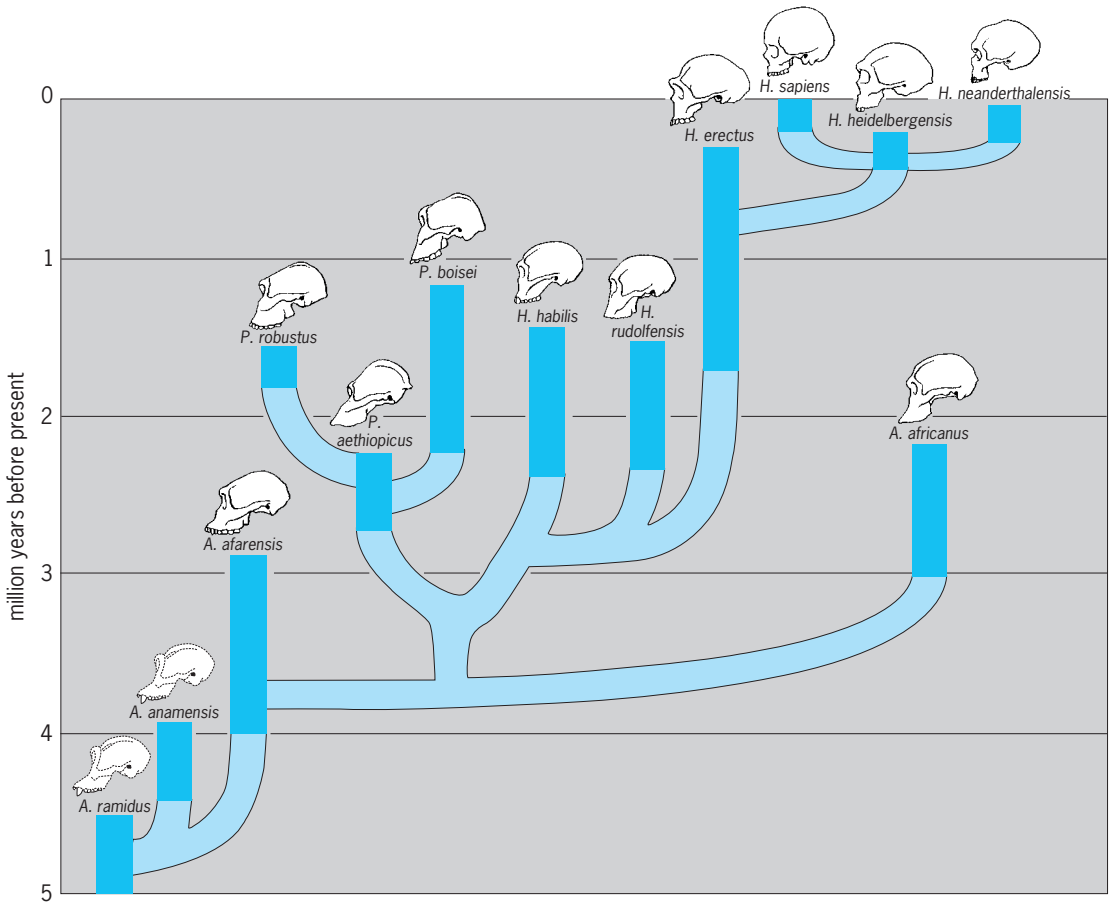

Fig. 3. Phylogenetic tree showing hypothesized evolutionary relationships among species of *Australopithecus*, *Parathropus*, and *Homo*.

Paranthropus aethiopicus is considered to be a likely candidate for the ancestry of both *P. boisei* and *P. robustus*. The reason is that *P. aethiopicus* shares a number of primitive features with *A. afarensis*, but at the same time it shares a host of derived features with the later species of *Paranthropus*, namely *P. robustus* and *P. boisei*.

The lineage leading to *Paranthropus* shares a number of morphological features with that leading to *Homo*. For example, all species of *Homo* and *Paranthropus* share a coronally oriented petrous temporal bone, a foramen magnum that is roughly horizontal in disposition, and a vertically oriented mandibular symphysis. Thus, it is most parsimonious to assume that these two lineages shared a common ancestor at some time prior to 2.8 Ma.

The apparent increase in cranial capacity that is shown by some species of *Paranthropus* (*P. robustus* and *P. boisei*) would appear to parallel the increase in brain size that characterized the evolutionary history of the lineage that led to *Homo sapiens*. Tools made of stone or bone are not known to be associated with *Ardipithecus* or the three species of *Australopithecus*. Stone tools are known from sites that contain *P. boisei* fossils, and both bone and stone tools are known from sites that preserve *P. robustus* remains. However, early members of the genus *Homo* are known also from these same localities. Thus, it is difficult to determine whether *Paranthropus* species may have been responsible for some of the late Pliocene and early Pleistocene archeological record. Indeed, it has been argued that the later species of *Paranthropus* may have been driven to extinction through competition with early *Homo*, because the latter possessed a distinct ecological advantage through the utilization of lithic technology in the procurement of food. While the evidence for this is not compelling, it is possible that ecological interactions between *Paranthropus* and early members of *Homo* may have influenced the evolutionary course of the human genus. *See* APES; FOSSIL HUMANS.

Frederick E. Grine

Bibliography. L. Aiello and C. Dean, *An Introduction to Human Evolutionary Anatomy*, Academic, New York, 1990; D. S. Strait, F. E. Grine, and M. A. Moniz, A reappraisal of early hominid phylogeny, *J. Human Evol.*, 32:17–82, 1997; M. H. Wolpoff, *Paleoanthropology*, 2d ed., McGraw-Hill, New York, 1999.

Authigenic minerals

Minerals that are formed in sediment or a sedimentary rock. Their in-place origin distinguishes them from minerals that are formed elsewhere and transported to the site of deposition (detrital minerals). Authigenic minerals form at the Earth's surface as well as during subsequent burial. The postdepositional processes are referred to as diagenesis, and the resulting minerals are important clues to postdepositional physical and chemical changes in the rock. *See* DIAGENESIS; SEDIMENTARY ROCKS.

Mode of formation. Authigenic minerals precipitate from the overlying water column, pore fluids in the sediment, recrystallization or alteration of preexisting minerals, and structural transformation of one mineral to another. The minerals change in an attempt to equilibrate to the physical and chemical conditions present at any given time. Critical factors in their formation are initial mineral assemblage, temperature, pressure, ionic concentration, pH, electron availability, and the fluid flux through the rock.

In sedimentary rocks it is common to find a record of multiple diagenetic events based on the authigenic minerals. For example, in sediments near the surface, meteoric water may displace original marine pore water, resulting in distinct types of cements. Iron oxide can result from oxidizing fluid. Depletion of oxygen by bacteria may result in the formation of iron sulfides, such as pyrite. During burial, the sediments respond to increasing temperature (up to 200°C; 390°F), pressure (up to 2.5 kilobars; 250 megapascals), and fluid movement from compaction-driven waters or influx of water from the basin flanks. As a result, the sedimentary rock may contain authigenic minerals that record a sequence of events ranging from processes occurring near the sediment-water interface to those forming during deep burial. Unlike metamorphic rocks, the preexisting (detrital) mineral assemblage is at least partially retained, in part due to the sluggish reaction rates at diagenetic conditions. Early cementation processes often seal up the rock, preventing subsequent diagenetic reactions and preserving the original detrital mineral assemblage.

Examples. Authigenic minerals occur in all sedimentary rock and can vary from trace amounts to virtually the total rock (see **table**). The carbonate minerals calcite, dolomite, and siderite are some of the most common types. They form in a wide range of depositional environments and at varying burial depths. Calcite and dolomite form the principal minerals in limestones and dolostones, respectively, as well as cements in sandstones or shales. Carbonate cements result from recrystallization of detrital carbonates and from dissolution of other calcium, iron, and magnesium minerals with carbon dioxide from organic reactions. Much of the calcite in limestones initially consisted of aragonite or magnesium-rich calcite, whereas most dolomite has been formed by the chemical alteration of calcite. Recrystallization may change aragonite to calcite. Aragonite (orthorhombic) is a naturally unstable form of calcium carbonate. With the passage of geologic time, aragonite normally inverts to the more stable calcite (hexagonal). The substitution of magnesium for calcium is responsible for the conversion of calcite or aragonite to dolomite, and it has been shown that dedolomitization (replacement of magnesium by calcium) is also possible. *See* ARAGONITE; CALCITE; CARBONATE MINERALS; CEMENT; DOLOMITE.

Sandstones and shales are often cemented near the sediment-water interface by carbonate, which forms subrounded shapes called concretions (**Fig. 1**). Early-formed iron oxides of hematite or

Common authigenic minerals

Mineral	Formula
Albite	$NaAlSi_3O_8$
Anatase	TiO_2
Anhydrite	$CaSO_4$
Apatite*	$Ca_5(PO_4)_3(F,Cl,OH)$
Aragonite (orthorhombic)	$CaCO_3$
Barite	$BaSO_4$
Boehmite	$AlO(OH)$
Calcite (hexagonal)	$CaCO_3$
Celestite	$SrSO_4$
Clay minerals	
Chlorites*	$(Mg,Fe^{2+},Fe^{3+})_6$-$(Al,Si_3)O_{10}(OH)_8$
Illites*	$K(Al)_2(AlSi_3)O_{10}(OH)_2$
Kaolinite	$Al_2Si_2O_5(OH)_4$
Smectites*	$(Na,0.5Ca)_{0.5}(Al,Mg,Fe)_2$-$(Al,Si_3)O_{10}(OH)_2 \cdot nH_2O$
Dolomite	$CaMg(CO_3)_2$
Gibbsite	$Al(OH)_3$
Glauconite*	$K(Al,Mg,Fe^{2+},Fe^{3+})_2$-$(Al,Si_3)O_{10}(OH)_2$
Goethite	$Fe_2O_3 \cdot n(H_2O)$
Gypsum	$CaSO_4 \cdot 2(H_2O)$
Halite	$NaCl$
Hematite	Fe_2O_3
Leucoxene	TiO_2
Limonite	$FeO(OH) \cdot n(H_2O)$
Opal (amorphous)	$SiO_2 \cdot n(H_2O)$
Orthoclase	$KAlSi_3O_8$
Pyrite (isometric)	FeS_2
Pyrolusite	MnO_2
Quartz	SiO_2
Siderite	$FeCO_3$
Zeolites*	$X_y^{1+,2+}Al_xSi_{1-x}O_z \cdot nH_2O$
Clinoptilolite*	$(Na,0.5Ca,K)_{3.5}\ Al_{3.5}\ Si_{14.5}O_{36} \cdot nH_2O$
Analcime	$NaAlSi_2O_6 \cdot H_2O$
Laumontite	$CaAl_2Si_4O_{12} \cdot 4H_2O$

*Group of minerals characterized by considerable chemical variation.

goethite may coat grains when sediments are exposed to oxidizing waters or iron sulfides (such as pyrite) if the water is low in dissolved oxygen. Quartz cements in sandstones form overgrowths on detrital quartz grains at advanced burial conditions. Potential reactions producing silica include quartz dissolution

Fig. 1. Calcite concretion formed in Paleocene marine siltstone, Moeraki beach, New Zealand. Uncemented siltstone has been eroded away by wave action. Concretion is about 2 m (7 ft) in diameter, which is unusually large.

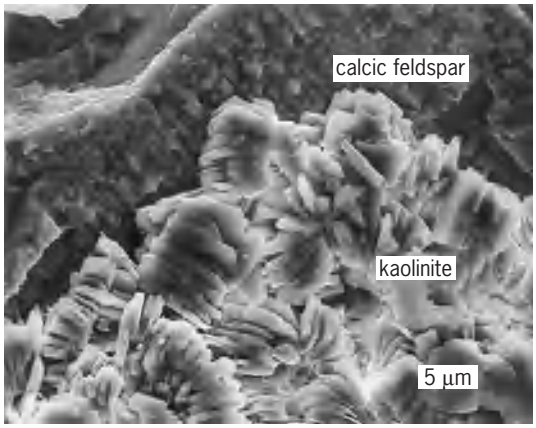

Fig. 2. Scanning electron micrograph of authigenic kaolinite. The clay mineral formed from dissolution of calcic feldspar in a Miocene sandstone, San Joaquin basin, California.

at quartz-clay grain contacts and breakdown of aluminosilicates such as clay minerals and feldspars. Overgrowth of authigenic feldspar on detrital feldspar is very common. At higher temperatures, calcic plagioclase is replaced by albite.

Authigenic clay minerals—chlorite, illite, and smectite—occur as grain coatings and cements. Clay minerals in sandstones commonly result from alteration of feldspars (**Fig. 2**). Commonly the dissolution within the feldspar will result in porosity, and the clay forms a pore-filling cement between the sand grains. In shales, mixed-layer smectite/illite clays are particularly common, and during progressive burial the increases in temperature and time cause the smectite to be converted to illite. Volcanic glass in sediments is often altered to zeolite and smectitic clay minerals (**Fig. 3**). With increasing geologic time the zeolite mineral assemblage becomes simpler due to the breakdown of relatively unstable early-formed zeolites. Pure fine-grain quartz rocks called chert (flint, jasper) crystallize from biogenic opaline silica during burial. *See* CHERT; CLAY; FELDSPAR; ZEOLITE.

Commonly, authigenic mineral assemblages are a result of a series of dehydration and desilicification reactions during progressive burial. Examples are reaction of gypsum to anhydrite, goethite to hematite, opaline silica to quartz, smectite to illite, and transformation of zeolites to feldspars. To balance these reactions, other processes and reactions consume silica and water. Recrystallization in authigenic mineral formation leads to an enlargement of crystal size through the process of dissolution followed by precipitation. Examples are large halite, gypsum, and anhydrite crystals in evaporite deposits and calcite crystals in some limestones. *See* GYPSUM; HALITE.

Identification. Since many of the authigenic minerals shown in the table may be deposited in sediments as detrital minerals, it is important that their authigenic origin be clearly established, especially when these minerals are utilized to interpret diagenetic chemical, temperature, and pressure conditions. When grown into pore space, authigenic minerals often exhibit crystalline shapes. In contrast,

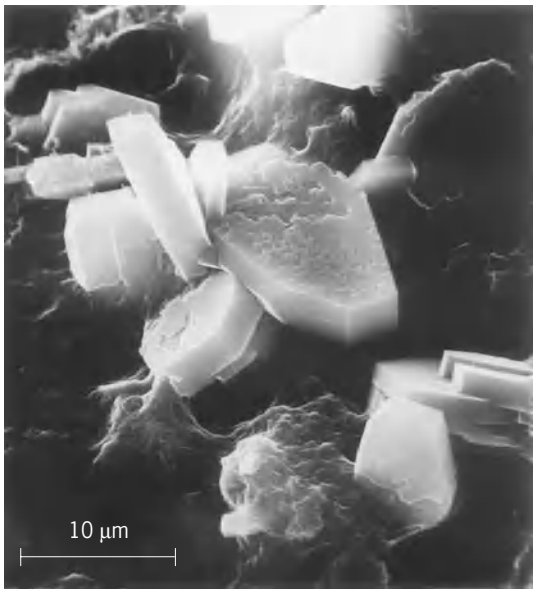

Fig. 3. Scanning electron micrograph of clinoptilolite zeolite (white crystals) on smectite-coated (dark background) sand grains. Sample is from Paleocene volcanic sandstone, Baja California, Mexico.

detrital mineral surfaces are irregular, rough, or rounded because they were leached in a weathering environment, were abraded during transportation, or have been leached further by pore fluids after deposition. Authigenic minerals also form obvious cements that coat grains and completely fill the original pore space in the sediment. In many cases, the authigenic minerals use preexisting crystal structures as templates for their growth. Examples are feldspars and zeolites forming on detrital feldspar, quartz overgrowths on quartz, and carbonates growing on shell tests. Although the authigenic origin of minerals can often be established from the study of thin sections, the scanning electron microscope and electron microprobe are particularly useful for their identification and chemical study. The identification of many fine-grain minerals such as clays and zeolites is possible only with x-ray diffraction and scanning electron microscope techniques (Fig. 2). The high magnification and resolution attainable with the scanning electron microscope make it possible to differentiate these minerals from like species of detrital origin by their crystal form. *See* MICROSCOPE; SCANNING ELECTRON MICROSCOPE; X-RAY DIFFRACTION.

James R. Boles

Autism

A neurodevelopmental condition that impairs the way that a person relates to and communicates with other people. Persons with autism also can have unusual behaviors, such as insistence on sameness, obsessions, or stereotypic (purposeless and repetitive) behaviors such as hand flapping, spinning, and toe walking. The condition varies greatly in the presenting symptoms, the timing of presentation, the range and severity of symptoms, and its association with other conditions. While recognition of autism is increasingly common, the cause and the reason for its increase remain unknown. With intensive early intervention, significant improvements in a large percentage of children with autism can be achieved.

Characteristics and diagnosis. Leo Kanner first described autism as a condition in 1943, but descriptions matching autism can be found throughout history. Based on his observation of 11 children, Kanner identified characteristic features of "early infantile autism" that are still considered essential in making an autism diagnosis. Three key features are: (1) social impairments that affect how a person relates to other people, apparent early in life; (2) peculiar use of language, or a failure to use language for interactive conversation; and (3) an obsessive desire for the maintenance of sameness associated with a "limitation in the variety of spontaneous activity." IQ ranged from 50 to 70 in classic Kanner autism, but currently autism is recognized over a broader range of IQ's.

Autism is diagnosed only through behavioral criteria. Typically, children with autism present with language delays, often first prompting a hearing assessment. Stereotypical behaviors such as arm flapping, rocking, and spinning usually develop later, after communication and social impairments are apparent. About a third of parents of autistic children report regression of previously achieved developmental milestones, and sometimes they attribute the loss of milestones to an identifiable stressor, such as a vaccination or the birth of a younger sibling. Two-thirds of the children whose parents reported regression had signs of social impairments or delays noted on review of videotape taken around 12 months of age.

The advantages of early diagnosis are many and generally outweigh concerns about potential harm from "labeling" a child. P. A. Filipek and colleagues authored a consensus statement on the assessment of children suspected of having autism. Developmental surveillance should occur at well-child visits. Concerns about speech, language, or social development warrant further evaluation. First-level evaluations include audiological assessments, testing for lead exposure if indicated by pica behaviors (eating of nonfood substances), and tests specific for autism [such as the Checklist for Autism in Toddlers (CHAT) screen or Autism Screening Questionnaire]. Children who screen positive for autism should be referred for a multidisciplinary assessment, which should include a detailed health and developmental history and a physical exam that looks for identifiable conditions that have autistic features. The history and physical findings should determine the necessity for additional laboratory and imaging studies. The assessment should include additional interviews and observations that specifically assess for autism. A complete evaluation also includes additional developmental screening and speech and language assessment, occupational therapy assessment, behavioral assessment, and a social work assessment.

Etiology. As recently as the late 1960s, autism was attributed to disturbed mother-child relationships.

The term "refrigerator mother" (implying destructive withholding of maternal nurturing) represents a dark period in the history of autism, when misplaced blame added to the anguish experienced by families with autistic children and delayed investigation into its true causes. The recognition that 25% of children with autism developed seizures began the shift in thinking to the more current view that autism is a neurodevelopmental disorder with a biological basis.

Genetics. The causes of autism remains elusive, but data support a strong genetic component. Families with one child with autism are more likely to have another child with autism than previously unaffected families, the risk being 3 to 5% in affected families compared to 0.1 to 0.2% in the general population. Autism occurs in males 3 to 4 times more frequently than females. Studies of families with multiply affected members have identified many chromosomes that are highly associated with autism; however, none are universally found in children with autism. Twin studies have found a concordance rate of 36–91% in identical twins (that is, when one twin has autism, the other twin has autism 36% to 91% of the time) compared with a less than 1% concordance rate in fraternal twins. Autism is more common in some conditions, such as Fragile X, neurofibromatosis, tuberous sclerosis, and Down syndrome, but no such associated condition can be identified in most children with autism. *See* CONGENITAL ANOMALIES; DOWN SYNDROME.

Prenatal and perinatal factors. Autism often manifests after the first year of life, but data from at least two sources suggest that autism often begins at or before the prenatal and perinatal periods. Analyses of neonatal blood spots taken from children later diagnosed with autism showed that 95% of a small sample of children with autism have elevated levels of four neuropeptides (short-chain proteins that can act as neurotransmitters) and neurotrophins (nerve growth factors). In this study, blood spot data for normal children and those with autism, mental retardation, or cerebral palsy were compared. Elevated neurotropin and neuropeptide levels were also found in children with mental retardation but not in normal children or those with cerebral palsy. Specifically, the damage that occurs during pregnancy resulting in cerebral palsy does not produce the changes in blood spot data that were found in autism or mental retardation. Another study found that 42% of children with autism have posteriorly rotated ears (compared with 10% of children without autism), representing changes that occur before birth in the first month of gestation. Other studies have shown neuroanatomic differences. Head sizes of autistic children are on average normal at birth, but become larger than average as the children age. Specific brain structures are differentially affected. Some are larger than normal in autistic children, such as the hippocampus and amygdala, and some are smaller, such as the cerebellar vermis. *See* MENTAL RETARDATION.

Autism spectrum disorders. Autism is a spectrum disorder that also includes Asperger disorder, Rett's disorder, childhood disintegrative disorder, and pervasive developmental disorder not otherwise specified (PDD-NOS). Other autism spectrum disorders (ASD) retain many of the core features of autism.

Asperger syndrome. Deficits in communication and social interactions occur in Asperger syndrome, but to a lesser degree than autism. The significant delay in the onset or early course of language that is seen in autism is not seen in Asperger syndrome. Overall, affected children have a better facility with language and more interest in social activities, but usually have autistic features, such as obsessive preoccupations.

Rett's disorder. Rett's disorder (also known as Rett syndrome), primarily affecting females, has normal early development, then regression some time after 5 months. Girls with Rett's disorder typically carry a single gene marker, the *MECP2* gene, and develop an unsteady gait, loss of or lack of language, constant hand wringing and loss of functional use of their hands, social deficits, and mental retardation. Head sizes of affected children are on average normal at birth, but slowing of head growth during early development results in small head size.

Childhood disintegrative disorder. Children with childhood disintegrative disorder have normal development for 2 to 10 years and then experience severe regression, resulting in complete loss of speech, social interaction, and self-help skills. Recovery has been quite limited in these children.

PDD-NOS. Children with PDD-NOS have some autistic features, such as communication and social impairments and repetitive behaviors, but do not meet the criteria for full syndrome autism and are not better described by one of the other ASD.

Epidemiology. Autism prevalence estimates have steadily increased over the last 20 years. Studies conducted prior to 1985 found that autism affected 4 to 5 per 10,000 children. Studies conducted between 1985 and 1994 show autism prevalence rates closer to 12 per 10,000 children. More recent data suggest autism is even more common, affecting 16 to 40 per 10,000 children. Straffordshire (United Kingdom) preschool children in the late 1990s had an autism prevalence rate of 16.2 per 10,000. In 1998 Brick Township, New Jersey, had autism rates estimated at 40 per 10,000 in 3- to 10-year-olds. Prevalence rates of autism and Asperger syndrome in Cambridgeshire (UK) were 1 in 175 (or ~57/10,000) for 5- to 11-year-olds. Children born in the Atlanta area in 1996 had an estimated autism prevalence rate of 34 per 10,000.

Controversy arises in the interpretation of the observed increase. A 1999 California Department of Developmental Services (DDS) report showed that from 1987 to 1998 autism cases in California's Regional Center System increased by 273%. This report was cited by some as evidence of an autism epidemic. Others challenged that the increase was due to changes in diagnostic criteria, increased awareness, increase in population, migration fluxes, earlier age at diagnosis, recording artifacts, and problems with cross-sectional data. However, a follow-up study that compared Regional Center clients who

were born between 1983 and 1985 with those born between 1993 and 1995 concluded that the increase in autism cases was not due to changes in diagnostic criteria, changes in misclassification of autism, or an influx of out-of-state autism cases. In addition, California DDS issued a 2003 report that documents a 100% increase in autism caseload since 1999.

Theories abound to explain the apparent increase in autism rates, given the lack of an identified cause for autism. These theories include: a bowel-brain connection that is altered by some potential insult, such as food allergies, antibiotic use, and immunizations; immunizations with MMR (measles, mumps, and rubella) or thimerosal-containing vaccines; heavy metal or other toxic exposure (such as to PCBs); defective hepatic detoxification capacity; autoimmune disorder or an imbalance between T-cell response; and yeast overgrowth in the bowel. The search for a cure or at least some improvement in their child's condition has led some families to unproven and potentially harmful interventions, which will likely continue until a cause and cure are found.

Epidemiology of other ASD is less well established. Some studies estimated Asperger syndrome as 3 to 4 times more common than autism. Based on a large population study conducted prior to the identification of the *MECP2* gene, Rett's disorder was estimated to occur in about 1 in 23,000. Childhood disintegrative disorder is felt to be rare: <5 per 10,000. PDD-NOS is believed to be more common that autism, but population estimates are hampered by the fuzziness in diagnostic criteria. One small-area estimate from the Brick Township investigation is that autism occurs in 1 in 250 and ASD occur in 1 in 150.

Treatment. Early identification and early intervention with intensive behavioral programs is the cornerstone of effective treatment for autism. In the past, dismal expectations led to minimal intervention resulting in discouraging outcomes. Expectations improved when Ivar Lovaas showed that intensive early intervention could make significant differences for about half of the children enrolled. Subsequently, a number of programs have been developed, including Auditory Integration Training, Daily Life Therapy (Higashi), EarlyBird Programme, LOVAAS/Applied Behavior Analysis (ABA), Treatment and Education of Autistic and Related Communication Handicapped Children (TEACCH), Options Son-Rise Programme, Picture Symbols (PECS) training, and Structure Positive (approaches and expectations) Empathy Low arousal Links (SPELL). Most of these programs are time-intensive and reward mastery of individual small steps that gradually lead to mastery of larger tasks. Adjuncts to intervention programs include diet modification, music therapy, facilitated communication, and computer applications. No one intervention is guaranteed to help every child with autism. Choice of programs is often dictated by program availability and funding. Even though early intensive interventions have been shown to be effective, many families have significant difficulties in obtaining services. There are no known medications that cure autism, but some are used to ameliorate some behaviors and improve functioning.

Long-term outcomes. A diagnosis of autism generally has lifelong implications. Outcomes in adults with autism reveal an overall poor prognosis for independent living and successful employment for the majority. Autistic adults with an IQ below 70 fare poorly, as might be expected by the cognitive limitations alone. In a study of autistic adults with IQ scores of at least 50, among those with more normal intelligence (IQ>70) only 17% achieved a high level of independence, having some friends and a job. Another 15% were working, but required some degree of support in daily living and had some friends or acquaintances. Compared with children currently diagnosed with autism, these adults were diagnosed much later (mean age at diagnosis was 7 years) and had fewer and less extensive therapeutic interventions. Studies of overall competence and language skills have demonstrated better outcomes for recently diagnosed individuals with autism. Continued advances in early diagnosis and development of effective treatments should further improve outcome for the increasing number of children diagnosed with autism. Robert Byrd

Bibliography. S. E. Bryson and I. M. Smith, Epidemiology of autism: Prevalence, associated characteristics, and implications for research and service delivery, *MRDD Res. Rev.*, 4:97–103, 1998; P. A. Filipek et al., The screening and diagnosis of autistic spectrum disorders, *J. Autism Develop. Disorders*, 29(6):439–484, 1999; S. Ozonoff, S. J. Rogers, and R. L. Hendren (eds.), *Autism Spectrum Disorders: A Research Review for Practitioners*, American Psychiatric Publishing, Arlington, VA, 2003; A. M. Wetherby and B. M. Prizant (eds.), *Autism Spectrum Disorders: A Transactional Developmental Perspective*, Paul H. Brookes Publishing, York, PA, 2001.

Autoimmunity

The occurrence in an organism of an immune response to one of its own tissues, that is, a response to a self constituent. Foreign substances to which an organism makes a protective immune response are called nonself. Efficient discrimination between self and nonself, the basis of normal immune function, depends upon a function known as immune tolerance, which means an inertness to substances that could be capable of provoking an immune response. Failure of immune tolerance to self constituents results in an autoimmune response which is often, although not invariably, associated with autoimmune disease. Autoimmune disease occurs when the autoimmune response to self constituents has damaging effects of a structural or functional character. *See* ACQUIRED IMMUNOLOGICAL TOLERANCE.

History. In 1901, P. Ehrlich proposed that, as a condition of life, organisms should not be reactive against their own tissues. This view influenced scientists until the 1950s, when it was found that autoimmunity could be experimentally created in animals

by vigorous immunizing procedures. Soon thereafter, laboratory tests in humans showed that immune responses to self tissues could occur spontaneously and could be associated with disease of the blood or thyroid gland. Procedures for laboratory diagnosis of autoimmunity developed rapidly, based on tests for detection of antibodies in the blood plasma. Antibodies are globulin molecules that react with a stimulating substance, the antigen. Autoantibodies are equivalent globulin molecules in blood that react with self antigens; their presence points to an autoimmune reaction in the body. By the 1960s, it was recognized that autoimmunity was a direct or contributing cause of a wide range of human diseases. *See* ANTIBODY; ANTIGEN; IMMUNOGLOBULIN.

Normal immunity mechanisms. When any foreign antigenic substance, such as a microorganism, protein, carbohydrate, nucleic acid, or a complex of these, is introduced into the body, it is recognized by a genetically determined process as being nonself. The foreign antigen is first degraded by phagocytic cells, which then present processed antigen to cells of the lymphoid system. The specialized cells with this function are part of the lymphoid reticuloendothelial system and include dendritic cells and macrophages. Lymphoid cells, after a recognition process based on surface receptor structures, undergo proliferation and generate one population of lymphocytes which reacts specifically with the antigen, and a second group which produces specifically reactive antibody molecules. Long-lived descendants of these populations of lymphoid cells provide for "memory," which allows for accelerated responses in the event of subsequent exposure to the same antigen. Immune responses are exemplified by the neutralization of toxic protein molecules, the killing of infecting microorganisms, the rejection of foreign tissue grafts, and memory by the protective effects of vaccination.

Lymphocytes that participate in immune responses belong to two major groups. One group, which matures in the thymus gland, comprises thymic or T lymphocytes, of which there are several subsets. These subsets have different functions and carry unique surface molecules which can be identified with monoclonal antibodies: (1) The subset of helper T lymphocytes which is marked by the CD4 molecule responds to antigens by releasing stimulatory cytokines (intercellular hormones) that can amplify the number and activity of lymphocytes participating in the immune response; (2) the subset of cytolytic T lymphocytes which is marked by the CD8 molecule can directly recognize and kill cellular targets, usually virus-infected cells; and (3) the subset of suppressor T lymphocytes, which also carries the CD8 molecule, releases molecules that reduce the intensity of immune responses, or switch these off altogether. The other major group of lymphocytes, which mature in the bone marrow, are B lymphocytes. After stimulation with antigen molecules, and under the influence of factors released by helper T cells, B lymphocytes proliferate and they later secrete the antibody molecules which, when circulating in the blood, provide for humoral immune responses. *See* CELLULAR IMMUNOLOGY; IMMUNITY.

Major histocompatibility complex (MHC). This gene complex on human chromosome 6 codes for cell-surface molecules which confer biological uniqueness on cells of an individual. The complex is called HLA in humans: it is important for the matching of transplanted organs, and also has fundamental importance for immune responsiveness. There are two classes of MHC (HLA) molecules. Class I molecules are expressed on all cells and are recognized by $CD8^+$ cytolytic T lymphocytes; this agrees with the fact that any cell may be infected with a virus and thus should be susceptible to killing. Class II molecules are expressed normally only on those cells, B cells or antigen-presenting cells, with the special function of interacting with and stimulating $CD4^+$ helper T lymphocytes. The expression of MHC (HLA) molecules on lymphocytes is amplified by cytokines; for example, the cytokine interferon-gamma can induce MHC class II molecules on cells which normally are incapable of presenting antigen to helper T lymphocytes. Such aberrant expression of MHC molecules is relevant to autoimmunity. *See* HISTOCOMPATIBILITY.

Recognition by the immune system. Concepts on the origin of specific recognition structures on the surface of cells of the immune system developed after the mid-1970s. During embryonic development there is generated information in the genome of cells of the immune system which provides for reactivity with all antigens ever likely to be encountered in later life. Thus there is genetic programming for the structure of the cell-surface receptor molecules that enable antigens to be recognized, and for the specificity of the antibody molecules that combine with antigens in the blood. Antigen selects lymphocytes with the appropriate surface receptor and stimulates these to proliferate as an identical clone of cells: this is the clonal selection concept of immune responsiveness. In the essentially random process of generation of antibody diversity, reactivity for self antigens will be created, and this must be dealt with, as described below.

Regulation of immune responses. The immune system remains in a state of balance conditioned by positive (on) signals and negative (off) signals. Positive signals are provided by antigen in low dose and the amplifying factors released by activated helper T lymphocytes, while negative signals are provided by antigen present in excess which causes an overload paralysis, and by suppressor T lymphocytes which are generated preferentially when self antigens are presented. There is still a lack of full understanding of the mechanism by which unwanted immune responses to self antigens are suppressed so as to provide for natural tolerance to self. The major processes are (1) permanent deletion, or functional inactivation in early life of cells capable of responding to self antigens; and (2) regulatory controls, which inhibit the activity of self-reactive lymphocytes that escape the deletion process. The relative contribution of these two mechanisms for specific self antigens appears to differ, and both probably operate to

control autoimmunity. Deletion of self-reactive cells would be the more economical means of achieving self-tolerance, but it is the less flexible. It is likely that these processes apply more particularly to T lymphocytes; since B lymphocytes are dependent on helper T lymphocytes for their activation and maturation, a lack of T-cell help is one way in which B-cell function can be reduced. There are low background levels of immunologic reactivity to many self antigens in healthy subjects, indicating that suppressor activity over immune responses to autoantigens must be continuously operative.

Causes. Failure of immune regulation is responsible for autoimmune disease. Inheritance may account for 25–50% of the risk for autoimmune diseases, since it is known that autoimmune disease, or at least the tendency to produce autoantibodies, runs in families. There are many genetic determinants, and they are poorly understood. One set is in some way associated with MHC (HLA). Since products of HLA genes normally function to direct T lymphocytes to cells with which they should interact, it is not surprising that autoimmune diseases are associated with the presence of particular HLA types; examples include B8 (thyrotoxicosis), DR4 (rheumatoid arthritis, type 1 diabetes mellitus), and DR2 (multiple sclerosis). The reason may be that the autoantigen readily associates with the MHC (HLA) molecule on cells which present antigen to helper T lymphocytes. The MHC influences the occurrence of autoimmunity in other ways. Thus release of cytokines by T lymphocytes may induce aberrant expression of class II molecules on tissue cells which then can present their own antigens, and these become inducers of and targets for an autoimmune response. In addition to the MHC, there are other inherited determinants of autoimmunity, including genes specifying immunoglobulin structure, and genes specifying weakness in the downregulation of immune reactions, thus allowing autoimmune responses to become amplified. *See* IMMUNOGENETICS.

There may also be somatic genetic causes of autoimmunity, meaning random mutations in later life among genes that code for immunoglobulins that function as recognition structures on the surface of B lymphocytes; such a mutation may generate a cell with a receptor structure with exquisite specificity for a self antigen which is resistant to regulation. Environmental causes could include infection with microorganisms that carry antigenic structures closely resembling those of self; these, when presented to the immune system, could provoke an uncontrolled response to the related self structures of the body.

Any autoimmune response must become self-sustaining, which implies coexisting failure of normal regulatory processes, either by reason of genetic predisposition or by an acquired disruption of immune function. Clearly the induction of autoimmunity is multifactorial, and occurs when there is a synchronous breakdown of several physiologic fail-safe functions that normally operate to maintain homeostasis in the immune system.

Effects of autoimmune reactions. Once self-sustaining, the autoimmune reaction can cause damage or dysfunction in one of several ways. First, autoantibody molecules circulate in the blood and, by attaching to self antigens on cell surfaces, either damage cells or interfere with important cell-surface receptor molecules. Second, antibodies can unite with their autoantigen, which results in the binding of a serum factor, complement, to form immune complexes that are capable of provoking inflammatory responses; for example, immune complexes deposited in the filtering tissue of the kidney cause progressive nephritis. Third, there may be generated T lymphocytes with the capacity for cellular destruction, and these may cause the progressive inflammatory damage that characterizes autoimmune reactions in solid organs.

Autoimmune diseases. Many human diseases can be attributed to autoimmune reactions. Circulating autoantibodies are responsible for diseases in which there is intravascular destruction of elements of the blood, for example, the red blood cells in hemolytic anemia. T lymphocytes may be responsible for some types of thyroid goiter, such as Hashimoto's disease; a stomach mucosal degeneration that results in nonabsorption of vitamin B_{12} and thus the blood disease pernicious anemia; the insulin-dependent or juvenile type of diabetes mellitus; and one type of chronic hepatitis. Immune complexes cause glomerulonephritis and most of the features of systemic lupus erythematosus, in which autoantibodies are formed to various constituents of cell nuclei. In Sjögren's disease, in which salivary and lacrimal glands are destroyed, damage by T lymphocytes within the glands may be accompanied by damage by immune complexes throughout the body. Some autoimmune diseases are caused by antibodies to cell receptors, which either block neuromuscular transmission, as in myasthenia gravis, or stimulate thyroid cells to overactivity, as in Graves' disease. Some important human diseases may be autoimmune disorders, although demonstration of an autoimmune basis is not yet adequate: these include rheumatoid arthritis, multiple sclerosis, and ulcerative colitis. In many autoimmune diseases the autoantigenic trigger has not been identified; perhaps autoimmunity supervenes after an unidentified primary cause has set the process going. *See* ANEMIA; ARTHRITIS; DIABETES; HEPATITIS; MULTIPLE SCLEROSIS; MYASTHENIA GRAVIS.

Animals. Autoimmune diseases are not confined to humans. The numerous examples in the animal world include hemolytic anemia and lupus, which occur as genetically dependent diseases in certain inbred strains of mice, and diabetes in inbred strains of rats; outbred domestic animals, notably dogs, develop autoimmune diseases, such as lupus, equivalent to those seen in humans. Study of these models in animals has been less informative than might have been expected in terms of clarifying causes of autoimmunity and autoimmune disease.

Treatment of autoimmune disease. Autoimmune diseases are alleviated by treatment, though these diseases are seldom actually curable. At the simplest

level, replacement of the specific secretions of tissues or organs damaged by autoimmune reactions may help: this applies to treatment of diseases resulting from atrophy of the thyroid gland (myxedema), stomach mucosal degeneration (pernicious anemia), and destruction of pancreatic islet tissue (diabetes mellitus). For multisystem autoimmune disease, such as lupus, there are drugs, particularly cortisone derivatives, that modify the harmful effects of humoral or cellular autoimmune attack on tissues and so allow the body to reestablish immunologic homeostasis. Also used are cytotoxic immunosuppressive drugs, which are given specifically to inhibit the activity of immunologically active cells responsible for autoantibody formation or for cytolytic damage to tissues. In the future, biological modifiers may allow facilitation of normal regulatory control by modifying various stages between the triggering of an autoimmune response and the destructive effect on the target organ. Finally, as techniques for molecular characterization of autoantigens are developed, it may become possible to inject these in such a way as to reestablish specific self-tolerance to the autoantigen responsible for an autoimmune disease. *See* IMMUNOLOGY.

Ian R. Mackay

Bibliography. C. A. Bona et al. (eds.), *The Molecular Pathology Autoimmune Disease*, 1993; I. R. Mackay and F. M. Burnet (eds.), *Autoimmune Diseases: Pathogenesis, Chemistry and Therapy*, 1963; N. R. Rose and I. R. Mackay (eds.), *The Autoimmune Diseases* 3d ed., 1998; N. Talal (ed.), *Molecular Autoimmunity*, 1991.

Automata theory

A mathematical model of computing. The versatility of modern computers and their applications to the study of complex and dynamic systems, such as financial markets, cellular growth, and communication networks, raise questions about the ultimate power of computers. For example, what are the limitations, if any, on the tasks that we can ask computers to perform? Can we use computers to carry out any computational task we want? To answer such questions, we generally make use of abstract mathematical models. One of the most important mathematical abstract models that has been widely used to simulate objects and processes such as computer and digital circuits is automata theory. *See* ABSTRACT DATA TYPE; COMPUTER.

Alan M. Turing, a British mathematician, first examined the notion of an automaton, or abstract machine, in 1936 while studying the limits of human ability to formalize methods of solving problems. At that time, Turing was trying to solve a decision problem to determine whether there is a general method that can be applied to any assertion to determine whether the assertion is true. In his approach to this problem, Turing invented the most general model of a computing machine, the Turing machine. Turing machines are universal in the sense that every known algorithm can be executed by a Turing machine. For purposes of explanation on how an automata works, we will use less general models such as the finite-state machine and pushdown automata. We will consider these two types of automata, along with the Turing machine.

Finite-state machines. A finite-state machine is the simplest of all automata theory models and serves as the control device for all other automata. The term "finite-state machine," depending upon the application, is defined in many different ways in the automata literature. The commonality of these definitions is that finite-state machines model computing devices that have a fixed and finite amount of memory and read sequences of symbols made up from a finite set of input symbols. The main differences in the definitions concern what the machines do in the way of output. Finite-state machines find applications in the computer field in areas such as spell checkers, sequential circuits, and compiler design.

Finite-state machines can be described by (1) a finite and nonempty set of states, with one of these states is designated as the starting state, (2) a finite set of input symbols (or input alphabet), and (3) a transition or next-state function that assigns to a current state, on a given input symbol, one or more states called the set of next states. If the transition function assigns only single next states, the finite-state machine is called deterministic; otherwise, it is called nondeterministic. Other types of finite-state machines may have a finite set of output symbols (or output alphabet) and an output function. The output function may associate an output value (from the output alphabet) with every transition (Mealy machines) or produce an output value based on the current state (Moore machines). Finite-state machines also have a subset of their states designated as final or accepting states.

As indicated in **Fig. 1**, a finite-state machine can be conceptualized as a control unit that has an input tape head that reads symbols from a linear tape in a sequential manner, one character at a time, from left to right. The control unit is made up of states. Initially the machine is in starting state and reads the leftmost symbol of an input string—a sequence of symbols or characters of a given alphabet. Each individual symbol of the input string occupies a single cell on the tape.

In mathematical terms, a finite-state machine, M, with no output can be defined as a five-tuple $M = (Q, \Sigma, \delta, q_0, F)$, where Q is a finite set of control states, Σ the set of input symbols, δ the transition function ($\delta: Q \times \Sigma \rightarrow Q$), q_0 the initial state, and F ($F \subseteq Q$) the set of final states. The term "five-tuple" is used to indicate that there are five components that

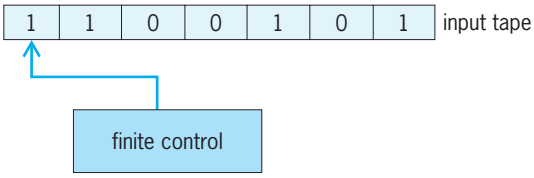

Fig. 1. Abstract representation of a finite-state machine.

define this type of mathematical structure. As long as these elements are clearly identified, the order in which they appear in the five-tuple is immaterial. The expression $\delta: Q \times \Sigma \rightarrow Q$ represents a function that maps pairs of elements of Q and Σ (in that order) into elements of Q. The interpretation of $\delta(q, a) = p$, for q and p in Q and a in Σ, respectively, is that M, in state q and scanning the input symbol a, moves its input head one cell to the right and goes to state p.

Finite-state machines are generally represented by state diagrams or transition tables as shown in **Fig. 2**. State diagrams are directed graphs with labeled edges, where circles represent states and arrows represent transitions. Input symbols associated with transitions are shown along the arrows. An incoming arrow identifies the starting state, and the accepting states are shown as double circles. The individual components of the finite-state machine of Fig. 2 are $Q = \{S_0, S_1, S_2\}$, $\Sigma = \{0, 1\}$, $q_0 = \{S_0\}$, $F = \{S_2\}$. The transition function, δ, determines the next state based on the current state and the current input symbol. For example, if the current state of the machine is S_0 and the current input symbol is 0, then S_1 is the next state. Likewise, if the current state is S_1 and the current input symbol is 1, then S_2 is the next state.

Transition tables convey the same information as state diagrams. In transition tables, there is a column for each input symbol and a row for every state. Table entries represent transitions. We will adopt the convention that the first row of the table corresponds to the starting state and that the accepting states are underlined.

The finite-state machine of Fig. 2 is an example of an acceptor. A finite-state machine accepts an input string when, after reading the last character of the string, the machine finds itself in one of its accepting states; otherwise, the machine rejects the input string. The set of all input strings accepted by a finite-state machine is called its language. The finite-state machine of Fig. 2 accepts an input string such as 1100101, and its language consists of all strings that end in 01.

Pushdown automata. Finite-state machines can be used only in computational tasks that require a fixed and finite amount of memory. Processes such as the recognition of valid algebraic expressions by a compiler cannot be carried out under these memory restrictions. Therefore, to perform tasks that require no memory limitations, it is necessary to augment the finite-state machine model and its memory capabilities. One model that accomplishes this is the pushdown automaton. This type of automata extends the finite-state machine with a pushdown list or stack (**Fig. 3**). A stack is a list in which insertions and deletions are possible. Both operations take place at one end of the list called its top. Symbols on a pushdown list can be inserted or deleted in the same way we can add to or take plates from a stack of plates in a cafeteria. We can only put a plate at the top and can only remove the topmost plate.

Formally a pushdown automaton, P, is a seven-

State	Input symbol	
	0	1
S_0	S_1	S_0
S_1	S_1	S_2
$\underline{S_2}$	S_1	S_0

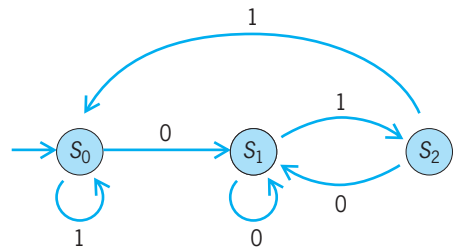

Fig. 2. Transition table and state diagram of a finite-state machine.

tuple $P = (Q, \Sigma, \Gamma, \delta, q_0, Z_0, F)$, where Q is a finite set of control states, Σ is a finite input alphabet, Γ is a finite alphabet of pushdown list symbols, and δ defines the control operations. The initial state of the machine is q_0. An element of Γ, Z_0, is the start symbol of the stack. And F ($F \subseteq Q$) is the set of final states.

A pushdown automaton can be in one of a finite set of states. Symbols from Γ are inserted on top of the stack by a "push" operation. The topmost symbol of the stack is deleted by a "pop" operation. As with any finite-state machine, the processing of an input string is carried out in a series of steps. Each step is determined by a control and its current configuration. The latter term refers to a 3-tuple formed by the current state, the current input symbol, and the symbol at the top of the stack. Each step may change the configuration of the machine by changing its state and pushing or popping the stack. However, a pushdown automaton, unlike a finite-state machine, may take several steps to process a single input symbol. At each step, the machine determines whether it is finished processing the current input symbol, in which case it reads the next input symbol, or whether the next step should keep the current input symbol. To process an input sequence, the pushdown automaton should be in its starting state, q_0, with Z_0 on top of the stack; the current symbol should be the first character of the string. The pushdown automaton then carries out operations called for by its control. If a transition is called for, a new current input, top stack symbol, and state are obtained and the control is called to carry out another operation. If an exit is called for, processing stops. The control operations cannot ask for a next symbol past the last character of the input string. Likewise, the stack cannot

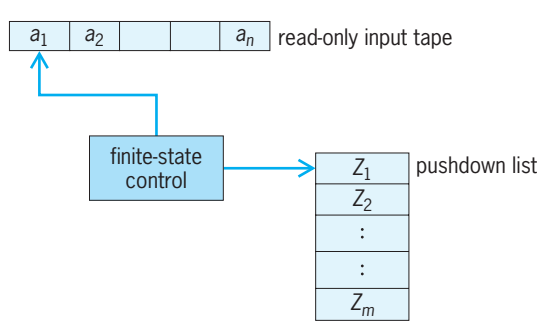

Fig. 3. Conceptual representation of a pushdown automaton.

be popped if Z_0 is at the top of the stack (an empty stack).

Turing machines. The Turing machine is the most powerful of all computational models. Since the publication of Turing's work, many other mathematical systems, such as the production systems of E. Post, the logical systems of A. Church, and the recursive functions of S. Kleene, have been yielded equivalent concepts. The results of these independent works support the argument that the computing power of the Turing machine represents a fundamental limit on the capability of realizable computing devices (Church's thesis). Put simply, Church's thesis states that a Turing machine can describe any computation that can be performed on a computer. In other words, a Turing machine can carry out any algorithm that can be executed by a computer. Thus, if there ever existed a procedure that could be mechanically carried out in a finite sequence of instructions but could not be described by a Turing machine, that procedure could not be programmed for any existing computer. *See* ALGORITHM.

Specifications for the Turing machine have been presented in different ways in the literature. We will consider a basic model that resembles the conceptualization of Fig. 1. It has a finite control with a tape head that, at any time, can read from a cell or write to a cell a single symbol from a finite set of tape symbols. The tape has a leftmost cell but it is infinite to the right. Initially, the n ($n \geq 1$) leftmost cells hold an input string whose symbols belong to a subset of the tape symbols called the input symbols. All cells to the right of the last character of the input string hold a noninput symbol called the blank. The Turing machine can make a move depending on the state of the control and the current symbol. In a move, the Turing machine may change its state, move the tape head left or right one cell, or write a nonblank symbol on the current cell. Any symbol written into a cell replaces its previous content. Pairs of current state and current symbol that do not have a specified move halt the machine. Whenever this happens, the remaining symbols on the tape form the output of the machine. A sequence of moves followed by a halt is called a computation. A Turing machine accepts some input string if it halts in it. The set of all accepted input strings is called the language accepted by the Turing machine.

Mathematically, a Turing machine, T, is a six-tuple, $T = (Q, \Sigma, \Gamma, \delta, q_0, F)$, where Q is a finite set of control states. Γ is a finite set of allowable tape symbols; one particular symbol, denoted by B, is the blank. Σ ($\Sigma \subseteq \Gamma$) is the set of input symbols; Σ does not include the blank. δ defines the control operations and is called the next-move function (δ: $Q \times \Sigma \rightarrow Q \times \Gamma \times \{\text{Left, Right}\}$). The function, δ, may be undefined for some arguments. The initial state of the finite control is q_0. And F ($F \subseteq Q$) is the set of final states.

Turing machines can be described by tables similar to that of the finite-state machines. **Figure 4** shows a table with input symbols 0, 1, and B and states q_1, q_2, and q_3. Entries in the tables represent moves of the machine. A blank entry indicates that the machine halts. The tape head, from its current position, can move to the right cell (R), to the left cell (L), or stay on the same cell (S). The entry, $1q_1$R, for state q_1 and input symbol 1 indicates that the machine writes 0 on the current cell, the next state is q_2, and the tape head moves to the right cell. This Turing machine, for example, for the input string 1010 will produce the output 0101.

	0	1	B
q_1	1 q_3 S	0 q_2 R	
q_2	1 q_1 R	B q_2 L	
q_3			

Fig. 4. Table representation of a Turing machine.

Ramon A. Mata-Toledo

Bibliography. J. E. Hopcroft, R. Motwani, and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, 2d ed., Pearson Addison-Wesley, 2000; P. Linz, *An Introduction to Formal Languages and Automata*, 3d ed., Jones & Bartlett, 2000.

Automated decision making

The use of computers to automate decision tasks. A decision task is any task requiring the generation or selection of options.

Related to automated decision making (ADM) are techniques for automated control. The term "control" is usually applied to continuously operating systems that are constantly monitored and adjusted. The term "decision making" usually refers to comparatively high-level tasks with discrete decision points. For example, continuous monitoring and adjustment of operating factory equipment may be considered a control task, while the scheduling of factory operations may be considered a decision task. *See* CONTROL SYSTEMS; PROCESS CONTROL.

Another related area is decision support systems. Decision support systems are intended to support human decision making, while automated decision making concentrates on completely computerautomated decision making. Nevertheless, there is a close connection between these two areas since many decision support systems are designed as advisory systems, where the decision support system recommends decisions. To generate a recommendation, the decision support system invokes an ADM model. Consequently, both automated decision making and decision support systems rely on the same underlying ADM models; and it is the maturity and reliability of the ADM models which determine whether the system is used to automate or just to support decision making. *See* DECISION SUPPORT SYSTEM.

Technological foundations. Techniques for automated decision making are drawn from several disciplines. Alternative approaches to automated decision making can be characterized in terms of

a trade-off between power and generality. An ADM technique is general to the extent that it can be applied to diverse problems. It is powerful to the extent that it quickly generates good answers.

Optimization. Mathematical optimization techniques can be used to build ADM models that are powerful but of limited generality. For example, linear optimization techniques can be used to assign resources optimally, but only if the resource allocation problem can be described by linear functions and inequalities. *See* OPERATIONS RESEARCH; OPTIMIZATION.

Artificial intelligence. Artificial intelligence techniques are more general but less powerful than optimization techniques. Many artificial intelligence techniques are designed to encode the diverse knowledge that a person might use to make decisions. However, the cost of this generality is that most artificial intelligence systems usually search for a satisficing solution, which is any solution that satisfies a set of constraints describing an acceptable solution. Artificial intelligence techniques have been applied to a broad range of decision-making problems, from medical diagnosis to automated game playing. *See* ARTIFICIAL INTELLIGENCE.

Decision theory. This is the science of decision making. Decision theory is partitioned into two fields. Normative decision theory searches for abstract characterizations of ideal decision-making behavior. The emphasis is on finding a theoretically justified approach to how judgments and decision should be made. Descriptive decision theory is a branch of psychology that performs research on human judgment and decision making. The emphasis is on characterizing how people actually make decisions.

Normative decision theory provides diverse hypotheses and option selection techniques. These techniques are particularly useful when the decision problem can be framed as a few discrete options. The use of techniques from game theory helps select options when the decision maker needs to consider the possible actions of other decision makers (for example, a business competitor). *See* GAME THEORY.

Descriptive decision theory is useful for characterizing the types of decision problems where people tend to make errors and may therefore benefit from good decision support or automated decision making. For example, one line of research in descriptive decision theory focuses on the use of linear models to predict and replace human judgment. This tradition is based on three findings. First, in probabilistic domains simple linear models (for example, adding up pros and cons) are often good predictors of human judgment. Second, this result holds true even for domains where people employ a complex expert knowledge to make judgments (for example, clinical judgment). Third, linear models of expert judgment usually perform better than the experts from which they were derived.

Thus, in the decision-theory approach to automated decision making, an ADM system considers the same variables that a human expert considers in making a decision, but does not emulate human expert decision-making processes. Rather, the variables are integrated into an overall decision procedure using techniques from normative decision theory. *See* DECISION THEORY.

Integration of techniques. During the 1990s there was a trend toward integration of ADM techniques. Decision-theory approaches now strongly influence artificial intelligence research in automated decision making, and there has also been an active interest in integrating optimization and artificial intelligence problem-solving approaches.

Application to simulation. Commercial computer games and many military training simulations involve an active interaction of "players" with other computer-controlled agents in a simulated world. In both cases, the value of the simulation (whether for entertainment or effective training) depends on the realism and intelligence of the simulated agents. Simulated agents must make decisions in their simulated worlds and are therefore executing automated decision making. *See* SIMULATION; VIDEO GAMES.

Autonomous systems. There is an increasing demand for autonomous systems that can perform various complex tasks without human control or intervention. In 1999, for instance, the National Aeronautics and Space Administration (NASA) released control of its *Deep Space One* spacecraft to an autonomous planning and decision-making system. As the demand for autonomy increases, so will the reliance on ADM capabilities. *See* SPACE PROBE.

Paul E. Lehner

Bibliography. D. Brown and C. White (eds.), *Operations Research and Artificial Intelligence: The Integration of Problem-Solving Strategies*, Kluwer Academic, 1990; E. Castillo, J. M. Gutiérrez, and A. S. Hadi, *Expert Systems and Probabilistic Network Models*, Springer-Verlag, New York, 1996; R. Dawes, The robust beauty of improper linear models in decision making, *Amer. Psychol.*, 34(7):571–582, 1979; P. Lehner and L. Adelman, Behavioural decision theory and its implications for knowledge engineering, *Knowl. Eng. Rev.*, 5(1):5–14, 1990; D. Olson and J. Courtney, Jr., *Decision Support Models and Expert Systems*, Dame, Houston, 1997.

Automatic frequency control (AFC)

The automatic control of the intermediate frequency in a radio, television, or radar receiver, to correct for variations of the frequency of the transmitted carrier or the local oscillator. In high-fidelity broadcast receivers AFC keeps distortion due to detuning to a low figure. In the reception of long-haul telegraph signals, AFC reduces the error rate due to signal pulse distortion or interference from lower intensity signals in the same frequency band.

Single-sideband receivers receive signals which are transmitted with a carrier level that is reduced to as small a proportion as 5%, or less, of the sideband (intelligence) amplitude. Proper demodulation requires the generation locally of a carrier-frequency signal synchronized to the transmitted carrier by AFC. Since propagation at frequencies of 3–30 MHz

is dependent upon reflections from the ionosphere, motion in this medium will speed up and retard the arrival of the wave, causing Doppler-effect frequency changes. Transmission to and from speeding aircraft will also suffer Doppler frequency drift. To reduce these effects, the carrier is transmitted for synchronization so that the frequency difference between the carrier and side frequency is maintained on reception.

AFC techniques are varied. One uses a discriminator to furnish a voltage whose magnitude and polarity are determined by the frequency change. This voltage is used to adjust the frequency of the local oscillator of the receiver, thereby keeping the intermediate frequency constant. A second technique uses two-polarity pulse accumulation which furnishes a dc potential proportional to frequency error. Another technique for spread-spectrum transmissions uses the recovered chipping frequency for AFC.

To select only the carrier for AFC, very narrow bandwidths are utilized. As an example, a bandwidth of 30 Hz at 70% of maximum response is quite common. This is accomplished in telemetry transmissions by using phase modulation with low index so a carrier component remains for receiver AFC. The response of the control circuits is usually designed to be slow and to be inactive below a determinable level of carrier input. These techniques reduce noise and interference as well as frequency-control capture by undesired adjacent carriers. *See* RADAR; RADIO RECEIVER; TELEVISION RECEIVER. Walter Lyons

Automatic gain control (AGC)

The automatic maintenance of a nearly constant output level of an amplifying circuit by adjusting the amplification in inverse proportion to the input field strength, also called automatic volume control (AVC). Almost all radio receivers in use employ AGC. In broadcast receivers AGC makes it possible to receive incoming signals of widely varying strength, yet have the sound remain at nearly the same volume. In communications receivers a type of AGC circuit called a squelch circuit is used to prevent noise during periods of no transmission, such as in the reception of on-off keying, frequency-shift keying, and phone. AGC is also useful in accelerating the switching action between receivers in diversity connection. *See* RADIO RECEIVER.

AGC action depends on the characteristic, possessed by most electronic tubes and transistors, of adjustment of gain by the variation of the applied bias voltage. If the dc voltage applied to the control grid of a vacuum tube is made more negative, the amplification of that stage will be reduced.

In most broadcast receivers the AGC voltage is taken from the detector. This dc voltage, proportional to the average level of the carrier, adjusts the gain of the radio-frequency (rf) and intermediate-frequency (i-f) amplifiers and the converter (**Fig. 1**). AGC tends to keep the input signal to the audio-frequency (af) amplifier constant despite variations

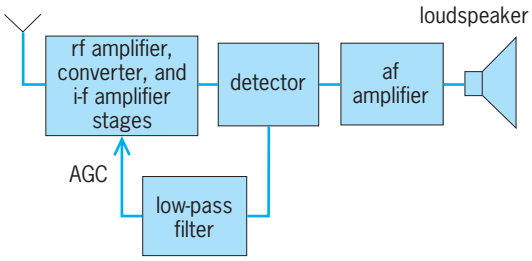

Fig. 1. Block diagram of broadcast receiver using AGC.

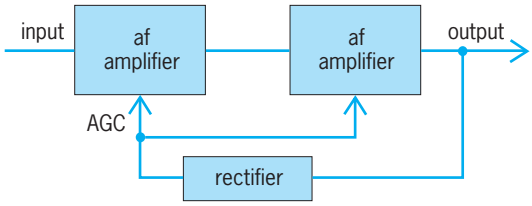

Fig. 2. Block diagram of AGC of an audio amplifier.

in rf signal strength. There are several modifications of this basic circuit.

Perfect AGC action would provide a constant output for all values of input signal strength. A slightly rising output characteristic with increased signal strength is generally desirable to facilitate proper tuning. The figure of merit applied to AGC action is given as the change in input required for a given output change. An example of a good figure of merit may be seen when an 80-dB change in input carrier signal results in an output change of no more than 3 dB. This applies to the unmodulated carrier strength only, since the modulation of the carrier must always vary as the modulation of the transmitter.

AGC circuits are also used in dictation recording equipment, public address systems, and similar equipment where a constant output level is desirable. **Figure 2** shows a typical block diagram for such equipment. *See* AMPLIFIER. Walter Lyons

Automatic landing system

The means for guiding and controlling aircraft from an initial approach altitude to a point where safe contact is made with the landing surface. Such systems differ from low-approach systems in three major respects: (1) They furnish not only guidance but control of the aircraft as well. (2) They furnish information on the aircraft's position with respect to the terrain, and the rate at which the landing surface is being approached. (3) They do not require the pilot to assume manual control near the ground.

Two automatic landing systems have been developed. One is a radar-beam type that was developed to assist pilots landing on an aircraft carrier. The landing guidance is derived on the ship and is communicated to the aircraft. This ground-derived landing system detects the position and rate of change in position of the landing aircraft by means of a radar beam emitted from a ground derived-control (aircraft carrier) complex. The other is a variant of an instrument landing

system (ILS) in which the landing solution is derived on the aircraft. The aircraft derives position and rate of change in position by using its onboard instrumentation in addition to radio signals transmitted by ILS-type equipment on the ground. The aircraft instrumentation includes accelerometers (which may be part of an inertial navigation system) and a radio altimeter. Essential to both systems is an autopilot in the aircraft, commanded by a computer on the ground in the radar-beam system and by a computer in the aircraft for the ILS-based system. The radar-beam and ILS-based systems are described below. *See* INSTRUMENT LANDING SYSTEM (ILS).

Ground-derived landing system. The ground-derived system was designed to place the burden of space, weight, and complexity on the aircraft carrier, or at the air station in the case of shore application. The carrier landing system is composed of a K_a-band fire-control-type radar, a stable element (mechanized by inertial instrumentation), a high-speed general-purpose digital computer, a display console, flight-path recorder, data-link coder/transmitter, and data-link monitor on the carrier. In the aircraft a data-link receiver/decoder, automatic-pilot coupler, automatic pilot, and radar augmentor are required. Automatic throttle control is also required in the aircraft, although not directly connected to the automatic landing system. The general system layout is shown in the **illustration**.

An aircraft is flown through a prescribed area, called a gate, where the radar locks on and tracks the aircraft to the touchdown. The angular tracking information of the radar is fed from digital encoders to the computer. The digitally encoded information from a stable element that tracks the ship motion is also fed to the computer. A data stabilization routine in the digital computer removes the effects of ship's roll, pitch, yaw, and heave from the angular tracking data. The corrected data, which precisely locate the aircraft in space, are fed to the flight computer routine in the digital computer, where they are compared with a stored flight path optimized for the type of aircraft under control. Error signals are generated which are transmitted to the aircraft by means of a ground-to-air data link and displayed to the pilot on a cross-pointer indicator or heads-up display. Also stored in the digital computer is sufficient aerodynamic control information for each aircraft type to convert error signals to commands to the aircraft autopilot, taking into account the response characteristics for different aircraft types. Autopilot commands are transmitted to the aircraft by means of a ground-to-air data link in the same message used to transmit error signals. To limit the dispersion in heavy seas, deck motion compensation is used during the last 12 s to fly the aircraft to the vertical position of the carrier deck touchdown point at impact.

Use of two systems. In practice, two complete systems are used on each carrier to double the landing rate and to provide a high systems availability in the event of component failure. Each system uses a conical scan radar antenna, with a choice of

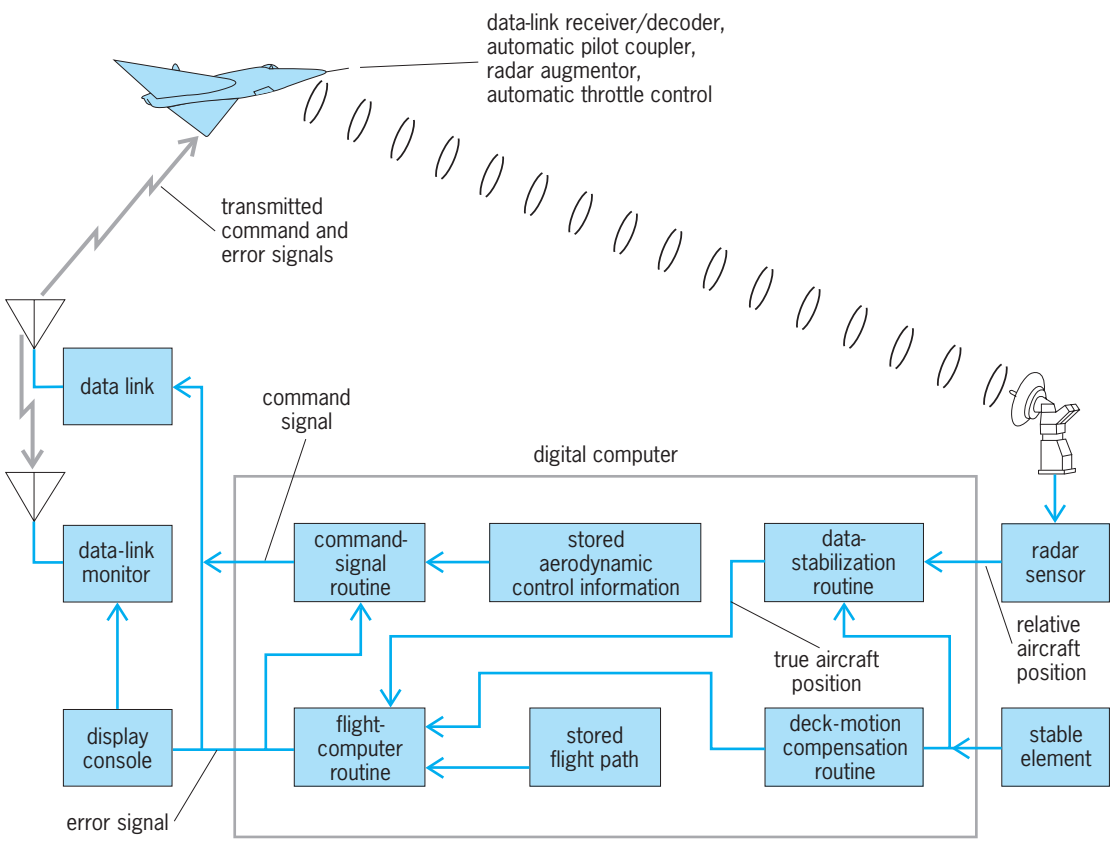

Schematic of radar-beam system for aircraft-carrier and shore applications.

vertical or circular polarization. A 4-ft (1.2-m) reflector provides a beam width of 0.5°. The radar operates in the 33.0–33.4-GHz frequency band, with a pulse repetition rate of 2000 pulses per second and a peak power output of 40 kW. A radar augmentor (radar beacon or corner reflector) eliminates target scintillation and aids in penetration of heavy rainfall. *See* RADAR.

The digital computer has an 18-bit-word and 32,000-word memory. The memory recycling time is 2 microseconds. Each computer performs all computations for two approaching aircraft at a rate of 20 times per second. At the same time the computer performs on-line diagnostics throughout both systems. The second computer serves as an off-line monitor. Addressed digital data-link messages containing error and command signals are transmitted to each aircraft at the rate of 10 times per second. In addition, discrete signals are transmitted to indicate to the pilot such things as radar lock-on, 10 s to touchdown, and wave-off.

A display console is provided to monitor automatic approaches. The display provides an AZ-EL presentation which permits the console operator to "talk down" the aircraft, as in ground-controlled-approach (GCA) systems. Talk-down can be provided if the aircraft is not equipped with a data link or if there is a failure in the data link, coupler, or automatic pilot. Flight path, airspeed, rate of descent, and ship's motion and impact velocity are recorded for each landing, and may be used for debriefing pilots when manual landings are made for training purposes. *See* PRECISION APPROACH RADAR (PAR).

The landing gate may be placed 2–8 mi (3–13 km) from touchdown. Assuming a landing speed of 120 knots (60 m/s), the landing rate for aircraft may vary from 1 to 4 min per radar. Provision of two radars permits the landing interval aboard a carrier to be varied from $^1/_2$ to 2 min. For carrier operations, 68% of the landings must be within ±20-ft (6-m) longitudinal and ±10-ft (3-m) lateral dispersion when there is no deck motion. For heavy seas, longitudinal dispersion can be degraded to ±40 ft (12 m).

Air-derived landing system. The air-derived system is designed to take advantage of the large number of ILS installations in worldwide use under the international standards and recommended practices of the International Civil Aviation Organization. Its use, largely in Europe, is confined to the larger multiengine aircraft, because of the size and complexity of the airborne installation. Part of this complexity is necessary to reduce the effect of electrical noise on the localizer and glide-slope beam signals caused by multipath. For example, an aircraft taking off over the localizer antenna may cause wide perturbation of the beam in an approaching aircraft for more than a second. A time-referenced scanning-beam microwave landing system (MLS) is under development and is expected to significantly reduce problems due to multipath.

In the ILS-based system the position of the localizer and glide-slope beams is detected in the aircraft. The autopilot is commanded to seek the ILS fixed-beam centers. This action will describe the runway centerline extension and a fixed glide slope of about 3°. At a flare-initiate height determined by the radio altimeter (about 50 ft or 15 m), the autopilot is commanded to execute a preprogrammed flare to reduce the rate of descent from as much as 10 ft/s (3 m/s) to 1.5–2 ft/s (0.4–0.6 m/s) at touchdown. An automatic throttle control is used (as in the radar-beam system) to maintain the pilot-selected airspeed during changes in the aircraft pitch and bank.

Requirements for the airborne digital computer are similar to those indicated for the ground digital computer of the radar-beam system. The radar sensor is replaced by the ILS receiver, and the stable element by an aircraft inertial navigation system or accelerometers. Measurement of vertical and lateral accelerations is required. The best estimate of aircraft position is determined in the computer by filtering and smoothing the ILS receiver horizontal and vertical inputs. In performing filtering and smoothing, true-motion inputs as determined by the accelerometers are used by the computer to reduce multipath and noise peaks from the ILS receiver inputs. A filter and smoothing routine replaces the data stabilization routine. Rate of change in position is based upon changes in true position after filtering and smoothing. The deck-motion routine of the radar-beam system is not required ashore and is replaced by a flare-path computer routine. During flare, glide-slope information from the ILS receiver is not used. Starting at the flare-initiate height, vertical accelerations are altered to provide a cumulative vertical deceleration along an exponential curve to touchdown. Lateral guidance is still referenced to the ILS localizer input signal. An additional computer routine is required to compensate for crosswinds experienced on the landing path, whereas no crosswind compensation is provided for landing aboard an aircraft carrier, because the carrier is headed into the wind during landing operations.

A pilot normally "crabs" (steers the aircraft toward the wind) to compensate for crosswind. This operation has limitations when used for automatic landings on land. First, the runway is not directly ahead in the pilot's field of vision and is more difficult to locate when trying to establish visual contact with the ground. Also, the aircraft response to rudder control, required to perform de-crab, is marginal to inadequate. This response is particularly poor for aircraft that land with a high-angle-of-attack (nose-up) attitude. ILS-based automatic landing systems use a forward slip maneuver to compensate for the crosswind. In this maneuver the computer controls the ailerons to bank the aircraft while maintaining the aircraft heading along the landing path. This form of compensation keeps the pilot's head pointed to the runway and takes advantage of the more rapid aircraft response to aileron control for removing the wing-down before touchdown. To eliminate a possibility of having a wing contact the ground before the wheels, the amount of wing-down is restricted; hence the forward slip can compensate only for a limited crosswind. If the restricted forward slip is not sufficient to compensate for the crosswind (perhaps above 30 knots or 15 m/s with a wide-body aircraft),

crab is added to the forward slip. This is a very unusual condition and crab is held to a minimum. The removal of forward slip and any de-crab must occur very close to the ground to ensure that the aircraft is not blown too far off centerline and that aircraft weight is evenly distributed on the landing gear at touchdown. Crosswind compensation is a complex computer routine and must be carefully tailored to the flight characteristics of each aircraft type.

Since each aircraft computer serves only one aircraft, its computation and memory requirements are about half of those required for the ground computer. The command signal routine directly interfaces the autopilot coupler. Aircraft displays are energized by ILS receiver outputs. *See* ACCELEROMETER; INERTIAL GUIDANCE SYSTEM.

To provide the degree of safety required for airline operations, the ILS-based automatic system uses triple-redundant control systems in the aircraft. This requirement is due in part to the lack of an independent landing monitor for the pilot or on the ground. The use of triple redundancy increases the complexity, size, weight, and cost of the airborne ILS-based system. In addition, civil systems are subject to a very rigorous certification process that further improves their reliability. The ground-based system does permit ground monitoring as a part of the system but not independently. Both classes of systems can provide adaptive computer programs to compensate for reduced aircraft or equipment performance. John L. Loeb; Douglas Hodgkins

Bibliography. T. A. Bombard et. al., Future DOD precision approach and landing architecture, *Proc. ION Annual Meeting*, *June 1998*, pp. 627–636, 1998; M. Kayton and W. Fried, *Avionics Navigation Systems*, 2d ed., 1995; R. J. Kelly and E. F. C. LaBerge, MLS: A total system approach, *IEEE AES*, 5(5):27–39, May 1990; B. R. Peterson, G. Johnson, and J. Stevens, Feasible architectures for Joint Precision Approach and Landing System (JPALS) for land and sea, *Proc. ION GNSS Meeting*, *September 2004*, pp. 544–554, 2000.

Automatic sprinkler system

An integrated arrangement of fixed facilities for protection from combustion by use of water extinguishment. The system comprises an adequate water supply, hydraulically designed internal piping, and sprinklers connected in a systematic pattern over the protected area; the system is activated by a fire to discharge a fine spray of water over the heat.

Essential features of a system are its self-detection of fire, prior installation, and built-in activation. In these respects, the automatic sprinkler system is among the earliest-used architectural features that contribute actively to maintenance of the internal environment (in contrast to the passive fire-resistant contribution of the static structure). Auxiliary to an automatic sprinkler system may be a fire alarm. *See* FIRE DETECTOR.

Effectiveness. Amenable to all occupancy classes of the National Building Code, automatic sprinklers are the most widely used fixed apparatus for fire protection. In extinguishment of class A fires and, in adapted system forms, of class B and class C fires, the water spray acts four ways: (1) it cools burning material by conversion of water to steam; (2) the steam displaces the oxygen supply, thereby tending to smother the fire; (3) by dampening materials in the area the spray limits the supply of new fuel; and (4) while the spray is falling, it lowers the temperature in the vicinity by evaporative cooling.

Through prompt response, an automatic sprinkler system generally requires less water to control a fire than does a hose. Also, the vertical spray produces less mechanical damage than does a horizontal hose stream. A sprinkler system is considered effective if it extinguishes or checks a fire until fire-fighting forces arrive. Failures of such systems are principally where they have been made inoperative during building alteration or disuse, or the occupancy hazard has been increased beyond initial system capability.

Not all hazards are accessible to sprinkler protection. For example, closely packed stock piled over two stories high sheds water so that the bottoms of these piles may not be wetted from overhead sprinklers. Thus, management of materials handling and storage and of housekeeping is necessary to stay within the capabilities of the installed sprinklers.

Elements of a system. The water supply for a sprinkler system is separate from that used by a fire department. Normally no water flows in the supply lines to sprinklers, thus freezing is a greater risk than in mains with continuous flow. Standards require sprinkler mains to be buried well below the frost line. The underground main enters the building in a heated area to supply a riser.

Connected at the riser are valves, meters, and often an alarm to sound when flow exceeds discharge from one sprinkler. At the top of this vertical riser, a horizontal network of pipes extends throughout the fire compartment in the building. Other risers feed distribution networks to systems in adjacent fire compartments. Compartmentalization divides a large building horizontally, on a single floor, and vertically, floor to floor. Thus, several sprinkler systems may serve one building.

In the distribution network, branch lines carry the sprinklers. A sprinkler may extend up from a branch line (upright configuration), placing the sprinkler close to the ceiling, or a sprinkler can be below the branch line (pendant configuration). For use with concealed piping, a flush-mounted pendant sprinkler extends only slightly below the ceiling.

For a high-rise building—one in which a fire must be fought internally because of height—hydraulic design must position a water source to supply sprinklers for not less than 30 min. The flow-actuated alarm produces a location indication at a ground-level register accessible to the fire department.

Sprinklers. The principal component of the system is a thermally sensitive sprinkler with a linkage assembly that holds closed the discharge opening. In various designs the assembly is disrupted through a low-melting-point chemical, a frangible bulb filled

with liquid, a bimetallic disk, or—usually—a low-melting-point alloy link. The linkage separates above the operating range, which may be any one of the six standard steps from 100°F (38°C) to 475°F (246°C). Then the sprinkler abruptly opens to discharge water against a deflector so that water falls in a hemispherical spray across the area below. Where water damage ranks in importance with fire damage, stop-and-go sprinklers with bimetallic heat sensors open in response to heat and close automatically when heat subsides; they recycle if necessary to track the fire cycle.

Clearance for the spray to spread is maintained between the top of storage, such as library stacks, and sprinkler deflectors. If clearance cannot be maintained, branch lines follow aisles (alternate aisles in the absence of vertical back-to-back shelf dividers). If ventilation extends through decks in a library, sprinkled aisles are staggered vertically.

Modes of operation. Water reaches the sprinklers variously in five basic setups. In the usual wet-pipe system, for heated buildings, all pipes contain water under pressure for immediate release through any sprinkler that opens.

In a dry-pipe system—for unheated buildings in freezing climates or for cold-storage rooms—branch lines and other distribution pipes contain dry air or nitrogen under pressure. This pressure holds closed a dry-pipe valve at the riser. When heat from a fire opens a sprinkler, the air escapes and the dry-pipe valve trips; water enters branch lines and fire suppression begins.

For occupancies where flash fires are possible, a deluge system is appropriate. All sprinklers are continuously open while the pipes are empty. Separate fire detectors throughout the sprinkled compartment are of a type appropriate to the occupancy hazard. In response to combustion, the detectors actuate a valve at the riser. All sprinklers then deluge the entire area, both to suppress the fire and to wet down unburned combustibles.

In a preaction system, both heat-sensitive sprinklers and separate detectors guard the area; the separate detectors open a preaction valve and sound an alarm in the event of fire. The alarm provides opportunity for local hand control of the fire; otherwise heat fuses the links of local sprinklers, designed to operate above the temperature of the auxiliary detectors. These self-selected sprinklers then discharge.

Another adaptation is the recycling preaction system, applicable to unattended buildings. For example: heat from a fire actuates an auxiliary detector at 140°F (60°C). The detector in turn operates a master valve at the riser to admit water to branch lines. If the temperature continues to rise, at 160°F (71°C), to which these sprinklers are fused, one or more sprinklers respond. After these sprinklers bring the fire under control, the temperature drops and the auxiliary detector closes the riser valve. The water flow recycles if necessary to follow the fire cycle.

For overall building protection, a water sprinkler system may serve as backup to a first-response extinguishment facility that protects fire loads not directly extinguishable by water spray. *See* FIRE EXTINGUISHER; FIRE TECHNOLOGY. Frank H. Rockett

Bibliography. W. K. Bare, *Fundamentals of Fire Protection*, 1977; E. G. Butcher and A. C. Parnell, *Smoke Control in Fire Safety Design*, 1978; M. D. Egan, *Concepts in Building Fire Safety*, 1978; G. P. McKinnon (ed.), *Fire Protection Handbook*, 18th ed., 1997; J. Morris, *Managing the Library Fire Risk*, 1979.

Automation

The process of having a machine or machines accomplish tasks hitherto performed wholly or partly by humans. As used here, a machine refers to any inanimate electromechanical device such as a robot or computer. As a technology, automation can be applied to almost any human endeavor, from manufacturing to clerical and administrative tasks. An example of automation is the heating and air-conditioning system in the modern household. After initial programming by the occupant, these systems keep the house at a constant desired temperature regardless of the conditions outside.

Reasons for automating. The complete or partial automation of any task is usually based upon the following considerations:

1. *Productivity*. Automating a task usually results in an increment in productivity, where productivity is defined as tasks performed per unit time.

2. *Efficiency*. Automation has the potential to increase the efficiency in completing a task, where efficiency is defined as inversely proportional to the amount of raw material (for example, electrical power) consumed.

3. *Quality*. Automating a task can significantly increase the quality of a final product or process. Quality can be defined in many ways with many attributes. An obvious definition is simply the degree of excellence which a thing possesses. Quality could refer to the consistency with which a product is manufactured to meet that degree.

4. *Safety*. By removing direct human intervention or activity, automation has the potential to significantly increase the safety of any task. This is particularly true in the area of manufacturing.

5. *Flexibility*. Automation can significantly enhance the flexibility of a process, where flexibility refers to the ability to alter the number or nature of the tasks completed per unit time.

6. *Cost*. While each of the previous items can be related directly or indirectly to the cost of completing tasks, cost is often the primary motivation behind automating tasks previously undertaken partially or completely by humans. Chief among these cost savings is that associated with labor, including salaries, benefits, training, and so forth.

Enabling technologies. Historically, advances in automation have depended upon advances in the requisite enabling technologies, such as digital computation. The fundamental constituents of any automated process are (1) a power source, (2) a feedback control mechanism, and (3) a programmable

command structure. Programmability does not necessarily imply an electronic computer. For example, the Jacquard loom, developed at the beginning of the nineteenth century, used metal plates with holes to control the weaving process. Nonetheless, the advent of World War II and the advances made in electronic computation and feedback have certainly contributed to the growth of automation. While feedback is usually associated with more advanced forms of automation, so-called open-loop automated tasks are possible. Here, the automated process proceeds without any direct and continuous assessment of the effect of the automated activity. For example, an automated car wash typically completes its task with no continuous or final assessment of the cleanliness of the automobile. *See* CONTROL SYSTEMS; DIGITAL COMPUTER.

Categorization of automated tasks. Because of the growing ubiquity of automation, any categorization of automated tasks and processes is incomplete. Nonetheless, such a categorization can be attempted by recognizing two distinct groups, automated manufacturing and automated information processing and control. Automated manufacturing includes automated machine tools, assembly lines, robotic assembly machines, automated storage-retrieval systems, integrated computer-aided design and computer-aided manufacturing (CAD/CAM), automatic inspection and testing, and automated agricultural equipment (used, for example, in crop harvesting). Automated information processing and control includes automatic order processing, word processing and text editing, automatic data processing, automatic flight control, automatic automobile cruise control, automatic airline reservation systems, automatic mail sorting machines, automated planet exploration (for example, the rover vehicle, *Sojourner*, on the *Mars Pathfinder* mission), automated electric utility distribution systems, and automated bank teller machines. *See* ASSEMBLY MACHINES; COMPUTER-AIDED DESIGN AND MANUFACTURING; COMPUTER-INTEGRATED MANUFACTURING; FLEXIBLE MANUFACTURING SYSTEM; INSPECTION AND TESTING; SPACE PROBE; WORD PROCESSING.

Elements of automation process. The relationships of the key elements in an automated task or process, particularly the three fundamental constituents outlined above, can be shown in a schematic diagram (**Fig. 1**). Nearly every automated task can be interpreted in terms of this diagram. If no feedback is associated with the task, that is, if it is open-loop in nature, then the feedback element (2) is simply omitted. The increasing pace of automation is attributable to the increased capabilities of digital computers, which not only provide the programmable command but also form the feedback element when implemented as a digital control system. The precise form and function of these elements depend on the particular automation process, as in the following examples. *See* DIGITAL CONTROL.

Robotic assembly machine. A robotic assembly machine includes the robot itself, a conveyor, an assembly pad, a manipulator, a camera, and associated signal processing equipment (**Fig. 2**). Guided by information from the camera, the robot takes parts from the conveyor, assembles them, and places the assembled parts back on the conveyor. The power source is assumed here to be electrical. The command programming serves as one input to the robot controller. The feedback activity is exemplified by the information received from the camera and fed back to the controller. If a human assembler were completing this task, activities paralleling those just described would occur. That is, the worker's vision system would replace the camera, his or her limbs and hands would replace the robot arm and manipulator, and a feedback loop would be effectively utilized between visual inputs and motion. *See* ROBOTICS.

Automatic flight control system. The constituent elements of an automatic flight control system include the pilot, aircraft, flight control computer, control surface actuator, and control surface (**Fig. 3**). The pilot can be replaced by a partially automated

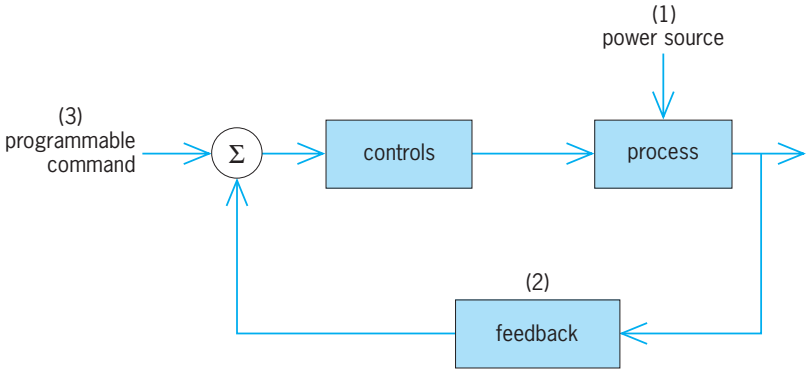

Fig. 1. Elements of an automated system.

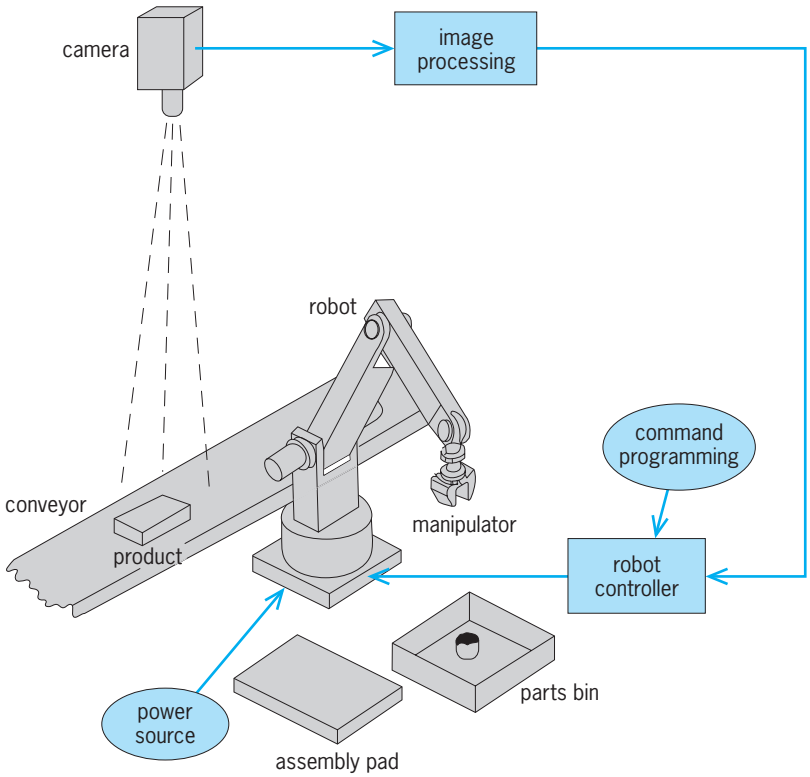

Fig. 2. Simplified representation of a robotic assembly machine.

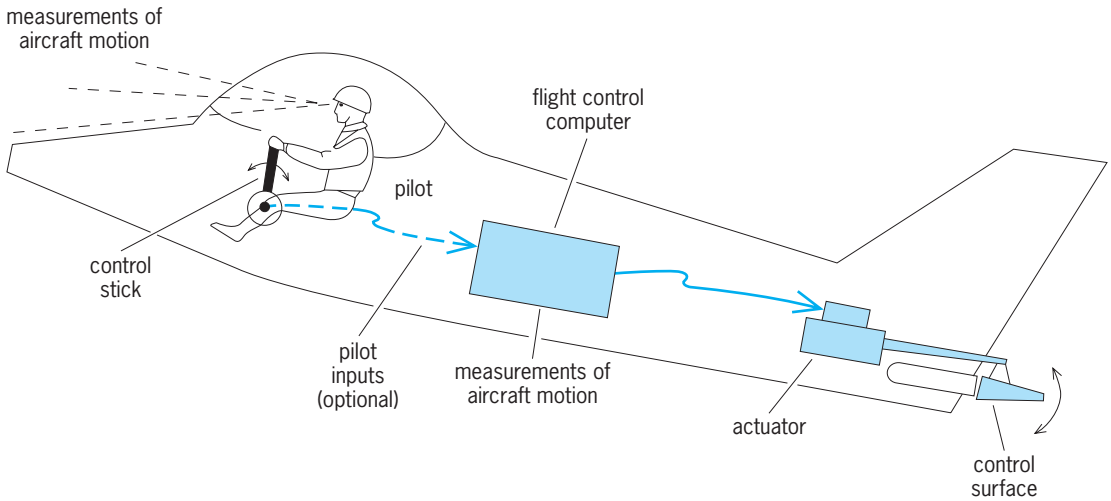

Fig. 3. Simplified representation of an automatic flight control system.

system in which pilot inputs are still possible. The flight control computer receives and processes information from onboard systems (possibly including the pilot), which sense aircraft angular and linear motion (for example, rate of change of pitch-attitude). The processed information is then used to drive the actuator (a hydraulic or electrohydraulic device), which moves the control surface to change the attitude of the aircraft. The power source (not explicitly shown in Fig. 3) would likely be electrical. The command programming is contained in the flight control law, which is coded into the flight control computer. The controller activity is also carried out in the flight control computer, and the feedback activity occurs when the aircraft motion is measured (by inanimate sensors or the pilot) and fed back to the computer. *See* FLIGHT CONTROLS.

Effects on human activity. One obvious societal implication of increasing automation is the elimination and displacement of human labor. While there are positive aspects of such displacement, for example, the elimination of dangerous tasks which required human participation, the negative aspects cannot be overlooked. For example, as automated manufacturing becomes more prevalent, the size of the human work force that supports such enterprises naturally shrinks and the skill requirements of the new work force increase considerably. For example, the type of human activity in an automated factory is significantly different from that in a nonautomated one and can be summarized as follows:

1. *Equipment maintenance*. In an automated manufacturing facility, equipment maintenance becomes a critical issue. Malfunctions or breakdowns in one part of the facility can halt the entire production line.

2. *Computer programming*. Since the digital, programmable computer forms the backbone of the automated manufacturing system, programming the computers, entering pertinent product and process information, and updating manufacturing databases constitute an integral part of the automated factory. *See* COMPUTER PROGRAMMING.

3. *Updating production technologies*. Economic competitiveness requires nearly constant improvement in the production systems in automated manufacturing facilities. This entails equipment redesign, replacement, and upgrades.

4. *Production supervision*. Managing automated manufacturing facilities requires personnel with skills considerably different than those employed in nonautomated plants. Greater emphasis is placed upon the technical skills of the automated plant supervisor as compared to those of the supervisor concerned primarily with personnel management.

The problems and opportunities just outlined are not limited to automated manufacturing. For example, the introduction of automated systems into airline cockpits has reduced the required crew from three to two (with the elimination of the flight engineer). In addition, the required skill set of the airline crew now includes cockpit resource management, which describes the ability of the crew to effectively utilize the capabilities of the modern cockpit, for example, the computerized flight management system. *See* AIRCRAFT INSTRUMENTATION.

Human-machine interaction. While the ability to automate manufacturing and information/control tasks is enhanced by advances in computer technology, the effective application of automation remains a challenge, particularly as it involves tasks and processes requiring the automated system to act in consort with humans. Before the advent of automation, the human often assumed the roles associated with controls and feedback in an automated system (Fig. 1). In the aircraft piloting example, and prior to the introduction of automation, it was the human pilot's visual and vestibular (motion) sensing which provided the feedback elements, and it was the pilot's operation upon the difference between these sensed variables and a command or desired state that resulted in an input to the aircraft through the control stick or throttle. As cockpit automation has progressed, however, the responsibilities for feedback and control have increasingly been assumed by the aircraft automatic flight control system. Human

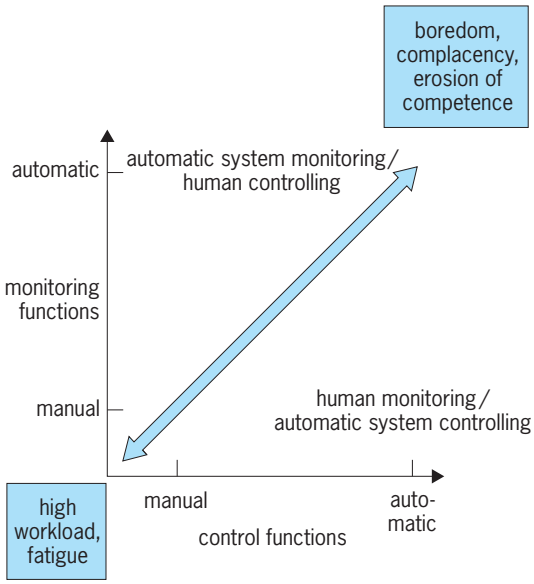

Fig. 4. Trade-off between manual (human) and automatic (machine) monitoring and control.

presence in the cockpit is still required from the standpoint of safety and passenger confidence. The crew must continuously monitor the performance of the automatic flight control system in order to reassume an active feedback and control role should the automatic system fail or situations arise which are beyond the capabilities of this system. In this context, the term "appropriate automation" describes the optimal distribution of responsibility between human and machine in tasks such as airline flight. The effects of this distribution of responsibility between human (pilot) and machine (automatic flight control system) can be demonstrated by means of a simple graph (**Fig. 4**): In the lower left (no automation) human workload is high and fatigue is a concern, while in the upper right (complete automation) human boredom and possible erosion of competence and piloting skills can occur.

In addition to workload and skill atrophy, the successful utilization of a completely or partially automated system depends upon the ease with which the human can communicate with the system. This communication can take the form of human supervision of the automated system, human "training" of the automated system, or humans' reasserting direct control of the task being performed by the automated system. The importance of this concern has led to the concept of human-centered automation, that is, an automation philosophy which considers from the outset that humans will, of necessity, be required to interact with the system. Such a discussion leads to the topic of function allocation between human and machine. *See* HUMAN-MACHINE SYSTEMS.

Function allocation. A major issue in the design of systems involving both human and automated machines concerns allocating functions between the two, that is, answering the question, "What should the (automated) machine do and what should the human do?" This allocation can be static or dynamic. Static allocation is fixed; that is, the separation of responsibilities between human and machine do not change with time. Dynamic allocation implies that the functions allocated to human and machine are subject to change. Historically, static allocation began with reference to lists of activities which summarized the relative advantages of humans and machines with respect to a variety of activities. For example, at present humans appear to surpass machines in the ability to reason inductively, that is, to proceed from the particular to the general. Machines, however, surpass humans in the ability to handle complex operations and to do many different things at once, that is, to engage in parallel processing. Dynamic function allocation can be envisioned as operating through a formulation which continuously determines which agent (human or machine) is free to attend to a particular task or function. In addition, constraints such as the workload implied by the human attending to the task as opposed to the machine can be considered. *See* HUMAN-FACTORS ENGINEERING.

Artificial intelligence. It has long been the goal in the area of automation to create systems which could react to unforeseen events with reasoning and problem-solving abilities akin to those of an experienced human, that is, to exhibit artificial intelligence. Indeed, the study of artificial intelligence is devoted to developing computer programs that can mimic the product of intelligent human problem solving, perception, and thought. For example, such a system could be envisioned to perform much like a human copilot in airline operations, communicating with the pilot via voice input and spoken output, assuming cockpit duties when and where assigned, and relieving the pilot of many duties. Indeed, such an automated system has been studied and named a pilot's associate. Machines exhibiting artificial intelligence obviously render the sharp demarcation between functions better performed by humans than by machines somewhat moot. While the early promise of artificial intelligence has not been fully realized in practice, certain applications in more restrictive domains have been highly successful. These include the use of expert systems, which mimic the activity of human experts in limited domains, such as diagnosis of infectious diseases or providing guidance for oil exploration and drilling. Expert systems generally operate by (1) replacing human activity entirely, (2) providing advice or decision support, or (3) training a novice human in a particular field. *See* ARTIFICIAL INTELLIGENCE; EXPERT SYSTEMS.

Ronald A. Hess

Bibliography. M. P. Groover, Automation, in R. C. Dorf and A. Kusiak (eds.), *Handbook of Design, Manufacturing and Automation*, Wiley, 1994; B. B. Morgan, Jr., et al., Implications of automation technology for aircrew coordination and performance, in W. B. Rouse (ed.), *Human/Technology Interaction in Complex Systems*, JAI Press, 1993; T. B. Sheridan, *Telerobotics, Automation, and Human Supervisory Control*, 1992; C. D. Wickens, *Engineering Psychology and Human Performance*, 2d ed., HarperCollins, 1992.

Automobile

A self-propelled land vehicle, usually having four wheels and an internal combustion engine, used primarily for personal transportation (**Fig. 1**). Other types of motor vehicles include buses, which carry large numbers of commercial passengers, and medium- and heavy-duty trucks, which carry heavy or bulky loads of freight or other goods and materials. Instead of being carried on a truck, these loads may be placed on a semitrailer, and sometimes also a trailer, forming a tractor-trailer combination which is pulled by a truck tractor. *See* BUS; TRUCK.

Design. The basic design of the automobile was standardized by 1908, when the Ford Motor Company began production of the Model T. A metal frame served as the main structural member that supported the power train and body. The frame was supported through springs by four wheels and tires. The engine was mounted at the front of the vehicle and transmitted power to the two rear wheels. The body was open, providing little protection from the weather.

By the early 1930s, most automobiles had a closed body made of stamped steel parts that were welded together before attachment to the frame. Improved versions of the front-engine rear-wheel-drive power train dominated automotive production into the 1970s. Until then, this basic configuration was never seriously challenged, although some vehicles had either front-engine front-wheel drive or rear-engine rear-wheel drive.

In the United States, the change from rear drive to front drive was largely necessitated by three acts passed by Congress which established new laws covering automotive air pollution, automotive safety, and automotive fuel economy. The solutions to problems of automotive air pollution and safety were often add-on devices that allowed the conventional automobile to be manufactured in compliance with the new regulations. However, the additional weight and action of some of these devices reduced fuel economy.

When compliance with fuel-economy standards also became a design objective, automotive stylists and engineers began redesigning the car by reducing vehicle weight, downsizing new models, using smaller engines, and changing to lighter, stronger materials. The result was an evolutionary redesign that allowed continuing production while the required new parts and vehicles were developed and phased into each manufacturer's passenger-car fleet.

Automotive air pollution. Historically, the design and manufacture of automotive vehicles in the United States was an unregulated industry. In 1955, the Department of Health, Education, and Welfare was instructed to begin a study of air pollution. By 1962, the automobile engine had been identified as the source of more than 40% of airborne pollutants, making the automobile the largest single contributor to air pollution. *See* AIR POLLUTION.

In 1963, Congress passed the Clean Air Act which, together with its subsequent amendments, directed the Environmental Protection Agency (EPA) to promulgate automotive emission standards and to regulate automotive engine fuels and fuel additives. Because of unusual air-pollution problems in Los Angeles, the Clean Air Act permitted the state of California to set more stringent standards than the rest of the nation. *See* DIESEL FUEL; GASOLINE.

To comply with automotive emission standards, emission-control devices were developed to reduce or eliminate air pollutants from the engine crankcase,

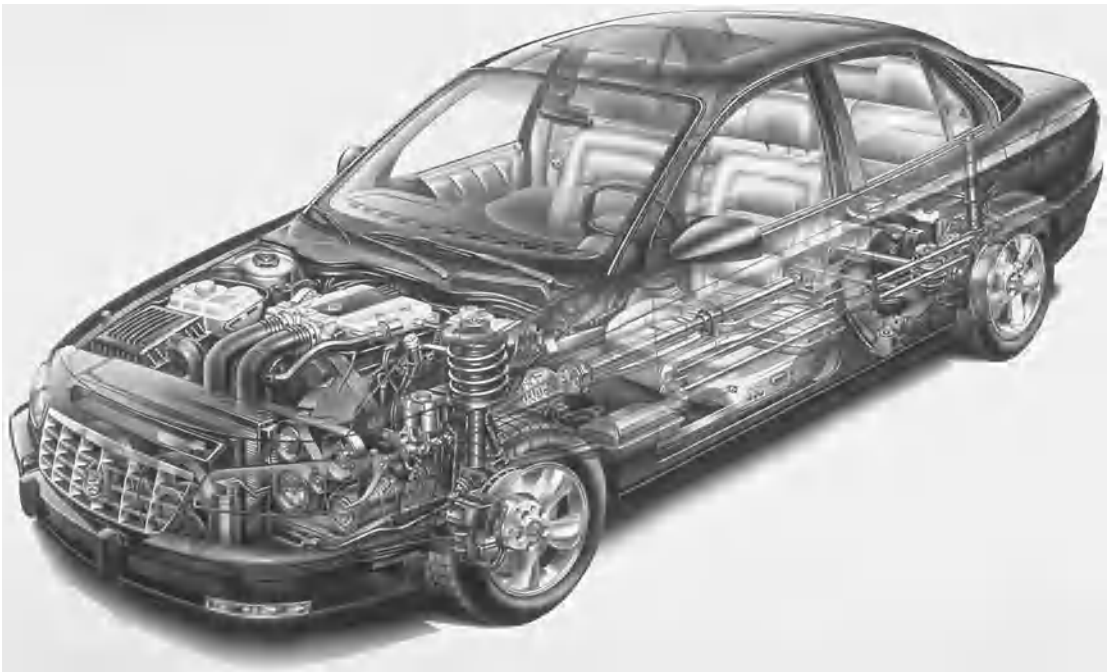

Fig. 1. Phantom view of a front-engine automobile with unitized-body construction. (*Cadillac Division, General Motors Corp.*)

engine exhaust gas, and fuel tank and carburetor. However, some devices reduced fuel economy as well as emissions, resulting in 1968–1974 automobiles having worse fuel economy than earlier models.

In 1973, however, the EPA moved toward requiring automobile engines to run on unleaded gasoline (tetraethyllead was formerly used as an additive for motor fuels to reduce engine knock). Beginning in 1975, a stepped phase-down limited the amount of tetraethyllead allowed in gasoline. Tetraethyllead impairs the operation of some engine emission-control devices and, after the additive becomes airborne in the engine exhaust gas, creates a human and environmental health hazard. *See* LEAD; TOXICOLOGY.

Automotive safety. During the early 1960s, traffic deaths in the United States approached 50,000 per year. Congress began passing legislation in an attempt to reverse the upward trend. The National Traffic and Motor Vehicle Safety Act (1966) required that all new motor vehicles sold in the United States be manufactured in compliance with Federal Motor Vehicle Safety Standards. The first 20 such standards were issued in 1967 and became mandatory on all 1968 models.

Automotive fuel economy. In 1973, Arab nations imposed a total ban on oil exports to the United States. This resulted in a shortage of gasoline and diesel fuel, and caused a consumer demand for smaller cars which were lighter and more fuel-efficient. As the limited amount of fuel that was available rapidly increased in price, the larger, more powerful rear-drive cars built by manufacturers in the United States were gradually replaced by smaller, more fuel-efficient cars imported from Europe and Japan.

To help reduce the dependence of the United States on imported oil, Congress passed the Energy Policy and Conservation Act (1975), which established corporate average fuel economy standards to be met by new cars produced by each automotive manufacturer. To essentially double the fuel economy of the nation's passenger-car fleet in 10 years, Congress established a 1985 corporate average fuel economy standard of 27.5 mi/gal (2 km/liter). This could be achieved by the manufacturers' producing smaller, lighter cars powered by smaller engines.

Weight was reduced by redesigning the basic vehicle structure, eliminating excess material, and continuing the change to lighter and stronger materials. Fuel economy was further improved by introducing more aerodynamic body styles, reducing rolling resistance, and reducing friction and power losses in engines and other components. To maximize these gains, most new car designs had front drive with a transversely mounted integrated power train in which the engine, transmission, final drive, and differential form a single unit.

By the 1980s, the downsized front-engine front-drive car had become the most fuel-efficient and widely produced automobile design. However, some data indicated that the new lighter and smaller car had higher personal injury claims and damage repair costs after a collision than the full-size car.

Manufacturing. In automobile manufacturing, production is the making of the thousands of parts, subsystems, and modular assemblies that comprise the automobile, while assembly is the fitting together of these components. Typical assembly of a passenger car begins when the steel pieces that form the side openings for mounting the doors are assembled and spot-welded together by robot welders. The side assemblies are then welded to the underbody, which consists of pieces of galvanized steel that have been welded together to form the floor pan, engine compartment, and luggage compartment. Welding helps prevent noise—squeaks, rattles, and vibration—that could be caused later by parts moving while the vehicle is on the road. Roof, quarter panels, doors, deck lid, and hood are also welded during fabrication so that a minimum of mechanical fasteners is required.

The body assembly is dipped in a phosphate bath to clean any debris from the surface and then submerged in an anticorrosion solution, such as zinc phosphate. Fillers and sealers are applied to body joints, seams, and cavities that might leak under stress and collect or admit moisture. Sound-deadening material and additional sealer and primer are applied where necessary. A chip-resistant urethane coating is sprayed onto the lower body to provide extra protection against stones flying up from the road. *See* METAL COATINGS.

Robot painters apply the main color, while workers paint spot areas on the body. The paint is sealed with a clear coating to produce a high gloss finish. The hard trim, which includes the instrument panel, steering column, weather stripping, and body glass, is installed. Then, after testing for water leakage, the soft trim is installed. This includes carpeting over a layer of sound-deadening material, seats, door pads, roof insulation, and upholstery. Adhesive bonding may be used, especially when joining various types of nonmetal parts.

The body moves along an assembly line that is above a conveyor line on which the engine and other chassis components are moving (**Fig. 2**). The chassis is the complete operable vehicle, but without the body. A manufacturer's designation for a specific chassis is often called a platform. Changing outside body panels and interiors allows different models to be built on the same platform. All models built on that platform have the same tread width, but wheelbase may vary.

As the engine, complete with transmission or transaxle and exhaust system, is prepared for installation, the fuel tank and bumpers may be fitted to the body. Final assembly occurs when the engine and chassis are raised into position, mounting bolts are installed, and the wheel-and-tire assemblies are attached. Wheel alignment is tested under simulated road conditions as the vehicle undergoes final inspection and drive-away for delivery to a dealer or distribution center.

Body. The automobile body is the assembly of sheet-metal, fiberglass, plastic, or composite-material panels together with windows, doors, seats, trim and upholstery, glass, and other parts that form

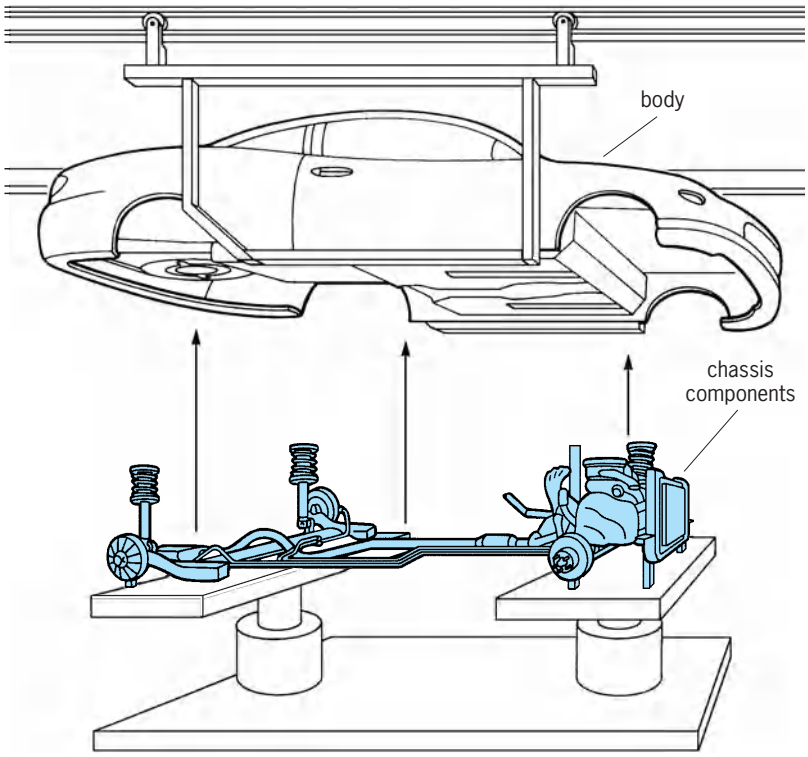

Fig. 2. Final assembly of the automobile, as the engine and chassis are raised into position and fastened to the body. (*Chevrolet Division, General Motors Corp.*)

enclosures for the passenger, engine, and luggage compartments. The assembled body structure may attach through rubber mounts to a separate or full frame (body-on-frame construction), or the body and frame may be integrated (unitized-body construction). In the latter method, the frame, body parts, and floor pan are welded together to form a single unit that has energy-absorbing front and rear structures, and anchors for the engine, suspension, steering, and power-train components. A third type of body construction is the space frame which is made of welded steel stampings. Similar to the tube chassis and roll cage combination used in race-car construction, non-load-carrying plastic outer panels fasten to the space frame to form the body. *See* COMPOSITE MATERIAL; SHEET-METAL FORMING; WELDED JOINT.

Size and style. Automobiles are built in several sizes and body styles and may be further classified according to the number of doors and the intended usage. Basic body styles include the hatchback, two-door coupe or sedan, roadster, convertible, four-door sedan, and station wagon. Passenger-car sizes range from the largest full-size car to the compact and small sizes. Light-duty trucks, which include pickups, suburbans, vans, minivans, and sport-utility vehicles, also are available in various sizes.

Many full-size cars are two- or four-door sedans that seat from four to six and have a separate luggage compartment. Some smaller cars are hatchbacks, which have two or four doors, no separate luggage compartment, and a rear lid that lifts up for access to the rear of the passenger compartment. For carrying cargo in a hatchback the available space may be enlarged by folding down the rear seat. The station wagon has two or three seats, and a cargo area that may be enlarged by folding down the rear seat in a two-seat wagon, and the center and rear seats in some three-seat wagons.

Roof options. Several roof options are available, including the vinyl roof, popup sunroof, electric slide sunroof, and T roof. The vinyl roof is a covering of colored vinyl that is laid over the finished roof of the car, primarily for decoration. The sunroof is a panel in the roof above the front seat that can be opened for ventilation; it may pop up, or it may slide open when moved by a hand crank or by an electric motor. A T roof has hatch panels above the driver and passenger that can be removed. Each panel fits between the top of the door glass and a center T bar, which runs from the front section of the roof to a point in back of the front seat. Convertibles are made with a soft fabric top and can be operated as open or closed cars. When the top is down, the convertible is basically a car without a roof.

Safety features. Automobile bodies include many safety features to help protect the occupants, such as the windshield, side windows, and rear window, which are made of laminated or tempered safety glass. The side doors have steel cross bars or beams to resist intrusion into the passenger compartment during side impacts. The body design may include an integral steel safety cage to surround and help protect the occupants, as well as crumple zones front and rear to absorb kinetic energy by deforming during an impact. Energy-absorbing front and rear bumpers, with standardized installed height, prevent or minimize vehicle damage resulting from a low-speed impact or collision. *See* SAFETY GLASS.

Automotive safety features are classified as helping to provide either crash avoidance or occupant protection. Crash-avoidance (active safety) features are often considered more beneficial because in a significant percentage of incidents they help the driver avoid an impending accident or crash. Examples are antilock-braking, traction-control, and vehicle stability control systems; headlights that turn on and off automatically as needed; and electrochromic rear-view mirrors that automatically dim to reduce glare. *See* AUTOMOTIVE BRAKE.

Occupant-protection (passive safety) features include the collapsible energy-absorbing steering column, head restraints, seat belts, air bags, and breakaway inside rear-view mirror. Some front-seat shoulder belts are motorized and move into position automatically. In some vehicles, the seat belt may automatically tighten in a crash to provide additional protection. Air bags are balloon-type supplemental restraints that inflate automatically to help protect the driver and front-seat passenger in a crash (**Fig. 3**). Rapid inflation of the air bag, usually by a pyrotechnic device, prevents the occupant from being thrown forward and injured by striking the steering wheel or windshield. The seat belt and air bag together provide maximum protection against injury in a collision. Additional air bags in the sides of the seatbacks or doors are also used to provide side-impact

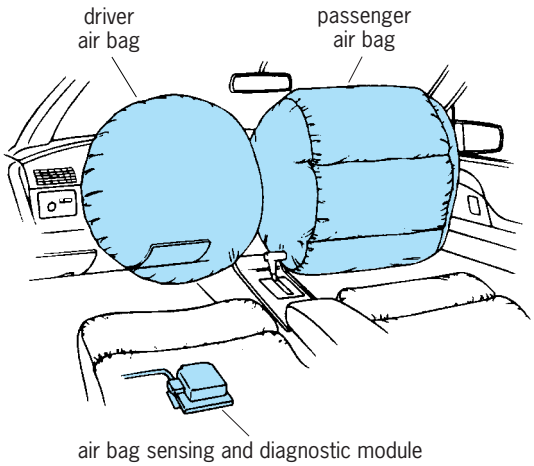

Fig. 3. Front air bags, which provide protection for the driver and front-seat passenger in a collision. (*Cadillac Division, General Motors Corp.*)

protection, primarily for the head and torso, of front- and rear-seat passengers. *See* PYROTECHNICS.

Comfort and convenience features. Comfort features include passenger-compartment ventilation air filter, heating and air conditioning, solar-control glass, electrically heated seats, and tilting and telescoping steering wheel. Convenience features include power windows, power sunroof, remote keyless entry, central locking systems, sound systems with tape players and disc players, voice-activated cellular telephone, lighted sun-visor vanity mirrors, and instrument-panel or console cup holders. *See* AUTOMOTIVE CLIMATE CONTROL; MOBILE RADIO.

Frame. This is the main structural member to which all other mechanical chassis parts and the body are assembled to make a complete vehicle. In older vehicle designs, the frame is a separate rigid structure made of heavy steel stampings welded together to resist the road loads imposed upon the vehicle. Newer passenger-car designs have the frame and body structure combined into an integral unit, or unitized body. The weight of the side rails and other structural members added in construction of the unitized body is less than the weight of the separate frame.

Short or stubby subframes and their assembled components attach to the side rails at the front and rear of the unitized body. The front subframe carries the engine, transmission or transaxle, lower front suspension, and other mechanical parts. The rear subframe, if used, carries the rear suspension and rear axle.

Suspension. The suspension supports the weight of the vehicle, absorbs road shocks, transmits brake-reaction forces, helps maintain traction between the tires and the road, and holds the wheels in alignment while allowing the driver to steer the vehicle over a wide range of speed and load conditions. The action of the suspension increases riding comfort, improves driving safety, and reduces strain on vehicle components, occupants, and cargo. The springs may be coil, leaf, torsion bar, or air. Most automotive vehicles have coil springs at the front and either coil or leaf springs at the rear. *See* AUTOMOTIVE SUSPENSION; SPRING (MACHINES).

The rear suspension includes the springs, control arms, shock absorbers or dampers, struts (if used), and other parts that act together to support the weight of the rear end and absorb road shocks. Similar parts and steering knuckles in the front suspension perform similar functions while also providing steering control.

To produce a safe, stable, and comfortable ride, a shock absorber is located at each wheel to dampen spring oscillations and wheel bounce. Some vehicles have electronic ride control, which allows the driver to electronically adjust the shock absorbers for a softer or firmer ride. An automatic setting enables the system controller to automatically change the firmness of each shock absorber to suit road conditions.

A stabilizer bar, which usually connects to the lower control arms, may be used with the front and sometimes the rear suspension. The stabilizer bar is a transverse-mounted spring-steel bar that helps reduce body roll when making turns and when driving on rough or uneven surfaces. With the bar ends attached to each side of the suspension and the center pivoting in bushings attached to the frame or underbody, the stabilizer bar resists being twisted when spring travel is unequal.

Steering. The steering system enables the driver to turn the front wheels left or right to control the direction of vehicle travel. The rotary motion of the steering wheel is changed to linear motion in the steering gear, which is located at the lower end of the steering shaft. The linear motion is transferred through the steering linkage to the steering knuckles, to which the front wheels are mounted. Steering systems are classed as either manual steering or power steering, with power assist provided hydraulically or by an electric motor.

The force that the driver exerts to turn the steering wheel is multiplied by the mechanical advantage of the steering gear, which usually is a rack-and-pinion or recirculating-ball type. Typical automotive steering ratios range from about 24:1 with manual steering to 14:1 with power steering. Most small cars have a rack-and-pinion steering gear; some full-size and older cars and many trucks have a recirculating-ball steering gear. Some vehicles have four-wheel steering that is either mechanically or electronically controlled. *See* ANTIFRICTION BEARING; AUTOMOTIVE STEERING.

Brake. A brake is a device that uses a controlled force to reduce the speed of or stop a moving vehicle, or to hold the vehicle stationary. The automobile has a friction brake at each wheel. When the brake is applied, a stationary surface moves into contact with a moving surface. The resistance to relative motion or rubbing action between the two surfaces slows the moving surface, which slows and stops the vehicle.

The four-wheel hydraulically operated brake system that the driver's foot controls by pressing and

releasing the brake pedal during normal driving conditions is called the service brake. Automobiles and trucks also have a mechanical parking brake, operated by a separate foot pedal or hand lever, which usually acts only on the rear wheels.

Automotive service brakes are either drum- or disc-type, with many vehicles having front disc brakes and rear drum brakes. Some vehicles have disc brakes at all four wheels. In addition, most vehicles have power brakes, which use the power assist of a hydraulic or vacuum booster to increase the service-brake pedal force. Many vehicles have an antilock-braking system that prevents the brake at any wheel from locking, which would cause the tire to skid. *See* AUTOMOTIVE BRAKE; BRAKE.

Engine. The engine supplies the power to move the vehicle. The power is available from the engine crankshaft after a fuel, usually gasoline, is burned in the engine cylinders. Most automotive engines are located at the front of the vehicle and drive either the rear wheels or the front wheels through a drive train or power train made up of gears, shafts, and other mechanical and hydraulic components. In some vehicles, the engine is located at the rear of the vehicle, or in back of the front seat in what is known as a mid-engine location. In vehicles with front engine and rear-wheel drive, the engine is positioned longitudinally; in vehicles with front engine and front-wheel drive, the engine is usually positioned crosswise or transversely, but may be installed longitudinally.

Most automotive vehicles are powered by a spark-ignition four-stroke-cycle internal combustion engine. The inline four-cylinder engine and V-type six-cylinder engine are the most widely used, with V-8 engines also common. Other automotive engines have three, five, ten, and twelve cylinders. Some passenger cars and trucks have diesel engines. Some automotive spark-ignition and diesel engines are equipped with a supercharger or turbocharger. *See* AUTOMOTIVE ENGINE; DIESEL ENGINE; ENGINE; IGNITION SYSTEM; INTERNAL COMBUSTION ENGINE; SUPERCHARGER; TURBOCHARGER.

To improve fuel economy, and to reduce air pollution from the engine exhaust gas and from escaping fuel vapors, most automotive engines have electronic fuel injection instead of a carburetor. A computer-controlled electronic engine control system automatically manages various emissions devices and numerous functions of engine operation, including the fuel injection and spark timing (**Fig. 4**). This allows optimizing power and fuel economy while minimizing exhaust emissions. *See* CARBURETOR; CATALYTIC CONVERTER; CONTROL SYSTEMS; FUEL INJECTION.

Power train. The power available from the engine crankshaft to do work is transmitted to the drive

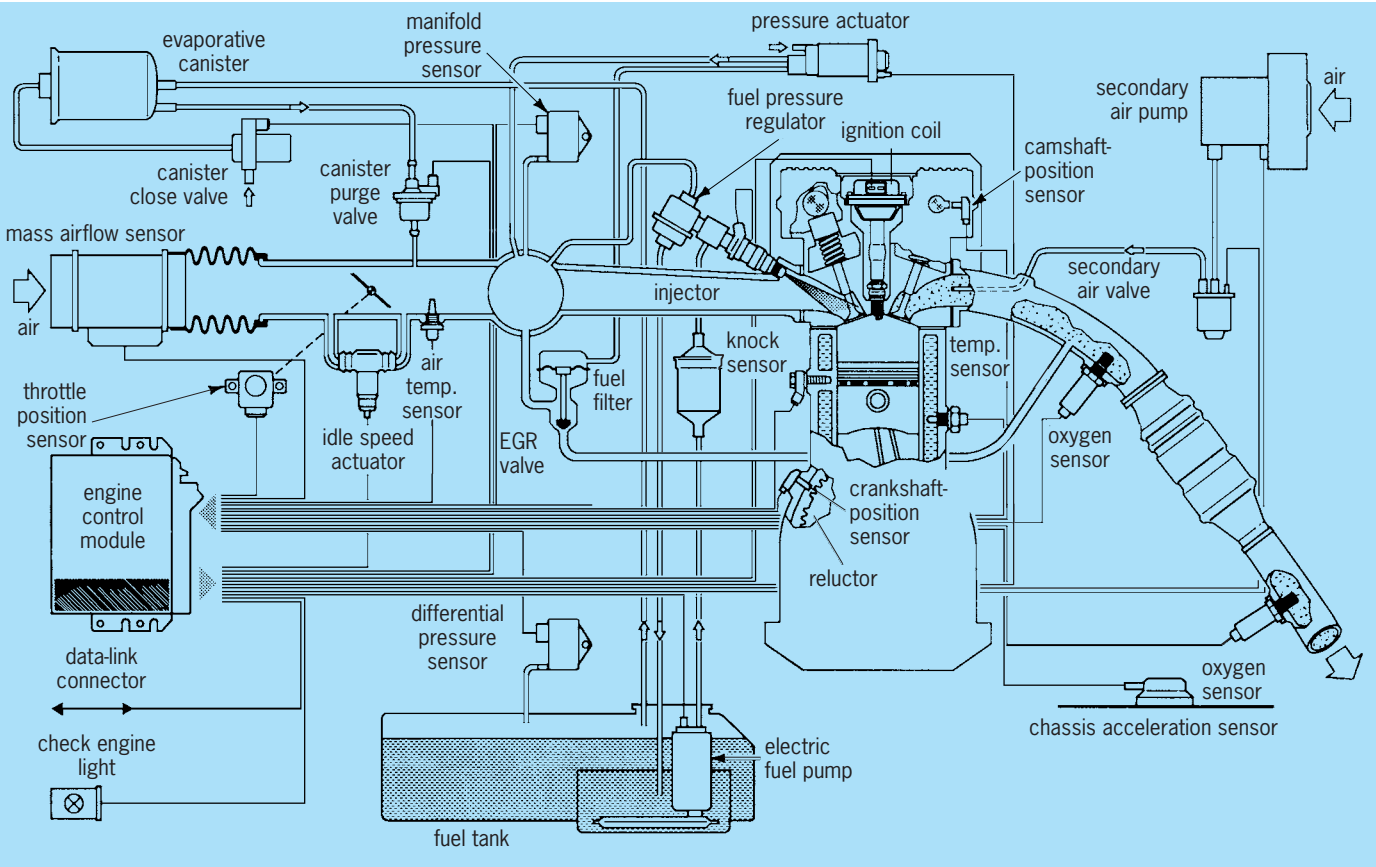

Fig. 4. Electronic engine control system that includes onboard diagnostics version II (OBD II), which can identify failure of certain emission-control components and detect deterioration in their performance throughout the life of the vehicle. (*Robert Bosch Corp.*)

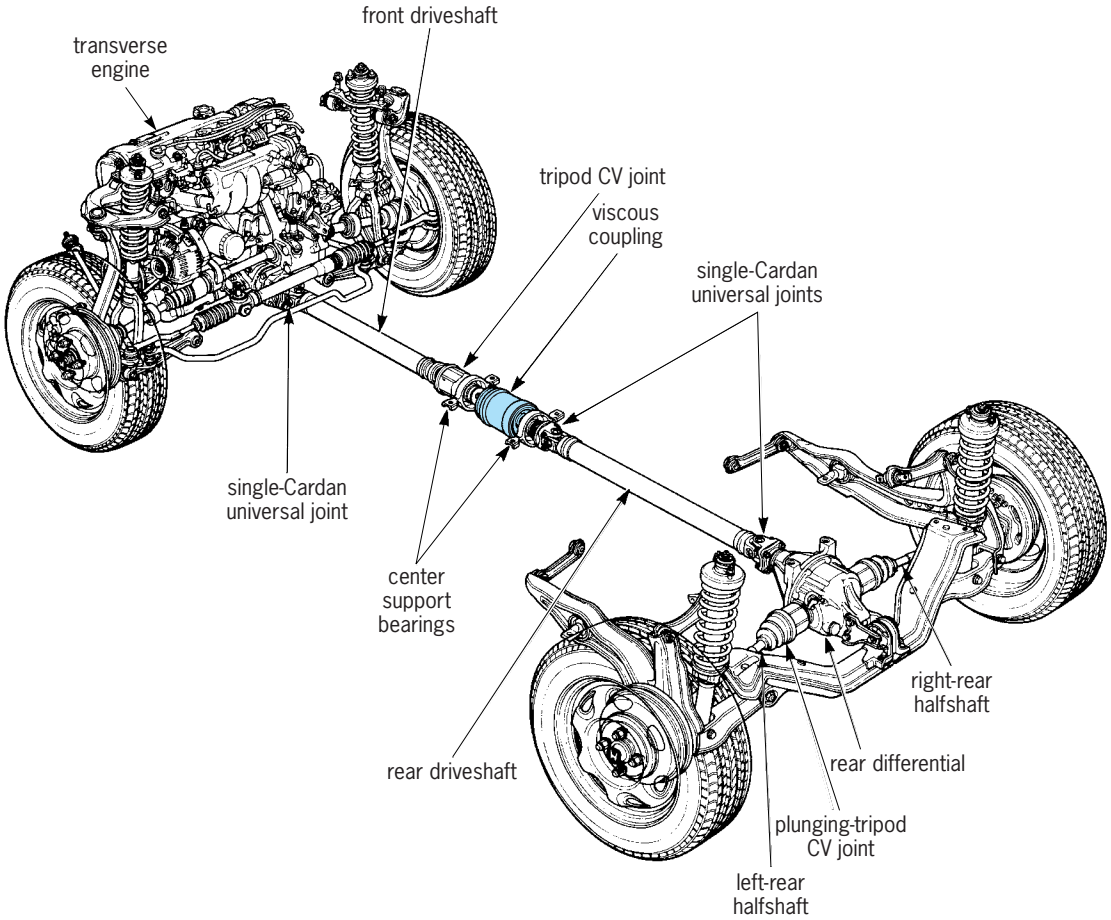

Fig. 5. Power train for an all-wheel-drive vehicle in which the front wheels are normally driven. When the front wheels spin, the viscous coupling locks and sends power to the rear axle. (*American Honda Motor Company, Inc.*)

wheels by the power train, or drive train (**Fig. 5**). In the front-engine rear-drive vehicle, the power train consists of a clutch and manual transmission, or a torque converter and an automatic transmission; driveshafts and Hooke (Cardan) universal joints; and rear drive axle that includes the final drive, differential, and wheel axle shafts. In the typical front-engine front-drive vehicle, the power train consists of a clutch and manual transaxle, or a torque converter and an automatic transaxle. The final drive and differential are designed into the transaxle, and drive the wheels through half-shafts with constant-velocity (CV) universal joints. *See* CLUTCH; GEAR; GEAR TRAIN; UNIVERSAL JOINT.

Most vehicles with four-wheel drive have a two-speed transfer case. This is an auxiliary transmission mounted behind the main transmission, and used to divide engine power and transfer it to both front and rear drive axles, either full or part time. The transfer case in most light-duty trucks has a high and a low range, doubling the number of forward (and reverse) speeds. Some vehicles have part-time four-wheel drive, which operates only when engaged by the driver. Other vehicles have full-time four-wheel drive, which engages and disengages automatically.

Passenger cars and vehicles such as station wagons and minivans may have all-wheel drive (Fig. 5), which is a type of four-wheel drive used to improve traction in on-road vehicles. If a transfer case is utilized, it has only high range. Many all-wheel-drive vehicles are front-drive with power to the rear wheels only when needed. Power transfer is controlled by a viscous coupling, which is a type of self-actuated fluid coupling that can be employed in the transfer case or drive line. If the primary drive wheels have normal traction, no power flows to the other axle. However, if the primary drive wheels lose traction and begin to spin, the viscous coupling locks, transferring power to the secondary drive wheels until the spinning wheels regain traction.

Transmission. The transmission is the device in the power train that provides different forward gear ratios between the engine and drive wheels, as well as neutral and reverse.

The two general classifications of transmission are manual transmission, which the driver shifts by hand, and automatic transmission, which shifts automatically. The transmission may be a separate unit, or it may be combined with the drive axle to form a transaxle which is used in front-drive vehicles. Most manual transmissions provide four, five, or six forward speeds. Automatic transmissions typically have four or five forward speeds.

To shift a manual transmission, the clutch must

first be disengaged. However, some vehicles have automatic clutch disengagement for manual transmissions, while other vehicles have a limited manual-shift capability for automatic transmissions. A continuously variable transmission (CVT) is used in some vehicles. It provides an infinite number of gear ratios by transmitting torque through a belt that runs between two variable-diameter pulleys. *See* AUTOMOTIVE TRANSMISSION.

Final drive. In the power train, the final drive is the speed-reduction gear set that drives the differential. The final drive is made up of a large ring gear driven by a smaller pinion, or pinion gear. This provides a gear reduction of about 3:1; the exact value can be tailored to the engine, transmission, weight of the vehicle, and performance or fuel economy desired.

Differential. In drive axles, the differential is the gear assembly between axle shafts that permits one wheel to rotate at a speed different from that of the other (if necessary), while transmitting torque from the final-drive ring gear to the axle shafts. When the vehicle is cornering or making a turn, the differential allows the outside wheel to travel a greater distance than the inside wheel; otherwise, one wheel would skid, causing tire wear and partial loss of control.

The standard or open differential delivers the same torque to each wheel. But if one tire begins to slip and spin, the open differential divides the rotary speed unequally. The tire with good traction slows and stops. This action may also stop the vehicle or prevent it from moving. To limit the amount of slip and provide traction to both drive wheels, some drive axles have a limited-slip differential. It usually has a cone clutch or multiple-disc clutch on the inner end of each axle shaft. When the clutch engages, it locks the inner end of the axle shaft to the differential case so both axles turn at the same speed. A viscous coupling in the drive axle between the axle shafts also can serve as a limited-slip differential by locking both axle shafts together if one wheel slips. Some vehicles have differential locks that can be engaged and disengaged by the driver.

A differential is used in the drive axle of a rear-drive vehicle and in the transaxle of a front-drive vehicle. A vehicle with four-wheel drive has a differential in both front- and rear-drive axles. In addition, a third center or interaxle differential which may be in the transfer case is used to compensate for any difference in front and rear wheel travel in vehicles with four-wheel drive and all-wheel drive. *See* DIFFERENTIAL.

Wheels and tires. A wheel is a disc or a series of spokes with a hub at the center and a rim around the outside for mounting of the tire. The wheels of a vehicle must have sufficient strength and resiliency to carry the weight of the vehicle, transfer driving and braking torque to the tires, and withstand side thrusts over a wide range of speed and road conditions. Wheel size is primarily determined by the load-bearing strength of the tire.

The high speeds encountered in highway driving demand low centers of gravity and have resulted in the adoption of relatively small wheels and tires of matching size. After the tires are mounted on the wheels, or after the wheels are installed on the vehicle, the wheel-and-tire assemblies are spin-balanced. Even a slightly unbalanced wheel-and-tire assembly can cause steering trouble, vibration, and rapid tire wear at high speeds.

Passenger-car wheels are usually the steel disc type. The wheel disc may be a solid plate or a slotted steel disc. Either type can be welded or riveted to the rim. The disc is dished in and usually offset to bring the center point of ground contact under the larger inner wheel bearing. Many vehicles have aluminum wheels, which are lighter than steel and reduce unsprung weight. Molded composite wheels are installed on some vehicles. Composite wheels are made of fiberglass, sheet-molding compound (SMC), and special resins, and are lighter than aluminum wheels. Steel wheels are usually trimmed with wheel covers made of steel or aluminum and attached to the wheel by clips or other fasteners.

Most passenger-car rims have a drop center which improves tubeless tire sealing and bead retention while permitting removal and installation of the tire. The wheels mount to the vehicle with three to five wheel lug bolts or lug nuts. The nose of the lug bolts or nuts usually has either a conical or spherical shape to clamp the wheel in the radial and axial directions. By removing the lug bolts or nuts, the wheel is demountable from the brake drum, axle-shaft flange, or hub which rotates on the axle or steering-knuckle spindle. This serves as the mounting for the wheel to the vehicle.

The tire is the casing-and-tread assembly (with or without a tube) that is mounted on a wheel, filled with pressurized air, and installed on the vehicle to provide pneumatically cushioned contact and traction with the road surface. New automotive vehicles are equipped with tubeless radial tires. In radial tire construction, the plies run parallel to each other and vertical to the tire bead. Tire pressure may be monitored electronically, regardless of vehicle speed, with low pressure in any tire causing a warning light on the instrument panel to illuminate. *See* TIRE.

Some tires are self-sealing to minimize air loss if the tread area is punctured by a nail or other object that has a diameter of 3/16 in. (5 mm) or less. Other tires have stiffened sidewalls that provide limited run-flat capability for up to 200 mi (320 km) after loss of air. Run-flat tires reduce vehicle weight by allowing the jack and spare tire to be eliminated from the vehicle, improving fuel economy.

Automotive electronics. The use of solid-state electronic devices in the automobile began during the 1960s, when the electromechanical voltage regulator of the alternator, with vibrating contact points, was replaced by a transistorized voltage regulator. This was followed in the 1970s by electronic ignition, fuel injection, and cruise control. Since then, electronic devices and systems on the automobile have proliferated. These include engine and power train control, air bags, antilock braking, traction control, suspension and ride control, remote keyless entry, memory seats, driver information and navigation

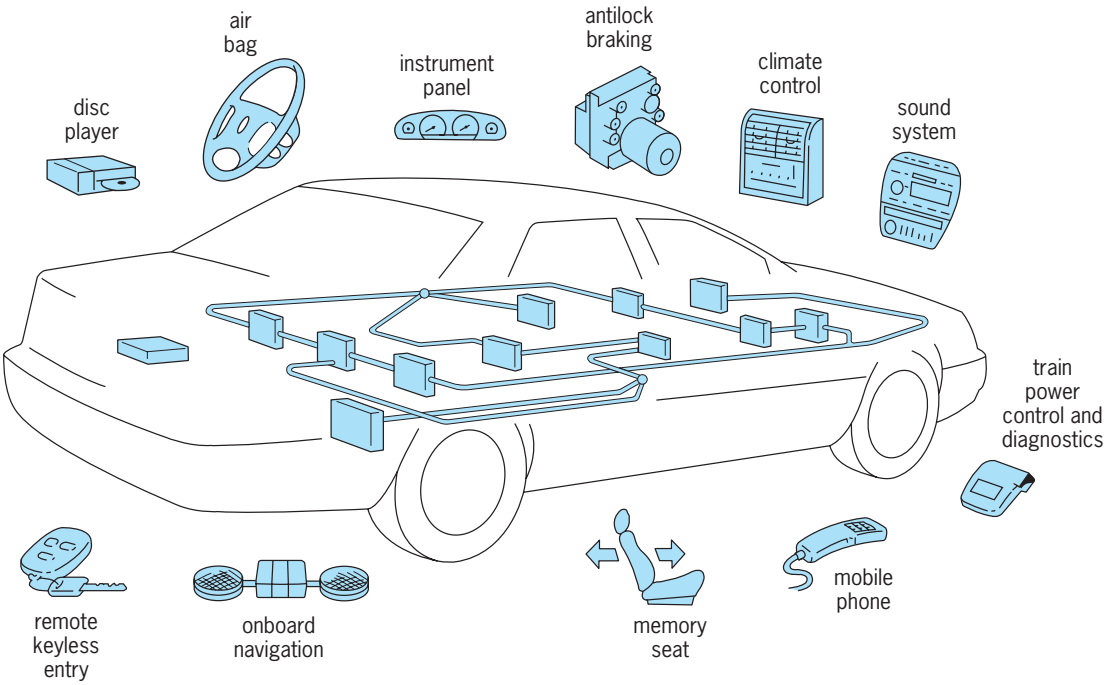

Fig. 6. Automotive multiplex system that allows data transfer and information sharing by all control modules attached to the wiring loom. (*Cadillac Division, General Motors Corp.*)

systems, cellular telephone and mobile communications systems, and onboard diagnostics. *See* ELECTRONIC DISPLAY; FEEDBACK CIRCUIT; SATELLITE NAVIGATION SYSTEMS.

The self-diagnostic capability of the vehicle computer, power-train or engine control module (Fig. 4), or system controller may be aided by a memory that stores information about malfunctions that have occurred and perhaps temporarily disappeared. When recalled from the memory, this information can help the service technician diagnose and repair the vehicle more quickly, accurately, and reliably.

As factory-installed electronics in the automobile has increased, so has the number of controllers or microprocessors, each of which may control only a separate stand-alone device or system. This creates duplication of wiring, sensors, and other system elements. To eliminate the redundancy, multiplex systems are being used which allow data transfer over a signal bus and sharing of this information with all other control modules that require it (**Fig. 6**). The multiplexed vehicle has fewer sensors, and the total length and weight of wiring, average wire gauge, and number of terminals and splices are greatly reduced while reliability is increased. *See* MULTIPLEXING AND MULTIPLE ACCESS.

Future progress in automotive electronics requires growth of multiplexing and continuing emphasis on systems integration, which combines separately developed devices so they work together as a complete system. Systems integration allows the signals from a wheel-speed sensor to be used in antilock braking, traction control, vehicle stability control, tire-pressure monitoring, and other systems that need wheel speed information. Drive-by-wire (electronic throttle control), brake-by-wire, steer-by-wire, obstacle detection, collision avoidance, and other automotive innovations will become less expensive and more functional as advances are made in the integration of mechanical, electronic, and data-processing systems.

Donald L. Anglin

Bibliography. K. Newton, W. Steeds, and T. K. Garret, *The Motor Vehicle*, 12th ed., Society of Automotive Engineers, 1996; Robert Bosch GmbH, *Automotive Handbook*, 4th ed., 1996; *SAE Handbook*, Society of Automotive Engineers, 3 vols., annually.

Automotive brake

An energy conversion device used to slow a vehicle, stop it, or hold it in position. The two systems are the service brake and the parking brake (**Fig. 1**), both of friction type. The service brake includes a hydraulically operated brake mechanism at each wheel. These wheel brakes are controlled by movement of the brake pedal, providing braking proportional to the applied pedal force. The parking brake is a mechanical brake operated through a separate hand lever or pedal; it applies parking-brake mechanisms usually at the two rear wheels. Most automotive vehicles have power-assisted braking, where a hydraulic or vacuum booster increases the force applied by the driver to the service-brake pedal. *See* BRAKE.

The two types of wheel-brake mechanisms are drum brakes and disc brakes (**Fig. 2**). Drum brakes are used at all four wheels on older vehicles, and at the rear wheels of many vehicles with front disc brakes. Some vehicles have disc brakes at all four wheels.

Hydraulic system. The four wheel brakes are hydraulically interconnected so they operate together

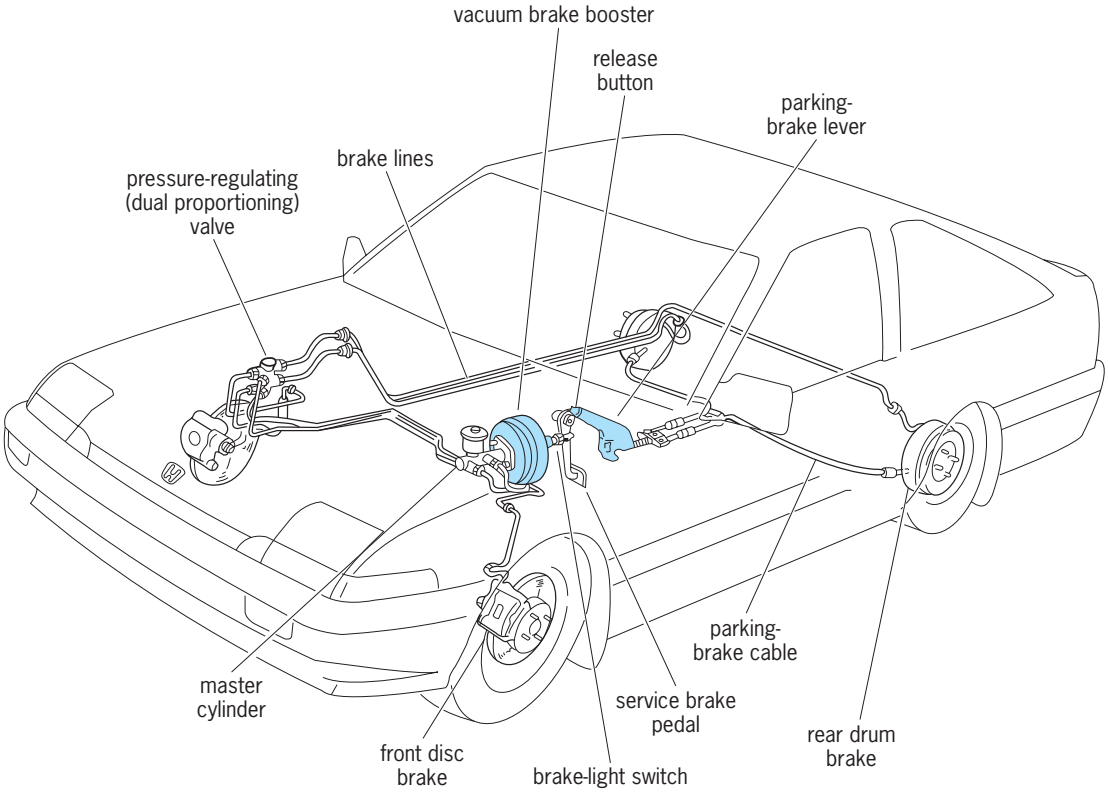

Fig. 1. Automotive brake system, showing a diagonally split hydraulic system with front disc brakes, rear drum brakes, vacuum booster, and hand-operated parking brake. (*American Honda Motor Co., Inc.*)

from one control. When the driver depresses the brake pedal, pistons are forced into fluid chambers in the master cylinder. The resulting hydraulic pressure is transmitted through steel pipe and rubber hose to hydraulic cylinders in the wheel brakes. The pressure forces pistons in the cylinders to move outward, pushing brake friction material, or lining, into contact with the rotating drum or disc to apply the brakes. *See* HYDRAULICS.

For safety, the master cylinder and hydraulic system are divided into two separate sections. Most rear-drive vehicles have a front-rear split, with one section of the hydraulic system operating the front brakes and the other operating the rear brakes. Most front-drive vehicles have a diagonal or X-type split, with each section operating one front brake and the diagonally opposite rear brake. Either arrangement allows some braking by one section if the other section fails because of damage or fluid loss.

In vehicles with front-disc and rear-drum brakes, a pressure-regulating valve is located in the hydraulic line to the rear brakes. During hard braking, this proportioning valve reduces pressure to the rear brakes, where less braking is needed, to avoid rear-wheel lockup and tire skid. A diagonally split hydraulic system may have two proportioning valves, one in each rear-brake line.

Drum brake. In a drum brake, two nonrotating curved steel shoes, faced with heat- and wear-resistant lining, are forced against the inner surface of a rotating brake drum as the driver depresses the brake pedal. When the pedal is released, return springs pull the shoes away from the drum.

Disc brake. In a disc brake, a nonrotating caliper containing one or more pistons and carrying two brake pads, or lined flat shoes, straddles the rotating disc. As the driver depresses the brake pedal, the piston and hydraulic reaction push the brake pads against each side of the disc. When the brake pedal is released, the piston seal, which was deflected as the piston moved out, provides piston retraction. Two types of caliper are the fixed or nonmoving, and the floating or sliding. The floating or sliding type depends on slight inward movement of the caliper, resulting from hydraulic reaction, to force the outer brake pad against the disc.

Power brake. Power-assisted braking is provided by a hydraulic or vacuum booster. As the brake pedal is depressed, the booster furnishes most of the force to push a pushrod into the master cylinder. The power piston in the hydraulic booster is operated by oil pressure from the power-steering pump or from a separate pump driven by an electric motor. In the vacuum booster, a diaphragm usually is suspended in a vacuum supplied from the engine intake manifold or from a vacuum pump driven by the engine or an electric motor. Depressing the brake pedal allows atmospheric pressure to act against one side of the diaphragm. The resulting pressure differential moves the diaphragm and power piston, which forces the pushrod into the master cylinder.

Antilock braking system. Braking is most efficient if lockup is prevented and the brakes are allowed to continue slowing the wheels. In a vehicle with a

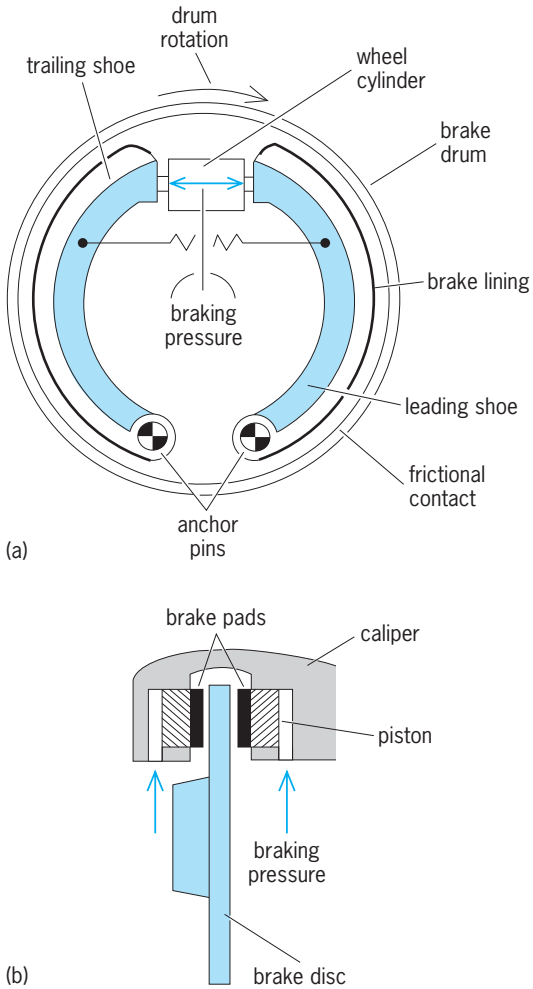

Fig. 2. Friction brakes of (*a*) drum type and (*b*) disc type used in automotive vehicles. (*Robert Bosch Corp.*)

typical antilock braking system (ABS), wheel-speed sensors signal the controller when a wheel is about to lock up. The controller then signals solenoid valves which open and close rapidly to prevent any increase, or begin to decrease, line pressure for that brake. This avoids lockup and tire slip. During normal braking, ABS has no effect on the service brakes.

Several types of ABS are in use. Some light-duty trucks, mostly those with primary rear-wheel drive and a front-rear hydraulic split, have a type of ABS that can prevent lockup of only the two rear wheels. The signal from the vehicle-speed sensor is used by the ABS controller to read the rate of rear-wheel deceleration. If pressure modulation is then necessary, both rear brakes will be affected equally because of the single hydraulic line to the rear axle.

Most vehicles have a wheel-speed sensor for each wheel. These ABS systems may be either three-channel or four-channel. In a three-channel system, the front brakes are controlled individually while both rear brakes are controlled together. In a four-channel system, each wheel brake can be individually controlled.

Traction-control system. When a tire receives more torque than it can transfer to the road, the tire loses traction and spins. To prevent this, a traction-control system (TCS) can be incorporated with ABS. When a wheel is about to spin, the controller applies the service brake at that wheel. This slows the wheel until the chance of wheel spin has passed. On many vehicles, TCS operation can reduce engine speed and torque if braking alone will not prevent wheel spin. Then, the engine controller momentarily retards ignition timing and reduces the amount of fuel delivered to the engine cylinders.

Stability-control systems. ABS and TCS provide braking and acceleration control when the vehicle is moving forward, or longitudinally. Some vehicles have an electronic stability-control system that can actively help control the vehicle during sideways or lateral movements such as cornering and skidding. These systems can automatically reduce engine torque while applying individual wheel brakes to stabilize the vehicle and maintain the driver's desired line of travel. *See* AUTOMOTIVE SUSPENSION.

Donald L. Anglin

Bibliography. Robert Bosch GmbH, *Automotive Handbook*, 4th ed., 1996.

Automotive climate control

A system for providing a comfortable environment within the passenger compartment of a vehicle. Controlled ventilation is utilized, along with a heater, an air conditioner, or an integrated heater and air-conditioner system. Linked to the setup is a windshield defrosting and defogging system capable of clearing the windshield as specified in a timed test, as required for all passenger cars manufactured for sale in the United States. Some vehicles have a ventilation-air filter which cleans the outside air that enters the passenger compartment through the fresh-air inlet. The increasing glass area of many passenger vehicles places an additional load on the air conditioner. Many vehicles incorporate solar-control glass to reduce solar transmission to the interior.

Heating. There are two types of passenger-compartment heaters: engine-dependent and engine-independent. The engine-dependent heater utilizes waste heat from the engine. The engine-independent heater includes a small combustion chamber in which fuel is burned.

Most vehicles have a liquid-cooled engine and an engine-dependent heater through which hot engine coolant flows. The coolant passes through the tubes of a tube-and-fin heater core (see **illus.**) while air flows between the fins. Heat output into the passenger compartment is regulated by controlling either the coolant flow or the airflow. An electric blower motor may run at various speeds to help move the air. When the heater is turned off, a coolant flow-control valve may close to stop the flow of coolant through the heater core. *See* DEWAR FLASK; ENGINE COOLING.

Air conditioning. When the outside temperature is above 68°F (20°C), the passenger compartment may be uncomfortable for the occupants unless

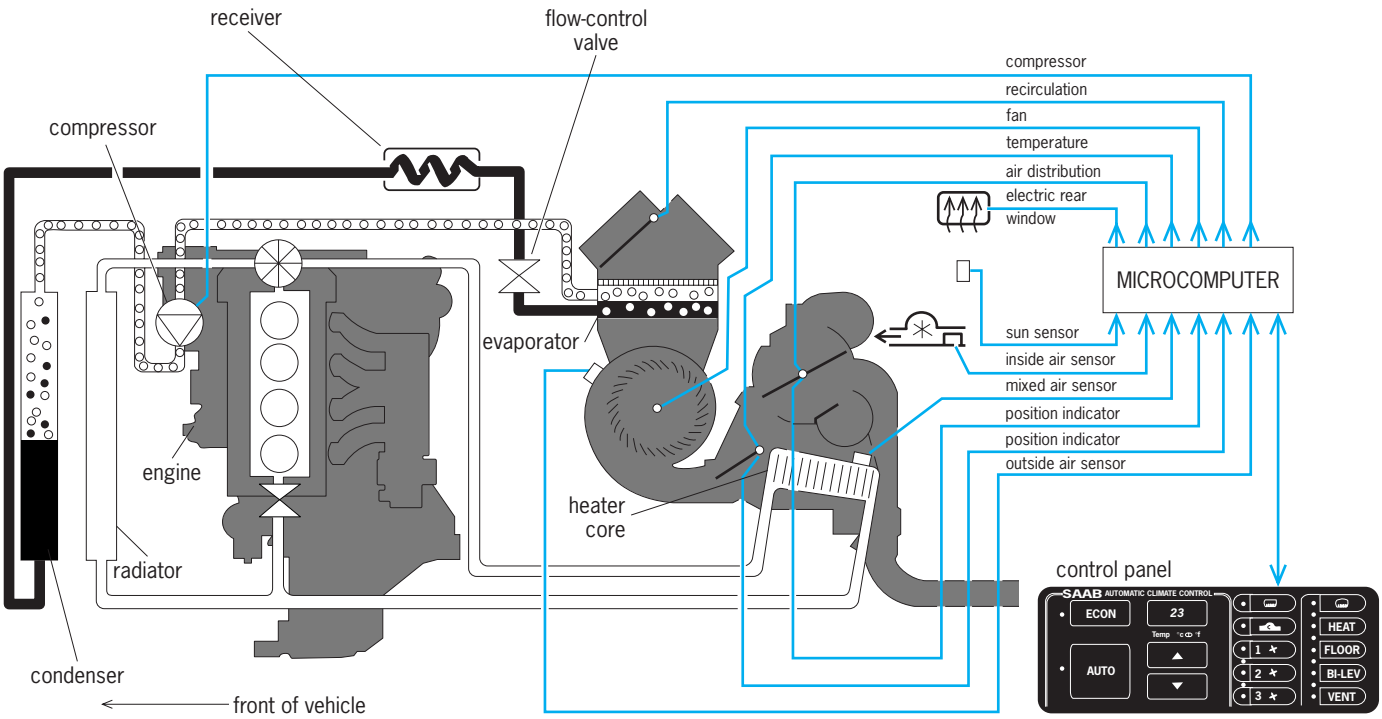

Schematic of an automotive air conditioner. (*Saab Cars USA, Inc.*)

the inside air is cooled. The cooling is provided by a mobile, vehicle-mounted refrigeration system known as an automotive air conditioner (see illus.).

The automotive air conditioner combines the refrigeration system with an air-distribution system and a temperature-control system to cool, clean, dry, and circulate passenger-compartment air. Cooling is provided by a mechanical vapor-compression refrigeration system with five major components: compressor, condenser, refrigerant flow-control valve, evaporator, and a receiver or accumulator that includes a desiccant. These parts are connected by hose and tubing, through which the refrigerant flows. Most systems use a fixed-orifice tube as the refrigerant flow-control valve. Other systems use a variable-orifice expansion valve to control refrigerant flow. The condenser and the evaporator are both tube-and-fin heat exchangers. The condenser usually is mounted ahead of the engine radiator at the front of the vehicle. The evaporator is located in the engine compartment. *See* COMPRESSOR; DESICCANT; EVAPORATOR; FLUOROCARBON; HEAT EXCHANGER; REFRIGERATION.

The compressor usually is driven by a belt from the engine crankshaft. When the engine is running and the compressor operating, heat is removed from the passenger-compartment air and transferred to the outside air. This process begins when a small amount of low-pressure liquid refrigerant flows into the evaporator. At the same time, warm passenger-compartment air flows between the evaporator fins. The liquid refrigerant absorbs heat from the warm air and vaporizes, cooling the evaporator and the air passing through it. The heat-carrying low-pressure vapor then flows from the evaporator into the accumulator (in an orifice-tube system), which traps any liquid refrigerant to avoid its damaging the compressor.

From the accumulator, the low-pressure vapor is drawn into the compressor, which compresses the vapor, raising its pressure and temperature. The hot, high-pressure vapor then flows into the condenser, where the hotter vapor loses heat to the cooler outside air flowing across the condenser. As the vapor loses heat, it condenses into a high-pressure liquid that leaves the condenser and flows to the refrigerant flow-control valve. The small opening in the valve allows passage of only a small amount of liquid refrigerant, which now has a lower pressure as it enters the evaporator. Passing through the evaporator, the low-pressure liquid refrigerant again absorbs heat and vaporizes. This cooling cycle repeats as long as the compressor runs.

Operation, air temperature, and air distribution through the passenger compartment may be controlled either automatically or manually by the driver. In some vehicles, conditioned air distribution can be controlled for each seat or seating position. *See* AIR CONDITIONING.

Refrigerant. For years, the refrigerant in automotive air conditioners was dichlorodifluoromethane (CCl_2F_2), a chlorofluorocarbon (CFC). Because of damage to the Earth's ozone layer caused by release of this compound into the atmosphere, R-12 has been replaced by less harmful compounds. The factory-installed refrigerant in most automotive air conditioners is the hydrofluorocarbon 1,1,1,2-tetrafluorethane (CF_3CH_2F), known as R-134a. *See* ATMOSPHERIC CHEMISTRY.

Clean Air Act. In the United States, the 1990 amendments to the Clean Air Act require that any work on an automotive air conditioner involving the refrigerant must be performed by a trained and certified technician using approved refrigerant recycling equipment. The refrigerant must not be discharged or vented into the atmosphere. Recycling the recovered refrigerant cleans it for reuse.

After recycling, contaminants in refrigerant R-12 should not exceed, by weight, 15 parts per million of moisture, 4000 ppm of refrigerant oil, and 330 ppm of noncondensable gases (air). Limits, by weight, for refrigerant R-134a are 15 ppm of moisture, 500 ppm of refrigerant oil, and 150 ppm of noncondensable gases. Donald L. Anglin

Bibliography. P. E. Anglin (ed.), *Automotive Climate Control Systems*, Society of Automotive Engineers, 1991; *SAE Handbook*, Society of Automotive Engineers, 3 vols., annually.

Automotive drive axle

A theoretical or actual crossbar or assembly that supports a motor vehicle and on which one or more wheels turn. The axle is either a live axle or a dead axle. A live axle, or drive axle, drives the wheels connected to it while supporting part of the weight of the vehicle. A dead axle, or nondrive axle, carries part of the weight of the vehicle but does not drive the wheels. *See* AUTOMOBILE.

A drive axle on which the wheels can pivot for steering, such as on the front axle of a four-wheel-drive vehicle, is a steerable drive axle (see **illus.**). The rear axle in most automotive vehicles is a nonsteering drive axle.

Automotive configurations include front-wheel drive, rear-wheel drive, and four-wheel or all-wheel drive. All of these include a transmission or transaxle, differential, and driveline or axle shafts. In vehicles with a manual transmission or transaxle, the power flows from the engine crankshaft through a clutch. *See* AUTOMOTIVE TRANSMISSION; CLUTCH.

The wheels that rotate to move the vehicle are on the outer ends of a drive axle. A vehicle with front-wheel drive has a front drive axle, while the rear-wheel-drive vehicle has a rear drive axle. In a vehicle with four-wheel drive or all-wheel drive, power can be sent to all four wheels. The difference is that four-wheel drive usually includes a two-speed transfer case that makes the vehicle suitable for off-road use. All-wheel drive does not have a two-speed transfer case, making the vehicle primarily for on-road use. *See* TRUCK.

Rear drive axle. The rear drive axle is suspended from the vehicle body or frame by springs attached to the axle housing. The housing encloses the final-drive gears, differential, and wheel axle shafts. *See* AUTOMOTIVE SUSPENSION.

The functions performed by the rear drive axle include changing the rotation of the driveshaft by 90° to rotate the wheel axle shafts; providing a final speed reduction between the driveshaft and the axle shafts through the final-drive gears; permitting one wheel to turn at a speed different from the other when necessary (differential action); driving the rear wheels through axle shafts, or halfshafts; and acting as thrust- and torque-reaction members during acceleration and braking. The differential may be either the open type or, for greater traction, the limited-slip type. Differential action between wheels, and between front and rear drive axles (interaxle

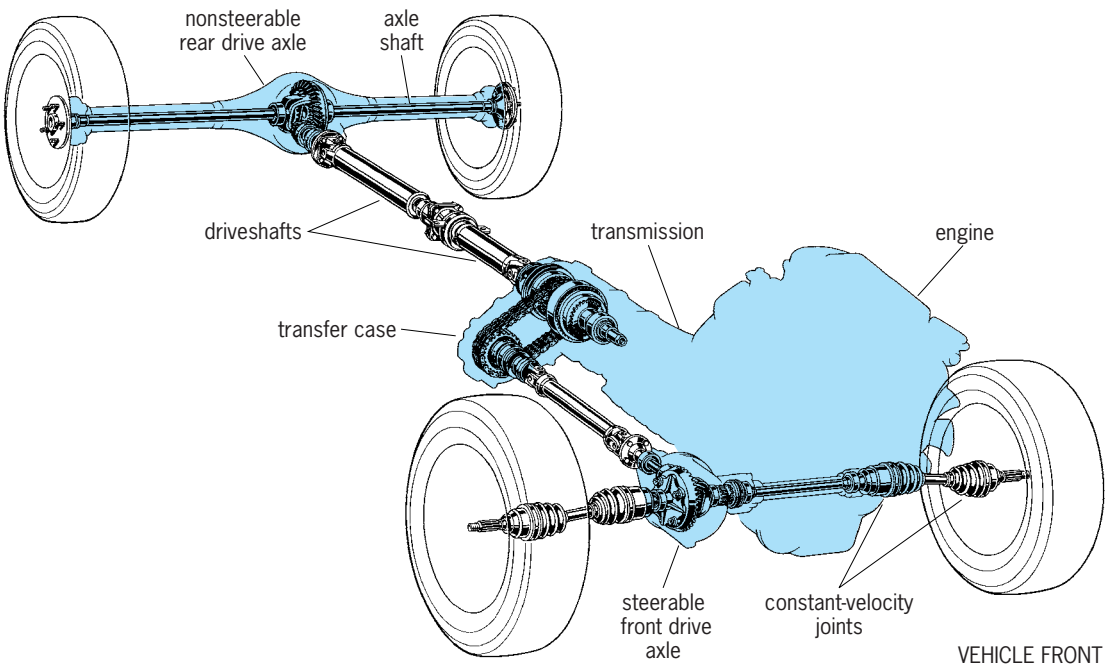

Drive train of a four-wheel-drive vehicle, showing the steerable front drive axle and nonsteerable rear drive axle. (*Mazda Motors of America Inc.*)

differential), may also be provided by a viscous coupling. *See* DIFFERENTIAL.

Instead of an axle housing, some vehicles have independent rear suspension. The differential carrier is attached to a body or frame crossmember, and the axle halfshafts are exposed. A constant-velocity universal joint (CV joint) is usually located at the inner and outer ends of each halfshaft. The nonsteerable rear wheels are held in position by rear-suspension control arms. *See* UNIVERSAL JOINT.

Front drive axle. Most four-wheel-drive vehicles have a steerable front drive axle. It is usually similar in construction and function to the rear drive axle, having a housing and enclosed axle shafts. The principal difference is in the provisions made for steering. Because the front wheels must turn in or out on the steering-knuckle pivots, universal joints are located at the outer ends of the axle shafts. This allows the wheels to be driven while moving in and out or up and down. *See* AUTOMOTIVE STEERING.

Automotive vehicles with front-engine and front-wheel drive have independent front suspension. A housing-type front drive axle is not used. Instead, the functions of the transmission and drive axle are combined into the transaxle. The axle halfshafts are exposed, extending from the sides of the transaxle to drive the front wheels. Constant-velocity universal joints are located at the inner and outer ends of each halfshaft.

Steerable rear drive axle. Some automotive vehicles have steerable drive axles on the front and rear which provide both four-wheel drive and four-wheel steering. In these vehicles, the rear wheels are steered in the opposite direction to that of the front wheels at low speed and in the same direction at higher speed. This increases vehicle maneuverability.

Donald L. Anglin

Bibliography. W. H. Crouse and D. L. Anglin, *Automotive Mechanics,* 1993.

Automotive electrical system

The system in an automotive vehicle that furnishes the electrical energy to crank the engine for starting, recharge the battery after cranking, create the high-voltage sparks to fire the compressed air-fuel charges, and power the various lights, motors, controllers and control systems, and other signaling and accessory systems.

The vehicle electrical system includes the battery, wiring, starting motor and controls, generator and voltage regulator, electronic ignition, and electronic fuel injection. Also included may be a computerized power-train control system, electronically displayed driver-information system, air bags, power seats and windows, and various types of radios and sound systems.

Most vehicles have a one-wire or ground-return system, with the metal of the vehicle serving as the common ground. To reduce wiring and duplication, some vehicles have interconnected controllers and components that form a multiplex network. By sharing wiring and information over a data bus, the vehicle has fewer sensors, and the amount of wiring and the number of connections are greatly reduced. *See* BATTERY; ELECTRIC SWITCH; ELECTRONIC DISPLAY; ELECTRONICS; FUSE (ELECTRICITY); GENERATOR; REGULATOR; SPEEDOMETER; VOLTAGE REGULATOR.

Donald L. Anglin

Bibliography. W. H. Crouse and D. L. Anglin, *Automotive Mechanics*, 1993.

Automotive engine

The component of the motor vehicle that converts the chemical energy in fuel into mechanical energy for power. The automotive engine also drives the generator and various accessories, such as the air-conditioning compressor and power-steering pump. *See* AUTOMOTIVE CLIMATE CONTROL; AUTOMOTIVE ELECTRICAL SYSTEM; AUTOMOTIVE STEERING.

Early motor vehicles were powered by a variety of engines, including steam and gasoline, as well as by electric motors. The flexibility of the gasoline engine operating on the four-stroke Otto cycle soon made this engine predominant, and it remains the dominant automotive power plant. The basic modern automotive engine (see **illus.**) is a gasoline-burning, liquid-cooled, spark-ignition, four-stroke-cycle, multicylinder engine. It has the intake and exhaust valves in the cylinder head, and electronically controlled ignition and fuel injection. *See* ENGINE.

Otto-cycle engine. An Otto-cycle engine is an internal combustion piston engine that may be designed to operate on either two strokes or four strokes of a piston that moves up and down in a cylinder. Generally, the automotive engine uses four strokes to convert chemical energy to mechanical energy through combustion of gasoline or similar hydrocarbon fuel. The heat produced is converted into mechanical work by pushing the piston down in the cylinder. A connecting rod attached to the piston transfers this energy to a rotating crankshaft. *See* GASOLINE; INTERNAL COMBUSTION ENGINE; OTTO CYCLE.

Cylinder arrangement. Engines having from 1 to 16 cylinders in in-line, flat, horizontally opposed, or V-type cylinder arrangements have appeared in production vehicles, progressing from simple single-cylinder engines at the beginning of the twentieth century to complex V-12 and V-16 engines by the early 1930s. Increased vehicle size and weight played a major role in this transition, requiring engines with additional displacement and cylinders to provide acceptable performance.

High-volume usage of the V-8 engine began in the mid-1930s and accelerated dramatically after World War II, until it was the predominant engine used in American-built vehicles by the late 1950s. Manufacturers in other countries continued large-volume production of smaller engines with four and six cylinders, primarily because of significantly higher fuel costs. As vehicle size and weight increased, average engine displacement also increased until the early 1970s, when V-8 engines approaching 500 in.^{3}

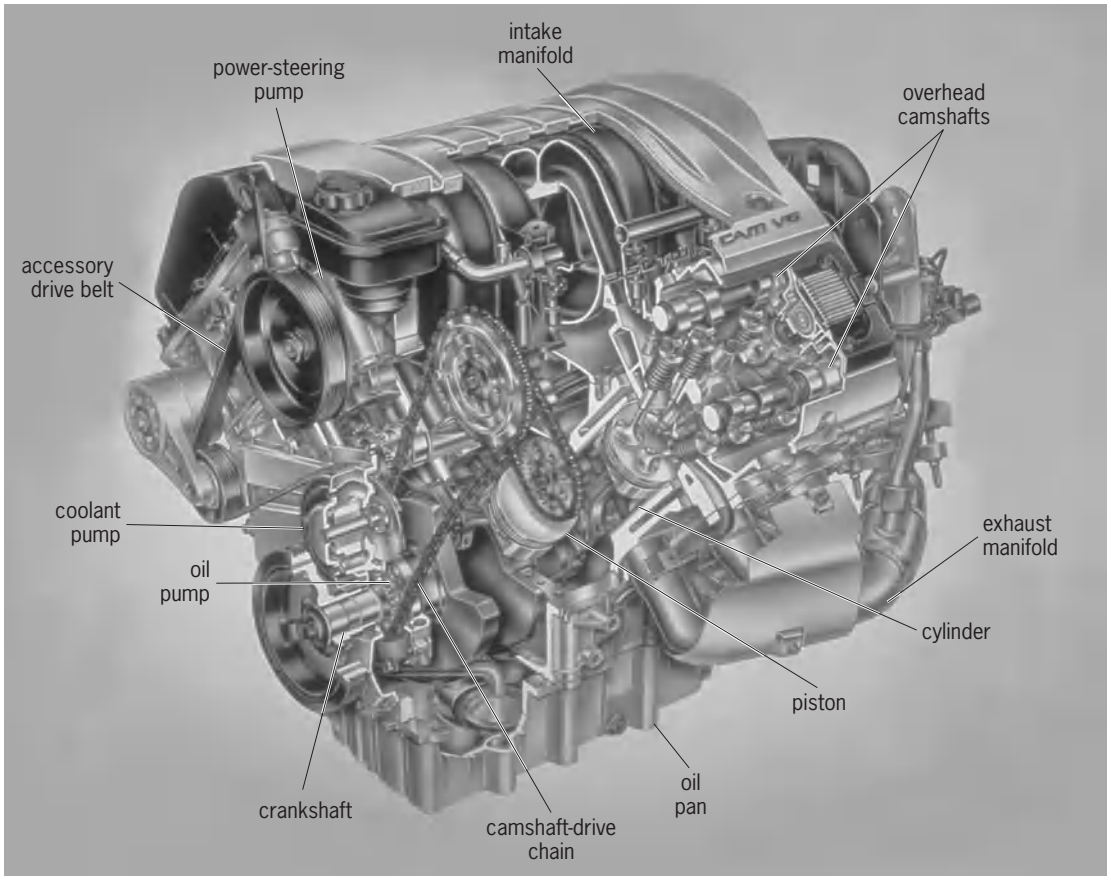

Automotive engine, which has six cylinders, double-overhead camshafts, 24-valve electronic coil-on-plug spark ignition, and multiport fuel injection. (*Oldsmobile Division, General Motors Corp.*)

(8 liters) displacement were in production. However, oil shortages in 1973–1974 and 1979–1980 reversed this trend, and V-8 engine usage dropped in favor of engines with four and six cylinders.

Turbocharger and supercharger. To provide acceptable vehicle performance with a smaller engine, forced induction may be used. A turbocharger or supercharger forces more air into the intake manifold, allowing the engine to burn more fuel and produce more power. The turbocharger is a centrifugal air compressor driven by an exhaust-gas-powered turbine mounted on a common shaft. The energy in the exhaust gas spins the turbine, which spins the compressor, forcing more air or air-fuel mixture into the combustion chambers. In a typical passenger car, this may increase engine power output by up to 40%.

A supercharger, which is belt-driven from the engine crankshaft, may be used instead of a turbocharger. The supercharger does not have the brief acceleration lag, or so-called turbo lag, that is found objectionable by many drivers of vehicles with turbocharged engines. *See* AUTOMOBILE; COMBUSTION CHAMBER; COMPRESSOR; MUFFLER; SUPERCHARGER; TURBINE; TURBOCHARGER.

Emissions. In the United States, passenger-car emission standards became effective in California in 1966 and in the other 49 states in 1968. These regulations began placing limits on crankcase, exhaust, and evaporative emissions into the atmosphere. The limits became increasingly stringent over the years, requiring the use of catalytic converters and unleaded gasoline beginning with 1975-model cars. Because more accurate fuel metering and ignition timing were required on engines to meet the tightening standards, electronic controls became necessary. As a result, fuel injection replaced the carburetor on automotive engines.

Electronic controls. Ignition, fuel, and emissions systems are integrated under an electronic engine control system. The system utilizes an onboard computer to provide management of various engine-operating parameters and emissions devices. The computer, known as the powertrain control module, may also control shifting of the automatic transmission or transaxle.

Engine design trends. In many automotive engines, the camshaft, which operates the intake and exhaust valves, has been moved from the cylinder block to the cylinder head (see illus.). This overhead-camshaft arrangement allows the use of more than two valves per cylinder, with various multivalve engines having three to five. Some overhead-camshaft engines have only one camshaft, while others have two camshafts, one for the intake valves and one for the exhaust valves. A V-type engine may have four camshafts, two for each bank of cylinders. Some multivalve overhead-camshaft engines have the power and performance of a turbocharged engine of similar size.

Most engines have fixed valve timing, regardless of number of camshafts or their location. Variable valve timing can improve fuel economy and minimize exhaust emissions, especially on multivalve engines. At higher speeds, volumetric efficiency can be increased by opening the intake valves earlier. One method drives the camshaft through an electrohydraulic mechanism that, on signal from the engine computer, rotates the intake camshaft ahead about 10°. Another system varies both valve timing and valve lift by having two cam lobes, each with a different profile, that the computer can selectively engage to operate each valve. Computer-controlled solenoids for opening and closing the valves will allow elimination of the complete valve train, including the camshaft, from the automotive piston engine while providing variable valve timing and lift.

Materials trends. Historically, major engine components have been made from ferrous metals, either by casting or by forging. However, emphasis on weight reduction for improved fuel economy has greatly increased the usage of aluminum for cylinder blocks, cylinder heads, and other engine components. Some engine covers and intake manifolds are made of magnesium. Internal engine parts, such as connecting rods, sprockets, oil-pump rotors, and valve guides, are cast or forged to nearly net shape using powder metallurgy. High-speed engines may use titanium connecting rods to reduce reciprocating mass. *See* POWDER METALLURGY.

Parts such as engine covers, intake manifolds, and oil pans also can be fabricated of plastic or composite materials. These materials provide weight savings while reducing engine noise and vibration. Ceramic engine parts and coatings will allow engine operation at higher temperatures, raising engine efficiency. Ceramic-lined exhaust ports in the cylinder head can lower its temperature while increasing the effectiveness of the catalytic converter.

Fuel-metering trends. With the introduction of electronic controls, a device was added to the carburetor to automatically adjust the air-fuel ratio in response to feedback from an exhaust-gas oxygen sensor. Demand for more accurate fuel metering resulted in the feedback carburetor being replaced by a similarly located throttle-body fuel-injection unit. It meters fuel through the computer-controlled pulsing of one or two solenoid-operated fuel injectors. Further improvements in engine power, fuel economy, and exhaust emissions are provided by multiport fuel injection, which places a fuel injector in each intake port. Solenoid-operated fuel injectors can be pulsed or energized in simultaneous, group, or sequential fashion—the last energizes each injector individually in firing-order sequence.

Ignition trends. On many automotive engines, the ignition distributor has been replaced with computer-controlled distributorless ignition; this in turn is being replaced with coil-on-plug or direct ignition, in which an ignition coil sits directly above, and is connected to, each spark plug. Some engines have two spark plugs per cylinder to provide higher power output with cleaner combustion and less tendency for spark knock, or detonation. Spark knock can be monitored by a knock sensor, which signals the computer for less spark advance to prevent engine damage. The knock sensor also is used, especially with a supercharger or turbocharger, to allow engine operation on a more economical, lower-octane-rated fuel than otherwise would be required.

Onboard diagnostic developments. An onboard computer with self-diagnostic capability has become standard equipment for automotive engine control. The first generation of onboard diagnostics (OBD I) identified the failure of certain emission-control components. The second generation (OBD II), required for 1996 and later model vehicles, has additional capability, including detection of deterioration in performance of emission-control components throughout the life of the vehicle.

Alternative engines. Alternative engine designs have been investigated as replacements for the four-stroke Otto-cycle piston engine, including the two-stroke, diesel, Stirling, Wankel rotary, gas turbine, and steam engines, as well as electric motors and hybrid power plants. However, only two engines are in mass production as automotive power plants: the four-stroke gasoline engine described above, and the diesel engine. Continuing improvements to the Otto-cycle piston engine, such as electronic controls and value actuation and other changes in design and materials, appear to assure its predominance in the short term. *See* BATTERY; DIESEL ENGINE; ELECTRIC VEHICLE; FUEL CELL; GAS TURBINE; MOTOR; POWER PLANT; ROTARY ENGINE; SOLAR CELL; STEAM ENGINE; STIRLING ENGINE.

Donald L. Anglin

Bibliography. H. Heisler, *Advanced Engine Technology*, Society of Automotive Engineers and Edward Arnold, 1995.

Automotive steering

The means by which a motor vehicle is controlled about the vertical axis. It allows the driver to control the course of vehicle travel by turning the steering wheel, which turns the input shaft in the steering gear. The steering system (**Fig. 1**) has three major components: (1) the steering wheel and attached shaft in the steering column which transmit the driver's movement to the steering gear; (2) the steering gear that increases the mechanical advantage while changing the rotary motion of the steering wheel to linear motion; and (3) the steering linkage (including the tie rod and tie-rod ends) that carries the linear motion to the steering-knuckle arms. *See* MECHANICAL ADVANTAGE.

When the only energy source for the steering system is the force that the driver applies to the steering wheel, the vehicle has manual steering. When the driver's effort is assisted by hydraulic pressure from an electric or engine-driven pump, or by an electric motor, the vehicle has power-assisted steering, commonly known as power steering.

Steering column. The steering column (Fig. 1) supports the steering wheel and encloses the steering

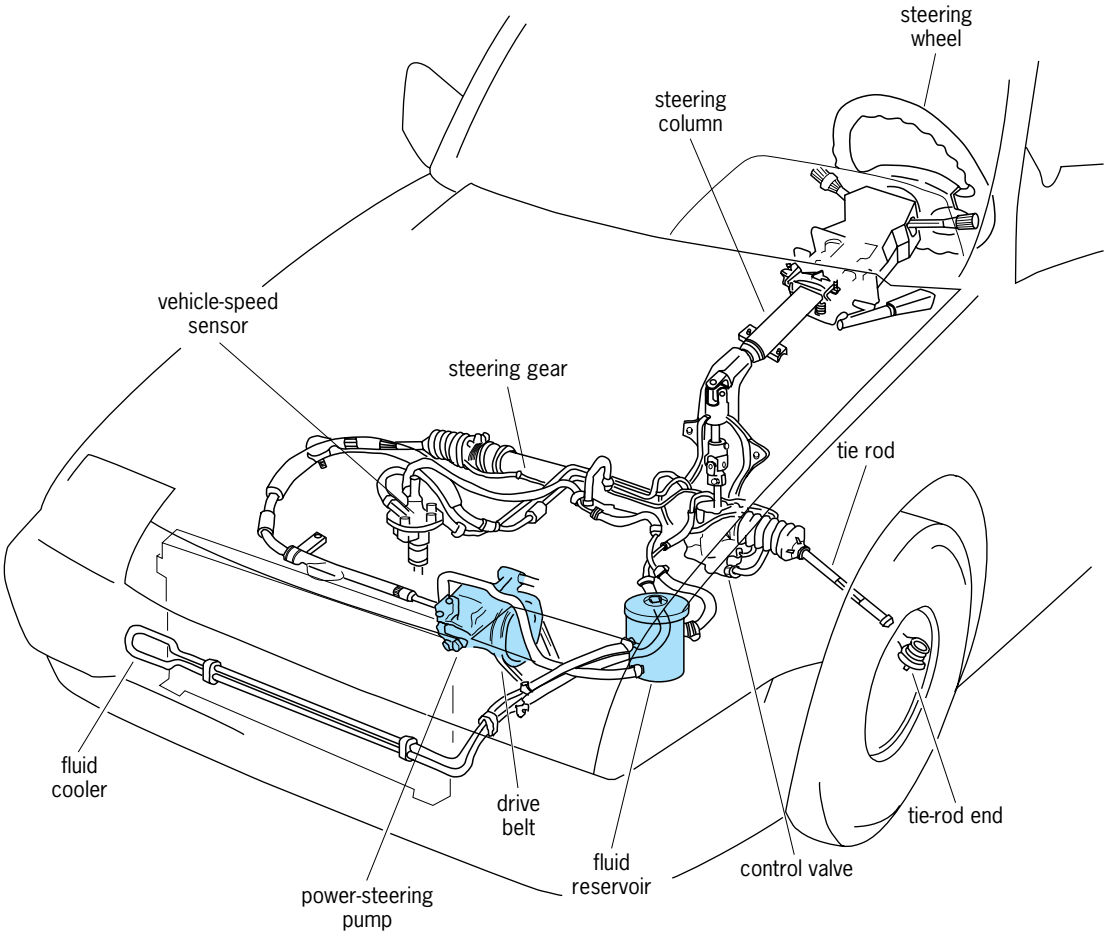

Fig. 1. Speed-sensitive rack-and-pinion power-steering system that provides variable assist. (*American Honda Motor Co., Inc.*)

shaft. Many vehicles have tilting steering wheels, some of which can also be moved in or out of the steering column. In some vehicles, the desired position of the steering wheel can be stored in memory and then automatically reset after the driver enters. The driver's side air bag and inflator are stored in the center of the steering wheel. *See* AUTOMOBILE.

Types of steering gears. Two types of automotive steering gears are rack-and-pinion and recirculating-ball. In a rack-and-pinion steering gear (**Fig. 2**), a tubular housing contains the toothed rack and a pinion gear. The housing is mounted rigidly to the vehicle body or frame to take the reaction to the steering effort. The pinion gear is attached to the lower end of the steering shaft, and meshes with rack teeth. Tie rods connect the ends of the rack to the steering-knuckle arms at the wheels. As the steering wheel turns, the pinion gear moves the rack right or left. This moves the tie rods and steering-knuckle arms, which turn the wheels in or out for steering.

In a recirculating-ball steering gear, a worm gear is attached to the lower end of the steering shaft. The worm gear turns inside a ball nut which rides on a set of recirculating ball bearings. These ball bearings roll in the grooves in the worm and inside the ball nut. Gear teeth on one outside flat of the ball nut mesh with a sector of teeth on the output or sector shaft to which the pitman arm is attached. As the steering wheel is turned, the rotary motion of the worm gear causes the ball nut to move up or down, forcing the sector shaft and pitman arm to rotate. This action moves the steering linkage to the right or left, turning the front wheels in or out for steering. *See* ANTIFRICTION BEARING; GEAR.

Many steering gears provide variable-ratio steering in which the mechanical advantage is less in the straight-ahead, on-center position and greater in any off-center position. This arrangement provides a fast steering ratio with improved response for highway driving, and a slow steering ratio with reduced steering effort for cornering and parking. Some vehicles have four-wheel steering that is controlled either mechanically or electronically.

Power steering. When the force applied to turn the steering wheel is produced by the driver and another energy source, the vehicle has power-assisted steering. This arrangement, commonly known as power steering, allows manual steering to always be available, even if the engine is not running or the power-assist system fails.

There are three types of hydraulic power steering: integral rack-and-pinion (Fig. 2), integral recirculating-ball, and nonintegral or linkage-type

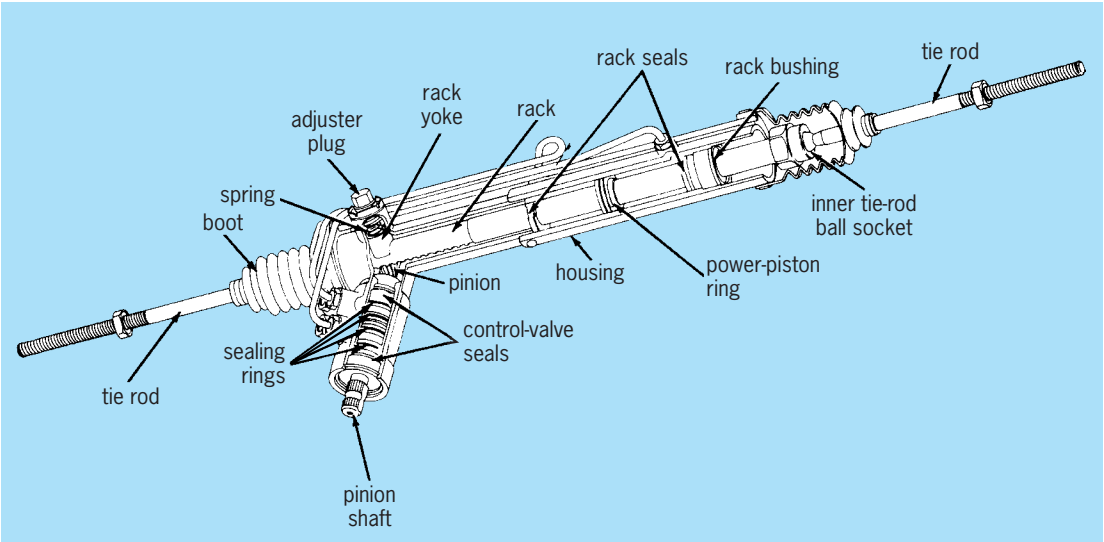

Fig. 2. Construction of a rack-and-pinion power-steering gear. (*Moog Automotive, Inc.*)

recirculating ball. Common components include the pump, fluid reservoir, control valve, power cylinder, and hoses. The fluid may be cooled by passing through a small heat exchanger or fluid cooler. Integral means the power cylinder is built into the steering gear. Nonintegral means a separate power cylinder is attached between the vehicle body or frame and the steering linkage.

Operation of hydraulic power steering is by two basic systems, one that generates and transmits the power assist and the other that controls it. Pressure is developed by the pump, which forces fluid from the reservoir through a pressure hose to the control valve. When the steering wheel is turned, the control valve senses the steering direction desired and directs oil accordingly to one side of the piston in the power cylinder. The fluid then returns under low pressure through a hose to the reservoir. Some systems are electronically controlled by a microprocessor that varies the assist based on vehicle speed and steering-wheel angle. *See* HYDRAULICS; SERVOMECHANISM.

An electronic rack-and-pinion power-steering gear is used on same vehicles. It replaces hydraulic power steering with electronic controls and an electric motor built into the rack housing. Steering-wheel movement is detected by a sensor mounted on the steering-gear input shaft. The electric motor is powered by the vehicle battery, which provides power assist even if the engine is not running.

Ackerman steering. Wheeled motor vehicles use the Ackerman system of steering, in which all wheels roll on circles with a common center. The front wheels are mounted on pivoted steering knuckles, and the steering linkage connects the steering-knuckle arms so the wheels swing together about their pivots. During a turn, the inwardly inclined arms move the inner wheel in more than the outer wheel. This toe-out on turns permits the front tires to round a turn without sideslip. *See* AUTOMOTIVE SUSPENSION.

Wheel alignment. Wheel alignment is the relationship among the wheels, steering parts, suspension angles, and the road that affect the operation and steering of a vehicle. Six basic wheel alignment factors are suspension height, caster, camber, toe, steering-axis inclination, and turning radius or toe-out on turns (**Fig. 3**). Alignment of the front wheels with the rear wheels should provide a common vehicle centerline, geometric centerline, and thrust line.

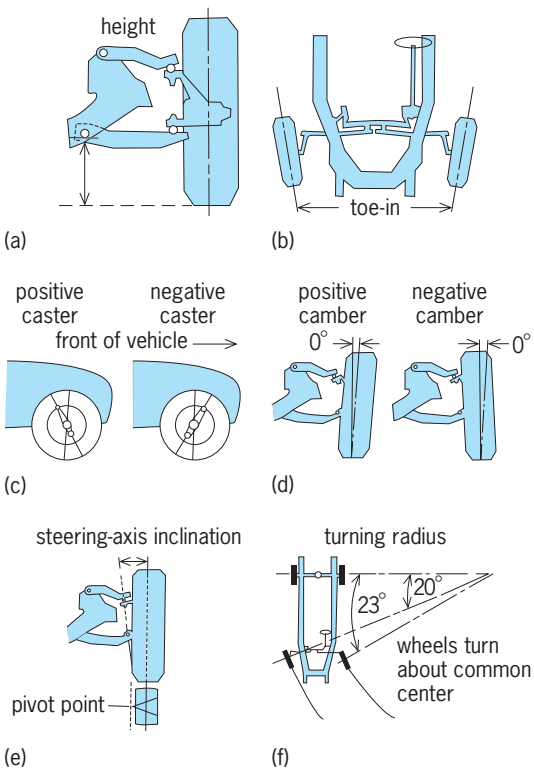

Fig. 3. Basic factors in wheel alignment. (*a*) Suspension height. (*b*) Toe. (*c*) Caster. (*d*) Camber. (*e*) Steering-axis inclination. (*f*) Turning radius. (*Chrysler Corp.*)

Suspension height is the distance measured from a specific point on the body, frame, or suspension to the ground. Caster is the angle formed by the forward or rearward tilt of the steering axis from the vertical, when viewed from the side of the wheel. The angle is positive when the steering axis tilts backward. Camber is the amount the top of a wheel tilts in (negative) or out (positive), when viewed from the front of the vehicle. Toe is the amount that the front of the wheels point inward (toe-in) or outward (toe-out). Steering-axis inclination is the angle between the vertical and a line drawn through the centers of the suspension ball joints, when viewed from the front of the vehicle. Turning radius, or toe-out on turns, is the difference between the angle that each steered wheel makes with the vehicle body or frame during turns, usually measured with the outside wheel turned 20°.

Other factors that affect steering are scrub radius or steering offset, wheel setback, and thrust angle. Scrub radius is the distance between the steering axis and the tire contact-area centerline at their intersections with the road surface. Wheel setback is the difference in vehicle wheelbase from one side to the other. Thrust angle is the angle formed between the vehicle and geometric centerlines, and the thrust line. The thrust line runs forward from the midpoint between the two rear wheels and determines the direction that the vehicle will travel if unaffected by the front wheels. If the thrust line does not coincide with the vehicle centerline, a thrust angle is formed and the thrust line then represents the course that the rear wheels will try to travel. Donald L. Anglin

Bibliography. Robert Bosch GmbH, *Automotive Handbook*, 4th ed., 1996.

Automotive suspension

The springs and related parts intermediate between the wheels and the frame, subframe, or side rails of a unitized body. The suspension supports the weight of the upper part of a vehicle on its axles and wheels, allows the vehicle to travel over irregular surfaces with a minimum of up-and-down body movement, and allows the vehicle to corner with minimum roll or loss of traction between the tires and the road. Four types of spring used in automotive suspension are coil, leaf, torsion bar, and air. *See* AUTOMOBILE; SPRING (MACHINES).

In a typical suspension system for a vehicle with front-engine and front-wheel drive (see **illus.**), the weight of the vehicle applies an initial compression to the coil springs. When the tires and wheels encounter irregularities in the road, the springs further compress or expand to absorb most of the shock. The suspension at the rear wheels is usually simpler than for the front wheels, which require multiple-point attachments so the wheels can move up and down while swinging from side to side for steering.

Shock absorber. A telescoping hydraulic damper, known as a shock absorber, is mounted separately or in the strut at each wheel to restrain spring movement and prevent prolonged spring oscillations. The shock absorber contains a piston that moves in a cylinder as the wheel moves up and down with respect to the vehicle body or frame. As the piston moves, it forces a fluid through an orifice, imposing a restraint on the spring. Spring-loaded valves open to permit quicker flow of the fluid if fluid pressure rises high enough, as it may when rapid wheel movements take place. Most automotive vehicles use gas-filled shock absorbers in which the air space above the fluid is filled with a pressurized gas such as nitrogen. The gas pressure on the fluid reduces the creation of air bubbles and foaming. *See* SHOCK ABSORBER.

Electronic ride control. With electronic ride control or variable-damping suspension, the shock absorbers at all four wheels can be electronically adjusted by the driver for softer or firmer operation. When the driver presses the SOFT or FIRM switch, a control module signals a solenoid or small electric motor in the top of each shock absorber. This opens or closes orifices of various sizes in the shock-absorber piston so that the resistance to fluid flow changes within the shock absorber. The switch may have a position that allows the shock absorbers to automatically select the proper damping for the road and driving conditions. Some vehicles have an electronic stability-control system that can selectively apply individual wheel brakes to help control the vehicle during cornering and skidding. *See* AUTOMOTIVE BRAKE.

Front suspension. Most automotive vehicles have independent front suspension, usually using coil springs as part of either a short-arm long-arm or a MacPherson-strut suspension system. In the short-arm long-arm system, two unequal-length control (suspension) arms attach through ball joints to the top and bottom of the steering knuckle. Typically, the coil spring is mounted between the lower control arm and the underbody or frame.

A MacPherson-strut suspension (see illus.) combines a coil spring and shock absorber into a strut assembly that requires only a beam-type lower control arm. The angled ends of a stabilizer bar attach to the outer ends of the lower control arms, holding the arms in position while helping to control the longitudinal loads. The top of the strut attaches through an upper strut bearing to the vehicle underbody, allowing the strut to turn with the steering knuckle. Some vehicles have a modified MacPherson suspension, in which the spring is mounted between the lower control arm and the underbody, instead of on the strut. Other strut-type front suspensions have the spring on the strut, with upper and lower control arms used to position the steering knuckle.

Some vehicles with short-arm long-arm front suspension use either longitudinal or transverse torsion bars for the front springs. One end of the torsion bar is attached to the lower control arm, and the other end is anchored to the vehicle body or frame. As the tire and wheel move up and down, the torsion bar provides springing action by twisting about its long axis. Turning an adjustment bolt at one end of the

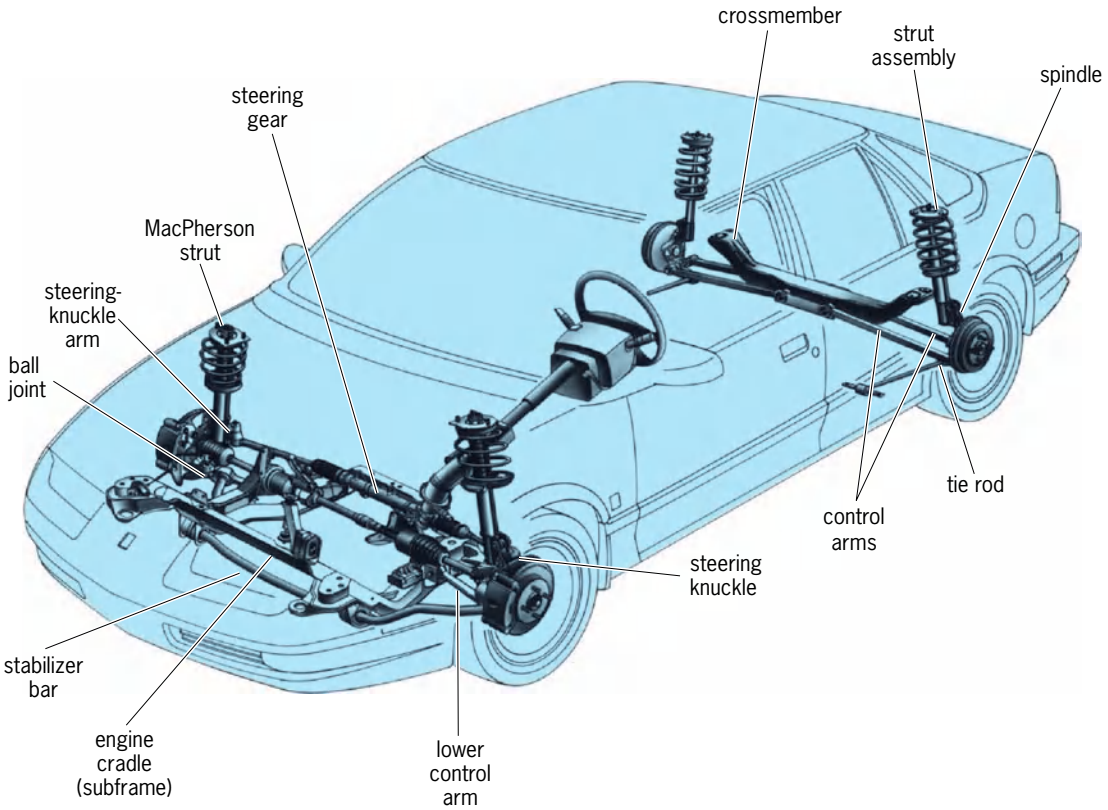

Front-wheel-drive car with MacPherson-strut front suspension and strut-type independent rear suspension. (*Saturn Corp.*)

torsion bar raises or lowers the vehicle ride height. *See* TORSION BAR.

Rear suspension. Most automobiles and many light trucks have coil springs at the rear. These may mount on the rear drive axle, on struts, or on various types of control or suspension arms in an independent suspension system. Some rear-drive vehicles have leaf springs at the rear. Others use transverse torsion bars.

Air suspension. The air-suspension system uses air springs at the front or rear of the vehicle, or at all four wheels. The air spring is a rubber cylinder or bag filled with compressed air which provides the springing action. As the load in the vehicle changes and the lower control arm moves up and down, a height sensor in the strut signals the change in position to the system controller. The controller then opens a valve in the air bag to add or release air. An electric air compressor on the vehicle connects to the valve and keeps the air springs properly inflated. The system controls spring rates and provides automatic height and level control.

Active suspension. An active suspension system uses computer-controlled hydraulic actuators (instead of springs and shock absorbers) to keep each tire pushing against the road with a constant force. As a tire meets an irregular road surface, the computer varies the hydraulic pressure to the actuator, slightly raising or lowering the wheel so the tire maintains the same force against the road. This action keeps the vehicle level while providing the best possible ride and handling. Since most of the tire-and-wheel movement is absorbed in the suspension, little shock and vibration reach the vehicle body and its occupants. *See* HYDRAULIC ACTUATOR.

Donald L. Anglin

Bibliography. J. C. Dixon, *Tires, Suspension and Handling*, 2d ed., Society of Automotive Engineers, Warrendale, 1996.

Automotive transmission

The device in the power train of a motor vehicle that provides different gear ratios between the engine and drive wheels, as well as neutral and reverse. An internal combustion engine develops relatively low torque at low speed and maximum torque at only one speed, with the crankshaft always rotating in the same direction. To meet the tractive-power demand of the vehicle, the transmission converts the engine speed and torque into an output speed and torque in the selected direction for the final drive. This arrangement permits a smaller engine to provide acceptable performance and fuel economy while moving the vehicle from standstill to maximum speed. The transmission may be a separate unit as in front-engine rear-drive vehicles or may be combined with the drive axle to form a transaxle as in most front-drive vehicles. *See* AUTOMOBILE; AUTOMOTIVE DRIVE AXLE; DIFFERENTIAL.

The two general classifications are manual transmissions that the driver shifts by hand after disengaging the foot-operated clutch, and automatic transmissions that shift with no action by the driver.

However, manual transmissions can have a clutch that is automatically disengaged by an actuator when the driver moves the shift lever, and automatic transmissions can have manual-shift capability which allows the driver to select the shift to the next lower or higher gear ratio by movement of the shift lever. *See* CLUTCH.

Automotive manual transmissions and transaxles typically have four, five, or six forward speeds. Manual transmissions in trucks can have up to 16 forward speeds. However, more than 85% of new cars manufactured in the United States have an automatic transmission or transaxle that has either four or five forward speeds. Both manual and automatic transmissions and transaxles usually have overdrive, a top forward gear ratio that causes the output shaft to overdrive or turn faster than the input shaft. In overdrive, the engine can run slower for improved fuel economy, reduced engine wear, and quieter engine operation while maintaining the same vehicle speed, but with reduced hill-climbing ability and acceleration. Some vehicles have a continuously variable transmission (CVT), which provides an infinite number of gear ratios. In one design, engine torque is transmitted through a belt that runs between two variable-diameter pulleys. By varying the ratio between the pulleys, the engine can operate at its optimum speed more of the time. *See* TRUCK.

Manual transmission. The manual transmission is an assembly of gears, shafts, and related parts contained in a metal case or gearbox partially filled with lubricant. The transmission input shaft connects through the clutch and flywheel to the engine crankshaft (**Fig. 1**). The transmission output shaft connects through a driveshaft to the final-drive gearing in the drive axle. To get the vehicle into motion, reduction or underdrive gearing in the transmission allows the engine crankshaft to turn fast while the drive wheels turn much more slowly but with greatly increased torque. As the vehicle accelerates, and less torque and more speed are needed, the driver shifts the transmission into successively lower numerical gear ratios, known as higher gears. In a typical five-speed manual transmission, gear ratios are approximately 3.35:1 for first gear, 2:1 for second gear, 1.35:1 for third gear, 1:1 (direct drive) for fourth gear, and 0.75:1 (overdrive) for fifth gear. Most transmissions with four or more forward speeds are operated by a floor-mounted shift lever. *See* GEAR; GEAR TRAIN.

Automotive manual transmissions are synchromesh (synchronized constant-mesh) transmissions. The teeth on the forward gears (and sometimes reverse) run in constant mesh, with no gear actually sliding in or out of mesh when a shift is made. Each gear is allowed to turn freely, or is locked to its shaft (usually the output shaft), by movement of the sleeve on a synchronizer clutch. The sleeve fits around the outside of the synchronizer hub, which is splined to the shaft.

When the driver moves the shift lever, the synchronizer clutch equalizes the speeds of the gear and the shaft. Then the synchronizer sleeve can slip over a smaller ring of external teeth which are formed around the face of the gear, locking that gear, through the synchronizer hub, to the shaft. This synchronizing action prevents gear clash during shifting. The sleeve is moved from its center position on the synchronizer hub, or neutral, toward the front or rear of the transmission by a shift fork that is mechanically connected to the shift lever. An interlock device in the transmission prevents two gears from engaging at the same time.

Manual transaxle. Most manual transaxles have either five or six forward speeds. The transmission section of the transaxle is basically the same in function and operation as the manual transmission. However, the final-drive pinion gear is on the transaxle output shaft and meshes with the helical ring gear bolted to

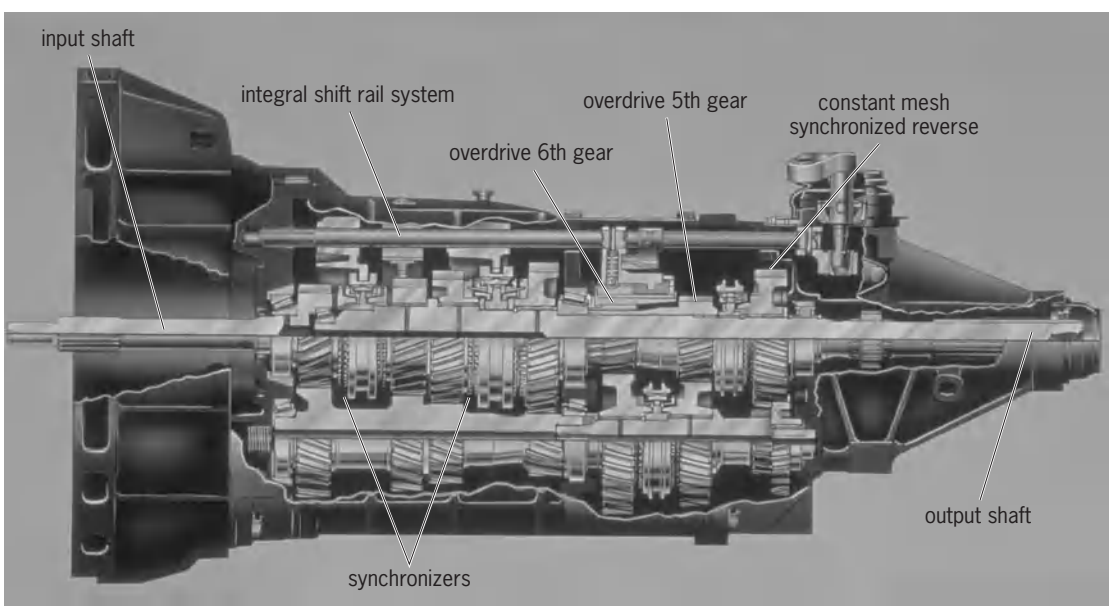

Fig. 1. Six-speed manual transmission for a rear-drive car. (*Pontiac-GMC Division, General Motors Corp.*)

the differential case. As the ring gear rotates, the differential pinion gears drive the differential side gears, which rotate the axle half-shafts to turn the wheels.

Automatic transmission. The automatic transmission provides automatic control of drive-away, gear-ratio selection, and gear shifting through four or five forward speeds. A typical automotive automatic transmission includes a hydrodynamic three-element torque converter with locking clutch, a planetary-gear system that provides overdrive in fourth or higher gear, and a hydraulic or electrohydraulic control system. Shifts are made without loss of tractive power. *See* HYDRAULICS; HYDRODYNAMICS; PLANETARY GEAR TRAIN; TORQUE CONVERTER.

The hydraulic torque converter uses an impeller, a turbine, and a stator to automatically and continuously vary the torque between the input and output shafts. The torque converter also provides for drive-away, and for damping of the shock and vibration caused by gear shifts. The amount of torque multiplication varies, depending on the difference in speed between the impeller (input shaft) and the turbine (output shaft). Torque multiplication decreases from a maximum of about 2.6:1 with the vehicle not moving (turbine stalled) until the ratio approaches 1:1, or the coupling point, as vehicle speed increases. Power loss through the torque converter can be eliminated by use of a centrifugal or electronically-controlled torque-converter clutch to lock the turbine to the converter cover. The clutch is normally unlocked at low speeds, during acceleration, and by a downshift. *See* TORQUE.

The automatic transmission has two or more planetary gear sets, which may be arranged in various ways to provide three or four forward speeds. These gear ratios provide further torque multiplication for acceleration and heavy road loads. In the power flow, the planetary gear sets are positioned after the torque converter. The individual elements of each planetary gear set are controlled by multiple-disc friction clutches and brake bands, which are actuated by oil pressure from an engine-driven oil pump in the front of the transmission. Some elements are controlled mechanically by a one-way clutch.

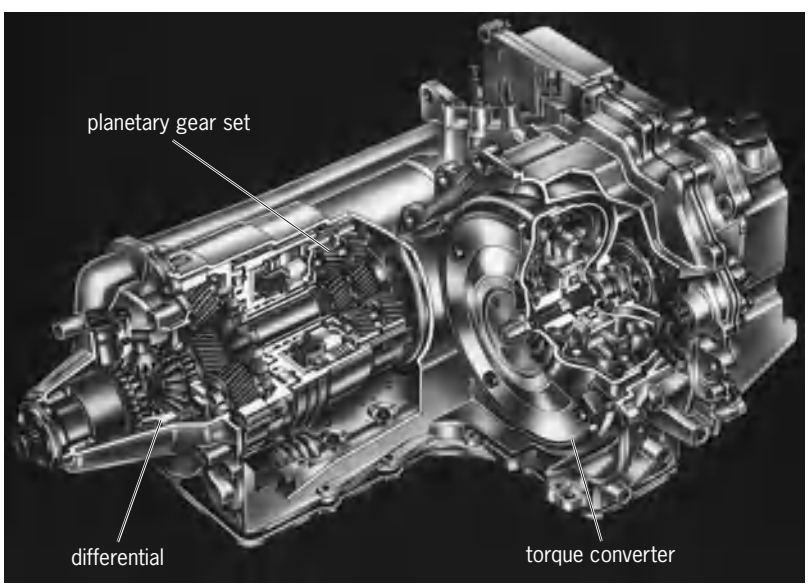

Fig. 2. Electronically controlled four-speed automatic transaxle for a front-drive car. (*Cadillac Division, General Motors Corp.*)

The oil pump supplies oil to fill the torque converter, hydraulic control system and valve body, and shift components. In addition, the oil lubricates and cools the transmission. The oil used in an automatic transmission is one of several types of automatic transmission fluid (ATF), usually dyed red for identification if leakage occurs. *See* LUBRICANT.

The typical automatic-transmission selector lever may be moved to Park, Reverse, Neutral, Drive, or Low (PRNDL). In a hydraulically controlled transmission, upshifts and downshifts are controlled by selector-lever position, engine vacuum, throttle position, and governor action. The governor is a speed-sensitive device that varies hydraulic pressure proportionally to output-shaft speed. *See* GOVERNOR.

Drive is the normal operating range, in which shifts occur primarily as a function of engine load and road speed. A vacuum modulator connected to engine intake-manifold vacuum, or the throttle position, allows the transmission to sense engine load. The governor senses road speed. With the selector lever in Drive and the vehicle at a standstill, the transmission starts out in first gear. As output-shaft speed increases, the governor allows increasing hydraulic pressure to the control-valve assembly. This causes spool valves to open and close oil passages to the hydraulic actuators and clutches, shifting the transmission to second gear. As vehicle speed increases, this basic sequence repeats until the transmission is in normal drive gear—direct drive (third) or overdrive (fourth or fifth). When the driver fully depresses the accelerator pedal, a wide-open-throttle forced downshift usually occurs for increased power or acceleration. *See* HYDRAULIC ACTUATOR.

In an electronically controlled automatic transmission, sensors monitor operating conditions such as vehicle speed, engine load, coolant temperature, and transmission oil temperature. This information is sent to the transmission controller or power-train control module (PCM). The controller decides when (shift timing) and how (shift quality) to make the shift. Signals are sent to electric shift solenoids on the control-valve assembly. The solenoids then open or close fluid passages, which send or release fluid pressure to the clutches and band servos controlling the planetary gear sets. Applying the brake pedal signals the controller to unlock the torque-converter clutch. *See* SOLENOID (ELECTRICITY).

The electronic controls are adaptive, automatically compensating for changes in engine or friction-element torque. Shift quality is not affected by a poorly performing engine. Under normal conditions, the controller "learns" to optimize shift quality during the first few shifts.

Some automatic transmissions and transaxles are similar in construction to manual transmissions and transaxles. Instead of planetary gears and a single centerline, two parallel shafts with spur and

helical gears are used. Synchronizers are replaced with hydraulically-operated multiple-disc clutches that are electronically controlled.

Automatic transaxle. Front drive vehicles usually have a four-speed or five-speed overdrive automatic transaxle with locking torque-converter clutch (**Fig. 2**). The final drive and differential arrangement are basically the same as in a manual transaxle. The transmission section is similar in construction and operation to the automatic transmission. Automatic transaxles may be hydraulically or electronically controlled.

To reduce overall length, many automatic transaxles have the torque converter and oil pump offset from the input shaft. Power is transferred from the torque-converter output shaft through sprockets and chain to the transaxle input shaft.

Donald L. Anglin

Bibliography. *Design Practices: Passenger Car Automatic Transmissions*, 3d ed., Society of Automotive Engineers, 1994; Robert Bosch GmbH, *Automotive Handbook*, 4th ed., 1996.

Autonomic nervous system

The part of the nervous system that controls visceral functions of the body. This system innervates smooth and cardiac muscle and the glands, and regulates visceral processes including those associated with cardiovascular activity, digestion, metabolism, and thermoregulation. The autonomic nervous system functions primarily at a subconscious level. It is traditionally partitioned into the sympathetic system and the parasympathetic system, based on the region of the brain or spinal cord in which the autonomic nerves have their origin. The sympathetic system is defined by the autonomic fibers that exit thoracic and lumbar segments of the spinal cord. The parasympathetic system is defined by the autonomic fibers that either exit the brainstem via the cranial nerves or exit the sacral segments of the spinal cord. *See* PARASYMPATHETIC NERVOUS SYSTEM; SYMPATHETIC NERVOUS SYSTEM.

The defining features of the autonomic nervous system were initially limited to motor fibers innervating glands and smooth and cardiac muscle. This arbitrary definition limited the autonomic nervous system to visceral efferent fibers and excluded the sensory fibers that accompany most visceral motor fibers. Although the definition is often expanded to include both peripheral and central structures (such as the hypothalamus), contemporary literature continues to define the autonomic nervous system solely as a motor system. This bias ignores the importance of the afferent pathways. Moreover, it confuses the study of the dynamic regulatory function of the autonomic nervous system, since the regulation of visceral state and the maintenance of homeostasis implicitly assume a feedback system with the necessary constituents of motor, sensory, and regulatory components. Thus, from a functional perspective, the autonomic nervous system includes afferent pathways conveying information regarding the visceral organs and the brain areas (such as the medulla and the hypothalamus) that interpret the afferent feedback and exert control over the motor output back to the visceral organs.

Phylogeny provides insights into the functional and anatomical organization of the autonomic nervous system. The autonomic nervous system may be phylogenetically organized by proceeding from primitive structures that conserve metabolic resources and regulate visceral homeostasis, to structures that mobilize for fight-flight behaviors, and finally to structures found only in mammals that promote social and emotional behavior. Paralleling these functional shifts are increases in the interaction between visceral and somatic neurons and increases in influences from higher brain structures. *See* HOMEOSTASIS; NERVOUS SYSTEM (VERTEBRATE).

Stephen W. Porges

Bibliography. J. N. Langley, *The Autonomic Nervous System*, Heffer and Sons, Cambridge, 1921; S. W. Porges, Love: An emergent property of the mammalian autonomic nervous system, *Psychoneuroendocrinology*, 23:837–861, 1998.

Autopilot

An automatic means for steering an aircraft or other vehicle. The original use of an autopilot, or automatic pilot, was to provide pilot relief during cruise modes. Autopilots now perform functions more rapidly and with greater precision than the human pilot. The functions, designs, and uses of autopilots vary widely depending on the type of vehicle. In addition to controlling various types of aircraft and spacecraft, autopilots are used to control ships or sea-based vehicles and in some cases land-based vehicles. This article discusses autopilots used in aircraft and space vehicles.

Development. Early autopilots were developed with very little theoretical foundation and were based primarily on the ingenuity of such aviators as L. Sperry, who flew the first automatically controlled aircraft in 1914. His designs evolved through the years to the point where a three-axis autopilot was employed in several commercial aircraft in the late 1930s. In 1933 W. Post used a three-axis autopilot in his around-the-world flight. The system included gyros with pneumatic pickoffs and three-axis control with proportional hydraulic servos. During World War II the United States developed electric autopilot designs that could accept maneuver commands either from the pilot or from guidance sensors. Also at that time, two important features evolved that are still essential for the continued growth of automatic flight controls: feedback control theory and electronic computers.

Basic concepts. An autopilot is unique equipment in that it is expected to make the aircraft fly in the same manner as a highly trained, proficient pilot. It must provide smooth control and avoid sudden and erratic behavior. The intelligence for control must

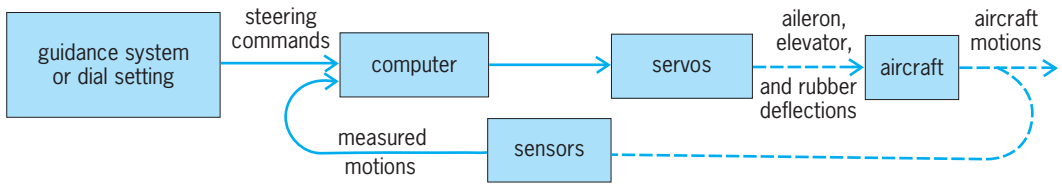

Fig. 1. Basic elements of an autopilot system.

come from sensors such as gyroscopes, accelerometers, altimeters, airspeed indicators, automatic navigators, and various types of radio-controlled data links. The autopilot supplies the necessary scale factors, dynamics (timing), and power to convert the sensor signals into control surface commands. These commands operate the normal aerodynamic controls of the aircraft. *See* ACCELEROMETER; AIRCRAFT INSTRUMENTATION; ALTIMETER; GYROSCOPE; INERTIAL GUIDANCE SYSTEM.

Autopilots come in varying degrees of sophistication. A simple attitude hold (wing leveler) just barely justifies the term autopilot, while a top-of-the-line system that automatically takes the aircraft from one location to another exceeds the normal capabilities of an autopilot. Sophisticated autopilots are no longer limited to military aircraft but are now common in commercial aircraft and are available for general aviation. In modern fly-by-wire aircraft the autopilot and the flight control system often reside together in the same digital computer, and it is difficult to separate their functions. These advanced systems provide the pilot relief functions plus help to stabilize the aircraft, protect the aircraft from undesirable maneuvers, and provide automatic landings (in some cases on a moving ship). Research aircraft are being tested with backup automatic control concepts that continue to control the aircraft even if the primary controls are damaged and no longer function. *See* FLIGHT CONTROLS.

Aircraft motion is usually sensed by a gyro, which transmits a signal to a computer (**Fig. 1**). The computer commands a control servo to produce aerodynamic forces to remove the sensed motion. The computer may be a complex digital computer, an analog computer (electrical or mechanical), or a simple summing amplifier, depending on the complexity of the autopilot. The control servo can be a hydraulically powered actuator or an electromechanical type of surface actuation. Signals can be added to the computer that supply altitude commands or steering commands. For a simple autopilot, the pitch loop controls the elevators and the roll loop controls the aileron. A directional loop controlling the rudder may be added to provide coordinated turns. *See* AILERON; AMPLIFIER; ANALOG COMPUTER; CONTROL SYSTEMS; DIGITAL COMPUTER; ELEVON; GUIDANCE SYSTEMS; SERVOMECHANISM.

Design. As autopilots become more complex, additional control loops are added. The pitch and roll axis of control must first be stabilized and controlled; then additional control loops can be added. Altitude is supplied by an altimeter; flight path commands come from a navigation system or another automatic system. Autopilots have provisions for pilot control or stick steering. A cruise control designed for fuel economy may command the aircraft to fly at the most economical altitude. An instrument landing system (ILS) or other automatic landing system can also command the autopilot. *See* AUTOMATIC LANDING SYSTEM; INSTRUMENT LANDING SYSTEM (ILS).

Any one of these commands could be selected to supply signals to the computer and are appropriately summed with one or more of the feedback loops. For longitudinal control (**Fig. 2**), pitch rate is the inner loop and is used for damping and stability. An accelerometer may also be used for damping and to limit the acceleration resulting from commands to the aircraft. Pitch angle is fed back to sum with flight-path commands so that the computer can calculate when the proper altitude is reached. The

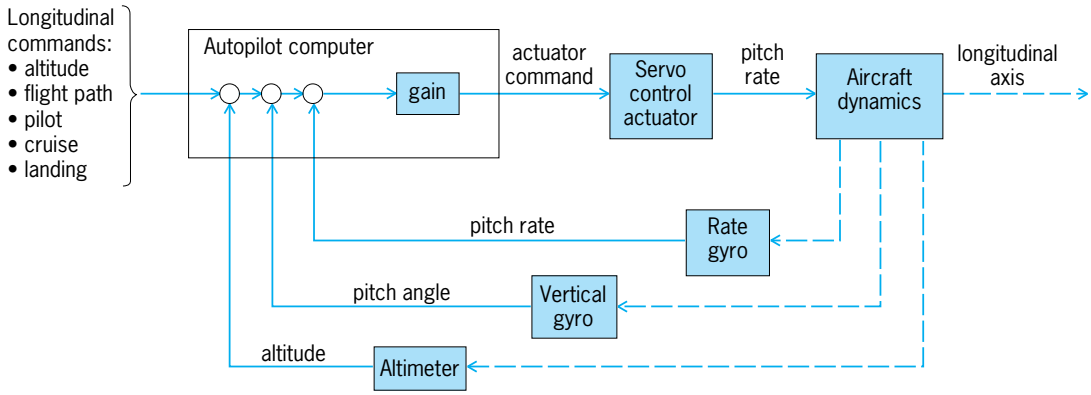

Fig. 2. Typical longitudinal autopilot control loop.

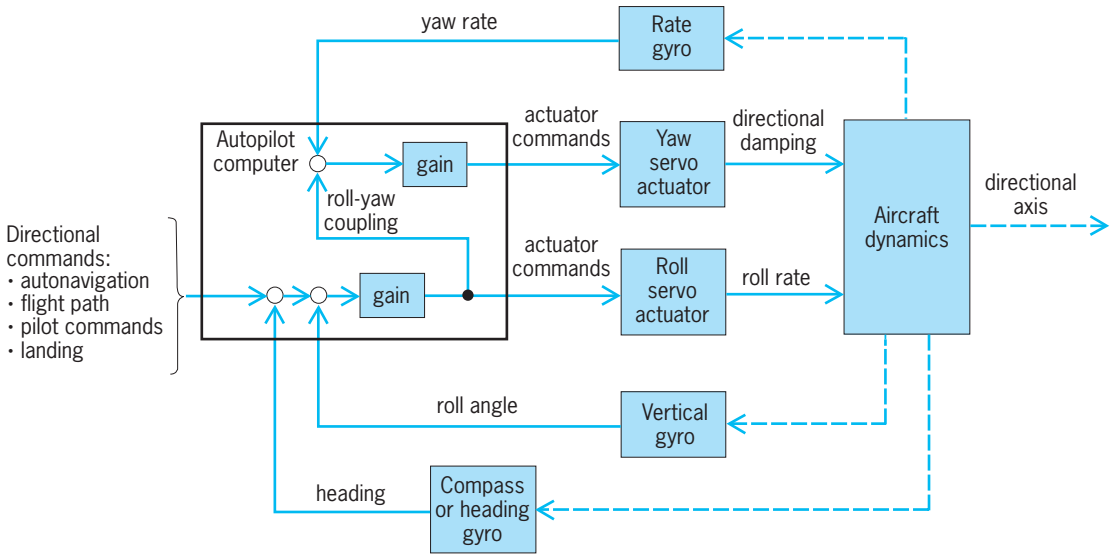

Fig. 3. Typical directional autopilot control loop.

altitude feedback is used to sum with altitude and cruise commands to permit the computer to calculate the proper flight altitude. After summing the appropriate signals, the computer produces an error signal that drives a servo actuator controlling a flight control surface. The control surface produces an aerodynamic force that causes a pitch rate to be generated on the aircraft. The appropriate feedback loop limits the pitch rate, pitch angle, and altitude to suitable values.

Directional control loops (**Fig. 3**) are slightly more complicated, since roll and yaw are coupled in the aircraft dynamics. Basically, the aircraft must be banked (roll) to produce a rate of change in the heading; therefore, roll and yaw must be controlled together. Normally a yaw rate gyro is used to control a vertical surface, such as a rudder, to provide directional damping. A lateral accelerometer may be used to help produce coordinated turns. In some aircraft an interconnecting link is provided between roll and yaw to aid in producing coordinated turns. Typical directional commands (Fig. 3) are similar to the longitudinal commands. The computer sums these signals to produce an error signal to command a roll control surface. This produces a roll rate until the proper roll angle is reached. The magnitude of bank angle is proportional to the required heading correction. As the heading error is removed, the aircraft returns to level.

Modern aircraft are becoming more complex and are using more control surfaces. Both the leading and trailing edges of the wing are now controllable, producing additional methods of pitch and roll control. The aircraft's horizontal surfaces move differentially, allowing the horizontal tail to be used for roll control. These additional control surfaces became practical with the development of fly-by-wire controls where hydraulic surface actuators are controlled with electrical signals. These signals are supplied from a computer, either analog or digital, which controls the aircraft.

Modern control theory. Automatic flight controls now use digital computers, which accept data from multiple sensors, to control an unstable aircraft with multiple feedback loops operating several aerodynamic control surfaces. With this type of system a control failure cannot be tolerated. The system is guarded against failures by having three or four channels operating in parallel with computers monitoring all the channels. When a failure is detected, the faulty unit is corrected or disconnected. This system monitoring adds further complexity to the system design because failures must be corrected without allowing undesirable inputs to the aircraft. Modern control theory enables systems of this type to be analyzed and designed by using computer synthesis. In modern control technique a desired performance criterion is selected and a unique design obtained. The performance of the system is then said to be optimal in terms of the defined criterion. On vehicles using automatic controls of this type the separation of flight controls and autopilot becomes obscure. For purposes of definition, the pilot relief functions are referred to as the autopilot, although the autopilot and flight controls are the same physical units of equipment. *See* OPTIMAL CONTROL THEORY.

Helicopter and VTOL autopilots. Helicopters and vertical takeoff and landing (VTOL) aircraft also use automatic control systems. Automatic control is especially helpful for hovering and low-speed flight since these aircraft are unstable at low speed. The control loops are similar to those discussed above except that the control servos control rotor blade angles on helicopters and vectoring thrust nozzles (or gas jets) on VTOL aircraft. A vertical axis of control may be used to control height. *See* HELICOPTER; VERTICAL TAKEOFF AND LANDING (VTOL).

Missile autopilots. The autopilots that control vehicles without flight crews use the same basic concepts as those described for aircraft autopilots. A missile autopilot is usually designed for use only once and therefore must be inexpensive. Since no crew is

involved, fail-safety is not a requirement; however, reliability is still important. Also, with no crew aboard it is possible to perform maneuvers involving higher accelerations. Turns do not have to be coordinated; therefore, the vehicle does not need to bank to turn. Maneuvering is limited only by the structure of the vehicle. The autopilot controls the missile through control surfaces that are powered with hydraulic, pneumatic, or electrical actuators. *See* GUIDED MISSILE; MISSILE.

A special class of uncrewed aircraft is the remotely piloted vehicle (RPV); the pilot stays on the ground or in another aircraft, and actually flies the vehicle with a radio link that controls the autopilot. Remotely piloted vehicles have been used for flight research. In the HI-MAT program of the U.S. Air Force, most of the autopilot components, including the computers, were on the ground. A data link transmitted signals between the computers and the aircraft's control servos and sensors. *See* DRONE.

Launch-vehicle autopilots. The design of launch-vehicle autopilots is very similar to that of aircraft autopilots except for the difference in the devices used to obtain control forces. The rocket-propelled launch vehicles or boosters, used to place spacecraft into orbit and ballistic missiles into trajectories, are controlled by vectoring (controlling the direction) of the thrust. This can be achieved by swiveling the rocket engines or deflecting their thrust with vanes powered by large hydraulic actuators. At low speeds, aerodynamic forces are negligible and all control must come from the thrust. As the vehicle gains speed, the aerodynamic forces increase and provide some damping. Launch-vehicle attitude is sensed by stable platforms, and rate gyros and accelerometers are used for damping. Trajectory commands are provided by the vehicle guidance system. *See* ROCKET PROPULSION.

Space-vehicle attitude control. Attitude-control systems are required in satellites and other space vehicles to maintain correct orientation. In the absence of an atmosphere there are no aerodynamic forces to use for control or to produce disturbing torques. The required control torques are several orders of magnitude lower than those for an aircraft. Attitude disturbances come from irregularities in the gravity and magnetic fields of nearby planets, the Sun's radiation pressure, and impacts of small meteorites.

The autopilot control-loop concept is very similar to that for aircraft. Attitude sensors such as stable platforms, gyros, and accelerometers are used. The principal position sensors are optical, and include star trackers, sun sensors, and horizon sensors. Active attitude control is obtained by optically establishing the pointing direction and measuring the error signal between the pointing desired and the actual vehicle position. The computer applies corrective torques on the vehicle by employing pairs of gas jets. Control by gas jets produces large forces, but gas supplies last only a few days or weeks. Thus gas jets are used typically on short-duration or crewed missions. Where small forces are required, they may be produced by inertial reaction against gyros or wheels, by solar radiation pressure, or by reaction with local gravity or magnetic fields. These devices are useful on missions lasting many months or years. *See* STAR TRACKER.

Passive attitude control is sometimes possible for uncrewed long-duration parts of complex missions and is then preferred because of its higher long-term reliability. The whole vehicle may be spun like a gyro; its spin axis then tends to remain fixed in space. The space vehicle may be designed with a suitable inertia distribution, for example, by deploying long booms that tend to be attracted to the gravity field of the Earth. This system is undamped and requires some form of inertia damping. Another type of passive attitude control is the use of vanes mounted on short booms that react to the photon pressure of the Sun and keep the vehicle pointing at the Sun. This system also requires damping. A system of this type was used on a *Mariner* Mars probe in 1964. *See* SPACE NAVIGATION AND GUIDANCE; SPACECRAFT STRUCTURE.

Lloyd L. Kohnhorst

Bibliography. J. H. Blakelock, *Automatic Control of Aircraft and Missiles*, 2d ed., 1991; J. J. D'Azzo and C. H. Houpis, *Linear Control System Analysis and Design*, 3d ed., 1988; P. Garrison, *Autopilots, Flight Directors, and Flight Control Systems*, 1985; N. Radovcich, Active control in tomorrow's marketplace, *Aerosp. Amer.*, 26(8):14–22, August 1988.

Autoradiography

A photographic technique used to localize a radioactive substance within a solid specimen; also known as radioautography.

Procedure. A photographic emulsion is placed in contact with the object to be tested and is left for several hours, days, or weeks, depending on the suspected concentration of the radioactive material to be measured. The emulsion, which is a gel containing silver halide, is then developed, fixed, and washed as in the usual photographic process. At sites where the emulsion was close enough to the radioactive substance, it appears dark because of the presence of silver grains (see **illus.**). When the number of grains is insufficient to darken the film to the unaided eye, the film may be examined with the aid of a microscope. The individual silver grains may then be seen. The pattern formed by the grains depends on the type of radiation and the nature of the photographic emulsion. Alpha particles produce short, straight rows or tracks of grains. Beta particles as well as x-rays and gamma rays, which affect film by producing beta particles, produce tortuous tracks whose lengths and grain densities depend on the energy of the beta particles. Low-energy particles produce shorter tracks with higher grain densities. Very low energy particles like those from tritium (hydrogen-3) may produce only a single grain very close to the site of decay.

Chromatographs are frequently used to prepare autoradiographs. X-ray film is applied in close contact with the surface. The surface is protected by a thin

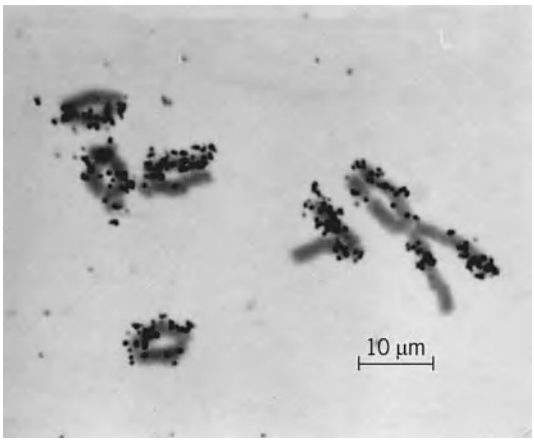

Photomicrograph of the chromosomes of a plant root cell and its autoradiogram. Black grains indicate the presence of tritium in certain parts of the chromosome.

plastic sheet in the case of wet gels. Such gels have been used extensively for analysis of nucleic acids and are the basis for the widely used techniques for the determination of the nucleotide sequences of nucleic acids. Of course, all molecules must have been made radioactive, usually by attaching a radioisotope or synthesizing the smaller molecules with a radioisotope such as phosphorus-32 or carbon-14 before their separation and use for autoradiography. *See* CHROMATOGRAPHY; NUCLEIC ACID; RADIOACTIVE TRACER; RADIOISOTOPE (BIOLOGY).

The principal advantage of autoradiography is that emulsions can be stored for weeks or months in the cold to reduce fading of the latent image and very small amounts of radioactivity can be detected. Even greater enhancement is possible by the use of scintillators which absorb the energy of radioactivity and emit light. At low temperatures (−70°F or −56°C) with tritium, the enhancement can be 60-fold or greater. The enhancement is useful in autoradiography of chromatograms containing tritium but is not desirable where high resolution is required. *See* BETA PARTICLES; GAMMA RAYS; X-RAYS.

Resolution. Autoradiograms (the picture obtained by autoradiography) can be made with many objects which can be brought in contact with photographic film. The resolution, that is, the precision with which the radioactive substances can be located, is good only when the specimen is flat and in close contact with the film or photographic plate. Whole plants, animals, rocks, or any other objects may be used, but the flat, cut surface of the object provides better contact.

The precision with which the source of radioactivity can be localized (resolution) is dependent on the nearness of the emulsion to the specimen. The highest resolution is obtained by embedding the biological specimen in the emulsion. Flattened cells or thin sections of cells on glass slides can be dipped in melted emulsion, or coated with thin strips of wet emulsion which is then dried before exposure. For very high resolution with the electron microscope, the grids containing the specimen are dipped in dilute melted emulsion, or a loop of emulsion is lowered over the grids. All procedures must be carried out in very low intensity red light or complete darkness. For autoradiography of chromatograms or gels produced by electrophoresis, the specimens are dried and placed in contact with a sheet of x-ray film, with or without a thin protective sheet between. The protective sheet may contain scintillators to enhance the image for specimens that emit low-energy beta particles. Under some conditions it may be possible to get the scintillators into the specimen to improve efficiency, but if the materials are embedded in gels, the scintillators penetrate poorly. Autoradiograms are useful in detecting deoxyribonucleic acid (DNA) for sequencing studies.

Electron microscopy. The technique has been extended to the location of radioactive molecules in specimens photographed with the electron microscope. The emulsion with very small crystals of silver salts is spread on a thin section of the type used for electron microscopy, somewhat as in the dipping technique described. To coat the section, a loop of wire containing a thin film of liquid emulsion is lowered onto the specimen mounted on the electron microscope grid. After drying, the processing is similar to that described for the light microscope except that development is regulated so that the developed silver grains are small enough not to obscure the tiny specimens which are being photographed in the electron microscope. *See* ELECTRON MICROSCOPE.

Applications. Autoradiography can be used to detect, and measure semiquantitatively, the radioactive materials in almost any object that can be placed in contact with film or photographic emulsion in some form. However, in biological research the object may be (1) a whole plant or animal that can be flattened against a film; (2) the cut surface of a plant or animal, or one of its organs; (3) thin sections of tissues or cells; (4) squashed or otherwise flattened cells; (5) surface films produced by spreading on water the protein monolayers containing DNA or ribonucleic acid (RNA) that are picked up on grids for electron microscopy; (6) sheets of paper or other materials on which radioactive substances have been separated by chromatography or electrophoresis; or (7) acrylamide gels in which DNA, RNA, or proteins have been separated by electrophoresis.

High-resolution autoradiograms remain a powerful tool for detecting and measuring small amounts of radioactivity that cannot be detected in any other way. Most radioactive isotopes produce radiations that penetrate the emulsions deeply enough to produce a somewhat out-of-focus picture, but tritium (3-hydrogen) is an exception. Its very low energy beta particles are stopped very close to the site of decay in emulsions, and when the beta particles are in a part of a molecule such as DNA, the site of the molecule can be visualized with high resolution. Tritium-thymidine of very high specific activity is used to label viral and bacterial DNAs so that individual molecules (chromosomes) can be visualized with a light microscope in emulsions by grain patterns. Similar autoradiograms of fragments

of DNA from chromosomes of higher cells are used to show patterns of replication. *See* RADIOGRAPHY.

J. Herbert Taylor

Bibliography. J. R. Baker, *Autoradiography: A Comprehensive Overview*, 1989; T. Maniatis et al., *Molecular Cloning: A Laboratory Manual* 2d ed., 1989; M. A. Williams, *Autoradiography and Immunochemistry*, 1978; M. A. Williams (ed.), *Practical Methods in Electron Microscopy: Autoradiography and Immunochemistry*, vol. 6, 1978.

Autotransformer

A special form of transformer having one winding, a portion of which is common to both the primary and the secondary circuits. The current in the high-voltage circuit flows through the series and common windings (see **illus.**). *See* WINDINGS IN ELECTRIC MACHINERY.

The current in the low-voltage circuit flows through the common winding and adds vectorially to the current in the high-voltage circuit to give the common winding current. Thus, an electrical connection exists between high-voltage and low-voltage windings. Because of this sharing of parts of the winding, an autotransformer having the same kilovolt-ampere output rating is generally smaller in weight and dimensions than a two-winding transformer. The equivalent size of a two-winding autotransformer without taps is given by the co-ratio times the kilovolt-ampere output, where the co-ratio equals $(HV - LV)/HV$. When the co-ratio is small, that is, when the high-voltage and low-voltage magnitudes are close together, the cost advantage in favor of an autotransformer is large. As the co-ratio becomes large, the equivalent size and the cost of an autotransformer approach that of a normal transformer.

One possible disadvantage of autotransformers is that the windings are not insulated from each other and that the autotransformer provides no isolation of the primary and secondary circuits.

Types. Autotransformers of large sizes are used for interconnecting high-voltage power systems. They are used in small sizes for intermittent-duty starting of motors. For this use the motor is connected for a short time to the common winding voltage, and then connected to the full line voltage. Small, variable-ratio autotransformers are used in testing and as components of other apparatus. Autotransformers are also used as the induction coil in the ignition system of spark-ignition engines. *See* IGNITION SYSTEM.

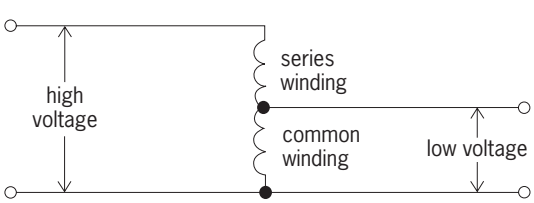

Typical autotransformer circuit.

Characteristics. Because of their smaller equivalent size, autotransformers generally have lower no-load loss, exciting current, and load loss than the corresponding transformers. In addition, the impedance and regulation normally are lower because of the connection between the two circuits.

Taps. Taps can be provided in the autotransformer to adjust the turns ratio. This provides control of the output voltage over the operating range of the transformer. These taps may be placed in the series winding or in the common winding. The problems of switching from tap to tap under load are similar to those encountered in tap changing on a two-winding transformer. The switching from tap to tap can be done without interruption of service.

Small variable-ratio autotransformers used for testing have a brush contact which serves as a common line. This contact may be moved across the turns to give a common line voltage from approximately 100 to 0% of the high voltage. *See* TRANSFORMER.

J. R. Sutherland

Autoxidation

The slow, flameless combustion of materials by reaction with oxygen; sometimes spelled autooxidation. Autoxidation is important because it is a useful reaction for converting compounds to oxygenated derivatives, and also because it occurs in situations where it is not desired (as in the destructive cracking of the rubber in automobile tires).

Although virtually all types of organic materials can undergo air oxidation, certain types are particularly prone to autoxidation, including unsaturated compounds that have allylic hydrogens or benzylic hydrogens; these materials are converted to hydroperoxides by autoxidation. This reaction gives particularly good yields (98%) where there is a single allylic hydrogen, a hydrogen atom located on a carbon atom that is adjacent to a double bond, as in cyclohexene, reaction (1).

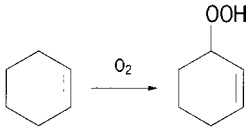

(1)

Certain saturated hydrocarbons with unique hydrogens can also be air-oxidized to good yields of hydroperoxides; one example is isobutane, which is oxidized to *tert*-butyl hydroperoxide in commercial practice, reaction (2).

$$\underset{\text{Isobutane}}{(CH_3)_3CH} \xrightarrow{O_2} \underset{\text{\textit{tert}-Butyl hydroperoxide}}{(CH_3)_3{-}OOH} \qquad (2)$$

Both plants and animals contain lipoxygenase enzymes that convert unsaturated fatty acids containing a *cis-cis*-1,4-pentadienoic system to the allylic hydroperoxide. This reaction is analogous to nonenzymatic reactions (1) and (2). For example, lipoxygenase-1, an enzyme isolated from

soybeans, converts linoleic acid primarily to the 13-hydroperoxide in which the original double bonds have shifted into conjugation, reaction (3). The en-

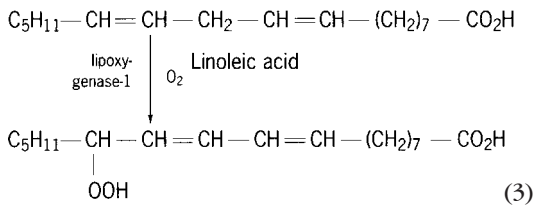

(3)

zymes that effect this conversion in animals are associated with the prostaglandin enzyme systems and normally have arachidonic acid, the 20-carbon, tetra-unsaturated fatty acid, as their substrate. The products of such reactions are hydroperoxy-eicosatetraenoic acids and are extremely important intermediates in the biosynthesis of potent physiological regulators called leukotrienes. *See* EICOSANOIDS; ENZYME.

Reaction mechanism. Autoxidation is a free-radical chain process. Such reactions can be divided into three stages: initiation, propagation, and termination. In the initiation process, some event causes free radicals to be formed. For example, free radicals can be produced purposefully by the decomposition of a free-radical initiator, such as benzoyl peroxide, as shown in reaction (4). (The free rad-

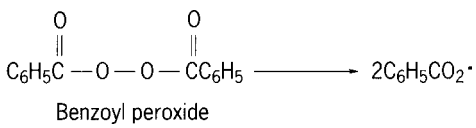

(4)

ical is indicated by the dot representing its unpaired electron.) In some cases, initiation occurs by a process that is not well understood but is thought to be the spontaneous reaction of oxygen with a material with a readily abstractable hydrogen, reaction (5). This random process is responsible for

$$RH + O_2 \rightarrow R\cdot + HOO\cdot \quad (5)$$

the slow discoloration and decay of paper, plastic, and rubber materials left in the air for long periods. Destructive autoxidation processes also are initiated by pollutants such as those in smog.

Once free radicals are formed, they react in a chain to convert the material to a hydroperoxide, ROOH. For example, reactions (6) and (7) show the autox-

$$R\cdot + O_2 \longrightarrow ROO\cdot \quad (6)$$

$$ROO\cdot + RH \longrightarrow ROOH + R\cdot \quad (7)$$

idation chain for the compound RH.

The chain is ended by termination reactions in which free radicals collide and combine their odd electrons to form a new bond; for example, reaction (8) shows the termination reaction of two

$$R\cdot + R\cdot \longrightarrow R{-}R \quad (8)$$

R· radicals. *See* CHAIN REACTION (CHEMISTRY); FREE RADICAL; ORGANIC REACTION MECHANISM.

Significance. Autoxidation is a process of enormous economic impact, since all foods, plastics, gasolines, oils, rubber, and other materials that must be exposed to air undergo continuous destructive reactions of this type. The slow rancidity that develops in milk and cheese is an example, as is the off-flavor and toxic compounds that develop in fatty meats with long standing in air. The food and plastics industries are actively involved in research to develop acceptable and efficient antioxidants to slow down these natural autoxidation reactions. All plastics and rubber and most processed foods contain antioxidants to protect them against the attack of oxygen. *See* ANTIOXIDANT; FAT AND OIL (FOOD); FOOD MANUFACTURING; PLASTICS PROCESSING; RUBBER.

Autoxidation also causes a slow decay of plant and animal tissue, especially tissue that is subjected to unusually high oxidation environments. For example, emphysema, cataract formation, some types of liver diseases, some types of arthritis, certain chemically induced cancers, and even aging itself are believed to be due in part to autoxidation reactions of susceptible tissue such as the polyunsaturated fatty acids in fats and lipids. *See* COMBUSTION; OXIDATION PROCESS.

William A. Pryor

Bibliography. J. K. Kochi (ed.), *Free Radicals*, vol. 2, pp. 3–62, 1973; W. A. Pryor, *Free Radicals*, 1966; W. A. Pryor (ed.), *Free Radicals in Biology*, vol. 1, 1976, vol. 5, 1982; C. Walling, *Free Radicals in Solution*, 1957.

Auxin

Any of a group of organic compounds which, when applied in low concentration, are able to promote elongation growth of plant shoots excised from a growing region of a young seedling. These substances were the first plant hormones to be studied, and were given the generic name auxins, indicating an increase, by F. Kögl and A. J. Haagen-Smit in 1931. The ability to increase the rate of shoot elongation is a key to the designation of a synthetic or naturally occurring compound as an auxin. However, auxins, and the other plant hormones, influence a variety of plant processes during various stages of plant development.

Discovery of auxin. Studies in the nineteenth century showed that plants were sensitive to light and gravity, and that their growth pattern could be modified by changing the plant's orientation toward such stimuli. These responses were given the name tropisms. This work led directly to the discovery of auxins. In 1872 T. Ciesielski showed that in roots the control of growth and of curvature in response to gravity depends on a stimulus originating from the root tip. His experiments were followed up in 1880 by C. Darwin with experiments on shoots and studies on the transmitted stimulus in phototropism.

In the 1920s and 1930s several investigators began to study the chemical nature of the growth-promoting stimulus coming from plant shoot tips. In 1926 Fritz Went developed the first quantitative

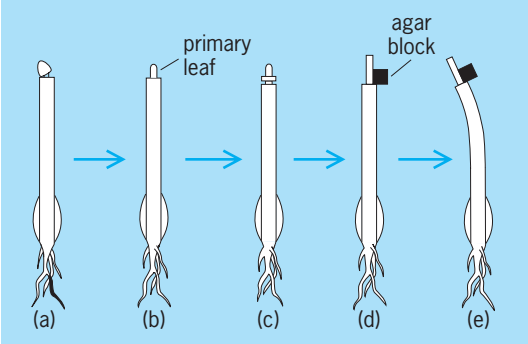

Fig. 1. Oat curvature test for auxin. (*a*) The tip is removed. (*b*) The plant is allowed to remain decapitated for 3 h. (*c*) A second short piece of stump is removed. (*d*) An agar block containing auxin is placed against the plant tissue. (*e*) After 90 min there is curvature due to asymmetric growth on two sides.

biological assay for auxins—the oat coleoptile curvature test (**Fig. 1**). Soon after, the structure and identity of the major naturally occurring auxin, indole-3-acetic acid (IAA), was described (**Fig. 2**). It was not until the late 1940s that the structure was confirmed by adequate chemical methods, and not until the 1970s that the structure of plant-derived indoleacetic acid was found by use of mass spectrometry. *See* MASS SPECTROMETRY; PLANT MOVEMENTS.

Occurrence and biosynthesis in plants. IAA has been found in almost all plant tissues that have been studied. It occurs in minute quantity, usually in less than micromolar amounts. This means that less than 1 oz (28 g) would be present in 150 tons (135 metric tons) of plant material. The understanding of auxins has been closely tied to the successful separation of such small amounts of active compounds from the great bulk of plant constituents. In particular, the development of chromatography has allowed auxins such as indoleacetic acid to be separated from plant material without chemical alteration, which is essential for subsequent biological and chemical tests. The coupling of chromatographic techniques with chemical and spectral analysis (in particular, mass spectrometry) has resulted in the characterization of a number of growth-promoting compounds and in the identification of several important metabolites of the known auxins. *See* CHROMATOGRAPHY.

Perhaps most significant has been the finding that auxins occur in plant tissue in several chemical forms. They may occur as the free active hormone (such as indoleacetic acid), and they may also be present in plant tissue as a number of so-called bound auxins (Fig. 2). The bound forms are auxins linked by a covalent chemical bond to some other chemical compound. These bound forms are thought to be important reserve forms which function to regulate the levels of free hormone in tissue, especially during certain developmental stages such as seedling growth. In addition, it is possible that the compound to which the auxin is linked (usually a sugar or an amino acid) is important in the transport of the hormone within the plant. Thus, auxin conjugates may serve as a type of biological "ZIP code" which determine the distribution of the hormone once it enters the transport system. Conjugation also renders the auxin immune to many of the enzymes which would normally degrade the free auxin. Thus, conjugation prevents further metabolism until the hormone has reached its target site and is released as the free active auxin. When synthetic auxins (for example, naphthalene acetic acid; 2,4-dichlorophenoxyacetic acid, or 2,4-D) are supplied to plant material, they will usually be conjugated to sugars or amino acids by the plant in a manner similar to that which occurs with the natural auxin.

In bean seeds, IAA has been shown to be conjugated to specific plant proteins. Early evidence also indicates that attachment of IAA to cell proteins may be a more general feature of auxin metabolism.

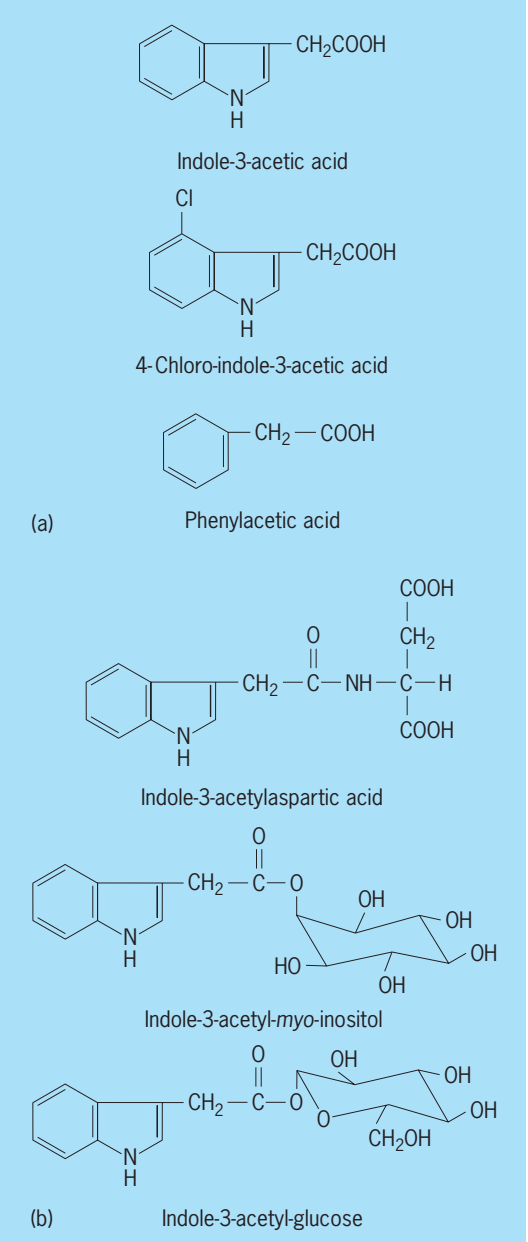

Fig. 2. Chemical structures of auxins. (*a*) Auxins isolated from plants. (*b*) Auxin conjugates found in plants.

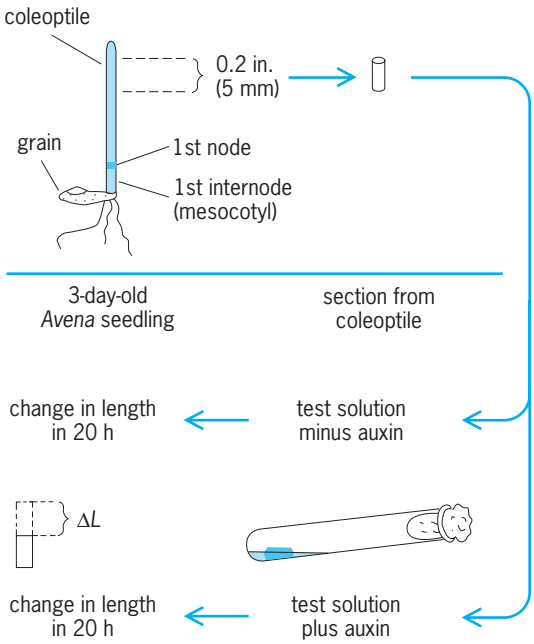

Fig. 3. Oat straight-growth bioassay.

While modifications of proteins after assembly by methylation and phosphorylation have been carefully studied and shown to be an important aspect of metabolic regulation in many organisms, the substitution of proteins by plant hormones is a novel discovery. The role played by these IAA-containing plant proteins is not yet known. Two general hypotheses have been proposed: First, IAA proteins represent an additional conjugate class with a role in controlling the availability of IAA in the plant or, second, IAA proteins are more directly involved in the mechanism by which auxins exert their regulatory function.

Biological activity. A number of experimental systems have been developed to study the auxin response. Most of these rely on the curvature of young seedlings treated on one side with the growth-active compound, or on the increase in growth of short tissue segments treated with the auxin. The use of the coleoptile of 3- or 4-day-old oat seedlings has most often been used, although most seedling stems show a pronounced response when treated with auxin and are thus suitable for such work. Several of the most common biological assays for auxin are shown in Fig. 1 and **Figs. 3** and **4**.

Auxin responses in plant tissue can be artificially divided into two groups, based on the time of their appearance. First, there are a number of very rapid responses which can be measured within minutes after the addition of auxin. Included in these responses are auxin-induced increases in protoplasmic streaming, cell elongation, and an increase in the acidity of the cell-wall free space. The second group of responses includes the long-term effects in which the observable response can be measured only after hours or even days of hormone treatment. Examples of these responses are auxin-induced increases in ribonucleic acid and protein synthesis, initiation of xylem differentiation, and an inhibition of lateral bud growth.

A major challenge to plant physiologists is understanding the mechanisms by which auxin elicits both these long- and short-term responses in plants and how the mechanisms are related. Short-term cell elongation responses to auxins are explained in part by the acid growth theory, which proposes that the first response to auxin is acidification of the cell wall region, resulting in the activation of wall-loosening enzymes involved in elongation growth. The general observations that support this theory are as follows: (1) Acid solutions will cause transient increases in the rate of cell elongation, both in intact tissue segments and in frozen-thawed segments placed under tension. (2) Auxin-sensitive tissues will acidify the cell wall region if given auxin treatment. (3) The rapid kinetics of auxin-induced cell elongation is similar to the kinetics for wall acidification.

Long-term responses to auxin have been attributed to direct interaction with plant deoxyribonucleic acid (DNA) resulting in increased transcription of specific messenger ribonucleic acids (mRNAs). Polysaccharide synthetases are proposed as one product of auxin-induced transcription resulting in increased production of cell wall material. The mechanism of auxin involvement with the plant genome has been theorized to be through binding of auxin, which passes through the plasma membrane to cytosol, with protein mediators. Such soluble auxin-binding proteins have been purified and shown to promote transcription.

Physiologically combining both short- and long-term effects of auxin has been the subject of speculation based on studies of intracellular signaling in cells involving calcium and metabolites of plasma membrane phospholipids known as phosphoinositides. These signals are released from storage into the cytosol by hormone–receptor interactions and effect growth-promoting cell processes. Theories for auxin-induced plant growth involving these signals are taken largely from animal cell systems, which have received more study; however, certain evidence obtained from plant cells has made such theories plausible and encourages research into these processes.

The generalized model for auxin action begins when a membrane receptor responds to auxin by releasing phosphoinositides from the pool of

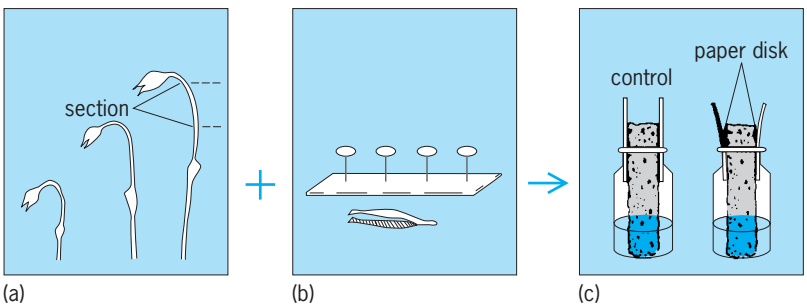

Fig. 4. Bean first-internode bioassay. (*a*) A section (1.5 in. or 4 cm) is cut from the stem between two nodes. (*b*) A test solution is applied to paper disks and placed against the side of the section. (*c*) After 3 h, bending is measured and compared with a control section.

membrane-bound phospholipids. These signals are carried through the cytosol and initiate release of calcium stored in vacuole and endoplasmic reticulum compartments possibly through their own receptor systems. This sudden increase in calcium initiates a response from at least two cell systems. First, the active removal of calcium from cytosol back to the vacuole is begun through the pumping of calcium in exchange for protons in the vacuole. These protons are in turn pumped out of the cytosol and into the cell-wall where acidification occurs. Second, the increase of calcium in cytosol acts to stimulate enzymes known as protein kinases which modify other proteins by phosphorylation. One protein that is modified by phosphorylation in this way is the proposed auxin-binding protein mediator that becomes receptive to auxin and can thus act on plant DNA to promote transcription of mRNAs that are critical for enzymes involved in cell growth. Evidence directly from plant systems for this model is as follows: (1) The release of calcium from subcellular compartments was enhanced by auxin or phosphoinositides in soybean and zucchini cell cultures. (2) The accumulation of calcium into vacuoles, that is, tonoplast vesicles, occurs quickly in carrot. (3) Protein phosphorylation occurs rapidly in the presence of calcium in zucchini and corn preparations. (4) Auxin enhances hydrolysis of the membrane-bound phospholipids into phosphoinositides that are vital in releasing calcium from intracellular storage in *Catharanthus*. This model proposes a direct relationship between the effects of auxin on cell-wall acidification and cell-wall matrix synthesis in growth-limiting epidermal cells. These effects can begin 30 min after auxin treatment. Continued research on the role of phosphoinositides and calcium in plants will undoubtedly test the connection between such signaling systems and the mechanism of action of auxin in plants.

Uses. The most widespread agricultural use of auxins is for weed and vegetation control. Synthetic auxins, such as 2,4-D and 2,4,5-trichlorophenoxyacetic acid (2,4,5-T) and their homologs, are commonly used as herbicides (**Fig. 5**). It has been estimated that the yearly savings to farmers achieved by use of these compounds in one midwestern state alone would more than pay for all of the auxin research in the United States to date. Synthetic auxins have found practical use for other agricultural needs as well. For example, auxin applications are effective for floral thinning of overproductive orchard trees, such as olive, plum, orange, and persimmon. Synthetic auxins have also been used to prevent premature fruit drop and to improve fruit quality in tree crops, such as bananas, apricots, and plums. *See* HERBICIDE.

The use of auxin for the propagation of plants from cuttings dates back to antiquity. Early farmers and gardeners used seed grain as an auxin source for such purposes. Today, commercial preparations, which usually contain the synthetic auxins indole-3-butyric acid or napthaleneacetic acid, are used. Auxin treatments have been used to enhance root-

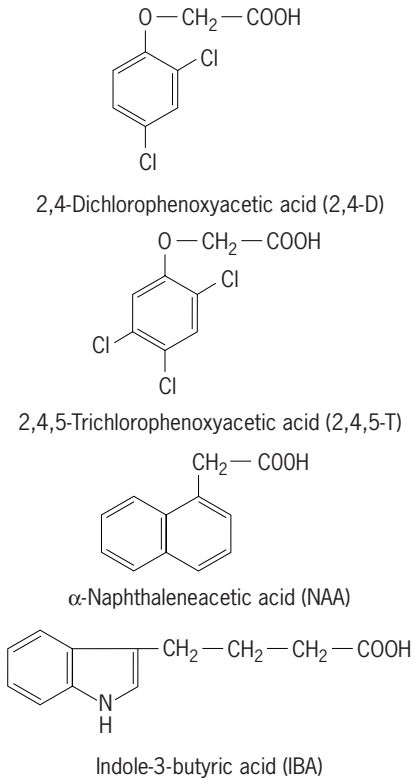

Fig. 5. Structure of synthetic auxins.

ing of over 1000 different plant species, and have been applied on a practical scale to over 30 different species throughout the world. Auxin preparations designed to enhance root formation in cuttings are also available for home use.

The production of large numbers of genetically identical (clonal) plants is now possible by use of plant cell culture techniques. Additions of auxins as well as another type of plant hormone, cytokinin, are usually necessary for the growth of such cultures. Both the type and amount of these two phytohormones are important for regeneration of new intact plants from the cell cultures and seem to differ for each plant type. Over 300 different plant species, ranging from ferns to forest trees, are propagated on a commercial scale by cell culture techniques. In addition, ongoing research on the production of new varieties of agriculturally important plants, using the emerging techniques of molecular biology, rely on cell culture methods and a detailed knowledge of the role of auxins and other plant hormones in the developmental life of plants. *See* CYTOKININS; MOLECULAR BIOLOGY; PLANT GROWTH; PLANT HORMONES; TISSUE CULTURE.

Jerry D. Cohen; Bruce G. Baldi

Bibliography. M. Bopp, *Plant Growth Substances 1985*, 1986; D. A. Brummell and J. L. Hall, Rapid cellular responses to auxin and the regulation of growth, *Plant Cell Environ.*, 10:523–543, 1987; J. D. Cohen and R. S. Bandurski, Chemistry and physiology of the bound auxins, *Annu. Rev. Plant Physiol.*, 33:403–430, 1982; P. J. Davies, *Plant Hormones and Their Role in Plant Growth and Development*, 1987; C. Ettlinger and L. Lehle, Auxin induces

rapid changes in phosphatidylinositol metabolities, *Nature*, 331:176–178, 1988; U. Kutschera and W. R. Briggs, Rapid auxin-induced stimulation of cell wall synthesis in pea internodes, *Proc. Nat. Acad. Sci.*, 84:2747–2751, 1987.

Avalanche

In general, a large mass of snow, ice, rock, earth, or mud in rapid motion down a slope or over a precipice. In the English language, the term avalanche is reserved almost exclusively for snow avalanche. Minimal requirements for the occurrence of an avalanche are snow and an inclined surface, usually a mountainside. Most avalanches occur on slopes between 30 and 45°.

Types. Two basic types of avalanches are recognized according to snow cover conditions at the point of origin. A loose-snow avalanche originates at a point and propagates downhill by successively dislodging increasing numbers of poorly cohering snow grains, typically gaining width as movement continues downslope. This type of avalanche commonly involves only those snow layers near the surface. The release mechanism in this case is analogous to that which would occur in dry sand. The second type, the slab avalanche, occurs when a distinct cohesive snow layer breaks away as a unit and slides because it is poorly anchored to the snow or ground below. A clearly defined gliding surface as well as a lubricating layer may be identifiable at the base of the slab, but the meteorological conditions which create these layers are complex. The thickness and areal extent of the slab may vary greatly, and those slab avalanches with larger dimensions pose the greatest threat to life and property. Both loose and slab types may occur in dry or wet snow. Dry avalanches often entrain large amounts of air within the moving mass of snow and are thus referred to as powder avalanches. Velocities for dry-snow avalanches may exceed 150 mi/h (67 m/s). Wet-snow avalanches occur when liquid water is present in the snow cover at the point of origin. While the wet avalanches move at lower velocities, they often involve greater masses of snow and therefore significant destructive forces. Theoretical calculations and empirical evidence indicate the general range of maximum impact forces to be between 7 and 70 lb/in.^{2} (5 and 50 metric tons/m^2), with extreme values reaching 142 lb/in.^{2} (100 metric tons/m^2). It is frequently the wet-snow avalanche which damages the soil and vegetation cover.

Release mechanism. In the case of the loose avalanche, release mechanisms are primarily controlled by the angle of repose, while slab releases involve complex strength-stress problems. A release may occur simply as a result of the overloading of a slope during a single snowstorm and involve only snow which accumulated during that specific storm, or it may result from a sequence of meteorological events and involve snow layers comprising numerous precipitation episodes. In the latter case, large avalanches may not necessarily be restricted to storms with large amounts of precipitation, but can result from lesser amounts of precipitation falling on older snow layers underlain by an extremely weak structure.

Defense methods. Where snow avalanches constitute a hazard, that is, where they directly threaten human activities, various defense methods have evolved. Attempts are made to prevent the avalanche from occurring by artificial supporting structures or reforestation in the zone of origin. The direct impact of an avalanche can be avoided by construction of diversion structures, dams, sheds, or tunnels. Hazardous zones may be temporarily evacuated while avalanches are released artificially, most commonly by explosives. Finally, attempts are made to predict the occurrence of avalanches by studying relationships between meteorological and snow cover factors. In locations where development has yet to occur, zones of known or expected avalanche activity can be mapped, allowing planners to avoid such areas entirely. Avalanche hazard is small when compared with certain other natural hazards such as floods and tornadoes, but it continues to rise as the popularity of wintertime mountain recreation increases.

Richard L. Armstrong

Bibliography. *Annals of Glaciology*, vol. 26: *Papers from the International Symposium on Snow and Avalanches, Chamonix, France, 26–30 May 1997*, 1998; B. R. Armstrong and K. Williams, *The Avalanche Book*, rev. ed., 1992; S. C. Colbeck, *Dynamics of Snow and Ice Masses*, 1980; D. McClung and P. Schaerer, *The Avalanche Handbook*, 2d ed., 1993.

Aves

The class of animals consisting of the birds. Modern birds are characterized by being feathered, warm-blooded (homeothermic), and bipedal (two-legged) with the forelimb modified into a wing which, together with the tail feathers attached to the short tail, forms the flight mechanism, and by having a very high metabolic rate. Such a characterization, however, as with any group of vertebrates, holds for the living forms and most fossil members of this class, but is blurred by the early fossil record, which contains species with characteristics closer to those of the reptilian ancestors of birds. The feathers of birds are lightweight modifications of the outer skin possessing remarkable aerodynamic qualities. They serve not only as surfaces to generate lift and thrust, and as a streamlined outer surface of the body, but also as insulation to maintain high body temperatures. In addition, birds have lightweight, hollow bones; a well-developed air-sac system and flow-through lungs; a wishbone or furcula (fused clavicles); and a hand reduced to three digits (comparable to digits 2, 3, and 4 of the human hand). Birds have most likely evolved from an ancestor within the large group of ancient diapsid reptiles known as archosaurs (including alligators, snakes and lizards, and dinosaurs, among others). However, debate still centers on whether

birds are derived from a basal archosaurian stock or arose later directly from the later and more derived theropod dinosaurs (carnivores such as *Allosaurus* and *Velociraptor*). The latter theory gained support by the discovery of fossils from the end of the Mesozoic Era ("Age of Reptiles") that some researchers reported to be feathered dinosaurs; however, other researchers think they were birds that had already became secondarily flightless and represent "Mesozoic kiwis."

Feathers. Feathers are unique to birds. These lightweight structures made of keratin are the most complex appendages produced by the skin of any vertebrate. The vanes of a feather are supported by a central shaft, or rachis, which is the structural backbone. They are composed of specialized filaments called barbs, which possess secondary, tiny filaments called barbules, or barbs. The barbs are bound together in "Velcro" fashion by small hooklets, or hamuli, located on the barbs. Body-contour feathers of many species have an associated smaller secondary shaft—an aftershaft—which results from the developmental process of the feather and which adds to the insulating property of the plumage. Feathers grow from specialized dermal follicles and are generally renewed once a year during the postbreeding molt. Body feathers of some birds are molted twice a year; the large wing feathers of some raptors are not replaced every year. In addition to the typical vaned form, feathers come in many other forms, such as down, power down, bristles, and filoplumes. However, the tuft of epidermal outgrowths, the beard present on the breast of male turkeys, are not feathers.

Body feathers in most birds are arranged in definite feather tracts, leaving large parts of the skin free of feathers. Muscles attaching to the base of the feather, the calamus, move the feathers, thereby allowing the bird, for example, to fluff the body feathers into a loose ball around its entire body for maximum insulation or to move courtship plumes in precise and complicated ways.

The wing feathers, the remiges, produce lift and forward thrust in flight. They are asymmetric, with a smaller outer vane and a larger inner vane, which are the building blocks of slotted wings. The bird wing comprises two main sets of flight feathers, the outer primary feathers which are attached to the hand, and the inner secondary feathers which are attached to the ulna. The several feathers attached to the moveable second finger constitute the ulna which can form a slot to prevent stalling at low speeds, similar to those in airplane wings.

Contour feathers provide smooth aerodynamic surfaces of the body, resulting in laminar airflow during flight. Typically, the body contour feathers have symmetrical vanes. Most of the vane is stiff and tightly bound, like flight feathers, and is known as the pennaceous portion. However, the part of the vane near the base, known as the plumaceous portion, lacks hooklets and is loosely bound. This basal portion can by fluffed up to trap body heat next to the skin. In warm conditions the body feathers can be flattened to allow heat to escape. Thus, feathers form an insulating plumage to cover the surface of the avian body.

Tail feathers (rectrices) resemble the flight feathers of the wing and provide lift in flight. The tail feathers of modern birds are attached to a vertically flattend bone known as the pygostyle, formed by a number of fused caudal vertebrae. The pygostyle (sometimes called the plowshare bone) also supports the uropygial or oil gland that provides a rich oil used to preen the feathers and to maintain their moistness and flexibility.

Not all feathers have stiff vanes. Typically, newly hatched birds are covered by a coat of plush down feathers, which are replaced by the adult contour feathers. Down feathers lack a central shaft but have long, loosely connected barbules that provide an insulating layer next to the skin.

In addition to flight and insulation, feathers can serve other functions, ranging from providing color patterns and structural forms serving for species recognition and courtship displays (including sound production by wing or tail feathers) to cryptic color patterns for protection (as is the case with the woodcock, which has feathers that blend with the dead leaves of the forest floor). Another unusual adaptation of feathers is seen in the sand grouse, a desert bird of Africa and Asia that makes long flights to a water hole to drink. The male sand grouse also obtains water for the young, storing it in his specialized flattened and coiled barbules on the contour feathers of the belly. This feature allows the male to hold water in flight and transport it to the nest, which may be some 20 mi (30 km) away. The young drink by squeezing the wet feathers in their bills. *See* FEATHER.

Beak. The original toothed jaws of ancestral birds have evolved into the light toothless beak in which the upper and lower jaws are covered by a horny rhamphotheca, and which may vary in texture from the strong beaks of predatory raptors and seed-eating parrots and finches to relatively soft beaks of shorebirds and ducks. Beaks have a great variety of adaptive forms, including the flesh-tearing hooked beaks of hawks and eagles, the filter-feeding straining beaks of flamingos and ducks, the fish-trapping beaks of pelicans, the climbing and nut-cracking beaks of parrots, the hammering beaks of woodpeckers, and the seed-eating beaks of finches. As in most vertebrates (many fish and reptiles), in addition to the usual moveable lower jaw, the upper jaw of all birds is kinetic, that is, moveable with respect to the brain case, which enables many important functional properties of the avian feeding apparatus in contrast to the akinetic jaw apparatus of mammals. The structure of the tongue also varies greatly depending on the diet of the species.

Birds have developed a muscular gizzard (also found in their relatives, crocodiles and dinosaurs) for grinding and processing food into small pieces. The grinding is often assisted by the addition of ingested gizzard stones. The ancient giant moas of New Zealand are preserved with masses of stones in the

stomach region, and are known to have subsisted on a diet of leaves.

Wing. Birds are characterized as flying vertebrates with the forelimb modified as a wing with strong flight feathers. They also possess a large, strong pectoral girdle and bony sternum for the attachment of the large flight muscles and for the transfer of the body weight to the wings when the bird is flying. Considerable variation exists in the morphology of the wing (size and shape of the aerodynamic surface formed by the flight feathers) and of the entire muscle-bone system of the pectoral girdle and limb. But it is a mistake to conclude that all activities of birds are associated with flight; indeed for most birds, flight occupies a minor part of their daily activities. The important biological roles of flight in most birds are associated with quick escape from predators, reaching secure nesting places, larger separation of the nesting and foraging sites, and long-distance migration. Yet a few birds (such as swifts and terns) are able to spend many days or even weeks in the air, apparently being able to sleep during flight. And some birds (such as waterfowl and shorebirds) are able to make very long distance migration flights of up to 3000–4000 km (1800–2400 mi).

Hind limb. The leg and foot structures of birds differ greatly depending on their nonflying locomotion: from the long legs of wading birds to the strong, shorter legs of predators and aquatic birds to the almost rudimentary legs of the swifts and hummingbirds. The most primitive avian foot, found in the earliest definitively known bird, *Archaeopteryx*, is the anisodactyl perching arrangement of toes, also found in most modern tree-dwelling birds. Three toes point forward, and a reversed first toe, or hallux, opposes them in perching on a branch.

Other modifications include the zygodactyl feet (fourth toe reversed) of woodpeckers, cuckoos, and parrots, and the heterodactyl foot (second toe reversed) of the trogons with two forward and two rearward-pointing toes. The webbed feet of ducks and many other swimming birds have three forward-pointing toes united by a web that serves as a paddle. Members of the order Pelecaniformes have a foot in which all four toes are united by webbing, a totipalmate foot. Ostriches are unique among living birds in having a foot with only two toes.

Other distinctive features. The avian neck is long and flexible, consisting of many vertebrae with unique saddle-shaped articulating surfaces, allowing the bird to reach most parts of its plumage with its beak for preening. Modern flying birds have a well-developed pectoral girdle including a large, bony sternum with a keel, or carina, for the attachment of the large flight musculature. The pelvic girdle is large with the bones fused together with the synsacrum (a number of completely fused sacral vertebrae) to provide firm support for the bipedal hind leg. The ankle and foot bones are fused and elongated, so that the avian leg consists of three long segments—the femur, the tibiotarsus plus the reduced fibula, and the tarsometatarsus (the fused ankle and foot bones, which is equivalent to the human foot)—and finally the toes. Thus, birds walk on their toes, which are four in most birds but have been reduced to three, with the loss (usually) of the posterior hallux, or two (hallux and fourth) in the ostrich.

In many respects birds have more advanced sensory and physiological systems than mammals. They have a keen sense of vision (color vision based on four different cone cells, allowing color vision into the ultraviolet and fluorescent colors; the plumage of many birds looks very different to a bird than to a human) and hearing. The sense of smell (olfaction) is not particularly well developed, although some birds possess a good olfactory sense (for example, the turkey vulture, *Cathartes aura*). Birds have developed a small, rigid, flow-through lung using tubular air passages for a greater oxygen exchange surface than in mammals, and an extensive air-sac system serving as the pumping mechanism and as an internal system to lose excess heat. Nitrogenous wastes are lost in the form of uric acid, which requires much less water and a simpler kidney than in mammals.

Migration. Birds are found over the entire Earth. One of the most intriguing aspects of avian biology is the ability to migrate exceptional distances. Birds possess highly specialized directional senses for orientation, navigation, homing, and migration, including the ability to detect the Earth's magnetic field. These uncanny abilities permit birds to occupy widely separated wintering and nesting grounds, thus expanding their usable habitats. Some migrations, such as that of the Arctic tern, involve a circum-Atlantic route from Alaska to the South Pole. *See* FLIGHT; MIGRATORY BEHAVIOR.

Taxonomy. There are about 10,000 species of birds living today, of which more than 5500 belong to the order Passeriformes. Many avian species are particularly well known; however, the relationships of the higher categories of birds (orders, etc.) are still debated. Recognition and arrangement of avian orders and of their contained families has traditionally been based on study of morphological, and to a much lesser extent behavioral, features. Over the last three decades, much new information has been added using molecular features and especially with comparison of nucleotide sequences of nuclear and mitochondrial DNA. The number of avian orders, and the contained families and their arrangement are still controversial; even today, the relationships of these orders are almost unknown. Many authors advocate different arrangements. Because the situation is in flux, a fairly conservative system is used below (fossil groups are designated by a dagger). *See* NEOGNATHAE; NEORNITHES; VERTEBRATA; articles on the different avian orders.

Class Aves
- Subclass Sauriurae†
  - Infraclass Archaeornithes†
    - Order Archaeopterygiformes (Late Jurassic reptile-birds, *Archaeopteryx*)†
    - Confuciusornithiformes (Lower Cretaceous, beaked reptile-birds, *Confuciusornis*)†

Infraclass Enantiornithes (archaic Mesozoic land birds)†
Subclass Ornithuriae
Infraclass Neornithes (or Carinata)
Superorder Incertae Sedis
Order Hesperornithiformes (Cretaceous toothed divers, *Hesperornis*)†
Ichthyornithiformes (gull-like, Mesozoic toothed birds, *Ichthyornis*)†
Order Incertae Sedis (*Gansus*, *Chaoyangia*, etc., archaic modern-type birds)†
Superorder Neognathae (modern birds)
Order Struthioniformes (58 species)
Galliformes (chickens and allies, 282)
Anseriformes (waterfowl, 161)
Gastornithiformes (giant ground-birds, 2)†
Sphenisciformes (penguins, 17 species)
Procellariiformes (tube-nose seabirds, 114)
Pelecaniformes (pelicans and allies, 66)
Ciconiiformes (storks and allies, 119)
Falconiformes (hawks, eagles, and vultures, 309)
Gruiformes (rails, cranes and allies, 214)
Podicipediformes (grebes, 22)
Charadriiformes (shorebirds, gulls and allies, 349)
Phoenicopteriformes (flamingos, 5)
Gaviiformes (loons, 5)
Columbiformes (pigeons and doves, 332)
Psittaciformes (parrots, 360)
Coliiformes (mousebirds, 6)
Cuculiformes (cuckoos, 165)
Opisthocomiformes (hoatzin, 1)
Strigiformes (owls, barn owls, 173)
Caprimulgiformes (nightjars and allies, 116)
Apodiformes (hummingbirds and swifts, 425)
Trogoniformes (trogons, quetzals, 39)
Coraciiformes (kingfishers, bee-eaters and allies, 219)
Piciformes (woodpeckers and allies, 407)
Passeriformes (perching birds, songbirds, 5739)

Fossil record. The classification system presented above agrees with those presented in most major treatises on birds. The subclass Sauriurae contains the archaic birds of the Mesozoic Era, the Age of Reptiles, which includes the toothed fossil *Archaeopteryx*, or Urvogel, the oldest known bird (about 145 million years ago). Sauriurae also contains the subsequently discovered *Confuciusornis*, a beaked bird from Chinese deposits probably of Lower Cretaceous age, but possibly as recent as 120 million years ago. The first recognized specimen of *Archaeopteryx* was discovered in 1861 from the Solnhofen lithographic limestone of Late Jurassic age in Bavaria, and was named *A. lithographica* in reference to the rocks in which it was found. The crow-sized specimen from the bottom of a shallow salt-water lagoon was preserved with its wing and tail feathers in place, illustrating that it was indeed a bird and not a reptile. Another specimen was found in 1876 and became known as the Berlin specimen, as it is housed there. Often called an avian "Rosetta Stone," the Berlin specimen is beautifully preserved, with outstretched wings, and is often cited as the best example of an animal perfectly intermediate between two classes of vertebrates, in this case reptiles and birds. It fulfilled Darwins expectation that such forms should exist, and this specimen has played a large role in the debate on evolution over the years. The jaws of *Archaeopteryx* possess rounded teeth with large root crowns, unlike the flattened, recurved, serrated teeth of theropod dinosaurs. The wing feathers are typical of modern birds, showing that feathers remained essentially unchanged structurally in the 150 million years since *Archaeopteryx* lived. As in modern birds, the wing feathers are divided into primary and secondary flight feathers, which exhibit asymmetric vanes, indicative of aerodynamic function, probably gliding because of the overall weakness of the pectoral girdle suggesting small flight muscles. The wings end in three sharply clawed fingers, the claws being virtually identical to the pedal claws of tree-trunk-climbing birds such as woodpeckers. Because the curvature of the foot claws of *Archaeopteryx* is very similar to that of perching birds, it has been assumed that *Archaeopteryx* was a tree-dweller that used its clawed hands and feet to ascend tree trunks; its flying ability was mainly gliding, with some very limited flapping flight. There are now seven skeletal specimens of *Archaeopteryx*. *See* ARCHAEOPTERYX; ARCHAEORNITHES; REPTILIA.

A controversial fossil, *Protoavis texensis*, was described from the Late Jurassic of Texas from very fragmentary, crushed material. It does have clearly cervical vertebrae with the typical heterocelous (saddle-shaped) articulating surfaces that have been considered to be unique to birds. If this fossil is confirmed as an ancestral bird, it would push the known record of birds back nearly 100 million years and the origin of birds to the earliest members of the archosaurian radiation. But more and better material is needed before the relationships of this fossil can be determined with any assurance.

Confuciusornis sanctus, discovered in 1994 in China, is not as old as *Archaeopteryx* and could be as recent as 130–120 million years, but it shares many features with the Urvogel, including a wing with three sharply clawed fingers. It is clearly a perching bird like *Archaeopteryx* but has better-developed flight architecture and exhibits a typical avian beak without teeth. Many hundreds of specimens of

C. sanctus have been found, some with an elongated pair of narrow central tail feathers and presumed to be males.

During the past 3 decades a large diversity of Cretaceous fossil birds have been discovered in China, including some controversial fossils from the same deposits that produced *Confuciusornis*; these include two feathered but flightless creatures named *Protarchaeopteryx* and *Caudipteryx*. They are proclaimed to be feathered dinosaurs by Chinese and some western scientists, and said to prove the dinosaurian ancestry of birds; however, others believe that these are merely flightless Mesozoic birds and have little to do with avian origins. Despite their superficial resemblance to small dinosaurs, they show many of the features associated with flightlessness in modern birds. Other birds had become flightless early in avian history such as *Hesperornis* (Cretaceous) and *Diatryma* (Paleocene).

The Sauriurae contains a group known as the enantiornithines, or opposite birds. These archaic land birds of the Mesozoic are now known throughout the world from the Cretaceous Period, and are called "opposite birds" because their foot bones fuse in the opposite direction to that of modern birds. They also have a different type of formation of the triosseal canal in the pectoral girdle which is produced by the bones associated with the flight apparatus and through which passes the tendon of the elevator muscle of the wing. These archaic birds are particularly well known from the Lower Cretaceous of Spain and China. All opposite birds became extinct at the end of the Cretaceous Period.

Other Mesozoic birds included the ancient ornithurine birds more closely allied with the modern radiation of birds, among them such forms as the hesperornithiforms (*Hesperornis*, *Baptornis*, etc.), the Cretaceous toothed divers, which superficially resembled loons. They became extinct at the end of the Cretaceous along with their gull-like contemporaries, *Ichthyornis* and *Apatornis*. The Lower Cretaceous *Ambiortus* from Mongolia was a fully volant (capable of flight) ornithurine bird about the size of a pigeon, possessing a well-developed sternal keel and other features of the pectoral region typical of modern birds, which showed that true flying birds existed some 12 million years after the appearance of *Archaeopteryx*.

The Struthioniformes contain the living ratites—flightless birds such as the ostrich, rhea, emu, cassowary, and kiwi—as well as their South American chickenlike relatives, the tinamous, which are fully capable of flight. Recently, a group of fossil birds, the Lithornithidae, has been described from the Paleocene and Eocene of North America and Europe. These chickenlike forms were fully volant, and are closely related to the living tinamous and the ratites. They now are thought to be the ancestral stock that gave rise to the Struthioniformes. The more reasonable explanation for the disjunct distribution of living ratites is that their ancestors, the lithornithids, flew to remote parts of the world and gave rise to various lineages of flightless birds. The huge elephant birds lived on Madagascar contemporaneously with the native peoples, and probably became extinct in historic times. The same was true of New Zealand, where 11 or 12 species of large moas lived. They too survived until the arrival of Polynesians in the late thirteenth century and subsequently became extinct years ago. *See* RATITES; STRUTHIONIFORMES.

By the Eocene, approximately 50 million years ago, all the major orders of modern birds were present. By the Oligocene, most of the families were present, and by the Miocene, some genera of modern birds were well established.

Economic significance. Birds have a huge economic importance in terms of domesticated species, such as chickens and turkeys, and hunting. Today, however, the economics of birds for entertainment, such as birdwatching, ecotourism, and simply backyard feeding and watching, is far more important than for hunting. The interest of humans in observing birds is perhaps the major driving force in conservation efforts for the past century.

Walter J. Bock

Bibliography. W. J. Bock, Aves, in S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, vol. 2, 1982; J. Cracraft et al., Phylogenetic relationships among modern birds (Neornithes), pp. 468–489, in J. Cracraft and M. J. Donoghue (eds.), *Assembling the Tree of Life*, Oxford University Press, 2004; E. C. Dickinson (ed.), *The Howard and Moore Complete Checklist of the Birds of the World*, 3d ed., 2003; J. del Hoyo et al. (eds.), *Handbook of Birds of the World*, vols. 1–9, 1992–2004; A. Feduccia, *The Origin and Evolution of Birds*, 2d ed., 1999; F. B. Gill, *Ornithology*, 2d ed., 1995; D. P. Mindell (ed.), *Avian Molecular Evolution and Systematics*, Academic Press, 1997; N. S. Proctor and P. J. Lynch, *Manual of Ornithology: Avian Structure and Function*, 1993; C. G. Sibley and J. E. Ahlquist, *Phylogeny and Classification of Birds: A Study in Molecular Evolution*, 1990; C. G. Sibley and B. L. Monroe, *Distribution and Taxonomy of Birds of the World*, 1990.

Avian leukosis

A complex of several related and unrelated viruses (both C-type retroviruses and herpesviruses) that are collectively responsible for a variety of benign and malignant neoplasms in chickens and, to a lesser extent, in other avian species. Although most neoplasms observed in avian species are induced by viruses, there are some that are of unknown etiology.

Viruses. The neoplastic diseases induced by the leukosis-sarcoma group of retroviruses include lymphoid leukosis, myeloid or erythroid leukemias or solid tumors, tumors of connective tissue origin (for example, sarcomas, fibromas, and chondromas), epithelial carcinomas, and endothelial tumors. The many viral strains involved have similar physical and chemical characteristics and share a group-specific antigen; some can cause more than one type of neoplasm. The viruses are about 100 nanometers in diameter; have a core composed of ribonucleic

acid; contain a reverse transcriptase; mature by budding from the cell membrane; and are divided into subgroups based on envelope glycoproteins. Some strains carry their own specific oncogenes that induce neoplasms within days or weeks. Others lack an oncogene and cause neoplasms less frequently and only after several months, probably by activating a specific cellular oncogene. *See* CANCER (MEDICINE); ONCOGENES; ONCOLOGY.

Diseases. Lymphoid leukosis is the most important of the leukosis sarcoma diseases. The lymphoid leukosis virus is transmitted vertically from hen to chick through the egg. Infection can result in leukotic neoplasms in various visceral organs following metastasis from primary tumors in the bursa of Fabricius. Horizontal transmission from bird to bird is inefficient, rarely causing tumors unless infection occurs within the first week. Large-scale transmission of the lymphoid leukosis virus can be eradicated by eliminating individual infected breeders.

Reticuloendotheliosis virus strains constitute another retrovirus group, unrelated to the lymphoid-sarcoma group, and also may carry a specific oncogene. They can cause a chronic neoplastic form of reticuloendotheliosis or other neoplasms in turkeys, chickens, ducks, geese, quail, and pheasants, and a runting disease has been seen in chickens after accidental contamination of vaccines with reticuloendotheliosis virus.

Marek's disease in chickens is caused by an oncogenic, cell-associated, lymphotrophic, highly contagious herpesvirus. Although the virus is biologically like lymphotrophic gamma herpesviruses, the structure of its genome is similar to alpha herpesviruses. The virus replicates fully only in the epithelial cells surrounding the feather shaft, and is released to the environment associated with dead cells and dander. Inhalation of the virus causes an active infection in lymphoid organs; after about 1 week, a latent infection develops in lymphocytes. T-cell lymphomas may develop within a few weeks or months, depending on age, genetic makeup, virus virulence, and other factors. Degenerative, inflammatory and lymphoproliferative lesions occur principally in the peripheral nerves (causing paralysis), lymphoid tissues, visceral organs, muscle, and skin. Eye involvement (gray eye) can cause blindness. The disease is of great economic importance in chickens, and several vaccines, injected at 1 day of age, have been in worldwide use since about 1970. The vaccines are prepared from attenuated strains of Marek's disease virus or of related, but naturally nononcogenic, herpes viruses isolated from turkeys or chickens. *See* ANIMAL VIRUS; TUMOR VIRUSES.

Bruce W. Calnek

Bibliography. B. W. Calnek et al. (eds.), *Diseases of Poultry* 10th ed., 1997.

Aviation

A general term including the science and technology of flight through the air. Aviation also applies to the mode of travel provided by aircraft as carriers of passengers and cargo, and as such is part of the total transportation system. Aviation also describes the employment of aircraft in such fields as military aviation. The world of the airplane, including the people who manufacture, market, and repair aircraft or who work in allied industries, is frequently spoken of as aviation. *See* AIRPLANE; MILITARY AIRCRAFT.

Aviation in modern usage differs from the gliding or soaring of unpowered aircraft, such as a glider or sailplane, and the flight of lighter-than-air craft, such as a balloon. The pilot of heavier-than-air craft is an aviator, whereas the pilot of a lighter-than-air craft is an aeronaut. *See* BALLOON; GLIDER.

Aerology is the branch of meteorology concerned with the study of free air, that portion of the atmosphere which is undisturbed by objects on the earth, such as mountains or buildings. Because aircraft move through free air, aerology is of great importance to aviation. A person professionally competent in the science of aerology is an aerologist. *See* AERONAUTICAL METEOROLOGY.

Aviation is broadly grouped into three classes: general aviation, air transport aviation, and military aviation. General aviation comprises all aviation not included in military or air-transport aviation. Military aviation includes all forms of aviation in military activities, and air-transport aviation is primarily the operation of commercial airlines essentially as a public utility for the movement of persons and commodities. *See* AIR TRANSPORTATION; GENERAL AVIATION.

Leslie A. Bryan

Bibliography. M. J. Taylor (ed.), *Jane's Encyclopedia of Aviation*, 1993.

Aviation security

A multilayered, multidimensional means to safeguard passengers, crew, ground personnel, and the general public against acts of unlawful interference perpetrated in flight or within the confines of the airport. It specifically refers to the techniques and methods used in protecting aviation from crime or terrorism.

Preventive Measures

Preventive measures are techniques and methods that are in place to prevent potential threats (such as terrorists, criminals, or explosive materials) from getting on the aircraft. All carry-on baggage is screened by explosive detection equipment or "other appropriate means of detection," which could include trace-chemical detection equipment, manual searches, and bomb-sniffing dogs. The high cost of explosive detection machines may account for the approval by the Transportation Security Administration (TSA) of the use of alternative detection systems. The screening of both people and carry-on luggage remains a controversial issue with security breaches occurring and travelers subjected to different rules from airport to airport. Changes being considered include replacing federal screeners with private screeners at some airports; implementing programs that use biometric systems to screen out potential terrorists; employing new explosion detection, trace-chemical detection, and bomb detection

systems; increasing video surveillance; and installing more stringent cargo security systems that track airport workers in luggage and loading areas. Many improvements have been made in aviation security in regards to screening passengers and cargo for potential hazards. Various technological advances have been made to develop more rigorous explosive detection systems with highly developed surveillance systems.

Biometric systems. Biometric systems are designed to recognize biological features to facilitate identity verification. The individual whose identity is being verified places his or her finger, hand, retina or face onto or near a scanner and provides data which are compared with a database. The systems can be based on fingerprints (optical scanning of a finger), signature recognition (including measurement of the motion and pressure used in writing the signature), hand geometry (the physical attributes of the hand, such as the length of the figures), speaker verification (utilizing the uniqueness of voice patterns), or the blood vessel pattern of the retina.

Explosive detection systems. An explosive detection machine operates and looks like a medical computerized axial tomography (CAT) scanner. The system first produces an x-ray scan similar to the conventional airport x-ray scanner. An automated inspection algorithm determines the locations within the baggage where the absorption indicates a suspicious area; cross- sectional computerized tomography (CT) slices then determine size, shape, mass density, and texture of any suspect object. If no high-density areas are detected, a single slice through the bag is made to look for any explosives that may not have been seen in the projection scan. After entering the CT chamber, a bag is divided virtually into three-dimensional units called "voxels." Based on their similarities, certain voxels are grouped together as volumes of the same objects within a bag, so that the density of the scanned object can be determined. Based on density and volume, software automatically correlates the mass characteristics of luggage contents to those of potential explosives. If the system finds a match, it alerts the operator by highlighting suspect areas within the CT slice. Since the CT scan produces true cross-sectional slices, it is able to identify objects that are surrounded by other materials or hidden by innocuous objects. Three-dimensional rendering may also be applied. The machine is programmed to recognize numerous chemical compositions. *See* COMPUTERIZED TOMOGRAPHY.

Another type of explosive detection system uses a document scanner to analyze samples collected by swiping the surface of a document over a collection disk and alerts the screener if an explosive's residue is detected. During the pilot program, passengers selected for secondary screening at particular checkpoints had their boarding passes scanned. If the document scanner raised an alarm, additional screening procedures were implemented.

Trace-chemical detection. Trace-chemical detection equipment can analyze a swipe or air sample, detecting and identifying minute traces of explosive substances or biohazards. The detection process takes 4 s to collect the trace particles and another 8 s to analyze them. As the person being inspected stands in the center of an archway, gradually stronger puffs of air come from four surrounding columns, positioned to direct the air from the lower to the upper parts of the body, accelerating the plume at a faster rate than it would naturally rise. The plume is collected in an overhead detector hood, and the collected particles are vaporized. The molecules are either positively or negatively charged, and the resulting ions are pulsed down a drift tube. The equipment measures how fast the ions travel from point to point. This acts as a thumbprint of the substance, since each specific type of ion has its own particular travel time. This enables the machine to identify a broad range of organic matter, including explosives and materials associated with chemical and biological weapons.

Passengers identified as needing additional screening will pass through the trace portals. Some equipment can access the human convection plume to collect any threatening particles. If a person has explosives strapped to their body or has even handled explosives, trace particles will contaminate clothing and register. A computerized voice indicates when a passenger may exit the portal. Screeners will take necessary and appropriate steps to resolve alarms.

Bomb-detection systems. A new type of bomb-detection system being considered is a terahertz radiation (t-ray) device. Visible light provides the naked eye with a very narrow slice of the electromagnetic spectrum. This technology is destined to reveal not only the shape of objects hidden but also their composition. Because different chemical structures absorb them differently, t-rays could be utilized to identify hidden materials, including explosives and dangerous objects. A t-ray machine looks much like a copying machine. An object is placed on the imaging window, the beam passes across it, and a detector measures the transmitted rays. A screen displays the image and a separate probe arm can scan objects that do not fit neatly into the image window.

Neutron technology is being researched to pinpoint chemicals. Pulsed fast neutron analysis (PFNA) technology is based on the detection of signature radiation (gamma rays) induced in material scanned by a beam of neutrons. The PFNA system has the potential to meet TSA goals as it has demonstrated some superior characteristics compared to existing x-ray systems in detecting explosives in cargo containers and passenger baggage. The PFNA system currently requires a long scan time (an average of 90 min per container), needs considerable radiation shielding, is significantly larger than current x-ray systems, and has high implementation costs. These factors are likely to limit installation at transportation facilities, even if the detection capability is improved. Nevertheless, PFNA currently has the good potential for detecting explosives in cargo as well as vehicles approaching a facility.

Video surveillance. New wireless video surveillance systems will enable separate security operations centers to simultaneously monitor distant sites, including secured airport areas, public parking lots,

and roadway tunnels. These systems provide a single wireless solution for viewing, storing, and managing real-time video.

Cargo security. The cargo hold remains the most vulnerable part of the aircraft as most cargo is loaded unscreened. In addition, many problems revolve around the shipment of known and unknown cargo. Known cargo is generically defined as cargo presented for carriage by a known freight forwarder who has been a customer for over a year.

Air cargo security programs are in use in Europe, which has a complete air cargo security program. Air cargo agents must have an approved air cargo security program, and regulated agents conduct security investigations not only of the delivery component (usually trucks) but also of the manufacturer. The entire supply chain has a security element authorized by law.

Defensive Measures

Defensive measures are systems and devices that are in place to defend against threats (such as terrorists, criminals, or explosive materials) that are on the aircraft or can attack the aircraft in flight. Many advances have been made in the development of sophisticated defense systems and devices. These advances are providing progressively safer service worldwide for passengers traveling by air transportation. *See* AIR TRANSPORTATION.

Cockpit doors. All U.S.-based passenger airplanes are required to install reinforced cockpit doors to prevent intruders from gaining access to the flight deck. Currently, reinforced and ballistic-resistant doors, composed of a variety of metals and composites, are designed to withstand extreme pounding and a hail of bullets, as well as heat and smoke. Cargo airlines are granted an exception if they file a security plan. *See* COMPOSITE MATERIAL.

Active Denial System (ADS). The Active Denial System is a nonlethal energy weapon that is designed to provide stand-off nonlethal capabilities at ranges beyond the effectiveness of small arms. The ADS uses a transmitter producing energy at a frequency of 950 GHz and an antenna to direct a focused, invisible beam towards a designated subject. The energy reaches the subject and penetrates the skin to a depth of less than 1/64 in. This energy beam induces an instantaneous heat sensation on the target's skin within seconds which is quite intolerable. The reaction is so intense that the targeted individual is repelled without injury. The tactical theory behind the weapon is to provide a means to stop an aggressor without causing injury. The sensation immediately stops when the individual steps out of the beam.

Containing blasts. Blast containment design concepts attempt to completely suppress the results of an explosion within a luggage container. Blast management design concepts consider the container as part of a placement system inside the cargo bay of the aircraft. Factors such as pressurization of the vehicle, load, and in the case of aircraft, aerodynamic effects can play crucial roles in the outcome of an explosion. The baggage can absorb a great deal of the energy of an explosion, lessening the shock waves, but the explosion may also generate a shower of projectiles entirely dependent on the specifics of the individual bags.

Tests have determined that the blast containment concept offers the best alternative for suppressing the potentially catastrophic effects of postblast fires. This system is an independent unit. It stands alone and needs no special handling or placement within the cargo bay. On the other hand, the blast management concept allows a container to essentially fail and bases its control on the ability to vent the detonation products into adjacent containers. The transporter must arrange the cargo appropriately for the system to have any usefulness. This could prove to be tedious and quite time-consuming. Consequently, early on it was decided to focus on a blast containment container which would be constructed of state-of-the-art high-strength composite materials with fragment-penetration-resistant and fire-retardant properties. Blast-resistant containers are only available for wide-body aircraft, which includes only 25% of the aircraft in service.

Fiber optic sensor systems. These systems have provided perimeter security for many types of facilities for many years. They are particularly useful in the protection of airports because of their stability, long-range applicability (optical fibers can carry light over distances of more than 160 km (or 100 mi), suitability for buried applications, and the fact that they are immune to lightning and other electrical interference. These systems are movement activated.

Antimissile defense. The concept of equipping the commercial carrier fleet with antimissile defenses presents an entirely different set of challenges. Officials are addressing the possibility that terrorists could use shoulder-fired missiles to bring down an airliner or attack other forms of transportation. Patrols of perimeters have increased, but facilities remain vulnerable.

An aircraft antimissile defense would be based on a Doppler radar system made up of four antennas at the front of the aircraft, two on the sides and four on the back. The antennas would be capable of giving 360° of radar coverage around the aircraft. Within seconds of a missile being detected, an on-board computer would release flares, firing at different angles to act as a diversion. The system would be completely automated, (i.e., there would be no involvement of the pilots). *See* DOPPLER RADAR.

Shoulder-fired missiles are currently being produced under license in several countries. Several thousand are in circulation and some are thought to have been acquired by terrorist groups. The threat is real but the feasibility of equipping aircraft with antimissile systems has a prohibitive cost in conjunction with significant technological challenges relating to the speed required to react and the fact that commercial planes are not as maneuverable as fighter military aircraft. However, an airborne defense system against shoulder-launched missiles aimed at civilian aircraft underwent a successful test at Palmahim, Israel in May 2004.

Kathleen M. Sweet

Bibliography. C. Combs, *Terrorism in the Twenty First Century*, 2d ed., Prentice Hall, Little Saddle River, NJ, 2000; N. M. Denkin, Fiber optics, World Book Online Reference Center, World Book, June 9, 2004; C. E. Simonson and J. R. Spindlove, *Terrorism Today, The Past, The Players, The Future*, Prentice Hall, Upper Saddle River, NJ, 2000; K. M. Sweet, *Terrorism and Airport Security*, Edwin Mellen Press, March 2002; K. M. Sweet, *Aviation and Airport Security: Safety and Terrorism Concerns*, Prentice Hall, Upper Saddle River, NJ, November 2003; K. M. Sweet, *Transportation Security: Threats and Solutions*, Prentice Hall, Upper Saddle River, NJ, August 2005.

Avocado

A tropical and subtropical fruit tree, *Persea americana*, in the Lauraceae family (containing also the laurel of antiquity and the trees that yield commercial cinnamon and camphor). It originated in Central America or adjoining regions of North or South America. It has now spread to much of the near-tropical world. *See* MAGNOLIALES.

Races. The species is divided into three horticultural races with differing commercial qualities. The so-called West Indian race is least tolerant of cold, the Mexican most tolerant, and the Guatemalan intermediate. This same gradation is found in salt tolerance (West Indian most) and oil content (West Indian lowest). But in some other respects the West Indian race is intermediate (skin thickness), or one of the races is different from the other two (the West Indian fruit is less tolerant of cold storage; the Mexican has smaller fruit with a unique anise-like odor; the Guatemalan has a smaller seed ratio, and takes twice as long to mature—14 months or more in California).

Cultivars. In warm and humid areas such as southeast Florida and all world lowlands near the Equator, only the West Indian race is well adapted; but West Indian × Guatemalan hybrids are doing well and becoming increasingly important in extending the harvest season. Cultivar (horticultural "variety") selection for these tropical regions has taken place chiefly in Florida. There the early season (pure West Indian) is dominated by Waldin and Pollock, with Simmonds increasing. Later-maturing cultivars, with an admixture of Guatemalan-race genes, include the Booth 8, first in production among all Florida cultivars and with increasing acreage, and the declining Booth 7. For the still later period, Lula has long predominated and still rivals Booth 8, but is being planted very little because of susceptibility to scab. Choquette and Monroe are two late cultivars of increasing importance. A dozen or so additional cultivars are also of some significance in Florida and other tropical areas.

Cultivar development for drier and colder-winter regions has occurred mostly in California. There, and in similar climates such as most of Mexico, Chile, South Africa, Israel, and Australia, Guatemalan and Guatemalan × Mexican hybrids are best adapted. Fuerte, containing about equal genetic contributions from the two races, was long the leading cultivar but is now declining in importance because of erratic production. Hass, chiefly of Guatemalan-race origin, is increasingly dominating California production, and its success in Mexico makes it easily the world's leading cultivar—in Michoacan State, the proportion planted (or top-worked) to Hass is estimated at a remarkable 90%. A heavy-setting cultivar with superb quality, Hass well deserves its preeminence (**Fig. 1**). Among hundreds of other selections named around the world, the following are most important—California: Zutano, largely Mexican; Bacon, intermediate in race traits, very cold-hardy; Reed, apparently pure Guatemalan; Pinkerton, intermediate, precocious, necked fruit, but very promising; Hawaii: Sharwil (from Australia), intermediate, plus several local selections; South Africa: Edranol (from California), mostly Guatemalan; Israel: Ettinger, mostly Mexican. Major breeding programs in Israel and California should soon produce superior cultivars.

Fig. 1. Hass, the world's leading avocado cultivar.

Culture and management. The avocado is limited especially by its climatic requirements, with their race differences. It is also highly susceptible to drought injury. But excess soil moisture is equally fatal, encouraging the dread Phytophthora root rot in most of the world, and where the fungus is unknown, as in Israel, promoting root asphyxiation. It is likewise unusually sensitive to salts in the root zone. The combined drought and salt susceptibilities mean that, over the 6 months or so without significant precipitation in California, regular irrigation is necessary. It may be needed weekly on light soils, especially when the available water is rather saline. The drip irrigation method is working well, especially with young trees.

Few California orchards need sprays for pest control—predators and parasites normally control such insects and mites as may develop. But in a few areas, during a few seasons, either of two caterpillar genera or thrips build up sufficient numbers to make spray control desirable. In other parts of the world, insects are more of a problem and require regular control measures.

All California orchards need nitrogen fertilizers and probably most need zinc. Laboratory analysis of leaves will reveal their nutritional status for these elements and others that are more rarely present in either deficiency or excess. The experienced grower can often identify nutritional problems from leaf symptoms. For other parts of the avocado world, the only safe guide is local observation and experiment.

Precocity and productivity. Precocity and productivity vary with both cultivar and location. The best California climate is more rigorous than is desirable for the avocado, both in terms of occasional freezes and, more seriously, in terms of average suboptimal temperatures for fruit set. Within California and other growing areas with comparable climate, a location providing good air drainage and southern exposure may well be sufficiently warmer to make the difference between success and failure of the enterprise.

One can expect a light crop set 2 years after planting Pinkerton, 3 years for Hass, 3 or 4 for Bacon, and 4 or 5 for Fuerte. Average production from mature trees is usually correlated with precocity; with spacing at about 7 m (24 ft) square, yields in kilograms per hectare (or pounds per acre) from a well-managed orchard in a good location will be perhaps 7000 for Fuerte in its limited zone of adaptation, and also for Bacon; perhaps 10,000 for Hass; perhaps 15,000 for Pinkerton—although it is not yet adequately tested. The Israeli cultivar Tova may even exceed the Pinkerton level, but Tova quality is lower. Both of these very heavy setters are about on a par with adapted cultivars in more favorable near-tropical regions. Planting arrangements to encourage cultivar cross-pollination will generally, and often strikingly, enhance fruit set because of the avacado's unusual flower behavior. This behavior makes self-pollination difficult by causing nearly all the open flowers of a cultivar, in any given location, to be functionally female one time of day and male a different time; this condition is called synchronous protogynous dichogamy. *See* POLLINATION.

Mexico is the world's leading producer, followed by Brazil and California, then Colombia and Venezuela, countries of eastern South America, Central America, Caribbean Islands, Florida, Philippines, and Zaire (central West Africa). South Africa and Israel have important export industries, primarily to Europe. Many other countries have begun development. The California industry has expanded rapidly, and the avocado has become one of the state's leading fruit crops.

Marketing. Seedling trees have fruits ranging from about 2 oz to 4 lb (50 to 2000 g), with most commercial markets preferring about 10 oz (280 g). However, some local markets, especially in more tropical areas, prefer a somewhat larger fruit. Shape varies from slightly oblate to highly elongate; an intermediate oval or pear shape is preferred. Mature-fruit skin color varies from greenish-yellow to purple or black, through different shades of green and even some red tints. Color preference is usually a reflection of what the consumer has become accustomed to.

Fig. 2. Avocado tree with Phytophthora root rot.

Marketing is made more difficult by the avocado's unusual flavor. Instead of being in the sweet-to-tart range of most fruits, the avocado flesh has a unique, subtle taste, due to its forming oils rather than carbohydrates. As a result, most consumers are not favorably impressed at their first exposure; only after several experiences does it usually become a prized food delicacy. Therefore, to effectively enter new markets, an extensive program of consumer education and advertising may be necessary. *See* FRUIT; FRUIT, TREE.

B. O. Bergh

Diseases. Phytophthora root rot (causal fungus *Phytophthora cinnamomi*) is the most serious and destructive disease of the avocado worldwide (**Fig. 2**). Resistant rootstocks, soil fungicides, and possibly biological control offer promise against this disease. Other relatively minor problems require action as follows. For Verticillium wilt, preplanting or interplanting avocado with other susceptible crops such as tomato should be avoided. For Armillaria root rot, trees should be removed and the soil fumigated with methyl bromide. Sunblotch, a viroidlike disease, is controlled by using certified seed and budwood sources. Phytophthora trunk cankers necessitate excision and painting affected areas with fungicide, and avoiding moisture accumulation around the trunk. Several leaf and fruit spots in humid tropical areas are caused by the fungi *Colletotrichum, Sphaceloma*, and *Cercospora*, and are controlled by spraying with fungicides. Black streak is a disease of unknown cause affecting Guatemalan varieties. *See* PLANT PATHOLOGY.

George A. Zentmyer

Bibliography. *California Avocado Society Yearbook*, annually; J. Janick and J. N. Moore (eds.), *Advances in Fruit Breeding*, 1975; *Proceedings of the*

1st International Tropical Fruit Short Course: The Avocado, 1977; G. A. Zentmyer et al., *Avocado Diseases*, Univ. Calif. Circ. 534, 1965; G. A. Zentmyer and H. D. Ohr, *Avocado Root Rot*, Univ. Calif. Leaf. 2440, 1978.

Avogadro number

The number of elementary entities in one mole of a substance. A mole is defined as an amount of a substance that contains as many elementary entities as there are atoms in exactly 12 g of ^{12}C; the elementary entities must be specified and may be atoms, molecules, ions, electrons, other particles, or specified groups of such particles. Experiments give 6.0221367×10^{23} as the value of the Avogadro number. In most calculations the coefficient is rounded off to 6.02. Thus, a mole of ^{12}C atoms has 6.02×10^{23} carbon atoms, a mole of water molecules contains 6.02×10^{23} H_2O molecules, a mole of electrons contains 6.02×10^{23} electrons, and so forth. *See* MOLE (CHEMISTRY).

Significance. The atomic weight (relative atomic mass) of ^{12}C is exactly 12, by definition. Consider 12 g of ^{12}C (which is one mole and contains the Avogadro number of atoms) compared with 4 g of He, whose atomic weight is 4. The 12 g to 4 g ratio of the masses of the two samples is the same as the 12 to 4 ratio of the masses of the atoms of ^{12}C and He. Therefore the two samples must contain the same number of atoms, and 4 g of He contains the Avogadro number of atoms. The same argument holds for any element. Thus, for an element with atomic weight x, a sample with mass x grams contains the Avogadro number of atoms. Similarly, for a substance with molecular weight y, a sample whose mass is y grams must contain the Avogadro number of molecules. For example, 18 g of water contains 6.02×10^{23} H_2O molecules. *See* ATOMIC MASS; RELATIVE ATOMIC MASS.

The Avogadro number is a dimensionless number. The Avogadro constant is defined as the Avogadro number divided by the unit "mole." The Avogadro constant is usually symbolized by N_A, N_0, or L. Since N_A gives the number of molecules per mole, $N_A = N/n$, where N is the number of molecules present in n moles of a substance.

The Avogadro number relates the mass of a mole of a substance to the mass of a single molecule. For example, for H_2O (whose molecular weight is 18) the mass of one mole is 18 g and the mass of one molecule is $(18\text{ g})/(6.02 \times 10^{23}) \approx 3 \times 10^{-23}$ g. The mass m of one molecule of a substance with molar mass M is $m = M/N_A$.

The Avogadro constant N_A is related to other fundamental physical constants. The Faraday constant F is the absolute value of the charge on one mole of electrons. Therefore $F = N_A e$, where e is the absolute value of the charge on one electron. Also, $R = N_A k$, where R is the gas constant and k is the Boltzmann constant. *See* BOLTZMANN CONSTANT; GAS CONSTANT.

Widespread use of the mole concept began only around 1900. The nineteenth-century concept most closely related to the Avogadro number is the number of molecules per unit volume in a gas at 0°C and 1 atm. [The ideal-gas law $PV = nRT = (N/N_A)RT$ gives $N/V = N_A P/RT$, so N/V, the number of gas molecules per unit volume, is proportional to the Avogadro constant N_A at fixed pressure P and temperature T.] Avogadro hypothesized in 1811 that at a fixed temperature and pressure the number of molecules per unit volume is the same for different gases, but he had no way of estimating this number. *See* AVOGADRO'S LAW.

Determination. The first estimate of the number of molecules per unit volume in a gas was given by J. Loschmidt in 1867. Using a kinetic-theory-of-gases equation for the gas viscosity and estimating NV_{molecule} (where V_{molecule} is the volume of a single molecule) from the volume of the liquid formed by condensing the gas, Loschmidt gave a value of N/V which is one-thirtieth the correct value. *See* KINETIC THEORY OF MATTER.

The number of molecules per cubic centimeter in an ideal gas at 0°C and 1 atm is called the Loschmidt number in English-speaking countries. (However, in Germany, the Loschmidt number means the number of molecules per mole.)

In 1900 M. Planck, using the hypothesis of energy quantization, derived the law for blackbody radiation. Planck's law contains the Boltzmann constant k and the Planck constant h. By fitting experimental data to his law, Planck found values for h and k, and from k (which equals R/N_A) he calculated the Avogadro number as 6.18×10^{23}. Using $F = N_A e$, he calculated $e = 4.69 \times 10^{-10}$ statC. Thus Planck had obtained the first reasonably accurate values of h, k, N_A, and e. *See* BOSE-EINSTEIN STATISTICS.

In 1906 A. Einstein used his molecular theory of diffusion and data for aqueous sugar solutions to estimate $N_A = 4 \times 10^{23}$ mol^{-1}; Einstein subsequently found an error in his work and recalculated this estimate as 6.56×10^{23} in 1911.

J. B. Perrin in 1908 observed the distribution in the Earth's gravitational field of colloidal gamboge particles of uniform size. According to the Boltzmann distribution law, the ratio n_2/n_1 of numbers of particles at heights h_2 and h_1 in the gravitational field is given by Eq. (1). The potential energies of the particles are

$$n_2/n_1 = \exp[-(E_2 - E_1)/kT] \qquad (1)$$

$E_1 = mgh_1$ and $E_2 = mgh_2$, where m is the mass of a particle (corrected to allow for the buoyancy of the suspending fluid) and g the gravitational acceleration. Observation of n_2/n_1 allows calculation of the Boltzmann constant k, and use of $R = N_A k$ then gives N_A. Perrin found $N_A = 7.0 \times 10^{23}$ mol^{-1}.

From 1908 to 1917, R. A. Millikan and H. Fletcher determined the electron charge by observing the motions of charged oil drops. Using their final result for e and the value of the Faraday constant known from electrolysis experiments, Millikan found $N_A = F/e = 6.06 \times 10^{23}$ mo1^{-1} in 1917. *See* ELECTRON.

A very accurate way to find N_A uses x-rays to measure the lattice spacing in a crystal. The crystal's density ρ equals the mass of a unit cell divided by the volume of a unit cell: $\rho = m_{cell}/V_{cell}$. The mass of one formula unit is M/N_A, where M is the molar mass; hence $m_{cell} = ZM/N_A$, where Z is the number of formula units per unit cell. If the unit cell is cubic, then $V_{cell} = a^3$, where a is the unit-cell edge length. The crystal density is then given by Eq. (2). Z is readily

$$\rho = MZ/a^3N_A \qquad (2)$$

found from x-ray diffraction study of the crystal. *See* X-RAY CRYSTALLOGRAPHY; X-RAY DIFFRACTION.

In the period 1930–1970, several determinations of N_A were done by using Eq. (2), with a a determined from the diffraction pattern produced by x-rays of known wavelength. The wavelength was determined by diffracting the x-rays with a grating having closely ruled lines of known spacing.

The accuracy of the x-ray crystal-density method was improved substantially by R. Deslattes and coworkers in 1974 by using a combination of x-ray and optical interferometry to determine the lattice spacing in a very pure silicon (Si) crystal. These workers also did very accurate measurements of the density and atomic weight of silicon. P. Becker and coworkers in 1981 repeated the interferometry measurement of the Si lattice spacing and obtained a slightly different result from that of Deslattes. When the Becker determination of a is combined with Deslattes group's determination of M and ρ, the value $N_A = 6.02213 \times 10^{23}\ \mathrm{mol}^{-1}$ is found. *See* FUNDAMENTAL CONSTANTS. Ira N. Levine

Bibliography. R. D. Deslattes, The Avogadro constant, *Annu. Rev. Phys. Chem.*, 31, 435–461, 1980; D. Kolb, The mole, *J. Chem. Educ.*, 55, 728–732, 1978.

Avogadro's law

The principle that equal volumes of all gases and vapors, under the same conditions of temperature and pressure, contain an identical number of molecules; also known as Avogadro's hypothesis. From Avogadro's law the converse follows that equal numbers of molecules of any gases under identical conditions occupy equal volumes. Therefore, under identical physical conditions the gram-molecular weights of all gases occupy equal volumes. Avogadro's law is not strictly obeyed by real gases at ordinary temperatures and pressures, although the deviations are only slight. At high pressures the deviations may be large. Avogadro's law can be shown to follow theoretically from the simple kinetic theory of gases. *See* GAS; KINETIC THEORY OF MATTER. Thomas C. Waddington

Axenic culture

The growth and maintenance of a single species in isolation, free from foreign or contaminating species. Isolation is usually achieved by growing the species in an environment that was previously sterilized, and was thereby rid of contaminating organisms. Since, from a practical viewpoint, the contaminating organisms usually encountered are microorganisms, axenic cultures, whether of invertebrates (for example, nematodes) or vertebrates (for example, rodents), are often referred to as germ-free. Indeed, the terms axenic and germ-free are occasionally used interchangeably. Gnotobiotic is also often used interchangeably with axenic; however, in common practice gnotobiotic specifically refers to germ-free conditions.

A principal goal of early studies with axenic cultures was the establishment of nutritional requirements for individual species. Historically, axenic cultures have also been employed to demonstrate that organisms growing in close association can have both direct and indirect effects on each other. Direct effects include competition for nutrients, attacks (for example, by a parasite), or production by one species of toxic compounds that affect a second species. Indirect effects include such phenomena as the production by intestinal microflora of vitamin K (germ-free animals therefore require it as a dietary supplement).

Technology. The two methodological challenges associated with developing axenic culture methods are formulating an appropriate culture medium and designing a functional habitat. Formulating a culture medium for small organisms, including microorganisms and many invertebrates, is the more difficult of the two challenges. For larger organisms, a culture medium is replaced by a carefully formulated diet, which is usually a relatively straightforward task. Initiating the axenic cycle and maintaining a sterile environment can, however, be a formidable endeavor for larger organisms.

For axenic culture of small organisms, such as bacteria, protozoa, and fungi, a typical strategy is to add to a minimal medium [which usually contains only a carbon source (such as glucose), nitrogen, sulfur, and phosphorus, as well as various salts and trace minerals] several supplements that are rich in a large number of growth factors. Such supplements often include water-soluble extracts of meat or yeast containing vitamins and other organic compounds. Once successful growth is achieved, it is usual to attempt to simplify the culture medium by deleting, in a stepwise fashion, each of the more esoteric ingredients (for example, individual vitamins).

Initiating the axenic cycle and designing an appropriate habitat for larger organisms require special effort. First, organisms free of contaminating species need to be obtained; in some cases, cesarean births are employed. Second, sterile containers that have entry and exit ports as well as a comfortable living environment must be constructed.

Methodology. Axenic culture has been achieved for a broad range of organisms. The method employed to remove contaminating organisms in a starter culture or for the initiation of an experiment depends upon the type of organism to be decontaminated and the nature of the contaminating organism.

Decontamination procedures for various types of organisms are summarized below.

Microorganisms. Most bacteria, fungi, and algae can be grown on a solid surface such as that provided by agar supplemented with a variety of nutrients. Initially, either single cells are handpicked and placed on the agar surface, or a collection of cells is first diluted to low density in an appropriate culture medium and then spread on the agar surface. After an appropriate culture period, those single cells will proliferate into a small mass of cells that represents an axenic culture.

Plants. Many types of plant cells can be cultured by using methods similar to those for microorganisms. Axenic culture of whole plants begins with sterilization of the starting material. For example, seeds or small whole plants can be washed with a sterilizing agent, such as alcohol, and then deposited on an agar surface containing several nutrients and a variety of antibiotics. Similarly, some hardy, small whole plants can be sterilized with alcohol or mercuric chloride solution and grown in appropriate liquid or solid media.

Protozoa. Since protozoa are often isolated from water sources laden with particulate soil components, grasses, and other bacteria-infested particles, washing by centrifugation is often employed to provide starting material for culturing. Protozoa are separated from debris and bacteria. Alternatively, a single cell can be isolated with a microscope and transferred to an appropriate culture medium. In both cases it is usually advisable to permit the isolated protozoa to incubate in culture medium for several hours before the final transfer to axenic conditions is made. This permits protozoa, which normally ingest bacteria, to clear any bacteria that they harbor in vacuoles or that cling to their surface structures.

Nematodes. The free-living nematode *Caenorhabditis* has been employed extensively as an experimental system for studies of the influence of gene action on the pattern of embryogenesis. Normally, it ingests bacteria as a food source. It can, however, be grown axenically in a buffered salt solution that contains tissue extracts from mammalian liver or chick embryos. Alternatively, the salt solution can be supplemented with autoclaved (killed) bacteria, sterols, and a heme-containing protein, such as hemoglobin, myoglobin, or cytochrome.

Insects. Various types of insects can be grown axenically, and thus free of both cellular and intestinal microbes. Axenic culture begins with disinfected eggs produced by bathing freshly fertilized eggs in a dilute solution of a sterilizing agent. Disinfected eggs are then raised on a sterilized diet, which can be either undefined (for example, animal parts) or completely synthetic (that is, containing pure organic ingredients). Once developed to the adult stage, axenic insects can be maintained in transparent glass culture containers to which fresh food is regularly added.

Eggs. Various marine invertebrate eggs, as well as the eggs of lower vertebrates such as fish and frogs, can also be disinfected and grown into adults to provide breeding stock for permanent axenic strains.

Mammals. To raise small mammals, such as mice and rats, for long-term (multiple-generation) studies, the whole room that houses the animals is made germ-free. Then cesarean-delivered animals are raised in the room to provide breeding stock. Sterilized bedding and food are delivered to the room; fresh feces are collected regularly and tested for the presence of bacteria. In addition, sterile food and beverage (usually water) must be routinely supplied.

Antibiotic therapy. Since bacteria represent the single most important potential contaminant in axenic cultures, antibiotic therapy is usually employed to maintain the cultures. There is no single antibiotic that inhibits all bacteria, however, so combinations of up to four to six antibiotics are usually added either directly to the culture fluid for axenic stocks of protozoa, plant cells, or nematodes, or to the food or beverage for larger, germ-free animals such as small mammals. Axenic cultures need to be continually monitored for the presence of contaminating microorganisms. *See* BACTERIOLOGY.

Uses. Axenic cultures have been used extensively for both basic research and practical application.

Basic research. Many microorganisms occupy subtle niches in host species. For example, the colon of mammals is coated with a mucous lining that is heavily infested with highly anaerobic bacteria. These bacteria outnumber other bacteria by up to 1000 to 1, but because of their extreme sensitivity to oxygen they earlier evaded detection. Once axenic hosts (for example, mice) became available, it was possible to inoculate them and to identify the interrelationships between microbe and host. An extension of that strategy has been successfully employed in dental research: to study the role of bacterial flora in tooth decay, a germ-free host, usually a mammal such as a small rodent or dog, has a pure culture of bacteria, for example, *Streptococcus*, applied as an oral swab.

Practical applications. Agriculture, medicine, and aerospace research have all benefited tremendously from application of axenic culture methods.

For plants, axenic cultures of peanut, tobacco, corn, and grass, as well as flowering plants, are now routine. Three areas of research are being exploited: (1) determination of plant growth requirements, (2) analysis of organic compounds released from certain plants, and (3) colonization of certain plant roots by bacteria as well as fungi.

Economically important agricultural animals, such as axenic chickens, have been the subject of intense research with the goal of separating the effects of the microflora, which abound in the domestic fowl's digestive tract, from the host's normal digestive enzyme action. Digestion is substantially modified in the axenic chicken. For example, in the crop, the first stages of starch digestion depend on bacteria, which secrete lactic acid. In axenic chickens, that phase of digestion is absent.

Silkworm culture became more practical because of the elucidation of the pathogenic microorganisms that affect the development and productivity of the silkworm.

Parasitology has also benefited. Several parasitic protozoa have been cultivated under axenic conditions and then inoculated into germ-free animals. It was discovered, for example, that when *Entamoeba histolytica* was inoculated into germ-free guinea pigs it failed to induce typical abscesses, as it normally would when injected into a conventional host. Further experiments revealed that in germ-free hosts this parasite was deprived of its natural food source—the bacterial flora of the intestine.

Finally, in aerospace research it is important that flights to other planets or the Moon do not contaminate their surfaces with microorganisms from Earth. Hence, any equipment involved in an actual landing needs to be free of contaminating organisms. Similarly, any extraterrestrial samples that are brought back to Earth should be treated as axenic cultures in order to facilitate their characterization and analyses.

Limitations. One problem is that it is not always possible to be certain that viruses and mycoplasmas are not present as a contaminant in axenic cultures. Another limitation is that since axenic cultures are designed to be pure it is almost impossible to construct a holistic or comprehensive view of an organism with data from axenic cultures alone.

Nevertheless, the benefits that have been obtained from the use of axenic cultures have far outweighed the limitations. Indeed, virtually all aspects of research on pharmaceuticals, nutrition, ecology, agricultural production, and parasitology have benefited enormously from the use of axenic cultures. *See* POPULATION ECOLOGY.

George M. Malacinski

Bibliography. S. S. Bhojwani (ed.), *Plant Tissue Culture: Applications and Limitations*, 1991; R. I. Freshney (ed.), *Animal Cell Culture: A Practical Approach*, 1992; L. Nuzzolo and A. Vellucci, *Tissue Culture Techniques*, 1983.

Aye-aye

A prosimian primate in the family Daubentoniidae inhabiting Madagascar. It is the most endangered of all mammals in Madagascar. A member of the order Primates and the suborder Prosimiae, aye-ayes (*Daubentonia madagascariensis*) are most closely related to the lemurs.

Morphology. Aye-ayes have a long, coarse, dark brown to black coat; long, naked, leathery ears; large eyes surrounded by distinctive dark rings; and an extremely long, slender middle finger with a long claw which is used not only to groom but also to extract insect larvae from beneath the bark of trees and to dig out the pith of bamboo, sugarcane, and coconuts. The nose, cheeks, chin, throat, and spots over the eyes are yellowish white. The hands and feet are black. At times, the aye-aye will suspend itself by its hind feet, thus freeing its hands for feeding or grooming. The tail is long and bushy. Aye-ayes have a head-plus-body length of 360–440 mm (14–17 in.), a tail length of 500–600 mm (19.5–23 in.), and they weigh 2–3 kg (4.5–6.5 lb). There is no sexual dimorphism. The mouth contains 18 teeth (I 1/1, C 0/0, PM 1/0, M 3/3 × 2). The incisors are large, curved, and chisel-like with enamel on only the front surface; they continue to grow throughout the life of the animal in the same manner as the incisors of a rodent. A second, larger species (*D. robusta*), known only from skeletal remains and a few teeth, became extinct several centuries ago.

Ecology. Although aye-ayes are primarily arboreal and inhabit rainforests, mangrove forests, and bamboo thickets, they frequently descend to the ground. They are the largest nocturnal primate and live at very low densities. They are either solitary or live in small family groups. They construct elaborate sleeping nests approximately 500 cm (20 in.) in diameter in the forks of trees. The nests are constructed of twigs and interlaced leaves and are located 10–15 m (30–45 ft) above the ground. Each individual may have 5 to 10 such nests in their territory of approximately 5 hectares (12.5 acres). Male home ranges are overlapping and may be over 200 hectares (500+ acres), although female home ranges have been found to be considerably smaller.

When searching for food, aye-ayes will listen intently for the sounds of insect larvae in decaying wood. They often tap on the wood with their long middle finger, an action thought to help them locate the extensive system of galleries constructed by insect larvae as well as by ants and termites. When a food source is located, the aye-aye will bite through the wood using its strong incisors and extract the insect with its long, clawed middle finger.

Reproduction and development. Females breed every 2 to 3 years and give birth to a single young. The only two recorded gestation periods have been 158 and 172 days. The aye-aye is the only species of the order Primates in which the paired mammary glands and nipples are located in the abdominal, rather than the thoracic, region. Weaning occurs at about 1 year of age, and they reach sexual maturity at 3 years of age. Individuals have lived for more than 23 years in captivity.

Distribution. The aye-aye occurs in a few rainforests on the eastern coast of Madagascar, a rainforest in east-central Madagascar, in several localities in humid northwestern forests, and even in deciduous forests in western Madagascar. These regions contain the older, tall trees which harbor beetle larvae upon which aye-ayes feed. In 1996, 11 aye-ayes were taken to Nosy Mangabe Island off the northeastern coast of Madagascar. This island is a protected wildlife reserve and now contains a relocated population of aye-ayes. It is estimated that the total remaining number of wild aye-ayes numbers between 1000 and 10,000. Their decline has been due mainly to forest destruction and killing by humans. They are designated as endangered by the International Conservation Union (IUCN) and the USDI, and they are on Appendix 1 of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES). *See* MAMMALIA; PRIMATES.

Donald W. Linzey

Bibliography. *Grzimek's Encyclopedia of Mammals*, vol. 2, McGraw-Hill, 1990; D. Macdonald,

The Encyclopedia of Mammals, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Azeotropic distillation

Any of several processes by which liquid mixtures containing azeotropes may be separated into their pure components with the aid of an additional substance (called the entrainer, the solvent, or the mass separating agent) to facilitate the distillation. Distillation is a separation technique that exploits the fact that when a liquid is partially vaporized the compositions of the two phases are different. By separating the phases, and repeating the procedure, it is often possible to separate the original mixture completely. However, many mixtures exhibit special states, known as azeotropes, at which the composition, temperature, and pressure of the liquid phase become equal to those of the vapor phase. Thus, further separation by conventional distillation is no longer possible. By adding a carefully selected entrainer to the mixture, it is often possible to "break" the azeotrope and thereby achieve the desired separation. *See* AZEOTROPIC MIXTURE; DISTILLATION.

Entrainers fall into at least four distinct categories that may be identified by the way in which they make the separation possible. These categories are: (1) liquid entrainers that do not induce liquid-phase separation, used in homogeneous azeotropic distillations, of which classical extractive distillation is a special case; (2) liquid entrainers that do induce a liquid-phase separation, used in heterogeneous azeotropic distillations; (3) entrainers that react with one of the components; and (4) entrainers that dissociate ionically, that is, salts. *See* SALT-EFFECT DISTILLATION.

Within each of these categories, not all entrainers will make the separation possible, that is, not all entrainers will break the azeotrope. In order to determine whether a given entrainer is feasible, a residue curve map for a mixture undergoing simple distillation is created.

Residue curve maps. The least complicated of all distillation processes is the simple distillation, or open evaporation, of a mixture from an open vessel. The liquid is boiled in such a way that the vapor is removed from contact with the liquid as soon as the vapor is formed. The composition of the liquid will change continuously with time since the vapor is always richer in the more volatile components than the liquid from which it came. The path of liquid compositions starting from some initial point is known as a simple distillation residue curve, or simply, a residue curve. The collection of all such curves for a given mixture is known as a residue curve map. These maps contain the same information as the corresponding phase diagram for the mixture, but they represent it in a way that it is more useful for understanding and designing distillation systems.

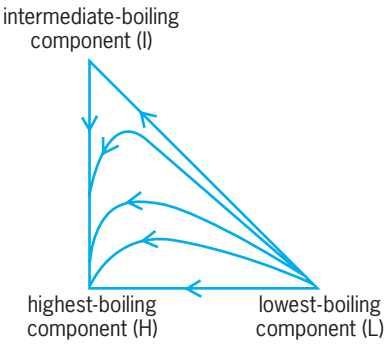

Fig. 1. Schematic representation of the residue curve map for nonazeotropic ternary mixtures.

Mixtures that do not contain azeotropes have residue curve maps that all look the same. For ternary mixtures of this sort, the map looks like the one shown in **Fig. 1**. The residue curves all start at the lowest-boiling pure component (L), they then move toward the intermediate-boiling component (I), and they end at the highest-boiling component (H). These curves show that when such a mixture is boiled the liquid initially gets richer in each of the heavier components but eventually gets richer in only the heaviest component until only the pure heavy component remains. One of the most important properties of residue curves is that they must move in such a way that the boiling temperature of the mixture increases along every curve. The shape of the curve in Fig. 1 reflects this property. The arrows on the edges point in the direction of increasing boiling temperature, and the map in the interior is the only structure that is consistent with the sides of the triangle. Typical systems for nonazeotropic ternary mixtures include hexane (L) + heptane (I) + nonane (H) and acetaldehyde (L) + methanol (I) + water (H).

The presence of even one binary azeotrope destroys the structure shown in Fig 1. If the mixture contains a single minimum-boiling binary azeotrope, three residue curve maps are possible, depending on whether the azeotrope is between the lowest- and highest-boiling components (**Fig. 2***a*), between the intermediate- and highest-boiling components

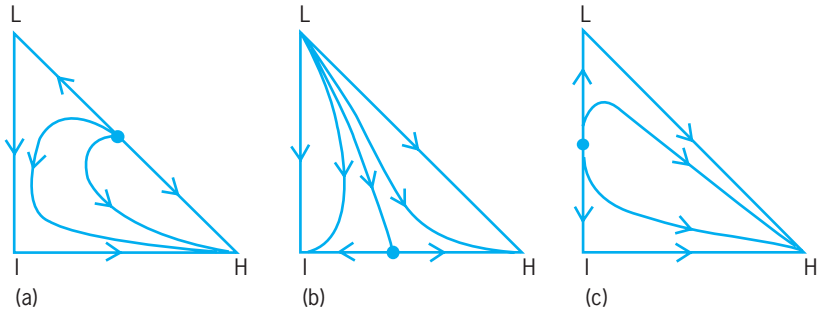

Fig. 2. Schematic representation of the residue curve maps for ternary mixtures with one minimum-boiling binary azeotrope. (*a*) Azeotrope between the lowest- and highest-boiling components. (*b*) Azeotrope between the intermediate- and highest-boiling components. (*c*) Azeotrope between the intermediate- and lowest-boiling components.

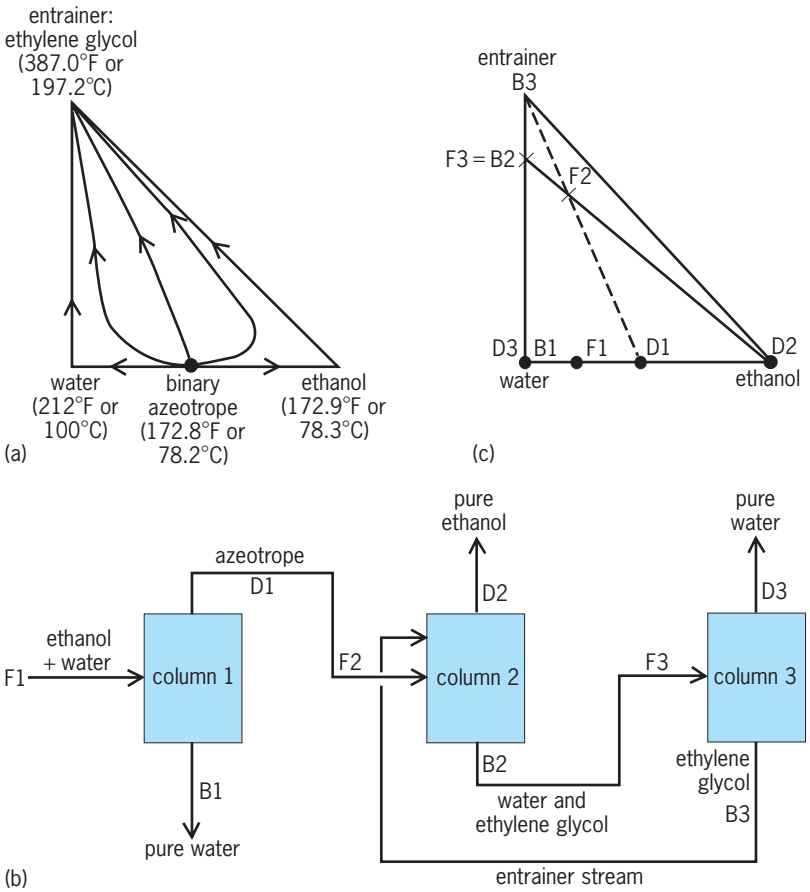

Fig. 3. Classical extractive distillation of ethanol and water using ethylene glycol as entrainer. (*a*) Residue curve map. Temperatures shown are the normal boiling points of the pure components and the azeotrope. Binary azeotropic composition is not drawn to scale, for clarity. (*b*) Distillation sequence showing column configuration. (*c*) Overall material balance lines. F = feed. D = distillate. B = bottoms.

(Fig. 2*b*), or between the intermediate- and lowest-boiling components (Fig. 2*c*). It is significant that the residue curves in Fig. 2*a* and *c* all end up at the highest-boiling pure component, but in Fig. 2*b* the curves end at either the I vertex or the H vertex depending on the initial composition. The special curve that divides these two distillation regions is called a distillation boundary. Mixtures with more than one azeotrope exhibit more complex residue curve maps with either single or multiple distillation boundaries. Typical systems include methanol (L) + i-propanol (I) + water (H), as in Fig. 2*b*, and ethanol (L) + water (I) + ethylene glycol (H), as in Fig. 2*c*.

Homogeneous azeotropic distillation. When a mixture is separated continuously in a distillation column, the overall material balance demands that the feed composition, the distillate composition, and the bottoms composition lie on a straight line in the composition triangle. It can be argued that these overall material balance lines cannot cross a simple distillation boundary to any appreciable or useful extent. Thus, nonazeotropic mixtures may be separated into their pure components by using a sequence of distillation columns because there are no distillation boundaries to get in the way. Of course, the situation is quite different when azeotropes are present, as can be seen from Fig. 2. It is possible to separate mixtures that have residue curve maps similar to those shown in Fig. 2*a* and *c* by straightforward sequences of distillation columns. This is because these maps do not have any distillation boundaries. These, and other feasible separations for more complex mixtures, are referred to collectively as homogeneous azeotropic distillations. Without exploiting some other effect (such as changing the pressure from column to column), it is impossible to separate mixtures that have residue curve maps like Fig. 2*b*.

Extractive distillation. A large number of mixtures have residue curve maps similar to Fig. 2*c*, and therefore the corresponding distillation is given the special name extractive distillation. As can be seen from this residue curve map, the entrainer, H, that must be added to the binary azeotropic mixture of components L and I must have the properties of being the heaviest pure component in the mixture and of introducing no new azeotropes. For example, ethylene glycol acts as just such an entrainer for separating ethanol and water. The residue curve map, the corresponding distillation sequence, and the overall material balance lines for this extractive distillation are shown in **Fig. 3**. Column 1 (Fig. 3*b*) separates the binary mixture of ethanol and water into a pure water stream (B1) and an azeotrope (D1). This azeotrope and the ethylene glycol entrainer are fed into column 2, which is designed to produce pure ethanol as distillate (D2) and a binary mixture of water and ethylene glycol as bottoms (B2). Finally, column 3 performs the easy separation between water (D3) and ethylene glycol (B3), which is recycled to the middle column (2). Typical overall energy requirements to separate a feed mixture consisting of 10 wt % ethanol and 90 wt % water by this method is 22,000 Btu/gallon (6100 kilojoules/liter) of anhydrous ethanol. A well-designed, thermally integrated sequence of distillation columns can carry

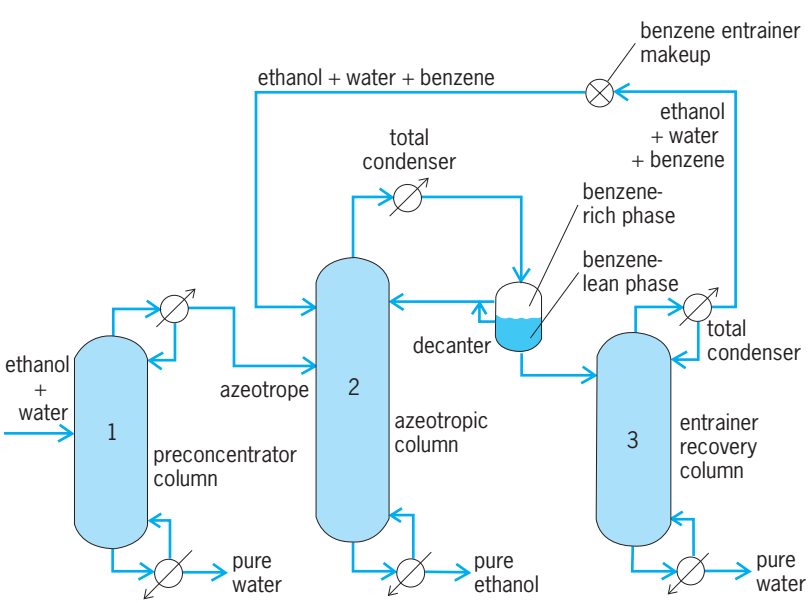

Fig. 4. Typical sequence of columns for heterogeneous azeotropic distillation, for example, ethanol + water + benzene (entrainer).

out this same separation for about 8000 Btu/gallon (2200 kJ/liter) of ethanol. Other azeotropic mixtures that are separated by this technique include (entrainers are in parentheses) acetone + methanol + (water), methanol + methyl acetate + (acetic acid), propanol + allyl alcohol + (water), and methanol + acetone + (tetrahydrofuran).

Heterogeneous azeotropic distillation. Heterogeneous entrainers cause liquid-liquid phase separations to occur in such a way that the composition of each phase lies on either side of a distillation boundary. In this way, the entrainer allows the separation to "jump" over a boundary that would otherwise be impassable. Classic examples include ethanol + water plus; (benzene), or any one of the entrainers pentane, hexane, or heptane instead of benzene; and isopropanol + water + any one of the previous entrainers. The residue curve maps associated with all these mixtures are essentially the same, having two binary minimum-boiling homogeneous azeotropes, one binary minimum-boiling heterogeneous azeotrope, and one ternary minimum-boiling heterogeneous azeotrope. The corresponding sequence of distillation columns is shown in **Fig. 4**.

The first column (preconcentrator) concentrates the binary feed mixture (ethanol + water) into an azeotrope as distillate and water as bottoms. The binary azeotrope is fed to the second column (azeotropic), together with an ethanol + water + benzene stream which is recycled from the third column (entrainer recovery). The bottoms stream is ethanol product, and the overhead vapor is close to the composition of the minimum-boiling ternary azeotrope. The vapor is condensed, whereupon it splits into two liquid phases. The benzene-rich layer is returned to the column as reflux, and the water-rich layer is fed to the final column (entrainer recovery), where it is separated into a bottoms stream of pure water and a distillate stream containing all three components. This stream is recycled to the second column. A typical overall energy requirement to separate a feed mixture consisting of 10 wt % ethanol and 90 wt % water by this method is 23,000 Btu/gallon (6400 kJ/liter) of anhydrous ethanol.

While this configuration is the most common, other heterogeneous residue curve maps have been exploited. The separation of diethoxymethane from a mixture of ethanol and water, has been developed commercially. M. F. Doherty

Bibliography. *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., 1999.

Azeotropic mixture

A solution of two or more liquids, the composition of which does not change upon distillation. The composition of the liquid phase at the boiling point is identical to that of the vapor in equilibrium with it, and such mixtures or azeotropes form constant-boiling solutions. The exact composition of the azeotrope changes if the boiling point is altered by a change in the external pressure. A solution of two components which form an azeotrope may be separated by distillation into one pure component and the azeotrope, but not into two pure components. Standard solutions are often prepared by distillation of aqueous solutions until the azeotropic composition is reached. At 1 atm (760 mmHg or approximately 100 kilopascals) pressure, hydrogen chloride and water form an azeotrope containing 20.24% by weight of HCl. *See* DISTILLATION; SOLUTION. Francis J. Johnston

Azide

A compound containing the group $—N_3$, which can be represented as a resonance hybrid of two structures, as shown in expression (1).

$$—\overset{\ominus}{N}—\overset{\oplus}{N}\equiv N \longleftrightarrow —N=\overset{\oplus}{N}=\overset{\ominus}{N} \qquad (1)$$

Sodium azide, from which most other azides are prepared, is manufactured by passing nitrous oxide over heated sodium amide, reaction (2). It is a water-

$$2NaNH_2 + N_2O \rightarrow NaN_3 + NaOH + NH_3 \qquad (2)$$

soluble, stable compound. Heavy-metal azides are highly explosive and very shock-sensitive; lead azide, $Pb(N_3)_2$, is used as a detonator to set off explosives. Sodium azide, in combination with an oxidizing agent, may be used as a gas generator in motor-vehicle passive-restraint systems. *See* EXPLOSIVE.

A variety of organic azides are known. The most important of these are aryl azides (ArN_3), azido-formates ($ROCON_3$), and sulfonyl azides (RSO_2N_3). These lose N_2 when heated or exposed to ultraviolet light to generate species known as nitrenes, which are so reactive that they will react with almost any organic compound to form amine derivatives, as shown in reaction (3). Aryl azides are widely used

$$RN_3 \xrightarrow[-N_2]{\text{heat or light}} R\ddot{\underset{..}{N}} \xrightarrow{R'CH_3} RNHCH_2R' \qquad (3)$$

to probe the active sites of biological targets by photoaffinity labeling. Difunctional aryl azides are used commercially to prepare photoresists, the nitrenes reacting with a polymer containing double bonds to insolubilize the polymer in the light-struck areas. The insoluble polymer protects the underlying metal from being attacked by an etching solution. Difunctional azido-formates and sulfonyl azides can be used to cross-link polymers, to prepare polymeric foams, and to adhere tire cord to rubber in the manufacture of automobile tires. Compounds containing a sulfonyl azide group and a hydrolyzable silane group, such as $—Si(OCH_3)_3$, in the same molecule are used to bond silaceous fillers (glass fibers, silica, mica, and so forth) to almost any organic polymer. Such coupled systems have much superior properties to simple mechanical mixtures.

Simple acyl azides (R = alkyl or aryl) undergo a Curtius rearrangement on heating to form an isocyanate, reaction (4).

$$RCON_3 \xrightarrow{heat} RN{=}C{=}O + N_2 \qquad (4)$$

See NITROGEN; POLYMERIZATION; RESONANCE (MOLECULAR STRUCTURE); RUBBER. David S. Breslow

Bibliography. W. Lwowski (ed.), *Nitrenes*, 1970; S. Patai (ed.), *The Chemistry of the Azido Group*, 1971; E. F. V. Scriven (ed.), *Azides and Nitrenes*, 1984.

Azurite

A basic carbonate of copper with the chemical formula $Cu_3(OH)_2(CO_3)_2$. Azurite is normally associated with copper ores and often occurs with malachite. Azurite is monoclinic. It may be massive or may occur in tabular, prismatic, or equant crystals (see **illus.**). Invariably blue, azurite was originally used extensively as a pigment. Hardness is 3.5–4 (Mohs scale) and specific gravity is 3.8. It can be synthesized by gentle heating of cupric nitrate or sulfate solutions with calcium carbonate in a closed tube. Notable localities for azurite are at Tsumeb, Southwest Africa, and Bisbee, Arizona. *See* COPPER; MALACHITE. Robert I. Harker

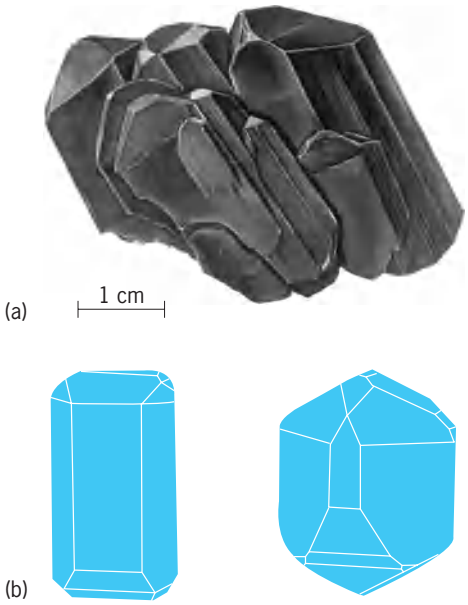

Azurite. (*a*) Crystals from Tsumeb, Southwest Africa (*specimen from Department of Geology, Bryn Mawr College*). (*b*) Crystal habits (*after C. Klein, Manual of Mineralogy, 21st ed., Wiley, 1993*).

B

Bacillariophyceae — Binomial theorem

Bacillariophyceae

A class of nonflagellate unicellular algae, commonly called diatoms, with boxlike silicified walls. Diatoms range in maximum dimension from 4 micrometers to more than 1 millimeter. The diatom wall or frustule (**Fig. 1**) comprises several interlocking, usually elaborately sculptured, lightly or heavily silicified pieces overlying a thin polysaccharide layer. The two largest pieces are the upper and lower valves, which fit together like the top and bottom of a petri dish or shoe box. Between the valves (along the side or girdle of the cell), several smaller pieces—hooplike girdle bands—are intercalated. Depending upon which dimension is larger, breadth or depth, a diatom tends to lie on the valve side or on the girdle side.

Valves are honeycombed by perforate chambers arranged in patterns characteristic of individual species. The overall symmetry of the valves and details of their structure and ornamentation provide the basis for classification. More than 200 genera and 12,000 species of living diatoms have been described. Two main groups are recognized: those in which structural features of the valve are arranged with reference to a central pole (centric valve; **Fig. 2**) or to two or more poles (gonioid valve); and those in which the features are arranged with respect to a line, often symmetrically (pennate valve). The first group are usually called centric diatoms (although not all members have centric valves) and are circular, semicircular, elliptical, or polygonal in valve view; the second group, usually called pennate diatoms, are naviculoid (boat-shaped), crescent-shaped, or needle-shaped in valve view.

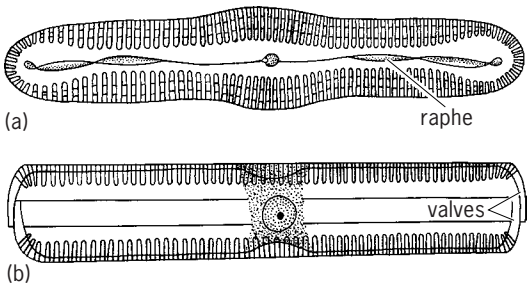

Fig. 1. Pennate diatoms. *Pinnularia*: (*a*) top, or valve view, showing longitudinal slit, or raphe; (*b*) side, or girdle view, showing overlapping halves or valves. (*After H. J. Fuller and O. Tippo, College Botany, rev. ed., Holt, 1954*)

The diatom cell contains a nucleus and, except in the few colorless species, one to many chloroplasts, which may be minute disks or expanded plates. The chloroplasts are yellow to dark brown as a result of the masking of chlorophylls *a* and *c* by β-carotene and various xanthophylls, especially fucoxanthin. Pyrenoids are present in some species. Thylakoids are grouped in threes, and chloroplast deoxyribonucleic acid is distributed in a ring. A noteworthy feature of the diatom cell is the spindle, which is composed of highly organized microtubular elements and is readily visible with a compound microscope during cell division.

Reproduction. Vegetative reproduction consists of binary fission in the plane of the valves. Following mitosis, two new valves are formed, each lying inside an old valve. One of the daughter cells is thus the same size as the parent cell, while the other is smaller. The average cell size of a population gradually diminishes to a value as small as one-fifth maximum. Eventually, maximum size is restored by sexual reproduction or by an asexual regenerative process.

Vegetative cells of both pennate and centric diatoms are diploid, with meiosis occurring during gametogenesis. Centric diatoms are oogamous, with one or two eggs and up to 128 sperm being produced by separate cells. Each sperm has only one flagellum, which is anteriorly inserted and lacks the central microtubular doublet characteristic of all other flagella. Pennate diatoms are isogamous or anisogamous, with one or two ameboid gametes being produced by each cell. The zygote in both centric and

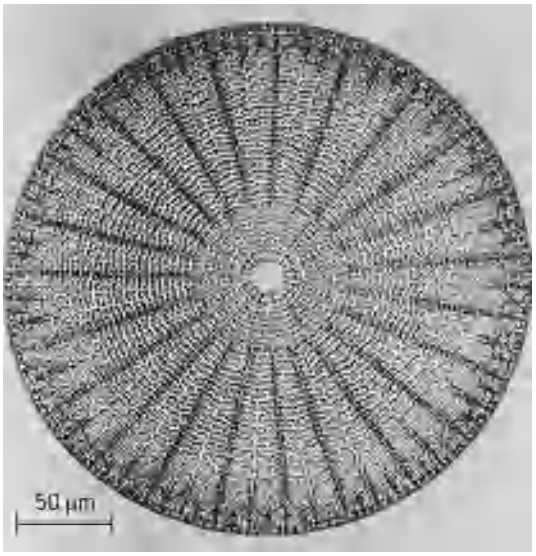

Fig. 2. ***Arachnoidiscus ehrenbergii*, a concentric diatom with radial symmetry. (*After H. J. Fuller and O. Tippo, College Botany, rev. ed., Holt, 1954*)**

pennate diatoms enlarges into an auxospore, with an organic wall that may be covered by small siliceous scales or rings. A frustule pattern characteristic of the species is restored in the second generation of daughter cells. The primary inductor of sexuality is a critical minimum size, but high light intensity may also be important.

Motility of vegetative cells is found only in pennate diatoms that possess a raphe, which is a complex longitudinal slit or end-to-end pair of slits passing through the wall of one or both valves. Movement, which can occur only if the raphe-bearing valve is in contact with the substrate, is effected by secretion of polysaccharide along the length of the raphe. A bundle of actin microfilaments, assumed to play a role in movement, parallels the raphe in the adjacent cytoplasm.

Ecology. Most diatoms are photosynthetic. Many, however, are auxotrophic, requiring an external source of certain vitamins. The few colorless species are found where high concentrations of nutrients are available, as in sewers or in the mucilage of brown algae. Silicon, which in the hydrated form $SiO_2 \cdot H_2O$ makes up the frustule, is of critical importance. Diatoms cannot divide in the absence of silicate. Germanium is a competitive inhibitor to the uptake of silicate and is thus a toxin that is selective for diatoms. Food reserves include lipids and the polysaccharide chrysolaminaran.

Diatoms are likely to occur wherever there is moisture. They are free-living or attached, solitary or colonial. Planktonic forms (primarily centric diatoms) bear long spines and bristles that aid flotation. Free-living benthic forms include many motile pennate diatoms. When fixed to a substrate, diatoms may be cemented individually by the lower valve, attached by a gelatinous pad or stalk, or enveloped in a gelatinous film. Colonies take the form of chains of cells linked valve to valve by interlocking spines, mucous or chitinous threads, or mucilaginous pads, or of pseudofilaments or blades composed of numerous individuals (often motile) in a common tubular or flattened mucilaginous matrix.

Terrestrial habitats, which generally support pennate rather than centric diatoms, include wet rocks in the spray zone at the seashore, along streams, and, if there is sufficient light, in caves. Diatoms are also found in the upper layers of soil, in old snow, and among mosses. They inhabit all bodies of still and running fresh water, including hot springs with temperatures as high as 118°F (48°C). They are abundant in both plankton and benthos and may be attached to other algae, higher plants, animals, or sediment particles, or free-living in the sediment. In marine habitats they are often the main constituent of the phytoplankton and the mudflat flora. They are the most abundant eukaryotic organisms on the lower surface of sea ice in the Arctic and Antarctic, where they survive temperatures below 28°F (−2°C). Diatoms grow on the surface of most marine macroalgae, and several species occur on the skin of whales. A few live symbiotically as naked cells within marine invertebrates and foraminifera.

Marine planktonic diatoms have been important primary producers for at least 100 million years, and over the millennia their frustules have accumulated on the ocean floor. From the abundant fossil record, it is known that centric diatoms evolved first (Cretaceous), followed by pennate diatoms (Paleocene). Uplifted deposits (called diatomite, diatomaceous earth, or kieselguhr) are mined at several locations, including Santa Barbara County, California. Diatomite is used as a filter, abrasive, catalyst carrier, absorbent, and insulator.

Because certain species are characteristic of clean water while others are characteristic of organically polluted water, diatoms may be used as an environmental index. *See* ALGAE; CHRYSOPHYCEAE.

Paul C. Silva; Richard L. Moe

Bibliography. P. Bourrelly, *Les Algues d'Eau Douce . . .*, tome 2: *Les Algues Jaunes et Brunes, Chrysophycées, Phéophycées, Xanthophycées et Diatomées*, 1968; N. I. Hendey, *An Introductory Account of the Smaller Algae of British Coastal Waters*, pt. 5: *Bacillariophyceae (Diatoms)*, 1964; D. Werner (ed.), *The Biology of Diatoms*, 1977.

Bacillary dysentery

A highly contagious intestinal disease caused by rod-shaped bacteria of the genus *Shigella*. Bacillary dysentery is a significant infection of children in the developing world, where it is transmitted by the fecal-oral route. The global disease burden is estimated as 165 million episodes and 1.3 million deaths annually. Common-source outbreaks occasionally occur in developed countries, usually as a result of contaminated food. The most common species isolated in developed countries is *S. sonnei*, while *S. flexneri* serotypes predominate in endemic areas.

Epidemics of *S. dysenteriae* 1 occur in equatorial regions, and these outbreaks can involve adults as well as children.

When ingested even in very small numbers, shigellae multiply in the intestine and invade the epithelial lining of the colon. Infection of this tissue elicits an acute inflammatory response (colitis) that is manifested as diarrhea or bloody, mucoid stools (dysentery). The virulence of all *Shigella* species, and *Shigella*-like enteroinvasive *Escherichia coli*, depends on an extrachromosomal genetic element (virulence plasmid) that encodes four invasion plasmid antigen (Ipa) proteins and a secretory system (Type III) for these proteins. Secreted Ipa proteins help shigellae to initiate colonic invasion through specialized endocytic intestinal cells (M cells). After shigellae pass through these M cells, they are phagocytized by tissue macrophages in the underlying lymphoid tissue. Ipa proteins then induce apoptosis (programmed cell death) in infected macrophages, releasing cytokines (primarily IL-1) that initiate an acute, localized inflammatory infiltrate. This infiltrate of polymorphonuclear leukocytes destabilizes tight junctions between absorptive epithelial cells (enterocytes), making the tissue more susceptible to additional *Shigella* invasion. Secreted Ipa proteins induce uptake of shigellae by the colonic enterocytes. The virulence plasmid also encodes an intercellular spread protein (IcsA) that recruits mammalian cytoskeletal elements (primarily actin) to the bacterial surface. This actin is organized into a cytoplasmic motor that facilitates spread of shigellae to adjacent enterocytes. *See* DIARRHEA; ESCHERICHIA.

In otherwise healthy individuals, bacillary dysentery is typically a short-term disease lasting less than a week. The symptoms can be truncated by appropriate antibiotic therapy (such as oral ampicillin or cyprofloxacin) that rapidly eliminates shigellae from the intestinal lumen and tissues. When *S. dysenteriae* 1 is the etiologic agent, however, hemolytic uremic syndrome (HUS) can be manifested as a serious consequence of disease. This species produces a cytotoxin (Shiga toxin or Stx) that is functionally identical to the toxin of enterohemorrhagic *E. coli* (for example, O157:H7). Stx inhibits protein synthesis, damaging endothelial cells of the intestinal capillary bed; the toxin may also damage renal tubules, causing acute renal failure with chronic sequela in up to one-third of HUS patients. *See* MEDICAL BACTERIOLOGY.

Thomas L. Hale

Bibliography. R. Bhimma et al., Post-dysenteric hemolytic uremic syndrome in children during an epidemic of *Shigella* dysentery in Kwazulu/Natal, *Pediat. Nephrol.*, 11:560–564, 1997; E. Egile et al., SopA, the outer membrane protease responsible for polar localization of IcsA in *Shigella flexneri, Mol. Microbiol.*, 23:1063–1073, 1997; R. Menard et al., The secreted Ipa complex of *Shigella flexneri* promotes entry into mammalian cells, *Proc. Nat. Acad. Sci. USA*, 93:1254–1258, 1996; P. J. Sansonetti et al., Role of IL-1 in the pathogenesis of experimental shigellosis, *J. Clin. Invest.*, 96:884–892, 1995.

Background count

The number of counts recorded by a radiation detector from background radiation. The term background radiation refers to the natural ionizing radiation on the Earth. Ionizing radiation refers to all radiations, waves, and particles that are energetic enough to remove electrons from stable atoms; they are stronger than infrared radiation, radio waves, or visible light, which cannot separate electrons from stable atoms. Radiation strong enough to cause ionization of atoms is measured in electrical units which range from 32 electronvolts up to millions of electronvolts. *See* ELECTROMAGNETIC RADIATION; IONIZATION POTENTIAL.

Sources. Background radiation comes from two sources, cosmic and terrestrial. The sum of these radiations is called the background radiation dose. Cosmic radiation comes from outer space and consists of gamma rays and alpha, neutron, and beta particles. Alpha particles from outer space are very energetic, and each consists of two neutrons coupled to two protons. They are relatively massive but do not penetrate deeply into matter such as human tissue. Beta particles are energetic electrons and are more penetrating. However, it is the gamma-ray flux, that is, concentration, that constitutes most of the cosmic-ray dose impacting on the Earth. This dose, the collection of all ionizations, is attenuated as it passes through the Earth's atmosphere; the dose rate roughly doubles for each 2000 m (6500 ft) increase in altitude. At thinner mountain atmospheres the cosmic-ray dose is higher, and on jet flights the dose rate can be 100 times that at sea level. Some of the cosmic rays are so energetic that they can still be detected in the deepest mines. *See* ALPHA PARTICLES; BETA PARTICLES; COSMIC RAYS; GAMMA RAYS.

The Earth is naturally radioactive with the decay products of natural elemental uranium and thorium which are present (at about 1 part per million on average) in every gram of the Earth's mantle. These primordial radionuclides decay through a complex chain, ultimately ending as stable elemental lead. Each step in the decay process results in the emission of some ionizing particles and gamma rays.

Dosage. A radiation detector at sea level, for example, would detect the radiation from both cosmic and terrestrial sources. The measurement of background radiation is in absorbed dose units known as sieverts (Sv). (Previously the unit was the millirem, and 100 millirem = 1 mSv.) The chief concern is with the amount of background radiation absorbed by people's tissues. Most people absorb about 0.3 mSv per year at sea level, roughly 1 microsievert per day. About 10% of this is external radiation from cosmic radiation, and 10% is external radiation from terrestrial sources. Internal radiation from inhaled and ingested radionuclides constitutes the rest. Some 13% of the background radiation comes from the natural terrestrial radionuclide (radioisotope) potassium-40, which is uniformly concentrated in all living cells and is present as 1 in every 2000 potassium atoms. The remaining two-thirds of the

background dose comes from inhalation of radon daughter products. Radon is a radioactive noble gas which is derived from the decay of radium, a decay product of uranium. *See* RADON; UNITS OF MEASUREMENT.

In regions where the geology is such that there is a relatively high concentration of uranium, the radon concentration may be elevated, and the background dose rate can be quite high. About 0.1% of the population of the United States may receive radon doses 10 times the average, and some 10–15% may be subjected to levels 5 times higher than the average. One such area is the so-called Reading Prong in Pennsylvania, where the levels of radon in some buildings have been found to be so high as to constitute a possible radiation hazard for the occupants.

Similarly, at high altitudes, such as in Denver, Colorado, at 1.6 km (1 mi), the increased cosmic-ray dose raises the total background dose rate about 10% above that at sea level, but this increase is of minimal significance in terms of radiation risk. There are many areas where the background radiation rate varies because of local geology or altitude. Background radiation is not constant since cosmic flux is influenced by sunspots and radon levels are influenced by humidity and the degassing of soil.

Typically, an average person will receive a lifetime background dose of about 0.25 Sv. The level can be a somewhat lower or up to about 10 times higher, as mentioned above, but it can never be zero.

Instrument correction. All radiation detection instruments are influenced by background radiation from the sources mentioned above. Thus any measurement to determine the level of radiation in a specific environment must correct for the natural level of ionization reaching the detector. Analytical radiation-measuring devices frequently employ a means of automatically subtracting the background level so as to provide information on the net amount of radiation in the sample or environment in question. For very sensitive measurements, the detector is frequently sheathed in a heavy shield of lead, and a lead-collimated tube may also extend from the sample to the detector to shield the sensitive detector from external background radiation and minimize the background count. *See* ENVIRONMENTAL RADIOACTIVITY; RADIATION SHIELDING; RADIOACTIVITY.

Marvin Goldman

Bacteria

Extremely small—usually 0.3 to 2.0 micrometers in diameter—and relatively simple microorganisms possessing the prokaryotic type of cell construction. Although traditionally classified within the fungi as Schizomycetes, they show no phylogenetic affinities with the fungi, which are eukaryotic organisms. The only group that is clearly related to the bacteria are the blue-green algae.

Bacteria are found almost everywhere, being abundant in soil, water, and the alimentary tracts of animals. Each kind of bacterium is fitted physiologically to survive in one of the innumerable habitats created by various combinations of space, food, moisture, light, air, temperature, inhibitory substances, and accompanying organisms. Dried but often still living bacteria can be carried into the air.

One of the few locations in which bacteria are not usually found is within the cells of other healthy organisms, though even this is subject to exceptions, as there are many bacteria that do live intracellularly in a number of eukaryotic organisms.

Bacteria have a practical significance for humans. Some cause disease in humans and domestic animals, thereby affecting health and the economy. Some bacteria are useful in industry, while others, particularly in the food, petroleum, and textile industries, are harmful. Some bacteria improve soil fertility. *See* FOOD ENGINEERING; FOOD MICROBIOLOGY; INDUSTRIAL MICROBIOLOGY; MEDICAL BACTERIOLOGY; PETROLEUM MICROBIOLOGY; SOIL MICROBIOLOGY; TEXTILE MICROBIOLOGY; ZOONOSES.

As in higher forms of life, each bacterial cell arises either by division of a preexisting cell with similar characteristics or through combination of elements from two such cells in a sexual process. The earlier idea, that full-fledged bacteria arise from nonliving material by spontaneous generation, has been disproved by careful elimination of living bacteria from the nonliving material. This does not eliminate the possibility that, sometime during evolution of the universe, life was derived from the nonliving. Separation of matter into living and nonliving is arbitrary, though useful and unambiguous when transitional states are not under consideration. *See* PREBIOTIC ORGANIC SYNTHESIS; REPRODUCTION (ANIMAL).

Cultures

Descriptions of bacteria are preferably based on the studies of pure cultures, since in mixed cultures it is uncertain which bacterium is responsible for observed effects. Pure cultures are sometimes called axenic, a term denoting that all cells had a common origin in being descendants of the same cell, without implying exact similarity in all characteristics. Pure cultures can be obtained by selecting single cells, but indirect methods achieving the same result are more common.

If conditions are suitable, each bacterium grows and divides, using food diffused through the gel, and produces a mass of cells called a colony. Colonies always develop until visible to the naked eye unless toxic products or deficient nutrients limit them to microscopic dimensions. *See* AXENIC CULTURE; BACTERIOLOGY; CULTURE.

Classification

The morphology, that is, the shape, size, arrangement, and internal structures, of bacteria can be distinguished microscopically and provides the basis for classifying the bacteria into major groups. Three principal shapes of bacteria exist, spherical (coccus), rod (bacillus), and twisted rod (spirillum). The coccus may be arranged in chains of cocci as in

Streptococcus, or in tetrads of cocci as in *Sarcina*. The rods may be single or in filaments. Stains are used to visualize bacterial structures otherwise not seen, and the stain reaction with Gram's stain provides a characteristic used in classifying bacteria. *See* STAIN (MICROBIOLOGY).

Many bacteria are not motile. Of the motile bacteria, however, some move by means of tiny whirling hairlike flagella extending from within the cell. Others are motile without flagella and have a creeping or gliding motion. Spiral forms are usually polarly flagellated, that is, with flagella at the end of the cell. Cocci (spheres) are rarely flagellated. Rod-shaped bacteria may lack flagella or have polar or peritrichous (around the entire surface of the cell) flagella.

Many bacteria are enveloped in a capsule, a transparent gelatinous or mucoid layer outside the cell wall. Some form within the cell a heat- and drought-resistant spore, called an endospore. Cytoplasmic structures such as reserve fat, protein, and volutin are occasionally visible within the bacterial cell.

The nucleus of bacteria is prokaryotic, that is, not separated from the rest of the cell by a membrane. It contains the pattern material for forming new cells. This material, deoxyribonucleic acid (DNA), carrying the information for synthesis of cell parts, composes a filament with the ends joined to form a circle. The filament consists of two DNA strands joined throughout their length. The joining imparts a helical form to the double strand. The double-stranded DNA consists of linearly arranged hereditary units, analogous and probably homologous with the "genes" of higher forms of life. During cell division and sexual reproduction, these units are duplicated and a complete set is distributed to each new cell by an orderly but as yet unelucidated mechanism. *See* BACTERIAL GENETICS.

The submicroscopic differences that distinguish many bacterial genera and species are due to structures such as enzymes and genes that cannot be seen. The nature of these structures is determined by studying the metabolic activities of the bacteria. Data are accumulated on the temperatures and oxygen conditions under which the bacteria grow, their response in fermentation tests, their pathogenicity, and their serological reactions. There are also modern methods for determining directly the similarity in deoxyribonucleic acids between different bacteria.

Temperature relationships. Bacteria are said to be psychrophilic if their optimum temperature is below 60°F (20°C), mesophilic if it is 60–113°F (20–45°C), and thermophilic if it is above 113°F (45°C). Some bacteria can grow at temperatures as high as 167°F (75°C). Others, which are not killed but which cannot grow at high temperatures, are called thermoduric.

Oxygen relationships. Bacteria are said to be aerobic if they require oxygen and grow best at a high oxygen tension, usually 20% or more. Microaerophilic bacteria need oxygen, but grow best at, or may even require, reduced oxygen tensions, that is, less than 10%. Anaerobic bacteria do not require oxygen for growth. Obligatorily anaerobic bacteria can grow only in the complete absence of oxygen.

Fermentation and respiration. Fermentation is a term used to indicate processes in which foodstuffs are decomposed in the absence of oxygen. Respiration is the comparable aerobic process, in which oxygen is one of the foods. Some oxygen-utilizing microorganisms cannot completely oxidize the food to water and carbon dioxide, and often form acids as a product of this type of aerobic food utilization. These incomplete oxidative processes are sometimes called fermentations, though they are actually examples of limited respiration. As in all chemical rearrangements, some of the available energy in both respiration and fermentation is dissipated as heat. The remainder is stored in the form of the materials that make up living cells. *See* FERMENTATION.

In respiration as much as 50% of the food material and energy appears as bacterial cell material and the remainder as carbon dioxide, water, and heat. In fermentation the lack of O_2 decreases the energy supply available through rearrangement of the food, and less food (up to 15%) is converted to cells. Less heat is dissipated, and fermentation products are formed, such as CO_2, hydrogen, methane, ethanol, acetone, glycerol, and formic, acetic, propionic, butyric, lactic, and succinic acids. These products, when combined with oxygen by aerobes to form water and carbon dioxide, yield energy equal to the difference between the energy available in respiration and in fermentation.

Fermentation tests. Fermentation tests, which use liquid media, each medium containing a different nutrient, aid in classifying bacteria. Gas or acids or both are always formed when carbohydrates are fermented. Acid is detected by including an indicator of acidity in the culture medium; a small inverted tube is used to trap any gas emitted. Other useful tests measure the acidity (pH) developing during fermentation and the range of acidity permitting growth.

Fermentation of proteins yields products similar to those from carbohydrates plus large quantities of nitrogenous products, such as ammonia and amines. Since ammonia and amines are weak bases, a protein fermentation causes alkalinity instead of acidity. The ability of an organism to ferment carbohydrate or protein is tested by inoculating the organisms into milk containing litmus indicator, called litmus milk. The culture is incubated and then examined for color changes denoting acidity or alkalinity.

Digestion tests. Tests for digestion of protein, starch, fats, cellulose, pectin, and many other insoluble materials disclose other physiological characteristics useful in classification. Ability to digest protein (peptonization) is often tested by examining litmus milk cultures for an increase in transparency, caused by digestion of casein, the protein responsible for the white opacity of milk. The digestion of gelatin, another protein, may be detected by liquefaction.

Other metabolic reactions of bacteria include the oxidation of ammonia to nitrate by nitrifying

bacteria, oxidation of sulfur to sulfates by sulfur bacteria, and oxidation of ferrous to ferric iron by iron bacteria. Some of the products formed within the cell in these oxidations react with carbon from carbon dioxide, with hydrogen from water, and with other elements to form new cells.

The purple sulfur bacteria and the green sulfur bacteria also form new cells from inorganic compounds, but the hydrogen is obtained by splitting water with light (photolysis) instead of with chemical energy. These bacteria form an oxidized substance as the second product of the photolysis of water, as do green plants. But the former cannot convert this to oxygen, which the latter do; hence the photosynthetic bacteria can photosynthesize only if an oxidizable compound, such as hydrogen (H_2), hydrogen sulfide (H_2S), or a suitable organic substance, is present with which the oxidized moiety is continually reduced.

Some bacteria obtain energy from the oxidation of reduced substances with compounds other than oxygen (O_2). The sulfate reducers use sulfate, the denitrifiers nitrate or nitrite, and the methanogenic bacteria carbon dioxide as the oxidizing agents, producing H_2S, nitrogen (N_2), and methane (CH_4), respectively, as reduction products. *See* BACTERIAL PHYSIOLOGY AND METABOLISM.

Pathogenicity. Pathogenicity, the ability to cause disease, is another property used in establishing the relationship between various groups of bacteria. Some bacteria produce disease only in certain species; for example, *Neisseria gonorrhoeae* will cause gonorrhea only in humans. Some bacteria cause only one disease, while others may cause several diseases. An example of the former is *Corynebacterium diphtheriae*, which causes diphtheria; *Staphylococcus aureus* belongs to the latter category and may cause boils, osteomyelitis, and pneumonia. *See* PATHOGEN.

Serological reactions. Serological reactions are very useful in distinguishing closely related bacteria. If two bacteria, A and B, differ, some of their proteins and other complex molecules also differ. When cells of A are injected into an experimental animal, such as a rabbit, some of their constituent molecules (especially proteins) cause production in the rabbit's blood of special proteins called antibodies. Each of these can combine specifically with the molecular species that caused its production. If, after a suitable incubation period, blood is drawn from the animal (in the case of a rabbit usually from an ear vein) and allowed to clot, a clear yellow liquid (blood serum) is extruded as the clot shrinks. It contains antibodies against each protein in the injected A cells. If B cells in excess are added to this antiserum, each B protein occurring also in A reacts with its corresponding antibody, thereby removing from the serum all antibodies against proteins common to both A and B. Addition of A cells gives a further reaction if A contains proteins not found in B. With reciprocal absorption of B antiserum with A cells, and testing with B cells for antibodies restricted to B, any differences in the A and B cells can be detected. *See* BACTERIAL TAXONOMY; SEROLOGY.

Bacterial pathogens have been observed that form a protective capsule which may belong to one of several different serological types. Noncapsulated mutant cells occasionally arise. In the animal they are destroyed by phagocytes, but on artificial media they survive and produce colonies with a rough surface. Such strains are called rough or *R* forms to distinguish them from the capsulated, smooth or *S* types. Heavy inoculation of an *R* strain into a susceptible animal causes a change to an *S* strain since any reverse mutant (that is, a mutation from *R*) *S* cell can multiply, whereas the *R* cells are consumed by phagocytes. Pretreatment of the injected *R* cells with an extract of *S* cells or growing *R* cells in a medium with an extract of *S* cells induces reversion to an *S* strain, serologically identical with the one used to prepare the extract. *See* LYSOGENY; PNEUMOCOCCUS.

Natural defenses against infections depend in part on serological mechanisms. When bacteria enter animal hosts containing antibodies against them, the bacteria become coated with antibodies and are then susceptible to engulfment and digestion (phagocytosis) by host cells. Chicks, mice and rats, aseptically removed from the shell or uterus, can be reared bacteria-free. These axenic animals, when mature, are highly susceptible to infection by bacterial types harmless to normal animals. Antibodies carried over from the mother protect very young animals and, by the time maternal antibodies are depleted, normally reared offspring have developed their own. The isolation of axenic animals deprives them of the bacterial antigens necessary for development of protective antibodies. Every ancestor of living organisms survived bacterial attacks to reach maturity. The resistance, selected in this manner, is the factor most commonly concerned in defense against infections. *See* DISEASE.

Bacterial Variation

Variation in the characteristics of a single cell occurs during cell division, but since cells in a culture divide at different times, unless artificially synchronized, the average for the entire cell population is constant, as long as the environment is constant. A change in any limiting factor in the environment causes a change in the population. In nature environmental changes are often cyclic, and bacteria undergo accompanying changes in morphology interpreted by some as a life cycle. The applicability of this term to bacteria has been disputed.

An extremely small proportion of living organisms undergo sudden genetic changes, usually involving only one characteristic of the cell, which are transmitted through many generations. Because of the tremendous number of bacteria (1 in.^{3} or 16 ml of a culture may contain 50 billion bacteria), their mutations are common in cultures as well as in nature. If, in a given environment, a mutation enables its possessor to grow and divide more rapidly than the type from which it arose, the mutants ultimately predominate. The characteristics of the population are changed by the environment through selection of cells most fitted to survive, rather than by a direct

action on all cells as discussed in the preceding paragraph. *See* MUTATION.

Interrelationships

Interrelationships may be close and may involve particular species. Examples are the parasitic association of many bacteria with plant and animal hosts, and the mutualistic association of nitrogen-fixing bacteria with leguminous plants, of cellulolytic bacteria with grazing animals, and of luminous bacteria with certain deep-sea fishes. *See* NITROGEN FIXATION; POPULATION ECOLOGY.

Bacteria are also active in other less intimate, but no less important, natural interrelationships. Bacterial decomposition of the dead bodies of animals, and especially plants, releases for reuse by living plants the carbon dioxide needed in photosynthesis. Many other chemical activities relate bacteria to other organisms through the world pool of materials essential to life, to which all organisms contribute and from which they draw their food. Robert E. Hungate

Gas Vesicles and Vacuoles

Gas vesicles are submicroscopic structures of cylindrical shape with conical ends. They were discovered in electron micrographs of gas vacuole–containing cyanobacteria. These studies revealed that the gas vesicles are the structural units which make up the gas vacuoles recognizable within the cells by ordinary light microscopy. Gas vesicles occur exclusively in prokaryotic microorganisms. *See* VACUOLE.

Characteristics. The gas vacuoles were first described for purple sulfur bacteria in 1888 and called hollow cavities; they were named gas vacuoles in 1895, when it was observed that they conferred buoyancy to the cells containing them. In the light microscope, gas vacuoles appear in the cytoplasm as refractile hollow cavities of irregular shape and pinkish shine.

The gas vesicles are homologous structures in all prokaryotic organisms. The size of the vesicles varies considerably in different systematic groups. In the cyanobacteria and green sulfur bacteria, the vesicles are about 70 nanometers in diameter and, on the average, 400 nm long (maximum length up to 1 micrometer). The gas vesicles of the purple sulfur bacteria and of several species and genera of chemotrophic bacteria are 100–200 nm wide but only up to 300 nm long. The widest gas vesicles, with a diameter of up to 300 nm, occur in the halobacteria.

The gas vesicle membrane is 2 nm thick and exhibits a ribbed fine structure from both the outside and the inside. The ribs are 4.5 nm wide and lie perpendicular to the long axis of the vesicle cylinder. The ribs apparently represent turns of a shallow spiral rather than stacks of concentric rings. The vesicle membrance, permeable to gases but impermeable to water, encloses a hollow space into which gases diffuse freely.

The gas vesicles are fairly rigid structures. However, when suspensions of cells with gas vacuoles are subjected to a sudden increase in pressure, the gas vesicles collapse and lose their gas-filled hollow spaces. Consequently, the gas vacuoles are no longer detectable by light microscopy and the cells lose their buoyancy.

Composition and molecular structure. Intact gas vesicles can readily be isolated from gently lysed cells by centrifugation, which causes the vesicles to float to the supernatant surface, where they are skimmed off. Chemical analysis of isolated gas vesicles from cyanobacteria and halobacteria shows that the vesicles consist exclusively of protein. Only one type of protein was found in the vesicles of cyanobacteria, while the vesicles of *Halobacterium* consisted of two very similar protein types. The molecular weights of the vesicle proteins are between 13,000 and 15,000. The amino acid composition of the vesicle proteins is fairly similar in all species. The proteins have in common a high proportion (more than 50%) of hydrophobic aliphatic amino acids and a low proportion of aromatic amino acids; both cysteine and methionine are absent. *See* CENTRIFUGATION.

It was established by x-ray and neutron diffraction studies that the gas vesicle membrane consists of a monolayer of the vesicle protein. Since the outside of the vesicles is hydrophilic and wettable while the inside is hydrophobic, the protein molecules must be positioned so that their hydrophobic aliphatic amino acids are located toward the inner surface of the vesicles. It is assumed that the hollow space of the vesicle arises simultaneously with formation and enlargement of the vesicle. The gases in the surrounding cytoplasm diffuse into the hollow space, while the hydrophobic inner surface of the vesicle prevents the penetration of water as well as the formation of water droplets inside the structure. *See* CELL MEMBRANES.

Function. Good evidence exists for only one of several suggested biological functions for gas vesicles. The hollow space of the vesicles effectively decreases the specific weight of the cells and, therefore, provides buoyancy. Buoyancy can be of selective advantage only in nonturbulent aquatic habitats. In agreement with this, gas-vacuolated microorganisms are mainly encountered as planktonic cells in lakes and stagnant water bodies, while they are rarely found in soils.

Since the vesicle membrane is permeable to gases, the pressure of the gas in the vesicles is always equal to the atmospheric pressure at the surface of the aquatic habitat. The rigid structure of the vesicle membrane maintains the hollow space of the vesicle against both the turgor pressure of the cells and the hydrostatic pressure of the water at the depth in which the cells float. Depending on the critical pressure that is tolerated by the vesicle membrane of a given species, a kind of buoyancy regulation is feasible by either formation or collapse of gas vesicles. For certain planktonic cyanobacteria (for example, *Oscillatoria agardhii*) which develop their maximum population density at a certain depth in stratified lakes, buoyancy regulation has been experimentally established. When at growth-limiting light

intensities the turgor pressure of the cells is low, newly forming gas vesicles cause the cells to rise in the water column. At higher light intensities in the vertical light gradient, a rising turgor pressure may cause part of the gas vesicles to collapse and, consequently, cause the cells to stop rising or even to sink. At the population maximum, gas vesicles render the cells neutrally buoyant.

Under certain conditions in nutrient-rich eutrophic lakes, gas-vacuolated cyanobacteria may become overbuoyant by an excess of gas vesicles. In this case, masses of cells appear at the surface and form a water bloom. *See* EUTROPHICATION.

Norbert Pfennig

Endospores

Endospores are resistant and metabolically dormant bodies produced by the gram-positive rods of *Bacillus* (aerobic or facultatively aerobic), *Clostridia* (strictly anaerobic), by the coccus *Sporosarcina*, and by certain other bacteria. Sporeforming bacteria are found mainly in the soil and water and also in the intestines of humans and animals. Some sporeformers are found as pathogens in insects; others are pathogenic to animals and humans. Endospores seem to be able to survive indefinitely. Spores kept for more than 50 years have shown little loss of their capacity to germinate and propagate by cell division.

The endospore appears as a light-refractile body inside another cell (sporangium), as shown in **Fig. 1**. Each sporangium produces one endospore with a characteristic size, shape, and position within the cell.

Structure and constituents. The mature spore has a complex structure which contains a number of layers. The outermost envelope, surrounding the spore, is called the exosporium and is a thin, membranous covering. Beneath the exosporium lies the spore coat, which is composed of several layers, largely of a protein nature; each is about 2–2.5 nm thick. Beneath these is a thin membrane which separates the spore coat from an area of low electron density called the cortex. The cortex is primarily a modified peptidoglycan structure. It occupies approximately half the volume of the spore. A wall and a thin membrane separate the cortex from the cytoplasm of the dormant spore. The internal structures of the spore appear similar to those seen in the cytoplasm of vegetative cells, as shown in **Fig. 2**.

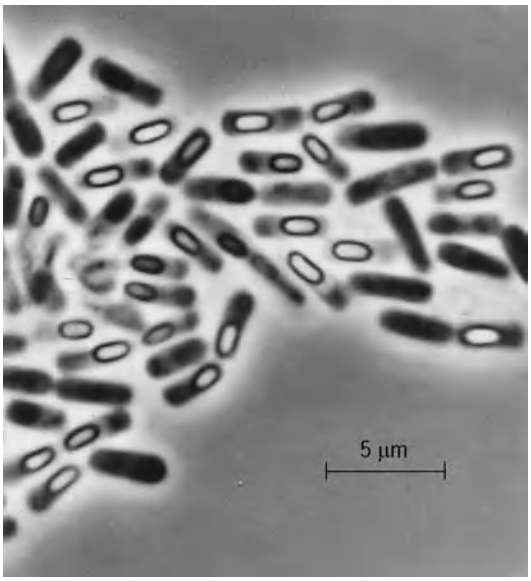

Fig. 1. Sporulating cells of *Bacillus cereus*. (*Photomicrograph by F. C. Fitz-James*)

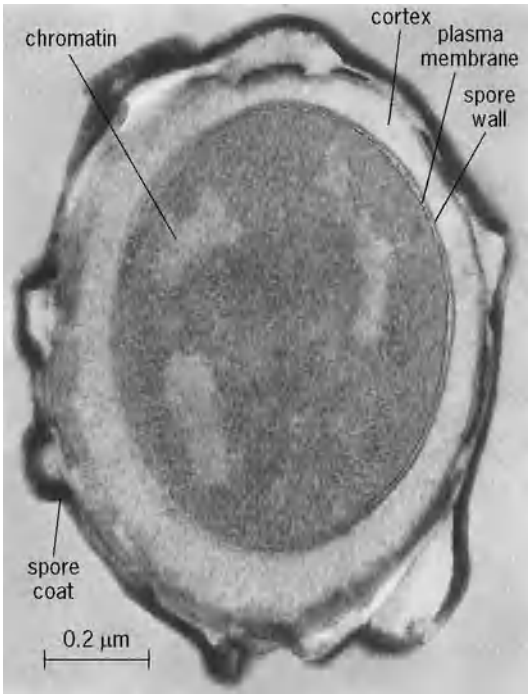

Fig. 2. Section of resting spore of *Bacillus megaterium*. (*Electron micrograph by C. F. Robinow*)

The unique properties of bacterial spores are their extreme resistance to heat, radiation from ultraviolet light and x-rays, organic solvents, chemicals, and desiccation. The most conspicuous chemical component is a chelating agent called dipicolinic acid (2,6-pyridine dicarboxylic acid), which constitutes 5–15% of the dry weight. This compound is absent from vegetative cells. Spores also differ from vegetative cells in containing higher levels of divalent metals and disulfide bonds and little, if any, free water.

Formation. The capacity of a bacterial cell to form a spore is under genetic control, although the total number of genes specific for sporulation is not known. The actual phenotypic expression of the spore genome depends upon a number of external factors. For each species of sporeforming bacteria, there exist optimum conditions for sporogenesis which differ from the optimal conditions for vegetative growth. These conditions include pH, degree of aeration, temperature, metals, and nutrients. Limitations in a variety of substances in the medium can initiate the process of sporulation. Sporulation is an ordered sequence of morphological and

biochemical events leading to the formation of a mature spore. Commonly observed stages of sporulation are, in order of appearance, antibiotic release, cortex synthesis, increased light refractility, dipicolinic acid synthesis, formation of coats, resistance to octanol, resistance to heat, and release of the spore from the mother cell.

Sporulation is strongly influenced by the carbon and nitrogen sources available. Metabolizable nitrogen compounds generally repress sporulation. Glucose, in the presence of an available nitrogen source, effectively represses the initiation of spore formation; in general, carbon and nitrogen sources which are rapidly metabolized favor vegetative growth, whereas carbon and nitrogen sources which are more slowly metabolized stimulate spore formation.

Breaking of dormancy. The three processes involved in the conversion of the spore into a vegetative cell are (1) activation (usually by heat or aging), which conditions the spore to germinate in a suitable environment; (2) germination, an irreversible process which results in the loss of the typical characteristics of a dormant spore; and (3) outgrowth, in which new classes of proteins and structures are synthesized so that the spore is converted into a new vegetative cell.

Germination. Germination is an irreversible process in which a number of simultaneous events take place, shortly after the exposure of activated spores to specific stimulants (amino acids, sugars, and nucleotides). Germination is accompanied by a swelling of the spore, either rupture or absorption of the spore coat, and loss of a number of typical properties of the spore. Among these last-mentioned events are a loss of resistance to environmental stress, a loss of refractility, an increase in permeability, a release of spore components (dipicolinic acid, calcium, spore peptides), and an increase in metabolic activity. As a whole, the process is degradative and probably involves a number of enzymes.

Outgrowth. Germination is followed by a period of biosynthetic activity called outgrowth. The endospore coat breaks, and the cortex disappears (degraded by hydrolytic enzymes). A new cell emerges from the spore coat and eventually matures into a vegetative form, as shown in **Fig. 3**. During this period, new proteins and structures characteristic of vegetative cells are synthesized. Outgrowth terminates at the time of cell division and return to vegetative growth. The conditions for outgrowth are usually different from those supporting germination. Germination and outgrowth have different optimum temperatures, and most spores need nutrients for outgrowth that are not required for germination.

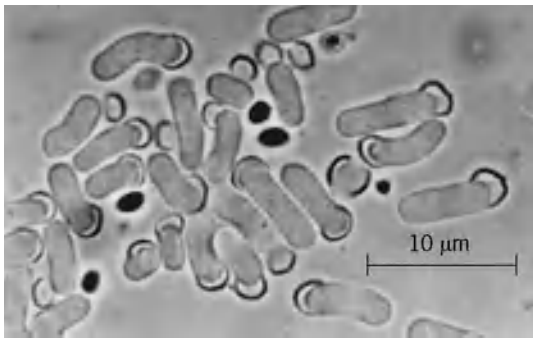

Fig. 3. Germination of spores of *Bacillus megaterium*. (*Photomicrography by F. C. Fitz-James*)

Microcycle sporulation. Spores may be germinated under nutrient-limited conditions. When faced with insufficient nutrients for proliferation, the new vegetative cell does not divide but instead enters the sporulation process, a shortcut leading to the formation of a mature spore.

Harlyn O. Halvorson; Keith Hutchison; Cristian Orrego

Appendages

On the basis of structure and function, bacteriologists recognize several types of bacterial appendages: flagella, pili, acellular stalks, and prosthecae.

Flagella. A bacterial flagellum has three parts, a long helical filament (about 0.01 micrometer in diameter), a short proximal hook (about 0.05 μm long), and a basal body, composed of a rod and a set of rings (about 0.02 μm in diameter) embedded in the cell wall and cytoplasmic membrane. The filament is a polymer or crystal composed of several thousand subunits per turn. Its shape can be changed by altering the structure of the subunit (the protein flagellin), the pH, ionic strength, temperature, or dynamic load. It dissolves in acids or in bases or when heated. Cultured in a solution containing pure monomeric flagellin, the filament grows in one direction at a constant rate. When attached to the living cell, it grows at the free end at a rate that decreases exponentially with length. Evidently, monomeric flagellin passes through the center of the filament and crystallizes on it at the distal end. The proximal hook is made up of a second protein, and the basal body is built up of another dozen or so different proteins. In a gram-positive bacterium, the basal body has two rings, an M ring embedded in the cytoplasmic membrane and an S ring found just outside this membrane. Common gram-negative bacteria have two additional rings, thought to serve as a bushing that carries the rod through the outer membrane. The M ring appears to be the rotor, the S ring the stator, the rod the drive shaft, the proximal hook a flexible coupling, and the filament a propeller.

Howard C. Berg

Pili. The pilus is also a proteinaceous appendage but differs from the flagellum in that it has a hollow core, is generally finer (ranges in width from 3 to 30 nm), and does not cause motility. Nonetheless, several important functions have been attributed to this structure. The most thoroughly documented function performed by pili is their role in bacterial sexual conjugation. The "male" cell of strains capable of conjugation produces a sex pilus that enables it to attach to an appropriate "female" cell containing the specific receptor site for the pilus. Only when cells of these two mating types are physically attached by the pilus can genetic material be transferred from the male to the female cell. *See* BACTERIAL GENETICS.

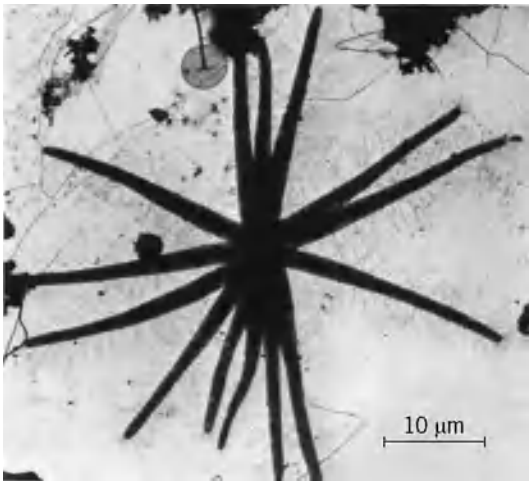

Fig. 4. Electron micrograph of a rosette-forming bacterium found in ponds and lakes.

Pili have also been implicated in the attachment of bacteria to unrelated organisms. For example, piliated strains of *Neisseria gonorrhoeae*, the bacterium which causes the venereal disease gonorrhea, are more frequently pathogenic than nonpiliated strains. It is thought that the piliated strains can use the pili to attach more strongly to host tissues. Certain viruses that infect bacteria attach to specific pili during the initial stages of infection.

Immotile aquatic bacteria commonly possess pili. **Figure 4** shows an unnamed aquatic bacterium that forms rosettes (the rosette shown had 14 cells which are joined together at a common center). Each cell has more than 100 very fine pili emanating from it. The role of these structures in such aquatic bacteria is not known.

Acellular stalks. Flagella and pili are too fine to be observed using ordinary light microscope techniques of observation. Most acellular stalks, on the other hand, are sufficiently wide to be seen by these classical procedures. For this reason it should not seem surprising that stalks were observed long before pili and flagella were seen. Indeed, the stalked bacterium *Gallionella ferruginea* was one of the first bacteria described. Interestingly, however, early investigators thought that the stalk was the bacterium because it was enormous compared to the small, bean-shaped cell which was readily dislodged from its position at the tip of the stalk. *Gallionella* is commonly found in iron springs, where the stalk becomes heavily encrusted with iron oxides that impart a rust-colored appearance. Although these bacteria have been grown in pure culture, the difficulties encountered in cultivating them have precluded extensive studies of their biology. Pure-culture studies indicate that the stalk is made up of several small fibrils, possibly pili, that are extruded during growth from the concave side of the cell. The iron appears to be precipitated in a sheath that surrounds the fibers.

The *Blastocaulis-Planctomyces* group also has members with acellular stalks. One example is shown in **Fig. 5**. This is a rosette containing nine ovoid cells borne at the tips of acellular stalks which are connected together at a common center. The cells reproduce by budding at the opposite, nonstalked pole of the cell. Note also that each cell has numerous pili in addition to the single stalk.

In both of the examples cited above, the appendage is an excretion of the cell and for this reason is termed extracellular or acellular.

Prosthecae. Unlike acellular stalks, the prostheca is an appendage that is actually part of the cell; that is, it is bounded by some or all of the layers of the cell envelope (cell wall and cell membrane). Frequently, species that have these appendages also have nonprosthecate cells that serve as stages in the life cycle of the organism. These nonprosthecate cells are usually motile by flagella and undergo predictable developmental stages during which prosthecae differentiate and daughter cells are formed. Thus, these bacteria are among the simplest unicellular organisms in which cellular developmental processes can be studied at the molecular level.

There are two hypotheses which have been advanced by workers in this field to explain the function of prosthecae: (1) These structures serve to act as "wings" by preventing the cells that have them from settling out of the water column in aquatic habitats; and (2) by increasing the membrane surface area of the organism, prosthecae enable it to take up nutrients more quickly in the dilute environments in which they reside.

To provide some insight into the nature of these bacteria, the life cycles of several of those most thoroughly studied are described below.

Caulobacter. The best-known genus of prosthecate bacteria is *Caulobacter*. These bacteria have a single prostheca, termed a stalk, that extends from one end of the cell (**Fig. 6**). *Caulobacter* cells undergo division at their nonstalked pole by binary fission to

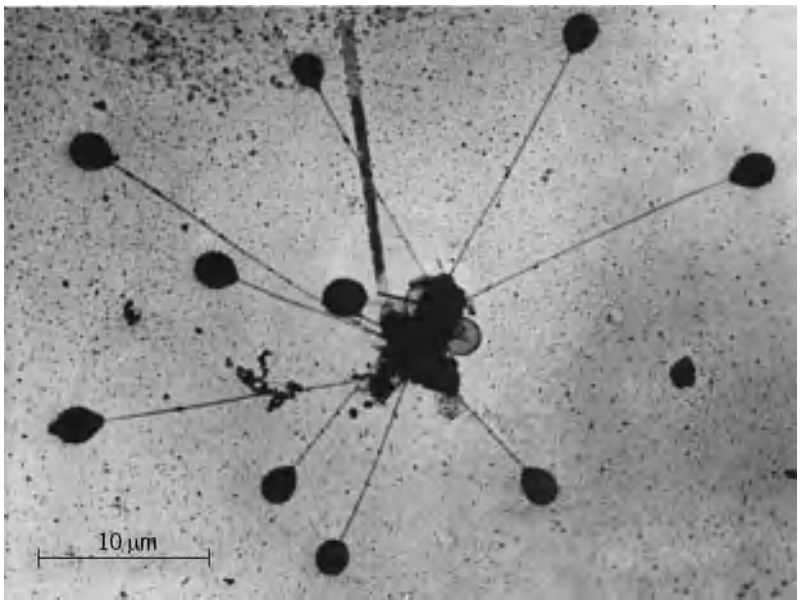

Fig. 5. Electron micrograph of a rosette of a budding bacterium of the *Blastocaulis-Planctomyces* group; note small bud on cell at lower left. Pili can be seen emanating from each cell.

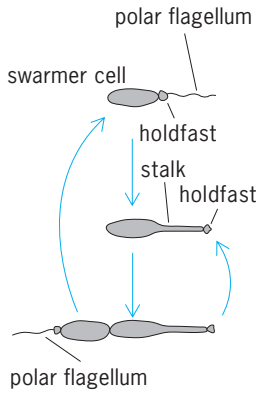

Fig. 6. Life cycle of *Caulobacter* species.

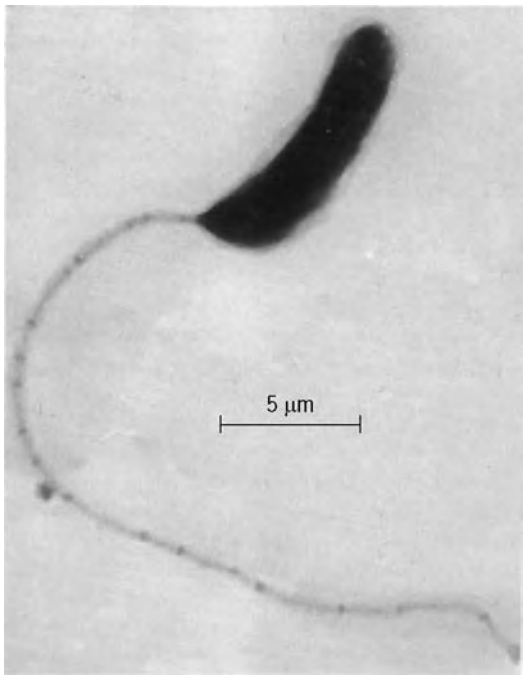

Fig. 7. Electron micrograph of a *Caulobacter* cell showing numerous crossbands on the prostheca; each crossband was formed when a daughter cell was produced.

produce a daughter cell that has a single flagellum. The daughter cell, which has no stalk, becomes motile and separates from the mother cell. After separation, the daughter cell, called a swarmer cell because of its motility, normally attaches to a solid substratum by a sticky holdfast material located at the base of the flagellum. In time the cell loses its motility and synthesizes a stalk at the flagellum-base position so that the holdfast is borne at the tip of the stalk. It is now a mature cell that elongates and produces swarmer cells of its own.

It has been discovered that the age of a *Caulobacter* cell can be estimated by counting the number of crossbands in the stalk of the cell. These crossbands can be seen when cells are observed with the electron microscope (**Fig. 7**). Apparently, each time a cell undergoes binary fission, the mother cell synthesizes a crossband in its stalk. Therefore the age of the cell can be estimated by counting these structures in the same manner that the annual rings of a tree can be used to determine its age. This is the only bacterium whose age can be determined by direct examination of the organism.

Hyphomicrobium. The genus *Hyphomicrobium* is distinctive because of its prosthecae and because of its division by budding. Like *Caulobacter*, it undergoes, for a bacterium, a rather complex life cycle including flagellated, nonprosthecate cells and prosthecate, nonmotile cells (**Fig. 8**). Newly formed oval buds are motile by subpolar flagella. Some species have holdfast material associated with the cell that permits them to attach to particulate material. Eventually the buds lose their motility and develop prosthecae, invariably from one of the cell ends. The tip of the prostheca enlarges to form a bud which separates from the mother cell and undergoes the same cycle. The mother cell, however, has several options available to it: (1) It may produce another bud from the same site at the tip of its prostheca; (2) it may produce another prostheca at the same or opposite pole of the cell and form a bud at its tip; or (3) it may produce a branch from its already existing prostheca and form a bud at its tip. Therefore, the life cycle of this bacterium is more complex than that of *Caulobacter*.

Prosthecomicrobium. Bacteria in the genus *Prosthecomicrobium* have approximately 20 conical prosthecae that extend in all directions from the cell, giving the organism the appearance of a bur of a cocklebur plant when observed in the microscope (**Fig. 9**). The appendages vary in length from one species to another, being as short as 0.2 μm or as long as 1.0 μm, or more. The bacteria reportedly divide by binary fission. Although both motile and nonmotile cells are found in some strains, not all strains have

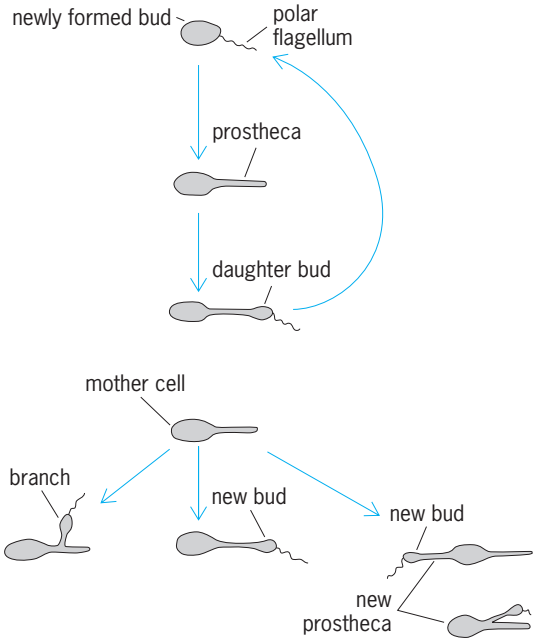

Fig. 8. Life cycle of *Hyphomicrobium* species.

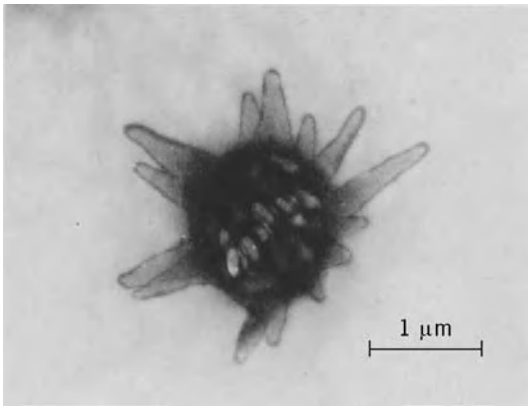

Fig. 9. Electron micrograph of *Prosthecomicrobium pneumaticum*. Note the 14 prosthecae extending from the cell, and the transparent gas vesicles inside the cell.

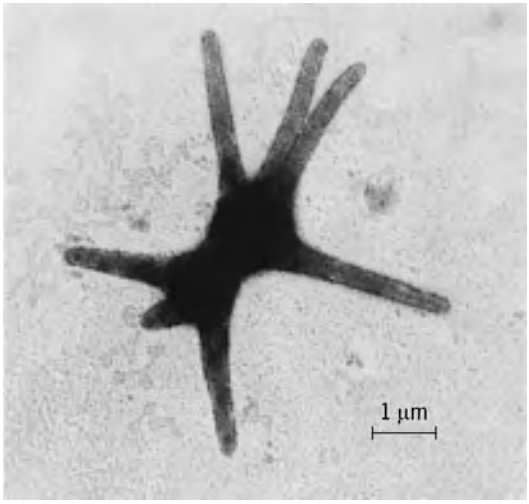

Fig. 10. Electron micrograph of *Ancalomicrobium adetum*. The cell has seven prosthecae.

motile stages, and in those that do, it is not known whether the motile stage is part of the life cycle, as in the case of *Caulobacter* and *Hyphomicrobium*.

Ancalomicrobium. Like *Prosthecomicrobium* cells, cells of the genus *Anacalomicrobium* have several prosthecae per cell, although the number rarely exceeds eight (**Fig. 10**). This nonmotile bacterium produces an outgrowth from one position on the cell surface. This outgrowth, or bud, differentiates to form from two to four prosthecae, each about 3 μm long. These appendages occasionally form branches analogous to those of *Hyphomicrobium*, but buds are not produced at the tips. James T. Staley

Motility

Motile bacteria swim or glide. Common gram-positive and gram-negative organisms swim by rotating flagellar filaments that project from the surface of the cell and extend several micrometers into the surrounding medium. Spirochetes, another group of gram-negative organisms, swim by rotating filaments that run between the protoplasmic cylinder and the outer membrane. Gliding bacteria, a third major group of gram-negative organisms, do not swim or have flagella but creep when in contact with solid surfaces.

Flagellar motion. Each flagellum is driven at its base by a reversible rotary motor; the filament rotates rigidly clockwise or counterclockwise; it does not wave or beat. To demonstrate this, a cell can be fixed to a glass slide by one of its flagellar filaments: the cell body spins alternately clockwise and counterclockwise. Mutants exist that spin in only one direction. Their motion has been followed for millions of revolutions; therefore, a bacterial flagellum does not wind up and unwind—it truly rotates. The flagella are not powered by the high-energy phosphate compounds that drive muscles, but rather by a proton flux. The passage of about a thousand protons from the outside to the inside of the cell carries the motor through one revolution. When connected to the proximal hook, the motor spins about 100 revolutions per second; when also connected to the filament, it spins about half as fast. When loads are light, the motor appears to run at constant speed; when loads are heavy, it runs at constant torque.

The style of locomotion of a bacterium depends on the size and shape of its body and its mode of flagellation. A rod-shaped cell with one or more flagella at either pole (monotrichous or lophotrichous flagellation, either monopolar or bipolar) swims forward or backward in the direction of its long axis. The body rotates (rolls) in a direction opposite to the direction of rotation of the flagella. Thrust generated by the flagellar filaments is balanced by viscous drag due to translation of the cell body, and torque generated by the filaments is balanced by viscous drag due to the roll.

If the cell body is corkscrew-shaped (as in *Spirillum*), the roll also contributes to the thrust. This is the primary mechanism for locomotion in most spirochetes (for example, the genus, *Spirochaeta*), which are long thin helical organisms with loosely attached outer membranes. The rotation of the filaments underneath this membrane is thought to cause the surface of the cell to move circumferentially. Viscous drag, in turn, causes the cell body to roll about its long axis, and the cell screws its way through the medium. This mode of locomotion is particularly effective in gel-like media. *Leptospira*, having a more tightly fitting outer membrane, generates torque by gyrating its ends—the anterior end in a spiral configuration, the posterior end in a hook configuration; additional thrust is gained from propagation of the anterior spiral wave.

Rod-shaped cells with flagella arising at random points on their sides (peritrichous flagellation) swim as in a random walk. When their flagella spin counterclockwise, they coalesce into a synchronous bundle that pushes the cell steadily forward: the cell is said to run. When the flagella spin clockwise, the bundle flies apart, and the flagella turn independently, moving the cell this way and that in a highly erratic manner: the cell is said to tumble. These modes alternate, and the cell executes a three-dimensional random walk. A variety of species swim in this way, including

Bacillus subtilis, *Salmonella typhimurium*, and *Escherichia coli*.

Taxis. Bacteria change the direction in which they swim in response to changes in their environment: they accumulate in regions that they find favorable. Different cells respond to light, heat, oxygen, and a variety of chemicals; and are said to be phototactic, thermotactic, aerotactic, and chemotactic, respectively. Cells of *E. coli*, for example, swim into capillary tubes filled with dilute solutions of simple sugars and amino acids. They do this by biasing their random walk. Runs that happen to carry a cell up the gradient are extended, while those that happen to carry it down the gradient are not. The cells monitor concentrations as a function of time; they compare the occupancy of specific receptors over the past second or so with the occupancy a few seconds before that. The memory required for this comparison involves receptor carboxymethylation. If current receptor occupancy exceeds past occupancy, as indicated by a relatively low level of methylation, the probability that the flagella spin counterclockwise increases and the probability that they spin clockwise decreases.

Gliding. Cells of some gliding bacteria, for example, *Myxococcus*, aggregate to form fruiting bodies. This behavior requires intracellular communication; the way in which one cell glides depends on its contacts with other cells. *Myxococcus* glides very slowly, about 2 μm per minute; other gliding bacteria, for example, *Cytophaga*, glide more rapidly, up to 2 μm per second. The organelles of locomotion for these bacteria have not been identified. Howard C. Berg

Bibliography. J. Adler, The sensing of chemicals by bacteria, *Sci. Amer.*, 234(4):40–47, 1976; R. P. Burchard, Gliding motility of prokaryotes: Ultrastructure, physiology and genetics, *Annu. Rev. Microbiol.*, 35:497–529, 1981; E. Canale-Parola, Motility and chemotaxis of spirochetes, *Annu. Rev. Microbiol.*, 32:69–99, 1978; I. I. Gunsalus and R. Y. Stanier (eds.), *The Bacteria*, 7 vols., 1960–1979; G. L. Hazelbauer and S. Harayama, Sensory transduction in bacterial chemotaxis, *Int. Rev. Cytol.*, 81:33–70, 1983; J. G. Holt (ed.), *The Shorter Bergey's Manual of Determinative Bacteriology*, 8th ed., 1977; J. M. Parry et al., *A Colour Atlas for the Identification of Bacillus Species*, 1983; E. M. Purcell, Life at low Reynolds number, *Amer. J. Phys.*, 45:3–11, 1977; R. Y. Stanier et al., *An Introduction to the Microbial World*, rev. ed., 1979; A. Sussman and H. O. Halvorson, *Spores: Their Dormancy and Germination*, 1966.

Bacterial genetics

The study of gene structure and function in bacteria. Genetics itself is concerned with determining the number, location, and character of the genes of an organism. The classical way to investigate genes is to mate two organisms with different genotypes and compare the observable properties (phenotypes) of the parents with those of the progeny. Bacteria do not mate (in the usual way), so there is no way of getting all the chromosomes of two different bacteria into the same cell. However, there are a number of ways in which a part of the chromosome or genome from one bacterium can be inserted into another bacterium so that the outcome can be studied. *See* GENETICS.

Characteristics of bacteria. The ways of classifying organisms are rapidly changing. Similarities in nucleotide sequences, particularly of the ribosomal ribonucleic acid (rRNA) sequences, have revealed new aspects of the evolutionary tree. All organisms have diverged from a common ancestral prokaryote whose precise location in the evolutionary tree is unclear. This has resulted in three primary kingdoms, the Archaebacteria, the Eubacteria, and the Eukaryotae. All bacteria are prokaryotes, that is, the "nucleus" or nucleoid is a single circular chromosome, without a nuclear membrane. Bacteria also lack other membrane-bounded organelles such as mitochondria or chloroplasts, but they all possess a cytoplasmic membrane. Most bacteria have a cell wall that surrounds the cytoplasmic membrane, and some bacteria also contain an outer membrane which encompasses the cell wall. Duplication occurs by a process of binary fission, in which two identical daughter cells arise from a single parent cell. Every cell in a homogeneous population of bacterial cells retains the potential for duplication. Bacteria do not possess the potential for differentiation (other than spore formation) or for forming multicellular organisms. *See* ARCHAEA; BACTERIA; PROKARYOTAE; RIBONUCLEIC ACID (RNA).

One of the most frequently used organisms in the study of bacterial genetics is the rod-shaped bacillus *Escherichia coli*, whose normal habitat is the colon. Conditions have been found for growing *E. coli* in the laboratory on defined media, and it is by far the best understood of all microorganisms. The single circular chromosome of *E. coli* contains about 4.5×10^6 base pairs, which is enough to make about 4500 average-size genes (1000 base pairs each). The relative positions of almost half of these genes are known. In regions where mapping studies are reasonably complete, the impression is obtained of an efficiently organized genome. Protein coding regions are located adjacent to regulatory regions. There is no evidence for significant stretches of nonfunctional deoxyribonucleic acid (DNA), and there is no evidence for introns in the coding regions. The coding regions from eukaryotic cells frequently contain introns, that is, regions that are removed by splicing the messenger RNA (mRNA) before it is translated into protein. Very little repetitive DNA exists in the *E. coli* chromosome other than the seven sequence-related rRNA genes that are dispersed at different locations on the chromosome. *See* CHROMOSOME; DEOXYRIBONUCLEIC ACID (DNA); GENETIC CODE.

Frequently, genes with a related function are tightly clustered. This clustering results from the fact that genes with a related function are commonly transcribed as a single polygenic (polycistronic) mRNA molecule. A chromosomal region transcribed in this

way is called an operon. It contains a single site at one end called the promoter, which is where RNA polymerase binds just prior to the initiation of transcription. *See* OPERON.

Genetic methodology. The first step in performing genetic research on bacteria is to select mutants that differ from wild-type cells in one or more genes. Then crosses are made between mutants and wild types, or between two different mutants, to determine dominance-recessive relationships, chromosomal location, and other properties. Various genetic methods are used to select bacterial mutants, antibiotic-resistant cells, cells with specific growth requirements, and so on.

Selection for bacterial mutants. Spontaneous mutation frequencies in *E. coli* vary from about 10^{-7} to 10^{-9} mutations per genome per generation. Since it is possible to plate 10^7 to 10^9 bacteria on a single standard-size growth plate, a few growth plates should contain a number of mutants of the desired type. The next consideration is to find a suitable selection procedure˜ for detecting and isolating the desired mutants from a large, principally wild-type population. The two most popular selection techniques involve selecting for mutants with altered antibiotic resistance or nutritional requirements. *See* MUTATION.

Selection for antibiotic-resistant cells. Ordinary wild-type cells are sensitive to a number of bacteriostatic and bacteriocidal factors, including antibiotics, toxins produced by certain plasmid-containing cells, and bacterial viruses (bacteriophages). Mutants that are resistant to antibiotics are very useful in the analysis of biochemical pathways because the sites of action, that is, the target sites of such antibiotics, are often highly specific. For example, the target site for streptomycin action is the ribosome, as evidenced by the fact that streptomycin-resistant mutants usually have an altered ribosomal protein. *See* ANTIBIOTIC; BACTERIOPHAGE; DRUG RESISTANCE; PLASMID; TOXIN.

The usual procedure for isolating antibiotic-resistant mutants is simple. Approximately 10^8 cells grown in a liquid medium are spread onto a semisolid agar plate containing growth medium and the antibiotic. Only a few cells will survive and grow on the surface of the agar. These will give rise to clones, that is, a population of genetically identical cells. *See* CULTURE.

Selection for cells with specific growth requirement. Wild-type *E. coli* cells that can grow on minimal medium, such as one containing only glucose ammonium chloride, and other salts, are called prototrophs. By contrast, auxotrophs are cells that require supplementation of the minimal medium with one or more factors. The penicillin enrichment technique can be used to isolate an auxotroph by selecting against the prototroph. For example, a tryptophan auxotroph (trp^-) will grow on glucose-minimal medium to which the amino acid tryptophan is added. To select for a trp^- strain, a large number of cells are grown on glucose-minimal medium in the presence of penicillin. Penicillin is a bacteriocidal agent that kills only growing cells by inhibiting one or more reactions in cell-wall synthesis. The trp^- cells, or for that matter any auxotrophs that cannot grow on glucose-minimal medium, will survive penicillin treatment even though they cannot grow in its presence. In fact, they survive precisely because they cannot grow. The wild-type cells that do grow in this medium will be killed by the penicillin treatment. After a suitable period, usually overnight, the cells are removed from the penicillin by centrifugation. They are resuspended in growth medium containing tryptophan and spread on an agar plate. Under these conditions the trp^- auxotrophs, which have survived the exposure to penicillin on the minimal medium, will grow.

Back mutations or revertants from an auxotrophic to a prototrophic state may be directly selected for by growth on minimal medium. Starting from a large population of auxotrophs, the only cells capable of forming clones on minimal medium would be revertants. Revertants possess either the exact reverse mutation or another mutation that compensates for the initial mutation and permits growth in the absence of added amino acid or other growth requirements.

Regulatory gene mutations. Certain genes that have the function of modulating the expression of other genes are known as regulatory genes. Mutations that affect the action of regulatory proteins are of two types: those that occur in the genes that encode the regulatory proteins, and those that affect the genetic loci where the regulatory protein interacts to modulate the level of gene expression. *See* PROTEIN.

In the case of tryptophan synthase, which is one of the enzymes required for tryptophan synthesis, a detailed examination has shown that the genes for tryptophan synthase are part of a multigene cluster, or operon. Expression of the operon is regulated at the promoter which is located at one end of the *trp* operon. In the *trp* operon the promoter overlaps with another site known as the operator. Whereas the promoter serves as the locus for initial binding of RNA polymerase and initiation of transcription, the operator serves as the binding site for a so-called repressor protein. When the repressor protein is bound to the operator locus, it prevents the RNA polymerase from binding to the promoter locus, thereby inhibiting the initiation of transcription and gene expression. The repressor is encoded by an unlinked gene located at a distinct site called *trpR*.

Mutations that lead to decreased or enhanced expression of the *trp* operon have been observed and analyzed in great detail. Such mutations result from changes in the DNA at the promoter-operator locus or from changes in the structure of the repressor protein, but they never affect the structure of the enzymes encoded by the operon, only their amounts. Some regulatory gene mutations cause overproduction and some cause underproduction of gene products. This is the hallmark of a mutation that influences the functioning of a regulatory protein or regulatory factor–binding site; it affects the quantity but not the quality of other gene products. Furthermore, regulatory gene mutations are frequently pleiotropic, that is, they influence the rate

of synthesis of several gene products simultaneously. *See* GENE ACTION.

Induced mutagenesis. Frequently, geneticists want to increase the number or types of mutants that can be obtained as a result of spontaneous mutagenesis. In such instances, they treat a bacterial population with a mutagenic agent to increase the mutation frequency. This is called induced mutagenesis. The simplest techniques of induced mutagenesis involve measured exposure of the bacteria to a mutagenic agent, such as x-rays or chemical mutagenic agents. Such procedures have a general effect on the increase in the mutation rate. *See* MUTAGENS AND CARCINOGENS.

More sophisticated procedures involve isolating the gene of interest and making a change in the desired location. This is called site-directed mutagenesis. The goal is usually to determine the effects of a change at a specific gene locus. The gene in question is isolated, modified, and reinserted into the organism. Discrete alterations can be made in a variety of ways on any DNA in cell-free culture, and the effect of such alterations can be subsequently tested in the organism. Most of the techniques used rely upon a site-specific restriction enzyme at some stage. The most popular procedures for directed mutagenesis are all variations of a particular procedure. This procedure begins with the limited degradation of one DNA chain of a duplex to yield a gapped molecule. The gapped DNA is annealed to a synthetic oligonucleotide containing discrete differences from the original DNA. After completing the ligation process, the resulting heteroduplex can be reinserted into the organism, and by suitable selection procedures mutants can be isolated that contain specific changes dictated by the sequence in the synthetically produced oligonucleotide. *See* GENETIC ENGINEERING.

Gene mapping. Bacteria do not mate to form true zygotes, but they are able to exchange genetic information by a variety of processes in which partial zygotes (merozygotes) are formed. The first type of genetic exchange between bacteria to be observed was transformation. Naturally occurring transformation involves the uptake of DNA. This phenomenon is observed only for a limited number of bacterial species and is a relatively difficult technique to use for gene manipulation. For these reasons and because more powerful approaches to gene mapping were discovered soon afterward, transformation has received relatively narrow use as a means for gene mapping. In 1946 direct chromosomal exchange by conjugation between *E. coli* cells was discovered by J. Lederberg and E. Tatum, and in 1951 transduction, the virus-mediated transfer of bacterial genes, was discovered. Both conjugation and transduction provide facile, generally applicable methods for moving part of the bacterial chromosome from one cell to another. The discovery of bacterial transposons in the 1970s has been useful in marking and mobilizing genes of interest. The purely genetic approaches to mapping have been supplemented by the biochemical approaches of hybrid plasmid construction and DNA sequence analysis. A detailed genetic and physical map of the entire bacterial chromosome is within grasp. *See* GENETIC MAPPING; TRANSFORMATION (BACTERIA).

Conjugation. The naturally occurring exchange of genetic information in *E. coli* bacteria takes place between cells of different mating types. Donor cells are defined as bacteria capable of transferring genes, and recipient cells are defined as bacteria capable of accepting and integrating such genes. The difference between a donor and a recipient cell is due to a plasmid known as F factor. The F factor is a circular DNA molecule with a molecular weight of about one-fortieth that of the bacterial host chromosome. Donor cells (F^+ cells) arise from bacteria that contain the F factor. Within an F^+ population, cells that transfer bacterial genes represent distinct variants. These variants, called Hfr's, have integrated the F factor into the host chromosome, and the site of integration determines the order of transfer of bacterial genes when an Hfr is mated with an F^- cell. The order of transfer of markers may be determined by interrupting the mating process at different times after mixing the two cell types. The so-called interrupted mating procedure has been carried out on many different Hfr's that have integrated the F factor at different locations and in either location. The results from such conjugations have permitted a rough mapping of the bacterial chromosome. Indeed, it was studies of this sort that led to the prediction that the bacterial chromosome was circular, a prediction that was verified by physical measurements many years later.

On rare occasions, the integrated F factor in the Hfr can reversibly excise to give back the original F^+ cell. On even rarer occasions, the F factor excises imperfectly, taking with it a part of the bacterial chromosome. Such F′ factors have been used as a valuable source of discrete segments of the bacterial chromosome; they can readily be transferred to F^- cells. They have been most valuable in determining dominance-recessive relationships for alleles of the same genes.

Transduction. Transfer of bacterial segments can be done by sexduction, that is, the transfer of F′ particles, but a far more convenient procedure for most applications involves transport by bacteriophages. Transfer of bacterial DNA by bacteriophages is called transduction. Transduction has provided the most valuable technique for bacterial gene mapping and for the isolation of discrete segments of the bacterial chromosome. A number of bacteriophages are used in transduction. They have different properties, and each one has particular advantages for gene manipulation. Some bacteriophages, such as the λ bacteriophage, transfer only specialized regions of the bacterial genome. For this reason, λ is referred to as a specialized transducing virus. Other bacteriophages, such as P1, which can transfer virtually any part of the bacterial genome are referred to as generalized transducing viruses. To obtain a P1 preparation that carries parts of a particular bacterial chromosome, the bacterial strain is infected with P1. The infection results in the production of a large number of

normal P1 virus particles and a small mixed population of P1 transducing particles. The P1 transducing particles carry segments of the bacterial genome instead of the viral genome. When they infect a bacterial cell, the segment of bacterial genome enters the cell.

For a high-resolution mapping of bacterial genes, generalized transduction is a virtually indispensable tool. Bacterial genes that are transduced into *E. coli* by P1 transducing particles cannot replicate or survive unless they become integrated into the host genome. In so doing, they replace the homologous segment in the bacterial genome. Integration requires a minimum of two crossovers. In three-factor crosses, the middle marker may be identified by the least frequent recombinant class. Exclusive exchange of the middle marker requires two crossovers also. It occurs with a lower than average probability because the crossover points are restricted so that only the middle marker exchanges. *See* CROSSING-OVER (GENETICS); TRANSDUCTION (BACTERIA).

Transposons. Chromosomes might be thought of as being composed of two types of genetic elements: those that maintain a fixed location in a larger stable structure, and those that are designed to move from one location to another. The latter type is known by the general term mobile genetic element. Transposons are a class of mobile genetic elements that are commonly found in bacterial populations. Most transposons have three characteristic properties: they carry readily selectable marker genes; they insert at many different locations in the bacterial chromosome; and they are mutagenic in regions adjacent to the site of insertion. *See* TRANSPOSONS.

Expression of bacterial genes. At any given time, only a small percentage of the *E. coli* genome is being actively transcribed. The remainder of the genome is either silent or being transcribed at a very low rate. When growth conditions change, some active genes are turned off and other, inactive genes are turned on. The cell always retains its totipotency, so that within a short time (seconds to minutes), and given appropriate circumstances, any gene can be fully turned on. The maximal activity for transcription varies from gene to gene. For example, a β-galactosidase gene makes about one copy per minute, and a fully turned-on biotin synthase gene makes about one copy per 10 min. In the maximally repressed state, both of these genes express less than one transcript per 10 min.

The level of transcription for any particular gene usually results from a complex series of control elements organized into a hierarchy that coordinates all the metabolic activities of the cell. For example, when the rRNA genes are highly active, so are the genes for ribosomal proteins, and the latter are regulated in such a way that stoichiometric amounts of most of the ribosomal proteins are produced. When glucose is abundant, most genes involved in processing more complex carbon sources are turned off in a process called catabolite repression. If the glucose supply is depleted and lactose is present, the genes involved in lactose breakdown (catabolism) are expressed. In *E. coli* the production of most RNAs and proteins is regulated exclusively at the transcriptional level, although there are notable exceptions. Rapid response to changing conditions is ensured partly by a short mRNA lifetime, on the order of 1 to 3 min for most mRNAs. Some mRNAs have appreciably longer lifetimes (10 min or more), and the consequent potential for much higher levels of protein synthesis per mRNA molecule. These atypical mRNAs, at least in some instances, also may be subject to translational control. Finally, the fine-level control for any particular enzyme system is subject to regulation by activators or inhibitors directly modulating the enzyme activities.

Rel gene. The ribosome is the site of protein synthesis. Bacterial ribosomes are composed of 3 rRNAs and about 50 ribosomal proteins. *Escherichia coli* cells, under conditions of rapid growth, contain about 10,000 ribosomes. There are seven operons located at distinct and separate points on the bacterial chromosome which make similar rRNA transcripts. The transcription of the rRNA operons is controlled by the compound guanosine tetraphosphate, ppGpp, known as magic spot. This compound has pyrophosphates at the 3′ and the 5′ carbons of the ribose sugar. Under conditions of amino acid deprivation, the level of magic spot rises. It combines with the RNA polymerase, changing the structure of the RNA polymerase so that the latter has a low affinity for the promoters of the rRNA genes. The result is that less rRNA is made. The ppGpp compound has a similar effect on the expression of genes for the rRNA proteins. The mRNAs for the ribosomal protein genes are also controlled at the translation level by a feedback inhibition mechanism. These proteins normally combine with the rRNA to make ribosomes. However, if there is an excess of ribosomal proteins over the available rRNA, the ribosomal protein will combine with its own message so as to prevent further translation of the message. In this way the amount of ribosomal protein never reaches large excess but always is synthesized in slight excess over the available level of rRNA. *See* RIBOSOMES.

Gene regulation for DNA repair. The cell possesses an intricate system (the *din* genes) for regulating the synthesis of genes involved in DNA repair. The *din* genes are under the negative control of the lex repressor. Unlike most repressors, including the lac repressor, the lex repressor action is not reversibly controlled by complexing with a small molecule. Rather, it is broken down by a highly specific protease. This protease is carried by the recA protein and is not active under normal growth conditions. However, when the growing cells are exposed to ultraviolet light or other agents that damage DNA, the recA protease becomes active. It is believed that this is due to a complex formed between the recA protease and a DNA fragment. Once the protease becomes active, it breaks down all of the lexA protein in the cell. The genes repressed by lexA then become active and stay active until the DNA damage is repaired. After the DNA damage has been repaired, the recA protease

returns to its normal inactive state, and the concentration of lexA repressor reaches a level where it silences the *din* genes and everything returns to normal. *See* BACTERIAL PHYSIOLOGY AND METABOLISM.

Geoffrey Zubay

Bibliography. C. R. Woese, Bacterial evolution, *Microbiol. Rev.*, 51:221–271, 1987; G. Zubay, *Genetics*, 1987.

Bacterial growth

The processes of both the increase in number and the increase in mass of bacteria. Growth has three distinct aspects: biomass production, cell production, and cell survival. Biomass production depends on the physical aspects of the environment (water content, pH, temperature), the availability of resources (carbon and energy, nitrogen, sulfur, phosphorus, minor elements), and the enzymatic machinery for catabolism (energy trapping), anabolism (biosynthesis of amino acid, purines, pyrimidines, and so forth), and macromolecular synthesis [proteins, ribonucleic acid (RNA), and deoxyribonucleic acid (DNA)]. Cell production is contingent on biomass production and involves, in addition, the triggering of chromosome replication and subsequent cell division. The cells may or may not separate from each other, and the division may partition the cell evenly or unevenly. Alternatively, growth may occur by budding (unequal division). Most cells so produced are themselves capable of growing and dividing; consequently, viability is usually very high when growth conditions are favorable. Moreover, in many cases the incidence of death is surprisingly low in the absence of needed nutrients. Many bacteria differentiate into resistant resting forms (such as spores); others may simply reduce their rate of metabolism and persist in the vegetative state for long times.

Exponential balanced growth. If bacteria have been growing under well-mixed, low-cell-density, constant conditions for many generations, then cell-cell interactions are negligible and all aspects of the growth process come to have rates balanced so that all cell components double in a doubling time. However, because bacteria can grow so fast, it is usually necessary to keep diluting the culture to achieve such balanced growth in order to prevent the population from entering the stationary phase. The mathematics of balanced growth are readily expressed. Let B stand for the amount of bacteria per milliliter of culture. Clearly, if the bacteria behave independently of each other, growth would be twice as fast if B were twice as large. Therefore, as shown in Eq. (1),

$$\frac{dB}{dt} = rB \tag{1}$$

the rate of production is proportional to B. Thus bacteria grow in an autocatalytic way because the product of the reaction is the catalyst for more growth. The rate constant r is usually called the specific growth rate. (The symbol μ is used in certain subfields of microbiology.)

Integration of Eq. (1) between the limits $B = B_0$ and $B = B_t$, where B_0 is the cell concentration at time zero and B_t is the cell density at time t, yields Eq. (2), where e is the base of natural logarithms.

$$B_t = B_0 e^{rt} \tag{2}$$

Taking natural logarithms of both sides results in Eq. (3). The specific growth rate can be estimated

$$\ln B_t = \ln B_0 + rt \tag{3}$$

from plots of data according to Eqs. (2) or (3), and sophisticated statistical procedures can be employed. Often, the doubling time t_d is of interest. Setting B_t equal to $2B_0$ in Eq. (2) yields Eq. (4) or (5). The latter

$$2 = e^{rt_d} \tag{4}$$

$$\ln 2 = rt_d \tag{5}$$

provides the way to convert values of r into values of t_d, and vice versa.

Evidently, many bacteria have been selected in nature for the ability to grow rapidly in their natural habitat. Doubling time may be many hours; however, wild-type strains of *Escherichia coli* under optimal conditions can double once every 16 min ($t_d = 16$ min, $r = 2.6$/h). At a rate this fast, one cell would grow in one day into 10^{27} organisms. It is no wonder that exponential growth lasts only a short time.

Culture cycle. After an extended culturing period, cells have exhausted their external and internal resources and are usually in the stationary phase. If a culture in this phase is diluted in fresh nutrient-rich medium at a temperature at which growth can occur (**Fig. 1**), there is often a lag phase during which the cell number does not increase and cells are preparing

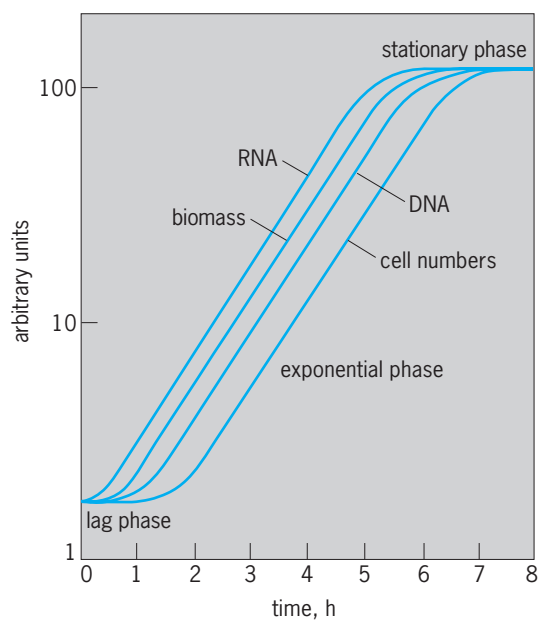

Fig. 1. Growth cycle, showing the consequence of diluting a stationary-phase culture into fresh medium.

for active growth. During the lag phase, cells are accumulating some metals and detoxifying others and synthesizing cofactors (containing vitamins). They are also synthesizing ribosomes and the factors necessary for protein synthesis and the enlargement of the cell. After some time, chromosomal DNA synthesis is initiated, followed by cell division. Eventually, the exponential balanced phase is reached. Some nutrient becomes limiting or some toxic product accumulates, and growth again slows and stops. If this transition is relatively gradual or if the cells have internal reserves, the bacteria may be able to respond morphologically to the impending "hard times" by forming spores or cysts, which are characteristic of the stationary phase. *Escherichia coli* continues DNA replication and cell division after the rate of protein synthesis slows; this results in more but smaller cells. In some cases, ribosomes and certain proteins are torn down to make components of the stationary cells. In early stationary-phase cultures, the vast majority of the countable cells are viable in the sense that they can give rise to a colony. Subsequently, a phase occurs in which viability is lost and the biomass is reduced by metabolism and autolysis. In nature, bacteria would be subject to digestion by protozoans and metazoans, or would be killed by physical conditions or host immunity mechanisms. *See* CELL CYCLE.

Dependence of growth rate on substrate. Many bacteria have very complex nutrition, but others grow on simple media. For the latter types to grow in varying environments, they must be capable of many molecular transformations to make all the cell's building blocks from available resources. However, if the concentration of a single irreplaceable nutrilite (nourishing compound) is lowered, eventually the growth rate must fall. An approximate kinetic description (called the Monod model) used extensively in microbiology, ecology, and biotechnology is drawn from an analogy to enzyme kinetics [Eq. (6)]. K_m is the concentration supporting half-

$$r = r_{\max} \frac{S}{K_m + S} \tag{6}$$

maximal growth and S is the concentration of of substrate, or, in this case, necessary nutrilites.

It follows from this relationship that r is proportional to S at low S and is independent at high S, when $r = r_{\max}$. In the intermediate region, studies have shown that a better approximation is given by Eqs. (7). S_c is the critical concentration where

$$\begin{aligned} r &= \alpha S \qquad S \le S_c \\ r &= r_{\max} \quad S \ge S_c \end{aligned} \tag{7}$$

growth becomes independent of nutrilite concentration, and $\alpha = r_{\max} S_c$.

A more elaborate form is needed in other cases, and additional modifications are needed for oligotrophs, that is, organisms living and growing in very dilute environments. Oligotrophs have very effective uptake systems so that they can scavenge scarce nutrients; sometimes they have extra appendages (prosthecae) to give them extra surface. But even *E. coli*, an inhabitant of the mammalian gut, has the capability of taking up nearly every molecule of glucose that impinges on its cytoplasmic membrane. For organisms capable of coping with both abundance and scarcity, the uptake systems must be adequate when the concentration is low, but would be excessive when the substrate is not limiting. This is one of the reasons that Eqs. (7) are, in many cases, a better representation of cellular growth than Eq. (6).

Alternative carbon resources. When presented with two suitable carbon (and energy) sources, and under certain circumstances, a bacterium may elect to consume one compound and commence synthesizing the enzymes and other metabolic machinery needed to use the second substance, only when the first one has been depleted. For example, *E. coli* can be given both glucose and lactose (milk sugar). After the glucose is consumed, there is a lag in growth while β-galactosidase and galactoside permease are formed, allowing renewed growth. β-Galactosidase is an enzyme that splits lactose into glucose and galactose; galactoside permease is a membrane protein that functions in the entry of lactose into the cell. This growth pattern is called diauxic (double-fed) growth.

Catabolite repression, a concept in molecular biology, is now understood in molecular terms through the study of diauxic growth. **Figure 2** shows the DNA coding for the regulatory and structural part of *E. coli* lactose metabolism. In the presence of high levels of cyclic adenosine monophosphate (cAMP) resulting from glucose shortage, the catabolite activator protein (CAP protein) binds the "upstream" region of the regulatory site of the *lac* operon about 60 base pairs before the site of transcription initiation. This allows the promoter region to increase the binding and functioning of RNA polymerase. RNA polymerase synthesizes a polycistronic message that specifies β-galactosidase and galactoside permease. More directly favoring lactose metabolism is the effect of a lactose derivative binding to the repressor and unblocking the operator region.

This regulatory system is sufficiently complex that it functions appropriately and sometimes quite differently under a series of other circumstances. When two substrates are each present in low concentration and when one serves as a nitrogen source as well as a carbon and energy source, metabolism of both types can occur simultaneously. *See* ALLOSTERIC ENZYME; ENZYME INHIBITION; GENE; GENE ACTION; OPERON.

Cell division. The small size of bacteria provides a high surface-to-volume ratio, which assures effective uptake of substances from the environment. Since small size increases the number of propagules per unit of biomass, cell division is an important aspect of growth. Seemingly, it is not necessary for cell division to be very precise, because the important thing is to keep the living "packages" small. However, in some cases cell division is quite accurate; the bacteria divide nearly in the middle and, during balanced growth, nearly at the same size. Except in laboratory-created mutants, cells almost never form without receiving a complete chromosome. But even in the

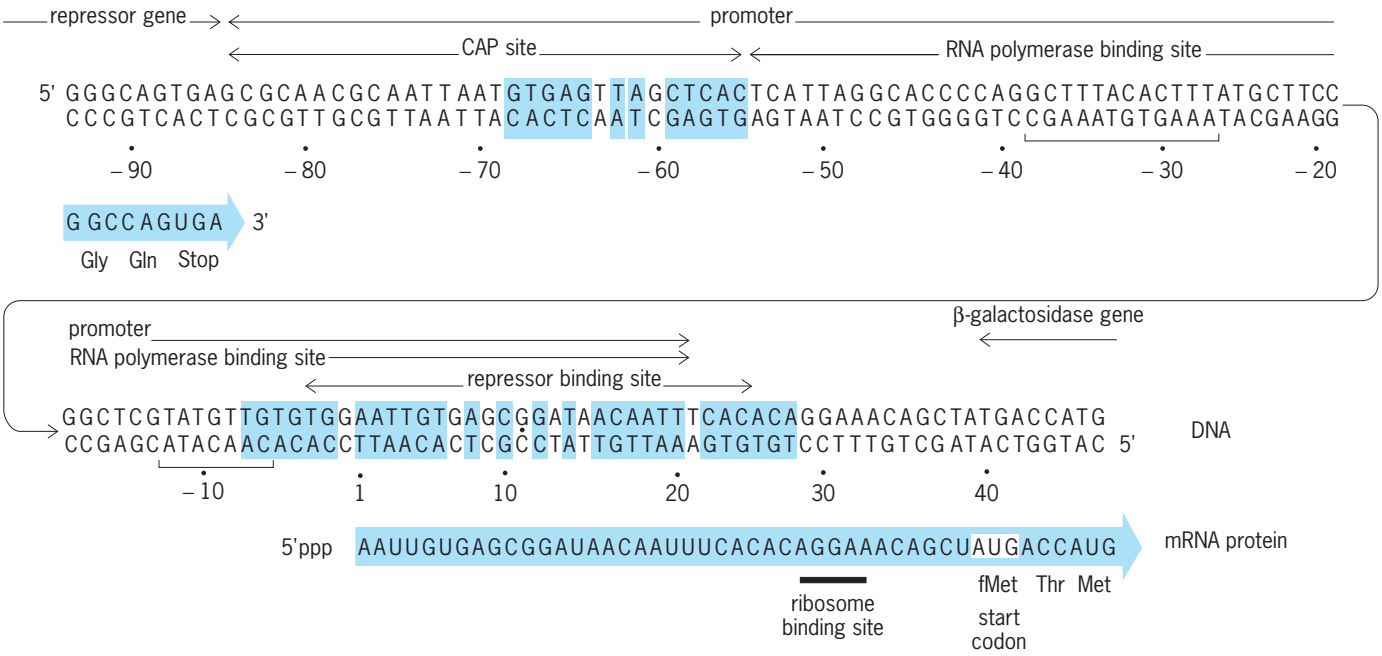

Fig. 2. The *lac* operon. The genetic region of *Escherichia coli* K12 is shown indicating the region interacting with effector molecules and enzymes. The promoter region contains the part of the DNA binding to the CAP protein, the RNA polymerase, and a portion of the repressor-binding site. The part binding to repressor is called the operator. Various elements of symmetry are noted with the boxes. The underlined regions in the RNA polymerase–binding site are functionally important in initiating transcription. On the upper right, the tail end of the repressor gene is shown. On the lower right, the messenger RNA and the first three translated codons of β-galactosidase are shown.

most precise cases of cell division there is some variation in size and evenness at division, which translates into a larger variability of the age that cells attain before they divide. The doubling time is a suitable average of the individual division times of cell cycles.

Continuous culture. Bacteria in nature are usually engaged in activities similar to the culture cycle described above. A spore or stationary cell finds its way to a new habitat, grows, and divides until the clone has utilized the available resources as effectively as its metabolic capabilities permit. Later, as spores or stationary-phase cells become dispersed to new habitats, the cycle starts over. Another major growth mode occurs where there is flow of nutrient through the ecosystem and growth is more or less continuous. This second mode is idealized by systems of continuous culture. One device that is used for continuous culture is the chemostat. It operates by the introduction of sterile medium to a constant-volume culture vessel at a constant rate. When growth is adequate to deplete the medium (usually of one limiting nutrilite), the system becomes self-regulating. For a specified culture volume, V, and flow rate of medium, ω, the bacteria deplete the limiting nutrilite until its concentration, S, becomes such a value that r from Eqs. (1)–(6) equals ω/V. This is a stable situation, because if the biomass were to temporarily increase, S would decrease and then r would decrease until the biomass returned to its original steady-state value. *See* CHEMOSTAT.

In another form of continuous culture, the turbidostat, the biomass concentration is monitored photoelectrically. New medium is added and an equal volume of culture is removed, so that the turbidity stays constant and the steady-state rate of growth on that particular medium is maintained. *See* CULTURE.

Growth on surfaces. Most organisms, even if they grow well in dispersed cultures, are capable of binding to living cells and dead solids. This is advantageous in that they are not easily removed from a favorable environment. In addition, the surface may serve either to provide nutrients or to sequester nutrients. In continuous cultures, adherence provides a way to remain near the inflow of nutrient-rich medium. In fact, most systems of continuous culture are plagued by adherence of the strain under study, or of its mutants or contaminants.

Secondary growth phase. Another aspect of the culture cycle is that many organisms produce special chemicals at certain times and conditions. This has led microbiologists, particularly those working in the fermentation industries, to divide the growth process into the trophophase (feeding phase) and the idiophase (special or personal phase). It is during the latter phase that bacteria and fungi produce large amounts of antibiotics and other chemicals of economic importance. It appears that when the stationary phase approaches, different organisms redirect their metabolism toward the production of different exotic chemicals, which probably have some benefit to the organism. For industrial purposes, the idiophase is extended by strain selection and by careful feeding of appropriate resources (for example, glucose) at a rate that allows very slow growth to continue and the culture to persist for a long time at the border of the exponential and stationary phase.

In some cases, the production at idiophase depends on the concentration of specific metals and, in all cases, the production depends on the absence of exogenous inorganic phosphate. In some, it precedes the processes of differentiation into spores or conidia.

Organismal response to stress. The known organizational functions of the cell together with the stimulatory and regulatory processes use about 10% of the coding capacity of the genome of an organism, such as *E. coli*. Sequencing the entire genome shows that there are many regions that code for unknown proteins. Although the genes for a wide range of functions have been studied, these open reading frames have been encountered; some can be assigned probable functions, while some cannot.

There are two interpretations for the roles of open reading frames. First, they are genes that were needed in the past, then silenced, and in the future may eventually be removed by deletion. Second, they are genes that are functional but are part of systems that function only when stimulated. Such functions are hard to study because the challenges that elicited them are not known. There are, however, three stimulons which are being actively studied that may serve as prototypes for many of the unassigned genes. As distinct from an operon, such as the lactose operon shown in part in Fig. 2, a stimulon is a system of genes not physically linked together. These systems can be studied because they have long-term and specialized functions for the cells. Therefore, stimulons that protect the cell against high temperature, genetic damage, starvation, and oxidation can all be studied by looking for those genes being uniquely transcribed when the cells are challenged. *See* OPERON.

Heat stimulon. Approximately 46 genes are involved in protecting *E. coli* against temperature shocks. Some have been known for decades with regard to how they function and relate to the corresponding genes in flies, humans, and other organisms. Many are synthesized when new sigma subunits (σ^{32} or σ^{e}) for the RNA polymerase are formed in response to the temperature challenge. The sigma factor, σ^{e}, is formed at higher temperatures than σ^{32}. Besides the functions of some of these genes, it is known how some respond to other types of challenges. Thus, treatment with alcohol elicits the formation of a subset of the heat shock genes.

SOS stimulon. When the DNA of the cell is damaged, a set of processes that behave in different ways is initiated. Some are involved in repair; for example, they correct lesions. Photoreactivation, for example, is catalyzed by a gene that uses ultraviolet-A-light (UVA) and visible-light energy to undo the changes in adjoining pyrimidine stretches in the DNA that have been chemically cross-linked by ultraviolet B light (UVB) and ultraviolet C light (UVC). In addition, there are excision repair processes, in which DNA is removed from one strand and then reconstructed complementary to the other partner. Also, there is postreplication repair that can occur if there are two copies of the genome in the cell by recombining fragments between them until a complete, functional genome is generated.

Besides the repair function of this stimulon, the SOS signal inhibits cell division. The inhibition may give time for the repair to take place before the signaling is reversed and the cell attempts to replicate the DNA and to divide. *See* ULTRAVIOLET RADIATION (BIOLOGY).

Starvation stimulon. Some bacteria sporulate when all of an available resource has been consumed, but others such as *E. coli* elicit the formation of up to 50 different proteins. Cells prevented from forming these proteins die very rapidly when starved. These 50 proteins are probably only the tip of the iceberg. For the long-term survival of the species, the death of most of the cells in a colony is irrelevant. However, the creative change of one in a million or a trillion cells so that one cell can cope with a severe environment and grow is all that is needed. Quite likely, such processes involve cell mutations making extreme changes in their DNA that mostly lead to even more immediate death, but may (although very rarely) provide an answer to the cause of the starvation. *See* ANTIBIOTIC; BACTERIAL PHYSIOLOGY AND METABOLISM; INDUSTRIAL MICROBIOLOGY. Arthur L. Koch

Bibliography. M. DePedro, J. V. Holtje, and W. Loffelhardt, *Bacterial Growth and Lysis*, 1993; P. Gerhardt (ed.), *Methods for General and Molecular Bacteriology* 2d ed., 1994; A. L. Koch, *Bacterial Evolution, Growth and Form*, 1995; F. C. Neidhardt, J. L. Ingraham, and M. Schaechter, *Physiology of the Bacterial Cell*, 1990; D. White, *Microbial Physiology*, 1995.

Bacterial luminescence

The production of visible light by bacteria; with very few exceptions this light is blue-green. The phenomenon is seen in many species of several genera, including *Vibrio, Photobacterium, Alteromonas*, and *Xenorhabdus*.

Habitats. Luminous bacteria are primarily marine, but there are some genera with terrestrial (*Xenorhabdus*) and fresh-water (*Vibrio*) species. In the marine environment the bacteria are found in various habitats, including planktonic (free-floating), saprophytic (on a variety of marine proteinaceous materials), parasitic (on a number of marine invertebrates), and symbiotic. Marine shellfish left at room temperature will often begin to glow after a few hours due to the colonies of saprophytic luminous bacteria growing on them.

The symbiotic habitat can take one of several forms. The symbiotic bacteria may be loosely associated as gut symbionts in many different marine organisms; they may be specifically and more tightly associated in the light organs of marine fishes and squids; or they may be very tightly associated as intracellular symbionts in luminescent pyrosomes (light-omitting organelles). When associated as light-organ symbionts, the bacteria are used by the host fish or

squid as a biological light bulb. Under these conditions the bacteria are maintained in specialized organs where they are cultured by the host organism, kept free from contaminants, and continuously emit light. The actual light emission is then controlled physically by the host's use of shutters, chromatophores, or other mechanisms. These symbiotic relationships are probably the most common habitats in which bacterial luminescence is observed in the ocean.

Biochemistry and physiology. The chemistry of bacterial luminescence is unique among luminous organisms. The enzyme that catalyzes light emission is luciferase; it combines with a riboflavinlike substance called flavin mononucleotide ($FMNH_2$). This complex then reacts with a long-chain aldehyde, and with molecular oxygen to form an excited state capable of emitting light. The molecule that actually emits the light is an altered form of the flavin. This unique biochemistry has been used as an indicator of the presence of luminous bacteria in cases where the symbiotic bacteria could not be obtained in pure culture or could not be grown free from their host. *See* RIBOFLAVIN.

Physiological studies have shown that the luminous activities are regulated by many different factors, including the carbon and energy sources used for growth, oxygen, salt (osmotic) concentration, and iron. These controls can differ markedly between bacterial species. Autoinduction, another control system, makes the bacteria luminous whenever they are in high concentration, and shuts down the luminescence when the cells are dilute. This mechanism is apparently found in all species of luminous bacteria.

Many mutants have been isolated after treatment with chemical mutagens, and these have been used to elucidate the physiology and biochemistry of light emission. Mutants have been found in the synthesis of the luminescent system, in the enzyme luciferase, and in several other key components. For three different species of luminous bacteria, the genes for bioluminescence have been cloned and shown to express functionally in *Escherichia coli*, a common intestinal bacteria of humans. Cloning the genes into *E. coli* has allowed many genetic and biochemical manipulations that were previously impossible, and has greatly enhanced the knowledge of bacterial luminescence. Cloning as few as seven genes into *E. coli* can make these organisms luminous and in many ways analogous to the marine luminous bacteria. This raises questions about why bacterial luminescence is not more widespread, and what the use of luminescence is to the bacteria. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; BIOLUMINESCENCE; GENETIC ENGINEERING. Kenneth H. Nealson

Bibliography. G. Leisman, D. Cohn, and K. H. Nealson, Bacterial origin of luminescence in marine animals, *Science*, 208:1271–1273, 1980; K. H. Nealson and J. W. Hastings, Bacterial bioluminescence: Its control and ecological significance, *Microbiol. Rev.*, 43:469–518, 1979; M. P. Starr (ed.), *The Prokaryotes*, 1981.

Bacterial physiology and metabolism

The biochemical reactions that together enable bacteria to live, grow, and reproduce. Strictly speaking, metabolism describes the total chemical reactions that take place in a cell, while physiology describes the role of metabolic reactions in the life processes of a bacterium.

The study of bacteria has significance beyond the understanding of bacteria themselves. Since bacteria are abundant, easily grown, and relatively simple in cellular organization, they are ideal for the study of basic life processes and thus have been used extensively in biological research. Functional analyses of bacterial systems have provided a foundation for much of the current detailed knowledge about molecular biology and genetics.

Bacteria are prokaryotes, lacking the complicated cellular organization found in higher organisms; they have no nuclear envelope and no specialized organelles. Yet they engage in all the basic life processes—transport of materials into and out of the cell, catabolism and anabolism of complex organic molecules, and the maintenance of structural integrity. To accomplish this, bacteria must obtain nutrients and convert them into a form of energy that is useful to the cell.

This article discusses cell chemistry, enzymes, coenzymes, catabolism including oxidative metabolism and fermentative metabolism, and anabolism.

Cell Chemistry

Bacterial cells contain a number of distinctive chemical constitutents, many of which have been found in the cells' capsules, walls, or membrane structures. Despite the presence of the variety of unusual compounds in the major classes of chemical structures in bacteria, the quantitative chemical analyses are broadly similar to those for other cell types. Thus, on a weight basis, bacteria contain 70–90% water, and a mineral content of 1–10%. Much of the mineral matter is accounted for by potassium, sodium, calcium, magnesium, and phosphorus. Proteins and nucleic acids, the major organic constituents of bacteria, are present in more constant amounts than other organic compounds, such as the lipids and carbohydrates. Protein, determined by the percentage of nitrogen, accounts for 40–60% of the dry weight of the bacterial cell. The nucleic acid content is 10–20%. The amounts of carbohydrate and lipid vary widely and are dependent on the growth conditions. Carbohydrate content is 10–30% and that of lipids 1–50%.

Chemical composition. The overall chemical composition of the bacterial cell is very similar to that of all other types of cells of animal, plant, and microbial origin, which are capable of growth and replication. Bacteria thus possess the proteins ribonucleic acid (RNA) and deoxyribonucleic acid (DNA), which are the major classes of chemical constituents. These are required for the biochemical activities and the continuity and expression of the genetic characteristics of the cell. All cells possess a plasma membrane structure responsible for permeability properties, and one

of the universal features of cellular membranes is their protein-lipid nature.

Bacterial cytoplasmic membranes generally have a high ratio of protein to lipid (3:1), a feature similar to that of eukaryotic mitochondrial membranes. Thus, bacteria contain lipids of the phospholipid class as essential chemical constituents of their cell membranes. Lecithin is one of the commonest phospholipids in higher organisms, and it is found in several bacterial groups, but much less frequently. Although sterols, such as cholesterol, are widely distributed in nature, occurring in membranous structures from animal, plant, and higher microbial cells, they have not been found in bacterial cells (except the mycoplasma group of pathogenic bacteria where sterols are sequestered into the membranes from an exogenous supply in the complex growth medium) or in blue-green cyanobacteria. Hopanoids of the triterpenoid family have been found in a wide variety of bacteria, and are structurally related to sterols and may be the structural equivalents to sterols of eukaryotic organisms. On the other hand, carotenoids, which are also lipid-soluble compounds, are found in many bacteria and in the closely related cyanobacteria where they are localized in the cell membranes and photosynthetic membranes (thylakoid equivalents in prokaryotes). Isoprenoid quinones, a class of terpenoid lipids, occur in the membranes of bacteria, cyanobacteria, and Archaebacteria where they function as respiratory quinones; and as chemical entities they have chemotaxonomic relevance. The other major class of chemical components of cells, the carbohydrates or polysaccharides, also occurs in a wide variety of bacterial groups. The polysaccharides may occur in the bacterial cell as starch- or glycogenlike storage granules, or as surface capsular or slime layers or polymers covalently linked to peptidoglycan in gram-positive bacteria, or to lipid structures of lipopolysaccharides in gram-negative bacteria.

Many of the unusual compounds of biological origin have been found in bacterial cell structures and products. Although bacterial proteins appear to be quite normal with respect to the variety and isomeric form of the amino acid building blocks, the cell walls, certain capsules, and antibiotics formed by bacteria possess peptides containing D isomers of amino acids, such as alanine, glutamic and aspartic acids, and phenylalanine. The amino acid diaminopimelic acid, first detected and isolated from the bacterium *Corynebacterium diphtheriae* in 1950, was subsequently found to be one of the characteristic cell-wall amino acids of many bacterial species. It occurs in the wall peptidoglycan but is not an amino acid component of the bacterial proteins. *See* AMINO ACIDS; ANTIBIOTIC; CAROTENOID; CELL MEMBRANES; STEROL.

Cell walls of bacteria contain one of the most distinctive bacterial products, the amino sugar muramic acid (3-*O*-carboxyethyl-glucosamine), in combination with glucosamine in the peptidoglycan polymer forming the rigid wall structure. The distribution of muramic acid in nature is confined to the bacteria, cyanobacteria, and rickettsias. Muramic acid has not been detected in higher microorganisms or in plant and animal cells. It has also not been found in Archaebacteria, but certain members of this group possess an analogous wall structure which contains talosaminuronic acid in its glycan; this amino sugar serves as the linkage molecule to the peptides. *See* ARCHAEA.

The amino sugars generally found in higher organisms are glucosamine, galactosamine, and the sialic acids. In addition to these amino sugars and muramic acid, bacteria also contain a variety of other amino sugars in their polysaccharide and lipopolysaccharide structures. Polysaccharides of the pneumococcus capsular substances and lipopolysaccharides of gram-negative bacteria have been shown to contain the following variety of amino sugars: mannosamine, fucosamine, pneumonsamine (an aminodideoxytalose), amino uronic acids, amino dideoxy sugars, and diamino sugars.

The bacterial lipopolysaccharides possess a unique lipid tail consisting of lipid A, a glucosamine moiety to which β-hydroxy fatty acids are linked, with the amino sugar linked to the core and polysaccharide chains. They may also contain a unique class of sugars, the dideoxy sugars—colitose, abequose, paratose, and tyvelose. The other very characteristic product of the lipopolysaccharides is the sugar acid ketodeoxyoctonic acid. Bacterial cell walls also possess a group of polymers called the teichoic acids, whose occurrence so far is confined to bacteria. The teichoic acids are polymers of ribitol or glycerol phosphate with additional compounds such as glucose, *N*-acetylglucosamine, and D-alanine linked to the polyol backbone. The D-alanine in the teichoic acid is linked through an ester bond and thus represents another example of the unusual features of the bacterial wall compounds. In addition to wall teichoic acids, gram-positive bacteria have either a glycerol lipoteichoic acid or a lipomannan anchored in their membranes. This chemically unique class of compounds is amphiphilic, possessing a glyceride lipophilic tail. *See* AMINO SUGAR.

Proteins and peptides. Most of the cellular proteins formed by bacteria are similar in amino acid composition to those of other organisms. The extracellular proteins secreted by gram-positive bacteria show one conspicuous difference from many other proteins in that they are devoid of, or exceptionally low in, the amino acid cysteine. Thus it appears that these proteins lack the S-S type of bridges linking the peptide chains together. In addition to the polypeptide antibiotics, the bacterial cell produces unusual peptide structures. The substance forming a thick, viscid layer, or capsule, surrounding the cells of *Bacillus anthracis* is a polymer of D-glutamic acid.

Although glycoproteins similar to those of animal cells have not been found in bacteria, some of the Archaebacteria possess cell envelopes assembled from glycoprotein subunits, and certain eubacteria have surface layers composed of glycoproteins. *See* PROTEIN.

Nucleic acids. As in other cells, the bacterial cell contains both RNA and DNA. The RNA is largely in the form of ribosomes (RNA-protein particles) of somewhat smaller dimensions than those found in higher organisms. The DNA of the bacterial cell gives the Feulgen reaction, which is the cytochemical basis for the demonstration of the bacterial nucleus. The DNA in bacteria is not enclosed within a nuclear membrane as it is in other types of cells. The base composition (G + C, or guanine + cytosine) values vary from one bacterial species to another and range from 20 to 74%. *See* NUCLEIC ACID.

Carbohydrates. A great variety of carbohydrates are synthesized by bacteria. They range from simple polymers of glucose, such as the cellulose produced by *Acetobacter xylinum*, to complex substances composed of a number of sugar units, such as glucose, galactose, rhamnose, and amino sugars. Unusual features of certain bacterial carbohydrates include the presence of the D isomer of arabinose in some, the occurrence of heptose sugars, and the detection of new amino sugars. Neuraminic acid, an important constituent of carbohydrate-protein complexes of animal cells and tissues, has been reported in polymeric form in *Escherichia coli* and is also found in the surface polysaccharide of group B streptococci and in meningococcal capsular polysaccharides. *See* CARBOHYDRATE.

Lipids. Bacteria produce a variety of lipids and lipid complexes, waxes, fats, glycolipids, and peptidoglycolipids. Although phospholipids constitute the major class of lipid, smaller amounts of lipid-soluble compounds including carotenoids, isoprenoid quinones, sterollike hopanoids, and hydrocarbons are found in bacteria. Despite the lower amounts of these compounds, they are functionally important for the bacteria.

Phospholipids found in bacteria include lecithin (phosphatidyl choline); phosphatidyl derivatives of ethanolamine, serine, and inositol; phosphatidic acid; phosphatidyl glycerol; and cardiolipin. An unusual lipid first detected in *Bacillus megaterium* is a polymer of β-hydroxybutyric acid. It is present in the form of granules that may be stained with the fat-soluble dye Sudan black, and has been found in many different kinds of bacteria.

The lipids of Archaebacteria are chemically quite unique and differ from the glycerol–fatty acid lipids characteristic of eubacteria and eukaryotic cells. The archaebacterial lipids are isopranyl glycerol ether lipids and are distinctive chemical markers for the identification of this group of bacteria.

The lipids of bacteria contain the common fatty acids found in lipids from other types of cells. However, bacteria lack the ability to form polyunsaturated fatty acids and thus differ from higher microorganisms, plants, and animals. Certain bacteria contain high proportions of branch-chain fatty acids in their lipids. Other unusual features of the bacterial lipids include the occurrence of lipoamino acids in a variety of species and the derivatives of glycerol diether in the extremely halophilic bacterium *Halobacterium cutirubrum*. *See* LIPID.

Chemical anatomy. Anatomy, the science of bodily and cellular structure, has been applied to the bacterial cell. The major structural elements of bacteria have been isolated and chemically characterized. This method of studying the chemical anatomy of the bacterial cell avoids the difficulties of applying cytochemical tests, such as the Feulgen reaction, to such small cells.

Flagella and pili. The filamentous, locomotive appendages of bacteria are composed of flagellin, a protein related chemically and physically to the hairlike proteins of other organisms. Common enteric bacteria and other groups of gram-negative organisms possess long, hairlike appendages (pili, or fimbriae) composed of proteins called pilins that function in the adhesion of bacterial cells. *See* CILIA AND FLAGELLA.

Capsules. The thick, mucous envelope, or capsules, may be composed of polypeptide or carbohydrate material. The viscous capsular carbohydrates may be composed of a variety of sugars, including hexoses, pentoses, methylpentoses, uronic acids, and amino sugars.

Cell walls. Bacteria possess a rigid wall structure that defines the shape of each cell. The principal structural component of the wall is a covalently bonded peptidoglycan, the glycan portion of which consists of alternating molecules of *N*-acetylmuramic acid and *N*-acetylglucosamine linked by $\beta 1 \rightarrow 4$ glycosidic bonds. The peptide component is linked by an amide bond to the muramic acid. Alanine, D-glutamic acid, lysine, or diaminopimelic acid form the basic part of the peptide, and in some bacterial walls these peptide strands may be cross-linked by either glycine, alanine, or threonine peptide bridges or sometimes all three. In the gram-positive bacteria, the teichoic acids and polysaccharides may be covalently linked to the rigid peptidoglycan.

The wall, or more strictly speaking, the cell envelope, of the gram-negative bacteria is a more complex structure with a rigid layer of peptidoglycan upon which the lipopolysaccharide-protein-lipid complexes are anchored together with a cysteinyl glycerol lipoprotein covalently linked to peptidoglycan. The peptidoglycan has the basic building blocks of the structure found in the walls of gram-positive organisms. Certain Archaebacteria have a rigid layer or sacculus which is chemically different from the peptidoglycan layer of the common bacteria (eubacteria) in that the former layer has talosaminuronic acid instead of muramic acid and only L-isomers of amino acids in the peptide moiety.

Protoplasmic membrane and mesosome. The plasma membrane and its intrusions within the cell (called the mesosome) are composed of 20–30% lipid and 50–70% protein. In some instances, small amounts of carbohydrate have been found in isolated membranes. Virtually all of the cell's lipid, phospholipid, carotenoid, and electron-transport components are localized in the membrane-mesosome structures, in addition to which the membranes of gram-positive organisms contain the membrane amphiphiles, lipoteichoic acids, or lipomannans.

The membranes of bacteria perform mitochondrial, transport, and biosynthetic functions and thus contain a variety of proteins separable by polyacrylamide gel electrophoresis. M. R. J. Salton

Enzymes

A list of bacterial enzymes (organic catalysts) would include many of the enzymes found in mammalian tissues, as well as many enzymes not found in higher forms of life. By combining with such enzymes, many antibiotics are able to exert a selective killing or inhibition of bacterial growth without causing toxic reactions in the mammalian host. The great capability of the bacterial cell to metabolize a wide variety of substances, as well as to control to some extent the environment in which the cell lives, is reflected in its ability to form inducible enzymes. The majority of bacterial enzymes require cofactors for activity. These cofactors may be inorganic cations of organic molecules called coenzymes. *See* COENZYME; ENZYME.

Classification. Bacterial enzymes may be classified in numerous ways, for example, on the basis of (1) whether they are inducible or constitutive; (2) whether they are degradative (catabolic; resulting in the release of energy) or synthetic (anabolic; using energy to catalyze the formation of macromolecules); or (3) whether they are exoenzymes (enzymes secreted from the cell to hydrolyze insoluble polymers—wood, starch, protein, and so on—into smaller, soluble compounds which can be taken into the cytoplasm of the bacterium).

In addition, bacterial enzymes are involved in the transport of substrates across the cell wall, in the oxidation of inorganic molecules to provide energy for the cell, and in the destruction of a large number of antibiotics.

Inducible and constitutive enzymes. Constitutive enzymes are defined as those enzymes formed by the bacterial cell under any or all conditions of growth, whereas inducible enzymes are formed by the bacterial cell only in response to an inducer. In nature the inducer is usually the substrate for the specific inducible enzymes involved. However, in the laboratory it is possible to use inducers which are not substrates, provided the molecular configurations are similar to those of true substrates. Galactokinase, an enzyme necessary for the phosphorylation of galactose, is an example of an inducible enzyme, since it is formed only if the bacteria are grown on a medium containing D-galactose or D-fucose as a carbon source. If a culture of bacteria grown on galactose is subsequently transferred to a galactose-free medium but otherwise sufficient for protein synthesis, the enzyme (galactokinase) will no longer be synthesized by the bacterial cell. Although bacteria vary considerably in their ability to utilize various substrates, most are versatile in their ability to form inducible enzymes. However, such enzymes are not ordinarily synthesized by the bacterium because a repressor substance produced by the cell reacts with the operator gene to prevent the transcription of the specific gene involved and, hence, the ultimate synthesis of the enzyme. When inducer is present (normally the substrate), it reacts directly with the repressor, preventing its reaction with the operator gene. The gene is then able to function, and transcribes its message to messenger RNA for synthesis of the enzyme or enzymes being induced. Since this process involves the removal of a repressor, it is also referred to as derepression.

An additional control was described when it was observed that, when cells were actively metabolizing a substrate such as glucose, it was not possible to induce them to synthesize inducible enzymes even in the presence of the specific inducer. This phenomenon, called catabolite repression, can be overcome by the addition of adenosine 3′,5′-cyclic monophosphate (cAMP); it is now known that before the operon can be transcribed a catabolite repressor protein (CR-protein) must first bind to cAMP and then react with the promotor gene for the enzyme. When a cell is rapidly metabolizing glucose, cAMP is decreased, and the CR-protein cannot react with the promoter to permit transcription. Thus, these controls accomplish two functions: (1) inducible enzymes are not synthesized unless the specific substrate is available; and (2) inducible enzymes are not synthesized when the cell is already rapidly metabolizing a better substrate.

Catabolic and anabolic enzymes. Many bacteria can oxidize various organic compounds in the presence of oxygen to CO_2 and H_2O, and other organisms carry out incomplete oxidations or fermentations resulting in the formation of stable oxidation or fermentation products. Examples of a few such end products of bacterial metabolism would include acetic acid, lactic acid, formic acid, butyric acid, ethanol, isopropanol, acetone, butanediol, butanol, hydrogen, and carbon dioxide. These degradative reactions yield energy which can be trapped as adenosine triphosphate (ATP) to supply the energy for macromolecular syntheses.

Anabolic enzymes are necessary to synthesize proteins, nucleic acids, lipids, cell walls, and so on. Many of these enzymes carry out reactions which are similar to those occurring in mammalian cells, but others are extremely specific for a bacterial cell, such as those enzymes involved in cell wall biosynthesis and capsule synthesis. Furthermore, many bacterial enzymes involved in the transcription of DNA (DNA-dependent RNA polymerase) and synthesis of proteins either are not found in mammalian cells or are sufficiently different from their mammalian counterparts that they can be inhibited by various antibiotics without harm to the mammalian cell.

Some of the Common reaction types catalyzed by bacterial enzymes are listed in **Tables 1** and **2**; the reactions given as examples illustrate either a general or a specific case.

Many enzymes are recognized as composed of a protein moiety, also called carrier or apoenzyme, and a small molecular species, the coenzyme, also known as the prosthetic or active group. Neither component acts as a catalysts by itself; only the combination, or holoenzyme, does. The coenzyme is responsible for

TABLE 1. Reactions involving carbohydrates

Enzyme type	Type of reaction catalyzed	Example
Kinases	Phosphorylation of substrate	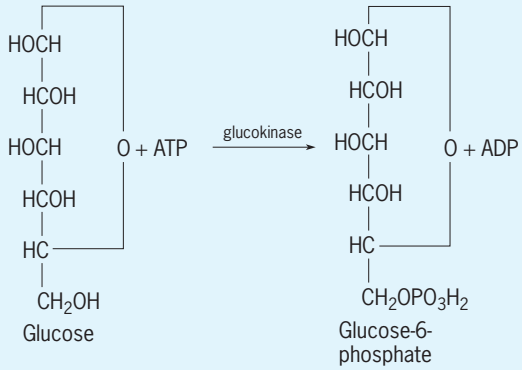
Dehydrogenases	Oxidation or reduction of substrate	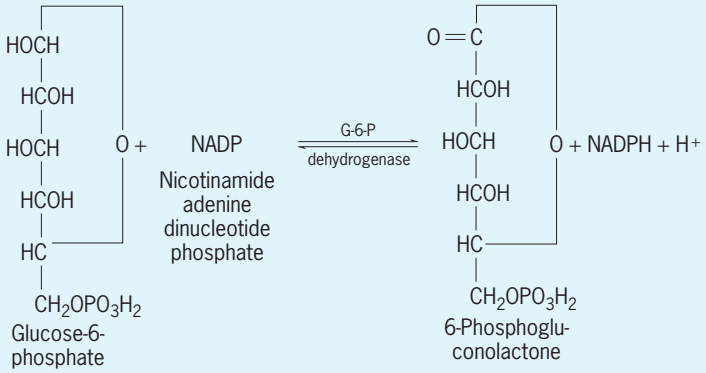
Isomerases	Molecular rearrangement of substrate to form ketose from aldose or aldose from ketose	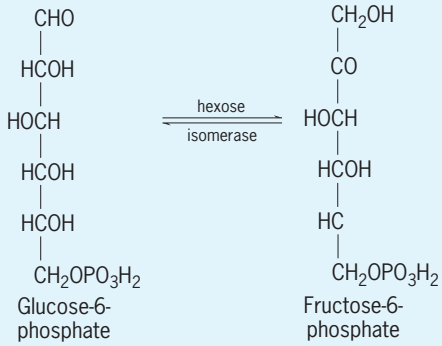
Epimerases	Rearrangement of hydroxyl groups on substrtae	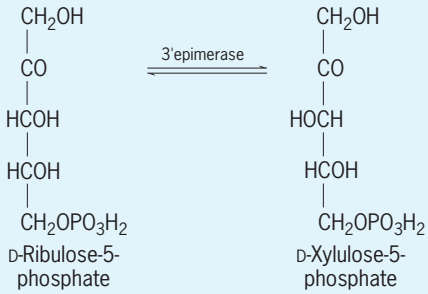
Aldolases	Cleavage of substrate into an alcohol and an aldehyde	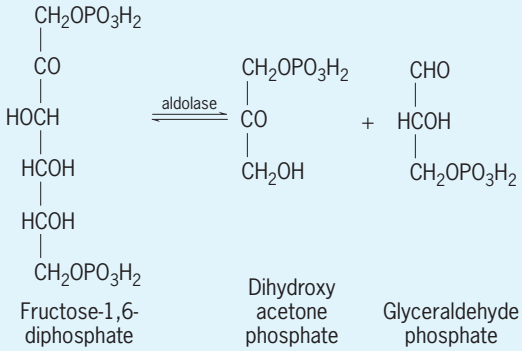

-s

TABLE 1. Reactions involving carbohydrates (*cont.*)

Enzyme type	Type of reaction catalyzed	Example
Ketolases	Cleavage of substrate at position of carbonyl carbon	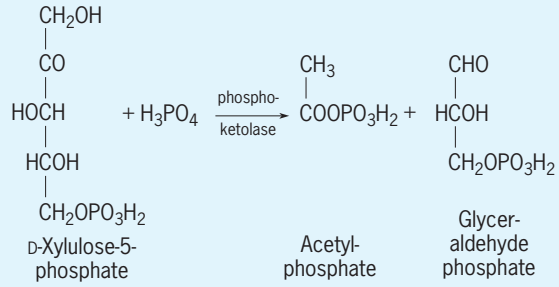
Enolase	Convert keto form of substrate to enol form by addition of H_2O	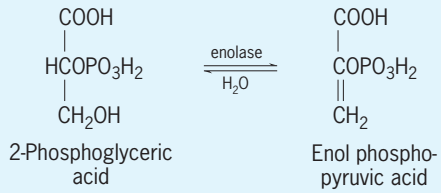
Decarboxylases	Remove carboxyl carbon from substrate	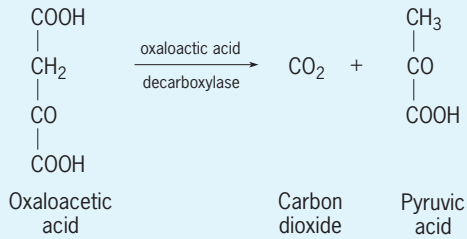
Mutases	Result in transposing phosphate from one carbon of substrate to a different carbon atom of same substrate molecule	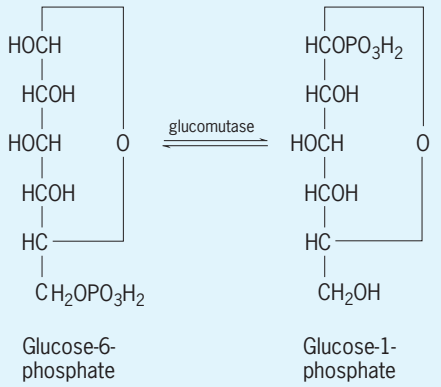
Transketolase	Similar to ketolase except 2-carbon fragment must be transported by the enzyme to an acceptor compound to form a new compound	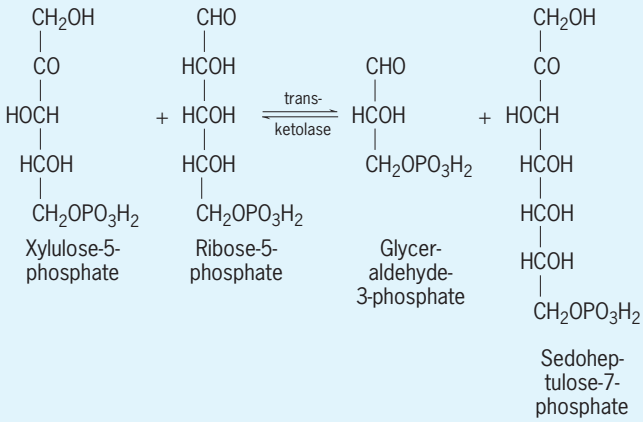

TABLE 1. Reactions involving carbohydrates (*cont.*)

Enzyme type	Type of reaction catalyzed	Example
Synthetase	Condensation of phosphoenol pyruvate with CO_2 or other phosphorylated carbohydrates; each synthetase is specific for one condensation	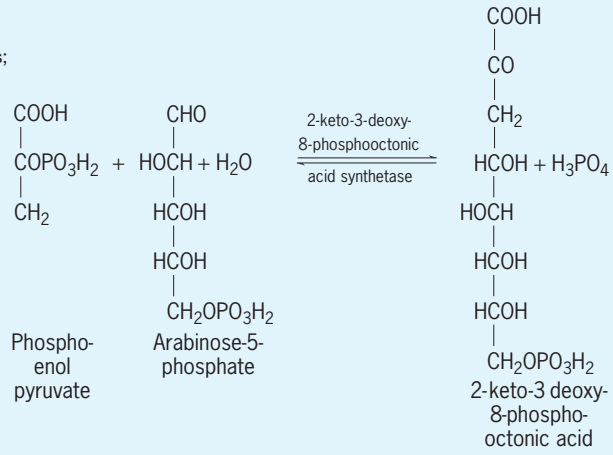
Oxidases	Yield same overall result as dehydrogenases; however, no coenzyme is involved required and molecular oxygen is	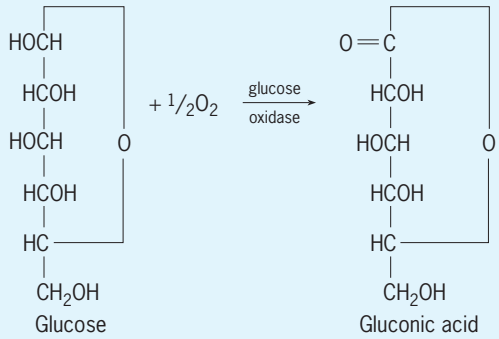
Phosphorylases	Condenses smaller phosphorylated sugars into polysaccharides	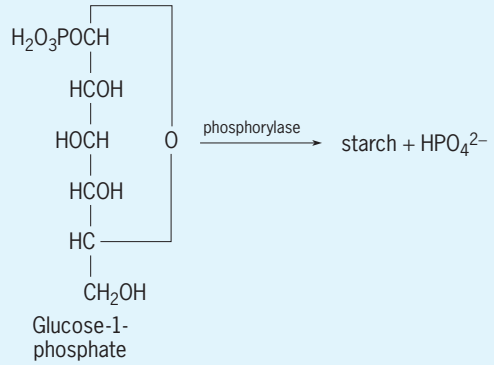
Phosphatases	Remove phosphate from aldehyde or acid substrate	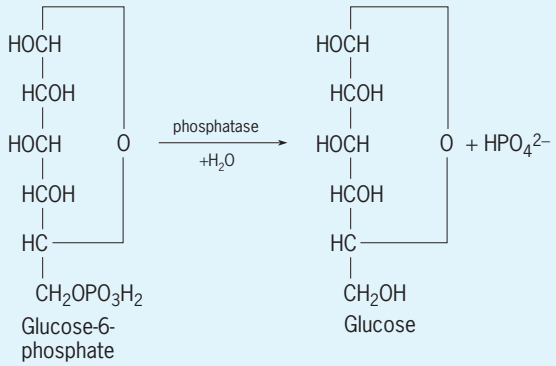

TABLE 2. Enzymes involving amino acid metabolism

Enzyme type	Type of reaction catalyzed	Example
Amino acid deaminases	Remove amino group from amino acid; different deaminases may carry out a reductive, or oxidative, deamination; members of genus *Clostridium* carry out an anaerobic dismutative deamination referred to as the Stickland reaction; it represents a combination of the two kinds	*Reductive:* $RCH(NH_2)COOH \xrightarrow[+2H]{\text{deaminase}} RCH_2COOH + NH_3$ *Oxidative:* $RCH(NH_2)COOH \xrightarrow{\text{deaminase}} RCOCOOH + NH_3 + 2H$ *Stickland reaction:* $RCH(NH_2)COOH + XCH(NH_2)COOH \xrightarrow{+H_2O} RCOCOOH + XCH_2COOH + 2NH_3$
Amino acid transaminases	Removes amino group from one amino acid and carries it to an α-keto acid to form a new amino acid and a new α-keto acid	$HOOC(CH_2)_2CH(NH_2)COOH$ (Glutamic acid) + $CH_3CO(COOH)$ (Pyruvic acid) $\xrightarrow{\text{transaminase}}$ $CH_3CH(NH_2)COOH$ (Alanine) + $HOOC(CH_2)_2CO(COOH)$ (α-Keto-glutaric acid)
Amino acid decarboxylases	Remove carboxyl group leaving basic amine	$RCH(NH_2)COOH \longrightarrow RCH_2NH_2 + CO_2$ Amino acid → Amine
Amino acid oxidases	Same as oxidative deaminases (see amino acid deaminases above)	
Amidases	Remove amide group from some amino acids	$HOOC(CH_2)_2CH(NH_2)CONH_2$ (Glutamine) $\xrightarrow[H_2O]{\text{glutaminase}}$ $HOOC(CH_2)_2CH(NH_2)COOH$ (Glutamic acid) + NH_3

the particular type of catalytic activity displayed by the holoenzyme, such as the dehydrogenation of an alcohol to an aldehyde or of an aldehyde to a carboxyl group or the transfer of an amino group from one molecule to another. The apoenzyme, on the other hand, confers on the holoenzyme its pronounced specificity. Therefore, a single kind of coenzyme can act in combination with any one of a variety of apoenzymes, so that each combination of coenzyme and apoenzyme is functional in the conversion of one specific substrate.

Exoenzymes. In addition to the enzymes of bacteria which participate directly in the metabolic cycles of the cell, there are numerous enzymes which function as depolymerizing enzymes, making micromolecules available to the cell. The most common examples include the cellulases which degrade cellulose to cellobiose, amylases which degrade starch to dextrins and maltose, and the proteinases which hydrolyze proteins to polypeptides and amino acids. The ability of any given organism to use a particular substance of high molecular weight depends upon its ability to produce hydrolytic enzymes that are capable of attacking the substance in question. The exoenzymes are generally referred to as hydrolases because they catalyze the hydrolytic breakdown of their substrates (**Table 3**). *See* AMYLASE.

Inducible exoenzymes. Many exoenzymes are constitutive. Nevertheless, their production may be

TABLE 3. Extracellular bacterial enzymes

Organism	Enzyme	Substrate	End products
Clostridium	Cellulase	Cellulose	Cellobiose
Beneckea	Chitinase	Chitin	*N*-Acetyl glucosamine derivatives
Erwinia	Pectinase	Pectin	Galacturonic acid derivatives
Bacillus	Amylase	Starch Amylose Glycogen	Limit dextrins, maltose
Bacillus	Proteinase	Protein	Polypeptides
	Lipases (esterases)	Fats	Glycerol and fatty acids

enhanced by cultivation of the appropriate bacteria in the presence of the corresponding di-, oligo-, or poly-saccharides; hemicelluloses; and proteins. In some cases exoenzymes are formed only in response to the presence of the specific substrate.

Medical importance. Many pathogenic microorganisms excrete enzymes which may play an important role in pathogenesis in some cases. The α-toxin (lecithinase) of *Clostridium perfringens* illustrates a highly active enzyme which is responsible for the necrotizing action associated with gas gangrene infections due to this microorganism. *Streptococcus pyogenes* excretes hyaluronidase which degrades ground substance (polymer of hyaluronic acid), and streptokinase which activates plasmin resulting in a system that lyses fibrin. Other examples include coagulase of the *Staphylococcus* which activates clotting of plasma, urease of *Proteus vulgaris* which splits urea to ammonia and carbon dioxide, and collagenase of *Clostridium* which hydrolyzes collagen. A summary of the more common enzymes excreted by pathogenic microorganisms is presented in **Table 4**. *See* DIPHTHERIA; GANGRENE; HYALURONIDASE; STAPHYLOCOCCUS.

Enzymes destroying antibiotics. Many bacteria are able to synthesize enzymes which will hydrolyze or modify an antibiotic so that it is no longer effective. Essentially all of these enzymes are coded by DNA that exists in bacterial plasmids. As a result, the ability to produce enzymes which destroy antibiotics can be rapidly passed from one organism to another either by conjugation in gram-negative organisms or by transduction in both gram-negative and gram-positive bacteria. Bacterial plasmids can code for enzymes that (1) destroy penicillin by hydrolysis of the β-lactam bond; (2) destroy streptomycin by adding adenyl groups to the antibiotic; (3) inactivate neomycin and kanamycin by phosphorylating the antibiotics; (4) destroy chloramphenicol by adding acetyl groups; and (5) render tetracycline ineffective, probably by reducing the permeability of the bacterial cell to the antibiotic.

Autolytic enzymes. Autolytic enzymes seem to be a necessary component of the cell to hydrolyze covalent bonds in the rigid cell wall, allowing insertion of new structural units into the growing wall. Autolysis (dissolution of the cell) occurs when autolytic enzymes and cell wall synthetic enzymes do not work in unison. Thus, if too many bonds are opened in the rigid cell wall before new subunits are ready to occupy these spaces, internal osmotic pressures within the bacteria cause the cell to burst. In some organisms the autolytic system takes over following injury to the cell, and in others (such as *Streptococcus pneumoniae*) this system can be easily activated by treatment with surface-active agents such as bile salts.

Control of enzymes. In the absence of a control mechanism, a bacterial cell would synthesize much more of a specific metabolite than is necessary for sustained growth. Therefore the cell has developed two types of control for the regulation of bacterial enzyme activity: enzyme repression and feedback inhibition.

Enzyme repression. This is a rather coarse control that acts at the genetic level. This system acts in a manner that is somewhat the reverse of enzyme induction. Thus, if an organism that is synthesizing its own amino acid, histidine (which requires 20 different enzymes), is transferred to a medium containing adequate histidine, the organism will cease to synthesize the enzymes needed for histidine biosynthesis.

TABLE 4. Enzymes excreted by microorganisms of medical importance

Organism	Enzyme	Substrate	End products
Clostridium	Lecithinase	Lecithin	Diglyceride, phosphoryl choline
Clostridium	Collagenase	Collagen	?
Streptococcus	Hyaluronidase	Hyaluronic acid polymer	Hyaluronic acid
Streptococcus	Streptodornase	Deoxyribonucleic acid	Nucleotides
Streptococcus	Streptokinase	Activates plasminogen to plasmin	Results in lysis of fibrin clots
Staphylococcus	Coagulase	Coagulase reacting factor	Results in coagulation of plasma
Proteus	Urease	Urea	Ammonia and carbon dioxide
Corynebacterium diphtheriae	Diphtheria toxin	Nicotinamide adenine dinucleotide (NAD)	Splits NAD and adds ADP-ribose to elongation factor 2 to prevent protein synthesis by freezing ribosome movement

In other words, the presence of histidine has stopped the transcription of the entire portion of DNA which is responsible for histidine biosynthesis. A model for this effect supposes that the histidine reacts with an inactive repressor, resulting in the formation of an active repressor which can now react with the operator gene of the chromosome so as to prevent the transcription of the operon. Should the histidine level fall too low, the repressor is released, and the operon will again be transcribed. Many bacterial enzymes are subject to enzyme repression at the genetic level.

Feedback inhibition. This is a much finer control in which the end product of a series of reactions reacts directly with the enzyme of the first reaction and temporarily inactives it so that no more end product can be made. Note that in this case the end product reacts directly with an enzyme, and not with a repressor or operator gene on the chromosome. Thus, feedback inhibition does not prevent synthesis of the enzyme, but only inhibits its activity. As the end product is used up by the bacterium, the enzyme is freed to make more product. Enzymes that are subject to feedback inhibition are called allosteric enzymes since they appear to have two active sites, that is, one for their enzymatic activity, and one which reacts with a final end product, causing the inhibition of the enzymatic site. *See* ALLOSTERIC ENZYME; BACTERIAL GENETICS.

Coenzymes

Organic molecules that participate directly in a bacterial enzymatic reaction and may be chemically altered during the reaction are called coenzymes. Although many enzymes do not need specific cofactors, the majority of bacterial endoenzymes do. These cofactors may be simple inorganic cations such as magnesium, manganese, or calcium, which are commonly referred to as activators, or more complex organic molecules (coenzymes).

Coenzymes are the functional units, also called prosthetic or active groups, of an enzyme. Their catalytic activity depends on their association with a protein moiety, the apoenzyme or carrier, which is responsible for the high degree of specificity of the complex, or holoenzyme. Many of these combinations are readily dissociable: holoenzyme $\rightleftharpoons$ coenzyme + apoenzyme. Good examples of such enzymes are the pyridine nucleotide enzymes whose coenzymes are nicotinamide adenine dinucleotide (NAD) and nicotinamide adenine dinucleotide phosphate (NADP); and the flavin nucleotide enzymes, whose coenzymes are flavin mononucleotide and flavin adenine dinucleotide (FMN and FAD). These coenzymes act as hydrogen acceptors and can, in combination with a wide variety of specific apoenzymes, catalyze the dehydrogenation of a large number of substrates.

Some holoenzymes, particularly the iron-porphyrin enzymes such as catalases, peroxidases, and cytochromes, can be dissociated only by drastic means; their coenzymes are far more firmly bound to the protein. This is in keeping with the much more restricted range of their activity, which appears to be limited to specific reactions that are common to the majority of bacteria and other living organisms. *See* ENZYME; NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD); NICOTINAMIDE ADENINE DINUCLEOTIDE PHOSPHATE (NADP).

Coenzymes and growth factors. The universal occurrence in the metabolism of all living organisms of hydrogenations, dehydrogenations, transhydrogenations, aminations, deaminations, transaminations, and various other general reaction types implies that the coenzymes for these processes must be present in the cells. This has been important in discovering the often-encountered special nutrient requirements, especially for vitamins of the B group. These vitamins are substances that are structurally closely allied to coenzymes. Thus the ability of a particular bacterium to develop only if supplied with a certain B vitamin can be readily interpreted to mean that the organism cannot synthesize the corresponding coenzyme from other substances. For example, *Hemophilus influenzae* cannot synthesize and must consequently be supplied with the intact coenzyme before it can grow. Other organisms can synthesize the pyridine nucleotides if they are provided with simpler building blocks, such as nicotinamide or nicotinic acid. If a bacterium can grow in a medium devoid of such building blocks, the implication is that it can perform the synthesis of the coenzyme from still more remote, and often quite simple, ingredients. Where this has been experimentally tested, the results have been uniformly positive. *See* CULTURE.

Many organisms have become obligatory parasites because of a spontaneous loss in their ability to synthesize a needed vitamin. As a result, such organisms become completely host-dependent for a source of growth factors. This has been the basis for the development of ideas on physiological specialization and evolution. *See* VITAMIN.

Function. The mode of action of coenzymes generally involves their role as temporary acceptors for particular atoms or atom groups derived from a substrate, and subsequently as donors of the same entities to another kind of acceptor under the influence of specific apoenzymes. This is shown in **Fig. 1**, where H_2A represents an appropriate oxidizable substrate and A its oxidation product.

Similar diagrams can be constructed to illustrate the role of ATP as a phosphorylating agent with the production of adenosine diphosphate (ADP), and the subsequent reformation of ATP from ADP with phosphate groups produced in special positions during metabolism; of thiamine pyrophosphate as the coenzyme of certain decarboxylations; of pyridoxal

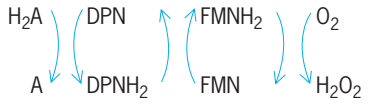

Fig. 1. Diagram of role of pyridine nucleotide and flavin nucleotide coenzymes in hydrogen transfer.

TABLE 5. Some organic coenzymes with their functions and related nutritional factors

Coenzyme	Function	Related nutritional factor or vitamin
Nicotinamide adenine dinucleotide (NAD)	Hydrogen acceptor in dehydrogenases	Nicotinamide
Nicotinamide adenine dinucleotide phosphate (NADP)	Hydrogen acceptor in dehydrogenases	Nicotinamide
Flavin mononucleotide (FMN)	Hydrogen acceptor in dehydrogenases	Riboflavin
Flavin adenine dinucleotide (FAD)	Hydrogen acceptor in dehydrogenases	Riboflavin
Coenzyme A (CoA)	Condensing enzymes	Pantothenic acid
Pyridoxal phosphate	Transaminases, amino acid decarboxylases	Pyridoxine, pyridoxal, pyridoxamine
Thiamine pyrophosphate	Cocarboxylase	Thiamine
Tetrahydrofolic acid	Transformylation	Folic acid
6,8-Dithio-*n*-octanoic acid (lipoic acid)	Oxidative decarboxylation	Required by some microorganisms
Iron porphyrin	In catalase, peroxidase, cytochromes	Protoporphyrin
Glucose-1,6-diphosphate	Phosphoglucotransferase	None
Cobalamin	Methylmalonyl CoA mutase	B_{12}

phosphate in amino group transfer; of tetrahydrofolic acid in the transfer of formyl groups; and of coenzyme A in the transfer of acyl groups. In the last-mentioned case acetyl groups, derived from the oxidation of alcohol or lactic, pyruvic, or fatty acids, may be transferred to oxaloacetate with the production of citrate, which in turn is subject to further oxidation by way of the tricarboxylic acid or Krebs cycle. *See* ADENOSINE DIPHOSPHATE (ADP); ADENOSINE TRIPHOSPHATE (ATP); BIOLOGICAL OXIDATION; CITRIC ACID CYCLE; COENZYME; TRANSAMINATION.

Reaction specificity. As already mentioned, a particular coenzyme may act in concert with any one of a large number of substrate-specific apoenzymes. The nature of the coenzyme then determines the general type of reaction that the holoenzymes can catalyze (**Table 5**).

Wesley A. Volk

Bacterial Catabolism

Bacterial catabolism comprises the biochemical activities concerned with the net breakdown of complex substances to simpler substances by living cells. Substances with a high energy level are converted to substances of low energy content, and the organism utilizes a portion of the released energy for cellular processes.

Endogenous and exogenous catabolism. Endogenous catabolism relates to the slow breakdown of nonvital intracellular constituents to secure energy and replacement building blocks for the maintenance of the structural and functional integrity of the cell. This ordinarily occurs in the absence of an external supply of food. Exogenous catabolism refers to the degradation of externally available food. The principal reactions employed are dehydrogenation or oxygenation (either represents biological oxidation), hydrolysis, hydration, decarboxylation, and intermolecular transfer and substitution. The complete catabolism of organic substances results in the formation of carbon dioxide, water, and other inorganic compounds and is known as mineralization. Catabolic processes may degrade a substance only part way. The resulting intermediate compounds may be reutilized in biosynthetic processes, or they may accumulate intra- or extracellularly. The degree of breakdown depends on the genetic constitution of the particular bacteria, and also can be markedly influenced by environmental conditions. Catabolism also implies a conversion of the chemical energy into a relatively few energy-rich compounds or "bonds," in which form it is biologically useful; also, part of the chemical energy is lost as heat.

Intermediary metabolism. Bacterial intermediary metabolism relates to the chemical steps involved in metabolism between the starting substrates and the final product. Normally these intermediates, or precursors of subsequent products, do not accumulate inside or outside the bacterial cell in significant amounts, being transformed serially as rapidly as they are formed. The identification of such compounds, the establishment of the coenzymes and enzymes catalyzing the individual reaction steps, the identity of active forms of the intermediates, and other details of the reaction mechanisms are the objectives of a study of bacterial intermediary metabolism.

In studying intermediary metabolism, the usual objective is to interrupt the normal reaction sequence artificially so that the formation of one or more intermediates from their precursors continues. Since the intermediates cannot then be further transformed, they accumulate. Intermediates may be detected qualitatively and quantitated by fairly specific procedures. In some cases, they are isolated in pure form and identified. A number of chemical, biological and physical methods are applied to this kind of study and usually are complementary. Frequently it is desirable to simulate the reaction and study it with enzymes extracted from the living cell. The respiratory quotient RQ (moles CO_2 produced/moles O_2 consumed) provides a rough index of the chemical nature (relative state of reduction) of a material being oxidized by a population of live bacteria. By comparison with the theoretical values for complete oxidation, valuable information relative to the extent of oxidation may be obtained. The instrument universally employed for measuring the gas exchange on a micro scale is the Warburg respirometer. Changes in pressure of CO_2 and O_2 are measured at constant gas volume and temperature in individual manometers attached to vessels containing substrate and suspensions of cells.

Uniform physiochemical environment is ensured by constant shaking in a constant-temperature water bath. Chemical or enzymatic reactions in which any gas is formed or consumed can also be studied kinetically and quantitatively.

The Thunburg technique is used to study oxidation of a substrate occurring by dehydrogenation reactions. A reversibly oxidizable indicator, methylene blue, substitutes for molecular oxygen as the ultimate hydrogen acceptor (oxidant), becoming reduced to the colorless leuco form. Rigorous exclusion of gaseous oxygen is essential.

Triphenyl tetrazolium compounds can be used instead of methylene blue. The oxidized form is colorless; the reduced form is a red formazan. Tetrazolium indicators have the advantage over methylene blue that they are reoxidized by air only very slowly; hence anaerobic conditions are not essential.

Metabolite utilization determines the probable involvement of a supposed intermediate in a metabolic pathway. If the hypothetical intermediate is metabolized, it must be examined further. However, negative results may only indicate the failure of an intermediate to penetrate the permeable membrane of the cell. In this event, cell-free enzyme preparations of the bacteria are helpful.

Metabolite replacement by analogs or antimetabolites, chemicals structurally related to the natural intermediates, is a method for interfering fairly specifically with a given reaction or pathway. Usually a degree of inhibition of a reaction can be achieved and the effect on the formation of an end product or of a function studied.

Inhibition of a specific reaction on a pathway by any of a number of means, including metabolite analogs or nonspecific poisons, frequently results in the accumulation of a precursor at the blocked reaction. The accumulation may be intracellular, extracellular, or both. *See* POISON.

Isolation and determination of the identity of the accumulated intermediates is then possible, whereas normally the intermediates exist in unisolable concentrations.

Paper, vapor phase, and column chromatographic techniques are invaluable physical procedures for the separation, purification, isolation, and sometimes identification of compounds from multicomponent mixtures such as extracts of bacterial cells or cultural filtrates. Tracer techniques employ metabolites labeled uniformly or in particular atoms with radioactive or stable isotopes. The metabolic path of the metabolite is traced by following the distribution and fate of the labeled elements, usually in various other metabolites, with suitable instruments. Concentrations of intermediates below the sensitivity of chemical isolation or detection procedures can readily be studied with tracers. *See* CHROMATOGRAPHY; RADIOISOTOPE (BIOLOGY).

Biochemical mutation is a tool for securing an inheritable block in a metabolic pathway. A gene controlling the synthesis or function of a specific enzyme is permanently altered mutationally, usually via irradiation or treatment with nitrogen mustard gas. This alteration implies that the corresponding enzyme does not function in the mutated cell. Since each of the steps in a metabolic chain reaction is mediated by its own specific enzyme, the result is that the chain is interrupted. If, for example, a biochemical process normally proceeds by a consecutive conversion of substance S_1 to product S_n via the steps

$$S_1 \xrightarrow[E_1]{} S_2 \xrightarrow[E_2]{} S_3 \xrightarrow[E_3]{} S_4 \xrightarrow[E_4]{} S_5 \xrightarrow[E_5]{} S_6 \xrightarrow[E_6]{} \cdots \xrightarrow[E_{(n-1)}]{} S_n$$

and the individual step reactions involve the enzymes E_1, E_2, E_3, E_4, E_5, . . . , E_{n-1} respectively, then a mutated cell lacking a functioning E_4 can convert S_1 only as far as S_4. Hence this substance, now no longer subject to conversion to S_5, tends to accumulate, often in sufficient quantity to permit its isolation and hence identification as an intermediate product in a reaction chain. *See* MUTATION.

Adaptive or induced enzymes are synthesized by the bacterial cell only in the presence of a specific inducing substance. The product of the action of the first induced enzyme in turn functions as an inducer of a second enzyme, and so on, until the original substrate is converted to products acted on by the central constitutive enzymatic machinery of the cell. If the hypothetical intermediate is utilized nonadaptively by a bacterium preadapted to a given substrate, that is, if the cell is simultaneously adapted to attack it, this is taken as presumptive evidence that the postulated intermediate is on the pathway of utilization of the original substrate.

Coenzyme action can be deduced by determining the biochemical or metabolic alterations caused in bacteria by affecting the function of specific enzymes through interference with the participation of essential inorganic or organic coenzymes. This may be done by creating a deficiency, by limiting the amount of coenzyme furnished as exogenous nutrients, or through vitamin or coenzyme analogs. In cell-free extracts, the coenzyme deficiency may be obtained by means of dialysis.

Catabolism: Oxidative Metabolism

Bacterial oxidative metabolism is manifested by the decomposition of a tremendous number of different substances by one or another species of bacteria. Virtually any naturally occurring substance theoretically capable of yielding energy through oxidation is utilized for that purpose by some bacterial organism or other. Some bacteria are extremely versatile and can utilize any one of a great number of different organic substrates as sources of energy and carbon for growth. Other bacteria attack limited numbers of compounds, and some are extremely specific for a single compound or for very few compounds. Some are obligately linked to the oxidation of the reduced forms of a single inorganic element. In the great majority of cases, the substance, or substrate, undergoing oxidative attack is converted for the most part to bacterial cell, carbon dioxide, and complex mucoid substances. The accumulation in appreciable yields of incomplete oxidation or degradation products

by aerobic bacteria is not common, although several instances are known. In terms of substrate utilized the energy requirements of bacteria are considerably in excess of the carbon requirements for cell synthesis, and so it is characteristic of bacterial oxidations that relatively large amounts of substrates are oxidized in relation to the amount of cellular material synthesized. The more completely a substrate is decomposed by bacteria, the less of it needs to be oxidized to yield a given amount of bacterial growth.

Aerobic oxidations. Aerobic oxidations are reactions in which the substrate is oxidized with concomitant reduction of oxygen to water. The oxidizable substrate may be represented by the generalized symbol H_2A. Most bacterial oxidations take place via dehydrogenation of the hydrogen donors. This occurs by transfer of two electrons *e* and two protons, resulting in the formation of the oxidized product A as shown in reaction (1). The dehydrogenation

$$\underset{\text{Hydrogen donor (substrate)}}{H_2A} \xrightarrow[\text{(oxidation)}]{\text{dehydrogenation}} A + 2H^+ + 2e \qquad (1)$$

reaction will proceed only when the equilibrium is shifted to the right by removal of one or more products of the reaction. This is done by hydrogen acceptors, or oxidants, which take up the protons and electrons, thereby becoming reduced. The oxidant may be denoted by the symbol B, and its reduction by reaction (2).

$$B + 2H^+ + 2e \xrightarrow[\text{(reduction)}]{\text{hydrogenation}} H_2B \qquad (2)$$

Coupling the two reactions, the net transformation for the bacterial oxidation of any substance may be represented as the sum of these reactions, that is, reaction (3). The transfer of protons (active hydrogen) and electrons is mediated by a series of carriers present in catalytic concentrations, and the equation merely represents the overall transformation by which all bacteria secure energy.

$$H_2A + B \longrightarrow A + H_2B + \text{energy} \qquad (3)$$

Another less common mechanism of bacterial oxidation is oxygenation, the direct incorporation of molecular oxygen in certain substrates in the initial oxidative step. These reactions are catalyzed by oxidases and may be represented by the generalized reactions (4) and (5).

$$\underset{\text{Substrate}}{A} + O_2 \longrightarrow \underset{\text{Oxidized substrate}}{AO_2} \qquad (4)$$

$$\underset{\text{Substrate}}{AH} + 2[H] + O_2 \longrightarrow \underset{\text{Oxidized substrate}}{AOH} + H_2O \qquad (5)$$

Depending on the bacterial species, or even strain, and the environmental conditions, H_2A may be any one of a great variety of organic and a lesser number of inorganic substances.

Different organisms are capable of employing different hydrogen acceptors, or oxidants, for the oxidation of a substrate, and this provides a basis for classification of the bacteria as well as for distinguishing types of bacterial oxidations. Most bacteria utilize molecular oxygen as the ultimate hydrogen acceptor; these are, accordingly, aerobic bacteria, or aerobes. In the absence of oxygen, that is, under anaerobic conditions, they can use other hydrogen acceptors in lieu of O_2, such as methylene blue or tetrazolium compounds. Some of them can also use nitrate, nitrite, or nitrous oxide (N_2O) in that capacity. Some bacteria that cannot live in the presence of oxygen (anaerobic bacteria) use sulfate or carbon dioxide as hydrogen acceptors (oxidants). As in the case of oxygen, which yields H_2O, the reduced forms of these oxidants accumulate as the result of the bacterial transformations; each plays the role of B under appropriate conditions, reaction (6).

$$B + 2H^+ + 2e \xrightarrow[\text{(reduction)}]{\text{hydrogenation}} H_2B \qquad (6)$$

See BIOLOGICAL OXIDATION; FERMENTATION.

Metabolism of inorganic substances. When the oxidation of the reduced forms of inorganic substances is utilized as the sole source of energy for growth at the expense of carbon dioxide as the sole source of carbon, the process is known as autotrophy. The use of organic substances for energy and carbon assimilation is known as heterotrophy.

1. Molecular hydrogen. Molecular hydrogen is used autotrophically by certain bacteria, all of which are capable of growing either autotrophically or heterotrophically; they are facultative autotrophs. The growth-supporting reaction (sometimes called the Knallgasreaktion, that is, the explosive gas reaction) which these bacteria carry out is shown in reaction (7).

$$2H_2 + O_2 \longrightarrow 2H_2O + \text{energy} \qquad (7)$$

The best-known hydrogen bacteria belong to the genera *Hydrogenomonas* and *Mycobacterium*. Many heterotropic bacteria can also carry out the previous reaction, but they are incapable of growing at the expense of the energy derived from that reaction exclusively.

2. Iron compounds. Iron may occur in nature in the form of salts and complexes of organic compounds. Bacteria that decompose the organic moiety will therefore cause the formation of ferric carbonate or hydroxide. Such organisms have often been designated as "iron bacteria" because their activities lead to the precipitation of $Fe(OH)_3$.

Nevertheless, this term was first proposed for the members of a special physiological group of bacteria, characterized by their ability to oxidize ferrous to ferric ion and to live at the expense of the energy liberated in this oxidation; they can therefore grow as strict autotrophs. Thus far, this ability has been definitely established for only one bacterial species, *Ferrobacillus ferrooxidans*, and made probable for another, *Gallionella ferruginea*. Still others,

such as the filamentous, sheathed bacteria known as *Leptothrix ochracea* and related species, that regularly occur along with *Gallionella* in localities where ferrous ion–containing seepage water runs into an oxygen-rich environment and $Fe(OH)_3$ precipitates may also belong to this physiological group. However, the *Leptothrix* species are now recognized as special environmental forms of *Sphaerotilus* species, which can also thrive in media not containing ferrous iron where they grow at the expense of the oxidation of organic substances.

The oxidation of ferrous to ferric iron may be represented by reaction (8).

$$4FeCO_3 + O_2 + 6H_2O \longrightarrow 4Fe(OH)_3 + 4CO_2 + \text{energy} \quad (8)$$

3. Reduced sulfur compounds. Inorganic sulfur compounds at various levels of reduction, such as sulfite, tetrathionate, thiosulfate, elementary sulfur, and hydrogen sulfide, can be used as oxidizable substrates by many kinds of bacteria. Generally the oxidation yields sulfate as the end product, as in reactions (9) and (10); if H_2S is the substrate, it is oxidized

$$2S + 3O_2 + 2H_2O \longrightarrow 2H_2SO_4 + \text{energy} \quad (9)$$

$$\underset{\text{Sodium thiosulfate}}{Na_2S_2O_3} + 2O_2 + H_2O \longrightarrow Na_2SO_4 + H_2SO_4 + \text{energy} \quad (10)$$

initially to elementary sulfur, as in reaction (11). In the case of the relatively large representatives, such

$$2H_2S + O_2 \longrightarrow 2S + 2H_2O + \text{energy} \quad (11)$$

as species of *Beggiatoa, Thiothrix, Thiovulum*, and *Chromatium*, this sulfur accumulates in the bacterial cells in the form of characteristic highly refractile droplets which apparently can serve as a reserve source of substrate and disappear if the supply of H_2S gets too low.

Many sulfur bacteria are obligately autotrophic. Physiologically the best known are the small, rod-shaped *Thiobacillus* species. *Thiobacillus thioparus* thrives in neutral media and *T. thiooxidans* in acid media; the latter can still grow in solutions containing as much as 5% sulfuric acid. *Thiobacillus novellus* is a facultatively autotrophic thiosulfate oxidizer.

The oxidation of sulfide is postulated to proceed via sulfur (hydrate), sulfoxylic acid, and sulfurous acid; that of thiosulfate via tetrathionate, trithionate, and sulfite.

4. Nitrogen compounds (nitrification). The best-known bacteria which oxidize reduced inorganic nitrogen compounds are the members of the obligately autotrophic rods comprising two genera, *Nitrosomonas* and *Nitrobacter*. *Nitrosomonas europaea* oxidizes ammonia to nitrous acid via hydroxylamine; nitroxyl (NOH) has been postulated as the subsequent intermediate oxidation stage, as shown by reactions (12*a*) and (12*b*). *Nitrobacter winogradskii* oxidizes nitrite to nitrate by way of reaction (13).

$$\underset{\text{Ammonia (hydrate)}}{NH_4OH} + {}^1/_2O_2 \longrightarrow \underset{\text{Hydroxylamine}}{NH_2OH} + H_2O + \text{calories} \quad (12a)$$

$$\underset{\text{Hydroxylamine}}{NH_2OH} + O_2 \longrightarrow \underset{\text{Nitrous acid}}{HNO_2} + H_2O + \text{calories} \quad (12b)$$

$$\underset{\text{Nitrous acid}}{HNO_2} + {}^1/_2O_2 \longrightarrow \underset{\text{Nitric acid}}{HNO_3} + \text{energy} \quad (13)$$

Nitrogen-free aliphatic compounds. Although the number of oxidizable compounds and bacteria capable of effecting their oxidation are almost infinitely large, patterns of oxidation common to particular types of substrates appear to be the rule. In general, the oxidation pathways shape the substrate so that it can feed into the central terminal oxidation mechanisms almost universal in bacteria. In the main, the product entering the terminal oxidation cycle is active acetate, that is, acetyl coenzyme A (acetyl-CoA). Thus the special pathways for the decomposition of the diverse organic substrates in reality are mechanisms for their conversion to acetic acid. Most carbon assimilated by the bacterial cell is drained from the oxidation pathway at or following the acetate stage.

1. *Hydrocarbons.* A variety of bacteria, especially in the genera *Mycobacterium* and *Nocardia*, oxidatively attack paraffinic and olefinic hydrocarbons, including gases, liquids, and solids. The available information indicates that paraffinic hydrocarbons are oxidized monoterminally, probably via the addition of molecular oxygen to form the hydroperoxide, which in turn yields the corresponding primary alcohol by way of reaction (14). The gaseous alkanes can

$$\underset{\text{Aliphatic paraffinic hydrocarbon}}{RCH_2CH_3} + O_2 \longrightarrow \underset{\text{Hydroperoxide}}{RCH_2CH_2(OOH)} \xrightarrow{+2H} \underset{\text{Primary alcohol}}{RCH_2CH_2OH} + H_2O \quad (14)$$

also be oxidized subterminally; in that case they yield methyl ketones. The proposed mechanism is shown in reaction (15). Palmitate waxes (esters) are

$$\underset{\text{Gaseous alkane}}{RCH_2CH_3} + O_2 \longrightarrow RCH(OOH)CH_3 \xrightarrow{+2H} \underset{\text{Secondary alcohol}}{RCHOHCH_3} \xrightarrow{-2H} \underset{\text{Methyl ketone}}{RCOCH_3} \quad (15)$$

produced by certain bacteria during growth on long-chain liquid alkanes. The alcohol moiety of the ester corresponds to the substrate hydrocarbon.

2. *Alcohols.* Primary low-molecular-weight alcohols are oxidized by many bacteria, the first stage, reaction (16), being the corresponding aldehyde. The

$$\underset{\text{Primary alcohol}}{RCH_2OH} \xrightarrow{-2H} \underset{\text{Aldehyde}}{RCHO} \qquad (16)$$

enzyme alcohol dehydrogenase is responsible. The aldehyde is almost invariably rapidly oxidized to the corresponding carboxylic acid by hydration and dehydrogenation, as in reaction (17).

$$\underset{\text{Aldehyde}}{RCHO} \xrightarrow{+HOH} \underset{\text{Aldehyde hydrate}}{RC(H)(OH)(OH)} \xrightarrow{-2H} \underset{\text{Carboxylic acid}}{RC(=O)OH} \qquad (17)$$

Although most bacteria that can carry out this type of oxidation rapidly convert the carboxylic acids to further oxidation products, some, notably members of the genera *Acetobacter* and *Acetomonas*, cause the accumulation of such acids, often in high yields. *See* VINEGAR.

Primary alcohols with more than five carbon atoms in their chain are toxic, and can be oxidized only if present in low concentration.

Secondary alcohols are typically oxidized to the corresponding ketones, as shown in reaction (18).

$$\underset{\text{Secondary alcohol}}{R-CHOH-R'} \xrightarrow{-2H} \underset{\text{Ketone}}{R-CO-R'} \qquad (18)$$

The ketones often accumulate. Tertiary alcohols are also decomposed, but nothing is known about the consecutive oxidation stages.

3. *Ketones.* Only the oxidation of the short-chain methyl ketones, particularly of acetone, has been studied in some detail. Acetone is photosynthetically carboxylated by *Rhodopseudomonas gelatinosa* to acetoacetic acid, which is then oxidized to acetic acid. This mechanism is represented by reaction (19).

$$\underset{\text{Acetone}}{CH_3COCH_3} + CO_2 \xrightarrow{\text{light}}$$

$$\underset{\text{Acetoacetic acid}}{CH_3COCH_2COOH} \longrightarrow \underset{\text{Acetic acid}}{2CH_3COOH} \qquad (19)$$

This mode of attack may also occur in the nonphotosynthetic bacterial oxidation of methyl ketones. However, a different type of mechanism, reaction (20), has been discovered with an acetone-

$$\underset{\text{Acetone}}{CH_3COCH_3} + {}^1/_2O_2 \longrightarrow \underset{\text{Acetol}}{CH_3COCH_2OH} \longrightarrow$$

$$\underset{\text{Acetaldehyde}}{CH_3CHO} + \underset{\text{Active carbon fragment}}{C_1} \qquad (20)$$

oxidizing diphtheroid bacterium. This causes an initial oxidation to acetol, which in turn is broken down to acetaldehyde and an active one-carbon fragment, presumably formaldehyde or formate.

4. *Fatty acids.* Fatty acids are decomposed by oxidative attack at the beta carbon atom, resulting in the formation of acetic acid (active acetate; acetyl-CoA). The acetate represents the carboxyl and the alpha carbons of the carbon chain. A fatty acid with two carbons less is always formed, and the process is repeated serially. The first step is the activation by formation of the fatty acid thioester with coenzyme A (RCO—SCoA). As in **Fig. 2**, two mechanisms are known for the activation of a substrate fatty acid.

The first involves enzymatic displacement of acyl kinase of the pyrophosphate group (PP) of ATP by the fatty acid, forming acyl adenylate. Magnesium ions are required. An enzymatic reaction with CoA displaces adenylic acid, resulting in the thioester (acyl-CoA). The second mechanism of fatty acid activation takes place by direct enzymatic transfer of CoA by CoA transphorase from a donor acyl-CoA which itself may be the product of fatty acid oxidation. The acyl-CoA is dehydrogenated between the α- and β-carbons by acyl-CoA dehydrogenase. At least two forms of this enzyme are known. Both are metalloflavoproteins. The prosthetic group of both is FAD. The one which attacks short-chain fatty acids is a copper enzyme; the one attacking longer chains may be an iron enzyme. The addition of the elements of water to the double bond is accomplished by a hydrase, specifically crotonase. This enzyme has no known coenzyme, and it also provides a means of reversible conversion of β,γ-unsaturated acids to α,β-unsaturated acids via a common β-hydroxy acid intermediate shown in reaction (21).

$$RCH=CHCH_2CO-SCoA \overset{HOH}{\rightleftharpoons}$$

$$RCH_2CHOHCH_2CO-SCoA \overset{HOH}{\rightleftharpoons}$$

$$RCH_2CH=CHCO-SCoA \qquad (21)$$

The β-hydroxy acyl-CoA is oxidized by β-hydroxyl acyl-CoA dehydrogenase, the enzyme-containing NAD. All acids containing 4–12 carbon atoms (C_4 to C_{12}) are oxidized at approximately the same rate. Only the *d*-isomer is oxidized; the *l*-isomer is inert. The β-ketoacyl-CoA is then cleaved by a β-ketothiolase (acetyl-CoA transacetylase) between the α- and β-carbons with the concomitant uptake of 1 mole of CoA. The original CoA remains on the two-carbon cleavage fragment as acetyl-CoA. The thiolase reaction adds the second CoA molecule to the remaining carbon chain, at the carbonyl group, forming an acyl-CoA chain which is shorter by two carbons than the original fatty acid. *See* OPTICAL ACTIVITY.

This shorter acyl-CoA moiety is further oxidized by the same general mechanism. Hence the fatty acids that contain an even number of carbon atoms and have the general composition $C_{2n-1}H_{4n-1}COOH$ eventually yield a corresponding number, n, of

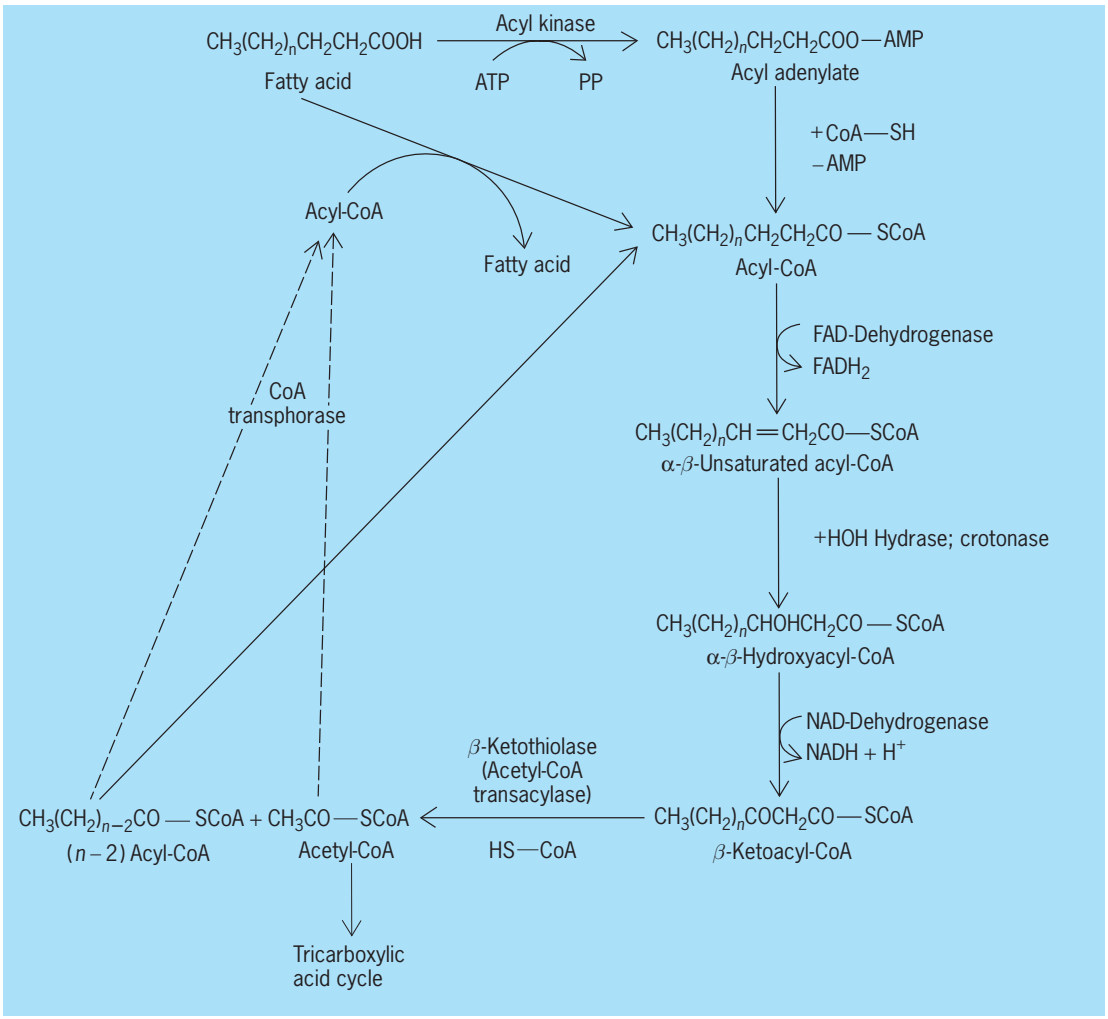

Fig. 2. The two known mechanisms for the oxidative decomposition of fatty acids.

acetyl-CoA molecules. The fatty acids with an odd number of carbon atoms, $C_{2n}H_{4n+1}$ COOH, yield one molecule of propionyl-CoA in addition to $n - 1$ molecules of acetyl-CoA.

Acetyl-CoA is further oxidized via the citric acid or glyoxylic acid cycle, discussed later. It is probable that propionyl-CoA is degraded by an initial carboxylation to methyl malonyl–CoA, which is next converted to succinyl-CoA, as in reaction (22). The succinate can then be oxidized to acetate via fumarate, malate, and oxaloacetate.

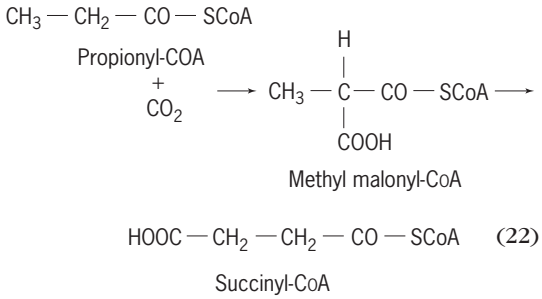

(22)

Certain bacteria produce acetyl phosphate, apparently as an energy reservoir. Acetyl phosphate is readily converted to acetyl-CoA by the enzyme phosphotransacetylase in reaction (23).

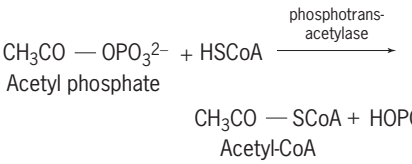

(23)

As in biological oxidations in general, the pattern of bacterial oxidations is to chip large and more complex substrates into small carbon chain fragments. As a rule, these fragments are then broken down in stepwise fashion, by fairly complex cyclic systems, to CO_2, water, and energy in a process known as terminal oxidation.

The Krebs tricarboxylic acid (TCA) system is the most common cyclic mechanism of terminal oxidation in biological systems. It was conceived and clarified by H. A. Krebs, a British biochemist who received a Nobel prize for this contribution. The mechanism appears to be identical in bacterial and

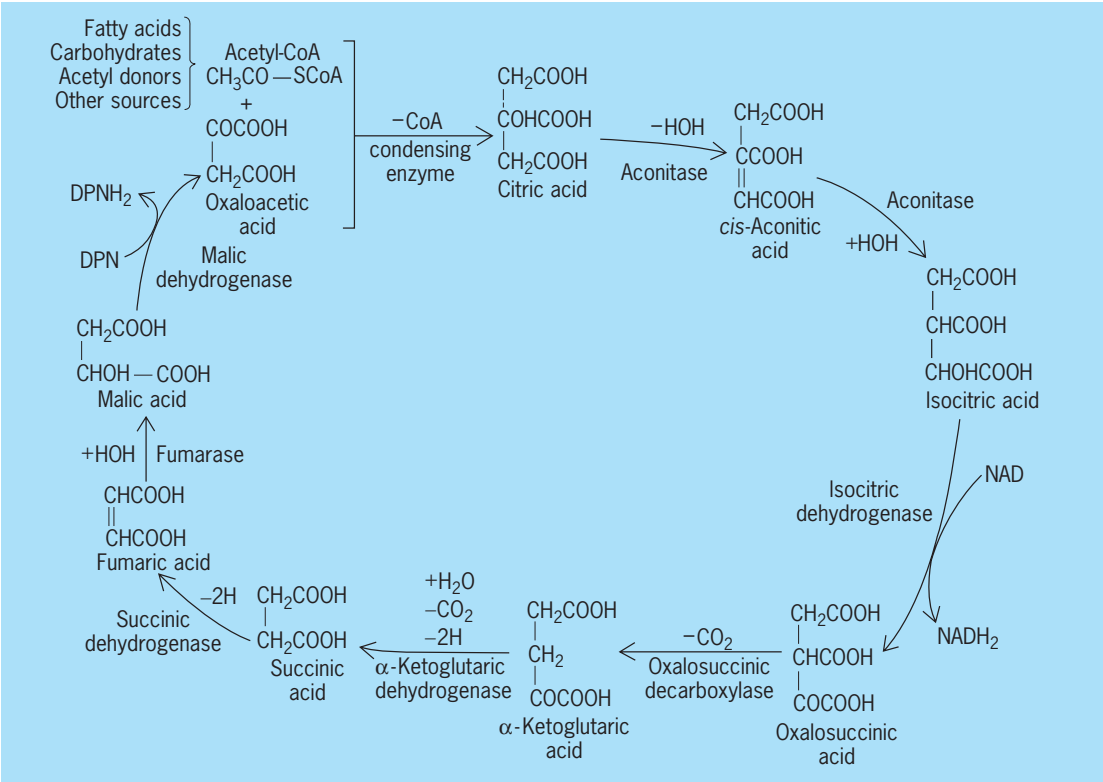

Fig. 3. Pathways of the Krebs tricarboxylic acid cycle, a common biological oxidative mechanism.

other microorganisms and in plant and animal cells. Minor modifications or adaptations of the Krebs cycle are known. The TCA system is a cyclic pathway for the oxidation of acetyl-CoA derived from fatty acid or carbohydrate breakdown, from acetyl donors, or from other types of metabolism (**Fig. 3**) to CO_2 and water, producing high-energy bonds. In the presence of assimilable nitrogen and mineral salts, various components are drained off from the cycle to serve as precursors of amino acid and other cellular components. Under resting cell conditions the substrate may be oxidized completely via the citric acid cycle, or variable amounts may be drained off and converted to products of oxidative assimilation. The oxidative formation of acetyl-CoA from pyruvic acid will be discussed later under triose pathway in connection with carbohydrate oxidation. Oxaloacetic acid is the key component of the cycle; it functions as the initial acceptor of acetate into the cycle. A transfer of acetate is effected by a specific enzyme, oxaloacetate transacetylase, better known as citrate-condensing enzyme, which links the methyl carbon of the acetyl group with the carbonyl carbon of oxaloacetate, splits out CoA, and forms citric acid.

In the next step, the elements of water are removed from citrate by the enzyme aconitase, leaving the asymmetrical unsaturated *cis*-aconitic acid. The elements of water are then added back to the double-bond carbons, but in the reverse positions, forming isocitric acid. There is some evidence for two different aconitases effecting the dehydration and the hydration. The simultaneous dehydrogenation of isocitrate at the carbinol group and decarboxylation of the tertiary carboxylic group are believed to be catalyzed simultaneously by an NAD-dependent enzyme, isocitric dehydrogenase, and by oxalosuccinic decarboxylase. The product, α-ketoglutaric acid, is oxidatively decarboxylated to succinic acid. The coenzymes diphosphothiamine, lipoic acid, and CoA are sequentially involved in this overall transformation, which comprises three intermediate enzymatic steps. This system is similar to that which oxidatively decarboxylates pyruvic acid and which is discussed below. Succinyl-CoA is the immediate precursor of the succinic acid; the acid is liberated by succinyl-CoA deacylase. Succinic acid is then oxidized to fumaric acid by succinic dehydrogenase, a flavin-dependent enzyme. The fumarate is next hydrated by fumarase, forming L-malic acid, which in turn regenerates oxaloacetate after its oxidation by means of the NAD-dependent malic dehydrogenase. The hydrogens, or electrons, are converted stepwise to water via the terminal electron transport system. High-energy phosphate bonds are thereby generated at various stages; their formation constitutes oxidative phosphorylation.

In the living cell, where the keto acid components are siphoned off for the synthesis of amino acids, the cycle would quickly run down because there would no longer be a continuous supply of oxaloacetic acid.

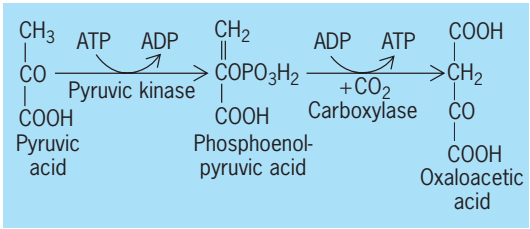

Fig. 4. Wood-Werkman reaction.

In addition, this substance itself can be readily decarboxylated to carbon dioxide and pyruvic acid, which constitutes an additional drain. The level of oxaloacetate can, however, be maintained by a carboxylation of pyruvic acid derived from other sources, mainly carbohydrate. One such reaction is the so-called Wood-Werkman reaction (**Fig. 4**), in which pyruvic acid is converted to phosphoenolpyruvic acid, the actual acceptor of CO_2. The fixation is coupled with the regeneration of an energy-rich pyrophosphate bond; either ADP or inosine diphosphate functions as the phosphate acceptor. The second carboxylation reaction is that accomplished by the malic enzyme; it represents a simultaneous reduction and carboxylation of pyruvic acid to malic acid by a NAD-dependent enzyme, and is shown in reaction (24).

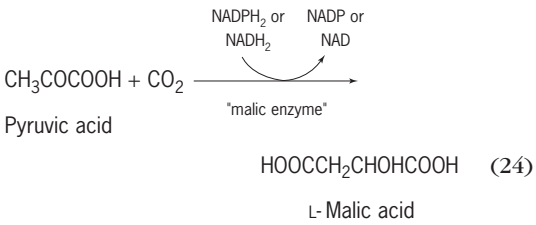

(24)

It has been recognized, however, that these carboxylation reactions are inadequate to generate sufficient oxaloacetate, particularly in those cases where acetate is the main or only oxidizable substrate. This has long presented a problem, which has been solved by the discovery of the glyoxylic acid cycle, which permits the formation of oxaloacetate, and hence also of citrate, from two-carbon fragments exclusively. The cycle comprises a condensation of acetyl-CoA with glyoxylic acid to malic acid under the influence of the enzyme malate synthetase, as in reaction (25). The malate is then oxidized to oxaloacetate.

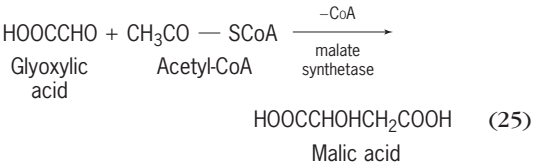

(25)

A continuous supply of glyoxylic acid is ensured by a conversion of isocitrate other than its oxidation to oxalosuccinate. It is a cleavage, under the influence of the enzyme isocitratase, to glyoxylate and succinate according to reaction (26). The cycle is

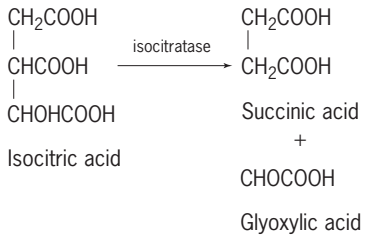

(26)

schematically represented in **Fig. 5**. The net change in one turn of the glyoxylic acid cycle is represented by reaction (27). The succinate is metabolized via

$$2\ \text{Acetate} + {}^{1}/_{2}O_2 \longrightarrow \text{succinate}$$

the Krebs cycle or other pathways.

5. *Hydroxy acids.* The β-hydroxy monocarboxylic acids are discussed under fatty acids. The α-hydroxy acids are oxidized to the corresponding α-keto acid by zinc-containing dehydrogenases, as in the conversion of L-lactic acid to pyruvic acid under the influence of lactic dehydrogenase in re-

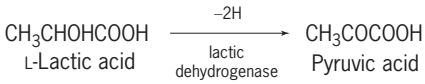

(28)

action (28). Lactate dehydrogenase acts on several α-hydroxy acids. The enzymes are stereospecific for the L forms. Three lactic dehydrogenases are known which differ in their coenzyme requirements. One utilizes NAD, a second uses flavin nucleotide, and the third cytochrome *c*; the last one is oxygen linked.

6. *Keto acids.* The most important keto acids are the α- and β-ketomonocarboxylic acids. The α-keto acids may be successively decarboxylated to the next smaller aldehyde by a diphosphothiamine

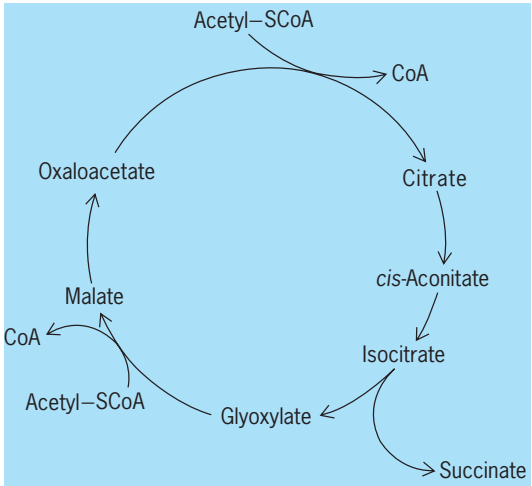

Fig. 5. Glyoxylic acid cycle.

decarboxylase, and then oxidized, reaction (29). The

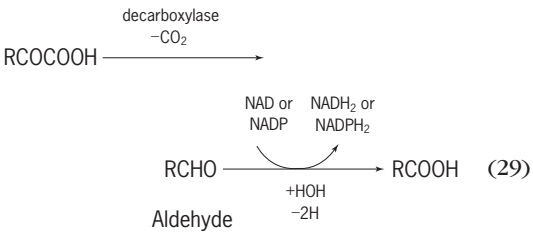

(29)

oxidative decarboxylation of pyruvic acid involves a series of coenzymes as shown in **Fig. 6**.

For β-ketomonocarboxylic acids oxidation, see the discussion of fatty acids; for β-ketoadipic acid oxidation, see the section on aromatic compounds.

7. *Carbohydrates.* The oxidation of carbohydrates occurs by different routes, all of which may occur in any one bacterial species. However, one or more may be lacking in some bacteria.

The triose pathway is responsible for the production of pyruvic acid by the Embden-Meyerhof mechanism of glycolysis, described later in this article. The pyruvic acid is oxidatively decarboxylated to acetyl-CoA as described above, which in turn is oxidized by way of the citric acid cycle.

In aerobic bacteria, the rupture of the carbon chain of glucose occurs after oxidation to 6-phosphogluconate. This compound may be oxidatively decarboxylated to ribulose-5-phosphate, which is further metabolized via the pentose phosphate pathway described later. Alternatively, the 6-phosphogluconate may be dehydrated at carbon atoms two and three, forming 2-keto, 3-deoxy, 6-phosphogluconic acid. This pathway is also described later (Entner-Doudoroff).

The net effect of six turns of the hexosemonophosphate cycle is the complete oxidation of one molecule of glucose, as represented by reactions (30*a*) and (30*b*).

$$6\ \text{Glucose} + 6O_2 \longrightarrow 5\ \text{glucose} + 6H_2O + 6CO_2 \qquad (30a)$$

$$\text{Net: Glucose} + 6O_2 \longrightarrow 6H_2O + 6CO_2 \qquad (30b)$$

The oxidative steps are those between glucose and ribulose-5-phosphate (RP) [**Fig. 7**]. Glucose is regenerated from RP by interconversion of five-, six-, and seven-carbon sugar phosphate esters. The terminal two carbons of one RP are transferred by transketolase (a diphosphothiamine enzyme) to a second RP, forming a seven-carbon sugar (D-sedulose-7-phosphate) and leaving glyceraldehyde-3-phosphate (G-3-P). By transaldolase, the terminal three carbons of the sedoheptulose are condensed with the G-3-P to produce fructose-6-phosphate. The tetrose remaining from the seven-carbon sugar is condensed with a terminal two-carbon fragment from another RP, forming fructose-6-phosphate and leaving G-3-P. The latter is converted to an equilibrium mixture with dihydroxyacetonephosphate by triose phosphate isomerase. These trioses are then condensed to hexose diphosphate by the action of aldolase.

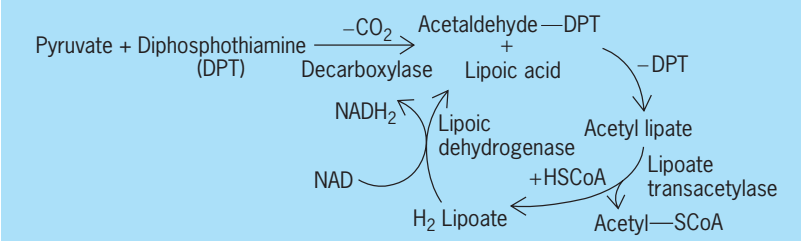

Fig. 6. Oxidative decarboxylation of pyruvic acid showing coenzymes involved.

Oxidation of aromatic compounds. This property is found mainly in aerobic bacteria belonging to the genera *Pseudomonas, Vibrio*, and *Mycobacterium*. As a rule, the aromatic compounds are converted to a benzene ring compound containing two adjacent carboxyls. The catechol and the protocatechuic acid pathways are the best known, both eventuating in β-ketoadipic acid (β-KAA). Phenol is oxidized directly to catechol. Mandelic acid, tryptophan, and indoleacetic acid are converted to catechol via the

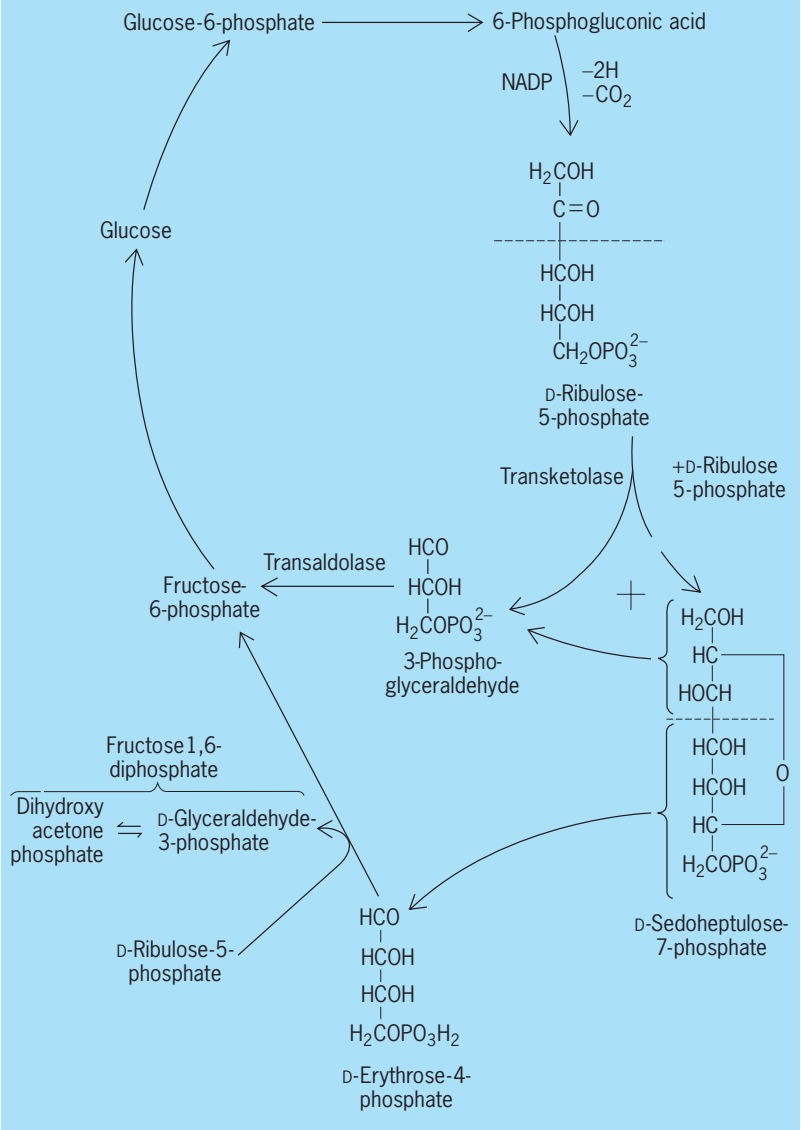

Fig. 7. Carbohydrate oxidation in aerobic bacteria.

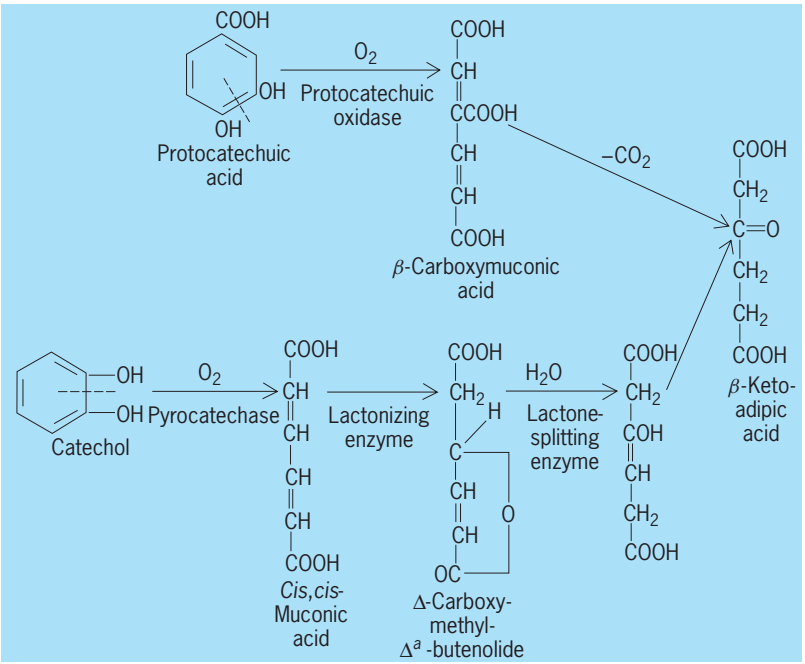

Fig. 8. Cleavage of dihydroxy ring structures in protocatechuic acid and catechol.

sequences shown in reactions (31), (32), and (33).

D(–)–Mandelic acid $\underset{\text{racemase}}{\rightleftharpoons}$

L(+)–mandelic acid $\xrightarrow[\text{L–mandelic dehydrogenase}]{-2H}$

benzoylformic acid $\xrightarrow[\text{benzoylformic carboxylase}]{-CO_2}$

benzaldehyde $\xrightarrow[\text{benzaldhyde dehydrogenase}]{+H_2O_3-2H}$

benzoic acid $\longrightarrow$ catechol (31)

Tryptophan $\xrightarrow[\text{tryptophan peroxidase}]{O_2,\ H_2O}$ formylkynurenine $\xrightarrow[\text{formylase}]{H_2O}$

kynurenine $\xrightarrow[\text{kynureninase}]{-\text{alanine}}$

anthranilic acid $\longrightarrow$ catechol (32)

Indoleacetic acid $\longrightarrow$ skatole $\longrightarrow$ indoxyl $\longrightarrow$

salicylic acid $\longrightarrow$ catechol (33)

The protocatechuic acid pathway is represented

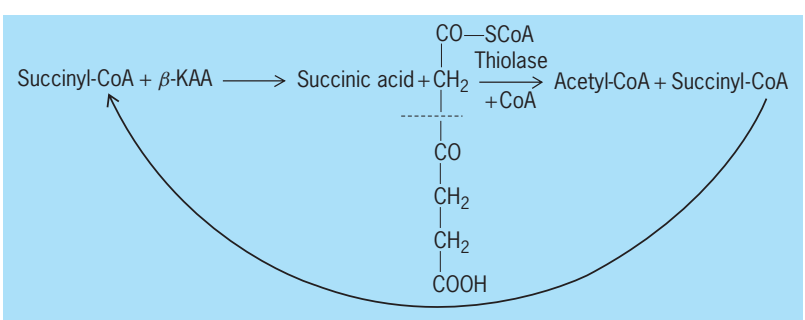

Fig. 9. Splitting of β-ketoadipic acid to form acetyl-CoA and succinyl-CoA.

by reaction sequence (34).

p-Nitrobenzoic acid $\longrightarrow$ *p*-aminobenzoic acid

$\downarrow$

p-cresol $\longrightarrow$ *p*-hydroxybenzyl alcohol $\longrightarrow$

p-hydroxybenzoic acid $\longrightarrow$ protocatechuic acid (34)

The dihydroxy ring structures are then cleaved between the adjacent hydroxyls to form aliphatic dicarboxylic acids. The two pathways converge at β-ketoadipic acid, as shown in **Fig. 8**. Beta-ketoadipic acid is split to acetyl-CoA and succinyl-CoA. This reaction sequence is shown in **Fig. 9**. The succinyl-CoA activates β-ketoadipic acid, and the process is a cyclic one whereby the net products are acetyl-CoA and succinic acid, which are then oxidized via the citric acid cycle.

The oxidative degradation of the amino acids phenylalanine and tyrosine proceeds by a different pathway, as shown in reaction (35).

Phenylalanine $\xrightarrow[\text{phenylalanine hydroxylase}]{}$ tyrosine $\xrightarrow[\text{L-amino acid oxidase}]{+H_2O,\ -2H,\ -NH_3}$

p-hydroxyphenylpyruvic acid $\xrightarrow[\text{oxidase}]{+O_1,\ -CO_2}$

homogentisic acid $\xrightarrow[\text{oxidase}]{O_2}$

4-maleylacetoacetic acid $\xrightarrow[\text{isomerase}]{}$

4-fumarylacetoactic acid $\xrightarrow[\text{β-diketonase}]{H_2O}$

fumaric acid + acetoacetic acid (35)

Polycyclic aromatic hydrocarbons are oxidized by certain bacteria by routes which have not been clarified enzymatically. As a rule, *o*-hydroxycarboxylic acids containing one ring less than the starting compound have been isolated and are considered to be intermediate breakdown products.

Nitrogen-containing organic substances. The complete breakdown of this class ordinarily results in release of all of the nitrogen as ammonia.

1. *Amino acids.* Enzymes are specific for L- and D-amino acids. The key reaction is the splitting out of ammonia by deamination. Four types of deamination are known, the applicable one depending on the amino acid and the bacterial species.

Amino acid oxidases are flavoproteins which attack a number of different amino acids according to reaction (36). The imino acid hydrolyzes spontaneously. The flavoprotein is linked to oxygen and produces hydrogen peroxide. Glycine oxidase is a separate specific amino acid oxidase.

$$\underset{\alpha\text{-Amino acid}}{RCHNH_2COOH} \xrightarrow[\text{amino acid oxidase}]{-2H} \underset{\text{Imino acid}}{RC{=}NHCOOH} \xrightarrow{H_2O} \underset{\alpha\text{-Keto acid}}{RCOCOOH} + \underset{\text{Ammonia}}{NH_3} \quad (36)$$

Glutamic and alanine dehydrogenase are NAD- or NADP-linked enzymes which convert the substrates to α-ketoglutaric acid and pyruvic acid, respectively, with the liberation of ammonia. Oxygen is not essential.

Aspartase catalyzes the deamination of aspartic acid to fumaric acid and ammonia.

Beta-substituted amino acids, that is, serine, threonine, and cysteine, are dehydrated through the process shown in reaction (37).

$$\underset{\substack{\beta\text{-Substituted}\\ \alpha\text{-Amino acid}}}{RCHOHCHNH_2COOH} \xrightarrow[\text{dehydrase}]{-H_2O}$$

$$(RCH{=}CNH_2COOH) \longrightarrow \underset{\text{Imino acid}}{RCH_2C{=}NHCOOH} \xrightarrow{H_2O}$$

$$\underset{\alpha\text{-Keto acid}}{RCH_2COCOOH} + \underset{\text{Ammonia}}{NH_3} \qquad (37)$$

Histidine and hydroxy L-proline are oxidized as shown by reactions (38) and (39).

$$\text{Histidine} \xrightarrow[\text{deamination}]{-NH_3} \text{urocanic acid} \xrightarrow[\text{uricanase}]{HOH}$$

$$\text{formimino-glutamic acid} \begin{cases} \longrightarrow \text{glutamic acid} + \text{formylglycine} \longrightarrow \text{formic acid} + \text{ammonia} \\ \longrightarrow \text{glutamic acid} + \text{formamide} \longrightarrow \text{formic acid} + \text{ammonia} \end{cases} \qquad (38)$$

$$\text{Hydroxly-L-proline} \longrightarrow \text{allohydroxy-D-proline} \longrightarrow$$
$$\Delta'\text{-4-hydroxypyrroline-2-carboxylic acid} \longrightarrow$$
$$\text{2-keto-4-hydroxy-5-aminovaleric acid} \longrightarrow$$
$$\text{glutamic acid} \qquad (39)$$

2. *Purines and pyrimidines.* The oxidation of purines is best known in the case of uric acid and is shown in reaction (40). The formic acid and CO_2

$$\text{Uric acid} \xrightarrow{+O,\ H_2O} \text{hydroxyacetylene} \longrightarrow$$
$$\text{allantoin} \xrightarrow{+H_2O} \text{allantoic acid} \longrightarrow$$
$$2\ \text{urea} + \text{glyoxylic acid} \longrightarrow$$
$$CO_2 + \text{formic acid} \qquad (40)$$

produced from glyoxylic acid are derived from the carboxyl and α-carbons, respectively.

The initial step, reaction (41), in pyrimidine

$$\text{Pyrimidine} \xrightarrow{+2H} \text{dihydropyrimidine} \longrightarrow$$
$$\beta\text{-carbamylamino acid} \longrightarrow$$
$$\beta\text{-amino acid} + CO_2 + NH_3 \qquad (41)$$

degradation is a reduction. From uracil the intermediates are β-carbamylaminopropionic acid and β-aminopropionic acid, respectively; from thymine they are β-carbamylaminoisobutyric acid and β-aminoisobutyric acid. The oxidase pathways shown in reactions (42) and (43) have been de-

$$\text{Uracil} \xrightarrow[\text{oxidase}]{+O} \text{barbituric acid} \xrightarrow[\text{barbiturase}]{+H_2O}$$
$$\text{urea} + \text{malonic acid}$$
$$\downarrow$$
$$\text{acetic acid} + CO_2 \qquad (42)$$

$$\text{Thymine} \xrightarrow[\text{oxidase}]{+O} \text{5-methylbarbituric acid} \qquad (43)$$

scribed in a species of *Mycobacterium*.

3. *Pyridines.* The scanty information on the mechanism of degradation of pyridine compounds is limited to the oxidation of pyridoxine (vitamin B_6), nicotinic acid, and nicotine. These appear to be oxidized by way of the pathways shown in reactions (44), (45), and (46).

$$\text{Pyridoxine} \longrightarrow \text{isopyridoxal} \longrightarrow \text{5-pyridoxic acid} \longrightarrow$$
$$a\text{-hydroxymethyl-}a\text{-(}N\text{-acetylaminomethylene)-succinic acid} \qquad (44)$$

$$\text{Nicotinic acid} \xrightarrow{H_2O}$$
$$\text{6-hydroxynicotinic acid} \xrightarrow{-CO_2,\ +O_2}$$
$$\text{2,5-dihydroxynicotine} \longrightarrow$$
$$\text{fumaric acid} + \text{formic acid} \qquad (45)$$

$$\text{Nicotine} \longrightarrow \text{3-(4-methylaminobutyryl)-pyridine} \longrightarrow$$
$$\text{3-succinyl-6-hydroxypyridine} \longrightarrow ? \qquad (46)$$

4. *Proteins.* Proteins are hydrolytically broken down to the constituent amino acids by means of proteolytic enzymes and peptidases. Polypeptides of gradually diminishing size are produced; free amino acids are ultimately liberated. These are oxidized as described above.

Anaerobic oxidations. A number of different bacteria are capable of utilizing inorganic oxidants as alternatives to oxygen or organic substances as the ultimate electron, or hydrogen, acceptors. The nature of the oxidant and the change it undergoes distinguish the various kinds of anaerobic oxidations. The reduction of the oxidant is part of the energy-liberating process as contrasted to reduction for purposes of assimilation as a nutrient. In the former case, the amount of oxidant that is reduced per unit of growth is very large, and the reduced product accumulates extracellulary.

Nitrate reduction. A great many aerobic bacteria, including species belonging to a variety of different genera, can facultatively utilize nitrate in the absence of oxygen. The enzyme nitrate reductase is a molybdenum-dependent flavin-linked inducible enzyme. Usually, the specific substrates that an organism can oxidize with nitrate as the ultimate oxidant are the same as those it can oxidize with oxygen. The nitrate is reduced to nitrite, which may accumulate or may be further reduced to ammonia by way of hydroxylamine. Nitrite reductase and hydroxylamine

reductase are also flavoproteins. Cytochrome *c* appears to mediate the flow of the electrons to nitrite.

Denitrification. The reduction of nitrate or nitrite to gaseous products constitutes denitrification. Nitrogen gas is the most common product, although nitrous oxide, N_2O, and nitric oxide, NO, are also formed under some conditions. The latter two may be further reduced to N_2, but they are not obligatory intermediates in the denitrification. Nitrite is reduced to nitric oxide by a flavoprotein that requires copper and iron. All of the enzymes in denitrification are capable of receiving electrons from pyridine nucleotide enzymes.

Sulfate reduction. The bacteria effecting sulfate reduction belong to the genus *Desulfovibrio, D. desulfuricans* being the most common species. *Clostridium nigrificans* also is a sulfate reducer. The organisms are obligate anaerobes which reduce sulfate to hydrogen sulfide during oxidation of various substrates. Blackening of the medium results from the interaction of the H_2S with iron. Sulfate reduction is mediated by a cytochrome-linked enzyme, and the active form of sulfate that undergoes reduction is adenosine-3′-phosphate-5′-phosphosulfate. Sulfate-reducing bacteria cause corrosion of iron because, by creating a strongly reduced environment, they locally create anodic conditions relative to the more oxidized surroundings, thus producing a corrosion cell.

Carbonate reduction. Two different types, distinguishable by the nature of the reduction products, are known. They are the following:

1. *CO_2 reduction to acetic acid.* This is performed by certain obligately anaerobic sporeforming bacteria that use CO_2 as the hydrogen acceptor for the oxidation of various substrates. *Clostridium aceticum* thus produces acetic acid from molecular hydrogen and CO_2 in the manner of reaction (47)

$$2CO_2 + 4H_2 \longrightarrow CH_3COOH + 2H_2O \qquad (47)$$

and, like *C. thermoaceticum*, oxidatively decarboxylates pyruvic acid with a concomitant reduction of CO_2 to acetic acid by way of reactions (48). These

$$\begin{array}{l} 4CH_3COCOOH + H_2O \longrightarrow CH_3COOH + CO_2 + 2[H] \\ \qquad 2CO_2 + 8[H] \longrightarrow CH_3COOH + 2H_2O \\ \hline 4CH_3COCOOH + 2H_2O \longrightarrow 5CH_3COOH + 2CO_2 \end{array} \qquad (48)$$

organisms decompose sugar in accordance with reaction (49), and tracer studies with *C. thermoaceticum* have shown that this conversion proceeds by an initial decomposition of the sugar via pyruvic acid, as shown in reactions (50). *Clostridium acidi-urici* carries out an anaerobic degradation

$$C_6H_{12}O_6 \longrightarrow 3CH_3COOH \qquad (49)$$

$$\begin{array}{l} C_6H_{12}O_6 \longrightarrow 2CH_3COCOOH + 4[H] \\ 2CH_3COCOOH \longrightarrow 2CH_3COOH + 4[H] + 2CO_2 \\ 2CO_2 + 8[H] \longrightarrow CH_3COOH \\ \hline C_6H_{12}O_6 \longrightarrow 3CH_3COOH \end{array} \qquad (50)$$

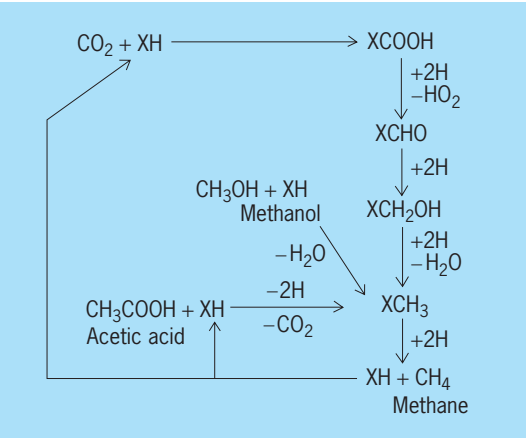

Fig. 10. A possible mechanism for methane fermentation. XH = intracellular reduced carrier or coenzyme.

of purines in which the oxidation of the substrate is likewise coupled with the reduction of CO_2 to acetic acid.

2. *CO_2 reduction to methane.* The anaerobic decomposition of organic matter by complex microbial populations invariably results in the production of methane from CO_2 by the activities of the so-called methane bacteria. One stage in the disposal of sewage is marked by an intense methane fermentation. The lower primary and secondary alcohols, as well as organic acids, are the most common substrates whose oxidation is linked to the reduction of CO_2 to methane. Hydrogen and carbon monoxide also undergo a methane fermentation. In some cases it has been found that carbon monoxide, formic acid, formaldehyde, and methanol are not intermediates in the bacterial reduction of CO_2 to methane. The bacteria that carry out this process are strict anaerobes and are dependent on CO_2. Four genera are known: *Methanobacillus, Methanobacterium, Methanosarcina,* and *Methanococcus*. In addition to morphological differences, the bacteria display substrate specificities. Some types do produce methane by reduction of methanol, or from the methyl group of acetate, proving that a mechanism independent of CO_2 reduction exists for formation of methane. Experiments with deuterium-labeled acetate have shown that the methane evolved may comprise the intact methyl groups, suggesting a reductive demethylation of acetate. *See* METHANOGENESIS (BACTERIA).

The scheme in **Fig. 10** represents a possible mechanism of the methane fermentation.

Jackson W. Foster; R. E. Kallio

Catabolism: Fermentative Metabolism

Many bacteria are able to decompose organic compounds and to grow in the absence of oxygen gas. Such anaerobic bacteria obtain energy and certain organic compounds needed for growth by a process of fermentation. This consists of an oxidation of a suitable organic compound, using another organic compound as an oxidizing agent in place of molecular oxygen. In most fermentations both the

compounds oxidized and the compounds reduced (used as an oxidizing agent) are derived from a single fermentable substrate. In other fermentations, one substrate is oxidized and another is reduced. Different bacteria ferment different substrates. Many bacteria are able to ferment carbohydrates such as glucose and sucrose, polyalchohols such as mannitol, and salts of organic acids such as pyruvate and lactate. Other compounds, such as cellulose, amino acids, and purines, are fermented by some bacteria.

Bacterial fermentations are rather arbitrarily classified according to either the nature of the substrate or the nature of the characteristic product. These two types of classification are not mutually exclusive and lead to different associations of genera and species catalyzing fermentative reactions. In the following discussion, the primary classification is based upon the nature of the substrate, which is either nonnitrogenous (carbohydrates and related compounds) or nitrogenous (amino acids, purines). The further subdivision of the carbohydrate fermentations is based largely on the nature of the characteristic fermentation product.

Bacterial carbohydrate fermentations. Some of the characteristic products of bacterial carbohydrate fermentations are ethyl alcohol, lactic acid, acetic acid, propionic acid, butyric acid, butyl alcohol, acetone, and butanediol. Because certain microorganisms form predominantly one or another of these compounds, the name of the product has come to be associated with the type of fermentation. In this way such terms as alcoholic fermentation, lactic acid fermentation, and propionic acid fermentation have become established, though often of more historical than scientific value.

Alcoholic fermentation. Alcoholic fermentation is the process by which certain yeasts decompose sugars in the absence of oxygen to form ethanol and carbon dioxide, according to reaction (51). This fer-

$$\underset{\text{Glucose}}{C_6H_{12}O_6} \longrightarrow \underset{\text{Ethanol}}{2CH_3CH_2OH} + \underset{\text{Carbon dioxide}}{2CO_2} \qquad (51)$$

mentation provides the basis for the commercial production of ethanol and alcoholic beverages such as wine and beer.

The conversion of glucose to ethanol and carbon dioxide is a complex enzymatic process involving many chemical steps. The main features of this process are the phosphorylation of the hexose (six-carbon) sugar to form a hexose diphosphate, a cleavage of this compound into two molecules of triose (three-carbon) phosphate, the oxidation of the latter and its further conversion to pyruvate, the decarboxylation of pyruvate to carbon dioxide and acetaldehyde, and the reduction of the latter to ethanol. An essential feature of this process is a coupled conversion of inorganic orthophosphate to a high-energy phosphate group in ATP. Two moles of phosphate are converted in this way for each mole of glucose fermented. The resulting ATP is used to provide energy for the biosynthesis of essential cellular components.

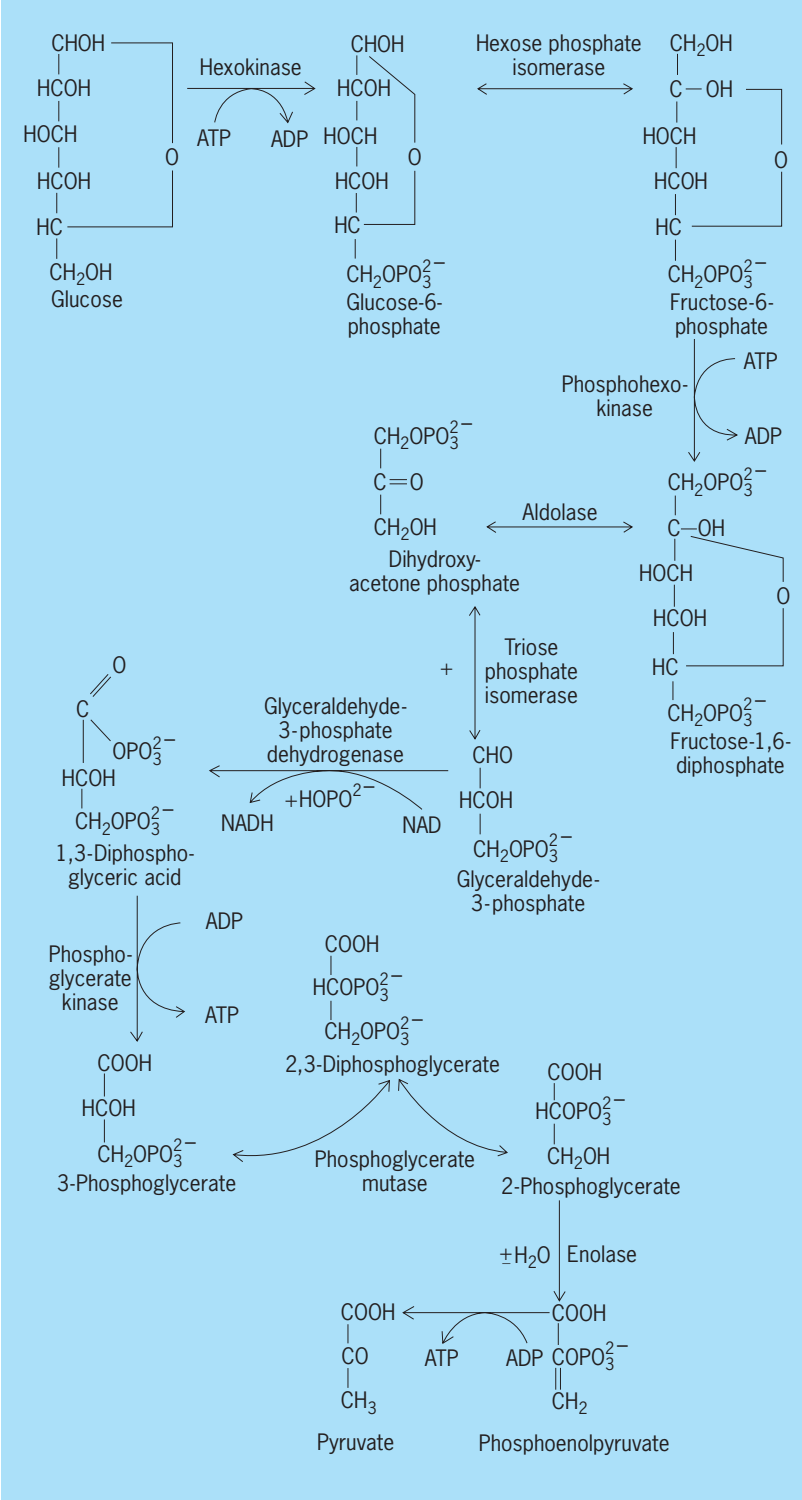

Fig. 11. Embden-Meyerhof pathway of glycolysis and alcoholic fermentation.

Glycolysis is the conversion of glucose to lactic acid. Glycolysis and alcoholic fermentation by yeast involve the same chemical steps from glucose to pyruvate. This sequence of enzymatic reactions, also known as the Embden-Meyerhof pathway, is shown in **Fig. 11**. The first step is a transfer of a phosphate group from ATP to glucose to

form glucose-6-phosphate. This is rearranged to form fructose-6-phosphate, which reacts with ATP to form fructose-1,6-diphosphate. This compound is cleaved by the enzyme aldolase into 1 mole each of dihydroxyacetone phosphate and D-glyceraldehyde-3-phosphate. The enzyme triose phosphate isomerase interconverts the two triose phosphates. After these preparatory reactions, glyceraldehyde-3-phosphate is oxidized and NAD is reduced. This oxidation-reduction reaction, catalyzed by the enzyme glyceraldehyde phosphate dehydrogenase, requires the presence of orthophosphate. During the reaction, orthophosphate disappears and 1,3-diphosphoglyceric acid is formed. The phosphate group attached to carbon atom 1 of this compound is transferred to ADP to form 3-phosphoglyceric acid and ATP. The enzyme phosphoglyceromutase transfers the phosphate group from the 3 to the 2 position of glyceric acid in a reaction involving 2,3-diphosphoglyceric acid. Enolase dehydrates 2-phosphoglyceric acid to phosphoenol pyruvic acid, a compound containing a "high-energy" phosphate group. Finally the phosphate group is transferred to ADP to yield pyruvic acid and ATP. The conversion of 1 mole of glucose to 2 moles of pyruvic acid is accompanied by an uptake of 2 moles of orthophosphate, P_i, and a net formation of 2 moles of ATP, according to reaction (52).

$$C_6H_{12}O_6 + 2NAD + 2ADP + 2P_i \longrightarrow 2CH_3COCOOH + 2ATP + 2NADH \quad (52)$$

In glycolysis by muscle tissue, the reduced diphosphopyridine nucleotide (NADH) is reoxidized back to the oxidized form (NAD) by reaction (53), in

$$\underset{\text{Pyruvate}}{CH_3COCOOH} + NADH + H^+ \longrightarrow \underset{\text{Lactate}}{CH_3CHOHCOOH} + NAD^+ \quad (53)$$

which pyruvic acid is simultaneously reduced to lactic acid.

In alcoholic fermentation by yeast enzymes the acetaldehyde which is formed in the decarboxylation of pyruvic acid, reaction (54*a*), reoxidizes the NADH, reaction (54*b*). The overall reaction is shown in reaction (55).

$$\underset{\text{Pyruvic acid}}{CH_3COCOOH} \longrightarrow \underset{\text{Acetaldehyde}}{CH_3CHO} + CO_2 \quad (54a)$$

$$\underset{\text{Acetaldehyde}}{CH_3CHO} + NADH + H^+ \longrightarrow \underset{\text{Ethanol}}{CH_3CH_2OH} + NAD^+ \quad (54b)$$

$$C_6H_{12}O_6 + 2ADP + 2P_i \longrightarrow 2CH_3CH_2OH + 2CO_2 + 2ATP \quad (55)$$

In living yeast the ATP (present in catalytic amounts) is used in various biosynthetic reactions. In cellfree yeast extract, in which conditions are not favorable for the occurrence of many biosynthetic reactions, ATP is used mainly to phosphorylate excess glucose. Both types of reaction continuously regenerate ADP, which is essential for the oxidation of glyceraldehyde phosphate.

The pentose phosphate pathway of glucose decomposition involves hexose monophosphates and pentose monophosphates but not fructose diphosphate. This pathway is used by some anaerobic bacteria. It is shown in **Fig. 12**.

The first step is the same as in glycolysis: the conversion of glucose to glucose-6-phosphate. The latter compound is oxidized by a dehydrogenase to 6-phosphoglucono-Δ-lactone; the oxidizing agent is NADP. The lactone is hydrolyzed to 6-phosphogluconic acid which is then oxidized by NADP and a specific dehydrogenase to ribulose-5-phosphate and carbon dioxide. A rearrangement at carbon atom 3 under the influence of an epimerase results in the formation of xylulose-5-phosphate. This compound is cleaved between carbon atoms 2 and 3 in an enzymatic reaction that consumes orthophosphate, P_i, and produces glyceraldehyde-3-phosphate and acetyl phosphate. The glyceraldehyde phosphate is converted by a portion of the Embden-Meyerhof pathway to pyruvate. The further

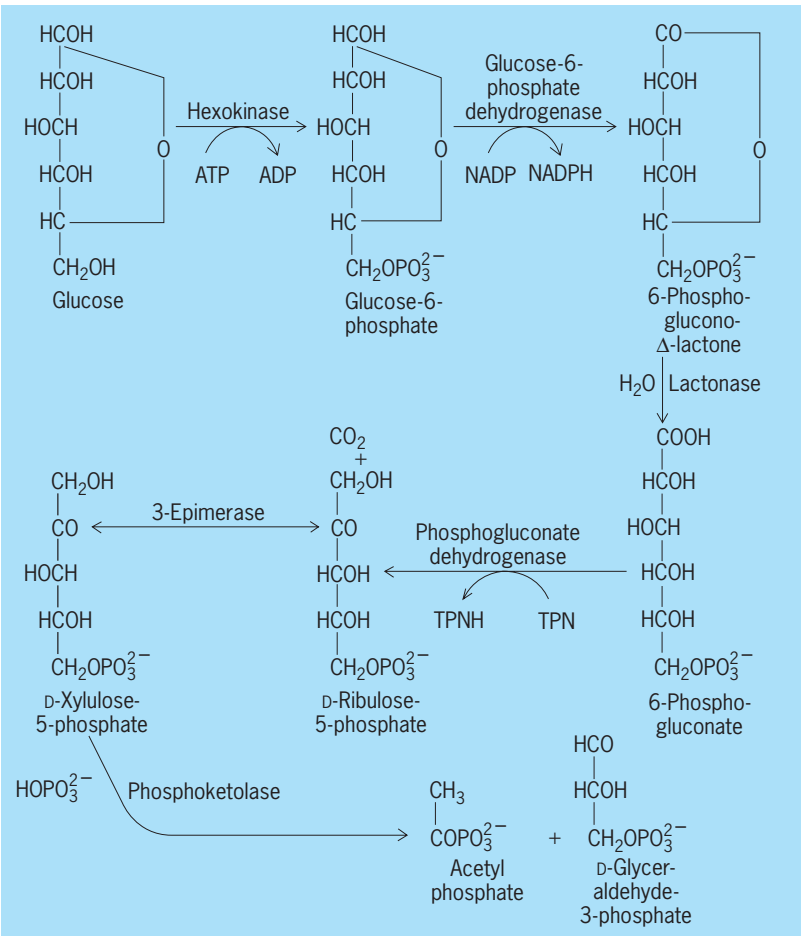

Fig. 12. Pentose phosphate pathway of glucose decomposition in some anaerobic bacteria. Pathway involves hexose and pentose monophosphates.

fate of pyruvate and acetyl phosphate depends upon the enzymatic makeup of the organism using this metabolic path.

The 2-keto-3-deoxy-6-phosphogluconate (Entner-Doudoroff) pathway represents another method of degrading glucose with the liberation of energy. This pathway has been studied mainly in an aerobic organism, *Pseudomonas saccharophila*, but probably also functions in some anaerobic bacteria. The distinguishing feature of the Entner-Doudoroff pathway of glucose metabolism is shown in reaction sequence (56).

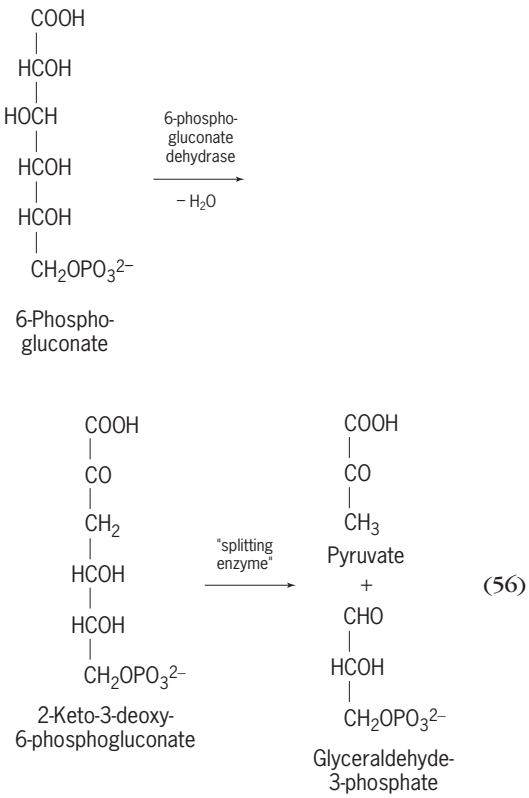

(56)

The initial steps, from glucose to 6-phosphogluconate, are the same as in the pentose phosphate pathway. The 6-phosphogluconate is then dehydrated and rearranged under the influence of a dehydrase to form the characteristic compound of this pathway, 2-keto-3-deoxy-6-phosphogluconate. A "splitting enzyme" cleaves this compound between carbon atoms 3 and 4 to yield pyruvate from glucose carbon atoms 1–3 and glyceraldehyde-3-phosphate from carbon atoms 4–6. The conversion of the triose phosphate to a second mole of pyruvate follows the Embden-Meyerhof pathway. The further reactions of pyruvate vary with the organism.

Alcoholic fermentation by bacteria is relatively uncommon. *Pseudomonas lindneri* (*Zymononas mobilis*), a bacterium found in pulque, ferments glucose or fructose to 45% carbon dioxide, 45% ethanol, and 7% lactic acid. The fermentation appears to follow the Entner-Doudoroff pathway. This is indicated by the fact that the carbon dioxide is derived from glucose carbon atoms 1 and 4, and the primary alcohol group of ethanol is derived from glucose carbon atoms 2 and 5 (**Fig. 13**).

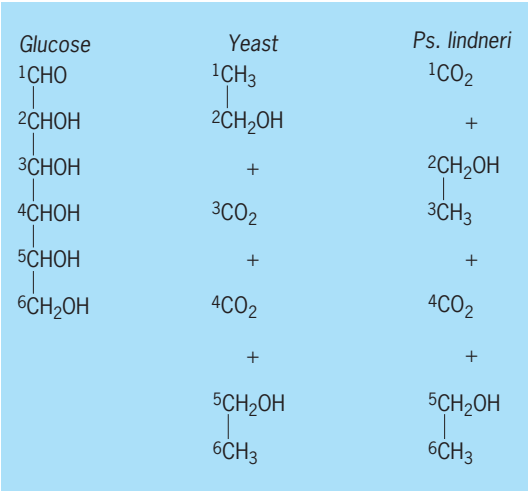

Fig. 13. Comparison of the products of alcoholic fermentation in yeast and *Pseudomonas lindneri*.

Lactic acid fermentation. Lactic acid fermentations are processes by which glucose and some other carbohydrates are converted to lactic acid and sometimes other products. These fermentations are caused by species of *Streptococcus, Leuconostoc, Lactobacillus*, and some other genera. These bacteria are responsible for the souring of milk and cream, for the preservation of silage, and for some types of food spoilage. *See* MILK.

Lactic acid fermentations may be divided into two types, the homolactic and heterolactic fermentations.

Homolactic fermentations, caused mainly by *Streptococcus* and certain *Lactobacillus* species, convert sugar almost entirely to lactic acid, as in reaction (57).

$$\underset{\text{Hexose}}{C_6H_{12}O_6} \longrightarrow \underset{\text{Lactic acid}}{2CH_3CHOHCOOH} \qquad (57)$$

Only when the fermentations occur in an alkaline medium do other products such as formate, acetate, and ethanol accumulate in considerable amounts. *Streptococcus* species usually form L(+)-lactic acid, the isomer metabolized most readily by humans. *Lactobacillus* species may form either L(+)-, D(−)-, or racemic lactic acid. Homolactic fermentations occur by the glycolytic (Embden-Meyerhof) pathway. The pyruvic acid is reduced to lactic acid by reduced nicotinamide adenine dinucleotide (NADH) under the influence of the enzyme lactic dehydrogenase by reaction (58).

$$\underset{\text{Pyruvic acid}}{CH_3COCOOH} + NADH + H^+ \longrightarrow \underset{\text{Lactic acid}}{CH_3CHOHCOOH} + NAD^+ \qquad (58)$$

Heterolactic fermentations, caused mainly by *Leuconostoc* and certain *Lactobacillus* species, produce lactic acid in smaller yield and convert a considerable part of the sugar to carbon dioxide and other products, such as ethanol, acetic acid, glycerol, and mannitol. *Leuconostoc mesenteroides* ferments glucose according to reaction (59). The glucolytic pathway

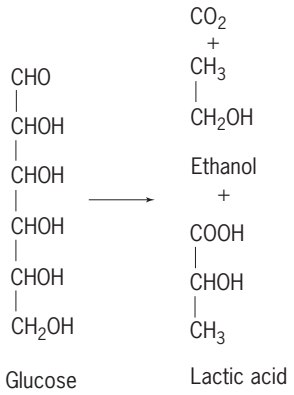

(59)

of sugar decomposition cannot account for the formation of equimolar amounts of the three products. The origin of the carbon atoms of the products, established by experiments with ^{14}C-labeled glucose, is consistent with the occurrence of the pentose phosphate pathway. The acetyl phosphate produced in this sequence is reduced first to acetaldehyde and then to ethanol by use of reducing agents (NADPH and NADH) produced during the oxidation of glucose to acetyl phosphate, pyruvate, and carbon dioxide. Apparently all typical heterolactic fermentations involve the pentose phosphate pathway. Heterolactic fermentations of fructose, but not glucose, characteristically produce mannitol in good yield.

Several homolactic and heterolactic species are able to ferment the pentose sugars, xylose, arabinose, and ribose. The products of pentose fermentation by both types of lactic acid, bacteria, reaction (60), are lactic and acetic acids.

$$\underset{\text{Pentose}}{C_5H_{10}O_5} \longrightarrow \underset{\text{Lactic acid}}{CH_3CHOHCOOH} + \underset{\text{Acetic acid}}{CH_3COOH} \qquad (60)$$

Propionic acid fermentation. The propionic acid fermentation converts sugars, polyalcohols, salts of organic acids such as lactate and pyruvate, and some amino acids such as alanine and serine to propionic acid and often other products, including acetate, succinate, propanol, and carbon dioxide. Most of the bacteria that cause the propionic acid fermentation are species of *Propionibacterium*. A few other bacteria, including *Clostridium propionicum* and *Micrococcus* (*Veillonella*) *lactilyticus*, cause a similar fermentation.

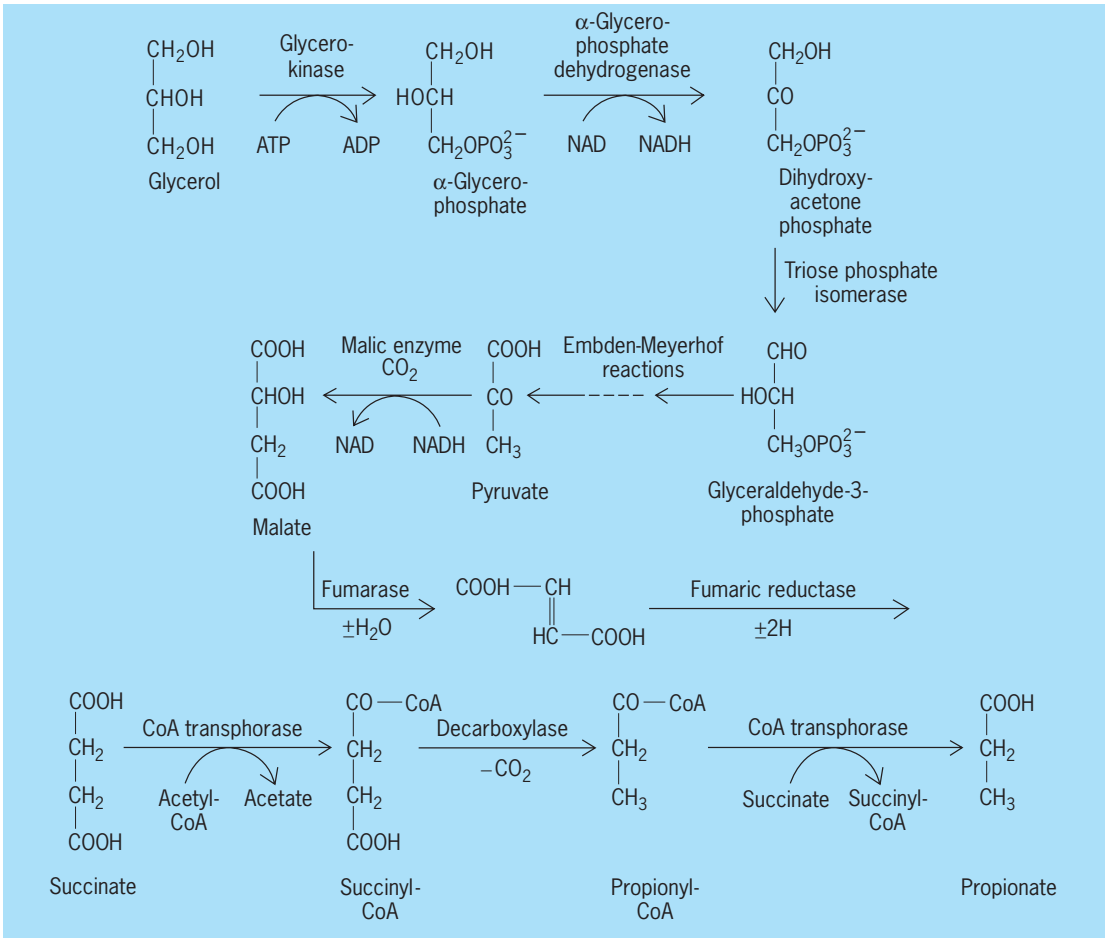

Fig. 14. Probable sequence of reactions for propionic acid fermentation from glycerol.

The simplest propionic acid fermentation is that of glycerol; 90% of the substrate may be converted to propionic acid, as shown in reaction (61). When

$$\underset{\text{Glycerol}}{CH_2OHCHOHCH_2OH} \longrightarrow \underset{\text{Propionic acid}}{CH_3CH_2COOH} + H_2O \qquad (61)$$

the medium contains a considerable concentration of carbon dioxide, succinic acid is also formed and the yield of propionic acid is reduced. For each mole of succinate accumulating, 1 mole of carbon dioxide is consumed. Experiments with ^{14}C-labeled carbon dioxide show that carbon dioxide carbon is incorporated into the carboxyl groups of both propionic and succinic acids. There is considerable evidence that succinic acid and succinyl coenzyme A are intermediates in the formation of propionic acid. Propionic acid is probably formed from glycerol by the sequence shown in **Fig. 14**.

The acetyl-CoA required for the formation of succinyl-CoA is probably formed by the reaction shown in reaction (62). Lactate, a favorite substrate

$$\underset{\text{Acetic acid}}{CH_3COOH} + ATP + CoA \longrightarrow \underset{\text{Acetyl-CoA}}{CH_3CO - CoA} + AMP + PP \qquad (62)$$

of propionic acid bacteria, is fermented approximately according to reaction (63). A little succinic

$$\underset{\text{Lactic acid}}{3CH_3CHOHCOOH} \longrightarrow \underset{\text{Propionic acid}}{2CH_3CH_2COOH} + \underset{\text{Acetic acid}}{CH_3COOH} + CO_2 + H_2O \qquad (63)$$

acid is also formed. Hexose and pentose sugars are converted to the same products.

The propionic acid fermentation is essential for the preparation of Swiss cheese. The characteristic flavor is partially attributable to propionic acid and the holes are caused by carbon dioxide formed by the propionic acid bacteria. *See* CHEESE.

Butyric acid fermentation. The butyric acid fermentation results in the formation of butyric acid and usually other products which may include hydrogen, carbon dioxide, acetic acid, and caproic acid. This type of fermentation is usually caused by anaerobic spore-forming bacteria such as *Clostridium butyricum* and *C. kluyveri*. A few nonsporulating bacteria, such as *Butyribacterium rettgeri* and *Sarcina maxima*, also produce butyric acid. Suitable substrates for the butyric acid fermentation are hexose and pentose sugars, polyalcohols, polysaccharides, salts of organic acids (tartrate, lactate), and occasionally ethanol plus a short-chain fatty acid.

The simplest butyric acid fermentation is caused by *C. kluyveri*, which converts ethanol and acetate to caproic acid. The fermentation involves two stages; in the first stage butyric acid is formed, reaction (64*a*), and in the second stage it is converted to caproic acid, reaction (64*b*). These reactions

$$\underset{\text{Ethanol}}{CH_3CH_2OH} + \underset{\text{Acetic acid}}{CH_3COOH} \longrightarrow \underset{\text{Butyric acid}}{CH_3CH_2CH_2COOH} + H_2O \qquad (64a)$$

$$\underset{\text{Ethanol}}{CH_3CH_2OH} + \underset{\text{Butyric acid}}{CH_3CH_2CH_2COOH} \longrightarrow \underset{\text{Caproic acid}}{CH_3CH_2CH_2CH_2CH_2COOH} + H_2O \qquad (64b)$$

provide the energy needed by the bacteria.

The conversion of ethanol and acetate to butyrate, like all fermentations, is a relatively complex process. The individual enzymatic steps are shown in the reaction sequence in **Fig. 15**.

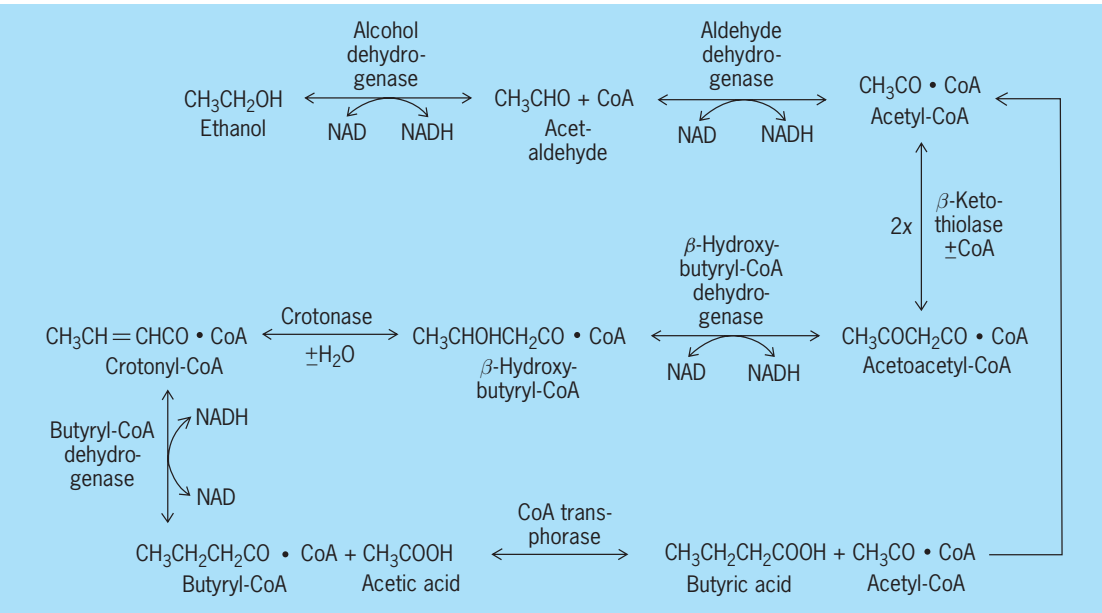

Fig. 15. Enzymatic steps in the conversion of ethanol and acetate to butyrate.

Caproic acid is formed by a similar sequence of reactions, in which a molecule of acetyl-CoA is replaced by butyryl-CoA in the β-ketothiolase reaction. Elongation of the carbon chain occurs by addition of two carbon moieties onto the carboxyl carbon of butyric acid.

Similar reactions occur in the butyric acid fermentation of glucose. The sugar appears to be converted to pyruvate by the Embden-Meyerhof pathway. Pyruvate is then oxidized to acetyl-CoA with the simultaneous formation of carbon dioxide and hydrogen according to reaction (65). As a consequence of this

$$\underset{\text{Pyruvic acid}}{CH_3COCOOH} + CoA \longrightarrow \underset{\text{Acetyl-CoA}}{CH_3CO - CoA} + CO_2 + H_2 \qquad (65)$$

reaction, 2 moles of carbon dioxide and 2 moles of hydrogen are formed per mole of glucose fermented. The acetyl-CoA is converted to butyric acid with the aid of NADH produced during the oxidation of triose phosphate.

Butanol-acetone fermentation. The butanol-acetone fermentation yields considerable amounts of the solvents butanol and acetone and smaller amounts of acetic and butyric acids along with the gases hydrogen and carbon dioxide. This fermentation, catalyzed by *C. acetobutylicum*, has been extensively used for the commercial production of butanol and acetone from starch and sugars.

The butanol-acetone fermentation results from a small modification of the chemistry of the butyric acid fermentation of carbohydrates. Initially the same products, mainly butyric and acetic acids, carbon dioxide, and hydrogen, are formed in both fermentations. Then, if the medium is permitted to become acid, a leak develops in the sequence of reactions for converting acetyl-CoA to butyrate. This leak develops as a result of the formation of an enzyme system that converts acetoacetyl-CoA to acetoacetate and then decarboxylates the latter to acetone, as shown in reaction (66). The destruction of acetoacetyl-

$$\underset{\text{Acetoacetyl-CoA}}{CH_3COCH_2CO - CoA} \longrightarrow CoA + \underset{\text{Acetoacetate}}{CH_3COCH_2COOH} \xrightarrow{\text{acetoacetate decarboxylase}} \underset{\text{Acetone}}{CH_3COCH_3} + CO_2 \qquad (66)$$

CoA both causes the accumulation of acetone and prevents the further formation of butyrate. The reducing agents, initially used to reduce acetoacetyl-CoA to butyryl-CoA, are then used for the reduction of butyryl-CoA via butyraldehyde to *n*-butanol, asshown in reaction (67).

$$\underset{\text{Butyryl-CoA}}{CH_3CH_2CH_2CO \cdot CoA} \xrightarrow{NADH \rightarrow NAD} CoA + \underset{\text{Butyraldehyde}}{CH_3CH_2CH_2CHO} \xrightarrow{NADH \rightarrow NAD} \underset{\text{Butanol}}{CH_3CH_2CH_2CH_2OH} \qquad (67)$$

Consequently, acetone and butanol are formed simultaneously. During the solvent-producing stage of the fermentation, a large increase in the formation of carbon dioxide relative to hydrogen results from acetoacetic acid decarboxylation.

Clostridium butylicum, a species closely related to *C. acetobutylicum*, forms an additional product, isopropanol. This branched-chain secondary alcohol is formed by reduction of acetone, reaction (68).

$$\underset{\text{Acetone}}{CH_3COCH_3} + 2H \longrightarrow \underset{\text{Isopropanol}}{CH_3CHOHCH_3} \qquad (68)$$

Ethanol-acetone fermentation. The ethanol-acetone fermentation of carbohydrates is catalyzed by *Aerobacillus macerans*. This species forms ethanol, acetone, acetic and formic acids, carbon dioxide, and hydrogen from glucose or starch. The chemistry of the fermentation has not been studied extensively but is probably similar to that of the butanol-acetone fermentation, except for the absence of reactions from acetoacetyl-CoA to butyrate. The acetone is formed by decarboxylation of acetoacetic acid; and the ethanol, a major product, originates by reduction of acetyl-CoA via acetaldehyde, a reaction analogous to the reduction of butyryl-CoA to butanol by *C. acetobutylicum*.

The 2,3-butanediol fermentation of carbohydrates is caused by *Aerobacter* and *Serratia* species, *Aerobacillus polymyxa*, and some strains of *Bacillus subtilis*. This type of fermentation produces large yields of carbon dioxide, ethanol, and 2,3-butanediol; moderate yields of formate and hydrogen; and small yields of acetate, lactate, and succinate. In *Aerobacter* species the coversion of glucose to pyruvate appears to follow the Embden-Meyerhof pathway. Some of the pyruvate undergoes a cleavage reaction involving coenzyme A with formation of acetyl-CoA and formate, reaction (69). The acetyl-

$$\underset{\text{Pyruvate}}{CH_3COCOOH} + CoA \longrightarrow \underset{\text{Acetyl-CoA}}{CH_3CO - CoA} + \underset{\text{Formate}}{HCOOH} \qquad (69)$$

CoA is partly converted to acetate and partly reduced via acetaldehyde to ethanol. The formate is partly broken down by the formic hydrogenylase enzyme system to carbon dioxide and hydrogen, as shown in reaction (70). The remainder of the pyru-

$$\underset{\text{Formate}}{HCOOH} \longrightarrow CO_2 + H_2 \qquad (70)$$

vate is mostly converted to acetolactate and carbon

dioxide through reaction (71). The acetolactate is

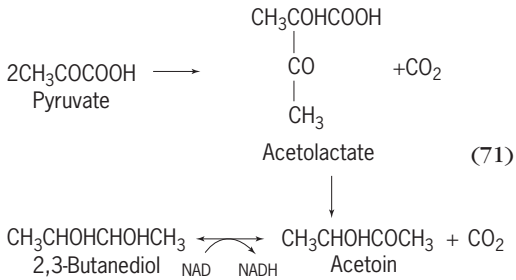

rapidly decarboxylated by a specific enzyme to acetoin. The latter compound is an effective oxidizing agent under anaerobic conditions; it is readily reduced under the influence of a specific dehydrogenase to 2,3-butanediol. Under aerobic conditions, some bacteria can slowly oxidize acetoin to diacetyl, reaction (72), a slightly volatile compound responsible for the flavor of butter.

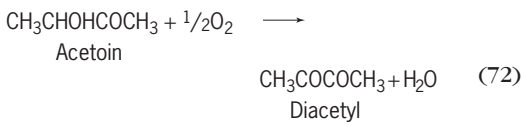

Fermentations of nitrogen compounds. Under anaerobic conditions, bacteria are able to decompose many organic nitrogenous compounds, including amino acids and various heterocyclic compounds such as the purines. Fermentations of such compounds are described as putrefactive because they frequently result in the formation of substances having unpleasant odors. The bacteria that ferment nitrogenous compounds are mostly species of *Clostridium* (*C. botulinum, C. perfringens, C. sporogenes, C. tetani*, and *C. tetanomorphum*), but several species belonging to the genera *Fusobacterium, Diplococcus*, and *Micrococcus* also cause such fermentation.

Amino acid fermentation. Amino acid fermentations involve either the anerobic decomposition of a single substrate or the coupled decomposition of two or more substrates. The latter type of decomposition is known as the Stickland reaction.

Most of the common amino acids can be fermented singly by certain anaerobic bacteria. Alanine, cysteine, glutamate, glycine, histidine, serine, and threonine are some of the amino acids known to be fermented. The products usually include one or more of the short-chain fatty acids (acetic, propionic, butyric), carbon dioxide, hydrogen, and ammonia.

The decomposition of alanine by *C. propionicum* is one of the simpler amino acid fermentations. The process, described by reaction (73), is similar to the

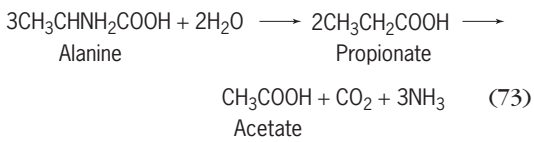

propionic acid fermentation of lactate by the same organism except for the additional formation of ammonia. The chemical steps in the fermentation of alanine have not been completely elucidated, but they include an oxidation of 1 mole of alanine to acetate, carbon dioxide, and ammonia, via pyruvate, and a reduction of 2 moles of alanine to propionate and ammonia, probably via the coenzyme A derivative of acrylate.

The fermentation of glutamate by *C. tetanomorphum* results in the formation of ammonia, acetate, butyrate, carbon dioxide, and hydrogen. This modified butyric acid fermentation involves the transient formation of several branched-chain dicarboxylic acids shown in reaction (74). These reactions result

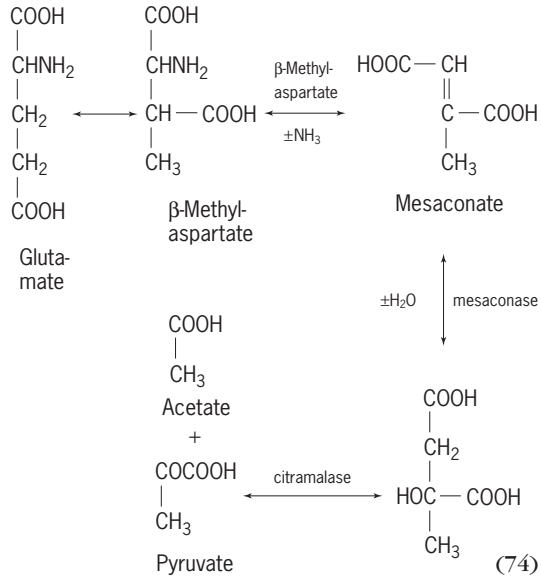

in the conversion of glutamate to acetate and pyruvate. The latter compound undergoes oxidative decarboxylation to form an acetyl compound which is converted partly to acetate and partly to butyrate. Energy for synthetic reactions is produced during the fermentation of pyruvate as in other butyric acid fermentations.

The aerobic decomposition of the sulfur-containing amino acid cysteine by *Proteus vulgaris* and other bacteria results in the formation of hydrogen sulfide (H_2S), ammonia, and pyruvate, as shown in reaction (75). By an analogous reaction

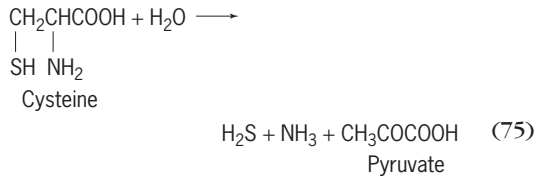

another sulfur-containing amino acid, methionine, gives methylmercaptan (CH_3SH), ammonia, and α-ketobutyrate as shown in reaction (76). Both

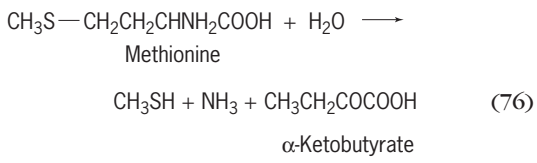

hydrogen sulfide and methylmercaptan contribute

to the putrefactive odor of decomposing proteins. The two keto acids are further fermented to fatty acids and carbon dioxide.

A decarboxylation of certain amino acids, according to reaction (77), is catalyzed by a number of

$$RCHNH_2COOH \longrightarrow RCH_2NH_2 + CO_2 \qquad (77)$$

bacteria. The amino acids most commonly decarboxylated are histidine, tyrosine, arginine, ornithine, lysine, tryptophan, and glutamate. The decarboxylation of lysine, for example, produces cadaverine, reaction (78), a basic substance having a very un-

$$\underset{\text{Lysine}}{NH_2(CH_2)_4CHNH_2COOH} \longrightarrow \underset{\text{Cadaverine}}{NH_2(CH_2)_4CH_2NH_2} + CO_2 \qquad (78)$$

pleasant odor. The decarboxylation reaction occurs mainly in bacteria exposed to a relatively acidic medium. Because the products of decarboxylation are more basic than the amino acids, the reaction helps to maintain a neutral environment.

The Stickland reaction, the fermentation of pairs of different amino acids, is caused by many *Clostridium* species such as *C. sporogenes* and *C. botulinum*. One amino acid in the pair is oxidized whereas the other is reduced. The simplest example is the coupled decomposition of alanine and glycine according to reaction (79). In this process the oxidation of

$$\underset{\text{Alanine}}{CH_3CHNH_2COOH} + \underset{\text{Glycine}}{2CH_2NH_2COOH} + 2H_2O \longrightarrow \underset{\text{Acetate}}{3NH_3 + 3CH_3COOH} + CO_2 \qquad (79)$$

1 mole of alanine to acetate, carbon dioxide, and ammonia is coupled with the reduction of 2 moles of glycine to acetate and ammonia. Other amino acids that are oxidized include leucine, isoleucine, valine, and histidine. The products are always ammonia, carbon dioxide, and a carboxylic acid with one less carbon atom. Other amino acids that are reduced include proline, hydroxyproline, ornithine, and tryptophan. The reduction of proline (or hydroxyproline) converts a ring compound to a straight-chain compound without formation of ammonia, for example, reaction (80). With ornithine or tryptophan,

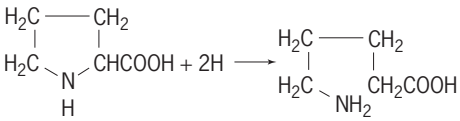

(80)

as with glycine, the α-amino group is removed by a reductive deamination.

Purine fermentation. Purine fermentations are caused by a few *Clostridium* and *Micrococcus* species. *Clostridium acidi-urici*, for example, rapidly ferments uric acid, xanthine, and certain other purines to ammonia, carbon dioxide, acetate, and very small amounts of glycine and formate. The fermentation of xanthine is approximately described by reaction (81). Other purines are converted to xanthine be-

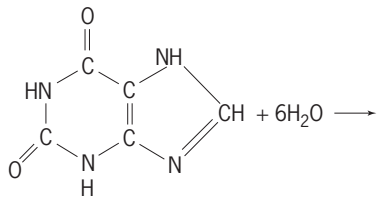

$$4NH_3 + 3CO_2 + CH_3COOH \qquad (81)$$

fore the ring system is ruptured. Chemical steps in xanthine fermentation are shown in **Fig. 16**.

The fermentation involves destruction of the six-membered ring, followed by conversion of the imidazolone ring to the straight-chain compound form-iminoglycine. The latter is converted to glycine, ammonia, and activated formic acid, actually a derivative of tetrahydrofolic acid. The activated formic acid is reduced to the formaldehyde level of oxidation and condensed with glycine to form serine. This amino acid is deaminated and rearranged to form pyruvate, which is oxidized to acetate and carbon dioxide. High-energy phosphate is probably formed during the oxidation of pyruvate and stored as ATP for use in biosynthetic reactions.

Horace A. Barker

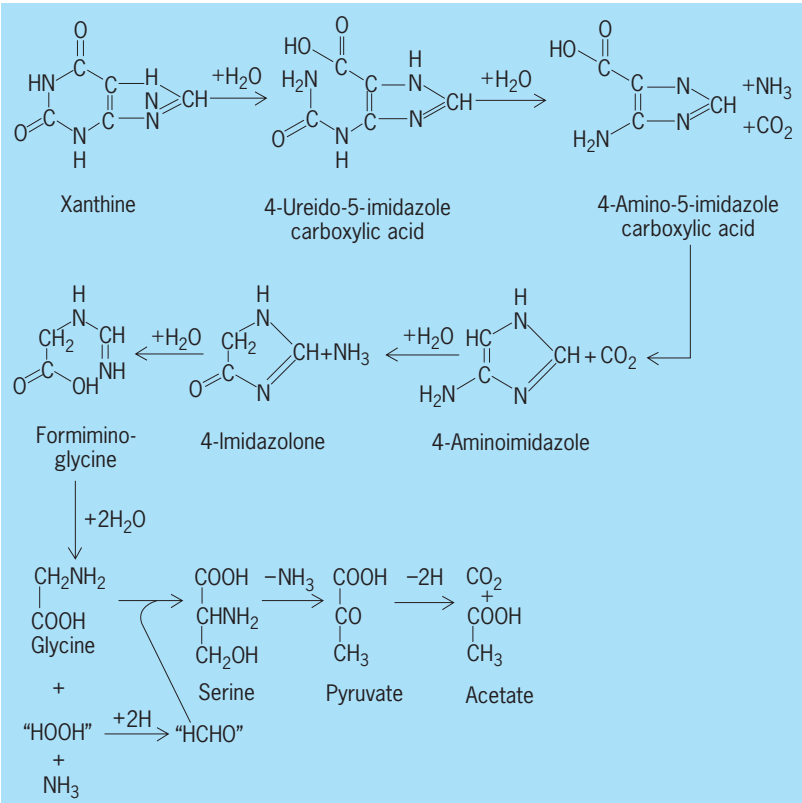

Fig. 16. Steps in the fermentation of the purine xanthine by *Clostridium acidi-urici*.

Bacterial Anabolism

Bacterial anabolism comprises the physiological and biochemical activities concerned with the acquisition, synthesis, and organization of the numerous and varied chemical constituents of a bacterial cell. Clearly, when a cell grows and divides to form two cells, there exists twice the amount of cellular components that existed previously. These components are drawn, directly or indirectly, from the environment around the cell, and (usually) modified extensively in the growth processes when new cell material is formed (biosynthesis). This build-up, or synthesis, begins with a relatively small number of low-molecular-weight building blocks which are either assimilated directly from the environment or produced by catabolism. By sequential and interrelated reactions, they are fashioned into different molecules (mostly of high molecular weight, and hence called macromolecules), for example, lipids, polysaccharides, proteins, and nucleic acids, and many of these molecules are in turn arranged into more complex arrays such as ribosomes, membranes, cell walls, and flagella. Other typical anabolic products, of lower molecular weight, include the pigments, vitamins, antibiotics, and coenzymes. The enzymes responsible for the sequential reactions in any one biosynthetic pathway or assembly sequence are often located on or in cellular structures and thus in physical proximity to the preceding and succeeding enzymes, and their products, and to the site(s) where cellular structures are to be formed. Anabolism also includes the transport of molecules into cells, of building blocks to reaction sites, energetic activations, and the transfer and incorporation of the finished products to their ultimate sites in or outside the cell.

Energy relations in metabolism. During the metabolism of a substrate, bacteria perform numerous biochemical transformations whose overall effect is the liberation of energy in the form of heat. This can be illustrated by the spectacular rise in temperature accompanying the microbial decomposition of large masses of plant material(s) which are sufficiently well insulated to prevent a rapid release of heat into the environment, such as in a compost pile. The temperature in such a mass may rise to 80°C (176°F) or more in a matter of hours. Similarly, in the large culture vats in which microbes are grown, for example, for the commercial production of antibiotics, heat production is so intense that the contents must be cooled continuously to keep the temperature within the physiologically tolerable range.

Prior to 1930, it was customary to attribute the energy liberation exclusively to those metabolic processes collectively called catabolism. The anabolic reactions were grouped together in a separate category, because they could all be considered as processes involving the uptake of energy.

Reactions that release free energy and thus can proceed spontaneously are called exergonic. Here the sum of the chemical (free) energy of the reactants is greater than that of the products. Reactions in which the total free energy of the products exceeds that of the reactants are called endergonic. They can occur only if enough energy is supplied to compensate for the difference. If a proper mechanism exists whereby the energy released in an exergonic reaction can be transferred to an endergonic one, the former can drive the latter, and in this case the two processes are said to be energetically coupled.

The sharp distinction between catabolic and anabolic processes in bacterial metabolism could be justified on this basis. For example, the degradation of a fatty acid to the corresponding number of acetic acid molecules is an energy-yielding process. Its opposite, the formation of fatty acids from acetate, which is quite common in bacterial metabolism, must therefore require energy. This kind of argument could also be applied to practically every aspect of anabolism. Anabolism is an immensely complex phenomenon. This is clear from the fact that many bacteria can live and multiply in a simple inorganic medium supplied with but a single organic substrate, such as acetate. This implies that these organisms can synthesize all of their cell constituents—carbohydrates, fats, proteins, enzymes, nucleic acids, and so on—from the substrate. So bewildering are the number and diversity of the components of a bacterial cell that it seemed hopeless to attempt an explanation of the manner in which the necessary syntheses are accomplished. The best that could be accomplished was to recognize the principle of energetic coupling, which could be expressed in the form of two general reactions (82*a*) and (82*b*).

$$\text{Substrate} \xrightarrow{\text{catabolism}} \text{breakdown products} + \text{energy} \qquad (82a)$$

$$\text{Substrate} + \text{energy} \xrightarrow{\text{anabolism}} \text{products of synthesis} \qquad (82b)$$

These equations show that catabolic reactions may occur without concomitant anabolic ones, while the reverse is obviously impossible; anabolism is dependent on catabolism. After 1930, however, when a sound beginning had been made with a more detailed analysis of catabolic processes, it gradually became clear that anabolic reactions might be amenable to this same approach, and after 1950 enormous strides were made in understanding the mechanism by which they proceed. It is now universally accepted that anabolic processes can also be interpreted as being composed of series of simple, consecutive-step reactions, and that each of these can be represented by reaction (83), which is fundamentally similar to the one given earlier as a general expression of catabolic step reactions, reaction (3).

$$H_2A + B \longrightarrow A + H_2B \qquad (83)$$

This significant change in outlook, which led to a progressively refined understanding of the mechanism of biosynthetic reactions, began with the recognition that certain step reactions in such a process

need not be regarded as endergonic. The seeming contradiction of this statement with the argument presented earlier can be readily explained. It was learned that the actual reactants in a biosynthetic reaction, the building blocks, are not the simple substances like the substrate molecules previously considered, but rather are derivatives of the latter, in which they occur combined with a reactive moiety. As a result of this combination, they become more reactive; this can also be expressed by saying that the substrate molecules have thus been lifted to a higher energy level. Consequently they can then participate in spontaneous reactions for which the unchanged substrate molecules are unfit.

Such reactive derivatives can be formed in either of two ways.

1. They may originate directly as intermediate products during a breakdown process. This happens if a substrate remains tied to the enzyme or coenzyme that mediates its decomposition. In that event the catabolic product first emerges as a coenzyme-bound substance.

2. They may also be formed by a reaction in which a general energy-storage product, such as ATP, is used to transfer an energy-rich (reactive) fragment to a substrate molecule and so to raise the reaction potentials of the latter.

The example of the degradation and synthesis of fatty acids can again serve to illustrate these two aspects. In the section on fermentative metabolism, it has been mentioned that the breakdown proceeds under the influence of enzymes in association with coenzyme A. The fatty acid first becomes attached to the latter as an acyl-CoA compound from which, after dehydrogenation, two-carbon fragments are split off as acetyl-CoA. This substance is far more energy-rich, and hence reactive, than acetate itself, and two molecules can readily be transformed without additional energy supply, to acetoacetyl-CoA. On the other hand, acetoacetic acid cannot be so formed from acetate; this is an endergonic process. Thus reaction (84) illustrates the direct catabolic formation

$$\text{Fatty acid} - \text{CoA complex} \xrightarrow[\text{genation}]{\text{dehydro-}} \text{acetyl-CoA} \qquad (84)$$

of a reactive entity that can function as a reactant in a synthesis that does not require an input of energy.

The second type of mechanism is illustrated by the spontaneous, enzyme-mediated reaction between acetate and ATP which yields acetyl phosphate and ADP. Acetyl phosphate, as a mixed anhydride comparable to acetyl chloride, is an unstable substance on the same energy level as acetyl-CoA, and its acetyl group can be enzymatically transferred, without extra energy supply, to CoA, yielding the genuine building block for the synthesis of the long-chain fatty acids by way of reactions (85*a*) and (85*b*). The

$$\text{Acetate} + \text{ATP} \longrightarrow \text{acetyl phosphate} + \text{ADP} \qquad (85a)$$

$$\text{Acetyl phosphate} + \text{CoA} \longrightarrow \text{acetyl-CoA} + \text{phosphate} \qquad (85b)$$

formation of acetyl phosphate from acetate represents an activation of the latter; it is one instance of many similar processes by which a relatively stable, that is, unreactive, substance can be converted into an unstable, reactive one. Because of this activation, the substance can then participate in a spontaneously occurring synthetic process; the unactivated one cannot.

This is the fundamental principle of the mechanism by which the energy-liberating catabolic processes can be coupled to the energy-requiring anabolic reactions.

Role of phosphorylated compounds. The above paragraphs have shown that ATP, by transfer of one of its terminal phosphate groups to another substance, can be used to raise the energy level of the latter. Such reactions are quite common in the metabolism of bacteria, as well as in that of other microbes, plants, and animals. In fact, most phosphorylations are accomplished in this manner, and in earlier parts of this article many examples have been given. This kind of reaction is possible because ATP itself is an unstable substance which contains two energy-rich (commonly called high-energy) phosphate bonds. Upon hydrolytic cleavage, each of these can release an amount of energy of the order of magnitude of 8000–10,000 cal/mole. If such a phosphate group is transferred to a metabolically unreactive molecule, the latter acquires some or all of the bond energy, which is thus not lost as heat. This fundamental notion of the preservation of bond energy was first formulated by F. Lipmann in 1939.

However, not all phosphorylations at the expense of ATP are equally effective in preserving bond energy. In addition to the energy-rich compounds, of which acetyl phosphate is an example, there are other types of phosphorylated substances in which the energy of the phosphate bond is only of the order of magnitude of 3000 cal/mole, as shown by the amount of heat liberated during their hydrolysis. These are the phosphate esters, such as glycerophosphate, glucose-6-phosphate, and fructose-1,6-diphosphate, all of which can be formed by transfer of a phosphate group from ATP according to reaction (86). The transfer of phosphate groups from

$$\text{R} - \text{CH}_2\text{OH} + \text{ATP} \xrightarrow[\text{kinase}]{\text{specific}} \text{R} - \text{CH}_2\text{O} - \text{PO}_3\text{H}_2 + \text{ADP} \qquad (86)$$

ATP to substrate molecules is mediated by enzymes that are called kinases. They are specific for a particular substrate, and are correspondingly named glucokinase (or hexokinase), 6-phosphofructokinase, and so on.

Because the phosphate-bond energy in the esters is far less than that in ATP, the esterification reactions involve the loss of a large proportion of the chemical energy as heat. Nevertheless, the esterifications are metabolically significant, for they also yield compounds with potentialities greater than those possessed by nonphosphorylated substances. These phosphate esters are used in polysaccharide biosynthesis.

In the activation of substrate molecules by transphosphorylation from ATP the latter is consumed, and it will be evident that in a metabolizing cell the stock of this material must be continuously replenished. This is accomplished by the catabolic formation of high-energy phosphate bonds, including ATP itself. It may occur in a number of different ways as discussed earlier in the section on bacterial catabolism.

Adenosine phosphates. Three adenosine phosphates, which are derivatives of the glucoside adenine riboside or adenosine, are known. They are adenosine monophosphate (AMP), also referred to as adenylic acid, ADP, and ATP. The structures and relations of these substances are shown in **Fig. 17**. It will be seen that the phosphate group of AMP is attached to adenosine by an ester (low-energy) linkage, while the additional phosphate groups of ADP and ATP represent anhydride (high-energy) bonds. Consequently ADP and ATP, but not AMP, can act as high-energy phosphate donors, and AMP and ADP are both high-energy phosphate acceptors.

Although ADP and, particularly, ATP are undoubtedly the most common high-energy phosphate donors in bacterial metabolism, there are others, such as the previously mentioned acetyl phosphate, phosphoenolpyruvic acid, and 1,3-diphosphoglyceric acid. The last two are known only as intermediate metabolic products, while ADP and ATP generally are, and acetyl phosphate occasionally may be, high-energy phosphate storage products as well. This also applies to guanosine and other di- and triphosphates (GDP, GTP) which play a role in the Krebs or tricarboxylic acid cycle.

Apart from being a high-energy phosphate donor, ATP can also participate in anabolic processes in another way. It can transfer the adenyl (AMP) moiety to a substrate molecule, with the formation of an adenylate and pyrophosphate (PP). The andenylates, too, represent activated substrates and are, as will be seen later, the genuine reactants in certain biosynthetic processes. The significance of such adenylations has been sought in the concomitant formation of pyrophosphate by the enzyme pyrophosphatase. In a situation where the equilibrium of the adenylase reaction is highly in favor of the starting materials (indicated by $\rightleftharpoons$) and thus where adenylate would be formed in only extremely low concentration, the removal of the pyrophosphate in a subsequent reaction would cause a shift in the equilibrium, with a larger concentration of adenylate being formed, as shown in reactions (87*a*) and (87*b*).

$$\text{Substrate} + \text{ATP} \xrightleftharpoons{\text{adenyltransferase}} \text{substrate-adenylate} + \text{PP} \qquad (87a)$$

$$\text{PP} + H_2O \xrightarrow{\text{pyrophosphatase}} \text{orthophosphate} \qquad (87b)$$

Energy-rich coenzyme-substrate. The adenylation reactions, mediated by enzymes known as adenylyltransferases, are a special instance of a more general type of substrate activation which yields an energy-rich substrate-enzyme complex. Such complexes may be formed in one of two ways.

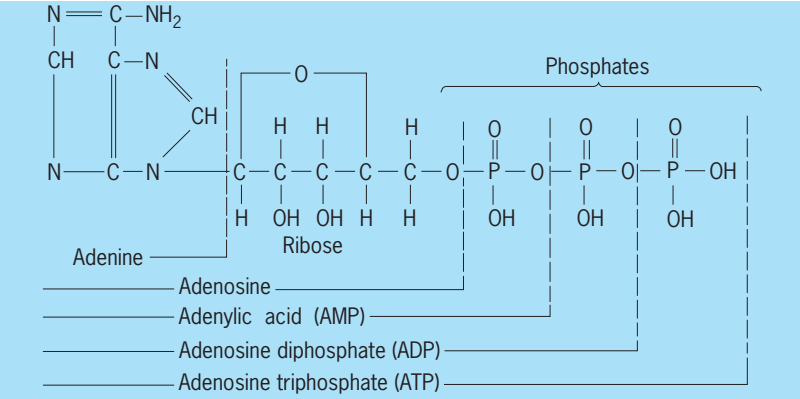

Fig. 17. Relationship of adenosine and its three adenosine phosphate derivatives.

1. They may arise directly from the metabolic conversion of a substrate-coenzyme compound that does not initially contain an energy-rich bond. This matter of origin is illustrated by the formation of the coenzyme complex of 3-phosphoglyceric acid, with its high-energy bond, with the addition product of 3-phosphoglyceraldehyde and the sulfhydryl-containing coenzyme by dehydrogenation.

2. Alternatively, they may result from the transfer of a moiety from an already activated to a nonactivated substance. Examples of this mode of formation are the above-mentioned transadenylations, and the transacetylation by which the acetyl group of acetyl phosphate is transferred to coenzyme A with the formation of acetyl-CoA and orthophosphate. The former represent cases where the energy-rich bond initially resides in the coenzyme, ATP, rather than in the substrate; the second exemplifies the reverse.

In all such reactions the principle of the preservation of bond energy is manifest. The same principle is further illustrated by those processes in which a metabolic waste product originates in a coenzyme-bound form with an energy-rich bond. If the waste product were split off from the complex by hydrolysis, the bond energy would be lost as heat. Generally, it is not so released, but rather through transfer of the coenzyme moiety to a substrate molecule which thereby becomes activated.

Catabolism and anabolism. The metabolic significance of the energy-rich coenzyme-substrate complexes is that they represent the genuine building blocks that take part in the spontaneous biosynthetic reactions. As shown in the preceding sections, these complexes are formed, directly or indirectly, as a result of processes by which substrates are degraded. Hence it now becomes possible to present a general picture of the connection between the catabolic and anabolic aspects of metabolism that is much more revealing than the earlier invoked concept of energetic coupling. However useful this concept may have been at a time when nothing was known about the mechanism of energy transfer, it was unsatisfactory precisely because it avoided this major problem.

Now that the general principles of that mechanism have been elucidated, the notion of energetic coupling has lost its element of mystery. At the same time, it has become evident that the two aspects of metabolism that had formerly been sharply distinguished need no longer be viewed in this manner. On the contrary, the connection between them can now be interpreted far more adequately by considering the breakdown of a substrate as the means by which the genuine building blocks for the synthetic processes are manufactured. This can be expressed by notation (88), which emphasizes the common origin of waste products and of products of synthesis from the same intermediate products rather than the energy relations.

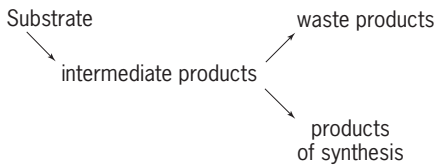

(88)

Essential for these advances in the understanding of metabolic processes has been the recognition that many of the enzyme-controlled biochemical reactions are reversible, and that the same intermediate products, required as building blocks for the synthesis of different classes of compounds, are produced during the metabolism of a wide variety of substrates. This has been amply documented in the sections on oxidative and fermentative catabolism of this article.

The reversibility of enzymatic reactions calls for an additional comment. The equilibrium of such processes is determined by the free-energy changes involved. Thus, if for a reaction of the type shown in reaction (89) the total free energy of A and B is greater

$$A + B \rightleftharpoons C + D \qquad (89)$$

than that of C and D, the equilibrium will be in favor of the latter. The greater the difference, the farther will the reaction proceed to the right. But if this system were combined with one representing an even greater free-energy change, such as reaction (90), the

$$A + E \rightleftharpoons F + G \qquad (90)$$

removal of A by this reaction may bring about an increased formation of B from C and D. This situation has been touched upon in discussing the adenylation of substrates by reaction (91). There it was suggested that the subsequent hydrolysis of pyrophosphate to orthophosphate by pyrophosphatase could drive the adenylase reaction to the right. The reason is that pyrophosphate contains an anhydride (that is, high-energy) bond whose energy is released as heat during hydrolysis. This makes the pyrophosphatase reaction for all practical purposes irreversible, so that its inclusion in a complex enzymatic processes can be an effective method of energetic coupling by a simple mechanism.

$$\text{Substrate} + \text{ATP} \rightleftharpoons \text{substrate-adenylphosphatelate} + \text{PP} \qquad (91)$$

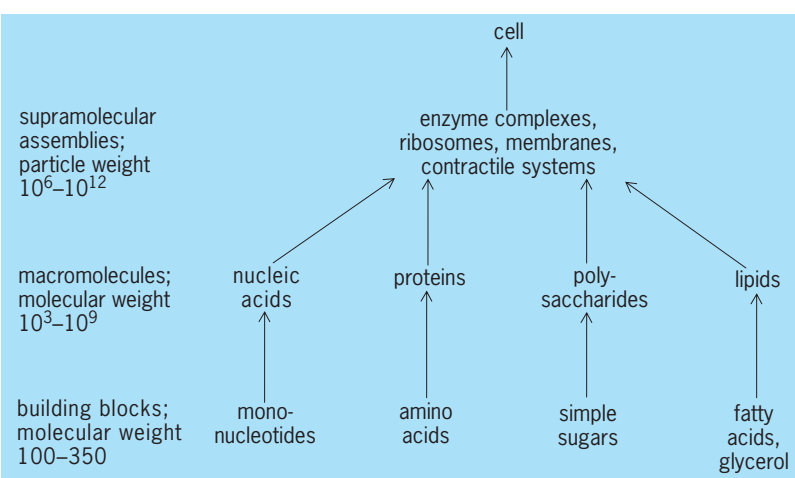

Fig. 18. Levels of organization in bacteria. (*After W. B. Wood, The Molecular Basis of Metabolism* [Unit 3, *Biocore*], *McGraw-Hill, 1974*)

Oxidative and fermentative assimilation. In a growing bacterial culture the number and diversity of synthetic processes comprising the formation of all cell constituents are so vast that a detailed analysis of the mechanism by which any one of them is formed is virtually impossible. This requires a simpler situation, which can be realized by using suspensions of bacteria in a medium in which they cannot multiply because it lacks some ingredient, such as a nitrogen source, necessary for the synthesis of a quantitively important cell material. If a small amount of an oxidizable or fermentable substrate is added to such a suspension, it will be quickly and completely metabolized. At this point the quantity of carbon that disappeared as substrate is, however, much greater than that recoverable as CO_2 or fermentation products, and the missing fraction is found as reserve material in the cells; it may amount to 20–80% of the substrate carbon. Thus the metabolism entails a partial conversion of the substrate into assimilation products; this is referred to an oxidative or fermentative assimilation, depending upon whether it occurred during an oxidative or fermentative decomposition of the substrate.

In the presence of substances such as 2,4-dinitrophenol or sodium azide, which prevent the metabolic generation of ATP, the substrate is entirely converted into metabolic waste products. The absence of assimilation under these conditions indicates that the assimilatory processes involve transphosphorylations. This conclusion is entirely consistent with current knowledge of biosynthetic mechanisms.

The nature of the assimilation product depends on the organism as well as on the substrate. For example, *Pseudomonas saccharophila* produces mainly poly-β-hydroxybutyric acid, a fatlike reserve material, from acetate, butyrate, and glucose; *Rhodospirillum rubrum* also forms this polymer from acetate and butyrate, but forms a glycogenlike polysaccharide when metabolizing propionate, pyruvate, succinate, or malate.

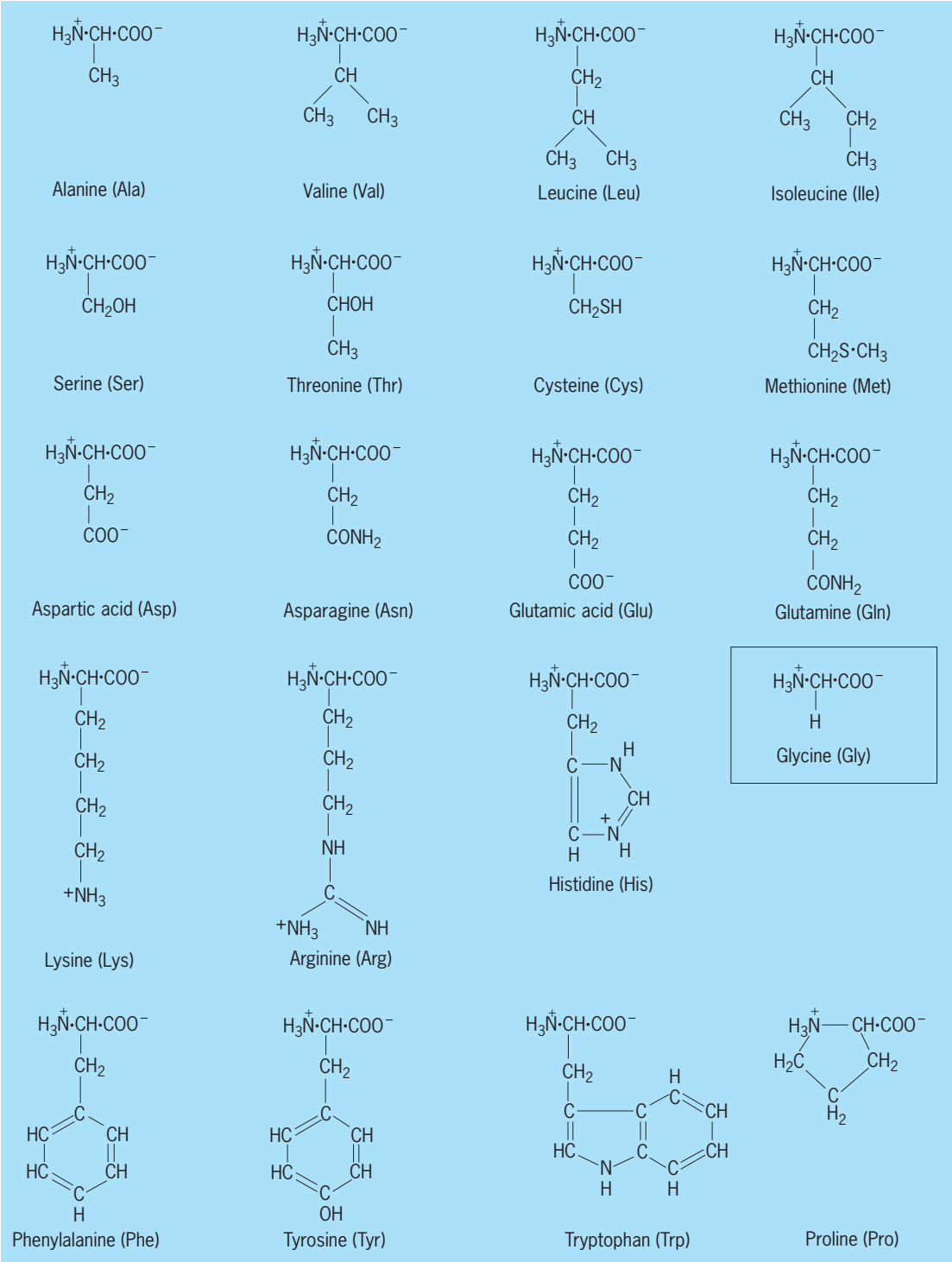

Fig. 19. The 20 amino acids commonly occurring in proteins. All the amino acids have the l configuration when they occur in proteins, except glycine which is not optically active. (*After J. Mandelstam and K. McQuillen, Biochemistry of Bacterial Growth, 2d ed., Halsted Press* [*John Wiley*], *1973*)

Group transfer. In the preceding sections several examples have been given of reactions in which smaller or larger groups of atoms are transferred from one molecule to another. In this manner special building blocks, arising as intermediate products during the metabolism of a particular substrate, can be used in biosynthetic processes for the modification of other compounds, and thus aid in the eventual synthesis of cellular constituents, many of which bear little or no resemblance to the substrates from which they can be manufactured.

The groups in question range from a pair of hydrogen atoms, used for the reduction of some intermediate product and generally donated by reduced pyridine nucleotides, to sulfhydryl, sulfate, amino, methyl-carbamyl, formyl, acetyl, succinyl, glucosyl,

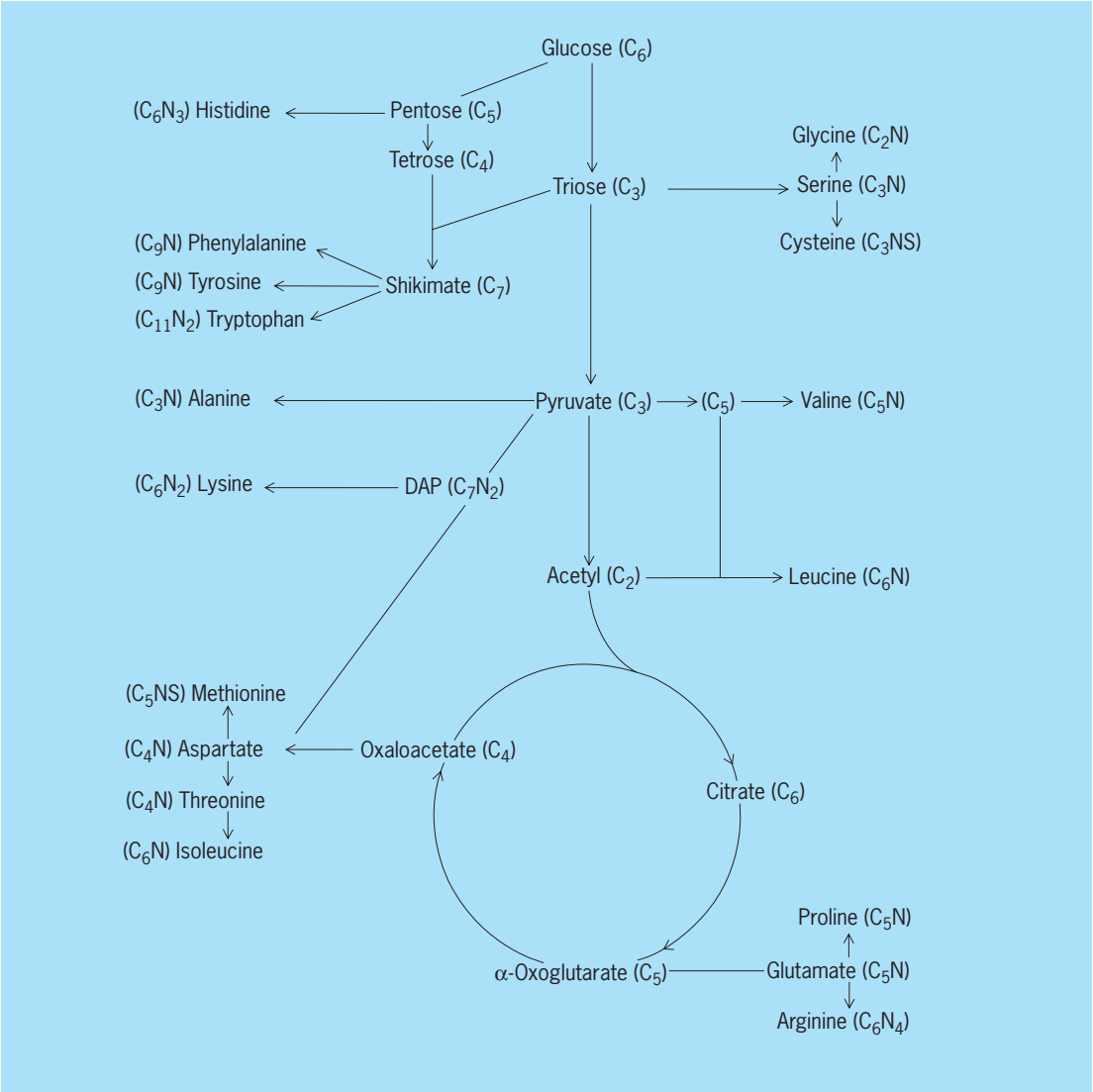

Fig. 20. Synthesis of amino acids. The carbon is supplied by glucose, the nitrogen by NH_4^+, and the sulfur by SO_4^{2-} (diaminopimelate is represented by DAP). (*After J. Mandelstam and K. McQuillen, Biochemistry of Bacterial Growth, 2d ed., Halsted Press* [*John Wiley*], *1973*)

fructosyl, and even entire amino acid groups.

Cellular biosyntheses. The following account of some specific biosynthetic processes contains several examples of group transfer.

Of what is a bacterial cell composed? In addition to water, which accounts for about 70% of the cell's total weight, there are present over 3000 different protein molecules (which account for some 15% of

TABLE 6. Some building blocks and macromolecules of bacterial cells

Building blocks	Macromolecules
Amino acids	Proteins
Nucleotides	Nucleic acids
Monosaccharides	Polysaccharides
Fatty acids, glycerol	Lipids
Amino sugars, amino acids	Peptidoglycan
Lipids, monosaccharides, amino sugars, glycerol or monosaccharides	Lipopolysaccharide

TABLE 7. The amino acid families

Precursor	Family name	Member acids
Oxaloacetic acid* or fumaric acid*	Aspartate	Asparagine, lysine, threonine, methionine isolecuine, aspartic
α-Ketoglutaric acid*	Glutamate	Glutamine, proline, arginine, glutamic
3-Phosphoglyceric acid†	Serine	Glycine, serine, cysteine
Pyruvic acid†	Pyruvate	Alanine, valine, leucine
Erythrose phosphate‡	Aromatic	Phenylalanine, tyrosine, tryptophan
Phosphoenol-pyruvate†		

*An intermediate of the Krebs (tricarboxylic acid) cycle.
†An intermediate in the Embden-Meyerhof-Parnas pathway for carbohydrate degradation.
‡An intermediate in the hexose monophosphate pathway for carbohydrate degradation.

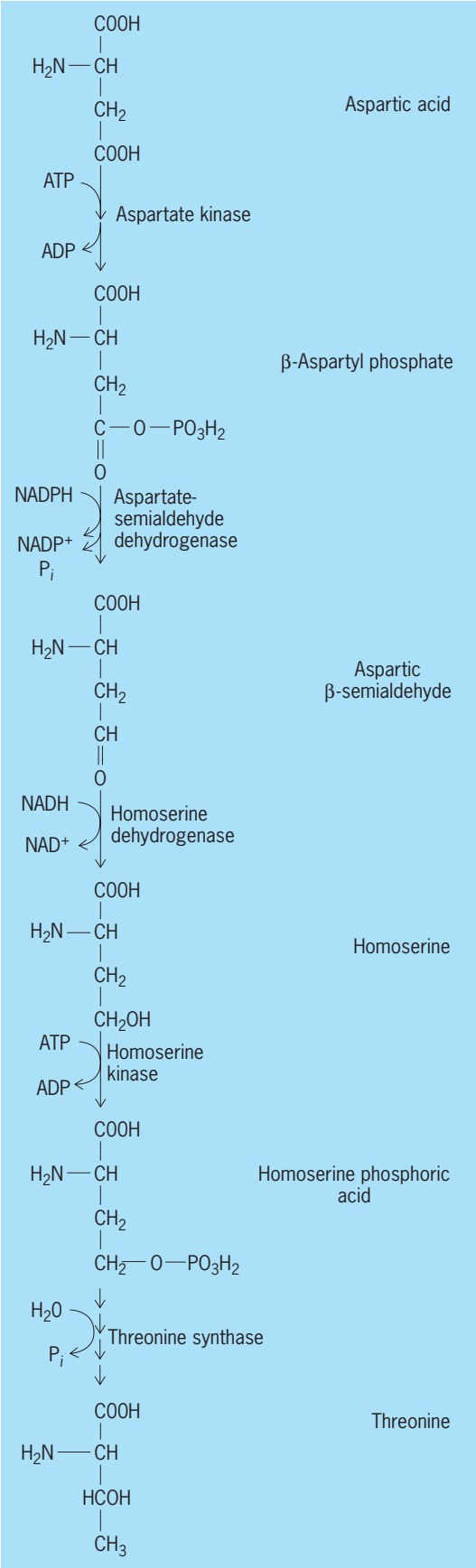

Fig. 21. Biosynthesis of threonine from aspartate. (*After A. Lehninger, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

the total weight), deoxyribonucleic acid (one or two kinds, about 1% of the total weight), ribonucleic acid (some 1000 different molecular species, totaling about 6% of the weight), carbohydrates (perhaps 50 different kinds, 3% of cell weight), lipids (about 50 different kinds, 2% of total cell weight), the building blocks or molecules of intermediary metabolism (500 kinds, 2% of total cell weight), and inorganic ions (about 12 kinds, 1% of cell weight).

Thus the task facing the colon bacterium *Escherichia coli* or the soil inhabitant *Pseudomonas putida* is to utilize the carbon and energy inherent in, for example, acetate

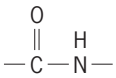

so as to form the diverse array of biological

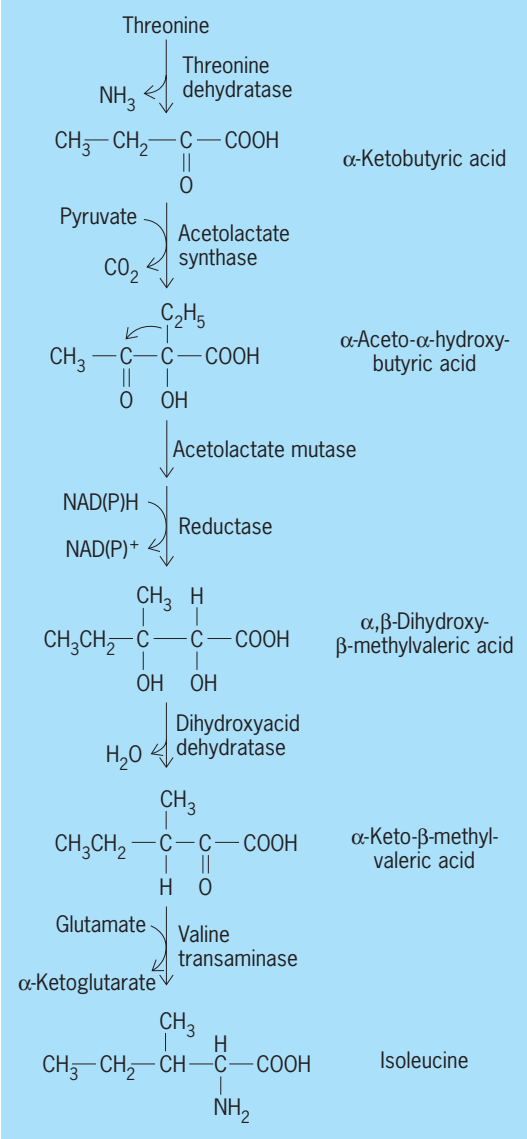

Fig. 22. Pathway for biosynthesis of isoleucine from threonine. (*After A. Lehninger, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

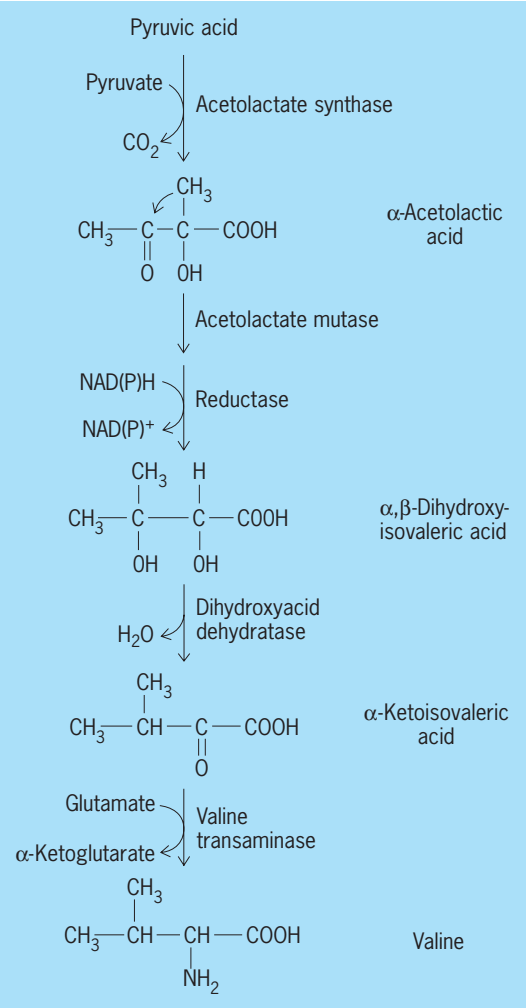

Fig. 23. Pathways to valine and isoleucine. (*After A. Lehringer, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

macromolecules which are present in a cell, and which must be duplicated in order for the cell to divide. In an environment with a sufficient supply of water and acetate, ammonium, phosphate, sulfate, and other ions, the growing cell assimilates these materials and fashions them into the building blocks (**Table 6**) from which the macromolecules, the supramolecular assemblies, and, ultimately, the cell is formed (**Fig. 18**). The energy required for these metabolic conversions is made available by catabolic reactions, and used in the biosynthetic, or anabolic, reactions. Of the latter, it is useful to speak of two main types: one concerned with the conversion of substrate molecules, or molecules derived from the substrate, into the building blocks and other small molecules (vitamins, pigments, coenzymes); and a second type, involved with conversion of the small (low-molecular-weight) molecules into macromolecule and molecular aggregates. All classes or reactions—those of catabolism, as well as those of anabolism—are, of course, specifically catalyzed by a highly efficient class of proteins called enzymes.

Amino acids. The proteins of an organism contain over 20 different amino acids, some of which are acidic, others basic, others essentially neutral, while others differ by virtue of the aromatic rings which are part of their structure (**Fig. 19**). The individual amino acids are formed by different biosynthetic routes, or pathways, as a result of the sequential action(s) of many different enzymes. However, because many of the carbon atoms in the different amino acids are derived from common precursor molecules, it is possible to group the amino acids into families so as to reflect their common biosynthetic origins (**Table 7**); histidine is not in this grouping, primarily because its biosynthesis is more akin to that of purines and pyrimidines than to that of any other amino acid.

A detailed examination of the individual, sequential reactions involved in the synthesis (**Fig. 20**) of each amino acid is beyond the scope of this section,

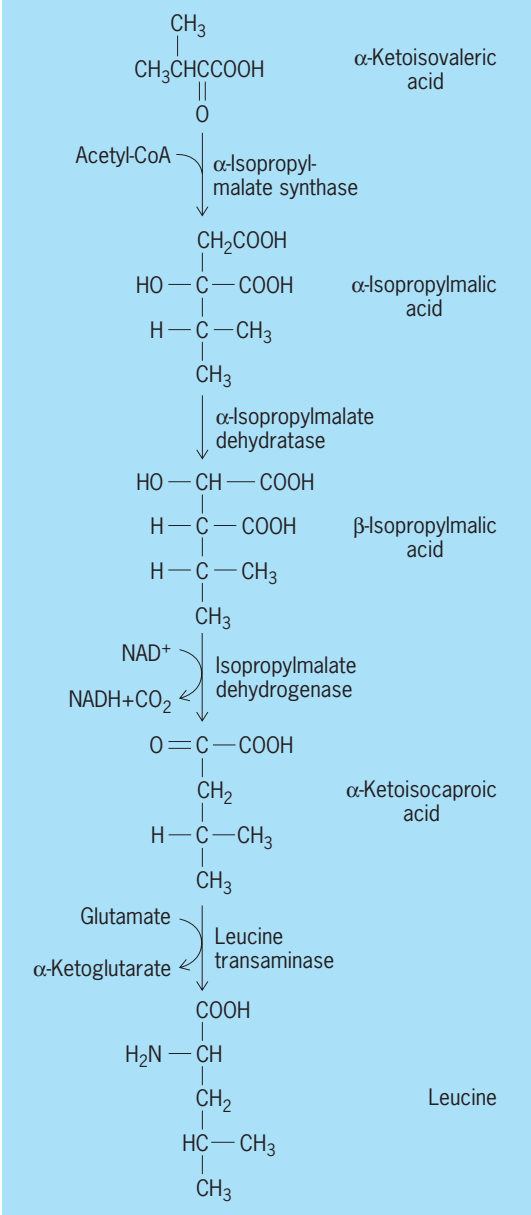

Fig. 24. Biosynthetic pathway to leucine. (*After A. Lehringer, Biochemistry, 2d ed., Worth Publishing Co. 1975*)

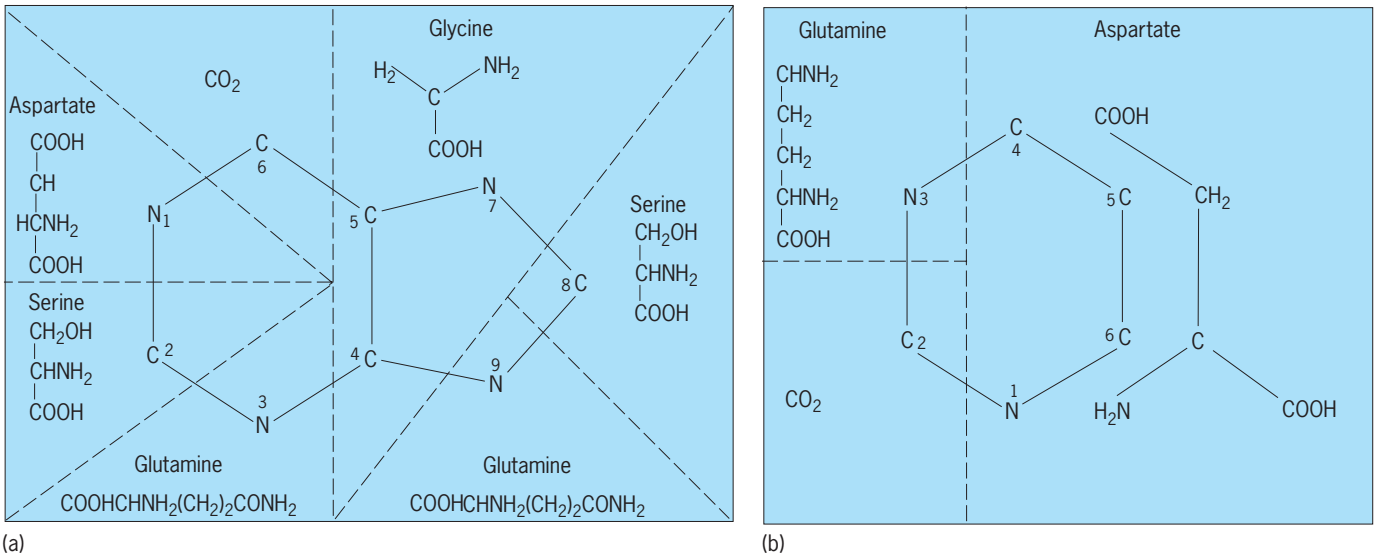

Fig. 25. Origins of the atoms of (*a*) purine rings and (*b*) pyrimidine rings. (*After J. Mandelstam and K. McQuillen, Biochemistry of Bacterial Growth, 2d ed., Halsted Press* [*John Wiley*], *1973*)

but several principles are especially relevant: (1) the energy currency of the cell, ATP molecules, or the so-called high-energy phosphate compounds such as phosphoenolpyruvate are used often in the stepwise synthesis of many amino acids, as are also reductants (electron donors) such as NADH or NADPH which, like the high-energy phosphate compounds, are generated as a result of catabolic activities; (2) several of the amino acids are synthesized in the form of alpha-keto acids, and only in the final enzymatic step in a pathway is the amino (—NH_2) group introduced into the molecule by a set of reactions termed transamination. *See* AMINO ACIDS.

In transamination an amino group of one amino acid (the amino donor) is transferred to an alpha-keto acid (the amino receptor). This receptor thus becomes an amino acid, and the donor is transformed into an alpha-keto acid. The amino donor is, quite

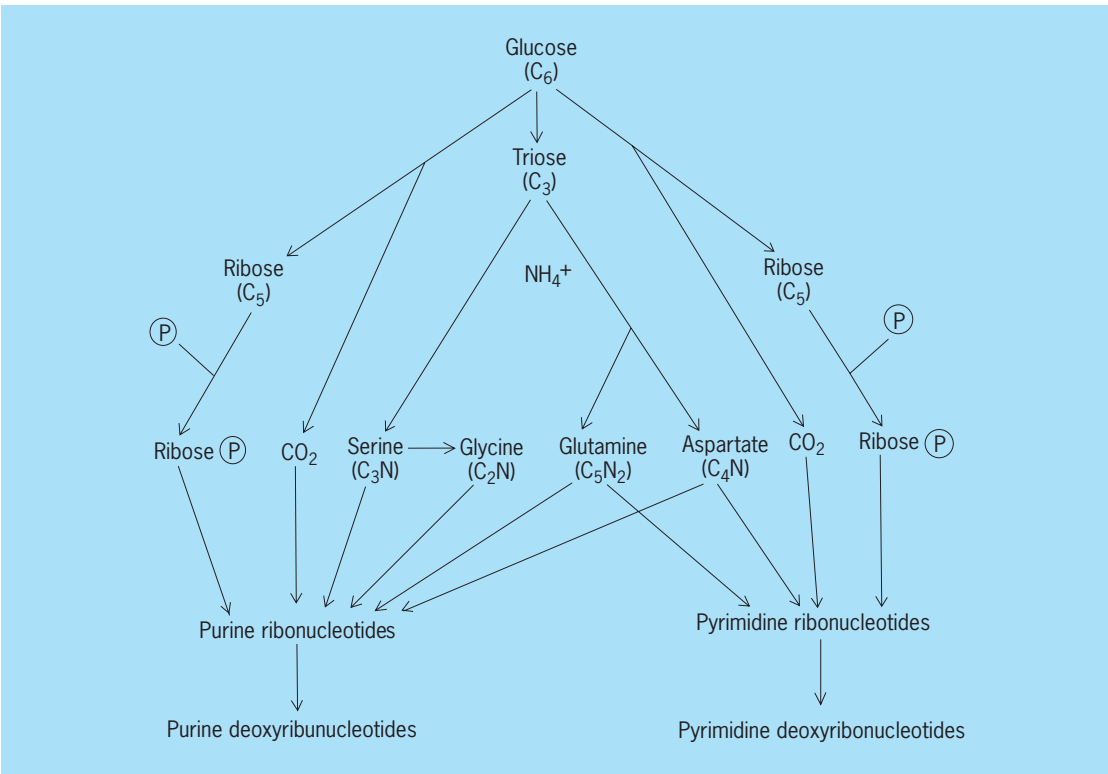

Fig. 26. Diagram of the pathways in the synthesis of nucleotides. (*After J. Mandelstam and K. McQuillen, Biochemistry of Bacterial Growth, 2d ed., Halsted Press* [*John Wiley*], *1973*)

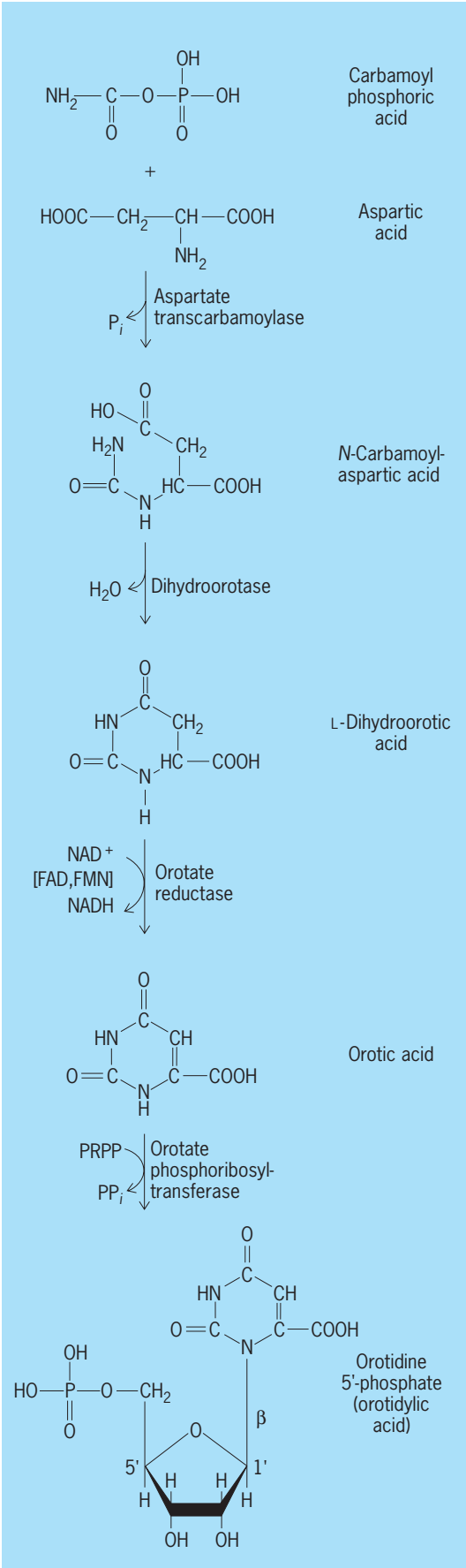

Fig. 27. Biosynthesis of uridylic acid. (*After A. Lehninger, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

often, glutamic acid. This amino acid is synthesized not by a transamination reaction, but by a reductive amination catalyzed by the enzyme glutamic dehydrogenase which has among its substrates the Krebs cycle intermediate, alpha-keto glutaric acid, ammonium ion, and NADH, as in reaction (92).

$$NH_4^+ + \alpha\text{-ketoglutarate} + NADH + H^+ \rightleftharpoons \text{L-glutamate} + NAD^+ + H_2O \qquad (92)$$

Since it is clear that the amino groups of alanine and aspartic acid are also introduced by transamination reactions (where pyruvic acid and oxaloacetic acids are the respective "acceptors") which use glutamic acid as the amino donor, obviously one of the

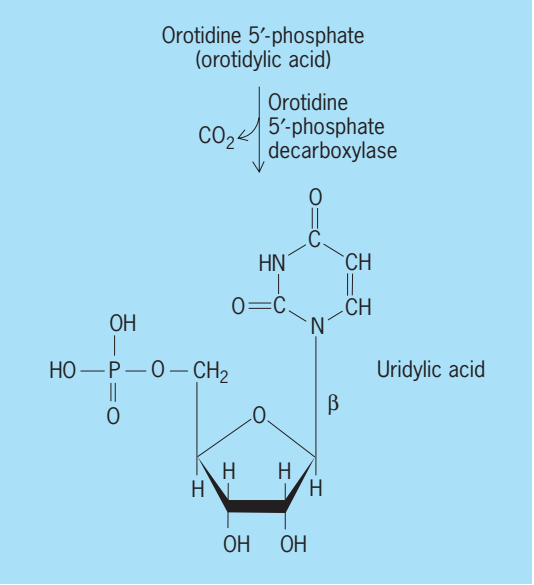

Fig. 28. Formation of uridylic acid (UMP). (*After A. Lehninger, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

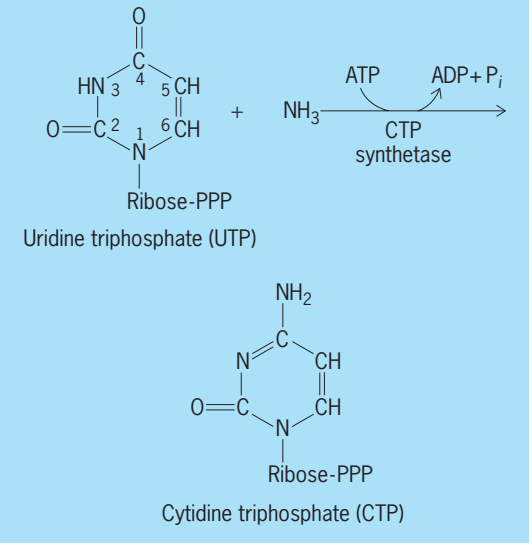

Fig. 29. Amination of UTP to CTP by CTP synthesis. (*After A. Lehninger, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

main pathways by which NH_4^+ is converted to organic nitrogen is via the reductive amination of alpha-keto glutaric acid.

Special enzymatic steps are involved in the introduction of the S atom in amino acids such as cystine; cysteine, and methionine, for the introduction of the methyl (CH_3—) group which is part of methionine, and for the introduction of the additional N and C atoms found in arginine.

Selected examples of amino acid biosynthesis will illustrate some key features. **Figure 21** depicts the pathway for biosynthesis of the amino acid threonine from the amino acid aspartate, and **Fig. 22** depicts the pathway by which some threonine molecules so formed undergo further sequential conversion(s) to form isoleucine. The biosynthetic pathway for valine is depicted in **Fig. 23**, while **Fig. 24** indicates that this pathway branches (after synthesis of alpha-keto valeric acid) and thus leads also to the synthesis of leucine. The synthesis of these and other amino acids, as well as other cellular components, is energetically expensive, but the activities of many enzymes are subject to rapid regulation, and thus unnecessary, wasteful syntheses may be avoided.

The amino acids either formed by these biosynthetic sequences or assimilated from the extracellular milieu are used largely for the biosynthesis of the

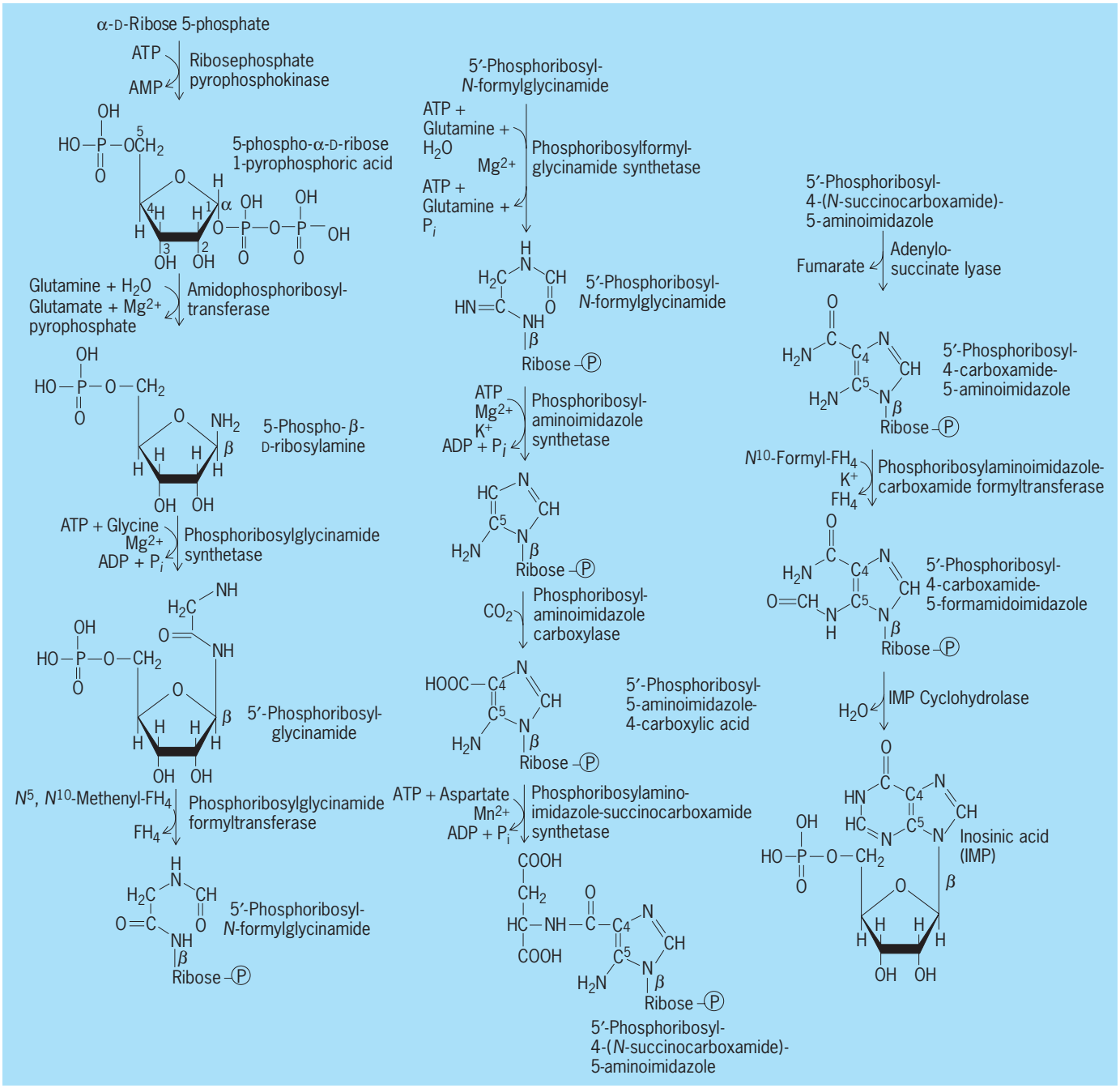

Fig. 30. Biosynthesis of inosinic acid (IMP). (*After A. Lehninger, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

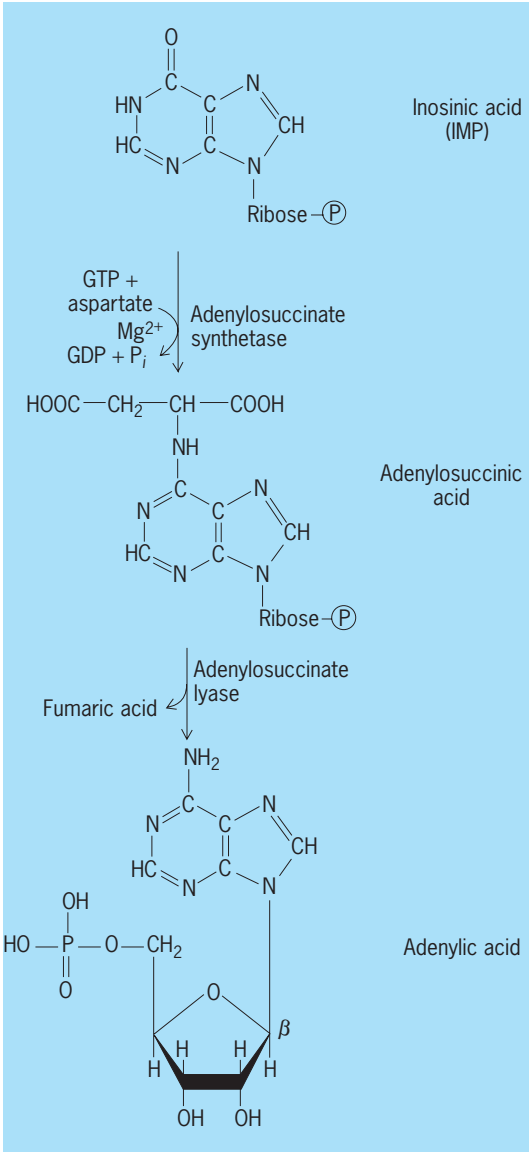

Fig. 31. The pathway from inosinic acid to adenylic acid. (*After A. Lehninger, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

many protein molecules of the cell by a series of endergonic reactions (discussed later).

Just as intermediate metabolites resulting (usually) from carbohydrate metabolism are the starting points for amino acid synthesis, so are other catabolic intermediates the raw materials for biosynthesis of the purine and pyrimidine bases found in the nucleotides, which themselves are precursors of the nucleic acids. Different multienzyme pathways are involved in the formation of the two bases. The diverse compounds which contribute specific atoms of the purine and pyrimidine rings are shown in **Figs. 25** and **26**.

Pyrimidine nucleotide biosynthesis. Largely as a result of the work of Arthur Kornberg and colleagues, the pathway for the conversion of aspartate to uridylic acid, and thence to uridine triphosphate (UTP) and cytidine triphosphate (CTP), was unraveled. As shown in **Fig. 27**, this pathway involves the condensation of aspartate and carbamyl phosphate, and the splitting out of water, for the synthesis of the first ring compound (L-dihydroorotic acid) in the reaction sequence; following an oxidation involving the coenzymes FAD, FMN, and NAD^+, the addition of an energy-rich form of the pentose sugar ribose (PRPP) results in the formation of the nucleotide, orotidine-5′-phosphate. This nucleotide is then decarboxylated to yield uridylic acid (UMP; **Fig. 28**), which is then phosphorylated with ATP to form UTP, as in reactions (93*a*) and (93*b*), and some of this

$$\text{UMP} + \text{ATP} \rightleftharpoons \text{UDP} + \text{ADP} \qquad (93a)$$

$$\text{UDP} + \text{ATP} \rightleftharpoons \text{UTP} + \text{ADP} \qquad (93b)$$

UTP aminated, with the expenditure of another ATP, to form cytidine triphosphate (CTP; **Fig. 29**).

Purine nucleotide biosynthesis. There is a much more complex biosynthetic pathway for synthesis of the purine nucleotides adenosine triphosphate (ATP) and guanosine triphosphate (GTP). Ribose-5-phosphate, formed from glucose by enzymes catalyzing the initial steps of the hexose monophosphate pathway, is activated, that is, converted to an energy-rich form (**Fig. 30**), and to the resulting product (PRPP) an amino group (donated by

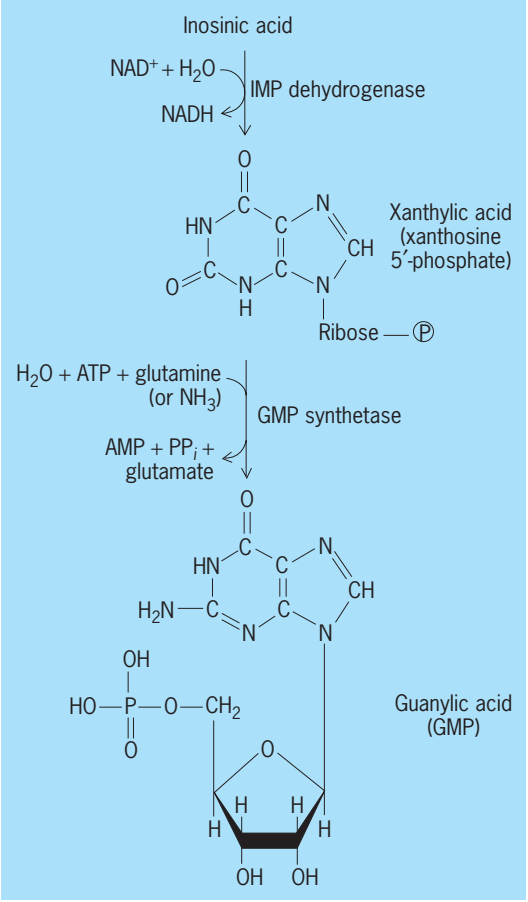

Fig. 32. The pathway from inosinic acid to guanylic acid. (*After A. Lehninger, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

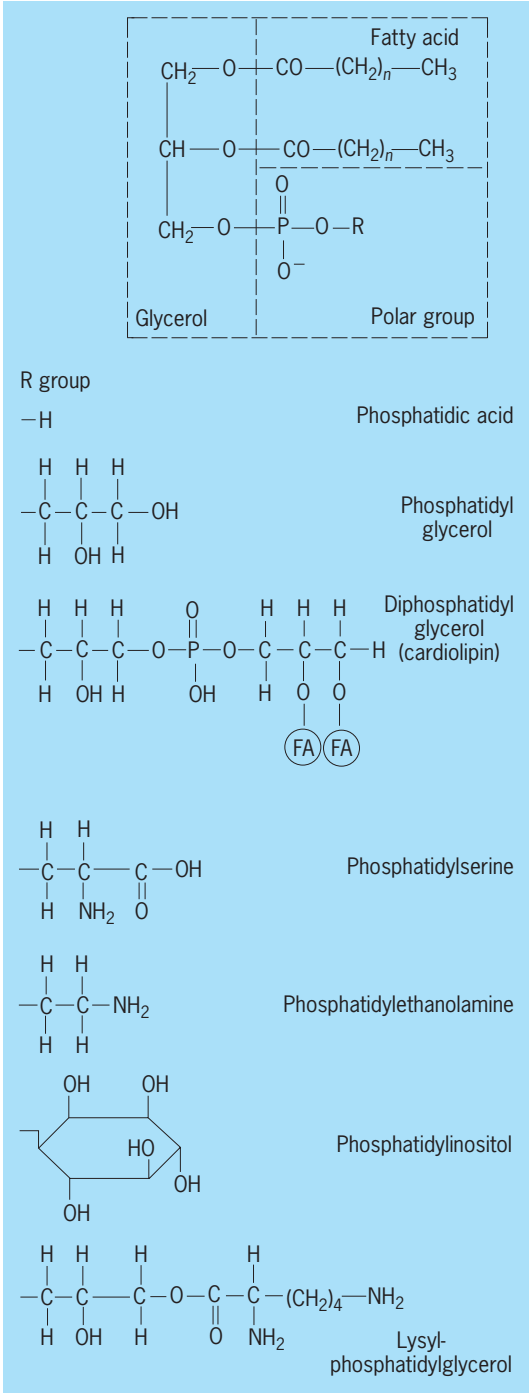

Fig. 33. Phospholipids found in bacteria. (*After J. Mandelstam and K. McQuillen, Biochemistry of Bacterial Growth, 2d ed, Halsted Press* [*John Wiley*], *1973*)

glutamine) is added, the amino acid glycine is linked via a peptide bond to that amino group, and then a

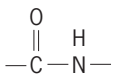

reduction (involving a reduced form of the cofactor folic acid) is carried out to form 5-phosphoribosyl-*N*-formyl glycinamide. Later steps in this enzymatic sequence involve addition of more N (from glutamine) at the expense of ATP, synthesis of the imidazole ring, addition of a carbon atom in the form of CO_2, incorporation of aspartic acid, further folate-involved reductions, and finally formation of the pyrimidine portion of the purine ring and the nucleotide inosinic acid (IMP; Fig. 30). Additional enzymes convert this IMP to adenylic acid (AMP; **Fig. 31**) or guanylic acid (GMP; **Fig. 32**) or both. These nucleotides are then converted to their triphosphate forms (ATP, GTP) again at the expense of ATP generated in catabolism; the triphosphates are used for synthesis of nucleic acids and proteins.

Lipid biosynthesis. In bacterial cells, lipids (**Fig. 33**) usually occur in two main forms: as the lipoproteins and phospholipids which are part of the cell's cytoplasmic membrane and as triglycerides which are storage forms. Triglycerides are formed from fatty acids and glycerol, the latter derived from glucose via reduction of the dihydroxyacetone phosphate (produced in glycolysis) to alpha-glycerol phosphate, as in reaction (94).

$$\text{Dihydroxyacetone phosphate} + \text{NADH} + \text{H}^+ \rightleftharpoons \underset{(\alpha\text{-glycerol phosphate})}{\text{L-glycerol-3-phosphate}} \qquad (94)$$

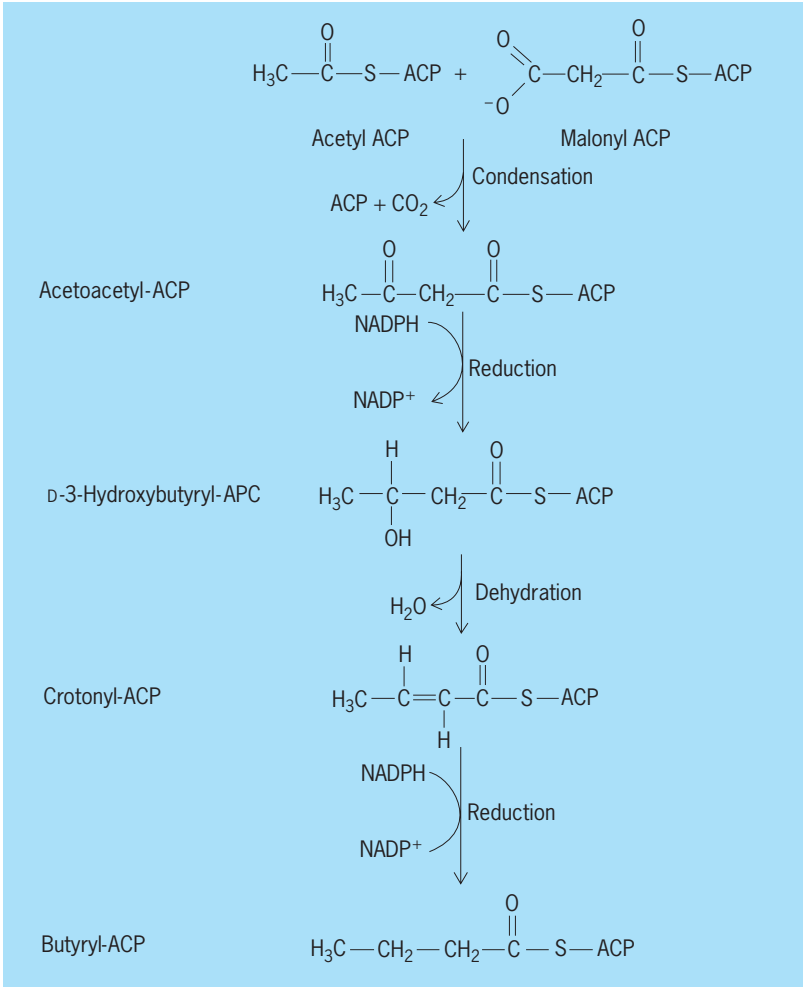

Fig. 34. Reaction sequence in the synthesis of fatty acids: condensation, reduction, dehydration, and reduction. The intermediates shown here occur in the first round of synthesis. (*After L. Stryer, Biochemistry, W. H. Freeman, 1975*)

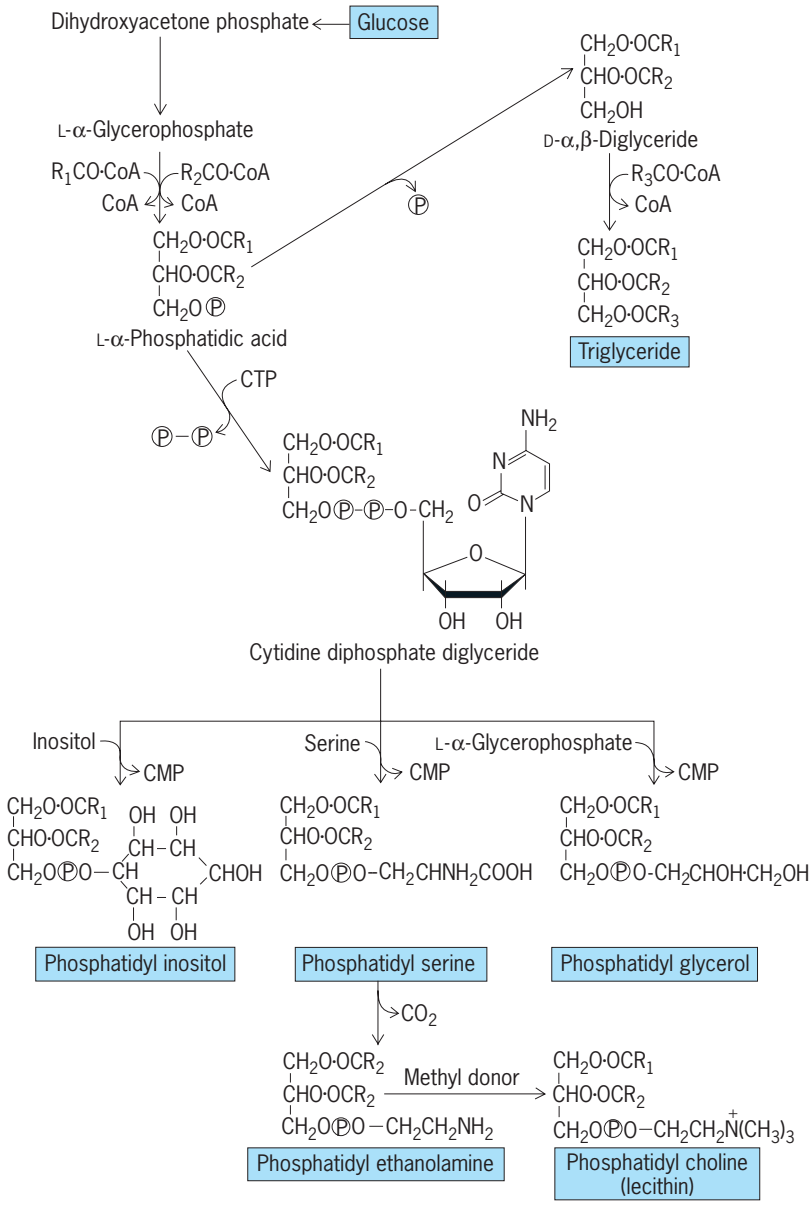

Fig. 35. Biosynthesis of triglycerides and phospholids. (Note that fatty acyl S · ACP rather than fatty acyl S · CoA functions as the fatty acyl donor in some microorganisms.) (*After J. Mandelstam and K. McQuillen, Biochemistry of Bacterial Growth, 2d ed., Halsted Press* [*John Wiley*], *1973*)

Fatty acid synthesis utilizes acetyl coenzyme A (acetyl-CoA), formed from pyruvate, which is then carboxylated, with the involvement of CO_2, ATP, and an enzyme which contains the cofactor, biotin, to form malonyl-CoA, as in reaction (95). The

$$\text{Acetyl} - \text{CoA} + CO_2 + H^+ + \text{ATP} \rightleftharpoons \text{malonyl} - \text{CoA} + \text{ADP} + P_{\text{inorganic}} \qquad (95)$$

malonyl-CoA and additional acetyl-CoA molecules are then converted to the forms in which they are active in fatty acid biosynthesis by reacting with a sulfhydryl protein (acyl-carrier-protein, ACP-SH); these ACP derivatives are enzymatically condensed to form acetoacetyl-ACP, CO_2, and ACP-SH (**Fig. 34**). Following the multistep reduction of the acetoacetyl-ACP to butyryl-ACP, to the latter another malonyl-ACP is added to form a six-carbon ACP derivative (and, again, CO_2 and an ACP-SH), which again is reduced to form the saturated fatty acid–ACP compound. Repetition of the malonyl-ACP addition and reduction events with 8, 10, 12, etc., carbon fatty acid–ACP molecules results in the formation of the long-chain saturated fatty acids. For the synthesis of palmitic acid (16 carbon atoms), 8 acetyl-CoA molecules, 7 of ATP, and 14 of NADPH are utilized. Formation of some of the long-chain monounsaturated fatty acids (those containing a double bond, —CH═CH—) diverges from the biosynthetic pattern at the 10-carbon–ACP-derivative stage: a specific dehydratase forms a double bond between the beta and gamma carbons, and then three additional two-carbon units are successively incorporated and reduced, but this addition is to the

$$-\overset{\overset{\displaystyle O}{\|}}{C}-\text{ACP}$$

portion of the molecule, rather than, as before, to the terminal methyl (CH_3—) carbon, thus eventually resulting in the formation of an acid such as palmitoleic (C_{16}, monounsaturated):

$$CH_3(CH_2)_5CH{=}CH(CH_2)_7\overset{\overset{\displaystyle O}{\|}}{C}-OH$$

Other pathways for synthesis of monounsaturated fatty acids may well exist.

The amounts of triglycerides—glycerol which is esterified with three different long-chain fatty acids—are rather small in bacteria. Enzymatic synthesis proceeds using a molecule of L-alpha-glycerophosphate and two different (depicted as R_1 and R_2 in **Fig. 35**) molecules of fatty acids in the CoA esterified form resulting in the formation of a phosphatidic acid which, by loss of phosphate, is converted to a diglyceride. Triglyceride formation involves addition of a third fatty acid, via its CoA derivative (Fig. 35).

The phosphatidic acid is a branching point in this biosynthetic pathway, for synthesis of phospholipids proceeds by the enzymatic linkage of part of a cytidine triphosphate molecule to the unesterified hydroxyl group of the phosphatide (Fig. 35). The resulting coenzyme (cytidine diphosphate glyceride) is the active intermediate which reacts with either inositol, serine, or L-alpha-glycerophosphate to form, finally, phosphatidyl-inositol, -serine, -glycerol, -ethanolamine, or -choline, which constitute the major cellular phospholipids, and are found mainly in the cell membrane or membrane wall (envelope) complex of gram-positive or gram-negative organisms, respectively.

Glycolipids, compounds which contain sugar(s) linked to the unesterified hydroxyl group of a diglyceride, are present in bacterial cells, again most likely in the membranes or in the membrane wall complexes. Mannose, glucose, galactose, glucosamine, and rhamnose are among the sugars present in different glycolipids. The sugars may be present singly

or linked to form di- or trisaccharide chains which are joined by glycosidic linkage to the free hydroxyl group of the diglyceride. Details of glycolipid biosynthesis remain unclear, although it seems likely that sugars in nucleotide-coenzyme form (linked to guanosine diphosphate) are intermediates, and the sugar (glycosyl) moiety of this nucleotide is transferred by the responsible enzyme to the diglyceride, with the concomitant release to guanosine diphosphate.

Glycophospholipids, which contain both carbohydrate and phosphate residues, are being found with increasing regularity in diverse bacterial groups, and may be as abundant and as significant as glycolipids. Details of the biosynthesis of these compounds is even less clear than for glycolipids, but the same general synthetic features and types of intermediates may be anticipated.

Other lipids of cellular significance are the terpenes, among which are the carotenoid pigments and the long-chain primary alcohols such as bactoprenol (a compound of importance in wall synthesis). Both types of compounds are synthesized

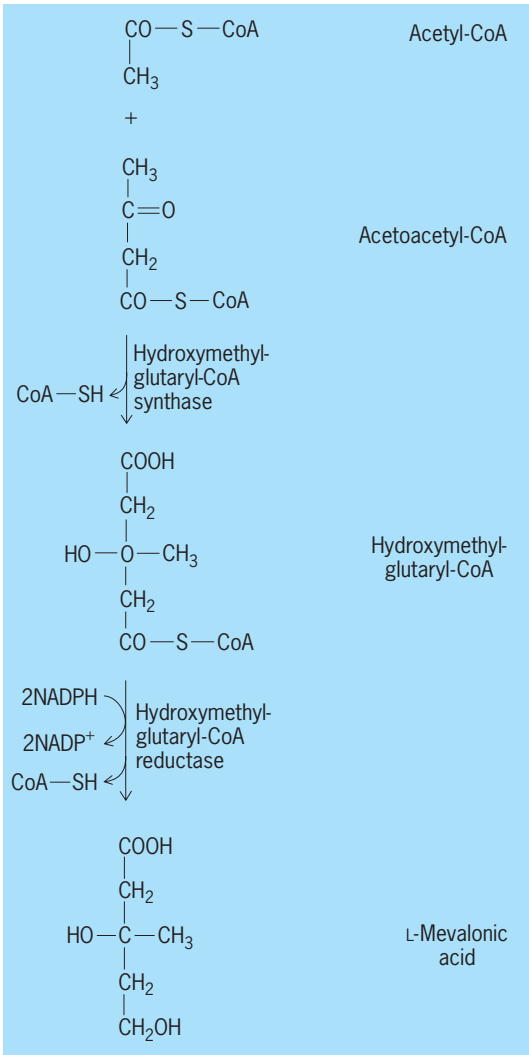

Fig. 36. Synthesis of mevalonic acid. (*After A. Lehninger, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

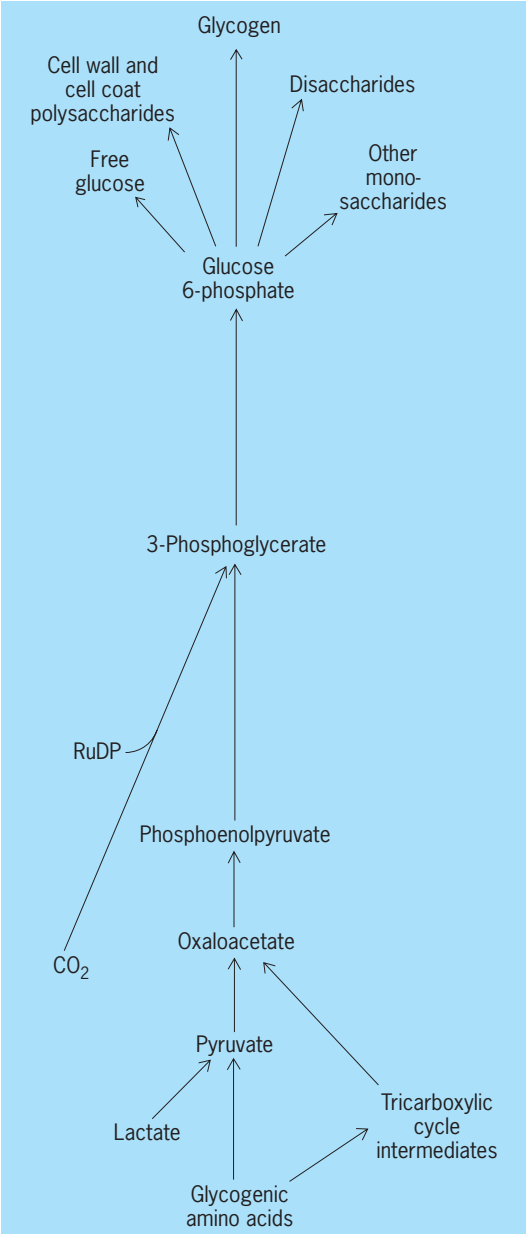

Fig. 37. Central pathway of hexose biosynthesis. Auxiliary pathways feed into the central pathway from CO_2, lactate, and amino acids; diverging pathways lead from glucose 6-phosphate to other carbohydrates. RuDP is ribulose 1,5-diphosphate, an important intermediate in autotrophic formation of hexose. (*After A. Lehninger, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

via condensation of acetyl-CoA molecules to form mevalonic acid (**Fig. 36**) which then undergoes extensive modification(s) to form isoprenyl-phosphate compounds which themselves are further modified and joined to form the terpenes. *See* CAROTENOID; CHOLESTEROL.

Acetyl-CoA molecules also undergo condensation to form acetoacetyl-CoA, which after reduction becomes beta-hydroxybutyric acid, several molecules of which may be joined by ester linkages (between the hydroxyl group of one molecule and the carboxyl group of another) to form the polymer, poly-beta-hydroxybutyrate; this is a storage form of lipid

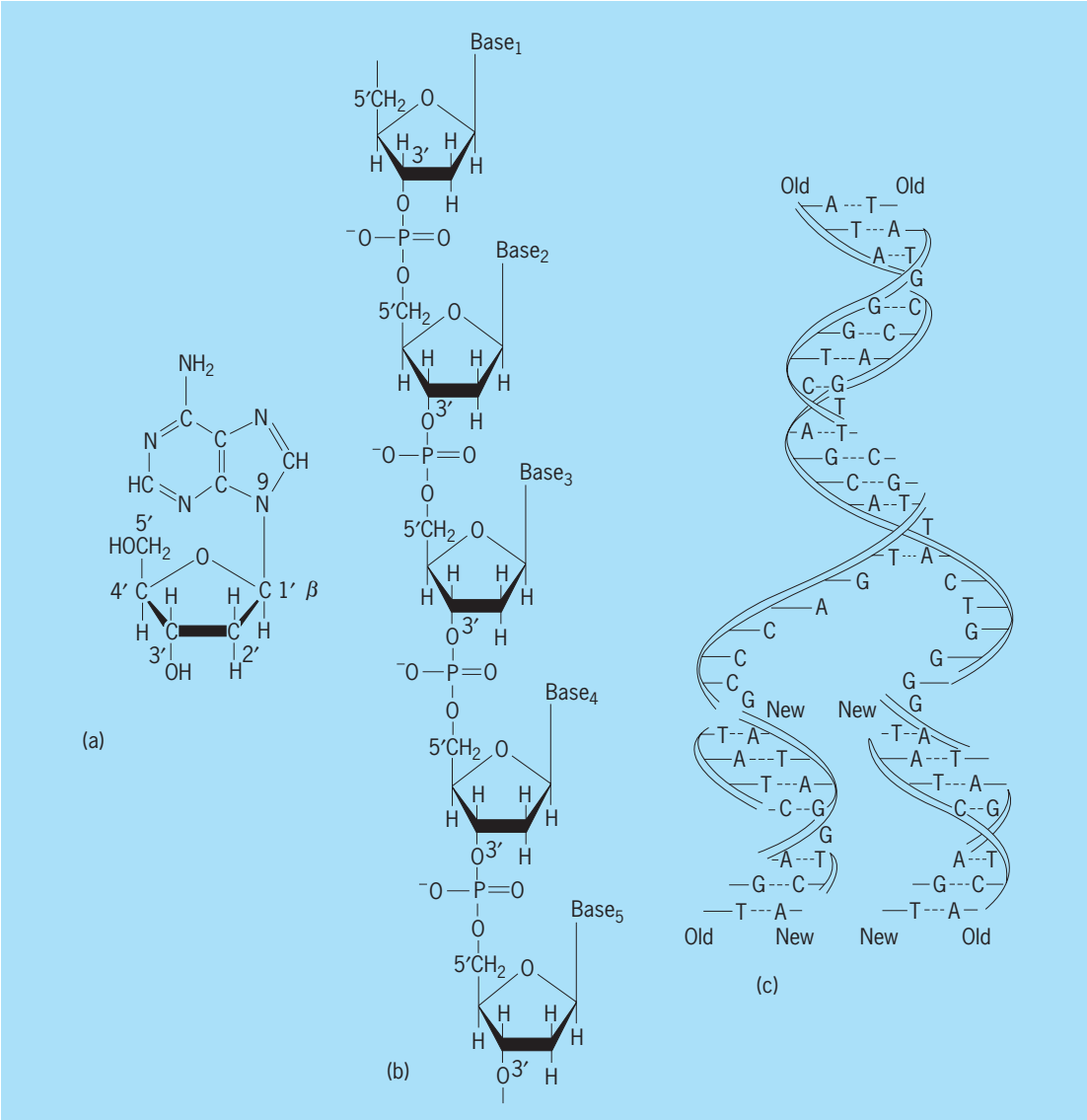

Fig. 38. Deoxynucleotide and polydeoxynucleotide structures. (*a*) 2′-Deoxyadenosine. (*b*) A polydeoxynucleotide chain. (*c*) Configuration of the DNA double helix, showing mechanism of self-duplication. (*After W. B. Wood, The Molecular Basis of Metabolism* [*Unit 3, Biocore*], *McGraw-Hill, 1974*)

in many bacteria, particularly in situations where substrate carbon is in ample supply, and nitrogen is limited.

A rather different type of storage polymer, polyphosphate, is often found within bacterial cells. Whether these polymers function primarily as an energy or as a phosphorus reserve is not yet certain. Clearly, they differ from other known storage polymers by their lack of carbon and nitrogen. They are, thermodynamically, rather high-energy phosphate compounds and synthesized by one of two ways, involving either a polyphosphate kinase which catalyzes transfer of the terminal phosphoryl group to polyphosphate, reaction (96) by a polyphosphate

$$(P_i)_n + \text{ATP} \rightleftharpoons (P_i)_{n+1} + \text{ADP} \qquad (96)$$

transferase which uses 1,3-diphosphoglycerate as phosphate donor, reaction (97).

$$\text{1,3-diphosphoglycerate} + (P_i)_n \rightleftharpoons \text{3-phosphoglycerate} + (P_i)_{n+1} \qquad (97)$$

Carbohydrate synthesis. Synthesis of this class of molecules has several aspects—the synthesis of simple sugars (monosaccharides) by organisms utilizing compounds such as CO_2, fatty, organic, or amino acids, or other noncarbohydrate materials as their growth substrate(s), and the synthesis of polysaccharides such as glycogen, starch, cell walls, and extracellular polymers.

As indicated in **Fig. 37**, glucose may be synthesized from noncarbohydrate substrates other than CO_2 by comparable, but not identical, intermediates and pathways which are involved in carbohydrate catabolism. Different bacterial types may vary in the precise details followed.

Of special significance is the synthesis of triose and hexose phosphates from CO_2 by photosynthetic

and chemosynthetic bacteria. In the former the required energy in the form of ATP and NADH or NADPH results from the trapping of radiant energy and its conversion to ATP and so forth, while in the latter group these compounds are formed as a result of the oxidation of compounds such as NH_4^+, NO_2^-, H_2, H_2S, SO, $S_2O_3^{2-}$, Fe^{2+}, and so forth. The biosynthetic pathways for CO_2 assimilation and conversion into sugars is analogous to the Calvin cycle involved in green plant photosynthesis. *See* PHOTOSYNTHESIS.

Macromolecular biosyntheses. The polymerization of the different monomeric building blocks, whose synthesis has just been sketched, into macromolecules such as polysaccharides, proteins, nucleic acids, cell walls, and extracellular polymers is invariably an energy-requiring process. As will be seen, the requisite energy for driving these endergonic reactions comes, again, from the large amount of ATP formed in catabolism.

Deoxyribonucleic acid. Deoxyribonucleic acid, the genetic material of the cell, exists as double-helical structures of complementary strands of polynucleotides. Four different purine and pyrimidine deoxyribonucleotides—dAMP, dGMP, dCMP, and dTMP—are present in each strand (**Fig. 38**). Replication of a preexisting DNA strand proceeds with a series of enzymes (DNA polymerases) which link the deoxyribonucleoside triphosphates, dATP, dGTP, dCTP, and dTTP, to form a polymer of complementary base composition to the "primer" strand serving as a template; pyrophosphate is eliminated from each triphosphate as this polynucleotide formation occurs. Each strand of the existing double-helical DNA is, presumably, being separated and copied simultaneously, thus resulting, when separation is complete, in the formation of two identical molecules of DNA, each of which contains a single strand of the original, parental DNA. The genetic information contained in the original DNA has thus been duplicated. *See* DEOXYRIBONUCLEIC ACID (DNA).

Ribonucleic acid. The information inherent in the sequence of nucleotides in DNA is transcribed into the polyribonucleotides of an RNA molecule (**Fig. 39**) when one is synthesized from the purine and pyrimidine ribonucleoside triphosphates (ATP, GTP, CTP, UTP), by a polymerizing enzyme (DNA-directed RNA polymerase) which "copies" only one strand of DNA.

The RNA formed by RNA polymerase using DNA as a directive template is of three types: messenger, transfer, and ribosomal (mRNA, tRNA, rRNA). The mRNA serves as a template for protein biosynthesis, and hence each mRNA molecule must contain the information transcribed from a gene, or group of genes; thus this class of molecules is of heterogeneous composition. Transfer RNAs are rather specific for each of the some 20 amino acids with which they interact, and in addition have a portion (the anticodon) of the molecule which is complementary to a specific region(s) (the codons) of mRNA with which they will eventually combine. Ribosomal RNA is incorporated along with proteins into the subcellular particles, ribosomes, which serve as the sites of protein synthesis. *See* RIBONUCLEIC ACID (RNA).

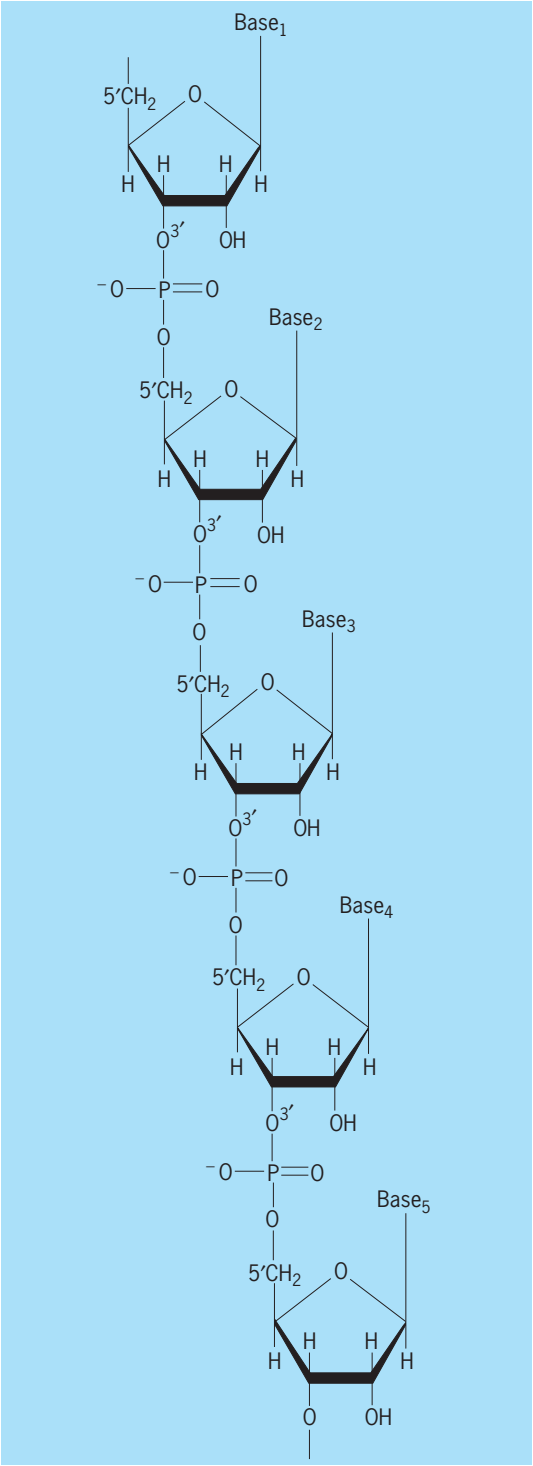

Fig. 39. Structure of a polyribonucleotide. (*After W. B. Wood, The Molecular Basis of Metabolism* [*Unit 3, Biocore*], *McGraw-Hill, 1974*)

Proteins. Protein biosynthesis proceeds after each of the some 20 types of amino acids is activated by means of an enzyme specific for both the amino acid and its corresponding species of tRNA. In the reaction, ATP is cleaved into AMP and pyrophosphate,

and the energy is thus provided to drive the reaction.

The resulting amino-acyl-tRNA molecule is endowed with a specificity (in the tRNA moiety) which has its complement(s) in the mRNA strand; thus each amino acid (in the form of its tRNA derivative) is ensured of being located next to the correct neighboring amino acids, that is, the genetic information of DNA and RNA is translated into the correct positioning of amino acids in the polypeptide chain (protein) to be formed.

The site of protein synthesis is the ribosomes which are complex aggregates of RNA ($^2/_3$ by weight) and protein ($^1/_3$ by weight). Once the mRNA, the tRNAs, and the ribosome have been appropriately complexed, the peptide bond formation between adjacent amino acids (that is, protein elongation) occurs as the mRNA is translocated with respect to the ribosome in a complex manner of which all the details are not yet clear. However, the complete process ensures the collinearity of the base sequence (information) in a gene and the amino acid sequence of the product of the gene—a polypeptide (protein). *See* RIBOSOMES.

Although the bulk of cellular protein is synthesized on ribosomes, some rather small proteins, such as polypeptide antibiotics, are not; instead, steps analogous to certain of these fatty acid syntheses occur in the cytoplasm.

Polysaccharides. The bacteria synthesize a variety of polysaccharides, including intracellular reserve, or storage material such as glycogen or glycogenlike materials, as well as extracellular polymers such as the homopolymers levan and dextran, a variety of heteropolysaccharides, and structural polymers such as the cell wall components peptidoglycan and lipopolysaccharide.

The synthesis of intracellular homopolymers such as glycogen usually occurs when growth is limited by the supply of a suitable nitrogen source, and where there is an ample supply of extracellular carbon substrate (note the parallel with the conditions which favor formation of the lipid reserve polymer, polybeta-hydroxybutyrate). Glycogen biosynthesis proceeds by enzymatic conversion of glucose to glucose-1-phosphate, and the latter compound then is a substrate for the enzyme ADP-glucose pyrophosphorylase which utilizes ATP and the glucose-1-phosphate to form the nucleoside diphosphate derivative of the sugar, adenosine diphosphate glucose (ADP-glucose) and pyrophosphate. The ADP-glucose is the activated, or carrier, form of the glucose and substrate for the enzyme glycogen synthase which repetitively adds the glucose moiety to an existing (primer) bit of glycogen and thus extends the chain by one glucose unit, and releases the ADP.

Nucleoside diphosphate sugar derivatives (**Fig. 40**) are widely used for biosynthesis of other polysaccharides, but more often than not the carrier, or activated, form is a derivative of uridine diphosphate (UDP) rather than ADP as was the case for glycogen synthesis. This is true, for example, for synthesis of a last part of the lipopolysaccharides,

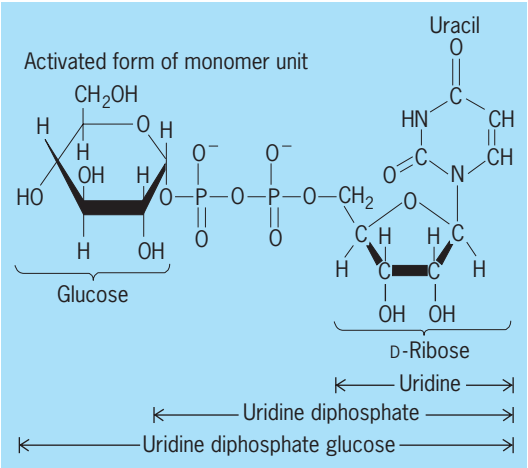

Fig. 40. Structure of the nucleoside diphosphate sugar, urdine diphosphate glucose. (*After W. B. Wood, The Molecular Basis of Metabolism* [*Unit 3, Biocore*], *McGraw-Hill, 1974*)

and the cell wall polymer, peptidoglycan. But, by contrast, the synthesis of another type of polymer found in bacterial cell walls, the teichoic acids, involves the cytidine diphosphate (CDP) derivatives of the sugar alcohols, glycerol or ribitol, to form glycerol-teichoic acid or ribitol-teichoic acid respectively. In other instances guanosine diphosphate (DP) appears to serve as the carrier, or activated, form. These differences in the particular nucleoside triphosphate used in polysaccharide biosynthesis should not, however, obscure the significant role that these compounds play in biosynthesis of macromolecules other than the nucleic acids.

Because such a major portion (up to 15%) of a cell's total dry weight may be in the form of the bag-shaped macromolecule peptidoglycan, and because this molecule is unique to the bacterial world, it is important to consider the general features of its synthesis (**Fig. 41**). Peptidoglycan, as the name implies, consists of peptides and sugars; the linear chains of polysaccharides are crosslinked to one another by means of short peptides. The repeating units of the polysaccharide chains are two glucose derivatives, *N*-acetyl glucosamine and *N*-acetyl muramic

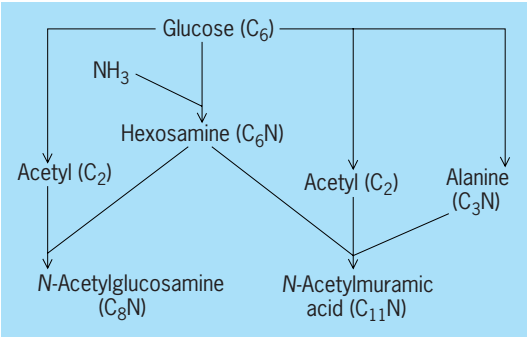

Fig. 41. Synthesis of hexosamines. The acetylated hexosamines are constituents of the peptidoglycans which are essential components of bacterial cell walls. (*After J. Mandelstam and K. McQuillen, Biochemistry of Bacterial Growth, 2d ed., Halsted Press* [*John Wiley*], *1973*)

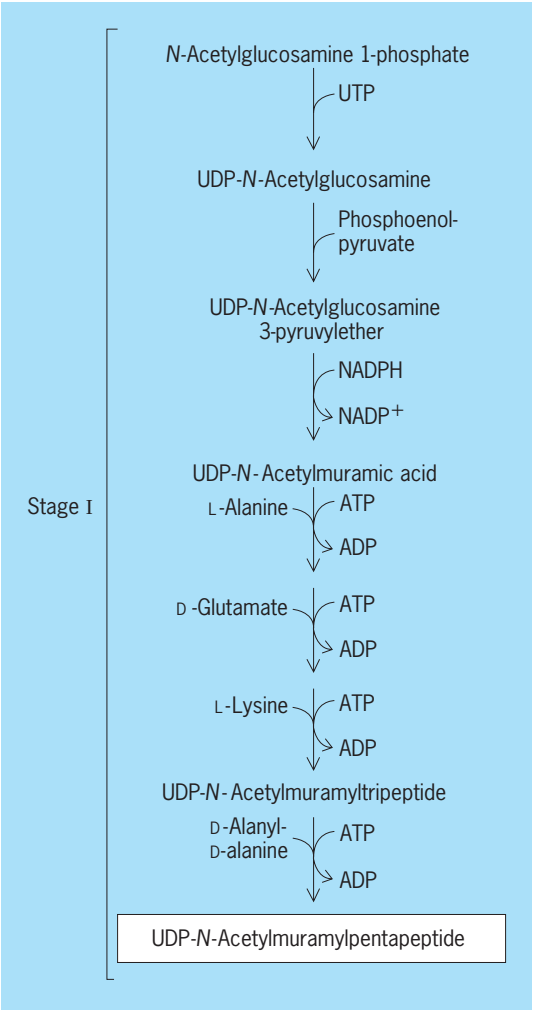

Fig. 42. Steps in the biosynthesis of the peptidoglycan of a bacterial cell wall. (*After A. Lehninger, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

acid which, as their UDP derivatives, are linked to form a disaccharide. To the UDP-*N*-acetyl muramic acid is added (**Fig. 42**), successively, the amino acids L-alanine, D-glutamate, L-lysine, and finally a dipeptide, D-alanyl-D-alanine, thus forming a pentapeptide chain linked to the UDP-*N*-acetyl muramic acid (the formation of each peptide link requires the expenditure of an ATP; ribosomes, however, are not involved).

The next stage in synthesis involves enzymatic transfer of the UDP–muramic acid–pentapeptide to a terpene, undecaprenyl phosphate, which is bound in the cell membrane. This lipid acts as a carrier molecule which is able to effect the transport of the polar (charged) UDP across the lipid-rich cell membrane. To the *N*-acetyl muramic acid portion of the complex is first added *N*-acetyl glucosamine (as its UPP derivative), and then to the resulting dissacharide-pentapeptide complex is added short peptide chains which will act as cross links between these peptide chains attached to the muramic acid moieties (**Fig. 43**).

The final step in peptidoglycan formation is cleavage of this entire disaccharide unit from the lipid carrier and concomitant transfer to a preexisting (primer) polysaccharide chain. Cross-linking between parallel peptidoglycan chains occurs via the short peptide chains (**Fig. 44**). Repetitive addition of the disaccharide repeating unit (including side chains and cross-linking chains) and repeated cross-linking (it is this latter reaction that is sensitive to the antibiotic penicillin) brings about the encircling of the cell with the single molecule of enormous size and strength. To this rigid backbone are attached the various teichoic acids, and in gram-negative organisms an additional wall layer, the complicated lipopolysaccharide.

Extracellular polymers. Although the synthesis of many extracellular polysaccharides involves nucleoside diphosphate carriers, others, usually homopolymers (such as dextran and levan), are synthesized without involvement of these coenzymes. Particularly when dextran or levan is synthesized by cells growing on sucrose (a disaccharide of glucose and fructose), a rather different principle is used: the energy of the glycosidic bond linking the two

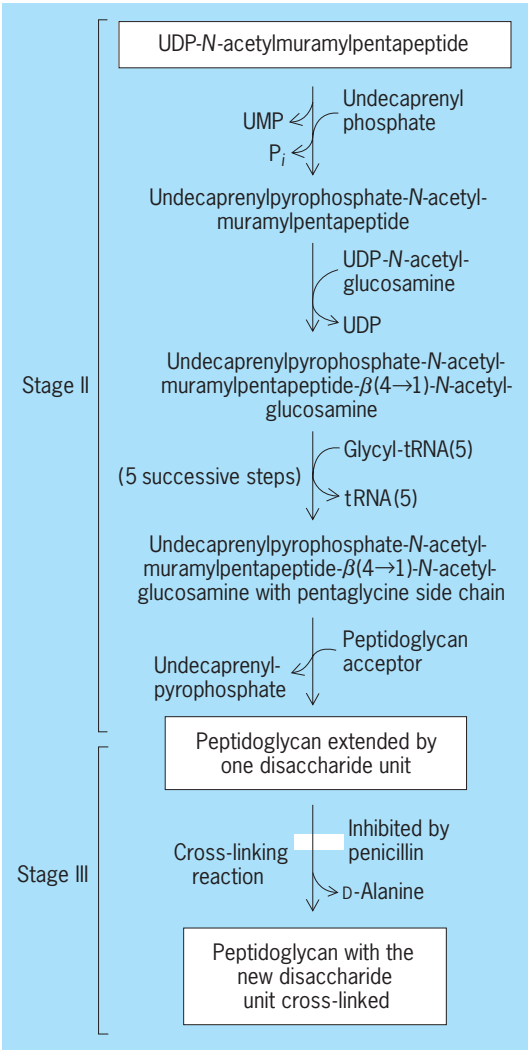

Fig. 43. Stages II and III in peptidoglycan biosynthesis. (*After A. Lehninger, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

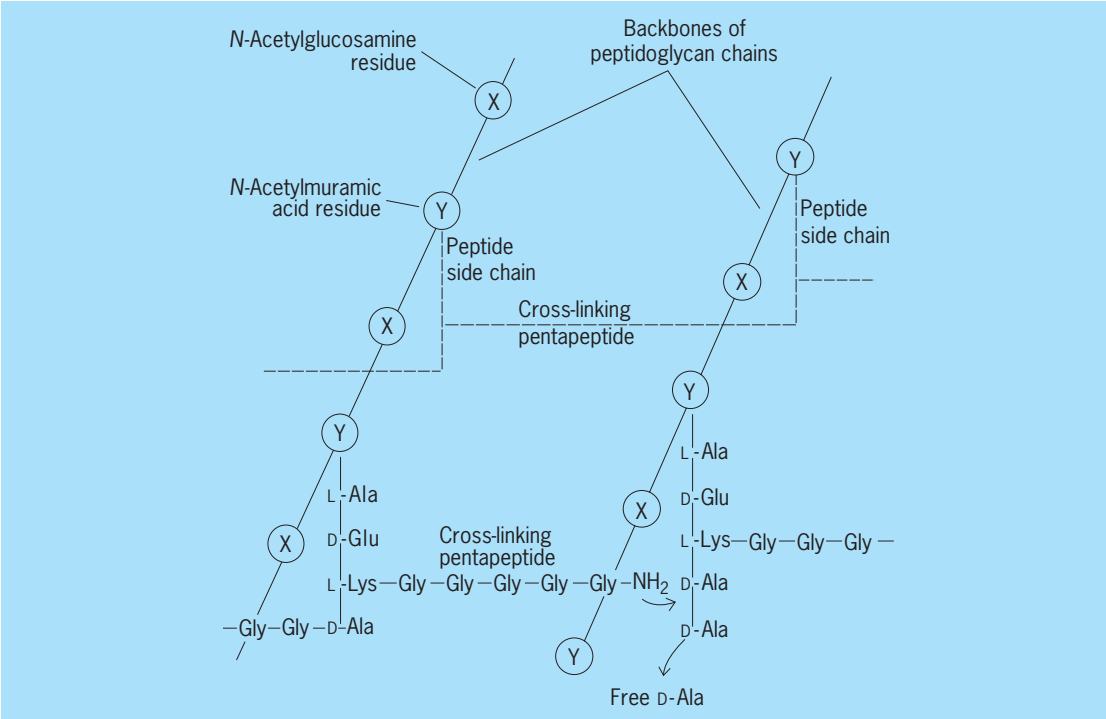

Fig. 44. Completion of a cross link between two adjacent peptidoglycan chains in the bacterial cell wall. This reaction is blocked by penicillin. (*After A. Lehninger, Biochemistry, 2d ed., Worth Publishing Co., 1975*)

monosaccharides is preserved in an enzyme-glucose complex formed when the enzyme dextransucrase, for example, cleaves sucrose with the release of fructose. The energy inherent in the enzyme-bound glucose moiety is then utilized when this glucosyl unit is added to a dextran molecule, thus increasing the latter's size. Similarly when the enzyme levansucrase cleaves sucrose, with the release of glucose, the enzyme-bound fructosyl moiety is in an activated form and may be added to a levan molecule without utilization of an additional energy source.

Synthesis of extracellular heteropolysaccharides, where two or more different carbohydrates or carbohydrate-derivatives are polymer constituents, invariably involves the successive, stepwise addition of the monomers as nucleoside diphosphate and complexes.

Energy and growth. It should thus be clear that synthesis of cell materials (anabolism), an endergonic process, utilizes massive amounts of ATP. Although it would be very difficult to calculate the amount of ATP needed by an organism synthesizing all its cell material from a single substrate such as acetate, it is a somewhat easier task to calculate the ATP requirements for synthesis of macromolecules from their monomeric units, where only polymerization reactions are involved. Such calculations indicate that for every 0.1 g (dry weight) cell material formed, over 3 millimoles of ATP are needed, and that, accordingly, 1 mole of ATP should be sufficient energy to form over 30 g cell material from preformed monomers.

Experimental assessment of the molar growth yields of bacteria and yeasts reveals that only a scant 10 g dry weight of cell material is actually formed per mole of ATP generated in catabolism. The basis for the discrepancy between this 10 g, as measured experimentally, and the some 30 g, as calculated, is not yet clearly understood, but the "loss" may represent energy expenditure for cellular repairs, active transport, motility, and, of course, as heat.

Transport across membranes. The very composition of the cell or plasma membrane, rich in phospholipids, proteins, and combinations thereof, renders it an effectively impermeable barrier particularly to highly polar molecules and ions, which are of course characteristic intermediate metabolites in the cell; the membrane thus serves a useful purpose in preventing diffusion of these metabolic intermediates, formed intracellularly, from the cell. The same membrane feature simultaneously excludes a variety of polar organic molecules (such as glucose, acetate and amino acids) from gaining entrance to the cell interior by simple diffusion across the cell membrane, and in addition, prevents the loss or release, by diffusion, of metabolic end products which may be toxic to the cell if appreciable amounts of them accumulate. The rapid growth rate of bacteria demands a very high metabolic rate, and hence of substrate availability and end-product removal. These problems are met by bacteria by the use of transport mechanisms which effect a movement of glucose, amino acids, and so on, not only across the cell membrane but against a concentration gradient. Because such activity results in an increase in free energy, the process must be driven by energy-yielding reactions which in bacteria are known to be of several types.

Located in the periplasmic space—the area between the cell membrane and cell wall—are a number of different proteins that bind, selectively, different inorganic ions, sugars, or amino acids. These binding proteins are thought to play an important role in transport processes.

One type of sugar transport mechanism known to be present in some anaerobic bacteria involves the phosphorylation of incoming sugar by an enzyme located in the cell membrane using as the energy source a specific phosphorylated protein formed in the interior cytoplasm. Transported across the membrane, thus, is not the free sugar, but the phosphorylated form. In some aerobic bacteria a different tactic is employed: amino acids and sugars are not chemically modified (for example, phosphorylated), but their transport is intimately coupled to transfer of electrons from substrate to oxygen. The details of this coupling are not fully known, but one view, that of the chemoosmotic hypothesis, presumes a proton-gradient generated by electron transport as the driving force for such active transport.

Cell organization and enzyme location. It has been mentioned for enzymes involved in transport process and for final stages of peptidoglycan synthesis that some enzymes occupy particular locales in the bacterial cell. It is roughly accurate to say that most enzymes of substrate catabolism and monomer, or building-block, biosynthesis are located in the cytosol, and hence termed "soluble" enzymes, while many of those concerned with polymerization or synthetic reactions of macromolecules involved in

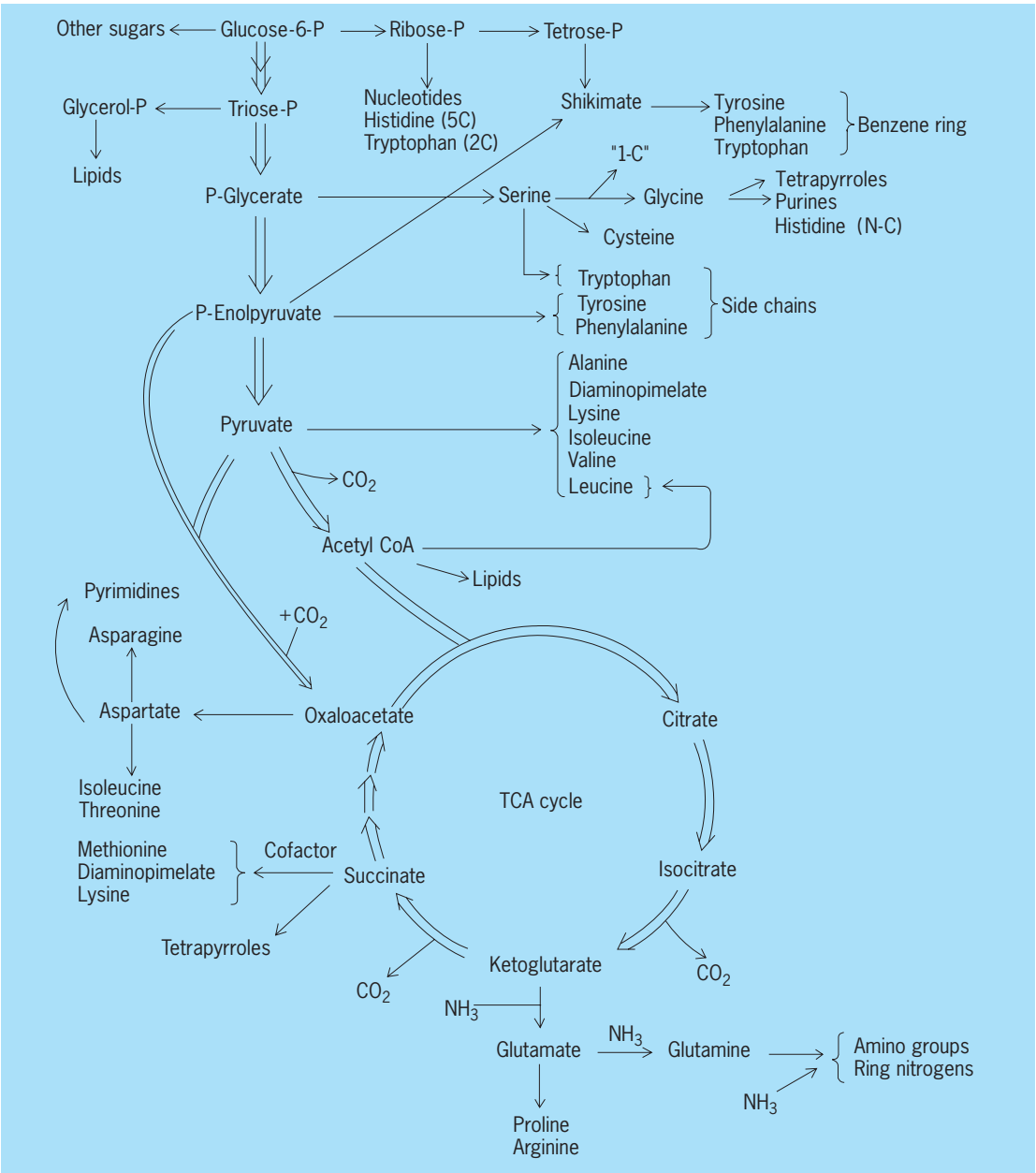

Fig. 45. Relation of the amphibolic pathways (indicated by the double arrows) to the main anabolic pathways. (*After B. Davis et al., Microbiology, 2d ed., Harper and Row, 1973*)

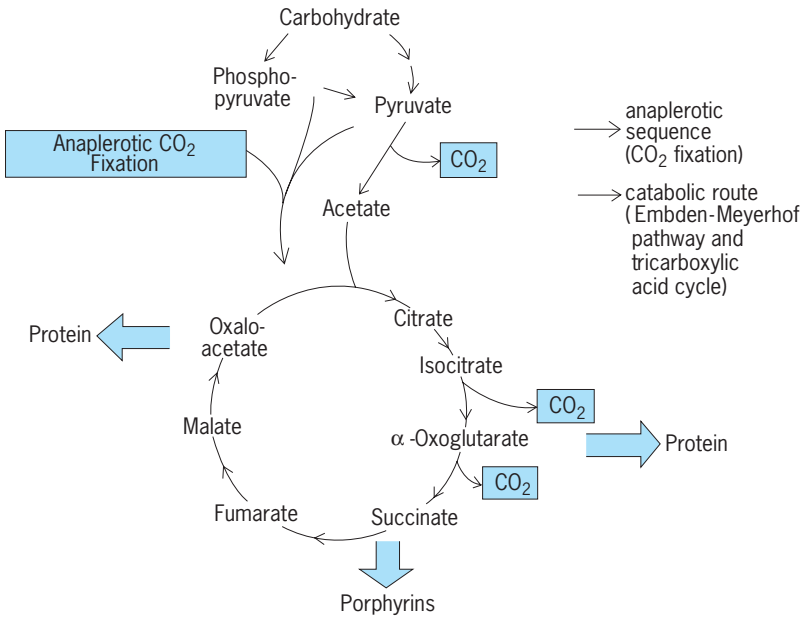

Fig. 46. Routes for the provision of energy and of cell components during the utilization of carbohydrates. (*After P. N. Campbell and G. D. Greville, eds., Essays in Biochemistry, vol. 2, Academic Press, 1966*)

structural features of the cell are bound to, or located in, those structural components. In addition to enzymes involved in synthesis and degradation of phospholipids, biosynthesis of glycolipids, peptidoglycans, lipopolysaccharides, exopolysaccharides, teichoic acids, and other macromolecules significant in cell structure, there are membrane-associated enzymes involved in other processes: ATPases, electron transport components (such as oxidases, dehydrogenases and cytochromes) and, perhaps, enzymes for initiation of DNA synthesis. In contrast, enzymes involved in carbohydrate degradation, biosynthesis of amino acids, purines and pyrimidines, and so on, are thought not to be bound to cell structural components. Some enzyme activities involved with synthesis or degradation of storage polymers such as poly-beta-hydroxybutyrate or glycogen may be associated with particles composed of these polymers which are found in the cytosol.

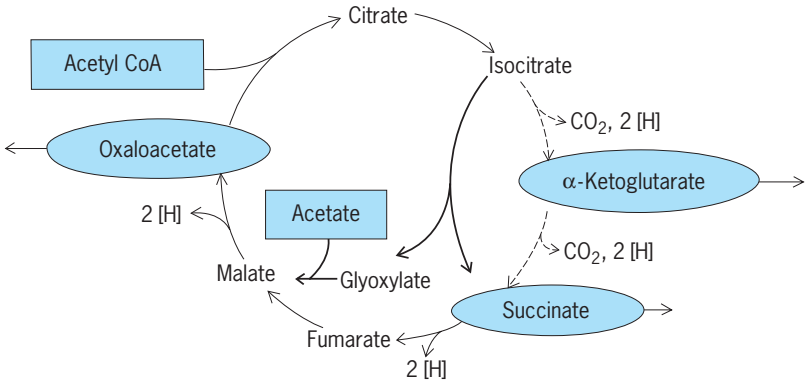

Fig. 47. Krebs cycle, with the glyoxylate cycle reactions (heavy arrows) inserted. The broken arrows indicate reactions which may be bypassed when the glyoxylate cycle is being used by organisms growing with acetate as sole source of carbon and energy. (*After B. Wood et al., Microbiology, 2d ed., Harper and Row, 1973*)

Anaplerotic sequences. It should be clear from the discussions of catabolism and anabolism that a great deal of substrate carbon is used in the formation of one of a cell's major products—another cell—as well as in the compounds more commonly regarded as end products, CO_2 for aerobes, acids or alcohols for anaerobes. The source(s) of the cellular macromolecules are, of course, the building blocks derived from the pathways of carbohydrate breakdown or Krebs cycle activities. But how, for example, is the latter cyclical series of reactions to continue operation if intermediates in this cycle (such as ketoglutarate, oxaloacetate, and succinate) are being depleted as they are used, as they almost constantly are, for amino acid and purine and pyrimidine biosynthesis?

Clearly there must be ancillary reactions, or pathways, which effect a net synthesis and replenishment of certain metabolic intermediates and thus make possible the maintenance of cyclical activities. Such replenishment reactions have been termed "anaplerotic" (from the Greek, for "filling up") by the British biochemist Hans Kornberg.

One such key reaction(s) in organisms metabolizing carbohydrates or other compounds which are degraded to pyruvate is the enzymatic condensation of CO_2 and either pyruvate or phosphoenolpyruvate to form oxaloacetate, which thus generates an acceptor molecule for acetyl-CoA and keeps the cycle functioning, as in reaction (98). The interrelation-

$$\text{Pyruvate} + CO_2 + \text{ATP} \longrightarrow \text{oxaloacetate} + \text{ADP} + P_i \qquad (98)$$

ship of this synthesis and other related activities is shown in **Fig. 45**.

For organisms growing on substrates which are not degraded to pyruvate (for example, long-chain fatty acids) or on compounds containing only one, or no, carbon-to-carbon bonds (for example, acetate and methanol), the same problem, that of providing a sufficient supply of intermediates for the Krebs cycle, as well as the additional one of providing precursors for carbohydrate synthesis, must be solved. In the case of organisms growing at the expense of acetate, or compounds degraded to acetate, this is done by means of the glyoxylate cycle, which brings about the net synthesis of a four-carbon dicarboxylic acid, malate, from two acetate molecules (**Fig. 46**). For organisms growing at the expense of one-carbon compounds such as methane, methanol, or methylamine, oxidized forms of these substrate molecules undergo condensation reactions with, on the one hand, glycine, to form serine, and eventually, phosphoglycerate (which is of course readily convertible to pyruvate), or on the other hand, with a pentose phosphate, to form, eventually, a hexose phosphate which can be cleaved to give rise to phosphoglyceraldehyde and thence pyruvate. In both instances some of the triose phosphates formed must be diverted in a cyclical fashion to regenerate the initial acceptor molecules (glycine, or a pentose phosphate) for the oxidized one-carbon compounds.

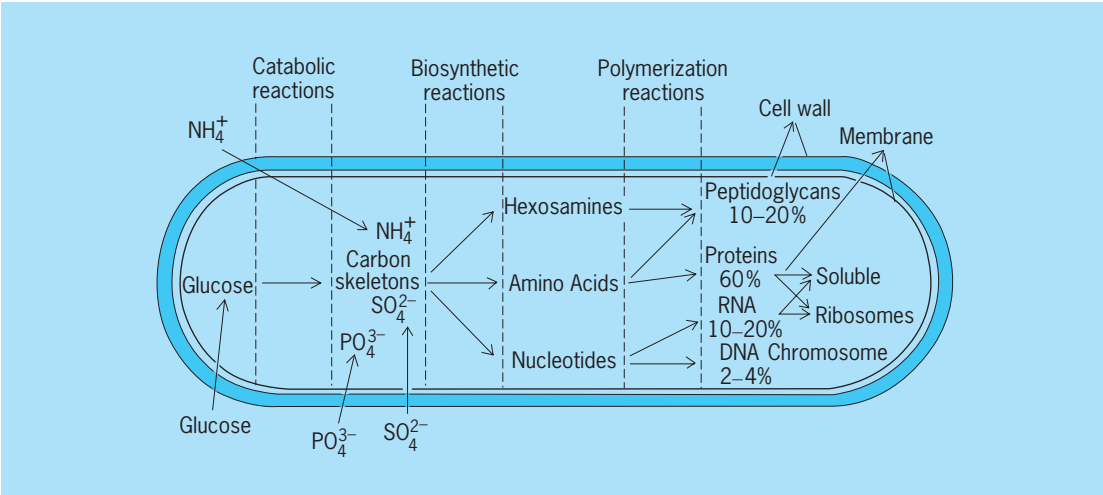

Fig. 48. Generalized flow diagram for the synthesis of the bacterial cell components. (*After J. Mandelstam and K. McQuillen, Biochemistry of Bacterial Growth, 2d ed., Halsted Press* [*John Wiley*], *1973*)

In a larger sense, the reductive pentose pathway, or Calvin cycle, operative in photo- or chemosynthesis may be considered an anaplerotic sequence, for the net result is to provide multicarbon compounds for conversion to building blocks for macromolecular biosynthesis.

Amphibolic pathways. Although initially convenient for purposes of categorization, the terms "catabolism" and "anabolism" may now be seen to have considerable overlap. For example, the pathways of carbohydrate degradation to pyruvate and the Krebs cycle act both to provide energy as well as key biosynthetic intermediates. Those metabolic sequences (**Fig. 47**) which function in this dual capacity are often termed amphibolic, to indicate that they may serve both roles, or under some conditions, either role. This convergence of biosynthetic and catabolic pathways, with common intermediates, is one expression of the close interrelationships which exist between energy generation and consumption in the functioning cell.

Other aspects of anabolism. It is not only multiplying cells that carry out endergonic syntheses. Cells about to enter dormant or resting states (endospores, cysts), into particular stages in a life cycle (stalk forms), or undergoing conversion to carry out specific functions (heterocysts) also undergo particular metabolic or intra- or extracellular structural changes, as do also dormant cells when germinating or others when becoming motile. The energy expenditures connected with these biosyntheses, as well as those connected with N_2 fixation, nitrate reduction, either swimming or gliding motility, photo- and chemotactic behavior, or colonial morphogenesis are no less significant to the economy of the cell than are those which have been enumerated in more detail in the foregoing paragraphs, and summarized in **Fig. 48**. E. R. Leadbetter

Bibliography. H. A. Barker, *Bacterial Fermentations*, 1957; V. H. Cheldenlin, *Metabolic Pathways in Microorganisms*, 1961; G. N. Cohen, *Biosynthesis of Small Molecules*, 1967; G. N. Cohen, Regulation of enzyme activity in microorganisms, *Annu. Rev. Microbiol.*, 19:105, 1965; B. D. Davis et al., *Microbiology*, 3d ed., 1980; H. W. Doelle, *Bacterial Metabolism*, 2d ed., 1975; I. C. Gunsalus and R. Y. Stanier (eds.), *The Bacteria*, 7 vols., 1960–1979; P. E. Hartman and S. R. Suskind, *Gene Action*, 1965; M. Inouye, *Bacterial Outer Membranes as Model Systems*, 1986; H. A. Krebs and H. L. Kornberg, Energy transformation in living matter, *Ergeb. Physiol.*, 49:212–298, 1957; M.T. Madigan et al., *Brock's Biology of Microorganisms*, 8th ed., 1996; J. Mandelstam et al., *Biochemistry of Bacterial Growth*, 3d ed., 1982; W. G. Murrell and I. R. Kennedy, *Microbiology in Action*, 1988; D. Nelson and M. Cox, *Lehninger Principles of Biochemistry*, 3d ed., 2000; E. L. Oginsky and W. W. Umbreit, *An Introduction to Bacterial Physiology*, 2d ed., 1959; M. J. Pelczar, Jr., et al., *Microbiology*, 5th ed., 1986; J. M. Reiner, *The Organism as an Adaptive Control System*, 1968; H. J. Rogers et al., *Microbial Cell Walls and Membranes*, 1980; M. R. J. Salton, *The Bacterial Cell Wall*, 1964; M. R. J. Salton, *Biochemistry of Bacterial Membranes*, 1975; R. Y. Stanier, E. A. Adelberg, and J. Ingraham, *The Microbial World*, 4th ed., 1976; L. Stryer, *Biochemistry*, 4th ed., 1995; W. A. Volk et al., *Essentials of Medical Microbiology*, 5th ed., 1995; W. A. Volk and M. Wheeler, *Basic Microbiology*, 8th ed., 1997; J. D. Watson, *Molecular Biology of the Gene*, 4th ed., 1987.

Bacterial taxonomy

The classification, nomenclature, and identification of bacteria; sometimes used as a term to indicate the theory of classification. The bacteria are members of the kingdom Prokaryotae, which is defined in terms of the unique structural and biochemical properties of their cells; more specifically, the organization of the deoxyribonucleic acid (DNA) in the nucleus, the lack of a nuclear membrane, the lack of independent membrane-bounded cytoplasmic organelles, the lack

of endocytosis and exocytosis, and the chemical nature of some components of plasma membrane and cell walls. *See* BACTERIA.

Bacteria inhabit a great diversity of ecologic niches. The individual kinds of bacteria within these bacterial populations exhibit a variety of physiologic capabilities and accomplish remarkable chemical transformations. The systematic study of the component bacteria in these populations demands recognition of these properties, and a classification of those bacteria that can be cultivated so that they may be assigned to already named species or be described as a new organism to be classified and named.

Classification involves the recognition of similarities and relationships as a basis for the arrangement of the bacteria into taxonomic groups or taxa. The basic taxon is the species. Identification involves the recognition of a bacterium as a member of one of the established taxa, appropriately named, by the comparison of a number of characters with those in the description. *See* TAXONOMIC CATEGORIES.

Species. A bacterial species is a conceptual entity that is hard to define, despite its role as the basic taxonomic grouping. Bacteriologists accept the imprecision and recognize that a species represents a cluster of clones exhibiting some variations in minor properties. They have developed a formal approach to the description of the taxon while trying to solve the problems encountered in the process of recognizing and naming species.

The description is an assembly of such structural, chemical, physiologic, genetic, and ecologic characteristics as can be determined for the available strains that closely resemble each other. A strain is any pure culture of an organism isolated from nature, and the collected strains may then be conserved as cultures in the laboratory for study and comparison. In addition to the description, one strain must be designated by the author and preserved in a culture collection as a type strain, or permanent example, of the species and available to all who study bacteria. If that type strain is lost or succumbs, a formal proposal of a substitute strain (neotype) must be published. In general, bacterial taxonomy is built around the living type specimen: a species consists of the type strain and, whenever available, all other strains sufficiently similar to the type strain to be considered as included in the species. There is a provision for the description and naming of a distinctive species that is not yet cultivable, with the requirement for a suitably preserved type specimen.

A new species, validly described, must be assigned to a genus in order to accord with the binomial system of nomenclature initiated by C. Linnaeus. Thus, a species assigned to the genus *Bacillus* would be referred to as, for example, *Bacillus subtilis*. Such formal names of taxa are italicized to indicate that they are considered to accord with the formal description. If there is no appropriate genus available, a new genus must be named in accord with the *International Code of Nomenclature of Bacteria* and provided with a description that circumscribes the included species, and a type species must be designated as the exemplary representative of the genus.

The lowest nomenclatural rank that is recognized by the *Code* is subspecies, which is a subdivision of the species recognizing consistent variations in otherwise stable characters in the species description, for example, *Bacillus cereus* ssp. *mycoides*. There are times, however, when even finer but unofficial subdivisions of the species are useful and contribute to science, for example, for the epidemiology of pathogenic species. Then, groups of strains may be recognized by some special character as a variety of the species. These may be based on a biological property (biovar), antigenic variation (serovar), pathogenicity (pathovar), or susceptibility to particular bacterial viruses (phagovar). These characters have no formal standing in nomenclature.

Higher taxa. The rules of nomenclature provide for higher taxa, and many have been created and are in use. However, it must be admitted that the relationships implied by their association in a practical scheme of classification, based on resemblances in characters (phenotype), can be very misleading. It may be better to admit that it is not known where some taxa belong. The result is an avoidance of a complete hierarchical classification because of insufficient data. Instead, recognized taxa are grouped within parts or chapters, using vernacular names as titles to express a sharing of substantial and readily determined characters—for example, "gram-negative facultatively anaerobic rods," or "the gliding bacteria," or "gram-positive cocci." Groupings such as these may be diverse in evolutionary or phylogenetic terms, but they are practical for the purpose of identification. It is necessary to include in virtually all such chapters one or more "genera of uncertain affiliation" that cannot be assigned to any family or higher taxon, but at least accord with the broad definition of the chapter. As more is learned to define the relationships at any level, a stable and useful scientific classification of the higher taxa will develop. It has started with the recognition of the Archaebacteria as being phylogenetically separate from all other bacteria. *See* ARCHAEA; BACTERIA.

A provisional arrangement of higher taxa has been proposed (see below), which utilizes some reliable and paramount characters: the cell wall and its constitution (in part recognized by the gram reaction) or its absence, and photosynthesis or its absence.

Kingdom Prokaryotae
- Division I: Gracilicutes (gram-negative bacteria)
  - Class I: Scotobacteria—Bacteria that do not utilize light energy and are not phylogenetically related to class II
  - Class II: Proteobacteria—Photosynthetic bacteria that do not produce oxygen together with their nonphotosynthetic, phylogenetic relatives
  - Class III: Oxyphotobacteria—Photosynthetic bacteria that produce oxygen, including those prokaryotes known as cyanobacteria (formerly called blue-green algae)

Division II: Firmicutes—Gram-positive bacteria
- Class I: Firmibacteria—Gram-positive bacteria of simple shapes
- Class II: Thallobacteria—Gram-positive, branching, rod-shaped bacteria

Division III: Tenericutes—Bacteria lacking a cell wall
- Class I: Mollicutes—Single class of Tenericutes, the mycoplasmas

Division IV: Mendosicutes—Bacteria with walls of unusual composition
- Class I: Archaebacteria—Bacteria with walls, membrane lipids, and ribosomes of unusual or novel composition (including methanogenic and halophilic bacteria)

It is apparent that, with one exception, shape or form is not used for these ranks. Modern taxonomic studies have shown a number of examples in which morphology is misleading. However, up to now, the gram reaction with the other attendant properties has proved to be relatively reliable in all walled groups, except the class Archaebacteria. Such an arrangement can be used as a bridge between a classification for determinative purposes and a phylogenetic arrangement that will surely arise from current taxonomic research. It is also apparent from the application of nucleic acid sequence data to the determination of relatedness that there are phylogenetic reasons for associating some nonphotosynthetic Gracilicutes with the purple anoxygenic photosynthesizing bacteria. This association defies, so far, a simple phenotypic characterization.

The inclusion of the blue-green algae (the Cyanophyta of the botanical literature) in this taxonomy as cyanobacteria is now less contentious than heretofore. There is general agreement to classify among the bacteria those that can appropriately be activated, described, and typified by bacteriological methods, and to avoid nomenclatural confusion by using whenever possible the names that are also maintained under the botanical code. The blue-green algae are prokaryotic in nuclear organization, have a gram-negative type of cell wall, possess a murein peptidoglycan, and have ribosomes of prokaryotic size and have ribosomal ribonucleic acid (rRNA) that is definitively related to that of the true bacteria. However, many are not yet cultivated and form complex associations or morphological transformations in nature, and are so recognized with preserved specimens as species in the botanical taxonomy. *See* ALGAE.

Each of the classes listed above contains one or more orders and the ranks appropriate to displaying the species included in them. So, for example, a complete taxonomic description of *Neisseria gonorrhoeae*, the cause of the venereal disease gonorrhoea, would be:

Kingdom	Prokaryotae	Prokaryotic
Division	Gracilicutes	Gram-negative
Class	Scotobacteria	Nonphotosynthetic
Order	Not assigned, but grouped with the gram-negative aerobic, rod-shaped, or spherical bacteria	
Family	Neisseriaceae	Included in the family typified by the genus
Genus	*Neisseria*	*Neisseria* Trevisan 1885, 105
Species	*Neisseria gonorrhoea*	*N. gonorrhoea* (Zopf 1885) Trevisan 1885, 106[AL]
	Type strain, American Type Culture, Collection #19424	

The ascription translates to mean it is the species described by W. Zopf in 1885 and amended by V. Trevisan in 1885 on page 106 of his paper, and the name was included in the *Approved Lists*. A name placed in any rank has a similar detailed ascription.

The higher taxa and the major ranking of organisms within them represent the practical groupings referred to above. But new understanding and changes in approach are producing a steadily accelerating change in bacterial taxonomy. Of greatest importance are systematic studies of the molecular constitution of the major biological polymers and their variations among the prokaryotes.

Nature of characters. Bacterial cells do not exhibit a sufficient variety of reliable (that is, stable) shapes for a useful classification or to make the fragmentary fossil record effective for integration into the higher taxa assigned to existent organisms. There is microscopic evidence of their presence and geological evidence of some of their activities for as much as 3.8×10^9 years before present. But all that can be observed are simple shapes; the rest is inference. Many fossil forms have been named, but usually in accord with the rules of the *Botanical Code*.

Reproduction is almost exclusively vegetative (either by binary fission or by budding); life cycles are absent or infrequent and simple in nature; gametes are not formed and conjugation is not a regular function in reproduction. The determinative features useful in the classification of plants and animals are not available. Bacteriologists were then forced to classify on the basis of any character that could be observed: shape, staining properties, motility, growth requirements, substrate utilization, fermentation products, reserve substances, enzymes, toxins, and so on. The number of characters so generated is rather great and, unfortunately for taxonomists, the selection determined for particular organisms was often dictated by the interests of the bacteriologist; for example, organisms of interest to the medical bacteriologist were subjected to tests that differed from those applied to similar organisms studied by plant pathologists, and vice versa. However, it is realized now that the great diversity of applicable tests represents an extensive sampling of the genome. In a sense this is one form of genetic analysis, and it records the expression of available genes, the phenotype. Formal

genetic analysis has been applied to only a few species of bacteria, but in many ways this function is now served by study and comparison of the sequences of nucleotides in the nucleic acids, by recognition of structural similarities in proteins, or by recognition of the functional similarities in metabolic pathways. The results of such comparisons may be interpretable in terms of relationships and phylogeny (evolutionary sequence). These newer approaches must now be discussed, because in them lie the seeds of change. *See* BACTERIAL GENETICS.

Modern approaches. Several new techniques are presently used in modern approaches to taxonomy.

Numerical taxonomy. Numerical taxonomy (taxometrics) is a first approach for the analysis of phenotype. It requires a larger number of tests (say 150) on each of a number of bacterial strains under investigation, together with a collection of type strains and veritable reference strains (preferably not less than 50), both closely and remotely resembling the grouping under study. An analysis of the similarities and dissimilarities requires the comparison of each of the strains (Operational Taxonomic Units or OTUs) with each of the other strains. The data can then be assembled in a similarity table based on the count, for any pair of OTUs, of the number of characters in which they are identical (positive or negative) expressed as a percentage. Other matching coefficients are possible, but this simple matching (S_{SM}) coefficient is the one most often used in bacteriology. Such analyses only became possible with the advent of computers, which allow the table of similarities to be further manipulated to enhance the appreciation of taxonomic structure. The clusters of closely similar OTUs can be assembled in a treelike diagram or dendrogram, in which the branch points represent the S_{SM} that separates each pairing of clusters. Another useful expression of the results is a sorted similarity matrix, in which the clusters are arranged in a two-dimensional format and the similarity is expressed by shading, allowing the relative proportion, size, and separation of the clusters to be recognized by eye. The major separable similarity clusters thus generated are termed phenons, which may or may not be equated with a taxon. But most phenons formed at the 80% similarity level include all or most of the representatives of a species, and so have taxonomic usefulness and form a method for objective analysis of phenotype. It is a useful exercise for the assessment of taxa, and a basis for further and more elaborate statistical analysis. Interpretation, however, requires well-founded scientific judgment despite the overt intention to apply the adansonian principle of not assigning undue weight to any character. In fact, one of the useful returns of taxometric analysis is the formal assessment of the reliability of selected characters for classificatory keys and identification.

Numerical taxonomy implies the existence of programs for computer-assisted identification, either as recognizable phenons or by relation into a computer-stored classification and key program. Either of these methods has found considerable application in dealing with masses of isolates (for example, in studies of pollution or of sediments), or in dealing with results of automated systems for identification of pathogens in clinical bacteriology. The artificial keys and the reference data included in the memory for computer analysis can be infinitely more complex than is practical for the unaided bench worker. Identification can include alternative diagnostic possibilities and a probability estimate. Thus, intuitive judgment in practical bacteriology can be supplemented with artificial keys, and in academic bacteriology with attempts to develop natural or phylogenetic classifications. *See* NUMERICAL TAXONOMY.

Chemotaxonomy. Chemotaxonomy applies systematic data on the molecular architecture of components of the bacterial cell to the solution of taxonomic problems. This has been a powerful tool since the 1950s, and a number of chemotaxonomic markers have been identified, ranging from molecules unique to the Prokaryotae (for example, the murein peptidoglycans) or specific groups of bacteria (for example, the ether-linked lipids of the Archaebacteria) to mechanisms or products of metabolism that characterize genera or species. The availability, and relative simplicity of techniques for amino acid analysis, for sequential analysis of polymers, for gas and thin-layer chromatography, for fermentation products and lipids, and so on, have made the systematic studies possible. These have led to a more effective definition of taxonomic groups based on biochemical assessment of cell wall composition, lipid composition of membranes, the types of isoprenoid quinones, the amino acid sequences of select proteins, and the characterization of proteins, such as the cytochromes and many other macromolecules.

Of all the various chemotaxonomic approaches, the most exemplary is that based on the chemistry of the peptidoglycans of the cell walls. These heteropolymers are unique to the Prokaryotae and are possessed by all taxa except those in the Tenericutes and the Mendosicutes, which have no walls or strangely constructed walls. The great majority of genera are consistent in the amino acids that make up the pentapeptides that are associated with the amino-sugar polymeric backbone, and the simpler peptides that cross-link the strands into a strong network. But some genera and groups, especially among the gram-positive Firmicutes, have peptidoglycan types defined by distinctive diamino acids and amino acids forming their cross-links. These are now found to be consistent with phylogenetic groups recognized on the basis of RNA analysis. *See* CHEMOTAXONOMY.

Nucleic acid studies. These studies have been by far the most potent generators and arbiters of data on relatedness, with distinct capability for applications to the phylogenetic assessment of taxonomic arrangements. *See* DEOXYRIBONUCLEIC ACID (DNA); NUCLEIC ACID; RIBONUCLEIC ACID (RNA).

The DNA base composition, which is most simply expressed as the content of one of the base pairs (as mol % guanine + cytosine), was first applied

to species and genera. It is useful information to this extent: an excessive spread of values for a species or a genus should warn the taxonomist that the group is probably more complex than a single taxon; a strain that is divergent in mol % guanine + cytosine from the type and reference strains is probably misidentified. Conversely, similar values are no measure of identity because the overall composition cannot be related to the sequential arrangement of the nucleotides.

Nucleic acid homology experiments are designed to give a measure of the overall similarity in the sequence of the bases in DNA or RNA from one organism compared to that from another organism. This can be done with genomic DNA because the complementary strands can be dissociated into single strands by an appropriately poised high temperature, and reassociated at a lower temperature with any single strands having appropriate sequences, either from the same organism or a related one. The reassociation is proportional to the overall similarity of the sequences. This is usually measured as a percentage of the uptake of radioactively labeled DNA from the test organisms compared to the uptake of labeled homologous DNA by the reference DNA (taken as 100%). There are several well-established methods which give comparable results. Single-stranded RNA molecules will also hybridize with a complementary DNA strand, and the degree of association can be similarly assessed. Consequently, DNA/DNA association allows the use of the whole genome for assessment of the genetic similarity of two organisms. A much smaller fraction of the genome is assessed by DNA/RNA homology, using either rRNA or transfer RNA (tRNA), because relatively few cistrons determine those RNA sequences. However, evolutionary changes in DNA (and messenger RNA, or mRNA) sequences have been far more rapid than in those of rRNA or tRNA, probably because the latter determine the assembly and function of a crucial and evolutionarily stable organelle, the ribosome, which has recognizable and only somewhat variable elements throughout the kingdom. As a result, DNA/DNA homology is useful only for assessing relatedness within a genus and for establishing or separating species (better than 70% homology) and subspecies. DNA/RNA homology is more useful in the assessment of the relationships of genera because the homologous elements involved have been more highly conserved over that evolutionary period. The experience with these techniques reinforced the realization that exceptionally stable determinants, universally available in all kinds of cells, are needed for the assessment of relatedness encompassing the broad scope of biological evolution.

Biochemical phylogenetic analysis is a thrilling development in bacterial taxonomy, and has been markedly stimulated by the introduction of methods allowing assessment of similarities in the 16S rRNA sequences of diverse taxa (see **illus.**). The 16S RNA is chosen rather than the 5S or the 23S because it is neither too large for analysis nor so small as to have limited information content. The approach has been developed and applied to many bacteria by C. R. Woese and his collaborators. In essence, the radioactively labeled and isolated 16S cRNA is cut into small nucleotides with T_1 ribonuclease. Each of these nucleotide sequences ends in a guanine residue and can be recognized by its radioactivity and position after two-dimensional electrophoresis. The very short pieces (five residues or less) have no useful comparative value, but those of six residues or more can be identified as to sequence and recorded in a catalog of oligomers. Each catalog of oligonucleotides is then compared individually to all other catalogs to provide a similarity coefficient (S_{AB}) for each pair of organisms. The table of S_{AB} values is then subjected to cluster analysis, giving the average linkage S_{AB} value between merged groups, and this can be expressed in the form of a dendrogram (see illus.). The most distantly related have about 50 oligonucleotides larger than hexamer in common (S_{AB} about 0.10), and the closely related have as many as 400 in common. It can be assumed that the differences in the average linkage S_{AB} values approximate to the phylogenetic separation, so that groups of organisms separated by deep clefts (that is, low S_{AB}) in the dendrograms have evolved independently for a long time. The precise relationship to real time is uncertain. The result is that an early divergence of the Archaebacteria from the mainstem of prokaryotic evolution can be identified. A third major line of

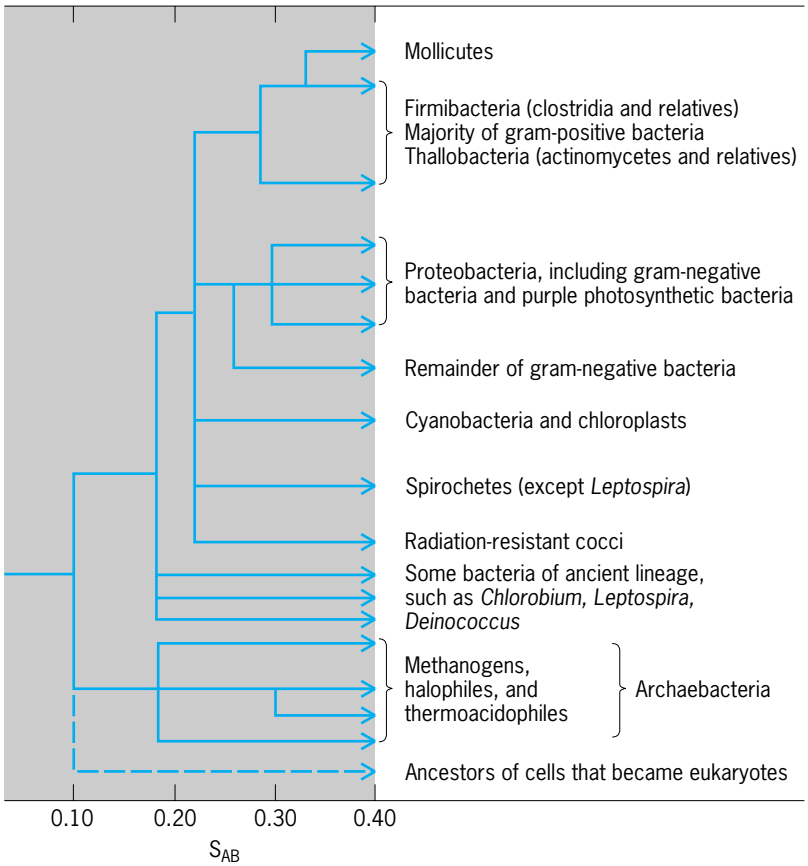

Composite dendrogram showing the major lineages of prokaryotes deduced from 16S rRNA oligonucleotide catalogs.

descent from the ancestral clones would be the as yet unidentified cell types that were later parasitized by chloroplast or mitochondrial prokaryotes and became the ancestors of eukaryotic plants and animals. Woese has called these three lineages "primary kingdoms"; however, others think that this is a premature assignment.

The lineages of the main phylogenetic groupings of the extant true bacteria (that is, all except the Archaebacteria and a few separate and insufficiently studied groups) go back to a major separation expressed as an S_{AB} value of about 0.22. This means that much higher resolution will be required to estimate the exact order of branching of the evolutionary lines. For the moment this is of small consequence; the fascinating and exciting fact is that a phylogenetic order can be established. The technique (albeit somewhat difficult and expensive at this time) is applicable to all the kingdoms, not just the Prokaryotae, and it gives useful data concerning the relationships of genera within groups as well as the genealogy of the groups.

Bacterial taxonomy in transition. Therefore, current taxonomic schemes are departing from the entirely artificial arrangements based on expressed phenotype, and are moving toward the ideal of a natural order expressing evolutionary relationships. This revolution derives from the developments in molecular biology and the availability of computer techniques to generate, systematize, and compare the data on molecular sequences. It depends on recognizing that the evolutionary record is preserved in the sequences of the component molecules of the most stable and most slowly changing complex macromolecules. These sequences are the "semantides" of E. Zuckerkandl and L. Pauling, which they called "the documents of evolutionary history." Few macromolecules can be expected to show the whole story; some of them express subtle systematic shifts of molecular arrangement in some groups of organisms (for example, in the structure of cytochrome *c*). Others show distinctive molecular changes in each of a number of arborizations of a specific major taxon (for example, the peptidoglycan types of the gram-positive bacteria) when other major segments of the classification show no clear distinctions in the same class of macromolecule; other information-laden macromolecules shift too much in sequences and in meaning (for example, chromosomal DNA and mRNA) to be more than informative about that species or genus; and only a few (notably the rRNA molecules) are so stable in evolutionary time that some resemblances can be followed and put into phylogenetic order, whatever the origin. It is fortunate that a large part of the rRNA molecules have to be stable and change only slowly with time in order to preserve the effective and vital organization of ribosomal proteins for effecting protein synthesis from coded nucleotide determinants.

It is now evident that a phylogenetic arrangement of taxa will be accomplished because effective arrangement of many of the broader groups is possible. In fact, the evolutionary resolution will improve, as will the capability (already evident) of the same sequence data to resolve inter- and intrageneric relationships. The integration of the data from 16S rRNA and from 5S rRNA will eventually be joined, as techniques of analysis and data handling improve, by data from the larger and more complex 23S rRNA. Other sources of evolutionary data are coming out of the study of proteins and enzyme pathways, and even if results are not as broad in scope as that from rRNA, they will (and do) provide powerful assistance in the resolution of problems presented by specific taxonomic groups, both small and large. So far, it is encouraging to find a correlation between the different approaches to molecular phylogeny of the bacteria.

Molecular systematics reveals genealogical associations that run counter to the arrangements in the artificial keys based on morphological and physiological characters. For instance, the 16S rRNA catalogs suggest that some species of *Micrococcus* are more closely related to the genus *Arthrobacter* than to other species of *Micrococcus*. Also, there are three well-separated phylogenetic groups of purple photosynthetic bacteria, and each of these shows clear relationships with representatives of a variety of well-known gram-negative genera, some of which are represented in at least two of the groupings. There is the strange possibility (one should remember the prejudice that photosynthesis was an early and unique development) that a great variety of nonphotosynthetic bacteria derived from photosynthetic stock. The resolution of these and other strange associations will require not only confirmatory data using different systems of analysis but the inclusion of many more reference cultures than has been possible up to now. But it is clear that the traditional characters for classification of morphology and physiology present some taxonomic hazards, despite their usefulness in diagnostic bacteriology.

It has to be recognized that a practical classification or a key that enables the identification of an organism may not be (and does not have to be) capable of expressing phylogeny. The essential thing is to be able to arrive at a correct diagnosis. In fact, at this time, the characters (sequences and the like) required for a phylogenetic assignment are too complex and slow in accomplishment for the practicalities of determination. So it is entirely possible that both practical and academic classifications will be needed; and in a way that time has already come. Taxonomy can and must serve all the requirements of bacteriologists and integrate new knowledge to best effect understanding and practical needs. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; BACTERIOLOGY.

R. G. E. Murray

Bibliography. M. J. Carlile, J. F. Collins, and B. E. B. Moseley (eds.), *Molecular and Cellular Aspects of Microbial Evolution*, 1981; P. Gerhardt et al. (eds.), *Methods for General and Molecular Bacteriology*, 1993; J. G. Holt and N. R. Krieg (eds.), *Bergey's Manual of Systematic Bacteriology*, vol. 1, 9th ed., 1993; S. P. Lapage et al. (eds.), *International Code of Nomenclature of Bacteria*, 1975; K. H. Schleifer

and E. Stackebrandt, Molecular systematics of prokaryotes, *Annu. Rev. Microbiol.*, 37:143–187, 1983; V. B. D. Skerman et al. (eds.), *Approved Lists of Bacterial Names*, amended ed., 1989; P. Sneath (ed.), *Bergey's Manual of Systematic Bacteriology*, vol. 2, 1986; J. T. Staley (ed.), *Bergey's Manual of Systematic Bacteriology*, vol. 3, 1989; S. T. Williams (ed.), *Bergey's Manual of Systematic Bacteriology*, vol. 4, 1989; C. R. Woese, Archaebacteria, *Sci. Amer.*, 244:98–122, 1981; E. Zuckerkandl and L. Pauling, Molecules as documents of evolutionary history, *J. Theoret. Biol.*, 8:357–366, 1965.

Bacteriology

The study of bacteria. While bacteria are different in some important respects from all other kinds of cells, their basic processes of physiology and genetics are the same as in all forms of life. One unusual property of bacteria as a whole is their physiological diversity: Some live in the total absence of oxygen and convert complex carbohydrates to acids and alcohols (fermentation), sulfate to hydrogen sulfide, nitrate to nitrogen gas, and hydrogen plus carbon dioxide to methane gas; others carry out photosynthesis by mechanisms nearly identical to plants; some bacteria can grow and multiply by using energy obtained from oxidation of sulfur, ammonia, hydrogen, or iron while obtaining carbon for cell synthesis from carbon dioxide; and some can obtain their needed nitrogen from the gas in air. *See* BACTERIA; BACTERIAL PHYSIOLOGY AND METABOLISM; FERMENTATION; METHANOGENESIS (BACTERIA).

Humans and the planet Earth are profoundly affected by bacteria. There is solid evidence that oxygen first appeared in the air some 2 billion years ago as a result of the activity of bacteria. The bacteria are critically important in the recycling of materials essential to plants and animals. The degradation of complex substances such as carbohydrates, proteins, and lipids to carbon dioxide allows plants to grow. Conversion of ammonia to nitrate and of nitrogen gas to amino acids is also essential to plants. Great numbers of bacteria live on human skin surfaces as well as in the mouth and intestinal tract; most of these are benign or even beneficial. Some bacteria, given the opportunity, can cause severe diseases of humans, other animals, and plants.

Spontaneous generation. The discovery of bacteria was made by Antony van Leeuwenhoek in 1676. As a result of examination of teeth scrapings and of an infusion of pepper (the purpose of the latter was to learn if hotness of pepper was due to something biting the tongue), van Leeuwenhoek saw and accurately described microscopic organisms that are now known to be bacteria. However, scientists did not realize the significance of bacteria for 200 years. The idea that these tiny organisms might be involved in disease occurred to some early scientists, but, curiously, the most intense interest and controversy during this period was over the question of where bacteria came from.

A group of scientists, particularly J. T. Needham and G. L. de Buffon, showed that a boiled infusion of meat would be teeming with bacteria when allowed to stand in a closed container for a day or two. They, and others, concluded that bacteria arose spontaneously from this dead organic matter. In 1776, Lazarro Spallanzani demonstrated that more careful boiling of the infusions followed by immediate hermetic sealing of the vessels produced a broth that remained free of organisms for days and weeks. In spite of this seemingly convincing demonstration, the battle of spontaneous generation in infusions continued for another hundred years. The question was finally settled by the careful and tedious experiments of Louis Pasteur and John Tyndall. They proved conclusively that life begets life and that bacteria do not originate spontaneously from dead organic matter.

It is important to note that the current consensus of scientific thinking is that bacteria were the first forms of cellular life on Earth and that they arose spontaneously as the culmination of a long period of chemical evolution. The important point is that life arose spontaneously in a very special primordial atmosphere over a 1- to 2-billion-year period and not overnight from organic materials such as a heated meat infusion.

Bacteria and disease. One of the most important events in the history of civilization was the discovery that bacteria cause disease. In 1876, Robert Koch provided convincing proof that anthrax, an important disease of cattle and sometimes of humans, was caused by bacteria. Koch had already developed methods for growth and isolation of pure cultures of bacteria. To prove that bacteria cause disease, he first observed rod-shaped bacteria in the blood of anthrax-diseased animals and then isolated the organism as a pure culture from the blood. Next, the bacteria were injected into a healthy animal, and when the symptoms of anthrax appeared, the same kind of bacteria were once again isolated. This series of procedures—observation and isolation of an organism from an animal suffering from a disease, inoculation into a healthy animal with development of the original symptoms, and reisolation of the same kind of organism—became known as Koch's postulates. These have since been accepted as the criteria necessary to prove the etiology of diseases not only caused by bacteria but also by viruses, protozoa, and fungi.

Soon after Koch's monumental discovery, Pasteur, while studying cholera in chickens, discovered the basis for acquired immunity to disease. He found that old cells of the disease-causing bacteria became attenuated, or changed in such a way that when they were inoculated into a healthy chicken the disease did not occur. Moreover, injection of young virulent bacteria no longer caused the disease. The chicken had become immune to cholera.

These discoveries by Koch and Pasteur were the impetus for a subsequent golden era of study of the cause and control of bacterial disease of humans and of immunization as a means to combat diseases. A result of the work in this era was a tripling of the life

expectancy of humans between 1876 and 1926. *See* IMMUNITY; MEDICAL BACTERIOLOGY.

Models for biochemistry and genetics. At the molecular level, the basic physiological processes associated with metabolism and growth are nearly the same for all forms of life. Because of their rapid growth rate (some bacteria divide every 20–30 min), and the ease of obtaining large amounts of material in a controlled atmosphere in a short time, bacteria were favorite subjects for studies of biochemistry and cell physiology. Much of the current knowledge of the properties and activities of enzymes, the mechanism and control of degradation of foodstuffs, and the biosynthesis of proteins, nucleic acids, carbohydrates, and lipids was gained from studies of bacteria. *See* BACTERIAL GROWTH; BIOSYNTHESIS.

The same properties that made bacteria useful for biochemical studies made them ideal subjects for studies of genetics. The ability to obtain population increases from a few hundred to millions of cells overnight and in a test tube made studies of mutation and selection feasible. It was studies of behavior of populations of bacteria that led Oswald Avery, M. McCarty, and C. M. MacLeod to the proof that deoxyribonucleic acid (DNA), and not protein, is the cell macromolecule of which genes are made. *See* GENETICS.

Relationship to other life forms. Bacteria were considered to be primitive plants until approximately 1955, when improved technology of electron microscopy made it possible to observe their intracellular structure. The organisms were embedded in plastic, then cut into slices as thin as 0.1 micrometer. The intracellular structures of bacteria were found to be much different and simpler than all other cells, both plant and animal. Bacteria have no discrete, membrane-bound nucleus and no special intracellular membrane-bound organelles. Animal and plant cells contain organized nuclei and organelles called mitochondria, and plant cells contain chloroplasts. The ribosomes, particulate structures involved in protein synthesis, are smaller in bacteria than are the ribosomes in the cytoplasm of plant and animal cells. From such observations, it became clear that bacteria are not plants but are as different from plants as plants are from animals. C. B. Van Neil and R. Y. Stanier in 1955 formalized these observations and ideas into a classic paper, "The Concept of a Bacterium." The bacteria were to be considered a completely different evolutionary line from other life. Bacteria were called prokaryotic cells, meaning primitive or primordial; all other cells were termed eukaryotic, denoting more advanced or true cells. *See* CELL (BIOLOGY).

Biochemical evolution. The idea is now generally accepted among biologists that the first cells to originate on Earth were prokaryotic. These arose spontaneously some 3–3.5 billion years ago in an environment completely devoid of free oxygen. Geochemical evidence verifies that the Earth was anaerobic (there was no oxygen in the atmosphere) until 2–2.5 billion years ago. The prokaryotic cells went through a period of biochemical evolution for a billion years or so, wherein they adapted by mutation and selection to become increasingly adept at obtaining substrates and energy for growth. It was the evolution of oxygenic photosynthetic bacteria some 2.5 billion years ago that brought about the appearance of free oxygen in the air. Bacteriologists theorize that much of the biochemical diversity of present-day bacteria represents processes that have changed but little for billions of years. Thus, some anaerobic bacteria found in soil, sediments, and intestines where there is no free oxygen are thought to be living fossils of early biochemical processes that occurred before free oxygen existed. *See* ATMOSPHERE, EVOLUTION OF; PREBIOTIC ORGANIC SYNTHESIS.

Bacteria and endosymbiosis. The theory now most widely accepted for explaining the origin of eukaryotic cells is most intriguing. Comparisons of the basic physiological properties and cell structure of eukaryotic and prokaryotic cells continued to show how different the cells were. A surprising result of these studies was the demonstration of how similar bacteria are to the eukaryotic cell organelles, that is, mitochondria and chloroplasts. The ribosomes of the bacteria and of the organelles are the same size, and function virtually identically in protein synthesis. The nuclear genetic material of bacteria and of the organelles was found to be organized in similar fashion. In fact, the genetic material, DNA, of organelles is more similar to DNA of bacteria found in nature than it is to the DNA comprising the central nuclear region in the cytoplasm of the eukaryotic cells.

These observations are the basis for the endosymbiosis theory of evolution of eukaryotic cells. The theory suggests that some prokaryotic cells during the period of the primordial anaerobic atmosphere were predators. These cells became larger and engulfed and digested smaller cells. Eventually, some of the engulfed cells were not digested but instead took up permanent residence in the cytoplasm and supplied energy to the host cell while remaining in this protected environment. The cell line with a symbiotic relationship with respiratory bacteria developed into animals. Another cell line, with a symbiotic relationship with both the respiratory bacteria and oxygenic photosynthetic bacteria, became plants. The purpose of mitochondria in animals and plants is to produce energy by coupling oxidation of food to reduction of oxygen to water, a process called respiration. The function of chloroplasts is photosynthesis—the conversion of solar energy to biologically useful energy with formation of oxygen. If the theory of endosymbiosis is correct, bacteria not only are all around but are an essential and integral part, as mitochondria, of every cell of the human body. *See* CELL PLASTIDS; MITOCHONDRIA.

Archaebacteria. Biologists had barely become comfortable with the concept of three major avenues of cellular evolution (prokaryotic cells and eukaryotic plant and animal cells), when Carl Woese discovered the existence of microorganisms that do not fit into this scheme. In an attempt to develop a phylogenetic history of evolution, Woese focused

on the constituent nucleotide bases of the ribosomal 16-S ribonucleic acid molecule. He reasoned that this molecule is present in all cells and performs the same function for all. Changes, caused by occurrence of mutations, in the kinds and sequences of the nucleotide sequence would be expected to provide a record of genetic relatedness with respect to evolutionary time. Very different organisms should have very different 16-S RNA nucleotide sequences, while similar groups of organisms should have fewer differences.

The initial data verified the theory and reinforced the concept that prokaryotic cells are more ancient than eukaryotic cells. However, some anaerobic methane-forming bacteria were found to have 16-S RNA with a greatly different nucleotide sequence from that of other prokaryotic cells. Woese deduced that these methane-forming bacteria are as different from the prokaryotic bacteria as are prokaryotic cells from eukaryotic plant and animal cells. These different organisms were named archaebacteria and were considered to be a fourth major evolutionary cell line that originated very early.

Some thermophilic bacteria, growing at near-boiling temperatures, and halophilic bacteria, growing at saturated salt concentrations, were also found to be archaebacteria. Physiological studies of members of this group of bacteria have verified their difference from other bacteria. The true bacteria are surrounded by cell walls containing an unusual sugar—muramic acid—and their membranes are composed of ester-linked straight-chain aliphatic fatty acid molecules. Archaebacteria cell walls contain no muramic acid, and their membranes are unique in nature in being made of ether-linked branched aliphatic fatty acids. The current concept of bacteria is that they actually belong to two widely different and evolutionary distinct groups, the eubacteria, or true bacteria, and archaebacteria. Both are considered to be prokaryotic cells. *See* ARCHAEA.

Biotechnology and genetic engineering. Bacteria are the basis of many important industrial processes. They are involved in production of cheese, fermented food products such as sauerkraut, pickles, and sausage, and formation of methane gas from sewage and other wastes. Most medically important antibiotics are produced by bacteria. Exploitation of bacteria for detoxification of environmental pollutants and for production of useful materials promises to increase in the future. New industries are based on biotechnology, where bacteria play a key role. *See* ANTIBIOTIC; FOOD MICROBIOLOGY; INDUSTRIAL MICROBIOLOGY.

The basis for the new biotechnology is genetic engineering. This involves the ability to transfer specific genes from one kind of organism to another. As a consequence, new biological properties are produced, or engineered. It was discovered that many bacteria possess small pieces of DNA called plasmids that exist in the cell cytoplasm and are not associated with the cell chromosome or nucleus. Plasmids generally contain genetic information for a relatively few cell properties. Often, the ability of bacteria to become resistant to antibiotics is due to plasmid genes. The plasmids are capable of being transmitted rapidly between bacterial populations. Plasmids have been successfully isolated from one kind of bacterium and then inserted into a different kind of bacterium. Often, the cell receiving the plasmids acquires properties from the donor cell due to the plasmid-borne genetic information.

In order to genetically engineer a cell, the plasmid molecule is cut by using a specific enzyme called restriction endonuclease. Next, the genetic material of another cell containing information for formation of the desired product is cut into fragments by the endonuclease. The plasmid and gene-fragment molecules are mixed together and treated with an enzyme that joins fragments of DNA together. The result is reformation of the plasmids containing pieces of the genetic information of the cell. The recombinant plasmid is next introduced into a bacterium, usually the one from which it was originally isolated. The bacteria are allowed to grow and multiply, and analyses are made to determine if the desired property is being expressed.

The gene for production of insulin has been transferred in this way from animal cells into the bacterium *Escherichia coli*, and insulin is being produced on a commercial scale by using the bacteria. The amino acids phenylalanine and aspartic acid that constitute the sugar substitute aspartame are produced by genetically engineered bacteria. The ability to transfer specific genes from one kind of organism to another has opened virtually unlimited possibilities for creation of new kinds of bacteria tailored to perform some useful function. It is now possible to consider large-scale production of anticancer and antiviral agents and of hormones that are present in the human body in infinitesimally small amounts on a commercial scale by using genetically engineered bacteria. *See* BACTERIAL GENETICS; GENETIC ENGINEERING.

J. C. Ensign

Bacteriophage

Any of the viruses that infect bacterial cells. They are discrete particles with dimensions from about 20 to about 200 nanometers. A given bacterial virus can infect only one or a few related species of bacteria; these constitute its host range. Bacteriophages consist of two essential components: nucleic acid, in which genetic information is encoded (this may be either ribonucleic acid, in the RNA phages, or deoxyribonucleic acid, in the DNA phages), and a protein coat (capsid), which serves as a protective shell containing the nucleic acid and is involved in the efficiency of infection and the host range of the virus.

The description of a bacterial virus involves a study of its shape and dimensions by electron microscopy (**Fig. 1**), its host range, the serological properties of its capsid, the kind of nucleic acid which it contains,

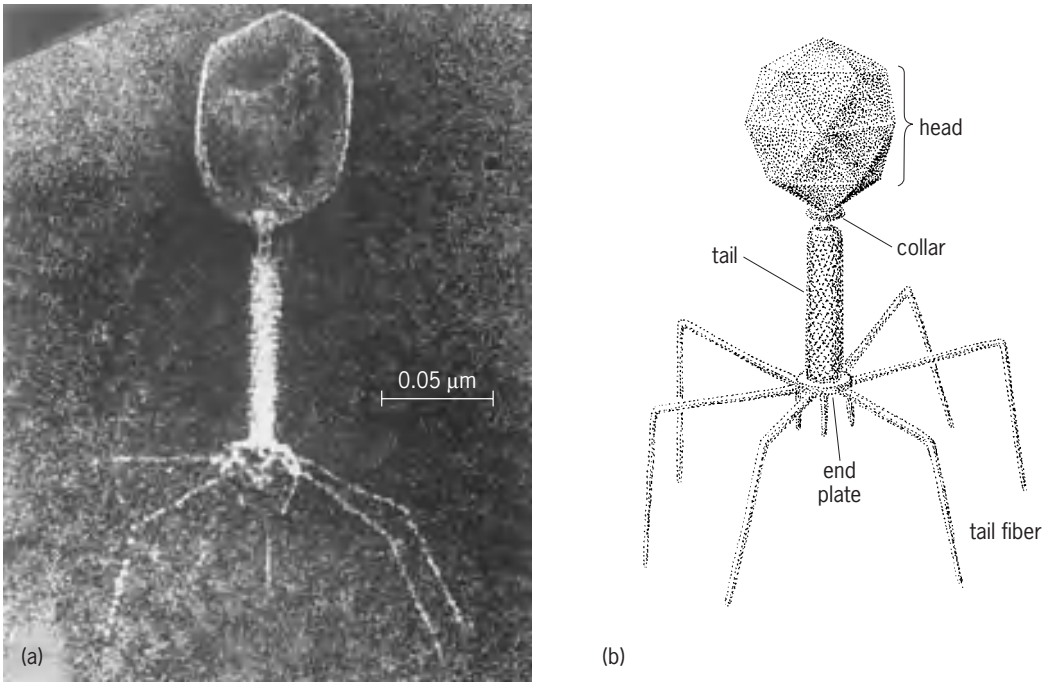

Fig. 1. Bacteriophage structure. (*a*) Electron micrograph of negatively stained T2 bacteriophage showing head and tail components (*courtesy of H. Fernandez-Moran*). (*b*) Diagram of similar T4 bacteriophage.

and the characters of the plaques which it forms on a given host.

Both the nucleic acid and the capsid proteins are specific to the individual virus; in the case of the capsid proteins this specificity is the basis for serological identification of the virus.

The most striking form of phage infection is that in which all of the infected bacteria are destroyed in the process of the formation of new phage particles. This results in the clearing of a turbid liquid culture as the infected cells lyse. When lysis occurs in cells fixed as a lawn of bacteria growing on a solid medium, it produces holes, or areas of clearing, called plaques (**Fig. 2**). These represent colonies of bacteriophage. The size and other properties of the plaque vary with individual viruses and host cells (**Fig. 3**).

The most systematically investigated of the bacteriophages are a series of DNA phages designated T1 through T7, which lyse strains of the gram-negative bacterium *Escherichia coli* and its relatives. These are referred to as coliphages. Certain facts established with these viruses and their hosts have had general application for the study of all viruses. *See* COLIPHAGE.

The effect of phage infection on a population of susceptible host cells is not always the same. RNA phages invariably lyse the cells they infect, with the production of new phage progeny. A few cells in a population may survive; these are found to be mutants resistant to phage infection. DNA phages are of two types. The obligately lytic phages, like the RNA phages, give rise following infection to productive lysis, and surviving bacterial cells prove to be resistant mutants. Other DNA phages, called temperate, are capable of a more complex relationship with their host cells. Infection of a population by these phages results in productive lysis of most of the bacteria, but the few survivors are now of two types: (1) the noninfected resistant mutants and (2) cells in which the phage DNA is integrated in the bacterial

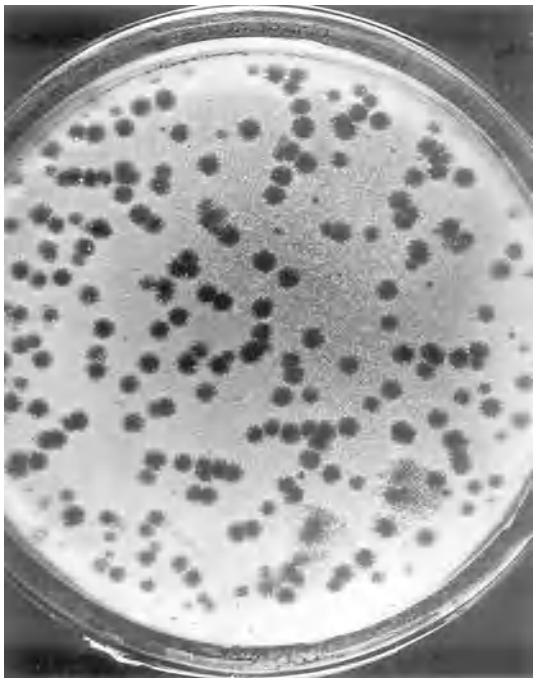

Fig. 2. Plaques of *Rhizobium* bacteriophage; clear, dark areas produced by lysis of infected bacteria. (*From J. Kleczkowska, The production of plaques of Rhizobium bacteriophage on poured plates and its value as a counting method, J. Bacteriol., 50:71–79, 1945*)

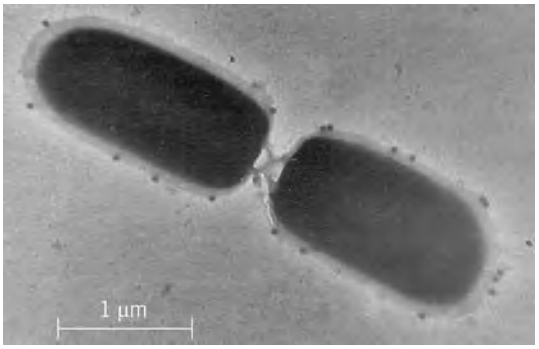

Fig. 3. Dividing *Escherichia coli* cells with 19 particles of coliphage adsorbed to the cell walls. (*From S. E. Luria, M. Delbrück, and T. F. Anderson, Electron microscope studies of bacterial viruses, J. Bacteriol., 46:75,1943*)

chromosome, replicates with it, and endows the bacterial cell with new properties. *See* LYSOGENY.

The lytic cycle involves five steps: (1) adsorption, which involves a receptor site (a specific molecular configuration) on the surface of the host bacterium which reacts with an equally specific taillike structure of the phage particle (Fig. 3); (2) penetration, in which the germinal substance (nucleic acid) of the virus reaches the interior of the bacterium; (3) multiplication, in which the genetic information encoded in the nucleic acid is transcribed and translated to direct the synthetic processes of the cell toward the making of phage components; (4) maturation, involving the assembly of the newly synthesized phage components into mature particles, with nucleic acid enclosed in capsids; (5) lysis of the infected cell with liberation of phage progeny. The specific details of productive lysis differ among the coliphages and other phages as well. The time from infection to lysis with several of the T phages is as short as 13 min. For other phage-host cell systems the latent period is longer; under standard conditions it is characteristic for each system. The average number of new infectious particles produced by *E. coli* infected by a T phage is from 100 to 300 for each infected cell. The average yield for other combinations may be lower or higher: Some corynebacteriophage (*Corynebacterium diphtheriae*) systems produce an average of only 30 virus particles for each infected cell, while *E. coli* infected with some RNA phages may produce more than 10,000 progeny in each host cell. *See* LYTIC INFECTION.

Rare mutants occur among the virus particles produced by phage infection. Mutants of temperate phages, known as virulent, are no longer capable of giving rise to lysogenic bacteria but are obligately lytic. Other mutants are detected by extensions of host range among related bacteria, plaque characteristics, ability to multiply at elevated temperatures, and so on. Related bacteriophages and their mutants are capable of undergoing genetic recombination in mixed infected cells, allowing the construction of genetic maps of a number of bacterial viruses. Studies of the translation of genetic information and its products during the process of phage multiplication have been of inestimable value in establishing the various interactions of DNA, RNA, and enzymes in biosynthetic processes. *See* ACTINOPHAGE; BACTERIAL GENETICS; RECOMBINATION (GENETICS); VIRUS.

Lane Barksdale

Bibliography. E. A. Birge, *Bacterial and Bacteriophage Genetics: An Introduction*, 1983; J. Douglas, *Bacteriophages*, 1975; M. S. DuBow, *Bacteriophage Assembly*, 1981; R. Hendrix et al. (eds.), *Lambda II*, 1983; G. S. Stent, *Molecular Biology of Bacterial Viruses*, 1963.

Badger

A carnivorous mammal in the family Mustelidae along with weasels and otters. Badgers originated in Asia and are currently found from Ireland to Japan and the Philippine Islands as well as in North America. They occupy a wide range of habitats including open plains, semidesert, and the boreal forests of Scandinavia and Russia. In parts of Russia, they occur as far north as the Arctic Circle. Worldwide there are ten species of badgers classified in six genera (see **table**).

American badger. The American badger (*Taxidea taxus*) is found from northern Alberta and southern British Columbia to Ohio, central Mexico, and Baja California. Throughout most of its range, this species prefers open country, living in the prairies and plains where its principal foods—ground squirrels, prairie dogs, and other burrowing animals—abound. The American badger is a heavy-bodied, medium-sized mammal with a broad head; a short, thick neck; short legs; and a short, bushy tail. The ears are low and rounded. Five toes are present on each foot. A prominent nonretractile claw is present on each toe, with the claws on the front feet being very long and exceeding 1 in. in length. The thick fur is rather short on the back but moderately long on the sides, thus giving the badger a characteristic

Names and geographic distribution of badgers

Name	Range
Mellivora capensis (Honey badger, or ratel)	Africa, India
Meles meles (Old World badger)	Europe, Asia
Arctonyx collaris (Hog badger)	Southern Asia, Sumatra
Mydaus javanensis (Malay stink badger, teledu)	Malay Archipelago
Mydaus marchei (Palawan stink badger)	Philippines
Taxidea taxus (American badger)	North America
Melogale moschata (Chinese ferret badger)	China
Melogale personata (Indian ferret badger)	Nepal, Indochina
Melogale orientalis (Oriental ferret badger)	Java, Bali
Melogale everetti (Everett's ferret badger)	Borneo

American badger (*Taxidea taxus*). (*Photo by Gerald and Buff Corsi; © 1999 California Academy of Sciences*)

squat appearance (see **illustration**). The general coloration is grayish to reddish with a slight yellowish cast. The brownish face is marked with a white stripe reaching from near the nose to the crown of the head and sometimes onto the neck and back. Paired white areas extend from around the mouth onto the cheeks and insides of the ears, and a prominent, vertical black bar, or "badge," occurs in front of each ear. The backs of the ears and the feet are black; the tail is yellowish-brown; and the underparts are predominately yellowish-brown. Both sexes are colored alike. The dental formula is I 3/3, C 1/1, PM 3/3, and M 1/2 × 2 for a total of 34 teeth. Adults have a head and body length between 420 and 720 mm (16–28 in.), a tail length of 100–155 mm (4–6 in.), and a weight of 4–12 kg (9–26 lb). Males are heavier than females. *See* DENTITION.

Badgers have two pairs of scent glands, one on the belly and the other pair near the anus. The anal scent glands secrete a strong, sometimes pungent, odor and are probably used mostly during the breeding season for sexual attraction and at other times for defense. Badgers hiss, grunt, growl, and snarl when fighting or cornered by an aggressor. They can walk and run, attaining a topmost speed of 10 to 15 miles per hour.

Badgers are excellent diggers, as evidenced by their heavy body, powerful muscles, strong front feet, and long claws. They can dig at a faster rate than a human can dig with a shovel. When digging, badgers loosen the soil with their front feet, pass the dirt under their belly, and kick it out of the hole with their hind feet. Sometimes they use their mouth in digging. Some ferret badgers are good climbers and may sleep in trees.

Reproduction and development. Breeding in American badgers generally occurs in August or September. Several days following fertilization, the embryos become dormant in the uterus for several months before implantation occurs. This process of delayed implantation is common in several other members of the family Mustelidae. In February or March, the embryos become implanted in the uterus and complete their development in about 5 weeks. Thus, while the entire gestation period may last from 6 to 9 months, the period of actual development is completed in about 6 weeks. A single annual litter consisting of 1 to 7 young (usually 2 or 3) is born in late spring. At birth, the altricial young are furred and blind. Their eyes open between 4 and 6 weeks of age. Weaning occurs when the young are about half-grown. The young stay in the vicinity of their home burrow until fall before dispersing.

Ecology. Most badgers are solitary mammals coming together only to mate. European badgers, however, live in groups and are territorial. American badgers usually spend their entire lives within a home range of 1 or 2 square miles and normally do not go far from some hole into which they can make a hasty retreat if danger threatens. They are active mostly at night but sometimes forage in early morning or late evening. They often sunbathe near the entrance of their burrows.

Badgers are active during all seasons, although in more northern portions of their range they may spend considerable time sleeping in their burrows during the colder winter months. During most of the year, the badger's home is a shallow burrow (known as a sett) about 1 ft in diameter which it digs in its daily quest for food. Ground squirrels are a staple food of the American badger during the warmer months of the year, but they also consume mice, rabbits, insects, moles, lizards, snakes, turtle eggs, ground-nesting birds and their eggs, snails, and fish. Both live animals and carrion may be consumed. At times, American badgers and coyotes may forage together for their mutual benefit. During the badger's digging operations, a rodent may escape only to be pounced upon by the agile coyote. Hawks have also been observed following a badger, presumably on the alert for some escaping rodent. While the American badger is a specialist predator, the other nine species are generalist omnivores feeding on a variety of insects and invertebrates as well as small vertebrates, cereals, and tubers.

Indiscriminate poisoning and destruction of their rodent prey has caused a marked decline in American badgers. Dogs and coyotes also prey on badgers. Badgers are a valuable part of the ecosystem because they serve as a control on rodent populations, and their digging aerates and mixes the soil. *See* CARNIVORA; HIBERNATION AND ESTIVATION; MUSTELIDAE.

Donald W. Linzey

Bibliography. *Grzimek's Encyclopedia of Mammals*, McGraw-Hill, 1990; D. Macdonald, *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Balance

An instrument for weighing. The word "balance," derived from the Latin *bilanx* ("having two pans"), is often used interchangeably with "scale" or "scales" (Old English, meaning dishes or plates). However, balance is the preferred term for an instrument used for the precise measurement of small weights or masses in amounts ranging from micrograms up to a few kilograms. *See* MASS; WEIGHT.

Classification. Balances are differentiated according to design, weighing principle, and metrological criteria (see **table**). For a given weighing task, a balance is selected primarily for its maximum weighing load (Max) and for the finest graduation or division (d) of its weight-reading device (scale dial, digital display, readout).

The jewelry trade uses balances graduated in carat (1 carat = 200 mg), troy ounce (1 troy ounce = 1.1 avoirdupois ounce), and pennyweight (1 pennyweight = 1.555 g) for weighing diamonds and precious metals. Except for the special weight units, these instruments are identical with macroanalytical and precision balances.

Balances can be roughly differentiated from scales by their resolution or number of scale divisions, $n = \text{Max}/d$. Balances typically have a resolution of more than 10,000 divisions, and scales for the most part have less.

Traditional mechanical balances. In its classic form a balance consists of a symmetric lever called a balance beam, two pans suspended from its ends, and a pivotal axis (fulcrum) at its center (**Fig. 1**). The object to be weighed is placed on one pan, whereupon the balance is brought into equilibrium by placing the required amount of weights on the opposite pan.

Classification of balances

Type	Division (d)	Typical capacity (max.)
Ultramicroanalytical	0.1 μg	3 g
Microanalytical	1 μg	3 g
Semimicroanalytical	0.01 mg	30 g
Macroanalytical	0.1 mg	160 g
Precision	$\geq$1 mg	160 g–60 kg

Thus the weight of an object is defined as the amount represented by the calibrated standard masses that will exactly counterbalance the object on a classic equal-arm balance. Although this is not self-evident with modern balances and scales, the measurement of weight continues to be based on this original understanding.

In the design of an equal-arm balance (**Fig. 2***a*), it is critical that the pan suspension end pivots (E) be equidistant from, and in a straight line with, the center fulcrum (F). A rigid, truss-shaped construction of the beam minimizes the amount of bending when the pans are loaded. The center of gravity (C) of the beam is located a fraction of an inch below the center fulcrum, which gives the balance the properties of a physical pendulum, letting the balance swing about its equilibrium position. With a slight difference in pan loads, the balance will come to rest at an inclined position, the angle of inclination being proportionate to the load differential (Fig. 2*b*). By reading the pointer position on a graduated angular scale, it is possible to determine differential amounts of weight in between the discrete step values of the weight set.

In so-called trip balances, the beam has its center of gravity above the fulcrum. This type of balance

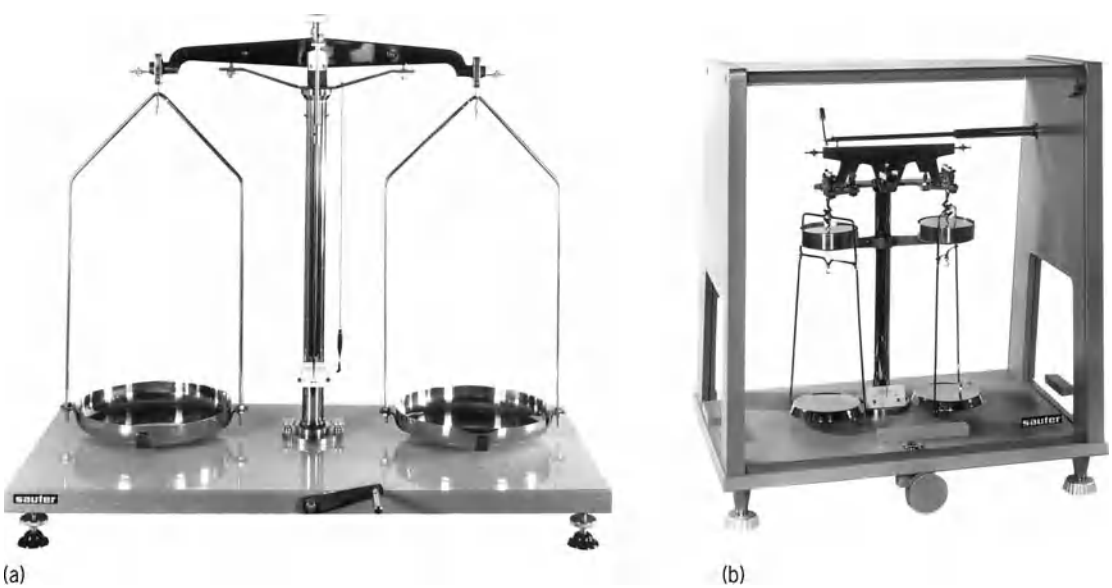

Fig. 1. Mechanical balances. (*a*) Traditional type: crank at center of base plate operates arrestment device for protection of knife-edge pivots, and thumbscrews at ends and center top of beam serve to adjust horizontal equilibrium position and vertical location of center of gravity of balance beam. (*b*) Mechanical analytical balance with several refinements: enclosure to protect balance from air draft disturbance; remote-operated rider poise on top of beam; air dashpots for oscillation damping. (*Mettler Instrument Corp.*)

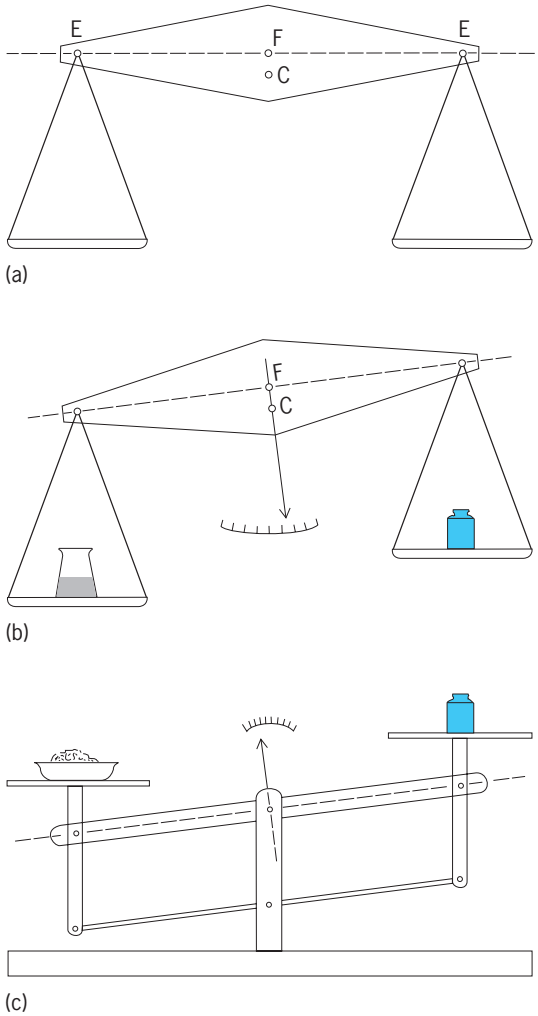

Fig. 2. Balance design considerations. (*a*) Critical design aspects for an equal-arm balance. (*b*) Weighing small weight differentials with an equal-arm balance. (*c*) Top-loading equal-arm balance.

trips to the side of the heavier load, but will not seek or maintain a horizontal equilibrium position with equal loads.

Notable variations and refinements of the mechanical balance include knife-edge bearings of hardened steel, agate, or synthetic sapphire; air damping or eddy-current (magnetic) damping of beam oscillations; sliding-weight poises or riders; built-in weight sets operated by dial knobs; microscope or microprojector reading of the angle of beam inclination; arrestment devices to disengage and protect pivots; and pan brakes to stop the swing of the balance pans.

Top-loading balances (Fig. 2*c*) are based on the same principles and equipped with many of the same features and refinements. With the pans stabilized by lever linkages [so-called Roberval (parallelogram) linkages] they remain horizontal as the pans move up and down. Thus, top-loading balances are more convenient to operate and give unobstructed access for tall or wide weighing loads.

Single-pan substitution balance. First described by the French physicist Jean Charles de Borda (1733–1799), and introduced in industrially manufactured balances by Erhard Mettler in 1946, the substitution principle represented the conclusive step in the evolution of the mechanical balance. Substitution balances (**Fig. 3***a* and *c*) have only one hanger assembly, incorporating both the load pan and a built-in set of weights on a holding rack. The hanger assembly is balanced by a counterpoise which is rigidly connected to the other side of the beam.

The weight of an object is determined by lifting weights off the holding rack until the balance returns to an equilibrium position within its angular, differential weighing range. Small increments of weight in between the discrete dial weight steps are read from the projected screen image of a graduated optical reticle which is rigidly connected to the balance beam.

With only two pivot axes, this design disposes of the problem of keeping pivots in a straight line with, and equidistant from, the fulcrum. Furthermore, the elastic deformation of the beam stays unchanged as the loads on the beam remain essentially constant.

Substitution balances were built as analytical balances with a hanging pan (Fig. 3*b*) and as top-loading precision balances with a stabilized pan (Fig. 3*d*). With minor exceptions, the technical development and production of substitution balances ended because electronic balances have superior accuracy and operating convenience. However, substitution balances will remain in widespread use in laboratories until replacement by new models.

Electronic balances. The evolution of electronic (more accurately, electromechanical) balances (**Fig. 4**) started in the late 1960s and has extended over several generations of electronic technology. Among a number of technical possibilities, one operating principle, electromagnetic force compensation, emerged early as the standard in high-precision weighing. First described by K. Ångström in 1895, the principle of electromagnetic force compensation became feasible for technical application as a result of the advancements in solid-state electronic components. Besides improved accuracy, reliability, and speed of operation, the main benefits from this technology are human-engineered design for optimized interaction between operator and instrument, and numerous operating conveniences such as push-button zero setting, automatic calibration, built-in computing capabilities for frequently used work procedures, and data output to printers and computers.

In every electromechanical weighing system, there are three basic functions (**Fig. 5**): (1) The load-transfer mechanism, composed of the weighing platform or pan, levers, and guides, receives the weighing load on the pan as a randomly distributed pressure force (P) and translates it into a measureable single force (F). (2) The electromechanical force transducer, often called load cell, converts the mechanical input force into an electrical output, for example, voltage, current, or frequency. (3) The electronic signal-processing part of the balance receives

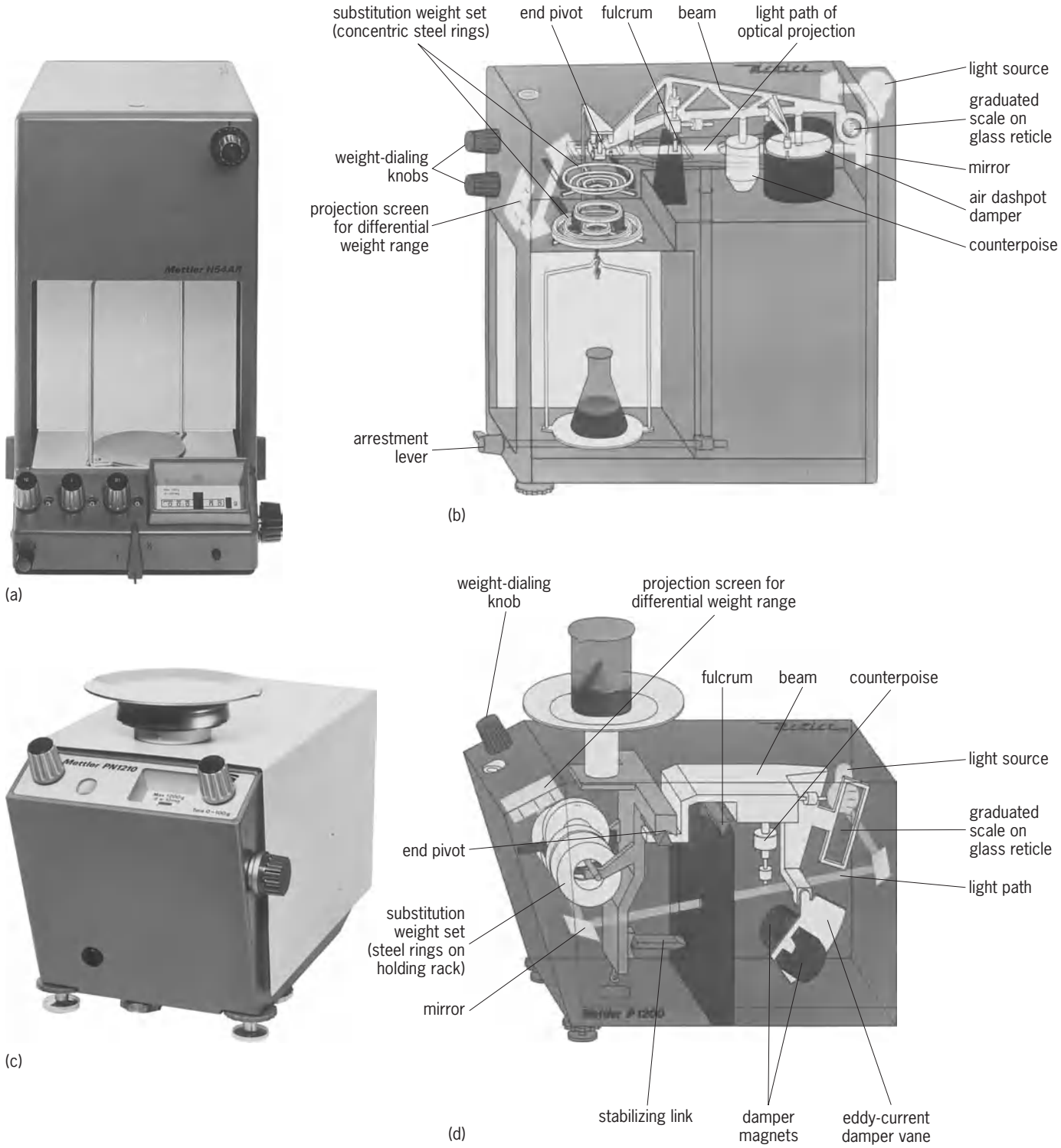

Fig. 3. Substitution balances. (*a*) The standard in analytical laboratory weighing from about 1950 to 1980. (*b*) Cutaway side view of a substitution analytical balance. (*c*) A top-loading substitution balance, readable to 1 mg or coarser. (*d*) Cutaway side view of a top-loading substitution balance. (*Mettler Instrument Corp.*)

the output signal, converts it to numbers, performs computation, and displays the final weight data on the readout.

Load-transfer mechanism. In electronic balances load-transfer mechanisms are lever linkages similar to those used in mechanical substitution balances. As a refinement, the traditional knife-edge pivots have been replaced by elastic flexure pivots (**Fig. 6**). Levers and other parts that rotate relative to each other are connected by flexibly bending strips or bands of high-grade spring metal. Flexure pivots for balances were first described by Wilhelm Weber in 1841, but the breakthrough in their application came with the development of electronic scales and

balances. Pivot points in precision mechanisms are more accurately defined and maintained by flexure pivots than was previously possible with knife edges. As a result, precision balances can be built in an ergonomically correct, low-profile form, and all balances, except for the microgram range, use the top-loading pan arrangement, the pan being nonpendulous and supported and stabilized by flexure-pivoted levers.

Electromechanical force transducer. In a laboratory balance the transducer in many cases resembles the configuration of an electrodynamic loudspeaker. A sound-modulated current in the speaker coil interacts with the magnetic field in the circular gap of the permanent magnet, generating an oscillating force on the coil and voice cone assembly and thus producing sound waves.

Fig. 4. Electronic balances. (*a*) Analytical balance. The weighing pan is nonpendulous and stabilized by a parallelogram linkage. Operating functions are controlled by single touch bar (integrated in readout panel) and include full-scale tare, automatic calibration, selection of weighing range, averaging period, and stability detector setting. (*b*) Electronic precision balance. Typically, balances reading to 1 mg or coarser require no enclosed weighing compartment, as they are less affected by air drafts than analytical balances. Operating features are similar to those of the analytical balance. The balance is operated through control bar along the front. (*Mettler Instrument Corp.*)

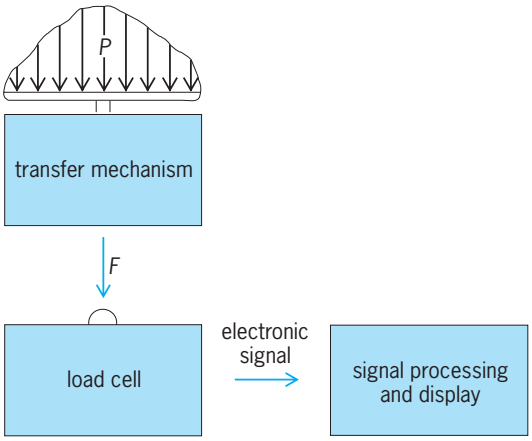

Fig. 5. Three-part structure of an electronic balance.

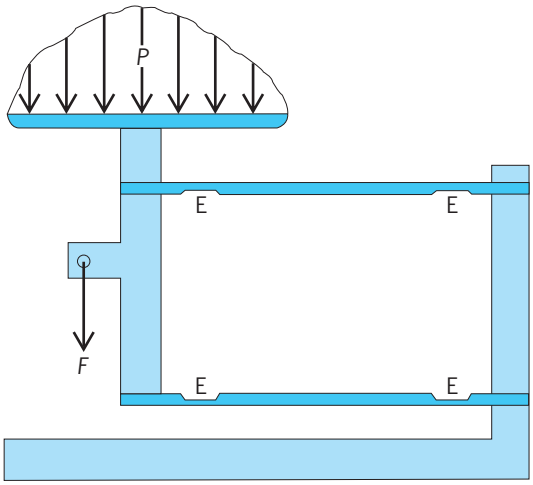

Fig. 6. Diagram of the transfer mechanism of electronic balance. Random force action of weighing load (*P*) is converted into a single force (*F*). Platform is stabilized by flexure-pivoted guides. Pivots are formed by the elastically flexible sections (*E*) in the horizontal guide members (called check links).

In the electromagnetic force compensator (**Fig. 7**), a unidirectional current generates a static force on the coil which, in turn, counterbalances the weight force from the object on the balance pan, usually aided by one or more force-reduction levers. The amount of coil current is controlled by a closed-loop servo circuit which monitors the vertical deflections of the pan support through a photosensor and adjusts the coil current as required to maintain equilibrium between weighing load and compensation force.

The interactive force (F) between a magnetic field and a current-carrying conductor (at a right angle to each other) is determined by the equation below,

$$F = IlB$$

whereby I represents the current, l is the total length of coil wire, and B is the magnetic flux density in the air gap. The force F points in the third direction in space, that is, perpendicular to the directions of current and magnetic field. The force F, and with it the weight of the object on the balance pan, now results from a measure of the current I, for example, by measuring the output voltage V across a precision resistor R wired in series to the force coil, and using the relationship $V = IR$.

Electronic signal processing. This part of a balance can be thought of as a voltmeter whose dial face or digital indicator reads in units of weight. However, the accurate resolution of the measuring range into several million parts, for example, 30 to 0.00001 g, exceeds by far the capabilities of all but the most sophisticated voltmeters. Signal processing in balances involves the most advanced methods of electrical measurement, as well as special computation routines that are applied to each weight value before it is displayed. There are five basic computations: (1) Tare subtraction; subtracts container weight or other preload from each weight value. (2) Automatic

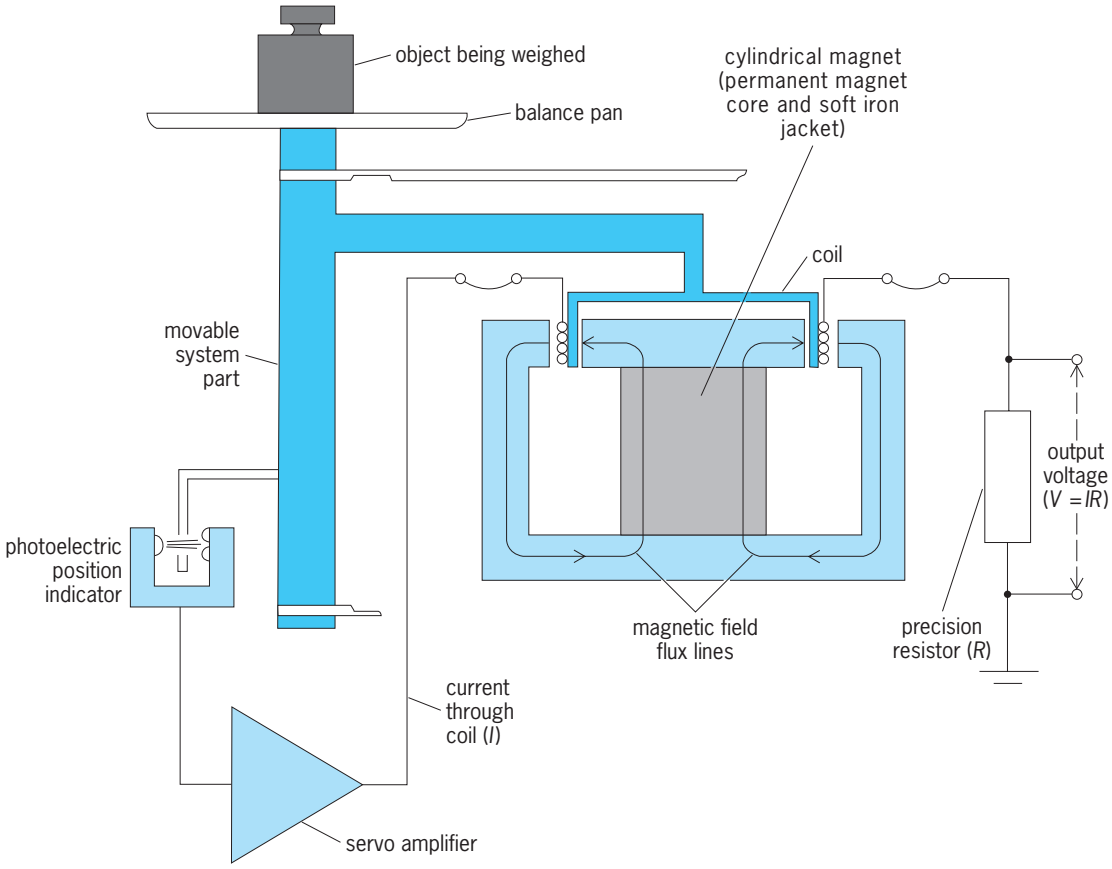

Fig. 7. Diagram of the electromagnetic force compensation principle. The output voltage is proportionate to the weight on the pan.

zero correction; compensates for tendencies of the measuring system to gradually wander off a set zero, even by a fraction of a display division. (3) Vibration filter; applies averaging or other noise-reduction techniques to measurement values in order to produce stable weighing results in the presence of building vibrations or other disturbances, for example in weighing live animals. (4) Stability detector; compares consecutive weighing results with each other, and blocks data output and signals a warning light when weighing results are transient or fluctuating. (5) Automatic calibration; as the operator presses a key and deposits a built-in or external mass standard on the balance, the instrument calibration factor is recalculated and stored in a nonvolatile (power-independent) memory.

In addition to these signal-processing features, state-of-the-art balances offer a number of useful conveniences either built in or available as options on some models, such as display in nonmetric weight units, parts counting, weight statistics, percentage calculation, or recall of net total.

Accuracy. As with all scientific measurements, weighing results should not be taken as unconditionally accurate. Manufacturers' specifications indicate tolerance limits for the most common instrument errors, that is, imprecision, nonlinearity, and temperature drift. However, the single largest error in determining absolute amounts of mass lies outside the balance. It is made when the effects of air buoyancy on the weighing object are neglected. For example, when weighing 1 liter (1.06 qt) of water at sea level, an amount of 1.05 g (0.037 oz) should be added to the net weight as read from the balance in order to obtain the mass of the water. This represents the difference between the buoyancy forces acting on a kilogram (2.2 lb) of water on the one hand, and a kilgram of steel (by which the balance was calibrated) on the other. While the buoyancy correction is necessary in the absolute determination of mass, it is of no concern in any other applications where weight, in the traditional sense, is taken as a direct measure for quantities of any weighable matter. Errors entering into experiments from other sources, for example, volume vessels, are generally much larger, making balance errors negligible in comparison. *See* PHYSICAL MEASUREMENT.

Walter E. Kupper

Bibliography. G. D. Christian, *Analytical Chemistry*, 5th ed., 1993; B. Kisch, *Scales and Weights: A Historical Outline*, 1965; R. O. Leonard, Electronic laboratory balances, *Anal. Chem.*, 48(11):879A–894A, September 1976; P. E. Pontius, *Mass and Mass Values*, Nat. Bur. Stand. Monogr. 133, 1974.

Ball-and-race-type pulverizer

A grinding machine in which balls rotate under pressure to crush materials, such as coal, to a fine consistency. The material is usually fed through a chute to

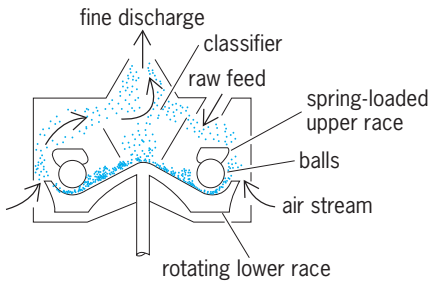

Coarse raw material is ground by crushing and attrition between balls and races and is then withdrawn from the pulverizer by an airstream.

the inside of a ring of closely spaced balls. In most designs the upper spring-loaded race applies pressure to the balls, and the lower race rotates and grinds the coarse material between it and the balls (see **illus.**). The finely ground material discharges along the outer periphery of the ball races. For the pulverization of coal, hot air, introduced between the lower race and the pulverizer housing, lifts or carries the fines to a cyclone classifier at the center of the pulverizer. There the finer particles discharge from the pulverizer while the larger particles return to the grinding zone for further reduction in size. Two or more rings of balls can be cascaded in one machine to obtain greater capacity or output. Counterrotating top and bottom rings also are used to increase pulverizer capacity. Such pulverizers are compact and the power required per ton of material ground is relatively low. *See* CRUSHING AND PULVERIZING. George W. Kessler

Ballast resistor

A resistor that has the property of increasing in resistance as current flowing through it increases, and decreasing in resistance as current decreases. Therefore the ballast resistor tends to maintain a constant current flowing through it, despite variations in applied voltage or changes in the rest of the circuit. The ballast resistor acts as a variable load on the system; therefore it differs from load resistors, which have a constant resistance. Ballast resistors are now mainly of historical interest; electronic devices fulfilling the same function have replaced them. *See* RESISTOR.

The ballast action is obtained by using resistive material that increases in resistance as temperature increases. Any increase in current then causes an increase in temperature, which results in an increase in resistance and reduces the current. Ballast resistors may be wire-wound resistors. Other types, also called ballast tubes, are usually mounted in an evacuated envelope to reduce heat radiation.

Ballast resistors have been used to compensate for variations in line voltage, as in some automotive ignition systems, or to compensate for negative volt-ampere characteristics of other devices, such as fluorescent lamps and other vapor lamps. *See* FLUORESCENT LAMP; VAPOR LAMP; VOLTAGE REGULATOR. Donald L. Anglin

Ballistic missile

A weapon that consists of integral rocket propulsion, means of pointing or guiding the weapon's velocity vector to a prescribed orientation at the position and time of rocket engine shutoff or burnout, and a warhead. In certain applications, means of deploying multiple warheads or submunitions may be incorporated. Ballistic missiles are conceptually simple weapons whose implementation becomes more complex with increasing accuracy, range, and defense penetration requirements.

Trajectory. The term ballistic means that part or most of the missile's trajectory is not subject to propulsion or control. In its ballistic phase of flight, a missile's motion is affected only by gravitation and uncontrolled aerodynamic interactions with the atmosphere.

The ballistic missile follows an elliptical path due to action of the Earth's gravitational field (see **illus.**). The vertical component of velocity after propulsion burnout decreases as kinetic energy is converted to potential energy through the action of gravity, and minor velocity losses occur due to momentum decrease in the presence of aerodynamic resistance. If both the burnout velocity and burnout altitude are large, then an upwardly slanted flight path will cause the missile's trajectory to rise high above the sensible atmosphere, thereby eliminating the retarding and disturbing influences of the Earth's atmosphere for most of the trajectory. When all of the vertical component of burnout kinetic energy has been converted to potential energy, a condition of minimum speed and maximum altitude associated with the apogee of the trajectory is achieved, and the missile begins to descend while transforming potential energy back into kinetic energy as the downward component of velocity increases because of gravity. If propulsion burnout occurs in the upper reaches of the atmosphere, then the maximum speed of the missile upon descent will be about the same as achieved at propulsion burnout because of conservation of energy in the absence of atmospheric retardation. *See* BALLISTICS; CELESTIAL MECHANICS.

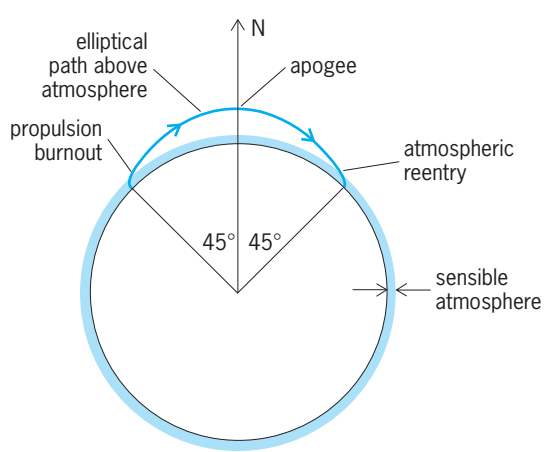

Typical ICBM trajectory superimposed on Earth's circumference.

Propulsion. The purpose of the propulsive function is to impart velocity to the missile so as to achieve an energy state at propulsion burnout that is sufficient for the missile to coast to its target under the influence of gravitation and aerodynamic resistance. Multiple stages of propulsion may be employed to maximize burnout velocity. Each stage of a ballistic missile is an autonomous propulsion unit consisting of propellants, propellant tanks and liquid propellant feed equipment (or casing if solid propellants are used), one or more exhaust nozzles, and means of controlling the direction of the thrust vector. Each stage is operated serially, and when the propellant of a given stage has been consumed the entire empty stage is jettisoned to reduce the system mass that must be accelerated by subsequent stages, and to enable ignition of the subsequent stage. Each stage is successively smaller in propellant mass and thrust magnitude, consistent with the reduced mass each step of the way. *See* ROCKET PROPULSION.

Guidance and control. All ballistic missiles incorporate means of pointing or guiding their velocity vectors so that their trajectories end coincidentally with the intended target. The simplest instance involves launching a missile from a guide rail or tube which the weapon operator points toward the intended target and upward at predetermined elevation angle that will result in a missile impact in the target area. The Army's Multiple Rocket Launching System (MRLS) is a good example of simple pointing as the means of initial-conditions guidance. The most complex guidance schemes are employed with intercontinental ballistic missiles (ICBMs) wherein propulsion durations of several minutes are typical and precise pointing of the velocity vector at propulsion burnout is essential to achieve the desired accuracy in hitting the target area after intercontinental flight. ICBMs and theater nuclear weapons typically employ control over the direction of thrust from their rocket motors to change the orientation of the missile in response to guidance commands. Guidance is based upon the principles of inertial navigation, wherein precise measurements of the missile's acceleration and rotation rate are used to very accurately compute its velocity, orientation, and position. The resulting information is used to apply precise corrections to the missile flight path so as to assure impact very close to its target, regardless of disturbances from the atmosphere or tolerable anomalies in propulsion performance. Control over the burnout velocity may be achieved by varying the thrust or terminating it when the desired velocity is achieved. Additional measurements, such as observing the positions of stars, or other known references, may be made to supplement inertial navigation. *See* GUIDANCE SYSTEMS; INERTIAL GUIDANCE SYSTEM; STAR TRACKER.

Warhead. The warhead of a ballistic missile consists of a single or multiple explosive devices designed to destroy or neutralize the usefulness of the intended target. The explosive devices may consist of conventional chemical explosives with fragmentation material, or nuclear or thermonuclear materials and equipment. When the warheads of ballistic missiles are released above the atmosphere at nearly orbital (3–4 mi/s or 5–6 km/s) speeds with flight path inclinations in the 17–30° range, their trajectories arch high above the sensible atmosphere, so that the vehicles that carry the warheads must be specially designed to protect them from high aerodynamic heating loads upon reentry into the atmosphere at such high speeds.

Modern ICBMs employ multiple reentry vehicles carrying thermonuclear warheads. Each of these vehicles is spin-stabilized and released from a so-called bus vehicle with an added velocity and resultant flight path angle so that each impacts a separate target within a specified region of the Earth's surface. When reentry vehicles have accomplished reentry and descend to altitudes where aerodynamic maneuvering is efficient for changing their flight paths, such maneuvers can be applied as a means of penetrating defenses or of reducing the time to impact. In some designs, on-board means of locating specific targets and guiding the reentry vehicles to them are employed. *See* ATMOSPHERIC ENTRY; ATOMIC BOMB; HYDROGEN BOMB.

Types. Ballistic missiles are either land-based or sea-based.

Land-based missiles. Land-based versions are commonly categorized according to the distance they can fly. *See* ARMY ARMAMENT.

1. Battlefield ballistic missiles can hit targets from 12 to 300 mi (20 to 500 km) from the launch point, and generally employ conventional (nonnuclear) or submunition warheads.

2. Intermediate-range ballistic missiles (IRBM), which are sometimes referred to as intermediate nuclear forces (INF) or theater nuclear weapons, come in a variety of sizes, and can hit targets 300 to 3000 mi (500 to 5000 km) from the launch site. The 3000-mi (5000-km) maximum range is determined by the geometry and extent of Europe, significantly exceeding the distance from London to Moscow. These missiles invariably carry one to three nuclear warheads, and modern versions emphasize vehicle-borne mobility.

3. ICBMs have flyout ranges between 5500 and 7500 mi (9000 and 12,000 km), the larger range being all that is necessary to reach the southern border of China from the northernmost United States, encompassing the former Soviet Union. ICBM survivability in nuclear war is a major issue in determination of basing; the viable options range from so-called hardened underground emplacements to measures that exploit mobility and deceit. Modern ICBMs carry from one to ten nuclear warheads in reentry vehicles that are independently targetable (multiple independently targeted reentry vehicles, or MIRVs).

Sea-based missiles. Sea-based ballistic missiles are invariably based on submarines (submarine-launched ballistic missiles, or SLBMs), providing considerable uncertainty as to their location as an important element of survivability. The flyout range of SLBMs has systematically increased from a few thousand

kilometers to more than 5500 mi (9000 km), equaling the capability of ICBMs. Nuclear-powered submarines are employed to maximize duration of deployment without surfacing. SLBM deployments of the early 1960s required submarines to approach coastal regions nearest their targets to reach interior targets. However, SLBM range improvements of the late 1970s and early 1980s allow deployments relatively close to the home ports of the submarines. SLBMs have employed MIRV warheads since the early 1970s, and accuracy enhancements of the 1980s provide SLBMs with effectiveness levels comparable to those of their land-based counterparts. *See* GUIDED MISSILE; MISSILE; NAVAL ARMAMENT; SUBMARINE.

M. Michael Briggs

Bibliography. R. H. Battin, *An Introduction to the Mathematics and Methods of Astrodynamics*, 1987; B. G. Levi, M. Sakitt, and A. Hobson (eds.), *The Future of Land-Based Strategic Missiles*, 1989; P. G. Savage, *Introduction to Strapdown Navigation Systems*, 1985; G. Siouris, *Aerospace Avionic Systems*, 1993.

Ballistics

That branch of applied physics which deals with the motion of projectiles and the conditions governing that motion. Commonly called the science of shooting, it is, for practical purposes, subdivided into exterior and interior ballistics.

Exterior ballistics begins at the instant the projectile leaves the muzzle of the gun barrel or the instant after impact; interior ballistics, logically, deals with the events preceding this instant, that is, the events inside the gun barrel or the details of the impact that produces the motion. In addition to "shooting," the field of ballistics also applies to many games such as football, baseball, golf, volleyball, and pool or billiards.

Exterior Ballistics

Exterior ballistics is the science dealing with the motions of projectiles not under propulsive power, as of a shell after it has left the cannon, of a rocket after burning has ceased, or of a baseball after being hit. If the Earth stood still and had no atmosphere, such trajectories would be simple to describe mathematically. The motion of a body projected at an angle to the Earth's surface is perturbed by several forces, so that the trajectory differs from the parabolic, hyperbolic, or elliptical form it would otherwise have. The perturbing forces are drag, cross wind forces, and the Coriolis effect. Their magnitude depends upon the shape and mass of the body, the density of the air, the wind currents, and the length of the path along the trajectory. These effects seriously complicate the calculation of the trajectory.

The trajectory of a projectile under the influence of the Earth's gravitational field and in vacuum would be an arc of an ellipse with the center of the Earth as one of the foci. This arc can be closely approximated by a parabola if the motion does not carry the projectile so high that the gravitational force differs significantly from the force at the Earth's surface. The gravitational force varies with height, as shown in Eq. (1), where g is the acceleration of gravity at

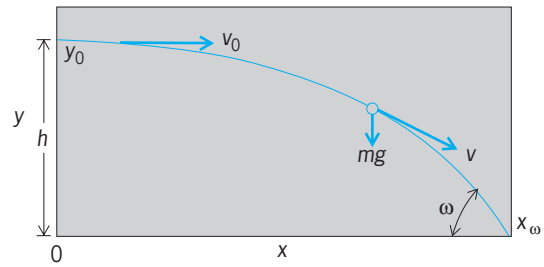

Fig. 1. Parabolic trajectory of a body projected horizontally, without resistance.

$$g = g_0\left(\frac{R}{R+y}\right)^2 \tag{1}$$

height y, g_0 is its value at the surface ($y = 0$), and R is the Earth's radius. In order that g fall to $0.99g_0$, y must be 20 mi (32 km).

Horizontal projectile. This trajectory is illustrated in **Fig. 1**. After the projectile of mass m leaves the muzzle at $x = 0$, the only force acting upon it will be gravity; this force will have a magnitude mg and will be directed downward. It is assumed here that the range is so short that the Earth may be considered flat, that the height is so low that g is the same as at the Earth's surface, and that the projectile is in a vacuum; complications due to the atmosphere will be dealt with later.

The motion relationships are given in Eqs. (2).

$$\begin{aligned} m\frac{d^2x}{dt^2} &= 0 \\ m\frac{d^2y}{dt^2} &= -mg \end{aligned} \tag{2}$$

The initial conditions (subscript 0) are given in Eqs. (3).

$$\begin{aligned} &x_0 = 0 && y_0 = h \\ &v_{x_0} = \left(\frac{dx}{dt}\right)_0 = v_0 && v_{y_0} = \left(\frac{dy}{dt}\right)_0 = 0 \\ &\left(\frac{d^2x}{dt^2}\right)_0 = 0 && \left(\frac{d^2y}{dt^2}\right)_0 = -g \end{aligned} \tag{3}$$

Solutions to these equations are given in Eqs. (4).

$$\begin{aligned} x &= v_0 t \\ y &= h - g\frac{t^2}{2} \end{aligned} \tag{4}$$

The curve of the trajectory, obtained by eliminating t, is the parabola shown in Eq. (5). The range of

$$y = h - \frac{g}{2v_0^2}x^2 \tag{5}$$

the point of fall x_ω (the value of x when $y = 0$) is given by Eq. (6).

$$x_\omega = v_0\sqrt{\frac{2h}{g}} \tag{6}$$

The time of flight t_ω (the value of t when $y = 0$) is independent of v_0 and is given by Eq. (7). The angle of fall ω is given by Eq. (8).

$$t_\omega = \sqrt{\frac{2h}{g}} \tag{7}$$

$$\tan\omega = -\left(\frac{dy}{dx}\right)_{y=0} = \frac{\sqrt{2gh}}{v_0} \tag{8}$$

Angular projectile. This trajectory is illustrated in **Fig. 2**. Let a body with mass m be projected in a vacuum with an initial velocity v_0 at an angle θ_0 to the horizontal, and assume a flat Earth with g constant. The origin of coordinates is chosen at the point of projection, the y axis is vertical, and the x direction is so chosen that the trajectory lies in the xy plane. The equations of motion are given by Eqs. (9). The initial conditions are given by Eqs. (10).

$$\begin{aligned} m\frac{d^2x}{dt^2} &= 0 \\ m\frac{d^2y}{dt^2} &= -mg \end{aligned} \tag{9}$$

$$\begin{aligned} \left(\frac{dx}{dt}\right)_0 &= v_0\cos\theta_0 \quad x_0 = 0 \\ \left(\frac{dy}{dt}\right)_0 &= v_0\sin\theta_0 \quad y_0 = 0 \end{aligned} \tag{10}$$

Integrating the equations of motion and inserting the initial conditions, one obtains Eqs. (11).

$$\begin{aligned} x &= v_0 t\cos\theta_0 \\ y &= v_0 t\sin\theta_0 - \tfrac{1}{2}gt^2 \end{aligned} \tag{11}$$

To obtain the shape of the trajectory, eliminate t, as shown in Eq. (12). This is the equation of a

$$y = x\tan\theta_0 - \frac{gx^2}{2v_0^2\cos^2\theta_0} \tag{12}$$

parabola whose vertex, the summit (x_s, y_s) of the tra-

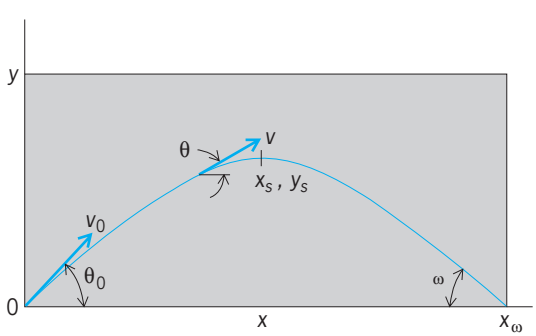

Fig. 2. Parabolic trajectory of a body projected at an angle, without resistance.

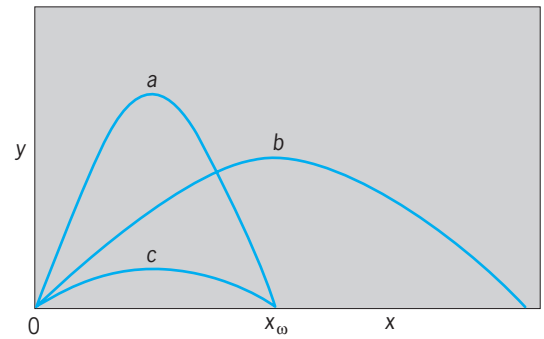

Fig. 3. Possible trajectories of body projected at same initial velocity but at different angle θ_0.

jectory, can be found by equating dy/dx to zero, as in Eqs. (13).

$$\begin{aligned} x_s &= \frac{v_0^2}{g}\sin\theta_0\cos\theta_0 \\ y_s &= \frac{v_0^2}{2g}\sin^2\theta_0 \end{aligned} \tag{13}$$

The time of flight to the summit is then given by Eq. (14). The range of fall x_ω can be determined by setting $y = 0$, as in Eq. (15). The time of flight t_ω, Eq. (16), is found by substituting $y = 0$ or $x = x_\omega$ in the equations involving t.

$$t_s = \frac{v_0}{g}\sin\theta_0 \tag{14}$$

$$x_\omega = \frac{2v_0^2}{g}\sin\theta_0\cos\theta_0 \tag{15}$$

$$t_\omega = \frac{2v_0}{g}\sin\theta_0 \tag{16}$$

Both the range and the time of fall are just twice the values at the summit, as would be expected from symmetry. To obtain the condition for maximum range for a given v_0, let $dx_\omega/d\theta_0 = 0$. Then Eq. (17) holds, and from it Eq. (18) is derived. With a fixed

$$\cos\theta_0 = \sin\theta_0 \tag{17}$$

$$\theta_0 = \frac{\pi}{4} = 45^\circ \tag{18}$$

initial velocity and for a less than maximum range, two trajectories are possible (**Fig. 3**), one with $\theta_0 > \pi/4$ and one with $\theta_0 < \pi/4$.

Motions of bodies with resistance. Resistance to the motion of a body through a fluid such as air derives from the fact that the body must move the fluid out of its way. The fluid particles are accelerated as the forward-moving body collides with them, and this results in a force, called drag, opposing the motion of the body and robbing it of some of its energy. The drag is a function of the density of the fluid ρ, of the area of the body, and of its velocity. The expression for the drag is given by Eq. (19), where

$$D = \rho d^2 v^2 K_D \tag{19}$$

d is the diameter of the body (assuming a circular

cross section), v the velocity, and K_D the drag coefficient. The drag coefficient itself is a function of velocity and undergoes a rapid change as the velocity approaches that of sound. Both D and K_D are functions of height, since both the density and temperature of air vary markedly with height. The velocity of sound varies with temperature according to Eq. (20),

$$a = a_s\sqrt{\frac{T}{T_s}} \tag{20}$$

where a is the velocity at absolute temperature T, and a_s is the velocity at standard temperature T_s (15°C or 59°F). The temperature in the atmosphere normally decreases with height for the first several miles.

The effect of the drag is to modify the theoretical vacuum trajectory so that the projectile falls short of the impact point it would have reached in a vacuum and impacts with a larger value of ω. Extensive tables have been determined empirically for the drag coefficient under wide variations of temperature and density. The equations of motion in the presence of drag are given by Eqs. (21).

$$\begin{aligned} m\frac{d^2x}{dt^2} &= -D\cos\theta \\ m\frac{d^2y}{dt^2} &= -D\sin\theta - mg \end{aligned} \tag{21}$$

The initial conditions are shown by Eqs. (22).

$$\begin{aligned} \left(\frac{dx}{dt}\right)_0 &= v_0\cos\theta_0 \\ \left(\frac{dy}{dt}\right)_0 &= v_0\sin\theta_0 \end{aligned} \tag{22}$$

In the equations of motion, the drag force D is a function of the velocity along the trajectory and is thus a complex function of x, y, and t. For this reason, these equations cannot be integrated by normal means, and approximate numerical methods have been developed.

The form of the drag coefficient, when plotted against the projectile velocity expressed in terms of the velocity of sound, is given in **Fig. 4**.

Prior to World War I, when shells were fired at low angle of departure (less than 15 or 20°), the Siacci method of calculating trajectories was commonly used. During that war the use of trajectories with large angles of departure (quadrant angles) became common, and new methods that were based on numerical integration were introduced—in the United States by F. Moulton. These methods were adapted to ballistics from earlier uses in astronomy.

The adequacy of the various approximations which can be used for the calculation of missile trajectories is determined, in the main, by the velocity of the missile and the elevation of the trajectory. In order of complexity, the methods are as follows.

Vacuum approximation. When the velocity is low, the quadrant angle small, and the projectile rather heavy (so that the range is short and atmospheric effects are small), the equations developed earlier for the vacuum trajectory may be used with good accuracy.

Didion-Bernoulli method. This method may be used when the quadrant angle is low (less than about 20°) and the velocity is less than about 800 ft/s (240 m/s), so that the drag may be considered proportional to the square of the velocity. The variables in the equations of motion are separated by employing, in certain critical places in the solution, the approximation that sec θ = constant. Once this assumption is admitted, the solution is exact and is found in formulas written around four functions. These functions and the appropriate values of the constant representing sec θ are tabulated and, in conjunction with the formulas representing the solution of the equations of motion, make possible the nearly exact calculation of low-velocity, low-angle trajectories.

Siacci method. This method is accurate and useful for trajectories of high-velocity missiles with low quadrant angles of departure. The basic assumptions are that the atmospheric density anywhere on the trajectory is approximately constant and that the angle of departure is less than about 15°. The equations of motion can then be put in a form permitting a solution in terms of formulas for t, x, y, and tan θ. These formulas, used in conjunction with the tables of four functions involved in the solution, yield the trajectories of projectiles fired at low angles.

Otto-Lardillon method. For trajectories of missiles with low velocities (so that the drag is proportional to the velocity squared) and quadrant angles of departure that may be high, exact solutions of the equations of motion are arrived at by numerical integration and are then tabulated.

Numerical integration. When projectiles with high velocities and high quadrant angles of departure were introduced, all earlier methods used in ballistics failed to give sufficiently accurate solutions. The introduction of high velocities and high angles meant increased complexity both in the variation of the drag coefficient, so that drag was no longer simply a function of the velocity squared, and in the variations

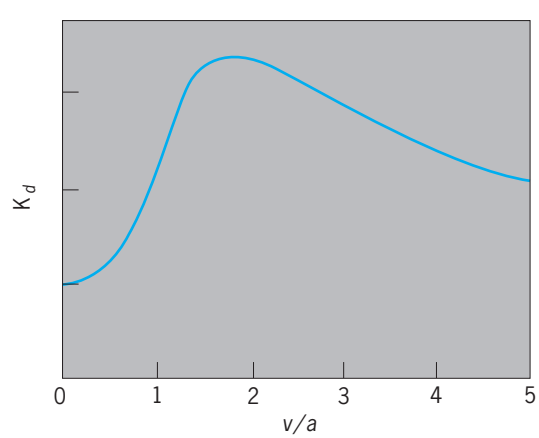

Fig. 4. Plot of drag coefficient K_D. Quantity v is velocity of projectile, and a is velocity of sound.

Methods of trajectory computation

Velocity	Maximum range	Quadrant angle	Method
Low	Short	Low	Vacuum approximation
Low	Medium	Low	Didion-Bernoulli
High	Medium-long	Low	Siacci
Low	Medium	Any	Otto-Lardillon
High	Long	Any	Numerical integration

of atmospheric density along the trajectory. Moulton and his associates introduced the method of approximate numerical integration, long used in astronomy, to obtain solutions for these trajectories. Computers are used extensively to perform the necessary calculations. *See* NUMERICAL ANALYSIS.

Since the advent of rockets which travel high in the atmosphere, the method of numerical integration has been very widely used. Although the problem is complex, even a desktop personal computer can obtain the solution.

Drag effects are very important in many sports. The atmosphere can be considered to have a constant density in sports such as football, golf, and baseball. Both golf balls and baseballs often have initial speeds in excess of their terminal speeds.

The basic approach in solving the rocket trajectory problem is to determine the conditions x, y, dx/dt, dy/dt, and θ at the end of burning and to go back along a hypothetical trajectory to determine the coordinates, velocity, and quadrant angle of departure of an equivalent cannon shell which will pass through the same point with the same velocity vector as the spent rocket. Any of the five methods outlined in the preceding discussion may then by used to calculate the resulting trajectory. An indication of the best method for use with various combinations of the initial conditions of velocity and quadrant angle is given in the **table**.

Effect of Earth's curvature. The above discussion of trajectories is based on the assumption that the Earth may be considered flat over the range of the projectile, and that the force of gravity is directed at all times parallel to its original direction. When the range of the missile or projectile becomes great, this assumption can lead to serious error; thus the analysis must be modified to accommodate the fact that the force of gravity varies with height, is directed toward the center of the Earth, and changes direction as the projectile moves forward. Also, the impact point is depressed because of the Earth's curvature, and the range therefore is extended.

These effects are customarily taken care of by the method of differential corrections but will be discussed here in a somewhat different fashion to clarify the point. Assume a nonrotating Earth for the time being. A projectile hurled along a trajectory away from the surface of the Earth is, while in motion, a satellite of the Earth and moves in its trajectory, which is really a section of an elliptical orbit about the center of the Earth. This is shown diagrammatically in **Fig. 5**. All trajectories or orbits with velocities less than about 25,000 ft/s (7600 m/s) intersect the surface of the Earth. The equations of motion for this two-body problem have been solved in celestial mechanics, except for the complications of drag and varying air density.

The ballistic problem could be solved by treating it as one in celestial mechanics and determining the intersection of the Earth's surface with the orbit to give the impact point. For projectiles of intermediate range, this is not done, but rather the technique of differential corrections on existing methods is used. However, for extremely long-range missiles, the approach of celestial mechanics employing numerical integration to take account of changing air densities and drag coefficients is used. *See* CELESTIAL MECHANICS.

Effect of Earth's rotation. Once the projectile is clear of the muzzle of the gun or the rocket is clear of the launching stand, it is a separate entity free of the Earth's rotation. One component of its initial velocity is the velocity of the surface of the Earth at the point of firing. For this reason, when the projectile arrives at the impact point, it will miss by an amount dependent on the azimuth of the trajectory, the range, and the height of the trajectory. This is called the Coriolis effect. For intermediate trajectories, such as those with pronounced Earth curvature, the Coriolis effect is handled by means of differential corrections. For very long-range missiles, the approach of celestial mechanics with nonrotating Earth-centered coordinates is used. Here the rotation is taken care of by adding the proper vector velocity for the latitude of launch and by taking account of the rotation of the Earth during the time of flight. *See* CORIOLIS ACCELERATION.

Stabilization by spinning. In order to stabilize a projectile in flight, the projectile is frequently spun about its longitudinal axis. The resulting gyroscopic

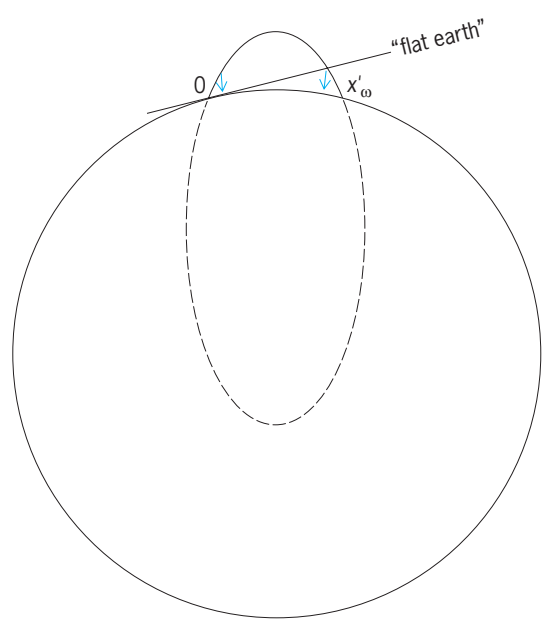

Fig. 5. Elliptical trajectory of a long-range missile. The range of fall is x'_w rather than x_w.

stabilization keeps the projectile from tumbling and hence maintains the proper orientation in flight to keep drag at a minimum. The spin of a shell is imparted by rifling the gun barrel, that of a rocket by canting the exhaust nozzles or the fins. While spin does tend to stabilize the orientation, in practice the spin axis will precess so that the nose of the projectile follows a spiral path. This causes troublesome complexities in the trajectory calculations, in that oscillating lift and yaw terms must be introduced. The equation of motion of a spinning projectile can be integrated by high-speed computing machines.

A further difficulty in spinning projectiles fired with large quadrant angles of departure is that the spin, in stabilizing the orientation along the original direction, causes the projectile to arrive at the point of impact at an inclination. that is different than the trajectory direction. Billiards and golf both depend upon understanding the effects of spin on the motion.

John P. Hagen; John L. Safko

Interior Ballistics

Interior ballistics deals with the motions of projectiles while they are still inside the gun barrel or cannon, and are accelerated by the pressure of expanding gases resulting from the burning of the change. The term is also used to refer to the similar events inside the casing of a solid-fuel rocket, in which case the exhaust blast is regarded as equivalent to the projectile.

Gas pressure changes. If the propelling charge of the cartridge is changed into gas instantaneously, the gas pressure curve, from the instant of ignition to the instant the projectile leaves the muzzle, would be very nearly an adiabatic curve, quite similar to the curve showing the diminishing potential of a gravitational field. In reality, it takes some time for the powder charge to burn; furthermore, various resistance factors must be overcome. Therefore, the gas pressure in the gun barrel behind the projectile does not reach its maximum value until after the projectile has started moving inside the barrel. The pressure should reach a minimum value at the instant the projectile leaves the gun barrel, as shown in **Fig. 6**; hence, there is a definite relationship between powder charge, gas pressure, and barrel strength. If the pressure is still rather high even when the projectile "unseals" the gun barrel, a portion of the propelling charge has been wasted. If the pressure drops too low before the projectile reaches the muzzle, the projectile is slowed down inside the barrel; in this case, shortening the barrel would actually improve the performance.

Minimizing of recoil. Even when the pressure drop has been properly calculated (and achieved), the exit velocity of the powder gases is still 2 to $2^1/_2$ times the muzzle velocity of the projectile. This fact can be utilized to prevent or to minimize the recoil of the weapon by means of the muzzle brake first developed by Schneider-Creusot in France. The muzzle brake deflects the direction of motion of the gases by 120–180°, thereby creating what might be called a gas recoil opposite in direction to the projectile recoil. Recoilless guns eliminate recoil by discharging a quantity of the propelling gases through the rear.

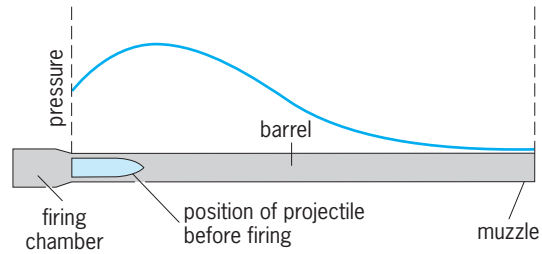

Fig. 6. Gas pressure in gun barrel.

Resistance factors. The expanding gases in the gun barrel have to overcome the following main resistance factors (in addition to their own inertia): (1) the inertia of the projectile, (2) the friction in the barrel caused by the copper driving bands against the lands and grooves of the rifling, and (3) the inertia of the air inside the barrel between the nose of the projectile and the muzzle of the gun. This last factor is rather unimportant for small guns or for short-barreled pieces of any caliber, and is often overlooked. However, it does influence performance in large and long-barreled pieces, especially when the muzzle velocity of the projectile is much larger than Mach 1, since in that case the air is not only accelerated but also compressed. The existence of this "plug" of compressed air often makes the experimental determination of the muzzle velocity difficult.

Further energy losses are caused by gas leakage, by the recoil motion of the gun barrel, and by the heating of the barrel. This last phenomenon not only absorbs energy but may increase the leakage. Obviously, there are many interdependent factors which make an exact calculation impossible. Empirical values must often be used.

Measuring devices. At first, the only tool of the ballistician was the ballistic pendulum (**Fig. 7**), which caught the projectile and indicated its momentum by the magnitude of its swing. This method was first suggested by J. D. Cassini in 1707 and was brilliantly utilized by Benjamin Robins some 30 years later. However, it permitted only indirect conclusions as to gas pressure inside the barrel. The first useful device for

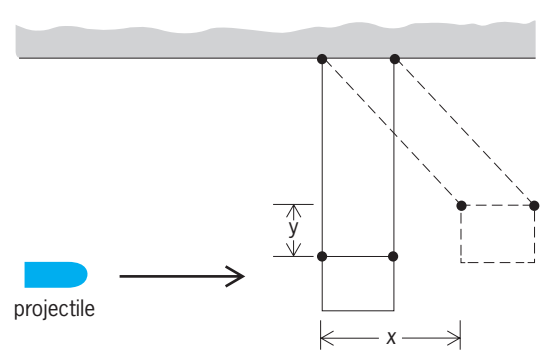

Fig. 7. Ballistic pendulum before (solid lines) and after (broken lines) projectile impact. Measurements of *x* and *y* and knowledge of mass and initial velocity of projectile give its momentum.

measuring gas pressure was A. Nobel's copper pressure gage, in which a copper plug was deformed by the gas pressure. The copper pellet (a lead pellet is used for weapons where the gas pressure is expected to be low) has remained the most useful tool since its introduction in 1860. Not until about 1940 was it supplemented by the piezoelectric gage (suggested originally by J. J. Thomson), which consists of a tourmaline disk that generates electric currents proportional to the pressure to which it is subjected.

In addition, strain gages have shown that a gun barrel momentarily expands slightly as the projectile moves toward the muzzle.

Basic equations. The main problems of interior ballistics have been defined by H. Athen as follows: "For a gun of which caliber, barrel length, firing chamber volume, weight of projectile, and type and weight of the propelling charge are known, the following factors must be determined: duration of burning of the powder charge, duration of travel of the projectile inside the barrel, gas pressure inside the barrel, and projectile velocity. Of these factors, the peak gas pressure and the muzzle velocity are especially important, since they can most easily be checked experimentally and are also of most practical value to the designer."

Investigations of the problems led to three equations, which are known as the energy equation, the combustion equation, and the equation of motion.

Energy equation. The derivation of the energy equation is based on the van der Waals equation (23),

$$\left(p + \frac{\beta}{s_v^2}\right) \cdot (s_v - \alpha) = RT \tag{23}$$

where p is the pressure, T the absolute temperature, R the gas constant, s_v the specific volume of the gas, α the volume of the gas for $T = 0$, and β a constant. It is empirically known that β/s_v^2 is nearly zero; it is further known that for each type of powder, RT_c (T_c is the combustion temperature) is a constant value which is designated f and is called the specific pressure. If the volume of the closed space in which the powder is burned is designated as S and the weight of the powder as W, W/S might be called D (density of the charge), and the much simplified Abel equation (24), is obtained. It has been shown that

$$p = f\frac{D}{1 - \alpha D} \tag{24}$$

this equation holds true for pressures from 17,000 to 64,000 lb/in.^{2} (120 to 440 megapascals or 1200 to 4500 kgf/cm^2). By introducing a few additional factors, Resal's energy equation is obtained.

Equations of combustion and motion. The combustion equation is shown in (25), where y is the distance

$$\dot{y} = A(y)p^k \tag{25}$$

that the explosion has propagated, $\dot{y} = dy/dt$ is the velocity of the pressure front, (y) is a function determined by the grain shape of the charge and the location of the pressure front, A is a constant characteristic for the propellant under investigation, and p is the gas pressure. The factor k is generally taken to be less than unity.

The equation of motion is derived from Newton's second law. It is shown, taking $k = 1$, in Eq. (26),

$$\dot{v} = v\frac{dv}{dx} = \frac{qp}{\mu} \tag{26}$$

where v is the velocity of the projectile inside the launcher, μ is the mass of the projectile, and q is the projectile cross section. *See* EXPLOSIVE; GAS.

If the interior ballistic is an impact such as a bat hitting a ball, the angle of impact and the structure of the projectile must be used to calculate the initial velocity and direction.

Willy Ley; David Wollersheim; John L. Safko

Bibliography. R. K. Adair, *Physics of Baseball*, 3d ed., Harper-Collins, New York, 2001; P. J. Brancazio, *Sport Science*, reprint edition, Touchstone, New York, 1985; C. L. Farrar and D. W. Leeming, *Military Ballistics: A Basic Manual (Land Warfare: Brassey's New Battlefield Weapons Systems & Technology Series into the 21st Century)*, Brassey's Publishers, London, 2002; R. L. McCoy, *Modern Exterior Ballistics: The Launch and Flight Dynamics of Symmetric Projectiles*, Schiffer Publications, Atglen, PA, 1999.

Balloon

A nonporous envelope of thin material filled with a lifting gas and capable of lifting any surrounding material and usually a suspended payload into the atmosphere. A balloon which is supported chiefly by buoyancy imparted by the surrounding air is often referred to as an aerostat. The balloon rises because of a differential displacement of air according to Archimedes' principle, which states that the total upward buoyant force is equal to the weight of the air displaced. The upper practical limit for useful ballooning is approximately 34 mi (55 km). Beyond this, the exponential nature of the atmosphere would require balloons of enormous size and delicately thin skin. *See* ARCHIMEDES' PRINCIPLE; ATMOSPHERE.

Balloons have been configured in many geometrical shapes, but the most common are spheres, oblate spheroids, and aerodynamic configurations. The materials used in the manufacture of the balloon envelope have been paper, rubber, fabric, and various plastics. Several types of lifting gases have been used to inflate balloons, but the most common in use are helium, hydrogen, and heated air.

History of ballooning. Crewed ballooning began on November 21, 1783, in Paris, France, when a balloon built by the brothers J. E. and J. Montgolfier carried the first two aeronauts, Pilâtre de Rozier and the Marquis François Laurent d'Arlandes, into the atmosphere. The balloon was constructed of paper and linen and used smoke and heated air for the lifting gas. Early balloons found their primary applications as a war weapon and for sport, and serious scientific

balloon experiments were not conducted until the late nineteenth century.

In 1934 and 1935 in the United States the National Geographic Society and the Army Air Corps cosponsored two crewed balloon flights to conduct research in the stratosphere by measuring cosmic radiation, vertical ozone content, and composition of the atmosphere. Following this activity, a surge of crewed ballooning was conducted, with records being established, only to be broken on upcoming flights. Altitude records in excess of 100,000 ft (30 km) were recorded in crewed pressurized capsules. The flurry of crewed scientific ballooning continued until 1961, when it abruptly ended and yielded to the space era.

During the early 1950s considerable research was conducted on the design of balloons and improved materials for balloon manufacture. This research was primarily sponsored by the joint military services of the United States and resulted in the development of the modern zero-pressure balloon constructed of thin polyethylene film. This balloon is by far the most common vehicle for carrying large scientific experiments to the fringes of space. The polyethylene balloon has increased in volume from 1×10^6 to 7×10^7 ft^3 (3×10^4 to 2×10^6 m^3). Payload weight increased from hundreds of pounds to over 11,000 lb (5000 kg).

Balloon types. The many types of balloons in use fall into two main categories, extensible (expandable) and nonextensible. There are three methods of balloon operation: free balloons which are released into the atmosphere, tethered or moored balloons, and powered or controlled free balloons. The various types of balloons in use are the hot-air balloon, meteorological balloon, zero-pressure balloon, superpressure balloon, tethered balloon, and powered balloon.

Hot-air balloon. Historically, this balloon was used for many of the earliest achievements, and contributed significantly to science and aerostatic research. However, since about 1960 its primary function has been as a free balloon in sport, to carry people aloft. The gas bag is constructed of a nonextensible material and is usually spherical. The heated air required to produce lift is generated as a hydrocarbon gas burner attached above the basket. This basket, or gondola, is used to carry the passengers into the atmosphere. Lift can be controlled by adjusting the burning rate of the gas. A valve at the top of the balloon has an attached rope that allows a passenger to control descent. A rip panel and rip cord are attached to the balloon to allow rapid release of the gas on landing, to prevent dragging the load on impact.

Meteorological balloon. The meteorological or sounding balloon is used to carry radiosondes and other lightweight instruments aloft. It is an extensible balloon made from natural or synthetic rubber and filled with either hydrogen or helium. Hundreds of these balloons are launched daily by meterological stations throughout the world in support of the upper-air program of the World Meteorological Organization, as well as local weather forecasting. These flights are vertical sounding probes of the atmosphere for measuring pressure, temperature, humidity, and wind velocity. They can carry up to 40 lb (80 kg) of useful payload. High-altitude types carrying much lighter payloads may exceed 25 mi (40 km) before expanding to the bursting point. *See* METEOROLOGICAL INSTRUMENTATION.

Zero-pressure balloon. This is the workhorse of scientific ballooning, with a worldwide launch rate of several hundred per year. This is a nonextensible balloon constructed of thin-film polyethylene in a tailored natural shape (inverted teardrop). At float, it assumes the theoretical shape of a load-bearing balloon with zero circumferential stress. The polyethylene is manufactured by a blown extruding technique which produces a continuous tube of thin film. The thickness of the film can be adjusted by varying the extruder die, temperature, and pressure to produce films as thin as 9 micrometers (0.00035 in.). The tube is slit, unfolded, and rerolled on spools which are approximately 10 ft (3 m) wide. The material is then laid out on an assembly table, which is the full length of the balloon, and is cut to conform to one vertical segment of balloon, which is referred to as a gore. Additional gores are then added, with a polyester load tape heat-sealed to each seam. When finished, the balloon consists of 60 to 200 polyethylene gores, with load-bearing tapes extending from the apex top-fitting to the bottom end-fitting. The bottom of the balloon is left open through an arrangement of ducts, and hence the term zero-pressure balloon.

During inflation, the lifting gas is contained in a bubble which occupies less than 1% of the total balloon volume (**Fig. 1**). Approximately 10% additional lift over that required for neutral buoyancy is added to the balloon to provide the necessary vertical force to achieve an ascent rate of 800 ft/min (250 m/min).

Typically, these balloons require 3 h to reach a float altitude of 100,000 to 140,000 ft (30 to 42 km). After reaching float, the excess lift is allowed to vent through the ducts. To maintain a constant altitude, this system requires the release of ballast whenever a decrease in heating due to solar radiation is experienced. This effect is most noticeable when entering sunset condition, but can also be significant when passing over extensive cloud cover or thunderstorm activity.

A dynamic launch technique is generally employed when releasing these balloons. The payload is suspended on a launch vehicle, while the balloon and deployed parachute are laid out on the ground. A restraining spool holds the inflated bubble of lifting gas at the far end of the launch site. After sufficient lifting gas has been transferred to the bubble, the polyethylene fill tubes are tied off. On command from the launch director, the spool is released, allowing the bubble of gas to accelerate in a vertical direction. When the entire system is in a vertical position, the payload is released from the launch vehicle, which has been positioned directly under the ascending balloon train.

The float duration of a typical zero-pressure balloon experiment is from 8 to 40 h, depending on the time of year of launch. Recovery of the scientific equipment is essential; therefore coastal regions and international borders determine the extent of usable range. The upper-altitude winds are somewhat

Fig. 1. Inflation of zero-pressure balloon. (*NASA*)

predictable and experience two 180° east-west reversals or "turn-arounds" during the year. The turnaround periods exist in May and September in the United States. During these periods the upper-altitude winds are less than 20 mi/h (10 m/s), and flights of several days duration can be accomplished. The winds become easterly during the summer and reach maximum velocities of 70 to 90 mi/h (30 to 40 m/s) during July. In the winter the winds become westerly and may exceed 60 m/s (136 mi/h).

The zero-pressure balloon is capable of carrying scientific instruments weighing thousands of pounds to an altitude of 140,000 ft (42 km), which is above 99.8% of the Earth's atmosphere. These scientific instruments have become very sophisticated and resemble a satellite in complexity. They utilize the latest technology in electronics; in fact, new technology is often proved in balloon experiments before being flown on satellites. The reasons are the shorter development time for the balloon experiment and the lower cost of the system, which allows a higher risk. The subsystems contained in a typical automated gondola are a high-bit-rate pulse-coded-modulated telemetry system for sending the scientific data to the ground tracking station, a digital command system used to convey information from the ground to the instrument, a battery power system, a temperature-control system, and a pointing system.

Superpressure balloon. A balloon that has been the subject of considerable research is the superpressure or constant-level balloon (**Fig. 2**). This is a nonextensible (constant-volume) free balloon which converts its free lift into superpressure by the time the float level is reached. As long as the balloon is superpressurized, it will continue to float at a constant density level. The main advantage of the superpressure balloon is the possibility of long-duration flights which can circumnavigate the globe. Small superpressure balloons 13 ft (4 m) in diameter have been released in the Southern Hemisphere, and some have remained aloft for over a year. Balloons having diameters of over 100 ft (30 m) have been flown with limited results. The objective of the research has been to produce a reliable superpressure balloon capable of carrying 700 lb (300 kg) to an altitude of 140,000 ft (42 km). The primary failure mode of the superpressure balloon is small pinholes as a result of manufacture and handling, which lead to the premature loss of superpressure. To minimize this effect, bilaminated materials usually of Mylar or some other strong polyester film are used.

Tethered balloon. This type of balloon is attached to a mooring line and allowed to ascend by releasing a winch mechanism (**Fig. 3**). The tethered balloon is generally configured as an aerodynamic body to provide lift, achieve lower drag, and increase lateral stability. Although 20,000 ft (6 km) is generally considered the upper altitude for this concept, the National Center for Space Studies in France reported flying a tethered balloon carrying an 800-lb (350-kg) gondola to an altitude of 55,000 ft (17 km) for 12 h.

The tethered balloon has primarily been used for military purposes to provide acoustic, radar, and optical surveillance, tall antennas, and communication relays. The inherent danger of the mooring cable to aircraft has prevented the application of the high-altitude tethered balloon for commercial and scientific uses. However, a commercial application is in

Fig. 2. Superpressure balloon. (*Raven Industries*)

Fig. 3. Tethered balloon. (*U.S. Air Force*)

the logging industry, where heavy-lift, low-altitude balloons provide a means of removing inaccessible timber. This balloon has a lift capability of 22,000 to 44,000 lb (10,000 to 20,000 kg).

Powered balloon. Many large, crewed, lighter-than-air ships or dirigibles of both the rigid and nonrigid type were produced in the United States and Germany prior to World War II. These craft can be propelled and steered. The rigid dirigible was popular until the *Hindenburg* disaster of 1937. The nonrigid airships or blimps were used by the U.S. Navy during World War II for antisubmarine detection. With the advancement of aircraft design providing all-weather operation and superior performance, the crewed airships were driven into obscurity, but these craft are again being seriously considered for a number of civil and military applications. *See* AIRSHIP; BLIMP.

Efforts were undertaken to develop an uncrewed high-altitude (65,000-ft or 20-km) powered balloon which would take advantage of the low wind field and remain stationary over given areas for long periods of time. These platforms would be used for communication purposes, earth resource survey sensors, and military applications.

Space application. The zero-pressure and superpressure balloons have been used for space-related programs. The most extensive use has been development and testing of new detectors, experiments, and sensors which are later flown on satellites or are to be flown on the space shuttle. The balloon provides an ideal vehicle to carry x-ray, gamma-ray, cosmic-ray, infrared, and ultraviolet experiments above the Earth's atmosphere to obtain scientific data. These instruments can be recovered, redesigned, and reflown many times before commitment to a satellite program.

NASA conducted a program using zero-pressure balloons to perform in-place and remote measurements of stratospheric constituents of ozone, nitric acid, water vapor, nitrogen dioxide, and temperature. These balloon measurements were conducted simultaneously with the overpass of the *Nimbus-G* satellite, with its sensors providing remote measurements of the same data. These correlative measurements provided a means of validating the satellite data.

Another NASA program, the Global Atmospheric Research Program, involves the release of hundreds of superpressure balloons to provide measurements of pressure, temperature, and wind direction of the atmosphere. These data are transmitted to the *Tiros-N* satellite and later retransmitted to satellite ground stations to provide worldwide meteorological data. *See* METEOROLOGICAL SATELLITES. Walter R. Nagel

Bibliography. T. D. Crouch, *The Eagle Aloft: Two Centuries of the Balloon in America*, 1983; W. Reidler and K. M. Torkar (eds.), *Balloon Technology and Observations*, 1993.

Balsa

A fast-growing tree, *Ochroma lagopus*, widely distributed in tropical America, especially in Ecuador. The leaves are simple, angled, or lobed, and the flowers are large and yellowish-white or brownish and they are terminal on the branches. *See* MALVALES.

With plenty of room for growth in a rich, well-drained soil at low elevations, trees may attain a height of 50–60 ft (15–18 m) and a diameter of 24–30 in. (60–75 cm) in 5–6 years. Under such conditions, the wood is very light and soft, weighing about 6–8 lb/ft^3 (95–130 kg/m^3). Culture is important, for if the trees are injured only slightly, the wood develops a hard and fibrous texture, thereby losing its commercial value. To secure a uniform product the trees must be grown in plantations.

The wood decays easily in contact with the soil and is subject to sap stain if not promptly dried. Seasoned lumber absorbs water quickly, but this can be largely overcome by waterproofing. Balsa became popular during World War I, and large quantities were used for life preservers and mine buoys.

Balsa owes most of its present commercial applications to its insulating properties. Balsa has sound-deadening qualities, and is also used under heavy machinery to prevent transmission of vibrations. Certain other woods are lighter, but they lack the strength of balsa and are not available in sufficient size or quantity to be commercially important. The heartwood of balsa is pale brown or reddish, whereas the sapwood is nearly white, often with a yellowish or pinkish hue. Luster is usually rather high and the wood is odorless and tasteless. *See* FOREST AND FORESTRY; TREE. Arthur H. Graves/Kenneth P. Davis

Baltic Sea

A semienclosed brackish sea located in a humic zone, with a positive water balance relative to the adjacent ocean (the North Sea and the North Atlantic). The Baltic is connected to the North Sea by the Great Belt (70% of the water exchange), the Øresund (20% of the water exchange), and the Little Belt. The total area of the Baltic is 147,414 mi^2 (381,705 km^2), its total volume 4982 mi^3 (20,764 km^3), and its average depth 181 ft (55.2 m). The greatest depth is 1510 ft (459 m), in the Landsort Deep.

The topography of the Baltic is characterized by a sequence of basins separated by sills and by two large gulfs, the Gulf of Bothnia (40,100 mi^2 or 104,000 km^2) and the Gulf of Finland (11,400 mi^2 or 29,500 km^2). More than 200 rivers discharge an average of 104 mi^3 (433 km^3) annually from a watershed area of 637,056 mi^2 (1,649,550 km^2). The largest river is the Newa, with 18.5% of the total freshwater discharge. From December to May, the northern and eastern parts of the Baltic are frequently covered with ice. On the average, the area of maximum ice coverage is 82,646 km^2 (214,000 km^2). The mean maximum surface-water temperature in summer is between 59 and 63°F (15 and 17°C).

Geologic history. The Baltic Sea in its present form is a relatively young postglacial sea which developed from the melting Fenno-Scandian glaciers about 14,000 years ago. Depending on the balance between the elevation of the Fenno-Scandian land area (due to the decreasing ice weight) and the rising of the oceanic water level, the connection to the open ocean was several times closed and then reopened. Among the stages were the Baltic Ice Lake (12,000–8000 B.C.; arctic, fresh water), the Yoldia Sea (8000–7250 B.C.; arctic, oceanic), the Ancylus Sea (7250–2000 B.C.; warm, fresh water), the Littorina Sea (2000 B.C–A.D. 500; moderate, brackish to saline), and the Mya Sea (A.D. 500 to present; moderate, brackish) [see **illus.**].

Temperature. As the Baltic stretches from the boreal to the arctic continental climatic zone, there are large differences between summer and winter temperature in the surface waters, ranging from about 68 to 30°F (20 to −1°C) in the Western Baltic and 57 to 32°F (14 to −0.2°C) in the Gulf of Bothnia and the Gulf of Finland. Usually one or more sharp thermoclines develop April through June between 16 and 80 ft (5 and 25 m), separating the brackish winter water from the surface layer. Below the permanent thermohalocline, at a depth of 190 to 240 ft (60 to 75 m), the water temperature is almost constant (42 to 44°F or 5.5 to 6.5°C), not influenced by seasonal warming or cooling. Vertical convection due to the cooling of the surface water in autumn and winter penetrates to the bottom only in the inner Gulf of Bothnia and the Gulf of Finland. A significant increase in the temperature in the deep water has been observed since 1900; the increase is about 3.6°F (2°C) in the Bornholm Basin and 2.7°F (1.5°C) in the Central Baltic.

Salinity. The salt content of the Baltic waters is characterized by two major water bodies; the brackish surface water and the more saline deep water. Salinities for the surface water range from 8 to 6‰ in the Western and Central Baltic and 6 to 2000 in the Gulf of Bothnia and the Gulf of Finland; salinities for the deep water range from 18 to 13‰ in the Western and Central Baltic and 10 to 4‰ in the Gulf of Bothnia and the Gulf of Finland. The boundary between these two major water bodies is marked by a relatively steep and permanent halocline which to a large extent prevents vertical exchange of water and dissolved substances. The total average annual outflow of brackish surface water from the Baltic is 220 mi^3 (942 km^3), and the salt balance of the Baltic is maintained by an average annual inflow of 113 mi^3 (471 km^3). Most of the oceanic water entering the Belt Sea does not reach the Baltic proper. Major intrusions of oceanic water happen only every 5 to 10 years; they are caused by the coincidence of several meteorologic factors during the cold season.

Circulation. The surface currents of the Baltic are dominated by a general counterclockwise movement and by local and regional wind-driven circulations. A complex system of small- and medium-scale gyres develops especially in the central parts of the Baltic. The currents in the Belt Sea are dominated by the topography; they are due to sea-level differences between the Baltic proper and the North Sea. Tides are of minor importance, ranging between 0.8 and 4.7 in. (2 and 12 cm). Water-level changes of more than 6 ft (2 m) occur occasionally as a result of onshore or offshore winds and the passage of cyclones over the Baltic Sea area. The frequency of longitudinal sea-level oscillations is about 13.5 h. *See* OCEAN CIRCULATION.

A significant increase of 1–2‰ in both surface and deep water salinities, probably caused by long-term changes in meteorological conditions and by the decrease in precipitation, has been observed since 1900.

Oxygen and nutrients. Due to the limited exchange of deep water with the North Sea and the permanent halocline, the oxygen content of the water below the halocline is considerably reduced. Long-term records show an average annual oxygen deficit of 4.9×10^{10} oz (1.4×10^{12} g) for the Baltic proper. Anaerobic conditions develop frequently in depths below 380 ft (120 m), wiping out all benthic and aquatic life except anaerobic bacteria in large parts of the Central Baltic deep waters. There is a significant increasing trend in the nutrient content of Baltic deep, intermediate, and surface waters, resulting in increased eutrophication. This increase is caused partly by increasing stagnation and partly by human activities. *See* EUTROPHICATION.

Ecology. The flora and fauna of the Baltic are those of a typical brackish-water community, with considerably reduced numbers of species compared to an oceanic community. The productivity is relatively low compared to other shelf seas. The major commercially exploited species are cod, herring, sprat, flounder, eel, and salmon, and some fresh-water species such as whitefish, pike, perch, and trout. The total annual catch amounts to about 880,000 tons (800,000 metric tons). The Baltic is completely

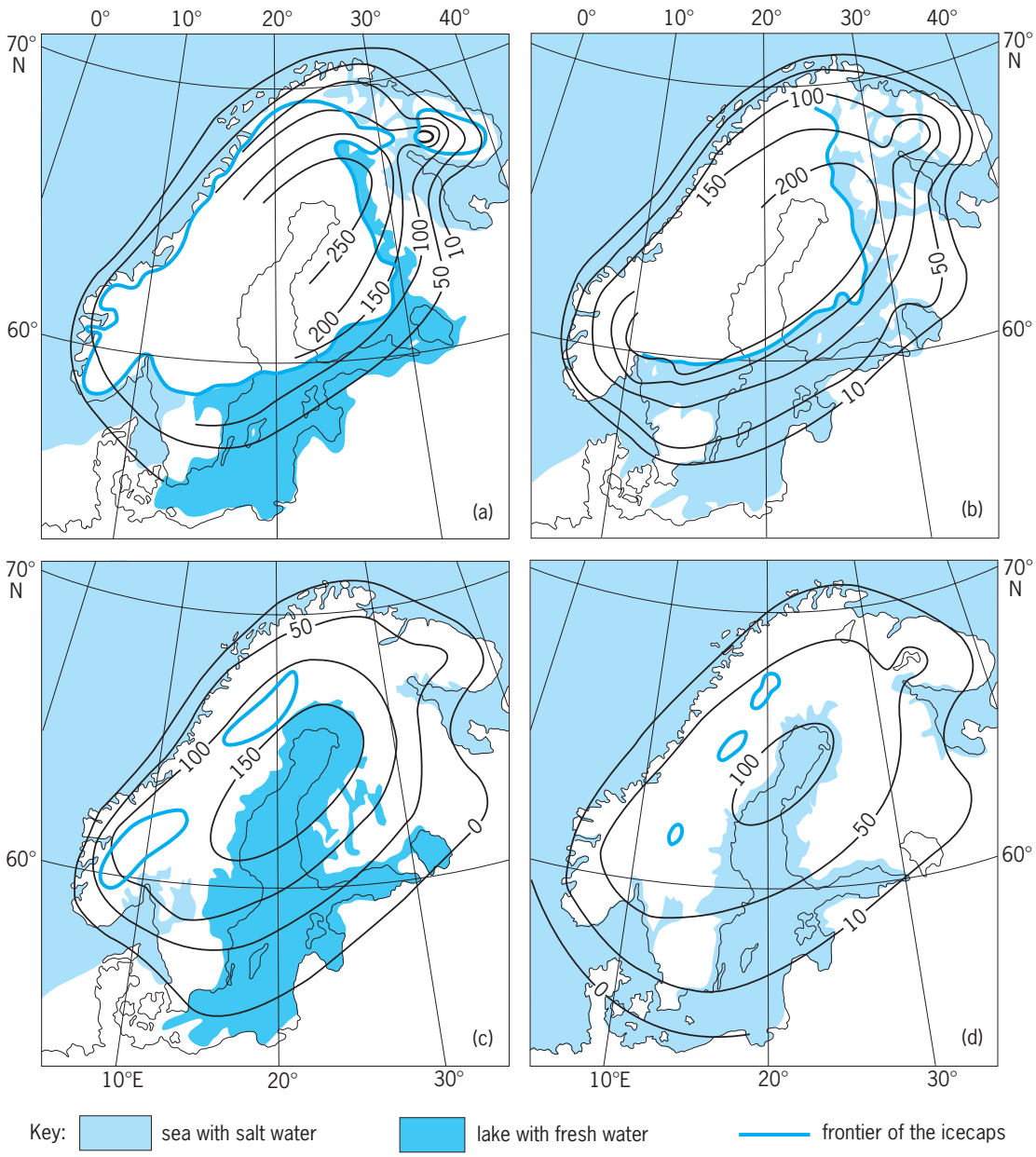

Postglacial development of the Baltic: (*a*) Baltic Ice Sea, 12,000–8000 B.C.; (*b*) Yoldia Sea, 8000–7250 B.C.; (*c*) Ancylus Sea, 7250–2000 B.C.; (*d*) Littorina Sea, 2000–A.D. 500. (*After K. Grasshoff, in S. P. Riley and G. Skirrow, eds., Chemical Oceanography, vol. 2: The Hydrochemistry of Landlocked Basins and Fjords, Academic Press, 1975*)

divided into fishery zones, with exclusive fishing rights belonging to the respective countries.

Pollution. The water pollution of the Baltic is generally decreasing as a result of the Convention for the Protection of the Marine Environment of the Baltic; this covered both land-based and air-borne pollution, and was signed in 1974 by all seven Baltic states. It came into power on May 3, 1980. The major sources of pollution are domestic and industrial wastewater. *See* WATER POLLUTION.

Resources. Other than fish the only major resources that have been exploited are sand and gravel in the Western Baltic Sea. It is believed that the deeper layer under the Gotland Basin contains mineral oil, but so far only exploratory drilling has been carried out in the near-coastal regions. Limited amounts of mineral oil have also been located in the Gulf of Kiel. K. Grasshoff

Bibliography. G. Alexandersonn, *The Baltic Straits*, 1982; S. P. Riley and G. Skirrow (eds.), *Chemical Oceanography*, vol. 2: *The Hydrochemistry of Landlocked Basins and Fjords*, 1975; A. Voipio (ed.), *The Baltic Sea*, 1981.

Bamboo

The common name of various perennial, ornamental grasses (Gramineae). There are five genera with approximately 280 species. They have a wide distribution, but occur mainly in tropical and subtropical parts of Asia, Africa, and America, extending

from sea level to an elevation of 15,000 ft (4572 m). Their greatest development occurs in the monsoon regions of Asia. The plants grow very rapidly. From the jointed rhizome, numerous straight, usually erect, stems arise, which at their full height produce dense masses of horizontal branches. The giant bamboo (*Dendrocalamus giganteus*), the largest known grass, attains a height of 120 ft (36 m) and a diameter of 8–12 in. (20–30 cm). Most are woody; a few are herbaceous or climbing. The economic uses of bamboo are numerous and varied. The seeds and young shoots are used as food and the leaves make excellent fodder for cattle. In varying sizes, the stems are used for pipes, timber, masts, bows, furniture, bridges, cooking vessels, buckets, wickerwork, paper pulp, cordage, and weaving. Entire houses are made of bamboo stems. Certain bamboos have been naturalized in California, Louisiana, and Florida. *See* CYPERALES.

Perry D. Strausbaugh/Earl L. Core

Banana

A large tropical plant; also its edible fruit, which occurs in hanging clusters, is usually yellow when ripe, and is about 6–8 in. (15–20 cm) long. The banana belongs to the family Musaceae. The banana of commerce (*Musa sapientum*), believed to have originated in the Asian tropics, was one of the earliest cultivated fruits. For commercial production the plant requires a tropical climate within the temperature range 50–105°F (10–40°C) and a constant supply of moisture by rainfall or irrigation. Bananas are subject to mechanical injury by strong winds which tear the leaves or blow down the plants. *See* IRRIGATION (AGRICULTURE); ZINGIBERALES.

The plant portion above the ground is a false stem (pseudostem) consisting of several concentrically formed leaves, from the center of which develops the inflorescence stalk (**Fig. 1**). The rhizome or true stem is underground. Near the tip of the flower stalk are several groups of sterile male flowers subtended by brilliant purple bracts. The lower female flower clusters on the same stalk give rise to the fruit and contain aborted stamens (male organs). The single fruits are called fingers, a single group of 8–12 fingers is termed a hand, and the several (6–18) hands of the whole inflorescence make up the stem.

Fig. 1. Pseudostem of commercial banana plant (*Musa sapientum*), showing characteristic foliage and single stem of bananas.

After the single fruiting stalk has produced the fruit bunch, the whole pseudostem is cut off at the ground, allowing one of several new buds or suckers from the underground stem to develop into a new plant. The fruit bunch requires 75–150 days to mature and must be removed from the plant to ripen properly. Chilled banana fruits do not soften normally; hence for best edibility the fruit is kept well ventilated at room temperature.

Kinds and distribution. Banana fruits of commerce set without pollination, by parthenocarpy, and hence are seedless. When mature, most varieties are yellow, although fine red-skinned types are well known. There are several hundred varieties grown throughout the world. The Cavendish banana (*M. nana*, variety Valery) is becoming important in the American tropics. The more starchy bananas, known as plantains, must be cooked before they can be eaten.

The yearly world production averages nearly 100,000,000 bunches, over 60% of which is consumed in the United States. Three-fourths of the world's bananas is grown in the Western Hemisphere. The greatest production occurs in Ecuador, Guatemala, Honduras, and other tropical Central and South American countries. Commercial production of bananas has been attempted in Florida, but climatic conditions there do not allow continuous economical cultivation of the fruit. *See* FRUIT; FRUIT, TREE.

Charles A. Schroeder

Diseases. F. L. Wellman indicates that there are between 200 and 220 diseases of bananas in the tropics and some 80 organisms have been associated with banana diseases, but only 3 or 4 have been involved in major epidemics. A new form of fungus leaf spot more damaging than the Sigatoka disease has occurred in several areas and has destroyed many plantings before adequate control measures could be undertaken. This disease is referred to as the black Sigatoka. The less destructive but important Sigatoka disease occurs in most banana-growing districts, and it becomes a major factor in banana losses in the high-rainfall areas. Without adequate controls and somewhat resistant cultivars, bananas could not be grown commercially. The Panama disease, caused by a soil fungus, has forced plant breeders to produce new clones so that growers could cope with the rapid spread and destruction of susceptible bananas. New varieties such as Valery have replaced the once popular Gros Michel in Central and South America. A bacterial disease called Moko has caused widespread damage in many banana-growing areas of Central and South America, and has made it necessary for growers to seek soils that have not been previously planted to bananas and to use disease-free

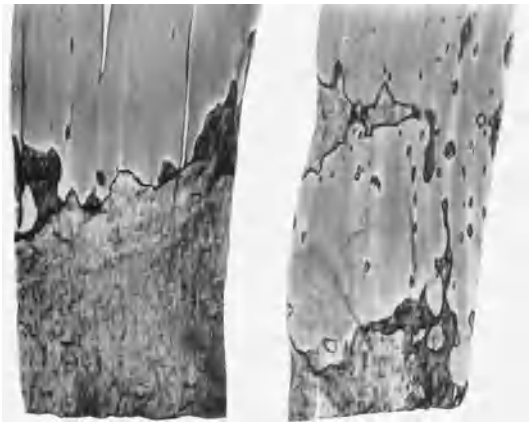

Fig. 2. Sigatoka disease, banana with spotting and necrosis of marginal areas on leaves of Gros Michel variety. (*Division of Tropical Research, United Fruit Co.*)

planting stock, and to follow strict sanitary cultural practices in order to maintain plantings. Tremendous efforts by all segments of the banana industry, coupled with millions of dollars spent on research and development, have made it possible to produce bananas economically despite the devastating effects of these diseases.

Etiological agents. Diseases of the banana roots, rhizome, pseudostem, leaves, and fruit are caused by 7 species of bacteria, 57 fungi, 2 viruses, and 14 nematodes. The most important and destructive diseases in commercial plantings include: Sigatoka disease or leaf spot, caused by the fungus *Mycosphaerella musicola* (**Fig. 2**) and its close but more virulent relative *M. fijiensis* var. *difformis*; Fusarium wilt or Panama disease, caused by the soil fungus *Fusarium oxysporum* f. sp. *cubense*; Moko disease or bacterial wilt, caused by *Pseudomonas solanacearum* (**Fig. 3**); bunchy top, caused by the aphid-transmitted bunchy top virus; infectious chlorosis, caused by a strain of cucumber mosaic virus; and rhizome and root rots, caused by the burrowing nematode (*Radopholus similis*) and other nematodes. Fruit spotting and decay are important in postharvest handling and shipping, often to distant markets. Many fungi are involved; some are wound parasites so that care should be taken to avoid injuries.

Fig. 3. Moko disease, with internal browning of banana fruit (Gros Michel variety).

Pathogenic forms and host reactions. Many of the organisms causing banana diseases are specific types that attack only bananas and their near relatives, plantains and abaca. Some of these organisms form distinct races that may occur in different geographic areas and may attack different banana cultivars or clones and some of their mutants.

Control methods. Plant breeders have been busy worldwide producing an acceptable fruit that is resistant to the devastating diseases. *Musa* Gros Michel AAA (a commercially excellent type of banana) and its tetraploid dwarf *Musa* Highgate AAAA are both susceptible to Panama disease, whereas *Musa* Giant Cavendish AAA and *Musa* Dwaft Cavendish AAA are somewhat resistant. All varieties of these genetic types are susceptible to the Sigatoka disease. If Gros Michel is crossed with a wild-type (*M. acuminata*), the progeny is resistant to both Panama and Sigatoka diseases but is not commercially acceptable. Some resistance to the burrowing nematode has been found in the cultivar Pisang Jari Buaya, and selections of the bluggoe-type plantain have been used in breeding because of its resistance to both the Panama and Moko diseases.

Despite early banana breeding programs and the awareness of resistance to Panama disease in the triploids of the Cavendish-Lacatan clones, these types were not used extensively because of their poor shipping and storage qualities. Studies in fruit physiology and handling procedures have made it possible to use these and other resistant clones, and Panama disease is no longer important in most banana-growing regions. Black Sigatoka and Sigatoka leaf spots continue to take their tolls unless protective fungicides are used on a regular basis, since all commercial banana types are susceptible to these leaf spots. In some humid areas, a foliar application of fungicide may be necessary on a weekly basis. Commercial companies have pioneered in the development of sophisticated equipment to apply control chemicals. Benzimidazoles, chlorothalonils, dithiocarbamates, copper compounds, imidazoles, and oils have all been used effectively against the leaf diseases, but fungicides do not weather well and the fungicidal activity is soon lost to heavy rainfall in many banana-growing areas. There is some hope that systemic fungicide formulations will give additional control in the future. Presently used banana clones are susceptible to Moko disease and many of the species of nematodes. These diseases are controlled

by sanitation and the use of disease-free planting stock. Nematicides have been effective in reducing populations in the field and eliminating them from the offshoots used as planting stock. The inability to "cure" virus diseases has made it necessary to eradicate diseased plants, eliminate weed hosts, and control the insect vectors. Storage and transit losses are reduced by controlling the post-harvest environment by using fungicidal dips and cut-surface applications, polyethylene wraps, wax coatings, and careful handling and boxing. *See* PLANT PATHOLOGY.

J. E. Dimitman

Bibliography. N. W. Simmonds, *Bananas*, 2d ed., 1982; R. H. Stover, *Banana, Plantain, and Abaca Diseases*, 1972.

Band spectrum

A spectrum consisting of groups or bands of closely spaced lines. Band spectra are characteristic of molecular gases or chemical compounds. When the light emitted or absorbed by molecules is viewed through a spectroscope with small dispersion, the spectrum appears to consist of very wide asymmetrical lines called bands. These bands usually have a maximum intensity near one edge, called a band head, and a gradually decreasing intensity on the other side. In some band systems the intensity shading is toward shorter waves, in others toward longer waves. Each band system consists of a series of nearly equally spaced bands called progressions; corresponding bands of different progressions form groups called sequences.

Six spectra of diatomic molecular fragments are shown in the **illustration**. The spectrum of a discharge tube containing air at low pressure is shown in illus. *a*. It has four band systems: the γ-bands of nitrogen oxide (NO, 230–270 nanometers), negative nitrogen bands (N_2^+, 290–350 nm), second-positive nitrogen bands (N_2, 290–500 nm), and first-positive nitrogen bands (N_2, 550–700 nm). The spectrum of high-frequency discharge in lead fluoride vapor in *b* has bands in prominent sequences. The spectrum in *c* shows part of one band system of SbF, and was obtained by vaporizing SbF into active nitrogen. Emission from a carbon arc cored with barium fluoride (BaF_2) and absorption of BaF vapor in an evacuated steel furnance are illustrated in *d*. These spectra were obtained in the second order of a diffraction grating, as were the spectra in *e* and *f*. The photograph *e* is that of the CN band at 388.3 nm from an argon discharge tube containing carbon and nitrogen impurities, and *f* is a band in ultraviolet spectrum of NO, obtained from glowing active nitrogen containing a small amount of oxygen.

When spectroscopes with adequate dispersion and resolving power are used, it is seen that most of the bands obtained from gaseous molecules actually consist of a very large number of lines whose spacing and relative intensities, if unresolved, explain the appearance of bands of continua (parts *e* and *f* of the illustration). For the quantum-mechanical explanations of the details of band spectra *See* MOLECULAR STRUCTURE AND SPECTRA.

W. F. Meggers; W. W. Watson

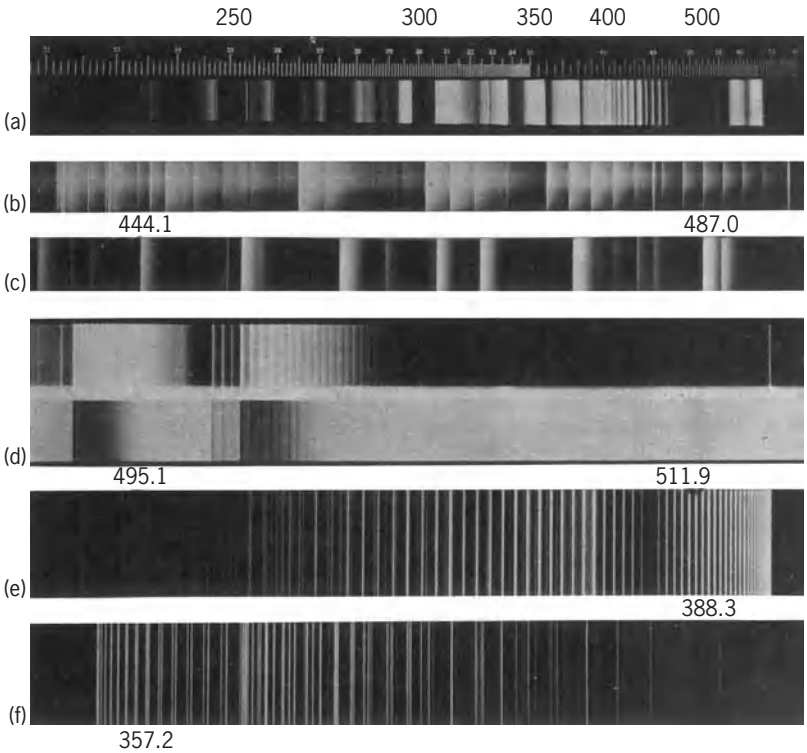

Photographs of band spectra of (*a*) a discharge tube containing air at low pressure; (*b*) high-frequency discharge in lead fluoride vapor; (*c*) SbF (*b* and *c* taken with large quartz spectrograph, after Rochester); (*d*) BaF emission and absorption; (*e*) CN; and (*f*) NO. The measurements are in nanometers. (*From F. A. Jenkins and H. E. White, Fundamentals of Optics, 3d ed., McGraw-Hill, 1957*)

Band theory of solids

A quantum-mechanical theory of the motion of electrons in solids. Its name comes from the fact that it predicts certain restricted ranges, or bands, for the energies of electrons in solids.

Suppose that the atoms of a solid are separated from each other to such a distance that they do not interact. The energy levels of the electrons of this system will then be those characteristic of the individual free atoms, and thus many electrons will have the same energy. Now imagine the atoms being slowly brought closer together. As the (equal) distance between atoms is decreased, the electrons in the outer shells begin to interact with each other. This interaction alters their energy and, in fact, broadens the sharp energy level out into a range of possible energy levels called a band. One would expect the process of band formation to be well advanced for the outer, or valence, electrons at the observed interatomic distances in solids. Once the atomic levels have spread into bands, the valence electrons are not confined to individual atoms, but may jump from atom to atom with an ease that increases with the increasing width of the band.

Although energy bands exist in all solids, the term

energy band is usually used in reference only to ordered substances, that is, those having well-defined crystal lattices. In such a case, a basic theorem due to F. Bloch asserts that an electron energy state can be classified according to its crystal momentum **p** or its electron wave vector $\mathbf{k} = \mathbf{p}/\hbar$ (where $\hbar$ is Planck's constant h divided by 2π). If the electrons were free, the energy of an electron whose wave vector is **k** would be as shown in Eq. (1), where E_0 is the energy of the lowest state of a valence electron and m_0 is the electron mass. In a crystal, although the electrons may move fairly independently, they are not free because of the effect of the crystal binding and the forces exerted on them by the atoms; consequently, the relation $E(\mathbf{k})$ between energy and wave vector is more complicated. The statement of this relationship constitutes the description of an energy band. $E(\mathbf{k})$ can be calculated as the eigenvalue of a one-electron Schrödinger equation. *See* BLOCH THEOREM.

$$E(\mathbf{k}) = E_0 + \frac{\hbar^2\mathbf{k}^2}{2m_0} \tag{1}$$

A knowledge of the energy levels of electrons is of fundamental importance in computing electrical, magnetic, optical, or thermal properties of solids.

Allowed and forbidden bands. The ranges of possible electron energy levels in a solid are called allowed energy bands. It often happens that there are also ranges of energy which it is impossible, for all directions of motion, for an electron to have in a given crystal. Such ranges are called forbidden bands, or gaps. The allowed energy bands sometimes overlap and sometimes are separated by forbidden bands. The presence of a forbidden band, or energy gap, immediately above the occupied allowed states (such as the region A to B in the **illustration**) is the principal difference in the electronic structures of a semiconductor or insulator and a metal. In the first two substances there is a gap for all directions of motion, between the valence band or normally occupied states and the conduction band, which is normally unoccupied. In a metal there is no overall gap between occupied and unoccupied states.

The presence of a gap means that the electrons cannot easily be accelerated into higher energy states by an applied electric field. Thus, the substance cannot carry a current unless electrons are excited across the gap by thermal or optical means.

Effective mass. When an external electromagnetic field acts upon the electrons in a solid, the resultant motion is not what one would expect if the electrons were free. In fact, the electrons behave as though they were free but with a different mass, which is called the effective mass. This effective mass can be obtained from the dependence of electron energy on the wave vector, $E(\mathbf{k})$, in the following way.

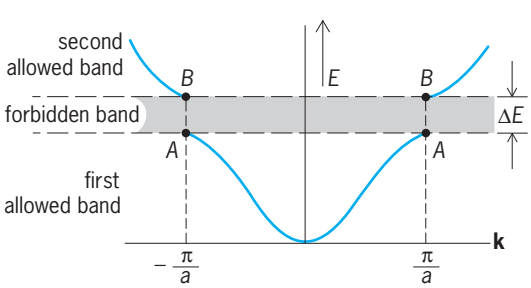

Electron energy *E* versus wave vector k for a monatomic linear lattice of lattice constant *a*. (*After C. Kittel, Introduction to Solid State Physics, 6th ed., Wiley, 1986*)

Suppose there is a maximum or minimum of the function $E(\mathbf{k})$ at the point $\mathbf{k} = \mathbf{k}_0$. The function $E(\mathbf{k})$ can be expanded in a Taylor series about this point. For simplicity, assume that $E(\mathbf{k})$ is a function of the magnitude of **k** only, that is, is independent of the direction of **k**. Then, by dropping terms higher than second order in the Taylor series, Eq. (2) results. By

$$E(\mathbf{k}) = E(\mathbf{k}_0) + {}^1/_2(\mathbf{k} - \mathbf{k}_0)^2 \left(\frac{d^2E}{dk^2}\right)_{\mathbf{k}_0} \tag{2}$$

analogy with Eq. (1), a quantity m^* with the dimensions of a mass can be defined by relation (3).

$$\frac{\hbar^2}{m^*} = \left(\frac{d^2E}{dk^2}\right)_{\mathbf{k}_0} \tag{3}$$

The quantity m^* is called the effective mass of electrons at $\mathbf{k}_0$. For many simple metals, the average effective mass is close to m_0, but smaller effective masses are not uncommon. In indium antimonide, a semiconductor, the effective mass of electrons in the conduction band is 0.013 m_0. In a semiclassical approximation, an electron in the solid responds to an external applied electric or magnetic field as though it were a particle of mass m^*. The equation of motion of an electron is shown in Eq. (4), where **v** is the electron velocity, **E** the electric field, **B** the magnetic induction, and e the charge of the electron.

$$m^*\frac{d\mathbf{v}}{dt} = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{4}$$

Quite frequently, the energy $E(\mathbf{k})$ does depend upon the direction of **k**. In such a case, the effective mass is a tensor whose components are defined by Eq. (5). Equation (4) remains valid with a tensor m^*. Bismuth is an example of a metal in which the effective mass depends strongly on direction.

$$\frac{\hbar^2}{m_{ij}^*} = \left(\frac{\partial^2E}{\partial k_i \partial k_j}\right)_{\mathbf{k}_0} \tag{5}$$

Transitions between states. Under external influences, such as irradiation, electrons can make transitions between states in the same band or in different bands. The interaction between the electrons and the vibrations of the crystal lattice can scatter the electrons in a given band with a substantial change in the electron momentum, but only a slight change in energy. This scattering is one of the principal causes of the electrical resistivity of metals. *See* ELECTRICAL RESISTIVITY.

An external electromagnetic field (for example, visible light) can cause transitions between different bands. In such a process, crystal periodicity leads to the requirement that the sum of the electron crystal momentum $\hbar\mathbf{k}$ and the photon momentum must be conserved. Because the momentum of a photon $h\nu/c$ (where ν is the frequency of the light and c its velocity) is quite small, the momentum of the

electron before and after collision is nearly the same. Such a transition is called vertical in reference to an energy band diagram. Conservation of energy must also hold in the transition, so absorption of light is possible only if there is an unoccupied state of energy $h\nu$ available at the same $\mathbf{k}$ as the initial state. These transitions are responsible for much of the absorption of light by semiconductors in the visible and near-infrared region of the spectrum.

Energy band calculation. As is the case for any quantum-mechanical system, the energy levels of electrons in solids are determined in principle by the solution of the Schrödinger wave equation for the system. However, the enormous number of particles involved makes exact solution of this equation impossible. Approximate methods must be employed. The atoms are regarded as fixed in their equilibrium positions. Each electron is regarded as having an individual wave function $\psi_n(\mathbf{k}, \mathbf{r})$, in which $\mathbf{k}$ is the wave vector and the index n designates a particular band. The wave function $\psi_n(\mathbf{k}, \mathbf{r})$ is frequently called a Bloch function. The wave function of the many-electron system is written as a determinant of Bloch functions to satisfy the requirements of the Pauli exclusion principle. The general variational method of quantum mechanics may be employed to derive an equation for the individual Bloch functions. This equation, which is known as the Hartree-Fock equation, is similar to a one-electron Schrödinger equation in which each electron moves in the field of all the nuclei of the system and the average field of all the other electrons. An additional term, exchange, takes account of the reduction of the average electronic repulsion, because the probability of close encounters of electrons of parallel spin is reduced by the Pauli exclusion principle.

The Hartree-Fock equations are still quite complicated, and must be solved by the self-consistent field method. In this approach, some reasonable distribution of electrons is assumed to exist. An effective potential $V(\mathbf{r})$ can be calculated from this distribution, including the contribution from the nuclei. Usually it is an adequate approximation to include the exchange terms in this potential. Then the Bloch functions may be formed by solving Eq. (6), in which

$$-\frac{\hbar^2}{2m_0}\nabla^2\psi_n(\mathbf{k}, \mathbf{r}) + V(\mathbf{r})\psi_n(\mathbf{k}, \mathbf{r}) = E_n(\mathbf{k})\psi_n(\mathbf{k}, \mathbf{r}) \quad (6)$$

$V(\mathbf{r})$ is the potential described above and $E_n(\mathbf{k})$ is the energy of an electron in band n having wave vector $\mathbf{k}$. The potential $V(\mathbf{r})$ is periodic in space with the periodicity of the crystal. The wave function $\psi_n(\mathbf{k}, \mathbf{r})$ obtained from Eq. (6) yields a new electron distribution from which the potential $V(\mathbf{r})$ may be calculated again. The process is repeated until the potential used in the solution of Eq. (6) agrees with that obtained from the solution $\psi_n(\mathbf{k}, \mathbf{r})$ to sufficient accuracy.

The local density approximation is frequently used to obtain this potential. In the simplest formula of this type, the Kohn-Sham potential, $V(\mathbf{r})$ is given by Eq. (7). The first term represents the potential energy of an electron in the field of all the nuclei of the system. These nuclei are located at sites $\mathbf{R}_\mu$ and have charge Z_μ. The second term contains the average electronic repulsion of the electron distribution, and $\pi(\mathbf{r})$ is the electron density. The third term approximately describes the exchange interaction, referred to above.

$$V(\mathbf{r}) = -e^2\sum_\mu \frac{Z\mu}{|\mathbf{r}-\mathbf{R}_\mu|} + e^2\int\frac{p(\mathbf{r}')d^3r'}{|\mathbf{r}-\mathbf{r}'|} - 2e^2\left(\frac{3\rho(\mathbf{r})}{8\pi}\right)^{1/3} \quad (7)$$

The self-consistent-field procedure is evidently quite complicated. Since it may not be practical to repeat the calculations to obtain a self-consistent solution, it is desirable to choose an initial $V(\mathbf{r})$ which approximates the actual physical situation as closely as possible. Choice of an adequate crystal potential is the chief physical problem in the calculation of energy levels of electrons in solids.

Several techniques are available for solving wave equation (6) with a given $V(\mathbf{r})$. Those in common use include the Green's function method, the augmented plane wave method, the orthogonalized plane wave method, and the linear combination of atomic orbitals method. It is also possible to use experimentally determined energy levels and effective masses to obtain a suitable potential for use in Eq. (6). This procedure, which bypasses many of the difficulties of the self-consistent-field approach, is known as the pseudopotential method and is now widely used. In simple metals (not transition metals) and wide-band semiconductors, such as germanium and silicon, the pseudopotential is rather weak, and many properties can be calculated by using perturbation theory.

Density of states. Many properties of solids, including electron specific heat and optical and x-ray absorption and emission, are related rather simply to a basic function known as the density of states. This function, here denoted $G(E)$, is defined so that the number of electronic energy states in the range between E and $E + dE$ is $G(E)dE$. It can be shown that $G(E)$ is given by Eq. (8), in which $E(\mathbf{k})$ is the energy band function, δ is the Dirac delta function, Ω is the volume of a unit cell of the solid, and the constant multiplying the integral has been chosen so that the number of states is two (one for each spin direction) for each energy band. The first integral in Eq. (8) is taken over the Brillouin zone, the second over a surface of constant energy E in the zone. The density of states will show structure in the neighborhood of energies corresponding to points where $|\nabla_k E(\mathbf{k})|$ vanishes. Such points, which are required to exist by reasons of crystal symmetry and periodicity, are known as Van Hove singularities. The energies of states at which Van Hove singularities occur can be

$$G(E) = \frac{\Omega}{4\pi^3}\int \delta[E - E(\mathbf{k})]\, d^3k = \frac{\Omega}{4\pi^3}\int\frac{dS_k(E)}{|\nabla_k E(\mathbf{k})|} \quad (8)$$

determined from optical and (sometimes) x-ray measurements.

Experimental information. A considerable amount of experimental information has been obtained concerning the band structures of the common metals and semiconductors. In metals, experiments usually determine properties of the Fermi surface (the surface in **k** space bounding the occupied states at zero temperature), since only electrons in states near the top of the Fermi distribution are able to respond to electric and magnetic fields. For example, the magnetic susceptibility of a metal exhibits a component which is an oscillatory function of the reciprocal of the magnetic field strength when the field is strong and the temperature is quite low, typically a few kelvins (the de Haas–van Alphen effect). Measurement of the period of these oscillations determines the area of a cross section of the Fermi surface in a plane perpendicular to the direction of the magnetic field.

Other properties of the Fermi surface are obtained by studying the increase of the electrical resistance in a magnetic field, and from the magnetic field dependence of ultrasonic attenuation. The density of states (number of electron states per unit energy) and effective mass at the Fermi energy can be found from measurements of the electron contribution to the specific heat. Effective masses of electrons and holes can be determined by cyclotron resonance measurements in both semiconductors and metals.

Optical and x-ray measurements enable the determination of energy differences between states. The smallest width of the characteristic energy gap between valence and conduction bands in a semiconductor can be determined by measuring the wavelength at which the fundamental absorption begins. Application of a strong uniform magnetic field will produce an oscillatory energy dependence of this absorption, from which a precise determination of effective masses is possible.

In both semiconductors and metals, states removed in energy from the Fermi level can be studied optically. Measurements of the reflectivity will show structure at energies corresponding to transitions at symmetry points of the Brillouin zone where Van Hove singularities are likely to occur. To reveal such structure clearly against a background which frequently may be large, it is useful to subject the solid to a periodic perturbation, possibly an electric field, stress, or a change in temperature. This modifies the band structure slightly, particularly near Van Hove singularities, which tend to be sensitive to perturbations. As a result, the optical properties of the sample are slightly modulated, and this modulation is detected directly. It is possible to apply this technique in the soft x-ray wavelength region as well. Modulation of the soft x-ray emission by alternating stress (the piezo soft x-ray effect) enables the determination of the energies of transitions between band and deep core states.

Photoemission measurements can also yield information about energy bands below the Fermi energy. In these experiments, an electron is excited by a photon from a state of energy, E_i, below the Fermi energy to a state of high enough energy, E_f, such that the electron can escape from the crystal, usually from within a few tenths of a nanometer from the surface. The relation between the energies is $E_f = E_i + \hbar\omega$, where ω is the angular frequency associated with the light. The number of emitted electrons is studied as a function of the angle between the direction of the outgoing electrons and that of the incident photon (often chosen to be perpendicular to the surface). The polarization of the light can also be controlled. If the surface of a single crystal sample is plane, the component of the electron momentum parallel to the solid ($\mathbf{k}_\parallel$) is conserved. The measurement is then sensitive to the distribution of electrons in energy along a line in the Brillouin zone parallel to the surface normal and with the specified value of $\mathbf{k}_\parallel$. The results can be interpreted to give the positions of the energy band in the initial state. It is also possible to determine energies of electrons localized in the surface region.

Limitations of band theory. The results of energy-band calculations for ordinary metals usually predict Fermi surfaces and other properties that agree rather well with experiment. In addition, cohesive energies and values of the lattice constant in equilibrium can be obtained with reasonable accuracy, and, in the case of ferromagnetic metals, calculated magnetic moments agree with experiment. However, there are significant discrepancies between theoretical calculations and experiments for certain types of systems, as follows.

1. There are a number of metallic compounds containing rare-earth or actinide elements ($CeCu_6$ and UPt_3 are examples) in which the electron contribution to the specific heat is extremely large. These materials are known as heavy-fermion systems. If the specific heat is interpreted in terms of an electron effective mass, the values obtained may range from several hundred to more than a thousand times the free-electron mass. Band calculations have been successful in predicting the shapes of the observed Fermi surfaces, but cannot account for the large masses.

2. Band calculations for semiconductors such as silicon, germanium, and gallium arsenide (GaAs) predict values for the energy gap between valence and conduction bands in the range one-half to two-thirds of the measured values.

3. The occurrence of superconductivity at low temperatures in a metal indicates a breakdown of single-particle physics due to the formation of a collective or condensed state of electron pairs. It cannot be predicted, or even be adequately described phenomenologically within energy band theory. However, the single-particle states furnish a useful framework for the construction of a theory of electron pairing, and they are useful in describing the normal state above the transition temperature.

4. There are many systems, especially compounds of $3d$ transition elements, for which band calculations predict that some bands are only partially filled, and therefore should exhibit metallic behavior, but which in fact are insulators with antiferromagnetic structures. Cobaltous oxide (CoO) is an example. Materials of this type are called Mott-Hubbard insulators.

They are better described in terms of localized electron states on the atoms or ions in the system, rather than in terms of extended, band states. *See* SUPERCONDUCTIVITY.

In all the cases discussed, the failures of band theory are attributed to an inadequate treatment of strong electron-electron interactions. For example, in the Mott-Hubbard insulators, electron interactions are believed to lead to electron localization, and the occurrence of atomic magnetic moments. In conventional superconductors, the interaction of electrons with lattice vibrations is believed to lead to an effective attractive interaction between electrons, which in turn leads to the formation of pairs. The origin of the interactions leading to high-temperature superconductivity (as in $YBa_2Cu_3O_7$) is not yet understood. *See* BRILLOUIN ZONE; COHESION (PHYSICS); CRYSTAL STRUCTURE; EXCITON; FERROMAGNETISM; FREE-ELECTRON THEORY OF METALS; HOLE STATES IN SOLIDS; KRONIG-PENNEY MODEL; NONRELATIVISTIC QUANTUM THEORY; PHOTOEMISSION; SEMICONDUCTOR; SPECIFIC HEAT OF SOLIDS. Joseph Callaway

Bibliography. P. W. Anderson, *Concepts in Solids*, 1964; N. W. Ashcroft, *Solid State Physics*, 2d ed., 2000; R. H. Bube, *Electrons in Solids: An Introductory Survey*, 3d ed., 1992; G. Burns, *Solid State Physics*, 1985; J. Callaway, *Quantum Theory of the Solid State*, 2d ed., 1991; J. R. Hook and H. E. Hall, *Solid State Physics*, 2d ed., 1995; W. Jones and N. H. March, *Theoretical Solid State Physics*, 1973, reprint 1985; C. Kittel, *Quantum Theory of Solids*, 2d ed., 1987.

Banded iron formation

A sedimentary rock that was commonly deposited during the Precambrian. It was probably laid down as a colloidal iron-rich chemical precipitate, but in its present compacted form it consists typically of equal proportions of iron oxides (hematite or magnetite) and silica in the finely crystalline form of quartz known as chert. Its chemical composition is 50% silicon dioxide (SiO_2) and 50% iron oxides (Fe_2O_3 and Fe_3O_4), to give a total iron content of about 30%. Banding is produced by the concentration of these two chemical components into layers about 1–5 cm (1/2–2 in.) thick; typical banded iron formation consists of pale silica-rich cherty bands alternating with black to dark red iron-rich bands (**Fig. 1**). These contrasting layers are sharply defined, so that the rock has a striped appearance; banded iron formation is normally a hard, tough rock, highly resistant both to erosion and to breaking with a hammer.

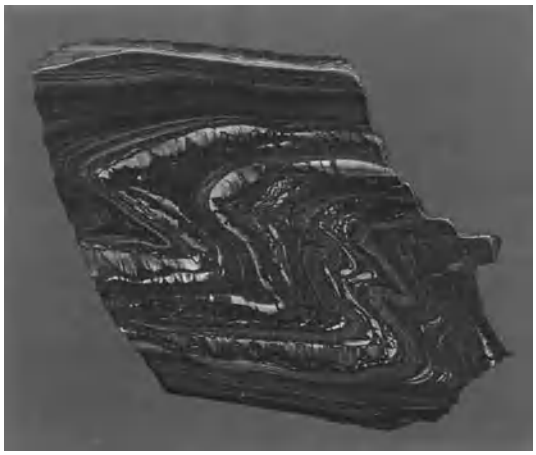

Fig. 1. Folded banded iron formation from the Ord Range, Western Australia. The distance between top and bottom of the polished face of the sample is about 15 cm (6 in.). Chert jasper bands alternate with dark magnetite-rich bands. The thin pale layers of irregular thickness are bands of asbestiform amphibole, now replaced by silica, to give the semiprecious material "tiger-eye." (*Photo courtesy of John Blockley*)

Varieties and nomenclature. Historically, different names were applied to banded iron formations in different regions. The itabirites of Brazil, the ironstones of South Africa, and the banded hematite quartzites of India are now generally identified as banded iron formations. The iron formations of the Lake Superior area ranges (such as Marquette, Gogebic, Cuyuna, Mesabi, and Gunflint), which were the first to be systematically described, are a special case. In these the banding is usually coarser and less regular, and much of the constituent material consists of fine (1–2 mm) rounded grains of both silica and iron oxides, giving rise to the name granule or pelletal iron formation; the extensive banded iron formations of Labrador are also of this type. This rock is sometimes called taconite by mining engineers.

Global distribution and ages. Banded iron formation occurs in the Precambrian of all continents. The larger individual occurrences lie mainly in the southern continents and include the Carajas and Minas Gerais banded iron formations of Brazil, the Transvaal Supergroup banded iron formations of South Africa, and the Hamersley Group banded iron formation of northwestern Australia. All of these are individually larger than such Northern Hemisphere banded iron formations as those of North America (Labrador and the Lake Superior ranges), Ukraine (Krivoi Rog), west Africa (Mauritania and Liberia), India (Bababudan), and China (An Shan). While banded iron formations are generally restricted to the Precambrian, more specifically they range in age from about 3800 million years ago (Ma; Isua, Greenland) to less than 1000 Ma (Rapitan, western Canada; Damara Belt, Namibia; Urucum, Brazil). An exceptional banded iron formation also occurs in rocks of Devonian age (350 Ma) in the Altai area of Kazakhstan. Although banded iron formation deposition took place intermittently throughout this vast time range, banded iron formation was particularly abundant, and formed individually larger deposits, between about 2500 Ma and 2000 Ma.

Associated rocks, deposit dimensions, and band continuity. Banded iron formation may be stratigraphically associated in the same depositional basin with a variety of other sedimentary rock types, such as shale, sandstone, and carbonate (dolomite), as well as with volcanic rocks. However, in any one basin, banded iron formation tends to be concentrated in one or two stratigraphic levels, rather than being deposited repetitively throughout the life of the basin. Individual banded iron formations may be as thin

as a few meters or over a kilometer thick, and may be traceable along their bedding from distances of one hundred to several hundred kilometers. The minimum depositional area of the Hamersley Group banded iron formations of northwestern Australia has been shown to be 150,000 km^2 (58,000 mi^2), and the banded iron formations have a total thickness of 700 m (2300 ft). Some thin (about 0.2–2.0 mm) repeated laminae, or microbands, of these banded iron formations have been shown to extend continuously over the whole area of the basin. It has been proposed that each microband, composed of an iron-rich and an iron-poor layer, represents a single year of deposition, and this hypothesis has now been shown to be consistent with the rate of banded iron formation deposition in the basin measured by isotope geochronology.

Origin. The mechanism of formation of most common sedimentary rocks, such as sandstone or limestone, is evident from observation of identical materials now being deposited. However, banded iron formation is not now occurring. A model for its origin must therefore be built up from the application of basic physical and chemical principles to detailed observation of its structure and composition. Early hypotheses mostly appealed to such concentrated sources of iron as local volcanic activity, or iron-rich river discharges resulting from an unusual Precambrian weathering regime. But recent studies suggest that the deposition of even very large banded iron formations could have been precipitated from ocean water containing as little as 10 parts per million iron, with about 30% of this precipitated each year from the water of local offshore basins. In the early Precambrian, when the oxygen content of the atmosphere and oceans was very low, upwelling could have provided an abundant supply of dissolved ferrous iron, leached from ocean-floor volcanic rocks. Seasonal oxidation by photosynthesizing algae in the near-surface water provides a credible precipitation mechanism, and the resultant insoluble ferric precipitate would have accumulated on the basin floor in periods of tectonic stability. This process may have provided a major sink for the oxygen produced by the earliest photosynthesizers, which would have lacked biochemical mechanisms to survive in an oxidizing environment, and whose biological productivity would have been restricted by the need to dispose of the oxygen they produced. The abundance of banded iron formation in the 2500–2000-Ma period may be related to the development of such mechanisms, while the later decrease in abundance probably reflects the fall of ocean water iron concentration to its present low level.

Economic importance. The world's iron and steel industry is based almost exclusively on iron ores associated with banded iron formation. Banded iron formation itself may be the primary ore, from which hematite or magnetite is concentrated after crushing. But the main ore now mined globally is high-grade (greater than 60% iron) material that formed within banded iron formation by natural leaching of its silica content.

Alec Francis Trendall

Bibliography. P. W. U. Appel and G. L. La Berge (eds.), *Precambrian Iron-Formations*, Theophrastus Publications, Athens, 1987; H. D. Holland and A. F. Trendall (eds.), *Patterns of Change in Earth Evolution*, Springer-Verlag, Berlin, 1984; A. F. Trendall and J. G. Blockley, Iron formations of the Precambrian Hamersley Group of Western Australia, with special reference to crocidolite, *Western Australia Geol. Surv. Bull.*, no. 119, 1970; A. F. Trendall and R. C. Morris (eds.), *Iron-Formation: Facts and Problems*, Developments in Precambrian Geology 6, Elsevier, Amsterdam, 1983.

Bandwidth requirements (communications)

The channel bandwidths needed to transmit various types of signals, using various processing schemes.

Propagation conditions for communication systems that transmit signals through the atmosphere or through free space are strongly dependent on the transmission frequency. Government regulations specify the modulation type, the bandwidth, and the type of information that can be transmitted over designated frequency bands. In the United States, the Federal Communications Commission (FCC) regulates radio, telephone, television, and satellite transmission systems. In North America, standards are developed by the Institute of Electrical and Electronics Engineers (IEEE), by Committee TI–Telecommunications (accredited by the American National Standards Institute, ANSI), and by the Telecommunications Industry Association (TIA). Internationally, frequency assignment and technical standards are developed by the International Telecommunications Union–Telecommunications Standardization Sector (ITU-T) and the International Telecommunications Union–Radiocommunication Sector (ITU-R), functioning under the auspices of the United Nations International Telecommunications Union (ITU). *See* RADIO SPECTRUM ALLOCATION.

Definition of bandwidth. Every signal observed in practice can be expressed as a sum (discrete or over a frequency continuum) of sinusoidal components of various frequencies. The plot of the amplitude versus frequency constitutes one feature of the frequency spectrum (the other being the phase versus frequency). The difference between the highest and the lowest frequencies of the frequency components of significant amplitudes in the spectrum is called the bandwidth of the signal expressed in the unit of frequency, hertz. Every communication medium (also called channel) is capable of transmitting a frequency band (spectrum of frequencies) with reasonable fidelity. Qualitatively speaking, the difference between the highest and the lowest frequencies of components in the band over which the channel gain remains reasonably constant (or within a specified variation) is called the channel bandwidth. *See* FOURIER SERIES AND TRANSFORMS; WAVEFORM.

The notions of significant amplitudes, reasonable fidelity, reasonably constant gain, and so forth, however, differ with different applications. Consequently, there are several quantitative definitions of bandwidths. The following bandwidth definitions

are commonly observed in practice:

1. *Half-power bandwidth*. This is the difference between the two frequencies at which the amplitude spectrum is 3 dB below the peak value (which is the same as saying that the power spectrum is half of the peak value).

2. *Specified power containment bandwidth*. This is the band of frequencies that contains 100 $(1-\epsilon)$% of the total signal power. The value of ϵ is usually of the order of 0.01 to 0.1 depending on the application.

3. *Bounded-power spectral density bandwidth*. This is the width of the frequency band outside which the power spectral density must fall at least a given level (such as 40 or 60 dB) below its maximum value.

4. *Null-to-null bandwidth*. This is equal to the width of the main lobe of the power spectral density.

5. *Equivalent noise bandwidth*. This is the width of a rectangle whose height is equal to the maximum value of the power spectral density of the signal and whose area is equal to the area under the power spectral density of the signal. For modulated signals, this figure is multiplied by 0.5.

In analog signals and systems, half-power bandwidth is commonly specified, whereas in digital communication the specification of the null-to-null bandwidth is the usual practice.

Signal conditioning. Clearly, to transmit a signal with reasonable fidelity over a communication channel, the channel bandwidth must match and be at least equal to the signal bandwidth. Proper conditioning of a signal, such as modulation or coding, however, can increase or decrease the bandwidth of the processed signal. Thus, it is possible to transmit the information of a signal over a channel of bandwidth larger or smaller than that of the original signal.

Amplitude modulation (AM) with double sidebands (DSB), for example, doubles the signal bandwidth. If the audio signal to be transmitted has a bandwidth of 5 kHz, the resulting AM signal bandwidth using DSB is 10 kHz. Amplitude modulation with a single sideband (SSB), on the other hand, requires exactly the same bandwidth as that of the original signal. Amateur radio and commercial and military services are examples of this scheme. Vestigial sideband (VSB) modulation is a compromise between DSB and SSB. It requires a bandwidth slightly larger (typically 20–30%) than that of the message signal. Broadcast television is an example of this scheme. The television signal (video and audio) bandwidth is 4.5 MHz. It is transmitted by VSB modulation that requires a bandwidth of 6 MHz. *See* AMPLITUDE MODULATION; AMPLITUDE-MODULATION RADIO; SINGLE SIDEBAND; TELEVISION.

DSB, SSB, and VSB modulation are examples of amplitude-modulation schemes. In broadcast frequency modulation (FM), on the other hand, audio signal bandwidth is 15 kHz (for high fidelity), but the corresponding frequency-modulated signal bandwidth is 200 kHz. In telephony, the audio signal bandwidth is approximately 3 kHz and is transmitted over local lines of 3 kHz bandwidth, or over long distances (using SSB) requiring a transmission bandwidth of 4 kHz per telephone channel (3 kHz for the signal and a guard band of 1 kHz). On the other hand, a digitized telephone signal (processed by analog-to-digital conversion) requires a bandwidth of 32–64 kHz. Facsimile typically requires a bandwidth of 5 kHz for AM and 25 kHz for FM. In telegraphy, the bandwidth for transmission of 100 words per minute is 170–400 Hz. *See* ANALOG-TO-DIGITAL CONVERTER; FACSIMILE; FREQUENCY MODULATION; FREQUENCY-MODULATION RADIO; TELEGRAPHY.

In pulse communication systems, the rate at which pulses can be transmitted over a channel is directly proportional to the channel bandwidth. The sampling theorem leads to the conclusion that in the absence of noise a channel can transmit correctly at most two independent pieces of information per second per unit bandwidth (in hertz). In most pulse communication systems, for example, pulse amplitude modulation (PAM) or pulse code modulation (PCM), the information carried by a pulse resides in the pulse amplitude. Consequently, the information from one pulse is equivalent to one piece of information. Hence for PAM and PCM signals, up to $2B$ pulses per second can be transmitted over a bandwidth of B Hz. This, however, is an upper (theoretical) limit, which can be attained only by using ideal filters (or noncausal pulses of infinite duration); this limiting case is, of course, unrealizable. In practice, the transmission rate remains below $2B$ pulses per second.

The rates discussed above result from the inherent limitations imposed by the channel bandwidth. In practice, a specific coding scheme imposes further restrictions on the rate. For example, the so-called bipolar (or pseudo-ternary) scheme of binary signaling requires a transmission bandwidth of B Hz to transmit B binary pulses per second. This is only half the rate that would be allowed by the channel bandwidth restriction alone. On the other hand, a duobinary (or partial response coding) scheme, represented in digital telephony by the alternate mark inverted (AMI) line code, can transmit $2B$ pulses per second over a channel of bandwidth B.

Advances in coding and modulation theory, as well as in the design of electronic circuitry, have made possible the implementation of modulation schemes more complex than those described above. The more complex schemes sometimes carry many bits of information per pulse and can adapt themselves automatically to best utilize the capacities of channels with impaired frequency responses.

Such modulation schemes lie at the heart of the growing family of digital subscriber systems (xDSL, where x is a letter representing the system type). Such systems allow the transmission of data along copper cable pairs of the telephone network's subscriber loop plant at relatively high rates, extending the useful life of the loop plant. *See* MODULATION; PULSE MODULATION; RADIO.

Signal bandwidth restriction. It can be shown that a signal cannot be simultaneously time-limited and band-limited. Thus the spectrum of a finite-duration

signal extends to infinity; consequently its bandwidth is infinite. Since all practical signals are of finite duration (time-limited), the bandwidth of any practical signal is infinite. The implication is that on any given channel, only one signal can be transmitted at any time; for example, on air waves, only one station can be broadcast at a time. Because of tremendous demand for communication in modern times, intolerable usage congestion would result. Fortunately, all practical signal spectra decay asymptotically with frequency and approach zero at infinite frequency. High-frequency components can be cut off, causing acceptable distortion in the signal waveform. Thus, every practical signal can be effectively band-limited to a certain essential bandwidth.

Effects of bandwidth restriction. Band-limiting a signal (cutting off high-frequency components) results in distortion of the signal. The effect of the distortion is often more serious for digital than analog signals. In analog systems, the distortion causes the signal quality to deteriorate somewhat and is not particularly significant. In digital communication systems, the signal is in the form of a sequence of pulses. It can be shown that the pulses are dispersed (spread out) when the channel bandwidth is restricted. Such a dispersion causes pulses to interfere with neighboring pulses. This interference (intersymbol interference) can be additive and leads to a higher probability of errors in pulse detection. In practice, equalizers (in the form of transversal filters) are used at the receiver to correct partly the damage to the signal resulting from channel bandwidth restrictions. *See* DISTORTION (ELECTRONIC CIRCUITS); ELECTRIC FILTER; EQUALIZER.

Shannon's equation. C. E. Shannon proved that over a channel of bandwith B the rate of information transmission, C, in bits/s (binary digits per second) is given by the equation below, where SNR is the

$$C = B\log_2(1 + \mathrm{SNR}) \quad \text{bits/s}$$

signal-to-noise power ratio. This result assumes a white gaussian noise, which is the worst kind of noise from the point of view of interference. For any other noise the rate of information transmission will be at least as large as that indicated by Shannon's equation. *See* INFORMATION THEORY; SIGNAL-TO-NOISE RATIO.

It follows from Shannon's equation that a given information transmission rate C can be achieved by various combinations of B and SNR. It is thus possible to trade B for SNR, and vice versa. In a communication system, one of the two resources (B or SNR) may be more precious than the other, and the communication scheme would be optimized accordingly. A typical voice-grade telephone circuit has a limited bandwidth (3 kHz) but a lot of power is available, whereas in communications satellites large bandwidths are available but power is scarce. Hence the communication schemes required in the two cases are usually different. *See* SPACE COMMUNICATIONS.

Broadbanding. A corollary of Shannon's equation is that, if a signal is properly processed to increase its bandwidth, the processed signal becomes more immune to interference or noise over the channel. This means that an increase in transmission bandwidth (broadbanding) can suppress the noise in the received signal, resulting in a better-quality signal (increased SNR) at the receiver. Frequency modulation and pulse-code modulation are two examples of broadband schemes where the transmission bandwidth can be increased as desired to suppress noise. The former transmits information in analog form (an analog scheme), while the latter digitizes the analog information and transmits the resulting digital data (a digital scheme). In general, digital schemes use increased bandwidth more efficiently than do analog schemes in suppressing noise. Amplitude modulation, on the other hand, is not a broadband scheme because its transmission bandwidth is fixed and cannot be increased.

Broadbanding is also used to make communication less vulnerable to jamming and illicit reception by using the so-called spread spectrum signal. The jamming signal can be considered as noise, and by using the proper technique of broadbanding (spread spectrum) the jamming signal can be rendered less effective (suppressed). Spread spectrum may use direct-sequence, frequency, time-hopping, and chirp systems. In spread spectrum systems, some signal function other than the information being sent is employed to determine the transmitted signal bandwidth. For this reason, conventional FM, where the signal bandwidth is determined only by the message signal, cannot be considered a spread spectrum system. *See* ELECTRICAL COMMUNICATIONS; ELECTRONIC WARFARE; SPREAD SPECTRUM COMMUNICATION.

B. P. Lathi; Maynard Wright

Bibliography. R. L. Freeman, *Reference Manual for Telecommunication Engineering*, 2d ed., Wiley, 1994; B. P. Lathi, *Modern Digital and Analog Communication Systems*, 3d ed., Oxford, 1998; T. Starr, J. M. Cioffi, and P. Silverman, *Understanding Digital Subscriber Line Technology*, Prentice Hall, 1999; M. E. Van Valkenburg, *Reference Data for Engineers: Radio, Electronics, Computer, and Communications*, 9th ed., 2001.

Barbiturates

A group of drugs widely used for the suppression of anxiety, the induction of sleep, and the control of seizures. Some of them, when injected intravenously, produce a general anesthesia.

Barbituric acid, from which the various barbiturate congeners come, was synthesized in 1864 by Adolf von Baeyer and named after St. Barbara. It is a malonyl urea. In the years that followed, a dozen or more closely related compounds obtained by adding or substituting various radicals to the general formula have been synthesized. The names of many of them have become familiar; examples are Phenobarbital, Meberal, Seconal, Nembutal, Amytal, and Pentothal. They vary in speed and duration of action and in potency.

These drugs have been shown to act on the

reticular formation of the upper brainstem and thalamus. Actually they suppress the excitability of all tissues; but all tissues are not equally sensitive. Low dosages induce drowsiness, and high dosages coma and death. The spread between the therapeutic and fatal doses varies with the different barbiturates.

The prolonged use of these drugs results in habituation; and insomnia, agitation, confusional psychosis, and seizures may occur within 24 to 36 h of withdrawal. Overdose is one of the commonest means of suicide in Western countries, and life can be saved only by admission to a hospital where respiration can be maintained and cerebral anoxia prevented. *See* ADDICTIVE DISORDERS; CENTRAL NERVOUS SYSTEM.

Raymond D. Adams

Bibliography. R. Rosen (ed.), *Barbiturates and Other Depressants*, 1993.

Barite

A mineral with chemical composition $BaSO_4$. It is isostructural with celestite, $SrSO_4$, and anglesite, $PbSO_4$. Barite is orthorhombic, space group Pnma, $a = 0.887$ nanometer, $b = 0.545$ nm, $c = 0.714$ nm. The **illustration** shows typical crystals. Barite possesses one perfect cleavage, and two good cleavages as do the isostructural minerals. Anhydrite, $CaSO_4$, although also orthorhombic, possesses a different space group and a different structure. Solid solution in the series $BaSO_4$–$SrSO_4$ occasionally occurs in the natural environments. Solid solution in $BaSO_4$–$PbSO_4$ is uncommon, although up to 30 mole % $PbSO_4$ in barite has been reported. Infrequency of natural solid solution between these two end members may stem from the different origins of the two minerals, anglesite being an oxidation and weathering product of galena ores. Since Ba^{2+} has the largest ionic radius of the isostructural series with the ionic radius of $Ba^{2+} \approx Pb^{2+} > Sr^{2+} \gg Ca^{2+}$, separation of these species may be a consequence of this geometrical distinction. Barite can be distinguished from celestite and anhydrite by its superior specific gravity of approximately 4.5. The mineral is relatively soft, approximately 3 on Mohs scale. The color ranges through white to yellowish, gray, pale blue, or brown, and a thin section is colorless. Barite is often an accessory mineral in hydrothermal vein systems, but frequently occurs as concretions or cavity fillings in sedimentary limestones. It also occurs as a residual product of limestone weathering and in hot spring deposits. It occasionally occurs as extensive beds in evaporite deposits.

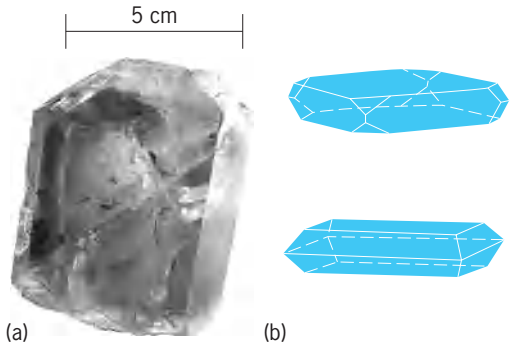

Barite. (*a*) Specimen from Dufton, Westmoreland, England (*American Museum of Natural History*). (*b*) Typical crystals; crystals may be very complex (*after C. S. Hurlbut, Jr., Dana's Manual of Mineralogy, 17th ed., John Wiley and Sons, Inc., 1959*).

Since barite is dense and relatively soft, its principal use is as a weighting agent in rotary well-drilling fluids. It is the major ore of barium salts, used in glass manufacture, as a filler in paint, and, owing to the presence of a heavy metal and inertness, as an absorber of radiation in x-ray examination of the gastrointestinal tract.

Occurrences of barite are extensive. It is found as a vein mineral associated with zinc and lead ores in Derbyshire, England. Typical associated minerals are calcite, dolomite, pyrite, quartz, and fluorite. Large deposits occur at Andalusia, Spain. Commercial residual deposits occur in limestones throughout the Appalachian states such as Georgia, Tennessee, and Kentucky. It also occurred in substantial amounts in galena ore in Wisconsin and Missouri. *See* BARIUM.

Paul B. Moore

Barium

A chemical element, Ba, with atomic number 56 and atomic weight of 137.34. Barium is eighteenth in abundance in the Earth's crust, where it is found to the extent of 0.04%, making it intermediate in amount between calcium and strontium, the other alkaline-earth metals. Barium compounds are obtained from the mining and conversion of two barium minerals. Barite, barium sulfate, is the principal ore and contains 65.79% barium oxide. Witherite, sometimes called heavy spar, is barium carbonate and is 72% barium oxide. *See* BARITE; PERIODIC TABLE; WITHERITE.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

The metal was first isolated by Sir Humphry Davy in 1808 by electrolysis. Industrially, only small amounts are prepared by aluminum reduction of barium oxide in large retorts. These are used in barium-nickel alloys for spark-plug wire (the barium increases the emissivity of the alloy) and in frary metal, which is an alloy of lead, barium, and calcium used in place of babbitt metal because it can be cast.

Properties of barium

Property	Value
Atomic number	56
Atomic weight	137.34
Isotopes (stable)	130, 132, 134, 135, 136, 137, 138
Atomic volume	36.2 cm^3/g-atom
Crystal structure	Face-centered cubic
Electron configuration	2 8 18 18 8 2
Valence	2+
Ionic radius (A)	1.35
Boiling point, °C	1140(?)
Melting point, °C	850(?)
Density	3.75 g/cm^3 at 20°C
Latent heat of vaporization at boiling point, kj/g-atom	374

The metal reacts with water more readily than do strontium and calcium, but less readily than sodium; it oxidizes quickly in air to form a surface film that inhibits further reaction, but in moist air it may inflame. The metal is sufficiently active chemically to react with most nonmetals. Freshly cut pieces have a lustrous gray-white appearance, and the metal is both ductile and malleable. The physical properties of the elementary form are given in the **table**.

For the manufacture of barium compounds, soft (easily crushable) barite is preferred, but crystalline varieties may be used. Crude barite is crushed and then mixed with pulverized coal. The mixture is roasted in a rotary reduction furnace, and the barium sulfate is thus reduced to barium sulfide or black ash. Black ash is roughly 70% barium sulfide and is treated with hot water to make a solution used as the starting material for the manufacture of many compounds.

Lithopone, a white powder consisting of 20% barium sulfate, 30% zinc sulfide, and less than 3% zinc oxide, is widely used as a pigment in white paints. Blanc fixe is used in the manufacture of brilliant coloring compounds. It is the best grade of barium sulfate for paint pigments. Because of the large absorption of x-rays by barium, the sulfate is used to coat the alimentary tract for x-ray photographs in order to increase the contrast. Barium carbonate is useful in the ceramic industry to prevent efflorescence on claywares. It is used also as a pottery glaze, in optical glass, and in rat poisons. Barium chloride is used in purifying salt brines, in chlorine and sodium hydroxide manufacture, as a flux for magnesium alloys, as a water softener in boiler compounds, and in medicinal preparations. Barium nitrate, or the so-called baryta saltpeter, finds use in pyrotechnics and signal flares (to produce a green color), and to a small extent in medicinal preparations. Barium oxide, known as baryta or calcined baryta, finds use both as an industrial drying agent and in the case-hardening of steels. Barium peroxide is sometimes used as a bleaching agent. Barium chromate, lemon chrome or chrome yellow, is used in yellow pigments and safety matches. Barium chlorate finds use in the manufacture of pyrotechnics. Barium acetate and cyanide are used industrially as a chemical reagent and in metallurgy, respectively. Reel F. Riley

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; D. F. Shriver and P. W. Atkins, *Inorganic Chemistry*, 3d ed., 1999.

Bark

A word generally referring to the surface region of a stem or a root. The term has a variable meaning, and sometimes part or all of the bark is called rind. Occasionally the word is used as a substitute for periderm or for cork only. Most commonly, however, bark refers to all tissues external to the cambium (**Fig. 1**). If this definition is applied to stems having only primary tissues, it includes phloem, cortex, and epidermis; the bark of roots of corresponding age would contain cortex and epidermis. More generally, the term bark is restricted to woody plants with secondary growth.

In most roots with secondary growth, the bark consists of phloem and periderm since the cortex and epidermis are sloughed off with the first cork formation in the pericycle, that is, beneath the cortex. In stems, the first cork cambium may be formed in any living tissue outside the vascular cambium, and the young bark may include any or all of the cortex in addition to the phloem and periderm. The first cork cambium may persist throughout the life of the plant. The resulting periderm has a relatively smooth surface. In most plants, however, new periderms are formed in successively deeper tissues. The new layers of cork are often shell-like in form and enclose islands of dead phloem (**Fig. 2**). In some plants, the successive cork layers occur in wide bands or even in whole cylinders. The region composed of the successive layers of periderm and the enclosed dead tissues is called outer bark. The outer bark gradually splits, wears, and weathers and acquires a characteristic surface pattern in relation to the form of origin of the cork cambium and the composition of the dead

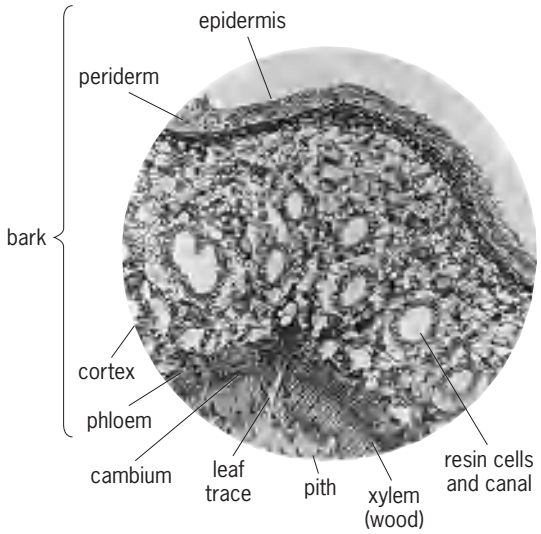

Fig. 1. Transverse section of young twig of balsam fir (*Abies balsamea*) showing tissues often included in the term bark. (*Forest Products Laboratory, USDA*)

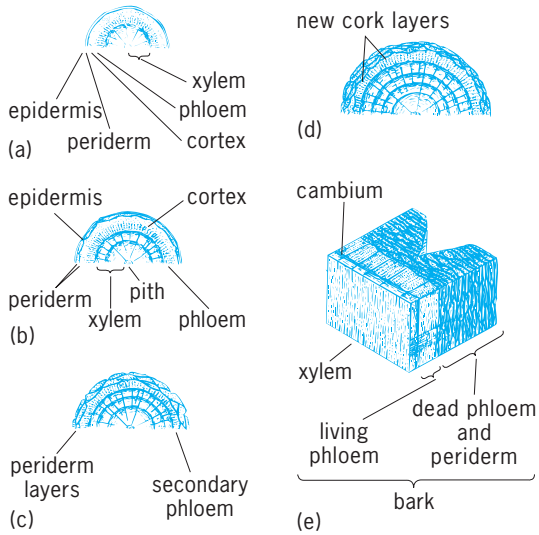

Fig. 2. Successive periderm layers in typical woody stem at (*a*) 1, (*b*) 2, (*c*) 3, (*d*) 4 years. (*e*) Outer tissues of old trunk. (*After A. J. Eames and L. H. MacDaniels, An Introduction to Plant Anatomy, 2d ed., McGraw-Hill, 1947*)

tissues. The surface may be fairly smooth or broken into strips conforming to the fibrous layers of the phloem, or into plates or strips conforming to the size and shape of the alternating zones of cork and phloem.

The outer bark composed of dead phloem alternating with bands of cork is called, technically, rhytidome. Both stems and roots may have rhytidome. The inner bark is living, and consists of phloem only. In this phloem only a narrow band next to the vascular cambium contains functioning sieve elements, but ray cells and parenchyma are alive throughout the inner bark. From these living cells, new cork cambia may develop. *See* CORTEX (PLANT); EPIDERMIS (PLANT); PARENCHYMA; PERICYCLE; PERIDERM; PHLOEM; ROOT (BOTANY); STEM.

H. Weston Blaser

Bibliography. J. D. Mauseth, *Plant Anatomy*, 1988.

Barkhausen effect

An effect, due to discontinuities in size or orientation of magnetic domains as a body of ferromagnetic material is magnetized, whereby the magnetization proceeds in a series of minute jumps. *See* FERROMAGNETISM; MAGNETIZATION.

Ferromagnetic materials are characterized by the presence of microscopic domains of some 10^{12} to 10^{15} atoms within which the magnetic moments of the spinning electrons are all parallel. In an unmagnetized specimen, there is random orientation of the various domains. When a magnetic field is applied to the specimen, the domains turn into an orientation parallel to the field, or if parallel to the field, the domains increase in size. During the steep part of the magnetization curve, whole domains suddenly change in size or orientation, giving a discontinuous increase in magnetization. If the specimen being magnetized is within a coil connected to an amplifier and loudspeaker, the sudden changes give rise to a series of clicks or, when there is a rapid change, a hissing sound. This is called the Barkhausen effect; it is an important piece of evidence in support of a domain theory of magnetism.

Kenneth V. Manning

Barley

A cereal grass plant whose seeds are useful to humans. It is grown in nearly all cultivated areas of the temperate parts of the world, and is an important crop in Europe, North and South America, North Africa, much of Asia, and Australia. Barley is the most dependable cereal crop where drought, summer frost, and alkali soils are encountered. In the United States, barley is grown to some extent in 49 states, with the most important production areas in North Dakota, Montana, and California. Principal uses of barley grain are as livestock feed, for the production of barley malt, and as human food. It is also used in some areas for hay and silage.

Origin and description. Barley is probably one of the most ancient cultivated cereal grains. Its origin is still debated. Archeological evidence in the form of heads, carbonized grains, and impressions of grains on pottery have been found in several areas of the world. Some in Egypt are dated over 5000 years ago. Barley was probably grown by the Swiss Lake Dwellers between 2000 and 3000 B.C., in China about 200 B.C., and in Greek and Roman areas about 300 B.C.

Taxonomically barley belongs to the family Gramineae, subfamily Festucoideae, tribe Hordeae, and genus *Hordeum*. Most of the modern cultivated barleys are *H. vulgare* (six-rowed) or *H. distichum* (two-rowed; **Fig. 1**). All cultivated barleys are annuals, are naturally self-pollinated, and have seven pairs of chromosomes. **Figure 2** shows a typical barley plant. Cultivated barleys can be readily crossed by emasculating before pollen is shed or through genetic male sterility. There are also two wild species having chromosomes homologous with the cultivated species which are crossable with each other and with the cultivated species. Both of these, *H. spontaneum* (two-rowed) and *H. agriocrithon* (six-rowed), have brittle rachis and readily reseed themselves in the wild state. There are also a number of wild grasslike species which belong to the genus *Hordeum*. The basic chromosome number of these species is 7, and $2n$ numbers of 14, 28, and 42 have been observed. These species do not cross readily with the cultivated species, although there has been marked advancement in artificial crosses with the use of colchicine techniques and in growing excised embryos on artificial media. Among this group of wild barleys are annuals including *H. murinum*, *H. gussoneanum*, *H. glaucum*, *H. leproinum*, *H. depressum*, and *H. geniculatum*. Perennial types are *H. jubatum*, *H. bulbosum*, and *H. montanese*. *See* CHROMOSOME; CYPERALES; GENETICS; POLLINATION; REPRODUCTION (PLANT).

Varieties. In the cultivated varieties, a wide diversity of morphological, physiological, and anatomical types are known. There are spring, facultative,

Fig. 1. Barley spikes: (*a*) six-rowed *Hordeum vulgare*; (*b*) two-rowed *H. distichum* (*USDA*)

and winter growth habits; hulled and naked grain; awned, awnless, and hooded lemmas; black, purple, and white kernels; and also a wide range of plant heights, spike densities, and resistances to a wide range of diseases and insects. There are in excess of 150 cultivars presently commercially grown in the United States and Canada alone, and many additional cultivars are grown in other parts of the world. New and improved varieties produced by barley breeders are constantly replacing older varieties. Several barley collections are being maintained in different countries as germplasm sources for breeding and research. These include both collections made by direct exploration in many barley-growing areas of the world and lines from barley-breeding programs. Among the largest of these collections is one maintained by the U.S. Department of Agriculture, which includes more than 17,000 individual strains.

Cultural practices. The best soils for growing barley are well-drained fertile loams and clay loams. This grain is the most tolerant of the cereal grains to alkaline soils (pH 6.0 to 8.0), and the least tolerant to acid soils (below pH 5.0), especially if there is soluble aluminum in the soil solution. Liming may be necessary to correct acidity and to precipitate aluminum, and lime should be applied at the time of seedbed preparation. Soil tests will determine if phosphorus or potash is needed. Nitrogen will generally be needed for good production, depending on the previous crop and climatic conditions. Time and rate of seeding, seedbed preparation, and rotations vary widely with geographical area. In the United States, spring varieties are grown in the Midwest, northern Great Plains, and Pacific Northwest. Seeding should be done as early as the seedbed can be prepared. This will vary from about March 1 to 15 in the Kansas-Nebraska area, the Great Basin, and the Pacific Northwest to late April or early May in the northern Great Plains. Spring varieties are commonly sown in the late fall or winter in the southwestern states, the Atlantic seaboard, the southern Great Plains, and in the milder areas of the Pacific Northwest. Generally the planting date will be 1 to 2 weeks ahead of the earliest fall frost-free date. Common seeding rates are 4–8 pecks/acre (0.087–0.174 m^3/ha), depending on expected moisture conditions. Barley can be grown in a variety of rotations with other crops common to a particular region. In areas with long growing seasons, it is possible to double-crop (growing two crops in 12 consecutive months on the same land) with summer crops such as soybeans or cotton. Barley is ideally suited to this purpose because it is usually the earliest maturing of the small grains. The variety of the alternate crop should also be early-maturing. *See* GRAIN CROPS. David A. Reid

Diseases. Barley is one of the cereals most subject to diseases, especially in humid areas. Consequently, grain yields of varieties grown in humid areas have not been consistently high. Because of diseases and the grain quality needed for malting purposes, barley production has shifted westward from the Corn Belt to northwestern Minnesota, through North Dakota, Montana, northern Idaho, and California, where the relative humidity is lower.

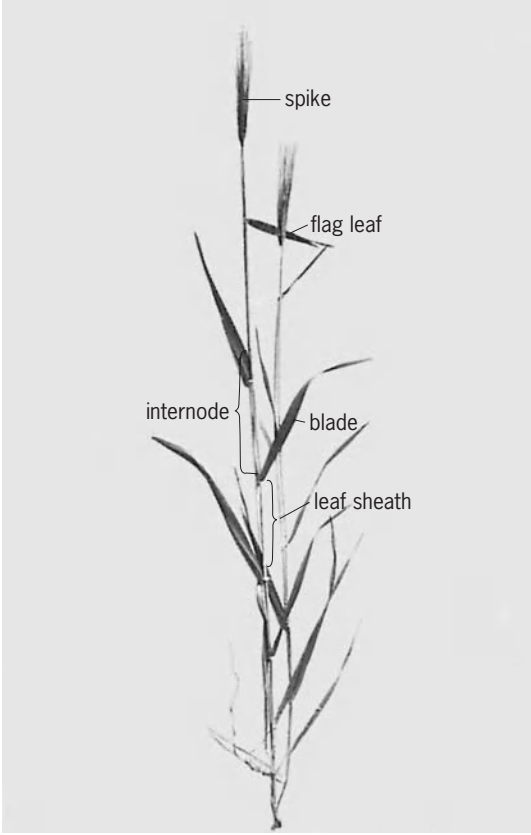

Fig. 2. Typical barley plant. (*USDA*)

Fig. 3. Barley stripe, caused by *Helminthosporium gramineum*. (*Minnesota Agricultural Experiment Station*)

Yield losses from some diseases result from reductions in stand and in plant vigor. Other diseases cause the quality of the grain to be reduced through shriveling, discoloration, and toxin production in the grain itself.

Seed-borne pathogens. Fungi carried on or in the seed [*Ustilago hordei*, which causes covered smut; *U. nuda*, which causes loose smut; *Helminthosporium gramineum*, the cause of stripe (**Fig. 3**); and *H. sorokinianum* and *H. sativum*, which cause seedling blight and spot blotch (**Fig. 4**)] are controlled to a large extent by growing varieties with some degree of resistance and by fungus-free seed. Where clean seed is not available, treatment of the seed with a mixture of Carboxin and Thiram will give virtually 100% control. *See* FUNGISTAT AND FUNGICIDE.

Fig. 4. Barley spot blotch, caused by *Helminthosporium sativum*. (*Minnesota Agricultural Experiment Station*)

Foliage diseases. Fungi that cause these diseases may be seed-borne, carried over on plant debris from the previous season, or blown in from distant sources. Varieties resistant to several of them have been developed. Seed treatment and crop rotation or disposal of debris (plowing-under) will also aid in their control. Examples of leaf diseases and their causes are spot blotch (*H. sorokinianum*), net blotch (*H. teres*), speckled leaf blotch (*Septoria passerinii*), and scald (*Rhynchosporium secalis*). Leaf rust (caused by *Puccinia hordei*, stem rust (caused by *P. graminis*), and powdery mildew (caused by *Erysiphe graminis*) are usually windborne from distant sources and are mainly controlled by the growing of resistant varieties. However, the occurrence of new pathogenic strains requires the development of new resistant varieties with different resistance genes.

Grain diseases. Kernel blights, caused by *H. sorokinianum*, *Gibberella zeae*, and *Alternaria* species (**Fig. 5**), shrivel and discolor the grain, thus reducing the marketing and particularly the malting quality. The scab disease, caused by *G. zeae*, also reduces the feeding value of barley, particularly for hogs, because of the toxin or toxins produced by infection. Because this fungus also infects corn as a stalk or ear rot, barley should not follow corn in a rotation. Saprophytic organisms on the kernels may cause deterioration of the grain in storage.

Seedling, root, and crown diseases. Several of the abovementioned fungi cause rotting of various plant parts. These organisms may be seed- or debris-borne. Others, such as *Pythium* species, may be soil-borne. Pathogen-free seed, chemical seed treatment, sanitation, crop rotation, and cultural practices may be of value in their control.

Other diseases. Barley stripe mosaic and barley yellow dwarf diseases are caused by viruses. The first is seed-borne and is controlled by the use of virus-free seed and resistant varieties. The second is aphid-transmitted from perennial grasses, and is controlled by the use of resistant varieties. The oat cyst nematode (cereal root eelworm) *Heterodera avenae* has not been important in the United States. In Europe, it is a serious problem in barley as well as in oats. *See* PLANT PATHOLOGY; PLANT VIRUSES AND VIROIDS.

Deane C. Arny

Pot and pearled barley. The firmly adhering husk of the barley grain must be removed to facilitate use of barley in cooking. Pot and pearled barley are particulate products produced by abrasive removal of the hull or outer portion of the kernel. Thoroughly cleaned barley grain in sound condition is essential to production of good-quality products. The grain is exposed to the abrasive action of a series of rotating carborundum-coated disks to grind the hull from the kernel. Air aspiration is used to facilitate removal of fine particles and hull fragments. To produce pot

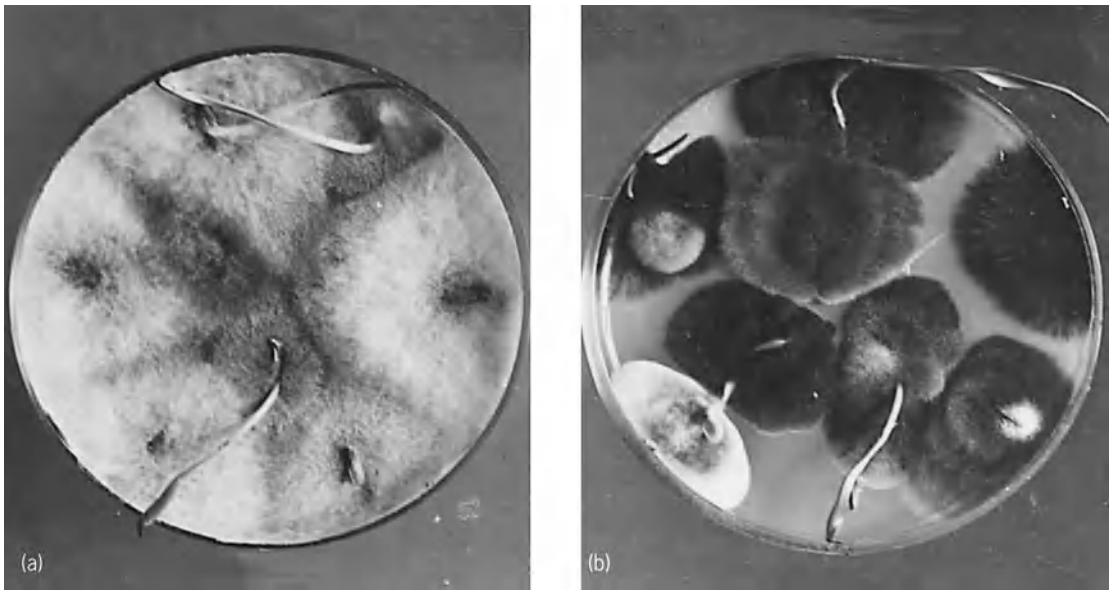

Fig. 5. Infected barley seed planted out on nutrient agar. (*a*) *Gibberella zeae*. (*b*) *Helminthosporium* and *Alternaria* sp. These kernel blights occur in more humid regions. (*Minnesota Agricultural Experiment Station*)

barley, the partially dehulled grain passes through a series of grinders until the kernel is mulled through the bran and aleurone layer. Removal of this outer portion results in about a 65% yield of clean pot barley suitable for cooking. Increasing the succession of passes through abrasive disks results in removal of greater portions of the kernel. White-pearled barley is obtained as a smooth and round kernel with a yield approximating 35%. The heating of grain during this attrition process must be controlled to avoid starch damage. Pot and pearled barley are cleaned, sized, and packaged for use in soups or porridges. Pearled barley may be further ground to produce high-quality grits or flour.

Malting. Malt is obtained from the controlled germination of barley grain. While malting procedures are an extremely varied art, precise scientific monitoring which utilizes numerous techniques or points of control is required. The characteristics of the resultant malt are influenced by the initial grain and the stages and conditions of the malting process. During the sprouting of cereal grains, numerous physiochemical changes occur, including increases of respiration (enzyme mobilization) and modifications of stored energy reserves (starch-sugars). The basic process for malting includes three stages: steeping, germination, and drying.

Steeping. Cleaned and sized barley is placed in large cylindrical steep tanks, covered with 50–60°F (10–16°C) water, aerated, and held there until the grain moisture content reaches 43–45%. The steeping time may vary from 30 to 70 h, during which time control of temperature, aeration rate, and steep-water quality are essential for maximizing germination potential.

Germination. Technically, germination is initiated during the steeping process; however, soaked grain is transferred to specialized germinating rooms or floors where the steeped grain is piled and turned by hand daily. Modern pneumatic malting equipment enables controlled mixing within drums or chambers. Grain possessing optimum moisture is maintained at 60–70°F (16–21°C) with adequate levels of aeration. Control of germinating temperature reduces losses to respiration and sprout growth and thus increases dry-matter yield. Germination time varies between 4 and 7 days depending on the type of barley and its intended use.

Drying or kilning. The germinated barley grain, known as green malt, is heat-inactivated at the appropriate stage of development. The green malt is heated to 110–130°F (43–54°C) and dried to about 6–12% moisture. This stabilized product will retain high levels of enzymatic activity and provide characteristic color and flavor. Kiln temperatures, ranging up to 195°F (90°C), may be used to develop highly flavored and colored malts with low enzyme activity. Barley sprouts are removed following drying, and malted grain is suitable for grinding and blending.

Use of malts. Malted barley is used in the brewing industry as an active source of starch-hydrolyzing enzymes (amylases) to improve the energy source, glucose, during fermentation and to impart the desired flavor and color to the brew. Food-grade malts used for selective enzyme activity or characteristic flavor and color are of increasing importance in the formulation of baked goods and breakfast cereals. *See* AMYLASE; FOOD MANUFACTURING; MALT BEVERAGE.

Mark A. Uebersax

Bibliography. D. E. Briggs, *Barley*, 1978; N. L. Kent, Barley: Processing, nutritional attributes, technological uses, *Technology of Cereals*, chap. 12, 1978; W. H. Leonard and J. H. Martin, *Cereal Crops*, 1963; D. E. Mathre (ed.), *Compendium of Barley Diseases*, no. 8, 1982; S. A. Matz, *Cereal Science*, 1969; Y. Pomeranz, Barley, *Elements of Food Technology*, pp. 173–178, 1977; Y. Pomeranz, Food uses of barley, *CRC Critical Reviews in Food Technology*, pp. 377–394, 1974; U.S. Department of Agriculture, *Barley: Origin, Botany, Culture, Winter Hardiness, Utilization, Pests*, 1979.

Baroclinic field

A distribution of atmospheric pressure and mass such that the specific volume, or density, of air is a function of both pressure and temperature, but not either alone. When the field is baroclinic, solenoids are present, there is a gradient of air temperature on a surface of constant pressure, and there is a vertical shear of the geostrophic wind. Significant development of cyclonic and anticyclonic wind circulations typically occurs only in strongly baroclinic fields. Fronts represent baroclinic fields which are locally very intense. *See* AIR PRESSURE; FRONT; GEOSTROPHIC WIND; SOLENOID (METEOROLOGY); STORM; WIND.

Frederick Sanders; Howard B. Bluestein

Barometer

An absolute pressure gage which is specifically designed to measure atmospheric pressure. This instrument is a type of manometer with one leg at zero pressure absolute. *See* MANOMETER.

The common meteorological barometer (see **illus.**) is a liquid-column gage with mercury. The top of the column is sealed, and the bottom is open and submerged below the surface of a reservoir of mercury. The atmospheric pressure on the reservoir keeps the mercury at a height proportional to that pressure. An adjustable scale, with a vernier scale, allows a reading of column height. Before each reading, the scale must be adjusted to correct for mercury level in the reservoir. Often a peg (the zero reference point) attached to the scale is adjusted by a micrometer to just touch the mercury surface. The apparent mercury height must be corrected for temperature and gravity, with standard conditions being 0°C and 980.665 cm/s^2.

Typically, the barometer scale is in millimeters or inches of mercury, with sea-level pressures in the order of 760 mm or 29.9 in. of mercury.

Aneroid barometers using metallic diaphragm elements are usually less accurate, though often more sensitive, devices, and not only indicate pressure but may be used to record. *See* PRESSURE MEASUREMENT.

John H. Zifcak

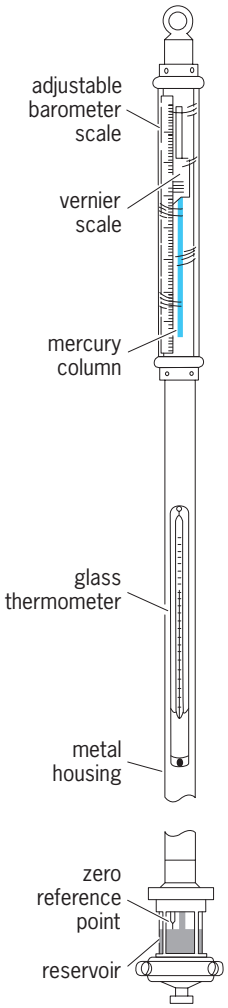

Type of mercury barometer used in meteorology.

Barotropic field

A distribution of atmospheric pressure and mass such that the specific volume, or density, of air is a function solely of pressure. When the field is barotropic, there are no solenoids, air temperature is constant on a surface of constant pressure, and there is no vertical shear of the geostrophic wind. Significant cyclonic and anticyclonic circulations typically do not develop in barotropic fields. Considerable success has been achieved, paradoxically, in prediction of the flow pattern at middle-tropospheric elevations by methods which are strictly applicable only to barotropic fields, despite the fact that the field in this region is definitely not barotropic. The subtropics, however, are to a large extent barotropic. *See* AIR PRESSURE; BAROCLINIC FIELD; GEOSTROPHIC WIND; SOLENOID (METEOROLOGY); WEATHER FORECASTING AND PREDICTION; WIND.

Frederick Sanders; Howard B. Bluestein

Barracuda

Perciform fishes in a single genus, *Sphyraena*, in the family Sphyraenidae. They are characterized by an elongate, slightly compressed body; a long, pointed head with a large horizontal mouth with nonprotractile premaxillae and the lower jaw projecting beyond

Great barracuda (*Sphyraena barracuda*). (*Photo © John E. Randall*)

Names, distributions, habitats, and characteristics of 25 species of barracuda (*Sphyraena*)

Species	Common name	Distribution	Habitat	Max.length/weight
S. acutipinnis	Sharpfin barracuda	Indo-Pacific	Lagoons and seaward reefs	80 cm (32 in.)
S. afra	Guinean barracuda	Eastern Atlantic	Continental shelf to 75 m, enters lagoons and estuaries	205 cm (81 in.)/50 kg (110 lb)
S. argentea	Pacific barracuda	Eastern Pacific	Usually near shore in small schools	145 cm (57 in.)/12 kg (26 lb)
S. barracuda	Great barracuda	Indo-Pacific, Eastern and Western Atlantic	Estuaries, reefs, and open ocean to 100 m depth	200 cm (79 in.)/50 kg (110 lb)
S. borealis	Northern sennet	Western Atlantic from Massachusetts to Gulf of Mexico	Reefs	46 cm (18 in.)
S. chrysotaenia	Yellowstripe barracuda	Indo-Pacific		30 cm (12 in.)
S. ensis	Mexican barracuda	Eastern Pacific, Mexico to Ecuador	Continental shelf	127 cm (54 in.)/9.5 kg (21 lb)
S. flavicauda	Yellowtail barracuda	Indo-Pacific	Lagoons, sheltered seaward reefs	60 cm (24 in.)
S. forsteri	Bigeye barracuda	Indo-Pacific	Outer reefs	75 cm (30 in.)
S. guachancho	Guaguanche	Eastern and Western Atlantic	Turbid coastal waters, often in river estuaries, seaward to 100 m	20 cm (8 in.)/1.75 kg (3.9 lb)
S. helleri	Heller's barracuda	Indian Ocean and islands of the Pacific	Bays and coral reefs	80 cm (32 in.)
S. idiastes	Pelican barracuda	Southeast Pacific, Cocos and Galápagos islands	Reefs	91 cm (36 in.)
S. japonica	Japanese barracuda	Western Pacific, southern Japan, and South China Sea	Near shore	35 cm (14 in.)
S. jello	Pickhandle barracuda	Indo-West Pacific	Bays and seaward reefs, to 200 m depth	150 cm (59 in.)/11.5 kg (25 lb)
S. lucasana	Cortez barracuda	Eastern Central Pacific		70 cm (28 in.)
S. novaehollandiae	Australian barracuda	Indo-Pacific	Open channels and semiprotected areas	100 cm (39 in.)
S. obtusata	Obtuse barracuda	Indo-Pacific	Seagrass beds and rocky reefs in bays and estuaries, depth to 120 m	55 cm (22 in.)
S. picudilla	Southern sennet	Western Atlantic, Bermuda, Florida, Uruguay	Rocky and coral reefs of coastal waters	61 cm (33 in.)
S. pinguis	Red barracuda	Northwest Pacific	Near shore on mud or rock bottoms	30 cm (12 in.)
S. putnamae	Sawtooth barracuda	Indo-West Pacific	Bays, current-swept lagoons, and seaward reefs	90 cm (35 in.)
S. qenie	Blackfin barracuda	Indo-Pacific and Mexico and Panama in the Eastern Pacific		170 cm (67 in.)
S. sphyraena	European barracuda	Eastern Atlantic and Bermuda and Brazil in Western Atlantic	Coastal and offshore waters	165 cm (65 in.)
S. tome	No common name	Southwest Atlantic, Brazil, and Argentina	Depth to 83 m	45 cm (18 in.)
S. viridensis	Yellowmouth barracuda	East Central Atlantic, including the Azores	Probably coastal and offshore waters	128 cm (50 in.)
S. waitii	No common name	Indo-West Pacific		31 cm (12 in.)

the upper; strong, sharp, uneven conical teeth in the jaws as well as in the roof of the mouth (palatine teeth) and usually a large canine tooth near the tip of the lower jaw; two dorsal fins, both short, the first located above the pelvic fins with five strong spines, the second usually with one spine and nine soft rays, well removed from the first and a mirror image of the anal fin; short pectoral and pelvic fins; a forked caudal fin; cycloid scales that are smooth to the touch; and coloration that is usually gray to green or bluish above, with silvery reflections, lighter below, and sometimes with dark vertical bars or chevrons (see **illustration**).

Barracudas inhabit the tropical and subtropical Atlantic, Pacific, and Indian oceans, usually near the surface, but they may descend to 100 m (330 ft)

or more. All are pelagic marine species; some frequently enter brackish waters, especially the juveniles. They are swift and powerful swimmers and voracious predators, feeding mostly on small fishes and to a lesser extent on shrimp and squid. For the most part barracudas are harmless; however, large individuals are known to attack humans, and in some parts of the world they are more feared than sharks. Although a food fish, some barracudas are known to contain ciguatera poison.

The known distribution, habitat, and characteristics of each of the 25 named species are given in the **table**. *See* PERCIFORMES. Herbert Boschung

Bibliography. D. P. de Sylva, Sphyraenoidei: Development and relationships, in H. G. Moser et al., *Ontogeny and Systematics of Fishes*, Amer. Soc. Ichthy. Herp. Spec. Pub. no. 1, 1984.

Barretter

A bolometer element with a positive temperature coefficient of resistance, used to detect and measure power at radio, microwave, infrared, and optical frequencies. The temperature of the barretter increases when electromagnetic energy is absorbed. Barretters are made of metal; therefore, the electrical resistance increases when the temperature increases. The resulting resistance change of the barretter is measured by using direct-current or low-frequency instruments. Early radio demodulators and current regulator tubes using fine iron wires also have been called barretters.

Barretters for measuring power from less than 1 microwatt to several milliwatts are made from Wollaston wire. Wollaston wire is a platinum wire less than 100 microinches (2.5 micrometers) in diameter drawn inside a silver wire with a diameter about 10 times larger. The silver permits handling the wire for mounting and is then etched off to leave a bare platinum wire of the desired resistance.

The barretter resistance is selected to absorb most of the power when the barretter is mounted as a termination in a waveguide or coaxial transmission line. A barretter can be made to detect power at optical and infrared frequencies by using a very thin metal ribbon blackened to absorb light. *See* TRANSMISSION LINES; WAVEGUIDE.

Barretters with less sensitivity and accuracy for use at radio frequencies can be made by using low-current fuses made with fine wires. Fuses intended to protect instruments and rated for about 10 milliamperes work well. For higher powers, small lamp filaments can serve as barretters. *See* FUSE (ELECTRICITY); INCANDESCENT LAMP.

Measurement of modulated power. A meter can be made to measure high-frequency signal amplitudes using a barretter (see **illus.**). The temperature and hence the resistance of a barretter can change at audio-frequency rates, but the time constant of a barretter is too great for the resistance to vary at radio-frequency rates. A radio- or microwave-frequency

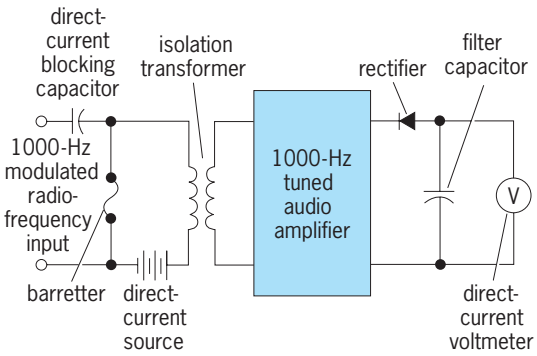

Circuit for measuring modulated radio-frequency power by using a barretter.

current modulated at a low frequency will cause the barretter resistance to follow the low-frequency signal. If a direct-current voltage is applied to the barretter while the modulated radio-frequency current is also applied, the varying resistance will produce a current which follows the modulation. The low-frequency current can be coupled to the input of an audio amplifier tuned to the modulation frequency by using an audio transformer. The output of the audio amplifier may be rectified to drive a direct-current meter. The meter then indicates the relative amplitude of the radio-frequency or microwave signal. *See* AUDIO AMPLIFIER.

Bridge circuit. For precision measurements, the barretter can be used as one arm of a Wheatstone bridge. The dc power dissipated in the barretter and the barretter resistance can be calculated from the resistances of the other three arms of the bridge and the voltage applied to the bridge. If the barretter and the bridge resistors have the same temperature coefficient of resistance, the measuring circuit will be self-compensating for ambient temperature variations. *See* BRIDGE CIRCUIT; RESISTANCE MEASUREMENT; WHEATSTONE BRIDGE.

High-frequency effects. The current path and heating effects in a barretter are not exactly the same at low frequencies and high frequencies. The simple cylindrical geometry of a barretter, however, permits accurate calculation of the current and heat distribution. When direct currents heat the barretter, the center of the barretter is at a higher temperature than the surface because the power is dissipated uniformly but the heat is dissipated primarily at the surface. At high frequencies, however, the power is dissipated primarily at the surface because of skin effect. Direct currents dissipate power uniformly along the length of the barretter, but at higher frequencies, as the wavelength of the radio-frequency wave approaches the length of the barretter, the current approaches a sinusoidal distribution and the heating becomes nonuniform along the wire. The ratio of direct current to radio-frequency heating must be carefully calculated or, for critical work, calibrated against known standards. *See* BOLOMETER; ELECTROMAGNETIC RADIATION; SKIN EFFECT (ELECTRICITY). Robert C. Powell

Bibliography. G. H. Bryant, *Principles of Microwave Measurements*, 1988; T. S. Laverghetta, *Modern Microwave Measurements and Techniques*, 1988.

Barrier islands

Elongate accumulations of sediment formed in the shallow coastal zone and separated from the mainland by some combination of coastal bays and their associated marshes and tidal flats. Barrier islands are typically several times longer than their width and are interrupted by tidal inlets. There are at least three possible explanations for their origin: (1) longshore spit development and subsequent cutting of inlets; (2) upward shoaling by wave action of subtidal longshore sand bars; and (3) drowning of coastal ridges. All three modes of origin may have occurred. The first two have been observed in recent times, the third is yet to be conclusively demonstrated.

Formation. Modern barrier islands extend along approximately 15% of the Earth's coasts. They are geologically quite young, having been formed about 7000 years ago or less, as sea level slowed or came to its present position during the Holocene transgression. Some are still in the process of forming and gaining definition. The primary requisites for their formation are a significant amount of available sediment, a place for it to accumulate, and significant wave-energy influence. Tectonic setting is an important control in the development of barrier islands, with most forming along the American-type trailing edge coasts of mature continental margins such as the east coasts of North and South America. The low-relief coastal plain and adjacent shelf are served by well-developed drainage systems with abundant sediment. Barriers are also well developed along marginal sea coasts such as the Gulf of Mexico. Both of these settings provide ideal conditions to produce long and continuous barrier island systems. This contrasts with leading-edge coasts on converging plate boundaries where there is high relief. This type of coasts tends to produce numerous but small barrier spits emanating from headlands, such as along the Pacific coast of the United States. *See* CONTINENTAL MARGIN; PLATE TECTONICS.

Barrier environments. It is not possible to discuss barrier islands without considering the adjacent and closely related environments within the coastal system. Beginning offshore and proceeding toward land, the sequence of environments comprises shoreface, including the offshore and nearshore; beach; dunes; washover fans, which comprise the back-island flats; marsh; tidal flat; and coastal bay, which may be a combination of estuarine and lagoonal environments leading to the mainland (**Fig. 1**). The barrier island proper includes the beach, dunes, and washover fans; however, the adjacent shoreface is well integrated with the barrier in terms of morphology, processes, and sediments.

Shoreface. The shoreface extends from the zone where storm waves influence bottom sediment transport, shoreward to the low-tide line. Typically it is subdivided into the outer shoreface, that area affected only by storm waves, and the inner shoreface or nearshore zone, where normal waves influence the bottom. The latter contains shallow longshore sand bars and troughs, varying in number from one coastal location to another. The depths of water marking the boundaries of the shoreface vary with location, depending upon wave climate.

Beach. The barrier island itself begins at the low-tide or outer limit of the beach. In many places the lower part of the beach is characterized by an ephemeral bar and trough generally called a ridge and runnel (Fig. 1). This small bar moves onto the upper beach under low-energy wave conditions, either repairing earlier storm erosion or adding to the beach, thereby causing the island to prograde. The upper intertidal or foreshore beach extends up to the berm, where the foreshore and backshore beach environments meet. The upper foreshore slopes toward the water and is the site of the uprush and backwash of waves—the swash zone. The backshore part of the beach is a dry, eolian environment except during storms. It tends to be horizontal or slightly landward-sloping. *See* EOLIAN LANDFORMS.

Dunes. Coastal dunes occupy a wide range of portions of the barrier island. The most common site of dune development is immediately landward of the beach in the form of shore-parallel ridges. These are called foredunes and may rise several meters above the beach. Barrier islands that receive abundant sediment and tend to prograde seaward may have numerous foredune ridges. *See* DUNE.

Low dunes or low areas between dunes may act as pathways for water and sediment during storm surges, producing washover fans on the landward, flat side of the barrier. This part of the barrier is generally only a meter or two above sea level and is covered with a floral community that extends from grasses in the high areas down to marsh grasses as the high

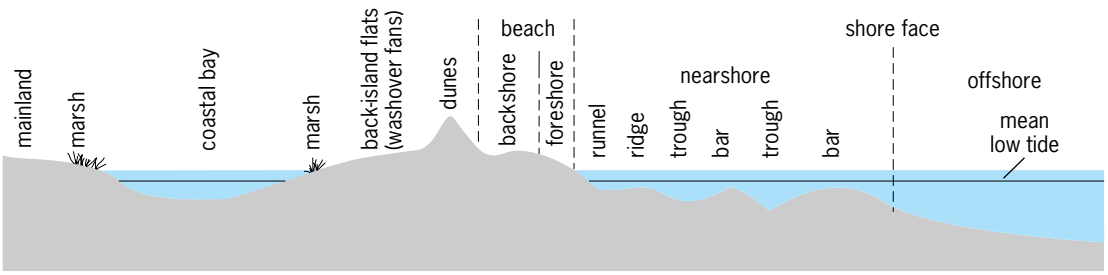

Fig. 1. Generalized profile across a barrier island complex showing various environments.

intertidal zone is reached. Sediment that supports this vegetation is delivered in the form of washover fans which coalesce to form a washover apron—the landward part of the barrier. An individual fan may be about a half meter in thickness and extend over several acres. Large storms may generate sufficient surge to produce washover fans that traverse the entire barrier and extend into the adjacent back-barrier bay.

Marshes and tidal flats. These are the landward continuation of the back-barrier environment. Their extent tends to be related to the combination of tidal range and morphology of the island; generally the higher the tidal range, the wider the back-barrier marsh and tidal-flat complex.

Coastal bay. The coastal bay that separates the barrier island from the mainland may take on various characteristics. In many areas it is an estuary, a bay that is diluted by fresh-water runoff from streams and that has good tidal circulation with the open ocean. Examples are found behind the Outer Banks barrier system of North Carolina and the barriers of the Florida Gulf Coast. The other extreme is a coastal bay that is a lagoon with no significant fresh-water input and very restricted tidal flux. Examples are Laguna Madre behind Padre Island on the southern Texas coast, and the Coorong lagoon in South Australia—both well over 100 km (60 mi) long. Some barrier systems include coastal bays that have important seasonal variation in their hydrography.

Inlets. The continuity of barriers is interrupted by inlets, which carry great quantities of water and sediment between the open marine environment and the coastal bays. Inlet spacing, size, and stability vary, depending upon the dominant processes controlling barrier island morphology. Inlets may be closely spaced such as along the Georgia and South Carolina coasts, or they may be widely spaced such as on the Outer Banks of North Carolina and on the Texas coast.

Processes. A variety of physical processes occur along the coast. These processes act to shape and maintain the barrier island system and also to enable the barrier to slowly migrate landward as sea level rises. The most important process in forming and maintaining barriers is the waves, which also produce longshore currents and rip currents. Waves and longshore currents dominate the outer portion of the barrier system, whereas tidal currents are dominant landward of the barrier, although small waves may also be present. Tidal currents also dominate inlet processes, although waves also may influence this environment. On the supratidal portion of the barrier island, the wind is the most dominant physical process.

Waves. Normal waves that cross the inner shelf and break in the surf zone do not cause significant sediment motion until they steepen and break in shallow water, typically over a longshore bar. Such breaking causes much sediment to be temporarily suspended and generates local currents that may carry sediment toward or away from the shore. Storm waves cause sediment motion in relatively deep water. They also tend to remove sediment from the beach and transport it seaward. The reverse situation exists during low-wave-energy conditions. Examples of this condition occur seasonally along much of the west coast of the United States. The winter storms produce large waves which remove sediment from the beaches and transport it offshore, where it is stored until the lower-wave-energy conditions of the spring. The sediment then returns and the beach is restored. This cycle of erosional and accretional beach is common globally and is related to stormy versus quiet wave conditions.

Waves rarely approach the coast with the crests parallel to the shore. Instead, they approach at some acute angle and are refracted, thereby generating longshore currents that flow away from the acute angle made by the wave and the coast. These currents are the primary mechanism for distribution of sediment along the barrier beach and surf zone. Longshore or littoral drift may cause a net transport of hundreds of thousands of cubic meters of sediment to pass a given point in a year's time. Waves and longshore currents, in concert with nearshore topography, cause formation of rip currents that may transport sediment seaward. These currents are produced by the temporary piling up of water between the inner longshore bar and the beach. This unstable condition causes water to flow seaward, and the water does so through low areas of the bar. These narrow currents are quite local but may be dangerous to swimmers. *See* OCEAN WAVES.

Tides. Tidal currents carry great quantities of water and sediment through the inlets and to a lesser extent in the estuaries landward of the barrier islands. Much sediment is deposited just landward of the inlet as a flood-tidal delta. The equivalent seaward accumulation, the ebb-tidal delta, varies greatly in size and shape, depending on wave and tidal interactions. Along many barriers, waves produce longshore currents that cause inlets to migrate or even close; these are wave-dominated inlets. By contrast, other inlets have strong tidal currents that cut deep and stable inlets which show little wave influence; these are tide-dominated inlets.

Tides are also important landward of the barriers in the estuaries, marshes, and tidal flats. Currents generated by tides are a primary mechanism for moving sediment both in suspension and as bedload. Flooding of low areas by tides provides a means for depositing sediment on the marsh or tidal flat. Typically the rate of tidal sediment flux is directly related to the tidal range. *See* OCEAN CIRCULATION.

Wind. The wind is also an important agent for sediment transport and therefore for molding barrier island morphology. Prevailing winds along the coast typically have some onshore component. The back-beach zone, or at low tide the entire beach, serves as a sediment source for onshore winds that accumulate this sediment behind the beach in the form of dunes. Wind also causes sediment to blow from the dunes back on the washover apron or marsh and also directly into the bay. Along parts of Padre Island in southern Texas, wind has blown so much sediment

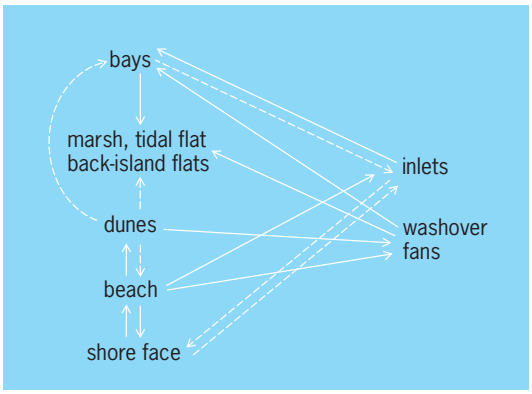

Fig. 2. Sediment budget for barrier island complex. Solid lines are major interactions; broken lines are relatively minor intersections.

into Laguna Madre that it is essentially filled. *See* NEARSHORE PROCESSES; WIND.

Sediment budget. The combined processes operating in the various sedimentary environments of the barrier island complex are best viewed through a sediment budget (**Fig. 2**). There are two distinct sources of sediment for this system (excluding biogenic or skeletal debris): the bays and the shoreface. Both of these environments receive feedback from other environments also. Shoreface sediment is transferred primarily to the beach, with small amounts going directly to the inlet. Beach sediment moves toward four sites: dunes, alongshore to inlets, over the barrier by storm washover, or seaward to the shoreface. Dunes also serve as a sediment source for washover fans, and they contribute as well to the back-island environments and bays through eolian transport of sediment. The marshes and other low-relief, back-island areas receive sediment from several sources, but do not provide sediment except in the case of severe storms, such as hurricanes, when even these low-energy environments are eroded.

The coastal bays show great variability in the sediment budget. Those bays that have little or no runoff from land or receive little tidal influence (lagoons) have little sediment input except from washover and blowover. Many bays receive much sediment runoff from the mainland and are subjected to pronounced tidal influence (estuaries). These bays provide sediment to marshes and tidal flats and some to inlets. They receive sediment from the same sources as lagoons, with the addition of inlets.

Morphodynamics. The response of the barrier island system to the above processes results in various sizes and shapes of barriers and their related inlets. These morphodynamics are dictated primarily by the interaction of wave- and tide-generated processes. Barrier islands are restricted to wave-dominated coasts and to coasts where there is a mixture of wave and tide influence. Tide-dominated coasts do not permit barrier islands to develop, due to the dominance of on- and offshore flux. Barrier islands on wave-dominated coasts are long and narrow with abundant and widespread washover fans (**Fig. 3***a*). Inlets tend to be widely spaced, small, and unstable due to the dominance of waves primarily through longshore currents.

Mixed-energy coasts produce barrier island systems that have short islands that are wide at one

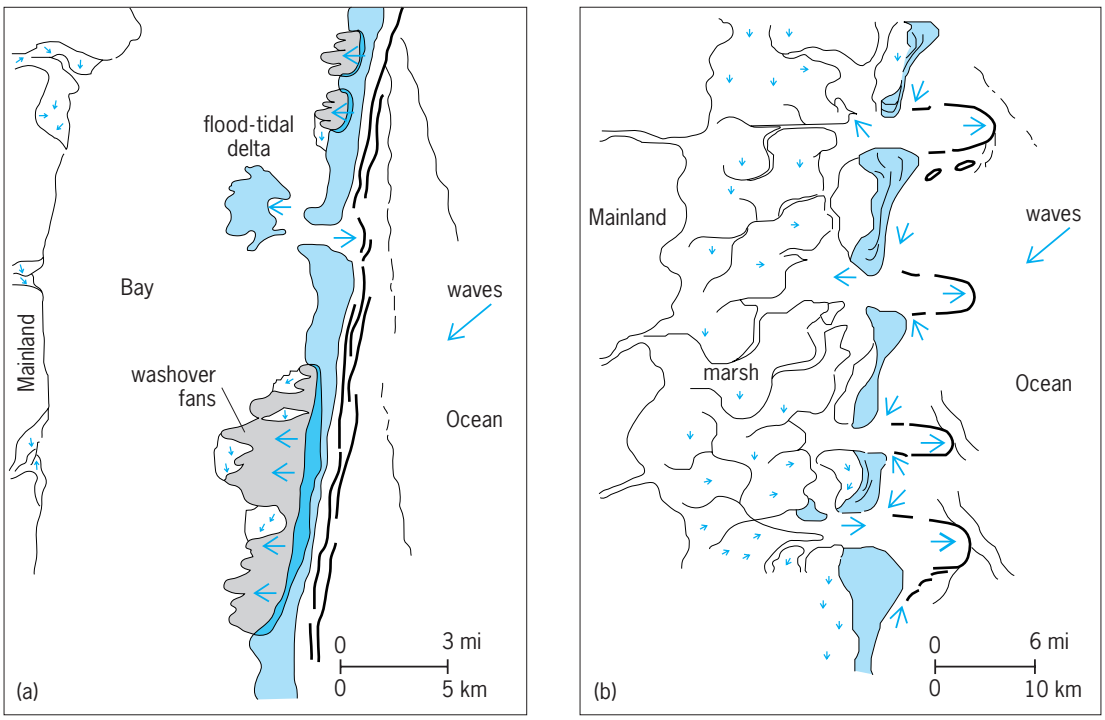

Fig. 3. Maps showing major coastal environments and directions, and directions of coastal processes of (*a*) wave-dominated barrier islands and (*b*) mixed-energy barrier islands. Large arrows indicate wave direction, and small arrows indicate direction of sediment transport. (*After S. P. Leatherman, ed., Barrier Islands from the Gulf of St. Lawrence to the Gulf of Mexico, Academic Press, 1979*)

-s

end with large, fairly stable inlets resulting from the combination of wave and tide influence (Fig. 3*b*). Strong tidal currents in the inlets develop a substantial ebb-tidal delta that has an outer lobe smoothed by wave action. As waves are refracted around the ebb delta, there is a zone of longshore current reversal downdrift of the regional littoral drift direction. This condition causes sediment to become trapped and accumulate as prograding beach ridges. The remainder of the island is wave-dominated and is served by longshore currents moving in the direction of regional littoral drift. This part of the barrier island tends to be narrow and characterized by washover fans. Mixed-energy barriers have been termed drumstick barriers because of this shape.

Sediments and sedimentary structures. Although the barrier island system contains a broad spectrum of sediment types and sedimentary structures, each sedimentary environment can be characterized by its own suite. Knowledge of how these features formed and their spatial relationships with one another is important in recognizing and interpreting ancient barrier island systems preserved in the rock record.

The outer shoreface is typically composed of muddy sand or sandy mud with thin layers of sand. Bioturbation is nearly ubiquitous, and the surface may have inactive ripples or an irregular undulating surface from a previous storm. Discrete sand layers may show small-scale ripple cross-stratification and are related to storm events. The inner shoreface or nearshore zone displays great variety in sedimentary structures and has generally moderately to well sorted sand which may be a mixture of terrigenous and biogenic particles. Bedforms include symmetrical, wave-formed, two-dimensional ripples in the outer part, with three-dimensional megaripples on the seaward side of the longshore bars, where waves steepen producing combined flow conditions, and plane beds near the bar crests, where waves break. This sequence is repeated as each of the longshore bars is traversed by waves. The result is a complex but predictable pattern of bedforms and corresponding stratification types. Rip channels have seaward-oriented bedforms and cross-stratification due to the seaward transport of sediment in this localized environmnent.

The distinct zones within the beach also are readily distinguishable by their sediments and structures. Foreshore sediments are generally moderately to well sorted sand or gravel. Sorting values increase from worst to best from the base of the foreshore to the berm crest. Stratification is pronounced and slopes seaward at angles generally less that 7 or 8°. The great dynamics of this environment and the pronounced temporal changes in energy levels cause a range in the size of sediment particles deposited at any given time; some layers are sand, some are shell hash, and some are a combination. Erosion on the beach is typically marked by lag accumulations of dark, heavy minerals, resulting in thin layers of such minerals as zircon, garnet, magnetite, and rutile.

The backbeach is more uniformly finer-grained and better sorted that the adjacent foreshore. This is a response to eolian-dominated conditions, which may produce a thin surface deflation pavement of shell debris. Stratification is nearly horizontal and may be disrupted by burrows of various arthropods.

Dunes on the barrier island are characterized by fine, well-sorted sand which is dominantly quartz but may also contain carbonate shell debris. Dunes show the highest sorting values in the barrier system. There is a general trend of decreasing grain size and increasing sorting from the lower foreshore to the dunes. The surface of dunes is commonly covered with small wind-generated ripples. Internal stratification is dominated by large-scale cross-stratification formed as the active portion of the dune migrates. Dips range widely from about 10° to greater than 30°; generally a particular barrier is characterized by a given range in cross-stratification dips. Azimuths of the cross-stratification typically reflect prevailing and predominant wind directions.

The low-lying, back-island area is composed of moderately sorted sand dominated by horizontal, plane-bed stratification developed under sheet flow conditions of washover deposition. Roots and benthic invertebrates may destroy some of the stratification. Algal-mat-covered wind tidal flats may be widespread in the supratidal zone of some back-barrier areas. The blue-green algae stabilize the sediment under normal conditions but are torn up during storms, thereby permitting erosion by storm surge waters.

Marshes act as efficient sediment traps and contain a combination of mud, plant debris, and sand. Suspended sediment is transported onto the marsh as the tide rises, especially during storm surge conditions. Particles settle, and are trapped by the grass, resulting in slow accumulation of sediment. Washover fans may encroach upon the marsh, providing sand layers within the peaty muds.

Tidal flats receive sediment as the tides rise and fall, bringing both bedload sand and suspended mud. These sediments commonly accumulate in thin alternations, forming tidal bedding. The subtidal portion of the estuaries associated with barrier islands tends to be dominated by mud with variable amounts of terrigenous sand and shell gravel. Oyster shell debris is an important constituent of many low-energy estuaries, such as along the Atlantic and Gulf coasts of the United States. Some estuaries are influenced by strong tidal currents and have coarser sediment with bedforms, such as the Bay of Fundy in Nova Scotia, Canada.

Lagoons accumulate little terrigenous sediment because of the lack of mechanisms to transport it. Biogenic material and the precipitation of carbonate and evaporite minerals in some lagoons provide most of the sediment, with some contribution from both washover and blowover.

Inlets contain a wide variety of sediment texture and bedforms. Cross-stratification ranges from small-scale to large-scale, depending on the size of the bedforms present. A generally bimodal direction to this cross-stratification may be present, representing flooding and ebbing tidal currents.

Management. Continued pressures of growing populations have caused increased development of barrier islands, primarily for residential, commercial, and tourism purposes. This development along with the apparent increase in the rate of sea-level rise over the past several decades is causing severe problems. Natural barrier island processes include landward migration as washover occurs, and many tidal inlets tend to migrate along the coast. Construction of buildings, roads, and other infrastructure tends to lend some degree of permanence to barrier islands in that these structures are not designed to move as the island environments migrate.

Stabilization caused by structures commonly causes problems in various ways. Seawalls are designed to protect buildings or other structures from wave attack and coastal erosion. They also prohibit washover and landward migration, and commonly fail as wave action scours underneath them. Because of their nature, they prevent beaches from developing as barrier island retreat takes place. Structures that are perpendicular to the coast, such as groins and jetties, interrupt the littoral drift of sediment along the coast. These structures cause considerable sediment to pile up on their updrift side and result in erosion on the downdrift side because of lack of sediment supply. This situation is much like the sediment that is impounded behind a dam in a reservoir. Jetties are similar in their nature and are typically large because of their role in inlet stabilization. These concrete and rip-rap structures do a good job, but in the process they are also effective sediment traps on the updrift side, and the downdrift side of the inlet suffers serious erosion problems due to lack of sediment. Several locations around the United States now have sediment-bypassing systems to counteract this problem, but they are costly and commonly inoperable due to mechanical problems. New technology is needed to provide more efficient sand transfer across structured inlets.

There are federal, state, and local zoning restrictions on construction type and location, and other management programs that are in place. Most states have zoning that restricts construction on barriers through the implementation of some type of minimum setback from the active beach and dune environment. Some also prohibit seawalls and have limitations on rebuilding after loss due to storms. Extensive permitting procedures are involved in nearly all barrier island construction. New strict construction codes have resulted in tremendous reduction in structural damage due to severe storms such as hurricanes. Conditions associated with Hurricane Opal along the Florida panhandle coast in 1995 offer a good example. Of the nearly 2000 homes and other buildings along the open coast that were in the path of this storm, about 1000 incurred serious damage or were destroyed. Not one of the destroyed homes was built under recent construction codes, whereas homes built under these codes experienced no major damage.

Soft construction has become the primary method for protection of upland barrier properties. The most utilized approach is through beach nourishment, but various methods of constructing dunes and extensive vegetation programs are also common. All are designed to provide protection and stabilization to barrier environments while maintaining the esthetic appearance of the area and permitting natural processes to act on the islands.

Beach nourishment is expensive and temporary but has been effective in most places. This involves taking beach sand from a source location, generally offshore, and placing it on the beach in a specific design configuration. This solution typically costs a few million dollars per mile and lasts 5–10 years before additional nourishment is required. The cost of such projects is typically shared among federal, state, and local levels in decreasing proportions.

Although great strides have been made in coastal management over the past few decades, there is still a long way to go. Current regulations need to be enforced, and exceptions—which are too common—minimized. Like so many current environmental problems, the coastal zone is subjected to ever-increasing pressures of population and the related development. Barrier island systems are among the most affected. *See* COASTAL ENGINEERING.

Richard A. Davis, Jr.

Bibliography. R. W. G. Carter and C. D. Woodroffe (eds.), *Coastal Evolution*, 1994; R. A. Davis, Jr. (ed.), *Geology of Holocene Barrier Island Systems*, 1994; M. L. Schwartz (ed.), *The Encyclopedia of Beaches and Coastal Environments*, 1982.

Bartonellosis

An arthropod-borne infection caused by the bacterium *Bartonella bacilliformis*. The disease, also known as Carrión's disease, is confined to South America. It is characterized by two distinct clinical forms, Oroya fever and verruga peruana. Oroya fever causes a high temperature accompanied by the destruction of red blood cells; verruga peruana is a benign, nodular wart that develops on the surface of the skin. *See* ANEMIA.

Epidemiology. Bartonellosis occurs in an area that extends from 150 mi (240 km) north of the Equator to 975 mi (1570 km) south of it. It affects people on the western slopes of the Andes in Peru, Columbia, and Ecuador, at altitudes between 2000 and 8000 ft (600 and 2430 m). The geographic restriction of the disease is attributed to the limited habitats of *Phlebotomus*, its sandfly vector. *Phlebotomus verrucarum*, a night-biting sandfly with a short flight range, is the usual vector, but other species of sandfly have been implicated.

In the restricted geographic area, the disease is endemic. Ten to fifteen percent of the population have been shown to have *B. bacilliformis* in their peripheral blood. These asymptomatic cases, in addition to long-term carriers of the organisms, serve as reservoirs of infection. Animal reservoirs have not been identified.

Prevention of bartonellosis requires control of the vector, generally by spraying exteriors and interiors of buildings with insecticide. Insect repellents and bed netting provide significant personal protection.

Pathogenesis. Infection of nonimmune individuals caused by the bite of an infected sandfly results in the parasitization of erythrocytes. Large numbers of bacteria adhere to and penetrate erythrocytes, causing deep and long-lasting invaginations in erythrocyte membranes with resultant mechanical fragility. In individuals with Oroya fever, the organisms may be seen in Giemsa stains of the peripheral blood and in the cytoplasm of endothelial cells lining the blood vessels and lymphatics, where they occur in rounded masses. As many as 90% of the erythrocytes may be destroyed, leading to the severe anemia that is characteristic of the disease. *See* ANEMIA.

Oroya fever. The distinctive symptoms of the Oroya fever form of bartonellosis appear after an incubation period of 2–6 weeks. They include intermittent high fever, tender enlarged lymph nodes, diffuse muscle and joint pain, and the systemic symptoms associated with severe hemolytic anemia. Concurrent infections with salmonellosis, amebiasis, malaria, and tuberculosis are frequent and contribute to the severity and poor prognosis of Oroya fever. The mortality rate is 50% in untreated cases. Chloramphenicol is the recommended treatment, but the disease also responds well to penicillin, tetracyclines, and streptomycin.

Verruga peruana. The second stage of the disease, verruga peruana, may develop in survivors of Oroya fever after a variable period of time. This phase may also occur in individuals who have had a subclinical infection or who have never had a systemic *Bartonella* infection. The cutaneous phase is characterized by nodules that develop in successive crops over a period of 1–2 months, and may last for months to years. The verrugas (warts) are red to purple, diffuse or confluent, with individual lesions reaching 0.4–0.8 in. (1–2 cm) in diameter. They occur most frequently on exposed parts of the body but may also appear on mucous membranes and viscera. Fever and anemia are usually not present during this stage. *Bartonella* organisms are more difficult to identify in the blood during the cutaneous stage, but may be cultured from the verrugous lesions. The cutaneous lesions exhibit a variable response to antibiotic therapy. In verruga peruana the mortality rate is less than 5%.

Etiologic agent. The causative agent of bartonellosis is a motile, gram-negative, pleomorphic bacillus. In erythrocytes, organisms usually appear as short rods ranging from 1 to 3 micrometers in length and 0.25 to 0.5 μm in width. Forms vary from small coccoid and ring-shaped structures to slender bacilliform bodies. They stain poorly with the bacteriologic aniline dyes, but with the Giemsa stain they assume a reddish purple appearance and can be easily identified. In culture, the organisms possess unipolar flagella, but no flagella have been demonstrated in fresh preparations of clinical specimens.

Bartonella bacilliformis is aerobic, and may be cultured in the laboratory on semisolid agar containing fresh rabbit serum and rabbit hemoglobin. The organism grows over a temperature range of 28–37°C (82–99°F), but maximal longevity occurs at 28°C (82°F). *See* MEDICAL BACTERIOLOGY. Hilda P. Willett

Bibliography. J. C. Bartone, *Bartonella Infections: Index of Modern Information*, 1990; J. P. Kreier and M. Ristic, The biology of hemotrophic bacteria, *Annu. Rev. Microbiol.*, 35:325–338, 1981.

Barycentric calculus

The application of the properties of the centroid of a set of weighted points to geometry. For certain new operations such as the addition of point and point or of point and vector, it is possible to give their vector equivalents and work within the framework of vector algebra.

Theorems. A point P associated with a real number λ (for example, its mass or electric charge) is called a weight point and is denoted by λP. With a given origin O, the point λP has the companion vector $\lambda\overrightarrow{OP}$.

The sum $\mathbf{S} = \Sigma\lambda_i\overrightarrow{OP}_i$ of the companion vectors of n given points will be independent of the origin when, and only when, their total weight $\Sigma\lambda_i = 0$.

$$\text{If} \qquad \mathbf{S}' = \Sigma\lambda_i\overrightarrow{O'P}_i$$

$$\text{then} \qquad \mathbf{S} - \mathbf{S}' = \Sigma\lambda_i(\overrightarrow{OP}_i - \overrightarrow{O'P}_i) = (\Sigma\lambda_i)\overrightarrow{OO'}$$

Hence $\mathbf{S} = \mathbf{S}'$ implies $\Sigma\lambda_i = 0$ and conversely.

The properties of $\Sigma\lambda_i\overrightarrow{OP}_i$ thus depend on whether $\Sigma\lambda_i = 0$ (case 1) or $\Sigma\lambda_i \neq 0$ (case 2).

Case 1. $\Sigma\lambda_i = 0$. $\mathbf{S}$ may be computed by taking O successively at $P_1, P_2, \ldots, P_n$; then all n vectors $\Sigma_{i\neq j}\lambda_i\overrightarrow{PP}_i$ $(j = 1, 2, \ldots, n)$ are equal to $\mathbf{S}$. This is one source of geometric theorems.

In particular, the vector $\mathbf{S}$ may be zero. When $n = 2, 3, 4$, the points are coincident, collinear, or coplanar, respectively. Writing $\overrightarrow{OA} = \mathbf{a}$, $\overrightarrow{OB} = \mathbf{b}, \ldots$, gives the important theorems I and II.

Theorem I. Three distinct points A, B, C are collinear when, and only when, there are three nonzero constants α, β, γ such that Eqs. (1) hold.

$$\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} = \mathbf{0} \qquad \alpha + \beta + \gamma = 0 \tag{1}$$

Theorem II. Four distinct points A, B, C, D, no three collinear, are coplanar when, and only when, there are four nonzero constants $\alpha, \beta, \gamma, \delta$ such that Eqs. (2) hold.

$$\alpha\mathbf{a} = \beta\mathbf{b} = \gamma\mathbf{c} = \delta\mathbf{d} = \mathbf{0} \quad \alpha + \beta + \gamma + \delta = \mathbf{0} \tag{2}$$

The point C is said to divide the segment AB of a straight line in the ratio β/α if $\alpha\overrightarrow{AC} = \beta\overrightarrow{CB}$ or $\alpha(\mathbf{c} - \mathbf{a}) = \beta(\mathbf{b} - \mathbf{c})$; then Eq. (3) holds.

$$\mathbf{c} = \frac{\alpha\mathbf{a} + \beta\mathbf{b}}{\alpha + \beta} \tag{3}$$

Thus Eq. (1) implies that A, B, C divide BC, CA, AB in the ratios $\gamma/\beta, \alpha/\gamma, \beta/\alpha$.

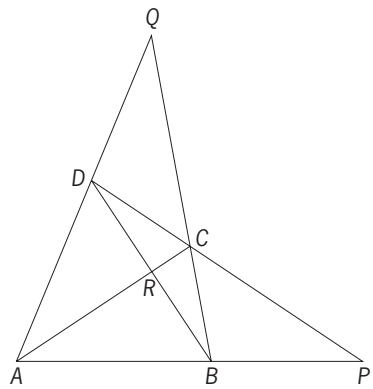

Complete plane quadrilateral and its diagonal points.

Equation (2) implies that $ABCD$ is a plane quadrilateral whose diagonal points P, Q, R (see **illus.**) are given by Eqs. (4). Metric and projective properties

$$\mathbf{p} = \frac{\alpha\mathbf{a} + \beta\mathbf{b}}{\alpha + \beta} \quad \mathbf{q} = \frac{\beta\mathbf{b} + \gamma\mathbf{c}}{\beta + \gamma} \quad \mathbf{r} = \frac{\gamma\mathbf{c} + \alpha\mathbf{a}}{\gamma + \alpha} \tag{4}$$

of the complete quadrilateral readily follow.

Case 2. $\Sigma\lambda_i = \lambda \neq 0$. The sum $\mathbf{S}_0 = \Sigma\lambda_i\overrightarrow{OP}_i$ now depends on O. But Eq. (5) defines the same point λP^*

$$\Sigma\lambda_i\overrightarrow{OP}_i - \lambda\overrightarrow{OP^*} = \mathbf{0} \tag{5}$$

for any choice of O; for the total weight of the points $\lambda_i P_i$ and $-\lambda P^*$ is zero. If O is taken at P^*, this gives the basic relation defining P^* according to Eq. (6). The

$$\Sigma\lambda_i\overrightarrow{P^*P_i} = \mathbf{0} \tag{6}$$

point λP^* is called the centroid of the points $\lambda_i P_i$. For an arbitrary origin the centroid is given by Eq. (7).

$$\lambda\mathbf{p}^* = \Sigma\lambda_i\mathbf{p}_i \tag{7}$$

On the right the sum extended over any subset of nonzero weight may be replaced by their centroid. This replacement principle is a second source of geometric theorems. It also reduces the problem of centroid location to that of two points.

If $\alpha = \beta \neq 0$, the centroid of αA and βB is the point $(\alpha + \beta)C$ given by $(\alpha + \beta)\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b}$; and Eq. (3) shows that C divides AB in the ratio β/α. The centroid of A and B is $2\mathbf{c} = \mathbf{a} + \mathbf{b}$, and C is the midpoint of AB.

The centroid $3D$ of A, B, C is $3\mathbf{d} = \mathbf{a} + \mathbf{b} + \mathbf{c}$. Since Eq. (8) holds, the medians of a triangle meet in

$$3\mathbf{d} = \mathbf{a} + 2\frac{\mathbf{b} + \mathbf{c}}{2} = \mathbf{b} + \frac{\mathbf{c} + \mathbf{a}}{2} = \mathbf{c} + \frac{\mathbf{a} + \mathbf{b}}{2} \tag{8}$$

a point which divides each in the ratio 2:1.

The meaning of $\Sigma\lambda_i\mathbf{p}_i = \mathbf{0}$ when $\Sigma\lambda_i = 0$ is now apparent. If such a set is divided into two groups of nonzero weight by transposing some $\lambda_i\mathbf{p}_i$, the subsets have centroids at the same weighted point, whose weights differ only in sign.

Barycentric coordinates. If ABC is a triangle of reference, weights α, β, γ can be applied to A, B, C so that the centroid of αA, βB, γC is at any given point P in the plane of ABC; Eq. (9) holds; only the

$$(\alpha + \beta + \gamma)\mathbf{p} = \alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} \tag{9}$$

ratios of the barycentric coordinates α, β, γ are determined; and Eq. (10) can be written, where PBC

$$\frac{\alpha}{PBC} = \frac{\beta}{PCA} = \frac{\gamma}{PAB} = \frac{\alpha + \beta + \gamma}{ABC} \tag{10}$$

is the area of the triangle PBC taken + or − depending on whether the circuit PBC is taken in the same sense as ABC or the opposite.

The equation $l\alpha + m\beta + n\gamma = 0$ represents a straight line; and $\alpha + \beta + \gamma = 0$ is the line at infinity.

In 3-space the weights α, β, γ, δ given to the vertices of a reference tetrahedron $ABCD$ are the barycentric coordinates of the centroid of αA, βB, γC, δD. $l\alpha + m\beta + n\gamma + k\delta = 0$ represents a plane; and $\alpha + \beta + \gamma + \delta = 0$ is the plane at infinity. The ratios of α, β, γ, δ may now be given in terms of oriented tetrahedra. *See* CENTROIDS (MATHEMATICS).

Louis Brand

Baryon

Strongly interacting composite particle that accounts for almost all of the visible mass of the universe. The proton and neutron that make up the atomic nucleus are examples of baryons. Strongly interacting particles (hadrons) are either baryons, which are fermions with half-integral spin, or mesons, which are bosons with integer spin. *See* FUNDAMENTAL INTERACTIONS; HADRON; MESON; NEUTRON; PROTON; QUANTUM STATISTICS.

Although baryons are very small (roughly 10^{-15} m in diameter), they consist of even smaller, elementary particles called quarks, together with the exchange particles of the strong interaction between the quarks, the gluons. This strong interaction is based on a three-valued "color" degree of freedom carried by the quarks. Modern particle theory has an as-yet-unproven postulate called confinement, which states that the only strongly interacting particles that can emerge from interactions and exist in isolation are those that carry no net color. This explains why quarks have not been seen in isolation. As the smallest number of colored quarks which can make up a colorless state is three, just as colorless (white) light is made from the three primary colors, baryons must contain at least three quarks. These are referred to as valence quarks. By contrast, mesons are made up of a minimum of one quark-antiquark pair, which can also be colorless. *See* COLOR (QUANTUM MECHANICS); GLUONS; QUARKS.

Because the interaction that binds quarks into a baryon is very strong, any number of quark-antiquark pairs, which are generically called sea quarks, can also be present in a baryon. Any antiquark which exists in a baryon must be a sea quark, but it may not be possible to distinguish between a sea quark and a valence quark with the same intrinsic properties.

Baryons are identified by properties referred to as

The six flavors of quarks, their charges, and their masses

Charge = $+2e/3$	Flavor:	up (u)	charm (c)	top (t)
	Mass, MeV/c^2:	3	1250	174,000
Charge = $-e/3$	Flavor:	down (d)	strange (s)	bottom (b)
	Mass, MeV/c^2:	6	120	4250

flavor, spin, and parity, all of which have discrete values. They also have measurable properties which are continuous, such as their mass and, for all but the proton which is the lightest baryon and so stable against decay into a lighter baryon plus other particles, their decay lifetime. The eventual product of all baryon decays is a proton. If more than one type of decay chain is possible, the fraction of decays which proceed through one mode or another may also be measured. A description of how the discrete properties of baryons are found from those of their constituent particles follows, and involves some interesting physical ideas.

Flavor. There are six different kinds, or flavors of quarks, in three families (see **table**). The lightest family contains the up (u) and down (d) quarks, with masses of roughly 3 and 6 in convenient units of MeV/c^2 (where MeV is a million electronvolts, the kinetic energy gained by an electron that is accelerated by a million-volt potential, and use is made of Einstein's relation $E = mc^2$ or $m = E/c^2$), and charges of $+^2/_3$ and $-^1/_3$ in units of the electron charge (e), respectively. These light quark masses are very small, on the scale of the mass of the proton, 938 MeV/c^2. *See* ELECTRONVOLT; FLAVOR.

The proton is made up of two up quarks and a down quark, labeled *uud*, and the neutron has flavor *udd*. The neutron is only slightly more massive than the proton, by about 1.3 MeV/c^2. Roughly 3 MeV/c^2 of mass is added by changing an up quark to a down quark, but this is largely compensated for by a loss in electromagnetic self-energy made by the same change. This is the energy required to collect the on-average positively charged quarks into the proton against the electrostatic repulsion. The resulting small size of the neutron-proton mass difference prevents neutrons from decaying when bound into most nuclei. If the down quark were significantly more massive than the up quark, the visible universe would be made up of only hydrogen atoms, as the neutrons in all nuclei would decay to protons, which cannot bind into nuclei by themselves.

Because the difference between the masses of the up and down quarks and the electromagnetic self-energy of quarks in a baryon are so small compared to baryon masses, the properties of baryons are to a very good approximation unchanged by replacing an up quark with a down quark. This is referred to as isospin symmetry. In this language the proton and neutron are referred to collectively as different charge states of the nucleon. *See* I-SPIN.

There is a third quark that is light on the scale of the proton mass, the strange quark s, with a mass of about 120 MeV/c^2 and charge $-e/3$. Baryons that differ from the proton and neutron by the conversion of an up or a down quark into a strange quark are denoted Λ or Σ. There are three Σ charge states: Σ^+ with quark content *uus*, Σ^0 with quark content *uds*, and Σ^- with quark content *dds*. The Λ^0 also has quark content *uds*, but differs from Σ^0 by its isospin symmetry. The flavor-dependent part of its quantum-mechanical wave function changes sign under switching the up and down quarks, while that of the Σ states does not. Baryons that contain one up or down quark and two strange quarks are called cascade or Ξ baryons, with $\Xi^0 = uss$ and $\Xi^- = dss$. Baryons that contain three strange quarks, called Ω^-, have also been observed. Baryons containing strange quarks are collectively known as hyperons. *See* HYPERON; STRANGE PARTICLES.

The remaining three flavors of quarks, the charm c, bottom b, and top t quarks, are heavy on the scale of the proton mass (see table). Charmed baryons made up of a pair of up or down quarks and a charmed quark are called Λ_c and Σ_c, in analogy to the equivalent strange baryons. Baryons with a charmed quark, a strange quark, and an up or down quark have been observed, and are denoted Ξ_c, and those with a charmed quark and two strange quarks are denoted Ω_c. Some evidence for a doubly charmed *ccu* or *ccd* baryon, denoted Ξ_{cc}, has been found. The naming scheme for baryons containing a bottom quark is similar, although so far only Λ_b has been definitively observed in nature. The top quark is very massive, and this causes it to decay so rapidly that it does not have time to form a hadron by binding strongly to other quarks or to an antiquark. *See* CHARM.

Spin. Quarks carry an intrinsic degree of freedom called spin. This is related to the angular momentum carried by a macroscopic object when it rotates, felt by everyone who has tried to stop a rotating merry-go-round or bicycle wheel. In the very short distance, quantum world, angular momentum is quantized in units of $\hbar = h/(2\pi)$, where h is Planck's constant. The smallest possible angular momentum is $\hbar/2$, and this is the spin of each of the quarks making up a baryon. Such particles are referred to as spin-$^1/_2$ particles, or spinors. To find the spin of baryons made up of quarks and gluons, first consider the rules for combining the spins of three spin-$^1/_2$ particles. Two spin-$^1/_2$ particles can have their spins combined to 0 (antialigned along an axis called the quantization axis) or to 1 (aligned along the quantization axis), in units of $\hbar$. With the addition of a third quark there are two ways to get to a spin of $^1/_2$, namely, combining spin 0 with spin $^1/_2$ or spin 1 with spin $^1/_2$; and one way to get a spin of $^3/_2$, namely, combining spin 1 with spin $^1/_2$. This means that the contribution to the spin of a baryon from the spin of the three valence quarks that must be present is either $^1/_2$ or $^3/_2$. *See*

ANGULAR MOMENTUM; SPIN (QUANTUM MECHANICS).

The gluons, which mediate the color force between the quarks, like the photons, which mediate the electromagnetic force between particles, carry a spin angular momentum of $\hbar$ and are referred to as spin-1 or vector particles. In addition to the spin angular momentum carried by the quarks and gluons, these particles also have orbital angular momentum as they move around their common center of mass. In the simplest possible (naïve quark model) picture, the spin angular momentum of a baryon is made up of the orbital angular momentum and intrinsic spin of the three valence quarks, and the contribution of the sea quarks and gluons is neglected.

Just as the lightest states of the hydrogen atom are spherically symmetric and have zero orbital angular momentum between the electron and proton, in this simple picture the lightest states of a baryon have zero orbital angular momentum between the quarks, and so their total spin is just that of the quarks, either $^1/_2$ or $^3/_2$. The proton and neutron do indeed have spin $^1/_2$, and the lightest excited baryon containing only the *u* and *d* quarks which make up the proton and neutron, called $\Delta(1232)$, has spin $^3/_2$. The mass of the Δ is given in brackets after its name, in MeV/c^2. It has four charge states, Δ^{++}, Δ^{+}, Δ^{0}, and Δ^{-}, consisting of *uuu*, *uud*, *ddu*, and *ddd* quarks. The mass difference between the Δ and the nucleon can be thought of as due to a kind of magnetic interaction between the colored, spinning quarks, analogous to the magnetic interaction between charged, spinning electrons. These interactions are repulsive when the quark spins are aligned; this raises the mass of the spin-$^3/_2$ Δ, with all three quark spins aligned, relative to that of the nucleon, with only two of the quark spins aligned. *See* DELTA RESONANCE; ELECTRON SPIN.

Baryons with the same quark content as the proton, neutron, and four Δ charge states exist with spin angular momentum higher than $^1/_2$ or $^3/_2$. In most cases this is because the quarks have orbital angular momentum around their common center of mass, but it is possible for the gluons and sea quarks present in baryons to carry angular momentum. Since orbital angular momentum comes in integral multiples of $\hbar$, which is also true of the intrinsic spin of a gluon or quark-antiquark pair, all baryons must have half-integral spin $J = {}^1/_2, {}^3/_2, {}^5/_2$, and so forth. Particles with half-integral spin are called fermions. Nucleon and Δ-flavored baryons containing only *u* and *d* valence quarks have been clearly observed with spins up to $^{11}/_2$, and states with spins up to $^7/_2$ have been observed for baryons containing strange quarks.

Exchange symmetry and color. The lightest baryons with a particular flavor content and isospin symmetry [like the nucleon and the $\Delta(1232)$, containing only up or down valence quarks] have either spin $^1/_2$ or spin $^3/_2$. Such states are called ground-state baryons. Since quarks have spin $^1/_2$, they are fermions, and the Pauli principle implies that the quantum-mechanical wave function of a baryon made up of quarks must be antisymmetric (change sign) under the exchange of any pair of identical quarks. When this idea is put together with the (isospin) symmetry of the strong interactions under the exchange of up and down quarks, the result is that the wave functions of the nucleon and Δ must be antisymmetric under exchange of any pair of valence quarks. *See* EXCLUSION PRINCIPLE.

In the case of the ground state Δ^{++}, this leads to a paradox. The flavor content is *uuu*, which is exchange symmetric. The three quark spins are maximally aligned to spin $^3/_2$, which is also symmetric. The $\Delta(1232)$ is the least massive Δ-flavored baryon and so should not have orbital angular momentum, which would significantly increase its mass. Orbital angular momentum zero states are known to be spherically symmetric and exchange symmetric. How can the wave function which describes the Δ^{++} be antisymmetric, as required by the Pauli principle?

The resolution of this paradox led to the postulate that the quarks carry an additional degree of freedom. The minimal assumption required for antisymmetry is that this degree of freedom is three-valued. This degree of freedom was called color. The modern theory of the strong interactions, called quantum chromodynamics (QCD), grew out of the conviction that the interaction between colored quarks and antiquarks is responsible for binding them into baryons and mesons, just as the interaction between charged nuclei and electrons is what binds them into atoms. *See* QUANTUM CHROMODYNAMICS.

Parity. Another interesting discrete property of a baryon is its parity. The strong interactions between colored quarks have the property that they do not change when viewed in a mirror. This is essentially the same as spatial inversion symmetry, invariance under the operation of reflecting all of the spatial coordinates through the origin. As a consequence, under the operation of inversion of the coordinates, the quantum-mechanical wave functions of states made up of strongly interacting particles either remain the same, called positive parity, or change sign, called negative parity. Also, if a system goes from one state to another via a strong interaction, the parity of the entire system cannot change. By examining the strong decays of excited baryons it is possible to determine if they have positive parity, the same as that of the proton, or negative. All ground-state baryons have positive parity.

After ground-state baryons, the next most massive excited states are those with negative parity, for all flavors of baryon. These states have the lowest possible orbital angular momentum, $\hbar$, which can result from the quarks orbiting around their common center of mass. The quantum-mechanical wave function of such states is known to change sign under the operation of spatial inversion. Negative-parity excited states have been discovered for nucleons and Δ-flavored baryons, as well as for Λ, Σ, and Ξ hyperons, and charmed baryons Λ_c and Σ_c. *See* PARITY (QUANTUM MECHANICS).

Strong decays. The $\Delta(1232)$ decays to a proton and a pion in a very short time (roughly 10^{-24} s) via the strong interaction. The same is true of the equivalent spin-$^3/_2$ $\Sigma(1385)$ baryon, which decays

to the spin-$^1/_2$ Σ baryon. The pion is made up of a minimum of an up or down quark and an up or down antiquark, and is the lightest of the mesons, particles made up of a minimum of one quark-antiquark pair. Because the difference between the mass of the Δ and the nucleons is more than 140 MeV/c^2, the mass of the pion, it can decay very rapidly to a nucleon and a pion via the strong interaction. If the strong decay of any baryon into a less massive state consisting of a baryon and a meson is allowed by conservation of energy, then this will be how it decays.

In one simple picture of such strong decays, a quark-antiquark pair is created via the strong interaction, with the created quark ending up in the final baryon and the created antiquark in the final meson. The flavor of the created quark must be the same as that of the created antiquark, leading to no net change of flavor between initial and final states linked by a strong interaction. The strong decay of the Δ proceeds by the creation of an up or down quark-antiquark pair. More highly excited nucleon and Δ-flavored baryons can also decay via the creation of a strange-quark pair, creating a Λ or Σ hyperon and a K meson containing a strange antiquark, if their mass exceeds the sum of the masses of these final-state particles.

Electromagnetic and weak decays. Certain decays of a baryon would result in a state with the same quark content, but cannot proceed via the strong interaction because of conservation of energy. An example is the decay of the Σ^0 to the Λ^0, with masses that differ by less than a pion mass. Such decays proceed via the emission of an energetic particle of light, a gamma-ray photon, as the electromagnetic interaction between quarks, made possible by their electric charges, is the next strongest of the fundamental interactions between them. Electromagnetic decays are rapid; the lifetime of the Σ^0 is about 10^{-19} s. *See* GAMMA RAYS; PHOTON.

The baryon eventually resulting from a chain of strong and possibly electromagnetic decays will be a ground-state, spin-$^1/_2$ baryon whose quark content may differ from *uud*, that of the proton. In that case, to decay to the proton, this baryon must change the flavor of one of its quarks, but neither the strong nor the electromagnetic interactions can bring about such a change. Such decays proceed significantly more slowly via the weak interactions, with typical lifetimes of between 10^{-13} and 10^{-10} s, depending on the mass difference between the initial and final baryons involved. An exception is the weak decay of the down quark to the up quark, which allows an isolated neutron to decay to a proton, with an 886-s lifetime because of the tiny (1.3-MeV/c^2) neutron-proton mass difference.

Baryon decays that proceed via the weak interactions are mediated by the massive vector-boson particles W^+, W^-, or Z^0. Such decays can result in the creation of a charged lepton and a neutrino of the same family from a W boson, such as an electron and an electron antineutrino, or the electron's antiparticle (the positron) and an electron neutrino, and are known as semileptonic decays. An important example is the beta decay of the neutron to a proton, electron, and electron antineutrino. If allowed by conservation of energy, the more massive mu and tau leptons may be created. *See* ELECTRON; INTERMEDIATE VECTOR BOSON; LEPTON; NEUTRINO.

Alternatively, a quark-antiquark pair with mismatched flavors can be created from the decay of a weak vector boson, and the result is a flavor-changing decay involving a baryon and a meson in the final state. For example, the Σ^+ decays to a proton and a π^0, or a neutron and a π^+, removing a strange quark but maintaining the electric charge of the initial state, with roughly equal probability. Such decays are called nonleptonic weak decays. *See* WEAK NUCLEAR INTERACTIONS.

Baryon production. Since only protons and neutrons occur naturally in our environment, the knowledge that we have gained about other baryons has required their production in experiments involving collisions of particles. This production proceeds via the strongest of the mechanisms causing the decays of baryons, the strong and electromagnetic interactions. Nucleon and Δ-flavored baryons can be created by scattering a beam of relatively long-lived pions off nuclear targets which contain protons and neutrons, or by beams of energetic photons. Similarly, Λ- and Σ-flavored baryons can be made by scattering beams of K mesons, which contain the required strange quark, off nuclear targets. Baryons containing more strange quarks are produced by the strong decay of a Λ or Σ via the creation of a strange quark-antiquark pair. Those containing more massive quarks can be created in an electron-positron or proton-antiproton collider experiment, where these particles annihilate and form a charmed or bottom quark-antiquark pair. These states ultimately form a number of hadrons (with an equal number of baryons and antibaryons), which can include a heavy-quark baryon. *See* ELEMENTARY PARTICLE; PARTICLE ACCELERATOR. Simon Capstick

Bibliography. F. Close, *The Cosmic Onion: Quarks and the Nature of the Universe*, Heinemann Educational Secondary Division, 1984; S. Eidelman et al. (Particle Data Group), Review of particle physics, *Phys. Lett.*, B592:1, 2004; N. Isgur and G. Karl, Hadron spectroscopy and quarks, *Phys. Today*, 36(11):36–42, 1983.

Basalt

An igneous rock characterized by small grain size (less than about 0.2 in. or 5 mm) and approximately equal proportions of calcium-rich plagioclase feldspar and calcium-rich pyroxene, with less than about 20% by volume of other minerals. Olivine, calcium-poor pyroxene, and iron-titanium oxide minerals are the most prevelant other minerals. Certain rare varieties of basalts are poor in SiO_2 and rich in melilite, olivine, and nepheline and may be poor in calcium-rich plagioclase and pyroxene. Most basalts are dark gray or black, but some are light gray. Various structures and textures of basalts are useful in

inferring both their igneous origin and their environment of emplacement. Basalts are the predominant surficial igneous rocks on the Earth, Moon, and probably other bodies in the solar system. Several chemical-mineralogical types of basalts are recognized. The nature of basaltic rocks provides helpful clues about the composition and temperature within the Earth and Moon. The magnetic properties of basalts are responsible in large part for present knowledge of the past behavior of the Earth's magnetic field and of the rate of sea-floor spreading. *See* IGNEOUS ROCKS.

Extrusion of basalt. Basalt erupts out of fissures and cylindrical vents. Fissures commonly are greater than several feet wide. Cylindrical vents generally are many tens of feet in diameter and may pass downward into cylindrical shell-shaped or planar fissures. Repeated or continued extrusion of basalt from cylindrical vents generally builds up a volcano of accumulated lava and tephra around the vent. Fissure eruptions commonly do not build volcanoes, but small cones of tephra may accumulate along the fissure. *See* LAVA; VOLCANO; VOLCANOLOGY.

Form and structure. Basalts display a variety of structures mostly related to their environments of consolidation. On land, basalt flows form pahoehoe, aa, and block lava, while under water, pillow lava is formed. Basaltic lava flows on land have rugged to smooth surfaces. The rugged flows are called aa and have top surfaces made of spines and ridges up to several meters high. They are typically composed of three layers: a massive interior overlain and underlain by rubbly zones. Flows with smooth surfaces are called pahoehoe. Pahoehoe has been seen to form on fluid lava flows as the incandescent surface rapidly advances (3–33 ft/s or 1–10 m/s), cools below red heat, and stiffens into a rigid crust about an inch thick in about a minute. There are transitional types of surfaces, as well as a separate kind called block lava, which is similar to aa lava but has a surface made up of equidimensional blocks up to about 3 ft (1 m) in diameter.

Pillow lava is characterized by lens-shaped bodies (pillows) of basalt up to several feet in length which have upward convex tops and partially upward convex bottoms conforming to the convex tops of underlying pillows. Such structure is found in basalts dredged from the deep-sea floor, in basalts which extruded under ice caps, as in Iceland, and in basalts which have flowed into lakes, rivers, and the sea from vents on land.

Basalt occurs as pumice and bombs as well as lava flows. Commonly, basaltic pumice is called scoria to distinguish it from the lighter-colored, more siliceous rhyolitic pumice. Basaltic pumice can have more than 95% porosity. Bombs are relatively smooth, round bodies up to about 3 ft (1 m) in diameter whose surfaces commonly are cracked as a result of expansion of the effervescing interior which remains fluid longer than the quickly cooled surface. Bombs with cracked surfaces are called bread-crust bombs, the name attesting both to the fluid condition and to the gas content of the extruded material. Pele's hair and Pele's tears are terms applied to threads and drops of basaltic glass blown out of vents. Pasty material intermediate between pumice and bombs is called spatter and commonly forms agglutinate or vent breccia near vents of basalts. Accumulations of pumice, bombs, and spatter commonly develop around vents and form cinder cones, spatter cones, and spatter ramparts. Similar, but commonly much smaller, secondary vent deposits may develop on lava flows distant from the primary vent. Secondary vents are rooted within the flow and do not penetrate it; they are called hornitos. Littoral cones may form where a lava flow encounters abundant water and water vapor blows through the liquid lava. *See* VOLCANIC GLASS.

Basalt commonly contains tiny round voids called vesicles. In pumice, vesicles generally make up most of the volume. The vesicle porosity in flows is high at the top and bottom of flows, but it decreases dramatically a few meters away from cooling surfaces. In general, vesicles are smallest (0.04–0.4 in. or 1–10 mm in diameter) in the top and bottom couple of feet of flows, reach a maximum of a couple of inches in diameter several feet from the cooling surface, and then decrease in size and abundance toward the center of the flow. Vesicles in basalts on the sea floor (about 3 mi or 5 km depth) are very tiny (about 10 μm diameter), and the total porosity is commonly less than 1%. Many pillow basalts now on land lack vesicles or have only a few tiny vesicles and presumably formed in deep water. However, there are basalts on land which are vesicle-poor, and some sea-floor basalts have significant vesicles. In sum, the presence of vesicles indicates that basalts evolve gas in low-pressure environments, and the distribution of vesicles within flows provides a means of identifying the tops and bottoms of flows. Flows of basalt develop a characteristic kind of fracture called columnar joining, whereby the rock splits into polygonal columns oriented perpendicular to the cooling surfaces.

Physical properties. Many physical properties of basalt have been measured in the field and laboratory, making it one of the best-characterized rock types. The physical properties of basalts vary greatly because of the range of textures and compositions. Temperature ranges from 1830 to 2230°F (1000 to 1220°C) for flowing basalt. The viscosity for magma is 10^3–10^6 poise (10^2–10^5 N · s/m^2), increasing as the temperature falls and the crystal content of the magma increases. The surface tension for magma is 270–350 dynes/cm (0.27–0.35 N/m). For melt the density is 2.6–2.7 g/cm^3 (1.5–1.6 oz/in.^{3}) at 2190°F (1200°C), while for crystalline rock at standard temperature and pressure (STP), void-free (calculated from simple mixtures of pyroxene, plagioclase, and iron oxides consistent with the definition of basalt), density is 2.8–3.1 g/cm^3 (1.6–1.8 oz/in.^{3}). The velocity of sound is 1.4 mi/s (2.3 km/s) for a compressional wave in melt at 2190°F (1200°C), and 3.5 mi/s (5.7 km/s) for rock at 1290°F (700°C). Both isothermal and adiabatic compressibilities are closely equal to 7×10^{-12} cm^2/dyne for melt at 2190°F (1200°C),

and 2×10^{-12} for rock at 1470°F (800°C). The crushing strength and cohesive strength of cold rock are 1700–2200 atm (170–220 megapascals) and 320–440 atm (32–44 MPa), respectively. The coefficient for volumetric thermal expansion of melt is about 3.8×10^{-5} per °F (2×10^{-5} per °C). The electrical conductivity is about 100 per ohm per centimeter for melt at 2190°F (1200°C). Thermal conductivity is about 1–3 $\times 10^{-3}$ cal/cm s °C for melt at 2190°F (1200°C), and 4–6 $\times 10^{-3}$ cal/cm s °C (1 cal = 4.19 joules) for rock at STP. Heating capacity at constant pressure is about 0.2–0.3 cal per °C per gram. About 100 cal/g is generally assumed for the heat of melting, but the figure is possibly uncertain by 30%. Theoretical and observed diffusion coefficients for Mg^{2+} in basaltic melt at 2140°F (1170°C) are 0.62×10^{-9} in.^{2}/s (4×10^{-9} cm^2/s) and 11×10^{-9} in.^{2}/s (72×10^{-9} cm^2/s).

The magnetic susceptibility of basalts is commonly about 10^{-4} to 4×10^{-3} emu/g. The natural remanent magnetization intensity of basalts depends upon the amount of magnetite, the composition of the magnetite, the grain size, the cooling history, and the paleointensity of the Earth's field. Observed values of natural remanent magnetization intensity are commonly between 10^{-4} and 10^{-3} emu/g for subaerial basalts, but they increase to 10^{-2} to 10^{-1} emu/g for basalts from the sea floor. Thus, the remanent magnetization of submarine basalts is stronger than their induced magnetization—a relationship that has made it possible for the sea floor to preserve a "tape-recorded" history of the reversals of the Earth's magnetic field. The natural magnetization of many basalts is sufficiently strong to be detected with a magnetic compass. *See* MAGNETOMETER; ROCK MAGNETISM.

Mineralogy and texture. The mineralogy and texture of basalts vary with cooling history and with chemical composition. With slow cooling (few degrees Fahrenheit per day 33 ft or 10 m deep in a flow), crystals grow to large sizes (millimeters) and acquire equant shapes. In basaltic melt, diffusion may limit the rate at which a given crystal-liquid boundary can move because the crystals differ in composition from the melt. If crystallization is forced at a rapid rate by rapid cooling (about 180°F or 100°C per minute 4 in. or 10 cm deep in the flow), either the crystals have high ratios of surface area to volume (needles, plates) or many seeds (nuclei) develop. Crystals which grow rapidly under strongly supercooled conditions have different compositions from those which form if slower cooling prevails. Consequently, the same basalt flow may contain different minerals in variously cooled parts of the flow.

The size of crystals in basalt reflects the rate of production of seed crystals (or nuclei) as well as the rate of growth of individual crystals. Crystals of pyroxene are able to grow at an average linear rate of about 8×10^{-10} in./s (2×10^{-5} μm/s), or 8×10^{-5} in./day (2 μm/day). At higher temperatures, in melts rich in magnesium and iron, crystals may grow as rapidly as 24×10^{-8} in./s (6×10^{-3} μm/s), or about 0.02 in./day (0.5 mm/day). The large crystals (phenocrysts) present in most basalts possibly formed in a few hours or days during the ascent of the magma through the crust. Alternatively, many large crystals possibly formed slowly in a body of stored magma. Such stored magma bodies, if thicker than about 300 ft (100 m), take many decades to cool, and their crystals may remain suspended or be picked up from the walls when magma erupts. *See* PHENOCRYST.

As basalt crystallizes, both the minerals and the residual melt change in composition because of differences between the composition of the melt and the crystals forming from it. In basalts, because of the rapid cooling, there is little chance for crystals to react with the residual melt after the crystals have formed. Early-formed crystals tend to become armored with later overgrowths. Because diffusion is sluggish in crystals even at the high temperatures at which basalts crystallize, the minerals generally retain their compositional zoning during cooling. Completely solid basalts generally preserve a record of their crystallization in the zoned crystals and residual glass. **Figures 1** and **2** illustrate the development of crystalline texture in basalt. **Figure 3** shows the sequence of crystallization in basaltic lava on Hawaii.

Basalts may alternatively yield a silica-saturated residual melt (rhyolite) or one deficient in silica to form nepheline, depending upon the initial bulk composition of the basalt. Basalts yielding nepheline upon crystallization are termed alkaline; those yielding rhyolite are called subalkaline (or tholeiitic).

Most basalts contain minor amounts of chromite, magnetite, ilmenite, apatite, and sulfides in addition to the minerals mentioned above. Chromite and the sulfides commonly form early in the sequence of crystallization, whereas magnetite, ilmenite, and apatite form late. Magnetite in basalts contains a history of the strength and orientation of the Earth's magnetic field at the time of cooling. Therefore, although magnetite is minor in amount, it is probably the most important mineral in terrestrial basalts, because it enables earth scientists to infer both the magnetic history of the Earth and the rate of production of basaltic ocean floor at the oceanic ridges.

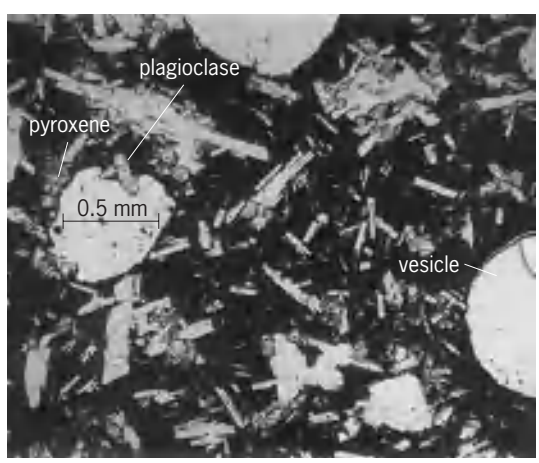

Fig. 1. Thin section of upper crust of lava about 1.2 in. (3 cm) from top; black is microcrystalline groundmass. 0.5 mm = 0.02 in.

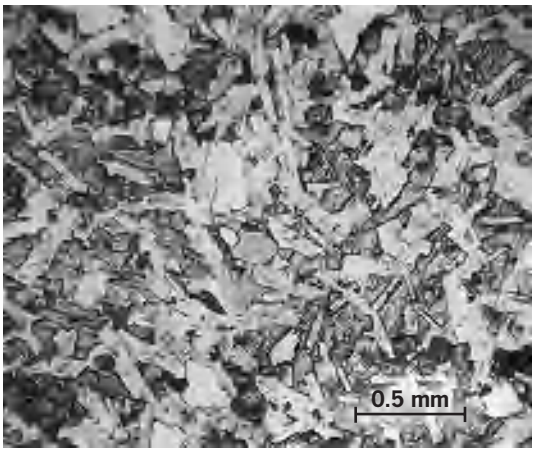

Fig. 2. Thin section of interior of lava shown in Fig. 1, but 26 ft (7.9 m) below the top of the flow. Clear plagioclase crystals are ophitically enclosed in large pyroxene grains up to 0.04 in. (1 mm) in diameter. Crystals of black magnetite are minor. A few irregular voids are present.

Chemical compositions of basalts, in weight percent*

	1†	2	3	4	5	6
SiO_2	49.92	49.20	49.56	51.5	45.90	45.5
TiO_2	1.51	2.32	1.53	1.1	1.80	2.97
Al_2O_3	17.24	11.45	17.48	17.1	15.36	9.69
Fe_2O_3	2.01	1.58	2.78	n.d.	1.22	0.00
Cr_2O_3	0.04	n.d.	n.d.	n.d.	n.d.	0.50
FeO	6.90	10.08	7.26	8.9	8.13	19.7
MnO	0.17	0.18	0.14	n.d.	.08	0.27
MgO	7.28	13.62	6.97	7.0	13.22	10.9
CaO	11.85	8.84	9.99	9.3	10.71	10.0
Na_2O	2.76	2.04	2.90	4.3	2.29	0.33
K_2O	0.16	0.46	0.73	0.80	0.67	0.06
P_2O_5	0.16	0.23	0.26	n.d.	0.62	0.10
Sum	100.00	100.00	100.00	100.00	100.00	100.02
H_2O	0.4	0.3	n.d.	2	n.d.	0.0
CO_2	0.02	0.01	n.d.	n.d.	n.d.	n.d.
F	0.02	0.03	n.d.	n.d.	n.d.	0.002
Cl	0.02	0.03	n.d.	0.09	0.01	0.0005
S	0.08	n.d.	n.d.	0.19	n.d.	0.07

*n.d. = not determined.
†Columns show (1) average of 10 basalts from oceanic ridges; (2) submarine basalt, Eastern Rift Zone, Kilauea Volcano, Hawaii; (3) average high-alumina basalt of Oregon Plateau; (4) initial melt of Pacaya Volcano, Guatemala; (5) alkali basalt, Hualalai Volcano, Hawaii; (6) average *Apollo 12* lunar basalt.

Chemical composition. Significant differences exist in the composition of basalt which relate to different tectonic environments. The **table** lists typical compositions of basalts from oceanic ridges, island arcs, continental platforms, and oceanic islands and of alkalic basalts which have no particular tectonic association. In addition to the common elements, the table gives the concentrations of H_2O, CO_2, F, Cl, and S typical of the various basalts. It is common practice to reduce chemical analyses of igneous rock to a volatile-free basis for comparison purposes, and this practice has been followed in the table.

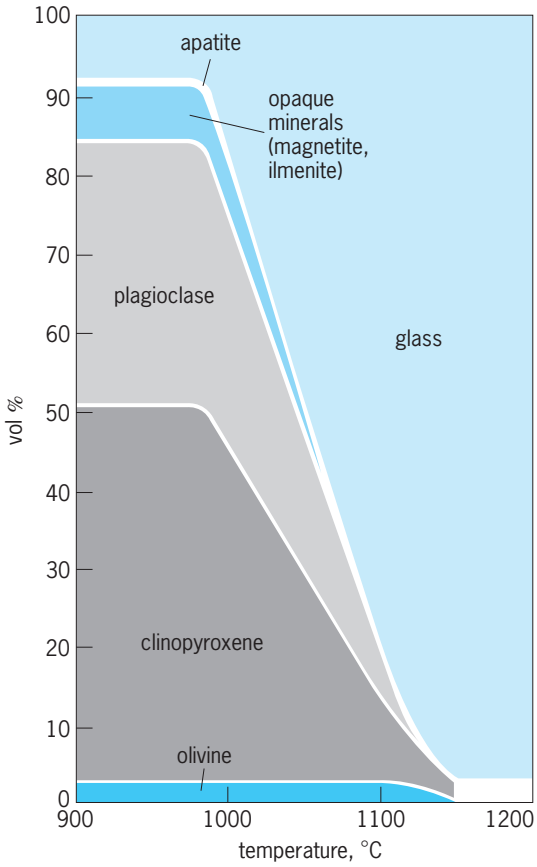

Fig. 3. Growth of crystals as a function of temperature in basalt in Hawaiian lava. °F = (°C × 1.8) + 32. (*After D. L. Peck, T. L. Wright, and J. G. Moore, Crystallization of tholeiitic basalt in Alae lava lake, Hawaii, Bull. Volcanol., 29:629–656, 1966*)

Chemical analyses of basalts are now used instead of, or together with, the textural criteria as a basis of classification. Geologists customarily recast the chemical analysis into a set of ideal minerals according to a set of rules. The result is called the norm of the rock. The principal normative minerals of basalts are the feldspars; diopside, $Ca(MgFe)Si_2O_6$; olivine, $(MgFe)_2SiO_4$; hypersthene, $(MgFe)SiO_3$; nepheline, $NaAlSiO_4$; and quartz, SiO_2. Because quartz reacts with olivine to form hypersthene and with nepheline to form feldspar (albite), basalts can be classified in order of increasing SiO_2 saturation by the presence of normative nepheline (lowest SiO_2), olivine, or quartz. Diagrams illustrating these relations and the particular names associated with individual combinations are shown in **Fig. 4**. In general, there is rather close correspondence between the normative minerals and the observed minerals in basaltic rocks.

Lunar basalts. Lunar basalts belong exclusively to the subalkaline class. They are similar to subalkaline basalts on Earth but have metallic iron rather than ferric-iron-bearing magnetite. Many lunar basalts are exceptionally rich in iron and titanium. They also differ compositionally from terrestrial basalts in being poor in alkali elements and siderophile (metal-seeking) elements such as nickel. *See* MOON.

Meteoritic basalts. Some meteorites are basaltic rocks. They differ significantly from lunar basalts and appear to have originated elsewhere in the solar system at a time close to the initial condensation of the solar nebula. Meteoritic basaltic rocks are called eucrites and howardites. *See* METEORITE.

Tectonic environment and origin. Basalts occur in four major tectonic environments: ridges in the sea floor, islands in ocean basins, island arcs and mountainous continental margins, and interiors of continents. The principal environment is the deep sea floor, where the rate of production of basalt is about 1.5×10^{16} g/year. Exploration of the Mid-Atlantic Ridge revealed that the ridge trough contains longitudinal depressions and mounds. It has been suggested that the floor of the ridge crest cracks open and is subsequently filled with a mound of pillow basalt on a time scale of 10,000–20,000 years. Basalts of the oceanic ridges typically are subalkaline in composition. Their chemical composition, tectonic environment, and the melting behavior of mantle rocks as a function of pressure suggest that ridge basalts evolve from primary melts that separate from partially melted mantle rocks at depths of a few tens of kilometers beneath the ridge (1 km = 0.6 mi).

Basaltic islands in the ocean basins contain diverse basalts. Major oceanic centers of volcanism (hot spots) such as Hawaii, Galápagos, and Réunion generally extrude subalkaline basalts. Although less abundant than subalkaline basalts, alkaline basalts are common on oceanic island volcanoes, and form mainly during the later stages of volcano development when rates of extrusion are less.

Island arc and continental margin basalts are predominantly subalkaline, but alkaline basalts are common away from the main volcanic front toward the continental interior. Island arc basalts differ compositionally from those at ocean ridges and islands. The differences, such as a higher concentration of alumina and more abundant feldspar phenocrysts, probably reflect the subduction and partial melting of siliceous and hydrous rocks mixed with altered oceanic crust and relatively little modified mantle rock. Water dissolved in magma modifies its crystallization behavior and stabilizes silica-poor, early-crystallizing minerals such as olivine and hornblende. *See* CONTINENTAL MARGIN; MARINE GEOLOGY; OCEANIC ISLANDS; PLATE TECTONICS.

Basalts in continental interiors fall into two classes: (1) major plateaus of lava flows extruded from giant fissures (Columbia River, Deccan Traps, Siberian Traps, Paranas Basin, Karoo, Tasmania), and (2) volcanoes (mostly small) of alkaline basalt commonly containing nodules of olivine-rich rocks apparently derived from the mantle.

The basalts of the major plateaus vary in chemical composition but are, with rare exception, subalkaline. Some of the plateau basalts are strongly similar to basalts from oceanic ridges, but are slightly richer in some of the elements normally found in high concentration in the granitic continental crust. Other plateau basalts are similar to the subalkaline basalts of oceanic islands. Still others are rather unique (Tasmania). The plateau basalts are unquestionably derived from the mantle, because they are erupted at too high a temperature to have been generated within the crust. The tectonic history and compositional aspects of plateau basalt provinces suggest that the provinces form in association with upwellings

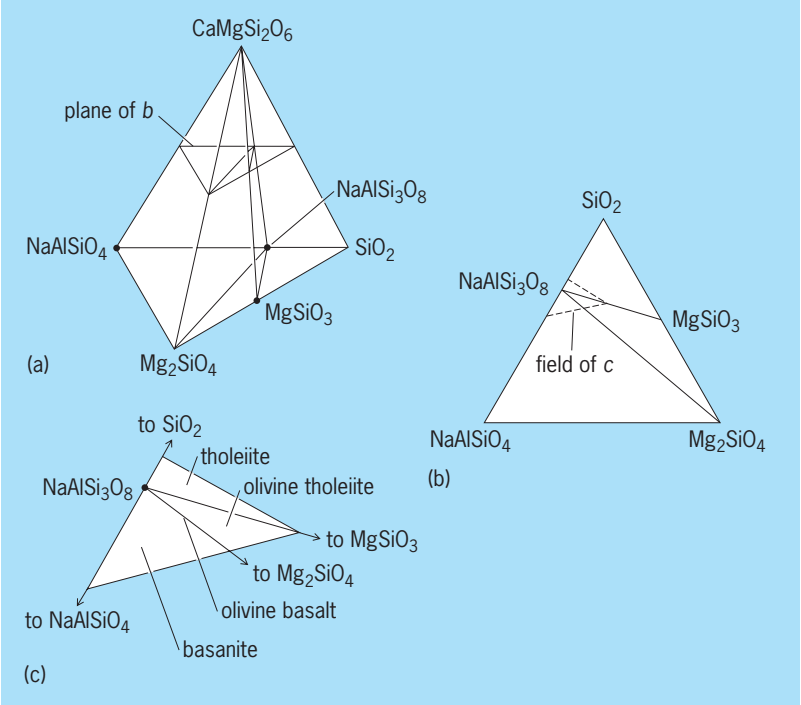

Fig. 4. Diagrams illustrating the compositions of basalts in terms of normative mineral formulas. Customarily, $CaAl_2Si_2O_8$ (anorthite) is grouped together with $NaAlSi_3O_8$ (albite), and Fe_2SiO_4, $FeSiO_3$, and $CaFeSi_2O_6$ are grouped with the corresponding Mg formulas. (*a*) The position of the basalt plane in a tetrahedron with $CaMgSi_2O_6$ (representing calcium-rich pyroxene) at the apex and plagioclase (represented by $NaAlSi_3O_8$) and other minerals indicates SiO_2 saturation on the base. (*b*) Diagram shows that natural basalts occupy only a small portion of the basalt plane. (*c*) Basalt plane from *b*.

(plumes) of relatively pristine mantle rock from great depths.

Chemical and mineralogical variations. The variations which occur within individual basalt flows can mostly be explained by separation of crystals and their original associated melts. In some cases certain elements have been removed, gained, or redistributed by oxidizing and acid gases which permeate the cooling body of rock from surrounding air and the variably wet and vegetated ground.

In general, as the space and time frame of compared basalts is enlarged to include flows from other vents and eruptive episodes, explanation of chemical and mineralogical differences becomes more difficult. Two principal avenues of interpretation overlap: The first uses the crystallization process evident in individual lava flows broadened to include the effects of increasing pressure on the crystallizing (or melting) minerals as well as the complex flow patterns within large bodies of magma. The second uses a variable bulk-chemical composition of the parental material (heterogeneity of the mantle).

Because the dynamical and crystallization behaviors of basaltic magmas are incompletely predictable, only those compositional attributes insensitive to the melting and crystallization process can be taken as evidence for variation in the bulk composition of source. For the most part, these attributes have been limited to ratios of isotopes. As long as the bulk chemical compositions of sources are identical, it is expected that melting will yield isotopically identical

primary basaltic liquids. Isotopically distinct liquids could form if there are major compositional differences or ages of source rocks.

Large ratios of $^3He/^4He$ in gases emitted near Yellowstone and elsewhere (larger than for ocean ridge basalts) suggest derivation of associated basaltic and other magmas from undifferentiated mantle source rocks. This is because helium is lost from the Earth's atmosphere and consequently is not available to be derived from an average collection of terrestrial source rocks. In such cases it is necessary to place some reliance on meteorites and other solar system data and even nucleosynthetic theory in order to justify an estimate of an initial undifferentiated composition. Therefore basalts eventually play an important role in concepts of the evolution of the solar system and its planets. *See* MAGMA. Alfred T. Anderson

Bibliography. P. Floyd (ed.), *Oceanic Basalts*, 1991; H. H. Hess (ed.), *Basalts*, 2 vols., 1968; A. R. McBirney, *Igneous Petrology*, 2d ed., 1992; H. Williams and A. R. McBirney, *Volcanology*, 1979, reprint 1982.

Base (chemistry)

In the Brönsted-Lowry classification, any chemical species, ionic or molecular, capable of accepting or receiving a proton (hydrogen ion) from another substance. The other substance acts as an acid in giving up the proton. A substance may act as a base, then, only in the presence of an acid. The greater the tendency to accept a proton, the stronger the base. The hydroxyl ion acts as a strong base. Substances that ionize in aqueous solutions to produce the hydroxyl ion (OH), such as potassium hydroxide (KOH) and barium hydroxide [$Ba(OH)_2$], are also conventionally called bases. Water (H_2O) acts as a weaker base in reaction (1), where H_3O^+ is the hydronium ion.

$$H_2O + H^+ \rightleftharpoons H_3O^+ \qquad (1)$$

Anions of weak acids, such as acetic and formic, act as bases in reacting with solvent water to form the molecular acid and hydroxyl ion, for example, the acetate ion (CH_3COO^-), as in reaction (2).

$$CH_3COO^- + H_2O \longrightarrow CH_3COOH + OH^- \qquad (2)$$

Ammonia (NH_3) and amines react similarly in aqueous solutions, as in reaction (3), where NH_4^+ is the

$$NH_3 + H_2O \longrightarrow NH_4^+ + OH^- \qquad (3)$$

ammonium ion. In these examples, the acetate ion and acetic acid (CH_3COOH) and NH_3 and the ammonium ion are conjugate base-acid pairs. The basicity constant, K_b, is the equilibrium constant for the proton transfer reaction, and it is a quantitative measure of base strength. For reaction (2), the basicity constant is given by expression (4), and its value at

$$K_b = \frac{[CH_3COOH][OH^-]}{[CH_3COO^-]} \qquad (4)$$

298 K is 5.7×10^{-9}. For formate ion ($HCOO^-$), the corresponding K_b is 5.6×10^{-10}. Acetate is, therefore, the stronger base. While these concepts are most useful in aqueous solutions, they are equally valid in other solvent systems.

The Lewis classification involves the concept of a base as a substance that donates an electron pair to an acid acceptor. In the gas phase [reaction (5)], NH_3 acts as a base in contributing an electron

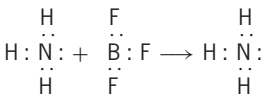

(5)

pair to the formation of a covalent bond with the boron trifluoride (BF_3) molecule. *See* ACID AND BASE. Francis J. Johnston

Bibliography. B. W. Jensen, *The Lewis Acid-Base Concepts: An Overview*, 1979; D. R. Lide (ed.), *CRC Handbook of Chemistry and Physics*, 1993; K. W. Whitten, K. D. Gailey, and R. E. Davis, *General Chemistry*, 4th ed., 1992.

Basidiomycota

A phylum in the kingdom fungi; commonly known as basidiomycetes. Basidiomycetes traditionally included four artificial classes: Hymenomycetes, Gasteromycetes, Urediniomycetes, and Ustilaginomycetes. They are mostly filamentous fungi characterized by the production of basidia. These are microscopic, often club-shaped end cells in which nuclear fusion and meiosis usually take place prior to the maturation of external, typically haploid basidiospores, which are then disseminated. Common basidiomycetes are the rusts and smuts, which cause severe plant diseases, mushrooms (edible and poisonous), boletes, puffballs, stinkhorns, chanterelles, false truffles, jelly fungi, bird's-nest fungi, and conk or bracket fungi. Basidiomycetes are the most important decayers of wood, living or dead, in forests or buildings, causing either brown rot (for example, dry rot) or white rot. Many, especially mushrooms and boletes, are the primary fungal partners in symbiotic ectomycorrhizal associations with tree roots. Plant litter and soil are other major habitats. A few basidiomycetes are cultivated for food. Some are luminescent, hallucinogenic, lichenized, nematophagous, or aquatic. Some are cultivated by ants or termites, or are symbiotic with captured scale insects. Some can convert to a yeast (or single-cell) phase, one of which causes cryptococcosis in humans and animals. *See* MUSHROOM; MYCORRHIZAE; WOOD DEGRADATION.

In a typical life cycle, haploid basidiospores are forcibly discharged from projections on the basidia into the air, disperse, and germinate on a suitable substrate to form filaments (hyphae), which make up a vegetative stage (mycelium). This mycelium must anastomose with a compatible mycelium. Mating is controlled by one or two allelic genes resulting

in two (bipolar) or four (tetrapolar) morphologically indistinguishable mating types. The mycelium resulting from mating has two compatible nuclei in each cell, and is said to be dikaryotic. This nuclear phase is unique to basidiomycetes and can last for years. Nuclear fusion eventually takes place in a basidium, either on the mycelium or on or in a fruit body (for example, mushroom or puffball), pustule, or sorus. Specialized sexual structures (pycnia, spermatia, protoaecia) are formed only by rusts. Many other basidiomycetes form characteristic short bridges (clamp connections) which act to bypass the wall (septum) between cells for one of the paired nuclei during mitosis. Many variations occur. Basidiospores may be dikaryotic when discharged. They may be passively released, formed in chains; dispersed dry or in slime by vectors such as insects or other agents such as water, either singly or in special organs; dispersed by special organs; or lacking entirely. The haploid phase may be yeast-like or lacking. The dikaryotic phase may be yeast-like. Nuclear fusion may occur in structures other than the basidia. The fungus may reproduce asexually by conidia or, if a rust, by aeciospores and urediospores; conidia may anastomose with mycelia to form dikaryons. *See* FUNGI; GASTEROMYCETES; HYMENOMYCETES; UREDINOMYCETES (RUST); USTILAGINOMYCETES (SMUT). Scott A. Redhead

Bibliography. G. C. Ainsworth, F. K. Sparrow, and A. S. Sussman, *The Fungi: An Advanced Treatise*, vol. 4B, 1973; D. L. Hawksworth et al., *Ainsworth & Bisby's Dictionary of the Fungi*, 8th ed., 1996; D. S. Hibbett et al., *Proc. Nat. Acad. Sci.*, 94:12002-12006, 1997; E. C. Swann and J. W. Taylor, *Can. J. Bot.*, Suppl. 1:S862-S868, 1995.

Basil

An annual herb (*Ocimum basilicum*) of the mint family (Labiatae), grown from seed for its highly scented leaves. It is sold fresh or dried, and its essential oil is used in pharmaceuticals and flavoring. The genus *Ocimum* has approximately 150 species, many of which are also grown as "basil." *See* LAMIALES.

Growth habit varies with type, but it is usually upright and branching from a square central stem up to 24 in. (61 cm) in height. Leaf shape and odor vary considerably from type to type, with "sweet basil," the most popular, having broad, flat, shiny green leaves to 3 in. (7.5 cm) long. Wild forms of basil are found in South America, Africa, India, and Europe. Cultivated basil differs little from its wild forms, although breeding and selection work have been performed on some of the varieties grown for essential oil production. Some of the more popular cultivated types are: sweet basil (used fresh or dehydrated), lemon-scented basil, opal basil (a purple-leaved type), and large- or lettuce-leaf basil. Many unusual types can be obtained which have a scent similar to camphor, lemon, licorice, or nutmeg.

Basil is grown commercially in many areas of the world. In the United States, the largest producer of dried leaf products is California, where the drier climate aids dehydration. Basil grown for oil production is centered in the eastern states, though some is also produced in Washington and Oregon. A considerable quantity of the basil consumed in the United States is imported from Europe.

All basil types have similar cultural requirements, which include high light intensity, 20–60% relative humidity, and a high level of soil fertility. Because of the large leaf suface, basil also requires a good supply of water. Deviation from these requirements results in either disease or low essential oil content. Since all propagation is by seed, basil is replanted annually. Large commercial farms utilize direct seeding into the field, planted in 3–6-in. (7.5–15-cm) rows, 24–42 in. (61–122 cm) apart.

Approximately 60 days after seeding, harvesting is begun, and may be repeated up to five times during the growing season. As is typical of nearly all herbs, regrowth occurs after trimming the plant. Harvesting is accomplished mechanically, utilizing machinery similar to that used in the mint industry. After transport from the field, either basil is dehydrated by forced air dryers or the oil is extracted with steam or solvent. Leaves are stripped from the stem after dehydration and then processed to obtain a uniform product.

Traditionally associated with use in Italian and southern European cuisines, basil is most often used in conjunction with other herbs for flavoring sauces, and is the primary ingredient in Italian pesto sauce. Basil oil is used in perfumery and as a source of eugenol. *See* SPICE AND FLAVORING. Seth Kirby

Bibliography. S. Arctander, *Perfume and Flavor Materials of Natural Origin*, 1960; L. H. Bailey, *Manual of Cultivated Plants*, rev. ed., 1975; J. C. Uphof, *Dictionary of Economic Plants*, 1968.

Basin

A low-lying area which is wholly or largely surrounded by higher land. An example is the Hudson Bay in northeastern Canada, which was formed by depression beneath the center of a continental ice sheet 18,000 years ago. Another example, the Qattara depression, is 150 mi (240 km) long and the largest of several wind-excavated basins of northern Egypt. Depressions in the ocean floor are also basins, as the Canary Basin, west of northern Africa, or the Argentine Basin, east of Argentina. These basins occur in regions where cold, dense oceanic crust lies between the topographically elevated ocean ridges and the continental margins. *See* CONTINENTAL MARGIN; MARINE GEOLOGY.

Drainage basins. A drainage basin is the entire area drained by a river and its tributaries. Thus, the

Mississippi Basin occupies most of the United States between the Rocky Mountains and the Appalachians. Interior drainage basins consist of depressions that drain entirely inward, without outlet to the sea. Examples may be quite small, such as the Salton Sea of southern California or the Dead Sea of central Asia. One of the most remarkable examples of an interior drainage basin is the Chad Basin in northern Africa, the center of which is occupied by Lake Chad. The fresh waters of the lake drain underground to feed oases in the lowlands 450 mi (720 km) to the northeast.

Geologic basins. In the geologic sense, a basin is an area in which the continental crust has subsided and the depression has been filled with sediments. Such basins were interior drainage basins at the time of sediment deposition but need not be so today. As these basins subside, the layers of sediment are tilted toward the axis of maximum subsidence. Consequently, when the sedimentary layers are observed in cross section, their geometry is a record of the subsidence of the basin through time and contains clues about the origin of the basin.

Characteristics. Many geologic basins are nearly circular in plan view and symmetrical in cross section, such as the Michigan Basin (see **illus.**) and the Paris Basin. The maximum subsidence in these basins occurred at a central point, producing a saucer-shaped depression. Other basins are elongate or trough-shaped in plan view, with lengths in some cases reaching thousands of miles. Some elongate basins have symmetrical cross sections, such as the rift basins of eastern Africa and the Newark Basins of eastern North America, while others are asymmetric in cross section and typically become deeper or thicker toward mountain ranges against which they abruptly terminate. Examples of elongate asymmetric basins are the Denver Basin, the Appalachian Basin (see illus.), and the Ganges Basin.

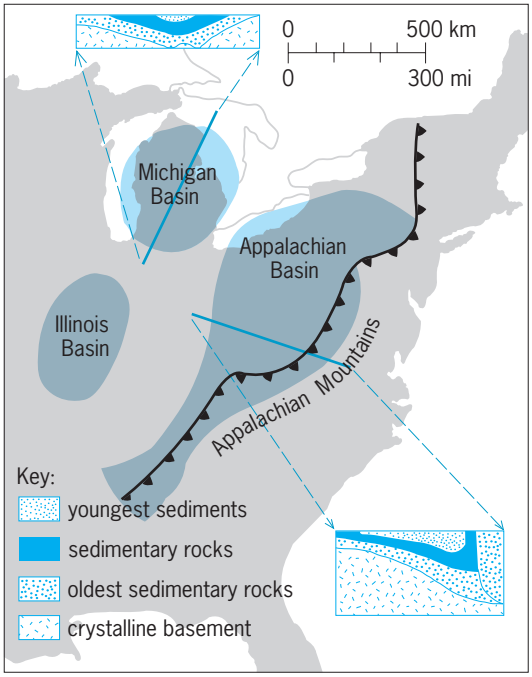

Paleozoic (approximately 250–570 million years old) basins of the eastern United States. The cross sections (insets) show the thickness and general pattern of deposition of sediments in the Michigan Basin and Appalachian Basin. The arrows show where that edge of the cross section is located on the plan-view map.

Origin. The different types of basins form in response to significantly different geologic processes. Elongate rift basins form where the Earth's lithosphere has been stretched in response to tensional forces, commonly, but not always, associated with the rifting and breakup of continents. Such basins are bounded by steeply dipping faults that were active as the basins subsided and filled with sediment. Many elongate asymmetrical basins form where the basin floor has been bent downward under the weight of slices of the Earth's crust that have been faulted, one on top of the other, along the edges of mountain ranges. Such basins, called foreland basins, commonly form where lithospheric plates have collided and the crust is under compression. *See* FAULT AND FAULT STRUCTURES; PLATE TECTONICS.

The most enigmatic basins are those that are symmetrical and have diameters approximately equal in all directions, such as the Michigan Basin. In contrast to the other types of basins, no geologic evidence of a mechanism has been found adjacent to or beneath these basins. They are thought to be the product of processes operating at a deep level within the Earth, but the exact nature of those processes is poorly understood.

The origin of geologic basins is a topic of continuing interest in both applied and basic geological studies. They contain most of the world's hydrocarbon reserves, and they are regarded as some of the best natural laboratories in which to understand the thermal and mechanical processes that operate deep in the interior of the Earth and that shape the Earth's surface. Gerard Bond; Michelle Kominz

Bibliography. P. A. Allen and J. R. Allen, *Basin Analysis: Principles and Applications*, 2d ed., 2005; M. Leeder, *Sedimentology and Sedimentary Basins: From Turbulence to Tectonics*, 1999; A. D. Miall, *Principles of Sedimentary Basin Analysis*, 2d ed., 1990.

Basommatophora

A superorder of the molluscan subclass Pulmonata containing about 2500 species that are grouped into 11–15 families. Only a few members of the family Ellobiidae are terrestrial, with the other species today being tidal to supratidal or estuarine in habitat. The families Siphonariidae and Trimusculidae are marine limpets with caplike shells. The families Otinidae and Amphibolidae also are marine taxa. The remaining families inhabit a great variety of fresh-water

situations and are quite varied in shell structure and shape. Because of this great variation in habitat and form, it is difficult to find structures that are common to all taxa and thus diagnostic of the group, which indeed may not have a common origin. The most easily observable characters are having the eyespots normally located at the base of two slender tentacles that are neither contractile nor retractile, and usually having two external genital orifices. Most features of the anatomy are shared with other pulmonate superorders, or are specializations related to major changes in habitat or body form.

Some of the few marine taxa are phylogenetic relicts (Otinidae and Amphibolidae) that show unusual features such as retaining the operculum characteristic of prosobranch snails, are highly specialized in habitat, and are restricted in distribution. Only the Siphonariidae and Trimusculidae, intertidal limpets, are widely distributed and form a significant part of the littoral fauna. The Ellobiidae are scattered in distribution, occasionally abundant near the high-tide mark, as in *Melampus* in the Atlantic coast saltmarshes of North America. One genus, *Carychium*, is truly terrestrial and common in very wet areas along streams in Eurasia and North America.

The remaining families of the Basommatophora are fresh-water dwellers, usually highly visible, and ecologically significant, both in terms of serving as a food resource for vertebrates and in their browsing on the shallow-water encrusting organisms (algae, fungi, protozoa, and so forth) that grow on rocks and plants. Only a very few species are of medical significance, but these are part of very serious problems. The snails serve as intermediate hosts for trematode parasites of both humans and domestic animals. Some planorbids transmit schistosomiasis in Africa, the West Indies, and South America, while lymnaeids are involved in both sheep and cattle liver fluke life cycles. In North America, some bird schistosomes attempt to penetrate human skin, with only partial success, and then die. The rash produced by their decaying bodies produces the complaint known as swimmer's itch. The annual cost of diseases transmitted by these few snails is very great. Killing all the snails in a body of water has not proved effective, as many of the snails survive in nooks and crannies, and new specimens are carried in by birds and insects from neighboring bodies of water.

Indeed, one of the most important aspects of basommatophoran snails is the extent to which their biology and evolution have been shaped by the fact that most bodies of fresh water are small and isolated, and exist for only a few thousand years. When new ponds or lakes appear, they must be colonized by dispersal across land barriers. Basommatophorans appear in the geological record at the Jurassic-Cretaceous boundary, shortly after the origin of birds. Frequently pond snails are carried on the feet or feathers of birds from one body of water to another. So that one such successful colonization is enough to populate a new body of water, most species are capable of self-fertilization, although cross-fertilization is the normal mode of reproduction. The eggs are laid in gelatinous masses on rocks or plants. Generation time usually is very short, and enormous numbers of snails can build up very quickly during the favorable weather seasons. As a result of this ability to start a colony from one individual and to multiply rapidly, basommatophoran snails show a great deal of minor variation among populations. This makes identification of species very difficult and has led to much confusion as to the proper names for a species and in determining exactly what is a species in these families.

By far the vast majority of fresh-water basommatophorans live in water of less than 12 ft (4 m) depth, and only rarely have species been found living at relatively great depths. These few species must depend upon cutaneous respiration or an air bubble used as a physical gill for oxygen exchange, while those living in shallow water normally come to the surface at regular intervals to breathe.

Human activities, particularly polluting waters with industrial chemicals, excess fertilizers, and silt, have had profound effects upon the fresh-water snail faunas, greatly reducing both the diversity of species and numbers of individuals present.

The major families of fresh-water basommatophorans are readily separated. The South American Chilinidae have peculiarities in the nervous system and incomplete separation of the male and female ducts in the pallial region. The probably related Latiidae from New Zealand and Acroloxidae from Eurasia and the Rocky Mountain area of North America represent specializations toward limpet form. The nearly worldwide Lymnaeidae have completely separated pallial reproductive ducts, generally a high, spiral shell, rather flat tentacles, and several other minor anatomical differences. They are common in temporary ponds, lakes, and sluggish streams. The sinistrally coiled Physidae have peculiar radular teeth and a reduction in size of the jaw. They are common in temporary ponds and other isolated bodies of water in the Northern Hemisphere, and are thriving where introduced into Southern Hemisphere localities.

The Planorbidae and Ancyliade typically are readily separable by the limpet-shaped shell of the latter, but discoveries of intermediate taxa and studies of anatomy have demonstrated that the limpet shape was evolved several times, and some authors propose uniting them into one family. The Planorbidae are the most diverse in size and form of all basommatophoran groups. Both groups live in many types of fresh-water habitats. *See* PULMONATA.

G. Alan Solem

Bibliography. V. Fretter and J. Peake (eds.), *Pulmonates*, vol. 2A, 1978; L. Hyman, *The Invertebrates*, vol. 6: *Mollusca I*, 1967; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Basses

A common name applying to a large group of fishes, mostly in the order Perciformes, generally having the following characteristics: body about 2.5 or 3 times longer than its depth; two dorsal fins, the first composed of stiff spines; pectoral fins on the sides of the body just behind the head; pelvic fins thoracic in position, with one spine and five rays; ctenoid scales; a relatively large mouth; and a predatory lifestyle. Because "bass" is not a taxonomic unit, however, the term cannot be precisely described or defined. In the western North Atlantic, the eastern North Pacific, and freshwaters of Central and North America alone, the word "bass" appears in the vernacular of 53 species and six hybrids in 17 genera and six families. A generally accepted limited taxonomy of the basses follows. (A indicates a North Atlantic range; P, North Pacific; F, fresh waters of Central and North America.) Fishes called "basslet" (primarily family Grammatidae) are not included.

Order Perciformes
- Suborder Percoidei
  - Family Moronidae—temperate basses
    - *Morone americana*—white perch (A-F)
    - *Morone chrysops*—white bass (F)
    - *Morone mississippiensis*—yellow bass (F)
    - *Morone saxatilis*—striped bass (A-F)
    - ♂ *Morone americana* × ♀ *M. saxatilis* hybrid—Virginia bass (F)
    - ♀ *Morone americana* × ♂ *M. saxatilis* hybrid—Maryland bass
    - ♀ *Morone chrysops* × ♂ *M. saxatilis* hybrid—sunshine bass
    - ♂ *Morone chrysops* × ♀ *M. saxatilis* hybrid—palmetto bass
    - ♂ *Morone mississippiensis* × ♀ *M. saxatilis* hybrid—paradise bass
  - Family Acropomatidae—lanternbellies
    - *Synagrops bellus*—blackmouth bass (A)
    - *Synagrops spinosus*—keelcheek bass (A)
    - *Synagrops trispinosus*—threespine bass (A)
  - Family Polyprionidae—wreckfishes
    - *Stereolepis gigas*—giant sea bass (P)
  - Family Symphysanodontidae—slopefishes
    - *Symphysanodon berryi*—slope bass (A)
  - Family Serranidae—sea basses and groupers
    - *Anthias nicholii*—yellowfin bass (A)
    - *Anthias tenius*—threadnose bass (A)
    - *Anthias woodsi*—swallowtail bass (A)
    - *Bathyanthias mexacana*—yellowtail bass (A)
    - *Bullisichthys caribbaeus*—pugnose bass (A)
    - *Centropristis fuscula*—twospot sea bass (A)
    - *Centropristis ocyurus*—bank sea bass (A)
    - *Centropristis philadelphica*—rock sea bass (A)
    - *Centropristis striatus*—black sea bass (A)
    - *Dermatolepis dermatolepis*—leather bass (P)
    - *Hemanthias aureorubens*—streamer bass (A)
    - *Hemanthias leptus*—longtail bass (A)
    - *Hemanthias peruanus*—splittail bass (P)
    - *Hemanthias signifier*—hookthroat bass (P)
    - *Paralabrax auroguttatus*—goldspotted sand bass (A)
    - *Paralabrax clathratus*—kelp bass (P)
    - *Paralabrax loro*—parrot sand bass (P)
    - *Paralabrax maculatofasciatus*—spotted sand bass (P)
    - *Paralabrax nebulifer* barred sea bass (P)
    - *Paralabrax atrimanus*—bank bass (A)
    - *Plectranthias garrupellus*—apricot bass (A)
    - *Pronotogrammus eos*—bigeye bass (P)
    - *Pronotogrammus martinicensis*—roughtongue bass (A)
    - *Pronotogrammus multifasciatus*—threadfin bass (P)
    - *Pronotogrammus gregoryi*—reef bass (A)
    - *Pronotogrammus thaumasium*—Pacific reef bass (P)
    - *Serraniculus pumilo*—pygmy sea bass (A)
    - *Serraniculus annularis*—orangeback bass (A)
    - *Serraniculus atrobranchus*—blackear bass (A)
    - *Serraniculus baldwini*—lantern bass (A)
    - *Serraniculus chionaraia* snow bass (A)
    - *Serraniculus flaviventris*—twinspot bass (A)
    - *Serraniculus luciopercanus*—crosshatch bass (A)
    - *Serraniculus notospilus*—saddle bass (A)
    - *Serraniculus tigrinus*—harlequin bass (A)
    - *Serraniculus tortugarum*—chalk bass
  - Family Centrarchidae—sunfishes
    - *Ambloplites ariommus*—shadow bass (F)
    - *Ambloplites cavifrons*—Roanoke bass (F)
    - *Ambloplites constellatus*—Ozark bass (F)
    - *Ambloplites rupestris*—rock bass (F)
    - *Micropterus cataractae*—shoal bass (F)
    - *Micropterus coosae*—redeye bass (F)
    - *Micropterus dolomieu*—smallmouth bass (F)
    - *Micropterus notius*—Suwannee bass (F)
    - *Micropterus punctulatus*—spotted bass (F)
    - *Micropterus salmoides*—largemouth bass (F)
    - *Micropterus treculii*—Guadalupe bass (F)
    - *Lepomis macrochirus* × *Micropterus salmoides* hybrid—blue bass (F)

Acropomatidae. Acropomatidae is a family of about 40 species but only three bear the name bass. Adults of these fishes attain lengths of 11 to 23 cm (4 to 9 in.) and they are bathypelagic, occurring at depths to at least 910 m (3000 ft), in the Eastern and Western Atlantic and elsewhere.

Polyprionidae. Of the nine species in Polyprionidae only one is a bass, the giant sea bass, with a length of 250 cm (8 ft 2 in.) and weighing 256 kg (564 lb). This giant fish is known in the eastern Pacific from coastal Mexico and California, and in the western North Pacific from Japan. Juveniles live in and around kelp beds, whereas adults live near kelp in deeper water over rocky substrates. The giant sea bass is

listed by the World Conservation Union as a critically endangered species.

Symphysanodontidae. The slope bass is one of six species in the family Symphysanodontidae. It is a small bass of about 13 cm (5 in.) that inhabits rocky substrates at depth of 220 to 467 m (720–1500 ft) in the Bahamas and off the coasts of Central America and Venezuela.

Centrarchidae. Centrarchidae is a fresh-water family indigenous to North America east of the Continental Divide, except for the Sacramento perch in California. Centrarchids are actually sunfishes; only the species of genera *Ambloplites* and *Micropterus* are called basses. *See* SUNFISHES.

Moronidae. Members of the family Moronidae (temperate basses) are fresh-water or euryhaline (tolerating a wide range of salinities) fishes that spawn in fresh water. The family consists of six species in two genera, *Morone* in North America and *Dicentrarchus* off Europe and North Africa. The family Moronidae has also been, and still is by some authors, placed in the Percichthyidae, a family of fresh-water fishes limited to southern South America and Australia. Temperate basses differ from other basses in the following combination of characters: they have a dorsal fin in two distinct parts, the first with 8–10 spines and the second with one spine and 9–12 soft rays; an opercle with one or two flat spines; seven branchiostegal rays; a lateral line that is complete and extending onto the caudal fin; and a series of black horizontal stripes on the sides of the body, except in *Morone americana*, the white perch.

The white bass and yellow bass are small species, whose adult sizes rarely exceed 1 and 3 kg (2 and 6 lb), respectively. Both are popular food and game fishes, occupying clear lakes and reservoirs, as well as backwaters of large rivers. Most famous of the temperate basses is the striped bass, *Morone saxatilis*. The names *Roccus saxatilis* and *Roccus striatus* appear in the literature but they are synonyms of *Morone saxatilis*.

The native range of the striped bass is coastal waters and fresh-water streams from the St. Lawrence River south to the St. Johns River, Florida, disjunctly in Gulf tributaries from west Florida to Lake Pontchartrain, Louisiana. The species was introduced to the west coast in 1879 and now also occurs from southern California to Washington and possibly Alaska. Adults migrate up rivers in search of suitable spawning habitats. Their fertilized eggs, when water hardened, are about 3.5 mm in diameter and slightly denser than water. The eggs require several days in sunlit water to develop fry. Currents of about 60 cm/s keep the eggs suspended or nearly so. Should the eggs settle on a muddy substrate they will die; therefore, long stretches of clear, free-flowing water and hard substrates over which the eggs tumble are the essentials of a suitable habitat. Striped bass have also been introduced on the Atlantic coast, where they attain weights exceeding 35 kg (78 lb) and are a very popular target of surf fishers.

Chalk bass (*Serranus tortugarum*). (*Photo © John E. Randall*)

Serranidae. Certain species of the large family Serranidae are collectively known as the sea basses. The family is characterized by three flat spines on the opercle; generally ctenoid scales, cycloid in some species; a complete lateral line, not extending onto the caudal fin; a dorsal fin that is usually continuous and may be notched, usually with 10 or fewer spines; three anal fin spines; pelvic fins, each with one spine and five rays; usually seven branchiostegals; and usually 24 vertebrae (see **illustration**). Serranids are hermaphroditic, usually developing first as females and then becoming males; however, some (such as genus *Serranus*) are functional hermaphrodites (that is, individuals are both sexes simultaneously). Serranids vary in adult length from only 3 cm to 300 cm (1 in. to 10 ft) and attain weights of 455 kg (1003 lb). Sea basses occur worldwide in tropical and temperate seas, typically on reefs and objects that provide shelter. Depending on the species, their food may be plankton, crustaceans, or fishes. Many members of the family are important food and game fishes. Worldwide their commercial importance is enormous.

Of 62 genera and 449 species in the family, it is difficult to determine on a worldwide basis the number of serranid "basses." In the geographical area described above, 37 of 138 species of serranids bear the name bass. Some other serranid vernaculars are grouper, basslet, hamlet, soapfish, and sand perch.

Herbert Boschung

Bibliography. G. R. Allen and D. R. Robertson, *Fishes of the Tropical Eastern Pacific*, University of Hawaii Press, Honolulu, 1994; J. Baillie and B. Groombridge (compilers and editors), *1996 IUCN Red List of Threatened Animals*, IUCN, Gland, Switzerland, 1996; W. N. Eschmeyer, E. S. Herald, and H. Hammann, *A Field Guide to Pacific Coast Fishes of North America*, Houghton Mifflin, Boston, 1983; G. D. Johnson, Percoidei: Development and relationships, pp. 464–498, in H. G. Moser et al. (eds.), *Ontogeny and Systematics of Fishes*, Amer. Fisheries Soc. Spec. Publ. 1, Bethesda, MD, 1984; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994; J. S. Nelson et al., *Common and Scientific Names of Fishes from the United States, Canada, and Mexico*, Amer. Fisheries Soc. Spec. Publ. 29, Bethesda, MD, 2004.

Basswood

A member of the linden family (Tiliaceae) in the order Malvales. One species is known as the American linden or basswood (*Tilia americana*). It is a timber tree of the northeastern quarter of the United States and the adjacent area of Canada and grows to a height of 120 ft (36 m). The leaves are heart-shaped, coarsely toothed, long, pointed, and alternate (see **illus.**). *See* MALVALES.

American basswood (*Tilia americana*).

All species of *Tilia* can be recognized by the peculiar winter buds, which have a large outer scale that gives the bud a humped, asymmetrical appearance, and by the small, spherical, nutlike fruits borne in clusters. *Tilia* is also an ornamental tree. *Tilia europea*, on lime tree of Europe, is often cultivated along the streets. The lindens are also important as bee trees. A species of the southern United States, *T. heterophylla*, known as the bee tree linden, has leaves that are white beneath, or brownish with a dense, hairy coat.

There are about 30 species in the temperate regions of the Northern Hemisphere, in North America south to Mexico, but none in the western part. In Asia lindens grow south to central China and southern Japan. The wood of basswood, also known as whitewood, is white and soft and is used for boxes, venetian blinds, millwork, furniture, and woodenware. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Batales

A small order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Dilleniidae of the class Magnoliopsida (dicotyledons). It consists of two small families and fewer than 20 species, of no economic significance. The plants are trees, shrubs, or subshrubs with simple, entire leaves and small, unisexual flowers. The perianth consists of small, separate or fused sepals with no petals. The stamens are four to numerous; the pistil has (1)2–many carpels, fused to form a plurilocular, superior ovary. The fruit is a dry, dehiscent nutlet or a drupe. The taxonomic disposition of the Batales has long been disputed, and it is often included in the Caryophyllidae. However, the presence of mustard oils as well as certain morphological features suggest that the order is best included near the Capparales in the Dilleniidae. *See* CAPPARALES; DILLENIIDAE; MAGNOLIOPHYTA; PLANT KINGDOM.

T. M. Barkley

Bathynellacea

An order of syncarid crustaceans found in subterranean waters in central Europe and England. The body is elongate and segmented. Organs and limbs are primitive. Each of the thoracic limbs has two epipodites which function as gills. The thorax and abdomen have a similar form (see **illus.**). These animals have no metamorphosis. *See* SYNCARIDA.

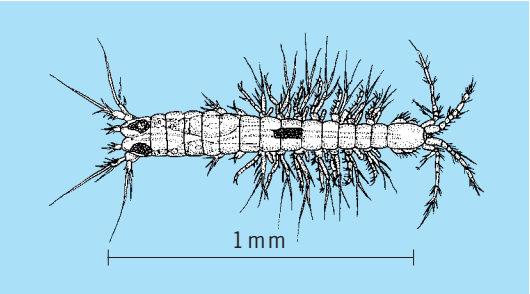

***Bathynella natans*, female.**

Hans Jakobi

Batoidea

One of two subdivisions of the subclass Elasmobranchii, consisting of four extant orders, 17 families, and about 456 species, known collectively as rays or batoids. The other subdivision, Selachii, includes the sharks. Some authors group the modern sharks and several families of fossil sharks plus the modern rays in a superorder, Euselachii, supposedly a monophyletic unit. Others group all Recent sharks and rays, each in a single subdivision, Selachii and Batoidei, as presented herein.

Class Chondrichthyes
  Subclass Elasmobranchii
    Subdivision Selachii
      Batoidea
        Order Torpediformes (electric rays)
          Pristiformes (sawfishes)
          Rajiformes (skates)
          Myliobatiformes (stingrays, round rays, butterfly rays, and eagle rays)

See separate articles on each order.

Most batoids are easily recognizable by the flattened head and trunk, with the tail more or less distinct from the body, and eyes and large spiracles on the dorsal surface of the head (see **illustration**). They differ from sharks in having ventral gill slits, the edge of the pectoral fins attached to the side of the head anterior to the gills, the eyeball attached to the upper margin of the orbit, and by lacking an

anal fin (most sharks have one). Also, spiracular respiration is much more important to batoids than it is to sharks. A method of reproduction common to all batoids, internal fertilization is accomplished by means of intromittent organs called claspers, which are modified medial lobes of the pelvic fins. Most batoids are ovoviviparous, retaining the eggs in the oviducts until birth, but skates are oviparous, depositing eggs that are protected by horny cases. The early embryos are more sharklike than batoidlike, as the body is slender and the pectoral fins are posterior to the gill openings. Also in the early embryonic stage an anal fin is present as in sharks, but is soon lost. By full term the pectoral fins have migrated forward as well as rearward to form the characteristic disk. Batoids vary in size from less than 0.3 m (1 ft) in length as adults to a width of over 6 m (20 ft) in the giant mantas.

By wiggling their disk, the bottom-dwelling batoids will throw sand over their back, leaving only the eyes and spiracles exposed. All are carnivorous, feeding on a wide variety of marine worms, mollusks, crustaceans, and other invertebrates, as well as small fishes. They are represented in tropical to arctic latitudes of the Atlantic, Pacific, and Indian oceans of both hemispheres. Some are limited to fresh water (such as the subfamily Potamotrygoninae of South America) and a few marine species, especially stingrays and sawfishes, penetrate estuaries of rivers or even live for extended periods far upstream, but the vast majority are strictly marine species. Electric rays and skates may live at depths of 3000 and 9000 feet, respectively. However, members of both groups are more numerous in kind and numbers at depths between 120 to 600 feet. Typically the myliobatids are abundant in relatively shallow waters around islands. Although batoids are abundant in many seas, they are of minor commercial importance.

Where known, the body of extinct skates and rays appears to have resembled that of modern types in its depressed shape, expanded pectoral fins, and slender tail. In general, however, remains consist of teeth only. Many Paleozoic sharks appear to have been bottom dwelling, feeding on shelled invertebrates; these sharks became extinct, however, and the first, relatively rare, batoids are found in the Late Jurassic. With the exception of the eagle rays and torpedoes, all major groups now extant were present in the Upper Cretaceous and had assumed the prominent position that persist today. *See* ELASMOBRANCHII; SELACHII.

Reeve M. Bailey; Herbert Boschung

Bibliography. H. B. Bigelow and W. C. Schroeder, Sawfishes, guitarfishes, skates and rays, in *Fishes of the Western North Atlantic*, Mem. Sears Found. Mar. Res. Mem. 1(pt. 2), pp. 1–514, 1953; L. J. V. Compagno, Interrelationships of living elasmobranchs, pp. 15–61, in P. H. Greenwood et al. (eds.), *Interrelationships of Fishes*, Zool. J. Linn. Soc. 53, Suppl. 1, Academic Press, New York, 1973; L. J. V. Compagno, Checklist of living elasmobranchs, pp. 471–498, in W. C. Hamlett (ed.), *Sharks, Skates and Rays: The Biology of Elasmobranch Fishes*, John Hopkins University Press, 1999; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 2006.

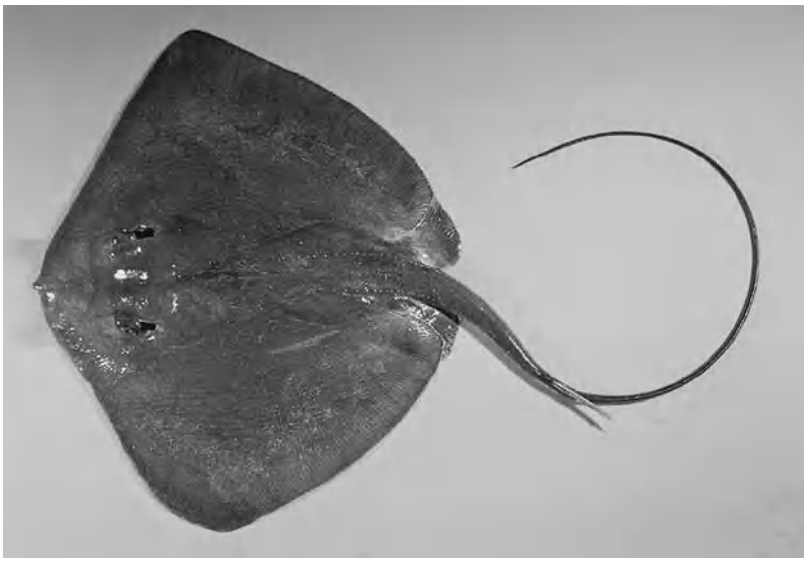

Brown stingray (*Dasyatis latus*). (*Photo © John E. Randall*)

Batrachoidiformes

The toadfishes, an order of actinopterygian (ray-finned) fishes consisting of a single family, Batrachoididae. Batrachoididae comprises three subfamilies, 22 genera, and 73 species. Also known as the Haplodoci, the toadfishes are most closely related to the Lophiiformes, and the two groups collectively are known as the Pediculati. These are rather robust fishes, tapering posteriorly, with a broad flattened head, often with barbels or a fleshy flap; undivided posttemporal bone; dorsal, skyward-looking eyes; a large mouth and powerful jaws equipped with blunt teeth (some species have canine teeth); only three gill arches and reduced gill openings; pelvic fins in the throat area (jugular), with one spine and two or three soft rays; fanlike pectoral fins; a short, spinous dorsal fin consisting of two or three strong sharp spines; long soft dorsal and anal fins; a usually scaleless body covered with mucus; one to several lateral lines on head and body; and mostly drab brown color variously marked with dark saddles and blotches.

Toadfishes occur in the Atlantic, Pacific, and Indian oceans, primarily as benthic coastal inhabitants; however, a few species enter brackish water, and several are limited to freshwater. These are sluggish fishes that ambush mollusks and crustacean prey, and they are known to bite when handled.

Members of the subfamily Batrachoidinae occur worldwide but have the greatest diversity in the New World, in the Western Atlantic and Eastern Pacific. They have three solid dorsal spines, an opercular spine, and no venom glands; a subopercular spine is present; canine teeth are absent; the body is scaleless or not; and there are no photophores. Most species occur in mud or sand substrates, but one

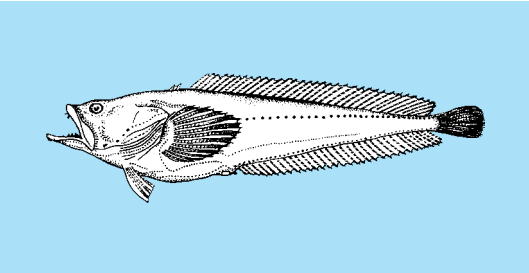

Atlantic midshipman (*Porichthys porosissimus*), about 8 in. (20 cm) long.

genus (*Sanopus*) inhabits coral reefs and is brightly colored. The Porichthyinae are New World species that have two solid dorsal spines, solid opercular spine, and no venom glands; a subopercular spine is present as are canine teeth; the body is scaleless with several lateral lines; and there are photophores in all but two species (*Aphos*). Members of the genus *Porichthys* (=*Nautopaedium*) have lines of bright yellow photophores (buttons) in the lower sides and belly, hence the common name midshipmen (see **illustration**). It is very unusual for a shallow-water fish to possess photophores. The Thalassophryninae are venomous toadfishes and also New World in distribution. They have a scaleless body; no subopercular spine; no photophores; no canine teeth; the lateral line is single or absent. The two dorsal spines and opercular spine are hollow, hypodermic-needle-like weapons that deliver a venom capable of producing excruciating wounds. *See* ACTINOPTERYGII; BIOLUMINESCENCE; TELEOSTEI. Herbert Boschung

Bibliography. B. B. Collette, A review of the venomous toadfishes, subfamily Thalassophryninae, *Copeia*, 1966(4):846–864, 1966; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994.

Battery

An electrochemical device that stores chemical energy which can be converted into electrical energy, thereby providing a direct-current voltage source. Although the term "battery" is properly applied to a group of two or more electrochemical cells connected together electrically, both single-cell and multicell devices are called battery. *See* ELECTROCHEMISTRY; ELECTROMOTIVE FORCE (CELLS).

Types. The two general types are the primary battery and the secondary battery. The primary battery delivers current as the result of a chemical reaction that is not efficiently reversible. Practically, this makes the primary battery nonrechargeable. Only one intermittent or continuous discharge can be obtained before the chemicals placed in it during manufacture are consumed. Then the discharged primary battery must be replaced. The secondary or storage battery is rechargeable because it delivers current as the result of a chemical reaction that is easily reversible. When a charging current flows through its terminals in the direction opposite to the current flow during discharge, the active materials in the secondary battery return to approximately their original charged condition.

Components. The cell is the basic electrochemical unit. It has three essential parts: (1) a negative electrode (the anode) and (2) a positive electrode (the cathode) that are in contact with (3) an electrolyte solution. The electrodes are metal rods, sheets, or plates that are used to receive electrical energy (in secondary cells), store electrical energy chemically, and deliver electrical energy as the result of the reactions that occur at the electrode-solution surfaces. Solid polymer or plastic active materials have been developed that can serve as the cathode in rechargeable batteries. The electrolyte is a chemical compound (salt, acid, or base) that when dissolved in a solvent forms a solution that becomes an ionic conductor of electricity, but essentially insulating toward electrons—properties that are prerequisites for any electrolyte. In the cell or battery, this electrolyte solution is the conducting medium in which the flow of electric current between electrodes takes place by the migration of ions. When water is the solvent, an aqueous solution is formed. Some cells have a nonaqueous electrolyte, for example, when alcohol is used as the solvent. Other cells have a solid electrolyte that when used with solid electrodes can form a leak-free solid-state cell or battery.

During charging of a secondary cell, the negative electrode becomes the cathode and the positive electrode becomes the anode. However, electrode designation as positive or negative is unaffected by the operating mode of the cell or battery. Two or more cells internally connected together electrically, in series or parallel, form a battery of a given voltage. Typical are the rectangular 9-V primary battery, which has six flat 1.5-V zinc-carbon or alkaline "dry" cells connected in series, and the 12-V automotive or secondary battery, which has six 2.1-V lead-acid "wet" cells connected in series.

Size. Both primary and secondary cells are manufactured in many sizes, shapes, and terminal arrangements, from the miniature coin- or button-shaped battery (which has a diameter greater than its height) and the small cylindrical penlight battery to the large submarine battery, where a single rectangular cell has weighed 1 ton (0.9 metric ton). For optimum performance, the battery must be constructed for its particular application consistent with cost, weight, space, and operational requirements. Automotive and aircraft batteries are secondary batteries that have relatively thin positive and negative plates with thin, porous, envelope separators to conserve space and weight and to provide high rates of current discharge at low temperatures. Standby batteries are secondary batteries that use thick plates and thick separators to provide long life. Solid-state batteries can be constructed with unusual features and in irregular sizes and shapes. Size and weight reductions in all types of batteries continue to be made through use of new materials and methods of construction.

Selection and applications. Batteries are probably the most reliable source of power known. Most critical electrical circuits are protected in some

manner by battery power. Since a battery has no moving parts, tests, calculations, or comparisons are made to predict the conditions of the cells in some batteries. Growing battery usage reflects the increased demand for portable computers; mobile voice, data, and video communications; and new or redesigned/repowered products in the consumer, industrial, and transportation sectors. Further growth may result from significant increases in dc system operating voltage, such as from 12 V to 24 V or 48 V, which can provide much higher power generally with less weight, greatly broadening the range of potential battery applications.

For most applications, the basic choice in selection is whether to use either a primary (nonrechargeable) or a secondary (rechargeable) cell or battery. Electrical characteristics affecting selection include maximum and minimum voltage, current drain, and pulse current (if any), its duration and frequency of occurrence. Other factors such as performance in the specific application, operating environment, and final packaging of the cell or battery also must be considered.

Primary battery usage. Primary batteries are used as a source of dc power where the following requirements are important:

1. Electrical charging equipment or power is not readily available.
2. Convenience is of major importance, such as in a hand or pocket flashlight.
3. Standby power is desirable without cell deterioration during periods of nonuse for days or years. Reserve-electrolyte designs may be necessary, as in torpedo, guided missile, and some emergency light and power batteries.
4. The cost of a discharge is not of primary importance.

Secondary battery usage. Secondary batteries are used as a source of dc power where the following requirements are important:

1. The battery is the primary source of power and numerous discharge-recharge cycles are required, as in wheelchairs and golf carts, industrial hand and forklift trucks, electric cars and trucks, and boats and submarines.
2. The battery is used to supply large, short-time (or relatively small, longer-time), repetitive power requirements, as in automotive and aircraft batteries which provide power for starting internal combustion engines.
3. Standby power is required and the battery is continuously connected to a voltage-controlled dc circuit. The battery is said to "float" by drawing from the dc circuit only sufficient current to compensate automatically for the battery's own internal self-discharge. Computers and communications networks and emergency light and power batteries are in this category.
4. Long periods of low-current-rate discharge followed subsequently by recharge are required, as in marine buoys and lighthouses, and instrumentation for monitoring conditions such as earthquakes and other seismic disturbances.
5. The very large capacitance is beneficial to the circuit, as in telephone exchanges.

Ratings. Two key ratings of a cell are its voltage and ampere-hour (Ah) capacity. The voltage is determined by the chemical system created by the active materials used for the negative and positive electrodes (anode and cathode). The ampere-hour capacity is determined by the amount of the active materials contained in the cell. The product of these terms is the energy output or watthour capacity of the battery. In actual practice, only one-third to one-half of the theoretical capacity may be available. Battery performance varies with temperature, current drain, cutoff voltage, operating schedule, and storage conditions prior to use, as well as the particular design.

Many primary batteries are rated by average service capacity in milliampere-hours (mAh). This is the number of hours of discharge that can be obtained when discharging at a specified temperature through a specified fixed resistance to a specified final or cutoff voltage, which is either at the point of rapid voltage drop or at minimum usable voltage.

Secondary batteries, such as automotive batteries, have been rated by ampere-hour capacity. Typically, this is the amount of current that the battery can deliver for 20 h without the temperature-corrected cell voltages dropping below 1.75 V per cell. A battery capable of giving 2.5 A for 20 h is rated at 50 ampere-hours at the 20-h rate. This same battery may provide an engine-cranking current of 150 A for only 8 min at 80°F (27°C) or for 4 min at 0°F (−18°C), giving a service of only 20 and 10 ampere-hours, respectively. By multiplying ampere-hours by average voltage during discharge, a more practical watt-hour rating is obtained. Automotive batteries also are rated by reserve capacity (RC) and cold-cranking amps (CCA). Reserve capacity is the length of time that a fully charged battery at 80°F (27°C) can deliver 25 amperes before the voltage falls to 10.5 V. A typical rating is 125 min, which is the length of time the battery could carry a minimum electrical load after failure of the vehicle's charging system. Cold-cranking amps is a measure of the ability of a battery to crank an engine when the battery is cold. It is measured by the number of amperes that a 12 V battery can deliver for 30 s when it is at 0°F (−18°C) without the battery voltage falling below 7.2 V. A typical CCA rating for a battery with a reserve capacity of 125 min is 430 amperes. The CCA rating for most automobile batteries is between 300 and 600 amperes.

Life. The life expectancy of a cell or battery depends on its design and materials, as well as its application and operating conditions. Life expectancy is measured by shelf life and service life. Shelf life is the expected time that elapses before a stored battery becomes inoperative due to age or deterioration, or unusable due to its own internal self-discharge. Service life is the expected length of time or number of discharge-charge cycles through which a battery remains capable of delivering a specified percentage of its capacity after it has been put into service. This can vary from the one shot or single discharge

TABLE 1. Comparison of principal chemical systems of primary batteries

Principal chemical systems	General characteristics under optimum operating conditions	Configurations and rated capacity*	Nominal cell voltage, V
Zinc-carbon (Zn/MnO_2)	Popular, common, low-cost battery; moderate shelf life; sloping discharge characteristic	Cylindrical cells, to about 40 Ah	1.5
Zinc-alkaline-manganese dioxide (Zn/MnO_2)	Popular, general-purpose premium battery; good low-temperature and high-rate performance; good shelf life; sloping discharge characteristic	Cylindrical cells to 20 Ah; also available in small button cells	1.5
Zinc-mercuric oxide (Zn/HgO)	High volumetric capacity; good shelf life; flat discharge characteristic	Button and cylindrical cells from 40 mAh to 13 Ah; also available in special larger sizes	1.35 or 1.40
Zinc-silver oxide (Zn/Ag_2O)	High gravimetric capacity; good shelf life; flat discharge characteristic	Button cells to 180 mAh for specialized applications	
Zinc-air (Zn/O_2)	Highest energy density on continuous discharge; excellent shelf life (unactivated); limited rate capability and shelf life when activated; flat discharge characteristic	Button cells to 1150 mAh; larger cells to 6.5 Ah	1.4
Lithium-manganese dioxide (Li/MnO_2)	High energy density; good rate capability and low-temperature performance; excellent shelf life; relatively flat discharge characteristic	Coin cells to 500 mAh; spiral-wound cylindrical cells to 1.25 Ah; bobbin-type cells to 2.5 Ah	3.0
Lithium-sulfur dioxide (Li/SO_2)	High energy density; excellent high-rate and low-temperature performance; pressurized, hermetically sealed cell; excellent shelf life; flat discharge characteristic	Cylindrical cells from 700 mAh to 19 Ah	3.0

*Button cells are rated at 500–1000 h at 70°F (21°C), cylindrical cells at 50–100 h. The cutoff voltage is approximately 60% of the nominal voltage. Typical capacities can be higher.

obtainable from primary cells to 10,000 or more discharge-charge cycles obtainable from some secondary batteries. Automotive batteries may last as little as 18 months in hot climates and 10–12 years in cold climates, but typically they have an average life of 3 years and may last for 6000 cycles. Industrial batteries have a 10–20-year service life. Standby sizes may be expected to float across the dc bus 8–30 years. Generally the most costly, largest, and heaviest cells have the longest service life.

Many batteries experience an abrupt loss of voltage without warning when the active materials are depleted. However, in some batteries, open-circuit voltage serves as a state-of-charge indicator, decreasing slightly but continuously with loss of capacity. Many electronic devices now include a battery that will last for the life of the product. This eliminates the need for the consumer or user to replace the battery and risk causing damage to the device or the battery by inadvertently shorting the terminals, reversing the polarity, or installing a similar-size battery having the wrong chemistry.

To obtain the maximum life and ensure reliability of batteries, the manufacturer's recommendations for storage and maintenance must be followed. The stated shelf life and temperature of primary cells must not be exceeded. For dry reserve-electrolyte primary cells and secondary cells of the dry construction with charged plates, the cell or battery container must be protected against moisture, and storage must be within specified temperature limits. Wet, charged secondary batteries may require periodic charging and water addition, depending upon the construction.

Primary Batteries

A primary cell or battery is not intended to be recharged and is discarded when it has delivered all its electrical energy (**Fig. 1**). Several kinds of primary cell are widely used, particularly in portable devices and equipment, providing freedom from the dependence on alternating-current line power. They are convenient, lightweight, and usually relatively inexpensive sources of electrical energy that provide high energy density (long service life) at low-to-moderate or intermittent discharge rates, good shelf life, and ease of use while requiring little or no maintenance.

Primary cells are classified by their electrolyte, which may be described as aqueous, nonaqueous, aprotic, or solid. In most primary cells the electrolyte is immobilized by a gelling agent or mixed as a paste,

	Energy density †					Shelf life, % loss per year at 20°C (68°F)	Typical operating temperature,† °F (°C)
	Wh/lb	Wh/kg	Wh/in.^{3}	Wh/L	Typical applications		
Cylindrical:	34	75	2.3	140	General purpose for lighting, radio, novelties	10	23 to 113° (−5 to 45°)
Button: Cylindrical:	25 60	55 130	2.4 5.8	145 350	General purpose, such as lighting, photographic equipment, toys, radio and TV, tape recorders, watches and instruments; popular for high-drain applications	3	−4° to 130° (−20° to 54°)
Button: Cylindrical:	50 55	110 120	7.3 7.4	445 450	Hearing aids, medical and photographic equipment, communication devices, papers, detection and emergency equipment, and other applications requiring steady output voltage	4	15° to 130° (−9° to 54°)
Button:	60	130	8.2	500	Hearing aids, watches, photographic equipment, and special electronics requiring small, high-capacity batteries	6	−4° to 130° (−20° to 54°)
Button:	140	310	19.0	1150	Hearing aids, medical monitoring instruments, pagers, communication devices, emergency lighting, and other relatively short-term applications	2 (unactivated)	32° to 122° (0° to 50°)
Coin: Cylindrical: spiral-wound bobbin	80 105 135	175 230 300	8.3 8.3 10.1	505 505 61.5	General purpose in photographic and electronic applications requiring small, high-capacity batteries, such as watches, calculators, computers, clock/calendar, memory backup, photoflash, motor-driven cameras, and instruments	0.5−1	−4° to 140° (−20° to 60°) −4° to 140° (−20° to 60°) −4° to 160° (−20° to 71°)
Cylindrical:	275	125	440	7.1	Military and special industrial applications (such as process control equipment), requiring high-capacity, high rate, and extreme temperature operation	0.5−1	−40° to 160° (−40° to 71°)

†Batteries can perform over a wider temperature range, under certain discharge conditions.

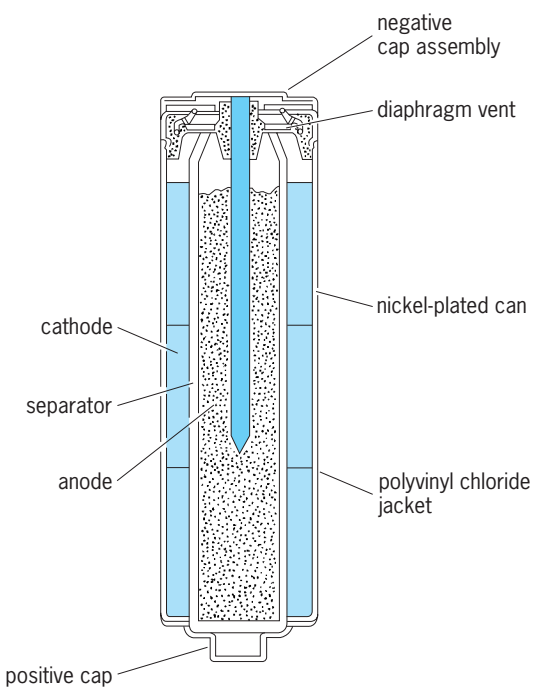

Fig. 1. Diagram of a zinc-alkaline-manganese dioxide cylindrical cell.

with the term "dry cell" commonly applied to the zinc-carbon Leclanche cell and sometimes to other types. An aqueous electrolyte or electrolyte system is used in zinc-carbon, magnesium, alkaline-manganese dioxide, mercuric oxide, silver oxide, and zinc-air cells. Nonaqueous electrolyte systems are used in lithium cells and batteries. *See* ELECTROLYTE.

Typical characteristics of different types of primary batteries and their applications are summarized in **Table 1.** Performance characteristics of primary batteries at various temperatures are shown in **Fig. 2.**

Zinc-carbon cells. The zinc-carbon or LeClanche dry cell was invented in 1866 and continues to be popular. It is made in sizes of varying diameter and height, and batteries are available in voltages ranging from 1.5 to 510 V. Common cell construction uses the cylindrical zinc container or can as the negative electrode and manganese dioxide (mixed with carbon black to increase conductivity and retain moisture) as the positive active material, with the electrical connection made through a center carbon electrode. The slightly acidic electrolyte is an aqueous solution of ammonium chloride and may also contain zinc chloride, immobilized in a paste or the paper separator. The electrochemical reaction between the cathode, anode, and electrolyte in a

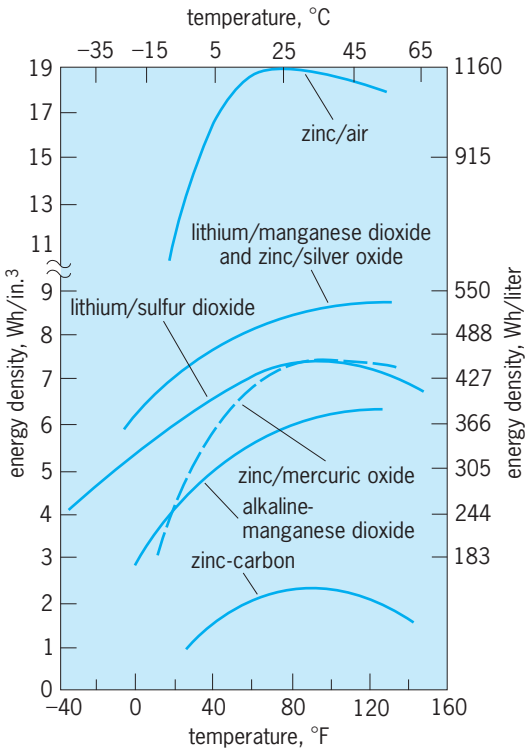

Fig. 2. Performance characteristics of primary batteries at various temperatures.

zinc-carbon LeClanche cell is shown in reaction (1).

$$2MnO_2 + 2NH_4Cl + Zn \rightarrow ZnCl_2 \cdot 2NH_3 + Mn_2O_3 + H_2O \quad (1)$$

Typical open-circuit voltage of a fresh LeClanche cell is over 1.55 V. The closed-circuit voltage gradually declines as a function of the depth of discharge.

Significant improvements in capacity and shelf life of zinc-carbon batteries have been made through the use of new cell designs, such as the paper-lined cell, and new materials. Upgrading to the use of purer beneficiated manganese dioxide and an electrolyte that consists mainly of zinc chloride and water has created the heavy-duty or zinc chloride cell. Its construction is similar to the zinc-carbon cell, but having an anode of high-purity zinc alloy and typically a higher proportion of carbon to manganese dioxide and a greater volume of slightly more acidic electrolyte. The electrochemical reaction in a zinc chloride cell is shown in reaction (2). Open-circuit

$$8MnO_2 + 4Zn + ZnCl_2 + 9H_2O \rightarrow 8MnOOH + ZnCl_2 \cdot 4ZnO \cdot 5H_2O \quad (2)$$

voltage of a fresh zinc chloride cell typically is over 1.60 V. As in the LeClanche cell, closed-circuit voltage gradually declines as a function of the depth of discharge.

Compared to zinc-carbon batteries, zinc chloride batteries have improved high-rate and low-temperature performance and are available in voltages ranging from 1.5 to 12 V. A common design is the flat cell, used in multicell batteries such as the 9-V battery, which offers better volume utilization and, in some designs, better high-rate performance. Characteristics of standard zinc-carbon and zinc chloride cells are given in **Table 2.**

Magnesium cells. The magnesium-carbon primary cell is basically a cylindrical zinc-carbon cell in which the can or container is made of magnesium or its alloy instead of zinc. Magnesium has a greater electrochemical potential than zinc, providing an open-circuit voltage of about 2.0 V. Developed for military use in radios and other equipment, the magnesium

TABLE 2. Characteristics of 1.5-V zinc-carbon unit cells

Cell designation*			Weight, g (oz)	Maximum dimensions, mm (in.)		Capacity: 2 h/day			
						Standard cell		Zinc chloride cell	
ANSI	IEC	NEDA		Diameter	Height	Starting drain, mA	Service hours	Starting drain, mA	Service hours†
N	R-1	910	6.2 (0.22)	12 (0.47)	30.2 (1.19)	1.5	320		
						7.5	60		
						15	20		
AAA	R-03	24	8.5 (0.30)	10.5 (0.41)	44.5 (1.75)	2	350		
						10	54		
						20	21		
AA	R-6	15	17 (0.60)	14.5 (0.57)	50.9 (1.99)	3	450		
						15	80	187.5	5.8
						30	32	375	1.9 (cont.)
C	R-14	14	40 (1.4)	26.2 (1.03)	50.0 (1.97)	5	520	187	17
						25	115	375	4.6 (cont.)
						50	53		
D	R-20	13	94 (3.3)	34.1 (1.34)	61.5 (2.42)	10	525	9	1100
						50	125	187	42
						100	57	667	5.3 (cont.)
F	R-25	60	160 (5.6)	34.1 (1.34)	91.4 (3.60)	15	630	187	10
						75	135	375	36
						150	60	667	18

*ANSI = American National Standards Institute: IEC = International Electrotechnical Commission: NEDA = National Electrical Distributors Association.
†cont. = continuous.

cell has twice the capacity or service life of a zinc-carbon cell of equivalent size, and longer shelf life, particularly at elevated temperatures. However, when the load is applied to a magnesium cell, a characteristic voltage drop occurs briefly before the cell recovers to its usable voltage. This may make the magnesium cell unacceptable for use in some applications. The magnesium cell has a vent for the escape of hydrogen gas, which forms during discharge.

Alkaline-manganese dioxide cells. These cells became the most popular battery during the 1980s. Available in voltages ranging from 1.5 to 12 V and in a variety of sizes and shapes, they have higher energy density, better low-temperature and high-rate performance, and longer shelf life than a zinc-carbon battery. Alkaline and zinc-carbon cells are chemically similar, both using zinc for the anode and manganese dioxide for the cathode active materials. However, the cells differ significantly in construction, electrolyte, and formulation of the active materials.

The container for the cylindrical alkaline cell is a steel can that does not participate in the electrochemical reaction. A gelled mixture of zinc powder and electrolyte is used for the anode. The highly alkaline electrolyte is an aqueous solution of potassium hydroxide, which is more conductive than the salt of the zinc-carbon cell. The cathode is a highly compacted, conductive mixture of high-purity manganese dioxide and graphite or carbon. To prevent contact between the anode and cathode, which would result in a very active chemical reaction, a barrier or separator is placed between the two. A typical separator is paper or fabric soaked with electrolyte. The separator prevents solid-particle migration, while the electrolyte promotes ionic conductivity. The electrochemical reaction in an alkaline-manganese dioxide cell is shown in reaction (3). Open-circuit voltage of a fresh cylin-

$$2Zn + 3MnO_2 \rightarrow 2ZnO + Mn_3O_4 \qquad (3)$$

drical alkaline cell typically is 1.58 V. Closed-circuit voltage gradually declines as a function of the depth of discharge.

Modifying the internal design of zinc-alkaline cells has resulted in cells that have an average of 20–50% higher amperehour capacity or service life. This newer cell design provides a significantly greater volume for the active materials, the factor most responsible for improved performance. The modified cell is manufactured with a nickel-plated steel can and an outside poly(vinyl chloride) jacket (Fig. 1). A diaphragm vent is incorporated into the assembly in case of pressure buildup. Alkaline-manganese dioxide cells also are made in miniature sizes with flat, circular pellet-type cathodes and homogeneous gelled anodes.

Mercuric oxide cells. The zinc-mercuric oxide cell (**Fig. 3**) has high capacity per unit charge, relatively constant output voltage during discharge, and good storage qualities. It is constructed in a sealed but vented structure, with the active materials balanced to prevent the formation of hydrogen when

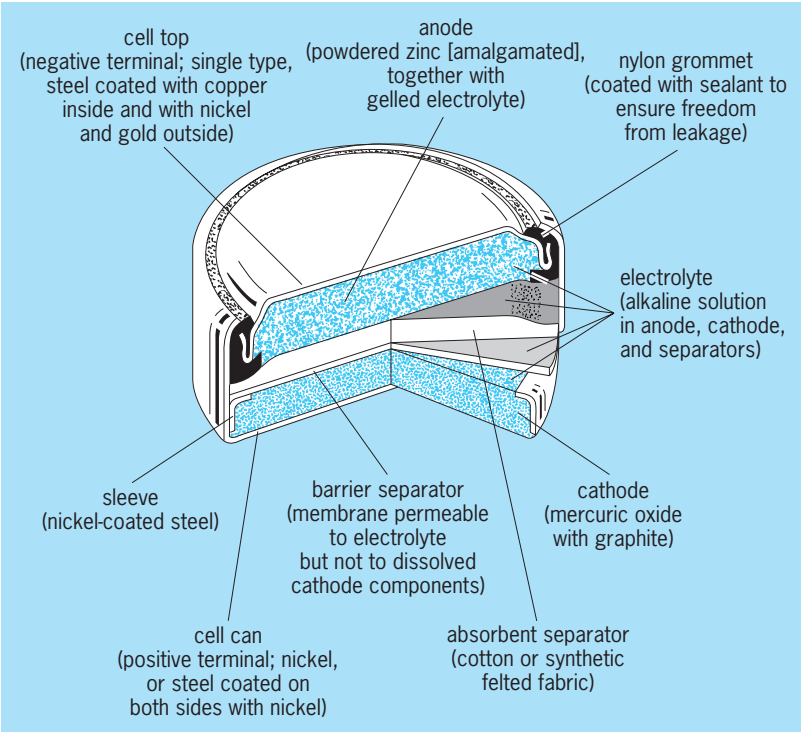

Fig. 3. Diagram showing the design of a mercury button cell.

discharged. Three basic structures are used: the wound anode, the flat pressed-powder anode, and the cylindrical pressed-powder electrode. Typically, the anode is zinc, the cathode is mercuric oxide, and the electrolyte is alkaline. The electrochemical reaction in a mercuric oxide cell is shown in reaction (4).

$$Zn + 2MnO_2 \rightarrow ZnO + Mn_2O_3 \qquad (4)$$

Open-circuit voltage of a mercuric oxide cell is 1.35 V, with some batteries delivering up to 97.2 V.

Mercuric oxide batteries have been made in various sizes and configurations, from the miniature button 16 mAh to the large 14 Ah cylindrical cell, and are suited for use at both normal and moderately high temperatures. However, the low-temperature performance is poor, particularly under heavy discharge loads. Use at temperatures below 32°F (0°C) is generally not recommended. Mercuric oxide batteries have been used for hearing aids, cameras, watches, calculators, implanted heart pacemakers and other medical applications, emergency radios, and military equipment. Today concerns about health hazards and environmental regulations related to disposal generally prevent manufacture and sale of mercuric oxide batteries. They are replaced with newer, less expensive designs such as zinc-air that do not contain mercury or its compounds.

The cadmium-mercuric oxide battery is similar to the zinc-mercuric oxide battery. The substitution of cadmium for the zinc anode lowers the cell voltage but offers a very stable system, with a shelf life of up to 10 years, and improved performance at low temperatures. Its watthour capacity, because of the lower voltage, is about 60% of the zinc-mercuric

oxide battery. However, health concerns and environmental regulations related to disposal also limit the availability and usage of batteries containing cadmium.

Silver oxide cells. The zinc-silver oxide primary cell is similar to the zinc-mercuric oxide cell, but uses silver oxide in place of mercuric oxide (Fig. 3). This results in a higher cell voltage and energy. In the small button-cell configuration, this is a significant advantage for use in hearing aids, photographic applications, watches, and calculators. The silver oxide battery has a higher voltage than a mercury battery; a flatter discharge curve than an alkaline battery; good resistance to acceleration, shock, and vibration; an essentially constant and low internal resistance; and good performance characteristics at temperature extremes.

Silver oxide batteries generally have a flat circular cathode and a homogeneous gelled anode. The cathode is a mixture of silver oxide with a low percentage of manganese dioxide and graphite, and the anode is a mixture of amalgamated zinc powder. A highly alkaline electrolyte of either sodium hydroxide or potassium hydroxide is used, although potassium hydroxide makes the battery more difficult to seal. The separator is a material that prevents migration of any solid particles in the battery.

The chemical reaction in a silver oxide cell is shown in reaction (6). Open-circuit voltage of a silver oxide cell is 1.6 V, with batteries available having a nominal voltage of 6 V.

$$Zn + Ag_2O \rightarrow ZnO + 2Ag \qquad (6)$$

Zinc-air cells. Zinc-air cells (**Fig. 4**) have the highest energy density of the commercially available primary cells. They use atmospheric oxygen for the active cathode material, so there is no need to include the cathode material in the cell. This allows the cathode electrode to be very thin. The remaining space can be used for increasing the amount of zinc, which is the active anode material, resulting in a higher cell capacity (Fig. 4*a*).

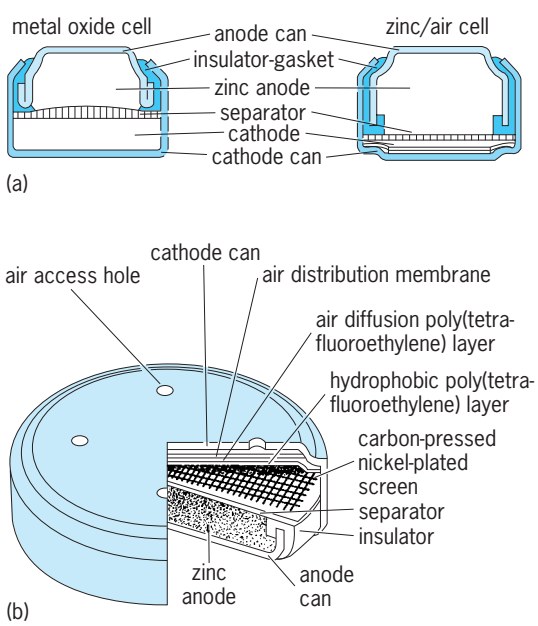

Fig. 4. Zinc-air cell. (*a*) Cross section compared with metal oxide cell. (*b*) Major components.

The zinc-air cell is usually constructed in a button configuration (Fig. 4*b*). The cell consists of two cans, isolated from each other by an annular insulator. One can contains the zinc anode and the other the air or oxygen cathode. The anode can is fabricated from a triclad metal; the external material is nickel for good electrical conductivity, the middle layer is stainless steel to provide strength, and the inner surface is copper, which is compatible with the cell components. The cathode can is fabricated of nickel-plated steel and contains holes that allow air to enter the cell. A layer of permeable polytetrafluoroethylene (Teflon) serves as a separator to assure the proper distribution of air and limit the entrance or exit of moisture. The anode is an amalgamated gelled zinc powder and electrolyte. Although the cathode uses atmospheric oxygen as the active material, a mixture of carbon, polytetrafluoroethylene, and manganese dioxide is impressed on a nickel-plated screen. The carbon and manganese dioxide serve as catalysts for the oxygen reaction. The electrolyte is an aqueous solution of potassium hydroxide with a small amount of zinc oxide.

The air holes of the zinc-air cell are sealed until the cell is used, to inhibit the entrance of air. In this condition, the cells can retain more than 95% of their rated capacity after 1 year of storage at room temperature. The cells are activated by removing the seal which permits the flow of air into the cell. The chemical reaction in a zinc-air cell is shown in reaction (7).

$$2Zn + O_2 \rightarrow 2ZnO \qquad (7)$$

The open-circuit voltage of a zinc-air cell is about 1.4 V, providing a flat discharge with the operating voltage between 1.35 and 1.1 V, depending on the discharge conditions. The cell is capable of low to moderately high discharge rates, and is used in applications requiring a relatively short operating time before replacement.

Zinc-air cells are manufactured in sizes from 50 to 6500 mAh. Multicell batteries are available in a wide range of voltage and capacity ratings, including a nominal 8.4-V zinc-air battery that is interchangeable in some applications with the zinc chloride and alkaline 9-V battery.

Lithium cells. Lithium is a silvery-white element that is the lightest metal, having a density only about half that of water, with which the alkaline lithium is highly reactive. Lithium has good conductivity, and its standard potential and electrochemical equivalence are higher than those of any other metal. It is widely used in various forms as battery anode material (negative electrode), especially in coin-size primary cells. There are several different types of lithium cells, primary as well as reserve and secondary, similar to the variety of cells using a zinc anode. Each type of lithium cell differs in the cathode material for the positive electrode, electrolyte, and cell chemistry as well as in physical, mechanical,

TABLE 3. Classification of lithium primary cells*

Cell classification	Typical electrolyte	Power capability	Size, Ah	Operating range, °F (°C)	Shelf life, years	Typical cathodes	Nominal cell voltage, V	Key characteristics
Soluble cathode (liquid or gas)	Organic or inorganic (w/solute)	Moderate to high power, W	0.5–20,000	−67 to 158 (−55 to 70)	5–10	Sulfur dioxide Thionylchloride Sulfuryl chloride	3.0 3.6 3.9	High energy output; high power output; low temperature operation; long shelf life
Solid cathode	Organic (w/solute)	Low to moderate power, mW	0.03–5	−40 to 122 (−40 to 50)	5–8	Silver chromate Manganese dioxide Carbon monofluoride Copper(I) sulfide Iron disulfide Iron sulfide Copper(II) oxide	3.1 3.0 2.6 1.7 1.6 1.5 1.5	High energy output for moderate power requirements; nonpressurized cells
Solid electrolyte	Solid state	Very low power, μW	0.003–0.5	32 to 212 (0 to 100)	10–25	I_2 poly-2-vinylpyridine	2.8	Excellent shelf life; solid-state—no leakage; long-term microampere discharge

*From D. Linden (ed.), *Handbook of Batteries and Fuel Cells*, McGraw-Hill, 2d ed., 1995.

and electrical features (**Table 3**). Performance advantages of lithium cells over other primary cells include high voltage (may be above 3 V), high energy density, operation over a wide temperature range, and long shelf life. Lithium cells are manufactured in different sizes and configurations, ranging in capacity from milliamperehours to over 20,000 Ah, with characteristics matched to the general application. *See* LITHIUM.

In a lithium cell, nonaqueous solvents must be used for the electrolyte because of the solubility and reactivity of lithium in aqueous solutions. Organic solvents, such as acetonitrile and propylene carbonate, and inorganic solvents, such as thionyl chloride, are typical. A compatible solute is added to provide the necessary electrolyte conductivity. Solid-cathode cells are generally manufactured as cylindrical cells in sizes up to about 30 Ah in both low- and high-rate constructions. The low-rate, or high-energy, designs use a bobbin-type construction to maximize the volume for active materials (**Fig. 5***a*). These cells have the highest energy density of cells of similar size. The high-rate cells use a jelly-roll (spiral-wound) construction, which provides the large electrode surface required for the high discharge rates (Fig. 5*b*). These cells deliver the highest rates of all primary batteries and are widely used in military applications. Early ambient-temperature lithium systems used lithium-sulfur dioxide chemistry, and found applications in medical implants, weapons systems, and spacecraft. These cells provide 3 V and have seven times the energy of a typical alkaline battery. However, other lithium chemistry has been developed. Four types of cells are lithium-carbon monofluoride, lithium-manganese dioxide, and lithium-iron disulfide, all of which have solid cathodes, and lithium-thionyl chloride, which has a liquid cathode.

Lithium-carbon monofluoride batteries have a solid carbon monofluoride cathode and characteristics that allow the battery to be used for the life of the product in which the battery is installed, such as computer real-time clock and memory backup applications, instead of having scheduled replacement intervals. The cell has a storage and operational temperature range of −40 to +85°C, and a long life made possible by a self-discharge rate that may be

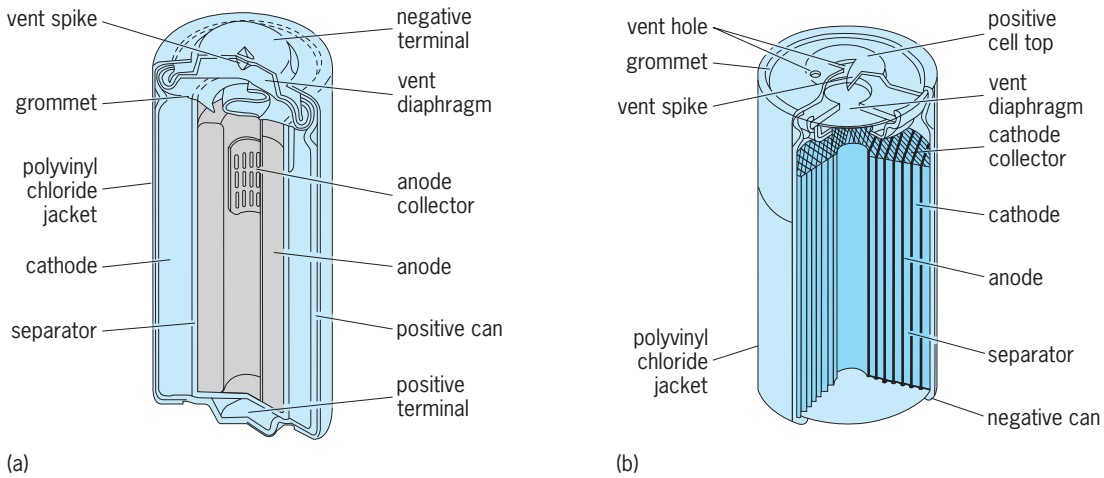

Fig. 5. Solid cathode cells. (*a*) Lithium-manganese dioxide bobbin cell. (*b*) Lithium-manganese dioxide spiral-wound cell. (*Duracell, Inc.*)

less than 0.2% of capacity loss per year. As a lithium-carbon monofluoride cell discharges while powering a specified load, internal resistance generally remains low and level while the closed-circuit voltage profile remains high, flat, and stable until the depth of discharge exceeds approximately 85%. When completely discharged, this type of lithium cell may be disposed of as nonhazardous waste. Batteries containing unreacted lithium metal are considered reactive hazardous waste and must be disposed of following the applicable local, state, and federal regulations.

Lithium-manganese dioxide cells have a solid manganese dioxide cathode and characteristics that make the preferred use of the cell to be where periodic replacement is routinely performed. The cell can supply both pulse loads and very small current drains, typical needs for microprocessor applications. As the cell discharges, its internal resistance increases because of the manganese dioxide, causing a tapered discharge profile and a declining closed circuit voltage. The manganese dioxide also limits the maximum temperature rating of the cell to 140°F (60°C), a temperature at which self-discharge may increase to a rate of over 8% per year. However, the cell costs less than a lithium-carbon monofluoride cell, and is widely used especially for applications requiring intermittent pulsing such as remote keyless entry systems. When completely discharged, this type of lithium cell may be disposed of as non-hazardous waste.

Lithium-iron disulfide cells have an anode of lithium foil in contact with a stainless-steel can and a cathode-material mixture of graphite and iron disulfide. These cells have an open-circuit voltage of 1.8 V, a nominal voltage of 1.5 V, and can be used in any application that uses other AA-size 1.5 V batteries. Each battery contains two safety devices: a thermal switch which acts as a current limiter if the battery begins to overheat, and a pressure-relief vent that opens at a predetermined temperature. This type of lithium cell may be disposed of using the same procedures approved for other cells.

Lithium-thionyl chloride cells were designed for the military and have been used for nuclear surety applications, which require long-life power sources for command and control systems. These cells have a voltage of about 3.6 V, making possible the use of fewer cells for the same application. They are classed as high-energy, high-current-drain-rate cells, with twice the energy density of lithium-sulfur dioxide cells, and will last twice as long. Lithium-thionyl chloride cells have a soluble cathode which is liquid, toxic, and corrosive. Each cell is hermetically sealed, with a designed-in fuse to avoid rupture or explosion if the battery is inadvertently charged or shorted. A lithium-thionyl chloride cell should be replaced only by a trained technician. Because thionyl chloride is highly toxic and corrosive, this battery is an environmental hazard which requires special handling and disposal.

Solid-electrolyte cells. These may be classified as (1) cells using a solid crystalline salt (such as lithium iodide) as the electrolyte and (2) cells using a solid-polymer electrolyte. With either type of electrolyte, the conductivity must be nearly 100% ionic as any electronic conductivity causes a discharge of the cell, which limits shelf life. Solid-polymer electrolytes are based on the characteristic that solutions of alkali metal salts dissolved in certain polymers, such as poly(ethylene oxide), form solid polymers that have reasonable ionic conductivity. The solid polymers can be fabricated in thin cross sections, are more stable than liquid electrolytes, and can serve as the separators. Most development work on thin-film polymer electrolyte batteries has been for rechargeable batteries, using a lithium anode and solid cathode materials. Energy densities of 5–10 times that of lead-acid cells have been calculated. The technology can also be applied to the design of primary cells and batteries. *See* ELECTROLYTIC CONDUCTANCE; IONIC CRYSTALS.

The most widely used solid-electrolyte cells have lithium iodide as the solid electrolyte. These batteries have a solid lithium foil anode and a cathode that is largely iodine. The iodine is made conductive by the addition of an organic compound, poly(2-vinylpyridine). This iodine-polyvinylpyridine cathode has a tarlike consistency in the fresh battery, then solidifies gradually as the battery is discharged. Lithium and iodine are consumed during the discharge, and the reaction product, lithium iodide, forms in place in the region between the two reactants where it serves as the cell separator. Because the electrolyte continues to form as the battery discharges, the overall resistance of the cell continually increases with discharge. This results in a drop in cell voltage for a given current drain. The nominal cell voltage is around 2.8 V. These batteries have a very low rate of self-discharge, a long storage life, and high reliability, but can be discharged only at low rates because of the low conductivity of the electrolyte. Applications for these batteries include memory backup and implanted cardiac pacemakers.

A typical implantable battery has a volume of 0.4 in.^{3} (6 cm^3), weighs 0.8 oz (23 g), and has a 2 Ah capacity. The lifetime of this battery is 5–10 years since pacemakers typically draw only 15-30 microamperes. The battery is considered to be discharged when the output voltage drops to 1.8 V. This may be detected by monitoring the patient's pulse rate. *See* MEDICAL CONTROL SYSTEMS.

Cells that use only solid components, including an ion-conducting solid electrolyte, are called solid-state cells or batteries. Some of these have electrodes made of fast-ion conductors, materials that are good conductors for both ions and electrons. Examples include the layered-structure disulfides of titanium and vanadium, TiS_2 and VS_2, which can be used as the cathode or sink for lithium, and aluminum-lithium alloys, which can be used as the anode or source of lithium in lithium batteries. These materials can sustain high reaction rates because of their unique crystalline structures which allow the incorporation of ions into their crystalline lattices without destruction of those lattices. *See* ENERGY STORAGE; INTERCALATION COMPOUNDS; SOLID-STATE CHEMISTRY.

In a solid-state cell, the polycrystalline pressed

electrolyte is interspaced between a metallic anode and the solid cathode material. The electrodes are applied to the electrolyte by mechanically pressing the materials together, or in some cases the electrolyte is formed in place by reaction between the two electrodes. These cells are then stacked together to form a battery of the required voltage. A carbon current collector is often used on the cathode side, and this is frequently admixed with the cathode material. If the cathode material is sufficiently conductive, for example titanium disulfide, no carbon conductor is needed.

Reserve batteries. Most primary batteries are ready for use at the time of manufacture. However, high-energy primary batteries are often limited in their application by poor charge retention resulting from self-discharge between the electrolyte and the active electrode materials. The use of an automatically activated battery, or reserve battery, can overcome this problem, especially when the battery is required to produce high currents for relatively short periods of time (seconds or minutes), typical of military applications. A reserve battery is manufactured as a complete but inert battery that is stored in an inactive condition by keeping one of the critical cell components, such as the electrolyte, separated from the rest of the battery. The battery is activated just prior to use by adding this component manually or automatically. An important design consideration is to ensure that the electrolyte or other withheld component is delivered as quickly as possible at the time of activation while avoiding chemical short-circuiting of the cells.

Zinc-silver oxide reserve batteries. Automatically activated batteries have been used in missile and spacecraft applications, where the battery has to be activated from a remote source. The package contains a mechanism that drives the electrolyte out of the reservoir and into the cells of the battery. Most of these automatically activated batteries use zinc-silver oxide chemistry with a solution of potassium hydroxide for the electrolyte in order to achieve the required high discharge rates.

Magnesium water-activated batteries. Several types of water-activated (dunk type) batteries use magnesium anodes and silver chloride or cuprous chloride cathodes. The batteries are assembled dry, with the active elements separated by an absorbent. Activation occurs by pouring water into the battery container or by immersing the battery in water or seawater. Magnesium-silver chloride batteries have many marine applications, including buoys, beacons, flares, and safety lights. They also power the pingers which are attached to and help locate the cockpit

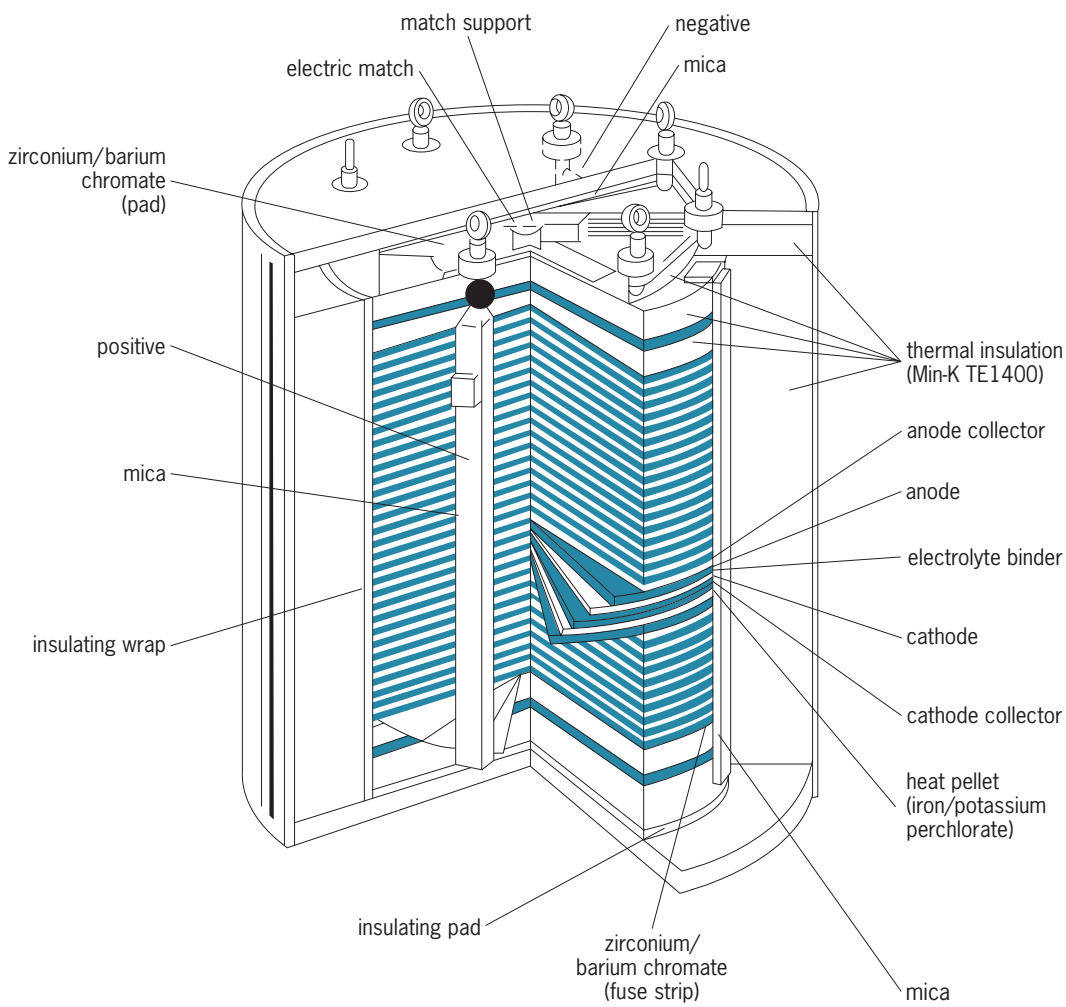

Fig. 6. Lithium/iron disulfide thermal battery. (*Sandia Laboratories*)

voice recorder and flight data recorder in downed aircraft, and emergency lighting for submarine escape hatches. Magnesium-copper iodide and magnesium-cuprous chloride batteries have similar capabilities.

Lithium-anode reserve batteries. Several types of batteries having a lithium anode are available as reserve batteries. These include batteries using lithium-sulfur dioxide, lithium-thionyl chloride, and lithium-vanadium pentoxide chemistry. Lithium-sulfur dioxide and lithium-thionyl chloride reserve batteries may be gas-activated, with activation triggered by electrical, mechanical, or pressure signals. The lithium-vanadium pentoxide battery is an electrolyte-activated battery. The electrolyte is contained in a glass ampule or vial, which is broken by mechanical shock, allowing the electrolyte to disperse within the cell.

Thermal batteries. A thermal battery, or fused-electrolyte battery, is a type of reserve battery that is activated by the application of heat. At room temperature, the electrolyte is a solid and has a very low conductivity, rendering the battery essentially inert. When the temperature is raised above the melting point of the electrolyte, the molten electrolyte becomes ionically conductive, and the battery is then capable of delivering electrical energy. Thermal batteries were introduced in 1955 to solve the wet-stand time limitation of aqueous systems, and subsequently were used as the main power source for nuclear weapon systems.

Early thermal batteries used calcium-calcium chromate chemistry. This was replaced in 1980 by thermal batteries that used a lithium anode, an iron sulfide cathode, and a lithium chloride-potassium chloride electrolyte. **Figure 6** shows a typical multicell lithium-iron disulfide thermal battery, which is activated by an electrical pulse to an integral electric match (squib). Other thermal batteries are activated mechanically using a percussion-type primer. The principal means for heating are pyrotechnic heat sources such as zirconium barium chromate heat paper or a heat pellet containing fine iron powder and potassium perchlorate. When the temperature rises to about 750°F (400°C), the electrolyte becomes molten and conductive. For other sources of electric energy known as batteries or cells *See* FUEL CELL; NUCLEAR BATTERY; SOLAR CELL.

Donald L. Anglin

Secondary Batteries

Secondary batteries (also known as accumulators) are rechargeable. This means that the electrochemical reactions in the cell must be reversible so that if the load in the external circuit is replaced by a power supply, the reactions in the cell can be forced to run in reverse, thereby restoring the driving force for reaction and hence recharging the cell. In contrast, primary batteries cannot be recharged because the reactions that produce current cannot be made to run in reverse; instead, totally different reactions occur when the cell is forcibly charged, and in some instances the reaction products are dangerous, that is, explosive or toxic.

The paradigm of battery design is to identify a chemical reaction with a strong driving force and then to fashion a cell that requires the reaction to proceed by a mechanism involving electron transfer, thereby making electrons available to a load in the external circuit. The magnitude of the driving force will determine cell voltage; the kinetics of reaction will determine cell current.

There are many chemistries that will serve as the basis for secondary batteries. Distinctions can be made on the basis of the following metrics: voltage, current (maximum, steady-state, and peak), energy density (Wh/kg and Wh/L), power density (W/kg and W/L), service life (cycles to failure), and cost ($/kWh). **Table 4** summarizes performance characteristics of some of the secondary battery types. Such comparisons are made difficult by the fact that many performance characteristics are functions of battery size and service conditions. For example, service life is greatly affected by depth of discharge and discharge current. However, both of these operating parameters can be very different for two different battery types that were designed for very different applications.

Another classification scheme considers the state of aggregation of the principal cell components, that is, whether the electrolyte is a solid or liquid, and whether the electrode active material is a solid, liquid, or gas. Most batteries have solid electrodes and a liquid electrolyte. However, there are examples of batteries in which the anode and cathode are both liquid, and the electrolyte is solid.

What follows is a series of descriptions of the significant secondary battery chemistries.

Lead-acid batteries. The lead-acid battery is the dominant secondary battery, used in a wide variety of applications, including automotive SLI (starting, lighting, ignition), traction for industrial trucks, emergency power, and UPS (uninterruptible power supplies). The attributes of lead-acid batteries

TABLE 4. Comparison of performance characteristics of secondary batteries

Characteristics	Pb-acid	Ni-Cd	Ni-MH	Zn-Ag	Na-S	Zn-air*	Li-ion	Li-SPE†
Nominal voltage, V	2	1.2	1.2	1.5	1.9	1.5	3.6	3.6
Specific energy, Wh/kg	35	40	90	110	80	280	125	500
Specific energy, kJ/kg	126	144	324	396	288	1008	450	1800
Volumetric energy, Wh/L	70	100	245	220		385	440	900
Volumetric energy, kJ/L	252	360	882	792		1386	1584	3240

*Based upon commercial prototype notebook computer battery.
†Projections based upon thin-film microbattery test results in the laboratory.

include low cost, high discharge rate, and good performance at subambient temperatures. The first practical lead-acid cell was made by Planté in France in 1860.

The anode is metallic lead. The cathode active material is lead dioxide, which is incorporated into a composite electrode also containing lead sulfate and metallic lead. The electrolyte is an aqueous solution of sulfuric acid, 37% by weight when the battery is fully charged.

The half-cell reactions at each electrode are shown in reactions (8) and (9). The overall cell reaction is the formation of lead sulfate and water, shown in reaction (10). When the battery is discharging,

$$\text{Anode}: Pb + SO_4^{2-} \rightarrow PbSO_4 + 2e^- \quad (8)$$

$$\text{Cathode}: PbO_2 + H_2SO_4 + 2H^+ + 2e^- \rightarrow PbSO_4 + 2H_2O \quad (9)$$

$$\text{Overall}: Pb + PbO_2 + 2H_2SO_4 \rightarrow 2PbSO_4 + 2H_2O \quad (10)$$

the reactions proceed from left to right; upon charging, the reactions proceed from right to left. The elementary electrochemical reactions are shown in reactions (11) and (12). On discharge, electrons are

$$\text{Anode}: Pb \rightarrow Pb^{2+} + 2e^- \quad (11)$$

$$\text{Cathode}: Pb^{4+} + 2e^- \rightarrow Pb^{2+} \quad (12)$$

produced at the anode as metallic lead is oxidized to Pb(II) present in lead sulfate; the complementary reaction at the cathode is the reduction of Pb(IV) present in lead dioxide to Pb(II) present in lead sulfate with the attendant consumption of electrons.

Sulfuric acid functions not only as the electrolyte but also as one of the reactants. Consequently, the acid concentration is a function of the instant state of charge and attains maximum value when the battery is fully charged.

Nominal voltage is 2 V, which decreases as the battery is discharged. The lower limit is governed by the depth of discharge. The theoretical specific energy is 170 Wh/kg. Commercially available lead-acid batteries deliver about 35 Wh/kg, which is about 20\% of theoretical. However, lead-acid batteries are capable of high specific power, have relatively flat discharge profiles, are comparatively insensitive to voltage changes with temperature, can be made for calendar lives exceeding 20 years, and are the least costly of all rechargeable technologies. Thus, when mobility is not a consideration and hence there is no penalty for the high density of lead, lead-acid technology can be found. The disadvantages of lead-acid technology include low specific energy and possibility of hydrogen evolution.

Nickel-cadmium batteries. The nickel-cadmium battery is the dominant alkaline secondary battery and is used in many of the same heavy industrial applications as are lead-acid batteries. At the same time, nickel-cadmium batteries are found in applications requiring portable power such as electronics and power tools. The first nickel-cadmium battery was made by Jungner in Sweden in 1900.

The anode is metallic cadmium. The cathode-active material is nickel oxide present as nickel oxyhydroxide (NiOOH). The active materials of both electrode are incorporated into porous sintered nickel plaques. The electrolyte is an aqueous solution of potassium hydroxide, 31% by weight when the battery is fully charged.

The half-cell reactions at each electrode are shown in reactions (13) and (14). The overall cell reaction is the formation of nickel hydroxide and cadmium hydroxide, shown in reaction (15). The elementary

$$\text{Anode}: Cd + 2OH^- \rightarrow Cd(OH)_2 + 2e^- \quad (13)$$

$$\text{Cathode}: 2NiOOH + 2H^+ + 2e^- \rightarrow 2Ni(OH)_2 \quad (14)$$

$$\text{Overall}: Cd + 2NiOOH + 2H_2O \rightarrow 2Ni(OH)_2 + Cd(OH)_2 \quad (15)$$

electrochemical reactions are shown in reactions (16) and (17). On discharge, electrons are produced

$$\text{Anode}: Cd \rightarrow Cd^{2+} + 2e^- \quad (16)$$

$$\text{Cathode}: 2Ni^{3+} + 2e^- \rightarrow 2Ni^{2+} \quad (17)$$

at the anode as metallic cadmium is oxidized to Cd(II) present in cadmium hydroxide; the complementary reaction at the cathode is the reduction of Ni(III) present in nickel oxyhydroxide to Ni(II) present in nickel hydroxide with the attendant consumption of electrons.

Nominal voltage is 1.2 V. The theoretical specific energy is 220 Wh/kg. Commercially available nickel-cadmium batteries deliver about 40 Wh/kg, although cells with plastic-bonded plates can achieve 55 Wh/kg. Nickel-cadmium batteries are capable of withstanding considerable mechanical abuse. They are also chemically quite stable owing to the fact that none of the cell components is attacked by the electrolyte. The disadvantages of nickel-cadmium technology include low specific energy, and the so-called memory effect in which the cell loses capacity as it seemingly conforms to a duty cycle that does not involve full discharge.

Nickel-metal hydride batteries. The nickel-metal hydride battery, commonly designated NiMH, is a comparatively new technology that has found uses in laptop computers, cellular telephones, and other portable devices where high specific energy is sought. Larger cells have been built for use in electric vehicles.

The electrochemistry shares some features with the nickel-cadmium battery. For example, the cathode active material is nickel oxyhydroxide. In contrast, the anode active material is hydrogen, which is present as the hydride of a proprietary metal alloy. The electrolyte is an aqueous solution of potassium hydroxide, roughly 30% by weight.

The half-cell reactions at each electrode are shown in reactions (18) and (19). The overall cell reaction is

the formation of nickel hydroxide and extraction of the hydrogen storage alloy, shown in reaction (20).

$$\text{Anode}: MH + OH^- \rightarrow M + H_2O + e^- \quad (18)$$

$$\text{Cathode}: NiOOH + H_2O + e^- \rightarrow Ni(OH)_2 + OH^- \quad (19)$$

$$\text{Overall}: MH + NiOOH \rightarrow Ni(OH)_2 + M \quad (20)$$

Here M denotes the hydrogen storage alloy and MH is the hydride compound of same. The elementary electrochemical reactions are shown in reactions (21) and (22). On discharge, electrons are produced at the anode as hydrogen stored in the alloy is

$$\text{Anode}: H \rightarrow H^+ + e^- \quad (21)$$

$$\text{Cathode}: Ni^{3+} + e^- \rightarrow Ni^{2+} \quad (22)$$

oxidized to protons, which combine with hydroxyl ions in the electrolyte to form water; the complementary reaction at the cathode is the reduction of Ni(III) present in nickel oxyhydroxide to Ni(II) present in nickel hydroxide with the attendant consumption of electrons.

Nominal voltage is 1.2 V. The theoretical specific energy is 490 Wh/kg. Commercially available nickel-cadmium batteries deliver about 90 Wh/kg, which is substantially higher than nickel-cadmium or lead-acid batteries. However, materials costs associated with the hydrogen storage alloy are high, and thus the final cost per unit charge ($/kWh) is higher for NiMH than for the other two cited technologies.

Silver-zinc batteries. The silver-zinc battery boasts the highest specific energy and specific power of any secondary battery containing an aqueous electrolyte. However, the high cost of silver has restricted the battery's use to military and aerospace applications primarily.

The anode is metallic zinc. The cathode active material is silver oxide. The electrolyte is an aqueous solution of potassium hydroxide, 45% by weight.

The discharge reaction occurs in two steps, shown in reactions (23) and (24). The overall cell reaction is

$$2AgO + Zn \rightarrow Ag_2O + ZnO \quad (23)$$

$$Ag_2O + Zn \rightarrow 2Ag + ZnO \quad (24)$$

the production of silver by metallothermic reduction of silver oxide by zinc. The elementary electrochemical reactions are shown in reactions (25), (26), and (27). On discharge, electrons are produced at the

$$\text{Anode}: Zn \rightarrow Zn^{2+} + 2e^- \quad (25)$$

$$\text{Cathode}: 2Ag^{2+} + 2e^- \rightarrow 2Ag^+ \quad (26)$$

$$2Ag^+ + 2e^- \rightarrow 2Ag \quad (27)$$

anode as zinc is oxidized to Zn(II) present as zinc oxide; the complementary reactions at the cathode are the reduction of Ag(II) present in silver oxide first to Ag(I) present in Ag_2O, followed by the reduction of Ag(I) to metallic silver, each step accompanied by the attendant consumption of electrons.

Nominal voltage is 1.5 V. The theoretical specific energy is 590 Wh/kg. Commercially available nickel-cadmium batteries deliver about 120 Wh/kg. However, silver-zinc suffers from the high costs associated with silver, loss of cathode active material through dissolution of silver oxide in the electrolyte, and cell shorting by zinc dendrites on charging, all of which lead to poor cycle life. However, when huge currents are needed (and cost is no object), silver-zinc is the battery of choice.

Sodium-sulfur batteries. The sodium-sulfur battery is unique in several ways. Most batteries have solid electrodes and a liquid electrolyte; the sodium-sulfur battery has molten electrodes separated by a solid membrane serving as the electrolyte. In order to keep the electrodes molten and the electrolyte sufficiently conductive, superambient temperatures are required; the sodium-sulfur battery operates at 300°C (572°F). Sodium-sulfur batteries were once prime candidates for electric vehicle power sources and electric utility energy storage for load leveling. However, in recent years, commercial production has ceased.

The electrolyte is a U-shaped tube made of the ceramic material aluminium oxide, specifically β''-Al_2O_3 which is a sodium-ion conductor. The tube is filled with molten metallic sodium which is the anode active material. A metal rod immersed in the sodium serves as current collector. The anode/ electrolyte assembly in positioned inside a larger metal container. The space between the metal container and the ceramic electrolyte tube is occupied by carbon felt filled with elemental sulfur which is the cathode active material. The cathode current collector is the carbon felt which is in electrical contact with the outer metal container.

The half-cell reactions at each electrode are shown in reactions (28) and (29). The overall cell reaction is the formation of sodium sulfide, shown in reaction (30). The elementary electrochemical reactions

$$\text{Anode}: 2Na \rightarrow 2Na^+ + 2e^- \quad (28)$$

$$\text{Cathode}: xS + 2e^- \rightarrow S_x^{2-} \quad (29)$$

$$\text{Overall}: 2Na + xS \rightarrow Na_2S_x \quad (30)$$

$$2.7 < x < 5$$

are identical to the above. The key to the cell's operation is the fact that β''-Al_2O_3 acts as a separator to prevent direct contact of molten sodium and sulfur while allowing only Na^+ to pass from the anode chamber to the cathode chamber.

On discharge, electrons are produced at the anode as metallic sodium is oxidized to Na(I) present in the β''-Al_2O_3 electrolyte as free sodium ions; the complementary reaction at the cathode is the reduction of elemental sulfur to S(II) present as polysulfide ion, S_x^{2-}, which reacts immediately to form sodium sulfide. The sodium ions are delivered to the cathode chamber by the solid electrolyte. Thus, as the cell discharges, the sulfur content of the cathode chamber decreases and the sodium sulfide concentration rises.

Nominal voltage is 1.9 V. The theoretical specific energy is 750 Wh/kg. The last commercially available sodium-sulfur batteries delivered about

80 Wh/kg. While sodium-sulfur batteries have performance characteristics comparable to the best batteries with aqueous electrolytes, thermal management requirements (keeping the cell at temperature) and the indeterminacy of the mode of failure, which in turn lead to safety concerns, have limited the successful deployment of this battery technology.

Zinc-air batteries. The combination of a metal anode and an air electrode results in a battery with an inexhaustible supply of cathode reactant. Recharging involves restoring the anode either electrochemically or mechanically, that is, by direct replacement. In principle, extremely high specific energies should be attainable as the theoretical energy density is 1310 Wh/kg.

The anode is metallic zinc. The cathode active material is oxygen, which is present as a component of air. The current collector is high-surface-area carbon. The electrolyte is an aqueous solution of potassium hydroxide 45% by weight, which may be gelled to prevent leakage.

The half-cell reactions at each electrode electrode are shown in reactions (31) and (32). The overall cell reaction is the oxidation of zinc to form tetrahydroxyzincate [$Zn(OH)_4^{2+}$] or zinc oxide, depending upon the instant composition of the electrolyte, shown in reaction (33). When the concentration of $Zn(OH)_4$

$$\text{Anode}: Zn + 4OH^- \rightarrow Zn(OH)_4^{2-} + 2e^- \quad (31)$$

$$\text{Cathode}: {}^1/_2O_2 + H_2O + 2e^- \rightarrow 2OH^- \quad (32)$$

$$\text{Overall}: Zn + {}^1/_2O_2 + H_2O + 2OH^- \rightarrow Zn(OH)_4^{2-} \quad (33)$$

$^{2-}$ exceeds the solubility limit, precipitation of zinc oxide occurs according to reaction (34). Under these

$$\text{Precipitation}: Zn(OH)_4^{2-} \rightarrow ZnO + H_2O + 2OH^- \quad (34)$$

conditions the overall reaction is effectively the oxidation of zinc according to reaction (35). The ele-

$$\text{Overall}: Zn + 1/_2O_2 \rightarrow ZnO \quad (35)$$

mentary electrochemical reactions are shown in reactions (36) and (37). On discharge, electrons are

$$\text{Anode}: Zn \rightarrow Zn^{2+} + 2e^- \quad (36)$$

$$\text{Cathode}: {}^1/_2O_2 + H_2O + 2e^- \rightarrow 2OH^- \quad (37)$$

produced at the anode as zinc is oxidized to Zn(II) present as $Zn(OH)_4^{2-}$ or zinc oxide; the complementary reaction at the cathode is the reduction of oxygen present in air to O(II) present in hydroxide.

Nominal voltage is 1.5 V. For applications such as notebook computers demonstration cells have been built that deliver 280 Wh/kg. Other attributes include a relatively flat discharge curve, long shelf life thanks to water activation (battery is stored "dry"), and visual inspection of anode condition to determine state of charge. Among the disadvantages is the complexity of a system that comprises a solid anode, liquid electrolyte, and a gaseous cathode. Because the oxygen electrode does not stand up well to operation as an anode (which must happen on charging), a third electrode must be incorporated into the cell. Alternatively, provision must be made for easy disassembly for replacement of a fresh zinc anode and subsequent flawless reassembly of the "recharged" battery.

Lithium-ion batteries. The lithium-ion battery is fast becoming the dominant battery in applications requiring portable power such as electronics and power tools. As such, it is displacing nickel-cadmium and nickel-metal hydride from some of their traditional uses.

The term "lithium-ion" derives from the fact that there is no elemental lithium in the battery; instead, lithium shuttles between hosts, a behavior that has earned this battery the nickname "rocking chair." In the anode, the chemical potential of lithium is high, that is, desirably near that of pure lithium; in the cathode, the chemical potential of lithium is low. The anode is a carbonaceous material (graphite or coke) chosen for its electronic conductivity and its ability to intercalate lithium at potentials near that of pure lithium. An example is LiC_6. In commercial cells the cathode active material is a lithiated transition-metal oxide such as lithium cobalt oxide. Because lithium is more electropositive than hydrogen, the electrolyte must be nonaqueous and aprotic. A representative formulation is a solution (1:1 by volume) of ethylene carbonate and propylene carbonate containing a suitable lithium salt (at a concentration of about 1 *M*) such as lithium hexafluorophosphate, $LiPF_6$, which has been introduced in order to raise the conductivity of the electrolyte. For safety, a separator made of a polyolefin such as microporous polypropylene is placed between the electrodes. If the electrolyte temperature exceeds a certain value, the separator melts and current flow ceases.

The half-cell reactions at each electrode are shown in reactions (38) and (39). The overall cell reaction

$$\text{Anode}: xLi \rightarrow xLi^+ + xe^- \quad (38)$$

$$\text{Cathode}: xLi^+ + xe^- + LiCoO_2 \rightarrow Li_1 + xCoO_2 \quad (39)$$

is the formation of lithium cobalt oxide, shown in reaction (40). The elementary electrochemical reactions are shown in reactions (41) and (42). On dis-

$$\text{Overall}: xLi + LiCoO_2 \rightarrow Li_1 + xCoO_2 \quad (40)$$

$$\text{Anode}: xLi \rightarrow xLi^+ + xe^- \quad (41)$$

$$\text{Cathode}: xCo^{4+} + xe^- \rightarrow xCo^{3+} \quad (42)$$

charge, electrons are produced at the anode as elemental lithium is oxidized to Li(I) present in the electrolyte as free lithium ions; the complementary reaction at the cathode is the reduction of Co(IV) to Co(III) with the attendant consumption of electrons. Note that lithium intercalates into carbon as Li^0, a neutral species, whereas lithium intercalates into $LiCoO_2$ as Li+, a charged species. However, the presence of Li^+ in the cathode triggers a valence shift in the host itself: the compensating electron does not neutralize Li^+ but is localized on Co^{4+}, converting it to Co^{3+}.

Nominal voltage is 3.6 V. The theoretical specific energy of this cell fitted with a $LiCoO_2$ as cathode active material is 770 Wh/kg. Commercially available lithium-ion batteries deliver about 125 Wh/kg.

Lithium-solid polymer electrolyte (SPE) batteries. Since the 1970s it has been known that, upon addition of appropriate salts, polymers can be rendered lithium-ion conductors. Such materials can serve as electrolytes in lithium batteries. When full measure is given to the capacity for miniaturization of a fully solid-state battery, it becomes evident that, in principle, this has the potential to attain the highest specific energy and specific power of any rechargeable technology. In addition, a lithium-SPE battery offers other advantages: ease of manufacture, immunity from leakage, suppression of lithium dendrite formation, elimination of volatile organic liquids, and mechanical flexibility. It is this last attribute that makes lithium-SPE batteries most intriguing: a multilayer laminate of thin films of metal, polymer, and ceramic measuring some tens of micrometers in thickness, totally flexible, with a specific energy in excess of 500 Wh/kg and capable of delivering power at rates in excess of 1000 W/kg at room temperature, operable over a temperature range from -30°C to $+120^\circ$C, and costing less than \$500/kWh. Realization awaits advances in research in materials science and battery engineering.

Outlook. What is in store for secondary power sources, remains speculative. Perhaps higher-capacity batteries for automobile traction will be developed. There may be a push for thin-film, all-solid-state microbatteries, which will enable distributed power sources in combination with their own charger and electrical appliance or load. An example of such integration might be an electrical device powered by a secondary battery connected to a photovoltaic charger. Donald R. Sadoway

Bibliography. J. O. Besenhard (ed.), *Handbook of Battery Materials*, Wiley-VCH, Weinheim, 1999; T. R. Crompton, *Battery Reference Book*, 2d ed., 1996; D. Linden, *Handbook of Batteries*, 2d ed., 1995; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., Wiley, New York, 1992; *Ullmann's Encyclopedia of Industrial Chemistry*, 5th ed., VCH, Weinheim, 1985.

Bauxite

A rock mainly comprising minerals that are hydrous aluminum oxides. These minerals are gibbsite, boehmite, and diaspore. The major impurities in bauxite are clay minerals and iron oxides. Bauxite is a weathering product of aluminous rock that results from intense leaching in tropical and subtropical areas, a process called laterization. Bauxite deposits are generally found on plateaus in stable areas where they had sufficient geologic time to form and were protected from erosion. *See* ALUMINUM.

Bauxite is the primary ore of aluminum. The two types of bauxites that are used commercially as aluminum ores are laterite and karst. Lateritic bauxites constitute more than three-fourths of the world's bauxite resources. Karstic bauxites are formed on a carbonate terrain and are concentrated in sinkholes and solution depressions on the surface of carbonate rocks. *See* LATERITE.

Commercial sources. Bauxite is mined and processed in many tropical and subtropical areas. Lateritic deposits are mined mainly in Australia, Brazil, West Africa, India, China, Surinam, Guyana, and Indonesia. The largest producer is Australia, from which about one-third of the world's bauxite production, about 125 million metric tons annually, is mined. Karstic bauxites are mined in Jamaica, Dominican Republic, Haiti, Hungary, Greece, Yugoslavia, Rumania, and Turkey. All the mines are open pit, and the benefication involves washing and screening in order to remove some of the clay and iron minerals. The beneficiated bauxite is transported to locations where alumina is produced from the bauxite. Four to seven tons of bauxite, depending on the available alumina content, are required to produce 2 tons of alumina, and approximately 2 tons of alumina are used to produce 1 ton of aluminum.

Uses. Bauxite used to produce alumina is called metallurgical grade; approximately 85% of the world's production is for this purpose. Other major uses are in refractories, abrasives, chemicals, and aluminous cements. The compositional requirements are much more rigid for these uses. The alumina content must be higher, and the iron, silica, and titanium contents significantly lower, than for metallurgical-grade bauxite. World resources of bauxite are many tens of billions of tons, so an adequate supply is assured for hundreds of years. However, since no metallurgical-grade bauxite is being mined in the United States, the nation is completely dependent on foreign sources. Haydn H. Murray

Bibliography. G. Bardossy, *Karst Bauxites*, 1982; P. W. Harben, *The Industrial Minerals Handybook*, 1998; L. Jacob Jr. (ed.), *Bauxite*, 1984.

Bayesian statistics

An approach to statistics in which estimates are based on a synthesis of a prior distribution and current sample data. Bayesian statistics is not a branch of statistics in the way that, say, nonparametric statistics is. It is, in fact, a self-contained paradigm providing tools and techniques for all statistical problems. In the classical frequentist viewpoint of statistical theory, a statistical procedure is judged by averaging its performance over all possible data. However, the bayesian approach gives prime importance to how a given procedure performs for the actual data observed in a given situation. Further, in contrast to the classical procedures, the bayesian procedures formally utilize information available from sources other than the statistical investigation. Such information, available through expert judgment, past experience, or prior belief, is described by a probability distribution on the set of all possible values of the unknown parameter of the statistical model at hand.

This probability distribution is called the prior distribution. The crux of the bayesian approach is the synthesis of the prior distribution and the current sample data into a posterior probability distribution from which all decisions and inferences are made. This synthesis is achieved by using a theorem proved by Thomas Bayes in the eighteenth century.

Example. The basic principles of bayesian statistics may be understood by considering an example in which a quality control inspector regularly receives lots of a manufactured product, each lot consisting of 10 units. In the past, 70% of the lots have had no defective units, 20% have had one defective unit, and 10% have had two defective units. The quality control inspector has just received a lot of 10 units, but has no time for 100% inspection and must estimate the quality of the lot by testing five units randomly selected. The inspector finds one defective unit among the five tested, and must make an estimate of θ, the proportion of defective units in the lot. If the inspector ignores past experience, a reasonable estimate of θ is 1/5. However, the past experience about θ can be incorporated into the estimation process by basing the estimate on the posterior distribution of θ given the observed data x, namely, $\pi(\theta|x)$ defined by Eq. (1), where α indicates "is proportional to,"

$$\pi(\theta|x) \propto \pi(\theta)f(x|\theta) \tag{1}$$

$\pi(\theta)$ describes the prior belief in the value θ, and $f(x|\theta)$ assesses the chances of the actual observed x when θ is the actual value of the parameter. In the present case, $x = 1$ (the observed number of defectives), $\pi(\theta) = 0{\cdot}70$ if $\theta = 0$, $\pi(\theta) = 0{\cdot}20$ if $\theta = 1/10$, and $\pi(\theta) = 0{\cdot}10$ if $\theta = 2/10$. Furthermore, it can be argued that the chance of there being one defective unit among the five tested when θ is the proportion of defective units in the lot is given by Eq. (2), where

$$f(1|\theta) = \frac{\begin{bmatrix} 10\,\theta \\ 1 \end{bmatrix}\begin{bmatrix} 10\,(1-\theta) \\ 4 \end{bmatrix}}{\begin{bmatrix} 10 \\ 5 \end{bmatrix}} \tag{2}$$

the brackets indicate binomial coefficients. It follows from Eqs. (1) and (2), and the fact that the sum of the posterior probabilities must equal unity, that $\pi(\theta|1)$ is given by $\pi(\theta|1) = 0$ for $\theta = 0$, $\pi(\theta|1) = 9/14$ for $\theta = 1/10$, and $\pi(\theta|1) = 5/14$ for $\theta = 2/10$. (In this case, the factor of proportionality in Eq. (1) equals 90/14.) According to the bayesian approach, it is the posterior distribution $\pi(\theta|1)$ which should form the basis for estimating θ. *See* BINOMIAL THEOREM.

The posterior distribution combines the prior information about θ with the information contained in the observed data to give a composite picture of the final judgments about θ. In order to arrive at a single number as the estimate of θ, it may be necessary to bring in the notion of the loss suffered by the decision maker as a result of estimating the true value θ by the number $\hat{\theta}$. Depending on the choice of the loss function, the mode, the mean, and the median of the posterior distribution $\pi(\theta|x)$ are all reasonable estimates of θ.

Conjugate priors. The choice of the prior distribution for the unknown parameter θ is of crucial importance in bayesian statistics. The selected prior distribution for θ must be at least a reasonable approximation to the true beliefs about θ. In addition, the prior distribution $\pi(\theta)$ must be such that the posterior distribution $\pi(\theta|x)$ is tractable. Frequently, conjugate priors provide the answer. If (1) $\Im$ is the class of probability density functions $f(x|\theta)$ [indexed by θ] chosen to model the data-generating process, and (2) $\wp$ is a class of prior probability density functions $\pi(\theta)$ with the property that the posterior probability density function $\pi(\theta|x)$ belongs to the class $\wp$ whenever f belongs to $\Im$ and π belongs to $\wp$, then $\wp$ is said to be a family of conjugate priors for $\Im$. Thus, a conjugate family is closed under sampling in the sense that if the sample data are used to update a prior distribution in a conjugate family by Bayes' theorem the resulting posterior distribution will also belong to the same family. A prior density $\pi(\theta)$ is called a natural conjugate to the likelihood function $L(\theta|x) = f(x|\theta)$ if $\pi(\theta)$ and $L(\theta|x)$ are proportional as functions of θ.

A simple example of the usefulness of conjugate priors is provided by random sampling from a normal distribution. If $x = (x_1, \ldots, x_n)$ is a random sample from a normal distribution with unknown mean θ and known standard deviation σ, and if the prior distribution of θ is normal with mean μ and standard deviation λ, then the posterior distribution of θ given x is normal with mean $\alpha\mu + (1-\alpha)\bar{x}$ and standard deviation $\lambda\alpha^{1/2}$, where α and $\bar{x}$ are given by Eqs. (3) and (4). Thus, the family of normal priors is a conjugate family for the normal distribution with unknown mean and known standard deviation. It updates the prior parameters μ and λ to $\alpha\mu + (1-\alpha)\bar{x}$ and $\lambda\alpha^{1/2}$, respectively, thereby making transparent the effect of prior and sample information. *See* DISTRIBUTION (PROBABILITY).

$$\alpha = \frac{\sigma^2}{\sigma^2 + n\lambda^2} \tag{3}$$

$$\bar{x} = \frac{x_1 + \cdots + x_n}{n} \tag{4}$$

Noninformative and improper priors. The bayesian approach remains applicable even when little or no prior information is available. Such situations can be handled by choosing a prior density giving equal weight to all possible values of θ. Priors that seemingly impart no prior preference, the so-called noninformative priors, also arise when the prior is required to be invariant under certain transformations. Frequently, the desire to treat all possible values of θ equitably leads to priors with infinite mass. Such noninformative priors are called improper priors.

An example of the usefulness of improper priors is that of a random quantity whose pattern of variation is describable by a normal distribution with unknown positive mean θ and unit standard deviation. If no specific prior knowledge is available, then the improper prior given by $\pi(\theta) = 1$ for all $\theta > 0$ may be used. Although this prior has infinite mass,

it results in a perfectly legitimate posterior density given by Eq. (5). If the observed x equals zero and if

$$\pi(\theta|x) \propto \exp\left\{\frac{-(\theta - x)^2}{2}\right\} \qquad \theta > 0 \qquad (5)$$

the mean of the posterior distribution is used to estimate θ, then the estimate of θ can be easily shown to be $(2/\pi)^{1/2}$. The classical estimate of θ is zero, which is unsuitable because θ is given to be positive.

Bayesian estimation. Optimal estimation of θ by a single point $\hat{\theta}$ requires specification of $l(\hat{\theta}, \theta)$, the loss suffered when the actual value θ is estimated by $\hat{\theta}$. If θ is unidimensional and $l(\hat{\theta}, \theta)$ is the squared error loss function $(\hat{\theta} - \theta)^2$, then the mean of the posterior distribution is the optimal point estimate of θ. If $l(\hat{\theta}, \theta)$ is the absolute error loss function $|\hat{\theta} - \theta|$, then the median of the posterior distribution is used as the estimate of θ. Point estimates of θ are usually accompanied by a measure of the accuracy of the estimate. When θ is unidimensional and the loss function is the squared error, the measure of accuracy of the Bayes estimate of θ is the variance of the posterior distribution. In situations involving noninformative priors, the Bayes estimates and their measures of accuracy are sometimes numerically very similar to the corresponding non-bayesian quantities.

If it is desired to estimate θ by a range of values in a parametric space Ω rather than by a single point $\hat{\theta}$, the bayesian approach is to select a subset C of Ω such that the probability assigned to C by the posterior distribution is at least a prespecified number $1-\alpha$. Such a subset C is called a $100(1-\alpha)\%$ credible set for θ. It is the bayesian analog of a classical confidence set. Bayesian credible sets are generally easier to find than classical confidence sets. In bayesian practice, the credible set selected is the highest posterior density (HPD) credible set C defined to consist of all values of θ such that $\pi(\theta|x)$ is equal to or greater than $\kappa(\alpha)$, where $\kappa(\alpha)$ is the largest constant for which C is a $100(1-\alpha)\%$ credible set. Often, the posterior distribution can be approximated by a normal distribution which yields an approximate HPD credible set without any computational difficulties.

Bayesian hypothesis testing. When concerned with the problem of deciding between two hypotheses, H_0, stating that θ is a member of Ω_0, and H_1, stating that θ is a member of Ω_1, the bayesian approach is to simply calculate the posterior probabilities $\alpha_i = \text{Prob}(\Omega_i|x)$, $i = 0,1$, and choose between H_0 and H_1 accordingly. This approach is in sharp contrast to the classical non-bayesian approach in which a test procedure is evaluated in terms of the so-called type I and type II error probabilities. The bayesian approach to multiple hypothesis testing is equally simple, namely, to calculate the posterior probability of each hypothesis and decide in favor of the hypothesis having the largest posterior probability.

The posterior probability of a hypothesis H_0, stating that θ is a member of Ω_0, is zero when Ω_0 is a singleton and the prior distribution is a continuous distribution. It would appear that the bayesian approach would not work for testing such hypotheses. However, the problem can be resolved by assigning a positive probability π_0 to the simple hypothesis $H_0 : \theta = \theta_0$ and a density $(1 - \pi_0)\pi(\theta)$ to $\theta \neq \theta_0$. Here $\pi(\theta)$ is a proper prior for θ, and π_0 is chosen as the probability mass that would be assigned to the hypothesis $|\theta - \theta_0| < \delta$ for some positive δ determined by the practical problem at hand. The rationale for this approach is the fact that the hypothesis $\theta = \theta_0$ is rarely reasonable and that, for suitably chosen $\delta > 0$, all θ in $(\theta_0 - \delta, \theta_0 + \delta)$ are of equal practical importance.

Bayesian tests of hypotheses are frequently reported in terms of Bayes factors. The Bayes factor in favor of a hypothesis H_0 is defined by Eq. (6). A

$$B = \frac{\text{posterior odds ratio of } H_0 \text{ to } H_1}{\text{prior odds ratio of } H_0 \text{ to } H_1} \qquad (6)$$

person can determine his or her personal posterior odds by simply multiplying the reported Bayes factor by his or her personal prior odds.

Criticism. Proponents of the bayesian paradigm claim a number of significant and fairly convincing advantages over the classical non-bayesian approach. However, bayesian statistics is itself criticized for its apparent lack of objectivity. The use of a prior distribution introduces subjectivity into a bayesian analysis so that decision makers having different priors may come to different conclusions even when working on the same problem with the same observed data. This, however, is the price that must be paid in order to incorporate prior information or expert judgment. Moreover, the bayesian approach seems to be the most logical approach if the performance of a statistical procedure for the actual observed data is of interest, rather than its average performance over all possible data. *See* ESTIMATION THEORY; PROBABILITY; STATISTICS.

Syed N. U. A. Kirmani

Bibliography. J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, 3d ed., 1993; J. M. Bernardo and A. F. M. Smith, *Bayesian Theory*, 1994; G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*, 1973, reprint 1992; P. M. Lee, *Bayesian Statistics: An Introduction*, 2d ed., 1997; D. V. Lindley, *Bayesian Statistics: A Review*, 1972, reprint 1984; J. S. Press, *Bayesian Statistics: Principles, Models, and Applications*, 1989.

Bdelloidea

A class of the phylum Rotifera characterized by paired gonads and obligately asexual reproduction, found in diverse freshwater habitats worldwide. Bdelloid rotifers are among the most common microinvertebrates, particularly in ephemerally aquatic environments such as rain gutters, bird baths, moss, lichen, and detritus. First observed by A. Leeuwenhoek more than 300 years ago, bdelloids have long been a favorite subject of amateur microscopists and

can easily be isolated and observed with a low-power microscope.

Morphology. Bdelloids are small; adults of most species are between 0.1 and 0.3 mm long. Despite their minute size, they contain about 1000 nuclei (cell membranes are ill-defined and many cells are syncytial). The body contains a nervous system, complex musculature, and complete digestive system, along with the corona (anterior ring of cilia), mastax (muscular pharynx), toes, and syncytial intracytoplasmic lamina (a specialized integument) characteristic of rotifers (**Fig. 1**). Bdelloids propel themselves through water by rhythmic beating of their coronal cilia. When attached to a substrate by their toes, they may withdraw their corona and extend an apical structure called the rostrum. Bdelloids crawl in a leechlike fashion by extending the rostrum. *See* ROTIFERA.

Classification. Although the systematics of Rotifera and allied phyla have been undergoing revision, Bdelloidea is usually afforded the class status, and most morphological and molecular evidence indicates that it is most closely related to the rotifer group Monogononta. There are 370 described species classified in four easily distinguished families. Philodinidae and Habrotrochidae are the most species-rich families and contain the most common, morphologically typical species. Members of the family Adinetidae rarely swim and have a corona and mastax that are specialized for scraping food particles off the substrate. Philodinavidae is a poorly explored family that includes the only known predatory bdelloids. *See* MONOGONONTA.

Ecology and physiology. Bdelloids can be found in almost any freshwater ecosystem at temperatures ranging from 32 to 113°F (0 to 45°C) and pH levels from 2.5 to 10, and are important components of benthic and littoral ecosystems. Most eat bacteria, fungi, algae, and detritus, and are in turn eaten by cladocerans, copepods, insect larvae, and fish fry (newborn fish). Some species (particularly *Embata*) are epizoic, living on the bodies of freshwater arthropods. Bdelloids are particularly common in environments that are only temporarily hydrated and are capable of surviving desiccation at any stage of their life: as water evaporates from the local environment, a bdelloid exudes all free water, contracts into a form called a tun, and enters a metabolically quiescent state called anhydrobiosis (**Fig. 2**). In the anhydrobiotic state bdelloids can survive vacuum, high levels of radiation, and temperatures from −454 to 392°F (−270 to 200°C). When rehydrated, bdelloids resume activity with little or no decrease in lifespan or fecundity. There are apocryphal accounts of bdelloids revived after years of anhydrobioisis.

Reproduction. Males and hermaphrodites are unknown in Bdelloidea, and reproduction is parthenogenetic: females produce embryos from oocytes after two mitotic divisions without chromosome pairing or reduction (ameiotic, or apomictic, thelytoky). Bdelloid rotifers are the largest, oldest, most successful animal group for which observational, cytological, and molecular evidence are consistent with the long-term asexuality of the group. They are a model system for the study of the paradox of the maintenance of sexual reproduction.

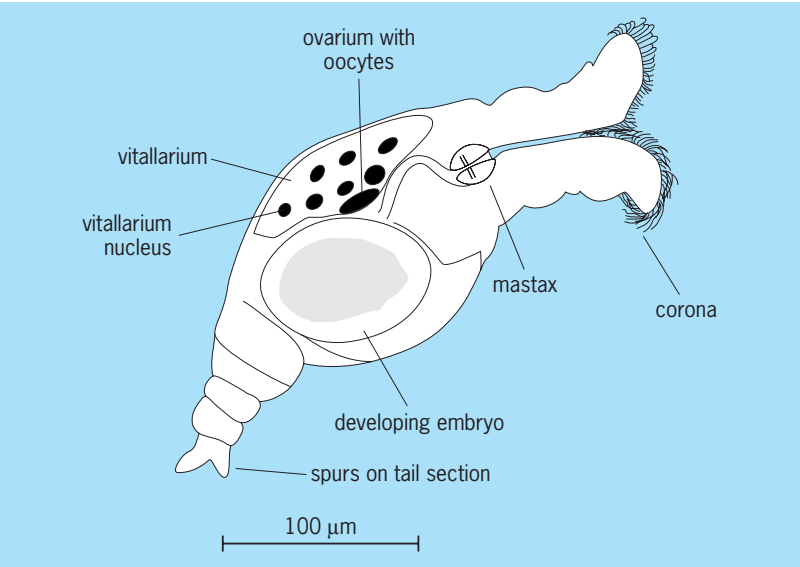

Fig. 1. Bdelloid rotifer *Macrotrachela quadricornifera*. The vitallarium provides nutrients to the developing egg.

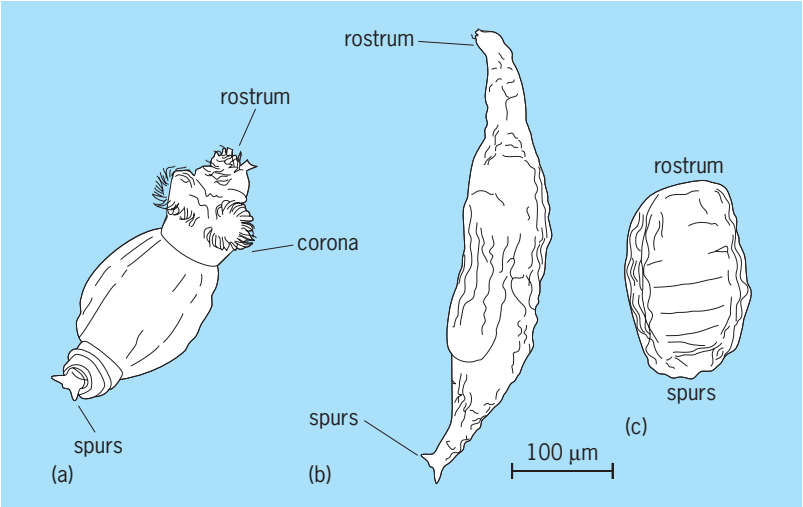

Fig. 2. Bdelloid rotifer *M. quadricornifera* contracts into an organized, compact tun as it undergoes desiccation and enters anhydrobiosis. (*a*) Active adult. (*b*) Adult contracting and entering anhydrobiosis. (*c*) Anhydrobiotic tun.

David Mark Welch

Bibliography. J. Donner, *Ordnung Bdelloidea*, Akademie-Verlag, Berlin, 1965; J. Donner, *Rotifers*, Frederick Warne, London, 1966; L. H. Hyman, *The Invertebrates*, McGraw Hill, New York, 1951; J. Maynard Smith, Contemplating life without sex, *Nature*, 324:300–301, 1986; R. L. Wallace and T. W. Snell, Rotifera, pp. 187–248, in *Ecology and Classification of North American Freshwater Invertebrates*, ed. by J. H. Thorp and A. P. Covich, Academic, New York, 2001.

Bdellonemertini

An order of the class Enopla in the phylum Rhynchocoela. The proboscis is unarmed. The alimentary system comprises the mouth, papillate foregut, sinuous intestine without diverticula, and anus. Members are characterized by the presence of a posterior adhesive disk and dorsoventral flattening of the body, which lacks eyes, cephalic slits, and cerebral organs. The order contains the single genus *Malacobdella*, with three species ectocommensal in the mantle cavity of marine bivalves and one in the pulmonary sac of a fresh-water gastropod. *Malacobdella grossa*, in the bivalve *Zirfaea crispata*, abstracts particulate food such as planktonic organisms from the host's feeding current, collects these by a unique filtration system, and digests them in the intestine by a combination of extracellular and intracellular processes. Carbohydrases predominate in the digestive physiology, in contrast to other, carnivorous rhynchocoelans, where proteases are dominant. *See* ANOPLA; ENOPLA; HOPLONEMERTINI; RHYNCHOCOELA. J. B. Jennings

Beam

A structural member that is fabricated from metal, reinforced or prestressed concrete, wood, fiber- reinforced plastic, or other construction materials and that resists loads perpendicular to its longitudinal axis. Its length is usually much larger than its depth or width. Usually beams are of symmetric cross section; they are designed to bend in this plane of symmetry, which is also the plane of their greatest strength and stiffness. This plane coincides with the plane of the applied loads. Beams are used as primary load-carrying members in bridges and buildings.

Internal forces. The internal resisting forces developed in the plane of the cross section are known as shear forces, and those normal to the cross section are called bending moments. **Figure 1***a* shows a steel beam cross section subjected to a bending moment. The resulting deformation is a linear rotation of the plane of the cross section from its original vertical position. The distance between a point on the original plane and the same point on the deformed plane is the strain (displacement/unit length). The inclination of the plane is the curvature of the beam. The corresponding stress (force/unit area) is related to the strain via the stress-strain relationship that is obtained by experiment for the material from which the beam is fabricated (Fig. 1*b*). Three stress distributions are shown in Fig. 1*a*: when the cross section is elastic, part of it has yielded, and the whole section is yielded. Yielding occurs in portion DBC in Fig. 1*b*. The relationship between the curvature and the internal bending moment is obtained by multiplying the stress at any point with its distance to the centroid of the cross section and integrating over the whole cross section. The resulting moment-curvature relationship (Fig. 1*c*) is nonlinear, and it attains its maximum at the plastic moment level. Under many practical conditions this is the capacity that can be utilized in design (point C). Sometimes it is necessary to curtail deformation to the elastic region, and then the maximum usable moment is the yield moment (point D). The moment-curvature relations for different materials and cross sections are obtainable in a similar manner. *See* SHEAR; STRESS AND STRAIN.

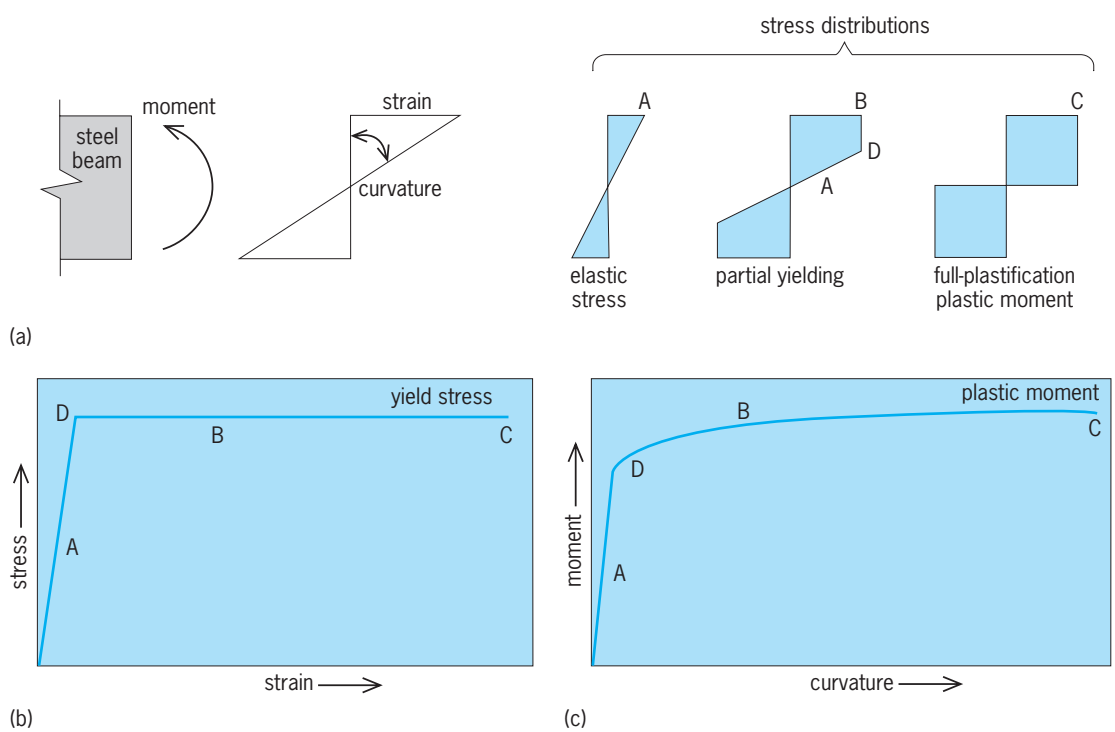

Fig. 1. Forces on a cross section of a beam. (*a*) Bending moment. (*b*) Stress-strain relationship. (*c*) Moment-curvature relationship.

-s

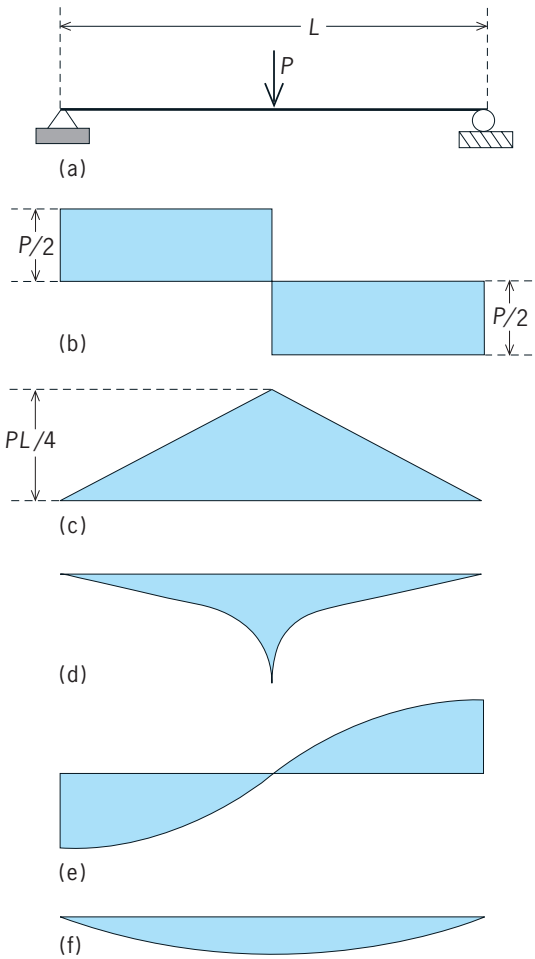

Fig. 2. Shear, moment, and deformation diagrams. (*a*) Loading. (*b*) Shear force. (*c*) Bending moment. (*d*) Curvature. (*e*) Slope. (*f*) Deflection.

Applied loads and resulting forces and deformations. In a simple beam of length L, a concentrated load P is applied at the center of the beam (**Fig. 2***a*). Such a beam has a support at one end that cannot move downward or sideways, but the beam end can freely rotate about the point of support. At the other end the beam may roll sideways also. There can be variations of the planar forces, or shears, along the length of the beam (Fig. 2*b*), or a bending moment distribution (Fig. 2*c*). The force diagrams are obtained by using the concepts of static equilibrium. Knowing the bending moment, the corresponding curvature can be obtained from the moment-curvature curve (Fig. 1*c*). The slope of the deflected beam at any location along its length (Fig. 2*e*) can then be determined from the curvature diagram by integration, since the curvature is the change in slope at any point along the beam. Similarly, the deflection (Fig. 2*f*) can be obtained by integrating the slope, since it is the change in deflection per unit length. These relationships can be found by integration downward or differentiation upward. Other load configurations can be handled in a similar manner. For beams that are not simply supported (clamped ends or multispan beams), additional conditions of compatibility of slopes and deflections at the supports must be introduced to obtain a solution. *See* LOADS, TRANSVERSE; STATICS; STRUCTURAL ANALYSIS.

Limiting conditions. Once the forces and deformations of the beam are known, they can be compared to the corresponding capacities of the beam. These are usually the shear strength, the bending strength, or the buckling strength. If these are close to the values that are imposed by the loading, the design is satisfactory. If the demand is much below the capacity or if the capacity is exceeded, a new beam must be tried and checked. It is most economical when the plastic moment can be used as the capacity. This can be achieved, in the case of metal or composite beams, if they are braced to prevent lateral buckling, or if stiffeners are provided to prevent local buckling of the plate components. The preferred failure of concrete beams is yielding of the reinforcing steel. To account for unavoidable uncertainties in the loads and capacities, load factors (larger than unity) and resistance factors (less than unity) are utilized.

Theodore V. Galambos

Bibliography. American Institute of Steel Construction, *Manual of Steel Construction: Load and Resistance Factor Design*, 2d ed., 1994; T. V. Galambos, F. J. Lin, and B. G. Johnston, *Basic Steel Design with LRFD*, Prentice Hall, 1996; R. C. Hibbeler, *Structural Analysis*, 4th ed., Prentice Hall, 1999; J. G. MacGregor, *Reinforced Concrete: Mechanics and Design*, 3d ed., Prentice Hall, 1997.

Beam column

A structural member that is subjected to axial compression and transverse bending at the same time. The combined compression and bending may be produced by an eccentrically applied axial load (**Fig. 1***a*), a concentrically applied axial load and an end moment (Fig. 1*b*), or an axial load and transverse load between the two ends. (Fig. 1*c*). A beam column differs from a column only by the presence of the

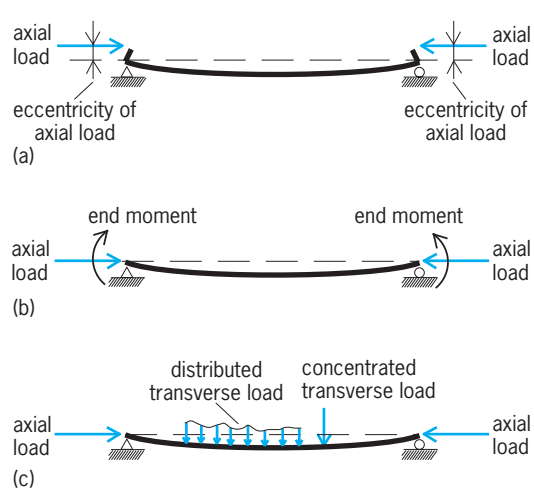

Fig. 1. Loadings that create beam-column effects. (*a*) Eccentrically applied axial load. (*b*) Concentrically applied axial load and end moment. (*c*) Axial load and transverse load or loads between the two ends.

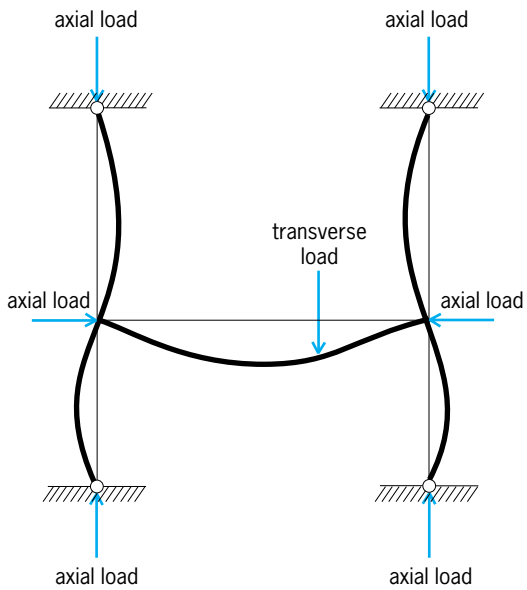

Fig. 2. Frame-type structure with one beam and four columns acting as beam columns.

eccentricity of the load application, end moment, or transverse load.

Beam columns are found in frame-type structures where the columns are subjected to other than pure concentric axial loads and axial deformations, and where the beams are subjected to axial loads in addition to transverse loads and flexural deformations (**Fig. 2**).

The other-than-axial loads create bending of the column immediately upon their application. Because of this initial bending and the consequent flexural stresses in addition to the pure axial stresses, a beam column fails at a lower axial force than a column of identical dimensions and end-support conditions.

The mathematical formulas representing bending deformation and variation of stress along the length of a beam column are generally more complex than those for a column. Structural design codes and specifications present simplified methods for the design of members under combined compression and bending. These methods take the forms of interaction formulas and interaction curves that limit the magnitudes of axial compression and transverse bending when the two are acting in tandem.

One form of the interaction formulas, which assures that the beam column's strength under axial load and bending is not exceeded, is given by the expression below, where P_u = ultimate axial load without the presence of bending, M_u = ultimate bending moment without the presence of the axial load, P = actual axial load when applied together with the bending moment, and M = actual bending moment when applied together with the axial load. The expressions for P_u and M_u can be quite complex, depending on the material, cross-sectional shape, slenderness, and types of end supports of the beam column. *See* BEAM; COLUMN; STRESS AND STRAIN; STRUCTURAL DESIGN.

$$\frac{P}{P_u} + \frac{M}{M_u} \leq 1$$

Robert T. Ratay

Bibliography. E. H. Gaylord and C. N. Gaylord, *Design of Steel Structures*, 3d ed., 1992; A. H. Nilson and G. Winter, *Design of Concrete Structures*, 12th ed., 1997.

Beam-foil spectroscopy

A technique used in atomic physics to study the structure and dynamics of atomic ions of any element in any state of ionization. For this purpose, a beam of fast ions is sent through a very thin foil. The ion-foil interaction shakes up the electronic shells of the projectile and, after leaving the foil, the ions shed surplus energy by emitting photons, and sometimes electrons. The energies and intensities of these particles yield spectral information on the projectile.

Beam-foil spectroscopy is a late offspring of the canal-ray studies of the early twentieth century. The technique became possible only after the advent of accelerators. Such machines were originally developed to produce and accelerate charged atomic particles as projectiles in nuclear reactions. The target material is usually prepared as a thin foil—typically a few thousandths of the thickness of writing paper (a few hundred atomic layers thick)—through which the ions pass without significant energy loss. Most of the projectiles, however, do not undergo nuclear reactions, but collide only with the many target electrons (and therefore largely maintain their original speed and direction) and suffer a shake-up of their own electron shells. Excited states can be studied by monitoring the photons and electrons that the fast projectiles emit in deexcitation after their passage through the foil. *See* NUCLEAR REACTION; PARTICLE ACCELERATOR.

Exit beam composition. The multitude of collisions inside the foil changes the complement of electrons that travel with the projectile ion; some are ejected and others are captured from the target atoms. The ion beam therefore has a different charge-state composition after passage through the foil. (Higher exit charge states are produced at higher incident beam energies.) The beam-foil interaction efficiently populates atomic levels with a high degree of excitation such as multiply excited and core-excited levels. Such levels are rare in conventional light sources. The richness of the resulting spectra yields a great deal of information on atoms and ions, but it also causes problems. For example, it is often difficult to resolve the details of such line-rich spectra that reflect the complexity of multiply excited systems.

Timing measurements. The ion beam travels in a high vacuum before and after transiting the target foil. This environment minimizes collisional perturbation of the ions. The sudden termination of the ion-foil interaction provides an inherently good time resolution to beam-foil spectroscopy. This property of the source permits measurements of lifetimes of

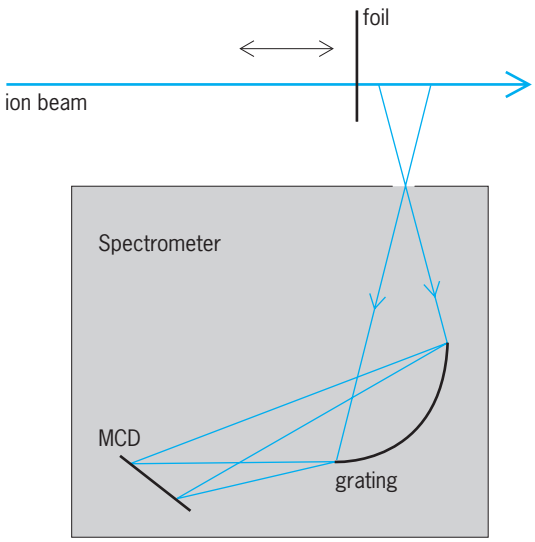

Schematic of a beam-foil spectrometer equipped with a multichannel detector (MCD). Light emitted by the ions after traversing the foil is incident on the grating and MCD. The spectrometer permits simultaneous spectral coverage of a selected wavelength range.

atomic levels populated during the beam-foil interaction as well as the observation of coherent-excitation phenomena such as quantum beats. Because the ion velocity is constant and measurable, it is sufficient to trace the change in intensity of the fluorescence from the ion beam as a function of distance from the foil in order to determine atomic level lifetimes (see **illus.**). At an ion energy of 0.5 MeV/nucleon, the ions travel 1 cm in 1 nanosecond (10^{-9} s). An easily measured mechanical displacement of 10 micrometers then corresponds to a time interval of only 1 picosecond (10^{-12} s). This "magnification" allows short time intervals to be measured without any need for fast electronic timing. The typical size of a high-vacuum experimental chamber sets a practical upper limit of about 100 ns. An extension of the beam-foil technique, however, uses a heavy-ion storage ring to measure atomic lifetimes in the millisecond range. *See* FLUORESCENCE.

Intensity decay curves are single-exponential if a single level is excited. A laser, which is highly selective, can excite a single level. Laser excitation is, however, not generally applicable to highly charged ions, since excitation energies are large. For many levels within the higher ionization stages, nonselective beam-foil excitation is one of the few methods available. In nonselective excitation, often a level of interest may be repopulated by the decay from one or more higher-lying excited levels that have also been populated. However, if all of these repopulating cascades are measured as well, the time constant of the decay of the level of interest can be obtained with precision (1% or better) by a correlated analysis. If the cascade information is incomplete, the precision of the desired atomic lifetime can be 10% or more. *See* LASER.

Metastable states. The properties of highly charged ions can be predicted by scaling the known properties of low-charge ions with the same number of electrons (that is, members of the same isoelectronic sequence). With increasing ion charge, the wavelengths of transitions and lifetimes of excited states become shorter. In most highly charged ions the level lifetimes become so short that the postfoil fluorescence cannot be measured with sufficient time resolution. An exception is the decay of the much longer lived metastable states. For example, transitions involving a change of spin or higher-order multipole radiation only decay at a much slower rate. These unusual transitions, however, are of fundamental interest in understanding atomic structure, and are often used in the diagnostics of terrestrial plasmas, such as those in fusion research, or in astrophysical observations. For example, the ultraviolet part of the spectrum of the solar corona shows many lines that, for a long time, had not been observed on Earth. In sources such as laser-produced and tokamak plasmas, where they might be expected to appear, they were probably collisionally quenched. A number of such lines were eventually identified, however, by beam-foil spectroscopy. The spectra were studied at a distance of a few centimeters from the foil (or after delay times of a few nanoseconds). After such a delay, the fast decays have died out and the weak decays associated with long-lived or metastable levels become apparent. *See* ISOELECTRONIC SEQUENCE; METASTABLE STATE; PLASMA (PHYSICS); PLASMA DIAGNOSTICS; SOLAR CORONA; SUN.

Fast-beam spectroscopy. Beam-foil spectroscopy has developed into many variants which now go under the name of fast-beam spectroscopy. For example, a gas target may be used, a laser, a combination of gas or foil and laser, or a target of free electrons in a heavy-ion storage ring. The ion-foil interaction is capable of producing all ionization stages of all elements from negative ions to U^{91+}. The production of the highest ionization stages, however, requires a beam energy of about 500 MeV/nucleon, which can be reached only at the most energetic accelerators. However, since only the relative motion of electrons and ions is important, the same degree of ionization can nowadays be reached by use of 250-keV electrons in an electron-beam ion trap (EBIT). The device offers easier ways to attain high spectroscopic precision because the ions are practically at rest. In beam-foil spectroscopy the ions are rapidly moving, which shifts and broadens the spectral lines. This, in turn, causes problems in wavelength calibration and spectral resolution. However, the inherent time resolution of the foil-excited fast-ion-beam source is unique and remains a great asset in time-resolved spectroscopic measurements. *See* ATOMIC STRUCTURE AND SPECTRA; ION SOURCES; SPECTROSCOPY.

Elmar Träbert

Bibliography. G. W. F. Drake (ed.), *Atomic, Molecular, and Optical Physics Reference Book*, AIP Press, Woodbury, 1996; S. M. Shafroth and J. C. Austin (eds.), *Accelerator-Based Atomic Physics: Techniques and Applications*, AIP Press, Washington, DC, 1997.

Bean

Any of several leguminous plants, or their seeds, long utilized as food by humans or livestock. Some 14 genera of the legume family contain species producing seeds termed beans and useful to humans. Twenty-eight species in 7 genera produce beans of commercial importance, implying that the bean can be found in trade at the village level or up to and including transoceanic commerce.

Varieties. The principal Asiatic beans include the edible soybeans, *Glycine* sp., and several species of the genus *Vigna*, such as the cowpea (*V. unguiculata*) and mung, gram, rice, and adzuki beans. The broad bean (*Vicia faba*) is found in Europe, the Middle East, and Mediterranean region, including the North African fringe. Farther south in Africa occur *Phaseolus* beans, of the *vulgaris* (common bean) and *coccineus* (scarlet runner) species. Some *Phaseolus* beans occur in Europe also. The cowpea, used as a dry bean, is also found abundantly in Nigeria. *See* COWPEA; SOYBEAN.

In the Americas, *P. vulgaris* and *P. lunatus* (lima bean) are the principal edible beans, although the blackeye cowpea, mung bean, and chick pea or garbanzo (*Cicer arietinum*) are grown to some extent. *Phaseolus coccineus* is often grown in higher elevations in Central and South America, as is *Vicia faba*. The tepary bean (*P. acutifolius*) is found in the drier southwestern United States and northern Mexico. *See* ROSALES.

Characteristics. Bean plants may be either bush or vining types, early or late in maturity, with white, yellow, red, or purple flowers. The seed itself is the most differentiating characteristic of bean plants. It may be white, yellow, black, red, tan, cream-colored, or mottled, and range in weight from 0.0045 to over 0.025 oz (125 to over 700 mg) per seed. Seeds are grown in straight or curved pods (fruit), with 2–3 seeds per pod in *Glycine* to 18–20 in some *Vigna*. Pods vary in fiber content; the immature pods of low-fiber types are often used as a green vegetable. Beans are normally self-pollinated, although a small amount of natural crossing may occur. Scarlet runner bean is normally highly cross-pollinated.

Production. Most navy beans in the United States are produced in Michigan; great northern beans in Nebraska, Wyoming, and Idaho; pinto beans in Colorado, North Dakota, and Idaho; red kidneys in New York, Michigan, and California; and large and baby limas in California. Such regional distribution of bean classes is partly historical and partly due to climatic and disease factors.

Culture. Most beans are a temperate-season crop, with major producing areas in the United States having a mean August temperature of 70°F (21°C) or less and a mean July temperature of between 72°F (22°C) and 74°F (23°C), except for the central valley of California. With higher temperatures there is extensive flower and pod abscission, except for those species such as *P. acutifolius* and *Vigna* spp. that require the higher temperature for seed development. Flower initiation in beans may be affected by the length of day, there being long-day-, day-neutral-, and short-day-requiring types. Many types suffer a delay in flowering with lower mean temperature. *See* PHOTOPERIODISM.

Beans are grown on a wide variety of soil types, the most productive being the well-drained loams, silt loams, and clay loams abundantly supplied with organic matter. It is important that the soils not be compacted. Sandy soils are satisfactory where rainfall is sufficient and well distributed throughout the season. Beans grow best in soils of pH 6.5–7.2 that are also high in phosphorus and potassium. Zinc and manganese are minor elements that may be the first to become limiting in the lake bed or outwash soils of the United States. Except where a special nutrient deficiency exists, beans do not respond well to added fertilizer. *See* FERTILIZING.

Beans are normally planted after danger of frost is past and when soil temperature has reached 65°F (18°C). Inoculation with nitrogen-fixing bacteria (*Rhizobium* spp.) is generally unnecessary since most bean soils are naturally supplied with nodulating bacteria. Beans under irrigated or humid conditions are usually grown in rows 20 to 28 in. (50 to 70 cm) apart, with plants 2–4 in. (5–10 cm) apart in the row. Under dryland conditions the spacing is greater. In much of the tropics a considerable portion of the production is from beans grown in association with another crop, such as maize.

Weed control. One of the main cultural requirements for successful bean growing is weed control. Cultivation by hand or by mechanical means is the time-honored method, but selective chemical herbicides are available which, under favorable temperature and soil moisture conditions, are capable of giving almost 100% control of most annual grasses and broadleaved weeds. Combinations of a preplant herbicide with a preemergence herbicide are capable of excellent long season weed control with a high safety factor for the bean plant. However, perennial weeds are still a serious problem in bean fields. Plant diseases continue to be a major hazard in production, although progress toward disease control by developing resistant varieties is being made. *See* HERBICIDE.

Harvesting. Most bean crops will mature in 80 to 120 days; in higher elevations or cooler seasons they may take up to 150 days. Beans are ready for harvest when the pods and upper portion of the plants have changed color from green to tan or brown and seeds have dried to under 20% moisture. Bean seeds split or break easily when threshed mechanically at under 14% seed moisture. The ripened plants are pulled from the soil by hand or by a machine called a puller which has two sets of heavy steel knives set at an angle to the row and positioned to intercept the roots about 2 in. (5 cm) below the soil surface. The pulled plants are formed into windrows and cured for a few hours in bright clear weather before threshing with a grain combine or special bean harvester. In some countries all the harvesting operations are carried out by hand, and winnowing of the seed is still practiced. A product of high physical quality usually

results from this process. In the United States field-run beans require screening, and, for highest quality, picking, that is, the removal of seed-sized foreign matter and off-colored, stained or damaged seeds. This was formerly done by hand but presently is accomplished by electric-eye machines in the cleaning plants or elevators. *See* AGRICULTURAL MACHINERY.

Utilization. Beans are consumed as food in several forms. Lima beans and snap beans are used by canning or freezing. Limas are also used as a dry bean. Mung beans are utilized as sprouts.

Usage of dry beans (*P. vulgaris*) for food is highly dependent upon seed size, shape, color, and flavor characteristics, and is often associated with particular social or ethnic groups. Popular usage includes soups, mixed-bean salads, pork and beans, beans boiled with meat or other vegetables or cereals, baked beans, precooked beans, and powder, and, in the Peruvian Andes, parched or roasted beans. Per capita consumption of dry beans in the United States is 6.7 lb/year (3 kg/year).

Dry beans range in protein content from 17 to 32%. Commercial types average around 23–25%. The sulfur-containing amino acids methionine and cystine are the first limiting factors in bean protein, from a nutritional standpoint. The dry beans have less than 2% oil (except soybeans) and some 61–63% carbohydrate, and are rich sources of fiber, calcium, and phosphorus. Because of the presence of heat-labile growth retardants, beans must be thoroughly cooked before being used as food. Newer developments are in the areas of cooking time, the manufacture of precooked dry powders that can be readily rehydrated to make a palatable soup, and new research impetus to identify the flatulence factors in the dry beans. *See* AGRICULTURAL SOIL AND CROP PRACTICES; BREEDING (PLANT); LEGUME. M. W. Adams

Origin and evolution. Molecular techniques that allow analysis of specific compounds such as seed storage proteins, enzymes, or deoxyribonucleic acid sequences have helped to clarify the origin of various bean species. Specific variants of these molecules are identified in the cultivars, and by tracing these variants among the existing wild populations of the crop species the actual progenitor populations of the crop can be identified and a domestication area can be proposed. It has been determined, for example, that the common bean (*Phaseolus*) was domesticated in at least three areas, Mexico, Central America, and the Andes mountains in South America. These domestications gave rise to two groups of cultivars. A group with small or medium-sized seeds originated in Middle America (Mexico and Central America) and is represented by navy, pinto, great northern, and black beans. The other group, with large seeds, came from the Andes and includes the kidney, yelloweye, cranberry, and snap beans.

Subsequent to the arrival of Columbus in the Americas, both cultivar groups were dispersed worldwide. The Middle American group was introduced into the southwestern and Great Plains areas of the United States. The Andean group was introduced into Europe, from where it was further disseminated to North America and possibly Africa. The lima bean resulted from a similar domestication pattern, leading to a small-seeded, Middle American and a large-seeded, Andean cultivated group. The cowpea was domesticated in Africa but was introduced as early as 1000 B.C. into the Indian subcontinent, where additional cultivar types were selected. During the slave trade the cowpea was introduced into the United States and Brazil. Soybean originated in northern China, and spread to Japan, Korea, and southeastern Asia. It is only in the last century that the cowpea has actively been grown outside Asia, mainly because it is an important source of oil and protein.

Molecular analyses show that domestication has induced a reduction in genetic diversity in beans. Wild progenitors can therefore be a source of useful traits. For example, the arcelin seed protein was identified in wild common bean from Mexico. This protein confers resistance to seed weevils and has now been introduced into bean cultivars. P. Gepts

Diseases. Beans are affected by nonparasitic and parasitic diseases. The latter are more widespread and serious and are caused by fungi, bacteria, and viruses. The former result from unfavorable environment, including improper soil-mineral relationships and impure air. Although there are more than 40 known diseases which affect field-grown beans, only half of them occur consistently and of these about 9 are considered important on a worldwide basis. Some common diseases kill young seedlings, others cause injury or death of older plants, and some produce spots on pods and seeds that make the product unsalable.

Bacterial diseases. Bacteria-induced bean diseases are especially important where rain falls frequently during the growing season, or where overhead irrigation is compulsory. The blight bacteria, *Xanthomonas phaseoli* and *Pseudomonas phaseolicola*, are seed-borne. Symptoms of blight appear most conspicuously on leaves as water-soaked spots which may

Fig. 1. Bacterial brown spot on leaves and pods of bean.

later merge to form brown, dead areas. The diseased areas caused by *X. phaseoli* often have narrow yellow margins; those caused by *P. phaseolicola* have large yellow halos. Diseased pods also develop water-soaked spots which later turn brown and form sunken lesions with reddish-brown margins. Bacterial brown spot is caused by *P. syringae*, which incites small, brown lesions on both leaves and pods (**Fig. 1**). The centers of leaf lesions fall out, producing the "shot-hole" symptom. Diseased pods are often strikingly misshapen. The bacterial wilt pathogen, *Corynebacterium flaccumfaciens*, causes yellowing, wilting, and shriveling of infected plants. Seedlings produced by infected seeds may be stunted and may not emerge from the soil.

Fungal diseases. The most important fungal diseases of beans are rust, root rots, and anthracnose. Rust, caused by *Uromyces phaseoli* f. sp. *typica*, primarily attacks leaves and produces numerous small, rust-colored pustules or blisters which may lead to premature leaf drying and defoliation. The spores formed within the blisters may be spread by wind for hundreds of miles. Root rots are caused by several fungi, including *Fusarium solani* f. sp. *phaseoli, Pythium* ssp., and *Rhizoctonia solani*. All produce reddish-brown lesions which often coalesce to form large rotted areas on the roots, resulting in stunting of diseased plants. Extensive root decay (**Fig. 2**) causes death of the plant. Lesions caused by *R. solani* are often confined in size but are characteristically sunken. If moisture and nutrition are adequate, root-rot–diseased plants may respond by production of many new fibrous roots just beneath the soil; these new roots maintain the weakened plant. Anthracnose, induced by *Colletotrichum lindemuthianum*, causes dark, sunken cankers on pods (**Fig. 3**) and angular dead spots on leaves. The pathogen is spread by infected seed and by wind, rain, and insects.

Fig. 3. Anthracnose on bean pods (fruits).

Fig. 2. Root rot of a bean plant.

Viral diseases. Bean diseases caused by viruses are troublesome throughout the world and are often serious locally. They are manifested by leaf mottling and distortion and plant stunt, malformation, and sometimes death. The important bean common mosaic virus is aphid- and seed-borne. Yellow mosaic virus is aphid-transmitted. The curly top virus is carried by leafhoppers. The golden mosaic virus of Latin America is spread by white flies. *See* PLANT VIRUSES AND VIROIDS.

Disease control. The most desirable control for bean diseases is the use of resistant cultivars. This will be possible for almost all of the diseases mentioned above in the near future. At present, there are many bean cultivars resistant to rust and to common bean mosaic. Use of disease-free seed produced in the semiarid western states is another inexpensive and practical control measure of major significance in the United States. Crop rotation and field sanitation should also be practiced. Treating seeds with chemical protectants and proper spraying of plants with safe, effective chemicals are also very useful control measures. *See* PLANT PATHOLOGY. D. J. Hagedorn

Bibliography. R. Hall (ed.), *Compendium of Bean Diseases*, 1991; R. J. Summerfield and E. H. Roberts (eds.), *Grain Legume Crops*, 1985.

Bear

The bear family (Ursidae) contains the largest terrestrial carnivores. Most taxonomists agree that there are eight extant species of bears in three subfamilies: Ursinae (American black, Asiatic black, brown, polar, Malayan Sun, and sloth bears); Tremarctinae (Andean or spectacled bear); and Ailuropodinae (giant panda). Bears currently inhabit North America, South America, Europe, and Asia. Habitats range from polar ice floes to tropical forests. They are absent from Australia, Africa, and Antarctica.

General characteristics. The fur of bears is long, shaggy, and generally unicolored brown, black, or white, although some genera possess a prominent crescent-shaped white mark on their chests. *Tremarctos*, the spectacled bear of South America, typically has a patch of white hairs encircling each eye. *Ailuropoda*, the giant panda, has a striking black-and-white color pattern. The head and body length of adults ranges between 1000 and 2800 mm (39–109 in.) with a tail length between 65 and 280 mm (3–8 in.). Adults weigh between 27 kg (60 lb) (Malayan sun bear, *Ursus malayanus*) and 775 kg (1700 lb) (Alaska brown bear, *Ursus arctos*). Males average about 20% larger than females.

Bears have a heavily built body with short, powerful limbs and a short tail. The head is large, the eyes are small, and the ears are small, rounded, and erect. Five toes are present on each foot, and each has a long, strong, moderately curved, nonretractile claw that is used for tearing and digging. The soles of the feet are hairy in species that are mainly terrestrial but naked in species such as the sun bear, which spends a great deal of time climbing. Locomotion is plantigrade (both heel and toe contact the ground). Most bears possess 42 teeth (I 3/3, C 1/1, PM 4/4, M 2/3 × 2). The sloth bear, as an adaptation for sucking up termites, has only two upper incisors and a total of 40 teeth. *See* DENTITION.

Bears do not have particularly good eyesight, but their senses of hearing and smell are excellent. They rely on their keen sense of smell, color vision, and a good memory to locate food sources. Most bears are omnivorous with the exception of the polar bear, which feeds mainly on seals and fish. Bears are among the few mammals able to survive for half a year or more without eating, drinking, urinating, or defecating. In many regions, bears are known to enter dormancy when cold weather approaches. They become fat during autumn, but with the approach of cold weather they cease eating, go into a den in a protected location, and sleep through the winter, living mainly off stored fat reserves. Dens may be in hollow trees, in excavated depressions in the ground, or in caves. Their bones do not deteriorate despite complete inactivity, and fluids and amino acids are derived by recycling products that are normally excreted as urine. Although the heart rate and respiration rate are lower than when they are active, their metabolic processes do not drop as dramatically as in a true hibernating mammal such as a woodchuck. Therefore, there is disagreement as to whether bears should be classified as true hibernators. They can be easily aroused during the winter and may awaken on their own during periods of mild weather. *See* HIBERNATION AND ESTIVATION.

Bears are normally solitary animals, with the exception of females accompanied by their young. Litters of 1 to 6 young are produced at intervals of 1 to 4 years. Births occur while the mother is in her den during the winter. Although mating occurs in late spring or summer, the fertilized egg does not implant in the uterus immediately (delayed implantation). Thus, the period of pregnancy usually extends 6 to 9 months. The young are relatively tiny at birth, ranging from 225 to 680 g (8–24 oz). They are smaller in proportion to the size of the mother (less than 1%) than is the case with any other placental mammal. They remain with their mother at least through the first autumn. They become sexually mature between 2.5 and 6 years of age and normally live 15–30 years in the wild. Populations of many species are affected by excessive killing, habitat loss, and population fragmentation from road building.

American black bear (Ursus americanus). The most numerous of all bears, this inhabits virtually all of the forested areas of North America from northern Canada to central Mexico. This species is hunted as a game animal in many areas, with over 40,000 being killed annually in North America. Black bears are extremely adaptable, a characteristic which has aided them in becoming tolerant of human presence. Although most have long, shaggy, black fur, many are brown (**Fig. 1**). Black bears are distinguished from grizzly bears by the straight (versus dished) profile of their face and by their lack of a shoulder hump. They also have smaller claws, which are adapted more to climbing than digging. Adults are 1.2–1.9 m (4–6.2 ft) in length and stand 0.7–1.0 m (2.3–3.3 ft) at the shoulder. Males usually weigh 60–225 kg (130–500 lb), females 40–150 kg (90–330 lb). Mating usually occurs from May to July with young being born in January. Vegetable matter such as fruits, berries, nuts, acorns, grass, and roots constitutes at least 75% of the diet. Other foods may include insects, fish, rodents, carrion, and occasionally, young ungulates.

Asiatic black bear (Ursus thibetanus). This bear is found from Iran and the Himalayas to Japan, China, and eastern Siberia. The coat is black with a white V-shaped marking on the chest. The chin is white. A ruff of longer hair is present on the neck. Adults are 1.2–1.9 m (3.9–6.2 ft) in total length and stand 0.7–1.0 m (2.3–3.3 ft) at the shoulder. Males weigh 60–200 kg (130–440 lbs), females 40–140 kg (90–310 lb). This species is mainly herbivorous. Asiatic bears have been severely affected by the commercial trade in bear parts; The meat, fat, and paws are popular foods in these countries. The bones, brain, blood, spinal cord, and gallbladder are sought for medicinal uses in Chinese medicine.

Fig. 1. Black bear (*Ursus americanus*). (*Photo by Tom Brakefield; © 1999 California Academy of Sciences*)

Brown bear (Ursus arctos). Also known as the grizzly or Kodiak bear, this species is both the most widespread bear and the largest living carnivore. It ranges from northwestern North America and Scandinavia through Russia to Japan. Scattered populations also exist in parts of Europe, the middle east, the Himalayas, China, and Mongolia. These bears are highly adaptable and can live in forests, tundra, and deserts. The pelage (coat, including all fur and hairs) is uniform brown or blonde to near black. Adults are 1.5–2.8 m (5.0–9.2 ft) in total length and stand 0.9–1.5 m (3–5.0 ft) at the shoulder. Males weigh 135–545 kg (300–1200 lb), females 80–250 kg (175–550 lb). Many populations are very small and in danger of extirpation. For example, in the United States outside of Alaska, only five isolated populations survive, totaling about 1000 animals. These populations are on the Endangered Species list.

Polar bear (Ursus maritimus). Polar bears are among the largest of the world's bear species. Despite their different appearance, polar bears and brown bears are closely related. The polar bear's body is stocky but lacks the shoulder hump of the brown bear. The neck is longer in relation to the rest of the body than in other bears, and the large, oarlike forepaws are adapted for swimming (**Fig. 2**). The ears are small, the fur is whitish to yellowish, and the skin beneath the coat is black. Polar bears mainly inhabit arctic regions around the North Pole and the coastlines of continents and islands where there are high densities of ringed seals, their primary prey. Adults are 180–250 cm (5.8–8.3 ft) in total length. Males weigh 400–600 kg (880–1320 lb), females 200–350 kg (440–770 lb). Polar bears are not endangered, but are threatened by increasing pollution and global warming. The Agreement on the Conservation of Polar Bears (1976) requires that this species be managed according to "sound conservation practices," and mandates protection of their habitat.

Sloth bear (Ursus ursinus). This bear inhabits India, Nepal, Bhutan, Sri Lanka, and possibly Bangladesh. The coat is long, shaggy, and black with a white U on the chest. Termites are the primary food, but other insects, honey, eggs, carrion, and berries are also eaten. Adaptations for its unusual diet include long, slightly curved claws; large, protrusible lips; a broad, hollowed palate for sucking; and the absence of two front upper incisors. Sloth bears are largely nocturnal, and females frequently carry their young on their back, two behaviors that are unique among bears. Sloth bears are 1.4–1.9 m (4.6–6.2 ft) in total length and stand 60–90 cm (2–3 ft) at the shoulder. Adult males weigh 80–145 kg (176–320 lb), adult females 55–95 kg (120–210 lb).

Malayan sun bear (Ursus malayanus). These are the smallest of the bears and have the shortest hair. They inhabit the dense forests of southeast Asia including Sumatra, Borneo, eastern India, and southern China. The coat is black with an ocher or white circle or semicircle on the chest. Both their snout and ears are short. These nocturnal bears are highly arboreal and often sleep in nests in trees. They are omnivorous, feeding primarily on insects and fruit. Adults are 1.1–1.5 m (3.6–4.9 ft) in total length and stand approximately 70 cm (2.3 ft) at the shoulder. Adults weigh 27–65 kg (60–143 lb).

Andean bear (Tremarctos ornatus). Also known as the spectacled bear because of the light markings that often encircle the eyes, these bears are the only South American ursid and are the second largest land mammal in South America. They inhabit the

Fig. 2. Polar bear (*Ursus maritimus*).

mountainous regions from eastern Panama to northern Argentina. They may be found from high altitudes to coastal lowlands and are found in a wide variety of habitats including cloud forests, rainforests, high-altitude grasslands, and scrub deserts. Together with Malayan sun bears, these are the only bears that live on both sides of the equator. The coat is black with a creamy white, biblike marking on the chin, neck, and/or chest. These relatively slim and agile bears build tree nests for sleeping and feed primarily on plant matter, especially bromeliads. Adults are 1.3–2.0 m (4.3–6.6 ft) in total length and 70–90 cm (2.3–3.0 ft) at the shoulder. Males weigh 100–175 kg (220–385 lb), females 60–80 kg (132–176 lb).

Giant panda (Ailuropoda melanoleuca). The current range of this species has been restricted to six isolated mountain ranges in central and western China. Giant pandas live in cool, damp bamboo forests at altitudes of 1500–3400 m (4900–11,000 ft). It is one of the most endangered mammals in the world; only about 1000 exist in the wild. Although formerly classified with the Procyonidae (raccoons and their relatives), recent genetic studies show that the giant panda represents an early divergence from the bear family. It differs from other bears in several ways: it does not hibernate, one of its wrist bones has evolved into a "pseudothumb"; and a newborn is remarkably small, weighing just 100–200 g, about 0.001% of its mother's weight. The ears, eye patches, legs, and shoulders are black; the remainder of the body is white. Giant pandas are 1325–1625 mm (57–70 in.) in total length, including a 120–130 mm (4–5 in.) tail. They stand 700–800 cm (2.7–3.2 ft) at the shoulder and weigh 100–150 kg (220–330 lb). They feed almost exclusively on bamboo. *See* PANDA.

Donald W. Linzey

Bibliography. G. A. Feldhamer, B. C. Thompson, and J. A. Chapman (eds.), *Wild Mammals of North America: Biology, Management, and Conservation*, 2d ed., Johns Hopkins University Press, 2003; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., John's Hopkins University Press, 1999; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Beat

A variation in the intensity of a composite wave which is formed from two distinct waves with different frequencies. Beats were first observed in sound waves, such as those produced by two tuning forks with different frequencies. Beats also can be produced by other waves. They can occur in the motion of two pendulums of different lengths and have been observed among the different-frequency phonons in a crystal lattice. *See* CRYSTAL STRUCTURE; PENDULUM.

As a simple example, consider two waves of equal amplitudes A_1 and different frequencies ω_1 and ω_2 in hertz at the same spatial point given by $\psi_1(t) = A_1 \sin(\omega_1 t)$ and $\psi_2(t) = A_1 \sin(\omega_2 t)$. The sum of these waves at this point at an instantaneous time is given by Eq. (1), assuming that they are coherent. The last

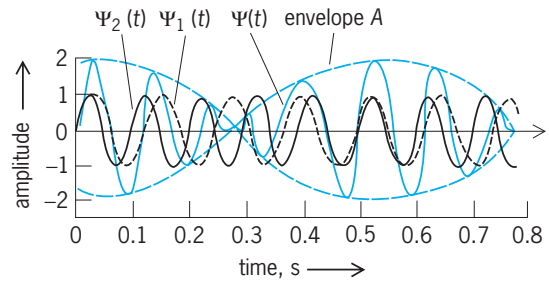

Production of beats. Two waves, $\psi_1(t)$ with frequency $\omega_1 =$ 8 Hz and $\omega_2(t)$ with frequency $\omega_2 = 10$ Hz, produce composite wave $\psi(t)$ with amplitude A, an envelope that encloses ψ.

$$\psi = \psi_1(t) + \psi_2(t) = A_1(\sin \omega_1 t + \sin \omega_2 t)$$

$$= \{2A_1 \cos[{}^1/_2(\omega_1 - \omega_2)t]\} \sin[{}^1/_2(\omega_1 + \omega_2)t] \quad (1)$$

form of the wave can be written as $\psi(t) = A \sin(\omega t)$, providing that the amplitude of the composite wave $\psi(t)$ is given by $A = 2A_1 \cos[\frac{1}{2}(\omega_1 - \omega_2)t]$ and the frequency is $\omega = \frac{1}{2}(\omega_1 + \omega_2)$. The **illustration** shows the case $A_1 = 1$, $\omega_1 = 8$ Hz, and $\omega_2 = 10$ Hz. The amplitude A of ψ is the envelope enclosing the curve of the total wave. The beat pulsations occur because of this low-frequency envelope.

There is a simple and interesting approximation to Eq. (1) when $\omega_1 - \omega_2 = \epsilon$ is small and short times are considered. The amplitude of $\psi(t)$ is approximately $B \simeq 2A_1 - \frac{1}{2}A_1\epsilon^2 t^2$, and the composite wave is given by Eq. (2) for sufficiently short times. Of course, the

$$\psi_{1f}(t) \approx (2A - {}^1/_2 A_1 \epsilon^2 t^2) \sin[{}^1/_2(\omega_1 + \omega_2)t] \quad (2)$$

beat phenomenon is lost in this limit. Similarly, the wave in Eq. (1) would reduce to Eq. (3), and only the

$$\psi_{bf} \approx A_1 \cos[{}^1/_2(\omega_1 - \omega_2)t] \quad (3)$$

beat frequency would appear if the sum frequency $\frac{1}{2}(\omega_1 + \omega_2)$ was too high to be observed.

The individual waves $\psi_1(t)$ and $\psi_2(t)$ could represent a model for monochromatic sound waves, the simple harmonic motion of two pendulums, two acoustic phonons in a crystal, or two electromagnetic waves at a point in space. For sound waves, the beats produce a throbbing sound at a frequency $(\omega_1 - \omega_2)/2$. In the example given in the illustration, $\omega_1 - \omega_2 = 2$ Hz, and one beat per second is produced; this is the envelope. If the difference in frequencies is greater than 12–16 Hz, the beats are difficult to distinguish from one another unless the difference frequency is audible, and, in this case, a difference tone is produced. It is also possible to produce audible summation tones. *See* ELECTROMAGNETIC RADIATION; HARMONIC MOTION; LATTICE VIBRATIONS; SOUND.

One application of beat phenomena is to use one object with accurately known frequency ω_1 to determine the unknown frequency ω_2 of another such

object. The beat-frequency or heterodyne oscillator also operates by producing beats from two frequencies. *See* OSCILLATOR; WAVE MOTION. Brian De Facio

Beaver

The common name for the most primitive living rodent, which has an ancestry that goes back to the time when modern mammals evolved. Two families and three species of beavers exist in the world. The family Aplodontidae contains a single species, the mountain beaver (*Aplodontia rufa*). The family Castoridae contains two species: the American beaver (*Castor canadensis*; see **illustration**) and the Eurasian beaver (*Castor fiber*).

Mountain beaver. This species is limited to the coastal conferous forests from southern British Columbia to northern California. It has no close living relatives. Its name is a misnomer as it is not particularly montane (being more often found in lowland forests, although it does occur at all elevations from sea level to the treeline), nor does it swim like American beavers.

Mountain beavers, also known as sewellels or boomers, are medium-sized rodents with blackish to reddish-brown fur on their dorsal surface and tawny hair on their ventral surface. The head is broad and flattened, the ears and eyes are small, and the vibrissae are long, stiff, and whitish. A distinctive white spot is present beneath each ear. The furred tail is short and vestigial. The senses of sight and hearing are poor, but these mammals have excellent olfactory and tactile sensations. The teeth are rootless and grow throughout life. The dental formula is I 1/1, C 0/0, PM 2/1, M 3/3 × 2 for a total of 22 teeth. Adults are 238–470 mm (9–18 in.) in total length, including a 19–55 mm (0.7–2 in.) tail. They weigh 806–1.325 g (28–46 oz). *See* DENTITION.

Mountain beavers inhabit moist forests and forest openings with a dense understory. These nocturnal, mostly solitary mammals spend most of their time in a series of underground burrows with interconnecting feeding and nesting chambers. They do not hibernate but may remain in their burrow for extended periods during cold weather. They are strictly vegetarians and eat a variety of plant foods including ferns, grasses, and roots. If necessary, they are capable of climbing trees in search of food. Mating occurs in late winter and a single litter of 2 or 3 young are born in the spring after a gestation of about 1 month. Sexual maturity is attained in the second year.

American beaver. This is the largest rodent in the Northern Hemisphere and one of the largest in the world, second only to the capybara of South America. Beavers are stocky mammals with scaly, nearly naked, horizontally flattened, paddle-shaped tails and webbed hind feet (see illustration). The dense pelage consists of fine, short underfur overlaid with long, coarse, shiny guard hairs. The coat is brown to blackish-brown, becoming slightly paler on the underparts. The tail and feet are black. The ears are short and black and are set far back on the broad, rounded head. The small ears, like the nose, are equipped with valves that shut when the animal is underwater. Although a beaver's eyesight is not exceptional, its senses of hearing, smell, and touch are excellent. Each foot has five clawed toes. The claws of the two inner toes on each hind foot have split nails and are modified as combs for grooming. Both sexes possess paired anal scent glands known as "castors." These glands are largest in males and secrete a bitter, orange-brown, pungent oil known as castoreum. This secretion is deposited on scent mounds usually near the edge of an individual's territory and appear to serve as a population regulation mechanism to warn other beavers away from occupied territory. Adult beavers are 900–1200 mm (35–47 in.) in total length, including a 225–300 mm (9–12 in.) tail. Adults weigh 16–30 kg (35–65 lb).

American beaver (*Castor canadensis*).

Beavers formerly ranged along streams and lakes throughout most of the forested regions of North America. They have been extirpated from many areas primarily by overtrapping, but with regulation of trapping and translocations, populations have increased dramatically. They have been introduced into South America, Europe, and Asia.

Beaver signs include the familiar dams, lodges, canals, and tree cuttings. Dams are constructed across slow, meandering streams that pass through flat, moist, wooded valleys. With their powerful, chiselllike incisor teeth, beavers can fell a tree 10 cm (4 in.) thick in 15 min. Lodges are dome-shaped and are formed by a haphazard piling and interlacing of thick sticks and heavier poles that are plastered together with mud. A small air hole is left at the top for ventilation. Beaver lodges usually rise 0.9 –1.5 m (3–5 ft) above the surface of the water and may be 12 m (40 ft) in diameter at the base. Entrances are located beneath the water level and a crude platform in the center of the pile is used for a place to rest. Bank burrows, also with entrances below the surface of the water, may be used in some areas.

Beavers, which are mainly nocturnal, feed primarily on the bark and adjacent outer layers (cambium) of deciduous trees such as aspen, cottonwood, birch,

willow, and alder. In areas of freezing winter temperatures, beavers cache branches in the mud at the bottom of the pond as a winter food supply. Their fur-lined lips can be closed and sealed behind the incisors so that they can carry branches underwater without taking water into their lungs.

Beavers are monogamous and normally mate for life. Mating usually occurs in January and February. After a gestation of approximately 106 days, a single litter of 3 or 4 kits is born. Young beavers mature slowly. Even though they learn to swim when a month old and are weaned after 6 weeks, they remain under their parents' care for about 2 years. Thus, a typical beaver colony may consist of the parents, kits, and yearlings. Sexual maturity is usually attained during the beaver's second winter. Life span in the wild is 10–12 years, although some wild individuals have been known to live over 20 years.

Eurasian beaver. Once found throughout Europe and Asia, this species was reduced to only eight relic populations by the early 1900s. Populations have grown and the range has expanded due to reintroductions and translocations. The description and ecology of this species is similar to *Castor canadensis.*

Donald W. Linzey

Bibliography. G. A. Feldhamer, B. C. Thompson, and J. A. Chapman (eds.), *Wild Mammals of North America: Biology, Management, and Conservation*, 2d ed., Johns Hopkins University Press, 2003; D. W. Linzey, *The Mammals of Virginia* McDonald & Woodward, 1998; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Beech

A genus, *Fagus*, of deciduous trees of the beech family, Fagaceae, order Fagales. They can best be distinguished by their long (often more than 1 in. or 2.5 cm), slender, scaly winter buds; their thin, gray bark, smooth even in old trees; and their simple, toothed, ovate or ovate-oblong, deciduous leaves. *See* FAGALES.

The American beech (*F. grandifolia*) grows to a height of 120 ft (36 m) and is native in the United States east of the Mississippi River and in the lower Mississippi Valley. The hard, strong wood is used for furniture, handles, woodenware, cooperage, and veneer. The small, edible, three-sided nuts, called beechnuts, are valuable food for wildlife. They are borne in pairs, surrounded by a spiny involucre, the homolog of the spiny covering in the chestnut bur and the cup of the acorn, which also belong to the beech family. *See* FRUIT.

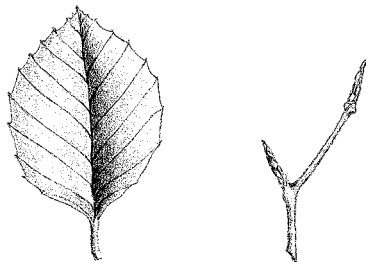

European beech (*Fagus sylvatica*).

The European beech (*F. sylvatica*; see **illus.**) is more popular as an ornamental tree than the American species. Its leaves are smaller, with 5–9 pairs of primary veins compared with 9–14 pairs in the American beech. The leaf teeth are also shorter. Important ornamental varieties are *F. sylvatica purpurea*, the copper or purple beech; var. *incisa*, the cut-leaved or fern-leaved beech; and *F. pendula*, the weeping European beech. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Beef cattle production

Raising of cattle for meat. The muscle from cattle over 6 months of age is beef and from younger cattle, veal. Beef is a highly nutritious food, containing approximately 20% protein; it is rich in essential amino acids and minerals, and is a good source of the B vitamins (B_6, B_{12}, niacin, and riboflavin). Beef also supplies a generous portion of the daily dietary iron requirements.

Cattle, members of the family Bovidae, are ruminants because they have a four-compartment stomach filled with microorganisms which allow digestion of forages (grasses) and other fibrous food sources. Most domestic breeds common to the temperate zone originated from European *Bos taurus* ancestry, whereas Zebu or Brahman originated from Asian *B. indicus*, indigenous to the tropics. *See* ARTIODACTYLA; ZEBU.

World beef production. Beef cattle numbers have either stayed the same or increased slightly in almost all areas of the world since 1970. As countries improve their standards of living and become increasingly industrialized and as populations grow, there will be increased demand for beef cattle worldwide. Competition from pork and poultry products has resulted in decreased beef consumption in many developed countries. A slow decrease of beef cattle numbers is expected to continue in the United States due to a stable population growth and cost competitiveness of other food commodities.

The United States is the major beef producer. Australia, New Zealand, Argentina, and Brazil are leading beef exporters, followed by France and Denmark. Leading importing countries include the United States, Great Britain, Germany, and Japan. China has also increased beef imports markedly.

World beef production is based on the use of grass and harvested forages. Many land areas are unsuitable for growing crops but are suitable for forage production. Cattle can graze these areas and convert vegetation into high-quality protein food. The major phases of beef production involve the use of cow herds for raising calves and then calves grown out and fattened for market. Calves weigh 65–100 lb (30–45 kg) at birth and nurse until they are 6–7 months of age, when they are weaned. Under good management, weaned calves can weigh 500 lb (225 kg) at

7 months. In some countries, weaned calves are grown on grass or forages until becoming yearlings, at which time they are fattened on grain until marketed at the age of 15–24 months. The pampas of Argentina provide year-round grazing and produce slaughter beef directly from grass at 2 years of age. Many New Zealand, Australian, and Argentinean cattle are marketed from pasture alone, while many cattle in the United States, Canada, or England are grain-fed. Grain feeding allows cattle to be marketed at young ages and with a more palatable flavor. Cattle could be marketed successfully from forages alone, though it would take longer to get cattle to acceptable market weights.

United States beef production. Approximately 3% of United States beef cattle are in purebred registered herds. These cattle are identified and selected by using extensive record systems in order to take advantage of genetic improvements in cattle production. The superior purebred bulls are used by commercial producers to increase market acceptability of the calves. The rank of the six leading purebred breeds are Angus (**Fig. 1**), Hereford (**Fig. 2**), Limousin Simmental (**Fig. 3**), Polled Hereford, and Charolais (**Fig. 4**). In the southern region of the United States, Brahman (**Fig. 5**) or breeds with Brahman blood such as Santa Gertrudis, Brangus, and Beefmaster are more tolerant of humidity, heat, and insects. A systematic cross of two or three breeds is popular with producers because the hybrid vigor of mating different breeds results in lower mortality, improved growth, and reproductive efficiency. Purebred producers use measures at birth, at weaning (205 days), and at yearling to project the growth rate of cattle; they also measure milk production and fertility in producing cows. National sire summaries which statistically calculate expected progeny differences for purebred cattle provide genetic predictions of cattle performance and are used to improve accurate matings of animals by purebred producers. *See* BREEDING (ANIMAL).

Beginning in the late 1960s, many commercial producers seeking greater size and other improvements included one or more of the European breeds of Charolais, Simmental, Limousin, Maine Anjou, Chianina, Beef Fresian, and others in their crossing systems. As a result, over 40 breeds of cattle are used by producers in the United States. The climate, market demands, and management capabilities determine which breeds or crosses are best adapted for production. Due to the rough topography and limited rainfall of the Rocky Mountains and sections of the central and western plains, these areas are adapted to the production of feeder calves, yearlings, and some 2-year-olds. In the range areas of the Nebraska Sandhills, Kansas, and Oklahoma Flint Hills, approximately 8–15 acres (3–6 ha) of grass are required annually for a cow and calf, while in more arid areas of Arizona and New Mexico 150–200 acres (60–80 ha) are needed. In the Midwest, a cow and calf can be maintained on 2–4 acres (1–2 ha).

Fig. 1. Angus bull. (*American Angus Ass.*)

Fig. 2. Hereford bull. (*American Hereford Ass.*)

Fig. 3. Simmental bull. (*American Simmental Ass.*)

Feeding beef cattle. Feed is the major cost item in beef production. It is estimated that roughages furnish 80% of the total feed for cattle. From birth to a market weight of 1000 lb (450 kg), it takes less than 3 lb of grain for every pound of live weight gain. Most grain is fed in the final stages before marketing to shorten the feeding period and ensure a juicy, highly palatable meat product.

The mature cow is physiologically adapted to utilize, in addition to quality feeds, roughages such as straws, corn stover, cottonseed hulls, and byproducts of the sugar and canning industries. The rumen compartment of a mature bovine stomach can hold 40–60 gal (150–230 liters) and has microorganisms, such as bacteria, yeasts, and protozoa, to aid roughage digestion. The rumination processes of

Fig. 4. Charolais bull. (*American International Charolais Ass.*)

cud chewing, regurgitation, and rumen microbial fermentation aid in the release of digestible energy in the form of glucose and acetic and propionic acids from the fibrous cellulose stems of roughage. Cattle have an advantage over monogastric swine or poultry because they can utilize forages and synthesize their necessary B vitamins and essential amino acids. The nutrient requirements of cattle are water, proteins, carbohydrates, fats, minerals, and vitamins. Energy from carbohydrates and fats along with proteins is a major dietary and cost consideration. Quality forages and good pastures, except in soil-deficient areas, can furnish adequate nutrients for maintenance, growth, and reproduction.

With population increases and greater human consumption of concentrate grains, research efforts have been directed to maximizing the use of corn and milo stovers as well as by-product straws of grain harvest in beef rations. Improved storage techniques and the processing of low-quality forages with nitrogen (protein) additives allow producers to get feed from waste materials. A popular feed for cattle is corn gluten feed, which is a high-moisture by-product of the distilling industry. Mature cattle eat about 2 lb daily for each 100 lb of weight; younger cattle eat 2.5–3.0 lb of air-dry feed per 100 lb (cured forages and grains of approximately 10–15% moisture).

The National Research Council gives the maintenance requirement of cattle as 0.75 lb of total digestible nutrient (TDN) energy per 100 lb liveweight.

Fig. 5. Brahman, tolerant of humidity, heat, and insects. (*American Brahman Breeders Ass.*)

The TDN of a feed is defined as the sum of all digestible organic nutrients, protein, fiber, carbohydrate, plus 2.25 times the fat content. The council also expresses energy requirements as digestible energy, with 1 lb of TDN equivalent to 2000 kilocalories of digestible energy. One pound of weight gain requires 3.53 lb of TDN above maintenance. Air-dry grains, such as oats, barley, corn, and milo, vary in energy but average 72–80% TDN, and cured roughages, such as straw and hays, average 40–50% TDN. Balanced grain rations have more digestible energy and produce larger daily weight gains. Energy utilization of grains is improved by steaming and cracking, rolling, or flaking. The average daily gain of calves (over 6 months), yearlings, and 2-year-old steers fed a full feed of 1.75 lb grain per 100 lb are about 2.10, 2.25, and 2.50 lb, respectively, Compared to steers, weight gains are 5–10% slower for heifers and 5–10% faster for bulls. Efficiency of gains (energy per pound of gain) favors calves over yearlings and yearlings over 2-year-olds. As an animal grows, more fat is deposited and the animal becomes less efficient, requiring more feed to produce a pound of weight gain.

The recommended crude protein percent of fattening rations is approximately 12 and 11%, respectively, for calves and yearlings, and is somewhat greater for breeding bulls, lactating cows, and very young calves. Digestible protein needs are 75% of crude protein. Grains and many hays and straws are low in protein and are often supplemented with soybean and linseed and cottonseed oilmeal. Green pastures and legume hays are good sources of protein.

Nonprotein nitrogen sources, such as urea, biuret, and starea (supplements synthetically crystallized from nitrogen in the air), have found wide use as cheap protein substitutes for oilmeals. The rumen microorganisms combine the nonprotein nitrogen with carbohydrate and mineral elements to provide essential amino acids for use by cattle.

Ensiling is the method of preserving grass forages with 40–50% moisture (haylage), or grain forages (silage) with 60–70% moisture, in airtight structures. The exclusion of air with consequent fermentation of carbohydrates in the forages produces lactic and acetic acids, which act as preservatives. Silages are palatable and are widely used not only because they reduce curing and harvest losses but also because they provide more nutrients per acre. Certain hormonelike feed additives (such as ionophores) can alter rumen digestion and improve feed efficiency, while other hormonelike compounds, fed or implanted under the skin, can increase gains and efficiency 8–15%. All feed-additives use is extensively regulated to prevent any residue from remaining in the meat. Low-level feeding of antibiotics, such as aureomycin and terramycin, improves roughage utilization and helps minimize disease. *See* ANIMAL FEEDS; ANTIBIOTIC.

Marketing and grading. Cattle of varying weights, ages, and sexes are slaughtered annually. For clarity of market reporting, cattle are classified as steers, heifers, cows, bulls, stags, veals, and slaughter calves.

In 1973 a bullock market class was included for slaughter bulls under 2 years of age.

Consumers prefer beef from young grain-fed cattle averaging 900–1200 lb (400–540 kg) alive. Steer (castrated males) beef at a young age is generally the highest priced. Young fat heifers, if not pregnant, yield carcasses comparable to steers; however, heifers mature and fatten sooner and must be sold at lighter weights. Pregnancy reduces carcass yield and so heifers undersell steers. Cows (mature females) yield less muscular carcasses and less tender beef. Bulls and stags (late-castrated males) have coarser-textured meat. A veal is a fleshy young calf averaging 150–300 lb (65–135 kg), while older and heavier calves—up to 550 lb (248 kg)—are slaughter calves. The meat of young cattle has more moisture and a lighter-colored lean.

Slaughter cattle are purchased primarily on the quality of the carcass, probable dressing percent, and lean yield. They are graded according to government standards in descending value as prime, choice, select, standard, commercial, utility, and cutter. The age of the carcass plus the quality of the lean determines the final quality grade. The yield grade is determined by the relation of the amount of muscle to the amount of fat in the carcass. Dressing percent is the carcass yield based upon liveweight (the ratio of liveweight to carcass weight) and varies from 65% for prime cattle to 40% for cutter cattle. Fat increases dressing percent, whereas water, feed fill remaining in the paunch, coarse bone, and heavy hides reduce dressing percent.

Marbling, an interspersion of fat with lean, contributes to flavor, whereas texture (diameter of muscle fibers) and age determine tenderness. Young carcasses have softer, red-colored bones with cartilage tips on the vertebrae as contrasted to flinty white bone of older beef. Meat from older cattle is less tender because of intermuscular collagen and connective tissue. In the United States, cattle are marketed at young ages (12–20 months), and as the age of market cattle has been reduced, cattle are marketed with less fat waste. Old producing cows and bulls are used primarily as ground beef. Because of a vast increase in fast-food restaurants, ground beef makes up over half of the beef consumed.

Aging of beef (connective tissue breakdown) occurs naturally in beef for the first 7 days after slaughter. Electrical stimulation applied to the carcass shortly after slaughter can also improve tenderness. Provisioners have inproved efficiency and reduced their transportation costs by boning and trimming carcasses at the slaughter sites. A majority of beef is trimmed, packaged, and shipped in boxes to improve efficiency and standardize packages. *See* FOOD MANUFACTURING.

Beef economics. Production and marketing of cattle in the United States changes periodically. In the mid-1800s, slaughter plants were established in the East and were supplied with western grass cattle. In the 1920s and through the early 1940s, the corn belt states were important feeding areas. Many farmers used their corn to fatten 50–100 head annually and marketed them at major slaughter plants in Chicago, Omaha, and Kansas City. Annual per capita beef consumption increased to over 100 lb (45 kg) after World War II, and the feeding and slaughtering centers moved westward to the High Plains. Feedlots with 1000 to 10,000 head were established in Texas, Oklahoma, Colorado, and Nebraska, close to feeder cattle supplies.

Beef numbers, because of supplies and prices, have fluctuated in cycles every 10–12 years. Low cattle numbers cause an increase in prices, which in turn encourages herd increases until an oversupply occurs. A surplus in turn results in lower prices. To minimize risks in cattle feeding, operators are using computers to formulate least-cost diets and are using future contracts to hedge purchases and marketings. Since imports affect supplies and prices, producers have promoted legislation to control imports in proportion to domestic production.

Higher living standards made steak a prestige restaurant meat. Changes in life styles popularized hamburger food chains throughout the world. Shortages of cow slaughter and of grass cattle for hamburger have increased blendings of soybean protein with beef. Westernization and industrialization created greater demands for fed beef in Europe, in Japan, and in the Mideast and other developing areas.

Slaughter. In slaughtering, carcasses are split through the backbone into sides. For wholesale trade, sides are divided between the twelfth and thirteenth ribs into fore- and hindquarters and then retailed primarily as steaks and roasts. The major forequarter cuts are chuck and rib; the hindquarter cuts are loin and round.

Approximately 75% of fresh beef is merchandised by large grocery chains and is mostly choice and select grade. The emphasis for lean, marbled beef has resulted in further scoring carcasses within grades for cutability with 1 as top yield and 5 as the poorest. Cutability or lean yield is determined by measuring the loin area and fat thickness over the longissimus rib-eye muscle exposed between the twelfth and thirteenth ribs, plus accounting for kidney fat and carcass weight.

An attempt is made to utilize all offal of slaughter. Edible offal, such as heart, liver, and tongue, are called meat specialties. The hide, used as leather, represents about 8% of liveweight. Use is made of inedible offal for production of bone meal, blood meal, and meat scraps for livestock feeds; organs and glands for pharmaceutical products; hooves and horns for glue and fertilizer; and intestines for sausage casings.

Reproduction. Cattle are polyestrous, or breed at all seasons. Heifers reach puberty or sexual maturing, depending upon nutrition, at about 5–12 months of age. The estrous period, or acceptance of the male, averages about 18 h, and the estrous cycle occurs at intervals of 18–24 days. Well-grown heifers can be bred to calve at 2 years of age, but some cattle raisers prefer to calve later at $2^1/_2$–3 years. Average gestation or pregnancy is about 283 days. Twin births occur about 1 in 200, and other multiple births are less common. A heifer born twin to a bull is usually infertile and is called a freemartin. Young bulls are virile at

about 10 months and should be used sparingly until 16 months. In range areas one active bull over 2 years is kept for 20–30 cows. Artificial insemination (AI) is not used as extensively with beef as with dairy cattle. Synchronizing or controlling estrus by feeding or injecting progesterone hormones is used. Estrous synchronization permits breeding of many cows (mostly by AI) close together and results in more uniform age calves. Paralleling AI use, the techniques of fertilized ova transplants and hormonal superovulation have proved fairly successful. These techniques permit annual multiple-offspring production from superior genetic cows through surgical transfer of fertilized ova to compatible recipient mothers. *See* ESTRUS.

Progressive cattle ranchers use performance tests to select cows that will wean 500–600-lb (225–270-kg) calves at 7 months and feeders that will gain $2^1/_2$–3 lb (1.1–1.4 kg) daily and weigh 1000–1100 lb (450–495 kg) at 1 year. Rate of gain and efficiency are highly correlated and heritable, meaning that fast-gaining parents should produce fast- and efficient-gaining offspring. Genetic engineering has attempted to change the characteristics of the beef animal by altering or rearranging its deoxyribonucleic acid (DNA). Mapping of the beef genome would allow producers to identify beef cattle that produce more efficiently, yield a more nutritious product, and develop a greater resistance to disease.

Diseases. The most universally destructive diseases of cattle are tuberculosis, brucellosis or abortion, foot-and-mouth disease, anthrax, and blackleg. Competent veterinarian diagnosis and treatment are important.

Tuberculosis is highly contagious and is transmissible to humans through milk. It is caused by a specific bacterium that may attack all parts of the body but particularly the lymphatic glands. There is no specific treatment, but the identification by an annual tuberculin skin test and disposition of reactors has markedly reduced tuberculosis. *See* TUBERCULOSIS.

Contagious abortion, brucellosis or Bang's disease, is a serious reproductive disease caused by *Brucella abortus*. Bang's disease causes birth of premature or dead calves. It has been reduced by blood-testing females (agglutination test) annually and slaughtering carriers. Calfhood vaccination of heifers between 4 and 8 months can provide immunity. *See* BRUCELLOSIS.

Other reproductive diseases, such as leptospirosis, vibriosis, and trichomoniasis, are prevalent, but they can be minimized by testing, immunization, and good health management.

Foot-and-mouth disease is caused by a virus and produces blister lesions on the feet and in and on the mouth. The disease is more debilitating than fatal and causes great economic losses due to weight reductions when cattle cannot or refuse to eat. A 1947 outbreak in Mexico was eradicated through inspection, vaccination, quarantine, and slaughter. Outbreaks in Europe and particularly in England in 1967 caused banning of European cattle imports into the United States. Outbreaks of foot-and-mouth disease in Argentina and other countries have limited imports to only canned beef from South America.

In 1984, a new vaccine for foot-and-mouth disease was developed by using biotechnology. The vaccine involves a live bacterin and is not used in the United States since it is free of foot-and-mouth disease. *See* FOOT-AND-MOUTH DISEASE; VIRUS.

Anthrax (splenic fever) is very infectious and usually fatal. The anthrax organism, a bacillus, can live many years and is transferable to other species and to humans. Diagnosis can be made by microscopic examination of a blood sample for the specific bacillus. *See* ANTHRAX.

Blackleg is an acute disease of young cattle. Annual calfhood vaccination is practiced in areas where the disease prevails. When outbreaks occur, the best preventive measure for blackleg and anthrax are burning or deep burial of infected carcasses and prompt vaccination with antiserum. *See* BLACKLEG.

Bloat is a digestive disturbance which causes a loss of millions of dollars annually. Bloat is recognized by distention of the animal's left side due to excessive gas formation in the rumen. Pasture forages such as alfalfa, clover, and other legumes eaten wet oftentimes produce heavy fermentation. A gradual change from dry roughage to legume pasture, as well as feeding hay on pasture, may help minimize bloat. The use of miscible oils in the drinking water and the use of poloxalene in mineral mixes and in protein blocks help reduce pasture bloat. BSE (bovine spongiform encephalopathy, or mad cow disease) is a progressive neurological disorder of cattle that results from infection by an unconventional transmissible agent termed a prion; it is an invariably fatal brain disease with an unusually long incubation period (measured in years). *See* PRION DISEASE.

Many other diseases affect the beef cow herd, though most can be minimized by a good health management program. Diseases include anaplasmoisis, which is transmitted by insect vectors and causes anemia, weakness, and dehydration; bovine respiratory disease, which is caused by several viruses and bacteria and produces respiratory problems; bovine virus diarrhea, which is caused by direct contact with infected animals and results in respiratory and alimentary problems; and bovine rhinotracheitis (rednose), which is a viral disease resulting in respiratory problems, abortion, conjunctivitis, enteritis, and encephalitis.

Feedlot health. To reduce shipping fever (hemorrhagic septicemia flu), incoming feedlot cattle are fed for 2 weeks after arrival daily high levels (300–500 mg) of antibiotics. New feeder cattle are usually wormed, sprayed for lice and grubs, and vaccinated for leptospirosis, infectious bovine rhinotracheitis, and bovine virus diarrhea.

There are two basic types of vaccines, killed (inactivated) and live. Killed products are more stable to use in feedlot situations.

There are programs of preconditioning feeder cattle on ranches for parasites and feedlot disease immunity. Foot rot or foul foot, common in wet feedlots, is controlled by low levels of antibiotics and iodine compounds in rations or in salt-mineral mixtures. Vitamin A deficiencies in feedlot cattle are prevented through intramuscular injections of vitamins

A, D, and E or by supplemental mixes in the rations. *See* AGRICULTURAL SCIENCE (ANIMAL); DAIRY CATTLE PRODUCTION. Douglas F. Parrett

Bibliography. B. A. Anderson and I. M. Hoke, *Composition of Foods: Beef Products*, Agric. Handb. 8-13, 1990; H. M. Briggs, *Modern Breeds of Livestock*, 1980; National Research Council, *Recommended Nutrient Allowances for Beef Cattle*, 1990; R. E. Taylor, *Beef Production*, 1992; U.S. Department of Agriculture, *Official Standards for Grades of Carcass Beef*, 1973.

Beekeeping

The management and maintenance of colonies of honeybees. Beekeeping is an ancient art. Although the commonly known honeybee species is native to Europe and Africa only, humans have transported them to other continents, and in most places they have flourished. The natural home for a honeybee colony is a hollow tree (**Fig. 1**), log, or cave. European strains of the honeybee build a nest only in locations which are dry and protected from the wind and sunlight (**Fig. 2**). African bees are less selective and may nest in hollowed-out termite mounds, rock piles, and locations which are less well protected. It is not known when humans first started to husband bees. However, there is much archeological and historical evidence that there were beekeepers among all of the ancient European and African civilizations.

Bees are important plant pollinators, but early beekeepers were not concerned with pollination and cross pollination of agricultural crops since their fields were small and they conducted diversified farming. For them, the importance and value of honeybees lay in the various hive products, especially honey, which was used as a sweetener, in making a fermented beverage, in medicine, and in conducting rituals. Beeswax was civilization's first plastic, and was used as a polish, in the preparation of molds for metal casting, for waterproofing, and in many other ways. Some ancient peoples ate honeybee brood, the developing larvae and pupae in the hive. Brood eating has never been widely practiced, but even today the habit is found among some African and Asian peoples.

Fig. 1. A hollow tree, the original, natural home for honeybees; this tree trunk has been split apart to expose the nest. The bees used today are not domesticated and can easily revert to their old ways. Beekeeping is possible because of present-day understanding of bee biology. (*New York State College of Agriculture, Cornell University, Ithaca, New York*)

Honey. Honey was a precious commodity in early Roman, Greek, and Egyptian civilizations. For people 2000 to 4000 years ago, it was the main source of sugar. Thus, in the ancient Mediterranean world, honey and beekeepers were held in extremely high esteem. Today, honey is still valued as a sweetener and is used to some extent in medicine.

Production. The honey which beekeepers harvest is made from nectar, a sweet sap or sugar syrup produced by special glands in flowers, collected from both wild and cultivated plants. Nectar, the honeybees' source of sugar or carbohydrate, and pollen, their source of protein and fat, make up their entire diet.

Nectar contains 50–90% water, 10–50% sugar (predominantly sucrose), and 1–4% aromatic substances, coloring material, and minerals. To transform nectar into honey, bees reduce its moisture content, so that the final honey produced contains between 14 and 19% water, and also add two enzymes which they produce in their bodies. One enzyme inverts the sucrose, a sugar with 12 carbon atoms, into two 6-carbon sugars, levulose and dextrose. The second enzyme added by bees is glucose oxidase, which is inactive when honey has a normal moisture content (14–19%). However, when the honey is diluted, as when it is fed to honeybee larvae, the enzyme becomes active, resulting in the breakdown of dextrose into hydrogen peroxide and gluconic acid. Hydrogen peroxide prevents or slows the growth of bacteria and thus protects the diluted honey from attack by microbes. Glucose oxidase is also active while the bees are in the process of reducing the nectar's moisture content, a process which may take 24 h or more. The gluconic acid produced during this time, along with other acids present in honey, gives honey a pH of 3.5 to 4.0 which, among food products, is quite acid and makes honey an inhospitable medium for bacterial growth. Also, since it is a supersaturated sugar solution, honey has a high osmotic pressure. Therefore, any microorganisms present in honey are plasmolyzed and die.

Use in medicine. Honey has been used in medicines for thousands of years. Since honey is about half levulose, the sweetest of the common sugars, it is sweeter than cane sugar or corn syrup. Even today, honey is used in certain medicinal compounds because the sweet levulose covers up harsh, bitter flavors better than other sugars. The most popular honey for use by drug manufacturers is tupelo honey, which is produced only along the banks of the Apalachicola River in west Florida. Tupelo honey is the sweetest of the commercially produced honeys in the United States. Honeys high in levulose granulate much more slowly than do honeys high in dextrose; this is an added

advantage for the drug manufacturer, who does not want the product to granulate or contain sugar crystals.

Before the discovery of modern salves and creams, honey was widely used as a wound dressing. It was an ingredient in over half of the medicinal compounds and cures made by Egyptian physicians 2000 to 5000 years ago. Honey is an effective dressing, not only because bacteria cannot live in it, but also because honey on a wound creates a physical barrier through which bacteria cannot pass. Some of the ancient Egyptian salves called for a mixture of honey and grease. Lint was sometimes added to help bind the mixture. If the honey was diluted with water, the glucose oxidase became active, producing hydrogen peroxide and again giving the wound some protection. Honey also prevented dried blood from adhering to a bandage and kept the bandage soft. Thus, dressings could be removed from a wound periodically, and the wound treated with more salve and rebandaged.

Use in baking. About half the honey produced in the world is sold as table honey and is used as a sweetener on bread, cereal, or desserts. The remaining half is used in the baking trade. Honey is hygroscopic, that is, it absorbs moisture. When honey is added to a baked product some of this property is retained, helping to keep the bread or cake moist. Bakers are also concerned with the browning qualities of their product, and the sugar levulose in honey has the desired browning qualities. Because the price of honey is many times that of sugar, it is too expensive to use in all baked products. Thus, honey is added only in specialty products.

Biology of the honeybee. Honeybees are social animals. A minimum of about 200 worker bees and a queen must live together for social order to exist; however, such a small number of bees could not survive for long without human assistance.

A colony of bees in an artificial hive, or living wild, contains about 10,000 worker bees, all females, in midwinter. The same colony would have a peak population of 30,000 to 60,000 workers in early summer. In addition to the workers, a normal colony has a single queen and 0 to 3000 drones (males). The sole function of the drones is to mate, and their presence in the hive is not tolerated by the workers during nonmating seasons. In the fall, the drones are deprived of food by the workers and are driven from the hive. The rearing of new drones in the spring starts well after bees have initiated the rearing of workers.

It is generally accepted that, in most situations, all the bees in a colony develop from eggs laid by the single queen. The size and composition of the colony population is regulated by how many eggs are reared during each season. In the Northern Hemisphere, bees rear the least number of young, called brood, in October and November. The least number of adult bees is found in the hive about 3 to 4 months later. Pollen first becomes available in the spring, February in the southern states and April in the northern states. The availability of pollen controls brood rearing, for without large quantities only a small number of young are reared. The natural seasonal population fluctuation which takes place in a colony is referred to as the cycle of the year.

Fig. 2. A swarm which escaped from a beekeeper's hive and took up residence in the side of a house. The house boards have been removed to photograph the swarm. A swarm can live successfully in the side of a house, though the bees may be a nuisance.

Honeybees separate their young and their food, which is one of the factors which makes beekeeping possible. Honey and pollen are usually stored above the brood-rearing area, which is called the brood nest. The brood nest, which contains the eggs, larvae, and pupae, is a compact area more or less the shape of a ball. Brood are reared in many combs at the same time, but always in combs which are side by side. Honeybees maintain a constant brood-rearing temperature of about 92–95°F (24–36°C). This allows them to rear their young in a set period of time (see **table**) regardless of seasonal and climatic temperature variations. A few other species of social insects exercise some control over brood-rearing temperature, but none so precisely as the honeybee.

When cool weather prevails, the bees form an insulating shell of live bees around the outside of the brood nest. Bees in the center of the brood-rearing area make physical movements to generate heat. If the temperature drops, more bees generate heat and the shell of bees around the area becomes more compact. In this manner the heat is retained within the cluster; bees do not heat the inside of the whole hive. Bees can survive severe winters, even in Alaska, if they have sufficient honey. However, because so much honey is required for energy, usually in excess of 100 lb (45 kg) per colony in Alaska, it is considered

Development time for honeybee castes, in days

Stage	Queen	Worker	Drone
Egg	3	3	3
Larva	$5\frac{1}{2}$	6	$6\frac{1}{2}$
Pupa	$7\frac{1}{2}$	12	$14\frac{1}{2}$
Total	16	21	24

uneconomical to keep bees there. In the summer when the temperature is high, bees collect water, deposit it in droplets around the hive, and fan their wings, causing masses of air to pass through the hive, evaporating the water and preventing excessive temperatures in the brood nest. Colonies of bees have survived in the Arizona desert, where temperatures may reach 120°F (49°C). In such cases, a colony may need to collect and evaporate as much as 1 or 2 gal (3.8 or 7.6 liters) of water per day.

Remarkably, the largest, most important individual in the colony, the queen, is reared in the shortest period of time. Evolution has favored this fast development. If a queen dies, a replacement must be produced rapidly or the colony will perish. Queens and workers develop from identical eggs. In fact, a queen can be made from a larva that is from 24 to about 48 h old. Larvae destined to become queens are fed a special food, royal jelly, which is produced by glands in the heads of worker bees; larvae destined to become worker bees receive less of this glandular secretion and more honey and pollen in late larval life. Thus, in a colony that loses its queen but has young larvae present, a queen may be produced in 11 or 12 days. *See* SEX DETERMINATION.

Pollination. Pollination is the process by which pollen, comprising the male germ cells, is transferred from the male part of a flower (anthers) to the female part (stigma) on the same plant. The process of transferring the pollen from the flower of one plant to that of another is called cross pollination. Pollen may be carried by wind, water, or animals. While only a small number of plants are water-pollinated, wind pollination is a fairly common occurrence.

Many plants, however, require animal agents for cross pollination. Certain birds, especially hummingbirds, are especially adapted for this purpose, but insects, especially bees, are the most efficient pollinators. Most bees are solitary insects, living and making their nests in hollow twigs, under bark, in stone crevices, or in tunnels in the ground. These insects pollinate many of the flowers which produce seeds, nuts, and fruit. Some species of social bees, however, including bumblebees and honeybees, are important pollinators. In many areas in the United States the solitary twig- and ground-nesting bees and the bumblebees outnumber honeybees.

The honeybee is especially well adapted to gathering large quantities of pollen and nectar and, in the process, bringing about efficient cross pollination. A honeybee is a very hairy creature, the hairs on its body being plumose, or branched, so that pollen grains are easily caught. When a honeybee collects pollen, it actually wallows in the flower parts, dusting itself thoroughly with pollen. It then hovers above the flower and uses its legs to clean its body of pollen; the pollen is raked into areas of longer hairs, called pollen baskets, one on each of its hindlegs. Bees pack pollen into the baskets until they become large, round balls; bright colored pollen balls may be seen with the naked eye on foraging bees, both as they collect more pollen in the field and as they later enter their hive.

When the pollen balls on the two hindlegs are large enough, the bee carries the pollen back to the hive, where it is deposited in cells and becomes food for adult bees and larvae. A bee never cleans its body so thoroughly that all the pollen grains are removed, and does not clean itself after visiting every flower. Thus, in the process of moving from one flower to another, large numbers of pollen grains are transferred. *See* FLOWER; POLLINATION.

Modern beekeeping. Scientific beekeeping started in 1851 when an American, L. L. Langstroth, discovered bee space and the movable frame hive. Bee space is the open space which is about 0.4 in. (1 cm) wide and maintained around and between the combs in any hive or natural nest and in which the bees walk. If this space is smaller or larger than 0.4 in. (1 cm), the bees will join the combs. When the combs are stuck together, the hive is not movable, and it is not possible for beekeepers to manipulate a colony or to examine a brood nest. Prior to 1851, beekeeping methods were crude and the quantity of honey harvested per hive was probably one-quarter of that obtained today.

It was found, in 1857, that bees could be forced to build a straight comb in a wooden frame by giving them a piece of wax, called foundation, on which the bases of the cells were aready embossed. Bees use these bases to build honeycomb, the cells of which are used for both rearing brood and for storing honey. When a hive of bees is given a frame of foundation, they are forced to build the comb where the beekeeper wants it and not where they might otherwise be inclined to build it.

Another discovery, made in 1865, was that honey can be removed from the comb by placing a comb full of honey in a centrifugal force machine, called an extractor. Approximately 16 lb (7 kg) of honey must be consumed by the bees in order for them to make 1 lb (0.45 kg) of wax comb. If the beekeeper can return an intact comb to a hive after removing the honey from it, the bees are saved the time and trouble of building a new comb, and the honey harvest is increased.

The next discovery, in 1873, was the modern smoker. When bees are smoked, they engorge with honey and become gentle. Without smoke to calm a hive, normal manipulation of the frames would not be possible. While smoke has been used to calm bees for centuries, the modern smoker has a bellows and a nozzle which allows a beekeeper to use smoke more effectively by directing the smoke where it is needed.

Within the short span of 22 years, these four discoveries laid the basis for the modern beekeeping industry. Prior to 1851, only a few individuals in the United States owned a hundred or more colonies of bees. By 1880, several people owned a thousand or more colonies each, and honey, which had once been a scarce commodity, became abundant.

Hives. All beehives used in the industry today are made of wooden boxes (**Fig. 3**) with removable frames of comb, and 90% of these are 10-frame Langstroth hives. Beekeepers have standardized their

Fig. 3. A typical modern-day commercial apiary. Beekeepers often select apiary sites surrounded by woods so as to protect bees from damage by wind; hidden apiaries are also less likely to be vandalized.

equipment so that parts will be interchangeable within an apiary and from one commercial operation to another. The dimensions of the modern-day hive (**Fig. 4**) were determined in 1851 by Langstroth, who made a box based on the size of lumber available to him and on what he thought was an easy unit to lift and handle. The honeybee is remarkably adaptable. While it may show a preference for a nest of about 1.4 ft^3 (40 liters) in volume, it will live almost anywhere it can keep its combs dry and protected.

Commercial practice. To be successful in commercial beekeeping, beekeepers must move to those areas where nectar-producing plants abound. The best-known honey in the United States is clover honey. Clovers grow in almost every state and, where abundant, they may produce large quantities of nectar. Alfalfa is also a good nectar-producing plant and it too is a major source of honey. Oranges grown in Florida, Arizona, and California are also known for the high quantity and quality of nectar which they produce. Just as each flower is different in its color, shape, and design, so each flower produces a unique nectar. Thus, orange honey does not taste like clover honey, and clover honey is different from the honey which is produced by buckwheat, sage, gallberry, and so on. Bees collect nectar from hundreds of kinds of flowers, and thus there are a great variety of honey flavors.

In agricultural regions where land use is intense, large areas of land are cleared, removing hedgerows and woodlots in which many insects nest and grow. Large acreages of such monocrops as apples, oranges, alfalfa, almonds, cantaloupes, and watermelons are planted, decreasing the diversity of the vegetation and eliminating many indigenous insect pollinators. In these areas there are few solitary bees.

Thousands of colonies of honeybees are rented each year by growers of crops needing cross pollination. When the plants are coming into bloom the beekeeper moves the bees, usually at night when all the bees are in the hive, and places them in groups in groves, orchards, and fields. While beekeepers make most of their living producing honey, the real importance of their bees in the agricultural economy

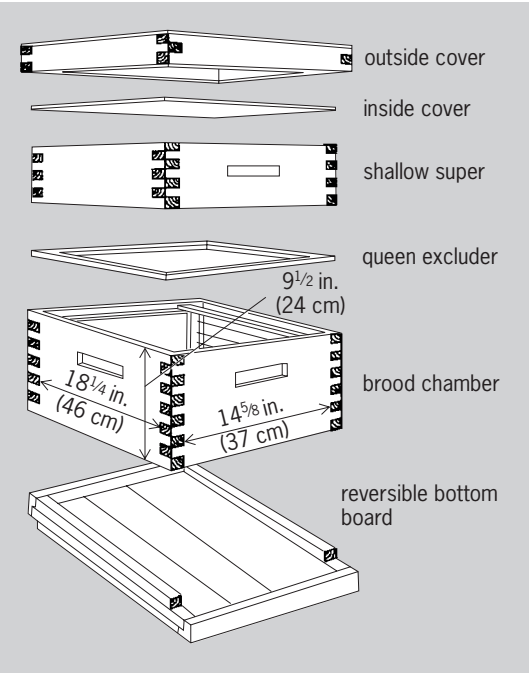

Fig. 4. Construction and dimensions of a 10-frame beehive. (*Apiculture Research Branch, Plant Industry Station, Agricultural Reseach Service, USDA, Beltsville, Maryland*);

is as cross pollinators. Without cross pollination the abundance and variety of food and flowers would not exist.

It is estimated that there are about 2000 commercial, full-time beekeepers and 200,000 persons for whom beekeeping is a part-time job or hobby in the United States. As mentioned, to be successful in beekeeping one must move to areas where nectar-producing plants are numerous. One may plant some flowering trees, bushes, and ground plants on which bees may forage, but a colony needs so much foraging area during the course of a year that a beekeeper cannot plant all that is needed. Most commercial beekeepers agree they need to harvest an average of 60 to 100 lb (27 to 45 kg) of honey per colony per year to be successful. In addition 30 to 60 colonies are needed in an apiary to be economically efficient. Not all locations have sufficient nectar forage to support that number of colonies. Commercial beekeepers feel they need to own 500 to 2000 colonies to make a living, depending on how they market their honey. Some beekeepers sell their honey on the wholesale market only, while others pack their honey themselves and devote a great deal of time to sales. *See* AGRICULTURAL SCIENCE (ANIMAL). Roger A. Morse

Beet

The red or garden beet (*Beta vulgaris*), a cool-season biennial of Mediterranean origin belonging to the plant order Caryophyllales (Centrospermales). This beet is grown primarily for its fleshy root, but also for its leaves, both of which are cooked fresh or canned as a vegetable (see **illus.**). The so-called seed is a fruit containing two to six true seeds. Detroit Dark Red strains predominate. Beets are only slightly tolerant of acid soils and have a high boron requirement. Cool weather favors high yields of dark red roots. Exposure to temperatures of 40–50°F (4–10°C) for 2 weeks or more encourages undesirable seed-stalk development. Roots are harvested when they are 1–3 in. (2.5–7.5 cm) in diameter, generally 60–80 days after planting. Wisconsin, New York, and Texas are important beet producing states. *See* CARYOPHYLLALES; SUGARBEET. H. John Carew

Fleshy roots and leaves of beet plants. (*Asgrow Seed Co., Subsidiary of the Upjohn Co.*)

Behavior genetics

The hereditary factors of behavior may be studied in animals and humans. Charles Darwin, who originated the theory that natural selection is the basis of biological evolution, was persuaded by Francis Galton that the principles of natural selection applied to behavior as well as physical characteristics. Members of a species vary in the expression of certain behaviors because of variations in their genes, and these behaviors have survival value in some environments. One example of such a behavior is curiosity—some organisms are more curious than others, and in some settings curiosity is advantageous for survival. Therefore, in those environments more of the organisms that express curiosity survive to reproduce than do those that lack this trait.

The science of behavior genetics is an extension of these ideas and seeks (1) to determine to what extent the variation of a trait in a population (the extent of individual differences) is due to genetic processes, to what extent it is due to environmental variation, and to what extent it is due to joint functions of these factors (heredity-environment interactions and correlations); and (2) to identify the genetic architecture (genotypes) that underlies behavior.

Traditionally, some of the clearest and most indisputable evidence for a hereditary influence on behavior comes from selective-breeding experiments with animals. Behavior genetic research has utilized bacteria, paramecia, nematodes, fruit flies, moths, houseflies, mosquitoes, wasps, bees, crickets, fishes, geese, cows, dogs, and numerous other organisms. Breeding of these organisms allows genetically useful types of relationships, such as half-sibs, to be produced easily. Although not all of this work involves breeding experiments, enough of it does to demonstrate conclusively the importance of genetic processes on behavior in every species examined. Artificial selection (selective breeding) can be used to obtain a population that scores high or low on specific traits. Inbred strains of many animals (populations that are made up of individuals which are nearly identical genetically as a result of inbreeding), particularly rodents, are readily available, and the study of various types of crosses among them can provide a wealth of information. An experimental design using the recombinant inbred-strain method shows great promise for isolating single-gene effects. This procedure derives several inbred strains from the F_2 generation (grandchildren) produced by a cross between two initial inbred strains. Since it is possible to exert a great deal of control over the rearing environments, the experimenter can manipulate both heredity and environment, a feat virtually impossible in human studies but required for precise answers to behavior genetic questions.

Other work has focused on the effects of the environment and genotype-environment interactions.

For example, experiments with mice have shown that, with respect to several learning tasks, early environmental-enrichment effects and maternal effects were quite small, relative to the amount of normal genetic variation found in the strains of mice tested. Only a few genotype-environment interactions were found. Still other work has shown that early experiences affect later behavior patterns for some strains but not others (a genotype-environment interaction).

An increasing role for animals in genetic research is to provide models of human genetic diseases, many of which have behavioral features. Such animal models may occur naturally or may be engineered in the laboratory. For example, a mouse model was produced which is deficient in the enzyme HPRT. Deficiency of this enzyme in humans is due to a recessive X-linked gene and leads to a disorder known as Lesch-Nyhan syndrome, which is characterized by self-injurious behavior and mental retardation. Animal models are available for many other neurobehavioral disorders, including narcolepsy, various epilepsies, and alcoholism. The availability of animal models allows researchers to obtain information about the development of genetic disorders and the effects of different environments on this development, as well as to explore treatment options. While it is not always prudent or desirable to generalize from animal results to humans, it is assumed that basic genetic systems work in similar ways across organisms, and it is likely that these types of animal studies will play a key role in elucidating the ways in which environment influences phenotypic variation. With advances in genetic technology, it is possible to observe genetic variation more directly by locating, identifying, and characterizing genes themselves.

Chromosomes and Genes

In most organisms, chromosomes are linear structures comprising strands of deoxyribonucleic acid (DNA) and accompanying proteins. A gene is a specific segment of DNA representing the genetic information for making a needed protein or for regulating other genes. Each gene may come in several variations, called alleles, that may produce somewhat different phenotypic outcomes. An organism's phenotype, or complement of observable properties, is due to the combined effects of its genetic information and the environmental conditions it experiences. *See* CHROMOSOME; GENE.

Each human inherits 23 paternal chromosomes and 23 maternal chromosomes at fertilization. Genes occupy (fairly) specific positions on these chromosomes. Located in the nucleus of each cell, the 46 chromosomes can be thought of as comprising 23 pairs, with one member of each pair coming from each parent; for example, the maternal chromosome 1 pairs with the paternal chromosome 1, the maternal chromosome 15 pairs with the paternal chromosome 15, and so on. These are called homologous pairs. Thus, at the same location on each member of the pair, there is a gene for a particular kind of genetic information, or two genes of that kind per individual. Since each chromosome in the pair comes from one parent, each gene in the pair comes from one parent. This is true for the first 22 pairs of chromosomes, called the autosomes. The twenty-third pair is the sex chromosomes, and they may or may not be the same. A female has two X chromosomes (designated 46, XX), but a male has one X chromosome and one Y chromosome (46, XY). The Y chromosome contains crucial genetic information that causes an embryo to develop as a male. Thus, a mother can contribute only an X chromosome to her offspring, but a father can contribute either an X or a Y chromosome. It is, therefore, the father who determines the sex of the offspring. *See* HUMAN GENETICS; SEX DETERMINATION.

The Y chromosome is smaller than the X chromosome and contains less genetic material. Consequently, few genes that reside on the X and Y chromosomes are paired. Traits influenced by genes carried only on the X chromosome are called X-linked, while genes carried only on the Y chromosome are called holandric genes. Red-green color blindness is under the control of a recessive variant (allele) of a gene on the X chromosome. This allele is called recessive, and the gene for normal color vision is called dominant because, in females, the gene for red-green color blindness must be present on both X chromosomes in order for the female to be red-green color-blind. One dominant normal color vision gene can override one defective gene. In the male, however, the situation is different, because this X-chromosome gene does not have a matching gene on the Y chromosome. If the X chromosome carries the defective version of the gene and is matched with a Y chromosome, that individual is red-green color-blind. Females can only be red-green color-blind if both genes at that location are recessive, so this behavioral trait is more frequent among males than females. *See* DOMINANCE; SEX-LINKED INHERITANCE.

Chromosome abnormalities. There are a number of abnormalities involving entire chromosomes or pieces of chromosomes, rather than isolated genes, that influence behavior, and many important ones involve the sex chromosomes. The reason is that most abnormalities of the autosomes are lethal during fetal development. It is estimated that as many as 50% of all conceptuses have gross chromosomal abnormalities, while these defects are found in only 0.5% of newborns. It seems that an excess of chromosomal material is better tolerated than a deficit; even so, when a newborn is found to have an extra autosome, it is usually one of the smaller chromosomes, such as 13, 18, or 21. *See* CHROMOSOME ABERRATION; CONGENITAL ANOMALIES.

Autosomal chromosome anomalies. The abnormal presence or absence of autosomes is usually lethal in the early stages of fetal development. There are, however, several exceptions. The best known is Down syndrome, or trisomy 21, in which there is an extra twenty-first chromosome (47, +21), occurring in about 1 in 700 live births. The physical features of these individuals are striking: round face with

broad head, small medial epicanthal fold, flattened bridge of the nose, small ears and nose, and protruding tongue. Skeletal and congenital heart defects are common, as is increased susceptibility to respiratory infections. Down children have a reduced life expectancy, although modern antibiotics and heart surgery have improved this somewhat. Although the IQ of these individuals ranges from less than 25 to about 75, with a mean of 40–50, many can be taught to read and write and to care for themselves.

Other kinds of chromosomal abnormalities exist, including duplications, rearrangements, and deletions of various amounts of chromosomal material. In most cases, the more chromosomal material involved, the more deleterious the outcome. However, it is becoming apparent that even very small deletions and duplications can lead to serious disorders, although these usually are considered single gene disorders rather than chromosomal abnormalities. For example, the abnormal repetition of three of the units (base pairs) that make up the gene's DNA has been implicated in, among others, Fragile X, Huntington's disease, a progressive neurological disorder, and myotonic dystrophy, characterized by muscle spasms and muscle wasting, cataract, and other features.

Sex chromosome anomalies. Most sex chromosome anomalies involve the presence of one or more extra X or Y chromosomes, such as XXX, XXY, XXXX, or XYY. An exception is Turner syndrome, in which only one complete X chromosome is present (45, X). [In some cases, part of a second X may be present. In a few individuals, some cells have two X chromosomes while others have only one; these individuals are referred to as mosaics.] Each type of sex chromosome anomaly has specific features or abnormalities associated with it. For example, most individuals with Turner syndrome have only one sex chromosome, an X, but are phenotypically female. Physical features commonly found in these individuals include webbed neck, short stature (usually less than 5 ft or 150 cm), streak gonads, infantile genitalia, shield chest, and widely spaced nipples. Pubic and underarm hair may or may not be present. Since the ovaries do not develop, the women do not menstruate and are sterile. However, not all of these features are present in all individuals with Turner syndrome. There are certain behavioral characteristics that have been associated with women having Turner syndrome, such as a lack of aggressive tendencies, reduced libido, and strong inhibitions; today, however, there is some question about whether these characteristics are due in part to the way these short, immature-looking women are treated by others. Sex hormone therapy can help to increase height somewhat and to induce physical maturation; such changes may have psychological benefits as well. Intelligence is usually in the normal range in Turner syndrome, although the individuals may have deficits in space-form recognition and directional sense. *See* DOWN SYNDROME.

Another type of sex chromosome abnormality leads to the Fragile X syndrome. The name derives from a fragile site on the long arm of the X chromosome which can be seen when cells are grown under special conditions. Affected males usually have some characteristic facial features, enlarged testes, and mental retardation. The Fragile X syndrome is the most frequent known cause of mental retardation in males. Females with one Fragile X chromosome and one normal X chromosome may have a lesser degree of mental retardation.

Variations in genes. Far more common than chromosome abnormalities are variations in the genes themselves. Most of these variations lead to differences among individuals which are considered normal, for example, variations in eye color. Some gene variations, however, lead to differences which are not considered normal, such as Huntington's disease, albinism (lack of pigment in hair, skin, and eyes), or hemophilia (lack of blood clotting factors). It is now known that some differences among individuals are due to variations in single genes, while others are due to the combined effects of two or more genes (polygenic or continuous characters). It must also be remembered that the environment is always involved in the expression of a trait.

Single-gene effects. The effects of a single gene on behavior have been most extensively studied in the domain of mental retardation. Research has shown that there are a large number of metabolic pathways (series of chemical reactions used by cells to make or break down molecules) which have defects due to a single gene. Over 100 of these defects influence mental ability. *See* MENTAL RETARDATION.

One such single-gene defect is classic phenylketonuria (PKU), an autosomal recessive disorder, which also illustrates the role that environment can play in the expression of a trait. Individuals who are homozygous (having two copies of the PKU allele) are unable to make the enzyme phenylalanine hydroxylase, which converts the essential amino acid phenylalanine to tyrosine, a nonessential amino acid. Instead, the excess phenylalanine builds up in the blood and is converted into phenylpyruvic acid, which is toxic to the developing nervous system in large amounts. The main effect of untreated PKU is severe mental retardation, along with a distinctive odor, light pigmentation, unusual gait and posture, and seizures. Many untreated individuals with PKU show fearfulness, irritability, and violent outbursts of temper. The treatment for PKU consists of a diet that severely restricts phenylalanine intake. If the diet is carefully maintained from the first few weeks of life through about 8–10 years of age, PKU individuals have a good chance of developing normal or near-normal intelligence. Most states in the United States now require that newborns be tested for PKU before leaving the hospital. *See* PHENYLKETONURIA.

Polygenic effects. Most of the traits of interest to behavioral scientists are continuous (ranging from low to high, such as intelligence) rather than discontinuous (for example, affected or unaffected with PKU). They are continuous because they are under the control of many genes, each having a small effect (polygenic). Thus, a trait governed by a single gene usually

shows two or three discrete phenotypic categories. A polygenic trait shows a distribution of phenotypes that approximates a normal curve. For some types of behavior, the question of whether it is caused by a single gene or whether it is a polygenic trait has not been settled, and the role played by the environment is at issue.

Nature or Nurture Question

Every organism develops in a particular environment, and both genes and environment control development. It is, therefore, not possible to state that a particular behavioral trait is either genetic or environmental in origin. It is possible, however, to investigate the relative contributions of heredity and environment to the variation among individuals in a population. With humans, it is possible to obtain approximate results by measuring the similarity among relatives on the trait of interest.

The most widely used measure of similarity is the correlation coefficient. This measure varies from -1 to $+1$. If a large number of pairs of people were selected at random from a population and then compared, each pair's scores on some measure (for example, a physical variable such as height, or a psychological variable, such as extroversion), the degree of similarity among them, as measured by a correlation coefficient, would be zero. The reason is that, on the average, they would be genetically unrelated and would have been reared in different environments, and therefore would be unlikely to be similar on that measure, regardless of the influence of heredity or environment. In order to obtain information about these influences, scientists compare individuals who are related or who have shared environments. Because related individuals usually share some environments as well as genes, it is useful, but not absolutely necessary, to compare the results with those from the relatives that have been reared apart.

Twin and adoption studies. Twins are often used in behavior genetic studies. One method compares the similarity within pairs of both identical (monozygotic or one-egg) twins and fraternal (dizygotic or two-egg) twins reared together. Identical twins have all their genes in common by descent, since they arise from a single fertilized egg. Fraternal twins arise from two fertilized eggs and so, like any pair of nontwin siblings, share on average one-half of their genes. (Parents and offspring share exactly half of their genes, as described earlier.) If it is assumed that the effects of the shared environments of the two types of twins are equal (a testable assumption), greater resemblance between identical twins than fraternal twins should reflect the proportion of genes they share, and the difference between the correlations of the two twin types should represent about one-half the genetic effect.

A second type of twin study compares not only twins reared together but twins who have been reared apart. The degree of similarity between identical twins reared in the same home would reflect the fact that all their genes are identical and that they share a common family environment. On the other hand, if identical twins can be located who had been adopted by different families chosen at random (an unlikely event, since adopted children tend to be selectively placed), a measure of their degree of similarity would reflect only the effect of their common genes. If it were true that an individual's level on a measure (for example, height or extroversion score) is determined in large part by the characteristics of his or her family and the opportunities that the family makes available to him or her, reared-apart identical twins should be no more alike than pairs of individuals chosen at random. If they do exhibit some degree of similarity, it would reflect genetic effects alone. The existence of even very large genetic effects, however, would in no way imply that the environment was unimportant in the development of the trait; it would simply imply that environment was less important than genes in determining the variation among individuals on the trait in question at the time of measurement. That is, the individuals would differ more because of the genes they carry than because of the particular environments to which they were exposed. In another range of environments, the results might be different.

Another method of determining the relative importance of genes and environment is to compare the degree of similarity of unrelated individuals raised in the same family: adopted children and their adoptive parents and siblings. Any similarity between these individuals would be due entirely to common family environment and not to genes (barring placement in similar families). Family, twin, and adoption studies have their respective advantages and limitations; when results from different kinds of studies are similar regarding a particular trait, they can be viewed with increased confidence.

In addition, biometric and other quantitative methods allow information from many types of family relationships to be processed simultaneously. These approaches estimate the different genetic and environmental components of variance for a trait. They do this by estimating the phenotypic variances and covariances of measurements from different types of biological and adoptive relationships, for example, biological parent-offspring, adoptive parent-offspring, biological sib-sib, identical twins, fraternal twins, and so on. These estimates then allow the fitting or testing of models which predict certain amounts of genetic and environmental variance. Finding the best-fitting model leads to a better understanding of the underlying genetic (and environmental) architecture of a trait. Finally, modern genetic technology is being applied to single-gene conditions with behavioral components. One approach analyzes DNA from populations of affected and unaffected individuals to determine if a known, or candidate, gene (chosen because its function conceivably could contribute to the observed phenotype) could be the gene in question. Abnormal alleles of genes affecting neurons in the brain, such as those coding for neurotransmitters and their receptors and for ion channels, among others, are being explored as candidate genes for certain behavioral disorders. Several

candidate genes for schizophrenia have been ruled out. *See* BIOMETRICS; TWINS (HUMAN).

Gene mapping. Gene mapping or linkage analysis is a technology in which a gene is localized to a specific position on a chromosome. The DNA from affected and unaffected members of families is analyzed to find an association between the disease state and the presence of a gene or a piece of DNA whose position is known, called a marker. The assumption is that if the association is very high it is due to the fact that the gene involved in the trait under investigation and the marker DNA are physically close (linked) on the chromosome. Various techniques can then be used to identify the DNA that makes up the gene itself. One of the very first genes to be localized to a particular chromosome by this method was that for the progressive neurological disorder Huntington's disease (located on chromosome 4). Once a gene has been identified, its function can be determined and studied and treatment options considered. Linkage analysis is applied most often to the study of genetic disorders, also can be used to learn more about genes contributing to normal variation.

Gene mapping is a powerful tool that is widely applied. However, it cannot always provide answers to genetic questions. Its success depends upon several factors, including understanding of the mode of transmission of the gene (for example, whether the allele of interest is dominant or recessive, on an autosome or the X chromosome), having the correct diagnosis, and assuming that the trait in question is due to the same genes in all the individuals who have the trait (genetic homogeneity). Alzheimer's disease provides an illustration of genetic heterogeneity, being found in familial and nonfamilial as well as early- and late-onset forms. Three chromosomal locations have been identified among groups of families with inherited Alzheimer's disease—chromosome 19 for late-onset, and 14 and 21 for early-onset types. In addition, gene mapping requires that markers for the appropriate chromosomal locations are available and that the families being studied show some variation in allelic forms of these markers. *See* ALZHEIMER'S DISEASE.

Behavioral traits often present special challenges for this type of analysis. The mode of inheritance is often not known with confidence. It is widely believed that many, if not most, behavioral traits are polygenic in nature. If each gene involved has only a small effect, it will be difficult to identify the genes by mapping alone. However, biometric techniques developed for analysis of polygenic traits in plants (quantitative trait loci, or QTL, methods) can be used in animals, sometimes in conjunction with recombinant inbred strains. Another potential difficulty with behavioral traits is making an accurate, or at least consistent, diagnosis. Inclusion of too many or too few characteristics in a diagnosis will lead to incorrect categorization of some individuals in each family studied and will produce spurious results. Finally, it is thought that genetic heterogeneity may be common in behavioral traits; if it is not identified, results of linkage analyses may be inconclusive.

Environmental variables. Developmental psychologists are finding that differences in children's behavioral phenotypes are due more to their different genotypes than to their different rearing environments, as long as those environments are within a normal range of experiences. Identifying environmental variables from this normal range that have an important effect on the behavioral phenotype may be even more difficult than identifying contributing genes. Advances in theory and new technologies, combined with information from more traditional methodologies, will continue to provide insight into the contributions of genes and environment to behavior.

Intelligence. The results of over 50 years of research on the similarity in intelligence between relatives (as estimated by intelligence, or IQ, scores) indicate that identical twins reared apart are remarkably similar relative to adopted individuals reared in the same home (see **illus.**). Identical twins reared together are much more similar than fraternal twins reared together, and dramatically more similar than unrelated individuals reared together. This general evidence clearly suggests that heredity is an important determinant of population variation in intelligence. It has been determined that approximately40–70% of the variation in intelligence can be explained by hereditary factors. *See* INTELLIGENCE.

Special mental abilities. Intelligence (IQ) tests, while controversial, are felt by many to be a measure of general intelligence. Other special mental abilities exist, such as spatial orientation (the ability to rotate three-dimensional figures mentally) and perceptual speed and accuracy (the ability to detect errors in text quickly). It is possible to ask if profiles of abilities (patterns of strengths and weaknesses) are influenced by heredity.

Scores on the various subtests of the Wechsler intelligence scales can be thought of as reflecting measures of some of these special mental abilities. Two studies examined the correlations of the patterns of subtest scores in young twins. The first used the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) to measure 70 pairs of identical and 46 pairs of fraternal 6-year-old twins. The correlations are 0.43 and 0.27, respectively. The second study measured 69 identical and 35 fraternal pairs, with a mean age of 8 years, on the Wechsler Intelligence Scale for Children (WISC-R) and obtained correlations of 0.45 and 0.24, respectively. Data from the Minnesota Study of Twins Reared Apart, using the Wechsler Adult Intelligence Scale (WAIS) score from 32 pairs of identical twins reared apart and 13 pairs of fraternal twins reared apart, provide correlations of 0.52 and 0.12. The finding that the fraternal twins are approximately half as similar as the identical twins, and the similarity of results across age groups, despite the use of different instruments, strongly support the conclusion of a significant genetic influence on the patterning of mental abilities.

Developmental genetics and behavior. The Louisville Twin Study followed nearly 500 pairs

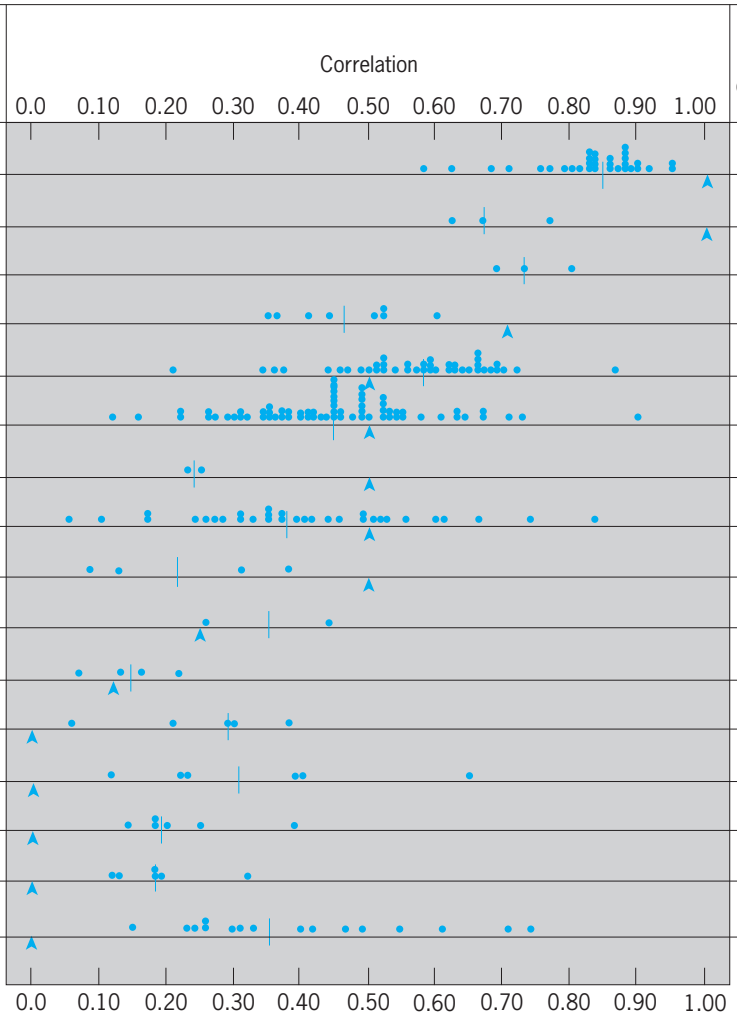

	Number of correlations	Number of pairings	Median correlation	Weighted average
Monozygotic twins reared together	34	4672	0.85	0.86
Monozygotic twins reared apart	3	65	0.67	0.72
Midparent-midoffspring reared together	3	410	0.73	0.72
Midparent-offspring reared together	8	992	0.475	0.50
Dizygotic twins reared together	41	5546	0.58	0.60
Siblings reared together	69	26,473	0.45	0.47
Siblings reared apart	2	203	0.24	0.24
Single parent–offspring reared together	32	8433	0.385	0.42
Single parent–offspring reared apart	4	814	0.22	0.22
Half-siblings	2	200	0.35	0.31
Cousins	4	1,176	0.145	0.15
Nonbiological sibling pairs (adopted/natural pairings)	5	345	0.29	0.29
Nonbiological sibling pairs (adopted/adopted pairings)	6	369	0.31	0.34
Adopting midparent–offspring	6	758	0.19	0.24
Adopting parent–offspring	6	1397	0.18	0.19
Assortative mating	16	3817	0.365	0.33

Familial correlations for IQ. The vertical bar in each distribution indicates the median correlation; the arrowhead, the correlation predicted by a simple polygenic model. (*After T. J. Bouchard, Jr., and M. McGue, Familial studies of intelligence: A review, Science, 212(4498):1055–1059, 1981*)

of twins since infancy as part of a longitudinal study of mental development. The results showed that developmental profiles—the spurts and lags in mental development over time, as measured by several age-appropriate tests—are more similar in identical twins than in fraternal twins. This evidence suggests that genes turn on and off during the course of mental development and influence behavior much the same way as they influence physical development (for example, children also show spurts and lags in growth). The Louisville data also show that the degree of similarity between identical twins, and between parents and offspring, increases over time; the degree of similarity between fraternal twins, while high very early in life, decreases. The correlations between both types of twins at age 15 are very close to those expected under a simple polygenic model.

These synchronous patterns of change show that mental development is under the control of genetic processes that unfold throughout childhood and adolescence. While the environment provided by parents is important in allowing children to maximize their developmental potential, the pattern of growth and the ultimate degree of similarity between relatives strongly reflect the effects of their shared genes. Identical twins reared in different families from a very young age are almost as similar on some traits as identical twins reared together.

Personality. Personality inventories measure such characteristics as introversion and extroversion, well-being, aggression, risk-taking and traditionalism. Many psychologists, while accepting the possibility of genetic influences on intelligence and mental abilities, believe that common family environmental factors are the principal determinants of personality traits. The evidence is, however, quite to the contrary. A number of adoption studies have demonstrated only modest degrees of similarity between genetically unrelated individuals who are reared together. A study comparing reared-apart and reared-together identical twins on a number of personality, psychological interest, and social attitude scales found that the correlations on each measure were very similar for both groups of twins (see **table**). These findings suggest that genetic differences

Intraclass correlations for Minnesota identical twins reared apart and together*

Variables	Reared apart		Reared together		
	Ratio	Number of pairs	Ratio	Number of pairs	Reared apart/ reared together ratio†
Personality variables					
Mean of 11 Multidimensional Personality Questionnaire scales	0.50	44	0.49	217	1.02
Mean of 18 California Psychological Inventory scales	0.48	38	0.49	99	0.98
Psychological interests					
Mean of 23 Strong Campbell Interest Inventory scales	0.39	52	0.48	116‡	0.81
Mean of 17 Minnesota Occupational Interest scales	0.40	40	0.49	376	0.82
Social attitudes					
Mean of 2 religiosity scales	0.49	31	0.51	458	0.96
Mean of 14 nonreligious social attitude items	0.34	42	0.28	421	1.21
Traditionalism scale	0.53	44	0.50	217	1.06

* After T. J. Bouchard, Jr., et al., Sources of human psychological differences: The Minnesota Study of Twins Reared Apart, *Science*, 250:223–228, 1990.
† Ratios of < 1.0 suggest some effect of shared environment.
‡ This value is for 116 studies, not pairs.

account for about half the group variability for these characteristics and that common family environment did not make the reared-together twins any more similar than the reared-apart twins. The remainder of the variability (and the reason the identical twins do not have a correlation of 1.0) must be due largely to the unique or unshared environments experienced by each twin. *See* PERSONALITY THEORY.

Psychopathology. While twin, family, and adoption studies support the role of the genes in the development of schizophrenia and the affective disorders (depression and manic-depression), the environment plays an important role: among identical twin pairs in which one twin is affected, the cotwin is often unaffected. However, if the offspring of members of such discordant twin pairs are examined, the risk of being affected is the same in the offspring of the unaffected twin as in those of the affected twin, demonstrating that the unaffected twin did indeed carry the disease alleles. Early excitement generated by reports of mapping schizophrenia to chromosome 5 and manic-depressive illness (or bipolar disorder) to chromosomes 11 and X has been curbed by failure to replicate findings in other affected families or even in the same families at a later time. While reports of linkage of manic-depressive illness to the X chromosome have seemed more robust than linkage of the disorder to chromosome 11, a reanalysis of several large families finds weakened support for a gene on the X chromosome in two of the three families. Reasons for the revised findings in these studies of psychopathologies include the possibility of incorrect diagnoses in previous studies; individuals becoming affected who were not previously, since many behavioral disorders often do not show up until adulthood; and the use of DNA markers for the later analyses (earlier ones were often based on phenotypic markers). Availability of informative DNA markers for all regions of all the chromosomes will eventually be realized, in large part through the Human Genome Initiative. Researchers will continue to look for genes that contribute to these disorders. *See* AFFECTIVE DISORDERS; SCHIZOPHRENIA.

Alcoholism. As in the psychopathologies, data accumulated over many years from twin, family, and adoption studies point to genetic influences on alcoholism. This view is also supported by the ability to breed selectively for sensitivity to ethanol and for ethanol-related behaviors in animals such as rats and mice. In humans, certain alleles of genes that code for the alcohol-metabolizing enzymes, alcohol dehydrogenase and aldehyde dehydrogenase, are found in different frequencies in alcoholic and nonalcoholic Chinese men. In spite of these findings, there has been little evidence for linkage of alcoholism to a particular gene or genes. Alcoholism is a very complex disorder, and it is expected that many genes, some conferring primary risks and others that interact with them, as well as numerous environmental variables and gene–environment interactions, are contributory. *See* ALCOHOLISM; PSYCHOLOGY. Kimerly J. Wilcox; Thomas J. Bouchard, Jr.

Bibliography. E. L. Gershon and R. O. Reider, Major disorders of mind and brain, *Sci. Amer.*, 267(3):126–133, 1992; R. Plomin, J. Owen, and P. McGuffin, The genetic basis of complex human behavior, *Science*, 264:1733–1739, June 1994; S. Scarr, Developmental theories for the 1990s: Development and individual differences, *Child Dev.*, 63:1–19, 1992; R. S. Wilson, The Louisville twin study: Developmental synchronies in behavior, *Child Dev.*, 54(2):298–316, 1983.

Behavioral ecology

The branch of ecology that focuses on the evolutionary causes of variation in behavior among populations and species. Thus it is concerned with the adaptiveness of behavior, the ultimate questions of why animals behave as they do, rather than the proximate questions of how they behave. The

principles of natural selection are applied to behavior with the underlying assumption that, within the constraints of their evolutionary histories, animals behave optimally by maximizing their genetic contribution to future generations. Several aspects of both individual and social behavior have been analyzed from this perspective as the following examples illustrate.

Behavioral homeostasis. Animals must maintain their internal physiological conditions within certain limits in order to function properly, and often they do this by behavior. Small organisms may avoid desiccation by living under logs or by burrowing. Many insects must raise body temperatures to 86–95°F (30–35°C) for effective flight, and achieve this by muscular activity, such as the shivering of butterflies in the early morning, or by orienting to the Sun. Butterflies spread their wings to bask in the sun and close them when warmed up. Dragonflies, grasshoppers, and lizards orient themselves perpendicular to the Sun's rays to warm up and in parallel to cool down. Such behavior may raise body temperatures 18–27°F (10–15°C) above the surrounding air. Desert and grassland mammals burrow to avoid high or low temperatures. The elaborate burrow system of prairie dogs provides an efficient cooling system. Some burrow mouths are in mounds so that as warm air rises from them, cool air is sucked in via lower entrances a short distance away. Honeybees regulate hive temperature by fanning at the entrance or fetching water for evaporative cooling.

Habitat selection. Animals can be very selective about where they live. In England, for example, the peppered moth occurs in two forms; a black form lives in sooty industrial areas and a mottled or "peppered" form in pollution-free woodlands. The black form prefers to rest on dark tree trunks and the peppered form on gray patches of lichens. By making these choices, each matches its background and is protected from predators. The mechanisms of habitat choice are often genetically influenced, but may be modified by experience. Cases in point are prairie and woodland deermice, which prefer their own habitats, but less strongly so if each is reared in the habitat of the other. *See* PROTECTIVE COLORATION.

An extreme case of habitat selection is migration, which evolves largely in impermanent habitats, allowing animals to escape deteriorating conditions, colonize new areas, and exploit seasonal breeding sites. Insects, such as water striders, common in ponds and streams, have evolved winged and wingless forms. The former occur in temporary ponds or intermittent streams, while the latter are found on permanent ponds or lakes. Some species have a wingless summer generation which channels energy to reproduction rather than flight, and a winged fall generation which flies to protected overwintering sites. Monarch butterflies that breed in eastern North America fly thousands of miles to overwinter in the mountains of Mexico. Insect migrants are adapted for colonizing by their ability to fly and to reproduce rapidly.

The migrations of birds may also have evolved to allow escape from impermanent habitats. Birds overwintering in the tropics usually occupy marginal or stressed environments in their winter ranges. They also exploit temperate zone habitats only seasonally available. The latter behavior is also true of fish. The breeding sites of oceanic herring, for example, occur in seasonal shallow-water habitats, and the fish migrate to nursery and feeding areas. The streams to which salmon migrate are also seasonally limited. The explosive and once-only reproduction of Pacific salmon may result because repeated migrations up long steep rivers are too costly. *See* MIGRATORY BEHAVIOR.

Foraging. All animals must find food to survive and reproduce. Caterpillars graze on plants, predatory starfish pull mollusks from their shells, and chimpanzees use sticks as tools to draw ants from their nests. Despite the diversity of foods and feeding behavior, ecological theory predicts that natural selection should favor animals that forage efficiently. Those who get the most food energy in the least time will have the most time and energy available for other activities.

Animals are constantly faced with choices about what to eat and where and how to forage. Northwestern crows, for example, feed on shellfish during low ocean tides. Crows have been observed picking up whelks, flying 15 ft (5 m) above the ground, and dropping their prey on the rocks to break open the shells and expose the meat. Crows pick only the largest whelks, which break more easily and provide more food than small whelks. Moreover, 15 ft (5 m) is the minimum height to ensure that whelks will break on the first drop. Northwestern crows, then, apparently choose a foraging strategy that provides them with the maximum energy for the minimum effort.

Nutrient requirements and predation pressures may alter an animal's feeding strategy. Moose, for example, need sodium in their diet. Sodium is most abundant in energy-poor aquatic plants; moose divide their foraging time between aquatic plants and energy-rich terrestrial plants. Seed-eating dark-eyed juncos must constantly stop foraging and scan the skies for predatory hawks. The time that juncos divide between feeding and watching reflects the trade-off between getting necessary food energy and ensuring that they will survive.

Groups. Animals of all sizes and from many taxa form groups such as herds, troops, schools, and pods (of whales). A major theory as to why they do this is the "selfish herd" hypothesis; that is, animals will group in order to keep another individual between themselves and a predator. It has been found that guppies in Trinidad streams are influenced by predation. Fish from streams with predators respond more quickly to a strange object by schooling and school more than fish from streams without predators. The behavior is at least partly genetic since it persists in laboratory descendants of the original guppies, implying that the schooling is a product of natural selection. Among birds, single pigeons are more likely to be caught by a hawk than are pigeons in a flock

even though the single birds spend more time on the alert. *See* POPULATION ECOLOGY.

Animals may also feed more efficiently in groups as several individuals searching for food increases the likelihood of finding some. Predators catch larger prey by cooperative hunting. Solitary leopards take prey about half their own size, while the highly social Cape hunting dog can bring down wildebeest and zebra 10 times larger than themselves. *See* ECOLOGICAL COMMUNITIES.

Territoriality. When resources, such as food and mates, are limited, members of a species will compete for access to them. One form of competition is when animals defend resource-rich areas, called territories, from conspecifics. Territorial behavior is very widespread. For example, female poplar aphids defend tiny territories at the base of poplar leaves. These are prime sites for establishing galls (in which offspring are produced), and there are many more females than available leaves. Similarly, male songbirds form spring territories in food-rich habitats. Females are attracted to these territories because abundant food supply enables them to raise offspring. Not all males manage to hold territories, and most females mate only with territorial males.

Territory holders gain exclusive access to valuable resources, but they also incur costs. Territorial defense requires time and energy, as exemplified by the calls and chases that male songbirds use to repel intruders. The conspicuous behavior of defenders may also increase predation risks. Animals should be expected to form territories only when the benefits of territoriality exceed the costs. Sandpipers, for example, are shorebirds which form feeding territories when their prey (small crustaceans) are at moderate densities but not when the prey are at very low or very high densities. At low densities, prey are too scarce to be worth defending, and at very high densities there are so many sandpipers trying to feed that territorial defense may be too costly. *See* TERRITORIALITY.

Reproduction. Males produce many small gametes (sperm), while females produce fewer, large gametes (eggs). As a result, males and females differ in their strategies of reproduction. Males can potentially produce many offspring by mating with several females, while females may benefit by ensuring that their relatively few offspring are of high quality. Males, then, generally compete for females. Selection for male mating ability, termed sexual selection, is a powerful evolutionary force.

Sexual selection underlies the elaborate courtship displays seen in many species. Male sage grouse, for example, have evolved bright plumage which they display in an effort to attract females. Sexual selection may favor females which choose high-quality males. Male hangingflies, for example, present gifts of prey (small insects) to prospective mates. It has been found that the prey vary in size and nutritional quality and that females preferentially copulate with males that bring the most valuable prey.

Animals have a variety of mating systems. In monogomous animals, males and females mate with one partner per breeding season, and parental care is often provided by both parents. Monogamous animals are frequently found in habitats with scattered resources. Many species are polygynous; that is, males mate with several females per breeding season; females mate with one male; and parental care is usually provided by females. Polygynous animals are frequently found in habitats with rich patches of resources. In polyandrous animals, females mate with several males per breeding season, while males mate with one female; parental care is usually provided by males. Polyandrous species, although generally rare, are most common in habitats with rich resources and high predation on offspring. In promiscuous animals, males and females mate with several partners per breeding season, and parental care is provided by males or females. Promiscuous animals are found both in habitats with scattered resources and highly synchronized female breeding and in habitats with unpredictable resources and high predation on offspring.

Most mammals are polygynous. A few males outcompete their rivals by defending groups of females or the resources which attract females to territories. The dominant males mate frequently, most males do not mate at all, and each female mates only once. Ecological constraints, however, often limit the ability of a few males to monopolize females. In most bird species, for example, females breed synchronously, and males have little opportunity to mate more than once. Young birds often require feeding by both parents, so the best male strategy is to stay with one female throughout the breeding season.

Parental care is common in birds and mammals, but examples can be found in almost all taxa. Parental care may include both provision of food and protection from predators. Mother lace bugs, for example, guard their nymphs against predators and greatly increase the chance that the nymphs will survive to adulthood. In general, females more frequently provide care than do males. Male parental care is most common when males are certain of their paternity. Thus, in bony fishes, species with male parental care tend to have external fertilization, and those with female parental care have internal fertilization.

In several species, nonbreeding animals help raise the offspring of their conspecifics. Breeding pairs of Florida scrub jays, for example, are aided by up to six helpers. Helpers are usually older siblings of the nestlings and are unable to establish their own breeding territories. They may benefit by gaining breeding experience and by ensuring the success of their relatives. The theory of kin selection, an extension of natural selection, predicts that animals should preferentially help kin, because close relatives share in common a large proportion of their genes. *See* BEHAVIOR GENETICS; ETHOLOGY; REPRODUCTIVE BEHAVIOR.

Hugh Dingle; Peter Frumhoff

Bibliography. J. Alcock, *Animal Behavior: An Evolutionary Approach*, 6th ed., 1997; J. R. Krebs and N. B. Davies, *An Introduction to Behavioral Ecology*, 1981; D. H. Morse, *Behavioral Mechanisms in Ecology*, 1980.

Behavioral psychophysics

The use of behavioral methods to measure the sensory capacities of animals and nonverbal humans such as infants. The observations are analogous to those obtained from adult human subjects who are asked to observe appropriate stimuli and report what they see. The behavioral methods differ primarily in that stimulus-response associations are established by means other than verbal instructions, either as unlearned reflexes or through conditioning.

Measures. Any sense or species may be studied, but most work has been done on the vision and hearing of primates, cats, pigeons, and rats. Typical investigations determine (1) the absolute threshold (the minimum intensity needed to elicit a standard response); (2) the difference threshold (the minimum change in a stimulus needed to elicit a standard response); and (3) points of apparent equality (values of stimuli that elicit no response because a change in one aspect compensates for a change in another). A few investigations have determined stimulus scales that express quantitative relations between the physical stimulus and the perceptual effect over a range of stimulus values. These various measures provide a picture of the sensory function, such as visual acuity, color sensitivity, loudness or pitch perception, and odor discrimination.

Methods. Efficiency and sensitivity to the sensory function of interest are the major factors that govern the choice of method in behavioral psychophysics.

Reflex methods. These methods are the most convenient since no training is required. An example is the use of the optokinetic reflex to estimate visual acuity. Here, the subject is exposed to a large field of moving stripes. Such a field elicits in most species a reflex turning of the head and eyes to follow the motion of the stripes. The stripes are made finer and finer until the turning response ceases, defining an acuity threshold. Another example is the preferential looking response in infants. Without training, infants spend more time looking at patterned stimuli than at blank fields. Two stimulus fields are placed before the infant, and the relative time spent looking at each is determined. A preference for one pattern according to this measure indicates that the infant can detect a difference between the two.

Unconditioned reflex methods are limited to sensory functions for which appropriate reflexes may be found; also, they usually impose severe limits on the specific stimulus conditions that may be used. Conditioning methods add considerable flexibility. In Pavlovian conditioning, the stimulus of interest becomes the conditioned stimulus through its association with a stimulus that elicits a clear-cut reflexive response (the unconditioned stimulus). For example, the sensitivity of an animal to air-pressure changes may be determined by causing pressure changes to precede the application of a mild electric shock. Since they signal shock, the pressure changes come to elicit heart-rate acceleration as the indicator response. The absolute threshold is that pressure change just strong enough to elicit a heart-rate acceleration.

Operant conditioning. This method offers still more flexibility. Typically the subject is rewarded for making an indicator response in the presence of one stimulus value; reward is withheld (or another response rewarded) in the presence of other stimulus values. The responses are selected to suit the species and the stimulus under study. For example, sucking and head turning have been employed in visual experiments with infants; bats have been taught to fly to sound sources; rats have pushed panels with their nose or jumped to platforms marked by the stimulus value in question.

A specific experiment indicates general aspects of behavioral psychophyics as well as the capabilities of the operant method. The purpose of this experiment was to determine the absolute sensitivity of the pigeon's eye continuously through time. The **illustration** diagrams the procedure. The pigeon was trained to work in the dark, positioning its head through a hole. Two response keys, A and B, sensed the pigeon's pecks. When the stimulus light was turned off, the bird was rewarded with food for pecking key B. Pecks on key A were rewarded indirectly: they periodically turned the stimulus off by closing a shutter, thus producing the condition under which pecking key B produced food. In this way, the pigeon learned to peck A only when the light was visible and B only when the light was invisible.

In addition to producing reward, pecks to A and B were made to control the intensity of the stimulus light through an automatic circuit. Between scheduled shutter closures, pecks on key A reduced stimulus intensity by moving a variable density optical wedge. Thus, when the bird could see the light, it pecked key A, reducing the light's intensity until this fell below the bird's absolute threshold. This disappearance of the light could not be discriminated from total darkness (shutter closure), so the bird

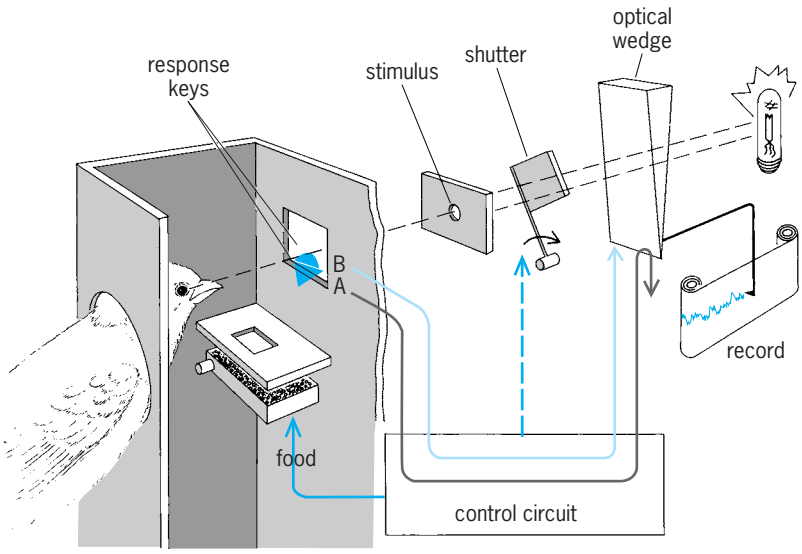

Diagram of the apparatus for measuring the pigeon's absolute visual threshold. (*After D S. Blough, A method for obtaining psychophysical thresholds from the pigeon, J. Exp. Anal. Behav., 1:31–43, 1958*)

now pecked key B, raising the stimulus to visibility once again. The pigeon's pecks thus caused the stimulus intensity to oscillate up and down across the bird's absolute threshold. Periodic shutter closures and feeding maintained the appropriate response pattern. This means of tracking the threshold through time enables one to study the changes in visual sensitivity, such as those that are induced by exposure to bright light or to certain drugs.

Significance. Full understanding of a sensory system can be achieved only when the relevant structures (such as the eye) and physiological mechanisms (such as neural activity in the visual system) are related to the overall functioning of the system as expressed in the behavior of the whole animal (such as visually guided behavior). In most instances, structure and physiology may best be studied in a nonhuman subject. Behavioral psychophysics in this species then provides coordinate information on the operation of the system as a whole. By far the largest number of studies using the methods described here have been motivated by this desire to understand a sensory system at several levels in a single species.

Interest in the development of sensory function has spurred the use of behavioral methods in infants and young children, while the effects of controlled sensory input during development have been extensively monitored in nonhuman subjects. Applications to the prevention and control of sensory disorders also are increasing. For example, a number of toxicants, drugs, and environmental stressors have been related to sensory disorders; the methods of behavioral psychophysics are used to follow the development of these disorders under controlled conditions, and to uncover potential preventive and therapeutic measures. *See* PSYCHOLOGY; SENSATION.

Donald Blough

Bibliography. H. Geissler et al., *Cognition, Information Processing and Psychophysics: Basic Issues*, 1989.

Behavioral toxicology

The study of behavioral abnormalities induced by exogenous agents such as drugs, chemicals in the general environment, and chemicals encountered in the workplace. Just as some substances are hazardous to the skin or liver, some are hazardous to the function of the nervous system. In the case of permanent effects, changes in sensation, mood, intellectual function, or motor coordination would obviously be undesirable, but even transient alterations of behavior are considered toxic in some situations. For example, operating room personnel accidentally exposed to very small doses of anesthetic do not exhibit altered performance on an intelligence test, a dexterity test, or a vigilance task. However, significant decrements in performance occur in recognizing and recording visual displays, detecting changes in audiovisual displays, and recalling series of digits.

By comparing the behavior of exposed subjects and control subjects, behavioral toxicologists seek to identify agents capable of altering behavior and to determine the level of exposure at which undesirable effects occur. When the agent under study is one in common use, and there is no evidence of its being hazardous to health, experiments may be carried out on human volunteers, or comparisons may be made from epidemiologic data. More frequently, safety considerations dictate the use of laboratory animals in toxicology research. Primates are favored for their wide behavioral repertoire, trainability, and long lifespan, but they are extremely expensive to purchase and maintain. Rodents can be produced and maintained more economically, and their short gestation period and rapid development are advantageous in some types of research. Scientists often investigate toxic effects in several species to take advantage of the useful characteristics of each and to increase the generality of their conclusions.

Humans. Perhaps the best-known example of toxicity in humans is methyl mercury poisoning (Minimata disease), which occurred in epidemic proportions in a Japanese coastal town where the inhabitants ate fish contaminated with mercury from industrial pollution. Although mercury affects a variety of behaviors, the most obvious symptoms are tremors and involuntary movements.

A different set of functional problems is exemplified by the effects of ethyl alcohol, a single agent with direct and indirect, short- and long-term consequences. The short-term, low-dose effects of alcohol include sensory disturbances, motor problems, and difficulties with processing information. Neurologically, alcohol is usually described as a central nervous system depressant which is general, in the sense that it disrupts many functions.

In some individuals, large quantities of alcohol consumed over a long period lead to permanent damage to the nervous system. Behaviorally, individuals with Korsakoff's syndrome exhibit severe memory deficits. Anatomically, their brains are found to have degenerative changes in the thalamus. This syndrome is not just an extension of the short-term effects. In fact, it is thought to arise from alcohol-induced malnutrition rather than as a direct effect of the alcohol itself.

Lasting injuries to the nervous system have been reported to occur in children exposed to alcohol before birth. The behavioral problems associated with fetal alcohol syndrome do not appear to be related to either Korsakoff's syndrome or the immediate effects of alcohol. Rather, they constitute a third set of effects, including learning deficits, problems with inhibiting inappropriate behavior, and fine motor dysfunction, along with some visible physical abnormalities. While malnutrition may play a role in this congenital syndrome, the mechanism and locus of damage are not known. *See* ALCOHOLISM; CONGENITAL ANOMALIES; FETAL ALCOHOL SYNDROME.

Mature animals. In the controlled environment of the laboratory, researchers can deliver carefully measured doses of test agents and manage the subjects' physical condition and experiential history. In addition, many aspects of behavior can be quantified. For

example, whereas reports of family members can be used to judge the activity level of a child at home, a laboratory rat can be housed in a cage designed to record the animal's movements, giving the experimenter an objective measure of activity. However, the specificity of laboratory studies creates a major problem for the behavioral toxicologist, for it is not possible to test all behaviors of an animal. Therefore, the researcher must decide which behaviors are to be evaluated. The problem of test selection is serious, because alterations of function can be very selective, leaving many behaviors normal while a few are grossly abnormal. A common strategy is to sample a variety of functions that appear to be unrelated, in the hope of covering as many functional systems as possible. For example, a series that included tests of vision, hearing, coordination, and strength would be more likely to detect toxic effects than a series of four visual tasks.

Another approach to the problem of test selection is to include some tests that require the integrity of many systems. For example, scientists at the National Institute of Occupational Safety and Health employ an elaborate operant task for primates. The animal sits in a chamber from which it can see a visual display panel. Lights signal that a trial is about to begin. To receive a food reward, the animal must grasp a spring-mounted lever, move it to a specified position, and hold it there for several seconds. A series of lights signals when the lever is in the proper position and tells the subject how long it must wait for the reward. While this task was designed as a very sensitive measure of motor control and tremor, abnormalities of attention, time perception, memory, or vision could also interfere with performance. A task like this requires a long period of training, but the trained animal can be used over and over, as long as it sustains no permanent injury from exposures to suspected toxic agents. When exposure is likely to cause permanent alterations of behavior, it is more economical to employ tests which require less investment of time and less expensive animals.

Developing animals. The study of congenital behavioral deficits is known as behavioral teratology. In these experiments, animals are exposed before birth or soon after birth and are later tested for behavioral abnormalities. As in the example of fetal alcohol syndrome, it is often impossible to predict the effects of an agent on the developing brain from its effects on the adult brain. Embryos and neonates survive some insults, such as anoxia, better than adults. Conversely, some agents that are well tolerated by mature animals are disastrous to developing ones. For example, agents that interfere with cell proliferation are not hazardous to the mature brain because neurons do not multiply in adults. But in the developing brain, interference with proliferation, as by irradiation with x-rays, can permanently alter the number of neurons and consequently the adult behavior of exposed animals. Regulations regarding exposure to radiation reflect this difference, limiting the exposure of pregnant women. Some agents which have reversible effects in adults may have permanent effects when administered to young animals. For example, several psychoactive drugs are believed to alter the developing nervous system of a young animal and lead to lasting behavioral changes.

Because an animal can be treated only once in studies of developmental toxicity, the typical experimental paradigm in this field employs many animals, often rodents. A behavioral test battery might begin with tests of reflex development in neonates. Other popular tests include open-field measures of emotionality and general activity, maze-learning problems, avoidance tasks, and sensory measures.

Significance of behavioral effects. When toxicity is considered only in terms of direct risk to survival, behavioral toxicity may seem to be of minor importance. However, survival is not the only criterion of good health. In a complex society that places heavy demands on an individual's educability, alertness, and emotional stability, even small deviations in behavior are potentially hazardous. Severe disabilities, as in the gross motor malfunctions of Minimata disease, have drawn attention to behavioral toxicology. Such incidents represent failures of control of toxic substances. Successes are difficult to measure, for they can be seen only in reduction of risk—the ultimate goal of toxicology. *See* TOXICOLOGY.

Patricia M. Rodier

Bibliography. Z. Annau (ed.), *Neurobehavioral Toxicology*, 1987; O. Hutzinger (ed.), *Reactions and Processes*, 1984; W. R. Notton et al. (eds.), *Health Surveillance of Individual Workers Exposed to Chemical Agents*, 1988; G. Zbinden et al. (eds.), *Application of Behavioral Pharmacology in Toxicology*, 1983.

Belemnoidea

An order of extinct dibranchiate cephalopods. These mollusks ranged from the Upper Mississippian to the upper Eocene. The oldest known representative is highly specialized, although the belemnoids are considered to be the most primitive of the dibranchiates. Belemnoids have a chambered

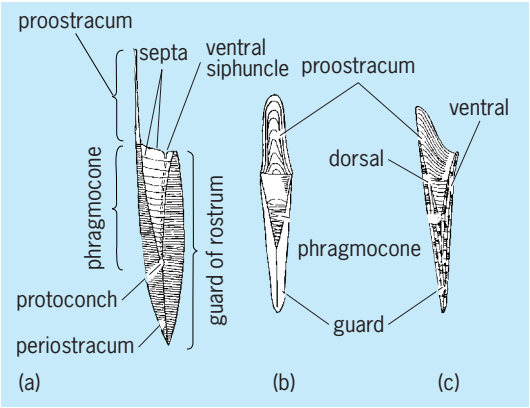

Belemnoidea. (*a*) Diagram of a belemnitid shell. (*b*) Ventral aspect and (*c*) lateral aspect of a belemnite. (*After R. R. Shrock and W. H. Twenhofel, Principles of Invertebrate Paleontology, 2d ed., McGraw-Hill, 1953*)

shell or phragmocone which fits into a guard or rostrum, the apical portion of the shell (see **illus.**). *Belemnites* is an important index fossil. *See* CEPHALOPODA; INDEX FOSSIL. Charles B. Curtin

Bell

A hollow metallic cylinder closed at one end and flared at the other. A bell is used as a fixed-pitch musical instrument or a signaling device, frequently in clocks. A bell is set vibrating by a clapper or tongue which strikes the lip. The bell has two tones, the strike note or key, and the hum tone, which is a major sixth below the strike note. Both tones are rich in harmonics.

The finer bells are made for carillons. Small carillons have two full chromatic octaves and large ones have four or more chromatic octaves, the bell with the lowest tone being called the bourdon. The Riverside Church in New York has 72 bells, the largest or bourdon weighing 18.25 tons (16.6 metric tons). All bells of a carillon are shaped for homogeneity of timbre.

Bell metal is approximately four parts copper and one part tin, although zinc, lead, or silver may be used. The shape of a bell—its curves and full trumpet mouth—evolved from experience. The bow or edge has a thickness of one-twelfth to one-fifteenth its diameter. The contour of a new bell is formed in the molds by rotating strickle boards, one for the core or inner mold and one for outer surface (**Fig. 1**). A dummy or false bell is made of brittle clay over the core; any required inscription is placed in relief with wax on the dummy. The cope or outer mold is then built up on the dummy, first with a thin soup clay mixture to fit the inscription, then with thicker mixtures.

Fig. 1. Internal core of bell, which is shaped by revolving a strickle board. (*Les Fils de G. Paccard*)

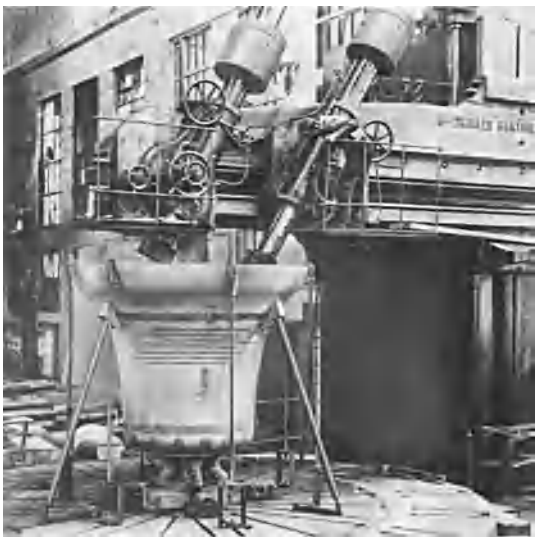

Fig. 2. A 10-ton (9-metric-ton) bell being milled after the molding process. (*Les Fils de G. Paccard*)

Molds and dummy are dried and baked to melt out the wax and harden the core and cope. The dummy is removed, the mold is locked in place, and the melted bronze is poured in and allowed to cool. An 18-ton (16-metric-ton) bourdon may take a quarter hour to pour and 36 to cool. The mold is removed and the bell trimmed, sandblasted, and possibly polished.

A bell can be pitched slightly. The tone is raised by grinding the outer surface, effectively decreasing the diameter, or lowered by grinding the inner surface (**Fig. 2**). Too much tuning of a well-shaped bell is apt to degrade its voice or timbre. Frank H. Rockett

Belladonna

The drug and also the plant known as the deadly nightshade, *Atropa belladonna*, which belongs to the nightshade family (Solanaceae). This is a coarse, perennial herb native to the Mediterranean regions of Europe and Asia Minor, but now grown extensively in the United States, Europe, and India. During the blooming period, the leaves, flowering tops, and roots are collected and dried for use. The plant contains several important medicinal alkaloids, the chief one being atropine, which is much used to dilate the pupils of the eye. *See* ATROPINE; SOLANALES.

Perry D. Strausbaugh; Earl L. Core

Beloniformes

An order of actinopterygian fishes that comprise, with the Atheriniformes and Cyprinodontiformes, the series Artherinomorpha. These fishes are identified by the interarcural cartilage that connects the first and second gill arches; small second and third epibranchials; absence of interhyal; a nonprotrusible upper jaw; far posterior placement of the dorsal

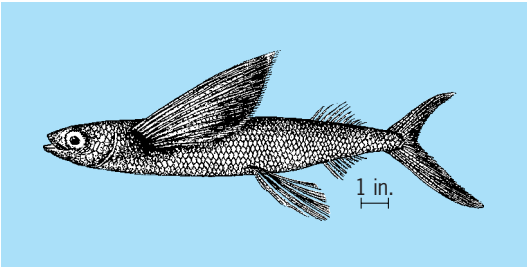

California flyingfish (*Cypselurus californicus*). 1 in. = 2.5 cm. (*After G. B. Goode, Fishery Industries of the U.S., 1884*)

and anal fins; abdominal pelvic fins; and the lower caudal fin lobed with more principal rays than the upper lobe and significantly longer in needlefishes and flyingfishes. The order consists of 191 species in 38 genera and 5 families. Each family has distinctive features that make identification simple.

Adrianichthyidae. Adrianichthyids (ricefishes) are small fishes that inhabit fresh and brackish waters from India to Japan and to the Indo-Australian archipelago. Adults vary from 3 to 20 cm (1.2 to 8 in.) in length. They are best described by skeletal structures that they have lost. Absent are certain head bones, such as vomer, meta- and ectopterygoids, rostral cartilage, supracleithrum, and lateral lines. In one species, the smallest, the maxillae are absent as well as the right pelvic fin of females, and the mouth is very small, whereas in several of the larger species the jaws are greatly enlarged and the mouth functions like a scoop shovel.

Exocoetidae. This family, the flyingfishes, is familiar to seafarers in the tropical and subtropical Atlantic, Indian, and Pacific oceans. It is readily recognized by having unusually large pectoral fins that can be used for gliding above the ocean surface. Some species gain additional gliding ability from greatly enlarged pelvic fins. A flyingfish does not flap its "wings"; instead it glides on them after being launched into the air by powerful thrusts of the tail. Sustained "flight" is accomplished by repeated contacts of its elongate lower caudal fin lobe with the water. Worldwide there are 8 genera and about 52 species, 26 of which occur in blue waters of North America, including all of Mexico.

Hemiramphidae. Known as halfbeaks, these fishes are readily distinguished from other beloniforms by having the upper jaw much shorter than the lower; premaxillae pointed anteriorly, forming a triangular upper jaw; short pectoral and pelvic fins; and the tip of the long lower jaw bright red or orange in most species. Of the 13 genera and more than 100 species, most occur in the warm waters of the Atlantic, Indian, and Pacific oceans, but several genera, some of which are viviparous, live in freshwaters of Southeast Asia. Four genera and 12 species are known from the marine waters of Mexico and the United States, and a single freshwater species occurs in Mexico.

Belonidae. The family Belonidae (needlefishes) differs from the preceding three families in having both upper and lower jaw elongated (very rarely the upper jaw is short) and equipped with numerous sharp teeth. It differs from the Scomberesocidae in lacking dorsal and anal finlets. At hatching, both jaws are short and of equal length (as in ricefishes and adult flyingfishes). As development progresses, the lower jaw elongates first (halfbeak stage, as in adult halfbeaks), followed by elongation of the upper jaw (needlenose stage, typical of adult belonids). Some of the larger belonids, such as the hound fish, *Ablennes hians*, can skitter over the water surface at great speeds, potentially posing a danger to waders, especially at night when the fish are attracted by lights or when they find themselves trapped by a net. There are 10 genera represented by 21 epipelagic species in tropical and temperate seas, and 11 species restricted to freshwater. Ten species in 4 genera occur in North American waters and only one, the Maya needlefish of Mexico (*Strongylura hubbsi*), is restricted to freshwater. Another, the Atlantic needlefish (*S. marina*), ascends rivers of the Atlantic and Gulf coasts to the Fall Line, including rivers of Mobile Basin, where it reproduces.

Scomberesocidae. This family, called sauries, consists of only four monotypic genera. They are closely related to the needlefishes, but finlets following the dorsal and anal fins easily distinguish them from other beloniforms. The jaws are delicate and vary from long and slender to relatively short. Sauries are epipelagic in tropical and temperate seas. *See* ACTINOPTERYGII; ATHERINIFORMES; CYPRINODONTIFORMES; TELEOSTEI.

Herbert Boschung

Bibliography. B. B. Collette, *Family Belonidae Bonaparte 1832—Needlefishes*, Calif. Acad. Sci. Annotated Checklists of Fishes No. 16, 2003; B. B. Collette et al., Beloniformes: Development and relationships, pp. 334–354, in H. G. Moser et al., (eds.), *Ontogeny and Systematics of Fishes*, Amer. Soc. Ich. Herp. Spec. Publ. No. 1, 1984; L. R. Parenti, *Relationships of Atherinomorph Fishes (Teleostei)*, *Bull. Mar. Sci.*, 52(1):170–196, 1993; D. E. Rosen and L. R. Parenti, *Relationships of Oryzias, and the Groups of Atherinomorph Fishes*, Amer. Mus. Novit. 2719, 1981.

Belt drive

The lowest-cost means for transmitting power between shafts that are not necessarily parallel. Belts run smoothly and quietly, and they cushion motor and bearings against load fluctuations. Belts typically are not as strong or durable as gears or chains. However, improvements in belt construction and materials are making it possible to use belts where formerly only chains or gears would do.

Advantages of belt drive are: They are simple. They are economical. Parallel shafts are not required. Overload and jam protection are provided. Noise and vibration are damped out. Machinery life is prolonged because load fluctuations are cushioned (shock-absorbed). They are lubrication free. They require only low maintenance. They are highly efficient (90–98%, usually 95%). Some misalignment is

tolerable. They are very economical when shafts are separated by large distances. Clutch action may be obtained by relieving belt tension. Variable speeds may be economically obtained by step or tapered pulleys.

Disadvantages include: The angular-velocity ratio is not necessarily constant or equal to the ratio of pulley diameters, because of belt slip and stretch. Heat buildup occurs. Speed is limited to usually 7000 ft/min (35 m/s). Power transmission is limited to 500 hp (370 kW). Operating temperatures are usually restricted to −31 to 185°F (−35 to 85°C). Adjustment of center distance or use of an idler pulley is necessary for wear and stretch compensation. A means of disassembly is provided to install endless belts.

Belt types and general use. There are four general types of belts, each with its own special characteristics, limitations, advantages, and special-purpose variations for different applications.

Flat belts. Flat belts, in the form of leather belting, served as the basic belt drive from the beginning of the Industrial Revolution. They can transmit large amounts of power at high speeds—up to 500 hp (370 kW) and 10,000 ft/min (50 m/s)—if wide belts are used over large pulleys. Such drives are cumbersome; they require high belt tensions, which in turn produce large bearing loads. For these reasons V-belts have generally replaced flat belts in power delivery applications. Flat belts find their widest application where high-speed motion, rather than power, is the main concern. Flat belts are very useful where large center distances and small pulleys are involved. They can engage pulleys on both inside and outside surfaces, and both endless and jointed construction is available.

V-belt. V-belts are the basic power-transmission belt. They provide the best combination of traction, operating speed, bearing load, and service life. The belts are typically endless, with a trapezoidal cross section which runs in a pulley with a V-shaped groove (**Fig. 1**). The wedging action of the belt in the pulley groove allows V-belts to transmit higher torque at less width and tension than flat belts. V-belts are far superior to flat belts at small center distances and high reduction ratios. Center distances greater than the largest pulley diameter and less than three times the sum of the two pulley diameters are preferred. Optimum speed range is 1000–7000 ft/min (5–35 m/s). V-belts require larger pulleys than flat belts because of their greater thickness. Several individual belts running on the same pulley in separate grooves are often used when the power to be transmitted exceeds that of a single belt. These are called multiple-belt drives.

Fig. 1. V-belt in cross section.

Fig. 2. Timing belt and pulleys. (*Morse Chain Division of Borg-Warner Corp.*)

Jointed and link V-belts are available when the use of endless belts is impractical. Jointed belts are not as strong and durable as endless and are limited to speeds of 4000 ft/min (20 m/s). A link V-belt is composed of a large number of rubberized-fabric links jointed by metal fasteners. The belt may be disassembled at any point and adjusted to length by removing links.

Film belts. Film belts are often classified as a variety of flat belt, but actually they are a separate type. Film belts consist of a very thin (0.5–15 mils or 100–4000 micrometers) strip of material, usually plastic but sometimes rubber. They are well suited to low-power (up to 10 hp or 7 kW), high-speed application where they provide high efficiency (as high as 98%) and long life. One of their greatest advantages is very low heat buildup. Business machines, tape recorders, and other light-duty service provide the widest application of film belts.

Timing belts. Timing belts have evenly spaced teeth on their bottom side which mesh with grooves cut on the periphery of the pulleys to produce a positive, no-slip, constant-speed drive, similar to chain drives (**Fig. 2**). They are often used to replace chains or gears, reducing noise and avoiding the lubrication bath or oiling system requirement. They have also found widespread application both in miniature timing applications and in automobiles. Timing belts, known also as synchronous or cogged belts, require

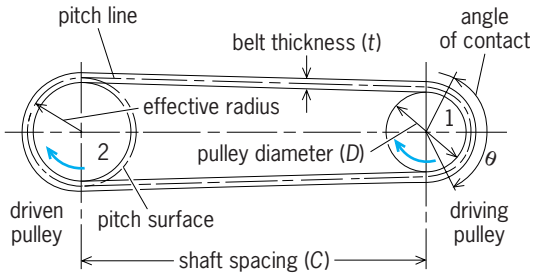

Fig. 3. Basic elements of open belt drive.

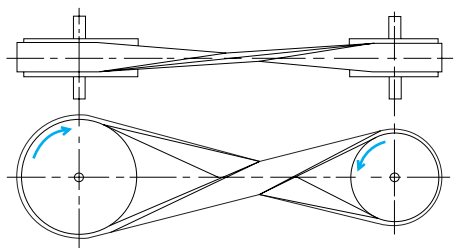

Fig. 4. Two views of crossed belt drive.

the least tension of all belt drives and are among the most efficient. They can transmit up to 200 hp (150 kW) at speeds up to 16,000 ft/min (80 m/s). There is no lower limit on speed. Disadvantages are high first cost, grooving the pulleys, less overload and jam protection, no clutch action, and backlash.

Belt-pulley arrangements. The most common arrangement, by far, is the open belt drive (**Fig. 3**). Here both shafts are parallel and rotate in the same direction. The cross-belt drive of **Fig. 4** shows parallel shafts rotating in opposite directions. Timing and standard V-belts are not suitable for cross-belt drives because the pulleys contact both the inside and outside belt surfaces. Nonparallel shafts can be connected if the belt's center line, while approaching the pulley, is in the pulley's central plane (**Fig. 5**). *See* ROLLING CONTACT.

Belt materials. Industrial belts are usually reinforced rubber or leather, the rubber type being predominant. Nonreinforced types, other than leather, are limited to light-duty applications. General belt material properties are given in the **table**.

Speed relationship between pulleys. Accurate calculation of the speed relationship between pulleys is based on pitch diameter, the effective pulley diameter. As a belt bends around a pulley, the outer fibers are in tension, the inner in compression. In between is the pitch line, which is subject to neither stress nor change in length (Fig. 3). For film and flat belts, the pitch line is midway through the belt surface. For timing and V-belts, the pitch line depends on cross-section shape and size. The factors to convert pulley diameters to pitch diameters are typically listed in engineering handbooks and manufacturer's literature. The speed relationship between pulleys is given by Eq. (1), where N_1 is the angular speed of pulley 1,

$$\frac{N_1}{N_2} = \frac{PD_2}{PD_1} \tag{1}$$

N_2 is the angular speed of pulley 2, PD_1 is the pitch diameter of pulley 1, and PD_2 is the pitch diameter of pulley 2.

The actual pulley speeds are often 0.5–1% less than predicted by Eq. (1), because of belt slip and stretch. The exact speed relationship for timing belts is also the inverse ratio of the number of teeth on the pulleys.

The speed in m/s of the belt itself is given by Eq. (2), where PD is the pitch diameter of the given pulley in rpm.

$$V = \pi(PD)(N) \tag{2}$$

Selection criteria. To design a belt drive, the following information is required: speeds of drive and driven unit; power transferred between drive and driven unit: power (kW) = torque (newton-meters) × rpm × 1.05×10^{-4}; desired distance between shafts; and operating conditions.

The drive horsepower is corrected for various conditions such as speed ratio; distance between shafts (long or short); type of drive unit (electric motor, internal combustion engine); service environment (oily, wet, dusty); driven unit loads (jerky, shock, reversed); and pulley-belt arrangement (open, crossed, turned). The correction factors are listed in engineering handbooks and manufacturer's literature. The corrected horsepower is compared to rated horsepowers of standard belt cross sections at specific belt speeds to find a number of combinations that will do

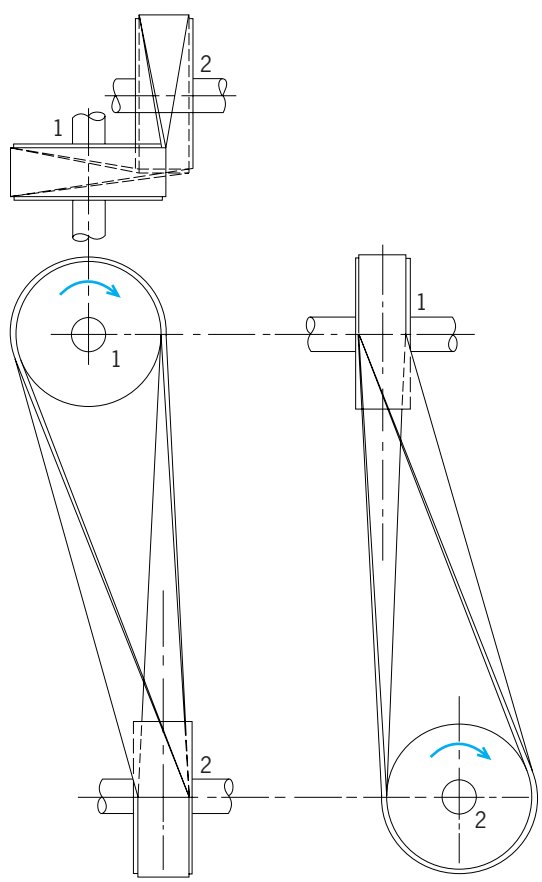

Fig. 5. Quarter-turn belt.

Belt material properties

Material	Properties	Uses
Nonreinforced rubber	Low-power and low-speed applications; very useful for fixed-center drives	Flat and film belts, vulcanized joints
Nonreinforced plastics	Low-power and low-speed application	Flat, film, and miniature timing belts, molded endless
Nonreinforced leather	High power; long service life; high cost; must be cleaned and dressed; prone to stretching and shrinking; limited to low and moderate speeds	Flat belts only; often laminated, with glued, riveted, or wire-laced joints
Fabric	Ability to track uniformly; low friction is a disadvantage	Endless flat belts constructed from cotton or synthetic fabric, with or without chemical or rubber coatings
Steel	Very high-speed, low-power applications; high tension required	Thin steel bands with riveted or silver-soldered joints
Reinforced rubber, leather, and plastics	Nylon, polyamide, glass fiber, or steel wire tensile members covered with rubber, leather, or plastic; speed and power capabilities are greatly improved by the reinforcement	All but film belts
Rubberized fabric	Consists of plies of cotton or synthetic fabric impregnated with rubber; low cost; good abrasion resistance, but has a lower power-transmitting ability and service life than leather; limited to speeds below 6000 ft/min (30 m/s)	All but film belts
Rubberized cord	Constructed with rubber-impregnated cords running lengthwise; carries 50% more power than rubberized fabric belts; fair shock-absorption capacity	Timing, flat, and V-belts, endless only
Rubberized cord and fabric	Combines strength of cord with abrasion resistance of fabric; well suited to minimum-stretch, high-speed operations	All but film belts, endless only

the job. At this point, the pulley diameters are chosen. A trade-off must be made between larger pulley diameters or larger belt cross sections, since larger belts transmit this same power at low belt speeds as smaller belts do at high speeds. To keep the drive unit as small as possible, minimum-diameter pulleys are desired. Minimum pulley diameters are limited by the elongation of the belt's outer fibers as the belt bends around the pulleys. Small pulleys increase this elongation, greatly reducing belt life. Minimum pulley diameters are often listed with each cross section and speed, or listed separately by belt cross section. Once the most economical pulley diameters and belt section are chosen, the belt length is calculated. If endless belts are used, the desired shaft spacing may need adjusting to accommodate standard length belts. It is often more economical to use two or more V-belts side by side, rather than one larger belt.

If a large speed ratio or a small distance between shafts is used, the angle of contact between the belt and pulley may be less than 180°. If this is the case, the drive power must be further increased, according to manufacturer's tables, and the selection process repeated. The drive power must be increased because belt power capacities are specified assuming 180° contact angles. Smaller contact angles mean less area for the belt to obtain traction, and thus the belt carries less power.

Belt specifications. To fully specify a belt, the material, length, and cross-section size and shape are required. Timing belts, in addition, require that the size of the teeth be given.

For all belt types, and for open and crossed belt drives, an accurate approximation of belt length is Eq. (3). (Symbol definitions are given in Fig. 3.) The

$$L = 2C + 1.57(D_2 + D_1) + \frac{(D_2 \pm D_1)^2}{4C} \qquad (3)$$

sign of the second D_1 is plus for crossed belt drives and minus for open belt drives.

The angle of contact in radians between belt and pulley (Fig. 3) is given by Eq. (4); the plus sign is

$$\theta = \pi \pm 2 \arcsin\left(\frac{D_2 - D_1}{2C}\right) \qquad (4)$$

for the larger pulley, the minus sign for the smaller pulley.

For the crossed belt drive, both angles of contact are calculated by Eq. (5).

$$\theta = \pi + 2 \arcsin\left(\frac{D_2 + D_1}{2C}\right) \qquad (5)$$

When correcting for angle of contact, the smallest angle calculated should always be used.

Belt tension. Belts need tension to transmit power. The tighter the belt, the more power transmitted. However, the loads (stress) imposed on the belt and supporting bearings and shafts are increased. The best belt tension is the lowest under which the belt will not slip with the highest load condition. Belt manufacturers recommend operating belt tensions according to belt type, size, speed, and pulley diameters. Belt tension is determined by measuring the

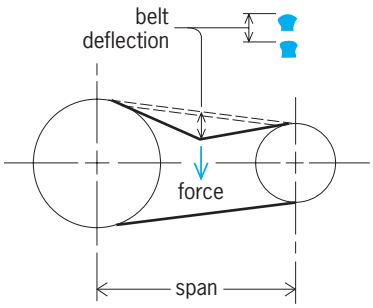

Fig. 6. Tension measurement by deflection.

force required to deflect the belt a specific distance per inch of span length (**Fig. 6**). Timing belts require minimum tension, just enough to keep the belt in contact with the pulleys.

Belt wear. Belts probably fail by fatigue more often than by abrasion. The fatigue is caused by the cyclic stress applied to the belt as it bends around the pulleys. Belt failure is accelerated when the following conditions are present: high belt tension; excessive slippage; adverse environmental conditions; and momentary overloads caused by shock, vibration, or belt slapping. *See* CHAIN DRIVE; PULLEY; WEAR.

Arthur Erdman; Raymond Giese

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; W. S. Miller (ed.), Drive systems, *Mach. Des.*, 50 (15):26–33, June 29, 1978; C. R. Mischke and J. E. Shigley, *Mechanical Engineering Design*, 6th ed., 2000; P. B. Schulbert (ed.), *Machinery's Handbook*, 21st ed., 1979.

Bentonite

The term first applied to a particular, highly colloidal plastic clay found near Fort Benton in the Cretaceous beds of Wyoming. This clay swells to several times its original volume when placed in water and forms thixotropic gels when small amounts are added to water. Later investigations showed that this clay was formed by the alteration of volcanic ash in place; thus, the term bentonite was redefined by geologists to limit it to highly colloidal and plastic clay materials composed largely of montmorillonite clay minerals, and produced by the alteration of volcanic ash in place. Many mineralogists and geologists now use the term bentonite without reference to the physical properties of the clay. On the other hand, the term has been used commercially for any plastic, colloidal, and swelling clays without reference to a particular mode of origin. *See* CLAY; GEL; MONTMORILLONITE.

Origin. The occurrence of shard structures of ash as pseudomorphs in the clay often indicates the volcanic ash parent material of bentonites. The presence of nonclay minerals characteristic of igneous material also provides evidence for the origin from ash.

Evidence strongly indicates that the transformation of volcanic glass to montmorillonite takes place either during or soon after accumulation of the ash. The alteration process is essentially a devitrification of the natural glass of the ash and the crystallization of montmorillonite. The ash often contains excess silica which may remain in the bentonite as cristobalite. *See* VOLCANIC GLASS.

Occurrence. Bentonites have been found in almost all countries and in rocks of a wide variety of ages. They appear to be most abundant in rocks of Cretaceous age and younger. In such material in older rocks the montmorillonite is often almost completely collapsed by compaction and metamorphism. This altered bentonite swells very little and does not have the usual high colloidal properties of bentonite; it is also called metabentonite.

Beds of bentonite show a variable thickness and can range from a fraction of an inch up to as much as 50 ft (15 m) thick. The color of bentonite is also variable, but perhaps most often it is yellow or yellowish-green. When observed in the weathered outcrop, bentonite tends to develop a characteristic cellular structure. It shrinks upon drying and often shows a characteristic jigsaw-puzzle set of fractures. Not all beds of ash in the geologic column have altered to bentonite. When present, bentonite is most frequently found in marine beds; thus, it seems certain that the alteration from ash to bentonite is favored in sea water.

In the United States, bentonites are mined extensively in Wyoming, Arizona, and Mississippi. England, Germany, former Yugoslavia, Russia, Algeria, Japan, and Argentina also produce large tonnages of bentonite.

Uses. Bentonites are of great commercial value. They are used in decolorizing oils, in bonding molding sands, in the manufacture of catalysts, in the preparation of oil well drilling muds, and in numerous other relatively minor ways. Since the properties of bentonites vary widely, they are not all suitable for commercial use. The properties of a particular bentonite determine its economic use. For example, the bentonite found in Wyoming is excellent for drilling muds and for foundry use, but it is not useful for the making of catalysts or for oil decolorizing. *See* CLAY, COMMERCIAL.

Ralph E. Grim; Floyd M. Wahl

Benzene

A colorless, liquid, inflammable, aromatic hydrocarbon of chemical formula C_6H_6 which boils at 80.1°C (176°F) and freezes at 5.4–5.5°C (41.7–41.9°F). In the older American and British technical literature benzene is designated by the German name benzol. In current usage the term benzol is commonly reserved for the less pure grades of benzene.

Benzene is used as a solvent and particularly in Europe as a constituent of motor fuel. In the United States the largest uses of benzene are for the manufacture of styrene and phenol. Other important outlets are in the production of dodecylbenzene, aniline, maleic anhydride, chlorinated benzenes (used in making DDT and as moth flakes), and benzene hexachloride, an insecticide.

X-ray and electron diffraction studies indicate that the six carbon atoms of benzene, each with a hydrogen atom attached, are arranged symmetrically in a plane, forming a regular hexagon. The hexagon symbol (**1**), commonly used to represent the structural formula for benzene, implies the presence of a carbon atom at each of the six angles and, unless substituents are attached, a hydrogen at each carbon atom. Whereas the three double bonds usually included in formula (**1**) are convenient in accounting for the addition reactions of benzene, present evidence is that all the carbon-to-carbon bonds are identical.

(**1**)

Prior to World War II most of the benzene commercially produced was a by-product of the steel industry. High-temperature pyrolysis of coal to produce metallurgical coke for the steel industry yields gases and tars from which benzene may be recovered.

In the present era nearly all commercial benzene is a product of petroleum technology. The gasoline fractions obtained by reforming or steam cracking of feedstocks from petroleum contain benzene and toluene, which can be separated economically. Unfortunately the presence of benzene and toluene are important for the high octane number of these same gasoline fractions, with the result that the prices of benzene and toluene are affected by that of lead-free gasoline. Benzene may also be produced by the dealkylation of toluene. *See* PETROLEUM PROCESSING AND REFINING.

Despite its low hydrogen-to-carbon ratio, benzene reacts chiefly by substitution. Because of the symmetry of the molecule, a given reagent will react to afford only one monosubstituted benzene. Some typical substitution products are included in the **table**. *See* SUBSTITUTION REACTION.

A large number of disubstitution products of benzene are known. For any two substituents there are only three possible arrangements. The groups may be adjacent to each other (ortho, *o*-). They may be located on carbons separated from each other by a single CH, that is, at positions 1 and 3 of formula (**2**)

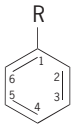

(**2**)

(meta, *m*-), or they may be at any two opposite corners of the hexagon (para, *p*-). The three possible dichlorobenzenes are shown below.

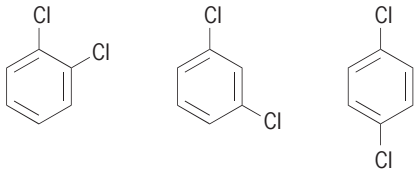

o-Dichlorobenzene *m*-Dichlorobenzene *p*-Dichlorobenzene

The position of substituents in trisubstituted or polysubstituted benzenes is usually indicated by numbers. For instance, the trichlorobenzene in which all three chlorines are adjacent is called 1,2,3-trichlorobenzene.

Benzene may also react by addition. Hydrogenation in the presence of a catalyst yields cyclohexane, C_6H_{12}, an intermediate in the manufacture of fibers, while addition of chlorine (usually in the presence of ultraviolet light) yields benzene hexachloride, $C_6H_6Cl_6$, an insecticide.

The largest commercial uses for benzene are in the manufacture of ethylbenzene (a styrene intermediate), cumene (an intermediate in the manufacture of phenol), and cyclohexane (an intermediate in the manufacture of fibers).

Benzene is a toxic substance, and prolonged exposure to concentrations in excess of 35–100 parts per million in air may lead to symptoms ranging from nausea and excess fatigue to anemia and leukopenia. In addition, the Occupational Safety and Health Administration has listed benzene as a potential carcinogen. *See* AROMATIC HYDROCARBON. Charles K. Bradsher

Bibliography. T. W. G. Solomons, *Organic Chemistry*, 7th ed., 1999; K. Weissermel and H. J. Arpe, *Industrial Organic Chemistry*, 3d ed., 1999.

Substitution products of benzene

Reagents	Formula (**2**), R =	Name of product
H_2SO_4	—SO_3H	Benzenesulfonic acid
$HNO_3 + H_2SO_4$	—NO_2	Nitrobenzene
$Cl_2 + Fe$	—Cl	Chlorobenzene
$C_2H_4 + AlCl_3$	—C_2H_5	Ethylbenzene

Benzoic acid

An organic acid, also known as benzene carboxylic acid, with the formula below. Melting point is

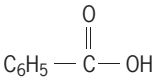

250.2°F (121.2°C), and the acid sublimes at 212°F (100°C). Benzoic acid is only slightly soluble in water but is soluble in most organic solvents, and reacts with bases to form the corresponding benzoate salts. Benzoic acid was first obtained by sublimation from gum benzoin. It occurs both free and combined in nature, being found in many berries (cranberries, prunes, cloves) and as the end product of phenylalanine metabolism.

Preparation. Benzoic acid is prepared in the laboratory by the Grignard reaction, hydrolysis of benzonitrile (C_6H_5CN), or prolonged oxidation of alkyl benzenes with potassium permanganate regardless of the length of the alkyl group. Commercially it was previously prepared by the chlorination of toluene ($C_6H_5CH_3$) with the subsequent hydrolysis of the

benzotrichloride ($C_6H_5CCl_3$), and by the monodecarboxylation of phthalic anhydride (from naphthalene). Modern preparation is by the catalytic oxidation of toluene at elevated temperatures with a cobalt catalyst, and purification by sublimation. *See* GRIGNARD REACTION.

Reactions. Benzoic acid is stable to mild oxidizing agents, but alkyl derivatives of benzene are readily oxidized to the acid. If *o*-phenyl fatty acids contain an odd number of carbon atoms in the chain, benzoic acid is formed, but if the chain contains an even number of carbon atoms, then phenylacetic acid is formed, proving degradation occurs two carbon atoms at a time (beta oxidation).

Benzoic acid undergoes the normal reactions of the aromatic ring (nitration, sulfonation, halogenation, alkylation). Groups are inserted in the meta position due to the directive influence of the carboxyl group. Substitution occurs less readily than with ortho- or para-directing groups due to the deactivating effect of the meta-directing group. Ortho or para derivatives can be obtained with some starting materials other than the acid. Benzoic acid also undergoes the usual reactions of the carboxyl group, forming acyl halides, anhydrides, amides, esters, and salts. *See* HALOGENATION; NITRATION; SUBSTITUTION REACTION; SULFONATION AND SULFATION.

Salts. Sodium benzoate is the only salt of importance. It is water-soluble, has antipyretic and antiseptic properties, is useful as a corrosion inhibitor with sodium nitrite if used for iron, and is also used to modify alkyd resins by increasing hardness, adhesion, and gloss. It is nearly tasteless, only slightly toxic, and is widely used as one of the few materials permitted by law (maximum 0.10%) as a food preservative, for which it is used particularly with acid foods, such as pickles, cider, catsup, and sodas. It is most effective at pH 2.5–4.5, with its effectiveness decreasing as the solution becomes more basic. It is removed from the body (detoxified) in the liver, where it reacts to form benzoyl coenzyme A, which then combines with glycine and is excreted as hippuric acid. Large doses (0.2–0.3 oz/day or 6–8 g/day) of sodium benzoate may cause nausea and vomiting; however, some people are not affected by 0.9 oz/day (25 g/day). In this way, it acts similarly to salicylates. *See* PH.

Esters. Esters of benzoic acid are also found in nature. They are almost universally fragrant. Methyl benzoate is the fragrant principle in tuberose. Some esters of benzoic acid are used in the perfume industry, for example, benzyl ester as a fixative. The butyl ester is used as a dye carrier because of its desirable biodegradable properties; and glycol esters are used as plasticizers. *See* ESTER.

Derivatives. Benzoic acid is a weak acid ($K_a = 6.3 \times 10^{-5}$), being slightly stronger than simple aliphatic acids. However, electrophilic groups substituted on the aromatic ring will increase the acidity; for example, *o*-chlorobenzoic acid has a $K_a = 130 \times 10^{-5}$ and *p*-nitro derivative has a $K_a = 620 \times 10^{-5}$. Hydroxybenzoic acid (salicylic acid) is used to prepare aspirin and salol, the latter as an enteric coating for pills which allows the medicine to pass through the acid stomach essentially unchanged but to be released in the alkaline intestine. *p*-Aminobenzoic acid (PABA) is included in the vitamin B complex and makes up a portion of the folic acid molecule. *o*-Aminobenzoic acid (anthranilic acid) is often used in the manufacture of aniline dyes. Benzoic acid reacts with benzotrichloride to produce benzoyl chloride, from which the peroxide can be made, this being a catalyst for free-radical reactions. *See* PARA-AMINOBENZOIC ACID; ASPIRIN; FREE RADICAL.

Uses. General class of uses for both benzoic acid and its derivatives include the pharmaceuticals and synthetic polymers. Benzoic acid is used in preservatives and many cosmetics. The derivatives are used in the dyeing industry, with some applications in the cosmetic industry. Pure benzoic acid is a standard for bomb calorimetry because of its ease of purification by sublimation. *See* CALORIMETRY; CARBOXYLIC ACID.

Elbert H. Hadley

Bibliography. R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; *Kirk-Othmer Encyclopedia of Chemical Technology*, vol. 3, 4th ed., 1992; *Merck Index*, Merck and Co., 11th ed., 1989; J. D. Morrison and R. N. Boyd, *Organic Chemistry*, 5th ed., 1987.

Benzoyl peroxide

A chemical compound, sometimes called dibenzoyl peroxide, with the formula shown. It is a color-

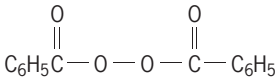

less, crystalline solid, melting point 106–108°C (223–226°F), that is virtually insoluble in water but very soluble in most organic solvents. Benzoyl peroxide is an initiator, a type of material that decomposes at a controlled rate at moderate temperatures to give free radicals. *See* AUTOXIDATION; FREE RADICAL.

Production. Benzoyl peroxide is synthesized in about 70% yield from the reaction of benzoyl chloride with hydrogen peroxide in alkaline aqueous solution or with sodium peroxide in aqueous acetone. It can be purified by precipitation from cold chloroform by addition of methanol. Recrystallization from hot solvents is hazardous; peroxides should be heated only in dilute solution.

Chemical uses. Benzoyl peroxide has a half-life (10 h at 73°C or 163°F in benzene) for decomposition that is very convenient for many laboratory and commercial processes. This, plus its relative stability among peroxides, makes it one of the most frequently used initiators. Its primary use is as an initiator of vinyl polymerization, used in amounts of 0.1 mole % of the vinyl monomer. It also is the preferred bleaching agent for flour; is used to bleach many commercial waxes and oils; and is the active ingredient in commercial acne preparations. *See* BLEACHING; POLYMERIZATION.

Hazards. Benzoyl peroxide itself is neither a carcinogen nor a mutagen. However, when coapplied

with carcinogens to mouse skin it is a potent promoter of tumor development. All peroxides should be treated as potentially explosive, but benzoyl peroxide is one of the least prone to detonate. However, it should be stored in paper or plastic containers and not in rigid, screw-top containers since unscrewing a cap that contains benzoyl peroxide crystals in its threads can result in detonation. *See* MUTAGENS AND CARCINOGENS.

Reactions. Like all peroxides, benzoyl peroxide decomposes by rupture of its oxygen-oxygen peroxide bond, reaction (1), to give two benzoyloxyl

$$C_6H_5\overset{\overset{\displaystyle O}{\|}}{C}-O-O-\overset{\overset{\displaystyle O}{\|}}{C}C_6H_5 \longrightarrow 2\,C_6H_5\overset{\overset{\displaystyle O}{\|}}{C}-O\cdot \qquad (1)$$

radicals. (The odd electron in the free radicals produced in this process is usually represented by a dot.) The benzoyloxyl radical rapidly decarboxylates (loses CO_2) to give the phenyl radical (C_6H_5), reaction (2). Thus, benzoyl peroxide provides a source

$$C_6H_5\overset{\overset{\displaystyle O}{\|}}{C}-O\cdot \longrightarrow C_6H_5\cdot + CO_2 \qquad (2)$$

of phenyl and benzoyloxyl radicals that can perform chemical reactions. For example, the phenyl radicals can add to the double bond of vinyl chloride (C_2H_2Cl) to form a new radical and thus initiate the polymerization of the vinyl monomer, reaction (3).

$$C_6H_5\cdot + CH_2=CHCl \longrightarrow C_6H_5-CH_2-\dot{C}HCl \qquad (3)$$

Accelerated and induced decomposition. Many materials cause benzoyl peroxide to decompose at an accelerated rate over that observed in pure benzene. For example, solvents with high reactivity toward free radicals (such as diethyl ether) cause a greatly accelerated decomposition. For this reason, concentrated solutions of benzoyl peroxide should be prepared only by experienced workers. Compounds that can act as nucleophiles, reducing agents, or electron donors also can greatly accelerate the decomposition; examples are amines and sulfides. Transition-metal ions such as iron or copper also cause an increased rate of decomposition; this can be utilized in commercial practice to achieve a lower temperature or more rapid decomposition.

All free radicals, including radicals from the peroxide itself, attack the peroxide at the oxygen-oxygen bond and cause it to decompose in a radical-induced decomposition. For example, phenyl radicals cause the induced decomposition shown in reaction (4).

$$C_6H_5\cdot + C_6H_5\overset{\overset{\displaystyle O}{\|}}{C}-O-O-\overset{\overset{\displaystyle O}{\|}}{C}C_6H_5 \longrightarrow C_6H_5O-\overset{\overset{\displaystyle O}{\|}}{C}C_6H_5 + C_6H_5CO_2\cdot \qquad (4)$$

See ORGANIC REACTION MECHANISM. William A. Pryor

Bibliography. W. Ando, *Organic Peroxides*, 1992; D. C. Nonhebel and J. C. Walton, *Free Radical Chemistry*, 1974; S. Patai, *The Chemistry of Peroxides*, 1984; W. A. Pryor, *Free Radicals*, 1966.

Bering Sea

A water body north of the Pacific Ocean, 875,900 mi^2 (2,268,000 km^2) in area, bounded by the Siberian and Alaskan landmasses to the west and east respectively, and the Aleutian-Komandorskiye island arc to the south. The Bering Sea is very productive biologically, and the ecosystem supports large populations of fish, shellfish, and marine birds and mammals.

The Bering Sea consists of a large, deep basin in the southwest portion, with depths as great as 10,000 ft (3000 m). To the north and east, an extremely wide, shallow continental shelf extends north to the Bering Strait. These two major regions are separated by a shelf break, the position of which coincides with the southernmost extent of sea ice in a cold season (see **illus.**). Ice is a prominent feature over the Bering Sea shelf during the winter months. Coastal ice begins to form in late October, and by February coastal ice may be found as far south as the Aleutians, such that the ice edge in the eastern Bering Sea advances and retreats seasonally over a distance as great as 600 mi (1000 km). Ice-free conditions can be expected throughout the entire region by early July. Extraordinary interannual variability in ice cover, air, sea-surface temperature, and winds have been observed in the region over the past decades.

The main water connections with the Pacific Ocean are in the west part of the Aleutian Islands: the 6000-ft-deep (2000-m) pass between Attu and the Komandorskiye Islands and the 14,000-ft-deep (4400-m) pass between the Komandorskiyes and Kamchatka. The Bering Sea connection with the Arctic Ocean (Chukchi Sea) is the Bering Strait, 53 mi (85 km) wide and 140 ft (45 m) deep. This is the only avenue of exchange for ice, water, heat, and nutrients between the Pacific and Arctic oceans. *See* ARCTIC AND SUBARCTIC ISLANDS; ARCTIC OCEAN; PACIFIC OCEAN.

Temperature and salinity. Over the continental shelf, the waters receive an excess of precipitation over evaporation, and about 1% of their volume is provided annually by river runoff. The hydrographic structure over the shelf is determined principally by the mechanical stirring of tidal action and wind mixing, and there is a negligible mean flow of water. Large seasonal variations in temperature occur in this area. In winter, the water temperatures are extremely cold, and under the ice they approach the freezing point (29°F or −1.7°C). Salinities range between 31 and 33 parts per thousand (‰). In summer, surface temperature rise to 43–50°F (6–10°C), and occasionally higher near the coast.

The shelf waters are separated from the adjacent oceanic waters by semipermanent fronts. These fronts are detectable as a relatively strong horizontal gradient in upper-layer salinity. The location of these

Map of the Bering Sea showing the position of the shelf and shelf break.

fronts appears to be tied to the continental-shelf break. They are centered over about 480 ft (150 m) depth, with a typical width of about 31 mi (50 km). These are areas of extremely high phytoplankton production and are critical to the southeast Bering Sea ecosystem. The oceanic water abutting the front on the seaward side is relatively invariant in temperature and salinity compared with the waters over the shelf, although the surface waters show some seasonal variation. A temperature maximum is found between 600 and 1300 ft (200 and 400 m), below which the temperature decreases to bottom values of 34.7–34.8°F (1.5–1.6°C). Between 300 and 600 ft (100 and 200 m) there is a halocline where salinity increases from the relatively low upper-layer values to over 34‰. The deep water exhibits a gradual increase in salinity to bottom values of 34.6–34.7‰. At the top of the halocline, about 30 ft (10 m), is a temperature minimum in summer, possibly in part due to water cooled locally during the previous winter. In winter, water of this density is also produced in large quantities on the shelf and its embayments, notably the Gulf of Anadyr, and is advected over the basin beneath the surface layer. *See* CONTINENTAL MARGIN; SEAWATER.

Tides and currents. Tides in the Bering Sea are semidiurnal, with a strong diurnal inequality typical of North Pacific tides. The mean range varies from about 4.8 ft (1.5 m) in the Aleutians to 12 in. (30 cm) in the Bering Strait, but are much larger in Bristol Bay (greater than 13 ft or 4 m). Three water masses are associated with the Bering Sea—Western Subarctic, Bering Sea, and the Alaska Stream. The general circulation of the Bering Sea is counterclockwise, with many small eddies superimposed on the large-scale pattern. The currents in the Bering Sea are generally an inch (a few centimeters) per second except along the continental slope, along the coast of Kamchatka, and in certain eddies, where somewhat higher values have been found. *See* OCEAN CIRCULATION; TIDE.

The direction of the overall flow over the Bering Sea shelf is to the north, ultimately through the Bering Strait into the Chukchi Sea. This flow is driven by a mean sea-level slope that results from the difference in sea level between the Atlantic and Pacific oceans. St. Lawrence Island forms two straits: the

Anadyr Strait to the west and Shpanberg Strait to the east. The most rapid flow over the Bering Sea shelf takes place through the Anadyr Strait, where the flow rate can reach 6 in. (15 cm) per second. This water originates as a northward offshoot of the Bering Slope Current, which flows along the Bering Sea outer shelf from east to west, and bifurcates when it reaches the Siberian coast. This northward-flowing Anadyr water is very rich in nutrients and is responsible for the high productivity of the northern Bering Sea, whereas the water farther to the east over the shelf is relatively unproductive. Three water masses have been identified in the north Bering Sea: Alaska Coastal Water, Bering Shelf Water, and Anadyr Water.

Biological production. Primary production in the Bering Sea is highly variable seasonally and spatially. High primary production in the form of a spring bloom is found at the retreating ice edge and may be more sustained along the frontal areas along the shelf. The highest sustained primary production rates occur in the Anadyr Water in the Gulf of Anadyr, the northern Bering Sea, and the Bering Strait. A few sites on the shelf have been identified as particularly effective in translating the high biological productivity into higher trophic levels. For example, the Chirikov Basin to the north of St. Lawrence Island is a major deposition site for biological material, as the Anadyr Current flows through the region. As a result, the basin is a major summer feeding ground for gray whales, which feed on the abundant tube-dwelling amphipods that live in the bottom. In the southeast portion of the shelf, the sediment-dwelling fauna (infauna) is dominated by polychaete worms, small clams, and brittle stars. The sediment-surface-dwelling animals are dominated by echinoderms in the northern Bering Sea and by commercially important crabs in the southeast. Total annual primary productivity varies between 0.48 and 1.3 oz carbon/ft^2 (150 and 400 g C/m^2), depending on the specific location. This input, along with input of organic matter from the river systems and from the coastal ecosystems (especially sea-grass beds in the Aleutians), supports a large population of organisms at higher trophic levels. Benthic (bottom-living) organisms vary in abundance, depending on the area.

A large proportion of southeastern Bering Sea bottom-living organisms are of direct use as food for humans. King crabs, snow crabs, and snails of the genus *Neptunea* are among the most abundant invertebrates, and the area may also support clam resources of commercial magnitude. Shrimp, especially pink shrimp, are also of potential commercial use. The most abundant fish species are walleye pollock and yellowfin sole. Pacific cod, sablefish, Pacific salmon, perch, herring, capelin, sculpin, plaice, halibut, other sole species, and flounder also occur.

Marine mammals, numbering about a million on the ice pack alone, include seals, whales, walruses, sea otters, and polar bears. A variety of marine birds are also associated with sea ice in the Bering Sea, including three species of gulls, old squaws, king eiders, kittiwakes, murres, and guillemots. Shearwaters and auklets are also abundant in the Bering Sea. *See* BIOLOGICAL PRODUCTIVITY; MARINE ECOLOGY.

Fisheries. The major groundfish fishery of the Bering Sea is pollock (*Theragra chalcogramma*). With a yearly harvest approaching 2.2 million tons (2 million metric tons), pollock constitutes 60% of the fishery. There is great concern about the impacts of this pollock fishery on other species in the Bering Sea ecosystem, especially the Steller sea lion (*Eumetopias jubata*). International issues have come into play in connection with the exploitation of the stocks. Fishing activity in the so-called Donut Hole, a region of international waters surrounded by United States and Russian waters, has been of particular concern, and a treaty was negotiated to conserve and manage pollock in the central Bering Sea. The pollock resource is capable of very rapid increases in abundance as well as decreases, related to variations in year-class abundance. Salmon also represents a major segment of the southeast Bering Sea fishery, with major runs in the Bristol Bay area. The herring fishery is small and targets roe for export to Japan.

Most of the fishing over the shelf is confined to the southern portion. The commercially important crabs dominate the southeast portion of the surface-sediment-dwelling fauna (epifauna). The decline of the king crab population and its associated fishery between 1980 and the mid-1990s was dramatic. The red king crab (*Paralithoides camschatica*) is most abundant in the middle shelf area, and in the Bering Sea one species of Tanner crab, *Chionoecetes bairdii*, is found only in the southeastern portion. The crab population of the northern Bering Sea is dominated by a second species of Tanner crab, *C. opilio*. *See* MARINE FISHERIES.

Environmental impact. Major changes in the Bering Sea environment over recent decades have included a long-term decline in the Steller sea lion population; their listing as an endangered species has resulted in regulatory impacts on the pollock fishery. In the late 1960s, a climatic regime shift produced increased temperatures and reduced sea ice. The most dramatic biological changes took place in 1997 and 1998, with a major die-off of marine birds, drastically reduced salmon runs in western Alaska, declines in sea otters, and a dense bloom of coccolithophorids which persisted over the Bering Sea shelf throughout the entire summer season (coccolithophorids had not previously been reported in large numbers in the Bering Sea). While the causes and interrelationships for these biological events remain obscure, physical factors and a reduced nutrient availability are implicated. *See* ENDANGERED SPECIES; FISHERIES ECOLOGY.

Vera Alexander

Bibliography. L. K. Coachman, Circulation, water masses, and fluxes on the southeastern Bering Sea shelf, *Continental Shelf Res.*, 5:23–108, 1986; L. K. Coachman and D. A. Hansell (eds.), ISHTAR: Inner Shelf Transfer and Recycling in the Bering and Chukchi Sea, *Continental Shelf Res.* (spec. iss.), 13:473–704, 1993; National Research Council, Committee on the Bering Sea Ecosystem, *The Bering Sea Ecosystem*, National Academy of Sciences

Press, Washington, DC, 1996; H. J. Niebauer and V. Alexander, Oceanographic frontal structure and biological production at an ice edge, *Continental Shelf Res.*, 4:367–388, 1985; North Pacific Marine Science Organization, *Dynamics of the Bering Sea*, ed. by T. Loughlin and K. Ohtani, University of Alaska Sea Grant, AK-SG-99-03, Fairbanks, 1999; J. D. Schumacher and P. J. Stabeno, The continental shelf of the Bering Sea, in K. H. Brink and A. R. Robinson (eds.), *The Sea*, vol. 11, Wiley, New York, 1998; P. J. Stabeno et al., Under-ice observations of water column temperature, salinity and spring phytoplankton dynamics: Eastern Bering Sea shelf, *J. Mar. Res.*, 56:239–255, 1998; J. J. Walsh et al., The role of the Bering Strait in the carbon/nitrogen fluxes of polar marine ecosystems, *6th Conference of the Comité Arctique International*, Fairbanks, Alaska, 1989; J. J. Walsh and C. P. McRoy, Ecosystem analysis in the southeastern Bering Sea, *Continental Shelf Res.*, 5:259–288, 1986.

Berkelium

Element number 97, symbol Bk, the eighth member of the actinide series of elements. In this series the 5*f* electron shell is being filled, just as the 4*f* shell is being filled in the lanthanide (rare-earth) elements. These two series of elements are very similar in their chemical properties, and berkelium, aside from small differences in ionic radius, is especially similar to its homolog terbium. *See* PERIODIC TABLE; RARE-EARTH ELEMENTS; TERBIUM.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**	
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**	

Berkelium does not occur in the Earth's crust because it has no stable isotopes. It must be prepared by means of nuclear reactions using more abundant target elements. These reactions usually involve bombardments with charged particles, irradiations with neutrons from high-flux reactors, or production in a thermonuclear device.

Berkelium metal is chemically reactive, exists in two crystal modifications, and melts at 986°C (1806°F). Berkelium was discovered in 1949 by S. G. Thompson, A. Ghiorso, and G. T. Seaborg at the University of California in Berkeley and was named in honor of that city. Nine isotopes of berkelium are known, ranging in mass from 243 to 251 and in half-life from 1 hour to 1380 years. The most easily produced isotope is ^{249}Bk, which undergoes beta decay with a half-life of 314 days and is therefore a valuable source for the preparation of the isotope ^{249}Cf. The berkelium isotope with the longest half-life is ^{247}Bk (1380 years), but it is difficult to produce in sufficient amounts to be applied to berkelium chemistry studies. *See* ACTINIDE ELEMENTS; TRANSURANIUM ELEMENTS.

Glenn T. Seaborg

Bibliography. S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; G. T. Seaborg (ed.), *Transuranium Elements: Products of Modern Alchemy*, 1978; G. T. Seaborg and W. D. Loveland, *The Elements Beyond Uranium*, 1990.

Bermuda grass

A long-lived perennial (*Cynodon* spp.) that originated in Africa. It is believed that in the early 1500s Spanish explorers unintentionally brought the first common Bermuda grass seeds to the Western Hemisphere with the hay that fed their horses. The weedy common type, *C. dactylon* (the type species), can now be found throughout the tropical and subtropical regions of the world. *See* CYPERALES.

Cynodon dactylon is an extensively creeping grass with both stolons and rhizomes. It has short leaves borne on 8–12-in. (20–30-cm) upright culms with several slender seed-bearing spikes digitate at their summit (see **illus.**). It is propagated by planting seed, stolons, rhizomes, or green stems cut at an advanced hay stage. It is well adapted to a wide range of soil types, tolerates a pH range from 4.0 to over 8.0, prefers well-drained soils, is very drought-tolerant, and responds well to fertilization. It grows best during high summer temperatures, grows little at temperatures below 65°F (18°C), and goes dormant when temperatures drop below 30°F (−1°C). Although a weed in cultivated fields, it controls erosion, makes excellent turf, and supplies good grazing for all classes of livestock. *See* LAWN AND TURF GRASSES.

Bermuda grass is a highly variable, heterozygous, cross-pollinated species. Breeding work, begun in

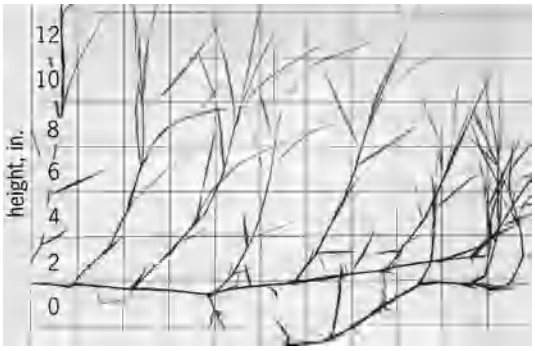

Coastal Bermuda grass, showing its surface runners, characteristic seedheads, and short leaves. (*From G. W. Burton, Coastal Bermuda grass, Ga. Agr. Exp. Sta. Bull., N.S. 2, 1954*)

1936, has produced a number of superior F_1 hybrids that are sterile and must be propagated vegetatively. Vegetative propagation is facilitated by the rapid spread of the aboveground stolons of the new hybrids. Plants set on 6 × 6 ft (2 × 2 m) centers with adequate fertilization, water, and weed control cover a pasture in two or three summer months and produce up to a ton of hay per acre (2.5 metric tons per hectare).

In developing countries, the improved Bermuda grass hybrids can be established by pushing the stolons into moist soil with a pointed stick. In developed countries, tractor-drawn machines are usually used to dig and plant the hybrids.

The first hybrid to be released, in 1943, was the best of 5000 hybrids between a very good common Bermuda grass and one introduced from South Africa. It was named Coastal for the Georgia Coastal Plain Station where it was bred, and it has been planted on more than 10 million acres (4 million hectares) across the South. Compared with common Bermuda grass, Coastal yields twice as much, is more disease-resistant and drought-tolerant, uses fertilizer and water more efficiently, and grows tall enough to be cut for hay. Coastal Bermuda has retained its superiority for 40 years and has been an excellent parent to produce improved varieties. Coastal crossed with a highly digestible, cold-susceptible *C. nlemfuensis* plant from Kenya produced the sterile F_1 hybrid Coastcross-1, which is 12% more digestible than Coastal, and when grazed by cattle gives 30–40% better daily gains and gains per acre. Coastcross-1 is less cold-tolerant than Coastal. Tifton 44 is the best of several thousand F_1 hybrids between Coastal and a winter hardy common Bermuda grass from Berlin, Germany. It is more cold-tolerant, more digestible, and when grazed gives up to 19% better daily gains than Coastal. Because it is male-sterile and equal or superior to Coastal in other important characteristics, Tifton 44 has replaced Coastal as the female for new forage hybrids being developed at Tifton, Georgia.

The top turf Bermuda grasses, Tifgreen, Tifway, and Tifdwarf, are the best of a number of F_1 hybrids between selected plants of *C. dactylon* and *C. transvaalensis*. The parents have 36 and 18 chromosomes as their respective $2n$ number and the Tif hybrids have 27. This chromosome number indicates sterility, making vegetative propagation necessary. Tifdwarf, tolerant of daily mowing at $^3/_{16}$ in. (0.5 cm), is unsurpassed for use as top-quality turf for golf greens and bowling greens. Tifgreen, also bred for golf greens, must be mowed at a higher cut. Tifway, more resistant to pests, weeds, and frost, is the best Bermuda grass for golf fairways and tees, lawns, and athletic fields. Tifway II and Tifgreen II are improved mutants of Tifway and Tifgreen created by exposing dormant rhizome sections to gamma radiation. Most of the mutants so created were inferior to their parents. The sterile turf hybrids are planted by machines as sprigs or sod. When sod of these hybrids is stripped from nurseries, enough rhizomes (underground stems) are left in the soil to reestablish the sod.

People suffering from asthma and hay fever find Bermuda grass pollen to be one of the worst offenders. The sterile turf Bermuda grass hybrids solve this problem by producing no pollen. *See* ASTHMA; GRASS CROPS; POLLEN.

Glenn W. Burton

Bernoulli's theorem

An idealized algebraic relation between pressure, velocity, and elevation for flow of an inviscid fluid. Its most commonly used form is for steady flow of an incompressible fluid, and is given by the equation below, where p is pressure, ρ is fluid density

$$\frac{p}{\rho} + \frac{V^2}{2} + gz = \text{constant}$$

(assumed constant), V is flow velocity, g is the acceleration of gravity, and z is the elevation of the fluid particle. The relation applies along any particular streamline of the flow. The constant may vary across streamlines unless it can be further shown that the fluid has zero local angular velocity. *See* KELVIN'S CIRCULATION THEOREM.

The above equation may be extended to steady compressible flow (where changes in ρ are important) by adding the internal energy per unit mass, e, to the left-hand side. *See* COMPRESSIBLE FLOW.

The equation is limited to inviscid flows with no heat transfer, shaft work, or shear work. Although no real fluid truly meets these conditions, the relation is quite accurate in free-flow or "core" regions away from solid boundaries or wavy interfaces, especially for gases and light liquids. Thus Bernoulli's theorem is commonly used to analyze flow outside the boundary layer, flow in supersonic nozzles, flow over airfoils, and many other practical problems. *See* AERODYNAMICS; BOUNDARY-LAYER FLOW.

According to the above equation, if velocity rises at a given elevation, pressure must fall. This principle is used in Bernoulli-type flow meters, such as orifice plates, venturi throats, and pitot tubes. A flow obstruction is deliberately created to cause a local pressure drop which can be calibrated with the flow rate. *See* FLOW MEASUREMENT.

By invoking additional restrictive assumptions, the Bernoulli theorem can be extended to unsteady flow with zero or constant angular velocity, flow in rotating machinery, and piping systems with frictional losses. *See* PIPE FLOW; TURBINE.

It is named after Daniel Bernoulli, who hinted at a "proportionality" between pressure and velocity in a 1738 hydrodynamics textbook. But the above equation was actually derived by Leonhard Euler in a series of papers in 1755. *See* FLUID FLOW.

Frank M. White

Bibliography. V. L. Streeter et al., *Fluid Mechanics*, 9th ed., 1998; F. M. White, *Fluid Mechanics*, 4th ed., 1998; C. S. Yih, *Fluid Mechanics*, 2d ed., rev. reprint, 1988.

Beroida

An order of the phylum Ctenophora (comb jellies) comprising two genera, *Beroë* and *Neis*. The single species of *Neis (N. cordigera)* is known only from Australia, but many species of *Beroë* have been described from polar to tropical waters all over the world. Beroids are predators on other ctenophores and occur in all parts of the water column where ctenophores are found.

Beroids are the only ctenophores that lack tentacles throughout their life, and their phylogenetic relationship to other ctenophores is uncertain. The conical or cylindrical body is flattened in the tentacular plane (see **illus.**) and ranges in size from a few millimeters to about 16 in. (40 cm). Most of the body is occupied by the stomodeum. *Neis* differs from *Beroë* in having prominent aboral keels. The body is mostly transparent or translucent but may have bluish, orange, pink, or red coloration of the meridional canals or stomodeum. Pigmentation can vary with age, diet, or environment and is probably not a stable species characteristic. Like other ctenophores, beroids are strongly bioluminescent, producing blue-green light along the meridional canals when disturbed.

Locomotion is produced by the well-developed comb (ctene) rows, which extend along the entire body length in some species. Beroids generally swim mouth first and can turn or reverse quickly. The meridional canals have side branches, which in some species anastomose with each other or with branches of the paragastric canals to form a complex gastrovascular network in the mesoglea. On either side of the aboral statocyst are papillae, which may have a sensory function but do not occur in other ctenophores. Beroids are simultaneous hermaphrodites and have ovaries and testes along the meridional canals. Eggs and sperm are shed through numerous gonopores directly to the outside. A fertilized egg develops directly into a miniature adult, with no trace of the cydippid larva stage characteristic of other ctenophores. Beroids are capable of rapid exponential increase in their local populations.

The mouth is large and extensile, capable of engulfing prey larger than the beroid itself. The lips of some *Beroë* are kept closed by special zipper-like strips of tissue while the animal swims. Upon contact with prey, the mouth unzips instantly to swallow it. The inner lips are lined with macrocilia, bundles of ciliary shafts (axonemes) up to 60 micrometers long surrounded by a common membrane. The tips of the macrocilia are formed into pointed teeth, which vary in shape among species. Prey are gripped by these teeth and drawn into the stomodeum by the metachronal beating of the macrocilia. Generally, whole prey are engulfed on contact, but pieces of large prey can be bitten off with the macrocilia. There is morphological and behavioral evidence that chemoreception is involved in detecting and recognizing prey organisms.

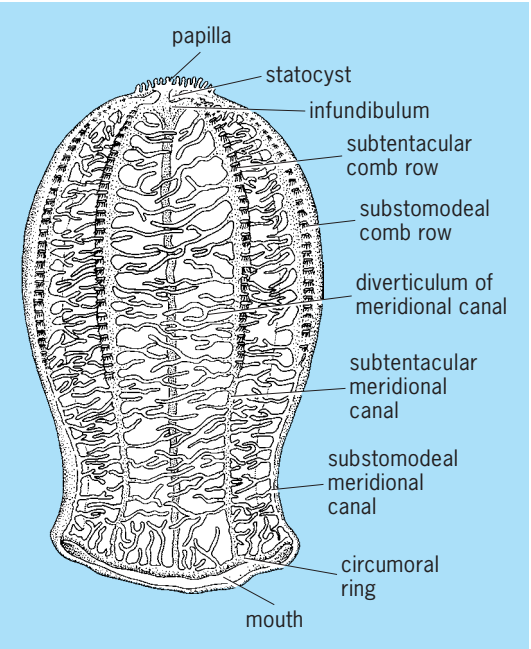

Beroid, *Beroë cucumis*.

Beroids feed mainly on other ctenophores, and in some situations may control the population of species such as *Pleurobrachia*. Some beroids are highly specific, consuming only a single prey species, but other are more general, sometimes eating siphonophores and salps as well as ctenophores. Predators on beroids include medusae, heteropod mollusks, hyperiid amphipods, and various fishes. *See* CTENOPHORA. L. P. Madin

Bibliography. G. R. Harbison, L. P. Madin, and N. R. Swanberg, On the natural history and distribution of oceanic ctenophores, *Deep-sea Res.*, 25:233–256, 1978; F. W. Harrison and J. A. Westfall (eds.), *Microscopic Anatomy of Invertebrates*, vol. 2, 1991; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Beryciformes

An order of somewhat intermediate position among actinopterygian fishes. It shares with the perciforms fin spines and ctenoid scales, an upper jaw bordered by the premaxillae, a ductless swim bladder, and absence of a mesocoracoid. An orbitosphenoid bone is present, as is typical of many lower teleosts. The pelvic fins are thoracic or subabdominal in position, and each has a spine and 3 to 13 (usually more than 5) soft rays; the pelvic girdle is most often attached to the cleithra; and there are 18 or 19 principal caudal rays. *See* PERCIFORMES; SWIM BLADDER.

Classification. Beryciforms are classified in 7 families as shown below, 28 genera, and 123 species.

Anoplogastridae (fangtooths). Body short and deep; numerous long fanglike teeth; small eyes; small or minute scales; lateral line an open groove; fins

Yellowfin soldierfish (*Myripistis chryseres*). (*Photo © John E. Randall*)

without spines; maximum length 16 cm (6.3 in.). Bathypelagic in the Atlantic, Pacific, and Indian oceans to depths of 5000 m (16,500 ft).

Diretmidae (spinyfins). No lateral line; dorsal and anal spines absent, pelvic fin with laminar spine; abdomen with sharp edge formed by ventral scutes; maximum length 37 cm (15.6 in.). Found in the Atlantic, Pacific, and Indian oceans.

Anomalopidae (flashlight fishes or lanternfishes). Light organ located beneath the eye with a shutter mechanism for controlling the emission of light, which is produced by symbiotic bacteria. Maximum length 27 cm (10.6 in.). Primarily found in tropical waters of the Indo-Pacific, from the surface to 300 m (1000 ft) depth.

Monocentridae (pinecone fishes). Body covered with large, heavy, platelike scales; organs on the lower jaw to emit light produced by luminous bacteria; maximum size 21 cm (8.3 in.). Found in tropical and subtropical inhabitats in Indo-Pacific ocean at depths of 30 to 300 m (100 to 1000 ft).

Trachichthyidae (roughies or slimeheads). Angle of the preopercle and posttemporal bone, each possessing distinct spine; median ridge of scutes on abdomen; maximum length 55 cm (22 in.). Found in the Atlantic, Indian, and Pacific oceans at depths of 100 to 1500 m (330 to 5000 ft).

Berycidae (alfonsinos). Compressed body; spiny scales; very large eyes; a large and oblique mouth; a single, relatively short dorsal fin, without a notch; maximum length has not been determined, but western Atlantic species probably do not exceed 60 cm (24 in.). Found in the Atlantic, Indian, western and central Pacific oceans at depths to 200 to 600 m (660 to 2000 ft).

Holocentridae (squirrelfishes and soldierfishes). Largest family of beryciform fishes (65 species), more perchlike than other families of the order (see **illustration**). Body compressed; opercle with spines, large eyes; the ctenoid scales large and very rough; long dorsal fin long, divided by a notch, with 10 to 13 sharp spines and 11 to 17 rays, reddish in color; a maximum body length of 61 cm (24 in.). Found in tropical to temperate waters of the Atlantic, Indian, and Pacific oceans, mostly from shore to a depth of 100 m (330 ft), rarely over 200 m (660 ft).

Squirrelfishes feed on benthic invertebrates and small fishes, soldierfishes on large zooplankton; both are nocturnal. These are important food fishes in many parts of the world. Fifteen species are found in Atlantic and Pacific waters of Mexico and North America. *See* ACTINOPTERYGII; OSTEICHTHYES; TELEOSTEI.

Reeve M. Bailey; Herbert Boschung

Bibliography. G. D. Johnson and R. H. Rosenblatt, Mechanism of light organ occlusion in flashlight fishes, family Anomalopidae (Teleostei: Beryciformes), and the evolution of the group, *Zool. J. Linn. Soc.*, 94:65–96, 1988; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994; L. P. Woods and P. M. Sonoda, Order Berycomorphi (Beryciformes): Fishes of the western North Atlantic, *Sears Found. Mar. Res. Mem.*, no. 1, (6):263–396, 1973.

Beryl

The most common beryllium mineral. Beryl, $Al_2[Be_3Si_6O_{18}]$, crystallizes in the hexagonal system with space group P6/mcc. The crystal structure consists of six-membered rings of corner-sharing silicon-oxygen (SiO_4) tetrahedra cross-linked by corner-sharing beryllium-oxygen (BeO_4) tetrahedra to make a three-dimensional honeycomb structure; aluminum-oxygen (AlO_6) octahedra lie between the Si_6O_{18} rings (see **illus.**). Cordierite, $(Mg,Fe)_2[Al_4Si_5O_{18}]$, has a similar structure. Open channels parallel to the sixfold axis contain variable amounts of alkalies (up to 5–8 wt %), such as lithium (Li), sodium (Na), potassium (K), rubidium (Rb), and cesium (Cs), and water molecules (up to ~2.5 wt %). Minor amounts of iron (Fe), chromium (Cr), manganese (Mn), and other metals substitute for aluminum, contributing to the various colors of natural beryl. Beryl has a vitreous luster and is

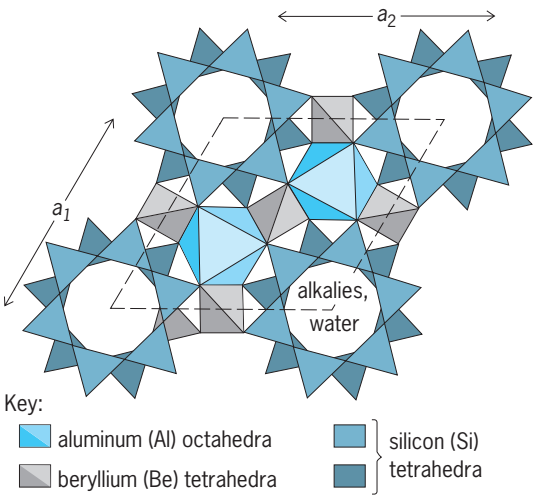

Crystal structure of beryl looking parallel to the sixfold axis. (*Modified from C. Klein, Manual of Mineralogy, 21st ed., John Wiley and Sons, Inc., 1993*)

typically white to bluish- or yellowish-green, but it can also be shades of yellow, blue, and pink. Its hardness is 7.5–8 on Mohs scale; it has an imperfect basal cleavage and a specific gravity of 2.7–2.9 (increasing with alkali content). Weakly colored varieties can be confused with quartz or apatite. *See* CRYSTAL STRUCTURE; HARDNESS SCALES.

Beryl is stable over a wide range of conditions from room conditions to more than 2200°F (1200°C) and more than 8 kilobars (800 megapascals). At high temperatures, beryl melts to silica-rich liquid, phenakite (Be_2SiO_4), and chrysoberyl ($BeAl_2O_4$). Under hydrous conditions at low temperatures (less than about 550°F or 300°C), beryl breaks down to a variety of minerals, including bertrandite, phenakite, kaolinite, and muscovite. Beryl has been synthesized from oxides and carbonates by dry and hydrothermal methods over a large range of temperature and pressure.

Beryl is a minor accessory mineral in many natural environments, most commonly in granites and associated hydrothermally altered rocks. Granitic pegmatites constitute the major source of beryl (used for beryllium and gemstones); rarely, single crystals weigh many tons. Alkali-rich beryl occurs in complex pegmatites which contain abundant rare-element minerals such as spodumene, lepidolite, and tourmaline. Alkali-poor beryl occurs in mineralogically simple pegmatites, tin and tungsten deposits, and hydrothermal veins. The gem varieties of beryl, aquamarine (blue), emerald (deep green), and morganite (pink to red), are produced from pegmatites (aquamarine, morganite, some emerald), veins (some aquamarine, most emerald), and, rarely, rhyolites (ruby-red morganite). Important pegmatite provinces occur in Brazil, Russia, Madagascar, New England, and California. Beryl is recovered from mineralized granites in Russia and is widespread in granite-related ore deposits from all continents. Emerald occurs in mica schists in Egypt, Austria, the Urals, and North Carolina; the largest, most important emerald deposits are albite-bearing quartz veins in bituminous limestone in Colombia. *See* BERYLLIUM; BERYLLIUM MINERALS; PEGMATITE; SILICATE MINERALS.

Mark D. Barton

Bibliography. J. Sinkankas, *Emerald and Other Beryls*, 1981, reprint 1989.

Beryllium

A chemical element, Be, atomic number 4, with an atomic weight of 9.0122. Beryllium, a rare metal, is one of the lightest structural metals, having a density about one-third that of aluminum. Some of the important physical and chemical properties of beryllium are given in the **table**. Beryllium has a number of unusual and even unique properties. *See* PERIODIC TABLE.

The largest volume uses of beryllium metal are in the manufacture of beryllium-copper alloys and in the development of beryllium-containing moderator and reflector materials for nuclear reactors. Addition of 2% beryllium to copper forms a nonmagnetic alloy which is six times stronger than copper. These beryllium-copper alloys find numerous applications in industry as nonsparking tools, as critical moving parts in aircraft engines, and in the key components of precision instruments, mechanical computers, electrical relays, and camera shutters. Beryllium-copper hammers, wrenches, and other tools are employed in petroleum refineries and other plants in which a spark from steel against steel might

1												13	14	15	16	17	18
1 H	2																2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Physical and chemical properties of beryllium

Property	Value
Atomic and mass properties	
Mass number of stable isotopes	9
Atomic number	4
Outer electronic configuration	$1s^2 2s^2$
Atomic weight	9.0122
Atomic diameter	0.221 nm
Atomic volume	4.96 cm^3/mole
Crystal structure	Hexagonal close-packed
Lattice parameters	$a = 0.2285$ nm $c = 0.3583$ nm
Axial ratio	$c/a = 1.568$
Field of cation (charge/radius^2)	17
Density*, 25°, x-ray (theoretical)	1.8477 ± 0.0007 g/cm^3
Density, 1000°, x-ray	1.756 g/cm^3
Radius of atom (Be^0)	0.111 nm
Radius of ion, Be^{2+}	0.034 nm
Ionization energy ($Be^0 \rightarrow Be^{2+}$)	27.4 eV
Thermal properties	
Melting point	1285°C (2345°F)
Boiling point†	2970°C (5378°F)
Vapor pressure ($T = K°$)	$\log P$ (atm) $= 6.186 + 1.454 \times 10^{-4}T - (16{,}700/T)$
Heat of fusion	250–275 cal/g
Heat of vaporization	53,490 cal/mole
Specific heat (20–100°)	0.43–0.52 cal/(g)(°C)
Thermal conductivity (20°)	0.355 cal/(cm^2)(cm)(s)(°C) (42% of copper)
Heat of oxidation	140.15 cal
Electrical properties	
Electrical conductivity	40–44% of copper
Electrical resistivity	4 microhms/cm (0°C) 6 microhms/cm (100°C)
Electrolytic solution potential, Be/Be†	$E^0 = -1.69$ volts
Electrochemical equivalent	0.04674 mg/coulomb

*Measured values vary from 1.79 to 1.86, depending on purity and method of fabrication.
†Obtained by extrapolation of vapor pressure data, not considered very reliable.

lead to an explosion or fire. *See* ALKALINE-EARTH METALS; BERYLLIUM ALLOYS.

Beryllium has found many special uses in nuclear energy because it is one of the most efficient materials for slowing down the speed of neutrons and acting as a neutron reflector. Consequently, much beryllium is used in the construction of nuclear reactors as a moderator and as a support or alloy with the fuel elements. *See* NUCLEAR REACTOR.

The following list shows some of the principal compounds of beryllium. Many of the compounds listed are useful as intermediates in the processes for the preparation of ceramics, beryllium oxide, and beryllium metal. Other compounds are useful in analysis and organic synthesis.

Acetylacetonate
Ammonium beryllium fluoride
Aurintricarboxylate
Basic acetate, $BeO \cdot Be_3(CH_3COO)_6$
Basic beryllium carbonate
Beryllate, BeO_2^{2-}
Beryllium ammonium phosphate
Bromide, BeB_2
Carbide, Be_2C
Chloride, $BeCl_2$
Dimethyl, $Be(CH_3)_2$
Fluoride, BeF_2
Hydroxide
Nitrate, $Be(NO_3)_2 \cdot 4H_2O$
Nitride, Be_3N_2
Perchlorate, $Be(ClO_4)_2 \cdot 4H_2O$
Oxide, BeO
Plutonium-beryllium, $PuBe_{13}$
Salicylate
Silicates (emerald)
Sulfate, $BeSO_4 \cdot 4H_2O$
Uranium-beryllium, UBe_{13}

Beryllium is surprisingly rare for a light element, constituting about 0.005% of the Earth's crust. It is about thirty-second in order of abundance, occurring in concentrations approximating those of cesium, scandium, and arsenic. Actually, the abundances of beryllium and its neighbors, lithium and boron, are about 10^{-5} times those of the next heavier elements, carbon, nitrogen, and oxygen. At least 50 different beryllium-bearing minerals are known, but in only about 30 is beryllium a regular constituent. The only beryllium-bearing mineral of industrial importance is beryl. Bertrandite substitutes for beryl as a domestic source in the United States. *See* BERYL.

Jack Schubert

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; D. F. Shriver and P. W. Atkins, *Inorganic Chemistry*, 3d ed., 1999.

Beryllium alloys

Dilute alloys of base metals which contain a few percent of beryllium in a precipitation-hardening system. They are the principal useful beryllium alloys manufactured today. Although beryllium has some solid solubility in copper, silver, gold, nickel, cobalt, platinum, palladium, and iron and forms precipitation-hardening alloys with these metals, the copper-beryllium system and, to a considerably lesser degree, the nickel-beryllium alloys are the only ones used commercially.

TABLE 1. Composition of principal beryllium-copper alloys

Alloy grade	Be, %	Other, %
25	1.80–2.05	0.20–0.35 Co
165	1.60–1.79	0.20–0.35 Co
10	0.40–0.70	2.35–2.70 Co
50	0.25–0.50	1.40–1.70 Co 0.90–1.10 Ag
275 C	2.60–2.85	0.35–0.65 Co
245 C	2.30–2.55	0.35–0.65 Co
20 C	2.00–2.25	0.35–0.65 Co
165 C	1.60–1.85	0.20–0.65 Co
10 C	0.55–0.75	2.35–2.70 Co
50 C	0.40–0.65	1.40–1.70 Co 1.00–1.15 Ag

Other than the precipitation-hardening systems, small amounts of beryllium are used in alloys of the dispersion type wherein there is little solid solubility (Al and Mg). Various amounts of beryllium combine with most elements to form intermetallic compounds. Development of beryllium-rich alloys has been chiefly confined to the ductile matrix Be-Al, Be-Cu solid solution alloy with up to 4% Cu, and dispersed-phase type alloys having relatively small amounts of compounds (0.25–6%), chiefly as BeO or intermetallics, for dimensional stability, elevated temperature strength, and elastic limit control. *See* ALLOY.

Beryllium-copper alloys. Commercial alloys of the beryllium-copper group are divided into cast or wrought types, usually in ternary combination such as Cu-Be-Co (**Table 1**). Alloys with higher beryllium content have high strength while those with low beryllium have high electrical and thermal conductivity.

Age-hardenable Cu-Be-Co alloys offer a wide range of properties, because they are extremely ductile and workable in the annealed condition and are strong and hard after precipitation or aging treatment. Cobalt is added to inhibit grain growth and provide uniform heat-treatment response. These alloys also have inherent characteristics of substantial electrical and thermal conductivity and resistance to corrosion and wear, being protected by BeO films which impart this property to all materials containing Be. These age-hardenable alloys resist dimensional change and fatigue.

Properties of beryllium-copper alloys, because of composition and thermal treatment, may vary greatly to provide applications with a wide range of useful characteristics, as shown in **Table 2**.

Applications. Primary applications are found in the electronics, automotive, appliance, instrument, and temperature-control industries for electric current-carrying springs, diaphragms, electrical switch blades, contacts, connectors, terminals, fuse clips, and bellows (foil, strip, and wire), as well as resistance-welding dies, electrodes, clutch rings, brake drums, and switch gear (rod, plate, and forgings). With 1.5% Be or more, the melting point of

TABLE 2. Range of properties of beryllium-copper alloys

Property	Variation possible in commercial alloy
Tensile strength	35,000–215,000 lb/in.^{2} (240–1480 MPa)
Yield strength	10,000–200,000 lb/in.^{2} (70–1400 MPa)
Elongation	0.5–60%
Hardness, Rockwell B and C scale	B_{10}–C_{47}
Thermal conductivity	600–1600 Btu-in./h-ft^2-°F
Electrical conductivity	15–60% IACS
Melting point	1570–2040°F (845–1115°C)

copper is severely depressed and a high degree of fluidity is encountered, allowing casting of intricate shapes with very fine detail. This characteristic is important for plastic injection molds.

For special applications specific alloys have been developed. Free machining and nonmagnetic alloys have been made, as well as high-purity materials. A precipitation-hardening beryllium monel for oceanographic application, containing about 30% Ni, 0.5% Be, 0.5% Si, and the remainder copper, illustrates one of a series of alloys having strength, corrosion resistance to seawater, and antifouling characteristics.

New applications in structural, aerospace, and nuclear fields are submarine repeater cable housings for transoceanic cable systems, wind tunnel throats, liners for magnetohydrodynamic generators for gas ionization, and scavenger tanks for propane-freon bubble chambers in high-energy physics research. Important developing applications for beryllium-copper are trunnions and pivot bearing sleeves for the landing gear of heavy, cargo-carrying aircraft, because these alloys allow the highest stress of any sleeve bearing material.

Production. Beryllium-copper master alloys are produced by direct reduction of beryllium oxide with carbon in the presence of copper in an arc furnace. Because beryllium carbide forms readily, the master alloy is usually limited to about 4.0–4.25% Be. The master alloy is remelted with additional copper and other elements to form commercial alloys. Rolling billets up to 1500 lb (680 kg) have been made by continuous-casting techniques.

Beryllium-copper alloys can be fabricated by all the industrial metalworking techniques to produce principally strip, foil, bar, wire, and forgings. They can be readily machined and can be joined by brazing, soldering, and welding. Annealing, to provide high plasticity at room temperature, is accomplished by heat-treating from 1450 to 1475°F (790 to 802°C) for the high-Be alloys (1650–1700°F or 900–930°C for the low-Be alloys) and water quenching. Precipitation hardening is accomplished by heating to 750–900°F (400–480°C; low Be) and 550–650°F (290–340°C; high Be). *See* COPPER ALLOYS.

Alloys with nickel and iron. Nickel containing 2% Be can be heat-treated to develop a tensile strength of 240,000 lb/in.^{2} (1700 megapascals) with 3–4% elongation. Little commercial use is made of the hard nickel alloys although they have been employed, principally as precision castings, for aircraft fuel pumps, surgical instruments, and matrices in diamond drill bits. Another nickel alloy (Brush M 220C), having 2.0–2.3% Be, 0.5–0.75% C, and the balance of nickel and refractory metals, has been used for mold components and forming tools for glass pressware of optical and container quality. Thermal conductivity, wear resistance, and strength, coupled with unusual machinability for a Ni-Be alloy, make this alloy particularly advantageous for glassworking tooling.

Attempts have been made to add beryllium to a number of ferrous alloys. Small amounts are used to refine grains and deoxidize ferritic steels (Japan), while promising properties have been developed for austenitic and martensitic steels. Stainless steels (Fe-Ni-Cr) may be made maraging by adding 0.15–0.9% Be, developing strengths as high as 260,000 lb/in.^{2} (1800 MPa) as cast (330,000 lb/in.^{2} or 2300 MPa as rolled) while retaining their oxidation- and corrosion-resistant characteristics. Amounts of 0.04–1.5% Be have been added to various Fe-base alloys containing Ni, Cr, Co, Mo, and W for special applications such as watch springs. *See* IRON ALLOYS; NICKEL ALLOYS.

Additions to nonferrous alloys. In materials for nuclear reactors small amounts of beryllium have been added to Mg, Al, Cr, and Zr in Russia for grain refinement and for decreasing surface reactivity. In the United Kingdom a family of reactors (Haddon Hall) used a Mg-Th-Be alloy for fuel element cans.

About 0.1–0.5% Be in Al-Mg alloys refines grains, promotes fluidity, and premits better workability and increased strength. Protection of magnesium from oxidation, burning, and dross formation during casting is achieved by adding as little as 0.005% Be.

Beryllium-base alloys. Three types of beryllium-base alloys are of interest. These consist of dispersed-phase types containing up to 4% BeO; ductile-phase or duplex alloys of Be and Al, particularly 38% Al in Be; and solid solution alloys of up to 4% Cu in Be.

Dispersed-phase alloys containing oxides, carbides, nitrides, borides, and intermetallic compounds in a beryllium matrix are chiefly characterized by increased strength and resistance to creep at elevated temperature. Major commerical alloys in this series are of the fine-grain, high-BeO (4.25–6%), hot-pressed types such as materials used for inertial guidance instruments characterized by high dimensional stability, high-precision elastic limit, 8000–15,000 lb/in.^{2} (55–103 MPa), and good machinability (Brush I-400).

62 Be-38 Al (Lockalloy) was developed as a structural aerospace alloy to combine high modulus and low density with the machining characteristics of the commoner magnesium-base alloys. This alloy in sheet form has at room temperature about 50,000-lb/in.^{2} (344-MPa) ultimate strength, 47,000-lb/in.^{2} (324-MPa) yield strength, and about 8% elongation. It is also produced as extrusions. It has a duplex-type microstructure characterized

TABLE 3. High-temperature, oxidation-resistant beryllides

General category	Example	Wt % of Be	Melting point, °F (°C)	X-ray density, g/cm^3 (oz/in.^{3})
Nb-Be	$NbBe_{12}$	53.8	3070 (1688)	2.92 (1.69)
	Nb_2Be_{19}	48.0	~3100 (1704)	3.17 (1.83)
	Nb_2Be_{17}	45.2	3100 (1704)	3.28 (1.90)
Ta-Be	$TaBe_{12}$	37.4	3360 (1849)	4.18 (2.42)
	Ta_2Be_{17}	29.8	3610 (1988)	5.05 (2.92)
Mo-Be	$MoBe_{12}$	53.2	~3100 (1704)	3.03 (1.75)
Ti-Be	$TiBe_{12}$	69.3	2900 (1593)	2.26 (1.31)
	Ti_2Be_{17}	61.5	2970 (1632)	2.46 (1.42)
Zr-Be	$ZrBe_{13}$	56.2	3500 (1927)	2.72 (1.57)
	Zr_2Be_{17}	45.7	3600 (1982)	3.08 (1.78)

by a semicontinuous aluminum phase. Other alloys of the Be-Al type have been reported, but the 62% Be–38% Al alloy is the most used.

Solid solution alloys containing up to 4% Cu have been intensively investigated because of higher indicated fracture strength, and toughness compared with beryllium.

Intermetallic compounds. Beryllium, combined with most other elements, forms intermetallic compounds having high strength at high temperature—up to 80,000-lb/in.^{2} (552-MPa) modulus of rupture at 2300°F (1260°C), good thermal conductivity, high specific heat, and good oxidation resistance. The BeO film formed by surface oxidation is protective to volatile oxides (Mo) and to elements of high reactivity (Zr and Ti) at temperatures 2000–2700°F (1100–1500°C) and for short times up to 3000°F (1650°C). Some of the lighter refractory beryllides are compared in **Table 3**, showing their composition and low density. Beryllides are also formed with actinide and rare metals, as well as with the transition metals.

The beryllides are of interest to the nuclear field, to power generation, and to aerospace applications. Evaluation of the intermetallics as refractory coatings, reactor hardware, fuel elements, turbine buckets, and high-temperature bearings has been carried out.

Wallace W. Beaver; W. Dean Trautman

Bibliography. G. E. Darwin and J. H. Buddery, *Beryllium*, 1960; H. H. Hausner, *Beryllium: Its Metallurgy and Properties*, 1965; O. Kubaschewski (ed.), *Beryllium: Physiochemical Properties of Its Compounds and Alloys*, 1973; H. Okamoto and L. E. Tanner (eds.), *Phase Diagrams of Binary Beryllium Alloys*, 1987, reprint 1990; D. Webster et al. (eds.), *Beryllium Science and Technology*, 2 vols., 1979; C. G. Wilber, *Beryllium: A Potential Contaminant*, 1980.

Beryllium metallurgy

Production of beryllium and its compounds uses 4000–10,000 tons (3600–9000 metric tons) per year of the mineral beryl. The variant amounts of raw materials reflect wide differences in production rates for the beryllium industry. Beryl, a beryllium-aluminum silicate containing (in commercial grades) about 11–13% beryllium oxide (BeO), 17–19% aluminum oxide (Al_2O_3), 64–70% silicon dioxide (SiO_2), and 1–2% alkali metal oxides, is the only mineral used as a raw material for beryllium. Although beryl ore is widely dispersed throughout the world, it is not concentrated in deposits sufficient to justify direct mining. *See* BERYL.

Large crystals of beryl found in pegmatites are a primary source of raw material; beryl is also found as a by-product of the mining of other minerals, such as mica, feldspar, and spodumene. Although concentration processes involving flotation have been tried, none has provided a means of producing quantities of commercial importance. The major producers of beryl ore are Mozambique and Brazil. *See* PEGMATITE.

Little beryl is produced in the United States, although large reserves of low-grade ores appear to be available. The discovery in Utah of the mineral bertrandite dispersed in clay deposits provided a source of domestic raw material. Bertrandite, $4BeO \cdot 2SiO_2 \cdot H_2O$, is dispersed in montmorillonite clays containing less than 0.5% Be. Processes have been developed to recover the beryllium content as BeO of a quality comparable to that resulting from beryl.

Two processes are used to extract beryllium oxide or hydroxide from beryl ore. There are also two methods employed to reduce BeO to beryllium metal. The two extraction methods are based on dissolving beryl as either a fluoride or a sulfate. Reduction is accomplished thermally by means of magnesium with beryllium fluoride, and electrolytically with beryllium chloride. Over 90% of the metal is made from the thermal process. *See* PYROMETALLURGY, NONFERROUS.

Extraction of oxide or hydroxide. In fluoride extraction finely powdered beryl is mixed with sodium fluoferrate (Na_3FeF_6), briquetted, and heated to 1380°F (750°C). The chemical reaction is given below.

$$3BeO \cdot Al_2O_3 \cdot 6SiO_2 + 2Na_3FeF_6 \rightarrow 6SiO_2 + Fe_2O_3 + Al_2O_3 + 3Na_2BeF_4$$

Water-soluble sodium beryllium fluoride (Na_2BeF_4) is removed from the insoluble oxides by filtration after leaching. The filtrate containing Na_2BeF_4 is treated with caustic soda or ammonium hydroxide to precipitate beryllium hydroxide. The resulting hydroxide is decomposed to BeO by firing at 1470°F (800°C). After removal of the beryllium hydroxide

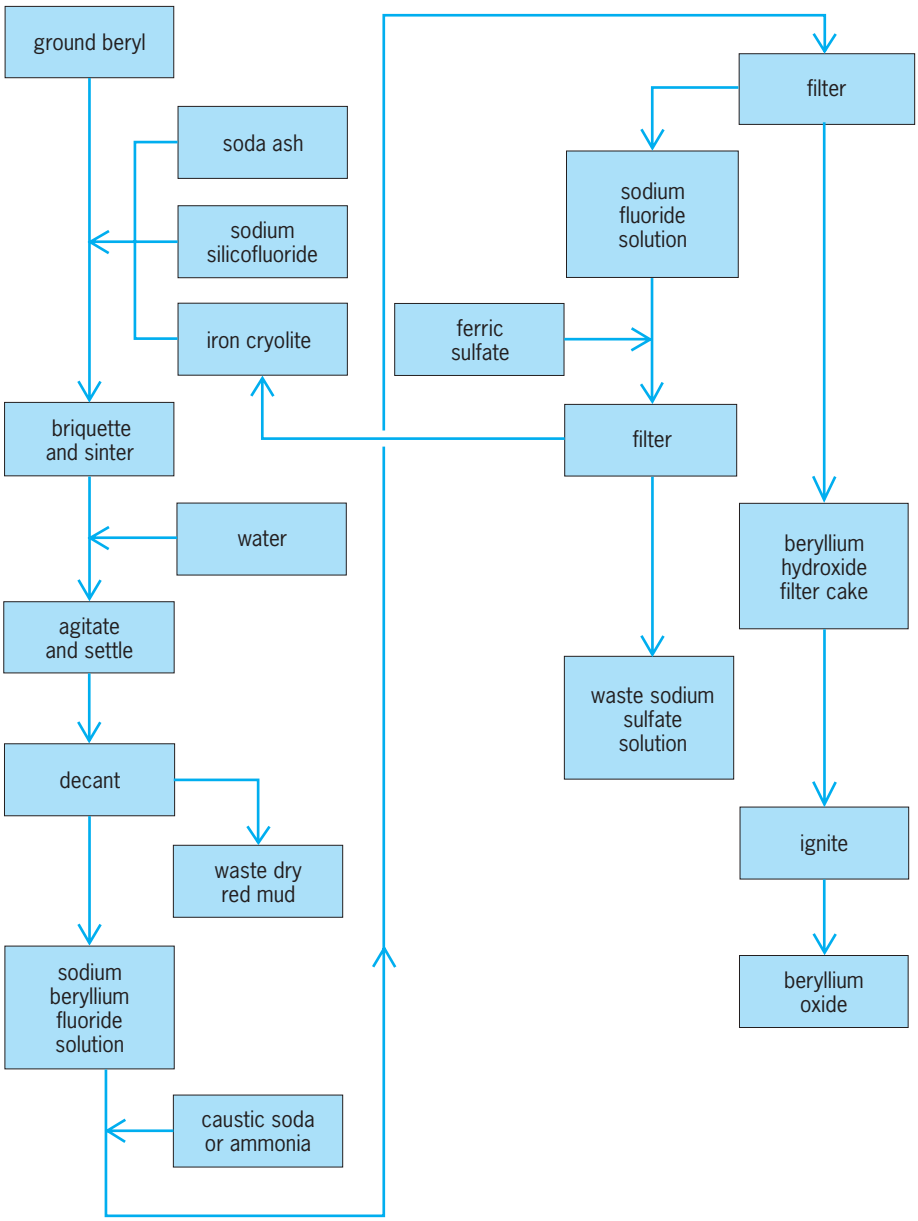

Fig. 1. Flow sheet for the extraction of BeO from beryl by means of the fluoride process. (*Beryllium Corp.*)

[$Be(OH)_2$], the dilute filtrate containing sodium fluoride (NaF) is treated with ferric sulfate [$Fe_2(SO_4)_3$] to produce Na_3FeF_6 for further beryl extraction (**Fig. 1**).

In sulfate extraction beryl ore is melted at 2957°F (1625°C) and quenched in water to destroy the original crystal structure, enhancing reactivity with sulfuric acid. This beryl glass is heat-treated further to precipitate the BeO out of solid solution and thus to increase leaching yield. The heat-treated frit is then ground and treated with sulfuric acid at 390–570°F (200–300°C). The soluble beryllium and aluminum sulfates are leached out with water and removed from the insoluble oxides. Aluminum is separated by converting it to ammonium aluminum sulfate through addition of NH_4OH and crystallizing. Chelating compounds are added to hold iron and nickel in solution. Sufficient sodium hydroxide is then added to convert the beryllium to sodium beryllate, which is hydrolyzed to precipitate beryllium hydroxide.

The use of solvent extraction for the purification of beryllium-containing solutions that result from solubilizing beryl and other beryllium minerals is also used. Liquid-liquid extraction is especially attractive for purifying highly impure liquors with beryllium contents of 0.13 oz/gal (0.2–0.4 g/liter). When this process is used with beryl, the steps between Be-Al sulfate solution and hydrolysis are replaced (**Fig. 2**). *See* EXTRACTION; HYDROMETALLURGY; LEACHING.

Beryllium oxide. In the manufacture of ceramic-grade beryllium oxide, impure beryllium hydroxide is generally converted to either the sufate, nitrate, oxalate, basic carbonate, or basic acetate and is purified by recrystallization. The purified salt is then decomposed thermally to yield beryllium oxide. The sinterability of the resulting oxide is chiefly a function of temperature, which should be sufficiently high

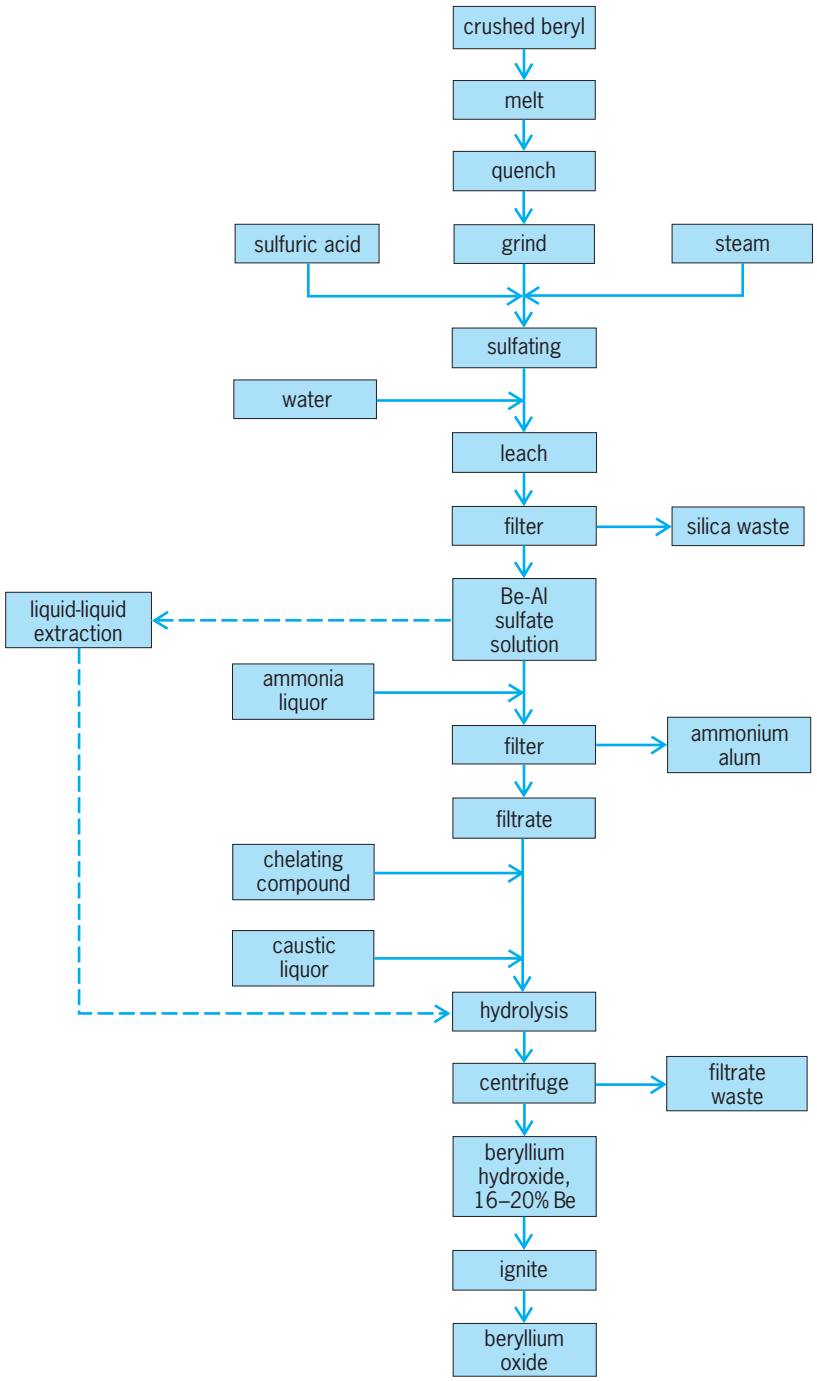

Fig. 2. Flow sheet for the extraction of beryllium oxide, BeO, from beryl by the sulfate process. (*Brush Beryllium Co.*)

to complete decomposition, but also sufficiently low to produce active surfaces. Generally, calcining at 1650–1830°F (900–100°C) gives sinterable-grade oxide, whereas higher temperatures produce a more refractory oxide. Beryllium oxide is primarily employed as a refractory for the melting of reactive metals, for high-temperature nuclear applications, and as a thermally conducting, electrically resisting ceramic.

Beryllium metal. In thermal reduction beryllium hydroxide or oxide is treated with ammonium bifluoride to form $(NH_4)_2BeF_4$. This solution is purified by adding lead dioxide to remove manganese and chromium, calcium carbonate to precipitate aluminum, and after removal of these impurities, ammonium polysulfide to precipitate remaining heavy elements. The purified solution is evaporated and crystallized to salt, which is decomposed to NH_4F and BeF_2 at temperatures of 1650–2000°F (900–1100°C).

Beryllium fluoride is mixed with an amount of magnesium corresponding to about 75% of that stoichiometrically required, and is reduced at 1650°F (900°C) in a graphite crucible employing a high-frequency electric furnace. After the reaction is complete, the temperature is raised to about 2550°F (1400°C) to allow the beryllium to collect. Excess BeF_2 is leached from the slag and the beryllium pebble and MgF_2 are separated. The resulting pebbles, containing slag and some magnesium, are vacuum-melted for purification in a BeO crucible. The metal is poured into graphite molds to produce castings of 120 lb (54.5 kg) each.

For electrolysis beryllium chloride can be made by passing chlorine gas over BeO and carbon at 1830°F (1000°C). $BeCl_2$ is collected as a sublimate in a condenser maintained below 750°F (400°C). The electrolytic cell may consist of a nickel or stainless-steel pot serving as a cathode and of a graphite electrode serving as an anode. A cover is provided to control atmospheric conditions in the pot.

The electrolyte consists of beryllium chloride mixed with sodium chloride. Electrolysis is carried out at about 700°F (370°C). Beryllium flakes, resulting from electrolysis, are leached in water and dried in alcohol. As in the case of pebbles thermally reduced by magnesium from BeF_2, the electrolytic flake is generally vacuum-cast to remove chlorides. *See* ELECTROLYSIS.

Fabrication. Vacuum-cast beryllium ingots either are produced as raw material for the powder-metal process, which accounts for over 90% of the material being used, or are cast directly for fabrication. In the latter process square billets are usually cast by bottom pouring, using a chill plate at the bottom of the mold and a hot top for directional cooling. This process produces sound ingots.

For powder metallurgy round vacuum-cast ingots are converted into chips, or swarf, by multiple turning and cutting operations. The chips are attritioned between beryllium plates in a controlled atmosphere. Powders from 200 mesh down to 15 micrometers maximum particle size are produced for various applications. For finer powders the attritioned particles are comminuted by ball milling.

The major portion of the powdered metal is hot-pressed at 100–2000 lb/in.^{2} (0.69–13.8 megapascals) in steel or graphite dies under vacuum at temperatures of 1920–2100°F (1050–1150°C). Sintering of slip-cast, isopressed, or axially pressed powders can be accomplished at 2057–2190°F (1125–1200°C), and sintering of loose powder employs temperatures up to 2255°F (1235°C). Porous bodies are made by sintering after mixing with a blowing agent for low densities; by temperature, time, and particle-size

control for high densities; and with fugitive binders for intermediate densities.

The sintered product is characterized by thin films of BeO of varying amounts defining the grain boundaries and by fine grains (5–20 micrometer average) with little tendency for gain growth at elevated temperatures. Ingot beryllium has a low oxide content, but tends to have coarser grain structure (greater than 50 μm) with high grain-growth potential after working. Consequently, machined parts are usually made from powder-metal products, while ingot beryllium is usually rolled or otherwise worked.

Rolling, forging, extrusion, and drawing are performed between the temperature of maximum ductility (750°F or 400°C) and the average recrystallization temperature (1470°F or 800°C). The bulk of rolling, forging, and extrusion is carried out at 1110–1470°F (600–800°C), while drawing is carried out at the lower end of the fabrication range. Because of the lower impurity level in ingot, greater reduction or lower temperatures of fabrication are possible than with the finer-grained sintered metal. Usually, fabrication processes are carried out by using protective lubricants or metal covers to prevent damage to the beryllium or work tooling surfaces. Hot rolling may be carried out in steel frames and jackets for support. Above 1470°F (800°C) it is mandatory to fully jacket to provide support and prevent oxidation; only a few extrusion operations use temperatures above 1470°F (800°C). Beryllium sheet can be formed, shaped, spun, and punched. Adequate surface protection and slow forming speeds are advantageous.

Beryllium can be welded, soldered, brazed, or plastically bonded, with brazing and plastic bonding, as well as mechanical fasteners, being the usual production methods. Surface protection of finished beryllium surfaces can be provided by anodizing, plating, optical polishing, or using conversion coatings. Chemical machining and chemical milling are used to provide patterned and low-damage surfaces.

Physical metallurgy. Extensive research has been conducted on the close-packed hexagonal beryllium crystal. Failure takes place by basal plane cleavage at relatively low stresses, but it can be increased with solid solution alloying elements. The critical resolved shear stress in beryllium crystals is largely a function of purity in an inverse relation to the glide strain slip. Prism slip occurring at room temperature is much greater than basal slip. When basal slip is increased by extensive purification, prism slip is increased even more, promoting severe crystal anisotropy even with extensive basal and prism slip available. *See* BERYLLIUM; BERYLLIUM ALLOYS.

In polycrystalline aggregates the plasticity is low at temperatures up to about 480°F (250°C) at which other slip planes become active. The metal becomes quite plastic up to temperatures where grain boundary failure becomes predominant, this type of failure defining the elevated temperature plastic working range. Strength is increased by smaller grain sizes and the presence of solid solution and dispersion hardening elements. Ductility is favored by preferred crystal alignment and finer grains. On working polycrystalline beryllium, severe structural anisotropy occurs, providing little plasticity in the direction of sheet thickness, but as much as 30% elongation in the plane of the sheet. Beryllium has poor notch sensitivity and tends to fracture under high rates of loading.

For design purposes beryllium has excellent compressive and shear properties with good tensile strength, especially at elevated temperatures. Its modulus of 42,000,000 lb/in.^{2} (290 gigapascals) in tension gives the highest stiffness-to-weight ratio of any metallic element. Fatigue properties are high. *See* METAL, MECHANICAL PROPERTIES OF.

Applications. Because of its ability to reflect, moderate, and react (n, α; n, $2n$ reactions) with neutrons, beryllium has had a long history of use in atomic energy as neutron sources, reflectors, and moderators in thermal and intermediate reactors. About all high-power engineering and test reactors utilize beryllium, as do package airborne power systems, such as SNAP and other thermionic-type reactors, and direct thermal systems, such as Phoebus and Nerva. Much auxiliary nuclear equipment also is made of beryllium, including photoneutron sensitizers, monochromatic neutron spectroscopes, and supporting hardware. A number of experimental power systems (France, United Kingdom, Germany, and United States) have employed beryllium as a fuel jacket cover.

Application of beryllium to aerospace usually provides weight saving up to 60% of that of a competing material on the same design basis. Other applications of beryllium are based on properties such as thermal stability (infrared mirrors), creep resistance and thermal capacity (Mercury heat shields and Mercury and Gemini cabin sections), and thermal conductivity (C5-A and other heavy aircraft brakes). *See* ATMOSPHERIC ENTRY.

Extensive use of beryllium in reentry structures (Ranger, Mariner, Polaris, and Poseidon) depends on high-temperature strength and thermal capacity. The metal's transparency to radiation (as used in x-ray windows) also is important to missile structures, which must withstand electromagnetic pulse conditions created by antimissile tactics. Finally, ability to machine to high tolerance coupled with dimensional stability has created almost exclusive employment of beryllium in inertial guidance navigation and control devices. *See* ACOUSTIC RADIOMETER; ELECTROMAGNETIC RADIATION.

Inasmuch as beryllium powder has one of the highest heats of reaction of any element, it may be mixed with oxidizers and binders in liquid, solid, and hybrid types of high-energy propellants. Solid-propellant motors containing up to 11,000 lb (5000 kg) of fuel mixtures of beryllium have been made.

Experimental evaluation of beryllium metal has been conducted for jet engine compressor blades, disks, shafts, and related hardware; actuators, tube supports, and fittings for space vehicles; vertical and horizontal stabilizers; skins for elevator control surfaces and trailing edges; and internal support hardware for aircraft. Applications are in high-speed

process machinery, for stiffening other metals or plastics in composites or as honeycomb, and as a cryogenic conductor for transmission of electrical energy. *See* SUPERCONDUCTIVITY.

Wallace W. Beaver; W. Dean Trautman

Beryllium minerals

Minerals containing beryllium as an essential component. Over 50 beryllium minerals have been identified, even though beryllium is a scarce element in the Earth's crust. The unusual combination of low charge (+2) and small ionic radius (0.035 nanometer) of the beryllium ion accounts for this diverse group of minerals and their occurrence in many natural environments. *See* BERYLLIUM.

Principal minerals. The **table** lists many of the geologically and mineralogically significant beryllium minerals. Nearly all beryllium minerals can be included in one of three groups: compositionally simple oxides and silicates with or without aluminum; sodium- and calcium-bearing silicates; and phosphates and borates. The first group is by far the most abundant; it contains beryl, the most common beryllium mineral, plus the common minerals phenakite, bertrandite, chrysoberyl, and euclase. Of this group, only beryl shows a wide compositional variation; it can have percent quantities of the alkali elements, volatile components, and transition metals. The other minerals, although widely distributed, rarely constitute more than a fraction of a percent of any rock. *See* BERYL.

Some significant beryllium minerals

Name	Formula*
Barylite	$Ba[Be_2Si_2O_7]$
Behoite	$Be(OH)_2$
Bertrandite	$Be_4Si_2O_7(OH)_2$
Beryl	$Be_3Al_2Si_6O_{18}\cdot nH_2O$
Beryllonite	$Na[BePO_4]$
Bityite	$CaLiAl_2[AlBeSi_2O_{10}](OH)_2$
Bromellite	BeO
Chkalovite	$Na_2[BeSi_2O_6]$
Chrysoberyl	$BeAl_2O_4$
Euclase	$BeAlSiO_4(OH)$
Eudidymite	$Na[BeSi_3O_7](OH)$
Gadolinite	$Y_2FeBe_2Si_2O_{10}$
Gugiaite	$Ca_2BeSi_2O_7$
Hambergite	$Be_2BO_3(OH)$
Helvite group	$(Mn,Fe,Zn)_4[BeSiO_4]_3S$
Herderite	$CaBe(PO_4)(F,OH)$
Hurlbutite	$Ca[BePO_4]_2$
Phenakite	Be_2SiO_4

* Brackets indicate framework component of formula.

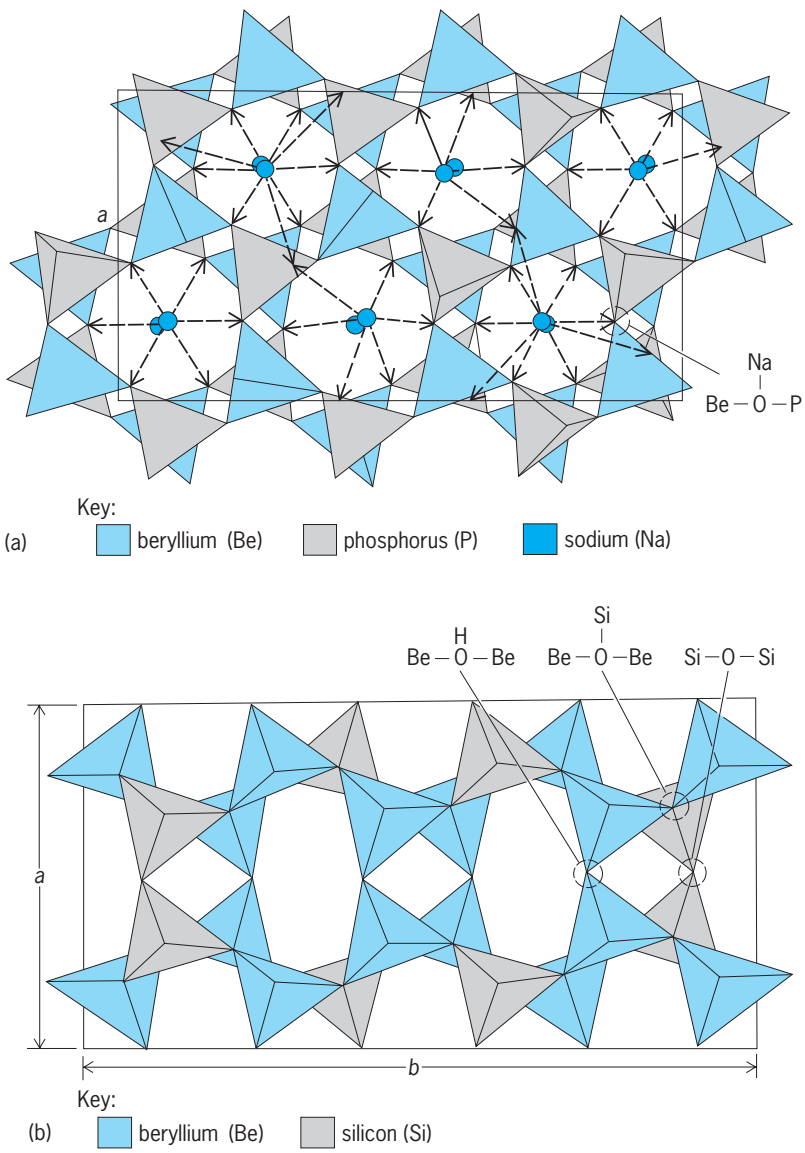

Arrangements of tetrahedra in beryllium minerals. (*a*) Beryllonite; oxygen shared by two tetrahedra. (*b*) Bertrandite; oxygen shared by three tetrahedra. (*After M. Ross, Crystal chemistry of beryllium, USGS Prof. Pap. 468, 1964*)

Crystal structures. The beryllium minerals have many structural characteristics similar to the major rock-forming silicate minerals, but are distinguished by containing large quantities of tetrahedrally coordinated beryllium ion (Be^{2+}) in place of, or in addition to, tetrahedrally coordinated aluminum ion (Al^{3+}) and silicon ion (Si^{4+}). In order to maintain electrical neutrality in the crystal structures, the lower charge on the beryllium ion requires either simultaneous substitution of other elements with higher charge or modification of the structural arrangement. Nearly all beryllium minerals have extensive or complete sharing of corners of beryllium-oxygen (BeO_4^{6-}) tetrahedra with other BeO_4^{6-}, and/or SiO_4^{4-}, AlO_4^{5-}, or phosphorus-oxygen (PO_4^{3-}) tetrahedra to form continuous two- or three-dimensional sheet or framework structures. In many minerals, the apical oxygen ions are shared by only two tetrahedra, in which case the excess charge on the oxygen is balanced by nontetrahedral cations (see **illus.**). These structures can closely resemble those of common silicates (for example, beryl = cordierite; beryllonite = nepheline; helvite = sodalite; bityite = margarite; gugiaite = melilite; hurlbutite = feldspar). In other minerals (for example, phenakite, bertrandite, and euclase), the apical oxygen ions are shared by two beryllium and one silicon tetrahedra, leading to local charge balance and also leading to structural linkages

distinct from the common silicates (see illus.). A few beryllium minerals do not fall in these categories but have other structures, including chrysoberyl (olivine structure) and hambergite. *See* CHRYSOBERYL; SILICATE MINERALS.

Stabilities. Only minerals in the BeO-Al_2O_3-SiO_2-H_2O system have well-determined stability relationships. They are typified by the progressive replacement of hydrous phases by anhydrous equivalents with increasing temperature. Bertrandite dehydrates to phenakite at 460–520°F (240–260°C). Euclase is stable to 750–1050°F (400–600°C), depending on the water pressure. Beryl has a wide stability range, which is enhanced by the incorporation of alkalies and water into its structure. Chrysoberyl and phenakite are stable to more than 2500°F (1400°C) and pressures of more than 20 kilobars (2 gigapascals). Equilibria in more complex systems are conveniently cast in terms of the activities (thermodynamic concentrations) of silica (SiO_2) and alumina (Al_2O_3). In so doing, it is seen that the minerals order bromelite < chrysoberyl < euclase < bertrandite = phenakite = beryl for compatibility with silica, and bromellite < bertrandite = phenakite < beryl < euclase < chrysoberyl for compatibility with alumina. In natural environments, the more abundant, rock-forming minerals govern the activities of silica and alumina and thus which beryllium mineral is stable.

The relative stabilities of other beryllium minerals can be inferred from their natural occurrences; no experimental studies have been done. The phosphates and borates (for example, beryllonite, hambergite, and herderite) are stable only at geologically high activities of phosphate and borate. The aluminum-free sodium- and calcium-bearing silicates (for example, chkalovite, eudidymite, and gugiaite) are stable only with high alkali activities and low alumina activities. The helvite group requires moderately, but not strongly, sulfidizing conditions.

Geological occurrence. Beryllium minerals occur in many geological environments, where they are generally associated with felsic (abundant feldspar ± quartz) igneous rocks and related, metasomatically altered rocks. Minerals from the BeO-Al_2O_3-SiO_2-H_2O system occur in virtually all environments where separate beryllium phases exist. The minerals bearing sodium (Na), calcium (Ca), sulfur (S), phosphorus (P), boron (B), manganese (Mn), and fluorine (F) have more restricted occurrences, reflecting the special conditions required for their stabilization. Quartz-bearing (silica-saturated) settings include the feldspar- and mica-bearing environments typical of granites, pegmatites and many ore deposits, and the alkali-deficient vein and greisen environments. Beryl, phenakite, and bertrandite occur in all settings; restricted to particular types are euclase (alkali-deficient settings, because it is incompatible with feldspar), chrysoberyl (high-temperature only, because it is incompatible with quartz at low temperature), and the phosphates and helvite-group minerals (in pegmatites and ore deposits, where phosphorus and sulfur are concentrated by magmatic or hydrothermal processes). Quartz-absent (silica-undersaturated) environments include skarns (silicate replacement of carbonate rocks), desilicated pegmatites, and some low-temperature volcanogenic deposits. Chrysoberyl, phenakite, and bertrandite are typical of the silica-undersaturated environments; less commonly beryl, bromellite, behoite, and helvite-group minerals occur. Alkali-excess igneous environments, including many syenites, commonly contain phenakite, bertrandite, and the alkali-beryllium silicates (for example, chkalovite and eudidymite).

Uses. Beryl and bertrandite, mined from granitic pegmatites and altered volcanic rocks, are the principal ores of beryllium; deposits of chrysoberyl and phenakite may become economically significant in the future. The colored varieties of beryl (emerald, aquamarine, morganite) are valued gemstones; chrysoberyl, phenakite, and a few of the other minerals are less common gemstones. Synthetic bromellite has specialized applications in the ceramic and nuclear industries. *See* EMERALD; GEM. Mark D. Barton

Bibliography. M. D. Barton, Phase equilibria and thermodynamic properties of minerals in the BeO-Al_2O_3-SiO_2-H_2O system, with petrologic applications, *Amer. Mineralog.*, 71:277–300, 1986; P. Cerny (ed.), *Granitic Pegmatites in Science and Industry*, Mineralogical Association of Canada Short Course Notes, pp. 329–346, 1982; M. Ross, *Crystal Chemistry of Beryllium*, USGS Prof. Pap. 468, 1964.

Bessel functions

By definition the solutions of Bessel's differential equation, Eq. (1). Bessel functions, also called cylin-

$$z^2 d^2y/dz^2 + z\,dy/dz + (z^2 - \nu^2)y = 0 \qquad (1)$$

der functions, are examples of special functions which are introduced by a differential equation. Bessel functions are of great interest in purely mathematical concepts and in mathematical physics. They constitute additional functions which, like the elementary functions z^n, sin z, or e^z, can be used to express physical phenomena.

Applications of Bessel functions are found in such representative problems as heat conduction or diffusion in circular cylinders, oscillatory motion of a sphere in a viscous fluid, oscillations of a stretched circular membrane, diffraction of waves by a circular cylinder of infinite length or by a sphere, acoustic or electromagnetic oscillations in a circular cylinder of finite length or in a sphere, electromagnetic wave propagation in the waveguides of circular cross section, in coaxial cables, or along straight wires, and in skin effect in conducting wires of circular cross section. In these problems Bessel functions are used to represent such quantities as the temperature, the concentration, the displacements, the electric and magnetic field strengths, and the current density as function of space coordinates. The Bessel functions enter into all these problems because boundary values on circles (two-dimensional problems), on

circular cylinders, or on spheres are prescribed, and the solutions of the respective problems are sought either inside or outside the boundary, or both.

Definition of Bessel functions. The independent variable z in Bessel's differential equation may in applications assume real or complex values. The parameter ν is, in general, also complex. Its value is called the order of the Bessel function. Since there are two linearly independent solutions of a linear differential equation of second order, there are two independent solutions of Bessel's differential equation. They cannot be expressed in finite form in terms of elementary functions such as z^n, $\sin z$, or e^z unless the parameter is one-half of an odd integer. They can, however, be expressed as power series with an exception for integer values of ν. The function defined by Eq. (2)

$$J_\nu(z) = \left(\frac{z}{2}\right)^\nu \sum_{l=0}^{\infty} \frac{(-z^2/4)^l}{l!\Gamma(\nu + l + 1)} \tag{2}$$

is designated as Bessel's function of the first kind, or simply the Bessel function of order ν. $\Gamma(\nu + l + 1)$ is the gamma function. The infinite series in Eq. (2) converges absolutely for all finite values, real or complex, of z. In particular, Eqs. (3) may be expressed. Along

$$\begin{aligned} J_0(z) &= 1 - \frac{1}{1!1!}\left(\frac{z}{2}\right)^2 + \frac{1}{2!2!}\left(\frac{z}{2}\right)^4 \\ &\qquad - \frac{1}{3!3!}\left(\frac{z}{2}\right)^6 + \cdots \\ J_1(z) &= \frac{z}{2} - \frac{1}{1!2!}\left(\frac{z}{2}\right)^3 + \frac{1}{2!3!}\left(\frac{z}{2}\right)^5 - \cdots \end{aligned} \tag{3}$$

with $J_\nu(z)$, there is a second $J_{-\nu}(z)$. It is linearly independent of $J_\nu(z)$ unless ν is an integer of n. In this case $J_{-n}(z) = (-1)^n J_n(z)$. *See* GAMMA FUNCTION.

The Bessel function of the second kind, also called Neumann function, is defined by Eq. (4).

$$Y_\nu(z) = \frac{\cos \nu\pi J_\nu(z) - J_{-\nu}(z)}{\sin \nu\pi} \tag{4}$$

If $\nu = n$, this expression is indeterminate, and the limit of the right member is to be taken. In particular, Eq. (5) may be expressed, with $\ln \gamma = 0.577215\ldots$

$$\begin{aligned} Y_0(z) &= \frac{2}{\pi} J_0(z) \ln\left(\frac{1}{2}\gamma z\right) \\ &\quad + \frac{2}{\pi}\left[\frac{1}{1!1!}\left(\frac{z}{2}\right)^2 = -\frac{1}{2!2!}\left(1 + \frac{1}{2}\right)\left(\frac{z}{2}\right)^4\right. \\ &\qquad \left. + \frac{1}{3!3!}\left(1 + \frac{1}{2} + \frac{1}{3}\right)\left(\frac{z}{2}\right)^6 - \cdots\right] \end{aligned} \tag{5}$$

(Euler's constant). There are two Bessel functions of the third kind, designated as first and second Hankel functions. They are defined as Eqs. (6).

$$\begin{aligned} H_\nu^{(1)}(z) &= J_\nu(z) + iY_\nu(z) \\ H_\nu^{(2)}(z) &= J_\nu(z) - iY_\nu(z) \end{aligned} \tag{6}$$

Elementary properties. For $z = 0$, then $J_0(0) = 1$, $J_n(0) = 0$ if $n = 1, 2, \ldots$, while $Y_n^{(1)}$ and therefore also $H_n^{(1)}$ and $H_n^{(1)}$ infinite for all n. The behavior of Bessel functions for small values of z is readily seen from the power series expansion in Eq. (2), if ν is not an integer. For integer $\nu = n = 0$, one has for small z the relations in notation (7) (the sign $\approx$

$$\begin{aligned} J_n(z) &\approx \left(\frac{z}{2}\right)^n \Big/ n! \\ Y_0(z) &\approx \frac{2}{\pi} \ln \frac{\gamma z}{2} \\ Y_n(z) &\approx -\frac{1}{\pi}\left(\frac{z}{2}\right)^{-n} (n-1)! \\ &\quad (n = 1, 2, \ldots) \end{aligned} \tag{7}$$

means approximately equal). Here $n! = 1 \cdot 2 \cdot 3 \ldots n$, $0! = 1$. The behavior of the functions for large values of $z(|z|\Gamma|\nu|^2)$ is given for all values of ν by the formulas in notation (8). They hold not only for real

$$\begin{aligned} J_\nu(z) &\approx \left(\frac{2}{\pi z}\right)^{1/2} \cos\left(z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) \\ Y_\nu(z) &\approx \left(\frac{2}{\pi z}\right)^{1/2} \sin\left(z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) \\ H_\nu^{(1)}(z) &\approx \left(\frac{2}{\pi z}\right)^{1/2} \exp\left[i\left(z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right)\right] \\ H_\nu^{(2)}(z) &\approx \left(\frac{2}{\pi z}\right)^{1/2} \exp\left[-i\left(z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right)\right] \end{aligned} \tag{8}$$

positive values of z but also for complex values in the angular domain $-\pi < \arg z < \pi$. These formulas are said to be asymptotic. This means that the difference of the left and right member, multiplied by a certain power of z, which is here $z^{3/2}$, is bounded as z goes to infinity along a straight line in the angular domain mentioned. The asymptotic formulas can be used to obtain approximate numerical values of the functions for near large $|z|$. They reveal that the Bessel functions are closely related to the functions $\cos z$, $\sin z$, $\exp(iz)$, $\exp(-iz)$, respectively.

The general behavior of Bessel functions can to a certain extent be inferred from the behavior near $z = 0$ and from the asymptotic behavior. The functions $J_\nu(z)$ and $Y_\nu(z)$ of real positive argument z and real order ν are oscillatory in z with a period that is not quite a constant but approaches 2π as $z \to \infty$. The oscillations have, however, an amplitude which decreases in accordance with $(2/\pi z)^{1/2}$ as $z \to \infty$. They are used in applications to represent standing cylindrical and spherical waves. The behavior of Hankel functions follows then from Eqs. (6). They are used to represent progressing cylindrical and spherical waves.

Figures 1 and **2** give the Bessel and Neumann functions for $\nu = 0, 1, 2$ and positive values of z.

The large zeros $z_{\nu,s}$ of the function $J_\nu(z) \cos \alpha + Y_\nu(z) \sin \alpha$ with α a constant are given by the asymptotic expression $z_{\nu,s}^2 \approx [(s + \nu/2 - 1/4)\pi + \alpha]^2 + 1/4 - \nu^2$, in which s assumes large integer values. Even the first positive zero of $J_0(z)$ is very accurately given by this formula, namely, $z_{0,1} \approx 2.409$, as compared with the accurate value $z_{0,1} = 2.4048\ldots$.

There exist two relations between Bessel functions whose orders differ by one. If $C_\nu(z)$ stands for any one of the four kinds of functions $J_\nu(z)$, $Y_\nu(z)$, $H_\nu^{(1)}(z)$, $H_\nu^{(1)}(z)$, then Eqs. (9) may be written, and

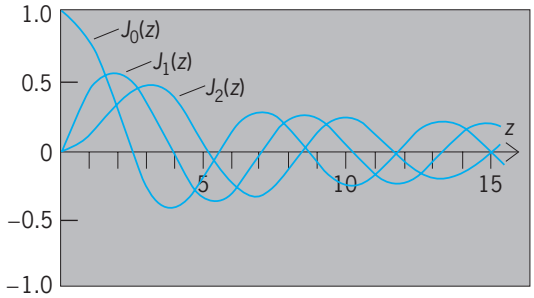

Fig. 1. Bessel functions $J_0(z)$, $J_1(z)$, $J_2(z)$ for $\nu = 0, 1, 2$ and positive values of z.

hence Eqs. (10) hold.

$$\begin{aligned}(2\nu/z)C_\nu(z) &= C_{\nu-1}(z) + C_{\nu+1}(z)\\ 2dC_\nu(z)/dz &= C_{\nu-1}(z) - C_{\nu+1}(z)\end{aligned} \tag{9}$$

$$\begin{aligned}z^\nu C_{\nu-1}(z) &= d[z^\nu C_\nu(z)]/dz\\ z^{-\nu}C_{\nu+1}(z) &= -d[z^{-\nu}C_\nu(z)]/dz\end{aligned} \tag{10}$$

Further properties. Some differential equations can be solved in terms of Bessel functions. The differential equation (11) has the solutions $y = z^{\beta\nu-\alpha}C_\nu(\gamma z^\beta)$, while $d^2y/dz^2 + (e^{2z} - \nu^2)y = 0$

$$z^2\frac{d^2y}{dz^2} + (2\alpha - 2\nu\beta + 1)z\frac{dy}{dz} + [\beta^2\gamma^2z^{2\beta} + \alpha(\alpha - 2\nu\beta)]y = 0 \tag{11}$$

has the solutions $y = C\nu(e^2)$. There are series with terms containing Bessel functions, whose sum can be expressed by elementary functions, for example, Eqs. (12). In the last formula the functions $P_n(u)$ are

$$\begin{aligned}\exp(iz\cos\theta) &= J_0(z) + 2\sum_{n=1}^{\infty} i^n J_n(z)\cos n\theta\\ 1 &= J_0^2(z) + 2\sum_{n=1}^{\infty} J_n{}^2(z)\\ e^{iuz} &= (\pi/2z)^{1/2}\sum_{n=0}^{\infty} i^n(2n+1)J_{n+1/2}(z)P_n(u)\end{aligned} \tag{12}$$

the Legendre polynomials. There are also addition theorems of the type in Eq. (13).

$$J_0(x \pm y) = J_0(x)J_0(y) + 2\sum_{n=1}^{\infty}(\mp 1)^n J_n(x)J_n(y) \tag{13}$$

Bessel functions are also useful because many definite integrals can be expressed in terms of Bessel functions. Examples are shown as Eqs. (14). The

$$\pi J_n(z) = i^{-n}\int_0^\pi e^{iz\cos\alpha}\cos n\alpha\, d\alpha \qquad (n = \text{integer})$$

$$\frac{\pi}{2}x^{1/2}\left[J_{1/3}(2x^{3/2}) + J_{-1/3}(2x^{3/2})\right] = \int_0^\infty \cos(t^3 - 3tx)\, dt \tag{14}$$

last integral is called Airy's integral. Definite integrals containing Bessel functions can often be expressed by elementary functions, for example, Eq. (15),

$$\int_0^\infty e^{-t}J_0(at)\, dt = (1 + a^2)^{-1/2} \tag{15}$$

when the imaginary part of a is between -1 and $+1$.

Expansions in terms of Bessel functions. In applications there is often the problem of expanding a function in terms of Bessel functions or expressing it as an integral with Bessel functions in the integrand.

The Fourier-Bessel series is a generalization of the Fourier series. Let $z_1, z_2, z_3, \ldots$ be the positive zeros of $J_\nu(z)$ for some real $\nu = -\frac{1}{2}$ arranged in increasing order of magnitude. Then a continuous function $f(x)$ defined in the interval $0 \leqq x \leqq 1$ can be expanded into the Fourier-Bessel series, Eq. (16), where the coefficients a_m are given by Eq. (17).

$$f(x) = \sum_{m=1}^{\infty} a_m J_\nu(z_m x) \qquad (0 < x < 1) \tag{16}$$

$$\left[J_{\nu+1}(z_m)\right]^2 a_m = \int_0^1 t f(t) J_\nu(z_m t)\, dt \tag{17}$$

The Fourier-Bessel integral is a generalization of the Fourier integral. Let $f(x)$ be a continuous function, such that the integral

$$\int_0^\infty f(x)\, dx$$

exists and is absolutely convergent. Furthermore, let ν be a real number $= -{}^1/_2$. Then the function $f(x)$ can be represented by an integral containing Bessel functions, the Fourier-Bessel integral, Eq. (18), where

$$f(x) = \int_0^\infty t^{1/2}J_\nu(xt)g(t)\, dt \tag{18}$$

$g(t)$, the Fourier-Bessel transform of $f(x)$, is given by Eq. (19). Both representations hold also if $f(x)$ has a

$$g(t) = \int_0^\infty x^{1/2}J_\nu(xt)f(x)\, dx \tag{19}$$

finite number of discontinuities, except at points of discontinuity themselves.

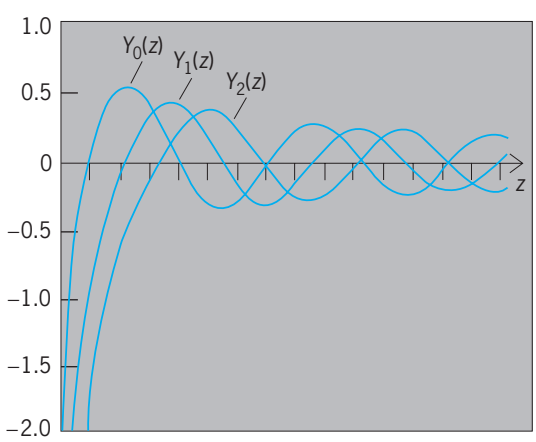

Fig. 2. Neumann functions $Y_0(z)$, $Y_1(z)$, $Y_2(z)$ for $\nu = 0, 1, 2$ and for positive values of z.

Functions related to Bessel functions. In some applications, Bessel functions occur with a fixed value of $\arg z \neq 0$, for instance with $\arg z = \pi/2$, that is, positive imaginary z. For convenience, special notations are used in such cases. For positive x the functions in Eqs. (20) are defined. These functions are called

$$\begin{aligned} I_\nu(x) &= \exp(-i\nu\pi/2)J_\nu(ix) \\ K_\nu(x) &= (i\pi/2)\exp(i\nu\pi/2)H_\nu^{(1)}(ix) \end{aligned} \tag{20}$$

modified Bessel function and modified Hankel function or MacDonald functions of order ν. For real values of $\nu = 0$, the functions $I_\nu(z)$ are monotonically increasing to infinity while the functions $K_\nu(z)$ are monotonically decreasing from infinity to zero as z goes through real values from 0 to ∞.

Particularly in problems for which the use of spherical coordinates is appropriate, there occur Bessel functions with a factor $z^{-1/2}$. They are called spherical Bessel, Neumann, and Hankel functions and defined by Eqs. (21). These functions can be expressed

$$\begin{aligned} \psi_\nu{}^{(1)}(z) &= (\pi/2z)^{1/2}J_{\nu+1/2}(z) \\ \psi_\nu{}^{(2)}(z) &= (\pi/2z)^{1/2}Y_{\nu+1/2}(z) \\ \psi_\nu{}^{(3)}(z) &= (\pi/2z)^{1/2}H^{(1)}{}_{\nu+1/2}(z) \\ \psi_\nu{}^{(4)}(z) &= (\pi/2z)^{1/2}H^{(2)}{}_{\nu+1/2}(z) \end{aligned} \tag{21}$$

by elementary functions in finite form, if ν is an integer. In particular, Eqs. (22) may be written.

$$\begin{aligned} \psi_0{}^{(1)}(z) &= \frac{\sin z}{z} \qquad \psi_{-1}{}^{(1)}(z) = \frac{\cos z}{z} \\ \psi_1{}^{(1)}(z) &= \frac{-\cos z}{z} + \frac{\sin z}{z^2} \\ \psi_{-2}{}^{(1)}(z) &= \frac{-\sin z}{z} - \frac{\cos z}{z^2} \end{aligned} \tag{22}$$

See FOURIER SERIES AND TRANSFORMS. Josef Meixner

Bibliography. M. Abramowitz and I. A. Stegun (eds.), *Handbook of Mathematical Functions*, 10th ed., 1972; L. C. Andrews, *Special Functions of Mathematics for Engineers*, 2d ed., 1992; A. Erdelyi (ed.), *Higher Transcendental Functions*, 3 vols., 1953–1955, reprint 1981; A. Erdelyi (ed.), *Tables of Integral Transforms*, 2 vols., 1954; A. N. Kiforov and V. Uvarov, *Special Functions of Mathematical Physics*, 1988; Z. X. Wang and D. R. Guo, *Special Functions*, 1989; G. N. Watson, *Treatise on the Theory of Bessel Functions*, 2d ed., 1944.

Beta particles

The name first applied in 1897 by Ernest Rutherford to one of the forms of radiation emitted by radioactive nuclei. Beta particles can occur with either negative or positive charge (denoted β^- or β^+) and are now known to be either electrons or positrons, respectively. Electrons and positrons are now referred to as beta particles only if they are known to have originated from nuclear beta decay. Their observed kinetic energies range from zero up to about 5 MeV in the case of naturally occurring radioactive isotopes, but can reach values well over 10 MeV for some artificially produced isotopes. *See* ALPHA PARTICLES; ELECTRON; GAMMA RAYS; POSITRON; RADIOACTIVITY.

Properties. When a nucleus beta-decays, it emits two particles at the same time: One is a beta particle; the other, a neutrino or antineutrino. With this emission, the nucleus itself undergoes a transformation, changing from one element to another. In the case of isotopes that β^+-decay, each decaying nucleus emits a positron and a neutrino, simultaneously reducing its atomic number by one unit; for those isotopes that β^--decay, each nucleus emits an electron and an antineutrino while increasing its atomic number by one. In both classes of decay, the energy released by the nuclear transformation is shared between the two emitted particles. Though the energy released by a particular nuclear transformation is always the same, the fraction of this energy carried away by the beta particle is different for each individual decaying nucleus. (The neutrino always carries away the remainder, thus conserving energy overall.) When observed collectively, the decaying nuclei of a simple radioactive source emit their beta particles with a continuous distribution of kinetic energies covering the range from zero up to the total nuclear decay energy available (**Fig. 1**).

Radioactive samples often contain several radioactive isotopes. Since each isotope has its own decay energy and beta-particle energy distribution, the energy spectrum of beta particles observed from such a sample would be the sum of a number of distributions like the one in Fig. 1, each with a different end-point energy. Indeed, many isotopes, especially those artificially produced with accelerators, can themselves beta-decay by additional paths that also release part of the energy in the form of gamma radiation. In general then, observed beta-particle spectra are more complex than that shown in Fig. 1, but they are always smooth distributions over a range of energy.

Under certain conditions, nuclear decay can result in the emission of monoenergetic electrons. These are called conversion electrons and are atomic

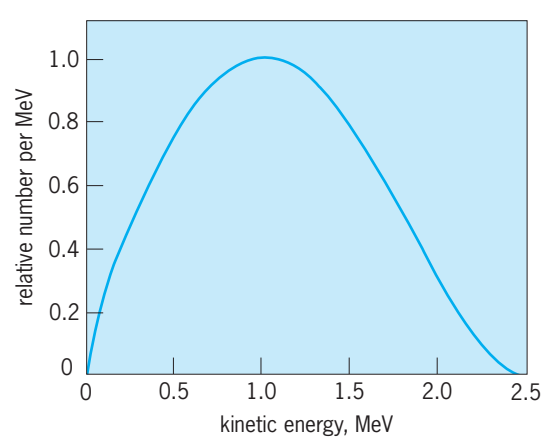

Fig. 1. Spectrum of energies of beta particles emitted from a radioactive source. The distribution shown corresponds to a light nucleus that releases a maximum kinetic energy of 2.5 MeV.

electrons ejected from their orbit by the energy released from a nucleus that does not change its atomic number. In the early 1900s, they were confused with the electrons emitted in beta decay, but their appearance is actually quite independent of the beta-decay process.

Interactions with matter. As a beta particle penetrates matter, it loses its energy in collisions with the constituent atoms. Two processes are involved. First, the beta particle can transfer a small fraction of its energy to the struck atom. Second, the beta particle is deflected from its original path by each collision and, since any change in the velocity of a charged particle leads to the emission of electromagnetic radiation, some of its energy is lost in the form of low-energy x-rays (called, in this context, bremsstrahlung). Though the energy lost by a beta particle in a single collision is very small, many collisions occur as the particle traverses matter, causing it to follow a zigzag path as it slows down. *See* BREMSSTRAHLUNG.

Because of their random zigzag paths, even monoenergetic electrons will not all travel through the same thickness of material before stopping. However, the thickness that is just sufficient to stop all the beta particles of a particular energy is called the range of those particles. For the continuous energy distribution normally associated with a source of beta particles, the effective range is the one that corresponds to the highest energy in the primary spectrum. That thickness of material stops all of the beta particles from the source. The range depends strongly on the electron energy and the density of the absorbing material (**Fig. 2**).

The slowing-down processes have the same effect on both β^- and β^+ particles. However, as antimatter, the positron (β^+) cannot exist for long in the presence of matter. It soon combines with an atomic electron, with which it annihilates, the masses of both particles being replaced by electromagnetic energy. Usually this annihilation occurs after the positron has come to rest and formed a positronium atom, a bound but short-lived positron-electron system. In that case, the electromagnetic energy that is emitted from the annihilation takes the form of two 511-keV gamma rays that are emitted in opposite directions to conserve momentum. *See* POSITRONIUM.

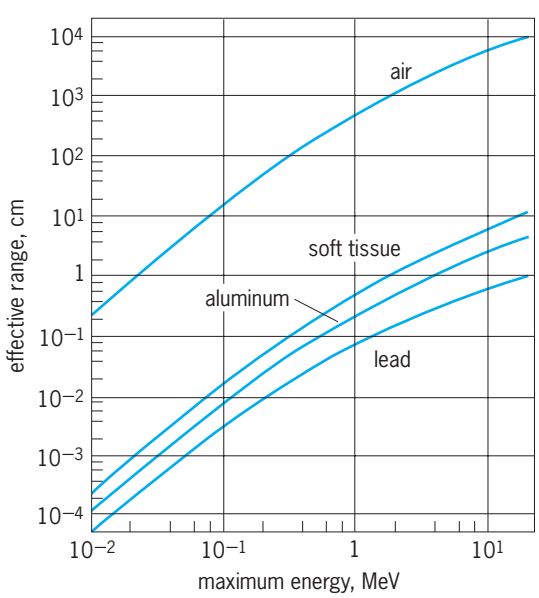

Fig. 2. Ranges of beta particles in representative substances, plotted as a function of their maximum energy. (*After L. Pages et al., Atomic Data, vol. 4, pp. 1–127, 1972*)

Detection. Beta particles are detected through their interaction with matter, and a wide variety of detection techniques are now available. One class of detectors employs gas as the detection medium. Ionization chambers, proportional counters, and Geiger-Müller counters are of this class. In these detectors, after entering through a thin window, the beta particles produce positive ions and free electrons as they collide with atoms of the gas in the process of their slowing down. An electric field applied across the volume of gas causes these ions and electrons to drift along the field lines, causing an ionization current that is then processed in external electronic devices. Depending on the application, the current can be integrated over time or, instead, the output can be recorded pulse by pulse to give information—energy and time of arrival, for example—on each individual beta particle. *See* GEIGER-MÜLLER COUNTER; IONIZATION CHAMBER; PARTICLE DETECTOR.

More precise energy information can be achieved with scintillation detectors. In certain substances, the ion-electron pairs produced by the passage of a charged particle result in the emission of a pulse of visible or near-ultraviolet light. If a clear plastic scintillator is used, it can be mounted on a photomultiplier tube, which converts the transmitted light into a measurable electrical current pulse whose amplitude is proportional to the energy deposited by the incident beta particle. *See* SCINTILLATION COUNTER.

Even better energy information comes from semiconductor detectors, which are effectively solid-state ionization chambers. When a beta particle enters the detector, it causes struck electrons to be raised into the conduction band, leaving holes behind in the valence band. The electrons and holes move under the influence of an imposed electric field, causing a pulse of current to flow. Such detectors are compact and easy to use, but have the disadvantage that they are relatively small and have a lower detection efficiency. They are useful mainly for low-energy beta particles. *See* JUNCTION DETECTOR.

If the ultimate in energy information is required, albeit at the further expense of efficiency, any one of these detectors can be combined with a magnetic spectrometer. Beta particles, like any charged particles, follow curved paths in a perpendicular magnetic field, their radius of curvature being proportional to the square of their energy. Their detected position on exiting the magnetic field can be precisely related to their energy. The best current measurement of the electron antineutrino mass comes from a spectrometer measurement of the tritium beta-decay spectrum. *See* NEUTRINO. John Hardy

Bibliography. K. Heyde, *From Nucleons to the Atomic Nucleus*, 1998; N. A. Jelley, *Fundamentals of Nuclear Physics*, 1990; W. S. C. Williams, *Nuclear and Particle Physics*, 1991.

Betatron

A device for accelerating charged particles in an orbit by means of the electric field E from a slowly changing magnetic flux Φ. The electric field is given by $E = -(^1/_2\pi r_o)\ d\Phi/dt$ (in SI or mks units), where r_o is the orbit radius. The name was chosen because the method was first applied to electrons. In the usual betatron both the accelerating core flux and a guiding magnetic field rise with similar time dependence, with the result that the orbit is circular. However, the orbit can have a changing radius as acceleration progresses. For the long path (usually more than 100 km), variations of axial and radial magnetic field components provide focusing forces, while space charge and space current forces due to the particle beam itself also contribute to the resulting betatron oscillations about the equilibrium orbit. In many other instances of particle beams, the term betatron oscillation is used for the particle oscillations about a beam's path.

Although there had been a number of earlier attempts to make a betatron, the complete transient theory guided the construction of a successful 2.3×10^6 eV accelerator, producing x-rays equivalent to the gamma rays from 2 g of radium, at the University of Illinois in 1940.

It can be shown that the flux change within an orbit of fixed radius must be twice that which would exist if the magnetic field were uniform. The flux can be biased backward by a fixed amount ϕ_0 to avoid iron saturation, as it was in the University of Illinois 320-MeV betatron.

The beam must be focused in both the radial direction and the axial direction (the direction perpendicular to the plane of the orbit). For radial focusing, the centripetal force F_o required to hold a particle in a circular orbit of radius r must be supplied by the magnetic field B exerting a force F_M. At an equilibrium radius r_o, $F_C = F_M$, and if F_M decreases less rapidly than $1/r$ with increasing radius, there will be a net focusing force directed back toward r_o when $r \neq r_o$. This focusing force will cause the radius r to undergo a simple harmonic oscillation about r_o. If the axial magnetic field B_z falls off less rapidly than $1/r$, say $B_z \sim 1/r^n$, then it can be shown that this oscillation has a frequency $\omega_r = \sqrt{1-n}\ \Omega$, where Ω is the angular velocity in the particle's orbit. Thus $n < 1$ is necessary for a stable radial oscillation giving a frequency less than the frequency of rotation.

In addition, axial focusing results from requiring the magnetic field to decrease with radius so that the lines of force are curved, giving a component of F_M on a displaced particle back toward the orbital plane, as shown in the **illustration**. It can be shown that this focusing force gives rise to an axial oscillation with frequency $W_z = \Omega\sqrt{n}$. Thus $0 < n < 1$ is necessary for complete focusing. This is the so-called weak focusing conditions for betatron oscillation in accelerators.

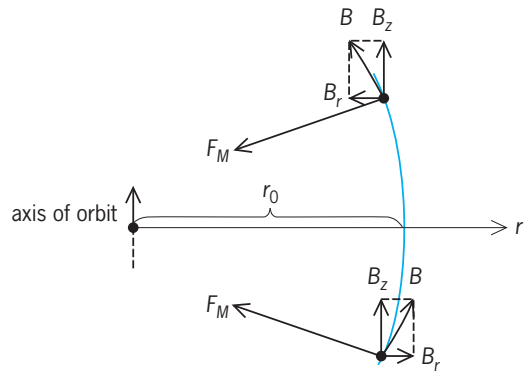

Magnetic fields and their resulting forces in the axial plane of a betatron orbit, which result in axial focusing.

By dividing the focusing structure into sectors of large and alternating gradients, $+n \gg 1$, so-called strong focusing results. This allows a small volume for the beam path, or direct-current focusing magnets with a finite change in orbit radius as the particle gains great energy. Such betatrons have been constricted.

Injection into an orbital oscillation is possible because the restoring force is proportional to B^2, which rises relatively rapidly at injection time. This increasing focusing force causes the oscillation amplitude to decrease as $1/\sqrt{B}$ and enables the particles to escape subsequent collision with a properly designed injection structure.

Collective effects from self-fields of the beam have been found important and helpful in injecting. Circulating currents of about 3 A are contained in the numerous industrial and therapeutic betatrons, although the average currents are below 10^{-7} A. Such beams have been extracted by using a magnetic shunt channel where $n = {}^3/_4$ so that $\omega/\Omega = {}^1/_2$. This field bump excites a half-integral resonance which throws the particles into the channel. *See* PARTICLE ACCELERATOR.

Donald W. Kerst

Betel nut

The dried, ripe seed of the palm tree *Areca catechu* (Palmae), a native of Sri Lanka and Malaya, which attains a height of 30–50 ft (9–15 m). The nuts, slightly larger than a chestnut, when broken open have a faint odor and a somewhat acrid taste. Betel nuts are chewed by the natives together with the leaves of the betel pepper, *Piper betle* (Piperaceae), and lime. The mixture discolors and eventually destroys the teeth. Frequently, habitual chewers of the nut are toothless by the age of 25. The seeds contain a narcotic that produces some stimulation and a sense of well-being. *See* NARCOTIC; PIPERALES.

Perry D. Strausbaugh; Earl L. Core

Betelgeuse

A cool and highly luminous star, prominently located in the right shoulder of the constellation Orion and noticeably red in color. Betelgeuse, or α Orionis, is a supergiant star about 130 parsecs (430 light-years) from the Sun. Its spectral type of M2 indicates an effective temperature of approximately 3500 K (5800°F). This temperature would result in a low overall luminosity were it not for the enormous diameter of the star, about 1100 times that of the Sun. Replacing the Sun, Betelgeuse would fill the solar system to beyond the orbit of Mars. Betelgeuse is a supergiant star with approximately 150,000 times the Sun's luminosity. Betelgeuse varies in brightness with a period of 6.4 years. Its mean density is extremely low, less than one-millionth the density of water, and the low surface gravity of the rarefied outer layers results in a continual loss of matter, presently at the rate of 1 solar mass every 250,000 years. Observations at infrared and radio wavelengths show a complex system of dust and gas shells extending out as far as a thousand stellar radii. If this material were visible to the eye, Betelgeuse would appear as large as Mars through a telescope. These shells appear to originate from occasional outbursts in which material is ejected and eventually merges with the interstellar medium. *See* SPECTRAL TYPE; STELLAR EVOLUTION; SUPERGIANT STAR; VARIABLE STAR.

Because of its enormous size, Betelgeuse is one of the few stars resolvable with special techniques, such as speckle interferometry, at large optical telescopes, where a diameter of about 0.055 arc-second has been measured. These techniques have also produced the first images of the disk of a star and show asymmetric brightness variations over the surface that change with time and are possibly due to large-scale convection of material within the star's atmosphere. Betelgeuse has also been imaged by the Hubble Space Telescope at ultraviolet wavelengths showing the asymetric inner atmosphere of the star and a central hot spot. In a million years or so, the star's central core will collapse in a supernova explosion, leaving behind a neutron star or a white dwarf star remnant. *See* SPECKLE; STAR; SUPERNOVA.

Harold A. McAlister

Bibliography. A. Frankoi, D. Morrison, and S. C. Wolff, *Voyages Through the Universe*, 3d ed., Brooks Cole, 2004; R. L. Gilliland and A. K. Dupree, First image of the surface of a star with the Hubble Space Telescope, *Astrophys. J.*, 463:L29–L32, 1996; J. B. Kaler, *The Hundred Greatest Stars*, 2002; J. M. Pasachoff and A. Filippenko, *The Cosmos: Astronomy in the New Millennium*, 3d ed., Brooks Cole, 2007.

Bias (electronics)

The establishment of an operating point on the transistor volt-ampere characteristics by means of direct voltages and currents.

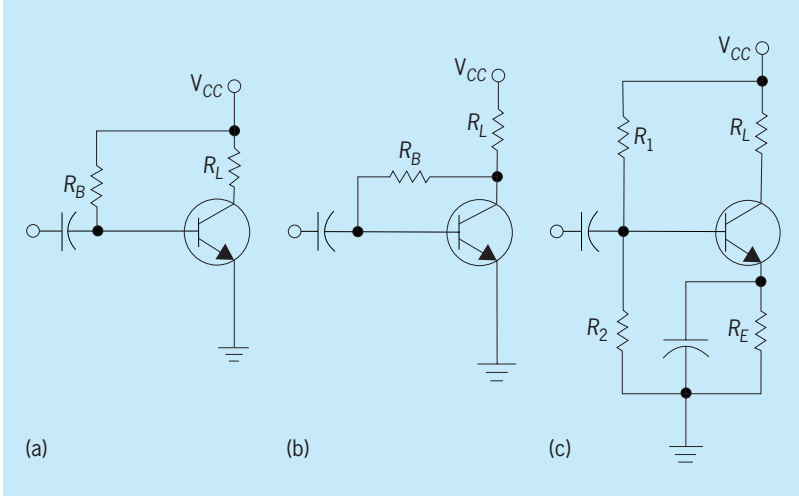

Fig. 1. Translator circuits. (*a*) Fixed-bias. (*b*) Collector-to-base bias. (*c*) Self-bias.

Biasing biopolar transistors. Since the transistor is a three-terminal device, any one of the three terminals may be used as a common terminal to both input and output. In most transistor circuits the emitter is used as the common terminal, and this common emitter, or grounded emitter, is indicated in **Fig. 1***a*. If the transistor is to used as a linear device, such as an audio amplifier, it must be biased to operate in the active region. In this region the collector is biased in the reverse direction and the emitter in the forward direction. The area in the common-emitter transistor characteristics to the right of the ordinate $V_{CE} = 0$ and above $I_C = 0$ is the active region. Two more biasing regions are of special interest for those cases in which the transistor is intended to operate as a switch. These are the saturation and cutoff regions. The saturation region may be defined as the region where the collector current is independent of base current for given values of V_{CC} and R_L. Thus, the onset of saturation can be considered to take place at the knee of the common-emitter transistor curves. *See* AMPLIFIER.

In saturation the transistor current I_C is nominally V_{CC}/R_L. Since R_L is small, it may be necessary to keep V_{CC} correspondingly small in order to stay within the limitations imposed by the transistor on maximum-current and collector-power dissipation. In the cutoff region it is required that the emitter current I_E be zero, and to accomplish this it is necessary to reverse-bias the emitter junction so that the collector current is approximately equal to the reverse saturation current I_{CO}. A reverse-biasing voltage of the order of 0.1 V across the emitter junction will ordinarily be adequate to cut off either a germanium or silicon transistor.

The particular method to be used in establishing an operating point on the transistor characteristics depends on whether the transistor is to operate in the active, saturation or cutoff regions; on the application under consideration; on the thermal stability of the circuit; and on other factors.

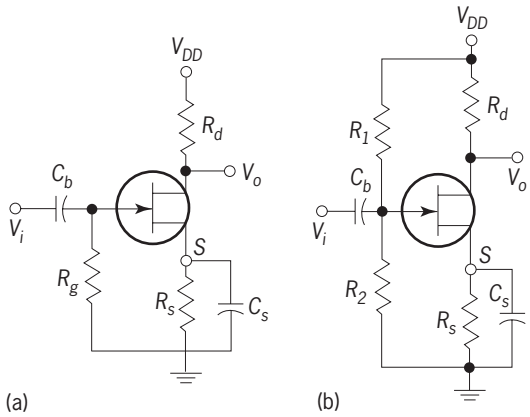

Fig. 2. Biasing circuits for FETs and MOSFETs. (*a*) Source self-bias circuit for an FET and a depletion-type MOSFET. (*b*) Biasing circuit for an FET and an enhancement-type MOSFET. (*After J. Millman Microelectronics McGraw-Hill, 1979*)

Fixed-bias circuit. The operating point for the circuit of Fig. 1*a* can be established by noting that the required current I_B is constant, independent of the quiescent collector current I_C, which is why this circuit is called the fixed-bias circuit. Transistor biasing circuits are frequently compared in terms of the value of the stability factor $S = \partial I_C/\partial I_{CO}$, which is the rate of change of collector current with respect to reverse saturation current. The smaller the value of S, the less likely the circuit will exhibit thermal runaway. S, as defined here, cannot be smaller than unity. Other stability factors are defined in terms of dc current gain h_{FE} as $\partial I_C/\partial h_{FE}$, and in terms of base-to-emitter voltage as $\partial I_C/\partial V_{BE}$. However, bias circuits with small values of S will also perform satisfactorily for transistors that have large variations of h_{FE} and V_{BE}. For the fixed-bias circuit it can be shown that $S = h_{FE} + 1$, and if $h_{FE} = 50$, then $S = 51$. Such a large value of S makes thermal runaway a definite possibility with this circuit.

Collector-to-base bias. An improvement in stability is obtained if the resistor R_B in Fig. 1*a* is returned to the collector junction rather than to the battery terminal. Such a connection is shown in Fig. 1*b*. In this bias circuit, if I_C tends to increase (either because of a rise in temperature or because the transistor has been replaced by another), then V_{CE} decreases. Hence I_B also decreases and, as a consequence of this lowered bias current, the collector current is not allowed to increase as much as it would if fixed bias were used. The stability factor S is shown in Eq. (1). This value

$$S = \frac{h_{FE} + 1}{1 + h_{FE}R_L/(R_L + R_B)} \tag{1}$$

is smaller than $h_{FE} + 1$, which is the value obtained for the fixed-bias case.

Self-bias. If the load resistance R_L is very small, as in a transformer-coupled circuit, then the previous expression for S shows that there would be no improvement in the stabilization in the collector-to-base bias circuit over the fixed-bias circuit. A circuit that can be used even if there is zero dc resistance in series with the collector terminal is the self-biasing configuration of Fig. 1*c*. The current in the resistance R_E in the emitter lead causes a voltage drop which is in the direction to reverse-bias the emitter junction. Since this junction must be forward-biased (for active region bias), the bleeder R_1-R_2 has been added to the circuit.

If I_C tends to increase, the current in R_E increases. As a consequence of the increase in voltage drop across R_E, the base current is decreased. Hence I_C will increase less than it would have had there been no self-biasing resistor R_E. The stabilization factor for the self-bias circuit is shown by Eq. (2), where

$$S = (1 + h_{FE})\frac{1 + R_B/R_E}{1 + h_{FE} + R_B/R_E} \tag{2}$$

$R_B = R_1R_2/(R_1 + R_2)$. The smaller the value of R_B, the better the stabilization. Even if R_B approaches zero, the value of S cannot be reduced below unity.

In order to avoid the loss of signal gain because of the degeneration caused by R_E, this resistor is often bypassed by a very large capacitance, so that its reactance at the frequencies under consideration is very small.

Biasing the FET and MOSFET. The selection of an appropriate operating point (I_D, V_{GS}, V_{DS}) for a field-effect transistor (FET) amplifier stage is determined by considerations similar to those given to transistors, as discussed previously. These considerations are output-voltage swing, distortion, power dissipation, voltage gain, and drift of drain current. In most cases it is not possible to satisfy all desired specifications simultaneously.

Source self-bias. The configuration shown in **Fig. 2*a*** can be used to bias junction FET devices or depletion-mode metal-oxide-semiconductor FETs (MOSFETs). For a specified drain current I_D, the corresponding gate-to-source voltage V_{GS} can be obtained analytically or by using the plotted drain or transfer characteristics. Since the gate current (and, hence, the voltage drop across R_g) is negligible, the source resistance R_s can be found as the ratio of V_{GS} to the desired I_D.

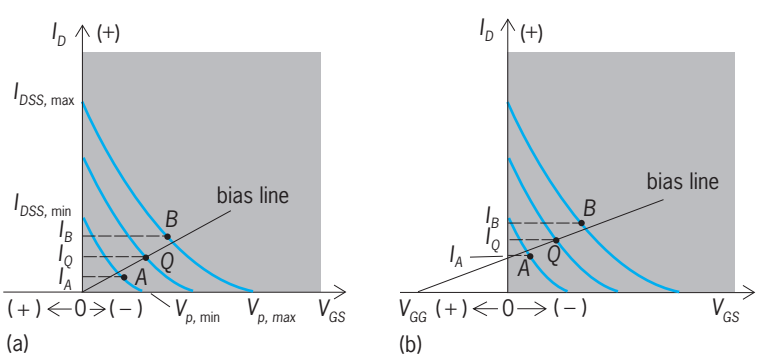

Fig. 3. Maximum and minimum transfer curves for biasing an *n*-channel FET. (*a*) Bias line passes through origin. (*b*) Bias line does not pass through origin. (*After J. Millman, Microelectronics, McGraw-Hill, 1979*)

Biasing against device variation. FET manufacturers usually supply information on the maximum and minimum values of I_{DSS} and V_P at room temperature, where I_{DSS} is the value of the drain current I_{DS} with $V_{GS} = 0$, and V_P is the pinch-off voltage. They also supply data to correct these quantities for temperature variations. The transfer characteristics for a given type of *n*-channel FET may appear as in **Fig. 3***a* where the top and bottom curves are for extreme values of temperature and device variation. Assume that it is necessary to bias the device at a drain current which will not drift outside $I_D = I_A$ and $I_D = I_B$. Then the bias line $V_{GS} = -I_D R_s$ must intersect the transfer characteristics between the points *A* and *B*, as indicated in Fig. 3*a*. The slope of the bias line is determined by the source resistance R_s. For any transfer characteristic between the two extremes indicated, the current I_Q is such that $I_A < I_Q < I_B$, as desired.

Consider the situation indicated in Fig. 3*b*, where a line drawn to pass between points *A* and *B* does not pass through the origin. This bias line satisfies Eq. (3).

$$V_{GS} = V_{GG} - I_D R_s \tag{3}$$

Such a bias relationship may be obtained by adding a fixed bias to the gate in addition to the source self-bias, as indicated in Fig. 2*b*. This circuit requires only one power supply and satisfies Eqs. (4), where it

$$V_{GG} = \frac{R_2 V_{DD}}{R_1 + R_2} \qquad R_g = \frac{R_1 R_2}{R_1 + R_2} \tag{4}$$

is assumed that the gate current is negligible. It is also possible for V_{GG} to fall in the reverse-biased region, so that the line in Fig. 3*b* intersects the axis of abscissa to the right of the origin. Under these circumstances, two separate supply voltages must be used.

Biasing the enhancement MOSFET. The self-bias technique of Fig. 2*a* cannot be used to establish an operating point for the enhancement-type MOSFET because the voltage drop across R_s is in a direction to reverse-bias the gate, and a forward gate bias is required. The circuit of **Fig. 4***a* can be used, and for this case $V_{GS} = V_{DS}$, since no current flows through R_f. If, for reasons of linearity in device operation or maximum output voltage, it is desired that $V_{GS} \neq V_{DS}$, then the circuit of Fig. 4*b* is suitable. In this case, $V_{GS} = [R_1/(R_1 + R_f)] \cdot V_{DS}$. Both circuits discussed here offer the advantages of dc stabilization through the feedback introduced with R_f.

The circuit of Fig. 2*b* is often used with the enhancement MOSFET. The dc stability introduced in Fig. 4 through the feedback resistor R_f is then missing, and is replaced by the dc feedback through R_s. *See* TRANSISTOR.

Christos C. Halkias

Bibliography. P. Horowitz and W. Hill, *The Art of Electronics*, 2d ed., 1989; J. Millman, *Microelectronics*, 2d ed., 1987; J. Singh, *Semiconductor Devices: An Introduction*, 1994.

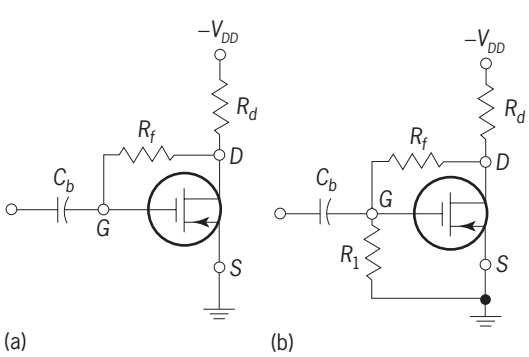

Fig. 4. Drain-to-gate bias circuits for enhancement-type MOSFET. (*a*) Circuit with $V_{GS} = V_{DS}$. (*b*) Improved version, with $V_{GS} \neq V_{DS}$. (*After J. Millman, Microelectronics, McGraw-Hill, 1979*)

Bicosoecida

An order of Zoomastigophorea (Protozoa). They are colorless, free-living cells, each with two flagella, one of which is used for attaching the organism to its exoskeleton (lorica). Although the attachment is normally from the front end, the flagellum emerges alongside the primary or vibrating one (see **illus.**). The anterior end in many species

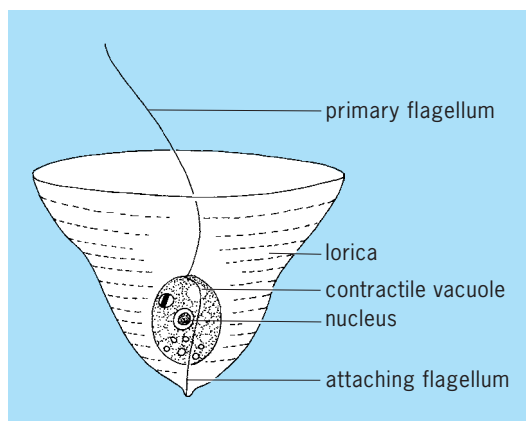

Bicosoecid, *Codomonas annulata.*

(*Codomonas annulata, Poteriodendron petiolatum*, and *Stephanocodon stellatum*) is ameboid, or at times appears to be formed into a lip which can turn in and engulf a bacterium. *Stenocodon* has the same sort of lip but neither of the two flagella is used for attachment. These Bicosoecida are common in fresh water, often seen attached to desmids or other algae. *Bicosocca mediterranea* is common at times in salt water, where assemblages of it are found on diatoms. The order has very few genera and species, and individuals easily escape observation because of small size and transparency. *See* CILIA AND FLAGELLA; PROTOZOA; ZOOMASTIGOPHOREA.

James B. Lackey

Big bang theory

The theory that the universe began in a state of extremely high density and has been expanding since some particular instant that marked the origin of the universe. The big bang is the generally accepted cosmological theory; the incorporation of developments in elementary particle theory has led to the inflationary universe version. The predictions of the inflationary universe and older big bang theories are the same after the first 10^{-35} s. *See* INFLATIONARY UNIVERSE COSMOLOGY.

Two observations are at the base of observational big bang cosmology. First, the universe is expanding uniformly, with objects at greater distances receding at a greater velocity. Second, the Earth is bathed in the cosmic background radiation, an isotropic glow of radiation that has the characteristics expected from the remnant of a hot primeval fireball. Since the last years of the twentieth century, a third observation has come to the fore: the expansion of the universe is accelerating.

Cosmological theory in general and the big bang theory in particular are based on the theory of gravitation advanced by A. Einstein in 1916 and known as the general theory of relativity. Though the predictions of this theory have little effect in the limited sphere of the Earth, they dominate on as large a scale as the universe, and have been well tested in such sources as the binary pulsar. *See* RELATIVITY.

Expansion of the universe. In the 1920s it became clear that the "spiral nebulae," clouds of gas with arms spiraling outward from a core, were galaxies on the scale of the Milky Way Galaxy. This was established in 1925, when observations of variable stars in several galaxies by E. Hubble enabled the distance to these galaxies to be determined with some accuracy.

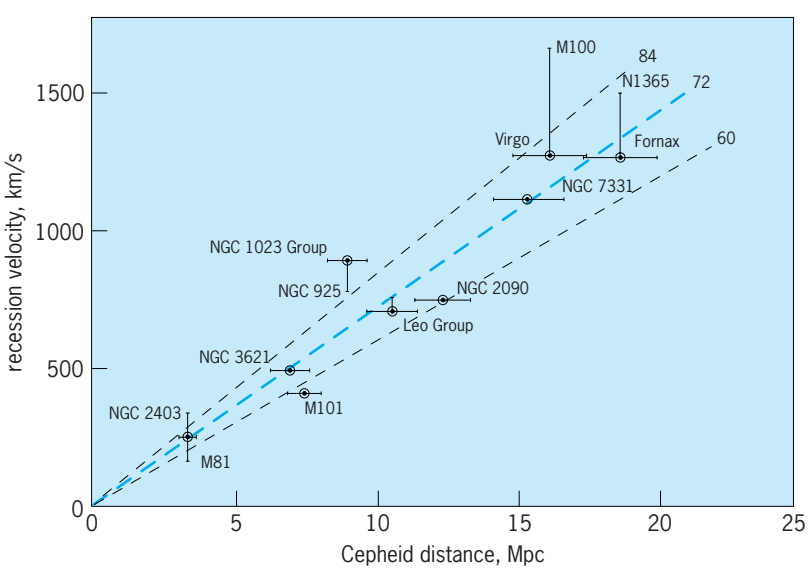

Fig. 1. Hubble diagram. Numbers next to the straight lines passing through the origin indicate corresponding values of Hubble's constant in km · s^{-1} · Mpc^{-1}. Data on this diagram give $H_0 = 72$ km · s^{-1} · Mpc^{-1}, with an uncertainty of ± 4 km · s^{-1} · Mpc^{-1} and a scatter of ± 17 k · s^{-1} · Mpc^{-1}. (*W. L. Freedman, B. Madore, and R. C. Kennicutt; NASA*)

Starting in 1912, the spectra of many of these spiral nebulae were found to have large redshifts. According to the Doppler effect, these large redshifts correspond to large velocities of recession from the Earth. Estimates of distances to these objects by Hubble and colleagues established a direct relation between the distance to a galaxy and its velocity of recession. It was soon interpreted as the result of an expanding universe. *See* DOPPLER EFFECT; REDSHIFT.

The relation known as Hubble's law is $v = H_0 d$, where v is the velocity of recession, d is the distance to the galaxy, and H_0 is a constant known as Hubble's constant. Determining Hubble's constant requires the independent measurement of the distances to galaxies; the redshift can easily be measured on spectra.

In the Hubble diagram (**Fig. 1**), the horizontal axis is the distance as derived from observations of the periods of Cepheid variable stars in distant galaxies and the vertical axis is the velocity of recession, calculated from the redshift $z = \Delta\lambda/\lambda$, where $\Delta\lambda$ is the shift in wavelength in radiation of wavelength λ measured on the spectra of those galaxies. The Hubble law is expressed by the fact that the data points in the diagram lie near a straight line. *See* CEPHEIDS.

If velocity is expressed in kilometers per second, and distance is expressed in megaparsecs (where 1 parsec is the distance from which the radius of the Earth's orbit would subtend 1 second of arc, and is equivalent to 3.26 light-years or 3.09×10^{13} km or 1.92×10^{13} mi), then Hubble's constant H_0 is given in km · s^{-1} · Mpc^{-1}. The determination of Hubble's constant was the aim of one of the three Key Projects of the *Hubble Space Telescope*, especially through studies of Cepheid variable stars in galaxies sufficiently distant to be beyond internal gravity effects in the Local Group, the cluster of galaxies to which the Milky Way belongs. The result is 72 km · s^{-1} · Mpc^{-1}. Tying the *Hubble Space Telescope* results to distances derived from supernovae extends Hubble's constant to still greater distances. Other methods involving planetary nebulae, the spread of rotational velocities of galaxies, the statistical fluctuations of each pixel of galaxy images, and supernovae, for example, also gave values close to 75 km · s^{-1} · Mpc^{-1}. The result from the National Aeronautics and Space Administration's (NASA) *Wilkinson Microwave Anisotropy Probe (WMAP)* closely confirms the Hubble Key Project value. *See* GALAXY, EXTERNAL; HUBBLE CONSTANT; LOCAL GROUP; SUPERNOVA.

Reversing the expansion rate, tracing the expansion of the universe back in time shows that the universe would have been compressed to infinite density approximately 8–16 $\times 10^9$ years ago (for $H_0 =$ 72 km · s^{-1} · Mpc^{-1}). The *WMAP* pinned this value down more precisely, to 13.7×10^9 years ago. In the big bang theory, the universe began at that time as a so-called big bang began the expansion. The big bang was the origin of space and time.

In 1917, Einstein found a solution to his own set of equations from his general theory of relativity that predicted the nature of the universe. His universe,

though, was unstable: it could only be expanding or contracting. This seemed unsatisfactory at the time, for the expansion had not yet been discovered, so Einstein arbitrarily introduced a special term—the cosmological constant—into his equations to make the universe static. The need for the cosmological constant seemed to disappear with Hubble's discovery of the expansion, though the cosmological constant has subsequently reappeared in some models. *See* COSMOLOGICAL CONSTANT.

Further solutions to Einstein's equations, worked out in the 1920s, are at the basis of the cosmological models that are now generally accepted. These solutions indicate that the original "cosmic egg" from which the universe was expanding was hot and dense. This is the origin of the current view that the universe was indeed very hot in its early stages.

Early universe. Modern theoretical work has been able to trace the universe back to the first instants in time. In the big bang theory and in related theories that also propose a hot, dense early universe, the universe may have been filled in the earliest instants with exotic elementary particles of the types now being studied by physicists with large accelerators. Individual quarks may also have been present. By 1 microsecond after the universe's origin, the exotic particles and the quarks had been incorporated in other fundamental particles. *See* ELEMENTARY PARTICLE; QUARKS.

Work in the early 1980s incorporated the effect of elementary particles in cosmological models. The research seems to indicate that the universe underwent a period of extremely rapid expansion in which it inflated by a factor of billions in a very short time. This inflationary universe model provides an explanation for why the universe is so homogeneous: Before the expansion, regions that now seem too separated to have been in contact were close enough to interact. After the inflationary stage, the universe was in a hot stage and was still dense; the models match the big bang models thereafter.

In the inflationary universe models, the universe need not have arisen from a single big bang. Rather, matter could have appeared as fluctuations in the vacuum.

It is not definitely known why there is an apparent excess of matter over antimatter, though attempts in elementary particle physics to unify the electromagnetic, the weak, and the strong forces show promise in explaining the origin of the matter–antimatter asymmetry. The asymmetry seems to have arisen before the first millisecond. The asymmetry in the decay of certain mesons may provide a clue to resolving this question. *See* ANTIMATTER; FUNDAMENTAL INTERACTIONS.

By 5 s after the origin of the universe, the temperature had cooled to 10^9 K (2×10^9 °F), and only electrons, positrons, neutrinos, antineutrinos, and photons were important. A few protons and neutrons were mixed in, and they grew relatively more important as the temperature continued to drop. The universe was so dense that photons traveled only a short way before being reabsorbed. By the time 1 min had gone by, nuclei of the light elements had started to form.

After about 400,000 years, when the universe cooled to 3000 K (5000°F) and the density dropped sufficiently, the protons and electrons suddenly combined to make hydrogen atoms, a process called recombination. Since hydrogen's spectrum absorbs preferentially at the wavelengths of sets of spectral lines rather than continuously across the spectrum, and since there were no longer free electrons to interact with photons, the universe became transparent at that instant. The average path traveled by a photon—its mean free path—became very large. The blackbody radiation of the gas at the time of recombination was thus released and has been traveling through space ever since. As the universe expands, the spectrum of this radiation retains its blackbody shape though its characteristic temperature drops. *See* BLACKBODY; HEAT RADIATION.

Background radiation. Between 1963 and 1965, observations with a well-calibrated horn antenna revealed an isotropic source of radio noise whose strength was independent of time of day and of season, and whose intensity at the observing wavelength of 7 cm (2.8 in.) was equivalent to that which would be emitted by a blackbody—an idealized radiating substance—at a temperature of about 3 K (−454 °F). Prior calculations that indicated that the remnant radiation from the big bang might be detectable led to the interpretation of the newly discovered radiation in terms of fossil radiation from the big bang. In 1975, final confirmation of the blackbody nature of the radiation was provided by measurements of the spectrum of the background radiation in a frequency range that included the peak of the blackbody curve and extended into the infrared.

The observations of the spectrum from the *Cosmic Background Explorer* (*COBE*) spacecraft during 1989–1993 were in complete agreement with a blackbody curve at 2.728 K (−454.7°F). The *WMAP*'s value, released in 2003, is in close agreement, and updates from *WMAP* will be released during and after its four-year lifespan. The radiation is very isotropic, once the dipole anisotropy caused by the Earth's motion is subtracted. Observations from another experiment on *COBE* discovered temperature fluctuations at the level of 5 parts per million. *WMAP* observed them in still finer detail. These fluctuations, largely statistical in nature, are revealing the seeds from which galaxies formed. Inflationary big bang models seem best able to explain how widely separated parts of the universe can be as similar as the *COBE* and *WMAP* observations have revealed them to be. Confirmations of these ripples in space have come from cosmic background telescopes on balloons and in Canada and Antarctica. Spacecraft missions with improved resolution and sensitivity, including the European Space Agency's *Planck* mission (2007), are planned. *See* COSMIC BACKGROUND RADIATION.

Nucleosynthesis. As the early universe cooled, the temperatures became sufficiently low for element

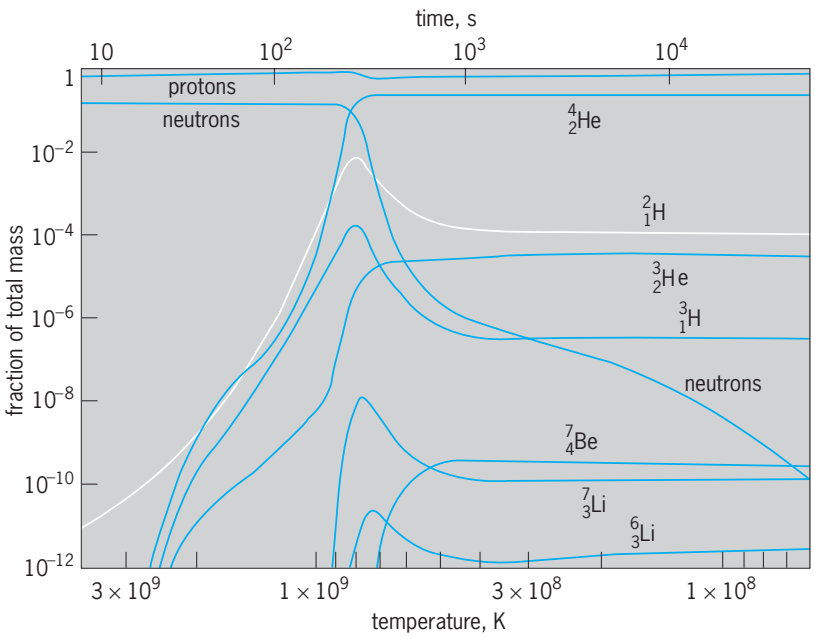

Fig. 2. Relative abundance of isotopes of the light elements in the first few minutes after the big bang according to the model of R. V. Wagoner, for a certain assumed density at some early time. (*After J. M. Pasachoff, Astronomy: From the Earth to the Universe, 6th ed., Brooks/Cole, 2002*)

formation to begin. By about 100 s, deuterium (comprising one proton plus one neutron) formed. When joined by another neutron to form tritium, the amalgam soon decayed to form an isotope of helium. Ordinary helium, with still another neutron, also resulted.

The relative abundances of isotopes of the light elements in the first few minutes have been calculated (**Fig. 2**). The calculations depend on a certain assumed density at some early time. They show that within minutes the temperature drops to 10^9 K (2×10^9 °F), too low for most nuclear reactions to continue. Most models give a resulting abundance of about 25% of the mass in the form of helium, regardless of the density of the universe. The helium abundance is hard to determine observationally. Current results are in rough agreement with the theoretical value.

The abundances of others of the light elements are more sensitive to parameters of matter in the early universe. Such abundances can be calculated as a function of the present-day density of matter (**Fig. 3**). From knowledge of the approximate rate of expansion of the universe, the density of matter in the early universe can be deduced from the current density, and abundances can then be calculated. In particular, the deuterium abundance is especially sensitive to the cosmic density at the time of deuterium formation, because the rate at which deuterium is "cooked" into tritium increases rapidly with increasing density.

Big bang nucleosynthesis, although at first thought to be a method of forming all the elements, foundered for the heavy elements at mass numbers 5 and 8. Isotopes of these mass numbers are too unstable to form heavier elements quickly enough. The gap is bridged only in stars, through processes worked out in 1957. Thus the lightest elements were formed as a direct result of the big bang (Fig. 3), while the heavier elements as well as additional quantities of most of the lighter elements were formed later in stars or supernovae. Measurements of the relative abundances of the light isotopes and elements have been used to set a limit of three on the number of neutrino types and thus on the number of quark-lepton families that exist. *See* LEPTON; NEUTRINO; NUCLEOSYNTHESIS.

Open versus closed universe. The two extreme possibilities for the future of the universe are that the universe will continue to expand forever, or that it will cease its expansion and begin to contract. It can be shown that the case where the universe will expand forever, in the absence of dark energy or other accelerating mechanisms, corresponds to an infinite universe. The term applied is the open universe. The case where the universe will begin to contract corresponds to a finite universe. The term applied is the closed universe. The inflationary universe scenario has the universe on the boundary between open and closed, as signified by the parameter Ω taking the value of 1. Such a universe would expand forever but at an ever-decreasing rate in the absence of accelerating mechanisms. However, according to more recent conclusions, discussed below, the expansion of the universe is now accelerating, having passed a transition point where dark energy has begun to dominate.

Deceleration parameter. One basic way to test whether the universe is open or closed is to determine the rate at which the expansion of the universe is slowing, that is, the rate at which it is deviating from Hubble's law. The slowing is measured with a term

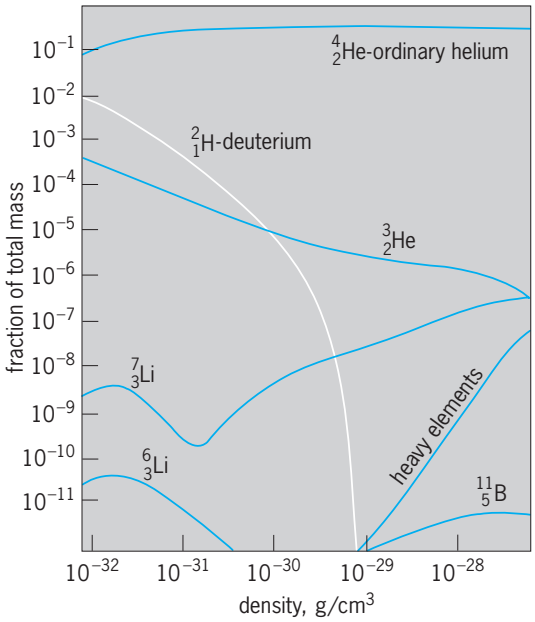

Fig. 3. Relative abundances of isotopes and elements as a function of present-day density of matter, according to the model of R. V. Wagoner. (*After J. M. Pasachoff, Astronomy: From the Earth to the Universe, 6th ed., Brooks/Cole, 2002*)

called q_0, the deceleration parameter. The dividing line is marked by $q_0 = {}^1/_2$, with smaller values corresponding to an open universe and larger values corresponding to a closed universe. The value $q_0 = {}^1/_2$ corresponds to $\Omega = 1$, where Ω is the ratio of the total density of matter in the universe to the critical density that divides open from closed universes.

The most obvious way to measure q_0 is by looking at the most distant galaxies or clusters of galaxies, measuring their distances independently from Hubble's law, and plotting Hubble's law in a search for deviations. But all such observations are subject to the considerable uncertainty introduced by the fact that observations deep into space also see far back into time. The galaxies then were surely different than they are now, so evolutionary effects must be taken into account. It is not even known for certain whether galaxies were brighter or dimmer in the distant past, and thus the method may reveal more about the evolution of galaxies than about cosmology.

Accelerating universe. The deceleration parameter is being studied from very distant supernovae, which are discovered in quantity through automated searches. Astonishingly, evidence seems to show that the expansion of the universe is accelerating, the opposite of what was expected. These data must be scrutinized for systematic errors before they can be definitely accepted, but enough independent confirmation has appeared (as of 2005) from different kinds of observations that doubt of this astonishing result is vanishing.

Observations of extremely distant supernovae have given astronomers views far enough back in time to see the era before the acceleration started. In that era, the universe's expansion was slowing. These discoveries gave confidence to astronomers that the conclusion that the universe is accelerating is correct. Results in 2004 from observations by the *Chandra X-ray Observatory* of extremely distant clusters of galaxies have confirmed that the universe's expansion is accelerating.

The results from *WMAP,* endorsing earlier results from studying ripples in the cosmic background radiation, have shown that the universe's content is exactly at the critical value between open and closed. At the same time, they show (endorsing nucleosynthesis observations summarized below) that the amount of ordinary matter (composed of atomic particles called baryons, like protons and neutrons) is only 4% of this critical value. Other studies, such as those of the gravitational attractions of objects in clusters of galaxies, show the existence of dark matter—with gravity but without emitting electromagnetic radiation—accounting for another approximately 30% of the critical value. The remaining two-thirds of the content of the universe remains unknown, and has been given the name "dark energy." Scientists have no definite idea of what the dark energy is composed, though Einstein's old "cosmological constant"—a seemingly arbitrary term in his equations—is one of the possibilities being considered. This dark energy now dominates the universe, causing the acceleration of the expansion. When the cosmological constant Λ is taken into account, $q_0 = {}^1/_2\,\Omega_m - \Omega_\Lambda$ (where Ω_m is associated with the matter in the universe and Ω_Λ is associated with the cosmological constant), so it is now negative. *See* ACCELERATING UNIVERSE; BARYON; DARK ENERGY; DARK MATTER.

Great Attractor. From studies of elliptical and spiral galaxies, a bulk streaming motion has been found toward a region called the Great Attractor. The region would be a concentration in the direction of the constellation Centaurus, 20 times more populous than the Virgo Cluster. Its distortion of the Hubble flow may have led to misevaluation of the Hubble constant. An even greater attractor farther out has also been reported. The presence of the Great Attractor and related regions of space mean that the measurements of velocities we made in local areas of the universe can be affected by the Great Attractor's mass, where it gave misleading results about the Hubble flow before the *WMAP* and supernova determinations of Hubble's constant. *See* VIRGO CLUSTER.

Cosmic deuterium abundance. Because deuterium is a sensitive indicator of the cosmic density at the time of its formation soon after the big bang, and because deuterium can only be destroyed and not formed in stars, the study of the cosmic deuterium abundance is one of the best methods for determining the future of the universe. Basically, it involves assessing whether there is enough gravity in the universe to halt the expansion.

Counts can be made of all the stars, galaxies, quasars, interstellar matter, and so forth, and the sum gives a density much too low to close the universe. But this method does not assess the amount of invisible matter, which could be in the form of intergalactic gas, black holes, and so forth. Indeed, studies of the motions of galaxies in clusters of galaxies often indicate that much more mass is present inside galaxy clusters than is visible. The amount of this missing mass (actually the mass is present and it is the light that is missing) may be 50 times the amount of ordinary mass (known as baryons). As previously noted, it is known as dark matter.

Assessing the density of the universe through studies of the deuterium abundance is independent of whether the ordinary matter is visible or invisible (detectable, for example, from its radio or x-ray emissions). Although deuterium is present even on Earth as a trace isotope, with an abundance 1/6600 that of normal hydrogen in seawater, it is difficult to detect in interstellar space. However, since 1972 a number of determinations of interstellar deuterium abundance have been carried out, through radio observations and studies in the ultraviolet made with telescopes in space, the most precise with the *Hubble Space Telescope*. Determinations of deuterium in the atmospheres of the planets have also been carried out. Though there are some discrepancies remaining between abundances determined in different locations and in different fashions, the deuterium observations seem to indicate clearly that the density of the universe is very low and hence that the

universe is open as far as baryons are concerned. This finding is supported by observations with one of the 10-m (400-in.) Keck telescopes of deuterium in a distant quasar, redshifted into the visible part of the spectrum. This result that the baryon density is low, only about 4% of the critical value that divides open and closed universes, matches the conclusions from *WMAP*'s observations of the ripples in the cosmic background radiation. *See* DEUTERIUM.

Hot and cold dark matter. Observations of a diffuse background of x-rays had led to the suspicion that a lot of hot material, previously undiscovered, may have been present between the galaxies. However, subsequent observations revealed that at least some of the x-ray background came from faint quasars, which appeared on long exposures of even fields that had been thought to be blank. Continued mapping and pointed observations have detected extensive halos around certain galaxies and even around some sparse clusters of galaxies. These observations indicated the presence of extensive dark matter.

Since neutrinos are known to be plentiful in the universe, if they each had enough mass they could dominate the universe. Since there are about 100 neutrinos in each cubic centimeter (1500 in each cubic inch) of the universe, much mass would be in that form if neutrinos had even a small rest mass. Indeed, observations of neutrinos with the Super-Kamiokande detector show that neutrinos change in type, which can occur only if they have mass. But the mass itself is not determined from these observations. Observations of the neutrinos from Supernova 1987A placed a sensitive limit on the amount of mass that neutrinos could have, since the arrival times of massive neutrinos would have spread out. Since neutrinos move so fast, they are known as hot dark matter. But observations from *WMAP* showed that the total mass of neutrinos does not dominate the universe.

Since hot dark matter would not explain satisfactorily the quickness with which matter clumped to form galaxies, most models include at least some so-called cold dark matter. The cold dark matter could be in the form of elementary particles of types yet unknown, such as axions or weakly interacting massive particles (WIMPs). Or it could be in the form of dim stars, brown dwarfs, or other macroscopic bodies known collectively as massive astrophysical compact halo objects (MACHOs). Supercomputer calculations involving millions of points interacting in a three-dimensional grid have computed the evolution of the universe for models with different values of Ω and for different combinations of cold and hot dark matter. *See* BROWN DWARF; WEAKLY INTERACTING MASSIVE PARTICLE (WIMP).

Inflationary scenarios. Although several lines of evidence, including studies of differential velocities of nearby galaxies caused by density perturbations and of the cosmic abundance of deuterium, indicate that the universe is open, it was not clear why the universe was so close to the dividing line between being open or closed that the subject was much in doubt. The inflationary universe model provides a natural explanation for the universe being on this dividing line. After expansion slows down at the close of the inflationary stage (thus causing a phase change, much like supercooled liquid water turning abruptly into ice when perturbed, a more extreme example of a phase change than the household boiling of water into steam), the universe necessarily approaches this line. Thus whether the universe was open or closed was the wrong question to ask. Further work on inflationary scenarios continues to assess certain problems, such as the inflationary model's predictions of the density fluctuations that lead to the coalescence of galaxies, now measured by the *WMAP* spacecraft. Though the simplest inflationary scenarios have been ruled out, more complex inflationary theories are viable and are widely accepted by cosmologists. *See* COSMOLOGY; UNIVERSE.

Jay M. Pasachoff

Bibliography. T. Ferris, *The Whole Shebang: A State of the Universe(s) Report*, Simon & Schuster, 1997; S. W. Hawking, *A Brief History of Time: The Updated and Expanded Tenth Anniversary Edition,* Bantam, 1998; J. Silk, *A Short History of the Universe*, Freeman, 1997; A. H. Guth and A. P. Lightman, *The Inflationary Universe: The Quest for a New Theory of Cosmic Origins*, Perseus, 1998; R. Kirshner, *The Extravagant Universe*, Princeton University Press, 2002; J. M. Pasachoff and A. Filippenko, *The Cosmos: Astronomy in the New Millennium*, 2d ed., Brooks/Cole, 2004, 3d ed., 2007; M. Rees, *Our Cosmic Habitat*, Princeton University Press, 2001; S. Weinberg, *The First Three Minutes*, 2d ed., BasicBooks, 1994.

Bilateria

A major division of the animal kingdom comprising all forms with a bilaterally symmetrical body plan (that is, their left and right sides are mirror images, or nearly so) and organ systems that develop from three tissue layers. The animal kingdom, Metazoa, is subdivided into about 30 major animal groups, or phyla, each constructed on a different architectural plan. Most of these phyla are grouped as the Bilateria, comprising a major early branch on the family tree of animals that may have originated nearly 600 million years ago. All bilaterians possess some features that are not perfectly symmetrical (for example, the human heart and liver), but all have descended from an essentially bilateral ancestor. The nonbilaterian phyla include sponges (which lack muscles and nerves), jellyfish, and sea anemones and their allies, which have quasiradial symmetry and organ systems that develop from only two tissue layers (see **illustration**). *See* ANIMAL KINGDOM; ANIMAL SYMMETRY; METAZOA.

Classification. Within Bilateria three main branches are usually recognized (see illustration). The first branch, Deuterostomia, includes the phyla Echinodermata (for example, sea stars) and Chordata (including vertebrates). A second branch, Ecdysozoa, includes Arthropoda (for example,

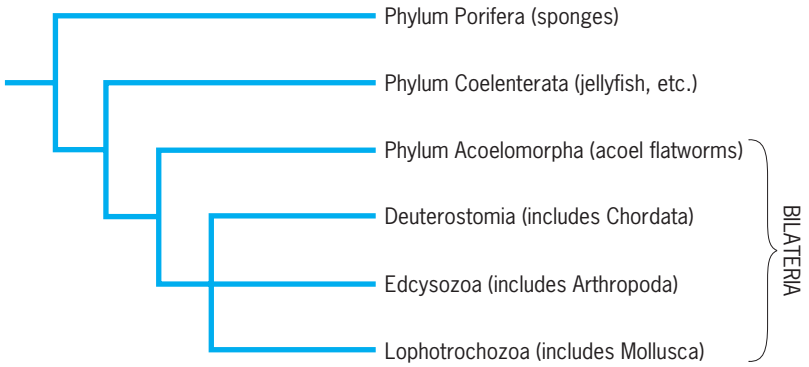

Position of Bilateria and of its principal subdivisions within the animal kingdom (Metazoa).

insects and crustaceans) and allied forms. The third branch, Lophotrochozoa, includes Mollusca (for example, clams and snails), Annelida (for example, earthworms), and allied forms. *See* ANNELIDA; ARTHROPODA; CHORDATA; ECHINODERMATA; MOLLUSCA.

Fossil record. The first metazoan fossils date back about 580 million years ago; they may include bilaterians, though their identities are not certain. The earliest evidence of bilaterian activity comes from 560- to 555-million-year-old trace fossils left by organisms crawling on muddy sea floors, leaving a trail in the sediment. Those organisms were small (less than a millimeter in width), and their precise body architecture is uncertain. Although the fossil record between 560 and 530 million years ago contains a few enigmatic forms that may be larger bilaterians, the recognizable ancestors of many living bilaterian phyla, especially those with durable skeletons, did not appear until about 530 to 520 million years ago. These appearances, collectively known as the Cambrian explosion, are our first glimpse of the wonderfully disparate bilaterian architectures known today. *See* FOSSIL; TRACE FOSSILS.

Living bilaterians. The earliest branch of living bilaterians comprises small simple worms called acoels. These worms probably represent the general architectural plan of the trace-makers of 560 million years ago. Acoels lack true body cavities, and instead of a brain they have a nerve plexus. Circulatory and excretory systems are also absent; since they are minute, diffusion is adequate for exchanging metabolic materials with the environment.

Simple worms are composed of approximately 15 different types of cells. In comparison, humans have a few hundred cell types and are vastly more complex, yet they have only about twice the number of genes as simple worms. Many of the same genes are present in most bilaterians, which have achieved their diversity and complexity by the use of different patterns of gene expression in increasingly complex networks to produce different body architectures in diverging lineages adapting to distinctive ways of life. The rise of this extensive morphological diversity and complexity is much of the story of animal evolution. *See* ANIMAL EVOLUTION; GENE.

James W. Valentine

Bibliography. R. C. Brusca and G. J. Brusca, *Invertebrates*, 2d ed., Sinauer Associates, Sunderland, MA, 2002; J. Cracraft and M. J. Donoghue (eds.), *Assembling the Tree of Life*, Oxford University Press, 2004; C. Nielsen, *Animal Evolution*, 2d ed., Oxford University Press, 2001; J. W. Valentine, *On the Origin of Phyla*, University of Chicago Press, 2004.

Bilirubin

The predominant orange pigment of bile. It is the major metabolic breakdown product of heme, the prosthetic group of hemoglobin in red blood cells, and other chromoproteins such as myoglobin, cytochrome, and catalase. The breakdown of hemoglobin from the old red cells takes place at a rapid rate in the reticuloendothelial cells of the liver, spleen, and bone marrow. The steps in this breakdown process include denaturation and removal of the protein globin, oxidation and opening of the tetrapyrole ring, and the removal of iron to form the green pigment biliverdin, which is then reduced to bilirubin by the addition of hydrogen. The formed bilirubin is transported to the liver, probably bound to albumin, where it is conjugated into water-soluble mono- and diglucuronides and to a lesser extent with sulfate. *See* LIVER.

In mammalian bile essentially all of the bilirubin is present as a glucuronide conjugate. The conjugation is catalyzed by the liver microsomal enzyme uridyl diphosphoglucuronyltransferase. Water-soluble bilirubin (conjugated) forms a purple substance when it is reacted with diazotized sulfanilic acid (van den Bergh reaction). Free bilirubin (unconjugated) must be made water-soluble, usually by the addition of alcohol, before the reaction will take place. Plasma bilirubin may be increased in diseases which result in increased hemoglobin destruction or diminished hepatic function.

Bilirubin glucuronide is passed through the liver cells into the bile caniculi and then into the intestine. The bacterial flora further reduce the bilirubin to colorless urobilinogen. Most of the urobilinogen is either reduced to stercobilinogen or oxidized to urobilin. These two compounds are then converted to stercobilin, which is excreted in the feces

and gives the stool its brown color. A small amount of urobilinogen is absorbed into the bloodstream through the portal circulation. Some is reconverted into bilirubin and a small portion is excreted unchanged into the urine. On standing, the urinary urobilinogen is oxidized to urobilin. The principal yellow pigment of urine, urochrome, is a compound of urobilin, urobilinogen, and a peptide.

The conjugation of bilirubin is reduced at birth and gradually becomes effective in a period of several days to 2 weeks. A genetic defect in an enzyme mechanism can result in excessive formation of unconjugated bilirubin. *See* HEMOGLOBIN.

Morton K. Schwartz

Bibliography. F. B. Armstrong and T. P. Bennet, *Biochemistry*, 1979; D. Bergsma (ed.), *Bilirubin Metabolism in the Newborn*, vol. 2, 1976; G. J. Dutton (ed.), *Glucuronic Acid, Free and Combined*, 1966.

Binary asteroid

A system composed of two almost equal-size asteroids orbiting around their center of mass, or one moonlet revolving around a larger primary body. The discovery of binary asteroids will aid the work of observers and theorists alike, since they help to determine the nature of these remnants of the solar system formation.

After the *Galileo* spacecraft discovered the first asteroid companion in 1993 [Dactyl, a 1.6-km (1.0-mi) companion of main-belt asteroid 243 Ida], it was realized that satellites might in fact be common around main-belt asteroids. A team led by W. J. Merline reported in 1999 the first direct detection of a satellite, Petit-Prince, of asteroid 45 Eugenia, using the Canada-Hawaii-France Telescope. From a study of the companion's orbit and via Kepler's third law, they derived a low bulk density for Eugenia and concluded that its structure either is that of a "rubble pile" or is composed of primitive, icy materials.

Occurrence. These fundamental findings have motivated several groups to conduct systematic searches for possible satellites of main-belt asteroids using adaptive optics imaging systems, and discoveries are accumulating (see **illus.**). At the begining of 2004, over 20 main-belt binary asteroids had been observed, and this number was growing rapidly. Based upon these surveys, it is estimated that about 2% of all main-belt asteroids are double. Studies suggest that less than 4% of the Trojan asteroids are binary. *See* TROJAN ASTEROIDS.

Evidence is also growing that near-Earth asteroids have companions. The binary nature of several near-Earth asteroids was established by radar observations. The complex visible-light curves of other near-Earth asteroids and the presence of doublet craters on the surface of terrestrial planets indicate that about 10–20% of near-Earth asteroids are binary systems.

After the binary trans-Neptunian object system consisting of Pluto and Charon, additional double trans-Neptunian objects were discovered, beginning in 2001. Hubble Space Telescope observations revealed that most of them are composed of similarly sized companions separated by more than 22,000 km (13,700 mi). *See* HUBBLE SPACE TELESCOPE; KUIPER BELT.

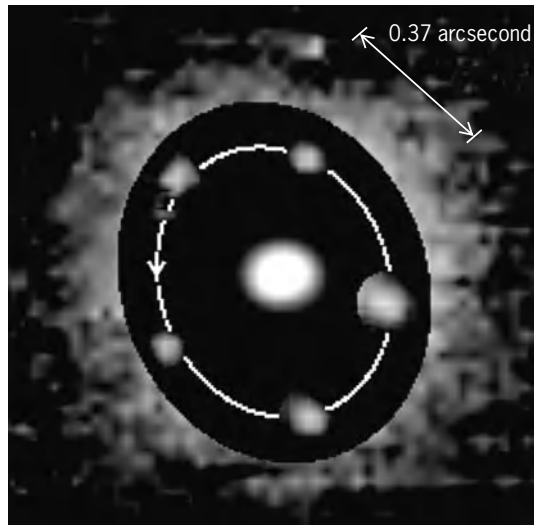

Asteroid 121 Hermione and its moonlet companion observed with the adaptive optics system mounted on the European Southern Observatory's Very Large Telescope [consisting of four 8.2-m (25-ft) telescopes] at six different times. The satellite revolves at a distance of about 790 km (490 mi) from the primary in 1.65 days. The angle of 0.37 arcsecond corresponds to a length of 750 km (466 mi) at the distance of the asteroid. (*After F. Marchis et al., On the mass and density of 121 Hermione from the analysis of its companion orbit, Icarus, in revision, 2004*)

Knowledge from observations. Until recently, little was known about the internal composition and structure of the asteroids. Binary asteroids provide unique information about the internal properties (mass, density, and porosity) of their components, as well as about the formation, history, and evolution of the solar system. This information otherwise may only be speculated upon from spacecraft mission flybys, which are extremely uncommon. In addition, the study of the companion orbits is a unique experiment to improve the knowledge of dynamical interactions. For example, both the irregular shape of the primary and strength of the solar tide influence the orbital distance and the system's stability. Also, the angular momentum distribution within the system (that is, the senses of rotation of the components, the separation of the components, the inclination of the orbiting system plane, and so forth) helps to distinguish between formation scenarios such as fast-rotating fission, fragment ejection, impact disruption, gravitational capture, or tidal force disruption. *See* ASTEROID; SOLAR SYSTEM.

Franck Marchis

Bibliography. W. J. Merline et al., Asteroids *do* have satellites, in W. F. Bottke et al. (eds.), *Asteroids III*, pp. 289–312, University of Arizona Press, 2003; S. J. Weidenschilling, P. Paolicchi, and V. Zappala, Do asteroids have satellites?, in R. P. Binzel, T. Gehrels, and M. S. Matthews (eds.), *Asteroids II*, pp. 643–658, University of Arizona Press, 1989.

Binary star

Two stars held together by their mutual gravitational attraction in a permanent (or long-term) association. The stellar universe is hierarchical. Stars exist singly, in binary pairs, in multiple systems, and in clusters. On large scales, roughly 10^5 light-years, astronomical objects congregate into galaxies. In fact, most stars are in binary systems. The Sun, with its collection of planets, is an exception. Stars in binaries revolve around their common center of mass (which itself moves through space). This star-star gravitational interaction makes possible the measurement of stellar masses and other basic properties. Stellar evolution in some binary systems can lead to spectacularly energetic activity. *See* GALAXY, EXTERNAL; MILKY WAY GALAXY; SOLAR SYSTEM; STAR CLUSTERS.

Types of binary systems. Binary systems are observed in various ways. They may be classified as visual binaries, spectroscopic binaries, or spectroscopic-eclipsing binaries, according to the means of observation.

Visual binaries. In these systems the stars are widely separated compared to their sizes and appear as separate points of light in the telescope. William Herschel first detected the curved path of one star relative to the other, proving their physical association (**Fig. 1**). After years of observation, he found that these motions obey Kepler's law of areas, showing that the law of gravitation applies in the stellar regime. Orbital periods run from years to hundreds of years, and orbits are usually noncircular. Binary stars provide the only proof that has been found so far of gravitation's validity beyond the solar system. (This proof has been extended to general relativity by observations of slow orbital decays in binary pulsars.) *See* CELESTIAL MECHANICS; GRAVITATION; KEPLER'S LAWS.

More detailed observations reveal the apparent orbit of each star relative to the center of mass, leading directly to the stellar mass ratio of the two components. If the binary's parallax (distance) can be measured, individual stellar masses can be found. Without a knowledge of masses, stellar evolution cannot be understood.

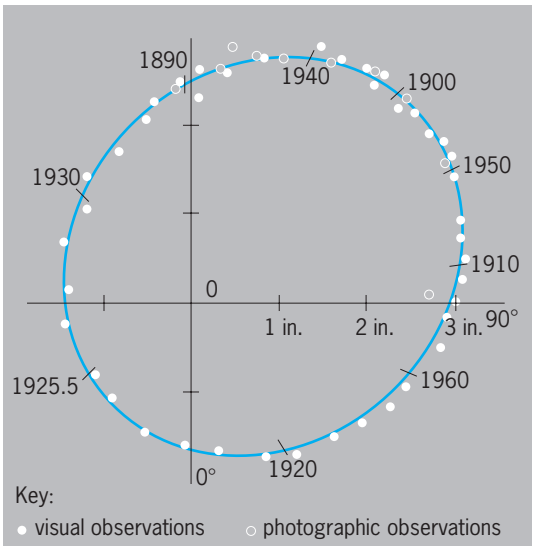

Fig. 1. Apparent orbit of visual binary Krüger 60.

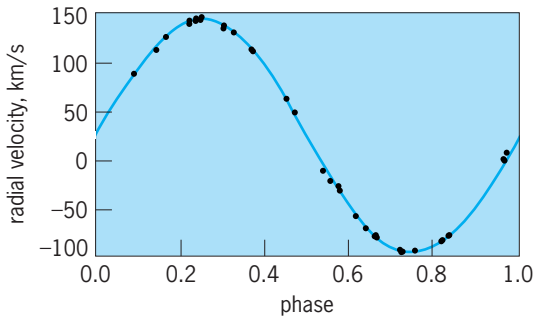

Fig. 2. Radial velocity curve of the loser in the eclipsing binary RX Geminorum. Phases 0.0 and 1.0 correspond to mideclipse of the hotter star. (*E. C. Olson and P. B. Etzel, Mount Laguna Observatory*)

Fewer than 1000 visual binaries have known orbits. Few professional astronomers now make these observations, and amateurs have been encouraged to take up this work. In some cases, orbital periods are so long that binary stars are recognized only by their parallel motions through space.

Atmospheric seeing and telescopic resolution prevent the measurement of visual binary orbits where stars are less than about 0.5 arc-second apart. Interferometric observations have made possible the measurement of more closely spaced binary pairs, opening a new vista on these objects.

Stars in visual binary systems comprise main sequence dwarfs and more evolved giant and supergiant stars. Observations with the Hubble Space Telescope reveal visual binaries composed of brown dwarf pairs, whose masses are too low to ignite hydrogen thermonuclear fusion in their cores. *See* BROWN DWARF; DWARF STAR; GIANT STAR; SUPERGIANT STAR.

Spectroscopic binaries. Large stellar separations in visual binaries result in very low orbital speeds. There is a large class of nonvisual binaries in which the stars are separated by only a few stellar diameters. They cannot be resolved in telescopes, and are referred to as close binaries. The closeness of the stars profoundly affects their evolution. It also creates high orbital speeds (tens to hundreds of kilometers per second), rendering them easily detectable through Doppler shifts of their spectral lines. During the orbital period, which may range from a fraction of a day to hundreds of days, the radial (sight-line) velocity varies. If the orbits are circular, as they often are in close binaries, then the velocity-time curve is sinusoidal (**Fig. 2**).

Radial velocity curve shapes for elliptical orbits are readily calculated. If the two stars are of comparable luminosity, both stars contribute equally to their combined spectra and stellar lines are double over much of the orbital cycle. These are double-line spectroscopic binaries. The inverse ratio of their velocity amplitudes is the stellar mass ratio. Some 1500 spectroscopic binaries are known, though only a small fraction have well-determined velocity curves.

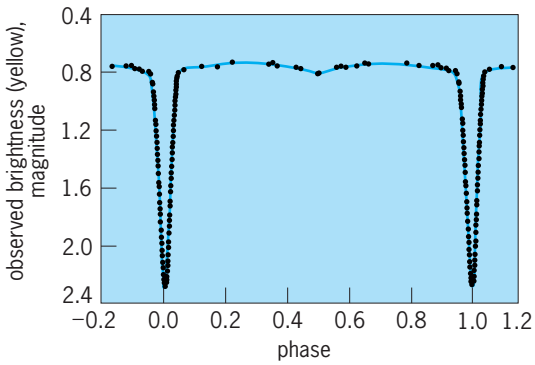

Fig. 3. Yellow light curve of RX Geminorum. Observed brightness is in magnitudes, where an increase of 1.0 magnitude corresponds to a brightness decrease by a factor of 2.5. (*E. C. Olson and P. B. Etzel, Mount Laguna Observatory*)

If the orbit plane is normal to the sight line, there is no radial motion and no Doppler shift. Where shifts are detected, the orbital inclination is still indeterminate. Thus, the information from spectroscopic binaries is limited. *See* ASTRONOMICAL SPECTROSCOPY; DOPPLER EFFECT.

Spectroscopic-eclipsing binaries. Here the orbital plane of a close binary is nearly in the sight line. Mutual stellar eclipses occur, yielding a brightness-time variation called the light curve (**Fig. 3**). The shape of this curve depends strongly on the inclination and relative stellar sizes and surface temperatures. That is, these quantities determine there are eclipses, and the nature of such eclipses (total, partial, or annular). Analysis of the light curve yields the inclination which, coupled with radial velocity curves of both stars, leads to both stellar masses and stellar sizes. (Radial velocities scale the orbital size, as parallaxes do in visual binaries.) Fairly good stellar parameters have been found for a few hundred eclipsing binaries. Interferometric techniques, lunar occultations, and such binary observations are the only source of stellar sizes (apart from the Sun). *See* OCCULTATION.

Modern eclipsing-binary light curves, usually obtained in various colors, reveal higher-order effects. These include mutual gravitational and rotational distortions of the stars, mutual heating of their facing hemispheres, limb darkening of the stellar disks, gravity brightening (a radiative transfer effect that brightens high-gravity parts of the surface), effects of star spots, and orbital eccentricity (though most orbits are nearly circular). Where orbits are noncircular, the major axis of the orbit slowly precesses, and in a few cases provides a critical test of general relativity. *See* ECLIPSING VARIABLE STARS; RELATIVITY.

Stellar evolution. To understand the evolution of binary systems, single stars will first be considered. Stars form from contracting clouds of hydrogen-rich interstellar gas and dust. Contraction releases gravitational energy which heats the forming star. When the core temperature reaches 10^7 K, thermonuclear fusion of helium nuclei from hydrogen nuclei (protons) begins. Fusion occurs only in the stellar core (about 10% of its total mass), and the star stops contracting. Stars spend most of their lives as main sequence dwarfs. Low-mass stars remain longest on the main sequence. The Sun, of 4.5×10^9 years age, is about halfway through this hydrogen-burning phase.

Red giants and white dwarfs. At the end of its main sequence life, the now nearly pure helium core is temporarily out of nuclear fuel. It reverts to gravitational contraction, bringing hydrogen-rich matter inward and heating it. Evolution now accelerates. A thin shell of hydrogen ignites around the contracting core, and the outer layers respond to the new energy source by expanding. A single star becomes a red giant (or supergiant, if massive enough). Its radius increases by 10–100 times and the surface cools (hence, its red color). In some famous visual binaries, the brighter component is a red giant and the fainter a hot main sequence star. Both stars formed together and are the same age. The red giant, of larger mass, reached the end of its main sequence phase first and subsequently expanded to its current state.

Contraction continues to heat the core of a red giant, until helium ignites in the so-called triple-alpha process, to fuse carbon (in repeated explosive flashes if the core has become electron-degenerate). Additional nuclear fusions may occur. The evolving star now has a highly compressed core and a very extended, rarefied envelope. If the star is not too massive, the envelope is eventually driven off to form a planetary nebula, leaving the hot, dense exposed core to cool and become a white dwarf of roughly terrestrial size. *See* PLANETARY NEBULA; WHITE DWARF STAR.

Supernova explosions. Alternatively, stars more massive than about 8 solar masses end their lives in supernova explosions, leaving a neutron-star remnant of about 10-km (6-mi) radius, or a black hole. Neutron stars are radio (or even optical and x-ray) pulsars by virtue of their highly magnetic, rapidly rotating state. The best known is the Crab Nebula pulsar, a remnant of the supernova explosion observed by the Chinese and others in 1054. The pulse interval is the neutron-star spin period. The Crab pulsar rotates about 30 times per second. As single neutron stars radiate photons and particles, they gradually slow down. *See* BLACK HOLE; CRAB NEBULA; NEUTRON STAR; PULSAR; SUPERNOVA.

End products. Thus, following post-main-sequence envelope expansion and loss, and possibly a supernova explosion, stars reach their graves as white dwarfs, neutron stars, or black holes. Electron-degenerate matter in white dwarfs is roughly 10^6 times as dense as water, while neutron-degenerate matter in neutron stars is at least 10^7 times denser, approaching the interior density of an atomic nucleus. White dwarfs and neutron stars have been described as gigantic nuclei.

Evolution of close binaries. Stellar evolution in close binaries is drastically changed by the second star. The system is already a binary in its pre-main-sequence formative stage. Initial conditions determine whether the system will be a visual binary or a close binary. In the latter, the more massive star will

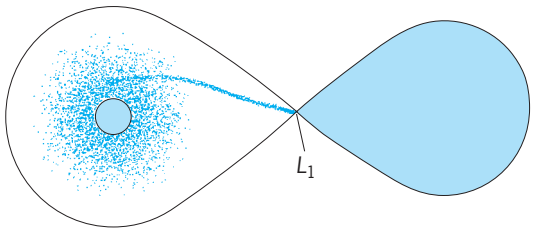

Fig. 4. Roche lobes surrounding the stars in a binary system. The contact point is L_1. The less massive star on the right already fills its lobe. A mass-transferring stream supplies an accretion disk around the other star. (*Matthew Senay*)

exhaust its core hydrogen first and begin its evolution to a red giant. Its expansion is halted when the star's surface reaches the so-called inner lagrangian surface. This is a teardrop-shaped equipotential surface, usually called the Roche lobe, that touches the lagrangian surface of the less massive companion at the inner lagrangian point L_1.

Instead of expanding through this surface, matter flows along the Roche lobe to L_1, evaporates from L_1, and forms a stream that accelerates hypersonically toward the companion (**Fig. 4**). In other words, the post-main-sequence star transfers matter to its companion. One star loses mass (the loser) while the other gains mass (the gainer). These surfaces and motions are determined by the gravitation of the stars and the fictitious centrifugal force in the coordinate frame that rotates with the orbital motion.

A visual binary, in which both stars are far from their Roche lobes, is referred to as detached. A system in which one star has evolved to fill its Roche lobe is called semidetached (Fig. 4). W Ursa Majoris stars are a class of main sequence binaries where both stars fill their lobes and share in effect a common envelope. Their periods are usually less than a day, and they are called contact binaries.

Algol eclipsing binaries illustrate evolution in close binaries. They contain a fairly massive main sequence gainer and a less massive, evolved subgiant (or giant) loser that fills its Roche lobe. Their existence posed the Algol paradox: how could the less massive star be the more evolved? The more evolved star must have originally been the more massive; mass transfer through L_1 to the companion has reversed the mass ratio. Transfer rates in Algols are of order 10^{-7} solar mass per year. Thus, evolution in these binaries is some 100 times more rapid than is solar evolution on the main sequence. RX Geminorum (Figs. 2 and 3) is a long-period Algol.

The Roche lobe prematurely removes the outer envelope of the post-main-sequence loser, precluding the planetary nebula phase of single star evolution. The loser may already have an electron-degenerate core. At some point the gainer will itself evolve from the main sequence. Its envelope expansion may well engulf the original loser, whose orbit will then decay because of friction with the envelope gas (the so-called common-envelope phase). This shortens the orbital period. If conditions are right, the envelope will eventually disperse, and the final product will be a binary white-dwarf system. Such binaries are known. Plausible evolutionary sequences that can explain a large variety of high- and low-mass binary star configurations are now known. *See* STELLAR EVOLUTION; SYMBIOTIC STAR.

Cataclysmic binaries. Stellar evolution with a common-envelope stage may produce cataclysmic variables, in which a white dwarf is paired with a Roche-lobe-filling main-sequence loser. Orbital periods are typically less than a day. The mass-transferring stream from the loser supplies an accretion disk rotating around the white dwarf. Matter in the stream is rapidly accelerated toward the white dwarf, producing significant shock heating on impact with the disk. Disk viscosity, necessary to transfer angular momentum outward to allow matter to fall onto the white dwarf, helps to maintain the disk temperature. Here the energy source is also gravitational.

The cataclysmic binary zoo includes novae, dwarf novae, recurrent novae, and novalike variables. Much of their light comes from the thick accretion disk. Thermal instabilities in the disk may trigger mass dumps onto the white dwarf, producing eruptive brightenings. Amateur astronomers have been enlisted to search for such events. In some cases, hydrogen-rich matter from the loser slowly accumulates on the high-gravity surface of the white dwarf. At a critical density this matter becomes electron-degenerate, igniting a thermonuclear runaway that is a nova outburst. The explosion does not destroy the star, and outbursts may recur after many centuries of quiescence. *See* CATACLYSMIC VARIABLE; NOVA.

High-energy close binaries. The most energetic binary systems are those that contain a neutron star (or rarely, a black hole) remnant. This can be seen from a comparison of the available energy sources.

Stellar energy sources. Stellar evolution taps both gravitational and thermonuclear energy sources. Nuclear sources are usually the most potent available. For example, gravitational energy release accompanying the formation of the Sun supplied its luminosity for a scant tens of millions of years. Core hydrogen burning yields a main sequence solar lifetime of some 10^{10} years. Thus, thermonuclear energy exceeds gravitational energy by about 3000 times in the Sun.

For this discussion a unit is chosen to be the amount of energy released when a kilogram of hydrogen is fused into helium in thermonuclear reactions. A kilogram of matter falling from a large distance onto the Sun's surface releases about 3×10^{-4} unit of gravitational energy. Thus, thermonuclear production is much larger than gravitational. (This ratio is consistent with the time scales just quoted, since only about 10% of the Sun's mass fuses hydrogen.)

The remnants of stellar evolution are not much larger than a solar mass, but are compressed to Earth size in white dwarfs and to a radius of about 10 km (6 mi) in neutron stars. Matter falling onto such compact objects releases much more gravitational energy. For a white dwarf it is about 3×10^{-2} unit per kilogram (nuclear energy still dominates), but for a neutron star it is roughly 50 units. By a wide margin,

gravitation wins the energy contest when neutron stars (or black holes) are involved.

Single neutron stars. Single neutron stars are rapidly rotating and highly magnetic, with typical polar fields of 10^8 tesla. (The field has been amplified by the collapse that formed the neutron star.) Spin periods run from fractions of a second to a few seconds, and slowly increase as rotational energy is lost by radiation. Most of the radiation originates near the magnetic poles. Magnetic and rotational axes are not quite aligned (the case even with the Earth). The magnetic pole facing the Earth is visible for only half the rotational period, producing the lighthouselike pulsing of the observed radiation (which is usually at radio wavelengths). This radiation is further beamed by the radiation mechanism.

X-ray binaries. Neutron stars in close binary systems, where the companion fills its Roche lobe and transfers mass, can tap the enormous gravitation energy of in-falling matter. The strong magnetic field interferes with the formation of an accretion disk and funnels transferred mass onto the magnetic poles of the neutron star. This matter, enormously accelerated, strikes the surface at about half the speed of light and heats the poles to 10^8 K. Thermal radiation at such enormous temperatures (hotter than the solar core) is mainly x-rays. An x-ray binary results. Her X-1 is a well-known x-ray binary. Its pulse period is about 1 second, and the binary orbital period is just under 2 days.

High- and low-mass x-ray binaries are now recognized. The former are born with stars of comparable mass, totaling more than 20 solar masses. The more massive component explodes as a supernova and leaves a neutron star (or a black hole) remnant, but the binary is not disrupted. As the second star evolves and expands, it feeds mass to the neutron star. The resulting x-radiation may reach 10^5 solar luminosities, making these high-mass binaries among the most luminous objects in the Milky Way Galaxy.

Low-mass x-ray binaries form with stars of differing mass (typically 10 and less than 2 solar masses). The more massive star suffers a supernova explosion and leaves a neutron star remnant. Mass accreted from the other star may spin up the neutron star to make a so-called millisecond pulsar. The enormous radiation from the neutron star may literally strip the companion to its degenerate helium core, as seems to be happening in the Black Widow Pulsar.

If the unseen x-ray source in a close binary has a mass significantly larger than about 3 solar masses, it cannot exist as a neutron star. According to current theory, it must have collapsed into a stellar-mass black hole. A handful of such candidates exist in the Milky Way Galaxy.

Type Ia supernovae. When the evolution of a moderate-mass binary results in two carbon-oxygen white dwarfs orbiting each other inside a common envelope, a remarkable event may occur. Viscous friction with the envelope causes the stars to spiral inward. If they merge into a total mass larger the about 1.4 solar masses (the Chandrasekhar limit), extremely intense thermonuclear reactions produce iron-peak elements, the structure abruptly runs out of thermonuclear energy, and an implosion-explosion occurs, producing a type Ia supernova. These highly luminous objects are seen in distant galaxies. They behave as standard candles (distance indicators), and are used to study the long-term evolution of the expansion rate of the universe. *See* COSMOLOGY; NUCLEOSYNTHESIS; UNIVERSE.

Gamma-ray bursts. The most energetic events in the universe are gamma-ray bursts, now suspected to occur in distant galaxies. Their origin is currently unknown, but one hypothesis suggests the merger of a pair of neutron stars to form a black hole. There are several binary pulsars (containing pairs of neutron stars) in the Milky Way Galaxy. Relativistic effects cause them to radiate their orbital energy, leading to possible coalescence. A gamma-ray burst in the Milky Way Galaxy could pose a serious hazard to life on Earth.

Frequency and origin of multiple systems. Observations show that the ratios of single:binary:triple:quadruple systems are about 58:35:6:1. That is, a randomly selected stellar system has about a 42% chance of being nonsingle, and most stars are in binary or multiple systems. The fraction of binaries may be even higher among forming, gravitationally contracting stars (known as T Tauri stars). Theoretical calculations suggest that fragmentation of contracting, rotating clouds may produce binary and multiple stars, if cloud rotation is sufficiently rapid. *See* PROTOSTAR; T TAURI STAR.

Extrasolar planetary systems. More than a dozen such systems are known. While not binary stars, most have been found using radial velocity observations like those used for spectroscopic binaries. However, while radial velocities of kilometers per second are seen in binary stars, the detection of planetary systems requires velocity measurements of meters per second.

Edward C. Olson

Bibliography. A. H. Batten, *Binary and Multiple Systems of Stars*, Pergamon Press, Oxford, 1973; I. Iben and A. Tutukov, The lives of binary stars, *Sky Telesc.*, 95(1):42–48, January 1998; F. H. Shu, *The Physical Universe*, University Science Books, Mill Valley, CA, 1981.

Binaural sound system

A sound-reproducing system in which sound is recorded or transmitted by using two microphones mounted at the ears of a dummy human head. To preserve the binaural effect, the sound must be monitored by a listener wearing a set of earphones identically spaced. In an ideal binaural transmission system, both the amplitude and phase of the sound waves incident on the dummy's ears are duplicated at the listener's ears.

Binaural systems have been made so perfect that the listener is unable to distinguish the monitored sound from the real sound. For example, when a

person walks around the dummy head, the listener, upon hearing the footsteps, has the compelling illusion of someone walking around him. No other sound system thus far devised can even approximate such an effect. *See* SOUND-REPRODUCING SYSTEMS; STEREOPHONIC SOUND; VIRTUAL ACOUSTICS.

Harry F. Olson

Bibliography. G. Alkin, *Sound Recording and Reproduction*, 3d ed., 1997; J. Eargle, *Handbook of Sound System Design*, 1989; F. A. Everst and R. Streicher, *The New Stereo Soundbook*, 1992.

Binoculars

Optical instruments consisting of a parallel pair of matched telescopes, used to extend the range of stereoscopic vision to far distances. Each half of a binocular consists of a telescope objective, a prism for inverting the image, and an eyepiece (see **illus.**). *See* EYEPIECE; STEREOSCOPY; VISION.

Characteristics. The characteristics of a binocular are stated by using a pair of numbers, such as 7×50 (seven by fifty), where the first number is the magnifying power of the binocular and the second is the diameter of the objective of the binocular in millimeters. The apparent field of view which the user sees is given by multiplying the angle in object space by the magnifying power. For example, an object appearing to subtend 5° by the unaided eye will appear to be enlarged to about 35° when viewed through the binocular. *See* MAGNIFICATION.

Image inversion. Since a lens forms an inverted image, the straight-through combination of an objective and eyepiece would provide an inverted field of view to the eye, as in an astronomical telescope. Almost all binoculars use prisms with an odd number of reflecting surfaces to invert the image correctly. The choice of prisms must also provide for the adjustment in pupil separation by having the optical axes on either side of the prism displaced but parallel. The most frequently used prism is a Porro prism, which is a combination of two 45°–90° prisms. These lead to a bulky mechanical construction, so that many modern binoculars use straight-through prisms. *See* MIRROR OPTICS; OPTICAL PRISM.

Modern prism binocular. (*Bausch and Lomb Optical Co.*)

Focusing and variable magnification. The magnification produced in binoculars reduces the apparent depth of focus because of the enlarged image. Therefore, binoculars require some ability to change the separation between the eyepiece and the objective to provide focusing to accommodate different object distances and possible refractive errors in the eye. Most binoculars provide for a joint focus adjustment of both tubes with one eye having an additional focus range to compensate for users who have differing refractive errors in each eye. The ability to adjust for focus becomes more important with the age of the user, since the eye's accommodative range is diminished as one grows older.

Another optical feature often available in binoculars is a variable magnification or zoom system. This can be accomplished by moving elements in either the eyepiece or the objective, although the smaller elements of the eyepiece are usually chosen for practical reasons. *See* ZOOM LENS.

Selection. Selection of binoculars should be made with some consideration of the intended use. A larger objective will permit use of the binoculars at lower light levels. However, binoculars with larger-diameter objectives and higher powers are heavier and less convenient to use. The jitter produced while holding the binoculars will be magnified by the binocular, so that very high power binoculars usually require a stable support, such as a tripod, to be used properly. A modest power such as 6 is usually more comfortable than a high power such as 10. The wider the field, the more complex and heavier the binocular, as well as the more costly. In addition, the eye relief, the distance that the eye may be placed behind the rear surface of the eyepiece, is important to user comfort. Since the full field of the binocular can be viewed only if the eye location coincides with this exit relief distance, long eye relief is important to users who wear corrective lenses.

The mechanical characteristics of binoculars are also of importance. Errors in alignment of the optical axes of the two halves of the binocular cause visual fatigue when the instrument is used for long periods. The instrument should be rugged, and the images through the eyepieces of a distant object should properly overlap, and not be rotated with respect to each other. There are, necessarily, some aberrations and distortion present in binoculars, especially compact and wide-field binoculars. A reasonable portion of the field of view should be free of obvious blurring or color fringes, and the distortion should be similar in each half of the binocular. *See* ABERRATION (OPTICS).

Opera glasses. Opera glasses are a type of low-power binoculars which use simpler optics. The use of a negative lens as an eyepiece, as in a Galilean telescope, limits the power and the field of view but permits a lighter and less expensive instrument. *See* TELESCOPE.

Robert R. Shannon

Bibliography. H. Paul, *Binoculars and All Purpose Telescopes*, 1980; L. J. Robinson, *Outdoor Optics*, 1989.

Binomial theorem

One of the most important algebraic identities, with many applications in a variety of fields. The binomial theorem, discovered by Isaac Newton, expresses the result of multiplying a binomial by itself any number of times, as given in Eq. (1), where the coefficients c_1,

$$(a+b)^n = a^n + c_1 a^{n-1} b + c_2 a^{n-2} b^2 + c_3 a^{n-3} b^3 + \cdots + c_r a^{n-r} b^r + \cdots + b^n \quad (1)$$

$c_2, c_3, \ldots, c_r, \ldots$, are given by Eqs. (2). The standard notation for these coefficients is given by Eqs. (3).

$$\begin{aligned} c_1 &= \frac{n}{1} \\ c_2 &= \frac{n(n-1)}{1 \cdot 2} \\ c_3 &= \frac{n(n-1)(n-2)}{1 \cdot 2 \cdot 3} \\ c_r &= \frac{n(n-1)(n-2)\cdots(n-r+1)}{1 \cdot 2 \cdot 3 \cdots r} \end{aligned} \quad (2)$$

$$c_1 = \binom{n}{1}, c_2 = \binom{n}{2}, \ldots, c_r = \binom{n}{r} \quad (3)$$

In this notation

$$\binom{n}{r}$$

is the coefficient of the term containing b^r in the expansion of $(a+b)^n$. It is a fraction with the numerator and denominator each containing r factors; those in the denominator begin with 1 and increase by 1; those in the numerator begin with n and decrease by 1. It is easily shown that Eq. (4) is valid.

$$\binom{n}{r} = \binom{n}{n-r} \quad (4)$$

Some examples of the binomial theorem are as follows:

1. $\binom{17}{3} = \frac{17 \cdot 16 \cdot 15}{1 \cdot 2 \cdot 3} = 680$

2. $\binom{12}{5} = \frac{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 792$

3. $$\begin{aligned} (a+b)^6 &= a^6 + \binom{6}{1} a^5 b + \binom{6}{2} a^4 b^2 \\ &\quad + \binom{6}{3} a^3 b^3 + \binom{6}{4} a^2 b^4 + \binom{6}{5} a b^5 + b^6 \\ &= a^6 + 6a^5 b + 15a^4 b^2 + 20a^3 b^3 + 15a^2 b^4 \\ &\quad + 6ab^5 + b^6 \end{aligned}$$

4. $$\begin{aligned} (2x-3y)^4 &= [(2x) + (-3y)]^4 \\ &= (2x)^4 + \binom{4}{1}(2x)^3(-3y) + \binom{4}{2}(2x)^2(-3y)^2 \\ &\quad + \binom{4}{3}(2x)(-3y)^3 + (-3y)^4 \\ &= 16x^4 - 96x^3 y + 216x^2 y^2 - 216xy^3 + 81y^4 \end{aligned}$$

Under suitable conditions the binomial formula is valid when n is not a positive integer. In this case the formula does not terminate, but generates an infinite series.

Much of the utility of the binomial theorem stems from the properties of the coefficients. In particular, the coefficient

$$\binom{n}{r}$$

gives the number of combinations of n distinct objects taken r at a time. The set of coefficients for any value of n form a distribution that has fundamental importance in the study of probability and statistics. *See* ALGEBRA; COMBINATORIAL THEORY; DISTRIBUTION (PROBABILITY); SERIES.

Hollis R. Cooley

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

3 BIO-CHA

www.MHEST.com

B

Bioacoustics, animal — Bytownite

Bioacoustics, animal

The study of the role of sounds in the life history of animals. Almost all animals, from insects to large mammals, depend on sounds for their well-being and survival. Sounds are used by animals for a wide variety of reasons such as advertisement for mating, territory demarcation, group cohesion, prey location, infant-mother identification and contact, predator detection, alarm warning, and social communication. The field of animal bioacoustics can be subdivided into the acoustics of terrestrial animals and aquatic animals. Each field of study can be subdivided into (1) auditory capabilities and processes, (2) characteristics of sound emissions and mechanisms of sound production, and (3) the function and meaning of specific vocalizations.

Airborne and underwater acoustics. Airborne acoustics, associated with terrestrial animals, and underwater acoustics, associated with aquatic animals, have three primary differences. First, the amount of acoustic energy that is absorbed and transformed into heat is much higher in air than underwater. Acoustic transmission loss in both media has a geometric spreading loss component plus an absorption loss component. The geometric spreading loss is dependent on the propagation geometry and not the medium, whereas absorption loss is directly dependent on the medium. Absorption loss in both media increases with frequency, although it is much higher in air than underwater (**Fig. 1**). Therefore, the acoustic range of most aquatic animals is considerably larger than for terrestrial animals. Animals interested in long-range communications naturally use lower-frequency sounds. *See* SOUND ABSORPTION.

Second, airborne acoustics and underwater acoustics differ greatly in the respective values of the acoustic impedance, ρc, where ρ is the density of the medium and c is the sound velocity in that medium. The density of water is approximately 1000 times greater than that of air, and the sound velocity in water is about 4.5 times greater than in air. Therefore, the acoustic impedance of water is approximately 3600 times greater than air. This difference in acoustic impedance has great ramifications on how sounds are produced and received in both media. For example, the middle-ear ossicular chain of terrestrial mammals serves as an impedance-matching transformer between the air in the external auditory meatus and the fluid in the cochlea, or inner ear. Such a chain is not needed for animals hearing underwater since the impedance of water is already matched to the fluid of the inner ear. The impedance difference issue becomes rather complex with regard to hearing and sound production in amphibians and pinnipeds, which must have good in-air and underwater hearing. *See* ACOUSTIC IMPEDANCE; AMPHIBIA; PINNIPEDS.

Third, between airborne acoustics and underwater acoustics there is a large pressure difference. Pressure underwater increases by 1 atmosphere for approximately every 10-m (34-ft) increase in depth. As animals in an underwater environment swim deeper, the pressure they experience will be considerably higher than at the surface and will cause the air in body cavities to compress and increase in density, which in turn affects the production and reception of

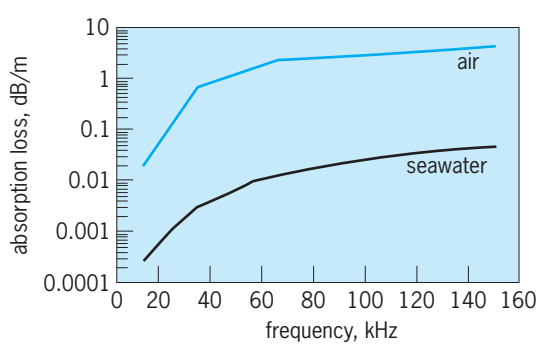

Fig. 1. Acoustic absorption loss in air and seawater as a function of frequency.

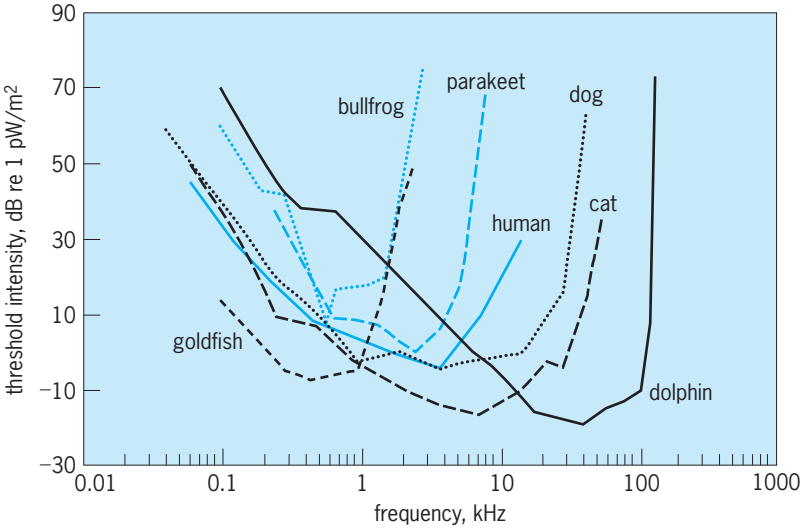

Fig. 2. Audiogram of different animals.

sounds. *See* ATMOSPHERIC ACOUSTICS; HYDROSTATICS; SOUND; UNDERWATER SOUND.

Auditory capabilities and processes. The structure and function of the ears of terrestrial mammals, including humans, are similar in nature; and although the ears of marine mammals are homologous to those of terrestrial mammals, there can be some distinct differences. For example, cetaceans do not have pinnas (external ears), and the auditory meatus in dolphins is but a fibrous tissue which is not capable of conducting sounds. Although there is a middle-ear ossicular chain, there is no direct connection between the tympanic membrane and the malleus. Sound enters into the dolphin's auditory system via the lower jaw and is probably conducted via fat channels directly into the cochlea. *See* CETACEA; EAR (VERTEBRATE).

Birds have a feather-covered auditory meatus and no pinnas. The feathers seem to minimize air turbulence and the subsequent flow noise associated with turbulent flow. The middle ear consists of a single major ossicle, the columella or stapes. The cochlea is elongated rather than curled in a shell. *See* ACOUSTIC NOISE; AVES.

Of the amphibians, the anurans, which include frogs and toads, have been studied the most. The outer and middle ears of anurans consist of a large tympanic membrane that is flush with the surface skin, and a columella which transmits sounds to the inner ear. The inner ear includes the basilar papilla and the amphibian papilla, operating independently in overlapping frequency ranges, with the amphibian papilla responding to sound from about 100 Hz to 1 kHz and the basilar papilla responding to frequencies of several hundred hertz to several kilohertz. The outer ear and middle ear of reptiles are similar to that of anurans except that the tympanic membrane may be recessed below the skin. *See* ANURA; REPTILIA.

The bodies of fishes are closely matched in impedance to water so that sounds can propagate through the body to the three otolith organs—the sacculus, lagena, and utricle—which contain hair cells that are excited by acoustic energy. In most fishes, the sacculus is the primary auditory organ; however, in some fishes the utricle is the primary receptor of sounds. Most fishes have swim bladders that can be compressed and expanded when impinged upon by an acoustic wave, and the subsequent motions are coupled to the auditory organs. Fishes also possess lateral lines which can sense particle motions caused by low-frequency acoustic waves in the water. *See* OSTEICHTHYES; PISCES (ZOOLOGY); SWIM BLADDER.

The structure and location of ears of insects tend to vary considerably from species to species. In some, the ears are located on different appendages, while in other the ears may be located on the abdomen or on other parts of the body. *See* INSECTA.

The hearing sensitivity of animals is determined by careful laboratory psychophysical experiments in which the subjects are trained to respond to an acoustic stimulus of varying frequency and amplitude. The audiograms of some representative species of terrestrial mammals are shown in **Fig. 2**. The human audiogram is included to serve as a reference. Intensity rather than pressure of the acoustic stimulus must be used to compare the hearing sensitivity of terrestrial and aquatic animals, because of the difference in the acoustic impedances of the respective media. A unit of acoustic pressure conveys different amounts of energy in the two media. Acoustic intensity for a plane acoustic wave is defined as $p^2/\rho c$, where p is the sound pressure. Allowance must also be made for the difference in reference units, 20 micropascals for airborne acoustics and 1 μPa for underwater acoustics. Dolphins have the highest upper-frequency limit of hearing and have the most acute hearing sensitivity of all animals. Fishes tend to have a limited frequency range of hearing. The best sensitivities of the species represented in Fig. 2, except for the bullfrog, are within 20 dB of each other. *See* AUDIOMETRY; HEARING (HUMAN); PHONORECEPTION; PSYCHOACOUSTICS; PSYCHOPHYSICAL METHODS.

Sound emissions and production. Animals use a wide variety of different sounds in different frequency ranges and for different purposes. Sounds can be very short events lasting less than 100 microseconds for some echolocating dolphins or can be very long duration events lasting several hours for singing humpback whales. The sounds can be infrasonic, with frequencies below the human hearing capability, or ultrasonic, with frequencies above the human hearing range. Most terrestrial animals, including mammals, birds, amphibians, and reptiles, produce sounds that are within the human frequency range, or sonic sounds. Many aquatic animals, including fishes and marine mammals, also produce sounds that are in the sonic range; however, some marine mammals also produce sounds that can extend either beyond (ultrasonic) or below (infrasonic) the human frequency range. Dolphins can produce very high

frequency echolocation sounds having frequency components up to 200 kHz. However, the animal that probably produces the highest-frequency sound is the snapping shrimp *Synalpheus parneomeris* (**Fig. 3**). The difference between the peak amplitude and the minimum amplitude in the spectrum of a snap is only 20 dB, an indication of a very broadband signal. *See* DECAPODA (CRUSTACEA); INFRASOUND; ULTRASONICS.

The mechanism of sound production in most animals is well known. Mammals and birds typically use their lungs as a source of air that is forced through a small orifice such as a larynx in mammals or a syrinx in birds. Amphibians such as frogs also use their lungs to force air through a larynx; however, the acoustic energy is coupled to an expandable air sac in the throats, which resonates to amplify sounds. For some frogs, the inflatable sac also strikes the substrate, causing seismic vibrations. Insects and crustaceans strike or rub certain appendages against other parts of their bodies. Some fishes produce sounds by inflating and compressing their swim bladder, causing vibrations in the water, while others may use the swim bladder as a resonator of sounds produced by stridulation between bony parts of their bodies.

The sound production mechanism of the odontocete, or toothed, whales baffled researchers for many years. In 1997, T. Cranford demonstrated that dolphins produce sounds with a pair of phonic lips previously called the monkey lips–dorsal bursae complex that are embedded in the nasal system. The simultaneous manipulation of these phonic lips and the production of echolocation clicks and whistles have been documented. Sperm whales probably use a similar set of phonic lips of the museau de singe, which are located in the forward portion of the animal's forehead. The exact mechanism of sound production by baleen whales is still a mystery.

Function and meaning of vocalizations. Careful studies in the field are often required in order to determine the function and meaning of specific vocalizations to conspecifics and to other species. Birds sing to mark territory and to attract potential mating partners. Some nonhuman primates use specific sounds as alarm calls indicating whether the threat is a predator bird or a mammal. Male frogs use their sound to attract females for mating. Dolphins and bats emit ultrasonic echolocation signals to detect and localize prey. Pinniped and dolphin mother-calf pairs use specific acoustic signals for identification purposes. Elephants use their infrasonic calls to maintain social contact and to coordinate family group movements. As sunset approaches in the savannahs of Africa, there is often a temperature inversion in the air mass where a layer of warm air is trapped by cooler denser air above. This inversion layer forms an acoustic duct to trap sounds and allows for good propagation of infrasonic sounds over tens of kilometers, making it possible for elephants in a family group to coordinate their movements over large distances. *See* ECHOLOCATION.

The meaning of many animal sounds still escapes human understanding. The infrasonic sounds of blue whales can propagate over hundreds of kilometers if trapped in a surface duct or in the deep sound channel of the ocean. Unfortunately, the aquatic environment has made it very difficult to study the behavior of blue whales and to determine how they respond to different sounds. Certain dolphins such as spinner and spotted dolphins often are found in groupings of up to several hundred animals. These dolphins may be swimming in the same direction but are spread out over hundreds of meters. Yet, some of these schools have been observed to suddenly change their course abruptly as a response to some kind of signal, probably acoustic in nature. However, the difficulty in conducting research in the ocean has made it next to impossible to determine the specific acoustic signals used in group coordination. Humpback whales have been known to "sing" for hours on end, yet there is no consensus on the role of the songs in their natural history. *See* ANIMAL COMMUNICATION.

Whitlow W. L. Au

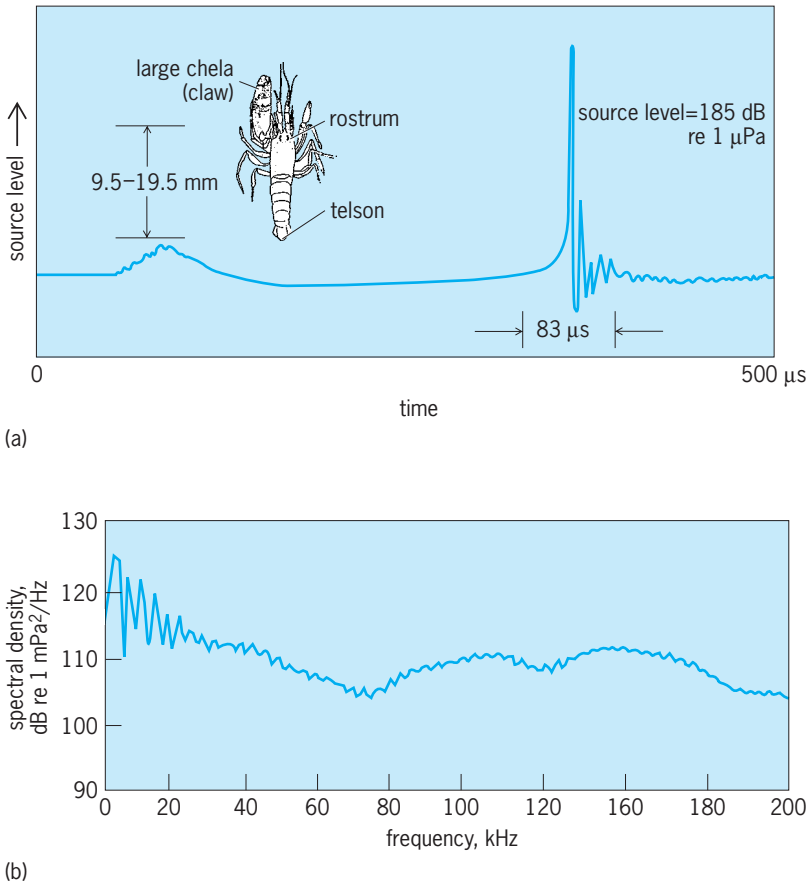

Fig. 3. Single snap of a snapping shrimp, *Synalpheus parneomeris* (shown in inset). (*a*) Waveform. Source level is sound pressure level 1 m (3.3 ft) from the source. (*b*) Spectrum.

Bibliography. W. W. L. Au, *The Sonar of Dolphins*, Springer-Verlag, New York, 1993; J. W. Bradbury and S. L. Vehrencamp, *Principles of Animal Communications*, Sinauer Associates, Sunderland, MA, 1998; R. R. Fay, *Hearing in Vertebrates: A Psychophysics Databook*, Hill-Fay Associates, Winnetka, IL, 1988.

Bioarcheology

The study of skeletal remains from archeological sites by biological (or physical) anthropologists. Bioarcheology differs in several ways from traditional skeletal research. Previous work focused on individual case studies (for example, individuals with identifiable diseases) or on typological analyses of cranial form, the object of which was to classify collections into racial or ethnic groups. Bioarcheology looks at populations rather than individuals, often highlighting variation within groups as much as differences between them. In addition, it considers the interaction of biology with human culture and behavior, and the effects of the latter upon skeletal morphology or form. Technological advances in computers and methodology have opened up new fields of study, such as biomechanics and paleonutrition, while revolutionizing older interests, such as biological distance studies (a measure of genetic relatedness). Finally, bioarcheology, in part because of the specialized nature of some subfields and in part because of its biocultural approach, emphasizes collaboration with other anthropologists as well as researchers in other disciplines. *See* ANTHROPOLOGY; ARCHEOLOGY.

Health. The field of bioarcheology is built in large part upon the traditional study of human disease in prehistoric remains, or paleopathology. Bioarcheologists are more interested in the effect of disease on populations, and interrelationships between disease and social systems. Only a few infectious diseases leave diagnostic lesions on the skeleton (including tuberculosis, treponemal diseases such as syphilis, and leprosy). A good deal of attention has centered on tracing the evolution of these diseases: leprosy in the Old World, syphilis (arguably) in the New World, and tuberculosis, which appears in both hemispheres. More common than these specific infections in skeletal collections are generalized indicators of health problems: osteitis or periostitis (bone infections due to a variety of agents), growth disruptions in tooth enamel (enamel hypoplasias) or bone (Harris lines), reduced growth rates (as seen in shorter bones in children than are expected for their age), and porotic hyperostosis and cribra orbitalia (bone porosity on the surface of the skull and the upper surface of the eye orbits, respectively, the result of anemia during childhood).

Infectious diseases increase in frequency in agricultural communities due to a variety of factors, including population growth, increasing sedentism, and an expansion in trade contacts with other societies. However, it would be an oversimplification to state that health always deteriorates with the advent of agriculture; for example, stored agricultural crops may alleviate seasonal fluctuations in nutritional availability. *See* DISEASE; EPIDEMIOLOGY; PATHOLOGY.

Paleonutrition. The study of prehistoric diet has been revolutionized by work on bone chemistry and isotopic variation. Essentially, this research focuses on the fact that some dietary ingredients leave a chemical or isotopic trace in bones; for example, carnivores retain less strontium than do herbivores, so that strontium can indicate the amount of meat in the diet. Carbon isotope analyses of bone also indicate when maize (corn), the main component of historical Native American diets, was introduced into North America. Such studies have also uncovered differences in the level of reliance on maize within populations based on sex and status, and between societies, which in turn have been tied to differences in general health. Other dietary components, such as marine resources, may also be determined utilizing nitrogen isotopes.

Injuries. Traumatic injuries and osteoarthritis (arthritis in the bony sections of joints) are extremely common in prehistoric populations. Although it is difficult to separate accidental trauma from that caused by interpersonal violence, bioarcheologists can identify examples of violent death in prehistory (for example, arrowpoints lodged in bone, around which there is no healing). These studies indicate that violence escalated in late prehistory in North America as societies became increasingly agricultural and population size rose. Osteoarthritis is common throughout prehistory and may be tied in part to levels of activities, although it is also caused by other factors, including injuries and infections of the joints. *See* ARTHRITIS.

Biomechanics. An area of research that has only recently become feasible through computer-aided technology is biomechanics. Like muscles, bones respond to higher-than-normal physical activities by increasing in mass, and can therefore indicate the usual level of physical exertion during life. Biomechanical studies have shown that physical activities changed as populations adopted agriculture, although the nature of the change varies in different regions of North America. In the Midwest and Deep South of the United States, activities increased when agriculture was adopted, especially in females, while the opposite is true in coastal Georgia. This variability may be linked with differences in the intensity of agricultural practices, but other activities may also be important. *See* BIOMECHANICS.

Biodistance. With more sophisticated computer technology, biodistance studies, which seek to establish genetic relationships between populations, are increasingly complex. To some extent, this means an elaboration of intricate typologies. However, biodistance studies can be used productively to answer questions about the evolution of regional societies, as well as to illuminate such practices as residence patterns.

Summary. New methods, technologies, and theoretical questions have transformed the fields of skeletal biology and paleopathology into modern bioarcheology. In spite of controversy concerning the proper treatment of human skeletal remains, bioarcheological research remains strong today. More information is contained in a single human burial than in any other archeological feature, and skeletons are the only source of data that focuses directly on individual humans. Burial information

provides a long-term perspective on changes in the human condition: for instance, the evolution of diseases and the effects of societal changes such as the adoption of agriculture. Therefore, human skeletal remains are a critical part of archeological research. With the innovations mentioned above, bioarcheological studies will continue to play a vital role in the interpretation of past societies. *See* BONE; PHYSICAL ANTHROPOLOGY.

Patricia S. Bridges; Clark Spencer Larsen

Bibliography. R. L. Blakely (ed.), *Biocultural Adaptation in Prehistoric America*, 1977; M. N. Cohen and G. J. Armelagos (eds.), *Paleopathology at the Origins of Agriculture*, 1984; C. S. Larsen, *Bioarchaeology: Interpreting Behavior from the Human Skeleton*, 1997; M. L. Powell, P. S. Bridges, and A. M. W. Mires (eds.), *What Mean These Bones?: Studies in Southeastern Bioarchaeology*, 1989; R. H. Steckel and J. C. Rose (eds.), *The Backbone of History: Health and Nutrition in the Western Hemisphere*, 2002.

Bioassay

A method for the quantitation of the effects on a biological system by its exposure to a substance, as well as the quantitation of the concentration of a substance by some observable effect on a biological system. The biological material in which the effect is measured can range from subcellular components and microorganisms to groups of animals. The substance can be stimulatory, such as an ion increasing taxis behavior in certain protozoans, or inhibitory, such as an antibiotic for bacterial growth. Bioassays are most frequently used when there is a number of steps, usually poorly understood, between the substance and the behavior observed, or when the substance is a complex mixture of materials and it is not clear what the active components are. Bioassays can be replaced, in time, by either a more direct measure of concentration of the active principle, such as an analytical method (for example, mass spectrometry, high-pressure liquid chromatography, radioimmunoassay), or a more direct measurement of the effect, such as binding to a surface receptor in the case of many drugs, as the substance or its mechanism of action is better characterized.

Assays to quantitate the effects of an exposure model the effect of a substance in the real world. Complex biological responses can be estimated by laboratory culture tests, which use, for example, bacteria or cells cultured in a petri dish (usually to model an effect either on the organism of interest, such as bacteria, or on some basic cellular function); by tissue or organ culture, which isolates pieces of tissue or whole organs in a petri dish (usually to model organ function); or in whole animals (usually to model complex organismic relationships). Tests which measure direct effects on an organism, for example, inhibition of mold growth by an antimold agent in a petri dish, are frequently easier to extrapolate to real-world situations (although they are still only laboratory simulations) than are more indirect tests, such as measurement of tumor induction in rats under restricted conditions as a model for tumor induction in people. However, each assay has advantages and disadvantages, including factors such as ease, cost, and feasibility which have to be weighted in using the test. The usefulness of an assay usually improves as the relationship of concentration and effect, that is, the dose-response curve, is characterized. Although this curve can be complex, many systems have a linear range—doses between which the biological effect is a constant times the dose—generally the most useful range for quantitation.

Assays to estimate concentrations of substances usually require a simple, reproducible end point which can be quantitated easily. The more complex the biological end point measured, the greater the number of factors that can modulate the expression of the phenomenon and, therefore, the greater the amount of control and cost necessary to have a reproducible measure. Usually this type of assay is used when the dose is in the linear range, and conditions are frequently adjusted to make the response a linear one.

Ronald W. Hart; Angelo Turturro

Microorganism Assay

Use of animals for bioassay antedated use of microorganisms. However, the expansion of knowledge of microbial nutrition which occurred during the period 1936–1945 stimulated the development of microbiological assay methods as a routine tool for determining vitamins and amino acids in natural materials. These methods possessed the required precision, and were found also to be rapid, convenient, and remarkably sensitive for detecting the presence and for following the isolation of new vitamins or vitaminlike materials present in trace amounts in crude materials. Many of these substances (nicotinic acid, pantothenic acid, pantetheine, inositol, biotin, pyridoxal, pyridoxamine) were found only subsequently to be important in animal nutrition. The development of microbiological assay methods for folic acid and vitamin B_{12} was of great aid in permitting isolation of these vitamins as well. In some cases microbiological assay remains today the only rapid and specific assay available for determination of individual vitamins.

Microbiological assay has several advantages over animal assay. It requires much less time, labor, and materials to perform. It is consequently less expensive than animal assay, and in most instances is much more precise. One disadvantage of microorganisms as compared with animals as assay tools is that extraction of the vitamin from the combined forms in which most of these naturally occur is usually required to make these available to the microorganism. It is generally assumed (sometimes on insufficient grounds) that animals utilize such combined forms completely. This disadvantage of the microbiological assay as an analytical tool has proved to be an asset in that it has permitted detection, isolation, and quantitative determination of several hitherto unknown

TABLE 1. Representative methods of assay for the water-soluble vitamins

Vitamin	Test organism
p-Aminobenzoic acid	*Neurospora crassa* (mutant)
Biotin	*Lactobacillus arabinosus*
	Saccharomyces cerevisiae
Choline	*Neurospora crassa* (mutant)
Folic acid	*Lactobacillus casei*
	Streptococcus faecalis
Inositol	*Saccharomyces carlsbergensis*
Lipoic acid	*Lactobacillus lactis*
Nicotinic acid	*Lactobacillus arabinosus*
Pantothenic acid	*Lactobacillus arabinosus*
Riboflavin	*Lactobacillus casei*
Thiamin	*Lactobacillus fermenti*
	Saccharomyces cerevisiae
Vitamin B_6	
Pyridoxine + pyridoxamine + pyridoxal	*Saccharomyces carlsbergensis*
Pyridoxamine + pyridoxal	*Streptococcus faecalis*
Pyridoxal	*Lactobacillus casei*
Vitamin B_{12}	*Lactobacillus leischmannii*

combined forms of the vitamins, for example, pantotheine, folinic acid, and pyridoxamine phosphate. *See* VITAMIN.

Vitamin assay. Microorganisms are known which require one or more of all of the water-soluble vitamins except ascorbic acid. Various yeasts and lactic acid bacteria have been most widely used for assay of these substances. The latter group includes species in the genera *Lactobacillus, Streptococcus*, and *Pediococcus* (or *Leuconostoc*), and has been used more widely than any other organisms for assay of both vitamins and amino acids. Although many modifications of the originally devised methods have been published, the methodology is similar in all cases. A single procedure for the vitamin nicotinic acid (niacin) describes the general method. Suitable methods for other vitamins are referred to in **Table 1**.

1. For extraction of nicotinic acid, an accurately known weight for finely ground or homogenized sample is autoclaved for 15 min in sufficient 1 *N* H_2SO_4 to give an estimated nicotinic acid content of 1 microgram per milliliter of extract. The cooled mixture is adjusted to pH 6.8, diluted to an estimated nicotinic acid content of 0.1–0.2 μg/ml, and filtered.

2. The assay organism, *Lactobacillus arabinosus* 8014, is maintained by monthly transfer of stab cultures on a medium containing 1% yeast extract, 0.25% glucose, and 1.5% agar.

3. See **Table 2** for the basal medium.

4. The inoculum is prepared the day before the assay is run. A transfer is made from a stab of the stock culture into a sterile 10-ml aliquot of basal medium (single strength) supplemented with 1 μg of nicotinic acid per tube. This culture is incubated for 16–24 h at 98.6°F (37°C). The cells are centrifuged, washed with sterile water, and resuspended in 10 ml of sterile water or saline. One drop of this suspension is used to inoculate each assay tube.

5. The assay is carried out in lipless 18- by 150-mm test tubes individually supported in a metal rack. To one series of tubes, the freshly diluted standard nicotinic acid solution (0.1 μ/ml) is added as follows with duplicate tubes at each level: 0, 0.25, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, and 5.0 ml. An extract of each material to be assayed, prepared as described in step 1, is similarly added in duplicate to a series of tubes in amounts estimated to supply 0.025–0.45 μg of nicotinic acid. The volume of extract added should not exceed 5 ml. All tubes are now diluted with distilled water to 5 ml; and 5 ml of the double-strength basal medium is added to each tube. The culture tubes are plugged with cotton or covered with glass or metal caps and autoclaved at 15 lb pressure (10.4 kilopascals) for 10–15 min. After cooling, each tube is inoculated with 1 drop of the inoculum suspension. The racks of tubes are incubated at 98.6°F (37°C) for 24 h, after which the turbidity of each culture tube is measured with a photoelectric colorimeter, preferably above 560 nm, where absorption of light due to color of the growth medium is minimized. The colorimeter reading is referred to a previously prepared curve relating turbidity of the cell suspension to dry cell weight. Such a curve is easily prepared by recording the optical density (OD) of serial dilutions of a heavy suspension of cells, a known aliquot of which is evaporated to dryness and weighed. A plot of optical density versus the corresponding calculated dry cell weight at each dilution gives the desired reference curve. Instead of measuring the growth response turbidimetrically, the incubation may be extended to 72 h and the acid produced in each assay tube measured by direct titration with 0.1 *N* NaOH.

6. In calculating results, a standard dose-response curve for nicotinic acid is prepared by plotting the average of the turbidity (recorded as optical

TABLE 2. Basal medium for determination of nicotinic acid with *Lactobacillus arabinosus*

Constituent	Amount per 100 ml of double-strength medium
Vitamin-free, acid-hydrolyzed casein	1.0 g
Glucose	2.0 g
Sodium acetate	1.2 g
L-Cystine	0.02 g
L-Tryptophan	0.02 g
Adenine	2.0 mg
Guanine	2.0 mg
Uracil	2.0 mg
Riboflavin	40 μg
Thiamin chloride	20 μg
Calcium pantothenate	20 μg
Pyridoxine	20 μg
Biotin	0.08 μg
Inorganic salts*	
Solution A	1.0 ml
Solution B	1.0 ml
pH adjusted to	6.8

*Solution A, 25 g K_2HPO_4, 25 g KH_2PO_4, water to make 250 ml solution. Solution B, 10 g $MgSO_4 \cdot 7H_2O$, 0.5 g NaCl, 0.5 g $FeSO_4 \cdot 7H_2O$, 0.5 g $MnSO_4 \cdot H_2O$, water to make 250 ml solution.

densities or, preferably, dry weight of cells produced) or titration values found at each level of the nicotinic acid standard against the amount of nicotinic acid present. The nicotinic acid content of the sample is determined by interpolating the response to known amounts of the test solution onto this standard curve. Only those values which fall on the rapidly ascending portion of the curve should be used. A typical standard curve obtained under these conditions is shown in the **illustration**.

The accuracy of the microbiological methods is usually given as ±10%. However, this limitation is set by procedural variations, and not by reproducibility of the growth response as such. Under proper conditions the reproducibility of assay results may surpass ±3%. Comparative data for the same material assayed by independent reliable methods are scarce. In evaluating an assay method, investigators have therefore relied on internal evidence, such as the ability to secure quantitative recoveries of the added test substance and also the reproducibility of assay values. In general, an assay for an essential nutrient is more accurate than one for a nutrient that is only stimulatory to growth of the test organism. *See* NIACIN.

Amino acid assay. These methods depend upon the use of microorganisms which have specific growth requirements for the individual amino acids. One microorganism may be used for the assay of several amino acids. *See* AMINO ACIDS.

The basal media used for assay of amino acids resemble those that are used for nicotinic acid (Table 2), but they contain a complete assortment of vitamins and an appropriate complete mixture of pure amino acids (save for the one to be determined) in place of the hydrolyzed casein. Procedures are then entirely similar to those described for the nicotinic acid assay, and dose-response curves similar to that of the graph are obtained for each of the amino acids.

The assay organisms useful for estimation of individual amino acids are listed in **Table 3**. L-Amino acids of demonstrated purity must be used as standards. With few exceptions, the D-isomers of the amino acids are inactive for most of the common assay organisms. The concentrations of the amino acids required to establish a standard curve are given in **Table 4**.

Peptides are widely utilized by certain microorganisms, and may, depending upon the assay conditions and the test organism, exhibit activity which is greater than, less than, or equal to that expected from their amino acid composition. In general, therefore, proteins should be hydrolyzed completely before

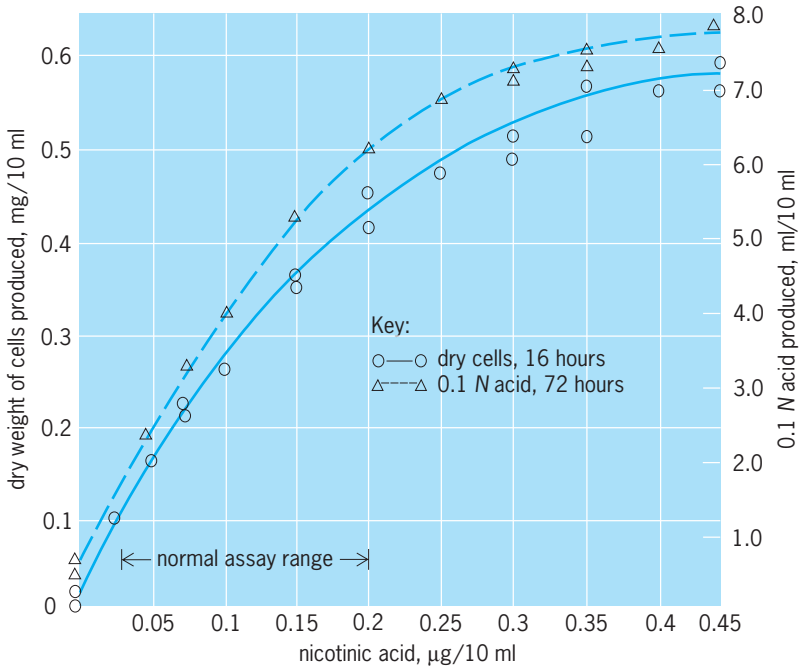

Growth and acid production of *Lactobacillus arabinosus* plotted as a function of the nicotinic acid concentration.

they are assayed microbiologically. The minimum time required for complete hydrolysis of any given protein can be determined experimentally. However, a general useful procedure for acid hydrolysis of protein samples for assay is as follows: The finely divided, weighed sample is placed in a test tube together with a 40-fold excess of 3*N* hydrochloric acid, and the tube is sealed under reduced pressure. It is autoclaved at 15 lb (10.4 kPa) pressure for 5–10 h, cooled, opened, and neutralized before assay. Tryptophan, tyrosine, and variable small amounts of other amino acids are destroyed by acid hydrolysis; modified procedures must be used to avoid this.

Antibiotic assay. Antibiotic assays depend upon the inhibition of growth of a susceptible microorganism. They are used to standardize batches of manufactured material, to test the potency of antibiotics in pharmaceutical preparations, for the estimation of antibiotics (which have been administered as chemotherapeutic agents) in body fluids, and for the detection of new antibiotics in culture filtrates of microorganisms and other source materials. The test organisms selected, in addition to being sensitive to the antibiotic being assayed, should be stable (not prone to changes in sensitivity or phase), easily

TABLE 3. Bacteria used for the determination of amino acids

Name and strain no. (ATCC)	For determination of
Lactobacillus arabinosus 8014	Glutamic acid, leucine, isoleucine, valine, phenylalanine, tryptophan
Lactobacillus delbrückii 9595	Phenylalanine, tyrosine, serine
Leuconostoc citrovorum 8081	Alanine; could also be used for most other amino acids
Leuconostoc mesenteroides 8042	All amino acids except alanine
Streptococcus faecalis 8043	Glutamic acid, histidine, lysine, arginine, leucine, isoleucine, valine, methionine, threonine, tryptophan

TABLE 4. Concentrations of various amino acids suitable for establishing a standard curve

Amino acid (L form)	Assay range, μg/ml
Alanine	0–25
Arginine	0–20
Aspartic acid	0–20
Cystine	0–5
Glutamic acid	0–25
Glycine	0–15
Histidine	0–5
Isoleucine	0–15
Leucine	0–15
Lysine	0–20
Methionine	0–5
Phenylalanine	0–5
Proline	0–10
Serine	0–10
Threonine	0–10
Tryptophan	0–2
Tyrosine	0–10
Valine	0–15

cultivated, and preferably nonpathogenic. The method selected depends upon the intended purpose.

Dilution methods. The substance to be tested is diluted to several concentrations and placed in a series of vessels containing a suitable liquid of solid nutrient medium. Sterile technique is observed except when the incubation period is as short as 3–5 h. The vessels are inoculated with a sensitive microorganism, and after incubation the lowest concentration of the substance which will just prevent, or reduce to one-half, growth or some other easily observable function of the microorganism is ascertained.

When a solid medium is used (agar streak method), several organisms are streaked on the surface of agar plates containing various dilutions of the antibiotic. This method is less accurate than the broth dilution method, but is useful as a rapid method of determining what organisms are inhibited and the approximate degree to which they are inhibited by the antibiotic (the antibiotic spectrum).

Diffusion methods. The agar plate (cup, cylinder, or paper disk) method is the most common horizontal diffusion method in use. A petri plate (or some variation thereof) containing nutrient agar is inoculated in bulk or on the surface (by spreading an aqueous suspension of cells, or a thin layer of inoculated agar) with vegetative cells or spores of a sensitive organism, and partially dried at 98.6°F (37°C). Four or five cylinders of glass, porcelain, or stainless steel are spaced on the plate and filled with the solutions to be tested. Absorbent paper disks may be used in place of the cylinders, or depressions or holes may be made in the agar to admit the solutions. The plates are incubated for a suitable period (12–16 h at 98.6°F or 37°C for penicillin). Clear zones appear around the point of application of the test substances where the antibiotic has diffused out to prevent growth of the microorganism. The diameters of the zones are measured (with a millimeter ruler, pointed dividers or calipers, or specially designed projection devices) and are compared with the dimensions of zones produced by known concentrations of the standard antibiotic.

Turbidimetric methods. These differ from dilution methods in that the response they give is a graded instead of an all-or-none response. The growth of the test organism in a liquid nutrient medium containing various known concentrations of the antibiotic is measured photoelectrically, and a standard curve is obtained by plotting the galvanometer readings versus the corresponding concentrations of antibiotic. Turbidity measurements of test solutions set up at the same time and under the same conditions are referred to the standard curve for evaluation.

Assay in body fluids. The assay of antibiotics in body fluids presents certain difficulties. The antibiotic may be present in low concentrations, and limited amounts of the body fluid may be available; therefore, the method must be sensitive and in some cases must be conducted on a microscale. Furthermore, the antibiotic may be destroyed or its potency changed by the presence of certain constituents in normal body fluids. Likewise, the growth of the test organism may be affected either favorably or adversely by the constituents in blood or serum. In general, the difficulties may be surmounted by employing one or more of the following procedures: (1) choosing a sensitive assay organism, avoiding those which are inhibited or stimulated by constituents of body fluids to be assayed; (2) when a micromethod is indicated, choosing a test organism which will grow on the undiluted body fluid; and (3) extracting the antibiotic from interfering materials.

The accuracy of these methods depends considerably upon the skill of the operator. Reproducibility of assay values is the only criterion of accuracy which has been extensively used.

Beverly M. Guirard; Esmond E. Snell

Disinfectant testing. A disinfectant serves to prevent infection. It must successfully kill microorganisms which occur on inanimate surfaces and can produce infection. These microorganisms include bacteria, fungi, yeasts, protozoa, and viruses. A knowledge of the effectiveness of disinfectants is important in many different areas. Most manufacturers of disinfectants require regular quality-control information; consumers such as hospitals have the responsibility of protecting their patients; regulatory agencies must make sure that all commercial products meet legal standards; and research scientists need to develop new products that will be superior to those already on the market. Therefore, standardized testing methods have been developed and improved in a century of experimentation.

Since disinfectants represent different chemical species and are used under different environmental conditions to act on different microorganisms, different methods must be employed as the specific situation requires. Some tests are strictly for use in the laboratory, others simulate practical conditions in the laboratory, and still others involve testing when in use under actual conditions. To illustrate the necessity for simulated and in-use tests, a disinfectant

was shown to be 20 times as effective on a ceramic tile wall as on a rubber tile wall; thus, a test unrelated to the actual surface in use would have given misleading results. The presence of organic matter is another of the many factors which may have a profound influence in reducing the activity of a disinfectant and must be taken into account in the testing process. The type of microorganism used is also very important to the results. Some disinfectants, for example, are active only on gram-positive bacteria, not on gram-negative organisms. The same organism may give different results if not treated in the same manner in laboratory culture; therefore, it is necessary to standardize the method of culture of the organism, as well as the other aspects of the test procedure. In the hopsital it may be necessary to additionally test wild species of the same organism that is used in the standard test, when such wild organisms are responsible for infection and may respond differently from the normal, standard culture grown in the laboratory.

While it would be desirable to test with a large number of different organisms and under practical conditions of use, the realistic considerations of cost and time most often dictate the use of repeatable tests performed in the laboratory with one to three organisms. Such standard tests have received many years of evaluation and improvement, and have been published under the auspices of authoritative professional organizations in the United States and Europe. There is no international set of standard tests, and the tests of different countries vary considerably from each other. In the United States the officially recognized methods are those published by the Association of Official Analytical Chemists (AOAC). Since not all microbiological testing methods are applicable to all types of disinfectants, different methods have been developed. The following briefly describes the more important AOAC methods for disinfectants.

Phenol-coefficient method. This is the oldest of the methods for testing disinfectants (dating from 1912) that is still in use. It is based upon a comparison of the disinfectant to be evaluated with a common standard, phenol. It is useful solely for chemicals which, by their type of action, are similar to phenol. It employs bacteria in liquid suspension and compares the concentration necessary for killing them in a specific time by the disinfectant and by an equal quantity of phenol, as determined by subculturing the bacteria in a special medium following contact with the chemical. Chemicals unlike phenol, such as mercurials and quaternary ammonium compounds, yield erroneous test results. Thus, the phenol-coefficient method is not widely in use.

TABLE 5. Example of test by the phenol-coefficient method employing *Salmonella typhosa**

Substance	Dilution	Growth† in subculture tubes in 5 min	10 min	15 min
Disinfectant X	1:300	0	0	0
	1:325	+	0	0
	1:350	+	0	0
	1:375	+	+	0
	1:400	+	+	+
Phenol	1:90	+	0	0
	1:100	+	+	+

*Phenol coefficient of Disinfectant X: 350/90 = 3.9.
†Symbol + indicates growth, and 0 no growth.

An example of the phenol-coefficient method is shown in **Table 5**. The phenol coefficient is calculated by dividing the greatest dilution of the disinfectant tested which is effective in killing the bacteria in 10 min, but not in 5 min, by the greatest dilution of phenol demonstrating the same result. This test method has been shown to yield a standard deviation of ±21%.

The phenol-coefficient method is restricted to disinfectants which are miscible with water or emulsifiable in water. Test organisms specified are *Salmonella typhosa, Staphylococcus aureus*, and *Pseudomonas aeruginosa*. They must be cultured in a specific manner on a stipulated medium in order to have a specific resistance, as shown in **Table 6**.

It was the commonly accepted criterion that disinfectants for general use be applied at a dilution equivalent in germicidal effectiveness against *Salmonella typhosa* to 5% phenol solution. This was determined for the phenolic-based disinfectant in question by multiplying the phenol-coefficient number against *Salmonella typhosa* by 20 to obtain the number of parts of water in which one part of the disinfectant should be incorporated. This is regarded as the highest possible dilution which could be considered for practical disinfection, and is valid only if the

TABLE 6. Resistance of test cultures to phenol as required by the phenol-coefficient method

Organism	Dilution	Growth* in subculture tubes in 5 min	10 min	15 min
Salmonella typhosa	1:90	+ or 0	+ or 0	0
	1:100	+	+	+ or 0
Staphylococcus aureus	1:60	+ or 0	+ or 0	0
	1:70	+	+	+
Pseudomonas aeruginosa	1:80	+ or 0	+ or 0	0
	1:90	+	+	+

*Symbol + indicates growth, and 0 no growth.

bactericidal adequacy of the solution can be confirmed by the use-dilution method.

Use-dilution method. Owing to the introduction of nonphenolic-based disinfectants, the use-dilution method is currently the test most commonly used to measure efficacy of disinfectants. This confirmation method tests effectiveness of disinfectants under semipractical conditions of use, namely on hard, nonporous surfaces where a substantial reduction of the bacterial contamination has not been achieved by prior cleaning. In the use-dilution method, the bacteria are applied to stainless steel cylinders, resulting in 1,000,000 bacteria per cylinder. Experience has shown this method to be much more reliable in indicating safe, practical dilutions for all types of disinfectants than the phenol-coefficient method, in which the bacteria are acted on in the form of suspensions. The use-dilution method simulates the action of the germicide on catheters, surgical instruments, and so forth. The steel cylinders are dipped into a 48-h culture of the bacteria and are allowed to drain and dry under set conditions. They are then immersed in dilutions of the disinfectant for 10 min at a standard temperature and transferred to subculture broth and incubated to determine whether complete kill of the test bacterium has occurred. Chemicals such as lecithin and thioglycollate are used in the subculture broth to neutralize any carry-over of disinfectant on the cylinder.

In this test the bacteria specified are *Salmonella choleraesius, Staphylococcus aureus*, and *Pseudomonas aeruginosa*. Disinfectants for hospital use or similar application require that all three bacteria be tested on three separate formulated product batches (one batch should be at least 60 days old), 60 carriers per bacterium, for a total of 540 determinations. Killing of 59 out of each set of 60 carriers or better is required.

Test for germicidal spray products. This method was developed to evaluate spray products, manual and pressurized, for disinfecting surfaces, and is adapted from the use-dilution method. It employs glass microscope slides instead of steel cylinders as the carriers for a suspension of *Salmonella choleraesius, Staphylococcus aureus, Pseudomonas aeruginosa*, and, if fungus control is claimed, spores of *Trichophyton mentagrophytes*. The inoculated slides are sprayed for 30 s and allowed 10 min contact time. Following broth culture of the treated slides and incubation, killing of all three test bacteria, as set forth for hospital disinfectants under the use-dilution method, is evidence of disinfecting action.

Germicidal and detergent sanitizer method. Public health standards require a percentage reduction of bacteria on precleaned surfaces that come in contact with food. Such surfaces could be dishes and utensils in restaurants, and dairies and other food-processing plants. The germicidal and detergent sanitizer test measures speed of kill with or without the presence of hardness in the test water. The test bacteria are *Escherichia coli* and *Staphylococcus aureus*, and are grown and diluted until a final concentration of 75,000,000 to 125,000,000 bacteria per milliliter is achieved. One milliliter of the bacterial suspension is added to 100 ml of the sanitizer diluted to use concentration at 77°F (25°C). After 30 to 60 s, the solution is sampled and the sample plated out. Colony-forming units are counted and compared to the original bacterial count. A 99.999% reduction in 30 s meets standards set by the U.S. Public Health Service Code, with or without hard water. This test should be clearly separated from the use-dilution method. In that test, complete kill of a target pathogenic bacterium is necessary. In this sanitizer test, a percentage reduction of bacteria found in the inanimate environment is necessary.

Available chlorine equivalent concentration method. Where sanitizers that are not solely chlorine-based are used on food contact surfaces, a separate test is used to measure efficacy. This test is designed to meet the requirements of the U.S. Public Health Service Code for sanitizers: at least 50 ppm of available chlorine and a 99.999% reduction of the test bacterium. This test shows whether or not a nonchlorine sanitizer is equivalent to a standardized concentration of chlorine. The test bacteria are *Salmonella typhosa* and *Staphylococcus aureus*. The bacteria are grown for broth for 24 h. For each concentration of chlorine (50, 100, and 200 ppm), 10 ml is placed in one tube. For the sanitizer under test, three concentrations are made and 10 ml of each concentration placed in a respective single tube. Over a period of 10 time-measured increments, each tube is dosed with one drop of bacterium, and after 1 min one drop is removed and placed in fresh media. The single tube is repeatedly dosed 10 times and subcultured 10 times. The media are incubated and observed for growth/no growth. If 25 ppm of iodine, for example, shows the same pattern of no growth for 1–5-min increments, and growth for 6–10-min increments, it would be equivalent to 200 ppm of chlorine, which would show the same growth/no growth pattern.

Sporicidal test. By definition, a disinfectant kills vegetative cells but not necessarily bacterial spores, which are the most resistant forms of life. When claims are made for sporicidal action, the test is made with resistant spores of *Bacillus subtilis, Clostridium sporogenes*, and spores of other species of particular interest, such as *Bacillus anthracis* and *Clostridium tetanis*. The sporicidal test is the same as a test for assessing sterilizers. (Sterilization is the killing of all life.) Much like the use-dilution method, this is a carrier test measuring disinfection at surfaces, in this case using silk surgical suture loops and porcelain cylinders impregnated with spores which have shown standard resistance to 2.5 *N* hydrochloric acid. The spore-treated carriers are exposed to the disinfectant (which may be a liquid or gas) for selected intervals at specified temperatures. This is followed by transfer to subculture media containing neutralizing compounds to prevent carry-over of bacteriostatic action, and incubation to indicate killing or survival of the spores. In actual tests, some 720 individual carriers are tested, with no failures in the final result.

Fungicidal test. A disinfectant may be employed to kill pathogenic fungi and fungal spores, which are less resistant to chemicals than bacterial spores. The test procedure is similar to the phenol-coefficient test, but employs a suspension of spores of the dermatophyte *Trichophyton mentagrophytes*, which has been shown to demonstrate the required resistance to phenol. The highest dilution that kills spores in 10 min exposure is considered effective.

Test for swimming pool disinfectants. This is a suspension test comparing bactericidal activity of the product to that of hypochlorite (chlorine). The organisms specified are *Escherichia coli* and *Streptococcus faecalis*. The test is based on the effectiveness of a swimming pool disinfectant in killing 1,000,000 test bacteria per milliliter of water in 30 s. After increasing contact times of 0.5 to 10 min with dilutions of the product, the culture is neutralized and subcultured in fresh medium. The concentration of the disinfectant under test that gives results equivalent to those of an NaOCl solution containing 0.58–0.62 ppm available chlorine at zero time and greater than 0.4 ppm after 10 min, is considered effective. Various swimming pool additives used to enhance the action of a disinfectant can be included in the test.

Test of tuberculocidal activity. A special test is required to determine tuberculocidal activity due to both the great resistance and slow growth of pathogenic mycobacteria. This test is divided into two parts: a screening test employing the more rapidly growing *Mycobacterium smegmatis* and a confirmatory test employing *Mycobacterium tuberculosis* var. *bovis*. The culture of organisms is applied to a carrier of cylinders made of porcelain, using 10 per trial, with 10 min exposure. In the confirmatory test, each carrier with tuberculosis bacteria is transferred to a tube with serum or neutralizer, followed by subculture in three different broth media with serum. Three different media are used because each furnishes to the damaged tuberculosis organism a growth factor that can be used to repair itself if the disinfectant did not kill the organism. The tubes are incubated for 60 days and, if there is no growth, for 30 days more. The maximum dilution of the disinfectant and the test organisms in the 10 carriers which shows no growth in the subculture represents the maximum safe use-dilution for practical tuberculocidal disinfection.

Virucidal methods. There is no official standard method for testing the effect of disinfectants on viruses; however, there are unofficial methods which are accepted by the U.S. Environmental Protection Agency on an individual case basis. These methods include the suspension method where the virus is suspended with the germicide in 5% blood serum and held for 10 min; the carrier method using dried viruses on inanimate surfaces; and the disinfectant spray method. In all cases, the virus is recovered and assayed in tissue culture, in embryonated egg, or in a living animal, depending on the virus under study. The work requires qualified laboratories and highly trained personnel.

Other methods. There are many other methods for testing disinfectants and antimicrobial chemicals. Some, for example, are practical tests for determining activity for dairy use where milk protein is an interfering substance; or tests on hospital linens and textiles, and for carpet sanitizers, where the disinfectant may be adsorbed on the fibers to reduce its activity. There are methods for testing antiseptics for use on the skin, and substances which are bacteriostatic or fungistatic, that is, which inhibit growth and reproduction but do not kill the cells. There are tests which determine the activity of chemicals to protect products such as foods, cosmetics, pharmaceuticals, paints, paper, wood, and petroleum from microbiological deterioration or destruction. There are methods which do not depend on observation of cell growth/no growth but determine microbial survival by such measurements as respiration and enzyme action. Seymour S. Block

Animal Assay

Animals, from hydra to humans (in humans, assays are termed clinical tests), are used in many ways in a wide range of different assays, from measuring clumping in slime molds for the estimation of minuscule levels of biological compounds, to sampling umbilical blood to evaluate human drug distribution in fetuses during pregnancy. Although there have been great strides in the development of computer simulation and predictive toxicological methods, it is still presently necessary to do animal tests because biological systems are so complex that the only way to be sure about the result is to study effects in animals.

Techniques. Animal bioassays are generally direct evaluations of agent efficacy (such as the number of dead mosquitoes in a standardized cage, or mice with cured cancers), indirect evaluations of agent efficacy using some marker (for example, reducing expression of an enzyme activity associated with AIDS progression), or estimations of agent toxicity using either short-term or long-term studies (mostly for regulatory purpose). Although similar whole-animal assays are sometimes used to determine substance concentrations, such tests, because of their high cost when compared to analytical methods or tests in microorganisms, are fairly rare and are usually confined to experimental situations where the active component is being sought or is in very low concentrations. *See* PHARMACOLOGY; TOXICOLOGY.

Bioassays, whether direct or indirect, for drug efficacy (or potency) also usually evaluate toxic effects. Animals are dosed with an agent, either acutely or chronically, under specified conditions, and the end result is evaluated. Tests using indirect evaluations of efficacy or toxicity tend to be of shorter duration than direct tests. Acute action of a single dose, such as lethality resulting from injection of digitalis, or chronic action of an acute or chronic dose (for example, long-term evaluation of the efficacy and toxicity of a single dose of carcinogen or lifelong treatment with diuretics for hypertension) can be evaluated. Drug potency evaluations are relatively short-term (especially in humans), since most drugs are usually intended to be given for a limited time, while toxicity evaluations tend to be both short- and

long-term. Acute effects can include vasodilation or heartbeat alterations when estimating drug potency, or weight loss and organ necrosis when estimating toxicity. Chronic action for potency can include changes in organ weight or blood composition, or for toxicity, carcinogenicity, teratogenicity, and mutational changes in the second and third generation following chemical exposure.

Although it is important to have a good experimental design in evaluating acute action, this becomes crucial in long-term tests, in part because of their cost. Some important factors can be illustrated by considering the chronic bioassay, a test to evaluate the change in tumor incidence and time-to-tumor in a rodent species as a result of chronic exposure to an agent. Issues which are constantly evaluated and debated include the number of animals used per dose, the strain and species used (although some strains are conventionally used), the maximum dose used, the number of other doses used, the dose route, the number of tissues evaluated for carcinogenicity, questions of tissue pathology, how long the assay should last (2 years is conventional), and the relevance of the test to aspects of human tumorigenesis (which the test is supposed to model). For example, the role that animal body weight has on influencing chronic bioassay sensitivity and outcome has been debated. This is especially true since the results of these assays have direct practical significance because they are presently used to set limits for human exposure to certain chemicals. This whole reevaluation process has led to basic questions, for example, about the extrapolation of data from animals given high doses to humans, especially in light of the low doses usually seen in living human subjects; and about what animal systems are proper models for humans.

Statistics. Because of the costs involved, while short-term tests tend to use large numbers of test animals, long-term animal assays are usually designed with the fewest animals possible. Another factor resulting in decreased animal numbers is concern about needless animal suffering. As a result, the tests tend to be fairly insensitive when used to model the response of a large human population. Because of this insensitivity, and biological variation between individuals and species, statistics (always a factor to consider in test analysis) assume special importance. Experimental design and data analysis are heavily influenced by statistical considerations, and special statistical techniques, such as methods for determining the effects of life-shortening and non-life-shortening tumor incidence, or determining the effect of agents on the time-to-tumor, have been developed which have subsequently been adapted for more general use (for example, factoring out the effect of life-shortening tumors on mortality in aging studies). *See* BIOMETRICS; STATISTICS.

Data management. Because of the vast amount of data generated by a large number of short-term tests, or even a single chronic bioassay, especially when good dose-response data are generated, there has arisen a need for automated systems for data acquisition and management. The costs of long-term tests has led to automation in the animal facility. An animal room without a computer terminal has become a rarity. Many short-term tests are automated, with robots doing the labor-intensive adding and mixing of solutions as well as the injection of samples into instruments for measurement. Integration of laboratory function with computers is increasing, with data being placed into databases. Extensive new data are becoming available on the Internet.

Angelo Turturro; Ronald Hart

Bibliography. G. E. Glass (ed.), *Bioassay Techniques and Environmental Chemistry*, Ann Arbor Science, Ann Arbor, MI, 1973; International Agency for Research in Cancer Working Group, Long-term and short-term screening assays for carcinogens: A critical appraisal, *IARC Monogr*. (Lyon, France), suppl. 2, 1980; *National Toxicology Program Specifications for the Conduct of Studies to Evaluate the Toxic and Carcinogenic Potential of Chemical, Biological, and Physical Agents in Laboratory Animals for the National Toxicology Program*, National Institutes for Environmental Health Sciences, Research Triangle Park, NC, 1992; R. L. Prentice, Surrogate measures in clinical trials: Definition and operational criteria, *Stat. Med.*, 8:963–974, 1989; R. Tennant and E. Zeigler, Genetic toxciology: Current status of methods of carcinogen identification, *Environ. Health Perspect.*, 100:307–315, 1993; A. Turturro et al., Body weight impact on spontaneous and agent-induced diseases in chronic bioassays, *Int. J. Toxicol.*, 17:79–100, 1998.

Biocalorimetry

The measurement of the energetics of biological processes such as biochemical reactions, association of ligands to biological macromolecules, folding of proteins into their native conformations, phase transitions in biomembranes, and enzymatic reactions, among others. Two different types of instruments have been developed to study these processes: differential scanning calorimeters and isothermal titration calorimeters.

Differential scanning calorimeters. These instruments measure the heat capacity at constant pressure of a sample as a continuous function of temperature. A typical high-sensitivity differential scanning calorimeter consists of two cells that are suspended inside an adiabatic shield (to prevent heat loss) and connected to each other through one or several thermopiles. One cell is filled with a solution containing the sample under study, and the other with a reference solution. The measurement is performed by heating both cells at a constant rate (approximately 1°C/min). When a thermally induced endothermic transition occurs in the sample cell, the temperature of this cell lags behind that of the reference cell. This phenomenon occurs because some of the applied thermal energy is utilized to induce the transition rather than to increase the temperature of the solution. This temperature difference is sensed by the thermopiles and used by a feedback circuit which adds additional energy to the sample cell in order to

maintain the temperature difference equal to zero. The additional electric power applied to the sample cell is the basic signal measured by a differential scanning calorimeter. Normalization of this signal by the scanning rate and the amount of solute in the sample cell permits a determination of the heat capacity function as a function of temperature. Modern instrumentation allows for increased sensitivity and accurate measurements of protein folding and unfolding transitions with approximately 1 mg of material. *See* THERMOCHEMISTRY.

Differential scanning calorimetry is the only technique that permits a direct measurement of the thermodynamics of thermally induced transitions. Thus it plays a unique role in the determination of the forces that stabilize the native structure of proteins, nucleic acids, and other biological structures. The combination of thermodynamic information with high-resolution molecular structures obtained by x-ray crystallography or nuclear magnetic resonance provides a powerful tool for the development of rational algorithms for protein engineering and drug design. At a most fundamental level, differential scanning calorimetry is also used to study the mechanisms of protein folding, since this technique permits a characterization of the energetics of partially folded structures along the folding pathway.

Isothermal titration calorimeters. Isothermal titration calorimeters measure directly the energetics (through heat effects) associated with biochemical reactions or processes occurring at constant temperatures. Experiments are performed by titrating reactants into sample solutions containing other reactants. After each addition, the heat released or absorbed as a result of the reaction is monitored by the instrument. The difference in heat effects between the sample solution and the reference solution is equal to the heat of reaction. In a power compensation instrument, the temperature difference between the sample and reference cells is monitored to maintain the temperature difference equal to zero. Instruments can measure heat effects down to about 4 microjoules (1 microcal) of total heat, the volume of the reaction cells is about 1 milliliter, and the volume introduced by each injection usually is 5–20 microliters. *See* TITRATION.

Isothermal titration calorimeters are typically used to measure the binding of a ligand to a macromolecule (for example, proteins) or larger macromolecular assemblies (for example, multisubunit proteins and membrane receptors). Usually, a sequence of injections are administered until all of the binding sites are saturated with a specific ligand. Since the heat effect for each injection is directly proportional to the amount of ligand bound, its magnitude decreases as the fractional saturation is titrated stepwise to completion. The output of a calorimetric titration experiment is the reaction heat as a function of the total amount of ligand added. A single calorimetric titration curve can be analyzed to yield the binding enthalpy, binding entropy, and therefore the Gibbs energy of binding. Performing the titration experiment at different temperatures allows estimation of the heat capacity change that accompanies the binding reaction. *See* THERMAL ANALYSIS.

Isothermal titration calorimeters can also be used to measure enzymatic reactions and, in general, biochemical reactions that can be initiated by the mixing of two or more reagents (for example, protein denaturation by the addition of chemical denaturants or by changes in pH). In all cases, the uniqueness of calorimetry resides in its capability to measure directly and in a model-independent fashion the heat energy associated with a process. *See* CALORIMETRY.

Ernesto Freire

Bibliography. A. E. Beezer (ed.), *Biological Microcalorimetry*, 1980; W. Hemminger and G. Hohne, *Calorimetry: Fundamentals and Practice*, 1984; J. A. McLean and G. Tobin, *Animal and Human Calorimetry*, 1988.

Biochemical engineering

The application of engineering principles to conceive, design, develop, operate, and/or use processes and products based on biological and biochemical phenomena. Biochemical engineering (also known as biomolecular engineering because of the emphasis on the molecular basis of biological phenomena) influences a broad range of industries, including health care, agriculture, food, enzymes, chemicals, waste treatment, and energy, among others. Historically, biochemical engineering has been distinguished from biomedical engineering by its emphasis on biochemistry and microbiology and by the lack of a health care focus. This is no longer the case. There is increasing participation of biochemical engineers in the direct development of pharmaceuticals and other therapeutic products. Biochemical engineering has been central to the development of the biotechnology industry, given the need to generate prospective products (often using genetically engineered microorganisms and cells) on scales sufficient for testing, regulatory evaluation, and subsequent sale. *See* BIOMEDICAL ENGINEERING; BIOTECHNOLOGY.

Historical perspective. Biochemical engineering takes its roots as a discipline in the fermentation of carbohydrates to produce alcohol, typically for beverages. This practice can be traced back to before 6000 B.C. The demand for large quantities of penicillin as part of the Allies' effort during World War II brought together microbiologists, biochemists, and chemical engineers to design and construct large reactors (which came to be known as fermenters) to cultivate the antibiotic-producing microorganism. Subsequently, this combination of expertise facilitated the development of a number of industries manufacturing products ranging from chemicals (organic acids and solvents) to pharmaceuticals (antibiotics and therapeutic proteins) to foods (alcoholic beverages, sweeteners, and cheeses). The advent of genetic engineering expanded the prospects for biological products and processes as well as the role that biochemical engineers have come to play. Thus, training in this discipline has evolved from a strong emphasis on chemical engineering

fundamentals with some elements of microbiology and biochemistry to a broader scope that includes molecular genetics, cell biology, pharmacology, materials science, and biophysics. *See* BIOPHYSICS; CHEMICAL ENGINEERING; GENETIC ENGINEERING; MATERIALS SCIENCE.

General principles. At the heart of biochemical engineering are the same principles that form the basis for chemical engineering, the discipline from which the engineering component is primarily derived. Whether the system is a large-scale bioreactor (a fermenter for microorganisms or a cell culture reactor for plant, insect, or mammalian cells) or a single cell, controlling the fluxes of heat, mass, and momentum between the system and its environment is central to effective operation. In fact, the biochemical engineer views a living cell as a minute but highly efficient bioreactor in which thousands of enzyme-catalyzed chemical reactions are orchestrated to serve the cell's needs to survive, reproduce, and possibly carry out a specific function in a multicellular organism. These reactions are subject to complex regulatory mechanisms that remain incompletely understood. Faced with the prospect of designing and operating a bioreactor in which billions of cells are cultivated, the biochemical engineer must often cope with overriding the intrinsic control system of the cell to produce a given substance. Thus, a biochemical engineer's task is often to promote inefficiencies in the cell's operation to produce a desired substance at higher levels than it would produce in its natural setting. This is particularly true for genetically engineered cells designed to overproduce a biochemical or protein, which may or may not be part of their normal metabolism.

Cultivation of microorganisms and cells. In the initial stages of the discipline, biochemical engineers were chiefly concerned with optimizing the growth of microorganisms under aerobic conditions, at scales of up to thousands of liters. While the scope of biochemical engineering has expanded in recent years, this remains a focus. More often, the aim is the development of an economic process to maximize production of a particular biomolecule (for example, a protein or metabolite), taking into consideration raw-material and other operating costs. The elemental constituents of biomass (carbon, nitrogen, oxygen, hydrogen, and to a lesser extent phosphorus, sulfur, mineral salts, and trace amounts of certain metals) are added to the bioreactor and consumed by the cells as they reproduce and carry out metabolic processes. Sufficient amounts of oxygen (usually supplied as sterile air) are also made available to the growing culture, usually involving mechanical agitation and gas sparging to overcome the low solubility of oxygen in aqueous media and to encourage the release of carbon dioxide formed from cellular respiratory processes. In some cases, cell morphology changes or extracellular metabolic products lead to increased viscosity of the media, thereby hindering gas-liquid mass transfer, a problem that must be addressed through specialized techniques. *See* CHEMICAL REACTOR; FERMENTATION; TRANSPORT PROCESSES.

The cultivation of anaerobic microorganisms, typically associated with fermentations in which organic acids or other solvents are produced, is usually characterized by slower growth rates and lower biomass yields; here, dissolved oxygen adversely affects microbial activity. The foremost application of anaerobic microorganisms is in waste treatment, where anaerobic digesters containing mixed communities of anaerobic microorganisms are used to reduce solids levels in industrial and municipal waste streams. *See* SEWAGE TREATMENT.

While the operation and optimization of large-scale microbial culture is still of major importance in biochemical engineering, the capability to cultivate a wide range of cell types has been developed. Biochemical engineers are often involved in the culture of cells derived from multicellular organisms, notably mammals, plants, and insects, for research purposes and to harness the unique biosynthetic capabilities of these cells. Most often, genetically engineered variants of these cell types are established for production of recombinant proteins. The ability of these cells to secrete proteins with proper glycosylation (attachment of sugar groups to proteins) and other modifications—a critical determinant of therapeutic protein quality—compensates for their slower growth, sensitivity to processing conditions (such as agitation, pH, and presence of metabolic waste products), and the need to formulate complex, defined media for their cultivation. *See* INDUSTRIAL MICROBIOLOGY.

Immobilized enzymes and cells. To harness the biocatalytic capacity of a cell or to make economic use of an enzyme needed for a particular biotransformation, immobilization of the cell or enzyme may be considered. In this case, the cells or enzymes are first produced on a large scale and then concentrated before immobilization. Next, the enzyme or cell is either attached to an inert support or embedded within a porous network or membrane. In doing so, a bioreactor can be operated such that the substrate is converted to products continuously by the immobilized biocatalyst, yielding enhanced throughput and obviating the separation of the biocatalyst from substrates and products. In some cases, living cells may be used to avoid the need for cofactor regeneration or to take advantage of a cascade of enzyme-catalyzed reactions that might be a component of the cell's metabolism. In this situation, some care must be taken to ensure cell viability, while discouraging unnecessary increases in biomass and in substrate consumption. In all cases, engineering aspects enter into the reactor design, such as the intensity of fluid agitation and the characteristic size of the insoluble support, which affect molecular transport of substrates and nutrients, as well as consideration of the material composition and structure and method of immobilization. *See* CELL (BIOLOGY); ENZYME.

Downstream processing. While the fermentation/cell culture and the associated reactor operation are central to bioprocess design, product recovery from solutions of cells and media is most often the more challenging and expensive task. Compared with the recovery of traditional chemicals, the needs for

concentrating the product and ensuring its purity are often far higher for biological products, and bioseparation processes must avoid damaging conditions such as heat and the addition of denaturing solvents. Approaches used for recovering biological products include various forms of liquid chromatography (such as ion exchange, hydrophobic interaction, size exclusion, and affinity), filtration, and electrophoresis. To the extent that the biological activity of the product must be maintained, care must be taken not to damage or adversely modify the molecule of interest during recovery. In the case of pharmaceuticals, this necessitates the development of purification protocols that must be strictly followed once regulatory approval of the drug and its manufacturing process is granted. Changes to the overall process necessitate a potentially long and expensive reevaluation from regulatory agencies. Thus, the design and operation of bioseparation processes are subject to constraints not encountered in chemical separations, such that they cannot be continuously optimized. *See* ELECTROPHORESIS; LIQUID CHROMATOGRAPHY.

New directions in biochemical/biomolecular engineering. Advances in modern biology and in the health sciences have led to the creation of new areas of interest that use biochemical engineering expertise. The focus on the molecular scale has given rise to the use of biomolecular engineering as the most appropriate descriptor for this engineering discipline. Within this field, there are a number of maturing and emerging specializations. Metabolic engineering uses the tools of molecular genetics, often in conjunction with whole-genome profiling of gene expression patterns and quantitative, mathematical models of metabolic pathways and bioreactor operation, to optimize cellular function for the production of specific metabolites and proteins. Protein engineering focuses on the identification of proteins with enhanced or novel biological activity for use as biocatalysts or proteins with therapeutic value, through functional screening of randomized protein variants or rational design based on protein structure. Cell and tissue engineering involves analysis and manipulation of the intra- and intercellular mechanisms and pathways that prompt and regulate cell functional responses in the context of an individual cell or multicellular tissue, as well as the design and characterization of artificial or composite biological/synthetic materials and tissues for biomedical applications.

Biochemical/biomolecular engineers are also actively involved in development of systems and processes for bioremediation and for the detection and clearance of hazardous/lethal biological agents, development of vaccines and antibodies, and the discovery and use of cells and enzymes capable of functioning in extreme environments. Biochemical and biomolecular engineering is viewed as a critical and enabling expertise in biological and biomedical research and in the further development of the biotechnology industry. Jason M. Haugh; Robert M. Kelly

Bibliography. J. E. Bailey and D. F. Ollis, *Biochemical Engineering Fundamentals*, 2d ed., 1986; I. B. Glowinski and G. Georgiou (eds.), *Research Opportunities in Biomolecular Engineering: The Interface between Chemical Engineering and Biology*, National Institutes of Health, Bethesda, MD, 1994; W. E. Goldstein et al. (eds.), *Biochemical Engineering VI: Annals of the New York Academy of Sciences*, vol. 589, 1990; D. A. Lauffenburger and J. J. Linderman, *Receptors: Models for Binding, Trafficking, and Signaling*, 1993; B. O. Palsson and S. N. Bhatia, *Tissue Engineering*, 2003; H. Pedersen et al. (eds.), *Biochemical Engineering VII: Annals of the New York Academy of Sciences*, vol. 665, 1992; M. L. Shuler and F. Kargi, *Bioprocess Engineering: Basic Concepts*, 2d ed., 2002; M. L. Shuler and W. A. Weigand (eds.), *Biochemical Engineering V: Annals of the New York Academy of Sciences*, vol. 506, 1987; G. N. Stephanopoulos, A. A. Aristidou, and J. Nielsen, *Metabolic Engineering: Principles and Methodologies*, 1998.

Biochemistry

The study of the substances and chemical processes which occur in living organisms. It includes the identification and quantitative determination of the substances, studies of their structure, determining how they are synthesized and degraded in organisms, and elucidating their role in the operation of the organism. Some processes of particular interest are the conversion of foods to energy, respiration, the synthesis of nucleic acids and proteins, and the regulation of the chemical activities of cells and organisms.

Substances. Carbohydrates are a class of substances which includes simple sugars such as glucose (also called dextrose) and large polysaccharides such as cellulose, a major structural material in plants, and starch, a storage form of glucose in plants. Carbohydrates are an important source of energy and structural materials in organisms, and may also be found linked to proteins in glycoproteins and linked to lipids in glycolipids. In addition, nucleic acids are composed partially of certain sugars. *See* CARBOHYDRATE; CELLULOSE.

Proteins are intimately involved in all life processes. Some function as biological catalysts called enzymes, speeding up chemical reactions which ordinarily proceed slowly at the temperatures of living organisms. Other proteins may act as carriers of material, as in the case of hemoglobin, which functions as a carrier of oxygen in many animals. Some proteins (such as collagen) act as structural material, for instance, as portions of cellular membranes and as the major component of hair, horn, skin, and feathers. Contractile proteins are involved in muscular activity and cell division. Proteins vary in size, but all are composed of essentially the same 20 amino acids; each protein molecule is composed of a definite sequence of 100 or more amino acids chemically linked one to the other. Normally, each chain of amino acids is folded into a specific three-dimensional structure. *See* AMINO ACIDS; ENZYME; PROTEIN.

There are two classes of nucleic acids, deoxyribonucleic acid (DNA) and ribonucleic acid

(RNA). The nucleic acids are large molecules composed of long chains of primarily four different nucleotides. Nucleotides are composed of sugar phosphates bonded to organic bases; they exist as small molecules in addition to providing the material for large nucleic acids. The sequential array of nucleotides is quite specific for each molecule and constitutes the form in which genetic information is stored and transferred. *See* DEOXYRIBONUCLEIC ACID (DNA); NUCLEIC ACID; RIBONUCLEIC ACID (RNA).

Lipids are a diverse group of "greasy" substances which are soluble in organic solvents but insoluble in water. Cells and tissues are composed predominantly of water, so cellular structures must be composed of materials which are insoluble in an aqueous environment. Cellular membranes are composed mainly of a class of lipids called phospholipids. The most common lipids are the fats, which serve as insoluble stores of biological fuel. Some other lipids are cholesterol and steroid hormones. *See* LIPID.

Minerals such as salts of sodium, potassium, calcium, and magnesium contribute to the ionic environment within and around cells. In addition, calcium phosphate is an important constituent of bone, and iron is a necessary part of respiratory proteins such as hemoglobin and cytochromes. Trace metals such as zinc, selenium, and molybdenum have been found to be constituents of certain enzymes. *See* CALCIUM METABOLISM; HEMOGLOBIN; PHOSPHATE METABOLISM.

Vitamins and hormones are substances which are biologically effective in small amounts. In living organisms, vitamins are converted into coenzymes, substances which collaborate with some enzymes in catalyzing reactions. Hormones are made in and secreted by endocrine glands, and they act in regulating the activities of other tissues. *See* COENZYME; ENDOCRINE SYSTEM (VERTEBRATE); HORMONE; VITAMIN.

Processes. Many of the chemical steps involved in the biological breakdown of sugars, fats, and amino acids are known. It is well established that living organisms capture the energy liberated from these reactions by forming a high-energy compound, adenosine triphosphate (ATP). In the absence of oxygen, some organisms and tissues derive ATP from an incomplete breakdown of glucose, degrading the sugar to an alcohol or an acid in the process. In the presence of oxygen, many organisms degrade glucose and other foodstuff to carbon dioxide and water, producing ATP in a process known as oxidative phosphorylation. In spite of many years of intensive investigation, the chemical reactions of oxidative phosphorylation are not well understood. *See* BIOLOGICAL OXIDATION; CARBOHYDRATE METABOLISM; LIPID METABOLISM; PROTEIN METABOLISM.

Structure and function. The relationship of the structure of enzymes to their catalytic activity is becoming increasingly clear. It is now possible to visualize atoms and groups of atoms in some enzymes by x-ray crystallography. Some enzyme-catalyzed processes can now be described in terms of the spatial arrangement of the groups on the enzyme surface and how these groups influence the reacting molecules to promote the reaction. It is also possible to explain how the catalytic activity of an enzyme may be increased or decreased by changes in the shape of the enzyme molecule, an important aspect of regulation. An important advance has been the development of an automated procedure for joining amino acids together into a predetermined sequence. This procedure permits the chemical synthesis of peptide hormones and enzymes. The use of recombinant DNA technology also makes possible the controlled biosynthesis of peptides and proteins of any amino acid sequence. This has permitted the chemical synthesis of the enzyme ribonuclease. This technology will permit the synthesis of slightly altered enzymes and will improve the understanding of the relationship between the structure and the function of enzymes. In addition, this procedure permits the synthesis of medically important polypeptides (short chains of amino acids) such as some hormones and antibiotics.

Genetic code. A subject of intensive investigation has been the explanation of genetics in molecular terms. It is now well established that genetic information is encoded in the sequence of nucleotides of DNA and that, with the exception of some viruses which utilize RNA, DNA is the ultimate repository of genetic information. The sequence of amino acids in a protein is programmed in DNA; this information is first transferred by copying the nucleotide sequence of DNA into that of messenger RNA, from which this sequence is translated into the specific sequence of amino acids of the protein. Each amino acid is specified by a sequence of three nucleotides (a triplet code); an amino acid may be specified by more than one triplet. This code has been deciphered and appears to be identical for all known organisms, thus supporting the concept of a unitary origin of life on Earth. Attempts have been made to simulate the conditions on Earth before life appeared; it has been possible to demonstrate, under these simulated conditions, the formation of substances such as sugars, amino acids, and nucleotides from the simpler substances which are presumed to have been present. *See* GENETIC CODE; MOLECULAR BIOLOGY; PREBIOTIC ORGANIC SYNTHESIS.

Disease. The biochemical basis for a number of diseases is becoming increasingly clear. This is particularly true for some genetically inherited diseases, in which the cause has been traced to the production of a defective protein. Sickle cell anemia is a particularly striking example; it is well established that the change of a single amino acid in hemoglobin has resulted in a serious abnormality in the properties of the hemoglobin molecule. *See* DISEASE.

Regulation. Increased understanding of the chemical events in biological processes has permitted the investigation of the regulation of these processes. An important concept in biological regulation is the chemical feedback circuit: the product of a series of reactions can itself influence the rates of the reactions. For example, the reactions which lead to the production of ATP proceed vigorously when the supply of ATP within the cell is low, but they slow down markedly when ATP is plentiful. These observations can be explained, in part, by the fact that ATP

molecules bind to some of the enzymes involved, changing the surface features of the enzymes sufficiently to decrease their effectiveness as catalysts. While it is possible to regulate a series of reactions by controlling the activities of the catalyzing enzymes, it is also possible to regulate these reactions by changing the amounts of the enzymes; the amount of an enzyme can be controlled by modulating the synthesis of its specific messenger RNA or by modulating the translation of the information of the RNA molecule into the enzyme molecule. Another level of regulation involves the interaction of cells and tissues in multicellular organisms. For instance, endocrine glands can sense certain tissue activities and appropriately secrete hormones which control these activities. The chemical events and substances involved in cellular and tissue "communication" have become subjects of much investigation. In addition, the process of biological development is also being examined. Chemical explanations are being sought, for instance, for the orderly changes that occur in an embryo as it develops.

Biological membranes. It has been known for some time that biological membranes exhibit selective permeability; some substances cross membranes easily, others are excluded completely. Indeed, some materials are actively pumped across membranes from a region of low concentration to a region of high concentration. A particularly striking example is the case of sodium ions which are actively transported out of most cells. Detailed explanations of selective permeability and active transport are not yet available and await a better understanding of membrane structure and function. A second aspect of membrane function that has received much attention involves the transfer of information across the cellular membrane. Some hormones, for instance, influence the metabolic activities of certain tissues without entering the affected cells and are thought to act by altering the activities of certain enzymes which are located on or within the cellular membrane. *See* CELL MEMBRANES.

Photosynthesis and nitrogen fixation. Two subjects of substantial interest are the processes of photosynthesis and nitrogen fixation. In photosynthesis, the chemical reactions whereby the gas carbon dioxide is converted into carbohydrate are understood, but the reactions whereby light energy is trapped and converted into the chemical energy necessary for the synthesis of carbohydrate are unclear. The process of nitrogen fixation involves the conversion of nitrogen gas into a chemical form which can be utilized for the synthesis of numerous biologically important substances; the chemical events of this process are not fully understood. *See* NITROGEN CYCLE; PHOTOSYNTHESIS.

Methods. The development of biochemistry closely parallels developments in methods of separately and quantitatively measuring small amounts of substances and of determining the structures of substances. The following are some methods employed in biochemistry.

1. Photometry for determining the absorption or emission of light by substances. Measurements conducted with ultraviolet light, below wavelengths of 400 nanometers, or with visible light, 400 to 700 nm, generally furnish information on the concentrations of substances. Infrared measurements, involving wavelengths greater than 700 nm, furnish information on chemical structure as well.

2. Chromatography for separating dissolved substances from one another. This may involve the differential binding of some substances to solid support, the partitioning of substances between two liquid phases, or separation by molecular sieving.

3. Gas chromatography for separation of substances in the gas phase.

4. Isotope labeling for following atoms through reactions, for locating elements in tissues by isotope measurements or radiography, and for ultramicroanalyses, including activation analysis by bombardment with neutrons.

5. Ultracentrifugation for separating large molecules in solution or cell particles in suspension by their differential migration in a centrifugal field. This method is also employed for the determination of molecular weight and for the separation of large molecules which differ from one another in their densities.

6. Electrophoresis for the separation of dissolved charged molecules on the basis of differential migration in an electric field.

7. X-ray crystallography for the determination of the three-dimensional structure of molecules. The method is based on the analysis of the diffraction patterns obtained when a crystal is placed in a beam of monochromatic x-rays.

8. Mass spectrometry for the identification of substances. The substance to be analyzed is volatilized, and the individual molecules are broken into charged fragments. The mass of each charged fragment is determined by its migration in an electric field; the number and characteristics of the fragments permit the deducing of the structure of the parent compound.

9. Nuclear magnetic resonance spectroscopy for the identification of chemical groups in substances and for evaluating changes in the environment around these groups. Certain atomic nuclei, for instance the nucleus of hydrogen, oscillate at radio frequencies when subjected to a strong constant magnetic field. When radio waves of varying frequencies are passed through the solution, absorption of the radio waves is observed at frequencies at which resonance occurs. The frequency at which resonance is observed is a function of the nucleus, the group of which the nucleus is a part, and the environment around the group. *See* ACTIVATION ANALYSIS; AUTORADIOGRAPHY; CHROMATOGRAPHY; ELECTROPHORESIS; GAS CHROMATOGRAPHY; HISTORADIOGRAPHY; MASS SPECTROMETRY; RADIOCHEMISTRY; SPECTROSCOPY; ULTRACENTRIFUGE; X-RAY CRYSTALLOGRAPHY. A. S. L. Hu

Bibliography. R. C. Bohinski, *Modern Concepts in Biochemistry*, 5th ed., 1987; T. M. Devlin, *Textbook of Biochemistry*, 4th ed., 1997; D. E. Schumm, *Essentials of Biochemistry*, 1988.

Biodegradation

The destruction of organic compounds by microorganisms. Microorganisms, particularly bacteria, are responsible for the decomposition of both natural and synthetic organic compounds in nature. Mineralization results in complete conversion of a compound to its inorganic mineral constituents [for example, carbon dioxide (CO_2; from carbon), sulfate or sulfide (from organic sulfur), nitrate or ammonium (from organic nitrogen), phosphate (from organophosphates), or chloride (from organochlorine)]. Since carbon comprises the greatest mass of organic compounds, mineralization can be considered in terms of CO_2 evolution. Radioactive carbon-14 (^{14}C) isotopes enable scientists to distinguish between mineralization arising from contaminants and soil organic matter. However, mineralization of any compound is never 100% because some of it (10–40% of the total amount degraded) is incorporated into the cell mass or products that become part of the amorphous soil organic matter, commonly referred to as humus. Thus, biodegradation comprises mineralization and conversion to innocuous products, namely biomass and humus. Primary biodegradation is more limited in scope and refers to the disappearance of the compound as a result of its biotransformation to another product. *See* HUMUS.

Enrichment culture. Compounds that are readily biodegradable are generally utilized as growth substrates by single microorganisms. Many of the components of petroleum products (and frequent ground-water contaminants), such as benzene, toluene, ethylbenzene, and xylene, are utilized by many genera of bacteria as sole carbon sources for growth and energy. Microorganisms can be isolated by a small environmental sample that is placed into a flask containing a mineral salts medium with the desired growth substrate (for example, benzene). Thus, the medium is enriched to select those microorganisms (prototrophs) able to utilize the substrate for growth and to synthesize all of their amino acids, nucleic acids, vitamins, and other cellular constituents from inorganic nutrients. Organisms that are capable of metabolizing the compound but require organic growth factors (auxotrophs) are not selected by this procedure. The addition of trace quantities of yeast extract or other growth supplements may enable selection of some auxotrophs, but a high concentration of growth supplements should be avoided as it will enrich for yeast extract utilizers instead of benzene utilizers in the example above. *See* BACTERIAL GROWTH; CULTURE.

Cometabolism and fortuitous metabolism. The process whereby compounds not utilized for growth or energy are nevertheless transformed to other products by microorganisms is referred to as cometabolism. Chlorinated aromatic hydrocarbons, such as diphenyldichloroethane (DDT) and polychlorinated biphenyls (PCBs), are among the most persistent environmental contaminants; yet they are cometabolized by several genera of bacteria, notably *Pseudomonas*, *Alcaligenes*, *Rhodococcus, Acinetobacter*, *Arthrobacter*, and *Corynebacterium*. Cometabolism is caused by enzymes that have very broad substrate specificity. *See* POLYCHLORINATED BIPHENYLS; PSEUDOMONAS.

Bacteria that have evolved pathways for the complete catabolism of aromatic hydrocarbons use these same enzymes to fortuitously transform many chlorinated analogs to more oxidized metabolites. However, not all enzymes in the catabolic pathway have broad specificity, so that a chlorinated product eventually accumulates. For example, bacteria that utilize biphenyl cometabolize polychlorinated biphenyls to chlorobenzoates, which are not acted upon by the enzymes that process benzoate. However, chlorobenzoates are further metabolized by other bacteria that have enzymes which act upon them. Although cometabolism of polychlorinated biphenyls in the environment is a very slow process, it can be greatly enhanced by the addition of analogous growth-promoting substrates, such as biphenyl, a process referred to as analog enrichment. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; CITRIC ACID CYCLE.

Microbial consortia and catabolic genes. Recalcitrant compounds are not degraded by single microorganisms but by a consortium of different organisms. In some cases, the interactions may be loosely associated or asymbiotic (protocooperation) whereby each organism carries out a different step in the catabolic pathway, yet is not dependent upon the other. In other cases, the association may be symbiotic where both benefit (mutualism) or only one benefits (commensalism). Biodegradation of polychlorinated biphenyls in soil is an example of commensalism in which the organism carrying out the initial oxidation receives no benefit while another organism benefits from utilizing the products of cometabolism (chlorobenzoates) for growth. Mutualism normally involves cross-feeding (syntrophism) in which all members are dependent upon each other.

It was discovered in the mid-1970s that dichlorodiphenyltrichloroethane was dehalogenated anaerobically and then mineralized aerobically by different bacteria. It has been shown that polychlorintated biphenyls that are inert to aerobic oxidations (particularly those containing six or more chlorine atoms per molecule) are dehalogenated to lesser chlorinated congeners by anaerobic bacteria in sediments. Thus, the potential exists for complete biodegradation of chlorinated aromatic hydrocarbons by a sequence of anaerobic/aerobic catabolic reactions.

An alternative strategy to mineralization by microbial consortia is to combine the genes of parental strains having the complementary parts of the catabolic pathway into a single organism by recombinant deoxyribonucleic acid (DNA) technology or by natural enhanced mating of parental strains. Catabolic genes that are encoded on plasmids, which are small extrachromosomal elements, are much more readily transferred and expressed in other bacteria than genes coded on the chromosome. Genetic exchange of catabolic plasmids among indigenous bacteria occurs naturally in the environment.

Bioremediation. The use of microorganisms to remediate the environment of contaminants is referred

to as bioremediation. This process is most successful in contained systems such as surface soil or ground water where nutrients, mainly inorganic nitrogen and phosphorus, are added to enhance growth of microorganisms and thereby increase the rate of biodegradation. The process has little, if any, applicability to a large open system such as a bay or lake because the nutrient level (that is, the microbial density) is too low to effect substantive biodegradation and the system's size and distribution preclude addition of nutrients. Bioremediation of oil spills on beaches and of ground water containing xylene, from leaking storage tanks can be enhanced by the addition of nitrogen (N) and phosphorus (P), which are essential macroelements of all cells.

Remediation of petroleum products from ground waters is harder to achieve than from surface soil because of the greater difficulty in distributing the nutrients throughout the zone of contamination, and because of oxygen (O_2) limitations. The latter problem is caused by the low solubility of oxygen and by the high biological oxygen demand that occurs when the organisms are present in high numbers and actively mineralizing the contaminant. Hydrogen peroxide (H_2O_2), which is completely miscible in water and rapidly decomposes to oxygen, has been added with nutrients to overcome this problem. Another approach involves a venting system whereby a stream of air is continuously introduced at injection wells at the site. This process also brings about volatilization of the contaminants, which may be metabolized by the indigenous microorganisms as they move upward through the soil. The removal of volatile organic compounds by air stripping and enhancement of biodegradation is referred to as bioventing. This process has proved to be successful in removing chlorinated organic solvents such as trichloroethylene that are major ground-water contaminants.

Fungal enzymes. Bacteria are the primary biocatalysts of organic contaminants, and no eukaryotic organism has ever been isolated that can utilize an aromatic hydrocarbon for growth. Nevertheless, fungi have superior ability to degrade lignin, the primary and most recalcitrant constituent of wood fiber. The extracellular enzymes, in the presence of oxygen, create highly oxidative radicals which break many of the carbon-carbon bonds and the aromatic ether bonds of lignin. The ligninases of *Phanerochaete chrysosporium* fortuitously oxidize many aromatic compounds, including diphenyldichloroethane, polychlorinated biphenyls, polyaromatic hydrocarbons, and trinitrotoluene.

Because this mode of biodegradation is nonspecific and does not proceed in a sequential manner as with bacterial consortia, it is less predictable, less controllable, and less likely to proceed to completion. It is also far less successful in soil than in culture because soil organic matter and other soil constituents also react with the enzyme-catalyzed radicals. Nevertheless, this obstacle has been partially overcome in land-farming operations by diluting soil with massive amounts of plant residues upon which the fungus is maintained or by adding the contaminated soil to reactors containing the fungus. *See* ENZYME; FUNGI.

Dennis D. Focht

Bibliography. R. M. Atlas, *Petroleum Microbiology*, 1984; R. M. Atlas and R. Bartha, *Microbial Ecology: Fundamentals and Applications*, 4th ed., 1997; D. T. Gibson, *Microbial Degradation of Organic Compounds*, 1984; T. Leisinger et al. (eds.), *Microbial Degradation of Xenobiotic and Recalcitrant Compounds*, 1982; W. W. Mohn and J. M. Tiedje, Microbial reductive dehalogenation, *Microbiol. Rev*, 56:482–507, 1992; M. J. Pelczar, Jr., E. C. S. Chan, and N. R. Krieg, *Microbiology: Concepts and Applications*, 1993.

Biodiversity

The variety of all living things; a contraction of biological diversity. Biodiversity can be measured on many biological levels ranging from genetic diversity within a species to the variety of ecosystems on Earth, but the term most commonly refers to the number of different species in a defined area.

Global biodiversity. Recent estimates of the total number of species range from 7 to 20 million, of which only about 1.75 million species have been scientifically described. The best-studied groups include plants and vertebrates (phylum Chordata), whereas poorly described groups include fungi, nematodes, and arthropods (see **table**). Species that live in the ocean and in soils remain poorly known. For most groups of species, there is a gradient of increasing diversity from the Poles to the Equator, and the vast majority of species are concentrated in the tropical and subtropical regions.

Human activities, such as direct harvesting of species, introduction of alien species, habitat destruction, and various forms of habitat degradation (including environmental pollution), have caused dramatic losses of biodiversity. The sixth major extinction event in geologic history is well under way. Indeed, current extinction rates are estimated to be 100–1000 times higher than prehuman extinction rates. This rapid loss of species has spurred a great deal of scientific interest in the topic of biodiversity. Currently, many biologists are working to catalog and describe extant species before they are lost. In addition, much research is focused on understanding the importance of biodiversity, particularly whether high levels of biodiversity are essential for proper functioning of ecosystems.

Importance. Ethical and esthetic arguments have been offered regarding the value of biodiversity and why it is necessary to guard against its reduction. Scientists, however, focus on issues such as the biological or ecological functions of biodiversity that can be addressed with experiments rather than debates about values. Certainly, some measure of biodiversity is responsible for providing essential functions and services that directly improve human life. For example, many medicines, clothing fibers, and industrial products and the vast majority of foods are derived from naturally occurring species. In addition, species are the key working parts of natural

Numbers of extant species for selected taxanomic groups

Kingdom	Phylum	Number of species described	Estimated number of species	Percent described
Protista		100,000	250,000	40.0
Fungi	Eumycota	80,000	1,500,000	5.3
Plantae	Bryophyta	14,000	30,000	46.7
	Tracheophyta	250,000	500,000	50.0
Animalia	Nematoda	20,000	1,000,000	2.0
	Arthropoda	1,250,000	20,000,000	5.0
	Mollusca	100,000	200,000	50.0
	Chordata	40,000	50,000	80.0

* With permission modified from G. K. Meffe and C. R. Carroll, *Principles of Conservation Biology*, Sinauer, Massachusetts, 1997.

ecosystems. They are responsible for maintenance of the gaseous composition of the atmosphere, regulation of the global climate, generation and maintenance of soils, recycling of nutrients and waste products, and biological control of pest species. Ecosystems surely would not function if all species were lost, although it is unclear just how many species are necessary for an ecosystem to function properly. Thus, the current extinction crisis has provoked many scientists to ask how many species can be lost from an ecosystem before the system is negatively affected.

Ecosystem function. Since species are the key working parts of ecosystems, biodiversity must be related to ecosystem function. Studies have assessed this relationship in ecosystem functions such as biogeochemical processes; the flow of nutrients, water, and atmospheric gases; and the processing of energy. Evidence of the importance of biodiversity for ecosystem function is derived from comparing ecosystems that differ in the number of species present. More recently, ecologists have undertaken manipulative experiments in which the number of species has been directly varied. Two notable experimental efforts include manipulations of plant diversity in a Minnesota grassland and manipulations of species and functional diversity in microbial communities. In general, these studies have demonstrated that various measures of ecosystem function such as production of biomass and nutrient uptake increase as the number of species present increases. However, some studies report no effect or even negative relationships between biodiversity and ecosystem processes.

Although some evidence supports the hypothesis that biodiversity increases or improves the overall functioning of ecosystems, the underlying mechanisms remain unclear. For example, a positive relationship between species diversity and productivity could result because including more species to increases the chance of encompassing particularly productive or fast-growing species. Alternatively, a diverse group of species may use the available resources more efficiently, since each species has slightly different requirements, resulting in higher overall growth.

It is unclear whether the number of species or the number of different functional types of species is driving these effects. This distinction is important because if the number of species matters most, every species that is added to an ecosystem should cause an improvement in ecosystem function (**illus.** *a*). In contrast, if the diversity of functional types is more important than the number of species per se, there will be initial increases in ecosystem function as species number rises, but these effects should level off once all of the functional types are represented (illus. *b*). Indeed, a nonlinear or satiating effect of species number on ecosystem processes is frequently observed (illus. *b*), suggesting that ecosystem function may be relatively unaffected by initial losses of species but may become severely impaired after some critical number of species is lost.

Ecosystem stability. A second purported benefit of biodiversity is that more diverse ecosystems may be more stable or more predictable through time when compared to species-poor ecosystems. Stability can be defined at the community level as fewer invasions and fewer extinctions, meaning that a more stable community will contain a more stable composition of species. However, stability can also be defined at the population level as reduced fluctuations in population size, meaning that a more stable population will contain a more constant number of individuals.

The idea that biodiversity confers stability upon ecosystems has a long and controversial history. Early ecologists used several lines of reasoning to argue that diverse ecosystems are more stable than those with fewer species. First, attempts to create simple, low-diversity ecosystems in the laboratory tended to fail, with most or all of the species declining to extinction. Second, unpredictable or strongly cyclical population dynamics are often observed in animals that live at high latitudes, and high-latitude

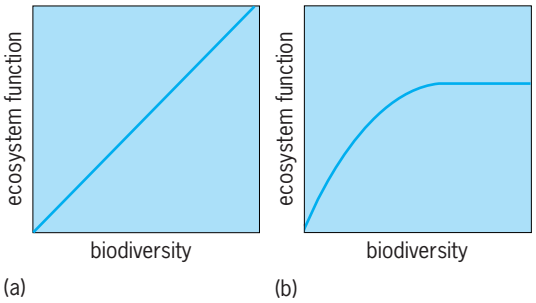

Ecosystem function (*a*) as a positive, linear function of biodiversity and (*b*) as a nonlinear, satiating function of biodiversity.

ecosystems generally include relatively few species. Finally, islands, which generally have fewer species than mainlands, tend to be more easily invaded by introduced species. In 1955, an additional argument was proposed in favor of a positive relationship between biodiversity and ecosystem stability. With more species in an ecosystem, there are more paths through which energy and nutrients can flow; therefore, in diverse ecosystems each species should be less affected by changes in the abundance of other species, leading to higher overall stability. Thus, the general consensus among early ecologists was that more diverse ecosystems should be more stable. *See* ECOSYSTEM; POPULATION ECOLOGY.

In 1973, mathematical models of many species interacting simultaneously were used to explore the relationship between biodiversity and population stability. The major outcome was that higher species diversity led to less stable population sizes of individual species. However, the apparent conflict between these modeling results and the intuitions of earlier ecologists remained unresolved for many years.

Recent studies in which the number of species has been experimentally manipulated have helped to resolve this long-standing controversy. For example, manipulations of plant diversity were used to examine not only the productivity of a grassland ecosystem but also the stability of ecosystem productivity over time. This and other studies have shown that although the abundance of individual species fluctuates more dramatically in high-diversity ecosystems, the total abundance or productivity of all species combined is actually more stable. High biodiversity decreased the stability of each species' population, lending support to mathematical modeling results, whereas the positive relationship between biodiversity and the stability of overall ecosystem productivity supports the proposals of the earlier ecologists.

Although the relationship between biodiversity and ecosystem stability is fairly clear, the mechanisms generating this pattern are not. In particular, diverse groups of species may be more stable because complementary species compensate for changes in one another's abundance. Alternatively, variation in aggregate measures such as total productivity may increase with richness due to averaging of random fluctuations in the growth of each species. Based on simple probability theory, it is expected that the more independently varying species added together, the more stable the sum of their abundances. The strength of this averaging effect depends on correlations among the species' fluctuations, but a positive relationship between biodiversity and the stability of aggregate measures of ecosystem function should usually be expected, simply due to averaging.

Clearly, biodiversity is (at least sometimes) related to both the overall rates and the stability of ecosystem functions. However, documenting a relationship between biodiversity and some measure of ecosystem function or stability does not reveal its underlying cause. Current ecological research continues to explore the mechanisms by which species diversity and functional diversity contribute to ecosystem function.

Species importance. Some species clearly play very important roles in ecosystems. In some cases, the addition or deletion of a single species can lead to dramatic changes in ecosystem functions such as productivity or nutrient uptake. For example, the introduction of a nitrogen-fixing tree species to Hawaii resulted in substantially altered productivity and nutrient dynamics in submontane forest ecosystems. Species that exert such strong control over ecosystems are termed keystone species.

It is not at all clear that most species in an ecosystem have such important effects. In other words, it may be possible to lose a number of species from an ecosystem and yet observe little overall impact on ecosystem function. This could be the case if several species that perform approximately the same function are present in the original ecosystem. The situation where multiple species play a similar role has been termed species redundancy. If species redundancy is a common phenomenon, ecosystem function should be largely independent of species diversity as long as major functional types are represented. Thus, when one species is lost from an ecosystem, some other species with a similar function may become abundant and compensate for the lost species, leaving the ecosystem as a whole relatively unaffected. Indeed, ecosystem processes often do remain stable despite large fluctuations in the abundance of the various species involved. In addition, the relationship between ecosystem function and biodiversity is often observed to be nonlinear (illus. *b*), suggesting that, at least initially, the loss of species would have little overall effect.

The term species redundancy may seem to imply that all species are not necessary for an ecosystem to function properly. However, species redundancy may be an essential feature for the long-term health of ecosystems. Just as engineers include multiple structures with redundant functions to increase overall reliability of the final structure, ecosystems with sets of functionally redundant species may have a built-in safety net that is lacking in species-poor ecosystems.

Rare species (those that occur in low abundance) may also appear to contribute relatively little to overall ecosystem functioning. However, during dramatic environmental changes, such as acidification of a lake, rare species can become very abundant, thereby compensating for reductions in other species. Even species that appear relatively unimportant because they are rare and functionally redundant with others may in fact be important in stabilizing ecosystem function during periods of rare but intense stress.

The overwhelming variety of life has captivated the human imagination for centuries, so it is surprising how much scientific uncertainty currently surrounds the role of biodiversity. Ignorance probably reflects the fact that biodiversity has been taken for granted; only over the last few decades, as biodiversity's staggering decline became more apparent, did ecologists start investigating what exactly is being lost. Most experiments provide compelling evidence that at some point the erosion of biodiversity will impair ecosystem function and stability. However,

these same experiments also show that a great deal of biodiversity can typically be lost with minimal effects. The value of biodiversity may well be revealed only on huge time scales that incorporate extremely infrequent but dramatic environmental challenges. If this is the case, standard short-term experiments will be unable to document the value of many species. Clearly, any failure of short-duration, small-scale experiments to identify a function for total biodiversity or for every species should not be used as a disingenuous argument to excuse human-caused extinctions. *See* EXTINCTION (BIOLOGY). Michelle A. Marvier

Bibliography. F. S. Chapin III et al., Ecosystem consequences of changing biodiversity, *Bioscience*, 48:45–52, 1998; M. A. Huston, *Biological Diversity: The Coexistence of Species on Changing Landscapes*, Cambridge University Press, 1994; C. G. Jones and J. H. Lawton (eds.), *Linking Species and Ecosystems*, Chapman & Hall, 1995; G. H. Orians, R. Dirzo, and J. H. Cushman (eds.), *Biodiversity and Ecosystem Processes in Tropical Forests*, Springer-Verlag, 1996.

Bioelectromagnetics

The study of the interactions of electromagnetic energy (usually referring to frequencies below those of visible light; **Fig. 1**) with biological systems. This includes both experimental and theoretical approaches to describing and explaining biological effects. Diagnostic and therapeutic uses of electromagnetic fields are also included in bioelectromagnetics.

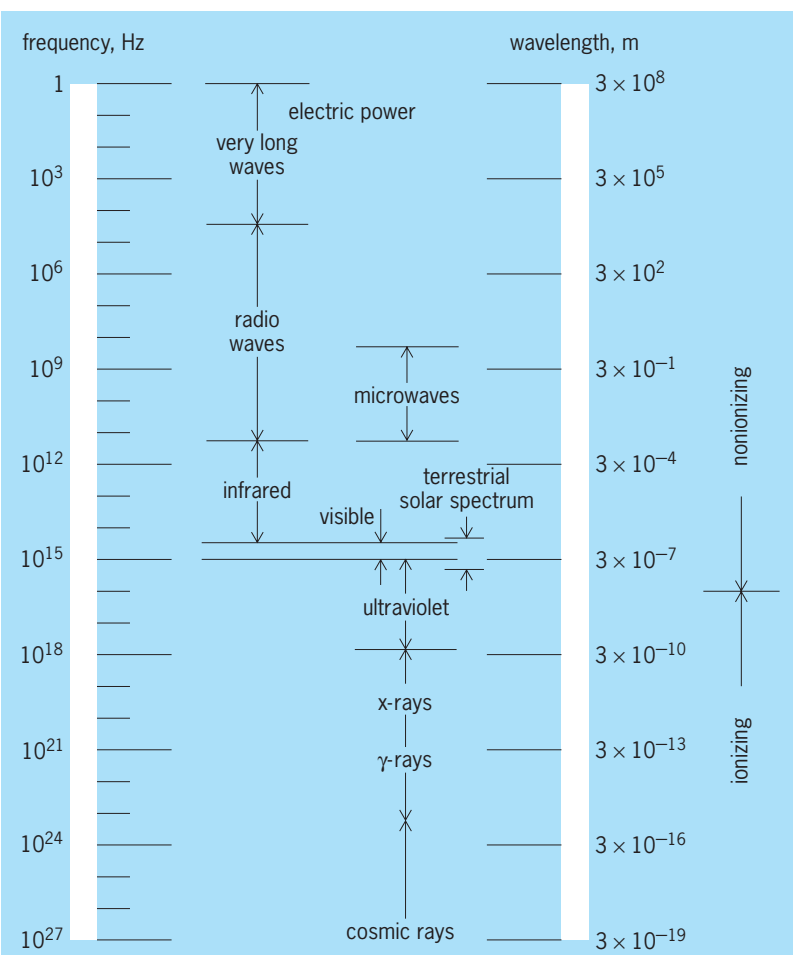

Fig. 1. Electromagnetic spectrum. Speed of light = 3×10^8 m/s = frequency (Hz) × wavelength (m).

Background. The interaction of electromagnetic fields with living organisms has intrigued both physicians and engineers since 1892 when J. A. d'Arsonval, a French physician and physicist, applied an electromagnetic field to himself and found that it produced warmth without muscle contraction. Subsequently, the use of electromagnetic energy to heat tissue became a common therapy, and in 1908 Nagelschmidt introduced the term diathermy to describe this process. During the 1930s "short-wave" diathermy (27 MHz) was in common use by physicians. World War II spurred the development of high-power microwave sources for use in radar systems. Shortly thereafter, concern over the safety of radar was voiced, leading to investigation of the biological effects of microwave radiation. Detailed study of the therapeutic potential of diathermy at microwave frequencies began after World War II as high-power equipment became available for medical and other civil applications.

Rapid growth in electronic systems for industrial, military, public service, and consumer use occurred during the 1970s and 1980s. Much of this equipment has the capability of emitting significant levels of electromagnetic radiation. The most extensive exposure to radio-frequency energy is from the thousands of transmitters authorized by the Federal Communications Commission (including commercial broadcast stations, cellular telephone transmitters, walkie-talkies, and microwave and satellite links). In addition, there are millions of unregulated citizen-band stations in use. The National Institute for Occupational Safety and Health estimates that more than 20,000,000 Americans are occupationally exposed to radio-frequency sources mainly from the heating and drying of plastics, textiles, wood products, and other manufactured goods. The use of electromagnetic fields in medicine increased as radio-frequency-induced hyperthermia was applied to cancer therapy. *See* AMATEUR RADIO.

Energy absorption. Electromagnetic energy is not absorbed uniformly across the geometric cross section of a biological organism. The total quantity of energy absorbed and the sites of maximum energy absorption depend on the frequency and polarization of the electromagnetic field, as well as on the electrical characteristics (the dielectric constant and conductivity—two properties of the tissue that control its interaction with electromagnetic radiation—which are frequency-dependent and vary with type of tissue), mass, and geometry of the absorbing object. In principle, the distribution of absorbed energy in an animal can be calculated from classical electromagnetic theory. However, the problem of energy distribution has not been solved for an object as complex as an animal. Simple calculations that assume the exposed system is of regular shape (for example, spheroidal) and of homogeneous composition allow some generalizations to be made. For an average person, maximal absorption (resonance) is predicted at

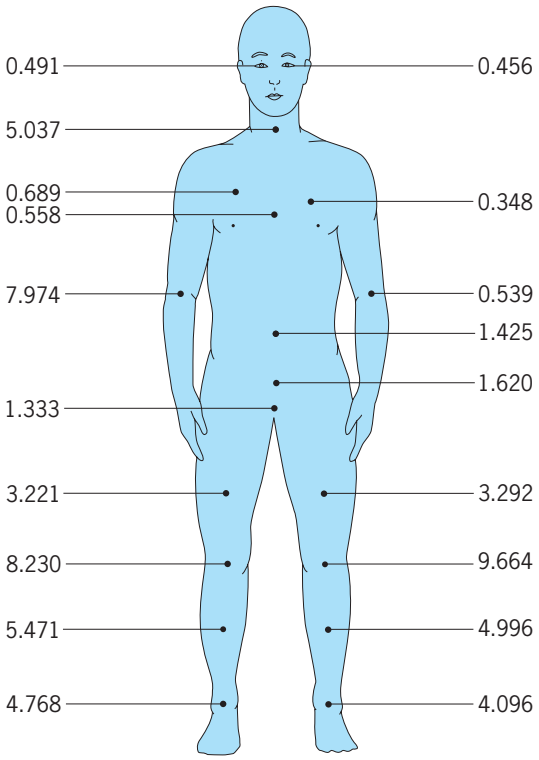

Fig. 2. Distribution of energy deposited in a human being exposed to electromagnetic radiation near the resonant frequency. The numbers provide an indication of the differences in energy absorbed by different parts of the body. (*After O. P. Gandhi, Dosimetry: The absorption properties of man and experimental animals, Ann. N.Y. Acad. Med., 55(11):999–1020, December 1979*)

approximately 70 MHz. When a complex target, such as a human being, is considered, resonant frequencies are also found in anatomically distinct portions of the body, such as head or leg. The internal energy absorption pattern has been calculated for several simple models: semi-infinite slab, homogeneous sphere, multishell sphere, and multiblock model of a human. While none of these models actually simulates a real human being, local regions of higher-than-average levels of energy absorption have been calculated (**Fig. 2**). Development of instruments that do not interfere with the field permit measurement of partial body resonances and the pattern of energy deposition within some experimental systems. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION.

Biological effects. The induction of cataracts has been associated with exposure to intense microwave fields. Although heating the lens of the eye with electromagnetic energy can cause cataracts, the threshold for cataract production is so high that, for many species, if the whole animal were exposed to the cataractogenic level of radiation, it would die before cataracts were produced.

In 1961 it was reported that people can "hear" pulsed microwaves at very low averaged power densities (50 microwatts/cm^2). It is now generally accepted that the perceived sound is caused by elastic-stress waves which are created by rapid thermal expansion of the tissue that is absorbing microwaves. The temperature elevation occurs in 10 microseconds, so that the rate of heating is about 1.8°F/s (1°C/s). However, the temperature increase is only about 0.00009°F (0.00005°C).

There are reports that microwave irradiation at very low intensities can affect behavior, the central nervous system, and the immune system, but many of these reports are controversial. Animals exposed to more intense electromagnetic fields that produce increases in body temperature of 1.8°F (1°C) or higher (thermal load equal to one to two times the animal's basal metabolic rate) demonstrate modification of trained behaviors and exhibit changes in neuroendocrine levels. Exposure of small animals to weaker fields has been shown to produce some changes in the functioning of the central nervous system and the immune system. While the mechanism is not yet known, a thermal hypothesis is not ruled out. Reports of the effect of very low-level, sinusoidally modulated electromagnetic fields on excitable cell and tissue systems have raised questions about the basic understanding of how those systems function.

Exposure of pregnant rodents to intense electromagnetic fields can result in smaller offspring, specific anatomic abnormalities, and an increase in fetal resorption. However, fields that produce significant heating of the gravid animal have been shown to be teratogenic.

Most biological effects of microwaves can be explained by the response of the animal to the conversion of electromagnetic energy into thermal energy within the animal. However, a few experiments yield results that are not readily explained by changes of temperature.

In an industrial society, an appreciation of the effects of stationary electric and magnetic fields, and of extremely low-frequency fields are important because of the ubiquitous nature of electricity. When the body is in contact with two conductors at different potentials, current flows through it. Typical adult-human thresholds for 50- or 60-Hz currents are:

Reaction	*Total body current*
Sensation	1 mA
"Let go"	10 mA
Fibrillation	100 mA

The "no contact" case, such as that experienced by an individual or animal under a high-tension transmission line (the field strength directly under a 765-kV line is 4–12 kV/m), has been investigated under controlled experimental conditions.

The possibility of hazard from occupational exposure to 50- or 60-Hz electric and magnetic fields has not been documented and is a subject of debate. Epidemiological studies have been undertaken to determine the health implications for workers and the general public exposed to these possible hazards. In addition, laboratory experiments have been designed to study the interactions of electric and magnetic fields with biological systems.

Some bacteria swim northward in stationary magnetic fields as weak as 0.1 gauss (the Earth's magnetic field is about 0.5 gauss at its surface). These bacteria contain iron organized into crystals of

Activation energies of molecular effects in biological systems*

Effect	Activation energy			Radiation parameters	
	kcal/mole	kJ/mole	eV	Frequency, GHz	Wavelength, μm
Thermal or brownian motion (at 30°C)	0.60	(2.50)	0.026	6.3×10^3	47.6
Ionization	230	(961)	10	2.4×10^6	0.12
Covalent bond disruption	115	(481)	5	1.21×10^6	0.25
London–van der Waals interactions	23	(96)	1	2.4×10^5	1.25
Hydrogen bond disruption	1.8–4.6	(7.5–14.2)	0.08–0.2	1.9×10^4–4.8×10^4	15.8–6.25
Proton tunneling	16.1	(67.3)	0.7	1.71×10^5	1.76
Disruption of bound water	12.9	(53.9)	0.56	1.4×10^5	2.14
Rotation of polar protein molecules	0.92–9.2	(3.8±38)	0.04–0.4	9.7×10^3–9.7×10^4	30.9–3.1
Reversible conformational changes in protein molecules	9.2	(38)	0.4	9.7×10^4	3.1
Charge transfer interaction	138–69	(577–288)	6.3	1.45×10^6–7.25×10^5	0.2–0.4
Semiconduction	23–69	(96–288)	1–3	2.4×10^5–7.25×10^5	1.2–0.41
Microwave radiation	2.7×10^{-6}	(1.12×10^{-7})	1.2×10^{-7}		
	0.03	(0.12)	1.2×10^{-3}	0.03±300	10^7–10^3

*From S. F. Cleary, Uncertainties in the evaluation or the biological effects of microwave and radiofrequency radiation, *Health Phys.*, 25:387–404, 1973.

magnetite. Magnetite is also found in the brains and Harderian glands of birds that can use local variations in the Earth's magnetic field for navigation and orientation. *See* BIOMAGNETISM; MICROWAVE; RADIATION BIOLOGY.

Mechanisms. The energy associated with a photon of radio-frequency radiation is orders of magnitude below the ionization potential of biological molecules (see **table**). Covalent bond disruption, London–van der Waals interactions, and hydrogen-bond disruption, as well as disruption of bound water or reversible conformational changes in macromolecules, all require more energy that is contained in a single microwave photon. There is a possibility that absorption of microwaves (or slightly shorter millimeter waves) may produce vibrational or torsional effects in bio-macromolecules. Thus, if radio-frequency electromagnetic fields have a specific action on molecules that alter their biological function, it will be through a route more complicated and less understood than that associated with ionizing radiation.

The most common mechanism by which electromagnetic fields interact with biological systems is by inducing motion in polar molecules. Water and other polar molecules experience a torque when an electric field is applied. In minimizing potential energy, the dipole attempts to align with the ever-changing electric field direction, resulting in oscillation. Both free and oriented (bound) water undergo dielectric relaxation in the radio-frequency region. The excitation of water, or other polar molecules, in the form of increased rotational energy is manifest as increased kinetic energy (elevation of temperature), but molecular structure is essentially unaltered if elevations are not excessive.

Alternating electromagnetic fields can cause an ordering of suspended particles or microorganisms. This effect, often called pearl-chain formation because the particles line up like a string of beads, results from a dipole-dipole interaction. Nonspherical particles may also be caused to orient either parallel or perpendicular to an applied field. These effects have been observed only at a very high field strength.

Electromagnetic energy absorbed by biological material can be converted into stress by thermal expansion. This phenomenon is caused by a rapid rise of temperature either deep within or at the surface of the material, and thus creates a time-varying thermal expansion that generates elastic stress waves in the tissue.

Medical applications. The therapeutic heating of tissue, diathermy, has been used by physicians for many years. "Short-wave" diathermy has been assigned the frequencies 13.56, 27.12, and 40.68 MHz, while microwave diathermy has been assigned 915, 2450, 5850, and 18,000 MHz. Short-wave diathermy provides deeper, more uniform heating than does diathermy at higher frequencies. *See* THERMOTHERAPY.

High-intensity radio-frequency fields have been used to produce hyperthermia in cancer patients. If a part of the body containing a tumor is heated, the difference of temperature between tumor and surrounding tissue can be increased, at times producing elevations to tumoricidal temperatures of 110–122°F (43–45°C) within the tumor while the surrounding tissue is below the critical temperature at which normal cells are killed. In addition, radio-frequency hyperthermia is often used in combination with x-ray therapy or with chemotherapy. In these cases, the

tumor temperature is kept at 114–116°F (41–42°C) to enhance the effectiveness of radiation or chemotherapy.

The development of bone tissue (osteogenesis) can be stimulated electrically either with implanted electrodes or by inductive coupling through the skin. The noninvasive technique uses pulsed magnetic fields to induce voltage gradients in the bone. This therapy has been used successfully to join fractures that have not healed by other means.

Electromagnetic fields were first used for medical diagnosis in 1926 when the electrical resistance across the chest cavity was used to diagnose pulmonary edema (water in the lungs). At frequencies below 100 kHz the movement of ions through extracellular spaces provides the major contribution to conductivity through the body. Thus, fluid-filled lungs can be detected by their lower resistivity. *See* RADIOLOGY.

Another diagnostic tool, magnetic resonance imaging, uses the behavior of protons, or other nuclei, in an electromagnetic field to obtain images of organs in the body. These images are as good as, and sometimes better than, those obtained with x-ray computed tomography. Magnetic resonance imaging is especially useful when viewing the head, breast, pelvic region, or cardiovascular system. In magnetic resonance imaging, a static magnetic field is applied to align nuclear magnetic moments, which then precess about the field direction with a characteristic frequency. When a radio-frequency field is applied transverse to the direction of the magnetic field, nuclei are driven to the antiparallel state. As they return to the ground state, they radiate at their resonant frequency. The strength of the signal is proportional to the concentration density of the nuclei being studied. Other information for image analysis can be obtained from the relaxation time constants of the excited nucleus. *See* COMPUTERIZED TOMOGRAPHY; MEDICAL IMAGING.

Internally generated fields associated with nerve activity (electroencephalography) and with muscle activity (electrocardiography, magnetocardiography) are used to monitor normal body functions. There may be other uses of electric currents or fields in growth differentiation or development which have not yet been explored. *See* CARDIAC ELECTROPHYSIOLOGY; ELECTROENCEPHALOGRAPHY; ELECTROMAGNETIC RADIATION; ELECTROMYOGRAPHY.

Electric fields now play a role in biotechnology. Intense, pulsed electric fields (about 60,000 V/m) produce short-lived pores in cell membranes by causing a reversible rearrangement of the protein and lipid components. This permits the entrance of deoxyribonucleic acid (DNA) fragments or other large molecules into the cell (electroporation). The fusion of cells using electric fields (electrofusion) is easy to control and provides high yields. A weak, nonuniform alternating-current field is used to bring the cells into contact (dielectrophoresis), and then a brief, intense direct-current field is applied, causing the cell membranes to fuse.

Elliot Postow

Bibliography. S. M. Michaelson and J. C. Lin, *Biological Effects and Health Implications of Radiofrequency Radiation*, 1987; C. Polk and E. Postow, *Handbook of Biological Effects of Electromagnetic Fields*, 2d ed., 1996.

Bioelectronics

A discipline in which biotechnology and electronics are joined in at least three areas of research and development: biosensors, molecular electronics, and neuronal interfaces. Some workers in the field include so-called biochips and biocomputers in this area of carbon-based information technology. They suggest that biological molecules might be incorporated into self-structuring bioinformatic systems which display novel information processing and pattern recognition capabilities, but these applications—although technically possible—are speculative.

It is an interesting reversal of the past in that bioelectronics uses the biological sciences in the development of the physical sciences and technologies, whereas the prior direction was to use the tools of the physical sciences to investigate and advance the biological sciences, life sciences, and medicine.

Of the three disciplines—biosensors, molecular electronics, and neuronal interfaces—which collectively constitute bioelectronics the most mature is the burgeoning area of biosensors. The term biosensor is used to describe two sometimes very different classes of analytical devices—those that measure biological analytes and those that exploit biological recognition as part of the sensing mechanism—although it is the latter concept which truly captures the spirit of bioelectronics. Molecular electronics is a term coined to describe the exploitation of biological molecules in the fabrication of electronic materials with novel electronic, optical, or magnetic properties. Finally, and more speculatively, bioelectronics incorporates the development of functional neuronal interfaces which permit contiguity between neural tissue and conventional solid-state and computing technology in order to achieve applications such as aural and visual prostheses, the restoration of movement to the paralyzed, and even expansion of the human faculties of memory and intelligence. The common feature of all of this research activity is the close juxtaposition of biologically active molecules, cells, and tissues with conventional electronic systems for advanced applications in analytical science, electronic materials, device fabrication, and neural prostheses.

Biosensors. A biosensor is an analytical device that converts the concentration of an analyte in an appropriate sample into an electrical signal by means of a biological sensing element intimately connected to, or integrated into, a transducer. Biosensors differ from existing analytical technologies in several important respects. First, there is intimate contact between the biological component, whether it be an enzyme, sequence of enzymes, organelle, whole cell, tissue slice, antibody, or other receptor or

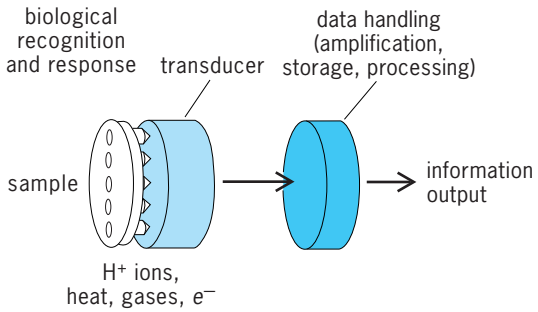

Fig. 1. Schematic representation of the essential features of a typical biosensor. (*After I. J. Higgins and C. R. Lowe, Introduction to the principles and applications of biosensors, Phil. Trans. Roy. Soc. Lond., B316:3–11, 1987*)

binding protein, and the transducer. Second, most new-generation biosensors are functionally small in size, thereby permitting small sampling volumes with minimum interruption of bodily functions following implantation, or, if used with process streams, following insertion in-line. Third, the biological material may be tailored to suit the medical or industrial needs and operate at various levels of specificity. Finally, biosensors are simple to use; they are single-step, reagentless devices which are inexpensive, disposable, and fully compatible with conventional data-processing technologies.

Figure 1 illustrates the general principle of a biosensor. Intimate contact between the biological and electrochemical systems is usually achieved by immobilization of the biosensing system on the transducer surface by physical restraint behind a polymer membrane or within a gel matrix, by chemical cross-linking with a bifunctional agent, or by direct covalent attachment. The biological system is responsible for the specific recognition of the analyte and subsequently responds with a concomitant change in a physicochemical parameter associated with the interaction. For example, if the biological interaction results in a change in pH, uptake or release of gases, ions, heat, or electrons, or a perturbation of an optical parameter proximal to the transducer, the biological signal may be converted into an electrical signal prior to being amplified, digitized, and outputted in the desired format.

Transducers. These are physicochemical devices that respond to the products of the binding or biocatalytic process. The choice of the most appropriate transducer configuration will be conditioned largely by the nature of the biocatalyst system, the secondary products to be monitored, and the potential application of the final device. The ideal transducer should display a moderately fast response time, typically 1–60 s, be amenable to the facile fabrication and miniaturization, be reliable, and should compensate for adverse environmental effects such as temperature dependency and drift. *See* TRANSDUCER.

One of the best-known transducers is the potentiometric ion-sensitive electrode which measures the accumulation of charge density at the electrode surface under equilibrium conditions. The potential of an ion-sensitive electrode is a logarithmic function of ionic activity, with a 59.2-mV change in electrode potential per tenfold change in concentration of a monovalent ion. Since the late 1970s, major improvements in electrode design and performance have been achieved, and ion-sensitive electrodes have been developed for monitoring a variety of cations, anions, metabolites, solvents, gases, mutagens, amino acids, drugs, and other analytes. This diversity has been made possible by coupling enzymes; whole plant, animal, and microbial cells; organelles; tissues; and immuno-enzymatic assay systems to conventional ion-sensitive electrodes to generate appropriate enzyme and immuno-enzyme electrodes. *See* ION-SELECTIVE MEMBRANES AND ELECTRODES.

The analyte-sensitive membranes of the ion-sensitive electrodes have been integrated with monolithic solid-state field-effect transistor (FET) technology to introduce a range of ion-selective and substrate-specific FETs. For example, an FET sensitive to penicillin has been constructed which responded to penicillin concentrations up to 50–60 millimoles in less than 30 s, displayed a lifetime of approximately 2 months, and permitted automatic compensation for temperature and ambient pH. The device comprised a matched pH-responsive ion-selective FET pair with an active gate film of cross-linked albumin-penicillinase and a reference gate film of cross-linked albumin. The enzyme present in the active gate polymer catalyzed the hydrolysis of penicillin with the concomitant release of hydrogen ion (H^+) and subsequent detection in the silicon nitride (Si_2N_4) gate material. When operated in differential mode, the sensor pair was reported to be relatively free from thermal effects and from changes in the ambient pH of the analyte solution. However, the buffer capacity of the analyte was found to have a profound influence on the sensitivity, range, and linearity of the response.

Similar limitations were experienced with other enzyme-modified FET devices, such as those responsive to glucose, urea, acetylcholine, adenosine triphosphate (ATP), and lipid, with the complementary enzymes glucose oxidase, urease, acetylcholinesterase, adenosine triphosphatase (ATPase), and lipase immobilized to pH-responsive Si_3N_4 or iridium oxide (Ir_2O_3) gate materials. A co-integrated coulometric feedback system may circumvent these limitations. The electrolysis of water at a noble-metal electrode spatially positioned close to a urea-sensitive FET generates H^+, which can be used to balance the uptake engendered by the enzyme activity. However, in addition to the problems encountered by immobilization of small quantities of enzyme for the determination of substrates in samples containing a high buffering capacity, the difficulty in fabricating small, well-defined immobilized enzyme membranes on the ion-selective gate of an FET is a hurdle which must also be circumvented prior to manufacturing monolithic enzyme-modified FETs. *See* TRANSISTOR.

Photoactivatable *p*-nitrophenylazides, photosensitive polyvinylalcohol with pendant stilbazolium groups as a photo-cross-linkable entrapment matrix,

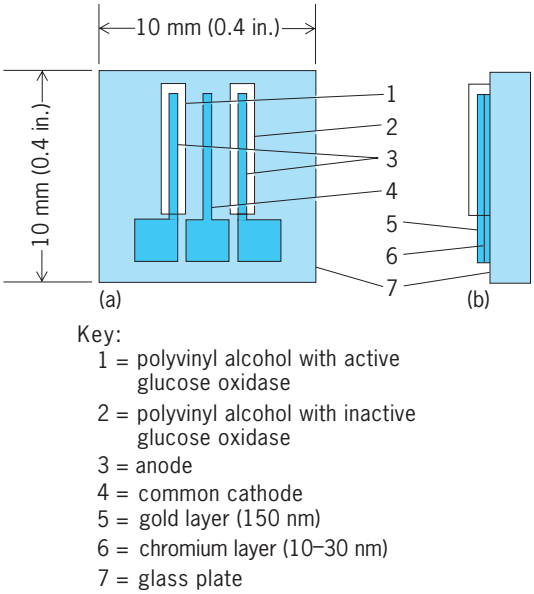

Fig. 2. Configuration of a thin-film glucose biosensor, (*a*) front view and (*b*) side view, and glucose oxidase-immobilized polyvinyl alcohol membranes. (*After T. Moriizumi, Solid state biosensors, Proceedings of International Symposium on Future Electron Devices, Tokyo, pp. 73–80, 1985*)

and piezoelectric ink-jet devices have all been used successfully to generate active small-area enzyme membranes. For example, it has been demonstrated that it is feasible to generate monolothic multienzyme-modified FET biosensors by using photolithographically patterned, enzyme-loaded polyvinyl alcohol films, and a triple-function, silicon-on-sapphire array, with an on-chip pseudo-reference electrode for the measurement of urea, glucose, and potassium ion (K^+). Thus, despite some unresolved problems, biosensors based on integrated solid-state devices display considerable potential for miniaturization of multifunction configurations.

Amperometric devices. Also known as current-measuring devices, these offer a wider scope of applications than potentiometric techniques and are linearly dependent on analyte concentration, thereby giving a normal dynamic range and normal response to errors in the measurement of current. A solid-state hydrogen peroxide (H_2O_2) sensor has been fabricated and used in a glucose sensor. **Figure 2** shows the electrode pattern on the sensor chip; it comprises three double-layer (150-nm-thick gold, 10 to 30-nm-thick chromium) electrodes deposited on a glass substrate with a bare gold-film constituting the central common cathode. Glucose oxidase-loaded membranes were spin-coated over one of the gold-film electrodes and photopolymerized, while denatured enzyme was deposited over the other pair of electrodes. The electrodes were connected to a potentiostat which followed the difference in anode currents from the active and reference electrodes caused by the reoxidation of enzymatically generated H_2O_2 when the anodes were biased at +0.8 V; with respect to the cathode. **Figure 3** shows a typical calibration curve for the miniaturized amperometric glucose sensor. Thin-film microsensors of this type are promising devices for H_2O_2 detection since they are widely applicable and display good sensitivity.

Solution conductance. An alternative measuring principle which is also widely applicable to biological systems is the exploitation of solution conductance. The development and operation of an accurate microelectronic conductance biosensor which operates in a differential mode by monitoring the change in conductance occasioned by the catalytic action of enzymes immobilized proximal to a planar microelectronic conductance cell of defined geometry has also been described. Each conductance device comprised a pair of serpentined and interdigitated parallel trimetallic (gold, platinum, titanium) tracks, 1000–3500 nm thick and 4000 nm wide, deposited onto an oxidized silicon surface by using standard sputter deposition processes or thermal evaporation under vacuum followed by appropriate photolithographic processes and metal etching. An enzyme-loaded cross-linked albumin membrane was cast over the interdigitated sample pair of electrodes, but not over the reference pair, so that subtraction of the amplified signals emanating from the conductimeter electronics continuously corrected for nonspecific variations in basal conductivity of the buffers and biological fluids in which measurements were being made. The output response of the differential enzyme microconductimeter with urease immobilized over the sample cell was linear up to 5 min at all urea concentrations in the range 0.1–10 millimoles. There was a linear correlation between serum urea concentrations determined with the microelectronic device and those determined by conventional procedures at a major hospital laboratory. The microelectronic enzyme conductimeter can be used with many other enzymes that create changes in solution conductance, thus providing a most useful and versatile addition to available analytical techniques.

Other measurement devices. Other on-chip measuring principles have also been exploited in the fabrication

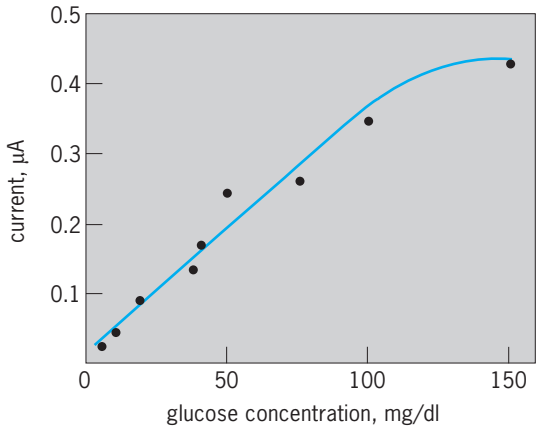

Fig. 3. Calibration curve for glucose of the thin-film biosensor at pH 7.0 and 37°C (98.6°F). (*After T. Moriizumi, Solid state blosensors, Proceedings of International Symposium on Future Electron Devices, Tokyo, pp. 73–80, 1985*)

of biosensors. For example, it has been demonstrated that two integrated-circuit temperature-sensitive devices, each composed of 3 Darlington connected *npn* transistors and a complementary metal-oxide semiconductor constant-current circuit, can be used for the differential calorimetric determination of glucose concentrations. The difference in steady-state output voltage of the enzyme-modified sensor compared to the unmodified sensor after addition of 2–100 millimoles glucose was related to the catalytically induced temperature change and thus to the glucose concentration.

Furthermore, miniature piezoelectric sensors, for example, quartz crystal and surface-acoustic-wave devices, can be used directly in aqueous solution as enzyme substrate and immunochemical sensors. Organophosphorus pesticides and parathion were detected with coatings of immobilized cholinesterase and anti-parathion antibodies respectively. Excellent reproducibilities, coating lifetimes, response times, and sensitivities were observed in the research. The most likely application of this type of biosensor technology lies in the detection of illicit drugs, of explosives in airport and post office surveillance, and of toxic pollutants.

Potential for the technology. Clearly biosensors have evolved into miniaturized, disposable, solid-state devices with the theoretical capacity to cointegrate the signal-processing and -conditioning circuitry directly on-chip and thereby obviate the requirement for traditional external instrumentation. A multifunction chip comprising an array of biologically sensitized gates deposited on a monolithic silicon chip could also incorporate sufficient signal-conditioning capability to interrogate each sensor in turn, assess the outputs, compare with a calibration standard, and release the information as the concentration of the analyte together with the date, batch code, expiration date, and operator code. However, the obstacles that remain from the commercial applications of this technology are formidable. Passivation of the sensitive electronic components to the rigors of aqueous solutions and the lability of the biologically active areas remain problems in the long-term acceptability of these devices.

Molecular electronics. Biology has many examples of organized structures on an intracellular, cellular, or intercellular level, and, at the molecular level, of complex biological analogs of conventional electronic data processing. Biological systems, in fact, perform all of the functions of interest to the modern electronics industry (sensing, input/output, memory, computation) as well as functions not yet achieved (rapid complex pattern recognition and learning). However, the lability and general prerequisite for an aqueous environment militate against the use of biological molecules as active components in nonbiological microsystems. It is more likely that the self-assembly properties of proteins will be exploited to form a template or matrix for the proper assembly of complex architectures from conventional electronic components. In addition, biological analogies are likely to suggest the development of novel structures and algorithms to achieve functions not readily accomplished by computing devices of present-day design.

Neuronal interfaces. The term biochip is sometimes used to describe an implantable system that would enable the interconnection of nervous tissue with conventional computer devices. However, like the construction of a biosensor, the fabrication of a functioning neuronal interface or artificial synapse will require the development of appropriate reversible chemical-to-electrical transduction processes. The development of a neuronal interface is thus likely to require greater knowledge about the chemical mechanisms which govern synaptic communication. *See* NEURON.

Christopher R. Lowe

Bibliography. N. C. Foulds and C. R. Lowe, Biosensors: Current applications and future potential, *Bioessays*, 3:129–132, 1985; C. R. Lowe, Biosensors, *Trends Biotechnol.*, 2:59–65, 1984; C. R. Lowe, An introduction to the concepts and technology of biosensors, *Biosensors*, 1:3–16, 1985; A. Sibbald, Recent advances in field-effect chemical microsensors, *J. Mol. Electr.*, 2:51–83, 1986.

Bioeroding sponges

Marine sponges that are able to erode calcareous substrates such as coral, mollusk shells, and limestone, also called boring or excavating sponges. Confirmed eroders belong to the families Samidae, Thoosidae, Clionaidae, Spirastrellidae, Acarnidae, and Phloeodictyidae, which are all demosponges. The best-known species belong to the genera *Cliona*, *Pione*, and *Aka*. Species are identified by their skeletal elements, morphology, tissue characteristics, and erosion traces. *See* DEMOSPONGIAE.

Description and characteristics. Bioeroding sponges occur in three growth forms: alpha, beta, and gamma (**Fig. 1**). Sponges in alpha form live almost entirely inside the substrate and fill the cavities and tunnels they produce. They reach the surrounding water with small tubes of tissue protruding from the substrate surface, the fistules (centimeter scale; Fig. 1*a*) or papillae (millimeter scale; Fig. 1*b*). The beta form additionally covers the substrate surface with a coherent layer of tissue (Fig. 1*c*). Gamma specimens resemble noneroding sponges and can be massive or cup-shaped. They may contain remnants of calcareous matter. The growth forms are thought to represent stages during the life of some bioeroding sponges, but by far most species occur only in the alpha form.

The life inside hard substrate provides the sponge with shelter against predators and adverse conditions such as desiccation at low tide or extreme temperatures. The sponge larva settles on a suitable surface and starts to erode into the substrate. In order to grow, the sponge forms or enlarges chambers and tunnels (**Fig. 2***a*). Specialized cells with fine, thread-like extensions secrete chemicals, which dissolve

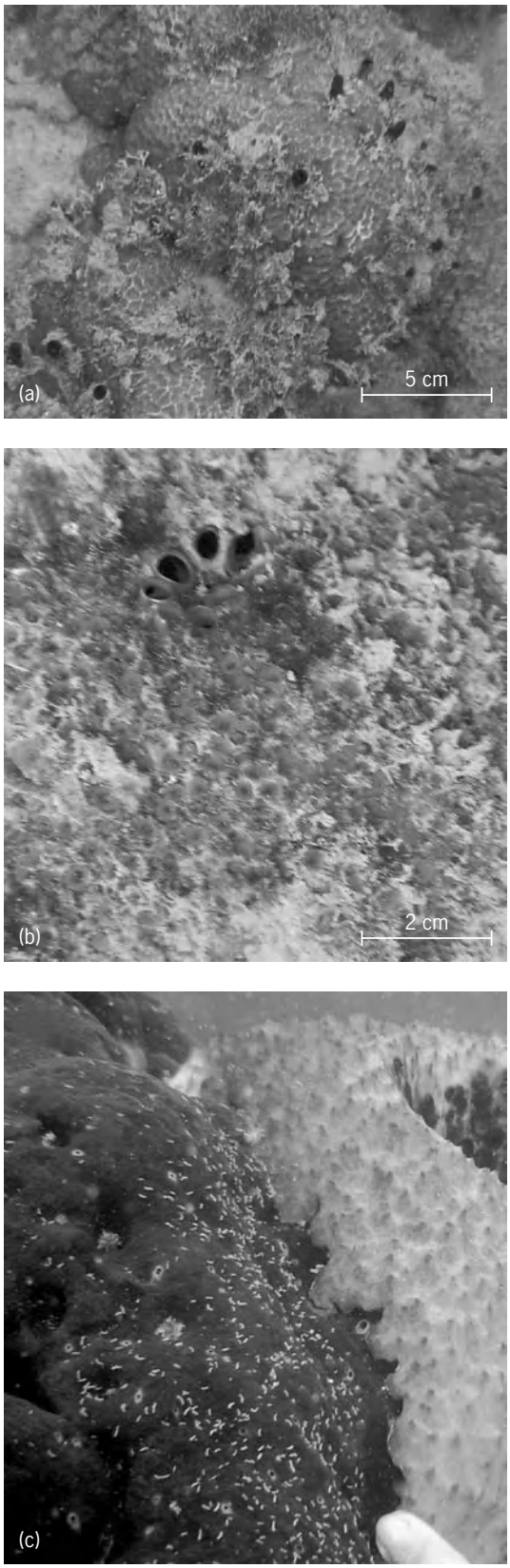

Fig. 1. Growth forms of bioeroding sponges. (*a*) Alpha form. Fistules of *Aka mucosa* (Phloeodictyidae) breaking through the surface of a live coral. (*b*) Alpha form. *Cliona celata* papillae (Clionaidae). Sievelike inhalant papillae, four open exhalant papillae in upper half of photograph. (*c*) Beta form. *Cliona orientalis* (Clionaidae, left) competing with a free-living sponge. Lighter structures on the surface of *C. orientalis* are exhalants. Finger in lower right for scale.

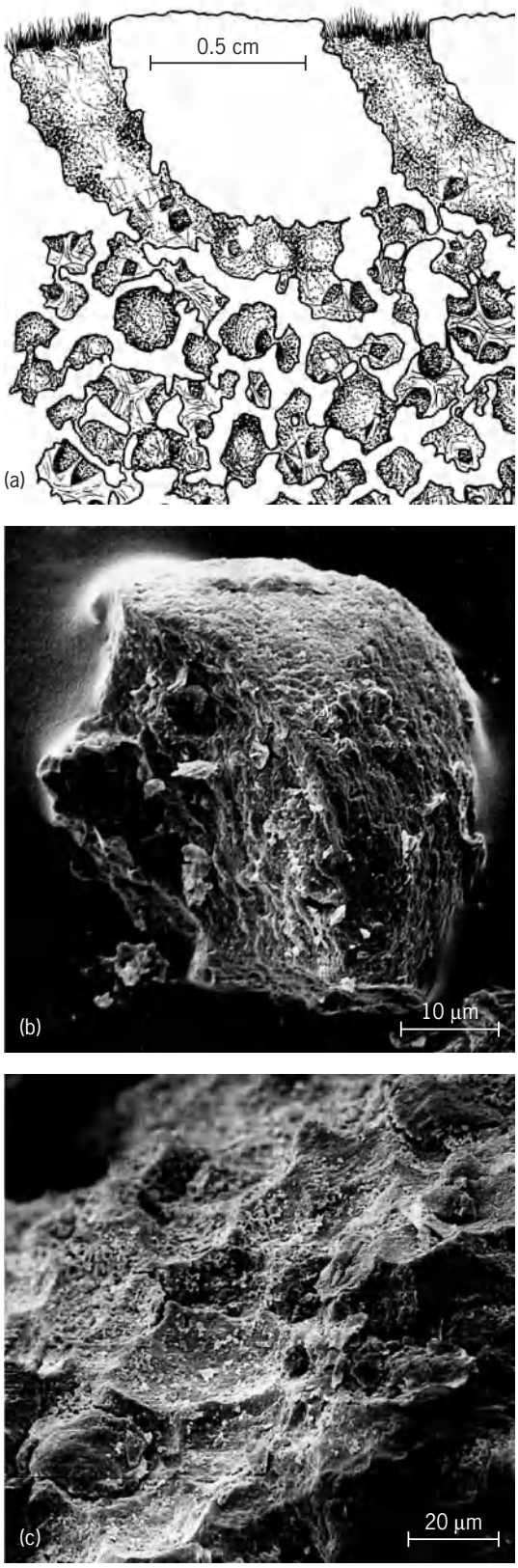

Fig. 2. Sponge bioerosion. (*a*) Cross section through a colony of *Cliona celata* in coral. Gray areas are bioerosion chambers filled with sponge tissue and pin-shaped skeletal elements. The latter are arranged in pallisade in the two papillae that break through the substrate surface (*drawing courtesy of Sula Blake*). (*b*) Sponge chip from *C. orientalis*. (*c*) Pitted structure of chamber walls in a clam shell eroded by *C. orientalis*. In lower left and upper right corners, sponge chips are still in the process of being cut.

2–3% of the eroded material. The sponge etches cup-shaped fissures into the material, cutting out lentil-shaped chips with a diameter of 20–100 μm (Fig. 2*b*). The activity of many cells results in a typical pitted structure on the walls of the eroded material (Fig. 2*c*). The silt-sized chips are expelled with the water that is pumped through the sponge and can contribute 30–40% of the sediments on tropical coral reefs. Erosion rates depend on environmental factors such as nutrient and light levels, currents, and substrate density.

Effects of bioerosion. Bioeroding sponges recycle calcareous materials, create new three-dimensional microhabitats in which other organisms can hide, and enhance coral dispersal by fragmentation. However, coral fragments may also be moved to environments with unfavorable conditions. In normal circumstances, coral reef growth is larger or equal to bioerosion. However, this balance may shift due to environmental disturbances such as overfishing of predators, increased nutrient levels, and the reduction of live surface cover. High levels of sponge bioerosion on coral reefs would result in an altered carbonate budget, a significant decline of three-dimensional structure provided by healthy coral; a reduction of biodiversity of organisms that need a high degree of structural diversity, including commercially used species; and the destruction of what we perceive as a beautiful reef, and consequently losses for the tourism industry. *See* REEF.

Bioeroding sponges are also destructive parasites of mollusks. Oysters in aquaculture are coated with protective paint or regularly cleaned to prevent the sponge larvae settling and developing on the shells. *See* MOLLUSCA.

Diversity and range. The guild of bioeroding sponges presently consists of about 250 valid species. The highest biodiversity of bioeroding sponges is recorded from tropical coral reefs in the Indian Ocean. Slightly fewer are known from the Atlantic and Pacific oceans and only two or three species from polar waters. The sponges mainly occur in shallow water 0–50 m (0–165 ft) deep, but have been sampled from as deep as 2030 m (6300 ft) [*Alectona mesatlantica*, Thoosidae]. The fossil record confirms the occurrence of bioeroding sponges as early as the Devonian (*Entobia devonica*, in fossil stromatoporoids and bivalves).

Christine Schönberg

Bibliography. J. N. A. Hooper and R. W. M. van Soest (eds.), *Systema Porifera: A Guide to the Classification of Sponges*, vol. 1, *Introductions and Demospongiae*, Kluwer Academic/Plenum, New York, 2002; K. Rützler and C. Rieger, Sponge burrowing: Fine structure of *Cliona lampa* penetrating calcareous substrata, *Mar. Biol.*, 21:144–162, 1973; C. H. L. Schönberg, Growth and erosion of the zooxanthellate Australian bioeroding sponge *Cliona orientalis* are enhanced in light, *Proc. 10th Int. Coral Reef Symp., Okinawa*, Japan, in press; C. H. L. Schönberg, Substrate effects on the bioeroding demosponge *Cliona orientalis*, 1. Bioerosion rates, *PSZNI: Mar. Ecol.*, 23(4):313–326, 2002.

Biofilm

An adhesive substance, the glycocalyx, and the bacterial community which it envelops at the interface of a liquid and a surface. When a liquid is in contact with an inert surface, any bacteria within the liquid are attracted to the surface and adhere to it. In this process the bacteria produce the glycocalyx. The bacterial inhabitants within this microenvironment benefit as the biofilm concentrates nutrients from the liquid phase. However, these activities may damage the surface, impair its efficiency, or develop within the biofilm a pathogenic community that may damage the associated environment. Microbial fouling or biofouling are the terms applied to these actual or potential undesirable consequences.

Microbial fouling affects a large variety of surfaces under various conditions. Microbial biofilms may form wherever bacteria can survive; familiar examples are dental plaque and tooth decay. Dental plaque is an accumulation of bacteria, mainly streptococci, from saliva. The attachment of bacteria to the tooth is not random. Certain types of bacteria are specific to certain locations on the tooth: anaerobic bacteria are found deep within the biofilm, whereas aerobes are found near the plaque-air interface. The ability of the biofilm to concentrate nutrients is not important in this environment because the mouth itself is rich in nutrients; rather, the plaque protects the bacterial inhabitants from being transported by swallowing to a more hostile environment in the stomach. The process of tooth decay begins with the bacteria colonizing fissures in and contact points between the teeth. Dietary sucrose is utilized by the bacteria to form extracellular glucans that make up the glycocalyx and assist adhesion to the tooth. Within this microbial biofilm or plaque the metabolic by-products of the bacterial inhabitants are trapped; these include acids that destroy the tooth enamel, dentin, or cementum. *See* PERIODONTAL DISEASE; TOOTH DISORDERS.

Nature and formation. The formation of a biofilm commences with the transportation of bacteria from the liquid phase to a contiguous surface. The mechanisms involved are sedimentation, chemotaxis, Brownian motion, cell surface hydrophobicity, and fluid dynamic forces. A hypothesis explaining the process of attraction postulates that different combinations of force affect the bacteria according to their distance from the surface. Van der Waals forces operate at a distance greater than 50 nanometers from the surface and give a rapid but weak attraction; both van der Waals and electrostatic interactions occur together between 10 and 20 nm from the surface. At a distance of less than 1.5 nm from the surface, van der Waals forces, electrostatic interactions, and specific interactions occur, producing an irreversible binding between the bacteria and the surface. This binding is associated with the activation of glycogen-producing genes. Then production by the bacteria of adhesive materials, such as exopolysaccharides, that form the glycocalyx begins. Not all of the different adhesive media are known. Many factors affect the initial colonization of the surface, including surface

roughness, the nature of the surface, the velocity of liquid next to the surface, and the temperature. *See* INTERMOLECULAR FORCES.

After initial attachment, the organisms grow and divide to form microcolonies within the glycocalyx. Microbial products and organic and inorganic matter are trapped within the microbial biofilm. This occurrence creates a concentration of nutrients in the biofilm. In addition, other microorganisms join and are protected within the glycocalyx. In this way, the structure of the microbial biofilm changes; as the biofilm matures, it is influenced by interrelated physical, chemical, and biological components. But the bacterial inhabitants exert the largest influence on biofilm development because they create a favorable environment and can alter the physical and chemical properties of the microbial biofilm.

Of less importance are the chemical and physical factors that influence biofilm development. Biofilm stability is promoted by the presence of calcium, magnesium, or iron, and the growth of the biofilm microorganisms is associated with the availability of nutrients. Physical factors include the volume and mass of the biofilm and the diffusion of gases, nutrients, and heat within it.

The concentration of bacteria within a microbial biofilm is associated with the concentration of nutrients to support growth. This is a useful response for survival in low-nutrient environments such as the so-called clean industrial processes, treated drinking water systems, and clear mountain streams. Microbial biofilms have been found in all such environments. In low-nutrient environments, such as an oil reservoir, studies show that bacteria adopt starvation responses, for example, they may become smaller, and that under long-term starvation regimes bacteria may produce less glycocalyx.

Microbial fouling. The process of microbial fouling involves the secretion by the biofilm inhabitants of metabolites that damage the surface or the production of organic or inorganic deposits upon the surface. The consequences of microbial fouling are (1) physical damage to the surface as a result of microbial growth and metabolite activity; (2) the reduction in the proper function of the surface because of the presence of the biofilm; and (3) the creation of a reservoir of potential pathogens within the biofilm. These consequences are not mutually exclusive. Microbial fouling in such places as drinking water, food industries, marine environments, and industrial water systems is of great concern, and it is therefore carefully monitored.

Drinking water and food industries. In these industries there is considerable concern with the potential for the microbial biofilm to act as a reservoir for potential pathogens. The presence of pathogens is often detected by sampling the bacteria in the liquid phase. This approach, however, neglects the fact that bacteria are concentrated within the biofilm, so that readings from the liquid sample may be misleading. In such cases, contamination is detected only when pathogens become detached from the mature biofilm and appear in the liquid phase.

In the water industry, biofouling is difficult to monitor and control because the supply system is made up of an extensive network of treatment plants and delivery pipes. Thus the absence of pathogens within a biofilm in one part of the system does not mean that other parts are not contaminated. The use of copper pipes in distribution does not lessen the chances of biofouling; at one time it was considered that bacteria would not grow on such materials, but later research showed that this was not the case. For example, a study of a drinking water utility showed that bacterial counts were above permitted levels. Samples of biofilms were taken as pipe scrapings from various parts of the system. While the treatment plant was found to contain bacteria within permitted levels, unacceptable levels were found in biofilms in the distribution system and along the way to the consumer. In addition, treatment by free chlorine did not control the problem because the pathogens were protected from the antimicrobial agents by the biofilms. *See* WATER SUPPLY ENGINEERING.

In the food industry the food comes into contact with a large number of surfaces during processing and packaging. This situation raises concerns regarding the prevention of contamination that are similar to those in the water supply industry. Some research has been carried out to find processes to reduce the opportunities for bacterial detachment from stainless steel or nylon; results so far, however, have been inconclusive because the type of surface and the type of bacteria studied were shown to affect the possibility of detachment. *See* FOOD MICROBIOLOGY.

Marine environments. Microbial fouling in seawater begins with the submerged surface absorbing organic material from the water before being colonized by a succession of different bacteria. The final event is confluent glycocalyx production. The biofouling is generally followed by macrofouling as various creatures are attracted to the surface to feed off the biofilm.

The fouling of hulls reduces vessel speed and increases the cost of transport. The fouling of metal marine structures may result in corrosion. Sulfate in seawater acts as an electron acceptor for sulfate-reducing bacteria. This leads to the production of sulfide, which enhances the corrosion of the surface. Antifouling paints are not always effective and may be damaging to the marine environment. Treated surfaces can be colonized within 2 weeks of submersion, and the glycocalyx may protect the biofilm inhabitants from the paint.

Industrial water systems. Water is used in industry for cleaning, as a coolant, or for a specific part of a process. The large surface areas associated with cooling systems and heat exchangers, such as radiators, give ample opportunity for biofouling, provided the water supports bacteria. This is not always so; pure water is used in the pharmaceutical industry to combat biofouling, but there is a risk of subsequent contamination.

Metal corrosion is the result of an electrochemical process whereby an anode site on the metal yields a supply of electrons that travel through the metal

to a cathode site elsewhere on the surface. This process forms anodic pits and cathodic depositions on the surface. Microorganisms are attracted to and colonize a metal surface, and activity within the biofilm generates conditions that are favorable for the sulfate-reducing bacteria, anaerobic bacteria that are the principal corroding organisms. That is, aerobic microorganisms in the upper biofilm deplete oxygen, producing anaerobic conditions in the lower biofilm contiguous to the metal surface. Sulfate-reducing bacteria produce hydrogen sulfide, which is highly corrosive, but in addition the electrochemical corrosion process is intensified by the activity of sulfate-reducing bacteria. The bacteria utilize hydrogen by means of hydrogenase enzymes, and this generates the anodic reaction. In addition, within the biofilm, production of bacterial metabolites and biopolymers with differing electrical charges and composition is sufficient to develop a corrosion potential. This process of corrosion by pitting of the metal means that failure tends to occur at points on the surface. *See* CORROSION. Hilary Lappin-Scott; J. W. Costerton

Bibliography. W. G. Characklis and P. A. Wildrer (eds.), *Structure and Function of Biofilms*,1989; M. Fletcher and B. Capdeville (eds.), *Biofilms: Science and Technology*, vol. 223, 1992; S.P. Gorman and M. Sussman (eds.), *Microbial Biofilms: Formation and Control*, 1994; M. W. Mittleman and G. G. Geesey, *Biological Fouling of Industrial Water Systems*: *A Problem Solving Approach*, 1987; D. C. Savage and M. Fletcher, *Bacterial Adhesion Mechanisms and Physiological Significance*, 1985.

Biogeochemistry

The study of the cycling of chemicals between organisms and the surface environment of the Earth. The chemicals either can be taken up by organisms and used for growth and synthesis of living matter or can be processed to obtain energy. The chemical composition of plants and animals indicates which elements, known as nutrient elements, are necessary for life. The principal element composition of land plants is shown in the **table**. The most abundant nutrient elements, carbon (C), hydrogen (H), and oxygen (O), supplied by the environment in the form of carbon dioxide (CO_2) and water (H_2O), are usually present in excess. The other nutrient elements, which are also needed for growth, may sometimes be in short supply; in this case they are referred to as limiting nutrients. The two most commonly recognized limiting nutrients are nitrogen (N) and phosphorus (P).

Biogeochemistry is concerned with both the biological uptake and release of nutrients, and the transformation of the chemical state of these biologically active substances, usually by means of energy-supplying oxidation-reduction reactions, at the Earth's surface. Emphasis is on how the activities of organisms affect the chemical composition of natural waters, the atmosphere, rocks, soils, and sediments. Thus, biogeochemistry is complementary to the science of ecology, which includes a concern with how the chemical composition of the atmosphere, waters, and so forth affects life. *See* ECOLOGY; OXIDATION-REDUCTION.

Elemental composition of plants*

Element	Concentration, % dry weight of tissue
Carbon	45
Oxygen	45
Hydrogen	6
Nitrogen	1.5
Potassium	1.0
Calcium	0.5
Phosphorus	0.2
Magnesium	0.2
Sulfur	0.1
Chlorine	0.01
Iron	0.01
Manganese	0.005
Zinc	0.002
Boron	0.002
Copper	0.0006
Molybdenum	0.00001

*From W. Stumm (ed.), *Global Chemical Cycles and Their Alterations by Man*, Dahlem Konferenzen, 1977.

The two major processes of biogeochemistry are photosynthesis and respiration. Photosynthesis involves the uptake, under the influence of sunlight, of carbon dioxide, water, and other nutrients by plants to form organic matter and oxygen. A generalized expression is shown in reaction (1), where the expression ($xN + yP + \cdots$) represents other nutrient elements in various chemical forms and $CH_2ON_xP_y\ldots$ represents organic matter. Respiration is the reverse of photosynthesis and involves the oxidation and breakdown of organic matter and the return of nitrogen, phosphorus, and other elements, as well as carbon dioxide and water, to the environment. *See* PHOTOSYNTHESIS; PLANT RESPIRATION.

$$CO_2 + H_2O + (xN + yP + \cdots) \rightarrow CH_2ON_xP_y + \cdots + O_2 \quad (1)$$

Biogeochemistry is usually studied in terms of biogeochemical cycles of individual elements. This gives rise to expressions such as the biogeochemical cycle of carbon and the terrestrial cycle of phosphorus. Time is an important consideration in such cycles: there are short-term cycles ranging from days to centuries and long-term (geological) cycles ranging from thousands to millions of years.

There has been increasing interest in biogeochemistry because the human influence on short-term biogeochemical cycling has become evident. Perhaps the best-known example is the changes in the biogeochemical cycling of carbon due to the burning of fossil fuels and the cutting and burning of tropical rainforests. The cycles of nitrogen and phosphorus have been altered because of the use of fertilizer and the addition of wastes to lakes, rivers, estuaries, and the oceans. Acid rain, which results from the addition of sulfur and nitrogen compounds to the atmosphere by humans, affects biological systems in

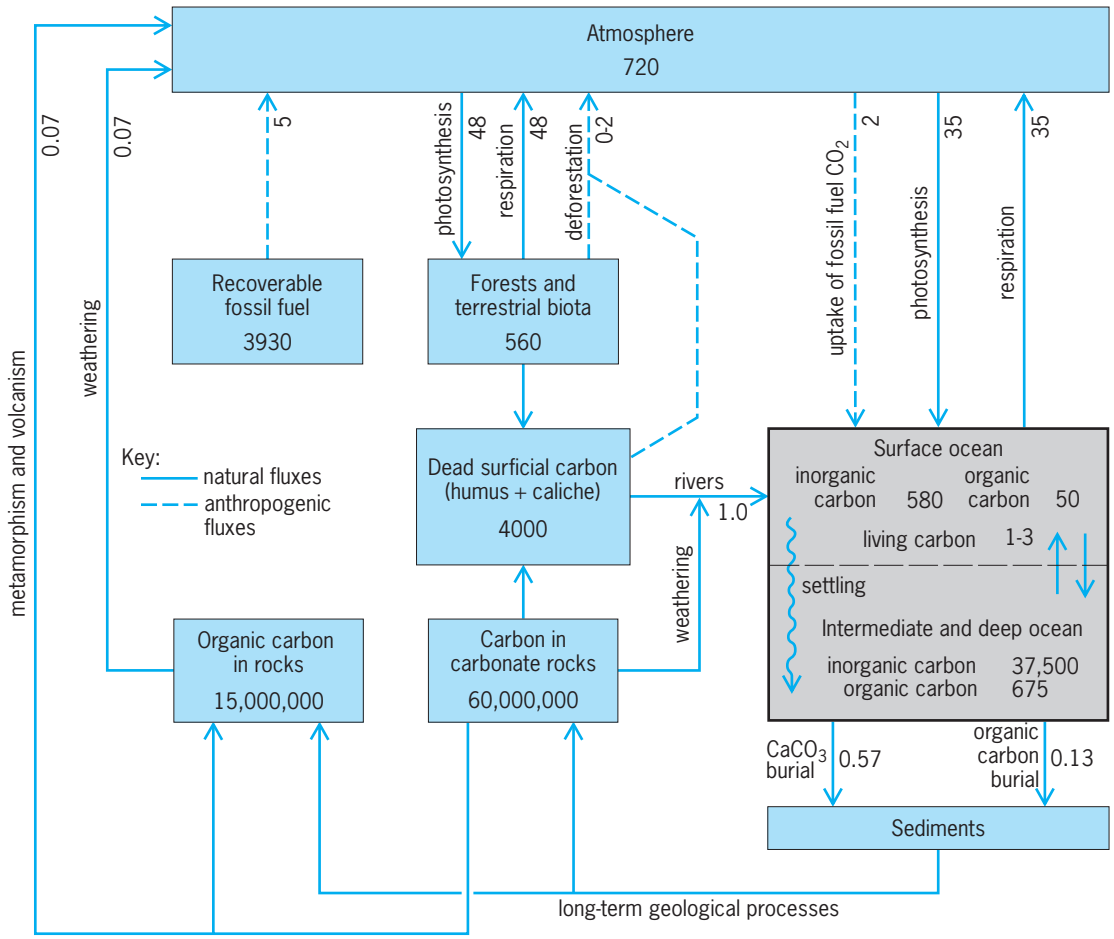

Fig. 1. Carbon cycle. Photosynthetic fluxes between the atmosphere and oceans, and the atmosphere and land, represent net primary productivity. Reservoir units, gigatons (10^{15} g) C; flux units, gigatons C/yr.

certain areas. A solid understanding of the biogeochemical cycles of the major nutrient elements is, therefore, basic to dealing with current and future problems caused by human impact on the environment. *See* HUMAN ECOLOGY.

Carbon cycle. Carbon is the basic biogeochemical element. The carbon cycle shown in **Fig. 1** provides a basis for understanding biogeochemical cycling. The atmosphere contains carbon in the form of carbon dioxide gas. There is a large annual flux of atmospheric carbon dioxide to and from forests and terrestrial biota, amounting to nearly 7% of total atmospheric carbon dioxide. This is because carbon dioxide is used by plants to produce organic matter through photosynthesis, and when the organic matter is broken down through respiration, carbon dioxide is released to the atmosphere. The concentration of atmospheric carbon dioxide shows a yearly oscillation (**Fig. 2**) because there is a strong seasonal annual cycle of photosynthesis and respiration in the Northern Hemisphere.

Photosynthesis and respiration in the carbon cycle can be represented by reaction (2), which is a simplified version of reaction (1). Breakdown of organic matter via respiration is accomplished mainly by bacteria that live in soils, sediments, and natural waters. As can be seen from Fig. 1, there is a very large reservoir of terrestrial carbon in carbonate rocks, which contain calcium carbonate ($CaCO_3$), and in rocks such as shales which contain organic carbon. Major exchange of carbon between rocks and the

$$CO_2 + H_2O \underset{\text{respiration}}{\overset{\text{photo-synthesis}}{\rightleftharpoons}} CH_2O + O_2 \qquad (2)$$

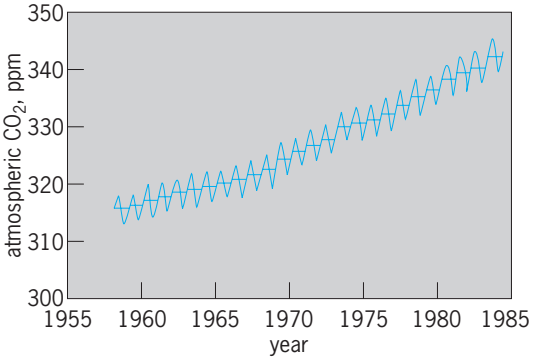

Fig. 2. Concentration of atmospheric carbon dioxide (CO_2) at Mauna Loa Observatory, Hawaii; ppm stands for parts per million volume fraction. Horizontal bars represent yearly averages. *(After J. R. Trabalka, ed., Atmospheric Carbon Dioxide and the Global Carbon Cycle, U. S. Department of Energy, DOE/ER-0239, 1985)*

atmosphere is very slow, on the scale of thousands to millions of years, compared to exchange between plants and the atmosphere, which can even be seasonal. *See* MICROBIAL ECOLOGY; SOIL MICROBIOLOGY.

Organic carbon burial on land is not important now, but was important in the geologic past when coal deposits formed in swamps. Carbon is lost from the land by river transport of dissolved bicarbonate ion [$(HCO_3)^-$] and of dissolved and particulate organic carbon. Overall, about two-thirds of the river load is derived from the biological cycle and from atmospheric carbon dioxide, and the rest comes from the weathering of carbonate rocks. However, the river transport of carbon (1 gigaton or 10^{15} g per year) is small compared to the amount of carbon cycled between the atmosphere and the land (48 gigatons/yr) or the atmosphere and the oceans (35 gigatons/yr). *See* CARBONIFEROUS.

The oceans taken as a whole represent a major reservoir of carbon. Carbon in the oceans occurs primarily as dissolved $(HCO_3)^-$ and to a lesser extent as dissolved carbon dioxide gas and carbonate ion [$(CO_3)^{2-}$]. The well-mixed surface ocean (the top 250 ft or 75 m) rapidly exchanges carbon dioxide with the atmosphere. However, the deep oceans are cut off from the atmosphere and mix with it on a long-term time scale of about 1000–2000 years. Most of the biological activity in the oceans occurs in the surface (or shallow) water where there is light and photosynthesis can occur. *See* MARITIME METEOROLOGY.

The main biological process in seawater is photosynthetic production of organic matter by phytoplankton. Some of this organic matter is eaten by animals, which are in turn eaten by larger animals farther up in the food chain. Almost all of the organic matter along the food chain is ultimately broken down by bacterial respiration, which occurs primarily in shallow water, and the carbon dioxide is quickly recycled to the atmosphere. However, some occurs in deeper waters which can accumulate excess dissolved carbon dioxide because of a lack of exchange with the atmosphere. Overall, of the large annual carbon flux (35 gigatons/yr) between the atmosphere and the oceans, only a small fraction (0.13 gigaton/yr) escapes destruction as organic matter that falls to the bottom and is buried in ocean sediments; part of this sedimentary organic matter is marine and part is transported by rivers from land. Most organic carbon is buried nearshore, for example, in marine deltas where the sediment deposition rate is high. *See* FOOD WEB; NEARSHORE PROCESSES; PHYTOPLANKTON; SEAWATER.

Another major biological process is the secretion of shells and other hard structures by marine organisms. Bicarbonate is removed from the water in the form of solid calcium carbonate (calcite and aragonite minerals), as shown in reaction (3).

$$Ca^{2+} + 2(HCO_3)^- \rightarrow CO_2 + H_2O + CaCO_3 \qquad (3)$$

A biogeochemical cycle of calcium and bicarbonate exists within the oceans, linking the deep and shallow water areas. Bottom dwellers in shallow water, such as corals, mollusks, and algae, provide calcium carbonate skeletal debris. Since the shallow waters are saturated with respect to calcium carbonate, this debris accumulates on the bottom and is buried, providing the minerals that form carbonate rocks such as limestone and dolomite. Calcium carbonate is also derived from the shells of organisms inhabiting surface waters of the deep ocean; these are tiny, floating plankton such as foraminiferans, pteropods, and coccoliths. Much of the calcium carbonate from this source dissolves as it sinks into the deeper ocean waters, which are undersaturated with respect to calcium carbonate. The undissolved calcium carbonate accumulates on the bottom to form deep-sea limestone. The calcium and the bicarbonate ions [Ca^{2+} and $(HCO_3)^-$] dissolved in the deep ocean water eventually are carried to surface and shallow water, where they are removed by planktonic and bottom-dwelling organisms to form their skeletons. *See* CARBONATE MINERALS; LIMESTONE.

The long-term biogeochemical carbon cycle occurs over millions of years when the calcium carbonate and organic matter that are buried in sediments are returned to the Earth's surface. There, weathering occurs which involves the reaction of oxygen with sedimentary organic matter with the release of carbon dioxide and water (analogous to respiration), and the reaction of water and carbon dioxide with carbonate rocks with the release of calcium and bicarbonate ions. This latter process is the reverse of that shown in reaction (3). *See* WEATHERING PROCESSES.

Human perturbation of carbon cycle. Fossil fuels (coal and oil) represent a large reservoir of carbon (Fig. 1). Burning of fossil fuels releases carbon dioxide to the atmosphere, and an increase in the atmospheric concentration of carbon dioxide has been observed since the mid-1950s (Fig. 2). While much of the increase is attributed to fossil fuels, deforestation by humans accompanied by the decay or burning of trees is another possible contributor to the problem.

When estimates are made of the amount of fossil fuels burned from 1959 to 1980, only about 60% of the carbon dioxide released can be accounted for in the atmospheric increase in carbon dioxide. The remaining 40% is known as excess carbon dioxide. The surface oceans are an obvious candidate for storage of most of the excess carbon dioxide by the reaction of carbon dioxide with dissolved carbonate to form bicarbonate. Because the increase in bicarbonate concentration in surface waters due to excess carbon dioxide uptake would be small, it is difficult to detect whether such a change has occurred. Greater quantities of excess carbon dioxide could be stored as bicarbonate in the deeper oceans, but this process takes a long time because of the slow rate of mixing between surface and deep oceans.

An increase in atmospheric carbon dioxide is of concern because of the greenhouse effect. The carbon dioxide traps heat in the atmosphere; notable increases in atmospheric carbon dioxide should cause

an increase in the Earth's surface temperature by as much as several degrees. This temperature increase would be greater at the poles, and the effects could include melting of polar ice, a rise in sea level, and changes in rainfall distribution, with droughts in interior continental areas such as the Great Plains of the United States. *See* DROUGHT; GREENHOUSE EFFECT.

Terrestrial nitrogen cycle. Nitrogen is dominantly a biogenic element and has no important mineral forms. It is a major atmospheric constituent with a number of gaseous forms, including molecular nitrogen gas (N_2), nitrogen dioxide (NO_2), nitric oxide (NO), ammonia (NH_3), and nitrous oxide (N_2O). As an essential component of plant and animal matter, it is extensively involved in biogeochemical cycling. On a global basis, the nitrogen cycle is greatly affected by human activities.

Nitrogen gas (N_2) makes up 80% of the atmosphere by volume; however, nitrogen is unreactive in this form. In order to be available for biogeochemical cycling by organisms, nitrogen gas must be fixed, that is, combined with oxygen, carbon, or hydrogen. There are three major sources of terrestrial fixed nitrogen: biological nitrogen fixation by plants, nitrogen fertilizer application, and rain and particulate dry deposition of previously fixed nitrogen. Biological fixation occurs in plants such as legumes (peas and beans) and lichens in trees, which incorporate nitrogen from the atmosphere into their living matter; about 30% of worldwide biological fixation is due to human cultivation of these plants. Nitrogen fertilizers contain industrially fixed nitrogen as both nitrate and ammonium. *See* FERTILIZER; NITROGEN FIXATION.

Fixed nitrogen in rain is in the forms of nitrate [$(NO_3)^-$] and ammonium [$(NH_4)^+$] ions. Major sources of nitrate, which is derived from gaseous atmospheric nitrogen dioxide (and nitric oxide), include (in order of importance) combustion of fossil fuel, especially by automobiles; forest fires (mostly caused by humans); and lightning. Nitrate in rain, in addition to providing soluble fixed nitrogen for photosynthesis, contributes nitric acid (HNO_3), a major component of acid rain. Sources of ammonium, which is derived from atmospheric ammonia gas (NH_3), include animal and human wastes, soil loss from decomposition of organic matter, and fertilizer release. *See* ACID RAIN.

The basic land nitrogen cycle (**Fig. 3**) involves the photosynthetic conversion of the nitrate and ammonium ions dissolved in soil water into plant organic material. Once formed, the organic matter may be stored or broken down. Bacterial decomposition of organic matter (ammonification) produces soluble ammonium ion which can then be either taken up again in photosynthesis, released to the atmosphere as ammonia gas, or oxidized by bacteria to nitrate ion (nitrification).

Nitrate ion is also soluble, and may be used in photosynthesis. However, part of the nitrate may undergo reduction (denitrification) by soil bacteria to nitrogen gas or to nitrous oxide which are then lost to the atmosphere. Compared to the land carbon cycle, the land nitrogen cycle is considerably more complex, and because of the large input of fixed nitrogen by humans, it is possible that nitrogen is building up on land. However, this is difficult to determine since the amount of nitrogen gas recycled to

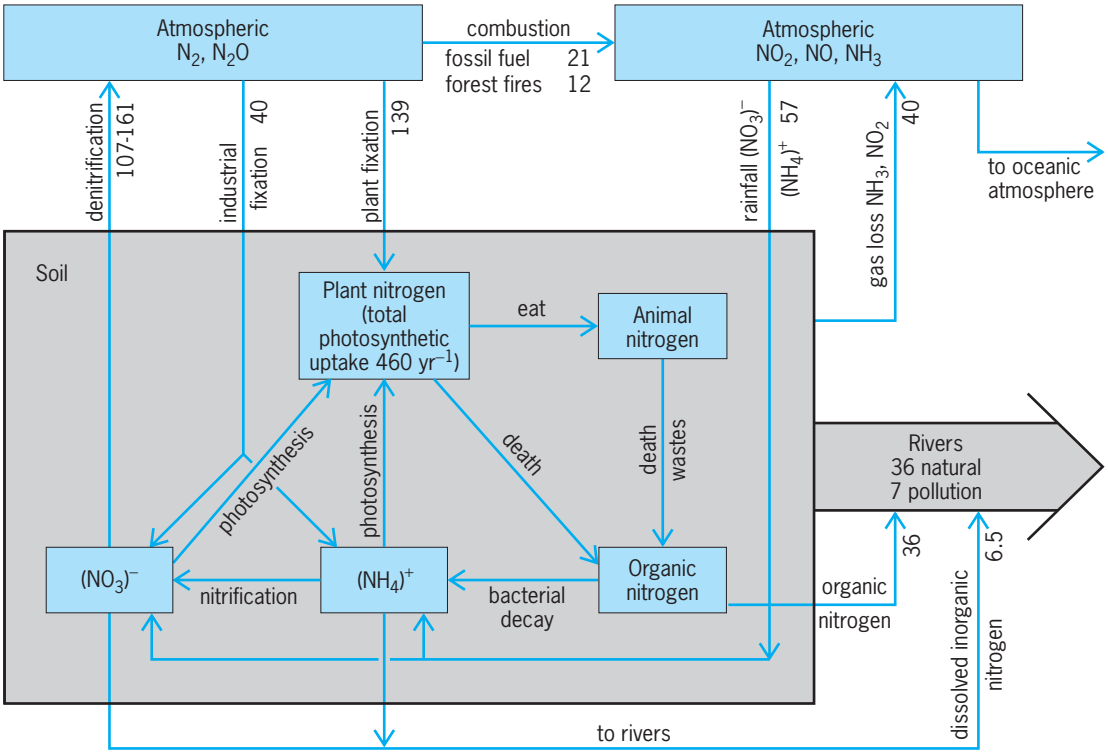

Fig. 3. Terrestrial nitrogen cycle; organic nitrogen includes both particulate and dissolved material. Flux units, teragrams (10^{12} g) N/yr.

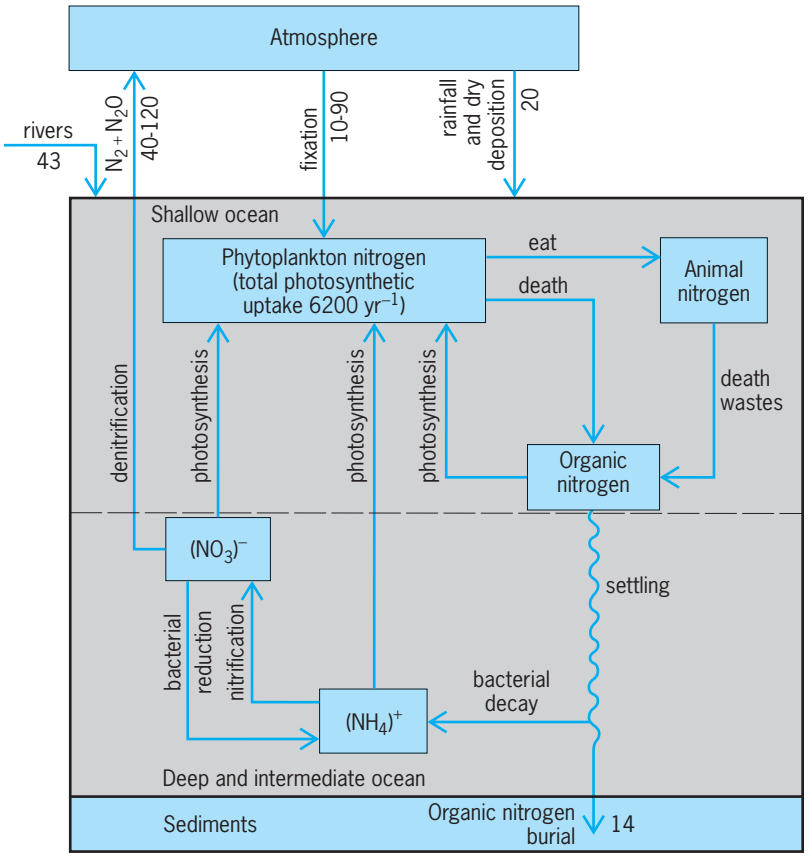

Fig. 4. Marine nitrogen cycle; organic nitrogen includes both particulate and dissolved material. Flux units, teragrams (10^{12} g) N/yr.

the atmosphere is not known and any changes in the atmospheric nitrogen concentration would be too small to detect. *See* NITROGEN CYCLE.

Oceanic nitrogen cycle. The oceans are another major site of nitrogen cycling (**Fig. 4**): the amount of nitrogen cycled biogenically, through net primary photosynthetic production, is about 13 times that on land. The main links between the terrestrial and the oceanic nitrogen cycles are the atmosphere and rivers. Nitrogen gases carried in the atmosphere eventually fall as dissolved inorganic (mainly nitrate) and organic nitrogen and particulate organic nitrogen in rain on the oceans. The flux of river nitrogen lost from the land is only about 9% of the total nitrogen recycled biogeochemically on land each year and only about 25% of the terrestrial nitrogen flux from the biosphere to the atmosphere.

River nitrogen is an important nitrogen source to the oceans; however, the greatest amount of nitrogen going into ocean surface waters comes from the upwelling of deeper waters, which are enriched in dissolved nitrate from organic recycling at depth. Dissolved nitrate is used extensively for photosynthesis by marine organisms, mainly plankton. Bacterial decomposition of the organic matter formed in photosynthesis results in the release of dissolved ammonium, some of which is used directly in photosynthesis. However, most undergoes nitrification to form nitrate, and much of the nitrate may undergo denitrification to nitrogen gas which is released to the atmosphere. A small amount of organic-matter nitrogen is buried in ocean sediments, but this accounts for a very small amount of the nitrogen recycled each year. There are no important inorganic nitrogen minerals such as those that exist for carbon and phosphorus, and thus there is no mineral precipitation and dissolution. *See* UPWELLING.

Phosphorus cycle. Phosphorus, an important component of organic matter, is taken up and released in the form of dissolved inorganic and organic phosphate. Phosphorus differs from nitrogen and carbon in that it does not form stable atmospheric gases and therefore cannot be obtained from the atmosphere. It does form minerals, most prominently apatite (calcium phosphate), and insoluble iron (Fe) and aluminum (Al) phosphate minerals, or it is adsorbed on clay minerals. The amount of phosphorus used in photosynthesis on land is large compared to phosphorus inputs to the land (**Fig. 5**). The major sources of phosphorus are weathering of rocks containing apatite and mining of phosphate rock for fertilizer and industry. A small amount comes from precipitation and dry deposition. *See* PHOSPHATE MINERALS.

Phosphorus is lost from the land principally by river transport, which amounts to only 7% of the amount of phosphorus recycled by the terrestrial biosphere; overall, the terrestrial biosphere conserves phosphorus. Humans have greatly affected terrestrial phosphorus: deforestation and agriculture have doubled the amount of phosphorus weathering; phosphorus is added to the land as fertilizers and from industrial wastes, sewage, and detergents. Thus, about 75% of the terrestrial input is anthropogenic; in fact, phosphorus may be building up on the land.

In the oceans, phosphorus occurs predominantly as dissolved orthophosphates [PO_4^{3-}, $(HPO_4)^{2-}$ and $(H_2PO_4)^-$]. Since it follows the same cycle as do carbon and nitrogen, dissolved orthophosphate is depleted in surface ocean waters where both photosynthesis and respiration occur, and the concentration builds up in deeper water where organic matter is decomposed by bacterial respiration. The major phosphorus input to the oceans is from rivers, with about 5% coming from rain. However, 75% of the river phosphorus load is due to anthropogenic pollutants; humans have changed the ocean balance of phosphorus. Most of the dissolved oceanic orthophosphate is derived from recycled organic matter. The output of phosphorus from the ocean is predominantly biogenic: organic phosphorus is buried in sediments; a smaller amount is removed by adsorption on volcanic iron oxides. In the geologic past, there was a much greater inorganic precipitation of phosphorite (apatite) from seawater than at present, and this has resulted in the formation of huge deposits which are now mined.

Nutrients in lakes. Biogeochemical cycling of phosphorus and nitrogen in lakes follows a pattern that is similar to oceanic cycling: there is nutrient depletion in surface waters and enrichment in deeper waters. Oxygen consumption by respiration in deep water

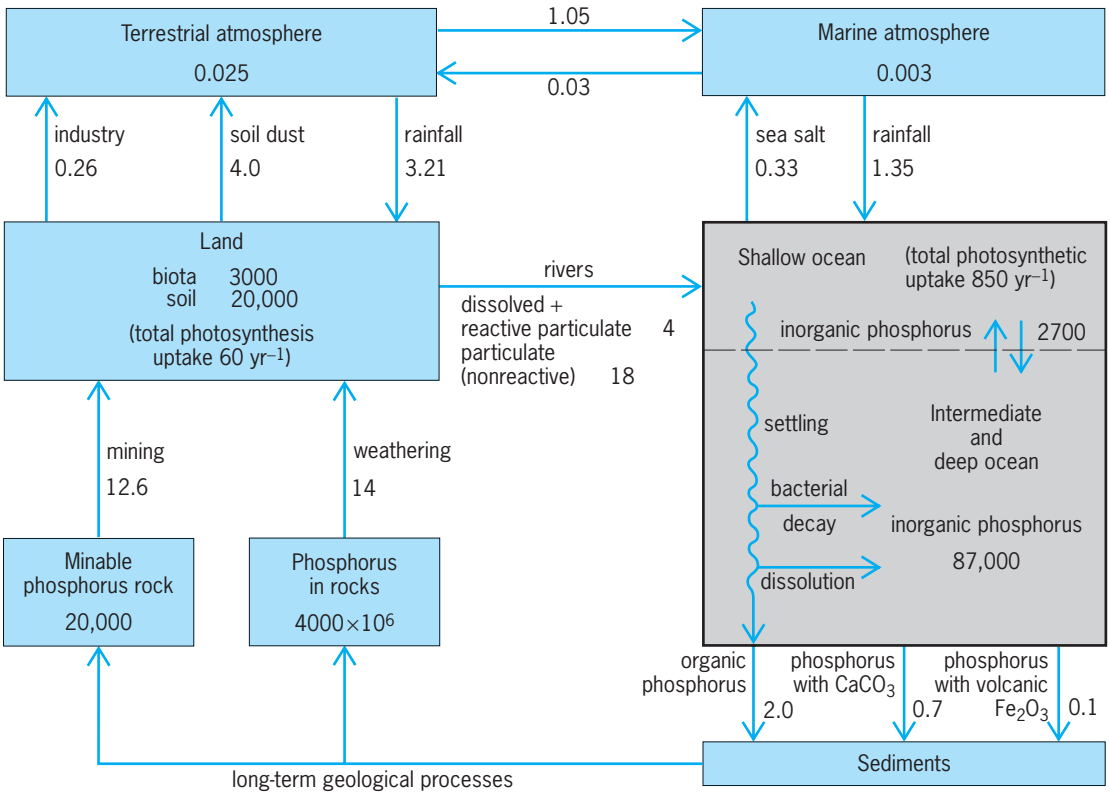

Fig. 5. Phosphorus cycle. Reservoir units, teragrams (10^{12} g) P; flux units, teragrams P/yr.

sometimes leads to extensive oxygen depletion with adverse effects on fish and other biota. In lakes, phosphorus is usually the limiting nutrient.

Many lakes have experienced greatly increased nutrient (nitrogen and phosphorus) input due to human activities. This stimulates a destructive cycle of biological activity: very high organic productivity, a greater concentration of plankton, and more photosynthesis. The result is more organic matter falling into deep water with increased depletion of oxygen and greater accumulation of organic matter on the lake bottom. This process, eutrophication, can lead to adverse water quality and even to the filling up of small lakes with organic matter. *See* EUTROPHICATION; LIMNOLOGY.

Biogeochemical sulfur cycle. A dominant flux in the global sulfur cycle (**Fig. 6**) is the release of 65–70 teragrams of sulfur per year to the atmosphere from burning of fossil fuels. Sulfur contaminants in these fuels are released to the atmosphere as sulfur dioxide (SO_2) which is rapidly converted to aerosols of sulfuric acid (H_2SO_4), the primary contributor to acid rain. Forest burning results in an additional release of sulfur dioxide. Overall, the broad range of human activities contribute 75% of sulfur released into the atmosphere. Natural sulfur sources over land are predominantly the release of reduced biogenic sulfur gases [mainly hydrogen sulfide (H_2S) and dimethyl sulfide] from marine tidal flats and inland waterlogged soils and, to much lesser extent, the release of volcanic sulfur. The atmosphere does not have an appreciable reservoir of sulfur because most sulfur gases are rapidly returned (within days) to the land in rain and dry deposition. There is a small net flux of sulfur from the atmosphere over land to the atmosphere over the oceans.

Ocean water constitutes a large reservoir of dissolved sulfur in the form of sulfate ions [$(SO_4)^{2-}$]. Some of this sulfate is thrown into the oceanic atmosphere as sea salt from evaporated sea spray, but most of this is rapidly returned to the oceans. Another major sulfur source in the oceanic atmosphere is the release of oceanic biogenic sulfur gases (such as dimethyl sulfide) from the metabolic activities of oceanic organisms and organic matter decay. Marine organic matter contains a small amount of sulfur, but sulfur is not a limiting element in the oceans.

Another large flux in the sulfur cycle is the transport of dissolved sulfate in rivers. However, as much as 43% of this sulfur may be due to human activities, both from burning of fossil fuels and from fertilizers and industrial wastes. The weathering of sulfur minerals, such as pyrite (FeS_2) in shales, and the evaporite minerals, gypsum and anhydrite, make an important contribution to river sulfate. The major mechanism for removing sulfate from ocean water is the formation and burial of pyrite in oceanic sediments, primarily nearshore sediments. (The sulfur fluxes of sea salt and biogenic sulfur gases do not constitute net removal from the oceans since the sulfur is recycled to the oceans.)

Biogeochemical cycles and atmospheric oxygen. The main processes affecting atmospheric oxygen, as shown by reaction (2), are photosynthesis and

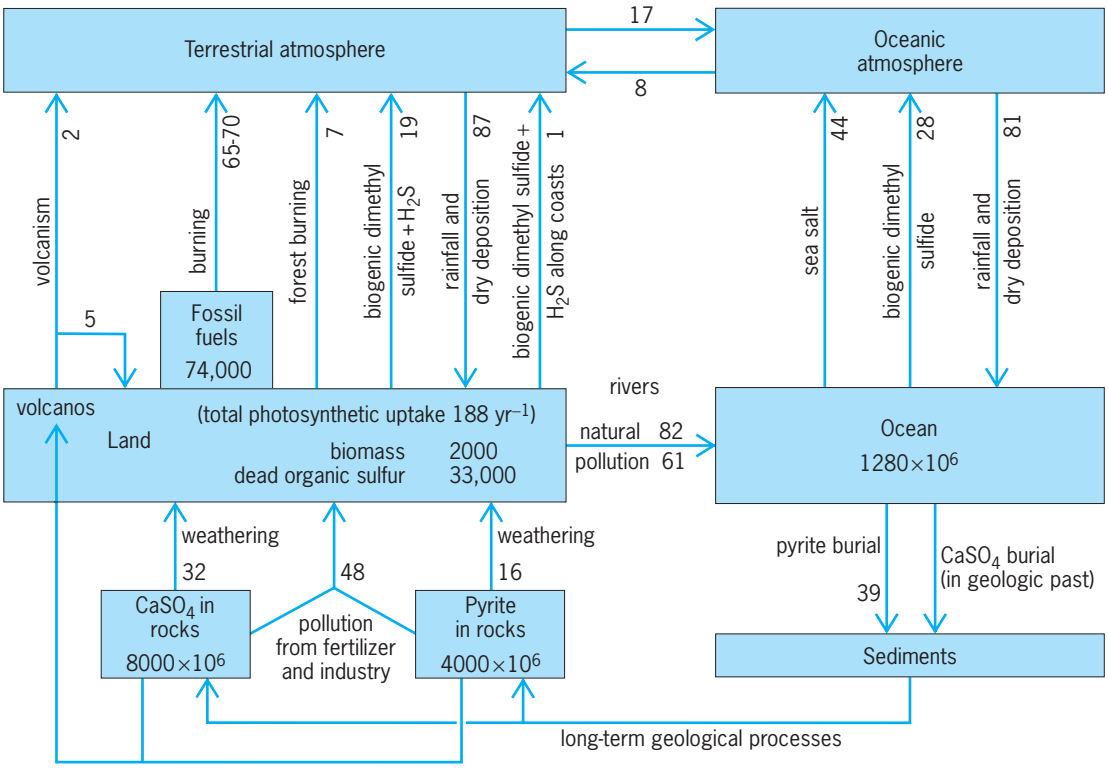

Fig. 6. Sulfur cycle. Reservoir units, teragrams (10^{12} g) S; flux units, teragrams S/yr.

respiration; however, these processes are almost perfectly balanced against one another and, thus, do not exert a simple effect on oxygen levels. Only the very small excess of photosynthesis over respiration, manifested by the burial of organic matter in sediments, is important in raising the level of oxygen. This excess is so small, and the reservoir of oxygen so large, that if the present rate of organic carbon burial were doubled, and the other rates remained constant, it would take 5–10 million years for the amount of atmospheric oxygen to double. Nevertheless, this is a relatively short time from a geological perspective. *See* ATMOSPHERE; ATMOSPHERE, EVOLUTION OF; ATMOSPHERIC CHEMISTRY; BIOSPHERE; GEOCHEMISTRY; HYDROSPHERE; MARINE SEDIMENTS.

Elizabeth K. Berner; Robert A. Berner

Bibliography. E. K. Berner and R. A. Berner, *The Global Water Cycle: Geochemistry and Environment*, 1987; R. A. Berner (ed.), Geochemical cycles of nutrient elements, *Amer. J. Sci.*, 282:401–542, 1982; C. B. Gregor et al., *Chemical Cycles in the Evolution of the Earth*, 1988; G. E. Likens, F. H. Bormann, and N. M. Johnson, *Biogeochemistry of a Forested Ecosystem*, 2d ed., 1995; J. R. Trabalka (ed.), *Atmospheric Carbon Dioxide and the Global Carbon Cycle*, U. S. Department of Energy, DOE/ER-0239, 1985.

Biogeography

A synthetic discipline that describes the distributions of living and fossil species of plants and animals across the Earth's surface as consequences of ecological and evolutionary processes. Biogeography overlaps and complements many biological disciplines, especially community ecology, systematics, paleontology, and evolutionary biology. *See* PLANT GEOGRAPHY; ZOOGEOGRAPHY.

Development. Early biogeographers were natural historians who collected data regarding the geographic locations of different species. Initial research focused on Europe. Data from Africa, Asia, and the New World were compiled as explorations of these regions by Europeans progressed. These early compilations of data on geographic distributions played a key role in C. Darwin's demonstration that species were not generated by independent events, but were derived from other species. Darwin showed that despite being designed in different ways, most organisms found together in a geographic region were closely related. For example, features of reproductive biology indicated that marsupial "moles" in Australia were more closely allied with kangaroos than with European moles, which looked superficially similar. Similar patterns, found within many different groups of plants and animals, could not be reconciled with arguments from design that assumed similarity in appearance was a basis for classifying species into natural groups. Biogeographic patterns supported Darwin's alternative view that natural groups of organisms were formed by genealogical relationships.

Based on relatively complete compilations of species within well-studied groups, such as birds and mammals, biogeographers identified six different realms within which species tend to be closely related and between which turnovers in major groups

Biogeographic realms

Realm	Continental areas included	Examples of distinctive or endemic taxa
Palearctic	Temperate Eurasia and northern Africa	Hynobiid salamanders
Oriental	Tropical Asia	Lower apes
Ethiopian	Sub-Saharan Africa	Great apes
Australian	Australia and New Guinea	Marsupials
Nearctic	Temperate North America	Pronghorn antelope, ambystomatid salamanders
Neotropic	Subtropical Central America and South America	Hummingbirds, antbirds, marmosets

of species are observed (see **table**). The boundaries between biogeographic realms are less distinct than was initially thought, and the distribution of distinctive groups such as parrots, marsupials, and southern beeches (*Nothofagus* spp.) implies that modern-day biogeographic realms have been considerably mixed in the past. *See* ANIMAL EVOLUTION; PALEOBOTANY; PALEOECOLOGY; PALEONTOLOGY; PLANT EVOLUTION; SPECIATION.

Historical biogeography. In the 1920s, data on the distribution of certain groups of living and fossil plants and animals were used by the German meteorologist Alfred L. Wegener to argue that continents had once been connected with one another. When the physical basis for plate tectonics was discovered, a revolution in biogeography ensued. Previously, biogeographic distributions of taxa were explained by dispersal of species away from centers of origin. For example, marsupials were thought to have originated in Australia and then dispersed by some means to South America. After it was realized that continents had moved across the face of the Earth in the past, and are in fact currently moving, explanations for disjunct distributions of taxa like marsupials could be explained as the initial evolutionary radiation of a group on a continent, followed by a splitting of that continent into two or more pieces.

A key to explaining the existence of disjunct distributions within a taxon is having an estimate of the evolutionary history of that taxon. The discovery that such estimates could be obtained using phylogenetic systematics ushered in yet another revolution in biogeography. Using the methods of phylogenetic systematics, estimates of phylogenies called cladograms are obtained (**Fig. 1**). A cladogram is a bifurcating network, where nodes in the network represent hypothesized common ancestors. When the geographic distributions of different taxa are superimposed upon a cladogram, the resulting area cladogram provides an estimate of the geological history of the areas (Fig. 1). Geographic regions with shallow branches in an area cladogram are predicted to have shared a recent geological history, while those with deep branches are hypothesized to have shared a more ancient history. Thus, area cladograms generate predictions about unobservable and unique historical events that can be examined with independent geological, paleontological, or biogeographic data sets.

Ecological biogeography. Historical approaches to biogeography are well suited for developing an understanding of geographic patterns in the distribution of taxa, but they have provided relatively little insight into patterns of species diversity within and among different geographic regions. Two such patterns have stimulated a great deal of progress in developing ecological explanations for geographic patterns of species richness. The first is that the number of species increases in a regular fashion with the size of the geographic area being considered. The second is the nearly universal observation that there are more species of plants and animals in tropical regions than in temperate and polar regions.

In order to answer questions about why there are a certain number of species in a particular geographic region, biogeography has incorporated many insights from community ecology. Species number at any particular place depends on the amount of resources available there (ultimately derived from the amount of primary productivity), the number of ways those resources can be apportioned among

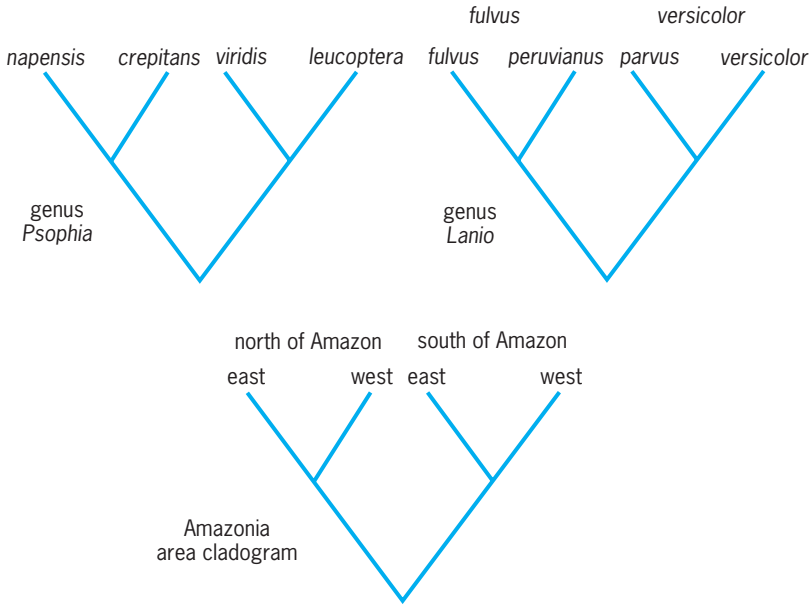

Fig. 1. Cladograms for two genera of birds living in the Amazon Basin, and the area cladogram most consistent with the data from these genera. Within each genus, taxa have disjunct, nonoverlapping distributions. *Psophia napiens* and *P. crepitans* live north of the Amazon, with *napiens* farthest west. *Psophia viridis* and *P. leucoptera* are south of the Amazon, with *leucoptera* farthest west. Two species in *Lanio* are sister taxa, with *L. fulvus* distributed north of the Amazon and *L. versicolor* south of the Amazon. Each species has two subspecies. *L. f. peruvianus* and *L. v. versicolor* are the westernmost subspecies in each case. The area cladogram implies that the Amazon marked a boundary that isolated populations of a single species in each genus on its north and south banks. Subsequently, barriers formed that isolated eastern and western populations within the northern and southern populations. (*Data from J. Cracraft, Deep-history biogeography: Retrieving the historical pattern of evolving biotas, Sys. Zool., 37:221–236,1988*)

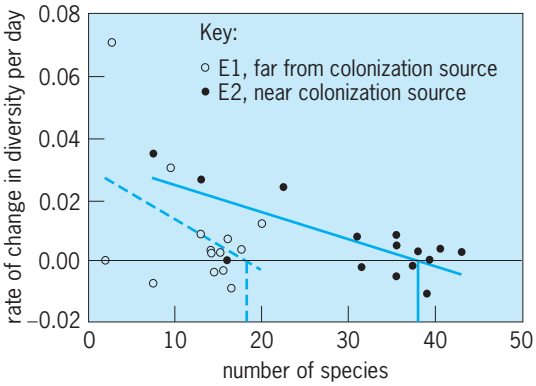

Fig. 2. Patterns of colonization by arthropod species for mangrove islands that were experimentally defaunated show colonization rates that correspond to those predicted from the equilibrium theory of island biogeography. Here, the rate of change in number of species as the islands recovered from the defaunation is plotted against the number of species on two islands. Least-squares regression lines were fitted to the rate data for each island. The equilibrium number of species is the number for which the rate of change in diversity is zero, shown here by the *x*-intercepts of the regression lines. The theory predicts that islands closer to a colonization source (island E2) should have a higher equilibrium diversity than islands farther from the source (island E1). This prediction is clearly substantiated by the experiment. (*Data from D. S. Simberloff and E. O. Wilson, Experimental zoogeography of islands: The colonization of empty islands, Ecology, 50:278–296, 1969*)

species, and the different kinds of ecological requirements of the species that can colonize the region. The equilibrium theory of island biogeography arose as an application of these insights to the distribution of species within a specified taxon across an island archipelago. This theory generated specific predictions about the relationships among island size and distance from a colonization source with the number and rate of turnover of species. Large islands are predicted to have higher equilibrium numbers of species than smaller islands; hence, the species area relationship can be predicted in principle from the ecological attributes of species. Experimental and observational studies have confirmed many predictions made by this theory (**Fig. 2**).

Applying the insights from the equilibrium theory of island biogeography to patterns in species diversity across continents has been much more difficult. Islands are discrete geographic features. Analogous "habitat" islands exist in some places on continents, but most terrestrial habitats are connected together within landscapes so that dispersal from one segment of habitat to another is relatively common. Populations on continents that might normally become extinct if they were located on islands (with relatively small likelihood of receiving colonists) can persist on a continent by frequent immigration from nearby populations. A consensus theory of species diversity on continents has not yet arisen, but current work on the subject indicates that the ability of organisms to disperse from one population to the next, and how this varies among species and landscapes, will be a primary component of the extension of island biogeographic theory to continents.

Aquatic organisms are exposed to unique physical and biotic conditions that make it difficult to extend ecological biogeographic theory from terrestrial to aquatic ecosystems. Fresh-water organisms inhabiting lakes can be viewed as inhabiting aquatic islands in a sea of terrestrial ecosystems, but species distributed in rivers exist within effectively linear biogeographic realms where physical flow tends to be in a single direction (such as downstream). Marine organisms often have larval dispersal strongly constrained by major oceanographic features such as directional ocean currents and gyres. Population sizes and species diversity may depend less on available local resources and more strongly on successful transportation of larvae across large distances. Such unique features of aquatic biology place unique challenges to constructing a general-theory ecological biogeography applicable to both aquatic and terrestrial biomes. *See* ECOLOGICAL COMMUNITIES; ECOLOGICAL SUCCESSION; ECOLOGY; ECOSYSTEM; ISLAND BIOGEOGRAPHY; TERRESTRIAL ECOSYSTEM.

The latitudinal gradient in species richness has generated a number of explanations, none of which has been totally satisfactory. One explanation is based on the observation that species with more temperate and polar distributions tend to have larger geographic ranges than species from tropical regions. This is not an artifact of the distribution of tropical regions on continents because similar patterns are seen in marine environments. It is thought that since species with large geographic ranges tend to withstand a wider range of physical and biotic conditions, this allows them to penetrate farther into regions with more variable climates at higher latitudes. If this were true, then species with smaller geographic ranges would tend to concentrate in tropical regions where conditions are less variable. While this might be generally true, there are many examples of species living in high-latitude regions that have small geographic regions. *See* ALTITUDINAL VEGETATION ZONES.

Modern developments. Historical biogeography and ecological biogeography have developed essentially independent of one another, involving research from entirely different groups of scientists that seldom communicate. This has traditionally hindered the development of a robust evolutionary approach to biogeography, which must necessarily incorporate both ecological and historical explanations. Consider the problem of understanding why a particular number of species inhabit a particular geographic region. The current number of species will be determined by a complex interaction of a number of factors operating at different temporal and spatial scales. Ecological factors might include the effects of climate on the spatial population dynamics across a species' ranges, spatial variation in appropriate biotic and abiotic resources, and the geographic distribution of competitors and natural enemies. Evolutionary factors might include the rate and mode of speciation and extinction, the rate of adaptive divergence, and the nature of barriers to gene flow across species' ranges. Attempts to resolve the effects of these kinds

of processes must incorporate an understanding of the natural history of the species involved, estimates of phylogenetic relationships among them, and the structure of the environment across which species are arrayed.

Recent attempts to integrate the perspectives described in the previous paragraph under the rubrics of macroecology, dynamic biogeography, and areography suggest promising directions for future research. Areography was a term devised by the Argentinian plant ecologist Eduardo Rapoport in 1982 to describe his approach to studying the properties of geographic ranges of species. His approach was explicitly synthetic, and he attempted to devise quantitative models that described variation in the properties of geographic ranges within and among species. His insights were later expanded by the concept of dynamic biogeography. This approach focused not only on the static properties of geographic ranges but also on changes in ranges due to invasions, climate change, and extinction. Dynamic biogeography sought to explain how biological processes operated across a geographic range to determine spatial variation in local abundance, and how this affects the size, shape, and dynamics of the range boundary for a species. Additional insights to the explanations of biogeographic patterns were developed by showing that patterns of body size variation among species were closely related to variation in range size, average abundance, trophic level, and phylogeny. This approach, termed macroecology, emphasized that most explanations for biogeographic phenomena required a comparative approach. It was found that patterns in the variation among species implied the operation of relatively general evolutionary processes, such as nonrandom speciation and extinction.

Biogeography is entering a phase where data on the spatial patterns of abundance and distribution of species of plants and animals are being analyzed with sophisticated mathematical and technological tools. Geographic information systems and remote sensing technology have provided a way to catalog and map spatial variation in biological processes with a striking degree of detail and accuracy. These newer technologies have stimulated research on appropriate methods for modeling and analyzing biogeographic patterns. Modern techniques of spatial modeling are being applied to geographic information systems data to test mechanistic explanations for biogeographic patterns that could not have been attempted without the advent of the appropriate technology. These innovations promise to make biogeography a vital and exciting field in the future. *See* GEOGRAPHIC INFORMATION SYSTEMS. Brian A. Maurer

Bibliography. J. H. Brown and M. V. Lomolino, *Biogeography*, Sinauer, Sunderland, MA, 1998; A. Hallam, *Outline of Phanerozoic Biogeography*, Oxford University Press, Oxford, 1994; R. H. MacArthur, *Geographical Ecology*, Harper and Row, New York, 1972; A. A. Myers and P. S. Giller, *Analytical Biogeography*, Chapman and Hall, London, 1988.

Bioinorganic chemistry

The field at the interface between biochemistry and inorganic chemistry; also known as inorganic biochemistry or metallobiochemistry. This field involves the application of the principles of inorganic chemistry to problems of biology and biochemistry. Because most biological components are organic, that is, they involve the chemistry of carbon compounds, the combination of the prefix bio- and inorganic may appear contradictory. However, organisms require a number of other elements to carry out their basic functions. Many of these elements are present as metal ions that are involved in crucial biological processes such as respiration, metabolism, cell division, muscle contraction, nerve impulse transmission, and gene regulation. The characterization of the interactions between such metal centers and biological components is the heart of bioinorganic chemistry. *See* BIOCHEMISTRY; INORGANIC CHEMISTRY.

Metal ions influence biological phenomena by interacting with organic functional groups on small and large biomolecules, forming metal complexes. From this perspective, much of bioinorganic chemistry may be considered as coordination chemistry applied to biological questions. In general, bioinorganic chemists tackle such problems by first focusing on the elucidation of the structure of the metal complex of interest and then correlating structure with function. The attainment of solutions usually requires a combination of physical, chemical, and biological approaches. Biochemistry and molecular biology are often used to provide sufficient amounts of the system for investigation. Physical approaches such as crystallography and spectroscopy are useful in defining structural properties of the metal site. Synthetic methods can be used for the design and assembly of structural, spectroscopic, and functional models of the metal site. All these approaches then converge to elucidate how such a site functions. *See* COORDINATION CHEMISTRY; CRYSTALLOGRAPHY; MOLECULAR BIOLOGY; SPECTROSCOPY.

Low-molecular-weight compounds. A number of coordination compounds found in organisms have relatively low molecular weights. Ionophores, molecules that are able to carry ions across lipid barriers, are polydentate ligands designed to bind alkali and alkaline-earth metal ions; they span membranes and serve to transport such ions across these biological barriers. Molecular receptors known as siderophores are also polydentate ligands; they have a very high affinity for iron. Bacteria and fungi use them to scavenge for this crucial element in the environment. Siderophores can thus be used as growth factors and antibiotics and for complexing excess iron in cases of iron overload. *See* CHELATION; IONOPHORE.

Other low-molecular-weight compounds are metal-containing cofactors that interact with macromolecules to promote important biological processes. Perhaps the most widely studied of the metal ligands found in biochemistry are the porphyrins; iron protoporphyrin IX (**Fig. 1**) is an example of the all-important complex in biology known as heme.

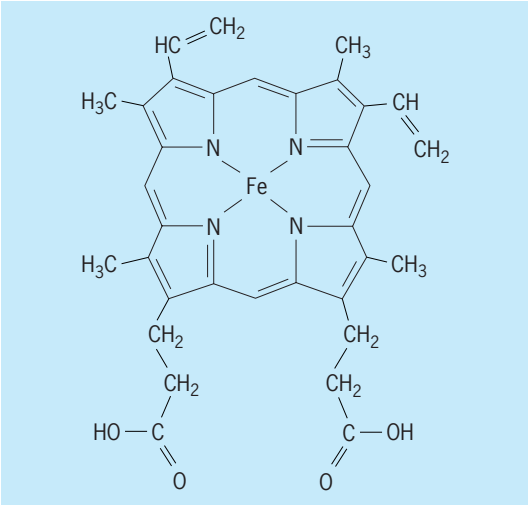

Fig. 1. Iron complex of protoporphyrin IX, or heme.

The iron (Fe) is attached to the four nitrogen (N) atoms of the porphyrin. Since iron can bind as many as six ligands, there are two other sites perpendicular to the heme plane that are available for coordination. One or both sites can be used to attach the heme to a protein via a coordinating amino acid residue. For example, in hemoglobin the iron is attached to only one protein residue. Thus the remaining coordination site opposite the protein attachment site is available for dioxygen (O_2) binding (**Fig. 2***a*). This ability to bind O_2 is the basis for the biological function of hemoglobin, which is the red substance in blood cells that transports oxygen to the tissues. Both available coordination sites in cytochrome *c*, however, are attached to protein residues (Fig. 2*b*). Cytochrome *c* participates in electron-transfer processes that are important in supplying energy for cellular functions. The electron-transfer function does not require the iron atom to bind exogenous ligands (like O_2); indeed, the efficiency of electron transfer is facilitated by the presence of the two protein ligands. In the preceding two examples, it can be seen that the structure of the metal site is intimately correlated with its function. *See* CYTOCHROME; ELECTRON-TRANSFER REACTION; HEMOGLOBIN; PORPHYRIN.

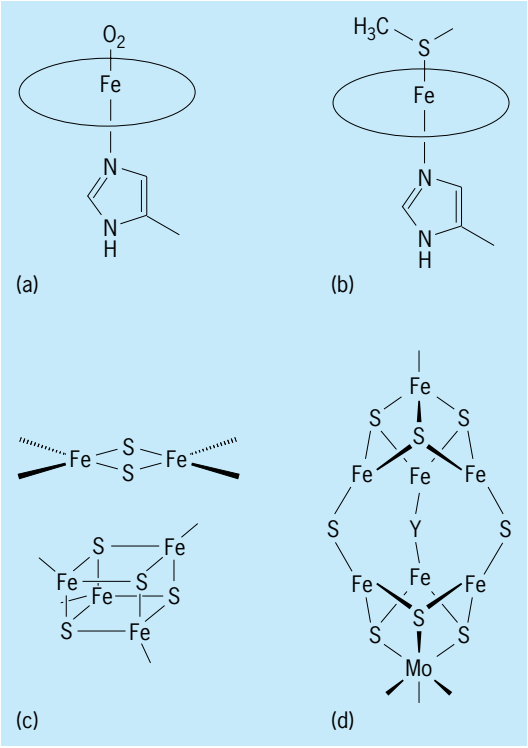

Fig. 2. Active sites of some metalloproteins. (*a*) Hemoglobin and (*b*) cytochrome *c*; ovals represent the porphyrin ring. (*c*) Ferredoxins. (*d*) Nitrogenase; the identity of the Y atom is unknown.

Chlorophyll and vitamin B_{12} are chemically related to the porphyrins. The carbon-nitrogen skeleta are similar, and a metal ion coordinates to the four nitrogen atoms of the macrocycle. Magnesium is the central metal ion in chlorophyll, which is the green pigment in plants used to convert light energy into chemical energy. Cobalt is the central metal ion in vitamin B_{12}; it is converted into coenzyme B_{12} in cells, where it participates in a variety of enzymatic reactions. *See* CHLOROPHYLL; VITAMIN B_{12}.

Another example of a metal-containing cofactor is molybdopterin, which can be extracted from a number of molybdenum (Mo)-containing enzymes that catalyze the transfer of oxygen atoms to and from substrates, for example, nitrate reductase and sulfite oxidase. When bound to the enzyme, the molybdenum is believed to be the locus of the oxo-transfer activity. *See* ENZYME.

Metalloproteins and metalloenzymes. These are metal complexes of proteins. In many cases, the metal ion is coordinated directly to functional groups on amino acid residues. In some cases, the protein contains a bound metallo-cofactor such as heme or molybdopterin. In metalloproteins with more than one metal-binding site, the metal ions may be found in clusters. Examples include ferredoxins, which contain iron-sulfur clusters (Fe_2S_2 or Fe_4S_4), and nitrogenase, which contains both Fe_4S_4 units and a novel $MoFe_7S_8$ cluster (Fig. 2*c*). *See* PROTEIN.

Some metalloproteins are designed for the storage and transport of the metal ions themselves—for example, ferritin and transferrin for iron and metallothionein for zinc (Zn). Others, such as the yeast protein Atx1, act as metallochaperones that aid in the insertion of the appropriate metal ion into a metalloenzyme. Still others function as transport agents. Cytochromes and ferredoxins facilitate the transfer of electrons in various metabolic processes. Hemoglobin is required for oxygen transport in humans and higher animals, while hemocyanin (containing a dicopper active site; **Fig. 3**) and hemerythrin (containing a diiron active site) perform the analogous function in crustaceans and sipunculids, respectively.

Many metalloproteins catalyze important cellular reactions and are thus more specifically called metalloenzymes. Enzymes are biological catalysts; when a metal ion is present, it is often the site at which

catalytic activity occurs. The metal ion can serve as a redox site to facilitate the oxidation or reduction of substrate. Superoxide dismutase, which removes potentially toxic superoxide by disproportionating it to O_2 and hydrogen peroxide (H_2O_2), can be found in nature as a copper-zinc enzyme, an iron enzyme, or a manganese enzyme. Cytochrome oxidase is the respiratory enzyme in mitochondria responsible for disposing of the electrons generated by mammalian metabolism; it does so by reducing O_2 to water with the help of both heme and copper centers. In contrast, the conversion of water to O_2 is carried out in the photosynthetic apparatus by manganese centers. *See* SUPEROXIDE CHEMISTRY.

Other metalloenzymes are involved in the transformation of organic molecules in cells. For example, tyrosine hydroxylase (an iron enzyme) and dopamine β-hydroxylase (a copper enzyme) carry out oxidation reactions important for the biosynthesis of neurotransmitters. Cytochrome P450 (a heme enzyme) and methane monooxygenase (a nonheme iron enzyme) hydroxylate unactivated aliphatic carbon-hydrogen (C-H) bonds in molecules such as steroids and methane; these reactions are difficult to carry out efficiently and specifically outside the biological milieu.

Alternatively, the metal center can serve as a Lewis acidic site to activate substrates for nucleophilic displacement reactions (that is, hydrolysis). Carboxypeptidase, which catalyzes the cleavage of peptide bonds, and alkaline phosphatase, which catalyzes the cleavage of phosphate ester bonds, are examples of zinc-containing hydrolytic enzymes. Metal ions have also been found to be required for the hydrolytic reactions catalyzed by some ribozymes (biological catalysts based on polyribonucleotides). *See* RIBOZYME.

Bioinorganic models. Insights into the metal sites of macromolecules can also be obtained from studies of much simpler metal complexes that can be designed to mimic some aspect of the active site structure and function. For example, a number of metal complexes can act like carboxypeptidase and hydrolyze peptide bonds. Though the metalloenzyme itself does this more efficiently, a comparison of the simpler complexes with their more complicated biochemical relatives can often provide clues to the behavior of the macromolecular catalyst. Modeling efforts have developed to the extent that a number of important metalloprotein active sites (like some of those in Fig. 2) can be synthesized independent of the protein backbone and can be studied in great detail. The greater challenge is to design models that can carry out the same reactions as those carried out by their macromolecular counterparts. Examples of functional models that have been synthesized are heme and dicopper complexes that model the reversible oxygen-binding ability of hemoglobin and hemocyanin, respectively. *See* SUPRAMOLECULAR CHEMISTRY.

Metal regulation of gene expression. Metalloproteins also play roles in gene expression. In some cases, the metal ion coordinates to several residues in a protein, imposing a particular structure on the polypeptide chain that enables it to interact with deoxyribonucleic acid (DNA); for example, zinc is a crucial component of the so-called zinc finger proteins that recognize and bind particular sequences of DNA and activate transcription. In other cases, the metal ion acts as a signal triggering the expression or repression of genes involved in respiration and metabolism. For example, the binding of nanomolar concentrations of mercuric ion to the merR protein activates the genetic machinery to produce enzymes responsible for the detoxification of mercury. In mammalian cells, a single protein that senses the presence of ferrous ion interacts with ferritin and transferrin receptor messenger ribonucleic acids (mRNAs) to regulate translation of the corresponding ferritin and transferrin receptor proteins. *See* DEOXYRIBONUCLEIC ACID (DNA); GENE ACTION; GENETIC CODE; RIBONUCLEIC ACID (RNA).

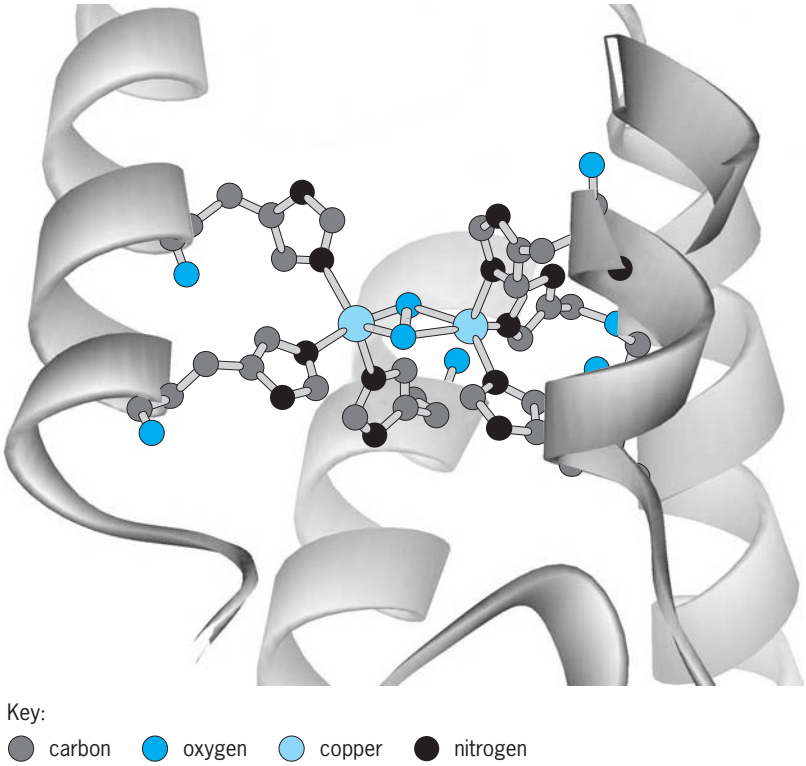

Fig. 3. Active site structure of oxyhemocyanin, the oxygen-bound form of hemocyanin, the oxygen carrier protein in arthropods and mollusks.

Metals in medicine. Metal complexes have also been found to be useful as therapeutic or diagnostic agents. Prominent among metal-based drugs is cisplatin, which is particularly effective in the treatment of testicular and ovarian cancers; second- and third-generation variants of this drug have been developed to mitigate undesirable side effects of drug treatment. Gold, gallium, and bismuth compounds are used for the treatment of rheumatoid arthritis, hypercalcemia, and peptic ulcers, respectively.

In clinical diagnosis, metal complexes can be used as imaging agents. The convenient half-life and

radioemission properties of technetium-99 make its complexes very useful for a number of applications; by varying the ligands bound to the metal ion, diagnostic agents have been developed for imaging the heart, brain, and kidneys. Complexes of paramagnetic metal ions such as gadolinium(III), iron(III), and manganese(II) are also used as contrast agents to enhance images obtained from magnetic resonance imaging (MRI). *See* COORDINATION COMPLEXES; MEDICAL IMAGING; ORGANOMETALLIC COMPOUND.

Lawrence Que, Jr.

Bibliography. Bioinorganic chemistry (spec. issue), *Science*, 261:699–730, 1993; J. A. Cowan, *Inorganic Biochemistry: An Introduction*, 2d ed., 1997; S. J. Lippard and J. M. Berg, *Principles of Bioinorganic Chemistry*, 1994; J. Reedijk (ed.), *Bioinorganic Catalysis*, 2d ed., 1999; R. M. Roat-Malone, *Bioinorganic Chemistry: A Short Course*, 2002; Spec. issues, *Chem. Rev.*, 96:2237–3042, 1996, and 99:2201–2842, 1999.

Bioleaching

The use of naturally occurring microorganisms to dissolve metals from minerals. This biotechnical process, involving bacteria, archaea, fungi, and yeast, can be applied to both sulfide and nonsulfide minerals. At present, bioleaching is employed on a large commercial scale to extract metals from sulfide minerals. *See* HYDROMETALLURGY; LEACHING.

Bioleaching of Sulfide Minerals

Sulfide minerals bioleaching has its technical origins in the 1940s, when researchers discovered the acid-loving bacterium *Thiobacillus ferrooxidans* (now reclassified as *Acidithiobacillus ferrooxidans*) and found that it obtained energy for reproduction and cellular functioning by oxidizing (removing electrons from) inorganic substances, principally ferrous iron (Fe^{2+}). Subsequently, other microorganisms capable of dissolving metals were characterized and their biochemical mechanisms defined, while engineered systems to commercially exploit these organisms were developed. *See* BACTERIA; SULFIDE AND ARSENIDE MINERALS.

Principles of bioleaching. Microorganisms involved in bioleaching of sulfide minerals are chemolithotrophic: they use carbon dioxide (CO_2) from the atmosphere as their carbon source for building cellular constituents and derive energy for their existence and reproduction from inorganic materials. To obtain energy, the microorganisms oxidize iron [reaction (1)] and certain chemically reduced sulfur

$$\underset{\text{Ferrous ion}}{4Fe^{2+}} + \underset{\text{Oxygen}}{O_2} + \underset{\text{Hydrogen ions}}{4H^+} \longrightarrow \underset{\text{Ferric ion}}{4Fe^{3+}} + \underset{\text{Water}}{2H_2O} \quad (1)$$

compounds [reaction (2)], which are sulfur species

$$\underset{\text{Elemental sulfur}}{S} + \underset{\text{Water}}{H_2O} + \underset{\text{Oxygen}}{1\tfrac{1}{2}O_2} \longrightarrow \underset{\text{Hydrogen ions}}{2H^+} + \underset{\text{Sulfate ion}}{SO_4^{2-}} \quad (2)$$

having readily removable electrons such as the sulfide ion (S^{2-}) and elemental sulfur (S).

These oxidation reactions require oxygen (O_2), so the microorganisms are aerobic. The microorganisms also need ammonium ions (NH_4^+), phosphate ions (PO_4^{3-}), and a few trace elements, such as magnesium (Mg^{2+}) and potassium (K^+). The bioleaching microorganisms are also acidophilic (acid loving), requiring the solution to be less than pH 2.5 and preferably higher than pH 1.0, created by sulfuric acid. This pH range ensures that ferrous iron (Fe^{2+}) is dissolved and readily available for oxidation. The bioleaching microorganisms tolerate very high concentrations of dissolved metals in their acidic environment (for example, 30 g/L of dissolved copper ions, Cu^{2+}). *See* OXIDATION-REDUCTION; SULFUR.

Bioleaching microorganisms can be grouped into temperature ranges at which they grow and function (see **table**). Mesophilic (middle-loving) bacteria function in the range 15–45°C (60–113°F). Moderately thermophilic (heat-loving) bacteria thrive in the range 40–65°C (104–150°F). Extremely thermophilic microorganisms, which are not bacteria but archaea (organisms that evolved from ancient Earth life forms), grow in the range 60–95°C (140–203°F). *See* ARCHAEBACTERIA.

The mechanism used by the bioleaching microorganisms to actually extract metals—such as copper, zinc, cobalt, and iron, from solid sulfide minerals—and cause these metals to dissolve has been studied and debated by scientists for decades. Free-floating microorganisms oxidize the dissolved Fe^{2+} and chemically reduced sulfur compounds. Some organisms attach to mineral surfaces using a biofilm composed of polymers that complex (bind) ferric iron (Fe^{3+}) to very high concentrations (53 g/L). The biofilm-mineral interface is the reaction zone where metal dissolution takes place. *See* BIOFILM.

Fe^{3+} is a very strong oxidizing agent that chemically attacks sulfide minerals, releasing metal ions and sulfur as either elemental sulfur (S) [reaction (3)] or other sulfur compounds such as sulfate

$$\underset{\text{Chalcopyrite}}{CuFeS_2} + \underset{\text{Ferric ion}}{4Fe^{3+}} \longrightarrow \underset{\text{Copper ion}}{Cu^{2+}} + \underset{\text{Ferrous ion}}{5Fe^{2+}} + \underset{\text{Elemental sulfur}}{2S} \quad (3)$$

(SO_4^{2-}) [reaction (4)]. Having removed an electron from the metal sulfide mineral in these reactions, Fe^{3+} is chemically reduced to Fe^{2+} (gaining an electron), which is quickly oxidized by the microorganisms, regenerating the oxidizing agent for dissolution

Examples of mesophilic, moderately thermophilic, and extremely thermophilic bioleaching microorganisms, their characteristics, and natural habitats

Temperature range	Genus and species examples and organism characteristics	Natural habitats
Mesophilic	*Acidithiobacillus ferrooxidans*: rod-shaped bacterium, approximately 0.5 μm × 1.0 μm; oxidizes ferrous iron (Fe^{2+}) and chemically reduced sulfur compounds	Acid springs; geological outcrops of rocks bearing sulfide minerals; acid mine waters; sulfide mineral rock piles exposed to the atmosphere
	Acidithiobacillus thiooxidans: rod-shaped bacterium, approximately 0.5 μm × 1.0 μm; oxidizes chemically reduced sulfur compounds	
	Leptospirillum ferrooxidans: rod-shaped bacterium, approximately 0.5 μm × 1.0 μm; oxidizes Fe^{2+} only	
Moderately thermophilic	*Sulfobacillus thermosulfidooxidans*: rod-shaped bacterium, approximately 1 μm by 2–3 μm; oxidizes Fe^{2+} and chemically reduced sulfur compounds	Acid, hot springs; volcanic areas; warm, acid mine waters; warm sulfide mineral rock piles exposed to the atmosphere
	Thiobacillus caldus: rod-shaped bacterium, approximately 1 μm × 2–3 μm; oxidizes chemically reduced sulfur compounds only	
Extremely thermophilic	*Acidianus brierleyi, Sulfolobus metallicus*; spherical-shaped archaea about 1 μm in diameter; oxidize Fe^{2+} and chemically reduced sulfur compounds	Hot, acid springs; volcanic areas

$$\underset{\text{Sphalerite}}{ZnS} + \underset{\text{Ferric ion}}{Fe^{3+}} + \underset{\text{Water}}{4H_2O} \longrightarrow \underset{\text{Zinc ion}}{Zn^{2+}} + \underset{\text{Sulfate ion}}{SO_4^{2-}} + \underset{\text{Ferrous ion}}{Fe^{2+}} + \underset{\text{Hydrogen ion}}{8H^+} \quad (4)$$

of the sulfide minerals [reaction (1)]. Chemically reduced sulfur compounds are also used by the microorganisms for energy [reaction (2)]. These redox (oxidation-reduction) reactions are catalyzed (accelerated) by the microorganisms [reactions (1)–(4)].

The oxidation of many mineral sulfides, including pyrite (FeS_2) by Fe^{3+}, results in acid (H^+) production, establishing the acidic environment for the microorganisms and keeping the metals dissolved for effective concentration and recovery [reaction (5)].

$$\underset{\text{Pyrite}}{FeS_2} + \underset{\text{Ferric iron}}{14Fe^{3+}} + \underset{\text{Water}}{8H_2O} \longrightarrow \underset{\text{Ferrous iron}}{15Fe^{2+}} + \underset{\text{Sulfate ion}}{2SO_4^{2-}} + \underset{\text{Hydrogen ion}}{16H^+} \quad (5)$$

The oxidation of some sulfide minerals releases heat (exothermic reaction). This is an important consideration in the type(s) of microorganisms used and the design of industrial-scale bioleaching operations. *See* BACTERIAL PHYSIOLOGY AND METABOLISM.

Industrial applications of bioleaching. The first industrial bioleaching applications were in the 1950s, when several major copper producers exploited microorganisms to scavenge copper from rocks containing very low amounts of copper in the form of sulfide minerals. This practice, called copper dump leaching, is used today.

In the 1980s, the fundamental principles of bioleaching were combined with standard metallurgical and dump bioleaching practices to create heap bioleaching and stirred-tank bioleaching processes. To this day, copper dump leaching, heap bioleaching, and stirred-tank bioleaching are employed industrially at very large scales throughout the world. The drivers for commercial bioleaching are (1) significantly reduced production costs owing to lower energy requirements, lower labor costs, and lower reagent usage; (2) diminished capital costs due to simpler equipment and rapid construction; (3) increased mineral reserves, because lower grade ores can be economically treated and varied mineral types can be processed; and (4) better environmental conditions and worker safety because there are no toxic gas emissions and, in some cases, no aqueous (water-based) discharges.

Copper dump bioleaching. Many of the world-class copper deposits are very large, low-grade (typically less than 1% copper), sulfide-mineral ore bodies, which are usually mined by open-pit methods. Ore grading higher than 0.5% copper is customarily processed by smelting. The vast tonnage of material grading lower than about 0.5% cannot be treated by conventional crushing, grinding, and smelting technologies, as the processing cost exceeds the value of the copper in the rock. Instead, these materials are blasted in the mine and hauled as large rock fragments, often exceeding 1 m (3.3 ft) in size, to dumps, where piles contain millions of tons of fragmented rock and reach 60 m (200 ft) high. Dilute acidic water is applied to the top surface of the dump, resulting in the growth of naturally occurring *Acidithiobacillus ferrooxidans*, *Acidithiobacillus thiooxidans, Leptospirillum ferrooxidans,* and other similar microorganisms throughout the rock pile. The microorganisms oxidize certain sulfur compounds, as well as Fe^{2+} to Fe^{3+}. In turn, Fe^{3+} oxidizes copper sulfide minerals, such as chalcocite (Cu_2S), covellite (CuS), to a lesser extent chalcopyrite ($CuFeS_2$), and pyrite (FeS_2), usually associated with copper deposits. Heat is generated and the temperature increases in the

dump. The hot gases rise and air for the microorganisms is pulled into the dump interior via convection from the porous sides of the rock pile. As the rock pile heats above the functional temperature (about 45°C) of the mesophilic bacteria, the moderately thermophilic microorganisms begin to grow (see table). When the temperature exceeds the upper limit (about 45°C) for the moderately thermophilic) microorganisms, they die. Insufficient studies have been done to determine whether the extremely thermophilic archaea replace the moderate thermophiles when the temperature in the dump exceeds 65°C (150°F). *See* OPEN PIT MINING; ORE AND MINERAL DEPOSITS.

Dissolved Cu^{2+}, resulting from bioleaching, is flushed from the rock pile by irrigation. The Cu^{2+}-containing solution is directed to a solvent extraction circuit, where the Cu^{2+} is stripped from the acidic water and the water is returned to the dump surface. Because less than optimal conditions exist for the microorganisms in the dump, the degradation of the sulfide minerals takes several decades. Nevertheless, dumps are a highly economical means of recovering copper, because an enormous volume of rock is processed and the production costs are low. *See* COPPER METALLURGY.

Heap bioleaching. The principles of heap bioleaching (also called bioheap leaching) are similar to dump leaching. However, heap bioleaching is done on ores with a higher copper grade or some gold content. This higher intrinsic value allows the system to be more highly engineered than dumps, resulting in greater metal recovery in less time.

Since the 1980s, heap bioleaching of the copper sulfide minerals Cu_2S and CuS has been widely practiced commercially for ores grading 1% Cu or higher. The ore is typically crushed to less than 19 mm (0.75 in.) and agglomerated (formed into a ball-shaped mass) with acidic water to bind finely ground rock into larger rock particles, hence improving air and water distribution in the heap. Large, earthen pads are covered with plastic and two arrays of perforated plastic pipe (one for air and the second for solution collection) are placed on the plastic. The ore is stacked 6–10 m (20–33 ft) deep on top of the plastic. The heap is drip irrigated with slightly acidic water and forced air is provided through the pipes under the crushed ore. Naturally occurring microorganisms proliferate, reaching 10^6 to 10^7 (1–10 million) per gram of ore and per milliliter of solution. Temperatures in the heap generally stay within the range for mesophilic bacteria (see table). Copper extraction of 85–90% is achieved in about a year.

When gold (Au) is associated with sulfide minerals, usually FeS_2 or arsenopyrite (FeAsS), micrometer-sized gold particles are embedded within the sulfide minerals. These gold ores are not amenable to direct cyanide treatment for gold extraction. The sulfide minerals first must be oxidized to liberate the gold. Since the mid-1990s, heap bioleaching of these ores has been practiced. Gold bioheaps increase in temperature when FeS_2 with embedded gold is oxidized [reaction (5)]. Temperatures can reach 70°C (170°F) or higher. To accommodate this temperature rise, the ore is agglomerated with all three groups of microorganisms (see table), with each group becoming active as its temperature range is achieved. *See* GOLD METALLURGY.

Gold is not dissolved when FeS_2 and FeAsS are oxidized; rather, the gold particles are liberated and remain as solids. After sufficient sulfide is oxidized (which can take up to 1.5 years depending on the ore particle size and other factors), the pretreated ore is removed from the pad, neutralized with lime, and chemically leached, usually with a dilute cyanide solution to dissolve the gold. The dissolved gold is processed into bullion. Bioleaching typically results in gold recoveries of 75–85%. Without bioleaching pretreatment, the gold recovery from these same ores would be 0–30%.

Aerated, stirred-tank bioleaching. Bioleaching is also accomplished in highly engineered, aerated, continuously stirred tank reactors (CSTR). Because of the high capital and operating costs of the process, the minerals must be of high value. Therefore, ores are preprocessed to remove worthless material, increasing the concentration of the valuable metal.

Stainless-steel tanks, each as large as 1380 m^3 (364,500 gallons), are equipped with agitators that maintain the finely ground sulfide mineral concentrate in suspension and ensure that oxygen is efficiently transferred to the microorganisms. Each tank is cooled to maintain the bioleach solution at the proper temperature for the microorganisms used. The **illustration** depicts a typical CSTR bioleach industrial plant.

The mineral concentrate is added to acidified water containing NH_4^+ and PO_4^{3-} ions and fed continuously to the first tank (stage) in a series of tanks. Because conditions are optimized for the microorganisms, numbers reach 10^9 to 10^{10} (one billion to

Industrial-scale, aerated, continuously stirred tank plant for bioleaching sulfidic-refractory gold concentrate. The large tanks are the bioleach circuit for pretreating the mineral concentrate and in front of them is the circuit for extracting the gold from the bioleached solids. (*Photo courtesy of BacTech Mining Corp.*)

10 billion) organisms per ml of solution. Continuously stirred tank reactors can be operated with any of the three microbial groups by controlling the temperature. Mineral concentrates move from the first-stage reactor(s) on a continuous basis through three to four CSTRs in series until the sulfide minerals are oxidized. This usually requires about a five-day residence time across the entire bioleaching circuit.

If the metal is a base metal, such as Cu^{2+} or Zn^{2+}, it dissolves when the mineral is oxidized. The oxidized slurry exits the final reactor and undergoes solid/liquid separation, after which the solid is discarded and the liquid is processed to recover the metal. If the mineral concentrate is gold bearing, the liquid from the last-stage reactor is neutralized and stabilized for safe disposal and the solid is processed to recover the gold. Metal recoveries in CSTR bioleach circuits are in the 95–98% range.

Using extremely thermophilic archaea, CSTR bioleaching of $CuFeS_2$ has recently been commercially developed. Unlike mesophilic and moderately thermophilic microorganisms, archaea have no cell walls. Improved agitator design has been incorporated to minimize shear forces and enhance oxygen mass transfer in the hot solutions, which diminish O_2 solubility. Highly resistant construction materials are used to resist the corrosive conditions due to the high temperatures and strongly oxidizing conditions.

Bioleaching of Nonsulfide Minerals

The dissolution of metals from minerals can be accomplished using microorganisms other than the acid-loving, chemolithotrophic microorganisms. Heterotrophic bacteria, yeast, and fungi require organic carbon for growth and produce organic acids or other substances that dissolve metals. Some heterotrophic organisms use inorganic compounds, such as metal oxides, rather than oxygen as electron acceptors. *See* FUNGI; YEAST.

One example of nonsulfide mineral bioleaching is the use of fungal strains of *Aspergillus* and *Penicillium* to produce bioacids that dissolve cobalt (Co^{2+}) and nickel (Ni^{2+}) from laterite, an iron and aluminum oxide material containing nickel and other base metals. The fungi oxidize organic compounds (such as sugars) and produce citric or lactic acids, which dissolve the laterite and release the metals. *See* FUNGAL BIOTECHNOLOGY; LATERITE.

Shewanella oneidensis, a bacterium found in soils, sediments, ground water, and surface water, oxidizes organic compounds and in the absence of oxygen (anaerobic conditions) passes the electrons to iron oxyhydroxides, such as hematite (Fe_2O_3) and goethite (FeOOH). This reaction dissolves the mineral. Although the thrust of this research is toward bioremediation of subsurface contaminants, the same approach might be considered for bioleaching a manganese dioxide (MnO_2) ore that hosts silver. The reduction of the manganese (Mn^{4+}) by the bacterially mediated electron transfer to the solid phase MnO_2 releases manganese as Mn^{2+} and allows the silver to be extracted.

Heterotrophic bioleaching requires large quantities of organic matter, and the process design of such systems poses significant problems, because unwanted microorganisms grow even in the absence of oxygen and consume the organic matter without releasing metals. Nonsulfide mineral bioleaching is not currently economical at commercial scale because of high costs, engineering complexities, and slow rates of metal dissolution.

Future Developments

Sulfide-mineral bioleaching relies on naturally occurring microorganisms. Despite revolutionary developments in molecular biology, its application in bioleaching is currently confined to microorganism identification. During bioleaching operations, occasional toxicity problems arise from certain anions (negatively charged ions), which might be mitigated by molecular modification of the microorganisms. However, if engineered microorganisms are used, the problem of competition from naturally occurring microorganisms must be resolved.

Chalcopyrite ($CuFeS_2$) is effectively bioleached in CSTRs using extremely thermophilic archaea and finely ground mineral concentrate. However, the greatest copper resource in the world is $CuFeS_2$ that is too low grade to be economically concentrated or extracted by any method other than heap bioleaching or in-situ (extraction of the metal without physical removal of the rock from the mineral deposit) bioleaching. Although we understand the chemical conditions under which $CuFeS_2$ is effectively bioleached in CSTRs, the challenge is to technically and cost-effectively establish and maintain suitable conditions for $CuFeS_2$ bioleaching in a bioheap or in-situ leach system.

In addition, more research is needed on the bioleaching of nonsulfide minerals to enhance reaction rates, find more effective microorganisms, and develop more effective engineering designs.

Corale L. Brierley

Bibliography. C. L. Brierley and A. P. Briggs, Selection and sizing of biooxidation equipment, *Mineral Processing Plant Design, Practice and Control*, Society for Mining, Metallurgy, and Exploration, Inc., Littleton, CO, pp. 1540–1568, 2002; H. L. Ehrlich, How microbes mobilize metals in ores: A review of current understandings and proposals for further research, *Miner. Metallurg. Process.*, 19:220–224, 2002; G. J. Olson, J. A. Brierley, and C. L. Brierley, Bioleaching review part B—Progress in bioleaching: Applications of microbial processes by the minerals industries, *Appl. Microbiol. Biotechnol.*, 63:249–257, 2003; G. J. Olson and T. R. Clark, Fundamentals of metal sulfide biooxidation, *Mining Eng.*, 56:40–46, 2004; D. E. Rawlings, Heavy metal mining using microbes, *Annu. Rev. Microbiol.*, 56:65–91, 2002; T. Rohwerder et al., Bioleaching review part A—Progress in bioleaching: Fundamentals and mechanisms of bacterial metal sulfide oxidation, *Appl. Microbiol. Biotechnol.*, 63:239–248, 2003.

Biological clocks

Self-sustained circadian (approximately 24-hour) rhythms, which regulate daily activities such as sleep and wakefulness, were described as early as 1729 by the French geologist de Mairan. He placed a plant in constant darkness and noticed that leaf movements continued to occur at specific times of day despite the absence of the day-night cycle. The notion that this rhythmicity was still driven in some fashion by the Earth's rotation was resolved in the midtwentieth century when it became clear that the period of self-sustained (free-running) oscillations usually does not match that of the environmental cycle, therefore the expression "approximately 24 hours." Moreover, the free-running period varies among species and also somewhat from one individual to another. Circadian rhythmicity is often referred to as the biological clock. *See* PHOTOPERIODISM; PLANT MOVEMENTS.

Almost all organisms display circadian rhythms, indicating an evolutionary benefit, most likely facilitating adaptation to the cyclic nature of the environment. Physiological processes that occur with a circadian rhythm range from conidiation (spore production) in the bread mold, *Neurospora crassa*, and leaf movements in plants to rest-activity behavior in animals. Despite the diversity of these phenomena, the basic properties of the rhythms are the same—they synchronize to environmental cues, predominantly light, but are maintained in the absence of such cues, and they display a constant periodicity over a wide temperature range.

Circadian rhythms in humans. In humans, circadian rhythmicity is manifested in the form of sleep-wake cycles, and control of body temperature, blood pressure, heart rate, and release of many endocrine hormones (melatonin, growth hormone, cortisol, and so on). This list is by no means complete, and as more is learned about human physiology and circadian rhythms, it is increasingly apparent that temporal ordering is a fundamental aspect of physiological processes. In fact, several disorders such as asthma, stroke, and myocardial infarction also tend to occur more frequently at certain times of the day. Awareness of circadian control has led to the concept of chronotherapeutics, which advocates drug delivery timed to the host's circadian rhythms. This concept is most popular in the treatment of cancer. Circadianly timed chemotherapy (chronochemotherapy), radiation therapy, or even surgery can enhance tumor regression and reduce toxic side effects.

While the importance of circadian rhythms in the disorders mentioned above is still under investigation, their role in jet lag or other problems related to light-dark synchronization is unquestionable. Shift work, travel to a different time zone, or any other activity that results in a lack of synchrony between the body circadian clock and the environment produces disturbances not only in the sleep-wake cycle but also in other body functions. In recent years, melatonin, a hormone produced by the pineal gland, has become the drug of choice for alleviating jet-lag symptoms, but its efficacy is debatable. While melatonin can induce temporary sleepiness and also produce a small shift in the circadian clock, it does not appear to be sufficient to synchronize all circadian rhythms. As a matter of fact, the function of melatonin in humans is unknown. In seasonal mammals such as hamsters, whose reproductive cycles are driven by day length (photoperiod), melatonin serves to convey length-of-day information. Basically the production of melatonin is restricted to the dark phase and thus defines the length of the night. In nonseasonal mammals, including humans, the release of melatonin cycles in a similar manner, but the functional significance of this is unknown. *See* PINEAL GLAND.

In humans, circadian rhythms have a potential role in psychiatric disorders, particularly those that cause depression. The disorder in which involvement of circadian rhythms has been best studied is seasonal affective disorder (SAD). Patients with SAD suffer from fall-winter depression, and there is some evidence to suggest that they have deficits in circadian rhythms, including poor entrainment to light. This is supported by the fact that SAD patients respond well to light therapy. Finally, circadian rhythms tend to be disorganized in older people, the effects being especially pronounced in aged individuals with some type of dementia such as Alzheimer's disease. The underlying basis has not been elucidated.

Localization of circadian clocks. In mammals the "master clock" controlling circadian rhythms is located in the hypothalamus, within a small group of neurons called the suprachiasmatic nucleus. Lesions of the suprachiasmatic nucleus in rodents eliminate all the rhythms mentioned above such as rest-activity cycles and rhythmic hormone release. Transplantation of a fetal suprachiasmatic nucleus will restore rest-activity cycles in a lesioned host, and the periodicity of the restored rhythm corresponds to that of the donor animal. How exactly the suprachiasmatic nucleus imparts temporal information to the rest of the organism is still unknown. Available data suggest that the suprachiasmatic nucleus transmits signals in the form of humoral factors as well as neural connections. The humoral hypothesis is best supported by experiments in which a suprachiasmatic nucleus encapsulated in a porous membrane was shown to be effective in restoring activity rhythms in a lesioned host. Since hormonal rhythms are not restored by these or by regular transplants, neural connectivity must be important also, probably more so for certain outputs. It should be stressed that while coordinated clock activity presumably requires a group of cells, approximately 24-h oscillations are generated within individual suprachiasmatic nucleus cells through the action of clock genes. *See* NERVOUS SYSTEM (INVERTEBRATE); NERVOUS SYSTEM (VERTEBRATE).

For many years the suprachiasmatic nucleus was thought to be the only site of a clock in mammals. This was in contrast to several other vertebrates where clocks were known to be present in the pineal gland and the eye as well. However, it is now clear that the mammalian eye also contains an oscillator (something that generates an approximately 24-h

cycle) whose activity can be assayed by measuring melatonin release in isolated retinas. Although the function of the eye clock is unknown, it is not difficult to envision a need for circadian control of visual sensitivity or some other related property. Perhaps more surprising is the accumulating evidence on circadian oscillators in other body regions.

Research done in fruit flies (*Drosophila melanogaster*) suggests that independent clocks are present in many different parts of the body. The rest-activity cycle in the fruit fly is controlled by a group of neurons in the central brain referred to as lateral neurons which are essentially the fly's suprachiasmatic nucleus. While these neurons may well constitute the master clock in flies, many other tissues also contain oscillators in that they express clock genes in a cyclic fashion even when severed from the lateral neurons. Since the identification of these oscillators is based upon the assay of molecular markers, as opposed to physiological processes, their functions are largely unknown. However, as in the case of the eye, it is easy to conceive of possible functions. For instance, the oscillator in the fly's prothoracic gland is most likely required to regulate eclosion (hatching of adult flies from pupae) which occurs predominantly at dawn and is known to be under circadian control. The oscillator in the malpighian tubule (kidney equivalent) may ef fect rhythmic excretion or osmoregulation.

The fruit fly study raises the tantalizing possibility that independent body clocks also occur in mammals, perhaps even humans. This idea is supported by the finding that recently identified mammalian clock genes are expressed in many tissues other than the suprachiasmatic nucleus and the retina. Several of these tissues also express molecules that are thought to serve as circadian photoreceptors. Finally, treatment of Rat-1 fibroblast cells with serum leads to the induction and subsequent cyclic expression of clock genes, again supporting the notion that oscillators can be generated outside the known clock tissues.

Genetic basis. The genetic basis of circadian rhythms was established through the identification of altered circadian patterns that were inherited. Such mutants were found first in *Drosophila* and then in *Neurospora* in the early 1970s. These two organisms, which are favored for their genetic amenability, became the systems of choice for those interested in pursuing the molecular dissection of circadian clocks. The earlier work in each of these organisms ended up focusing on a single gene that was the target of many mutations—*period* (*per*) in *Drosophila* and *frequency* (*frq*) in *Neurospora*. Both these genes are central components of the circadian clock.

The *Drosophila* and *Neurospora* studies set the stage for subsequent genetic analyses in these two systems and also in others. Mutations affecting circadian rhythms are now known in *Chlamydomonas*, *Arabidopsis thaliana*, cyanobacteria, hamsters, and mice. In addition, there is now an impetus to identify circadian abnormalities or naturally occurring variations in human populations. For instance, the difference between people that wake up and function most effectively in the early morning hours as opposed to those who prefer to sleep late into the morning may well lie in polymorphisms within clock genes.

Molecular basis. The question of how a clock is generated intrigued chronobiologists for many years. Theoreticians developed models to explain how oscillations could be generated and sustained, and these models invariably involved a feedback mechanism. It is now known that a feedback loop composed of cycling gene products that influence their own synthesis underlies overt rhythms in at least three organisms (*Drosophila*, *Neurospora*, and cyanobacteria) and most likely in a fourth (mammals). Similar feedback loops have also been found in plants, although it is not clear that they are part of the clock.

A genetic screen for new circadian rhythm mutations in *Drosophila* led to the identification of the *timeless* (*tim*) gene in the early 1990s. Like *per*, *tim*, can be mutated to yield diverse rhythm phenotypes, and loss of either gene product results in a loss of eclosion and behavioral activity rhythms. For example, the per^{s} mutation shortens the circadian period of eclosion and rest-activity rhythms to approximately 19 h, per^{l} lengthens period to approximately 29 h, and the per^{0} mutation eliminates rhythms altogether. Likewise, the tim^{0} mutation abolishes rhythms, the tim^{SL} mutation shortens circadian period by 0.5 h, and other lesions in *tim* can lengthen period to less than 30 h.

While the biochemical function of these genes has been difficult to assess, molecular analysis of their mode of regulation has been very informative. The two genes are coregulated and encode physically interacting and mutually dependent proteins that function in a common feedback loop (see **illus.**). The *per* and *tim* messenger ribonucleic acid (mRNA) levels cycle with a circadian rhythm, such that RNA levels are high at the end of the day or beginning of the night and low at the end of the night. The two proteins (PER and TIM) also cycle, with protein accumulation starting in the early evening and peaking in the middle of the night. As they are accumulating in the cytoplasm, PER and TIM associate to form heterodimers. This association stabilizes PER and permits nuclear entry of both proteins. Thus, while TIM does not require PER for stability, it is dependent on PER for nuclear transport. In the nucleus, either one or both proteins repress the synthesis of their own mRNAs. Turnover of the two proteins, TIM in the late night and PER in the early morning, allows RNA levels to rise once again and the cycle continues.

Advances in 1999 added three components to this feedback loop. Two of these proteins, *Clock* (CLK) and *Cycle* (CYC), are positive regulators of *per* and *tim*. Both these proteins are transcriptional activators, which means that they stimulate the production of *per* and *tim* RNA. They do so by interacting with specific deoxyribonucleic acid (DNA) sequences called E-boxes that are upstream of protein-coding regions in the *per* and *tim* genes. As might

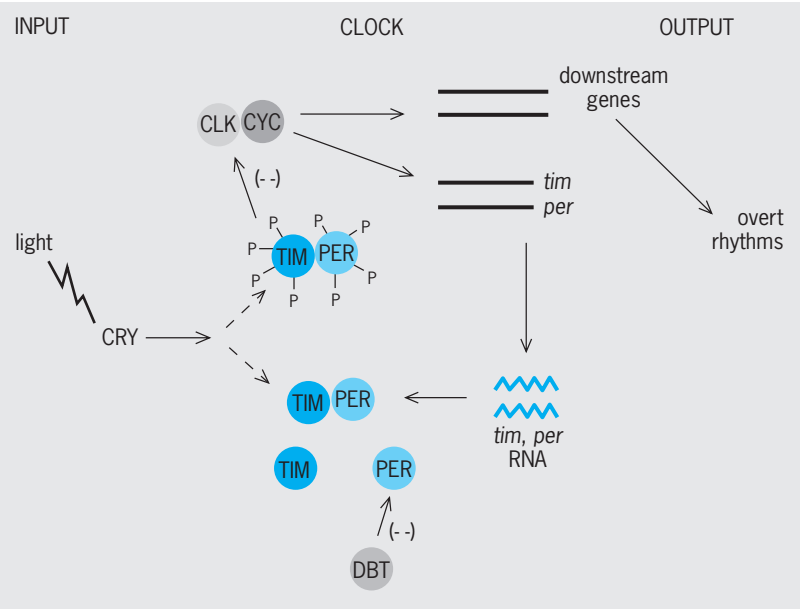

Model for the molecular clock with possible sites of interaction with input and output pathways. The negative signs reflect inhibitory influences of the proteins TIM and PER on CLK/CYC transcriptional activity, as well as the decrease in levels of PER monomers brought about by activity of the DBT casein kinase. While TIM levels are known to be reduced by light, negative notations are not used here because the relationship between cryptochromes (CRY) and TIM is not known.

be predicted, flies that are deficient in either CLK or CYC express very low levels of *per* and *tim* RNA as well as protein and, as a result, are arrhythmic. Since neither PER nor TIM appears to bind DNA, the current model for how they repress synthesis of their own mRNAs is based upon their interactions with CLK/CYC. Indeed, both PER and TIM bind CLK, and apparently this interaction inhibits transcriptional activation by CLK-CYC.

The third component added to the loop is a casein kinase produced by the *double-time* (*dbt*) gene. This kinase phosphorylates PER and renders it unstable in the absence of TIM. In a day-night cycle, as well as in free-running conditions, PER is increasingly phosphorylated as it accumulates, and maximally phosphorylated forms are found prior to the daily decay of the protein. Mutations in *dbt* affect period length by acclerating or delaying the rate of PER phosphorylation, which in turn affects stability of PER. Like PER, TIM is cyclically phosphorylated, but the mechanisms are not known.

In *Neurospora*, as well as in cyanobacteria, a similar transcription-based feedback loop underlies overt rhythms. However, the genes involved are very different, suggesting that clocks evolved more than once, but each time the same mechanism was used to assemble them. Thus, this type of feedback loop would appear to be the most efficient way of making a clock. Perhaps the only common feature of clocks in diverse organisms is a tendency to use proteins that contain PAS domains. PAS, named after the first three genes in which it was found (per, aryl hydrocarbon receptor nuclear translocator, single-minded), is a protein-protein interaction motif present in the PER, CLK, and CYC proteins. The *Neurospora frq* gene, which is the *per-tim* equivalent and functions in a feedback loop, does not contain a PAS domain, but each of the *white collar 1* (*wc1*) and *white collar 2* (*wc2*) genes, that are also important for *Neurospora* rhythms, contains a single PAS repeat. Given that the *wc* proteins are required for the light response of the *frq* clock gene and other PAS-containing proteins in lower organisms are associated with photoreception, it is conceivable that PAS proteins link photoreception pathways to central clocks.

In mammals, circadian rhythms appear to use homologs of the *Drosophila* genes with some additional complexity. For instance, there are three *per* genes, all of which cycle in the suprachiasmatic nucleus. Cyclic expression of mammalian *tim* (mtim) is controversial, but it is expressed in the suprachiasmatic nucleus and in a cell culture system, mammalian TIM, together with mammalian PER, can inhibit the transcriptional activity of a CLK-BMAL1 (mammalian version of CLK-CYC) heterodimer. Thus, while a feedback loop has not been demonstrated within an intact mammal, culture studies support its existence. In vivo studies are much more difficult in mammals. *Clock* is the only mammalian gene that was isolated as a result of a forward genetic screen for circadian rhythm mutations.

Synchronization to light. Light is the predominant environmental cue that synchronizes the circadian clock. The mechanisms used by the clock to entrain to light invariably involve a change in the levels of a clock component: in *Neurospora* levels of *frq* RNA are increased in response to light, in *Drosophila* levels of TIM protein are reduced, and in mammals levels of *mPer1* mRNA are increased (the other *mPer* and *mTim* RNAs can also be induced by light, but only at certain times of night). If one accepts that these molecules are timekeepers and their levels constitute time-of-day cues, then it follows that light would act by driving the levels in a particular direction, that is, up or down. It should be mentioned that in the mammalian suprachiasmatic nucleus other signaling pathways as well as several immediate early genes (for example, *c-fos*, *c-jun*) are also induced in response to light. The relevance of these molecules to the clock is not yet understood.

An area that is of great interest is the nature of the photoreception pathway used by the clock. In flies, ocular photoreceptors are not essential for circadian entrainment, although they are probably used, and may even dominate, when present. In mammals, there is general consensus that the eyes are required. However, the conventional photoreceptor cells, rods and cones, that mediate vision are dispensable for purposes of circadian entrainment. Again, this is not to say that rods and cones are incapable of transmitting signals to the clock; rather, other cell types, which presumably contain a novel photoreceptor, can carry out this function.

The nature of the novel photoreceptor in mammals is still debatable, although considerable attention is focused on the cryptochromes. Cryptochromes (CRY) are flavin-binding molecules that belong to a family whose members include

photolyases, which are DNA repair enzymes, and homologous proteins which are not photolyases and were first shown to be blue light photoreceptors in *Arabidopsis*. The latter are called cryptochromes, and it is now known that they function as circadian photoreceptors in plants and flies. In mammals, cryptochromes appear to be essential for free-running clock function, and their identity as photoreceptors remains to be proven. Expression of the cryptochromes cycles in both flies and mammals; in mammals this may be explained by their role in clock activity, but in flies it is of unknown significance. Interestingly, cryptochromes, like the clock genes, are expressed in many mammalian body tissues that, to date, are not associated with clock function.

Output signals. How the circadian clock conveys time-of-day information to the rest of the organism and produces rhythmic molecular, behavioral, and physiological outputs is a big unknown. At a molecular level, it is believed that clock genes affect output through transcriptional regulation. Thus, when PER-TIM enter the nucleus to regulate their own transcription, they presumably also regulate downstream genes that are targets of CLK/CYC. Genes that cycle under control of the clock have, in fact, been identified in several organisms, but the links between these genes and the clock or between these genes and the overt rhythm are yet to be determined.

The most direct link between clock genes and a known clock output comes from analysis of the gene encoding vasopressin, a peptide whose levels cycle in the suprachiasmatic nucleus. Suprachiasmatic nucleus vasopressin acts locally to regulate neuronal activity and is also released to other brain regions. A recent study showed that the vasopressin gene is transcriptionally regulated by the CLK, BMAL1, PER, and TIM proteins in a manner analogous to that described above for *per* and *tim*; that is, an E-box in the vasopressin promoter mediates transcriptional activation by CLK-BMAL1, and this activation is inhibited by PER-TIM. While these studies were conducted in a cell culture system, CLK-dependence of vasopressin cycling is supported by the observation that vasopressin does not cycle in CLK/CLK mutant mice.

The area of clock output is clearly going to be the focus of intense research for many years. It is likely that these studies will have relevance not only for chronobiology but also for a general understanding of physiological processes, such as sleep and hormone action, which happen to be controlled by the clock. *See* MIGRATORY BEHAVIOR; REPRODUCTIVE BEHAVIOR.

Amita Sehgal

Bibliography. J. Arendt, Complex effects of melatonin, *Therapie*, 53(5):479–488, 1998; A. R. Cashmore et al., Cryptochromes: Blue light receptors for plants and animals, *Science*, 284:760–765, 1999; J. C. Dunlap, Molecular bases for circadian clocks, *Cell*, 96:271–290, 1999; X. Jin et al., A molecular mechanism regulating rhythmic output from the suprachiasmatic circadian clock, *Cell*, 96:57–68, 1999; R. Lydic and H. A. Baghdoyan (eds.), *Handbook of Behavioral State Control, Cellular and Molecular Mechanisms*, Sec. 1: Mammalian circadian (24-hour) rhythms, CRC Press, 1999; D. Whitmore and P. Sassone-Corsi, Cryptic clues to clock function, *Nature*, 398:557–558, 1999; P. A. Wood and W. J. Hrushesky, Circadian rhythms and cancer chemotherapy, *Crit. Rev. Eukaryotic Gene Expression*, 6(4):299–343, 1996.

Biological oxidation

A biochemical reaction involving the transfer of a negatively charged electron from one organic compound to another organic compound or to oxygen. When a compound loses an electron, or is oxidized, another compound gains the electron, or is reduced. Oxidation-reduction (redox) reactions represent the main source of biological energy. Redox reactions occur simultaneously, and one does not occur without the other. The principal sources of reductants for animals are the numerous breakdown products of the major foodstuffs: carbohydrates, fats, and proteins. Energy release from these substances occurs in a stepwise series of hydrogen and electron transfers to molecular oxygen. *See* OXIDATION-REDUCTION.

Biological oxidation-reduction reactions. There are four prevalent varieties of biological redox reactions.

Direct transfer of electron. In one type of redox reaction, an electron is transferred directly from a donor to an acceptor [reaction (1)]. In this process, an elec-

$$\text{Cytochrome } c\,(Fe^{3+}) + \text{cytochrome } b\,(Fe^{2+}) \rightarrow \text{cytochrome } c\,(Fe^{2+}) + \text{cytochrome } b\,(Fe^{3+}) \quad (1)$$

tron is removed from the ferrous iron (Fe^{2+}) of a protein called cytochrome *b* and is transferred to the ferric iron (Fe^{3+}) of cytochrome *c*. Cytochrome *c* is the oxidizing agent, or oxidant, and cytochrome *b* is the reducing agent, or reductant. The cytochromes are a family of heme proteins; the heme group is an iron-porphyrin system similar to that found in hemoglobin, the protein that gives red blood cells their color. *See* CYTOCHROME.

Electrons transferred with protons. In a second type of redox reaction, electrons are transferred with protons in a process formally equivalent to the transfer of a hydrogen atom, which is made up of two positively charged protons and two negatively charged electrons ($2H^+ + 2e^-$). The enzyme pyruvate dehydrogenase functions in this fashion. This enzyme complex contains tightly bound flavin adenine dinucleotide (FAD). Pyruvate dehydrogenase mediates reaction (2).

$$\text{Pyruvate} + \text{coenzyme A} + \text{pyruvate dehydrogenase-FAD} \rightarrow \text{acetyl-coenzyme A} + \text{carbon dioxide} + \text{pyruvate dehydrogenase-FADH}_2 \quad (2)$$

Coenzyme A (for acylation) functions in the transfer of acyl [R—C(═O)—] groups, and pyruvate is a breakdown product of glucose. In reaction (2), pyruvate donates the equivalent of a hydrogen molecule to FAD, the oxidizing agent. *See* COENZYME; ENZYME.

Transfer of hydride ion. In a third type of redox reaction, a hydride ion ($H:^+$) is transferred; the hydride ion consists of a proton and two electrons and bears a negative charge. This is illustrated in the final portion of the pyruvate dehydrogenase reaction (3) where the hydride ion is transferred from

$$\text{Pyruvate dehydrogenase-FADH}_2 + \text{NAD}^+ \rightarrow \text{pyruvate dehydrogenase-FAD} + \text{NADH} + \text{H}^+ \quad (3)$$

tightly bound $FADH_2$ to nicotinamide adenine dinucleotide (NAD^+). NAD^+ functions as the oxidizing agent. FAD is alternately reduced by pyruvate and reoxidized by NAD^+. In order for such reactions to continue, the reduced coenzyme (NADH) has to be reoxidized by another substance. NAD^+ and nicotinamide dinucleotide phosphate ($NADP^+$) are derivatives of the vitamin niacin. Some enzymes use NAD^+ as oxidant, others use $NADP^+$, and still others can use either NAD^+ or $NADP^+$. *See* NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD^+); NICOTINAMIDE ADENINE DINUCLEOTIDE PHOSPHATE ($NADP^+$).

Electron transfer via direct reaction with oxygen. In a fourth type of oxidation-reduction, molecular oxygen (O_2) reacts directly with the reductant. This can be a monooxygenase or a dioxygenase process. In a monooxygenase reaction, one atom of O_2 is reduced to H_2O and the other is incorporated in the product [reaction (4)]. Phenylalanine, an amino acid, is

$$\text{Phenylalanine} + \text{tetrahydrobiopterin} + \text{O}_2 \rightarrow \text{tyrosine} + \text{H}_2\text{O} + \text{dihydrobiopterin} \quad (4)$$

oxidized to tyrosine, an amino acid that contains a hydroxyl group. The reaction is catalyzed by phenylalanine hydroxylase.

In a dioxygenase reaction, both atoms of molecular oxygen can be incorporated into an organic molecule [reaction (5)]. Both atoms of oxygen occur

$$\text{Homogentisate} + \text{O}_2 \rightarrow \text{4-maleylacetoacetate} \quad (5)$$

in 4-maleylacetoacetate. Homogentisate is a breakdown product of the amino acids phenylalanine and tyrosine.

Molecular oxygen exhibits two properties that make it a good oxidant. First, oxygen is very electronegative (second only to fluorine) and tends to attract electrons to it. Second, oxygen is electron-deficient. Each oxygen atom in the molecule lacks a complete octet of electrons. When oxygen reacts with two molecules of hydrogen, two molecules of water form, and the oxygen in water has a complete octet of electrons [reaction (6)], which is a more

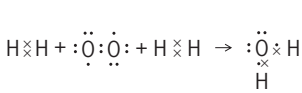

(6)

stable condition. (The electrons in molecular oxygen are represented by the dots, and the electrons in molecular hydrogen are represented by x's.)

Oxidative phosphorylation. Oxidative phosphorylation is responsible for most of the adenosine triphosphate (ATP) generated in animals, plants, and many bacteria. The transfer of electrons from NADH or $FADH_2$ to oxygen sustains the synthesis of ATP from adenosine diphosphate (ADP) and inorganic phosphate (P_i) by oxidative phosphorylation. Oxidative phosphorylation requires an electron transport chain in a membrane, O_2, and an ATP synthase. *See* ADENOSINE TRIPHOSPHATE (ATP).

Mitochondrial electron transport. The components of the electron transport chain are found in the inner mitochondrial membrane of animal and plant cells. The electron transport chain consists of four independent protein complexes (I, II, III, and IV). Coenzyme Q and cytochrome *c* transport electrons between the complexes. The transfer of electrons from reductant to oxidant is energetically favorable.

NADH, which is generated by pyruvate dehydrogenase [reaction (3)] and many other processes, donates its electrons to complex I (see **illus.**). Electrons are transferred from complex I to coenzyme Q to produce coenzyme QH_2. The flavin dehydrogenases, such as succinate dehydrogenase (complex II), transport electrons to coenzyme Q. (Succinate dehydrogenase is one of the enzymes of the Krebs tricarboxylic acid cycle.) Coenzyme QH_2, in turn, donates its electrons to complex III. Cytochrome *c* carries an electron from complex III to complex IV, or cytochrome oxidase. The cytochrome oxidase reaction of complex IV accounts for more than 95% of all oxygen consumed by humans. *See* CITRIC ACID CYCLE; MITOCHONDRIA.

Proton translocation in oxidative and photosynthetic phosphorylation. The chemiosmotic theory proposed by Peter Mitchell in 1961 provides the conceptual framework for understanding the mechanism of oxidative phosphorylation. The redox reactions of electron transport provide energy for the translocation of protons (H^+) from the inside (matrix) of the mitochondria to the space between its inner and outer membranes. Electron transport generates and maintains an electrochemical proton concentration difference across the inner membrane. An analogous

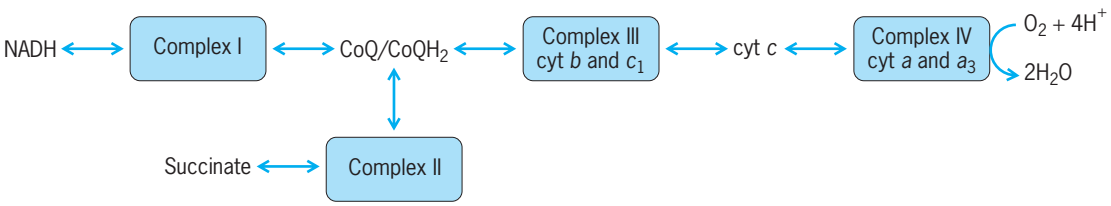

Overview of the mitochondrial electron transport chain.

situation occurs in aerobic bacteria, in which protons are translocated from the interior to the exterior of the cell membrane. The protons return into mitochondrial matrix, or bacterial cells, through a membrane-bound ATP synthase, and this energetically downhill process drives the energetically uphill conversion of ADP and P_i to ATP. Photophosphorylation in photosynthetic cells occurs by an analogous mechanism, except that light provides the energy to generate reductants for an electron transport chain. *See* PHOTOSYNTHESIS.

Reactive oxygen species. The reaction of oxygen with a single electron results in the formation of the superoxide anion radical [reaction (7)]. Superoxide

$$O_2 + e^- \rightarrow O_2^- \quad (7)$$

formation can occur by the spontaneous (nonenzymatic) reaction of oxygen with ferrous ion or reactive oxidation-reduction intermediates. Superoxide is a reactive substance that can cause the modification of cellular proteins, nucleic acids, and lipids in membranes, and is therefore toxic.

Superoxide dismutase is the enzyme that catalyzes the conversion of two superoxide anions and two protons to oxygen and hydrogen peroxide [reaction (8)]. Superoxide dismutase plays a pivotal role in pro-

$$2O_2^- + 2H^+ \rightarrow O_2 + H_2O_2 \quad (8)$$

tecting cells against oxygen toxicity.

Hydrogen peroxide is a reactive substance that can modify cellular macromolecules, and it is toxic. Hydrogen peroxide is converted to oxygen and water in a process mediated by the enzyme catalase [reaction (9)].

$$H_2O_2 \rightarrow H_2O + {}^1/_2O_2 \quad (9)$$

A third toxic oxygen derivative is the hydroxyl free radical. The hydroxyl free radical is more reactive and toxic than superoxide and peroxide. The hydroxyl free radical (denoted by the dot) can be formed through a nonenzymatic reaction of hydrogen peroxide with ferrous iron [reaction (10)]. The

$$H_2O_2 + Fe^{2+} \rightarrow Fe^{3+} + OH^- + OH^\bullet \quad (10)$$

hydroxyl free radical is also formed by the reaction of superoxide with hydrogen peroxide [reaction (11)].

$$H_2O_2 + O_2^- \rightarrow O_2 + OH^- + OH^\bullet \quad (11)$$

In addition, the hydroxyl free radical can be formed by the radiolysis of water produced by cosmic rays, x-rays, and other energetic electromagnetic radiation. The hydroxyl free radical reacts with and modifies cellular macromolecules. A widespread enzyme, glutathione peroxidase, mediates the destruction of hydroxyl free radicals [reaction (12)]. Glutathione

$$2OH^\bullet + 2GSH \rightarrow 2H_2O + GSSG \quad (12)$$

(GSH) is a peptide consisting of three amino acid residues, and its —SH group is the reductant. Moreover, a hydroxyl free radical can be destroyed by its reaction with ascorbate (vitamin C), β-carotene (vitamin A), or vitamin E. These actions may play a role in the beneficial and protective antioxidant effects of these vitamins. *See* ANTIOXIDANT.

Robert Roskoski, Jr.

Bibliography. J. M. Berg, J. L. Tymoczko, and L. Stryer, *Biochemistry*, 5th ed., 2001; C. Mathews, K. Van Holde, and K. Ahern, *Biochemistry*, 3d ed., 2000; D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 3d ed., 2000; D. G. Nicholls and S. J. Ferguson, *Bioenergetics 3*, 2002; R. Roskoski, *Biochemistry*, 1996.

Biological productivity

The amount and rate of production which occur in a given ecosystem over a given time period. It may apply to a single organism, a population, or entire communities and ecosystems. Productivity can be expressed in terms of dry matter produced per area per time (net production), or in terms of energy produced per area per time (gross production = respiration + heat losses + net production). In aquatic systems, productivity is often measured in volume instead of area. *See* BIOMASS.

Ecologists distinguish between primary productivity (by autotrophs) and secondary productivity (by heterotrophs). Plants have the ability to use the energy from sunlight to convert carbon dioxide and water into glucose and oxygen, producing biomass through photosynthesis. Primary productivity of a community is the rate at which biomass is produced per unit area by plants, expressed in either units of energy [joules/(m^2)(day)] or dry organic matter [kg/(m^2)(year)]. The following definitions are useful in calculating production: Gross primary production (GPP) is the total energy fixed by photosynthesis per unit time. Net primary production (NPP) is the gross production minus losses due to plant respiration per unit time, and it represents the actual new biomass that is available for consumption by heterotrophic organisms. Secondary production is the rate of production of biomass by heterotrophs (animals, microorganisms), which feed on plant products or other heterotrophs. *See* PHOTOSYNTHESIS.

Productivity is not spread evenly across the planet. For instance, although oceans cover two-thirds of Earth's surface, they account for only one-third of the Earth's productivity: it is estimated that the global terrestrial primary productivity is approximately 120×10^9 tons of dry weight per year, while the sea primary productivity is estimated at approximately 60×10^9 tons per year. Furthermore, the factors that limit productivity in the ocean differ from those limiting productivity on land, producing differences in geographic patterns of productivity in the two systems. In terrestrial ecosystems, productivity shows a latitudinal trend, with highest productivity in the tropics and decreasing progressively toward the Poles; but in the ocean there is no latitudinal trend, and the highest values of net primary production are found along coastal regions.

In the ocean, primary productivity is controlled by light and nutrient availability. Oceans do not show a gradient of productivity from the Equator to the Poles. Instead, there is a vertical productivity gradient, depending on the rate at which light decreases with depth. Primary production in the oceans is relatively low in the surface, due to excessive light, and is highest at 10–30 m (33–100 ft) depth. If light were the main factor controlling productivity in the ocean, an increasing productivity gradient should be seen in moving toward the Equator. However, some parts of the tropics and subtropics are very unproductive, while the Antarctic Ocean is one of the most productive oceanic regions. In the ocean, nutrients are often the factor limiting primary productivity. Nutrient concentration tends to be high in the deep water and low in the surface layers, where plankton lives. The highest-net-productivity areas in the ocean are found along coastal regions, where seawater mixes with nutrient-rich estuarine water, and in regions of upwelling, where nutrient-rich deep-ocean water is brought to the surface. As a consequence, in the ocean there are areas of high productivity even at high latitudes and low temperatures. The nutrients that most commonly limit productivity in the ocean are phosphorus, nitrogen, and iron. Different areas in the ocean are limited by different nutrients. An experiment in the Sargasso Sea demonstrated that nutrients are the limiting factor in this area of low productivity. If nitrogen and phosphorus but no iron were added, primary productivity barely increased. The addition of iron stimulated primary productivity but only for a short time. The addition of phosphorus, nitrogen, and iron increased productivity for prolonged periods, suggesting that although iron is the most limiting nutrient, availability of phosphorus and nitrogen in the region is also very low.

In terrestrial communities, primary productivity is correlated to actual evapotranspiration: the amount of water lost to the atmosphere by evaporation from the ground and by transpiration from the vegetation. Evapotranspiration is a measure of solar radiation, temperature, and rainfall, and primary productivity increases with actual evapotranspiration. Direct radiation from the Sun and temperature tend to increase toward the Equator. When water is not limited, photosynthetic rate increases with radiant energy and with temperature, and this translates into higher primary productivity. Other factors, such as leaf area and water-holding capacity of the soil, will affect evapotranspiration, and thus primary production. For example, in areas with no water shortage, plants will tend to have broad leaves, which will allow them to capture solar radiation and increase the circulation of air inside the leaves, where photosynthesis occurs. In these areas, primary productivity will tend to be high. In contrast, in areas where water is a limiting resource, plants will tend to have smaller leaf area to reduce water loss, reducing transpiration rate but also reducing photosynthetic rate. Finally, primary productivity in terrestrial communities can be limited by nutrients—nitrogen being the most common limiting nutrient.

To illustrate productivity at an ecosystem level, a terrestrial ecosystem is used in **Table 1**. For a lightly grazed shortgrass prairie, the gross rate of photosynthesis was 21,900 kJ/m^2 over 180 days. After accounting for respiration losses from the plants (7450 kJ), a net production of 14,455 kJ remained. This net primary production, partitioned as above- and belowground, flows principally into the decomposition pathways. The predominant flow of energy of primary production thus goes to the decomposers (bacteria, fungi, and detritivores). Although animal heterotrophs account for only a small fraction of the total energy flow (2–4%), they may play a synergistic or catalytic role in subsequent new plant production.

TABLE 1. Energy flow of autotrophic and heterotrophic production and respiration in lightly grazed shortgrass prairie at Pawnee Site, Colorado, 1972*

Energetic parameters	Production, kJ/m^2	Respiration, kJ/m^2
Autotrophic activity		
Gross primary production	21,900	
Net primary production	14,455	7,445
Aboveground	2,165	5,725
Belowground	12,290	1,720
Heterotrophic activity		
Herbivory	99	242
Decomposers	2,930	9,640

*"Lightly grazed" indicated 1 steer per 10.8 hectares or per 4.32 acres over 180 years.
SOURCE: Modified from D. C. Coleman et al., Energy flow and partitioning in selected man-managed and natural ecosystems, *Agroecosystems*, 3:45–56, 1976.

Production of agroecosystems and natural systems can be compared using net ecosystem production (NPP − respiration of heterotrophs) and ecosystem metabolism (GPP/net ecosystem respiration = autotrophic + heterotrophic respiration). Production and respiration in more mature ecosystems are nearly in balance, while there is more harvestable production in an agroecosystem (production greater than respiration) [**Table 2**]. Unfortunately, there is little information for many agroecosytems concerning belowground energy costs. Much of the existing information on production refers to readily harvested fibrous material; yet there may be considerable amounts of root exudates and sloughed cells produced in addition to the fibrous roots. These exudates represent an additional "cost of doing business" for plants in either old-growth forest or annual field crops. For example, it has been estimated that perhaps 23–25% of the carbon fixed by photosynthesis in a rice crop goes to furnish root exudates for rhizosphere (root-associated) microflora active in nitrogen fixation. This is an important link between carbon-energy flow and nutrient cycling. *See* AGROECOSYSTEM.

The production and activity of certain microflora may be of considerable importance to plant growth in a wide variety of terrestrial ecosystems. Thus various beneficial root-associated fungi (mycorrhizae) may facilitate uptake of a limiting nutrient, such as

TABLE 2. Annual production and respiration [kJ/(m^2)(year)] in growth-type and more mature ecosystems*

Energetic parameters	Types of ecosytems			
	Alfalfa field (U.S.)	Young pine plantation (U.K.)	Shortgrass prairie (Table 1)	80-year deciduous forest (Tennessee)
Gross primary production (GPP)	102,190	51,095	21,900	117,165
Autotrophic respiration (R_A)	38,530	19,685	7,445	76,220
Net primary production (NPP)	63,660	31,410	14,455	40,940
Heterotrophic respiration (R_H)	12,565	19,265	9,965	38,410
Net ecosystem production (NEP = NPP$-R_H$)	51,095	12,145	4,495	2,530
Net ecosystem respiration ($R_E = R_A + R_H$)	51,095	38,950	17,410	114,635
Ecosystem metabolism (GPP/R_E)	2.00	1.31	1.26	1.02

*Systems are arranged in decreasing magnitudes of net ecosystem production.

phosphorus, permitting a doubling of harvestable shoot production of plants such as clover. However, the net carbon energy cost to the plant is quite low: approximately 1% of the total carbon allocation of the host plant. *See* MYCORRHIZAE.

It is apparent that the important aspects of biological productivity must be defined carefully. Depending on management techniques, the amount and activity of beneficial microorganisms, such as mycorrhizae, may have an importance equal to that of the crop species of interest. A holistic approach should be important to manage ecosystems using lower inputs of fossil fuels for tillage and fertilizers. *See* ECOLOGICAL ENERGETICS; ECOSYSTEM; FOOD WEB.

David C. Coleman; Ellen Gryj

Bibliography. M. Begon, J. L. Harper, and C. R. Townsend, *Ecology*, Blackwell Science, Cambridge, MA, 1996; K. Blaxter, *Energy Metabolism in Animals and Man*, Cambridge University Press, New York, 1989; R. L. Hanson, Evapotranspiration and droughts, in R. W. Paulson et al. (eds.), *National Water Summary 1988-89: Hydrological Events and Floods and Droughts*, 1991; C. J. Krebs, *Ecology*, Harper Collins, New York, 1994; S. J. McNaughton, Serengeti migratory wildebeest: Facilitation of energy flow by grazing, *Science*, 191:92–94, 1976; W. H. Schlesinger, *Biogeochemistry: An Analysis of Global Change*, Academic Press, San Diego, 1997.

Biological specificity

The orderly patterns of metabolic and developmental reactions giving rise to the unique characteristics of the individual and of its species. It is one of the most widespread and characteristic properties of living organisms. Biological specificity is most pronounced and best understood at the cellular and molecular levels of organization, where the shapes of individual molecules allow them to selectively recognize and bind to one another. The main principle which guides this recognition is termed complementarity. Just as a hand fits perfectly into a glove, molecules which are complementary have mirror-image shapes that allow them to selectively bind to each other.

This ability of complementary molecules to specifically bind to one another plays many essential roles in living systems. For example, the transmission of specific hereditary traits from parent to offspring depends upon the ability of the individual strands of a deoxyribonucleic acid (DNA) molecule to specifically generate two new strands with complementary sequences. Similarly, metabolism, which provides organisms with both the energy and chemical building blocks needed for survival, is made possible by the ability of enzymes to specifically bind to the substrates whose interconversions they catalyze. *See* DEOXYRIBONUCLEIC ACID (DNA); ENZYME; METABOLISM.

In addition to metabolism and DNA replication, many other attributes of living systems depend upon specific recognition between molecules. During embryonic development, for example, individual cells associate with each other in precise patterns to form tissues, organs, and organ systems. These ordered interactions are ultimately dependent upon the ability of individual cells to recognize and specifically bind to other cells of a similar type. The existence of such preferential cell-cell interactions can be demonstrated experimentally by isolating individual cells from two different tissues and then mixing the two populations together in a test tube. If liver cells are mixed with cartilage cells, for example, the liver cells will preferentially bind to other liver cells, while the cartilage cells will adhere to other cartilage cells. This ability of cells to selectively bind to other cells of the same type is based upon specific recognition between molecules located at the cell surface. *See* CELL ADHESION; DEVELOPMENTAL BIOLOGY.

In addition to binding to one another, cells can also interact by releasing hormones into the bloodstream. Though all of an organism's cells are exposed to hormones circulating in the bloodstream, only a small number of target cells respond to any particular hormone. This selectivity occurs because the specific receptor molecules to which hormones bind are restricted to certain cell types. Thus each hormone exerts its effects on a few selected cell types because only those cells contain the proper receptor. Specific receptors are also involved in interactions between neurotransmitters and the cells they stimulate or

inhibit, between certain types of drugs and the cells they affect, and between viruses and the cells they infect. This last phenomenon has an important influence on the susceptibility of individuals to virally transmitted diseases. The reason that humans are resistant to infection by viruses which destroy bacteria, for example, is that human cells do not contain the cell-surface receptors present in bacterial cells. *See* ENDOCRINE MECHANISMS; HORMONE.

Although most examples of biological specificity are based upon interactions occurring at the molecular level, such phenomena affect many properties manifested at the level of the whole organism. The ability of individuals to defend against infectious diseases, for example, requires the production of antibody molecules which specifically bind to bacteria and viruses. The fertilization of an egg by a sperm is facilitated by specific recognition between molecules present on the surfaces of the sperm and egg cells. Finally, even communication between organisms can be mediated by specific chemical signals, called pheromones. Such chemical signals are utilized in trail marking by ants and bees, in territory marking by certain mammals, and as sexual attractants. Specific molecular interactions thus exert influences ranging from the replication of genes to the behavior of organisms. *See* CHEMICAL ECOLOGY; FERTILIZATION (ANIMAL); IMMUNOLOGY; MOLECULAR BIOLOGY; PHEROMONE. Lewis J. Kleinsmith

Biologicals

Biological products used to induce immunity to various infectious diseases or noxious substances of biological origin. The term is usually limited to immune serums, antitoxins, vaccines, and toxoids that have the effect of providing protective substances of the same general nature that a person develops naturally from having survived an infectious disease or having experienced repeated contact with a biological poison. As a matter of governmental regulatory convenience, certain therapeutic substances which have little to do with conferring immunity have been classified as biological products primarily because they are derived from human or animal sources and are tested for safety by methods similar to those used for many biological products. *See* IMMUNITY.

Immune serums. One major class of biologicals includes the animal and human immune serums. All animals, including humans, develop protective substances in their blood plasma during recovery from many (but not all) infectious diseases or following the injection of toxins or killed bacteria and viruses. These protective substances called antibodies usually are found in the immunoglobulin fraction of the plasma and are specific since they react with and neutralize only substances identical or closely similar to those that caused them to be formed. Thus serum containing antibodies induced by injections of diphtheria toxin neutralizes diphtheria toxin, but not tetanus toxin. *See* ANTIBODY; IMMUNOGLOBULIN; SERUM.

Antibody-containing serum from another animal is useful in the treatment, modification, or prevention of certain diseases of humans when it is given by intramuscular or intravenous injection. The use of these preformed "borrowed" antibodies is called passive immunization, to distinguish it from active immunization, in which each person develops his or her own antibodies. Passive immunization has the advantage of providing immediate protection, but it is temporary because serum proteins from other animals and even from other humans are rapidly destroyed in the recipient.

Types. Serums which contain antibodies active chiefly in destroying the infecting virus or bacterium are usually called antiserums or immune serums; those containing antibodies capable of neutralizing the secreted toxins of bacteria are called antitoxins. Immune serums have been prepared to neutralize the venoms of certain poisonous snakes and black widow spiders; they are called antivenins.

Because all products used for passive immunization are immune serums, or globulin fractions from such serums, they are named to indicate the diseases that they treat or prevent, the substances that they inhibit or neutralize, the animal from which they came, and whether they are whole serums or the globulin fractions thereof. Thus there is, for example, antipertussis immune rabbit serum, measles immune globulin (human), diphtheria antitoxic globulin (horse), tetanus immune globulin (human), and anti-Rh_0 (D) gamma globulin (human).

Preparation. The general methods of preparation of various immune serums and antitoxins differ only in details. Horses, cows, rabbits, or humans are injected with slowly increasing amounts of either the virus, bacterium, toxin, erythrocyte, or venom as antigens against which immunity is to be developed. These agents are usually treated with chemicals to render them noninfectious or nontoxic before injection. Injections are given over several months and the animals respond by producing relatively large amounts of specific antibody against the injected material. When the antibody concentration has reached a high level in the blood, the animals are bled and the blood is allowed to clot. The liquid serum can be separated from the blood cells after the clot has contracted. This serum is usually fractionated chemically to obtain the concentrated antibody-globulin fraction free of albumin and other nonantibody fractions. The antibody-globulin fraction may be digested with enzymes to remove other extraneous components. The antibody solution is then sterilized by filtration through a bacteriaproof filter, and a small amount of preservative antibacterial chemical is added to protect the solution from bacterial contamination during administration.

Standardization. Each lot of immune serum or globulin concentrate is tested to make certain that it is sterile, free of fever-producing impurities (pyrogens), and free of accidentally introduced poisonous materials. In addition, each serum is tested to determine the antibody concentration in it. Although methods vary, depending upon the type of immune

serum, potency is usually measured by determining the amount of antibody that is required to protect mice or guinea pigs from the fatal effects of a standard amount of the living infectious agent or of the corresponding toxin. When the protective dose of the unknown serum is compared to the protective dose of a standard antitoxin or reference immune serum, it is possible to calculate the relative potency in the serum or globulin concentrate that is being tested. The dosages of antitoxins are usually specified in units. The dosages of antibacterial or antiviral immune serums or globulin concentrates are usually expressed in milliliters.

Disadvantages. The use of animal immune serums for prevention of therapy in humans has certain disadvantages. The serum proteins themselves may cause the production of specific antibodies in the recipient of the immune serum, and thus the person may become allergically sensitized to the serum protein of this animal species. Therefore, a second dose of serum from the same species of animal given weeks, months, or years later may cause an acute anaphylactic reaction that may be fatal. It is advisable, particularly in cases in which serum may have been given to a patient previously, to test for hypersensitivity by injecting a tiny amount of serum into the skin of the patient. A positive reaction indicating hypersensitivity is evidenced by the rapid development of itching, inflammation, and swelling at the site of the injection. *See* ANAPHYLAXIS; HYPERSENSITIVITY.

A second and less acute type of reaction resulting from the injection of animal immune serum is so-called serum sickness. About 4–10 days after the administration of serum, the patient develops fever, urticarial (hivelike) rash, and sometimes joint pains. These reactions are not ordinarily fatal.

Many people have significant amounts of measles and poliomyelitis antibodies in their serums as a result of recovery from either the clinical diseases or inapparent infections. Therefore, the globulin concentrates from normal pooled human serums collected at random have sufficient antibody to be useful in preventing or moderating these infections in persons who have been exposed. Such human serum proteins rarely produce the allergic hypersensitivity problems that often follow the passive transfer of animal serum. Enzyme-digested globulins of human origin have the added advantage of intravenous administration.

Active immunity production. Products used to produce active immunity constitute the other large class of biological products. They contain the actual toxins, viruses, or bacteria that cause disease, but they are modified in a manner to make them safe to administer. Because the body does not distinguish between the natural toxin or infectious agent and the same material when properly modified, immunity is produced in response to injections of these materials in a manner very similar to that which occurs during the natural disease. Vaccines are suspensions of the killed or attenuated (weakened) bacteria or viruses or fractions thereof. Toxoids are solutions of the chemically altered specific bacterial toxins which cause the major damage produced by bacterial infections. Biological products producing active immunity are usually named to indicate the disease they immunize against and the kind of substance they contain: thus typhoid vaccine, diphtheria toxoid, tetanus toxoid, measles vaccine, mumps vaccine, and poliomyelitis vaccine. *See* VACCINATION.

Antigen modification. It is most important that vaccines and toxoids be altered in a way to make them safe, but not so drastically altered that they will no longer produce immunity. Certain strains of smallpox, measles, yellow fever, mumps, and polio viruses, and the tubercle bacillus have been attenuated so that, although they can still undergo limited multiplication in the body, they do not produce progressive infection or symptoms of disease. Vaccines made from these strains are called attenuated vaccines. Other viruses are rendered completely noninfectious by the addition of small amounts of formaldehyde, and bacteria may be killed with formaldehyde, phenol, or mercurial compounds. Vaccines made from these killed infectious agents are called killed vaccines. Typhoid vaccine, pertussis (whooping cough) vaccine, and Salk polio vaccine are examples of this type of product. Many bacterial toxins can be made nontoxic by incubation at 98.6°F (37°C) with small amounts of formaldehyde. These altered toxins are called toxoids and stimulate the body to produce antitoxin capable of specifically neutralizing the toxin. Diphtheria toxoid and tetanus toxoid are the two most common toxoids used to immunize humans.

Vaccine preparation. Bacterial vaccines are usually prepared by growing specially selected strains in suitable culture media, harvesting the bacteria from the media, and suspending them in saline. A killing agent is added to the bacterial suspension and is allowed to act at about 98.6°F (37°C) until no bacteria remain alive. The bacteria may be extracted or chemically fractionated to purify the immunizing antigen. The suspensions or purified antigens are then diluted with physiological saline to the desired potency. These vaccines, after testing for sterility, lack of abnormal toxicity, and immunizing potency, are ready for use.

Because viruses must be grown in the presence of living cells, these agents are inoculated into embryonated eggs, live animals, or cultures of living cells growing in bottles in the laboratory. The egg fluids, animal tissue juices, or the supernatant fluid of tissue cultures then contain the viruses. Treatment with formaldehyde destroys the infectivity of viruses, and these virus-containing solutions may be used for immunization after suitable tests to ensure safety, sterility, and potency. Live virus vaccines are made with viruses attenuated by serial passage in various species of animals or tissue cultures. They are grown in living cells and may be used for immunization after suitable tests to ensure safety, bacterial sterility, potency, and freedom from adventitious agents. Live virus vaccines may be given orally (Sabin poliomyelitis vaccine) or injected (measles vaccine). Live virus vaccines produce subclinical infections that confer

a durable immunity equivalent to that which occurs following the natural disease.

Toxoids. With certain toxoids the addition of a small amount of alum, aluminum hydroxide, or aluminum phosphate precipitates the toxoid molecules and changes the solution to a suspension that is more effective when injected than the original toxoid solution. These precipitated toxoids may be given in smaller amounts to produce equal or greater immunity than that caused by the plain toxoid. Alum-precipitated diphtheria toxoid is widely used in preference to plain toxoid.

Multiple immunization. Because it is frequently expedient to immunize people simultaneously against several diseases, various immunizing agents have been combined so that they can be given in a single series of two or three doses. A combination of pertussis vaccine, diphtheria toxoid, and tetanus toxoid with alum added has become very widely used as an immunizing agent for babies. Vaccines against three types of poliomyelitis have been added to this mixture. Combinations of typhoid vaccine and tetanus toxoid and of measles and smallpox vaccines have been used.

The combination of a large number of different immunizing agents within a single immunization dosage is an exceedingly complex problem because dosage volume must be kept small and the various agents must be compatible. This has resulted in much research on purification and concentration of antigens to reduce their volumes and eliminate toxic impurities. Tetanus and diphtheria toxoids and influenza vaccine have been considerably purified, and chemical fractions of various pathogenic bacteria (pertussis and cholera vaccines) are used as immunizing agents.

Veterinary use. Because livestock and pets as well as human beings require protection from infectious diseases, there is an extensive assortment of veterinary biological products. Many of these are vaccines and toxoids prepared in the same way as are products for human use.

Blood derivatives. In addition to products for active and passive immunization, a few other substances of human or animal origin are also classed (mostly for regulatory convenience) as biological products. One group of these is normal human blood plasma and its derivatives. Whole plasma and serum albumin are used to combat hemorrhage and shock, and other fractions are used in the treatment of hemophilia and for their clotting properties. The use of anti-Rh_0 (D) gamma globulin in Rh-negative women at parturition inhibits Rh sensitization and erythroblastosis fetalis in subsequent offspring. Although they do not produce immunity to infectious diseases, these substances require testing for sterility, safety, and freedom from pyrogens as to the immune serums. *See* BLOOD GROUPS; RH INCOMPATIBILITY.

Diagnostic agents. Another group of biological products consists of a miscellany of reagents used in the diagnosis of infectious diseases. These include immune serums for the serological typing and identification of pathogenic bacteria and viruses, and various antigens for the detection of antibodies in patients' serums as a means of assisting in diagnosis. In addition to these substances used in the laboratory diagnosis of disease, certain extracts of bacteria and fungi have been prepared for injection into the skin to detect allergic hypersensitivity. In certain diseases such as tuberculosis and brucellosis, the patient usually becomes hypersensitive to certain components of the infecting organism. Extracts of these organisms injected into the skin of such a person cause inflammation, whereas no reaction occurs in a person who has not been infected. These skin tests therefore are of assistance in the diagnosis of various infectious diseases.

Allergenic products are administered to humans for the diagnosis, prevention, or treatment of allergies to various plant and animal substances. They usually are suspensions of the allergenic substance (pollens, feathers, hairs, or danders) in 50% glycerin. *See* IMMUNOLOGY. Lee F. Schuchardt

Bibliography. E. A. Kabat, *Experimental Immunochemistry*, 1971; M. J. Pelczar, Jr., E. C. Chan, and N. R. Krieg, *Microbiology*, 5th ed., 1986; M. M. Wintrobe, *Blood, Pure and Eloquent*, 1980.

Biology

A natural science concerned with the study of all living organisms. Although living organisms share some unifying themes, such as their origin from the same basic cellular structure and their molecular basis of inheritance, they are diverse in many other aspects. The diversity of life leads to many divisions in biological science involved with studying all aspects of living organisms. The primary divisions of study in biology consist of zoology (animals), botany (plants), and protistology (one-celled organisms), and are aimed at examining such topics as origins, structure, function, reproduction, growth and development, behavior, and evolution of the different organisms. In addition, biologists consider how living organisms interact with each other and the environment on an individual as well as group basis. Therefore, within these divisions are many subdivisions such as molecular and cellular biology, microbiology (the study of microbes such as bacteria and viruses), taxonomy (the classification of organisms into special groups), physiology (the study of function of the organism at any level), immunology (the investigation of the immune system), genetics (the study of inheritance), and ecology and evolution (the study of the interaction of an organism with its environment and how that interaction changes over time).

The study of living organisms is an ongoing process that allows observation of the natural world and the acquisition of new knowledge. Biologists accomplish their studies through a process of inquiry known as the scientific method, which approaches a problem or question in a well-defined orderly sequence of steps so as to reach conclusions. The first step involves making systematic observations, either directly through the sense of sight, smell, taste,

sound, or touch, or indirectly through the use of special equipment such as the microscope. Next, questions are asked regarding the observations. Then a hypothesis—a tentative explanation or educated guess—is formulated, and predictions about what will occur are made. At the core of any scientific study is testing of the hypothesis. Tests or experiments are designed so as to help substantiate or refute the basic assumptions set forth in the hypothesis. Therefore, experiments are repeated many times. Once they have been completed, data are collected and organized in the form of graphs or tables and the results are analyzed. Also, statistical tests may be performed to help determine whether the data are significant enough to support or disprove the hypothesis. Finally, conclusions are drawn that provide explanations or insights about the original problem. By employing the scientific method, biologists aim to be objective rather than subjective when interpreting the results of their experiments. Biology is not absolute: it is a science that deals with theories or relative truths. Thus, biological conclusions are always subject to change when new evidence is presented. As living organisms continue to evolve and change, the science of biology also will evolve. *See* ANIMAL; BOTANY; CELL BIOLOGY; ECOLOGY; GENETICS; IMMUNOLOGY; MICROBIOLOGY; PLANT; TAXONOMY; ZOOLOGY.

Lee Couch

Bibliography. N. A. Campbell, *Biology*, 3d ed., 1993; H. Curtis and N. S. Barnes, *Biology*, 5th ed., 1989; R. H. Saigo and B. Saigo, *Careers in Biology II*, 1988.

Bioluminescence

The emission of light by living organisms that is visible to other organisms. The enzymes and other proteins associated with bioluminescence have been developed and exploited as markers or reporters of other biochemical processes in biomedical research. Bioluminescence provides a unique tool for investigating and understanding numerous basic physiological processes, both cellular and organismic.

Luminous organisms. Although rare in terms of the total number of luminous species, bioluminescence is phylogenetically diverse, occurring in many different groups (**Table 1**). Luminescence is unknown in higher plants and in vertebrates above the fishes, and is also absent in several invertebrate phyla. In some phyla or taxa, a substantial proportion of the genera are luminous (for example, ctenophores, about 50%; cephalopods, greater than 50%; echinoderms and annelids, about 4%). Commonly, all members of a luminous genus emit light, but in some cases there are both luminous and nonluminous species.

Bioluminescence is most prevalent in the marine environment; it is greatest at midocean depths, where some daytime illumination penetrates. In these locations, bioluminescence may occur in over 95% of the individuals and 75% of the species in fish, and about 85% of the individuals and 80% of the species in shrimp and squid. A midwater luminous fish (*Cyclothone*) is considered to be the most abundant vertebrate on the planet. Where high densities of luminous organisms occur, their emissions can exert a significant influence on the communities and may represent an important component in the ecology, behavior, and physiology of the latter. Above and below midocean depths, luminescence decreases to less than 10% of all individuals and species; among coastal species, less than 2% are bioluminescent.

TABLE 1. Major groups having luminous species

Group	Features of luminous displays
Bacteria	Organisms glow constantly; system is autoinduced
Fungi	Mushrooms and mycelia produce constant dim glow
Dinoflagellates	Flagellated algae flash when disturbed
Coelenterates	Jellyfish, sea pansies, and comb jellies emit flashes
Annelids	Marine worms and earthworms exude luminescence
Mollusks	Squid and clams exude luminous clouds; also have photophores
Crustacea	Shrimp, copepods, ostracods; exude luminescence; also have photophores
Insects	Fireflies (beetles) emit flashes; flies (diptera) glow
Echinoderms	Brittle stars emit trains of rapid flashes
Fish	Many bony and cartilaginous fish are luminous; some use symbiotic bacteria; others are self-luminous; some have photophores

Although luminescence is prevalent in the deep sea, it is not associated especially with organisms that live in total darkness. There are no known luminous species either in deep fresh-water bodies, such as Lake Baikal, Russia, or in the darkness of terrestrial caves. Luminous dipteran insect larvae live in caves near the mouths of New Zealand, but they also occur in the undercut banks of streams, where there is considerable daytime illumination. Firefly displays of bioluminescence are among the most spectacular, but bioluminescence is rare in the terrestrial environment. Other terrestrial luminous forms include millipedes, centipedes, earthworms, and snails, but the display in these is not very bright.

Functions. The functional importance of bioluminescence and its selection in evolution are based largely on detection by another organism; the responses of that organism then favor in some way the luminous individual. In some organisms, such as the fungi, it is difficult to determine what the function of the light emission is. Both the mycelium and fruiting body emit a dim continuous light day and night, but its function is not evident.

While not metabolically essential, light emission can confer an advantage on the organism. The light can be used in diverse ways. Most of the perceived functions of bioluminescence fall into four categories: defense, offense, communication, and dispersal to enhance propagation (**Table 2**).

Defense. Several different defensive strategies are associated with bioluminescence: to frighten, to serve as a decoy, to provide camouflage, and to aid in

TABLE 2. Functions of bioluminescence

Function	Strategy	Method and examples
Defense	Frighten or startle	Bright, brief flashes
	Decoy or divert	Glow, luminous cloud, sacrificial lure
	Camouflage	Ventral luminescence during the day, disrupting or concealing the silhouette seem from below
Offense	Frighten or startle	Bright flash may temporarily immobilize prey
	Lure	Glow to attract, then capture prey
	Enhance vision	See and capture prey
Communication	Signals for courtship and mating	Specific flashing signals; patterns of light emission are recognized by the opposite sex
Dispersal, propagation	Glow to attract feeders	Bacteria ingested by the feeder pass through the gut tract alive and are thereby dispersed

vision. Organisms may be frightened or diverted by flashes, which are typically bright and brief (0.1 s); light is emitted in this way by many organisms. A glowing object in the ocean often appears to attract feeders or predators. While a luminous organism would evidently be at risk by virtue of this attraction, an organism may create a decoy of exuded light to attract or confuse the predator, and then may slip off under the cover of darkness. This is done by luminous squid, which in the dark squirt a luminescent substance instead of ink, which would be useless in such a case. Some organisms sacrifice more than light: in scaleworms and brittle stars a part of the body may be shed and left behind as a luminescent decoy to attract the predator. In these cases the body part flashes while still intact but glows after detachment—evidence that a flash deters while a glow attracts.

A unique method for evading predation from below is to camouflage the silhouette by emitting light that matches the color and intensity of the downwelling background light. This has been called counterillumination, and it is used by many luminous marine organisms, including fish, to help escape detection.

Offense. Luminescence can aid in predation in several ways. Flashes, which are typically used defensively, can be used offensively in order to temporarily startle or blind prey. A glow can also be used offensively: it can serve as a lure, as in the angler fish. The prey organism is attracted to the light but is then captured by the organism that produced the light. Camouflage may also be used offensively, allowing the luminous predator to approach its prey undetected. Vision is certainly useful offensively: prey may be seen and captured under conditions that are otherwise dark.

Communication. The use of light for information exchange between individuals of a species occurs in several different ways. Best known is the use of flashing for courtship and mating in fireflies. There are also numerous examples in the ocean, such as the Bermuda fireworm, where mating occurs just posttwilight a few days after the full moon. The females swim in a tight circle while a male streaks from below and joins a female, with eggs and sperm being shed in the ocean in a brilliant luminous circle. Among ostracod crustaceans over shallow reefs in the Caribbean, males produce complex species-specific trains of secreted luminous material—"ladders of light"—which attract females.

Propagation. A continuous light emission serves as an attractant for feeding organisms, and saprophytic luminous bacteria typically grow on suitable substrates. After ingestion, the bacteria propagate as they pass through the gut tract and are then dispersed.

Physical and chemical mechanisms. Bioluminescence does not come from or depend on light absorbed by the organism. It derives from an enzymatically catalyzed chemiluminescence, a reaction in which the energy released is transformed into light energy. One of the reaction intermediates or products is formed in an electronically excited state, which then emits a photon. *See* CHEMILUMINESCENCE.

Chemiluminescence. Light emission due to a chemical reaction is a special case of the general phenomenon of luminescence, in which energy is specifically channeled to a molecule. Other kinds of luminescence include fluorescence and phosphorescence, in which the excited state is created by the prior absorption of light, or triboluminescence and piezoluminescence, involving crystal fracture and electric discharge, respectively. Electron bombardment, as in a television screen, also produces luminescence. The color is a characteristic of the excited molecule, irrespective of how the molecule was excited. *See* LUMINESCENCE.

Incandescence. In incandescence, excited states are produced by virtue of the thermal energy. An example is the light bulb, in which a filament is heated, and the color of the light depends on the temperature ("red hot" reflecting a lower temperature than "white hot"). *See* INCANDESCENCE; LIGHT.

Energy. The energy (E) of the photon is related to the color or frequency of the light: $E = h\nu$, where h is Planck's constant and ν the frequency. In the visible light range, E is very large in relation to most biochemical reactions. Thus, the energy released by a mole of photons (6.02×10^{23}) in visible wavelengths is about 50 kcal, which is much more than the energy from the hydrolysis of a mole of adenosine triphosphate (ATP), about 7 kcal. A visible photon is therefore able to do a lot of work (photosynthesis) or a lot of damage (mutation; photodynamic action, which can kill). Conversely, it takes a reaction which releases a large amount of energy to create a photon.

Mechanism. Chemiluminescence in solution generally requires oxygen which, in its reaction with a substrate, forms an organic peroxide. The energy from the breakdown of such peroxides, which should generate up to 100 kcal per mole, is ample to account for a product in an electronically excited state, and such reactions are typically responsible for bioluminescence.

A model of the mechanism in such chemiluminescent reactions is referred to as chemically initiated electron exchange luminescence (CIEEL). In this mechanism, peroxide breakdown involves electron transfer with chemiexcitation. It is initiated by an electron transfer from a donor species (D) to an acceptor (A), which is the peroxide bond in this case. After electron transfer the weak O-O bond cleaves to form products B and C^-. The latter is a stronger reductant than A^-, so the electron is transferred back to D with the concomitant formation of a singlet excited state D^* and emission.

It is useful to think of chemiluminescence as the reverse of a photochemical reaction. In the primary step of photosynthesis, for example, the energy of the excited state of chlorophyll (chl^*) gives rise to an electron transfer, with the consequent formation of an oxidant and reductant. Most of these redox species give rise to stable products and ultimately to CO_2 fixation. However, some species recombine and reemit a "delayed light." This is essentially the reverse of the reaction, with the formation of the singlet excited state of chlorophyll and its subsequent emission of a photon. *See* PHOTOSYNTHESIS.

Biology and biochemistry. Bioluminescence originated and evolved independently many times, and is thus not an evolutionarily conserved function. How many times this may have occurred is difficult to say, but it has been estimated that present-day luminous organisms come from as many as 30 different evolutionarily distinct origins. In the different groups of organisms, the genes and proteins involved are unrelated, and it may be confusing that the substrates and enzymes, though chemically different, are all referred to as luciferin and luciferase, respectively. To be correct and specific, each should be identified with the organism (**Table 3**). These differences are indicative of the independent evolutionary origins of the different systems.

Bacteria. Luminous bacteria typically emit a continuous light, usually blue-green. When strongly expressed, a single bacterium may emit 10^4 or 10^5 photons per second. A primary habitat where most species abound is in association with another (higher) organism, dead or alive, where growth and propagation occur. Luminous bacteria are ubiquitous in the oceans and can be isolated from most seawater samples, from the surface to depths of 1000 m (3280 ft) or more. However, they do not grow in seawater, as it is a poor growth medium; they occur there by spilling over from primary habitats.

Associations and symbiosis. The most exotic specific associations involve specialized light organs (for example, in fish and squid) in which a pure dense culture of luminous bacteria is maintained. In teleost fishes, 11 different groups carrying such bacteria are known, an exotic example being the flashlight fish. Exactly how symbioses are achieved—the initial infection, exclusion of contaminants, nutrient supply, restriction of growth but maintenance of bright light emission—is not understood. In such associations, the host receives the benefit of the light and may use it for one or more specific purposes; the bacteria in return receive a niche and nutrients. However, many fish are self-luminous, not dependent on bacteria.

Direct (nonspecific) luminous bacterial associations include parasitizations and commensals. Intestinal bacteria in marine animals, notably fish, are often luminous, and heavy pigmentation of the gastrointestinal tract is sometimes present, presumably to prevent the light from revealing to predators the location of the fish. Indeed, the light emission of fecal material, which can be quite bright, is thought to benefit the bacteria more directly, for example by attracting feeders and thereby promoting bacterial dispersion and propagation.

Terrestrial luminous bacteria are rare. The best-known are those harbored by nematodes parasitic on insects such as caterpillars. The nematode carries the bacteria as symbionts and injects them into the host along with its own fertilized eggs. The bacteria

TABLE 3. Luciferases and related proteins from four groups

Group	Molecular weight, kDa	GenBank™
Bacteria		
Vibrio fischeri luciferase	~80 (α, 41; β, 39)	VIBLUXAB (*Vibrio fischeri*)
Vibrio fischeri YFP	~28	VIBLUXY (yellow fluorescent protein)
Photobacterium phosphoreum LUMP	~28	PHRLUMP (LUMP proteins)
Dinoflagellates		
Gonyaulax luciferase	~135	GONLUCA (*Gonyaulax polyhedra*)
Gonyaulax LBP	~75	GONLUCIFER (luciferin-binding protein)
Coelenterates		
Renilla luciferase	~25	RELLUC (*Renilla reniformis*)
Aequorea luciferase	~25	AEVAEQA (*Aequorea victoria*)
Aequorea GFP	~21	AEVGFPA (green fluorescent protein)
Firefly		
Photinus pyralis luciferase	~60	PPYLUC (*Photinus pyralis*)

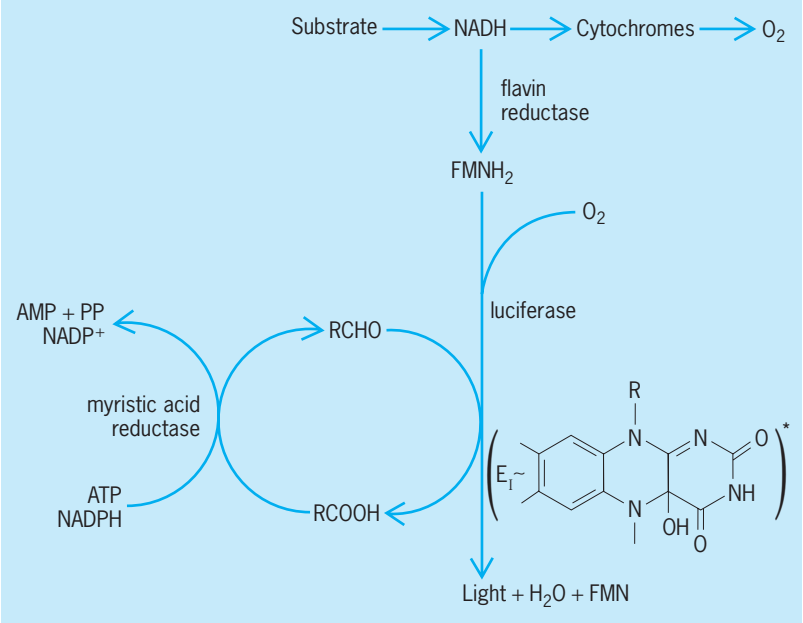

Fig. 1. Pathways and intermediates in the bacterial luciferase reaction. Luciferase reacts with reduced flavin from the electron transport pathway using molecular oxygen to form a flavin-peroxy intermediate. In the subsequent mixed function, oxidation with long-chain aldehyde hydroxy-FMN is formed in the excited state, and light is thereby emitted. The acid and FMN products are recycled.

grow, and the developing nematode larvae feed on them. The dead but now luminous caterpillar attracts predators, which serves to disperse the nematode offspring, along with the bacteria.

Biochemistry. The bacterial system is biochemically unique and is diagnostic for bacterial involvement in the luminescence of a higher organism, such as endosymbionts. The pathway itself (**Fig. 1**) constitutes a shunt of cellular electron transport at the level of flavin, and reduced flavin mononucleotide is the substrate (luciferin) that reacts with oxygen in the presence of bacterial luciferase to produce an intermediate peroxy flavin. This intermediate then reacts with a long-chain aldehyde (tetradecanal) to form the acid and the luciferase-bound hydroxy flavin in its excited state. The light-emitting steps have been modeled in terms of an electron exchange mechanism, and the quantum yield is estimated to be 20–30%. Although there are two substrates in this case, the flavin is the one called luciferin, since it forms, or bears, the emitter. There are enzyme systems that serve to maintain the supply of tetradecanal, and the oxidized flavin is likewise recycled (Fig. 1).

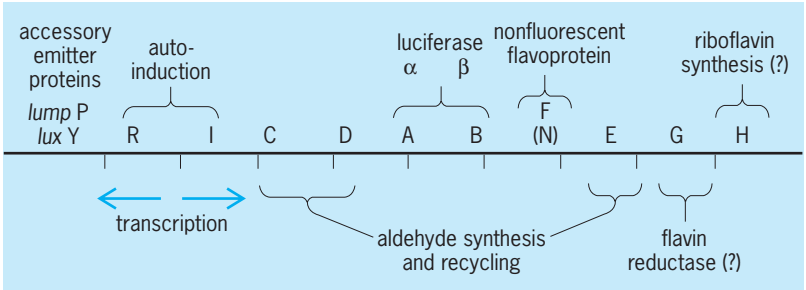

Fig. 2. Organization of the *lux* genes in *Vibrio fischeri*.

Antenna proteins. The biochemistry of the reaction is the same in all luminous bacteria, which typically emit light peaking at about 490 nanometers. However, in some bacterial strains the color of the light emitted by the living cell is blue- or red-shifted, even though the isolated luciferases still peak at about 490 nm. In both cases a second (antenna) protein with its own chromophore is responsible—lumazine and flavin, respectively (Table 3). A functional importance for such spectral shifts has not yet been elucidated, although strains with a blue-shifted emission occur at depths of approximately 600 m (1970 ft) in the ocean.

Luciferases. The proteins themselves are homologous heterodimeric (α-β) proteins (80 kilodaltons) in all bacterial species. They possess a single active center per dimer, mostly associated with the α subunit. Structurally, they appear to be relatively simple: no metals, disulfide bonds, prosthetic groups, or non-amino acid residues are involved. The three-dimensional x-ray crystal structure was reported in 1995, but without the flavin substrate, so the details of the active center are not yet known. However, a deep pocket on the α subunit extending to the β subunit may be the catalytic site. The two subunits fold in a very similar manner, assuming a single-domain eight-stranded motif with alternating α helices and β sheets.

The mechanism of the bacterial luciferase reaction has been studied in great detail. An interesting feature is its inherent slowness: at 20°C (68°F) about 20 s is required for a single catalytic cycle. The luciferase peroxy flavin itself has a long lifetime; at low temperatures [0° to −20°C (32° to −2°F)] it has been isolated, purified, and characterized. It can be further stabilized by aldehyde analogs such as long-chain alcohols and amines, which bind at the aldehyde site.

Genes. Lux genes cloned from several different bacterial species exhibit sequence similarities indicative of evolutionary relatedness and conservation. In *Vibrio fischeri*, the species most extensively studied, the *lux* operon has five structural and two regulatory genes (**Fig. 2**) with established functions. These include *lux* A and B that code for the α and β subunits of luciferase, and *lux* C, D, and E that code for the reductase, transferase, and synthetase components of the fatty acid reductase complex. Upstream, in the same operon, is the regulatory gene *lux* I, and immediately adjacent but transcribed in the opposite direction is *lux* R, whose product is responsible for the transcriptional activation of *lux* A through E. Other genes are also subject to this regulation, including those coding for the antenna proteins responsible for color shifting.

Regulation. Luminous bacteria are unable to regulate emission on a fast time scale (milliseconds, seconds), as in organisms that emit flashes. However, they typically control the development and expression of luminescence at both physiological and genetic levels. Physiologically, luciferase and aldehyde synthesis genes of the *lux* operon are regulated by a substance produced by the cells called autoinducer (a homoserine lactone) such that these genes are

transcribed only after the substance has accumulated at a high cell density. Autoinducer is a product of the *lux* I gene and stimulates *lux* R. The mechanism has been dubbed quorum sensing, and similar mechanisms occur in other types of bacteria with different genes. The ecological implications are evident: in planktonic bacteria, a habitat where luminescence has no value, autoinducer cannot accumulate, and no luciferase synthesis occurs. However, in the confines of a light organ, high autoinducer levels are reached, and the luciferase genes are transcribed.

Transcription of the lux operon is also regulated by glucose, iron, and oxygen, each of which has ecological implications. But there is also control at the genetic level: in some species of bacteria, dark (very dim) variants arise spontaneously, and the synthesis of the luminescent system does not occur. However, the *lux* genes are not lost and revertants occur, so that under appropriate conditions, where luminescence is advantageous, luminous forms will be available to populate the appropriate habitat.

Dinoflagellates. Dinoflagellates occur ubiquitously in the oceans as planktonic forms in surface waters, and contribute substantially to the luminescent flashing commonly seen at night (especially in summer) when the water is disturbed. In phosphorescent bays (for example, in Puerto Rico and Jamaica), high densities of a single species (*Pyrodinium bahamense*) occur; this is also true in red tides, which are blooms of dinoflagellates. Many species are photosynthetic.

All luminous dinoflagellates are marine, but only about 30% of the marine species are luminous. As a group, dinoflagellates are important as symbionts, notably for contributing photosynthesis and carbon fixation in animals; but unlike bacteria, no luminous dinoflagellates are known from symbiotic niches.

Dinoflagellates are stimulated to emit light when predators (such as crustaceans) are feeding, which can startle and divert the predator, resulting in a reduced predation. The response time to stimulation (milliseconds) is certainly fast enough to have this effect. Predation on dinoflagellates may also be impeded more indirectly, since predators of the crustaceans might be alerted, thereby increasing predation on the crustaceans themselves.

Cell biology. Luminescence in dinoflagellates is emitted from many small (about 0.4 micrometer) cortical locations. The structures have been identified in *Gonyaulax polyhedra* as novel organelles, termed scintillons (flashing units). They occur as outpocketings of the cytoplasm into the cell vacuole, like a balloon, with the neck remaining connected. However, scintillons contain only dinoflagellate luciferase and luciferin (with its binding protein), other cytoplasmic components somehow being excluded. Ultrastructurally, they can be identified by immunolabeling with antibodies raised against the luminescence, proteins, and visualized by their bioluminescent flashing following stimulation, as well as by the fluorescence of luciferin.

Biochemistry. Dinoflagellate luciferin is a novel tetrapyrrole related to chlorophyll (**Fig. 3**). The luciferase is an unusual 135-kDa protein with three

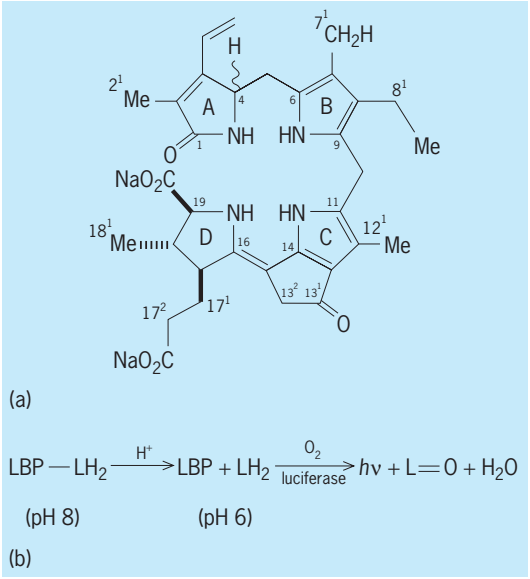

Fig. 3. Dinoflagellate luciferin (LH_2). (*a*) Structure of this tetrapyrrole. Oxygen addition occurs at position 13^2. (*b*) Schematic of reaction pathway. Luciferin, bound to the luciferin-binding protein (LBP) at pH 8, is released at pH 6 and is oxidized by luciferase to form an exicited state emitter (L=O*).

conserved active sites within the single molecule. The reaction involves the luciferase-catalyzed oxidation of luciferin triggered by a change of the pH from 8 to 6, which releases the luciferin from its binding protein and activates the luciferase. The flashing of dinoflagellates in vivo is postulated to result from a rapid and transient pH change in the scintillons, triggered by an action potential in the vacuolar membrane which, while sweeping over the scintillon, opens ion channels that allow protons from the acidic vacuole to enter.

Circadian clock control. Cell density and the composition of the medium have no effect on the development and expression of luminescence in dinoflagellates. However, in *G. polyhedra* and some other dinoflagellates, luminescence is regulated by day-night light-dark cycles and an internal circadian biological clock mechanism. The flashing response to mechanical stimulation is far greater during the night than during the day phase, and a dim glow exhibits a rhythm (**Fig. 4**). The regulation is due to an

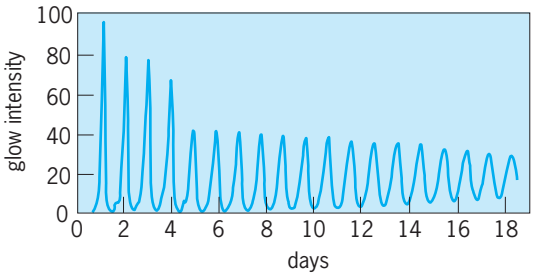

Fig. 4. Circadian rhythm of the steady glow of bioluminescence in *Gonyaulax*. The culture, grown in a 24-h light-dark cycle, was transferred to constant light on the first day. Measurements of the spontaneous light emission were made about once every hour for the subsequent 18 days. The period of the rhythm is about 25 h.

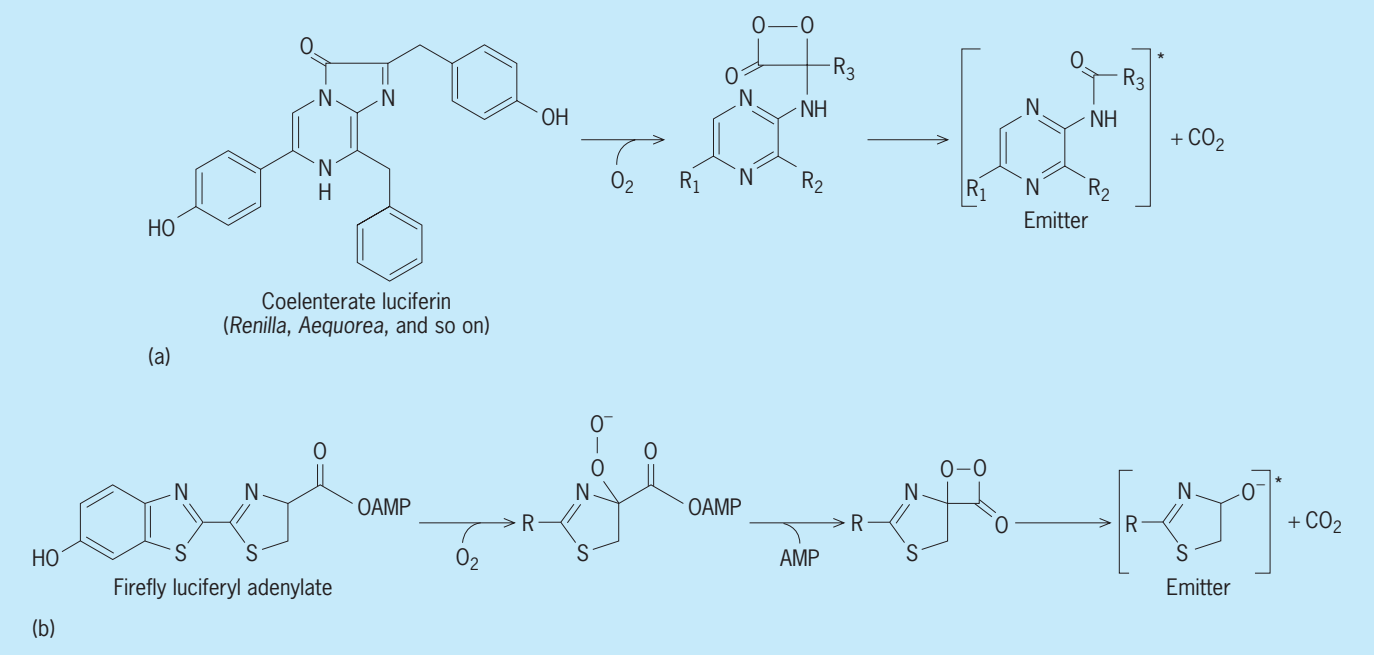

Fig. 5. Luciferin structures. (*a*) Coelenterate luciferin (coelenterazine) and its reaction with oxygen to give a postulated cyclic peroxide intermediate and then an excited emitter in the light-emitting reaction. (*b*) Firefly luciferin and the analogous intermediates in the firefly luciferase reaction.

endogenous mechanism; cultures maintained under constant conditions continue to exhibit rhythmicity for weeks, with a period that is not exactly 24 h but about one day (*circa diem*).

The nature of the circadian clock remains an enigma. In humans and other higher animals, where it regulates the sleep-wake cycle and many other physiological processes, the mechanism involves the nervous system. But it also occurs in plants and unicellular organisms, which lack neural control. In the case of *G. polyhedra*, daily changes occur in the cellular concentrations of both luciferase and luciferin and its binding protein: the proteins are synthesized and destroyed each day. Hence, the biological clock exerts control at a very basic level, by controlling gene expression.

Coelentrates and ctenophores. Luminescence is common and widely distributed in these groups. In the ctenophores (comb jellies), luminous forms make up over half of all genera, and in the coelenterates (cnidarians) it is about 6%. Luminous hydroids, siphonophores, sea pens, and jellyfish are mostly sessile or sedentary and, upon stimulation, emit flashes. Sea anemones and corals are not known to be bioluminescent.

Hydroids occur as plantlike growths, typically adhering to rocks below low tide level in the ocean. Upon touching them, there is a sparkling emission conducted along the colony; repetitive waves from the origin may occur. In luminous jellyfish (such as *Pelagia noctiluca* and *Aequorea victoria*) the bright flashing comes from photocytes along the edge of the umbrella at the base of the tentacles. *Aequorea* is the source of a photoprotein that emits light upon the addition of calcium and is widely used in research for calcium detection and measurement. The sea pansy, *Renilla*, which occurs near shore on sandy bottoms, has also figured prominently in studies of bioluminescence.

Cell biology. Photocytes occur as specialized cells located singly or in clusters in the endoderm. They are commonly controlled by epithelial conduction in hydropolyps and siphonophores and by a colonial nerve net in anthozoans. The putative neurotransmitter involved in luminescent control in *Renilla* is adrenaline or a related catecholamine.

Biochemistry. Coelenterate luciferase has a similarity to calmodulin in its amino acid sequence. The luciferin, coelenterazine (**Fig. 5***a*), possesses an imidazopyrazine skeleton. It has widespread phylogenetic distribution in nonluminous species, but whether the reason is nutritional or genetic (hence, possible evolutionary relatedness) has not yet been elucidated. In some cases (such as *Renilla*), the sulfated form of luciferin may occur as a precursor or storage form, and is convertible to active luciferin by sulfate removal with the cofactor 3′5′-diphosphadenosine. The active form may also be sequestered by a calcium-sensitive binding protein, analogous to the dinoflagellate binding protein. In this case, calcium triggers the release of luciferin and flashing.

Another type of control of the same luciferin in other cnidarians ultilizes a luciferase-peroxy luciferin intermediate poised for the completion of the reaction. The photoprotein aequorin, isolated from the jellyfish *Aequorea* (in the presence of EDTA to chelate calcium), emits light simply upon the addition of calcium, which is presumably the trigger in vivo. The luciferin and luciferase react with oxygen to form a peroxide in a calcium-free compartment (the photocyte), where it is stored. An action

potential allows calcium to enter and bind to the protein, changing its conformational state and allowing the reaction to continue, but without the need for free oxygen at this stage. An enzyme-bound cyclic peroxide, a dioxetenone (Fig. 5*a*), is a postulated intermediate; it breaks down with the formation of an excited emitter, along with a molecule of CO_2. Early literature reported that ctenophores could emit bioluminescence without oxygen. The explanation is now evident: the animal contains the luciferase-bound peroxy-luciferin in a stored and stable state, and only calcium is needed to emit light.

Antenna proteins and GFPs. In several coelenterates the light emission in vivo occurs at a longer wavelength than in vitro, similar to that described above for bacteria. In this case the different protein with a second chromophore is called green fluorescent protein (GFP), widely used as a reporter in molecular biology. Radiationless energy transfer is believed to be involved, with emission from a chromophore whose structure is unusual; it is part of the primary structure of the protein, three residues of which are chemically modified posttranslationally to produce a fluorescence with a peak emission at about 408 nm. Different groups appear to have different GFP proteins, but the chromophores are the same.

Fireflies. Of the approximately 70,000 insect genera, only about 100 are classed as luminous. But their luminescence is impressive, especially in the fireflies and their relatives. Fireflies possess ventral light organs on posterior segments; the South American railroad worm, *Phrixothrix*, has paired green lights on the abdominal segments and red head lights; while the click and fire beetles, Pyrophorini, have both running lights (dorsal) and landing lights (ventral). The dipteran cave glow worm, in a different group and probably different biochemically, exudes beaded strings of slime from its ceiling perch, serving to entrap minute flying prey, which are attracted by the light emitted by the animal.

Function. The major function of light emission in fireflies is for communication during courtship, typically involving the emission of a flash by one sex as a signal, to which the other sex responds, usually in a species-specific pattern. The time delay between the two may be a signaling feature; for example, it is precisely 2 s in some North America species. But the flashing pattern (such as trains distinctive in duration and intensity) is also important in some cases, as is the kinetic character of the individual flash (duration; onset and decay kinetics).

Fireflies in Southeast Asia are noteworthy for their synchronous flashing. Congregations of many thousands form in trees, where the males produce an all-night-long display, with flashes every 1–4 s, dependent on species. This may serve to attract females to the tree.

Biochemistry. The firefly system was the first in which the biochemistry was characterized. It had been known before 1900 that cold water extracts could continue to emit light for several minutes or hours, and that after the complete decay of the light, emission could be restored by adding a second extract, prepared by using boiling water to extract the cells, then cooling the liquid. The enzyme luciferase was assumed to be in the cold water extract (with all the luciferin substrate being used up during the emission), whereas the enzyme was denatured by the hot water extraction, leaving some substrate intact. This was referred to as the luciferin-luciferase reaction; it was already known in the first part of the twentieth century that luciferins and luciferases from the different major groups would not cross-react, indicative of their independent evolutionary origins.

In 1947 it was discovered that bioluminescence was restored in an exhausted cold water extract simply by the addition of adenosine triphosphate (ATP). But ATP could not be the emitter, since it does not have the appropriate fluorescence, and could not provide the energy for light emission, since the energy available from ATP hydrolysis is only about 7 kcal per mole, whereas the energy of a visible photon is 50 kcal per mole. It was concluded that luciferin had not actually been used up in the cold water extract, and it was soon shown that ATP functions to form a luciferyl adenylate intermediate. This then reacts with oxygen to form a cyclic luciferyl peroxy species, which breaks down to yield CO_2 and an excited state of the carbonyl product (Fig. 5*b*).

The crystal structure of firefly luciferase has been determined recently, but without substrate, so the binding site is not established. The gene has been cloned and expressed in other organisms, including *Escherichia coli* and tobacco. In both cases, luciferin must be added exongenously; tobacco "lights up" when the roots are dipped in luciferin. Luciferase catalyzes both the reaction of luciferin with ATP and the subsequent steps leading to the excited product. There are some beetles in which the light from different organs is a different color, but no second emitter protein appears to be utilized. Instead, the same ATP-dependent luciferase reaction with the same luciferin occurs in the different organs, but the luciferases are slightly different, coded by different (but homologous) genes. They are presumed to differ with regard to the structure of the site that binds the excited state, which could thereby alter the emission wavelength.

Cell biology. The firefly light organ comprises a series of photocytes arranged in a rosette, positioned radially around a central tracheole, which supplies oxygen to the organ. Organelles containing luciferase have been identified with peroxisomes on the basis of immunochemical labeling.

Regulation of flashing. Although flashing is initiated by a nerve impulse that travels to the light organ, photocytes are not stimulated directly. The nerves in the light organ terminate on the tracheolar end cells, accounting for the considerable time delay between the arrival of the nerve impulse at the organ and the onset of the flash. The flash might be controlled by the availability of oxygen.

Mollusks. Snails (gastropods), clams (bivalves), and squid (cephalopods) have bioluminescent members. Squid luminous systems are numerous and diverse in both form and function, rivaling the fishes in

these respects. As is also true for fishes, some squid utilize symbiotic luminous bacteria, but most are self-luminous, indicating that bioluminescence had more than one evolutionary origin even within the class.

Compound photophores have associated optical elements such as pigment screens, chromatophores, reflectors, lenses, and light guides. Many squid possess such photophores, which may be used in spawning and other interspecific displays (communication). They may emit different colors of light, and are variously located near the eyeball, on tentacles, or on the body integument, or associated with the ink sac or other viscera. In some species, luminescence intensity has been shown to be regulated in response to changes in ambient light, indicative of a camouflage function.

Along the coasts of Europe a clam, *Pholas dactylus*, inhabits compartments that it makes by boring holes into the soft rock. When irritated, these animals produce a bright cellular luminous secretion, squirted out through the siphon as a blue cloud. This luminescent animal has been known since Roman times, and the system was used in the discovery and description of the luciferin-luciferase reaction in the 1880s. The reaction involves a protein-bound chromophore, and the luciferase is a copper-containing large (greater than 300 kDa) glycoprotein. Well ensconced in its rocky enclosure, the animal presumably used the luminescence to somehow deter or thwart would-be predators.

The New Zealand pulmonate limpet, *Latia neritoides*, is the only known luminous organism that can be classed as a truly fresh–water species. It secretes a bright luminous slime (green emission, λ_{max} = 535 nm), whose function may be similar to that of the *Pholas* emission. Its luciferin is an enol formate of an aldehyde, and a "purple protein" is also required, but only in catalytic quantities, suggesting that it may be somehow involved as a recycling emitter.

Annelids. *Chaetopterus* is a marine polychaete that constructs and lives in U-shaped tubes in sandy bottoms and exudes luminescence upon stimulation to deter or thwart predators. Other annelids include the Syllidae, such as the Bermuda fireworm, and the polynoid worms, which shed their luminous scales as decoys. Extracts of the latter have been shown to emit light upon the addition of superoxide ion.

Terrestrial earthworms, some of which are quite large, over 24 in. (60 cm) in length, exude coelomic fluid from the mouth, anus, and body pores upon stimulation. This exudate contains cells that lyse to produce a luminous mucus, emitting in the blue-green region. In *Diplocardia longa* the cells responsible for emission have been isolated; luminescence in extracts involves a copper-containing luciferase (about 300 kDa), and *N*-isovaleryl-3 amino-1 propanal. The in vitro reaction requires hydrogen peroxide, not free oxygen. But the exudate from animals deprived of oxygen does not emit, but will do so after the admission of molecular oxygen to the free exudate.

Crustaceans. Cyprindinid ostracods such as *Vargla hilgendorfii* are small organisms that possess two glands with nozzles from which luciferin and luciferase are squirted into the seawater, where they react and produce a spot of light useful either as a decoy or for communication. Cypridinid luciferin is a substituted imidazopyrazine nucleus, similar to coelenterazine, which reacts with oxygen to form an intermediate cyclic peroxide, breaking down to yield carbon dioxide and an excited carbonyl. However, the cypridinid luciferase gene has no homology with the gene for the corresponding coelenterate proteins, and calcium is not involved in the cypridinid reaction. Thus, the two different luciferases reacting with similar luciferins have apparently had independent evolutionary origins, indicative of convergent evolution at the molecular level.

Blue ventrally directed luminescence is emitted by compound photophores in euphausiid shrimp. The system is unusual because both its luciferase and luciferin cross-react with the dinoflagellate system. This cross-taxon similarity indicates a possible exception to the rule that luminescence in distantly related groups had independent evolutionary origins. The shrimp might obtain luciferin nutritionally, but the explanation for the functionally similar proteins is not known, and the shrimp luciferase gene has not been cloned. One possibility is lateral gene transfer; convergent evolution is another.

Fishes. Both teleost (bony) and elasmobranch (cartilaginous) fishes have luminous species. Many coastal and deep-sea fishes culture luminous bacteria in special organs, but most are self-luminous. These include *Porichthys*, the midshipman fish, having linear arrays of photophores distributed along the lateral lines, as with buttons on a military uniform. Its luciferin and luciferase cross-react with the cypridinid ostracod crustacean system described above, and the fish apparently obtain luciferin nutritionally.

Open-sea and midwater species include sharks (elasmobranchs), which may have several thousand photophores; the gonostomatids, such as *Cyclothone*, with simple photophores; and the hatchet fishes, having compound photophores with elaborate optical accessories to direct luminescence downward, indicative of a camouflage function of the light.

Different kinds of photophores located on different parts of the fish evidently have different functions. One arrangement, known in both midwater squid and myctophids, makes use of a special photophore positioned so as to shine on the eye or a special photoreceptor. Its intensity parallels that of the other photophores, so it provides information to the animal concerning the organism's brightness, thus allowing the animal to match the intensity of its counterillumination to that of the downwelling ambient light. A different functional use is in *Neoscopelus*, where photophores on the tongue attract prey to just the right location.

Luminescence in many fish is sexually dimorphic, and the use of the light differs in male and female. Some self-luminous fish eject luminous material; in the searsid fishes this material is cellular in nature, but it is not bacterial, and its biochemical nature is

not known. Such fish may also possess photophores.

Applications. The measurement of light emission is typically more sensitive than many other analytical techniques. A typical photon-counting instrument can readily detect an emission of about 10^4 photons per second, which corresponds to the transformation of 10^5 molecules per second (or 6×10^6 per minute) if the quantum yield is 10%. A substance at a concentration of 10^{-9} *M* could readily be detected in a single cell 10 μm in diameter by this technique.

Bioluminescence and chemiluminescence have thus come into widespread use for quantitative determinations of specific substances in biology and medicine. Luminescent tags have been developed that are as sensitive as radioactivity, and now replace radioactivity in many assays.

Analytical measurements. The biochemistry of different luciferase systems is different, so many different substances can be detected. One of the first, and still widely used, assays involves the use of firefly luciferase for the detection of ATP. Since many different enzymes use or produce ATP, their activities may be followed using this assay. With bacterial luciferase, any reaction that produces or utilizes NAD(H), NADP(H), or long-chain aldehyde either directly or indirectly, can be coupled to this lightemitting reaction. The purified photoprotein aequorin has been widely used for the detection of intracellular calcium and its changes under various experimental conditions; the protein is relatively small (about 20 kDa), is nontoxic, and may be readily injected into cells in quantities adequate to detect calcium over the range of 3×10^{-7} to 1×10^{-4} *M*.

The amount of oxygen required for bioluminescence in luminescent bacteria is small and therefore the reaction readily occurs. Luminous bacteria can be used as a very sensitive test for oxygen, sometimes in situations where no other method is applicable. An oxygen electrode incorporating luminous bacteria has been developed.

Reporter genes. Luciferase and antenna proteins have also been exploited as reporter genes for many different purposes. Analytically, such systems are virtually unique in that they are noninvasive and nondestructive: the relevant activity can be measured as light emission in the intact cell and in the same cell over the course of time. Examples of the use of luciferase genes are the expression of firely and bacterial luciferases under the control of circadian promoters; and the use of coelenterate luciferase expressed transgenically to monitor calcium changes in living cells over time.

Green fluorescent protein is widely used as a reporter gene for monitoring the expression of some other gene under study, and for how the expression may differ, for example at different stages of development or as the consequence of some experimental procedure. As with the luciferases, this is done by placing the GFP gene under the control of the promoter of the gene being investigated.

Studies of bioluminescence have thus contributed to knowledge in ways not foreseen at the time of the investigations, illustrating how basic knowledge can be a powerful tool for advancing understanding in unrelated areas.

J. Woodland Hastings

Bibliography. T. Baldwin et al., Structure of bacterial luciferase, *Curr. Biol.*, 5:798–809, 1995; M. Chalfie and S. Kain (eds.), *Green Fluorescent Protein: Properties, Applications and Protocols*, Wiley-Liss, New York, 1998; E. Conti, N. P. Franks, and P. Brick, Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes, *Structure*, 4:287–298, 1996; J. W. Hastings, Chemistries and colors of bioluminescent reactions: A review, *Gene*, 173:5–11, 1996; J. W. Hastings and J. G. Morin, Bioluminescence, in C. L. Prosser (ed.), *Neural and Integrative and Animal Physiology*, Wiley Interscience, New York, 1991; J. W. Hastings and D. Tu (eds.), Molecular mechanisms in bioluminescence, *Photochem. Photobiol.*, 62:599–673, 1995; J. W. Hastings, L. J. Kricka, and P. E. Stanley (eds.), *Bioluminescence and Chemiluminescence: Molecular Reporting with Photons*, Wiley, Chichester, 1997; P. J. Herring (ed.), *Bioluminescence in Action*, Academic, London, 1978; M. K. Nishiguchi, E. G. Ruby, and M. J. McFall-Ngai, Competitive dominance among strains of luminous bacteria provides an unusual form of evidence for parallel evolution in sepiolid squid-*Vibria* symbioses, *Appl. Environ. Microbiol.*, 64:3209–3213, 1998; T. Wilson and J. W. Hastings, Bioluminescence, *Annu. Rev. Cell Dev. Biol.*, 14:197–230, 1998.

Biomagnetism

The production of a magnetic field by a living object. The living object presently most studied is the human body, for two purposes: to find new techniques for medical diagnosis, and to gain information about normal physiology. Smaller organisms studied include birds, fishes, and objects as small as bacteria; many scientists believe that biomagnetics is involved in the ability of these creatures to navigate. This article deals only with the human body. *See* MIGRATORY BEHAVIOR.

Magnetic field production. The body produces magnetic fields in two main ways: by electric currents and by ferromagnetic particles. The electric currents are the ion currents generated by the muscles, nerves, and other organs. For example, the same ion current generated by heart muscle, which provides the basis for the electrocardiogram, also produces a magnetic field over the chest; and the same ion current generated by the brain, which provides the basis for the electroencephalogram, also produces a magnetic field over the head. Ferromagnetic particles are insoluble contaminants of the body; the most important of these are the ferromagnetic dust particles in the lungs, which are primarily Fe_3O_4 (magnetite). Magnetic fields can give information about the internal organs not otherwise available.

These magnetic fields are very weak, usually in the range of 10^{-14} to 10^{-9} tesla; for comparison, the Earth's field is about 10^{-4} T (1 T = 10^4 gauss, the older unit of field). The fields at the upper end of

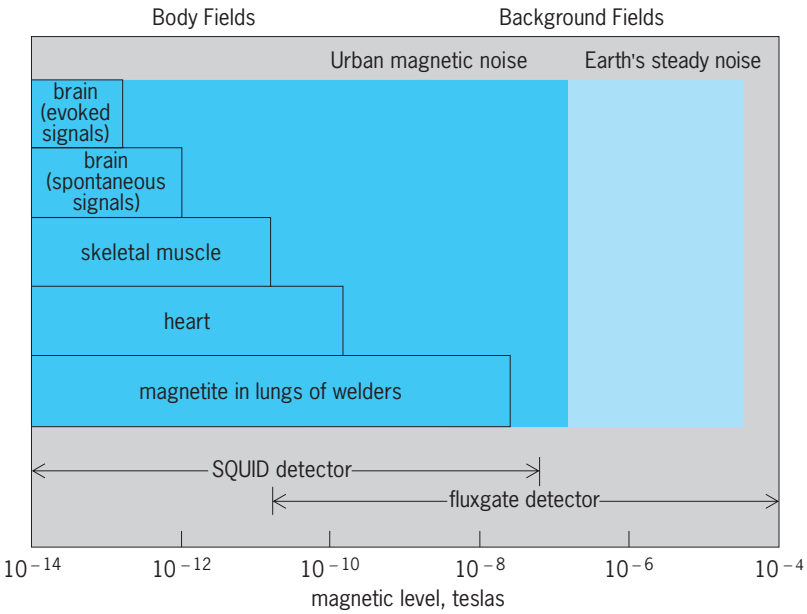

Fig. 1. Levels of the body's magnetic fields and background magnetic fields, as well as the range encompassed by the detectors.

this range, say stronger than 10^{-4} T, can be measured with a simple but sensitive magnetometer called the fluxgate; the weaker fields are measured with the extremely sensitive cryogenic magnetometer called the SQUID (superconducting quantum interference device). The levels of the body's fields are shown in **Fig. 1**. These fields, whether they are fluctuating or steady, are seen to be orders of magnitude weaker than the fluctuating or steady background fields. They can, however, be measured by using either a magnetically shielded room **(Fig. 2)** or two detectors connected in opposition so that much of the background is canceled, or a combination of both methods.

Fig. 2. Magnetic field of the brain being measured in a magnetically shielded room. The SQUID detector is in the tail of a Dewar cylinder containing liquid helium, which maintains the cryogenic temperature of the SQUID.

The organs producing magnetic fields which are of most interest are the brain, the lungs, and the liver. These are discussed below.

Brain measurement. Theory suggests that a measurement of the brain's magnetic field, the magnetoencephalogram (MEG), could reveal new information on the brain that is not available from the electroencephalogram (EEG). In the first place, the MEG sees only the electrical sources in the head which are oriented tangentially to the skull, while the EEG sees both tangential sources and sources oriented radially. The MEG therefore sees a subgroup of the electrical sources in the brain, that is, those tangential sources not seen on the EEG because they are "swamped" by the radial sources, which dominate the EEG. Second, the MEG pattern on the head due to a tangential source is oriented perpendicularly to the EEG pattern on the head, due to the same source **(Fig. 3)**. Because of this difference, the MEG is complementary to the EEG in localizing a source in the brain, such as an epileptic focus. The EEG localizes better in one direction on the scalp (x direction in Fig. 3), while the MEG localizes better in the other direction (y directional). Third, the MEG pattern on the head is somewhat smaller than the EEG pattern (Fig. 3). This is because the skull, through which the current must pass in order to be recorded by the EEG electrodes, diffuses the current, while the MEG sees only the deep currents at the source and therefore is unaffected by the skull; without the skull, the MEG and EEG patterns would be of about equal size (although rotated by 90°). This smaller pattern on the MEG allows it to localize better in its best (y) direction, in comparison to the EEG in its best (x) direction.

These theoretical differences have received experimental confirmation. Further, MEGs are recorded on many of the same types of electrical events seen on the EEG, including evoked responses (where the stimuli, such as periodic clicks to the ear, are controlled) and spontaneous events such as epileptic discharges. In the evoked responses and spontaneous events, new information appears to be seen in the MEG, of use in clinical diagnosis and in research of the normal brain. *See* ELECTROENCEPHALOGRAPHY.

Lung measurement. To measure ferromagnetic dust in the lungs, the dust in a subject's lungs is first magnetized by an externally applied magnetic field. Then, after the external field is removed, the remanent field of magnetized dust is measured around the torso with a fluxgate or a SQUID; this yields the amount of dust in the lungs. These measurements are made for three main reasons.

First, when Fe_3O_4 dust is inhaled through occupational exposure, a magnetic measurement of the amount of dust in the lungs provides a measure of the dust exposure in that worker. For example, a magnetic measurement of the amount of Fe_3O_4 itself is relatively harmless, but other associated dusts could be quite harmful; therefore knowledge of the presence of Fe_3O_4 is useful.

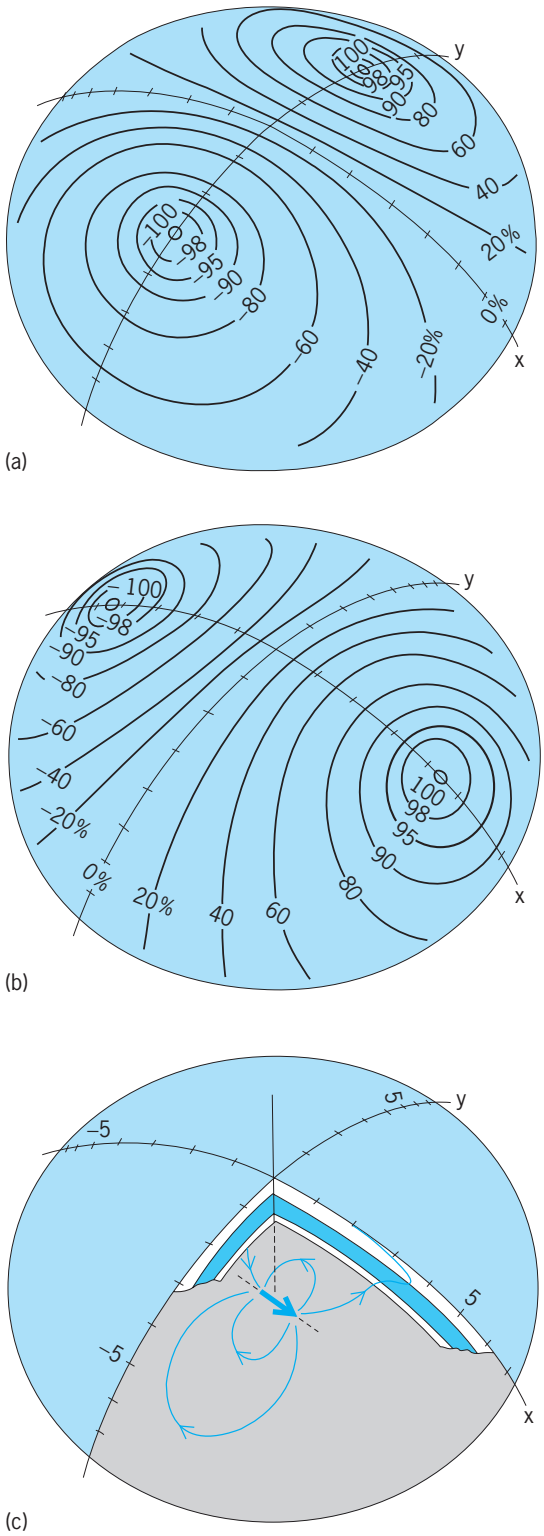

Fig. 3. Theoretical calculations showing two of the three differences between (*a*) the MEG pattern and (*b*) the EEG pattern, due to a tangential source. The theoretical calculation was performed by using a four-layer spherical model of the head. (*c*) A section of the sphere produced by computer modeling, where the electrical source (heavy arrow) produces currents (light arrows) in the conducting medium of the head. The tick marks are in centimeters on the scalp surface. The MEG and EEG patterns are shown in percent of their maximum, on the same spherical cap. The MEG pattern is seen to be rotated by 90° to that of the EEG, and is somewhat smaller.

The second reason involves the clearance of dust from the lungs. Inhaled dust is deposited either in the airways (bronchi) or in the terminal chambers (alveoli). The dust deposited in the airways is known to be cleared out by the mucociliary carpet in a matter of hours; this rate, called short-term clearance, has been accurately measured in the human lung by using short-lived radioactive dust. The dust deposited in the alveoli is cleared out much more slowly, in a matter of weeks or months. This rate, called long-term clearance, had not previously been accurately measured, because of the hazard of using long-lived radioactive dust. However, this rate can now be measured by using only Fe_3O_4, which is nonhazardous. In one study the clearance rate in smokers was compared with that in nonsmokers, and was found to be considerably slower in the smokers. This impaired clearance of carcinogenic dust may contribute to the high incidence of lung cancer in smokers.

The third reason involves the phenomenon called relaxation. This takes place when the external field, which first magnetizes the lung, is removed. The remanent field at the chest due to the particles is not steady, but decreases during the first hour to about 10% of its initial value. This is because the particles, originally aligned by the applied field, receive physiological impulses which randomly rotate them; as they become scrambled in direction, the net remanent field they create decreases in strength. The source of the impulses is now known to be motion within the macrophage cells. Measurement of the rate of relaxation is useful as a new index of internal cell function; that is, the rate of decrease of the remanent field gives information about the internal activity of the cell.

Liver measurement. Unlike the lung or the brain, the magnetic field of the liver is measured via its magnetic susceptibility. This is a third way the body produces a magnetic field. When an external magnetic field is applied to the body, each element of the body produces its own weak magnetic field in response to the applied field. The susceptibility measurement consists in determining the amount and polarity of this weak field, in the presence of the applied field; the weak field can be in the same or in the opposite direction as the applied field, and the element producing the response is called paramagnetic or diamagnetic, respectively. The liver is normally a diamagnetic organ; however, it is the main localized storage site of iron in the body, and its susceptibility depends on the amount of iron it contains within its organic molecules. In the normal liver, a few parts in 10^4 consist of iron, but this concentration may rise a hundredfold in iron-overload disease, or drop by a similar large factor in iron-deficiency disease; the iron concentration changes the liver susceptibility so that it becomes paramagnetic for iron concentrations greater than 6 parts in 10^3, which is within the overload range. While departures from normal are associated with widespread medical problems, no reliable noninvasive method has been found to assess this iron store. Thus an estimation of the iron

concentration from a magnetic susceptibility measurement is of clinical value.

In SQUID measurements in a hospital, where this method is being evaluated, it was found that the iron concentration estimated from this susceptibility was in agreement with the iron concentration determined by a direct chemical analysis of liver tissue obtained in biopsy. The two estimates were in substantial agreement over a range from normal to some 30 times normal; hence, the measurement of the magnetic field of the liver, via its susceptibility, appears to be a promising technique. *See* BIOELECTROMAGNETICS.

David Cohen

Bibliography. D. Cohen, Magnetic fields of the human body, *Phys. Today*, pp. 34–43, August 1975; D. Cohen and B. N. Cuffin, Demonstration of useful differences between MEG and EEG, *Electroencephalography and Clinical Neurophysiology*, 56:38–51, 1983; J. L. Kirschvink, Birds, bees, and magnetism: A new look at the old problem of magnetoreception, *Trends Neurosci.,* 5:160–167, 1982; S. J. Williamson and L. Kaufman, Biomagnetism, *J. Magnetism Magnetic Mater.*, 22:129–201, 1981; S. Willamson et al. (eds.), *Biomagnetism: An Interdisciplinary Approach*, 1983.

Biomass

The organic materials produced by plants, such as leaves, roots, seeds, and stalks. In some cases, microbial and animal metabolic wastes are also considered biomass. The term "biomass" is intended to refer to materials that do not directly go into foods or consumer products but may have alternative industrial uses. Common sources of biomass are (1) agricultural wastes, such as corn stalks, straw, seed hulls, sugarcane leavings, bagasse, nutshells, and manure from cattle, poultry, and hogs; (2) wood materials, such as wood or bark, sawdust, timber slash, and mill scrap; (3) municipal waste, such as waste paper and yard clippings; and (4) energy crops, such as poplars, willows, switchgrass, alfalfa, prairie bluestem, corn (starch), and soybean (oil). *See* BIOLOGICAL PRODUCTIVITY.

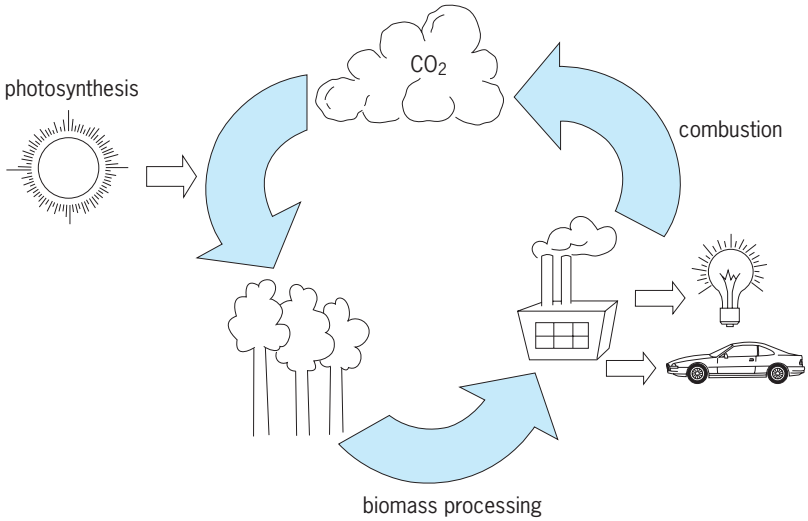

Fig. 1. Carbon cycle, showing conversion of solar energy to fuel and electricity.

Composition. Biomass is a complex mixture of organic materials, such as carbohydrates, fats, and proteins, along with small amounts of minerals, such as sodium, phosphorus, calcium, and iron. The main components of plant biomass are carbohydrates (approximately 75%, dry weight) and lignin (approximately 25%), which can vary with plant type. The carbohydrates are mainly cellulose or hemicellulose fibers, which impart strength to the plant structure, and lignin, which holds the fibers together. Some plants also store starch (another carbohydrate polymer) and fats as sources of energy, mainly in seeds and roots (such as corn, soybeans, and potatoes). Starch and cellulose are homopolymers (long chains of a single component) of glucose. Glucose is the main sugar used by living systems as a source of biochemical energy. Hemicellulose is a structural polymer made of sugars, such as xylose, glucose, arabinose, and mannose. Lignin is a much more complex mixture of organic materials, composed of highly cross-linked phenylpropane units with other organic components attached. Heterogeneous biomass materials, such as hemicellulose and lignin, are more difficult to decompose to make useful chemicals than homogeneous materials, such as starch and cellulose. Fats are triacylglycerides, three long-chain fatty acids (hydrocarbonlike structures) attached to a glycerol molecule. The major part of the fat molecule is fatty acids, which have a chemical structure very similar to hydrocarbons (petroleum). The general chemical structures are hydrocarbons, C_nH_{2n+2}; carbohydrates, $C_nO_nH_{2n}$; and triacylglycerides, $3(C_{n+1}H_{2n+1}O)+C_3H_5O_3$ [fatty acids plus glycerol].

Hydrocarbons are composed of only carbon and hydrogen. Upon combustion with oxygen, carbon dioxide and water are formed and thermal energy is released. The more oxygen reacted, the more energy released. Since carbohydrates already contain oxygen, their combustion releases less thermal energy than hydrocarbons. With fats, the fatty acids are usually separated from glycerol, then processed to make fuels and chemicals. Fatty acids contain nearly the same amount of energy as hydrocarbons. *See* CELLULOSE; LIGNIN.

Renewability. A major advantage of using biomass as a source of fuels or chemicals is its renewability. Utilizing sunlight energy in photosynthesis, plants metabolize atmospheric carbon dioxide to synthesize biomass. An estimated 140 billion metric tons of biomass are produced annually.

Combustion of biomass forms carbon dioxide, and plants metabolize carbon dioxide to form more biomass via photosynthesis. The overall process is an ecological cycle called the carbon cycle, since elemental carbon is recycled. From an energy viewpoint, the net effect of the carbon cycle is to convert solar energy into thermal energy that can be converted into more useful forms of energy, such as electricity and fuels (**Fig. 1**). *See* BIOGEOCHEMISTRY.

Utility. Historically, humans have used biomass for energy (burning wood) and a variety of consumer products (paper, lumber, cloth). Current industrial biomass utilization is focused on the production of fuels or energy and chemical intermediates (**Fig. 2**).

Energy. Coal and petroleum have gradually replaced biomass fuels, due to higher energy content and handling convenience. Coal and petroleum were actually created from biomass deep in the earth over millions of years. However, due to the long time periods needed to form these materials, they are considered nonrenewable. With global petroleum consumption expected to peak during the next 30 years, attention has focused on developing renewable biomass energy alternatives. Increases of atmospheric carbon dioxide attributed to petroleum combustion (global warming) have also contributed to interest in renewable biomass fuels. The amount of energy available from biomass is enormous, an estimated 2.89×10^4 exajoules (1 EJ = 10^{18} J). This is approximately eight times the total annual world consumption of energy from all sources (about 3600 EJ). At present, the world population obtains about 7% of its annual energy needs from biomass. *See* COAL; PETROLEUM.

Production of electricity from biomass combustion is being explored. Approximately 9 gigawatts of electricity are currently produced in the United States, primarily from forest products. For example, California power generation companies are burning scrap wood pellets to generate "green" electricity for interested consumers. Another power company is developing 27,000 acres of fast-growing hybrid poplar trees as a fuel source to generate 25 megawatts of electricity for Minnesota homes and industries.

Major limitations of solid biomass fuels are difficulty of handling and lack of portability for mobile engines. To address these issues, research is being conducted to convert solid biomass into liquid and gaseous fuels. Both biological means (fermentation) and chemical means (pyrolysis, gasification) can be used to produce fluid biomass fuels. For example, methane gas is produced in China for local energy needs by anaerobic microbial digestion of human and animal wastes. Ethanol for automotive fuels is currently produced from starch biomass in a two-step process: starch is enzymatically hydrolyzed into glucose; then yeast is used to convert the glucose into ethanol. About 1.5 billion gallons of ethanol are produced from starch each year in the United States.

Since cellulose is available in much larger quantities than starch, it would be a preferable source of biomass for making ethanol. However, cellulosic biomass is difficult to utilize since it is tightly combined with lignin and hemicellulose. It must be physically pretreated by grinding, steam explosion, chemical pulping, or solvent extraction to make the cellulose more accessible to enzymes. Additionally, the natural enzymes that break down cellulose are less efficient than starch enzymes, so research has focused on genetically engineering cellulose enzymes to improve efficiency and reduce costs. For example, a facility has been built in Jennings, Louisiana, to produce 20 million gallons of ethanol per year from cellulosic biomass (sugarcane wastes and rice hulls) using novel recombinant genetic microbial technology. Another facility has been built in Middletown, New York, to recycle municipal solid wastes by recycling plastics, glass, and metal and converting cellulosic wastes into 8 million gallons of ethanol annually.

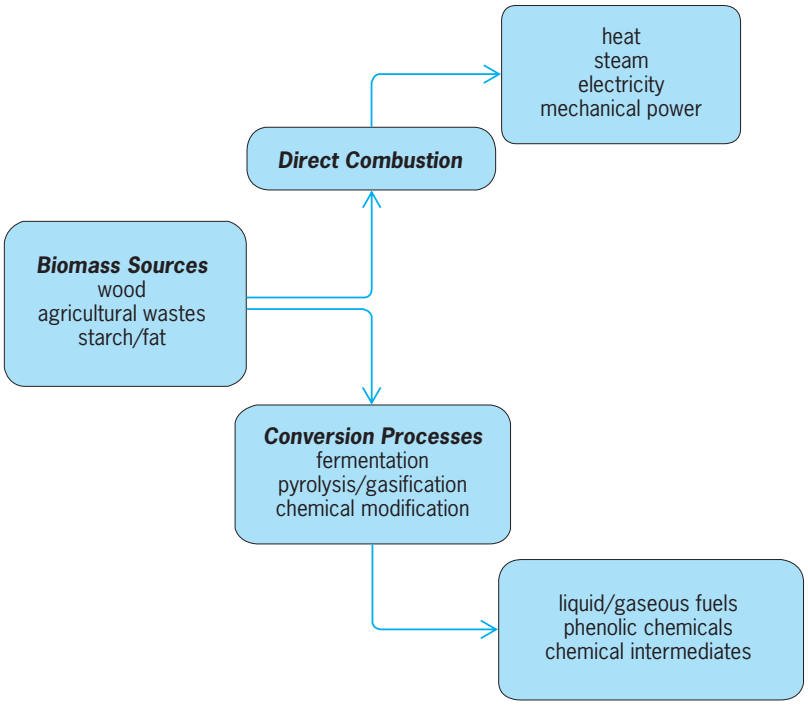

Fig. 2. Methods of converting biomass to energy, fuels, and chemicals.

Pyrolysis and gasification. Pyrolysis involves heating biomass to high temperatures in the absence of oxygen. It rearranges the chemical bonds of lignin and hemicellulose to produce chemicals for plastics, adhesives, and resins, or short-chain hydrocarbons for fuels and chemical intermediates. Given the highly variable composition of lignins and hemicelluloses, pyrolysis biomass products are very complex mixtures. Research is focused on developing reaction conditions and catalysts that make reproducible product compositions from biomass materials.

Synthetic fuel gases (mainly methane, ethane, and small hydrocarbons) can be produced by adding hydrogen and carbon monoxide during pyrolysis. This process is called gasification and converts about 65–70% of the biomass energy content into combustible gas fuels. This fuel is used like natural gas to create electricity, is used to power vehicles, or is chemically converted to other synthetic fuels. *See* PYROLYSIS.

Biodiesel fuel. Since fatty acids from plant fats and petroleum hydrocarbons have very similar chemical structures, plant fats can also be converted into a liquid fuel similar to diesel fuel. Fatty acids obtained from soybeans or other oil-rich plants are reacted with methanol to form fatty methyl esters. Since these methyl esters have nearly equivalent energy content and similar chemical structure to petrochemical diesel fuels, they have been called biodiesel

fuels. Biodiesel fuels have been demonstrated to work well in diesel engines and have reduced exhaust particulate emissions. Methyl esters have also been used as industrial cleaning solvents, in cosmetics, and for other industrial applications. *See* ALCOHOL FUEL; ETHYL ALCOHOL; GASOLINE. Bernard Y. Tao

Bibliography. J. P. Henderson, Anaerobic digestion in rural China, *BioCycle*, 39(1):79, 1997; C. M. Kinoshita et al., *Power Generation Potential of Biomass Gasification Systems: Proceedings of the American Power Conference*, Chicago, 1996; J. Lee, Biological conversion of lignocellulosic biomass to ethanol, *J. Biotechnol.*, 56:1, 1997; R. G. Lugar and R. J. Woolsey, The new petroleum, *Foreign Affairs*, 78(1):88, 1999; L. Salisbury, Customers go for the green in power supply, *Reuter Business Rep.*, February 17, 1998.

Biome

A major community of plants and animals having similar life forms or morphological features and existing under similar environmental conditions. The biome, which may be used at the scale of entire continents, is the largest useful biological community unit. In Europe the equivalent term for biome is major life zone, and throughout the world, if only plants are considered, the term used is formation. *See* ECOLOGICAL COMMUNITIES.

Each biome may contain several different types of ecosystems. For example, the grassland biome may contain the dense tallgrass prairie with deep, rich soil, while the desert grassland has a sparse plant canopy and a thin soil. However, both ecosystems have grasses as the predominant plant life form, grazers as the principal animals, and a climate with at least one dry season. Additionally, each biome may contain several successional stages. A forest successional sequence may include grass dominants at an early stage, but some forest animals may require the grass stage for their habitat, and all successional stages constitute the climax forest biome. *See* DESERT; ECOLOGICAL SUCCESSION; ECOSYSTEM; GRASSLAND ECOSYSTEM.

Distributions of animals are more difficult to map than those of plants. The life form of vegetation reflects major features of the climate and determines the structural nature of habitats for animals. Therefore, the life form of vegetation provides a sound basis for ecologically classifying biological communities. Terrestrial biomes are usually identified by the dominant plant component, such as the temperate deciduous forest. Marine biomes are mostly named for physical features, for example, for marine upwelling, and for relative locations, such as littoral. Many biome classifications have been proposed, but a typical one might include several terrestrial biomes such as desert, tundra, grassland, savanna, coniferous forest, deciduous forest, and tropical forest. Aquatic biome examples are freshwater lotic (streams and rivers), fresh-water lentic (lakes and ponds), and marine littoral, neritic, upwelling, coral reef, and pelagic. *See* FRESH-WATER ECOSYSTEM; MARINE ECOLOGY; PLANTS, LIFE FORMS OF; TERRESTRIAL ECOSYSTEM. Paul Risser

Bibliography. M. G. Barbour, J. H. Burk, and W. D. Pitts, *Terrestrial Plant Ecology*, 3d ed., 1998; S. J. McNaughton and L. L. Wolf, *General Ecology*, 1973; P. G. Risser et al., *The True Prairie Ecosystem*, 1981; R. H. Whittaker, Classification of natural communities, *Bot. Rev.*, 28:1–239, 1962.

Biomechanics

A field that combines the disciplines of biology and engineering mechanics and utilizes the tools of physics, mathematics, and engineering to quantitatively describe the properties of biological materials. One of its basic properties is embodied in so-called constitutive laws, which fundamentally describe the properties of constituents, independent of size or geometry, and specifically how a material deforms in response to applied forces. For most inert materials, measurement of the forces and deformations is straightforward by means of commercially available devices or sensors that can be attached to a test specimen. Many materials, ranging from steel to rubber, have linear constitutive laws, with the proportionality constant (elastic modulus) between the deformation and applied forces providing a simple index to distinguish the soft rubber from the stiff steel.

Living tissues. While the same basic principles apply to living tissues, the complex composition of tissues makes obtaining constitutive laws difficult. Most tissues are too soft for the available sensors, so direct attachment not only will distort what is being measured but also will damage the tissue. Devices are needed that use optical, Doppler ultrasound, electromagnetic, and electrostatic principles to measure deformations and forces without having to touch the tissue.

All living tissues have numerous constituents, each of which may have distinctive mechanical properties. For example, elastin fibers give some tissues (such as blood vessel walls) their spring-like quality at lower loads; inextensible collagen fibers that are initially wavy and unable to bear much load become straightened to bear almost all of the higher loads; and muscle fibers contract and relax to dramatically change their properties from moment to moment. Interconnecting all these fibers are fluids, proteins, and other materials that contribute mechanical properties to the tissue.

The mechanical property of the tissue depends not only upon the inherent properties of its constituents but also upon how the constituents are arranged relative to each other. Thus, different mechanical properties occur in living tissues than in inert materials. For most living tissues, there is a nonlinear relationship between the deformations and the applied forces, obviating a simple index like the elastic modulus to describe the material. In addition, the complex arrangement of the constituents leads to material properties that possess directionality; that is,

unlike most inert materials that have the same properties regardless of which direction is examined, living tissues have distinct properties dependent upon the direction examined. Finally, while most inert materials undergo small (a few percent) deformations, many living tissues and cells can deform by several hundred percent. Thus, the mathematics necessary to describe the deformations is much more complicated than with small deformations. Frank C. P. Yin

Cells and organs. Many functions of cells and organs are biomechanical and the understanding of these functions requires application of the principles and techniques of mechanics for analysis and quantitation. Biomechanics is a discipline that forms the basis for the understanding of biological processes such as the contraction of the heart, circulation of blood, gas exchange in the lung, peristalsis of the intestine and the ureter, load bearing by bones and cartilages, lubrication of joints, contraction of skeletal muscle, and mechanical aspects of vision and hearing. Biomechanics is also the foundation for clinical applications such as the design of orthopedic prostheses and cardiac assist devices. *See* PROSTHESIS.

The biomechanical properties and behaviors of organs and organ systems stem from the ensemble characteristics of their component cells and extracellular materials, which vary widely in structure and composition and hence in biomechanical properties. An example of this complexity is provided by the cardiovascular system, which is composed of the heart, blood vessels, and blood. *See* CARDIOVASCULAR SYSTEM.

Blood is a suspension of blood cells in plasma. The mammalian red blood cell consists of a membrane enveloping a homogeneous cytoplasm rich in hemoglobin, but it has no nucleus or organelles. While the plasma and the cytoplasm behave as fluids, the red blood cell membrane has viscoelastic properties; its elastic modulus in uniaxial deformation at a constant area is four orders of magnitude lower than that for areal deformation. This type of biomechanical property, which is unusual in nonbiological materials, is attributable to the molecular structure of the membrane: the lipid membrane has spanning proteins that are linked to the underlying spectrin network. The other blood cells (leukocytes and platelets) and the endothelial cells lining the vessel wall are more complex in composition and biomechanics; they have nuclei, organelles, and a cytoskeletal network of proteins. Furthermore, they have some capacity for active motility. *See* BLOOD; CYTOSKELETON.

Cardiac muscle and vascular smooth muscle cells have organized contractile proteins that can generate active tension in addition to passive elasticity. Muscle cells, like other cells, are surrounded by extracellular matrix, and cell-matrix interaction plays an important role in governing the biomechanical properties and functions of cardiovascular tissues and organs. The study of the overall performance of the cardiovascular system involves measurements of pressure and flow. The pressure-flow relationship results from the interaction of the biomechanical functions of the heart, blood, and vasculature. To analyze the biomechanical behavior of cells, tissues, organs, and systems, a combination of experimental measurements and theoretical modeling is necessary. *See* MUSCLE.

Other organ systems present many quantitative and qualitative differences in biomechanical properties. For example, because the cardiovascular system is composed of soft tissues whereas bone is a hard tissue, the viscoelastic coefficients and mechanical behaviors are quite different. Cartilage is intermediate in stiffness and requires a poroelastic theory to explain its behavior in lubrication of joints. In general, living systems differ from most physical systems in their nonhomogeneity, nonlinear behavior, capacity to generate active tension and motion, and ability to undergo adaptive changes and to effect repair. The biomechanical properties of the living systems are closely coupled with biochemical and metabolic activities, and they are controlled and regulated by neural and humoral mechanisms to optimize performance. While the biomechanical behaviors of cells, tissues, and organs are determined by their biochemical and molecular composition, mechanical forces can, in turn, modulate the gene expression and biochemical composition of the living system at the molecular level. Thus, a close coupling exists between biomechanics and biochemistry, and the understanding of biomechanics requires an interdisciplinary approach involving biology, medicine, and engineering. *See* BIOMEDICAL ENGINEERING. Shu Chien; Richard Skalak

Humans. A few examples of how mechanics is applied to human problems illustrate the scope of this field.

Forces on foot. Biomechanicists often want to know the forces that humans and animals exert on the ground when walking, running, or jumping. They use force-sensitive plates set into the floor which record the magnitude of any force that acts on them, together with its direction and the position on the plate at which it acts. **Figure 1** shows a human foot at the midpoint of a running step, the stage when the force F_g is greatest. Because the feet are off the ground for much of the stride, the force while they are on the ground rises well above body weight, to about 1900 newtons (430 pounds force) on the foot of a 700-N (158-lbf) person.

This force acts on the ball of the foot, so a force F_t is needed behind the ankle to balance the moments

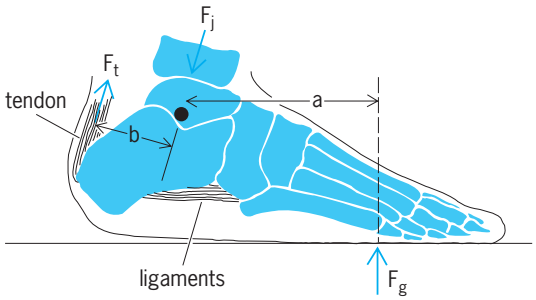

Fig. 1. Forces on a runner's foot; symbols are explained in the text.

on the foot. This is transmitted from the calf muscles through the Achilles tendon to the calcaneus (heel bone). The force F_t in the tendon can be calculated from the ground force F_g by taking moments about the ankle, as in Eq. (1). It is found to be about 4700 N

$$F_t = \frac{a}{b} F_g \tag{1}$$

(1060 lbf, or almost a half ton) in the case here considered. Finally, a downward force F_j of about 6400 N (1440 lbf) must act at the ankle joint to balance forces F_t and F_g. *See* BONE.

That calculation illustrates the large forces that act in the human body, and leads to calculations and experiments that demonstrate a basic principle of running.

Elastic mechanisms in running. The elasticity of a bouncing ball carries energy forward from one bounce to the next. When it hits the ground it is decelerated, losing kinetic energy, which is stored as elastic strain energy in the distorted ball. Elastic recoil throws the ball back into the air, reconverting strain energy to kinetic energy. The same principle operates in running, greatly reducing the energy that has to be supplied by the metabolism of foodstuffs.

The elastic properties of tendons are investigated by stretching them in machines such as engineers use for testing metals and plastics. Tendons are moderately strong (stronger than the same thickness of rubber but less strong than nylon), and they stretch by about 8% of their length before breaking. When allowed to recoil after a smaller stretch, they return no less than 93% of the energy previously used to stretch them; the rest is lost as heat. In this respect, the tendon is a very good elastic material, as good as most rubbers.

The peak force in a runner's Achilles tendon has been calculated. By knowing the force and the dimensions and properties of the tendon, it can be calculated that it must be stretched by about 0.6 in. (15 mm); the strain energy stored in it can be found by measuring the area under a graph of the force F_t in the tendon against the length to which it is stretched. This is expressed mathematically by Eq. (2).

$$\text{Strain energy} = \int F_t \, dt \tag{2}$$

This energy is found to be about 35 joules, which is about one-third of the kinetic energy lost and regained at each step. The elasticity of the tendon reduces the energy needed for running by about one-third.

A second elastic structure gives a further saving. The three forces shown in **Fig. 2** tend to flatten the arch of the foot. (Motion pictures of barefoot runners show the arch quite noticeably flattened at this stage of the stride.) This stretches the ligaments of the arch of the foot, which are made of material similar to tendons, storing up strain energy in them. Mechanical tests on feet that have had to be amputated for medical reasons show that this energy amounts to some 17 J. The elasticity of the Achilles tendon and foot ligaments together reduces the metabolic energy required for running to about half of what would otherwise be needed.

Forces in leg muscles. In the example described above, the force in the Achilles tendon was easily calculated. Many other problems in biomechanics cannot be solved unambiguously, because the body is so complex. For instance, an attempt was made to estimate the forces exerted by the various muscles of the human leg during standing in various postures. The muscles do not lie in a single plane, so the problem had to be tackled in three dimensions. The human body was treated as an assembly of seven segments (two feet, two lower legs, two thighs, and a trunk). Six equations describe the equilibrium of a rigid body, so the seven segments yielded 42 simultaneous equations, 21 for each leg. These equations could have been solved to evaluate up to 21 unknown quantities, for instance, the forces exerted by 21 muscles. Unfortunately, there were 29 muscles which had to be considered in each leg, and also various ligaments and joint surfaces which exerted unknown forces. The problem could not be solved unambiguously because, with so many muscles, there were many possible solutions involving different combinations of forces in the muscles. A unique solution could be obtained only by imposing some additional criterion, for instance, by choosing the solution which minimized the total of the forces exerted by individual muscles. It is not immediately apparent whether this or any other criterion is realistic.

Further information was obtained by electromyography, a technique which is often useful in biomechanics. When a muscle is active, electrical events (action potentials) occur in it. They can be detected by means of electrodes, either metal plates attached to the skin or fine wires inserted into the muscles. The amount of electrical activity increases as the force exerted by the muscle increases, but it is unfortunately not possible to make reliable calculations of force from records of electrical activity. However, some useful information was obtained by

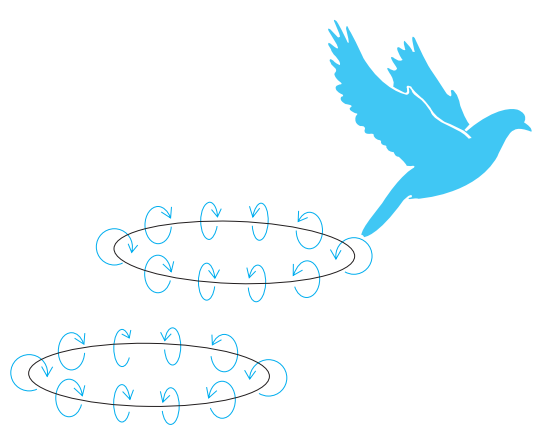

Fig. 2. Air flow in the wake of a slow-flying pigeon (*Columba livia*).

the following experiment. Electrodes were stuck to a person's skin, over six of the principal leg muscles. Electrical activity was recorded as the subject stood successively in various positions. The relative strengths of electrical activity in the various positions were compared to the muscle forces calculated by using various criteria. The electrical records were most easily reconciled with the estimates of force when the criterion for choosing a solution was to minimize the moments at joints which had to be resisted by ligaments. Even this elaborate technique gave some results which seem unlikely on other grounds. For instance, it indicated that one of the calf muscles (the soleus) exerted more than 6 kilonewtons (0.6 metric ton weight) in a deep crouch position. It is very doubtful whether the muscle is strong enough to do so.

Animals. The examples presented above concern the human body and use only solid mechanics. A final example will show that biomechanics deals also with animals, and uses fluid mechanics.

Formerly, biomechanicists tried to understand bird flight by using the aerodynamics that engineers apply to fixed-wing aircraft. It was realized that this approach could not give reliable estimates of the forces on a flapping wing, and they set out to observe the air movements in the wake of flying birds so that the forces could be calculated from first principles. Air movements were made visible by releasing small soap bubbles, filled with helium; this gas is lighter than air and so bubbles of a certain size have exactly the same density as the air and move with it. Birds were flown through clouds of helium-filled bubbles and multiple-flash photographs were taken, using two cameras to record the movement of the bubbles in three dimensions. *See* AERODYNAMICS.

Figure 2 shows the pattern of air flow behind a slow-flying bird. Each downstroke of the wings drives a downward puff of air that is surrounded by a vortex ring, like the smoke rings formed around puffs of air from a smoker's mouth. Only the downstroke drives air downward, and so it alone is responsible for supporting the bird's weight. In fast flight, however, air is driven downward throughout the down- and up-strokes, and the wake is not broken up into discrete vortex rings. Instead, the wings leave trailing vortices behind them, like the vortices that sometimes become visible as vapor trails behind the wing tips of aircraft. *See* BIOPHYSICS; FLIGHT.

R. McNeill Alexander

Bibliography. B. Alberts et al., *Molecular Biology of the Cell*, 3d ed., 1994; R. McN. Alexander, *Elastic Mechanisms in Animal Movement*, 1988; J. Currey, *The Mechanical Adaptations of Bones,* 1984; Y. C. Fung, *Biomechanics: Mechanical Properties of Living Tissues,* 2d ed., 1993; J. D. Murray, *Mathematical Biology,* 2d ed., 1993; R. Skalak and S. Chien, *Handbook of Bioengineering,* 1987; J. Valenta (ed.), *Biomechanics,* 1993; S. Vogel, *Life in Moving Fluids,* 2d revised ed., 1994.

Biomedical chemical engineering

The application of chemical engineering principles to the solution of medical problems due to physiological impairment. Knowledge of organic chemistry is required of all chemical engineers, and many study biochemistry and molecular biology. This training at the molecular level gives chemical engineers a unique advantage over other engineering disciplines in communication with life scientists and clinicians in medicine, since the technical language of modern medicine is based in biochemistry and molecular biology. Practical applications include the development of tissue culture systems, the construction of three-dimensional scaffolds of biodegradable polymers for cell growth in the laboratory, and the design of artificial organs. *See* BIOCHEMISTRY; MOLECULAR BIOLOGY.

Vascular biology. An example of where the interaction between biomedical chemical engineers and life scientists has been extremely beneficial is the area of vascular biology. The leading causes of death in the western world are coronary thrombosis (heart attack), stroke, and cancer. All these pathological states involve alterations in the cardiovascular system, eventually leading to cessation of blood flow to a vital organ. Understanding the complex interplay between blood flow, cell metabolism, cell adhesion, and vessel pathology leading to these diseases is possible only with an interdisciplinary approach.

The vascular system is naturally dynamic, with fluid mechanics and mass transfer closely integrated with blood and vascular cell function. Bioengineers are able to understand how local wall shear stress and strain modulate endothelial cell metabolism at the level of the gene. Tissue culture systems make it possible to control both the biochemical and mechanical environment—usually not possible in animal models (**Fig. 1**). This knowledge may help explain the

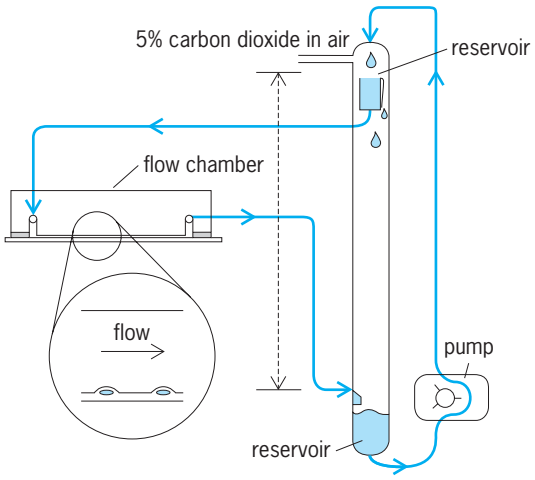

Fig. 1. Schematic of a flow loop used to study the effect of controlled flow on the metabolism of cultured endothelial cells. Cells form a monolayer on one surface of a parallel plate flow chamber that is mounted in a loop for media circulation. The flow rate is determined by the vertical distance between the upper and lower reservoirs.

focal nature of many vascular pathologies, including atherosclerosis. Understanding mechanical control of gene regulation, at the level of specific promoter elements and transcription factors involved, permits development of novel constructs for localized delivery of specific gene products in regions of high or low shear stress or strain in the vascular system. In addition, local fluid mechanics can alter receptor specificity in cell-cell and cell-matrix protein adhesion and aggregation. Knowledge of the specific molecular sequences involved in cell-cell recognition is necessary for development of targeted therapeutics, with applications in thrombosis, inflammation, cancer metastasis, and sickle cell anemia. *See* ARTERIOSCLEROSIS; TISSUE CULTURE; VASCULAR DISORDERS.

Cell transplantation therapy. Cell transplantation is explored as a means of restoring tissue function. With this approach, individual cells are harvested from a healthy section of donor tissue, isolated, expanded in culture, and implanted at the desired site of the functioning tissue. Also, cell-based therapies involve the return of genetically altered cells to the host by using gene insertion techniques.

Cell transplantation has several advantages over whole-organ transplantation. Because the isolated cell population can be expanded in the laboratory by using cell culture, only a small number of donor cells are needed to prepare an implant. Consequently, the living donor need not sacrifice an entire organ. The need for a permanent synthetic implant is eliminated through the use of natural tissue constituents, without the disruption and relocation of a whole piece of normal tissue. Isolating cells allows removal of other cell types that may be the target of immune responses, thus diminishing the rejection process. Major surgery on the recipient and donor, with its inherent risks, is avoided. Finally, the cost of the transplantation procedure can be significantly reduced.

Isolated cells cannot form new tissues and require specific environments that often include the presence of supporting material to act as a template for growth. Three-dimensional scaffolds can be used to mimic their natural counterparts, the extracellular matrices of the body. These scaffolds serve as both a physical support and an adhesive substrate for isolated parenchymal cells during cell culture and subsequent implantation. Because of the multiple functions that these materials must fill, the physical and chemical requirements are numerous. In order to accommodate a sufficient number of cells for functional replacement, a cell transplantation device must have a large surface area for cell adhesion. High porosity provides adequate space for cell seeding, growth, and production of extracellular matrices. A uniformly distributed and interconnected pore structure is important so that cells are easily distributed throughout the device and so that an organized network of tissue constituents can be formed, allowing for cell-cell communication through direct contact and via soluble factors. Also, nutrients and waste products must be transported to and from differentiated groups of cells, often in ways that maintain cell polarity. In the reconstruction of structural tissues such as bone and cartilage, tissue shape is integral to function. Therefore, these scaffolds must be made of varying thickness and shape. Furthermore, because of eventual human implantation, the scaffold must be made of biocompatible materials. As the transplanted cell population grows and the cells function normally, they will begin to secrete their own extracellular matrix support. The need for an artificial support will gradually diminish; and thus if the implant is biodegradable, it will be eliminated as its function is replaced. The development of processing methods to fabricate reproducibly three-dimensional scaffolds of biodegradable polymers that will provide temporary scaffolding to transplanted cells until they secrete their own matrix components to form a completely natural tissue replacement (**Fig. 2**) will be instrumental in engineering tissues.

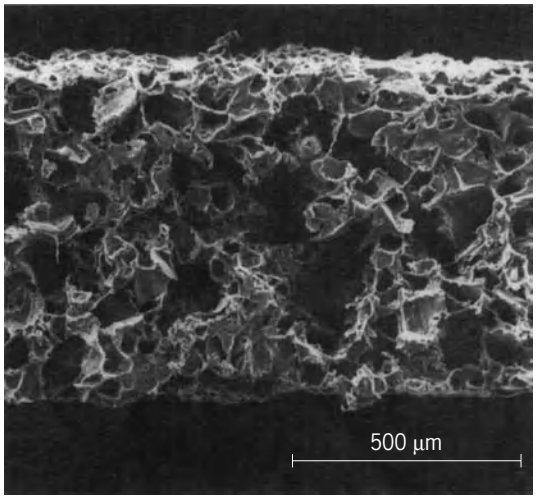

Fig. 2. Scanning electron micrograph of a biodegradable poly(L-lactic acid) foam of 91% porosity and 94-micrometer median pore diameter prepared by a particulate-leaching method which involved the casting of polymer/salt composite membranes followed by the dissolution of the salt.

The study of the adhesive interactions between cells and both synthetic and biological substrates is pivotal in determining the effect of different physical and chemical factors on cell and tissue growth and function. Much biomaterials research has focused on minimizing biological fluid and tissue interactions with biomaterials in an effort to prevent either fibrous encapsulation from foreign-body reaction or clotting in blood that contacts artificial devices. Innovations utilizing the inverse approach, that of programmed extensive interaction of the material with biological tissue, is an alternative focus for biomaterials research. One concept is novel biomaterials that incorporate specific peptide sequences to enhance cell adhesion and promote differentiated cell growth by releasing growth factors, angiogenic factors, and other bioactive molecules. *See* CELL (BIOLOGY); CELL BIOLOGY; TRANSPLANTATION BIOLOGY.

Artificial organs. When the primary organ cells cannot be grown in culture to provide the critical mass required for cell transplantation and functional

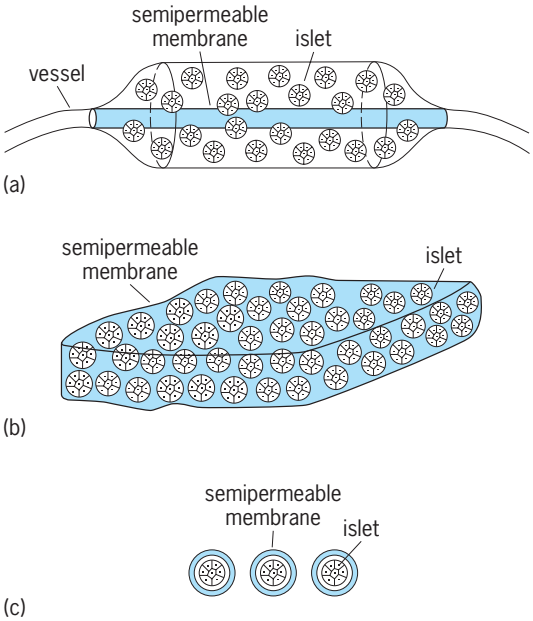

Fig. 3. Designs for islet immunoisolation devices. (*a*) Intravascular devices. (*b*) Extravascular macrocapsules. (*c*) Extrovascular microcapsules.

replacement, xenografts (tissue from other species) may be used to replace tissue function. Xenogeneic cells can be encapsulated in membranes and implanted to restore metabolic regulation and activity. These membranes are designed to isolate the cells from the body, thereby avoiding the immune responses that the foreign cells could initiate, and to allow the desired metabolites, such as insulin and glucose for pancreatic islet cells, to diffuse in and out of the membrane. Islet immunoisolation devices to replace the pancreas are classified as intravascular and extravascular (**Fig. 3**). Intravascular devices consist of a tubular membrane connected directly to the vascular system. Extravascular devices include macrocapsules and microcapsules. Macroencapsulation involves the envelopment of a large mass of cells or tissue in planar membranes, hollow fibers, or diffusion chambers. Microencapsulation involves the envelopment of single cells, islets, or a small number thereof into microspheres of diameter smaller than 1 mm. The critical issues for the development of bioartificial organs include membrane biocompatibility, diffusion limitations, device retrieval in the event of failure, and mechanical stability.

Chemical engineers have made significant contributions to the design and optimization of many commonly used devices for both short-term and long-term organ replacement. Examples include the artificial kidney for hemodialysis and the heart-lung machine employed in open-heart surgery. The artificial kidney removes waste metabolites (such as urea and creatinine) from blood across a polymeric membrane that separates the flowing blood from the dialysis fluid. The mass transport properties and biocompatibility of these membranes are crucial to the functioning of hemodialysis equipment. The heart-lung machine replaces both the pumping function of the heart and the gas exchange function of the lung in one fairly complex device. While often life saving, both types of artificial organs only partially replace real organ function. Long-term use often leads to problems with control of blood coagulation mechanisms to avoid both excessive clotting initiated by blood contact with artificial surfaces and excessive bleeding due to platelet consumption or overuse of anticoagulants. *See* BLOOD; DIALYSIS; MEMBRANE SEPARATIONS.

Other applications include methodology for development of artificial bloods, utilizing fluorocarbon emulsions or encapsulated or polymerized hemoglobin, and controlled delivery devices for release of drugs or of specific molecules (such as insulin) missing in the body because of disease or genetic alteration. *See* BIOMEDICAL ENGINEERING; POLYMER.

Larry V. McIntire; Antonios G. Mikos

Bibliography. S. Bhatia, *Microfabrication in Tissue Engineering and Bioartificial Organs*, 1999; J. D. Bronzino, *Tissue Engineering and Artificial Organs*, 3d ed, 2006; J. Enderle et al., *Introduction to Biomedical Engineering*, 1999; J. A. Frangos (ed.), *Physical Forces and the Mammalian Cell*, 1992; J. K.-J. Li, *Dynamics of the Vascular System* (*Series on Bioengineering & Biomedical Engineering*, vol. 1), 2004; A. G. Mikos et al., Islet transplantation to create a bioartificial pancreas, *Biotechnol. Bioeng.*, 43:673–677, 1994; G. Vunjak-Novakovic and R. I. Freshney, *Culture of Cells for Tissue Engineering*, 2006.

Biomedical engineering

An interdisciplinary field in which the principles, laws, and techniques of engineering, physics, chemistry, and other physical sciences are applied to facilitate progress in medicine, biology, and other life sciences. Biomedical engineering encompasses both engineering science and applied engineering in order to define and solve problems in medical research and clinical medicine for the improvement of health care. Biomedical engineers must have training in anatomy, physiology, and medicine, as well as in engineering.

Biomedical engineering entails the design and development of tools and instrumentation to enhance the ability of physicians and life scientists to obtain information on physiological and biological processes, and to aid in the treatment of diseases or other abnormalities. It is also an important component in the design and development of artificial organs, limbs, and other prosthetic devices that replace missing body parts or that enhance deteriorated physiological functions. Biomedical engineering also contributes to medical and biological research aimed at learning more about the functional processes in various physiological systems. An important part of biomedical engineering research is the use of computer techniques to develop mathematical models of physiological functions and processes to expand available theoretical knowledge. *See* SIMULATION.

The impact of biomedical engineering is manifested in the many instruments and devices that surround a patient in the modern hospital and in the many sophisticated artificial organs and other prosthetic devices available. In addition, biomedical engineering contributes greatly to medical and biological research.

Instrumentation. A wide variety of instrumentation is available to the physician and surgeon to facilitate the diagnosis and treatment of diseases and other malfunctions of the body. Instrumentation has been developed to extend and improve the quality of life. A primary objective in the development of medical instrumentation is to obtain the required results with minimal invasion of the body. Invasion refers to the penetration of the skin or external integument, usually involving increased risk and trauma to the patient. A device in the stomach accessed through the throat is not invasive, in contrast to a needle penetrating the skin to withdraw blood from a vein, or a blood pressure transducer at the tip of a catheter introduced into a blood vessel. In a broader sense, exposing the body to x-radiation or radioisotopes can also be considered invasive.

Diagnostic instrumentation. Medical diagnosis requires knowledge of the condition of specified organs or systems of the body, and a vast array of instrumentation has been developed. Bioelectric potentials, that is, electrical signals generated within the body, provide a useful source of information. Much can be learned about the functioning of the heart by using electrocardiography (EKG) to measure and analyze the electrical activity of the heart. Electrocardiographic signals are used to measure heart rate and to identify certain types of abnormalities within the heart. They also serve as a means of synchronizing other measurements with cardiovascular activity. Neurological activity within the brain can be assessed by using electroencephalography (EEG) to measure and analyze the electrical activity of the brain. Muscular activity can be measured by using electromyography (EMG), which utilizes the electrical activity of muscles. These bioelectric potentials are generally obtained noninvasively, but localized measurements may require placement of electrodes at sites that require invasive procedures. Neurological measurements usually require invasive placement of electrodes within or adjacent to specific nerves or neurons. Although bioelectric potentials have been used clinically for several decades, improved electrode design and instrumentation have made these measurements more reliable and have increased their usefulness. *See* BIOPOTENTIALS AND IONIC CURRENTS; CARDIAC ELECTROPHYSIOLOGY; ELECTRODIAGNOSIS; ELECTROENCEPHALOGRAPHY.

Sounds generated within the body constitute another source of information. Notable are heart sounds corresponding to the pumping action of the heart and operation of the heart valves; and sounds generated by blood pulsating through a partially occluded vessel, used in indirect blood pressure measurements.

Electrical impedance measurements are utilized to obtain information noninvasively. Impedance plethysmography is a technique by which changes in the volume of certain segments of the body can be determined by impedance measurements. These changes are related to factors involving the mechanical activity of the heart such as cardiac output, or conditions of the circulatory system such as blood volume and flow in various regions, or to respiratory flow and other physiological functions.

Ultrasound is utilized in many types of diagnostic measurements. In most applications, periodic bursts of ultrasound are introduced into a region of the body, and echoes from various structures and organs are measured. Echocardiography is an ultrasound procedure which provides an indication of the physical movements of the walls and valves of the heart. Echoencephalography involves ultrasonic echoes from the midline of the brain to detect tumors that would shift that midline. Electronically produced images of various internal organs and body structures are obtained by using different types of ultrasonic scans. All of these measurements are performed noninvasively. *See* ECHOCARDIOGRAPHY; MEDICAL ULTRASONIC TOMOGRAPHY.

X-rays have long been utilized to visualize internal body structures, but computerized tomographic methods and other improvements have greatly increased their diagnostic potential, permitting measurements with reduced radiation exposure to the patient. Also, better contrast media have been developed to make organs or regions of interest more clearly identifiable in the x-ray image. *See* COMPUTERIZED TOMOGRAPHY; RADIOGRAPHY; RADIOLOGY.

Another imaging process is magnetic resonance, which requires excitation of the site of measurement by an alternating magnetic field in the radiofrequency range. Magnetic resonance utilizes the principle that the nuclei of certain elements resonate when excited by a given radio frequency. Different elements have different resonance characteristics for identification. Like ultrasound, the excitation required for magnetic resonance measurements has not shown any indication of harming the body. *See* MEDICAL IMAGING; NUCLEAR MAGNETIC RESONANCE (NMR).

Imaging of organs and body structures has been greatly enhanced by computerized tomography. A number of linear scans through a given cross section of the organ or body are taken from different vantage points and combined mathematically by computer to produce a cross-sectional image in which all features are clearly shown. Tomography applied to x-ray images produces computerized tomography or computerized axial tomography (CAT) scans. When radioisotopes are used, the result is called positron emission tomography (PET) or single photon emission computerized tomography (SPECT). Computerized tomography is also applied to obtain images from other types of measurements. For example, tomographic imaging from electrical impedance measurements within the body is called electrical impedance tomography (EIT).

Monitoring equipment found at the bedside of each patient in the intensive care unit of a hospital combines a number of biomedical measurement devices to perform continuous noninvasive measurements of body temperature, heart rate, blood pressure, respiration, blood gases, and other variables that indicate the condition of the patient and that alert hospital personnel to an impending emergency. Noninvasive transcutaneous sensors permit reliable measurement of blood gases, oxygenation, and other factors in the blood. Telemetry makes it possible to monitor many of these physiological variables from an ambulatory person.

There are still many clinical situations in which noninvasive measurements are unable to provide adequate data. For example, systolic and diastolic blood pressure can be measured noninvasively by using an indirect method involving occluding cuffs and detection of the sounds of blood flowing through the occlusions. However, measurement of the actual blood pressure waveform itself is often required, and a blood pressure transducer must be carried by a catheter to a specified location within the cardiovascular system. Similarly, blood flow through a major artery or region of the body can be estimated or traced qualitatively from a probe outside the body, but reliable quantitative measurements still require invasive procedures. These examples from the cardiovascular system are typical of similar problems in nearly all of the physiological systems of the body.

Therapeutic instrumentation. Instrumentation is used in many ways for the treatment of disease. Devices are available to provide control for the automated administration of medication or other substances such as insulin for diabetics. In some cases, the rate at which the substance is administered is controlled by the level of certain constituents of the blood or other body fluids. Examples of this equipment may be found in the hospital, where the medication is administered to a patient who is in bed or whose movement is otherwise restricted. Other devices are designed to be worn on the body or even implanted within the body. *See* DRUG DELIVERY SYSTEMS.

Implanted pacemakers have been available for some time, but improvements have made them more sensitive to the natural function of the heart and the demands of the body. Such pacemakers monitor the patient's electrocardiogram and other physiological data and stimulate the heart only when the natural functions become insufficient to meet demands. The size and life of the batteries powering such devices have been improved, thus reducing the need for surgery to replace pacemaker batteries. Implantable defibrillators are available for special situations in which periodic automatic defibrillation is required to maintain functionality of a heart.

Another therapeutic technique involves the application of electrical stimulation to muscles or nerves to block involuntary muscle contractions or pain. The stimulating electrodes can be implanted or can be placed on the surface of the skin, depending on the depth of the site to be stimulated.

Surgical equipment. A number of tools are available to the surgeon in the operating room in addition to the monitoring instruments. The family of electrosurgical tools uses highly concentrated electric currents to precisely cut tissue and at the same time control bleeding. Lasers, which deliver highly concentrated light energy, are also used for making incisions and for fusion of tissue, as in the reattachment of the retina in the eye. *See* SURGERY.

Clinical engineering. Responsibility for the correct installation, use, and maintenance of all medical instrumentation in the hospital is usually assigned to individuals with biomedical engineering training. This phase of biomedical engineering is termed clinical engineering, and often involves providing training for physicians, nurses, and other hospital personnel who operate the equipment. Another responsibility of the clinical engineer is to ensure that the instrumentation meets functional specifications at all times and poses no safety hazard to patients. In most hospitals, the clinical engineer supervises one or more biomedical engineering technicians in the repair and maintenance of the instrumentation.

Biomedical research. The application of engineering principles and techniques has a significant impact on medical and biological research aimed at finding cures for a large number of diseases, such as heart disease, cancer, and AIDS, and at providing the medical community with increased knowledge in almost all areas of physiology and biology. The data obtained by medical and biological research are also used to develop more realistic and sophisticated prosthetic devices, making possible better methods of control and better understanding of the natural systems that the prostheses are designed to replace.

Biomedical engineers are involved in the development of instrumentation for nearly every aspect of medical and biological research, either as a part of a team with medical professionals or independently, in such varied fields as electrophysiology, biomechanics, fluid mechanics, microcirculation, and biochemistry. A number of fields, such as cellular engineering and tissue engineering, have evolved from this work.

Physiological modeling. A significant role for biomedical engineers in research is the development of mathematical models of physiological and biological systems. A mathematical model is a set of equations that are derived from physical and chemical laws and that describe a physiological or biological function. Experimental data obtained under conditions for which the model is expected to be valid can be used to test the model, which can then be used in several ways to represent the function or process for which it was created. By utilizing computer implementation, mathematical models can be used to predict the effect of given events, such as the administration of a given drug or treatment.

Modeling can be done at various physiological levels, from the cellular or microbiological level to that of a complete living organism, and can be of various

degrees of complexity, depending on which kinds of functions they are intended to represent and how much of the natural function is essential for the purpose of the model. Mechanical, neurological, neuromuscular, electrochemical, biochemical, thermal, biological, metabolic, pneumatic (pulmonary), hydraulic (cardiovascular), and behavioral (psychological) systems are among the many types of systems for which models have been developed and studied. The models involve transport, utilization, and control of mass, energy, momentum, and information within these systems. A major objective of biomedical engineering is to create models that more closely approximate the natural functions they represent and that satisfy as many of the conditions encountered in nature as possible. *See* MATHEMATICAL BIOLOGY.

Artificial organs and prosthetics. A highly important contribution of biomedical engineering is in the design and development of artificial organs and prosthetic devices which replace or enhance the function of missing, inoperative, or inadequate natural organs or body parts. Hemodialysis units, which perform the functions of the kidney, have been in use for a long time, but most are large external devices that must be connected to the patient periodically to remove wastes from the body. A major goal in this area is to develop small, self-contained, implantable artificial organs that function as well as the natural organs, which they can permanently supersede.

Artificial hips, joints, and many other body structures and devices designed to strengthen or reinforce weakened structures must be implanted and must function for the life of the patient. They must be made of biocompatible materials that will not cause infection or rejection by the body. They must be built with sufficient reliability that the need for adjustment, repair, or replacement will be minimized.

Sophisticated artificial limbs for amputees and neural prostheses to facilitate or strengthen the use of nonfunctional limbs is another area in which biomedical engineers are involved. Prosthetic limbs attempt to replicate the vital features of natural limbs with miniature electric motors acting as muscles to operate the moving parts. Control of the movements is a major challenge, especially for the arm and hand, where there is a particular need to provide dexterity for the user. Electromyographic signals from muscles remaining in the stump and elsewhere in the body are utilized where possible. However, the limitations are severe, and the ability of the user to learn to operate the device effectively poses a further problem. The problem is compounded by the need for sensory feedback from the prosthesis in order to provide smooth control. Simple on-off switching operations can be accomplished by moving a part of the body such as the shoulder, by placing a device in the mouth that can be operated by the tongue, by a puff-sip action, or by other motions within the physical capability of the patient.

Neural prostheses bypass a portion of the nervous system and provide electrical stimulation to existing muscles in situations where paralysis has interrupted the natural control pathways. The electrical stimulation can be applied either along the surface of the skin over the muscles to be stimulated or via electrodes implanted within or immediately adjacent to the muscle. Control of the muscle action when using these prostheses is even more difficult than with artificial limbs, especially if the sensory path from the muscle has also been interrupted. Some success has been achieved in providing functional electrical stimulation to specified muscles in the foot to overcome a condition called drop foot. The stimulation is controlled by a switch activated by the position of the foot at a specified portion of the walking cycle. Neural prostheses for the lower limbs have been successful in helping paralyzed patients walk to a limited degree. *See* PROSTHESIS.

Rehabilitation engineering. The goal of rehabilitation engineering is to increase the quality of life for the disabled. One major part of this field is directed toward strengthening existing but weakened motor functions through use of special devices and procedures that control exercising of the muscles involved. Another part is devoted to enabling disabled persons to function better in the world and live more normal lives.

Included in this area are devices to aid the blind and hearing-impaired. Hearing aids are available to assist persons in which the acoustical signal has weakened. Cochlear implants that bypass the hair cells and directly stimulate the auditory nerve make it possible to restore some hearing ability to many individuals. Reading machines that either produce greatly enlarged images or convert written words into sound have been developed to assist the visually impaired. All of these devices have their shortcomings, and work to create improved visual and auditory aids continues. *See* HEARING AID.

Human-factors engineering is utilized in modifying the home and workplace to accommodate the special needs of disabled persons. Improved wheelchairs and wheelchair control devices and modified automobile control mechanisms permit greater mobility for many who are unable to walk or drive a conventional vehicle. Special modifications to the control mechanisms of television sets and other devices and appliances in the home allow for use by people with disabilities. *See* BIOMECHANICS; HUMAN-FACTORS ENGINEERING; MEDICAL CONTROL SYSTEMS.

Fred J. Weibell

Bibliography. L. A. Geddes and L. E. Baker, *Principles of Applied Biomedical Instrumentation*, 3d ed., 1989; D. C. Mikulecky and A. M. Clarke (eds.), *Biomedical Engineering: Opening New Doors*, 1990; D. J. Schneck, *Engineering Principles of Physiologic Function*, 1990; R. B. Stein, P. H. Peckham, and D. B. Popovic (eds.), *Neural Prostheses: Replacing Motor Function after Disease or Disability*, 1992; J. G. Webster (ed.), *Medical Instrumentation: Applications and Design*, 3d ed., 1997; D. L. Wise (ed.), *Bioinstrumentation: Research, Developments and Applications*, 1990.

Biomedical ultrasonics

The applications to medicine and biology of sound waves that have a frequency higher than the audible spectrum. Biomedical applications of ultrasound range from cell sonicators using frequencies in the kilohertz range to ultrasonic imaging in the megahertz range. The best-known application, ultrasonic imaging, is the second most utilized diagnostic imaging modality, after x-rays. High-intensity ultrasound has been used for therapeutic applications.

Ultrasonic imaging possesses numerous advantages over other imaging modalities of similar capabilities such as x-ray computed tomography, radionuclide imaging, and magnetic resonance imaging. It uses radiation that is noninvasive to the human body at the diagnostic intensity level, produces images at a very fast rate of 30 frames per second, and can be used to yield blood flow information by applying the Doppler principle. It has been used in a variety of medical disciplines, including cardiology, obstetrics, and radiology. *See* MEDICAL IMAGING.

Notable disadvantages of ultrasound are that organs containing gases and bony structures cannot be adequately imaged, and only a limited acoustic window (space or opening between bones) is available for examination of organs such as the heart and the neonatal brain.

Ultrasound at high intensity levels has been used for hyperthermia treatment of tumors, frequently in combination with chemotherapy or radiotherapy, and for tissue ablation. *See* ONCOLOGY; THERMOTHERAPY.

Ultrasonic wave propagation. Ultrasound is a form of acoustic wave with a frequency higher than 20 kHz. Ultrasound parameters include pressure, velocity, displacement, density, and temperature. Ultrasound velocity and wavelength in water are approximately 1.48×10^5 cm/s and 0.03 cm at 5 MHz and 20°C (68°F). The sound speeds in soft biological tissues are similar to that in water.

An acoustic disturbance in a fluid propagates as a longitudinal or compressional wave, which has a displacement in the same direction as the direction of wave propagation. On the other hand, the displacement of a shear or transverse wave is perpendicular to the wave direction. The sound speeds of shear and longitudinal waves are different. Biological tissues other than bone can be treated as fluidlike media, and as such they cannot support shear-wave propagation. *See* SOUND.

Characteristic acoustic impedance. The characteristic acoustic impedance of a medium is defined as the ratio of the pressure to medium velocity. Since these quantities are analogous to the voltage and current in an electrical circuit, the acoustic impedance carries the same physical significance as the electrical impedance. For a fluid medium, the characteristic acoustic impedance Z is given by Eq. (1),

$$Z = \rho c \tag{1}$$

where ρ is the density and c is the sound speed. The unit of acoustic impedance, the rayl, is 1 kg/(m^2-s). The acoustic impedance of water is approximately 1.48×10^6 rayl, and the acoustic impedances of tissues are of similar magnitude. *See* ACOUSTIC IMPEDANCE.

Intensity. Analogous to the power in electrical circuits, the intensity of an acoustic wave is defined as the power carried by the wave per unit area, propagating in a medium of acoustic impedance Z. It is related to the pressure, p, by Eq. (2).

$$I(t) = \frac{p^2(t)}{Z} \tag{2}$$

A conventional pulse-echo ultrasonic image is formed by launching an ultrasonic pulse into the body, and detecting and displaying the echoes reflected or scattered from the tissues. The pulse duration of the scanner is very short, typically less than 1 microsecond, but the period between successive pulses is relatively long, typically greater than 1 millisecond. Consequently, the instantaneous pressure when the pulse is on can be quite high, reaching a few megapascals, but the time-averaged pressure over the whole cycle is much lower. *See* SOUND PRESSURE.

Attenuation, reflection, and scattering. As an ultrasonic wave penetrates a tissue, part of the ultrasonic energy is lost through absorption of the energy by tissues and part by scattering or reflection, which diverts the energy from the incident direction whenever the wave encounters a discontinuity in acoustic properties. For a plane wave, the sound pressure decreases as a function of the propagation distance, x, exponentially, as given by Eq. (3), where $p(0)$ is

$$p(x) = p(0)e^{-\alpha x} \tag{3}$$

the pressure at $x = 0$ and α is the attenuation coefficient in neper/cm. The attenuation coefficient in tissues has been found to be linearly proportional to frequency in the range from 1 to 30 MHz. In soft tissues it is approximately 0.03 neper/(cm-MHz). *See* SOUND ABSORPTION.

The behavior of a plane wave at a flat boundary between two media depends on the acoustic impedances of the media and on the angles of reflection and transmission of the sound wave. Most of the incident energy is reflected at the interface of two media of very different acoustic impedances. Ultrasound is not useful for imaging anatomical structures that contain air, such as lung and bones, because their acoustic impedances are very different from those of the surrounding body tissues, so that most of the ultrasound energy is reflected at these boundaries and does not penetrate them.

Echoes that result from reflection or scattering in the tissues are used by pulse-echo ultrasonic scanners to form an image. Scattering is described by a parameter called the scattering cross section, defined as the power scattered by the object when the incident beam has unit incident power per unit area. The scattering cross section of a scatterer of simple

geometry, for example, a sphere, whose dimensions are much smaller than a wavelength, can be solved analytically, and is called Rayleigh scattering. At 10 MHz, the scattering cross section of a human red blood cell is approximately $1.5 \times 10^{-12}\text{cm}^2$, which is extremely small.

Doppler effect. The Doppler principle has been used to estimate blood flow in a blood vessel noninvasively. For an ultrasound beam making an angle θ relative to blood flowing at velocity V, the Doppler frequency shift f_{Dop} is related to V by Eq. (4), where f

$$f_{\text{Dop}} = \frac{2V\cos\theta}{c} f \tag{4}$$

is the frequency of the sound. The Doppler frequencies associated with blood flow typically are in the audible range for medical ultrasound frequencies. *See* DOPPLER EFFECT.

Ultrasonic imaging. An ultrasonic image is formed from echoes returned from tissues. Although there are a number of ways of displaying the image, the B (brightness) mode is the most common. In this mode, the echo amplitude is displayed in gray scale as a function of the distance into the tissue. The center of the B-mode imaging system (**Fig. 1**) is an energy-conversion device, the ultrasonic transducer, which converts the electric signal into ultrasound and vice versa. *See* MEDICAL ULTRASONIC TOMOGRAPHY; TRANSDUCER.

Transducers. An ultrasonic transducer contains one or more elements made from a piezoelectric ceramic such as lead zirconate titanate. Depending upon how a transducer is driven, ultrasonic scanners are classified as mechanical-sector and electronic-array scanners. Electronic-array systems produce images of better quality and are faster. Mechanical-sector scanners are no longer used except for special applications. A typical array has dimensions of 1×8 cm (0.4×3.2 in.), and consists of 128 or more small rectangular piezoelectric elements (**Fig. 2***a*). The gaps between the elements, called kerfs, are filled with an acoustic insulating material. In a linear sequenced array, or simply a linear array, a subaperture of 32 elements or more is fired simultaneously to form a beam. The beam is electronically focused and swept across the aperture surface from one end to the other (Fig. 2*b*). Electronic dynamic focusing can be achieved by appropriately controlling the timing of the signals to each individual element. In a linear phased array, the beam is also electronically steered, resulting in a pie-shaped image. Dynamic focusing can be used to better focus the beam throughout the field of view during reception or transmission. *See* PIEZOELECTRICITY.

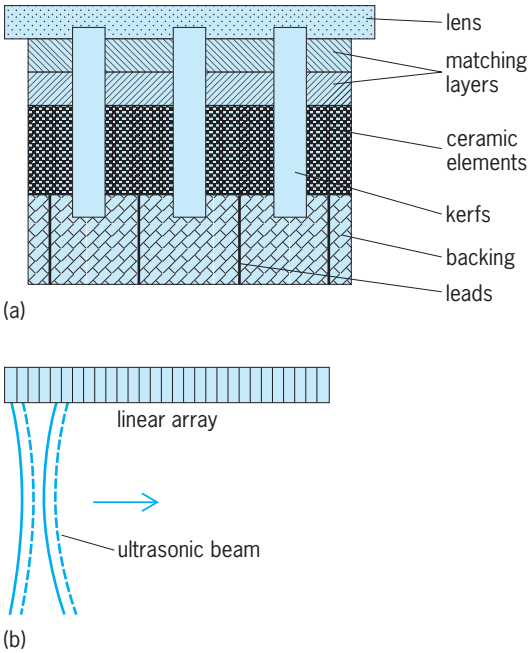

Fig. 2. Ultrasonic transducer consisting of linear array of piezoelectric elements. (*a*) Cross section of array. (*b*) Electronic focusing and sweeping of a beam by a linear sequenced array (or simply linear array). Generation of beam indicated by the solid line is followed by that indicated by the broken line.

The array may include two matching layers, a lens, and a backing material (Fig. 2*a*). The matching layers are needed to reduce the acoustic impedance mismatch between the piezoelectric ceramic and water or soft tissues. The lens is used to focus the beam to control the thickness of the image slice. The backing modifies the duration and shape of the excited acoustic pulse. A transducer consisting of only one piezoelectric element is the simplest type used in biomedical ultrasonics.

The array is excited by a pulser (Fig. 1). The returned echoes from the tissues are detected by the same transducer. Beam sweeping and position are controlled and encoded by the beam former, which may be either analog, consisting of a matrix of delay lines and amplifiers, or digital, consisting of an array of preamplifiers and analog-to-digital converters. The signal-processing chain performs functions

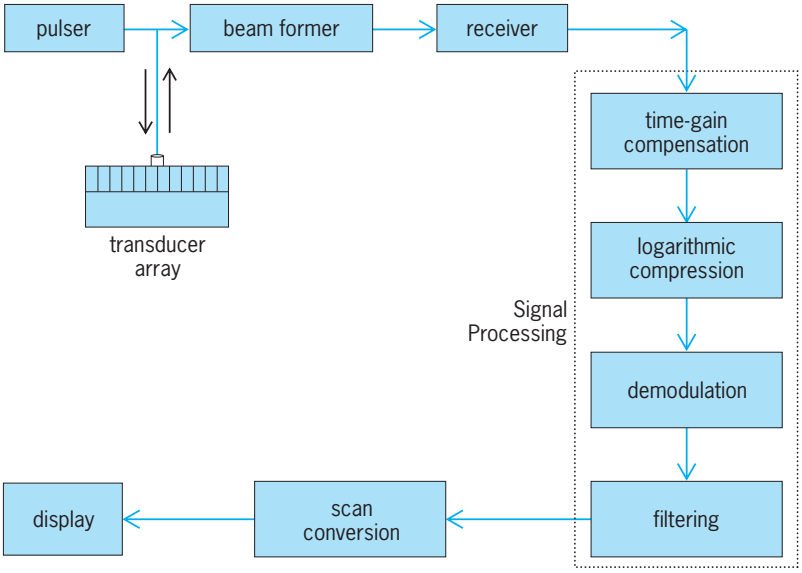

Fig. 1. Block diagram of a B-mode ultrasonic imaging system.

including time-gain compensation, signal compression, demodulation, and filtering. Time-gain compensation is used to compensate for the signal loss due to tissue attenuation and beam divergence as the beam penetrates the body. Logarithmic signal compression accentuates lower-level signals. Following signal processing, the signal may be digitized by an analog-to-digital converter. In high-end systems, analog-to-digital conversion is done in the beam former. The scan converter is a digital memory device that stores data that have been converted into a format that is displayable by a monitor. Before display, the video data may be processed again. Modern B-mode scanners can acquire images at a rate of 30 frames per second, allowing the motion of an organ to be monitored (**Fig. 3**).

The B-mode ultrasonic images exhibit a granular appearance, called a speckle pattern, which is caused by the constructive and destructive interference of the wavelets scattered by the tissue components as they arrive at the transducer surface. Ultrasound imaging is diffraction-limited because focusing can be achieved only within a limited field of view, not in the regions near and very far from the transducer. *See* SPECKLE.

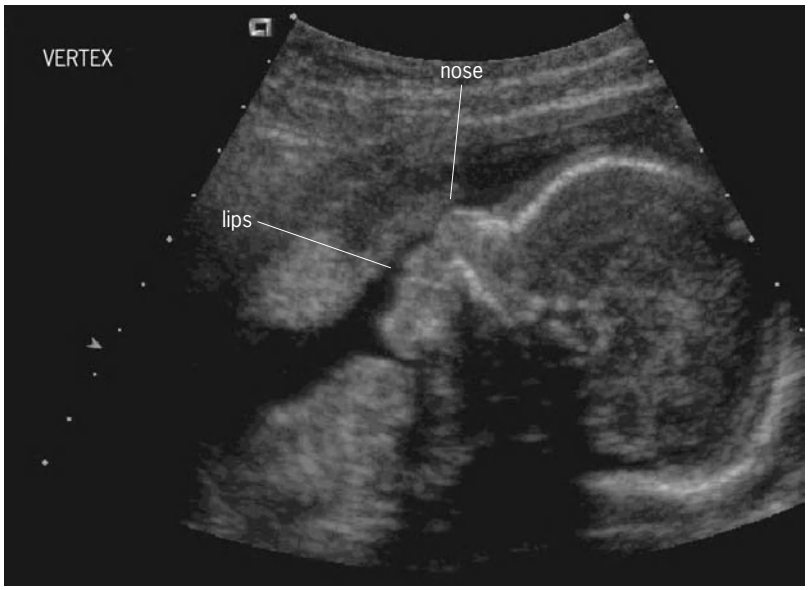

Fig. 3. B-mode image of a fetus in the uterus

Flow measurements. Two approaches have been used for ultrasonic Doppler flow measurements: continuous-wave and pulsed-wave Doppler. A probe consisting of two piezoelectric elements, one for transmitting the ultrasound signal and one for receiving the echoes returned from blood, is used in continuous-wave Doppler. The returned Doppler-shifted echoes are amplified, demodulated, and low-pass-filtered to remove the carrier frequency. The demodulated signal can be heard with a speaker or analyzed with a spectrum analyzer. Alternatively, a zero-crossing counter can be used to estimate the mean Doppler frequency.

A pulsed-wave Doppler may be used to alleviate the problem that a continuous-wave Doppler is unable to differentiate the origins of the Doppler signals produced within the ultrasound beam. Ultrasound bursts of 5–10 cycles are used, and the time of flight of the returned signal yields information about the signal's origin. *See* FLOW MEASUREMENT.

Color Doppler flow imaging. Color Doppler flow imaging scanners are capable of obtaining both B-mode images and Doppler blood flow data simultaneously in real time. The Doppler information is encoded in color. Red and blue are assigned to indicate flow toward and away from the transducer, respectively. The magnitude of the velocity is represented by the shade of the color. The color Doppler image is superimposed on the gray-scale B-mode image.

The basic concept of the color Doppler imager is similar to that of the pulsed Doppler instruments, which extract the mean Doppler-shift frequency from a sample volume. The only difference is that the color Doppler instruments are capable of estimating the mean Doppler shifts of many sample volumes along a scan line in a very short period of time, of the order of 30–50 ms. A fast algorithm has been developed to do so. It calculates the mean Doppler shift and variance from the autocorrelation function of the signal, reducing significantly the processing time. If, instead of the mean Doppler shift, the power of the signal is displayed, this mode of imaging is called color Doppler power imaging. It has the advantages that (1) the data can be averaged to achieve a better signal-to-noise ratio and (2) signal distortion is no longer a problem since only the power is detected. The disadvantages are that it is more susceptible to motion artifacts and the image contains no information on flow velocity and direction.

Multidimensional arrays. The linear arrays used by current scanners are capable of dynamic focusing in only one dimension, in the azimuthal or imaging plane. Beam focusing in the elevational plane, the plane perpendicular to the imaging plane, which determines the slice thickness, must be achieved by a lens. This means that the slice thickness varies in the field of view and is the smallest only near the focus. To further improve the image quality, multidimensional arrays are necessary to control the beam in the elevational plane. The design and fabrication of such arrays have proved difficult. The technology used in microelectromechanical systems (MEMS) to miniaturize electronic components as well as novel interconnection technology may have to be incorporated in these systems.

3-D imaging and parallel processing. Three dimensional (3-D) reconstruction from multiple two-dimensional ultrasonic images can be readily achieved off-line by spatially encoding the imaging plane. Ultrasonic imaging in three dimensions in real time is feasible only with two-dimensional arrays which allow beam steering and dynamic focusing in three dimensions and parallel processing of the data. Only one acoustic pulse is propagated in the tissue at any one time in conventional ultrasonic image processing. The scanning time can be considerably shortened if several scan lines of information

can be collected and processed at the same time by transmitting a broad beam.

Contrast media and harmonic imaging. Ultrasonic contrast agents have been successfully developed to enhance the image contrast of anatomic structures that contain the agents. A majority of these agents utilize microscopic encapsulated gas bubbles, which are strong scatterers because of the acoustic impedance mismatch. Moreover, the bubbles can resonate when insonified by an ultrasonic wave, and therefore the echoes from the bubbles can be further enhanced if the incident wave is tuned to the resonant frequency of the bubbles. These bubbles have a mean diameter of less than 5 micrometers and can be injected intravenously.

An important application of the gas-containing agents is in harmonic imaging and Doppler measurements, where the effect of the surrounding stationary structures on the image and results is minimized because only microbubbles resonate and emit ultrasound at harmonic frequencies. Blood flowing in a blood vessel containing the contrast agent will produce harmonic signals, but the blood vessel will not. Various types of oral contrast media which absorb or displace stomach or bowel gas for ultrasonic imaging have also been developed for better visualization of the abdominal organs.

Harmonic imaging has also been found to be useful even without a contrast agent to improve the image quality, especially in patients who are difficult to image with conventional ultrasound devices. This technique is possible because harmonic signals are produced from the nonlinear interaction of ultrasound and tissues.

Biological effects of ultrasound. Although ultrasound is known as a form of noninvasive radiation, biological effects inevitably result if its intensity is increased beyond some limit. Two types of ultrasound bioeffects can be produced: thermal and mechanical. The thermal effect results from an increase in temperature when the ultrasound energy is absorbed by the tissue, whereas the mechanical effect is caused by the generation of microscopic air bubbles or cavitation. No bioeffects have been found for ultrasound with a spatial peak temporal average intensity less than 100 mW/cm^2. *See* CAVITATION.

Therapeutic ultrasound. Ultrasound has been used for therapeutic purposes since high-intensity ultrasound produces biological effects. Based upon the duration and intensity level of the exposure during therapy, ultrasonic therapeutic applications can be categorized as hyperthermia and noninvasive surgery. Tissues are exposed to ultrasound of longer duration at lower intensity levels in ultrasound hyperthermia. The duration is 10–30 min at an intensity level of a few watts per square centimeter, elevating tissue temperature to 43–45°C (109–113°F). For noninvasive surgery, the duration is shorter, on the order of a few seconds, at an intensity level of a few hundred watts per square centimeter, raising the temperature in the focused region of the transducer to 70–90°C (158–194°F) within a short period of time. Cell death results immediately at this temperature. High-intensity focused ultrasound surgery has been shown to be clinically efficacious in the treatment of benign prostate hyperplasia, ocular tumor, and retinal tear. *See* ULTRASONICS.

K. Kirk Shung

Bibliography. D. H. Evans et al., *Doppler Ultrasound: Physics, Instrumentation, and Signal Processing*, 2d ed., Wiley, New York, 2000; W. N. McDicken, *Diagnostic Ultrasonics: Principles and Use of Instruments*. 3d ed., Churchill Livingstone, Edinburgh, 1991; J. A. Zagzebski, *Essentials of Ultrasound Physics*, Mosby, St. Louis, 1996.

Biometeorology

A branch of meteorology and ecology that deals with the effects of weather and climate on plants, animals, and humans.

Thermal equilibrium. The principal problem for living organisms is maintaining an acceptable thermal equilibrium with their environment. Mechanisms used to accomplish this equilibrium vary widely. Organisms have natural techniques for adapting to adverse conditions. These techniques include acclimatization, dormancy, and hibernation, or in some cases an organism can move to a more favorable environment or microenvironment. Humans often establish a favorable environment through the use of technology. *See* DORMANCY; HIBERNATION AND ESTIVATION; MICROMETEOROLOGY.

Homeotherms, that is, humans and other warm-blooded animals, maintain relatively constant body temperatures under a wide range of ambient thermal and radiative conditions through physiological and metabolic mechanisms. Poikilotherms, that is, cold-blooded animals, have a wide range in body temperature that is modified almost exclusively by behavioral responses. Plants also experience a wide range of temperatures, but because of their immobility they have less ability than animals to adapt to extreme changes in environment. *See* THERMOREGULATION.

Healthy humans, for example, maintain a body temperature of about 98.6°F (37°C). The thermal equilibrium is maintained through metabolism and through heat exchange and radiative exchange with the environment. Heat is gained by metabolism, gained or lost by radiative exchange, and generally lost by convection, conduction, and evaporation. The body surface loses heat when it radiates heat away from the body toward a colder environment, such as a cold wall or a clear sky, and gains radiative heat from the Sun or from a warm surface. Heat is exchanged directly to the air or through physical contact with a surface by conduction or convection. If the surfaces are cooler than the body temperature, heat is removed from the body, and if the surfaces are warmer, the body gains heat. Evaporation always causes a heat loss, while metabolism is always a source of heat. The amount of heat gained by metabolism is a function of the level of activity. *See* HOMEOSTASIS; METABOLISM.

TABLE 1. Insulating qualities of clothing

Environment	Sex	Garments	clo units
Tropical heat	m	Open-necked shirt, shorts, sandals	0.20
	f	Bra and panties, short-sleeved blouse, shorts, sandals	0.25
Warm summer	m	Open-necked shirt, slacks, shorts, ankle socks, shoes	0.40
	f	Bra and panties, short-sleeved blouse, light skirt, short stockings, shoes	0.30
Comfortable weather	m	Business suit, cotton underwear, long-sleeved shirt, socks, shoes	1.0
	f	Bra and panties, slip, dress, panty-hose, pumps	1.0
Cool weather	m	Business suit, light underwear, socks, shoes, light overcoat	1.5
	f	Bra and panties, stockings, slip, dress, sweater, shoes	1.5
Cold weather	m	Business suit, long underwear, woolen socks, shoes, hat, overcoat	2–2.5
	f	Warm underwear, skirt or slacks, slip, long-sleeved blouse, heavy sweater, hat, overcoat	2–2.5
		(For very cold weather, add lined gloves, fur or leather garment for both m & f)	3–3.5
Polar cold	m & f	Woolen underwear, coveralls, quilted parka with hood, fur mittens, fur-lined boots, face mask	4–4.5

Humans are physically adapted to a narrow range of temperature, with the metabolic mechanism functioning best at air temperatures around 77°F (25°C). There is a narrow range above and below this temperature where survival is possible. To regulate heat loss, warm-blooded animals developed hair, fur, and feathers. Humans invented clothing and shelter. The amount of insulation required to maintain thermal equilibrium is governed by the conditions in the atmospheric environment. There are a limited number of physiological mechanisms, controlled by the hypothalamus, that regulate body heat. In cold environments these mechanisms increase metabolic heat production, as in shivering, and lower the temperature at the skin surface by restricting blood flow. This mechanism, known as vasoconstriction, lowers the temperature at the body surface and therefore reduces the radiative heat loss. In warm environments the blood flow to the surface is increased, raising the temperature and thus increasing the radiative heat loss. In addition, evaporative cooling is enhanced by perspiration and respiration. *See* RESPIRATION.

The earliest humans had only the natural defenses available to counteract the atmospheric environment; hence their habitat was restricted to the warm climates of the tropics. The development of clothing permitted settlement of other climates. The insulating quality of clothing is measured by a unit (clo) derived from the metabolic heat losses and predicated on maintenance of skin temperature of 91.4°F (33°C). The heat flux from the body can be calculated with Eq. (1), where n = number of degrees

$$\text{Heat flux} = \frac{n}{0.18(x)} \qquad (1)$$

(°C) and x = number of clo units (**Table 1**). Many animals exhibit winter fur that has significant insulating qualities (**Table 2**).

Cold stress. Clothing, shelter, and heat-producing objects can largely compensate for environmental cold, but with extensive exposure to cold, vasoconstriction in the peripheral organs can lead to chilblains and frostbite on the nose, ears, cheeks, and toes. This exposure is expressed quantitatively as a wind chill equivalent temperature that is a function of air temperature and wind speed. The wind chill equivalent temperature is a measure of convective heat loss and describes a thermal sensation equivalent to a lower-than-ambient temperature under calm conditions, that is, for wind speeds below 4 mi/h (1.8 m/s). *See* COMFORT TEMPERATURES.

Persons exposed to extreme cold develop hypothermia, which may be irreversible when the core temperature drops below 91°F (33°C). The core temperature defines the temperature of the internal vital organs, for example, the cardiopulmonary system, in contrast to that of the peripheral organs, for example, hands or skin. On the whole, there is a clear correlation between winter death rates and temperature: the colder the winter, the more deaths occur. Research findings seem to indicate that outdoor deaths may be caused by breathing cold air, which lowers the temperature of the cardiopulmonary system to a lethal level. When the ambient temperature is 68°F (20°C), exhaled air has a temperature of about 91°F (33°C). It has a value of 73°F (23°C) when the ambient air is at 23°F (−5°C). *See* HYPOTHERMIA.

Heat stress. The combination of high temperature with high humidity leads to a very stressful thermal environment. The combination of high temperature with low humidity leads to a relatively comfortable thermal environment, but such conditions create an environment that has a very high demand for water.

Conditions of low humidity exist principally in subtropical deserts, which have the highest daytime temperatures observed at the Earth's surface. Temperatures of 86–104°F (30–40°C) are common, and occasionally even temperatures of 122°F (50°C) or higher are observed in these areas. Simultaneous rel-

TABLE 2. Insulating characteristic of winter fur of selected mammals

Species	Approximate thickness, in. (cm)	Relative insulation quality
Weasel	0.4 (1)	2
Squirrel	0.8 (2)	2.5
Rabbit	1.2 (3)	5
Dog	1.6 (4)	6
Beaver	1.6 (4)	5
Sheep	2.8 (7)	8

ative humidities are usually below 10% and rarely exceed 30% in daytime. Human and animal bodies are also exposed to strong solar radiation and and radiation reflected from the surface of the sand. This combination makes extraordinary demands on the sweat mechanism. Water losses of 1.06 quarts (1 liter) per hour are common in humans and may be even greater with exertion. Unless the water is promptly replaced by fluid intake, dehydration sets in. In humans a 2% weight loss by dehydration is slight and causes no symptoms other than thirst. If the loss rises to 6%, scanty urine and a rapid pulse accompanying a temperature rise of 3.6°F (2°C) will ensue; a greater weight loss by dehydration will cause circulatory failure and brain damage. Death occurs when the weight loss reaches 10–15% of body weight. Desert populations use the cooler night hours, with absence of incoming radiation, for travel and other activities. *See* DESERT; HUMIDITY; SOLAR RADIATION.

Animals that live in deserts are mostly nocturnal. Lizards and rodents spend the day in hollows and burrows because temperatures are 18–36°F (10–20°C) cooler than the extreme temperatures near the surface. Desert animals have also developed various mechanisms to maintain thermal equilibrium. The most water-efficient desert animal is the camel, which can store fluid in its system and survive several days without water. Camels can survive a 25% weight loss by dehydration. In other species, 10–15% weight loss is the limit.

Dehydration can occur in heated buildings during winter because the cold outside air has very little moisture, and as it infiltrates into the building the relative humidity drops drastically. Household activities such as cooking, laundering, bathing, and growing house plants add moisture, but unless humidifiers are used, water losses of occupants can be substantial. At a room temperature of 70°F (21°C) and a relative humidity of 50%, a resting individual loses about 1.06 quarts (1 liter) of water in 24 h. At a relative humidity of 30% and the same temperature, the water loss is about 1.16 quarts (1.1 liters), mostly from insensible perspiration and exhaled air.

Both outdoor and indoor comfort is governed by temperature and humidity conditions. However, there is a great variability among individuals in terms of comfort zones. Acclimatization plays a role, but in general, older persons and infants require higher temperature values for comfort. On the average, there is a spread of 2.6–9°F (2–5°C) for comfort sensations for persons living in middle or higher latitudes compared with residents of subtropical or tropical climates, where there is a narrower spread of 2–5°F (1–3°C). In the warmer climates, the absolute comfort values for individuals are 5–7°F (3–4°C) higher than for those in the cooler climates.

When humans are exposed to warm environments, the first physiological response is dilation of blood vessels, which increases the flow of blood near the skin. The next response occurs through sweating, panting, and evaporative cooling. Since individuals differ in their physiological responses to environmental stimuli, it is difficult to develop a heat stress index based solely on meteorological variables. Nevertheless, several useful indices have been developed.

Several indices describe comfort sensations as functions of ambient temperatures and humidities. These include the humiture index, the humidex, and the temperature-humidity index (THI). The THI value is computed from Eqs. (2), using wet- (T_w)

$$\mathrm{THI} = 0.4(T_d + T_w) + 15 \quad [\text{for } T_d, T_w \text{ in } ^\circ\mathrm{F}] \qquad (2a)$$

$$\mathrm{THI} = 0.4(T_d + T_w) + 4.78 \quad [\text{for } T_d, T_w \text{ in } ^\circ\mathrm{C}] \qquad (2b)$$

and dry-bulb (T_d) temperatures measured with a psychrometer. *See* PSYCHROMETER.

Only a few people will be comfortable when THI = 53°F (21°C) or less, and almost everyone will be uncomfortable when THI = 80°F (26.5°C). When THI exceeds 91°F (33°C), heat stroke leading to death may occur. Under exertion, even at lower THI values, panting, high pulse rate, and cardiopulmonary symptoms may develop. Indices based solely on temperature and humidity data are applicable for buildings and other confined conditions where wind and solar radiation loads are at low levels.

Since wind moves body heat away and increases the evaporation from a person, it should be accounted for in developing comfort indices describing the outdoor environment. One such index, used for many years by heating and ventilating engineers, is the effective temperature (see **illustration**). People will feel uncomfortable at effective temperatures above 81°F (27°C) or below 57°F (15°C); between 63°F (17°C) and 77°F (25°C) they will feel comfortable.

Statistics for the United States show increased mortality during summer heat waves. Older persons living in non-air-conditioned quarters are particularly afflicted. Individuals with congestive heart disease are especially endangered by heat. Victims of multiple sclerosis also find their symptoms aggravated by hot atmospheric environments.

The use of certain drugs is inadvisable during hot weather. For example, phenothiazines, which include several major tranquilizers, reduce the ability to sweat and can cause heat stroke. The use of diuretics should be restricted, and drugs used to alleviate Parkinson's disease must be avoided during hot periods. Many drugs that inhibit sweating have been banned from use in the tropics. Coffee or caffeine-containing preparations should also be avoided because they constrict the peripheral blood vessels. The use of fans, shade, and especially air conditioning can reduce the acute troubles caused by excessive heat. *See* AIR CONDITIONING.

Weather sensitivity. Both physiological and psychological responses to weather changes (meteorotropisms) are widespread and generally have their origin in some bodily impairment. Reactions to weather changes commonly occur in anomalous skin tissue such as scars and corns, changes in atmospheric moisture cause differential hygroscopic expansion and contraction between healthy and abnormal skin, leading to pain. Sufferers from

rheumatoid arthritis are commonly affected by weather changes; both pain and swelling of affected joints have been noted with increased atmospheric humidity. Sudden cooling can also trigger such symptoms. Clinical tests have shown that in these individuals the heat regulatory mechanism does not function well, but the underlying cause is not understood.

Weather is a significant factor in asthma attacks. Asthma as an allergic reaction may, in rare cases, be directly provoked by sudden changes in temperature that occur after passage of a cold front. Often, however, the weather effect is indirect, and attacks are caused by airborne allergens, such as air pollutants and pollen. An even more indirect relationship exists for asthma attacks in autumn, which often seem to be related to an early outbreak of cold air. This cold air initiates home or office heating, and dormant dust or fungi from registers and radiators are convected into rooms, irritating allergic persons. *See* ALLERGY; ASTHMA.

A variety of psychological effects have also been attributed to heat. They are vaguely described as lassitude, decrease in mental and physical performance, and increased irritability. Similar reactions to weather have been described for domestic animals, particularly dogs, hence hot, humid days are sometimes known as dog days.

Many of the physiological or pathological reactions to weather in the middle and higher latitudes have been related to a typical sequence of weather events that are common in the cold season. In such a sequence there are six weather phases: (1) high pressure: clear sunny sky, light winds, cool; (2) falling pressure: few clouds, dry moderate winds, solar warming; (3) falling pressure: increasing cloudiness; (4) passage of warm front: rain (or snow in winter), rising temperatures, fresh winds, occasional fog, high humidity; (5) passage of cold front: precipitation, sharp drop in temperature, rapidly rising pressure, brisk winds; (6) high pressure, cool temperature, low humidity. A weather map can be used to show these phases with specific notations for each phase. Weather phase 4, the warm front passage, is the one most frequently related to meteorotropic symptoms. It also has an apparent relation to the frequency of industrial accidents and suicides. In schools, disciplinary infractions are higher and students' test scores are lower. In phases 4 and 5, weather elements change most rapidly, and there is an increase in cardiovascular deaths; also, phase 5 has been associated with the initiation of labor at the end of pregnancies. *See* WEATHER MAP.

There is speculation concerning biological effects of small atmospheric ions that are continuously created by cosmic radiation and by decay of natural radioactive substances. Their lifetime is very short, and they either recombine or attach themselves to larger aerosols, most of which in settled regions are air pollutants. These ions can cause allergic reactions. *See* AIR POLLUTION.

Meteorological and seasonal changes in natural illumination have a major influence on animals. Photoperiodicity is widespread. The daily cycle of illumination triggers the feeding cycle in many species, especially birds. In insectivores the feeding cycle may result from the activities of the insects, which themselves show temperature-influenced cycles of animation. Bird migration may be initiated by light changes, but temperature changes and availability of food are also involved. In the process of migration, especially over long distances, birds have learned to take advantage of prevailing wind patterns. In humans, light deprivation, as is common in the cold weather season in higher latitudes, is suspected as a cause of depression. Exposure to high-intensity light for several hours has been found to be an effective means of treating this depression. *See* AFFECTIVE DISORDERS; MIGRATORY BEHAVIOR; PHOTOPERIODISM.

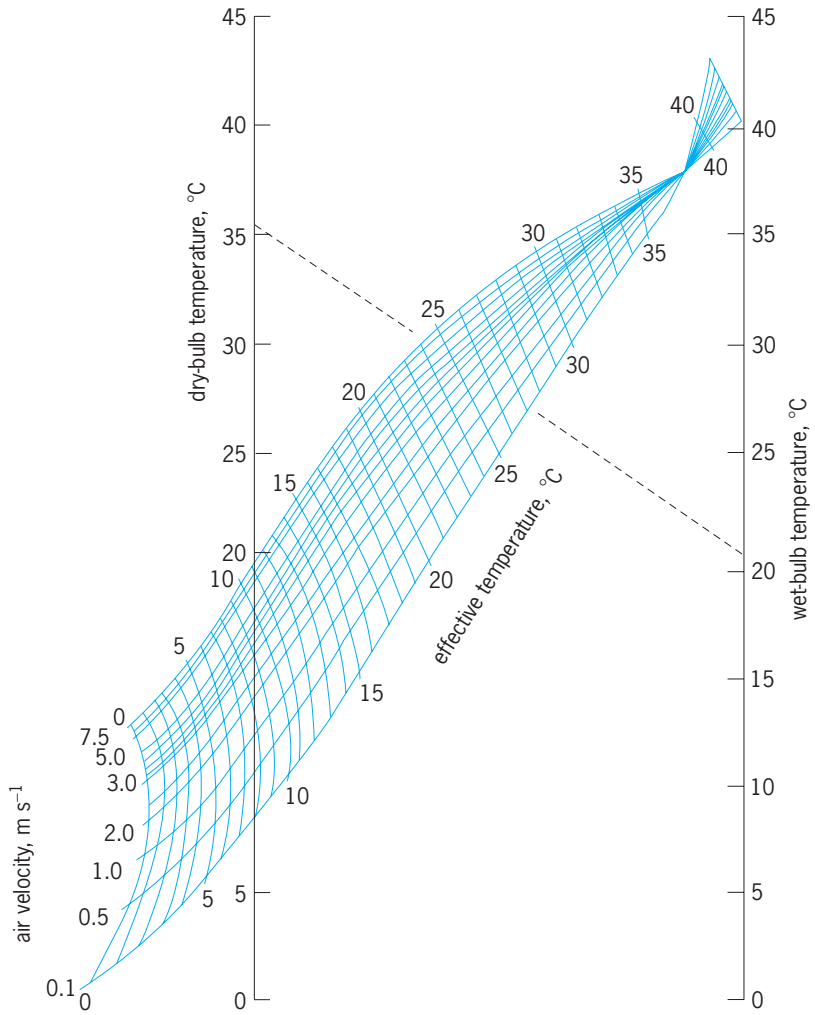

Chart of the effective temperature as a function of air velocity and wet-bulb and dry-bulb temperatures °F = (°C × 1.8) + 32°; 1 m s^{-1} = 0.21 mi/h.

Ultraviolet radiation. The atmosphere forms a protective shield from cosmic radiation. This energetic, lethal, shortwave radiation is absorbed in the highest layers of the atmosphere and is involved in the formation of ozone. Fortunately, ozone absorbs all ultraviolet radiation less than about 290 nanometers wavelength and much of the radiation between 290 and 320 nm, known as UV-B radiation, which causes sunburn and with repeated lengthy exposure can lead to skin cancer. Solar radiation in the 320–400-

nm range, known as UV-A, is less dangerous, but it can produce sunburn. All ultraviolet radiation can produce thickening and aging of the skin. Exposure of the human eye to ultraviolet radiation, especially UV-B, can lead to cataracts and lesions of the cornea and retina. Discovery of an ozone hole over the polar regions led to widespread concern that manufactured chemicals such as the chlorofluorohydrocarbons are interacting with the ozone and reducing the protective ozone shield, thereby permitting increasing amounts of ultraviolet radiation to reach the Earth's surface. *See* RADIATION INJURY (BIOLOGY); ULTRAVIOLET RADIATION (BIOLOGY).

Altitude. With an increase in altitude there is a decrease in the amount of oxygen, and individuals who are accustomed to the amount of oxygen in the air near sea level may suffer from hypoxia (lack of oxygen) at higher elevations, especially upon exertion. The result may be mountain or altitude sickness. For this reason, cabins of high-flying aircraft are pressurized. However, humans and animals can become acclimatized to higher elevations through changes in the cardiopulmonary system. This acclimatization is accomplished by large increases in the oxygen-carrying red blood corpuscles and in lung alveoli. *See* HYPOXIA.

Weather and animals. Wild animals have adapted to their environment through natural selection, but domesticated animals, because of their selective breeding and confinement, are often poorly adapted. The maintenance of a proper body temperature and an acceptable water status is critical for the growth, production, and survival of animals. Production of meat, eggs, and milk by farm animals is often closely linked to weather conditions. There has been considerable research to determine the response of animals to their atmospheric environment, and to develop management practices to help minimize losses and maximize production.

Many breeds of domesticated animals are adapted to temperate climates and do not do well under hot, humid conditions. Not only does production fall under hot, humid conditions, but care must be taken to prevent death. Animals should be placed where air circulates freely, where water is readily available, and where shade is provided. In many cases, animals are kept in confined conditions where temperature, humidity, and ventilation are controlled. In other instances, they may be in partially confined conditions, or the microclimate may be beneficially modified. *See* AGRICULTURAL BUILDINGS; AGRICULTURAL SCIENCE (ANIMAL).

Adaptation and acclimatization. Humans and animals often exhibit a remarkable ability to adapt to harsh or rapidly changing environmental conditions. An obvious means of adaptation is to move to areas where environmental conditions are less severe, examples are birds and certain animals that migrate seasonally, animals that burrow into the ground, and animals that move to shade or sunshine depending on weather conditions. Animals can acclimatize to heat and cold. The acclimatization process is generally complete within 2–3 weeks of exposure to the stressful conditions. For example, in hot climates heat regulation is improved by the induction of sweating at a lower internal body temperature and by the increasing of sweating rates. The acclimatization to cold climates is accomplished by increase in the metabolic rate, by improvement in the insulating properties of the skin, and by the constriction of blood vessels to reduce the flow of blood to the surface. *See* ADAPTATION (BIOLOGY); BURROWING ANIMALS.

Weather and insects. Weather plays a major role in the development and movement of insects. Winds carries some insects long distances. The primary control of body temperature by insects is accomplished through behavioral responses. They regulate their thermal balance by moving to and from the sunshine, by burrowing, by flying or gliding, or by attaching themselves to objects such as rocks or vegetation. Temperature plays a significant role in determining the rate of growth and development of insects. The timing of transitions from one stage of growth to another is determined in large measure by the thermal environment. Temperature is the most important environmental variable that influences insect development, but humidity and moisture content also play a role. *See* INSECT PHYSIOLOGY.

Weather and flora. Unlike humans and animals, plants cannot move from one location to another; therefore, they must adapt genetically to their atmospheric environment. Plants are often characteristic for their climatic zone, such as palms in the subtropics and birches or firs in regions with cold winters. Whole systems of climatic classification are based on the native floras. *See* ALTITUDINAL VEGETATION ZONES; PLANT GEOGRAPHY.

In plants, photosynthesis, respiration, and transpiration are governed by the atmospheric environment. In these processes, water is usually limiting, and so the ratio of rainfall to evapotranspiration is a governing factor in plant development. In addition, growth requires certain soil and air threshold temperatures; soil temperature and moisture regulate seed germination; and photosynthetic rates are also governed by light and temperature. *See* HYDROMETEOROLOGY; PHOTOSYNTHESIS; PLANT GROWTH; PLANT RESPIRATION.

In responding to their environment, plants experience various cycles, so that at certain levels of light and temperature, specific growth responses such as leaving, flowering, and fruit development take place. These seasonal phases can be disrupted by adverse weather conditions, such as drought and frosts. *See* COLD HARDINESS (PLANT); VERNALIZATION.

Agriculture has adopted many management practices to help modify microclimatic conditions to improve the atmospheric environment to enhance crop production. For example, irrigation is used to supplement natural precipitation; windbreaks are established to reduce mechanical damage to plants, to improve plant-water relations, and to increase plant photosynthesis; mulches are used to retard soil evaporation, thereby making more water available to plants and plastic mulches warm soils more rapidly in the spring. Numerous frost protection practices

have been developed. *See* AGRICULTURAL METEOROLOGY; ECOLOGY; IRRIGATION (AGRICULTURE); METEOROLOGY. Blaine L. Blad; H. E. Landsberg

Bibliography. G. S. Campbell, *An Introduction to Environmental Biophysics*, 2d ed., 1998; G. E. Folk, *Introduction to Environmental Physiology*, 2d ed., 1974; D. M. Gates, *Biophysical Ecology*, 1980; N. J. Rosenberg, B. L. Blad, and S. B. Verma, *Microclimate: The Biological Environment*, 1983; S. W. Tromp, *Biometeorology*, 1980.

Biometrics

The application of mathematical and statistical methods to describe and analyze data concerning the variation of biological characteristics obtained from either observation or experiment. The concept of a population and the samples derived from it are discussed in this article.

A population, or "universe," is a technical term meaning a collection of events or objects with certain characteristics. A population may be finite or infinite. It may exist in physical form or may be abstract and conceptual. Examples of a population are income of all families in the United States, height of all college males, outcome of throwing a die an infinitely large number of times, and all possible offspring of a specified type of mating. The population must be precisely identified and defined in any study.

An observation technically means recording the actual value of the characteristic of a member, or element, of the population under investigation. Such a value is an observed value. A complete recording of the values of all the members of a (finite) population is called a census, while a partial recording consisting of only a limited number of observations (for example, n) is called a sample. A sample may be large (n large) or small (n small). In most practical cases, a complete examination of a population is unattainable, and one must be satisfied with the study of a limited number of observations based on one or more samples. From any specified population, many samples can be drawn which, although from the same population, may consist of quite different observed values. These actually observed sets of values are usually referred to as data.

Random sample refers to a sample wherein each member is drawn at random from a population. At random means the complete lack of any systematic selection scheme in drawing members of the samples. In particular, members of a sample are not chosen with reference a priori to their values for each characteristic. Various mechanical devices and tables of random numbers have been prepared to facilitate random sampling. Such a sample is most important in theory and practice because long-range properties of such samples, based on the totality of all possible samples, are known and may be described by mathematical expressions.

Sample size is the number n of observations in a sample. Under a random sampling scheme, each sample would be more or less different from other similarly drawn samples. As sample sizes increase, however, the random variation from sample to sample decreases, and the sample itself resembles the population more and more. Size, therefore, is an important feature of a sample. It determines essentially the reliability or degree of resemblance between the sample and the population.

Parameter and statistics. If y is the measurement of a characteristic (height, weight, age, income, blood pressure, or some other quantitative trait) of an individual member, and there are n such values in a sample, then the value, as in Eq. (1), is called the

$$\bar{y} = (y_1 + \cdots + y_n)/n = \Sigma y/n \qquad (1)$$

arithmetic mean of the sample. The mean value of y for the entire population is denoted by $\mu = E(y)$, which is a fixed value. $E(y)$ is read as the expected value of y. It can be said that μ is a parameter of the population and $\bar{y}$ is an estimate of μ. As a general term, any value calculated from the n observations of a sample is called a statistic. The arithmetic mean illustrates simply the meaning and relationship between a population parameter and a sample statistic. If the $\bar{y}$ is calculated from n random observations, then the long-range average value of $\bar{y}$ from a large number of such samples is equal to the population mean μ. Therefore, one can state that $e(\bar{y}) = \mu$. In such a case, $\bar{y}$ is an unbiased estimate of μ. Various other parameters of a population may be defined, and the corresponding sample estimates may be constructed. Next to the mean, the most important parameter is variance.

Variance and its estimates. With each y value is associated a value $y - \mu$, known as the deviation from the mean. Population variance σ^2 is defined as the average value (expected value) of the squared deviations in the population, as shown by Eq. (2), where

$$\sigma^2 = E[(y - \mu)^2] \qquad (2)$$

E is "expectation of . . . ," y is a measurement of a characteristic, and μ is mean value of y. Since the magnitude of σ^2 depends on the magnitude of the deviations $y - \mu$, variance is employed as a measure of variability or dispersion of the y values about their mean. In usual cases, when the true population mean μ is unknown, the value of σ^2 may be estimated by the statistic, sample variance, as in Eq. (3).

$$s^2 = \frac{\Sigma(y - \bar{y})^2}{n - 1} \qquad (3)$$

The numerator $\Sigma(y - \bar{y})^2$ is the sum of squares of deviations from the sample mean. The estimate is so constructed that the long-range average value of s^2 from repeated random sampling is equal to the population variance σ^2; namely, $E(s^2) = \sigma^2$. Thus, s^2 is an unbiased estimate of σ^2. The square root of a variance is called the standard deviation.

Degrees of freedom. The denominator of s^2 is $n - 1$. Although there are n independently observed ys in the sample, the sum of squares in the numerator is calculated from the sample mean $\bar{y}$, not from the fixed population mean μ. Thus, one constant has been fitted to the data before the deviations can be obtained. This fitting causes the loss of 1 degree of freedom for estimating the errors so that the sum of squares has only $(n - 1)$ degrees of freedom. Hence, the above expression for s^2. Generally, each linear restriction on the observed values takes up one degree of freedom. The estimate of a variance is always a certain sum of squares divided by its corresponding degree of freedom.

Variance of sample mean. The sample mean varies from sample to sample, although it does not vary as such as the single ys. For random samples of n observations each, the variate $\bar{y}$ has a mean μ (as stated previously) and the variance of $\bar{y}$ is expressed by Eq. (4), or more simply as Eq. (5). The variance of $\bar{y}$ is only $1/n$ of

$$E[(\bar{y} - \mu)^2] = \sigma_{\bar{y}}^2 = \sigma^2/n \tag{4}$$

$$\sigma_{\bar{y}} = \sigma/\sqrt{n} \tag{5}$$

the variance for single values. When n is large, the sampling variance of $\bar{y}$ is small, so that the sample mean is close to the population mean.

Data presentation and description. When the amount of data is large, whether from a large sample or an entire population, it is convenient to arrange such data into tabular form to facilitate grasping the salient features of the data. This is done by collecting observations of more or less the same value into one class, and the result is a frequency table.

Frequency table. In **Table 1** the measurements of height of 400 college males are grouped into nine classes. There are 12 individuals with height between 61 and 63 in., which are one set of class limits. The width of a class, known as the class interval, is 2 in. The number of individuals in a class is called the frequency of the class. The sum of all the class frequencies is the total number of observations: $\Sigma f = n$. The fraction f/n is the relative frequency. Thus, for the 61–63 class, the relative frequency is $12/400 = 0.030$, or 3% of the total observations. Relative frequencies always add up to unity of 100%. In condensing the 400 single observations into a compact frequency table, some of the details of individual measurements are lost, but a quick recognition of the distribution of height is gained. In practical calculations, the height of the f students in a class is taken to be the midpoint of that class. The mean and standard deviation are then calculated by using Eqs. (6), (7), and (8). The standard error of the sample mean from $n = 400$ observations is $s_{\bar{y}} = s/\sqrt{n} = 2.66/20 = 0.133$ in. In calculating s^2 in this example, n is used for simplicity instead of the correct $n - 1$, since n is large and it makes little difference whether 400 or 399 is being employed.

$$\bar{y} = \frac{1}{n}\Sigma fy = 67.50 \text{ in.} \tag{6}$$

$$s^2 = \frac{1}{n}\Sigma f(y - \bar{y})^2 = \frac{1}{n}(\Sigma fy^2) - \bar{y}^2 = 7.07 \text{ in.}^2 \tag{7}$$

$$s = \sqrt{7.07} = 2.66 \text{ in.} \tag{8}$$

TABLE 1. Frequency table

Class limits, in.	Midpoint, y	Frequency, f	Relative frequency, f/n
59–61	60	4	0.010
61–63	62	12	0.030
63–65	64	50	0.125
65–67	66	102	0.255
67–69	68	120	0.300
69–71	70	78	0.195
71–73	72	28	0.070
73–75	74	4	0.010
75–77	76	2	0.005
Total		400	1.000

Histogram and frequency polygon. A frequency table may be converted into a diagram for easy visual inspection of the distribution. The diagram commonly used is a series of rectangles of equal width corresponding to the equal class intervals and height proportional to the class frequency or relative frequency. Such a series of rectangles is called a histogram (**Fig. 1**). If all individuals in a class are assumed to have the same midpoint value, these midpoints may be plotted against their corresponding frequencies. These points lie in the middle at the top of the rectangles. Two adjacent points may be connected by a straight line. The resulting series of straight lines is called a frequency polygon. If the variate, such as height, takes only certain discrete values, a bar or solid column may be erected at each value with the height

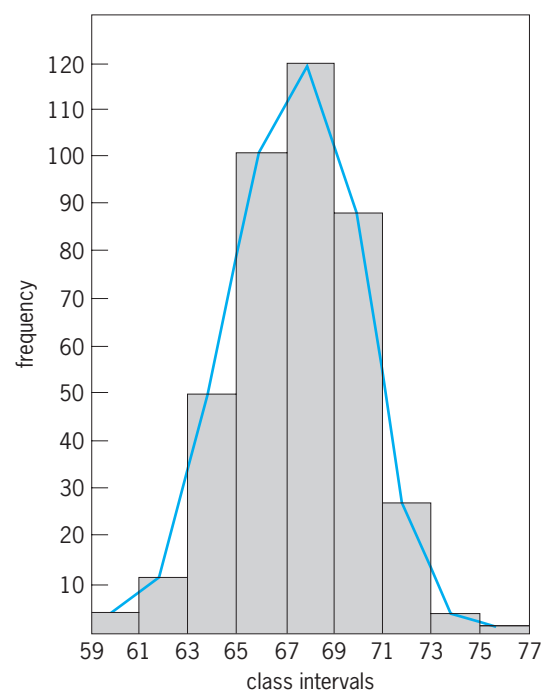

Fig. 1. Histogram and frequency polygon.

proportional to the frequency. The result is a bar graph. Even in such a case, one may still use the frequency polygon to help trace the change in frequency from class to class.

Frequency curve. Although the data are condensed into a frequency table which consists of a limited number of classes with arbitrary class limits, it is recognized that such a variate as height really changes continuously and may assume any value within a certain range. If there are a very large number of observations (theoretically, infinitely large) and the class interval becomes infinitely narrow, the frequency polygon will approach a smooth frequency curve as a limit (**Fig. 2**). Therefore, in all theoretical work dealing with a continuous variate, the distribution (height, for example) in the entire population will be described by a mathematical expression which, when plotted, will assume the shape of the frequency curve. Such a mathematical expression is called the density function. The relative frequency between two values a and b is the area under the frequency curve. This relative frequency is bounded by the ordinates $y = a$ and $y = b$.

Mode and median. The mode is the value of the variate at which the frequency is the highest, that is, the most frequently occurring value of the variate. For example, for the height of the college male students the modal value is 68 in., taken as the midpoint of the modal class 67–69 in. For a continuous distribution, the modal value is that at which the frequency curve reaches its peak. If there is one peak in the frequency curve, the distribution is said to be unimodal. If there are two peaks, the distribution is bimodal. The median is such a value in the variate that half of the observations are greater and half are smaller than it. Thus, the median height in the example is approximately 67.53 in., assuming that the 120 individuals are uniformly distributed in the 67–69 class and applying linear interpolation. Then 200 students are taller and 200 students are shorter than 67.53 in. For a continuous distribution, the vertical line erected at the median divides the entire area under the frequency curve into two equal halves. The ratio of standard deviation to mean is called the coefficient of variability (c.v.).

Normal distribution. Normal distribution refers to a symmetrical bell-shaped frequency distribution (**Fig. 3**). Many distributions in nature are approximately of the normal shape. It may be shown that if a variate y, such as height, is influenced by a large number of factors, each exerting independently a small effect on y, and if these effects are additive, the resulting distribution of y will approach the normal shape as a limit. If a large number of measurements are taken repeatedly of an object, these measurements will not be identical no matter how carefully taken, but will assume more or less varying values which are, again, normally distributed. For this reason, the normal distribution is also known as the distribution of errors. Large deviations are comparatively few, while small deviations may occur quite frequently. It is the most useful distribution in both theoretical and practical statistics, for, in many cases, a normal population may be assumed; and when the population is not normal, the statistics calculated from repeated random samples (sampling distribution) assume approximately the normal form. A normal distribution has two parameters: the mean μ and the variance σ^2 (or standard deviation σ). When these two values are specified, the normal curve is uniquely determined. In terms of standardized units, the mean is zero and variance is unity whatever the actual values of μ and σ^2 are. Since the normal distribution is symmetrical, mean = mode = median; the area under the normal curve between the mean μ and $\mu + \sigma$ is approximately 0.34, and so the relative frequency of the y values within the interval $\mu \pm \sigma$ is approximately 0.68, or 68%. At the y values $\mu \pm \sigma$, the curve twists from a concave to a convex one. The relative frequency within the interval $\mu \pm 1.96\sigma$ (slightly less than 2σ on either side of the mean) is 0.95, and thus, that outside the same interval is 0.05 with 0.025 under each tail. The areas or relative frequencies of the normal distribution have been tabulated in various ways, but always with the standardized unit as the independent variable.

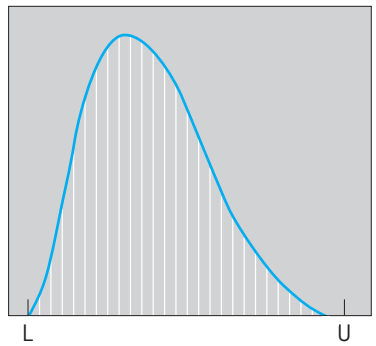

Fig. 2. Frequency curve. L and U represent the lower and upper limits of the range of values.

The sample mean $\bar{y}$ from n observations is almost normally distributed, even though the y population may not be normal itself. This is true especially when the population is unimodal, and the size of the sample is moderately large. The facts that $E(\bar{y}) = \mu$ and $\sigma^2_{\bar{y}} = \sigma^2/n$ make possible certain statements about the population mean. The normal deviate for the

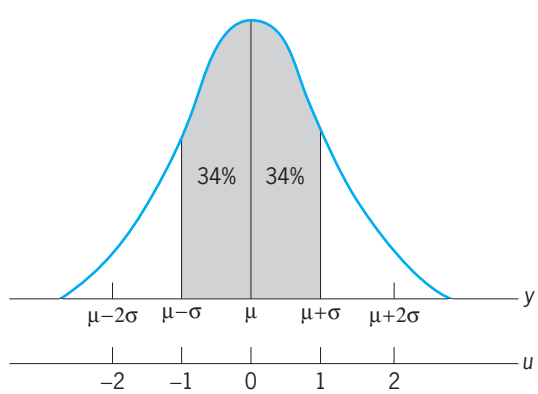

Fig. 3. Normal, symmetrical distribution.

variate $\bar{y}$ (not y) is expressed by Eq. (9). Then the

$$u = \frac{\bar{y} - \mu}{\sigma_{\bar{y}}} \tag{9}$$

probability is 0.95 that inequality (10) holds. The inequality may be written as (11), in which the limits

$$-1.96 < \frac{\bar{y} - \mu}{\sqrt{\sigma^2/n}} < +1.96 \tag{10}$$

$$\bar{y} - 1.96\sqrt{\sigma^2/n} < \mu < \bar{y} + 1.96\sqrt{\sigma^2/n} \tag{11}$$

vary from sample to sample while the central item μ remains constant. In repeated sampling, these varying limits will "catch" or cover the population mean 95% of the time. This gives an interval estimate of μ. For practical calculation, s^2 may be substituted for σ^2. The values $\bar{y} \pm 1.96;\sigma^{\bar{y}}$ are called the 95% confidence limits, and the width between these two limits is called the 95% confidence interval. The 95% confidence in this example is arbitrary. Other degrees of confidence may be chosen.

Test of hypothesis. This is also known as a test of significance. Assume that $\bar{y}$ and s^2 have been calculated from a moderately large sample and that, instead of using $\bar{y}$ as an estimate of the population mean, one may want to test the hypothesis that the population mean is equal to a certain fixed value, for example, m. If m is indeed the population mean, then the normal deviate for the variate $\bar{y}$ is calculated from Eq. (12), where s^2 may be substituted for σ^2 in

$$u = \frac{\bar{y} - m}{\sqrt{\sigma^2/n}} \tag{12}$$

practical calculations for large samples. The probability that this u value will lie between -1.96 and 1.96 is 0.95 and that it will lie outside this interval is $\alpha = 0.05$, if m is the true population mean. The rule of testing is that if $|u| > 1.96$, it is regarded as an unusual phenomenon hardly consistent with the hypothesis. In such an event, the hypothesis is abandoned and the difference $\bar{y} - m$ is said to be statistically at the $\alpha = 5\%$ significance level. Instead of the normal deviate 1.96 and the corresponding probability of 0.05, used as an example here, other significance levels may be chosen. The testing of a hypothesis results in either its acceptance or rejection. In so doing, either a type I or a type II error may be committed.

Type I and type II errors. If $|u| > 1.96$ is employed as the criterion to reject the hypothesis $\mu = m$ which happens to be true, the probability of this wrong rejection is $\alpha = 0.05$. This type of error is called the type I error. It is the error of attribution of false significance. On the other hand, the hypothesis itself may be untrue. Suppose that the true mean is M, and yet is is hypothesized that it is m. Ideally, the untrue hypothesis should be rejected. However, the distribution with mean M overlaps with the distribution with mean m, and the sample may well yield a value of u within the interval -1.96 and 1.96. In such a situation, the untrue hypothesis would not be rejected. This error is called type II error, the error of failure to detect significance. The probability of committing a type II error depends upon, among several other things, the value of M. The greater the difference between M and m, the less probability of committing the second type of error. The two types of errors cannot be committed at the same time. In any particular test, if significance is declared, only the risk of committing an error of type I is run; and if the hypothesis is retained, only the risk of committing an error of type II is run. With proper balancing of the two types of error in a long series of similar tests, the first kind of error is sometimes committed and the second kind of error committed other times, but one will be correct in rejecting the wrong hypothesis and accepting the right hypothesis most times. The increase of sample size can reduce both types of error (**Table 2**).

TABLE 2. Decision about hypothesis

	Reject	Accept
Hypothesis $\mu = m$ true	Type I error	Correct conclusion
Hypothesis $\mu = m$ false	Correct conclusion	Type II error

Difference between two means. Suppose that there are two independent samples of sizes s_1^2 and s_2^2. To determine if these two means can be considered significantly different, one sets up the hypothesis that the two population means are equal: $\mu_1 = \mu_2$. A hypothesis of this nature is called a null hypothesis, because it assumes the absence of differences. Now consider the new variate $d = \bar{y}_1 - \bar{y}_2$. If the null hypothesis is true, the expected value of d is shown by Eq. (13), and the variance of d is shown by Eq. (14). Hence,

$$E(\bar{y}_1) - E(\bar{y}_2) = \mu_1 - \mu_2 = 0 \tag{13}$$

$$\sigma_d^2 = \sigma_{y1}^2 + \sigma_{y2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \tag{14}$$

the normal deviate for the variate d is represented by Eq. (15). The procedure of testing the significance

$$u = \frac{d}{\sqrt{\sigma_d^2}} = \frac{(\bar{y}_1 - \bar{y}_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \tag{15}$$

is the same as discussed previously except that the variate d is being considered and the particular hypothesis being tested is that its mean value is zero.

Student's t distribution. Since σ is unknown, s may be used in its place for large samples without introducing serious error. However, in small samples the value of s varies from sample to sample too much to be taken as a constant. Student's t is the ratio of a deviation to its varying sample s. In other words, it is a normal variate standardized in terms of the sample standard deviation. The distribution of t closely resembles that of u, being bell-shaped with zero mean, but is lower in the middle portion and higher in the two tail regions than the normal curve. The exact dispersion of the t distribution depends on the number

of degrees of freedom, which is $n - 1$ for samples of size n. The areas under the t curve have been tabulated for various degrees of freedom at various significance levels.

Student's test. If the σ in a previous expression for u is replaced by the corresponding s, the resulting standardized unit is the t value, Eq. (16). To test the

$$t = \frac{(\bar{y} - m)}{\sqrt{s^2/n}} \tag{16}$$

hypothesis that the population mean is equal to m based on a sample of only $n = 14$ observations, where s^2 has 13 degrees of freedom, the absolute value of t has to be equal to or greater than 2.16 (instead of 1.96 for the u test) at the same $\alpha = 0.05$ significance level. Confidence intervals may be constructed in the same way as before, using the appropriate t value instead of the u value.

If the two population variances σ_1^2 and σ_2^2 are not equal, there is no simple method to test the significance of the difference $\bar{y}_1 - \bar{y}_2$ in comparing two sample means. If the two variances and the two sample sizes do not differ too much, it may be assumed that the population variances are equal, and a pooled estimate of the common variance from the two samples may be obtained by calculating Eq. (17). Then $t$can be expressed as Eq. (18), with degree of free-

$$s^2 = \frac{\Sigma(y_1 - \bar{y}_1)^2 + \Sigma(y_2 - \bar{y}_2)^2}{(n_1 - 1) + (n_2 - 1)} \tag{17}$$

$$t = \frac{(\bar{y}_1 - \bar{y}_2)}{s\sqrt{1/n_1 + 1/n_2}} \tag{18}$$

dom $= n_1 + n_2 - 2$. For example, if $n_1 = 15$ and $n_2 = 12$, so that the degree of freedom is 25, the absolute value of t has to be equal to or greater than 2.06 (instead of 1.96 for the u test) at the same $\alpha = 0.05$ significance level. In the special case where the observations of the two samples are paired, the n differences for the n pairs are then regarded as the n observations of one sample and the t value (with $n - 1$ degrees of freedom) may be calculated to test the null hypothesis that the true difference is zero. When the number of degrees of freedom is large, the t distribution becomes the u distribution.

Two-tail and one-tail tests. The examples given above for both u and t tests are two-tail tests because the significance of a difference is being tested, whether one is larger than the other or vice versa. If one is interested only in testing if $\bar{y}$ is greater than m, and uses $u = +1.645$ as the criterion, the probability of reaching a wrong rejection of the null hypothesis is $\alpha = 0.05$, which is the area under the right-hand tail of the normal curve. Testing the significance of difference in one direction only is called a one-tail test. Both types of tests are in common practice, depending upon the nature of the problem.

Analysis of variance. Analysis of variance is a method of comparing the differences between several sample means; it is a more general tool than the t test. This method is used most frequently in analyzing experimental data. Suppose there are k groups (or k samples), and in the first group there are n_1 observations with mean $\bar{y}_1$; in the second, n_2 observations with mean $\bar{y}_2$; and so on, the total number of observations being $n_1 + \cdots + n_k = N$ with the general mean y. If the k populations from which the k samples are derived have the same variance, a pooled estimate of the common variance σ^2 from the k samples is given by Eq. (19), which has $N - k$

$$s^2 = \frac{\Sigma(y_1 - \bar{y}_1)^2 + \cdots + \Sigma(y_k - \bar{y}_k)^2}{(n_1 - 1) + \cdots + (n_k - 1)} \tag{19}$$

degrees of freedom. On the null hypothesis that the k population means are equal, another estimate of the common variance may be obtained from the k sample means by calculating the quantity shown in Eq. (20), which has $k - 1$ degrees of freedom. The

$$s_m^2 = \frac{n_1(\bar{y}_1 - \bar{y})^2 + \cdots + n_k(\bar{y}_k - \bar{y})^2}{k - 1} \tag{20}$$

sum of squares in the numerator of s^2 and s_m^2 add up to the total sum of squares for the N observations as a whole $\Sigma\Sigma(y - \bar{y})^2$; and so the degrees of freedom in the denominators are $(N - k) + (k - 1) = N - 1$. If the variates are normally distributed, the ratio of two independent estimates of the same variance has a sampling distribution known as the F or variance-ratio distribution. Thus, to test the null hypothesis that the k population means are equal, the ratio $F = s^2{}_m / s^2$ is calculated, and the table is consulted for the F value with $k - 1$ degrees of freedom for the numerator and $N - k$ degrees of freedom for the denominator. If the observed F is greater than the tabulated F, the null hypothesis is discarded, and the significance of the differences between the k means is assumed. When there are only two samples, the F test is equivalent to the t test, for $F = t^2$ in such a case.

Linear regression. Very frequently the variate y under investigation is influenced by another (sometimes uncontrollable) variate x. For example, the yield y of a crop may be influenced by the amount of rainfall x during the growing season. When the pairs of values (y, x) are plotted in rectangular coordinates, the points form a scatter diagram (**Fig. 4**). If these points show a linear trend, a straight line running among these points may be fitted to represent the trend. Such a line is called the linear regression line. The problem is how to draw that line other than freehand.

Method of least squares. This is a method for fitting a straight line to a series of points. For the given value $x = x_i$, the observed $y = y_i$, while the y value on the straight line is $y' = y_i'$. Then $d = y - y'$ is the vertical distance from the observed y to the y' on the straight line. The method of least squares requires that the straight line should be drawn such that $\Sigma d^2 = \Sigma(y - y')^2$ is as small as possible. The equation of the straight line is $y' = a + bx$, where a is the intercept on the y axis and b, the slope of the line, is known as the regression coefficient of y on

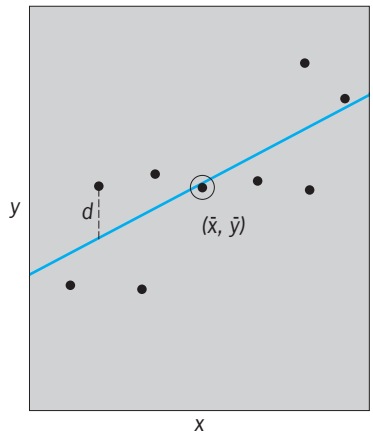

Fig. 4. Scatter diagram and linear regression.

x. The least-square solutions for a and b are derived from Eqs. (21) and (22), respectively. The regression

$$a = \bar{y} - b\bar{x} \tag{21}$$

$$b = \frac{\Sigma(y - \bar{y})(x - \bar{x})}{\Sigma(x - \bar{x})^2} \tag{22}$$

coefficient is in terms of the number of y units per x unit, for example 2 lb/in. This line passes through the central point $(\bar{y}, \bar{y})$.

Residual variation of y. Substituting these values of a and b in the straight-line equation, y' is obtained. Since these y' values are determined by the regression line, their variation is measured by $\Sigma(y' - \bar{y})^2$, with one degree of freedom accounted for by the variation of x. The deviation $d = y - y'$ from the regression line, the "residual," is regarded as the random error of y. Hence, the estimate of $\sigma^2 = \sigma^2_{y-x}$ free of x influence is expressed as Eq. (23). For practical

$$s^2 = s^2_{y \cdot x} = \Sigma(y - y')^2/(n - 2) \tag{23}$$

calculation, the individual yy' values and the individual distances d need not be obtained, because of the relationship in Eq. (24).

$$\begin{aligned} \Sigma d^2 &= \Sigma(y - y')^2 \\ &= \Sigma(y - \bar{y})^2 - b^2\Sigma(x - \bar{x})^2 \end{aligned} \tag{24}$$

Significance test for b. The sampling variance of b is $\sigma^2/\Sigma(x - \bar{x})^2$. To test the hypothesis that the true regression coefficient is β, the expression shown in Eq. (25) may be calculated. This follows Student's t

$$t = \frac{b - \beta}{\sqrt{s^2/\Sigma(x - \bar{x})^2}} \tag{25}$$

distribution with $n - 2$ degrees of freedom. To test the hypothesis that y is not linearly influenced by x, β is set equal to 0.

Analysis of covariance. This is a combined application of linear regression and an analysis of variance. Suppose that the effects of three fertilizers on corn yield y are studied. The significance of the treatment effects may be tested by the analysis of variance method. Now suppose that a complicating factor intervenes: the number of plants x, known as the concomitant variate in general, varies from plot to plot for reasons other than the treatments applied, so that the comparison of the yield between the three fertilizer groups may be partially blurred or distorted by the influence of x. To correct for the influences of x, the regression coefficient of y on x and the residual variation of y free from the influence of x are first found and then the F, or variance ratio, test is applied.

Correlation. A correlation problem considers the joint variation of two measurements, neither of which is restricted by the experimenter. Examples of correlation problems are studies of the relationship between school grades and IQ; yield and cornstalk height; and metabolism and blood pressure. Both variables are observed as they naturally occur, neither variable being set at prescribed levels. For n pairs of observed values (y, x), the quantity in Eq. (26) is defined as the correlation coefficient between y and x. This value ranges from $+1$ to -1. If the points in the scatter diagram are close to a straight line, r will be close to $+1$ or -1, depending on the direction of the line. If the points are scattered without a linear trend, r will be close to zero. The r and b coefficients are related and always have the same sign. This may be expressed by Eqs. (27) and (28). The quantity $\Sigma(y - \bar{y})(x - \bar{x})/(n - 1) = s_{xy}$

$$r = r_{xy} = \frac{\Sigma(y - \bar{y})(x - \bar{x})}{\sqrt{\Sigma(y - \bar{y})^2 \cdot \Sigma(x - \bar{x})^2}} \tag{26}$$

$$b = r\frac{s_y}{s_x} \tag{27}$$

$$r = b\frac{s_x}{s_y} \tag{28}$$

is called the covariance of y and x. Thus, $r = s_{xy}/s_x s_y$ for a sample, and the population correlation is defined as $\rho = \sigma_{xy}/\sigma_x\sigma_y$. Testing the null hypothesis $\rho = 0$ is equivalent to testing $\beta = 0$, and the same t test may be used. Or, from the relation $\Sigma(y - y')^2 = (1 - r^2)\Sigma(y - \bar{y})^2$, the test may be rewritten as Eq. (29). The definition of r, being symmetrical with

$$t = \frac{r}{\sqrt{(1 - r^2)/(n - 2)}} \tag{29}$$

respect to y and x, measures the mutual association of the two variates. The interpretation of r is another problem. The existence of a significant r value does not imply a causal relationship between y and x. Both variates may be influenced by a common factor or factors omitted from consideration in this discussion.

Path coefficients. These are regression coefficients with respect to standardized variates. When y and x are standardized, the regression equation of y on x will be $y = p_{y-x}x + e$, where p_{y-x} is called the path coefficient from x to y. In this simple case, p_{y-x} is simply r_{yx}, since the latter is the standardized covariance. One application of the path method is the

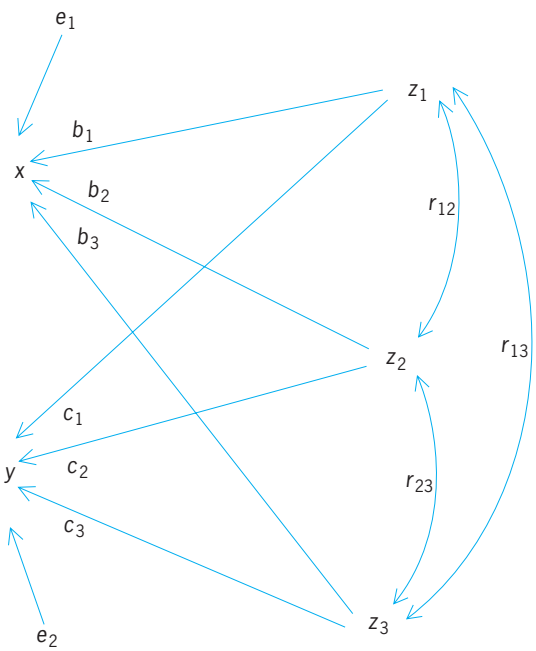

Fig. 5. Analysis of correlation of two variates, *x* and *y*, which are linear functions of the *e*'s and *z*'s. The *b*'s and *c*'s are path coefficients.

deduction of the chain property of correlation under certain simple conditions. Suppose that $y = p_1x + e_1$ where $x = p_2z + e_2$, the p's being path coefficients and the e's random errors. Substitution gives $y = p_1p_2z + e$. Hence, the path coefficient from z to y is the product of that from z to x and that from x to y. Also, $r_{yz} = r_{yx}r_{xz}$. Next, consider a standardized multiple regression equation $y = p_1z_1 + p_2z_2 + e$. Since every variate has unit variance, one obtains $1 = p^2{}_1 + p^2{}_2 + 2p_1p_2r_{12} + \text{Var}\,(e)$. The term p_1 is called the coefficient of determination of z_1 on y, and so on. Another application is the analysis of correlation between two variates which are linear functions of other common variables. Suppose that in **Fig. 5**, x and y were expressed as Eqs. (30) and (31),

$$x = b_1z_1 + b_2z_2 + b_3z_3 + e_1 \tag{30}$$

$$y = c_1z_1 + c_2z_2 + c_3z_3 + e_2 \tag{31}$$

respectively, where the e's are uncorrelated with z's or among themselves, and the b's and c's are path coefficients. The total correlation can then be written as Eq. (32),where for brevity r_{12} is the correlation be-

$$\begin{aligned} r_{xy} = \text{cov}(x, y) = {} & b_1c_1 + b_1r_{12}c_2 + b_1r_{13}c_3 \\ & + b_2c_2 + b_2r_{21}c_1 + b_2r_{23}c_3 \\ & + b_3c_3 + b_3r_{31}c_1 + b_3r_{32}c_2 \end{aligned} \tag{32}$$

tween z_1 and z_2 (cov indicates covariance). Thus the total correlation r_{yx} has been analyzed into nine components corresponding to the nine different ways x and y are connected in the diagram. A correlation line has two arrowheads and can be used in both directions. When the common causes (z_1, z_2, z_3) are uncorrelated, $r_{xy} = b_1c_1 + b_2c_2 + b_3c_3$. The general theorem is that the total correlation between two variates is the sum of all paths by which they are connected. To apply the method of path coefficients in practical problems the relationship of the variates must be specified prior to the evaluation of the various paths. The path method has been highly successful in certain types of genetical problems where the causal relationship is known. *See* POPULATION GENETICS.

Binomial population and sampling. In biological and medical research, variates are frequently dealt with that fall into either of two categories, for example, alive or dead, ill or well, male or female, and, generally, with or without a certain trait. This is known as a binomial variate. A complete description of a binomial population requires only one parameter, such as the proportion P of individuals with the trait; for then the proportion of those without the trait is $Q = 1 - P$. The statistical device of handling such a variate is to let $y = 1$ for individuals with the trait, and $y = 0$ for those without it. Then the population mean is $\mu = P \times 1 + Q \times 0 = P$, and the population variance is $\sigma^2 = P(1 - \mu)^2 + Q(0 - \mu)^2 = P(1 - P) = PQ$.

If a random sample of n observations is taken from a binomial population, the observations will take the form $y = 0, 1, 1, 0, 1, 0, 0, 0, 1, \ldots$. The sample total is a, the number of ones in the sample, and the sample mean is $\bar{y} = a/n = p$. The variance of $\bar{y}$ is σ^2/n, which may be estimated by pq/n. When n is large, the sampling distribution of $\bar{y}$ is nearly normal if P is not close to zero or unity. In determining how large a sample is "large," the usual working rule is that nP and nQ should be greater than 5, so the largeness of a sample depends on both n and P. For large samples, the test of significance follows the general procedure of the normal test. For example, to test the hypothesis that the population mean is P, Eq. (33) is used.

$$u = \frac{p - P}{\sqrt{PQ/n}} \tag{33}$$

The u test is appropriate here because the hypothesis of mean P implies the population variance PQ.

Suppose that there are two samples. The observed results are $a/n_1 = p_1$ for the first sample and $b/n_2 = p_2$ for the second. The null hypothesis $P_1 = P_2$ implies equal variance for the two populations, and the pooled estimate of the common variance is pq, where p equals the value of Eq. (34) and repre-

$$p = \frac{a + b}{n_1 + n_2} \tag{34}$$

sents the overall proportion from the two samples. Then the null hypothesis ($P_1 = P_2$) may be tested by calculating the normal deviate, Eq. (35). On the

$$u = \frac{p_1 - p_2}{\sqrt{pq(1/n_1 + 1/n_2)}} \tag{35}$$

other hand, if the hypothesis to be tested is $P_1 - P_2 = D$, then the u test should be modified, Eq. (36),

$$u = \frac{(p_1 - p_2) - D}{\sqrt{p_1q_1/n_1 + p_2q_2/n_2}} \tag{36}$$

since the two populations with different mean values necessarily have different variances.

For small samples from a binomial population, the sampling distribution is given by the expansion of $(P + Q)^n$. The exact probabilities of obtaining certain sampling results may be thus calculated. The values of these probabilities have been extensively tabulated for various values of P and n.

Poisson distribution. If P of a binomial population is very small, then the occurrence of a member with the trait is considered a rare event. A rare event can be observed only in a very large sample. If n is so large that $nP = m$ is some finite value, then the sampling distribution is given by the expansion of e^m in powers of m, each term being divided by e^m so that the total probability is unity:

$$e^{-m}\left\{1, m, \frac{m^2}{2!}, \frac{m^3}{3!}, \cdots\right\}$$

Many empirical observations of natural and experimental events conform with this distribution. The mean number of events is m and its variance is also m. When m is 5 or less, the distribution is skewed, but when it is larger, it becomes more symmetrical. In some fields of application, if the observed number of events is a, $\sqrt{a}$ is taken as the standard error of the observed number.

Chi-square test. The quantity χ^2 is the sum of several independent squared normal deviates. In the analysis of frequency data, the χ^2 test is a generalization of the previous test for binomial data in two respects. (1) It extends to a multinomial population where each observation may fall into one of several classes. (2) It furnishes a comparison between several samples instead of only two. The χ^2, as a generalization of u, is analogous to F as a generalization of t.

Calculation. The calculation of χ^2 for testing the goodness of fit may be illustrated as shown in **Table 3**. According to a certain genetic theory, the breeding should yield 9/16 red, 6/16 sandy, 1/16 white. To test this theory, the number of pigs is divided into three classes in the proportions 9:6:1. Then the value of χ^2 is as shown in Eq. (37). If there

$$\chi^2 = \sum \frac{(f - f')^2}{f'} = \frac{25}{63} + \frac{49}{42} + \frac{4}{7} = 2.135 \qquad (37)$$

are k classes, the χ^2 has $k - 1$ degrees of freedom because the total number of observations is the same for the two sets of frequencies. In this example, degrees of freedom $= 3 - 1 = 2$. If the genetic hypothesis is true, the probability that the χ^2 with two degrees of freedom should exceed 5.99 (read from χ^2 table) is 0.05. The statistical reasoning for the test of significance is the same as before. Since the χ^2 in question is smaller than 5.99, the hypothesis is retained, the observed results not being contradictory to the genetic theory. The χ^2 test is a one-tail test. It does not prove that the theory is the correct one. It is conceivable that the observed results are also consistent with some other hypothesis. However, as the class frequencies become larger, the test becomes more discriminating. When an expected class frequency is 5 or less, the χ^2 test is not applicable, and a more exact method, if available, should be employed. Several small classes may be pooled to form one larger class.

TABLE 3. Data for χ^2 calculation

Color of pig	Observed number, f	Expected number, f'	Difference, $f - f'$
Red	58	63	–5
Sandy	49	42	7
White	5	7	–2
Total	112	112	0

TABLE 4. Contingency table

Category	Inoculated	Not inoculated	Total
Dead	a	b	$a + b$
Alive	c	d	$c + d$
Total	$a + c = n_1$	$b + d = n_2$	$a + b + c + d = N$

Contingency table (2 × 2). If the observed frequencies are classified dichotomously according to two different characteristics simultaneously, as exemplified by the fourfold **Table 4**, it may often be desirable to know if the two characteristics are independent of each other. If the characteristics are independent, the proportions of inoculated and not inoculated should be the same among the dead and the alive, or, stated alternatively, the proportions of dead and alive should be the same among those inoculated and those not inoculated. In such a case, the expected frequency in the inoculated dead cell should be $a' = (a + b) \times (a + c)/N$, and so on. Then $\chi^2 = (a - s')^2/a' - \cdots$ which simplifies to $\chi^2 = N(ad - bc)^2/(a + b)(c + d)(a + c)(b + d)$. Since the expected frequencies are calculated on the basis of the marginal totals, this χ^2 has only one degree of freedom. At the 0.05 significance level and for 1 degree of freedom the critical value is $\chi^2 = u^2 = (1.96)^2 = 3.84$. The χ^2 test for independence in a 2×2 table is equivalent to the u test for the hypothesis $P_1 = P_2$ for two binomial samples.

The test for independence may be extended to contingency tables of any size, for example m rows times k columns. The expected frequency of the cell in the third row and the second column, for example, is equal to third row total times second column total over the grand total. The χ^2 may then be calculated in the usual manner. It has $(m - 1) - (k - 1)$ degrees of freedom. For a $2 \times k$ table, the χ^2 test is equivalent to testing the hypothesis that the means of the k binomial populations are equal.

Maximum likelihood method. There are many methods of estimating a parameter which underlies the probability of an event. The most efficient method

TABLE 5. Estimation by maximum likelihood method

Category	Probability	Observed number of occurrences
1st trial success; no 2d trial	p	a
1st trial failure; 2d success	qp	b
1st and 2d trials failure	q^2	c
Total	1	n

of estimating is known as the maximum likelihood method, the principle of which may be illustrated by the following example.

Suppose that p is the (unknown) probability of success (and $q = 1 - p$) in a trial (for example, hitting a ball, recovery from an illness, or germination of a seed), and that the individual is allowed a second trial if the first one is a failure. The observed frequencies for the three possible outcomes are a, b, c, as shown in **Table 5**. The principle of maximum likelihood says that the value of p should be so chosen that the probability of observing such a set of frequencies is a maximum. The probability of observing such a set of occurrences (Table 5) is expressed as Eq. (38).

$$P(S) = \text{const}(p)^a(qp)^b(q^2)^c \tag{38}$$

This expression without the constant is called "likelihood," the logarithm of which is calculated from Eqs. (39*a*) and (39*b*).

$$L = (a + b)\log p + (b + 2c)\log q \tag{39a}$$

$$\frac{dL}{dp} = \frac{a+b}{p} - \frac{b+2c}{q} \tag{39b}$$

The value of p that maximizes the likelihood (and hence its logarithm L) is given by the solution of the equation $dL/dp = 0$. Solving, $\hat{p}$ gives the value shown in Eq. (40), where $T = a + 1b + 2c$ is the total

$$\hat{p} = \frac{a+b}{a+2b+2c} = \frac{a+b}{T} \tag{40}$$

number of trials, of which $a + b$ are successful ones. The variance $V(\hat{p})$ is given by Eq. (41). The chief

$$\frac{1}{V(\hat{p})} = \frac{-d^2L}{dp^2} = \frac{a+b}{p^2} + \frac{b+2c}{q^2} = \frac{T}{pq} \tag{41}$$

drawback of this powerful method is that, in most cases, and especially when they are two or more parameters to be estimated, explicit expressions for the estimates are unobtainable, and for a solution, one must resort to iteration methods. Calculation of the variances is even more tedious.

Transformation. The four statistical tests u, t, χ^2, F are all based on the normal distribution. If the variate y is not normally distributed, often a change of scale renders the new variate $z = f(y)$ approximately normal, so that the standard tests may be employed on the new variate. Such a change of scale is called a transformation of variables. There is no rule of thumb as to how a variate scale should be changed to render it normal and to equalize the variances of two or more samples. The following are some of the commonly used transformations.

Logarithmic. Probably the most frequently used transformation is $z = \log y$ or $\log(y + c)$. If the new variate z is normal, then y is said to be log-normal, a special name for the original distribution of y. It may be shown that if the measurement y is influenced by a large number of factors acting multiplicatively (instead of additively), then y has a log-normal distribution. A logarithmic transformation also changes the original exponential relationship with another variable into a linear one which is much easier to handle statistically.

Angular. For samples from a binomial population, the mean $y = p$ and its variance pq/n are correlated The p's of several samples cannot be averaged in the usual manner because each has a different variance. The angular transformation is to replace the proportion p by an angle θ such that $p = \sin^2\theta$. The variance of θ depends on n only. Then the angular values can be treated by standard statistical methods. The square-root transformation accomplishes the same purpose for a Poisson distribution.

Probit. This is the normal deviate plus five $(u + 5)$ to avoid negative numbers. Suppose that $p = 12/20 = 0.65$ is the observed proportion of deaths in a group of 20 animals treated with a certain dose of drug. The assumption is that the individual lethal dose level is normally distributed. Then the area $p = 0.65$ under the left side of the normal curve is set off by an ordinate at $u = 0.385$ or a probit value 5.385. The probit transformation thus replaces a proportion (area) by its corresponding normal deviate (abscissa value) plus 5. The reason for doing this is that the mortality curve for proportions against dosage is sigmoid in shape, but the mortality curve for probits against dosage is linear. (The dosage may be replaced by log dosage, but this is not a part of the probit transformation.) Probits are widely used in toxicology, biological assay, and similar problems where response is of the all-or-none type. *See* ANALYSIS OF VARIANCE; HUMAN GENETICS; STATISTICS. C. C. Li

Bibliography. C. C. Li, *Introduction to Experimental Statistics*, 1964; G. W. Snedecor and W. G. Cochran, *Statistical Methods*, 8th ed., 1989; R. Sokal and F. Rohlf, *Biometry: The Principles and Practice of Statistics in Biological Research*, 3d ed., 1994.

Bioorganic chemistry

The science that describes the structure, interactions, and reactions of organic compounds of biological significance at the molecular level. It represents the meeting of biochemistry, as attempts are made to describe the structure and physiology of organisms on an ever smaller scale, with organic chemistry, as attempts are made to synthesize and understand the behavior of molecules of ever-increasing size and complexity. Areas of research include enzymatic catalysis, the structure and folding of proteins, the structure and function of biological membranes, the chemistry of poly(ribonucleic acids)

and poly(deoxyribonucleic acids), biosynthetic pathways, immunology, and mechanisms of drug action.

Being at the interface of two disciplines, bioorganic chemistry utilizes experimental techniques and theoretical concepts drawn from both. Important experimental techniques include organic synthesis, kinetics, structure-activity relationships, the use of model systems, methods of protein purification and manipulation, genetic mutation, cloning and overexpression (engineered enhancement of gene transcription), and the elicitation of monoclonal antibodies. Theoretical concepts important to bioorganic chemistry include thermodynamics, transition-state theory, acid-base theory, concepts of hydrophobicity and hydrophilicity, theories of stereocontrol, and theories of adaptation of organisms to selective pressures.

Enzyme catalysis. Historically, a major focus of bioorganic research has been the study of catalysis by enzymes. Enzymes are proteins that are responsible for catalyzing nearly every chemical reaction that occurs in living systems. Without enzymes, the reactions that support living organisms would occur too slowly to maintain life.

Mechanisms of rate increase. Enzymes have a dramatic ability to increase the rates at which reactions occur. Enzyme-catalyzed reactions typically proceed 10^6–10^{12} times faster than they would otherwise. Enzymes accomplish these enormous rate increases for a wide variety of reactions, despite the fact that most biological reactions occur between 0 and 38°C (32 and 100°F) and at essentially neutral pH.

One of the ways that enzymes increase the rates of bimolecular reactions is to overcome the entropic barrier associated with bringing two particles together to form one. At 25°C (77°F), the loss of entropy experienced by two molecules combining to form one transition state accounts for an energetic barrier of 10–11 kcal mol^{-1}. Overcoming this entropic barrier could increase the rate of a reaction by a factor of 10^8. Enzymes use favorable noncovalent interactions (ionic, dipole-dipole, hydrogen bonding, or hydrophobic) to bind the two reactants in proximity and in the right orientation for a reaction to occur. Thus, the reaction is said to occur intramolecularly. Although the reaction is strictly intermolecular, the entropy of the reaction is like that of an intramolecular reaction because the reaction occurs within a single molecular complex. Intramolecular reactions occur much more rapidly than similar intermolecular reactions. For example, the rate of acid-catalyzed lactonization of hydroxyacid [(**1**) in reaction (1)] is

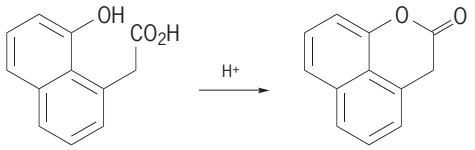

(1)

(**1**)

5.3×10^8 times faster than the rate of ester formation by phenol and acetic acid [reaction (2)].

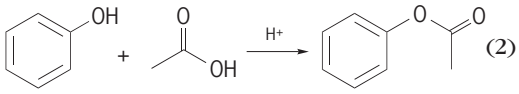

(2)

See CHEMICAL DYNAMICS.

Enzymes also catalyze reactions by facilitating proton transfers. Many of the reactions catalyzed by enzymes, such as the formation and hydrolysis of esters and amides, require the deprotonation of a nucleophile (base catalysis) or the protonation of an electrophile (acid catalysis). However, enzymes function in an aqueous environment that is thermodynamically essentially neutral. Enzymes overcome this limitation by using their own amino acid side-chain functional groups to donate or accept protons as necessary. These proton-transferring groups are called general acids and general bases, respectively. They interact with enzyme substrates so that proton transfers occur simultaneously with the cleavage or formation of bonds between heavy (nonhydrogen) atoms. Thus, no large equilibrium concentration of hydronium or hydroxide ion is required. *See* ACID AND BASE.

Intramolecular general acid catalysis by the carboxylic acid group [reaction (3)] causes acetal (**2**)

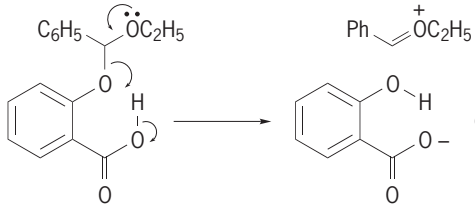

(3)

(**2**)

to hydrolyze 10^7 times faster than the related acetal [(**3**) in reaction (4)] at pH 7.0. In compound (**4**), the

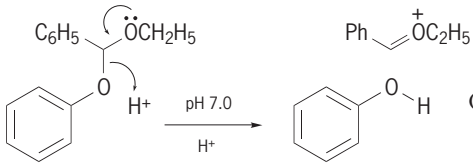

(4)

(**3**)

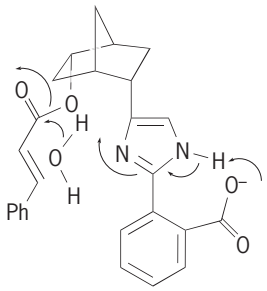

(**4**)

general base-catalyzed ester hydrolysis is 10^5 times faster at pH 7.0 than the uncatalyzed hydrolysis. The carboxylate and imidazole groups in compound (**4**) mimic the charge relay system composed of aspartate and histidine side chains in the active site of the proteolytic enzyme α-chymotrypsin.

Enzymes bind metal ions in their active sites by coordinating them to the side-chain thiol groups of cysteine or imidazole groups of histidine. The metal ions catalyze reactions by acting as Lewis acids or as sources of nucleophilic hydroxide ion. As Lewis acids, metal ions coordinate substrate molecules, making them more electrophilic or easier to deprotonate. Compound (**5**) is a model of the active site

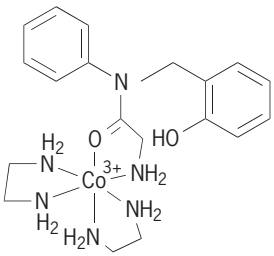

(**5**)

of the proteolytic enzyme, carboxypeptidase A. Electrophilic coordination of the amide carbonyl group by the cobalt(III) cation (Co^{3+}) makes the amide more susceptible to hydrolysis than it is in the absence of cobalt(III).

Metal ions can also bind water molecules to their vacant coordination sites. Metal-bound water is more acidic and thus dissociates more readily to form hydroxide ion than unbound water. Metal-bound hydroxide ion, however, is nearly as nucleophilic as free hydroxide ion. Metals thus serve to increase the local concentrations of hydroxide ion in the active sites of enzymes. In compound (**6**), the cobalt(III)-bound hy-

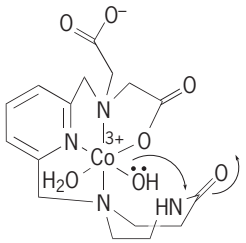

(**6**)

droxide ion attacks the amide carbonyl group, causing hydrolysis to occur 5×10^8 times faster than it does in the absence of cobalt. *See* COORDINATION CHEMISTRY.

The transition state of a chemical reaction has a fleeting existence. Its lifetime is shorter than the time required for a single molecular vibration ($\leq 10^{-13}$ s). Furthermore, the structure of a transition state (shape, bond lengths, extent of charge separation) always differs from that of a ground-state reactant. In solution, a transition state is always less than optimally solvated, because solvent molecules cannot reorganize quickly enough to accommodate the shape, size, and dipole moments of the transition state. Enzymes, however, may act as preorganized solvation shells for transition states. If the active site of an enzyme has just the right size, shape, and arrangement of functional groups to bind a transition state, it will automatically bind the reactants less well. Selective binding of the transition state lowers the energy of the transition state relative to that of the reactants. Because the energy barrier between reactants and transition state is reduced, the reaction proceeds more rapidly.

Experiments with monoclonal antibodies have demonstrated that catalysis by selective binding of transition states is possible. Antibodies are proteins that can bind small molecules tightly. Unlike enzyme proteins, antibodies usually show no catalytic behavior. However, monoclonal antibodies that have been raised against ground-state molecules that "look" like transition states of reactions can act as catalysts. A monoclonal antibody raised against phosphonate ester (**7**) increases the rate of hydrolysis of cocaine to

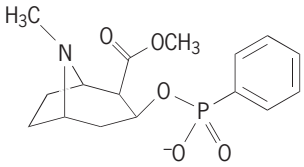

(**7**)

ecgonine methyl ester 540-fold. In this case, compound (**7**) resembles the transition state leading to the tetrahedral intermediate formed during hydrolysis of the carboxylic ester in cocaine [reaction (5)].

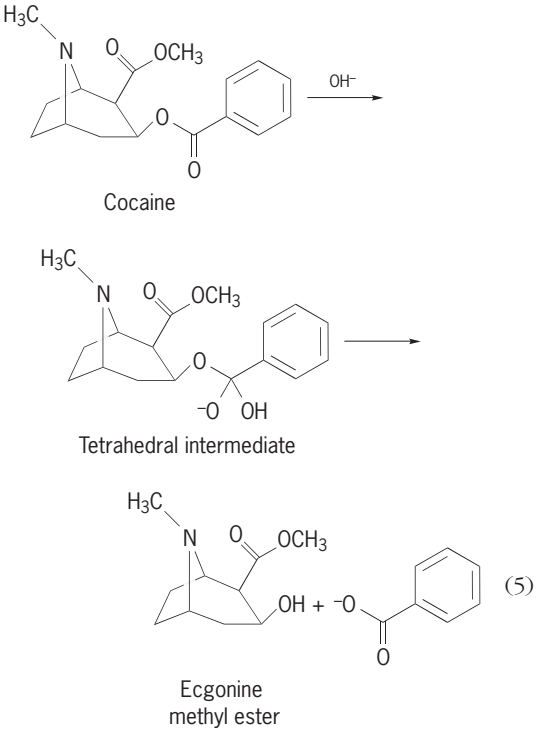

(5)

The antibody that binds compound (**7**) tightly also binds the transition state for hydrolysis of cocaine tightly. *See* MONOCLONAL ANTIBODIES.

This example also illustrates two important features of catalytic antibodies. First, by appropriate choice of a "transition-state mimic," it is possible to create custom-designed catalysts. Second, some of these catalysts may be therapeutically useful. Because ecgonine methyl ester lacks the pharmacological activity of cocaine, the monoclonal antibody

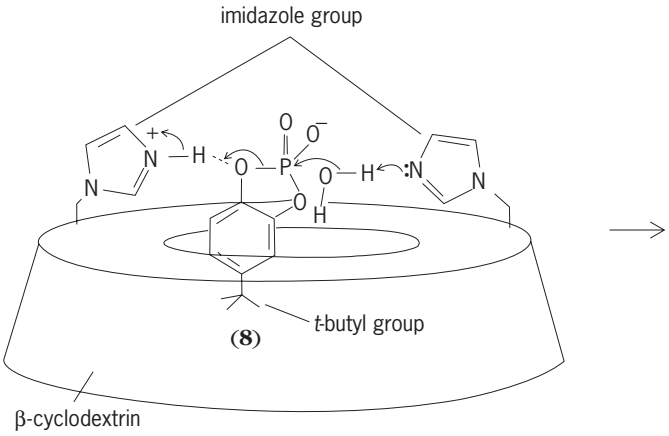

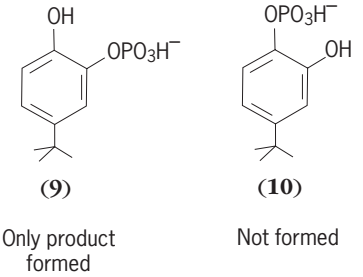

(6)

catalyzing the hydrolysis of cocaine may be useful as a source of passive immunity for those who are trying to overcome addiction to cocaine.

Selectivity of enzymatic catalysis. Another property of enzymes is the tremendous selectivity with which they catalyze reactions. For example, fumarase catalyzes the addition of water across the double bond of fumaric acid to form L-malic acid. It does not catalyze the addition of nucleophiles other than water to fumaric acid, nor does it catalyze the hydration of compounds that are structurally similar to fumaric acid. The solvent-derived proton adds to the carbon-carbon double bond of fumaric acid from the face opposite to that to which the hydroxyl group adds (the addition is diastereoselective), and the hydroxyl group adds to only one of the two enantiotopic faces of fumaric acid, forming only the L enantiomer of malic acid (the addition is enantioselective).

Few synthetic catalysts exhibit enzymelike selectivities. However, the hydrolysis of regioselective phosphodiester (**8**) catalyzed by a derivatized cyclodextrin [reaction (6)] is an example of the mechanism by which selective catalysis occurs. β-Cyclodextrin (the cyclic α-1,4-linked heptamer of glucose) can be viewed as a large molecular "doughnut" with a hydrophobic interior (**8**). When imidazole groups are attached to the first and fourth glucose residues of β-cyclodextrin, they act as general acid and base catalysts for the hydrolysis of compound (**8**). This hydrolysis mimics the cleavage of the phosphodiester linkages of ribonucleic acid (RNA) by the enzyme ribonuclease (RNase). Because the *t*-butyl group of compound (**8**) first binds inside the hydrophobic cleft of the cyclodextrin, the interaction with the imidazole catalysts is regioselective, leading to the formation of compound (**9**), not compound (**10**), as the only product of hydrolysis. The key to selective catalysis is thus the binding of specific substrates in a unique orientation so that they interact selectively with functional groups appropriately positioned within the active site of the catalyst. *See* RIBONUCLEIC ACID (RNA).

Enzymes in organic synthesis. The selectivity of enzymes makes them useful as catalysts for organic synthesis. Surprisingly, enzymes are able to catalyze not only the reactions that they mediate in living systems but also similar, selective transformations of unnatural substrates. Enzyme-catalyzed reactions that have been scaled up as industrial processes include the conversion of corn starch to high-fructose corn syrup [reaction (7)], the transamidation of porcine insulin to form human insulin [reaction (8)], and the kinetic chiral resolution of glycidol [reaction (9)].

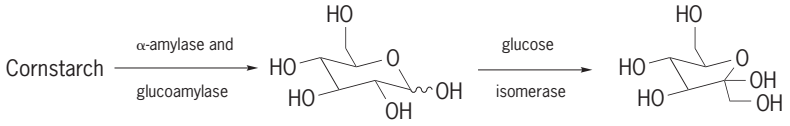

(7)

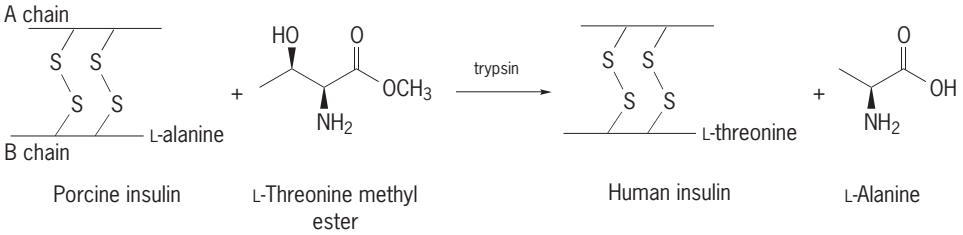

(8)

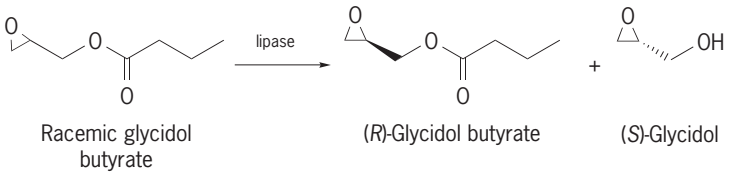

(9)

The first process utilizes enzymes simply to act upon their natural substrates. The latter two processes, however, require enzymes to act upon unnatural substrates that would not be encountered by the enzymes in the living system.

The application of enzymes to organic synthesis is made possible by advances in genetic cloning and overexpression, protein purification, and enzyme immobilization. Immobilization of enzymes on solid, insoluble supports or on soluble polymers allows the enzymes to be retained within flow-through reactors and to be recovered for reuse. Immobilization often has the added benefit of increasing the stability of the enzyme protein.

Enzymes can catalyze reactions in organic solvents as well as in water. Changing the solvent from water to an organic solvent often has profound effects on the equilibria of reactions. Because the solvent affects the conformation of an enzyme and thus the shape of its active site, a change in solvent may also change the substrate specificity or stereoselectivity of an enzyme.

Because of their selectivity, several enzyme-catalyzed reactions may run simultaneously in the same vessel. Thus, a reactant can undergo several reactions in series without the need for isolation of intermediates. Enzymes can be combined so as to reconstitute within a reaction vessel naturally occurring metabolic pathways or to create new, artificial metabolic pathways. Sequences of up to 12 serial reactions have been executed successfully in a single reaction vessel. *See* CATALYSIS; ENZYME; ORGANIC SYNTHESIS.

Protein folding. The biological activity of a protein, whether binding, catalytic, or structural, depends on its full three-dimensional or conformational structure. The linear sequence of amino acids that make up a protein constitutes its primary structure. Local regions of highly organized conformation (α-helices, β-pleats, β-turns, and so on) are called secondary structure. Further folding of the protein causes regions of secondary structure to associate or come into correct alignment. This action establishes the tertiary structure of the native (active) protein. Sometimes, covalent disulfide bonds between cysteine side chains help to "lock in" the correct tertiary structure.

A goal of bioorganic chemistry is to achieve an understanding of the process of protein folding. The primary structure of a protein contains all of the information necessary to determine its tertiary structure in the native state. Some proteins, after being unfolded by heat or chemical treatment, spontaneously refold into their native conformation. Others, however, do not. Some proteins may fold up as they are coming off the ribosome during synthesis into a conformation that differs from the one that they would adopt if the fully formed protein were allowed to fold. In fact, the native conformation of a protein may be only a metastable state and not its thermodynamically most stable conformation. Some experiments indicate that regions of secondary structure form spontaneously and subsequently control the formation of tertiary structure. Other experiments have shown that hydrophobic interactions between nascent α-helices facilitate helix formation. *See* RIBOSOMES.

Necessary to an understanding of protein folding is a knowledge of the factors that stabilize secondary and tertiary structures. Certain amino acids, such as alanine and valine, are known to be α-helix stabilizing. Proline promotes β-turn formation. Bonding interactions between amino acid side chains on successive turns of α-helices have been used to "crosslink" the α-helices in unnatural peptides. Both noncovalent (hydrogen bonding, ionic, or metal chelation) and covalent (disulfide or amide bond) interactions have been used. Nonpeptide prosthetic groups having rigid conformations have been incorporated into peptides and found to nucleate secondary structure formation. Attempts to stabilize tertiary structures by introduction of genetically engineered disulfide bonds have, in general, been unsuccessful. *See* AMINO ACIDS; PROTEIN.

DNA binding and cleavage. Understanding mechanisms of DNA binding and cleavage is important to understanding the activities of many antibiotics and to the development of new therapeutic agents. The design of agents that would cleave DNA selectively at specific nucleotide sequences would also benefit genetic research and molecular biology. Compounds that bind DNA usually interact with exposed nucleotide bases and often bind to either the major or the minor groove of double-stranded α-helical deoxyribonucleic acid (DNA). *See* ANTIBIOTIC; DEOXYRIBONUCLEIC ACID (DNA).

Mechanisms of drug action. A major success of bioorganic chemistry is its ability to elucidate the chemical mechanisms responsible for the biological activities of drugs. Invariably, drugs react by the same mechanisms by which other organic molecules react. An example is the mechanism of action of the enediyne antibiotics. Calicheamicin, esperamicin, and dynamicin are antitumor antibiotics that possess at their core an enediyne structure. In the case of calicheamicin (**11**), reductive or nucleophilic

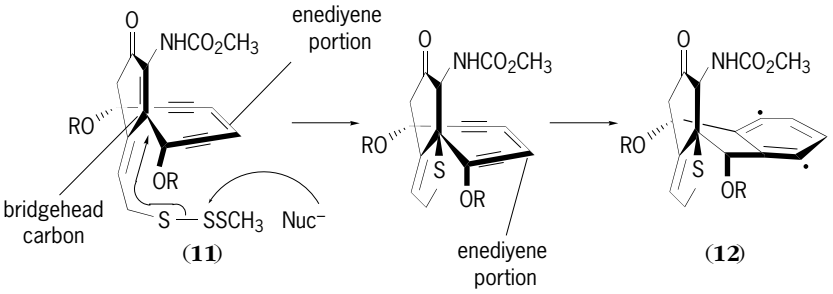

(10)

cleavage of the pendant trisulfide by a reductant or nucleophile (shown as Nuc^- and a curved arrow) leads to a Michael addition at a bridgehead carbon. A Michael addition is the addition of a nucleophile to one end of a carbon-carbon double bond that has an electron-withdrawing double bond attached at the other end. In this case, the C═O attached to the C═C is the electron-withdrawing double bond.

Rehybridization of the bridgehead carbon allows the enediyne to undergo ring closure, generating a benzendiyl diradical [(**12**) in reaction (10)]. If the antibiotic is bound to DNA, the diradical can abstract a hydrogen atom from the sugar phosphate backbone, leading to DNA strand scission. Although the first enediyne antibiotic was discovered in 1987, the mechanism of enediyne cyclization to form a benzenediyl diradical was elucidated in 1972. *See* BIOCHEMISTRY; ORGANIC CHEMISTRY.

H. Keith Chenault

Bibliography. R. Breslow (ed.), International Symposium on Bioorganic Chemistry, *Ann. N.Y. Acad. Sci.*, vol. 471, 1986; R. Jaenicke, Protein folding: Local structures, domains, subunits, and assemblies, *Biochemistry,* 30:3147–3161, 1991; A. J. Kirby, Effective molarities for intramolecular reactions, *Adv. Phys. Org. Chem.*, 17:183–278, 1980; R. A. Lerner, S. J. Benkovic, and P. G. Shultz, At the crossroads of chemistry and immunology: Catalytic antibodies, *Science,* 252:659–667, 1991; J. F. Liebman and A. Greenberg (eds.), *Mechanistic Principles of Enzyme Activity,* 1988; G. M. Whitesides and C.-H. Wong, Enzymes as catalysts in synthetic organic chemistry, *Angew. Chem. Int. Ed. Engl.,* 24:617–638, 1985.

Biophysics

A hybrid science involving the overlap of physics, chemistry, and biology. A dominant aspect of biophysics is the use of the ideas and methods of physics and chemistry to study and explain the structures of living organisms and the mechanisms of life processes.

Many eminent scientists of previous generations, such as A. Volta, H. von Helmholtz, and J. L. M. Poiseuille, were really biophysicists because they made important contributions to both biology and physics. Helmholtz and his contemporaries explicitly planned to explain biological phenomena in terms of the physical sciences. Nevertheless, the recognition of biophysics as a separate field is relatively recent, having been brought about, in part, by the invention of physical tools such as the electron microscope, the ultracentrifuge, and the electronic amplifier, which greatly facilitate biophysical research. These tools are peculiarly adapted to the study of problems of great current importance to medicine, problems related to virus diseases, cancer, heart disease, and the like.

Furthermore, with the advent of the space and atomic age, new problems of survival, which involve hitherto unthought-of reactions with the physical environment, have been thrust upon humans. Study of biological processes at every level, molecular or macroscopic, continues to reveal the complex physical and chemical basis of life, demanding the development of physical thought for the special task of interpreting physical measurements on living systems. As a result, biophysics departments and research programs in many universities, which provide opportunities for study at undergraduate, graduate, and postdoctoral levels, were established, and the Biophysical Society was organized.

There are several areas within biophysics, the major ones being molecular biophysics, radiation biophysics, physiological biophysics, and mathematical or theoretical biophysics.

Molecular biophysics has to do with the study of large molecules and particles of comparable size which play important roles in biology. Favorite physical tools for such research are the electron microscope, the ultracentrifuge, and the x-ray diffraction camera. Conspicuous success has been achieved in the study of viruses and in the investigation of nucleic acid, the substance which is involved in hereditary mechanisms. *See* ELECTRON MICROSCOPE; MOLECULAR BIOLOGY; ULTRACENTRIFUGE; X-RAY DIFFRACTION.

Radiation biophysics consists of the study of the response of organisms to ionizing radiations, such as alpha particles, beta particles, gamma rays, and x-rays, and to ultraviolet light. The biological responses are death of cells and tissues, if not of whole organisms, and mutation, either somatic or genetic. Somatic mutations represent changes in nonreproductive cells and might lead to uncontrolled or cancerous growth. Genetic mutations involve changes in the germ cells which can be transmitted to future generations. *See* MUTATION; RADIATION BIOLOGY.

Physiological biophysics, called by some classical biophysics, is concerned with the use of physical mechanisms to explain the behavior and the functioning of living organisms or parts of living organisms and with the response of living organisms to physical forces. The questions of how nerve impulses are transmitted; what the mechanism of muscle contraction is; what the visual mechanism is; how one hears, tastes, smells, or feels; and why living cells are able to ingest some substances and excrete others have been studied and, in part at least, answered by physiological biophysicists. Closely related matters are the changes in normal functioning of organisms, including the human body, when the physical environment is changed. Of particular importance in the space age are the changes brought about in the human body by subjecting the individual to high accelerations.

Mathematical and theoretical biophysics deals primarily with the attempt to explain the behavior of living organisms on the basis of mathematics and physical theory. Biological processes are being examined in terms of thermodynamics, hydrodynamics, and statistical mechanics. Mathematical models are being investigated to see how closely they simulate biological processes. *See* BIOMECHANICS; BIOPOTENTIALS AND IONIC CURRENTS; BONE; MATHEMATICAL

BIOLOGY; MICROMANIPULATION; MICROSCOPE; MUSCLE; MUSCULAR SYSTEM; OXIMETRY; SKELETAL SYSTEM; SPACE BIOLOGY; THERMOREGULATION; THERMOTHERAPY.

Max A. Lauffer

Biopolymer

A macromolecule derived from natural sources; also known as biological polymer. Nature has developed effective means for efficiently preparing and utilizing biopolymers. Some are used as structural materials, food sources, or catalysts. Others have evolved as entities for information storage and transfer. Familiar examples of biopolymers include polypeptides, polysaccharides, and polymers derived from ribonucleic acid (RNA) and deoxyribonucleic acid (DNA). Other examples include polyhydroxybutyrates, a class of polyesters produced by certain bacteria, and cis-1,4-polyisoprene, the major component of rubber tree latex.

Polypeptides. Amino acids are the monomers from which polypeptides (**1**) are derived. Only one of the

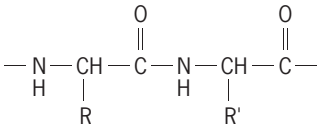

(**1**)

two enantiomers of amino acids, the S or levo isomer, is used to make polypeptides. Polypeptides may be polymers of a single amino acid, but far more commonly they are polymers of many (up to about 20) amino acids arranged in a specific sequence in the polymer chain. This specific sequence is called the primary structure. The side chains (R's) of the amino acids differ in their size and charge, and their placement in a chain determines how the chain will interact with itself (intramolecularly) and with other chains (intermolecularly). For example, if the side chain is simply a hydrogen atom, a polypeptide will tend to form extended sheets (a so-called β-pleated sheet) via intermolecular hydrogen bonding between amides. If the side chain is large, steric interactions force hydrogen bonding intramolecularly, leading to a helix (more formally termed an α-helix). The extended sheets and helices are examples of secondary structure, the regular, recurring arrangement in space of the polypeptide chain in one dimension. *See* AMINO ACIDS; HYDROGEN BOND; STEREOCHEMISTRY.

Many polypeptides, however, fold into three-dimensional structures through association of different parts of the chain, and this arrangement is known as tertiary structure. Polypeptide chain folding occurs by association of oppositely charged side chains or association of nonpolar (hydrophobic) side chains that tend to be clustered rather than individually surrounded by polar water molecules. In many cases, tertiary structures are held in place through the formation of covalent disulfide (—S—S—) by reaction of a pair of thiol (—S—H) side chains.

More than one polypeptide chain can associate through primarily noncovalent association to yield quaternary structure. Polypeptides alone, as well as multipolypeptide complexes or complexes with other molecules, are known as proteins, and each has a specific biological function. Examples of proteins are enzymes, which are highly efficient and selective catalysts for many chemical reactions, and collagen, the fibrous material of skin and cartilage. *See* CATALYSIS; COLLAGEN; ENZYME; PROTEIN.

Polysaccharides. In polysaccharides, sugar units, predominantly glucose (**2**), are linked together by acetal (glycoside) bonds. Amylose is a polysaccharide having α-1,4 linkages between repeat units (**3**). It is

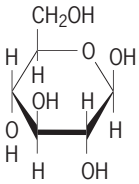

(**2**)

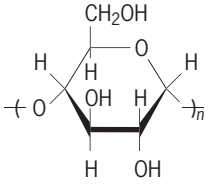

(**3**)

a component of starch, the other component being amylopectin, a branched version of amylose. Glycogen is another important polysaccharide and is even more branched than amylopectin. Both starch (in plants) and glycogen (in mammals) serve as sources of glucose, the food used to produce adenosine triphosphate (ATP), the energy source that promotes many biological reactions. An interesting contrast is provided by cellulose (**4**), which is a linear

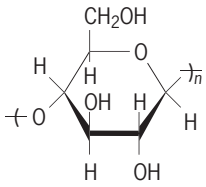

(**4**)

homopolymer of glucose but with exclusively β-1,4 linkages. Cellulose is a water-insoluble, structural polymer found in the cell walls of plants. The β-1,4 linkages encourage an extended-chain conformation and aggregation into fibrils. The α-1,4 linkages of glycogen, however, favor a coiled helix. Besides these differences, cellulose cannot be metabolized by humans, and therefore it is not a useful energy source. However, ruminants such as cattle have in their digestive tracts bacteria that produce the enzymes necessary to catalyze the breakdown of cellulose into glucose. *See* ADENOSINE TRIPHOSPHATE (ATP); CARBOHYDRATE; CELLULOSE; GLUCOSE; GLYCOGEN; GLYCOSIDE; POLYSACCHARIDE; STARCH.

Polypeptide/polysaccharide hybrids. There are numerous examples of biopolymers having a polysaccharide and polypeptide in the same molecule, usually with a polysaccharide as a side chain in a polypeptide, or vice versa. Covalent bonds between the polysaccharides and polypeptides are made via a

glycoside bond to side chains of the latter, which are usually serine, threonine, or asparagine. Examples are proteoglycans, which are primarily polysaccharides, and glycoproteins, which are primarily polypeptides. Proteoglycans are frequently found in connective tissue. Many of the hydroxyls of the sugars are sulfated, and hence these molecules are highly anionic (negatively charged). Several glycoproteins are known wherein the protein part of the molecule embeds itself in a cell membrane, with the polysaccharide (glyco) part projecting outward from the membrane surface. These sugar residues act as signals to the immune system that identify the type of cell with which they are associated. In this way, foreign cells are recognized, and the immune system is activated. Proteoglycans are frequently aggregated into network structures, an example of which is a bacterial cell wall. Here polysaccharide chains form a gridlike network crosslinked by short peptide segments. *See* BACTERIAL PHYSIOLOGY AND METABOLISM.

DNA and RNA. The macromolecules DNA (**5**) and RNA (**6**) are polymers of nucleotides, which com-

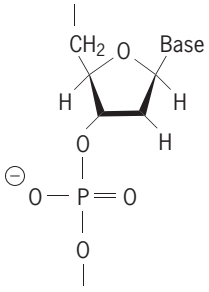

(**5**)

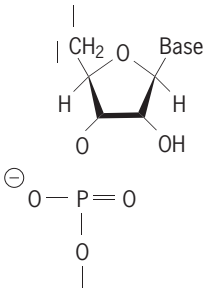

(**6**)

prise a sugar, a phosphate group, and an appendage referred to as a base. The bases are either purines, such as adenine (**7**), or pyrimidines, such as thymine (**8**). Intermolecular hydrogen bonding between

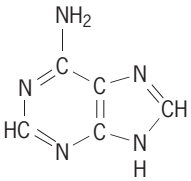

(**7**)

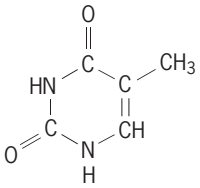

(**8**)

guanine-cytosine or adenosine-thymine pairs (called Watson-Crick base pairs) is responsible for the formation of the DNA double helix. The order of appearance of bases (primary structure) along a DNA chain translates into a set of instructions for the synthesis of a particular polypeptide. There are three-base codes associated with each amino acid. A sequence of bases representing the sequence of amino acids in a particular polypeptide is termed a gene. This information is copied in the base sequence of a special RNA molecule (messenger RNA, or mRNA) which forms the template upon which polypeptides are synthesized. The correct amino acid is brought to the correct location on mRNA via a transfer RNA (tRNA) molecule which has bases that are complementary to those on the mRNA. The correctly positioned amino acid is then linked to the polypeptide with the aid of enzymes. Without the mRNA template, amino acids would react to form polypeptides with random and hence useless sequences.

Polynucleotides of any base sequence, sometimes referred to as designer genes, can be readily prepared by automated synthesis in the laboratory and inserted into the DNA of simple organisms. The organisms with the new piece of DNA spliced into their own DNA can be easily cloned. These organisms will make the polypeptide coded by the new DNA sequence in addition to all of the others in their DNA program. By using this recombinant DNA technology, valuable polypeptides, such as hormones, as well as completely new ones, can be produced. The isolation of key proteins from a broth of many proteins represents a major challenge in the biotechnology industry. *See* DEOXYRIBONUCLEIC ACID (DNA); RIBONUCLEIC ACID (RNA).

Polyesters. Certain bacteria, when grown under conditions of limited nutrient supply, synthesize polyesters that can be broken down later and used for food and, therefore, energy. The principal class of polyesters that the bacteria produce is known as poly(β-hydroxyalkenoates), structure (**9**). The most common monomer is *d*-(−)-3-hydroxybutyric acid (**10**), and the resulting polymer is known as

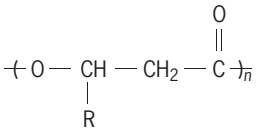

(**9**)

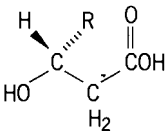

(**10**)

poly(3-hydroxybutyrate), or PHB (R = methyl). Commercial interest in this material stems from the fact that it is biodegradable.

Copolyesters can be readily prepared by controlling the nature of the nutrient medium. For example, *d*-(−)-3-hydroxybutyrate is synthesized in bacterial cells by utilizing acetate as the carbon source. However, addition of propionate results in the synthesis of a copolyester having both hydroxybutyrate and hydroxyvalerate (R = ethyl) units. The copolyesters are biodegradable and in addition have good mechanical properties, making them attractive for many applications. This approach has been extended to the synthesis of a whole series of novel copolyesters. PHB and related polymers also can be prepared in the laboratory by ring-opening polymerization of cyclic esters known as lactones. However, the purity of the biological product and the reliability of the biological "machines" make biosynthesis an attractive route. *See* BIODEGRADATION.

Polyisoprenes. The latex from rubber trees consists of tiny spheres of *cis*-1,4-polyisoprene (**11**)

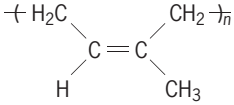

(**11**)

stabilized from coagulation by a proteinaceous coating. The latex from *Gutta percha* and balata is primarily *trans*-1,4-polyisoprene (**12**). Both are polymers of the monomer isoprene (**13**).

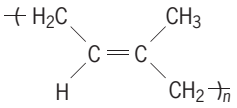

(**12**)

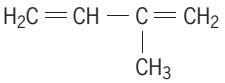

(**13**)

The latex can be coagulated in aqueous salt solutions. The resulting material from the rubber tree is known as natural rubber and is particularly important, finding extensive use in crosslinked form as an elastomer for tires, gaskets, and other applications. Demand for natural rubber decreased after chemists discovered how to prepare high-quality *cis*-polyisoprene in the laboratory. The *trans*-isomer is a hard, semicrystalline polymer that is used for golf ball covers, and also can be easily prepared abiologically.

There seems to be no agreement as to a biological motivation for certain plants to produce polyisoprenes. Some suggestions are that polyisoprenes are nutrient sources or the waste products of metabolic processes, or that they provide protection from infection when the tree is damaged. Gary E. Wnek

Bibliography. A. L. Lehninger, D. L. Nelson, and M. M. Cox, *Principles of Biochemistry*, 4th ed., 2004; A. K. Mohanty et al. (eds.), *Natural Fibers, Biopolymers, and Biocomposites*, 2005; J. D. Watson et al., *Molecular Biology of the Gene,* 5th ed., 2003.

Biopotentials and ionic currents

The voltage differences which exist between separated points in living cells, tissues, organelles, and organisms are called biopotentials. Related to these biopotentials are ionic charge transfers, or currents, that give rise to much of the electrical changes occurring in nerve, muscle, and other electrically active cells. Electrophysiology is the science concerned with uncovering the structures and functions of bioelectrical systems. This study involves a search for pertinent molecular entities directly involved with or related to biological potentials and currents. According to their function, these structures are given descriptive names such as channels, carriers, ionophores, gates, and pumps. Several of these operational entities have been identified with molecular structures whose precise chemical composition and configuration are known at the atomic level.

Electrophysiology has become increasingly a major subdiscipline of biophysics, and with the advent of studies relating molecular structure to physiological activity it has leaned heavily on the methods and insights of molecular biology and biochemistry. As the methods and techniques used in electrophysiology involve sophisticated instrumentation, this endeavor has resulted in the development of a new area of biomedical engineering. Much of what is now possible technically has resulted from rapid advances in solid-state physics and in computer sciences. Many bioelectrical systems lend themselves readily to mathematical descriptions. These applied mathematical models continue to be a significant concern of mathematical biophysicists. *See* BIOCHEMISTRY; BIOMEDICAL ENGINEERING; CARDIAC ELECTROPHYSIOLOGY; MATHEMATICAL BIOLOGY.

Electric fishes developed a bioelectrical system that became one of the first biological systems studied systematically. Animal electricity was known by its effects to ancient Egyptians, who made observations of the electric catfish, *Malapterurus*, about 2600 B.C. The early Egyptians, Greeks, and Romans were fascinated by these creatures, and recommended using the electric discharges of the fishes for the treatment of such disorders as headaches and epilepsy.

While A. Galvani is usually credited with having performed in 1790 the first laboratory experiments on the stimulation to activity of a bioelectrical system, J. Swammerdam performed somewhat similar experiments in 1658, and P. van Musschenbroeck of Leyden described in the mideighteenth century the effects of the electric discharge of a Leyden jar (an early type of capacitor for storing electric charge) on his own nervous system. *See* ELECTRIC ORGAN (BIOLOGY).

Galvani performed two basic experiments. In the first he showed that electricity from any source could stimulate (or galvanize) muscle to contract even when the electricity was applied to the nerve innervating the muscle. In the second experiment he connected nerves in the spinal cord of a frog to the leg muscles through a bimetallic arc and caused the leg muscles to contract. He hypothesized that electricity in the spinal neurons passed through the metallic arc to the leg muscles, indicating that this electricity was of biological origin. He made an analogy to a Leyden jar discharged through a short circuit. Volta believed that the source of electric current in Galvani's second experiment was the bimetallic arc itself, and ran a series of experiments examining different metals, which culminated in the discovery of the electromotive force series (for metals) and the invention of the voltaic pile (or electrophorus). Volta compared his electrochemical battery with the plates in the electric organ of an electric fish. By the midnineteenth century, in his popular physics text, Ganot began his section on dynamic electricity with an introduction featuring what was then known about bioelectricity. Galvani's successors, C. Matteucci, E. Dubois-Reymond, H. L. F. von Helmholz, and Hermann, over the half-century from 1830 to 1880, extended Galvani's observations and founded the science of electrophysiology.

Resting potentials. The potential difference measured with appropriate electrodes between the interior cytoplasm and the exterior aqueous medium of the living cell is generally called the membrane potential or resting potential (E_{RP}). This potential is usually in the order of several tens of millivolts and

is relatively constant or steady. The range of E_{RP} values in various striated muscle cells of animals from insects through amphibia to mammals is about −50 to −100 mV (the voltage is negative inside with respect to outside). Nerve cells show a similar range in such diverse species as squid, cuttlefish, crabs, lobsters, frogs, cats, and humans. Even in the coenocytic marine algae, *Valonia, Halicystis,* and *Nitella,* potentials up to about −140 mV are measured between the cell sap and the outside medium. Similar potentials have been recorded in single tissue-culture cells.

Steady potential differences also can be measured across various tissues, such as the frog skin and the gastric mucosa of the stomach, and often have been shown to exist between cells. These potentials are often of complex cellular origins, but it is thought that their origin is predicted from the resting potentials of the cells making up the tissues, or involve mechanisms similar in kind if not in detail to those described here. *See* CELL (BIOLOGY).

The first indication of a cellular resting potential involved the measurement of an injury or demarcation potential by employing two nonpolarizable electrodes and a galvanometer. Dubois-Reymond showed in 1848 that a potential difference existed between the cut ends of muscle fibers and the intact external surface that was bathed in saline.

However, it was not until the use of true internal electrodes that accurate measurements of cellular resting potentials were made. These measurements demanded certain requirements: the measurement should not injure the cell; the measurement should not draw any appreciable current from the cell; and the electrode (properly the salt bridge) should not contribute any large and unknown potentials of its own to the measurement. Many investigators tried a variety of intracellular electrode measurements during the 1920s and the early 1930s, but the modern approach began with the use of the giant axons radiating from the stellate ganglia of the squid *Loligo*. *See* ELECTRODE.

In the late 1930s, H. J. Curtis and K. S. Cole in the United States and A. L. Hodgkin and A. F. Huxley in England used these giant axons to make resting potential measurements by using internal recording electrodes. The giant axons were large enough (0.02–0.04 in. or 0.5–1 mm in diameter) to allow penetration well into an axon of an axial capillary tube filled with potassium chloride solution. The potential difference across the membrane was determined by connecting the capillary (salt bridge) to a reversible electrode and positioning its open end opposite an external reversible electrode. These measurements supported and confirmed much of the earlier work done with indirect techniques.

Microelectrodes. In the late 1940s Judith Graham and R. W. Gerard perfected the microelectrode, which was a glass pipet pulled out to a fine tip with an opening in the order of the wavelength of visible light. This pipet was filled with concentrated potassium chloride and was connected as a salt bridge to a reversible electrode, such as a calomel half-cell or a silver–silver chloride wire. This electrode allowed the measurement of the resting potentials of a wide variety of cells whose diameter prohibited the use of capillary electrodes. Because of its very small opening the microelectrode has a very high resistance that necessitates the use of a preamplifier with a fixed gain and high input impedance, such as obtained by using a field-effect transistor (FET) at the input. In practice the microelectrode and the external reference electrode are first placed in the medium external to the cell. Any potential difference between the two electrodes is balanced to zero, and then the microelectrode is pushed through the cell surface. If done properly, an immediate jump in potential is recorded as soon as the microelectrode makes contact with the interior of the cell. By using a calibrated chart recorder, this potential is monitored as a function of time or as a function of a set of alterations in the environment of the cell. Electrochemically, because of the similarity in the mobilities of potassium and chloride ions in aqueous solution, use of concentrated KCl solutions for salt bridges reduces to a minimum extraneous potentials due to junctions between the electrodes and media they contact.

Voltage source. Biopotentials arise from the electrochemical gradients established across cell membranes. In most animal cells, potassium ions are in greater concentration internally than externally, and sodium ions are in less concentration internally than externally. Generally, chloride ions are in less concentration inside cells than outside cells, even though there are abundant intracellular fixed negative charges. While calcium ion concentration is relatively low in body fluids external to cells, the concentration of ionized calcium internally is much lower (in the nanomolar range) than that found external to the cells. The earliest detailed explanations of the origin of the resting potential involved the application of ideas that came from electrochemistry. In 1888 and 1889 the physicist W. Nernst proposed that the potential difference measured with reversible electrodes between regions of a solution having different concentrations of the same univalent electrolyte could be explained on the basis of the ions diffusing passively toward an equilibrium (equilibration of concentration throughout the solution). Nernst considered that ions moved with a velocity proportional to the potential and the concentration gradients. The potential arose as a result of the tendency for the more mobile ion species to get ahead, even though electrostatic attraction prevented any appreciable separation of charges. The relation describing this condition is represented by Eq. (1), where E is the potential, F is the Faraday constant [96,500 coulombs/(g)(ion)], R is the gas constant, T is the absolute temperature, Z is the valence, C_1 and C_2 are different concentrations of the electrolyte, and U and V are the aqueous mobilities of

$$E = \frac{U - V}{U + V}\frac{RT}{ZF}\ln\frac{C_1}{C_2} \qquad (1)$$

the cation (+) and the anion (−), respectively. This relation was subsequently modified by substituting activities for concentrations in order to take account of the intermolecular attractions.

A system may be set up in which two solutions of the same electrolyte at different concentrations (C_1, C_2) are separated by a membrane or barrier which allows only one ion species to permeate through it, and there is zero net current through the membrane. It is possible to derive from thermodynamics that the potential difference between the two solutions measured with reversible electrodes is given by Eq. (2),

$$E = \frac{RT}{ZF} \ln \frac{a_1}{a_2} \tag{2}$$

where a_1 and a_2 are the ion activities of the permeant species in the respective solutions. For a membrane permeable to a single ion species and separating two solutions of a salt, this relation predicts a 58-mV potential difference for a 10-fold difference in ion activity at about room temperature. This type of potential is called a concentration potential.

By 1900 it was well recognized that living cells contained higher concentrations of potassium internally than externally, that they had lower concentrations of sodium internally than externally, and that they contained internal anions which were large and thus assumed to be impermeant. In 1902 J. Bernstein proposed that the resting potential (demarcation potential) could be explained as a potassium ion concentration potential, assuming that cell membranes are impermeable to sodium ions and that potassium ions neutralize the internal impermeant anions (Gibbs-Donnan equilibrium).

During the late 1930s and early 1940s considerable evidence was accumulated showing that the resting potential was indeed a function of the external potassium ion concentration. Curtis and Cole in 1942 showed that over a given range of concentrations the relation between changes in E_{RP} and the log of the external K^+ concentration was linear with a slope given by RT/F. However, at low external potassium concentrations, deviations from predictions derived from consideration of a simple potassium concentration potential were obtained which could not be explained by experimental error.

Actual measurements of ionic movements through cell membranes of muscle fibers by H. B. Steinbach using chemical analysis and by L. A. Heppel using tracer techniques did not support the idea that membranes are impermeable to sodium ions. In fact, it was found that radioisotopically labeled sodium ion movement through the cell membrane from inside to outside seemed to depend upon the metabolism of the cell. I. M. Glynn showed that the sodium efflux from red cells depended on the ambient glucose concentration, and Hodgkin and R. D. Keynes demonstrated in squid and *Sepia* giant axons that the sodium efflux could be blocked by a variety of metabolic inhibitors (cyanide, 2,4-dinitrophenol, and azide). Inasmuch as the impermeant internal anions could not account for the low internal sodium ion concentration (via a Gibbs-Donnan equilibrium), it was proposed that a metabolic process (sodium pump) located in the cell membrane extruded sodium from the cell interior against an electrochemical gradient.

P. C. Caldwell's experiments on the squid's giant axon in the late 1950s indicated that there was a close relation between the activity of the sodium pump and the intracellular presence of high-energy compounds, such as adenosine triphosphate (ATP) and arginine phosphate. Caldwell suggested that these compounds might be directly involved in the active transport mechanism. Evidence by R. L. Post for red cells and by Caldwell for the giant axon also suggested that there was a coupling between sodium extrusion and potassium uptake. Experiments involving internal perfusion of these axons with various phosphagens support these ideas. Convincing evidence has been presented that ATP breakdown to adenosine diphosphate and phosphorus (ADP + P) provides the immediate energy for sodium pumping in the squid giant axon. It seems that the sodium pump is a sufficient explanation to account for the high internal potassium and the low internal sodium concentrations in nerve, muscle, and red blood cells. *See* ABSORPTION (BIOLOGY); ADENOSINE TRIPHOSPHATE (ATP).

It has been suggested that the key part of the sodium pump is the enzyme adenosinetriphosphatase (ATPase). This enzyme has been found in the membranes of most animal cells. The squid axon internal dialysis experiments have provided the stoichiometry for the action of membrane ATPase in catalyzing ATP breakdown promoting sodium extrusion and potassium uptake in the axon.

The existence of measurable resting sodium and chloride permeabilities in cell membranes, however, would make less acceptable the idea that the resting potential is solely due to a pure concentration potential for potassium ion. In some respects, the resolution of this was made less difficult by the earlier introduction of the constant field equation by D. E. Goldman in 1943. Assuming that the potential gradient across the thickness of the cell membrane was linear, Goldman determined a relation for the resting ionic conductances in the membrane. This relationship under the condition of zero net current flow was expressed as Eq. (3) by Hodgkin and B. Katz, where

$$E_{RP} = \frac{-RT}{F} \times \ln \frac{[K^+]_{in} + P_{Na}[Na^+]_{in} + P_{Cl}[Cl^-]_{out}}{[K^+]_{out} + P_{Na}[Na^+]_{out} + P_{Cl}[Cl^-]_{in}} \tag{3}$$

the P values indicate the ionic membrane permeabilities relative to potassium ion permeability, and the brackets indicate the concentrations (more correctly, the activities) of the enclosed ions. Assuming anion impermeability (and zero net current), the relation reduces to Eq. (4). Hodgkin and P. Horowicz

$$E_{RP} = \frac{-RT}{F} \ln \frac{[K]_{in} + P_{Na}[Na]_{in}}{[K]_{out} + P_{Na}[Na]_{out}} \tag{4}$$

were able to fit the experimentally determined curvilinear relation between E_{RP} of muscle fibers and the logarithm of external potassium ion concentration by using this equation. *See* CELL MEMBRANES.

Ionic channels. Most of the foregoing discusses systems and relationships which applied only to macroscopic conditions (many ions and large areas) and often invoked the restriction of zero net current. A zero net ionic current is a balance between equal and opposite ionic movements. In living cells there are two general types of transport processes. In the first, the transported ionic species flows down the gradient of its own electrochemical potential. In the second, there is a requirement for immediate metabolic energy. Ionic currents flowing during such active events as the nerve impulse apparently belong to the first category, even though they are ultimately dependent on energy-requiring pump systems that maintain the ionic gradients across the membrane of the living cell. This first category of bioelectrical events is associated with a class of molecules called channels, embedded in living cell membranes.

It is now known that cell membranes contain many types of transmembrane channels. Channels are protein structures that span the lipid bilayers forming the backbones of cell membranes. The cell membranes of nerve, muscle, and other tissues contain ionic channels. These ionic channels have selectivity filters in their lumens such that in the open state only certain elementary ion species are admitted to passage, with the exclusion of other ion species. Other tissue cells such as epithelial cells, endocrine gland cells, lymphocytes, and sperm cells require channel activity for their functioning. In some cases, specific channels have been implicated in diseases such as cystic fibrosis and diabetes. There is now evidence of a specialized chloride channel called the cystic fibrosis transmembrane conductance regulator (CFTR). In cystic fibrosis, epithelial cells produce abnormally thick secretions, particularly in airways to mammalian lungs. The CFTR channel is regulated by cyclic adenosine monophosphate (cAMP) and can transport many substances vital to cell function as well as chloride ions.

Nerve impulse. It is now widely accepted that a two-state ion-selective membrane channel, capable of being gated between closed and open states, is the molecular mechanism responsible for excitation in neuromuscular systems. It is also accepted that these channels are specific protein molecules capable of undergoing conformational changes in their structure that account for the observed gating properties.

There are two general types of channels, and these are classified according to the way in which they respond to stimuli. Electrically excitable channels have opening and closing rates that are dependent on the transmembrane electric field. Chemically excitable channels (usually found in synaptic membranes) are controlled by the specific binding of certain activating molecules (agonists) to receptor sites associated with the channel molecule.

The long conducting parts of nerve cells, such as axons, transmit information by trains of constant-amplitude nerve impulses, or action potentials. The information content is encoded into the frequency of impulse transmission. The generation of any one of these impulses is brought about by a sequence of movements of specific ions through an ensemble of protein molecules bridging the nerve cell membrane.

Calcium channels are involved in synaptic transmission. When a nerve impulse arrives at the end of a nerve fiber, calcium channels open in response to the change in membrane potential. These channels admit calcium ions, which act on synaptic vesicles, facilitating their fusion with the presynaptic membrane. Upon exocytosis, these vesicles release transmitter molecules, which diffuse across the synaptic cleft to depolarize the postsynaptic membrane by opening ionic channels. Transmitter activity ceases from the action of specific transmitter esterases or by reabsorption of transmitter back into vesicles in the presynaptic neuron. Calcium channels inactivate and close until another nerve impulse arrives at the presynaptic terminal. Thus biopotentials play an important role in both the regulation and the genesis of synaptic transmission at the membrane channel level. There are many neurotransmitters. Glutamate is the most prevalent transmitter in central nervous system synapses in the human brain. Acetylcholine is a common transmitter at nerve muscle junctions.

The propagation of the conducted action potential in nerve cells (**Fig. 1**) results from fast, accurately voltage-dependent permeability changes to specific ions that take place across the nerve cell membrane. In the case of the squid giant axon, the permeation of sodium ions occurs in pores in the cell membrane that have been called sodium channels. Similarly, potassium ion permeation takes place through pores called potassium channels. These channels have been identified as molecular entities with distinct operational characteristics. These membrane proteins or channels can exist in either of two states.

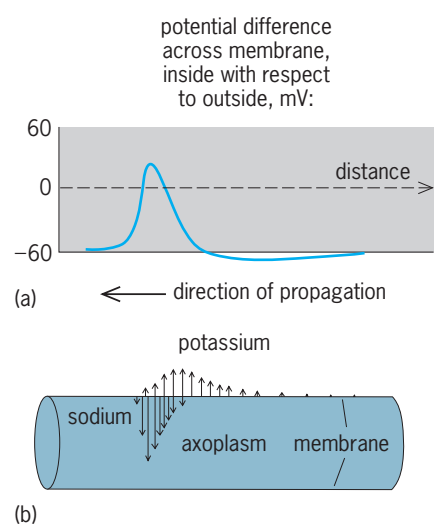

Fig. 1. Action potential in squid giant axon. (*a*) Voltage waveform. (*b*) Direction of flow of sodium and potassium ions corresponding to the action potential in *a*. The lengths of the arrows are proportional to the conductances to sodium and potassium ions. (*After W. J. Adelman, Jr., and R. J. French, The squid giant axon, Oceanus, 19(2):6–16, 1976*)

In the open state they admit ion flow, and in the closed state ion flow is prohibited. Transitions between resting, open, and closed states, or vice versa, are referred to as channel gating. It is now clear that such gating is the result of channel molecule conformational changes.

In order to describe the sequence of events that occur during an action potential, and to explain the macroscopic properties of electrically excitable membranes (such as threshold, propagation velocity, refractoriness, and anesthetic blockade), it is necessary to understand that these properties arise from the ensemble characteristics of many channels activated in some form of recruited synchrony. The most accurate measurement of these ensemble characteristics has involved voltage-clamping a reasonably large area of excitable membrane and recording the ionic current flow through thousands of parallel channels.

In 1952 Hodgkin and Huxley had described in quantitative detail the ionic currents flowing through a large area of nerve membrane containing an ensemble of many channels. From this description, they were able to reconstruct the action potential and predict many of its characteristics. *See* NERVE.

Single channels. Irwin Neher and Bert Sakmann were awarded a Nobel prize in 1991 for their development of the patch electrode and the description of the function of single channels in cells. **Figure 2** shows a glass pipette electrode sealing to the membrane of a cell such that a single channel is held in the orifice of the pipette. Ionic currents flowing through the channel can be recorded, and the membrane potential across the channel can be clamped to different values (patch clamp). With improvements in techniques such as the use of ultraclean or specially coated pipettes and use of negative pressures (reambient) inside the pipette, gigohm seals of the pipette electrode to the cell membrane could be achieved. These high-resistance seals meant that the background noise generated in the shunt resistance across the seal was reduced to a minimum (the higher the resistance, the lower the Johnson noise). This made the recording of single-channel ionic currents more distinct from background noise.

In 1980, R. Horn and J. Patlak showed that an isolated patch of membrane with its functioning membrane channels could be torn off from the cell and remain attached to the gigaseal electrode. The channel activity was the same as in the intact cell, and manipulation of internal solutions was made much easier.

Figure 3 shows records of single-channel potassium currents recorded with a patch electrode sealed to the inside of a squid giant axon. As the patch membrane was voltage-clamped, the membrane voltage and hence the voltage across the channel could be set to different values. The records clearly show that the dwell time in the conducting and nonconducting states is a function of voltage. In 1980 the results of using the patch electrode technique were published, indicating that single-channel sodium currents in muscle cells are related to macroscopic

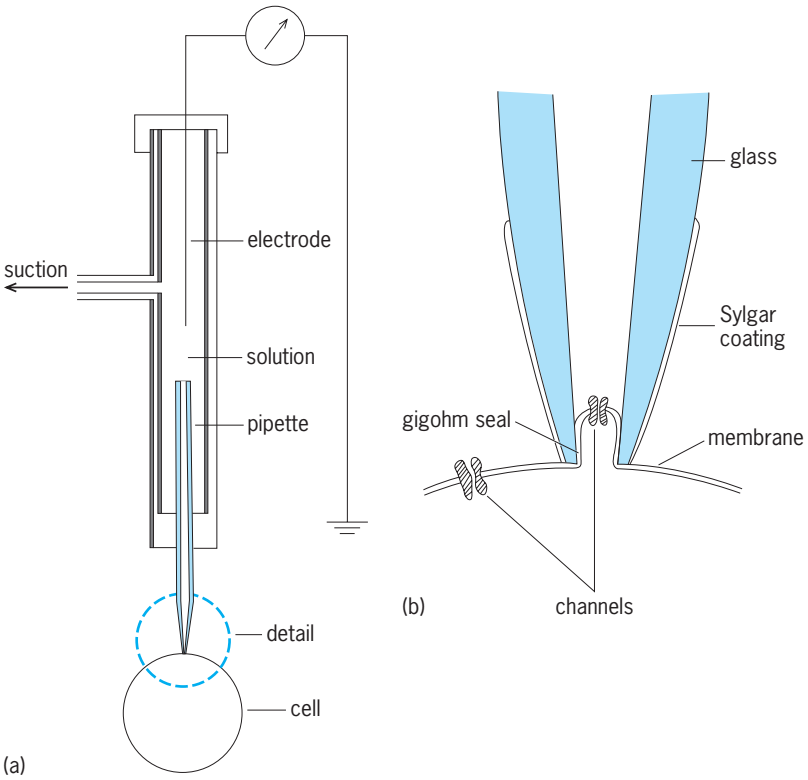

Fig. 2. Single-channel recording. (*a*) Gigaseal electrode holder, gigaseal glass pipette attached to a living cell, silver-silver chloride electrode in the test solution, and side arm for attaching tubing to produce the negative pressure required to establish the gigaseal. (*b*) Enlargement of detail in *a*. Not to scale.

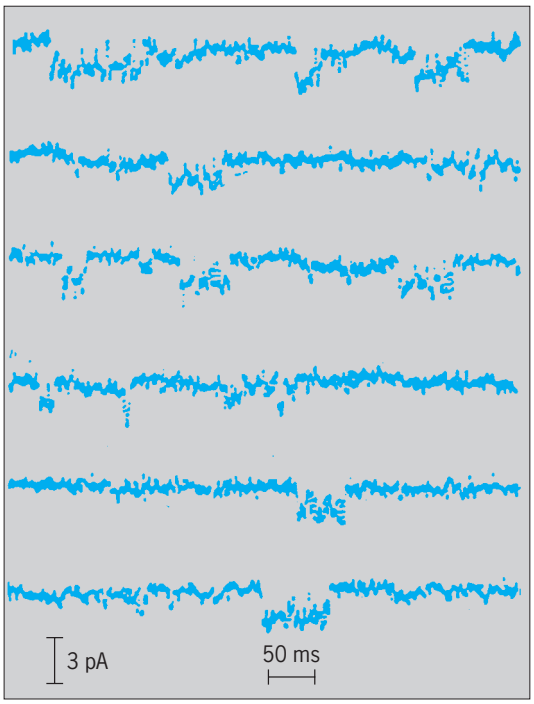

Fig. 3. Discrete current step measured for single potassium channels in the squid giant axon membrane. The upper three recordings were obtained with a −25-mV membrane potential. The lower three recordings were obtained with a −35-mV potential. The current steps are less frequent in the lower recording, indicating that the number of steps is voltage-dependent. (*After F. Conti and E. Neher, Single channel recordings of K^+ currents in squid axons, Nature, 285:140–143, 1980*)

measurements of sodium currents from the cells. The response of a single sodium channel to a step in membrane potential repeated and summed many times has the same shape as the macroscopic membrane response to a single voltage clamp step, which presumably arose from the spatial summation of a set of similar unit channels. This superposition summation experiment has been taken by many investigators as the most convincing evidence that the phenomena giving rise to the nerve impulse have their origin in single-channel kinetics.

Mechanisms exist whereby specific ion channels can be opened and closed. Fluctuation between open and closed states among a population of channels determines both the magnitude and the sign of the biopotential existing across the cell membrane at any time. In general, there are two classes of channels: those that respond to the electric field across the membrane, and those that respond to chemical stimuli. Even channels that admit a specific ion can be of these two general classes, for example, voltage-activated potassium channels and acetylcholineactivated potassium channels. Acetylcholine released by vagus nerve activity has been shown to activate potassium channels in nodes and atrial muscles of the heart, which are different from the usual cardiac potassium channels. In human vascular smooth muscle cell membranes, at least four different types of potassium channels have been found (calcium-sensitive, voltage-dependent, adenosine triphosphate–sensitive, and inward rectifying channels). In 1984, B. Hille suggested that the human genome probably codes for as many as 50 different types of membrane channels.

Ionic currents flow through open channels. The ion impermeable membrane lipid bilayer acts as a dielectric separating two highly conductive salt solutions. Ionic channels have the electrical property of a conductance between these solutions. In 1939, Cole and Curtis showed that this conductance could be viewed as existing in parallel with the membrane capacitance arising from the lipid bilayer. The membrane conductance at any moment depends on the total number of channels, the type of channels, the fraction of channels found in the open state, and the unit conductances of these open channels. This unit conductance, g_{open}, is expressed in Eq. (5),

$$g_{\text{open}} = I_{\text{open}}/f(V - E_i) \tag{5}$$

where V is the voltage across the channel, and E_i is the reversal potential for the ionic current; $f(V - E_i)$ considers that often current voltage relations are not linear and the possibility of competition among ion species for selectivity filter sites is of concern. The average conductance of a single channel fluctuating between open and closed states, g_{aver}, is given by Eq. (6), where p is the probability of finding the channel in the open state. When p is 1 the channel is open all the time; when p is zero the channel is closed all the time. For a given area of membrane containing N channels of the same type, the conductance is given by Eq. (7), where p may be a function of the electric

$$g_{\text{aver}} = pg_{\text{open}} \tag{6}$$

$$g = Npg_{\text{open}} \tag{7}$$

field or of agonist binding to a channel receptor or both, depending on the type of channel.

The most common channels directly giving rise to biopotentials are those admitting mainly sodium ions, potassium ions, chloride ions, or calcium ions. These channels are named after the predominant charge carrier admitted in the open state, such as potassium channels. It is now known that there are charged amino acid groups lining the channel lumen that determine the specificity of the channel for particular ions. These selectivity filters admit only ions of the opposite charge.

For a membrane area containing a variety of ionic channels, the overall ionic conductance, G_m, is given by Eq. (8), where the subscripted conductances are

$$G_m = g_{\text{Na}} + g_{\text{K}} + g_{\text{Cl}} + g_{\text{Ca}} + \cdots + g_i \tag{8}$$

the specific channel conductances. Each is described in Eq. (9). The ionic current, I, through a parallel array of channel conductances may then be readily expressed as Eq. (10). The membrane current flowing

$$g_i = N_i p g_{i,\text{open}} \tag{9}$$

$$\begin{aligned} I_i = {} & g_{\text{Na}} f(V - E_{\text{Na}}) + g_{\text{K}} f(V - E_{\text{K}}) \\ & + g_{\text{Cl}} f(V - E_{\text{Cl}}) + g_{\text{Ca}} f(V - E_{\text{Cl}}) \\ & + \cdots + g_i (V - E_i) \end{aligned} \tag{10}$$

during a step change in membrane potential is given by Eq. (11), where C_m is the membrane capacity.

$$I_m = C_m(dV/dt) + I_i \tag{11}$$

When dV/dt is zero (constant voltage), the capacity current is zero, and the membrane current is the ionic current. In principle, this is the basis for step voltage clamping. In practice, step voltage clamping requires that changes in membrane voltage be virtually instantaneous so that the capacity current settles to zero before ionic currents begin to flow.

In order to explain how channels gate from closed to open states and conversely, Hodgkin and Huxley proposed in 1952 that there were charged molecular entities responsible for the opening and closing of the ionic conductance pathways. These structures had to be charged to be able to move in response to changing electrical forces when the membrane voltage changed. Any movement of the gating structures would require a movement of charge and hence should have a detectable component of current flow across the membrane. It was not until 1973 that the existence of a gating current in squid axon sodium channels was demonstrated. By eliminating ionic charge carriers from the preparation, researchers measured a small but finite current whose characteristics had the properties of the expected

gating charge movement. This was confirmed, and gating currents and their significance became a lively endeavor in membrane biophysics.

Channels appear to exist in all living excitable membranes. There are at least 10 classes of channels, and unit channels found in human muscle membranes bear a striking resemblance to those found in lower species. Many laboratories have attempted to isolate, purify, and then reconstitute biological channels into artificial lipid membranes.

The study of channel molecular structure has been extended into developmental biology. Channels develop, mature, differentiate, and are suppressed throughout the development of an organism from the initial egg cell to the mature organism. This process can be followed by electrophysiological methods (such as patch clamping) and by molecular biological methods.

Channel proteins. It has become almost routine to be able to isolate channel proteins from natural membranes and reconstitute these channels into artificial bilayer membranes formed between aqueous salt solutions. Acetylcholine receptor channels, sodium channels from the electric eel (*Electrophorus electricus*), sodium channels from mammalian brain, calcium-activated potassium channels from the sarcoplasmic reticulum of skeletal muscle, calcium channels from several sources, and potassium channels from muscle have all been isolated and reconstituted into planar bilayers.

More importantly, it has become possible to make use of gene expression to produce specific channel proteins in certain egg cells, such as eggs from the toad (*Xenopus laevis*). By obtaining the channel messenger ribonucleic acid (mRNA) from its appropriate complementary deoxyribonucleic acid (cDNA), and injecting this mRNA into an egg with proper substrates and activators, channel proteins may be synthesized in the egg and the channels reconstituted into the egg membrane.

Patch-clamping these channels has shown that they have ionic currents and closed-open transition kinetics that are similar to those measured in the original channel systems from which the cDNA was obtained. It is now possible to use genetic engineering techniques to alter the channel DNA and obtain new messengers which can be used to synthesize and incorporate in living membranes specifically altered channel proteins.

One of the major advances in the application of the techniques of molecular biology and genetic engineering to the study of channels took place in Japan in the early 1980s. Following the successful sequence of several pituitary hormones and significant active peptides such as beta lipotropin, the gene sequencing for the different subunits of the acetylcholine receptor channel was determined. From the nucleic acid sequence it was possible to determine the precise amino acid sequence of this protein. Following this success, the sequence of the sodium channel protein, of the transducer protein in visual cells which is linked to activated rhodopsin, and of the sodium-potassium adenosinetriphosphatase protein (the sodium pump molecule) was also determined.

Several neurotoxins have been shown to bind specifically to sodium channels in chick cardiac muscle cells, skeletal muscle fibers of the rat, rat brain cells, and the electric organ of the eel (*E. electricus*). By making use of these specific bindings, highly purified sodium channel protein has been obtained. The molecular weight of the polypeptide sodium channel protein ranges from 200,000 to 300,000, depending on the species. Recombinant DNA methods were used to clone DNA sequences complementary to mRNA coding for the sodium channel polypeptide of the eel electroplax. The complete amino acid sequence of this polypeptide was obtained by making use of a nucleotide sequence analysis of this cloned cDNA.

The eel electroplax sodium channel protein was shown to contain 1820 amino acid residues with a molecular weight of 208,321. Late in 1984 the nature of this primary structure of the sodium channel of the eel electroplax as based on cDNA sequencing was described. It was suggested that this primary sequence was composed of four domains of six subgroups each. A theoretical space-filling molecular model of the tertiary structure of the membrane-bound sodium channel based on the primary sequence and energy considerations was produced.

Use has been made of mRNAs derived from acetylcholine channel cDNAs in which the code for the synthesis of the channel had been changed for only one key amino acid at a time. By making use of the expression of genetically engineered altered channel proteins in *Xenopus* eggs, researchers have been able to show which specific amino acids in the acetylcholine channel protein are involved in specific channel functions. Once the expressed channels were incorporated in the oocyte membrane, these channels were patch-clamped and their characteristics were measured. This has allowed a direct correlation between the structural chemistry of the channel subunits and the biopotentials and ionic currents generated through the channels. *See* BIOPHYSICS.

William J. Adelman, Jr.

Bibliography. K. S. Cole, *Membranes, Ions, and Impulses*, University of California Press, Berkeley, 1968; R. J. French and W. J. Adelman, Jr., Competition, saturation and inhibition: Ionic interactions shown by membrane ionic currents in nerve, muscle and bilayer systems, *Curr. Top. Membranes Transport*, 8:161–207, 1976; A. D. Grinnell, D. Armstrong, and M. B. Jackson (eds.), *Calcium and Ion Channel Modulation*, Plenum Press, New York, 1988; A. L. Hodgkin and A. F. Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve, *J. Physiol. (London)*, 117:500–544, 1952; H. Lecar, C. Morris, and B. S. Wong, Single-channel currents and the kinetics of agonist-induced gating, in D. C. Chang et al., *Structure and Function in Excitable Cells*, Plenum Press, New York, 1983; V. Milesi et al., Role of a Ca^{2+}-activated K^+ current in the maintenance of resting membrane potential of isolated human saphenous

vein smooth muscle cells, *Pflugers Arch. Eur. J. Physiol.*, 437:455–461, 1999; B. Sakmann and E. Neher (eds.), *Single-Channel Recording*, 2d ed., Plenum Press, New York, 1995; E. M. Schweibert et al., CFTR is a conductance regulator as well as a chloride channel, *Physiol. Rev.*, 79 (suppl. 1):S145–S173, 1999; K. Takahashi and Y. Okamura, Ion channels and early development of neural cells, *Physiol. Rev.*, 78:307–337, 1998; H. Wheal and A. Thomson (eds.), *Excitatory Amino Acids and Synaptic Transmission*, 2d ed., Academic Press, London, 1995.

Biopyribole

A member of a chemically diverse, structurally related group of minerals that comprise substantial fractions of both the Earth's crust and upper mantle. The term was coined by Albert Johannsen in 1911 to describe collectively the rock-forming pyroxene, amphibole, and mica minerals [biopyribole is a contraction of biotite (a mica), pyroxene, and amphibole]. In addition to these common and long-recognized biopyriboles, several other members of the group have been recognized since the mid-1970s.

The pyroxene minerals contain single chains of corner-sharing silicate (SiO_4) tetrahedra (**Fig. 1***a*), and the amphiboles contain double chains (Fig. 1*b*). Likewise, the micas and other related biopyriboles (talc, pyrophyllite, and the brittle micas) contain two-dimensionally infinite silicate sheets, which result in their characteristic sheetlike physical properties. In the pyroxenes and amphiboles, the silicate chains are articulated to strips of octahedrally coordinated cations, such as magnesium and iron; and in the sheet biopyriboles, the silicate sheets are connected by two-dimensional sheets of such cations. *See* AMPHIBOLE; MICA; PYROXENE.

In addition to the classical single-chain, double-chain, and sheet biopyriboles, several biopyriboles that contain triple silicate chains (Fig. 1*c*) have been discovered. These minerals had been overlooked previously due to small crystal size and similarity in physical properties to those of the more common pyriboles (pyribole is a term applied to the chain-silicate subset of the biopyribole group).

Structurally disordered types. Triple silicate chains of the type shown in Fig. 1*c* were first found as isolated faults in amphibole asbestos. Such triple chains, as well as quadruple, quintuple, and wider chains, occur as defects in many amphibole specimens. The intergrowth of these triple and wider chains in amphibole is a type of structural disorder. In addition, similar structural disorder occurs in some pyroxenes, where double, triple, and wider silicate chains are intergrown with the predominantly single-chain structure.

Much of the knowledge about structurally disordered biopyriboles has been obtained by the use of high-resolution transmission electron microscopy. An electron micrograph of structurally disordered anthophyllite (an amphibole) is shown in **Fig. 2**. In this electron image, the crystal is viewed parallel to the silicate chains, which define the direction of the *c* crystallographic axis. In addition to the double-chain structure characteristic of amphibole, the crystal contains slabs with triple, quadruple, and sextuple silicate chains.

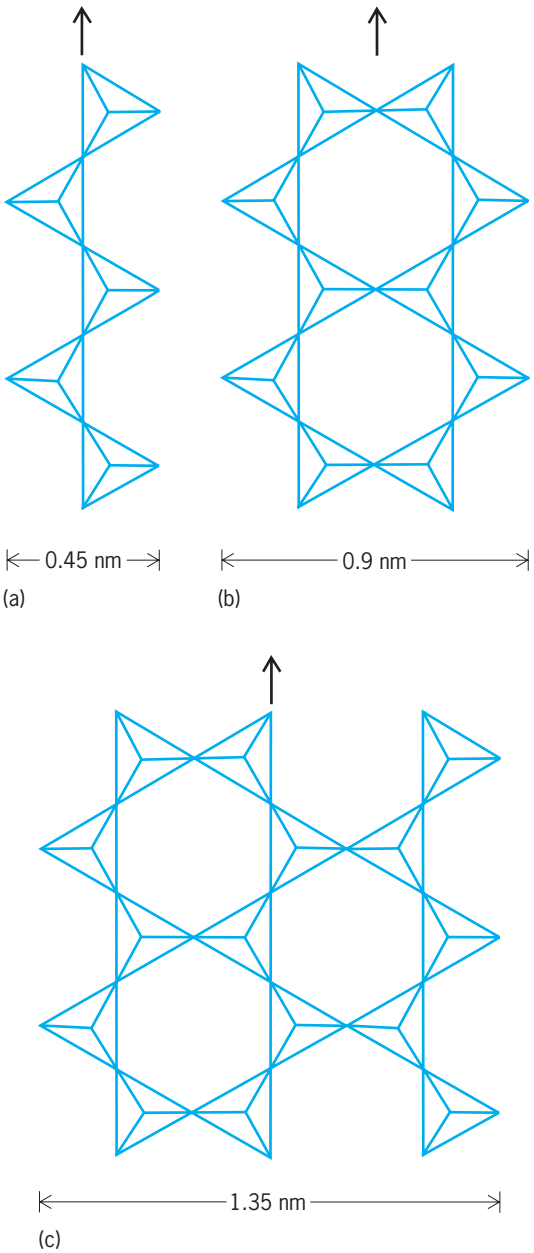

Fig. 1. Silicate-chain structures for biopyribole minerals. (*a*) Single silicate chain characteristic of pyroxenes. (*b*) Double silicate chain of amphiboles. (*c*) Triple silicate chain as found in the new biopyriboles. (*After D. R. Veblen et al., Asbestiform chain silicates: New minerals and structural groups, Science, 198:359–365, 1977*)

Structurally ordered wide- and mixed-chain types. The wide-chain biopyribole minerals containing ordered arrays of triple silicate chains were discovered by using single-crystal x-ray diffraction methods. The minerals jimthompsonite and clinojimthompsonite consist purely of triple-silicate-chain structure. The mineral chesterite and an unnamed mineral contain both triple and double silicate chains that alternate

Chemical formulas, unit-cell parameters, and space group symmetries of ferromagnesian biopyriboles*

System	Pyroxenes, $(Mg,Fe)_2Si_2O_6$	Amphiboles, $(Mg,Fe)_7Si_8O_{22}(OH)_2$	Triple-chain silicates, $(Mg,Fe)_{10}Si_{12}O_{32}(OH)_4$	Mixed-chain silicates, $(Mg,Fe)_{17}Si_{20}O_{54}(OH)_6$
Orthorhombic	Enstatite $a = 1.82$ nm $b = 0.88$ $c = 0.52$ *Pbca*	Anthophyllite $a = 1.86$ nm $b = 1.81$ $c = 0.53$ *Pnma*	Jimthompsonite $a = 1.86$ nm $b = 2.72$ $c = 0.53$ *Pbca*	Chesterite $a = 1.86$ nm $b = 4.53$ $c = 0.53$ $A2_1ma$
Monoclinic	Clinoenstatite $a = 0.96$ nm $b = 0.88$ $c = 0.52$ $\beta = 108.3°$ *C2/m*	Cummingtonite $a = 0.95$ nm $b = 1.81$ $c = 0.53$ $\beta = 109.6°$ *I2/m*	Clinojimthompsonite $a = 0.99$ nm $b = 2.72$ $c = 0.53$ $\beta = 109.5°$ *C2/c*	Unnamed mineral $a = 0.99$ nm $b = 4.53$ $c = 0.53$ $\beta = 109.7°$ *A2/m*

*The space groups are for high-temperature forms.

rigorously parallel to the *b* crystallographic axis. Chesterite is the first known example of a mixed-chain silicate in that it contains two topologically distinct types of silicate chains.

The properties of some ferromagnesian pyroxenes and amphiboles are listed in the **table**, along with those of the more recently recognized wide-chain ordered pyriboles. The differences in chain width among these minerals are revealed by the lengths of the *b* crystallographic axes, which are parallel to the widths of the chains ($b \cong 0.9$ nanometer for single chains, 1.8 nm for double chains, 2.7 nm for triple chains, and 4.5 nm for alternating double and triple chains). All of the monoclinic chain silicates listed in the table possess *a* axes of approximately 1.0 nm, because their chains are stacked in an identical fashion parallel to this axis. Likewise, all of the orthorhombic minerals in the table have *a* axes of about 1.8 nm, the result of their identical stacking parallel to *a* that differs from that of the monoclinic members of the group. These relationships among unit-cell parameters are the direct result of the close structural relationships among all the biopyriboles. *See* CRYSTAL STRUCTURE.

The sequences of silicate chains in biopyriboles can be represented as follows: pyroxene [1], amphibole [2], chesterite [23], jimthompsonite [3], and mica or talc [∞]. As an example, in this notation [2] implies that double silicate chains repeat infinitely with the sequence . . . 222222222. . . in the direction of the *b* axis. Of these macroscopically occurring, ordered biopyriboles, only chesterite is a mixed-chain silicate because it is the only structure that contains more than one type of silicate chain. However, many additional ordered, mixed-chain structures have been observed with electron microscopy. These include biopyriboles with more complicated chain sequences, such as [2233], [233], [2333], [234], and [222222333333]. In addition, mixed-chain biopyriboles containing single chains have been observed, such as [1112], and the structure [12] has been synthesized at very high pressures as single crystals large enough for analysis by x-ray diffraction methods. *See* X-RAY CRYSTALLOGRAPHY.

Theoretical basis of structures. Amphibole structures consist of slabs of pyroxene (P) structure and mica or talc structure (M). These slabs alternate rigorously parallel to the (010) planes of amphibole in the sequence . . . MPMPMP . . . , or simply (MP). Any amphibole can thus be thought of as a 1:1 mechanical mixture of pyroxene and mica. The pyriboles containing triple silicate chains also have structures made of the same types of slabs, with the sequence (MMP) in pure triple-chain minerals such as jimthompsonite and the sequence (MPMMP) in the mixed-chain members of the group such as chesterite.

Because all of the biopyriboles consist of such mixtures of two different kinds of structural slabs, or of pure M slabs (mica) or P slabs (pyroxene), they all have stoichiometries that are collinear with the pyroxene and mica (or talc) stoichiometries. This chemical relationship is shown in **Fig. 3** for the idealized pure magnesium compositions. A group of structures that possesses this kind of structural and chemical relationship has been called a polysomatic series.

The biopyribole crystal structures are very compliant and, within the constraints of their stoichiometries, can adapt to wide variations in chemical composition, pressure, and temperature. Depending on the exact geochemical and physical conditions, the

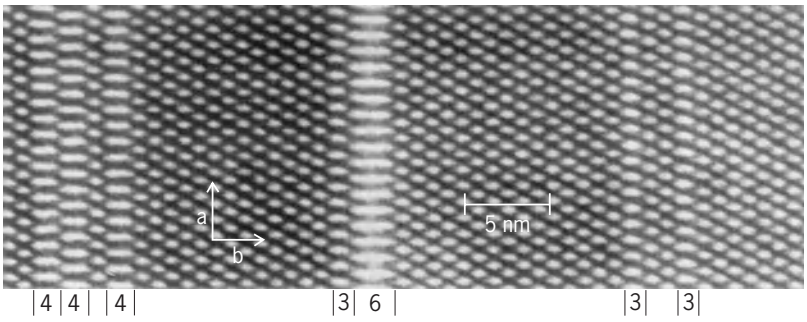

Fig. 2. Transmission electron micrograph of a structurally disordered crystal of the amphibole anthophyllite. Normal double-chain structure is unlabeled, and planar defects with triple-, quadruple-, and sextuple-chain structure are labeled 3, 4, and 6, respectively. (*From D. R. Veblen, ed., Amphiboles and Other Hydrous Pyriboles: Mineralogy, Mineral. Soc. Amer. Rev. Mineral., vol. 9a, 1981*)

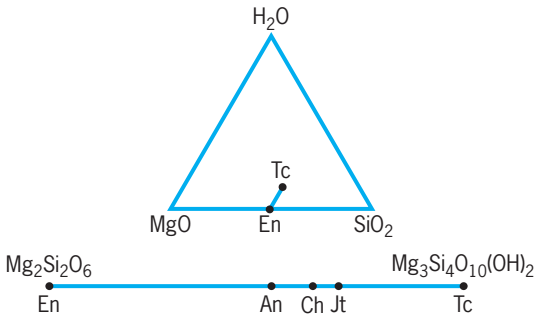

Fig. 3. Chemical compositions of ideal pure magnesium biopyriboles. En = enstatite (pyroxene), An = anthophyllite (amphibole), Ch = chesterite, Jt = jimthompsonite, and Tc = talc (sheet silicate). (*After D. R. Veblen et al., Asbestiform chain silicates: New minerals and structural groups, Science, 198:359–365, 1977*)

various naturally occurring biopyribole minerals can contain major amounts of hydrogen, lithium, oxygen, fluorine, sodium, magnesium, aluminum, silicon, chlorine, potassium, calcium, titanium, magnesium, and iron, in addition to other elements that may occur less commonly. Furthermore, compositions containing a much broader range of elements can be synthesized in the laboratory. In some biopyriboles, vacancies are also abundant on some crystallographic sites, and sites that are filled for some compositions may be vacant for others. *See* ELEMENT (CHEMISTRY).

Occurrence and stability. Pyroxenes, amphiboles, and micas of various compositions can occur in igneous, metamorphic, and sedimentary rocks. Pyroxenes are the second most abundant minerals in the Earth's crust (after feldspars) and in the upper mantle (after olivine). Amphiboles occur as at least minor constitutents in many common igneous rocks and may make up the bulk of some metamorphic rocks. The sheet-silicate biopyriboles are major phases in some sedimentary rocks such as shales, occur in many igneous rocks, can make up the bulk of some metamorphic rocks, and may even be important volatile-bearing minerals in the upper mantle. These biopyriboles clearly have fields of thermodynamic stability. *See* IGNEOUS ROCKS; METAMORPHIC ROCKS; SEDIMENTARY ROCKS.

Unlike the pyroxenes, amphiboles, and micas, the wide-chain biopyriboles do not occur as abundant minerals in a wide variety of rock types. However, some of them may be widespread in nature as components of fine-grain alteration products of pyroxenes and amphiboles and as isolated lamellae in other biopyriboles (but single crystals up to several millimeters long and a few tenths of a millimeter wide have been observed for jimthompsonite and chesterite). Also, unlike the classical biopyriboles, it is not known whether or not any of the wide-chain pyriboles possess true fields of thermodynamic stability. Possibly, they always form as a result of metastable reaction or growth processes. There has been no experimental confirmation of the stability relations of these wide-chain structures. *See* BIOTITE; SILICATE MINERALS.

David R. Veblen

Bibliography. J. Akai, Polymerization process of biopyribole in metasomatism at the Akatani ore deposit, Japan, *Contrib. Mineral. Petrol.*, 80:117–131, 1982; S. W. Bailey (ed.), Micas, *Mineral. Soc. Amer. Rev. Mineral.*, no. 13, 1984; W. A. Deer, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, 2d ed., Geological Society, London, vol. 2A, 1978, vol. 2B, 1997; B. H. Grobéty, New short-range biopyribole polysomes from the Lepontine Alps, Switzerland, *Amer. Mineral.*, 81:404–417, 1996; B. H. Grobéty, The replacement of anthophyllite by jimthompsonite: A model for hydration reactions in biopyriboles, *Contrib. Mineral. Petrol.*, 127:237–247, 1997; B. E. Leake et al., Nomenclature of amphiboles, *Amer. Mineral.*, 82:1019–1037, 1997; N. Morimoto et al., Nomenclature of pyroxenes, *Amer. Mineral.*, 73:1123–1133, 1988; C. T. Prewitt (ed.), Pyroxenes, *Mineral. Soc. Amer. Rev. Mineral.*, no. 7, 1980; M. Rieder et al., Nomenclature of the micas, *Can. Mineral.*, 36:905–912, 1998; J. B. Thompson, Jr., Biopyriboles and polysomatic series, *Amer. Mineral.*, 63:239–249, 1978; D. R. Veblen (ed.), Amphiboles and Other Hydrous Pyriboles: Mineralogy, *Mineral. Soc. Amer. Rev. Mineral.*, no. 9A, 1981; D. R. Veblen, Polysomatism and polysomatic series: A review and applications, *Amer. Mineral.*, 76:801–826, 1991; D. R. Veblen and P. H. Ribbe (eds.), Amphiboles: Petrology and Experimental Phase Relations, *Mineral. Soc. Amer. Rev. Mineral.*, vol. 9B, 1982.

Biorheology

The study of the flow and deformation of biological materials. As all living things flow and deform, biorheology is involved in the study of every aspect of biology. The behavior and fitness of living organisms depend partly on the mechanical properties of their structural materials. Humans are healthy when the rheological properties of their tissues are normal; deviations lead to illness. For example, a thinner than normal fluid in the knee joint results in pain and arthrosis, and softened or weakened bones result in fractures.

Thus, biologists are interested in biorheology from the point of view of evolution and adaptation to the environment. Physicians are interested in it in order to understand health and disease. Bioengineers devise methods to measure or to change the rheological properties of biological materials, develop mathematical descriptions of biorheology, and create new practical applications for biorheology in agriculture, industry, and medicine.

The rheological behavior of most biological materials is more complex than that of air, water, and most structural materials used in engineering. Air and water are viscous fluids; all fluids whose viscosity is similar to that of air and water are called newtonian fluids. Biological fluids such as protoplasm, blood, and synovial fluid behave differently, however, and they are called non-newtonian fluids. For example, blood behaves like a fluid when it flows, but when

it stops flowing it behaves like a solid with a small but finite yield stress. When it does flow, its viscosity decreases when the shear-strain rate increases. Thus, in a tube the faster the blood flows, the less viscous it becomes. These features are non-newtonian. *See* NON-NEWTONIAN FLUID; VISCOSITY.

Most materials used in engineering construction, such as steel, aluminum, or rock, obey Hooke's law, according to which stresses are linearly proportional to strains. These materials deviate from Hooke's law only when approaching failure. Solids obeying Hooke's law are said to be Hookean. A structure made of Hookean materials behaves linearly: load and deflection are linearly proportional to each other in such a structure. Some biological materials, such as bone and wood, also obey Hooke's law in their normal state of function, but many others, such as skin, tendon, muscle, blood vessels, lung, and liver, do not. These materials, referred to as non-Hookean, become stiffer as stress increases. *See* BONE; ELASTICITY; HOOKE'S LAW; STRESS AND STRAIN.

Microbiorheology. To gain a fundamental understanding of the rheological properties of biological tissues, a modern trend is to probe deeper into the cellular and molecular aspects of the system. Since the knowledge gained is applicable at the basic level, this approach often provides important practical uses.

Mucus. The rheological properties of mucus may be determined by a high-molecular-weight glycoprotein that is covered by carbohydrate side chains. Variations of this molecule are introduced by the secretory cells of the epithelium. In the case of ciliated epithelia, the mucus forms a layer in contact with the cilia and is rheologically adjusted so as to cause optimal transport and clearance. There are reasons to think that mucus is not normally present over the ciliary epithelium and that the cilia are not normally in motion. Secretion is stimulated when loads impinge over a particular section of the epithelium. Ciliary motion is initiated by the presence of mucus and ceases when the patch of mucus has moved over and beyond the cells in question. In this view, mucociliary transport action is initiated by mechanical events. Further, it is recognized that mucus production results from concerted function of not one but several cellular types: secretory cells that produce the raw materials and other epithelial cells which are responsible for the postsecretory formation of a hydrogel. *See* EPITHELIUM.

Soft tissue. Soft tissues are often very complex in composition. A connective tissue is composed of structural proteins such as collagen and elastin, ground substances such as glycosaminoglycans, and cellular components. If the properties of individual components are known and the structure of the tissue is determined, then the mechanical properties of the tissue can be computed according to appropriate models. In practical applications, physical, chemical, and biological interventions may be introduced to influence the individual components or the structure as a whole in order to control the properties of the tissue. These studies have important applications to surgery, rehabilitation, and control of hypertrophy and resorption. *See* COLLAGEN; CONNECTIVE TISSUE.

Muscle. The rheological properties of muscles are of fundamental importance to medicine. Constitutive equations are not available, but with a constitutive equation for the active contraction of the heart muscle, a rigorous theory of cardiology could be developed. Similarly, a constitutive equation for vascular smooth muscle would allow hypertension and circulatory regulation to be analyzed. A phenomenological approach has not been very successful for these muscles. Since the advent of the sliding filament model, most researchers have focused their attention on the study of the actin and myosin molecules and cross bridges between them in order to understand the generation of force and motion. Other researchers study the collagen fibers in the muscle bundle which control the mechanical properties of the muscle when it is in a resting state, and the constitutive equation when it is contracting. The difficulty lies in the nonlinear behavior of these fibers, so that the contributions from different sources cannot be simply added. *See* MUSCLE; MUSCLE PROTEINS.

Motility of leukoytes. The interior structure of leukocytes is nonhomogeneous. Under suitable stimulation, sol-gel transformation takes place in the cell at various localities. The basic mechanism seems to be closely associated with actin molecules, that is, stress fibers, in the cells. Leukocytes and platelets contain a protein, gelsolin, which binds calcium ion (Ca^{2+}) and has powerful effects on actin: in the presence of micromolar Ca^{2+}, gelsolin severs actin filaments. Leukocyte rheology is of great importance in medicine. *See* BLOOD.

Thrombus formation and dissolution. Blood may form a thrombus in response to injury on the inner wall of a blood vessel or the presence of a cardiovascular device. This function is normally a life-saving process in that it seals up blood vessels that are injured as a result of cuts and bruises. It is, however, also implicated in atherosclerosis and stroke, and it is a major problem for successful implantation of artificial internal organs. The molecular mechanisms of thrombus formation and dissolution involve a change of blood rheology and are very complex. *See* ARTERIOSCLEROSIS; THROMBOSIS.

Cell membrane. Since a living cell contacts other cells, the importance of cell membrane biorheology is evident. The subject includes not only the elasticity, viscoelasticity, and viscoplasticity of the cell membrane but also adhesion and other interaction between membranes. *See* CELL ADHESION; CELL MEMBRANES.

Macrobiorheology. On the other end of the scale from microbiorheology are the macroscopic phenomena, such as blood flow, breathing, micturition, heartbeat, and intestinal peristalsis. Many important aspects of these phenomena depend on the rheological properties of the tissues. For example, the flow separation in arteries at points of bifurcation is a phenomenon intimately related to plaque

formation in atherosclerosis. In this case the elasticity of the wall and the non-newtonian rheology of the blood have considerable influence on the separation process. Such problems are often treated in biomechanics.

Constitutive equations. In biorheology, so-called constitutive equations are used to describe the complex mechanical behavior of materials in terms of mathematics. At least three kinds of constitutive equations are needed: those describing stress-strain relationships of material in the normal state of life; those describing the transport of matter, such as water, gas, and other substances, in tissues; and those describing growth or resorption of tissues in response to long-term changes in the state of stress and strain. The third type is the most fascinating, but there is very little quantitative information available about it except for bone. The second type is very complex because living tissues are nonhomogeneous, and since mass transport in tissues is a molecular phenomenon, it is accentuated by nonhomogeneity at the cellular level. The best-known constitutive equations are therefore of the first kind. Some examples of these equations are described below.

Blood. If blood is tested in a Couette viscometer (a rotational viscometer) or a cone-plate viscometer, it is found that in a steady flow the shear stress τ is approximately related to the shear strain rate by Casson's equation (1), where the constant τ_y has the

$$\sqrt{\tau} = \sqrt{\tau_y} + \sqrt{\eta\dot{\gamma}} \tag{1}$$

significance of a yield stress and the constant η is called Casson's coefficient. This relation reflects the deformability of blood cells in shear flow, and the ability of the red cells to stick to each other and form rouleaux (stacks of erythrocytes) when the shear gradient decreases. Equation (1) does not account for the normal stress due to shear, nor for other viscoelastic features of the blood. The mathematical description of blood viscoelasticity is quite complex, but it is important because the viscoelasticity affects flow separation of blood at points of bifurcation of the arteries or at sites of severe stenosis. Hence it is implicated in atherogenesis, and is relevant to technical problems of prosthetic heart valves and implantable artificial organs.

Red blood cells are filled with hemoglobin solution, which in normal blood is a newtonian fluid with a viscosity of about 6 centipoise. These cells are severely deformed when flowing through capillaries, but when stationary they assume the regular shape of a biconcave circular disk. Normal human red cells have an average disk diameter of about 7.8 micrometers, a maximum thickness of about 2.6 μm, a minimum thickness of 0.8 μm, and a wall (membrane) thickness of about 0.005 μm. For such a thin shell, mechanical analysis shows that the discoid shape implies that the membrane stress resultant in the cell membrane is zero in the normal condition, and that the membrane stress remains small in deformed states. Hence, a great engineering marvel is achieved by a simple design of geometric shape: the discoid shape makes the cells flexible and unstressed in severe deformation. Strength is achieved by flexibility.

By extensive study of cell deformation in response to external loading, it has been determined that the shear modulus of elasticity of the red cell membrane is about 10,000 times smaller than the elastic modulus for the change of membrane surface area. Hence the area of the membrane does not change significantly when the cell deforms. In fact, the membrane breaks when its area is increased by approximately 5% above the normal, but the cell membrane can be stretched several hundred percent in any direction if the area is preserved.

The interior material of the white blood cells is viscoelastic, and sol-gel transformation takes place when pseudopods form. Chemotaxis of leukocytes is related to the change in rheological properties of the cellular material, and it is most likely that the formation, movement, and desolution of the vesicles and their membranes cause these rheological changes.

Blood vessels. All living tissues are pseudoelastic; that is, they are not elastic, but under periodic loading a steady-state stress-strain relationship exists which is not very sensitive to strain rate. The loading and unloading branches can be treated separately, and a pseudoelastic potential can be introduced to describe the stress-strain relationship in either loading or unloading. The pseudoelastic potential $\rho_0 W$ (where W is defined for unit mass of the vessel wall and ρ_0 is the density in the initial state) is a function of Green's strain components E_{ij}, and is the strain energy per unit initial volume. The partial derivatives of $\rho_0 W$ with respect to E_{ij} give the corresponding stresses S_{ij} (Kirchhoff stress). For arteries and veins the pseudostrain-energy functions can be expressed either in the form of a polynomial, or as an exponential function, Eq. (2), where Q is a function of the strain components shown in Eq. (3). The subscripts

$$\rho_0 W = Ce^Q \tag{2}$$

$$Q = a_1E_1^2 + a_2E_2^2 + a_3E_3^2 + 2a_4E_1E_2 + 2a_5E_2E_3 + 2a_6E_3E_1 \tag{3}$$

1, 2, 3 in Eq. (3) refer to the circumferential, longitudinal, and radial directions respectively, and E_1, E_2, E_3 are shorthand for the strain components E_{11}, E_{22}, E_{33} respectively. Cylindrical symmetry is assumed in this expression for simplicity; otherwise one should generalize the quadratic form Q to include all six strain components. If the vessel wall can be considered homogeneous, then the constants $C, a_1, \ldots, a_6$ are independent of the location in the blood vessel. If, in addition, the arterial wall material is assumed to be incompressible, then the condition of incompressibility, Eq. (4), may be introduced through a la-

$$(1 + 2E_1)(1 + 2E_2)(1 + 2E_3) = 1 \tag{4}$$

grangian multiplier H. A new pseudoelastic function

is defined by Eq. (5) so that the Kirchhoff stresses are expressed by Eq. (6). It is well known that H has

$$\rho_0 W^* = Ce^Q + H[(1+2E_1)(1+2E_2)(1+2E_3)-1] \quad (5)$$

$$S_i = \frac{\partial(\rho W^*)}{\partial E_i} \qquad (i = 1, 2, 3) \quad (6)$$

the significance of a hydrostatic pressure which can assume an arbitrary value. The true value is determined by the equations of motion (or equilibrium) and boundary conditions.

Other soft tissues. The elastic behavior of many soft tissues, such as the skin, tendon, cartilage, mesentery, bladder, ureter, muscles, and taenia coli, is similar to that of the blood vessels in the sense that Eqs. (2)–(6) may be made to fit the experimental data. These materials become stiffer when stressed, and their incremental elastic moduli are almost linearly proportional to the stresses.

All of these soft tissues also have a fairly similar viscoelastic behavior. They all show hysteresis when subjected to cyclic loading and unloading but the hysteresis loop is fairly insensitive to the strain rate. Typically, in a thousandfold change in strain rate, the stress at a given strain in a loading (or unloading) process does not change by more than a factor of 2. When a tissue is stretched suddenly to a new dimension, which is then held constant, it will show the phenomenon of stress relaxation, that is, the stress gradually decreases with increasing time. The degree of relaxation depends on the tissue. In time, an extended ureter can relax off almost all the stress, whereas an artery can relax off only 20–30%, and a ligament nuchae relaxes off only a few percent. When the tissues are stressed suddenly by a step function and then the stress is kept constant, the tissue creeps; that is, its strain increases with increasing time. The degree of creep also depends on the tissue; but is usually fairly small in tendons, ligaments, skin, and arteries. These features, relaxation, creep, hysteresis, and strain-rate insensitivity, can be condensed in a mathematical formulation called the quasilinear viscoelasticity theory. It is nonlinear in elastic stress response to strain, linear in relating the stress at present to the elastic stress of the past, with a continuous relaxation spectrum.

Synovial and other body fluids. Most of these fluids, containing large molecules such as hyaluronic acid, behave rheologically in a manner similar to the blood. They are shear thinning, that is, their coefficients of viscosity decrease with increasing shear strain rate, and viscoelastic.

Mucus and saliva. Mucus is a complex mixture of water, electrolytes, glycoproteins, epithelial cells, and white blood cells. It is a viscoelastic material. When mucus is subjected to a small oscillatory deformation in a testing machine, the shear stress τ can be related to the shear strain γ by Eq. (7), where $G(i\omega)$

$$\tau = G(i\omega)\gamma \quad (7)$$

is the complex modulus of elasticity, i is the imaginary number $\sqrt{-1}$, and ω is the circular frequency of the oscillation. The real part of $G(i\omega)$ is called the storage modulus, and the imaginary part of $G(i\omega)$ is called the loss modulus. These names are associated with the fact that if the material obeys Hooke's law, then the strain energy stored in the material is proportional to the storage modulus; whereas if the material behaves like a newtonian viscous fluid, then the energy dissipated is proportional to the loss modulus.

Saliva is also viscoelastic, in many ways similar to bronchotracheal mucus. It has been proposed that the testing of drugs to combat the common cold, which usually aim at modifying the rheological property of mucus, may well be carried out by using saliva. *See* BIOMECHANICS; RHEOLOGY. Y. C. Fung

Bibliography. A. L. Copley and A. Silberberg (eds.), Symposia on cellular and molecular basis of mucus rheology and molecular biorheology, *Biorheology J.*, 24:535–687, 1987; E. A. Evans and R. Skalak, *Mechanics and Thermodynamics of Biomembranes*, 1980; E. Fukada and T. Azuma (eds.), Perspectives in biorheology, *Biorheology J.*, 25:1–401, 1988; Y. C. Fung, *Biomechanics: Mechanical Properties of Living Tisssues*, 2d ed., 1993; S. Oka, *Cardiovascular Hemorheology*, 1981; G. H. Pollack and H. Sugi, *Contractile Mechanisms in Muscle*, 1984.

Biosensor

An integrated device consisting of a biological recognition element and a transducer capable of detecting the biological reaction and converting it into a signal which can be processed. Ideally, the sensor should be self-contained, so that it is not necessary to add reagents to the sample matrix to obtain the desired response. There are a number of analytes (the target substances to be detected) which are measured in biological media: pH, partial pressure of carbon dioxide (pCO_2), partial pressure of oxygen (pO_2), and the ionic concentrations of sodium, potassium, calcium, and chloride. However, these sensors do not use biological recognition elements, and are considered chemical sensors. Normally, the biological recognition element is a protein or protein complex which is able to recognize a particular analyte in the presence of many other components in a complex biological matrix. Recently this definition has been expanded to include oligonucleotides. The recognition process involves a chemical or biological reaction, and the transducer must be capable of detecting not only the reaction but also its extent. An ideal sensor should yield a selective, rapid, and reliable response to the analyte, and the signal generated by the sensor should be proportional to the analyte concentration.

Biosensors are typically classified by the type of recognition element or transduction element employed. A sensor might be described as a catalytic biosensor if its recognition element comprised an enzyme or series of enzymes, a living tissue slice (vegetal or animal), or whole cells derived from microorganisms such as bacteria, fungi, or yeast. The sensor

might be described as a bioaffinity sensor if the basis of its operation were a biospecific complex formation. Accordingly, the reaction of an antibody with an antigen or hapten, or the reaction of an agonist or antagonist with a receptor, could be employed. In the former case, the sensor might be called an immunosensor.

Catalytic biosensors. Historically, the first biosensor was based on the measurement of an enzyme-catalyzed reaction. This device, described by L. C. Clark and C. Lyons in 1962, involved the incorporation of glucose oxidase into a membrane. This enzyme catalyzes the highly selective oxidation of glucose by oxygen to form gluconic acid and hydrogen peroxide, as in reaction (1). The reaction rate is ren-

$$\text{Glucose} + O_2 \rightarrow \text{gluconic acid} + H_2O_2 \qquad (1)$$

dered proportional to the concentration of glucose, and it is measured by monitoring the formation of hydrogen peroxide. This is accomplished by applying a controlled potential between two electrodes. There is produced a current which is proportional to the peroxide concentration and hence to the concentration of glucose. This device is called an enzyme electrode, because an enzyme is the recognition element and the transducer an electrode. This is by far the most important biosensor, largely because it is used extensively for clinical monitoring of blood glucose in the detection and treatment of diabetes. A potential is applied between a silver and a platinum electrode, and the resulting current from the peroxide produced in the oxidase enzyme membrane is measured. A cellulose acetate membrane and a polycarbonate membrane (see **illus.**) are used to screen out species that might interfere with analyte detection. Additional so-called oxidase enzymes catalyze the oxidation of other analytes (enzyme substrates) in a manner analogous to glucose. The two most important are lactate and glutamate. Lactate is used clinically in monitoring patients following myocardial infarction, and in sports medicine to assess exercise which can delay the buildup of lactate, a source of the pain of exertion. Monitoring of glucose, lactate, and glutamate are important in the food processing industry and in aiding the control of fermentation processes in bioreactors. *See* ANALYTICAL CHEMISTRY.

Alternatively, reaction (1) can be coupled with reaction (2) [catalyzed by a second enzyme, per-

$$H_2O_2 + \text{MBTH} + \text{DMAB} \rightarrow \text{MBTH} - \text{DMAB (blue)} \qquad (2)$$

oxidase], which generates a color to which the concentration of glucose can be related. MBTH is 3-methyl-2-benzothiazolinone hydrazone, and DMAB is 3-dimethylaminobenzoic acid. This principle forms the basis for the majority of "fingerstick" home monitoring blood glucose systems used by diabetic patients. The intensity of the color development is evaluated using a simple device which measures the light reflected off the surface where the reaction occurs.

Enzyme-based biosensors are not restricted to oxidases. A large class of enzymes known as dehydrogenases also catalyze the oxidation of analyte substrates without using oxygen. For example, a glucose dehydrogenase is used in one of the home monitoring blood glucose systems. Sometimes more than one enzyme is needed to convert the analyte into a form suitable for detection. The determination of acetylcholine, a key neurotransmitter, can be effected in two steps: the conversion of acetylcholine to choline, catalyzed by the enzyme acetylcholinesterase, followed by the oxidation of choline by oxygen using choline oxidase. Organophosphates such as Malathion (insecticide) and organophosphonates such as Sarin (nerve gas) are toxic to living organisms because they inhibit acetylcholinesterase activity. Thus this same system can be used to determine the inhibition of enzyme activity instead of substrate concentration, making it suitable for the detection of insecticides or nerve gas. Enzyme-based sensors are not restricted to oxidation-reduction reactions. Urease is used to hydrolyze urea to yield ammonia and carbon dioxide. The urea concentration can be related to the resulting pH change due to the reaction.

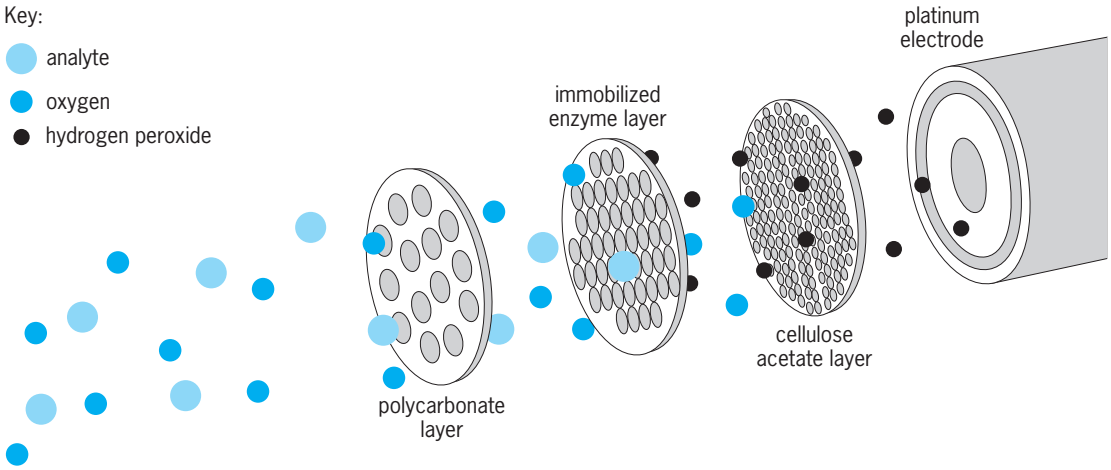

Biosensor employed in a commercially available clinical glucose analyzer. (*Yellow Springs Instrument Co. Inc.*)

Tissue slices and bacterial cultures have also formed the basis for electrochemical biosensors. Because they comprise many enzymes, the reactions are complex. However, they are useful for estimating the toxicity of a sample (inhibition of activity) and for such measurements as biochemical oxygen demand (BOD), an important measure of wastewater quality.

Since enzyme-based sensors measure the rate of the enzyme-catalyzed reaction as the basis for their response, any physical measurement which yields a quantity related to this rate can be used for detection. The enzyme may be immobilized on the end of an optical fiber, and the spectroscopic properties (absorbance, fluorescence, chemiluminescence) related to the disappearance of the reactants or appearance of products of the reaction can be measured. Since biochemical reactions can be either endothermic (absorbing heat) or exothermic (giving off heat), the rate of the reaction can be measured by microcalorimetry. Miniaturized thermistor-based calorimeters, called enzyme thermistors, have been developed and widely applied, especially for bioprocess monitoring.

Affinity biosensors. The highly selective and high-affinity binding of a molecule to a receptor forms the basis for a relatively new class of biosensors. The most common bioaffinity sensors are based on an antibody (Ab) reaction with an antigen (Ag), such as a protein, peptide, or oligonucleotide, or with a hapten, which is a small organic molecule such as a drug. Accordingly, the purpose of the biosensor is to detect the extent of the binding process and relate it to the concentration of the analyte. Unfortunately, detection of binding is not simple. An increase (albeit extremely small) in the mass of the sensor can be used as the basis for detection of binding, and indeed sensors based on the quartz crystal microbalance have been developed. The antibody is immobilized on the sensor surface, and when the antibody-antigen reaction takes place, an increase in mass results, forming the basis for an immunoassay. A second approach, called surface plasmon resonance (SPR), exploits the change in refractive index of a bioaffinity coating on a gold film when binding takes place. Biospecific interaction analysis (BIA), based on this approach, has proven useful for fundamental studies of interactions between well-defined reactants in a simple matrix. Unfortunately, it shares with the microbalance the problem that any interaction with the surface can result in a change in signal, thus making the detection of binding an inherently nonspecific method.

The alternative to the above approaches is to label one of the reactants with a fluorescent tag or an enzyme which generates a product that can be detected by its fluorescence, absorbance, chemiluminescence, or electroactivity. For example, an antibody may be immobilized on an optical-fiber outer surface and reacted with a fluorescent-labeled antigen. An appropriately chosen light beam is passed down the fiber so that it couples optically with the fluorescent label on the outer surface (the evanescent wave). Next, the sample containing the antigen is introduced, and it displaces the labeled antigen, causing a reduction in signal. The extent of this reduction can be related to the concentration of the antigen (analyte) in the samples. Since the evanescent wave samples only the immediate vicinity of the surface, species present in the bulk sample do not interfere. Other bioaffinity systems include deoxyribonucleic acid (DNA) hybridization and purified and reconstituted membrane receptors, including the glutamate receptor ion channel and the nicotinic acetylcholine receptor.

As in the case of the catalytic biosensors, many physical techniques can be used to detect affinity binding: microcalorimetry (thermometric enzyme-linked immunosorbent assay, or TELISA), fluorescence energy transfer, fluorescence polarization, or bioluminescence.

The quality of the results obtained from sensors based on biological recognition elements depends most heavily on their ability to react rapidly, selectively, and with high affinity. Antibodies and receptors frequently react with such high affinity that the analyte does not easily become unbound. To reuse the sensor requires a time-consuming regeneration step. Nonetheless, if this step can be automated, semicontinuous monitoring may be possible. Successful use of biosensors requires detailed understanding of their operating principles.

George S. Wilson

Bibliography. A. E. G. Cass (ed.), *Biosensors: A Practical Approach*, Oxford University Press, 1990; A. J. Cunningham, *Introduction to Bioanalytical Sensors*, John Wiley, 1998; E. A. H. Hall, *Biosensors*, Prentice Hall, 1991; A. P. F. Turner, I. Karube, and G. S. Wilson (eds.), *Biosensors: Fundamentals and Applications*, Oxford University Press, 1987.

Biosphere

The thin film of living organisms and their environments at the surface of the Earth. Included in the biosphere are all environments capable of sustaining life above, on, and beneath the Earth's surface as well as in the oceans. Consequently, the biosphere overlaps virtually the entire hydrosphere and portions of the atmosphere and outer lithosphere.

Neither the upper nor lower limits of the biosphere are sharp. Spores of microorganisms can be carried to considerable heights in the atmosphere, but these are resting stages that are not actively metabolizing. A variety of organisms inhabit the ocean depths, including the giant tubeworms and other creatures that were discovered living around hydrothermal vents. Evidence exists for the presence of bacteria in oil reservoirs at depths of about 6600 ft (2000 m) within the Earth. The bacteria are apparently metabolically active, utilizing the paraffinic hydrocarbons of the oils as an energy source. These are extreme limits to the biosphere; most of the mass of living matter and the greatest diversity of organisms are within the upper 330 ft (100 m) of the lithosphere

and hydrosphere, although there are places even within this zone that are too dry or too cold to support much life. Most of the biosphere is within the zone which is reached by sunlight and where liquid water exists.

Origin. For over 50 years after Louis Pasteur disproved the theory of spontaneous generation, scientists believed that life was universal and was transplanted to other planets and solar systems by the spores of microorganisms. This is the theory of panspermia. In the 1920s the modern theory of chemical evolution of life on Earth was proposed independently by A. I. Oparin, a Soviet biochemist, and J. B. S. Haldane, a British biochemist. The basic tenet of the theory is that life arose through a series of chemical steps involving increasingly complex organic substances that had been chemically synthesized and had accumulated on the prebiotic Earth. The first organisms originated under anaerobic conditions in an atmosphere devoid of free molecular oxygen. Because they were incapable of aerobic photosynthesis, a complex process which evolved undoubtedly after life originated, the oxygen in the present atmosphere arose as a secondary addition.

The first organisms appeared probably prior to 3.5 billion years ago (Ga) and possibly earlier than 3.8 Ga. Carbon (C) isotopes provide one piece of evidence pinpointing the time of emergence of the first cells. The $^{13}C/^{12}C$ ratio of the insoluble organic carbon, or kerogen, in sedimentary rocks from Greenland extending about 3.8 Ga is within the range of reduced carbon found in all younger rocks; this is an indication that organisms, which are responsible for the isotopic fractionation of carbon, had already appeared.

These approximately 3.8-billion-year-old sedimentary rocks from the Isua region of western Greenland contain no evidence that life existed at this time other than that provided by their carbon isotopic composition. These rocks, however, demonstrate that liquid water was present and that the processes of erosion, transportation, and deposition of sediments were occurring, all indicative of conditions on the Earth conducive to life-forms. If life did exist by this time, the preceding period of chemical evolution leading up to the first cells lasted for about 0.3–0.5 billion years, or about the time span between 3.8 Ga and the cooling of the Earth's surface sufficiently to contain liquid water, an event that may have occurred within a few hundred million years of the formation of the Earth about 4.6 Ga. *See* ORGANIC EVOLUTION; PREBIOTIC ORGANIC SYNTHESIS.

The fossil record provides more direct evidence for the existence of early life-forms. The oldest known microfossils are from the early Archean Apex Chert and Towers Formation (both about 3.5 billion years old) in Western Australia. The microfossils from the Apex Chert are present in siliceous grains up to a few millimeters in diameter that were apparently eroded from an even older deposit and subsequently incorporated into the Apex Chert. This deposit contains a diverse assemblage of filamentous microfossils resembling cyanobacteria that were apparently capable of oxygen-producing photosynthesis. The discovery of these fossils establishes that the early Archean biota was quite diverse, and it suggests that cyanobacterial photoautotrophs capable of oxygen production may have already existed this early in Earth history. The time required for these advanced prokaryotic organisms to evolve means that the first primitive cells must have appeared much earlier. *See* ARCHEAN; CYANOBACTERIA; FOSSIL; PROKARYOTAE.

Evolution. The changes wrought during the evolution of the biosphere had profound effects on the atmosphere, hydrosphere, and outer lithosphere.

Photosynthesis. Probably the first significant event in the evolution of the biosphere was the development of photosynthesis, a process that led to autotrophic organisms capable of synthesizing organic matter from inorganic constituents. This evolutionary advance freed the early organisms from a dependence on the coexisting abiogenic organic matter, which was being depleted by heterotrophic anaerobes. These autotrophs were anaerobic organisms similar to modern photosynthetic bacteria, and they did not release oxygen as a by-product. Some time later, probably by about 3.5 Ga with the appearance of cyanobacteria, oxygen-releasing photosynthesis, another major milestone in the development of the biosphere, evolved. *See* PHOTOSYNTHESIS.

Eukaryotes. A profound evolutionary event was the emergence of eukaryotic organisms. All early life-forms were prokaryotic, as are modern bacteria and cyanobacteria; all other organisms are eukaryotic, characterized by the presence of a nucleus and capable of reproducing by meiosis and mitosis. The origin of eukaryotic organisms was an important evolutionary event, as it represented the emergence of sexual reproduction and set the stage for the later evolution of multicellular organisms. Recent evidence, based on molecular fossils, now places the time of first appearance of eukaryotic organisms as long as 2.7 Ga. This is about 0.5–1.0 billion years before eukaryotic remains are found in the fossil record. *See* EUKARYOTAE.

Animals. The final major evolutionary advance leading to the development of the modern biosphere was the rise of multicellular animals (Metazoa). Although a diverse metazoan assemblage (the so-called Ediacaran fauna) first appeared in the fossil record in the latest Precambrian Eon about 600 million years ago (Ma), the origin and early evolutionary stages of these organisms must predate this time by a few hundred million years. Because these early metazoans had soft bodies that are not readily preserved as fossils, it may never be known precisely how and when the animal ancestors of humans evolved. Within a short span of geologic time after their first appearance, however, metazoans developed the ability to form shells of calcium carbonate and calcium phosphate, which can be easily preserved as fossils. The dramatic increase in fossils of multicellular organisms at the beginning of the Cambrian Period (about 570 Ma) is due to the preservability of their exoskeletons. Eventually, both plants and animals escaped the confines

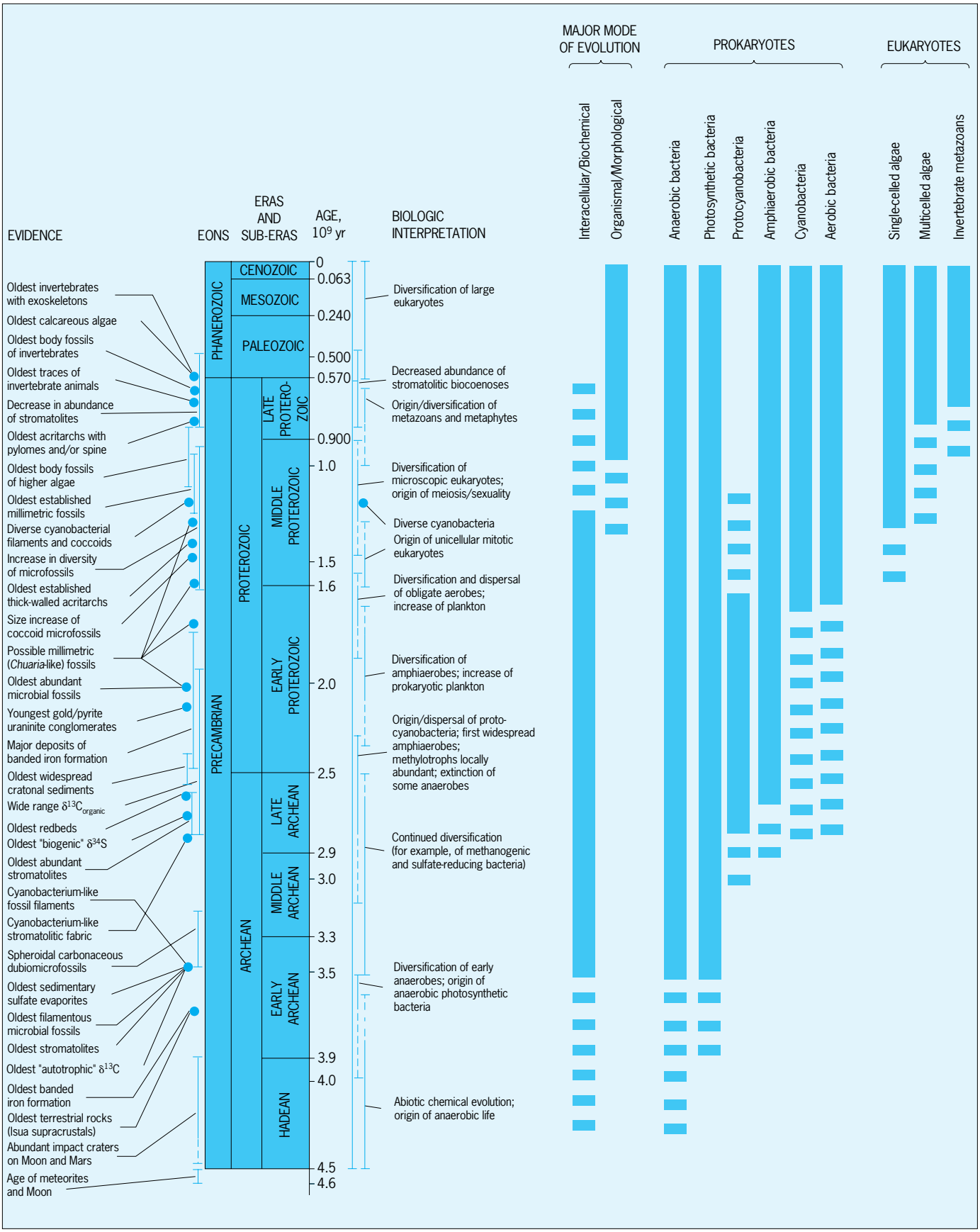

Geochronological summary of evidence relating to the origin and evolution of the earliest biosphere. (*After J. W. Schopf et al., eds., Earth's Earliest Biosphere: Its Origin and Evolution, Princeton University Press, 1983*)

of the oceans and colonized the land surface during the Silurian-Devonian periods (about 400 Ma). *See* CAMBRIAN; DEVONIAN; EDIACARAN BIOTA; METAZOA; PRECAMBRIAN; SILURIAN; STROMATOLITE.

Geochemical cycles. Organisms have had an effect on the geochemistry of their environment for at least 3.5 billion years. Prior to the development of oxygen-releasing photosynthesis, the Earth's surface was anoxic, all organisms were anaerobic, and chemical elements existed in their reduced forms. The gradual release of free oxygen permanently changed the Earth's surface from a reducing environment to an oxidizing one. Elements such as iron and sulfur were converted to an oxidized state, and their solubilities were changed. The oxidation of iron resulted in precipitation of extensive deposits of banded iron formations from 3.8 to 1.8 Ga, although most of these formed between 3.0 and 2.0 Ga. Scientists consider that cessation of banded iron formation deposition about 1.8 Ga represents the time when the Earth's surface became fully oxic and all the reduced iron dissolved in the oceans was completely oxidized. Major events in Precambrian evolution are shown in the **illustration.** *See* BIOGEOCHEMISTRY.

Extinctions. Virtually from the time of appearance of the first organisms, evolution of new genera and species was inevitably accompanied by extinction of others that were not able to adapt to different environments, to changing climatic conditions, or to new predators. Superimposed on this general background of continuous extinctions during the Phanerozoic Eon, the time represented by the best fossil record, are periodic occurrences of mass extinctions. During the two largest mass extinctions, at the ends of the Permian and Cretaceous periods, about 60% and 35%, respectively, of marine genera perished. Other mass extinctions had a less severe impact on the biosphere, but they resulted in extensive mortalities of genera. The cause of these mass extinctions is the subject of considerable debate among geologists and paleontologists. The mass extinction at the end of the Cretaceous Period, which included the demise of the dinosaurs, has been attributed to a catastrophic impact of a large comet or meteor with the Earth. Mass extinctions occurring at other times may have had different causes. *See* CRETACEOUS; EXTINCTION (BIOLOGY); PERMIAN.

Organisms. Although the biosphere is the smallest in mass, it is one of the most reactive spheres. In the course of its reactions, the biosphere has important influences on the outer lithosphere, hydrosphere, and atmosphere. *See* ATMOSPHERE; HYDROSPHERE; LITHOSPHERE.

Ecosystem. The biosphere is characterized by the interrelationship of living things and their environments. Communities are interacting systems of organisms tied to their environments by the transfer of energy and matter. Such a coupling of living organisms and the nonliving matter with which they interact defines an ecosystem. An ecosystem may range in size from a small pond, to a tropical forest, to the entire biosphere. Ecologists group the terrestrial parts of the biosphere into about 12 large units called biomes. Examples of biomes include tundra, desert, grassland, and boreal forest. *See* BIOME; ECOLOGICAL COMMUNITIES; ECOSYSTEM.

Metabolic processes. The major metabolic processes occurring within the biosphere are photosynthesis and respiration. Green plants, through the process of photosynthesis, form organic compounds composed essentially of carbon (C), hydrogen (H), oxygen (O), and nitrogen (N) from carbon dioxide (CO_2) and water (H_2O), and nutrients, with O_2 being released as a by-product. The O_2 and organic compounds are partially reconverted into CO_2 and H_2O through respiration by plants and animals. Driving this cycle is energy from the Sun; respiration releases the Sun's energy which has been stored by photosynthesis. *See* PHOTORESPIRATION; PHOTOSYNTHESIS.

Types. The fundamental types of organisms engaged in these activities are producers, or green plants that manufacture their food through the

Annual and seasonal Net Primary Productivity of the major units of the biosphere*

	Ocean		Land
Seasonal			
April to June	10.9		15.7
July to September	13.0		18.0
October to December	12.3		11.5
January to March	11.3		11.2
Biogeographic			
Oligotrophic	11.0	Tropical rainforests	17.8
Mesotrophic	27.4	Broadleaf deciduous forests	1.5
Eutrophic	9.1	Broadleaf and needleleaf forests	3.1
Macrophytes	1.0	Needleleaf evergreen forests	3.1
		Needleleaf deciduous forests	1.4
		Savannas	16.8
		Perennial grasslands	2.4
		Broadleaf shrubs with bare soil	1.0
		Tundra	0.8
		Desert	0.5
		Cultivation	8.0
Total	48.5		56.4

*From satellite data. Values are in petagrams of carbon (1 Pg = 10^{15} g).
SOURCE: C. B. field et al., Primary production of the biosphere: Integrating terrestrial and oceanic components, *Science*, 281:239, 1998.

process of photosynthesis and respire part of it; consumers, or animals that feed on plants directly or on other animals by ingestion of organic matter; and decomposers, or bacteria that break down organic substances to inorganic products.

Productivity, the rate of formation of living tissue per unit area in a given time, is one of the fundamental attributes of an ecosystem. All organisms in the community are dependent on the energy obtained through gross primary productivity. The productivity of an ecosystem is determined by a number of environmental variables, particularly temperature and the availability of water and nutrients such as phosphate and nitrate. In marine environments, availability of light and nutrients is the limiting factor, while the productivity in terrestrial environments is limited by the availability of water and nutrients. The productivity of different ecosystems is given in the **table**. *See* BIOLOGICAL PRODUCTIVITY.

The portion of gross primary productivity that is not respired by green plants is the net primary productivity that is available to consumer and decomposer organisms. Organic matter and energy are passed from one organism to another along food chains. The steps, or groupings of organisms, in a chain are trophic levels. Green plants, the producers, occupy the first trophic level, followed by herbivores (primary consumers), primary carnivores, secondary carnivores, and tertiary carnivores. The number of links in a food chain is variable, but three to five levels are common. Only about 10% or less of the matter and energy is passed from one trophic level to the next because of the utilization of energy in respiration at each level. *See* FOOD WEB.

Not all ecosystems depend on photoautotrophs or green plants for their food and energy. One of the most exotic ecosystems consists of giant tubeworms, clams, and other organisms that thrive around hydrothermal vents, the jets of hot water gushing from portions of mid-oceanic ridges. These communities, discovered only in the late 1970s, are based on chemosynthetic bacteria that live inside the tubeworms. Since then, similar communities of organisms, although living in cold water, have been found near oil and gas seeps on the sea floor. *See* HYDROTHERMAL VENT; TUBEWORMS.

Living and dead organic matter that is not consumed by higher trophic levels is decomposed as dead tissue through bacterial action. Decomposers play a major role in the flow of matter and energy in ecosystems, because they are the final agents to release the photosynthetic energy from the organic compounds that they utilize. Decomposers recycle the chemical components, but not energy, back into the ecosystem. *See* POPULATION ECOLOGY.

The biomass, or the total mass of living organisms at one time in an ecosystem, may or may not correlate with productivity. In fact, the total terrestrial biomass, composed largely of tropical, temperate, and boreal forests, greatly exceeds that of the total oceanic biomass; yet their net primary productivities are nearly equal. An accurate tally of forested biomass, of importance for balancing the global carbon budget, has been lacking, however, because of the difficulty in estimating the total areas of forest cover and the average biomass per unit area for each forest type. Use of remote-sensing data from satellites will improve biomass estimates and will provide baseline measurements for determining whether any of the forested biomasses are increasing or decreasing as a consequence of human activities. *See* BIOMASS.

Human impact. Human beings are part of the biosphere, and some of their activities have an adverse impact on many ecosystems and on themselves. Ecologists are concerned that another mass extinction, caused by human activity, may be under way. As a consequence of deforestation, urban sprawl, spread of pollutants, and overharvesting, both terrestrial and marine ecosystems are being destroyed or diminished, populations are shrinking, and many species are dying out. Whales are one group of marine mammals that have been extensively hunted, leading to near extinction of some species such as the blue whales and the humpback whales. The decrease in commercial whaling since the late 1980s may enable these populations to recover, although decades will be required because whales have a low birthrate. In the United States, many organisms threatened with extinction are federally protected through the Endangered Species Act.

In addition to causing extinctions of some species, humans are expanding the habitats of other organisms, sometimes across oceanic barriers, through inadvertent transport and introduction into new regions. Cargo ships, for example, discharge their water ballast, teeming with plankton from the port of origin, into the waters of another port at a distant location, thereby introducing many new organisms into a similar habitat where they can thrive.

Humans also add toxic or harmful substances to the outer lithosphere, hydrosphere, and atmosphere. Many of these materials are eventually incorporated into or otherwise affect the biosphere, and water and air supplies in some regions are seriously fouled. The major types of environmental pollutants are sewage, trace metals, petroleum hydrocarbons, synthetic organic compounds, and gaseous emissions. *See* HAZARDOUS WASTE.

Sewage. Bulk sewage is generally disposed of satisfactorily and is normally not a problem unless it contains one or more of the other types of substances. Under certain conditions, either when the load of organic matter is at a high level or nutrients lead to eutrophication, dissolved oxygen is depleted, creating undesirable stagnation of the water. Just such an occurrence exists during the summer months off the mouth of the Mississippi River in the Gulf of Mexico, where a large volume of water near the bottom (the "dead zone") has a dissolved oxygen concentration that is too low to support shellfish and fin fish. Scientists consider the cause of the dead zone to be high productivity in the surface waters of the Gulf due to runoff of fertilizer nutrients from farms in the Mississippi River drainage basin, combined with lower dissolved oxygen concentration in the

water column during warmer months. *See* EUTROPHICATION; SEWAGE TREATMENT.

Trace metals. Some metals are toxic at low concentrations, while other metals that are necessary for metabolic processes at low levels are toxic when present in high concentrations. Examples of metals known to be toxic at low concentrations are lead, mercury, cadmium, selenium, and arsenic. Industrial uses and automobile exhaust are the chief sources of these metals. Additions of these metals through human activity are, in some cases, about equal to the natural input. Great increases of metals are observed near major sources. Plants and soils near highways or lead-processing facilities contain as much as 100 times the unpolluted value of lead. Introduction of unleaded gasoline and passage of legislation requiring safer disposal practices and clean-up of contaminated land should reduce the danger from trace-metal contamination.

Petroleum hydrocarbons. Petroleum hydrocarbons from runoff, industrial and ship discharges, and oil spills pollute coastal waters and may contaminate marine organisms. Oil slicks smother marine life by preventing the diffusion of oxygen into the water or by coating the organisms with a tarry residue. Aromatic constituents of petroleum are toxic to all organisms. Large doses can be lethal, while sublethal quantities can have a variety of physiological effects, especially carcinogenicity. The hydrocarbons are lipophilic and, because they are not metabolized or excreted, tend to accumulate in fatty tissues. Even when marine life has not been subjected to massive contamination by an oil spill, the organisms may be gradually contaminated by accumulating the hydrocarbons to which they have been continuously exposed at low concentrations. Contamination of shellfish, in particular, is of concern because they are consumed in great quantities by humans. *See* WATER POLLUTION.

Synthetic organic compounds. Manufacture and use of synthetic organic compounds, such as insecticides and plasticizers, are commonplace worldwide. Many of these chemicals are chlorinated hydrocarbons, a category of substances that is known to be generally toxic to organisms. Of even greater concern is that chlorinated hydrocarbons are carcinogenic and long-term exposure to sublethal doses increases the risk of cancer. The widespread use of chlorinated hydrocarbons in industrial society leads to an increasing background of these compounds in natural waters, including drinking water.

Synthetic organic compounds and petroleum hydrocarbons are generally not metabolized or excreted by organisms. Both groups tend to be retained in the fatty tissues and passed on to other organisms in the next level of the food chain. Consequently, predators high in the food chain are exposed to much greater concentrations of toxic or carcinogenic substances through their ingested food than is present at background levels in the environment. Prior to banning its widespread use in the United States (use is still legal in many other countries), DDT was found in concentrations as high as tens of parts per million in fish-eating birds such as gulls and grebes, even though the concentration of DDT in the lower organisms of their food chain and in the bottom sediment was as low as 0.01 ppm. The high concentrations led to thinning of egg shells and decline in reproduction rate in many predator bird species, including the bald eagle. Due to the ban on DDT use in the United States, as well as other protective measures of the Endangered Species Act, the eagle population has increased and is no longer considered endangered, although the bird remains on "threatened" status. *See* ENDANGERED SPECIES; INSECTICIDE; MUTAGENS AND CARCINOGENS.

Gaseous emissions. Gaseous emissions to the atmosphere from industrial sources and automobiles represent both short-term and long-term hazards. Short-term problems include the addition of sulfurous gases leading to more acidic rainfall, the emission of nitrogen oxides which form smog, and carbon monoxide from automobile exhaust which leads to impairment of breathing in cities. *See* SMOG.

Of more long-term concern is the potential impact on global climate due to the increase in CO_2, methane, chlorofluorocarbons, and nitrous oxide in the atmosphere. These gases absorb heat radiated from the Earth—the so-called greenhouse effect—which will lead to a gradual warming of the Earth's surface and the attendant problems of shifting of climatic zones, upsetting of ecosystem balances, partial melting of polar icecaps, and possibly other unanticipated effects. Combustion of fossil fuels and deforestation are the major sources of the CO_2, the concentration of which has been increasing at the rate of about 1 part per million by volume each year since measurements began in the late 1950s. Since the early 1960s the increase in CO_2 correlates precisely with the growth of world population.

Only about one-half of the estimated annual release of CO_2 stays in the atmosphere. A portion of the other half dissolves in the ocean. The fate of the remainder, which cannot be found either in the atmosphere or in the ocean, has been the subject of considerable dispute. One view is that portions of the biosphere, largely temperate or boreal forests, are converting the balance of the CO_2 into new growth. Another view is that clearing of tropical forests adds a significant additional amount of CO_2 to the atmosphere, so that the biosphere is a net source of CO_2 rather than a possible sink for CO_2. The problem with both scenarios is that estimates of total forest biomass cannot be made accurately enough to establish whether this biomass is increasing or decreasing.

Atmospheric methane, another so-called greenhouse gas, has doubled during the last few hundred years. Methane is produced naturally by bacterial decomposition of organic matter in the absence of oxygen, with major sources being rice paddies, natural wetlands, and animal digestive tracts (termites, for example, are a significant source). Like that of CO_2, the increase in atmospheric methane may also be due to population-related environmental pressures as forests are cleared and rice paddies are expanded.

The synthetic organic compounds called chlorofluorocarbons, also known as the freons, used in aerosol sprays and refrigeration equipment, are both greenhouse gases and a threat to the Earth's ozone layer, which shields organisms from lethal ultraviolet radiation. Scientists predict that a partial depletion of the ozone layer will cause an increase in the incidence of skin cancer and a decrease in crop and seafood yields due to the increased dosage of damaging ultraviolet radiation. Concern about the potential destruction of the ozone layer has escalated with the discovery of a large hole in it centered over Antarctica. This concern led to the signing in 1987 of an international ozone protection treaty, known as the Montreal Protocol, since revised and updated three times, in which the developed countries agreed to phase out the production of chlorofluorcarbons and other chemicals that destroy ozone by the end of the twentieth century.

Nitrous oxide (N_2O), derived by nitrification and denitrification reactions in fertilized soils, sewage, and animal wastes, is another agent leading to depletion of ozone. Increasing reliance on chemical fertilizers to increase crop productivity is generating a larger N_2O flux to the atmosphere. *See* FERTILIZER.

Most, if not all, of the additions of potentially harmful substances to the environment are a result of the population growth and the technological advances of industrial society. The impact of these pollutants will be felt by future generations. Humans are conducting a global ecological experiment with uncertain consequences by altering the fragile equilibrium between the biosphere and its environment. *See* ACID RAIN; AIR POLLUTION; ATMOSPHERIC CHEMISTRY; GREENHOUSE EFFECT; HUMAN ECOLOGY.

Fate of biosphere. The survival of the biosphere depends on the continued availability of liquid water, CO_2, and sunlight. Insufficient amounts of any one of these will cause productivity to cease. Scientists estimate that lack of CO_2 will ultimately doom the biosphere on Earth. On a long-term basis, the amount of CO_2 in the atmosphere is maintained within broad but sufficient limits by input through volcanic emanations and removal by weathering of rocks. Scientists also estimate that solar luminosity is increasing, and that the Sun's luminosity when the Earth formed was about 30% lower than it is now. As the Sun's energy output continues to rise, the temperature of the Earth will increase, as will the rate of rock weathering. Consequently, atmospheric CO_2 will decline when the rate of rock weathering exceeds the rate of return of CO_2 to the atmosphere by volcanic emanations. Scientists estimate that in about 1 billion years the CO_2 concentration will be too low to support plant productivity and the biosphere on Earth will cease to exist. Richard M. Mitterer

Bibliography. J. J. Brocks et al., Archean molecular fossils and the early rise of eukaryotes, *Science*, 285:1033–1036, 1999; R. P. Detwiler and C. Hall, Tropical forests and the global carbon cycle, *Science*, 239:42–47, 1988; C. B. Field et al., Primary production of the biosphere: Integrating terrestrial and oceanic components, *Science*, 281:237–240, 1999; D. M. Raup and G. E. Boyajian, Patterns of generic extinction in the fossil record, *Paleobiology*, 14:109–125, 1988; M. T. Rosing, ^{13}C-depleted carbon microparticles in >3700-Ma sea-floor sedimentary rocks from West Greenland, *Science*, 283:674–676, 1999; J. W. Schopf (ed.), *Major Events in the History of Life*, 1992; J. W. Schopf, Microfossils of the Early Archean Apex Chert: New evidence for the antiquity of life, *Science*, 260:640–646, 1993; J. W. Schopf et al. (eds.), *Earth's Earliest Biosphere: Its Origin and Evolution*, 1983; T. Takahashi et al., Balancing the budget: Carbon dioxide sources and sinks, and the effects of industry, *Oceanus*, 35:18–28, 1992.

Biosynthesis

The synthesis of more complex molecules from simpler ones in cells by utilizing biosynthetic pathways consisting of a series of reactions mediated by enzymes. The overall economy and survival of the cell is governed by the interplay between the energy gained from the breakdown of compounds and that supplied to biosynthetic pathways for the synthesis of compounds having a functional role, such as deoxyribonucleic acid (DNA), ribonucleic acid (RNA), and enzymes. Biosynthetic pathways (see **illus.**) give rise to two distinct classes of metabolite, primary and secondary. Primary metabolites (DNA, RNA, fatty acids, α-amino acids, chlorophyll in green plants, and so forth) are essential to the metabolic functioning of the cells. Secondary metabolites (antibiotics, antifeedants, alkaloids, pheromones, skunk scent, and so forth) aid the functioning and survival of the whole organism more generally. Unlike primary metabolites, secondary metabolites are often unique to individual organisms or classes of organisms. *See* ENZYME; METABOLISM.

The selective pressures that drive evolution have ensured a rich and diverse array of secondary metabolite structures. Despite this diversity, secondary metabolites can be grouped (for example, as terpenes or polyketides) to some extent by virtue of their origin from key biosynthetic pathways. It is often in the latter stages of these pathways that the structural diversity is introduced. All terpenes, for example, originate from the C_5 intermediate isopentenyl pyrophosphate via mevalonic acid. The mammalian steroids, such as cholesterol, derive from the C_{30} steroid lanosterol, which is constructed from six C_5 units. Alternatively, C_{10} terpenes (for example, menthol from peppermint leaves) and C_{15} terpenes (for example, juvenile hormone III from the silk worm) are derived after the condensation of two and three C_5 units, respectively, and then with further enzymatic customization in each case. *See* CHOLESTEROL; ORGANIC EVOLUTION; STEROID; TERPENE; TRITERPENE.

The polyketide pathway has its origin in fatty acid biosynthesis. Fatty acids are assembled in a similar manner in all cells by fatty acid synthase enzymes, and they are built by the condensation of C_2 fragments derived from acetyl coenzyme A (acetyl CoA).

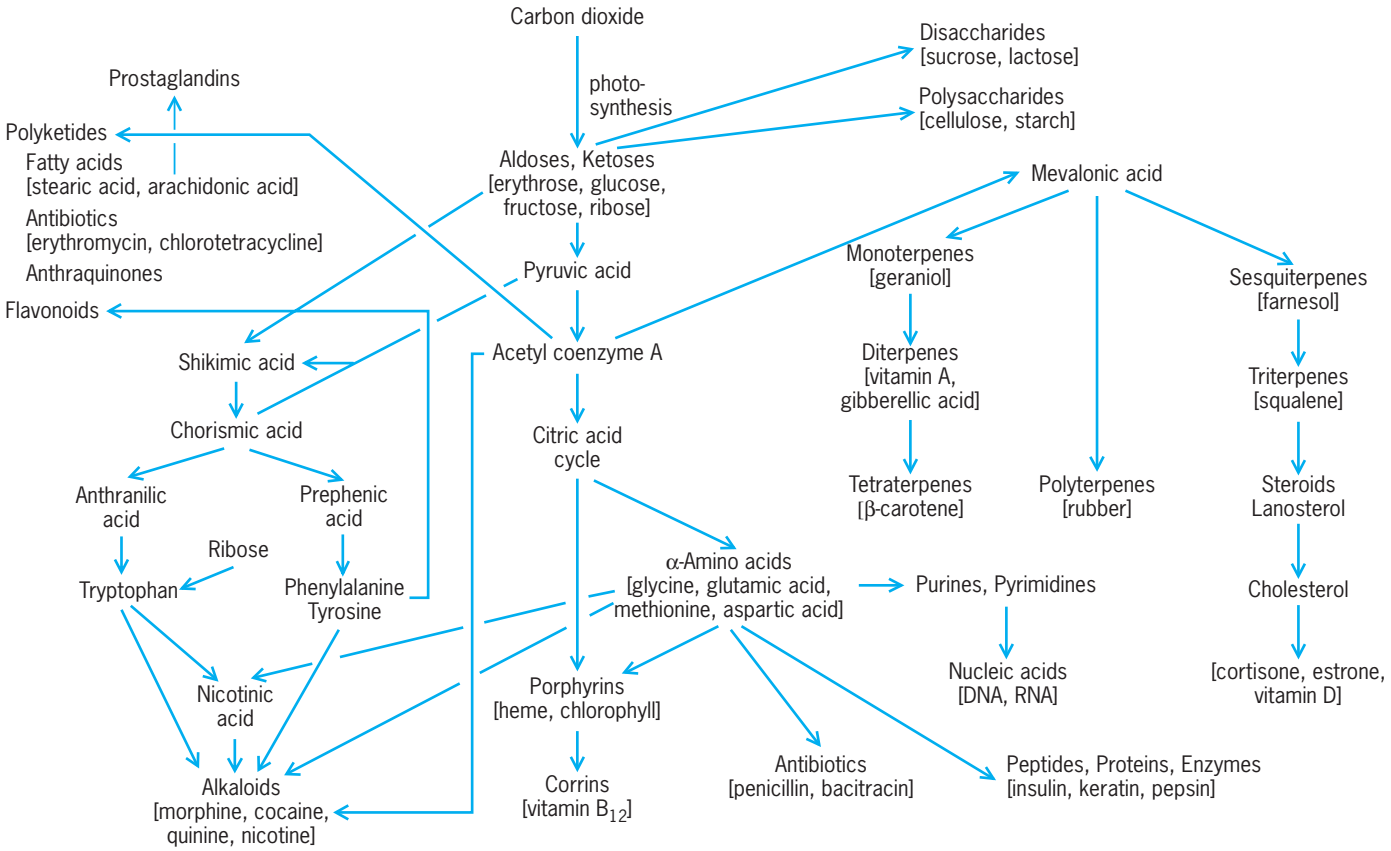

Biosynthetic relationships of some natural products. Specific examples which belong to the various classes are given in brackets.

First, acetyl CoA and malonyl-SEnz (itself derived from acetyl CoA) condense to give acetoacetyl-SEnz, an enzyme-bound C_4 intermediate (SEnz stands for S-enzyme and indicates that the acetoacetyl group is linked to the enzyme by a sulfur atom). The ketone group (C═O) of this intermediate is then removed by the action of three enzymes to generate a saturated C_4 fatty acid. This process repeats in a cyclical manner, adding C_2 units until the synthase releases a saturated fatty acid (such as C_{16}-palmitic acid or C_{18}-stearic acid) of optimal length for that synthase. *See* COENZYME; LIPID METABOLISM.

Secondary metabolic pathways have evolved that modify the fatty acid pathway, giving rise to many highly complex structures. The antibiotics chlorotetracycline and erythromycin-A, which are metabolites of Actinomycetes (soil bacteria), are such polyketides. Chlorotetracycline derives from the condensation of acetate units, but unlike fatty acids, all of the ketone groups are retained in the early stages, and are not removed. The result is an enzyme-bound polyketo-intermediate which can cyclize to form four fuzed aromatic rings. Further enzymatic elaborations then generate chlorotetracycline. Similarly, erythromycin-A is a structurally elaborate polyketide antibiotic. It is also biosynthesized by a pathway that has evolved from classical fatty acid biosynthesis. In this case, C_3 propionate units are condensed, in place of the C_2 acetate units in fatty acid assembly, and the cyclical processing of fatty acid biosynthesis is modified appropriately during each cycle to generate ketones, alcohols, double bonds, and methylene (—CH_2—) groups along the backbone of a complex precursor "fatty acid" thioester. This thioester cyclizes to a 14-membered lactone, which is then further modified to the antibiotic.

α-Amino acids combine to generate functional peptides and proteins (for example, enzymes) and are important primary metabolites for this reason. α-Amino acids also contribute to the biosynthesis of many secondary metabolites (such as alkaloids or snake venoms). There are fungi (for example, *Penicillium chrysogenum, Cephalosporium acremonium*) in which a pathway operates from three α-amino acids to the antibiotic penicillin-N. *See* ANTIBIOTIC.

Generally, biosynthetic pathways are elucidated by isotopic labeling techniques, often termed tracer studies. This technique involves introducing an early precursor or intermediate on a pathway, enriched with an isotope at a specific atomic site, into the relevant biological system. The successful or unsuccessful incorporation of the isotope into the metabolite can then give information on true intermediates on the biosynthetic pathway. Radioactive isotopes, such as carbon-14 and hydrogen-3 (tritium), emit β particles during their radioactive decay and can thus be detected. However, it has become more common to probe the details of biosynthetic

pathways with the stable isotopes of carbon-13 and hydrogen-2 (deuterium), as the location of these isotopes can be probed directly by nuclear magnetic resonance (NMR) spectroscopy. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; NUCLEAR MAGNETIC RESONANCE (NMR); PHOTOSYNTHESIS; PLANT METABOLISM; RADIOACTIVE TRACER. David O'Hagan

Bibliography. R. B. Herbert, *The Biosynthesis of Secondary Metabolites*, 2d ed., 1989; G. Zubay, *Biochemistry*, 4th ed., 1999.

Biot-Savart law

A law of physics which states that the magnetic flux density (magnetic induction) near a long, straight conductor is directly proportional to the current in the conductor and inversely proportional to the distance from the conductor. There is also a law in fluid dynamics bearing this name; it is concerned with vortex motion, and bears a close analogy to the law discussed in the present article.

The field near a straight conductor can be found by application of Ampère's law. Consider any element *dl* of the current-carrying conductor (see **illus.**). The contribution of this element to the flux density at point *P* is perpendicular to the plane of the paper and directed into the paper. *See* AMPÈRE'S LAW.

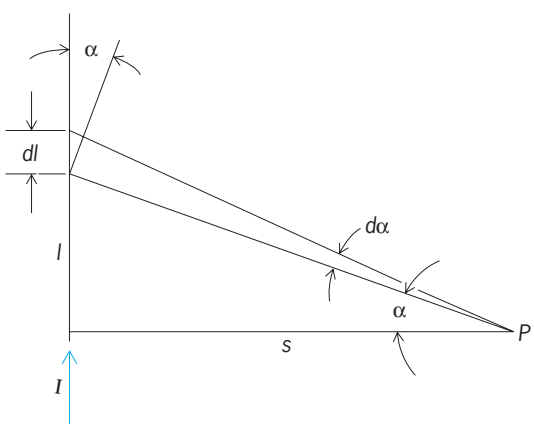

Diagram of the Biot-Savart law.

The flux density *B* at point *P* is found by summing up the contributions of all elements of the conductor from one end to the other. All these contributions are in the same direction at *P*, and hence the vector sum is the simple integral taken from one end of the wire to the other. The limits for the angle α are chosen to represent this sum. For a wire of any length, the lower end is represented by $\alpha = \alpha_1$ and the upper end by $\alpha = \alpha_2$. If *I* is the current, μ_0 the permeability of free space, and *s* the distance of *P* from the conductor, then the flux density is expressed by Eq. (1). For the special case of an infinite straight conductor, that is, one that is long in comparison to *s*, $\alpha_2 = 90^\circ$ and $\alpha_1 = -90^\circ$. Then Eq. (2) holds. This

$$B = \int dB = \frac{\mu_0 I}{4\pi s}\int_{\alpha_1}^{\alpha_2} \cos\alpha \, d\alpha$$

$$= \frac{\mu_0 I}{4\pi s}(\sin\alpha_2 - \sin\alpha_1) \qquad (1)$$

$$B = \frac{\mu_0 I}{4\pi s}\left[\sin 90^\circ - \sin(-90^\circ)\right] = \frac{\mu_0 I}{4\pi s}(1+1)$$

$$\text{or} \qquad B = \frac{\mu_0 I}{2\pi s} \qquad (2)$$

relation was originally developed from experimental observations by J. Biot and F. Savart, and was named after them as the Biot-Savart law.

The magnetic flux density *B* near the long, straight conductor is at every point perpendicular to the plane determined by the point and the line of the conductor. Therefore, the lines of induction are circles with their centers at the conductor. Furthermore, each line of induction is a closed line. This observation concerning flux about a straight conductor may be generalized to include lines of induction due to a conductor of any shape by the statement that every line of induction forms a closed path. Kenneth V. Manning

Bibliography. E. M. Purcell, *Electricity and Magnetism*, 2d ed., 2000; W. M. Schwartz, *Intermediate Electromagnetic Theory*, 1964; H. D. Young et al., *Sears and Zemansky's University Physics*, 10th ed., 1999.

Biotechnology

Generally, any technique that is used to make or modify the products of living organisms in order to improve plants or animals or to develop useful microorganisms. By this definition, biotechnology has actually been practiced for centuries, as exemplified by the use of yeast and bacteria in the production of various foods, such as wine, bread, and cheese. However, in modern terms, biotechnology has come to mean the use of cell and tissue culture, cell fusion, molecular biology, and in particular, recombinant deoxyribonucleic acid (DNA) technology to generate unique organisms with new traits or organisms that have the potential to produce specific products. The advances and products in the biotechnology arena have been developing at a rapid pace. Some examples of products in a number of important disciplines are described below.

Genetics. Recombinant DNA technology has opened new horizons in the study of gene function and the regulation of gene action. In particular, the ability to insert genes and their controlling nucleic acid sequences into new recipient organisms allows for the manipulation of these genes in order to examine their activity in unique environments, away from the constraints posed in their normal host. Transformed plants, animals, yeast, and bacterial genes may be examined in this way. *See* GENE.

Microbiology. Genetic transformation normally is achieved easily with microorganisms; new genetic material may be inserted into them, either into their chromosomes or into extrachromosomal elements, the plasmids. Thus, bacteria and yeast can be

created to metabolize specific products or to produce new products. Concomitant technologies have been developed to scale up the production of the microorganisms to generate products, such as enzymes, carbohydrates, and proteins, in great quantity. *See* BACTERIAL GENETICS; PLASMID.

Immunology. Genetic engineering has allowed for significant advances in the understanding of the structure and mode of action of antibody molecules. Practical use of immunological techniques is pervasive in biotechnology. Notably, antibodies are used in diagnostic procedures for detecting diseases of plants and animals, and in detecting minute amounts of such materials as toxic wastes, drugs, and pesticides. The antibodies themselves are being employed to target therapeutic agents to specific cellular sites. Antibodies are bivalent molecules that bind to their target molecules at one or both of their two combining sites. Hybrid antibodies are being produced in which one of the sites contains a drug or a poison while the other site directs the antibody to its target, that is, a cancerous cell. The ability to artificially combine subunits of antibodies produced in different species also will tailor them for specific targets. Antibodies are also being developed with enzymatic properties, thereby enabling them to deactivate target molecules with which they combine in a cell. *See* ANTIBODY.

Monoclonal antibodies respond to a single antigenic site, allowing for great specificity. They have been most important in the diagnostic arena, where tests have been developed for human, plant, and animal diseases and for pregnancy and ovulation prediction. *See* MONOCLONAL ANTIBODIES.

Agriculture. Few commercial products have been marketed for use in plant agriculture, but many have been tested. Interest has centered on producing plants that are tolerant to specific herbicides. This tolerance would allow crops to be sprayed with the particular herbicide, and only the weeds would be killed, not the genetically engineered crop species. Some herbicide-tolerant crop species have been isolated by selection of tissue culture variants, but in other cases plants have been transformed with genes from bacteria that detoxify the herbicide. For example, a gene isolated from the soil bacterium *Klebsiella ozaenae* has been used to create plants that are resistant to the herbicide bromoxynil. Other strategies have been to alter plants so that they will overproduce the herbicide-sensitive biochemical target, or so that the biochemical target will be altered, thereby reducing the affinity of the herbicide for its biochemical target in the crop species. *See* HERBICIDE.

Tolerance to plant virus diseases has been induced in a number of crop species by transforming plants with portions of the viral genome, in particular the virus's coat protein and the replicase enzyme used for virus multiplication. Plants exhibiting this tolerance are available for commercial use once regulatory constraints are removed. Genes coding for protein toxins from the bacterium *Bacillus thuringiensis* have been introduced into a number of plant species to provide insect tolerance. Potato plants tolerant to the Colorado potato beetle and cotton plants tolerant to the cotton boll worm were the first commercial products of this technology.

To enhance the quality and nutrition of foods, a number of techniques have been developed. Recombinant DNA strategies can be used to retard the softening of tomatoes so they can reach the consumer with better flavor and keeping qualities. A technique was developed to increase the starch content of the potato to enhance the quality of french fries and potato chips, including reduction of their capacity to absorb oils during processing. With both selective plant breeding and genetic engineering, oil seed rape (also known as canola or *Brassica napus*) has been produced to yield plant-based oils (oleochemicals) for the specialty chemicals industry. Such oils are used to produce amides, fatty acids, esters, and other chemical intermediates. Oils high in erucic acid are used as lubricants and as additives to automatic transmission fluid. Oil seed rape is also an important source of oil for margarine production and cooking. Genetic engineering techniques can be used to lower the proportion of saturated fat by inserting the gene for the stearoyl–acyl carrier protein desaturase enzyme gene into oil seed rape and other oil-producing crop plants. *See* FAT AND OIL (FOOD).

Biotechnology also holds great promise in the production of vaccines for use in maintaining the health of animals. However, because of the economics of the market, most emphasis has been placed on race horses and pets, such as cats and dogs. Subunit vaccines are being developed to replace live-virus or killed-whole-virus vaccines. Feline leukemia virus is an important disease, estimated to infect 1.5 million cats annually in the United States. Conventional vaccines are ineffective against the virus and may actually spread the disease. Subunit vaccines have been developed that overcome these disadvantages. Pseudorabies virus is the causal agent of an important disease affecting hogs. The conventional killed-virus vaccine gives only partial protection, and a live-virus vaccine is only slightly more effective. A genetically engineered vaccine has been developed in which the gene producing the enzyme that enables the virus to escape from nervous tissue has been deleted, thus reducing virus virulence. Another vaccine incorporates genes for pseudorabies proteins into viable vaccinia virus as a vector to immunize animals. Recombinant vaccines are being produced for many other diseases, including foot-and-mouth virus and parvovirus.

Interferons are also being tested for their use in the management of specific diseases. They show some promise in treating bovine shipping fever, a complex of respiratory infections manifested during the crowded conditions experienced by cattle when shipped. Recombinant produced bovine growth hormone has been shown to be effective in increasing the production of milk in cows and in reducing the ratio of fat to lean meat in beef cattle.

Animals may be transformed to carry genes from other species, including humans, and are being used to produce valuable drugs. This technology has been termed biopharming. For example, goats are being used to produce tissue plasminogen activator, which has been effective in dissolving blood clots. Transgenic animals have been produced that carry the acquired immune deficiency syndrome (AIDS) virus, providing a practical cost-effective experimental model system to test control measures for the disease, which infects only primates in nature.

Plant science. Plant scientists have been amazed at the ease with which plants can be transformed to enable them to express foreign genes. This field has developed very rapidly since the first transformation of a plant was reported in 1982, and a number of transformation procedures are available. Most widely used as a transformation vector is the plasmid derived from the plant pathogenic bacterium *Agrobacterium tumefaciens*. Another method involves shooting DNA-coated tungsten (or gold) particles into cells. Controlling genetic elements have been identified that allow for the insertion of genes into specific tissues or plant organs. For example, nucleic acid sequences may be targeted to the pollen grain to specifically inactivate genes involved in producing pollen, thereby allowing for the production of sterile male plants. Such a technique will be very useful in the commercial production of hybrid seeds of crop species.

Coupled with these transformation procedures has been the development of tissue culture techniques to enable the transformed cells to be regenerated into whole plants (termed totipotency). Many plant species have been regenerated, thereby facilitating the transfer of useful genes to the most important crops (rice, maize, wheat, and potatoes). Cultures of plant cells, roots, and tissues are used to produce secondary plant products, such as the anticancer drug taxol. *See* BREEDING (PLANT); TISSUE CULTURE.

Medicine. Genetic engineering has enabled the large-scale production of proteins which have great potential for treatment of heart attacks. Most promising of these new drugs is tissue plasminogen activator, which has been shown to be effective in dissolving blood clots. Active tissue plasminogen activator, a very minor constituent of human blood vessels and some other tissues, may now be produced by recombinant technology in transformed tissue culture cells or in the filamentous fungus *Aspergillus nidulans*. Another substance produced by recombinant technology that has been tested for treatment of heart attacks is urokinase.

Many human gene products, produced with genetic engineering technology, are being investigated for their potential use as commercial drugs. Cloned human growth hormone is being used for the treatment of childhood dwarfism. Epidermal growth factor is a protein that causes the replication of epidermal cells and has applications to wound healing. Interleukins produced by blood cells are being tested for treatment of specific cancers. Granulocyte macrophage colony–stimulating factor is being tested as a treatment for ovarian cancer.

Recombinant technology has been employed to produce vaccines from subunits of viruses, so that the use of either live or inactivated viruses as immunizing agents is avoided. Conventional vaccines are sometimes infectious in themselves, and most require refrigeration which makes their use in tropical countries a problem. Poliovirus vaccines involving attenuated live virus are a source of the perpetuation of the disease due to rare mutations of the virus to virulence. Recombinant technology is being used to modify the viral genome to prevent reversion. Other viruses being investigated are hepatitis B and herpes simplex, where surface antigens are being produced in yeast cells. Malaria, one of the most important parasitic diseases of humans, is another case in which biotechnology is being applied to produce a vaccine. Proteins from the causal organism, *Plasmodium* sp., are poorly immunogenic. To enhance immunogenicity, they are transferred to the attenuated bacterium *Salmonella typhimurium*, which is then used as an immunizing agent. Another strategy has been to fuse the gene for the immunological determinant to that of a hepatitis B virus protein, to express the fusion protein in yeast, and to use that chimeric protein as a vaccine. *See* VACCINATION.

Cloned genes and specific, defined nucleic acid sequences can be used as a means of diagnosing infectious diseases or in identifying individuals with the potential for genetic disease. The specific nucleic acids used as probes are normally tagged with radioisotopes, and the DNAs of candidate individuals are tested by hybridization to the labeled probe. The technique has been used to detect latent viruses such as herpes, bacteria, mycoplasmas, and plasmodia, and to identify Huntington's disease, cystic fibrosis, and Duchenne muscular dystrophy. In many cases, restriction-length polymorphisms are being utilized. When DNA is cut into small fragments by specific restriction enzymes and then is probed with specific genes or nucleic acids, differences between individuals in a population can be identified, and the relationships of specific patterns to specific diseases or traits can be determined. This technology is being used in many other useful ways, such as identifying important genes in plants and animals as an aid in breeding improved stocks, and as a forensic tool by assigning specific identity to individuals through their DNA in much the same way that fingerprints are now used. The technique is called DNA fingerprinting, and it is as reliable as conventional fingerprinting. The technique has the potential of distinguishing the DNA from one individual in a population of 10 billion. Tissue samples containing DNA left at the scene of a crime, such as bone, blood, semen, skin, hair (if it is attached to its root), saliva, and sweat, can be used in the procedure. If the amount of DNA found is too small to be useful, a new technique, the polymerase chain reaction, has been developed to amplify it to practical levels. *See* FORENSIC MEDICINE.

Gene functions can often be blocked by attacking them with complementary or antisense sequences

of the same gene. This technology has been used by molecular biologists to define functions for specific genes, but it has also been shown to have a number of practical applications. In agriculture it has been used to generate male sterile plants, enabling the production of hybrid varieties more easily, and to slow the ripening of tomatoes. Most importantly, antisense technology presents the possibility of useful gene therapy. For example, the human immunodeficiency virus (considered to be the casual agent of AIDS) can be inhibited by transforming T lymphocytes with antisense nucleic acids directed against a virus enzyme, reverse transcriptase.

It is now possible to put foreign genes into cells and to target them to specific regions of the recipient genome. This presents the possibility of developing specific therapies for hereditary diseases, exemplified by sickle-cell anemia, which is caused by a defect in the β-globin gene which results in defective hemoglobin in affected individuals.

Environment. Microorganisms, either genetically engineered or selected from natural populations, are used to degrade toxic wastes in the environment. For example, polycyclic aromatic compounds, such as polychlorinated biphenyls, and petroleum products which contaminate soil and ground-water supplies may be degraded by populations of microorganisms. These technologies have the potential to solve some significant environmental problems. Waste products of industry and agriculture are being composted, with added microorganisms selected for their capacity to degrade organic materials. *See* BIODEGRADATION; GENETIC ENGINEERING; MOLECULAR BIOLOGY.

Milton Zaitlin

Bibliography. M. Butler, *Mammalian Cell Biotechnology: A Practical Approach*, 1991; P. N. Cheremisinoff and L. M. Ferrante (eds.), *Biotechnology: Current Progress*, vol. 1, 1991; J. Farrington, *Agriculture Biotechnology*, 1990; G. LeWandowski, *Biotechnology Application*, 1990; P. M. Peters, *Biotechnology: A Guide to Genetic Engineering*, 1993.

Biotelemetry

The use of telemetry methods for sending signals from a living organism over some distance to a receiver. Usually, biotelemetry is used for gathering data about the physiology, behavior, or location of the organism. Generally, the signals are carried by radio, light, or sound waves. Consequently, biotelemetry implies the absence of wires between the subject and receiver.

The basic equipment required for biotelemetry includes a transmitter, a receiver, and their antennas. The receiver usually converts the incoming signal (for example, radio frequency) to an electronic signal that can be recorded or stored (for example, in a computer), or to a sound that can be monitored by a human attendant. *See* ANTENNA (ELECTROMAGNETISM); RADIO RECEIVER; RADIO TRANSMITTER.

Generally, biotelemetry techniques are necessary in situations when wires running from a subject to a recorder would inhibit the subject's activity; when the proximity of an investigator to a subject might alter the subject's behavior; and when the movements of the subject and the duration of the monitoring make it impractical for the investigator to remain within sight of the subject. Biotelemetry is widely used in medical fields to monitor patients and research subjects, and now even to operate devices such as drug delivery systems and prosthetics. Sensors and transmitters placed on or implanted in animals are used to study physiology and behavior in the laboratory and to study the movements, behavior, and physiology of wildlife species in their natural environments.

Applications. There is tremendous diversity of applications of biotelemetry, and as new electronics technology develops, biotelemetry will be applied even more broadly.

Medicine. Biotelemetry is an important technique for biomedical research and clinical medicine. Holter monitoring, or storage telemetry, was one of the earliest applications of biotelemetry to human medicine. Holter monitoring involves transmitting a signal, such as heart rate, to a receiver on a person who is free to move about. This idea is still widely used in medicine; however, with new technology many other biotelemetry methods are used in research and during treatment.

Another example of medical telemetry is the placement of a pressure transducer on the biting surface of a tooth and radio transmission of the resulting signal to provide data about tooth contact and bruxing, that is, unconscious tooth grinding that can cause severe headaches and muscle pain. The use of pressure sensors and small transmitters permits dental researchers to monitor bruxing without disrupting the individual's behavior.

Sensors placed on muscles can detect contractions, and those data can be transmitted to a remote receiver, allowing monitoring of muscle tension. To study digestive disorders, telemetry is used to monitor the muscular contractions of the intestinal tract, the movements of gut contents, and aspects of the digestive tract environment (for example, pH). Urinary incontinence can be studied by using telemetry to monitor urethral and bladder pressure, urine flow, and the effects of drug therapy. Biotelemetry allows this to be done while the individual conducts normal daily activities.

Electrode sensors with leads to transmitters are used to study the nervous system. By using telemetry, records of the brain's activity (electroencephalograms) can be obtained from epileptics. Sometimes electroencephalograms are measured in conjunction with telemetry of muscle activity to learn when epileptics experience neural dysfunction during regular activity. Also, there has been research with biotelemetry to develop nerve stimulation, including biofeedback mechanisms for people who have suffered nerve damage. Phrenic nerve stimulation, scoliosis correction, and

epilepsy control have all been investigated with the aid of biotelemetry. *See* ELECTROENCEPHALOGRAPHY.

Furthermore, telemetry methods are employed in studies of breathing abnormalities during sleep, including sleep apnea, snoring, and sudden death syndrome. Fetal and neonatal brain patterns and heart rates can easily be monitored with telemetry, thereby circumventing restrictions to the person's movements, while facilitating nursing care.

Perhaps cardiovascular research and treatment have benefited the most from biotelemetry. Heart rate, blood flow, and blood pressure can be measured in ambulatory subjects and transmitted to a remote receiver-recorder. Telemetry also has been used to obtain data about local oxygen pressure on the surface of organs (for example, liver and myocardium) and for studies of capillary exchange (that is, oxygen supply and discharge). Biomedical research with telemetry includes measuring cardiovascular performance during the weightlessness of space flight and portable monitoring of radioactive indicators as they are dispersed through the body by the blood vessels. *See* SPACE BIOLOGY.

For applied medicine, anesthesia can be evaluated during surgery by telemetric monitoring of electroencephalograms. Fetal heart rates and the mother's uterine contractions can be monitored in preparation for child delivery. After birth, the cardiopulmonary function of neonates and premature babies can be checked with telemetry. After treatment by drug therapy or coronary surgery, telemetry is useful for monitoring cardiovascular performance so that recovery and the efficacy of the treatment can be evaluated.

Telemetry has significant potential applications for controlling drug dosages from implanted delivery systems (for example, insulin pumps), adjusting cardiac pacemakers, operating prosthetics, and assessing the function of orthopedic implants.

In sports medicine, by measuring heart rate, physiologists can assess the differences in conditioning caused by interval and continuous training regimens. Telemetry also permits monitoring of athletes during actual conditions of training and competition, so it has been applied to runners, swimmers, rowers, and others. In one study the heart rate of cyclists and the physical parameters from the bicycle wheels were simultaneously recorded by telemetry so that performance could be evaluated over the turns and hills of the race course.

Animal behavior. Just as telemetry has been used in many ways to study human subjects, it has been applied widely to animal research. To avoid disturbing animals or restricting their movements unnecessarily, telemetry has been used to record electroencephalograms, heart rates, heart muscle contractions, and respiration, even from sleeping mammals and birds (**Fig. 1**). Telemetry and video recording have been combined in research of the relationships between neural and cardiac activity and behavior. Using miniature electrodes and transmitters, ethologists have studied the influence of one bird's song on the heart rate and behavior of a nearby bird.

Fig. 1. Location transmitter glued to the back of a 1-oz (30-g) thrush. Such transmitters, usually used in collars for mammals, have revolutionized animal ecology studies by allowing investigators to find and identify individual animals at will. (*From L. D. Mech, Handbook of Animal Radio-Tracking, University of Minnesota Press, 1983*)

Many studies of domestic animals have included techniques of biotelemetry. The performance of race horses has been monitored by using heart rate and respiratory data obtained with telemetry. In this way, the effects of different diets, different training methods, and the age or sex of the horse can be evaluated. Telemetry can help to provide information about the social behavior and energetics of livestock that is raised for food. For example, biotelemetry was used to monitor heart rate and body temperature, to learn that a breed of pigs was less disturbed and exhibited more control of body temperature when kept in pens with other pigs than when housed individually. These conditions resulted in better swine growth and productivity for farmers.

Biotelemetry has been applied in several ways to animals such as cattle that are allowed to wander over large areas in herds. Implanted temperature-sensing transmitters are used to monitor the breeding condition of female animals so that the likelihood of successful artificial insemination can be maximized. Individually coded transmitters are used to mark animals to census their presence in the herd, as well as to automatically dispense a specialized portion of food when each animal's signal is received at the feeding trough. Such techniques permit some attention to individual animals, without the effort needed to observe or capture each animal every day.

Ecology. Many species of wildlife are difficult to find and observe because they are secretive, nocturnal, wide-ranging, or move rapidly. Since the 1960s, biotelemetry has been a valuable technique for many animal ecology studies. Most commonly, a transmitter is placed on a wild animal so that biologists can track or locate it by homing toward the transmitted signal or by estimating the location by plotting the intersection of two or more bearings from the receiver toward the signal. For some purposes, after homing to a transmitter-marked animal, the biologists observe its behavior. For other studies, successive estimates of location are plotted on a map to describe movement patterns, to delineate the amount

of area the animal requires, or to determine dispersal or migration paths. Ecologists can associate the vegetation or other features of the environment with the locations of the animal.

Sensors can be placed on animals' transmitters to provide data about the temperature, humidity, and air pressure around the animal. The general activity of transmitter-marked animals can be related to signal modulation caused by transmitter antenna movements or to a motion-sensing mercury switch placed inside the transmitter. Implanted sensors, or entirely implanted transmitters equipped with sensors, are used to record body temperature, heart rate, pH, and muscular activity. The depths at which eels and whales swim and penguins feed have been studied with telemetry. The temperatures around transmitters on grouse have been monitored to learn when the birds roost under the snow and emerge to fly into the treetops to eat the buds from leafless winter branches. Biologists have also used telemetry to study the responses of fish to thermal pollution caused by electrical and nuclear generating plants.

Ecologists can mark numerous animals, each with a unique frequency or coded signal, and thus keep track of many individuals simultaneously without having to observe them all at once. In this way, biotelemetry can be useful for studying the social interaction among animals, such as the seasonal gathering of adults for breeding, and the proximity and duration of parent-offspring relationships. The survival of many transmitter-marked animals can be monitored, and the dead animals can be located in order to determine the cause of death. Consequently, with the tools of biotelemetry ecologists are learning more about wildlife diseases, predator-prey relationships, and the productivity and longevity of many wildlife species.

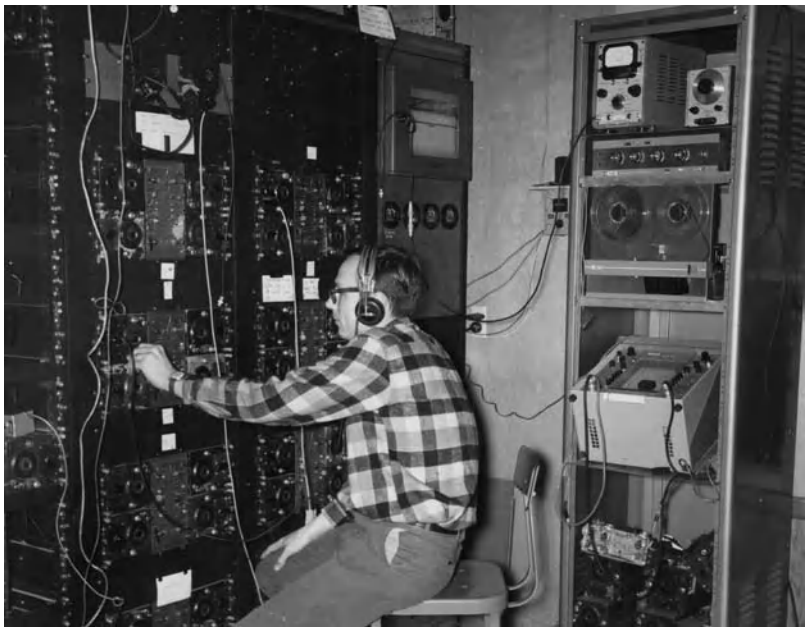

Fig. 2. Basic receiving system for biotelemetry signals uses a portable receiver and a hand-held directional antenna. (*From L. D. Mech, Handbook of Animal Radio-Tracking, University of Minnesota Press, 1983*)

Workplace. Biotelemetry has been used to study the performance, the comfort, and the health of workers. For example, the cardiovascular strain on glassblowers has been measured, as have the changes in body temperature and heart rate of factory employees as they move among different work stations. Telemetry has been used to monitor the effects of changing air pressure on workers who descend into subway tunnels. Video recordings of workers' behavior and activity have been combined with telemetry to identify those situations that cause stress to employees. For example, changes in heart rate or oxygen uptake, monitored by telemetry, and combined with a video of eye blinking, have been used to detect stress. Also, measurements of the environment, such as humidity, ambient temperature, and gas concentrations, can be monitored at many sites, sent by telemetry to a central receiver, and ultimately combined with biotelemetry to help evaluate working conditions.

Technical aspects. Biotelemetry is used to measure biological attributes over some distance without disrupting the subject's normal behavior. Consequently, there are usually two concerns associated with the use of biotelemetry: the distance over which the signal can be received, and the size of the transmitter package. Often, both of these concerns depend on the power source for the transmitter. Generally, the greater the radiated power from the transmitter antenna, the greater the distance of reception, when all other factors remain constant. But most transmitters are powered by primary batteries using mercuric oxide, silver oxide, or lithium chemical reactions. These batteries are available in a variety of sizes, with larger sizes having more energy capacity. Larger batteries can be used to make a transmitter last longer or to produce a more powerful signal. Integrated circuits and surface mount technology allow production of very small electronic circuitry in transmitters, making batteries the largest part of the transmitter package. However, the more powerful transmitters with their larger batteries are more difficult to place on or implant in a subject without affecting the subject's behavior or energetics. *See* BATTERY.

The mere presence of a transmitter can bother subjects, thus distracting them from routine behavior. Transmitters can interfere with normal movements. The additional mass of a transmitter package increases the energy required by the subject to maintain posture and to move about. On flying or swimming subjects, externally mounted transmitters can increase drag. The placement of the transmitter can affect a subject's balance and maneuverability. External transmitters, antenna, and attachment materials can become ensnared on objects in the subject's environment. The presence of the transmitter can affect the way that other organisms behave toward the subject, and on some animals a transmitter can increase the risk of the subject being attractive to a predator. Concern about these factors has led biologists to develop and test many attachment and implant methods for different sizes of transmitters. Nevertheless, these factors limit the power, and

thus the operational life and reception distance, of biotelemetry transmitters.

In addition, there are several factors related to the receiver and antennas that are important for functional aspects of biotelemetry. For instance, comparatively simple and insensitive radio receivers, like an AM/ FM transistor radio, can be used with powerful transmitters or at relatively short distances (for example, less than 30 ft or 10 m), but more sensitive receivers are capable of picking up less powerful transmitter signals or those from larger distances. Transmitting and receiving antennas also come in many designs, some more sensitive or more directional than others. As signals (radio, sound, or light) travel farther from the transmitter, they become weaker and can be detected only by the more sensitive receiving equipment.

There are many environmental factors that influence signals and consequently the reception range. Sources of signals at or on the operating frequency will cause interference that can reduce, or eliminate, reception range of target transmitters. Radio signals can be absorbed and scattered by the Earth's surface, vegetation, moisture in the air, and the subject's body. Therefore, many studies are designed to maximize the relationship between the transmitting and receiving antennas. For example, a single-element wire transmitting antenna is most efficient when positioned at an angle away from the subject's body. Also, whenever possible, the receiving antennas are raised above the ground, or vegetation, by placing them on a mast or the wing of an airplane or, for some applications, in an orbiting satellite.

Transmitter power, antenna (transmitting and receiving) design and position, receiver sensitivity, atmospheric conditions, and interference all interact to affect the quality of biotelemetry signals and the distance over which they can be detected. Radio transmitters that are small, a few millimeters in diameter (25 mm = 1 in.) and weighing less than a gram (28g = 1 oz), might transmit only 3 ft (1 m) from within a subject to a receiver in a laboratory, or less than 300 ft (100 m) from an animal's body surface to a receiver and antenna held by a wildlife biologist. Infrared light signals are usually limited to an area within a room. The radio signal from a 1-oz (30-g) transmitter around the neck of a mammal might be received over 12 mi (20 km), while the same package on the back of a flying bird could be received over 120 mi (200 km) away in an airplane. Signals from 0.4–11 lb (0.2–5 kg) transmitters are sent to satellites that are 520 mi (860 km) in space; some transmitters send sound waves more than 60 mi (100 km) through water. In some applications, signals are sent a relatively short distance to local receivers, from which they are amplified and sent much greater distances to another receiver.

Receiving systems (**Fig. 2**) can function simply to transform an incoming signal from one form (for example, radio) to another (for example, sound) or can be combined with a recording device (for example, magnetic tape recorder or paper chart recorder). Receiving systems can function to decode complex digital and modulated signals. Receivers can be combined with computing equipment to sample various frequencies at preprogrammed intervals, to sample incoming signals, to perform analyses, and to display and sort data derived from transmitted signals. *See* TELEMETERING.

Mark Roy Fuller

Bibliography. C. J. Amlaner (ed.), *Biotelemetry: Proceedings of the 10th International Symposium on Biotelemetry*, 1989 (and also earlier proceedings); C. J. Amlaner and D. W. McDonald (eds.), *A Handbook on Biotelemetry and Radio Tracking*, 1980; R. Kenward, *Wildlife Radio Tagging*, 2d ed., 2000.

Biotin

A vitamin, widespread in nature, whose deficiency causes dermatitis and other disturbances in animals and humans. It is only sparingly soluble in water (0.03 g/100 ml H_2O at 25°C or 77°F). It is stable in boiling water solutions, but can be destroyed by oxidizing agents, acids, and alkalies. Under some conditions, it can be destroyed by oxidation in the presence of rancid fats. Its structural formula is shown below.

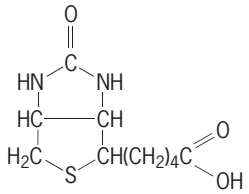

Occurrence. Biotin's occurrence in nature is so widespread that it is difficult to prepare a natural deficient diet. Studies of egg-white injury resulted in the discovery of biotin deficiencies in animals. Raw egg white contains a protein, avidin, which combines with biotin to form an indigestible complex. The feeding of large amounts of raw egg white has been used in producing experimental biotin deficiencies. Heating of egg white denatures avidin, destroying its biotin complexing ability. Although the avidin-biotin complex is stable to intestinal enzymes, it is hydrolyzed when administered parenterally.

Bioassay. Chemical and physical methods for biotin analysis are not available. Rat and chick assays are not very good and are seldom used. Microbiological assays using yeast or *Lactobacillus arabinosus* are commonly used.

Nutritional value. Biotin deficiency in animals is associated with dermatitis, loss of hair, muscle incoordination and paralysis, and reproductive disturbances. Biotin deficiency produced in human volunteers by feeding large amounts of egg white resulted in dermatitis, nausea, depression, muscle pains, anemia, and a large increase in serum cholesterol.

Biochemistry. Biotin occurs in nature mainly in bound forms. One of these isolated from yeast, biocytin, is ϵ-*N*-biotynyl-L-lysine. Biocytin can replace biotin as a growth factor for many organisms. Although the metabolic role of biocytin is unknown, its structure and concentration in some materials

suggest that it is more than an oddity. Biotin functions as a coenzyme in carbon dioxide fixation and decarboxylation enzymes, as illustrated by the reversible reaction: pyruvic acid + $CO_2 \rightleftharpoons$ oxaloacetic acid. The amino acid aspartic acid can replace biotin for some microorganisms because the bacteria require aspartic acid for growth, and the oxaloacetic acid produced by biotin-containing enzymes can be converted by transamination to aspartic acid. Biotin is also involved in deamination of some amino acids: aspartic acid $\rightleftharpoons NH_3$ + fumaric acid. Some microorganisms require oleic acid if biotin is not supplied. Biotin may be important for oleic acid synthesis, or vice versa. Under normal conditions, biotin is synthesized to such a degree by intestinal bacteria that a dietary source is probably not necessary. Most diets contain 25–50 micrograms per day, but 2–5 times as much is excreted in the urine and feces. *See* COENZYME; VITAMIN B_{12}. Stanley N. Gershoff

Industrial synthesis. The structure of biotin was established and the vitamin synthesized by V. du Vigneaud and coworkers in a masterpiece of technical management of a highly complicated synthesis. *meso*-Dibenzylaminosuccinic acid is reacted with phosgene to close the imidazolidone ring. The two carbocyclic groups enable closing the adjacent thiophanone ring by reduction and sulfonation. A three-carbon chain is then introduced by reaction with ethoxypropylmagnesium bromide. Upon closing a third (thiophanium) ring, resolution of the optical antipodes is possible. A malonic ester condensation providing the necessary side chain is followed by dibenzylation, saponification, and decarboxylation, giving the naturally occurring dextrorotatory biotin. *See* VITAMIN. Roger A. Mercanton

Bibliography. K. Dakshinamurti and H. N. Bhagavan, *Biotin*, 1985.

Biotite

An iron-magnesium-rich layer silicate; it is also known as black mica. Biotite is the most abundant of the mica group of minerals. The name is derived from that of the French chemist J. Biot. The formula for the ideal end member, phlogopite, is $KMg_3AlSi_3O_{10}(OH)_2$. The more general formula is $AX_3Y_4O_{12}(Z)_2$, where A (interlayer cation) = K, Na, Ca, Ba, or vacancies; X (octahedral cations) = Li, Mg, Fe^{2+}, Fe^{3+}, Al, Ti, or vacancies; and Y (tetrahedral cation) = Fe^{3+}, Al, Si; Z = (OH), F, Cl, O^{2-}. This formula is more indicative of the wide range of compositions known for this mineral. Biotite has no commercial value, but vermiculite, an alteration product of magnesium-rich biotite, is used as insulation, as packing material, and as an ingredient for potting soils. *See* VERMICULITE.

The crystal structure of biotite is based on a sheet of octahedrally coordinated iron or magnesium ions sandwiched between two sheets of tetrahedrally coordinated silicon and aluminum ions to form a layer. The layers are bonded by interlayer alkali or alkaline-earth ions. Over 20 different polytypes have been discovered, all based on multiple layers of the basic unit which has monoclinic symmetry. The observed polytypes belong to the triclinic, monoclinic, and trigonal crystal systems. *See* CRYSTAL STRUCTURE.

The hardness of biotite ranges between 2.5 and 3 on Mohs scale, and the specific gravity between 2.7 and 3.5. The color ranges from colorless (phlogopite) through shades of brown and green to black. Oxidized and Ti-rich biotites are usually reddish brown. Iron-rich varieties of biotite have been referred to as lepidomelane, and biotites rich in both iron and aluminum are called siderophyllites. When Fe^{3+} substitutes in the tetrahedral layers for Al and Si, the orientation of the optic plane changes, and such varieties of biotite are referred to as anomites.

Biotites are found commonly in igneous and metamorphic rocks. They are the common ferromagnesian phase in most granitic rocks, and are also found in some siliceous and intermediate volcanic rocks. In basaltic rocks biotite sometimes occurs in the crystalline groundmass, and is a common late interstitial phase in gabbroic rocks. It has been recognized in samples of the Earth's mantle found as inclusions in volcanic rocks. Biotite has not been observed in lunar rocks. *See* IGNEOUS ROCKS; METAMORPHIC ROCKS.

Biotites are not stable at the surface of the Earth, as they decompose by both hydrolysis and oxidation when exposed to the Earth's atmosphere. They alter to vermiculite, chlorite, and iron oxides, and thus are uncommon in sedimentary rocks. Biotites in erupted lavas will undergo dehydrogenation, a process in which hydrogen leaves the crystal structure, and the charge balance is maintained by the conversion of Fe^{2+} to Fe^{3+} and (OH) to O^{2-}. Such crystals are known as oxybiotites.

Biotites are important constituents of metamorphic rocks such as schist and gneiss, and the first appearance of biotite is an important marker in metamorphism. Biotite persists to very high grades of metamorphism, where it reacts with quartz to form granulites made up of potassium feldspar and orthopyroxene, garnet, or cordierite, in addition to quartz and plagioclase. Under conditions of ultrametamorphism, biotite reacts with quartz, plagioclase, and alkali feldspar to form siliceous melts. Biotite is also a common gangue mineral in ore deposits. The mineral has been used as an indicator of H_2O, HF, O_2, and S_2 activities in both rock- and ore-forming processes. Its absence in lunar rocks can be attributed to the extremely low contents of H_2O and F in lunar materials. *See* METAMORPHISM; MICA; SILICATE MINERALS. David R. Wones

Birch

A deciduous tree of the genus *Betula* which is distributed over much of North America, in Asia south to the Himalaya, and in Europe. About 40 species are known. The birches comprise the family Betulaceae in the order Fagales. The sweet birch, *B. lenta* (**illus.** *a*), the yellow birch, *B. alleghaniensis* (illus. *b*), and the paper birch, *B. papyrifera*, are

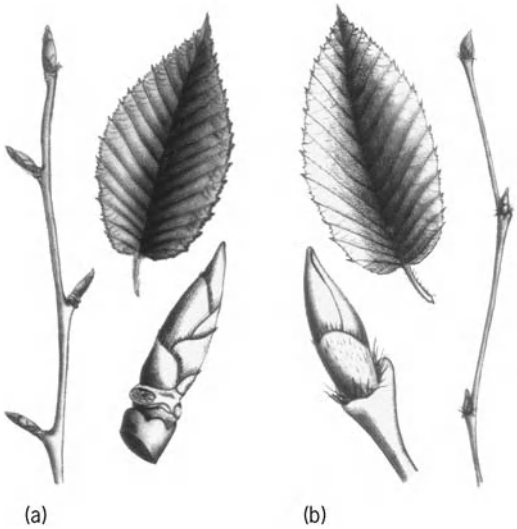

Birch twigs, leaves, and buds: (*a*) sweet birch (*Betula lenta*); (*b*) yellow birch (*B. alleghaniensis*).

all important timber trees of the eastern United States. The yellow and the paper species extend into Canada. The gray birch, *B. populifolia*, is a smaller tree of the extreme northeastern United States and adjacent Canada. Both sweet (black) and yellow birches can be recognized by the wintergreen taste of the bark of the young twigs. A flavoring similar to oil of wintergreen is obtained from the bark of the sweet birch. The paper and gray birches can be easily identified by their white bark. The bark of the paper birch peels off in thin papery sheets, a characteristic not true of the gray birch. The river birch, *B. nigra*, is a less common tree of wet soils and banks of streams and is important as an ornamental and for erosion control. The hard, strong wood of the yellow and the sweet birches is used for furniture, boxes, baskets, crates, and woodenware. The European birches, *B. pubescens* and *B. pendula*, are the counterparts of the paper and gray birches in the United States. European birches are also cultivated in America. *See* FAGALES; FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Birefringence

The splitting which a wavefront experiences when a wave disturbance is propagated in an anisotropic material; also called double refraction. In anisotropic substances the velocity of a wave is a function of a displacement direction. Although the term birefringence could apply to tranverse elastic waves, it is usually applied only to electromagnetic waves.

In birefringent materials either the separation between neighboring atomic structural units is different in different directions, or the bonds tying such units together have different characteristics in different directions. Many crystalline materials, such as calcite, quartz, and topaz, are birefringent. Diamonds on the other hand, are isotropic and have no special effect on polarized light of different orientations. Plastics composed of long-chain molecules become anisotropic when stretched or compressed. Solutions of long-chain molecules become birefringent when they flow. This first phenomenon is called photoelasticity; the second, streaming birefringence. *See* PHOTOELASTICITY.

For each propagation direction with linearly polarized electromagnetic waves, there are two principal displacement directions for which the velocity is different. These polarization directions are at right angles. The difference between the two indices of refraction is also called the birefringence. When the plane of polarization of a light beam does not coincide with one of the two principal displacement directions, the light vector will be split into components that are parallel to each direction. At the surface of such materials the angle of refraction is different for light that is polarized parallel to the two principal directions.

For additional information on befringence and birefringent materials *see* CRYSTAL OPTICS; POLARIZED LIGHT; REFRACTION OF WAVES. Bruce H. Billings

Birth control

Methods of fertility control, including contraception, that are intended to prevent pregnancy, and means of interrupting early pregnancy. The efficacy of the various methods and consistency of use vary widely. Factors associated with degree of effectiveness include user age, income, marital status, and intention (that is, whether contraception is used to delay or to prevent pregnancy). The available methods consist of hormonal methods (including oral contraceptives, subdermal implants, and injectable formulations), sterilization, intrauterine devices, barrier and chemical methods, and fertility awareness methods.

Hormonal contraceptives. Hormonal contraception relies upon the use of various synthetic hormones that are either identical or similar in action to those normally produced within a woman's body. However, by varying the dosages and combinations of these hormones, ovulation is generally suppressed, producing temporary infertility. These contain one or both of two compounds (estrogen and progestin) similar to the hormones that regulate the menstrual cycle. Each monthly series of pills either suppresses ovulation or alters the uterine lining and the cervical mucus, or both. Oral contraceptives can help protect users against anemia, ectopic pregnancy, pelvic inflammatory disease, benign breast disease, ovarian cysts, and carcinoma of the ovary and of the endometrium. Complications include blood clots, which can cause stroke or heart attack; their occurrence, although rare, is far more likely for women over age 35 who smoke heavily or have certain preexisting health problems, for example, hypertension, than for other users. Cardiovascular risk may be lower with use of low-dose oral contraceptives. Those containing the progestins norgestimate, desogestral, or gestodene may even offer some

protective effects. *See* ENDOCRINE MECHANISMS; ESTROGEN; MENSTRUATION; PROGESTERONE.

Postcoital contraception is another hormonal method. In emergency situations (for example, rape) high dosages of oral contraception can be used. One dose is given within 72 h after the episode of unprotected intercourse, and an additional dose is given 12 h later.

Subdermal implants. The subdermal implant consists of small hollow rods that are placed under the skin of a woman's upper arm and release a low, continuous dose of a progestin. It is more effective than the oral contraceptives and, because it lacks estrogens, does not pose a risk of cardiovascular complications. It is reversible, lasts for 5 years, is nearly as reliable as sterilization, and is less expensive than birth control pills. Women using this method experience more irregular bleeding than with other hormonal methods. No serious complications have been documented; however, it is essential that insertion and removal of subdermal implants be performed by a physician fully trained in this procedure.

Injectables. An injection of progestin suppresses ovulation and can be given every 3 months. It can be used by women who should not take estrogens. Women experience irregular bleeding in the first 6 months of use, often followed by cessation of menses with continuing use. It has been shown to be as safe and reliable as sterilization, yet is readily reversible.

Intrauterine devices. The main mode of action for the intrauterine device (IUD) is considered to be prevention of fertilization. Of the two commercially available IUDs in the United States, the one containing copper is designed to remain in place for 10 years; for users over the age of 25, the pregnancy rate is less than 1%. The other one releases a daily dosage of the natural hormone progesterone to suppress the uterine lining and requires annual replacement.

Use of an IUD can seriously increase the risk of pelvic infection and subsequent infertility if a user's relationship with her partner is not monogamous; thus, the IUD is not suited for women who have more than one sexual partner, or whose partner does. Other problems associated with IUD use can include cramps, increased menstrual bleeding, expulsion of the device, and, more rarely, perforation of the uterus or tubal (ectopic) pregnancy. However, since the risk of pregnancy is low among IUD users, the absolute risk of ectopic pregnancy is significantly less for users of the copper IUD than among women not using contraception.

Barrier methods. Such methods include the male condom, female intravaginal pouch, diaphragm, cervical cap, vaginal contraceptive sponge, and various chemical preparations. The condom is a sheath of thin latex (sometimes coated with spermicide) or animal tissue that covers the penis. The intravaginal pouch, also known as a female condom, is a loose-fitting vaginal liner with an external rim designed to hold it in place. The diaphragm is a shallow rubber cup with a ring rim that fits securely in the vagina to cover the cervix. The cervical cap is a smaller, thimble-shaped latex device that fits over the cervix. The diaphragm and cervical cap are used with spermicides. The vaginal contraceptive sponge is a soft, synthetic, disposable sponge that fits over the cervix and, when moistened, continuously releases spermicide. Among the chemical preparations available as vaginal contraceptives are foams, creams, jellies, and suppositories.

Fertility awareness methods. Fertility awareness methods enable a woman to estimate when she is fertile so that she can practice abstinence or use a barrier method during those times. Techniques used to determine fertility include cervical mucus observation, and body signs with temperature tracking. Such methods are often less effective for contraception.

Sterilization. Sterilization is the most commonly used method of birth control for women and men both in the United States and worldwide. It is preferred by women over the age of 30 who have had the desired number of children. Less than 1% of males who have a vasectomy or females who undergo tubal sterilization are thereafter involved in pregnancy, and so the individuals no longer need to practice a method of birth control. The procedures do not adversely affect the production of male or female hormones, so that individual sexual characteristics such as sex drive and menses usually remain unchanged.

Vasectomy is a minor male surgical procedure that occludes the vas deferens by various means (such as by cautery or suture).

In the United States, tubal sterilization in the female is an operation commonly performed through the laparoscope, a long, thin, tubular instrument with magnifying mirrors that is inserted through the navel for visualization. The instrument that performs the tubal occlusion may be either attached to the laparoscope or inserted through the lower abdomen. Clips, bands, or cautery are used to close off the fallopian tubes. Minilap, an operation to tie the tubes, may also be performed through a small incision in the lower abdomen. Serious complications such as hemorrhage or infection are very rare.

The individual needs to be sure that children are no longer wanted before opting for sterilization, because the operation should be considered as permanent. Only in a small number of instances has it been successfully reversed.

Future methods. New systems to deliver long-acting steroidal contraceptives are likely to become available. Examples of systems that are in the research stage are injectable microcapsules, vaginal rings, and transdermal patches. Vaccines that would be injected annually are also being tested. Antiprogesterone compounds, which were developed in Europe, block receptors for progesterone, a key hormone in establishing and maintaining pregnancy. *See* FERTILIZATION (ANIMAL); PREGNANCY.

Louise B. Tyrer

Bibliography. L. Corson, R. J. Derman, and L. B. Tyrer (eds.), *Fertility Control,* 2d ed., 1995.

Bismuth

The metallic element, Bi, of atomic number 83 and atomic weight 208.980 belonging in the periodic table to group 15. Bismuth is the most metallic element in this group in both physical and chemical properties. The only stable isotope is that of mass 209. It is estimated that the Earth's crust contains about 0.00002% bismuth. It occurs in nature as the free metal and in ores. The principal ore deposits are in South America. However, the primary source of bismuth in the United States is as a by-product in refining of copper and lead ores. *See* PERIODIC TABLE.

1												13	14	15	16	17	18
1 **H**	2																2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

The main use of bismuth is in the manufacture of low-melting alloys which are used in fusible elements in automatic sprinklers, special solders, safety plugs in compressed gas cylinders, and automatic shutoffs for gas and electric water-heating systems. Some bismuth alloys, which expand on freezing, are used in castings and in type metal. Another important use of bismuth is in the manufacture of pharmaceutical compounds.

Bismuth is a gray-white, lustrous, hard, brittle, coarsely crystalline metal. It is one of the few metals which expand on solidification. The thermal conductivity of bismuth is lower than that of any metal, with the exception of mercury. The **table** cites the chief physical and mechanical properties of bismuth. Bismuth is inert in dry air at room temperature, although it oxidizes slightly in moist air. It rapidly forms an oxide film at temperatures above its melting point, and it burns at red heat, forming the yellow oxide, Bi_2O_3. The metal combines directly with halogens and with sulfur, selenium, and tellurium; however, it does not combine directly with nitrogen or phosphorus. Bismuth is not attacked at ordinary temperatures by air-free water, but it is slowly oxidized at red heat by water vapor.

Almost all compounds of bismuth contain trivalent bismuth. However, bismuth can occasionally be pentavalent or monovalent. Sodium bismuthate and bismuth pentafluoride are perhaps the most important compounds of Bi(V). The former is a powerful oxidizing agent, and the latter a useful fluorinating agent for organic compounds. Samuel J. Yosin

Physical and mechanical properties of bismuth

Property	Value	Temperature
Melting point, °C	271.4	
Boiling point, °C	1559	
Heat of fusion, kcal/mole	2.60	
Heat of vaporization, kcal/mole	36.2	
Vapor pressure, mm Hg	1	917°C
	10	1067°C
	100	1257°C
Density, g/cm^3	9.80	20° (solid)
	10.03	300° (liquid)
	9.91	400° (liquid)
	9.66	600° (liquid)
Mean specific heat, cal/g	0.0294	0–270°C
	0.0373	300–1000°C
Coefficient of linear expansion	13.45×10^{-6}/°C	
Thermal conductivity, cal/(s)(cm^2)(°C)	0.018	100° (solid)
	0.041	300° (liquid)
	0.037	400° (liquid)
Electrical resistivity, μohm-cm	106.5	0° (solid)
	160.2	100° (solid)
	267.0	269° (solid)
	128.9	300° (liquid)
	134.2	400° (liquid)
	145.3	600° (liquid)
Surface tension, dynes/cm	376	300°C
	370	400°C
	363	500°C
Viscosity, centipoise	1.662	300°C
	1.280	450°C
	0.996	600°C
Magnetic susceptibility, cgs units	-1.35×10^{-6}	
Crystallography	Rhombohedral, $a_0 = 0.47457$ nm	
Thermal-neutron absorption cross section, barns	0.032 ± 0.003	
Modulus of elasticity, lb/cm^2	4.6×10^6	
Shear modulus, lb/cm^2	1.8×10^6	
Poisson's ratio	0.33	
Hardness, Brinell	4–8	

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; D. F. Shriver and P. W. Atkins, *Inorganic Chemistry*, 3d ed., 1999.

Bison

The name for two species of the Bovidae, the cattle family, in the mammalian order Artiodactyla, found in North America and Europe. The European bison (*Bison bonasus*) is commonly known as the wisent. The American species (*B. bison*; see **illus.**), is often called buffalo but should not be confused with the African buffalo.

The American Bison Society was formed in 1905 to preserve this species. A small herd was established

North American bison (*Bison bison*).

in the Wichita Forest Reserve in Oklahoma. This herd prospered, and many specimens of this stock were sent to zoos, where they adjusted and reproduced. Another herd was established in South Dakota; it is culled annually, and the excess animals are sold at auction for food. Experimental crosses have been made with range cattle. The resulting offspring forage and survive adverse environmental conditions much better than cattle do.

Wisent. The European bison was originally abundant in the forested areas of western Europe during the late Cenozoic Era. The wisent is a browsing, woodland animal which congregates in relatively small herds. It is almost extinct in its natural range and, although a few herds exist, it is known mainly from a few hundred specimens preserved in zoological gardens and zoos. The few protected wild herds, which were maintained on the large landed estates of Poland and Lithuania prior to World War I, were killed off for food, thus reducing the numbers drastically. An intensive effort has been made to preserve the species by breeding and maintaining this animal in captivity.

After World War II a concerted effort was made to locate these scattered animals and determine their pedigree. At the present time, it is estimated that more than 400 pureblood wisents have been recorded. In an effort to maintain the animals in Europe, they had frequently been crossed with the American bison. The hybrids have been excluded from the pedigree stock lists.

Although there are differences, the wisent is closely allied to the American species. The wisent has a small head carried in a high position and a short mane, and is more graceful and less massive than the North American species. The hump is less noticeable, the legs are longer, the horns are more slender, and the body is not so shaggy. A wary animal, the wisent has poor vision and a keen sense of smell. A single calf is born after a gestation period of 39–40 weeks. Cows mature at about 4 years and bulls at about 6–8 years.

American bison. Enormous herds of the American bison existed on the Plains area of North America in western Canada and the western United States during the nineteenth century. It is estimated that there are still about 20,000 bison in Canada and the United States on preserves and national parks. These bovines are massive, with the males attaining a length of 9 ft (3 m), a height of 6 ft (2 m) at the shoulder, and a weight of up to 3000 lb (1350 kg).

These animals are herbivorous, and migrated originally when the grass or forage became scarce. The bison chews a cud, and the teeth are typically those of grazing animals with a dental formula of I 0/3 C 0/1 Pm 3/3 M 3/3 for a total of 32 teeth. The large head is held in a low position, and behind the neck region is a characteristic hump. The forequarters are shaggy with the hair more pronounced in the male than the female. The senses of smell and hearing are well developed, while vision is poor. Both the male and female have horns which are small and set far apart. The bison is unusual in having 14 pairs of ribs in contrast to the usual 13 pairs. The female comes into heat in late summer, and after a gestation period of over 9 months, a single calf is born in the spring. The maximum lifespan of the bison is about 30 years. *See* ARTIODACTYLA; BUFFALO; MAMMALIA.

Charles B. Curtin

Bit

A binary digit. In the computer, electronics, and communications fields, "bit" is generally understood as a shortened form of "binary digit." In a numerical binary system, a bit is either a 0 or 1. Bits are generally used to indicate situations that can take one of two values or one of two states, for example, on and off, true or false, or yes or no. If, by convention, 1 represents a particular state, then 0 represents the other state. For example, if 1 stands for "yes," then 0 stands for "no." *See* BOOLEAN ALGEBRA; NUMBERING SYSTEMS.

In a computer system a bit is thought of as the basic unit of memory where, by convention, only either a 0 or 1 can be stored. In a computer memory, consecutive bits are grouped to form smaller or larger "units" of memory (see **table**). Depending upon the design of the computer, units up to 64 bits long have been considered. Although there is common agreement as to the number of bits that make up a byte, for larger memory units the terminology depends entirely on the convention used by the manufacturer. In all of these units the leftmost bit is generally called

Units of computer memory

Unit name	Number of consecutive bits
Nibble	4
Byte	8
Word or half-word	16
Double word or full-word	32

the most significant bit (msb) and the rightmost the least significant bit (lsb).

Alphanumeric representation. Bytes and larger units can be used to represent numerical quantities. In these cases the most significant bit is used to indicate the "sign" of the value being represented. By convention a 0 in the msb represents a positive quantity; a 1 represents a negative quantity. Depending on the convention used to represent these numbers, the remaining bits may then be used to represent the numerical value. In addition to numerical quantities, bytes are used to represent characters inside a computer. These characters include all letters of the English alphabet, the digits 0 through 9, and symbols such as comma, period, right and left parentheses, spaces, and tabs. Characters can be represented using the ASCII code (American Standard Code for Information Interchange) or the EBCDIC code (Extended Binary Coded Decimal Interchange Code). The latter is used by some mainframe computers. Computers are set up to handle only one of these two character codes. Generally, the internal representation of a character is different in the two codes. For instance, in ASCII the plus sign is represented by the numerical sequence 00101011, and in EBCDIC, by 01001110. *See* NUMERICAL REPRESENTATION (COMPUTERS).

Terminology. The word "bit" seems to have been coined during a meeting convened by John Von Neumann and Norbert Wiener during the winter of 1943–1944 in Princeton, New Jersey. The participants found it convenient to measure information in terms of numbers of "yes" and "no" and to call this unit of information a bit. In 1975 the International Organization for Standardization (OSI) defined a new unit, the shannon (honoring C. E. Shannon). A shannon is a unit of measure equal to the decision content of a set of two mutually exclusive events. In information theory, the word "bit" is sometimes used as a synonym for this preferred term. *See* INFORMATION THEORY.

Ramon A. Mata-Toledo

Bibliography. W. Stallings, *Computer Organization and Architecture*, 3d ed., 1993; R. White, *How Computers Work*, Ziff-Davis, 1993.

Bitumen

A term used to designate naturally occurring or pyrolytically obtained substances of dark to black color consisting almost entirely of carbon and hydrogen with very little oxygen, nitrogen, and sulfur. Bitumen may be of variable hardness and volatility, ranging from crude oil to asphaltites, and is largely soluble in carbon disulfide. *See* ASPHALT AND ASPHALTITE.

Irving A. Breger

Bivalve, rudist

An extinct group of epifaunal bivalves that constructed reefs during the late Mesozoic. The valves of rudists are coiled to cylindrical in external form, and the shapes and sizes of rudist shells are extreme relative to other members of the class Bivalvia. Rudists lived as solitary individuals or in large clusters on predominantly carbonate substrates in low to mid-paleolatitudes until their demise near or at the end of the Cretaceous. *See* BIVALVIA; MESOZOIC.

Evolution. The earliest rudists are found in rocks more than 150 million years old. Rudist ancestry can be traced to thick-shelled megalodontid bivalves through a series of comparative internal features such as ligament, dentition, and muscle supports, and through external shape of the asymmetrical valves. Rudists are animals classified in the paleontologic literature as mollusks under the superfamily Hippuritoidea. There are more than 200 genera and more than 1500 named species of rudists. *See* MOLLUSCA.

Morphology. Rudists are similar to bivalves with which they are classified, yet they evolved unique features in their epifaunal shells. Rudists and all bivalves have two calcium carbonate valves hinged together. The valves protect the heart, kidneys, and other organs vital to the biological processes of the animal. The calcium carbonate valves of rudists have distinct outer calcite and inner aragonite layers, and the animal lived inside the shell, protected by these hard, mineralized deposits. The internal part of the rudist shell is partitioned to accommodate growth, and shell layers were accreted as the animal grew. Unlike most members of the Bivalvia, size variation among rudist individuals is extreme and ranges from a few centimeters to more than 1 m (3.3 ft). The rudist animals with shells of doubly coiled morphologies tended to dwell as individuals or in small clusters, whereas the cylindrically shaped valves allowed for closer packing of individuals and construction of reefs.

Paleoecology. The substrates on which rudists dwelled consisted of silt-sized carbonate particles that shifted with bottom currents, and carbonate mud that was relatively nonmobile. Rudist larvae attached themselves to mobile grains, and the rudist animal secreted cements that adhered the shell to the substrate early in life. Immobilization of the grains was likely aided by microbial biomineralization. In all substrates, rudists grew their shells at a rate faster than the rate of sediment accumulation. In this manner, they elevated the commissure—the opening between the valves—above the substrate and kept it free from sediments entering the valves and impairing vital life functions. Although rudists lived primarily on carbonate sediments, they adapted through time to substrates that contained rock particles derived from land and from reworked submarine volcanic rocks. Within these carbonate to mixed carbonate-siliciclastic and volcaniclastic substrates, rudists lived on all parts of the shallow-water marine platforms. Rudists are known from isolated lagoons with relatively poor water circulation through reef crests and forereefs where they constructed the largest buildups in relatively higher energy surf conditions. The deepest water in which rudists lived is not known but it is shelfal, and

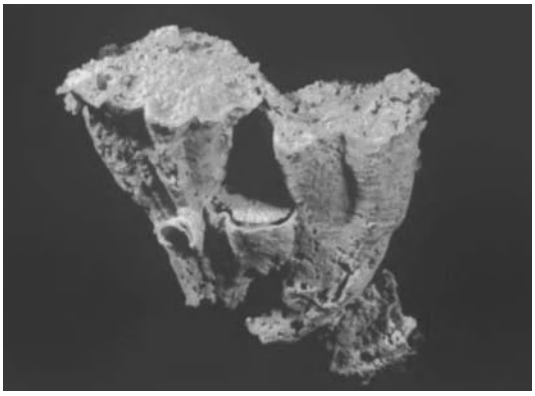

Fig. 1. Rudist bivalve morphology. Cluster of individual rudists showing the conical shape of the erect-growing lower valves and the operculiform shape of the upper valves. Two generations of rudists preserved as silicified specimens illustrate the bivalved nature of this group that allows for their classification within the class Bivalvia. These hippuritid rudists are from Maastrichtian strata of Puerto Rico.

shelfal in present-day measurements indicates less than 200 m (660 ft). On the platforms and shallower parts of the shelves, rudists were subjected to both calm water and hurricane-force energy conditions, as are the reefs of today. Rudists reefs were thus destroyed frequently by high-energy waves and were rebuilt during calmer conditions, resulting in thick accumulations of numerous generations of buildups. *See* PALEOECOLOGY; REEF.

Global distribution in tropical seas. Although rudists dominated the reefs through most of the Cretaceous, their numerical and ecological development can be traced through time. During the earliest Cretaceous, reefs and biotic accumulations of low relief were constructed by rudist bivalves and corals in the shallower marine waters, whereas corals, silica-secreting sponges, and microbes built biotic structures in deeper waters. By the Middle Cretaceous, about 100 million years ago, rudist bivalves diversified and dominated the shallow-water marine ecosystems both numerically and ecologically across low paleolatitudes in tropical seas. In later Cretaceous times, reefs composed of rudist bivalves were distributed east to west between the Arabian Peninsula and the Americas, and north to south from the low to mid-paleolatitudes. Westward-flowing surface currents accounted for the low- to mid-latitude distribution pattern of reefs, whereas northward surface currents accounted for northern occurrences in the European and North American regions, especially during sea-level highstands when shelfal areas were flooded. A global correspondence exists between the development of Upper Cretaceous reefs and highstands of the seas, but there is only a regional, not global, correlation between reefs and sea-level fluctuations of lesser magnitudes. Fluctuations in the levels of the seas are responsible for both fluxes in the amount of sediment spread across carbonate platforms and for the pathways of surface currents that bring larvae for settlement and subsequent colonization by reefal biota. *See* SEDIMENTATION.

Fig. 2. Rudist reef. Ecological association of rudists in proximity. Variation in valve shapes and sizes is illustrated in this photo from the latest Cretaceous Maastrichtian stage of Jamaica.

Large Upper Cretaceous reefs containing rudist bivalves as the dominant constructor also contained corals, sponges, stromatoporoids, and algae in lesser numbers and as secondary ecological constituents. North of the rudist reefs in the Northern Hemisphere, corals dominated the shallow-water ecosystems, and north of the corals were bioconstructions of silicious sponges or bryozoans. In the Southern Hemisphere, corals dominated the reefs in higher paleolatitudes. A zonation in reefs by a variety of dominant organisms thus occurred along paleolatitudinal bands, suggesting a thermal and/or chemical signature to the control of the reef-building biota.

Thermal conditions of the Middle and Late Cretaceous oceans were warmer than those of today in both the high and low paleolatitudes, with warmth extending to the polar regions as evidenced by freshwater lizards in the high northern paleolatitudes. The expanded tropics, the realm of the rudist bivalves, were hot, with ocean temperatures exceeding those of today by several degrees. Thermal and salinity conditions were probably higher in the shallowest waters in comparison to the open seas. Rudists were adapted to a wide range of thermal and saline conditions, even to temperatures and salinities that were not tolerated by corals, algae, and other biota. *See* CRETACEOUS.

Extinction. From their origins in the Late Jurassic through their terminal extinction in the latest Cretaceous, rudists as a group persisted through several regional and global extinction events. The exact timing of their terminal extinction is debatable. The extinction occurred either prior to the Cretaceous-Paleogene boundary or right at the boundary. The causes of the extinction remain unclear but are attributed to the thermal and chemical conditions of the latest Cretaceous oceans, perhaps in combination with effects of the extraterrestrial impact that formed the Chicxulub crater in the Yucatan Peninsula in the Caribbean ocean. *See* EXTINCTION (BIOLOGY).

Claudia C. Johnson; Erle G. Kauffman

Bibliography. R. Hofling and R. W. Scott, Early and mid-Cretaceous buildups, in W. Kiessling, E. Flugel,

and J. Golonka (eds.), *Phanerozoic Reef Patterns*, pp. 521–548, Soc. Sed. Geol. Spec. Publ. 72, Tulsa, 2002; C. C. Johnson, The rise and fall of rudist reefs, *Amer. Scientist*, 90:148–153, 2002; C. C. Johnson et al., Patterns and processes influencing Upper Cretaceous reefs, in W. Kiessling, E. Flugel, and J. Golonka (eds.), *Phanerozoic Reef Patterns*, pp. 549–585, Soc. Sed. Geol. Spec. Publ. 72, Tulsa, 2002; C. C. Johnson and E. G. Kauffman, Maastrichtian extinction patterns of Caribbean Province rudists, in N. MacLeod and G. Keller (eds.), *Cretaceous-Tertiary Mass Extinctions: Biotic and Environmental Changes*, W. W. Norton, New York, 1996; T. Steuber et al., Catastrophic extinction of Caribbean rudist bivalves at the Cretaceous-Tertiary boundary, *Geology*, 30:999–1002, 2002.

Bivalvia

One of the five classes in the phylum Mollusca; sometimes known as Pelecypoda. All bivalves are aquatic, living at all depths of the sea and in brackish and fresh waters. With about 25,000 living species (**Fig. 1**), Bivalvia is second to class Gastropoda (over 74,000) in molluscan species diversity. However, the total biomass of bivalves is much greater, and certain bivalve species are numerically dominant in many benthic ecosystems. The most primitive bivalves are infaunal, burrowing into soft sediments, but many families are epifaunal, attached to rocks or shells or residing on the sediment surface. Bivalves are well represented in the fossil record from the Early Paleozoic because of their calcareous shells. *See* MOLLUSCA.

Anatomy and physiology. In general, bivalves are bilaterally symmetrical and laterally compressed. They have a fleshy mantle that secretes the shell enclosing the body (**Fig. 2**). The mouth is located anteriorly in bivalves; and in the Lamellibranchiata, the largest subclass, the mouth is flanked by paired labial palps that act to sort food prior to ingestion. Sensory organs are located on the outer mantle margin that has the closest contact with the environment. Frequently these sensory organs are borne on tentacles, and they are sensitive to tactile and chemical stimuli. Certain species of scallops have highly developed light-sensing organs or "eyes" on their mantle tentacles. *See* LAMELLIBRANCHIA.

The shell consists of two valves with a noncalcified connecting ligament holding them together at a hinge plate. The shell layers consist of an outer horny periostracum (protective layer) that can be either absent or eroded in some species, a middle prismatic layer consisting of crystalline calcium carbonate, and an inner lamellar or nacreous layer. In some families such as the Mytilidae (mussels) or the Pteriidae (winged or pearl oysters), the nacreous layer can exhibit a beautiful iridescent sheen, whereas in most bivalves the inner layer is smooth but with a chalky appearance. Hinge ligament tension holds the valves in a gaping position, with valve closure effected by adductor muscles. Paired adductor muscles can be of equal size (isomyarian) such as those found in clams of the family Mactridae or Veneridae, or of unequal size as in the Mytilidae (heteromyarian); the adductor muscle also may be single as in oysters of the family Ostreidae (monomyarian). There are two muscle fiber types. Fast muscle fibers cause rapid closure of the valves, and catch fibers maintain closure for long periods of time without a great expenditure of metabolic energy. The ciliated molluscan gills, properly called ctenidia, are enlarged in the subclass Lamellibranchiata and occupy a substantial portion of the mantle cavity. The ctenidia consist of layered filaments which function primarily to pump water into the mantle cavity and to filter particulate food from the incurrent water stream. Observations using fiber-optic endoscopy in actively feeding bivalves have shown that both food and nonfood (silt) particles are trapped on mucous sheets on the ctenidia during the filter-feeding process. From the ctenidial face, trapped particles are passed ventrally to a ctenidial food groove that acts as a conveyor belt transferring the particles anteriorly to the labial palps, where they are sorted prior to ingestion. Rejected particles encased in mucus are dropped onto the mantle and eventually released as pseudofeces. There is evidence that large populations of filter-feeding bivalves can increase the clarity of natural waters by actively removing phytoplankton and other particles.

The ctenidia of bivalves of the subclass Protobranchia also serve to pump water, but they are smaller and less developed than in the lamellibranchs and do not serve to filter food particles. Protobranch bivalves are deposit feeders that gather food by extending thin muscular palp proboscides to probe soft sediments and entrap organic detrital particles. Cilia on the proboscides transfer particles to labial palps for sorting prior to ingestion. Bivalves of the subclass Septibranchia (sometimes called Anomalodesmata) have highly modified ctenidia that lack filaments. A septum divides the mantle cavity into dorsal and ventral chambers, and water is pumped by muscular contraction of the septum wall. Septibranch bivalves often feed on zooplankton or water-borne detritus particles. Their labial palps are reduced in size and do not have a sorting function. *See* PROTOBRANCHIA; SEPTIBRANCHIA.

In the protobranchs and lamellibranchs, ingested food particles pass through an esophagus to the gut, which in some species contains the crystalline style, a clear worm-shaped gelatinous worm-shaped organ that is continually rotated by cilia in a stomach pouch (style sac). The rotation of the style acts to release digestive enzymes and mix the stomach contents. Septibranch stomachs lack a crystalline style but are often stiff and chitinous to grind food particles. The bivalve stomach also contains a number of tiny sacs (digestive diverticula) that serve as sites for nutrient absorption. From the stomach, the digestive tract consists of an intestine that frequently passes through the heart of many bivalves and leads to the anus, which releases feces into the excurrent water stream. In the cases in which the bivalve intestine passes through the heart, the heart

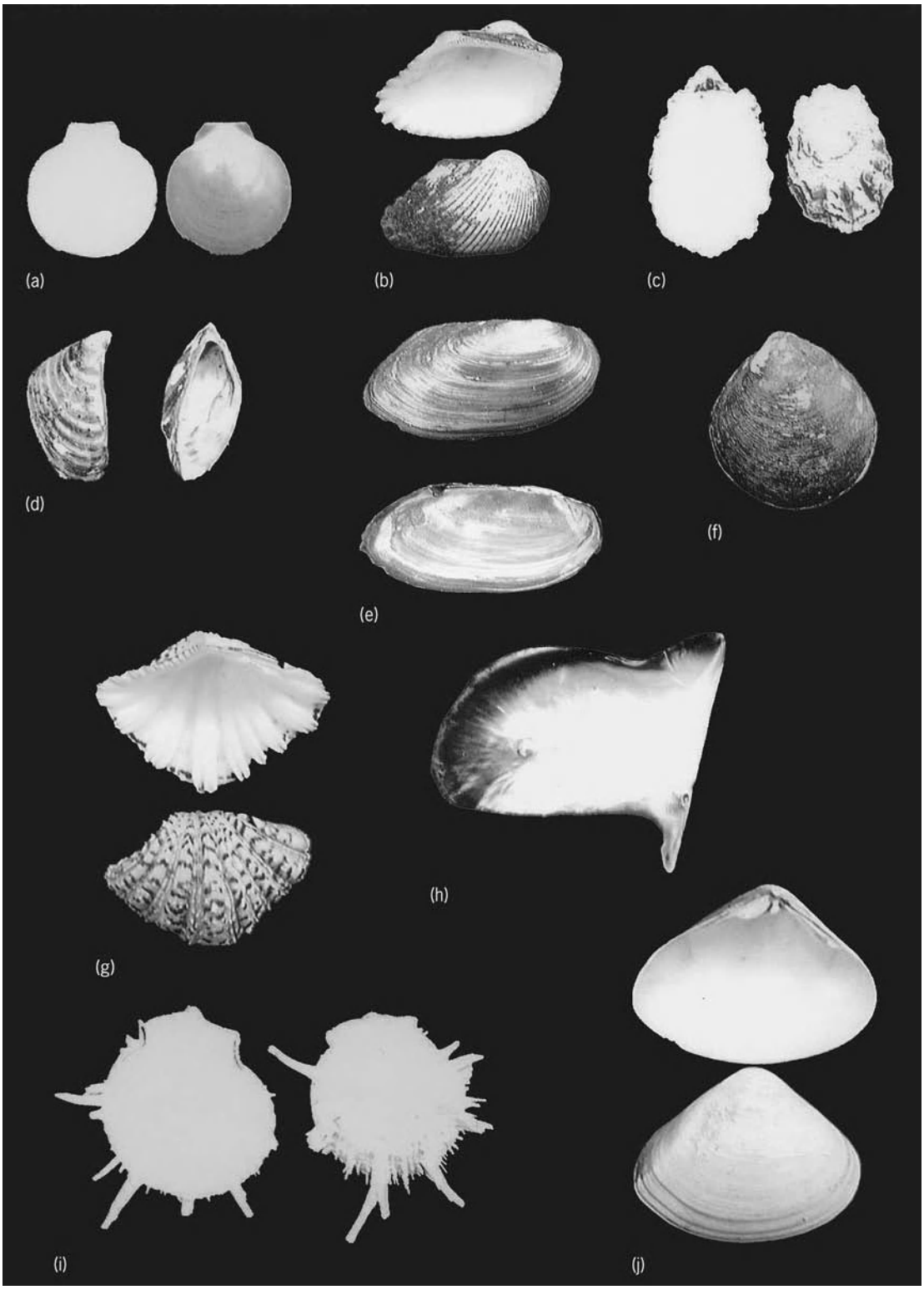

Fig. 1. Some bivalve species. (*a*) Scallop, *Amusium pleuronectes*, Pectinidae, Philippines. (*b*) Arkshell, *Barbatia reevana*, Arcidae, Ecuador. (*c*) Mangrove oyster, *Crassostrea rhizophorae*, Ostreidae, West Indies. (*d*) Zebra mussel, *Dreissena polymorpha*, Dreissenidae, Michigan. (*e*) Freshwater clam, *Eliptio complanata*, Unionidae, Rhode Island. (*f*) Estuary clam, *Geloina coaxans*, Corbiculidae, Philippines. (*g*) Giant clam, *Hippopus hippopus*, Tridacnidae, Samoa. (*h*) Pearl oyster, *Pinctada margaritifera*, Pteriidae, Indonesia. (*i*) Spiny oyster, *Spondylus princeps*, Spondylidae, Baja California, Mexico. (*j*) Pismo clam, *Tivela stultorum*, Veneridae, California.

performs the dual function of pumping blood and passing food particles down the digestive tract.

Alternative feeding mechanisms include the direct uptake of dissolved organic molecules from seawater across the ctenidial and mantle surfaces of actively pumping animals; and in some species, symbiotic organisms in the mantle or ctenidia serve to supply nutrients. For example, in giant clams of the family

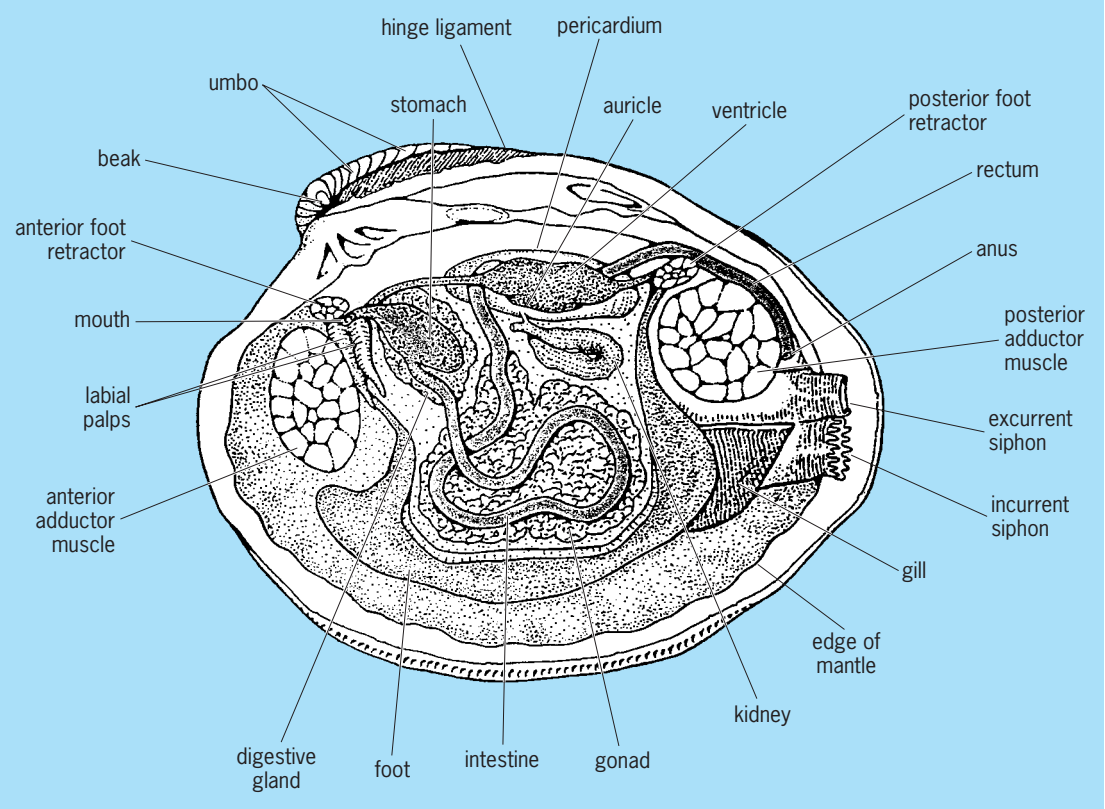

Fig. 2. Bivalve anatomy as represented by the quahog clam, *Mercenaria mercenaria*.

Tridacnidae, photosynthetically active symbiotic phytoplankton called zooxanthellae reside in the mantle tissue, allowing the clams to thrive in clear tropical waters that have little particulate food. Similarly, other bivalves in harsh sulfide-rich environments have symbiotic bacteria in their ctenidia that can chemically synthesize organic molecules that nourish the host. For example, mussels of the genus *Bathymodiolus* that reside near deep-ocean hydrothermal vents are totally dependent upon chemosynthetic symbiotic bacteria for their nutrition.

The body of bivalves is often attached to the shell by one or two paired pedal retractor muscles. Some bivalves have a foot for locomotion. If present, the foot can be extended from the shell by blood pressure and dilated to act as an external anchor while movement is effected by contraction of the retractor muscles. Some bivalves of the family Pectinidae (scallops) lack a foot, but are highly active swimmers through clapping their valves and jetting water through orifices (openings) near the hinge. Some bivalves, such as oysters and giant clams, are sedentary and lack a foot as adults.

Reproduction and development. Bivalves exhibit a wide range of reproductive strategies. Most bivalves are dioecious or have separate sexes, while others exhibit various forms of hermaphrodism. For example, as mature adults, scallops carry both eggs and sperm, while oysters exhibit protandric hermaphrodism in which they first develop as males and in subsequent years change sex to develop ovaries. Most species of bivalves shed eggs and sperm directly into the water, where fertilization occurs; in others, eggs may be held in a brood chamber, where they are fertilized by sperm in incurrent water and released as well-developed larvae into the water. Most bivalves go through several planktonic stages prior to settlement and metamorphosis to their benthic form. As a notable exception to this planktonic larval pattern, freshwater mussels of the family Unionidae produce larvae that are parasitic on fish gills.

Economic importance. A number of bivalve species are harvested for human consumption. According to the United Nations Food and Agriculture Organization, in 2003 world bivalve shellfishery landings (harvest of wild stock) were about 2.6 million metric tons with an estimated value of $2 billion. Many species of bivalves are actively farmed either for human consumption of the meats or for shell products. Worldwide production of farmed bivalves for food consumption in 2003 was 11.1 million metric tons with a value of about $10.8 billion. Most of the gem-quality pearls sold in the world originate from farmed pearl oysters of the genus *Pinctada* in Japan, Australia, and islands of the tropical Pacific. Freshwater pearls are produced from freshwater mussels in the United States, China, and Japan. Estimates of the worldwide trade in farmed pearls exceed $1.5 billion annually. Some species of bivalves are of economic concern as pest organisms or biological invaders. Historically, shipworms of the genus *Teredo* were the bane of wooden sailing ships, and they still pose problems by boring into wooden pilings and weakening piers or other coastal structures. Exotic

species of bivalves intentionally or accidentally introduced into North America have caused millions of dollars worth of damage to public and commercial infrastructure. In the 1950s, Asian freshwater clams of the genus *Corbicula* spread extensively through the waterways of the Tennessee Valley, causing damage to irrigation dams and power generation facilities. In 1985, zebra mussels (genus *Dreissena*) native to the Caspian Sea area in Eurasia were discovered in Lake St. Clair between Ontario and Michigan. Planktonic zebra mussel larvae were probably transported from Europe in the ballast water of a freighter entering the Great Lakes. In less than 10 years, zebra mussels had spread throughout the Great Lakes and down the Mississippi River drainage to New Orleans. Zebra mussels have few natural enemies in North America and have caused millions of dollars worth of damage by clogging power plants, water pumping stations, and other industrial facilities. There is also evidence that the prolific zebra mussels have supplanted native North American freshwater unionid bivalves in many water bodies. *See* SHIPWORM.

Fossils. The fossil record of the Bivalvia can be traced to the Lower Cambrian *Fordilla*, which had a small, round, laterally compressed shell. The early origins of bivalves remain murky, but recent molecular analysis of molluscan DNA indicates that the Bivalvia and other molluscan classes are probably derived from an extinct conchiferan ancestor with bivalves diverging relatively late in the Cambrian. The Ordovician was a major period of bivalve speciation, but throughout the Paleozoic the molluscan Bivalvia remained second to brachiopodan bivalves in species diversity and abundance. A number of ancient molluscan bivalve families first appearing during the Paleozoic Era still have representative species, which include Nuculidae, Nuculanidae, Pteriidae, Arcidae, Pectinidae, Mytilidae, Lucinidae, and Carditidae.

During the Mesozoic Era, the brachiopods declined in importance. Various adaptations of the Bivalvia that developed during the era include active swimming by scallops, deep burrowing by many sediment-dwelling infaunal species, and extremely thick-shelled oysters. It is probable that the diverse adaptations of the Bivalvia to avoid predatory gastropods, arthropods, and fish evolving during the Mesozoic were a major factor in the replacement of the more exposed brachiopods as the dominant bivalves. The evolutionary radiation occurring during the Mesozoic includes the emergence of many species of bivalves that bore into rocks, hard corals, and wood. The emergent family Ostreidae, which includes oysters, remains to the present. The transition from the Mesozoic to Cenozoic began with the extinction of many ancient families and the emergence of several modern families. Important emergent modern families of Bivalvia during the Cenozoic Era include Chamidae, Cardiidae, Mactridae, Solenacidae, Tellinidae, Veneridae, Myidae, and Pandoridae. Michael A. Rice

Bibliography. J. T. Carlton, Introduced marine and estuarine mollusks of North America: An end-of-the-20th-century perspective, *J. Shellfish Res.*, 11:489–505, 1992; C. M. Cavanaugh et al., Symbiosis of methylotrophic bacteria and deep-sea mussels, *Nature*, 325:346, 1987; C. R. Fassler, Pearl farming, *World Aquacult.*, 29(1):6–13, 1998; E. Gosling, *Bivalve Molluscs: Biology, Ecology and Culture*, 2003; R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, Pt. N: *Mollusca 6*, 1969; J. A. Schneider, Bivalve systematics in the 20th century, *J. Paleontol.*, 75:1119–1127, 2001; R. K. Trench et al., Observations on the symbiosis with zooxanthellae among the Tridacnidae, *Biol. Bull.*, 161:180–198, 1981; P. R. Walne, *Culture of Bivalve Mollusks*, 1974.

Black hole

One of the end points of gravitational collapse, in which the collapsing matter fades from view, leaving only a center of gravitational attraction behind. Any other given point in space has a past and a future. A black hole is a region that has a past but no future, or at least none that is accessible to an external observer. The center of a black hole is said to be a singular region of space-time. The existence of such a singularity is predicted by the theory of general relativity. If a star of more than about 3 solar masses has completely burned its nuclear fuel, it should collapse to a black hole. The resulting object is independent of the properties of the matter that produced it and can be completely described by specifying its mass, spin, and charge. The most striking feature of this object is the existence of a surface, called the horizon, which completely encloses the collapsed matter. The horizon is an ideal one-way membrane: that is, particles and light can go inward through the surface, but none can go outward. As a result, the object is dark, that is, black, and hides from view a finite region of space (a hole). Arguments concerning the existence of black holes originally centered on the fact that there are many stars of over 3 solar masses and that there seemed to be no other outcome of collapse than the formation of a black hole. In 1971, however, some direct observational evidence was obtained for a black hole in the binary x-ray system Cygnus X-1. Since that time, black holes have been identified by similar evidence in a number of other x-ray binaries, two notable examples being A0620−00 in the Milky Way Galaxy and LMC X-3 in the Large Magellanic Cloud. There are more than 15 identified black holes, in the Milky Way Galaxy. In addition, supermassive black holes may be responsible for the large energy output of quasars and other active galactic nuclei, and there is growing evidence that black holes exist also at the center of many other galaxies, including the Milky Way Galaxy. *See* GRAVITATIONAL COLLAPSE; RELATIVITY.

Theory. Shortly after Albert Einstein formulated the general theory of relativity in 1916, the solution of the field equations corresponding to a nonrotating black hole was found. For many years this solution, called the Schwarzschild solution, was used to

describe the gravitational attraction outside a spherical star. However, the interpretation of the Schwarzschild solution as a solution for a black hole was not made at the time. More than 20 years elapsed before it was shown that such a black hole could, and probably would, be formed in the gravitational collapse of a nonrotating star of sufficient mass. It was not until 1963 that the solution for a spinning black hole, the Kerr solution, was found. This was particularly important, since most stars are rotating, and the rotation rate is expected to increase when such stars collapse. Although some collapsing, rotating stars might avoid becoming black holes by ejecting matter, thereby reducing their mass, many stars will evolve to a stage of catastrophic collapse in which the formation of a black hole is the only conceivable outcome. However, unlike the case of nonrotating black holes, no one has shown that a collapsing, rotating star of sufficient mass must form a Kerr black hole. On the other hand, it has been shown that if the collapse of a star proceeds past a certain point, the star must evolve to a singularity, that is, an infinitely dense state of matter beyond which no further evolution is possible. Such singularities are found inside black holes in all known black hole solutions, but it has only been conjectured that the singularity produced in a collapse must be inside a black hole. However, the existence of such a naked singularity would have undesirable consequences, such as allowing the violation of fundamental physical laws that appeal to the conservation of mass-energy and to causality. The so-called cosmic censorship theorem is based on the conjecture that formation of a naked singularity is impossible. *See* STELLAR EVOLUTION.

Black hole solutions have also been found for the case in which the black holes have a charge, that is, an electrical as well as a gravitational influence. However, since matter on the large scale is electrically neutral, black holes with any significant charge are not expected in astronomy. Similarly, black hole solutions allow black holes to possess magnetic charge, that is, a magnetic single-pole interaction. Although some elementary-particle theories predict that there should exist particles with magnetic charge, called magnetic monopoles, sufficient experimental evidence is not yet available to confirm their existence. Even if monopoles did exist, they would play little part in the formation of black holes, and so astronomical black holes are expected to be both electrically and magnetically neutral. *See* MAGNETIC MONOPOLES.

Uniqueness theorems about black holes make it likely that at least some Kerr black holes would be formed. Uniqueness theorems address the question of how many kinds of black holes could exist and how complicated their structure could be. These theorems show that black holes must have a simple structure. In fact, the mass, spin, charge, and magnetic charge are all that are needed to specify completely a black hole. Further, any distortion of a black hole, such as is caused by a chunk of matter falling inside, is removed by a burst of radiation. Therefore, although the collapse of a rotating star would be quite complicated, it appears that the final system, the Kerr black hole, would be relatively simple and independent of the details of collapse.

The possible formation of black holes depends critically on what other end points of stellar evolution are possible. Chunks of cold matter can be stable, but their mass must be considerably less than that of the Sun. For masses on the order of a solar mass, only two stable configurations are known for cold, evolved matter. The first, the white dwarf, is supported against gravitational collapse by the same quantum forces that keep atoms from collapsing. However, these forces cannot support a star whose mass exceeds about 1.2 solar masses. (A limiting value of 1.4 solar masses was first found by S. Chandrasekhar and is known as the Chandrasekhar limit. More realistic models of white dwarfs, taking into account nuclear reactions, lower this number somewhat, but the actual value depends on the composition of the white dwarf.) The second stable configuration, the neutron star, is supported against gravitational collapse by the same forces that keep the nucleus of an atom from collapsing. There is also a maximum mass for a neutron star, estimated to be between 1 and 3 solar masses, the uncertainty being due to the poor knowledge of nuclear forces at high densities. Both white dwarfs and neutron stars have been observed, the former for many years at optical wavelengths and the latter more recently in the studies of radio pulsars and binary x-ray sources. *See* NEUTRON STAR; PULSAR; WHITE DWARF STAR.

It would appear from the theory that if a collapsing star of over 3 solar masses does not eject matter, it has no choice but to become a black hole. There are, of course, many stars with mass larger than 3 solar masses, and it is expected that a significant number of them will reach the collapse stage without having ejected sufficient matter to take them below the 3-solar-mass limit. Further, more massive stars evolve more rapidly, enhancing the rate of formation of black holes. It seems reasonable to conclude that a considerable number of black holes should exist in the universe. One major problem is that, since the black hole is dark, it is itself essentially unobservable. Fortunately, some black holes may be observable in the sense that the black hole maintains its gravitational influence on other matter, and thus it can make its presence known. This is precisely how, in a binary system in which a black hole has a massive stellar companion, some of the ejecta from the evolving massive star are accreted by the black hole and yield a distinct x-ray signature. Otherwise, the detection of a black hole would be limited to observations of the collapse of a star to a black hole, a rare occurrence in the Milky Way Galaxy and one which happens very quickly. Astronomers expect that such collapses would occur about once every thousand years in our galaxy. By continuously surveying many thousands of galaxies, one could therefore hope to detect black hole formation. This is the goal of detectors under construction to search for gravitational waves from stars collapsing and binary neutron stars

merging to form black holes. *See* GRAVITATIONAL RADIATION.

Structure. For a nonrotating black hole, the radius of the horizon (Schwarzschild radius) is determined entirely by the mass. Defining R so that the surface area of the spherical horizon is $4\pi R^2$, the equation relating R to the mass M is $R = 2GM/c^2$, where G is the constant of gravity and c is the speed of light. Classical general relativity would allow M to take on all possible values, but quantum effects suggest that the lowest possible value of M is about 10^{-8} kg (2×10^{-8} lb). However, the lower-mass black holes may not be astronomically relevant, since collapsing stars with masses less than about a solar mass (2×10^{30} kg or 4×10^{30} lb) would become white dwarfs or neutron stars. It is thought that low-mass black holes could exist only if they were created at the time of the origin of the universe.

An astronomical black hole of 5 solar masses would have a radius of about 20 km (12 mi). This size is comparable to that expected for neutron stars. The density to which matter would have to be compressed in order to form such a black hole is comparable to that found in neutron stars or in the nuclei of atoms. Black holes with a mass larger than 1000 solar masses may be formed from the collapse of supermassive stars. The existence of such supermassive stars has not been proven, but supermassive stars and supermassive black holes are thought to be at the cores of quasars and active galactic nuclei. Infall of matter released by colliding stars in the dense nuclei of galaxies feeds such black holes, releasing gravitational energy as the matter falls in, and provides the energy that powers active galactic nuclei. A supermassive black hole with a mass of a few thousand solar masses would have a radius comparable to the radius of the Earth. Black holes of 10^6 to 10^9 solar masses are likely to exist in the center of some galaxies, including the Milky Way Galaxy. These could be formed either in the collapse of a supermassive star with a possible subsequent accretion of matter, or in the coalescing of a large number of black holes of more modest mass. The density required to form these very massive black holes is low, approaching that of ordinary terrestrial densities. There is now considerable evidence that black holes exist in the cores of many galaxies. *See* GALAXY, EXTERNAL; MILKY WAY GALAXY; QUASAR; SUPERMASSIVE STARS.

For nonrotating black holes, the horizon is also a surface of infinite redshift; that is, light emitted from just above the surface reaches a distant observer with a much lower frequency and energy than it had when it was emitted. As a result, an object that falls into a black hole appears to an observer to take an infinite time to reach the Schwarzschild radius, with the observed light coming from the object redshifting quickly to darkness as the approach is made. The picture would be quite different for an observer associated with the falling object. The "rider" would reach the Schwarzschild radius in a finite time, feeling nothing more unusual than gravitational tidal forces. However, once inside, the person would be trapped since even light which moves outward cannot escape. The theory predicts that this individual, as well as everything else within the horizon, would be crushed to infinite density within a short time.

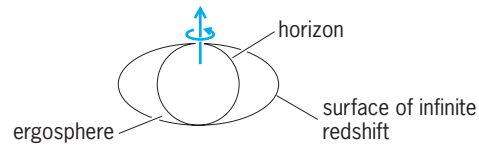

Fig. 1. The Kerr, or rotating, black hole. (*After P. C. Peters, Black holes: New horizons in gravitational theory, Amer. Sci., 62(5):575–583, 1974*)

For rotating black holes, the surface of infinite redshift lies outside the horizon except at the poles, as illustrated in **Fig. 1**. The region between the two surfaces is called the ergosphere. This region is important because it contains particle trajectories that have negative energy relative to an outside observer. Roger Penrose showed that it is possible, though perhaps unlikely astronomically, to use these trajectories to recover even more than the rest mass energy of matter sent into a rotating black hole, the extra energy coming from the slowing down of the rotation of the black hole. Others have proposed that radiation incident on a rotating black hole could be similarly amplified.

The black hole solutions of general relativity, ignoring quantum-mechanical effects as described below, are completely stable. Once massive black holes form, they will remain forever; and subsequent processes, for example, the accumulation of matter, only increase their size. Two black holes could coalesce to form a single, larger black hole, but a single black hole could not split up into two smaller ones. This irreversibility in time led researchers to consider analogies between black holes and thermal properties of ordinary matter, in which there is a similar irreversibility as matter becomes more disordered as time goes on. In 1974, Steven Hawking showed that when quantum effects are properly taken into account, a black hole should emit thermal radiation, composed of all particles and quanta of radiation that exist. This established the black hole as a thermal system, having a temperature inversely proportional to its mass. Since a radiating system loses energy and therefore loses mass, a black hole can shrink and decay if it is radiating faster than it is accumulating matter. For black holes formed from the collapse of stars, the temperature is about 10^{-7} K (2×10^{-7} °F above absolute zero, −459.67°F). Regardless of where such black holes are located, the ambient radiation incident on the black hole from other stars, and from the big bang itself, is much larger than the thermal radiation emitted by the black hole, implying that the black hole would not shrink. Even if the ambient radiation is shielded from the black hole, the time for the black hole to decay is much longer than the age of the universe, so that, in practice, black holes formed from collapse of a star are essentially as stable as they were thought to be before the Hawking radiation was predicted.

Theoretically, black holes of any mass could have been created at the beginning of the universe in the big bang. For smaller-mass black holes, the Hawking radiation process would be quite important, since the temperatures would be very high. For example, a black hole created with mass of 10^{12} kg (2×10^{12} lb), about the mass of a mountain, would have just radiated away all of its mass, assuming that no mass had been accreted in the meantime. Black holes created with a mass smaller than 10^{12} kg would have disappeared earlier, and those with a larger mass would still exist. The final stage of evaporation would be quite violent and would take place quickly. As a black hole radiates, it loses mass and its temperature rises. But a higher temperature means that it radiates and loses mass at a faster rate, raising its temperature even further. The final burst, as the black hole gives up the remainder of its mass, would be a spectacular event. The final emission would contain all radiation and particles that could exist, even those not generated by existing accelerators. At present, there is no evidence that points to the existence of black holes with small mass or to their evaporation by the Hawking radiation process. Philip C. Peters; Joseph Silk

Observation. Observations of black holes—which, by definition, are not directly detectable—rely on indirect methods using the gravitational interaction of the black hole with its surroundings. The first successful detections were accomplished in binary x-ray systems. These systems are thought to be close binary systems composed of a massive and very dense object—a black hole or a neutron star—and a companion star. Due to strong gravitational forces, the companion star loses mass that settles down into an accretion disk around the compact object (**Fig. 2**). Frictional forces cause the matter in the disk to spiral inward and also heat up the disk, causing it to emit thermal x-rays.

A neutron star is distinguished from a black hole as the compact object by analyzing the internal motion of the binary and calculating the mass of the compact body. In the notable case of the x-ray binary system Cygnus X-1, the mass determination gives a value of about 10 solar masses, a value much larger than the maximum mass for a gravitationally stable neutron star and therefore strong evidence for the existence of a black hole in this system. About 30 candidates for black hole x-ray binaries have been detected in the Milky Way Galaxy, of which about half have secure mass estimates that support the black hole assumption.

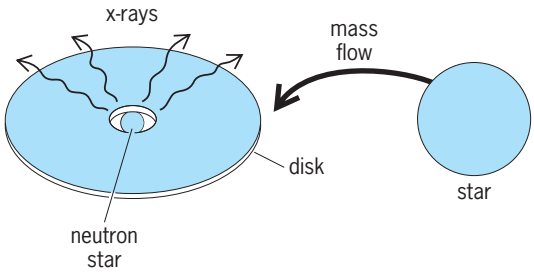

Fig. 2. Schematic diagram of a binary x-ray source. (*After P. C. Peters, Black holes: New horizons in gravitational theory, Amer. Sci., 62(5):575–583, 1974*)

Another line of observational evidence for the existence of black holes is drawn from active galactic nuclei (AGNs), whose luminosities are several times larger than those of the largest nonactive galaxies. Variations on time scales of tens of hours in the luminosities of AGNs indicate a central engine whose diameter should be of the order of the product of the variation time scale and the speed of light. To generate observed AGN luminosities in a spherical volume defined by this diameter in nuclear processes like those taking place in stars, one would need such large quantities of matter that the gravitational potential energy would far exceed the nuclear energy. Therefore, the origin of the luminosity is attributed to the transformation of potential energy into radiation.

The arguments in favor of a scenario in which potential energy is released in the accretion of gas onto a supermassive black hole are twofold. First, optical and ultraviolet emission lines of AGNs show so-called Doppler broadening corresponding to velocities of 2000–10,000 km/s (1200–6000 mi/s), indicating gas moving around a massive compact region. Second, AGNs display jets that are well collimated and straight, two properties that a central engine consisting of many bodies cannot accomplish easily, favoring a single black hole as the central engine. *See* DOPPLER EFFECT.

Mass determinations of black holes in AGNs rely on measurements of the properties of the accretion disk, such as the rotational velocity of this disk at different separations from the central engine after correcting for projection effects. In the case of a Keplerian rotation around the center, as for the galaxy M87, the rotational velocity should become smaller with increasing radius, and should be proportional to the square root of the enclosed mass. The mass of the central black hole is derived to be 3×10^9 solar masses for M87. Another frequently used technique is to measure the spectrum of the accretion disk. Since the accretion disk is optically thick, radiation on its way through the disk will be scattered many times before it leaves the surface of the disk, generating a disk nearly in thermodynamic equilibrium. The spectra of such disks have the form of a black body spectrum with characteristic temperature depending on the black hole mass and the mass accretion rate onto the black hole. Typical values derived using this method find black holes with masses in the range 1×10^6 to 3×10^9 solar masses.

The majority of estimates of the masses of black holes in the centers of normal galaxies are based on techniques using stellar dynamical methods. The orbits of stars in galaxies are affected significantly only by gravitational forces exerted on them. Unlike gas in an accretion disk, stars have significant random motions. A measure of the random motions of the stars is the so-called velocity dispersion, which characterizes the distribution of the different stellar velocities. Random motions of stars contribute to their overall velocity in addition to their systemic rotational

velocity. Therefore, it is necessary to measure not only the rotational velocity at different radii, as in the case of a gas accretion disk; but also the velocity dispersion at different radii. As a further observational input to measure the black hole mass, the galaxy's luminosity density at different radii is needed. This is necessary because the black hole mass determination involves modeling the stellar structure in the galaxy. In practice, one assumes a model for the orbital structure of the stars in the galaxy and calculates the density of the stars at each radius from the luminosity density assuming a constant mass-to-light ratio. From the model, it is now possible to calculate the velocity dispersion from the rotational velocity and vice versa. In the case of a black hole at the center, it is not possible to use a constant mass-to-light ratio at all radii to recover the observed velocity dispersion from the observed rotational velocity of the galaxy using the model. Instead, one finds a steep increase in the mass-to-light ratio when going to small radii. Using the mass-to-light ratio, it is then possible to calculate the black hole mass. Approximately 30 black hole masses in the centers of galaxies have been obtained using this and similar techniques, and more are being calculated.

Another stellar dynamical measurement of a black hole mass is the one in the center of our Milky Way Galaxy. In this case the orbits of individual stars have been measured in the vicinity of the center over a period of more than 15 years. The orbits of the stars are Kepler ellipses, which allow calculation of the mass of the central body with high precision. The calculations reveal a black hole mass of 3×10^6 solar masses at the center of the Milky Way Galaxy; it is believed to be associated with the radio source Sagittarius A*. *See* ASTROPHYSICS, HIGH-ENERGY; BINARY STAR; X-RAY ASTRONOMY.

Sadegh Khochfar; Joseph Silk

Bibliography. J. S. AI-Khalili, *Black Holes, Wormholes and Time Machines*, 1999; M. Begelman and M. Rees, *Gravity's Fatal Attraction: Black Holes in the Universe*, 1998; S. W. Hawking, *A Brief History of Time: From the Big Bang to Black Holes*, 2d ed., 1996; J.-P. Luminet, *Black Holes*, 1992; F. Melia, *The Black Hole at the Center of Our Galaxy*, 2003; S. L. Shapiro and S. A. Teukolsky, *Black Holes, White Dwarfs, and Neutron Stars: The Physics of Compact Objects*, 1983; E. F. Taylor and J. A. Wheeler, *Exploring Black Holes: Introduction to General Relativity*, 2000; K. S. Thorne, *Black Holes and Time Warps*, 1994; R. Wald, *Space, Time, and Gravity: The Theory of the Big Bang and Black Holes*, 2d ed., 1992.

Black pepper

One of the oldest and most important of the spices. It is the dried, unripe fruit of a weak climbing vine, *Piper nigrum*, a member of the pepper family (Piperaceae), and a native of India or Indomalaysia. The fruits are small one-seeded berries which, in ripening, undergo a color change from green to red to yellow. When in the red stage, they are picked, sorted, and dried. The dry, wrinkled berries (peppercorns) are ground to make the familiar black pepper of commerce. White pepper is obtained by grinding the seed separately from the surrounding pulp. *See* PIPERALES; SPICE AND FLAVORING.

Perry D. Strausbaugh; Earl L. Core

Black Sea

A semienclosed marginal sea with an area of 420,000 km^2 (160,000 mi^2) bounded by Turkey to the south, Georgia to the east, Russia and Ukraine to the north, and Romania and Bulgaria to the west (**Fig. 1**). The physical and chemical structure of the Black Sea is critically dependent on its hydrological balance. As a result, it is the world's largest anoxic basin. It has recently experienced numerous types of environmental stress.

The Black Sea consists of a large basin with a depth of about 2200 m (7200 ft). The continental shelf is mostly narrow except for the broad shelf in the northwest region. Fresh-water input from rivers (about 350 km^3 y^{-1}; 84 mi^3 y^{-1}), especially the Danube, Dniester, and Don, and precipitation (about 300 km^3 y^{-1}; 72 mi^3 y^{-1}) exceeds evaporation (about 350 km^3 y^{-1}; 84 mi^3 y^{-1}). Low-salinity surface waters are transported to the Mediterranean as a surface outflow (flux of 600 km^3 y^{-1}; 144 mi^3 y^{-1}). High-salinity seawater from the Mediterranean enters the Black Sea as a subsurface inflow through the Bosporus (flux of 300 km^3 y^{-1}; 72 mi^3 y^{-1}). This estuarine circulation (seawater inflow at depth and fresh-water outflow at the surface) results in an unusually strong vertical density gradient determined mainly by the salinity. Thus the Black Sea has a two-layered structure with a lower-salinity surface layer (0–100 m or 328 ft; volume = 42,000 km^3 or 10,000 mi^3; salinity = 18.5) and a higher-salinity deep layer (100–2200 m or 328–7218 ft; volume = 780,000 km^3 or 187,000 mi^3; salinity = 22.3).

The vertical stratification has a strong effect on the chemistry of the sea. Respiration of particulate organic carbon sinking into the deep water has used up all the dissolved oxygen. Thus, conditions favor bacterial sulfate reduction and high sulfide concentrations. As a result, the Black Sea is the world's largest anoxic basin and is commonly used as a modern analog of an environment favoring the formation of organic-rich black shales observed in the geological sedimentary record.

Circulation. Representative vertical profiles of temperature (T), salinity (S), and density (kg/m^3 and σ_t) for the central, western Black Sea are shown in **Fig. 2**. There are two important factors that control these distributions: (1) The only source of high-salinity water is through the Bosporus, which has a maximum sill depth of only 35 m (115 ft). (2) The temperature of the Bosporus inflow is about 15°C (59°F); thus the only source of cold water ($T < 15$°C or 59°F) is the surface of the Black Sea. The main

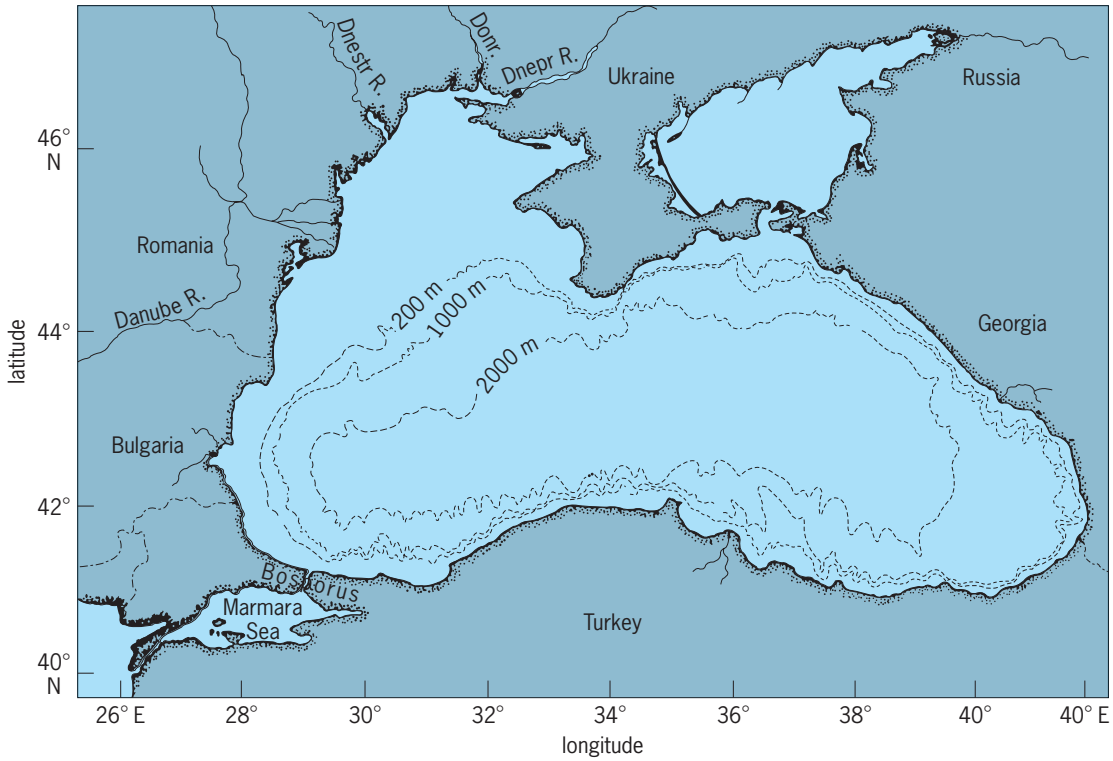

Fig. 1. Map of the Black Sea showing contiguous countries, major rivers, and bathymetry.

features in the vertical profiles are the seasonally and spatially variable surface layer, a cold intermediate layer (CIL; $T < 8^\circ$C or 46.4°F) at about 50 m (164 ft) and a deep layer with slowly increasing salinity with depth. The cold intermediate layer forms annually by the convective processes associated with winter cooling of surface waters in the central gyre areas and on the northwestern shelf. It has a residence time of about one year. The deep water ($S =$ 22.3) forms from an average mixture of about 1 part Bosporus inflow ($S = 35.0$) with 3.3 parts of cold intermediate layer ($S = 18.5$) which mix together on the shelf region just east of the Bosporus. This water is then injected into the deep basin. The resulting residence time of the deep layer is about 600 years. Since the residence time is the average time that a parcel of water spends beneath the surface layer of the ocean, its temperature and salinity are no longer influenced by the Sun, evaporation, or fresh-water run-off. As a result, T and S in the parcel of water can change only by mixing with other parcels.

The primarily cyclonic (counterclockwise) surface circulation results from the cyclonic nature of the wind field. The main current is the Rim Current, which flows along the edge of the continental slope completely around the basin. A series of anticyclonic eddies are confined between the coast and the Rim Current. A consequence of the surface circulation is that the isopycnal (constant-density) surfaces in the center of the Black Sea rise closer to the surface (called the doming effect). Thus the first appearance of hydrogen sulfide (H_2S), which always occurs at a density of about 1016.15 kg/m^3 ($\sigma_t = 16.15$), is found at depths of 80–100 m (262–328 ft) in the central basin and >150–200 m (492–656 ft) along the margins.

Chemical distributions. As a consequence of the hydrologic balance and circulation, dissolved oxygen is confined to the surface layer while the deep water is anoxic. The distributions of oxygen, sulfide, nitrate, nitrite, manganese, iron, ammonia and phosphate from a density of 1015 to 1016.5 kg/m^3 ($\sigma_t =$ 15.0 to 16.5) [approximately the upper 300 m or 984 ft] are shown in **Fig. 3**. The distributions are plotted versus density rather than depth, because while the depth of specific features may vary, due to the doming effect, they always occur at the same density. Anthropogenic effects have had a significant impact on many of the nutrients in the surface layer. For example, the inventory and concentration of nitrate has increased by about a factor of 5 since 1969 due to eutrophication. Dissolved silicate concentrations have decreased by 50% over the same time period due to dam construction.

In the deep anoxic layer, hydrogen sulfide increases to values over 400 μM and methane increases to over 10 μM.

At the oxic-anoxic interface, there is a suboxic zone. The suboxic zone has a thickness of about 10 to 50 m (33 to 164 ft) and is found between the oxygen-containing surface layer and the sulfide-containing deep layer. In the suboxic zone, both oxygen and sulfide have very low concentrations and no perceptible horizontal or vertical gradients. This zone appears to be a site of oxidation-reduction reactions involving species of nitrogen, manganese,

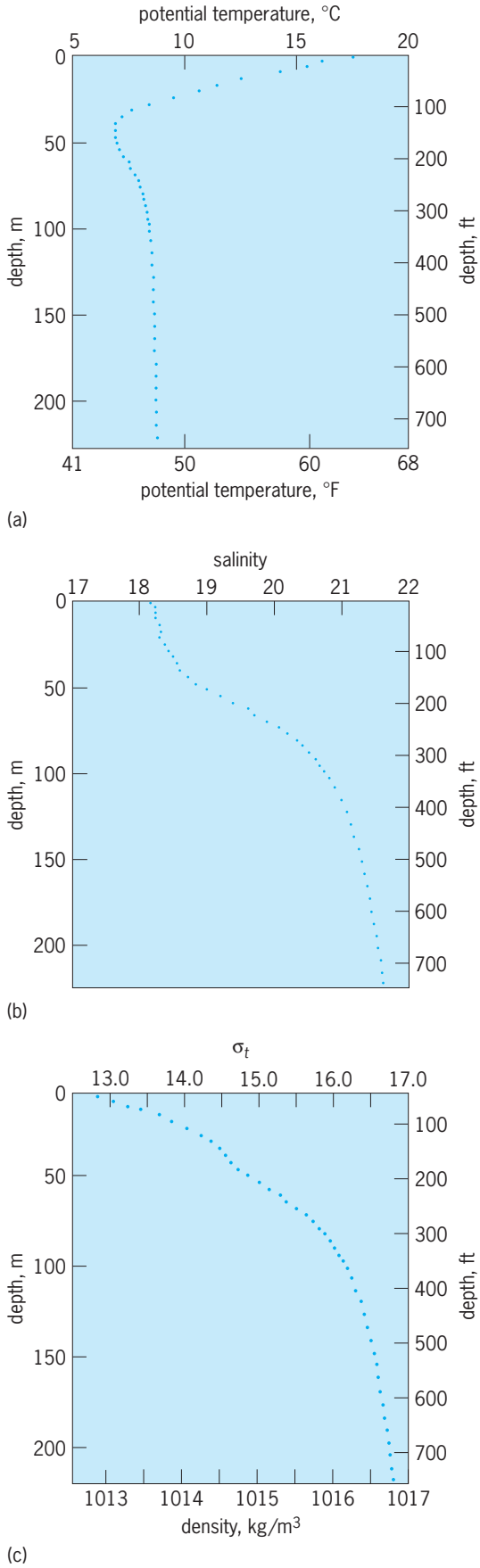

Fig. 2. Representative vertical profiles for 0–200 m (0–656 ft) of (*a*) temperature, (*b*) salinity, and (*c*) density (kg/m^3 and σ_t).

and iron, such as reactions (1)–(4).

$$3NO_3^- + 5NH_4^+ \rightarrow 4N_2 + 9H_2O + 2H^+ \quad (1)$$

$$2NO_3^- + 5Mn^{2+} + 4H_2O \rightarrow N_2 + 5MnO_2(s) + 8H^+ \quad (2)$$

$$2NO_3^- + 10Fe^{2+} + 24H_2O \longrightarrow N_2 + 10Fe(OH)_3(s) + 18H^+ \quad (3)$$

$$2NH_4^+ + 3MnO_2(s) + 4H^+ \rightarrow N_2 + 3Mn^{2+} + 6H_2O \quad (4)$$

Consequences of these reactions are: (1) The upward flux of sulfide appears to be oxidized by oxidized forms of manganese and iron rather than directly by oxygen. (2) The downward flux of nitrate (NO_2^- and the upward flux of ammonium (NH_4^+) are consumed at the same density level and are converted to nitrogen gas (N_2). (3) Manganese cycling between the II, III, and IV oxidation states appears to play a key catalytic role in these reactions. (4) Phosphate is adsorbed by particulate forms of manganese and iron, resulting in an intense phosphate minimum. (5) The suboxic zone reactions act as a nutrient trap preventing the upward flux of nutrients from the deep water from entering the surface layer and euphotic zone. The suboxic zone also contains high concentrations of bacteriochlorophylls and bacteriocarotenoids characteristic of the brown sulfur bacteria *Chlorobium*. Anaerobic photosynthesis, in which H_2S oxidation supports the phototrophic reduction of carbon dioxide (CO_2) to organic carbon, is also a strong possibility.

Biology. Before the 1970s the Black Sea had a highly diverse and healthy biological population. Its species composition was similar to that of the Mediterranean but with less quantity. The phytoplankton community was characterized by a large diatom bloom in May-June followed by a smaller dinoflagellate bloom. The primary zooplankton were copepods, and there were 170 species of fish, including large commercial populations of mackerel, bonito, anchovies, herring, carp, and sturgeon.

Since about 1970 there have been dramatic changes in the food web due to anthropogenic effects and invasions of new species. It is now characterized as a nonequilibrium, low-diversity, eutrophic state. The large increase in input of nitrogen due to eutrophication and decrease in silicate due to dam construction have increased the frequency of noxious algal blooms and resulted in dramatic shifts in phytoplankton from diatoms (siliceous) to coccolithophores and flagellates (nonsiliceous). The most dramatic changes have been observed in the northwestern shelf and the western coastal regions, which have the largest anthropogenic effects. The water overlying the sediments in these shallow areas frequently go anoxic due to this eutrophication. In the early 1980s the grazer community experienced major increases of previously minor indigenous species such as the omnivorous dinoflagellate *Noctilluca scintillans* and the medusa *Aurelia aurita*. The ctenophore *Mnemopsis leidyi* was imported at the end of the 1980s from the east coast of

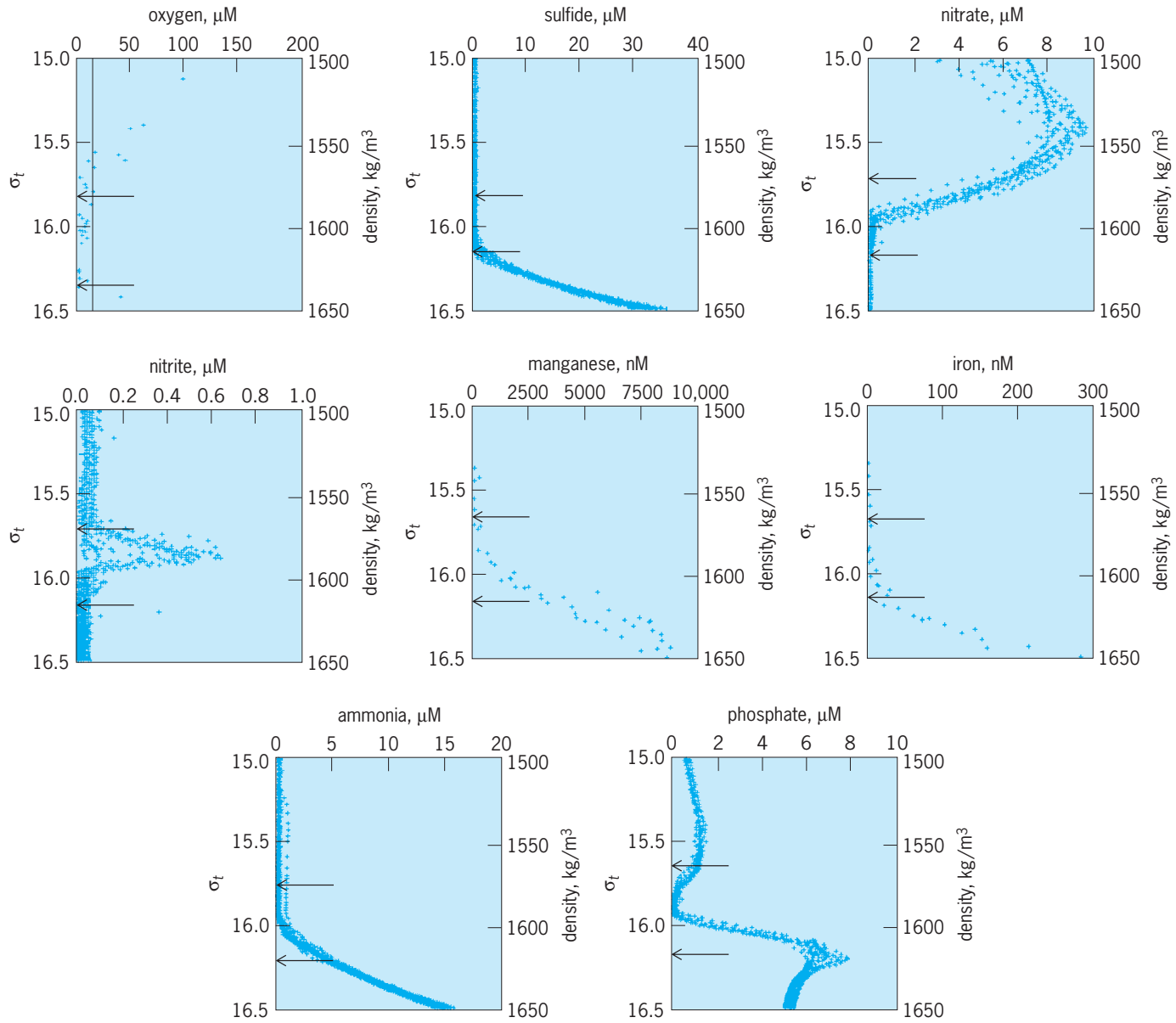

Fig. 3. Distributions of oxygen, sulfide, nitrate, nitrite, manganese, iron, ammonia, and phosphate from 1500 to 1650 kg/m^3 ($\sigma_t = 15.0$ to 16.5).

the United States as ballast water in tankers and experienced an explosive unregulated growth. These changes plus overfishing resulted in a collapse of commercial fish stocks during the 1990s.

Geological history. The geological evolution of the Black Sea has been recorded in the sediments. Prior to 7800 years ago, the sediment record suggests that the Black Sea was a fresh-water lake that formed during a period of lower sea level, when the Black Sea was isolated from the Mediterranean (Unit III). At about 7800 years ago, sea level rose to the point where it began to enter the Black Sea. Water column anoxia in the deep water was initiated, and the present two-layer structure developed. The sediment record at this time is characterized by a microlaminated, organic-rich (up to 20%) sapropel (Unit II). The onset of Unit II sapropel deposition was rapid and synchronous across the basin and at all water depths greater than 200 m (656 ft). At about 2000 years before present, the salinity distribution in the Black Sea had reached its present distribution, and since then the sediment record has been characterized by a laminated, moderately organic rich (~2%), coccolith-rich marl (Unit I). Remains of the calcium carbonate ($CaCO_3$)-secreting coccolithophorid *Emiliania huxleyi* are used to define this layer as the surface water salinity had to be at least $S = 11$ for their production. *See* EUTROPHICATION; SAPROPEL.

James W. Murray

Bibliography. M. A. Arthur and W. E. Dean, Organic-matter production and preservation and evolution of anoxia in the Holocene Black Sea, *Paleoceanography*, 13:395–411, 1998; L. Codispoti et al., Chemical variability in the Black Sea: Implications of continuous vertical profiles that penetrated the oxic/anoxic interface, *Deep-Sea Res.*,

38:S691–S710, 1991; C. Humborg et al., Effect of Danube River dam on Black Sea biogeochemistry and ecosystem structures, *Nature*, 386:385–388, 1997; J. W. Murray et al., Hydrographic properties and ventilation of the Black Sea, *Deep-Sea Res.*, 38:S663–S689, 1991; J. W. Murray et al., Oxidation-reduction environments: The suboxic zone in the Black Sea, in *Aquatic Chemistry: Interfacial and Interspecies Processes*, American Chemical Society, 1995; T. Oguz et al., Circulation in the surface and intermediate layers of the Black Sea, *Deep-Sea Res.*, 140:1597–1612, 1993; D. J. Repeta et al., Evidence for anoxygenic photosynthesis from the distribution of bacteriochlorophylls in the Black Sea, *Nature*, 342:69–72, 1989; Y. Zaitsev and V. Mamaev, *Marine Biological Diversity in the Black Sea: A Study of Change and Decline*, GEF Black Sea Environmental Programme, United Nations, 1997.

Black shale

A distinctive, fine-grained mudrock that derives its dark color from an unusually high content of organic matter, ranging from around 2% to over 50% of the weight of the rock (typical shales contain less than 1% organic matter). Most black shales can split into paper-thin sheets, and these are sometimes called paper shales. *See* SHALE.

Formation. The organic content of black shales is derived primarily from plankton and, to a lesser extent, small woody particles. Under normal deposition conditions in seas and lakes, planktonic organic matter is rapidly degraded by dissolved oxygen in the water column with little, if any, surviving to be buried. Two models explain the special conditions needed to produce organic-rich rocks. One model considers that black shales are likely to form only where the water column is stagnant such that the dissolved oxygen is rapidly consumed by decaying organic matter, producing oxygen-free (anoxic) conditions. Anoxia is considered to hinder the decay of organic matter and promote its burial. The modern-day Black Sea is an example of an anoxic basin in which organic-rich sediments are accumulating. The other model considers that areas of extremely high marine productivity, such as the upwelling zones off the coast of North and South America, are likely to produce organic-rich sediments. In these settings, the oxygen available in the water column is simply consumed by the large amount of decaying organic matter, some of which survives and is buried in the sediments. Both models involve an anoxic condition for deposition. In the former, it is due to a poor oxygen supply, whereas in the latter it is due to a large oxygen demand. *See* ANOXIC ZONES; BLACK SEA; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; MARINE SEDIMENTS; SEDIMENTOLOGY; UPWELLING.

Fossil content. The lack of oxygen in black shale depositional environments often ensures excellent fossil preservation due to the lack of scavengers and currents on the seabed or lakebed. Fossil diversity is often restricted to organisms, such as fish and ammonites, which lived above the uninhabitable seabed but sank to it after death. However, fossils of bottom-living organisms are also encountered in black shales. Their origin is controversial, although it seems likely that they colonized during brief periods with improved levels of bottom-water oxygen.

Global significance. Black shales are of unrivaled economic importance because they are sources of nearly all the world's hydrocarbons (coals are the only other significant source). During burial, alteration of organic matter releases complex hydrocarbons, which migrate and are trapped in oil and gas reservoirs. In addition, black shales are important for understanding many of the mass extinction events of the fossil record. Many of these extinctions coincided with intervals when black shale deposition was especially widespread. It appears that many marine extinctions were caused by the widespread development of stagnant (anoxic) conditions. Such conditions typify two of the largest mass extinctions of all time—the Late Devonian (365 million years ago) and the Late Permian (250 million years ago). *See* DEVONIAN; EXTINCTION (BIOLOGY); PERMIAN.

Paul B. Wignall

Bibliography. A. Hallam and P. B. Wignall, *Mass Extinctions and Their Aftermath*, Oxford University Press, 1997; B. J. Katz (ed.), *Petroleum Source Rocks*, Springer, 1995; P. B. Wignall, *Black Shales*, Oxford University Press, 1994.

Blackberry

Any of several species of the genus *Rubus* (family Rosaceae) having fruit consisting of many drupelets attached to a common fleshy, elongated core (receptacle) which is removed with the fruit. Ripe fruit is usually black or dark purple, and often sweet and flavorful. The bushy plants have perennial roots from which arise long, often thorny, biennial stems (canes) with compound leaves (see **illus.**). Many species are native to temperate regions, especially in the Northern Hemisphere, to which they are best

Blackberry fruit on a thorny, biennial cane.

adapted. They are commonly found on the edges of forests, along streams, and in clearings. Because of their thorns and prolific growth habit, blackberries are a nuisance in some areas. *See* ROSALES.

Blackberry cultivation began early in the nineteenth century, mainly in North America. Several species have been domesticated and used in the development of cultivars. Horticulturally important cultivars are separated into trailing (prostrate) types which require a trellis for support, and erect or semierect types which do not. Blackberries grow in most well-drained soils, except very sandy types, which may be too dry. Erect blackberries are more winter-hardy and grow best in cool, humid summers, whereas some trailing types, also called dewberries, can be grown in hot, dry climates if irrigated. Canes of hardy cultivars can withstand air temperatures below −20°F (−29°C) but require protection from more severe cold or drying winds. Bees are attracted to blackberry flowers and are required for good fruit production of cross-pollinated cultivars. Most blackberries are self-pollinating but still benefit from visits by bees and other pollinating insects.

Blackberries are vegetatively propagated from root suckers or root cuttings, or by burying the cane tips of trailing types, causing them to form new plants. *See* PLANT PROPAGATION; REPRODUCTION (PLANT).

Commercial blackberry production occurs mainly in the United States, but appreciable quantities are grown in the United Kingdom and in New Zealand. In commercial plantings in the United States, harvesting is often done by machines which shake the canes and catch the ripe fruit, most of which is frozen or canned for use in bakery products and yogurt or made into jelly, jam, or wine. Some fruit is hand-harvested and sold fresh. The Pacific Coast states account for about 80% of the annual North American production, with Oregon the major producer. The leading cultivars in this region are Thornless Evergreen, Marion, Olallie, and the blackberry–raspberry hybrids, Boysenberry and Loganberry. Cultivars grown in other producing regions include Cherokee, Comanche, and Dirksen Thornless (central United States), and Brazos and Humble (Texas). Blackberries, especially thornless cultivars such as Thornfree, Smoothstem, Dirksen Thornless, Black Satin, and Hull Thornless, are popular in home gardens and in "you-pick" enterprises.

Common diseases of blackberry are crown gall, a bacterial disease, and fungus diseases such as verticillium wilt, anthracnose, leaf and cane spot, double blossom (rosette), and orange rust. A prevalent viral disease (Rubus sterility) has symptoms of increased plant vigor, poor fruit set, and misshapen fruit. Although blackberries are often comparatively free of injury due to insects, canes are occasionally attacked by borers and the roots by white grubs. *See* FRUIT; PLANT PATHOLOGY. Patrick J. Breen

Bibliography. G. M. Darrow, The cultivated raspberry and blackberry in North America: Breeding and improvement, *Amer. Hort. Mag.*, 46:202–218, 1967; J. W. Hull and F. J. Lawrence, *Growing Blackberries*, USDA Farmer's Bull. 2160, 1972; J. N. Moore, Blackberry production and cultivar situation in North America, *Fruit Varieties J.*, 34:36–42, 1980; J. S. Shoemaker, *Small Fruit Culture*, 5th ed., 1978.

Blackbody

An ideal energy radiator, which at any specified temperature emits in each part of the electromagnetic spectrum the maximum energy obtainable per unit time from any radiator due to its temperature alone. A blackbody also absorbs all the energy which falls upon it. The radiation properties of real radiators are limited by two extreme cases—a radiator which reflects all incident radiation, and a radiator which absorbs all incident radiation. Neither case is completely realized in nature. Carbon and soot are examples of radiators which, for practical purposes, absorb all radiation. Both appear black to the eye at room temperature, hence the name blackbody. Often a blackbody is also referred to as a total absorber. Such a total absorber constitutes a standard for the radiation of nonblackbodies, since Kirchhoff's law demands that the blackbody radiate the maximum amount of energy possible in any wavelength interval. For an extended discussion of blackbody radiation and Kirchhoff's law *See* GRAYBODY; HEAT RADIATION. Heinz G. Sell; Peter J. Walsh

Blackleg

An acute, usually fatal, disease of cattle and occasionally of sheep, goats, and swine, but not humans. The infection is caused by *Clostridium chauvoei (C. feseri)*, a strictly anaerobic, spore-forming bacillus of the soil. The disease is also called symptomatic anthrax or quarter-evil. The characteristic lesions in the natural infection consist of crepitant swellings in involved muscles, which at necropsy are dark red, dark brown, or blue black. Artificial immunization is possible by use of blackleg aggressin or whole-culture bacterin. Animals surviving an attack of blackleg are permanently immune to its recurrence. *See* IMMUNITY. Leland S. McClung

Bibliography. I. S. Ross (ed.), *Veterinary Handbook for Cattlemen*, 5th rev. ed., 1980; P. Summanen et al., *Wadsworth Anaerobic Bacteriology Manual*, 5th ed., 1993.

Blastoidea

A class of extinct Pelmatozoa in the subphylum Blastozoa which arose in the Ordovician and flourished in the Carboniferous and Permian. Blastoids did not survive the Permo-Triassic mass extinction event. Historically, two orders of the Blastoidea have been recognized: the Fissiculata with exposed hydrospire slits and the Spiraculata with concealed hydrospires. In recognition of their polyphyletic origins, the Spiraculata are now subdivided into the

Troosticrinida, Nucleocrinida, Granatocrinida, and Pentremitida. *See* PELMATOZOA.

Echinoderms, including blastoids, have very complex mesodermal skeletons with many thousands of individual ossicles. The typical blastoid had approximately 190,000 ossicles. The blastoid theca has 18 major plates (three basals, five radials, five deltoids, five lancets) which are attached to a stem typically 10 cm (4 in.) in length. A few blastoid genera were stemless, had asymmetrical thecae, and sat or lay directly on the sea floor. Although the presence of tube feet in blastoids is controversial, these animals fed by capturing food in a brachiolar filtration fan, passing it down to the ambulacra into the mouth through a system of food grooves. Waste products were eliminated from a separate anus. Each of the five ambulacra is composed of a lancet flanked by rows of side plates and outer side plates which carried a branch of the ambulacral food groove to a brachiole. A typical blastoid had approximately 100 side food grooves on each of five ambulacra, which supported one brachiole each. The vast majority of the skeletal ossicles in a blastoid theca is located in this brachiolar system (see **illustration**). *See* ECHINODERMATA.

Respiration in blastoids occurred in the hydrospires, a series of thin calcite folds underneath the ambulacra hanging into the coelom. In fissiculate blastoids, the hydrospires are exposed to seawater. In the other orders, the hydrospires are internal. Water entered through hydrospire pores at the lateral margins of the ambulacra, traversed the hydrospire folds, and exited through spiracles juxtaposed to the mouth. The anal opening is often contained within one spiracle, producing an anispiracle. A reproductive structure, the gonopore, has been identified in some blastoid genera. In life, the mouth and spiracles were covered by a pyramid of small ossicles.

Johnny A. Waters

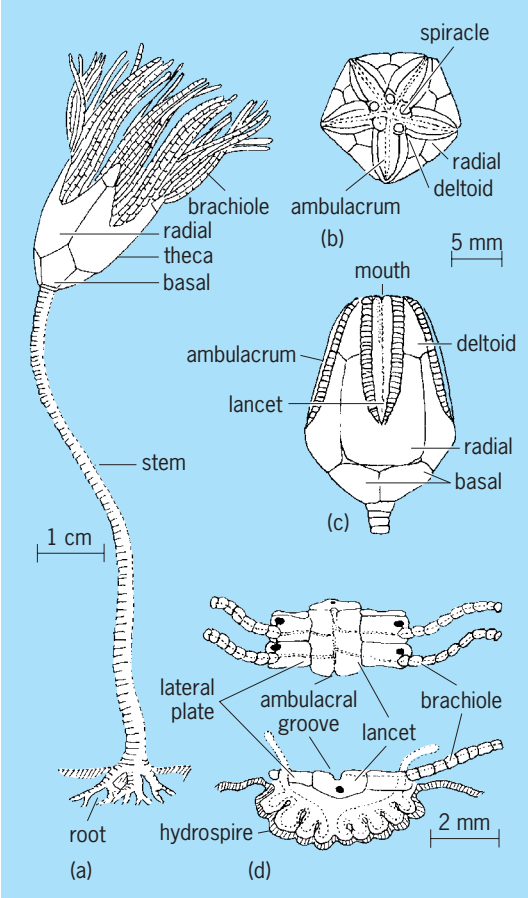

A blastoid, *Pentremites*. (*a*) Whole animal. (*b*) Adoral aspect of theca. (*c*) Lateral aspect of theca. (*d*) Details of ambulacral region in surface view and section.

Bibliography. T. W. Broadhead, *Blastozoa*, 1980; T. W. Broadhead and J. A. Waters (eds.), *Blastoidea in Echinoderms*, 1980; A. S. Horowitz and J. A. Waters, Ordinal level evolution in the Blastoidea, *Lethaia*, 26:207–213, 1993; C. R. C. Paul and A. B. Smith (eds.), *Echinoderm Phylogeny and Evolutionary Biology*, 1988.

Blastomycetes

A group of fungi consisting of yeasts that reproduce asexually by budding or splitting of single cells. In spite of simple morphology, these yeasts are quite diverse in terms of evolutionary origins and environmental roles. The group contains the yeasts used in baking and brewing. Some other yeasts are responsible for infections in humans and animals, while others occur on leaves or fruits. A few species are used for biological control of microbial plant pathogens. A very small number are pathogens of plants. *See* FUNGI; PLANT PATHOLOGY; YEAST.

Classification and identification. "Blastomycetes" was formerly accorded recognition as a class within a larger group consisting of all fungi reproducing asexually, the Deuteromycotina (deuteromycetes). Even at the time, it was recognized that the Blastomycetes contained fungi that were only distantly related: the Cryptococcales (belonging to ascomycetes) and the Sporobolomycetales (belonging to basidiomycetes). A later scheme divided yeasts into three classes: Saccharomycetes (for yeast-forming ascomycetes, notably *Saccharomyces*), Holobasidiomycetes, and Teliomycetes (both including other basidiomycetes in addition to yeast-forming basidiomycetes). Current practice places much less emphasis on a single classification scheme for yeasts per se, and "Blastomycetes" is seldom employed by contemporary taxonomists. Classification systems have changed rapidly with new information available from molecular-genetic techniques. However, certain characters continue to be informative. Formation of new cells via splitting (as opposed to budding) indicates the fission yeasts (such as *Schizosaccharomyces*, Schizosaccharomycetes), which are ascomycetes. The presence of reddish or orange carotenoid pigments, as in *Rhodotorula*, is evidence for affinity with basidiomycetes. An ability to produce ballistospores (spores which are "shot off" as in the "mirror yeast" *Sporobolomyces*) creates a mirror image of a colony on the lid of its petri dish. These mirror yeasts are considered basidiomycetes.

Most yeast-forming fungi are members of Saccharomycetes sensu stricto, are ascomycetes, and reproduce asexually by budding. The term "yeast" is itself without taxonomic significance, except that all yeasts are fungi. *See* ASCOMYCOTA ; BASIDIOMYCOTA; DEUTEROMYCOTINA; EUMYCOTA.

It is relatively easy to identify a microbe as a yeast because the single budding or splitting cell is large relative to bacteria. Assignment of a genus or species name has conventionally been accomplished by constructing a substrate utilization profile based on the ability to use a range of nutrient sources. Multiwell plates have a different substrate in each well, and the yeast growth in each well can be read automatically by computer and an identification returned. Several commercial systems are available. Recent molecular-genetic techniques use sequence analysis of one or more regions of DNA or RNA to identify yeasts as to genus or species.

Economic, medical, and ecological importance. *Saccharomyces cerevisiae*, the brewer's and baker's yeast, has arguably been the most important biotechnology organism in history. Other yeasts play similar but more minor roles in brewing and baking, and undesirable yeasts may impart bad flavors to juices or fermented beverages. The genera *Candida* and *Cryptococcus* contain species presenting medical problems which are especially severe in persons with compromised immune systems. Entirely different species of *Candida* and *Cryptococcus* have been used successfully for biological control of postharvest diseases of fruits. Perhaps the most important role for yeasts in the environment is their ubiquitous presence on plant surfaces, where they act as a major constraint on the growth of other microbes. *See* YEAST INFECTION.

Frank Dugan

Bibliography. J. A. Barnett, A history of research on yeasts 8: Taxonomy, *Yeast*, 21:1141–1193, 2004; J. A. Barnett, R. W. Payne, and D. Y. Yarrow, *Yeasts: Characteristics and Identification*, 3d ed., Cambridge University Press, 2000; C. P. Kurtzman and J. W. Fell (eds.), *The Yeasts: A Taxonomic Study*, 4th ed., Elsevier, 1998.

Blastulation

The formation of a segmentation cavity or blastocoele within a mass of cleaving blastomeres and rearrangement of blastomeres around this cavity in such a way as to form the type of definitive blastula characteristic of each species. The blastocoele originates as an intercellular space which sometimes arises as early as the four- or eight-cell stage. Thus blastulation is initiated during early cleavage stages, and formation of the definitive blastula is thought to terminate cleavage and to initiate gastrulation. Accordingly, cleavage and blastulation are simultaneous events which follow activation of the egg and precede the next major step in development, namely, gastrulation. Initially the diameter of the blastula is no greater than that of the activated egg; subsequently it increases. *See* CLEAVAGE (DEVELOPMENTAL BIOLOGY); GASTRULATION.

Blastula. The blastula is usually a hollow sphere. Its wall may be only one cell thick in many species, but in others it is several cells in thickness. In eggs which contain considerable amounts of yolk the blastocoele may be eccentric in position, that is, shifted toward the animal pole. The animal portion of its wall is always completely divided into relatively small cells, whereas the vegetative portion tends to be composed of relatively large cells and may be incompletely cellulated in certain species. The blastocoele contains a gelatinous or jellylike fluid, which originates in part as a secretion by the blastomeres and in part by passage of water through the blastomeres or intercellular material, or both, into the blastocoele.

The chemical nature of this secretion appears to vary in different species. In some it is said to be albuminous, in others to contain polysaccharides. It also contains certain salts. It possesses a viscosity which renders it visible when released from the blastocoele by puncture of the wall of the latter; however, the viscosity changes with time. The pH of this substance may be the same as that of seawater, or it may be more alkaline. It usually stains in preserved material, and stains most vividly during concentration of blastocoele fluid in those species where the blastocoele becomes reduced in size during gastrulation until it is eventually crowded out of existence. If fluid is withdrawn from the blastocoele with a hypodermic needle, the wall of the blastula crumples, but the quantity of blastocoele fluid is quickly restored to normal, as is the shape of the blastula. The volume of blastocoele fluid can also be reduced osmotically by placing blastulas in sucrose seawater, whereupon the diameter of blastulas decreases. Both of these observations indicate that blastocoele fluid exerts considerable pressure on the wall of the blastula and that the coherence of blastomeres must be considerable to resist this pressure. Cells and intercellular substance provide this structural stability, although the latter is often reinforced by an external elastic hyaline membrane. Permeability of the wall of the blastula changes during blastulation in certain species. At first it is freely permeable to all dissolved substances, but as the cells become smaller with each succeeding cleavage, permeability of the wall is reduced and the fluid in the blastocoele becomes increasingly cut off from the external environment. For example, the wall of the developing blastula of sand dollar eggs is freely permeable to large molecules such as those of sucrose up to the tenth cleavage, but thereafter this permeability is lost.

Wall of the blastula. The wall of the blastula is a mosaic of cellular areas, each of which will normally produce a certain structure during subsequent development. In other words, each area of cells in the wall of the blastula has a certain prospective fate which will be realized in normal development. In blastulas of some species the outlines of these several areas can be determined directly because the cytoplasm of cells within one area may be colored orange-yellow,

that in another area light gray, and that in still another area dark gray. This is the case in the blastula of the tunicate *Styela*, but this is an exceptional situation. In most cases it is necessary to determine the prospective fate of each cellular area indirectly by marking experiments, that is, by staining cells in each area with a vital stain, determining by observation what organ or organs are formed from each stained area, and with the aid of this information, constructing a prospective fate map secondarily on the surface of the blastula. This does not mean that each of these areas would develop independently into the structure in question if isolated from other areas of the blastula. It merely means that these several cellular areas are so related to one another spatially in the blastula stage that they can be shifted subsequently in an orderly way by the morphogenetic movements of gastrulation into appropriate positions which will enable them to cooperate in the formation of tissues and organs through processes of self-differentiation or interaction or both. Thus the events of blastulation appear to be necessary prerequisites for the succeeding period in development, namely, gastrulation. The blastocoele appears to be essential in many species to provide a space into which certain of the special cellular areas of the wall of the blastula can move en masse or individually during gastrulation. *See* FATE MAPS (EMBRYOLOGY).

Ray L. Watterson

Blattaria (Blattodea)

An order included in the Dictyoptera, and comprising 3500–4000 insect species, commonly known as roaches. The fossil record extends back some 400 million years, and roaches were found in abundance in the Carboniferous age. Blattaria are grouped with orders Isoptera (termites) and Mantodea (mantids) within the Dictyoptera. In fact, recent studies support the notion that termites are social roaches; Isoptera appear to be a sister group of roach species in the genus *Cryptocercus* that, like termites, digest wood with the aid of symbiotic intestinal protozoans. *See* INSECTA; ISOPTERA; ORTHOPTERA.

Characteristics. In general, cockroaches range in length from a few millimeters to 10 cm (4 in.), and have a flat oval body, a darkened cuticle, and a thoracic shield, the pronotum, which extends dorsally over the head. A pair of sensory organs, the whip-like antennae, are located on the head, and a much shorter pair of sensory organs, the cerci, are located at the end of the abdomen. Although some species are wingless, most have two pairs of wings but seldom fly. Most species are nocturnal and live in narrow crevices and under stones and debris. A few species actively burrow and others live on plants. The mouthparts are of the biting-chewing type but are relatively weak. Cockroaches are omnivorous, and a few species have adapted to the human environment, becoming household pests since they can chew foodstuffs, clothing, paper, and even plastic insulation. They also can emit a highly disagreeable order.

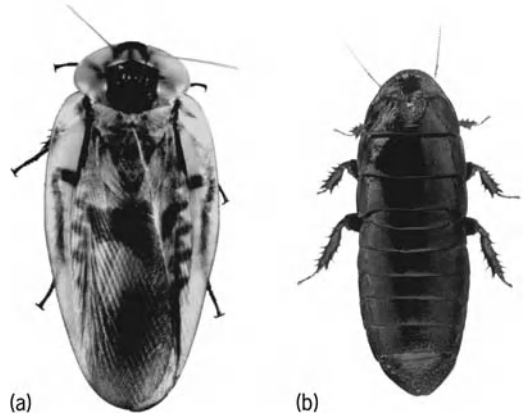

Adult cockroaches. (*a*) *Blaberus giganteus*. (*b*) *Cryptocercus punctulatus*.

Reproduction. Most species are sexual, although a few reproduce asexually. Eggs are typically laid in capsules (oothecae) which can be carried on the end of the abdomen. In some species the eggs hatch just as the ootheca is produced, whereas in others hatching occurs inside the female resulting in live birth. Nymphs go through several molts as they grow. The time from egg to adult varies from species to species, but can be longer than a year.

Pests. Although over 99% of species are not pests, several well-known roaches infest households. *Periplaneta americana*, the American cockroach, is a rust-colored insect 1.2–2.4 in. (30–60 mm) long. *Blatta orientalis*, the Asian cockroach, and *Blattella germanica*, the German cockroach, are smaller, darker-brown roaches which can infest houses in many parts of the world. German roaches have been found in heated buildings just a few hundred kilometers from the North Pole.

Other cockroaches. Other cockroaches of special interest are *Blaberus giganteus* (**illus.** *a*), a South American roach that can grow to 3.2 in. (80 mm); *Cryptocercus punctulatus*, a wood roach which has unicellular organisms in its gut to digest the consumed wood (illus. *b*); and *Gromphadorhina portentosa*, the Madagascar roach, which can compress its body and force air through modified spiracles (breathing pores), producing a loud hissing sound.

Use in research. Cockroaches have been claimed as vectors for numerous human diseases. There is little evidence to support this claim; however, there is evidence that they can produce allergenic reactions in humans. Because of their large size, cockroaches have proved to be very useful animals for scientific research on general problems of insect behavior, physiology, and biochemistry. *See* ALLERGY.

Darryl T. Gwynne; Charles R. Fourtner

Bibliography. D. G. Gordon, *The Compleat Cockroach: A Comprehensive Guide to the Most Despised (and Least Understood) Creatures on Earth*, Ten Speed Press, Berkeley, CA, 1996; I. Huber, E. P. Masler, and B. R. Rao, *Cockroaches as Models for Neurobiology*, CRC Press, Boca Raton, FL, 1990.

Bleaching

The process in which natural coloring matter is removed from a fiber to make it white. The process may be used on fiber, yarn, or fabric. Prior to the bleaching of a fabric, preliminary purification processes should be used. These processes remove applied encrustants (desizing) and natural encrustants (scouring or boil-off) so that the bleaching agent may act uniformly on the material free of impediment.

Bleaching is also classified as a purification process and varies with the content of the substrate or fibrous content of the material. It should not be confused with the stripping process, which is the removal of applied color.

The fabric off the loom is called gray, grey, or greige goods to distinguish it from the partially or completely finished fabric.

Desizing. This process removes sizing agents (slasher compounds) which have been added to the warp yarns (lengthwise yarns) prior to weaving to protect them against the abrasive action of loom parts during the formation of gray fabric. Cellulosics are slashed primarily with starches which may be removed with enzymes of the malt (low-temperature–resistant) or bacteriostatic (high-temperature–resistant) types and are called amylolytic enzymes. Enzymes break down or become inactive on a time basis at room temperature. This inactivation is accelerated by increasing the temperature, which decreases the time. The use of enzymes is thus a time-temperature inverse relationship activity which may accomplish the desizing process at low temperatures for lengthy times or at high temperatures in much shorter periods with equal efficiency.

Other effective agents are dilute mineral acids at high temperatures and sodium bromite in cold solution. Protein warps are sized or slashed or pretreated with proteinaceous agents such as gelatin, casein, albumin, or glue and require the use of proteolytic enzymes for their removal. Synthetics are slashed with water-soluble agents such as poly(vinyl alcohol) or sodium salts of poly(acrylic acids) which are removed by hot detergent solutions. Use of poly(vinyl alcohols) and carboxymethyl celluloses is increasing with cotton and other cellulosic fiber materials because of their biodegradable characteristics and as a means of decreasing stream pollution. They are removed by detergent solutions, as with synthetics. Elimination of the desizing operation creates insurmountable problems in subsequent processes, including bleaching.

Scouring. This process removes natural encrustants which are of a resistant nature from the fibers. In natural fibers such as cotton or linen, very resistant and water-repellent waxes are the main concern. These waxes may vary from $^1/_2$ to 1% by weight and require relatively severe treatment with caustic soda and other auxiliaries at high temperatures to form a wax soap (saponification); thus the waxes are emulsified prior to being rinsed or washed out. Meanwhile pectoses, proteins, and mineral matter in the order of 4–6% by weight are simultaneously removed. With protein fibers such as wool, the repellent is a natural grease. This also must be removed by emulsification with a slightly alkalinized soap or detergent, but with much lower temperature and alkaline content because of the greater susceptibility of wool to alkali damage at high temperatures as well as the greater receptivity of the grease to emulsification. Manufactured fibers, including rayons and synthetics, contain no natural encrustants but have so-called producer oils added to them as protective agents. These are self-emulsifiable oils and are removed concurrently with any sizing agent used in the desizing process, thus eliminating the need for a scour or boil-off. Conversely, if water-soluble sizes are used on manufactured fibers, desizing can be eliminated and the scour or boil-off will accomplish both simultaneously.

The equipment to be used in the process of scouring also depends upon the fiber content in the fabric. Continuous processing (the forerunner of automated processes) is utilized most extensively for cotton with kiers next in productive rate and finally smaller lots by kettles, becks, and jigs. Kettles, becks, jigs, and Hinnekin-type full-width units are used for synthetics. The Hinnekin units are used on filament rayons to prevent cracking or crease formation. Wool and wool blends with synthetics are scoured in rotary (Dolly) washers or kettles.

Continuous ranges for desizing, scouring, and peroxide bleaching of cotton dominate with the material in the rope or full-width form, based upon weight and susceptibility to irremovable wrinkle formation. These ranges consist of saturators, J boxes, and washers in sequence with no variance in feed of fabric and delivery rate of the fabric. Ranges are individual units used in sequence and synchronized in productive rates which allow this continuity of processing. The caustic saturator contains the scouring solution and is composed of caustic soda (emulsifier), sodium silicate (buffer), detergent, and a sequestrant or chelating agent to immobilize metallic salts, especially copper and iron. These two metals may create problems in bleaching and subsequent processing by degradation of the cellulose to oxycellulose. The presence of caustic soda in quantities necessary for saponification indicates that a high pH is present. This pH must be under strict control. The temperature range for this scouring is 206–214°F (96.7–101°C).

The kier is a pressurized, vertical autoclave with capacities as high as 4–5 tons (3.5–4.5 metric tons) of fabric with essentially the same reagents being used in much the same quantity but with cycles of 6–8 h. Pressure is 15 pounds per square inch gage (103 kilopascals), yielding a processing temperature of 240°F (116°C) and with continuous circulation of the solution through the mass and the unit as compared with 1–2-h cycles in the J box in the continuous unit. Care must be exercised in the kier to have uniform packing of the roped fabric to prevent channeling (free solution passage through the mass rather than seepage). Channeling and exposure to air (oxygen which develops oxycellulose) are the major side effects which may create problems in the uniform

emulsification and purification of a fabric free of damage. Kiers, with the exception of the horizontal kier used for heavy goods, are limited to the processing of soft goods which are amenable to roping. Other less productive units for full-width processing of heavy cotton goods are continuous jigs and individual jigs which are not pressurized. Horizontal kiers, continuous jigs, and jigs are used to prevent the heavy-weight goods from producing irremovable wrinkles.

Subsequent developments resulted in the use of solvent scouring with trichlorethylene or perchlorethylene at the boil of the solvent. This is a continuous scour at full width and is claimed to allow 96–97% solvent recovery, which is a necessity to make it economically feasible to compete with other systems.

Scouring of wool is accomplished in rotary or Dolly washers and, therefore, is a batch process having relatively low productive rates. The pH is much lower (8.2–8.4) and the scouring solution consists of soap or deteregent built to the proper pH with soda ash. The temperature is much lower with an average of 120–130°F (49–54°C). The use of caustic soda is prohibited because of the high decomposition rate of wool when exposed to this agent, in particular at elevated temperatures. The use of sequestrants or chelating agents is recommended, especially with soaps, to prevent the formation of insoluble metallic soaps, especially on the fabric. An alternative to this process of scouring is the use of the Derby dry cleaner unit, using solvent as the extractant for the wool grease and oils. Much wool is scoured in the fibrous state, but lubricants are subsequently added and must be removed during the processing of the eventual fabric. The lubricants are self-emulsifiable, and a detergent or soap solution will readily remove them with lesser time being necessary for this scouring or washing.

Scouring or the preferential term boil-off on rayons and synthetics is readily accomplished by using detergents built with soda ash or sodium polyphosphates. The producer oils are readily emulsified, as they are produced to be self-emulsifiable. Boil-off is the term acquired from the degumming of silk, which is accomplished in a soap solution at or near boil temperatures.

Bleaching processes. The three most prominent commercial bleaching processes are the peroxide, the chlorine, and the chlorite, in that order. In home bleaching the predominant bleach is the chlorine bleach, followed by the use of peroxygen compounds such as persulfates and perborates. The latter two are recommended for minimum-care and permanent-press fabrics to preclude the yellowing of the whites and the stripping of colored fabrics, which are always potentials when chlorine bleaches are used.

Peroxide bleaching. Hydrogen perioxide is the most universal of commercial bleaches, despite its higher cost compared with chlorine. This is compensated for by the need of less time and higher productive rates. More than 85% of the commercial bleaching of cotton today is done with prepared concentrates of 27.5, 35, and 50% hydrogen peroxide being properly diluted with water to appropriate low peroxide content for use in the bleaching. The designation of strength used in bleaching is most frequently by volume rather than by percentage peroxide strength, although they are interrelated proportionally. For example, 27.5% hydrogen peroxide is 100 volume, 35% is 130 volume, and 50% is 197 volume. Volume relates to the volume of oxygen released by one volume of the designated hydrogen peroxide. Bleaching of cotton with hydrogen peroxide solution may have an average composition of 34 to 2 volume peroxide, sodium silicate as a buffer to control and maintain pH, sodium polyphosphate as a sequestrant for iron or a chelating agent for copper and iron, plus sodium hydroxide to obtain an optimum pH. There is always the danger of too rapid evolution of nascent oxygen unless the solution is carefully controlled in pH and the potential of developing oxycellulose. The degraded and brittle cellulose shows as tender areas, pinholes, and so on in the bleached cotton. The presence of iron or copper accelerates the oxygen release and oxycellulose formation; thus the need exists for agents to inhibit any action on the part of these metals. The release or control of the release of nascent oxygen is closely associated with the pH according to reaction (1). A chart indicates the optimum con-

$$H_2O_2 \rightarrow H_2O + O \qquad (1)$$

ditions which control the evolution of oxygen and thus deter oxycellulose development.

Peroxide bleaching is controlled by the accurate consideration of the following factors: concentration or volume of peroxide used, pH of the bath, temperature of the bath, and time of exposure. Time and temperature are inverse in relationship.

Production rates are also closely dependent upon the weight and construction of the fabric. Heavy-weight goods processed full width have production rates of 45–50 yd/min (36–45 m/min) continuously, whereas soft and lighter-weight fabrics which may be roped may be processed as high as 250 yd/min (225 m/min). It has become feasible with pressurized and sealed equipment to process even heavy-weight fabrics at 100 yd/min (90 m/min) continuously at 250–300°F (121–149°C).

Inasmuch as the release of nascent oxygen is controlled in an alkaline bath, it is evident that for shipping and storage an acid condition must prevail. Further precaution against breakdown of hydrogen peroxide, especially under nonacidic conditions, is exemplified by the use of brown bottles for 3% or 10-volume medicinal peroxide. Commercial shipments of concentrates are in tank cars or in large drums which are vented to release any pressure buildup within the container.

Wool. The percentage of wool bleached is much less than the percentage of cotton bleached. The wool which is bleached is processed through a peroxide system. As a result of the wool's nature and its susceptibility to alkali, especially in hot solutions, both the pH and the temperature are lower

than for cellulosics with a compensating increase in time required to complete the process. This time may be shortened, however, by incorporating optical bleaches which are compatible with the peroxide. By substituting ammonia for the soda ash or tetrasodium pyrophosphate, the pH may be raised to 10 rather than the normal pH of 8–9 with a potential decrease in the peroxide time cycle. Present methods utilize batch rather than continuous systems. Precautions to be taken consist of acidification with an organic acid to a pH of 6.5–7.0 prior to the immersion of the fabric in the peroxide solution and again subsequent to the postrinsing following the bleach to reduce the alkali retention by the wool.

Wool cannot be bleached with hypochlorite since it contains amino components which cause a permanent yellowing of the fiber due to the formation of yellow chloramines. Silk also is bleached by peroxide but at a somewhat higher temperature and with a compensating shorter cycle for the bleaching. Wool and silk, being protein fibers, were initially bleached by the stoving process. This was a reduction bleach in which the uniformly wet fabric was exposed to sulfur dioxide gas in an enclosed chamber for an extensive period of time. Sulfur on burning released SO_2, which caused color reduction and bleaching. The process has been abandoned because the white obtained soon reverted to the prebleached color.

Synthetics such as polyamides, polyesters, and acrylics do not lend themselves to peroxide bleaching because of the alkalinity and the potential damage to the fibers. The chlorite bleach is used for these fibers. Manufactured fibers made from cellulose, such as viscose and Bemberg, are primarily white on production but, if necessary, peroxide bleach may be used for very short cycles, or a cold saturation followed by dwelling on the roll overnight when required. The exception is acetate rayon, whose lack of resistance is such that is cannot withstand alkalinity. Acetate, therefore, is better bleached with another peroxygen compound, peracetic acid, in an acid medium.

Hypochlorite bleaching. This bleaching process is most commonly called the chlorine bleach. This was the foremost bleaching process for cellulosics until after World War II, when the continuous peroxide system, which was restrained during the war, gained favor rapidly. Hypochlorite, however, is still the predominant agent used in household or home bleaching. The sodium rather than the calcium hypochlorite (bleaching powder) is dominant. Sodium hypochlorite may be readily made by the electrolysis of sodium chloride, but in actual plant manufacture it emanates from the bubbling of chlorine gas from pressurized tanks into dilute caustic soda or soda ash. During the process of bleaching with sodium hypochlorite in water solution, hydrolysis occurs with the production of hypochlorous acid. This is represented by reaction (2). This acid is very unstable and breaks

$$NaOCl + HOH \rightarrow HClO + NaOH \qquad (2)$$

down, releasing nascent oxygen as shown by reaction (3). Excess NaOH and the pH condition con-

$$HClO \rightarrow HCl + O \qquad (3)$$

trol the nascent oxygen release and thus the ultimate whitening. Unless the controls are exercised, oxycellulose may develop as in the peroxide system. Care must also be used commercially to antichlor with acid and sulfur dioxide or sodium bisulfite to preclude later damage to the fiber caused by the retention of chlorine which will hydrolyze the cellulose. It is also good household practice to use the sodium bisulfite in an automatic washer rinse, but this care is seldom exercised. Developments have resulted in a proposal to use a continuous hot chlorine bleach commercially. This requires very strict control of conditions such as varying the concentration, time, pH, and temperature.

Chlorite bleach. The use of sodium chlorite (Textone) as a bleaching agent first emerged about 1941, subsequent to the introduction of the continuous peroxide bleach in 1939. Unlike the other bleaches, the effective bleaching agent is ClO_2, chlorine dioxide. The use of $NaClO_2$ is quite distinctive also because it is accomplished as an acid system at an optimum pH of 3.5. The temperature recommended for the bleaching is 185°F (85°C). A significant characteristic of this bleaching process is the freedom from the degradation of cellulose and, consequently, no oxycellulose is produced as is true in the peroxide and hypochlorite systems. There is objection to its use commercially because its crystals are subject to explosive tendencies under suitable conditions and the ClO_2 gas is itself an obnoxious and toxic gas. Another factor is that even stainless steel is pitted when the acid bleaching medium is present. Despite these characteristics, it is the most effective bleaching agent for synthetics and is used on cotton in a few continuous systems. In a continuous system, a pitting inhibitor, a special J box with a seal to prevent ClO_2 escape, and a siphon to remove excess $NaClO_2$ and ClO_2 are components. Some consideration has been given to the use of fiberglass and ceramic equipment with this bleach. The ClO_2 release may be represented by reaction (4).

$$5NaClO_2 + 4HCl \rightarrow 4ClO_2 + 5NaCl + 2H_2O \qquad (4)$$

Optical bleaching. There are organic compounds which are capable of absorbing waves shorter than visual waves and emitting waves within the visible range. The most notable of the short waves absorbed is the ultraviolet light present in daylight. These compounds absorb ultraviolet light which is less than 400 nanometers and emit blue light which is in the visible range of 400–700 nanometers. This emitted light will counteract the yellow on the fabric surface and by the subtractive theory of color will produce a white. The agents are also characterized as brighteners and are extensively used in household detergents. They are, however, effective only in daylight because of the ultraviolet light therein. They are also subject in varying degrees to loss of effect upon exposure to sunlight as a result of fading. Materials

Characteristics of textile bleaching

Bleach	Fibers	pH	Temperature, °F (°C)	Time
Hypochlorite	Cellulosics	9.5–10.0	70–80 (21–27)	4–12 h
Hypochlorite	Cellulosics	10–11	180 (82)	5–10 s
Peroxide	Cellulosics	10.4–11.5	180 (82)	2–4 h
Peroxide	Cellulosics	10.4–11.5	208 or greater(98 or greater)	1–1$^1/_2$ min
Peroxide	Wool	9.0–10.0	120 (49)	2–6 h
Peroxide	Silk	9.0–10.0	160 (71)	$^1/_2$ h
Chlorite	Cellulosics and synthetics	3.5	185 (85)	$^3/_4$–1 h
Optical	Cellulosics and synthetics	5.0 or less	180 (82)	$^1/_4$– 1 h
Optical	Wool	8.0–8.5	160 (71)	$^1/_4$–$^1/_2$ h
$NaOCl/NaClO_2$	Cellulosics	8.7–9.0	70–80 (21–27)	1 h
$NaOCl/H_2O_2$	Cellulosics	9.5/10.5	70–80/200 or greater (21–27/93 or greater)	$^1/_3$–$^1/_2$ h

treated with these optical bleaches or agents fluoresce brilliantly in an ultraviolet light source. They are often used to supplement fabrics bleached by other methods, especially the hypochlorite and peroxide systems where selectively they can be incorporated in the normal bleaching solution. This may result in the cost of the bleaching being reduced because less bleach is used or the cycle is shortened. Furthermore, losses in strength may be minimized because less oxygen bleach is used and the opticals do not affect the strength of fibers.

Combination bleaching. There are several bleaching systems which are classified as combination bleaches because two different bleaching agents are used together or in tandem. These have been projected to lower the cost, lower the degradation, or shorten the cycle as well as to preclude equipment damage in certain cases.

Solvay activated peroxide bleach. This consists of first exposing cotton to cold hypochlorite. After this fewer chemicals and less time and heat exposure are necessary than is true with peroxide alone.

Hypochlorite-chlorite bleach. This process shortens the hypochlorite bleach time, releases ClO_2 rather than O, and prevents equipment damage since the process takes place under alkaline conditions. No heat is necessary as is true for the chlorite bleach. The **table** summarizes conditions for this and all the other bleaching processes discussed. *See* NATURAL FIBER; OXIDIZING AGENT; TEXTILE CHEMISTRY.

John J. McDonald

Bibliography. E. R. Trotman, *Dyeing and Chemical Technology of Textile Fibers*, 6th ed., 1985; T. L. Vigo, *Textile Processing and Properties*, 1994.

Blimp

A name originally applied to nonrigid airships and usually implying small size. Early blimps contained about 100,000 ft^3 (2800 m^3) and were only a fraction of the size of the rigid airships of that period. The nonrigid pressure-type airship, however, did not stay small; each succeeding model built was substantially larger than its predecessor until ships with volumes of approximately 1,500,000 ft^3 (42,000 m^3) were operational in the U.S. Navy. With the advent of these larger sizes it appeared that the name blimp would be replaced by the more general term airship. However, the principles of construction of these ships are basically the same regardless of the size.

Main envelope. The airship envelope is a symmetrical airfoil in the form of a body of revolution, circular in cross section and elliptical along its longitudinal axis. The fabric of the main envelope or pressure hull is usually made up of two or three plies of cloth impregnated with an elastomer; at least one of these plies is placed in a bias direction with respect to the others. This tends to make the fabric more resistant to shear loads and results in a stabilized structure. The materials used for the envelope must be lightweight, extremely strong, and resistant to gas diffusion. Chronologically the fibers and elastomers used were (1) cotton and rubber, (2) cotton and neoprene, (3) Fortisan rayon and neoprene, (4) nylon and neoprene, (5) Dacron and neoprene, (6) polyester, urethane, and Tedlar laminates. Each newly developed material improved some property required of envelope structures; and these improved fabrics, along with advancements in assembly techniques, permitted the increase in size of the pressure-type airship to fulfill any possible need. *See* MANUFACTURED FIBER; POLYMER; RUBBER.

A pattern is made for each panel and gore so that, when the fabric pieces are cemented, sewed, or heat-sealed together, the envelope will have the required shape when filled with gas. The outside of the envelope is sometimes coated with aluminized paint for protection against the sunlight. In later versions an ultraviolet shield is incorporated in the top surface of the fabric.

Inside the envelope are catenary curtains (**Fig. 1**). They support the weight of the car by distributing the loads imposed by it into the envelope fabric. This suspension system can be arranged in a number of ways, depending upon the particular airship configuration (**Fig. 2**). Basically, however, they all consist of cable systems attached to the car which terminate in fabric curtains, which in turn are cemented and

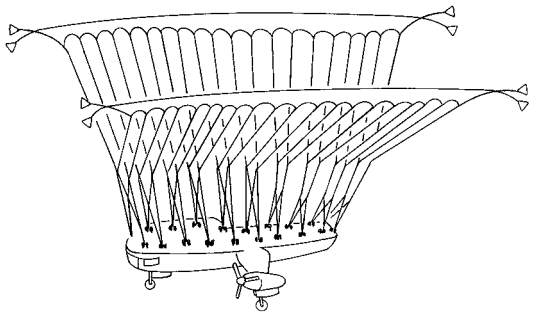

Fig. 1. Typical internal rigging of nonrigid airships.

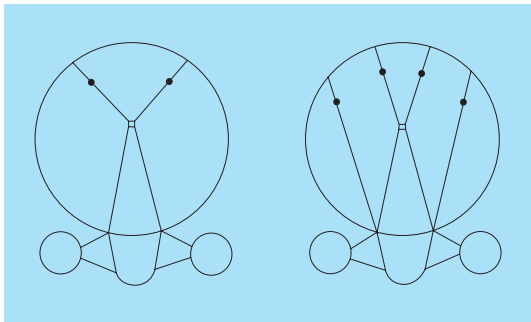

Fig. 2. Two types of internal suspension for nonrigid airship cars.

sewed or otherwise sealed to the envelope proper. The envelope also contains one or more air cells, fastened to the bottom or sides of the envelope, which are used to maintain the required pressure in the envelope without adding or valving of gas as the ship ascends or descends. These air cells are called ballonets and are usually made of a fabric much lighter in weight than that of the envelope, because they must merely retain gas tightness and do not have to withstand the normal envelope pressure.

The fabric envelopes, when not in use, can be folded or shipped or stored in a space which takes up less than 1% of the inflated volume of the envelope. This feature and the ability to erect the envelope by merely inflating it are two of the advantages of the fabric pressure-type airship over the rigid one.

Other components. The airship car is of conventional aircraft construction, earlier of a fabric-covered tubing framework but now of a metal or plastic monocoque design. Much like an airplane, it contains the crew compartments, the engines, the landing gear, and all the related instruments, controls, and fuel. A few additional accessories required for the airship car are the pressure controls, air scoops behind the propellers, valves, and the fan system necessary for maintaining the pressure in the envelope by pumping air in and out of the ballonets.

The airship tail surfaces have taken on three typical configurations, named from the placement of the surfaces on the envelope. They are the cruciform (+), the X, and the inverted Y. These tails are made up of a fixed main surface and a controllable smaller surface on the aft end. The construction consists primarily of lightweight metal structural beams covered with doped fabric. These are held to the envelope by cables which distribute the load into fabric patches cemented to the envelope proper. These surfaces are so efficiently designed that they weigh less than 0.9 lb/ ft^2 (4.4 kg/m^2) of tail area.

The nose cone, made up of metal, wood, or plastic battens, is laced to the envelope to serve the dual purpose of providing a good point of attachment for mast mooring and of adding extra rigidity to the nose, because it encounters the greatest dynamic pressure loads in flight. Sometimes the nose battens are covered to provide streamlining.

Erection for flight. The erection of a fabric pressure-type airship is a spectacular sight and takes place in a relatively short time. The envelope is spread out on the floor of the airship hangar and a net is placed over it. As gas is fed into the envelope from tank cars, each containing about 200,000 ft^3 (5700 m^3) of 99.9% pure helium compressed to 2100 lb/in.^{2} (14.5 megapascals), the net which is held down by sand bags is permitted to rise slowly, care being taken that the envelope does not slip out from under the net. Fins, nose cone, battens, air valves, and helium valves are attached, and the envelope is then allowed to rise high enough from the floor to permit rolling the car underneath. After attachment of the car, the net is removed and the airship is rigged for flight.

The strength required of the material used in the fabric pressure-type airship envelope is a function of the internal gas pressure, the distribution of the gas in the envelope, the location of the weights attached to it, and the aerodynamic and ground-handling loads imposed on the airship in flight and at the mast. One of the most important properties of the fabric pressure-type airship is its ability to accept an overload condition, which in a nonpressurized airship would result in a structural failure. In other words, if the airship envelope should encounter, for example, a gust of greater velocity than it had been designed to withstand, there would be a momentary buckle or wrinkle in the envelope with a complete structural recovery the moment the force was removed. This "forgiveness" is a special feature of the pressure-type structure utilizing a flexible material for the load-carrying members.

Applications. Up until the late 1980s, the only commercial airships operating in the United States were the Goodyear Tire and Rubber Company advertising blimps of about 150,000 ft^3 (4200 m^3). Fundamentally, they were communications platforms used for public service where airship characteristics are superior to other forms of flight.

Because of its ability to practically hover over one spot, the airship forms an excellent airborne platform for a TV camera and operator viewing major sports events such as football games, golf tournaments, or automobile or boat races. In times of emergency, it is also used to inspect the consequences of natural disasters such as floods or forest fires. Lightweight TV cameras with zoom lenses can

Typical operating blimps

Characteristic	Goodyear GZ-20	Goodyear GZ-22	Airship Industries Skyship 500	Airship Industries Skyship 600
Length, ft (m)	192 (58.5)	205.5 (62.6)	170.6 (52.0)	193.6 (59.0)
Width, ft (m)	46 (14.0)	47 (14.3)	45.93 (14.0)	49.86 (15.2)
Volume, ft^3 (m^3)	202,700 (5740)	247,800 (7017)	180,355 (5107)	233,100 (6600)
Height, ft (m)	57.5 (17.5)	60.2 (18.3)	61.2 (18.7)	66.6 (20.3)
Maximum speed, mi/h (km/h)	50 (80)	65 (105)	60 (97)	65 (105)
Maximum car length, ft (m)	22.75 (6.93)	34.75 (10.59)	29.44 (8.97)	37.44 (11.41)

readily pick up an overall color picture of a live happening or pinpoint a special detail.

For television operations the airship carries a pilot, a full-sized transistorized color TV camera, and a camera operator. The airship also carries monitoring equipment, encoding and microwave transmitting equipment, and an antenna. Usually the TV camera and its related equipment are operated by network personnel. Signals from the camera are monitored and corrected for picture quality and color balance before being fed to the conditioning equipment for transmission to the microwave antenna and for radiation from the airship antenna to a ground parabolic tracking antenna. These signals are then fed into a network connection or are transmitted to a station by a microwave radio system. *See* TELEVISION CAMERA; TELEVISION NETWORKS.

Another form of communication for which the airship is used is the night sign. The sign consists of a matte of multicolor incandescent lamps permanently fixed to the sides of the airship envelope. By programming the operating of these lamps, stationary or moving letters or pictures can be displayed against a night background.

Originally, these sign messages were developed by electromechanical relays. Now magnetic tapes are prepared on the ground by composing equipment. The tapes are then carried along on the flight and fed into an airborne reader, which plays the taped information back through a computer to the lamp driver circuits. The resulting message, which usually supports a civic endeavor such as a bond drive or water-saving campaign, is then readily seen by people from a substantial area of the ground. Because the airship can move quite slowly, extended news messages can be transmitted this way.

During the 1980s, several companies other than Goodyear began to produce blimps for advertising purposes, passenger operations, and even possible military missions. Typical of these was Airship Industries of England, which produces the Skyships 500 and 600, used in Europe, the United States, Australia, and Japan, primarily as advertising billboards. In 1988, Goodyear introduced the GZ-22, the largest and most advanced blimp flying at the time, featuring a unique X-configuration tail and shrouded, swivable turboprop power plants (Fig. 5). The relative characteristics of four widely used blimps are shown in the **table**.

Because of their endurance, slow flight capability, and stability, these vehicles were also tested for possible use in patrol of national borders, control of shipping, pollution monitoring, fishery protection, and search-and-rescue operations. A boarding boat can be lowered from a particular blimp model for possible use against drug smuggling. Applications to mine countermeasures as well as airborne early warning were also studied by the military services in order to define possible use with the U.S. Navy of a much larger version with exceptional range, endurance, and payload capabilities. *See* AIRSHIP. Robert S. Ross

Bibliography. W. E. Althoff, *Sky Ships: A History of the Airship in the United States Navy*, 1990, reissue 1998; B. Bailey and N. J. Mayer, *Military and Civil Application of Airships*, 1986; L. Gerkin, *Airships: History and Technology*, 1990; R. A. Grossnick (ed.), *Kite Balloons to Airships ... The Navy's Lighter-Than-Air Experience*, 1988; L. Payne, *Lighter Than Air: An Illustrated History of the Airship*, 1977, reprint 1991.

Bloch theorem

A theorem that specifies the form of the wave functions that characterize electron energy levels in a periodic crystal. Electrons that move in a constant potential, that is, a potential independent of the position $\mathbf{r}$, have wave functions that are plane waves, having the form $\exp(i\mathbf{k} \cdot \mathbf{r})$. Here, $\mathbf{k}$ is the wave vector, which can assume any value, and describes an electron having momentum $\hbar\mathbf{k}$. (The quantity $\hbar$ is Planck's constant divided by 2π.)

Electrons in a crystal experience a potential that has the periodicity of the crystal lattice. Lattice translations $\{\mathbf{L}\}$, are the vectors (in space) that describe the relative locations of any two primitive cells (for example, molecular units). The electron's potential, $V(\mathbf{r})$, is periodic, meaning that it is unchanged by all translations of the form $\mathbf{r} \rightarrow \mathbf{r} + \mathbf{L}$.

The Bloch theorem states that wave functions that characterize electron energy levels in a periodic crystal may always be written as in the equation below,

$$\psi_n(\mathbf{k}, \mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_n(\mathbf{k}, \mathbf{r})$$

where n is an energy-band index. Here, $u_n(\mathbf{k},\mathbf{r})$ is a function that has the same periodicity as the potential $V(\mathbf{r})$. The wave vector $\mathbf{k}$ can no longer assume any value. Instead, it is restricted to a polyhedral volume

called the Brillouin zone. For example, the Brillouin zone of a cubic crystal having lattice constant a is a cube (in **k** space) having edge length $2\pi/a$. The loss in allowed **k** space is compensated by the band index n, which can have integer values from 0 to ∞. *See* BRILLOUIN ZONE.

Electrons described by a Bloch function, as above, do not have momentum $\hbar\mathbf{k}$, but their mean momentum is the same in every cell of the crystal. Thus, the Bloch theorem resolved the paradox of how an electron can travel indefinitely in a perfect crystal without being scattered randomly by the atomic potentials. The practical utility of the Bloch theorem is that **k** and n are the indices needed to enumerate electron states when equilibrium or nonequilibrium phenomena are described. *See* BAND THEORY OF SOLIDS. Albert Overhauser

Block and tackle

Combination of a rope or other flexible material and independently rotating frictionless pulleys; the pulleys are grooved or flat wheels used to change the direction of motion or application of force of the flexible member (rope or chain) that runs on the pulleys. As in a lever, the summation of torques about the axis of rotation of the pulley equals zero for static equilibrium. Tension T in the rope is the same on both sides of the pulley. *See* LEVER.

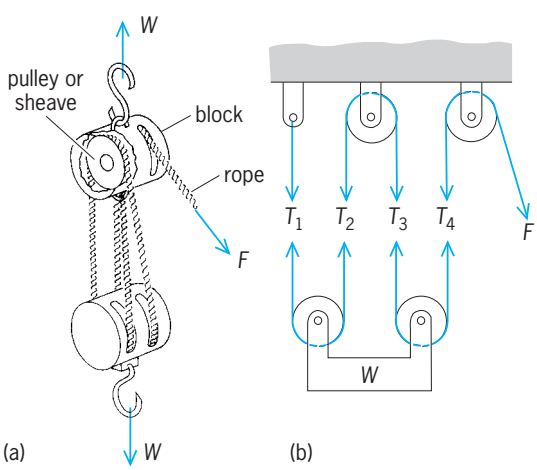

Block and tackle. (*a*) Actual view. (*b*) Schematic.

For example, in the particular block and tackle in the **illustration**, at static equilibrium, the summation of forces in any direction equals zero. Each vertical rope carries one-fourth the weight; algebraically $T_1 = T_2 = T_3 = T_4 = W/4$, and therefore the applied force F also equals one-fourth the weight W. For this case mechanical advantage MA is shown by the equation below. The block and tackle is used

$$MA = \frac{W}{F} = \frac{W}{W/4} = 4$$

where a large multiplication of the applied forces is desirable. Examples are lifting weights, sliding heavy machinery into position, and tightening fences. *See* SIMPLE MACHINE. Richard M. Phelan

Block diagram

A convenient graphical representation of input-output behavior of a system or subsystem, where the signal into the block represents the input and the signal out of the block represents the output. The flow of information (the signal) is unidirectional from the input to the output. A complex system, comprising several subsystems, is represented by the interconnection of blocks that represent the subsystems. The primary use of the block diagram is to portray the interrelationship of distinct parts of the system.

A block diagram consists of two basic functional units that represent system operations. The individual block symbols portray the dynamic relations between the input and output signals. The relationship between input and output signals of a subsystem is typically represented by what is known as the transfer function of the subsystem. In management science, criminal justice systems, and economic systems, a block diagram may represent a system function describing a set of activities. The second type of unit, called a summing point, is represented by a circle with arrows feeding into it. The operation that results is a linear combination of incoming signals to generate the output signal. The sign appearing alongside each input to the summing point indicates the sign of that signal as it appears in the output.

Block diagrams are widely used in all fields of engineering, management science, criminal justice, economics, and the physical sciences for the modeling and analysis of systems. In modeling a system, some parameters are first defined and equations governing system behavior are obtained. A block diagram is constructed, and the transfer function for the whole system is determined.

If a system has two or more input variables and two or more output variables, simultaneous equations for the output variables can be written. In general, when the number of inputs and outputs is large, the simultaneous equations are written in matrix form. *See* MATRIX THEORY; MULTIVARIABLE CONTROL.

Block diagrams can be used to portray nonlinear as well as linear systems, such as a cascade containing a nonlinear amplifier and a motor (see **illustration**).

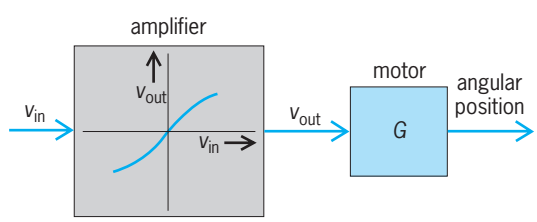

Block diagram of a motor and an amplifier with a nonlinear gain characteristic. v_{in} = input time variable, v_{out} = output time variable, G = transfer function.

If the motor is assumed to be linear, it can be represented by a transfer function. The nonlinear amplifier gain can describe only the relationship between the input time variable and the output time variable, and no transfer function exists between them. An overall transfer function relation cannot be defined for this cascade system unless the amplifier is represented by a linear gain model. *See* COMPUTER PROGRAMMING; CONTROL SYSTEMS; GAIN; NONLINEAR CONTROL THEORY; SYSTEMS ENGINEERING. G. V. S. Raju

Bibliography. J. J. D'Azzo, C. H. Houpis, and S. N. Sheldon, *Linear Control System Analysis and Design with MATLAB*, 5th ed., 2003; J. DiStefano, A. Stubberud, and I. Williams, *Feedback and Control Systems*, 2d ed., 1990; E. Kamen, *Introduction to Signals and Systems*, 2d ed., 1989; M. Sain, Miniaturization of large criminal justice systems, *J. Interdis. Model. Simul.*, 1(2):97–122, 1978; S. M. Shinners, *Modern Control System Theory and Design*, 2d ed., 1998.

Blood

The fluid that circulates in the blood vessels of the body. Blood consists of plasma and cells floating within it. The chief components of plasma are proteins (albumin and globulins), anions (mainly chloride and bicarbonate), and cations (mainly sodium, with smaller concentrations of potassium, calcium, and magnesium). The cells are derived from extravascular sites and then enter the circulatory system. They frequently leave the blood vessels to enter the extravascular spaces, where some of them may be transformed into connective tissue cells. The fluid part of the blood is in equilibrium with the tissue fluids of the body. The circulating blood carries nutrients and oxygen to the body cells, and is thus an important means of maintaining the homeostasis of the body. It carries hormones from their sites of origin throughout the body, and is thus the transmitter of the chemical integrators of the body. Blood plasma also circulates immune bodies and contains several of the components essential for the formation of blood clots. Finally, blood transports waste products to excretory organs for elimination from the body. Because of its basic composition (cells surrounded by a matrix), development, and ability to modify into other forms of connective tissues, blood can be regarded as a special form of connective tissue. *See* CONNECTIVE TISSUE.

Formed Elements

The cells of the blood include the red blood cells and the white blood cells. In all vertebrates, except nearly all mammals, the red blood cells or corpuscles contain a nucleus and cytoplasm rich in hemoglobin. In nearly all mammals the nucleus has been extruded during the developmental stages.

Erythrocytes. In normal adult men the blood contains about 5,000,000 red blood corpuscles or erythrocytes per cubic millimeter; in normal adult women, about 4,500,000. The erythrocytes in humans are about 8 micrometers in diameter and about 2 micrometers at their thickest and have a biconcave shape. They contain hemoglobin, which imparts to them their color, and possess an envelope, which when viewed with the electron microscope appears highly structured. The hemoglobin is involved in the transport of oxygen and carbon dioxide and plays a role in maintaining a constant pH in the blood. When circulating in the blood vessels, the red blood cells are not evenly dispersed. In the narrowest vessels, the capillaries, the erythrocytes are often distorted. In certain conditions they may be densely aggregated. This is known as a sludge. The erythrocytes respond to changes in osmotic pressure of the surrounding fluid by swelling in hypotonic fluids and by shrinking irregularly in hypertonic fluids. Shrunken red blood cells are referred to as crenated cells. The average life of the mature red blood cell is surprisingly long, spanning about 120 days. *See* HEMATOLOGIC DISORDERS; HEMOGLOBIN.

Leukocytes. In humans the white blood cells in the blood are fewer in number. There are about 5000–9000/mm^3. In general, there are two varieties, agranular and granular. The agranular cells include the small, medium, and large lymphocytes and the monocyte (**Fig. 1**). The small lymphocytes are spherical, about the diameter of erthyrocytes or a little larger, and constitute about 20–25% of the white blood cells. The medium and large lymphocytes are relatively scarce. In all lymphocytes the nucleus occupies nearly the whole volume of the cell, and the cytoplasm which surrounds it forms a thin shell (**Fig. 2***a*). The cytoplasm stains deeply with basic dyes as a result of its high content of ribonucleic acid, which exists in soluble form and in larger aggregates with protein which are known as ribosomes. Ribosomes are about 18 nanometers in diameter, and are grouped in functional clusters called polysomes, which play an important role in protein synthesis. The typical monocyte is commonly as large as a large lymphocyte (12 μm), and constitutes 3–8% of the white blood cells. The nucleus is relatively small, eccentric, and oval or kidney-shaped. The cytoplasm is relatively larger in volume than that in lymphocytes and does not stain as deeply with basic dyes. *See* RIBOSOMES.

The granular white blood cells or granular leukocytes are of three varieties: neutrophil, eosinophil, and basophil. Their structure varies somewhat in different species, and the following applies to those of humans.

The neutrophils make up 65–75% of the leukocytes. They are about as large as monocytes with a highly variable nucleus, consisting of three to five lobes joined together by threads of chromatin (Fig. 2*b*). The cytoplasm contains numerous minute granules which stain with neutral dyes and eosin. In certain conditions the neutrophils leave the blood vessels and wander into the connective tissue ground substance, where they may then disintegrate, releasing their granules. These granules are rich in certain hydrolytic enzymes, which become active and take part in some phases of the defense mechanisms of

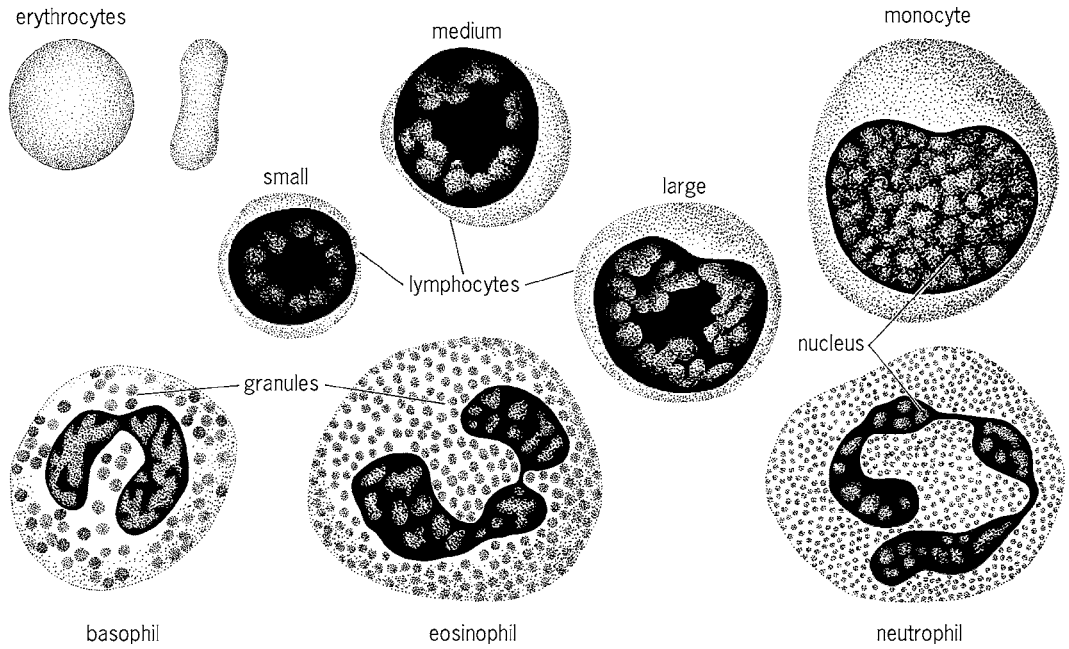

Fig. 1. Diagrams of human blood cells.

the body. The granules are considered to be pure lysosomes which are released from the cell during inflammation and in other conditions, thus liberating many enzymes which are important in combating infections, in cleaning up cell debris, and so on. *See* LYSOSOME.

The eosinophils (also called acidophils) are about the same size as the neutrophils but are less numerous, constituting about 1% of the leukocytes. The nucleus commonly contains but two lobes joined by a thin thread of chromatin. The granules which fill the cytoplasm are larger than those of the neutrophils and stain with acid dyes.

The basophils are about the same size as the other granular leukocytes. The nucleus may appear elongated or with one or more constrictions. The granules are moderately large, stain with basic dyes, and are water-soluble.

The functions of the leukocytes while they are circulating in the blood are not known. However, when they leave the blood vessels and enter the connective tissue, they constitute an important part of the defense mechanism and of the repair mechanism. Many of the cells are actively phagocytic and engulf debris and bacteria. Lymphocytes are of two major kinds, T cells and B cells. They are involved in the formation of antibodies and in cellular immunity. Lymphocytes develop into plasma cells, which form antibodies even more effectively than lymphocytes. The lymphocytes and plasma cells may act in conjunction with macrophages in promoting phagocytosis. In addition, lymphocytes and monocytes may develop extravascularly into macrophages (which are also phagocytic) and subsequently into fibroblasts. The fibroblasts are important in the formation of new connective tissue in

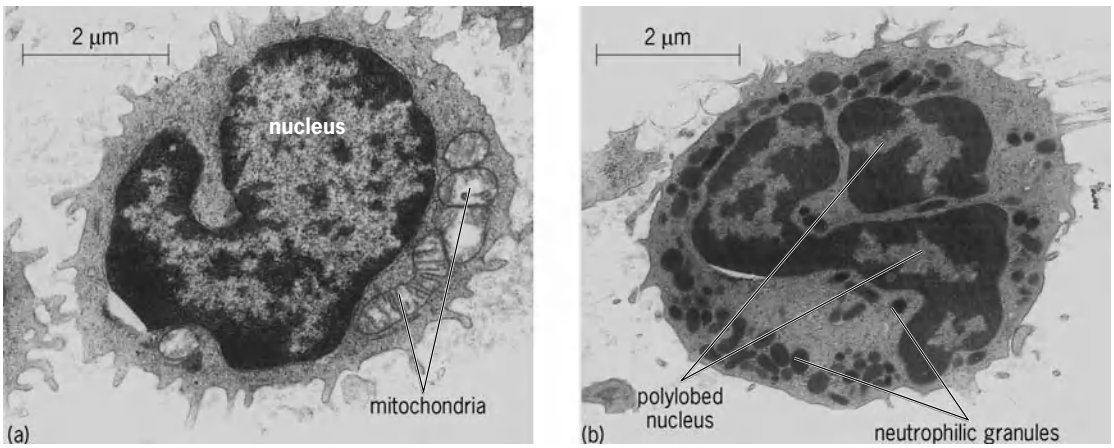

Fig. 2. Low-power electron micrograph of (*a*) lymphocyte and (*b*) neutrophil from the blood of a calf.

Range of variation in numbers and relative proportions of blood cells for some vertebrates

Vertebrate	Red blood cells, $\times 10^6$	Blood platelets, $\times 10^4$	Total leukocytes, $\times 10^3$	Relative proportions of leukocytes, %*				
				Lymphocytes	Monocytes	Neutrophils	Eosinophils	Basophils
Frog	0.2–0.7	0.34–0.39	2.4–25	24–69	2–20	17–72	6–26	1–37
Chicken	3.0–3.8 (3.3)†	1.7–5.0 (3.1)	9.3–32.3 (20)	37–90	0–1	13–49	2–14	1–7
Cat	7–10 (8.5)	15–36.5 (25.8)	8.6–27 (17.8)	10–64	1–3	31–85	1–10	0–2
Dog	5.6–6.8 (6.1)	16.5–37.8	7–11.4 (10)	9–50	1–6	42–77	0–14	0–1
White rat	5.5–10 (8)	43–48	5–25.6 (12.5)	62–75	1–6	18–36	1–4	0
Rabbit	4–6.4 (5.3)	12.6–100 (22)	5.2–12 (8)	20–90	1–4	8–50	1–3	0.5–30
Cow	5–10 (7)	10–80 (50)	4–12	45–75	2–7	15–47	2–20	0–2
Sheep	8–12 (10.8)	26.3–59.8 (40)	6.4–9.8 (8)	38–72	1–3	21–49	0–20	0.5–2
Pig	5.8–8 (6.5)	32.5–71.5 (52)	11–22 (16)	39–62	2–10	28–51	0.5–11	0.2
Horse	6.5–12.5 (9.5)	10–60 (33)	5.5–12.5 (9)	15–70	0.5–10	30–77	0–12	0–3
Elephant	2–4 (2.8)		6.4–14 (10.2)	40–60	0–5	22.50	6–15	0–1
Camel	3.8–12.6 (8.2)		12.9–27.2 (11.7)	27–65	0.1	21–56	0–19	0–1
Lion	6.9–10.9 (9.3)		8.2–19.8 (14.2)	7–37	0–2	54–97	0–6	0
Humans	4.6–6.2 (5.4)	14–44 (25)	5–9 (7)	20–25	3–8	65–75	0–1	0–0.5

*Relative proportion of leukocytes given as a range of total leukocytes per cubic millimeter of blood.
†Number of cells (average in parentheses) given as a range per cubic millimeter of blood.

regions of injury. *See* CELLULAR IMMUNOLOGY; PHAGOCYTOSIS.

Platelets. The blood platelets are small spindle-shaped or rodlike bodies about 3 μm long and occur in large numbers in circulating blood. In suitably stained specimens they consist of a granular central portion (chromomere) embedded in a homogeneous matrix (hyalomere). They change their shape rapidly on contact with injured vessels or foreign surfaces and take part in clot formation. During this process numerous fibrils of fibrin radiate from the platelets. The platelets are not to be regarded as cells and are thought to be cytoplasmic bits broken off from their cells of origin in bone marrow, the megakaryocytes.

Hematopoiesis. Some of the blood cells are terminal cells in the sense that they are fully differentiated and have little or no capacity for division. These are the erythrocytes and the granulocytes. The agranular leukocytes may develop extravascularly into other cell types (macrophages, fibroblasts, or plasma cells). All blood cells are formed in the hematopoietic organs (chiefly lymph nodes and spleen for agranular leukocytes, and red bone marrow for erythrocytes and granular lukocytes in the adult). In these organs the blood cells develop from precursor cells which themselves are thought by many to be derived eventually from the least differentiated cells of this strain, the hemocytoblast or primitive mesenchymal cell. *See* HEMATOPOIESIS.

Comparative aspects. Though blood cells occur in all vertebrates, they differ in density, in relative proportions, in morphology to some extent, and in their sites or origin (see **table**).

Although all cell types occur in all vertebrates, there are some morphological variations among the cells. The red blood cells of fish are nucleated, are ovoid or spindle-shaped, and may be larger than in any other vertebrate. They are also large in frogs, where they are somewhat less oval than in birds, and nucleated in both. Notable differences in internal structure of leukocytes are most marked in the size, shape, and staining properties of the granular leukocytes. In the chicken the granules of the neutrophils stain like those of eosinophils and appear rodlike. In the cat the neutrophilic granules are minute, and the eosinophilic granules are rodlike. In the rabbit neutrophilic granules are rather large spheres and stain eosinophilic, while the granules of eosinophils are elongated.

In adult mammals cells of the circulating blood normally arise from bone marrow (red blood cells, granular leukocytes, and platelets), lymph nodules of spleen, lymph nodes, and other lymphoid organs (lymphocytes). But in lower animals (fishes) the connective tissue of the kidney may be important in the origin of all types of blood cells. In other fish the connective tissue of the genitalia may be the site of origin of granulocytes, and the connective tissue of the intestine may be the site of origin of lymphocytes. The liver of some amphibians may be important for the origin of blood cells. Isidore Gersh

Plasma

Plasma is the residual fluid of blood left after removal of the cellular elements. Serum is the fluid which is obtained after blood has been allowed to clot and the clot has been removed. Serum and plasma differ only in their content of fibrinogen and several minor components which are in large part removed in the clotting process. *See* SERUM.

Plasma proteins. The major constituents of plasma and serum are proteins. The total protein concentration of human serum is approximately 7 g/ml, and most other mammals show similar levels. Birds, reptiles, amphibia, and most fishes have lower protein concentrations of approximately 3–5 g/100 ml. By various methods it can be demonstrated that serum protein is a heterogeneous mixture of a large number of constituents. Only a few are present in higher concentrations, the majority being present in trace amounts. The principal methods for demonstration of heterogeneity and for separation of defined

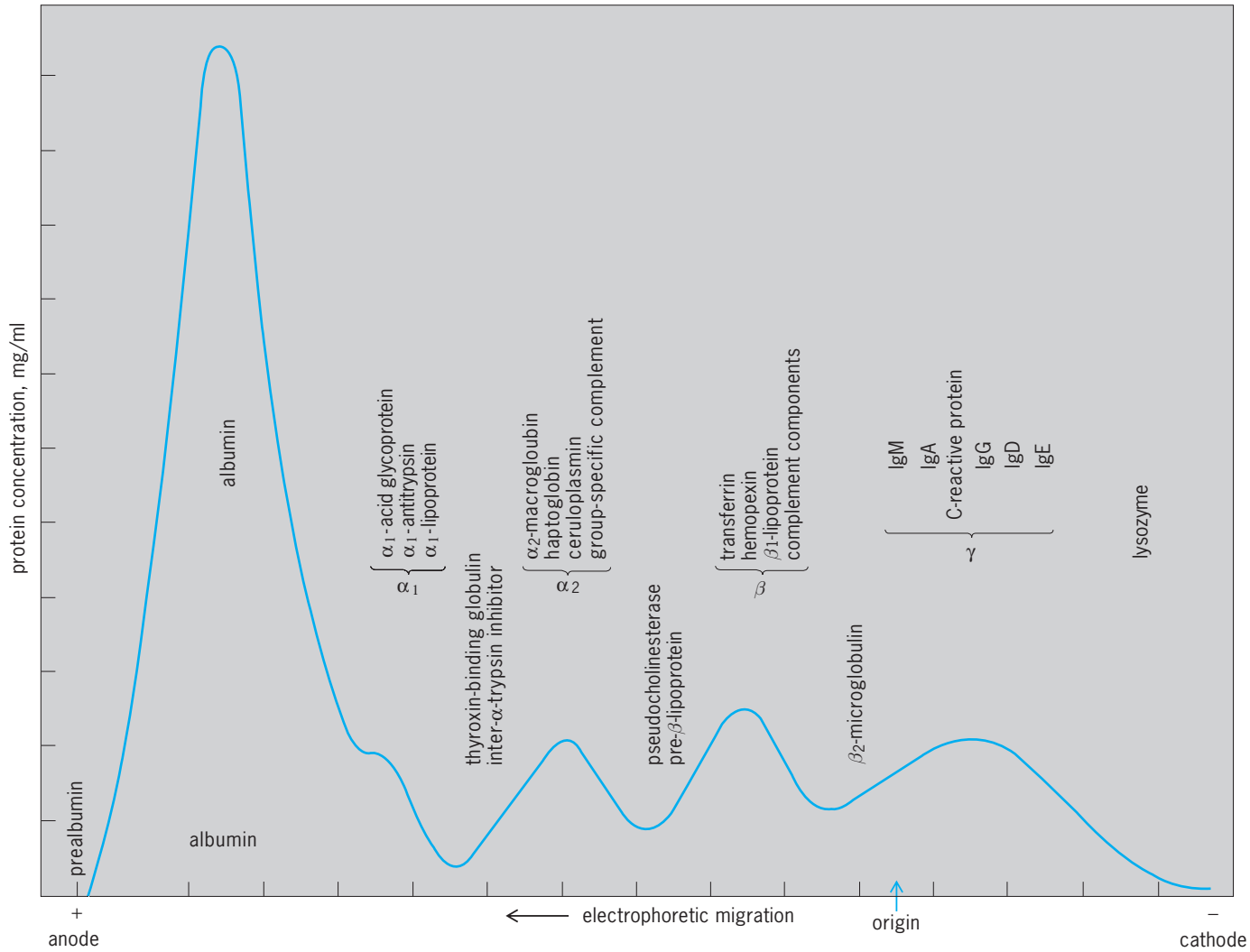

Fig. 3. Paper electrophoresis of normal human serum. Demonstrated is the distribution of several serum proteins in five major electrophoretic fractions, that is, albumin and α_1, α_2, β, and γ-globulins.

components in a high degree of purity are precipitation by various salts or organic compounds, electrophoresis, ultracentrifugation, ion-exchange chromatography, filtration on gel columns, and immunoprecipitation with antibody-containing antisera. Electrophoresis of serum in free solution or on paper strips enables separation of five major fractions called albumin and α_1, α_2, β, and γ-globulins (**Fig. 3**). Electrophoresis on starch or polyacrylamide gels and agar immunoelectrophoresis reveal the presence of more than 30 different protein components in serum. More than 60 protein components have been identified and characterized.

Albumin. This protein makes up more than one-half of the total plasma proteins and has a molecular weight of 69,000. Because of its relatively small molecular size and its high concentration, albumin contributes to 75–80% of the colloid osmotic pressure of plasma. It plays, therefore, a major role in the regulation of the intravascular volume and the fluid exchange between the vascular system and the extravascular space. In addition, albumin serves as a transport protein for various substances. It carries small ions, such as calcium and iodine, as well as organic compounds like bilirubin.

Immunoglobulins. The immunoglobulins, which represent approximately one-sixth of the total protein, largely constitute the γ-globulin fraction. The immunoglobulins are antibodies circulating in the blood, and thus are also called humoral antibodies. They are of great importance in the organism's defense against infectious agents, as well as other foreign substances. Of the different classes of immunoglobulins that can be distinguished, the principal ones are IgG, IgM, IgA, IgD, and IgE. The molecular structure, including the amino acid sequence, of several antibodies has been elucidated, at least in part. Antibodies are formed under the stimulus of a foreign substance, the antigen, and are specificially directed against this antigen. The specificity of an antibody reaction appears to be ensured by a specific molecular structure in certain parts of the antibody molecule. In addition, inherited structural variations are recognized in molecular regions not concerned directly with the antigen-antibody reaction. Immunoglobulins are thus a complex mixture

of a large number—in the order of 10,000–100,000—structurally different antibody molecules. Humoral antibodies are present in most vertebrates, apparently lacking only in certain primitive fishes. Comparative analysis of antibody structure in different species becomes increasingly important for the understanding of evolution. *See* IMMUNOGLOBULIN.

Other proteins. The lipoproteins are another class of serum proteins and make up slightly less than 10% of the total. Presumably they serve as transport proteins for lipids, such as cholesterol, cholesterol esters, and phospholipids. Three major classes can be differentiated: the α_1-, prebeta-, and β_1-lipoproteins.

A number of other serum proteins function as carriers for specific substances. These include transport proteins for metal ions such as the iron-binding protein, transferrin, and the copper-binding protein, ceruloplasmin. The thyroxin-binding globulin transports the thyroid hormone, and transcortin the steroid hormones. Hemoglobin is eliminated from the circulation by haptoglobin, and heme is bound to hemopexin.

Several serum proteins have the capacity to inhibit proteolytic enzymes, thereby protecting serum and possibly tissue proteins against enzymatic degradation. Among others, these include α_1-antitrypsin and the inter-α-trypsin inhibitor. Present in appreciable concentrations are several other glycoproteins, for instance the α_2-macroglobulin and α_1-acid glycoprotein, whose biological functions are not clearly established. A variety of enzymes are also present in plasma, in addition to an array of clotting factors.

Of considerable interest are the complement components. Complement is an important effector system in immune reactions, the target being the cell surface membranes. Its action may result in cell lysis, directed migration of cells, and preparation of damaged cells for uptake and removal by other cells, a process called phagocytosis. *See* COMPLEMENT.

On the one hand, the principal sites of synthesis appear to be the liver cells, where albumin, fibrinogen, α_1-acid glycoprotein, and others are synthesized, and on the other hand, the plasma and lymphoid cells, which are the sites of formation of immunoglobulins and several complement components.

Many plasma proteins of humans and other vertebrates are present in different molecular forms in different individuals. These structural variations are genetically determined. Thus, in addition to the well-known blood groups, there also exist inherited serum protein groups. Examples are the haptoglobins, the group-specific components, which are α_2-globulins with the capacity to bind and transport vitamin D, the different genetic groups of immunoglobulins, and variants of the enzyme pseudocholinesterase. *See* BLOOD GROUPS.

Apart from these inherited qualitative differences, genetically controlled defects of serum proteins exist, that is, the complete or almost complete absence of one or another protein. Some deficiencies are well tolerated by the affected individuals, as analbuminemia and ahaptoglobinemia, the inherited defects for albumin and haptoglobin, respectively. Other defects, however, lead to severe diseases. Immunoglobulin defects (the most common form is agammaglobulinemia, or the lack of γ-globulin) result in increased susceptibility for bacterial infections. These individuals suffer, if left untreated, from severe and recurrent infections. Another example is the defect of the copper-binding protein ceruloplasmin in Wilson's disease, a grave inherited condition with copper deposition in brain, liver, and kidney.

Many serum proteins show changes in their concentrations during disease states. The determination of serum levels of certain proteins is, therefore, important for diagnostic purposes. The most striking abnormalities are observed in a malignant disorder of the plasma cell system called myeloma and in a disease of the lymphoid cell system known as macroglobulinemia Waldenström. The former is commonly associated with the presence of large amounts of a homogeneous protein of the IgG, IgA, IgD, or IgE class in the patient's serum. The latter is characterized by the presence of an increased homogeneous IgM (γ-macroglobulin) fraction. Other serum proteins show altered concentrations in diseases connected with inflammations, with tissue destruction, or with increased loss of proteins into the urine due to kidney damage. In many infections the so-called C-reactive protein appears in measurable quantities in serum. Damage of liver cells leads often to impairment of protein synthesis. Subsequently, serum levels of albumin and clotting factors may decrease. This results in disordered fluid regulation with accumulation of extravascular fluid, particularly in the abdomen, and in bleeding tendency. *See* CLINICAL PATHOLOGY.

Other constituents. In addition to the proteins, many other important classes of compounds circulate in the blood plasma. Most of these are smaller molecules which diffuse freely through cell membranes and are, therefore, more similarly distributed throughout all the fluids of the body and not as characteristic for plasma or serum as the proteins.

In terms of their concentration and their function, the electrolytes are most important. They are the primary factors in the regulation of the osmotic pressure of plasma, and contribute also to the control of the pH. The chief cations are sodium, potassium, calcium, and magnesium. The chief anions are chloride, bicarbonate, phosphate, sulfate, and organic acids. Sodium and chloride are present in the highest concentrations. Potassium, although present in low concentrations, is of considerable importance, and alterations of potassium metabolism may result in hypo- or hyperpotassemia. These occur in a variety of diseases, in which symptoms involve chiefly the muscles and particularly the heart muscle, with characteristic changes in the electrocardiogram. *See* PH REGULATION (BIOLOGY).

The circulating blood, the system which connects the different parts of the body, also contains the many small compounds which are trans-

ported to the sites of synthesis of larger molecules in which they are incorporated, or which are shifted as products of metabolic breakdown to the sites of their excretion from the body. Urea, uric acid, creatinine, and pyruvic acid, important organic metabolites of plasma, are significantly increased in kidney diseases and bring on the symptoms of uremia. Glucose is also an important constituent of plasma because it is a major source of energy for cells throughout the body. Nonesterified fatty acids may, in addition, serve as a source of energy. Free amino acids, the constituents of peptides and proteins, circulate in plasma. Furthermore, bilirubin, vitamin A, steroid hormones, and small peptides like angiotensin and bradykinin are among the many substances present in plasma. Many of these compounds are bound to plasma proteins while they are transported in the blood. This complex formation results generally in a reversible, temporary inactivation of the compound that is being circulated. Hartwig Cleve

Blood Coagulation

When mammalian blood is shed, it congeals rapidly into a gelatinous clot of enmeshed fibrin threads which trap blood cells and serum. Modern theories envision a succession of reactions leading to the formation of insoluble fibrin from a soluble precursor, fibrinogen (factor I). *See* FIBRINOGEN.

Thrombin. The agent ultimately responsible for the formation of fibrin is a proteolytic enzyme, thrombin, which splits four small polypeptide fragments, the fibrinopeptides, from each molecule of fibrinogen. The remainder of the fibrinogen molecule, now called fibrin monomer, polymerizes to form insoluble fibrin, the structure of the clot. The strands of fibrin are given added strength through covalent bonds between adjacent fibrin monomers, brought about by the action of a second enzyme, the fibrin-stabilizing factor (fibrinase, factor XIII).

Thrombin does not exist as such in circulating blood but is generated during clotting from an inactive precursor, prothrombin (factor II). Under physiologic conditions thrombin is generated through one of two mechanisms, known as the extrinsic and intrinsic pathways. The extrinsic pathway comes into action when blood comes into contact with injured tissues, which provide a protein-phospholipid complex, tissue thromboplastin. The interactions of tissue thromboplastin, calcium ions, and several plasma proteins, including factor VII (formerly called proconvertin or pro-SPCA), Stuart factor (factor X), and proaccelerin (labile factor or factor V), lead to the formation of a prothrombin-converting principle which transforms prothrombin into thrombin.

Other clotting factors. Blood also clots when it touches glass or other negatively charged surfaces, through reactions described as the intrinsic pathway. The steps in the intrinsic pathway are complex and involve the participation of at least eight soluble protein factors, leading to the formation of the prothrombin-converting principle. These factors include, in the apparent order of their participation, Hageman factor (factor XII, the agent affected by glass), Fletcher factor (a plasma prekallikrein), high-molecular-weight kininogen (Fitzgerald factor), PTA (plasma thromboplastin antecedent, factor XI), Christmas factor (plasma thromboplastin component, factor IX), anti-hemophilic factor (factor VIII), Stuart factor, proaccelerin, and prothrombin. Several of the steps in this process are dependent upon the presence in blood of calcium ions and of phospholipids, the latter derived principally from blood platelets.

The coagulation of blood can also be induced by certain snake venoms which either promote the formation of thrombin or clot fibrinogen directly, accounting in part for their toxicity. Bacterial enzymes, such as staphylcoagulase, derived from *Staphylococcus aureus*, may also induce clotting.

Liquid-state maintenance. The liquid state of the blood is maintained in the circulation by inhibitors in the blood itself and by the action of the liver and the reticuloendothelial system. It is also possible that small clots formed inadvertently within blood vessels may be dissolved by the plasma proteolytic enzyme, plasmin. *See* PLASMIN.

Platelets. Platelets, besides furnishing phospholipids for the clotting process, help to stanch the flow of blood from injured blood vessels by accumulating at the point of injury, forming a plug. Platelets participate in the phenomenon of clot retraction, in which the blood clot shrinks, expelling liquid serum. Although the function of retraction is unknown, individuals in whom this process is impaired have a bleeding tendency.

Function of clotting process. The principal function of the clotting process is to check the flow of blood from severed vessels. The clotting mechanism is also concerned with the process of thrombosis, that is, the formation of a clot (thrombus) within blood vessels. The earliest step in the intrinsic pathway, the alteration of Hageman factor by contact with glass, is intimately associated with inflammation, initiating chemical processes which lead to the liberation of small polypeptides known as kinins which induce pain, dilate and increase the permeability of small blood vessels, and promote the accumulation of leukocytes. *See* THROMBOSIS.

Hereditary defects. Hereditary deficiencies of the function of each of the protein-clotting factors have been described, notably classic hemophilia (in which antihemophilic factor is present in a nonfunctional form) and Christmas disease (the deficiency of Christmas factor), which are disorders of males and clinically indistinguishable. The various hereditary functional deficiencies are associated with a bleeding tendency with the inexplicable exception of Hageman trait, the inherited deficiency of Hageman factor, Fitzgerald trait (the inherited deficiency of high-molecular-weight kininogen), and Fletcher trait, the inherited deficiency of Fletcher factor. Acquired deficiencies of clotting factors, sometimes of great complexity, are also recognized. Therapy for bleeding due to deficiencies of clotting factors often includes the transfusion of blood plasma or fractions of plasma rich in particular substances

the individual may lack. *See* HEMOPHILIA; HUMAN GENETICS.

Coagulability tests. Clinical tests of the coagulability of the blood include (1) determination of the clotting time, that is, the time elapsing until shed blood clots; (2) the prothrombin time, the time elapsing until plasma clots in the presence of tissue thromboplastin (and therefore a measure of the extrinsic pathway of clotting); (3) the partial thromboplastin time, the time elapsing until plasma clots in the presence of crude phospholipid (and therefore a measure of the intrinsic pathway of clotting); (4) the enumeration of platelets; and (5) crude quantification of clot retraction and of the various plasma protein-clotting factors. The specific diagnosis for deficiency of each of the clotting factors requires an analysis of the characteristics of the person's plasma and, in many cases, direct comparison of such plasma with that of individuals known to be deficient in the factor in question.

Role of vitamin K. Prothrombin, factor VII, Stuart factor, and Christmas factor are synthesized in the parenchymal cells of the liver in the presence of vitamin K, supplied largely by leafy vegetables and intestinal bacteria. Deficiency of vitamin K, or its inhibition by ingested coumarin compounds, impairs the synthesis of the factors and decreases the coagulability of blood. Fibrinogen and proaccelerin are also synthesized in the liver, and antihemophilic factor is synthesized at least partly in endothelial cells, but the site of synthesis of other factors is not known. *See* VITAMIN K.

Heparin. Heparin, a polysaccharide–sulfuric acid complex found particularly in the liver and lungs, impairs coagulation by interfering with the formation of the prothrombin-converting principle and with the action of thrombin; its presence in normal blood is disputed. Both coumarin and heparin are used clinically to impede coagulation in thrombotic states, including thrombophlebitis and coronary heart disease. *See* CIRCULATION; HEPARIN. Oscar D. Ratnoff

Bibliography. P. Astrup and J. W. Severinghaus, *The History of Blood Gases, Acids and Bases*, 1986; S. Gosh, *Handbook of Blood and Blood Products*, 1988; J. M. Jandl, *Blood: Textbook of Hematology*, 2d ed., 1996; D. R. Philips and M. A. Schulman (eds.), *Biochemistry of Platelets*, 1986; S. I. Rapaport, *Introduction to Hematology*, 1987; M. M. Wintrobe, *Blood, Pure and Eloquent*, 1980.

Blood groups

Genetically determined markers on the surface of cellular blood elements (red and white blood cells, platelets). They have medical, legal, and anthropologic importance. In medicine, the matching of ABO and Rh groups of recipients and donors before blood transfusion is of paramount importance; other blood groups also can be implicated in incompatibility. Markers on white cells (histocompatibility antigens) are shared by a number of body tissue cells; these markers are important to the survival of transplanted organs and bone marrow. In law, the recognition of identity between bloodstains found at the scene of a crime and those on clothing of a suspect has resulted in many convictions, and blood typing has served to resolve paternity disputes. From an anthropologic standpoint, some blood groups are unique to specific populations and can be a reflection of tribal origin or migration patterns. Blood groups are also valuable markers in gene linkage analysis, and their study has contributed enormously to the mapping of the human genome.

TABLE 1. Characteristics of blood group antibodies

Characteristic	Naturally acquired	Immune
Immunoglobulin	IgM	IgG
Sedimentation constant	19 S	7 S
Molecular weight	900–1000 kDa	150 kDa
Electrophoretic mobility	Between β and γ	γ
Bind complement	Often	Some
Placental transfer	No	Yes
Direct agglutinin	Yes	Rarely
Hemolytic in vitro	Often	Rarely
Example	Anti-A, anti-B	Rh antibodies

Antibodies. Antibodies are proteins made in response to exposure to blood group antigens. They can be acquired either naturally or through immunization with foreign antigen. The physical and biological properties of these two antibody types are shown in **Table 1**. Human blood can be classified into different groups based on the reactions of red blood cells with blood group antibodies. *See* ANTIBODY; ANTIGEN.

Naturally acquired antibodies. These antibodies are normally found in serum from persons whose red blood cells lack the corresponding antigen. They are of the IgM type, since IgM production is the immune response to carbohydrate antigens and they are stimulated by antigens present in the environment. Many bacteria that constitute normal intestinal flora carry blood group A- and B-like polysaccharides. In newborn infants, as the normal bacterial population of the gut is established, these organisms provide the immune stimulus for anti-A and anti-B production. Three observations support this hypothesis: (1) The serum of newborn infants does not contain naturally acquired anti-A and anti-B; these antibodies begin to appear at about age 3 months. (2) Animals raised in a germ-free environment do not acquire natural antibodies. (3) Infants treated with antibiotics from birth maintain a bacteria-free gut and do not develop anti-A and anti-B until the antibiotic treatment is stopped. Because anti-A and anti-B occur naturally, and because these antibodies can cause rapid life-threatening destruction of incompatible red blood cells, blood for transfusion is always selected to be ABO compatible with the plasma of the recipient.

Immune antibodies. Most blood group antibodies are immune in origin and do not appear in serum or plasma unless the host is exposed directly to foreign red blood cell antigens. The most common stimu-

lating event is blood transfusion or pregnancy. Since many blood group antigens are proteins, the immune response tends to involve first the formation of IgM antibodies and then IgG.

Because of the large number of different blood group antigens, it is impossible, when selecting blood for transfusion, to avoid transfusing antigens that the recipient lacks. However, these foreign antigens can be immunogenic; a single-unit transfusion of Rh D-positive to an Rh D-negative recipient causes production of anti-D in about 85% of cases. Consequently, in addition to matching for ABO types, Rh D-negative blood is almost always given to Rh D-negative recipients. In about 3% of all transfusions, antibodies to other blood group system antigens are formed.

In pregnancy, fetal red blood cells cross the placenta and enter the maternal circulation, particularly at delivery. The fetal red blood cells may carry paternally derived antigens that are foreign to the mother and stimulate antibody production. These antibodies may affect subsequent pregnancies by destroying the fetal red blood cells and causing a disease known as erythroblastosis fetalis. *See* RH INCOMPATIBILITY.

Reactions with red blood cells. Blood group antigen–antibody interactions are detected either directly by agglutination or complement-mediated hemolysis or indirectly by use of the antiglobulin test. The stages in these tests can be summarized as follows:

Direct tests, for the detection of IgM antibodies:

1. Mix serum and red blood cells.
2. Incubate (room temperature, optional).
3. Centrifuge.
4. Examine for agglutination and hemolysis.

Indirect tests, for the detection of IgG antibodies:

1. Mix serum and red blood cells (an enhancement reagent, to promote antibody uptake, may be incorporated here).
2. Incubate (37°C).
3. Examine for agglutination and hemolysis (optional).
4. Wash to remove unbound globulins.
5. Add antihuman globulin reagent.
6. Centrifuge.
7. Examine for agglutination.
8. Confirm negative tests with IgG-coated red blood cells.

Direct reactions occur in two phases: a coating phase in which antibody attaches to corresponding antigens on red blood cells; and an agglutination (clumping) phase during which red blood cells come close enough together for antibody molecules to span intercellular distances and bind to adjacent red blood cells to form an agglutination lattice.

Immunoglobulin M (IgM) antibodies directly agglutinate saline-suspended red blood cells; for this reason, they have been called complete antibodies. The ability of IgM antibodies to react in this manner is a reflection of both size and valency, that is, the number of antigen-binding sites (maximum of 10) per molecule. IgM antibodies are large enough to span the distance between red blood cells in suspension. Being multivalent, they can attach to antigens on adjacent red blood cells. Many IgM antibodies, especially naturally acquired anti-A and anti-B, also cause complement-mediated lysis of red blood cells.

Immunoglobulin G (IgG) antibodies, with only two antigen-binding sites per molecule, do not regularly cause direct agglutination of saline-suspended red blood cells, although they may do so if high-molecular-weight compounds or proteins (such as polyvinylpyrrolidone or bovine albumin) are added to the reaction medium. Because an additive is required to complete the agglutination of IgG-coated red blood cells, IgG antibodies have been called incomplete antibodies. In simple terms, albumin neutralizes the effect of the negative charge on red blood cells. This negative charge serves to keep red blood cells apart in suspension. Reduction in surface charge reduces intercellular repulsion, permitting red blood cells to come close together so that antigen-binding sites on IgG molecules can attach to adjacent red blood cells. *See* IMMUNOGLOBULIN.

Antigens, genes, and blood groups. Almost 300 distinct blood group antigens have been identified on human red blood cells. Biochemical analysis has revealed that most antigen structures are either protein or lipid in nature; in some instances, blood group specificity is determined by the presence of attached carbohydrate moieties. The human A and B antigens, for example, can be either glycoprotein or glycolipid, with the same attached carbohydrate structure. With few exceptions, blood group antigens are an integral part of the cell membrane.

A number of different notations are used, and different concepts have been put forth to explain the genetics of the human blood groups. The presence of a gene in the host is normally reflected by the presence of the corresponding antigen on the red blood cells. Usually, a single locus determines antigen expression, and there are two or more forms of a gene or alleles (for example, *A* and *B*) that can occupy a locus. Each individual inherits one allele from each parent. For a given blood group, when the same allele (for example, allele *A*) is inherited from both parents, the offspring is homozygous for *A* and only the antigen structure (A) will be present on the red blood cells. When different alleles are inherited (that is, *A* and *B*), the individual is heterozygous for both *A* and *B*, and both A and B antigens will be found on the red blood cells.

In some blood group systems, two genes (loci) govern the expression of multiple blood group antigens within that system. These loci are usually closely linked, located adjacent to each other on the chromosome. Such complex loci may contain multiple alleles and are referred to as haplotypes. *See* HUMAN GENETICS.

Blood group systems. Approximately 250 red cell antigens have been assigned to 29 different blood group systems (**Table 2**). For an antigen to form a new blood group system, the antigen must be defined by a human alloantibody, it must be an inherited

TABLE 2. Human blood group systems

System name	ISBT* symbol	System number	Antigens in system†	Chromosome location‡	Gene products	CD number§
ABO	ABO	001	4	9q34.1-q34.2	A = α-*N*-acetylgalactos-aminyl transferase B = α-galactosyl transferase	
MNS	MNS	002	43	4q28-q31	*GYPA* = glycophorin A; 43-kDa single-pass glycoprotein *GYPB* = glycophorin B; 25-kDa single-pass glycoprotein	CD235
P	P1	003	1	22q11-qter	α-Galactosyl transferase	
Rh	RH	004	50	1p36.13-p34	*RHD* and *RHCE*, 30–32-kDa multipass polypeptides	CD240
Lutheran	LU	005	19	19q13.2	78- and 85-kDa single-pass glycoproteins	CD239
Kell	KEL	006	27	7q33	93-kDa single-pass glycoprotein	CD238
Lewis	LE	007	6	19p13.3	α-Fucosyl transferase	
Duffy	FY	008	6	1q22-q23	38.5-kDa multipass glycoprotein	CD234
Kidd	JK	009	3	18q11-q12	43-kDa multipass glycoprotein	
Diego	DI	010	21	17q12-q21	95–105-kDa multipass glycoprotein	CD233
Yt	YT	011	2	7q22	160-kDa dimeric GPI-linked glycoprotein	
Xg	XG	012	2	Xp22.33, YP11.32	22–29-kDa single-pass glycoprotein	CD99
Sciana	SC	013	5	1p36.2-p22.1	60–68-kDa glycoprotein	
Dombrock	DO	014	5	12p13.2-12p12.1	47–58-kDa GPI-linked glycoprotein	
Colton	CO	015	3	7p14	28- and 40–60-kDa multipass glycoproteins; aquaporin 1	
Landsteiner-Weiner	LW	016	3	19p13.3	37–43-kDa single-pass glycoprotein	CD242
Chido/Rogers	CH/RG	017	9	6p21.3	Glycoproteins adsorbed onto red blood cells	
Hh	H	018	1	19q13	α-Fucosyl transferase	
Kx	XK	019	1	Xp21.1	37-kDa multipass glycoprotein	CD173
Gerbich	GE	020	8	2q14-q21	*GYPC* = glycophorin C; 40-kDa single-pass glycoprotein. Or glycophorin D; 30-kDa single-pass glycoprotein.	CD236
Cromer	CROM	021	13	1q32	CD55; GPI-linked glycoprotein	CD55
Knops	KN	022	8	1q32	CD35; single-pass glycoprotein	CD35
Indian	IN	023	2	11p13	CD44; 80-kDa single-pass glycoprotein	CD44
Ok	OK	024	1	19pter-p13.2	CD147; 35–69-kDa glycoprotein	CD147
Raph	RAPH	025	1	11p15.5	MER2; glycoprotein	
John Milton Hagen	JMH	026	1	15q24.1	GPI-linked protein	CD108
I	I	027	1	6p24.2	*N*-acetylglucosaminyl transferase	
Globoside	GLOB	028	1	3q26.1	β-galactosyl transferase	
Gill	GIL	029	1	9p13.3	Aquaporin 3	

*International Society of Blood Transfusion.
†Nonglycolipid antigens on red blood cells are either single-pass/multipass proteins or glycosylphosphatidylinositol (GPI)-linked proteins.
‡Chromosome locations of genes/loci are identified by the arm (p = short; q = long), followed by the region, then by the band within the region, in both cases numbered from the centromere; ter = end.
§CD = cluster designation.

character, the gene encoding it must have been identified and sequenced, and its chromosome location must be known. Further, the gene involved must be distinct from other blood group system genes, and not a closely linked homologue of a known blood group gene.

ABO blood group. Discovered by Karl Landsteiner in 1901, ABO was the first human blood group system to be described (**Table 3**). Three major alleles at the *ABO* locus on chromosome 9 govern the expression of A and B antigens. Gene *A* encodes for a protein (α-*N*-acetylgalactosaminyl transferase) that attaches a blood group–specific carbohydrate (α-*N*-acetyl-D-galactosamine) and confers blood group A activity to a preformed carbohydrate structure called H antigen. Gene *B* encodes for an α-galactosyl transferase that attaches α-D-galactose and confers blood group B activity to H antigen. In both instances, some H remains unchanged. The *O* gene has no detectable product; H antigen remains unchanged and is strongly expressed on red blood cells.

TABLE 3. ABO blood group system

Blood group	RBC antigens	Possible genotypes	Plasma antibody
A	A	*A/A or A/O*	Anti-B
B	B	*B/B or B/O*	Anti-A
O	—	*O/O*	anti-A and anti-B
AB	A and B	*A/B*	—

These three genes account for the inheritance of four common phenotypes: A, B, AB, and O. Blood types A and O are the most common, and AB the least common. The *A* and *B* genes are codominant; that is, when the gene is present the antigen can be detected. The *O* gene is considered an amorph since its product cannot be detected. When either A or B antigens are present on red blood cells, the corresponding antibody or antibodies should not be present in the serum or plasma. In adults, when A or B or both are absent from the red blood cells, the corresponding naturally acquired antibody is present in the serum. This reciprocal relationship between antigens on the red blood cells and antibodies in the serum is known as Landsteiner's law. Other ABO phenotypes do exist but are quite rare. Further, the A blood type can be subdivided, based on strength of

antigen expression, with A_1 red blood cells having the most A antigen.

At the molecular level, *A* and *B* allelic complementary deoxyribonucleic acids (cDNAs) are identical except for seven nucleotide substitutions resulting in four amino acid substitutions between A and B transferases. The majority of O alleles have a single nucleotide substitution close to the N-terminal of the coding sequence that shifts the reading frame of codons and results in translation of a protein lacking A or B transferase activity.

In addition to the ABO locus, expression of A and B is influenced by *H/h* and *Se/se* genes on chromosome 19. The *H* gene–specified transferase attaches fucose to paragloboside, a red blood cell membrane glycolipid. The H antigen thus formed is the structure

TABLE 4. Function and disease associations of blood group antigens

Antigens	Function/disease associations
A/B	Weakened in leukemia, especially cases involving chromosome 9 translocations, and in stress-induced hemopoiesis
	Altered (A to B) in acquired-B phenomenon due to modification by bacterial enzymes, especially in patients with septicemia or lesions of the gastrointestinal tract
MNSsU	Receptor for some *Escherichia coli* and for *Plasmodium falciparum*
	Receptor for complement bacteria and viruses
	Glycosylation imparts a negative electrical charge to red blood cells
Rh	May be expressed weakly in leukemia and myeloproliferative disorders
	Common target for IgG autoantibodies causing immune hemolytic anemia
	May play a role in transport of ammonia across red blood cell membranes
Lu	Adhesion molecule, integrin
Kel	Possible endopeptidase activity
	Target antigen in 0.4% of IgG autoantibodies causing immune hemolytic anemia
	Weakly expressed on red blood cells, in some cases X-linked chronic granulomatous disease, associated with acanthocytosis
	Absence in Ko (null phenotype) associated with cardiomyopathy, hypertension
Le	Le^b antigen is receptor for *Helicobacter pylori*
Fy	70% of African-Americans are Fy(a-b-); such red blood cells are resistant to invasion by *Plasmodium vivax*
Jk	Urea transporter
Di	Synonymous with band 3 or anion exchanger (AE1), involved with HCO_3^-, Cl^- transport across red blood cell membrane
Yt	Synonymous with red blood cell membrane acetylcholinesterase
	Absent on cells deficient in glycosylphosphatidylinosityl-linked proteins (paroxysmal noctumal hemoglobinuria III cells)
	Reduced levels in myelodysplasia associated with chromosome 7 abnormalities and in some cases of systemic lupus erythematosus
Xg	Adhesion molecule
Do	Absent on cells deficient in glycosylphosphatidylinosityl-linked proteins (paroxysmal nocturnal hemoglobinuria III cells)
Co	Synonymous with CHIP-1, water transport protein
	Absence associated with monosomy 7, congenital dyserythropoietic anemia
	Weak expression seen in some chromosome 7 rearrangements, leukemia
LW	Intercellular adhesion molecule (ICAM) that binds to leukocyte integrins
CH/RG	Synonymous with C4 (C3, C4, and C5 are components of human complement)
	Patients with inherited low levels of C4 are prone to develop autoimmmue disorders (insulin-dependent diabetes, autoimmune chronic active hepatitis)
	Null genes and specific allotypes are associated with Graves' disease and rheumatoid arthritis
	C4B-deficient (Ch-negative) children have increased susceptibility to bacterial meningitis
	Rg-negative individuals are predisposed to develop systemic lupus erythematosus
H	Increased expression in hemopoietic stress
	Weakened in leukemia and on malignant tissue cells
Kx	Possible transport protein
	Essential for red blood cell membrane integrity; absence on red blood cells associated with McLeod syndrome and acanthocytosis
	Absence on granulocytes associated with X-linked chronic granulomatous disease
GE	Interact with red blood cell membrane protein 4.1
	Reduced expression associated with hereditary elliptocytosis
CROM	Inhibits assembly and accelerates decay of C3 and C5 convertases
	Absent on cells deficient in glycosylphosphatidylinosityl-linked proteins (paroxysmal nocturnal hemoglobinuria III cells)
	Null phenotype (INAB) associated with intestinal disorders
KNOPS	Synonymous with complement receptor 1 (CR1)
	Mediates phagocytosis of cells coated with C3/C4
	Possible site for invasion by babesia
IN	Adhesion molecule
Raph	Raph-negative individuals predisposed to renal failure
Ii	Altered in diseases resulting in hemopoietic stress
	Anti-I associated with cold agglutinin disease and mycoplasma pneumonia
	Anti-i associated with infectious mononucleosis and lymphoproliferative disorders
P	Synonymous with globoside, receptor for Parvovirus B19
	Target for IgG autoantibody causing paroxysmal cold hemoglobinuria
	Cytotoxic IgM and IgG3 antibodies to globoside are associated with spontaneous abortions
AnWj	Receptor for *Haemophilus influenzae*
JMH	Absent on cells deficient in glycosylphosphatidylinosityl-linked proteins (paroxysmal nocturnal hemoglobinuria III cells)
	JMH-negative phenotype associated with congenital dyserythropoietic anemia

on red blood cells that serves as the substrate for *A* and *B* gene–specified transferases described above. In rare individuals, absence of *H* (that is, homozygous *h*) results in absence of H structures on red blood cell membranes and no substrate on which *A* and *B* gene products can act. The resulting phenotype is referred to as Bombay (or O_h). The *Se* gene governs the addition of fucose to type 1 chains in the secretions (such as saliva, plasma, urine) to form water-soluble H-active structures. *Se* is present in about 80% of the population; depending on genes present at the ABO locus, the secretions of *Se* individuals will contain A or B antigens or both in addition to H antigens.

Rh blood group. Currently 50 antigens are assigned to the Rh blood group system, although D is the most important. Red blood cells that carry D are called Rh-positive; red blood cells lacking D are called Rh-negative. Other important Rh antigens are C, c, E, and e. Rh antigen expression is controlled by two adjacent homologous structural genes on chromosome 1 that are inherited as a pair or haplotype. The *RhD* gene encodes D antigen and is absent on both chromosomes of most Rh-negative subjects. The *RhCE* gene encodes CE protein. Nucleotide substitutions account for amino acid differences at two positions on the CE protein and result in the Cc and Ee polymorphisms. Point mutations or rearrangement of exons from *D* with those of *CE* appear to account for other Rh polymorphisms.

Other blood groups. Prior to the discovery of Rh, Landsteiner and Philip Levine discovered three more blood group antigens (M, N, and P1) following injection of human red blood cells into rabbits. After adsorption of the rabbit serum to remove species antibodies, anti-M and anti-N were found to agglutinate 78% and 72% of human red blood cell samples, respectively. Family studies showed that M and N were inherited as the products of alleles. Later, through use of the antiglobulin test, S and s antigens were assigned to the MN system, and now over 40 MNS system antigens have been described. Anti-P1 reacted with red blood cells from 78% of Caucasians. The *P1* gene was shown to be inherited independently from the *ABO* and *MN* loci. Human examples of antibodies to M, N, and P1 were soon discovered, but they tend not to react at body temperatures and are, therefore, not usually of concern in blood transfusion practice.

Use of the antiglobulin test led to the discovery of most other blood group antigens and systems. Many of the antibodies defining the antigens were encountered as the cause of hemolytic disease of the newborn. Since such antibodies also cause accelerated destruction of transfused incompatible red blood cells, it is necessary to test a patient's serum for unexpected antibodies prior to blood transfusions.

Antigens that do not meet the criteria for assignment to a specific blood group system have been placed into collections, based primarily on biochemical data or phenotypic association, or into a series of either high- or low-frequency antigens.

Biological role. The function of blood group antigens has been increasingly apparent. Single-pass proteins such as the LU and XG proteins are thought to serve as adhesion molecules that interact with integrins on the surface of white blood cells. Multipass proteins such as band 3, which carries the DI system antigens, are involved in the transportation of ions through the red blood cell membrane bilipid layer. GPI-linked proteins such as CROM, also known as decay accelerating factor (DAF), are involved with complement regulation on the red blood cell surface. Some blood group antigens are essential to the integrity of the red blood cell membrane, for their absence results in abnormal surface shape; for example, absence of KEL protein leads to the formation of acanthocytes, and absence of RH protein results in stomatocytosis and hemolytic anemia (**Table 4**).

Many membrane structures serve as receptors for bacteria and other microorganisms (Table 4). For example, the FY or Duffy protein is the receptor on red blood cells for invasion by *Plasmodium vivax*, the cause of benign tertian malaria. Particularly significant is the fact that Fy(a-b-) phenotype is virtually nonexistent among Caucasians but has an incidence of around 70% among African-Americans. Presumably, the Fy(a-b-) phenotype evolved as a selective advantage in areas where *P. vivax* is endemic. Similarly, the S-s-U-red blood cell phenotype in the MNS blood group system affords protection against *P. falciparum*, or malignant tertian malaria. Yet other blood group antigens can be altered in disease states; A, B, and H antigens are sometimes weakened in leukemia or may be modified by bacterial enzymes in patients with septicemia. *See* BLOOD; IMMUNOLOGY.

W. John Judd

Bibliography. G. L. Daniels, *Human Blood Groups*, 2d ed., Blackwell Science, Oxford, 2002; P. D. Issitt and D. J. Anstee, *Applied Blood Group Serology*, 4th ed., Montgomery Scientific Publications, Durham, NC, 1998; M. E. Reid and C. Lomas-Francis, *Blood Group Antigen Factsbook*, 2d ed., Academic Press, San Diego, 2003.

Blood vessels

Tubular channels for blood transport, of which there are three principal types: arteries, capillaries, and veins. Only the larger arteries and veins in the body bear distinct names.

Arteries carry blood away from the heart through a system of successively smaller vessels. The three major layers which exist in the arterial wall are the outer adventitia of connective tissue, the middle elastic or muscular layer or both, and the inner lining or endothelium of squamous cells. The proportionate thickness of these layers varies according to the size and location of the artery. In the largest arteries close to the heart, such as the aorta, elastic fibers are prominent, allowing expansion and recoil with each systolic contraction of the heart. The medium-sized arteries have a predominantly muscular wall (smooth muscle fibers) with some elastic fibers. In the smaller arterioles the middle coat is entirely muscle, and even this is gradually lost as the arterioles become smaller.

Capillaries are the smallest but most extensive blood vessels, forming a network everywhere in the body tissues. Capillary beds form a huge total surface area of thin endothelium, serving as a semipermeable membrane for passage of materials between the circulatory system and body tissues. Normally, only a small portion of a capillary bed carries blood, the remaining portion being temporarily bypassed. Such alterations allow for shifts of blood, as well as increases or decreases in volume, of a particular region or organ, and constitute one of the homeostatic mechanisms of the body. *See* HOMEOSTASIS.

Veins carry blood from the capillary beds back to the heart through increasingly larger vessels. The walls of veins and venules superficially resemble those of arteries, but with several marked differences. Veins have larger diameters and thinner walls and are more distensible than corresponding arteries; they have much fewer muscle fibers and elastic fibers in their walls. Veins differ from arteries in having valves which prevent backflow of blood. The pattern of branching of veins is less regular and less constant than that of their arterial counterparts. The larger veins and arteries contain small blood vessels (vasa vasorum) for nourishment of the walls, and are innervated by a complex system of nerves from the autonomic nervous system, which innervates the smooth muscles. Contraction of these muscles is also under control of various hormones, such as norepinephrine.

In certain locations blood vessels are modified for particular functions, as the sinusoids of the liver and the spleen and the choroid plexuses of the brain ventricles. *See* LYMPHATIC SYSTEM. Walter Bock

Blowout coil

A coil that produces a magnetic field in an electrical switching device for the purpose of lengthening and extinguishing an electric arc formed as the contacts of the switching device part to interrupt the current. The magnetic field produced by the coil is approximately perpendicular to the arc. The interaction between the arc current and the magnetic field produces a force driving the arc in the direction perpendicular to both the magnetic flux and the arc current (**Fig. 1**).

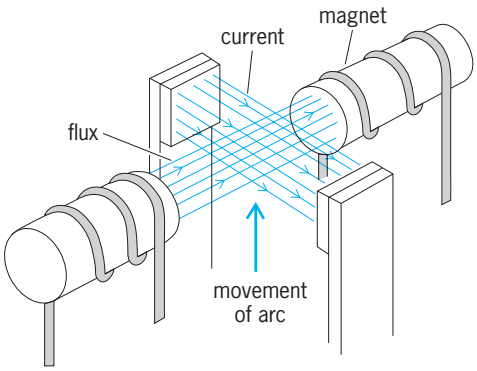

Fig. 1. Relation of directions of current, magnetic flux, and movement of arc.

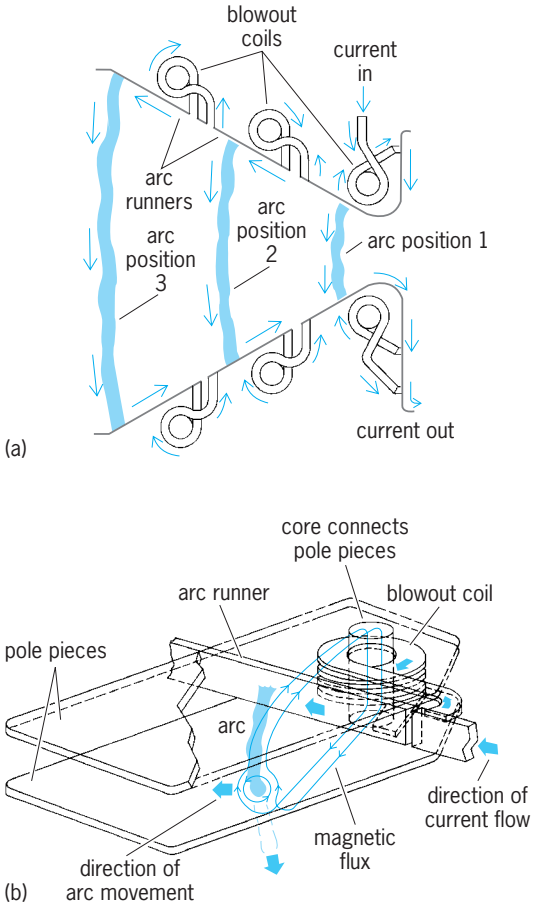

Fig. 2. Arc suppressor: (*a*) cross section and (*b*) cut-away view. As the arc progresses along the runners, it draws current through the blowout coils whose magnetic fields then drive the arc further out until it extinguishes.

In alternating-current circuits the arc is usually extinguished at in instant of time when the current passes through zero before reversing its direction. If, within a short time around current zero, sufficient energy can be removed from the arc, conduction will cease. The function of the blowout coil is to move the arc into a region inside a circuit interrupter, such as arc chutes in circuit breakers, where the energy removal process takes place. *See* CIRCUIT BREAKER.

In direct-current circuits there are no natural current zeros. An arc strikes when a switching device opens. Current will continue to flow as long as the voltage across the open contacts is sufficient to sustain the arc. To interrupt this dc arc, the arc must be converted to a form that, to continue, requires more than the available voltage. Thus, in a dc switching device, the function of the blowout coil is to increase the length of the arc, which increases the arc voltage, and also to move the arc into regions, inside the interrupter, where the arc voltage can be further increased.

The blowout coil is either connected in series with contacts or inserted as the arc moves along an arc runner (**Fig. 2**). Thomas H. Lee

Blowpipe

In glass blowing, a long straight tube on which molten glass is gathered and worked, partly by blowing into the tube. The blowpipe is spun to shape the glass object further by centrifugal force, or by a tool, in which case the blowpipe acts as a spindle for turning.

In analytical chemistry, a blowpipe is a small, tapered, and frequently curved tube that directs a jet, usually of air, into a flame to concentrate the flame onto an unknown substance. Coloration of the flame and other characteristic reactions in the reducing and in the oxidizing portions of the flame created by the blowpipe aid in identifying the substance. Such a blowpipe may be blown directly by the analyst, from a bellows, or from a pressurized line.

Frank H. Rockett

Blue straggler star

A star that is a member of a stellar association and is located at an unusual position on the association's color-magnitude diagram, above the turnoff from the main sequence.

Characteristics. A stellar cluster is observationally characterized by its color-magnitude diagram, which represents each member's luminosity plotted against its color, which in turn is related to its effective temperature. The location and structure of the various regions of a cluster's color-magnitude diagram can be made to yield crucial information on its age and chemical composition by exploiting the knowledge of the evolutionary tracks of stars with a specified range of initial masses and abundances. In order to make this method work properly, however, it is assumed that all the stars in a cluster are coeval (born at the same time), single, and isolated objects and that they are internally stable, nonrotating, nonmagnetic, and unmixed. The remarkable overall consistency between the observed color-magnitude diagrams and stellar evolutionary theory provides the most convincing evidence for the basic soundness of both the theory and the assumptions. *See* COLOR INDEX; HERTZSPRUNG-RUSSELL DIAGRAM; MAGNITUDE (ASTRONOMY).

The existence of a special group of stars known as blue stragglers represents a challenge to this tidy scenario. They were discovered by A. Sandage in 1953 in the galactic globular cluster M3 (see **illustration**). The blue stragglers are located below the horizontal branch and above the turnoff from the main sequence. They form a new sequence extending from the extrapolation of the main sequence to higher luminosities on the left to the red giant branch on the right. While the majority of the cluster members fall rather precisely on the expected isochrones for the approximately 15×10^9-year age of the cluster, the 50 or so blue stragglers in M3 are located on much younger isochrones. It seems, in other words, as if the blue stragglers somehow never made it across the gap between the main sequence and the red giant branch as expected from stellar evolutionary theory and are, therefore, true stragglers with respect to the main body of stars in the cluster.

Occurrence. More than 3000 objects of this type have been found in practically every known type of dense stellar association, including dwarf spheroidal galaxies. They represent, therefore, a ubiquitous phenomenon. In the case of the globular clusters where this phenomenon is most pronounced, blue stragglers are found in every cluster observed so far in numbers ranging from a few dozen to a few hundred. *See* GALAXY, EXTERNAL.

In most globular clusters, they are found mainly in or very near the dense core, but in a few they have also been detected in the outer regions. This suggests that their radial distribution is bimodal with a high peak in the center decreasing in the middle and rising again in the periphery of the cluster. The total number of blue stragglers in a cluster, however, does not seem to be correlated with either the stellar density in the core or its total mass.

All indications point to these objects physically having masses consistent with their observed position above and blueward of the turnoff point of the cluster color magnitude diagram. Thus, blue stragglers in globular clusters can be clearly defined as objects having the following general range of physical parameters: temperatures between 6000 and 10,000 K, masses between 1 and 2.5 solar masses, luminosities between 3 and 30 solar luminosities, and low metal abundances consistent with the rest of the stars in the cluster.

Origin. Since such stars cannot exist in the cluster under the assumptions made above, some or all of the assumptions must be wrong. The simplest explanation of the phenomenon, of course, is violation of the coeval hypothesis; that is, the blue stragglers really are younger than the rest of the cluster members,

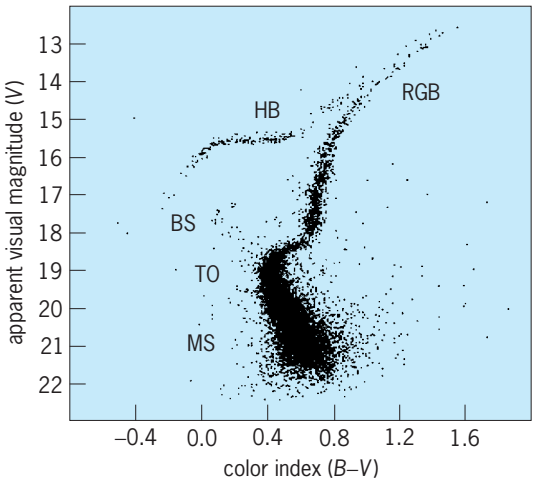

Color-magnitude diagram [apparent visual magnitude *(V)* versus color index *(B-V)*] of 10,637 stars in the galactic globular cluster M3 (NGC 5272). The important evolutionary stages are marked by MS (main sequence), TO (turn-off), RGB (red giant branch), HB (horizontal branch), and BS (blue stragglers). (*After R. Buonanno et al., High precision photometry of 10,000 stars in M3, Mem. Soc. Astron. It., 57:391-393, 1986*)

perhaps because of a recent burst of star formation. This delayed-formation scenario is possible in young open clusters and some dwarf spheroidal galaxies. However, since there is no evidence at all for recent star formation episodes in the galactic globular clusters, some other assumption must be flawed if the blue stragglers in these systems are to be explained. Three possibilities related to delayed evolution or extended lifetime in a strictly coeval ensemble have been analyzed: internal mixing that supplies new fuel from the outer layers to the core, mass transfer or merging of two low-mass stars in a binary system, or a collision between two or more unrelated stars. These last two events would lead to the formation of a relatively unevolved, more massive object only apparently younger than the rest of the cluster population. Apart from the lack of a plausible physical mixing mechanism in a single star, observations of the surface gravity and luminosities of the more massive blue stragglers seem to be inconsistent with the first hypothesis. *See* BINARY STAR.

Therefore, most astronomers in the field favor the last scenarios of the mass transfer or merging of stars in a primordial binary system and the formation of a relatively unevolved massive star by direct collision of two stars that involves a resonant 3- or 4-body encounter. The objects formed by the latter mechanism would be expected to be found in the inner denser regions where the collision probability between stars is very high, while the former mechanism would be expected to give rise preferentially to objects outside the core where primordial binaries are not so easily disrupted by encounters with neighboring objects.

All the characteristics described above can be reasonably well understood if it is assumed that both types of blue stragglers coexist in a globular cluster. The possibility of observationaly distinguishing these two populations is the subject of intense debate. The cluster-to-cluster variability in numbers of blue stragglers in the core may be the best indication available of the cluster's dynamic state, since the cluster must use its binaries to reduce the catastrophic effects of the runaway collapse that many clusters face at some point in their evolution.

By their very existence, blue stragglers sharpen the understanding of the nature of the stellar populations and their interactions in all clusters. They represent, therefore, a crucial tool in deciphering the origin and history of these objects. *See* STAR CLUSTERS; STELLAR EVOLUTION. Francesco Paresce

Bibliography. M. B. Davies et al., Blue straggler production in globular clusters, *Mon. Not. Roy. Astron. Soc.*, 349:129–134, 2004; M. Mapelli et al., The contribution of primordial binaries to the blue straggler population in 47 Tucanae, *Astrophys. J.*, 605:29–32, 2004.

Blueberry

Several species of the genus *Vaccinium*, plant order Ericales, ranging from low-growing, almost prostrate plants to vigorous shrubs reaching a height of 12–15 ft (3.5–4.5 m). The fruit, a berry, is usually black and covered with bluish bloom, generally occurring in clusters and has numerous small seeds, a characteristic that distinguishes the blueberry from the huckleberry, which has 10 rather large, gritty seeds. Although there are blueberry species on other continents, all cultivated varieties in the United States are American in origin. *See* ERICALES.

Types. The dryland blueberry (*V. ashei*) is adapted to relatively dry soils and has been brought under cultivation in Florida and Georgia.

In the Northeast the lowbush blueberry (*V. lamarckii*) grows wild over thousands of acres of dry hillsides, where it is harvested commercially, especially in Maine, but also in other New England states, Michigan, Minnesota, and a few others. Some berries are sold fresh, but most are canned or frozen to be used for pies. In some areas the plants are burned over every 3 years to kill certain weeds and to remove the old wood, thus stimulating vigorous new growth. Lowbush blueberries are harvested with a small scoop, resembling a cranberry scoop, which combs the berries from the plants, whereas most of the highbush and dryland berries are picked by hand.

In the Northwest fruit of the evergreen blueberry (*V. ovatum*) is harvested in the wild, and tons of the leafy twigs are shipped for use as florists' greens.

The highbush blueberry, represented by *V. australe* and *V. corymbosum*, provides most of the cultivated plants. These are bushes reaching 6–8 ft (1.8–2.4 m) in height, and normally propagated by hardwood or softwood cuttings. The highbush blueberry is found in swampy land, usually on hummocks, so that its roots are not actually submerged. It does best in an acid soil, preferably sand and peat, and must have an ample water supply and good drainage. It can be grown on ordinary mineral soils if the water is properly regulated and the soil acidity is kept in the range pH 4–6. Nitrogen is utilized as ammonium rather than nitrate. The highbush blueberry has been under cultivation since 1910, when the first experiments on its culture were reported. Production in the United States is greatest in New Jersey, followed by Michigan, North Carolina, and Washington. The fruit is sold fresh, canned, and frozen. The blueberry is becoming increasingly popular in home gardens, although its cultural requirements are rather exacting. For garden culture, mulching with sawdust or other organic matter is desirable because the roots are shallow. *See* FRUIT. J. Harold Clarke

Diseases. Blueberry diseases and their control are a major concern of commercial growers and home gardeners. Blueberries are affected by numerous diseases; however, the severity and economic importance vary from one blueberry growing region to another.

In the southeastern United States, fungal diseases such as stem canker (caused by *Botryosphaeria corticis*), stem blight (*B. dothidea*), and leaf spot (*Gloesporium minus*) are very prevalent and destructive, but are of little or no importance in the northern areas. Godronia canker, caused by the fungus *Godronia cassandrae*, is important in the northern

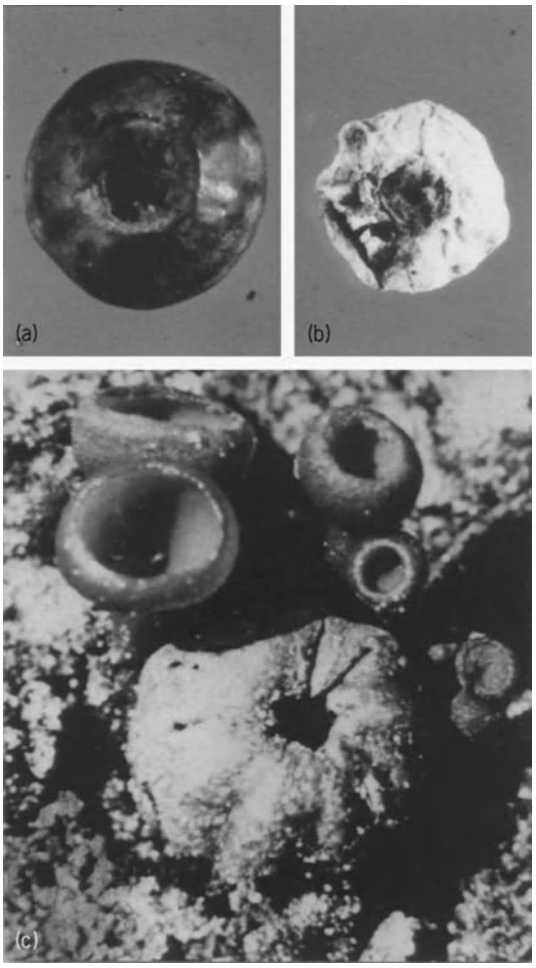

Mummy berry disease caused by the fungus *Monilinia vaccinii-corymbosi*. (*a*) Healthy fruit. (*b*) Diseased fruit. (*c*) Mummified fruit with spore-bearing structures called apothecia. (*Parts a and b from R. D. Milholland, Sclerotium germination and histopathology of Monilinia vaccinii-corymbosi on highbush blueberry, Phytopathology, 67:848–854, 1977; part c from R. D. Milholland, Factors affecting apothecium development of Monilinia vaccinii-corymbosi from mummified highbush blueberry fruit, Phytopathology, 64:296–300, 1974*)

regions of British Columbia, Michigan, Maine, Massachusetts, and Nova Scotia, but is not found in the south. Resistant varieties are the only means of controlling stem cankers and blights. Leaf spotting diseases can be controlled with fungicides.

Blueberry stunt is a destructive disease caused by a primitive bacterialike organism, a mycoplasma. Characteristic symptoms are small leaves, yellowing along the margins and between lateral veins, and reduced plant vigor. Viral diseases such as mosaic, shoestring, and necrotic ringspot are normally of minor importance; however, the red ringspot virus is a serious problem in New Jersey. The use of certified or virus-free plants, as well as roguing diseased bushes, is required for effective control of stunt and viral diseases.

Mummy berry, caused by the fungus *Monilinia vaccinii-corymbosi*, is one of the most widespread and devastating diseases of blueberry. The disease first appears as a blighting of new growth, and prior to harvest the infected berries become a light cream color rather than normal blue and drop to the ground. The pathogenic cycle of the fungus is repeated the following spring when small spore-bearing structures called apothecia are produced from overwintered mummies (see **illus.**). Spores from these structures initiate primary cycles of the pathogen. In general, the rabbiteye blueberry, although severely affected by mummy berry, is more tolerant of stem cankers and root rots than high bush blueberry. The most effective control practice for mummy berry would combine the use of an eradicant ground treatment or clean cultivation and the application of protectant fungicides. Since recommended control practices vary from one region to another, it is important to consult local agricultural authorities before applying any chemicals. *See* PLANT PATHOLOGY.

Robert D. Milholland

Bibliography. T. A. Chen, Mycoplasma-like organisms in sieve tube elements of plants infected with blueberry stunt and cranberry false blossom, *Phytopathology*, 61:233–238, 1971; R. D. Milholland and G. J. Galletta, Pathogenic variation among isolates of *Botryosphaeria corticis* on blueberry, *Phytopathology*, 59:1540–1543, 1969; D. C. Ramsdell, J. W. Nelson, and R. Myers, Interaction or eradicant and protectant treatments upon the epidemiology and control of mummy berry disease of highbush blueberry, *Phytopathology*, 66:350–354, 1976; D. P. Weingartner and E. J. Klos, Etiology and symptomatology of canker and dieback diseases on highbush blueberries caused by *Godronia (Fusicoccum) cassandrae* and *Diaporthe (Phomopsis) vaccinii, Phytopathology*, 65:105–110, 1975.

Bluefish

A predatory and voracious species of fish that ranges throughout the tropical and temperate seas of the world, except for the eastern and central Pacific areas. A single species, *Pomatomus saltatrix* (see **illus.**), makes up the family Pomatomidae. This fish,

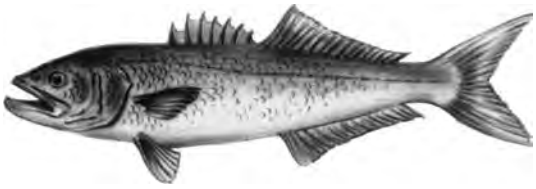

The bluefish (*Pomatomus saltatrix*).

also known as the skipjack, is bluish-gray with an average length of 3 ft (1 m) and weight of about 5 lb (2.25 kg). The mouth is large with sharp, strong teeth. The bluefish form schools and migrate north along the Atlantic coast, following schools of smaller fish upon which they prey. The bluefish continue to kill and destroy their prey even after feeding. About June they reach the New England coast, where the young can be found in estuaries and bays. *See* PERCIFORMES.

Charles B. Curtin

Bluegrass

Grass of genus *Poa* (also called meadowgrass), of the family Graminae (Poaceae). About 50 species are natural to the United States, and 4 are of economic importance: Kentucky bluegrass (*P. pratensis*), used in lawns and pasture; Canada bluegrass (*P. compressa*), used for erosion control; roughstalk bluegrass (*P. trivialis*), adapted for turf in cool wet shade; and annual bluegrass (*P. annua*), a weed of cool moist sites. Kentucky bluegrass was introduced from Europe to Illinois by French missionaries, and spread rapidly throughout the Midwest, thriving where mean July temperatures are below 68°F (20°C). *See* GRASS CROPS; LAWN AND TURF GRASSES.

Kentucky bluegrass (Junegrass) provides nutritious spring pasture, but tends to semidormancy and sparse pasture in summer. Replaced by other grasses for pasture, it is now valued as grass for lawns in temperate North America. Bluegrass lawn turf is planted during fall or spring from sod or seed. Seed is sown at about 2 lb/1000 ft^2 (1 kilogram/are). In fertile soil, plants spread by rhizomes to form a dense sod, favoring a soil of pH 5.8–6.8, good levels of phosphorus and nitrogen, and a mowing of not lower than 2 in. (5 cm) with a sharp blade. Fertilizer needs depend on soil, length of growing season, and whether clippings are removed or left so nutrients are recycled. Irrigation is needed in regions of low rainfall or during periods of drought.

Kentucky bluegrass is a facultative apomict, with most embryos arising from maternal tissue of the ovule instead of a fertilized egg. Chromosome numbers range from the twenties to the hundreds. As apomictic plants reproduce true from seed, single plant selections can be rapidly increased and seed sold as named varieties. Merion, the first of these varieties, was released about 1950. Seed production from Midwest pastures then moved to rogued fields in Oregon, Washington, and Idaho for production of certified varieties of bluegrass seed as a major crop. Today, many varieties are offered, selected for high seed yield, disease resistance, growth habit, and so on. Packaged lawn seed is often a blend of several bluegrass varieties, avoiding weaknesses of any single variety, or a mixture of bluegrasses with other grasses, such as red fescue or perennial ryegrass.

Several diseases that injure bluegrass include mildew in shade (caused by *Erysiphe graminis*), rusts (caused by *Puccinia* spp.), smuts (caused by *Ustilago* spp.), *Fusarium* blight, *Rhizoctonia* brown patch, and *Helminthosporium* diseases. Some 15 other fungi cause minor diseases. Disease is subject to predisposing factors such as overcast skies, short mowing, herbicide use, overfertilizing, and hot humid weather. *See* PLANT PATHOLOGY.

Insects that may reach harmful levels include moth caterpillars (such as lawn moth, cutworms), beetle grubs (such as Japanese beetle or June beetle), bugs (such as chinch bug or bluegrass bill bug), fly maggots (frit fly), and aphids. Weeds occur, especially if lawns are neglected, overwatered, or mowed too short. The most difficult weeds to control are other grasses.

John H. Madison

Bibliography. H. T. Hartmann, W. J. Flocker, and A. M. Kofranek, *Plant Science*, 1988; J. H. Madison, *Principles of Turfgrass Culture*, 1983.

Blueschist

Regional metamorphism with the highest pressures and lowest temperatures, commonly above 5 kilobars (500 megapascals) and below 750°F (400°C). Metamorphic rocks of the relatively uncommon blueschist facies contain assemblages of minerals that record these high pressures and low temperatures. The name "blueschist" derives from the fact that at this metamorphic grade, rocks of ordinary basaltic composition are often bluish because they contain the sodium-bearing blue amphiboles glaucophane or crossite rather than the calcium-bearing green or black amphiboles actinolite or hornblende, which are developed in the more common greenschist- or amphibolite-facies metamorphism. This difference in amphiboles in metamorphosed basalts led J. P. Smith in 1906 to conclude that glaucophane-bearing metamorphosed basalts had experienced different temperature and pressure conditions from those of other metamorphic rocks. In 1939 P. Eskola proposed a glaucophane-schist facies of regional metamorphism. More recently the term "blueschist facies" has been preferred because glaucophane is not always developed in many important rock types, for example, graywacke sandstones, which may contain other minerals that are characteristic of high pressure and low temperature such as aragonite, lawsonite, and jadeite + quartz. "Blueschist facies" is more generally applicable than "glaucophane-schist facies." *See* GLAUCOPHANE.

Blueschist metamorphism developed almost exclusively in young Mesozoic and Cenozoic mountain belts. As shown in **Fig. 1**, blueschist-facies metamorphism has taken place primarily during the last 5 to 10% of geologic time, an observation of considerable importance in understanding the evolution of the Earth. During most of geologic time, high-pressure metamorphism has been dominated by high-temperature amphibolite-facies metamorphism as well as by very high-temperature granulite-facies metamorphism, the latter being uncommon during

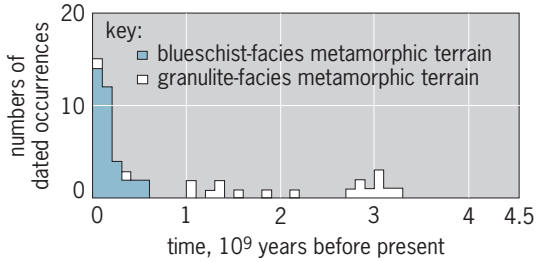

Fig. 1. Histogram showing the occurrence of blueschist-facies and granulite-facies metamorphism as a function of geologic time.

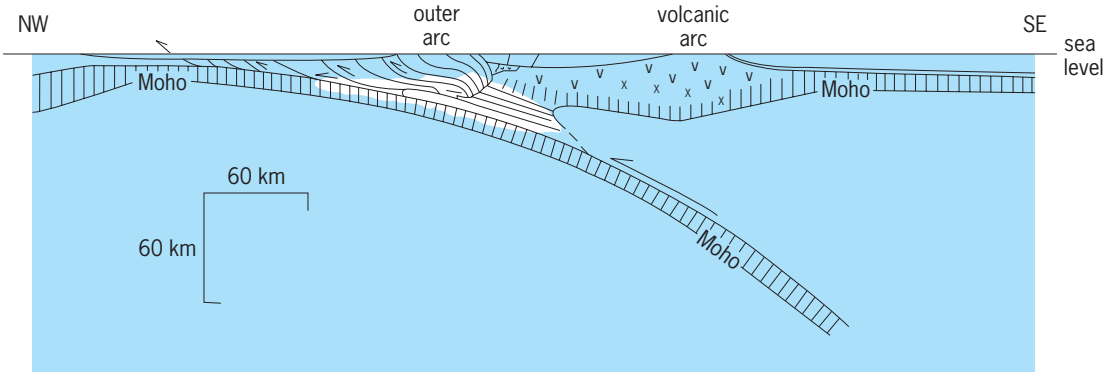

Fig. 2. Cross section of typical island arc and subduction zone with inferred region of blueschist metamorphism (white area). 1 km = 0.6 mi.

the last 5 to 10% of geologic time (Fig. 1). These phenomena are generally interpreted as being in some way an effect of the cooling of the Earth.

Mineral assemblages and physical conditions. A variety of mineral assemblages are characteristic of blueschist-facies metamorphism, depending upon the chemical composition of the rock and the actual temperature and pressure within the field of blueschist metamorphism. Detailed consideration of these assemblages and their mineral composition allows the determination of the physical conditions of metamorphism. Rocks of basaltic composition near the middle of the field of blueschist metamorphism commonly contain glaucophane + lawsonite + phengitic (silica-rich) white mica + sphene + aragonite or calcite; with increasing temperature, epidote, almandine garnet, actinolite, and rutile may appear in glaucophane-bearing assemblages; and at the very highest temperatures the sodic pyroxene omphacite or the sodic hornblende barrowisite may be present. At the lowest grades of blueschist metamorphism, blue amphibole may be missing entirely in metamorphosed basalts, with albite + pumpellyite + chlorite mineral assemblages being present. A common mineral assemblage in metamorphosed graywacke sandstone is albite + lawsonite + phengitic white mica + quartz + aragonite. At higher pressures jadeitic pyroxene or glaucophane crystallizes, and at lower pressures and temperatures pumpellyite may be present instead of lawsonite. Epidote may appear at higher temperatures. *See* EPIDOTE; GRAYWACKE.

The physical conditions of blueschist metamorphism are most easily determined from the phase relations of calcium carbonate ($CaCO_3$) and plagioclase feldspar. With increasing pressure, calcite, the trigonal polymorph of $CaCO_3$, breaks down to form aragonite, the more dense orthorhombic polymorph.

Oxygen isotope geothermometry, based on the variation in coexisting minerals of $^{18}O^{16}/O$ ratios with temperature, yields temperatures of about 570°F (300°C) for typical blueschists from western California. Higher-temperature garnet- and epidote-bearing blueschists yield temperatures of 750–930°F (400–500°C). *See* GEOLOGIC THERMOMETRY.

Tectonics. Blueschist metamorphic rocks are found almost exclusively in the young mountain belts of the circum-Pacific and Alpine-Himalayan chains. The rocks are usually metamorphosed oceanic sediments and basaltic oceanic crust. Previously continental rocks rarely exhibit blueschist metamorphism. The tectonic mechanism for blueschist metamorphism must move the rocks to depths of more than 6 to 12 mi (10 to 20 km) while maintaining relatively cool temperatures (390–750°F or 200–400°C). These temperatures are much cooler than for continental crust at those depths. Heat flow measurements above long-lived subduction zones, together with thermal models, suggest that the conditions of blueschist metamorphism exist today above subduction zones just landward of deep-sea trenches (**Fig. 2**). This tectonic setting at the time of blueschist metamorphism is independently inferred for a number of metamorphic terranes.

What is not well understood is how the blueschist metamorphic rocks return to the surface; clearly the mechanism is not simple uplift and erosion of 12–18 mi (20–30 km) of the Earth's crust. Blueschist metamorphic rocks are usually in immediate fault contact with much less metamorphosed or unmetamorphosed sediments, indicating they have been tectonically displaced relative to their surroundings since metamorphism. *See* METAMORPHIC ROCKS; METAMORPHISM; PLATE TECTONICS. John Suppe

Bibliography. A. Miyashiro, *Metamorphism and Metamorphic Belts*, 1973; A. R. Philpotts, *Principles of Igneous and Metamorphic Petrology*, 1990; J. Suppe, *Principles of Structural Geology*, 1985.

Bluestem grass

The common generic name often applied to the genera *Andropogon, Dichanthium, Bothriochloa*, and *Schizachyrium* in the grass tribe Andropogoneae. Photosynthesis in this entire tribe follows the C_4 pathway characterized by high activity of the NADP-malic enzyme, providing adaptation to environments with high light intensities and high temperatures during the growing season. The bluestems are medium

to tall, warm-season, perennial grasses with basic chromosome numbers of 9 or 10, often exhibiting polyploidy. Although asexual reproduction is common, reproduction is predominantly sexual, and ecotypic variation occurs in wide-ranging species. Prescribed burning is an important management tool, especially in subhumid and humid regions. Because of the bluestems' long season of growth, closely synchronized with decomposition rate, they usually compete successfully against cool-season grasses on soils that are low in fertility. The two most important forage species are big bluestem (*A. gerardi*) and little bluestem (*S. scoparius*). *See* CYPERALES; PHOTOSYNTHESIS.

Big bluestem. This is a tall (usually 3–6 ft or 1–2 m), deep-rooted grass with strong rhizomes; it occurs throughout the continental United States (except in the extreme western states), in southern Canada, and in northern Mexico. It is important in the true prairie, the savannas and forest openings that surround the central grassland, and even in the semiarid portion of the mixed prairie, wherever sufficient moisture compensation occurs due to subirrigation, run-in water, drifting snow, north-facing slopes, coarse-textured soil, or rock in the soil profile.

Big bluestem is one of the most palatable of all grasses when it is actively growing, but its nutritional value declines sharply with maturity. Because of its high palatability and tall growth habit, big bluestem has been reduced or eliminated by heavy grazing over much of its range. Nevertheless, it is quite grazing-resistant due to the presence of rhizomes, its large ratio of vegetative to reproductive shoots, its relatively large root-shoot ratio, and its rapid regrowth potential. Improved varieties are seeded as monocultures or in simple mixtures for warm-season grazing or hay and may be profitable for seed production, especially when grown in rows and fertilized with nitrogen. Big bluestem is used by many species of wildlife and is important as prairie chicken habitat. Along with little bluestem, it is responsible for much of the beauty of the prairies, especially late in the fall when the reddish-purple color of the foliage is fully developed.

Little bluestem. This is a medium-height bunchgrass (usually 12–24 in. or 30–60 cm, but taller in the south) which occurs in the same geographic area as big bluestem but extends farther north in the prairie provinces of Canada to the southern Yukon; it is more abundant on relatively dry sites, especially on weakly developed soils. It is an excellent grass for grazing and hay when actively growing, but nutritional value declines rapidly with maturity. Although old seedstalks are avoided, new shoots around the edge of the plant are usually grazed. When drought, heavy grazing, or fire removes the old growth, new herbage is grazed readily. Little bluestem is not very grazing-resistant because of the small ratio of vegetative to reproductive shoots and is often eliminated by heavy grazing before big bluestem. Improved varieties are available for seeding. This is the grass that provides the major aspect of the Flint Hills of Kansas and the Osage Hills of Oklahoma. *See* GRASS CROPS.

James K. Lewis

Bluetongue

An arthropod-borne disease of ruminant species. Its geographic distribution is dependent upon a susceptible ruminant population and climatic conditions that favor breeding of the primary vector, a mosquito (*Culicoides* species).

Virus. Blue tongue virus is the prototype of the *Orbivirus* genus (family Reoviridae). The viral genome exists as 10 segments of the double-stranded ribonucleic acid (RNA) that encode for seven structural and three nonstructural proteins. The viral particle has a double capsid, with the outer coat (morphologically poorly defined) being composed of two proteins. Twenty-four serotypes of blue tongue virus have been defined, and their distribution throughout the world is varied. *See* ANIMAL VIRUS.

Clinical disease. While multiple ruminant species can become infected, only sheep and deer typically display clinical blue tongue disease. Infection of sheep can result in signs ranging from subclinical to acute, with the latter being characterized by high fever, excessive salivation, nasal discharge, hyperemia (buccal and nasal mucosa, skin, coronet band), and erosions and ulcerations of mucosal surfaces in the mouth. Severity of the disease is dependent upon multiple factors, including virus strain, animal breed, and environmental conditions. Upon infection by a gnat bite, the virus apparently replicates in the local lymphatic system prior to the viral particles moving into the blood (viremia). Viral replication occurs in the endothelial cells of small vessels, resulting in narrowing of the vessel, release of proteinaceous material into the surrounding tissues, and possibly hemorrhage, with the respiratory tract, mucous membranes, cardiac and skeletal musculature, and skin being most affected. Animals experiencing acute clinical symptoms typically die from pneumonia or pulmonary failure; hemorrhage at the base of the pulmonary artery indicates the presence of a vascular lesion.

Blue tongue infection of cattle and goats is common and typically subclinical. Cattle tend to have a prolonged viremia (up to 100 days) compared with sheep (7–35 days). This extended viremia facilitates infection of numerous vectors and is thus important epidemiologically. The few reported cases of clinical blue tongue infection in cattle have been hypothesized to be the result of an allergic response. Affected cattle may develop lesions associated with mucosal surfaces in the mouth and nose, and dermatitis and thickening of skin in areas of the neck.

Blue tongue virus can also cross the placenta of pregnant animals, with the outcome being dependent upon gestational age. Infection during the first trimester may result in fetal death and resorption. As gestational age progresses, infection may result in abortion or birth of animals with hydranencephaly

(a cystic brain lesion). Late-stage infection results in the birth of apparently healthy, viremic animals. While wild-type virus results in fetal disease if inoculated directly into the fetus, cell culture–adapted blue tongue virus demonstrates the greatest ability to cross the placenta. Thus, administration of attenuated blue tongue virus vaccines (the only vaccines available) to pregnant animals should be avoided.

Control. Control of blue tongue disease requires the application of vaccines and modulation of the farm environment. While blue tongue virus vaccines are available, efficacy is often incomplete and variable, in part because of the multiplicity of serotypes active throughout the world and limited cross-serotype protection. Furthermore, use of polyvalent (multiple-serotype) vaccines in the United States has been discouraged because of potential genetic reassortment between vaccine viruses and wild-type viruses, a process that could possibly lead to pathogenic variants. Relative to environment, elimination of vector breeding sites can also facilitate control of virus transmission. With the multiplicity of serotypes typically active in an endemic area, and the minimal cross-serotype protection observed, administration of vaccine in the face of an outbreak may be of limited value. *See* VACCINATION.

Because blue tongue disease is vector borne, efforts to reduce *Culicoides* breeding habitats will minimize viral transmission. Housing sheep under covered facilities will also reduce the exposure to the virus as *Culicoides* has a tendency to avoid entry into enclosed structures. Application of insect repellents may also reduce exposure. A specific treatment for blue tongue is not available; therefore, practicing good animal husbandry techniques (reducing animal stress and maintaining a warm and dry environment) and treating secondary bacterial infections with antibiotics are suggested. Jeffrey L. Stott

Boat propulsion

The action of propelling a boat through water. A boat machinery plant consists principally of a propulsion engine, propulsor (propeller or jet pump), and driveline components. The engines are almost exclusively of the familiar internal combustion types: gasoline, diesel, or gas turbine. The gasoline engine has traditionally dominated the pleasure-boat field, while the diesel is favored for commercial and military craft. The gas turbine is comparatively rare and is found only in applications where high power from machinery of small weight and volume is essential. *See* PROPELLER (MARINE CRAFT).

Auxiliary items, such as bilge pumps, domestic water pumps and heaters, and electric generators and switchboards typically are found in the machinery plants of the larger boats. The generator (or alternator) is often driven by a belt from the propulsion engine, as in automotive practice, but is sometimes driven by a separate engine

Gasoline engines. The marine gasoline engine appears in inboard and outboard forms.

Outboard engines. The outboard engine is a unit assembly of engine, propeller, and vertical drive shaft that is usually clamped to the boat transom. It is traditionally the power plant for the smallest motor boats, or an auxiliary for rowboats and sailboats; powers as low as 1 hp are built. However, while a 50-hp (37-kW) engine driving a high-speed runabout is the most typical application, engines with power ratings as high as 300 hp (224 kW) are available. Most models use the two-stroke-cycle engine. The smaller sizes offer easy portability and adaptability to a variety of craft. The size and weight of the larger models somewhat cancel these advantages, though weight is significantly less than that of an inboard engine of the same power, and the outboard location has appeal because it frees usable space within the boat.

A cutaway view (**Fig. 1**) illustrates some of the construction features of outboard engines. The cut is taken through the engine centerline, so that pistons, flywheel, crankshaft, spark plugs, extension shaft to propeller, and gearing are visible in section. The flywheel is the cup-shaped gray component just under the top cover; the crankshaft and extension shaft run in direct line to the bevel gearing that turns the

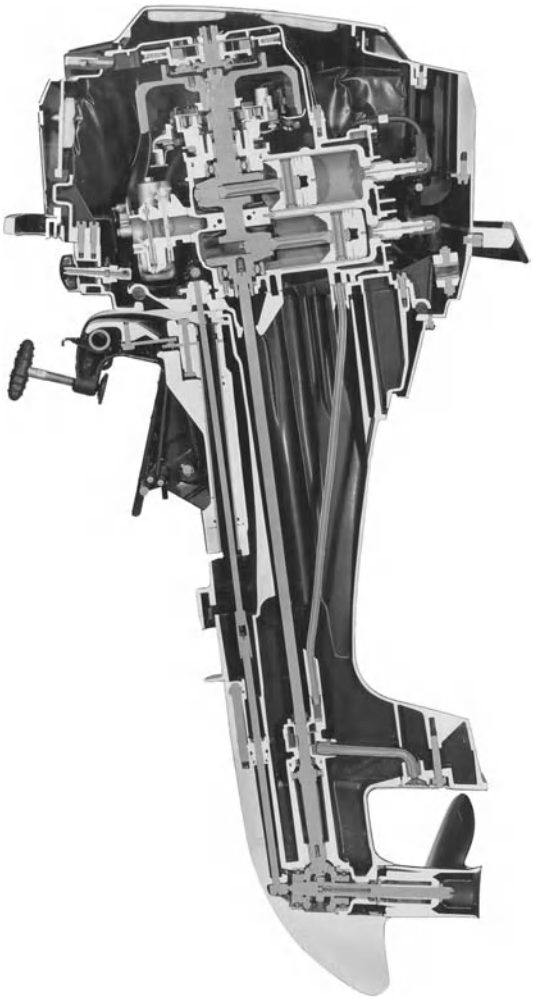

Fig. 1. Typical outboard engine. (*Kiekhaefer Mercury*)

shafting through almost 90° to the propeller. The forward-neutral-reverse actuating rod to the gearing is to the left of the shaft, and the cooling-water passage to the engine water jacket is visible to the right. Pistons, spark plugs, cylinder and crankcase walls, and part of the carburetor (left of the shaft) are shown in the figure. Engine operation is characterized by magneto ignition and by crankcase compression of the fuel-air mixture for scavenging purposes, as is typical of small two-cycle gasoline engines. Its unique marine features are principally the arrangement for mounting on the boat transom and its cooling system.

Inboard engines. The inboard form is almost exclusively adapted from one of the mass-produced automotive engines. Inboard engines are compact, light, lower in first cost than any competitor, and familiar to untrained users, and so they predominate among pleasure boats and the smaller fishing boats. **Figure 2** shows an engine of 140 hp coupled to an inboard-outboard drive unit.

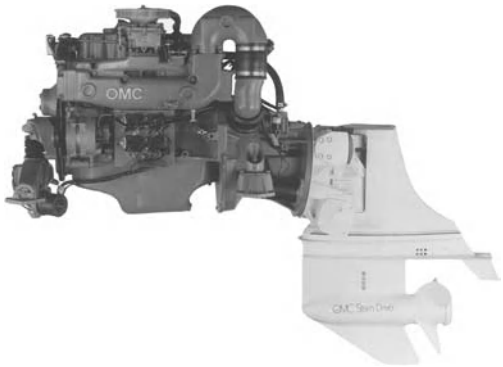

Fig. 2. Typical inboard gasoline engine coupled to an inboard-outboard drive unit. (*Outboard Marine Corp.*)

Engine adaptation to marine service is largely confined to externals. The cooling system is converted to use the surrounding water, either directly in the coolant passages or indirectly via a heat exchanger. Components that might be degraded rapidly by seawater corrosion are replaced with resistant items. Risers are fitted to the exhaust manifold to prevent water from backing up into the engine.

The most important changes, however, are dictated by safety considerations. Gasoline vapor that would be swept away by the wind in automotive use is trapped around the engine by the surrounding hull envelope. This vapor is a lethal hazard, since it can ignite explosively from a stray spark or contact with the hot exhaust manifold. For this reason, the exhaust manifold must be cooled by a water jacket, the engine air intake must be covered by a backfire trap, and the carburetor and fuel system must be designed to minimize the posibility of gasoline leakage. Also, the engine compartment must be well ventilated to promote the dissipation of vapors.

Diesel engine. The diesel engine is generally higher in first cost than the gasoline engine and is somewhat heavier for the same power, but it consumes less fuel. Fuel savings make the diesel attractive if the engine is to be used more than a few hundred hours a year

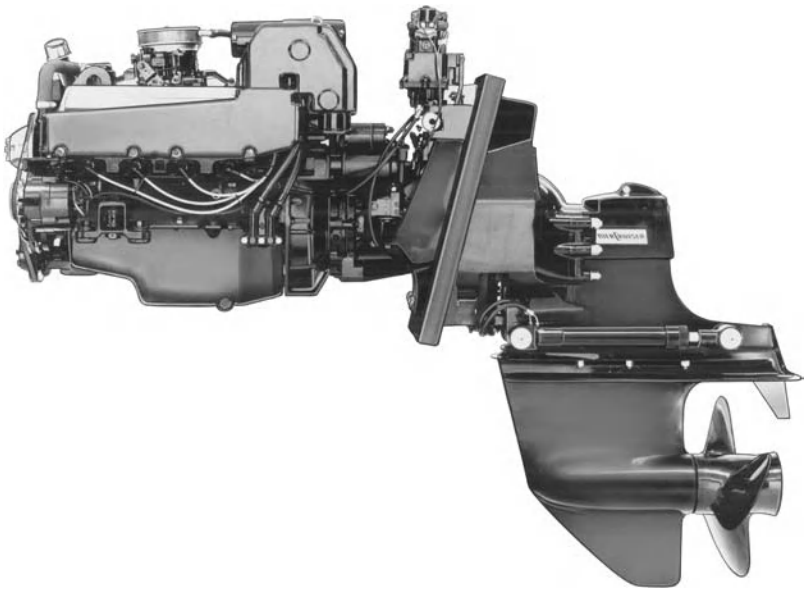

Fig. 3. Jet propulsion drive unit coupled to an inboard gasoline engine. (*Berkeley Pump Co.*)

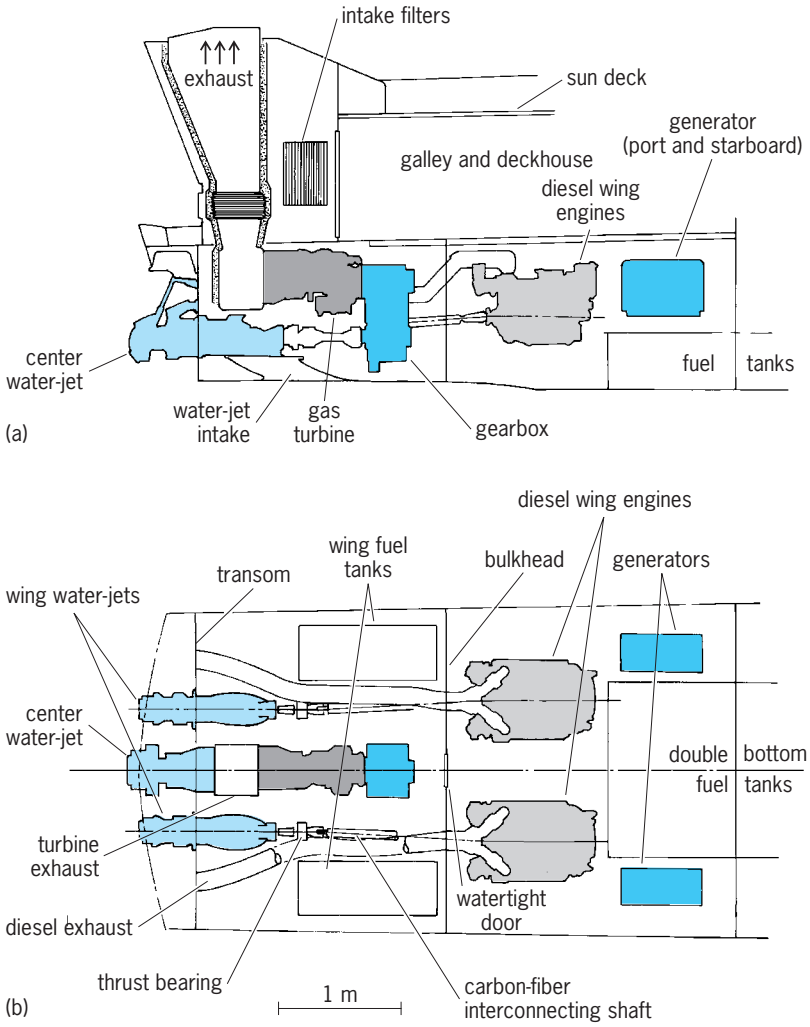

Fig. 4. Machinery layout for twin diesel engines combined with a gas turbine. (*a*) Side view. (*b*) Top view. (*Shead Design*)

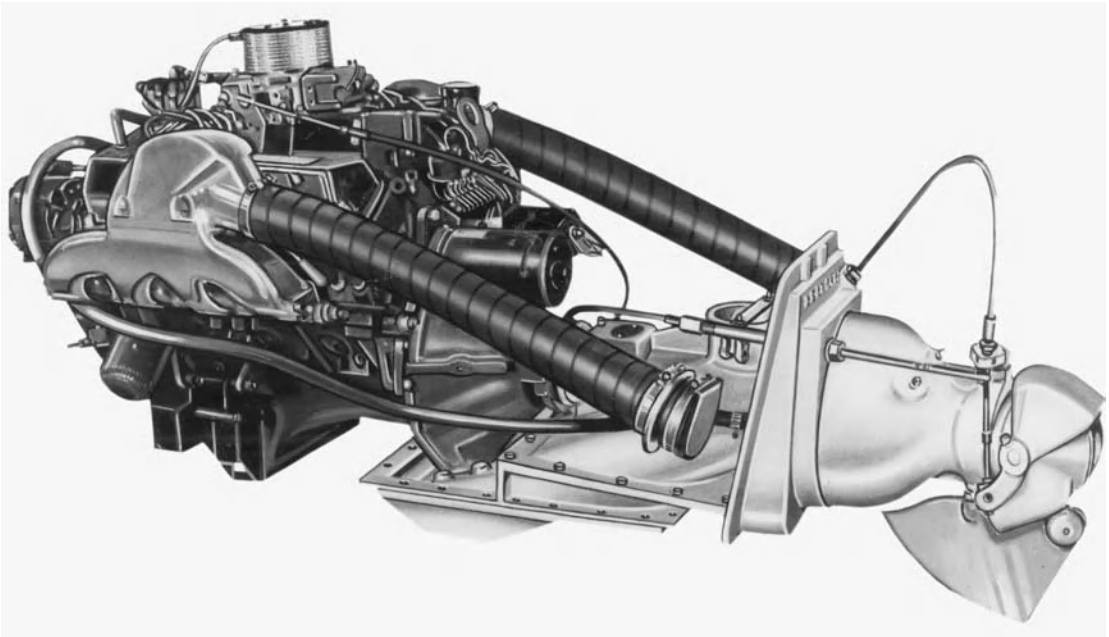

Fig. 5. Typical marine diesel engine. (*Detroit Diesel Allison, General Motors Corp.*)

or if the boat must have a long cruising range. Its principal market is thus in commercial and military craft, but some of the more compact models are used in pleasure craft.

The small-boat diesel is also almost exclusively adapted from the automotive industry, usually from a truck engine (**Fig. 3**). Both four-stroke and two-stroke models are popular. The conversion to marine use is similar to that for the gasoline engine, except that the precautions against explosion are not necessary because of the relatively high flash point of diesel fuel. *See* DIESEL ENGINE.

Gas turbine. The gas turbine engine is comparatively light and compact, making it attractive for high-speed, high-power boats. Its disadvantages are high first cost, high fuel consumption, and large exhaust and intake ducts. The last factor is due to its high rate of air consumption; the need for large volumes of air also makes it difficult, in a small vessel, to keep spray from being drawn into the engine, with consequent fouling and corrosion.

The small-craft gas turbine is typically adapted from a light-aircraft jet engine. Usually a second-stage turbine is added to the engine to drive the output shaft independently of the compressor. The combustors are modified to burn marine diesel fuel, and corrosion-resistant parts are substituted. **Figure 4** shows the propulsion machinery of a high-speed yacht. Two diesels of 1065 hp (794 kW) provide power for cruising at high speed; a single gas turbine of 4600 hp (3430 kW) provides power for sprinting at extreme speeds. *See* GAS TURBINE.

Drive systems. All of the engines usually require reduction gears to match their speeds to an efficient propeller speed. The gear boxes reduce shaft speed by factors of 1.5 up to perhaps 5, are mounted directly on the engine flywheel housing of the inboard engines, and incorporate reverse gears. A reduction gear of this type shows at the left end of the engine in Fig. 3. Speed reduction for the outboard engine occurs in the bevel gearing in its lower unit.

Inboard-outboard drives are frequently used with inboard engines, as shown in Fig. 2. (The lighter end is mounted outboard of the vessel stern.) These units contain the reduction gears, may be turned so that a rudder is not needed, and provide a path to an underwater exhaust, all similar to the underwater portion of outboard engines.

A water-jet pump sometimes replaces the propeller. Although it tends to be less efficient than a propeller and constitutes a bulky component to be accommodated within the boat, a water-jet pump is invulnerable to damage from underwater hazards, and is safe to operate in the vicinity of swimmers. **Figure 5** shows a typical jet pump coupled to an inboard engine. This unit includes a rudder and a movable gate by which the jet is redirected for reverse thrust. *See* INTERNAL COMBUSTION ENGINE; MARINE ENGINE; MARINE MACHINERY. John B. Woodward

Bibliography. P. Bowyer, *Diesel Boat Engines*, 1989; N. Calder, *Marine Diesel Engines*: *Maintenance, Troubleshooting and Repair*, 2d ed., 1991; L. Goring, *Marine Inboard Engines*: *Petrol and Diesel*, 1990; H. Stocker, *Marine Engines for Recreational Boats*, 1992; J. B. Woodward, *Low Speed Marine Diesel Engines*, 1981, reprint 1988.

Bog

Nutrient-poor, acid peatlands with a vegetation in which peat mosses (*Sphagnum* spp.), ericaceous dwarf shrubs, and to a lesser extent, various sedges (Cyperaceae) play a prominent role. The terms muskeg, moor, heath, and moss are used locally to indicate these sites. *See* MUSKEG.

Distribution. Bogs are most abundant in the Northern Hemisphere, especially in a broad belt including the northern part of the deciduous forest zone and the central and southern parts of the boreal forest zone. Farther south, and in drier climates farther inland, they become sporadic and restricted to specialized habitats. To the north, peatlands controlled by mineral soil water (aapa mires) replace them as the dominant wetlands.

Bogs are much less extensive in the Southern Hemisphere because there is little land in cold temperate latitudes. However, they do occur commonly on the South Island of New Zealand, in Tasmania, in southern Patagonia, and on islands in the temperate parts of the South Atlantic and South Pacific Oceans. In these Southern Hemisphere peatlands, *Sphagnum* is much less important, and Epacridaceae and Restionaceae replace the Ericaceae and Cyperaceae of the Northern Hemisphere.

Properties. Bogs have a fibric, poorly decomposed peat consisting primarily of the remains of *Sphagnum*. Decomposition is so incomplete that plant tissues, including those of mosses, are often still identifiable at a depth of more than 10 ft (3 m) and after thousands of years. Consequently, the vegetation history of bogs can be reconstructed by using time scales based on ^{14}C or other dating techniques. Well-preserved logs and bodies of animals and human beings have also been excavated from bogs. Pollen trapped and preserved in peat are commonly used to determine the vegetation history of surrounding upland areas. *See* HUMUS; PALYNOLOGY; PEAT.

Peat accumulation is the result of an excess of production over decomposition. Obviously, the very presence of bogs shows that production exceeded decay over the entire period of bog formation. However, in any given bog present production can exceed, equal, or be less than decomposition, depending on whether it is actively developing, in equilibrium, or eroding. In most bogs, production and decomposition appear to be in equilibrium at present.

Slow decay rather than high productivity causes the accumulation of peat. Decomposition of organic matter in peat bogs is slow due to the high water table, which causes the absence of oxygen in most of the peat mass, and to the low fertility of the peat. Bogs, in contrast to other peatlands, can accumulate organic matter far above the ground water table. This occurs only to any significant degree in ombrotrophic (rain-fed) bogs, the most nutrient-deficient bogs, which receive all their nutrients and water supply from the atmosphere. The level to which these bogs rise above the surroundings depends on the wetness of the climate and the size of the bog. In temperate, humid climates, the center of large bogs can be elevated more than 25 ft (8 m) above the mineral soil level in the bog border.

Permeability of peat decreases rapidly with depth, and within 1–2 ft (0.3–0.6 m) below the surface the hydraulic conductivity is so low that it can be considered impermeable for all practical purposes. As a result, almost all water movement takes place through the surface peat or over the surface. When peat accumulates above the ground water, a perched water table develops which is maintained by precipitation. In contrast to popular belief, the wetness of raised bogs is maintained by this perched water table and not by capillary water.

Types of bogs. Bogs are vegetation types as well as landforms. As vegetation types, bogs are often distinguished from fens. Bogs are ombrotrophic peatlands dependent on the atmosphere for their water and nutrient supply, whereas fens are minerotrophic peatlands influenced by water which has been in contact with the mineral soil. Hydrologically and ecologically, this is an important distinction. This restriction of the term bog to ombrotrophic peatlands is floristically a less fortunate choice because weakly minerotrophic fens have a floristic composition which is very similar to ombrotrophic peatlands and have little in common with nutrient-rich fens. The word bog is generally also used for all *Sphagnum*-dominated fens and swamps; it is used in this sense here.

As a landform, peatlands are complexes of plant communities, and all ombrotrophic peatlands include minerotrophic areas, at least in their margins. Therefore, the distinction between ombrotrophic and minerotrophic peatlands has no value here. A practical subdivision of peatlands as landforms can be based on the nature of the water which controls their development: ombrogenous, topogenous, soligenous, and limnogenous peatlands are distinguished (see **illus.**). Ombrogenous peatlands develop under the influence of rainfall and are restricted to humid, temperate climates. Topogenous peatlands develop in topographical positions with

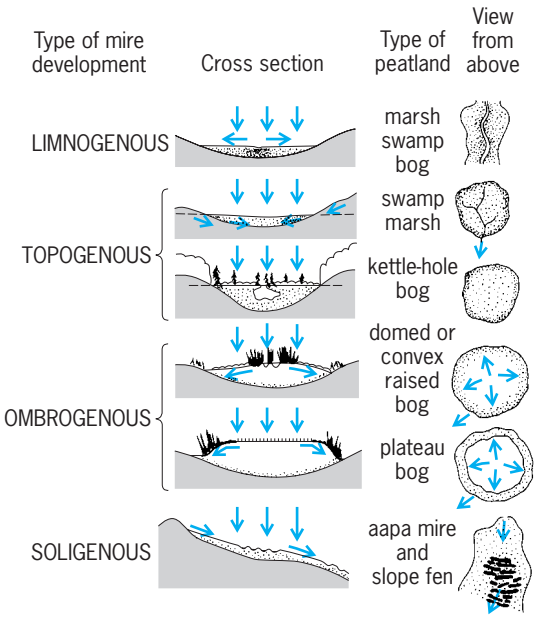

Relationship between source of water and major types of peatland development. Stippling indicates parts affected by mineral soil water; density of stippling shows relative fertility of the water. Arrows indicate source of water and major flow through peatlands. Note differences in direction of water movement.

a permanent ground-water table; they can occur under a wide range of climatic conditions but are most common in humid climates. Soligenous peatlands depend on a reliable source of minerotrophic seepage water; they are characterized by moving water, and are most common in regions with a high surplus of precipitation over evapotranspiration. Given a reliable water supply, they can also occur in warm climates. Limnogenous peatlands develop along lakes and slow-flowing streams. As landforms, bogs include all ombrogenous peatlands and some topogenous and limnogenous peatlands influenced by nutrient-poor water.

Geographic variation. Not surprisingly for a habitat which occupies such an extensive range, bogs show large geographic differences in floristic composition, surface morphology, and development. Blanket bogs, plateau bogs, domed bogs, and flat bogs represent a series of bog types with decreasing climatic humidity. Concentric patterns of pools and strings (peat dams) become more common and better developed northward. Continental bogs are often forest-covered, whereas oceanic bogs are dominated by dwarf shrub heaths and sedge lawns, with forests restricted to the bog slope if the climate is not too severe.

Uses. Bogs have long been used as a source of fuel. In Ireland and other parts of western Europe, the harvesting of peat for domestic fuel and reclamation for agriculture and forestry have affected most of the peatlands, and few undisturbed bogs are left. Other uses are for horticultural peat, air layering in greenhouses, litter for poultry and livestock, and various chemical and pharmaceutical purposes. Mechanical extraction of peat for horticultural purposes has affected large bog areas worldwide. The oil crisis stimulated an interest in peat bogs for generating electricity and gas. *See* BIOMASS; SWAMP, MARSH, AND BOG.

Antoni W. H. Damman

Bibliography. A. J. P. Gore (ed.), *Mires: Swamp, Bog, Fen and Moor*, 2 vols., 1983; H. A. P. Ingram, Size and shape in raised mire ecosystems: A geographical model, *Nature* (London), 297:300–303, 1982; K. E. Ivanov, *Water Movement in Mirelands*, 1981; P. D. Moore and D. J. Bellamy, *Peatlands*, 1974.

Bohrium

A chemical element, symbol Bh, atomic number 107. Bohrium was synthesized and identified in 1981 by using the Universal Linear Accelerator (UNILAC) of the Gesellschaft für Schwerionenforschung (GSI) at Darmstadt, West Germany, by a team led by P. Armbruster and G. Müzenberg. The reaction used to produce the element was proposed and applied in 1976 by Y. T. Oganessian and colleagues at Dubna Laboratories in Russia. A ^{209}Bi target was bombarded by a beam of ^{54}Cr projectiles.

The best technique to identify a new isotope is its genetic correlation to known isotopes through a radioactive decay chain. These decay chains are generally interrupted by spontaneous fission. In order

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

to apply decay chain analysis, those isotopes that are most stable against spontaneous fission should be produced, that is, isotopes with odd numbers of protons and neutrons. Not only does the fission barrier govern the spontaneous fission of a species produced, but also, in the deexcitation of the virgin nucleus, fission competing with neutron emission determines the final production probability. To keep the fission losses small, a nucleus should be produced with the minimum excitation energy possible. In this regard, reactions using relatively symmetric collision partners and strongly bound closed-shell nuclei, such as ^{209}Bi and ^{208}Pb as targets and ^{48}Ca and ^{50}Ti as projectiles, are advantageous.

Six decay chains were found in the Darmstadt experiment. All the decays can be attributed to ^{262}Bh, an odd nucleus produced in a one-neutron reaction. The isotope ^{262}Bh undergoes alpha-particle decay (10.38 MeV) with a half-life of about 5 ms.

Experiments at Dubna, performed in 1983 using the 157-in. (400-cm) cyclotron, established the production of ^{262}Bh in the reaction ^{209}Bi ^{54}Cr. *See* NUCLEAR FISSION; NUCLEAR REACTION; PERIODIC TABLE; RADIOACTIVITY; TRANSURANIUM ELEMENTS.

Peter Armbruster

Bibliography. S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; G. Münzenberg et al., Identification of element 107 by alpha correlations chains, *Z. Phys.*, A300:107–108, 1981; G. Münzenberg et al., The velocity filter SHIP, a separator of unsowed heavy-ion fusion products, *Nucl. Instrum. Methods*, 161:65–82, 1979.

Boiler

A pressurized system in which water is vaporized to steam, the desired end product, by heat transferred from a source of higher temperature, usually the products of combustion from burning fuels. Steam thus generated may be used directly as a heating medium, or as the working fluid in a prime mover to convert thermal energy to mechanical work, which in turn may be converted to electrical energy. Although other fluids are sometimes used for these purposes, water is by far the most common because of its economy and suitable thermodynamic characteristics.

The physical sizes of boilers range from small portable or shop-assembled units to installations comparable to a multistory 200-ft-high (60-m) building equipped, typically, with a furnace which can burn coal at a rate of 6 tons/min (90 kg/s). In terms of steam production capacities, commercial boilers range from a few hundred pounds of steam per hour to more than 6,000,000 lb/h (750 kg/s). Pressures range from 0.5 lb/in.^{2} (3.4 kilopascals) for domestic space heating units to 5000 lb/in.^{2} (34 megapascals) for some power boilers. The latter type will deliver steam superheated to 1100±°F (593±°C) and reheated to similar values at intermediate pressures. Large units are field-assembled at the installation site but small units (frequently referred to as package boilers) are shop-assembled to minimize the overall boiler price.

Boilers operate at positive pressures and offer the hazardous potential of explosions. Pressure parts must be strong enough to withstand the generated steam pressure and must be maintained at acceptable temperatures, by transfer of heat to the fluid, to prevent loss of strength from overheating or destructive oxidation of the construction materials. The question of safety for design, construction, operation, and maintenance comes under the police power of the state and is supplemented by the requirements of the insurance underwriters. The ASME Boiler Construction Code is the prevalent document setting basic standards in most jurisdictions.

Being in the class of durable goods, boilers that receive proper care in operation and maintenance function satisfactorily for several decades. Thus the types found in service at any time represent a wide span in the stages of development in boiler technology.

The earliest boilers, used at the beginning of the industrial era, were simple vats or cylindrical vessels made of iron or copper plates riveted together and supported over a furnace fired by wood or coal. Connections were made for offtake of steam and for the replenishment of water. Evolution in design for higher pressures and capacities led to the use of steel and to the employment of tubular members in the construction to increase economically the amount of heat-transferring surface per ton of metal. The earliest improvement was the passage of hot gases through tubes submerged in the water space of the vessel, and later, arrangements of multiple water-containing tubes which were exposed on their outer surface to contact with hot gases. *See* FIRE-TUBE BOILER; WATER-TUBE BOILER.

The overall functioning of steam-generating equipment is governed by thermodynamic properties of the working fluid. By the simple addition of heat to water in a closed vessel, vapor is formed which has greater specific volume than the liquid, and can develop increase of pressure to the critical value of 3208 lb/in.^{2} absolute pressure (22.1 MPa absolute pressure). If the generated steam is discharged at a controlled rate, commensurate with the rate of heat addition, the pressure in the vessel can be maintained at any desired value, and thus be held within the limits of safety of the construction. *See* STEAM.

Addition of heat to steam, after its generation, is accompanied by increase of temperature above the saturation value. The higher heat content, or enthalpy, of superheated steam permits it to develop a higher percentage of useful work by expansion through the prime mover, with a resultant gain in efficiency of the power-generating cycle. *See* SUPERHEATER.

If the steam-generating system is maintained at pressures above the critical, by means of a high-pressure feedwater pump, water is converted to a vapor phase of high density equal to that of the water, without the formation of bubbles. Further heat addition causes superheating, with corresponding increase in temperature and enthalpy. The most advanced developments in steam-generating equipment have led to units operating above critical pressure, for example, 3600–5000 lb/in.^{2} (25–34 MPa).

Superheated steam temperature has advanced from 500±°F (260±°C) to the present practical limits of 1050–1100°F (566–593°C). Progress in boiler design and performance has been governed by the continuing development of improved materials for superheater construction having adequate strength and resistance to oxidation for service at elevated temperatures. For the high temperature ranges, complex alloy steels are used in some parts of the assembly.

Steam boilers are built in a wide variety of types and sizes utilizing the widest assortment of heat sources and fuels. *See* BOILER ECONOMIZER; BOILER FEEDWATER; BOILER FEEDWATER REGULATION; MARINE ENGINEERING; NUCLEAR POWER; STEAM-GENERATING UNIT. Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; M. J. Bernstein et al., *Power Boilers: A Guide to Section I of the ASME Boiler and Pressure Vessel Code*, 1998; S. C. Stultz and J. B. Kitto (eds.), *Steam: Its Generation and Use*, 41st ed., 2005.

Boiler economizer

A component of a steam-generating unit that absorbs heat from the products of combustion after they have passed through the steam-generating and superheating sections. The name, accepted through common usage, is indicative of savings in the fuel required to generate steam. *See* BOILER.

An economizer is a forced-flow, once-through, convection heat-transfer device to which feedwater is supplied at a pressure above that in the steam-generating section and at a rate corresponding to the steam output of the unit. The economizer is in effect a feedwater heater, receiving water from the boiler feed pump and delivering it at a higher temperature to the steam generator or boiler. Economizers are used instead of additional steam-generating surface because the feedwater, and consequently the heat-receiving surface, is at a temperature below that corresponding to the saturated steam temperature; thus,

the economizer further lowers the flue gas temperature for additional heat recovery. *See* BOILER FEEDWATER; THERMODYNAMIC CYCLE.

Generally, steel tubes, or steel tubes fitted with externally extended surface, are used for the heat-absorbing section of the economizer; usually, the economizer is coordinated with the steam-generating section and placed within the setting of the unit.

The size of an economizer is governed by economic considerations involving the cost of fuel, the comparative cost and thermal performance of alternate steam-generating or air-heater surface, the feedwater temperature, and the desired exit gas temperature. In many cases it is economical to use both an economizer and an air heater. *See* AIR HEATER; STEAM-GENERATING UNIT.

George W. Kessler

Boiler feedwater

Water supplied to a boiler for the generation of steam. The water is heated to produce steam and acts to cool the heat-absorbing surfaces as it circulates through the boiler. The steam withdrawn from the steam drum is of high purity, and contaminants that enter with the boiler feedwater, even in small concentrations, accumulate in the boiler water.

Contaminants and impurities. Feedwater should be virtually free of impurities that are harmful to the boiler and its associated systems. Generally, natural waters are unsuitable for direct introduction into the boiler because they are contaminated from contact with the earth, the atmosphere, or other sources of pollution. The contaminants can be removed or altered by chemical treatment or other means to provide satisfactory feedwater. *See* RAW WATER; WATER TREATMENT.

The need for and type of feedwater treatment vary considerably in different plants, depending on the nature of the raw water supply and the percentage of makeup water required. The treatment of boiler feedwater is critical for steam-generating boilers operating at high pressures with high rates of heat transfer.

If not removed, the contaminants will interfere severely with boiler operation. They can cause corrosion, adversely affect boiler-water circulation or steam and water separation, or form deposits on the internal surfaces of the heat-absorbing components. Such deposits on the high-pressure parts of the boiler system result in high metal temperatures and hasten eventual failure of the parts.

Most soluble contaminants are readily removed by the blowdown of boiler water. Insoluble contaminants must be chemically treated to change their characteristics and either to ease their removal from the boiler water or to render them innocuous. *See* STEAM-GENERATING UNIT.

Components. Boiler feedwater is usually a mixture of relatively pure condensate and treated makeup water. The proportions vary with the amount of makeup water required to compensate for losses from the system served by the boiler. In a closed power-generating steam-boiler system, losses may be less than 1% of the cycle flow. When steam is used for industrial processes, the return of part or all of the condensate may be impractical. Likewise, in forced flow, once through boilers, the water is not circulated. In these cases, impurities entering with the feedwater must either leave with the steam or be deposited within the various components or on the surfaces traversed by the steam passing through the cycle. Such units require high-purity makeup water, which is either evaporated or demineralized. *See* STEAM CONDENSER.

Processing. Processing of raw water for use as treated makeup water entails the removal or conversion of suspended and dissolved solids that could form scale or insulating deposits on the heat-transfer surfaces or that could be transported with the steam to other parts of the system. Contaminants that may be corrosive under certain conditions existing in the cycle flow must be removed or neutralized. The effectiveness of the treatment of boiler feedwater varies with the methods used and in the care taken in operation, but water can be treated to meet the most exacting requirements dictated by the plant operating conditions.

Insoluble contaminants. For most applications, it is necessary to remove all but trace quantities of the troublesome insoluble contaminants from the feedwater before it is pumped into the boiler. Hardness in the form of calcium and magnesium salts is removed from raw water before it becomes suitable as makeup feedwater. Metallic oxides, mainly those of iron and copper, are frequently removed from condensate returns by a filtering process. Residual trace amounts of these deposit-forming contaminants are inevitably present, and additional contamination may occur from the leakage of raw cooling water into the steam or hot side of heat exchangers (the turbine condenser, for example).

Calcium hardness. At lower operating pressures (below 1500 lb/in.^{2} or 10 megapascals), residual calcium hardness is treated by introducing sodium phosphate to the water. The calcium precipitates as a phosphate sludge which is removed with the boiler-water blowdown before it can deposit on the heat-transfer surfaces. At high operating pressures (above 1500 lb/in.^{2} or 10 MPa), the rate of deposition is so rapid that blowdown is ineffective. In essence, the most effective method for minimizing boiler-water deposits is to reduce the amount of deposit-forming material in the boiler feedwater.

Silica contamination. Silica, which in combination with other contaminants may form an adherent and highly insulating scale, will vaporize at significant and increasing rates if the boiler-water temperature exceeds 500°F (260°C). The concentration of silica must be held within limits by blowdown, with the maximum allowable concentration decreasing as pressure increases. Silica in vapor solution with the steam exiting from the boiler is not arrested by mechanical means. The resulting buildup of silica scale on turbine blades leads to reduced

turbine output and costly outages for removal of the deposits.

Corrosion. Internal corrosion may be initiated by dissolved gases, such as oxygen and carbon dioxide, picked up by water at any point of contact with air. Accordingly, when treating feedwater, the dissolved oxygen and other gases are usually removed just before the feedwater is pumped to the boiler. Most of these gases can be removed by boiling the water in open heaters and discharging the noncondensable gases through a venting condenser. The most effective method of removal is provided by spray- or tray-type deaerating heaters arranged for countercurrent scavenging of released gases to prevent their going back into solution in the feedwater.

Care must likewise be taken to prevent corrosion of piping, heaters, and other parts of the water flow path which make up the preboiler systems. *See* CORROSION; COUNTERCURRENT TRANSFER OPERATIONS.

Corrosion prevention. Under normal operating conditions, internal corrosion of boiler steel is prevented by maintaining the boiler water in an alkaline condition. At lower operating pressures, the addition of sodium hydroxide to the boiler water will suffice to produce a pH within the range 10.5–11.5. At higher operating pressures, the presence of strong alkalies in the boiler water can cause metallic corrosion where local concentration cells become established. In addition, sodium hydroxide volatilizes at high pressure sufficiently to lead to its deposition on turbine blades, with the consequent reduction of turbine output. At higher operating pressures, modern practice seeks to maintain only a few parts per million of sodium phosphate or a volatile amine (ammonia, morpholine, or cyclohexylamine) in the water to keep the pH in the range 9.0–10.0.

Trace quantities of oxygen in the boiler water not removed by deaeration can cause corrosion and are usually scavenged chemically with hydrazine or sodium sulfite. Sodium sulfite is seldom used at pressures above 1500 lb/in.^{2} (10 MPa) because its thermal decomposition produces undesirable acidic gases. *See* BOILER.

Eugene A. Avallone

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers,* 10th ed., 2000; D. Gunn and R. Horton, *Industrial Boilers*, 1988; L. I. Pinkus, *Practical Boiler Water Treatment, Including Air Conditioning Systems*, 1962, reprint 1981; S. T. Powell, *Water Conditioning for Industry,* 1954; C. D. Schroeder, *Solutions to Boiler and Cooling Water Problems*, 1986.

Boiler feedwater regulation

Addition of water to a steam-generating unit at a rate commensurate with the removal of steam from the unit. The addition of water to a boiler requires a feedwater pump or some other device that will develop a pressure higher than that of the steam generated. Means also are required to control the rate at which water is added. *See* BOILER.

Pressurization. Reciprocating plunger pumps, driven by steam from the boiler itself, often are used for small-capacity boilers. They have simplicity and are not dependent on other sources of power. Such pumps may have single or multiple cylinders, with single- or double-acting plungers, and they can be operated at variable speeds to deliver water at a steady and controlled rate to maintain a specified water level in the steam drum of the boiler.

Pumps for high-pressure high-capacity boilers are usually of the centrifugal type, with multiple stages connected in series to develop the required feedwater pressure. They may be driven by electric motors, by auxiliary steam turbines, or directly from the main turbine shaft. When a pump is operated at constant speed, the flow of water to the economizer or boiler is controlled by regulating valves installed in the discharge piping, with a bypass around the pump to handle the excess flow from the pump.

Injectors are sometimes used to supply water to portable fire-tube boilers. The injectors utilize the energy of a jet of steam, supplied from the boiler itself, to develop a high velocity by expanding the steam through a nozzle. The kinetic energy imparted by the high-velocity steam to the entrained supply water is converted to pressure in a discharge tube which is connected to the boiler shell.

Regulation. The water level in the steam drum must be clearly indicated by an illuminated gage glass, which is required by law for reasons of safety. It is attached to a water column or vertical chamber which is connected, in turn, without obstruction, between upper and lower nozzle openings in the steam drum. In high-pressure boilers two water gages are directly connected to the steam drum, one at each end, thus eliminating the use of water columns.

Intermittent manual regulation of the feedwater flow satisfactorily maintains the steam drum water level within acceptable limits in boilers having relatively large water-storage capacity, such as shell-type or multiple drum water-tube units, but automatic control provides better regulation.

For water-tube boilers having a single steam drum of relatively small diameter, a continuous and exact regulation of the feedwater is required. Automatic single-element control devices can be used to operate the supply valve or to change pump speed in response to a change in water level.

Variations of water level in the steam drum usually are due to changes in the rate of steam generation because such changes affect the steam output, the volume of the steam below the water level, and the consequent displacement of water into or out of the drum. Changes in water level can be compensated by use of an automatic three-element control which, primarily, regulates the rate of feedwater flow to be equal to the rate of steam output, as determined by metering equipment; the control then readjusts the rate of feedwater flow to maintain the water level within the prescribed normal range. *See* BOILER FEEDWATER; STEAM-GENERATING UNIT.

George W. Kessler

Boiling

A process in which a liquid phase is converted into a vapor phase. The energy for phase change is generally supplied by the surface on which boiling occurs. Boiling differs from evaporation at predetermined vapor/gas–liquid interfaces because it also involves creation of these interfaces at discrete sites on the heated surface. Boiling is an extremely efficient process for heat removal and is utilized in various energy-conversion and heat-exchange systems and in the cooling of high-energy density components. *See* BOILER; EVAPORATION; HEAT EXCHANGER; HEAT TRANSFER.

Boiling is classified into pool and forced flow. Pool boiling refers to boiling under natural convection conditions, whereas in forced flow boiling the liquid flow over the heater surface is imposed by external means. Flow boiling is subdivided into external and internal. In external flow boiling, liquid flow occurs over heated surfaces, whereas internal flow boiling refers to flow inside tubes. Heat fluxes of 2×10^8 W/m^2, or three times the heat flux at the surface of the Sun, have been obtained in flow boiling. *See* CONVECTION (HEAT).

Pool boiling. Figure 1, a qualitative pool boiling curve, shows the dependence of the wall heat flux q on the wall superheat ΔT (the difference between the wall temperature and the liquid's saturation temperature). The plotted curve is for a horizontal surface underlying a pool of liquid at its saturation temperature (the boiling point at a given pressure). *See* BOILING POINT.

Several heat transfer regimes can be identified on the boiling curve: single-phase natural convection, partial nucleate boiling, fully developed nucleate boiling, transition boiling, and film boiling.

Boiling inception. As the heat input to the surface is increased, the first mode of heat transfer to appear in a gravitational field is natural convection (region I in Fig. 1). At a certain value of wall superheat (point A), vapor bubbles appear on the heater surface. The

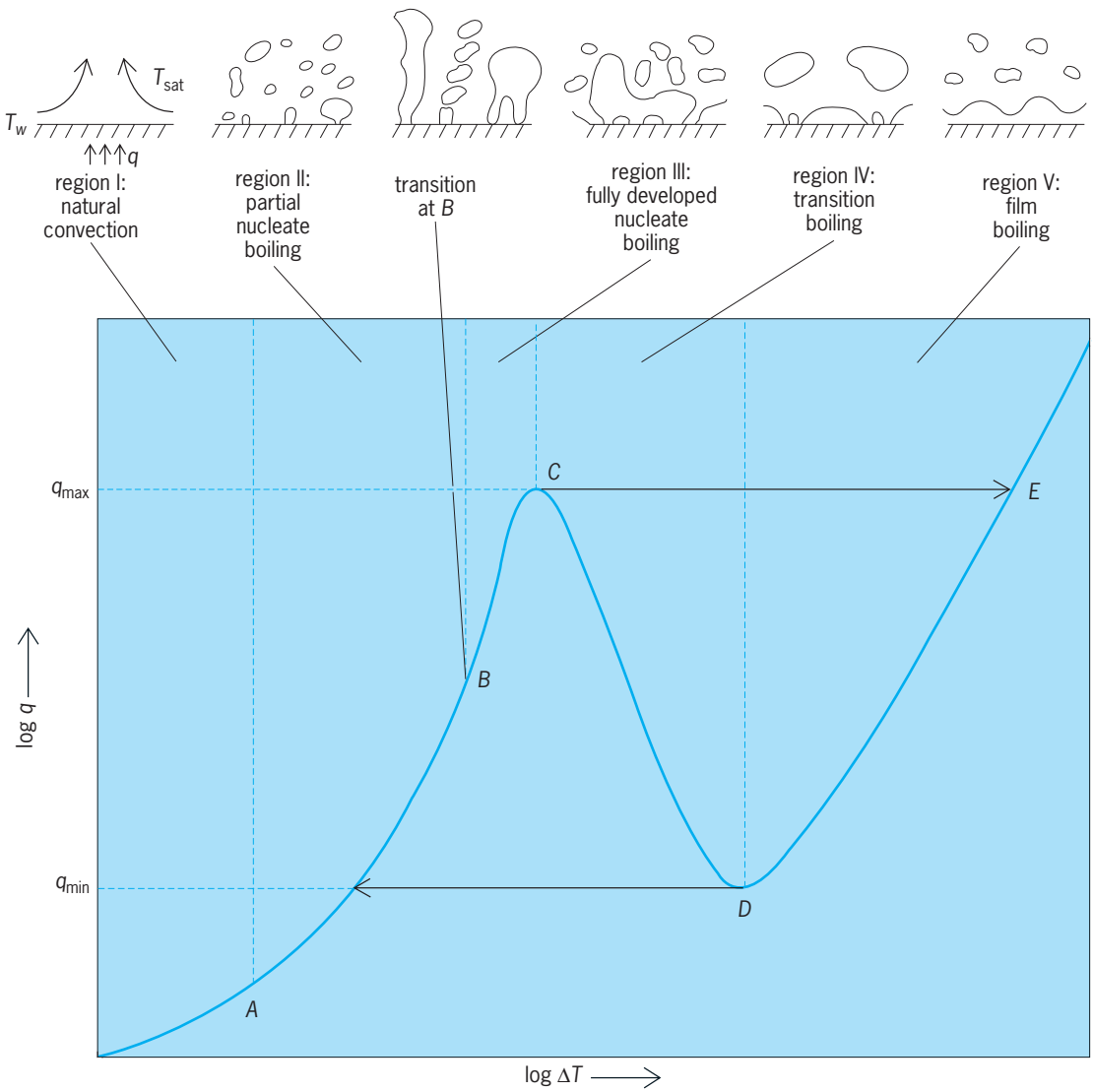

Fig. 1. Typical boiling curve, showing qualitatively the dependence of the wall heat flux q on the wall superheat ΔT, defined as the difference between the wall temperature and the saturation temperature of the liquid. Schematic drawings show the boiling process in regions I–V, and transition points A–E.

bubbles form on cavities or scratches on the surface because less surface energy is required there to form the vapor/gas–liquid interface. After inception, a bubble continues to grow until forces causing detachment exceed those holding the bubble against the wall. A complete analysis includes buoyancy, surface tension, fluid inertia, and liquid drag. *See* NUCLEATION.

Partial nucleate boiling. After inception a dramatic increase in the slope of the boiling curve is observed. In partial nucleate boiling, corresponding to region II (curve *AB*) in Fig. 1, discrete bubbles are released from randomly located active sites on the heater surface. The density of active bubble sites and the frequency of bubble release increase with wall superheat. Several parameters, such as the procedure used to prepare a surface, the wettability of the surface, the interaction between neighboring sites, and oscillations in temperature, affect the dependence of nucleation-site density on wall superheat.

Fully developed nucleate boiling. The transition (point *B* in Fig. 1) from isolated bubbles to fully developed nucleate boiling (region III) occurs when bubbles at a given site begin to merge in the vertical direction. Vapor appears to leave the heater in the form of jets. The condition of formation of jets also approximately coincides with the merger of vapor bubbles at the neighboring sites. Thus, vapor structures appear like mushrooms with several stems (**Fig. 2**). Now most of the heat transfer occurs because of evaporation at the vapor-liquid interface of the vapor stems implanted in the thermal layer adjacent to the heater surface. At present, no totally theoretical model exists for the prediction of nucleate boiling heat fluxes. However, the dependence of the heat flux on the wall superheat is given roughly by $q \propto \Delta T^{3 \text{ or } 4}$.

Maximum heat flux. The peak heat flux sets the upper limit on fully developed nucleate boiling or safe operation of equipment. For well-wetted surfaces, the upper limit on the nucleate-boiling heat flux is set by the vapor removal rate. Hydrodynamic theory based on vapor jet instability away from the heater has been successful in predicting the maximum heat flux on well-wetted surfaces. Maximum heat fluxes on partially wetted surfaces are lower than those on well-wetted walls.

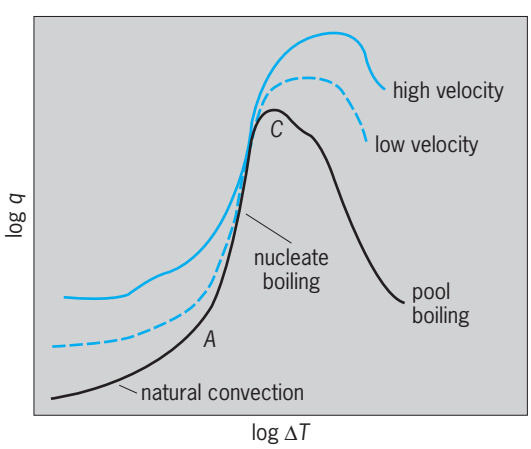

Fig. 2. Effect of external forced flow, at low and high velocity, on the boiling heat curve. The maximum heat flux is increased and the natural convection curve is shifted up due to forced convection. Locations of points *A* and *C* in Fig. 1 are indicated.

Film boiling. After the occurrence of maximum heat flux, most of the surface is rapidly covered by insulating vapor, and the surface temperature rises very rapidly. When heat flux is controlled, the surface will very quickly pass through regions IV and V in Fig. 1 and stabilize at point *E*. If the temperature at *E* exceeds the melting temperature of the heater material, the heater will fail (burn out). The curve *ED* (region V) represents stable film boiling, and the system can be made to follow this curve by reducing the heat flux. In stable film boiling, the surface is totally covered by vapor film. At low wall superheats just above point *D* in Fig. 1, the radiation contribution is negligible, and film-boiling heat flux varies roughly as $q \propto \Delta T^{3/4}$.

Minimum heat flux. The heat flux at which the stable film breaks down is called the minimum heat flux. Heater surface properties and geometrical parameters affect the minimum heat flux. From purely hydrodynamic considerations involving the instability of the interface, the minimum heat flux, point *D* in Fig. 1, can be predicted by theory. The heater temperature at point *D* is often called the Leidenfrost temperature, where a liquid droplet placed on the heater will dance on the vapor cushion that forms. It is named after J. C. Leidenfrost, who observed the phenomenon in 1756.

Transition boiling. Region IV in Fig. 1 (curve *CD*), falling between nucleate and film boiling, is called transition boiling. It is a mixed mode of boiling that has features of both nucleate and film boiling. Transition boiling is very unstable, since it is accompanied by a reduction in the heat flux with an increase in the wall superheat. As a result, it is extremely difficult to obtain steady-state data in transition boiling. In general, quenching or temperature-controlled systems are used to obtain transition-boiling data. Transition boiling is the least understood mode of boiling.

Forced flow boiling. Forced flow, both external and internal, greatly changes the boiling curve in Fig. 1. External flow effects are sketched in Fig. 2. The heat flux is increased by forced convection at temperatures below boiling inception, and after that the nucleate boiling region is extended upward until a flow-enhanced higher maximum flux (corresponding to point *C*) is achieved. Forced flow boiling in tubes is used in many applications; including steam generators, nuclear reactors, and cooling of electronic components. *See* NUCLEAR REACTOR; STEAM-GENERATING UNIT.

For flow in tubes, a distinction must be made between maximum or critical heat flux under low-flow, low-heat-flux and high-flow, high-heat-flux conditions. Under low-flow conditions, the liquid film can dry out and lead to a rise in wall temperature as the heat removal rate degrades. Under high-flow conditions, the critical heat flux condition corresponds to local dry-out of the tube surface even though the

tube core is full of liquid. Upon occurrence of the critical heat flux condition, the tube surface temperature rises rapidly to a very high value, and the tube can fail if the temperature exceeds the melting temperature of the heater material. V. K. Dhir

Bibliography. Y. Cengel, *Heat Transfer*, McGraw-Hill, 1997; J. G. Collier and J. R. Thome, *Convective Boiling and Condensation*, 3d ed., Oxford University Press, 1996; F. P. Incropera and D. P. Dewitt, *Fundamentals of Heat and Mass Transfer*, 4th ed., John Wiley, 1996; L. S. Tong and Y. S. Tang, *Boiling Heat Transfer and Two-Phase Flow*, 2d ed., Taylor and Francis, 1997.

Boiling point

The boiling point of a liquid is the temperature at which the liquid and vapor phases are in equilibrium with each other at a specified pressure. Therefore, the boiling point is the temperature at which the vapor pressure of the liquid is equal to the applied pressure on the liquid. The boiling point at a pressure of 1 atmosphere is called the normal boiling point.

For a pure substance at a particular pressure P, the stable phase is the vapor phase at temperatures immediately above the boiling point and is the liquid phase at temperatures immediately below the boiling point. The liquid-vapor equilibrium line on the phase diagram of a pure substance gives the boiling point as a function of pressure. Alternatively, this line gives the vapor pressure of the liquid as a function of temperature. The vapor pressure of water is 1 atm (101.325 kilopascals) at 100°C (212°F), the normal boiling point of water. The vapor pressure of water is 3.2 kPa (0.031 atm) at 25°C (77°F), so the boiling point of water at 3.2 kPa is 25°C. The liquid-vapor equilibrium line on the phase diagram of a pure substance begins at the triple point (where solid, liquid, and vapor coexist in equilibrium) and ends at the critical point, where the densities of the liquid and vapor phases have become equal. For pressures below the triple-point pressure or above the critical-point pressure, the boiling point is meaningless. Carbon dioxide has a triple-point pressure of 5.11 atm (518 kPa), so carbon dioxide has no normal boiling point. *See* TRIPLE POINT; VAPOR PRESSURE.

The normal boiling point is high for liquids with strong intermolecular attractions and low for liquids with weak intermolecular attractions. Helium has the lowest normal boiling point, 4.2 kelvin (−268.9°C). Some other normal boiling points are 111.1 K (−162°C) for CH_4, 450°C (842°F) for n-$C_{30}H_{62}$, 1465°C (2669°F) for NaCl, and 5555°C (10031°F) for tungsten.

The rate of change of the boiling-point absolute temperature T_b of a pure substance with pressure is given by the equation below. $\Delta H_{\text{vap,m}}$ is the molar

$$\frac{dT_b}{dP} = \frac{T_b \Delta V_{\text{vap,m}}}{\Delta H_{\text{vap,m}}}$$

enthalpy (heat) of vaporization, and $\Delta V_{\text{vap,m}}$ is the molar volume change on vaporization.

The quantity $\Delta H_{\text{vap,m}}/T_b$ is $\Delta S_{\text{vap,m}}$, the molar entropy of vaporization. The molar entropy of vaporization at the normal boiling point (nbp) is given approximately by Trouton's rule: $\Delta S_{\text{vap,m,nbp}} \approx$ 87 J/mol K (21 cal/mol K). Trouton's rule fails for highly polar liquids (especially hydrogen-bonded liquids). It also fails for liquids boiling at very low or very high temperatures, because the molar volume of the vapor changes with temperature and the entropy of a gas depends on its volume.

When a pure liquid is boiled at fixed pressure, the temperature remains constant until all the liquid has vaporized. When a solution is boiled at fixed pressure, the composition of the vapor usually differs from that of the liquid, and the change in liquid composition during boiling changes the boiling point. Thus the boiling process occurs over a range of temperatures for a solution. An exception is an azeotrope, which is a solution that boils entirely at a constant temperature because the vapor in equilibrium with the solution has the same composition as the solution. In fractional distillation, the variation of boiling point with composition is used to separate liquid mixtures into their components. *See* AZEOTROPIC MIXTURE; DISTILLATION; PHASE EQUILIBRIUM. Ira N. Levine

Bibliography. I. N. Levine, *Physical Chemistry*, 4th ed., 1995; J. E. Ricci, *The Phase Rule and Heterogeneous Equilibrium*, 1966.

Bolometer

A device for detecting and measuring small amounts of thermal radiation. The bolometer is a simple electric circuit, the essential element of which is a slab of material with an electrical property, most often resistance, that changes with temperature. Typical operation involves absorption of radiant energy by the slab, producing a rise in the slab's temperature and thereby a change in its resistance. The electric circuit converts the resistance change to a voltage change, which then can be amplified and observed by various, usually conventional, instruments.

Performance. The performance of a bolometer is measured in terms of its responsivity, or the electric signal generated per unit small change in incident radiation; its response time; its noise equivalent power, essentially the minimum detectable signal; and the spectral range over which the sensitive element produces a signal. Characteristics important for determining these quantities are the temperature coefficient of the sensitive element (how strongly its electrical properties change with temperature); the background electrical noise produced by the system; the thermal capacity of the element; and the thermal conductance between the element and a heat sink. These parameters can be controlled or improved by such means as encapsulating the element in a vacuum space, restricting its size, and cooling it to low temperature, such as to the boiling point of liquid helium, 4.2 K (−452.1°F), and below.

Sensitive element. The bolometer was invented in 1880 by S. P. Langley, who used a thin, blackened platinum strip as the sensitive element. Similar strips or fine wires are still used, but demands for improved performance have led to the deployment of materials with increased temperature coefficients, mainly semiconductors, such as carbon; mixed oxides of nickel, cobalt, and manganese (thermistor materials); germanium doped with gallium or indium; and indium antimonide. A class of bolometers incorporating a capacitive element uses thin films of materials with temperature-dependent electrical polarization properties. *See* SEMICONDUCTOR; THERMISTOR.

Applications. Langley used his bolometer to measure the spectral distribution of solar radiation in the infrared region. Although bolometers are useful in studying a variety of systems where detection of small amounts of heat is important, their main application is in measuring weak radiation signals in the infrared and far infrared, that is, at wavelengths from about 1 to 2000 micrometers from stars and interstellar material. *See* INFRARED ASTRONOMY; INFRARED RADIATION. W. E. Keller

Microwave bolometers. Bolometers are also used to detect and measure microwave energy or power. A bolometer constructed for this purpose is contained in a mount that protects the usually fragile element, guides the microwave energy to the bolometer, and has connection terminals for measuring the bolometer resistance. The bolometer resistance can be measured by using direct current or low-frequency instruments. By varying the direct-current or low-frequency power, the resistance of the bolometer can be adjusted so that most of the microwave power incident upon it will be absorbed. The direct current used to measure the resistance also dissipates power in the bolometer. Therefore, when the microwave power is applied, the direct-current or low-frequency power must be reduced to keep the bolometer resistance constant. The reduction of low-frequency power is a measure of the microwave power. The ratio of the low-frequency power change to the microwave power absorbed is the effective efficiency of the bolometer mount. The ratio of the low-frequency power change to the microwave power incident upon the bolometer is called the calibration factor of the bolometer mount.

A self-adjusting current loop (see **illus.**) can be designed to adjust the current that flows through a four-terminal resistor and a four-terminal bolometer so that the bolometer resistance is the same as that of the resistor. The power dissipated in the bolometer can be calculated from the voltage measured across the resistor.

Bolometers for measurement of microwave power are usually constructed by using one of four thermal-sensitive elements, as described below.

Fine metal wires. These bolometers are called barretters. They have positive temperature coefficients of resistance, are linear and stable, and are used for the most accurate measurements. Barretters are easily burned out and therefore are not used for routine measurements. *See* BARRETTER.

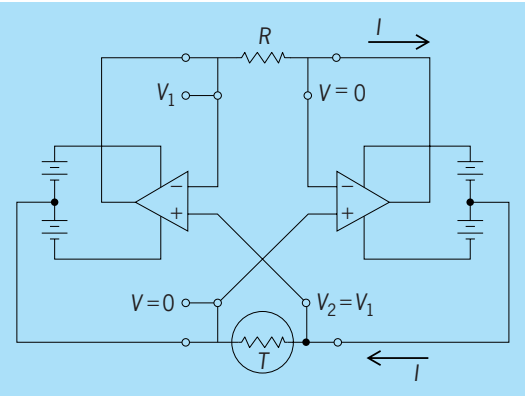

Self-adjusting current loop. The circuit adjusts currents so that the resistance of the bolometer *T* is the same as that of the resistor *R*. The circuit is shown with a semiconductor bolometer at *T*, but it can be used with a metal-type bolometer by interchanging *T* and *R*. (*After N. T. Larson, A new self-balancing DC-substitution RF power meter, IEEE Trans. Instrum. Meas., 25(4):343–347, December 1976*)

Semiconductor beads. These bolometers have negative temperature coefficients of resistance and are called thermistors. Thermistors are nonlinear and not as stable as barretters, but with proper techniques they can be used for accurate measurements. They are not as easily burned out as barretters and are therefore used extensively for microwave power measurements.

Thin metal films. Resistive-film bolometers are made by depositing a very thin metal film on a nonconducting substratum, and have characteristics similar to those of barretters. The microwave power dissipated in the resistance film also heats the substratum, and so these bolometers are not as sensitive or as subject to burnout as barretters. Resistive-film bolometers can be fabricated in special configurations to better fit the bolometer mount.

Ferroelectric capacitors. Thin sheets of ferroelectric material can be metallized on both sides to form a capacitor. These are called pyroelectric bolometers. The dielectric constant and the polarization of ferroelectric materials change with temperature, resulting in a change in capacity and a small current flow or voltage change. The metallized surfaces serve as electrodes and also as resistors to dissipate microwave power. Pyroelectric bolometers also may be fabricated in special shapes and have been used for pulse-power and millimeter measurements, but they are not common. *See* FERROELECTRICS; MICROWAVE POWER MEASUREMENT; PYROELECTRICITY; RADIOMETRY. Robert C. Powell

Bibliography. R. W. Boyd, *Radiometry and the Detection of Optical Radiation*, 1983; G. H. Bryant, *Principles of Microwave Measurements*, 1988; A. E. Fantom, A. E. Bailey, and A. C. Lynch (eds.), *Radio Frequency and Microwave Power Measurement*, Institution of Electrical Engineers, 1990; F. Grum and C. J. Bartleson (eds.), *Optical Radiation Measurements*, vol. 4: W. Budde (ed.), *Physical Detectors of Optical Radiation*, 1983; N. T. Larson, A new self-balancing DC-substitution RF power meter, *IEEE*

Trans. Instrum. Meas., 25(4):343–347, December 1976; M. Ou-Yang, C.-S. Sheen, and J.-S. Shie, Parameter extraction of resistive thermal microsensors by AC electrical measurement, *IEEE Trans. Instrum. Meas.*, 47(2):403–407, April 1998.

Bolt

A cylindrical fastener with an integral head on one end and an external screw thread on the other end designed to be inserted through holes in assembled parts and to mate with an internally threaded block, called a nut, which is turned to tighten or loosen the bolt. Tensioning the fastener by turning the nut differentiates a bolt from a screw, which is tightened by turning its head. *See* BOLTED JOINT; NUT (ENGINEERING).

Bolts are generally manufactured from metals, but bolts made of other materials, such as nylon, are commercially available. The properties of some bolting metals are modified by heat treatment and other means to increase yield strength. Bolt heads have various shapes to suit different applications (**Fig. 1**). Hexagon-headed bolts (hex bolts) are available in two head widths for the same body diameter—regular hex bolts and heavy hex bolts, which have wider heads. The heavy-series bolts are normally supplied with heavy nuts that are wider and thicker than regular nuts.

Manufacturing process. Modern bolt-forming processes require little machine finishing. The American National Standard requires none of the surfaces of square, hex, or heavy hex bolts to be machined except for the threads, providing that the other dimensions satisfy the Standard. Such bolts are commonly termed unfinished or rough. For the so-called finished bolts, only the undersides of the heads need to be finished; the bodies need not be machined if the forming process results in body dimensions that satisfy the Standard. Finishing of all bolt surfaces is required only for special applications. Ferrous bolts are obtainable with corrosion protection coatings such as zinc, cadmium, and organics, which are deposited on the bolts after machining.

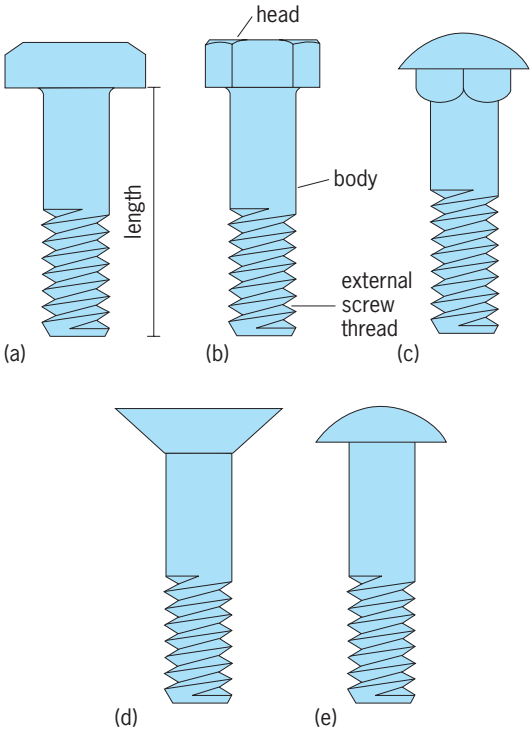

Fig. 1. Examples of bolts inserted through clearance holes: (*a*) square bolt; (*b*) hex bolt; (*c*) round-head square-neck bolt for connecting wood to metal; (*d*) countersunk bolt; and (*e*) round-head bolt.

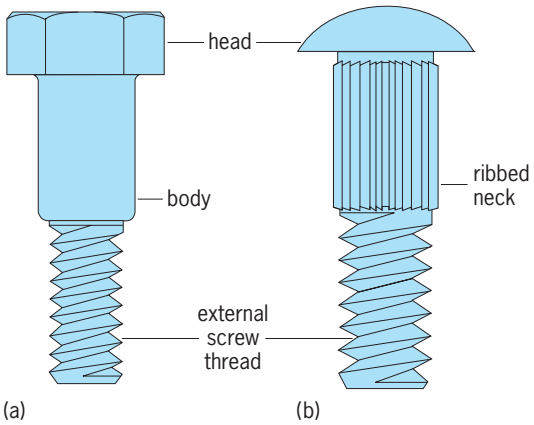

Fig. 2. Examples of bolts for interference body fits: (*a*) turned bolt and (*b*) round-head ribbed-neck bolt.

Fits and special bolts. The holes in the assembled parts through which the bolts are inserted are usually slightly larger in diameter than the bolt body, resulting in a clearance fit. For some applications the holes are made slightly smaller than the bolt body in order to obtain an interference fit; then it is necessary to force the bolt through the holes. Bolts for force fits usually have a body diameter larger than the major diameter of the thread in order to avoid damaging the threads during installation. Two examples of such special bolts are turned bolts and ribbed bolts (**Fig. 2**). All the surfaces of the turned bolt are normally finished, with the body finished smooth. Ribbed interference-body bolts are used in structural work because the hole size is not as critical as for turned bolts. The ribs of high-strength ribbed bolts are serrated. *See* LIMITS AND FITS.

Specifications. Bolts are made in a wide range of tensile strengths. Standard specifications such as those of the American Society for Testing and Materials (ASTM), the Society of Automotive Engineers (SAE), and the International Standards Organization (ISO) identify bolts by strength, grade, or property classes.

The following information is necessary to describe a bolt: external screw thread size, thread series, class of fit, and hand of thread; length; bolt head style; reference to standard specifications or material; and coating. (The bolt length is measured from the underside of the head to the tip of the thread, except for countersunk head and other special bolts.)

Bolt substitutes. Straight-thread screws are sometimes substituted for bolts. The common "stove bolt" fastener is not a bolt but a slotted round-head machine screw, with a square nut. Screws are also substituted for bolts because of the wider variety of heads available on screws, some of which are designed for tightening with small portable power tools.

Uses. Various types of bolts (Figs. 1 and 2) may be used for automobile, machinery, appliance, farm implement, and structural connections. For example, low-carbon-steel unfinished bolts with hex heads (ASTM A307, Grade A) are used in machinery, and with square heads for structural steel connections. Heat-treated medium-carbon-steel finished hex-head bolts (ASTM A449) are high-strength bolts used for connections in structures as well as machinery. However, there are two kinds of high-strength bolts made specifically for structural steel connections (ASTM A325 and ASTM A490), both kinds are heavy hex structural bolts (dimensions differ slightly from those of heavy hex screws). Other kinds of bolts are medium-carbon-steel or atmospheric-corrosion-resistant-alloy-steel, quenched and tempered bolts (A325); and alloy-steel or atmospheric-corrosion-resistant-alloy-steel, quenched and tempered bolts (A490). Both the A325 and A490 bolts are designed to be tightened by torquing either the head or the nut. A hardened washer is often specified to be placed under the part being turned. *See* SCREW FASTENER; WASHER. Charles Birnstiel

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 2000; E. Oberg et al., *Machinery's Handbook*, 26th ed., 2000.

Bolted joint

An assembly of two or more parts fastened together by inserting bolts through matching clearance holes in the parts and engaging nuts that are then tightened to clamp the assembly. The term bolted joint also colloquially denotes a screwed joint, for which screws are inserted through clearance holes in one part and tightened into internal threads in another part by turning the screw head. *See* BOLT.

External forces on bolts. The external force acting on a bolt of a joint depends on the manner in which the joint as a whole is loaded, the geometry of the bolt pattern, and the tension induced in the bolt during installation. A bolt may be subjected to direct tension, to shear, or to simultaneous tension and shear (**Fig. 1**). Cylinder head bolts are in direct tension; bolts uniformly spaced around the periphery pass through the head as well as the flange that is integral with the cylinder wall (Fig. 1*a*). If all bolts are tightened equally and the flange and head are sufficiently rigid, all the bolts will be subjected to direct tension of equal value. The external force F applied to the bolt is due to the pressure p inside the cylinder acting on the cylinder head and to the initial tension induced in the bolt during tightening of the nut.

In a double lap joint (Fig. 1*b*), the bolts are in shear

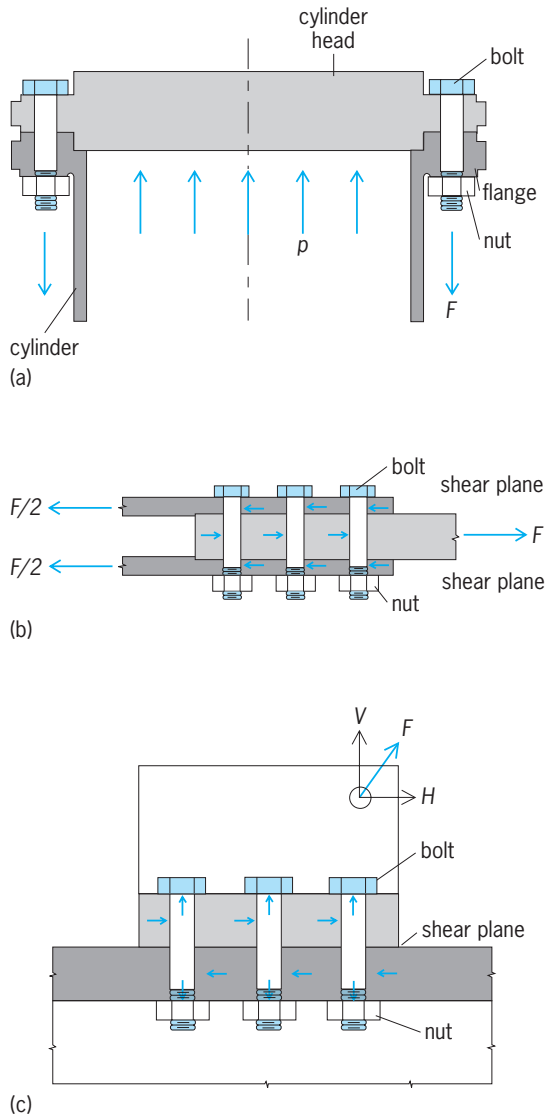

Fig. 1. Loadings on bolted joints. (*a*) Cylinder head bolts in direct tension. (*b*) Double lap joint, with bolts in double shear; small arrows represent the compressive forces applied by the plates to the bolts. (*c*) Tee bolted to a wide flange beam, with bolts in tension and shear.

at the planes between the outer and inner plates. The force F in the inner plate (due to the external loading) is equilibrated by the forces $F/2$ in the outer plates. Because each bolt passes through two shear planes, the bolts are said to be in double shear. The bolts will be subjected to different shearing forces. The two outer bolts (in the direction of loading) will be subjected to more shear than the inner bolt because of deformations of the plates. However, for design purposes the bolts are considered to be equally loaded unless the joint is long.

If the bolts are lightly tensioned (or loose), the bolt body will bear against the plates. The plates will slide with respect to each other (the joint is said to slip) until the plates bear on the bolts. Because of these bearing forces, the bolts will be subjected to bending in addition to shear. This bending is usually neglected in design. However, if the nuts on the

bolts are adequately torqued, the bolts will be pretensioned and the plates will be pressed together so that the shear force will be transmitted at the interface of the outer and inner plates by friction. If the friction is adequate, the joint will not slip and the plates will not bear on the bolts. The bolts of the tee joint (Fig. 1*c*) are subjected to tension due to the component *V* of the applied external force *F* and also to single shear due to the component *H*. A force, such as *F*, is a vector quantity; that is, it has magnitude and line of action. It can be replaced by two orthogonal forces, known as components, according to the parallelogram law (Fig. 1*c*). The bolts are in single shear because they pass through only one shear plane. If the bolts are not adequately pretensioned and the joint slips, the bolts may bear on the sides of the holes because of the component *H*.

Bolt tension. Bolts of machinery joints that are gasketed (gaskets are soft materials used to seal joints that need to be disassembled) are tightened to minimal initial tension to avoid damaging the gasket material. However, bolts in most rigid metal-to-metal machinery joints are tightened to about 75% of the yield strength of the fastener. The bolts are tightened to high initial tension so that the bolt tension remains constant as the external tensile load (Fig. 1*a* and *c*) fluctuates, assuming that the force applied to the bolt by the external load is less than the initial tension in the bolt. In this way the bolt is not subjected to fluctuating stresses that could lead to fatigue of the bolt. Fatigue is progressive, localized damage to the bolt material caused by fluctuating stress. The damage starts with cracks that may lead to fracture after a sufficient number of stress fluctuations. The time required to fracture depends on the character, magnitude, and rate of stress fluctuation as well as the bolt metallurgy. Bolts are also tensioned at joints subjected to shear (transverse loading) so that the shear will be transmitted between the assembled parts by friction, and the bolt body will not be subjected to shear and bending. Transverse loading of bolts is usually avoided in machinery design.

Connections of structural steel members (angles, channels, plates, and wide flanged beams) may be made with unfinished bolts of low carbon steel that are tightened with manual wrenches. Because the amount of pretension is uncertain, they are not considered pretensioned for design purposes and hence are denoted as bearing-type connections. For design, these joints are treated similarly to riveted connections, because the magnitude of the clamping force of hot-driven rivets is also variable.

High-strength structural bolts are normally tightened to a specific installation tension, which should be at least 70% of the tensile strength of the bolt. For a specific condition of the faying surface and the controlled clamping force, the force needed to cause slip of a shear joint may be computed. For certain service conditions the connection is designed such that the shear force acting on the most severely loaded fastener does not exceed that causing slip. Such joints are known as slip-critical joints.

Machinery joints. The design philosophy for bolted machinery joints (either permanent connections or those that are designed for disassembly) is that, generally, the fastener should pass through clearance holes and not be subjected to shear. Machinery joints subjected to shear are designed so that the shear is transferred between the parts by other mechanical means such as dowel pins, keys, and bearing surfaces configured into the joint. An exception is machinery mounting to structural steel (in some industries), where turned bolts in interference holes are utilized to resist shear and tension (Fig. 1*c*). These turned interference bolts may or may not be pretensioned. The general policy for bolt installation is to tighten

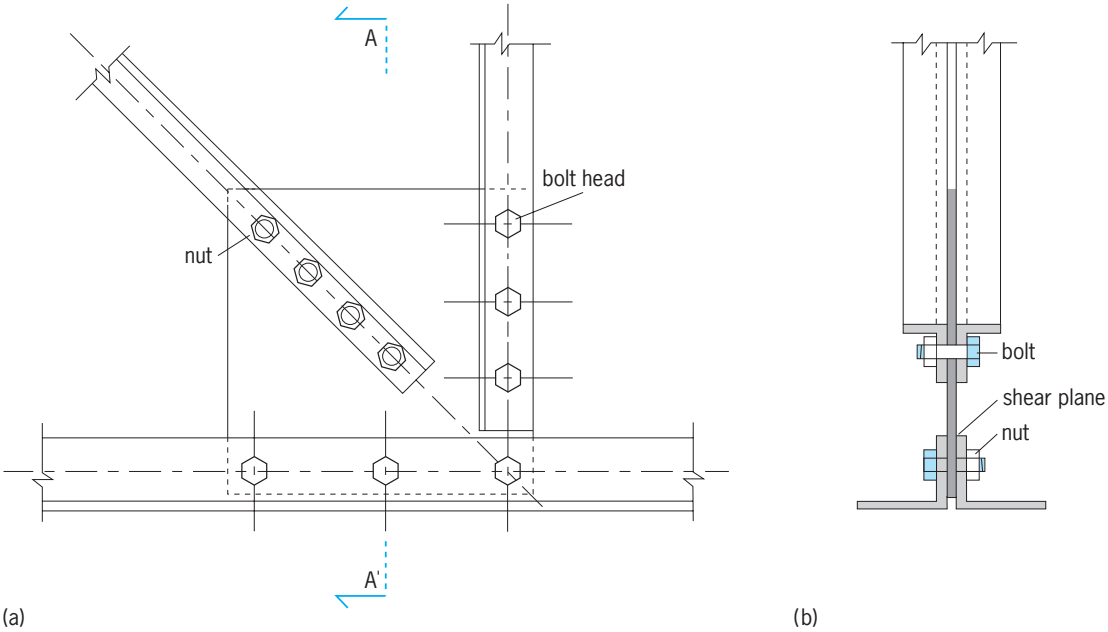

Fig. 2. Building roof truss joint. (*a*) High-strength bearing-type connection with all bolts in double shear. (*b*) Cross section A-A′.

the bolt to a specific minimum torque that is an index of the probable bolt pretension.

Structural joints. Lightly loaded beam-to-column connections in building frames are often made with unfinished bolts in clearance holes. The connections are usually configured so that the fasteners are in shear due to gravity forces. These joints are bearing-type without controlled pretension, and the connection is permitted to slip until the clearance is overcome and the bolt bodies bear against the structural steel parts.

High-strength bolts are used for connections in bridges and for building frame connections subjected to vibration and reversal of loading. The service condition is such that the connections are considered slip-critical. They are designed to transfer shear by friction, requiring controlled tensioning of the fasteners at installation.

For heavily loaded building frame joints that are not slip-critical, A325 and A490 bolts are designed as bearing-type connections. The allowable shear load per fastener is greater for the bearing-type connection than for the slip-critical connection, for the same fastener type and size. Hence, fewer bolts are required. High-strength bolts are required to be pretensioned for some bearing-type connections but need only be made snug tight for others. For example, a structural connection of a roof truss (**Fig. 2**) has high-strength bolts in double shear to transmit forces between the truss members. *See* STRUCTURAL CONNECTIONS.

High-strength bolt tightening. High-strength bolts for structural connections need to be tightened to at least 70% of their tensile strength in order to develop sufficient friction in the joint to be allowed loadings specified in building codes. The installation tensioning must be monitored. Three procedures for controlling the induced pretension for A325 and A490 bolts are the turn-of-the-nut-method, calibrated-wrench tightening, and direct-tension indicators (washers or ultrasonic transducers). For certain critical structural connections, and for special situations in machinery assembly, these three procedures are not considered sufficiently reliable, and the tension induced is monitored by measuring the bolt extension during tightening. *See* JOINT (STRUCTURES).

Charles Birnstiel

Bibliography. E. H. Gaylord and J. E. Stallmeyer, *Design of Steel Structures*, 3d ed., 2000; Industrial Fasteners Institute Staff, *Fastener Standards*, 6th ed., 1988; C. G. Salmon and J. E. Johnson, *Steel Structures: Design and Behavior*, 4th ed., 1996; J. E. Shigley and C. R. Mischke, *Mechanical Engineering Design*, 5th ed., 1989.

Boltzmann constant

A constant occurring in practically all statistical formulas and having a numerical value of 1.3807×10^{-23} joule/K. It is represented by the letter k. If the temperature T is measured from absolute zero, the quantity kT has the dimensions of an energy and is usually called the thermal energy. At 300 K (room temperature), $kT = 0.0259$ eV.

The value of the Boltzmann constant may be determined from the ideal gas law. For 1 mole of an ideal gas Eq. (1*a*) holds, where P is the pressure, V the

$$PV = RT \tag{1a}$$

$$PV = NkT \tag{1b}$$

volume, and R the universal gas constant. The value of R, 8.31 J/K mole, may be obtained from equation-of-state data. Statistical mechanics yields for the gas law Eq. (1*b*). Here N, the number of molecules in 1 mole, is called the Avogadro number and is equal to 6.02×10^{23} molecules/mole. Hence, comparing Eqs. (1*a*) and (1*b*), one obtains Eq. (2). Since k occurs explicitly in the distribution formula, Eq. (2),

$$k = R/N = 1.3807 \times 10^{-23}\ \text{J/K} \tag{2}$$

any quantity calculated using the Boltzmann distribution depends explicitly on k. Examples are specific heat, viscosity, conductivity, and the velocity of sound. Perhaps the most unusual relation involving k is the one between the probability of a state W and the entropy S, given by Eq. (3), which is obtained

$$S = k \ln W \tag{3}$$

by a process of identification similar to the one just described.

Almost any relation derived on the basis of the partition function or the Bose-Einstein, Fermi-Dirac, or Boltzmann distribution contains the Boltzmann constant. *See* AVOGADRO'S NUMBER; BOLTZMANN STATISTICS; BOSE-EINSTEIN STATISTICS; FERMI-DIRAC STATISTICS; KINETIC THEORY OF MATTER; STATISTICAL MECHANICS.

Max Dresden

Boltzmann statistics

To describe a system consisting of a large number of particles in a physically useful manner, recourse must be had to so-called statistical procedures. If the mechanical laws operating in the system are those of classical mechanics, and if the system is sufficiently dilute, the resulting statistical treatment is referred to as Boltzmann or classical statistics. (Dilute in this instance means that the total volume available is much larger than the proper volume of the particles.) A gas is a typical example: The molecules interacting according to the laws of classical mechanics are the constituents of the system, and the pressure, temperature, and other parameters are the overall entities which determine the macroscopic behavior of the gas. In a case of this kind it is neither possible nor desirable to solve the complicated equations of motion of the molecules; one is not interested in the position and velocity of every molecule at any time. The purpose of the statistical description is to extract from the mechanical description just those features relevant for the determination of the macroscopic properties and to omit others.

Distribution function. The basic notion in the statistical description is that of a distribution function. Suppose a system of N molecules is contained in a volume V. The molecules are moving around, colliding with the walls and with each other. Construct the following geometrical representation of the mechanical system. Introduce a six-dimensional space (usually called the μ space), three of its coordinate axes being the spatial coordinates of the vessel x, y, z, and the other three indicating cartesian velocity components υ_x, υ_y, υ_z. A molecule at a given time, having a specified position and velocity, may be represented by a point in this six-dimensional space. The state of the gas, a system of N molecules, may be represented by a cloud of N points in this space. In the course of time, this cloud of N points moves through the μ space.

Note that the μ space is actually finite; the coordinates x, y, z of the molecules' position are bounded by the finite size of the container, and the velocities are bounded by the total energy of the system. Imagine now that the space is divided into a large number of small cells, of sizes $w_1, \ldots, w_i, \ldots$. A certain specification of the state of the gas is obtained if, at a given time t, the numbers $n_1(t), \ldots, n_i(t), \ldots$ of molecules in the cells $1, \ldots, i, \ldots$ are given. The change in the state of the system in the course of time is now determined by the behavior of the functions $n_i(t)$ as functions of time. Strictly speaking, these functions may change discontinuously in the course of time. Just how detailed the description so obtained is depends, of course, on the sizes chosen for the cells w_i. One gets the crudest possible description by having just one cell of size V in the coordinate subspace, with N particles in it all the time. A very detailed description is obtained if cells the size of a molecule are chosen. In that case even a small change in the state of the molecules will cause a profound alteration in the numbers n_i. To apply statistical methods, one must choose the cells such that on the one hand a cell size w is small compared to the macroscopic dimensions of the system, while on the other hand w must be large enough to allow a large number of molecules in one cell. That this is possible stems from the fact that the molecular dimensions (linear dimension about 10^{-8} cm) are small compared to macroscopic dimensions (typical linear dimension 10^2 cm). In this case it is possible to have about 10^{15} cells, each one of linear dimension 10^{-3}, where in each cell there is "room" for 10^{15} molecules. If the cells are thus chosen, the numbers $n_i(t)$, the occupation numbers, will be slowly changing functions of time. The distribution functions $f_i(t)$ are defined by Eq. (1).

$$n_i(t) = f_i(t) w_i \tag{1}$$

The distribution function f_i describes the state of the gas, and f_i of course varies from cell to cell. Since a cell i is characterized by a given velocity range and position range, and since for appropriately chosen cells f should vary smoothly from cell to cell, f is often considered as a continuous function of the variables $x, y, z, \upsilon_x, \upsilon_z$. The cell size w then can be written as $dxdydzd\upsilon_x d\upsilon_y d\upsilon_z$.

In applications the continuous notation is often used; this should not obscure the fact that the cells are finite. L. Boltzmann called them physically infinitesimal.

Since a cell i determines both a position and a velocity range, one may associate an energy ϵ_i with a cell. This is the energy a single molecule possesses when it has a representative point in cell i. This assumes that, apart from instantaneous collisions, molecules exert no forces on each other. If this were not the case, the energy of a molecule would be determined by the positions of all other molecules.

Boltzmann equation; H theorem. Most of the physically interesting quantities follow from a knowledge of the distribution function; the main problem in Boltzmann statistics is to find out what this function is. It is clear that $n_i(t)$ changes in the course of time for three reasons: (1) Molecules located at the position of cell i change their positions and hence move out of cell i; (2) molecules under the influence of outside forces change their velocities and again leave the cell i; and (3) collisions between the molecules will generally cause a (discontinuous) change of the occupation numbers of the cells. Whereas the effect of (1) and (2) on the distribution function follows directly from the mechanics of the system, a separate assumption is needed to obtain the effect of collisions on the distribution function. This assumption, the collision-number assumption, asserts that the number of collisions per unit time, of type $(i,j) \rightarrow (k,l)$ [molecules from cells i and j collide to produce molecules of different velocities which belong to cells k and l], called A_{ij}^{kl}, is given by Eq. (2). Here

$$A_{ij}^{kl} = n_i n_j a_{ij}^{kl} \tag{2}$$

a_{ij}^{kl} depends on the collision configuration and on the size and kind of the molecules but not on the occupation numbers. Furthermore, for a collision to be possible, the conservation laws of energy and momentum must be satisfied; so if $\epsilon_i' = {}^1/_2 m\mathbf{v}_i^2$ and $\mathbf{p}_i = m\mathbf{v}_i$, then Eqs. (3*a*) and (3*b*) hold. It is possible to show

$$\epsilon_i' + \epsilon_j' = \epsilon_k' + \epsilon_l' \tag{3a}$$

$$\mathbf{p}_i + \mathbf{p}_j = \mathbf{p}_k + \mathbf{p}_l \tag{3b}$$

that the geometrical factor a_{ij}^{kl} has the property given by Eq. (4). Here a_{ij}^{kl} is the geometrical factor belong-

$$w_i w_j a_{ij}^{kl} = w_k w_l a_{kl}^{ij} \tag{4}$$

ing to the collision which, starting from the final velocities (k and l), reproduces the initial ones (i and j). Gains and losses of the molecules in, say, cell i can now be observed. If the three factors causing gains and losses are combined, the Boltzmann equation, written as Eq. (5), is obtained. Here $\Delta_x f_i$ is the gra-

$$\frac{\partial f_i}{\partial t} + (\mathbf{v}_i \cdot \Delta_x f_i) + (\mathbf{X}_i \cdot \Delta_v f_i) = \sum_{j,k,l} a_{ij}^{kl} w_j (f_k f_l - f_i f_j) \tag{5}$$

dient of f with respect to the positions, $\Delta_v f_i$ refers similarly to the velocities, and $\mathbf{X}_i$ is the outside force per unit mass at cell i. This nonlinear equation determines the temporal evolution of the distribution function. Exact solutions are difficult to obtain. Yet Eq. (5) forms the basis for the kinetic discussion of most transport processes. There is one remarkable general consequence, which follows from Eq. (5). If one defines $H(t)$ as in Eq. (6), one finds by straight manipulation from Eqs. (5) and (6) that Eqs. (7) hold.

$$H(t) = \sum_i n_i \ln f_i \tag{6}$$

$$\frac{dH}{dt} \leq 0 \qquad \frac{dH}{dt} = 0 \qquad \text{if } f_i f_j = f_k f_l \tag{7}$$

Hence H is a function which in the course of time always decreases. This result is known as the H theorem. The special distribution which is characterized by Eq. (8) has the property that collisions do not

$$f_i f_j = f_k f_l \tag{8}$$

change the distribution in the course of time; it is an equilibrium or stationary distribution

Maxwell-Boltzmann distribution. It should be stressed that the form of the equilibrium distribution may be determined from Eq. (8), with the help of the relations given as Eqs. (3*a*) and (3*b*). For a gas which as a whole is at rest, it may be shown that the only solution to functional Eq. (8) is given by Eqs. (9*a*) or (9*b*). Here A and β are parameters, not

$$f_i = Ae^{-\beta\epsilon i} \tag{9a}$$

$$f(\mathbf{x}, \mathbf{v}) = Ae^{(-1/2)\beta m v^2 - \beta U} \tag{9b}$$

determined by Eq. (8), and U is the potential energy at the point x, y, z. Equations (9*a*) and (9*b*) are the Maxwell-Boltzmann distribution. Actually A and β can be determined from the fact that the number of particles and the energy of the system are specified, as in Eqs. (10*a*) and (10*b*). From Eqs. (9) and (10)

$$\sum n_i = N \tag{10a}$$

$$\sum n_i \epsilon_i = E \tag{10b}$$

it can be shown immediately that Eqs. (11*a*) and (11*b*) hold. Therefore β is related directed to the

$$\frac{E}{N} = \frac{3}{2\beta} \tag{11a}$$

$$A = \frac{N}{V}\left(\frac{\beta m}{2\pi}\right)^{3/2} \tag{11b}$$

energy per particle while A is related to β and to the number density. Other physical quantities, such as pressure, must now be calculated in terms of A and β. Comparing such calculated entities with those observed, at a certain point one identifies an entity like β (a parameter in the distribution function) with a measured quantity, such as temperature. More precisely $\beta = 1/kT$, where k is the Boltzmann constant. This is the result of an identification. It is not a deduction. It is possible by specialization of Eq. (9*b*) to obtain various familiar forms of the distribution law. If $U = 0$, one finds immediately that the number of molecules whose absolute value of velocities lies between c and $c + dc$ is given approximately by Eq. (12). Equation (12) is useful in many applications.

$$4\pi V = Ae^{-(1/2)\beta mc^2} c^2 dc \tag{12}$$

If, on the other hand, a gas is in a uniform gravitational field so that $U = -mgz$, one finds again from Eqs. (9*a*) and (9*b*) that the number of molecules at the height z (irrespective of their velocities) is given by notation (13). If one uses the fact that $\beta = 1/kT$,

$$\text{Constant} \times e^{-\beta mgz} \tag{13}$$

then notation (13) expresses the famous barometric density formula, the density distribution of an ideal gas in a constant gravitational field.

Statistical method; fluctuations. The indiscriminate use of the collision-number assumption leads, via the H theorem, to paradoxical results. The basic conflict stems from the irreversible results that appear to emerge as a consequence of a large number of reversible fundamental processes. Although a complete detailed reconciliation between the irreversible aspects of thermodynamics as a consequence of reversible dynamics remains to be given, it is now well understood that a careful treatment of the explicit and hidden probability assumptions is the key to the understanding of the apparent conflict. It is clear, for instance, that Eq. (2) is to be understood either in an average sense (the average number of collisions is given by $\overline{n_i n_j a_{ij}^{kl}}$, where $n_i n_j$ is an average of the product) or as a probability, but not as a definite causal law. The treatment usually given actually presupposes that $n_i n_j = \overline{n_i}\,\overline{n_j}$; that is, it neglects fluctuations. *See* CHEMICAL THERMODYNAMICS.

To introduce probability ideas in a more explicit manner, consider a distribution of N molecules over the cells w_i in μ space, as shown in notation (14). Let

$$\begin{gathered} w_1, w_2, \ldots, w_i, \ldots \\ n_1, n_2, \ldots, n_i, \ldots \\ \epsilon_1, \epsilon_2, \ldots, \epsilon_i, \ldots \end{gathered} \tag{14}$$

the total volume of the μ space be Ω. Suppose that N points are thrown into the μ space. What is the probability that just n_1 points will end up in cell 1, n_2 in cell 2, n_i in cell i, and so on? If all points in the space are equally likely (equal a priori probabilities), the probability of such an occurrence is shown in Eq. (15). One now really defines the probability of a

$$W(n_1, \ldots, n_i, \ldots) = \frac{N!}{n_1! n_2! \ldots n_i! \ldots} \cdot \left(\frac{w_1}{\Omega}\right)^{n_1} \cdots \left(\frac{w_i}{\Omega}\right)^{n_i} \cdots \tag{15}$$

physical state characterized by notation (14) as given

by the essentially geometric probability shown in Eq. (15). The probability of a state so defined is intimately connected with the H function. If one takes the ln of Eq. (15) and applies Stirling's approximation for ln $n!$, one obtains Eq. (16). The H theorem,

$$\ln W = -H \tag{16}$$

which states that H decreases in the course of time, therefore may be rephrased to state that in the course of time the system evolves toward a more probable state. Actually the H function may be related (for equilibrium states only) to the thermodynamic entropy S by Eq. (17). Here k is, again, the Boltzmann constant.

$$S = -kH \tag{17}$$

Equations (16) and (17) together yield an important result, Eq. (18), which relates the entropy of a

$$S = k \ln W \tag{18}$$

state to the probability of that state, sometimes called Boltzmann's relation. *See* PROBABILITY.

The state of maximum probability, that is, the set of occupation numbers $n_1, \ldots, n_i \ldots$, which maximizes Eq. (15), subject to the auxiliary conditions given in Eqs. (10*a*) and (10*b*), turns out to be given by the special distribution written as Eqs. (19*a*), (19*b*), and (19*c*). This again is the Maxwell-Boltzmann dis-

$$n_i = Ae^{-\beta\epsilon_i} \tag{19a}$$

$$\sum n_i = N \tag{19b}$$

$$\sum n_i\epsilon_i = E \tag{19c}$$

tribution. Hence the equilibrium distribution may be thought of as the most probable state of a system. If a system is not in equilibrium, it will most likely (but not certainly) go there; if it is in equilibrium, it will most likely (but not certainly) stay there. By using such probability statements, it may be shown that the paradoxes and conflicts mentioned before may indeed be removed. The general situation is still not clear in all details, although much progress has been made, especially for simple models. A consequence of the probabilistic character of statistics is that the entities computed also possess this characteristic. For example, one cannot really speak definitively of the number of molecules hitting a section of the wall per second, but only about the probability that a given number will hit the wall, or about the average number hitting. In the same vein, the amount of momentum transferred to a unit area of the wall by the molecules per second (this, in fact, is precisely the pressure) is also to be understood as an average. This in particular means that the pressure is a fluctuating entity. In general, it may be shown that the fluctuations in a quantity Q are defined by expression (20),

$$\frac{\overline{Q}^2 - (\overline{Q})^2}{(\overline{Q})^2} \cong \frac{1}{N} \tag{20}$$

where $\overline{Q}$ = the average of Q, and N = number of particles in the systems. When the fluctuations are small, the statistical description as well as the thermodynamic concept is useful. The fluctuations in pressure may be demonstrated by observing the motion of a mirror, suspended by a fiber, in a gas. On the average, as many gas molecules will hit the back as the front of the mirror, so that the average displacement will indeed be zero. However, it is easy to imagine a situation where more momentum is transferred in one direction than in another, resulting in a deflection of the mirror. From the knowledge of the distribution function the probabilities for such occurrences may indeed be computed; the calculated and observed behavior agree very well. This clearly demonstrates the essentially statistical character of the pressure. *See* BOLTZMANN TRANSPORT EQUATION; BROWNIAN MOVEMENT; KINETIC THEORY OF MATTER; QUANTUM STATISTICS; STATISTICAL MECHANICS. Max Dresden

Boltzmann transport equation

An equation which is used to study the nonequilibrium behavior of a collection of particles. In a state of equilibrium a gas of particles has uniform composition and constant temperature and density. If the gas is subjected to a temperature difference or disturbed by externally applied electric, magnetic, or mechanical forces, it will be set in motion and the temperature, density, and composition may become functions of position and time; in other words, the gas moves out of equilibrium. The Boltzmann equation applies to a quantity known as the distribution function, which describes this nonequilibriium state mathematically and specifies how quickly and in what manner the state of the gas changes when the disturbing forces are varied. *See* KINETIC THEORY OF MATTER; STATISTICAL MECHANICS.

Equation (1) is the Boltzmann transport equation,

$$\frac{\partial f}{\partial t} = \left(\frac{\partial f}{\partial t}\right)_{\text{force}} + \left(\frac{\partial t}{\partial t}\right)_{\text{diff}} + \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \tag{1}$$

where f is the unknown distribution function which, in its most general form, depends on a position vector $\mathbf{r}$, a velocity vector $\mathbf{v}$, and the time t. The quantity $\partial f/\partial t$ on the left side of Eq. (1) is the rate of change of f at fixed values of $\mathbf{r}$ and $\mathbf{v}$. The equation expresses this rate of change as the sum of three contributions: first, $(\partial f/\partial t)_{\text{force}}$ arises when the velocities of the particles change with time as a result of external driving forces; second, $(\partial f/\partial t)_{\text{diff}}$ is the effect of the diffusion of the particles from one region in space to the other; and third, $(\partial f/\partial t)_{\text{coll}}$ is the effect collisions of the particles with each other or with other kinds of particles.

The distribution function carries information about the positions and velocities of the particles at any time. The probable number of particles N at the time t within the spatial element $dxdydz$ located at (x,y,z) and with velocities in the element $d\upsilon_x d\upsilon_y d\upsilon_z$

at the point ($\upsilon_x,\upsilon_y,\upsilon_z$ is given by Eq. (2) or, in vector notation, by Eq. (3). It is assumed that the particles

$$N = f(x, y, z, v_x, v_y, v_z, t)dxdydzdv_xdv_ydv_z \quad (2)$$

$$N = f(\mathbf{r}, \mathbf{v}, t)d^3rd^3v \quad (3)$$

are identical; a different distribution function must be used for each species if several kinds of particles are present.

Specific expressions can be found for the terms on the right side of Eq. (1). Suppose that an external force F_x acts on each particle, producing the acceleration $a_x = F_x/m$ and hence changing the velocity by $\Delta\upsilon_x = a_x\Delta t$ in the time interval Δt. If a group of particles has the velocity υ_x at time t, the same particles will have the velocity $\upsilon_x + \Delta v_x$ at the later time $t + \Delta t$. Therefore the distribution functions $f(\upsilon_x,t)$ and $f(\upsilon_x + \Delta\upsilon_x, t + \Delta t)$ satisfy the equality of Eq. (4),

$$f(v_x, t) = f(v_x + \Delta v_x, t + \Delta t) \quad (4)$$

when only the effect of acceleration a_x is taken into account. Multiplying Eq. (4) by -1 and adding $f(\upsilon_x,t + \Delta t)$ to both sides yields Eq. (5).

$$f(v_x, t + \Delta t) - f(v_x, t) = -[f(v_x + \Delta v_x, t + \Delta t) - f(v_x, t + \Delta t)] \quad (5)$$

Multiplying the left side of Eq. (5) by $1/\Delta t$ and the right side by the equal quantity $a_x/\Delta\upsilon_x$, Eq. (6) is

$$\frac{f(v_x, t + \Delta t) - f(v_x, t)}{\Delta t} = -a_x\left[\frac{f(v_x + \Delta v_x, t + \Delta t) - f(v_x, t + \Delta t)}{\Delta v_x}\right] \quad (6)$$

obtained. In the limit as $\Delta t \to 0$, Eq. (6) becomes Eq. (7). The generalization of this result for acceleration in an arbitrary direction is then given by Eq. (8). The quantity $(\partial f/\partial t)_{\text{force}}$ therefore depends on

$$\left(\frac{\partial f}{\partial t}\right)_{\text{force}} = -\left[a_x\frac{\partial f}{\partial v_x} + a_y\frac{\partial f}{\partial v_y} + a_z\frac{\partial f}{\partial v_z}\right] = -\mathbf{a}\cdot\frac{\partial f}{\partial \mathbf{v}} \quad (8)$$

$$\left(\frac{\partial f}{\partial t}\right)_{\text{force}} = -a_x\frac{\partial f}{\partial v_x} \quad (7)$$

both $\mathbf{a}$, the rate of change of velocity of the particles, and $\partial f/\partial\mathbf{v}$, the variation of the distribution function with velocity.

In a similar way $(\partial f/\partial t)_{\text{diff}}$ depends on both $\mathbf{v}$, the rate of change of position of the particles, and $(\partial f/\partial\mathbf{r})$, the variation of the distribution function with position. One writes $\Delta x = \upsilon_x\Delta t$ in place of $\Delta\upsilon_x = a_x\Delta t$. Then the form of Eqs. (4)–(7) is unchanged, except that υ_x is replaced by x and a_x by υ_x. The final expression is given by Eq. (9).

$$\left(\frac{\partial f}{\partial t}\right)_{\text{diff}} = -\left[v_x\frac{\partial f}{\partial_x} + v_y\frac{\partial f}{\partial_y} + v_z\frac{\partial f}{\partial_z}\right] = -\mathbf{v}\cdot\frac{\partial f}{\partial \mathbf{r}} \quad (9)$$

If Eqs. (8) and (9) are substituted into Eq. (1), one gets Eq. (10), which is the usual form of the Boltzmann equation. Before it can be solved for f, a specific expression must be found for $(\partial f/\partial t)_{\text{coll}}$, the rate of change of $f(\mathbf{r},\mathbf{v},t)$ due to collisions. The calculation of $(\partial f/\partial t)_{\text{coll}}$ begins with a mathematical description of the forces acting between particles. Knowing the type of statistics obeyed by the particles (that is, Fermi-Dirac, Bose-Einstein, or Maxwell-Boltzmann), the manner in which the velocities of the particles are changed by collisions can then be determined. The term $(\partial f/\partial t)_{\text{coll}}$ is expressed as the difference between the rate of scattering from all possible velocities $\mathbf{v}'$ to the velocity $\mathbf{v}$, and the rate of scattering from $\mathbf{v}$ to all possible $\mathbf{v}'$. For example, if the particles are electrons in a metal or semiconductor and if they are scattered elastically by imperfections in the solid, it is found that $(\partial f/\partial t)_{\text{coll}}$ obeys Eq. (11),

$$\frac{\partial f}{\partial t}\bigg)_{\text{coll}} = \sum_{\mathbf{v}'} W_{\mathbf{vv}'}f(\mathbf{r}, \mathbf{v}', t) - \sum_{\mathbf{v}'} W'_{\mathbf{vv}}f(\mathbf{r}, \mathbf{v}, t) \quad (11)$$

where $W_{\mathbf{vv}'}$ is the rate of scattering of one particle from $\mathbf{v}$ to $\mathbf{v}'$ or the reverse. A basic requirement for obtaining a meaningful expression for $(\partial f/\partial t)_{\text{coll}}$ in this way is that the duration of a collision be small compared to the time between collisions. For gas of atoms or molecules this condition implies that the gas be dilute, and for electrons in a solid it implies that the concentration of imperfections must not be too high.

Discussion. The Boltzmann expression, Eq. (10), with a collision term of the form in Eq. (11), is irreversible in time in the sense that if $f(\mathbf{r},\mathbf{v},t)$ is a solution then $f(\mathbf{r},-\mathbf{v},-t)$ is not a solution. Thus if an isolated system is initially not in equilibrium, it approaches equilibrium as time advances; the time-reversed performance, in which the system departs farther from equilibrium, does not occur. The Boltzmann equation therefore admits of solutions proceeding toward equilibrium but not of time-reversed solutions departing from equilibrium. This is paradoxical because actual physical systems are reversible in time when looked at on an atomic scale. For example, in a classical system the time enters the equations of motion only in the acceleration $d^2\mathbf{r}/dt^2$, so if t is replaced by $-t$ in a solution, a new solution is obtained. If the velocities of all particles were suddenly reversed, the system would retrace its previous behavior.

An actual system does not necessarily move toward equilibrium, although it is overwhelmingly probable that it does so. From a mathematical point of view it is puzzling that one can begin with the exact equations of motion, reversible in time, and by making reasonable approximations arrive at the irreversible Boltzmann equation. The resolution of

this paradox lies in the statistical nature of the Boltzmann equation. It does not describe the behavior of a single system, but the average behavior of a large number of systems. Mathematically, the irreversibility arises from the collision term, Eq. (11), where the approximation has been made that the distribution function $f(\mathbf{r},\mathbf{v},t)$ or $f(\mathbf{r},\mathbf{v}',t)$ applies to particles both immediately before and immediately after a collision.

A number of equations closely related to the Boltzmann equation are often useful for particular applications. If collisions between particles are disregarded, the right side of the Boltzmann equation, Eq. (10), is zero. The equation is then called the collisionless Boltzmann equation or Vlasov equation. This equation has been applied to a gas of charged particles, also known as a plasma. The coulomb forces between particles have such a long range that it is incorrect to consider the particles as free except when colliding. The Vlasov equation can be used by including the forces between particles in the term $\mathbf{a} \cdot (\partial f/\partial \mathbf{v})$. One takes $\mathbf{a} = q\mathbf{E}/m$, q being the charge of a particle and m the mass. The electric field $\mathbf{E}$ includes the field produced by the particles themselves, in addition to any externally produced field.

The Boltzmann equation is a starting point from which the equations of hydrodynamics, Eqs. (12*a*) and (12*b*), can be derived. The particles are now con-

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial \mathbf{r}} \cdot (\rho \mathbf{v}) = 0 \qquad (12a)$$

$$\rho \left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \frac{\partial \mathbf{v}}{\partial \mathbf{r}} \right) = \mathbf{F} - \frac{\partial}{\partial \mathbf{r}} \cdot \mathbf{p} \qquad (12b)$$

sidered as a continuous fluid with density ρ and mean velocity $\mathbf{v}$. $\mathbf{F}$ is the external force per unit volume and $\mathbf{p}$ is the pressure tensor. Equation (12*a*) states mathematically that the mass of the fluid is conserved, while Eq. (12*b*) equates the rate of change of momentum of an element of fluid to the force on it. An energy conservation equation can also be derived. *See* HYDRODYNAMICS; NAVIER-STOKES EQUATIONS.

Applications. The Boltzmann equation can be used to calculate the electronic transport properties of metals and semiconductors. For example, if an electric field is applied to a solid, one must solve the Boltzmann equation for the distribution function $f(\mathbf{r},\mathbf{v},t)$ of the $\mathbf{E}$ electrons, taking the acceleration in Eq. (10) as $\mathbf{a} = q\mathbf{E}/m$ and setting $\partial f/\partial \mathbf{r} = 0$, corresponding to spatial uniformity of the electrons. If the electric field is constant, the distribution function is also constant and is displaced in velocity space in such a way that fewer electrons are moving in the direction of the field than in the opposite direction. This corresponds to a current flow in the direction of the field. The relationship between the current density $\mathbf{J}$ and the field $\mathbf{E}$ is given by Eq. (13), where

$$\mathbf{J} = \sigma \mathbf{E} \qquad (13)$$

σ, the electrical conductivity, is the final quantity of interest. *See* FREE-ELECTRON THEORY OF METALS.

With the Boltzmann equation one can also calculate the heat current flowing in a solid as the result of a temperature difference, the constant of proportionality between the heat current per unit area and the temperature gradient being the thermal conductivity. In still more generality, both an electric field $\mathbf{E}$ and a temperature gradient $\partial T/\partial \mathbf{r}$ can be applied, where T is the temperature. Expressions are obtained for the electrical current density $\mathbf{J}$ and the heat current density $\mathbf{U}$ in the form of Eqs. (14*a*) and (14*b*). L_{11} is

$$\mathbf{J} = L_{11}\mathbf{E} + L_{12}\partial T/\partial \mathbf{r} \qquad (14a)$$

$$\mathbf{U} = L_{21}\mathbf{E} + L_{22}\partial T/\partial \mathbf{r} \qquad (14b)$$

the electrical conductivity and $-L_{22}$ the thermal conductivity. Equations (14*a*) and (14*b*) also describe thermoelectric phenomena, such as the Peltier and Seebeck effects. For example, if a thermal gradient is applied to an electrically insulated solid so that no electric current can flow ($\mathbf{J} = 0$), Eq. (14*a*) shows that an electric field given by Eq. (15) will appear.

$$\mathbf{E} = -(L_{12}/L_{11})(\partial T/\partial \mathbf{r}) \qquad (15)$$

The quantity (L_{12}/L_{11}) is called the Seebeck coefficient.

Finally, if a constant magnetic field $\mathbf{B}$ is also applied, the coefficients L_{ij} in Eqs. (14*a*) and (14*b*) become functions of $\mathbf{B}$. It is found that the electrical conductivity usually decreases with increasing $\mathbf{B}$, a behavior known as magnetoresistance. These equations also describe the Hall effect, the appearance of an electric field in the y direction if there is an electric current in the x direction and a magnetic field in the z direction, as well as more complex thermomagnetic phenomena, such as the Ettingshausen and Nernst effects. The conductivity σ in Eq. (13) and the L_{ij} in Eqs. (14*a*) and (14*b*) are tensor quantities in many materials. If the field $\mathbf{E}$ and the temperature gradient $\partial T/\partial \mathbf{r}$ are in a given direction, the currents $\mathbf{U}$ and $\mathbf{J}$ are not necessarily in the same direction. *See* CONDUCTION (ELECTRICITY); CONDUCTION (HEAT); GALVANOMAGNETIC EFFECTS; HALL EFFECT; MAGNETORESISTANCE; THERMOELECTRICITY; THERMOMAGNETIC EFFECTS.

Nonequilibrium properties of atomic or molecular gases such as viscosity, thermal conduction, and diffusion have been treated with the Boltzmann equation. Although many useful results, such as the independence of the viscosity of a gas on pressure, can be obtained by simple approximate methods, the Boltzmann equation must be used in order to obtain quantitatively correct results. *See* DIFFUSION; VISCOSITY.

If one proceeds from a neutral gas to a charged gas or plasma, with the electrons partially removed from the atoms, a number of new phenomena appear. As a consequence of the long-range Coulomb forces between the charges, the plasma can exhibit oscillations in which the free electrons move back and forth with respect to the relatively stationary heavy positive ions at the characteristic frequency known as the plasma frequency. This frequency is proportional to the square root of the particle density. If the propagation of electromagnetic waves through a plasma is studied, it is found that a plasma reflects

an electromagnetic wave at a frequency lower than the plasma frequency, but transmits the wave at a higher frequency. This fact explains many characteristics of long-distance radio transmission, made possible by reflection of radio waves by the ionosphere, a low-density plasma surrounding the Earth at altitudes greater than 40 mi (65 km). A plasma also exhibits properties such as electrical conductivity, thermal conductivity, viscosity, and diffusion. *See* PLASMA (PHYSICS).

If a magnetic field is applied to the plasma, its motion can become complex. A type of wave known as an Alfvén wave propagates in the direction of the magnetic field with a velocity proportional to the field strength. The magnetic field lines are not straight, however, but oscillate like stretched strings as the wave passes through the plasma. Waves that propagate in a direction perpendicular to the magnetic field have quite different properties.

When the plasma moves as a fluid, it tends to carry the magnetic field lines with it. The plasma becomes partially trapped by the magnetic field in such a way that it can move easily along the magnetic field lines, but only very slowly perpendicular to them. The outstanding problem in the attainment of a controlled thermonuclear reaction is to design a magnetic field configuration that can contain an extremely hot plasma long enough to allow nuclear reactions to take place. *See* MAGNETOHYDRODYNAMICS.

Plasmas in association with magnetic fields occur in many astronomical phenomena. Many of the events occurring on the surface of the Sun, such as sunspots and flares, as well as the properties of the solar wind, a dilute plasma streaming out from the Sun in all directions, are manifestations of the motion of a plasma in a magnetic field. It is believed that a plasma streaming out from a rapidly rotating neutron star will radiate electromagnetic energy; this is a possible explanation of pulsars, stars emitting regularly pulsating radio signals that have been detected with radio telescopes.

Many properties of plasmas can be calculated by studying the motion of individual particles in electric and magnetic fields, or by using the hydrodynamic equations, Eqs. (12*a*) and (12*b*), or the Vlasov equation, together with Maxwell's equations. However, subtle properties of plasmas, such as diffusion processes and the damping of waves, can best be understood by starting with the Boltzmann equation or the closely related Fokker-Planck equation. *See* MAXWELL'S EQUATIONS.

Relaxation time. It is usually difficult to solve the Boltzmann equation, Eq. (10), for the distribution function $f(\mathbf{r},\mathbf{v},t)$ because of the complicated form of $(\partial f/\partial t)_{\text{coll}}$. This term is simplified if one makes the relaxation time approximation shown in Eq. (16),

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = -\frac{f(\mathbf{r},\mathbf{v},t) - f_0(\mathbf{v})}{\tau(v)} \tag{16}$$

where $f_0(v)$ is the equilibrium distribution function and $\pi(v)$, the relaxation time, is a characteristic time describing the return of the distribution function to equilibrium when external forces are removed. For some systems Eq. (16) follows rigorously from Eq. (11), but in general it is an approximation. If the relaxation time approximation is not made, the Boltzmann equation usually cannot be solved exactly, and approximate techniques such as variational, perturbation, and expansion methods must be used. An important situation in which an exact solution is possible occurs if the particles exert forces on each other that vary as the inverse fifth power of the distance between them. *See* RELAXATION TIME OF ELECTRONS.

Example. As an illustration of how the Boltzmann equation can be solved, consider a spatially uniform electron gas in a metal or a semiconductor. Let the scattering of electrons be described by Eq. (16). One can calculate the electric current which flows when a small, constant electric field E_x is applied. The spatial uniformity of the electron gas implies that $\partial f/\partial \mathbf{r} = 0$; since the electric field is constant, one requires also that the distribution function be constant $(\partial f/\partial t = 0)$ so that the current does not depend on time. If one writes the acceleration of the electron as $a_x = F_x{-}m = -eE_x{-}m$, where $(-e)$ is the charge and m is the mass of the electron, Eq. (10) becomes Eq. (17). The difference between f and f_0 is small;

$$\frac{-eE_x}{m}\frac{\partial f}{\partial v_x} = -\frac{f(\mathbf{v}) - f_0(\mathbf{v})}{\tau(v)} \tag{17}$$

therefore f can be replaced by f_0 in the term containing E_x, since E_x itself is small. Thus Eq. (18) is

$$\begin{aligned}\partial f/\partial v_x \approx \partial f_0/\partial v_x &= (\partial f_0/\partial v)(\partial v/\partial v_x)\\ &= (\partial f_0/\partial v)(v_x/v)\end{aligned} \tag{18}$$

obtained, where $v\ (v_x{}^2 + v_y{}^2 + v_z{}^2)^{1/2}$ is the magnitude of the electron velocity; Eq. (17) then becomes Eq. (19). Here $f_0(\mathbf{v})$ is the equilibrium distribution

$$f(\mathbf{v}) = f_0(\mathbf{v}) + \frac{eE_x\tau}{m}\frac{v_x}{v}\frac{\partial f_0}{\partial v} \tag{19}$$

function describing the state of the electrons in the absence of an electric field. For a high electron density, as in metals, $f_0(\mathbf{v})$ is the Fermi-Dirac distribution function given by Eq. (20), where ζ is the Fermi energy, k is the Boltzmann constant, and T is the absolute temperature. In writing $f_0(\mathbf{v})$ in the form of Eq. (20), the simplifying assumption that the energy of an electron is $mv^2/2$ has been made. If the electron density is low, as it often is in semiconductors, $f_0(\mathbf{v})$ can be approximated by the Maxwell-Boltzmann distribution given in Eq. (21). In either case $f_0(\mathbf{v})$ depends only on the magnitude and not on the direction of the velocity. *See* BOLTZMANN STATISTICS; FERMI-DIRAC STATISTICS.

$$f_0(\mathbf{v}) \propto \frac{1}{e^{(mv^2/2-\zeta)/kT} + 1} \tag{20}$$

$$f_0(\mathbf{v}) \propto e^{-mv^2/2kT} \tag{21}$$

Therefore $f_0(\mathbf{v})$ is centered at $\mathbf{v} = 0$ in such a way that the net flow of electrons in every direction is zero. The nonequilibrium distribution function $f(\mathbf{v})$

similar to $f_0(\mathbf{v})$, except that is shifted in the $(-v_x)$ direction in velocity space. Consequently more electrons move in the $(-x)$ direction than in the $(+x)$ direction, producing a net current flow in the $(+x)$ direction. The current density J_x is calculated from $f\mathbf{v}$ by Eq. (22), where the spatial integral is to be taken

$$J_x = -e \int v_x f(\mathbf{v}) d^3v d^3x \tag{22}$$

over a unit volume, so that $\int d^3x = 1$. $f\mathbf{v}$ as given by Eq. (19) is substituted into Eq. (22), the term involving the equilibrium distribution function $f_0(\mathbf{v})$ by itself gives zero current, since no current flows at equilibrium. Hence one obtains Eq. (23).

$$J_x = \frac{e^2 E_x}{m} \int \frac{\tau(v) v_x{}^2}{v} \left(-\frac{\partial f_0}{\partial v}\right) d^3v \tag{23}$$

If n, the density of electrons, is introduced using the equation $n = \int f_0(\mathbf{v})\, d^3v$, Eq. (23) can be written as Eq. (24), where σ is given by Eq. (25), and τ given by Eq. (26), is an average relaxation time.

$$J_x = \sigma E_x \tag{24}$$

$$\sigma = ne^2\overline{\tau}/m \tag{25}$$

$$\overline{\tau} = \frac{\int \tau \frac{v_x^2}{v} \left(-\frac{\partial f_0}{\partial v}\right) d^3v}{\int f_0 d^3v} \tag{26}$$

Equation (25) is the final expression for the electrical conductivity σ.

The above treatment is applicable only if the electrons in the solid behave as if they were nearly free. In general, the wave vector $\mathbf{k}$, rather than the velocity $\mathbf{v}$, must be used as a variable in the distribution function, and the final expression for σ depends on the electronic band structure of the solid. Ronald Fuchs

Bibliography. R. S. Brodkey and H. C. Hershey, *Transport Phenomena: A Unified Approach*, 2000; S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases*, 3d ed., 1970 , reprint 1991; C. Cercignani, *The Boltzmann Equation and Its Applicaions*, 1988; R. O. Dendy (ed.), *Plasma Physics: An Introductory Course*, 1995.

Bond angle and distance

The angle between two bonds sharing a common atom is known as the bond angle. The distance between the nuclei of bonded atoms is known as bond distance. The geometry of a molecule can be characterized by bond angles and distances. The angle between two bonds sharing a common atom through a third bond is known as the torsional or dihedral angle (see **illus.**)

Certain pairs of atoms in a molecule are held together at distances of about 0.075–0.3 nanometer, with energies of about 150–1000 kilojoules/mol, because of a balance between electrostatic attraction and repulsion among the electrons and nuclei, subject to the electron distributions allowed by quantum mechanics. Such interactions are called chemical bonds. *See* CHEMICAL BONDING.

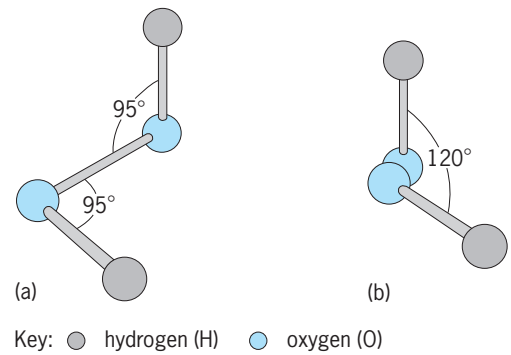

Bond angle versus torsional angle. (*a*) Two O—O—H bond angles in hydrogen peroxide (H_2O_2). (*b*) The same molecule when viewed along the O—O bond, revealing the torsional angle, that is, the angle between the two O—H bonds.

Molecular properties. Bond angles and distances are important because they determine the shape and size of a molecule and therefore affect many physical properties of a substance. The shape of a molecule influences its charge distribution, often determining whether a molecule is polar or nonpolar; polar molecules exhibit stronger intermolecular attraction and thus have higher boiling points. For example, the structures of *cis* and *trans*-1,2-dichloroethylene differ only in the torsional angle about the C═C bond. The polar cis isomer has a normal boiling point (60.3°C) considerably higher than that of the nonpolar trans isomer (47.5°C). Molecular shapes also influence the packing density of molecules in solids and liquids. For example, spherical molecules pack differently from long, rod-shaped molecules. For biological molecules, the detailed shape of the molecule ultimately governs its function for example, deoxyribonucleic acid (DNA), or its specificity (for example, enzymes). *See* DEOXYRIBONUCLEIC ACID (DNA); ENZYME; MOLECULAR STRUCTURE AND SPECTRA; POLAR MOLECULE.

Bonding data. Bond angles and distances depend on the identity of the atoms bonded together and the type of chemical bonding involved. For example, carbon–hydrogen bonds are remarkably similar in length in a wide range of compounds. Bonds vary in length depending on the bond multiplicity: they are shorter for a triple bond than for a single bond (**Table 1**). Single bonds vary in length slightly, depending on whether they are adjacent to multiple bonds (**Table 2**). Bond angles vary considerably from molecule to molecule; the variation in bond angles and distances depends primarily on the electronic structures of the molecules. **Table 3** shows the variation in bond angles for a number of triatomic molecules.

The bond lengths and angles of an electronically excited molecule usually differ from the ground-state values. For example, carbon dioxide is linear in its ground state but is bent in its lowest excited state; that is, the O—C—O angle is 180° in the ground state

TABLE 1. Dependence of bond length on bond multiplicity

Bond	Molecule	Name	Length, nm*
C—C	$H_3C—CH_3$	Ethane	0.153
C═C	$H_2C═CH_2$	Ethylene	0.134
C≡C	HC≡CH	Acetylene	0.120
N—N	$H_2N—NH_2$	Hydrazine	0.145
N═N	N═N═O	Nitrous oxide	0.113
N≡N	N_2	Nitrogen	0.110
C—O	$H_3C—OH$	Methanol	0.142
C═O	$H_2C═O$	Formaldehyde	0.121
C≡O	CO	Carbon monoxide	0.113

*Uncertainties are $\leq$ 0.001 nm in bond distances and $\leq 1^\circ$ in bond angles, unless otherwise stated.

TABLE 2. Variation of carbon–hydrogen bond length with neighboring carbon–carbon bond multiplicity

Molecule	C—H bond length, nm
$H_3C—CH_3$	0.110
$H_2C═CH_2$	0.108
HC≡CH	0.106

TABLE 3. Bond angles for three triatomic molecules

Molecule	Name	Angle	Value
H_2O	Water	H—O—H	105°
SO_2	Sulfur dioxide	O—S—O	119°
CO_2	Carbon dioxide	O—C—O	180°

but is $122 \pm 2^\circ$ in the first excited state. The C═O bond distance is 0.116 nm in the ground state but is 0.125 nm in the first excited state.

Experimental determination. Much experimental effort has gone into measuring bond angles and distances, and values for individual molecules have been extensively compiled. The experimental methods can be divided into diffraction methods and spectroscopic methods.

Diffraction methods. These employ the wave properties of x-rays, electrons, and neutrons. When the wavelength of x-rays or the de Broglie wavelength of electrons or neutrons is comparable to certain interatomic distances in a sample, then the waves or particles will be diffracted through an angle determined by the ratio of the wavelength to the interatomic distance. Gases are studied with electron diffraction, while solids and liquids are studied with x-ray and neutron diffraction. *See* DE BROGLIE WAVELENGTH; ELECTRON DIFFRACTION; X-RAY DIFFRACTION.

Spectroscopic methods. These rely primarily on the rotational motion of molecules in the gas phase. Rotational energies are quantized and depend on the molecule's moments of inertia. These moments depend on bond distances and angles, as well as nuclear masses. Rotational transitions are observed directly in microwave absorption spectroscopy, or as rotational fine structure in infrared, visible, and ultraviolet absorption spectroscopy. For many nonpolar molecules that do not absorb microwave radiation, the Raman effect can be employed to detect rotational or vibrational transitions. *See* RAMAN EFFECT; SPECTROSCOPY.

In liquids and solids, where molecules cannot rotate freely, spectroscopic structural information comes primarily from infrared and vibrational Raman spectra. Often, a comparison of the two provides qualitative symmetry information. For example, a mutual exclusion principle states that if a molecule has a center of symmetry, no vibrational mode will appear in both the infrared and Raman spectrum.

Theoretical calculation. Quantum mechanics can, in principle, be used to calculate bond lengths and angles accurately by solving the appropriate quantum-mechanical equations, although such calculations for large molecules are extremely difficult and require a great deal of computer time. A number of simple theoretical concepts have been developed that have some approximate correlation with experimentally determined bond angles and distances. One is called the electrostatic or valence-shell electron-pair repulsion model. It treats a chemical bond as a shared pair of electrons which repels other pairs associated with the same atom. An alternative model, valence-bond theory, treats a chemical bond as an overlap of electronic orbitals centered on the bonded atoms. Equivalent hybrid orbitals centered on the same atoms have a characteristic geometry. *See* QUANTUM CHEMISTRY; STRUCTURAL CHEMISTRY; VALENCE.

Bruce A. Garetz

Bibliography. K. Kuchitsu (ed.), *Structure Data of Free Polyatomic Molecules*, vol. 7 of Landolt-Bornstein, Group II, *Atomic and Molecular Physics*, 1995; I. Levine, *Physical Chemistry*, 4th ed., 2000; J. N. Murrell et al., *The Chemical Bond*, 2d ed., 1985; L. Pauling, *The Nature of the Chemical Bond*, 3d ed., 1960.

Bonding

A method of holding the parts of an object together with or without the aid of an adhesive such as an epoxy or a glue. Composite materials such as fiber-reinforced plastics require strong interfacial bonding between the reinforcement and the matrix. In the case of atomic or optical contact bonding, interatomic forces hold the parts together. In optical contact bonding, surface flatness and cleanliness between the mating parts determine the bonding strength, and the atoms at the surface provide the necessary forces. The number of valence electrons in the atoms of a material determines the bonding strength between the group of atoms which constitutes a molecule. In these cases the term chemical bonding is used. *See* ADHESIVE; CHEMICAL BONDING; COMPOSITE MATERIAL.

Wire bonding is an interconnect technique widely used in microchip manufacturing to provide electrical continuity between the metal pads of an integrated circuit chip and the electrical leads of the package housing the chip. The two common methods of wire bonding are thermocompression and

ultrasonic bonding. In these, a fine aluminum or gold wire is bonded at one end to the metal pad of the integrated circuit chip, and at the other to the electrical lead of the package. There are three types of thermocompression bonds: wedge, stitch, and ball. In thermocompression bonding, a molecular metallurgical bond is formed at the two metal junctions—bond wire and IC metal pad, and bond wire and package lead metal—by applying heat and pressure without melting. In ultrasonic bonding, the molecular metallurgical bond is achieved through a combination of ultrasonic energy and pressure. The bonding operation is done under pressure to break the few surface layers of the material and form the bond between the contamination-free surfaces. Thermocompression bonding has higher throughput and speed than ultrasonic bonding. The bonding wire is usually aluminum, which does not introduce any intermetallic problems. *See* CIRCUIT (ELECTRONICS); INTEGRATED CIRCUITS. Lakshmi Munukutla

Bone

The hard connective tissue that, together with cartilage, forms the skeleton of humans and other vertebrates. It is made of calcium phosphate crystals arranged on a protein scaffold. Bone performs a variety of functions: it has a structural and mechanical role; it protects vital organs; it provides a site for the production of blood cells; it serves as a reserve of calcium. Bone tissue is constantly renewing itself and has great capacity to respond to altered stresses and loads and to repair fractures. The formation and remodeling of bone is accomplished by the action of bone cells of the osteoblast and osteoclast lineages. The activities of these cells are directed by hormones and other molecules that circulate in the blood as well as regulators that are produced locally in the bone. There are a number of bone diseases which are the subject of vigorous research. *See* CONNECTIVE TISSUE; SKELETAL SYSTEM.

Composition and structure. There are two types of bone in the skeleton: the flat bones (for example, the bones of the skull and ribs) and the long bones (for example, the femur and the bones of the hand and feet). Although the growth and development of the two bone types are different, both types are characterized by an outer layer of dense, compact bone, known as cortical bone, and an inner spongy bone material made up of thin trabeculae, known as cancellous bone. Cortical bone consists of layers of bone (lamellae) in an orderly concentric cylindrical arrangement around tiny Haversian canals. These interconnecting canals carry the blood vessels, lymph vessels, and nerves through the bone and communicate with the periosteum and the marrow cavity. The periosteum is a thin membrane covering the outer surface of bone and consisting of layers of cells that participate in the remodeling and repair of bone. The cancellous bone is in contact with the bone marrow, in which much of the production of blood cells takes place. The interface between the cancellous bone and the marrow is called the endosteum, and it is largely at this site that bone is removed in response to a need for increased calcium elsewhere in the body. A dramatic example is seen in the egg-laying fowl, where trabecular bone is rapidly removed to provide calcium for the egg shell and is then replaced in a cyclical manner.

Bone is formed by the laying down of osteoid (the young hyaline matrix of truebone) by osteoblasts, the bone-forming cells, and the mineralization of the osteoid by the development and deposition of crystals of calcium phosphate (in the form of hydroxyapatite) within it. It is the mineral, organized in a regular pattern on a collagen scaffold, that gives bone its stiffness. Osteoid contains largely fibers of type I collagen (a member of a large family of collagen molecules) and lesser amounts of numerous noncollagenous proteins. These proteins include osteonectin, osteopontin, bone sialoprotein, biglycan, decorin, and osteocalcin. Although the role of these proteins in bone is not well understood, it is thought that their particular combination in bone gives this tissue the unique ability to mineralize. It is clear that these proteins interact with each other and that collagen and several of the noncollagenous proteins, such as osteopontin and bone sialoprotein, can bind to specialized receptors on the surface of bone cells. This binding is important for the adhesion of the cells to the bone matrix, and it also delivers behavioral signals to the cells. *See* APATITE; COLLAGEN.

Bone cells. Bone is rich in blood vessels and nerves and thus contains the cell types associated with these structures. However, the primary cell types in bone are those that result in its formation and maintenance (osteoblasts and osteocytes) and those that are responsible for its removal (osteoclasts). Osteoblasts form from the differentiation of multipotential stromal cells that reside in the periosteum and the bone marrow. Under the appropriate stimuli, these primitive stromal cells mature to bone-forming cells at targeted sites in the skeleton. Under different stimuli, they are also capable of developing into adipocytes (fat cells), muscle cells, and chondrocytes (cartilage cells). Several developmental stages can be recognized in the differentiation of osteoblasts from their precursor cells: preosteoblasts; synthetic-phase osteoblasts that secrete the proteins that constitute osteoid; flattened lining cells that cover the endosteal surface when no bone formation or resorption is taking place; and osteocytes, which are osteoblasts that become incorporated within the bone tissue itself. Osteocytes are the most numerous cell type in bone. They reside in spaces (lacunae) within the mineralized bone, forming numerous extensions through tiny channels (cannaliculi) in the bone that connect with other osteocytes and with the cells on the endosteal surface. Osteocytes are therefore ideally placed to sense stresses and loads placed on the bone and to convey this information to the osteoblasts on the bone surface, thus enabling bone to adapt to altered mechanical loading by the formation of new bone. Osteocytes are also thought to be the cells that detect and direct the repair of microscopic damage

that frequently occurs in the bone matrix due to wear and tear of the skeleton. Failure to repair the cracks and microfractures that occur in bone, or when this microdamage accumulates at a rate exceeding its repair, can cause the structural failure of the bone. This is seen, for example, in athletes where repeated loading of the bone leads to stress fractures.

A large number of molecules that regulate the formation and function of osteoblastic cells have been identified. The products of particular homeobox genes, which determine what tissue type the primitive cells will become, are especially important in the embryological development of the skeleton. Subsequently, circulating hormones, such as insulin, growth hormone, and insulinlike growth factors, combine with growth factors within the bone itself, such as transforming growth factor beta (TGFβ) and bone morphogenetic proteins (BMPs), to influence the differentiation of osteoblasts for the growth and repair of bone. During bone growth and in fracture repair, the sequential response of osteoblasts and their precursors to these influences controls each step in the cell differentiation pathway as well as the secretory activity of the mature osteoblasts.

Osteoclasts are typically large, multinucleated cells, rich in the intracellular machinery required for bone resorption. This is accomplished when the cells form a tight sealing zone by attachment of the cell membrane against the bone matrix, creating a bone-resorbing compartment. Into this space, the cell secretes acid to dissolve the bone mineral, and enzymes to digest the collagen and other proteins in the bone matrix. The removal of bone by osteoclasts is necessary to enable the repair of microscopic damage and changes in bone shape during growth and tooth eruption. Osteoclast-mediated bone resorption is also the mechanism for releasing calcium stored in bone for the maintenance of calcium levels in the blood. The level of calcium in the blood is monitored by the parathyroid gland and by the cells of the thyroid. If the level drops too low, parathyroid hormone is released to prevent the loss of calcium through the kidneys and to mobilize calcium from bone. If the blood calcium level is too high, the hormone calcitonin is released from the thyroid, which rapidly inhibits the resorptive activity of osteoclasts. Osteoclasts arise from precursor cells of the monocyte–macrophage series that are found in the bone marrow and circulate in the blood. Most agents that promote bone resorption act on osteoblastic cells, which in turn convey signals to osteoclast precursors to differentiate into mature osteoclasts. These agents include the active form of vitamin D, parathyroid hormone, interleukin-1, interleukin-6, and interelukin-11, and prostaglandins such as prostaglandin E_2. Differentiation to fully functional osteoclasts also requires close contact between osteoclast precursors and osteoblastic cells. This is due to a molecule called osteoclast differentiation factor (ODF) which is located on the surface of osteoblasts, binds to receptors on the surface of osteoclast precursor cells, and induces their progression to osteoclasts. The activity of ODF can be inhibited by a second molecule found in bone and elsewhere and termed osteoprotegerin or osteoclastogenesis-inhibitory factor (OCIF), which inhibits both the formation of osteoclasts and the activity of mature osteoclasts. Since OCIF acts as a naturally occurring ODF antagonist, it is most likely the local ODF/OCIF ratio that determines the local number and activity of osteoclasts. Thus, control of the concentrations of ODF and OCIF in various parts of the skeleton represents a possible mechanism for targeting the removal of bone to appropriate sites.

Bone formation, growth, and remodeling. The two types of bone (flat bones and long bones) are formed by different embryological means. Formation of flat bones occurs by intramembranous ossification, in which primitive mesenchymal cells differentiate directly into osteoblasts and produce bony trabeculae within a periosteal membrane. The initial nature of this bone is relatively disorganized and is termed woven bone. Later, this woven bone is remodeled and replaced by the mature lamella bone, consisting of layers of calcified matrix arranged in orderly fashion. Lamella bone has much greater strength characteristics than woven bone. Long bones are formed by intracartilaginous development in which the future bone begins as cartilage. The cartilage template is gradually replaced by bone in an orderly sequence of events starting at the center of the growing bone (the diaphysis). Cartilage remains at the ends of long bones during growth, forming a structure at each end termed the growth plate. Cartilage cells (chondrocytes) that arise in the growth plates proliferate and add to the length of the bone. This occurs during a complex series of events, with expansion away from the center of the bone to create the epiphyses, and expansion toward the center of the bone to create the metaphyses. (see **illus.**). When the bone achieves its final length in maturity, expansion from the growth plate ceases. Cartilage persists at the ends of the long bones in a specific form called articular cartilage, which provides the smooth bearing surfaces for the joints.

Bone is a dynamic tissue and is constantly being remodeled by the actions of osteoclasts and

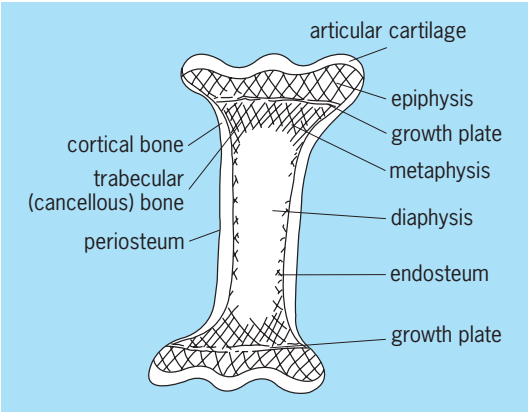

Diagramatic representation of a long bone.

osteoblasts. The removal of bone by osteoclasts occurs both on the surface of trabecular bone and within cortical bone, at random sites and at sites of microscopic damage. After bone removal, the osteoclasts either move on to new resorption sites or die; this is followed by a reversal phase where osteoblasts are attracted to the resorption site. It is thought that growth factors that are sequestered in an inactive form in the bone matrix are released and activated by the osteoclast activity and that these in turn promote fresh osteoid production by the recruited osteoblasts. The new osteoid eventually calcifies, and in this way the bone is formed and replaced in layers (lamellae), which are the result of these repeated cycles. In growing bone, the activities of bone cells is skewed toward a net increase in bone. However, in healthy mature bone there is an equilibrium between bone resorption and bone formation, due to communication between osteoblasts and osteoclasts. The nature of the signals between these two cell types is not completely understood, although, osteoclast formation and activation is directed in large part by osteoblasts. When the equilibrium between these two cell types breaks down, skeletal pathology results.

The most common bone disease is osteoporosis, in which there is a net loss of bone due to osteoclastic bone resorption that is not completely matched by new bone formation. The best-understood cause of osteoporosis is that which occurs in women due to the loss of circulating estrogen after menopause. A prolonged loss of bone in postmenopausal osteoporosis reduces bone strength, and it is the cause of a great deal of morbidity in the form of fractures, largely of the spine and femur. It is known that the frequency of this condition differs among racial groups. Exogenous estrogen, high-calcium diets, and exercise are some of the current treatments to counter the bone loss of postmenopausal osteoporosis. Another cause of osteoporotic bone loss is seen in disuse osteoporosis and in space flight. Just as bone can respond to increased loading with the production of additional bone, bone is also dependent on regular loading for its maintenance. Significant bone loss can occur during prolonged bed rest or, for example, in paraplegia and quadriplegia. Likewise, an unloading of the skeleton (due to a lack of gravitational pull) in space flight results in severe bone loss in astronauts unless the effects of gravity are simulated by special exercises and devices. *See* OSTEOPOROSIS.

Many metabolic and genetic diseases can affect the amount and quality of bone. Metabolic diseases such as diabetes, kidney disease, oversecretion of parathyroid hormone by the parathyroid glands, anorexia nervosa, and vitamin D–dependent rickets may cause osteopenias (the reduction in bone volume and bone structural quality). In addition, immunosuppressive therapy in organ transplant patients can lead to reduced bone mass. Tumors of bone and other sites can also lead to bone loss. Bone tumors include those that arise in bone and those that metastasize to bone from primary tumors in sites such as breast and prostate. Tumors can produce substances that cause the activation of osteoclastic bone resorption, which can result in the destruction of bone and elevated levels of calcium in the blood. Two examples of genetically based diseases of bone are osteopetrosis and osteogenesis imperfecta. In the latter, mutations in the gene for type I collagen result in the production of reduced amounts of collagen or altered collagen molecules by osteoblasts, with a reduced capacity to serve as a support for bone mineral. Osteopetrosis is a rare set of conditions characterized by increased bone density due to a defect in the number or activity of osteoclasts. Other common diseases of the skeleton are diseases of the joints. Rheumatoid arthritis is an autoimmune disease characterized by inflammation of the joints and subsequent destruction of the cartilage and bone in the joints. It can be treated by management of the inflammation. The causes of osteoarthritis are less clear, but this too is characterized by the gradual destruction of the articular cartilage. When the joints become nonfunctional in these disease states, it is now common practice to surgically replace the joints with artificial prostheses. *See* CALCIUM METABOLISM; PARATHYROID GLAND; THYROID GLAND.

David M. Findlay

Bibliography. M. Favus (ed.), *Primer on the Metabolic Bone Diseases and Disorders of Mineral Metabolism*, 3d ed., Lippincott-Raven Press, Philadelphia, 1996; R. Marcus, D. Feldman, and J. Kelsey (eds.), *Osteoporosis*, Academic Press, New York, 1996; R. B. Martin and D. B. Burr (eds.), *Structure, Function, and Adaptation of Compact Bone*, Raven Press, New York, 1989.

Bone disorders

Disturbances in the structure and function of the skeletal system. The musculoskeletal system provides a rigid structural frame that permits locomotion through muscle action and aids in the functioning of the other organ systems. It also houses blood-forming marrow tissue and stores calcium, phosphorus, magnesium, and a variety of other ionic species. Bone is unique among the tissues of the body in that it responds to injury by re-forming itself identically over time, without scarring.

Trauma. Trauma to the musculoskeletal system can produce a closed or open fracture, a joint dislocation, or a sprain. The incidence of these injuries is highest among males 17–24 years of age. In the United States the annual rate of these injuries in the general population exceeds 20 million, with fractures constituting about one-fourth of the total. Fractures are manifested clinically by significant pain, swelling, and false motion, and in long-bone injuries by crepitation. Dislocation occurs more commonly in the shoulder and fingers than in any other joints. Injuries to the ligaments that maintain joint alignment are known as sprains and occur most frequently at the ankle.

The diagnosis of dislocation can frequently be made on clinical examination, but roentgenograms (x-rays) provide definitive evidence. Fractures are

also diagnosed most frequently with routine roentgenograms, but special evaluation to determine the extent of the fracture or plan for surgical treatment requires computed tomography or magnetic resonance imaging. Operative treatment of long-bone fractures has become well accepted, particularly when multiple injuries are involved. Pelvic and spinal fractures are also treated with surgical correction and fixation. The traditional treatment approach of skeletal traction and casting has become less popular not only because it is more costly but also because the alternatives allow better results through earlier return to function. Even open fractures—after removal of foreign matter and damaged soft tissue and bone—are treated by internal fixation. *See* MEDICAL IMAGING; RADIOGRAPHY.

Bone healing begins with cellular proliferation and the formation of new blood vessels. These blood vessels penetrate the blood clot that fills the fracture area and provide provisional stabilization through the formation of a cartilaginous matrix. Calcification of this matrix produces a disorganized bony architecture (callus) that is larger than the original bone and provides mechanical stability at the fracture site until final internal remodeling takes place. Depending on the bone involved, this remodeling process requires weeks, months, or years for completion. Ultimately, the fracture callus is resorbed and the bone is reshaped so that it conforms to its original architecture. The extent and speed of this process depend on both mechanical stress and the age of the individual. *See* BONE.

Neoplasms. Primary neoplasms of bone and connective tissue are relatively uncommon compared with other musculoskeletal conditions; approximately 60% of all primary bone neoplasms are malignant. The most frequently occurring malignant neoplasms are multiple myeloma, osteosarcoma, chondrosarcoma, and Ewing's tumor. In young individuals the most common site of a neoplasm is the tibia or femur in the vicinity of the knee, while in older individuals the flat bones, such as the ilium and those of the axial skeleton, are frequently affected. The annual incidence of osteosarcoma, chondrosarcoma, and Ewing's tumor is 1.4–2.5 per 1 million population. Multiple myeloma, which arises from plasma cells in the bone marrow and has a higher annual incidence of 6–7 per 100,000 population, is also more prevalent among older persons. The incidence of benign bone tumors is less certain because individuals with these tumors frequently do not undergo treatment.

Benign lesions have many names and are also classified according to the type of tissue primarily affected (bone, cartilage, fibrous tissue). For example, cartilaginous tumors include the enchondroma and the exostosis, both of which may either results from hereditary factors or occur sporadically. The enchondroma is a lesion of cartilage within the bone; an exostosis is a lesion of cartilage that grows away from bone on a bony pedicle, usually near a joint. Bony lesions include the osteoblastoma and the osteoid osteoma; both of these have the histologic appearance of bone and are benign. Fibrocystic lesions include unicameral bone cysts, fibrous cortical defects, fibrous dysplasia, and giant-cell tumors. Fibrous dysplasia is a developmental lesion that often is characterized by multiply involved bones and is associated with café-au-lait spots on the skin. Unicameral bone cysts occur preferentially on the proximal humerus in adolescent children. They cause thinning of the bone and, occasionally, pathologic fractures.

Frequently, malignant tumors are identified when the individual complains of pain or swelling, or has a pathologic fracture. In contrast, many of the benign lesions are noted when routine roentgenographic examination is performed after trauma. Evaluation with various imaging techniques, including roentgenography, computed tomography, magnetic resonance imaging, and radionuclide bone scanning, is used to determine the type and extent of the tumor. If malignancy is suspected, a biopsy is performed to confirm the diagnosis. Benign tumors are treated by surgical removal and bone grafting when necessary, but malignant tumors require more aggressive treatment. If possible, the tumor is removed either by amputation or by limb salvage procedures. In the latter, the tumor and a margin of normal tissue are completely removed surgically and replaced with a prosthesis. This prosthesis, which is used to restore the associated joint or joints, is usually constructed from metal and plastic and is sometimes accompanied by bone allograft, that is, bone obtained from an unrelated donor. Chemotherapy and irradiation are used in conjunction with surgical treatment to provide additional protection against recurrence of the tumor.

Tumors arising in other organ systems, such as the gastrointestinal tract or the lungs, frequently spread to bone. These malignancies, while not bone tumors in the strict sense, affect the structural integrity of the bone and may therefore cause fractures and require treatment. In persons over 40 years of age, they occur more frequently than primary bone tumors. Because metastasis has such serious consequences, treatment of these lesions prior to fracture frequently involves radiation therapy. After fracture, treatment consists of internal fixation with the adjunct of acrylic bone cement, followed by irradiation. *See* CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; ONCOLOGY; PROSTHESIS; TUMOR.

Metabolic bone disease. Metabolic bone disease is an alteration in the normal bone metabolism. Bone structure is maintained by two types of cells: osteoclasts, responsible for removing damaged or old bone, and osteoblasts, responsible for building new bone. In various metabolic bone diseases, there are disturbances in the balance between osteoblastic and osteoclastic activity that alter the stability of skeletal structures. For example, a deficiency of vitamin C results in an inability to produce the protein matrix (osteoid) upon which bone mineral is laid by the osteoblast. Vitamin D deficiency results in an inability to mineralize the osteoid, the consequences of which are rickets in children and osteomalacia in adults. The incidence of vitamin D deficiency is quite

low in developed countries, but there are various syndromes of rickets that do not respond to vitamin D. Rickets is characterized by shortness of stature and angular deformities of the weight-bearing bones. Osteomalacia is rare but can result when vitamin D is not absorbed through the gastrointestinal tract. It is associated with significant and widespread pain and tenderness of the bones and deformity of the spine and limbs. *See* ASCORBIC ACID; VITAMIN D.

Osteoporosis, the most common metabolic bone disease, occurs in a variety of diseases, including hyperparathyroidism, hyperthyroidism, hyperpituitarism, and hyperadrenocorticism. Osteoporosis in older persons can lead to fracture; thoracic vertebral bodies, the hip, and the wrist are the sites most commonly affected. While postmenopausal and senile osteoporosis results from insufficient bone production (decreased osteoblastic activity) and excessive bone destruction (increased osteoclastic activity), the precise etiology of this disorder is unknown. Hormonal factors and mechanical stress are known to play a role. *See* CALCIUM METABOLISM; OSTEOPOROSIS.

Osteomyelitis. Osteomyelitis is an infection of bone that is the result of either the blood-borne spread of an infectious agent or the secondary contamination of a fracture that has penetrated the skin. The blood-borne variety is a disease of growing bones, is more common in boys than in girls, and occurs most frequently in the end of the femur, tibia, or humerus. The bacterium *Staphylococcus aureus* is the most common causative agent; other bacteria, including the tuberculosis mycobacterium, as well as fungi also (rarely) cause osteomyelitis. The source of the infection is usually a skin lesion, although respiratory or urinary tract infections can also be a source. Acute osteomyelitis is characterized by severe bone tenderness and an unwillingness to use the limb, accompanied by loss of appetite and a fever. Osteomyelitis that results from an open fracture develops more slowly and results in a chronically draining wound, and healing of the fracture is either delayed or fails to occur at all. Treatment of either entity requires antibiotic therapy for a period of up to 6 weeks. Frequently, surgical intervention is needed to reduce an abscess and remove necrotic tissue.

Septic arthritis. Although infection of a joint can result from direct inoculation with an infectious agent, septic arthritis is most often a blood-borne condition. The incidence of this disease entity is higher in children than in adults, with the former affected primarily in the hip and elbow. As with osteomyelitis, *S. aureus* is most often responsible, while streptococci, pneumococci, and other bacteria are present less frequently. In adults, gonorrhea may be a source of joint infection. An individual with an infected joint usually experiences muscle spasm, severe tenderness and pain with the slightest movement, elevated temperature, and swelling of the joint. Immediate treatment is necessary to prevent permanent damage to the joint through the activity of enzymes generated by the infection. Surgical irrigation of the joint is frequently necessary, and optimal results are obtained if this is combined with appropriate antibiotic therapy.

Osteoarthrosis. Osteoarthrosis is a premature or excessive deterioration of the cartilage surface of a joint. This type of change occurs in all adults, but usually not to the extent that it causes symptoms. The local deterioration of the cartilage surface is associated with remodeling of the contiguous bone and subsequent secondary inflammation of the membrane lining the joint. When it reaches this stage, it is referred to as osteoarthritis (or degenerative arthritis, degenerative joint disease, or senescent arthritis).

Rheumatoid arthritis. Rheumatoid arthritis is one of a family of inflammatory polyarthritides. The condition is thought to be caused by an abnormality in the immune system in which the body produces antibodies against its own tissues. In addition to the joints, other organ systems are often involved. Nonsteroidal anti-inflammatory medications such as aspirin form the cornerstone of treatment; other drugs that may cause serious side effects are used less frequently only to control the symptoms not adequately managed by the former medications. Surgical treatment is aimed primarily at replacing the joints that have been damaged and do not respond to medication. The unpredictable exacerbations, remissions, and variable response to medication make rheumatoid arthritis difficult to treat. *See* ARTHRITIS; JOINT DISORDERS; SKELETAL SYSTEM. Harry B. Skinner

Bibliography. J. H. Herndon, Orthopaedic surgery, *American College of Surgeons Bull.*, 74(2):18–24, 1989; T. L. Holbrook et al., *The Frequency of Occurrence, Impact, and Cost of Selected Musculoskeletal Conditions in the United States*, American Academy of Orthopedic Surgeons, 1984; D. J. McCarty (ed.), *Arthritis and Allied Conditions: A Textbook of Rheumatology*, 11th ed., 1989; R. B. Salter, *Textbook of Disorders and Injuries of the Musculoskeletal System*, 2d ed., 1983.

Bonsai

The construction of a mature, very dwarfed tree in a relatively small container. Although bonsai is formally an esoteric branch of horticulture, it is firmly based in the social, historical, and cultural ethic of Asian peoples and its introduction to Western cultures is a development of the twentieth century; indeed, Westerners had not even seen bonsai plants prior to the early decades of this century.

At the fundamental level, the word bonsai means merely "plant in a pot," and its origin can be traced back to China's Chou dynasty (900–250 B.C.), when emperors built miniaturized gardens in which soil, rocks, and plants from each province were installed as dwarfed representations of the lands held by the central government. These gardens included *punching* (potted plants within landscapes) and *shea tse ching* (plants in tubs) or, more commonly, *pun-sai* (plants in small containers). Selected species were usually those with mythic, symbolic, and esthetic significance—plum, cypress, and pines

plus bamboo—or plants representing seasons of the year—peony, cherry, and maple. The introduction of Buddhism into China, with the religious stress on self-awareness and a concept of the eternal and agelessness as exemplified by trees—concepts in keeping with tenets of Confucianism—shifted the emphasis of dwarfed potted trees from its political and esthetic aspects toward that of the construction as a representation and visual image of the eternal. The Buddhist cult of Chan (the Japanese Zen) utilized *pun-sai* as objects of contemplation to be viewed during the meditations that constituted a basic part of Chan ritual.

Up to the eleventh century A.D., China was the cultural and governmental model for most Asian cultures, and other nations sent officials, priests, and artists to China to learn the intellectual and practical skills which could be introduced into their own countries. Early Japanese visitors commented on the *pun-sai*, and some were carried back to Japan by the twelfth century. During the Kamakura period (1192–1333), a significant commerce in dwarfed plants was noted in official Japanese documents, and bonsai (a direct transliteration of *pun-sai*) were included in painting. Books and plays of the fourteenth century included bonsai as main or subsidiary themes, facilitating the diffusion of knowledge and appreciation of the plants and their significance. Bonsai in Japan received a major impetus when, in the seventeenth century, a Chinese official became disaffected with the Ming court and decamped to Japan, taking an extensive collection of plants and manuals on their construction and maintenance. A fairly complete historical record of bonsai styles, species used, and types of containers can be followed in Japanese woodblock prints dating from the sixteenth century.

In spite of the long-standing interest in bonsai in Japan and growing interest in Western nations, there have been surprisingly few controlled scientific studies related to the plants. As has long been true of most ornamental horticulture throughout the world, techniques and concepts have been passed by word of mouth or through apprenticeships, rather than having been codified and tested under controlled scientific conditions.

Esthetic aspects. There are two major aspects involved in the construction and maintenance of a bonsai plant, only one of which is susceptible to direct study. Esthetic aspects, which partake of the cultures from which the form is derived plus individual tastes, are not easily analyzed, although they undoubtedly play a more important role than do technological aspects.

Over several centuries the Japanese codified not only the plants to be used as bonsai subjects, but also the styles and sizes of the plants in their final form. In general, styles are based on observable natural growth patterns in various ecological situations. A solitary tree growing in an open field will develop a much different form than a genetically identical plant surrounded by other trees in a forest, and a tree exposed to constant winds coming from one direction will grow differently from one not so exposed. Major styles are shown in the **illustration**, and others have also been recognized. The rules for styles are flexible and variations are common. In Japan, some individual bonsai plants have been known as that particular plant for several centuries and are

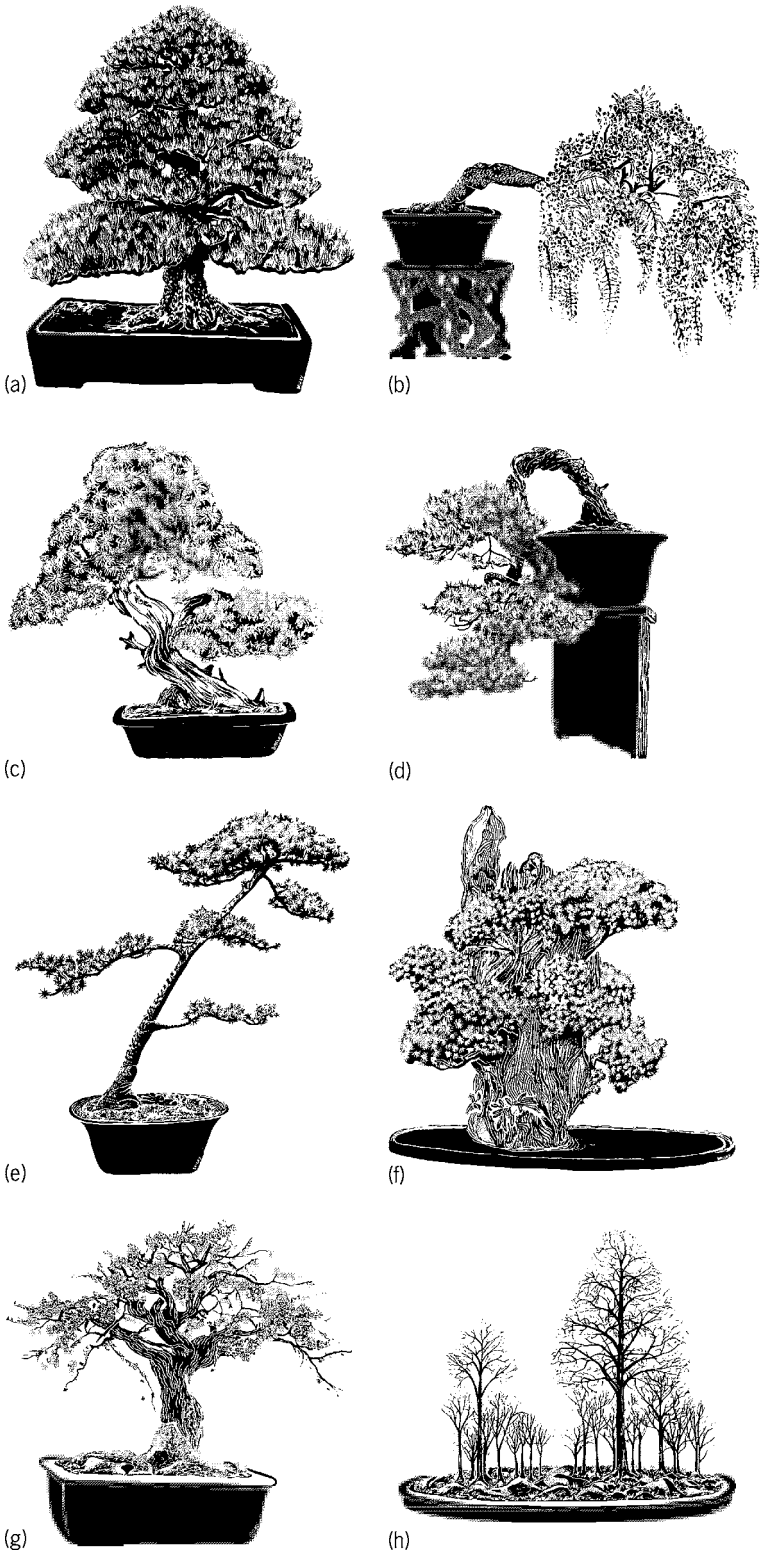

Major bonsai styles. (*a*) Formal upright (*Chokkan*); black pine. (*b*) Semicascade (*Han-Ken-gai*); wisteria. (*c*) Informal upright (*Moyogi*); juniper. (*d*) Cascade (*Kengai*); white pine. (*e*) Slanting (*Shakan*); spruce. (*f*) Rock clinging (*Ishezuke*); azalea. (*g*) Domed (*Gaito-Kengai*); plum. (*h*) Forest (*Yose-uye*); beech.

emulated and honored. Sizes, too, have been codified. Plants less than 15 cm (6 in.; pot included) are called *mame-bonsai* (hold-in-the-palm-of-one-hand), those from 15 to 30 cm (6 to 12 in.) tall are *ko-bonsai* (hold-in-one-hand), plants 30 to 60 cm (12 to 24 in.) tall are *chie-bonsai* (hold-with-two-hands), and those taller than 60 cm (24 in.) are *dai-bonsai* (two-people-needed).

The pot or container is also a matter of esthetic concern. Form, color, size, and decoration are selected to complement, enhance, or contrast with the plant. Until the eighteenth century, containers were usually Chinese antiques and correspondingly rare and expensive, but the growing enthusiasm for bonsai as both a religious form and a secular hobby caused the development of indigenous ceramic industries devoted to the manufacture of containers. These are usually clay or porcelain, high-fired and appropriately glazed, usually in earth tones of brown, blue, green, or rust. Sharp reds, purples, or other bright colors are usually avoided, although there are important exceptions.

Since, in Japan, bonsai was at first restricted to the priests and to associated nobility, monks were usually trained to make and maintain the plants. As interest rose, individuals adept at the procedures began to develop their own nurseries and to conduct classes on the technique. Some families have continued this tradition and, with formal examinations, can be designated as bonsai masters. Many have settled in Omiya, a suburb of Tokyo, where they not only keep large bonsai nurseries and teach, but also may keep precious individual plants for families who will have a plant delivered for viewing on a special holiday or for a few days when the plant is in bloom. Such viewings, mostly among the wealthy, are occasions for a tea ceremony. A national viewing and competition is held yearly under the patronage of the Emperor.

Technological aspects. To a large extent, the technology of bonsai is derived from classical horticultural procedures which are the product of long tradition firmed up with considerations of many areas of plant physiology and ecology. For optimum growth of any plant, adequate levels of water, air, nutrients, light, and temperature must be presented. Good horticultural practice is as necessary in bonsai as it is in any other aspect of gardening, although the restrictions of small containers and carefully controlled growth patterns necessitate additional care.

There are, however, some techniques which are unique to bonsai. To control and restrict top growth, nutrient levels, particularly those of nitrogen and phosphorus, require careful regulation. Most potted plants are viewed favorably when they are large and lush, while bonsai plants are admired when small, gnarled, and restrained. High levels of nitrogenous fertilizers, which facilitate rapid growth, are contraindicated, while higher P/N ratios promote woody growth with smaller leaves and better flowering. Because of small container volumes and restricted root systems, bonsai plants require a freely draining soil kept close to field capacity.

To maintain shape, not only must growth be controlled, but the branching patterns must be planned. This is done by selective pruning and by pinching out of terminal buds to release axillary buds from apical dominance. Root prunings, done at intervals as short as 2 or 3 months in some species, promote the development of a tight root ball with large numbers of young lateral roots whose ability to take up water and minerals is great.

Branch and stem position is controlled by selective pruning and by the wiring techniques developed by the Japanese in the eighteenth century. Copper wire is wound around the branch or stem, allowing the part to be moved into the desired position without cracking or breaking. Not only does the wire permit bending, but the stresses imposed by the wire and the new position induces the formation of reaction wood, xylem laid down by the cambium that fixes the branch in the new position, so that when the wire is removed, the branch remains in position. Weighting or tying is used for the same reasons. *See* PLANT GROWTH; PLANT MINERAL NUTRITION.

Richard M. Klein

Bibliography. T. Kawamota, *Saikei: Living Landscapes in Miniature*, Kodansha, Tokyo, 1967; J. Y. Naki, *Bonsai Techniques*, Bonsai Institute of California, Santa Monica, 1975; Y. Yoshimura and G. M. Halford, *The Japanese Art of Miniature Trees and Landscapes*, 1957.

Book manufacture

A series of professional operations including editing, designing, typography, platemaking, printing, and binding to transform a manuscript and its related illustrations into book form.

Books follow a distinctive technique in assembling and binding pages in sequential order and in fitting and sealing the pages into a separate cover. Those designed for relatively permanent usage are sealed into rigid or semirigid covers (called edition, case, or hardback binding; see **illustration**), and those designed for less permanent usage are sealed into paper covers (called soft or paperback binding).

The basic cell of a book's construction is called a signature or section and is composed of multiples of four pages. The folding of signatures printed on a web press is integrated with the printing units. In sheet-fed printing, folding is separate from presswork and is performed on special machines located in a book bindery.

Folding a sheet of paper once produces 4 pages, and each additional fold doubles the number of pages. A sheet folded three times yields a 16-page signature, the first page being the lowest numbered (a page number is called a folio). All right-hand pages are odd numbered, and left-hand pages are even numbered. A book of 160 pages has ten 16-page signatures, the first being folioed 1–16 and the last 145–160.

Signatures, assembled in sequence, are mechanically held together along one folded edge, called the

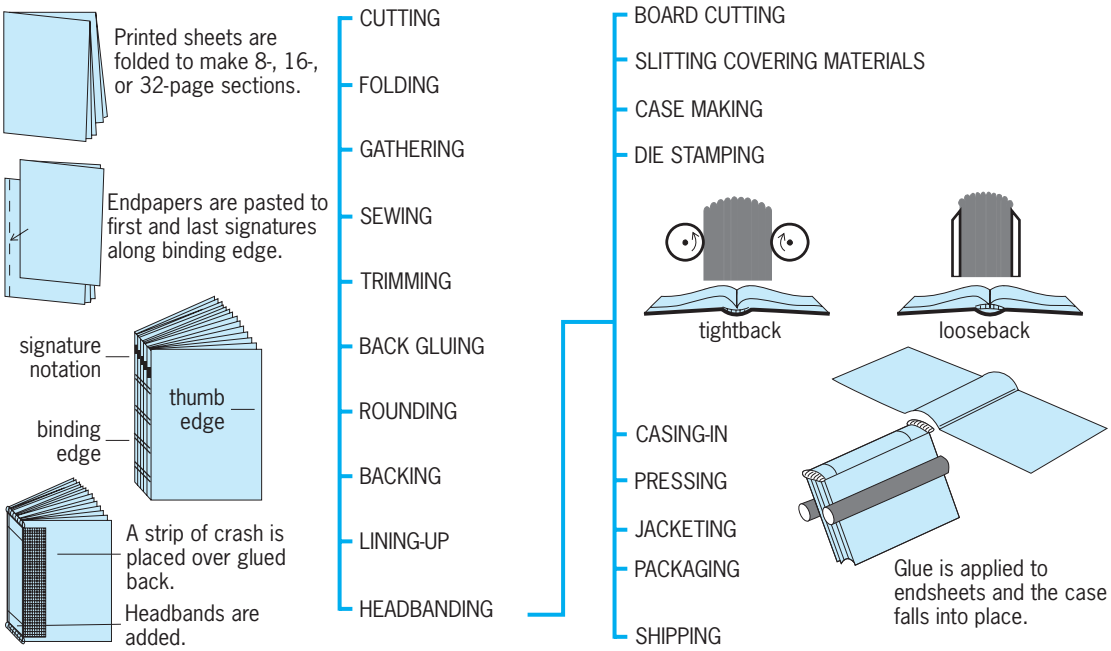

Steps in casebinding a book.

binding edge or spine, by continuous thread sewing through the fold of each signature; or by thread stitching through the side of the book from front to back; or by wire stapling; or by cutting off the binding-edge folds and applying a coating of strong flexible adhesive, sometimes adding an additional locking device of wide mesh textiles or stretch papers.

Purely decorative features include colorful bits of textiles added to the head and tail of the spine (called headbands) and tinted or gilded edges. One spectacular technical breakthrough was the mechanical gilding of book edges; for centuries this had been a closely guarded handicraft process.

The book cover serves to protect the pages against disintegration, to announce the title of the book, and to stimulate the visual interest of the prospective buyer. The degree of usage and permanency influences the selection of raw materials for a book cover (as in reference sets and text and library books), and visual appeal is influenced by the book's price and marketability. Permanent-type covers are made on special equipment in which the cover material and pulp-boards are assembled and joined with special adhesives. In embellishing the cover, the designer has numerous choices of special equipment and processes, such as flat ink, pigment or metal-foil stampings, embossing, silk-screening, and multicolor printing.

Preceding the final assembly of the book into its cover (called "casing in"), strong 4-page front and back end-sheets and one or more kinds of hinges (of textiles or paper) have been applied to strengthen the interlocking between book and cover (called "lining up"). The book also has acquired a concave and convex edge and a pair of ridges or joints along the binding edge (called "backing"). These features are added by special equipment and are specifically engineered to securely lock the book into its cover and to transmit the strains of opening the book throughout the entire body instead of only to the first and last pages.

For shelf display, a book is often wrapped in a colorful, eyecatching jacket (called dust wrapper), or in a transparent synthetic if the cover is highly decorative.

Books reach their markets packaged in many different ways. Single copies are packaged and mailed directly to the consumer; sets are assembled into cartons, one set per carton; textbooks and other single titles are bulk-cartoned or skid-packed and shipped to central warehouses.

Computer-based technologies have changed the way that publishers prepare copy and printers produce books. Book manufacturing technologies that use advanced techniques include digital prepress, electronic page imposition, computer-to-print technology, digital printing, and demand-print and custom publishing services. With demand-print and custom publishing services, product literature and books are produced on demand. Information is stored in a database, and then the publishers or even the reader (if the database is maintained at point of purchase) can select chapters, pages, or articles relevant to a segment of the market and publish them in very small quantities, even a single copy. With digital technology, books are custom-made for a single need. For college courses, for example, instructors may select specific chapters from a publisher-provided digital library. The output format can be

selected as well: from flexible prepunched looseleaf to high-quality patent or casebinding or in electronic form.

Book publishers and printers are on the threshold of a revolution in printing technology. Books printed by the conventional process of offset will continue to serve the needs of consumers and publishers. But decentralized multisite operations housing filmless, plateless digital presses will become more common as book publishers change from print-and-distribute to distribute-and-print, with information transmitted electronically to printing facilities or plants closer to markets. The book manufacturer takes on other functions for the publisher as well. Warehousing, inventory management, sophisticated bulk distribution, and individual fulfillment of orders of books to the consumer or reader are services offered by many book manufacturers. Printers can also be information managers for publishers, digitizing information and turning it into different media, such as electronic books in the form of CD-ROMs, on-line information services, or on-demand print products. *See* PRINTING.

Gregor W. Hendricks; Thomas M. Destree

Bibliography. H. M. Fenton and F. J. Romano, *On-Demand Printing*, 1998; J. Greenfield, *ABC of Bookbinding: A Unique Glossary with Over 700 Illustrations for Collectors and Librarians*, 1998; R. M. Romano and F. J. Romano (eds.), *The GATF Encyclopedia of Graphic Communications*, 1998.

Boolean algebra

A branch of mathematics that was first developed systematically, because of its applications to logic, by the English mathematician George Boole, around 1850. Closely related are its applications to sets and probability.

Boolean algebra also underlies the theory of relations. A modern engineering application is to computer circuit design. *See* DIGITAL COMPUTER; SWITCHING THEORY.

Set-theoretic interpretation. Most basic is the use of boolean algebra to describe combinations of the subsets of a given set I of elements; its basic operations are those of taking the intersection or common part $S \cap T$ of two such subsets S and T, their union or sum $S \cup T$, and the complement S' of any one such subset S. These operations satisfy many laws, including those shown in Eqs. (1), (2), and (3).

$$\begin{gathered} S \cap S = S \qquad S \cap T = T \cap S \\ S \cap (T \cap U) = (S \cap T) \cap U \end{gathered} \tag{1}$$

$$\begin{gathered} S \cup S = S \qquad S \cup T = T \cup S \\ S \cup (T \cup U) = (S \cup T) \cup U \end{gathered} \tag{2}$$

$$\begin{gathered} S \cap (T \cup U) = (S \cap T) \cup (S \cap U) \\ S \cup (T \cap U) = (S \cup T) \cap (S \cup U) \end{gathered} \tag{3}$$

If O denotes the empty set, and I is the set of all elements being considered, then the laws set forth in Eq. (4) are also fundamental. Since these laws are

$$\begin{gathered} O \cap S = O \quad O \cup S = S \quad I \cap S = S \\ I \cup S = I \quad S \cap S' = O \quad S \cup S' = I \end{gathered} \tag{4}$$

fundamental, all other algebraic laws of subset combination can be deduced from them.

In applying boolean algebra to logic, Boole observed that combinations of properties under the common logical connectives *and*, *or*, and *not* also satisfy the laws specified above. These laws also hold for propositions or assertions, when combined by the same logical connectives. *See* LOGIC CIRCUITS.

Boole stressed the analogies between boolean algebra and ordinary algebra. If $S \cap T$ is regarded as playing the role of st in ordinary algebra, $S \cup T$ that of $s + t$, O of 0, I of 1, and S' as corresponding to $1 - s$, the laws listed above illustrate many such analogies. However, as first clearly shown by Marshall Stone, the proper analogy is somewhat different. Specifically, the proper boolean analog of $s + t$ is $(S' \cap T) \cup (S \cap T')$, so that the ordinary analog of $S \cup T$ is $s + t - st$. Using Stone's analogy, boolean algebra refers to boolean rings in which $s^2 = s$, a condition implying $s + s = 0$. *See* RING THEORY; SET THEORY.

Boolean algebra arises in other connections, as in the algebra of (binary) relations. Such relations $\rho, \sigma, \ldots$ refer to appropriate sets of elements $I, J, \ldots$. Any such ρ can be defined by describing the set of pairs (x, t), with x, in I and y, in J, that stand in the given relation—a fact symbolized $x\ \rho\ y$, just as its negation is written $x\ \rho'\ y$. Because of this set theoretic interpretation, boolean algebra obviously applies, with $x(\rho \cap \sigma)y$ meaning $x\ \rho\ y$ and $x\ \sigma\ y$, and $x(\rho \cup \sigma)y$ meaning $x\ \rho\ y$ or $x\ \sigma\ y$.

Abstract relationships. Before 1930, work on boolean algebra dealt mainly with its postulate theory, and with the generalizations obtained by abandoning one or more postulates, such as $(p')' = p$ (brouwerian logic). Since $a \cup b = (a' \cap b')'$, clearly one need consider $a \cap b$ and a' as undefined operations. In 1913 H. M. Sheffer showed one operation only $(a|b = a' \cap b')$ need be taken as undefined. In 1941 M. H. A. Newman developed a remarkable generalization which included boolean algebras and boolean rings. This generalization is based on the laws shown in Eqs. (5) and (6). From these assump-

$$\begin{gathered} a(b + c) = ab + ac \\ (a + b)c = ac + bc \end{gathered} \tag{5}$$

$$\begin{gathered} a1 = 1 \quad a + 0 = 0 + a = a \quad aa' = 0 \\ a + a' = 1 \end{gathered} \tag{6}$$

tions, the idempotent, commutative, and associative laws (1) and (2) can be deduced.

Such studies lead naturally to the concept of an abstract boolean algebra, defined as a collection of symbols combined by operations satisfying the identities listed in formulas (1) to (4). Ordinarily, the phrase boolean algebra refers to such an abstract boolean algebra, and this convention is adopted here.

The class of finite (abstract) boolean algebras is easily described. Each such algebra has, for some nonnegative integer n, exactly 2^n elements and is algebraically equivalent (isomorphic) to the algebra of all subsets of the set of numbers $1, \ldots, n$, under the operations of intersection, union, and complement. Furthermore, if m symbols $a_1, \ldots, a_m$ are combined symbolically through abstract operations $\cap$, $\cup$, and $'$ assumed to satisfy the identities of Eqs. (1) to (4), one gets a finite boolean algebra with 2^{2m} elements—the free boolean algebra with m generators.

Infinite relationships. The theory of infinite boolean algebras is much deeper; it indirectly involves the whole theory of sets. One important result is Stone's representation theorem. Let a field of sets be defined as any family of subsets of a given set I, which contains with any two sets S and T their intersection $S \cap T$, union $S \cup T$, and complements S', T'. Considered abstractly, any such field of sets obviously defines a boolean algebra. Stone's theorem asserts that, conversely, any finite or infinite abstract boolean algebra is isomorphic to a suitable field of sets. His proof is based on the concepts of ideal and prime ideal, concepts which have been intensively studied for their own sake. Because ideal theory in boolean algebra may be subsumed under the ideal theory of rings (via the correspondence between boolean algebras and boolean rings mentioned earlier), it will not be discussed here. A special property of boolean rings (algebras) is the fact that, in this case, any prime ideal is maximal.

The study of infinite boolean algebras leads naturally to the consideration of such infinite distributive laws as those in Eqs. (7*a*) and (7*b*).

$$\begin{aligned} x \cap (\cup_B y_\beta) &= \cup_B (x \cap y_\beta) \\ x \cup (\cap_B y_\beta) &= \cap_B (x \cup y_\beta) \end{aligned} \qquad (7a)$$

$$\begin{aligned} \cap_C \left[\cup_{A_\gamma} u_{\gamma,\alpha}\right] &= \cup_F \left[\cap_C u_{\gamma,\phi(\gamma)}\right] \\ \cup_C \left[\cap_{A_\gamma} u_{\gamma,\alpha}\right] &= \cap_F \left[\cup_C u_{\gamma,\phi(\gamma)}\right] \end{aligned} \qquad (7b)$$

For finite sets B of indices $\beta = 1, \ldots, n$, if $\cap_B y_\beta$ means $\cup y_1 \cdots \cup y_n$, and so on, the laws (7*a*) and (7*b*) follow by induction from (1) to (3). Also, if the symbols x, y_β, and so in (7*a*) and (7*b*) refer to subsets of a given space I, and if $\cup_B y_\beta$ and $\cap_B y_\beta$ refer to the union and intersection of all y_β in B, respectively, then (7*a*) and (7*b*) are statements of general laws of formal logic. However, they fail in most infinite boolean algebras. This is shown by the following result of Alfred Tarski: If a boolean algebra A satisfies the generalized distributive laws (7*a*) and (7*b*), then it is isomorphic with the algebra of all subsets of a suitable space I. A related result is the theorem of L. Loomis (1947) which states: Every σ-complete boolean algebra is isomorphic with a σ-field of sets under countable intersection, countable union, and complement.

In general, such completely distributive boolean algebras of subsets may be characterized by the properties of being complete and atomic. These properties may be defined roughly as the properties that (*a*) there exists a smallest element $\cup_B y_\beta$ containing any given set B of elements y_β and (*b*) any element $y > 0$ contains an atom (or point) $p > 0$, such that $p > x > 0$ has no solution (from Euclid, "A point is that which has no parts"). Condition (*b*) is also implied by the "descending chain condition" of ideal theory.

Other forms. Nonatomic and incomplete boolean algebras arise naturally in set theory. Thus, the algebra of measurable sets in the line or plane, ignoring sets of measure zero, is nonatomic but complete. The field of Borel sets of space is complete as regards countable families B of subsets S_β, but not for uncountable B. Analogous results hold for wide classes of other measure spaces and topological spaces, respectively. In any zero-dimensional compact space, the sets which are both open and closed (which "disconnect" the space) form a boolean algebra; a fundamental result of Stone shows that the most general boolean algebra can be obtained in this way.

Many other interesting facts about boolean algebra are known. For instance, there is an obvious duality between the properties of $\cap$ and $\cup$ in the preceding discussion. However, so many such facts have natural generalizations to the wider context of lattice theory that the modern tendency is to consider boolean algebra as it relates to such generalizations. For instance, the algebras of n-valued logic, intuitionist (brouwerian) logic, and quantum logic are not boolean algebras, but lattices of other types. The same is true of the closed sets in most topological spaces. For further information about boolean algebra *see* LATTICE (MATHEMATICS); TOPOLOGY.

Garrett Birkhoff

Bibliography. G. Birkhoff, *Lattice Theory*, 3d ed., Amer. Math. Soc. Colloq. Publ., vol. 25, 1967, reprint 1993; G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, 1997; F. M. Brown, *Boolean Reasoning*, 1990; J. D. Monk (ed.), *Handbook of Boolean Algebras*, 3 vols., 1989; Research and Education Association Staff, *The Essentials of Boolean Algebra*, rev. ed., 1993.

Boötes

The Herdsman, a constellation perhaps linked in mythology with the invention of the plow, though no specific Greek myth is associated with the name. The herdsman can be thought of holding Canes Venatici, the Hunting Dogs, and using them to drive Ursa Major, the Great Bear. The kite-shaped outline often given to the brightest stars of this northern constellation (see **illustration**) is linked to the kite's string by Arcturus, a red giant and the fourth brightest star in the sky (the second brightest in the northern sky). The arc of the Big Dipper's handle leads to Arcturus: "arc to Arcturus." *See* ARCTURUS; URSA MAJOR.

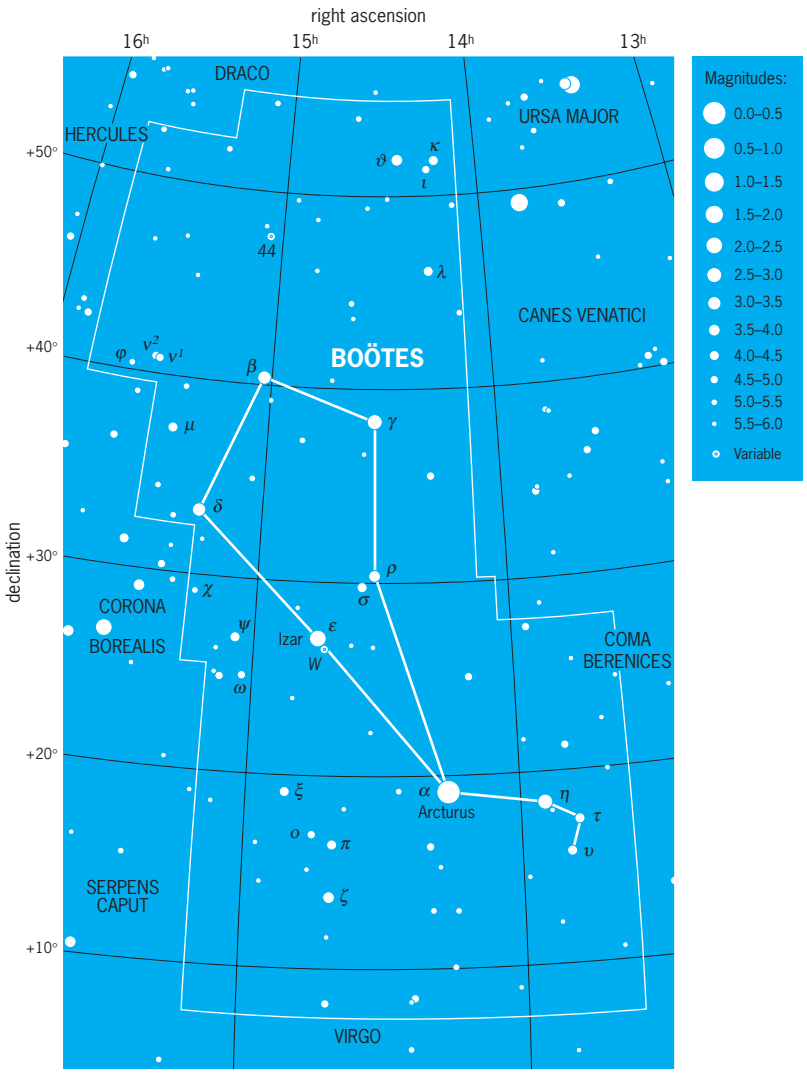

Modern boundaries of the constellation Boötes, the Herdsman. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Boracite

A borate mineral with chemical composition $Mg_3B_7O_{13}Cl$. It occurs in Germany, England, and the United States, usually in bedded sedimentary deposits of anhydrite ($CaSO_4$), gypsum ($CaSO_4 \cdot 2H_2O$), and halite (NaCl), and in potash deposits of oceanic type. The chemical composition of natural boracites varies, with Fe^{2+} or Mn^{2+} replacing part of the Mg^{2+} to yield ferroan boracite or manganoan boracite.

Boracite occurs in crystals which appear to be isometric in external form, despite the fact that the arrangement of the atoms in the crystal structure has only orthorhombic symmetry (see **illus.**). When this natural low-temperature form is heated to 510°F (265°C), a slight readjustment of the atoms takes place internally without change in external form. The atomic symmetry of this resulting high-temperature form, which does not occur naturally, is then consistent with the external habit.

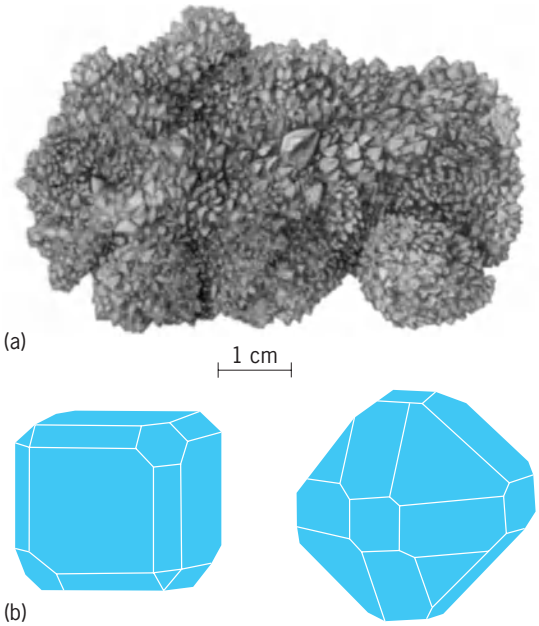

Boracite. (*a*) Specimen from Segeberg, Germany (*specimen from Department of Geology, Bryn Mawr College*). (*b*) Crystal habits (*after C. S. Hurlbut, Jr., Dana's Manual of Mineralogy, 17th ed., Wiley, 1959*).

The hardness is 7–7½ on Mohs scale, and specific gravity is 2.91–2.97 for colorless crystals and 2.97–3.10 for green and ferroan types. Luster is vitreous, inclining toward adamantine. Boracite is colorless to white, inclining to gray, yellow, and green, and rarely pink (manganoan); its streak is white; and it is transparent to translucent. It is strongly piezoelectric and pyroelectric and does not cleave. *See* BORATE MINERALS.

Charles L. Christ

Borane

The general name given to any one of the binary boron hydrides of the general formula B_nH_m. Boron is the only element other than carbon that exhibits an extensive series of molecular hydrides. In contrast to the chains and rings of carbon atoms that typify carbon hydride chemistry, the structures of the boron hydrides are based on polyhedral clusters of boron atoms. These borane clusters are characterized by triangular faces, and such polyhedra are known as deltahedra. The boranes typically have a surface of *exo* B-H bonds directed radically outward from the deltahedral cluster. Deltahedra are closed if they consist entirely of triangular faces, whereas open deltahedra can also typically incorporate four-, five-, and six-membered open faces. Associated with such open faces are additional BHB bridging and/or *endo* B-H hydrogen atoms, which are held in positions

approximately tangential to the cluster surface. These deltahedral cluster structures arise because assemblies of boron and hydrogen atoms have fewer electrons than bonding atomic orbitals available to hold them together. For this reason, they are sometimes described as electron-deficient. Carbon atoms have the same number of valence electrons as atomic orbitals available for bonding and can therefore form the more familiar electron-precise chains and rings of the hydrocarbons, or more electron-rich structures with active lone pairs if atoms of elements to the right of carbon in the periodic table, such as oxygen and nitrogen, are also incorporated. *See* BORON; HYDROCARBON; ORGANIC CHEMISTRY; STRUCTURAL CHEMISTRY.

The intracluster bonding framework of the boranes is typified by multicenter bonding, as there are insufficient framework electrons to provide two-electron two-center connectivity among all the atoms. In diborane, B_2H_6, the simplest borane, the two boron atoms are joined together by two BHB three-center bonds (structure **1**). There are insufficient electrons to provide the seven two-electron bonds for an ethane-like structure (**2**), although this

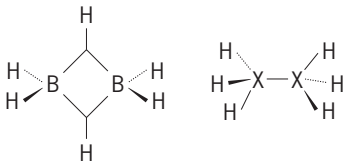

(**1**) (**2**)

is, in fact, the structure of the $[B_2H_6]^{2-}$ anion which, with two electrons more than neutral B_2H_6, is isoelectronic with ethane. Higher boranes incorporate additionally BBB three-center bonds. In the higher borane *nido*-decaborane, $B_{10}H_{14}$, a factorization of the bonding as in (**3**), in which circles represent

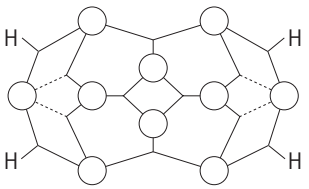

(**3**)

BH(*exo*) units, shows a typical higher-borane network of two-center bonds, three-center bonds, and partial three-center bonds. However, such renditions are simplistic, and represent a net valence-bond approximation. There is considerable multicenter delocalization in these larger clusters, best approached by molecular-orbital treatments. *See* MOLECULAR ORBITAL THEORY.

Classification. The known primary boranes fall naturally into three principal classes, the *closo* boranes, which are generally dianionic, of general formula $[B_nH_n]^{2-}$; the *nido* boranes, of general formula B_nH_{n+4}; and the *arachno* boranes, B_nH_{n+6}. These have progressively more open structures associated with a progressive increase in the number of electron pairs involved in the intracluster bonding. As an extension of the *closo-nido-arachno* sequence, *hypho* and *klado* boranes, B_nH_{n+8} and B_nH_{n+10}, can be postulated. Such structural types are not known for pure boranes, but are recognized in the carboranes, which are somewhat less electron-deficient, and in borane derivatives that incorporate electron-rich ligands. The carboranes (more rigorously, carbaboranes) are borane cluster compounds that also incorporate carbon atoms into the deltahedral clusters; they are closely related to the boranes and have the same general electronic and structural classifications. For a more detailed account of the electronic and geometric classification of boron-containing clusters *see* CARBORANE

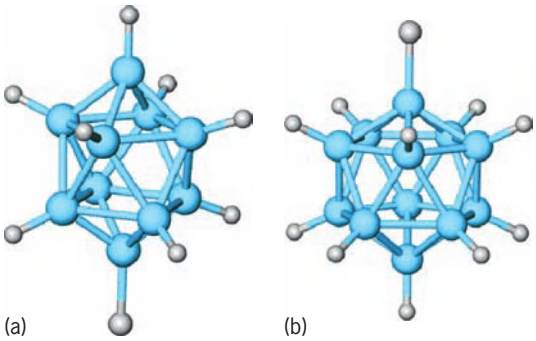

Fig. 1. Representation of the molecular structures of (*a*) the $[closo\text{-}B_{10}H_{10}]^{2-}$ dianion and (*b*) the $[closo\text{-}B_{12}H_{12}]^{2-}$ dianion. The cluster of $[B_{10}H_{10}]^{2-}$ is of regular bicapped-square-antiprismatic D_{4d} symmetry, and the cluster of $[B_{12}H_{12}]^{2-}$ is of regular icosahedral *I* symmetry.

The known single-cluster boranes are B_2H_6 (formally *nido*), *arachno*-B_4H_{10}, *nido*-B_5H_9, *arachno*-B_5H_{11}, *nido*-B_6H_{10}, *arachno*-B_6H_{12}, *nido*-B_8H_{12}, *arachno*-B_8H_{14}, *arachno*-B_9H_{15} (two isomers), and *nido*-$B_{10}H_{14}$, together with all the *closo* species $[B_nH_n]^{2-}$ from $n = 6$ through 12. The $[closo\text{-}B_{12}H_{12}]^{2-}$ ion (**Fig. 1**), with a very symmetrical icosahedral $\{B_{12}\}$ configuration, is particularly stable. This closed 12-vertex configuration is, in fact, very stable, and represents an effective upper barrier to single-cluster size that is not easy to cross. Boron-based single clusters with more than 12 vertices are rare. Often, for a particular value of n, the most stable, best characterized, and best examined boranes are anions, formally obtained by deprotonation of a real or hypothetical parent neutral borane. This obviously applies to the $[closo\text{-}B_nH_n]^{2-}$ series, but other anions of particular stability and/or importance are $[arachno\text{-}B_3H_8]^-$, $[nido\text{-}B_9H_{12}]^-$, $[arachno\text{-}B_9H_{14}]^-$, $[arachno\text{-}B_{10}H_{14}]^{2-}$, and $[nido\text{-}B_{11}H_{14}]^-$. Other known borane anions include $[nido\text{-}B_4H_9]^-$, $[hypho\text{-}B_5H_{12}]^-$, and $[arachno\text{-}B_7H_{12}]^-$, which are mentioned because the neutral undeprotonated parent boranes have not been established. When heteroatoms are incorporated into the borane clusters as constituents to give heteroboranes, or when one-electron and two-electron substituents are attached to the boron cages, then the extent and variety of structure and chemistry, and of potential application, are increased considerably, compared to the known basic binary boron hydride species listed above.

Stability and reactivity. The boranes vary considerably in stability and reactivity. The smaller boranes spontaneously decompose to give higher boranes and hydrogen. Ultimately, an intractable hydrogen-poor boron-hydride polymeric species is formed. The smaller boranes will generally inflame spontaneously in air, whereas the larger boranes tend to be more stable. The larger borane *nido*-$B_{10}H_{14}$ is as stable in air as NaOH, and salts of the [*closo*-$B_{12}H_{12}]^{2-}$ dianion are indefinitely stable in air. However, these are gross generalizations, and there are many exceptions.

Syntheses. The syntheses of the boranes generally derive from the pioneering work of A. Stock and coworkers in the early 1900s. The classical route involves the thermolysis of B_2H_6, the simplest borane, which may be derived by treatment of metal borides or borohydrides with acid [reaction (1)], or by treatment of boron halides with hydridic reagents such as NaH or $Li(AlH_4)$ [reaction (2)]. Thermolysis of B_2H_6

$$2M_3B_n + 6nHX \rightarrow nB_2H_6 + 6MX_n \qquad (1)$$

$$2BCl_3 + 6NaH \rightarrow B_2H_6 + 6Na^+ + 6Cl^- \qquad (2)$$

leads, via dihydrogen loss and a series of unstable intermediates, to the initial isolatable product *nido*-pentaborane, B_5H_9 (**Fig. 2*b***), a colorless liquid with a melting point of −47°C (−52.6°F) and a boiling point of 60°C (140°F). Subsequently, the much more stable *nido*-decaborane, $B_{10}H_{14}$ (Fig. 2*c*), is formed, as a colorless, volatile crystalline solid with a melting point of 99.5°C (211°F) and a boiling point of 213°C (415°F). Multiton quantities of these two substances were produced by the United States and the former Soviet Union during the Cold War, as their very high heats of combustion suggested their possible use as high-energy jet and rocket fuels. However, their combustion characteristics were far from ideal, and the problem of storing and handling the spontaneously flammable liquid B_5H_9 precluded their use as fuels. Nevertheless, the bulk of borane and carborane chemistry known to date originated largely from B_5H_9 and $B_{10}H_{14}$ as starting substrates.

These bulk sources of B_5H_9 and $B_{10}H_{14}$ no longer exist, but alternative syntheses of the important species $B_{10}H_{14}$, which is the principal entry into monocarbaborane chemistry, dicarbaborane chemistry, and most other borane and heteroborane chemistry in the 8-vertex and greater cluster-size region, are currently being developed. B_5H_9 may be prepared in the laboratory from the $[B_3H_8]^-$ anion by reaction with BF_3 [reaction (3)], the $[B_3H_8]^-$ anion being obtained from the $[BH_4]^-$ anion by reaction with elemental iodine [reaction (4)].

$$3[B_3H_8]^- + 4BF_3 \rightarrow 2B_5H_9 + 3[BF_4]^- + 3H_2 \qquad (3)$$

$$3[BH_4]^- + I_2 \rightarrow [B_3H_8]^- + 2I^- + 2H_2 \qquad (4)$$

See METAL HYDRIDES.

The volatility and reactivity of B_5H_9 dictates that the study of most of its chemistry has required the use of vacuum-line techniques and a rigorously oxygen-free atmosphere. The much greater stability of $B_{10}H_{14}$ along with its consequent easier handling has generally permitted the examination of much of its chemistry under experimentally much easier conditions.

Reaction chemistry. Reactions of B_5H_9 and $B_{10}H_{14}$ are typical of general reaction behavior. In general, electrophilic substitution occurs at BH(*exo*) units off the open face (so-called apical positions), and nucleophilic substitution occurs at BH(*exo*) units around the open face. An extensive substituent chemistry is thus available, but is as yet insufficiently well refined to permit designed site-specific multiple substitution in the general case. BHB bridging hydrogen atoms are protolytically acidic; for example, $B_{10}H_{14}$ has a p*K*a value in the region of 2.70, similar to that of ClH_2CCOOH at 2.85, giving the [*nido*-$B_{10}H_{13}]^-$ anion via proton loss from a bridging site. Most *nido* and *arachno* boranes that have BHB bridging hydrogen atoms similarly yield conjugate base anions on treatment with (non-nucleophilic) bases. Although the resulting species are still electron-deficient cluster species, the anionic sites generated can act as Lewis bases, and as such can coordinate to Lewis acid sites such as metal centers, resulting in the formation of metallaboranes. Although such species are ostensibly metal complexes with borane ligands, the metal center is generally bound very tightly. Consequently, such species are often better described as metallaborane cluster compounds in which the borane cluster has incorporated a metal center into the cluster structure.

Cluster closure along the *arachno-nido-closo* sequence can occur by the removal of electrons from the cluster, as exemplified by the treatment of *nido*-$B_{10}H_{14}$ with amines, such as NEt_3, followed by thermolysis in which the $[NHEt_3]^+$ salt of the [*closo*-$B_{10}H_{10}]^{2-}$ anion is formed, along with dihydrogen [reaction (5)]. Conversely, electron addition

$$B_{10}H_{14} + 2NEt_3 \rightarrow 2[NHEt_3]^+ + [B_{10}H_{10}]^{2-} + H_2 \qquad (5)$$

results in cluster opening along the *closo-nido-arachno* sequence, as in the treatment of *nido*-$B_{10}H_{14}$ with elemental sodium in tetrahydrofuran to give the [*arachno*-$B_{10}H_{14}]^{2-}$ dianion by simple electron transfer. Closely related to the latter is the effective addition of electrons to the cluster by the attachment of two-electron ligands, as in the reactions of organyl phosphines, PR_3, with *nido*-B_8H_{12}

Fig. 2. Representations of the molecular structures of (*a*) *arachno*-B_4H_{10}; (*b*) *nido*-B_5H_9, and (*c*) *nido*-$B_{10}H_{14}$. B_4H_{10} and $B_{10}H_{14}$ have C_{2v} symmetry and B_5H_9 has C_{4v} symmetry.

to give *arachno*-$B_8H_{12}(PR_3)$ and with *nido*-B_5H_9 to give *hypho*-$B_5H_9(PR_3)_2$. An important variation of this is the reaction of *nido*-$B_{10}H_{14}$ with two two-electron ligands L to give *arachno* $B_{10}H_{12}L_2$ species with dihydrogen loss, for example, when L = SMe_2 [reaction (6)]. The *arachno* $B_{10}H_{12}L_2$ species are important intermediates in dicarborane formation from *nido*-$B_{10}H_{14}$.

$$B_{10}H_{14} + 2SMe_2 \rightarrow B_{10}H_{12}(SMe_2)_2 + H_2 \qquad (6)$$

Cluster dismantling (that is, descent through a series to give lower boranes) is usually effected by stronger nucleophilic bases; for example, 10-boron *nido*-$B_{10}H_{14}$ converts to the 9-boron [*arachno*-$B_9H_{14}]^-$ anion by use of alcoholic aqueous KOH [reaction (7)]. The converse process of cluster *Aufbau* (that is, addition of vertices) is an important reaction type for the ascent of the series of boranes; for example, 10-boron *nido*-$B_{10}H_{14}$ with simple borohydride species, such as the $[BH_4]^-$ anion or neutral $BH_3(NR_3)$, will generate the 11-boron [*nido*-$B_{11}H_{14}]^-$ monoanion [reaction (8)] and then, under increasingly forcing conditions, the 12-boron [*closo*-$B_{12}H_{12}]^{2-}$ dianion [reaction (9)].

$$B_{10}H_{14} + 2[OH]^- + 2H_2O \rightarrow [B_9H_{14}]^- + [B(OH)_4]^- + H_2 \qquad (7)$$

$$B_{10}H_{14} + [BH_4]^- \rightarrow [B_{11}H_{14}]^- + 2H_2 \qquad (8)$$

$$B_{10}H_{14} + 2BH_3(NR_3) \rightarrow [B_{12}H_{12}]^{2-} + 2[NHEt_3]^+ + 3H_2 \qquad (9)$$

Atoms of elements other than boron can be incorporated in *Aufbau* processes, such as the incorporation of metal centers to generate metallaboranes and of sulfur to generate thiaboranes, with both falling into a general classification of heteroboranes. Many reactions are accompanied by cluster rearrangement processes. Although cluster rearrangements are difficult to recognize in pure boron-hydride chemistry, requiring site labeling for their elucidation, cluster rearrangements are established as quite common in heteroborane chemistry in which sites are effectively labeled by their occupation by the heteroatoms.

Many reaction systems incorporate two or more of the several processes defined above, such as in the reaction of *nido*-$B_{10}H_{14}$ with the polysulfide anion to give the *arachno* 10-vertex $[SB_9H_{12}]^-$ anion, an important entry into thiaborane cluster chemistry [reaction (10)].

$$B_{10}H_{14} + [S_x]^{2-} + 4H_2O \rightarrow [SB_9H_{12}]^- + \{S_{(x-1)}\} + [B(OH)_4]^- + 3H_2 \qquad (10)$$

Cluster fusion and macropolyhedral boranes. Although the cluster chemistry of boron hydrides is potentially very extensive, its widest development is inhibited by the experimental difficulty of progressing beyond the single-cluster stability barrier of the 12-vertex icosahedron that is typified by the [*closo*-$B_{12}H_{12}]^{2-}$ anion. Boron-based single clusters with more than 12 vertices are rare. Although clusters can be joined by simple two-electron, two-center B—B sigma bonds—for example, in (*nido*-$B_{10}H_{13})_2$, of which several isomers can be isolated in low yield from a thermolysis of *nido*-$B_{10}H_{14}$—there is difficulty in generating bigger boranes in which smaller clusters are fused together more intimately with boron atoms held in common between the two clusters, so that the multicenter bonding is contiguous across the conjunction. The development of this area is inhibited by the lack of transferable generic routes for intercluster fusion. However, the structures and chemistries of the known macropolyhedral higher boranes—$B_{12}H_{16}$, $B_{13}H_{19}$, $B_{14}H_{18}$, $B_{14}H_{20}$, $B_{15}H_{23}$, $B_{16}H_{20}$, $B_{18}H_{22}$ (two isomers), the $[B_{19}H_{22}]^-$ anion, and $B_{20}H_{16}$—indicate that future chemistry will be extensive, and will be rendered even more so by combining the structural features of the boranes with the extensive known types of structural features of electron-rich molecular architecture such as that of carbon-based chemistry. Of the macropolyhedral boranes, the $B_{18}H_{22}$ isomers, *n*-$B_{18}H_{22}$ and *iso*-$B_{18}H_{22}$, also known as *anti*-$B_{18}H_{22}$ and *syn*-$B_{18}H_{22}$ respectively, are the best examined (**Fig. 3**). *Anti*-$B_{18}H_{22}$ is a colorless, stable crystalline compound that may be obtained by the thermolysis of *arachno* 9-vertex $B_9H_{13}(OEt_2)$ [reaction (11)].

$$2B_9H_{13}(OEt_2) \rightarrow B_{18}H_{22} + 2OEt_2 + 2H_2 \qquad (11)$$

Structurally (for example, lower schematic cluster structure diagrams in Fig. 3), the $B_{18}H_{22}$ isomers can be viewed as consisting of two *nido*-decaboranyl clusters fused together with two boron atoms held in common (for example, lower schematic cluster

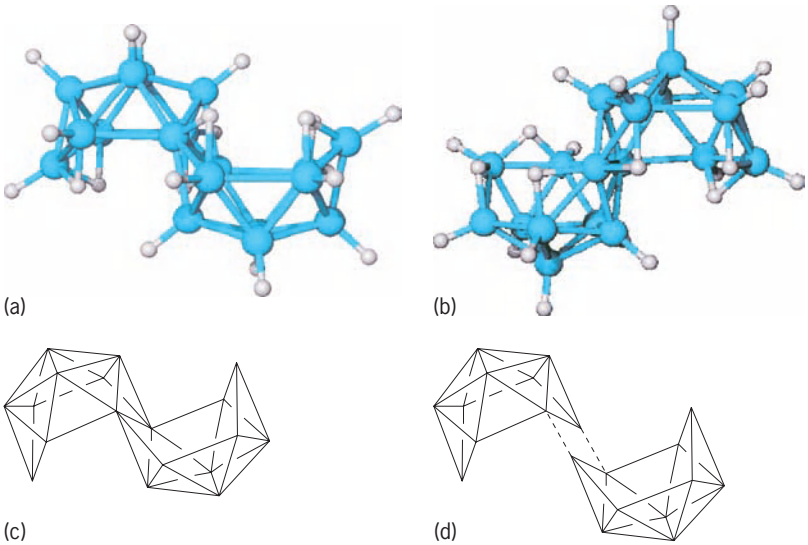

Fig. 3. Representations of the molecular structures of (*a*) *anti*-$B_{18}H_{22}$ and (*b*) *syn*-$B_{18}H_{22}$. Note the resemblance of each subcluster in each molecule to that of *nido*-$B_{10}H_{14}$ (Fig. 2*c*). Diagrams *c* and *d* illustrate for *anti*-$B_{18}H_{22}$ the manner of fusion, with two boron atoms held in common, of two *nido*-decaboranyl sub-cluster units: (*c*) a schematic representation of the boron cluster framework of *anti*-$B_{18}H_{22}$ and (*d*) a factorization of the structure into two *nido*-decaboranyl units.

structure diagrams in Fig. 3). This view is also reflected in some of their chemistry, with several, but by no means all, of their known reactions resembling those of *nido*-$B_{10}H_{14}$.

Heteroborane chemistry. The *closo/nido/arachno/* etc. geometrical and electronic principles apply to clusters that also contain atoms of elements other than boron as cluster constituents. R. Hoffman's isolobal principle for comparing molecular fragments is useful in rationalizing such heteroboranes. For example, the frontier orbitals and electron occupancy of a neutral {CH} fragment—three orbitals and two electrons—mimic those for an anionic $\{BH\}^-$ fragment, and so these fragments can notionally replace each other in molecular structures; that is, they are isolobal. Alternatively, cationic $\{CH^+\}$ can be regarded as isolobal with neutral {BH}. Application of this principle generates, for example, the [*closo*-$CB_{11}H_{12}]^-$ monanion and neutral *closo*-$C_2B_{10}H_{12}$ from the [*closo*-$B_{12}H_{12}]^{2-}$ dianion. Both these carboranes are in fact well-recognized species. The same principle applies to other main-group element fragments, such as {NH} and {S}, that can be regarded as isolobal with $\{BH\}^{2-}$. Thus, neutral 12-vertex $HNB_{11}H_{11}$ and $SB_{11}H_{11}$ are also known analogues of the stable [*closo*-$B_{12}H_{12}]^{2-}$ anion. The incorporation of such heteroatoms introduces more flexibility and variety than is available to pure boron-hydride chemistry. For example, the sulfur-containing thiaboranes have enabled the extension of fused-cluster macropolyhedral architectures significantly beyond those exhibited by the boranes themselves, with the elucidation of several neutral species, [*n*-$S_2B_{16}H_{16}$], [*iso*-$S_2B_{16}H_{16}$], [*iso*-$S_2B_{16}H_{14}(PPh_3)$], [$S_2B_{17}H_{17}$], [$S_2B_{17}H_{17}(SMe_2)$], and [$S_2B_{18}H_{20}$], and the monoanions [$S_2B_{16}H_{15}]^-$, [$S_2B_{16}H_{17}]^-$, [*iso*-$S_2B_{16}H_{17}]^-$, [$SB_{17}H_{19}]^-$, [$S_2B_{17}H_{16}]^-$, [$S_2B_{17}H_{18}]^-$, [$S_2B_{18}H_{19}]^-$, and [*iso*-$S_2B_{18}H_{19}]^-$.

The principle of isolobal replacement also applies to the rationalization of compounds resulting from the incorporation of metal centers into the boron hydride clusters to form metallaboranes. Thus, the neutral formally octahedral iron(II) fragment $\{Fe(CO)_3\}$ has three orbitals and two electrons available for bonding and thereby mimics a neutral {BH} fragment. The yellow low-melting [$(CO)_3FeB_4H_8$], isolated in low yield from the co-thermolysis of *nido*-B_5H_9 with [$Fe(CO)_5$], has a *nido* five-vertex $\{FeB_4\}$ cluster core that overall mimics *nido*-B_5H_9. The neutral $\{Ru(\eta^6\text{-Ar})\}$ fragment, where Ar is an aromatic ligand such as C_6H_6 or C_6Me_6, is another {BH} mimic. In accord with this, a series of isomers of [$(\eta^6\text{-}C_6Me_6)RuB_9H_{13}$], all with 10-vertex $\{RuB_9\}$ cluster structures analogous to the $\{B_{10}\}$ unit of *nido*-$B_{10}H_{14}$, are known. The area of metallaborane chemistry is extensive. The principle of isolobal substitution is general, so that the structures of many cluster compounds that do not contain boron, particularly the carbonyls of the Fe/Ru/Os group, can be rationalized on the basis of the borane structures and the isolobal concept. *See* METAL CLUSTER COMPOUND.

John D. Kennedy

Bibliography. L. Barton, Systemisation and structures of the boron hydrides, pp. 169–206 in *New Trends in Chemistry (Topics in Current Chemistry 100)*, ed. by F. L. Boschke, Springer-Verlag, 1982; J. Bould et al., *Boron Chemistry at the Millennium* (special edition), *Inorg. Chim. Acta*, 289:95–124, 1999; M. J. Carr et al., Macropolyhedral boron-containing cluster chemistry, a metallathiaborane from $S_2B_{17}H_{17}$: Isolation and characterisation of [$(PMe_2Ph)_2PtS_2B_{16}H_{16}$], a *neo-arachno* ten-vertex cluster shape, and the constitution of the [*arachno*-$B_{10}H_{15}]^-$ anion, *Collect. Czechoslovak Chem. Commun.* 70:430–440, 2005; J. Casanova (ed.), *The Borane-Carborane-Carbocation Continuum*, Wiley, 1998; M. G. Davison et al. (eds.), *Contemporary Boron Chemistry*, Royal Society of Chemistry, Cambridge, 2000; N. N. Greenwood and A. Earnshaw, Boron, pp. 139–215 in *The Chemistry of the Elements*, 2d ed., Butterworth-Heineman, Oxford, 1997/1998; M. F. Hawthorne, The role of chemistry in the development of boron neutron capture therapy of cancer, *Angewandte Chem. Int. Ed.*, 32:950–984, 1993; J. D. Kennedy, The polyhedral metallaboranes, *Prog. Inorg. Chem.* 32:519–679, 1984, and 34:211–434, 1986; E. L. Muetterties (ed.), *Boron Hydride Chemistry*, Academic Press, 1975; J. Plešek, Potential applications of the boron cluster compounds, *Chem. Rev.*, 92:269–278, 1992; S. L. Shea et al., Polyhedral boron-containing cluster chemistry: Aspects of architecture beyond the icosahedron, *Pure Appl. Chem.* 75:1239–1248, 2003; A. Stock, *Hydrides of Boron and Silicon*, Cornell University Press, 1933; K. Wade, Structural and bonding patterns in cluster chemistry, *Adv. Inorg. Chem. Radiochem.*, 18:1–66, 1976; A. J. Welch et al., Beyond the icosahedron: The first 13-vertex carborane, *Angewandte Chem. Int. Ed.*, 42:225–228, 2003.

Borate minerals

A large group of minerals in which boron is chemically bonded to oxygen. Boron is a fairly rare element. However, because of its chemical character, it is very susceptible to fractionation in Earth processes and can become concentrated to a degree not found in other elements of similar abundance. Boron is symbolized B, has the atomic number 5, and has the ground-state electronic structure $[He]2s^22p^1$. The very high ionization potentials for boron mean that the total energy required to produce the B^{3+} ion is greater than the compensating structure energy of the resulting ionic solid, and hence bond formation involves covalent (rather than ionic) mechanisms. However, boron has only three electrons to contribute to covalent bonding involving four orbitals, (s, p_x, p_y, p_z). This results in boron being a strong electron-pair acceptor (that is, a strong Lewis acid) with a very high affinity for oxygen. The structural chemistry of boron and silicon (Si), when associated with oxygen (O), is quite similar. The BO_3, BO_4, and SiO_4 groups have a marked tendency to polymerize in

the solid state, and this aspect of their behavior gives rise to the structural complexity of both groups. However, subtle differences in chemical bonding do give rise to differences in the character of this polymerization, particularly when H_2O is also involved. These differences result in the very different properties of the resultant minerals and their very different behavior in Earth processes. *See* BORON.

Structural classification. The boron cation is small and can occur in both triangular and tetrahedral coordination by oxygen: $(BO_3)^{3-}$ and $(BO_4)^{5-}$. Hydrogen (H) atoms readily bond to the oxygen atoms of the borate groups to form $[B(OH)_3]^0$ and $[B(OH)_4]^-$ groups or intermediate groups such as $[BO_2(OH)]^{2-}$ and $[BO_2(OH)_2]^{3-}$; all types of borate groups can be generally written as $(B\phi_3)$ or $(B\phi_4)$, where $\phi = O^{2-}$, OH^-. The borate groups readily polymerize to give great structural diversity in borate minerals. A specific structure may be characterized by the numbers of $(B\phi_3)$ triangles and $(B\phi_4)$ tetrahedra in the structure. These characteristics may be denoted by two symbols of the form *A:B*. *A* is the number of $(B\phi_3)$ and $(B\phi_4)$ groups, denoted as $A = x\triangle y\square$, where x is the number of $(B\phi_3)$ [$= \triangle$] groups and y is the number of $(B\phi_4)$ [$= \square$] groups. *B* is a character string that contains information on the connectivity of the polyhedra. The string is written such that adjacent $\triangle$ or $\square$ (or both) represent polyhedra that share corners, and the delimiters <> indicate that the included polyhedra share corners to form a ring. The sharing of polyhedra between rings is indicated by the symbols −, =, ≡, and so on, for one, two, three, or more polyhedra, respectively. For example, the *FBB* with the descriptor $2\triangle2\square{:}<\triangle2\square>=<\triangle2\square>$ contains two triangles and two tetrahedra. There are two three-membered rings of polyhedra, each of which contains one triangle and two tetrahedra, and the rings have two tetrahedra in common. The resulting symbol denotes the borate components and their connectivity in the crystal structure at the local level; such local clusters of polyhedra are called fundamental building blocks. These fundamental building blocks (*FBB*) link together to form the structural unit of the crystal, a strongly bonded (usually anionic) array of atoms; the structural units are then linked together by weakly bonding interstitial atoms (commonly sodium, potassium, calcium, and hydrogen in borate minerals) to form the complete crystal structure of the mineral. The structural classification of borate minerals is based on the character of the polymerization of the fundamental building blocks to form the structural unit. Polymerization may occur in 0, 1, 2, or 3 dimensions, forming structural units that are isolated polyhedra or finite clusters of polyhedra (0-dimensional polymerization), infinite chains (1-dimensional polymerization), infinite sheets (2-dimensional polymerization), and infinite frame-works (3-dimensional polymerization). In the **table** are shown the borate minerals of known structure, classified according to these principles. *See* POLYHEDRON; POLYMERIZATION.

Mixed-anion borate minerals. The $(B\phi_3)$ and $(B\phi_4)$ borate groups may combine with other geologically common oxyanions [for example, $(SO_4)^{2-}$, sulfate; $(PO_4)^{3-}$, phosphate; $(AsO_4)^{3-}$, arsenate; $(CO_3)^{2-}$, carbonate; $(SiO_4)^{4-}$, silicate; $(BeO_4)^{6-}$, beryllate] to produce mixed-anion minerals. Of these, the silicates are the most diverse (see table), and include tourmaline, a group of very common accessory minerals in a wide variety of geological environments.

Occurrence. Boron has an estimated primitive-mantle abundance of 0.6 part per million and a crustal abundance of 15 ppm. Despite this low abundance, fractionation in crustal processes results in concentration of boron to the extent that it forms an extensive array of minerals in which it is an essential constituent, and very complex deposits of borate minerals. Major concentrations of borate minerals occur in continental evaporite deposits (common in the desert regions of California and Nevada). Borate minerals are often very soluble in aqueous environments. In areas of internal drainage, saline lakes are formed, and continued evaporation leads to accumulation of large deposits of borate minerals. Borates may also occur in marine evaporites. Isolated-cluster borates (see table) are characteristic of metamorphosed boron-rich sediments and skarns. Most borosilicate minerals are characteristic of granitic-pegmatite environments, either as a pegmatite phase or as a constituent of their exocontact zone. In particular, tourmaline is the most widespread of the mixed-anion borate minerals, occurring in a wide variety of igneous, hydrothermal, and metamorphic rocks. *See* ORE AND MINERAL DEPOSITS; PEGMATITE; SALINE EVAPORITES.

Industrial uses. Despite its low crustal abundance, fractionation of boron in crustal processes leads to formation of deposits of borate minerals from which boron and (borates) can be easily extracted in large quantities. The easy availability and unique chemical properties result in boron being a major industrial material. It is widely used in soaps and washing powders. Boron combines well with silicon and other elements to form a wide variety of special-property glasses and ceramics; it also alloys with a variety of metals, producing lightweight alloys for specialty uses. Boron compounds usually have very low density; hence borates in particular are used as lightweight fillers in medicines, and also are used as insulation. The fibrous nature of some borate minerals results in their use in textile-grade fibers and lightweight fiber-strengthened materials. The mineral borax is used as a water softener and as a cleaning flux in welding and soldering. Boric acid has long been used as an antiseptic and a drying agent. Boron is also important as a constituent of inflammatory materials in fireworks and rocket fuel. Some mixed-anion borate minerals are used as gemstones. Tourmaline is of particular importance in this respect, forming pink (rubellite), blue-green (Paraiba tourmaline), green (chrome tourmaline), and pink + green (watermelon tourmaline) from a wide variety of localities. Kornerupine and sinhalite are also used

Borate minerals

Borate groups (*A*)	Name	Connectivity (*B*)	Formula
		Isolated borate-group minerals	
1△	Fluoborite	△	$Mg_3(BO_3)(F,OH)_3$
1△	Painite	△	$CaZr(BO_3)Al_9O_{15}$
1△	Jeremejevite	△	$Al_6(BO_3)_5F_3$
1△	Warwickite	△	$(Mg,Ti,Fe^{3+},Al)_2O(BO_3)$
1△	Azoproite	△	$(Mg,Fe^{2+})_2(Fe^{3+},Ti,Mg)O_2(BO_3)$
1△	Bonaccordite	△	$Ni_2Fe^{3+}O_2(BO_3)$
1△	Fredrikssonite	△	$Mg_2(Mn^{3+},Fe^{3+})O_2(BO_3)$
1△	Ludwigite	△	$Mg_2FeO_2(BO_3)$
1△	Vonsenite	△	$Fe_2Fe^{3+}O_2(BO_3)$
1△	Pinakiolite	△	$(Mg,Mn^{2+})_2(Mn^{3+},Sb^{3+})O_2(BO_3)$
1△	Hulsite	△	$(Fe^{2+},Mg)_2(Fe^{3+},Sn)O_2(BO_3)$
1△	Magnesiohulsite	△	$(Mg,Fe)_2(Fe^{3+},Sn)O_2(BO_3)$
1△	Chestermanite	△	$Mg_2(Fe^{3+},Mg,Al,Sb^{5+})O_2(BO_3)$
1△	Orthopinakiolite	△	$(Mg,Mn^{2+})_2Mn^{3+}O_2(BO_3)$
1△	Blatterite	△	$(Mn^{2+},Mg)_2(Mn^{3+},Sb^{3+},Fe^{3+})O_2(BO_3)$
1△	Takéuchiite	△	$(Mg,Mn^{2+})_2(Mn^{3+},Fe^{3+})O_2(BO_3)$
1△	Karlite	△	$Mg_7(OH)_4(BO_3)_4Cl$
1△	Wightmanite	△	$Mg_5O(OH)_5(BO_3)(H_2O)_2$
1△	Jimboite	△	$Mn_3(BO_3)_2$
1△	Kotoite	△	$Mg_3(BO_3)_2$
1△	Nordenskioldine	△	$CaSn(BO_3)_2$
1△	Tusionite	△	$MnSn(BO_3)_2$
1△	Sassolite	△	$[B(OH)_3]$
1□	Sinhalite	□	$AlMg(BO_4)$
1□	Bandylite	□	$Cu[B(OH)_4]Cl$
1□	Teepleite	□	$Na_2[B(OH)_4]Cl$
1□	Frolovite	□	$Ca[B(OH)_4]_2$
1□	Hexahydroborite	□	$Ca[B(OH)_4]_2 \cdot 2H_2O$
1□	Henmilite	□	$Ca_2Cu(OH)_4[B(OH)_4]_2$
		Finite-cluster borate minerals	
2△	Suanite	2△	$Mg_2[B_2O_5]$
2△	Szaibelyite	2△	$Mg_2(OH)[B_2O_4(OH)]$
2△	Sussexite	2△	$Mn_2(OH)[B_2O_4(OH)]$
2△	Kurchatovite	2△	$CaMg[B_2O_5]$
2□	Pentahydroborite	2□	$Ca[B_2O(OH)_6](H_2O)_2$
2□	Pinnoite	2□	$Mg[B_2O(OH)_6]$
2△1□	Ameghinite	⟨2△□⟩	$Na[B_3O_3(OH)_4]$
1△2□	Inderite	⟨△2□⟩	$Mg[B_3O_3(OH)_5](H_2O)_5$
1△2□	Kurnakovite	⟨△2□⟩	$Mg[B_3O_3(OH)_5](H_2O)_5$
1△2□	Inyoite	⟨△2□⟩	$Ca[B_3O_3(OH)_5](H_2O)_4$
1△2□	Meyerhofferite	⟨△2□⟩	$Ca[B_3O_3(OH)_5](H_2O)$
1△2□	Solongoite	⟨△2□⟩	$Ca_2[B_3O_4(OH)_4]Cl$
1△2□	Inderborite	⟨△2□⟩	$CaMg[B_3O_3(OH)_5]_2(H_2O)_6$
3□	Nifontovite	⟨3□⟩	$Ca_3[B_3O_3(OH)_6](H_2O)_2$
2△2□	Hydrochlorborite	⟨△2□⟩△	$Ca_2[B_3O_3(OH)_4][BO(OH)_3]Cl(H_2O)_7$
4□	Uralborite	⟨3□⟩□	$Ca_2[B_4O_4(OH)_8]$
4△1□	Sborgite	⟨2△□⟩–⟨2△□⟩	$Na[B_5O_6(OH)_4](H_2O)_3$
2△3□	Ulexite	⟨△2□⟩–⟨△2□⟩	$NaCa[B_5O_6(OH)_6](H_2O)_5$
2△2□	Borax	⟨△2□⟩=⟨△2□⟩	$Na_2[B_4O_5(OH)_4](H_2O)_8$
2△2□	Tincalconite	⟨△2□⟩=⟨△2□⟩	$Na_2[B_4O_5(OH)_4](H_2O)_3$
2△2□	Hungchaoite	⟨△2□⟩=⟨△2□⟩	$Mg[B_4O_5(OH)_4](H_2O)_7$
2△2□	Fedorovskite	⟨△2□⟩=⟨△2□⟩	$Ca_2Mg_2(OH)_4[B_4O_7(OH)_2]$
2△2□	Roweite	⟨△2□⟩=⟨△2□⟩	$Ca_2Mn_2(OH)_4[B_4O_7(OH)_2]$
3△3□	Mcallisterite	[ϕ]⟨△2□⟩|⟨△2□⟩|⟨△2□⟩|	$Mg_2[B_6O_7(OH)_6]_2(H_2O)_9$
3△3□	Aksaite	[ϕ]⟨△2□⟩|⟨△2□⟩|⟨△2□⟩|	$Mg[B_6O_7(OH)_6](H_2O)_2$
3△3□	Rivadavite	[ϕ]⟨△2□⟩|⟨△2□⟩|⟨△2□⟩|	$Na_6Mg[B_6O_7(OH)_6]_4(H_2O)_{10}$
12△3□	Ammonioborite	3(⟨2△□⟩–⟨2△□⟩)	$(NH_4)_3[B_{15}O_{20}(OH)_8](H_2O)_4$
		Infinite-chain borate minerals	
1□	Vimsite	□	$Ca[B_2O_2(OH)_4]$
1△2□	Colemanite	⟨△2□⟩	$Ca[B_3O_4(OH)_3](H_2O)$
1△2□	Calciborite	⟨△2□⟩	$Ca[B_2O_4]$
1△2□	Hydroboracite	⟨△2□⟩	$CaMg[B_3O_4(OH)_3]_2(H_2O)_3$
4△1□	Larderellite	⟨2△□⟩–⟨2△□⟩	$(NH_4)[B_5O_7(OH)_2](H_2O)$
2△3□	Probertite	⟨△2□⟩–⟨△2□⟩	$NaCa[B_5O_7(OH)_4](H_2O)_3$
3△2□	Ezcurrite	⟨△2□⟩–⟨2△□⟩	$Na_2[B_5O_7(OH)_3](H_2O)_2$
3△3□	Kaliborite	⟨△2□⟩–⟨△2□⟩△	$KMg_2H[B_7O_8(OH)_5](H_2O)_4$
3△4□	Kernite	⟨△2□⟩–⟨△2□⟩–⟨△2□⟩	$Na_2[B_4O_6(OH)_2](H_2O)_3$
3△3□	Aristarainite	[ϕ]⟨△2□⟩|⟨△2□⟩|⟨△2□⟩|	$NaMg[B_6O_8(OH)_4]_2(H_2O)_4$

(cont.)

Borate minerals (*cont.*)

Borate groups (*A*)	Name	Connectivity (*B*)	Formula
		Infinite-sheet borate minerals	
3△2□	Biringuccite	⟨2△□⟩–⟨△2□⟩	$Na_2[B_5O_8(OH)](H_2O)$
3△2□	Nasinite	⟨2△□⟩–⟨△2□⟩	$Na_2[B_5O_8(OH)](H_2O)_2$
3△2□,1△	Gowerite	⟨△2□⟩–⟨2△□⟩,△	$Ca[B_5O_8(OH)][B(OH)_3](H_2O)_3$
3△2□,1△	Veatchite	⟨△2□⟩–⟨2△□⟩,△	$Sr_2[B_5O_8(OH)]_2[B(OH)_3](H_2O)$
3△2□,1△	Veatchite-A	⟨△2□⟩–⟨2△□⟩,△	$Sr_2[B_5O_8(OH)]_2[B(OH)_3](H_2O)$
3△2□,1△	P-veatchite	⟨△2□⟩–⟨2△□⟩,△	$Sr_2[B_5O_8(OH)]_2[B(OH)_3](H_2O)$
3△2□,1△	Volkovskite	⟨△2□⟩–⟨2△□⟩,△	$KCa_4[B_5O_8\ (OH)]_4[B(OH)_3]_2Cl(H_2O)_4$
2△3□	Tuzlaite	⟨△2□⟩–⟨△2□⟩	$NaCa[B_5O_8(OH)_2](H_2O)_3$
3△3□	Nobleite	[ϕ]⟨△2□⟩|⟨△2□⟩|⟨△2□⟩|	$Ca[B_6O_9(OH)_2](H_2O)_3$
3△3□	Tunellite	[ϕ]⟨△2□⟩|⟨△2□⟩|⟨△2□⟩|	$Sr[B_6O_9(OH)_2](H_2O)_3$
5△3□	Strontioborite	[ϕ]⟨△2□⟩|⟨△2□⟩|⟨△2□⟩|2△	$Sr[B_8O_{11}(OH)_4]$
8△6□	Ginorite	[ϕ]⟨△2□⟩|⟨△2□⟩|⟨△2□⟩|–[ϕ]⟨△2□⟩|⟨△2□⟩|⟨△2□⟩|2△	$Ca_2[B_{14}O_{20}(OH)_6](H_2O)_5$
8△6□	Strontioginorite	[ϕ]⟨△2□⟩|⟨△2□⟩|⟨△2□⟩|–[ϕ]⟨△2□⟩|⟨△2□⟩|⟨△2□⟩|2△	$Sr_2[B_{14}O_{20}(OH)_6](H_2O)_5$
2△4□	Fabianite	⟨△2□⟩=⟨4□⟩=⟨△2□⟩	$Ca_2[B_6O_{10}(OH)_2]$
8□	Johachidolite	⟨6□⟩=⟨4□⟩	$Ca_2Al_2[B_6O_{14}]$
4△7□	Preobrazhenskite	□⟨△2□⟩–⟨△2□⟩–⟨△2□⟩–⟨△2□⟩□	$Mg_3[B_{11}O_{15}(OH)_9]$
		Infinite-framework borate minerals	
3□	Metaborite	⟨3□⟩	$B_3O_3(OH)_3$
2△2□	Diomignite	⟨△2□⟩=⟨△2□⟩	$Li_2[B_4O_6]$
2△3□	Hilgardite-1A	⟨△2□⟩–⟨△2□⟩	$Ca_2[B_5O_9]Cl(H_2O)$
2△3□	Hilgardite-4M	⟨△2□⟩–⟨△2□⟩	$Ca_2[B_5O_9]Cl(H_2O)$
2△3□	Hilgardite-3A	⟨△2□⟩–⟨△2□⟩	$Ca_6[B_5O_9]_3Cl_3(H_2O)_3$
2△3□	Tyretskite-1A	⟨△2□⟩–⟨△2□⟩	$Ca_2[B_5O_9](OH)(H_2O)$
1△6□	Boracite (low)	[ϕ]⟨3□⟩|⟨3□⟩|⟨3□⟩|△	$Mg_3[B_3O_{10}]_2(BO_3)I$
1△6□	Chambersite	[ϕ]⟨3□⟩|⟨3□⟩|⟨3□⟩|△	$Mn_3[B_3O_{10}]_2(BO_3)I$
1△6□	Congolite	[ϕ]⟨3□⟩|⟨3□⟩|⟨3□⟩|△	$Fe_3[B_3O_{10}]_2(BO_3)I$
1△6□	Ericaite	[ϕ]⟨3□⟩|⟨3□⟩|⟨3□⟩|△	$Fe_3[B_3O_{10}]_2(BO_3)I$
1△6□	Trembathite	[ϕ]⟨3□⟩|⟨3□⟩|⟨3□⟩|△	$Mg_3[B_3O_{10}]_2(BO_3)I$
4□	Boracite (high)	[ϕ]□|□|□|□	$Mg_3[B_7O_{13}]Cl$
7△8□	Pringleite	⟨△□·⟩=⟨△2□⟩△	$Ca_9[B_{20}O_{28}(OH)_{18}][B_6O_6(OH)_6]Cl_4(H_2O)_{13}$
7△8□	Ruitenbergite	⟨△□·⟩=⟨△2□⟩△	$Ca_9[B_{20}O_{28}(OH)_{18}][B_6O_6(OH)_6]Cl_4(H_2O)_{13}$
8△8□	Penobsquisite	⟨△□·⟩=⟨△2□⟩–⟨△2□⟩	$Ca_2Fe[B_9O_{13}(OH)_6]Cl(H_2O)_4$
		Sulfate-borates	
1□	Sulfoborite	□	$Mg_3(OH)[B(OH)_4]_2(SO_4)F$
2△3□	Heidornite	<△2□>–<△2□>–□	$Na_2Ca_3[B_5O_8(OH)_2](SO_4)_2Cl$
		Phosphate-borates	
1□	Seamanite	□	$Mn_3(OH)_2[B(OH)_4](PO_4)$
1△1□	Lüneburgite	△ □	$Mg_3(H_2O)_6[B_2(OH)_6](PO_4)_2$
		Arsenate-borates	
1□	Cahnite	□	$Ca_2[B(OH)_4](AsO_4)$
2△4□	Teruggite	[ϕ]<3□>|<△2□> <△2□>	$Ca_4Mg[B_6O_8(OH)_6]_2(AsO_4)_2(H_2O)_{14}$
		Carbonate-borates	
1△	Gaudefroyite	△	$Ca_4Mn_3(O,OH)_3(BO_3)_3(CO_3)$
1△	Sakhaite	△	$Ca_3Mg(BO_3)(CO_3)(H_2O)$
1△	Carboborite	□	$Ca_2Mg[B(OH)_4]_2(CO_3)_2(H_2O)_4$
1△	Moydite	□	$(Y,REE)[B(OH)_4](CO_3)$
4□	Borcarite	<4□>	$Ca_4Mg[B_4O_6(OH)_6](CO_3)_2$
		Beryllate-borates	
1△	Berborite	△	$Be_2(BO_3)(OH,F)(H_2O)$
1△	Hambergite	△	$Be_2(BO_3)(OH)$
1□	Rhodizite	□	$(K,Cs)Al_4Be_4(B,Be)_{12}O_{28}$
		Silicate-borates	
1□	Ferroaxinite	□	$Ca_2Fe^{2+}Al_2BSi_4O_{15}(OH)$
1□	Magnesioaxinite	□	$Ca_2MgAl_2BSi_4O_{15}(OH)$
1□	Manganaxinite	□	$Ca_2Mn^{2+}Al_2BSi_4O_{15}(OH)$
1□	Tinzenite	□	$(Ca,Mn^{2+},Fe^{2+})_3Al_2BSi_4O_{15}(OH)$
6□	Cappelenite-(Y)	6□	$Ba(Y,Ce)_6Si_3B_6O_{24}F_2$
2□	Danburite	2□	$CaB_2Si_2O_8$
1△	Reedmergnerite	□	$NaBSi_3O_8$
1□	Datolite	□	$Ca_2B_2Si_2O_8(OH)_2$
1□	Homilite	□	$Ca_2(Fe^{2+},Mg)B_2Si_2O_{10}$
1△	Dumortierite	△	$(Al,□,Ti,Mg)Al_6BSi_3O_{16}(O,OH)^2$
1△	Holtite	△	$Al_6(Al,Ta,□)(BO_3)[(Si,Sb)O_4]_3(O,OH)^3$
2△5□	Garrelsite	□<△2□>–<△2□>□	$NaBa_3Si_2B_7O_{16}(OH)_4$
1△	Grandidierite	△	$(Mg,Fe^{2+})Al_3(BO_3)(SiO_4)O_2$

Borate minerals (*cont.*)

Borate groups (*A*)	Name	Connectivity (*B*)	Formula
		Silicate-borates (cont.)	
2□	Hellandite	2□	$(Ca,Y)_6(Al,Fe^{3+})Si_4B_4O_{20}(OH)_4$
2□	Tadzhikite	2□	$Ca_3(Ce,Nd,La)_2(Ti,Al,Fe)B_4Si_4O_{22}$
1△4□	Howlite	2□,<△2□>	$Ca_2B_5SiO_9(OH)_5$
1□	Hyalotekite	□	$(Ba,Pb,K)_4Ca_2(B,Si)_2(Si,Be)_{10}O_{28}F$
1□	Kalborsite	□	$K_6[Al_4Si_6O_{20}][B(OH)_4]Cl$
□	Leucosphenite	□	$Na_4BaTi_2B_2Si_{10}O_{30}$
2□	Taramellite	2□	$Ba_4(Fe^{3+},Ti)_4Si_8B_2O_{27}O_2Cl_x$
2□	Titantaramellite	2□	$Ba_4(Ti,Fe^{3+})_4Si_8B_2O_{27}O_2Cl_x$
□	Nagashimalite	□	$Ba_4(V^{4+},Ti)_4Si_8B_2O_{27}(O,OH)_2Cl$
1□	Poudretteite	□	$KNa_2B_3Si_{12}O_{30}$
1□	Searlesite	□	$NaBSi_2O_5(OH)_2$
1□	Stillwellite-(Ce)	□	$(Ce,La)BSiO_5$
1△	Tourmaline	△	Group name (see the article Tourmaline)
2□	Tienshanite	2□	$Na_2BaMn^{2+}TiB_2Si_6O_{20}$
1△	Werdingite	△	$(Mg,Fe^{2+})_2Al_{14}B_4Si_4O_{37}$
2△	Wiserite	2△	$(Mn^{2+},Mg)_{14}[B_2O_5]_4(OH)_8[\{Si_xMg_{1-x}\}\{O_{1-x}(OH)_x\}_4]Cl_{2x}$
		Silicate-carbonate-borate	
1△	Harkerite	△	$Ca_{24}Mg_8Al_2(SiO_4)_8(BO_3)_6(CO_3)_{10}(H_2O)_2$

as gemstones but are far less common. *See* BORON; TOURMALINE. Frank C. Hawthorne

Bibliography. E. S. Grew and L. M. Anovitz (eds.), *Reviews in Mineralogy*, vol. 33: *Boron: Mineralogy, Petrology and Geochemistry*, 1996; J. D. Grice et al., Borate minerals, II. A hierarchy of structures based on the borate fundamental building block, *Can. Mineral.*, 37:731–762, 1999.

Bordetella

A genus of gram-negative bacteria which are coccobacilli and obligate aerobes, and fail to ferment carbohydrates. These bacteria are respiratory pathogens. *Bordetella pertussis*, *B. parapertussis*, and *B. bronchiseptica* share greater than 90% of their deoxyribonucleic acid (DNA) sequences and would not warrant separate species designations except that the distinctions are useful for clinical purposes. *Bordetella pertussis* is an obligate human pathogen and is the causative agent of whooping cough (pertussis). *Bordetella parapertussis* causes a milder form of disease in humans and also causes respiratory infections in sheep. *Bordetella bronchiseptica* has the broadest host range, causing disease in many mammalian species, but kennel cough in dogs and atrophic rhinitis, in which infected piglets develop deformed nasal passages, have the biggest economic impact. *Bordetella avium* is more distantly related to the other species. A pathogen of birds, it is of major economic importance to the poultry industry. *See* WHOOPING COUGH.

Infection by all four species is characterized by bacterial adherence to the ciliated cells that line the windpipe (trachea) [see **illus.**]. Proteins on the bacterial surface allow attachment to host cells. In *B. pertussis*, these proteins include filamentous hemagglutinin, pertactin, fimbriae, and BrkA. Ciliated cells infected by *B. pertussis* stop their beating and eventually die. This process is in part mediated by a factor called tracheal cytotoxin, derived from the bacterial peptidoglycan of the cell wall. Humans and other higher organisms detect and respond to peptidoglycan as a warning system for bacterial infections. The peptidoglycan of *Bordetella* is identical to that of other gram-negative bacteria. However, unlike most bacteria, *B. pertussis* releases massive amounts of peptidoglycan, causing an exaggerated immune response that is ultimately deleterious, resulting in self-induced death of the ciliated cells. *Bordetella* also produces protein toxins. The best-characterized is pertussis toxin, made only by *B. pertussis*. This toxin interferes with the mechanisms used by host cells to communicate with one another. Whooping cough symptoms attributed to pertussis toxin include altered insulin secretion, an elevated number of circulating white blood cells (primarily lymphocytes), and defects in the ability of the immune system to detect and fight infections. *Bordetella pertussis*, as well as *B. parapertussis* and *B. bronchiseptica*, produces two other toxins. The adenylate cyclase toxin can enter mammalian cells and generate cyclic adenosine monophosphate, an intracellular signaling molecule. Inappropriate production of this signaling molecule impairs the ability of the immune system to fight infections. The dermonecrotic toxin causes skin lesions following intradermal injection, but its role in disease is unclear.

Bordetella pertussis is spread by coughing and has no environmental reservoir other than infected humans. Whooping cough occurs in three stages: catarrhal, paroxysmal, and convalescent. During the initial, catarrhal phase, the bacteria damage the respiratory tract, resulting in symptoms similar to the common cold (cough, fever, and runny nose). After about a week, patients progress to the paroxysmal phase and develop the characteristic whooping cough, where prolonged expulsive coughing ends

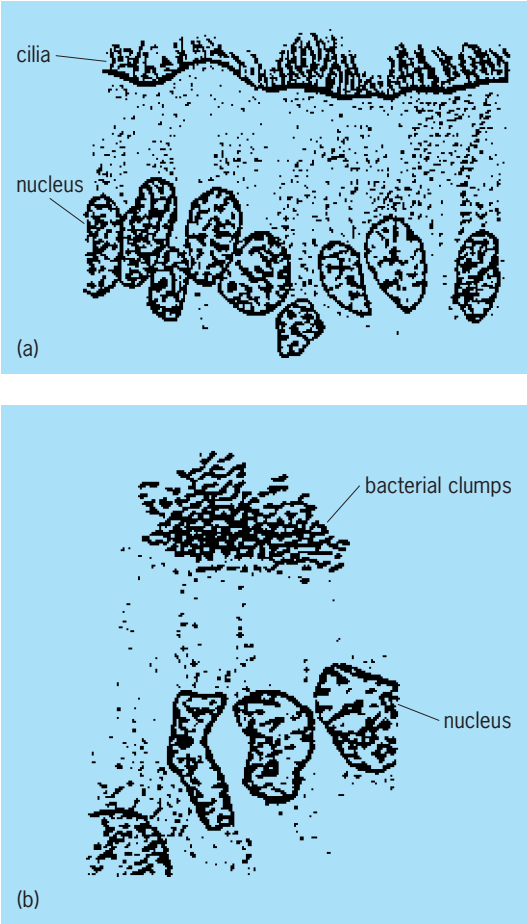

***Bordetella pertussis*–infected trachea. (*a*) Normal human tracheal cells; hairlike cilia project from the cell body, with the nucleus at the bottom. (*b*) Tracheal cells at autopsy from *B. pertussis*–infected child; clumps of bacteria occur between the cilia. (*After F. B. Mallory and A. A. Horner, Pertussis: The histological lesion in the respiratory tract, J. Med. Res., 134:115–123, 1912*)**

with a desperate effort to inhale. The vomiting of thick mucus often follows. The symptoms can last for weeks. After a prolonged illness, patients enter the convalescent stage, and their symptoms gradually improve. Culturing the organism is difficult. Diagnosis is often based on clinical presentation, especially whooping, and is frequently delayed until the paroxysmal phase. Erythromycin is the antibiotic used most frequently to treat whooping cough. Unfortunately, antibiotic treatment improves the patient's condition only if given early, when the disease is most difficult to diagnose, and does not help after whooping has begun. This is consistent with the concept that the early symptoms of the disease result from bacterial damage to the respiratory tract and the later symptoms are due to toxins released by the bacteria. Antibiotics can eradicate the microorganisms but cannot reverse the effects of toxins, which can cause damage far from the site of bacterial growth. Antibiotics can prevent transmission and secondary infections by other microorganisms, a major cause of mortality from whooping cough in the preantibiotic era.

Vaccines have been developed for whooping cough and kennel cough. Multicomponent pertussis vaccines consisting of inactivated pertussis toxin and various combinations of filamentous hemagglutinin, pertactin, and fimbriae are now replacing the older whole-cell vaccines consisting of killed bacteria, which were suspected but not proven to cause rare but serious side effects. Vaccination programs have greatly reduced the incidence of whooping cough in affluent nations, but worldwide nearly half a million deaths occur each year, most of which are vaccine-preventable. *See* ANTIBIOTIC; DRUG RESISTANCE; MEDICAL BACTERIOLOGY; VACUUM METALLURGY.

Alison Weiss

Bibliography. J. D. Cherry, Historical review of pertussis and the classical vaccine, *J. Infect. Dis.*, 174 (suppl. 3):S259–S263, 1996; E. L. Hewlett, *Bordetella* species, *in* G. L. Mandell, J. E. Bennett, and R. Dolin (eds.), *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Disease*, 4th ed., vol. 2, pp. 2078–2084, Churchill Livingstone, 1995; A. A. Weiss, The genus *Bordetella*, in A. Balows et al. (eds.), *The Prokaryotes*, 2d ed., vol. III, pp. 2530–2543, Springer-Verlag, 1991.

Boring bivalves

A variety of marine bivalve mollusks which penetrate solid substrata. They represent seven families and vary in the extent to which they are specialized, in the type of substrata they utilize, and in their method of boring. It is difficult to differentiate borers, burrowers, and nestlers because some species, such as *Hiatella arctica*, not only bore into hard limestones, calcareous sandstones, and chalks, but also burrow in peat or sand-filled crevices in hard rocks, or nestle in holdfasts of *Laminaria*. Undoubtedly the true borers evolved from burrowers via forms that were living in increasingly harder substrata.

Unlike other bivalves which use their shells as abrasive "tools," date mussels (*Lithophaga*, Mytilidae) penetrate calcareous rocks, corals, and shells by chemical means, possibly a weak carbonic acid. These borers are byssally attached and move back and forth in their burrows by alternate contractions of the anterior and posterior byssal retractor muscles. This action presses the thickened, fused, anterior mantle lobes against the end of the burrow, and when they are withdrawn they are covered with fine particles of rock which are ejected through the incurrent siphon along with the pseudofeces. Larvae of *Botula* (another small, elongate mytilid), which settle in crevices of soft rock, corals, or wood, abrade their way into the substrate by continued movement of their shells and siphons.

Rupellaria and *Petricola* (Petricolidae) are nonspecialized borers in peat, firm mud, and soft rock. *Petricola* are usually found in peat, but *Rupellaria* larvae settling in crevices of limestone and coral gradually enlarge their burrows as they grow. *Platyodon* (Myidae), closely related to the soft-shelled clam,

press their valves against the walls of the burrow by engorging the mantle and then abrade the soft rock by continued movement of their unspecialized valves.

Gastrochaena and *Spengleria* (Gastrochaenidae) are specialized for boring by having a closed mantle cavity, a large pedal gape, and a truncate foot, allowing them to press the foot and shell against the burrow wall. Despite their thin, fragile shells, they drill into limestone and coral blocks by mechanical means, making flask-shaped burrows which they line posteriorly with calcareous material to ensure a snug fit around the siphons.

The Pholadacea (Pholadidae and Teredinidae) are worldwide in distribution and are highly specialized for boring into hard substrata. The family Pholodidae (common name: piddocks) is composed of 17 genera, of which 5 are restricted to wood. All species of Teredinidae are obligate wood, nut, or plant-stem borers. The major difference between these two families is the presence of accessory plates in the pholads and pallets in the teredinids. For a discussion of the latter group *see* SHIPWORM.

Most pholads are marine, living at depths ranging from intertidal to about 1000 ft (300 m), with some of the deep-sea wood borers (Xylophagainae) extending to 16,500 ft (5000 m). Specializations for boring include filelike sculpture on the shell, a large pedal gape, a truncate foot, a closed mantle cavity, the reduction of the hinge and ligament, the insertion of the anterior adductor muscle on the outside of the valves anterior to the umbos so that it works in opposition to the posterior adductor muscle, and the pivoting of the valves on a dorsoventral axis passing through the umbos and ventral condyle (**Fig. 1**). The contraction of the anterior adductor muscle brings the valves together anteriorly; the attachment of the foot anteriorly and the contraction of the pedal muscles pull the shell forward against the anterior end of the burrow; and then the forceful contraction of the posterior adductor muscle spreads the valves anteriorly and scrapes the filelike ridges against the wall of the burrow. Water jetted anteriorly from the mantle cavity flushes the burrow, and the debris is ejected from the incurrent siphon as pseudofeces. Accessory plates, developed to protect the exposed anterior adductor muscle, are unique to the Pholadidae. Additional dorsal, ventral, and siphonal plates are produced, and the pedal gape is closed with a callum when adult Martesiinae cease boring (Figs. 1 and **Fig. 2***a*, *b*).

Species in the genera *Zirfaea* (Fig. 2*c*) and *Barnea* hasten the recycling of submerged peat beds, and the activity of rock borers such as *Penitella* (Fig. 2*b*), *Parapholas, Pholas*, and *Chaceia* contribute to the breakdown of intertidal and submerged cliffs and even cement bulkheading. *Pholas dactylus* of Europe, known to penetrate schistose rock, is also famous for its luminescent properties. *Diplothyra smithii* (Fig. 2*a*) and *Penitella parva*, both shell

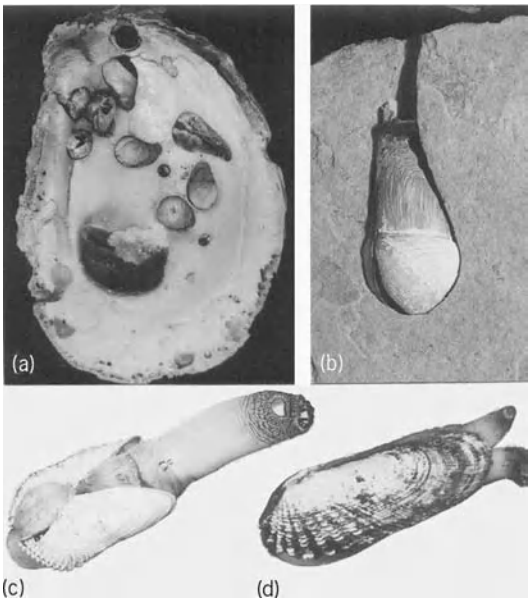

Fig. 2. Typical boring bivalves: (*a*) *Diplothyra smithii* boring in *Crassostrea virginica*; (*b*) *Penitella penita* boring in shale; (*c*) *Zirfaea crispata* removed from burrow, showing foot, anterior adductor muscle, and siphons (*from R. D. Turner, The family Pholadidae in the western Atlantic and eastern Pacific, Johnsonia*, 3(*33–34*):*1–160, 1954–1955*); (*d*) *Petricola pholadiformis* removed from burrow, showing extended siphons.

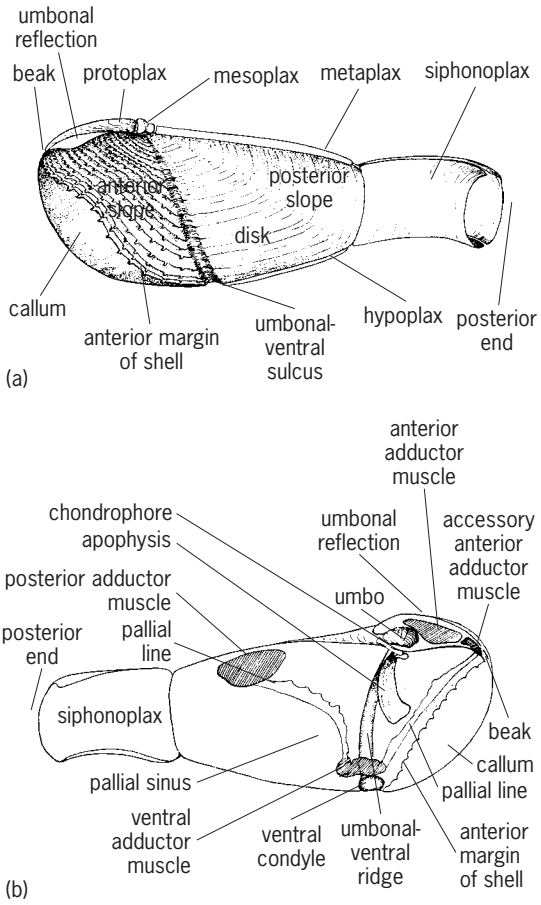

Fig. 1. A pholad shell: (*a*) external view, showing position of accessory plates and parts of the valve; (*b*) internal view, showing position of muscle and parts of the valves. (*After R. D. Turner, The family Pholadidae in the western Atlantic and eastern Pacific, Johnsonia*, 3(*33–34*):*1–160, 1954–1955*)

borers, are pests of oysters and abalones respectively. The wood-boring *Martesia* are important pests of waterfront structures and may even bore into polyvinyl chloride pipes in tropical and warm temperate areas. The Xylophagainae are the most important organisms involved in recycling wood in the deep sea. This subfamily includes three genera and about 30 species. *See* MOLLUSCA. Ruth D. Turner

Bibliography. G. L. Kennedy, *West American Cenozoic Pholadidae* (*Mollusca: Bivalvia*), San Diego Soc. Nat. Hist. Mem. 8, pp. 1–128, 1974; R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, pt. N, vol. 2, 1969; R. D. Purchon, The structure and function of British Pholadidae, *Proc. Zool. Soc. London*, 124:859–911, 1955; R. D. Turner, The family Pholadidae in the western Atlantic and eastern Pacific, *Johnsonia*, 3(33–34):1–160, 1954–1955.

Boron

A chemical element, B, atomic number 5, atomic weight 10.811, in group 13 of the periodic table. It has three valence electrons and is nonmetallic in behavior. It is classified as a metalloid and is the only nonmetallic element which has fewer than four electrons in its outer shell. The free element is prepared in crystalline or amorphous form. The crystalline form is an extremely hard, brittle solid. It is of jet-black to silvery-gray color with a metallic luster. One form of crystalline boron is bright red. The amorphous form is less dense than the crystalline and is a dark-brown to black powder. In the naturally occurring compounds, boron exists as a mixture of two stable isotopes with atomic weights of 10 and 11. *See* PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Many properties of boron have not been sufficiently established experimentally as a result of the questionable purity of some sources of boron, as well as of the variations in the methods and temperatures of preparation. A summary of the physical properties is shown in the **table**.

Physical properties of boron

Property	Temp., °C	Value
Density		
Crystalline	25–27	2.31 g/cm^3
Amorphous	25–27	2.3 g/cm^3
Mohs hardness		
Crystalline		9.3
Melting point		2100°C
Boiling point		2500°C
Resistivity	25	1.7×10^{-6} ohm-cm
Coefficient of thermal expansion	20–750	8.3×10^{-6} cm/°C
Heat of combustion	25	302.0 ± 3.4 kcal/mole
Entropy		
Crystalline	25	1.403 cal/(mole)(deg)
Amorphous	25	1.564 cal/(mole)(deg)
Heat capacity		
Gas	25	4.97 cal/(mole)(deg)
Crystalline	25	2.65 cal/(mole)(deg)
Amorphous	25	2.86 cal/(mole)(deg)

Boron and boron compounds have numerous uses in many fields, although elemental boron is employed chiefly in the metal industry. Its extreme reactivity at high temperatures, particularly with oxygen and nitrogen, makes it a suitable metallurgical degasifying agent. It is used to refine the grain of aluminum castings and to facilitate the heat treatment of malleable iron. Boron considerably increases the high-temperature strength characteristics of alloy steels. Elemental boron is used in the atomic reactor and in high-temperature technologies. The physical properties that make boron attractive as a construction material in missile and rocket technology are its low density, extreme hardness, high melting point, and remarkable tensile strength in filament form. When boron fibers are used in an epoxy (or other plastic) carrier material or matrix, the resulting composite is stronger and stiffer than steel and 25% lighter than aluminum. Refined borax, $Na_2B_4O_7 \cdot 10H_2O$, is an important ingredient of a variety of detergents, soaps, water-softening compounds, laundry starches, adhesives, toilet preparations, cosmetics, talcum powder, and glazed paper. It is also used in fireproofing, disinfecting of fruit and lumber, weed control, and insecticides, as well as in the manufacture of leather, paper, and plastics.

Boron makes up 0.001% of the Earth's crust. It is never found in the uncombined or elementary state in nature. Besides being present to the extent of a few parts per million in sea water, it occurs as a trace element in most soils and is an essential constituent of several rock-forming silicate minerals, such as tourmaline and datolite. The presence of boron in extremely small amounts seems to be necessary in nearly all forms of plant life, but in larger concentrations, it becomes quite toxic to vegetation. Only in a very limited number of localities are high concentrations of boron or large deposits of boron minerals to be found in nature; the more important of these seem to be primarily of volcanic origin. *See* BORATE MINERALS. F. H. May; V. V. Levasheff

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; J. J. Pouch and S. A. Alterovitz, *Synthesis and Properties of Boron Nitride*, 1990; K. Smith, *Organometallic Compounds*

of Boron, 1985; S. E. Thomas, *Organic Synthesis: The Role of Boron and Silicon*, 1992.

Borrelia

A genus of spirochetes that have a unique genome composed of a linear chromosome and numerous linear and circular plasmids. Borreliae are motile, helical organisms with 4–30 uneven, irregular coils, and are 5–25 micrometers long and 0.2–0.5 μm wide. Their locomotory apparatus consists of 15–22 fibrils coiled around the cell body and situated between the elastic envelope and cytoplasmic membrane. Motion is forward and backward, laterally by bending and looping, and corkscrewlike. Borreliae multiply by binary fission. *See* BACTERIA.

All borreliae are arthropod-borne. Of the 24 recognized species, 21 cause relapsing fever and similar diseases in human and rodent hosts; two are responsible for infections in ruminants and horses; and the remaining one, for borreliosis in birds.

The borreliae of human relapsing fevers are transmitted by the body louse (*Pediculus humanus humanus*) or by a large variety of soft-shelled ticks of the genus *Ornithodoros*. The species *B. burgdorferi*, the etiologic agent of Lyme disease and related disorders, is transmitted by ticks of the genus *Ixodes* (*I. scapularis* and *I. pacificus* in North America, *I. ricinus* in Europe, *I. persulcatus* and other species in Asian countries). *See* RELAPSING FEVER.

Investigations into the genetic (16S rRNA) and phenotypic characterization of the Lyme disease spirochete, *B. burgdorferi*, have shown genetic differences that led to the distinction of three genomic groups designated as *B. burgdorferi* sensu stricto, *B. garinii*, and *B. afzelii*. Tick or spirochete surveys in European and Asian countries produced at least eight new species of spirochetes that appear closely related to the above-cited genomic groups.

Borrelia anserina, which causes spirochetosis in geese, ducks, turkeys, pheasants, chickens, and other birds, is propagated by ticks of the genus *Argas*. Various species of ixodid ticks are responsible for transmitting *B. theileri* among cattle, horses, and sheep. *Borrelia coriaceae*, isolated from *O. coriaceus*, is the putative cause of epizootic bovine abortion in the western United States.

Borreliae stain well with nearly all aniline dyes and can be demonstrated in tissue sections by silver impregnation techniques. Dark-field microscopy is used for rapid examination and detection of spirochetes in peripheral blood or in vector tissues. Also useful for differentiating and identifying spirochetal isolates are molecular procedures such as polyacrylamide gel electrophoresis (SDS-PAGE) of whole-cell lysates for the detection of certain outer-surface proteins (Osp), deoxyribonucleic acid (DNA) hybridization probes, plasmid profiles, restriction endonuclease patterns of total DNA, and the polymerase chain reaction (PCR). The last offers the potential for detecting a very small number of spirochetes in clinical samples as well as in infected ticks and vertebrate tissues.

SDS-PAGE of spirochetes has shown that the outer surface of the microorganisms contains numerous variable lipoproteins of which at least two (Osp A, Osp B) are abundant. The antigenic variability is well known for the relapsing fever borreliae. A switch in the major outer-surface proteins leads to recurrent spirochetemias. The ability to produce new surface antigens results from a nonreciprocal exchange of sequences between numerous silent genes and a single expression locus. The Lyme disease spirochete, *B. burgdorferi*, contains related highly variable surface proteins whose role and mechanisms of its locus needs to be defined. In addition to antigenic variations in the mammalian hosts, *B. burgdorferi* alters its surface proteins in its tick vectors.

Microaerophilic borreliae can readily be cultivated in their tick vectors or animal hosts. Kelly's medium and modifications thereof are suitable for culturing several species of relapsing fever spirochetes and the agent of Lyme disease.

Tetracyclines, penicillins, and doxycycline are the most effective antibiotics for treatment of spirochetes.

Two vaccines consisting of recombinant *B. burgdorferi* Osp A (one vaccine with, the other without adjuvant) have been evaluated in subjects of risk for Lyme disease. Both proved safe and effective in the prevention of this disease. *See* ANTIBIOTIC; MEDICAL BACTERIOLOGY. Willy Burgdorfer; Patricia Rosa

Bose-Einstein condensation

When a gas of bosonic particles is cooled below a critical temperature, it condenses into a Bose-Einstein condensate. The condensate consists of a macroscopic number of particles, which are all in the ground state of the system. Bose-Einstein condensation (BEC) is a phase transition, which does not depend on the specific interactions between particles. It is based on the indistinguishability and wave nature of particles, both of which are at the heart of quantum mechanics.

Basic phenomenon in ideal gas. In a simplified picture, particles in a gas may be regarded as quantum-mechanical wave packets which have a spatial extent on the order of a thermal de Broglie wavelength, given by Eq. (1), where T is the temperature, m the

$$\lambda_{dB} = \left(\frac{2\pi^2\hbar}{mk_BT}\right)^{1/2} \tag{1}$$

mass of the particle, k_B is Boltzmann's constant, and $\hbar$ is Planck's constant divided by 2π. The wavelength λ_{dB} can be regarded as the position uncertainty associated with the thermal momentum distribution of the particles. At high temperature, λ_{dB} is small, and the probability of finding two particles within

this distance of each other is extremely low. Therefore, the indistinguishability of particles is not important, and a classical description applies (namely, Boltzmann statistics). When the gas is cooled to the point where λ_{dB} is comparable to the distance between particles, the individual wavepackets start to overlap and the indistinguishability of particles becomes crucial—it is like an identity crisis for particles. For fermions, the Pauli exclusion principle prevents two particles from occupying the same quantum state; whereas for bosons, quantum statistics (in this case, Bose-Einstein statistics) dramatically increases the probability of finding several particles in the same quantum state. In the latter case, as predicted by Einstein in 1925, the system undergoes a phase transition and forms a Bose-Einstein condensate, where a macroscopic number of particles occupy the lowest-energy quantum state (**Fig. 1**). The temperature and the density of particles n at the phase transition are related by Eq. (2).

$$n\lambda_{dB}^3 = 2.612 \tag{2}$$

Bose-Einstein condensation is a phase transition caused solely by quantum statistics, in contrast to other phase transitions (such as melting or crystallization) which depend on the interactions between particles. *See* BOLTZMANN STATISTICS; DE BROGLIE WAVELENGTH; EXCLUSION PRINCIPLE; NONRELATIVISTIC QUANTUM THEORY; PHASE TRANSITIONS; QUANTUM MECHANICS; UNCERTAINTY PRINCIPLE.

Bose-Einstein condensation can be described intuitively in the following way: When the quantum-mechanical wave functions of bosonic particles spatially overlap, the matter waves start to oscillate in concert. A coherent matter wave forms that comprises all particles in the ground state of the system. This transition from disordered to coherent matter waves can be compared to the step from incoherent light to laser light. Indeed, atom lasers based on Bose-Einstein condensation have been realized. *See* COHERENCE; LASER.

Bose-Einstein condensation occurs in thermal equilibrium. This is the way that nature maximizes entropy for a given number of atoms and a given amount of energy when the total energy is below the threshold for the phase transition to a Bose-Einstein condensate. Although the condensate itself does not contribute to the entropy, it allows the other particles to have a larger share of the total energy and thus increase the entropy.

Experimental techniques. The phenomenon of Bose-Einstein condensation is responsible for the superfluidity of helium and for the superconductivity of an electron gas, which involves Bose-condensed electron pairs. However, these phenomena happen at high density, and their understanding requires a detailed treatment of the interactions. *See* LIQUID HELIUM; SUPERCONDUCTIVITY; SUPERFLUIDITY.

The quest to realize Bose-Einstein condensation in a dilute weakly interacting gas was pursued in several directions: liquid helium diluted in vycor, a sponge-like glass; excitons in semiconductors, which consist of weakly bound electron-hole pairs; and atomic gases. Most of the research activities have focused on Bose-Einstein condensation in atomic gases.

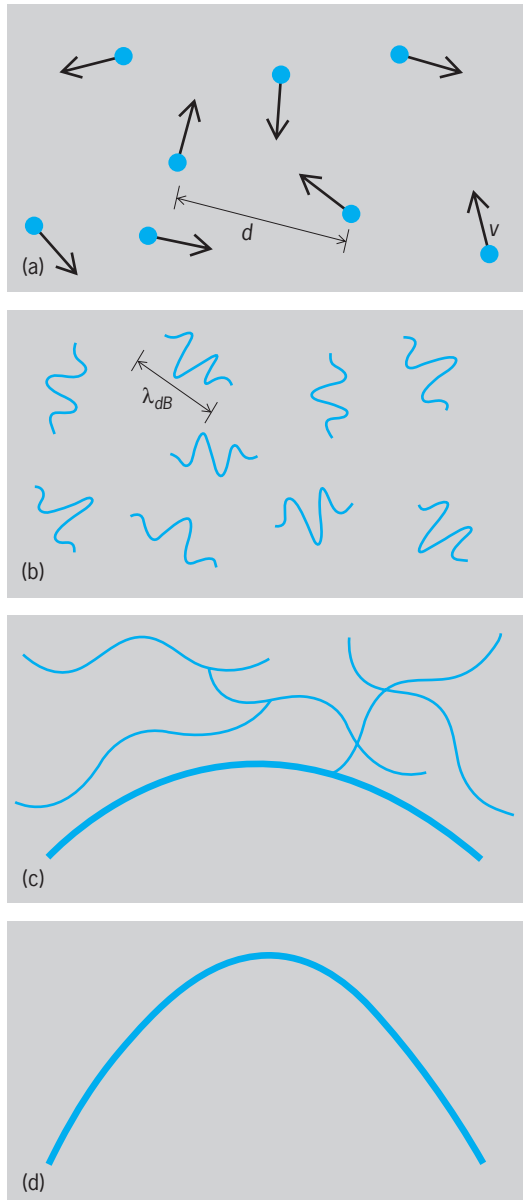

Fig. 1. Criterion for Bose-Einstein condensation in a gas of weakly interacting particles. (*a*) Gas at high temperature, treated as a system of billiard balls, with thermal velocity *v* and density d^{-3}, where *d* is the distance between particles. (*b*) Simplified quantum description of gas at low temperature, in which the particles are regarded as wave packets with a spatial extent of the order of the de Broglie wavelength, λ_{dB}. (*c*) Gas at the transition temperature for Bose-Einstein condensation, when λ_{dB} becomes comparable to *d*. The wave packets overlap and a Bose-Einstein condensate forms (in the case of bosonic particles). (*d*) Pure Bose condensate (giant matter wave), which remains as the temperature approaches absolute zero and the thermal cloud disappears. (*After D. S. Durfee and W. Ketterle, Experimental studies of Bose-Einstein condensation, Opt. Express, 2:299–313, Optical Society of America, 1998*)

At ultralow temperatures, all atomic gases liquefy or solidify in thermal equilibrium. Keeping the gas at

Properties of a gas during successive stages of cooling to a Bose-Einstein condensate*

Stage	Temperature	Number density of atoms, cm^{-3}	Phase-space density
Oven	500 K	10^{14}	10^{-13}
Laser cooling	50 μK	10^{11}	10^{-6}
Evaporative cooling	500 nK	10^{14}	2.612
Bose-Einstein condensate		10^{15}	10^{7}

*Numbers refer to cooling of sodium vapor in an experiment at the Massachusetts Institute of Technology. All numbers are approximate.

sufficiently low density can prevent this from occurring. Elastic collisions, which ensure thermal equilibrium of the gas, are then much more frequent than three-body collisions, which are required for the formation of molecules and other aggregates. As a result, the gaseous phase is metastable for seconds or even minutes and allows for the observation of Bose-Einstein condensation in a gas. Typical number densities of atoms between 10^{12} and $10^{15}/cm^3$ imply transition temperatures for Bose-Einstein condensation in the nanokelvin or microkelvin regime.

The realization of Bose-Einstein condensation in atomic gases required techniques to cool gases to such low temperatures, and atom traps to confine the gas at the required density and keep them away from the much warmer walls of the vacuum chamber. Most experiments use several laser-cooling techniques as precooling, then hold the atoms in a magnetic trap and cool them further by forced evaporative cooling. For atomic hydrogen, the laser-cooling step is replaced by cryogenic cooling. Evaporative cooling is done by continuously removing the high-energy tail of the thermal distribution from the trap. In a magnetic trap, this can be accomplished by flipping the atomic spins of the highest energy particles with radio-frequency radiation. Since this process reverses the magnetic force on the particles, they are expelled from the trapping region. The evaporated atoms carry away more than the average energy, which means that the temperature of the remaining atoms decreases. The high-energy tail must be constantly repopulated by collisions, thus maintaining thermal equilibrium and sustaining the cooling process. Evaporative cooling is a common phenomenon in daily life—it's how hot water cools down in a bathtub or hot coffee in a cup. In optical traps, evaporation is forced by lowering the laser power and thus the trap depth. In addition to the early observation of Bose-Einstein condensation in rubidium, sodium, lithium, and hydrogen, this condensation has now been observed also in potassium, cesium, ytterbium, chromium, metastable helium, and molecules of potassium and lithium.

In a typical cooling process (see **table**), the temperature of the gas is reduced by a factor of 10^9 though the combination of laser and evaporative cooling steps, while the density at the Bose-Einstein condensation transition is similar to the initial density in the atomic beam oven. In each step, the phase-space density, to which the ground-state population is proportional, increases by a factor of 10^6, resulting in a total increase of phase-space density by 20 orders of magnitude.

Observation. Bose-Einstein condensation was observed by suddenly switching off the atom trap, releasing the atom cloud into a ballistic expansion. The cloud was illuminated with resonant laser light, and its shadow was imaged onto a CCD (charge-coupled-device) camera (**Fig. 2**). Such images showed the sudden appearance, near the predicted transition temperature, of a bimodal cloud consisting of a diffuse normal component and a dense core (the condensate). The condensate expanded with a kinetic energy that was much larger than the zero-point energy of the trapping potential—clear evidence for repulsive interactions between the atoms, which accelerate the ballistic expansion. *See* CHARGE-COUPLED DEVICES.

The sudden reduction of the speed of ballistic expansion has been called condensation in momentum space. However, Bose-Einstein condensation is a condensation into the lowest energy state; in an inhomogeneous trapping potential, the lowest state has the smallest spatial extent. Condensation into this state thus results in the formation of a dense core of Bose condensed atoms within a diffuse thermal cloud of atoms. This spatial structure of a Bose condensate was directly observed using dispersive imaging techniques.

Weak interactions. Bose-Einstein condensation in gases allows for a theoretical description based

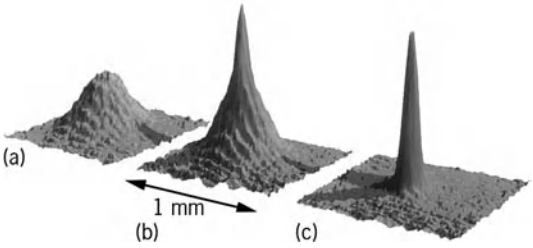

Fig. 2. Observation of Bose-Einstein condensation by absorption imaging. Absorption is plotted in two spatial dimensions. The Bose-Einstein condensate is characterized by its slow expansion, observed after a 6-millisecond time of flight. Shown are clouds expanding (*a*) after cooling to just above the transition point ($T > T_c$), (*b*) just after the condensate appeared ($T < T_c$), and (*c*) after further evaporative cooling has left an almost pure condensate ($T \ll T_c$). The total number of sodium atoms at the phase transition is about 7×10^5; the temperature, T_c, at the transition point is 2 μK. (*After D. S. Durfee and W. Ketterle, Experimental studies of Bose-Einstein condensation, Opt. Express, 2:299t–313, Optical Society of America, 1998*)

directly on first principles because there is a clear hierarchy of length and energy scales. In a gas, the separation between atoms, $n^{-1/3}$, is much larger than the effective range of the interatomic forces (characterized by the *s*-wave scattering length a, typically 1–5 nanometers for alkali atoms), that is, $na^3 \ll 1$. This inequality expresses the fact that binary collisions are much more frequent than three-body collisions. It is in this limit that the theory of the weakly interacting Bose gas applies. Interactions between the Bose-condensed particles are described by a mean field energy given by Eq. (3), which is the effective

$$U_{\text{int}} = 4\pi\hbar^2 na/m \tag{3}$$

potential energy of one particle due to the presence of the other particles in the condensate. This interaction is responsible for collective phenomena such as sound. The stability of large condensates requires repulsive interactions (positive a). For negative a, the interactions are attractive, and the condensate is unstable against collapse above a certain size.

Properties of the weakly interacting Bose gas have been studied both experimentally and theoretically. Of special interest are the collective excitations of the gas, different forms of sound (Bogoliubov, first and second sound), and the stability of vortices. *See* QUANTIZED VORTICES; QUANTUM ACOUSTICS.

When condensates were realized in a mixture of atomic hyperfine states, multicomponent condensates were observed. Their phase diagram, the spatial structure, possible metastable configurations, and properties such as miscibility and immiscibility are examples of new phenomena which have been studied. These examples show that gaseous Bose condensates are a new class of quantum fluids with properties quite different from the quantum liquids helium-3 and helium-4. Gaseous condensates provide a testing ground for aspects of many-body theories, which were developed many decades ago but never tested experimentally.

Macroscopic wave function. In an ideal gas, Bose condensed atoms all occupy the same single-particle ground-state wave function. The many-body ground-state wave function is then the product of N identical single-particle ground-state wave functions, where N is the number of condensed atoms. This single-particle wave function is therefore called the condensate wave function or macroscopic wave function. This picture is valid even when weak interactions are included. The ground-state many-body wave function is still, to a very good approximation, a product of N single-particle wave functions, which are now obtained from the solution of a nonlinear Schrödinger equation, which includes the interaction energy between atoms. The admixture of other configurations into the ground state is called quantum depletion. The theory of weakly interacting Bose gases shows that the fractional quantum depletion is given by expression (4), typically 1% or less for the

$$\frac{8}{3\pi^{1/2}}\sqrt{na^3} \tag{4}$$

alkali condensates. This means that even for the interacting gases, with 99% accuracy, all the atoms can be regarded as having the same single-particle wave function. This is in contrast to liquid helium, in which the quantum depletion is about 90%.

The phase of the condensate and the coherence properties of its wave function are relevant for quantum fluids because the gradient of the phase is proportional to the superfluid velocity. It also determines the properties of atom lasers based on Bose-Einstein condensation. In superconductors and liquid helium, the existence of coherence and of a macroscopic wave function is impressively demonstrated through the Josephson effect. In the dilute atomic gases, the coherence has been demonstrated even more directly by interfering two Bose condensates (**Fig. 3**). The interference fringes have a spacing of 15 micrometers, a huge length for matter waves. (In contrast, the matter wavelength of atoms at room temperature is only 0.05 nm, less than the size of the atoms.)

Strongly correlated quantum gases. The study of properties of the weakly interacting Bose gas have confirmed many long-standing predictions and led to a more profound understanding of the properties of Bose-Einstein condensates. Bose-Einstein condensates display the physics of single particles occupying a macroscopic wave function, and interactions are described by a mean-field potential. More recently, researchers have used the Bose-Einstein condensate as the starting point to "engineer" other new forms of matter, which are characterized by strong interactions and strong correlations. Included are atoms in optical lattices, ultracold atoms in one-dimensional geometries, and pairing and strong interactions near Feshbach resonances.

Optical lattices are created by several optical standing waves and provide an "egg carton" potential for neutral atoms. When the depth of the optical lattice is increased, the cold gas undergoes a quantum phase transition from a superfluid to an insulator. An atom trap with very tight transverse confinement realizes a one-dimensional geometry. At low density, the atoms

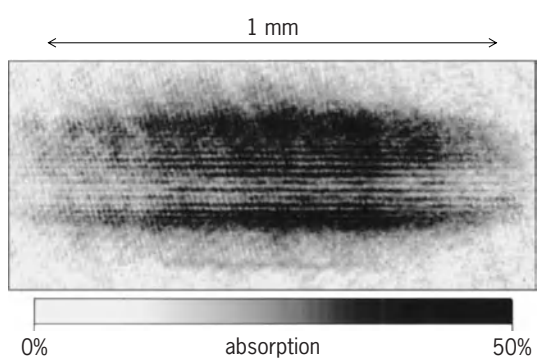

Fig. 3. Interference pattern of two expanding condensates, demonstrating the coherence of Bose-Einstein condensates. This absorption image was observed after a 40-millisecond time of flight. Interference fringes have a spacing of 15 μm. (*After D. S. Durfee and W. Ketterle, Experimental studies of Bose-Einstein condensation, Opt. Express, 2:299–313, Optical Society of America, 1998*)

are lined up like beads on a string and form a so-called Tonks gas, a gas of impenetrable bosons.

Two atoms interact most strongly when they couple resonantly to a molecular state. If the colliding atoms and the bound molecule have different spin configurations, the atoms and the molecule experience different energy shifts due to an external magnetic field (Zeeman shifts). It is possible to tune the magnetic field to a value where the bound molecular state has the same energy as the colliding atoms. This resonance is called a Feshbach resonance. Feshbach resonances have been used to create strongly interacting Bose-Einstein condensates and have triggered condensate explosions and implosions. The resonant recombination of atoms into molecules has led to the first molecular Bose-Einstein condensates. And finally, Feshbach resonances have induced the pairing of fermions in a regime of attractive interactions, where no molecular bound state exists in a vacuum. Those fermion pairs were shown to Bose-condense and to form a superfluid. A magnetic field variation across a Feshbach resonance transforms a molecular Bose-Einstein condensate into a condensate of fermion pairs, which can be regarded as tightly bound Cooper pairs and are described by an extension of the Bardeen-Cooper-Schrieffer (BCS) theory for superfluids. Feshbach resonances thus provide access to the long-sought BEC-BCS crossover from a condensate of tightly bound fermions (like molecules) to a BCS superfluid. *See* ATOMIC FERMI GAS; RESONANCE (QUANTUM MECHANICS).

Wolfgang Ketterle

Bibliography. J. R. Anglin and W. Ketterle, Bose-Einstein condensation of atomic gases, *Nature*, 416:211–218, 2002; E. A. Cornell and C. E. Wieman, The Bose-Einstein condensate, *Sci. Amer.*, pp. 40–45, March 1998; A. Griffin, D. W. Snoke, and S. Stringari (eds.), *Bose-Einstein Condensation*, Cambridge University Press, 1995; M. Inguscio, S. Stringari, and C. E. Wieman (eds.), Bose-Einstein Condensation in Atomic Gases, *Proc. Int. School Phys. Enrico Fermi*, Course CXL, IOS Press, Amsterdam, 1999; C. J. Pethick and H. Smith, *Bose-Einstein Condensation in Dilute Gases*, Cambridge University Press, 2002; L. P. Pitaevskii and S. Stringari, *Bose-Einstein Condensation*, Clarendon Press, Oxford, 2003.

Bose-Einstein statistics

The statistical description of quantum-mechanical systems in which there is no restriction on the way in which particles can be distributed over the individual energy levels. This description applies when the system has a symmetric wave function. This in turn has to be the case when the particles described are of integer spin.

Distribution probability. Suppose one describes a system by giving the number of particles n_i in an energy state ϵ_i, where the n_i are called occupation numbers and the index i labels the various states. The energy level ϵ_i is of finite width, being really a range of energies comprising, say, g_i individual (nondegenerate) quantum levels. If any arrangement of particles over individual energy levels is allowed, one obtains for the probability of a specific distribution, Eq. (1*a*). In Boltzmann statistics, this same probability would be written as Eq. (1*b*).

$$W = \prod_i \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!} \tag{1a}$$

$$W = \prod_i \frac{{g_i}^{n_i}}{n_i!} \tag{1b}$$

See BOLTZMANN STATISTICS.

The equilibrium state is defined as the most probable state of the system. To obtain it, one must maximize Eq. (1*a*) under the conditions given by Eqs. (2*a*) and (2*b*), which express the fact that the total number of particles N and the total energy E are fixed.

$$\sum n_i = N \tag{2a}$$

$$\sum \epsilon_i n_i = E \tag{2b}$$

One finds for the most probable distribution that Eq. (3) holds. Here, A and β are parameters to be determined from Eqs. (2*a*) and (2*b*); actually, $\beta = 1/kT$ where k is the Boltzmann constant and T is the absolute temperature.

$$n_i = \frac{g_i}{\frac{1}{A}e^{\beta\epsilon_i} - 1} \tag{3}$$

In the classical case, that is, when Boltzmann statistics is employed, the equilibrium distribution may be obtained from a specific assumption about the number of collisions of a certain kind. One assumes that the number of collisions per second in which molecules with velocities in cells i and j in phase space produce molecules with velocities in cells k and l is given by Eq. (4), where a_{ij}^{kl} is a geometrical factor. In the Bose case, Eq. (3) may be obtained in a similar way from a collision number assumption, which is written as Eq. (5).

$$A_{ij}^{kl} = n_i n_j a_{ij}^{kl} \tag{4}$$

$$A_{ij}^{kl} = a_{ij}^{kl} n_i n_j \left(\frac{g_k + n_k}{g_k}\right)\left(\frac{g_l + n_l}{g_l}\right) \tag{5}$$

One observes the interesting fact that the number of collisions depends on the number of particles in the state to which the colliding particles are going. The more heavily populated these states are, the more likely a collision is. Quite often one defines f_i, the distribution function, by Eq. (6).

$$n_i = f_i g_i \tag{6}$$

Applications. An interesting and important result emerges when one applies Eq. (3) to a gas of photons, that is, a large number of photons in an enclosure. (Since photons have integer spin, this is legitimate.) For photons one has Eqs. (7*a*) and (7*b*), where h is

$$\epsilon = h\nu \tag{7a}$$

$$p = \frac{h\nu}{c} \tag{7b}$$

the Planck constant, υ the frequency of the photon, c the velocity of light, ϵ the energy of the photon, and p the momentum of the photon. The number of photons in a given energy or frequency range is given by Eq. (8), where V is the volume of the enclosure con-

$$g = \frac{8\pi}{c^3} V \nu^2 \, d\nu \tag{8}$$

taining the photons. Actually, for a gas of photons, Eq. (2*a*) is not necessary (the number of photons is not fixed), and thus the distribution function depends on just one parameter, β; A can be shown to be unity. If one also uses the fact that $\beta = 1/kT$, one obtains from Eq. (3), for the number of photons in the frequency range $d\nu$, Eq. (9). The energy den-

$$n(\nu)\, d\nu = \frac{8\pi}{c^3} V \frac{\nu^2 \, d\nu}{e^{h\nu/kT} - 1} \tag{9}$$

sity (energy per unit volume) in the frequency range $d\nu$ is, by Eqs. (9) and (7*a*), Eq. (10). Equation (10)

$$\rho(\nu)\, d\nu = \frac{8\pi h}{c^3} \frac{\nu^3 \, d\nu}{e^{h\nu/kT} - 1} \tag{10}$$

is the Planck radiation formula for blackbody radiation. Thus, blackbody radiation must be considered as a photon gas, with the photons satisfying Bose-Einstein statistics. *See* HEAT RADIATION.

For material particles of mass m contained in a volume V, one may write Eq. (3) as (11), where

$$f(v_x v_y v_z)\, dv_x\, dv_y\, dv_z = \left(\frac{m}{h}\right)^3 V \frac{dv_x\, dv_y\, dv_z}{\frac{1}{A} e^{mv^2/2kT} - 1} \tag{11}$$

$\upsilon = \sqrt{\upsilon_x^2 + \upsilon_y^2 + \upsilon_z^2}$ is the total velocity. Equations (2*a*) and (2*b*) may now be written as integrals. If the so-called virial theorem, written as Eq. (12), is used

$$PV = {}^2/_3 E \tag{12}$$

where P is the pressure, V the volume, and E the energy, Eqs. (2*a*) and (2*b*) yield a pair of implicit equations, which give the equation of state of an ideal Bose gas, Eqs. (13*a*) and (13*b*). Here λ^3 is defined as

$$\frac{N}{V}\lambda^3 = B_{1/2}(A) \tag{13a}$$

$$\lambda^3 \frac{P}{kT} = B_{3/2}(A) \tag{13b}$$

in Eq. (14) while Eq. (15) holds. In Eq. (15) Γ is the

$$\lambda^3 = \frac{h^2}{2\pi mkT} \tag{14}$$

$$B_\rho(A) = \frac{1}{\Gamma(\rho + 1)} \int_0^\infty \frac{u^\rho \, du}{\frac{1}{A} e^u - 1} \tag{15}$$

usual Γ function and $u = m\upsilon^2/2kT$. From Eqs. (13*a*) and (13*b*), one obtains the relation between P, V, and T by eliminating A.

Now if one develops the numerator of the integral in Eq. (15) in an infinite series, one obtains the so-called Einstein equations, (16*a*) and (16*b*). It is

$$\frac{N}{V} = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} \frac{A^l}{l_{3/2}} \tag{16a}$$

$$P = \frac{kT}{\lambda^3} \sum_{l=1}^{\infty} \frac{A^l}{l^{5/2}} \tag{16b}$$

easy to verify that the sums of Eqs. (16*a*) and (16*b*) diverge for $A > 1$; however, they still converge for $A = 1$. For $A = 1$, $N/V = 2.61/\lambda^3$ and $P = 1.34kT/\lambda^3$. Einstein interpreted $N/V = 2.61/\lambda^3$ as a maximum possible density. If the gas is compressed beyond this point, the superfluous particles will condense in a zero state, where they do not contribute to the density or the pressure. If the volume is decreased, this curious Bose-Einstein condensation phenomenon results, yielding the zero state which has the paradoxical properties of not contributing to the pressure, volume, or density. The particles in the zero state are coherently matched with each other, analogous to the photons in a laser beam sharing a common quantum-mechanical wave function and losing their individual identities. Many of the superfluid properties exhibited by liquid helium are believed to be manifestations of such a condensation. In 1995, Bose-Einstein condensation was first observed directly in a cloud of rubidium atoms that had been cooled to 1.7×10^{-7} K through laser cooling followed by evaporative cooling in a magnetic trap. Since then it has been observed in a number of dilute gases. *See* BOSE-EINSTEIN CONDENSATION; COHERENCE; FERMI-DIRAC STATISTICS; LASER COOLING; LIQUID HELIUM; NONRELATIVISTIC QUANTUM THEORY; PARTICLE TRAP; QUANTUM STATISTICS; STATISTICAL MECHANICS.

Max Dresden

Botanical gardens

Gardens for the culture of plants collected chiefly for scientific and educational purposes. Such a garden is more properly called a botanical institution, in which the outdoor garden is but one portion of an organization including the greenhouse, the herbarium, the library, and the research laboratory. *See* HERBARIUM.

Famous botanical gardens. It was only in modern Europe, after the foundation of the great medieval universities, that botanical gardens for educational purposes began to be established in connection with the schools. The oldest gardens are those in Padua (established 1533) and at Pisa (1543). The botanical garden of the University of Leiden was begun in 1587 and the first greenhouse is said to have been constructed there in 1599. Louis XIII authorized the establishment of a royal garden in Paris "for the instruction of students"; it was opened to the public in 1640 under the name Jardin du Roy. Later the name was changed to the Museé National d'Histoire Naturelle. The Oxford University Botanic Garden, the first in Great Britain, was established in 1621. In Berlin the Botanischer Garten was established in 1646. In 1655

a garden was founded at Uppsala, Sweden, of which Carolus Linnaeus, the father of modern plant taxonomy, was director from 1742 to 1777. In Russia Peter the Great's Druggist's Garden was founded at St. Petersburg in 1713. In 1817 the Conservatoire et Jardin Botanique were established at Geneva and became one of the leading botanical centers of the world. The Royal Botanical Gardens at Kew, England, were officially opened in 1841. This institution came to be known as the botanical capital of the world.

The first of the great tropical gardens was founded at Calcutta in 1787. The original name, Royal Botanic Garden, was changed in 1947 to Indian Botanic Garden. Another great tropical garden, the Jardin Botanico of Rio de Janeiro, was founded in 1808. The great tropical botanical graden of Buitenzorg (Bogor), Java, which originated in 1817, has an area of 205 acres with an additional 150 acres in the Mountain Garden.

Gardens in North America. The first great garden of the United States was founded by Henry Shaw at St. Louis in 1859, and in now known as the Missouri Botanical Garden. The New York Botanical Garden was chartered in 1891 and the Brooklyn Botanic Garden in 1910. The Jardin Botanique of Montreal, the leading garden of Canada, was opened in 1936. *See* ARBORETUM.

Earl L. Core

Botany

That branch of biological science which embraces the study of plants and plant life. According to the specific objectives of the investigators, botanical studies may range from microscopic observations of the smallest and obscurest plants to the study of the trees of the forest. One botanist may be interested mainly in the relationships among plants and in their geographic distribution, whereas another may be primarily concerned with structure or with the study of the life processes taking place in plants.

Botany may be divided by subject matter into several specialties, such as plant anatomy, plant chemistry, plant cytology, plant ecology (including autecology and synecology), plant embryology, plant genetics, plant morphology, plant physiology, plant taxonomy, ethnobotany, and paleobotany. It may also be divided according to the group of plants being studied; for example, agrostology, the study of grasses; algology (phycology), the study of algae; bryology, the study of mosses; mycology, the study of fungi; and pteridology, the study of ferns. Bacteriology and virology are also parts of botany in a broad sense. Furthermore, a number of agricultural subjects have botany as their foundation. Among these are agronomy, floriculture, forestry, horticulture, landscape architecture, and plant breeding. *See* AGRICULTURE; AGRONOMY; BACTERIOLOGY; BREEDING (PLANT); CELL BIOLOGY; ECOLOGY; EMBRYOLOGY; FLORICULTURE; FOREST AND FORESTRY; GENETICS; LANDSCAPE ARCHITECTURE; PALEOBOTANY; PLANT ANATOMY; PLANT GROWTH; PLANT MORPHOGENESIS; PLANT PATHOLOGY; PLANT PHYSIOLOGY; PLANT TAXONOMY.

Arthur Cronquist

Botulism

An illness produced by the exotoxin of *Clostridium botulinum* and occasionally other clostridia, and characterized by paralysis and other neurological abnormalities. There are seven principal toxin types involved (A–G); only types A, B, E, and F have been implicated in human disease. Types C and D produce illness in birds and mammals. Strains of *C. barati* and *C. butyricum* have been found to produce toxins E and F and have been implicated in infant botulism. There is serologic cross-reactivity between *C. botulinum* and *C. sporogenes* and *C. novyi.* Botulinal toxin is among the most potent poisons known; it has a heavy chain (molecular weight about 100,000) and a light chain (about 50,000) joined by a disulfide bond. *See* ANAEROBIC INFECTION; VIRULENCE.

Mechanism of action. Botulinum neurotoxin consists of three domains with different functions: nonspecific binding, membrane translocation, and proteolysis for specific components of the neuroexocytosis apparatus. Botulinum neurotoxins act at the periphery by inducing a flaccid paralysis due to the inhibition of acetylcholine release at the neuromuscular junction. Serotypes B, F, and G toxins cleave a single peptide bond of the vesicle-associated membrane protein (VAMP) synaptobrevin. Types A and E cleave synaptosome-associated protein of 25 kilodaltons (SNAP-25) at different sites located within the carboxyl terminus. *See* ACETYLCHOLINE.

Forms. The three clinical forms are classic botulism, infant botulism, and wound botulism. Classic botulism is typically due to ingestion of preformed toxin, infant botulism involves ingestion of *C. botulinum* spores with subsequent germination and toxin production in the gastrointestinal tract, and wound botulism involves production of toxin by the organism's infecting or colonizing a wound. The incubation period is from a few hours to more than a week (but usually 1–2 days), depending primarily on the amount of toxin ingested or absorbed.

Symptoms and diagnosis. There is classically acute onset of bilateral cranial nerve impairment and subsequent symmetrical descending paralysis or weakness. Commonly noted are dysphagia (difficulty in swallowing), dry mouth, diplopia (double vision), dysarthria (a neuromuscular disorder affecting speech), and blurred vision. Nausea, vomiting, and fatigue are common as well. Ileus (impaired intestinal motility) and constipation are much more typical than diarrhea; there may also be urinary retention and dry mucous membranes. Central nervous system function and sensation remain intact, and fever does not occur in the absence of complications. Fever may even be absent in wound botulism. The diagnosis should be made clinically and confirmed by demonstrating the toxin in serum, stool, or epidemiologically implicated foods, or by recovery of the organism from the stool or wound of the patient. Toxin is

identified by mouse toxicity and neutralization tests with type-specific antitoxin. *See* TOXIN.

Epidemiology. In food-borne botulism, home-canned or home-processed foods (particularly vegetables) are commonly implicated, with commercially canned foods involved infrequently. Outbreaks usually involve only one or two people, but may affect dozens. In infant botulism, honey and corn syrup have been implicated as vehicles. Therapy involves measures to rid the body of unabsorbed toxin, neutralization of unfixed toxin by antitoxin, and adequate intensive care support. *See* FOOD POISONING; MEDICAL BACTERIOLOGY; POISON.

Pharmacological agent. Ironically, botulinum toxin has been used therapeutically in a variety of situations where there is excessive contraction of local muscles: strabismus (squint), blepharospasm (spasm of the eyelid), hemifacial spasm, cervical dystonia, torticollis (a contraction of muscle causing the head to be drawn toward the contracted side), spasmodic dysphonia, spastic hypertonia of calf muscles, focal dystonia (for example, writer's cramp, musician's cramp), essential palatal tremor, detrusor-sphincter dyssynergia, achalasia (failure of the lower esophageal sphincter to relax), nonrelaxing puborectalis syndrome, facial wrinkles, some types of headache, and chronic anal fissure. Botulinum toxin can also block activation of sweat glands which are cholinergically innervated; it has been used for patients with severe hyperhidrosis (usually palmar or axillary). *See* MUSCLE. Sydney M. Finegold

Bibliography. Centers for Disease Control, *Botulism in the United States, 1899–1977*, 1979; M. Hallett, One man's poison: Clinical applications of botulinum toxin, *New Engl. J. Med.*, 341:118–120, 1999; R. Pellizzari et al., Tetanus and botulinum neurotoxins: Mechanism of action and therapeutic uses, *Phil. Trans. Roy. Soc. Lond. B Biol. Sci.*, 354:259–268, 1999; R. L. Shapiro, C. Hatheway, and D. L. Swerdlow, Botulism in the United States: A clinical and epidemiologic review, *Ann. Intern. Med.*, 129:221–228, 1998.

Boundary-layer flow

That portion of a fluid flow, near a solid surface, where shear stresses are significant and the inviscid-flow assumption may not be used. All solid surfaces interact with a viscous fluid flow because of the no-slip condition, a physical requirement that the fluid and solid have equal velocities at their interface. Thus a fluid flow is retarded by a fixed solid surface, and a finite, slow-moving boundary layer is formed.

A requirement for the boundary layer to be thin is that the Reynolds number, $\mathrm{Re} = \rho UL/\mu$, of the body be large, 10^3 or more, where ρ and μ are the fluid density and viscosity, respectively, U is the stream velocity, and L is the body length. Under these conditions, as first pointed out by L. Prandtl in 1904, the flow outside the boundary layer is essentially inviscid and plays the role of a driving mechanism for the layer. *See* REYNOLDS NUMBER.

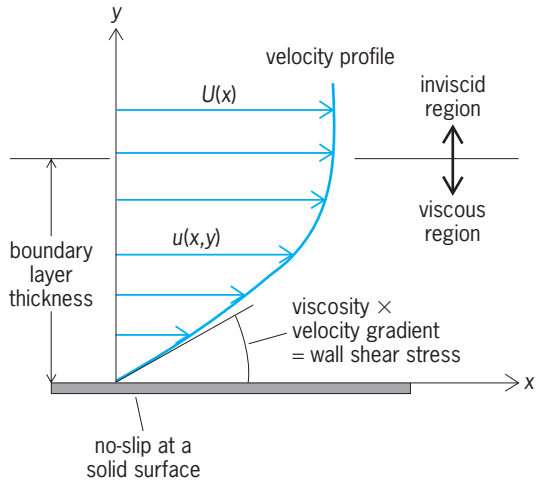

Fig. 1. Typical laminar boundary-layer velocity profile.

A typical low-speed or laminar boundary layer is shown in **Fig. 1**. Such a display of the streamwise flow vector variation near the wall is called a velocity profile. The no-slip condition requires that $u(x,0) = 0$, as shown. The velocity rises monotonically with distance y from the wall, finally merging smoothly with the outer (inviscid) stream velocity $U(x)$. At any point in the boundary layer, the fluid shear stress τ is related to the local velocity gradient by Eq. (1),

$$\tau = \mu(\partial u/\partial y) \tag{1}$$

assuming a newtonian fluid. The value of the shear stress at the wall, τ_W, is most important, since it relates not only to the drag of the body but often also to its heat transfer. At the edge of the boundary layer, τ approaches zero asymptotically. There is no exact spot where $\tau = 0$; therefore the thickness δ of a boundary layer is usually defined arbitrarily as the point where $u = 0.99U$. *See* LAMINAR FLOW.

Boundary-layer approximations. A boundary layer is mathematically different from a general viscous flow. In a steady, two-dimensional, incompressible flow (Fig. 1), the three unknowns would be pressure $p(x,y)$, streamwise velocity $u(x,y)$, and velocity normal to the wall, $\upsilon(x, y)$. These would be solved for simultaneously from the conservation of mass and the x- and y-momentum or Navier-Stokes equations. The solution would apply throughout the entire flow field.

If, however, the viscous region is thin, the boundary-layer approximations apply in the form of Eqs. (2). Many terms may then be dropped from

$$\begin{matrix} v & \partial u/\partial x u/\partial y \\ \partial v/\partial x v/\partial y & \partial p/\partial y \cong 0 \end{matrix} \tag{2}$$

the basic Navier-Stokes equations, and the resulting boundary-layer equations eliminate pressure as a variable and solve merely for u and υ by extending downstream between the wall, where no-slip obtains, and the inviscid outer stream, where $U(x)$ is assumed directly related to the pressure $P(x)$ by

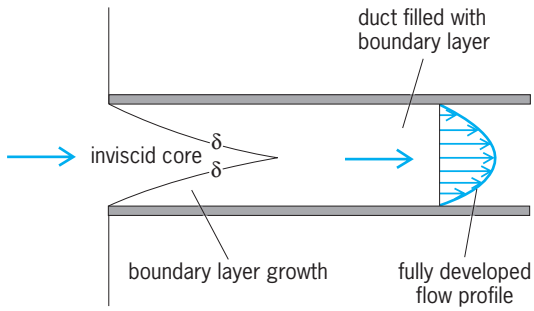

Fig. 2. Boundary-layer development in the entrance of a duct.

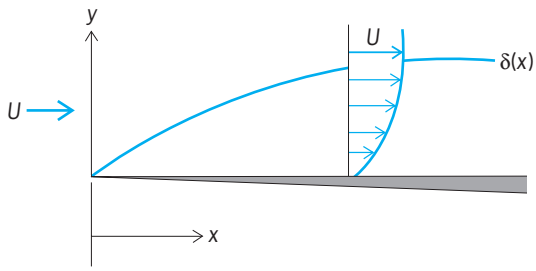

Fig. 3. Laminar boundary-layer flow past a sharp flat plate.

the Bernoulli relation (3). This is a profound math-

$$P(x) + 0.5\rho U^2(x) = \text{constant} \tag{3}$$

ematical simplification which permits designers to compute boundary-layer velocity, shear, temperature, and heat transfer without recourse to massive digital computer codes. *See* BERNOULLI'S THEOREM; NAVIER-STOKES EQUATION.

Boundary layers in duct flow. When a flow enters a duct or confined region, boundary layers immediately begin to grow on the duct walls (**Fig. 2**). An inviscid core accelerates down the duct center, but soon vanishes as the boundary layers meet and fill the duct with viscous flow. Constrained by the duct walls into a no-growth condition, the velocity profile settles into a fully developed shape which is independent of the streamwise coordinate. The pressure drops linearly downstream, balanced by the mean wall-shear stress. This is a classic and simple case of boundary-layer flow which is well documented by both theory and experiment.

Flat-plate flow. A classic incompressible boundary-layer flow is a uniform stream at velocity U, moving past a sharp flat plate parallel to the stream (**Fig. 3**). In the Reynolds number range $\rho UL/\mu = 1 \times 10^3$ to 5×10^5, the flow is laminar and orderly, with no superimposed fluctuations. The boundary-layer thickness δ grows monotonically with x, and the shape of the velocity profile is independent of x when normalized, that is, u/U is a function of y/δ. The profiles are said to be similar, and they are now called Blasius profiles, after the engineer who first analyzed them in 1908.

The Blasius flat-plate flow results in closed-form algebraic formulas for such parameters as wall-shear stress and boundary-layer thickness, given by Eqs. (4). Similar expressions can be formulated for

$$\begin{aligned} \delta &\approx 5.0(\mu x/\rho U)^{1/2} \\ \tau_W &\approx 0.332(\rho\mu U^3/x)^{1/2} \end{aligned} \tag{4}$$

temperature and heat-transfer parameters. These results are useful in estimating viscous effects in flow past thin bodies such as airfoils, turbine blades, and heat-exchanger plates.

Momentum and displacement thickness. As mentioned above, δ is defined arbitrarily as the point where $u = 0.99U$. This position is difficult to define experimentally. More definitive thickness measures are the two integral scales which arise from the conservation of mass and streamwise momentum. These are the displacement thickness δ^* and the momentum thickness θ, as defined in Eqs. (5). These two

$$\begin{aligned} \delta^* &= \int_0^\infty (1 - u/U)dy \\ \theta &= \int_0^\infty (u/U)(1 - u/U)dy \end{aligned} \tag{5}$$

terms are unambiguous and are often used to correlate data on local friction and heat transfer and on transition to turbulence of boundary layers. The displacement thickness represents the local deflection of the outer inviscid streamlines, whereas the momentum thickness is a measure of local wall-friction effects.

For flat-plate flow, if δ is given by Eq. (4), the Blasius theory predicts that $\delta^*/\delta = 0.344$ and $\theta/\delta = 0.133$.

Boundary-layer separation. The flat plate is very distinctive in that it causes no change in outer-stream velocity U. Most body shapes immersed in a stream flow, such as cylinders, airfoils, or ships, induce a variable outer stream $U(x)$ near the surface. If U increases with x, which means from Eq. (3) that pressure decreases with x, the boundary layer is said to be in a favorable gradient and remains thin and attached to the surface. If, however, velocity falls and pressure rises with x, the pressure gradient is unfavorable or adverse (**Fig. 4**). The low-velocity fluid near the wall is strongly decelerated by the rising pressure, and the wall-shear stress drops off to zero. Downstream of this zero-shear or separation point,

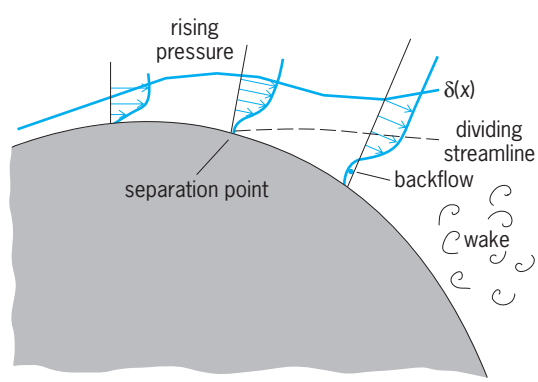

Fig. 4. Boundary-layer separation in a region of persistent rising pressure velocity.

there is backflow and the wall shear is upstream. The boundary layer thickens markedly to conserve mass, and the outer stream separates from the body, leaving a broad, low-pressure wake downstream. Flow separation may be predicted by boundary-layer theory, but the theory is not able to estimate the wake properties accurately. *See* WAKE FLOW.

In most immersed-body flows, the separation and wake occur on the rear or lee side of the body, with higher pressure and no separation on the front. The body thus experiences a large downstream pressure force called pressure drag. This happens to all blunt bodies such as spheres and cylinders and also to airfoils and turbomachinery blades if their angle of attack with respect to the oncoming stream is too large. The airfoil or blade is said to be stalled, and its performance suffers.

Instability and transition. All laminar boundary layers, if they grow thick enough and have sufficient velocity, become unstable. Slight disturbances, whether naturally occurring or imposed artificially, tend to grow in amplitude, at least in a certain frequency and wavelength range. The growth begins as a selective group of two-dimensional periodic disturbances, called Tollmien-Schlichting waves, which become three-dimensional and nonlinear downstream and eventually burst into the strong random fluctuations called turbulence. The critical parameter is the Reynolds number, usually based on boundary-layer displacement thickness, $\mathrm{Re}_{\delta^*} = \rho U \delta^* / \mu$. The process of change from laminar to turbulent flow is called transition. *See* TURBULENT FLOW.

In **Fig. 5**, the classic flat-plate flow is shown as an example. The leading-edge flow is the laminar, Blasius profile from Fig. 3. The point of first instability is at $\mathrm{Re}_{\delta *} \approx 520$ or $\mathrm{Re}_x \equiv \rho U x / \mu \approx 10^5$. Tollmien-Schlichting waves are formed with wavelengths of about six boundary-layer thicknesses and frequencies of approximately $0.1U/\delta$. These waves grow, are joined by other frequencies and wavelengths, become three-dimensional, and burst into turbulence further downstream, typically at $\mathrm{Re}_{x,\mathrm{tr}} \approx 5 \times 10^5$. The flow downstream of this transition point is fully turbulent and remains so in almost all cases. For very smooth flows, $\mathrm{Re}_{x,\mathrm{tr}}$ can be as high as 5×10^6.

The transition process is affected by many parameters. The transition is delayed by favorable pressure gradients, smooth walls, and a quiet outer stream. It is hastened by adverse pressure gradients, rough walls, and a noisy freestream. The transition can be triggered abruptly by inserting an unsteady, wake-producing tripping device, such as a wire or large particle, into the boundary layer near the instability point.

The prediction of the point of first instability for various boundary-layer flows is well developed and experimentally verified. The prediction of final transition, however, is quantitatively less successful because of the many important competing parameters affecting the process.

In rare cases, a fully turbulent flow can be made to relaminarize by subjecting the boundary layer to extreme deceleration.

Turbulent boundary layers. The turbulent flow regime is characterized by random, three-dimensional fluctuations superimposed upon time-mean fluid properties, including velocity, pressure, and temperature. The fluctuations are typically 3–6% of the mean values and range in size over three orders of magnitude, from microscale movements to large eddies of size comparable to the boundary-layer thickness. They are readily measured by modern instruments such as hot wires and laser-Doppler velocimeters.

The effect of superimposing a wide spectrum of eddies on a viscous flow is to greatly increase mixing and transport of mass, momentum, and heat across the flow. Turbulent boundary layers are thicker than laminar layers and have higher heat transfer and friction. The turbulent mean-velocity profile is rather flat, with a steep gradient at the wall (Fig. 5). The edge of the boundary layer is a ragged, fluctuating interface, as shown in Fig. 5, which separates the nonturbulent outer flow from large turbulent eddies in the layer. The thickness of such a layer is defined only in the time mean, and a probe placed in the outer half of the layer would show intermittently turbulent and nonturbulent flow.

No theory for turbulent flow is based on first principles, but many successful empirical models have

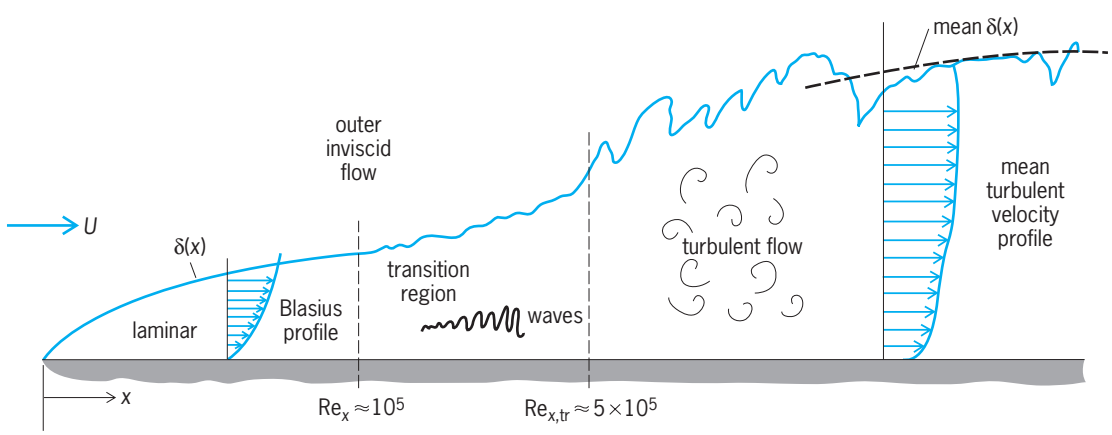

Fig. 5. Transition process in viscous flow past a sharp flat plate. The boundary-layer thickness is exaggerated. (*After F. M. White, Heat and Mass Transfer, Addison-Wesley, 1988*)

been proposed. A turbulent boundary layer may be broken into three regions: a narrow sublayer near the wall, of approximate thickness $5\mu/(\rho\tau_w)^{1/2}$, too small to be seen in Fig. 5, where eddies are damped out and molecular viscosity dominates; an intermediate layer, approximately logarithmic in shape, extending to about one-fourth of the layer thickness; and an outer wake layer, dependent on stream parameters, which rises from the log layer to merge smoothly with the stream velocity U.

The intermediate or logarithmic layer is dependent entirely on wall-related parameters and is approximated by Eq. (6). In flows with constant or

$$u/v^* \approx 2.44\log(\rho v^* y/\mu) + 5.0 \qquad (6)$$
$$v^* = (\tau_w/\rho)^{1/2}$$

falling stream pressure, the wake layer is small, and Eq. (6) can be employed to develop useful formulas for predicting turbulent boundary-layer friction and heat transfer. For example, formulas (7) result for flat-

$$\delta \approx 0.16(\mu x^6/\rho U)^{1/7}$$
$$\tau_W \approx 0.0135(\mu\rho^6 U^{13}/x)^{1/7} \qquad (7)$$

plate turbulent boundary-layer thickness and wall friction. They may be compared with their laminar-flow counterparts in Eqs. (4). For a given x, the turbulent layer is thicker and has a wall-shear stress several times larger; the wall-heat transfer is also higher.

The reason the turbulent-velocity profile in Fig. 5 is so flat is that its effective viscosity is very large. By analogy with the laminar shear-stress formula (1), an equivalent turbulent eddy viscosity may be defined to compute shear stress in a turbulent layer, Eq. (8). The eddy viscosity is much larger than the

$$\tau_{\text{turb}} \approx \mu_{\text{turb}}(\partial u/\partial y) \qquad (8)$$

molecular viscosity μ and is not a physical property, but depends upon flow parameters such as stream velocity and wall-shear stress. For example, the logarithmic layer of Eq. (6) may be used to infer the eddy viscosity formula of Eq. (9). Many digital computer

$$\mu_{\text{turb}} \approx 0.41\rho v^* y \qquad (9)$$

codes for predicting turbulent boundary layers use eddy-viscosity correlations to complete the equations of motion. Similarly, in turbulent heat-transfer analysis, an eddy thermal conductivity is often used.

Compressible boundary layers. As the stream velocity U becomes larger, its kinetic energy, $U^2/2$, becomes comparable to stream enthalpy, c_pT, where c_p is the specific heat at constant pressure and T is the absolute temperature. Changes in temperature and density begin to be important, and the flow can no longer be considered incompressible. An equivalent statement is that the stream Mach number Ma_∞, defined by Eq. (10), where a is the speed of

$$\text{Ma}_\infty = U/a \qquad (10)$$

sound of the stream, becomes significant when Ma_∞ is greater than 0.5. Liquids flow at very small Mach numbers, and compressible flows are primarily gas flows. For an ideal gas of absolute temperature T, the speed of sound is given by Eq. (11), where R is the gas constant, and γ, defined by Eq. (12) where c_p

$$a = \gamma RT^{1/2} \qquad (11)$$

$$\gamma = c_p/c_v \qquad (12)$$

is the specific heat at constant pressure, is approximately 1.4 for air. If Ma_∞ is greater than 1, the flow is said to be supersonic. *See* GAS; GAS DYNAMICS; MACH NUMBER.

In a flow with supersonic stream velocity, the noslip condition is still valid, and much of the boundary-layer flow near the wall is at low speed or subsonic. The fluid enters the boundary layer and loses much of its kinetic energy, of which a small part is conducted away although most is converted into thermal energy. Thus the near-wall region of a highly compressible boundary layer is very hot, even if the wall is cold and is drawing heat away. This situation is illustrated in **Fig. 6**, comparing an incompressible and a compressible ($\text{Ma}_\infty = 2$) thermal boundary layer. The basic difference between low and high speed is the conversion of kinetic energy into higher temperatures across the entire boundary layer.

In a low-speed (incompressible) boundary layer, a cold wall simply means that the wall temperature T_w is less than the free-stream temperature T_∞. The heat flow q_w is from high toward lower temperature, that is, into the wall (Fig. 6). For a low-speed insulated

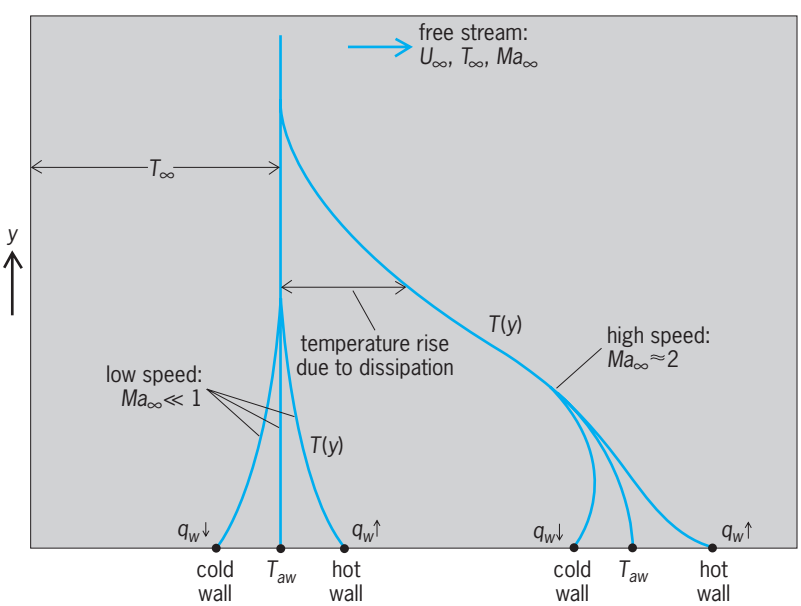

Fig. 6. Boundary-layer temperature profiles for incompressible and compressible (supersonic) stream velocities. (*After F. M. White, Heat and Mass Transfer, Addison-Wesley, 1988*)

wall, the boundary-layer temperature is uniform. For a high-speed flow, however, an insulated wall has a high surface temperature because of the viscous dissipation energy exchange in the layer. This adiabatic wall temperature T_{aw} may be estimated from Eq. (13). The parameter r is the recovery factor, in-

$$T_{aw} \approx T_\infty \left[1 + 0.5r\left(\gamma - 1\right)\mathrm{Ma}_\infty^2\right] \qquad (13)$$

dicating the efficiency of energy exchange; it varies from 0.85 to 0.89 for compressible air flows. For example, if $\mathrm{Ma}_\infty = 2.0$, as in Fig. 6, then $T_{aw} \approx 1.7T_\infty$. Equation (13) requires that temperatures be in absolute units. In such a flow, a cold wall means only that $T_w < T_{aw}$; the wall temperature itself may be considerably higher than the stream temperature, yet the fluid temperature drops off at the wall, into which heat flows.

Except for the added complexity of having to consider fluid pressure, temperature, and density as coupled variables, compressible boundary layers have similar characteristics to their low-speed counterparts. They undergo transition from laminar to turbulent flow (Fig. 5) but typically at somewhat higher Reynolds numbers. Compressible layers tend to be somewhat thicker than incompressible boundary layers, with proportionally smaller wall-shear stresses. They tend to resist flow separation (Fig. 4) slightly better than incompressible flows.

In a supersonic outer stream, shock waves can always occur. Shocks may form in the boundary layer because of obstacles in the layer or downstream, or they may be formed elsewhere and impinge upon a boundary. In either case, the pressure rises sharply behind the shock, an adverse gradient, and this tends to cause early transition to turbulence and early flow separation. Special care must be taken to design aerodynamic surfaces to accommodate or avoid shock-wave formation in transonic and supersonic flows. *See* COMPRESSIBLE FLOW; SHOCK WAVE.

Boundary-layer control. As boundary layers move downstream, they tend to grow naturally and undergo transition to turbulence. Boundary layers encountering rising pressure undergo flow separation. Both phenomena can be controlled at least partially. Airfoils and hydrofoils can be shaped to delay adverse pressure gradients and thus move separation downstream. Proper shaping can also delay transition. Wall suction removes the low-momentum fluid and delays both transition and separation. Wall blowing into the boundary layer, from downward-facing slots, delays separation but not transition. Changing the wall temperature to hotter for liquids and colder for gases delays transition. Practical systems have been designed for boundary-layer control, but they are often expensive and mechanically complex. *See* AIRFOIL; FLUID-FLOW PRINCIPLES; STREAMLINING; VISCOSITY.

Frank M. White

Bibliography. *Illustrated Experiments in Fluid Mechanics*, NCFMF Book of Film Notes, 1972; J. A. Schetz, *Boundary Layer Analysis*, 1997; H. Schlichting, *Boundary-Layer Theory*, 8th rev. ed., 1999; F. S. Sherman, *Viscous Flow*, 1990; M. Van Dyke, *An Album of Fluid Motion*, 1982; F. M. White, *Viscous Fluid Flow*, 2d ed., 1991.

Bovine viral diarrhea

An economically important, multisystemic viral disease of cattle that is common worldwide. The causal agent is bovine viral diarrhea virus (BVDV). Cattle of all ages may be infected with this virus, and a variety of acute or chronic disease processes occur that may affect the digestive, respiratory, and reproductive systems.

The virus. Bovine viral diarrhea virus is a member of the *Pestivirus* genus within the Flaviviridae family. Pestiviruses are an antigenically related group of viruses that includes classical swine fever virus, border disease virus of sheep, and bovine viral diarrhea virus. Mature viral particles are spherical with a diameter of 40 to 60 nm. A lipid envelope surrounds a 30-nm inner shell (capsid) that contains the viral genome. The viral genome is a single, positive-stranded ribonucleic acid (RNA) that is about 12,300 nucleotides in length. The viral RNA encodes four structural and seven nonstructural proteins. The virus replicates in the cytoplasm of the host cell, eventually forming new viral particles in vesicles within the cytoplasm. The new viral particles are transported to the cell surface and released to the environment. *See* CLASSICAL SWINE FEVER; VIRUS.

Analysis of the RNA from numerous viruses collected worldwide has resulted in the identification of two genetic groups, or genotypes, of bovine viral diarrhea virus. The viral genotypes are termed type 1 and type 2. Each genotype is further divided into two or more subgenotypes such as type 1a, type 1b, etc. The subgenotypes contain viruses that are genetically most closely related. Although bovine viral diarrhea viruses are divided into two genotypes based on their nucleic acid sequence, all of the viruses are related antigenically to each other. Members of either viral genotype cause similar disease processes. In addition to two viral genotypes, there are two biotypes of bovine viral diarrhea virus that are differentiated based on their effect on infected cell cultures. Both biotypes are found in each viral genotype. The viral biotypes are termed noncytopathic or cytopathic. The noncytopathic biotype is predominant in nature and grows in cell cultures without inducing overt harmful effects. The cytopathic biotype induces morphologic changes in infected cell cultures that include formation of cytoplasmic vacuoles of variable size and number. Soon after the appearance of cytoplasmic vacuoles, the infected cell dies. The harmful effects of cytopathic bovine viral diarrhea virus in cell culture do not equate with virulence of the virus in the infected host, as the most virulent viruses for cattle are noncytopathic in cell culture. In addition to cattle, bovine viral diarrhea virus infects most even-toed ungulates, including sheep, goat, swine, deer, bison, llama, and antelope. *See* ANIMAL VIRUS.

Clinical disease. Cattle usually are infected with bovine viral diarrhea virus in the first 2 to 3 years of life. Often, infection with the virus leads to a clinically inapparent or clinically mild disease. Signs of mild disease include low-grade fever, loose stool, low white blood cell count, mild depression, transient loss of appetite, and decreased milk production. Signs of mild disease last 1 to 3 days, and recovery from disease is rapid and complete. Some viruses cause clinically severe disease that is characterized by high fever, diarrhea, very low white blood cell count, depression, dehydration, and rapid respiration. Ulceration of the mouth, esophagus, and intestine may occur. Some cattle with clinically severe disease may experience a precipitous loss of platelets from the blood followed by mild to extensive hemorrhaging. The duration of clinical signs associated with severe disease may be 4 to 7 days, and the outcome may be death. Typical outbreaks of bovine viral diarrhea are associated with high morbidity and low mortality; however, clinically severe disease may have high mortality.

Bovine viral diarrhea virus replicates in lymphoid cells and may suppress normal function of those cells. By disrupting this function, the virus compromises the ability of the infected animal to defend itself against pathogens. Thus, bovine viral diarrhea virus is often associated with respiratory and enteric disease of cattle in which multiple other agents contribute to the disease process. Bovine viral diarrhea virus readily crosses the placenta to infect the fetus and cause reproductive failure. Embryonic resorption, abortion, stillbirth, premature birth, and congenital defects of the central nervous system, eye, and musculoskeletal system occur following fetal infection. Life-long persistent viral infection may be a sequel of fetal infection with noncytopathic, but not cytopathic, virus. For persistent infection to be established, fetal infection must occur in the first 4 months of pregnancy. At birth, persistently infected calves may be weak or stunted. However, persistently infected calves also may appear healthy and vigorous. Persistently infected cattle do not mount an immune response against the virus they carry, but are capable of producing an immune response against other infectious agents.

Acute and chronic mucosal disease are rare forms of bovine viral diarrhea that occur in cattle born with a persistent infection with noncytopathic virus. The signs of acute mucosal disease are fever, low white blood cell count, profuse diarrhea that is often bloody, ulcerations throughout the digestive tract, inappetence, dehydration, and death. The course of mucosal disease is one to a few days and recovery is rare. Persistently infected cattle also may develop a chronic from of mucosal disease that is characterized by intermittent diarrhea, lameness, wasting, and death. In the chronic form, death may not occur until several months after the initial signs of disease. Both the acute and chronic forms of mucosal disease occur when persistently infected cattle become superinfected with a cytopathic bovine viral diarrhea virus. Often the cytopathic virus arises from the noncytopathic persistent virus following a mutational event in the replication of the viral RNA. Cytopathic virus may be transmitted to persistently infected cattle from other persistently infected cattle that have mucosal disease. Also, cytopathic virus may be inoculated into a persistently infected animal through use of vaccine that contains live cytopathic virus. Mucosal disease is rare because it requires a persistently infected animal and persistent infection is relatively rare.

Pathology. Postmortem examination of cattle that have acute bovine viral diarrhea or acute mucosal disease often reveals ulcerative lesions of the mucosal surface (inner lining) of the digestive tract. Erosions and ulcerations in the mouth, esophagus, abomasum (true stomach of a ruminant), and small and large intestines are common lesions. In the intestines, ulcerations frequently are centered over submucosal lymphoid follicles. In chronic mucosal disease, there may be hair loss and thickening of the skin, usually in the neck area. Also, ulcers may be found in the skin between the claws of the hoof, around the dewclaws, or on the heel. Skin lesions often fail to heal, leading to secondary bacterial infections of the skin and occasionally hoof deformities. Some bovine viral diarrhea viruses induce thrombocytopenia (a decrease in platelets in the blood) during primary acute infection. This may lead to hemorrhages in the gums, in the sclera of the eye, and on serosal surfaces (outer surfaces) of the internal organs. The primary microscopic lesions for all forms of bovine viral diarrhea are depletion of lymphoid cells in the lymphoid follicles of the intestines. Viral antigen may be detected in lymphoid, neuronal, endocrine, skin, and bone marrow cells.

Mode of transmission. Oral or nasal exposure with virus is the primary mode of transmission of acute disease. Persistently infected cattle are efficient transmitters of virus, which is continually shed in saliva and nasal secretions. Susceptible cattle may become infected within 1 hour of contact with a persistently infected animal. Cattle afflicted with acute disease also shed virus in saliva and nasal secretions. Biting insects, veterinary instruments, or clothing contaminated with secretions or blood from an infected animal, and artificial insemination are reported to transmit disease. Because wild ruminants may be infected with bovine viral diarrhea virus, it is possible that disease may spread among farms with the movement of wildlife.

Diagnosis, treatment, and control. Clinical signs of disease and identification of mucosal ulcerations on postmortem examination support a tentative diagnosis of bovine viral diarrhea. Virus isolation in cell culture or polymerase chain reaction (PCR) using clinical specimens such as blood or nasal swabs confirm the diagnosis of acute or persistent infection with bovine viral diarrhea virus. Because most isolates of bovine viral diarrhea virus are noncytopathic in cell culture, virus isolation must be confirmed by immunohistochemistry or PCR. Antigen-capture enzyme-linked immunosorbent assays (ELISA) have proven useful for rapid detection of

persistent infection using blood or skin biopsy specimens. Immunohistochemistry using monoclonal antibody specific for the virus allows rapid detection of viral antigen in frozen or fixed tissues. Serologic tests, such as viral neutralization or ELISA, that detect antibody against virus are useful for detection of past infection or for monitoring response to vaccination.

Bovine viral diarrhea is treated with supportive care. Disease is prevented or diminished in severity through vaccination. Inactivated virus and modified-live virus vaccines for bovine viral diarrhea are available for use in cattle. Typically, vaccination is first done when calves are a few months old. Vaccination is often repeated at regular intervals during the life of the animal. Identification and elimination of persistently infected cattle is an important part of any control strategy for prevention of viral spread and disease. Temporary isolation of new additions to the herd and other farm biosecurity measures, including control of movement of equipment, animals, and personnel on the farm, also are important for control of bovine viral diarrhea. *See* VACCINATION.

Steven R. Bolin

Bibliography. S. R. Bolin, Bovine viral diarrhea and mucosal disease complex, *The Merck Veterinary Manual*, 9th ed., Merck, Whitehouse Station, NJ, 2005; K. V. Brock (ed.), Bovine viral diarrhea virus: Persistence is the key, *Vet. Clin. N. Amer.: Food Anim. Prac.*, vol. 20, no. 1, 2004; D. L. Grooms, J. C. Baker, and T. R. Ames, Diseases caused by bovine viral diarrhea virus, *Large Animal Internal Medicine: Disease of Horses, Cattle, Sheep, and Goats*, 3d ed., Mosby, St. Louis, 2002; D. M. Knipe and P. M. Howley (eds.), *Fundamental Virology*, 4th ed., Lippincott, Williams & Wilkins, Philadelphia, 2001; O. M. Radostits et al., *Veterinary Medicine: A Textbook of the Diseases of Cattle, Sheep, Pigs, Goats, and Horses*, 9th ed., W. B. Saunders, London, 2000.

Boyle's law

A law of gases which states that at constant temperature the volume of a gas varies inversely with its pressure. This law, formulated by Robert Boyle (1627–1691), can also be stated thus: The product of the volume of a gas times the pressure exerted on it is a constant at a fixed temperature. The relation is approximately true for most gases, but is not followed at high pressures. The phenomenon was discovered independently by Edme Mariotte about 1650 and is known in Europe as Mariotte's law. *See* GAS; KINETIC THEORY OF MATTER.

Frank H. Rockett

Brachiopoda

A phylum of solitary, exclusively marine, coelomate, bivalved animals, with both valves symmetrical about a median longitudinal plane. Brachiopods are typically attached to the substrate by a posteriorly located cuticle-covered stalk called a pedicle. Anteriorly, a relatively large mantle cavity is always developed between the valves, and the ciliated tentacular (filamentous) feeding organ, or lophophore, is suspended within, projecting anteriorly from the anterior body wall (**Fig. 1**). The name Brachiopoda means "arm foot" and refers to the morphology and location of the lophophore.

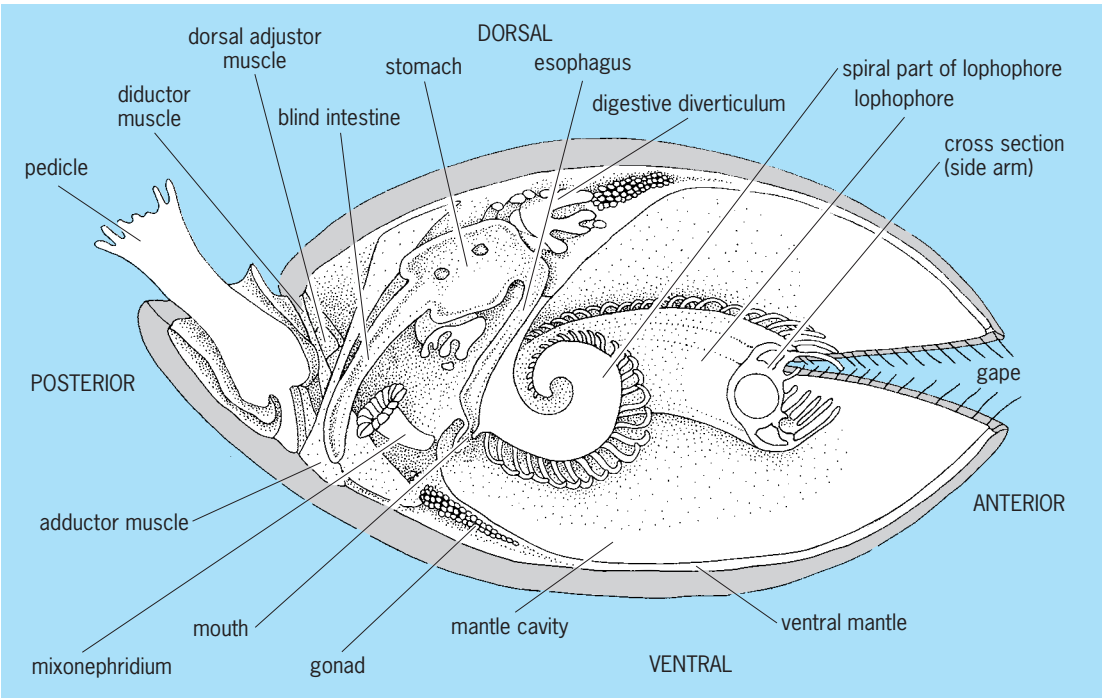

Fig. 1. Principal organs of a brachiopod as typified by *Terebratulina*. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. H, Geological Society of America, Inc., and University of Kansas Press, 1965*)

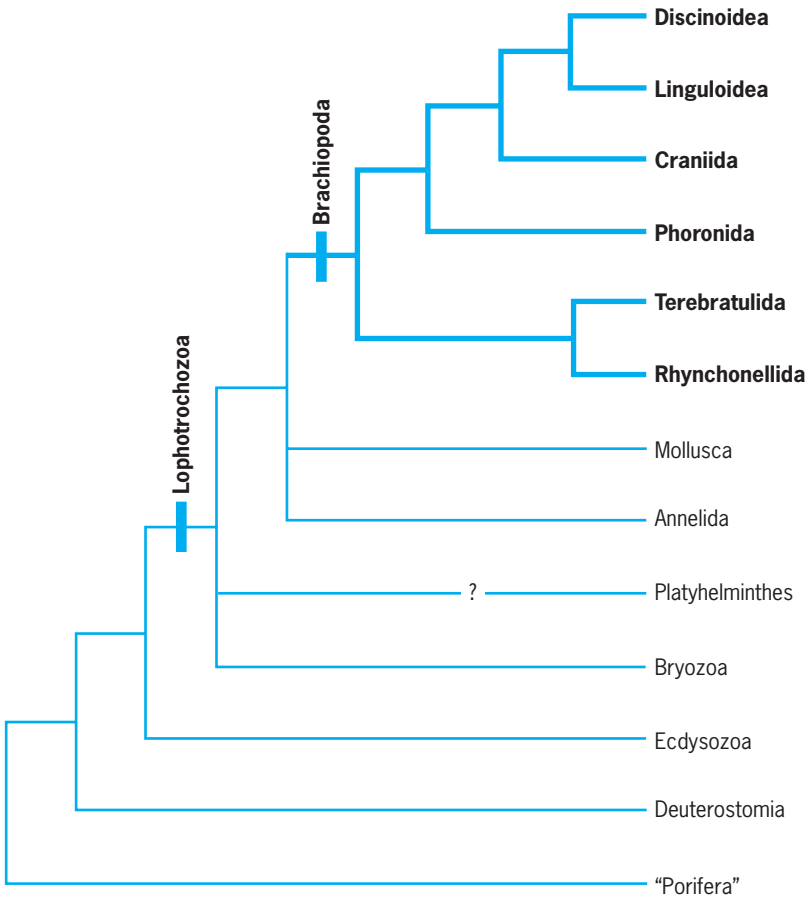

Fig. 2. Phylogenetic relationships among selected metazoans. (*Simplified from K. M. Halanych and Y. Passamaneck, Amer. Zool., 41:629–639, 2001; K. M. Halanych, Annu. Rev. Ecol. Evol. Sys., 35:229–256, 2004; and B. L. Cohen and A. Weydmann, Organisms, Diversity and Evolution, 5(4):253–273, 2005*)

This brief description serves to differentiate a brachiopod from any other animal. Within the Brachiopoda, three groups currently regarded as subphyla are recognized, each named after their most ancient living representative: Linguliformea, Craniiformea, Rhynchonelliformea. Phoroniformea are recognized by some as a fourth subphylum. The two classes (Inarticulata and Articulata), formerly distinguished by the presence or absence of articulation between the two valves of the shell, are no longer recognized as distinct taxa; rather, they appear to represent the ends of a spectrum of articulatory types, some well delineated, others less so. Rhynchonelliforms have impunctate, endopunctate, or pseudopunctate calcitic shells with a fibrous (or laminar) shell structure; most are articulated, with the valves typically hinged together by a pair of ventral teeth with complementary sockets in the dorsal valve; most have pedicles with a core of connective tissue. Craniiforms have punctate calcitic shells with a laminar (or tabular) shell structure; all lack articulation, with valves that are held together only by the soft tissue of the living animal, and all lack pedicles. Linguliforms have canaliculate organophosphatic shells with a stratiform shell structure; all lack articulation; and all have pedicles with a coelomate core. *See* INARTICULATA; LINGULIDA; RHYNCHONELLIFORMEA.

Phylogeny and classification. Molecular sequence data from nuclear and mitochondrial genes in extant brachiopods provide strong support for brachiopods as protostomous organisms, and not deuterostomes, as some earlier morphological and developmental data seemed to suggest. These data also indicate that the Phoronida is nested within the Brachiopoda, and is not the brachiopod sister group, as previously thought. Lophophorates alone do not appear to form a clade; the Bryozoa are more distantly related within the Lophotrochozoa, a rather poorly resolved clade that also includes mollusks and annelids, among other phyla (**Fig. 2**). Only approximately 5% of named brachiopod genera (and most likely, species) are extant; most of our current understanding of brachiopod phylogeny results from a composite of molecular data from a small taxonomic sample, and morphological (and stratigraphic) data from a much larger sample of extinct fossil taxa. *See* BRYOZOA; PHORONIDA; RHYNCHONELLIFORMEA.

The phylum is currently classified as follows:

Subphylum Linguliformea
- Class Lingulata
  - Order: Lingulida
    - Siphonotretida
    - Acrotretida
- Class Paterinata
  - Order Paterinida

Subphylum Craniiformea
- Class Craniata
  - Order: Craniida
    - Craniopsida
    - Trimerellida

Subphylum Phoroniformea (?)

Subphylum Rhynchonelliformea
- Class Chileata
  - Order: Chileida
    - Dictyonellida
- Class Obolellata
  - Order: Obolellida
    - Naukatida
- Class Kutorginata
  - Order: Kutorginida
- Class Strophomenata
  - Order: Strophomenida
    - Billingsellida
    - Orthotetida
    - Productida
- Class Rhynchonellata
  - Order: Protorthida
    - Orthida
    - Pentamerida
    - Rhynchonellida
    - Atrypida
    - Spiriferida
    - Spiriferinida
    - Thecideida
    - Athyridida
    - Terebratulida

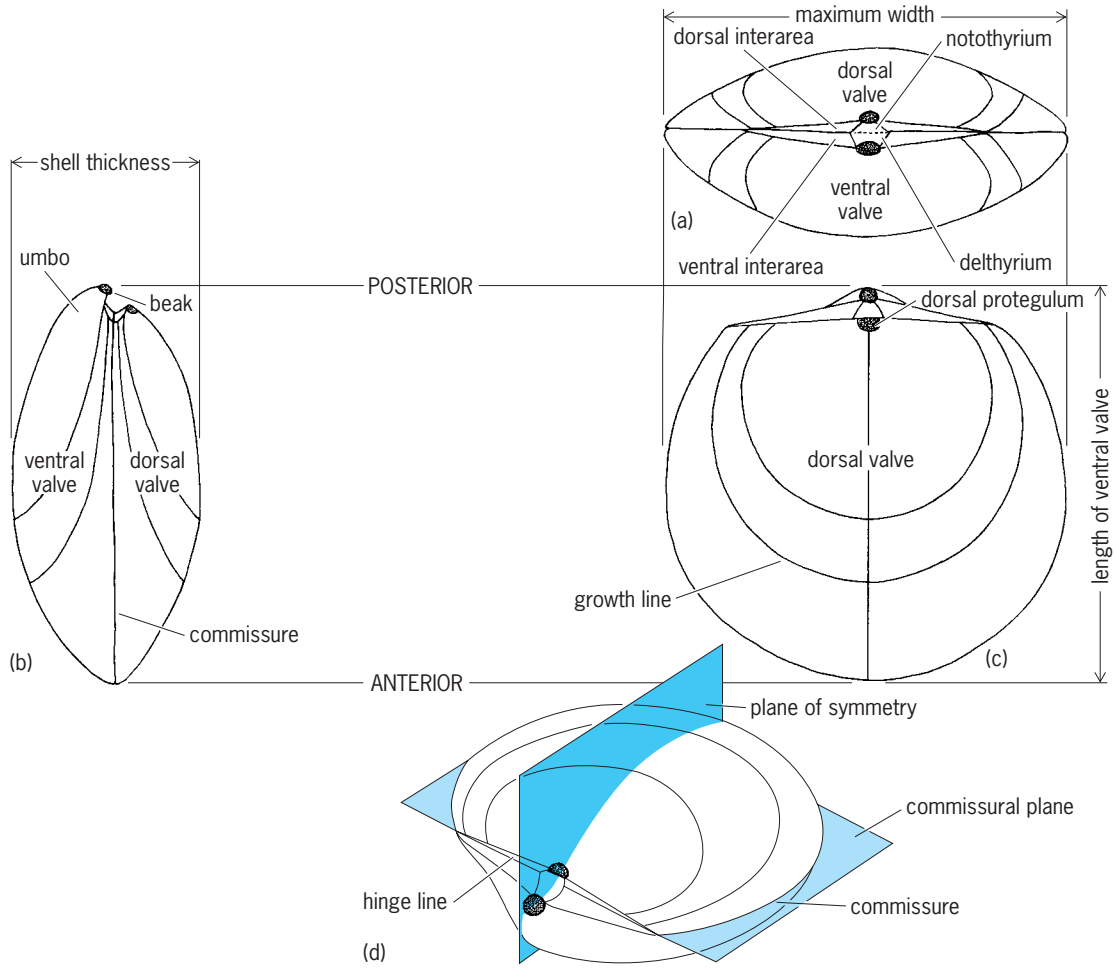

Fig. 3. Diagrammatic representation of the external features of a generalized brachiopod seen in (*a*) posterior, (*b*) left lateral, (*c*) dorsal, and (*d*) dorsolateral views. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. H, Geological Society of America, Inc., and University of Kansas Press, 1965*)

Orientation. The two brachiopod valves are currently referred to as the dorsal and ventral valves, not brachial and pedicle, because not all brachiopods possess pedicles. This body orientation distinguishes them from bivalved mollusks, which have valves on the left and right sides of the body. While topologically dorsal and ventral in adults, their developmental orientation is not entirely clear; both valves of some extant craniiforms may derive from the dorsal surface of the developing, modified trochophore larva. The pedicle either protrudes between the valves (through the delthyrium and notothyrium, which are small triangular apertures serving as pedicle openings), or (**Fig. 3**) more commonly emerges from a variably modified opening (pedicle foramen) in the ventral valve, which, in articulated forms, is the valve bearing the hinge teeth. Whatever the form of the pedicle opening, it is always at the posterior end of the animal and its enclosing shells, the opposite end being regarded as anterior. Both valves have a characteristic distribution of muscles, and any skeletal support for the lophophore is invariably developed from the dorsal valve (**Fig. 4**).

Anatomy. The pedicle is the only organ protruding outside the valves, for the remainder of the

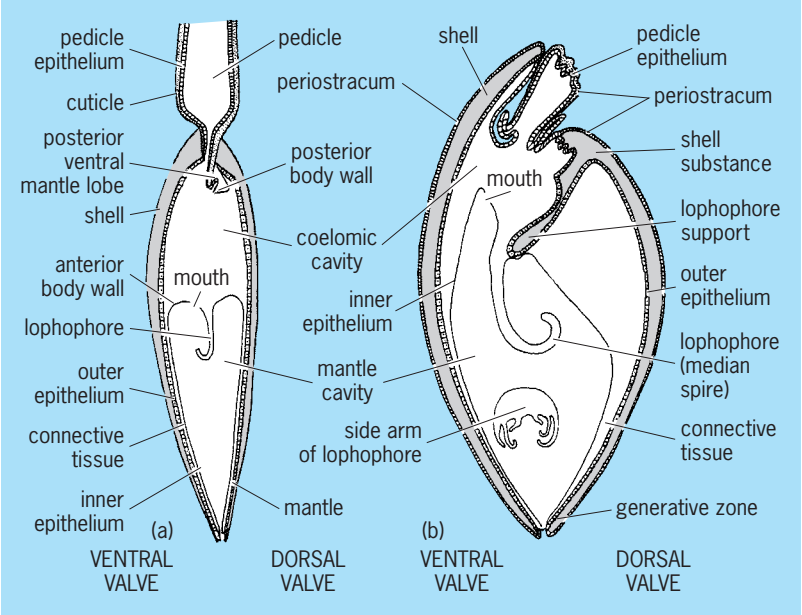

Fig. 4. Generalized representation of distribution of epithelium in relation to other tissues and organs in (*a*) lingulides and (*b*) terebratulides. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. H, Geological Society of America, Inc., and University of Kansas Press, 1965*)

animal is enclosed in the space between them. This space is divided into two unequal parts, a smaller posteriorly located body cavity and an anterior mantle cavity. The ectodermal outer epithelium underlying the shell bounding the body cavity is in a single layer; but anteriorly, laterally, and even posteriorly in most inarticulated brachiopods, it is prolonged as a pair of folds forming the ventral and dorsal mantles, from which the shells are mineralized (at the generative zones). In extant species, the two mantles approach each other and ultimately fuse along the posterior margin of rhynchonelliform brachiopods; in contrast, the mantles are invariably discrete in craniiforms and linguliforms, and are separated by a strip of body-wall inner epithelium (Fig. 4).

The body cavity contains the musculature; the alimentary canal; the mixonephridia, which are paired excretory organs also functioning as gonoducts; the reproductive organs; and the rather poorly understood circulatory and nervous systems. Except for the openings through the mixonephridia, the body cavity is enclosed, but the mantle cavity communicates freely with the sea when the valves are opened. The lophophore is a feeding and respiratory organ, typically suspended from the anterior body wall within the mantle cavity, and is always symmetrically disposed about the median plane. The lophophore consists of a ciliated, filament-bearing tube with two arms in varying configurations. The ciliary beat produces a laminar flow of water into and then out of the mantle cavity, flowing across the filaments while inside. The latter trap food particles which are carried along a groove in the lophophore to the mouth situated medially between the two arms. *See* LOPHOPHORE.

The alimentary canal of all brachiopods is broadly similar in structure, but differs in orientation in the body cavity. The mouth opens into a muscular tube, the esophagus, which continues to a stomach and intestine. In addition, there are a variable number of digestive diverticula that communicate with the stomach through narrow ducts (Fig. 1). In rhynchonelliform brachiopods, the intestine curves into a C-shape and ends blindly (with no anus) in a posteroventral location; in linguliforms, it curves into a U-shape, leaving the anus right-lateral or ventrolateral; in phoronids, it is also U-shaped, but the anus is anterodorsal; in craniiforms, the intestine is straight and does not curve or fold, leaving the anus medioposterior.

The diductor and adductor muscles that control the opening and closing of the valves are contained within the body cavity (Fig. 1), but their distribution varies among the three brachiopod subphyla. In most rhynchonelliforms, they are disposed to effect a rotation of the valves about a hinge axis located on the valves; in craniiforms and linguliforms, the hinge axis is located in the viscera between the valves; in selected linguliforms, the two valves can "scissor" past one another, allowing burrowing in a soft substrate. Adjustor muscles effect movement of the shell relative to the pedicle, and other muscles may be present that can move the lophophore slightly relative to the valves. Because of a differential rate of secretion of shell material by the epithelium at the bases of the muscles, the site of muscle attachment is commonly impressed in the valves, producing muscle scars. Rarely, the muscle scars may be elevated above the adjacent shell.

Although some small-bodied species are hermaphroditic and brood their larvae, the sexes are separate in the majority of brachiopods. The gonads, or reproductive organs, are located either within the body cavity (as in the lingulids and discinids) or more typically in slender tubelike extensions of the body cavity which project into the mantle, called the mantle canals (in craniids and rhynchonelliforms). The mantle canal pattern may be retained on the inner surface of the valves, even in fossils, by processes of differential secretion comparable with those producing the muscle scars.

Given the phylogenetic nesting of phoronids within brachiopods, the anatomy of Phoronida can be interpreted as having evolved from a brachiopod body plan; phoronid anatomy is modified largely as a reflection of the loss of the two mineralized valves. *See* PHORONIDA.

Embryology and ontogeny. Linguliforms, craniiforms, phoronids, and rhynchonelliforms share numerous aspects of embryogenesis and larval development, yet each has distinct characteristics not shared with any others (**Fig. 5**). Extant linguliforms and phoronids are planktotrophic and have likely remained so from their origin in the Cambrian (or earlier); extant craniiforms and rhynchonelliforms are lecithotrophic, and have each evolved independently from a planktotrophic ancestral state. Planktotrophic larvae remain in the water column for weeks to months, while lecithotrophic larvae remain in the water column only briefly, usually less than a week; these differences have important implications for the timing of differentiation of adult structures, and for the ability of the larvae to disperse biogeographically. Planktotrophic brachiopods develop more structures as larvae, prior to metamorphosis, and lecithotrophs develop more structures following metamorphosis, as young juveniles. Shell deposition begins at metamorphosis in all three subphyla, although mantle formation can begin during embryogenesis or, more commonly, during the larval growth period, just prior to metamorphosis.

Larvae in all three subphyla share a three-part body organization: an apical lobe, from which the lophophore differentiates in the larval stage of linguliforms and immediately following metamorphosis in craniiforms and rhynchonelliforms; a mantle lobe, from which the mantle (on which the periostracum and shells later mineralize), alimentary canal, and most of the viscera develop; and a pedicle lobe, from which the pedicle develops in linguliforms during the larval stage, and in rhynchonelliforms following settlement and metamorphosis. In rhynchonelliforms alone, the mantles are reversed after metamorphosis so that they are oriented anteriorly, partially covering what was the apical lobe of the larva, and with the original inner surface of the

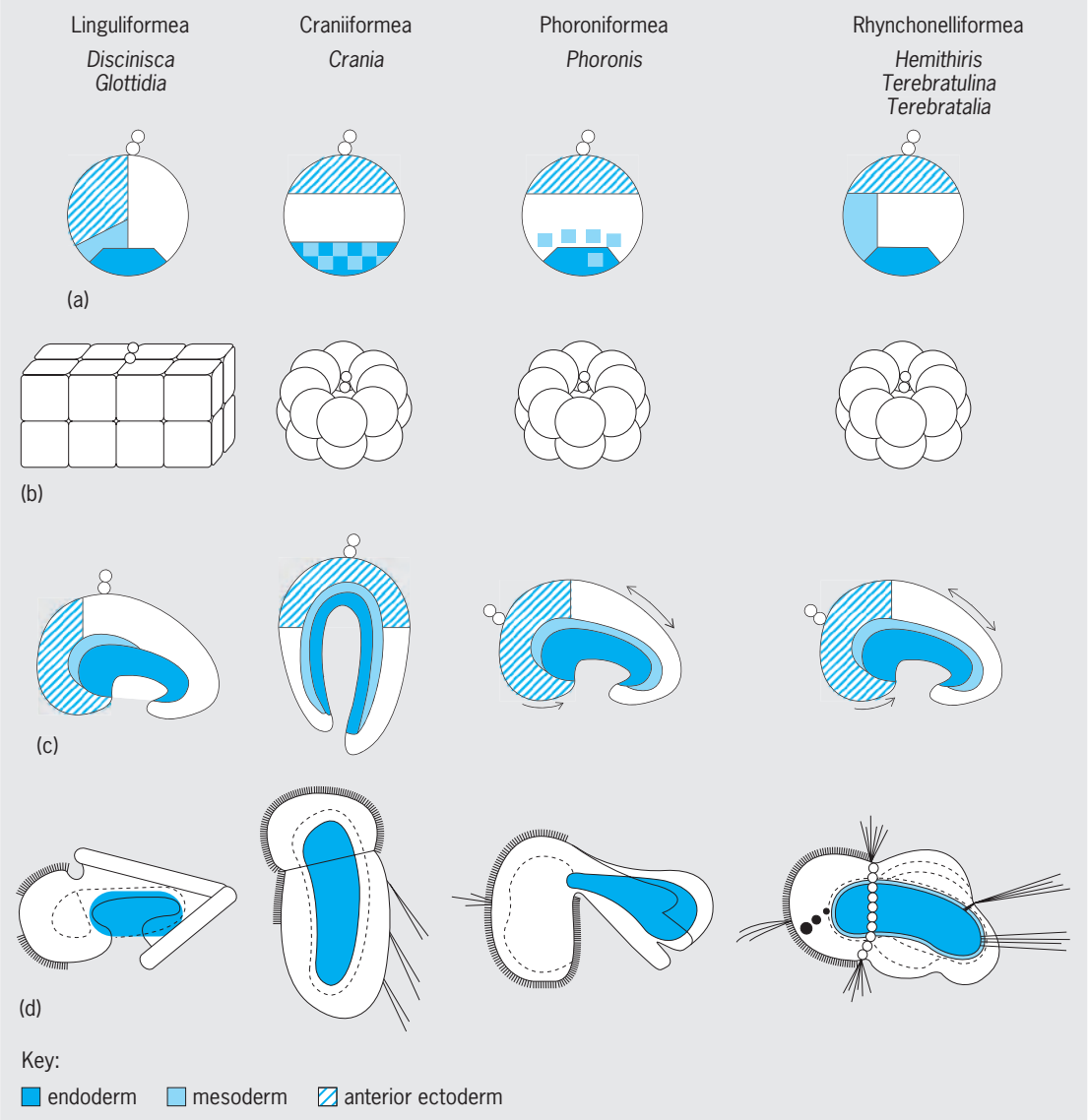

Fig. 5. Comparison of fate maps and selected developmental stages of the three brachiopod subphyla and Phoroniformea, considered by some to be a fourth brachiopod subphylum. (*a*) Fate maps of uncleaved eggs. (*b*) 16-cell embryos. (*c*) Late gastrula stage. (*d*) Lateral views of the early larva. (*Reprinted from G. Freeman, Developmental Biology, 261:263–287, 2003, with permission from Elsevier*)

mantle forming the outer surface of the organism. The mantles subsequently secrete the earliest, first-formed shell, the protegulum, which enlarges by terminal accretion during later shell growth. Only after metamorphosis does the rudimentary lophophore, alimentary canal, and adult musculature develop. Rhynchonelliforms and phoronids share many aspects of embryogenesis (Fig. 5*a–c*), providing further support for genetic data suggesting they evolved among brachiopods. Craniiforms do not develop pedicles or a larval pedicle lobe, but the larvae settle at their posteriormost ends, comparable to where a pedicle lobe would be located. Craniiforms do not undergo mantle reversal.

Linguliform larvae are much like miniature adults that undergo minimal morphological change at metamorphosis, having developed most of their adult features while larvae in the plankton. The two mantle rudiments are separated from each other early in larval life, and are not fused along the posterior margin in the manner characteristic of rhynchonelliform brachiopods.

Shell morphology. The two valves are commonly of unequal size, with the ventral valve typically larger. Because the shell increases in size by increments laid down at the mantle margin, ontogenetic changes in shape are faithfully recorded by growth lines. The valves may be further ornamented by concentric folds (rugae), growth lamellae, or radially disposed ribs of various amplitude and wavelength. Numerous extinct genera are characterized by extravagant development of spines. In rhynchonelliforms, the posterior region of one or both valves is commonly differentiated from the remainder of the valve as a somewhat flattened cardinal area. The ventral area is typically further modified by the pedicle

opening; among articulated brachiopods this consists of a triangular delthyrium which may be partially closed by a pseudodeltidium or a pair of deltidial plates. A corresponding triangular opening, the notothyrium, may be developed on the dorsal cardinal area. Internally, muscle scars and mantle canal impressions may be apparent, together with structures of varying complexity associated with articulation (for example, dental and socket plates) and support of the lophophore (crura, spiralia, loops). Calcite spicules are present in the mantle tissues of some extant terebratulides. Linguliform valves are less heavily mineralized and more organic-rich than rhynchonelliforms. Extant craniiforms, cemented to a hard substrate, have ventral valves that are flat and extremely thin, with gently cap-shaped dorsal valves.

Shell morphology varies considerably externally in overall shape and size, relative biconvexity, and degree of ornamentation. Shell morphology also varies internally in structures involved in articulation, lophophore support, and muscle position, from order to order, among the 26 orders recognized. This extensive variation makes it difficult to succinctly characterize overall trends. *See* RHYNCHONELLIDA; TEREBRATULIDA.

Ecology and biogeography. All modern brachiopods are marine, and there is little doubt from the fossil record that brachiopods have always been confined to the sea. A few genera, however, notably the closely related linguliform brachiopods *Lingula* and *Glottidia*, can tolerate reduced salinities and may survive in environments that would be lethal to the majority of forms. Recent brachiopods occur commonly beneath the relatively shallow waters of the continental shelves, which seem to have been the most favored environment in terms of diversity and abundance, but the bathymetric range of the phylum is very large. Some modern species live intertidally and, at the other extreme, some have been dredged from depths of over 16,500 ft (5000 m).

The majority of brachiopods form part of the sessile benthos and are attached by their pedicle during postlarval life. Paleozoic brachiopods not uncommonly exhibit boreholes penetrating their shells, suggesting death by boring predation; interestingly, modern brachiopods have very few known predators, boring or otherwise, even though most are epifaunal, typically living on hard substrates. *Glottidia* and *Lingula* are exceptional in being infaunal and making burrows with the help of complex musculature connecting the two valves that lack articulation. The loss of the pedicle has occurred multiple times over the course of brachiopod evolution, by either complete suppression or atrophy early in the life of the individual. Such forms lie free on the sea floor, are attached by cementation of part or all of the ventral valve, or are anchored by spines. The geographic distribution and geological setting of some fossil species suggest that they may have been epiplanktonic, attached to floating weed, but such a mode of life is unknown in modern faunas.

Biogeographically, brachiopods today exhibit an antitropical distribution, with their highest diversity closer to the Poles than to the tropics, and higher in the Southern Hemisphere than the Northern Hemisphere. Paleobiogeographical data from the distribution of fossil species indicate that Paleozoic brachiopods exhibited a more typical latitudinal diversity gradient, with high tropical diversity, decreasing toward the polar regions, such as is observed today in the majority of marine invertebrates.

Taxonomic diversity and stratigraphic distribution. Brachiopods formed a major part of skeletonized marine faunas throughout the Paleozoic; linguliforms follow a pattern of diversity through time more or less typical for the Cambrian Evolutionary Fauna described by J. J. Sepkoski, while rhynchonelliforms typify the Paleozoic Evolutionary Fauna. The oldest undoubted brachiopods, the paterinates, occur in the Lower Cambrian Tommotian stage. Representatives of seven of the eight classes are known from the Lower Cambrian, indicating that morphological diversity was high even very early in the history of the phylum. Throughout the Cambrian, linguliforms were more abundant than rhynchonelliforms; but during the Early Ordovician radiation, rhynchonelliform diversity and abundance both increased dramatically and remained high until the end-Permian. In the late Paleozoic, a number of major brachiopod groups became extinct, and the end-Permian extinction event caused diversity to plummet. Four of the five extant orders survived this extinction event, and rediversified, although much more modestly through the Mesozoic and Cenozoic than in the Paleozoic. In the present-day oceans, approximately 120 genera are recognized, dominated by rhynchonelliforms (terebratulides and rhynchonellides in particular). *See* EXTINCTION (BIOLOGY); PATERINIDA; RHYNCHONELLIDA.

S. J. Carlson; A. J. Rowell

Bibliography. S. J. Carlson and M. R. Sandy (eds.), *Brachiopods Ancient and Modern: A Tribute to G. Arthur Cooper*, Paleontol. Soc. Pap. 7, 2001; R. L. Kaesler (ed.), *Treatise on Invertebrate Paleontology*, pt. H: *Brachiopoda* (revised), 6 vols., Geological Society of America and University of Kansas, Boulder and Lawrence, 1997–2006; M. J. S. Rudwick, *Living and Fossil Brachiopods*, Hutchinson, London, 1970.

Brain

A collection of specialized cells (neurons) in the head that regulates behavior as well as sensory and motor functions. Neurons grow long thread-like structures—an axon and a dendritic tree—from their cell bodies, which provide them with a rapid communication network throughout the body. The axon uses pulses to transmit a signal to thousands of other neurons or to muscle or gland cells. The dendritic tree uses waves of electric current to integrate the pulses from thousands of other neurons. Groups of neurons form ganglia in chains along both sides of the body axis from "head to tail." The largest of these paired groups, the brain, is in the head, where the distance receptors (nose, eyes, and ears) are

located. These receptors respond to smells, sights, and sounds coming so far from the collective that the collective has time to receive the inputs, interpret them as signals, plan an action before being overtaken by circumstance, and act while monitoring and correcting its action. These are the minimal functions of a brain. The power of a brain lies not in its size but in the complexity of the connections among its functional parts. *See* NEURON.

Parts. The three main parts of the brain in vertebrates are the cerebrum, the cerebellum, and the brainstem that connects them with each other and with the spinal cord (**Fig. 1**). The two cerebral hemispheres are separated by a midline fissure that is bridged by a massive bundle of axons running in both directions, the corpus callosum. Each hemisphere has a core of groups of neurons (the basal ganglia); an outer shell of neurons in layers (the cerebral cortex); and massive bundles of axons for communication within the cerebrum and with the rest of the brain. These bundles are called white matter because of the waxy myelin sheaths surrounding the axons.

Basal ganglia. This component of the brain comprises three main groups. (1) The thalamus receives axons from all sensory systems and transmits information to the cortex. It also receives feedback from cortical neurons during sensory processing. (2) The striatum, comprising bundles of axons cutting through the groups of neurons, also has two-way communication with the cortex and assists in the organization of body movement. (3) The hypothalamus receives orders from the cortex and organizes the chemical systems that support body movement. One output channel is hormonal, and controls the pituitary gland (hypophysis) which in turn controls the endocrine system. The other channel is neural, comprising axons coursing through the brainstem and spinal cord to the motor neurons of the autonomic nervous system, which regulates the heart, blood vessels, lungs, gastrointestinal tract, sex organs, and skin. The autonomic and endocrine systems are largely self regulating, but they are subject to control by the cortex through the hypothalamus. *See* AUTONOMIC NERVOUS SYSTEM; ENDOCRINE SYSTEM (VERTEBRATE); NEUROBIOLOGY.

Cortex. This section of the brain is also called gray matter because it contains the axons, cell bodies, and dendrites of neurons but there is very little myelin. An index of the capacity of a brain is cortical surface area. In higher mammals, the cortical surface increases more rapidly than the volume during fetal development, and as a result the surface folds, taking the form of wrinkles, that is, convexities (gyri) and fissures (sulci) that vary in their details from one brain to another. However, they are sufficiently reliable to serve as landmarks on the cerebral hemisphere that it can be subdivided into lobes.

Lobes. Four lobes make up the shell of each hemisphere (**Fig. 2**), namely the frontal, parietal, temporal, and occipital lobes. Each lobe contains a motor or sensory map, which is an orderly arrangement of cortical neurons associated with muscles and sensory receptors on the body surface. The central sulcus delimits the frontal and parietal lobes. The precentral gyrus contains the motor cortex whose neurons transmit signals to motor neurons in the brainstem and spinal cord which control the muscles in the feet, legs, trunk, arms, face, and tongue of the opposite side of the body, in that order from medial to lateral position between Broca's area and the tongue area of the motor cortex. The number of neurons for each section is determined by the fineness of control, not the size of the muscle; for example, the lips and tongue have larger areas than the trunk. Within the postcentral gyrus is the primary somatosensory

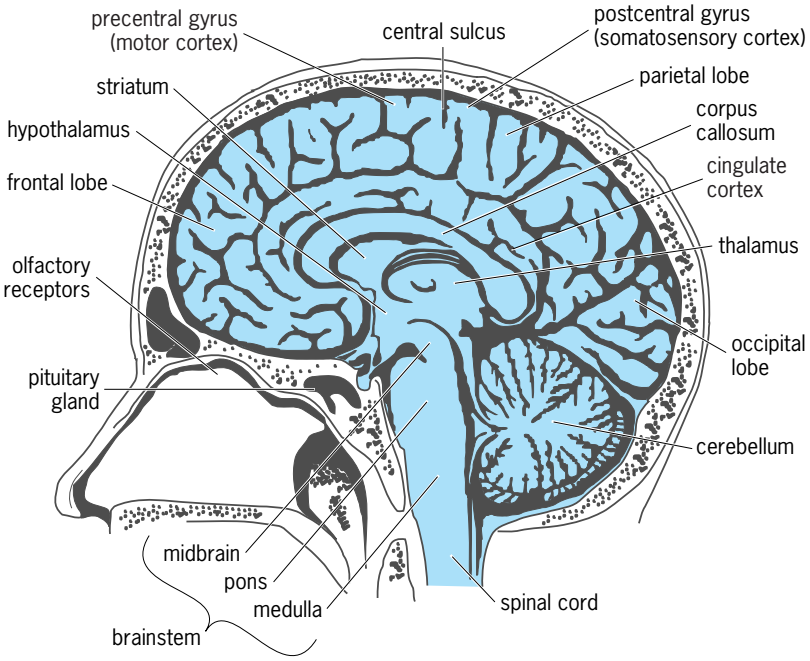

Fig. 1. Midsagittal (midline, medial) section through the human brain. (*After C. R. Noback, The Human Nervous System, 4th ed., McGraw-Hill, 1991*)

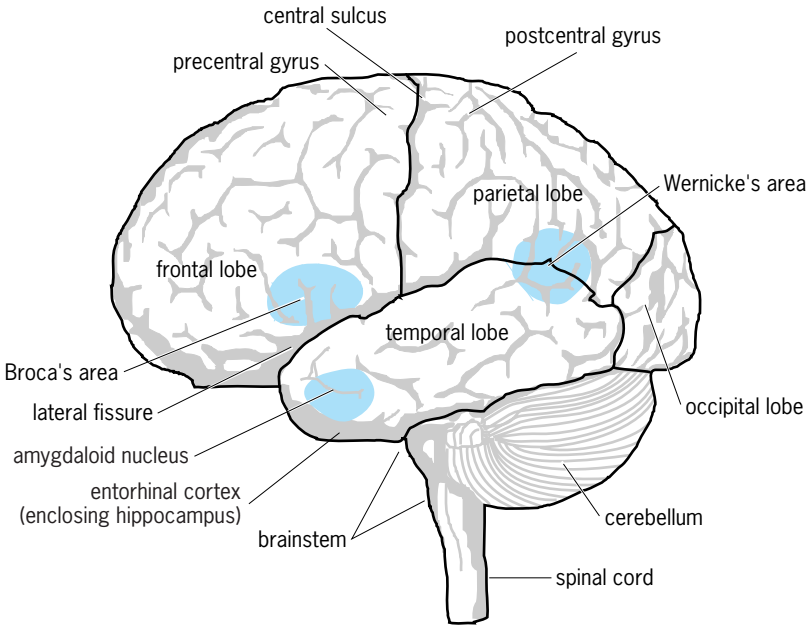

Fig. 2. Lateral surface view of the cerebral hemisphere. (*After C. R. Noback, The Human Nervous System, 4th ed., McGraw-Hill, 1991*)

cortex. Sensory receptors in the skin, muscles, and joints send messages to the somatosensory cortical cells through relays in the spinal cord and the thalamus to a map of the opposite side of the body in parallel to the map in the motor cortex. The lateral fissure separates the temporal lobe from the parietal and frontal lobes. The cortex on the inferior border of the fissure receives input relayed through the thalamus from the ears to the primary auditory cortex. The occipital lobe receives thalamic input from the eyes and functions as the primary visual cortex.

In humans, the association cortex surrounds the primary sensory and motor areas that make up a small fraction of each lobe. The occipital lobe has many specialized areas for recognizing visual patterns of color, motion, and texture. The parietal cortex has areas that support perception of the body and its surrounding personal space. Its operation is manifested by the phenomenon of phantom limb, in which the perception of a missing limb persists for an amputee. Conversely, individuals with damage to these areas suffer from sensory neglect, because parts of the body may no longer exist for them. The temporal cortex contains areas that provide recognition of faces and of rhythmic patterns, including those of speech, dance, and music. The frontal cortex provides the neural capabilities for constructing patterns of motor behavior and social behavior. It was the rapid enlargement of the frontal and temporal lobes in human evolution over the past half million years that supported the transcendence of humans over other species. This is where the capacity to create works of art, and also to anticipate pain and death is located. Insight and foresight are both lost with bilateral frontal lobe damage, leading to reduced experience of anxiety, asocial behavior, and a disregard of consequences of actions.

Motor systems. A small part of frontal lobe output goes directly to motor neurons in the brainstem and spinal cord for fine control of motor activities, such as search movements by the eyes, head, and fingers, but most goes either to the striatum from which it is relayed to the thalamus and then back to the cortex, or to the brainstem from which it is sent to the cerebellum and then through the thalamus back to the cortex. In the cerebellum, the cortical messages are integrated with sensory input predominantly from the muscles, tendons, and joints, but also from the eyes and inner ears (for balance) to provide split-second timing for rapid and complex movements. The cerebellum also has a cortex and a core of nuclei to relay input and output. Their connections, along with those in the cerebral cortex, are subject to modification with learning in the formation of a working memory (the basis for learned skills). *See* MEMORY; MOTOR SYSTEMS.

Limbic system. The cerebellum and striatum do not set goals, initiate movements, store temporal sequences of sensory input, or provide orientation to the spatial environment. These functions are performed by parts of the cortex and striatum deep in the brain that constitute another loop, the limbic system. Its main site of entry is the entorhinal cortex, which receives input from all of the sensory cortices, including the olfactory system. The input from all the sensory cortices is combined and sent to the hippocampus, where it is integrated over time. Hippocampal output returns to the entorhinal cortex, which distributes the integrated sensory information to all of the sensory cortices, updates them, and prepares them to receive new sensory input. This new information also reaches the hypothalamus and part of the striatum (the amygdaloid nucleus) for regulating emotional behavior. Bilateral damage to the temporal lobe including the hippocampus results in loss of short-term memory. Damage to the amygdaloid nucleus can cause serious emotional impairment. The Papez circuit is formed by transmission from the hippocampus to the hypothalamus by the fornix, then to the thalamus, parietal lobe, and entorhinal cortex. The limbic system generates and issues goal-directed motor commands, with corollary discharge to the sensory systems that prepares them for the changes in sensory input caused by motor activity (for example, when one speaks and hears oneself, as distinct from another).

Split brain. Each hemisphere has its own limbic, Papez, cortico-thalamic, cortico-striatal, and cortico-cerebellar loops, together with sensory and motor connections. When isolated by surgically severing the callosum, each hemisphere functions independently, as though two conscious persons occupied the same skull, but with differing levels of skills in abstract reasoning and language. The right brain (spatial)–left brain (linguistic) cognitive differences are largely due to preeminent development of the speech areas in the left hemisphere in most right- and left-handed persons. Injury to Broca's area (located in the frontal lobe) and Wernicke's area (located in the temporal lobe) [Fig. 2] leads to loss of the ability, respectively, to speak (motor aphasia) or to understand speech (sensory aphasia), appearing as a loss of declarative memory for facts and words. Studies of blood flow show that brain activity during intellectual pursuits is scattered broadly over the four lobes in both hemispheres. *See* CENTRAL NERVOUS SYSTEM; HEMISPHERIC LATERALITY. Walter J. Freeman

Bibliography. F. E. Bloom and A. Lazerson, *Brain, Mind and Behavior*, 1988; E. R. Kandel, J. H. Schwartz, and T. M. Jessell (eds.), *Essentials of Neural Science and Behavior*, Appleton & Lange, Norwalk, 1995; C. R. Noback, *The Human Nervous System*, 4th ed., 1991; H. K. Yusef, *Understanding the Brain and Its Development: A Chemical Approach*, 1992.

Brake

A machine element for applying a force to a moving surface to slow it down or bring it to rest in a controlled manner. In doing so, it converts the kinetic energy of motion into heat which is dissipated into the atmosphere. Brakes are used in motor vehicles, trains, airplanes, elevators, and other machines. Most brakes are of a friction type in which a fixed surface

is brought into contact with a moving part that is to be slowed or stopped. Brakes in general connect a moving and a stationary body, whereas clutches and couplings usually connect two moving bodies. *See* AILERON; CLUTCH; COUPLING.

The limitations on the applications of brakes are similar to those of clutches, except that the service conditions are more severe because the entire energy is absorbed by slippage which is converted to heat that must be dissipated. The important thing is the rate at which energy is absorbed and heat dissipated. With friction brakes, if the temperature of the brake becomes too high, the result is a lowering of the friction force, called fading.

There are also electrical and hydrodynamic brakes. The electrical type may be electromagnetic, eddy-current, hysteresis, or magnetic-particle. The hydrodynamic type works somewhat like a fluid coupling with one element stationary. Another type, used on electric trains, is the regenerative brake. This electrical machine can be used as a motor or a generator. As a generator, it brakes the train and stores the generated electricity in an accumulator. Another type of brake is the air brake (flaps) on an airplane.

Friction types. Friction brakes are classified according to the kind of friction element employed and the means of applying the friction forces. *See* FRICTION.

Single-block. The simplest form of brake consists of a short block fitted to the contour of a wheel or drum and pressed against its surface by means of a lever on a fulcrum, as widely used on railroad cars. The block may have the contour lined with friction-brake material, which gives long wear and a high coefficient of friction. The fulcrum may be located with respect to the lever in a manner to aid or retard the braking torque of the block. The lever may be operated manually or by a remotely controlled force (**Fig. 1***a*).

Double-block. Two single-block brakes in symmetrical opposition, where the operating force on the end of one lever is the reaction of the other, make up a double-block brake (Fig. 1*b*). External thrust loads are balanced on the rim of the rotating wheel.

External-shoe. An external-shoe brake operates in the same manner as the block brake, and the designation indicates the application of externally contracting elements. In this brake the shoes are appreciably longer, extending over a greater portion of the drum (Fig. 1*c*). This construction allows more combinations for special applications than the simple shoe, although assumptions of uniform pressure and concentrated forces are no longer possible. In particular, it is used on elevator installations for locking the hoisting sheave by means of a heavy spring when the electric current is off and the elevator is at rest. Self-energization is possible, as with block brakes, depending upon the arrangement of the supporting mechanism. *See* ELEVATOR.

Internal-shoe. An internal shoe has several advantages over an external shoe. Because the internal shoe works on the inner surface of the drum, it is protected from water and grit (Fig. 1*d*). It may be designed in a more compact package, is easily activated, and is effective for drives with rotations in both directions. The internal shoe is used in the automotive drum brake, with hydraulic piston actuation. *See* AUTOMOTIVE BRAKE.

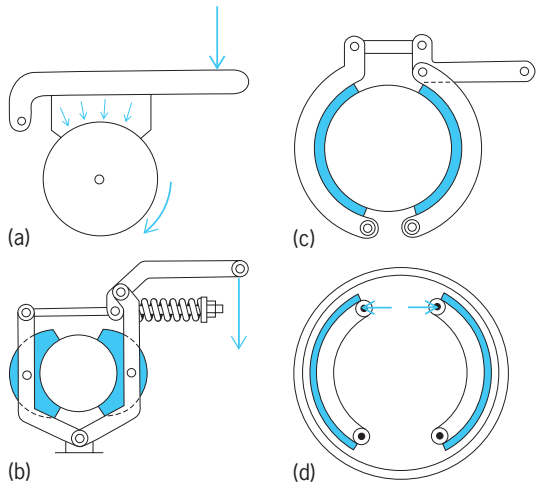

Fig. 1. Brakes. (*a*) Single-block brake. The block is fixed to the operating lever; force in the direction of the arrow applies the brake. *(b)* Double-block brake. The blocks are pivoted on their levers; force in the direction of the arrow releases the brake. (*c*) External shoe brake. Shoes are lined with friction material. (*d*) Internal shoe brake with lining.

Band. Hoists, excavating machinery, and hydraulic clutch-controlled transmissions have band brakes. They operate on the same principle as flat belts on pulleys. In the simplest band brake, one end of the belt is fastened near the drum surface, and the other end is then pulled over the drum in the direction of rotation so that a lever on a fulcrum may apply tension to the belt.

The belt may be a thin metallic strip with a friction lining. The method of applying the lever on the fulcrum and attaching the belt determines the structural operation of the brake. These variations make possible a sensitive differential brake, self-energizing brakes, and brakes operating with equal effectiveness in both directions. The radial force of the brake is proportional to the tension in the band. In automotive automatic transmissions, the bands are almost completely circular (**Fig. 2**). They are applied by

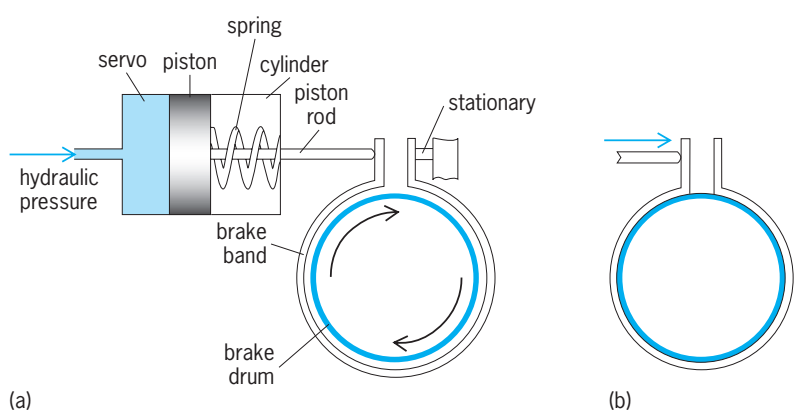

Fig. 2. Band brake around a brake drum. (*a*) Brake band released; the drum is free to rotate. *(b)* Brake applied; tightening the band stops rotation of the drum and holds it stationary. *(Automotive & Technical Writing, Charlottesville, Virginia)*

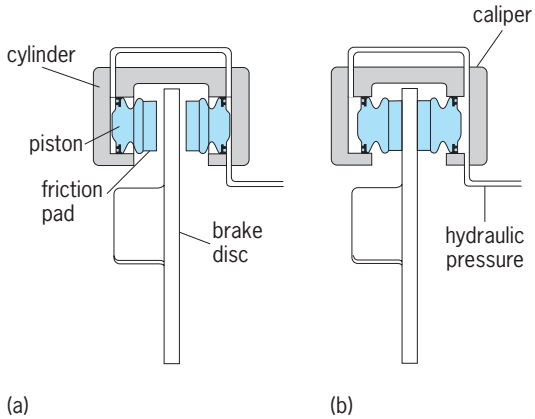

Fig. 3. Caliper disk brake. *(a)* Friction pads on either side of a disk that is free to rotate. *(b)* Brake applied, hydraulic pressure forces the pistons toward the disk to stop its rotation and hold it stationary. *(Automotive & Technical Writing, Charlottesville, Virginia)*

hydraulic pressure and released by spring force, usually aided by redirecting the hydraulic pressure.

Disk. Structurally similar to disk clutches, disk brakes have long been used on hoisting and similar apparatus. Because more energy is absorbed in prolonged braking than in clutch startup, additional heat dissipation must be provided in equivalent disk brakes. Disk brakes are used for the wheels of aircraft, where segmented rotary elements are pressed against stationary plates by hydraulic pistons. Flexibility, self-alignment, and rapid cooling are inherent in this design. Another application is the bicycle coaster brake.

The caliper disk brake (**Fig. 3**) is widely used on automotive vehicles. It consists of a rotating disk which can be gripped between two friction pads. The caliper disk brake is hydraulically operated, and the pads cover between one-sixth and one-ninth of the swept area of the disk. *See* AUTOMOTIVE BRAKE.

Actuators. A brake can be arranged so that it is normally either released or applied, a spring usually forcing it into the normal condition. A controllable external force then applies or releases the brake. This force can be, for example, mechanical through a linkage or cable to a lever, pneumatic or hydraulic through a cylinder to a piston, or electric through a solenoid to a plunger.

Air brake. Railway brakes are normally applied air brakes; if the air coupling to a car is broken, the brakes are applied automatically. To apply the brakes, the brake operator releases the compressed air that is restraining the brakes by means of a diaphragm and linkage. Over-the-road trucks and buses use air brakes. Another form of air brake consists of an annular air tube surrounding a jointed brake lining that extends completely around the outside of a brake drum. Air pressure expands the tube, pressing the lining against the drum. *See* AIR BRAKE.

Electromagnetic brake. In commonly encountered electromagnetic brakes, the actuating force is applied by an electric current through a solenoid. Direct current (dc) gives greater braking force than alternating current (ac), and is therefore used almost exclusively; ac is rectified if used. Usually the electromagnetic force releases the brake against a compression spring, which provides braking action if the power fails, and overspeed protection when the brake is used with dc series motors.

In a towed highway vehicle, such as a boat trailer, the trailer brakes are applied by electromagnets at each wheel. When the driver operates a controller, the electromagnets are attracted to disks on the rotating wheels. This causes the electromagnets to shift through a limited arc which forces the brake shoes into contact with the drum. Further movement of the controller allows more current to flow in the electromagnets, producing stronger magnetic fields and greater braking action. Donald L. Anglin

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; J. E. Shigley and C. R. Mischke, *Machine Design Fundamentals*, 1989.

Branch circuit

As defined by the National Electrical Code, the portion of an electrical wiring system that extends beyond the final, automatic overcurrent protective device (circuit breaker or fuse), which is recognized by the National Electrical Code for use as a branch-circuit overcurrent protector, and that terminates at the utilization equipment or outlet (such as a lighting fixture, motor, receptacle, or heater). Thermal cutouts, motor overload devices, and fuses in luminaires or plug connections, which are considered "supplementary overcurrent" devices, are not approved for branch-circuit protection and do not establish the point of origin of a branch circuit (see **illustration**).

Motor branch circuits must have conductors with current-carrying capacity at least 125% of the motor full-load current rating. Overcurrent protection must be capable of carrying the starting current of the motor. Maximum values are tabulated for

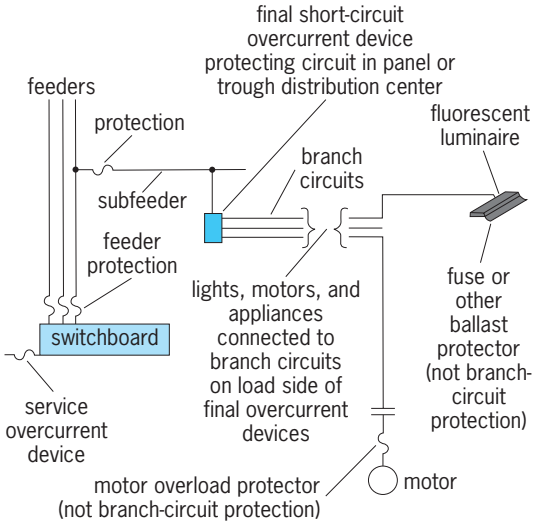

Typical electrical wiring system. Branch circuits are those portions of system beyond the final overcurrent device approved for branch-circuit protection.

various currents in the National Electrical Code. *See* ELECTRICAL CODES.

Branch circuits, other than motor branch circuits, are classified by the maximum rating of the overcurrent device. Thus a branch circuit with a 20-A circuit breaker is designated a 20-A branch circuit, even though the conductors may be capable of carrying higher current.

Individual branch circuits serving a single fixed appliance, such as a clothes dryer or a water heater, may be of any ampere rating if the overcurrent protection (fuse or circuit breaker) does not exceed 150% of the rated current of the appliance (where it is 10 A or more). For continuous loads of 3 h or longer duration, the branch circuit rating must be at least 125% of the continuous current plus the total of any noncontinuous load. Conductors supplying such loads must also have an ampacity (current-carrying capacity) equal to at least 125% of the continuous load plus any noncontinuous load.

Branch circuits serving more than one outlet or load are limited by the National Electrical Code to three types:

1. Circuits of 15 or 20 A may serve lights and appliances; the rating of one portable appliance may not exceed 80% of the circuit capacity; the total rating of fixed appliances may not exceed 50% of circuit capacity if lights or portable appliances are also supplied.
2. Circuits of 30 A may serve fixed lighting units with heavy-duty lampholders in other than dwellings or appliances in any occupancy.
3. Circuits of 40 or 50 A may serve fixed lighting with heavy-duty lampholders in other than dwellings, fixed cooking appliances, or infrared heating units.

Multiwire branch circuits consist of two or more ungrounded conductors having an identical potential difference (voltage) between each other and a grounded (neutral) conductor having an equal potential difference between it and the individual ungrounded conductors.

Conventionally the current-carrying capacity of conductors may be the same as the branch-circuit rating (for example, no. 12 conductors in a 20-A branch circuit); however, larger conductors may be required to avoid excessive voltage drop where the circuit must be run to supply remote loads. *See* WIRING.

Brian J. McPartland; J. F. McPartland

Bibliography. B. J. McPartland and J. F. McPartland, *National Electrical Code Handbook*, 25th ed., 2005; J. F. McPartland and B. J. McPartland, *The Handbook of Practical Electrical Design*, 2d ed., 2000; *National Electrical Code*, NFPA 70, 2005; H. P. Richter and H. P. Hartwell, *Practical Electrical Wiring*, 19th ed., 2005.

Branched polymer

A polymer chain having branch points that connect three or more chain segments. Examples of branched polymers include long chains having occasional and usually short branches comprising the same repeat units as the main chain (nominally termed a branched polymer); long chains having occasional branches comprising repeat units different from those of the main chain (termed graft copolymers); main chains having one long branch per repeat unit (referred to as comb polymers); and small core molecules with branches radiating from the core (star polymers). So-called starburst or dendritic polymers are a special class of star polymer in which the branches are multifunctional, leading to further branching with polymer growth. Star, comb, and starburst polymers, (**Fig. 1**), especially the last, represent interesting molecular structures that may lead to unusual supramolecular structures (for example, micelles and liposomes) that mimic the functions of complex biomolecules.

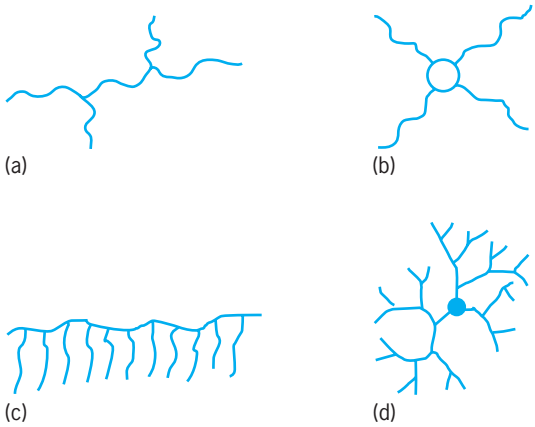

Fig. 1. Examples of branched polymers. (*a*) Branched polymer (if arms are of composition similar to backbone) or graft polymer (if compositions are different). (*b*) Star polymer. (*c*) Comb polymer. (*d*) Dendritic polymer.

Classic type. A classic example of a branched polymer where the branches have the same chemical composition as the backbone is low-density polyethylene (LDPE). When ethylene is polymerized by using free-radical initiators, short branches, typically ethyl and *n*-butyl along with lesser amounts of a few other types including very long branches, are generated by a process known as chain transfer to polymer. A propagating radical at the end of a growing polyethylene chain, which is a primary alkyl radical, is capable of abstracting a backbone hydrogen atom from the polymer backbone, leading to a more stable secondary radical. The preponderance of ethyl and *n*-butyl groups is thought to be the result of intramolecular chain transfer proceeding via conformations involving six-membered ring transition states, as in the reaction below.

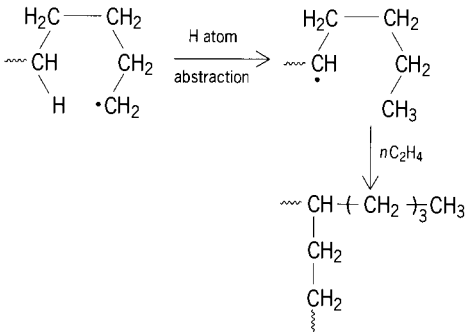

The short branches interfere to some degree with crystallization of the backbone and, since the chains cannot pack as efficiently, the density is lower than that of high-density polyethylene (HDPE). However, LDPE is somewhat tougher because of its greater amorphous content. The amorphous regions are above their glass transition temperature at room temperature, and the liquidlike nature of these regions provides for energy dissipation and hence toughness. Much LDPE is prepared by copolymerization of ethylene with α-olefins such as 1-hexene using organometallic initiators, a route that offers more control over branch type and concentration, and hence properties. This type is known as linear low-density polyethylene (LLDPE).

Branching can occur to some extent in the polymerization of other monomers, including vinyl acetate (radical polymerization), the base-initiated ring-opening polymerization of caprolactam, and the cationic ring-opening polymerization of ethyleneimine. Step-growth polymerizations using monomers with an average functionality greater than two can also lead to branching and, in many cases, crosslinking. An example of branching in biopolymers is the structure of amylopectin, a component of starch [monomer, structure (**1**)]; in amylopectin,

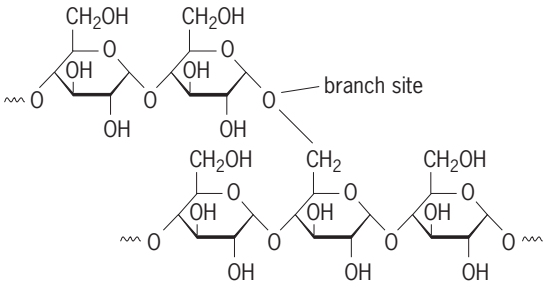

(**1**)

the branches on the main backbone are themselves branched [structure (**2**)]. Glycogen, the storage

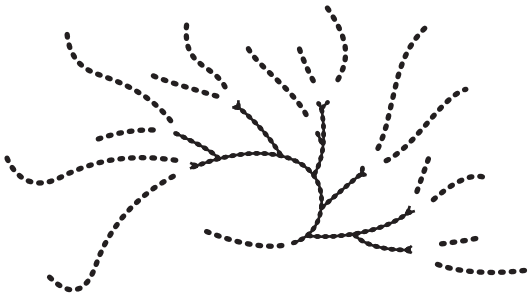

(**2**)

polymer of glucose in animals, is even more heavily branched than amylopectin. *See* BIOPOLYMER; GLYCOGEN; POLYOLEFIN RESINS; RING-OPENING POLYMERIZATION; STARCH.

Graft and comb. There are many examples of graft copolymers. They can be prepared by generation of a radical on a polymer backbone via ultraviolet or ionizing radiation in the presence of a monomer, or by polymerization of a monomer in the presence of a polymer with a polymerizable linkage such as a carbon-carbon double bond. An important example of the latter type is a graft copolymer of polystyrene and polybutadiene [structure (**3**)]. The resulting

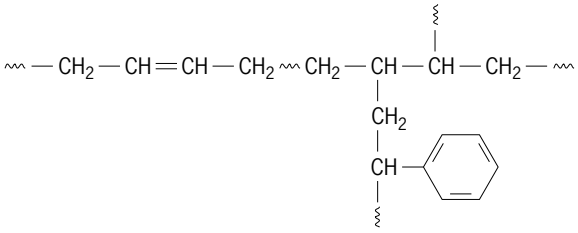

(**3**)

material is far tougher than polystyrene itself because of the presence of small domains of rubbery polybutadiene. Graft copolymers can also be prepared by coupling polymers having reactive groups at one chain end with different polymers having functional groups on their backbones. Preformed polymers with a polymerizable end (referred to as macromonomers) can be homopolymerized to form comb polymers or can be copolymerized with a variety of monomers to yield graft copolymers. Finally, graft copolymers can be prepared by mechanically blending different polymers. Homolytic bond cleavage resulting from shearing leads to the formation of free radicals on the polymer chains, and the coupling of these onto different polymer chains results in grafting. *See* COPOLYMER; FREE RADICAL.

Star and comb. Star polymers can be prepared by a core-first or arm-first approach. In the former, initiation of polymerization occurs from a multifunctional core molecule. In the latter, polymer chains are synthesized and then coupled to a core molecule via reactive end groups. Steric considerations dictate the efficiency of the coupling process. For example, reaction of polystyryl anions with tetrachlorosilane yields a mixture of three- and four-arm star polymers along with a coupled polymer and a homopolymer. A similar reaction with the less sterically demanding polybutadienyl anion gives an almost quantitative yield of four-arm star polymer. The arms need not be homopolymers. For example, star polymers with styrene-butadiene block copolymer arms have been prepared by reaction of polybutadienyl anionic ends with reactive core molecules such as tetrachlorosilane (**Fig. 2**). Star polymers are of interest because they are smaller in size than a linear polymer of comparable molecular weight and therefore will usually have lower softening points and solution viscosities, which can facilitate processing. Comb polymers can also be prepared by using the ideas employed for the synthesis of star polymers, as well as via polymerization of macromonomers, which are polymers terminated with polymerizable groups such as vinyl and epoxy.

Starburst and dendritic. The field of branched polymers also includes starburst, or dendritic, polymers (also referred to as arborols and cascade molecules). These unique materials are characterized by a highly branched, three-dimensional structure emanating from a central, multifunctional molecule. The difference between dendritic polymers and other so-called hyperbranched polymers such as amylopectin is that in the former each monomer unit has at least one branch point, and therefore the extent of branching is extraordinarily high. The pronounced branching allows the attainment of spherical molecules that resemble micelles.

Numerous syntheses of starburst polymers have been developed. An initial example involved the addition of ammonia to methyl acrylate. The resulting triester was then reacted with a diamine such as ethylenediamine, resulting in a new triamine that was effectively the core molecule. Repetition of these steps led to the production of a starburst polymer (**Fig. 3**). This approach, in which the synthesis is begun at the center of the polymer, is termed the divergent method, and the repetitions of the reactions are referred to as generations. The number of reactions that need to be performed increases geometrically with the number of generations. The molecule will tend toward a spherical shape as the number of generations increases. The size of the molecule can be predicted from the number of generations, and indeed experimentally determined diameters (via electron microscopy and size exclusion chromatography) are in excellent agreement with predictions. Also, the surface of the molecule will have a high concentration of functional groups (amines in this case).

An alternative approach to dendrimer synthesis is the convergent method, where the synthesis is begun at the outside of the dendrimer. Here portions of what will ultimately be a dendrimer, referred to as dendrons, are synthesized and then coupled to yield the dendrimer. Use is made of series of protection or deprotection steps. One advantage of this approach is that it does not limit the type and concentration of a particular functional group at the surface of the dendrimer. It has been successfully employed to prepare a wide variety of dendrimers, including polyesters and poly(aryl ethers), as well as novel block copolymers.

Dendrimers can be viewed as nanostructural compounds that can be used as building blocks for novel supramolecular assemblies. For example, it is possible to react dendrimers bearing nucleophilic surface groups (for example, amines) with others bearing electrophilic groups to form dendrimer clusters. Dendrimers can also be reacted with functionally terminal polymers to obtain networks having junction points of very high functionality. The high concentration of surface functionality and the ability to manipulate the functionality through simple organic reactions has spurred interest in their potential use as targeted drug carriers. Indeed, it may be that applications of dendrimers will be particularly significant in biological science in view of the ability to control functionality and supramolecular structure. *See* MICROMOLECULAR ENGINEERING; POLYMER; POLYMERIZATION.

Gary E. Wnek

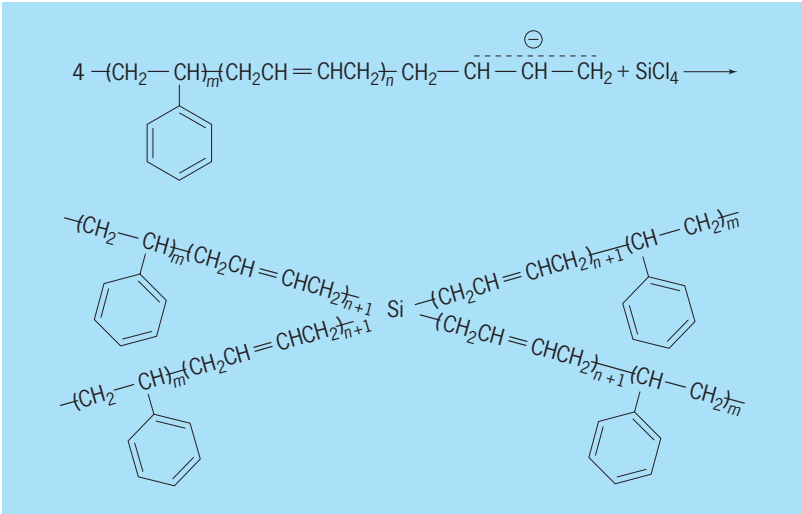

Fig. 2. Reaction for preparation of a star polymer with styrene-butadiene block copolymer arms.

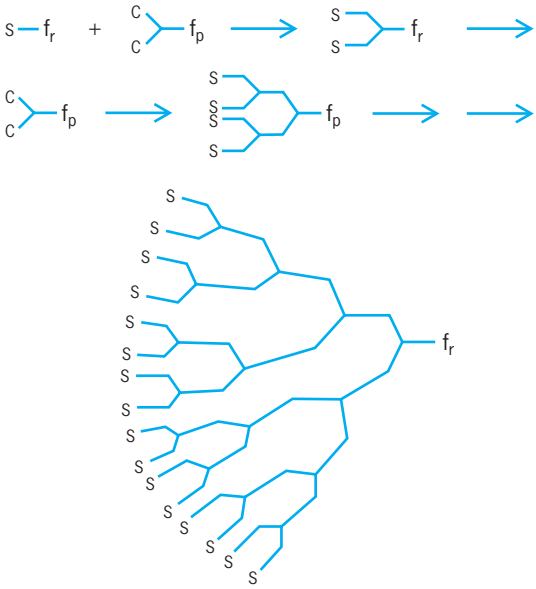

Fig. 3. Steps in the synthesis of a starburst polymer. s = surface. o = interior. f_r = reactive group. f_p = protected group.

Bibliography. N. Hadjichristidis, *Block Copolymers: Synthetic Strategies, Physical Properties, and Application*, 2002; K. Matyjaszewski and T. P. Davis (eds.), *Handbook of Radical Polymerization*, 2002; G. Odian, *Principles of Polymerization*, 4th ed., 2004.

Branchiopoda

A class of crustaceans. Formerly, four extant orders were recognized: Anostraca, Notostraca, Conchostraca, and Cladocera, the first two of which are well defined. The Conchostraca consist of two

groups which, although superficially similar, differ in so many fundamental features that they have been placed in separate orders, Laevicaudata and Spinicaudata. The Cladocera are in fact a heterogeneous assemblage of organisms that have now been split into four orders: Anomopoda, Ctenopoda, Onychopoda, and Haplopoda. Although the name Cladocera has to be abandoned as a taxonomic unit, it will doubtless continue to be used as a convenient descriptive term for members of these four orders. There are also two extinct orders, Lipostraca and Kazacharthra. Branchiopods have a long history. Both the Lipostraca and Spinicaudata were differentiated by Devonian times, and their ancestors doubtless originated even earlier. *See* ANOMOPODA; ANOSTRACA; HAPLOPODA; LAEVICAUDATA; NOTOSTRACA; ONYCHOPODA; SPINICAUDATA.

Living members of the Branchiopoda range from less than 0.5 mm (0.02 in.) to (exceptionally) 100 mm (4 in.) in length. Form is exceedingly diverse. The trunk may be abbreviated and of probably as few as 5 segments (although segmentation is sometimes obscure) or elongate and of more than 40 segments. Trunk limbs range from 5 to about 70 pairs. A carapace is often present, either as a dorsal shield or as a bivalved structure; anostracans lack a carapace. The eyes were primitively paired and are sometimes stalked and sometimes sessile; in several orders they have fused to give a single median sessile eye. Members of several orders use their antennae as organs of locomotion—a retention of the similar mechanism used by the nauplius larva. Others swim by means of trunk limbs. The mandibles are usually of the crushing, rolling type; only in a few cases have biting or piercing mandibles been acquired. Other mouthparts are usually small, but important. Trunk limb structure is extremely diverse. Many-limbed orders tend to have limbs that differ from each other only in degree; the Anomopoda have limbs that differ greatly from each other in form and from species to species in relation to habits.

Most species are microphagous. Microphagous forms collect their food either by direct scraping from surfaces, which may or may not be followed by filtration, or by abstracting suspended particles by the use of complicated filtering devices. A few species are omnivores or carnivores.

Reproductive habits are diverse. Parthenogenesis is widespread, and the production of highly resistant resting eggs that can withstand freezing and drying and retain their viability for several years is highly characteristic. Almost all species live in fresh water, but a few are marine and some frequent highly saline situations. Branchiopods have a worldwide distribution. *See* CRUSTACEA.

Geoffrey Fryer

Branchiura

A subclass of the Crustacea known as the fish lice. They are ectoparasites of fresh-water and marine fish on such areas as the base of fins and the gill chamber walls. *Argulus* (see **illus.**) is a common genus. They are a small homogeneous group, less than 100 species, with worldwide distribution, and very much alike in appearance. The disklike head and thorax (cephalothorax) is strongly flattened and bears a small unsegmented abdomen bilobed at its distal end. Larger species may exceed 1 in. (25 mm) in length.

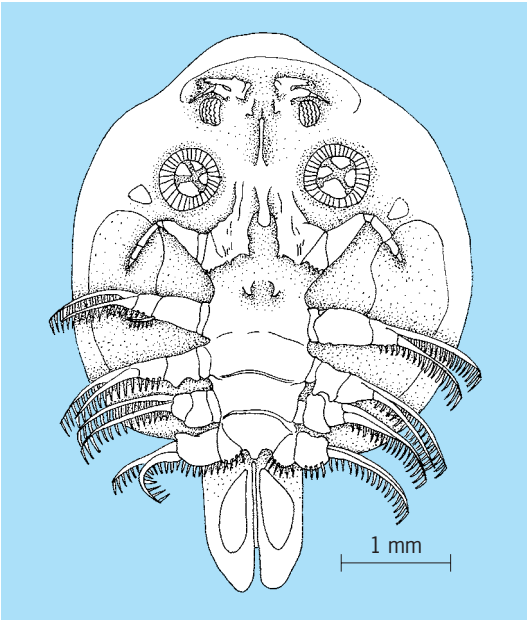

Argulus japonicus, **male fish louse of the Branchiura.**

Taxonomy. The Branchiura have often been included within the Copepoda either as the subclass Branchiura, opposed to the remainder grouped under Eucopepoda, or as the order Arguloida equivalent in rank to the Calanoida, Cyclopoida, and Harpacticoida. Though widespread acknowledgment is not yet apparent, critical appraisal of branchiuran morphology, ontogeny, and ethology has revealed several basic characteristics which have no homologs in the eucopepods. On the basis of these fundamental differences, linkage to the Copepoda is no more justified than placement within the Branchiopoda, Ostracoda, or even Cirripedia. *See* COPEPODA.

Some of the morphological characteristics distinguishing Branchiura from Copepoda include the presence of compound eyes, totally different segmentation of the swimming legs, location of the genital orifice on the thorax, absence of a spermatophore, and use of the second, third, and fourth pairs of legs in copulation. The absence of larval stages, either suppressed in the embryonic period or as a free-swimming nauplius followed by metamorphosis into the immature adult form, is noteworthy. Behavioral differences occur in the method of spawning; branchiurans attach small masses of eggs enveloped in gelatinous sheaths to stones and other firm objects on the bottom.

Morphology. The appendages, all on the cephalothorax, consist of two pairs of antennae, the

mouthparts which are one pair of mandibles and two pairs of maxillae, and four pairs of swimming legs. They originate on the underside of the body near the midline and extend laterally, perpendicular to the longitudinal axis. The wafer-thin cephalothorax with its appendages appressed to the underside permits fish lice to flatten themselves against a fish's skin. The resulting highly streamlined contour serves to minimize the considerable force of frictional drag found in a dense medium like water.

Other adaptations for holding fast to a mucus-covered swimming fish include numerous spinules and strategically placed hooks on the underside of the body and utilization of specialized portions of the body as suction cups. Despite these elaborate specializations, fish lice never lose the ability to abandon a host and swim. At least some of this activity is associated with mating and spawning.

Reproduction. Breeding males seek the female which remains passively on the host. Use is made of the second to fourth pairs of swimming legs in copulation, these appendages bearing pegs and corresponding sockets arranged in a manner unique to each species. Spawning females swim from the host to find firm objects, such as stones, on which to fix the mass of eggs encased in a gelatinous sheath. Following direct embryonic development the hatching juvenile resembles the parent although the appendages do not assume an adult form until later. Sexual maturity, occurring about a month after hatching, requires at least seven molt stages and does not seem to stop growth.

Nutrition. Branchiurans feed on tissue fluids, especially blood. The mouth, located at the end of a movable proboscis, is applied to the host's skin and the rasping mandibles make the necessary wound. Fish lice usually are no threat to fish populations. In restricted areas such as hatchery ponds, however, an infestation can increase to levels that bring about high fish mortality. Treatment may require such drastic measures as drainage and cleaning before restocking can be successful. *See* CRUSTACEA.

Abraham Fleminger

Bibliography. M. F. Martin, On the morphology and classification of *Argulus* (Crustacea), *Proc. Zool. Soc. London*, 1932:771–806, 1932; O. L. Meehean, A review of the parasitic Crustacea of the genus *Argulus* in the collections of the United States National Museum, *Proc. U.S. Nat. Mus.*, 88:459–522, 1940; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; C. B. Wilson, Parasitic copepods in the United States National Museum, *Proc. U.S. Nat. Mus.*, 94:529–582, 1944.

Brass

An alloy of copper and zinc. In manufacture, lump zinc is added to molten copper, and the mixture is poured either into castings ready for use or into billets for further working by rolling, extruding, forging, or a similar process. Brasses containing 75–85% copper are red-gold and malleable; those containing 60–70% are yellow and also malleable; and those containing 50% or less copper are white, brittle, and not malleable. Alpha brass contains up to 36% zinc; beta brass contains nearly equal proportions of copper and zinc. Specific brasses are designated as follows: gilding (95% copper: 5% zinc), red (85:15), low (80:20), and admiralty (70:29, with balance of tin). Naval brass is 59–62% copper with about 1% tin, less than that of lead and iron, and the remainder zinc. The nickel silvers contain 55–70% copper and the balance nickel. With small amounts of other metals, other names are used. Leaded brass is used for castings.

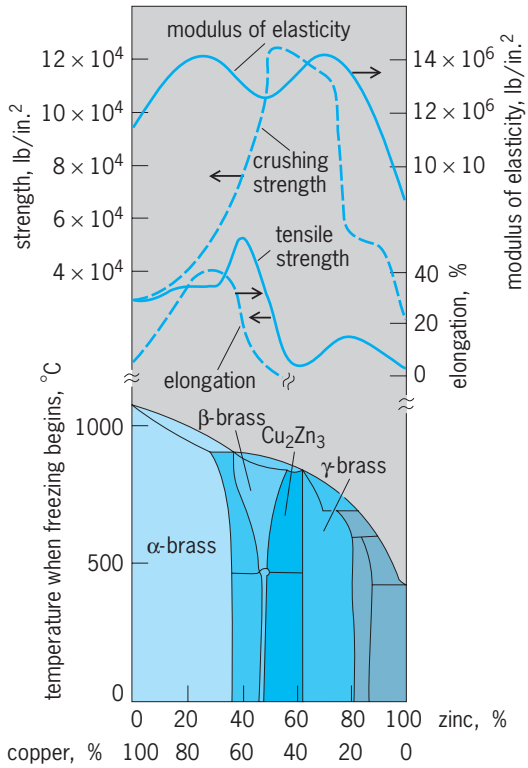

Approximate mechanical properties of cast brass in relation to alloy composition in simplified phase diagram.

Addition of zinc to copper produces a material that is harder and stronger than copper, as shown in the **illustration**, yet retains the malleability, ductility, and corrosion resistance of copper. Mechanical and heat treatment greatly vary the properties of the finished product. Because brass with 64% or more copper (α-brass) forms a single solid solution, it combines high strength, ductility, and corrosion resistance. The β solid solution is harder and more brittle. With 61% zinc the compound Cu_2Zn_3 forms; it is brittle, having the lowest tensile strength of the brasses. Higher percents of zinc produce other brittle constituents.

Brass stains in moist air; however, the oxide so formed is sufficiently continuous and adherent to retard further oxidation. When brass is required to remain bright, it is either washed in nitric acid and then coated with clear lacquer, or it is regularly polished

and waxed. For brightwork or other corrosion-resistant applications, α-brass is usually used. In atmospheres containing traces of ammonia, cold-worked brass or annealed brass highly stressed in tension may fail by cracking. This is called season cracking or stress corrosion cracking. Susceptibility of cold-worked brass to failure of this kind is reduced by proper annealing.

Brass is widely used in cartridge cases, plumbing fixtures, valves and pipes, screws, clocks, and musical instruments. *See* ALLOY; BRONZE; COPPER; COPPER ALLOYS; CORROSION; ZINC; ZINC ALLOYS.

Frank H. Rockett

Bibliography. R. M. Brick et al., *Structure and Properties of Engineering Materials*, 4th ed., 1977; K. G. Budinski, *Engineering Materials: Properties and Selection*, 6th ed., 1998.

Brayton cycle

A thermodynamic cycle (also variously called the Joule or complete expansion diesel cycle) consisting of two constant-pressure (isobaric) processes interspersed with two reversible adiabatic (isentropic) processes (**Fig. 1**). The ideal cycle performance, predicated on the use of perfect gases, is given by relationships (1) and (2). Thermal efficiency η_T,

$$V_3/V_2 = V_4/V_1 = T_3/T_2 = T_4/T_1 \qquad (1)$$

$$\frac{T_2}{T_1} = \frac{T_3}{T_4} = \left(\frac{V_1}{V_2}\right)^{k-1} = \left(\frac{V_4}{V_3}\right)^{k-1} = \left(\frac{p_2}{p_1}\right)^{\frac{k-1}{k}} \qquad (2)$$

the work done per unit of heat added, is given by Eq. (3). In these relationships V is the volume in

$$\eta_T = [1 - (T_1/T_2)] = \left[1 - \left(\frac{1}{r^{k-1}}\right)\right] \qquad (3)$$

cubic feet, p is the pressure in pounds per square foot, T is the absolute temperature in degrees Rankine, k is the c_p/c_v, or ratio of specific heats at constant pressure and constant volume, and r is the compression ratio, V_1/V_2.

The thermal efficiency for a given gas, air, is solely a function of the ratio of compression. This is also the case with the Otto cycle. For the diesel cycle with incomplete expansion, the thermal efficiency is lower. The overriding importance of high compression ratio for intrinsic high efficiency is clearly demonstrated by these data.

A reciprocating engine, operating on the cycle of Fig. 1, was patented in 1872 by G. B. Brayton and was the first successful gas engine built in the United States. The Brayton cycle, with its high inherent thermal efficiency, requires the maximum volume of gas flow for a given power output. The Otto and diesel cycles require much lower gas flow rates, but have the disadvantage of higher peak pressures and temperatures. These conflicting elements led to many designs, all attempting to achieve practical compromises. With a piston and cylinder mechanism the

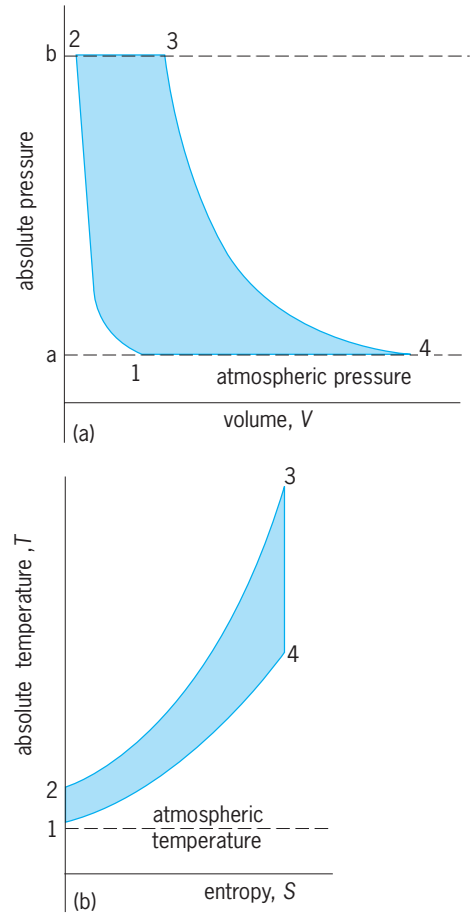

Fig. 1. Brayton cycle, air-card standard. (*a*) *p-V* cycle diagram. (*b*) *p-S* diagram. Phases: 1–2, compression; 2–3, heat addition; 3–4, expansion; and 4–1, heat abstraction.

Brayton cycle, calling for the maximum displacement per horsepower, led to proposals such as compound engines and variable-stroke mechanisms. They suffered overall disadvantages because of the low mean effective pressures. The positive displacement engine consequently preempted the field for the Otto and diesel cycles.

With the subsequent development of fluid acceleration devices for the compression and expansion of gases, the Brayton cycle found mechanisms which could economically handle the large volumes of working fluid. This is perfected today in the gas turbine power plant. The mechanism (**Fig. 2**) basically is a steady-flow device with a centrifugal or axial

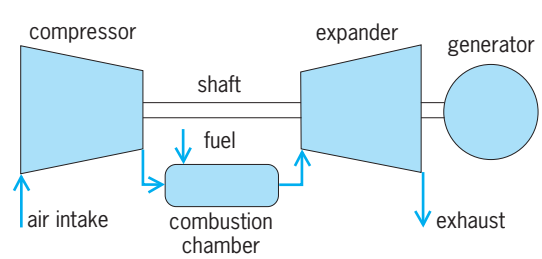

Fig. 2. Simple, open-cycle gas-turbine plant.

compressor, a combustion chamber where heat is added, and an expander-turbine element. Each of the phases of the cycle is accomplished with steady flow in its own mechanism rather than intermittently, as with the piston and cylinder mechanism of the usual Otto and diesel cycle engines. Practical gas-turbine engines have various recognized advantages and disadvantages which are evaluated by comparison with alternative engines available in the competitive marketplace. *See* AUTOMOBILE; GAS TURBINE; INTERNAL COMBUSTION ENGINE.

The net power output P_{net}, or salable power, of the gas-turbine plant (Fig. 2) can be expressed as shown by Eq. (4), where W_e is the ideal power output of

$$P_{net} = W_e \times \text{eff}_e - \frac{W_c}{\text{eff}_c} \qquad (4)$$

the expander (area b34a, Fig. 1), W_c the ideal power input to the compressor (area a12b, Fig. 1), eff_e the efficiency of expander, and eff_c the efficiency of compressor. This net power output for the ideal case, where both efficiencies are 1.0, is represented by net area (shaded) of the *p-V* cycle diagram of Fig. 1*a*. The larger the volume increase from point 2 to point 3, the greater will be the net power output for a given size of compressor. This volume increase is accomplished by utilizing the maximum possible temperature at point 3 of the cycle.

The difference in the two terms on the right-hand side of Eq. (4) is thus basically increased by the use of maximum temperatures at the inlet to the expander. These high temperatures introduce metallurgical and heat-transfer problems which must be properly solved.

The efficiency terms of Eq. (4) are of vital practical significance. If the efficiencies of the real compressor and of the real expander are low, it is entirely possible to vitiate the difference in the ideal powers W_e and W_c, so that there will be no useful output of the plant. In present practice this means that for adaptations of the Brayton cycle to acceptable and reliable gas-turbine plants, the engineering design must provide for high temperatures at the expander inlet and utilize high built-in efficiencies of the compressor and expander elements. No amount of cycle alteration, regeneration, or reheat can offset this intrinsic requirement for mechanisms which will safely operate at temperatures of about 1500°F (816°C). *See* CARNOT CYCLE; DIESEL CYCLE; OTTO CYCLE; THERMODYNAMIC CYCLE.

Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; J. B. Jones and G. A. Hawkins, *Engineering Thermodynamics*, 1960; H. Cohen et al., *Gas Turbine Theory*, 2d ed., 1975; R. T. Harman, *Gas Turbine Engineering: Applications, Cycles, and Characteristics*, 1981; G. N. Hatsopoulos and J. H. Keenan, *Principles of General Thermodynamics*, 1965, reprint 1981; M. J. Zucrow and J. D. Hoffman, *Gas Dynamics*, vol. 1, 1976, vol. 2, 1977.

Brazil nut

A large broad-leafed evergreen tree, *Bertholettia excelsa*, that grows wild in the forests of the Amazon valley of Brazil and Bolivia. The fruit is a spherical capsule, 3–6 in. (7.5–15 cm) in diameter and weighing 2–4 lb (0.9–1.8 kg) which, when mature, consists of an outer hard indehiscent husk about $^1/_2$ in. (1.2 cm) thick enclosing an inner hard-shelled container or pod filled with about 20 rather triangular seeds or nuts. *See* LECYTHIDALES.

Although there are a few plantations in Brazil, almost the entire production is gathered from wild trees. The quantity that gets into commercial channels varies with the price paid on the world market. The nuts are gathered either by local people who live in the forest or migrants who come into the area for the harvest season. The ripened capsules fall to the ground, and because of their weight and the height of the trees are a hazard to the harvesters. The fruits are carried to the stream bank and transported in small boats to local huts, where the capsules are split open to free the nuts. These are sold to traders who ship them downstream to brokers. They are mostly exported to Europe, Canada, and the United States. About one-fourth of the crop is shelled in Brazil before export.

Brazil nuts have a high oil and protein content and require careful handling and refrigeration to prevent spoilage. The nuts are used in confectionery, baked goods, and nut mixtures. *See* NUT CROP CULTURE.

Laurence H. MacDaniels

Brazing

A group of welding processes in which coalescence is produced by heating to suitable temperatures above 450°C and by using a filler metal that must have a liquidus temperature above 450°C and below the solidus temperature of the base metal. The filler metal is distributed between the closely fitted surfaces of the joint by capillary attraction. Brazing is distinguished from soldering in that the latter employs a filler metal having a liquidus below 450°C. *See* SOLDERING.

Brazing provides advantages over other welding processes, especially because it permits joining of dissimilar metals that, because of metallurgical incompatibilities, cannot be joined by traditional fusion processes. Since base metals do not have to be melted to be joined, it does not matter that they have widely different melting points. Therefore, steel can be brazed to copper as easily as steel to steel.

Brazing also generally produces less thermally induced distortion (warping) than fusion welding does, because an entire part can be brought up to the same brazing temperature, thereby preventing the kind of localized heating that can cause distortion in welding.

Finally, and perhaps most important to the

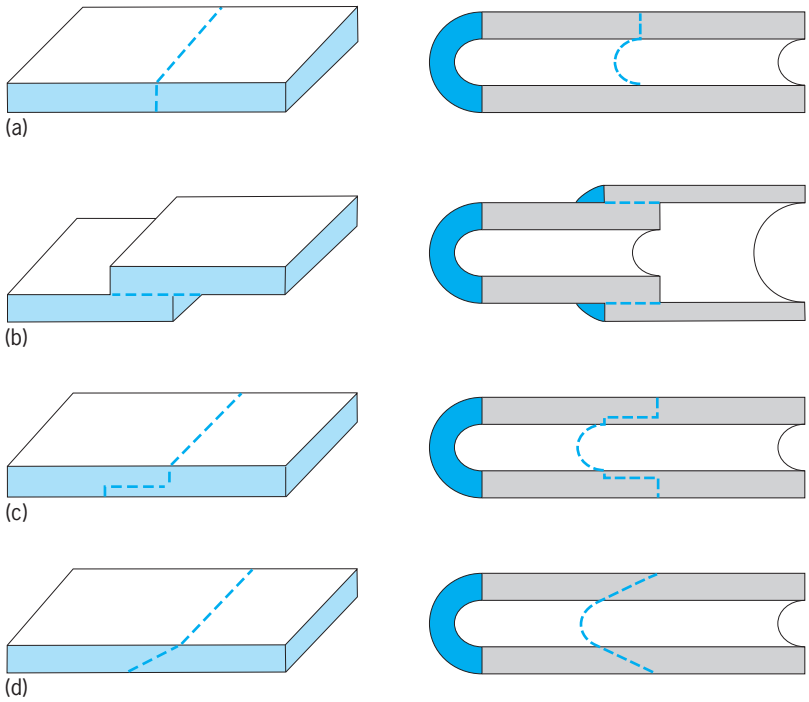

Fig. 1. Types of brazed joints, showing flat and tubular configurations. (*a*) Butt. (*b*) Lap. (*c*) Butt-lap. (*d*) Scarf.

manufacturing specialist, brazing readily lends itself to mass-production techniques.

Brazing process. The important elements of the brazing process are filler-metal flow, base-metal and filler-metal characteristics, surface preparation, joint design and clearance, temperature and time, and rate and source of heating.

A brazed joint is a heterogeneous assembly that is composed of different materials with different physical and chemical properties. In the simplest case, it consists of the base-metal parts to be joined and the added brazing filler metal. Diffusion processes, however, can change the composition and therefore the chemical and physical properties of the boundary zone formed at the interface between base metal and filler metal.

Small clearances should be used because the smaller the clearance, the easier it will be for capillarity to distribute the brazing filler metal throughout the joint area, and there will be less likelihood that voids or shrinkage cavities will form as the brazing filler metal solidifies. Small clearances and correspondingly thin filler-metal films make sound joints. The soundest joints are at least as high in tensile strength as the filler metal itself, and often higher. Clearances ranging from 0.03 to 0.08 mm are designed for the best capillary action and greatest joint strength.

Joint design. A number of factors must be considered in designing joints. First is composition and strength of the brazing filler metal. Generally, the bulk strength of the brazing filler metal is lower than that of the base metals, and so a correctly designed joint is required to obtain adequate mechanical strength. Second is capillary attraction. Because brazing depends on the principle of capillary attraction for distribution of molten brazing filler metal, joint clearance is a critical factor. Third is flux and air displacement. Not only must the filler metal be drawn into the joint, but flux and air must be displaced from it. Fourth is type of stress. In general, it is preferred that any load on a brazed joint be transmitted as shear stress rather than as tensile stress. Fifth is composition and strength of the base metals. In a joint made according to recommendations between high-strength members, the filler-metal film in the joint may actually be stronger than the base metal itself.

There are basically only two types of brazed joints, butt and lap, with all other joints being only modifications.

Butt and lap joints. The butt joint (**Fig. 1***a*) has the advantage of a single thickness at the joint. Preparation is relatively simple; however, the strength of any joint depends in part on the bonding area available. The thinnest member, therefore, dictates the maximum strength of the joint.

Strength is only one reason for the use of the lap design. Lap joints (Fig. 1*b*) can be readily designed to be self-jigging or, as in the case of tubing, self-aligning. Also, preplaced filler metal can be held in position better with such joints.

Butt-lap and scarf joints. The butt-lap joint (Fig. 1*c*) is an attempt to combine the advantage of a single thickness with maximum bonding area and strength. The scarf joint (Fig. 1*d*) represents another attempt to increase the cross-sectional area of the joint without increasing its thickness. Because the scarf joint is at an angle to the axis of tensile loading, its load-carrying capacity is similar to that of the lap joint and greater than that of the butt joint.

Joint clearance. The single most important design consideration in achieving good brazements is joint clearance—the distance between the faying surfaces to be joined. The ideal clearance for production work is frequently cited as 0.05–0.13 mm. However, some metals actually require interference fits, whereas others require clearances as great as 0.25 mm.

Selection of materials. Selection involves consideration of both the base metal and the filler metal.

Base metals. The base metal has a prime effect on joint strength. A high-strength base metal produces joints of greater strength than those made with softer base metals. In addition to the base-metal effects and the normal mechanical requirements of the base metal in the brazement, the effect of the brazing cycle on the base metal and the final joint strength must be considered. The brazing cycle by its very nature will usually anneal the cold-worked base metal, unless the brazing temperature is very low and the time at heat is very short. When it is essential to design a brazement having strength above the annealed strength of the base metals after the brazing operation, it will be necessary to specify a heat-treatable base metal.

Braze filler metals. The brazing filler metal is also known as the brazing alloy. A specific brazing filler metal cannot be chosen to produce a specific joint strength. Actually, strong joints can be brazed with

almost any commercial brazing filler metal if brazing methods and joint design are done correctly.

The degree to which brazing filler metal penetrates and alloys with the base metal during brazing is referred to as diffusion. In applications requiring strong joints for high-temperature, high-stress service conditions (such as turbine rotor assemblies and jet components), it is generally good practice to specify a brazing filler metal that has high diffusion and solution properties with the base metal. When the assembly is constructed of extremely thin base metals (as in honeycomb structures and some heat exchangers), good practice generally calls for a brazing filler metal with a low-diffusion characteristic relative to the base metal.

In choosing a brazing filler metal, the first criterion is the working temperature. Very few brazing filler metals possess distinct melting points. A brazing filler metal may have to melt below the temperature at which members of the joint lose strength or above the temperature at which oxides are reduced or dissociated. Joining may have to be carried out above the solution treatment temperature of the base metal, at a structure refining temperature, or below the remelt temperature of a brazing filler metal used previously in producing a brazed subassembly.

Methods of heating. There are a variety of heating methods available for brazing. Effective capillary joining requires efficient transfer of heat from the heat source into the joint. The seven commonly used methods are torch, furnace, induction, dip, resistance, infrared, and diffusion brazing.

Torch brazing. Manual torch brazing is most frequently used for repairs, one-of-a-kind brazing jobs, and short production runs as an alternative to fusion welding. Any joint that can be reached by a torch and brought to brazing temperature can be readily brazed by this technique. The adjustment of the torch flame is important. Generally, a slightly reducing flame is desirable.

In manual torch brazing, the brazing filler metal is usually face fed in the form of wire or rod, or preplaced. To ensure uniform heating throughout the joint, an important consideration, it may be advisable to use a multiple-tip torch or more than one torch.

Torch brazing can be automated relatively easily. Usually, such systems involve multiple-station rotary indexing tables. The part is fed into a holding fixture at the first station, then indexed to one or more preheating stations, depending on the heating time required. A brazing station is next, followed by a cooling station and an ejection station (**Fig. 2**).

Furnace brazing. Furnace brazing is used extensively where the parts to be brazed can be assembled with the brazing filler metal preplaced or in the joint. There are two basic types of furnaces used for brazing: the batch and the continuous. Furnace brazing is particularly applicable for high-production brazing in which continuous-conveyor-type furnaces are used. For medium-production work, batch-type furnaces are best. When continuous-type furnaces are used, several different temperature zones may be used to provide the proper preheating, brazing, and cooling temperatures. Heating is usually produced by electrical resistance.

The parts should be self-jigging or fixtured and assembled with brazing filler metal preplaced near or in the joint. The preplaced filler metal may be in the form of wire, foil, powder, paste, slugs, or preformed shapes. Fluxing is used except when a reducing atmosphere, such as hydrogen gas, can be introduced into the furnace. Pure, dry, inert gases, such as argon and helium, are used to obtain special atmospheric properties.

A large volume of furnace brazing is performed in vacuum, which prevents oxidation and often eliminates the need for flux. Vacuum brazing has found wide application in the aerospace and nuclear fields, where reactive metals are joined or where entrapped fluxes would be intolerable.

Induction brazing. The high-frequency induction heating method for brazing is clean and rapid, lends itself to close control of temperature and location, and requires little operator skill. The heat for induction brazing is created by a rapidly alternating electric current that is induced into the workpiece by an adjacent coil. The workpiece is placed in or near a coil carrying alternating current, which induces the heating current in the desired area. Brazing filler metal is normally preplaced in the joint, and the brazing can be done in air, in an inert atmosphere, or in a vacuum. Induction brazing is well suited for mass production, and mechanized systems for moving the assemblies to and from the coils are very common.

Dip brazing. This process involves immersion of assembled parts into a suitable molten bath to effect brazing. The bath can be either molten brazing filler metal or molten chemical flux. The molten-flux method is used extensively for brazing aluminum and its alloys. The brazing filler metal is preplaced and the assembly immersed in the flux bath, which has been raised to brazing temperature. The flux bath provides excellent protection against reoxidation of the metal, which can occur quite easily with aluminum.

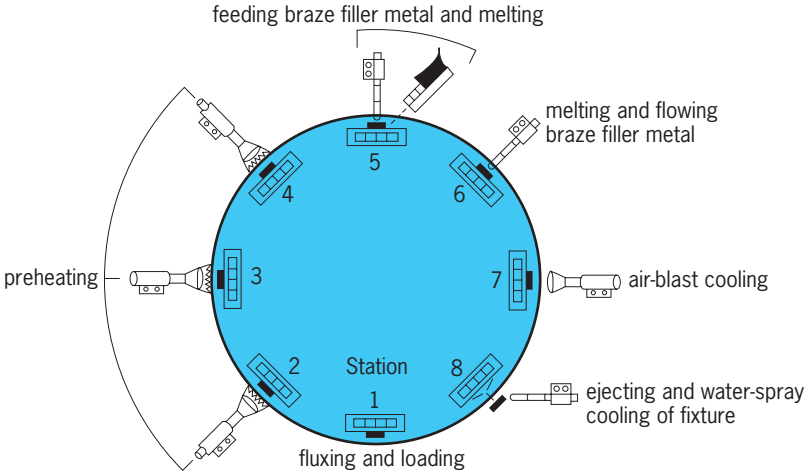

Fig. 2. Automatic eight-station torch brazing machine.

Resistance brazing. The process is most applicable to relatively simple joints in metals that have high electrical conductivity. The workpieces are heated locally, and brazing filler metal that is preplaced between the workpieces is melted by the heat obtained from resistance to the flow of electric current through the electrodes and the work. In the usual application of resistance brazing, the heating current, which is normally alternating current, is passed through the joint itself. Equipment is the same as that used for resistance welding.

Infrared brazing. The development of high-intensity quartz lamps and the availability of suitable reflectors have made infrared heat a commercially important generator for brazing. Lamps are often arranged in a toasterlike configuration, with parts traveling between two banks of lamps.

Diffusion brazing. This process is defined not by the method of heating but by the degree of mutual filler-metal solution and diffusion with the base metal resulting from the temperature used, and the time interval at heat. Temperature, time, in some cases pressure, and selection of base and filler materials are so controlled that the filler metal is partially or totally diffused into the base metal. For example, a nickel filler metal with a solidus temperature of 960°C, when partially diffused at a temperature of 1095–1150°C, will have a remelt temperature in excess of 1370°C.

Brazing environment. Fluxes, gas atmospheres, and vacuum conditions promote formation of brazed joints. They may be used to surround the work, exclude reactants, and provide active or inert protective atmospheres, thus preventing undesirable reactions during brazing.

Fluxes. The primary purpose of brazing fluxes is to promote wetting of the base metal by the brazing filler metal. The efficiency of flux activity, which is commonly referred to as wetting, can be expressed as a function of brazeability. The flux must be capable of dissolving any oxide remaining on the base metal after it has been cleaned, and also any oxide films on the liquid filler metal. To effectively protect the surfaces to be brazed, the flux must completely cover and protect them until the brazing temperature is reached. It must remain active throughout the brazing cycle.

Many chemical compounds are used in the preparation of fluxes. When a flux is heated, reactions take place between the various chemical ingredients, forming new compounds at brazing temperatures that are quite different chemically and physically from the unreacted constituents. Composition of the flux must be carefully tailored to suit all the factors of the brazing cycle.

Controlled atmospheres. The second way to control the formation of oxides during brazing and also reduce oxides present after precleaning is to surround the braze area with an appropriate controlled atmosphere. Like fluxes, controlled atmospheres are not intended to perform primary cleaning for removal of oxides, coatings, grease, oil, dirt, or other foreign materials.

In controlled-atmosphere applications postbraze cleaning is generally not necessary. Controlled atmospheres for brazing fall into three broad categories: reducing atmospheres, inert atmospheres, and vacuum conditions. They are used extensively for high-temperature brazing.

Cleaning. Cleaning of all surfaces that are involved in the formation of the desired brazed joint is necessary to achieve successful and repeatable braze joining. The presence of contaminants on one or both surfaces may result in formation of voids, restriction or misdirection of filler-metal flow, and inclusion of contaminants within the solidified braze area.

Inspection techniques. Inspection of the completed assembly or subassembly is the last step in the brazing process and is essential for ensuring satisfactory and uniform quality of the brazed unit. This operation also provides a means of evaluating the adequacy with which the prior steps in the process have been carried out with regard to ultimate integrity of the brazed joint.

The inspection methods chosen to evaluate the brazing procedure and the serviceability of the product will be largely dependent on the service requirements of the brazed assembly.

Applications. A myriad of applications of brazing occur in many fields, including automotive and aircraft designs; engines and engine components; electron tubes, vacuum equipment, and nuclear components, such as production of reliable ceramic-to-metal joints; and miscellaneous applications such as fabrication of food-service dispensers (scoops) used for ice cream and corrosion-resistant and leakproof joints in stainless steel blood-cell washers. *See* WELDING AND CUTTING OF MATERIALS.

Mel M. Schwartz

Bibliography. American Society for Metals, *Welding, Brazing, and Soldering*, vol. 6, 10th ed., 1994; American Welding Society, *Brazing Manual*, 4th ed., 1992; M. Schwartz, *Brazing for the Engineering Technologist*, 1995; M. Schwartz, *Joining Composite Materials*, 1994.

Breadboarding

Assembling an electronic circuit in the most convenient manner on a board or other flat surface, without regard for final locations of components, to prove the feasibility of the circuit and to facilitate changes when necessary. Standard breadboards for experimental work are made with mounting holes and terminals closely spaced at regular intervals, so that parts can be mounted and connected without drilling additional holes. Such breadboards are useful only for relatively simple circuits operating at audio or low radio frequencies.

Printed-circuit boards having similar patterns of punched holes, with various combinations of holes connected together by printed wiring on each side, are often used for breadboarding moderate radio frequency and digital logic circuits. *See* PRINTED CIRCUIT.

Advances in the capabilities of computer software for circuit simulation, component layout, and

interconnect routing, along with "rapid prototyping" capabilities provided by a number of vendors, have made it cost-effective for circuit designers to bypass conventional breadboarding and proceed directly to the construction of printed-circuit board prototypes of their designs. For even relatively simple designs, these prototypes can be delivered much more quickly than a breadboard can be constructed. These prototypes also have most, if not all, of the board-related idiosyncrasies that would be seen in production systems, so their performance will closely approximate that to be expected from the final product. *See* CIRCUIT (ELECTRONICS); COMPUTER-AIDED CIRCUIT DESIGN; PROTOTYPE.

Philip V. Lopresti; John Markus

Breadfruit

The multiple fruit of an Indo-Malaysian tree, *Artocarpus altilis*, of the mulberry family (Moraceae), now cultivated in tropical lowlands around the world (see **illus.**). Captain William Bligh was bringing plants from Tahiti to America when the mutiny on the *Bounty* occurred. The perianth, pericarp, and receptacle are all joined in a multiple fruit having a high carbohydrate content. The fruits vary considerably in size and are often borne in small clusters. Breadfruit is a wholesome food for both humans and animals. It is eaten fresh or baked, boiled, roasted, fried, or ground up and made into bread. There are many varieties, both with and without seeds.

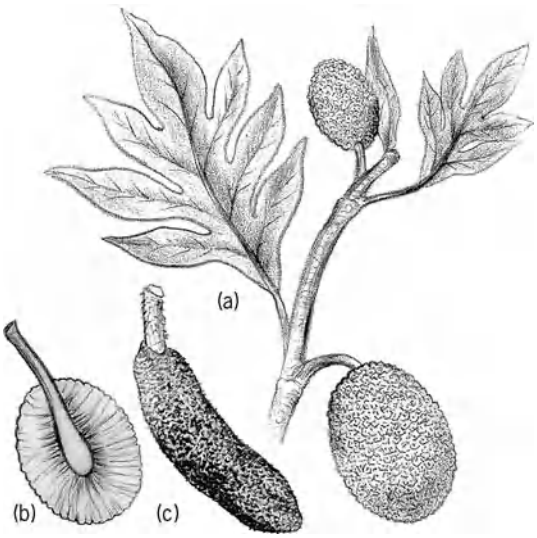

Breadfruit (*Artocarpus altilis*). (*a*) End of branch with multiple fruits. (*b*) Longitudinal section of multiple fruit. (*c*) Staminate flower cluster.

Earl L. Core

Breakdown potential

The potential difference at which an electrically stressed gas is tranformed from an insulator to a conductor. In an electrically stressed gas, as the voltage is increased, the free electrons present in the gas gain energy from the electric field. When the applied voltage is increased to such a level that an appreciable number of these electrons are energetically capable of ionizing the gas, the gas makes the transition from an insulator to a conductor; that is, it breaks down. The potential difference at which this transition occurs is known as the breakdown potential for the particular gaseous medium.

The breakdown potential depends on the nature, number density, and temperature of the gas; on the material, state, and geometry of the electrodes; on the type of voltage applied (steady, alternating, impulsive); and on the degree of preexisting ionization. Areas of surface roughness at the electrodes (especially the cathode) or the presence of conducting particles in the gas greatly reduces the breakdown potential because at such points the electric field is significantly enhanced, increasing the electron energies and thus gas ionization. The breakdown voltage varies considerably from one gaseous medium to another; it is very low for the rare gases, such as neon, for which a 0.394-in. (1-cm) uniform-field gap at 1 atm (101.3 kilopascals) breaks down at 0.53 kV; and very high for polyatomic, especially electronegative, gases such as sulfur hexafluoride (SF_6), for which a 0.394-in. (1-cm) uniform-field gap at 1 atm (101.3 kPa) breaks down at 88.9 kV.

The transition of a gas from an insulator to a conductor under an imposed electrical potential occurs in times ranging from milliseconds to nanoseconds, depending on the form of the applied field and the gas density. This transition depends on the behavior of electrons, ions, and photons in the gas, especially the processes which produce or deplete free electrons. Knowledge of these processes often allows prediction of the breakdown voltage of gases and the tailoring of gas mixtures which can withstand high electrical potentials for practical uses. *See* ELECTRICAL BREAKDOWN; ELECTRICAL CONDUCTION IN GASES.

Increasing the breakdown potential. Fundamental research on the microscopic physical processes which take place in an electrically stressed gas—especially those involving the interaction of slow (less than about 20 eV) electrons with atoms and molecules—has led to a better understanding of the electrical "breakdown" of gases. It has also led to the realization that an effective way of increasing the breakdown potential is to prevent the free electrons in the gas from initiating the breakdown by attaching them to gas molecules forming negative ions, which—being much heavier than the electrons—do not ionize the gas. (In such cases the breakdown potential can be defined as that value of the applied potential for which the rate of electron production by ionization is equal to the rate of electron depletion by attachment.) The ability of molecules to capture free electrons is large in the extreme low-energy range (less than about 1 eV) and decreases as the electron energy increases. For this reason, the gas itself—or an additive to it, as in a gas mixture—must be capable of reducing the energies of the free electrons and effectively scattering them into the

low-energy region, where they can be captured most efficiently. Knowledge of the electron-scattering properties of gases permits selection of suitable gases for the slowing down of free electrons in dielectric gas mixtures. Combinations of two or more electron-attaching and electron-slowing-down gases have been developed in this manner which effectively control the energies and number densities of free electrons in electrically stressed gases. This approach allows the systematic identification of unitary gases (for example, sulfur hexafluoride and perfluorocarbons) and the systematic development of multicomponent gaseous dielectrics (for example, sulfur hexafluoride–nitrogen; perfluorocarbon–nitrogen; and sulfur exafluoride–perfluorocarbon–nitrogen) which are able to withstand very high potentials before breaking down and are suitable for many applications.

Applications. The systematic development of gaseous dielectrics with high dielectric strength (that is, high breakdown potential) is most significant for high-voltage technology, which has a multiplicity of gas insulation needs. Dielectric gases are widely used as insulating media in high-voltage transmission lines, circuit breakers, transformers, substations, high-voltage research apparatus, and other electrical equipment (such as Van de Graaff accelerators). In transmission lines the high-voltage conductor is held in place by solid insulators and is enclosed in a vacuum-tight pipe, the earthed electrode (often buried underground), which is filled with the insulating gas at pressures usually ranging from 1 to 10 atm (100 to 1000 kPa) depending on the voltage. The voltages in such systems can be higher than 1,000,000 V, resulting in considerable reduction of resistive losses and hence substantial energy savings. The use of compressed-gas–insulated transmission lines and other gas-insulated high-voltage electrical equipment is expected to continue increasing. *See* CIRCUIT BREAKER; DIELECTRIC MATERIALS; TRANSMISSION LINES. L. G. Christophorou

Bibliography. L. G. Christophorou (ed.), *Gaseous Dielectrics VIII*, 1998; L. G. Christophorou, D. R. James, and R. Y. Pai, Gas breakdown and high voltage insulating gases, in *Applied Atomic Collisions*, vol. 5, pp. 87–167, 1982; E. E. Kunhardt and L. H. Luessen (eds.), *Electrical Breakdown and Discharges in Gases*, vols. 89a and 89b, NATO Advanced Science Institutes Series, 1983; J. M. Meek and J. D. Craggs (eds.), *Electrical Breakdown in Gases*, 1978.

Breast

The human mammary gland, usually well developed in the adult female but rudimentary in the male. Each adult female breast contains 15–20 separate, branching glands that radiate from the nipple. During lactation their secretions are discharged through separate openings at the base of the nipple.

In both sexes paired ridges (milk lines) develop in the 6-week-old fetus and extend from the axilla downward, converging toward the abdomen. Most portions of the milk lines regress, leaving only slight prominences in the region destined to become the breasts. During the middle part of fetal life clusters of epidermal cells grow inwardly in these areas, forming cords which develop into hollow ducts. After birth slow increase in size and branching of each primary duct occurs. Development in the male usually ends at this stage. Supernumerary nipples or breasts along the milk lines are not rare. They are of minor clinical significance because they are usually rudimentary, but may constitute a severe cosmetic or psychic problem, which is readily amenable to surgery.

In the female, hormonal changes in adolescence cause enlargement of breast tissue, but much of this is connective tissue although some glandular buds form. With the advent of full menstruation ovarian estrogenic hormones influence breast development. If pregnancy ensues, the glandular tissue reaches full development and full lactation begins shortly after birth. After cessation of lactation the breasts regress considerably and once again reflect cyclic regulation. *See* MENSTRUATION; PREGNANCY.

Occasionally the male breasts will be stimulated by estrogens, possibly of adrenal origin. Liver disease may prevent normal destruction of estrogen so that gynecomastia (enlargement of breasts in the male) results.

The breasts of infants may respond to maternal hormones during the first few days or weeks of life, but regression usually occurs spontaneously. *See* LACTATION. Walter Bock

Breast disorders

Benign and malignant lesions of the human mammary glands. Benign breast disorders are often symptomatic and bothersome but do not have malignant potential. Malignant disorders have the potential to grow locally in the breast and spread through the bloodstream to other parts of the body.

Physiologic changes. The breast is an organ that changes in response to fluctuations in hormone levels. Physiologic changes in the breast are often confused with disease. Fibrocystic disease, chronic cystic mastitis, and mammary dysplasia are terms that have been used to describe cyclical pain, tenderness, and lumpiness in the breast. These terms are imprecise and represent the normal physiologic responses to hormonal changes in the body rather than distinct clinical entities. Cyclic breast pain can occur in response to estrogens endogenous in premenopausal women or supplemental in postmenopausal women. Diffuse, palpable irregularities or lumps in the breast are also associated with this cyclical pain. Unlike malignant masses, however, these irregularities fluctuate in size and tenderness with the menstrual cycle, and are better described as physiologic nodularity of the breast. *See* MENOPAUSE.

Benign disorders. Common benign disorders include breast masses, cysts, gynecomastia, nipple

discharge, and breast infections. Breast masses or dominant lumps are different from lumpiness; they are persistent over time and are palpably distinct from the surrounding breast tissue. They can develop in any age group and should be carefully evaluated. Mammography and ultrasound can help to determine the character of the mass, depending on the age of the patient.

Dominant lumps include fibroadenomas, gross cysts, pseudolumps, and cancer. Fibroadenomas may occur in any age group but are most commonly seen in young women. They are benign tumors consisting of smooth, rounded masses that are easily palpable in the breast. The cause is unknown; however, there is evidence to support the presence of an imbalance in circulating hormone levels that might be responsible for tumor growth. Breast cysts may be difficult to identify by physical examination or mammography, but can be distinguished from solid masses by using ultrasound imaging. Treatment of a cyst involves draining it with a small needle and syringe. If the cyst completely disappears, no further treatments are required. However, if the cyst remains or recurs or the cyst fluid is bloody, further examinations are required to rule out an underlying carcinoma. *See* MAMMOGRAPHY.

Gynecomastia, a benign enlargement of the male breast which can occur at any age and in one or both breasts, is a physiologic response to hormones, drugs, or an underlying medical condition. Nipple discharge does not always indicate a pathologic process; the character of the discharge is significant. A watery bilateral discharge from multiple ducts is usually normal. A milky discharge (galactorrhea) is often physiologic but sometimes can be associated with a tumor of the pituitary gland, which secretes prolactin. Unilateral, and especially spontaneous, nipple discharge generally signifies underlying pathology. Bloody nipple discharge is of most concern. About 80% of the time, however, the underlying cause is a benign papilloma within a duct. Breast infections are a common problem seen in both lactating and nonlactating individuals. *See* LACTATION.

Proliferative breast disorders include ductal and lobular hyperplasias. It is thought that some of these disorders might represent precancerous changes since they signal an increased risk for the development of breast cancer. Women who have hyperplasia without atypical cell changes have a mildly elevated risk for the subsequent development of breast cancer when they are compared with the general population. Women with atypical hyperplasia have a risk of developing breast cancer 4.4 times that of women without identifiable risk factors. This lesion is seen in only about 4–10% of breast biopsies. It does not characteristically form lumps or show up on mammography. Other benign neoplasms of the breast include adenomas, intraductal papillomas, adenosis, and radial sclerosing lesions.

Precancer. Ductal carcinoma is a precancer. It does not have the ability to disseminate throughout the body but can progress to an invasive carcinoma if left untreated. Once discovered, this lesion can be treated with breast conservation surgery, that is, excision of the tumor with a margin of normal tissue, with or without radiation, or with total mastectomy. Lobular carcinoma, however, is a misnomer since it is not a premalignant lesion but a marker for subsequent cancer. Women with lobular carcinoma who subsequently develop cancer do so in either breast with a relative risk 5.7 times that of the general population. Treatment options include close observation with physical examination and mammography, or bilateral mastectomies.

Cancer. Breast cancer is the most common cancer in women and is the second leading cause of cancer deaths among all American women, particularly in the fifth and sixth decades of life. Almost 80% of invasive carcinomas of the breast are of ductal origin; the remainder are lobular carcinomas or other special histologic types. Invasive ductal carcinomas can be subtyped if they have one or more characteristics of a specific histologic type, including tubular, medullary, papillary, or mucinous differentiation. When a large part of the tumor is differentiated into one of these subtypes, they generally carry a more favorable prognosis.

There are no significant differences in the rates of local recurrence and survival when breast conservation surgery is combined with radiation therapy to the breast as compared with total mastectomy. Removal of axillary lymph nodes continues to be an important tool in the staging and prognosis of invasive carcinomas of the breast. If axillary lymph node metastases are detected, systemic therapy is indicated to decrease the incidence of distant metastases, and will usually decrease the risk of mortality by one-third. Combination chemotherapy is generally recommended for premenopausal women, and its use has been extended to healthy postmenopausal women. Hormonal therapy is used to treat postmenopausal women with axillary lymph node metastases and hormone receptor positive tumors. Tamoxifen (an antiestrogen) is a hormonal agent which has been shown to decrease the risk of recurrence and the development of second primary tumors in postmenopausal women. *See* CANCER (MEDICINE); CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS.

Thirty percent of all women with negative lymph nodes will have developed micrometastases at the time of diagnosis and will eventually die of breast cancer. Chemotherapy or hormone therapy is therefore often recommended even for women with negative nodes. The use of chemotherapy or hormone therapy in women without axillary lymph node metastases is determined by multiple factors. The most important determinant is tumor size. Women with tumors greater than 2 cm (0.8 in.) in diameter, with negative axillary nodes, should receive systemic therapy. For individuals without axillary lymph node metastases and tumors less than 1 cm (0.4 in.) in diameter, the probability of relapse 10 years after diagnosis is less than 10%. Therefore, systemic

chemotherapy is generally not recommended in this group.

Screening. Breast cancer screening is the most effective way of detecting breast cancer in its early stages. Screenings involve self-examination of the breast, physical examination, and mammography. The guidelines for when these examinations should be administered depend on the individual's age. A mammogram is the most reliable screening test for the early detection of breast cancer; mammograms are recommended every 2–3 years from age 40 to 50 and annually after age 50. As the density of the breast parenchyma changes with increasing age, mammography is better able to detect abnormalities within the breast. Kelly K. Hunt; Susan M. Love

Bibliography. I. M. Ariel and A. C. Cahan, *Treatment of Precancerous Lesions and Early Breast Cancer: Diagnosis and Management*, 1993; J. R. Harris et al., *Breast Diseases*, 1991; D. F. Hayes (ed.), *Atlas of Breast Cancer*, 1993; F. A. Tavassoli, *Pathology of the Breast*, 2d ed., 1999; A. K. Tucker (ed.), *Textbook of Mammography*, 1993.

Breccia

A clastic rock composed of angular gravel-size fragments; the consolidated equivalent of rubble. The designation gravel-size refers to a mean particle diameter greater than 0.08 in. (2 mm), which means that 50% or more of the particles (by volume) are this size or larger. Various classifications specify different values for the degree of angularity. One system specifies angular or subangular fragments (roundness ≤ 0.25), whereas another restricts the term breccia to aggregates with angular fragments (roundness ≤ 0.10). *See* GRAVEL.

Most classifications recognize three major types of breccia: sedimentary, igneous, and cataclastic. Sedimentary breccias are characterized as exogenic, that is, formed by processes at the surface of the Earth, while igneous and cataclastic breccias are characterized as endogenic, that is, formed within the Earth. A minor exogenic type is impact breccia, which is formed by the impact of extraterrestrial bodies, such as meteorites and comets. This type is rare on Earth, but it is abundant on the Moon, Mercury, and other heavily cratered planetary bodies.

Sedimentary breccia. These breccias, referred to as sharpstone conglomerates according to one classification, are significant because the angularity of their fragments indicates either proximity to the source or transportation by a mechanism that does not cause significant rounding of the fragments. Examples of the first condition are talus breccia formed at the base of a scarp, and reef breccia deposited adjacent to a reef margin. Transport mechanisms that can preserve the angularity of clasts over significant distances include debris flows, slumps, and glacial transport, although rounded fragments may also be carried. All of these mechanisms incorporate a large proportion of fine sediment in the transporting medium, which effectively cushions interparticle collisions and inhibits rounding. *See* CONGLOMERATE; REEF; SEDIMENTOLOGY.

Intraformational or intraclastic breccias are an important class of sedimentary breccias. They are formed by the breakup and incorporation of sediment aggregates from within the same formation, which requires either early cementation (for example, the formation of nodules or duricrusts) or uncemented aggregates sufficiently cohesive to be transported a short distance without disaggregation. Thus, uncemented aggregates are basically limited to sediments that are rich in clay or clay-size carbonates (calcilutites). The mechanisms for formation of intraformational breccias include bank slumping or desiccation fracturing of mud in river or tidal channels, and erosion and incorporation of mud blocks in mass flows such as slumps or turbidity currents. *See* SEDIMENTARY ROCKS.

Igneous breccia. These breccias are mainly of pyroclastic origin but may also form as intrusive breccias by forceful intrusion of magma. In the latter case the operative agent is fluid pressure; in the former it is the explosive escape of gas from solidifying viscous lava. These rocks, termed pyroclastic or volcanic breccias, are distinct from agglomerates, which accumulate mainly as lava bombs solidified during flight and which are commonly rounded. *See* PYROCLASTIC ROCKS.

Cataclastic breccia. These breccias result from the fracture of rocks by tectonic or gravitational stresses. However, since many tectonic processes are at least partly gravitational, the two processes can be considered together. Tectonic breccias include fault and fold breccias, the latter formed by fracturing of brittle layers within incompetent plastic strata during folding. In one classification, landslide and slump breccias are included in the gravitational category, but here they are considered to be sedimentary, commonly intraformational. Solution or collapse breccias are a type of nontectonic gravitational breccia. They result from the creation by ground-water solution of unsupported rock masses which collapse under their own weight to form breccia Brian Rust

Bibliography. H. Blatt, *Sedimentary Petrology*, 2d ed., 1991; W. C. Krumbein and F. J. Pettijohn, *Manual of Sedimentary Petrography*, 1938, reprint 1988; D. W. Lewis and D. M. McConchie, *Practical Sedimentology*, 2d ed., 1993.

Breeding (animal)

The application of genetic principles to improving heredity for economically important traits in domestic animals. Examples are improvement of milk production in dairy cattle, meatiness in pigs, growth rate in beef cattle, fleece weight in sheep, and egg production in chickens. Even after thousands of years of domestication, domestic animals respond readily to genetic selection. Selection permits the best parents to leave more offspring in the next generation than do genetically inferior parents. Many specialized breeds and strains have been developed for the

production of meat, fiber, and milk that are adapted to different environmental and economic conditions. *See* GENETICS.

The science of animal breeding is based on genetic principles; however, the current application of these principles relies heavily on statistical knowledge. Accurate measurement of performance of individuals, such as growth rate or milk production, is essential. Use of computers has become an integral part of animal breeding research and application. Modern computers allow the application of complicated statistical procedures to large volumes of performance data to identify the best individuals to become parents of the next generation.

Methods of genetic selection. Selection is the primary tool for generating directed genetic changes in animals. Selection may be (1) concentrated on one characteristic, (2) performed independently on several traits, or (3) conducted on an index or total score which includes information on several traits. The first method is most effective in generating improvement in the one chosen primary trait, but may generate a correlated response in other traits. The second method, independent culling levels, generates greater overall improvement in total merit than the first, but is less effective for making change in the primary trait. The third method, index selection, generates maximum improvement in several traits by weighting each trait by its relative economic importance and by the phenotypic and genetic relationships among the traits. In some species, index selection has the disadvantage of delaying selection decisions until measurements, taken at different times in the life of the animals, are completed on all traits.

In general, the third method is preferable when several important heritable traits need attention. In practice, selection is likely to be a mixture of the second and third methods.

Factors influencing genetic selection progress. Heritability, the fraction of the total variation in a trait that is due to additive genetic differences, is a key parameter in making selection decisions. Traits for which environmental effects cause most of the variation are only slightly heritable. Most economically important traits are strongly to moderately influenced by environmental or management conditions. Therefore, managing animals to equalize environmental influences on them, or statistically adjusting for environmental differences among animals, is necessary to accurately choose those with the highest genetic merit for various traits.

The accuracy with which the breeding values of the animals can be determined is directly related to the improvement achieved by selection. Accuracy, in turn, depends upon the heritabilities of the traits and whether they can be measured directly upon the subjects of selection (mass selection), their parents (pedigree selection), their brothers and sisters (family selection), or their progeny (progeny testing). Mass selection is effective for traits with high heritability which are expressed before breeding age. Pedigree selection is most useful for traits such as milk production which are expressed relatively late in life and for traits which are limited in expression to only one sex. Pedigree selection can never be highly accurate as compared with other kinds of selection because a subject can have only two parents, a maximum of four different grandparents, and so on, whereas the number of siblings or progeny can be large. For traits of medium heritability, the following sources of information are approximately equally accurate for predicting breeding values of animals: (1) one record measured on the animal itself; (2) one record on each ancestor for three previous generations; (3) one record each on five brothers or sisters where there is no environmental correlation between family members; and (4) one record each on five progeny having no environmental correlations, each from a different mate.

The intensity of genetic selection influences the rate of genetic improvement within a population of animals. As older females are removed (culled) from the breeding population, younger females (called replacements) must be added to the breeding herd in order to maintain herd size. If, for example, only the top 20% of females are chosen from the pool of potential replacements, their average level of genetic merit will be higher than instances when the top 50% of potential females are saved for replacements (this, of course, assuming that accuracy of selection is high).

The generation interval, or the length of time between the birth of the parents and the birth of the offspring which replace the parents in the breeding herd, alters the rate of genetic improvement. Typically, breeding males are replaced more frequently than breeding females, particularly in breeding programs where replacement females are generated from within the herd. One of the challenges faced by breeders of domestic animals is to attain the proper balance between rapid genetic improvement (attained by short generation intervals) and the economic sustainability of the herd (frequent purchase of breeding sires and/or replacement females costs money).

Purebred strains. Propagation of improved animal seed stocks is achieved primarily with purebred strains descended from imported or locally developed groups or breeds of animals which have been selected and interbred for a long enough period to be reasonably uniform for certain trademark characteristics, such as coat color. Each breed is promoted and sponsored by its farmer breeders organized as a purebred society. Because the number of breeding animals is finite and because breeders tend to prefer certain bloodlines and sires, some inbreeding occurs within the pure breeds, but this has not limited productivity in most breeds.

As a rule, each breed has certain easily identifiable characteristics such as coat color, size and shape of ear, and horns or polledness. Most of these breed-specific traits are encoded by one gene (or a group of linked genes), and these traits are called qualitative (or Mendelian) traits. The particular form, or allele, of each gene is almost identical among all animals within breeds that have been subjected to intense

selection for many generations. Conversely, other traits are controlled by numerous genes, each with a relatively small effect, and are known as quantitative (multigenic) traits. Traits such as growth rate or milk production are influenced by many genes, as are those concerned with appetite, ability to obtain and digest food, temperament, and energy metabolism. The effects of individual alleles upon such traits are likely to be small relative to the total variation. Hence, selection for improved performance changes the frequency of any one of the numerous desired alleles only slightly, although the cumulative improvement over many generations may be large. Selection has not always been consistent for the same goal, frequently changing intensity or direction as a consequence of changing styles or economic conditions. *See* MENDELISM.

Crossbreeding. One of the main reasons for crossbreeding is to make use of the genetic phenomenon of heterosis. Heterosis is improved performance of crossbred progeny, exceeding the average performance of their parents. Most commercial pigs, sheep, and beef cattle are produced by crossbreeding.

There are two basic types of crossbreeding systems—rotational and terminal. Rotational crossbreeding systems involve the rotation of different breeds of sires for mating of the herd females, and replacement females are chosen from the offspring produced within the herd. Rotational systems commonly utilize two or three different breeds, although four breeds may also be used. Terminal crossbreeding systems involve the mating of a sire breed, most often noted for superior meat quality, with herd females noted for maternal traits such as fertility, mothering ability, and/or milk production. All offspring produced are sent to harvest; hence the adjective "terminal." The herd females may be purebreds, but typically they are crossbred because crossbred females exhibit heterosis for reproductive traits. By mating a paternal sire breed with a female that excels in maternal traits, the resultant offspring should possess a combination of outstanding genetic characteristics not found in a single breed. Breeds should be chosen which complement one another. *See* GENETICS.

New developments. There have been new discoveries in four areas that have application, or potential application, for improvement of domestic animals: quantitative genetics, reproductive physiology, molecular genetics, and immunogenetics.

Quantitative genetics. Quantitative geneticists use statistical and genetic information to improve domestic animals. Typically a statistical procedure is used to rank animals based on their estimated breeding values for traits of economic importance. The breeding values are based on records of the animals themselves and on the records of their relatives. The animals that are ranked highest are used as parents of the next generation. The recording of animal performance, ranking the animals, and mating are repeated for successive generations of selection. The use of information on relatives allows more accurate prediction of the breeding values of the animals that are the subjects of selection. The statistical procedures used allow the ranking of animals across herds or flocks, provided the animals in different herds or flocks have relatives in common. This procedure allows for genetic comparisons of animals raised in different environments.

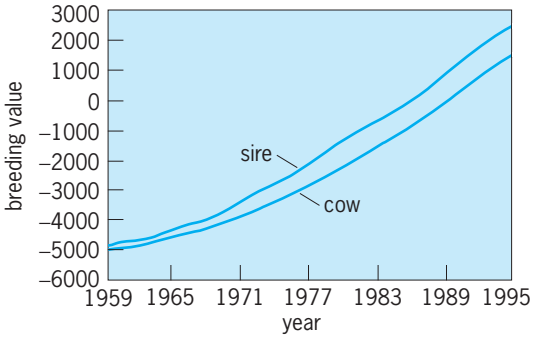

Fig. 1. Holstein average breeding value by birth year for milk. An organized breeding program in dairy cattle results in a genetic gain. The increase in breeding value of cows lags the gains made by sires; sires are more highly selected, and it takes a generation to pass the gain from a sire to offspring.

The statistical procedure called an animal model is used to compute breeding values for all dairy cattle in the United States. The well-organized breeding program in dairy cattle has resulted in a genetic gain (**Fig. 1**). In 1995–1999, the genetic trend in Holstein sires was an increase of 195.6 lb (88.9 kg) of milk, 6 lb (2.7 kg) of protein, and 5.8 lb (2.6 kg) of fat per cow per year. Breeding values are expressed relative to a base that is updated every 5 years. The actual production of Holsteins is 20,342 lb (9246 kg) per cow per year.

Reproductive physiology. The primary contribution of reproductive physiology to genetic improvement is development of techniques to increase the intensity of genetic selection and to reduce the generation interval. If genetic improvement is increasing at the same rate per generation, but generations are shorter, more genetic gain per unit of time can be attained.

The most important development of reproductive physiologists was artificial insemination, which allows extensive use of superior males. Artificial insemination allows one male to impregnate many more females than with natural mating. Artificial insemination has been used most widely in dairy cattle, where thoroughly planned young sire sampling programs are combined with carefully conducted progeny testing schemes to identify genetically superior sires whose semen is used to impregnate 70% of the dairy cows. These superior dairy sires produce many thousands of offspring. In the Holstein breed, approximately 500 sires produce enough semen to impregnate approximately 6,300,000 cows. Artificial insemination is used in all domestic animals in the United States. Its use is increasing rapidly in beef cattle, but currently less than 15% are bred artificially. The production of pork is increasingly done in large commercial operations. As a consequence, the use

of artificial insemination has increased dramatically. More than 50% of sows are now artificially inseminated. In chickens, both broilers for meat and layers for eggs, artificial insemination is used for matings in the elite seedstock companies; in turkeys, nearly all matings are conducted artificially. Artificial insemination is rarely practiced in sheep, but its use is growing in the horse industry as breed association regulations are being revised to allow it.

Another significant development was embryo transplantation, which allows more extensive use of genetically superior females. Embryo transplantation is nonsurgical. Females are treated hormonally to stimulate the release of several ova at a time; subsequently the superovulated cows are inseminated. The embryos (fertilized eggs) are flushed from the female and are implanted into recipient females or frozen for later use. The usual number of embryos recovered per flush varies from none to approximately 20, with 6 being average.

A fascinating contribution made by reproductive physiologists to the field of animal breeding is the ability to produce genetically identical individuals. Identical (monozygotic) twins have been reported in virtually all domestic animals, but the frequency of naturally occurring identical twins is low. In the 1970s, attempts to produce identical twins were made by separating the individual cells of a preimplantation embryo (called blastomeres), culturing them, and subsequently transferring them to recipient females. The task was laborious, time-consuming, and inefficient. In the 1980s, the technology of embryo splitting (or bisection) was pioneered. With this technology, morula or blastocyst stage preimplantation embryos were split into halves to produce identical twins. Although this technology is well developed and is available commercially to livestock producers, an embryo may be bisected only once, limiting the number of genetically identical animals that can be produced from a single embryo to two.

After blastomere separation techniques were developed, the technology of nuclear transfer began to evolve. Most people refer to nuclear transfer methodology as cloning . The earliest form of nuclear transfer used blastomeres as the source of genetic material for cloning. Isolated blastomeres were injected into enucleated oocytes (oocytes [or eggs] whose own genetic material [chromosomes, which contain DNA] has been removed), and the blastomere–oocyte complex was subjected to pulses of electricity (electrofusion) to cause fusion of the cells and activation of cell division. The resultant "cloned" embryos were cultured in vitro until they could be transferred into recipient females. Alternatively, the cloned embryos could undergo blastomere separation and a second round of nuclear transfer, a process known as serial nuclear transplantation.

In 1997, a new method of nuclear transfer was introduced with the announcement of Dolly the sheep (**Fig. 2**). Dolly was produced via a technique known as somatic cell nuclear transfer. A somatic cell is essentially any cell of the body other than a sex cell (sperm or egg). Scientist Ian Wilmut from the Roslin Institute near Edinburgh, Scotland, led a research team that took a cell from the mammary gland of an adult sheep, cultured that cell in a specially devised medium, injected the cultured cell into an enucleated oocyte, and subjected the somatic cell–enucleated oocyte complex to electrofusion. Transfer of that embryo led to the birth of Dolly. Somatic cell nuclear transfer in some regards was very similar to nuclear transfer using blastomeres as the donor cell, but in other ways it was fundamentally different because it used a differentiated cell (from the mammary gland) rather than an undifferentiated cell (a blastomere).

Although some people oppose somatic cell nuclear transfer due to religious or ethical beliefs, there have been some issues associated with the procedure that have concerned even the most enthusiastic scientist. Some of these problems have already been overcome by changes in methodology, but other issues remain. One lingering concern is that of shortened telomeres. Telomeres are a part of the chromosome which plays an integral role in cell division. When telomeres become too short, the cell cannot divide properly and dies. The shortening of telomeres occurs naturally, and many believe this is one of the major underlying causes of aging. Many animals produced via somatic cell nuclear transfer have exhibited signs of "premature" telomere shortening.

A second remaining challenge to the efficiency of somatic cell nuclear transfer is genomic imprinting. Genomic imprinting is a phenomenon where only one allele (either the maternal or the paternal) is expressed instead of two. Researchers have identified a number of genes that ordinarily are imprinted during development, including those regulating the development of the placenta and fetal growth. There

Fig. 2. Dolly and her first-born lamb, Bonnie. (*Courtesy of the Roslin Institute*)

is abundant documentation in cloned embryos and offspring produced via somatic cell nuclear transfer technology that the genetic reprogramming of the somatic cell is not complete. Many cloned embryos exhibit either an aberrant pattern of genomic imprinting or no imprinting in some genes where it should occur. This incomplete genetic reprogramming of somatic cells leads to a major reduction in the efficiency and reliability of somatic cell nuclear transfer.

The newest approach to nuclear transfer, called chromatin transfer, is being investigated as a potential means of overcoming some of the shortcomings of the traditional somatic cell nuclear transfer approach. In this procedure, somatic cells are permeabilized to enable disassembly of nuclei in a mitotic extract, and the resultant condensed chromatin is transferred into enucleated oocytes. This is followed by electrofusion, culture, and transfer of the cloned embryo to a recipient female. Initial reports of this approach are encouraging, as survival of cloned calves to one month of age is enhanced. Whether this approach circumvents problems with telomere shortening and/or genomic imprinting remains to be seen.

Molecular genetics. Four bases (adenine, thymine, cytosine, and guanine) make the genetic code, DNA, that determines heredity. Segments of DNA that are usually several hundred bases long make up the genes that control cellular functions and, of course, the traits of economic importance. There are many thousands of genes in humans and some other species, and knowledge is increasing regarding the location of genes on the chromosomes. Knowledge of which genes control traits of economic importance would be very useful to animal breeders.

In domestic animals, polymorphisms (changes in the order of the four bases) that are discovered in the DNA may be associated with economic traits. When the polymorphisms are associated with or directly code for economic traits, they are called quantitative trait loci (QTLs). When few QTLs are known that control a portion of the variability in a trait, increasing the frequencies of favorable alleles can enhance the accuracy of selection and augment production.

One of the ways in which animal breeders use knowledge of DNA sequences is via marker-assisted selection. Marker-assisted selection is the genetic selection of animals that is assisted by the knowledge of the specific nucleotide (DNA) sequence of a QTL. Marker-assisted selection is not the selection of an animal based solely on its genetic makeup at the QTL; instead, animals are selected using a combination of "traditional" information (such as breeding values estimated from performance records) and genetic marker DNA sequence. The reason for this combined approach is to ensure that an otherwise genetically inferior animal is not selected just because it possesses a desirable DNA sequence at one QTL. In the cattle industry, commercially available DNA tests help breeders select animals for increases in milk production, in marbling (intramuscular fat that lends greater flavor and palatability), and in meat tenderness. *See* DEOXYRIBONUCLEIC ACID (DNA).

Another use of molecular genetics is to detect the genes that code for genetic defects. One example is the bovine leukocyte deficiency gene, which does not allow white blood cells to migrate out of the blood supply into the tissues to fight infection. Affected calves perish at a young age. Screening all sires that enter artificial breeding organizations (and not using sires that transmit the defect) has effectively controlled this condition. Another example is determining the DNA sequence of the prion protein gene, as this is related to a sheep's susceptibility to scrapie disease. Scrapie is a transmissible spongiform encephalopathy similar to bovine spongiform encephalopathy (or "mad cow disease"). *See* SCRAPIE.

Immunogenetics. Immunogenetics is the science of genetic control of the immune system. All domestic animals have diseases that must be controlled in order for production of food to be safe and profitable. Information is accumulating about genetic control of the immune system in domestic animals, but there is much to be learned. Study of the immune system relies heavily on molecular techniques.

There is excellent evidence in domestic and laboratory animal species demonstrating some degree of genetic control of at least a part of the immune system. A group of genes that code for disease resistance are known as the major histocompatibility complex (MHC). The MHC genes (broken into classes I, II, and III) encode for proteins displayed on cell surfaces and for a group of disease-fighting proteins known as complement. The genes play a role in resistance and susceptibility to disease. Understanding these complexes of genes is expected to offer a genetic mechanism to aid in disease control. Difference among sire progeny groups has demonstrated disease resistance in several species. It is clear that there is some degree of genetic control in diseases, and the question is how to use genetic mechanisms to improve the health of domestic animals. Natural selection over many centuries has allowed the animals that could resist a variety of diseases to survive. Breeders of domestic animals must be careful to maintain resistance to a wide variety of diseases. *See* AGRICULTURAL SCIENCE (ANIMAL); GENETIC ENGINEERING; REPRODUCTIVE SYSTEM.

C. R. Youngs; A. E. Freeman

Bibliography. D. S. Falconer, *Introduction to Quantitative Genetics*, 2d ed., 1981; E. Li, Chromatin modification and epigenetic reprogramming in mammalian development, *Nat. Rev. Genet.*, 3:662–673, 2002; M. Lynch and B. Walsh, *Genetics and Analysis of Quantitative Traits*, Sinauer Associates, Sunderland, MA, 1997; E. J. Sullivan et al., Cloned calves from chromatin remodeled in vitro, *Biol. Reprod.*, 70:146–153, 2004; X. C. Tian et al., Meat and milk composition of bovine clones, *Proc. Nat. Acad. Sci. USA*, 102:6261–6266, 2005; L. D. Van Vleck, E. J. Pollack, and E. A. B. Oltenacu, *Genetics for the Animal Sciences*, 1987.

Breeding (plant)

The application of genetic principles to the improvement of cultivated plants, with heavy dependence upon the related sciences of statistics, pathology, physiology, and biochemistry. The aim of plant breeding is to produce new and improved types of farm crops or decorative plants, to better serve the needs of the farmer, the processor, and the ultimate consumer. New varieties of cultivated plants can result only from genetic reorganization that gives rise to improvements over the existing varieties in particular characteristics or in combinations of characteristics. In consequence, plant breeding can be regarded as a branch of applied genetics, but it also makes use of the knowledge and techniques of many aspects of plant science, especially physiology and pathology. Related disciplines, like biochemistry and entomology, are also important, and the application of mathematical statistics in the design and analysis of experiments is essential. *See* GENETICS.

Scientific Method

The cornerstone of all plant breeding is selection. By selection the plant breeder means the picking out of plants with the best combinations of agricultural and quality characteristics from populations of plants with a variety of genetic constitutions. Seeds from the selected plants are used to produce the next generation, from which a further cycle of selection may be carried out if there are still differences. Much of the early development of the oldest crop plants from their wild relatives resulted from unconscious selection by the first farmers. Subsequent conscious acts of selection slowly molded crops into the forms of today. Finally, since the early years of the century, plant breeders have been able to rationalize their activities in the light of a rapidly expanding understanding of genetics and of the detailed biology of the species studies.

Conventional breeding is divided into three categories on the basis of ways in which the species are propagated. Species that reproduce sexually and that are normally propagated by seeds occupy two of these categories. First come the species that set seeds by self-pollination, that is, fertilization usually follows the germination of pollen on the stigmas of the same plant on which it was produced. The second category of species sets seeds by cross-pollination, that is, fertilization usually follows the germination of pollen on the stigmas of different plants from those on which it was produced. The third category comprises the species that are asexually propagated, that is, the commerical crop results from planting vegetative parts or by grafting. Consequently, vast areas can be occupied by genetically identical plants of a single clone that have, so to speak, been budded off from one superior individual. The procedures used in breeding differ according to the pattern of propagation of the species. While conventional methods involving crossbreeding have been used very successfully for several decades, several innovative techniques have been explored to enhance the scope, speed, and efficiency of producing new, superior cultivars. Advances have been made in extending conventional sexual crossing procedures by laboratory culture of plant organs and tissues and by somatic hybridization through protoplast fusion.

Self-pollinating species. The essential attribute of self-pollinating crop species, such as wheat, barley, oats, and many edible legumes, is that, once they are genetically pure, varieties can be maintained without change for many generations. When improvement of an existing variety is desired, it is necessary to produce genetic variation among which selection can be practiced. This is achieved by artificially hybridizing between parental varieties that may contrast with each other in possessing different desirable attributes. All members of the first hybrid (F_1) generation will be genetically identical, but plants in the second (F_2) generation and in subsequent generations will differ from each other because of the rearrangement and reassortment of the different genetic attributes of the parents. During this segregation period the breeder can exercise selection, favoring for further propagation those plants that most nearly match the breeder's ideal and discarding the remainder. In this way the genetic structure is remolded so that some generations later, given skill and good fortune, when genetic segregation ceases and the products of the cross are again true-breeding, a new and superior variety of the crop will have been produced.

This system is known as pedigree breeding, and while it is the method most commonly employed, it can be varied in several ways. For example, instead of selecting from the F_2 generation onward, a bulk population of derivatives of the F_2 may be maintained for several generations. Subsequently, when all the derivatives are essentially true-breeding, the population will consist of a mixture of forms. Selection can then be practiced, and it is assumed, given a large scale of operation, that no useful segregant will have been overlooked. By whatever method they are selected, the new potential varieties must be subjected to replicated field trials at a number of locations and over several years before they can be accepted as suitable for commercial use.

Another form of breeding that is often employed with self-pollinating species involves a procedure known as backcrossing. This is used when an existing variety is broadly satisfactory but lacks one useful and simply inherited trait that is to be found in some other variety. Hybrids are made between the two varieties, and the F_1 is crossed, or backcrossed, with the broadly satisfactory variety which is known as the recurrent parent. Among the members of the resulting first backcross (B_1) generation, selection is practiced in favor of those showing the useful trait of the nonrecurrent parent and these are again crossed with the recurrent parent. A series of six or more backcrosses will be necessary to restore the structure of the recurrent parent, which ideally should be modified only by the incorporation of the single useful attribute sought from the nonrecurrent parent.

Backcrossing has been exceedingly useful in practice and has been extensively employed in adding resistance to disease, such as rust, smut, or mildew, to established and acceptable varieties of oats, wheat, and barley. *See* PLANT PATHOLOGY.

Cross-pollinating species. Natural populations of cross-pollinating species are characterized by extreme genetic diversity. No seed parent is truebreeding, first because it was itself derived from a fertilization in which genetically different parents participated, and second because of the genetic diversity of the pollen it will have received. In dealing with cultivated plants with this breeding structure, the essential concern in seed production is to employ systems in which hybrid vigor is exploited, the range of variation in the crop is diminished, and only parents likely to give rise to superior offspring are retained.

Inbred lines. Here plant breeders have made use either of inbreeding followed by hybridization (see **illustration**) or of some form of selection. During inbreeding programs normally cross-pollinated species, such as corn, are compelled to self-pollinate by artificial means. Inbreeding is continued for a number of generations until genetically pure, truebreeding, and uniform inbred lines are produced. During the production of the inbred lines rigorous selection is practiced for general vigor and yield and disease resistance, as well as for other important characteristics. In this way desirable attributes can be maintained in the inbred lines, which are nevertheless usually of poor vigor and somewhat infertile. Their usefulness lies in the vigor, high yield, uniformity, and agronomic merit of the hybrids produced by crossing different inbreds. Unfortunately, it is not possible from a mere inspection of inbred lines to predict the usefulness of the hybrids to which they can give rise. To estimate their values as the parents of hybrids, it is necessary to make tests of their combining ability. The test that is used depends upon the crop and on the ease with which controlled cross-pollination can be effected.

Tests may involve top crosses (inbred × variety), single crosses (inbred × inbred), or three-way crosses (inbred × single cross). Seeds produced from crosses of this kind must then be grown in carefully controlled field experiments designed to permit the statistical evaluation of the yields of a range of combinations in a range of agronomic environments like those normally encountered by the crop in agricultural use. From these tests it is possible to recognize which inbred lines are likely to be successful as parents in the development of seed stocks for commercial growing. The principal advantages from the exploitation of hybrids in crop production derive from the high yields produced by hybrid vigor, or heterosis, in certain species when particular parents are combined.

Economic considerations. The way in which inbred lines are used in seed production is dictated by the costs involved. Where the cost of producing F_1 hybrid seeds is high, as with many forage crops or with sugarbeet, superior inbreds are combined into a synthetic strain which is propagated under conditions of open pollination. The commercial crop then contains a high frequency of superior hybrids in a population that has a similar level of variability to that

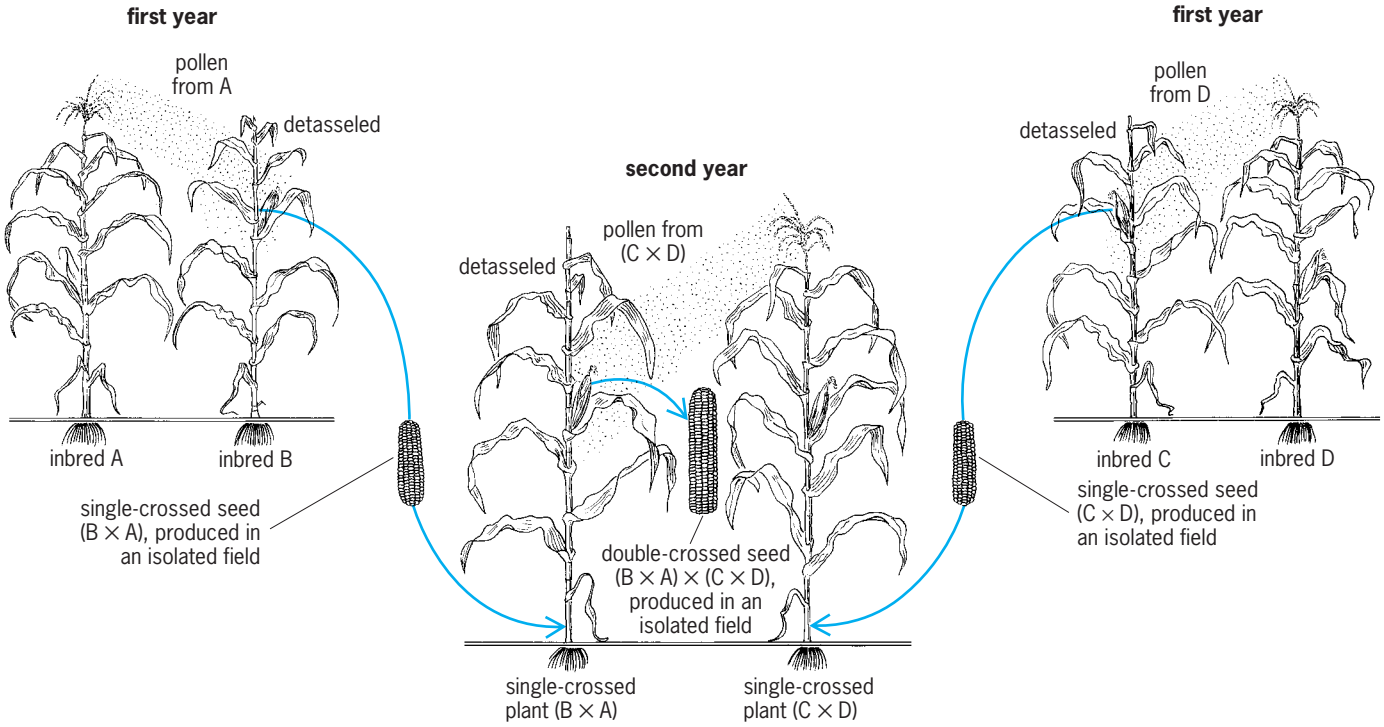

Sequence of steps in crossing inbred plants and using the resulting single-crossed seed to produce double-crossed hybrid seed. (*Crops Research Division, Agricultural Research Service, USDA*)

of an open-pollinated variety. However, because of the selection practiced in the isolation and testing of the inbreds, the level of yield is higher because of the elimination of the less productive variants.

When the cost of seed is not of major significance relative to the value of the crop produced, and where uniformity is important, F_1 hybrids from a single cross between two inbred lines are grown. Cucumbers and sweet corn are handled in this way. By contrast, when the cost of the seeds is of greater significance relative to the value of the crop, the use of single-cross hybrids is too expensive and then double-cross hybrids (single cross A × single cross B) are used, as in field corn. As an alternative to this, triple-cross hybrids can be grown, as in marrow stem kale, in the production of which six different inbred lines are used. The commerical crop is grown from seeds resulting from hybridization between two different three-way crosses.

Recurrent selection. Breeding procedures designated as recurrent selection are coming into limited use with open-pollinated species. In theory, this method visualizes a controlled approach to homozygosity, with selection and evaluation in each cycle to permit the desired stepwise changes in gene frequency. Experimental evaluation of the procedure indicates that it has real possibilities. Four types of recurrent selection have been suggested: on the basis of phenotype, for general combining ability, for specific combining ability, and reciprocal selection. The methods are similar in the procedures involved, but vary in the type of tester parent chosen, and therefore in the efficiency with which different types of gene action (additive and nonadditive) are measured. A brief description is given for the reciprocal recurrent selection.

Two open-pollinated varieties or synthetics are chosen as source material, for example, A and B. Individual selected plants in source A are self-pollinated and at the same time outcrossed to a sample of B plants. The same procedure is repeated in source B, using A as the tester parent. The two series of testcrosses are evaluated in yield trials. The following year inbred seed of those A plants demonstrated to be superior on the basis of testcross performance are intercrossed to form a new composite, which might be designated A_1. Then B_1 population would be formed in a similar manner. The intercrossing of selected strains to produce A_1 approximately restores the original level of variability or heterozygosity, but permits the fixation of certain desirable gene combinations. The process, in theory, may be continued as long as genetic variability exists. In practice, the hybrid $A_n \times B_n$ may be used commercially at any stage of the process if it is equal to, or superior to, existing commercial hybrids.

Asexually propagated crops. A very few asexually propagated crop species are sexually sterile, like the banana, but the majority have some sexual fertility. The cultivated forms of such species are usually of widely mixed parentage, and when propagated by seed, following sexual reproduction, the offspring are very variable and rarely retain the beneficial combination of characters that contributes to the success of their parents. This applies to such species as the potato, to fruit trees like apples and pears that are propagated by grafting, and to raspberries, grapes, and pineapples.

Varieties of asexually propagated crops consist of large assemblages of genetically identical plants, and there are only two ways of introducing new and improved varieties. The first is by sexual reproduction and the second is by the isolation of sports or somatic mutations. The latter method has often been used successfully with decorative plants, such as chrysanthemum, and new forms of potato have occasionally arisen in this way. When sexual reproduction is used, hybrids are produced on a large scale between existing varieties with different desirable attributes in the hope of obtaining a derivative possessing the valuable characters of both parents. In some potato-breeding programs many thousands of hybrid seedlings are examined each year. The small number that have useful arrays of characters are propagated vegetatively until sufficient numbers can be planted to allow the agronomic evaluation of the potential new variety.

Ralph Riley

Special Techniques

Breeding for new, improved varieties of crop plants is generally based on cross-pollination and hybrid production. Such breeding is limited to compatible plants, that is, plants that permit cross-pollination, fertilization of the egg cell, and development of an embryo and seedling. Compatibility lessens with increasing distance in the relationship between plants. Breeding would benefit from access to traits inherent in sexually noncompatible plants. For instance, it is desirable to breed disease resistance from mustard (*Brassica nigra*) or oil quality of Meadow foam (*Limnanthes douglasii*) into rapeseed (*B. napus*). Special techniques, collectively referred to as biotechnology, are employed to overcome incompatibility barriers. Cell biology, molecular genetics, and biological chemistry are the scientific disciplines which nurture the development of special breeding techniques.

True-to-type reproduction of plant stock is another goal of plant breeding. Traditionally it is accomplished by cuttings, tubers, and bulbs, that is, asexually. Such technique applies, for example, to propagation of sugarcane, banana, potato, or lilies. Cell technologies have been used to extend the range and efficiency of asexual plant propagation. The process is often referred to as micropropagation. *See* PLANT PROPAGATION.

Cell technologies. The regeneration of entire, fertile or mature plants from single cells excised from a source plant and cultured in a nutrient medium in an incubator constitutes the primary goal of plant cell culture. Cell development leads either to shoot and root formation or to embryo (as in seeds) and plantlet formation. Cell technologies underlie all of the following special techniques. *See* TISSUE CULTURE.

Micropropagation and cloning. Cell technologies have been applied to mass propagation of a great

number of species and varieties; particularly in the realm of horticulture. Conventionally, tissues rather than single cells are excised from root, stem, petiole, or seedling and induced to regenerate plantlets. All regenerants from tissues of one source plant constitute a clone. Nutrient formulas and growth conditions are specific for cells and tissues of each plant variety. Once sufficient growth has been achieved, plantlets are acclimatized to greenhouse and field conditions on location, or via air freight close to market. The Boston fern was one of the first plants to leave micropropagation labs and arrive in stores. High-priced ornamentals (for example, orchids) have been prime targets of micropropagation. Mass propagation of nursery stock has been aided by devices for automated processing (robotics).

Micropropagation has been used to eliminate virus from contaminated stock and is being practiced in quarantine stations. This technology has also been implemented in the restoration of natural habitats (propagation of *Cypripedium*) and of species on the brink of extinction. It complements other techniques whenever seed production of prototype plants is doubtful and asexual means of propagation are called for.

Microspore or anther culture. A technique of propagation particularly useful for plant breeders is the generation of plants from cells with but one set of chromosomes, haploid cells, as occurs in the development of pollen, referred to as microspore culture. Microspores are isolated from anthers and cultured on nutrient media, or entire anthers are cultured in this manner. Given properly conditioned source plants, a high-sugar nutrient medium, and growth conditions with specific temperature regimes, embryos may form directly or subsequent to callus formation. Doubling of chromosomes that may occur spontaneously or can be induced by treatment of callus, embryos, or plantlets with colchicine leads to the formation of homozygous di-haploid plants. Such material instantly stabilizes a (new) genotype, F-1plants, due to hybridization. *Brassica napus* (rapeseed), which responds to these techniques particularly well, enables several hundred microspores per macerated bud to develop into embryos. Anthers of certain barley varieties can yield up to 20 plantlets. Microspore culture relieves the plant breeder of several cycles of inbreeding to obtain pure lines. *See* POLLEN.

Cryopreservation. Regeneration of plants from cells and tissues opens new ways of preserving stock or germplasm. Treating pieces of tissue with cryoprotectants like sorbitol or dimethylsulfoxide followed by controlled freezing permits storage of such material in vials under low temperature (−280°F or −196°C) for years. Thereafter the tissues may be quickly thawed in a waterbath and, upon return to nutrient media, grown to plants. Shoot tips of a variety of horticultural species and of experimental plants are being preserved in this manner. Gene resources centers around the world have adopted cryopreservation as a standard practice. *See* CRYOBIOLOGY.

In vitro fertilization. Sexual incompatibility between plants can be caused by inadequate pollen tube growth and sperm cell development. Such problems could be overcome by in vitro fertilization. Compared to animal systems, in vitro fertilization in plants is compounded by the participation of the embryo sac, accessibility of the egg cell, and complementary function of auxiliary cells. Advances have led to the isolation of viable sperm cells from pollen tubes, the regeneration of plants from embryo sacs 1 day after fertilization, the electrofusion of sperm cells with isolated egg cells, and the injection of sperm cells into the embryo sac.

Embryo rescue. The excision and culture of embryos on nutrient media under in vitro conditions has become an important technique to overcome postfertilization problems, for example, deficiencies in the development of endosperm, the tissue that stores nutrients and surrounds the embryo. The younger the embryo, the more development depends on employing ovules rather than isolated embryos. The technique is used also to overcome natural seed dormancy or to accelerate seed germination. *See* SEED.

Somaclonal variation. Cells and tissues cultured on nutrient media may undergo spontaneous changes. Frequently these changes result in declining capacity for plant regeneration over time. Also, plants regenerated from cells cultured over several months may show increasing variation in phenotype. As a result, regenerated plants may become a pool for selection of plants with desirable traits. Introduction of such plants into a breeding program requires prior demonstration of the heritability of the variant trait by analysis of the progeny of the selected regenerant. A pool of di-haploid, homozygous plants with stable regenerants obtained by culturing microspores as in the case of rapeseed is preferred by breeders.

Somatic hybridization. Enzymatic removal of walls from cells of leaves and seedlings furnishes individual naked cells, that is, protoplasts. Naked cells permit the process of cell fusion. Exposure of protoplasts to electric current or to high concentrations of poly(ethylene glycol) induces adhesion followed by fusion of cell membranes. Similarity of membrane structure throughout the plant kingdom permits the fusion of distantly related protoplasts. Fusion may lead to nuclear fusion prior to and during nuclear division, resulting in amphi-diploid somatic hybrid cells. Such fusion products may divide. Fusion products of closely related, yet sexually incompatible plants have been grown to flowering plants; the most famous example is the potato + tomato hybrid = pomato (*Solanum tuberosum* + *Lycopersicon esculentum*). These plants did not set seeds. Rapeseed (*Brassica napus*) obtained by conventional breeding has been recreated by somatic hybridization using protoplasts from *B. campestris* and *B. oleracea* Furthermore, intergeneric fertile hybrids have been obtained via protoplast fusion of various rapeseed (*Brassica*) plants with *Eruca, Sinapis,* and *Diplotaxis* species. The difference between artificial cross-pollination and fertilization is the fusion not only

of nuclei but also of the cytoplasm with mitochondria and plastids of both parents. *See* PLANT CELL; SOMATIC CELL GENETICS.

DNA technologies. The development of technologies that enable the isolation of desirable genes from bacteria, plants, and animals (genes that confer herbicide resistance or tolerance to environmental stress, or encode enzymes and proteins of value to the processing industry) and the insertion of such genes into cells and tissues of target plants by direct or indirect uptake has led to the genetic transformation of plant cells. Desirable genes often are constructed to include regulatory deoxyribonucleic acid (DNA) sequences and genes which enable targeted expression, selection, and visualization of transformed cells and tissues.

Transformation technology began with the observation that nature practices this very process when crown-gall bacteria, which live in the soil, adhere to wounded tissue and release DNA with genes for hormone synthesis into neighboring nonwounded cells. The hormones subsequently drive the growth of the crown gall. In laboratories, this process is mimicked when the genes for hormones are excised from the extrachromosomal DNA of *Agrobacterium* and are replaced by desirable genes. Transformation of plant cells with this indirect method is widely used with plants which naturally respond to *Agrobacterium*, including dicotyledonous herbs and perennials, and trees like poplar. Cells and tissues of monocotyledonous plants (wheat, barley, corn, and rice) but also of conifers and a host of dicots are successfully transformed by an array of direct uptake and transformation processes, the most important of which is based on ballistic technology. Here, recombinant DNA is coated on microscopic particles of gold or tungsten, and these are introduced into cells under high velocity mediated either by gunpowder discharge or helium gas pressure. Initially, transformation success is measured by a response to products of selectable marker gene activity (such as resistance to antibiotics or pigmentation due to β-glucuronidase activity). *See* CROWN GALL; MOLECULAR BIOLOGY.

Assays for the expression of foreign genes in cells and tissues of target plants vary in technology: Southern blot analysis demonstrates the presence of the gene in a host genome; the Northern dot blot analysis, its transcription to ribonucleic acid (RNA); and the Western dot blot analysis, its translation into a gene product (protein or enzyme). The overall expression of a gene is regulated by associated genes that promote, enhance, or silence gene activity.

Transgenic plants. The regeneration of transformed plant cells and tissues results in new and novel genotypes to be assessed for transgenic phenotype. Contrary to hybrids obtained by cross-pollination, such plants are different from their parent by only one or two single, defined traits. Since the production of the first herbicide-resistant tobacco plant in 1985, transgenics of numerous species have successfully been grown and planted in field plots. Herbicide-resistant and insect pest–tolerant crops, both single-gene transgenics, have been entered into breeding programs and are close to commercialization.

Limitations to fast application of DNA technologies include the identification and isolation of useful genes, optimum combination with regulatory genes (promoters, enhancer sequences), the position of the transgene in the host genome, translation and final structure of the gene product, as well as survival of select transformed cells and refinement of nutrient and culture conditions for plant development and growth from these cells, avoiding of chimeric products. While DNA techniques generally add to or reinforce plant traits, suppression can also be valuable.

Trait changes based on the recombination of single genes represent state-of-the art biotechnology. Multigene traits such as yield or nitrogen fixation are too complex to manipulate and transfer. Still, genes can be introduced one at a time. Rapeseed plants with genes for male sterility and fertility restoration responding to different cues have been produced, and are being assessed for hybrid seed production.

DNA-based genetic diagnostics. Traditional plant breeding is based on selection of superior plants among segregating progeny of a cross. The selection is usually based on a visible or measurable phenotype (vigor, seed color, yield and disease resistance). Using DNA markers, which cosegregate with the genes of interest, allows for greater precision in breeding assessment. Analysis of DNA restriction fragment length polymorphisms (RFLPs) has become the method of choice. It is based on a variation in length of DNA fragments obtained after treatment of a given strand of DNA with restriction enzymes and visualized by electrophoresis. Molecular markers such as RFLPs have been tagged to specific agronomic traits such as disease resistance or male sterility. Generally, they are employed to map phenotypic traits, select superior plants, reduce the transfer of unwanted traits to hybrids (linkage drag), and fingerprint cultivars prior to patent application. The RFLP assay, however, is a laborious procedure, undesirable for breeding projects with high sample throughput. The random amplification of polymorphic DNA (RAPD) assay, however, is based on the use of short sequences of nucleotides as primers for the amplification through polymerase chain reaction of randomly selected segments of the target genome. RAPD markers have assisted in mapping a variety of traits with higher cost efficiency. The polymerase chain reaction is a technique for the enzymatic amplification of specific nucleotide sequences and results in an exponential increase in copies of such sequences. Apart from facilitating RAPDs, this technology promises to assist in reconstituting the genes of extinct species such as ancestors of present-day crop plants. *See* GENETIC ENGINEERING. Friedrich Constabel

Bibliography. A. M. R. Ferrie, C. E. Palmer, and W. A. Keller, *In Vitro Embryogenesis in Plants*, 1994; K. K. Kartha (ed.), *Cryopreservation of Plant Cells and Organs*, 1985; H. T. Stalker and J. P. Murphy, *Plant Breeding in the 1990s*, 1992; T. A. Steeves and I. M. Sussex, *Patterns in Plant Development*,

1989; I. K. Vasil (ed.), *Scale-up and Automation in Plant Propogation*, 1991; I. K. Vasil and T. A. Thorpe (eds.), *Plant Cell and Tissue Culture*, 1994.

Bremsstrahlung

In a narrow sense, the electromagnetic radiation emitted by electrons when they pass through matter. Charged particles radiate when accelerated, and in this case the electric fields of the atomic nuclei provide the force which accelerates the electrons. The continuous spectrum of x-rays from an x-ray tube is that of the bremsstrahlung; in addition, there is a characteristic x-ray spectrum due to excitation of the target atoms by the incident electron beam. The major energy loss of high-energy (relativistic) electrons (energy greater than about 10 MeV, depending somewhat upon material) occurs from the emission of bremsstrahlung, and this is the major source of gamma rays in a high-energy cosmic-ray shower. *See* COSMIC RAYS; ELECTROMAGNETIC RADIATION.

Properties. The spectrum of bremsstrahlung resulting from the collision of an electron with an atom is continuous and is roughly constant between $v = 0$ and $v = v_{\max}$; $v_{\max}$ is the maximum frequency of a photon which can be emitted; that is, $hv_{\max} = T$, where T is the initial kinetic energy of the electron and h is Planck's constant. The angular distribution of bremsstrahlung is roughly isotropic at low (nonrelativistic) electron energies, but is largely restricted to the forward direction at high energies. Very little bremsstrahlung is emitted at an angle much larger than $\theta_c = m_e c^2/T$ radians, where m_e is the electron mass and c is the velocity of light. Bremsstrahlung emitted at the angle θ_c is polarized with the electric vector perpendicular to the plane containing the direction of radiation and the incident electron velocity. It is difficult to observe the polarization, because it is small except near the angle θ_c. A longitudinally polarized electron (that is, one with its spin parallel to its velocity) emits circularly polarized bremsstrahlung; this effect is not sensitive to angle, and has proved useful in analysis of the longitudinal polarization of electrons emitted in beta decay.

Synchrotron radiation. In a broader sense, bremsstrahlung is the radiation emitted when any charged particle is accelerated by any force. To a great extent, as a source of photons in the ultraviolet and soft x-ray region for the investigation of atomic structure (particularly in solids), bremsstrahlung from x-ray tubes has been replaced by synchrotron radiation. Synchrotron radiation is an analog to bremsstrahlung, differing in that the force which accelerates the electron is a macroscopic (large-scale) magnetic field. Like bremsstrahlung, the synchrotron radiation spectrum is continuous and slowly varying. The main advantage of synchrotron radiation light sources is that the spectrum is accurately calculable and is uncontaminated by spectral lines from atomic transitions. In addition, they have a much higher efficiency, that is, a larger ratio of brightness to power input. This is partly because the electrons' only loss of energy is the emission of the synchrotron radiation, but mainly because the nuclear electric fields which accelerate the electrons to produce bremsstrahlung also scatter the electrons, thus broadening the angular distribution of subsequently emitted photons. *See* SYNCHROTRON RADIATION.

Nonelectronic bremsstrahlung. Because all other charged particles are much heavier than the electron, their accelerations are generally much smaller and so their bremsstrahlung is generally much weaker. But although nonelectronic bremsstrahlung is not a useful source of photons, its observation can be a useful indicator of the accelerations undergone in a particular process. An example is proton–proton scattering; the rate of the process $p + p \rightarrow p + p + \gamma$ yields information on the *p-p* nuclear force. *See* PLASMA (PHYSICS).

Charles Goebel

Bibliography. A. Bienenstock and H. Winnick, Synchrotron-radiation research: An overview, *Phys. Today*, 36(6):48–58, 1983; W. Heitler, *The Quantum Theory of Radiation*, 3d ed., 1954; J. D. Jackson, *Classical Electrodynamics*, 3d ed., 1998; E. M. Rowe and J. H. Weaver, The uses of synchrotron radiation, *Sci. Amer.*, 236(6):32–41, 1977; Special Issue on Synchrotron Radiation, *Phys. Today*, 34(5):27–71, 1983.

Brick

A construction material usually made of clay and extruded or molded as a rectangular block. Three types of clay are used in the manufacture of bricks: surface clay, fire clay, and shale. Adobe brick is a sun-dried molded mix of clay, straw, and water, manufactured mainly in Mexico and some southern regions of the United States. *See* CLAY; CLAY, COMMERCIAL.

The first step in manufacture is crushing the clay. The clay is then ground, mixed with water, and shaped. Then the bricks are fired in a kiln at approximately 2000°F (1093°C). A modern brick manufacturing plant includes a tunnel kiln that permits a nonstop firing process. Some older plants operate beehive kilns, which require periodic firing. After manufacture, the bricks are normally packaged into a steel-strapped cube with openings for handling by a forklift; this module generally contains 500 bricks. *See* REFRACTORY.

Substances in the clay such as ferrous, magnesium, and calcium oxides impart color to the bricks during the firing process. The color may be uniform throughout the bricks, or the bricks may be manufactured with a coated face. The latter are classified as glazed, claycoat, or engobe. Engobes are coatings, also called slurries, which are applied to plastic or dry body brick units to develop the desired color and texture. Claycoat is a type of engobe that is sprayed on as a coating of liquid clay and pigments.

Clay bricks are manufactured for various applications. In the United States the specifications are determined by the American Society for Testing and Materials (ASTM).

The most commonly used brick product is known as facing brick. In the United States the standard dimensions (modular brick size) are $3^5/_8$ in. × $2^1/_4$ in. × $7^5/_8$ in. (90 mm × 57 mm × 190 mm). However, 11 other brick sizes are used for specific applications.

In addition to standard bricks, decorative bricks molded in special shapes are available in both standard and custom sizes. They are used to form certain architectural details such as water tables, arches, copings, and corners. Bricks are also used to create sculptures and murals. Michael Gurevich

Bridge

A structure built to provide ready passage over natural or artificial obstacles, or under another passageway. Bridges serve highways, railways, canals, aqueducts, utility pipelines, and pedestrian walkways. In many jurisdictions, bridges are defined as those structures spanning an arbitrary minimum distance, generally about 10–20 ft (3–6 m); shorter structures are classified as culverts or tunnels. In addition, natural formations eroded into bridgelike form are often called bridges. This article covers only bridges providing conventional transportation passageways.

The longest single span provided by a bridge—the Humber suspension bridge in England—is 4626 ft (1410 m); the longest multiple-span bridge—the Lake Pontchartrain Causeway at New Orleans, Louisiana—126,055 ft (38,422 m).

History. Bridges undoubtedly have been built since the origin of humankind, perhaps first as trees felled over waterways and later as structures of timber or stone. The art of constructing stone bridges reached a high degree of development during the Roman era, and for a thousand years or so thereafter, it continued as an empirical art rather than a science. During the nineteenth century the theories of physics and mathematics were first applied to bridges in efforts to produce structures which would be rationally and economically proportioned to take the intended loads. During the midnineteenth century, with application of wrought iron as a material for construction (in the Brittania railway bridge in England), model testing and materials testing were initiated in a scientific manner. At this time, too, and continuing nearly until the end of the century, many firms in the United States developed proprietary bridges which were competitively peddled to railroads and governmental divisions. Eventually the failures of bridges, particularly railroad bridges, because of either faulty design or skimpy construction intended to lower cost, led in the 1880s to the establishment of consulting bridge engineering as a specialized discipline of civil engineering.

Parts. Bridges generally are considered to be composed of three separate parts: substructure, superstructure, and deck. The substructure or foundation of a bridge consists of the piers and abutments which carry the superimposed load of the superstructure to the underlying soil or rock. The superstructure is that portion of a bridge lying above the piers and abutments. The deck or flooring is supported on the bridge superstructure; it carries and is in direct contact with the traffic for which passage is provided.

Types. Bridges are classified in several ways. Thus, according to the use they serve, they may be termed railway, highway, canal, aqueduct, utility pipeline, or pedestrian bridges. If they are classified by the materials of which they are constructed (principally the superstructure), they are called steel, concrete, timber, stone, or aluminum bridges. Deck bridges carry the deck on the very top of the superstructure. Through bridges carry the deck within the superstructure. The type of structural action is denoted by the application of terms such as truss, arch, suspension, stringer or girder, stayed-girder, composite construction, hybrid girder, continuous, cantilever, or orthotropic (steel deck plate), prestressed, or segmental (concrete).

The main load-carrying member or members of a bridge are almost invariably parallel to the alignment of the bridge. When the alignment of the bridge and the obstacle being bridged are not square with one another, the main structural members may not be opposite one another, and the deck may be a parallelogram in plan; in this case the bridge is said to be a skewed bridge. Otherwise, it is known as square.

Many bridges are also designed on horizontally curved alignments to conform with curved approach roadways.

The two most general classifications are the fixed and the movable. In the former, the horizontal and vertical alignments of the bridge are permanent; in the latter, either the horizontal or vertical alignment is such that it can be readily changed to permit the passage beneath the bridge of traffic, generally waterbound, which otherwise could not pass because of restricted vertical clearance. Movable bridges are sometimes called drawbridges in reference to an obsolete type of movable bridge spanning the moats of castles.

A singular type of bridge is the floating or pontoon bridge, which can be a movable bridge if it is designed so that a portion of it can be moved to permit the passage of water traffic.

The term trestle is used to describe a series of stringer or girder spans supported by braced towers or bents, and the term viaduct is used to describe a structure of many spans, often of arch construction.

Fixed Bridges

Fixed-bridge construction is selected when the vertical clearance provided beneath the bridge exceeds the clearance required by the traffic it spans. For very short spans, construction may be a solid slab or a number of beams; for longer spans, the choice may be girders or trusses. Still longer spans may dictate the use of arch construction, and if the spans are even longer, stayed-girder bridges are used. Suspension bridges are used for the longest spans.

Each of the above types of bridge is generally

designed so that the substructure, superstructure, and deck are each considered to carry only the loads directly imposed upon them. In certain types of construction, the deck and the main load-carrying superstructure members are made to participate in carrying the load in order to make the bridge more economical than ordinary stringer or girder bridges with a concrete deck. This is known as composite or hybrid construction, depending upon the type of stringers. If there is stringer or girder-type construction, with a deck partly of steel topped by an asphalt surfacing material, the bridge is known as an orthotropic bridge. When the substructure and superstructure act together, the bridge is described as being of rigid frame construction.

The longer the span of a bridge, the greater is the relative cost per unit of deck area.

The choice of type of bridge superstructure may depend not only on the obstacle to be spanned but also on the substructure. Thus, if an expensive substructure is required because of water depth or unsatisfactory foundation conditions, the selection of a longer-span superstructure may be indicated even though it may not in itself be the economical choice.

Beam bridge. Beam stringer bridges consist of a series of beams, usually of rolled steel, supporting the roadway directly on their top flanges. The beams are placed parallel to traffic and extend from abutment to abutment. When foundation conditions permit the economical construction of piers or intermediate bents, a low-cost multiple-span structure can be built. Spans of 50 ft (15 m) for railroad beam bridges and 100 ft (30 m) for highway beam bridges may be economical.

Composite I-beam bridges are beam bridges in which the concrete roadway is mechanically bonded to the I beams by means of shear connectors, which develop horizontal shear between the concrete slab and the beam. Such a connection forces a portion of the slab to act with the beam, resulting in a composite T beam. This construction yields a saving in the weight of the beams. Rolled shapes such as the channel, angle, and I, bars in serpentine form or in the form of a longitudinal helix, and steel studs are used as shear connectors. These connectors are usually welded to the flange of the steel beam and should extend at least halfway into the slab. *See* COMPOSITE BEAM.

Plate-girder bridge. Plate-girder bridges are used for longer spans than can be practically traversed with a beam bridge. In its simplest form, the plate girder consists of two flange plates welded to a web plate, the whole having the shape of an I. The railroad deck plate-girder bridge consists of two girders which support the floor system for a single track directly on their top flanges. A double-track bridge usually consists of two single-track bridges placed side by side on common abutments or piers. Through plate-girder bridges are used when clearance below the structure is limited. For railroad traffic, the floor system consists of a number of transverse floor beams which are supported by the girders just above their lower flanges. *See* PLATE GIRDER.

Box-girder bridge. Steel girders fabricated by welding four plates into a box section have been used for spans from 100 to more than 850 ft (30 to 259 m). The Rhine River crossing at Cologne, Germany, is an example of an 850-ft (259-m) span. In the United States, a 750-ft (229-m) box-girder span is used in the San Mateo–Hayward Bridge in California. Conventional floor beams and stringers can be used on box-girder bridges, but the more economical arrangement is to widen the top flange plate of the box so that it serves as the deck. When this is done, the plate is stiffened to desired rigidity by closely spaced bar stiffeners or by corrugated or honeycomb-type plates. These stiffened decks, which double as the top flange of the box girders, are termed orthotropic. The wearing surface on such bridges is usually a relatively thin layer of asphalt. Single lines of box girders with orthotropic decks can be used for two-lane bridges, but when wider decks are required, two or more box-girder bridges can be placed parallel to each other.

Curved-girder bridge. Bridges on curved roadways present special problems. Such bridges must be deck bridges. For lightly curved roadways, or when the spans are short, straight stringers or girders usually are used and are positioned under the roadway parallel to a chord of the circular arc of the roadway. When the spans are long or when the curved roadway is of short radius, girders that are curved horizontally are used. These girders parallel the curved edges of the roadway. Sometimes curved girders are chosen for esthetic considerations even when straight girders might be practical, because straight girders on curved bridges are sometimes considered unattractive.

Prior to 1961 there were only a few curved-girder bridges. Both their design and fabrication are difficult, and their erection can be troublesome. Their design involves not only the usual vertical bending forces but also torsional forces caused by the unbalanced loading of a span whose supports are offset from its center of gravity. Early in their use, design was somewhat empirical and overly conservative. However, development of reliable computer programs for curved-girder design has made the use of curved girders more practical.

In addition to individual curved girders, curved steel box girders and curved concrete box girders are used.

Truss bridge. Truss bridges, consisting of members vertically arranged in a triangular pattern, can be used when the crossing is too long to be spanned economically by simple plate girders. Where there is sufficient clearance underneath the bridge, the deck bridge is more economical than the through bridge because the trusses can be placed closer together, reducing the span of the floor beams. For multiple spans, a saving is also effected in the height of piers. *See* TRUSS.

Through-truss. A through-truss bridge is illustrated in **Fig. 1**. The top and bottom series of truss members

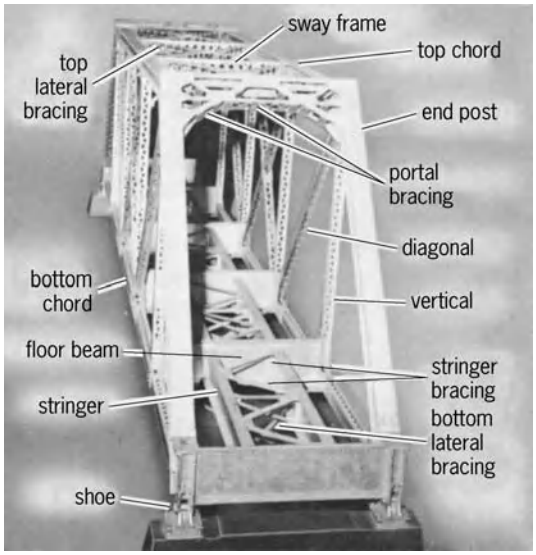

Fig. 1. Model of a through-truss railroad bridge.

parallel to the roadway are called top chords and bottom chords, respectively. The diagonals and verticals form the web system and connect the top and bottom chords. The point at which web members and a chord intersect is called a panel point. Gusset plates connect the members intersecting at a panel point.

Lateral bracing of a bridge ties the two trusses together and assures a stable and rigid structure. The top lateral bracing consists of cross struts connecting the top chords at opposite panel points and the diagonals joining the diagonally opposite ends of adjacent cross struts. It decreases the unsupported length of the top chord members, reducing the cross-sectional area they require. The floor beams and the diagonals connecting the opposite ends of adjacent floor beams constitute the bottom lateral system. Although the floor system of a highway bridge can take over the function of a lateral truss, the lateral bracing must be provided to stiffen the structure during erection and to furnish wind resistance until the steel or concrete floor slab is in place.

The stringers of an open-floor railroad bridge must be braced to relieve them of bending due to lateral forces from the train. In addition, the floor beams should be provided with bracing to relieve the bending due to tractional forces.

Portal and sway bracing are systems of bracing in transverse vertical planes of a bridge. Intermediate sway frames in the plane of opposite verticals give added rigidity.

The end posts of the through-truss bridge are tied together to form a rigid frame or portal capable of transferring the end reaction of the top lateral system to the abutments. To keep the bending stresses in the end posts as small as possible and to provide a rigid portal, portal bracing should be as deep as headroom allows. Also, the end post should be braced by brackets or diagonal members.

Simple-span trusses. Common types of simple-span bridge trusses are shown in **Fig. 2**. The Pratt truss, with its various modifications (Fig. 2*a* and *b*), has tension diagonals and compression verticals. The diagonal in every other panel of a Warren truss (Fig. 2*c*–*e*) is in compression. The depth of short-span trusses is usually determined by the depth necessary for clearance at the portal. For long spans, it is usually economical to make the depth of the truss greater at the

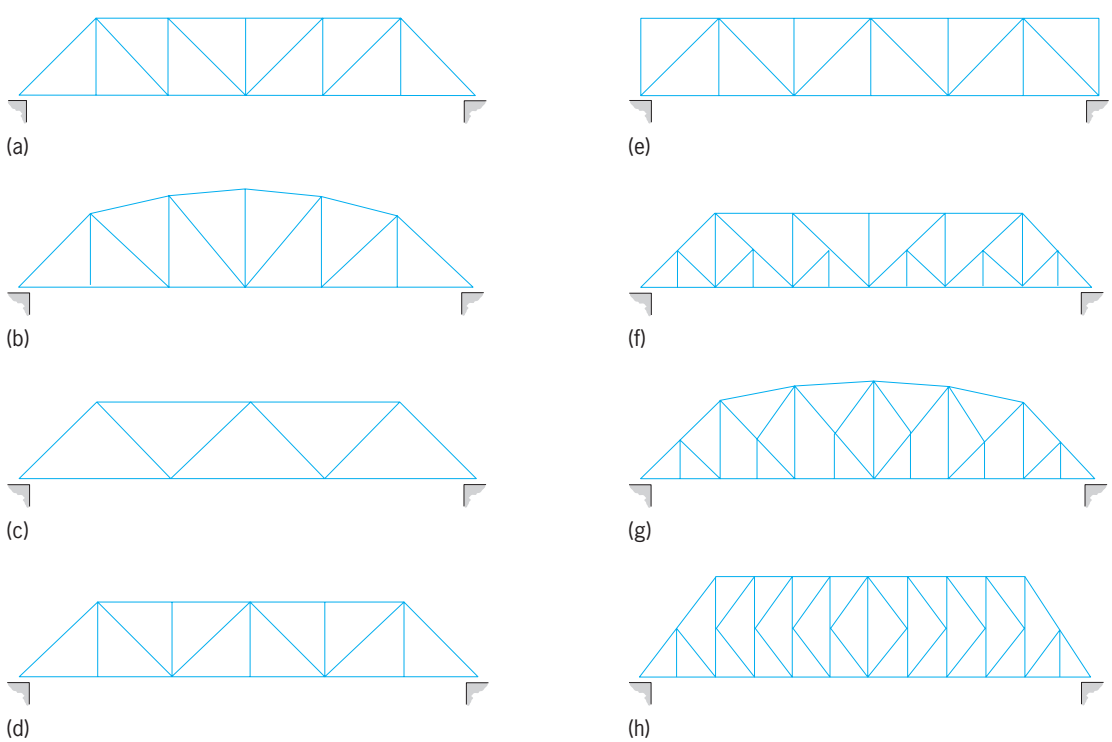

Fig. 2. Examples of simple-span trusses. (*a*, *b*) Pratt trusses. (*c*–*e*) Warren trusses. (*f*–*h*) Subdivided trusses.

center than at the ends. If the depth is increased in proportion to the increase in the forces tending to bend the bridge, the force in the chord members can be more nearly equalized. Figure 2*b* shows a curved-chord Pratt truss.

Trusses of economical proportions usually result if the ratio of depth of truss to length of span is approximately 1:5 to 1:8 and if the diagonals make angles of 45–60° with the horizontal. The panel length produced in long-span trusses when both of these factors are considered results in an uneconomical floor system. Subdivided trusses (Fig. 2*f–h*) are used to get reasonable length panels.

Arch bridge. In an arch bridge the main structural system supporting the deck is a curved member (or members), higher vertically at its center than its ends, acting almost entirely in compression, with this compressive load being maintained by thrust against immovable abutments. Arch bridges have been constructed of stone, brick, timber, cast iron, steel, and reinforced concrete, all of which can adequately take compressive loads. Sophistication of construction of stone masonry arch bridges reached a peak during the nineteenth century, spurred by road improvements, and the construction first of canals and then of railroads. Cast iron as a construction material was introduced at this time. *See* ARCH.

Stone masonry arch bridges are rarely constructed anymore. Almost all modern arch bridges are constructed of steel or reinforced concrete, and occasionally of timber. Also, the main structural load-carrying arch system in a modern bridge consists of a number of ribs, generally two, supporting a deck by columns, termed spandrel columns, or by suspenders, in place of the single barrel of stone masonry construction supporting the deck by walls and earth fill, termed spandrel fill. Thus the road carried by the arch no longer must be above the arch ribs, as in the stone masonry construction, but can pass between the ribs similar to through-truss spans. The choice of arch construction depends largely on foundation conditions at the site, as they must be suitable to economically take the tremendous thrust of the arch ribs. A deck-type arch generally would be used only if the ground level was high at the site, such as at a gorge, in order to obviate long approaches. *See* REINFORCED CONCRETE; STRUCTURAL STEEL.

Fig. 3. Lewiston-Queenston Bridge over Niagara River. (*Niagara Falls Bridge Commission and Hardesty & Hanover*)

Arches of steel have reached a maximum span of 1700 ft (518 m) in a bridge over the New River in West Virginia, and arches of concrete a span of 1000 ft (305 m) in the Gladesville Bridge over the Parramatta River at Sydney, Australia. Spans of less than about 150 ft (46 m) generally are not economical because the cost of construction of the curved arch ribs far exceeds any advantage of an arch over other types of bridges for such short spans.

Ribs. The ribs of an arch bridge terminate at inclined faces, termed skewbacks, of the bridge abutments. While arch ribs are essentially compression members, they are also subjected to bending due to partial live loadings, expansion or contraction due to temperature changes, contraction due to shrinkage in the case of concrete ribs, and the transfer of the deck load to the rib at discrete locations. However, the geometrical shape of the vertical curve of the arch ribs is selected to fit the conditions of the site and the loadings so that ribs act mainly as compression members, and then, for the bending, to make sure that no stress exceeds the allowable. The arch ribs of concrete arch bridges are of reinforced concrete, rectangular in section, varying in depth or width from crown to abutment. The ribs of steel arch bridges may be box-shaped in section, of constant width and depth, but varying in thicknesses of steel material from crown to abutment; or, like a truss bridge with bottom and top chord sections with web members, the ribs may be of constant depth, but varying in chord section from crown to abutment. In all cases, since the ribs are compression members, they generally are braced one to the other.

Hinges. If the arch rib extends continuously from skewback to skewback, and is fixed at the skewbacks so no rotation can occur, the bridge is called a fixed arch. An example is shown in **Fig. 3**. If devices known as hinges are introduced in the arch ribs, the bridge is called a two- or three-hinged arch. The three-hinge arch has hinges at the center (or crown of the rib) and at the skewbacks, and the two-hinged arch has hinges only at the skewbacks. These hinges permit the transfer of axial compressive loads from one section of arch rib to another or to the abutments without transfer of bending moments.

The choice of the number of hinges depends on several considerations. Hinges at the skewbacks simplify the abutment design; and hinges at crown and skewbacks simplify the design of the arch rib as well as the abutment, eliminate stresses due to change in temperature, and allow the arch to tolerate minor movements of the abutments. The fixed arch is more rigid, but much more difficult to analyze, and requires the abutment to take bending loads. Arch

bridges have been built in all three types, and also with a single hinge at the crown.

Tied arches. While most arch bridges have abutments to take the thrust of the arch action, there are a number in which one end of the arch rib is tied to the other with a structural tension member to take the thrust of the rib. These are called tied arches. Usually these are through-arches; the tie is at about deck level and the deck is suspended from the arch rib. These spans may also be a part of another type of construction, a bridge with cantilever and anchor spans, and the tied arch suspended from the cantilever spans, known as a tied arch cantilever.

Continuous bridge. The continuous bridge is a structure supported at three or more points and capable of resisting bending and shearing forces at all sections throughout its length. The bending forces in the center of the span are reduced by the bending forces acting oppositely at the piers. Trusses, plate girders, and box girders can be made continuous. The advantages of a continuous bridge over a simple-span bridge (that is, one that does not extend beyond its two supports) are economy of material, convenience of erection (without need for falsework), and increased rigidity under traffic. Its relative economy increases with the length of span. No increase in rigidity is obtained by making more than three spans continuous. The disadvantages are its sensitivity to relative change in the levels of supporting piers, the difficulty of constructing the bridge to make it function as it is supposed to, and the occurrence of large movements at one location due to thermal changes.

Cantilever bridge. The cantilever bridge consists of two spans projecting toward each other and joined at their ends by a suspended simple span. The projecting spans are known as cantilever arms, and these, plus the suspended span, constitute the main span. The cantilever arms also extend back to shore, and the section from shore to the piers offshore is termed the anchor span (**Fig. 4***a*). Trusses, plate girders, and box girders can be built as cantilever bridges.

The chief advantages of the cantilever design are the saving in material and ease of erection of the main span, both of which are due to the fact that no falsework is needed. By adding continuity of members after erection, the cantilever bridge is made to act as a continuous structure under live load. *See* CANTILEVER.

Cable-stayed bridge. A modification of the cantilever lever bridge which has come into modern use resembles a suspension bridge, and it is termed a cable-stayed bridge. It consists of girders or trusses cantilevering both ways from a central tower and supported by inclined cables attached to the tower at the top or sometimes at several levels (**Fig. 5**). Usually two such assemblies are placed end to end to provide a bridge with a long center span.

Suspension bridge. The suspension bridge is a structure consisting of either a roadway or a truss suspended from two or more cables which pass over two towers and are anchored by backstays to a firm foundation (Fig. 4*b* and *c*). If the roadway is attached directly to the cables by suspenders, the structure lacks rigidity, with the result that wind loads and moving live loads distort the cables and produce a wave motion on the roadway. When the roadway is supported by a truss which is hung from the cable, the structure is called a stiffened suspension bridge. The stiffening truss distributes the concentrated live loads over a considerable length of the cable.

Cables of the larger sizes, such as those up to 36 in. (91 cm) in diameter used for the George Washington Bridge in New York City, are assembled (or spun) in the field by using pencil-thick wires laid parallel. For smaller cables, strands of wire wound spirally in the factory are assembled into cables. Factory-made strands of parallel wires were used for the first time in 1968 for the 15-in.-diameter (38-cm) cables of the Newport Bridge in Rhode Island.

The longest bridge spans in the world are all of the suspension type. The longest main spans occur in the Humber Bridge of England (4626 ft or 1410 m), the Verrazano-Narrows Bridge in New York City (4260 ft or 1298 m; Fig. 4*c*), the Golden Gate Bridge in San Francisco (4200 ft or 1280 m), and the Mackinac Bridge in northern Michigan (3800 ft or 1158 m). The Akashi Strait Bridge in Japan, completed in April 1998, with the longest single span (6527 ft or 1989 m), is the world's longest suspension bridge.

Pontoon bridge. The pontoon bridge is a floating bridge supported by pontoons. The structure may be a temporary one, as for military usage, or a permanent one, if the level of the water can be carefully controlled. A pontoon bridge may be advantageous where deep water and adverse bottom conditions make piers expensive. Seattle's Lake Washington is crossed by three floating concrete bridges. Typically, the pontoons are 360 ft (110 m) long, 60 ft (18 m) wide, and 14.5 ft (4.4 m) deep. A 378-ft (115-m) pontoon is used as a floating draw span to give a clear channel opening 200 ft (61 m) wide. Adjacent pontoons are bolted together, and each is secured by a pair of anchors, one on each side. Cast-in-place reinforced concrete was used for the pontoons of the first bridge, but the later bridges are of prestressed concrete.

Concrete bridge. The bridges that have been discussed above are usually made of steel, although they may carry a concrete roadway. Increasingly, however, since the development of the prestressing method, bridges of almost every type are being constructed of concrete. Prior to the advent of prestressing, these bridges were of three types: (1) arch bridges, which were built in either short or long spans, even up to 1000 ft (305 m) for the Gladesville Bridge in Sydney, Australia; (2) slab bridges of quite short spans, which were simply reinforced concrete slabs extending from abutment to abutment; and (3) deck girder bridges, consisting of concrete slabs built integrally with a series of concrete girders placed parallel to traffic. The advent of prestressed concrete greatly extended the utility and economy of concrete for bridges, particularly by making the hollow box-girder type practicable. One of the longest such spans, 682 ft (208 m), is in a bridge over the

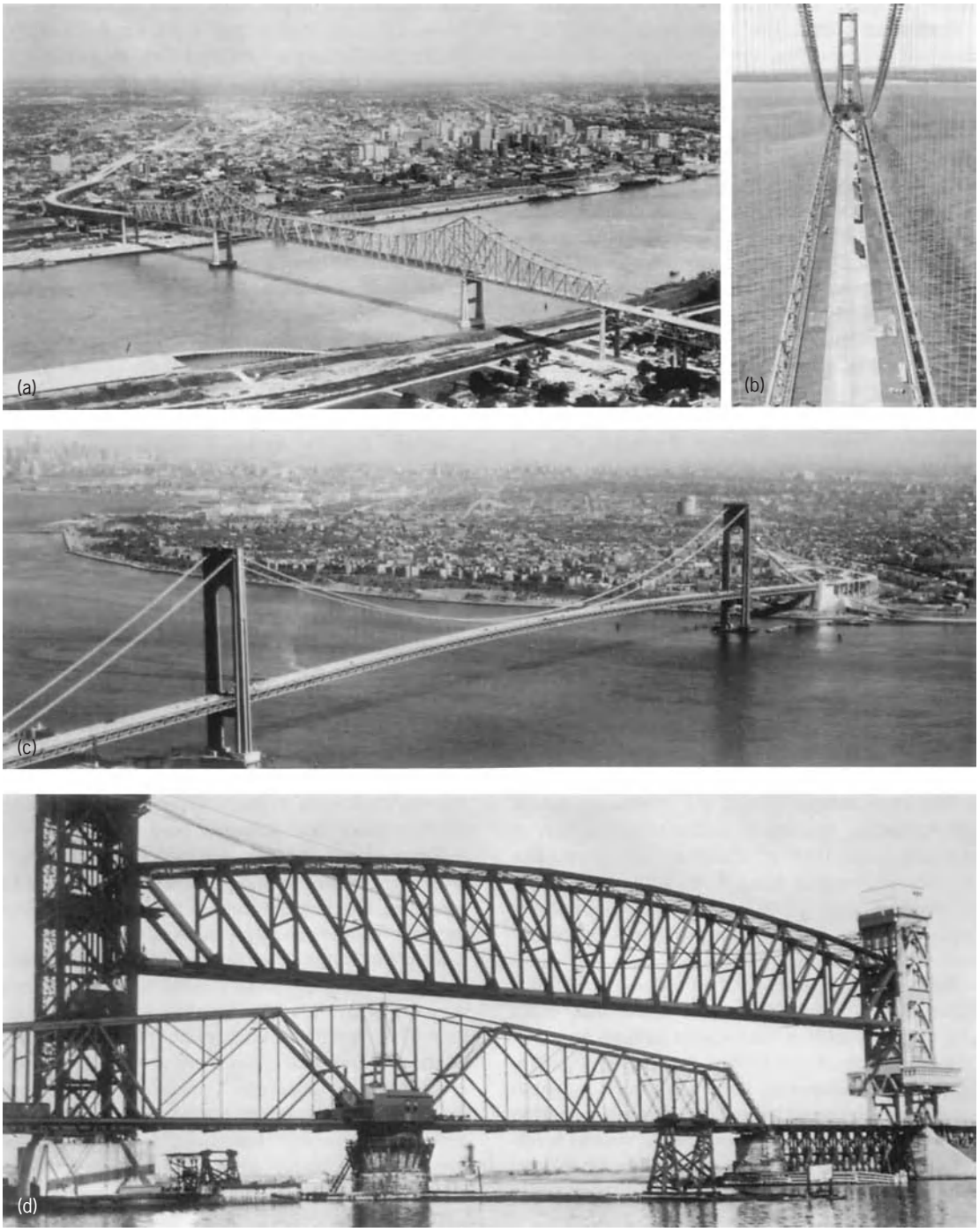

Fig. 4. Some major United States bridges. (*a*) Greater New Orleans cantilever bridge, across the Mississippi River (*Bethlehem Steel Co.*). (*b*) Looking from south tower during construction of Mackinac Bridge, connecting upper and lower peninsulas of Michigan (*Mackinac Bridge Authority*). (*c*) Verrazano-Narrows Bridge, New York City (*Triborough Bridge and Tunnel Authority*). (*d*) Vertical-lift span and swing bridge across Arthur Kill, between Staten Island and New Jersey. The vertical-lift bridge replaced the swing bridge, in use since 1888 (*Baltimore and Ohio Railroad Co.*).

Rhine River near Koblenz, Germany. *See* PRESTRESSED CONCRETE.

The objective in prestressing concrete is to reduce or eliminate tensile stresses in the concrete by applying a force that greatly increases internal compressive stress in the concrete member. The force is applied by stretching the reinforcing tendons, which may be either wires or bars, by means of hydraulic jacks that react against the ends of the concrete member. Prestressing may be done to precast members, or it may be done as the concrete is being placed in the field. It may be done by either pretensioning or posttensioning. In pretensioning, the reinforcing tendons are stretched prior to the placing of the concrete. In posttensioning, the reinforcing tendons are installed in tubes so that they are isolated from the concrete while it is being placed. When the concrete hardens, the tendons are stretched by jacks reacting

Fig. 5. Sunshine Skyway Bridge in Tampa, Florida. (*Florida Department of Transportation and Figg and Muller Engineers, Inc.*)

against and compressing the concrete member.

Bridge bearings. Almost all bridges have devices known as bearings or shoes where the superstructure transfers its vertical and horizontal loads to the substructure. Exceptions can be arch bridges, rigid frame bridges, and suspension or cable-stayed bridges, where other devices or actions obviate the need of bearings. These bearings are either fixed, expansion, or sliding. All the main structural members of a single-span bridge—the stringers, girders, and trusses—have a fixed bearing at one end of the support and an expansion bearing at the other. Multiple-span bridges have a single fixed bearing and one or two expansion bearings and sliding bearings between fixed and expansion bearings on each main structural member.

The fixed bearing transfers not only vertical loads to the substructure but also the longitudinal horizontal wind or traffic tractive forces. The expansion bearing takes vertical loads but allows longitudinal movements of the end of the member it supports. These movements are caused by temperature changes and changes in the length of the span caused by loadings. Sliding bearings function similarly. All bearings accommodate rotation of the ends of the span when the span or spans deflect under loading. Also, all bearings transfer transverse wind forces from the superstructure to the substructure of the bridge.

Depending on the span of the bridge, bearings can vary from simple steel plates curved on their undersides, one of which slides, to large cast-steel devices with a rocker, or nest of cylindrical rollers at the expansion bearing.

Expansion bearings are one of the most troublesome parts of the bridge as they are exposed to deterioration from moisture and debris sifting down from expansion openings in the bridge deck. In fact, a few single- and multiple-stringer bridges have been built experimentally without bearings, depending upon the restraint of the bridge abutments and the reversed stress induced in the superstructure to substitute for allowed expansion.

Also, since the mid-1960s bearings known as elastomeric bearings have been used for spans up to about 100 ft (30 m). These are blocks of resilient synthetic rubber which deform to accommodate the expansion and rotational forces.

Movable Bridges

Modern movable bridges are either bascule, vertical lift, or swing; with few exceptions, they span waterways. They are said to be closed when they are set for the traffic they carry, and open when set to permit marine traffic to pass through the waterway they cross.

Both bascule and vertical-lift bridges operate well and reliably. Sometimes the bascule is chosen because its appearance may be pleasing and creates

Fig. 6. Stanley Stroffolino Bridge in Norwalk, Connecticut. (*Connecticut Department of Transportation and Hardesty & Hanover*)

less of a visual impact than do the towers and ropes of vertical-lift bridges.

Swing bridges are now considered almost obsolete because the center pier, on which the span rotates, occupies the portion of the waterway that is most desirable from the standpoint of mariners. However, one feature of a swing bridge, which dictates its use near airports, is that it does not encroach on flight paths as would the towers of lift bridges or the raised leaves of bascule bridges.

Bascule, swing, and lift bridges may be of either stringer, girder, or truss construction, depending upon the length of their span.

Other types of movable spans, such as the retractable, are generally obsolete.

Bascule and swing bridges provide unlimited vertical clearance in the open position, whereas the vertical clearance of a lift bridge is limited by its design.

Bascule bridge. The bascule bridge consists primarily of a cantilever span, which may be either a truss or a plate girder, extending across the channel (**Fig. 6**). This type is generally chosen for spans up to about 175 ft (53 m) for highway use. Because of the large deflection of the cantilevered leaves of the double-leaf bascule, which is excessive under railway loadings and intolerable to railroad operation, bascule spans for railway usage are exclusively single-leaf bridges with maximum spans of about 250 ft (76 m).

Bascule bridges rotate about a horizontal axis parallel with the waterway. The portion of the bridge on the land side of the axis, carrying a counterweight to ease the mechanical effort of moving the bridge, drops downward, while the forward part of the leaf opens up over the channel much like the action of a playground seesaw. Bascule bridges may be either single-leaf, where the entire leaf rotates over the waterway about a single axis on one side of the waterway, or double-leaf, where two leaves over the waterway rotate about two different axes on opposite sides of the waterway. The two leaves of double-leaf bascule bridges are locked together where they meet when the bridge is closed. If the bridge actually rotates about an axis, it is called a trunnion-type bascule. If it rolls back on a track, it is called a rolling lift span.

Fig. 7. Stratford Avenue Bridge in Bridgeport, Connecticut. (*Connecticut Department of Transportation and Hardesty & Hanover*)

Vertical-lift bridge. The vertical-lift bridge has a span similar to that of a fixed bridge, and is lifted by steel ropes running over large sheaves at the tops of its towers to the counterweights, which fall as the lift span rises and rise as it falls. If the bridge is operated by machinery on each tower, it is known as a tower drive. If it is driven by machinery located on the lift span, it is known as a span drive. The 585-ft (178-m) span of the lift bridge over the Arthur Kill, an arm of New York harbor, is the longest of this type in the world (Fig. 4*d*). Another example of a vertical-lift bridge is shown in **Fig. 7**.

Swing bridge. Swing bridges rotate about a vertical axis on a pier, called the pivot pier, in the waterway (Fig. 4*d*). There are three general classes of swing bridges: the rim-bearing, the center-bearing, and the combined rim-bearing and center-bearing. A rim-bearing bridge is carried on a cylindrical girder on rollers, and a center-bearing on a single large bearing at the center of rotation. Swing bridges have been classified also as to the character of their main girders—that is, plate girder swing spans or truss swing spans.

Machinery. Almost all modern movable bridges are driven by electric motors which operate gear trains that convert the high-speed low-torque output of the motor to a low-speed high-torque output of the gear train at a pinion, acting on a rack.

Originally, movable bridges were operated by steam engines, and until recently, many still were. At least one former steam-operated bridge was operated for many years by compressed air driving the old steam engine. A very few bridges have been operated hydraulically. Some small bridges are operated by hand power. Provisions are still made on most movable bridges to operate by hand power in the event of power failure or malfunction of the bridge's electrical controls.

Machinery for early movable bridges consisted of simple custom-made components produced by the numerous small foundries, forge shops, and machine shops. Such components consisted of cast and roughcut gears of various tooth profiles, babbitt and bronze bushed bearings, custom-designed mounts for individual bearings, and common frames for the mounting of multiple bearings for open sets of reduction gears and various combination drive assemblies (**Fig. 8**). After World War II, a wide variety of standardized gear reducers, shaft bearings, and other machinery components suitable for use on bridge drives became available (**Fig. 9**). These components have generally replaced the custom-made components previously used.

Welded components have replaced cast components. Speed-reducer housings are essentially all

Fig. 8. Exposed gear reductions on an early bascule bridge.

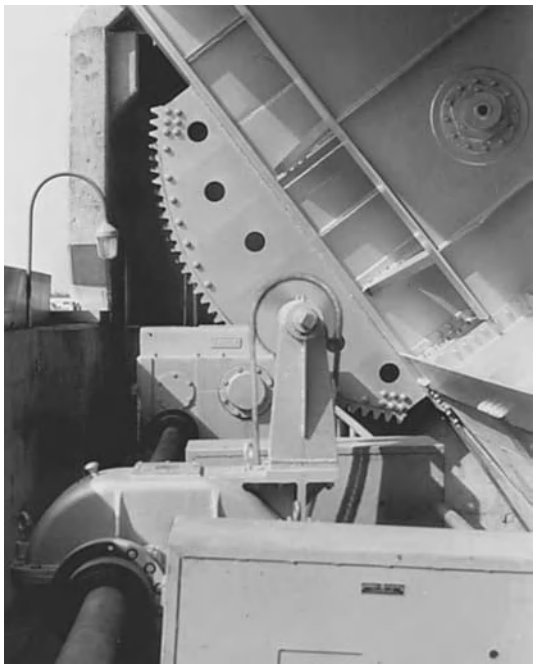

Fig. 9. Enclosed gear reducers on a later bascule bridge. (*Earle Gear and Machine Co.*)

welded. Even the huge sheaves for vertical-lift bridges are commonly manufactured as weldments instead of castings.

Electrical equipment. Except on the simplest small movable bridges, the electrical control of a movable bridge is so interlocked that the bridge cannot be moved until a series of prior operations have been made in correct order. Thus, on a highway bridge, first the traffic lights must be turned to red, then a set of traffic gates lowered, next the barrier gates set, and finally the locks or wedges pulled before the bridge can be moved.

The movable span generally can be completely opened or closed in from 1 to 2 min. However, the prior operations, particularly closing the bridge to traffic and making sure that the bridge is cleared of all vehicles between traffic gates, may run the cycle of operation up to 15 min or longer, depending upon the speed of passage through the span of the vessel for which the bridge has opened.

The drive systems on movable bridges have progressed from the early simpler forms of technology to the solid-state control devices permitting pushbutton operation.

Design

Bridges are designed according to the laws of physics pertaining to statics. The primary members of bridges act in tension, compression, shear, or bending, or in combinations of these. Secondary members act in the same fashion, sometimes participating in carrying the principal loads, especially if failure of any primary members should occur. Redundancy, so that failure of a single member or part of a member will not cause immediate collapse of the bridge, is now recognized as a desirable feature of design, particularly since collapse of the eyebar suspension bridge across the Ohio River at Point Pleasant, West Virginia, in 1967.

Modern design of bridges follows standards established by their principal users, the state highway or transportation departments and the railroads. In the United States, these standards are those of the American Association of State Highway and Transportation Officials and the American Railway Engineering Association, respectively. These specifications are developments of those for bridge design formulated by the early bridge engineers. While they are advisory, they are generally followed, albeit at times with modifications. Other countries have similar specifications.

The traffic load, called live load, is given in these specifications. It depends on the service to which the bridge will be subjected. The highway loadings are in terms of a simulated conventional truck or, for long spans, a uniform load with a roving concentrated load representing a line of average traffic. Railroad loadings are a simulated conventional locomotive load followed by a uniform load representing loaded freight cars. In both cases, in design, the live loads are positioned to give maximum load in the member being designed. In addition, dead load, the structure's own weight, and loads from impact, wind, temperature changes, ice, traction or braking of traffic, earthquake, and, in the case of highway bridges, pedestrians are specified. Allowable stresses and limiting deflections, too, are specified.

Other specifications followed in the design and construction of bridges include those of the American Society for Testing and Materials, the American Welding Society, and the American Institute of Steel Construction, and the construction specifications of the individual owners, and special specifications, when no other covers the situation. *See* STRUCTURAL DESIGN.

Vibration

Bridges are generally considered to be statically loaded structures under their own dead load, with dynamic loadings from the live load and from the wind.

Vibrations of bridges or individual components of a bridge occur when a resonant frequency of the bridge or a component is excited by one or more of

the applied dynamic loadings. Excitation from live loads is more apt to happen on bridges carrying rail traffic than on those carrying highway traffic because of the uniform spacing of railroad cars. Excitation from wind is caused by the repeated formation of eddies or vortices as the wind travels past nonstreamlined members. Such aerodynamically induced vibrations caused the failure of the Tacoma Narrows Bridge in the state of Washington in 1941. A similar failure is recorded for a suspension bridge over the Ohio River at Wheeling, West Virginia, in 1854. In both cases, the formation of vortices built up as the torsional movements of the deck increased, ultimately leading to failure. Long, thin, H-shaped truss members have failed from similar wind-induced torsional vibrations. Wind velocities necessary to excite such vibrations need not be excessively high. The Tacoma Narrows Bridge failed under a 40 mi/h (18 m/s) wind, less than half the equivalent static wind load for which it was designed, and less than 20% of the load which would have caused structural distress.

Light lateral bracing on many spans will vibrate under the passage of live load.

Sufficient data exist on aerodynamically induced vibration to predict the possibility of its occurrence and thus modify the design when necessary to prevent serious problems.

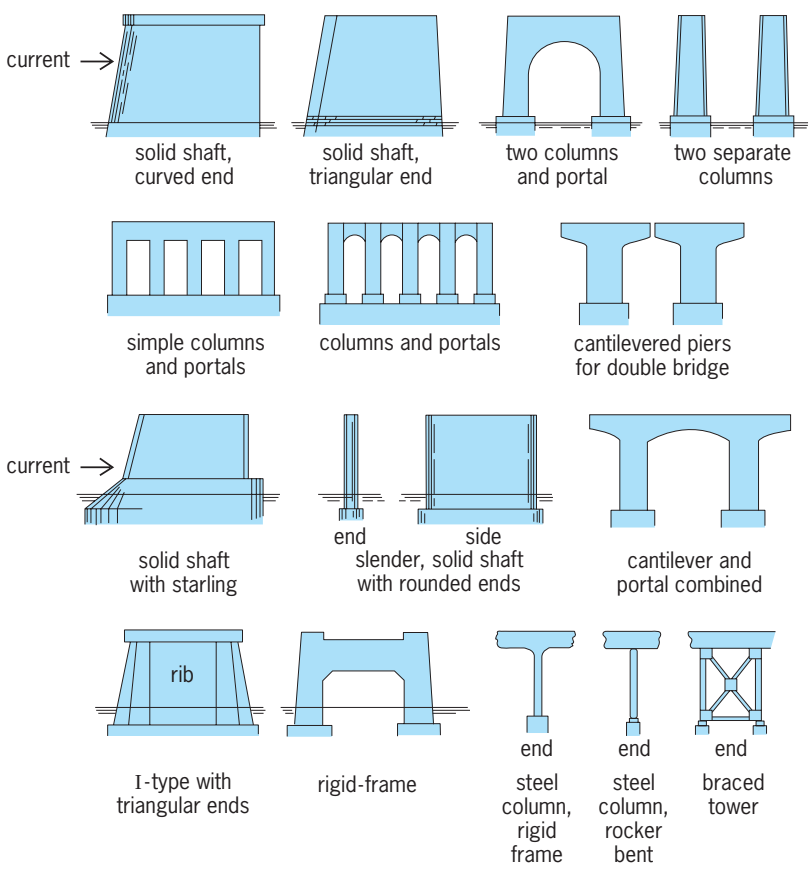

Fig. 10. General shapes of some piers for small bridges. (*After C. W. Dunham, Foundations of Structures, McGraw-Hill, 2d ed., 1962*)

Substructure

Bridge substructure consists of those elements that support the trusses, girders, stringers, floorbeams, and decks of the bridge superstructure. Piers and abutments are the primary bridge substructure elements. Other types of substructure, such as skewbacks for arch bridges, pile bents for trestles, and various forms of support wall, are also commonly used for specific applications.

The type of substructure provided for a bridge is greatly affected by the conditions of the site. Studies must be made on topography, stream currents, floating drift and ice, seismic potential, wind, and soil conditions. Forces and loads encountered in the design of substructure elements include dead load, live load, impact loads, braking forces, earth pressure, buoyancy, wind forces, centrifugal forces, earthquake loads, stream flow, and ice pressure.

Both piers and abutments are generally supported on either spread footings or pile footings. A spread footing is usually a concrete pad large enough in area to transmit all superimposed loads and forces directly to the soil on which it is founded. The size of the spread footing is related to the bearing capacity of the soil on which it rests and the external forces on the substructure which will be transmitted to the spread footing.

A pile footing is usually a large concrete block supported on piles, so that the superimposed substructure loads and forces are transmitted to the support piles through the footing. The footing piles are used primarily to transmit loads through soil formations having poor supporting properties into or onto formations that are capable of supporting the loads. Piles may be point-bearing or friction types or the two in combination, and they may be timber, steel, precast concrete, cast-in-place concrete, or prestressed concrete. They may be driven by the use of a pile driver equipped with a hammer; they may be augered, jetted or prebored, or predrilled and cast-in-place. *See* FOUNDATIONS.

Bridge piers. Bridge piers are the intermediate support systems of bridges and viaducts. They may be located in water or on dry land. When located in water, piers may be subjected to scour by current and collision by vessels. Bridge piers support the superstructure and must carry dead loads and live loads, and withstand braking forces and other induced forces peculiar to the location of the pier, such as wind, ice, earthquake, and stream flow. A major consideration in pier design is stability and the ability to support all loads without appreciable settlement.

The shape, type, and location of piers are based on many factors; the major ones are horizontal and vertical clearance requirements, subsurface conditions, architectural and esthetic considerations, political and urban planning factors, traffic, and cost.

The most common pier shapes are solid shafts, multiple columns and portal, two columns and portal, separate columns, T or hammerhead, and cantilever. There are many variations of these pier shapes, which are constructed using concrete, steel,

or wood (**Fig. 10**). A pier should have sufficient horizontal area at its top to receive the superstructure bearings. Architecturally, it should give the appearance of strength; it should not look weak and flimsy, although calculations may have shown the design to be adequate.

Piers in water are sometimes faced with stone, steel, or other protective devices below the high-water line to protect the pier from scour, ice, and other floating matter. At spans over navigable waterways, the piers of the channel span are provided with resilient timber fenders and clusters of timber piles for the protection of the piers and vessels.

Abutments. The abutments of a bridge are the substructure elements that support the ends of a bridge (**Fig. 11**). Bridge abutments are generally constructed of concrete or masonry and are designed to be pleasing esthetically as well as architecturally and functionally sound. Major loads to which abutments are subjected are dead loads, live loads, braking forces, ice, wind, earthquake, stream flow, and earth pressure. The last is applied at the rear of the abutment wall.

An abutment is generally composed of a footing, a wall with a bridge seat supporting the superstructure bearings, and a backwall to retain the earth. Abutments may have wingwalls to retain the earth of the approach fill to the bridge. Proper drainage behind abutment walls is essential to avoid increasing the lateral pressure forces.

Skewbacks. A skewback is a common expression for an abutment for an arch; it is practically nothing more than an inclined footing that receives the thrust from the arch superstructure (Fig. 11).

A skewback differs from common abutments in the type and direction of forces that are applied to it. The horizontal component of an arch reaction is usually very large and exceeds the vertical component.

The large thrusts which are supported by skewbacks require good foundations. The ideal foundations are rock gorges which have strong, sound, and suitably sloping rock.

In general, a skewback for a fixed arch must provide practically no yielding. Some hinged arches, which are statically determinate structures, can sustain slight yield of the foundation. Skewbacks for these may be supported on piles or on spread footings on earth.

Caissons. A caisson is a boxlike structure, round or rectangular, which is sunk from the surface of either land or water to the desired depth as excavation proceeds inside or under it. It is an aid in making excavations for bridge piers or abutments and remains in place as part of the permanent structure.

The most common types are open caissons, which have no top or bottom. Pneumatic caissons have permanent or temporary tops and are so arranged that people can work in the compressed air trapped under the structure.

An innovation in caissons was used for the Tappan Zee Bridge, which carries the New York State Thruway across the Hudson River 13 mi (22 km) north of New York City. Here eight buoyant precast

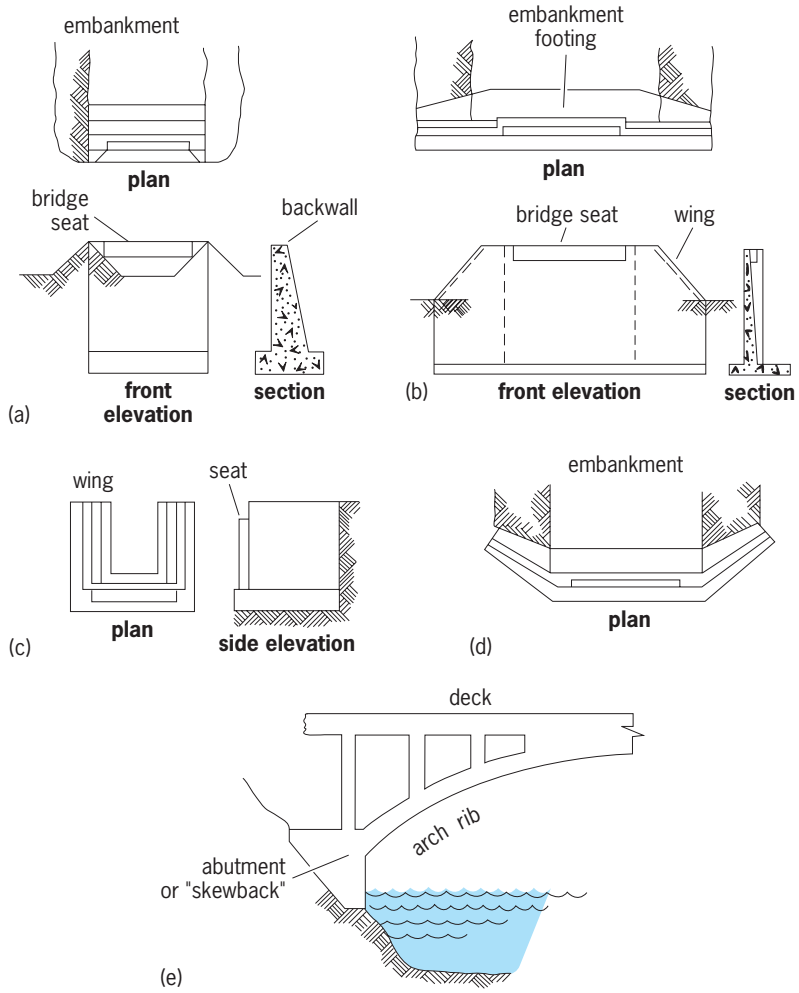

Fig. 11. Abutment designs for bridges. (*a*) Wingless. (*b*) Straight wing. (*c*) U-shaped. (*d*) Beveled wing. (*e*) Abutment for arch bridge.

concrete box caissons reduce the structure's dead load on supporting piles by 70%. The buoyant effect is maintained by keeping the boxes dewatered. Each of the two 15,000-ton (13,600-metric-ton) caissons supporting the main piers measures 100 ft (30 m) by 190 ft (58 m) and 40 ft (12 m) high.

Generally, the ultimate purpose of caisson construction is to reach a bearing stratum which will carry the load of supporting piers and abutments. *See* CAISSON FOUNDATION.

Cofferdam. A cofferdam, generally, is an enclosed temporary structure used to protect an excavation against lateral earth pressure or water pressure during construction. The material within the confines of the cofferdam is removed to allow the construction of piers and abutments below ground or water level.

In constructing cofferdams to counterbalance the upward pressure which may exist at the bottom of the caissons or cofferdams, a concrete seal, also known as tremie seal, is placed prior to dewatering. Concrete is placed in water using vertical pipes known as tremies. These pipes are continually filled with concrete. The bottom of the pipe is submerged in the plastic concrete while the concrete is poured through a funnellike top. During the pour, the tremie pipe is pounded, vibrated, and raised and lowered

to increase the hydraulic head of the concrete in the tube and to cause the concrete to flow. The depth of the tremie concrete pour is a function of the upward pressure encountered after dewatering. *See* COFFERDAM.

Repair and Rehabilitation

It is imperative that bridges are structurally able to carry the traffic for which they are intended. Consequently, they must be inspected regularly for any defects which may have developed and must be appropiately repaired.

Inspection. A bridge inspection and maintenance program resulted from the 1967 collapse of the U.S. Route 35 highway bridge over the Ohio River between Ohio and West Virginia. Subsequently, a policy for the inspection of highway bridges was developed by the Federal Highway Administration. Initially the program required interim inspections of bridges every 2 years, giving priority to those built before 1935, and in-depth inspections not to exceed 5-year intervals. The first round of inspections of approximately 600,000 highway bridges found that almost half of the nation's bridges were structurally deficient or functionally obsolete.

Bridge inspection entails a close-up examination of all parts of the structure, sometimes called a tactile examination, by a team of structural engineers supervised by a licensed professional engineer. At one time inspection was made by climbing and rigging or scaffolding for access. While the early methods are still used to some extent, special automotive equipment, such as snoopers and cherry pickers, are now employed for quick and easy access below the decks.

After inspection, the condition of a bridge is rated on a sliding scale. If a bridge is rated at the low end of the scale, the bridge may be closed or a load limit established until repairs can be made. If deterioration of primary structual members is found, the bridge is rated as to its load-carrying capacity related to the load-carrying cross-sectional area, known as loss of section, found in those members.

Maintenance. Over time, bridges rust, decay, crumble, fail, grow obsolete, and become functionally inadequate. Some 400,000 bridges in the United States were built prior to 1935. They were designed for vehicles lighter than those used in the 1990s and cannot sustain heavier truck loads without extensive, expensive strengthening. Most highway bridges were designed in compliance with standards developed in the 1950s and 1960s. In the 1950s, bridges were designed to carry a 36-ton (33-metric ton) vehicle. The more recent bridges are designed for 40-ton (36-metric ton) vehicles and more, which is one of the major causes for the rapid deterioration of older bridges designed for lighter loads.

Maintenance of bridges, much of which is developed from the ongoing inspection program, is a priority function. The most common maintenance problems are deterioration of reinforced concrete decks; corrosion of steel caused by salt usage for snow and ice removal; nonfunctioning bridge bearings and drainage; deterioration of substructure elements such as piers, abutments, and walls; damage to bridges by natural causes such as floods and earthquakes; and damage by fire and allision (striking of the bridge by a vessel).

Another common maintenance issue, metal fatigue, results from unforeseen high stresses that eventually cause failure. It occurs particularly in details that are unduly restrained or that have defects. Likewise, when a crack emanating from a defect in a member or a weld (existing, accidentally induced, or developing) reaches a critical length, it will propagate through the member and may cause failure.

Bridge superstructures. An alarming aspect of the condition of bridges is that many deteriorated portions are not readily visible. Hidden problems include the deteriorated structural reinforced concrete bridge slabs covered by wearing surfaces; minute cracks in steel structures, which are possible sources of fatigue failures; and the reduction in cross section of metallic members caused by corrosion.

The major maintenance and safety problem is the deterioration of bridge decks. Three principal defects are usually encountered—cracking, scaling, and spalling. Spalling is the most serious and the most difficult to control. It is usually caused by salt solutions that permeate through small cracks and pores in the concrete and corrode the reinforcing steel, thereby weakening the bond between the steel and the concrete. The weakened deck becomes progressively more susceptible to failure under truck loading.

Repair and rehabilitation of concrete decks usually involves the removal of the deteriorated concrete. Deterioration is determined primarily by deck cores establishing the salt content in the concrete, which affects the concrete strength and the degree of corrosion in the reinforcing steel. Often it is more economical to remove and replace the entire concrete deck. Generally the life of a bridge deck under modern usage is approximately 30 years. Many bridge decks are repaired and protected with epoxy surface coatings or waterproofing materials, covered with an asphalt overlay serving as the wearing course (riding surface). Another repair method is to mill 2–3 in. (5–7.5 cm) off the existing deck and overlay the remaining deck with a high-density concrete through which salt will not penetrate.

Steel repairs of heavily corroded members usually take the form of welding or bolting reinforcing plates to substitute for the deteriorated material. Often complete replacement of the deteriorated member is warranted. Many steel bridges built prior to 1970 had obsolete structural steel details such as intersecting welds, details that have reduced fatigue life, and details that do not provide for the effects of out-of-plane deformations. Repair details for these have been developed in the United States by the Federal Highway Administration (FHWA). All new bridges are designed with these considerations, and many old bridges are being retrofitted to meet the newer standards.

Other areas of bridge superstructure repair involve bridge bearings which must allow for temperature movements, structural rotation, and seismic movements. In the United States the standards for these

repairs have been developed by the FHWA and the various state departments of transportation, and they are defined in American Association of State Highway and Transportation Officials (AASHTO). Roadway bridge joints also need constant maintenance and repair. It is important that joints are frequently inspected and repaired or rehabilitated with new expansion material or devices.

Substructure. Piers and abutments are the foundations of the bridge. They may be steel or concrete. Their condition must be ascertained periodically and deterioration repaired. For concrete cracks it is possible to use epoxy inspection for the repair, or the crack can be opened to the first layer of reinforcing steel and new concrete can be bonded to the old. Similar repairs can be accomplished for spalls and other major deterioration. Scour in water piers must be repaired by rebuilding the undermined scoured area and protecting it with large stones (riprap) or a protective skirt (cofferdam). Steel substructure elements need the same degree of repair as superstructure elements. Corroded steel must be repaired through reinforcement or replaced.

E. R. Hardesty; H. W. Fischer; R. W. Christie; B. Haber

Degradation. Many factors can cause bridges to degrade and become structurally deficient and in need of repair. They include environmental factors, deicing salt, steel corrosion, vehicular damage, and special loads, such as earthquakes and wind.

Temperature changes. Two environmental factors that cause significant damage to primarily concrete components in bridges are excessive changes in temperature and freeze-thaw cycles in the presence of moisture. Such damage is aggravated if there was inadequate design for expansion and contraction due to temperature changes. In restrained members that cannot expand or contract to accommodate such changes, additional forces are developed that may cause some members to fail. Extreme temperature changes also alter the properties of some materials. Under very low temperatures, some types of steel become brittle and thus vulnerable to fatigue or repetitive load cycling from truck traffic. In concrete, freeze-thaw cycling in the presence of moisture causes cracks and spalling. Most of the degradation in concrete bridges in the northern United States is due to freeze-thaw cycling combined with corrosion of steel reinforcement.

Corrosion. Steel structures are vulnerable to corrosion, especially in prolonged moisture environments. Corrosion degrades the steel members and reduces their effective area, thus reducing their load capacity. The degree of corrosion depends on the chemical composition of steel, frequency of wetting and drying, and the location of the bridge (near salt water, near polluting plants, and so forth). Use of deicing salts on concrete pavements and bridge decks produces chemical reactions that accelerate the corrosion of reinforcing steel.

Vehicular damage. A significant cause of bridge damage is vehicular impact and fatigue from repeated truck loads. A large number of bridges are damaged by trucks driving under a bridge and hitting its girders or directly impacting other bridge members. Damage from repeated truck traffic loads is most visible in the form of potholes in the bridge deck. However, repeated truck loads also may cause cracking of steel members, especially at joints, a phenomenon known as fatigue.

Special loads. Special loads, such as seismic, wind, and snow, also may produce dramatic degradation of bridge structures. In California or other earthquake-prone regions, seismic damage to bridges has been extensive and dramatic, including the collapse of multiple bridge spans as occurred in the Cypress Freeway in Oakland during the 1989 earthquake. Earthquakes can also cause a severe loss of soil support under bridge abutments and piers, a phenomenon known as liquefaction and illustrated by the collapse of piers in the Kobe, Japan, earthquake in 1995. Wind is a special load that can damage or even collapse bridges, especially of flexible type, such as suspension bridges. Loads generated by wind pressure may produce direct damage and vibration or resonance, which results when the wind excites the bridge or sections of it to produce large oscillations. In 1940, Tacoma Narrows bridge in Washington collapsed due to such resonance. Snow is another special load on bridges. In northern regions of the United States, excessive snow and sleet can accumulate on a bridge, causing overloads and damage to the structure. This is especially true for long-span bridges in regions with heavy snowstorms. *See* EARTHQUAKE; MECHANICAL VIBRATION.

Strengthening. Some bridges need to be strengthened although they have not been damaged and are structurally sound. A bridge may need to be strengthened in order to upgrade its use for heavier truck loads. Many older bridges had been originally designed and built for truck loads of lower magnitude and number than those of today. Another cause for strengthening of bridges is to comply with current design codes. Continuous research on material behavior and effects of loads (such as earthquakes and winds) on bridges generates new knowledge, which translates into changes in codes and specifications that govern the design of bridges. Thus, older bridges may be considered structurally deficient according to new codes so that strengthening becomes necessary. The effect of new codes on the design of bridges is most evident in the earthquake regions of the United States. New data and observations collected in laboratory experimentation and actual seismic events are constantly reflected in new and improved seismic design codes. *See* SEISMIC RISK.

Strengthening techniques (concrete). The strengthening of concrete bridges is generally achieved by replacing the damaged material, incorporating additional structural members, as in external prestressing, or increasing the size and capacity of existing members.

Reinforcing bars. Reinforced concrete members are designed with compressive stresses absorbed by concrete and tensile loads taken by deformed steel bars. Concrete has very low tensile strength; thus, adding steel in the tension zones increases the load capacity of a reinforced concrete member. This strengthening

method consists in relieving the stresses from the member, removing concrete cover in the area where new reinforcement will be placed, adding the new steel bars, and finally, placing new concrete. *See* REINFORCED CONCRETE.

Increasing size of members. This technique is similar to the above technique with the difference that not only is new reinforcing steel added but also the overall size of the reinforced concrete section is increased with the addition of new concrete and new steel bars.

External steel plates. Reinforced concrete members, primarily beams, may be strengthened by attaching steel plates to their external surfaces. The basic technique consists of bonding steel plates to the concrete members with epoxy resins. The purpose is to increase the effective steel reinforcement of the reinforced concrete member.

Introducing new members. A reinforced concrete bridge can be strengthened by adding new members to the structure. These new members may be reinforced concrete or steel members. An example is the addition of new girders, beams, or stringers to a bridge deck system. When this strengthening technique is used, it is important to provide adequate details so that the new member will interact effectively within the existing structure.

External prestressing. Prestressing of concrete structures is a technique not only used to strengthen existing structures but also widely used in the construction of new bridges. This technique consists of high-strength steel twisted wire (or tendons) that are tensioned and anchored to the existing concrete members. The tension in the tendons is transferred to the concrete, generating forces and stresses that counteract the effects of applied loads. Existing bridges may be strengthened with this technique by modifying the tension in existing tendons or by introducing new tendons.

Column shells and wrapping. A technique widely used to strengthen bridges in primarily seismic areas consists in externally applying a steel (or composite) shell around the perimeter of columns or piers (**Fig. 12**). Steel shells are molded around the perimeter of columns, and the resulting gap is filled with a grout material. This shell provides additional strength and confinement to the concrete columns and increases the ability of the column to undergo extensive deflections without failure. In addition to steel shells,

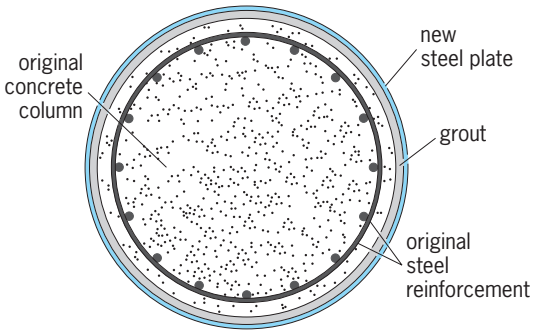

Fig. 12. Reinforcement of a circular reinforced concrete column using a steel shell.

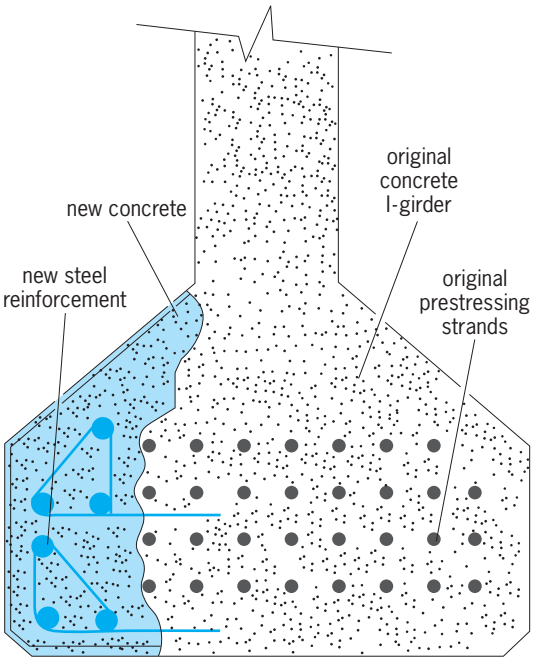

Fig. 13. Repair of a concrete I-girder by replacement of damaged concrete.

a fiber (graphite or glass)–reinforced composite can also be wet-wrapped around the column to achieve the same purpose. *See* COMPOSITE MATERIAL.

Repair techniques (concrete). Numerous repair techniques have evolved for concrete members in both bridges and buildings.

Replacement of damaged concrete. This method involves removing damaged concrete until a sound concrete surface is attained. This can be accomplished without damaging the existing steel reinforcement bars. After the damaged concrete is removed, the exposed surface is cleaned to attain an effective bonding surface, a primer is applied, and then new concrete is placed. Additional reinforcement may also be applied (**Fig. 13**).

Crack repair. Cracks that affect the structural integrity of a concrete member may be repaired by injection of materials such as cement grout (wide cracks) or epoxy resin. Cracks may both reduce the strength of elements and allow penetration of moisture to the steel bars, which will cause accelerated corrosion. Only stationary or nonpropagating cracks can be injected. Propagating cracks may require additional strengthening of the member to stop crack propagation. Crack injection with epoxy or grout does not increase the strength of the member; it may only restore its original integrity.

Corrosion of steel bars. Reinforced concrete members with excessive corrosion of the reinforcing steel bars may be repaired by removing the concrete in the affected area and restoring damaged steel. After exposing the corroded steel, the rust may be removed by sandblasting or wire-brushing. A corrosion protection coating may be applied to the reinforcement before placing the new concrete. If corrosion is too extensive, new steel reinforcement may have to be

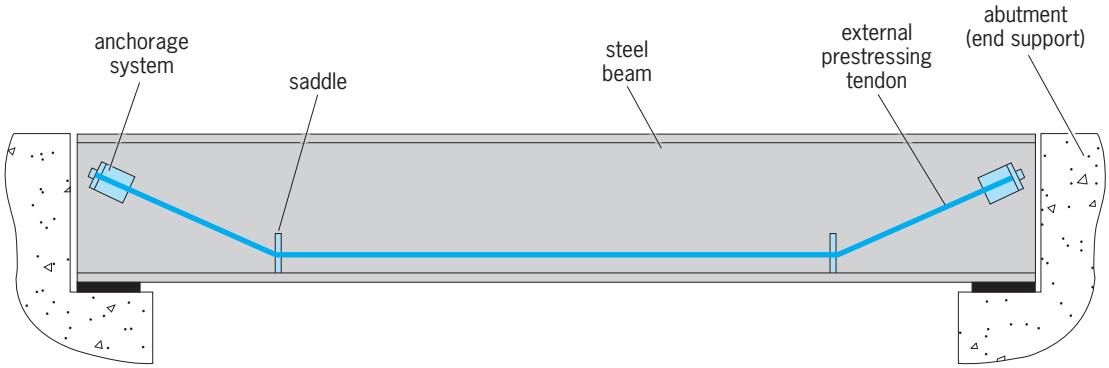

Fig. 14. Application of external prestressing force to strengthen a steel girder.

installed by either welding or mechanical attachment of new steel bars.

Strengthening techniques (steel). Steel bridges are most often strengthened by the addition of new steel members or smaller elements. Steel welding and bolting are well-developed techniques for steel connections. Thus, strengthening of steel bridges is perhaps more defined than for the concrete bridges.

External prestressing. Steel members, such as girders or truss members, can be strengthened by applying an external prestressing force (**Fig. 14**). The technique uses the same concept as described above for concrete members.

Adding supplementary members. Steel bridge structures or components can be strengthened by adding new members to the system at predetermined locations such as new girders, floor beams, or stringers. The new members are attached to the existing structure by connections that are designed to provide an effective load transfer without adversely affecting the integrity of the existing members.

Strengthening with external new plates. Steel girders or beams can be strengthened by increasing their section with the addition of plates or similar elements. An example is the addition of steel plates, usually by welding, to the flanges of an I-beam (**Fig. 15**).

Repair techniques (steel). Numerous techniques have evolved for the repair of steel bridge elements.

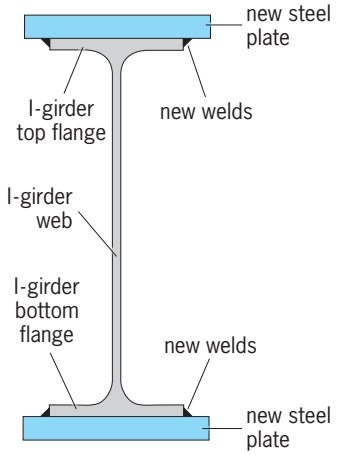

Fig. 15. Strengthening of a steel I-girder by welding new steel plates to the flanges.

Flame straightening. Steel members, such as I-girders or truss members damaged by vehicular impact, can be straightened using the flame straightening technique. Heat is applied at predetermined locations at and near the damaged region. The member is restrained and prevented from deforming in at least one direction while temperature is increased. The member is heated until permanent deformation occurs, and then the member is allowed to cool unrestrained. This will result in a permanent change of shape of the steel member. This process is repeated, if necessary, until the desired shape is attained.

Hot mechanical straightening. In this process, heat is applied to all sides of a bent member. While the member is still hot, it is straightened by applying a predetermined force. This method is used only when the heat will not adversely affect the mechanical properties of the steel.

Cold mechanical straightening. This repair method consists of straightening a bent member by applying a force without the use of heat. Cold mechanical straightening may affect the physical properties of the materials and should be used with care.

Welding. Welding of steel may be used for several types of repairs, such as crack repair, welding replacement segments into place, and addition of new plates. Special attention is required when welding members subjected to tensile loads in order to prevent the formation of cracks due to repetitive loading.

Bolting. Damaged steel members may be repaired by replacing the damaged element with a new piece of steel fastened with high-strength bolts. A combination of bolting and welding methods is sometimes effective for total repair of a member.

Partial replacement. If damage is serious, repair may be impossible, and it may be necessary to partially replace a member. This method consists of removing the damaged element and replacing it with either a welded or a bolted steel insert. This repair method may be utilized where steel members contain excessively wrinkled plates, excessive deformations and bends, tears, and large cracks.

Complete replacement. Excessively damaged members may have to be replaced entirely with a new member. Such replacement requires extensive evaluation

of load redistribution while the member is being replaced.

Joseph M. Plecnik; Oscar Henriquez; Andrew Pugel

Bibliography. E. DeLong, *Landmark American Bridges*, 1993; Guidelines for Evaluation and Repair of Damaged Prestressed Concrete Bridge Members, *Nat. Coop. Highway Res. Prog. Rep.*, no. 271 and 280, 1984, 1985; G. A. Hool and W. S. Kinne, *Movable and Long Span Steel Bridges*, 1943; D. Plowden, *Bridges: The Spans of North America*, 1974; V. K. Raina, *Concrete Bridges: Inspection, Repair, Strengthening, Testing and Load Capacity Evaluation*, McGraw-Hill, 1994; S. M. Shaker and R. D. Wakefield, *Modular Steel Bridges*, 1995; J. Toneas, *Bridge Engineering*, 1995; J. A. L. Waddell, *Bridge Engineering*, 1916; C. S. Whitney, *Bridges: A Study in Their Art, Science and Evolution*, 1929.

Bridge circuit

A circuit composed of a source and four impedances that is used in the measurement of a wide range of physical quantities. The bridge circuit is useful in measuring impedances (resistors, capacitors, and inductors) and in converting signals from transducers to voltage or current signals. *See* CAPACITOR; ELECTRICAL IMPEDANCE; INDUCTOR; RESISTOR; TRANSDUCER; WHEATSTONE BRIDGE.

The bridge impedances $Z_n = |Z_n|\angle\theta_n$, where $1 \le n \le 4$, are shown in **Fig. 1**. (Here, phasor notation is used: $|Z_n|$ is the amplitude of the impedance Z_n, and θ_n is its phase.) They may be single impedances (resistor, capacitor, or inductor), combinations of impedances, or a transducer with varying impedance. For example, strain gages are resistive transducers whose resistance changes when they are deformed. *See* ALTERNATING-CURRENT CIRCUIT THEORY; STRAIN GAGE.

Bridge circuits are often used with transducers to convert physical quantities (temperature, force displacement, pressure) to electrical quantities (voltage and current). High-accuracy voltmeters and ammeters are relatively inexpensive, and the voltage form of a signal is usually most convenient for information display, control decisions, and data storage.

Another important advantage of the bridge circuit is that it provides greater measurement sensitivity than the transducer. When a strain gage resistance increases from 120 to 120.10 ohms, the percentage change is small (less than 0.1%). If the same strain gage is placed in a bridge with 120-ohm resistors and a 10-V source, the output read by the meter changes from 0 to 2.082 mV. The circuit has two advantages: the percentage change may be very large (if the initial output was close to zero), and the small output voltage is more easily read. Much more precision is required of a meter to resolve small differences in larger voltages than to resolve small signals near zero.

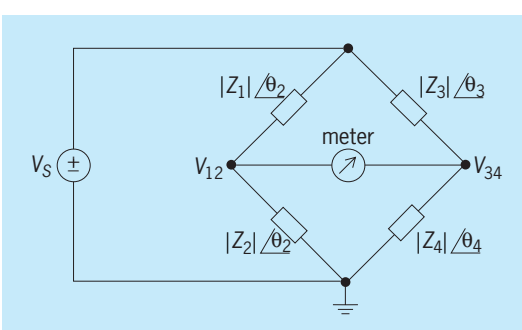

Fig. 1. Bridge circuit with source and impedances.

Balance condition. The bridge circuit is balanced when the output read by the meter is zero. In this condition the voltages on both sides of the meter are identical, as shown in Eqs. (1) and (2), and the

$$V_{12} = \frac{V_S Z_2}{Z_1 + Z_2} \tag{1}$$

$$V_{34} = \frac{V_S Z_4}{Z_3 + Z_4} \tag{2}$$

equality $V_{12} = V_{34}$ implies that the output is zero and that Eqs. (3) and (4), which are known as the balance

$$|Z_1| \cdot |Z_4| = |Z_2| \cdot |Z_3| \tag{3}$$

$$\theta_1 + \theta_4 = \theta_2 + \theta_3 \tag{4}$$

condition, are satisfied. The measurement sensitivity is highest when the bridge circuit is nearly balanced, because the values read on the meter are near zero.

Measurement methods. The bridge is used in two forms. The null adjustment method requires adjustment of a calibrated impedance to balance it. In this case the meter is usually a highly sensitive current-measuring galvanometer. The null adjustment method is often used to measure impedances, with the output read from a dial attached to the adjustable impedance. The deflection method requires an accurate meter in the bridge to measure the deviation from the balance condition. Assuming the meter has good linearity, the deviation (change in $V_{12} - V_{34}$) is proportional to the quantity being measured.

Special forms. There are many special forms of the bridge circuit. When all of the impedances are resistive and the bridge is operated in direct-current mode, it is commonly called a Wheatstone bridge. Other common forms use a current source in place of the voltage source, a sinusoidal source in place of a constant (dc) source, or branch impedances which are specific combinations of single passive impedances, for example, the Hayes bridge and the Owen bridge. The bridge circuit is also used in a variety of electrical applications varying from oscillators (Wien bridge oscillator) to instrumentation amplifier circuits for extremely accurate measurements. *See* INSTRUMENTATION AMPLIFIER; OSCILLATOR; RESISTANCE MEASUREMENT.

Bridge circuits based on transformers. It is possible to replace impedances Z_1 and Z_2 in Fig. 1 with either a voltage or a current transformer to provide instead an equivalent ratio of voltage or current. The remaining two impedances (Z_3 and Z_4) then become the

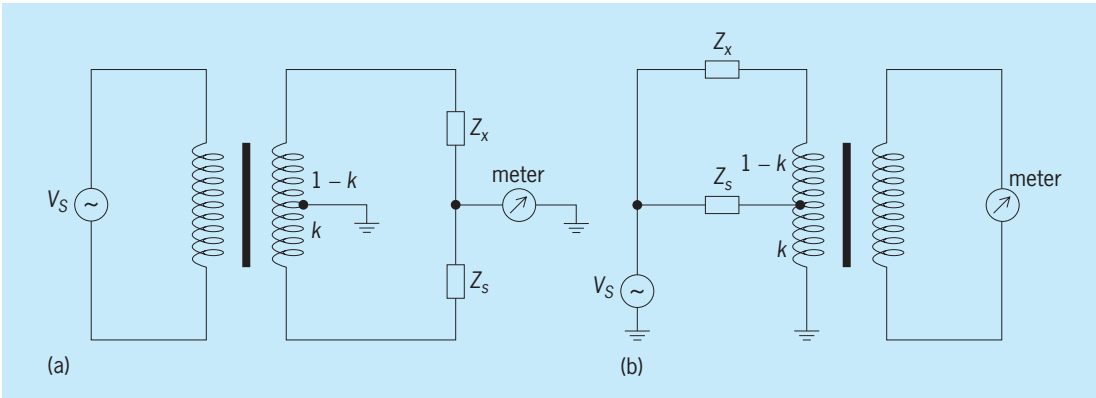

Fig. 2. Comparison of similar types of impedances based on (*a*) voltage ratio transformer and (*b*) current ratio transformer.

unknown (Z_x) and the reference (Z_s) impedances. Knowing the ratio of the voltage or current transformer and the value of Z_s then enables Z_x to be determined. The ratio of the transformer is simply given by the number of turns on the two output arms of the transformer. Although construction of the transformers is time-consuming, their voltage or current ratios can be better than a few parts-per-million accuracy, providing care is taken with the electrostatic and magnetic shielding between the primary and secondary windings of the transformer. These techniques are employed mostly by the national metrology institutes for the highest-accuracy impedance measurements.

Figure 2*a* shows a comparison of the reference impedance with the unknown impedance based on a voltage ratio transformer. The unknown impedance is then given by Eq. (5), where k is the ratio of

$$Z_x = Z_s\left(\frac{k}{1-k}\right) \tag{5}$$

the transformer windings, and lies in the range $0 \le k \le 1$. Similarly, Fig. 2*b* shows the equivalent circuit for comparing the two impedances based on a current transformer. The detector is nulled when equal and opposite flux (ampere-turns) is generated in the current transformer core by the two currents flowing through the impedances. The balance equation in this case also remains identical to that given in Eq. (5). A key difference between the two circuits is the arrangement of the source and detector. In Fig. 2*a* the source is isolated from the main bridge network, whereas the null detector is connected to the junction of the two impedances. In contrast, in Fig. 2*b* the source is directly connected to the main bridge network and the null detector is isolated. Finally, it is important to note that the circuits shown in Fig. 2 operate only with alternating current. *See* TRANSFORMER. Shakil A. Awan; Kirk D. Peterson

Bibliography. E. Barsoukov and J. R. Macdonald (eds.), *Impedance Spectroscopy: Theory, Experiment, and Applications*, 2d ed., 2005; L. K. Baxter, *Capacitive Sensors: Design and Applications*, 1997; T. G. Beckwith, R. D. Marangoni, and J. H. Lienhard V, *Mechanical Measurements*, 6th ed., 2007; E. O. Doebelin, *Measurement Systems: Application and Design*, 5th ed., 2003; J. P. Holman, *Experimental Methods for Engineers*, 7th ed., 2000; E. J. Kennedy, *Operational Amplifier Circuits: Theory and Applications*, 1988; A. K. Walton, *Network Analysis and Practice*, 1987.

Brillouin zone

In the propagation of any type of wave motion through a crystal lattice, the frequency is a periodic function of wave vector **k**. This function may be complicated by being multivalued; that is, it may have more than one branch. Discontinuities may also occur. In order to simplify the treatment of wave motion in a crystal, a zone in **k**-space is defined which forms the fundamental periodic region, such that the frequency or energy for a **k** outside this region may be determined from one of those in it. This region is known as the Brillouin zone (sometimes called the first or the central Brillouin zone). It is usually possible to restrict attention to **k** values inside the zone. Discontinuities occur only on the boundaries. If the zone is repeated indefinitely, all **k**-space will be filled. Sometimes it is also convenient to define larger figures with similar properties which are combinations of the first zone and portions of those formed by replication. These are referred to as higher Brillouin zones.

The central Brillouin zone for a particular solid type is a solid which has the same volume as the primitive unit cell in reciprocal space, that is, the space of the reciprocal lattice vectors, and is of such a shape as to be invariant under the symmetry operations of the crystal. *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Zone construction. Let $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ be the primitive translation vectors for some crystal lattice. New vectors $\mathbf{b}_i$ with $i = 1, 2, 3$, are defined by Eq. (1), where

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij} \tag{1}$$

δ_{ij} is unity when $i = j$, and zero for other values of i.

The vectors $\mathbf{b}_i$ are then given by Eqs. (2).

$$\mathbf{b}_1 = \frac{2\pi\,\mathbf{a}_2\times\mathbf{a}_3}{\mathbf{a}_1\cdot\mathbf{a}_2\times\mathbf{a}_3} \quad \mathbf{b}_2 = \frac{2\pi\,\mathbf{a}_3\times\mathbf{a}_1}{\mathbf{a}_1\cdot\mathbf{a}_2\times\mathbf{a}_3} \quad \mathbf{b}_3 = \frac{2\pi\,\mathbf{a}_1\times\mathbf{a}_2}{\mathbf{a}_1\cdot\mathbf{a}_2\times\mathbf{a}_3} \tag{2}$$

Now vectors $\mathbf{K}_i$ are defined by Eq. (3), where the

$$\mathbf{K}_i = h_1\mathbf{b}_1 + h_2\mathbf{b}_2 + h_3\mathbf{b}_3 \tag{3}$$

h_i's are arbitrary integers. The subscript i stands for some particular combination of the h. The end points of the vectors $\mathbf{K}_i$ form a lattice of points in reciprocal space. The vectors $\mathbf{K}_i$ have the property that plane waves of the form $e^{i\mathbf{K}_i\cdot\mathbf{r}}$ are periodic in the crystal lattice, since if some $\mathbf{r}' = \mathbf{r} + \mathbf{R}_n$ is considered, where $\mathbf{R}_n$ is a translation vector of the crystal, Eq. (4) holds.

$$e^{i\mathbf{K}_i\cdot\mathbf{r}'} = e^{i\mathbf{K}_i\cdot\mathbf{r}}e^{i\mathbf{K}_i\cdot\mathbf{R}_n} = e^{i\mathbf{K}_i\cdot\mathbf{r}} \tag{4}$$

The last step follows since $\mathbf{K}_i \cdot \mathbf{R}_n$ is an integer times 2π. Consequently, the plane waves $e^{i\mathbf{K}\cdot\mathbf{r}}$ are suitable functions for the Fourier expansion of any function which is periodic in the lattice.

Unit cells can be constructed in the reciprocal space of the lattice of the ends of the vectors $\mathbf{K}_i$, just as is done in the real crystal space. The lines connecting one point with the other lattice sites are drawn, and the planes which are the perpendicular bisectors of these lines are constructed. The smallest enclosed solid figure is the first Brillouin zone. It is the smallest unit cell in the reciprocal lattice which has the symmetry of the entire lattice. Higher Brillouin zones are also formed.

Brillouin zones for the body-centered-cubic, face-centered-cubic, and hexagonal-close-packed lattices are shown in **illus.** *a–c*.

Application to band theory. Each electron wave function in the crystal can be classified according to some $\mathbf{k}$ inside the first Brillouin zone. For if $\mathbf{k}'$ is a vector in reciprocal space whose end point lies outside the zone, then it can be written as Eq. (5), where $\mathbf{k}$

$$\mathbf{k}' = \mathbf{k} + \mathbf{K}_n \tag{5}$$

lies inside the zone and $\mathbf{K}_n$ is a reciprocal lattice vector. According to the Bloch theorem, Eq. (6*a*) holds.

(a) (b) (c)

Brillouin zone for (*a*) body-centered-cubic lattice, (*b*) face-centered-cubic lattice, (*c*) hexagonal-close-packed lattice.

But since $\mathbf{K}_n \cdot \mathbf{R}_j$ is an integer times 2π, Eq. (6*b*) is

$$\psi(\mathbf{k}', \mathbf{r} + \mathbf{R}_j) = e^{i\mathbf{k}'\cdot\mathbf{R}_j}\psi(\mathbf{k}', \mathbf{r}) \tag{6a}$$

$$e^{i\mathbf{k}'\cdot\mathbf{R}_j} = e^{i\mathbf{k}\cdot\mathbf{R}_j}e^{i\mathbf{K}_n\cdot\mathbf{R}_j} = e^{i\mathbf{k}\cdot\mathbf{R}_j} \tag{6b}$$

valid. Thus $\mathbf{k}'$ and $\mathbf{k}$ are equivalent in a certain sense. For this reason, it is possible to consider only the first zone in discussing the properties of solids.

A fundamental point in the application of Brillouin zone theory to the study of the properties of solids is that the energy $E(\mathbf{k})$ of the states ψ_k must be a continuous function of $\mathbf{k}$ inside the zone (although it will be multivalued if the reduced zone scheme is employed). Discontinuities can occur only across the faces of the zone. The number of allowed states inside each Brillouin zone can be determined as follows: The number of allowed states per unit volume in the space of the vector $\mathbf{k}$ is $1/(2\pi)^3$ for each spin per unit volume of the crystal, or $2/(2\pi)^3$ altogether. Thus, the number of states in the zone for each atom in the crystal is $2V/(2\pi)^3N$, where N is the number of atoms of the crystal and V is the volume. A substance which has just enough electrons per atom to fill some zone may be an insulator or semiconductor. If there are not enough electrons to fill a zone, it must be a metallic conductor.

Alkali metals. These metals are body-centered cubic at high temperatures. (Sodium and lithium undergo phase transitions at low temperatures.) The volume of the zone is $2(2\pi/a)^3$, where a is the lattice constant. The number of atoms per unit volume is $2/a^3$ so that the first zone can contain two electrons per atom. However, the atoms possess only one valence electron, so that these metals are good conductors.

Noble metals. Copper, silver, and gold are face-centered cubic metals with one valence electron per atom, and again the zone can hold two electrons per atom. Thus, these also are good conductors. If noble metals were adequately described by a free electron model, the Fermi surface (the surface bounding the occupied volume of reciprocal space) would be a sphere which would approach close to the surface of the zone. It is known, however, that the free electron model is not adequate for electrons near the Fermi surface, and that the Fermi surface touches the Brillouin zone boundary near the extremities of the (111) axes.

Divalent metals. Except for barium and mercury, the divalent metals have either hexagonal close-packed or face-centered cubic structures. The zones hold two valence electrons per atom, and on the basis of simple arguments, it would be expected that these materials should be insulators. However, some of the electrons go over into the next Brillouin zone, leaving pockets of holes behind in the first zone.

Semiconductors and insulators. In diamond, silicon, and germanium, all of which have the diamond structure of two interpenetrating face-centered cubic (fcc) lattices, the Brillouin zone has the same form as for the fcc. It can hold four electrons per atom. Occupied states are separated from higher states by an energy gap, since these materials are semiconductors

or insulators. Graphite, which has a hexagonal layer structure, can be regarded as a semiconductor with a vanishing energy gap. The Brillouin zone in a simple model is full at a temperature of 0 K, but there is no gap along the edges of the zone. In bismuth there is a higher Brillouin zone which holds five electrons per atom. There is a very small overlap of electrons across the faces of the zone, leading to the presence of a small number (about 10^{-5} per atom) of both electrons and holes at 0 K. *See* SEMICONDUCTOR.

Binary alloys. Many of the properties of binary alloys which form substitutional solid solutions can be explained on the hypothesis that the solute contributes its valence electrons to form a composite system whose fundamental parameter is the electron per atom ratio. The electronic structure is assumed to be essentially unaltered. Many changes of the properties, such as the lattice parameter, the concentration at which the various phases appear, and the Hall coefficient, can be explained in terms of overlap or contact of the Fermi surface with the Brillouin zone. *See* ALLOY STRUCTURES; BAND THEORY OF SOLIDS. Joseph Callaway

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 2d ed., 2000; W. Hume-Rothery and R. E. Smallman, *The Structure of Metals and Alloys*, 1988; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; M. N. Rudden and J. Wilson, *Elements of Solid State Physics*, 2d ed., 1993.

Brittleness

The catastrophic fracture of a material or structure without a significant amount of prior extension or plastic deformation. The energy absorbed during fracture is highly dependent on whether a material or structure behaves in a brittle or ductile manner. Materials or structures that fracture in a brittle manner exhibit small amounts of plastic deformation and extension. This requires a much lower amount of energy in the fracture process compared to a material or structure that fails in a ductile manner. **Figure 1** shows the effect of a change in test temperature on the energy absorbed during the fracture of some materials, where the energy absorbed during fracture at low temperatures is significantly less than that absorbed at higher temperatures. Materials or structures in the low-temperature regime are failing via brittle fracture. Increasing the test temperature produces a significant increase in the fracture energy, where the materials behave in a much tougher manner and exhibit a more ductile fracture appearance. At intermediate temperatures, a brittle-to-ductile transition is shown. Brittleness and brittle fracture are exhibited by a number of materials, including glass, minerals, some polymers, cast iron, concrete, some refractory metals, most ceramic materials, and a number of different steels tested below their brittle-to-ductile transition temperature.

The fracture surfaces produced during brittle fracture typically exhibit characteristic features related

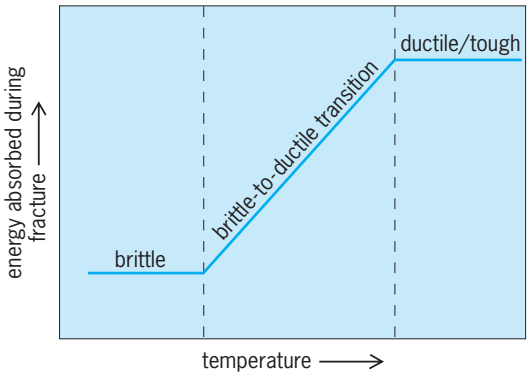

Fig. 1. Energy absorbed during fracture versus temperature. Brittle materials absorb much lower energy during fracture than ductile or tough materials. Some materials exhibit a brittle-to-ductile transition that depends on the test temperature.

to the fracture mechanisms that operate at a microscopic level. Brittle fracture that occurs in a crystalline metal (for example, iron) or alloy (for example, steel) typically occurs on a crystal plane of low atomic density, called the cleavage fracture plane. This fracture process occurs through the grains in the material, and is called transgranular cleavage fracture. This produces characteristic features that can be easily identified using a scanning electron microscope (SEM). An SEM can provide images with high resolution and high depth of field for magnifications exceeding 100,000×. An SEM image of a typical transgranular cleavage fracture surface of a steel tested at a temperature below the brittle-to-ductile transition temperature reveals some of these characteristic features (**Fig. 2**). The arrow shows the location of the suspected fracture origin, as all fracture surface features appear to radiate from this one site. Many of the features in Fig. 2 are less than 50 μm, the diameter of a human hair. The general lack of plastic deformation and elongation is reflected in the

Fig. 2. Scanning electron microscope image of a transgranular cleavage fracture surface in a steel tested at a temperature below its brittle-to-ductile transition. The arrow indicates the likely fracture initiation site.

brittle nature of the fracture surface observed at high magnification. Brittle fracture may also occur along the grain boundaries of certain materials, and this is called intergranular fracture. The fracture surface observed by eye will often exhibit a shiny appearance due to the reflection of light from the brittle features. In contrast, a material that fails in a ductile manner will exhibit a fracture surface with large amounts of elongation and plasticity. Ductile fractures examined at high magnification with an SEM typically exhibit dimples on the fracture surface and a dull gray appearance when viewed by eye. This dull gray appearance results from the diffraction of the light by the microscopic fracture features. In general, brittle fracture can be produced in susceptible materials by testing at low temperatures, at high (for example, impact) loading rates, and in the presence of stress concentrators such as notches, cracks, and defects. One example of this type of failure in a steel structure is the catastrophic fracture and sinking of the *Titanic* in 1912. *See* EMBRITTLEMENT; PLASTIC DEFORMATION OF METAL; PLASTICITY; SCANNING ELECTRON MICROSCOPE; STEEL; STRESS AND STRAIN.

John J. Lewandowski

Bibliography. *ASM Handbook*, vol. 19: *Fracture and Fatigue*, ASM International, 1996; R.W. Hertzberg, *Deformation and Fracture Mechanics of Engineering Materials*, 4th ed., Wiley, 1996; B. R. Lawn, *Fracture of Brittle Solids*, 2d ed., Cambridge University Press, 1993.

Broccoli

A cool-season biennial crucifer, *Brassica oleracea* var. *italica*, of Mediterranean origin, belonging to the plant order Papaverales. Broccoli is grown for its thick branching flower stalks which terminate in clusters of loose green flower buds (see **illus.**). Stalks and buds are cooked as a vegetable or may be processed in either canned or frozen form. Cultural practices for broccoli are similar to those used for cabbage. *See* CABBAGE; PAPAVERALES.

New varieties (cultivars) with greater disease resistance and higher quality are continually being developed. Broccoli is slightly tolerant of acid soils and has high requirements for boron and molybdenum. Terminal and axillary clusters are cut 80–140 days after planting. California and Texas are important broccoli-producing states.

H. John Carew

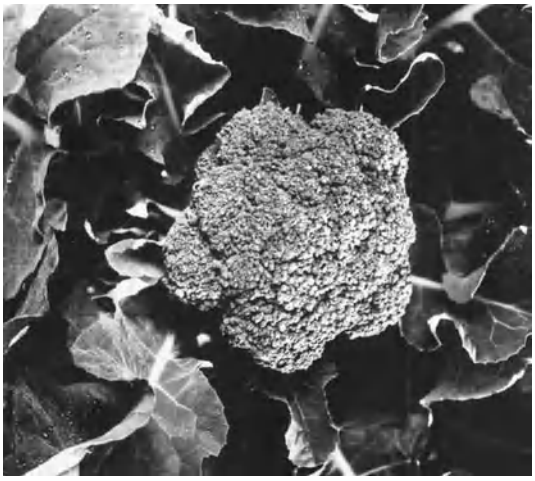

Flower buds of broccoli. (*Asgrow Seed Co., Subsidiary of the Upjohn Co.*)

Bromegrass

A common name designating a number of grasses found in the North Temperate Zone that produce highly palatable and nutritious forage. Of these, smooth bromegrass (*Bromus inermis*) is the most important (see **illus.**). This species was introduced to the United States during the 1880s from central Europe and Russia, where it is native. It proved to be well adapted to regions of cold winters and limited rainfall. Although first widely used in the eastern Great Plains and western Corn Belt regions, improved strains are now grown extensively for hay and rotation pastures north of the Mason-Dixon line, from the Plains to the Atlantic. Smooth bromegrass is a long-lived perennial, spreads by underground creeping stems, and is fairly deep rooted and drought-tolerant. Top growth is 2–4 ft (0.6–1.2 m) tall and is used for hay or pasture. Regional strains are available for Canada and the northern two-thirds of the United States. Bromegrass tolerates a wide range of soil conditions, but responds well to

Smooth bromegrass (*Bromus inermis*). (*a*) Entire plant. (*b*) Fruit. (*c*) Inflorescence.

higher levels of fertility. Where the moisture supply is adequate, smooth bromegrass grows well in mixtures with alfalfa and red clover. The seeds of bromegrass are large, light, and chaffy, and successful planting requires special seeders. *See* CYPERALES; GRASS CROPS. Howard B. Sprague

Bromeliales

An order of flowering plants, division Magnoliophyta (Angiospermae) in the subclass Zingiberidae of the class Liliopsida (monocotyledons). It consists of the single family Bromeliaceae, with about 45 genera and 2000 species, occurring chiefly in tropical and subtropical America. They are firm-leaved, terrestrial xerophytes (adapted structurally to live and grow with a limited water supply), or very often epiphytes (living on other plants nonparasitically), with six stamens and regular or somewhat irregular flowers that usually have septal nectaries and an inferior ovary. The order has often been associated with several related ones in a larger order, Farinosae, marked by its starchy endosperm. Spanish moss (*Tillandsia*) and the cultivated pineapple (*Ananas*) are familiar members of the Bromeliales, and many others attract attention as houseplants. *See* FLOWER; LILIOPSIDA; MAGNOLIOPHYTA; PIÑA; PINEAPPLE; PLANT KINGDOM; SECRETORY STRUCTURES (PLANT). Arthur Cronquist; T. M. Barkley

Bromine

A chemical element, Br, atomic number 35, atomic weight 79.909, which normally exists as Br_2, a dark-red, low-boiling but high-density liquid of intensely irritating odor. This is the only nonmetallic element that is liquid at normal temperature and pressure. Bromine is very reactive chemically; one of the halogen group of elements, it has properties intermediate between those of chlorine and iodine. *See* HALOGEN ELEMENTS; PERIODIC TABLE.

The most stable valence states of bromine in its salts are −1 and +5, although +1, +3, and +7 are known. Within wide limits of temperature and pressure, molecules of the liquid and vapor are diatomic, Br_2, with a formula weight of 159.818. There are two stable isotopes (^{79}Br and ^{81}Br) that occur naturally in nearly equal proportion, so that the atomic weight is 79.909. A number of radioisotopes are also known.

1												13	14	15	16	17	18
1 H	2																2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Physical properties of bromine

Property			Value
Flash point			None
Fire point			None
Freezing point, °C			−7.27
Density, 20°C			3.1226
Pounds per gallon, 25°C			25.8
Boiling point, 760 mm Hg, °C			58.8
Refractive index, 20°C			1.6083
Latent heat of fusion, cal/g			15.8
Latent heat of vaporization, cal/g, bp			44.9
Vapor density, g/liter, standard conditions (0°C, 1 atm)			7.139
Viscosity, centistokes, 20°C			0.314
Surface tension, dynes/cm, 20°C			49.5
30°C			47.3
40°C			45.2
Dielectric constant, 10^5 freq, 25°C			3.33
Compressibility, vapors, 25°C			0.998
Thermodynamic data, cal/(mole K)			
	T,K	*Entropy*	*Heat capacity*
Solid	265.9	24.786	14.732
Liquid	265.9	34.290	18.579

The solubility of bromine in water at 20°C (68°F) is 3.38 g/100 g (3.38 oz/100 oz) solution, but its solubility is increased tremendously in the presence of its salts and in hydrobromic acid. The ability of this inorganic element to dissolve in organic solvents is of considerable importance in its reactions. The **table** summarizes the physical properties of bromine.

Although it is estimated that from 10^{15} to 10^{16} tons of bromine are contained in the Earth's crust, the element is widely distributed and found only in low concentrations in the form of its salts. The bulk of the recoverable bromine, however, is found in the hydrosphere. Sea water contains an average of 65 parts per million (ppm) of bromine. The other major sources of bromine in the United States are underground brines and salt lakes, with commercial production in Michigan, Arkansas, and California.

While many inorganic bromides have found industrial use, the organic bromides have even wider application. Because of the ease of reaction of bromine with organic compounds and the ease of its subsequent removal or replacement, organic bromides have been much studied and used as chemical intermediates. In addition, any of the bromine reactions are so clean-cut that they can be used for the study of reaction mechanisms without complication of side reactions. The ability of bromine to add into unusual places on organic molecules has added to its value as a research tool.

Bromine and its compounds have found acceptance as disinfection and sanitizing agents in swimming pools and potable water. Certain bromine-containing compounds are safer to use than the analogous chlorine compounds due to certain persistent residuals found in the chlorine-containing materials. Other bromine chemicals are used as a working fluid in gages, as hydraulic fluids, as chemical intermediates in the manufacture of organic dyes, in storage batteries, and in explosion-suppressant and fire-extinguishing systems. Bromine compounds, because of their density, also find use in the gradation

of coal and other minerals where separations are effected by density gradients. The versatility of bromine compounds is illustrated by the commercial use of over 100 compounds that contain bromine.

Bromine is almost instantaneously injurious to the skin, and it is difficult to remove quickly enough to prevent a painful burn that heals slowly. Bromine vapor is extremely toxic, but its odor gives good warning; it is difficult to remain in an area of sufficient concentration to be permanently damaging. Bromine can be handled safely, but the recommendations of the manufacturers should be respected.

Randy C. Stauffer

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. Price, B. Iddon, and B. J. Wakefield (eds.), *Bromine Compounds: Chemistry and Applications*, 1988; D. F. Shriver and P. W. Atkins, *Inorganic Chemistry*, 3d ed., 1999.

Bronze

Usually an alloy of copper and tin. Bronze is used in bearings, bushings, gears, valves, and other fittings both for water and steam. Lead, zinc, silver, and other metals are added for special-purpose bronzes. Tin bronze, including statuary bronze, contains 2–20% tin; bell metal 15–25%; and speculum metal up to 33%. Gun metal contains 8–10% tin plus 2–4% zinc.

The properties of bronze depend on its composition and working. Phosphor bronze is tin bronze hardened and strengthened with traces of phosphorus; it is used for fine tubing, wire springs, and machine parts. Lead bronze may contain up to 30% lead; it is used for cast parts such as low-pressure valves and fittings. Manganese bronze with 0.5–5% manganese plus other metals, but often no tin, has high strength. Aluminum bronze also contains no tin; its mechanical properties are superior to those of tin bronze, but it is difficult to cast. Silicon bronze, with up to 3% silicon, casts well and can be worked hot or cold by rolling, forging, and similar methods. Beryllium bronze (also called beryllium copper) has about 2% beryllium and no tin. The alloy is hard and strong and can be further hardened and strengthened by precipitation hardening; it is one of the few copper alloys that responds to heat treatment, approaching three times the strength of structural steel.

Tin bronze is harder, stronger in compression, and more resistant to corrosion than the brasses. Bronze that will not be exposed to extremes of weather can be protected from corrosion by warming it to slightly over 212°F (100°C) in an oxygen atmosphere. A thin layer of oxide or patina forms to prevent further oxidation. A patina may be formed on art objects by exposure first to acid fumes and then drying as above. While still warm, the object can be further protected by a spray of wax in a solvent.

For bearings, sintered bronze is compacted from 10% tin, up to 2% graphite, and the balance by weight of copper. The aggregate is formed under pressure at a temperature below the melting point of its constituents but high enough to reduce their oxides. After forming, a bearing is repressed or sized and impregnated with oil, the pores retaining the lubricant until needed. In place of copper, alpha-bronze powder may be used; zinc may replace some of the tin. At forming temperatures below 1290°F (700°C), properties depend primarily on compacting pressure. At higher temperatures, properties depend first on temperature, although heat treatment beyond 30 min has minor influence. *See* ALLOY; ANTIFRICTION BEARING; BELL; COPPER; COPPER ALLOYS; TIN; TIN ALLOYS.

Frank H. Rockett

Brown dwarf

A starlike body whose mass is too small to sustain nuclear fusion reactions in its core. All stars, including the Sun, shine because they engage in nuclear fusion in their hot and dense cores. In the early 1960s, S. S. Kumar noted that, if they existed, stars with mass less than 8% that of the Sun would not have the high temperatures in their cores necessary to sustain nuclear fusion reactions. These objects, called brown dwarfs, would not truly be normal stars because their lack of nuclear fusion would inhibit their ability to shine. Indeed, brown dwarfs would grow dimmer as they aged. At even lower masses are the planets, such as Earth and Jupiter (which is approximately 0.1% the mass of the Sun). The distinction between planets and brown dwarfs has been debated; as of early 2006, no consensus on the scientific definition of the word "planet" had been reached. Some scientists maintained that planets ought to be distinguished from brown dwarfs by the way they form. Planets form in disks of dust and gas swirling about nascent stars, while brown dwarfs are thought to form like stars, out of the gravitational collapse of a huge cloud of gas in space. However, a basic physical distinction is that brown dwarfs are hot enough when they are young to host evanescent nuclear reactions while planets never host any fusion reactions. *See* PLANET; STAR; STELLAR EVOLUTION.

Largely because they did not know what these objects would look like, astronomers agreed in the 1970s to call them brown dwarfs. Brown is a mixture of many colors and represents the lack of agreement on the dwarf's appearance. Other proposed names include black dwarfs, infrared dwarfs, failed stars, super-Jupiters, and substellar objects.

Since the 1960s, brown dwarfs have commanded varying levels of importance in astronomical research. They have been invoked to explain the dark-matter problem of cosmology, while many astronomers believed that brown dwarfs did not exist because none had been found. In 1995, the first definitive detection of a cool brown dwarf and also the detection of three young brown dwarfs were announced. By the end of 2003, over 100 of varying ages and masses had been found, and two new spectral classes had been added to the stellar classification scheme for the first time in over 50 years. Brown dwarf research has become a mature and rich subfield of astrophysics. *See* COSMOLOGY.

Properties. Calculations by several research groups have established that the lowest-mass star (8% the mass of the Sun) will shine with a luminosity of about 10^{-4} times the luminosity of the Sun. (Luminosity here means the energy emitted per unit of time.) While young brown dwarfs can have luminosities larger than this value, they eventually cool to much smaller ranges of brilliance. These calculations also show that all brown dwarfs have essentially the same radius, about 10% that of the Sun. This is also approximately the radius of Jupiter. *See* RED DWARF STAR.

Young, hot brown dwarfs (L dwarfs). Because young brown dwarfs are hot and even more luminous than the oldest, lowest-mass stars, they are very difficult to distinguish from such stars. However, an important diagnostic exists. The fragile element lithium is transmuted in stars by high-temperature fusion reactions that are absent in most brown dwarfs' cores. Indeed, brown dwarfs below about 6% the mass of the Sun never transmute any lithium. Those between 6 and 8% the mass of the Sun transmute some or even all of the lithium because they can host the stellar nuclear fusion reactions, but only for the first 50–250 million years of their lives. All this suggests that although the youngest brown dwarfs might look identical to low-mass stars they will exhibit signs of lithium, which stars will not.

Two searches for brown dwarfs in the Pleiades, one of the youngest nearby star clusters, were undertaken in 1995. Sensitive spectroscopic observations with the Keck Telescope in Hawaii revealed for the first time the telltale features of lithium in three low-luminosity objects. These objects, the first young brown dwarfs identified, were called PPL 15, Teide 1, and Calar 3. *See* PLEIADES; TELESCOPE.

Since these discoveries, many more brown dwarfs have been identified in the Pleiades, in other star clusters, and in interstellar space. The majority of these brown dwarfs have been found due to improved astronomical imaging and spectroscopy technology, in particular, large-scale surveys of the sky in the infrared wavelengths. *See* ASTRONOMICAL SPECTROSCOPY; INFRARED ASTRONOMY.

Brown dwarfs slightly older than those found with the lithium criterion are now classified as L dwarfs, a spectral class lower in temperature than the M class, which is reserved for the lowest-temperature stars. Many examples of L dwarfs exist. The letter L has no particular significance except that, when the class was specified, L and T were among the few letters of the alphabet that were not already in use to designate classes of starlike celestial objects. The L class is generally characterized by temperatures below 1800 K (2780°F) and above 1400 K (2060°F). Spectra of L dwarfs reveal the complex chemistry present in these brown-dwarf atmospheres. Many molecules, such as chromium hydride, carbon monoxide, carbon dioxide, and water, are clearly identifiable.

L dwarfs exist in great abundance in the Milky Way Galaxy and may be twice as numerous as the number of stars. However, because they are not very massive, the population of L dwarfs in the galaxy cannot account for the "missing" or dark matter. *See* COSMOLOGY; MILKY WAY GALAXY.

Cool, old brown dwarfs (T dwarfs). In 1995, Tadashi Nakajima and coworkers at the Palomar Observatory, California Institute of Technology, discovered an object in orbit around the nearby star Gliese 229, located only 17 light-years (1.0×10^{14} mi or 1.6×10^{14} km) from the Sun (see **illus.**). With a luminosity of 6×10^{-6} that of the Sun, this object was the first unambiguous discovery of an old brown dwarf, one that had absolutely no similarity to any star. The discovery of Gliese 229B was the result of a systematic survey of nearby stars that used a camera tailored for imaging faint objects next to bright stars.

Gliese 229B has become the prototypical example of the newly invented T spectral class. Dozens of other T dwarfs have now been identified, primarily due to the work of the 2MASS infrared survey of the entire sky and spectroscopic observations

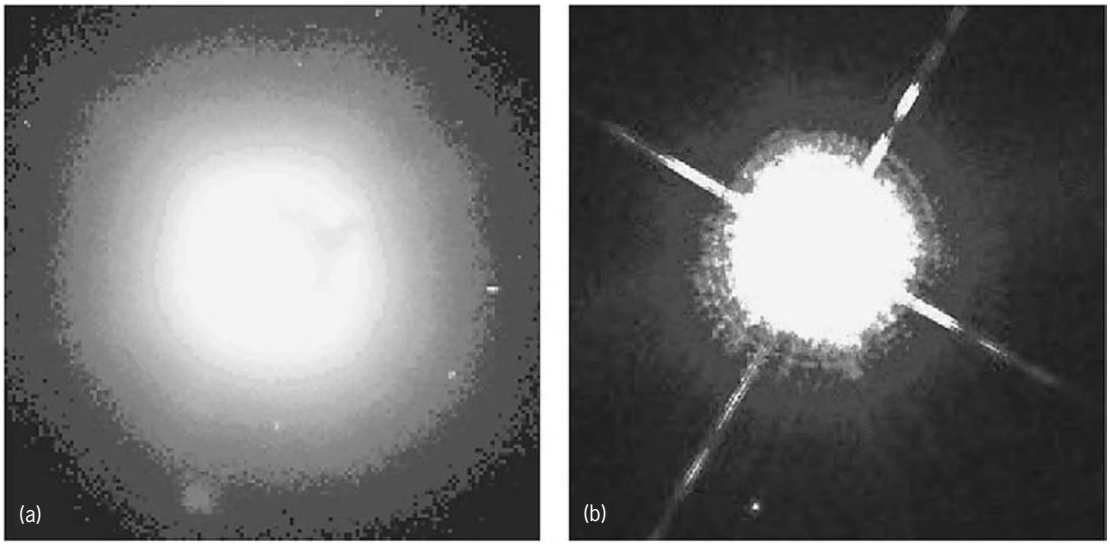

Images of the cool brown dwarf Gliese 229B. In each image, the bright primary star is at the center and the brown dwarf is the faint object near the bottom center. (*a*) Discovery image, taken with the Palomar 60-in. (1.5-m) telescope and an instrument designed to reduce the substantial glare of the nearby star. (*b*) Hubble Space Telescope image.

spearheaded by J. Davy Kirkpatrick and Adam Burgasser. The L dwarfs are believed to evolve into T dwarfs over a time span of around a billion years. The T dwarfs are distinct from the L dwarfs because they have telltale signs of methane in their spectra. The presence of methane in the atmospheres of these brown dwarfs indicates that they must be cooler than about 1200–1400 K (1700–2060°F). T dwarfs as cool as 700 K (800°F) had been identified by the end of 2003. Ben R. Oppenheimer

Bibliography. A. Burgasser et al., The spectra of T dwarfs, *Astrophys. J.*, 564:421–451, 2002; J. Davy Kirkpatrick et al., Dwarfs cooler than "M": The definition of spectral type "L" using discoveries from the Two Micron All Sky Survey (2MASS), *Astrophys. J.*, 519:802–833, 1999; T. Nakajima et al., Discovery of a cool brown dwarf, *Nature*, 378:463–465, 1995; B. R. Oppenheimer et al., Infrared spectrum of the cool brown dwarf Gl 229B, *Science*, 270:1478–1479, 1995; R. Rebolo, M. R. Zapatero-Osorio, and E. L. Martín, Discovery of a brown dwarf in the Pleiades star cluster, *Nature*, 377:129–131, 1995.

Brownian movement

The irregular motion of a body arising from the thermal motion of the molecules of the material in which the body is immersed. Such a body will of course suffer many collisions with the molecules, which will impart energy and momentum to it. Because, however, there will be fluctuations in the magnitude and direction of the average momentum transferred, the motion of the body will appear irregular and erratic, as shown in the **illustration**.

In principle, this motion exists for any foreign body suspended in gases, liquids, or solids. To observe it, one needs first of all a macroscopically visible body; however, the mass of the body cannot be too large. If its mass is M, one can estimate the root-mean-square velocity $\overline{v}$ by the equipartition law, written as Eq. (1). Here, k is the Boltzmann constant

$$\begin{aligned} {}^1/_2 M\overline{v^2} &= {}^3/_2 kT \\ \sqrt{\overline{v^2}} &= \sqrt{\frac{3kT}{M}} \end{aligned} \tag{1}$$

and T is the absolute temperature. Hence, for a large mass, the velocity becomes small. For example, a 0.01-g mass has at 300 K a root-mean-square velocity of 3.5×10^{-6} cm/s. *See* KINETIC THEORY OF MATTER.

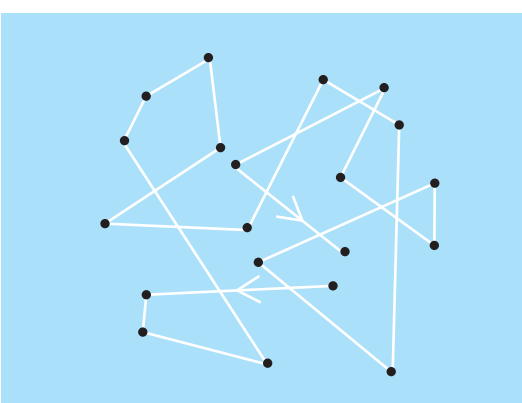

Random brownian movement of a particle.

Mirror in a gas. A simple observable example of brownian motion is the motion of a mirror in a gas. If a mirror of moment of inertia I is suspended by a fine fiber, this mirror will execute a simple harmonic motion of natural frequency ω_0. Its mechanical equation of motion is written as Eq. (2), where θ is the angular

$$\frac{d^2\theta}{dt^2} + \omega_0^2\theta = 0 \tag{2}$$

displacement and $\omega = d\theta/dt$ the angular velocity. If this mirror is placed in the gas, the unequal forces on the two sides of it, which are due to the fluctuations in the momentum transferred by the molecules, will cause fluctuations around the harmonic oscillations. After a time, one will have equipartition of energy. For the rotational kinetic energy, one has Eq. (3*a*). For the rotational potential energy, one has Eq. (3*b*).

$${}^1/_2 I\overline{\omega}^2 = {}^1/_2 kT \tag{3a}$$

$${}^1/_2 I\omega_0^2\overline{\theta}^2 = {}^1/_2 kT \tag{3b}$$

By careful measurement, relations such as Eqs. (3*a*) and (3*b*) may be checked. Brownian motion also yields a limit for the accuracy which may be obtained in a given experiment. *See* HARMONIC MOTION.

Langevin equation. If a small particle is suspended in a liquid, its displacement x in the course of time can be observed. The displacement will of course be erratic, but one can still calculate the average displacements, or root-mean-square displacements, at any time t. It is possible to compare these experimental results with theoretical calculations. If a small body moves through a liquid, it will the usual viscous drag, $-\beta Mv$, where M is the mass, v is the velocity, and β is a numerical coefficient depending on the viscosity. For a sphere of radius R, $M\beta$ has the value given by Stokes' law: $M\beta = 6\pi R\eta$, where η is the coefficient of viscosity. In addition, a fluctuating force $MA(t)$, varying rapidly in the course of time, but of average value zero, acts on the particle. One may therefore write an equation of motion as Eq. (4).

$$\frac{dv}{dt} = -\beta v + A(t) \tag{4}$$

See VISCOSITY.

If one now multiplies Eq. (4), the Langevin equation, by x, averages over time, and applies the equipartition law, one obtains for a Stokes case Eq. (5). This result was originally derived by Albert

$$\overline{x^2} = \frac{kTt}{3\pi R\eta} \tag{5}$$

Einstein; it expresses the root-mean-square displacement explicitly in terms of observable parameters. Equation (5) has been verified with an accuracy of about 0.5%. *See* COLLOID. Max Dresden

Bibliography. R. Durrett and H. Kesten (eds.), *Random Walks, Brownian Motion and Interacting*

Particle Systems, 2d ed., 1991; A. Einstein, *Investigations on the Theory of Brownian Movement*, 1926; D. Freedman, *Brownian Motion and Diffusion*, 1971, reprint 1983; T. Hida, *Brownian Motion*, 1980; I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, 2d ed., 1997; F. B. Knight (ed.), *Essentials of Brownian Motion and Diffusion*, 1981, reprint 1989; S. C. Port and C. J. Stone, *Brownian Motion and Classical Potential Theory*, 1978.

Brucellosis

An infectious, zoonotic disease of various animals and humans caused by *Brucella* species. Each species tends to preferentially infect a particular animal, but several types can infect humans. *Brucella melitensis* (preferentially infects goats and sheep), *B. suis* (infects pigs), and *B. abortus* (infects cattle) are the most common causes of human brucellosis. *Brucella melitensis* is the most virulent for humans, followed by *B. suis* and *B. abortus*. *Brucella canis* and *B. ovis*, which infect dogs and sheep respectively, rarely infect humans. Although brucellosis is found all over the world, in many countries the disease has been eradicated. The brucellae are small, gram-negative coccobacilli which are defined as facultative intracellular parasites since they are able to replicate within specialized cells of the host.

Animals. The brucellae often localize in the reproductive tract, mammary gland, and lymph node. They have a particular affinity for the pregnant uterus, leading to abortion and reduced milk production with resultant economic loss to the farmer. Males often undergo infection of the reproductive tract which is characterized by orchitis and epididymitis. Transmission of the disease may occur through the milk of infected animals, but especially through contact with aborted fetuses, afterbirth, and fluids which are very heavily contaminated with brucellae and discharged into the environment. Wildlife, including elk, feral pigs, bison, and reindeer, can become infected and can spread the disease to domestic livestock.

Humans. Brucellosis in humans is characterized by undulant fever, cold sweats, chills, muscular pain, and severe weakness. Some individuals may have recurrent bouts of the disease in which a variety of organs may be affected, sometimes resulting in death. The disease can be contracted by consuming unpasteurized milk or cheese, or via the introduction of organisms through small skin lesions or as an aerosol through the conjunctiva and the respiratory system. Treatment with tetracycline and other antibiotics is most successful if started early after symptoms occur. Development of the disease can be prevented if treatment is initiated immediately after contact with potentially infected material.

Diagnosis. Animals and humans infected with brucellae make antibodies to the bacterial component called the O-antigen. These antibodies are detected in the blood of animals and humans by a variety of serological tests; their presence indicates prior exposure. A test using milk (ring test) may also be used to diagnose the presence of the disease. The disease can also be diagnosed by culturing blood or tissue samples. However, negative culture does not mean absence of infection since numbers in small samples may be too few to be detectable.

Prevention. At present there are no effective vaccines for humans. The disease can be eliminated only by eradicating it in animals. A major source of brucellosis in humans is the consumption of *B. melitensis*–infected milk and cheese from goats. Incidence can be reduced by pasteurizing milk. Animals can be vaccinated to increase their immunity against brucellosis and therefore reduce abortions and disease transmission. Combining vaccination with the elimination of serologically positive animals can lead to disease eradication. In the past, vaccines available for cattle induced positive serological tests so that it was difficult to distinguish vaccinated animals from infected ones, complicating eradication programs. This problem has been eliminated with a new vaccine called strain RB51, which also protects swine. A vaccine, which does not avoid the serological problem, is available for goats and sheep. *See* EPIDEMIOLOGY; MEDICAL BACTERIOLOGY.

Gerhardt G. Schurig

Bibliography. G. G. Alton et al. (eds.), *Techniques for the Brucellosis Laboratory*, 1988; K. Nielsen and R. J. Duncan (eds.), *Animal Brucellosis*, 1990.

Brussels sprouts

A cool-season biennial crucifer (*Brassica oleracea* var. *gemmifera*), which is of northern European origin and belongs to the plant order Capparales. The plant is grown for its small headlike buds formed in the axils of the leaves along the plant stem (see **illus.**). These buds are eaten as a cooked vegetable.

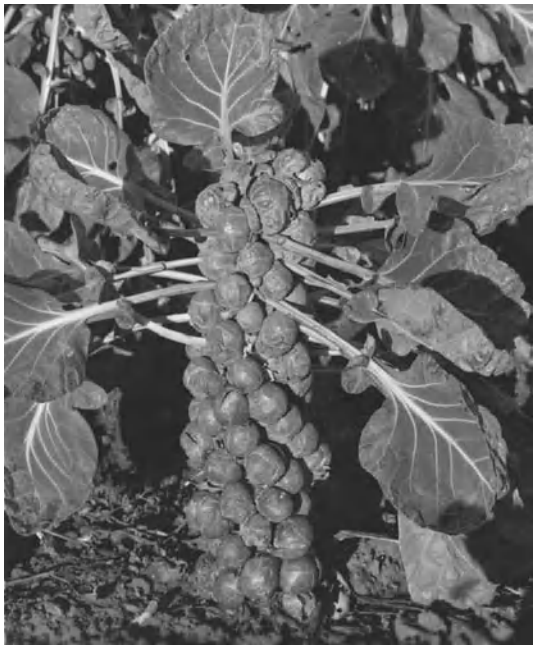

Brussels sprouts (*Brassica oleracea* var. *gemmifera*), Jade Cross. (*Joseph Harris Co., Rochester, New York*)

Cultural practices are similar to those used for cabbage; however, monthly mean temperatures below 70°F (21°C) are necessary for firm sprouts. Brussels sprouts are moderately tolerant of acid soils and have a high requirement for boron. Popular varieties (cultivars) are Half Dwarf and Catskill; however, hybrid varieties are increasingly planted. Harvesting begins when the lower sprouts are firm and 1–2 in. (2.5–5 cm) in diameter, usually 3 months after planting. California and New York are important producing states. *See* CABBAGE; CAPPARALES. H. John Carew

Bryales

An order of the subclass Bryidae. With 11 families and perhaps 44 genera, it is defined in terms of terminal inflorescences, with rare exceptions, and perfect, double peristomes which are papillose on the outer surface. The capsules are generally inclined and more or less pear-shaped. Erect capsules are associated with reduced peristomes.

These mosses often grow in disturbed places. They are perennial and grow in tufts, with stems erect and simple or forked and often densely covered with rhizoids. The leaves are generally bordered by elongate cells and often toothed. The midrib often ends in a hairpoint. The cells, usually smooth, are not differentiated at the basal angles. The sporophytes are nearly always terminal. The setae are generally elongate, and the operculate capsules are usually symmetric but generally inclined to pendulous and commonly pyriform owing to the development of a sizable neck. The peristome is normally double, with a well-developed endostome. The 16 lanceolate teeth of the exostome are evenly tapered and papillose on the outer surface. Keeled and perforate endostome segments alternate with the teeth and arise from a generally high basal membrane. Alternating with the segments are usually one to four cilia. The calyptrae are cucullate and naked. Chromosome numbers are 5, 6, 10, 11, 12, often in polyploid series. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA.

Howard Crum

Bryidae

A subclass of the class Bryopsida. Most genera of true mosses (Bryopsida) belong in the 16 orders of the Bryidae. The most characteristic feature is the peristome consisting of one or two series of teeth, derived from parts of cells rather than whole cells, as in the Tetraphididae, Dawsoniidae, and Polytrichidae. (The Buxbaumiidae have some resemblance in peristome structure to Bryidae.) The stems may be erect and merely forked, or prostrate and freely branched, with sporophytes produced terminally or laterally, respectively. The leaves are inserted in many rows, though sometimes flattened together and appearing two-ranked, but only rarely actually in two rows. The costa may be single or double, sometimes very short, and rarely lacking. The cells are short or elongate, thin- or thick-walled, and often papillose-roughened. The basal and especially the basal angular cells are often differentiated. The setae are generally present and elongate. The capsules dehisce by means of an operculum except in a few genera that show extreme reduction. The number of peristome teeth and segments of the inner peristome are usually 16. The spores are produced by two layers of spore mother cells; they are generally small. The calyptra may be cucullate or mitrate. *See* ARCHIDIIDAE; BRYALES; BRYOPHYTA; BRYOPSIDA; BRYOXIPHIALES; DAWSONIIDAE; DICRANALES; ENCALYPTALES; FISSIDENTALES; FUNARIALES; GRIMMIALES; HOOKERIALES; HYPNALES; ISOBRYALES; MITTENIALES; ORTHOTRICHALES; POLYTRICHIDAE; POTTIALES; SELIGERIALES; SPLACHNALES; TETRAPHIDIDAE.

Howard Crum

Bibliography. G. C. S. Clarke and J. G. Duckett, *Modern Approaches in Bryophyte Systematics*, 1979; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; H. Robinson, A revised classification for the orders and families of mosses, *Phytologia*, 21:289–293, 1971; R. J. Taylor and A. E. Leviton (eds.), *The Mosses of North America*, 1980.

Bryophyta

A division that consists of some 23,000 species of small and relatively simple plants commonly known as mosses, granite mosses, peat mosses, liverworts, and hornworts (see **illus.**). The bryophytes display a distinct alternation of sexual and asexual generations; the sexual gametophyte, with a haploid chromosome number, is the more diversified. The spore-bearing, diploid sporophyte is reduced in size and structure, attached to the gametophyte, and partially or almost completely dependent on it.

Structure. The gametophytes may consist of leafy stems or flat thalli. They have no roots but are anchored to the substrate by hairlike rhizoids. Vascular tissue is at best poorly differentiated, with no lignification of cells. Growth results from the divisions of single cells (rather than meristematic tissues) located at stem tips or in notches at the margins of thalli. The sex organs are multicellular and have a jacket of sterile cells surrounding either the single egg produced in flask-shaped archegonia or the vast number of sperms produced in globose to cylindric, stalked antheridia. The sperms swim by means of two flagella. The sporophyte develops, at least during the early stages, inside a calyptra of gametophytic tissue. The sporophyte commonly consists of a capsule that produces a large number of spores, a stalklike seta, and a swollen foot anchored in the gametophyte. The seta is sometimes lacking and, in a very few thallose liverworts, the foot as well. The spores, nearly always single-celled, are dispersed in the air, except in the case of a small number of aquatics. They germinate directly or produce a juvenile stage called a protonema. *See* REPRODUCTION (PLANT).

Diversity and phylogeny. The division can be divided into five classes: Sphagnopsida (peat mosses), Andreaeopsida (granite mosses), Bryopsida (true mosses), Hepaticopsida (liverworts), and Anthocero-

topsida (hornworts). The mosses have radially organized leafy gametophytes that develop from a protonema and have multicellular rhizoids with slanted crosswalls. The liverworts and hornworts are mostly flat and dorsiventrally organized and have no protonematal stage; the rhizoids are unicellular. Though obviously related, as evidenced by similar sex organs and attachment of a simplified sporophyte to a more complex and independent gametophyte, the classes differ greatly in structural detail. Swimming sperms and relative simplicity suggest an origin among the aquatic green algae, but the attachment of sporophyte to gametophyte and the frequent occurrence of trilete spores, stomata in the capsule wall, rudiments of vascular tissue of some similarity to xylem and phloem (especially in the gametophyte), protection of sex organs, and development of a sporangium that produces countless airborne spores are strong evidence for an origin on land, perhaps from a stock ancestral to the earliest vascular plants, or psilophytes. A significant difference from that group is that bryophytes never have branched sporophytes.

An outline of the Bryophyta is shown below. See separate articles on each group listed.

Class Hepaticopsida (liverworts)
- Subclass Jungermanniidae
  - Order: Takakiales
    - Calobryales
    - Jungermanniales
    - Metzgeriales
- Subclass Marchantiidae
  - Order: Sphaerocarpales
    - Monocleales
    - Marchantiales

Class Anthocerotopsida (hornworts)
Class Sphagnopsida (peat mosses)
Class Andreaeopsida (granite mosses)
Class Bryopsida (true mosses)
- Subclass: Archidiidae
- Subclass: Bryidae
  - Order: Fissidentales
    - Bryoxiphiales
    - Schistostegales
    - Dicranales
    - Pottiales
    - Grimmiales
    - Seligeriales
    - Encalyptales
    - Funariales
    - Splachnales
    - Bryales
    - Mitteniales
    - Orthotrichales
    - Isobryales
    - Hookeriales
    - Hypnales
- Subclass: Buxbaumiidae
  - Tetraphididae
  - Dawsoniidae
  - Polytrichidae

See PLANT EVOLUTION; PLANT KINGDOM.

Howard Crum

Bryophytes. (*a*) Moss plant, *Polytrichum juniperinum* (*General Biological Supply House*). (*b*) Male and female plants of liverwort, *Marchantia*. (*c*) Hornwort, *Anthoceros* (*Carolina Biological Supply Co.*).

Bibliography. N. S. Parihar, *An Introduction to Embryophyta*, vol. 1, 1961; P. Puri, *Bryophytes: A Broad Perspective,* 1973; D. H. S. Richardson, *The Biology of Mosses,* 1981; F. Verdoorn, *Manual of Bryology,* 1932; E. V. Watson, *The Structure and Life of Bryophytes,* 3d ed., 1971.

Bryopsida

The largest class of the division Bryophyta, the true mosses. Members of the class are best characterized by operculate capsules and a peristome that aids in the dispersal of spores, and are generally perennial. The class consists of about 14,000 species distributed in six subclasses based primarily on the structure and developmental history of the sporophyte and

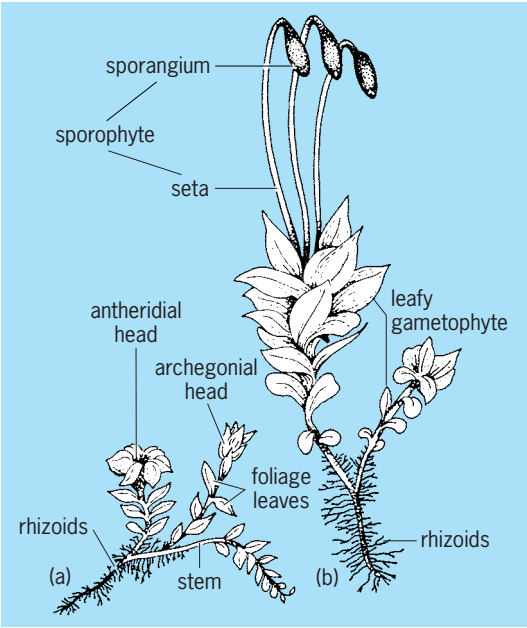

Fig. 1. True moss (*Mnium*). (*a*) Location of sex organs (male antheridia and female archegonia). (*b*) Gametophyte with attached sporophytes. (*After W. W. Robbins et al., Botany, 3d ed., 1964*)

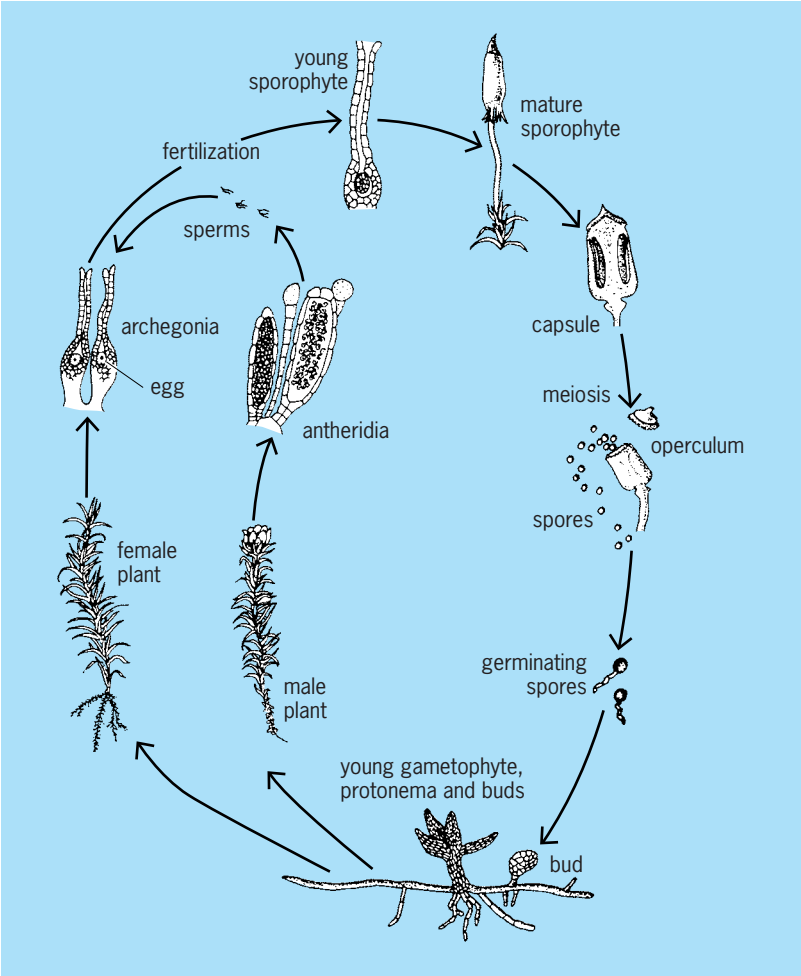

Fig. 2. Life cycle of a moss. (*After E. W. Sinnott and K. S. Wilson, Botany, 6th ed., McGraw-Hill, 1963*)

especially the peristome. The orders and families are likewise based primarily on stable sporophytic details, whereas genera and species are most often differentiated in terms of gametophytic features.

Gametophyte. The filamentous, freely branched protonema of Bryopsida gametophytes produces an abundance of leafy plants which may be erect, simple or sparsely forked, and growing in tufts and producing archegonia at the stem tips (**Fig. 1**); or, alternatively, plants may be prostrate, freely branched, growing in intertangled mats, and producing archegonia laterally. If the apical cell is used up by the formation of archegonia, growth usually continues by the formation of a new subapical branch, called an innovation. But if the terminal cell is not used up and the archegonia are produced laterally, the stems are indeterminate in growth and form numerous branches. Growth results from the activity of a single apical cell with three cutting faces, or in a few genera with two-ranked leaves the apical cell becomes, secondarily, two-faced. In large, erect-growing plants, the stems may have a central strand of vascular tissues similar to xylem and phloem but without lignification. In smaller plants, the vascular tissue is reduced or lacking. The rhizoids are multicellular, with slanted crosswalls. The inflorescences are usually enveloped in differentiated leaves, and the sex organs, of superficial origin, are often mingled with paraphyses, especially in the male inflorescence. The archegonia are flask-shaped; the antheridia are banana-shaped and stalked (**Fig. 2**). *See* PHLOEM; XYLEM.

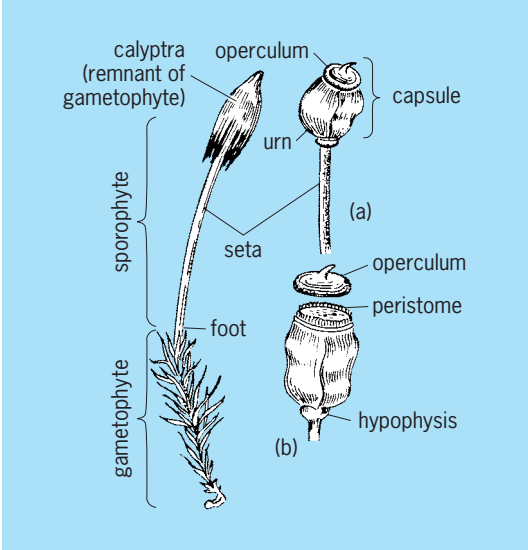

Fig. 3. Complete moss plant showing enlarged views of (*a*) tip of sporophyte with calyptra removed and (*b*) capsule (sporangium), with operculum and peristome and, beneath the capsule, the hypophysis. (*After F. W. Emerson, Basic Botany, 2d ed., Blakiston, 1954*)

Sporophyte. The long-lived sporophytes are abundantly green until maturity and largely self-supporting. They consist of foot and capsule, usually also a seta (**Fig. 3**). The capsules dehisce by means of a lidlike operculum, or rarely irregularly

by rupture. Operculate capsules are provided with a single or double fringe of peristome teeth in multiples of four at the mouth. The peristome is generally derived from the amphithecium (the outer tissue of the embryonic capsule). The capsule wall is usually spongy, especially in the neck portion below the spore sac. Stomata are usually present, especially in the neck or at the junction of capsule and seta; the stomata usually have two guard cells. The spore sac is derived from the endothecium; it surrounds but does not overarch the slender columella. The calyptra may be mitrate (conic and lobed) or cucullate (slit up one side and hoodlike). The chromosome numbers are exceedingly diverse; polyploidy is common, both within species and among related ones. *See* ARCHIDIIDAE; BRYIDAE; BRYOPHYTA; BUXBAUMIIDAE; DAWSONIIDAE; POLYTRICHIDAE; TETRAPHIDIDAE.

Howard Crum

Bibliography. F. Cavers, The inter-relationships of the Bryophyta, *New Phytol.,* 9:81–112, 157–186, 196–234, 269–304, 341–353, 1910, and 10:1–46, 84–86, 1911; E. V. Watson, *The Structure and Life of Bryophytes,* 3d ed., 1971.

Bryopsidales

An order of the green algae (Chlorophyceae), also called Caulerpales, Codiales, or Siphonales, in which the plant body (thallus) is a coenocytic filament (tube or siphon). The filaments may be discrete with free or laterally coherent branches, or organized into a dense plexus exhibiting distinctive morphological features. Septa, which are generally infrequent and incomplete, are formed by centripetal deposition of wall material. A large, continuous central vacuole restricts the cytoplasm to a thin layer just beneath the wall. The cytoplasm contains innumerable nuclei, discoid plastids, and other organelles. Pigments are similar to those in higher plants but include two distinctive xanthophylls, siphonein and siphonaxanthin. Some genera have distinct plastids for the storage of starch. The chief constituent of the wall is frequently xylan or mannan, polysaccharides in which the monomer is xylose and mannose, respectively. Vegetative reproduction is common, usually by rhizomes or fragments. The life history, known for only a few genera, comprises either

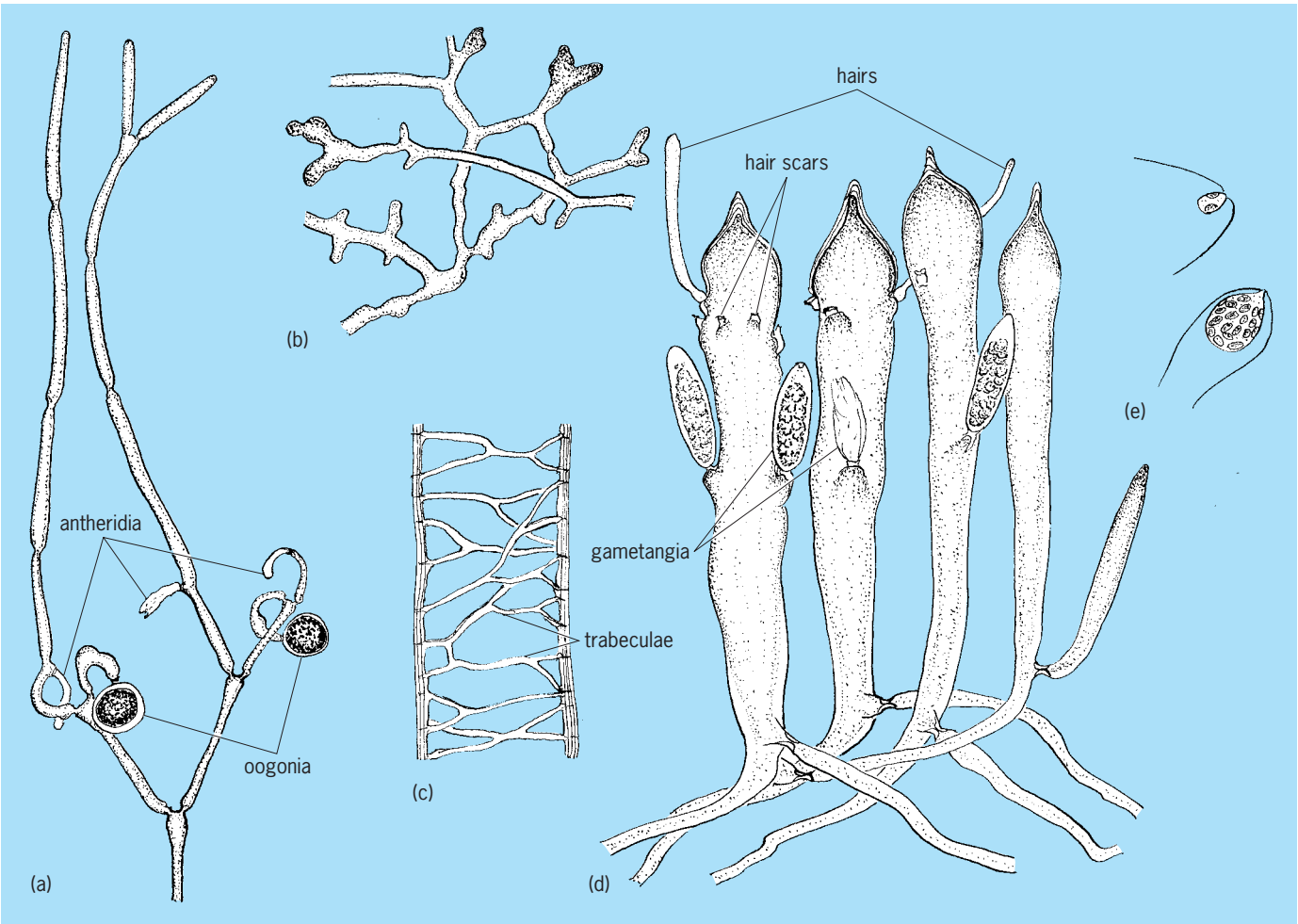

Fig. 1. Anatomical structures of selected members of Bryopsidales. (*a*) *Dichotomosiphon,* filament with oogonia and antheridia. (*b*) *Ostreobium,* filaments. (*c*) *Caulerpa,* section of filament showing trabeculae. (*d*) *Codium*, group of utricles with gametangia and hairs (or hair scars). (*e*) *Codium* gametes, male above and female below.

one somatic phase, with meiosis occurring in the production of gametes, or an alternation of strikingly dissimilar somatic phases, with meiosis occurring in the production of zoospores. Motile gametes are biflagellate, while zoospores have a subapical crown of flagella. Reproductive cells may be formed in unmodified or slightly modified portions of the filament or in special organs.

The order Bryopsidales comprises six families with about 24 genera. The Dichotomosiphonaceae, represented by a single species, *Dichotomosiphon tuberosus* (**Fig. 1***a*), occurs in fresh water, growing at the surface of pools and at depths of at least 50 ft (15 m) in lakes, where it becomes entangled in fish nets. The remaining families are marine and are most abundant and diverse in tropical and subtropical waters. *Bryopsis* (**Fig. 2***a*), an inhabitant of exposed rocky shores as well as wooden pilings in harbors, is the most commonly encountered member of the family Bryopsidaceae. The gametophyte of *Bryopsis* forms a dense tuft of profusely branched filaments, and is mosslike. The Ostreobiaceae includes a single, poorly defined genus (*Ostreobium;* Fig. 1*b*) with six species, all growing in calcareous substrates (old mollusk shells, corals, calcified red algae). *Caulerpa* (Fig. 2*b*), the only genus of the Caulerpaceae, has large filaments reinforced by strands of wall material (trabeculae; Fig. 1*c*) extending into and across the interior. The filaments may be threadlike, but usually they are coarsely cylindrical or flattened. The erect fronds may be very similar in general appearance to other plants, such as club moss and cypress. *Caulerpa* is extremely abundant and diverse in tropical and subtropical waters and (as an exception) in temperate southern Australia. Certain species prefer a rocky substrate, while others spread over large expanses of mud or fine calcareous sand by means of rhizoids. Some species produce a toxin, caulerpicin, which may enter the food chain of the tropical Pacific.

In Udoteaceae, the basic structural unit is a dichotomously or trichotomously branched threadlike filament that is often constricted in a regular pattern. The filaments often form a dense tuft borne on a stem (or stipe), as in *Penicillus* (Neptune's shaving brush; Fig. 2*c*). The most distinctive and ecologically important genus is *Halimeda* (Fig. 2*d*), in which the thallus is an articulate series of flattened or cylindrical segments that are heavily impregnated with aragonite, a form of calcium carbonate. *Halimeda* is nearly ubiquitous in tropical and subtropical seas, and its calcified remains have contributed to the buildup of coral reefs and the infilling of atoll lagoons since the Late Cretaceous. In *Codium* (Fig. 2*e*), the only genus of the Codiaceae, the basic structural unit is similar to that in the Udoteaceae, but the filaments are intertwined to form a spongy thallus that may be crustose, spheroidal, bladelike, or dichotomously branched and up to 33 ft (10 m) long. The chloroplasts are localized in swollen segments called utricles (Fig. 1*d*), which form a palisadelike surface layer. The genus is represented in all parts of the world except the Arctic and Antarctic, but is most diverse in temperate regions. *Codium fragile* ssp. *tomentosoides* is a notorious weed, spreading rapidly in the Atlantic and Pacific by fragmentation and parthenogenesis, and damaging shellfish plantings. *See* ALGAE; CHLOROPHYCEAE.

Paul C. Silva; Richard L. Moe

Bibliography. L. Hillis-Colinvaux, Ecology and taxonomy of *Halimeda*: Primary producer of coral reefs, *Adv. Mar. Biol.,* 17:1–327, 1980; J. Ramus, *Codium*: The invader, *Discovery* (Yale Univ.), 6(2):59–68, 1971.

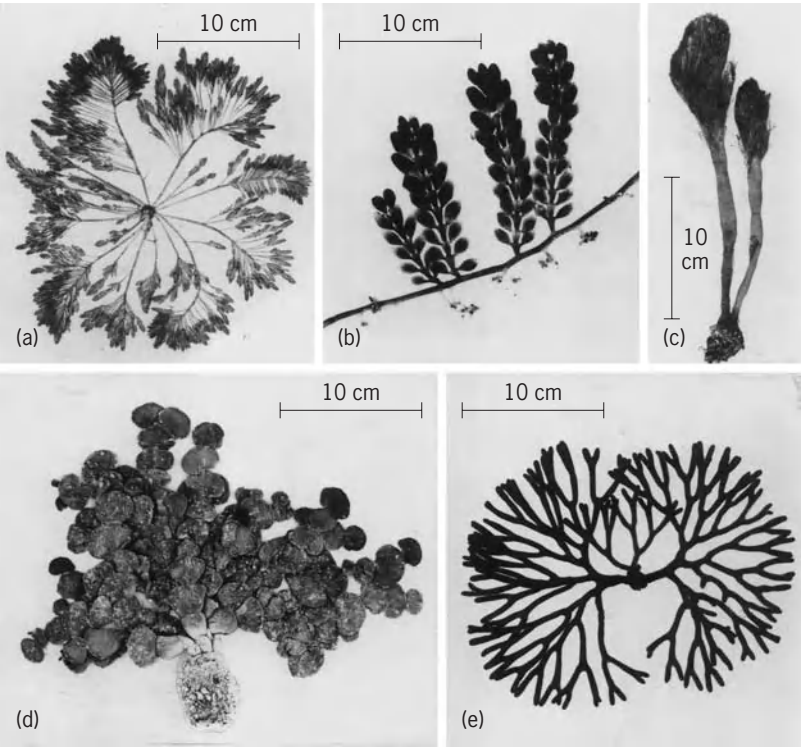

Fig. 2. Representative members of Bryopsidales. (*a*) *Bryopsis*; (*b*) *Caulerpa*; (*c*) *Penicillus*; (*d*) *Halimeda*; (*e*) *Codium*.

Bryoxiphiales

An order of the class Bryopsida in the subclass Bryidae. The order consists of a single genus and species, *Bryoxiphium norvegicum,* the sword moss. This order is characterized by a swordlike appearance owing to leaves overlapping in two rows. The shiny, rigid leaves are keeled and conduplicate-folded. The apex is long-awned at the stem tip and progressively shorter-pointed downward. The midrib bears at back a low ridge of one to four rows of cells. The leaf cells are smooth and subquadrate within, longer and narrower toward the margins. The plants are dioecious with terminal archegonia. The capsules are immersed and have an operculum but no peristome. The calyptra is cucullate, and the haploid

chromosome number 14. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA. Howard Crum

Bibliography. A. Löve and D. Löve, Studies on *Bryoxiphium, Bryologist,* 56:73–94, 183–203, 1953.

Bryozoa

A phylum of sessile aquatic invertebrates (also called Polyzoa) which form colonies of zooids. Each zooid, in its basic form, has a lophophore of ciliated tentacles situated distally on an introvert, a looped gut with the mouth inside the lophophore and the anus outside, a coelomic body cavity, and (commonly) a protective exoskeleton (**Fig. 1**). The colonies are variable in size and habit (**Figs. 2** and **3**). Some are known as lace corals and others as sea mats, but the only general name is bryozoans (sea mosses).

General Characteristics

Byrozoans form colonies by asexual budding from a primary zooid, or ancestrula, formed by metamorphosis of a sexually produced larva or from some kind of resting bud. Structurally the zooids are metazoan, triploblastic, unsegmented, and bilaterally symmetrical, with a regionated fluid-filled body cavity that is considered to be a coelom. The body wall comprises epidermis underlain by a feebly developed peritoneum, between which muscle may be present. The epidermis secretes a chitinlike cuticle or gelatinous layer, and the cuticle may become calcified as a rigid exoskeleton.

Much of the zooid consists of the lophophore, alimentary canal, and associated musculature, together known as the polypide. The lophophore comprises a circle or crescent of slender, ciliated tentacles plus their supporting ridge. When spread for feeding, the tentacles form a funnel with the mouth at its vertex. During withdrawal the tentacles close, and are pulled downward from their base through the simultaneously in-rolling introvert. They then lie within the introvert, called the tentacle sheath. The open distal end of the in-rolled tentacle sheath is termed the orifice, and it may be closed simply by a sphincter muscle or by more elaborate structures.

The alimentary canal is deeply looped, and regionated into pharynx, stomach, and rectum. The limb descending from the mouth consists basically of the pharynx and stomach cardia; the central stomach and its dilatation, the cecum, form the base of the loop; and the stomach pylorus and the rectum constitute the ascending limb. The anus opens outside the lophophore. The nerve ganglion, center of a system that may be complex, lies between the mouth and the anus. There are no special excretory or respiratory organs, and no circulatory system. Colonies, but not all zooids, are hermaphroditic. Zooids are generally not more than 0.04 in. (1 mm) long.

The colony may be minute, of not more than a single feeding zooid and its immediate buds, or substantial, forming masses 3 ft (1 m) in circumference, festoons 1.5 ft (0.5 m) in length, or patches 0.67 ft^2 (0.25 m^2) in area. Commonly the colonies form incrustations not more than a few square centimeters in area, small twiggy bushes up to about 1.2 in. (3 cm) in height, or soft masses up to about 4 in. (10 cm) in the largest dimension. In many colonies much of the bulk consists of the zooid exoskeletons, termed zooecia, which may persist long after the death of the organism and account for the abundance of fossilized bryozoan remains.

Many bryozoans display polymorphism, having certain zooids adapted in particular ways to perform specialized functions, such as protection, cleaning the surface, anchoring the colony, or sheltering the embryo. The evolution of nonfeeding polymorphs is dependent upon some form of intercommunication between zooids.

All bryozoans are epibenthic, and are generally attached to firm substrata, less often anchored in or resting on sand. Most are marine, but one complete class (Phylactolaemata) is confined to fresh water. In the latter habitat, statoblasts and hibernacula, resting bodies resistant to cold and desiccation, are produced. Sexual reproduction leads either to a distinctive, planktotrophic larva, called the

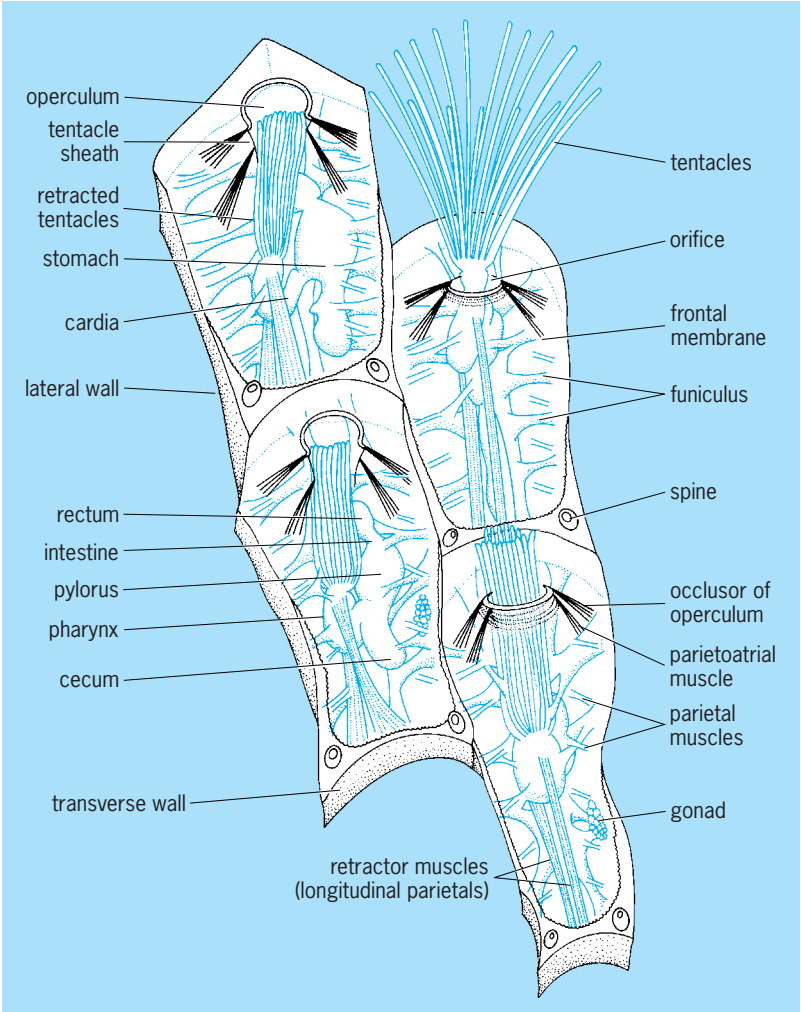

Fig. 1. Morphological features of autozooids, in frontal view.

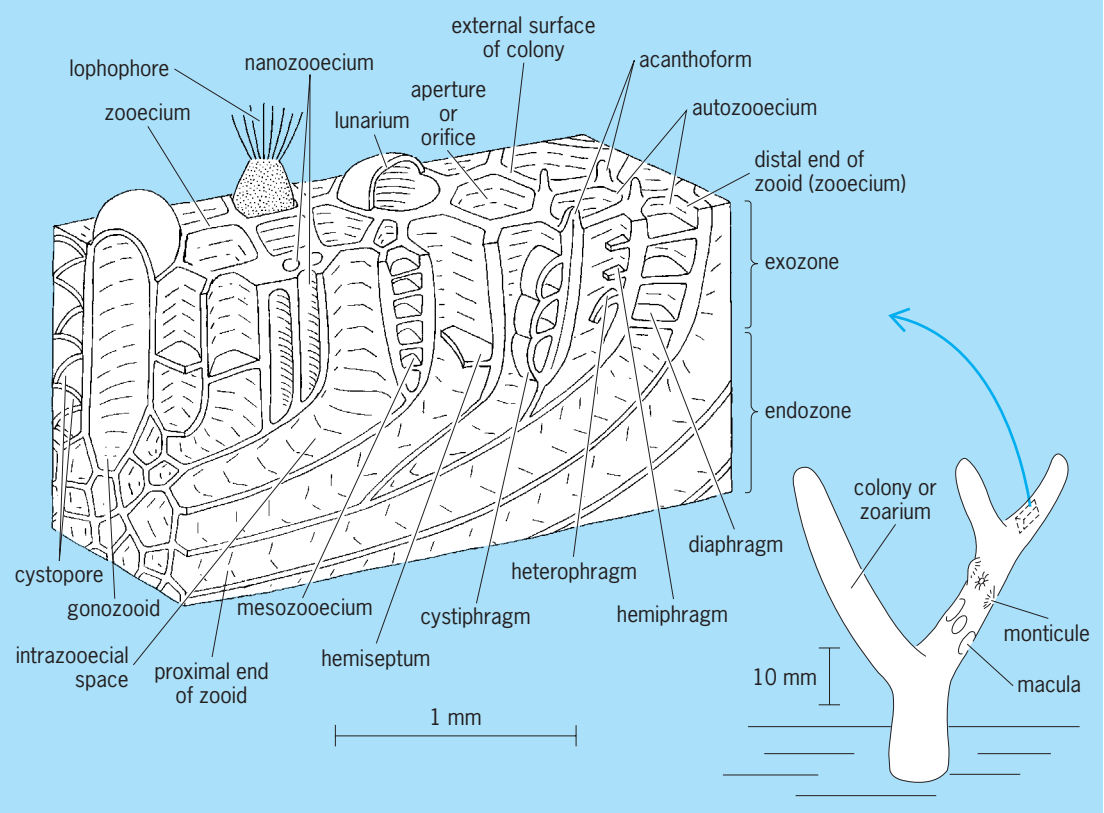

Fig. 2. Stenolaemate morphologic features, some of which are found only in living colonies and others only in fossil colonies. The enlarged oblique view is from the colony at the lower right.

cyphonautes, or more usually to a subspherical, nonfeeding, short-lived type that has had no particular name until recently distinguished as "coronate," because of its extensive ciliary corona.

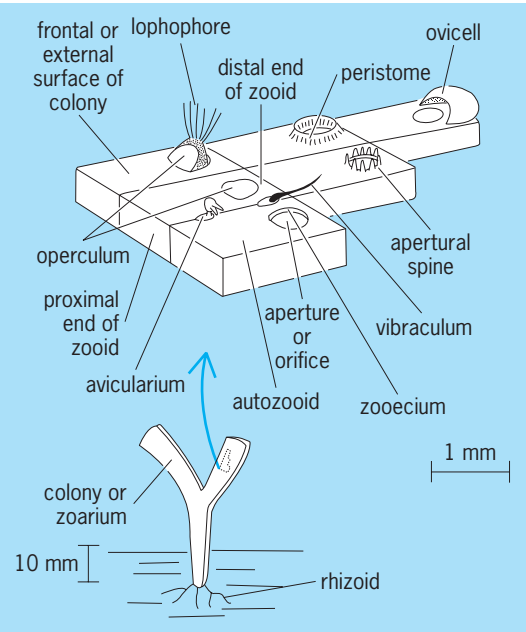

Fig. 3. Morphological features of gymnolaemate colony, and enlarged oblique view of block from it.

Taxonomy and classification. Bryozoa is the name of a phylum for which Ectoprocta is generally regarded as a synonym, these names being used by zoologists according to personal preference. Entoprocta (synonym Calyssozoa) is likewise regarded as an independent phylum. A minority regard Ectoprocta and Entoprocta as subphyla within the Bryozoa, while others maintain Ectoprocta and Entoprocta as phyla but link them under Bryozoa as a name of convenience. *See* ENTOPROCTA.

The phylum contains some 20,000 described species, one-fifth of them living. These are distributed among three classes and a somewhat variable number of orders.

Class Phylactolaemata
  Order Plumatellida
Class Gymnolaemata
  Order: Ctenostomata
    Cheilostomata
Class Stenolaemata
  Order: Cyclostomata
    Cystoporata
    Trepostomata
    Cryptostomata
    Hederellida

See GYMNOLAEMATA; PHYLACTOLAEMATA; STENOLAEMATA.

Until somewhat over 200 years ago bryozoans went unrecognized. When discovered they were

thought, with coelenterates, to be plants—a confusion that led Linnaeus to invent the term zoophyte. In 1830 J. Vaughan Thompson characterized bryozoans, separating them from other zoophytes, but at the same time ambiguously introducing a class Polyzoa. G. C. Ehrenberg's name Bryozoa followed in 1831. *See* CRYPTOSTOMATA; CTENOSTOMATA; CYCLOSTOMATA (BRYOZOA); CYSTOPORATA; HEDERELLIDA; TREPOSTOMATA.

Comparative morphologists link the Bryozoa with the Phoronida and the Brachiopoda, characterized especially by possession of a lophophore, which L. H. Hyman defined as "a tentaculated extension of the mesosome that embraces the mouth but not the anus, and has a coelomic lumen." The tentacle circlet of the Entoprocta is postanal and noncoelomic. The designation of the bryozoan lophophore as mesosomal, and its body cavity as mesocoelic, rests entirely on the interpretation of a supraoral flap of tissue, the epistome, present in phylactolaemates, as the protosome or first region of a tripartite body. In addition to the lophophore and presumed tripartite body, all three phyla have a U-shaped alimentary tract. The lophophore is crescentic in some members of all three groups. Even if the relationship is correct, however, it sheds little light on bryozoan origins. The soft-bodied Phylactolaemata have left no fossil record, but comparative anatomy suggests that they, the Stenolaemata, and the Gymnolaemata must already have been distinct by the Ordovician. The view that Bryozoa arose from entoprocts, perhaps by internal budding from the larva, seems tenuous, and implies at most no more than a similarity between larvae such as that which links Mollusca to Annelida. Certainly the immediate ancestor must have been vermiform and sedentary, for cylindrical zooids, terminal lophophores, and deeply recurved digestive tracts with the anus rather close to the mouth represent the primitive state in all classes.

Functional morphology of zooid. Zooids in all three bryozoan classes display the same ground plan: eversible lophophore, U-shaped gut, fluid-filled coelom, and deformable body wall. The lophophore is everted by increase in hydrostatic pressure caused by inflexion of the body wall; withdrawal is achieved by the direct pull of retractor muscles anchored to the wall. The flexible phylactolaemate wall incorporates circular and longitudinal muscle, and its contraction everts the lophophore. Recent work indicates that stenolaemates, as evidenced by the extant Cyclostomata, have a remarkably modified system. Throughout this class the zooid is cylindrical (often inaccurately described as tubular; only the exoskeleton is tubular). Except in the immediate vicinity of the orifice, the cuticle is calcified and totally rigid. The peritoneum, with a hypertrophied basal membrane and associated bands of circular muscle, has detached from the epidermis and lies freely as a bag, the membranous sac, anchored at some points to the body wall. The endosaccal cavity is coelomic, but the exosaccal cavity, external to the mesoderm, is not. The sac contracts to evert the lophophore; exosaccal fluid must be displaced proximally as the polypide is eased toward the orifice.

The simplest gymnolaemates (Ctenostomata) have cylindrical zooids and deformable walls, although the musculature is extrinsic. The circular (now renamed transverse parietal) muscles lie inside the peritoneum, and the longitudinal muscles are reduced and specialized. Mechanically the system works as in Phylactolaemata. An early trend of zooidal evolution in the Gymnolaemata was toward a flat, membranous, coffinlike shape, adherent to the substratum. It seems probable that the cheilostomes arose from ctenostomes like that by calcification of the lateral walls and the differentiation of a cuticular lid, the operculum, to close the orifice. Cheilostomes of this plan, classified as the suborder Anasca, preserve a wholly or partially flexible front wall. Transverse parietal muscles, relocated to span the coelom from the rigid lateral walls to the frontal membrane, bring about its inflexion.

Natural selection in cheilostomes has favored modifications which better protect the polypide while preserving the function of the frontal membrane. Thus the frontal membrane can be underlain by a shelf, at a distance sufficient to permit inflexion; or it can be overarched by flattened spines or by a meshwork of partially fused spines. In the subclass Ascophora the membrane may be overgrown or replaced by a calcified frontal wall, but the flexible membrane either remains or is replaced more deeply by a functional equivalent. A saclike cavity or ascus, opening just behind the orifice, is created either by overgrowth of the wall or by involution from the orifice underneath the wall. Transverse parietal muscles insert on the lower face of the ascus, which is pulled down to evert the lophophore, while water to compensate for the volume change enters the ascus through its opening.

Feeding. The expanded lophophore of marine bryozoans forms an almost radially symmetrical funnel of 8–30 tentacles, although fixed or transient bilateral symmetry may occur. In gymnolaemates the base of this funnel stands free above the surface of the zooid, but in stenolaemates it lies concealed within the orificial region of the exoskeleton. The top diameter is in the range of 0.01 to 0.05 in. (0.25 to 1.25 mm). In primitive phylactolaemates 20–30 tentacles form a funnel, but in most members of this class the large lophophore comprises 100 or more tentacles disposed in a horseshoe, the lobes of which project from the adanal side.

The base of the lophophore is the most complex part of the polypide. Essentially it is a hollow epidermal annulus, with thickened basement membrane. The lumen (mesocoel) opens into the main body cavity (metacoel) by an adanal pore. The cerebral ganglion partly occludes this pore, and a nerve tract parallels the mesocoel. Radial and circular muscles open and close the mouth. The tentacles rise from the annulus as slender hollow cylinders. Their central coelom is surrounded by a thick, flanged collagenous tube, the hypertrophied basement membrane

of the epidermal cells. The lumen, which opens terminally through a minute pore, is lined by peritoneum and contains frontal (that is, facing into the funnel) and abfrontal longitudinal muscles. The lateral epidermal cells bear long cilia (about 25 micrometers) and the frontal cells shorter cilia (about 15 μm). Laterofrontally is a line of static sensory cilia. The tentacles contain motor and sensory nerves.

The lateral cilia generate a water current that enters the top of the funnel and flows downward and outward between the tentacles. Particles such as bacteria and phytoplankton are separated from the water and directed at the mouth by the frontal cilia without the use of mucus. The mechanism of separation is uncertain, but a likely explanation is that particles passing out between the tentacles, perhaps sensed by the rigid laterofrontal cilia, are flicked back into the funnel by localized reversal of ciliary beat. Small particles accumulate just above the mouth; heavier particles may be projected down the funnel and directly into the partly open mouth.

Unwanted particles are rejected by closing the mouth, flicking the tentacles, closure or concerted movements of the funnel, and ejection from the pharynx (see below). A few bryozoans with large lophophores have rejection tracts leading centrifugally between the ventromedial (abanal) tentacles. Bryozoans are vigorous feeders, and it has been calculated that the lophophore of even a small zooid filters about 0.3 ml/h (and a colony may contain thousands of zooids).

Bryozoan colonies may have exhalant chimneys or set points of water outflow. The reason is clear: if the colony is large, solid, flat, and uniformly covered by tentacular funnels pumping toward its surface, the filtered water requires a means to get away. Small colonies may have a single, central chimney; large colonies may have a mamillate surface, the summit of each mamilla marking a chimney; and wholly flat colonies produce chimneys by an apparently coordinated bending away of lophophores from predetermined points. Filtered water flows toward the chimneys, through which it is expelled. Lophophores surrounding a chimney display strong bilateral symmetry, with the bordering tentacles standing tall and upright instead of curving into the chimney. A consequence of the posture is that the beat of the current-generating lateral cilia becomes centripetal to the chimney rather than being unhelpfully directed toward its base.

The pharynx is a remarkable organ. It is short, thick-walled, and situated immediately below the mouth. Its exterior is circular in section, its interior deeply furrowed. The frontal ciliation from the lophophore continues through the mouth and may cover the distal part of the pharynx: the midventral groove is ciliated, even when the rest of the pharynx is not, and its cilia beat upward, whether there is a ventromedial rejection tract or not. Between the grooves, the inwardly bulging walls are made up of vacuolated cells which incorporate intracellular myofibrils. The simultaneous contraction of all the myofibrils, acting against the fluid-filled vacuole, causes a brief convulsive dilation of the pharynx, engulfing the particles accumulated outside the mouth.

A sphincter and valve separate the pharynx from the cardia (descending arm of the stomach). In some ctenostomes, and a few other bryozoans, part of the cardia is differentiated as a gizzard, armed internally with teeth or keratinized plates, presumably for crushing the frustules of diatoms. The unmodified wall of the cardia contains secretory cells. The floor of the central stomach is ciliated, evidently to carry particles into the cecum, which is the main locus of extra- and intracellular digestion and absorption. The epithelium of the pylorus is ciliated and serves to compact the undigested remains into a revolving cord. A sphincter separates the pylorus from the rectum, which has an absorptive function and modifies the cord of reject material into fecal pellets. The pellet is moved through and expelled from the rectum by peristalsis. At summer temperatures (72°F or 22°C) the gut transit time is less than 1 h but at 43°F (6°C) it is 2 h.

Metabolic integration and colony evolution. The simplest bryozoan colonies presumably consisted of loosely interconnected, monomorphic feeding zooids. In the most ancient fossil stenolaemates, confluence of coeloms had been achieved either by means of internal pores or by the so-called double wall, in which a superficial (or hypostegal) coelom unites the colony outside the calcification. The hypostegal coelom must have been reasonably effective as a means of distributing metabolites, since the colonies of trepostomes were large by bryozoan standards, and the fenestrates evolved colony forms of considerable complexity. However, rapid growth and the proliferation of nonfeeding zooid polymorphs requires a transport system more efficient than aqueous diffusion. The funiculus provides such a system.

In living single-walled stenolaemates (cyclostomes) zooids communicate via small, open interzooidal pores. The funiculus is a small muscular strand linking the proximal end of the cecum with the base of the membranous sac at its attachment to the zooid wall. The testis is positioned where the funiculus joins the cecum. The phylactolaemate funiculus similarly is muscular, supports the testis, and links the cecum to the proximal end of the ventral body wall. More importantly, it contains a core of blastogenic tissue from which the statoblasts arise. These are the future resting buds and become richly supplied with yolk. The obvious inference is that the products of digestion are transported, perhaps to the testis but certainly to the developing statoblasts, along the funiculus. Communication between zooids in phylactolaemates is by open pores; indeed in many of them the dividing walls have almost completely broken down, so that the polypides are suspended in a communal coelom.

The adaptive radiation of zooid form has been very slight in both Cyclostomata and Phylactolaemata. Cyclostome zooids have remained as slender cylinders, while the colonies have evolved mainly in terms of the way in which the zooids are disposed.

Thus regular, alternating placement of zooid orifices is replaced by a pattern of separated short rows or fascicles, presumably to improve feeding efficiency; and erect or partially erect colonies have evolved from creeping ones. The reproductive zooids, at first clearly in series with the rest, become larger, more central, and often truly colonial by incorporating neighboring structures. Phylactolaemates have evolved larger, deeply crescentic lophophores, but the zooids have not diversified. Evolution of the colony has been toward unity by confluence of coeloms.

Primitive gymnolaemates have an organization very like that of the simpler phylactolaemates: elongate zooids connected by simple open pores, with the funiculus confined within each zooid. The radiation of stolonate ctenostomes has been facilitated by the development of a colonial, essentially nonmuscular funiculus which passes through the internal pores between zooids and along the stolons. A flux of lipid from feeding zooids into the stolonal funiculus occurs and spreads toward the growth and budding zones. The funiculus has become a colony-wide metabolic expressway.

When cheilostomes evolved from ctenostomes, they inherited the funicular system. They added mural calcification and developed complex communicating pores in the lateral walls, while retaining simple pores in the end walls. The main funicular trunks are longitudinal, passing from the proximal pores to the distal pores via the cecum, with branches to the lateral pores. The success of the Cheilostomata, in terms of species numbers, adaptive radiation, and polymorphism, is evidently the result of a high-potential module, plus the distributive efficiency of the funiculus in nourishing large numbers of specialized, nonfeeding zooids.

Nervous system. The cerebral ganglion is dorsally (adanally) situated in the lophophore base and has a diameter of 30–60 μm. It is extended circumorally as a ring or deep crescent. Apart from the fibrillar component, the anascan ganglion contains distal, central, and proximal groupings of cell bodies. The principal nerves are the following. (1) Sensory and motor lophophoral nerves. (2) Paired direct tentacle sheath nerves which unite with the first branch of the trifid nerve to form (3) the compound tentacle sheath nerves: these extend up the tentacle sheath to the superficial part of the zooid, including the sensitive region near the orifice. (4) The trifid motor nerve; of its three branches, the first, 4a, is visceral, meeting its opposite number at a small suprapharyngeal visceral ganglion; the second, 4b, innervates the paired retractor muscles of the polypide; and the third, 4c, joins the compound tentacle sheath nerve and innervates the tentacle sheath muscles, the orificial sphincter, occlusors of the operculum, and the transverse parietal muscles that deflect the frontal membrane. It will be noted that 4b and 4c innervate the antagonistic effectors in polypide withdrawal and eversion. (5) Paired parietal nerves parallel the direct tentacle sheath nerves, but pass deep to join a peripheral encircling ring which enters in turn each of the interzooidal pores. Here the parietal nerves of neighboring zooids make contact. (6) Paired visceral nerves originate in the suprapharyngeal ganglion.

Electrophysiological research has confirmed the existence of a functional colony-wide nervous system in anascans. Stimulation near the orifice results in the immediate withdrawal of all nearby expanded lophophores (while stimulation of a lophophore results only in its withdrawal).

Polymorphism. Phylactolaemates have achieved integration at the expense of zooid individuality, so that polymorphs have not evolved (unless the statoblasts are so regarded). Stenolaemates have metabolic integration dependent on diffusion gradients, and display weakly developed polymorphism. In living cyclostomes, four morphs are commonly encountered: (1) normal-feeding or autozooids; (2) empty or kenozooids of various kinds, which support colonies, make up attachment rhizoids and stolons, in-fill spaces, and project as spines; (3) female or gynozooids, also called brooding or gonozooids, that do not feed but produce ova and shelter the developing embryos; and (4) dwarf or nanozooids, containing a reduced unitentaculate polypide which sweeps and cleans the surface of the colony. All zooids which are not autozooids may be termed heterozooids.

Among gymnolaemates, which have metabolic integration via the funiculus, the ctenostomes have limited polymorphism. Apart from the autozooids, only kenozooids are commonly present, either as stolon segments or as spines attached to autozooids. The cylindrical type of zooid, whether calcified or cuticularized, has not provided the potential for extensive polymorphism. In cheilostomes, on the contrary, one sees a wide range of heterozooids, high numbers of heteromorphs per colony, and the grouping of polymorphs into functional units or cormidia (colonies within colonies). Of importance in contributing to such impressive adaptive radiation seem to be the funicular system, the retention of individualistic zooids, and the calcified, boxlike zooid with its partly membranous front and hinged, lidlike operculum closing the orifice.

Cheilostome zooid types merge with one another, but the following may usefully be recognized: (1) sterile or fertile autozooids, which may also brood embryos but do feed, at least for most of the time; (2) androzooids or nonfeeding males with reduced polypides; (3) gynozooids (gonozooids), nonfeeding females which produce ova and sometimes brood the embryo; (4) avicularia, in which the operculum is enlarged; (5) vibracula, in which the operculum is lengthened to whiplike proportions; (6) kenozooids for in-filling, or associated with attachment rootlets and stolons; (7) kenozooids in the form of spines; and (8) ooecia or brood chambers.

Avicularia appear to be zooids modified for grasping. They vary in form from something close to an autozooid, differing only in having an enlarged operculum, or mandible, to diminutive and highly modified structures. The large ones often obviously replace autozooids in the colony, and are termed vicarious; the smaller ones are attached to auto-zooids,

and are termed adventitious. Often a zooid may bear several adventitious avicularia of different kinds and in a variety of positions. Abductor muscles open the mandible by pulling down the small frontal membrane just behind the hingeline, whereas larger adductors snap it shut. In its most evolved form, the bird's-head avicularium, the zooid is pedunculate and mobile. The avicularian polypide is much reduced, and is believed to fulfill a sensory role.

Some avicularia have a long slender mandible, and may be precursors of vibracula, in which the operculum has become a whiplike seta. The vibracular seta, however, is mounted on asymmetrical condyles and can be swiveled by gyrator muscles, as well as being moved through a regular arc. In some species the vibracula remove sediment from the colony surface, and in nonattached sand-living forms they also act as struts or legs.

Spines are small kenozooids of varied form. They are normally adventitious and presumably have a protective role. The arguments for regarding cheilostome spines as zooids are basically three: they are separated from their bearing autozooid by a communication pore of standard morphology; they terminate with a flat membrane, seen as homologous to a frontal membrane; and they may replace an undoubted polymorph, for example, an autozooid or avicularium. The fundamental arrangement of spines in anascans is as a marginal ring around the frontal membrane. One worker sees this as the vestige of an ancient type of budding, each spine representing what was once a free-rising series of zooids. Even when this annular pattern is modified, as in certain erect, branching anascans and in ascophorans, the basic arrangement frequently persists in the founding zooid or ancestrula. In later zooids the spines are confined to the immediate region of the orifice.

Marginal spines may be modified, as when one or a series overarches the frontal membrane. One major fossil group (with a few extant species) has a porous, protective frontal shield formed by the fusion of such spines.

Most cheilostomes brood their embryos in special external chambers of standard plan: a hoodlike upfolding from the distal/proximal boundary of two zooids, which may become partly or wholly immersed in or between these zooids. It is called an ooecium and is double-walled, with outer ectooecium and inner entooecium separated by a coelomic lumen. In some species at least, the lumen is confluent with the coelom of the bearing zooid, and the ooecium appears to be a spinous polymorph. However, its development is inconsistent, for example, commencing with paired outgrowths from the proximal zooid or with a single outgrowth from the distal zooid. It is easy to interpret the former as representing the distal pair of marginal spines, especially since these are usually otherwise lacking in ooeciferous zooids. Development from the distal zooid is more difficult to explain, although it should be noted that proximally situated spinelike zooids are not unknown in cheilostomes.

The bryozoan colony is made up from replicated zooids. The simpler polymorphs are vicarious and dispersed through the colony in an ordered or apparently random manner. With the development of adventitious polymorphs, however, a second-order unit, the cormidium, becomes recognizable. Cormidia are groupings of dissimilar polymorphs. Siphonophora (Cnidaria) provide the standard example of colonies with cormidia and it seems generally unappreciated that cheilostomes have cormidia that are at least as complex and probably better integrated. Thus an autozooid may bear spines, an ooecium, adventitious avicularia, and perhaps a vibraculum. Little is known about behavioral coordination in cormidia; much more is known about the metabolic relationships between a female zooid and its ooecium.

Reproduction and growth. Propagation by statoblasts is important in Phylactolaemata. They develop in large numbers on the funiculi, and are liberated when the colony breaks up. Statoblasts have a protective coat, often a float, and sometimes hooked spines. They are resistant to cold and desiccation and give rise to a new colony under favorable conditions. Nonphylactolaemates in fresh and brackish water may produce resting zooids or hibernacula, and many marine species overwinter as nutrient-filled stolons. Asexual reproduction by fragmentation occurs in a few marine bryozoans, and is said to be important to the free-living sand forms. Lobulation of colonies is characteristic of the more evolved phylactolaemates.

Gonads in bryozoans are associated with the peritoneum: testes on the funiculus, ovaries more often on the zooid wall. Cyclostome spermatozoa develop in tetrads, but in phylactolaemates and gymnolaemates large numbers develop around a common mass, the cytophore. In marine bryozoans (at least) they escape from the body via the mesocoel and the terminal pore in two (the dorsomedial pair) or all of the tentacles. Special male zooids with nonfeeding polypides may be localized in chimneys, thereby taking advantage of the exhalant current to disperse the spermatozoa.

The fertilized egg in phylactolaemates develops in an invagination of the body wall. When released, it is already essentially a motile zooid ready to settle and found a colony. The cyclostome gynozooid is inflated and may occupy much of the colony. Its embryo, having reached the blastula stage, lobulates a series of secondary blastulae, which may repeat the process, until the brood chamber contains 100 or more embryos. Each develops into a simple, ovoid, coronate larva which swims for a short time and then metamorphoses into a proancestrula or incomplete primary zooid. This lengthens and becomes functional.

Sexual reproduction in gymnolaemates follows one of three general patterns. (1) Small eggs are discharged into the sea via the intertentacular organ, an inflated tube situated between the dorsomedial tentacles. Fertilization takes place there. The zygote develops into a planktonic, feeding, triangular, flattened, bivalved larva or cyphonautes. This probably lives for several weeks before it settles. (2) A

large, fully yolked egg is produced in the coelom, and extruded through the supraneural coelomopore (replacing the intertentacular organ) into an external brood chamber. If this is an ovicell, the embryo receives no nourishment, for its size is unchanged during development. A nonfeeding coronate larva is liberated. (3) Smallish yolky eggs are transferred to the lumen of an ovicell via the supraneural pore. The ovicell comprises the ooecium, previously described, and an internal membranous extension of the mother zooid, the inner vesicle. The fertilized egg lies between the ooecium and the inner vesicle. The vesicular epithelium, well supplied with funicular strands, hypertrophies where it is in contact with the embryo and becomes a pseudoplacenta. That food is supplied to the embryo is evidenced by its growth during development. A nonfeeding coronate larva is liberated.

The larva responds to light and other stimuli. It achieves limited dispersal and selects a settlement site. Prior to metamorphosis the larva explores a surface with special long cilia; at the moment of attachment a hollow organ, the adhesive sac, is everted, spreads as a disk, and flattens onto the surface. Metamorphosis is total and its details are complex, but the larval organization disappears and a reversed polarity emerges. (Phylactolaemate development does not include this reversal.) The ancestrular polypide differentiates from the upper wall, and colonial budding commences.

The method of budding differs between the three classes. In phylactolaemates new polypide rudiments appear, abanally to the parent, in an ordered manner which determines the branching pattern of the colony. Differentiation of cystids (the zooid walls), insofar as they are present, follows the polypides.

In stenolaemates polypide rudiments also appear first, but are adanal to the parent zooid. Since the cystids are essentially tubular, they grow around the polypide rudiments following division of the longitudinal walls. However, it is important whether these calcified walls meet or remain slightly distant from the terminal epithelium. In the latter case, coelomic confluence remains around their distal end (the double-walled structure), whereas in the former, cystids are separate once formed and communicate only through interzooidal pores.

Gymnolaemate budding is different again, and dependent from the outset on the connecting pores and funiculus. The cystid is produced first, adanally to the mother zooid, and the polypide develops within it.

The pattern of colony formation is termed astogeny. Early on, it is often characterized by sequential change in the size and morphology of the zooids; later astogeny is repetitive, with the production of virtually identical zooids, groupings of zooids, cormidia, or groups of cormidia. The continuous colony in cheilostomes arises from the apposition of repeatedly dividing lines of zooids. The walls between contiguous lines are therefore morphologically double, and the communication pores in them have differentiated from potential zooid buds, which is one reason why they are more complex than the pores in end walls. Frontal budding, of a layer of zooids above the existing layer or layers, may also occur.

Initial formation does not complete the zooid's development. Wall thickening may continue, phases of reproductive activity come and go, while a cycle of polypide aging and renewal continues until senescence. This cycle, in which the whole polypide regresses, and is histolyzed until only a compacted, fibrous "brown body" remains, is peculiar to bryozoans. Old zooids may consist only of an accumulation of brown bodies.

Ecology. Fresh-water bryozoans are present on submerged tree roots and aquatic plants in most lakes, ponds, and rivers, especially in clear water of alkaline pH. Their presence is best confirmed by dip-netting for statoblasts at the water's edge. Most other bryozoans are marine, although some gymnolaemates inhabit brackish water. They are common in the sea, ranging from the middle shore to a depth of over 26,000 ft (8000 m), and are maximally abundant in waters of the continental shelf. Most attach to firm substrata, so that their distribution is primarily determined by the availability of support. Mud is unfavorable and so is sand unless well provided with stone, dead shells, hydroids, or large foraminiferans, as in regions of strong current, where colony densities may reach 470/ft^2 (5000/m^2) of hard substrate.

Intertidal bryozoans shun the light, and are found under boulders, below overhangs, and on algae, particularly under conditions of high flow; they avoid turbid water and accumulating sediment. The seaweed dwellers are mainly fleshy ctenostomes and specially flexible cheilostomes, or are very small. Experiments with larvae have demonstrated high selectivity with regard to substrate, for example, algal species or shell surface. Settlement is influenced by texture and contour, with larvae often preferring concavities.

Colony form in bryozoans is to some extent related to habitat. Encrusting and bushy flexible species are adapted to wave exposure; brittle twiglike and foliaceous species are found deeper; some erect branching species tolerate sediment deposition. One group of tiny discoid species lives on sand in warm seas, and in one genus the colonies are so small that they live actually among the sand grains; a few species live anchored by long kenozooidal stems in mud. A number of stolonate ctenostomes bore into the substance of mollusk shells; other species are associated only with hermit crabs, and a few are commensal with shrimps or polychaete worms.

Primarily, however, bryozoans are inhabitants of shaded rock surfaces, where they compete with each other, sponges, and compound ascidians for space. They have no set size or shape, and reproductive capacity is simply a function of size. Some display adaptations, such as raised margins, long spines, or frontal budding, to resist being overgrown. When colonies meet, there is often a mutual cessation of growth, but with different species the one with larger zooids

may overgrow its rival or impair its feeding efficiency. Closely related species sometimes display slightly different habitat preferences, thereby avoiding direct competition. Pendent and outgrowing colonies escape the surface competition and become better placed for feeding; they include elegant lace corals and robust fans rising 6–12 in. (15–30 cm) in height.

Bryozoans have few serious predators. Nudibranch mollusks and pycnogonids (sea spiders) specialize in feeding on zooids but are rarely destructive of entire colonies. Loxosomatids (Entoprocta) and a hydroid (*Zanclea*) are common commensals.

Life spans vary. Small algal dwellers complete their life cycle in a few months. Many species survive a year but have two overlapping generations; others are perennial, with one known to survive for 12 years.

Bryozoans may be a nuisance in colonizing ship hulls and the insides of water pipes, and one species has caused severe dermatitis in fishers. Recently some delicate kinds have been used in costume jewelry, and green-dyed clumps of dried *Bugula* are often sold as "everlasting plants." John S. Ryland

Cystoporata

Within the Bryozoa is an extinct order, the Cystoporata, in the class Stenolaemata. Cystoporates had encrusting, mound-shaped, or basally attached, erect, calcified colonies made of sheets or branches composed of tubular feeding zooids and extrazooidal vesicular skeletal deposits.

Zooecia, the skeletal tubes delimiting the zooids, in many cystoporates are characterized by lunaria, which are thickened strips that extend up one side of each zooecium. Lunaria commonly have a different degree of transverse curvature than does the rest of the zooecium, and extend somewhat above the general surface of the colonies and arch slightly over the zooecial aperture. Polymorphic zooecia are uncommon and consist of small "space-filling" structures and, less commonly, enlarged, apparently female zooids.

Ceramoporine cystoporates are characterized by laminated wall structure, usually short zooecia, sparse diaphragms, and commonly by gaps of pores in the skeletal walls that may have served for nutrients to move between zooids. These organisms ranged from the Middle Ordovician to the Lower Devonian.

Fistuliporine cystoporates have laminated, granular, or granular-prismatic wall structure, short to long zooecia, and sparse to abundant diaphragms, and lack gaps or pores in the skeletal walls; and in many the extrazooidal vesicular skeletal deposits are replaced outwardly by continuous, dense skeletal material (stereom). Rodlike structures (acanthostyles) may occur in skeletal walls, oriented perpendicularly to the skeletal surface and projecting above the colony surface. These organisms ranged from the Lower Ordovician to the upper Permian.

Cystoporates were exclusively marine and lived in a wide variety of relatively low-energy environments on continental shelves and seas. Both subgroups had diverse growth habits, slightly greater in the fistuliporines. Erect colonies grew as two-layered sheets with zooids growing back to back (bifoliate), as narrow bifoliate branches that bifurcated to form bushes or that anastomosed to form perforated sheets, as radially symmetrical branches that could be several centimeters in diameter, and in other growth habits.

Most cystoporates, except those with narrow branches, have maculae—small, uniformly spaced areas composed of a group of zooecia or a similar-sized spot of extrazooidal skeleton. Maculae may be elevated, depressed, or surrounded by slightly enlarged zooecia that typically tilt away and have lunaria on the side closest to the maculae. Spacing, size, and disposition of zooecia around maculae suggest that they mark points where water, filtered by the surrounding zooids, flowed away from the colony surface in a focused, chimneylike flow, as seen in many living bryozoans with large surface areas.

Fossils

Fossil Bryozoa have a long geological history, from early in the Ordovician Period [500 million years ago (Ma)] to the Recent (Fig. 1). Individual fossils range in size from a few millimeters to several meters in maximum dimension. Various encrusting or erect growth forms are common, though some were free-living (Fig. 2). Representatives of the marine orders that secreted calcareous skeletons (Cryptostomata, Cyclostomata, Cystoporata, Trepostomata, and Cheilostomata) commonly are abundant in sedimentary rocks formed where benthic organisms flourished. Skeletons generally are calcite, though some are aragonite or mixed calcite and aragonite. Ctenostomata have nonmineralized skeletons, so they have been preserved only as excavations or borings in marine shells or on the undersides of other organisms that overgrew them (a preservation style termed bioimmuration). The fresh-water Phylactolaemata have gelatinous skeletons, but their tough statoblasts (dormant reproductive bodies) have been reported from sediments as old as the Jurassic (at least 150 Ma).

Fossil Bryozoa are most abundant in calcium carbonate–bearing rocks such as various limestones, calcareous shale, and calcareous siltstone, and they are especially common in some alternating thin-bedded limestones and shales. During the Ordovician, Carboniferous, and Permian periods, bryozoans were important parts of many fossil reefs, reef flanks, and other carbonate buildups in shallow (less than 100 m depth) tropical waters. Bryozoans commonly dominate and may reach very high diversities in post-Paleozoic cool-temperate carbonate deposits, indicating a shift in primary environment after the Paleozoic.

In the study of bryozoans, examination with a microscope is necessary. Although colonies of many species are large, the individual skeletons of each zooid (unit of the colony) range from less than 0.1 to about 1 mm in diameter. The smaller diameters are typical for cross sections of elongate tubes that

characterize zooids in stenolaemate bryozoans, and the larger diameters are typical for the more equidimensional zooids of cheilostomes. Identification is based on numerous external and, for most stenolaemates, internal features that require study with a microscope. Adequate classification of most stenolaemate bryozoans generally requires variously oriented thin sections through the colony interior (Fig. 2).

Classification. Features of the colonial skeletons (zoaria) as well as the morphology of the individual zooidal skeletons (zooecia) are used to classify bryozoans. Many fossil bryozoans had only one type of zooid (autozooids), which apparently could feed and carry out all other necessary biological functions of the colony. Others were polymorphic, with various types of specialized zooids supplementing the autozooids. Number, types, and morphology of polymorphs is important in classification. Other characters important in classification of fossil bryozoans are wall structure, reproductive chambers, general growth habit or specific shape of colonies, and for some, surface topography of the colony.

Geological history. The earliest bryozoans (representatives of the Cryptostomata, Trepostomata, Cyclostomata, and Cystoporata) are recorded from Lower Ordovician rocks (a few fossils from the Cambrian have been assigned erroneously to the Bryozoa). In Mid-Ordovician time, diverse groups of these various bryozoan orders, including the suborder Ptilodictyoidea of the order Cryptostomata, had a broad geographic distribution (**Fig. 4**). The phylum continued to evolve rapidly, and toward the end of the Ordovician Period the Trepostomata and the ptilodictyoid Cryptostomata were dominant. The Cystoporata, Cyclostomata, and Ctenostomata were less important faunal elements of the time period.

In the Silurian, the Trepostomata and Cryptostomata were still dominant. New groups diversified in other orders, including the Fenestellina (a suborder of Cryptostomata also referred to as fenestrates), which occurred sparsely in the Late Ordovician. The cystoporates also became a more distinctive part of the bryozoan faunas. Relatively little is known about the Silurian bryozoans of many regions because the principal sedimentary rocks are dolomites and evaporites and the bryozoans are either poorly preserved or they did not inhabit these environments.

Diversity of genera and species further increased in the Devonian Period. Fenestellinid and rhabdomesine cryptostomes were very abundant, and the number of fistuliporid cystoporates increased. The trepostomes continued to be well represented by new species and genera.

The Carboniferous and Permian bryozoans were dominated by various kinds of fenestellinids, rhabdomesines, and cystoporates. The trepostomes were less numerous, and cyclostomes remained uncommon. Toward the end of the Permian Period, almost all trepostomes, cryptostomes, and cystoporates became extinct. This drastic reduction in bryozoan groups extended throughout much of the late Permian and parallels similar patterns in other Paleozoic fossil groups.

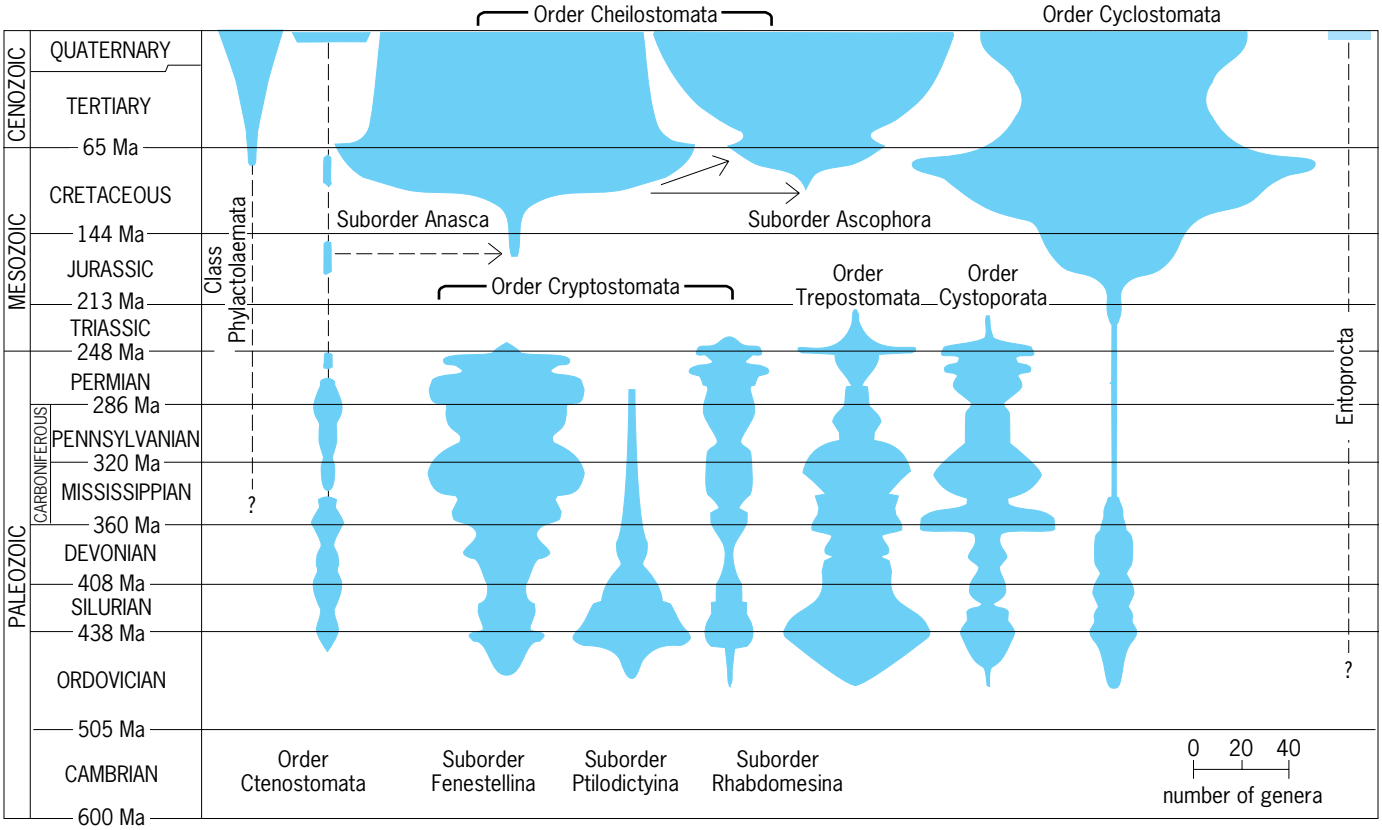

Fig. 4. Diagram of the geological distribution and abundance of Bryozoa.

Bryozoans are sparse in the Triassic Period. A few cystoporates, trepostomes, rhabdomesines, and cyclostomes occur in Triassic rocks. Representatives of the first three groups are Triassic "holdovers" from the Paleozoic fauna, and the oldest Mesozoic occurrence of cyclostomes is in Upper Triassic rocks.

The Jurassic Period witnessed a great expansion of the Cyclostomata. This order dominated bryozoan faunas until the Late Cretaceous. The Cheilostomata, initially represented by anascans, first appeared late in the Jurassic and rapidly became the dominant bryozoan group, undergoing explosive diversification during mid-Cretaceous. Many of the major cheilostome groups and major evolutionary innovations had developed before the end of the Late Cretaceous.

The cyclostomes continued to evolve a few new families during the Cenozoic Era, but their diversity was reduced by one-third at the end of the Cretaceous and the group has never rebounded. With the exception of the very beginning of the Cenozoic Era (the Danian Stage of the Paleocene Epoch), the Bryozoa of the Cenozoic Era, like the present-day fauna, were principally cheilostomes. Although cheilostome diversity was reduced by over one-fourth at the end of the Cretaceous, the group resumed rapid diversification by the Eocene Epoch (35–55 Ma). Some of the cheilostome groups, such as the cribrimorphs, which flourished in the Late Cretaceous, declined in numbers in the Cenozoic. The anascans, a dominant group in the Cretaceous, continued to maintain their importance in Cenozoic bryozoan faunas and to the present day. The ascophorans diversified especially rapidly during the Eocene, and they maintain high diversity to the present.

Bryozoans have experienced several vigorous radiations and two devastating extinction events. The first adaptive radiation occurred during the Ordovician and involved the Trepostomata and ptilodictyoid cryptostomes, and the second occurred during the Devonian and involved the fenestellinid cryptostomes, which were the dominant bryozoans during the late Paleozoic. All these stenolaemate groups, plus the rhabdomesines and cystoporates, were severely reduced by the end-Permian extinction and were completely extinct before the end of the Triassic Period. A third radiation, which involved only the cyclostome stenolaemates, occurred during the Jurassic, and the cheilostome gymnolaemates had the first of their two radiations during mid-Cretaceous. The end-Cretaceous extinction event strongly reduced the diversity of both cheilostomes and cyclostomes. The cyclostomes never recovered their Late Cretaceous diversity, but by the Eocene cheilostomes were experiencing a renewed interval of vigorous diversification and reached diversity levels higher than those of Late Cretaceous.

Evolutionary relationships. The fossil records of all orders of the class stenolaemate bryozoans (cyclostomes, cystoporates, trepostomes, and cryptostomes) begin in the Arenigian Stage (478–488 Ma) of the Lower Ordovician, and the soft-bodied gymnolaemate order Ctenostomata is represented back to the Late Ordovician Ashgillian Stage (438–448 Ma) by borings in shells. Thus, there are two problems with using the fossil record to interpret early evolutionary relationships among bryozoans. The first is that earliest representatives of the highly skeletonized stenolaemate orders appear fully differentiated and essentially simultaneously in the fossil record, so relationships among these highly mineralized groups cannot be inferred from the sequence of appearance. The other problem is that the soft-bodied ctenostomes must have had a substantially longer existence than indicated by the fossil record, because the oldest ctenostome fossils are specialized, apparently nonprimitive ctenostomes that left borings in shells. (Most ctenostomes live on surfaces of shells and other substrata and do not excavate chambers.)

Ctenostomes are widely considered to be a paraphyletic group (a group of taxa that contains the ancestar but not all of its descendants) that gave rise to the Stenolaemata during the early Paleozoic and to the cheilostome Gymnolaemata during the Jurassic. One hypothesis about the origin of the Stenolaemata concerns the method of protrusion of the tentacles of zooids to the feeding position. Ctenostomes have a completely flexible body wall. Their only defense against predators is either to remain small and cryptic or to bear allelochemicals, yet their completely flexible body wall can be easily contracted to translate fluid that pushes out the feeding apparatus (analogous with squeezing a toothpaste tube to extrude the toothpaste). Stenolaemates have rigidly calcified lateral body walls, and within the calcified tubular zooids there is a separate, mesodermal "membranous sac" that can be squeezed by transverse parietal muscles to protrude the feeding apparatus. It has been hypothesized that the stenolaemates originated when the ectodermal epithelium of a ctenostome-like ancestor secreted calcite as well as an outer organic membrane, along with separation of the muscle-bearing mesoderm from the ectoderm to form the membranous sac of the proto-stenolaemate. This condition would allow predator-resisting calcification to develop and still permit protrusion and retraction of the feeding apparatus. This evolutionary innovation apparently occurred early in the Ordovician Period, and stenolaemates rapidly diversified into several distinct types that are recognized as orders. Detailed phylogenetic analysis indicates that the Cyclostomata, as traditionally conceived, are paraphyletic and gave rise to the other, essentially Paleozoic, stenolaemate orders.

Origin of the cheilostomes is much clearer. A family of uniserial, encrusting ctenostomes with pyriform zooids is known from bioimmured (entombed while alive) representatives as old as Middle Jurassic; some Jurassic specimens have D-shaped orifices, which is normally a cheilostome rather than a ctenostome attribute. This family (arachnidiid ctenostomes) apparently gave rise to the cheilostomes. The earliest known cheilostomes were Late Jurassic, encrusting colonies of simple, uniserial

to loosely pleuriserial, pyriform to elliptical feeding zooids with mineralized lateral walls.

The Phylactolaemata are probably phylogenetically most closely related to the Ctenostomata. At present all phylactolaemates live in fresh water, and their earliest fossils consist of small asexual reproductive bodies found in Jurassic fresh-water deposits. The gelatinous nonmineralized colonies in a fresh-water environment have reduced probability of preservation as fossils. The only other fresh-water bryozoans are some ctenostomes, but it is unclear whether phylactolaemates evolved from fresh-water ctenostomes or whether they originated in the oceans and later migrated into fresh water and then disappeared from the oceans.

Evolutionary paleoecology. Maximum diversity and abundance of bryozoans has historically been on middle to outer shelves and similar conditions on epeiric sea floors, inferred to be several tens of meters deep. Bryozoan-rich deposits in the Paleozoic were built up by trepostomes, cystoporates, and cryptostomes and developed in shallow equatorial to cold-temperate (midlatitude) environments. However, post-Paleozoic bryozoan-rich deposits were generated early on by cyclostomes and then by cheilostomes, and almost all formed in middle latitudes, seldom in tropical environments. This shift in pattern is probably biologically driven, rather than driven by physical or chemical changes of the world oceans. The "Mesozoic marine revolution" saw the proliferation of several groups of vigorous predators in the tropics, as well as the rise in importance in the tropics of competing organisms such as scleractinian corals and certain algae that grew much more rapidly than bryozoans. While some cheilostomes and a few cyclostomes have reached enormous sizes during the post-Paleozoic, similar large sizes were reached by Paleozoic trepostomes and cystoporates, and on average Paleozoic bryozoans, especially trepostomes and cystoporates, were larger than their post-Paleozoic relatives.

During the Paleozoic, erect growth habits became progressively more abundant; whereas during the post-Paleozoic, encrusting growth habits became progressively more abundant. These contrasting patterns have been interpreted to result in part from an increase in predation during the post-Paleozoic, because slow-growing erect bryozoans cannot as easily recover from partial predation as can encrusting forms. Among erect bryozoans during the Paleozoic and also during the post-Paleozoic, species characterized by narrow, unilaminate branches increased in importance from a trivial percentage to the most numerous. The unilaminate, narrow-branched growth habit allows food–bearing water to be processed much more efficiently than in any of the other colony morphologies, and this efficiency has been inferred to provide an adaptive advantage.

Bryozoans are competitively inferior in overgrowth interactions (so they tend to be overgrawn by other taxa) to most other encrusting organisms such as sponges, ascidians, cnidarians, and algae, and this may have been part of the reason for their relative unimportance in the post-Paleozoic tropics. Among post-Paleozoic bryozoans, cheilostomes consistently have been superior in overgrowth interactions with cyclostomes, due to multiple differences such as, on average, higher growth rates, higher colony margins, and stronger feeding currents in cheilostomes. Simultaneously, cheilostomes have evolved a larger average colony size and have undergone two vigorous stages of diversification (Late Cretaceous and, after recovery from the end-Cretaceous mass extinction, Paleogene), while cyclostomes have declined slowly from their peak diversity in Late Cretaceous and now have on average smaller colony sizes than in the Cretaceous.

The relatively small size of most bryozoan colonies and their slow growth rates, plus the general competitive inferiority of encrusting bryozoans, has relegated them to a minor role in construction of reefs and other organic buildups through geological time. There have been exceptions. During the Mid-Ordovician, encrusting bryozoans were the primary framework builders on reefs and, in many instances, erect bryozoans contributed the bulk of the skeletal debris to the reef flanks. From Silurian through Carboniferous, reef frameworks were locally predominated by bryozoans, and fenestellinid bryozoans grew abundantly on certain organic buildups during the Carboniferous and Permian. In the post-Paleozoic, multilaminar encrusting to massive cyclostome bryozoans built a few small Mesozoic reefs, and cheilostomes similarly built small reefs a few meters wide and high during the Cenozoic. Beyond these exceptions in which reefs were largely or partly built by bryozoans, bryozoans occur widely in reefs as cryptic organisms inhabiting the undersides of corals and small-scale caves.

Throughout their geological history, bryozoans have been much more common on ephemeral substrata that are available for months or years than on persistent substrata that are available for several years and longer (such as reefs). They are widespread, and diverse on brachiopod shells in Paleozoic rocks and on bivalve shells in Jurassic to Neogene deposits. Various life history strategies have evolved to utilize the relatively ephemeral, often mobile shell substrata. Short-lived colonies may live preferentially on undersurfaces of shells and develop only a few zooids before brooding embryos. In contrast, colonies of some species can effectively occupy the entire exposed shell surface and continue to live and expand time after time as the shell or small rock is overturned, so that the original substrate is buried deeply in multiple layers of bryozoan zooids of a colony that can live and function in any orientation with respect to the sea floor.

Repetitive evolution of similar colony morphologies has been the hallmark of bryozoan history. Cheilostomes have evolved unique combinations of characteristics in their successful invasion of soft substrata, where they are either free-lying or live attached by specialized zooids that resemble roots. But otherwise, similar colony morphologies have evolved convergently in many different orders. The

limitations on bryozoan morphologies (including constraints on branch sizes and shapes) have been due to the basic organization of bryozoans. Zooids have always been small (almost always much less than 1 mm^2 in surface area), feeding bryozoans must be at the colony surface and have access to nutrient-bearing water, and the water that is drawn toward the colony surface must find a nearby outlet away from the colony after it is filtered. Paleozoic stenolaemates collectively, post-Paleozoic cyclostome stenolaemates, and cheilostome gymnolaemates have all developed similar-appearing encrusting, massive, and erect colonies. Consequently, any given bryozoan fossil commonly must be examined closely in order to place it in the correct taxonomic order.

Frank K. McKinney

Bibliography. D. P. Gordon, A. M. Smith, and J. A. Grant-Mackie (eds.), *Bryozoans in Space and Time*, 1996; P. J. Hayward, J. S. Ryland, and P. D. Taylor (eds.), *Biology and Palaeobiology of Bryozoans*, 1994; A. Herrera and J. B. C. Jackson (eds.), *Proceedings of the 11th International Bryozoology Association Conference*, 1999; F. K. McKinney and J. B. C. Jackson, *Bryozoan Evolution*, 1989; R. A. Robison (ed.), *Treatise on Invertebrate Paleontology, Part G*, rev.1983.

Buckeye

A genus, *Aesculus*, of deciduous trees or shrubs belonging to the plant order Sapindales. Buckeyes grow in North America, southeast Europe, and eastern Asia to India. The distinctive features are opposite, palmately compound leaves and a large fruit having a firm outer coat and containing usually one large seed with a conspicuous hilum. The Ohio buckeye (*A. glabra*), a small tree which may grow to a height of 30 ft (9 m), is found mainly in the Ohio valley and in the southern Appalachians. It can be recognized by the glabrous winter buds, prickly fruits 1–2 in. (2.5–5 cm) long, and compound leaves having five leaflets. *See* SAPINDALES.

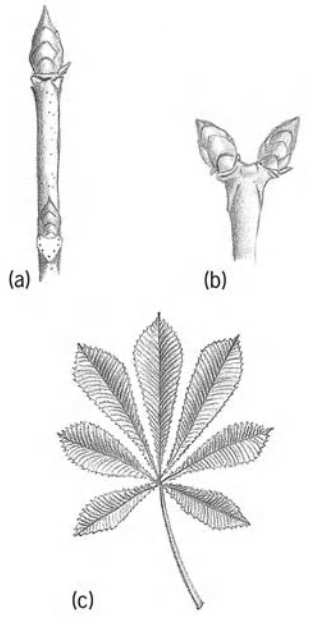

Horse chestnut (*Aesculus hippocastanum*): (*a*) twig, (*b*) buds, (*c*) leaf.

Another important species, the yellow buckeye (*A. octandra*), which may reach to 90 ft (27 m), is native in the Central states, has five leaflets and smooth buds, but differs in its smooth, larger fruit, which may be 2–$2^1/_2$ in. (5–6 cm) long. The horse chestnut (*A. hippocastanum*), which usually has seven leaflets and resinous buds (see **illus.**), is a native of the Balkan Peninsula. It is planted throughout the United States and is a beautiful ornamental tree bearing cone-shaped flower clusters in early summer. The seeds of all species contain a bitter and narcotic principle. The wood of the native tree species is used for furniture, boxes, crates, baskets, and artificial legs. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Buckwheat

A herbaceous, erect annual, the dry seed or grain of which is used as a source of food and feed. It is not a true cereal and is one of the very few plants, other than those of the Gramineae family, used for their starchy seed, which is processed as a meal or flour.

Buckwheat belongs to the Polygonaceae family, which also includes the common weeds dock, sorrel, knotweed, bindweed, smartweed, and climbing false buckwheat. Species of buckwheat that have been commercially grown are *Fagopyrum sagittatum* (*F. esculentum*), *F. emarginatum*, and *F. tataricum*. *See* POLYGONALES.

The plant grows to a height of 2–5 ft (0.6–1.5 m), with many broad heart-shaped leaves. It produces a single main stem which usually bears several branches, and is grooved, succulent, and smooth except for nodes (see **illus.**). Buckwheat is an indeterminate species in response to photoperiod, and produces flowers and fruits (so-called seeds) until the beginning of frost. *See* PHOTOPERIODISM.

Species differentiation. *Fagopyrum sagittatum* and *F. emarginatum* produce flowers that are densely clustered in racemes at ends of branches or on short pedicels that arise from the axils of the leaves. Individual flowers have no petals, but the calyx is composed of five petallike sepals which may be light green, white, pink, or red. Populations include plants typically of two floral types: the pin type which has flowers with long pistils and short stamens and the thrum type which has flowers with short pistils and long stamens. The pistil consists of a one-celled ovary, three-parted style with knoblike stigmas, and eight stamens. Glands (usually eight) which secrete nectar are located at the base of the ovary. Generally, self-fertilization is prevented by a heteromorphic incompatibility system, and seeds are produced only when cross-pollination occurs between the pin and thrum stylar types. The so-called seed of

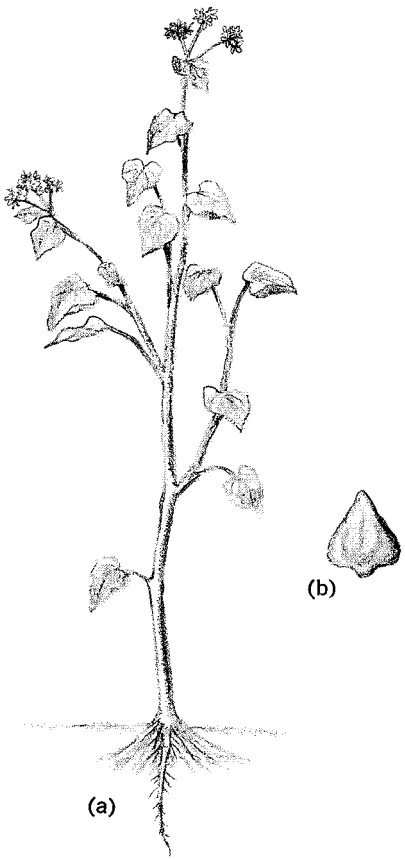

Buckwheat. (*a*) Mature plant. (*b*) Seed.

buckwheat (actually a fruit which is an achene) usually has three angles and varies in shape, size, and color. *See* FLOWER; FRUIT; SECRETORY STRUCTURES (PLANT).

Fagopyrum tataricum (commonly known as tartary buckwheat, rye buckwheat, duck wheat, or hulless) differs from the two previously described species. The leaves are narrower and arrow-shaped, and the flowers are smaller with inconspicuous greenish-white sepals. Plants are only of one flower type and are self-fertile.

Cultivation and use. Buckwheat is of minor importance as a grain crop in the United States, and is principally grown in areas of the Northeast where the weather is likely to be moist and cool. It is usually grown on land too rough or poor for other grain crops or, since it matures within 10–12 weeks, as an emergency crop where previous crops have failed. The crop is also used to smother weeds, as green manure, and as food and cover for various game birds and wildlife. Buckwheat is often used as bee pasture, and is the source of a dark, strong-flavored honey.

H. G. Marshall

Processing. The production of buckwheat flour requires cleaning, grinding, and fractionation in a manner similar to that used for wheat flour. The distinctive polygonal shape of the grain kernel enables the use of sieving screens with triangular openings to facilitate separation of weed seeds and other foreign material. Sieving by size and the use of air aspiration enables separation of good-quality grain. Poor-quality, low-density seed and trash are removed by air aspiration. Successive cleaning of the grain is essential for premium quality. *See* WHEAT.

Hot-water steeping or steam pretreatment of the grain may be used to improve the nutritional properties, yield, and market value of buckwheat products. Dry heating or steaming of the grain prior to drying with temperatures less than 158°F (70°C) improves the efficiency of hull removal and reduces cooking time of selected products. Hulling is accomplished by using roller mills adjusted to crack the husk with minimum damage to the kernel. This operation must be precisely controlled to avoid excessive kernel fracture. Hulls, constituting about one-fifth of the grain by weight, are separated by sieving. Buckwheat hulls provide little food value and are utilized as animal litter.

The products of dehulled buckwheat include whole groats, split kernels, middlings or farina, and flour. Whole groats, splits, and farina are obtained by selective sieving and may be utilized as breakfast cereals and porridges or as thickening agents. Dry roasting of whole groats may be employed to increase color and flavor while decreasing their preparation time. When milled as flour, buckwheat will yield 60–75% extraction. The flour is typically more coarse and more highly colored than wheat flour. Buckwheat middlings, which include the layer immediately below the hull and the germ, provide valuable animal feed stock. In the United States, buckwheat flour is used primarily in pancake mix formulations, blended with wheat, corn, rice, or oat flour. Japanese soba, a dried noodle product, is prepared from blends of buckwheat and wheat flours. *See* CEREAL; FOOD MANUFACTURING; GRAIN CROPS.

Mark A. Uebersax

Bibliography. H. G. Marshall and Y. Pomeranz, Buckwheat: Description, breeding, production, and utilization, in *Advances in Cereal Science and Technology*, vol. 5, American Association of Cereal Chemists, 1982; H. K. Wilson, *Grain Crops*, 2d ed., 1955.

Bud

An embryonic shoot containing the growing stem tip surrounded by young leaves or flowers or both, and the whole frequently enclosed by special protective leaves, the bud scales.

Position. The bud at the apex of the stem is called a terminal bud (**Fig. 1***a*). Any bud that develops on the side of a stem is a lateral bud. The lateral bud borne in the axil (angle between base of leaf and stem) of a leaf is the axillary bud (Fig. 1*a* and *d*). It develops concurrently with the leaf which subtends it, but usually such buds do not unfold and grow rapidly until the next season. Because of the inhibitory influence of the apical or other buds, many axillary buds never develop actively or may not do so for many years. These are known as latent or dormant buds. Above or beside the axillary buds, some plants regularly produce

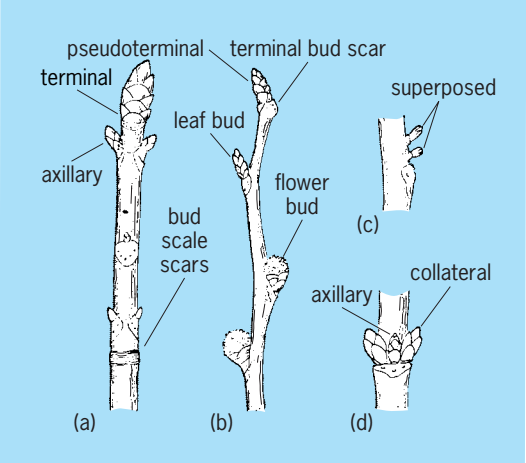

Fig. 1. Bud positions. (*a*) Terminal and axillary (buckeye). (*b*) Pseudoterminal (elm). (*c*) Superposed (butternut). (*d*) Collateral (red maple).

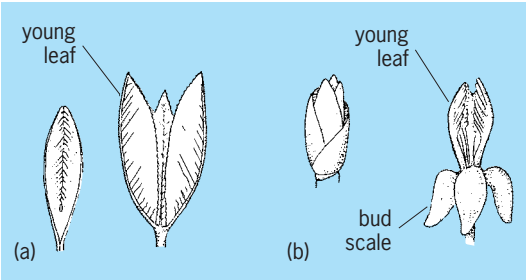

Fig. 2. Bud coverings. (*a*) Closed and open naked buds (hobblebush). (*b*) Closed and open scaly bud (hickory).

additional buds called accessory, or supernumerary, buds. Accessory buds which occur above the axillary bud are called superposed buds (Fig. 1*c*), and those beside it collateral buds (Fig. 1*d*). Under certain conditions, such as removal of terminal and auxillary buds, other buds may arise at almost any point on the stem, or even on roots or leaves. Such buds are known as adventitious buds. *See* PLANT GROWTH.

Composition. Buds that give rise to flowers only are termed flower buds (Fig. 1*b*), or in some cases, fruit buds. If a bud grows into a leafy shoot, it is called a leaf bud (Fig. 1*b*), or more accurately, a branch bud. A bud which contains both young leaves and flowers is called a mixed bud.

Covering. Buds of herbaceous plants and of some woody plants are covered by rudimentary foliage leaves only. Such buds are called naked buds (**Fig. 2***a*). In most woody plants, however, the buds are covered with modified protective leaves in the form of scales. These buds are called scaly buds or winter buds (Fig. 2*b*). In the different species of plants, the bud scales differ markedly. They may be covered with hairs or with water-repellent secretions of resin, gum, or wax. Ordinarily when a bud opens, the scales fall off, leaving characteristic markings on the stem (bud scale scars).

Vernation. The characteristic arrangement of young leaves within the bud is known as vernation or prefoliation (**Fig. 3**). Within the buds, the leaves may be folded on the midrib, plaited, rolled, coiled, or twisted in exact order and precision, each kind of plant having its distinctive inner bud pattern. Many ferns have a coiled leaf arrangement which is known as circinate vernation. *See* LEAF; PLANT ORGANS; PLANT TAXONOMY.

Neele Ammons

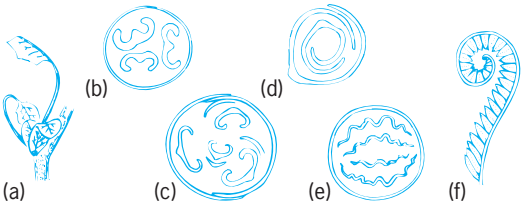

Fig. 3. Some types of vernation (prefoliation). (*a*) Conduplicate (tulip tree). (*b*) Revolute (dock). (*c*) Involute (poplar). (*d*) Convolute (cherry). (*e*) Plicate (sycamore). (*f*) Circinate (fern).

Buffalo

The name for members of the family Bovidae in the mammalian order Artiodactyla. The buffalo is an Old World species and resembles the oxen in general appearance. The North American bison is often called a buffalo, but is not related to the true buffalo. The Asiatic buffalo (*Bubalus bubalis*), known as the Indian or water buffalo and also as the carabao, is found as a domestic animal in the Balkans, Asia Minor, and Egypt. These buffalo exist in the wild state in southern Asia and Borneo, where they are considered to be ferocious and dangerous. Water buffalo are stocky, heavy-built animals that are not quite 6 ft (2 m) at shoulder height. They have very short hair and the shortest, most splayed horns of any variety (see **illus.**). Like all buffalo, they have a liking for marshes, where they wallow and become caked with mud that affords protection against insects. Their wide, flattened hooves allow them to stand firmly on the soft, marshy soil. They are hardy animals, being able to resist the infections and insect-borne diseases of the country, and they can withstand cold so well that they do not need shelters. Their diet consists entirely of reeds, rushes, and other aquatic plants. Domestic herds of 200–300 animals may provide 300 pints (1400 liters) of milk a year. This milk is rich in cream and is used to make butter and yogurt. *See* MILK.

Two other Asiatic species related to, but smaller than, the water buffalo are the tamarau (*Anoa mindorensis*), which is indigenous to the Philippines, and the still smaller (40 in. or 1 m high) anoa (*A. depressicornis*), or wild dwarf buffalo, found in the Celebes.

The African buffalo, classed in the genus *Syncerus*, was very numerous until the turn of the century, when the infectious disease rinderpest caused many deaths. They are still abundant though widely hunted by the natives. There are several varieties of African buffalo, and it is thought that all may be subspecies of *S. caffer*, the Cape buffalo, which attains a height of about 5 ft (1.5 m) at the shoulder and weighs more than 1500 lb (675 kg). They live in the open country of central, eastern, and southern Africa. Mating

Herd of Indian or water buffalo (*Bubalus bubalis*).

occurs early in the year, and a single calf is born after a gestation period of 11 months. Except for its size, this animal is difficult to distinguish from the rare dwarf or forest buffalo (*S. caffer nanus*), which is less than 4 ft (1.2 m) high at the shoulder and weighs a little more than 400 lb (180 kg). It lives in marshy, forested areas of western Africa, where it is known as the bush cow. This species has not been domesticated and occurs in herds of up to 1000. *See* ARTIODACTYLA; BISON; MAMMALIA; MUSK-OX; YAK. Charles B. Curtin

Buffers (chemistry)

Solutions selected or prepared to minimize changes in hydrogen ion concentration which would otherwise tend to occur as a result of a chemical reaction. In general, chemical buffers are systems which, once constituted, tend to resist further change due to external influences. Thus it is possible, for example, to make buffers resistant to changes in temperature, pressure, volume, redox potential, or acidity. The commonest buffer in chemical solution systems is the acid-base buffer.

Chemical reactions known or suspected to be dependent on the acidity of the solution, as well as on other variables, are frequently studied by measurements in comixture with an appropriate buffer. For example, it may be desirable to investigate how the rate of a chemical reaction depends upon the hydrogen ion activity (pH). This is accomplished by measurements in several buffer systems, each of which provides a nearly constant, different pH. Alternatively, it may be desirable to measure the effects of other variables on a pH-sensitive system, by stabilizing the pH at a convenient value with a particular buffer.

Effectiveness. Buffer action depends upon the fact that, if two or more reactions coexist in a solution, then the chemical potential of any species is common to all reactions in which it takes part, and may be defined by specification of the chemical potentials of all other species in any one of the reactions. To be effective, a buffer must be able to respond to an increase as well as a decrease of the species to be buffered. In order to do so, it is necessary that the proton transfer step of the buffer be reversible with respect to the species involved, in the reaction to be buffered. In aqueous solution the proton transfer between most acids, their conjugate bases and water, is so rapid and reversible that the dominant direct source of protons for a chemical reaction is H_3O^+, the hydronium ion.

An acid-base buffer reaction in water is defined by reversible reaction (1), and the equilibrium constant K_a shown by Eq. (2).

$$BH^+ + H_2O \rightleftharpoons B + H_3O^+ \tag{1}$$

$$K_a = \frac{[H_3O^+]\,[B]}{[BH^+]\,[H_2O]} \tag{2}$$

In Eq. (2) the square brackets designate the activity of the species involved. In normal concentrations of buffer (0.1 mole/liter) the activity of the solvent water is essentially constant and approximately that of pure water (55.5 *M*). Thus the position of the equilibrium may be defined by specifying the activity of any two of the three variable species in reaction (1). Normally this is by means of the equilibrium expression shown as Eq. (3) which, upon converting to a logarithmic form, can be reduced to Eq. (4). Here f

$$[H_3O^+] = K_a \frac{[BH^+]}{[B]} \tag{3}$$

$$\mathrm{pH} = \mathrm{p}K_a - \log\frac{(1-f)}{f} - \log\frac{\gamma_{BH^+}}{\gamma_B} \tag{4}$$

is the fraction of the total buffer concentration, $(BH^+) + (B)$ existing as B, and γ is the activity coefficient relating activity a to concentration X. This relation is shown by Eq. (5). Thus a buffer pH is

$$a_x = \gamma_x(X) \tag{5}$$

approximately defined by the dissociation constant K_a of the weak acid system and the ratio of acid to conjugate base concentrations. However, the third term in Eq. (4) indicates that the pH is dependent on the change in activity coefficients with concentration. Effects of this dependency may be eliminated in practice by providing a high and essentially invariant ionic environment in the form of an added pH-neutral strong electrolyte such as KNO_3 or NaCl.

Buffer capacity π is defined as the change in added H_3O^+ necessary to produce a given change in pH, $d[H_3O^+]/d\mathrm{pH}$. Since the buffer comes to equilibrium with added H_3O^+, $1/\pi$ may also be defined as $d\mathrm{pH}/df$. Inspection of Eq. (4) shows $1/\pi$ to be a minimum when $f = 1 - f$; hence a given buffer system has its highest capacity in a solution composed of equal parts BH^+ and B, and the capacity is directly proportional to the concentrations of BH^+ and B. For these reasons buffers are normally used at concentrations 10–100 times higher than the system to be controlled and, if possible, are selected so that the desired pH

is approximately equal to pK_a for the buffer system. As a general rule, weak acid systems are not used to stabilize solutions whose pH is more than 2 pH units removed from pK_a, to ensure that the ratio of BH^+ to B will fall in the range 100–0.01.

Water as solvent. Buffers are particularly effective in water, because of the unusual properties of water as a solvent. Its high dielectic constant (80) tends to promote the existence of formally charged ions (ionization). Because it has both an acidic (H) and a basic (O) group, it may form bonds with ionic species leading to an organized sheath of solvent surrounding an ion (solvation). Water also tends to self-ionize to form its own conjugate acid-base system as shown by Eq. (6*b*), in which K_{ap} is the autoprotolysis constant. The strength of an acid (or base) in solvent water cannot be separated from reaction (6*a*) and

$$2H_2O \rightleftharpoons H_3O^+ + OH^- \qquad (6a)$$

$$K_{ap} = [H_3O^+][OH^-] = 10^{-14} \qquad (6b)$$

the familiar acid (or base) dissociation equilibrium reaction (1). Strong acids are those for which the K_a of Eq. (2) is very large; weak acids do not completely transfer the proton to water. The strongest acid which may exist in water is H_3O^+; the strongest base is OH^-. Thus, the maximum range of acid level which a solvent can support is governed by its own acid-base properties. In water this range is 14 pH units, or 14 orders of magnitude change in activity of H_3O^+. *See* SUPERACID.

The mechanism of buffer action may be regarded as a sequence of the proton transfer steps implied in reaction (1) coupled with reaction (6*a*). For example, the result of the chemical production or deliberate addition of an acid, HA, is to cause the water autoprotolysis reaction and the buffer acid reaction to respond to the change shown by reactions (7) and (8). Addition of a base would be accommodated by

$$HA + H_2O \rightleftharpoons H_3O^+ + A^- \qquad (7)$$

$$H_3O^+ + B \longrightarrow BH^+ + H_2O \qquad (8)$$

the reverse of (7) and (8). The effect of adding HA depends on the position of the equilibrium shown in reaction (7); buffer capacity π is usually defined in terms of H_3O^+ added because H_3O^+ is the strongest possible acid in aqueous solution, and would tend to create the maximum possible change in solution pH per mole of added acid. If HA is relatively weak so its degree of dissociation, in reaction (1), is small, its effective H_3O^+ addition may be calculated through Eq. (9), where C_a is the concentration of added HA.

$$[H_3O^+] \cong \sqrt{K_a C_a} \qquad (9)$$

A simple calculation using Eq. (9) shows that a given buffer solution will undergo the same change in pH for the addition of 0.1 mole/liter of a weak acid such as acetic acid ($K_a = 10^{-5}$) as for the addition of 0.01 mole/liter of strong acids such as HCl, $HClO_4$, or HNO_3.

In studies of rates of chemical reactions at constant pH, it is necessary that the proton transfer processes of the buffer acid and base and the solvent be rapid with respect to the primary reaction. The phosphate ($HPO_4^{2-}-PO_4^{3-}$) and carbonate ($HCO_3^{-}-CO_3^{2-}$) systems, among others, sometimes give anomalous effects because this condition may not be obtained. Buffer rate effects are manifested in different reaction rates for a chemical system in two different buffers or otherwise identical ionic strength and nominal (equilibrium) pH. Later evidence seems to suggest that buffers of low-change type, for example, $NH_3-NH_4^+$, react more rapidly than high-charge types such as $HPO_4^{2-}-PO_4^{3-}$. *See* ACID AND BASE; ACID-BASE INDICATOR; IONIC EQUILIBRIUM.

A. M. Hartley

Bibliography. J. J. Cohen and J. P. Kassirer, *Acid-Base*, 1982; H. A. Laitinen, *Chemical Analysis*, 2d ed., 1975; D. D. Perrin and B. Dempsey, *Buffers for pH and Metal Ion Control*, 1979; D. A. Skoog, D. M. West, and F. J. Holler, *Fundamentals of Analytical Chemistry*, 7th ed., 1997.

Buffers (electronics)

Electronic circuits whose primary function is to connect a high-impedance source to a low-impedance load without significant attenuation or distortion of the signal. Thus, the output voltage of a buffer replicates the input voltage without loading the source. An ideal voltage buffer is an amplifier with the following properties: unity gain, $A_B = 1$; zero output impedance, $Z_{out} = 0$; and infinite input impedance, $Z_{in} = \infty$. For example, if the voltage from a high-impedance source, say a strain-gage sensor with 100 kΩ output resistance, must be processed by further circuitry with an input impedance of, say, 500 Ω, the signal will be attenuated to only 500/100,500 ≈ 0.5% of the sensor voltage if the two circuits are directly connected, whereas the full strain-gage voltage will be available if a buffer is used.

Operation. Buffers are generally applied in analog systems to minimize loss of signal strength due to excessive loading of output nodes (**Fig. 1***a*). Two kinds of circuits are frequently used: the operational-amplifier-based buffer and the transistor follower.

Operational-amplifier-based buffer. This circuit (Fig. 1*b*) is based on an operational amplifier (op amp) with unity-gain feedback. The open-loop gain, $A(s)$, of the operational amplifier should be very high. To form the buffer, the amplifier is placed in a feedback loop. The buffer gain, $A_B(s)$, is then given by Eq. (1).

$$A_B(s) = \frac{V_{out}}{V_{in}} = \frac{A(s)}{1 + A(s)} \qquad (1)$$

Here, $s = j\omega$ is the Laplace transform variable, $j = \sqrt{-1}$; $\omega = 2\pi f$ is the radian frequency in radians per second (rad/s); and f is the frequency in hertz (Hz). The magnitude of A_B is approximately equal to unity, that is, $|A_B|$ approaches 1, if $|A|$ becomes very large. A common representation of the frequency dependence of the operational-amplifier gain is given

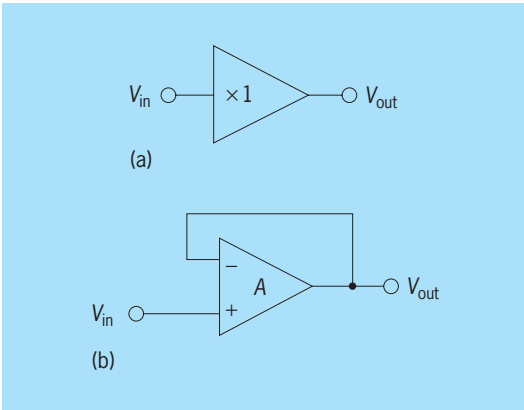

Fig. 1. Buffer circuit. (*a*) Circuit schematic symbol. (*b*) Operational-amplifier-based circuit.

by Eq. (2), where ω_t is the operational amplifier's

$$A(s) = \frac{\omega_t}{s} \tag{2}$$

unity-gain frequency. By using this notation, Eq. (1) becomes Eq. (3), which shows that the buffer's band-

$$A_B = \frac{\omega_t}{s + \omega_t} \tag{3}$$

width is approximately equal to the unity-gain frequency of the operational amplifier, typically 1 MHz or higher.

Under the assumption that the frequency of interest is much less than ω_t, it follows from Eq. (2) that the magnitude of the operational-amplifier gain, $A(s)$, is much greater than 1. In that case, it can be shown that, because of the feedback action, the buffer's input impedance is much larger than that of the operational amplifier itself [by a factor of $A(s)$]. Similarly, the buffer's output impedance is much smaller than that of the operational amplifier [again, by a factor of $A(s)$].

The very low output impedance of operational-amplifier-based buffers assures that a load impedance, $Z_L(s)$, does not affect the buffer's gain, A_B. Also, operational-amplifier-based buffers have no systematic offset. The high-impedance input node of a buffer may in practice have to be shielded to prevent random noise from coupling into the circuit.

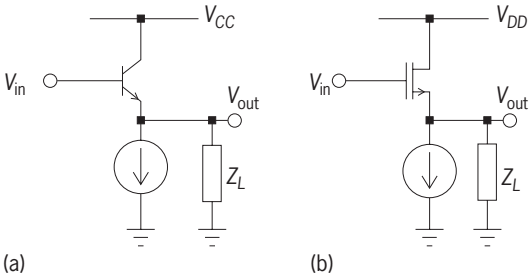

Fig. 2. Follower-type buffer circuits with load Z_L. (*a*) Bipolar emitter follower. An *npn* follower circuit is shown. V_{CC} = collector supply voltage. (*b*) Field-effect-transistor (FET) source follower. An *n*-type metal oxide semiconductor FET (MOSFET) follower circuit is shown. V_{DD} = drain supply voltage.

This shielding can be accomplished with a coaxial cable. To eliminate the capacitive loading of the source by the effective input capacitance of the cable, the shield can be driven with the output voltage of the buffer so that no voltage difference exists between the signal line and the shield. The driven shield is referred to as the guard. *See* AMPLIFIER; ELECTRICAL SHIELDING; ELECTROMAGNETIC COMPATIBILITY; OPERATIONAL AMPLIFIER.

Transistor followers. The bipolar junction transistor (BJT) emitter follower (**Fig. 2***a*) and the field-effect transistor (FET) source follower (Fig. 2*b*) are very simple but effective buffer circuits. Both consist of a single transistor and a bias-current source; they are used in applications where power consumption and circuit area must be reduced to a minimum or where specifications are not too demanding.

An important parameter for BJT operation is the current gain, β, a number of the order of 100. For large values of β, the small-signal voltage gain of the common-collector circuit or emitter follower is close to but smaller than 1. The other two parameters that are important for the buffer's operation are the input and output impedances. The input impedance is obtained approximately by multiplying the load impedance by the current gain, and the output impedance is approximately equal to the resistance at the base divided by the current gain.

The FET source follower also has an ac voltage gain that is less than 1. The output impedance is approximately the reciprocal of the transconductance of the FET transistor, g_m. The input impedance of the FET buffer is infinite, apart from a small gate capacitance.

The performance of a transistor follower circuit depends strongly on the source and load impedances, that is, on the surrounding circuitry. In fact, the transistors are so fast that the frequency response is usually determined by loading. In general, follower circuits exhibit a systematic direct-current (dc) offset equal to the base-to-emitter voltage, V_{BE}, in BJTs and equal to the gate-to-source voltage, V_{GS}, for FET. Only followers made with depletion-mode field-effect transistors can be biased with zero V_{GS} to avoid this offset. *See* EMITTER FOLLOWER; TRANSISTOR.

Other specifications. Specifications that apply to buffer circuits comprise dc and large-signal parameters on the one hand and ac and small-signal parameters on the other.

DC and large-signal parameters. Buffer circuits should have small dc offset voltages (dc outputs when no input is applied), small bias currents (to minimize the effect of high-impedance sources), large linear signal swing (to minimize distortion), and high slew rate (to handle fast transitions of the applied signals).

AC and small-signal parameters. Buffers should have a low-frequency gain of unity and wide bandwidth (to reproduce the applied signals faithfully), low phase margins (to prevent peaking and overshoots), and low equivalent input-referred noise (to have wide dynamic range). Field-effect-transistor input buffers

exhibit the lowest noise for high-impedance signal sources. *See* ELECTRICAL NOISE; GAIN.

Rolf Schaumann

Bibliography. P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3d ed., 1993; R. Gregorian and G. C. Temes, *Analog MOS Integrated Circuits for Signal Processing*, 1986; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Buhrstone mill

A mill for grinding or pulverizing, in which a flat siliceous rock, generally of cellular quartz, rotates against a stationary stone of the same material. The Buhrstone mill is one of the oldest types of mill and, with either horizontal or vertical stones, has long been used to grind grains and hard materials. Grooves in the stones facilitate the movement of the material. Fineness of the product is controlled by the pressure between the stones and by the grinding speed. A finely ground product is achieved by slowly rotating the stone at a high pressure against the materials and its mate (see **illus.**). The capacity,

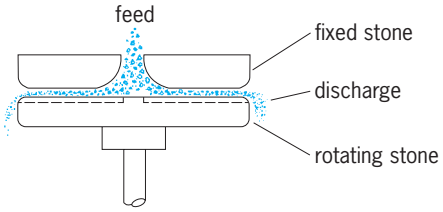

Buhrstone mill. Material moves toward outer edge of the stones where finely ground product is discharged.

or output, of a Buhrstone mill is low and its power requirements are high. The stones require frequent maintenance, even when grinding only slightly abrasive materials. *See* CRUSHING AND PULVERIZING; GRINDING MILL.

George W. Kessler

Buildings

Fixed permanent structures, more or less enclosed and designed to use as housing or shelter or to serve the needs of commerce and industry. The history of buildings is almost as old as human history. Most contemporary building techniques are possible because of materials and methods of construction that have been introduced since about 1800.

Iron and steel. With the use of cast and wrought iron early in the nineteenth century and the growth of the steel industry at the end of that century, large-scale buildings became common, and many were constructed. Notable were the Crystal Palace (London, 1851) where mass production and a demountable building were first demonstrated, the great London train sheds (Paddington Station, Victoria Station), the Eiffel Tower (Paris, 1889), and the first skyscrapers (Chicago, 1890s).

Iron and steel building components are noncombustible, and their basic strength is excellent compared to both masonry and timber of similar dimension. The strength-to-weight ratio of steel is also good. Steel is equally strong in tension and compression and possesses excellent ductility, a highly desirable quality in building design.

Contemporary applications of structural steel in building construction generally utilize rolled shapes in the form of wide flange and I beams, pipes and tubes, channels, angles, and plates. These are fabricated and erected into frameworks of beams, girders, and columns. Floors are usually concrete slabs cast on corrugated metal deck or on removable wood forms. *See* FLOOR CONSTRUCTION; STRUCTURAL STEEL.

Concrete. Another important building material, concrete, was known to builders in Roman times. Concrete is a mixture of cement, coarse and fine aggregate (usually stone and sand), and water, with other chemical admixtures sometimes introduced to impart special qualities. The material is inherently weak in tension and must be reinforced by means of steel bars embedded in and bonded to the concrete matrix. This combination of nonhomogeneous materials, called reinforced concrete, is utilized in many areas of building construction, including foundations, walls, columns, beams, floors, and roofs. *See* COLUMN; FOUNDATIONS; ROOF CONSTRUCTION; WALL CONSTRUCTION.

In its early applications in the late nineteenth century, reinforced concrete was primarily used in industrial buildings. However, because of its ability to be molded into any shape, its inherent resistance to fire and weather, and its relatively low cost, reinforced concrete has become a desirable material for all types of buildings. Construction of reinforced concrete structures is more labor intensive than steel structures because of the need to construct elaborate formwork, place the reinforcement accurately where it is required, convey the concrete materials from the mixer to the final destination, and finish and cure any exposed surfaces. Reinforced concrete structures of 40 stories and more are not uncommon. While beam and column frames are often used, reinforced concrete construction frequently involves the flat plate, a system in which the floor slab is of uniform depth with no beams projecting beneath it and with the plate supported only by columns, which may be randomly spaced rather than on a regular grid. *See* REINFORCED CONCRETE.

Concrete is subject to plastic flow (sometimes called creep) when subjected to compressive forces over a period of time. Therefore collateral building materials such as exterior walls, interior partitions, and vertical pipes and ducts must be detailed to accommodate this movement. *See* CONCRETE.

Timber. Two inventions made the use of timber extremely viable in small-scale structures: the machine-powered sawmill and the automatic nail-making machine. Previously, timbers were either hand hewed or sawed and nails were hand forged.

In North America, where large softwood forests

were plentiful, the milling of small-dimension lumber [2–3 in. (5–7.5 cm) wide by 4–12 in. (10–30 cm) deep] gave rise to the balloon frame house in the latter part of the nineteenth century. In this technique, closely space studs, joists, and rafters [no more than 24 in. (60 cm) on center] are fastened together with simple square cuts and nails. The balloon frame allowed relatively unskilled persons to erect simple frame houses. In the twentieth century, the balloon frame gave way to the platform frame, in which the studs were capped at each floor rather than running continuously for two stories (**Fig. 1**). The platform technique further simplified the framing system and compensated more uniformly for the drying shrinkage of newly milled wood with high moisture content that was often employed in construction.

Masonry. Masonry is a widely used construction technique, and perhaps the oldest building material. The three most common masonry materials are stone (quarried from natural geologic formations), brick (manufactured from clay that is exposed to high temperature in kilns), and concrete masonry units (solid or hollow blocks manufactured from carefully controlled concrete mixes). These materials are used alone or in combination, with each unit separated from the adjacent one by a bed of mortar. *See* BRICK; MORTAR; STONE AND STONE PRODUCTS.

The strength of a masonry wall depends greatly on the quality of construction. Since quality varies widely, it is desirable to introduce a relatively large factor of safety into the design. Masonry has been used in structural supporting walls built as high as 20 stories, although commonly the limit is 8–12 stories. Reinforcement can be introduced in the form of steel wires in horizontal bed joints as well as vertically in the form of rods set into grouted hollow cells of concrete masonry or in the grouted space between wythes of a brick wall. Reinforcement greatly enhances the strength of a masonry wall. *See* MASONRY.

Other materials. As new materials were developed, the technology of constructing buildings advanced, and it is anticipated that progress will continue. Newer metals include high-strength alloys of steel as well as products developed for space programs that have very high strength-to-weight ratios. Other desirable properties involve increased strength as well as resistance to corrosion, high temperature from fires, and fatigue.

Reinforced concrete with compressive strength increased from the older common range of 3000–5000 lb/in.^{2} (21–35 megapascals) up to 15,000–18,000 lb/in.^{2} (103–124 MPa) has been developed. This is of particular value in the design of columns in very tall buildings. Plastics are used in many building applications. However, these materials require improvements in strength and stiffness, long-term dimensional stability, resistance to high temperature and the degrading effects of ultraviolet radiation, and ease in being fastened and connected. Composite materials have been developed for application in buildings, and include sandwich panels in which the surfaces are bonded to a core. Combinations of steel and concrete (**Fig. 2**), masonry and steel reinforcement or prestress, timber and concrete, and timber and steel are in use. Other novel materials include high-performance fabric for roof coverings, structural adhesives, carbon fiber, and glass-fiber products.

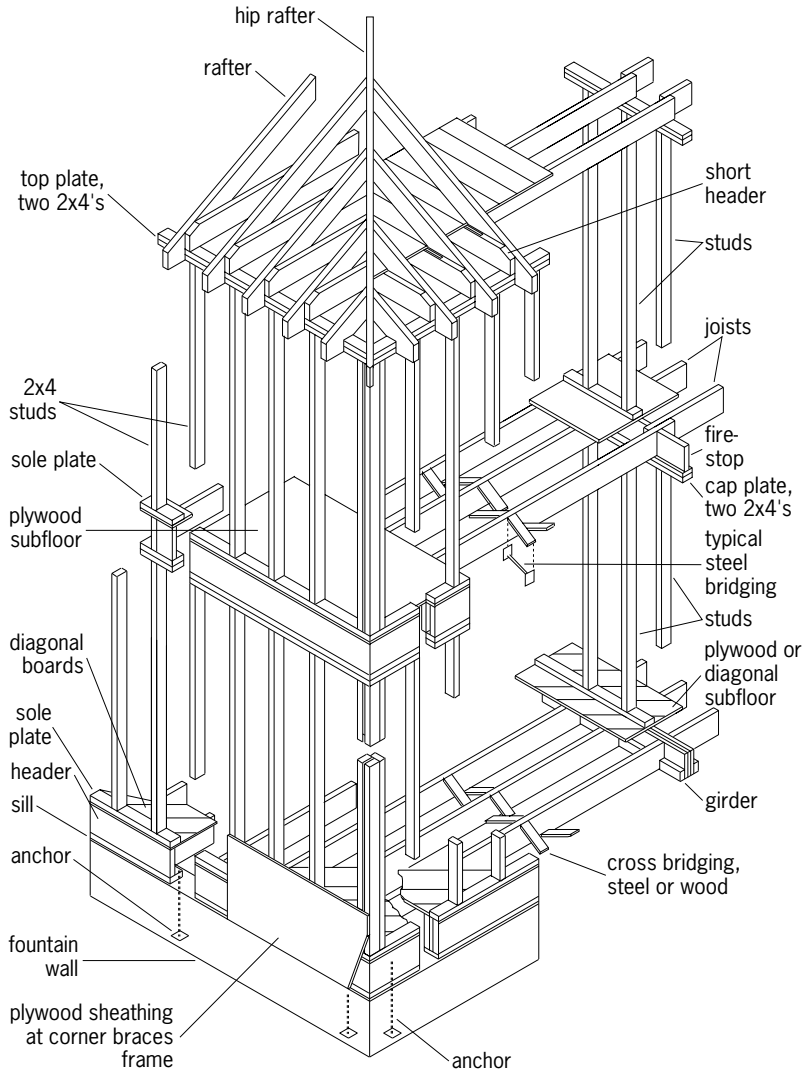

Fig. 1. Configuration of the platform frame construction. (*After C. Ramsey, Architectural Graphic Standards, 7th ed., Wiley, 1989*)

These modern materials are often lighter in weight than conventional products, facilitating long-span roofs that are made of steel or alloy cables and struts, coated fabric membranes, and reinforced plastic panels. An additional benefit of lightweight structures is their ability to be prefabricated in a factory and shipped to the job site, either as panels or as completed three-dimensional modules. Materials of high strength allow the construction of taller buildings, using columns of smaller dimensions. In addition, durability can be substantially improved in basic materials that will resist corrosion and deterioration due to atmospheric conditions including pollution and acid rain, as well as in secondary materials such as sealants (at building joints and junctures of dissimilar

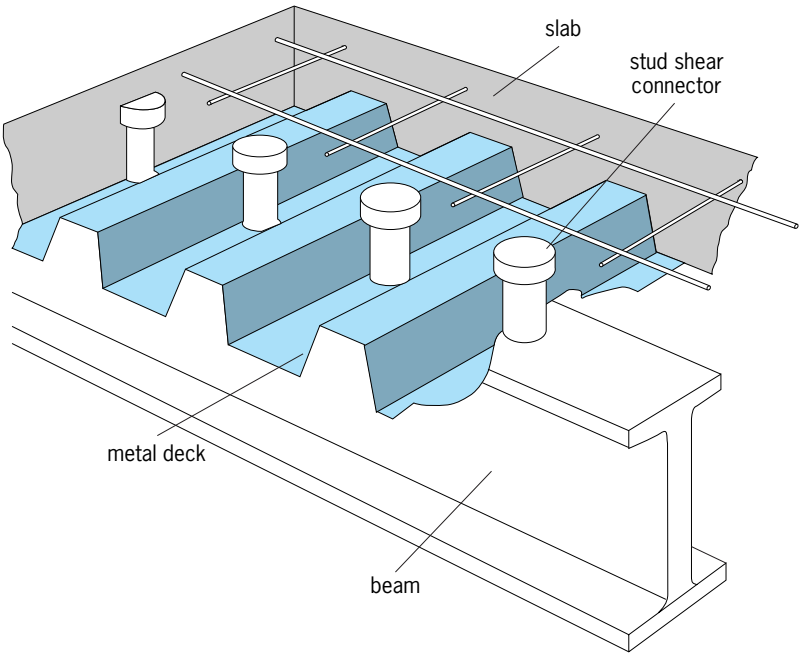

Fig. 2. Metal deck of composite design, providing a permanent form for concrete as well as a work platform. (*Nelson Studs Welding Systems*)

materials), adhesives, and connections (including bolts, screws, and nails).

Many building components are subassembled in factories prior to delivery to the job site. Wood structural members are made into glue-laminated beams using strip lumber, veneer lumber, particle board, or parallel strand lumber. Steel components include prefabricated open-web trusses and joists as well as three-dimensional space frames. Concrete components are factory precast to form hollow core planks and tee-shaped sections, often using a techniques for reinforcing called prestressing, in which steel wires or strands are stretched taut in the forms prior to concrete being cast around them.

Skyscrapers. Following the invention of the high-speed safety elevator, skyscrapers were developed at the end of the nineteenth and early in the twentieth century to maximize the economic return on parcels of land in urban environments. Earlier heavy-masonry-bearing-wall buildings had walls up to 6 ft (2 m) thick at their base to support as much as 16 stories of load. These walls occupied valuable space that could otherwise be rented to tenants. This drawback provided stimulus to the development of the skeleton steel frame, in which the thin exterior cladding does not participate in the support of the building but functions as a weather enclosure and a visual expression. These external skins (curtain walls) are often constructed of light aluminum or steel supports infilled with glass or metal panels. Curtain walls may also be fabricated of masonry veneer or precast concrete panels. They are designed to resist water and wind pressure and infiltration, and they are attached to the building frame for their primary support. Curtain walls may be built up of mullions and muntins, with the glass and metal in place on the building (the so-called stick system), or may be installed as a unit, with the framing members and glazing preassembled offsite.

The major structural problem that must be considered in the design of skyscrapers is the ability of the frame to resist lateral wind loads. The building must be both strong enough to resist the applied forces and stiff enough to limit the lateral displacement or drift to amounts that are tolerable to human comfort levels and to the performance of the collateral materials such as glass, interior partitions, and ceilings. In order to meet these criteria, methods have evolved

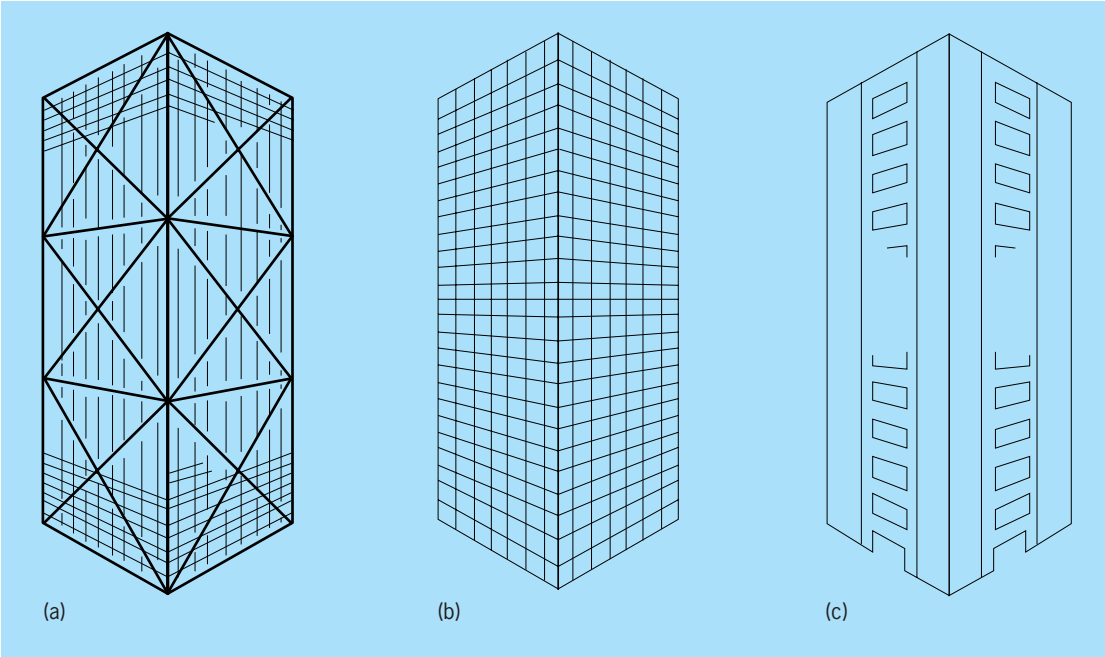

Fig. 3. Tube structural systems: (*a*) braced, (*b*) framed, and (*c*) shear wall.

since the skeleton frame was introduced. The simplest method of providing lateral rigidity is to ensure that the joints between girders and columns remain rigid, that is, their geometry remains unchanged. Thus if two members meet at a 90° angle, the design must ensure that this angle will not be changed when a load is applied. In steel frames, this requirement is accomplished by connecting the two members with adequately sized and spaced bolts, rivets, or welds. In reinforced concrete construction, the design must provide for adequately sized members and reinforcing steel that act together by virtue of the monolithic nature of the material. Rigid frame design is still the most economical method of framing buildings up to 20 stories tall. *See* STRUCTURAL ANALYSIS.

As buildings became taller than 20 stories, it was recognized that diagonal braces could be introduced between the top of one column and the bottom of an adjacent one to form a truss type of framework. The diagonals could form X's, V's, or K's and were found to be very efficient for buildings up to about 60 stories. The principle of a truss is that a series of triangles is interconnected, and these are found to be immutable shapes; that is, their geometry cannot be changed except by changing the lengths of the members. In reinforced concrete buildings, solid shear walls are used rather than diagonals. *See* TRUSS.

Special systems have been developed for constructing very tall buildings. One of the most effective is the tube system (used in the World Trade Center, New York), in which the perimeter framework of the building is composed of closely spaced columns (with small windows) and very rigid spandrel beams connected to form almost a solid tube with small penetrations. This is an efficient system because it maximizes the use of the framing material by locating it as far as possible from the center of the building, that is, at the perimeter (**Fig. 3**). Variations on the tube system include large-scale external X-bracing (as in the John Hancock Building, Chicago; **Fig. 4**) and the bundled tube (such as the Sears Tower, Chicago). In addition to these so-called passive framing systems, buildings can resist wind forces by means of tuned mass dampers. These are large weights, as much as 300–400 tons (270–360 metric tons), mounted on guided rollers or pendulums and located at the top of a skyscraper; they move counter to the direction of the wind, minimizing the deflection. The dampers are activated by the actual movement of the building, and they alter its dynamic response to wind.

Special-purpose types. Modern technology has created the need for different types of structures. Examples include aircraft maintenance hangars (the largest building in the world is the vertical assembly building at Cape Canaveral, Florida), communications and entertainment production centers, large stadiums (many of which were formerly outdoors but are now covered and climate controlled), transportation terminals, convention centers, specialized medical facilities and scientific laboratories, and high-technology manufacturing buildings. Some require long spans, others extremely stringent climate or vibration controls, and still others superclean environments.

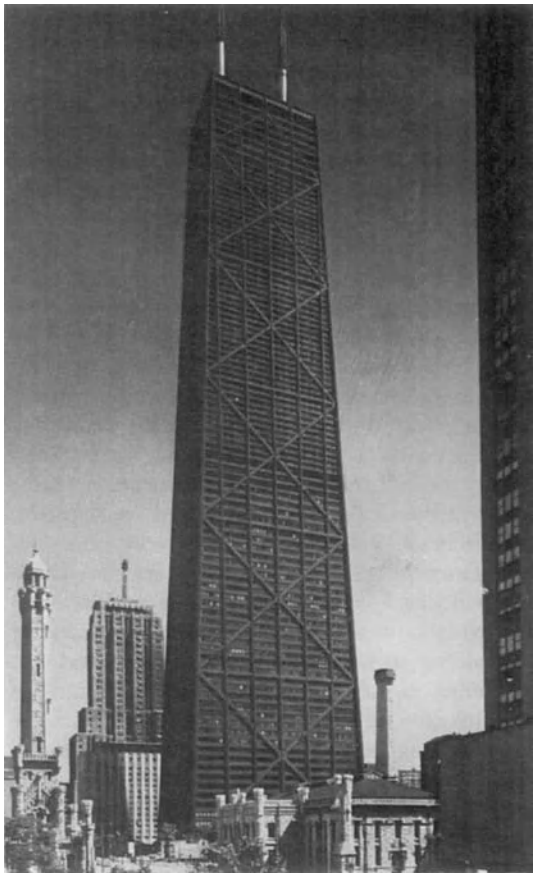

Fig. 4. Large-scale external X-bracing as used in the John Hancock Building, Chicago. (*ESTO*)

Building services. In order for buildings to be fully functional, they must be able to provide adequate levels of comfort and service. There are many methods used to supply the services of heating and cooling. Heating may be provided by radiation, conduction, or convection. Commonly, pipes containing steam or hot water from a central boiler are installed throughout a structure and connected to radiators or convectors in each room. Alternatively, air heated in a furnace may be delivered by means of fans blowing through sheet metal or fiber-glass ducts. Fuels commonly used in boilers are oil or gas. Cooling requires that warm air inside the building pass over coils containing cold water or refrigerant that exchange the heat, so that the cooled air is delivered to the interior while the heat is conveyed to the outside. *See* AIR CONDITIONING; COMFORT HEATING.

Electrical systems are installed throughout buildings to provide lighting as well as power to operate appliances and machinery. In large buildings, high-voltage power is reduced by means of transformers to usable voltage. Elaborate distribution systems of conduits and cables, wired to circuit-breaker panels for overload protection, are installed. Signal systems for telephones, computers, and alarms are also commonly specified and built. Finally, there

is plumbing service, which delivers hot and cold water and carries away wastewater as well as storm water into disposal systems such as sewers or septic systems.

Loads. Buildings are designed to resist loads due to their own weight, to environmental phenomena, and from the occupants' usage. The self-weight of a building, called dead load, is relatively easy to calculate if the composition and thickness of all of the materials are known. Included in the dead load are the building frame, walls, floors, roof, ceilings, partitions, finishes, and service equipment—that is, everything that is fixed and immovable.

Environmentally applied loads include rain, which may cause ponding, and snow and ice. Snow loads are calculated based on the weather history of the area in which the building is located. All of the loads mentioned thus far are vertical loads, but the effect of wind on a building is usually experienced as a lateral, horizontal force. Wind velocity records are kept by local weather bureaus; loads are then calculated and based on statistical probability of occurrence (usually a 50- or a 100-year return period) of peak wind speeds. These wind velocities are converted to force units, with gust loads as well as allowances for the shape of the building, its height, and exposure factored in.

Another significant load to which buildings are subjected is the force of earthquakes. The understanding of seismic activity is an emerging science. So too is the understanding of how buildings respond to seismic forces. Calculation of the probability of an earthquake occurring in any given area is based on past seismic experience as well as determination of the locations of fault lines and tectonic plates. However, it has not been possible to predict timing, location, or specific intensity of earthquakes. In fact, earthquake record-keeping, like that of snow and wind, has been effective for slightly more than 100 years, a very short period of time in the history of the Earth. Thus the models of statistical probability of the occurrence or the magnitude of an event or a storm are based on incomplete information. In many locations, building structures have been inadequately designed to meet unexpectedly large forces due to these environmental loads. *See* EARTHQUAKE.

Seismic loads, unlike most other loads except for wind, are dynamic in character rather than static. Furthermore, each seismic event produces a unique spectrum of vibrations to which a building may be subjected. The shaking of the ground causes both horizontal and vertical acceleration in a building, whose response may go through a number of vibration modes.

Engineers have devised a number of methods by which buildings can resist significant seismic loads. One is to design a maximum of energy absorption into the building by providing ductility in the frame and its connections. A second method involves an attempt to separate the superstructure of the building from ground-induced vibration—a method called base isolation. In this system, shock-absorbent material is inserted between the foundation and the superstructure to prevent vibrations from traveling up into the building. Other principles of seismic resistance require that all parts of the building be connected, that ornamental and other nonstructural features be securely attached to the building frame, and that mechanical and electrical systems be designed to accommodate anticipated movement. *See* SEISMIC RISK.

Fire protection. The danger of fire in buildings has several aspects. Of primary importance is the assurance that all occupants can exit safely and that firefighters can perform their work with minimal danger. The second consideration involves the protection of property, the building, and its contents. In the initial planning of a building, the location, number, and size of exits must be carefully considered in relation to the anticipated occupancy and the material of construction. Where it is not possible to provide sufficient access to exit doors at ground level, fire escapes (generally steel-bar platforms and stairs) are affixed to the sides of buildings.

Sophisticated fire detection systems are available that can sense both smoke and heat. These sound audible alarms and directly contact municipal fire departments and building safety officers. In addition, the alarm may automatically shut down ventilation systems to prevent smoke from spreading, may cause elevators to return to the ground floor where they remain until the danger is passed, and may close fire doors and dampers to compartmentalize the spread of smoke or flames. Supplementing this passive detection are automatic sprinkler systems, in which water is delivered to spray heads spaced to provide complete coverage of the area to be protected. The sprinklers are activated by heat or by the alarm system. *See* AUTOMATIC SPRINKLER SYSTEM; FIRE DETECTOR; FIRE TECHNOLOGY.

Building framing system are rated for their ability to withstand fires. Actual assemblages of building materials are tested in simulated fire conditions where they are exposed to flames of a certain temperature and a water-hose stream of a specified pressure. They are then rated based on how long (generally between 1 and 4 h) they can remain intact when exposed to the design fire conditions. Structural steel rapidly loses its ability to safely support loads when its temperature exceeds 800°F (427°C). Thus steel is protected by encasing the structural members with sprayed-on mineral fibers (formerly asbestos was used), concrete, masonry, plaster, or gypsum dry wall. In one technology a thin sprayed or troweled intumescent coating turns into foam when exposed to heat; the foam then serves as an insulator for a given period of time. Wood members can be impregnated with salts (zinc chloride, ammonium sulfate, borax, or boric acid) that render the wood resistant to combustion. Reinforced concrete and masonry are naturally resistant to fire for a minimum of several hours.

Certain building materials release toxic substances when exposed to high temperature. Thus while a building structure may be safe, the occupants can

be injured or killed by toxic emissions (smoke and fumes) before they have an opportunity to exit. Such materials are no longer allowed in many buildings.

In high-rise buildings the problems of fire safety are exacerbated by the inherent difficulties in avoiding a chimney stack effect for smoke and flames in elevator shafts, stairwells, ventilation shafts, and heating/cooling air ducts. Efforts are directed at compartmentalizing flames and smoke. This is often difficult to achieve, since doorways to exit stairs are frequently opened by persons fleeing the danger. These exit passages can be positively pressurized with air to deter the intrusion of smoke. Planning for evacuation of occupants without panic includes providing emergency lighting, smoke-free access routes, frequent fire drills, and positive communications during an emergency.

Building codes. The process of building is often regulated by governmental authorities through the use of building codes that have the force of law. In the United States there is no national code; rather there are regional, state, or even city building codes. Some of the more widely accepted codes are the Building Officials Conference of American (BOCA), used widely in the eastern United States; the Uniform Building Code (UBC), used in the western United States; and the Southern Building Code, used in the south. In Europe there are national codes and codes that transcend national boundaries (Eurocodes).

Codes establish classifications of buildings according to the proposed occupancy or use. Then, for any given type of construction (for example, wood, steel, or concrete), they establish minimum standards for exit and egress requirements, for height and area, and for fire resistance ratings. In addition, minimum loads are designated as well as requirements for natural light, ventilation, plumbing, and electrical services. Local codes are written to regulate zoning, stipulating items such as building type, occupancy, size, height, setbacks from property lines, and historic considerations.

Environmental concerns. The process of building raises large numbers of environmental issues. In many cases the owner must prepare an official environmental impact statement that considers the potential effect of the proposed building on traffic, air quality, sun and shadow, wind patterns, archeology, wildlife, and wetlands, as well as demands on existing utilities and services. Many building projects have had their original schemes severely altered or have been abandoned to avoid adverse effects on the environment.

One of the primary environmental concerns is energy conservation. In the initial design of a building, all systems are studied to obtain maximum efficiency. Heating and cooling are two of the largest consumers of energy, and a great deal of effort is directed toward minimizing energy consumption by techniques such as building orientation, sun shading, insulation, use of natural ventilation and outside air, recapture of waste heat, cogeneration (using waste heat to generate electricity), use of solar energy both actively and passively, and limiting heat generation from lighting. Efforts at reducing electric power consumption by designing more efficient lighting, power distribution, and machinery are also of high priority. Consumption of water and disposal of liquid and solid waste are additional concerns. Yet another aspect of energy conservation deals with the study of the quantity of embodied energy required to construct the original building. Considerations involve energy required to produce the material, transport it, and install it. *See* COGENERATION SYSTEMS; HEAT INSULATION; SOLAR HEATING AND COOLING; VENTILATION.

Planners are concerned with sustainable design and development. Included within this broad topic are not only considerations of energy conservation but also issues such as use of renewable resources, human scale, bioregionalism, and awareness of the need to nurture the planet for future inhabitants. Robert Silman

Bibliography. E. Allen, *Fundamentals of Building Construction*, 3d ed., 1998; L. S. Beedle (ed.), *Second Century of the Skyscraper*, Council on Tall Buildings and Urban Habitat, 1988; C. W. Condit, *American Building*, 2d ed., 1992; M. Salvadori, *Why Buildings Stand Up*, 1992.

Bulk-handling machines

A diversified group of materials-handling machines specialized in design and construction for handling unpackaged, divided materials.

Bulk material. Solid, free-flowing materials are said to be in bulk. The handling of these unpackaged, divided materials requires that the machinery both support their weight and confine them either to a desired path of travel for continuous conveyance or within a container for handling in discrete loads. Wet or sticky materials may also be handled successfully by some of the same machines used for bulk materials. Characteristics of materials that affect the selection of equipment for bulk handling include (1) the size of component particles, (2) flowability, (3) abrasiveness, (4) corrosiveness, (5) sensitivity to contamination, and (6) general conditions such as dampness, structure, or the presence of dust or noxious fumes.

Particles range in size from those that would pass through a fine mesh screen to those that would be encountered in earth-moving and mining processes. Fine granular materials are usually designated by their mesh size, which is an indication of the smallest mesh screen through which all or a specified percent of the particles will pass. As an example, a 100-mesh screen is one in which there are 100 openings per linear inch. Where the size of the bulk material varies, it is customary to indicate the percentage of each size in each mixture.

Flowability, corrosiveness, abrasiveness, and similar terms are relative and are usually modified by adjectives to indicate the degree of the characteristic, such as “mildly” corrosive or “highly” abrasive.

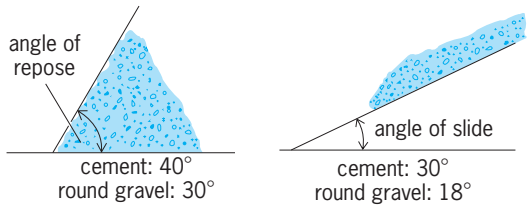

Fig. 1. Critical angles for bulk materials.

Many products are sensitive to contamination; this characteristic may be a determining factor in selecting bulk-handling equipment and its material of construction. For example, an enclosed conveyor is used to protect the material from exterior contamination; and construction from noncorrosive material, such as stainless steel, may be required to protect the handled material from interior contamination.

Consideration also must be given to the angle of repose and the angle of slide of the material (**Fig. 1**). The angle of repose is the maximum slope, expressed in degrees, which piled material will stand without sliding on itself. The angle of slide, also expressed in degrees, is the angle at which material flows freely on an inclined surface such as a chute.

Corrosiveness and other handling characteristics of materials are compiled and updated in literature provided by manufacturers of bulk-handling machines. Trade associations also publish technical bulletins on the relations of machines to jobs, applications, and operating costs. Related data are published by various governmental departments, especially the U.S. Department of Commerce.

Continuous bulk conveyor. Equipment that transports material continuously in a horizontal, inclined, or vertical direction in a predetermined path is a form of conveyor. The many different means used to convey bulk materials include gravity, belt, apron, bucket, skip hoist, flight or screw, dragline, vibrating or oscillating, and pneumatic conveyors. Wheel or roller conveyors cannot handle bulk materials.

Gravity chutes are the only unpowered conveyors used for bulk material. They permit only a downward movement of material. Variations such as chutes with steps or cleats are employed as a means of slowing product movement.

Belt conveyors of many varieties move bulk materials. Fabric belt conveyors have essentially the same operating components as those used for package service; however, these components are constructed more ruggedly to stand up under the more rigorous conditions imposed by carrying coal, gravel, chemicals, and other similar heavy bulk materials (**Fig. 2**). In the latter type, the belt runs on a bed of closely spaced rollers positioned to form a flat or troughed conveyor bed. Belts may also be made of such materials as rubber, metal, or open wire. Belt conveyors are only used within angles of 28° from horizontal. Materials feed onto belt conveyors from hoppers or from storage facilities overhead and may discharge over the end (**Fig. 3***a*) or, if the belt is flat, by being diverted off one side. Belt conveyors can handle most materials over long distances and up and down slopes. Their advantages include low power requirements, high capacities, simplicity, and dependable operation.

An apron conveyor is a form of belt conveyor, but differs in that the carrying surface is constructed of a series of metal aprons or pans pivotally linked together to make a continuous loop. The pans, which may overlap one another, are usually attached at each end to two strands of roller chain. The chain runs on steel tracks, movement being provided by suitable chain sprockets. Turned-up edges or side wings provide a troughed carrying surface (Fig. 3*b*). As in the case of the belt conveyor, the top strand of the apron is the carrying surface. This type of conveyor is suitable for handling large quantities of bulk material under severe service conditions and can operate at speeds up to approximately 100 ft (30 m) per minute, handling up to approximately 300 tons (270 metric tons) per hour; with the addition of cleats, an apron can convey up inclines to as steep as 60°. Apron

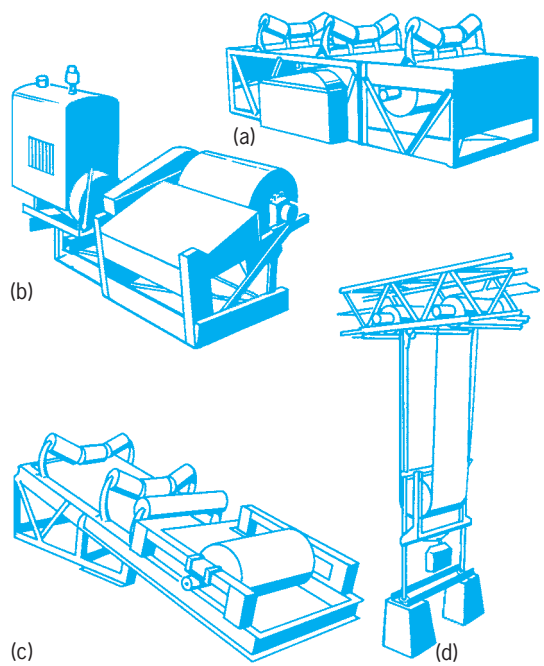

Fig. 2. Typical components of belt conveyor. (*a*) Electric motor wrap drive. (*b*) Gasoline engine head-end drive. (*c*) Screw type take-up. (*d*) Vertical gravity take-up.

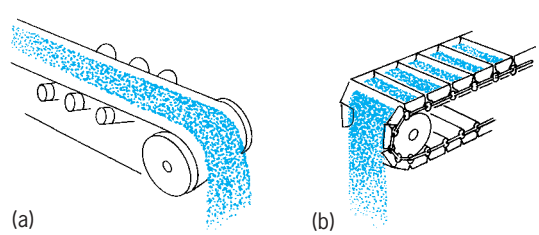

Fig. 3. Conveyors for abrasive materials, with material and conveying surface moving together. (*a*) Belt conveyor. (*b*) Apron conveyor.

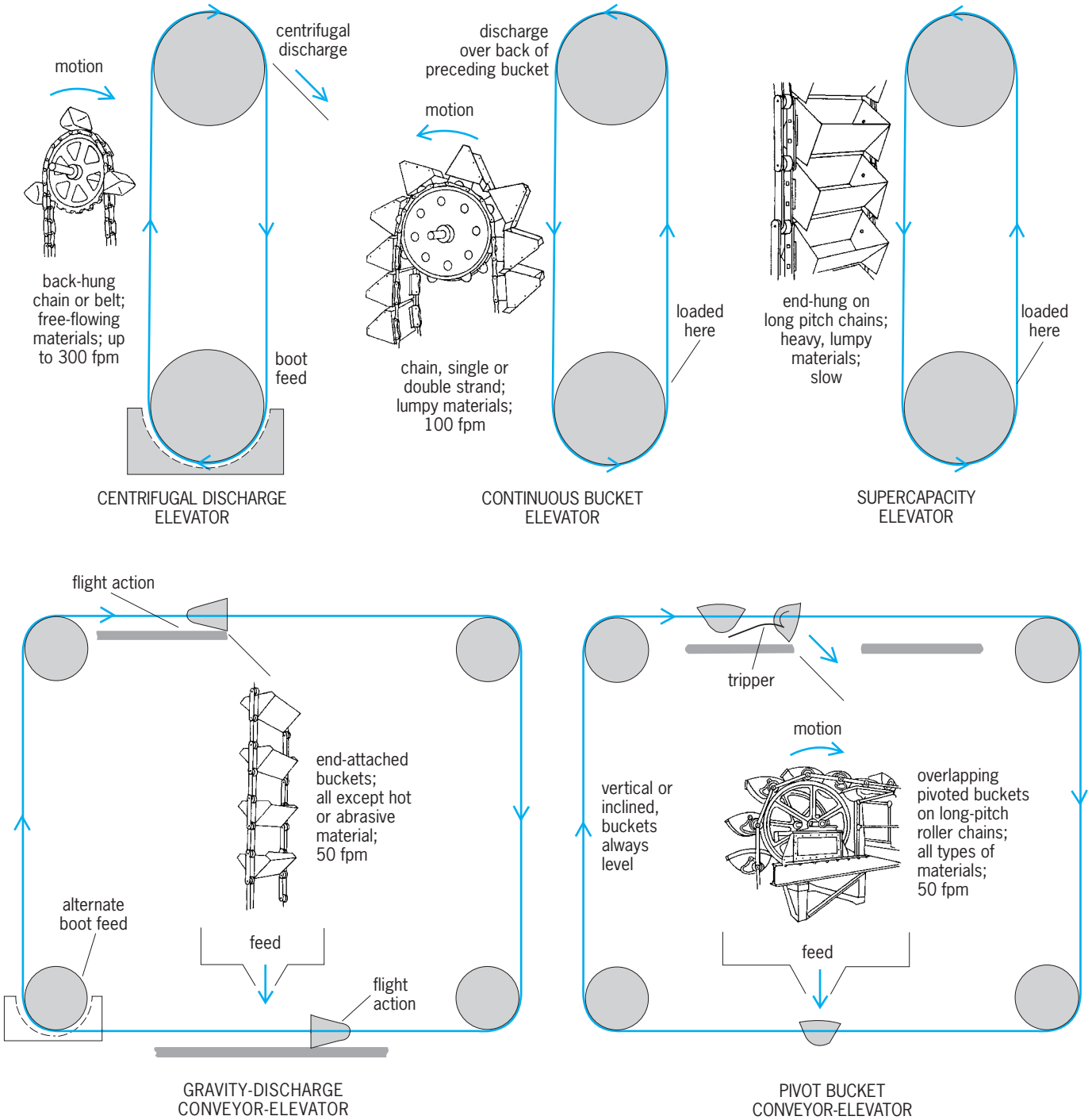

Fig. 4. Components of bucket conveyor-elevators and typical paths of travel for which they are used.

conveyors are most suitable for heavy, abrasive, or lumpy materials.

Bucket conveyors, as the name implies, are constructed of a series of buckets attached to one or two strands of chain or in some instances to a belt. These conveyors are most suitable for operating on a steep incline or vertical path, sometimes being referred to as elevating conveyors (**Fig. 4**). Bucket construction makes this type of conveyor most ideal for bulk materials such as sand or coal. Buckets are provided in a variety of shapes and are usually constructed of steel.

Flight conveyors employ flights, or bars attached to single or double strands of chain. The bars drag or push the material within an enclosed duct or trough. These are frequently referred to as drag conveyors. Although modifications of this type of conveyor are employed in a number of ways, its commonest usage is in a horizontal trough in which the lower strand of the flights actually moves the material. This type

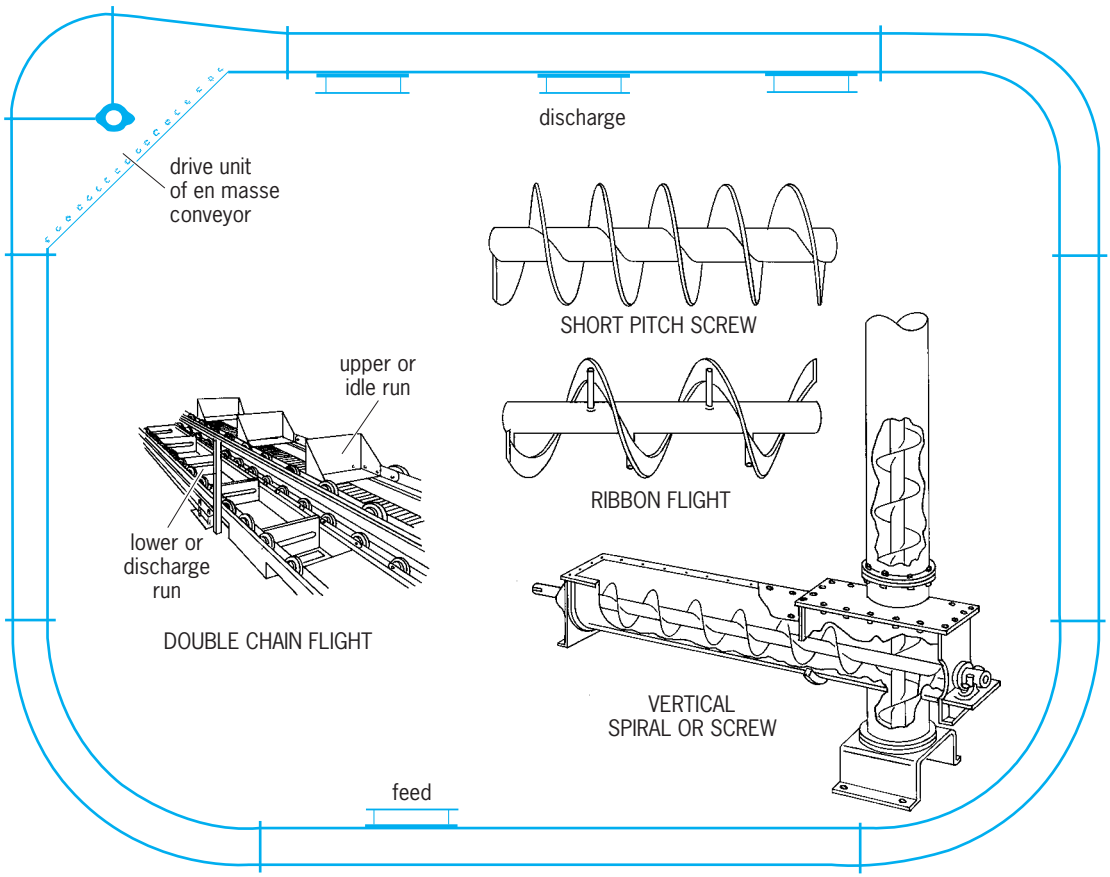

Fig. 5. Diagram of an en masse conveyor trough. The insets show some of the flight conveyors used to move bulk materials: double chain flight, and horizontal and vertical screw or spiral conveyors.

of conveyor is commonly used for moving bulk material such as coal or metal chips from machine tools (**Fig. 5**). Constant dragging action along the trough makes this type of conveyor unsuitable for materials which are extremely abrasive. Materials can be fed to the conveyor from any desired intermediate point and, by the use of gating, can be discharged at any desired point.

Dragline conveyors operate on basically the same principle as flight conveyors, as previously described. The chain, ruggedly made for this service, drags the material, such as clinkers and slag, along the bottom of a concrete trough. The dragging is done entirely by the links of the chain.

Spiral or screw conveyors rotate upon a single shaft to which are attached flights in the form of a helical screw. When the screw turns within a stationary trough or casing, the material advances (**Fig. 6***b*). These conveyors are used primarily for bulk materials of fine and moderate sizes, and can move material on horizontal, inclined, or vertical planes (Fig. 5). The addition of bars or paddles to the flight conveyor shaft also makes it ideal for mixing or blending the materials while they are being handled. In addition, enclosed troughs may be water- or steam-jacketed for cooling, heating, drying, and so forth.

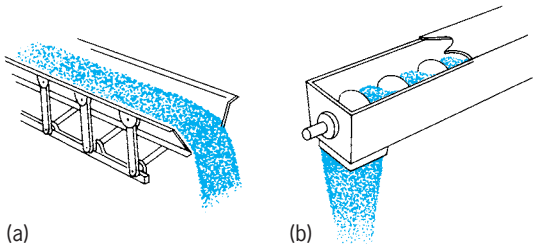

Fig. 6. Open and closed troughs. (*a*) Oscillating conveyor, with whole trough oscillating. (*b*) Screw conveyor, with helix driving material along trough.

Vibrating or oscillating conveyors employ the use of a pan or trough bed, attached to a vibrator or oscillating mechanism, designed to move forward slowly and draw back quickly (Fig. 6*a*). The inertia of the material keeps the load from being carried back so that it is automatically placed in a more advanced position on the carrying surface. Adaptations of this principle are also used for moving material up spiral paths vertically. Mechanical (spring), pneumatic (vibrator), and electrical (vibrator) devices provide the oscillating motion. These conveyors can handle hot, abrasive, stringy, or irregularly shaped materials. The trough can be made leakproof or, by enclosing it, dustproof.

Pneumatic, or air, conveyors employ air as the propelling media to move materials. One implementation of this principle is the movement within an air

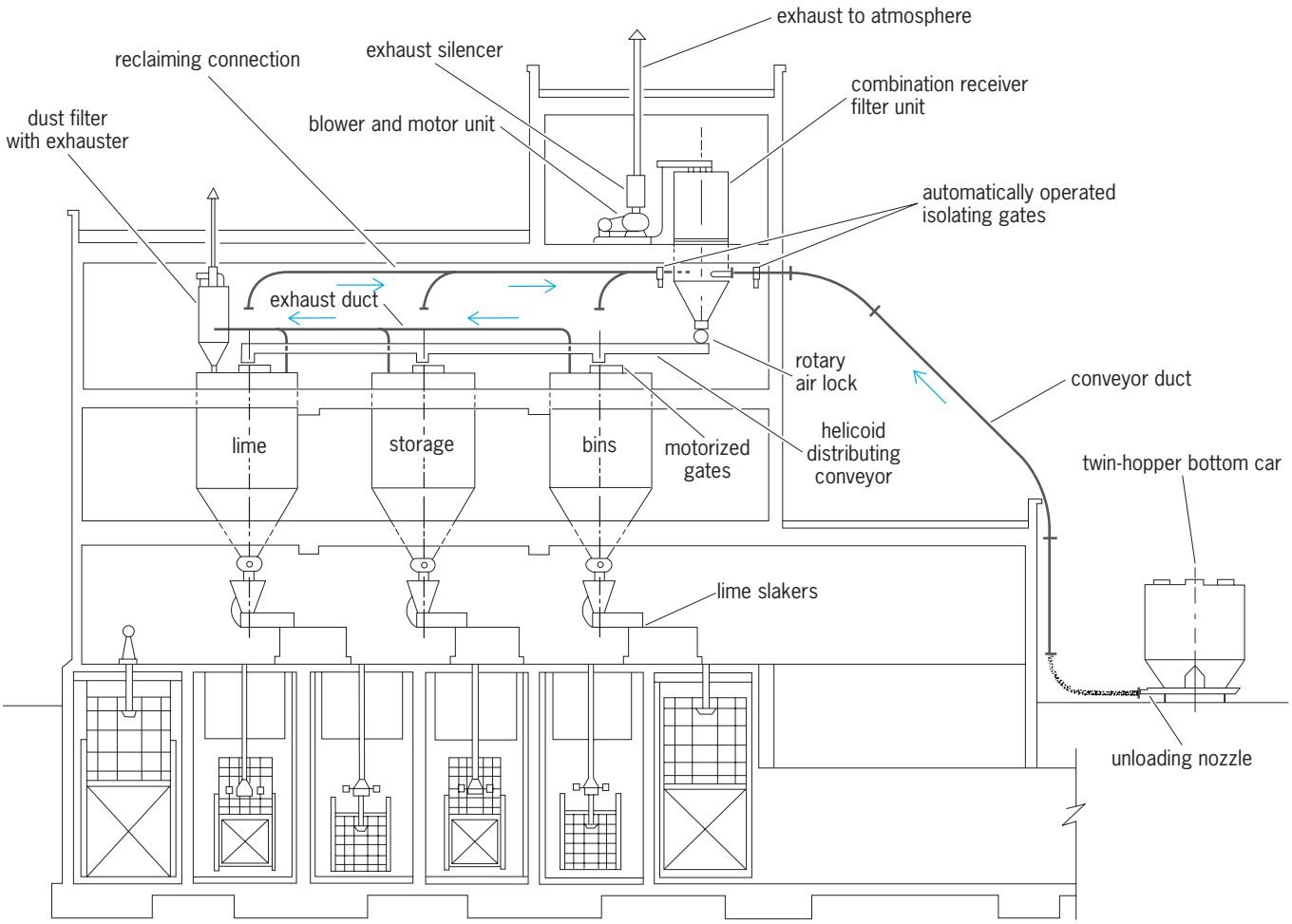

Fig. 7. Typical conveyor system, with hopper car delivering material, pneumatic lines unloading and lifting material, screw flight transferring it to bins, and gravity chutes discharging it. (*National Conveyors Co.*)

duct of cylindrical carriers, into which are placed currency, mail, and small parts for movement from one point to discharge at one of several points by use of diverters. Pneumatic pipe conveyors are widely used in industry, where they move granular materials, fine to moderate size, in original bulk form without need of internal carriers. An air compressor provides the air to move the material either by pressure or vacuum. Materials are introduced to the system by means of air locks, which are of the rotary or slide type. These locks are designed to permit the entry of the material with negligible loss of air. By means of diverted valves in the pneumatic pipe, a discharge can be effected at any one of a number of predetermined points. The small number of moving parts, the ability to move material in any direction, and the need for minimal prime plant floor space are among the advantages of this type of bulk-handling machine.

A pneumatic conveyor may form part of a conveyor system. The system comprises a network of machines, each handling the same product through its various processing stages; a single system may well employ numerous varieties of conveyors, such as pneumatic, screw, and vibrating conveyors. **Figure 7** shows a system for handling pebble lime in a water-treatment plant. The lime is received in railroad cars or trucks, unloaded and conveyed by a pneumatic vacuum system to a receiver filter, and discharged to a screw conveyor, from which it is selectively delivered to any desired storage bin. From the bins, the lime can be delivered on demand to the slakers below.

Another adaptation of the pneumatic conveyor is to activate a gravity conveyor. Such a conveyor handles dry pulverized materials through slightly inclined chutes. Air flows through the bottom of the chute, which is usually constructed of a porous medium, fluidizing the material and causing it to flow in the manner of a liquid. An advantage of this conveyor is that there are no moving parts.

Aerial tramways and cableways employ the use of a cab or carrier suspended by a grooved wheel on an overhead cable to transport materials over long distances, particularly where the terrain is such that truck or rail transportation is impractical. They are used primarily to meet the needs of such activities as dam construction, loading ships, bringing coal to

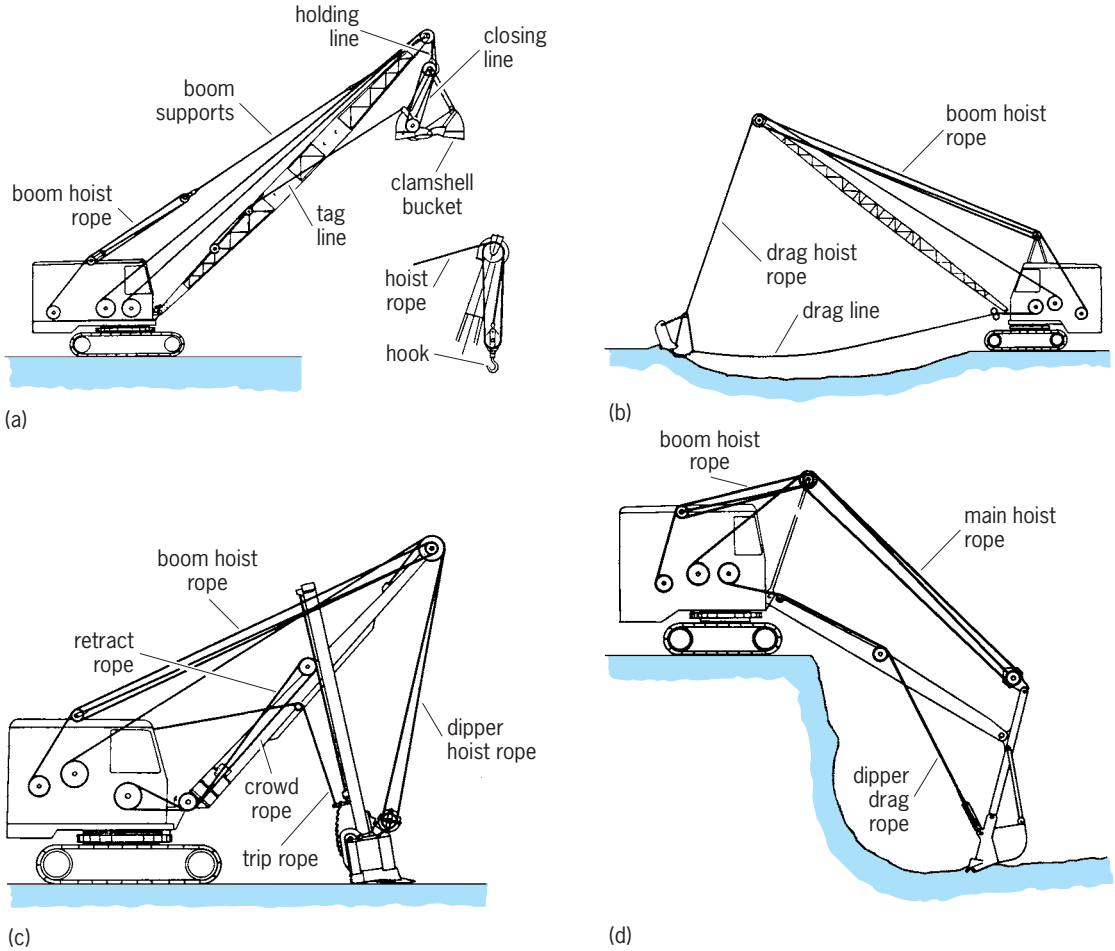

Fig. 8. Reeving diagrams for power crane fitted (*a*) with clamshell or (*b*) with scoop and for shovel fitted (*c*) with forward dipper or (*d*) with back hoe (backdigger or drag-shovel). (*United States Steel Corp.*)

and from power houses, and stock piling in open country.

Discontinuous bulk handlers. Power cranes and shovels perform many operations moving bulk materials in discrete loads. When functioning as cranes and fitted with the many below-the-hook devices available, they are used on construction jobs and in and around industrial plants. Such fittings as magnets, buckets, grabs, skullcrackers, and pile drivers enable cranes to handle many products.

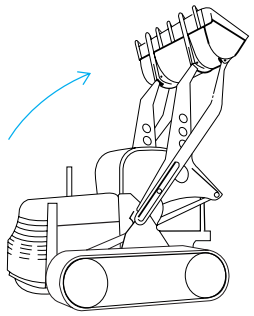

Fig. 9. Tractor fitted with shovel for up-and-over operation. (*Service Supply Corp.*)

The machines of the convertible, full-revolving type are mounted on crawlers, trucks, or wheels. Specialized front-end operating equipment is required for clamshell, dragline, lifting-crane, pile-driver, shovel, and hoe operations. Commercial sizes of these machines are nominally from $^1/_4$ to $2^1/_2$ yd^3 (0.19 to 1.9 m^3) as shovels and from $2^1/_2$ to 60 tons (2.25 to 54 metric tons).

The revolving superstructure consists of the rotating frame and the operating machinery thereon. It may be carried on a crawler mount, consisting of two continuous parallel crawler belts. A truck mount is a heavy-framed, rubber-tired carrier supported by two or more axles and having the general characteristics of a heavy-duty truck. The carrier may be controlled for road travel from a cab mounted on the carrier or located on the revolving superstructure. Machines of this type can also be secured on railroad mountings.

Six types of front-end operating equipment are standard: crane, clamshell, dragline, pile driver, shovel, and hoe. Common crane-boom equipment is used with crane, clamshell, dragline, and pile driver. The boom usually consists of two sections, between

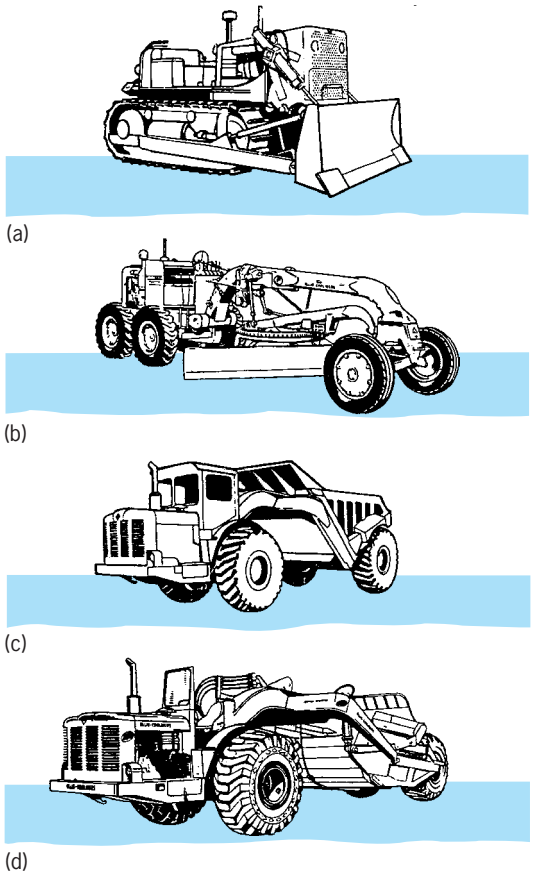

Fig. 10. Road-making machines. (*a*) Front-end shovel. (*b*) Motor grader. (*c*) Motor wagon. (*d*) Motor scraper.

which additional sections may be inserted. Shovel and hoe equipment have their distinctive mechanisms for operation (**Fig. 8**).

Scoops and shovels are used for handling bulk materials in plants and yards. Lighter models are usually wheel-mounted, while those for heavier duty are apt to be carried on tractors. Two basic types are recognized: those that load and dump only at the front or at the rear, and those with an up-and-over action which permits loading at the front end and discharging at the rear (**Fig. 9**). This last arrangement is frequently a time-saver in that it eliminates the need for maneuvering when loading wagons.

Specialized equipment for mechanized pit mining has been developed. Power cranes, shovels, and scoops are actively engaged in strip mines, quarries, and other earth-moving operations.

The rapid expansion of highway systems has led to the widespread use of road-making machines (**Fig. 10**). Some of these are modifications of equipment standard to this kind of work; others are highly specialized. Among the former are machines that have evolved by the addition of attachments to wheel and crawler tractors. Machines such as bulldozers and graders do not function as true handling equipment in that they do not pick up and transport materials but push them. On the other hand, self-loading scrapers do transport materials and are usually constructed so that they are either side- or end-dumping. Other general-purpose machines used in highway construction are trench diggers, hole diggers for utility poles, and cable-laying machines. Highly specialized equipment is used for surfacing the road with concrete or other material. *See* CONVEYOR; ELEVATING MACHINES; HOISTING MACHINES; INDUSTRIAL TRUCKS; MATERIALS-HANDLING EQUIPMENT; MONORAIL.

Arthur M. Perrin

Bibliography. S. C. Cowin (ed.), *Mechanics Applied to the Transport of Bulk Materials*, 1979; M. N. Kraus, *Pneumatic Conveying Systems for Bulk Materials*, 3d ed., 1991; F. J. Loeffler and C. R. Proctor (eds.), *Unit and Bulk Materials Handling*, 1980; J. S. Mason and A. S. Goldberg, *Bulk Solids Handling: An Introduction to the Practice and Technology*, 1986.

Buoy

An anchored or moored floating object, other than a lightship, intended as an aid to navigation. Buoys are the most numerous of all engineered aids to navigation, with some 20,000 in United States waters alone.

A great variety of buoys are in use. All are intended to serve as daymarks. Some buoys, particularly those at turning points in channels, are provided with lights of distinctive characteristics for location and identification at night. Some buoys are equipped with apparatus for providing distinctive sounds at intervals so they can be used as aids to navigation in fog and darkness. Some buoys are equipped with radio beacons, and some have reflectors to make them more conspicuous to radar.

Over the years, a number of different buoyage systems have been developed in various parts of the world. In an effort to reduce the differences, the International Association of Lighthouse Authorities (IALA) conducted an extensive study culminating in a recommended uniform system that has been adopted by most European, African, and Asian nations.

This system, identified as IALA system A, has one feature that has not been acceptable to most nations in the Western Hemisphere and some Asian countries. The objectionable feature is the use of red buoys to port while entering a channel from seaward. The United States and other nations were unwilling to give up their traditional "red right returning." To accommodate these nations, IALA proposed an optional system B, similar to system A but with the red buoys on the opposite side of the channel. The United States started installation of system B in 1983; with completion in 1989, the 6-year period coincides with the normal overhaul period of the buoys.

The principal features of IALA system B are shown in the **illustration**. All of these buoys are red and green. In addition, yellow buoys (with yellow lights

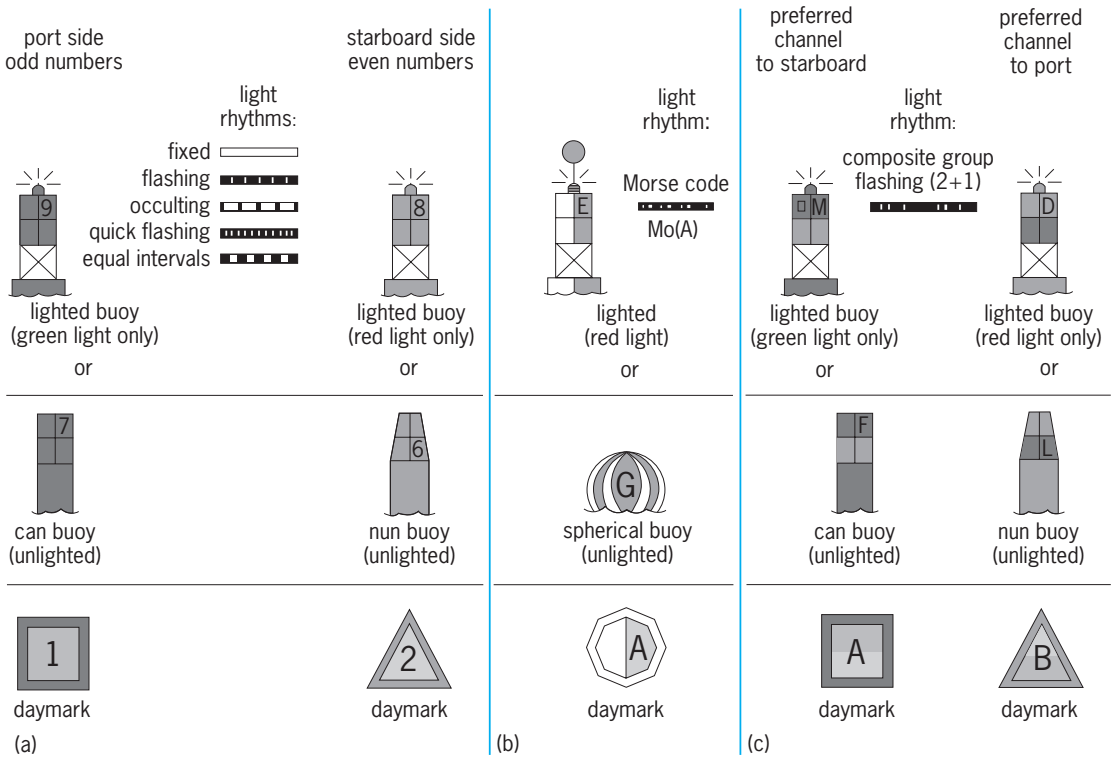

Various types of buoys in IALA system B. (*a*) Lateral aids, marking the sides of channels as seen when entering from seaward. These are marked with numbers. (*b*) Safe-water aids, marking midchannels and fairways. (*c*) Preferred-channel aids, marking bifurcations. Both safe-water and preferred-channel aids have no numbers but may be lettered. (*After U.S. Coast Guard, Modifications for a New Look in U.S. Aids to Navigation, ANSC 5N 3022, 1983*)

if such are used) are employed for ocean data acquisition systems, traffic separation schemes, dredging operations, fishnet areas, spoil grounds, military exercise zones, and anchorage areas. Orange and white buoys are used for regulatory and information markers. *See* PILOTING. Alton B. Moody

Bibliography. N. Bowditch, *American Practical Navigator, An Epitome of Navigation, 1995 edition*, 1995; B. Clearman, *International Marine Aids to Navigation*, 2d ed., 1988; U.S. Coast Guard, *Aids to Navigation Manual*, CG-222, 1982.

Buoyancy

The resultant vertical force exerted on a body by a static fluid in which it is submerged or floating. The buoyant force F_B acts vertically upward, in opposition to the gravitational force that causes it. Its magnitude is equal to the weight of fluid displaced, and its line of action is through the centroid of the displaced volume, which is known as the center of buoyancy. With V the displaced volume of fluid and γ the specific weight of fluid (weight per unit volume), the buoyant force equation becomes $F_B = \gamma V$. The magnitude of the buoyant force must also be given by the difference of vertical components of fluid force on the lower and upper sides of the body. *See* AEROSTATICS; HYDROSTATICS.

By weighing an object when it is suspended in two different fluids of known specific weight, the volume and weight of the solid may be determined. *See* ARCHIMEDES' PRINCIPLE.

A body floating on a static fluid has vertical stability. A small upward displacement decreases the volume of fluid displaced, hence decreasing the buoyant force and leaving an unbalanced force tending to return the body to its original position. Similarly, a small downward displacement results in a greater buoyant force, which causes an unbalanced upward force.

A body has rotational stability when a small angular displacement sets up a restoring couple that tends to return the body to its orignal position. When the center of gravity of the floating body is lower than its center of buoyancy, it will always have rotational stability. Many a floating body, such as a ship, has its center of gravity above its center of buoyancy. Whether such an object is rotationally stable depends upon the shape of the body. When it floats in equilibrium, its center of buoyancy and center of gravity are in the same vertical line. When the body is tipped, its center of buoyancy shifts to the new centroid of the displaced fluid and exerts its force vertically upward, intersecting the original line through the center of gravity and center of buoyancy at a point called the metacenter. A floating body is rotationally stable if the metacenter lies above the center of gravity. The distance of the metacenter above the center of gravity is the metacentric height and is a direct measure of the stability of the object. *See* SHIP DESIGN. Victor L. Streeter

Burgess Shale

Part of a clay and silt sequence that accumulated at the foot of a colossal "reef" during the Cambrian explosion, a dramatic evolutionary radiation of animals beginning about 545 million years ago. Although this explosion is most obvious from the geologically abrupt appearance of skeletons, the bulk of the radiation consisted of soft-bodied animal (**Fig. 1**). The Burgess Shale fauna, located near Field in southern British Columbia, is Middle Cambrian, approximately 520 million years old.

Diversity of fauna. The Burgess Shale fauna is remarkably diverse, with about 120 genera. Its approximate composition is arthropods 37%, sponges 15%, brachiopods 4%, priapulids 5%, annelids 5%, chordates and hemichordates 5%, echinoderms 5%, cnidarians and ctenophores 2%, mollusks 3%, and "other fauna" 19%. Although arthropods are the most important group, the trilobites, normally dominant among Cambrian arthropods, are entirely overshadowed both in number of species and in absolute number of specimens by a remarkable variety of other arthropods with delicate exoskeletons. The priapulids, which today are a more or less relict group of marine worms, also show a wide diversity of anatomical form, as do the polychaete annelids. Only one species of polychaete annelids has a close parallel among the Recent assemblages.

Highlights. The Burgess Shale has revealed many other aspects of the Cambrian explosion. First, a census of the collections reveals a marine ecology that is fundamentally unchanged to the present day. Predators, long thought to be insignificant in the Cambrian, are an important component. Second, groups with a minimal fossilization potential are preserved. One example is the gelatinous and delicate ctenophores, an important pelagic group in today's oceans but practically unknown as fossils. Third, although many of the species are a product of the Cambrian explosion, rare species are clear holdovers from the primitive Ediacaran faunas of late Precambrian age. Finally, some species are of particular evolutionary importance. Most significant is the worm *Pikaia*, which is interpreted as an early chordate, and as a predecessor of fish it lies near the beginning of the evolutionary path that ultimately leads to humans.

New discoveries. More than 35 other Burgess Shale–type faunas have been recognized, and Lower Cambrian examples from Chengjiang (South China) and Sirius Passet (North Greenland) are especially significant. The diversity of Chengjiang rivals that of the Burgess Shale, and the two faunas show many parallels. Significant discoveries include a relative of *Pikaia*, known as *Cathaymyrus* (**Fig. 2**), as well as primitive vertebrates. Despite paleogeographic proximity to the Burgess Shale, the Sirius Passet fauna is less similar, but it has yielded articulated halkieriids (**Fig. 3**).

Current controversies. A number of species appear so bizarre that earlier they were interpreted as new phyla, representing extinct body plans. They are, however, better interpreted as belonging to stem groups. This means that the fossils possess some, but not all, of the anatomical characters that define a particular phylum. These supposedly bizarre fossils potentially allow the investigation of the assembly of a given body plan. A good example of progress in this area concerns the arthropods. Once enigmatic taxa, such as *Hallucigenia* and *Microdictyon* from the Burgess Shale and Chengjiang respectively, are now identified as primitive lobopodians, related to *Aysheaia* (also from the Burgess Shale) and the living onychophores. Such lobopodians probably

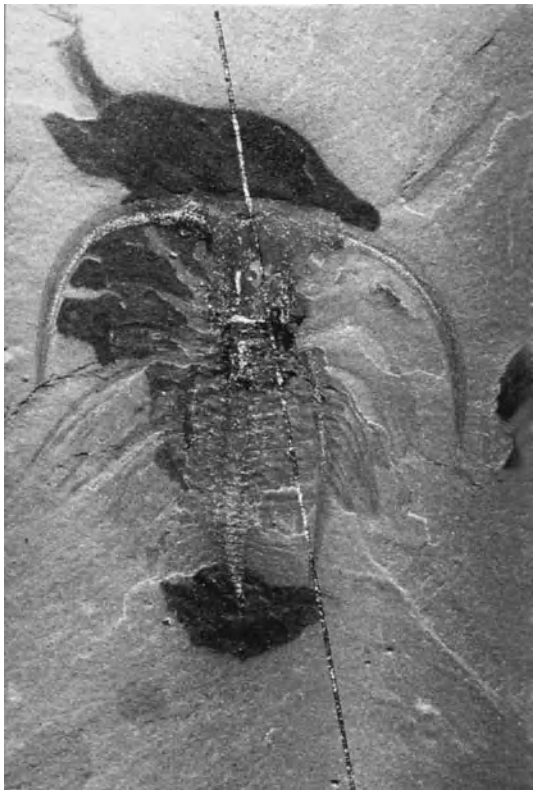

Fig. 1. ***Marrella splendens*, a characteristic Burgess Shale arthropod. The head shield bears two prominent pairs of spinose extensions; also visible are various appendages that include walking legs and gills. The prominent dark areas appear to represent body contents that oozed into the newly deposited sediment. This indicates that some decay occurred before an unknown factor intervened.**

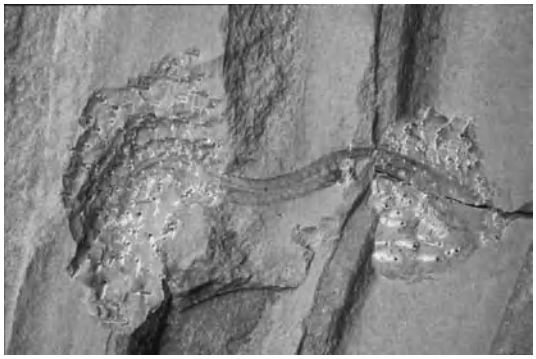

Fig. 2. ***Cathaymyrus diadexus*, a primitive chordate from the Chengjiang assemblage of China. The body is eellike, and at the anterior there are a pharynx and gill slits.**

Fig. 3. An articulated halkieriid, *Halkieria evangelista*, from the Sirius Passet fauna, North Greenland. The body was coated with platelike sclerites, and at either end carried a large shell.

descended from a wormlike ancestor. The succeeding stage is marked by the enigmatic anomalocaridids, which are now known to possess walking legs, either in the form of soft-bodied lobopods or sclerotized jointed appendages. From such animals it is suggested the more advanced arthropods evolved, which are exceptionally well represented in Burgess Shale–type faunas (Fig. 1). Puzzling combinations of anatomy within these various arthropods are being resolved by a series of new discoveries, as well as cladistic formulations that place these diverse faunas in a more ordered evolutionary context.

Another area of debate concerns the Burgess Shale as an exemplar of the principles of historical contingency. This view presupposes that Cambrian disparity—that is, the overall range of morphologies—was so much in excess of the present day that the whittling down was largely controlled by chance factors. Accordingly, it is argued any alternative history would have led to a radically different biosphere. Whether disparity in the Cambrian was so different from today is difficult to test. More importantly, the ubiquitous occurrence of biological convergence demonstrates that the options available to evolving life are constrained. Thus, historical contingencies are of only local relevance and unlikely to determine significantly the general history of life. *See* CAMBRIAN; FOSSIL; PALEONTOLOGY.

Simon Conway Morris

Bibliography. D. E. G. Briggs, D. H. Erwin, and F. J. Collier, *The Fossils of the Burgess Shale*, 1994; J.-Y. Chen, Y.-N. Cheng, and H. van Iten (eds.), The Cambrian explosion and the fossil record, *Bull. National Mus. Nat. Sci. Taiwan*, 101:1–318, 1997; S. Conway Morris, *The Crucible of Creation: The Burgess Shale and the Rise of Animals*, 1998.

Burn

The reaction that occurs when tissue (usually skin) receives more energy (heat, chemical energy, electrical energy, or radiation) than it can absorb without injury. Factors such as the duration of contact, temperature, volume of chemical, and voltage influence the severity of the injury. The notable pathologic changes are denaturation of protein and coagulation of blood vessels. Local destruction of tissue leads to bacterial invasion and infection; loss of fluid, electrolytes, and protein; loss of temperature control; and pain. Systemic reactions include altered blood flow and temperature regulation, fluid and electrolyte imbalances, shock, infections, and catabolism. Associated problems—other illnesses (particularly cardiopulmonary), trauma, and injuries caused by inhaling carbon monoxide, smoke, and occasionally heat—may be lethal or contribute significantly to mortality and morbidity from burn injuries. *See* SKIN.

Classification. Burns are classified according to location, percentage of total body surface involved, and depth of injury (see **illus.**).

First-degree burns, for example, sunburn, are superficial burns that result in some redness, pain, and swelling. Blister formation and necrosis (tissue death) do not extend beyond the epidermis (outer layer of skin). Healing is completed in a few days, and there is no scarring.

Second-degree burns, for example, scalds, are burns that show destruction of the entire epidermis and variable portions of the dermis (inner layer of skin); blistering and swelling occur and pain is severe. The degree of scarring depends on the depth of injury to the dermis. Skin grafting may be necessary if burns are deep or if trauma or secondary infection ensues and deepens the necrosis. *See* TRANSPLANTATION BIOLOGY.

Third-degree burns, for example, flame burns, are characterized by necrosis of the epidermis and dermis, including the dermal appendages (hair follicles and sweat glands). A pearly white, tan, or brownish-black eschar ("scab") with coagulated dermal blood vessels results. After 14 to 21 days, the eschar sloughs, leaving a raw surface. Skin grafting is usually necessary.

Fourth-degree burns, for example, severe electrical injuries, involve all layers of skin, as well as the underlying fat, muscle, nerves, or bone. These

injuries may require both amputation and skin grafting to heal.

Treatment. The initial step in management for a burn is to remove the injuring source from the burned patient. Attention should be directed to assessing the airway and breathing, particularly in individuals with inhalation or associated injuries. Patients with severe, extensive burns must be given large amounts of fluids and electrolytes intravenously, during the first 24 h and lesser amounts, with the addition of protein and other electrolytes, thereafter. Wound care involves removing dead tissue, preventing infection by applying antibiotics, and replacing the skin if indicated; the skin may be temporarily replaced by allografts (skin from other humans—for example, that obtained from cadavers), xenografts (skin from other species), or synthetic membranes or permanently replaced by autografts (skin from elsewhere on the patient's body). Nutritional support involves giving calories, protein, fat, and minerals by mouth and intravenously; requirements can be up to twice normal levels. Psychological support must overcome anxiety, denial, disbelief, guilt, depression, as well as fear of death, disfigurement, and disability, in order to provide coping mechanisms and enable planning for the future.

The more severe the burn or the more vulnerable the injured person (patients who are very young or very old, have other systemic diseases, or have other associated injuries are especially vulnerable), the more likely are complications. The most significant of these are systemic infections and the consequences of treating them. In addition, there may be specific or nonspecific injury to certain organs, such as Curling's ulcer (superficial ulceration of the stomach), renal damage related to inadequate fluid support, red cell or myoglobin destruction, adrenal changes, focal liver necrosis, or multiple organ failure.

Electrical burns. Electrical burns result from the amount of heat incident to the flow of a certain amount of electricity through the resistance offered by the tissues. The initial resistance is from dry skin. Large nerves and blood vessels carry electricity efficiently, but smaller blood vessels, muscle, and especially bone have higher resistance. Therefore, most electrothermal injuries consist of damage to the skin and immediately subjacent tissues, along with muscle necrosis due to the heat from bones and destruction of nutrient vessels. Deep penetration of the burn may follow large voltages. Small pitlike entrance and exit wounds may occur from arcing across adjacent flexed surfaces (for example, elbow, armpit, groin, and knee). Often the injury below a small entrance or exit wound is significantly larger and resembles tissue injured by crushing. Effects on the heart and brain are well documented and are evidenced in electric chair executions.

Chemical burns. Chemical burns result from corrosive agents that destroy tissues at the point of contact. Injury to the skin, eyes, and gastrointestinal tract are most common. The corrosives may be either acid or alkali; alkalis cause damage longer than acids. Extremes of pH above 11.5 or below 2.5 almost always result in irreversible damage. Protein coagulation, precipitation, and actual dissolution of tissue constituents may occur. The initial treatment is to dilute and thus remove the chemical by irrigating the area copiously with water; in addition, a few chemicals, for example, hydrogen fluoride, require specific neutralization. In general, chemicals should not be neutralized on the skin. In addition to causing burns, some chemicals injure specific organs; for example, phenol and carbon tetrachloride damage the liver.

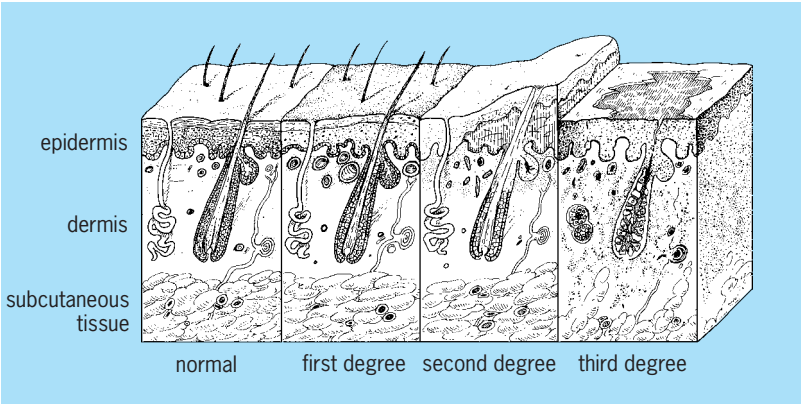

Anatomic diagram of normal and burned skin.

Epidemiology and prevention. Annually, 1% of the United States population is burned significantly enough to prevent participation in activities of daily living. Most burns occur in the kitchen, with scalds being the most frequent mechanism of injury; the next most frequent site of injury is the bathroom. Fire and flames cause most of the burn or inhalation injuries leading to hospital admissions and deaths. Most deaths from burns due to fire occur in the bedroom at night.

Burn prevention requires changes in behavior (conducting exit drills in the home and school, and not smoking in bed), in manufacturing (producing fire-retardant garments, fire-safe cigarettes, and lower-temperature thermostats on water heaters), in codes (requiring fire and smoke detectors, and sprinkling systems), and in laws. Thomas L. Wachtel

Bibliography. J. A. Boswick, *The Art and Science of Burn Care*, 1987; T. L. Wachtel, Burns, *Critical Care Clinics*, vol. 1, no. 1, 1985; T. L. Wachtel, V. Kahn, and H. A. Frank, *Current Topics in Burn Care*, 1983.

Burning velocity measurement

The velocity at which the combustible gases move through a flame in the direction perpendicular to the flame surface defines the burning velocity. The term flame as used here refers to the burning of combustible gases. The process of combustion, or burning, is described as a chemical reaction of two or more substances in which a large quantity of heat is evolved. This heat is self-absorbed and goes into

heating the products of combustion. Part of the light that is emitted is due to the heat of the flame, and part is due to the chemical reaction within the flame zones or thickness. It is the combustion of gases that forms the basis of a flame; and the zone in which chemical reactions occur, together with the concomitant temperature gradient, determines the location of the flame. An expression of the laminar flame speed for premixed homogeneous combustible gases has been obtained analytically, and indicates that the flame speed is proportional to the square root of the diffusivity and the reaction rate. The diffusion of heat and mass cause the flame to propagate. The reaction rate determines the temperature gradient by its effect on the thickness of the reaction zone: the slower the reaction, the thicker the zone.

A general definition of burning velocity consonant with measurement techniques cannot be formulated free from all possible objections. Ideally, if an infinite plane flame surface, considered stationary in a one-dimensional flow field, is cut perpendicularly by a stream tube, the area of the flame surface enclosed in the stream tube does not depend on how the flame is structured. The area of all parallel surfaces is the same no matter what property is chosen to define the surface, and this area is equal to the luminous surface of the flame. Where the flame surface is curved, the determination of the reference area within the flame, that is, the reaction zone, is much more difficult. For example, consider a combustible gas mixture which is supplied at a point at a steady rate and which flows radially outward in a laminar manner to a radially spherical flame. In this case the area of the surface of the combustion wave will depend on where the surface is located. Since the mass flow of the gases through the surface is constant, the flame speed will depend on the selected surface. In addition, since source flow with spherical symmetry exists, there is no constant asymptotic value for the velocity of the unburned gases. In this case, if it is assumed that the flame curvature is large compared with the flame thickness, then the burning velocity is the minimum value of the gas velocity upstream of the flame front.

Other difficulties are experimental and can arise because of the apparatus used to measure the flame speed. For example, probes cause the flow field to distort and can serve as heat sinks or catalytic surfaces for the reacting gases. In addition, where optical means are employed to measure temperature gradients in a flame, nonequilibrium effects can introduce errors in locating the reference surface for the determination of flame speed. Thus, measurement techniques that do not distort or interact with the chemistry of the flow field are preferred.

Idealized and actual flames. Experimental measurements of burning velocities have been made on both stationary and moving flames. Bunsen burners furnish a familiar example of stationary flames (**Fig. 1**). If the inner cone of a bunsen flame were a geometric cone whose base coincided with the port, the burning velocity S_u would be simply the volume of unburned gas entering per unit time per unit area of flame cone, as shown by Eq. (1). Expressed in terms

$$S_u = \frac{\text{area of port} \times \text{average gas velocity}}{\text{area of flame cone}} \tag{1}$$

of the port radius r, cone height h, angle α which one side of the cone makes with its axis, and average gas velocity U_u, burning velocity is given by Eq. (2).

$$S_u = \frac{rU_u}{\sqrt{r^2 + h^2}} = U_u \sin\alpha \tag{2}$$

In this idealized flame, all unburned gas must have the same velocity and must flow parallel to the axis of the burner; there must be no heating of the unburned gases; and the reaction zone must be infinitely thin. Actual flames deviate appreciably from this ideal. For example, consider the following methods of observation: shadowgraph, schlieren, and interferometry. Shadowgraph measures the derivative of the gas density gradient, and it is not well defined spatially. Schlieren measures the gas density gradient which corresponds more closely to the ignition temperature and is more readily defined and preferable. Interferometry measures the density, is very sensitive, and yields complicated graphical results which are difficult to define quantitatively. The flame cone defining the front of the combustion wave is not infinitely thin, yielding different locations for the flame surface and therefore different burning velocities. *See* FLAME.

Measurement techniques. The various methods used to measure flame speeds can be grouped under the following headings: conical stationary flames; flames in tubes; soap bubble methods; constant-volume explosions in spherical vessels; flat-flame burner methods; and laser diagnostic techniques. These are briefly described below.

Bunsen burner method. Premixed gases flow up a cylindrical tube long enough to ensure fully developed laminar flow. The gas burns at the mouth of the

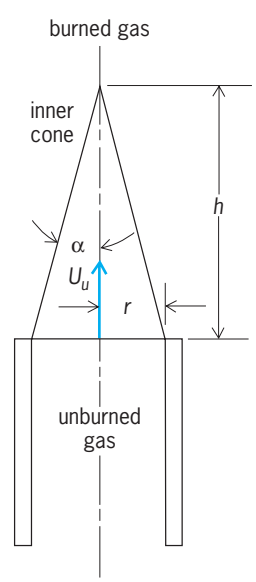

Fig. 1. Geometry of Bunsen flame.

tube, and the shape of the flame cone is recorded by photography or an equivalent method. The flame area is determined by various means of geometrical methods and the flame speed thus determined, as in Eq. (3). It is agreed that the schlieren cone is the

$$S_u = \frac{\text{volume of gas flow}}{\text{surface area of flame cone}} \tag{3}$$

most suitable for the determination of flame speed. The cone angle is measured at the central portion of the cone in about 30% of the internal portion of the cone. Wall effects which produce flame cooling, lower reaction rates, and lower flame speeds cannot be completely eliminated.

Cylindrical tube method. In this case, a premixed gas is placed in a horizontal or vertical tube. The gas is ignited at one end, and the rate of flame propagation into the unburned gases is measured. Provided that the tube is not too short, and the energy released by the spark used for ignition is not much larger than the minimum ignition energy, the combustion wave is observed to be nearly plane and to travel at almost constant speed between 2 and 10 tube diameters from the spark. This constant speed is the empirical burning velocity. This method is approximate and suffers from errors due to wall effects, buoyancy, combustion-wave distortion, and friction.

Soap bubble method. In an effort to eliminate wall effects, two spherical methods were developed. Soap bubbles blown with homogeneous combustible gas mixtures have been used as a containing medium. Spark ignition at the center of the bubble will propagate a combustion wave at constant pressure through the mixture. The flame speed is amplified by expansion of the burned gases behind the flame front.

Measurements on such a flame are made photographically by observing the growth of the flame radius through a slit as the film moves past the slit (**Fig. 2**). If an initial volume of gas of radius r_1 burns to a final volume of radius r_2 and is measured on a film moving with a velocity S_f and a magnification factor

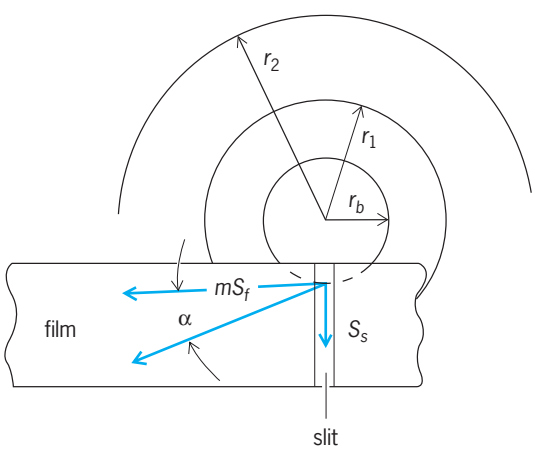

Fig. 2. Film moving past slit to record growth of flame by measurement of its radius.

m, the flame speed S_s in space is given by Eq. (4),

$$S_s = mS_f \tan\alpha \tag{4}$$

where α is the angle made with the film axis due to growth of the flame. The expansion ratio E is also a function of the burned gas density ρ_b and unburned gas density ρ_u, as expressed in Eq. (5). The mass of

$$E = \left(\frac{r_2}{r_1}\right)^3 = \frac{\rho_u}{\rho_b} \tag{5}$$

gas entering and leaving a unit area of combustion wave per unit time is derived from Eqs. (5) and (6) and shown in Eq. (7). Afterburning, which affects the

$$S_u\rho_u = S_s\rho_b \tag{6}$$

$$S_u = \frac{S_s\rho_b}{\rho_u} = \frac{S_s}{E} \tag{7}$$

precision of measurement of the final radius, can be eliminated by surrounding the bubble with inert gas. The method can be used at various pressures by placing the entire apparatus inside a pressure-controlled chamber.

Closed spherical bomb method. Another technique for measuring burning velocity in which adequate control of gas composition is maintained is combustion in spherical bombs with central ignition. Since reaction occurs under constant-volume conditions in this apparatus, the calculations to find the burning velocity are much more complex.

Flat-flame burner method. The flat-flame burner method is probably the most accurate because it offers the simplest flame front and one in which the area of the shadow schlieren and visible parts are the same. A combustible gas mixture flows up a tube in which a porous plug or grid is placed at the end. The grid stabilizes the flame. If the combustion gas speed is too great, a flame cone occurs; if too slow, flashback and quenching occur. Finally the tube is shrouded by inert gas to better define the flame. This method applies to mixtures having low burning velocities and can be applied as well to low pressures.

Laser diagnostic techniques. Through the combined use of laser Raman scattering and laser Doppler velocimetry, simultaneous nonintrusive, instantaneous, and pointwise measurements of temperature, specie concentration, and velocity can be obtained in flames. The nature of this diagnostic technique is also rather complicated. *See* LASER.

Maximum burning velocity in air mixtures of most hydrocarbons under atmospheric conditions varies from 10 to 40 in./s (25 to 100 cm/s). Typical gasoline hydrocarbons have burning velocities in the range of 12–16 in./s (30–40 cm/s). *See* COMBUSTION.

Vito Agosta

Bibliography. I. Glassman, *Combination*, 1977; S. Lederman, The use of laser Raman diagnostics in flow fields and combustion, *Progress of Energy and Combustion Sciences*, vol. 3, pp. 1–34, 1977; S. Lederman, A. Celentano, and J. Glaser, Temperature, concentration and velocity in jets, flames and

shock tubes, *Phys. Fluids*, 22(6):1065–1072, 1979; H. Taylor (ed.), *Physical Measurement in Gas Dynamics*, vol. 9, 1954; L. A. Vulis, *Thermal Regimes of Combustion*, 1961; F. A. Williams, *Combustion Theory*, 2d ed., 1985.

Burrowing animals

Some terrestrial and aquatic animals are capable of excavating holes in the ground for protection from adverse environmental conditions, as well as for storing food. Burrows vary from temporary structures of simple design (for example, the nesting burrows of some birds) to more permanent underground networks that may be inhabited for several generations (for example, rabbit warrens, badger sets, fox earths, and prairie dog burrows). They vary in structure from blind burrows with a single opening to extensive systems with several openings. Some animals (for example, some species of moles) live permanently underground, and their burrows have no obvious large openings to the surface. Burrows may be shared by a number of species, and abandoned burrows may be used by other species. Animals with limbs usually excavate their burrows by using their legs, but many burrowing animals are limbless and the mechanism of progression is not always obvious.

Terrestrial burrowers. Worms, slugs, many insects, and many vertebrates live in burrows. Earthworms are important soil organisms because their burrows improve drainage and aeration, their feeding habits enhance leaf decomposition, and their droppings increase soil fertility. Earthworms burrow by contracting circular muscles in their body wall to push forward, and contracting the longitudinal muscles to widen the burrow. Termites (Isoptera) and ants (Hymenoptera) are social insects, most of which live underground. Termites are major consumers of vegetation in warm climates, and many construct extensive underground galleries extending from the mound located on the surface. The conditions in the mound are carefully regulated by the termites. Most ants are predators or scavengers, but leaf-cutter ants feed on a fungus that grows on the harvested pieces of leaf in carefully tended underground galleries. Male mole crickets (*Gryllotalpa viniae*) attract mates by singing (they rub their wings together to generate sound) in their burrows; the burrow has a double, trumpet-shaped opening which acoustically focuses the sound, thereby increasing efficiency.

Some animals bury themselves to avoid adverse conditions, abandoning the burrow when conditions improve. Lungfish, for example, inhabit seasonal rivers in Australia, Africa, and South America. They bury themselves in the mud of drying rivers, breathe with the aid of an organ resembling a lung, and become active again when the next rainy season fills the river.

Several groups of burrowing vertebrates have small limbs or, in some cases, are limbless. For example, caecilians are amphibians that resemble earthworms. Burrowing skinks show varying limb reduction. Slow worms and snakes are burrowing, limbless reptiles, though many snakes live in other habitats.

Moles are highly adapted for a permanent life underground. European moles are solitary, feed mainly on earthworms, and have shovellike front limbs and well-developed tactile and olfactory senses, but their vision is reduced. Some moles are blind. The golden mole of Namibia effectively "swims," surprisingly quickly, through the desert sand by using its front feet and its head.

Most small and medium-sized mammals (for example, shrews, mice, voles, rabbits, prairie dogs, foxes, and badgers) live in burrows, though they have few anatomical adaptations to burrowing. Small animals are more susceptible to adverse environmental conditions than large ones. Very large mammals do not burrow.

Naked mole rats (*Heterocephalus glaber*) live in East Africa in underground colonies of up to 100 individuals; a colony may have 2.5 mi (4 km) of burrows. They feed on roots and tubers. They use their large incisor teeth to excavate the burrow and scrape away the soil to widen the opening. Uniquely among mammals, mole rats show a social order similar to that of social insects: each colony has only a single breeding female and several breeding males; all others are infertile and are called helpers. Foraging and burrowing are carried out mainly by small helpers, while defensive duties are performed by large helpers.

One problem all burrow dwellers face is ventilation. Diffusion alone is insufficient for extensive burrows. In order to overcome this difficulty, prairie-dogs (*Cynomys ludovicianus*) construct burrows up to 98 ft (30 m) long with two or three entrances and mounds of different shapes for each entrance. The force produced by wind creates a pressure differential between the entrances which causes air to flow through the burrow.

Aquatic burrowers. Many marine animals, including flatfish, crabs, and shrimps, take temporary refuge or live more permanently in sand or mud by burying themselves just below the surface. Aquatic sand and mud pose several problems for burrowing animals. First, the particles are usually tightly packed together, restricting movement and requiring these organisms to expend 10–1000 times as much energy to move a given distance compared to other forms of locomotion. Second, burrows readily collapse unless reinforced or consolidated. The wall of burrows may simply be consolidated with mucus, but some animals make a more permanent, substantial tube of particles stuck together with mucus. Finally, all burrowing animals must be able to create a current of water through the burrow so that they can breathe. Many animals also feed partly or wholly on particles carried in such currents.

Animals with limbs or similar appendages use them to burrow or move through sand. Crabs and shrimps use their legs. The burrowing crab (*Corystes cassivellaunus*) uses its unusually long antennae to breathe when it is buried. Interlocking setae hold the antennae together and form a tube, rather like

a snorkel. The burrowing sea urchin uses spatulate spines to burrow. Many burrowing worms and mollusks have no such appendages.

Saturated sediments have two properties that such animals exploit. When an intermittent force is applied to sediments, they become liquefied (thixotropic), and when a steady force is applied they become solid (dilatant). Many burrowing animals utilize these properties to burrow in a two-stage process that is clearly demonstrated in burrowing bivalved mollusks such as clams. The shell valves open against the sediment, forming a penetration anchor. Meanwhile, the foot repeatedly extends into the sediment. The siphons (breathing tubes) then close to prevent water leakage, and adductor muscles partly close the shell. These actions have several consequences: the penetration anchor is released; water is ejected from the shell margin, excavating a cavity in the sand; blood is forced from the body into the foot, which swells to form a terminal anchor; and the elastic hinge ligament of the shell is compressed. The muscles then pull the shell downward into the cavity. Finally, stored energy in the ligament is released so that the shell reopens. This cycle is repeated several times. *See* MOLLUSCA.

The alternation of penetration and terminal anchors can be recognized in many different kinds of burrowing animals. Burrowing worms and sea anemones use circular and longitudinal muscles in the body wall to elongate or thicken the body. The liquid contents of the body cavity act as a hydrostatic skeleton, allowing the two sets of muscle to antagonize.

Lugworms (*Arenicola*) are commonly found on sandy beaches, where each inhabits a U-shaped burrow. Coiled worm casts accumulate at the rear end of the burrow, and a small depression marks the front end. The worms pump water from back to front through the burrow to produce a current that provides for respiration and feeding. The sand in front of the head acts as a filter and is enriched with planktonic food before it is swallowed and digested. The indigestible sand is then deposited as casts on the surface at regular intervals.

Nereis diversicolor is a worm found in abundance in the mud of estuaries. It also inhabits a U-shaped burrow. It has a pair of jaws and can feed either as a predator or a scavenger or by filtering. It secretes a conical net of fine mucus strands in the front part of the burrow and pumps water through the net by undulating the body within the burrow. After about 7 min of pumping, the worm moves forward in the burrow, eating the net and the trapped food particles. It then secretes another net (the whole cycle takes about 10 min). A single worm may pump over half a quart of water in an hour.

Fiddler crabs (*Uca*) are common on many tropical and subtropical muddy shores such as mangrove swamps. They are air breathing and feed when the tide is out by sieving through surface mud for food. When the tide comes in, each crab retreats to its burrow and blocks the entrance with a plug of sand. Hugh D. Jones

Bibliography. J. Bailey, *The Life Cycle of a Crab*, 1990; M. L. Gorman and R. D. Stone, *The Natural History of Moles*, 1990; P. W. Sherman, J. U. M. Jarvis, and R. D. Alexander (eds.), *The Biology of the Naked Mole-Rat*, 1991.

Bursa

A simple sac or cavity with smooth walls and containing a clear, slightly sticky fluid interposed between two moving surfaces of the body to reduce friction. Subcutaneous bursae are found where the skin stretches around the greater curvature of a joint, as in the elbow or knee, and considerable chafing may occur; they may be single or multiple sacs. These bursae may enlarge as a result of continuous excessive irritation, as in housemaid's knee or miner's elbow. *See* BURSITIS; JOINT DISORDERS.

Synovial bursae are small closed sacs of fibrous tissue continuous with the joint cavity of a diarthrosis. They are lined with a complex membrane that secretes a clear lubricating fluid, serving to reduce friction between the opposing surfaces of the articulation.

Bursae may exist in the form of elongated sheaths surrounding tendons or ligaments, where these moving bands are in contact with another structure, such as a bone, muscle, or another tendon or ligament. Tendon sheaths are especially common where tendons bend around the ends of two bones at an articulation. *See* JOINT (ANATOMY); MUSCULAR SYSTEM; SKELETAL SYSTEM. Walter Bock

Bursitis

Any inflammation of a bursa. Bursae are synovial pouches, positioned to minimize friction between moving parts of the body. They are cystic in appearance and are filled with fluid. More than 80 bursae have been identified on each side of the body around joints and between tendons, muscles, and ligaments. In many cases, they are continuous with or near a regular joint cavity. It is generally thought that bursae exist as normal but irregularly apparent structures, but there is some evidence that they are pathologic alterations of connective tissue produced by irritation and friction. In any case, they rarely produce symptoms unless inflammation occurs. Bursitis most often occurs near the shoulder, hip, elbow, or knee. *See* BURSA; JOINT (ANATOMY).

Inflammatory changes in bursae produce acute or chronic swelling, an increase in the fluid content, and variable degrees of pain and tenderness. Acute bursitis may be septic (caused by microorganisms) or nonseptic. Nonseptic bursitis can be further subdivided into idiopathic (of unknown cause), traumatic, and crystal-induced bursitis. Septic bursitis may result from direct penetration by microorganisms through medical instrumentation or trauma; rarely, microorganisms may reach bursae through the blood. Most cases of bursitis are nonseptic; they may result from

trauma or physical stress. The fluid aspirated from an inflamed nonseptic bursa does not contain microorganisms.

In chronic bursitis, the wall of the bursa becomes thickened, shaggy, and irregular, with calcium deposits commonly being present. The fluid, which is normally clear, changes to a reddish-brown or black gritty mass as a result of repeated hemorrhages and precipitation of calcium. The treatment depends on whether the bursitis is septic or nonseptic. Septic bursitis is most commonly due to *Staphylococcus aureus* and requires prompt administration of appropriate antibiotics and repeated drainage of fluid containing pus. Nonseptic bursitis can be treated conservatively by withdrawal of fluid and administration of nonsteroidal, anti-inflammatory drugs. In nonseptic bursitis, withdrawal of the fluid can also be followed by injection of a steroid, but side effects, including chronic pain, may occur. Surgical incision and drainage are rarely necessary. Crystal-induced nonseptic bursitis is most frequently due to gout and usually responds well to drug therapy. Avoidance of trauma can help to prevent occupation-related cases of bursitis. Most cases of bursitis have a favorable prognosis. *See* CONNECTIVE TISSUE; GOUT.

Robert P. Searles

Bus

A motor vehicle for mass transit, built in various capacities and sizes, designed for carrying from 10 to 90 passengers or more on school, local, intercity, or interstate routes. In general, a bus has a long body with passengers sitting on benches or seats, while in local transit service additional passengers may be standing in the aisle. A school bus is used primarily to transport secondary school or younger students from home to school and back. Also called a coach or motor coach, the commercial bus for local transit and longer travel usually operates on a regular schedule and travels a fixed route, and each passenger pays a fare.

Motive power. Diesel engines are the principal source of motive power. Engine installation may be at the front, under the floor, at the rear (either laterally or transversely), or on the frame, according to capacity or operating specifications for the operator's unobstructed view of front entry and rear exit areas. *See* DIESEL ENGINE; INTERNAL COMBUSTION ENGINE.

In the United States, health concerns about exhaust emissions from diesel engines, especially nitrogen oxides and particulate matter, have prompted state and federal programs for retrofitting older buses with more effective emission controls, and replacing older buses with new and less polluting buses. Availability of ultralow-sulfur diesel fuel, having a sulfur content of less than 15 parts per million, allows installation of a diesel particulate-matter filter on new and in-use buses.

Some buses have a spark-ignition engine that is fueled with gasoline or natural gas, which is mostly methane. Compressed natural gas is the most common form, with far fewer buses using liquefied natural gas and propane. Natural gas is a cleaner-burning fuel that can reduce engine emissions. *See* LIQUEFIED NATURAL GAS (LNG); METHANE; PROPANE.

A battery electric bus obtains electricity from batteries stored onboard the vehicle. It is classed as a zero-emission vehicle because the bus produces zero emissions at the point of use. A hybrid electric vehicle has two onboard power sources—an internal combustion engine that provides energy conversion, and batteries or ultracapacitors that serve as energy storage devices for powering an electric motor on demand. Several transit systems have hybrid buses in service. Most are parallel hybrids in which either power source, or both, can be used to produce motive power to turn the wheels. With regenerative braking providing some recovery of the vehicle's kinetic energy, the engine can be downsized to reduce fuel consumption and exhaust emissions. Initial acceleration of a parallel hybrid bus is better than a conventional diesel-powered bus, which can outperform the hybrid bus on grades and at higher speeds. Emissions, fuel economy, and reliability of the hybrid bus are areas of continuing development.

Transit buses powered by a fuel cell have appeared in scheduled service and test programs. Some prototypes have onboard hydrogen storage, while others have an onboard reformer that converts liquid hydrocarbon fuel, such as gasoline or methanol, into hydrogen to feed the fuel cell. A liquid fuel avoids the need for expensive hydrogen-handling infrastructure at the refueling facilities. *See* FUEL CELL.

Design. Most transit buses have air brakes, air conditioning, air suspension, automatic transmission, public address system, and mobile phone for contact with the dispatcher. Location of the bus can be tracked by a global positioning system (GPS) that indicates to a control center in real time any deviation from the predetermined route. Buses for long-distance travel usually include restroom facilities. Electronic fare collection is common.

Basic bus design retains many truck components, including brake, steering, and suspension systems. Gross weight is kept at a minimum for economic operation and for maximum use of space for entry, aisle, and exit areas. Seating and safety devices must comply with federal regulations, which may require an onboard wheelchair lift or ramp. A truck-type ladder frame or integral body-and-frame construction performs the same functions as in other commercial vehicles for load-bearing capacity, rigidity, and impact resistance. *See* AUTOMOBILE; SATELLITE NAVIGATION SYSTEMS.

Body styles, seating, passenger capacity, power plant, and other equipment and passenger conveniences are available in great variety. Low boarding height and low floor height are typical of transit buses, while intercity and tour buses usually have a high floor over luggage compartments located between the axles. A double-deck bus has two separate passenger compartments, one above the other. An articulated bus is an extra-long bus (54 to 60 ft;

16.2 to 18.0 m) with two connected passenger compartments that bend at their connecting point as the bus turns. *See* MOTOR VEHICLE. Donald L. Anglin

Bibliography. M. Ehsani, *Modern Electric, Hybrid Electric, and Fuel Cell Vehicles: Fundamentals, Theory, and Design*, 2004; R. Hodkinson and J. Fenton, *Lightweight Electric/Hybrid Vehicle Design*, 2001; International Energy Agency, *Bus Systems for the Future: Achieving Sustainable Transport Worldwide*, 2002.

Bus-bar

An aluminum or copper conductor supported by insulators that interconnects the loads and the sources of electric power in an electric power system. A typical application is the interconnection of the incoming and outgoing transmission lines and transformers at an electrical substation. Bus-bars also interconnect the generator and the main transformers in a power plant. In an industrial plant such as an aluminum smelter, large bus-bars supply several tens of thousands of amperes to the electrolytic process. A single-line diagram of a typical substation bus-bar system is shown in **Fig. 1**. The bus-bar distributes the current of the two incoming lines between the outgoing line and the transformer. In this example, bus-bar failure interrupts the load connected to the transformer and the outgoing line. In a large station, the bus-bar failure may affect thousands of customers. The reliability of a system including bus-bars can be improved by using double bus-bars, where the loads automatically switch to the operating bus-bar in case of a fault. The ring bus-bar provides two supply routes for each load, which increases the system fault tolerance. A three-phase system has three bus-bars, which collectively are called a bus. *See* ELECTRIC POWER SUBSTATION.

Types. The major types are (1) rigid bus-bars, used at low, medium, and high voltage; (2) strain bus-bars, used mainly for high voltage; (3) insulated-phase bus-bars, used at medium voltage; and (4) sulfur hexafluoride (SF_6)–insulated bus-bars, used in medium- and high-voltage systems.

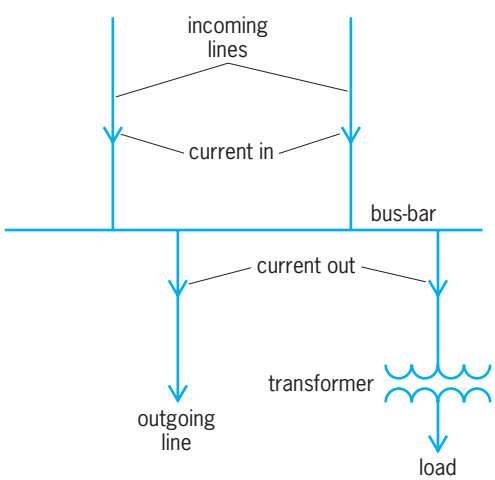

Fig. 1. Line diagram of a substation with a single bus-bar for distribution of current.

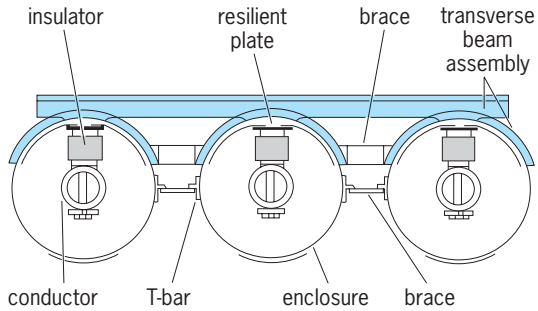

Fig. 2. Cross section of insulated-phase bus-bar.

The rigid bus-bar is an aluminum or copper bar, which is supported by porcelain insulators. The most frequently used bar types are rectangular, U-shape, and tube. In some cases, the bus-bars are protected by a metal duct, which surrounds all three phases.

The strain bus-bar is a flexible, stranded conductor which is strung between substation metal structures and held by suspension-type insulators. Its construction is similar to that of a common transmission line.

The insulated-phase bus-bar is a rigid bar supported by insulators and covered by a grounded metal shield (**Fig. 2**). The duct is often ventilated to assure better cooling of the conductor. The main advantage of this system is the elimination of short circuits between adjacent phases.

The sulfur hexafluoride–insulated bus-bar is a rigid aluminum tube, supported by insulators and installed in a larger metal tube, which is filled with high-pressure sulfur hexafluoride gas. The high-pressure sulfur hexafluoride has very high electric flashover strength, which permits the reduction of the bus-bar size at high voltage by a factor of 10 to 20. *See* ELECTRICAL INSULATION.

Design. Bus-bars are major components of electrical systems because they distribute large amounts of energy. A bus-bar failure is a major system disturbance which should be avoided, and bus-bars are therefore designed with extreme care. The major design considerations are that the bus-bar must carry the maximum load current continuously without overheating, and the mechanical forces generated by short circuits should not cause any damage. *See* CONDUCTOR (ELECTRICITY); WIRING. George G. Karady

Bushing

A removable metal lining, usually in the form of a bearing to carry a shaft. Generally a bushing is a small bearing in the form of a cylinder and is made of soft metal or graphite-filled sintered material. Bushings are also used as cylindrical liners for holes to preserve the dimensional requirements, such as in the guide bearings in jigs and fixtures for drilling holes in machine parts. *See* CYLINDER.

James J. Ryan

Butter

A food product made by churning cream. In the United States, butter is not defined by Standard of Identity in the Food, Drug and Cosmetic Act as are most foods, but instead it was defined in 1923 by an act of Congress. Butter must contain at least 80% milk fat and may contain added salt or color. In antiquity, butter and cheese were made as a means of preserving the surplus milk for later use. *See* MILK.

Butter is a water-in-fat emulsion. Cream is a fat-in-water emulsion. Cream consists of discrete fat globules, 6–16 micrometers in size, suspended in skim milk. Fat globules have a membrane or coating consisting of natural emulsifiers, lipoproteins, fat-soluble vitamins, cholesterol, and some other materials in lesser concentrations. The membrane provides stability for the globule and protects it from attack by lipase enzymes. *See* EMULSION.

Manufacturing. In the United States and Europe, negligible amounts of farm-separated cream are produced, and little butter is made by batch churning. Thus, the handling of cream and butter has advanced technologically to improve markedly the sanitary conditions under which it is produced and the quality of the finished food that reaches the consumer.

Cream that is to be used for making butter arrives at the manufacturer's processing plant from many sources, but the major source is those dairies where whole milk is separated into skim milk and cream by centrifugal force in a separator. Skim milk in these processing plants is converted to nonfat dry milk, and the cream is sold to other users. Another source of cream is fluid milk plants, where whole milk is partially or totally separated to produce standardized whole milk, 3.25% milk fat; low-fat milks, 1 and 2% milk fat; and skim milk, no more than 0.5% milk fat. The third source of cream is the cheese industry, where incoming milk supplies may be partially separated. Also, all cheese whey is separated to recover the milk fat that was lost in making the cheese. This is known as whey cream and is somewhat inferior to the sweet cream made directly from whole milk. These two sources of cream may be combined or stored separately. *See* CHEESE.

All cream is perishable and must be handled with care to retain high quality. After separation, cream is held in stainless steel tanks and refrigerated at 39–45°F (4–7°C). The fat content of creams ranges 35–45%.

Butter manufacturers buy and transport cream from many dairy plants. In the United States, some manufacturers buy cream in 20 or more states. Cream in large tankers is delivered to a butter manufacturer and is examined on arrival for flavor and odor, followed by taking of a sample for analysis. If the cream passes the flavor and odor test, it is pumped into the butter plant and is stored in another large tank. The next step is to pasteurize the cream. This heat treatment destroys bacteria, inactivates enzymes, and gives the cream a cooked or heated flavor. Pasteurization can either be a batch process where cream is heated to a minimum of 165°F (74°C) and held at the temperature for 30 min, or continuous at 185°F (85°C) where the hold is only 15 s. *See* PASTEURIZATION.

Following pasteurization, rapid cooling promotes fat crystals on the exterior and liquid fat on the interior of the fat globules. If the cream were churned after this step, the loss of fat to the buttermilk would be high. Thus, a tempering step is used in which the cream is held at about 50°F (10°C) to allow rearrangement of the fat crystals. Then, liquid fat is on the outside of the globules to allow rapid aggregation during churning.

In the summer season, consumption of green feed by dairy cattle elevates the oleic acid content of the milk fat. This phenomenon alters the melting point so that cream during tempering must be held 3.6°F (2°C) lower in summer to minimize fat loss to buttermilk and also to allow for suitable plasticity at wrapping and packaging.

Continuous churns produce as much as 15,000 lb (6800 kg) of butter per hour. They convert cream to butter in a few minutes; with the batch churn, about 45 min is needed to produce butter, and then at least 30 min to standardize the composition and get water dispersed in tiny droplets.

Churning. In the continuous churn the entering pasteurized and tempered cream is agitated vigorously by beater bars. This causes stripping of the fat globule membrane and aggregation of the fat globules into chunks 0.2–0.4 in. (0.5–1 cm) in diameter. At this point the emulsion has been inverted. The slope of the continuous churn allows the buttermilk to drain out the rear of the churn and the butter granules to continue through the churn barrel. The next flow-through position continues the kneading process to produce butter with finely dispersed droplets of moisture. If composition or color adjustment is required, it is done in this step; also, a salt solution is added to give the finished butter 1.2–1.5% salt. As the butter continues through the last step, more kneading is done with finer bars to complete the blending process and provide for fat crystallization that will yield optimum spreadability in the finished butter. A continuous ribbon of yellow butter streams from the end of the continuous churn. Butter drops into a hopper, where it is pumped to packaging machinery.

Packaging. Butter is delivered by stainless steel pipeline to a variety of packaging machinery. A paper parchment overwrap is used on $^1/_4$- and 1-lb prints (molded portions), then a waxed cardboard overwrap. Some product is delivered to machinery that makes individual patties and wrapped pieces. For a whipped product, nitrogen gas is injected into the butter; this product is sold in tubs of many sizes both retail and wholesale. In many plants, boxes fitted with plastic bags having a capacity of 68 lb (31 kg) are filled with butter either for government purchase from approved plants or for frozen storage.

Buttermilk. This by-product of butter manufacture has many uses. The composition of buttermilk is similar to that of skim milk except that it is higher in materials found in the fat globule membrane.

Buttermilk sold in supermarkets and labeled "cultured buttermilk" has been cultured with acid and flavor-producing bacteria and may be different in composition. Buttermilk from the churn is condensed and spray-dried; in this form it is in great demand by the ice cream and bakery industries.

Composition and grading. Butter in the package has a composition close to 80.0% milk fat, 1.2–1.5% salt, 17.5–17.8% water, and 1% milk solids. If butter is salt-free, the moisture and fat contents are adjusted to a slightly higher value to compensate. Salted butter has a shelf life of 3–6 months refrigerated; salt is a preservative and is about 7–8% of the water phase. Salt-free butter must be frozen to hold it for table use; otherwise, its shelf life is only 30 days. *See* COLD STORAGE.

In the United States, all butter in commercial channels is graded by licensed inspectors from the Department of Agriculture. This evaluation is organoleptic, using the senses of sight, taste, and odor, and requires highly trained technicians working in facilities that provide adequate room and excellent control of product temperature. Top-quality butter is graded as USDA-AA. If it is slightly off in texture or flavor, it receives USDA-A grade. Unacceptable product receives no grade or is rated below grade. Robert Bradley

Butter oil. This product is made by heating butter to break the emulsion and settling or centrifuging to separate the milk serum from the fat. It may also be prepared by deemulsifying cream. Moisture is reduced to a low level by drying, the butterfat content being over 99%. Although little butter oil is produced in the United States, it is of importance in Australia and New Zealand. It is usually canned, is much more stable than butter, and does not require refrigeration. Because it is practically anhydrous, spoilage by hydrolysis and microbiological action is eliminated. Butterfat consists principally of mixed glycerides of saturated and unsaturated fatty acids in approximately the following percentages: butyric, 3.5; caproic, 1.4; caprylic, 1.7; capric, 2.6; lauric, 4.5; myristic, 14.6; palmitic, 30.2; stearic, 10.5; longer-chain saturated, 1.6; decenoic, 0.3; dodecenoic, 0.2; tetradecenoic, 1.5; hexadecenoic, 5.7; octadecenoic (oleic), 18.7; octadecadienoic (linoleic), 2.1; and longer-chain unsaturated, 0.9.

Ghee. A common food fat in India, ghee is produced from boiled buffalo milk. Its manufacture is similar to that of butter oil. It can be kept for months, or years, without refrigeration, and has a more intense flavor than butter or butter oil. Frank G. Dollear

Microbiology. Seven major groups of microorganisms may be found in dairy products: bacteria, molds, yeasts, rickettsiae, viruses, algae, and protozoa. Only the first four are significant for butter.

Public health. The almost complete conversion from farm to factory manufacture of butter has all but eliminated public health hazards. All commercial butter is produced from pasteurized cream. When properly performed, pasteurization destroys all pathogens. Postpasteurization carelessness is, therefore, the only avenue for infection.

Pathogenic organisms which could, under some circumstances, enter and exist in butter include *Staphylococcus aureus* (which produces a toxin that survives pasteurization), some genera of Enterobacteriaceae (*Proteus* and *Salmonella*), *Clostridium*, *Corynebacterium diphtheriae*, *Brucella abortus* (undulant fever), *Mycobacterium tuberculosis*, and *Coxiella burneti* (Q fever).

Defects. Proper pasteurization destroys upward of 99.9% of all ordinary organisms present in milk or cream. Microbiologically induced flavors, developed prior to pasteurization, may carry over into butter. Attempts to minimize this carryover include the use of soluble food additives (illegal in the United States on the basis of the definition of butter by an Act of Congress in 1923) which are largely removed in the buttermilk, or by vacuum treatment following pasteurization. Neither method is completely effective.

Organisms responsible for flavor defects originate from dirty utensils, water, and air (indirectly from soil or plants). The name of an organism is often indicative of the flavor it produces. Among the organisms most responsible for flavor and other defects in butter are *Pseudomonas putrifaciens*; *P. nigrifaciens*; *P. fragi* and *P. fluorescens* (hydrolytic rancidity); acid-producing types; *Streptococcus lactis* var. *maltigenes*; such yeasts as *Saccharomyces*, *Candida mycoderma*, and *Torulopsis holmii*; and the molds *Geotrichum candidum*, *Penicillium*, *Alternaria*, and *Cladosporium*. Molds are responsible for musty flavors, but they, as well as some yeasts and *Pseudomonas nigrifaciens*, may also cause color defects.

All of the *Pseudomonas* groups found in butter are psychrotrophs; that is, they grow well at refrigerator temperatures of about 42°F (5.5°C). They may cause putrid or lipolytic flavors in 5–10 days.

Factors responsible for inhibiting the development of microbiological flavor defects in butter include common salt, pH, butter structure, and storage temperatures.

Butter with 1.5% salt and 17.5% water has an overall brine concentration of 8.6%, which is inhibitory to some organisms. The growth of psychrotrophs is inhibited by a lower pH. In properly worked (mixed) butter, fat is the continuous phase and surrounds the tiny moisture droplets. Unless the water used in the manufacture of butter is contaminated, the water-droplets are overwhelmingly sterile. Also, their size is such that growth of bacteria is impossible. Only molds whose mycelia can penetrate the fat phase can reproduce. Where the interval between manufacture and sale is less than 4–6 weeks, the butter is maintained at 35–55°F (1.7–12.8°C); hence some growth of organisms may be expected to continue, but where storage is from a month to a year or more, temperatures range from 0 to −15°F (−17.8 to −26.1°C) and growth ceases. The number of living organisms may actually decrease.

Palatability development. Not all nationalities prefer butter of the same flavor intensity. Butter made from sweet pasteurized cream is bland but may

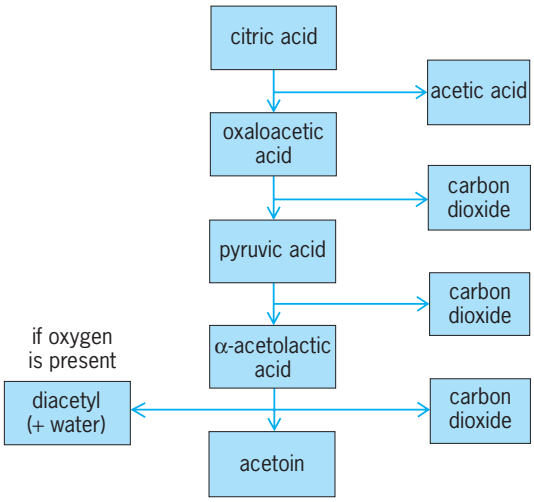

Decomposition of citric acid to diacetyl, chief component responsible for cultured butter aroma.

have a slight nutty or scalded-milk flavor. In general, Americans, Australians, and New Zealanders prefer this flavor. Europeans, Latin Americans, and some Asiatics prefer a more intense flavor. The desired flavor can be developed by the use of milk cultures of certain organisms. These cultures are referred to as starters. They may be added to the pasteurized and cooled cream at temperatures of 72°F (22.2°C) or less, or to the butter at about the time of salting. The latter practice is more economical and facilitates flavor control. The presence of citrates or citric acid is necessary for the development of diacetyl, $CH_3COCOCH_3$, recognized as the chief compound responsible for butter aroma (see **illus**.). Most unsalted butter is made by the use of starters, because they tend to inhibit the growth of psychrotrophs.

The organisms present in the starter determine how the flavor compounds develop, and their quantity and intensity. For use in buttermaking, a mixture of two or more of the following organisms is desirable: *Streptococcus lactis*, *Leuconostoc cremoris*, and *S. diacetilactis*; in recent years the last has been used alone or in combination with *S. lactis*. In some areas of Europe, *Candida krusii* (a yeast) has been tried in mixed cultures.

Starter distillates, made by steam distillation of starter cultures, are commercially available. Their advantages are economy, convenience, and uniformity. Their disadvantages are lack of delicate flavor and aroma and failure to inhibit the growth of some undesirable bacteria. Addition of lactic or other food-grade acids or of synthetic flavoring compounds could compensate for the latter disadvantage; however, such additions may not comply with the legal definition of butter.

Control. Standards and procedures for controlling the quality of butter have been developed collectively by the butter industry, dairy schools and experiment stations, the U.S. Department of Agriculture, and state regulatory agencies. By education, inspection, and the use of various tests standardized by the American Public Health Association for the maintenance of quality, butter has been greatly improved since the early 1940s.

Procedures for ascertaining the numbers of psychrotrophs (as represented by *Pseudomonas putrifaciens*) and for yeasts and molds are especially important because they are, to a great extent, an indication of postpasteurization handling practices.

The introduction of stainless steel churns and the use of metal in all other equipment have eliminated many flavor problems—especially those due to yeasts and molds. The practice of all-welded pipelines and cleaning in place, when combined with the proper use of cleaners and sanitizers, likewise has been an important factor in the production of high-quality butter. *See* FOOD MANUFACTURING; FOOD MICROBIOLOGY; INDUSTRIAL MICROBIOLOGY; MICROBIOLOGY. L. C. Thomsen

Bibliography. F. W. Bodyfelt, J. Tobias, and G. M. Trout, *The Sensory Evaluation of Dairy Products*, 1988; E. M. Foster et al., *Dairy Microbiology*, 1983; Y. H. Hui (ed.), *Dairy Science and Technology Handbook*, 3 vols., 1993; G. R. Richardson, *Standard Methods for the Examination of Dairy Products*, American Public Health Association, 16th ed., 1992; D. Swern (ed.), *Bailey's Industrial Oil and Fat Products*, vol. 1, 4th ed., 1986; N. P. Wong, *Fundamentals of Dairy Chemistry*, 3d ed., 1988.

Buxbaumiidae

A subclass of the class Bryopsida, the true mosses. It consists of a single family with four genera. The peristome shows some similarities to the double-peristome members of the subclass Bryidae, but the numerous series of teeth external to the endostomial cone are distinctive. *See* BRYIDAE.

The Buxbaumiidae is very distinctive in every way and most significantly in the structure of the peristome. The plants are small and occur especially on soil. The gametophyte is greatly reduced (*Buxbaumia*), with no stem and with leaves few

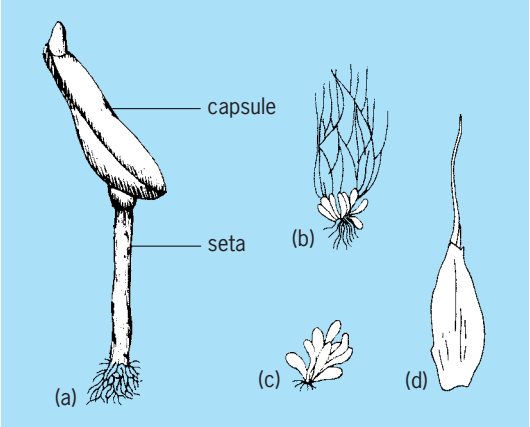

Morphological features of Buxbaumiidae. (*a*) Sporophyte of *Buxbaumia aphylla* (*after W. H. Welch, Mosses of Indiana, Indiana Department of Conservation, 1957*). *Diphyscium foliosum* (*b*) habit sketch, (*c*) leaves, and (*d*) perichaetial leaf (*after H. S. Conard, How to Know the Mosses and Liverworts, William C. Brown, 1956*).

and readily disappearing, or is better developed with well-formed leaves having a single costa and short cells. The sporophytes are terminal; the perichaetial leaves, if present, are conspicuous and exceed the capsules (see **illus.**). The capsules are disproportionately large and immersed or elevated on a seta (*Buxbaumia*). They are strongly inclined and asymmetric, tapered to a small mouth from a broad base. The operculum is small and conic. The exostome teeth of the peristome are greatly reduced in one to four series (representing several concentric series of amphithecial cells and consisting of thickenings deposited on the adjoining tangential walls between the series); the endostome consists of a tall, white, plicate cone. The calyptrae are small, conic, and naked. Chromosome numbers are 8 and 9. *See* BRYOPHYTA; BRYOPSIDA.

Howard Crum

Bibliography. M. R. Crosby, *Florschuetzia,* a new genus of Buxbaumiaceae (Musci) from southern Chile, *Bryologist*, 80:149–152, 1977; A. Engler and K. Prantl, *Die natürlichen Pflanzenfamilien,* 2d ed., vol. 2, 1925.

Bytownite

A member of the plagioclase feldspar solid-solution series with a composition ranging from $Ab_{30}An_{70}$ to $Ab_{10}An_{90}$ ($Ab = NaAlSi_3O_8$ and $An = CaAl_2Si_2O_8$) [see **illus.**]. In the high-temperature form, bytownite has the high-albite structure type in which Al and Si atoms are completely disordered; that is, neither element shows a preference for any of the four distinct tetrahedral sites. During cooling, natural material changes its Al/Si distribution, tending toward an ordered state in which Al and Si concentrate in different sites which alternate with one another in all directions. Bytownite is very abundant in basic igneous rocks, where it is the first plagioclase to crystallize under plutonic conditions; it forms the cores of zoned plagioclase phenocrysts in basaltic volcanics; and it sometimes occurs in anorthosites. *See* ANORTHITE; FELDSPAR; IGNEOUS ROCKS.

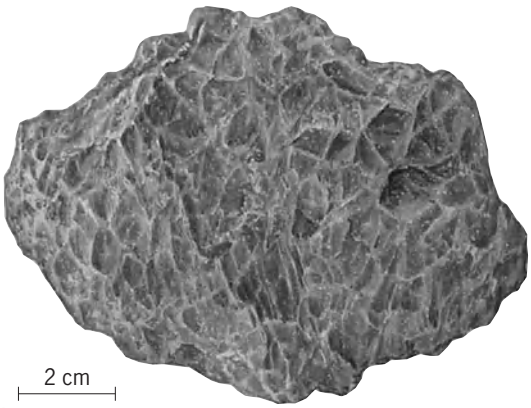

Bytownite, from Crystal Bay, Minnesota. (*Specimen from Department of Geology, Bryn Mawr College*)

Lawrence Grossman

C

C–H activation — Charophyceae

C—H activation

The cleavage of carbon-hydrogen bonds in organic compounds, leading to subsequent functionalization to introduce useful chemical groups. For example, petroleum and natural gas, important energy resources for the modern world, are both alkanes, C_nH_{2n+2}, with many C—H bonds. One problem is to convert these alkanes to useful alcohols, $C_nH_{2n+1}OH$. This is particularly difficult because of the lack of reactivity of alkanes as indicated by their older name, paraffins (from Latin, meaning low affinity). In C—H activation, metal catalysts are often used to mediate the hydrocarbon conversion reactions. *See* ALCOHOL; ALKANE; CHEMICAL BONDING; ORGANIC CHEMISTRY.

Radical reactions. The earliest C—H activation reactions, dating from the work of H. J. H. Fenton in the late nineteenth century, used iron salts as the catalyst and hydrogen peroxide as the oxidant. These are radical-based processes involving hydrogen atom abstraction from the hydrocarbon and subsequent reaction of the radical intermediates with oxygen from the air. The role of the metal is to catalytically decompose the peroxide to provide a continuous supply of reactive radicals, such as the hydroxyl radical (OH), capable of abstracting an H atom from the hydrocarbon. The reactions tend to be unselective and very inefficient. Improved Fenton chemistry has been developed in recent years, but the selectivity is still poor unless there is only one type of C—H bond that can be attacked. Such is the case for *p*-dimethylbenzene, where the CH_3 groups have particularly weak C—H bonds and are preferentially attacked. Commercial processes use Fenton-like conditions that lead to terephthalic acid, a component of polyester materials [reaction (1)]. *See* POLYESTER RESINS.

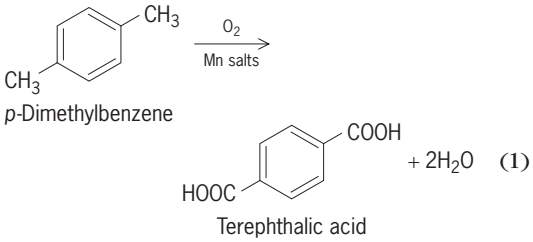

(1)

Superacid reactions. Another practically important reaction involves the rearrangement of the linear alkanes, often found in petroleum, to the more desirable branched chain isomers used in motor fuel. G. A. Olah (Nobel prize, 1994) found that this happens with very strong acids (proton sources) that abstract H^- from the hydrocarbon. These acids (superacids) can be formed by combining two acidic components. For example, HF and SbF_5 react to give the superacid $HSbF_6$, which is capable of protonating alkanes. Only a few such reactions are possible, at least on a commercial scale. These include cracking and reforming that are carried out at high temperature and are therefore unselective. *See* CRACKING; PETROLEUM PROCESSING AND REFINING; SUPERACID.

Methane-to-methanol conversion. Only certain reactions are possible using radical and superacid pathways. Additional methods are needed such as for converting natural gas, largely methane (CH_4), to methanol (CH_3OH). If realized, this would simplify the transport of natural gas from remote locations, methanol being an easily transported liquid.

The first example of this type of process was Shilov chemistry dating from 1970–1980 in the Soviet Union. In this work, Pt^{2+} salts break alkane C—H bonds under relatively mild conditions (hot aqueous solution). At first, isotopic substitution was seen (C—H → C—D), but with addition of an oxidizing agent, the desired conversion of C—H to C—OH was indeed observed. Unfortunately, the only efficient oxidant found was a very expensive platinic salt, and the rates of reaction were very slow.

R. A. Periana made the next advance by showing how a regenerable oxidant, sulfuric acid, could be

substituted for the platinic salt, and a much more efficient platinum-based catalyst (see structure) could

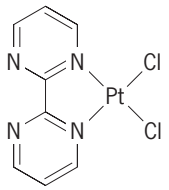

be substituted for the platinous salt of the original Shilov system. In this work, methane was converted to methyl bisulfate, CH_3OSO_3H, an ester from which methanol can be obtained by hydrolysis. Although a great improvement over the original system, the Periana system was still not commercially economic. In a later development, Periana saw the coupling of two methane molecules to give acetic acid, CH_3COOH, where both carbons of the product came from methane.

Alkane functionalization. A variety of ML_n fragments, consisting of a metal M and its associated ligands L_n, are capable of reaction with an alkane to give C—H bond breaking as shown in the first step of reaction (2). In principle, it should be possible to

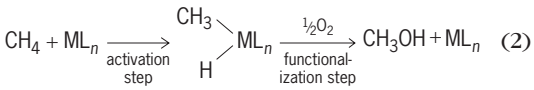

(2)

follow this with a second step to give the functionalized alkane, but this has proved difficult in practice. R. G. Bergman showed as early as 1982 that $(C_5Me_5)Ir(PMe_3)$, formed photochemically, is suitable for the first reaction, for example. *See* COORDINATION CHEMISTRY; LIGAND.

Other cases are known where the first step of reaction (2) is followed by a functionalization reaction. H. Felkin, R. H. Crabtree, and A. S. Goldman showed reactions of this kind where the intermediate alkyl hydride decomposes to give alkene and free H_2 or, in the presence of a second alkene as sacrificial oxidant, alkene and hydrogenated sacrificial oxidant [reaction (3)].

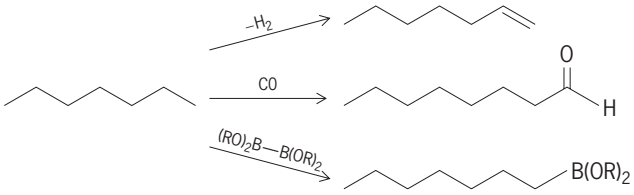

(3)

Y. Saito and M. Tanaka showed that the intermediate alkyl hydride can be trapped by carbon monoxide (CO) to give aldehyde (RCHO) as final product, and J. F. Hartwig trapped the alkyl with diborane derivatives to give alkyl boronic esters [reaction (3)]. In all of these cases, C—H activation preferentially occurs at the least hindered C—H bond, leading to products that are quite different from those formed in radical and acid pathways where the most substituted and most hindered C—H bonds are most reactive. In each case, appropriate transition-metal compounds are present and catalyze the reactions.

These do not yet form the basis of any practical process, but produce terminally functionalized products (the most desirable type) rather than the mixtures commonly found in other reactions.

Enzymes. Enzymes with iron-based reactive centers are known that catalyze alkane conversions. For example, methane monooxygenase, a bacterial enzyme, catalyzes the conversion of methane to methanol, and cytochrome P450 enzymes oxidize C—H to C—O—H in a wide variety of substrates. The mechanisms involved are not fully understood. *See* BIOINORGANIC CHEMISTRY; ENZYME. Robert H. Crabtree

Bibliography. R. H. Crabtree, Alkane C—H activation and functionalization with homogeneous transition metal catalysts: A century of progress—a new millennium in prospect, *J. Chem. Soc. Dalton*, 17:2437–2450, 2001; W. D. Jones, Isotope effects in C—H bond activation reactions by transition metals, *Acc. Chem. Res.*, 36:140–146, 2003; B. Meunier, *Biomimetic Oxidations Catalyzed by Transition Metal Complexes*, ICP, London, 2000; M. Newcomb, P. F. Hollenberg, and M. J. Coon, Multiple mechanisms and multiple oxidants in P450-catalyzed hydroxylations, *Arch. Biochem. Biophys.*, 409:72–79, 2003; R. A. Periana et al., Catalytic, oxidative condensation of CH_4 to CH_3COOH in one step via CH activation, *Science*, 301:814–818, 2003; A. E. Shilov and G. B. Shul'pin, *Activation and Catalytic Reactions of Saturated Hydrocarbons in the Presence of Transition Metal Complexes*, Kluwer, Dordrecht, 2000.

Cabbage

A hardy, cool-season crucifer (*Brassica oleracea* var. *capitata*) of Mediterranean origin and belonging to the plant order Capparales. Cabbage is grown for its head of overlapping leaves (see **illus.**), which are generally eaten raw in salads, cooked fresh, or processed into sauerkraut. Because it normally produces seed the second year, cabbage is considered to be a biennial by most authorities. Others regard it as a perennial because it will remain vegetative unless subjected to cold weather. *See* CAPPARALES.

Chinese cabbage is a related annual of Asiatic origin. Two species are grown in the United States, petsai (*B. pekinensis*) and pakchoi (*B. chinensis*).

Propagation of cabbage is by seed planted in the field, or in greenhouses and outdoor beds for the production of transplants. Field spacing varies; plants are usually 8–18 in. (20–46 cm) apart in 30–36 in. (76–91 cm) rows.

Cool, moist climate favors maximum yields of firm heads. However, exposure of young plants to prolonged low temperatures favors seed-stalk formation without production of normal heads.

Varieties. Cabbage varieties (cultivars) are generally classified according to season of maturity, leaf surface (smooth, savoyed, or wrinkled), head shape

Cabbage (*Brassica oleracea* var. *capitata*). (*Joseph Harris Co., Inc., Rochester, New York*)

(flattened, round, or pointed), and color (green or red). Round, smooth-leaved, green heads are commonest. Popular varieties are Golden Acre and strains of Danish Ballhead; hybrid varieties are increasingly planted. Principal varieties of other types are Chieftain (savoy), Jersey Wakefield (pointed), Red Danish (red), and Wong Bok (Chinese cabbage). Strains of Danish Ballhead are also used for sauerkraut manufacture and for late storing. Varieties differ in their resistance to disease and in the tendency for heads to crack or split in the field.

Harvesting. Harvesting of fresh market varieties begins when the heads are hard enough to be accepted on the market, generally 60–90 days after field planting. Buyer preference for heads weighing no more than 2–3 lb (0.9–1.4 kg) has encouraged close spacing and early harvesting. In the northern states production of cabbage for winter storage has declined as the Texas and Florida winter areas have expanded production, providing green cabbage as competition on the market for the white stored heads.

Texas and Florida are important winter crop producing states; Georgia, Mississippi, and North Carolina produce large acreages in the spring; and New York, North Carolina, and Wisconsin are important for the summer and fall crops. New York and Wisconsin are the important kraut cabbage states.

H. John Carew

Diseases. About 40 diseases of cabbage and other cole crops share the worldwide distribution of these important vegetables. The most important diseases are clubroot (caused by *Plasmodiophora brassicae*), blackleg (*Phoma lingam*), yellows (*Fusarium oxysporum* f. sp. *conglutinans*), black rot (*Xanthomonas campestris*), downy mildew (*Peronospora parasitica*), and Alternaria spot (*Alternaria* sp.).

Clubroot. The clubroot organism persists in infested soils for several years. When infected, young roots are enlarged and produce masses of abnormal root tissues instead of a normal root system. Proper absorption of water and minerals from the soil is prevented, and infected plants are stunted and often wilt during the day. Moist, acid soils favor the disease. Amendments such as lime to raise the soil pH can reduce disease damage. In some countries, fungicides can control the disease. Some cultivars that are resistant to clubroot are available.

Yellows. The fungus that causes yellows can also persist in the soil for many years. The fungus infects young roots and invades the entire vascular system, causing the leaves to wilt and drop off. Xylem vessels turn brown, and the plants are severely stunted or killed. Good heads are seldom produced. Cabbage cultivars resistant to yellows are available. The letters "YR" following the cultivar name signify that it is resistant to yellows. Yellows is seldom a problem in cole crops other than cabbage.

Blackleg and black rot. Blackleg and black rot are serious diseases in seedbeds and in the field. Both pathogens may be seed-borne and can spread rapidly from plant to plant by spattering or wind-blown rain. Thus, a very small amount of disease may soon develop throughout the entire crop. Neither organism persists in the soil more than a few years. Control is achieved by crop rotation and by using pathogen-free seed. Chemical and hot water treatments can eliminate these pathogens from infested seed.

Blackleg produces grayish-brown dead areas on leaves and stems and often extends to the roots. Plants may be killed or stunted. Black rot bacteria enter leaves through hydathodes at the leaf margin and invade the plant systemically. Yellow, triangular spots appear at the leaf margin, and the vascular system turns black. Leaves are killed, and affected heads are rotted.

Downy mildew. Downy mildew is destructive on young plants and mature heads. Small, dark flecks are produced in abundance on the undersides of cotyledons and leaves and on the outer head leaves. Young seedlings are frequently killed. Fungal structures and spores may be seen as downy-white masses on the undersides of affected leaves early in the day while dew is present.

Alternaria. Alternaria leaf spot is most destructive on cabbage approaching maturity. Dark spots up to 1 in. (2.5 cm) in diameter, with concentric rings, are formed on leaves and heads. Alternaria and downy mildew can be controlled with fungicides.

Other diseases. Virus diseases sometimes affect cabbage. Symptoms include mottling, ringlike dead spots, black flecking, and stunting. Many cultivars are resistant to common viruses. Black speck, a physiological disorder, produces abundant black specks throughout the head leaves, especially after a period of storage at low temperature. *See* CAULIFLOWER; PLANT PATHOLOGY; PLANT VIRUSES AND VIROIDS.

J. O. Strandberg

Bibliography. L. V. Boone et al., *Producing Farm Crops*, 4th ed., 1991; R. E. Nyvall, *Field Crop Diseases Handbook*, 1989.

Cable television system

A system that receives and processes television signals, as well as other digital forms of communications such as those related to high-speed Internet connectivity and cable telephony, from various sources and retransmits these signals through coaxial or optical-fiber (or fiber-optic) cables to subscribers' homes. The sources of the signals include broadcast transmissions, satellite-delivered programming, and local television studio productions. The facility that receives, processes, and retransmits the signals is called a headend.

Types of service. Cable television service is provided to subscribers by cable system operators that are awarded contracts or franchises by local or state governments to provide cable service to local communities. Cable television systems typically offer the customers in their franchise area four types of service options: basic service, expanded basic service, premium networks, and pay-per-view service. Modern cable systems provide a wealth of specialized programming, video-on-demand, subscription services, broadband high-speed Internet connectivity, cable telephony or voice over Internet Protocol (VoIP), and specialized interactive digital services. In exchange for being allowed to develop franchises, cable systems pay franchise fees and often set aside channels for public, educational, and governmental use, as well as providing studio and production facilities that enable community members, schools, and governments to produce and televise programs over these channels. *See* INTERNET; VOICE OVER IP.

Frequency spectrum. Compared to almost any other communications need, video requires extensive bandwidth. While telephone-quality voice transmissions require only 3 kHz of spectrum and high-fidelity sound requires about 20 kHz (double for stereo), the current United States analog video standard requires a bandwidth of 4.2 MHz. Uncompressed high-definition television requires about 30 MHz for each of the components of the color picture. Over-the-air broadcast high-definition television is limited to the same 6-MHz channel as the older analog standard, requiring complex compression algorithms. In order to carry numerous channels of programming, a separate enclosed spectrum is needed because the terrestrial broadcast frequency spectrum is too limited in bandwidth. *See* BANDWIDTH REQUIREMENTS (COMMUNICATIONS); DATA COMPRESSION.

Unlike broadcast television signals, which travel through free space, cable signals travel through coaxial cable or optical fiber, with different programs or channels traveling at different frequencies (much the same as frequency-division multiplex). In effect, the coaxial cable or optical fiber acts as a self-contained, closed, noninterfering frequency spectrum, created inside the cable by the reuse of the spectrum already in use for other purposes. Since cable television services are in an enclosed system and are not broadcast over the air, they can provide bandwidth enhancement techniques unavailable to other media providers. *See* MULTIPLEXING AND MULTIPLE ACCESS.

High-definition television (HDTV) and digital television (DTV) are two different types of service and should not be confused. High-definition television provides a sharper, larger picture using digital techniques, protocols, and algorithms. While digital television uses digital modulation techniques, it may not necessarily be of greater definition than standard analog television. Non-high-definition digital television may have several standard television programs compressed into one digital stream to be decoded at the subscriber node or set-top box. HDTV can require a greater bandwidth, depending upon compression ratios, than standard television. Likewise, DTV may also require greater bandwidth than standard television. However, DTV may have several video programs in the same spectral space, thus providing spectrum efficiency.

Transmission technology. Cable television is made possible by the technology of coaxial and optical-fiber cable and is subject to the principles of transmission-line theory. The primary disadvantage of coaxial-cable distribution systems is their relatively high loss or attenuation to television signals at the frequencies normally used in cable television systems (50–550 MHz, extending up to 1 GHz in newer systems). Amplifiers are required to overcome this signal loss, and the farther the subscriber is from the cable headend the more amplifiers are needed (**illus.** *a*). Noise and intermodulation distortions created by many cascaded amplifiers limit the practical length of any coaxial cable network. *See* COAXIAL CABLE.

Optical fiber does not have the same high attenuation or loss characteristics as coaxial cable, and optical-fiber networks can therefore be built without amplifiers (illus. *b*), or with few amplifiers for very large distribution areas. Most cable systems that use optical fibers do so in a hybrid fashion. Optical fiber is connected from the headend to some localized node or terminating location. The subscriber is then connected to the optical-fiber node by short distances of coaxial cable. The short coaxial cable subscriber connection in most cases eliminates the need for amplification, except in rare instances where one or at most two amplifiers are required. *See* COMMUNICATIONS CABLE; OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

System architecture. The system design or architecture is known as a tree-and-branch design. The tree-and-branch architecture is the most efficient way to transmit a package of multiple channels of programming from a headend to all subscribers.

There is one major difference between a switched telephone network and a cable television system. Cable systems are usually built on a non-switched basis, meaning that every subscriber receives the same channels unless physically restricted by the cable television operator. Since cable television systems are not a general-purpose communications mechanism but a specialized system for transmitting numerous television channels and data

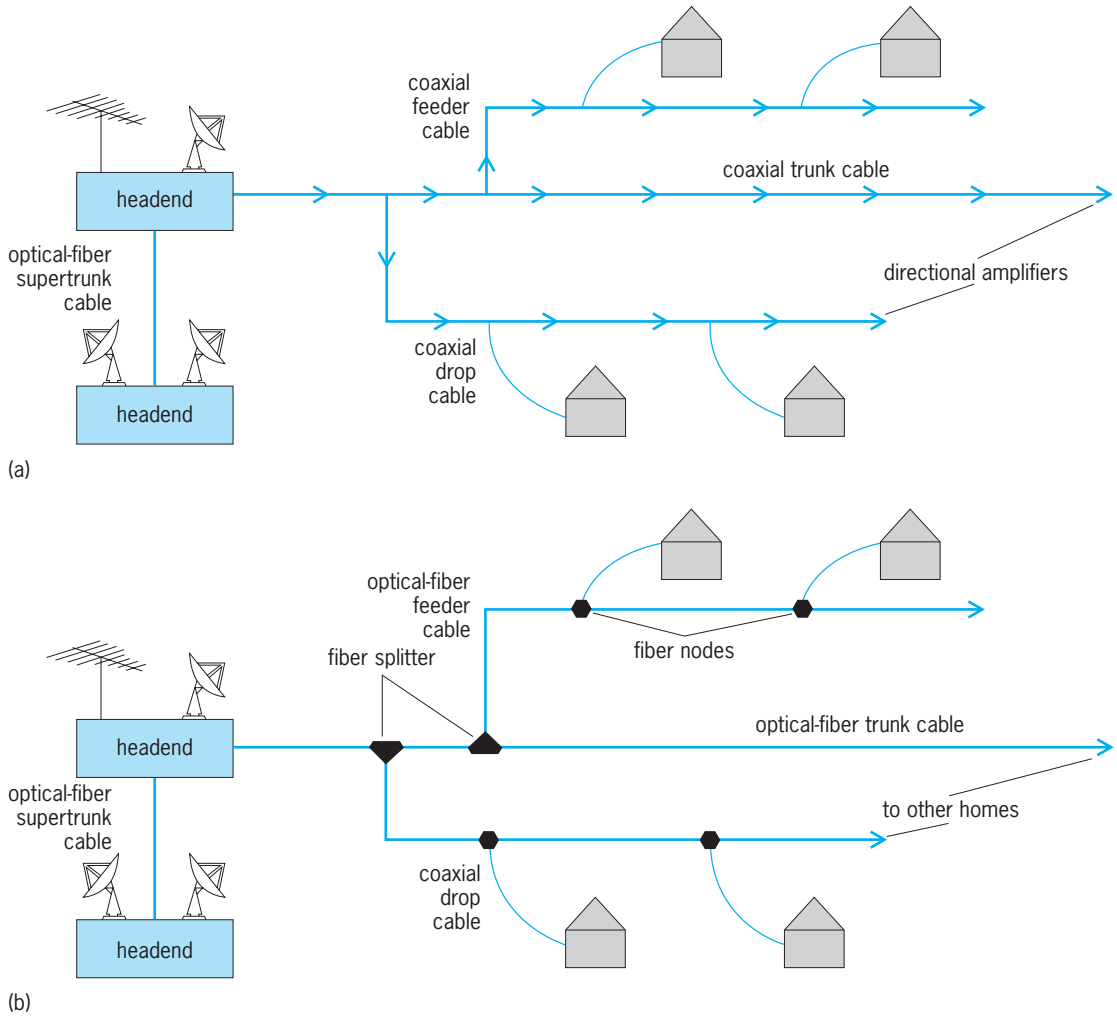

Cable television systems with (*a*) coaxial feeder cables and (*b*) optical-fiber feeder cables.

communications in a sealed spectrum, the topology can be customized for maximum efficiency. A cable television system comprises seven major parts: the headend, the fiber cable, the fiber node, the distribution (or feeder) cable in the neighborhood, the drop cable to the home, the in-home wiring, and the terminal equipment (that is, the set-top converter-descrambler). Generally, multiple-fiber strands carry the television and other communications signals from the headend to feeder or neighborhood distribution cables, and subscribers are connected through drop cables. Return or bidirectional data and voice signals, previously difficult to transmit in coaxial-only architecture, are possible with fiber cables. Multitudes of new services are possible with fiber architecture including broadband Internet connectivity, telephony services or VoIP, video on demand, interactive program guides, and video gaming.

Headend. The headend is the origination point or location where all of the cable television channels and data communications are received and assembled. Generally, the headend is centrally located in the cable television system for maximum efficiency. Multiple fibers emanate from the headend to serve many tens of thousands of subscribers.

Fiber cable. A fiber-optic cable contains a bundle of individual fiber strands (that is, individual optical conductors), a physical-strength member, and fillers, surrounded by a protective sheath. Replacing trunk cable and amplifiers, fiber-optic cables overcome most of the limitations of coaxial cable–based systems; they exhibit less attenuation per linear distance, require little or no amplification (thus reducing signal distortion), and provide greater spectral bandwidth with return or bidirectional signals on individual and dedicated fibers. Fiber strands are very small, and hundreds can be placed in the same space as one large coaxial cable.

Fiber node. A fiber node is the location in the distribution system where the fiber cable terminates and a coaxial distribution network, serving subscribers, begins. The node contains an optical receiver to convert the optical television and digital signals to radio-frequency (RF) spectrum energy for the coaxial system. Also at this node, return signals from subscriber terminals and signals related to telephony and Internet connectivity are separated from the coaxial cable for optical transmission back to the headend for

processing via a separate fiber strand. A fiber node may serve 100 to 250 subscribers depending upon neighborhood densities.

Distribution cable. The fiber-optic node interfaces with the coaxial distribution cable, which delivers signals to multiple subscribers. This distribution cable runs past homes and is tapped with passive directional couplers so that multiple subscribers may be connected. One or two specialized amplifiers may be included in this cable. *See* DIRECTIONAL COUPLER.

Drop cable. A small, flexible cable is used to connect an individual subscriber to the distribution cable through a tap or directional coupler. In the subscriber's home, this flexible drop cable delivers the signal to the terminal device, cable-ready television receiver, or video cassette recorder.

Terminal device. Commonly called a set-top converter, the terminal device may perform multiple functions. A converter allows non-cable-ready consumer devices to tune to one of the multiple channels delivered by the cable television system to the subscriber. A descrambler, however, may also perform the function of a converter, but its primary purpose is to allow subscribers who have paid for premium programming channels to receive them. Addressability is another function of the terminal device, which allows the cable operator to authorize or deauthorize subscriber access levels without requiring a technician to visit the subscriber premise. The set-top box may also interconnect with game systems, and provide high-quality (stereo FM–like) audio services, interactive program guides, digital video recorders, and other services.

Advanced technology. In the past, most cable television systems were designed to maximize delivery of analog video signals in only one direction. However, the minimum transmission quality criteria of optical fibers are well above the demands of digital transmissions. Optical fiber also improves the quality of the cable television signal, primarily because few or no amplifiers are needed. In order for cable television systems to provide full-duplex data communications, the system must be built for two-way transmissions by utilizing part of the spectrum for forward or downstream transmissions and another part for the reverse. Alternatively, a second and separate cable or fiber can be used to provide the duplex transmission path.

Cable systems today are mostly constructed with hybrid fiber-coaxial architecture to provide full-duplex communications capability. Because of the great amounts of bandwidth provided by optical fiber and coaxial cable as compared with other communications media, cable systems are in a unique position to provide expanded television programming and data services, and to support advanced technologies such as video on demand, high-definition television, digital communications such as Internet connectivity, and cable telephony.

Through the promulgation of new digital standards and specifications such as DOCSIS (Digital-Over-Cable Service Interface Specification), high-speed (megabytes per second) Internet connections are possible over cable television systems. A cable modem connected to a personal computer at the subscriber's location provides Ethernet reception and transmission with a central server at the cable company's headend. This central server is in turn connected to the Internet backbone for worldwide communications. Telephony and other voice-grade services are provided in the same model as Internet connectivity, with the headend providing hub or central-office facilities. Likewise, cable telephony can be delivered via a separate modem or by using VoIP multiplexed onto the Internet cable modem. *See* CLOSED-CIRCUIT TELEVISION; DATA COMMUNICATIONS; TELEVISION.

Roger D. Pience

Bibliography. W. Ciciora et al., *Modern Cable Television Technology: Video, Voice, and Data Communication*, 2d ed., Morgan Kaufmann, 2004; D. P. Dulchinos, *Twenty First Century Television: Cable Television in the Information Age*, 1993; B. Harrell, *The Cable Television Technical Handbook*, 1985; B. E. Keiser, *Broadband Coding, Modulation, and Transmission Engineering*, 3d ed., 1998; National Cable & Telecommunications Association, Science and Technology Department, *NCTA Technical Papers*, published annually.

Cable testing (electricity)

The testing of electric circuits to determine and locate any of the following circuit conditions: (1) an open circuit, (2) a short circuit with another conductor in the same circuit, (3) a ground, which is a short circuit between a conductor and ground, (4) leakage (a high-resistance path across a portion of the circuit, to another circuit, or to ground), and (5) a cross (a short circuit or leakage between conductors of different circuits). Circuit testing for complex systems often requires extensive automatic checkout gear to determine the faults defined above as well as many quantities other than resistance. *See* OPEN CIRCUIT; SHORT CIRCUIT.

In cable testing, the first step in fault location is to identify the faulty conductor and type of fault. This is done with a continuity tester, such as a battery and flashlight bulb or buzzer (**Fig. 1**), or an ohmmeter.

Murray loop test. Useful for locating faults in relatively low-resistance circuits, the Murray loop is shown in **Fig. 2** with a ground fault in the circuit under test. A known "good" conductor is joined to the faulty conductor at a convenient point beyond the fault but at a known distance from the test

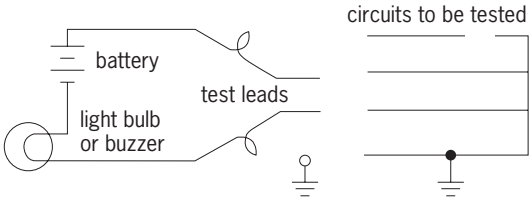

Fig. 1. Simple continuity test setup.

connection. One terminal of the test battery is grounded. The resulting Wheatstone bridge is then balanced by adjusting R_B until a null is obtained, as indicated by the detector in Fig. 2. Ratio R_A/R_B is then known. For a circuit having a uniform ratio of resistance with length, circuit resistance is directly proportional to circuit length. Therefore, the distance to the fault is determined from the procedure given by Eqs. (1)–(3). From Eq. (3) and a knowledge of

$$R_C \propto l + (l - x) \qquad R_D \propto x \tag{1}$$

$$\frac{R_A}{R_B} = r = \frac{R_C}{R_D} = \frac{2l - x}{x} \tag{2}$$

$$x = \frac{2l}{r - 1} \tag{3}$$

total length l of the circuit, once ratio r has been measured, the location of the fault x is determined. If circuit resistances are not uniform with distances, as when the known faultless conductor is different in size from the faulty conductor, additional calculation taking into consideration the resistance per unit length of the conductors is necessary.

If the fault is a short circuit or a cross instead of a ground, the battery is connected to the conductor to which the short or cross has taken place (**Fig. 3**). The test circuit is then equivalent to the one used to locate a ground (Fig. 2). The bridge is balanced and the calculations carried out as before. *See* BRIDGE CIRCUIT.

Varley loop test. This is similar to the Murray loop test except for the inclusion of the adjustable resistance R. The Varley loop (**Fig. 4**) is used for fault location in high-resistance circuits.

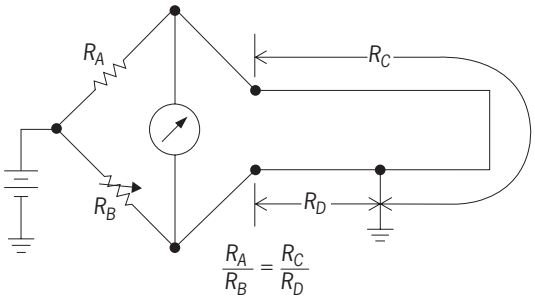

Fig. 2. Murray loop for location of ground fault.

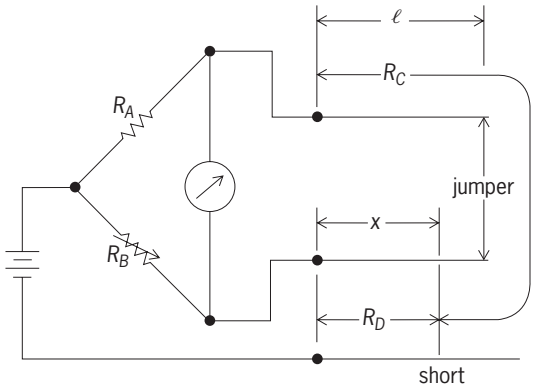

Fig. 3. Murray loop for location of short or cross fault.

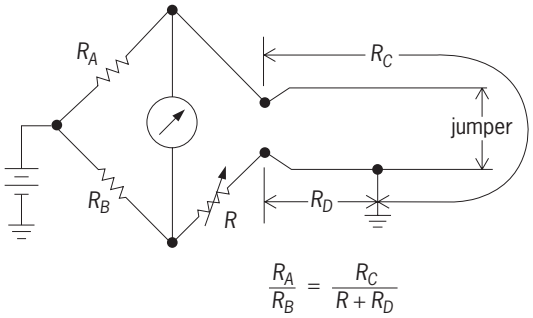

Fig. 4. Varley loop for location of leakage to ground.

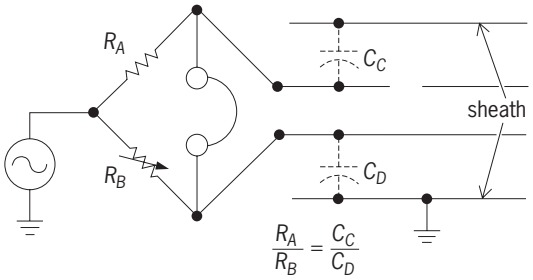

Fig. 5. Alternating-current capacitance bridge used in location of an open circuit in one conductor.

Open circuit. An alternating-current capacitance bridge can be used for locating an open circuit as shown in **Fig. 5**. One test terminal is connected to the open conductor and the other terminal to a conductor of known continuity in the cable. All conductors associated with the test are opened at a convenient point beyond the fault but at a known distance from the test connection. An audio oscillator supplies the voltage to the bridge, which is balanced by adjusting R_B for a null as detected by the earphones. Measured ratio R_A/R_B equals the ratio of capacitances between the lines and the grounded sheath. Because each capacitance is proportional to the length of line connected to the bridge, the location of the open circuit can be determined from Eq. (4).

$$\frac{R_A}{R_B} = \frac{C_C}{C_D} \tag{4}$$

For convenience in carrying out the tests for fault location, the fault location bridge has switches for setting up the various loop test circuits. Basically it is a bridge like other bridges used in circuit testing. *See* CIRCUIT (ELECTRICITY); OHMMETER; RESISTANCE MEASUREMENT; WHEATSTONE BRIDGE.

Charles E. Applegate

Bibliography. A. D. Helfrick and W. D. Cooper, *Electronic Instrumentation and Measurement Techniques*, 4th ed., 1990; L. Schnell (ed.), *Technology of Electrical Measurements*, 1993; L. M. Thompson, *Electrical Measurements and Calibration*, 2d ed., 1994.

Cacao

Theobroma cacao, a small tropical tree 13–19 ft (4–6 m) in height that is cultivated for the almond-shaped seeds, 0.8–1.2 in. (2–3 cm) long, which are used to make chocolate. It is a member of the tea family (Thealeae) in the order Theales. The species is native to the rainforest of the Amazon basin, and two regions of distribution in pre-Columbian times are recognized. The crop was first cultivated in Central America and northern South America, the varieties found there being known as Criollos. The second region comprises the Amazon and Orinoco basins, where the cacao populations are known as Amazonian Forastero. The second type is more commonly cultivated, particularly in Brazil, Ivory Coast, Ghana, and Nigeria, which produce 75% of the world annual output. *See* THEALES.

As a typical crop of the humid tropics, cacao is grown commercially between 15°N and 15°S, where mean annual temperatures vary from 73 to 79°F (23 to 26°C). It is generally accepted that the lower limits for successful growth are a mean monthly minimum temperature of 59°F (15°C) and an absolute minimum of 50°F (10°C). Annual rainfall in growing areas varies from 56 to 80 in. (1400 to 2000 mm).

Cultivation. Cacao can be grown in many different soils, provided they have good moisture retention and are well drained and aerated. Loose, crumbly, soft, and plastic soils facilitating root penetration are preferred. Cacao appears to have high soil fertility requirements, but can be cultivated in relatively poor soils if suitable management practices such as liming and fertilization are used.

In its natural forest habitat, cacao grows under taller trees. Traditional cultivation practices simulate such conditions by using thinned forest stands or planted shade trees. The main role of shade is to counteract unfavorable ecological factors, such as low soil fertility, wind damage, high transpiration leading to moisture stress, and increased insect attack. If all such factors can be controlled by other methods, the highest production is obtained without shade.

Cacao can be propagated by seeds or by vegetative methods (rooted cuttings or budding). The former is practically the only method used for commercial planting, with cultivars produced by hybridization. Seedlings are usually grown in shaded nurseries for 4–6 months and then transplanted to the field under temporary shade, such as that provided by banana plants. The recommended planting distance is 9 × 9 ft (3 × 3 m). The permanent shade trees are spaced at 77 ft (24 m).

Fruit and seeds. The first commercial crops are obtained in 4–5 years. In areas with even rainfall distribution, flowering may occur almost throughout the year. The flowers are borne directly on the trunk and larger branches. The fruit, an indehiscent drupe with a usually woody shell (**Fig. 1**), is red or green when unripe, and varies between 6 and 12 in. (15 and 30 cm) in length. The fruits ripen in 5–6 months and change in color to yellow or orange.

Fig. 1. Cacao (*Theobroma cacao*). (*USDA*)

The seeds, which may number up to 65 per fruit (average is 35–40), are surrounded by a mucilaginous pulp. The dry weight of the beans constitutes about one-third of the total dry weight of the pod. In most producing areas the mean annual yield varies from 660 to 1100 lb (300 to 500 kg) of dry beans per hectare, but yields of 4400–6600 lb (2000–3000 kg) may be obtained with superior cultivars and improved cultural practices. Harvesting consists of cutting the pods from the tree and opening them with a machete to remove the seeds. Following a period of 4–6 days of fermentation in heaps or "sweatboxes," the seeds (cocoa) are dried on sun-platforms or artificially until the final moisture content is about 7%.

Marketing. The produce is generally exported in the form of dry beans. The farmers' production is purchased by dealers and exported by registered exporters or government marketing boards. Sales are effected through contracts or futures markets, principally in New York and London. Growers' prices may be fixed at predetermined levels or may fluctuate according to market trends. International standards have been adopted to bring about uniformity. Accordingly, exported produce is subject to grading and treatment to avoid spoilage. The market distinguishes between bulk and fine cocoas. The latter have preferred flavor or other characteristics and receive a premium price. *See* COCOA POWDER AND CHOCOLATE.

Paulo de T. Alvim

Diseases. The most serious disease of cacao is black pod (**Fig. 2**), caused by the fungus *Phytophthora palmivora*, now recognized as having three morphological forms; two of these may be distinct species. This fungus causes great economic loss following wet weather favorable for disease development. The fungus can also attack leaves, branches, and trunks. Some control can be achieved with copper fungicides, and there is genetic material with

Fig. 2. The most widespread and most severe disease of cacao, caused by *Phytophthora palmivora*. (*Inter-American Institute of Agricultural Sciences*)

some resistance. Another pod rot, associated with the fungus *Botryodiplodia theobromae*, is occasionally a problem following wounding of pods by various vertebrates.

Witches'-broom, caused by the fungus *Crinipellis perniciosa* (formerly *Marasmius perniciosus*), occurs in tropical South America and in the West Indies. The fungus causes distorted growth of aerial plant parts, resulting in witches'-broom, and forms fruiting bodies ("mushrooms") on dead tissue. Some control is possible with sanitation and use of resistant varieties. Another New World (northwest South America, Panama, Costa Rica) disease is Monilia pod rot (watery pod rot), caused by the fungus *Monilia roreri*. Damage to pods can be severe. Control measures include removal of affected pods and use of resistant varieties. Swollen shoot, a virus disease transmitted by mealybugs, occurs only in western Africa and causes deformation of the tree and considerable economic loss. Eradication programs have been attempted. Some tolerance to the disease has been found. Other occasional problems include wilt, caused by the fungus *Ceratocystis fimbriata* and commonly spread by *Xyleborus* borers; and cushion galls, caused by a species of *Calonectria*. *See* PLANT PATHOLOGY; PLANT VIRUSES AND VIROIDS.

George A. Zentmyer

Bibliography. P. H. Gregory (ed.), *Phytophthora Disease of Cocoa*, 1974; C. A. Thorold, *Diseases of Cocoa*, 1975.

Cachexia

The severe wasting syndrome that accompanies such diseases as cancer, infection, or parasitic infestation. Occasionally, it is also observed in noninvasive conditions, such as severe cardiac failure. The causes of cachexia are only partially understood. However, it is clear that most cachexia is caused by diminished consumption of nutrients rather than by a hypermetabolic state.

Anorexia, the proximal cause of this problem, is thought to be related to the expression of endogenous factors collectively termed cytokines, some of which have now been identified. For example, tumor necrosis factor, a protein (also known as cachectin), when administered to animals for a long period of time, causes a syndrome of cachexia indistinguishable from that produced by chronic disease. Inhibitors of tumor necrosis factor activity may partially alleviate cachexia caused by experimental tumors. It is likely that other cytokines are also involved, and that together these agents cause wasting of such severity that it may lead to death in a wide variety of diseases. *See* ANOREXIA NERVOSA; CYTOKINE.

It is unclear whether cytokines suppress appetite through a direct action on satiety centers in the hypothalamus or through peripheral effects (such as effects on the bowel or other enteric structures), which then project to the central nervous system. Another clinical feature related to cachexia appears to be caused by a direct effect of cytokines (notably tumor necrosis factor and interleukin-1) on stem cells and stromal cells of the bone marrow. Tumor necrosis factor and interleukin-1 also appear to mobilize lipid through direct effects on adipocytes.

The cytokines that cause cachexia are produced mainly by cells of the immune system, especially macrophages. Synthesis is triggered by contact with molecules produced by microbial pathogens or tumor cells. Induction of synthesis depends upon both transcriptional and translational activation pathways. Rational strategies for alleviation of cachexia include eradicating the underlying infection or tumor, blocking cytokine synthesis with agents that specifically interrupt the requisite signaling pathways, or inhibiting cytokine activity with specific antibodies or other antagonists. Oral or intravenous administration of nutrients is likely to be effective if the process has not advanced to a point at which utilization of nutrients is impaired. *See* ONCOLOGY.

Bruce Beutler

Cactus

The common name for any member of the cactus family (Cactaceae). There are 120 genera with perhaps 1700 species, nearly all indigenous to America. The cacti are among the most extremely drought-resistant plants, and consequently they thrive in very arid regions. The group is characterized by a fleshy habit, presence of spines and bristles, and large, brightly colored, solitary flowers. There is a great

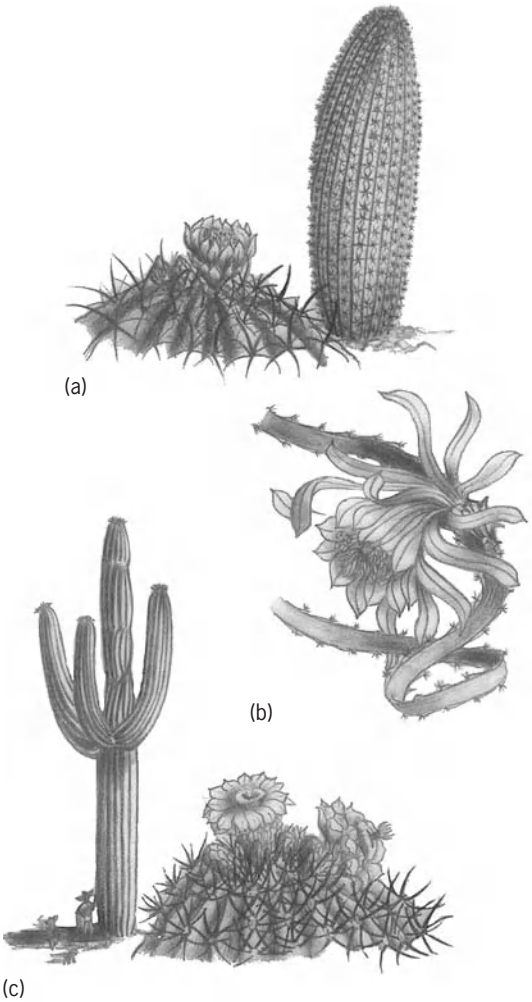

Fig. 1. Three examples of cacti. (*a*) Mule cactus, with closeup of flower, maximum height of 6 ft (2 m). (*b*) Flower of night-blooming *Cereus*, maximum height of 40 ft (12 m). (*c*) Saguaro, maximum height of 70 ft (21 m), with closeup of flower.

Fig. 2. *Mamillaria vaupellii* var. *flavispina*.

variety of body shapes and patterns, and many of the species are grown as ornamentals or oddities (**Figs. 1** and **2**). A few have edible fruits. The cochineal insect, which produces a valuable red dye, is cultivated chiefly on the cochineal cactus (*Nopalea coccinellifera*). The saguaro (*Cereus giganteus*) of Arizona and Sonora is the largest of the cacti, attaining a height of 70 ft (21m). *See* CARYOPHYLLALES.

Perry D. Strausbaugh; Earl L. Core

Cacti grow in many habitats, ranging from epiphytes living on trees in dense tropical forests to large, isolated plants in deserts. In all cases, various adaptations that lead to water conservation are apparent. For example, the stems of most cacti are massive and can store large amounts of water that sustain the plants during prolonged drought. Adaptations for water conservation also occur on a metabolic level. Cacti are increasingly recognized as important crops, especially in arid and semiarid regions. Perhaps the most widespread use of cacti is as ornamentals; collectors and hobbyists worldwide cherish their shapes, attractive flowers, and ability to tolerate long periods without watering.

Cactus roots are shallow [most occur within 20 cm (8 in.) of the soil surface], which allows the plants to respond rapidly to light desert rainfalls, as only a small depth of soil needs to be wetted. Moreover, roots represent only a small fraction of a cactus plant (generally only 7–12% of the total plant dry weight), suggesting that a relatively small root surface area is necessary to supply sufficient water from the soil. A conventional way to represent the water storage capacity of a plant is the volume available for water storage divided by the area across which water can be lost to the air by transpiration. This volume-to-area ratio is high for cacti compared with the majority of plants, whose photosynthetic organs are thin flat leaves. The leaves of most cacti are ephemeral or absent, so the photosynthetic organs are the relatively massive green stems. Besides the obvious succulence of the stems of cacti, their cells are relatively large with prominent vacuoles.

Another adaptation of cacti to arid conditions is their relatively slow loss of water to the air. Cacti have a thick waxy cuticle on their stems, which acts as a waterproofing skin. Also, the pores (called stomata or stomates) in the cactus skin that are necessary for the uptake of carbon dioxide (CO_2) from the atmosphere tend to open only at night. The opening of stomata allows the vaporized water within the plant to escape to the atmosphere. Because the stem and the air temperatures are lower and the relative humidity is higher at night than during the daytime, much less water (generally 80–90% less) is lost by transpiration during the nocturnal opening of stomata by cacti compared with the daytime stomatal opening of most other plants.

The opening of stomata at night presents a problem for photosynthesis, which requires light. Spec ifically, ph otosynthesis uses atmospheric carbon dioxide and the energy of sunlight to form sugars in the chloroplasts of the chlorenchyma (the green tissue that occurs outside the whitish water-storage tissue). Open stomata are necessary for appreciable uptake of carbon dioxide, which at night for cacti is attached to the three-carbon compound phosphoenolpyruvate to form a four-carbon acid such as malic acid. During the night the acid accumulates in the large vacuoles of cactus chlorenchyma cells; during the next daytime, when the stomata have closed, carbon

dioxide is released from the accumulated acid within the stems of cacti. Carbon dioxide is then fixed into sugars via photosynthesis when sunlight is available as the energy source. The opening of stomata and the increase in chlorenchyma acidity during the night followed by decarboxylation of organic acids and the formation of photosynthetic products during the daytime is known as crassulacean acid metabolism (CAM), because it was first discovered in the plant family Crassulaceae.

Crassulacean acid metabolism is crucial for the adaptation of cacti to the dry conditions characteristic of deserts. Its water-conserving attribute is also important in the increasing cultivation of cacti for their fruits and as fodder for livestock, particularly cattle. Currently just over 1 million hectares (2.5 million acres) of cacti are cultivated, mostly for the prickly pear cactus, *Opuntia ficus-indica* (the second most cultivated crassulacean acid metabolic species is pineapple, *Ananas comosus*). In particular, *O. ficus-indica* is extensively cultivated in Brazil (the leading country with over 300,000 ha) and northern Africa for fodder. In Mexico, fruit production is the main usage (about 60,000 ha); young stem segments of *O. ficus-indica* are also used in Mexico as a vegetable, termed nopalitos. Commercialization of its fruit, generally called cactus pears, requires the removal of the irritating and skin-penetrating glochids from the surface and even the application of a wax that improves appearance and prolongs shelf-life. Another use of mature stem segments of *O. ficus-indica* (and a few other cactus species), either on the plant or removed and placed in sheds, is for infestation by females of the scale insect *Dactylopius opuntiae*, which produce the brilliant red dye cochineal. Cochineal (carminic acid), 80% of which is currently produced in Peru, leads to colors ranging from yellow to purple for fabrics, foodstuffs, and even liqueurs.

Fruits termed pitayas, pitajayas, and various other names of more than 20 species of cacti are collected from the wild and sold locally throughout Latin America. These are produced by tall columnar cacti such as *Cereus peruvianus*, *Stenocereus griseus*, and *S. queretaroensis* as well as by vine-like cacti such as *Hylocereus polyrhizus*, *H. undatus*, *and Selenicereus megalanthus*. The pulp color varies from white to orange to deep purple; the seeds are small (like kiwi seeds) and readily ingested. The peels do not have the nasty glochids of *O. ficus-indica*, but certain postharvest problems such as a short shelf life require further investigation, and the market would also welcome greater year-round availability.

Cacti can grow in regions not suitable for other crops. Moreover, growth of cacti responds favorably to increasing levels of atmospheric carbon dioxide, and cacti can survive increasing atmospheric temperatures, two aspects of global climate change. Areas for cultivation of opuntias used for fodder are expanding, as are markets for fruits of many species. Indeed, because water is the principal limiting factor for agriculture in much of the world, the water-conserving qualities of cacti augur well for their increased commercialization in the future.

Park S. Nobel

Several diseases of wild cacti are known. The most conspicuous of these is probably bacterial soft rot of saguaro cacti, which is usua lly seen in older plants. Affected tissues appear as dark, soft lesions that usually develop a light-green border; these lesions frequently leak a dark, odorous fluid. Infected branches may fall, and the fleshy tissue of diseased plants may be destroyed in several weeks to a few months, leaving bare the woody supportive ribs (xylem elements) of the plant. Control by chemical sprays or soil drenches is not feasible. Therefore, when symptoms are first noted, rotting tissues should be removed. If the rot is limited, $^1/_2$ in. (13 mm) of surrounding healthy tissue should also be discarded. Immediately the cavity should be washed thoroughly with a solution of one part household bleach and nine parts water with the washing repeated in 18 to 24 h; the wound should be left open. Infections so treated frequently do not spread. Plants with numerous or large lesions should be removed and the debris thoroughly treated with the diluted bleach, or burned, or deeply buried.

The same controls apply to similarly infected organ-pipe cacti. Infected prickly-pear cacti initially have similar symptoms, but internal tissues can become so liquefied that pads bulge. The infected pad should be promptly destroyed, and the former point of attachment treated with bleach.

Recent studies in the United States have established that the soft-rot disease is caused by bacteria in the genus *Erwinia*, particularly *E. cacticida*. This species has also been isolated from rotting cacti in Australia and Mexico.

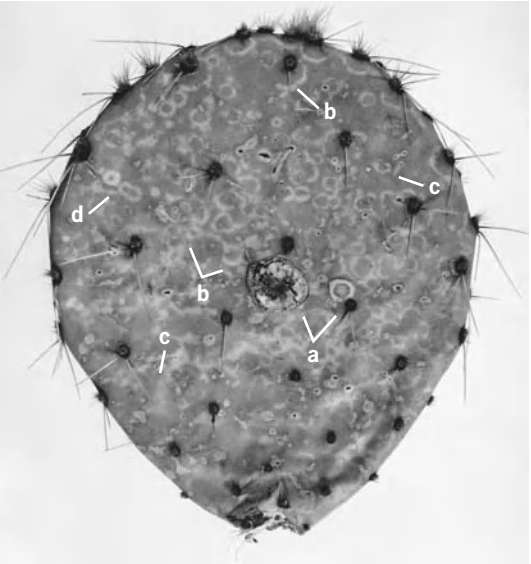

Fig. 3. Disease symptoms on prickly-pear pad: (*a*) lesions caused by *Phyllosticta concava*; (*b*) virus-induced symptoms; (*c*) injuries from penetrating cactus spines; (*d*) ring pattern caused by feeding insects. (*From S. M. Alcorn, R. L. Gilbertson, and M. R. Nelson, Some diseases of cacti in Arizona, Prog. Agr. Ariz., 27(3):3–5, 1975*)

Saguaro cacti may be toppled by heavy winds following soil-soaking rains. Roots of these plants usually have been previously weakened by the root- and wood-rotting fungus *Poria carnegieana*. Wind-thrown plants tend to be mature specimens; no control is known.

Prickly pears can be infected by other pathogens. Certain viruses cause whitish rings or irregularly shaped patterns to develop on pads (**Fig. 3**), Severely affected plants also are stunted. The light-colored lines usually are narrow. By contrast, rings caused by the feeding of sap-sucking insects have small, central green areas with broad, light-colored margins. Presumably, sap-sucking insects carry the viruses between plants, but "new," infected plants will develop if diseased pads root after their detachment from the parent plant. The fungus *Phyllosticta concava* causes the formation of generally circular, whitish to tan, depressed, dry lesions on pads; these areas usually become covered with black, slightly raised "dots." An infection may penetrate to the opposite side, where a similar lesion forms. Infected tissue can disintegrate, leaving a hole through the pad. Fungicidal sprays will minimize infections, which start in cool, rainy periods. *See* PLANT PATHOLOGY.

Stanley M. Alcorn

Bibliography. S. M. Alcorn, R. L. Gilbertson, and M. R. Nelson, Some diseases of cacti in Arizona, *Prog. Agr. Ariz.*, 27(3):3–5, 1975; S. M. Alcorn et al., Taxonomy and pathogenicity of *Erwinia cacticida* sp. nov., *Int. J. Sys. Bacteriol.*, 41(2):197–212, 1991; L. Benson, *The Cacti of the United States and Canada*, 1982; G. M. Milbrath, M. R. Nelson, and R. E. Wheeler, The distribution and electron microscopy of viruses of cacti in southern Arizona, *Phytopathology*, 63:1133–1139, 1973; Y. Mizrahi, A. Nerd, and P. S. Nobel, Cacti as crops, *Hort. Rev.*, 18:291–319, 1997; P. S. Nobel, *Remarkable Agaves and Cacti*, 1994.

Cadmium

A relatively rare chemical element, symbol Cd, atomic number 48, closely related to zinc, with which it is usually associated in nature. It is a silvery-white ductile metal with a faint bluish tinge. It is softer and more malleable than zinc, but slightly harder than tin. It has an atomic weight of 112.40 and a specific gravity of 8.65 at 20°C (68°F). Its melting point of 321°C (610°F) and boiling point of 765°C (1410°F) are lower than those of zinc. There are eight naturally occurring stable isotopes, and eleven artificial unstable radio isotopes have been reported. Cadmium is the middle member of group 12 (zinc, cadmium, and mercury) in the periodic table, and its chemical properties generally are intermediate between zinc and mercury. The cadmium ion is displaced by zinc metal in acidic sulfate solutions. Cadmium is bivalent in all its stable compounds, and its ion is colorless. *See* PERIODIC TABLE; TIN; ZINC.

Cadmium does not occur uncombined in nature, and the one true cadmium mineral, greenockite

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

(cadmium sulfide), is not a commercial source of the metal. Almost all of the cadmium produced is obtained as a by-product of the smelting and refining of zinc ores, which usually contain 0.2–0.4% cadmium. The United States, Canada, Mexico, Australia, Belgium-Luxembourg, and the Republic of Korea are principal sources, although not all are producers.

The fumes of cadmium, its compounds, and solutions of its compounds are very toxic, and cadmium-plated articles should not be used in food, nor should cadmium-coated articles be welded or used in ovens. Nickel-cadmium batteries are the second-largest application, with pigment and chemical uses third. Sizable amounts are used in low-melting-point alloys, similar to Wood's metal, and in automatic fire sprinklers. Cadmium compounds are used in the production of cadmium phosphors. Because of its great neutron-absorbing capacity, especially the isotope 113, cadmium is used in control rods and shielding for nuclear reactors. *See* ALLOY; CADMIUM METALLURGY.

Wilber Hague

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; R. R. Lauwerys and P. Hoet, *Industrial Chemical Exposure: Guidelines for Biological Monitoring*, 3d ed., 2001; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004.

Cadmium metallurgy

Cadmium resembles zinc in its chemical properties, and its occurrence for commercial purposes is with zinc ores. A relatively rare element in the Earth's crust, cadmium ranks in abundance between mercury and silver. Most cadmium occurs in solid solution in the zinc sulfide mineral called sphalerite. Although cadmium may be recovered from some lead and copper ores, it is associated with the zinc which is also found in these ores. Although a few zinc concentrates contain as much as 1–2% cadmium, typical content ranges from 0.06 to 0.5%. Since cadmium is entirely a by-product metal, the supply available is closely aligned with zinc production, averaging about 0.4% of zinc production. *See* SPHALERITE.

Recovery methods. All cadmium recovery processes involve the dissolution of cadmium-bearing feed material (sinter fume from the sintering of zinc ores or calcines, residue from the electrolyte purification of electrolytic zinc plants or from the purification of zinc sulfate solutions in the manufacture of zinc salts and pigments, and so on) followed by various purification and cadmium displacement steps. Methods of processing can be grouped conveniently into two basic categories, electrolytic and electromotive. In the former case, cadmium is recovered by electrolyzing purified solutions; in the latter case, cadmium in the form of a metallic sponge is displaced from purified solutions by a less noble metal, zinc being used in every known commercial application, and the sponge is melted or distilled, or both.

Electrolytic recovery. In electrolytic cadmium recovery, the feed material is dissolved in spent electrolyte either from the cadmium cells or from an electrolytic zinc circuit. The leach solution is neutralized to a pH of 5.0 or above, and troublesome impurities, such as copper, are removed by adding a small amount of zinc dust. Thallium can be removed from solution by precipitating with potassium permanganate or sodium chromate. Cadmium is electrolyzed from purified solutions at cell voltages of 2.4–2.8 V. Most electrolytic cadmium plants employ a current density of 3–8 A/ft^2 (32–86 A/m^2). Extensive refining of cathode cadmium is not required to produce metal that meets market specifications. In the usual case, the cathode metal is melted under a layer of caustic at a temperature of 716–950°F (380–510°C) for 3–18 h before it is cast into commercial shapes.

Electromotive recovery. Electromotive cadmium producers recover the metal from fume produced in the sintering of zinc calcines or zinc concentrates. Most fume contains appreciable concentrations of troublesome impurities, such as arsenic. Prior to leaching, fume may be oxidized by roasting or sulfated by baking with sulfuric acid. Roasted fume is leached in sulfuric acid solution; sulfated fume is leached with water. Leach solutions are purified to remove arsenic, iron, copper, thallium, and lead. For example, a small amount of zinc dust may be added to remove copper, and the filtered solution treated with potassium permanganate to oxidize iron (which is precipitated with caustic) and remove arsenic, following which a small amount of sodium chromate may be added to precipitate lead and thallium. Cadmium sponge is precipitated from purified solutions with zinc dust (zinc is higher in the electromotive series than cadmium is, and it displaces cadmium from solutions). Most electromotive producers densify the cadmium sponge by briquetting it in a press at pressures of 3000–12,000 lb/in.^{2} (20–80 megapascals). If the cadmium sponge is sufficiently pure, metal meeting ASTM specifications may be produced by melting the compacted sponge under caustic. Less pure sponge may require retorting (distilling) to meet product cadmium specifications.

Product specifications and applications. Product specifications are set forth in the **table**. Major end uses for cadmium are in corrosion-resistant plating and in cadmium compounds for use as pigments in paints, ceramics, and plastics. These two uses account for about 45% and 15%, respectively, of total cadmium production. The remainder is used in alloys, plastic stabilizers, batteries, and television picture tube phosphors. Other uses for cadmium include solar energy cells and energy storage systems.

Cadmium chemical requirements*

Element	Composition, %†
Cadmium	99.90
Zinc	0.035
Copper	0.015
Lead	0.025
Tin	0.01
Silver	0.01
Antimony	0.001
Arsenic	0.003
Thallium	0.003

*ASTM designation: B 440–69.
†The minimum percentage is indicated for cadmium, the maximum percentage for the other elements.

Cadmium is a toxic element. Care must be exercised to avoid breathing or ingesting it or its compounds. *See* CADMIUM; ZINC. Robert E. Lund

Bibliography. American Society for Testing and Materials, *Annual Book of ASTM Standards*, 02.0, *Nonferrous Metals*, 1992; R. A. Burkin, *Topics in Non-Ferrous Extractive Metallurgy*, 1980; C. H. Cotterill (ed.), *AIME World Symposium on Mining and Metallurgy of Lead-Zinc*, vol. 2, 1970; C. B. Gill, *Nonferrous Extractive Metallurgy*, 1988.

Caffeine

An alkaloid, formerly synthesized by methylation of theobromine isolated from cacao, but now recovered from the solvents used in the manufacture of decaffeinated coffee. Chemically, caffeine is 1,3,7-trimethylxanthine, and has the structural formula below. It is widely used in medicine as a stimulant

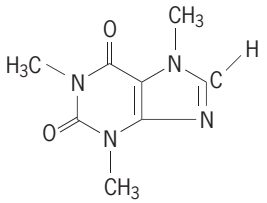

for the central nervous system and as a diuretic. It occurs naturally in tea, coffee, and yerba maté, and small amounts are found in cola nuts and cacao. Caffeine crystallizes into long, white needlelike crystals that slowly lose their water of hydration to give a white solid that melts at 235–237.2°C (455–459.0°F). It sublimes without decomposition at lower temperatures. Caffeine has an intensely bitter taste, though it is neutral to litmus. *See* ALKALOID. Frank Wagner

Cage hydrocarbon

A compound that is composed of only carbon and hydrogen atoms and contains three or more rings arranged topologically so as to enclose a volume of space. In general, the "hole" within a cage hydrocarbon is too small to accommodate even a proton. The carbon frameworks of many cage hydrocarbons are quite rigid. Consequently, the geometric relationships between substituents on the cage are well defined. This quality makes these compounds exceptionally valuable for testing concepts concerning bonding, reactivity, structure-activity relationships, and structure-property relationships. *See* CHEMICAL BONDING.

Platonic solids. The carbocyclic analogs of the platonic solids that are tenable are tetrahedrane (**1**, where X = —H), cubane (**2**), and dodecahedrane (**3**). According to theory, tetrahedrane is so highly

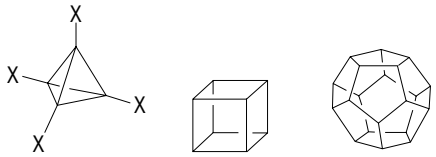

(**1**) (**2**) (**3**)

strained that it should dissociate spontaneously to two molecules of acetylene. Nevertheless, G. Maier was able to synthesize tetra-*t*-butyl tetrahedrane [**1**, X = —$C(CH_3)_3$] in 1978. The symmetry of the carbon skeleton of the tetrahedron permits the distances among all the bulky *t*-butyl groups [—$C(CH_3)_3$] to be maximized, and so their intramolecular repulsions are minimized. Cubane also requires the bond angles and bond lengths of its skeletal carbon atoms to be substantially distorted. Cubane was first prepared in 1964 by P. Eaton. Calculations suggest that octanitrocubane may be a high-density, high-energy material. Although the synthesis of dodecahedrane is difficult, independent preparations of this nearly globular cage hydrocarbon were reported in 1982 and 1987. *See* ALICYCLIC HYDROCARBON; BOND ANGLE AND DISTANCE.

Prismanes. An unsubstituted prismane has the general formula of $(CH)_n$, and the carbon atoms are located at the corners of a regular prism. Prismane (**4**), cubane, pentaprismane (**5**), and hexaprismane (**6**, unknown) are the simplest members of

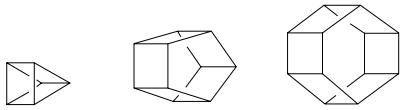

(**4**) (**5**) (**6**)

this family of cage hydrocarbons. Despite their high strain energies (which in most ring structures would facilitate ring opening), the prismanes have unusual kinetic stabilities due to the symmetry-imposed barriers that exist toward any ring opening. *See* MOLECULAR ORBITAL THEORY.

Adamantane. The monomer of the diamond carbon skeleton is adamantane (**7**, where X = —H). Although this compound has been isolated from petroleum, its availability is a result of a chance observation in 1957 that tetrahydrodicyclopentadiene (**8**),

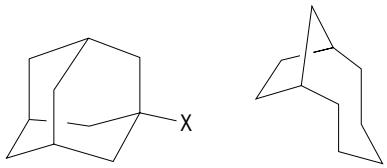

(**7**) (**8**)

an inexpensive compound, isomerizes to adamantane in the presence of aluminum halide catalysts. Eventually, it was recognized that drugs containing adamantyl groups are fat soluble and not readily degraded in the human body. Thus, such drugs are more persistent and longer acting than formulations based on long-chain hydrocarbons. Amantadine (**7**, where X = —NH_2) was developed commercially as the first orally active antiviral drug for the prevention of respiratory illness due to influenza A2-Asian viruses. Amantadine also has been found to be clinically effective for treating the symptoms of Parkinson's disease. *See* VIRUS CHEMOPROPHYLAXIS.

The other simple diamondoid hydrocarbons, diamantane (**9**) and triamantane (**10**), can be synthe-

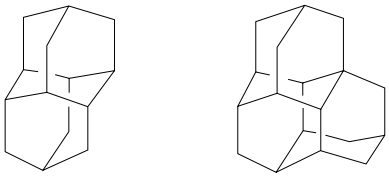

(**9**) (**10**)

sized by the Lewis acid–catalyzed rearrangement of isomeric hydrocarbons that possess a moderate degree of strain energy. Thus, treatment of (**11**) and (**12**) with appropriate catalysts give (**9**) and (**10**), respectively.

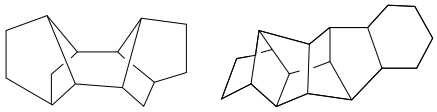

(**11**) (**12**)

Other structures. Organic chemists have prepared a wide variety of cage hydrocarbons that do not occur in nature. Among these compounds are triasterane (**13**), iceane or wurtzitane (**14**), and pagodane (**15**). Each of these compounds is prepared by

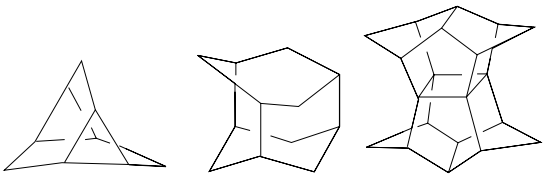

(**13**) (**14**) (**15**)

a complex multistep organic synthesis. *See* ORGANIC CHEMISTRY; ORGANIC SYNTHESIS. Roger K. Murray, Jr.

Bibliography. G. A. Olah (ed.), *Cage Hydrocarbons*, 1990; E. Osawa and O. Yonemitsu (eds.), *Carbocyclic Cage Compounds: Chemistry and Applications*, 1992.

Caisson foundation

A permanent substructure that, while being sunk into position, permits excavation to proceed inside and also provides protection for the workers against water pressure and collapse of soil. The term caisson covers a wide range of foundation structures. Caissons may be open, pneumatic, or floating type; deep or shallow; large or small; and of circular, square, or rectangular cross section. The walls may consist of timber, temporary or permanent steel shells, or thin or massive concrete. Large caissons are used as foundations for bridge piers, deep-water wharves, and other structures. Small caissons are used singly or in groups to carry such loads as building columns. Caissons are used where they provide the most feasible method of passing obstructions, where soil cannot otherwise be kept out of the bottom or where cofferdams cannot be used. *See* BRIDGE; PILE FOUNDATION.

The bottom rim of the caisson is called the cutting edge. The edge is sharp or narrow and is made of, or faced with, structural steel. The narrowness of the edge facilitates removal of ground under the shell and reduces the resistance of the soil to descent of the caisson.

Open caisson. An open caisson is a shaft open at both ends. It is used in dry ground or in moderate amounts of water. A bottom section, having a cutting edge (**Fig. 1**), is set on the ground and soil is removed from inside while the caisson sinks by its own weight. Sections are added as excavation and sinking proceed. After excavation is completed, the hollow interior of the shaft is filled with concrete.

Open caissons are usually constructed of reinforced concrete, but if steel shells are used, concrete may also be required to provide weight. When an open caisson is to be towed into place, hollow walls or false bottoms are provided to give it buoyancy; they are removed after the caisson is in position. Open caissons can be sunk in water to practically any depth.

In deep water an open caisson is sometimes sunk by the sand-island method. A cofferdam is constructed and filled with sand to form an artificial island. The island serves as a working platform and guide for sinking the caisson. This method also avoids the necessity of transporting a fabricated shell and holding it in position.

Figure 2 shows the major types of caissons in use. Chicago caissons (open well) are formed by excavating a few feet of soil, installing short vertical pieces of wood or concrete lagging beveled to form a circle, bracing with segmental hoops, and repeating. Such caissons are used in clay that is not strong enough to carry the permanent load alone but will permit excavation for the lagging without flowing in at the bottom. The clay is hoisted out in buckets or elevators. The bottom may be belled out. Lagging and bracing are generally left in place.

Fig. 1. Underside of open caisson for Greater New Orleans bridge over Mississippi River. (*Dravo Corp.*)

Sheeted caissons are similar to Chicago caissons, except that vertical sheeting is continuous and is driven by hand or pile-driving hammers either in advance of excavation or as it proceeds. Bracing rings are placed every few feet. Sheeting may be wood, concrete plank, or steel, and is left in place. Sheeted caissons may be round or square.

Small cylindrical caissons, 2–12 ft (0.6–3.6 m) in diameter, consist of a concrete shell or steel pipe sunk into the ground by its own or by a temporarily superimposed weight. Excavation is done by hand or mechanically, and the caisson is finally filled with concrete.

Drilled caissons are drilled holes that are filled with concrete. Rock socketing may be used at the bottom, or the bottom may be belled up to 30 ft (9 m) in diameter. The caisson may be as long as 150 ft (46 m). Cohesive soils are drilled with augers or bucket-type drills. Granular soils are drilled in the same manner with the aid of a binding agent or by rotary drills using drilling mud or water to keep the hole from caving in. Steel-shaft casings are used where needed to shut off water or keep soil out of the hole. Casings are removed before concreting, or during concreting in unstable ground.

Driven caissons are formed by driving a cylindrical steel shell with one or more pile-driving hammers. After excavation by grabs, buckets, or jetting, concrete is placed inside. The shells usually are withdrawn as the concrete rises, unless the soil is so soft or water conditions are such that they must be left in place.

Pneumatic caisson. A pneumatic caisson is like a box or cylinder in shape; but the top is closed and thus compressed air can be forced inside to keep water and soil from entering the bottom of the shaft. A pneumatic caisson is used where the soil cannot

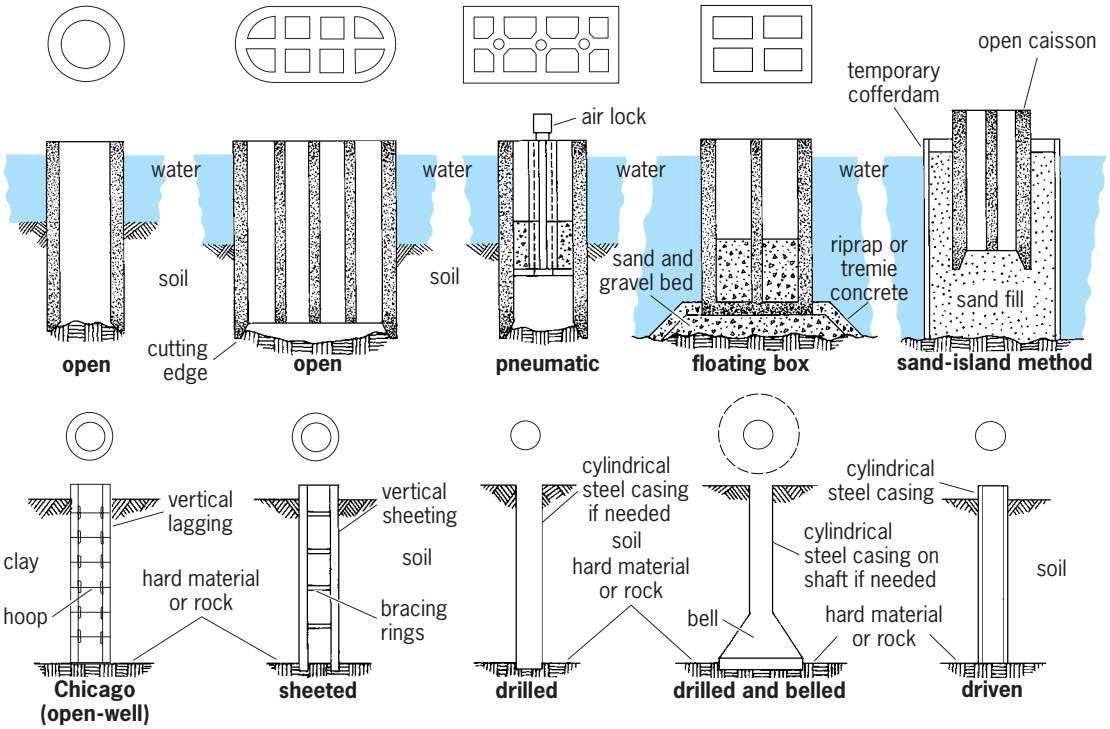

Fig. 2. Major types of caissons as they appear from top and in cross section from side.

be excavated through open shafts (for instance, where there are concrete, timbers, boulders, or masonry lying underwater) or where soil conditions are such that the upward pressure must be balanced. The air pressure must balance or slightly exceed the hydrostatic head and is increased as the caisson descends. The maximum depth is about 120 ft (37 m), corresponding to an air pressure of 52 lb/in.^{2} (360 kilopascals), which is about the limit of human endurance. Workers and materials must pass through air locks. Too rapid decompression may result in caisson disease or "bends," resulting from the expansion of bubbles of air trapped in joints, muscles, or blood. The length of time that workers can perform under pressure and the speed of decompression are regulated by law. Pneumatic caissons may be started as open caissons, then closed and air applied. They are made of reinforced concrete, which may be faced on one or both sides with steel plates. *See* DIVING.

Floating caisson. A floating or box caisson consists of an open box with sides and closed bottom, but no top. It is usually built on shore and floated to the site where it is weighted and lowered onto a bed previously prepared by divers. The caisson is then filled with sand, gravel, or concrete. This type is most suitable where there is no danger of scour. Floating caissons may be built of reinforced concrete, steel, or wood.

Small box caissons usually consist of a single cell. Large caissons are usually divided into compartments; this braces the side walls and permits more accurate control during loading and sinking. *See* FOUNDATIONS. Robert D. Chellis

Calamine

A term that may refer to either a zinc mineral, $Zn_4Si_2O_7(OH)_2{\cdot}H_2O$, which is also known as hemimorphite, or to zinc oxide, ZnO, which is used in medicinal or pharmaceutical products and in cosmetics. *See* HEMIMORPHITE; ZINC. E. Eugene Weaver

Calanoida

An order of Copepoda that includes the larger and more abundant of the pelagic species. Some authorities consider the Calanoida an order of the subclass Copepoda. In the food cycles of the sea these copepods are the most important group of marine animals because of their overwhelming numbers, ubiquitous distribution, and position at the base of the animal food chain. *See* COPEPODA.

The anterior part of the body is cylindrical with five or six segments, and much broader than the posterior part, usually with two, three, or four segments in the female and five segments in the male. A pair of caudal rami bear short setae, often of nearly equal length. The first antennae have 20 to 25 segments and usually extend to the ends of the caudal rami or beyond. They are not for locomotion, but are stabilizers and sinking retarders, and also have an olfactory function. In the male the right first antenna is often geniculate for grasping the female. The second antennae and mandibular palps are biramous and create water currents for feeding and slow movement. These and the first and second maxillae and maxillipedes have long setae and create feeding-current

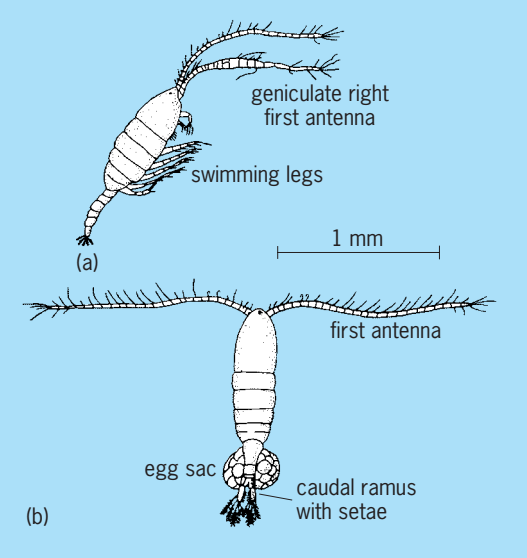

***Diaptomus*. (*a*) Lateral view of male. (*b*) Dorsal view of female.**

eddies and strain small food organisms from the water. The five pairs of swimming legs are biramous, but the last pair is sometimes reduced or absent in the female, and the male's right fifth leg may be modified for grasping the female (see **illus.**). There is a small oval, dorsal thoracic heart with an anterior aortic opening, two posterior venous openings, and one ventral venous opening. A pair of excretory glands are in the second maxillae, and an oil storage sac is usually found dorsal to the digestive tract in the thorax.

In this order there are 120 or more recognized genera, some with numerous species. In the fresh-water genus *Diaptomus* about 100 species are known. However, except in surveys over very wide geographic areas, the number of species to be found in any one genus is usually less than 10 or 15.

Calanus is the most outstanding genus in boreal and arctic seas. *Calanus finmarchicus* (Gunnerus) was the first copepod to be described. It is a cold-water species, 0.12–0.24 in. (3–6 mm) long, which ranks first among North Atlantic copepods, and several researchers have claimed that, feeding at the second trophic level, it may be the world's most numerous animal species.

In biomass, the calanoids exceed all other copepod groups. In the Gulf of Maine alone the total standing crop of calanoids has been calculated at 4,000,000 tons.

Nearly all calanoids are planktonic and, as a group, occur in all parts of the oceans from the surface to abyssal depths. The geographic and bathymetric ranges of many species are, however, restricted by the nature of water currents and the chemical and physical conditions of the water. In their southernmost range the northern species are found at greater depths.

Many species live at depths of 650–1000 ft (200–300 m) during the day, but migrate upward at night and return to usual levels again at dawn. These diurnal migrations are doubtless in response to conditions of light, but are modified by the temperature tolerances of the species and by other factors.

In keeping with their planktonic habit, many calanoids are provided with long plumose setae and long antennae to aid in flotation. Often the body or plumes are brightly pigmented. Common colors are pink, red, green, blue, and violet. Some are mostly black or entirely hyaline. A number of species, especially of the genera *Metridia* and *Pleuromomma*, are bioluminescent, emitting flashes of phosphorescent light. *See* BIOLUMINESCENCE; ZOOPLANKTON.

Harry C. Yeatman

Bibliography. J. Green, *Biology of Crustacea*, 1961; S. M. Marshall and A. P. Orr, *The Biology of a Marine Copepod, Calanus finmarchicus (Gunnerus)*, 1955; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; G. O. Sars, *Crustacea of Norway*, vol. 4, 1903; T. H. Waterman (ed.), *The Physiology of Crustacea*, 2 vols., 1960, 1961.

Calcarea

A class of the phylum Porifera, including exclusively marine sponges with a skeleton composed of spicules of calcium carbonate to which can be added a solid basal calcitic skeleton. Calcarea vary from radially symmetrical vase-shaped species, to aggregates made up of a reticulum of thin tubes, to irregular massive forms. Calcareous sponges are mostly small

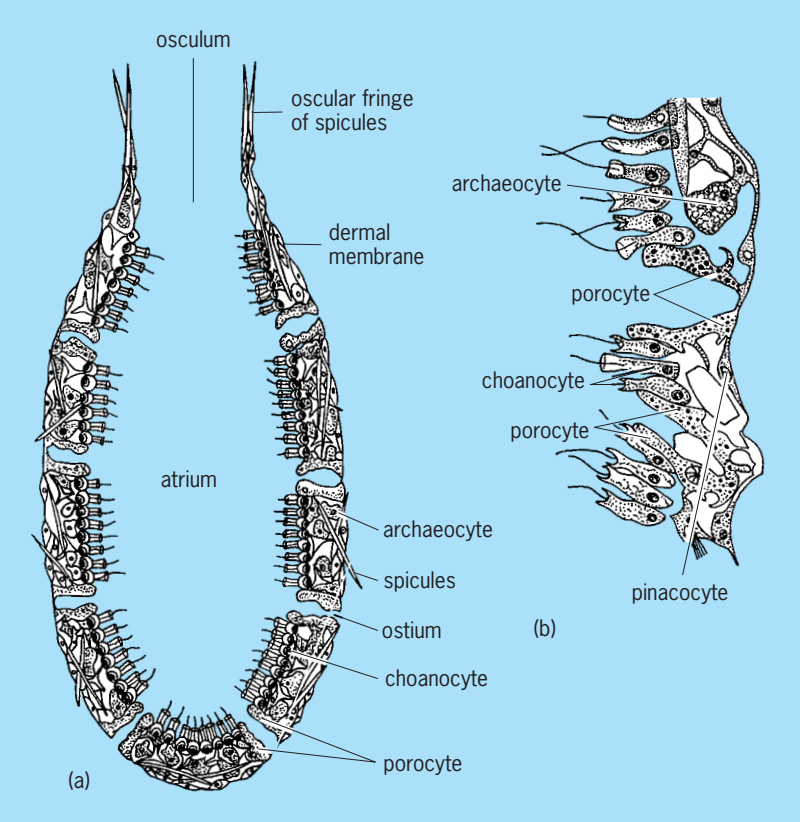

Fig. 1. Morphology of asconoid calcareous sponge. (*a*) Shown in longitudinal section. (*b*) Cross section through *Clathrina* showing cell types.

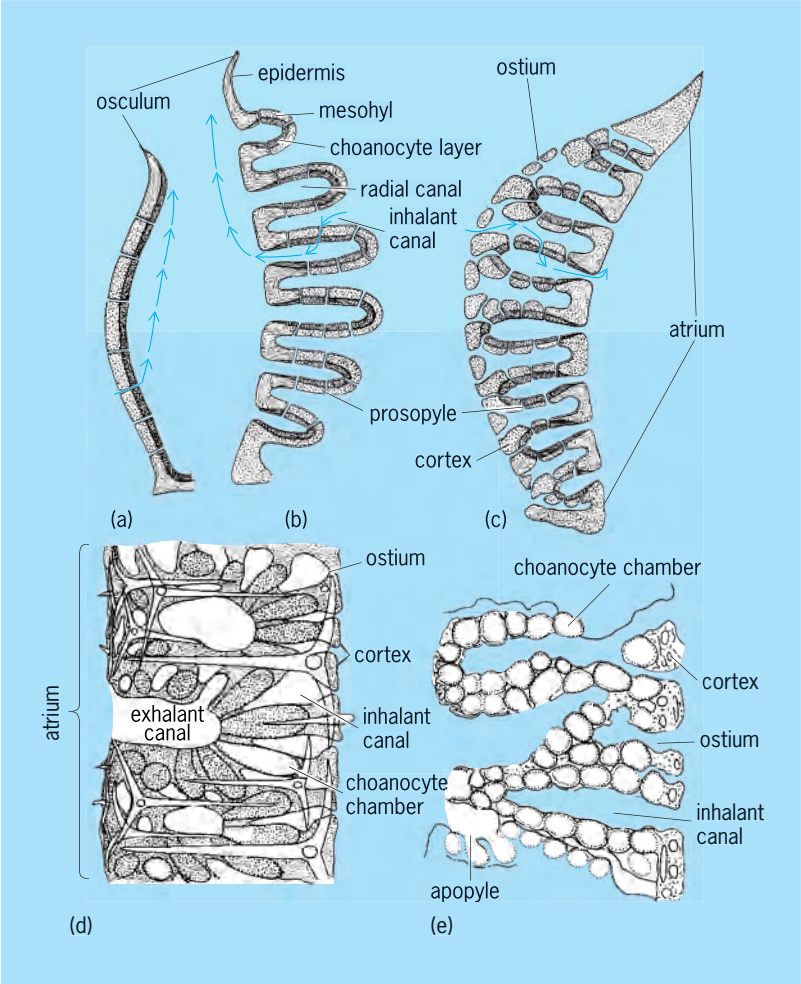

Fig. 2. Grades of construction of water-current system in Calcarea. (*a*) Asconoid type. (*b*) Syconoid type without cortex. (*c*) Syconoid type with cortex. (*d*) Intermediate or transitional stage between syconoid and leuconoid type. (*e*) Leuconoid type.

and inhabit the shallow waters of all seas, from tidal areas to depths of 660 ft (200 m), with a few species extending down to at least 13,200 ft (4000 m). *See* PORIFERA.

Morphology. The simplest calcareous sponges with an asconoid grade of construction consist of aggregates of upright tubes with unfolded walls made up of an outer epidermis of pinacocytes and an inner lining of choanocytes (**Fig. 1**). Between these layers of cells is a stratum of mesohyl including archaeocytes and spicules. Cells called porocytes, each perforated by a tubular canal, pierce the walls at intervals and allow water to enter the central cavity or atrium (**Fig. 2***a*). Water leaves by way of a terminal osculum.

A somewhat more complicated structure is seen in calcareous sponges of the syconoid grade of construction (Fig. 2*b*). Syconoid sponges are usually individual vase-shaped forms with a thick wall enclosing a large central atrium opening out through a terminal osculum. In the simplest forms the wall is pushed out at intervals into fingerlike projections, called radial canals, in which the choanocytes are localized. Water enters the radial canals directly through pores without the intervention of special inhalant canals. In most syconoid species, however, a dermal membrane made up of pinacocytes and mesohyl forms a cortex of greater or less thickness which joins the outer ends of the radial canals. Pores or ostia pierce the dermis and open into inhalant canals which are simply the spaces between the radial canals in some cases. In forms with a thick cortex (Fig. 2*c*), the inhalant canals run a course through the cortical mesohyl before reaching the outer ends of the radial canals (choanocyte chambers).

The leuconoid grade of construction has probably evolved independently among the several lines of calcareous sponges. In those with a syconoid ancestry, the radial canals subdivide into many small choanocyte chambers which arise as outpocketings of the radial canal wall. A common intermediate stage is seen in genera in which each radial canal is subdivided into elongate choanocyte chambers grouped around a common exhalant canal (Fig. 2*d*). Several lines of calcareous sponges have apparently reached the leuconoid grade of construction during the course of evolution through the anastomosis of ascon tubes, rather than by the branching of the radial canals of a syconoid sponge (Fig. 2*e*).

Spicules. Calcareous sponge spicules are monaxonid (diactines), triradiate (triactines), or quadriradiate (tetractines) in form. Diactines, where present, tend to occur in the cortex from which they may project to form a bristly surface. In some species, diactines project around the osculum. Either triactines or tetractines or both, depending on the species, are distributed in the mesohyl, often arranged in definite patterns in relation to the cortex and internal structures. The spicules provide increased rigidity for the tube and serve as protective elements.

Calcareous spicules are secreted extracellularly in a cavity formed by several sclerocytes. Two cells participate in the formation of a diactine or each ray of a triactine and tetractine (**Fig. 3**). The groups of sclerocytes are so oriented that the individual rays

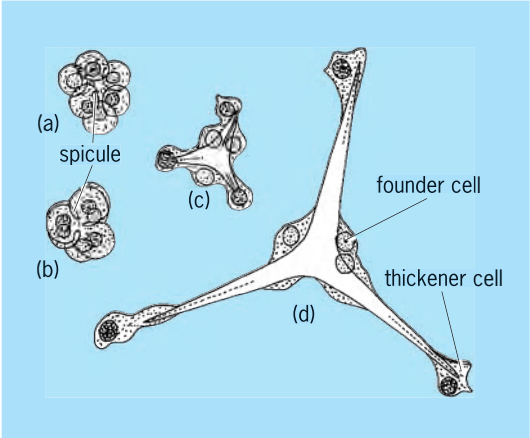

Fig. 3. Spicule formation in the calcareous sponges. (*a*-*c*) Early stages in secretion of triactine spicule, with founder and thickener cells. (*d*) Late stage of spicule formation.

fuse basally to form a triradiate or quadriradiate pattern.

Taxonomy. The monophyletic origin of calcareous sponges has been confirmed by recent molecular phylogenetic studies. Several of these studies suggest that Porifera are paraphyletic, with Calcarea appearing as a distinct phylum which would be the sister group of nonsponge metazoans. This point may be important in the knowledge of the origin of metazoans, but presently remains an open question. The Calcarea may be divided into two evolutionary lines based on morphology and development which have been recently supported by molecular phylogeny. In these two lines, corresponding to the subclasses Calcinea and Calcaronea, all the steps in increasing complexity of the aquiferous system organization and the architecture of the skeleton are present. These steps have been considered as stages in evolution, from simple forms with asconoid organization to more complex ones with leuconoid organization. Such a gradist view is now in the process of being replaced by a more phylogenetic conception, although the evolutionary pathways are presently not fully understood.

In the subclass Calcaronea, including the asconoid genus *Leucosolenia* and its syconoid and leuconoid relatives, the choanocytes have a nucleus apical in position, the triactine spicules are inequiangular, and the free-swimming larvae are amphiblastulae (**Fig. 4***a*). These are larvae with an anterior hemisphere of flagellated cells and a posterior mass of nonflagellated cells. The fate of these two categories of cells is determined at an early stage of cleavage. After a short free-swimming period, the amphiblastula larva settles down on its anterior end, and gastrulation occurs by invagination of the flagellated hemisphere or by downgrowth of the nonflagellated cells. A small tubular asconoid sponge called an olynthus results, and this may retain the asconoid structure or differentiate later into a syconoid or leuconoid structure according to species. Flagellated cells become choanocytes, while nonflagellated cells are totipotent and give rise to all other cell types of the adult sponge. Calcaronea includes three orders, Leucosoleniida, Lithonida, and Baerida. In Leucosoleniida, with nine families and 42 genera, the skeleton is composed exclusively of free spicules. In Lithonida, with one family and five Recent genera, the skeleton is reinforced by linked or cemented spicules. In Baerida, with four families and nine genera, the skeleton includes microdiactines and/or dagger-shaped small tetractines and may be reinforced by a rigid basal mass of calcite. *See* CALCARONEA; LEUCOSOLENIIDA.

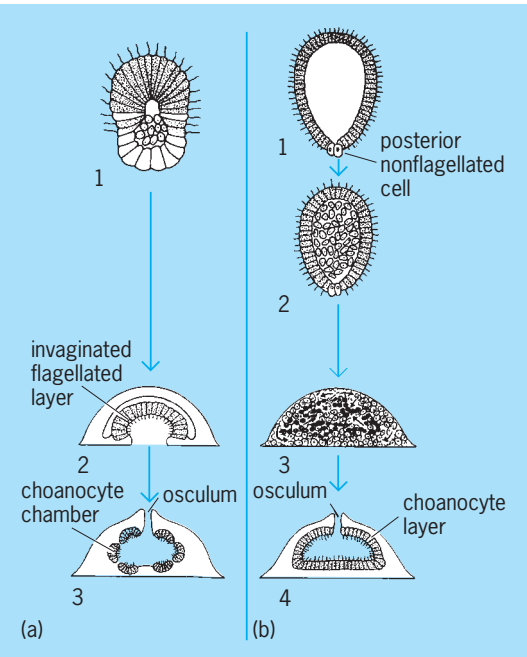

Fig. 4. Larval structure and metamorphosis in calcareous sponges. (*a*) Developmental stages characteristic of the Calcaronea: 1, amphiblastula larva; 2, newly settled larva showing metamorphosis; 3, young syconoid stage. (*b*) Developmental stages characteristic of the Calcinea: 1, coeloblastula larva; 2, solid larva; 3, newly settled larva; 4, young sponge of asconoid grade.

In the subclass Calcinea, including the asconoid genus *Clathrina* and its syconoid and leuconoid relatives, *Leucetta*, and others, the nucleus is located basally in the choanocytes, the triactine spicules are equiangular, and the free-swimming larva is a coeloblastula (Fig. 4*b*). In this case, cleavage results in a hollow blastula made up of a single layer of flagellated cells, with one to several nonflagellated cells at the posterior pole, some of which wander into the interior of the embryo. The larva so produced leads a short free-swimming life and then settles to the bottom, where it attaches by the anterior pole and flattens out. The cells of the larva differentiate according to their position in the metamorphosing sponge. Cells of the outer layer become pinacocytes; those of the interior become choanocytes that line the spongocoel. A young asconoid sponge or olynthus is thus formed that will develop into the adult condition. Calcinea includes two orders, Clathrinida and Murrayonida. In Clathrinida, with six families and 16 genera, the skeleton is composed exclusively of free spicules. In Murrayonida, with three Recent species classified in three families, the skeleton is reinforced either by a rigid network of calcite or by calcareous plates or spicules tracts. *See* CALCINEA; CLATHRINIDA; LEUCETTIDA.

Jean Vacelet; Willard D. Hartman

Fossils. Many fossil, and some living, sponges have a skeleton built largely of massive, nonspicular calcium carbonate; it is called a basal skeleton. Such sponges are said to be hypercalcified. Some belong to the class Calcarea, as demonstrated by the presence of typical calcarean triradiate spicules (three rays in one plane) and tetraradiate spicules (a fourth ray at right angles to the other three). Others have siliceous monaxons typical of the class Demospongea. Many have no spicules and, if fossil, cannot be assigned with certainty to a class; however, if living, they can be assigned on the basis of cell morphology. In general, basal skeletons of aragonite are found in ceractinomorph Demospongea, while basal skeletons of

(a) (b)

Fig. 5. ***Trachysinia*, a Jurassic pharetrone with inozoan morphology. (*a*) Entire sponge. (*b*) A fiber enlarged, showing the basal skeleton with embedded triradiates and tetraradiates. (*After G. J. Hinde, 1884, and R. M. Finks and J. K. Rigby, From Treatise on Invertebrate Paleontology, courtesy of and © 2004, The Geological Society of America and The University of Kansas*)**

magnesian calcite are found in Calcarea or in tetractinomorph Demospongea. *See* DEMOSPONGEA; SKELETAL SYSTEM.

All the fossil sponges of the class Calcarea belong to the subclass Calcaronea. Those with a basal skeleton are known as pharetrones. Their fossil record goes back to the Permian Period. Their morphology is either inozoan (**Fig. 5**), in which the basal skeleton is organized into a network of fibers, or sphinctozoan (**Fig. 6**), in which the basal skeleton coats the soft parts so as to form a series of perforated chambers. In both types, typical calcarean spicules—triradiates, tetraradiates, and "tuning-fork" spicules (triradiates in which two rays are subparallel)—are embedded in the basal skeleton. Forms in which the spicules are merely cemented together by calcite are known as lithonines. Their fossil record goes back to the Carboniferous Period.

There are, however, large groups of fossils of considerable abundance, often forming reefs (as do the pharetrones), which have a massive calcareous skeleton but whose living relatives appear to be demosponges. These groups include species of inozoan and sphinctozoan morphology, with aragonite basal skeletons of spherulitic, penicillate, or microgranular structure, assigned to the demosponge orders Agelasida and Vaceletida, as well as some with magnesian calcite basal skeletons assigned to the demosponge order Hadromerida. They are especially abundant in the Permian and Triassic periods, and also include most of the living species called sclerosponges. They probably include as well the stromatoporoids, the chaetetids, and the favositids, formerly considered to be corals, and perhaps also the Archaeocyatha (see below). Because many of them lack spicules (when spicules are present they are siliceous monaxons), they are sometimes difficult to distinguish from pharetrones. *See* COELENTERATA; REEF; STROMATOPOROIDEA.

Another group, the Heteractinida, which is confined to the Paleozoic Era, probably belongs to the class Calcarea. Their typical spicule has six rays in one plane, sometimes with two more at right angles to the six. This type of spicule can be thought of as a double triradiate (or tetraradiate), and the interpretation is strengthened by the simultaneous occurrence of triradiates and tetraradiates (typical of Calcarea), together with the six-rayed forms, in late Paleozoic representatives of the group, such as Wewokella.

Another possibly related group is the Archeocyatha. This exclusively Cambrian group has a pharetrone-like skeleton, but the double-walled conical growth form is not a typical pharetrone shape. *See* ARCHAEOCYATHA.

Although pharetrones have existed from the Permian to the present-day, they were most abundant and diversified in the late Paleozoic and Mesozoic. The lithonines are not known before the Jurassic. *See* ECHINODERMATA.

Robert M. Finks

Bibliography. R. M. Finks and J. K. Rigby, Hypercalcified sponges, in *Treatise on Invertebrate Paleontology*, pt. E: *Porifera Revised*, vol. 3, pp. 585–764, ed. by R. L. Kaesler, 2004; W. D. Hartman, A re-examination of Bidder's classification of the Calcarea, *Sys. Zool.*, 7(3):97–110, 1958; J. N. A. Hooper and R. W. M. V. Soest, *Systema Porifera: A Guide to the Classification of Sponges*, Plenum, New York, 2002; M. Manuel et al., Phylogeny and evolution of Calcareous sponges: Monophyly of Calcinea and Calcaronea, high level of morphological homoplasy, and the primitive nature of axial symmetry, *Sys. Bio.*, 52:311–333, 2003; R. Wood, *Reef Evolution*, Oxford University Press, 1999.

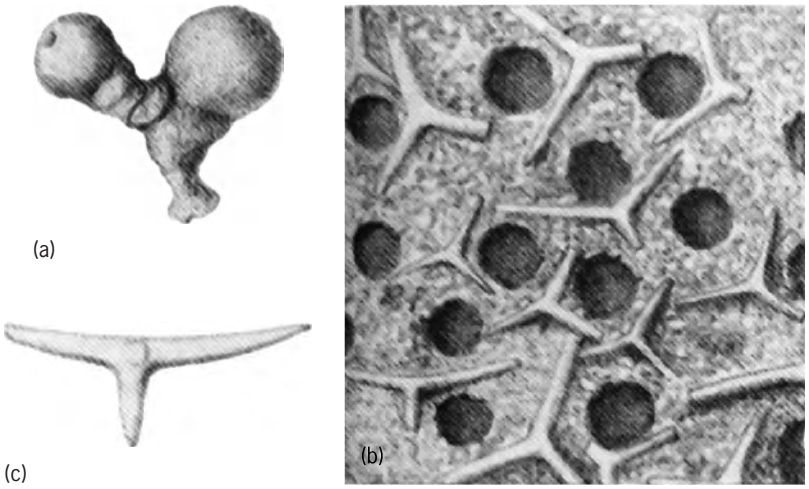

(a) (b) (c)

Fig. 6. ***Tremacystia*, a Cretaceous pharetone with sphinctozoan morphology. (*a*) Entire sponge. (*b*) Outer surface enlarged, showing perforated basal skeleton with embedded tetraradiates (fourth ray directed inward). (*c*) Isolated tetraradiate. (*After G. J. Hinde, 1884, and R. M. Finks and J. K. Rigby, From Treatise on Invertebrate Paleontology, courtesy of and © 2004, The Geological Society of America and The University of Kansas*)**

Calcichordates

Primitive fossil members of the phylum Chordata with a calcite skeleton of echinoderm type. They occur in marine rocks of Cambrian to Pennsylvanian age (530–300 million years old) and, because of their skeletons, have traditionally been placed in the phylum Echinodermata. They are shown to be chordates, however, by many chordate anatomical

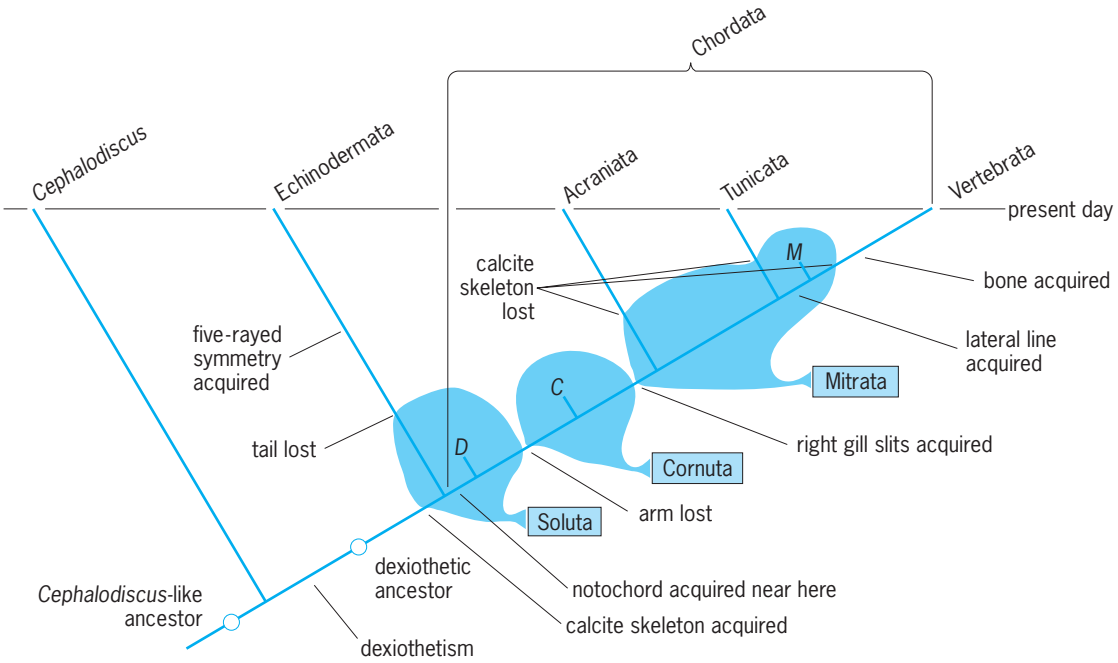

Fig. 1. Evolutionary relationships of calcichordates (Soluta, Cornuta, and Mitrata). The oldest events are at the bottom and the most recent are at the top. *D* = *Dendrocystoides*; *C* = *Cothurnocystis*; *M* = *Mitrocystites*.

features. Their calcite skeletons merely confirm an old view—that echinoderms and chordates are closely related. *See* CHORDATA; ECHINODERMATA.

There are three main groups of calcichordates—the Soluta, the Cornuta, and the Mitrata. **Figure 1** shows the relationships of these groups to each other, to the extant echinoderms and chordates, and to the recent hemichordate *Cephalodiscus*. The last is a small marine animal (**Fig. 2**) with a head shield, several pairs of arms, a trunk region with a pair of gill slits, and a locomotory stalk. *See* HEMICHORDATA.

In the common ancestry of echinoderms and chordates, a *Cephalodiscus*-like animal probably lay

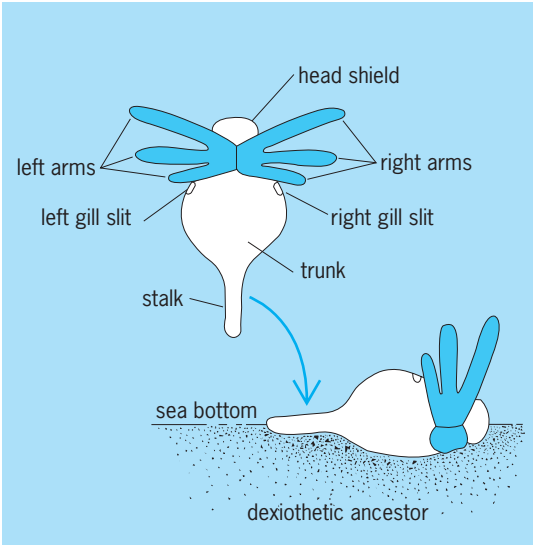

Fig. 2. The *Cephalodiscus*-like ancestor, by falling over on its right side, became the dexiothetic ancestor. This form acquired a calcite skeleton and gave rise both to the echinoderms and to the chordates.

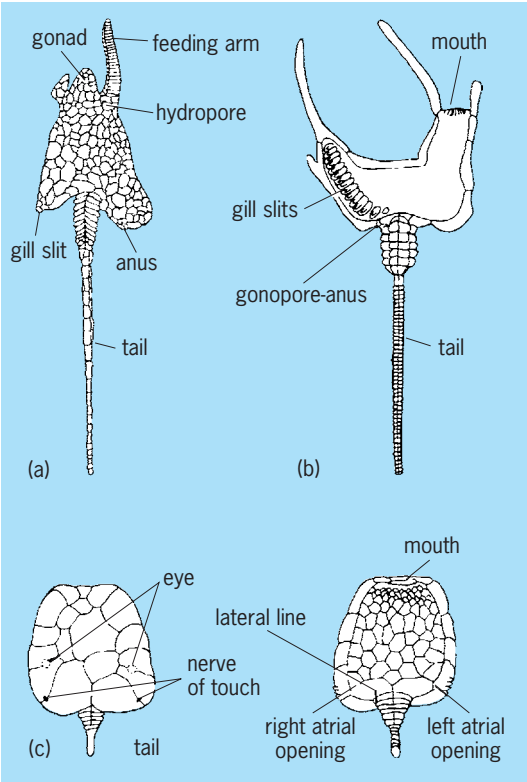

Fig. 3. The three main groups of calcichordates. (*a*) Soluta; dorsal view of *Dendrocystoides* (Ordovician, Scotland, about 440 million years old). (*b*) Cornuta; dorsal view of *Cothurnocystis* (Ordovician, Scotland, about 440 million years old). (*c*) Mitrata; dorsal and ventral views of *Mitrocystites* (Ordovician, Czechoslovakia, about 460 million years old). The atrial openings are external outlets for groups of gill slits inside the head. The lateral line is a sense organ found also in fishes.

down on its right side (dexiothetism) on the sea floor (Fig. 2). The right arms and right gill slit were lost as a result. Primitive right became ventral while primitive left became dorsal. The resulting animal, called the dexiothetic ancestor, acquired a calcite skeleton and gave rise both to the echinoderms and to the chordates.

The Soluta (**Fig. 3***a*) had an echinodermlike feeding arm and a tail homologous with the stalk of *Cephalodiscus* and with the tail of chordates. There was a gill slit at posterior left in the head, equivalent to the left gill slit of *Cephalodiscus*. The tail probably contained a notochord and would have pulled the head rearward across the sea floor, as it probably did in cornutes and mitrates also. By loss of the tail, solutes gave rise to echinoderms.

The Cornuta (Fig. 3*b*) arose from Soluta by loss of the feeding arm. Cornutes had many gill slits at left posterior in the roof of the head.

The Mitrata (Fig. 3*c*) arose from Cornuta by acquiring right gill slits. They had a complicated fishlike brain and nervous system. Right and left gill slits were inside the head. The lateral line of *Mitrocystites* is a special resemblance to vertebrates. By losing their calcite skeleton three times, mitrates gave rise to the extant groups of chordates. The phosphatic skeleton of vertebrates arose after the calcite skeleton had disappeared. *See* ANIMAL EVOLUTION.

R. P. S. Jefferies

Bibliography. A. P. Cripps, A new species of stem-group chordate from the Upper Ordovician of Northern Ireland, *Palaeontology*, 31:1053–1077, 1988; R. P. S. Jefferies, *The Ancestry of the Vertebrates*, British Museum (Natural History), 1986; R. P. S. Jefferies, The solute *Dendrocystoides scoticus* from the Upper Ordovician of Scotland and the shared ancestry of chordates and echinoderms, *Palaeontology*, 1990; R. L. Parsley, Feeding and respiratory strategies in Stylophora, in C. R. C. Paul and A. B. Smith (eds.), *Echinoderm Phylogeny and Evolutionary Biology*, Liverpool Geological Society, 1989; G. M. Philip, Carpoids: Echinoderms or chordates?, *Biol. Rev.*, 54:439–471, 1979; G. Ubaghs, Réflexions sur la nature et la fonction de l'appendice articulé des carpoides Stylophora (Echinodermata), *Ann. Paléontol.* (*Invert.*), 67:33–48, 1981.

Calcite

A mineral composed of calcium carbonate ($CaCO_3$); one of the most common and widespread minerals in the Earth's crust. Calcite may be found in a great variety of sedimentary, metamorphic, and igneous rocks. It is also an important rock-forming mineral and is the sole major constituent in limestones, marbles, and many carbonatites. Calcite in such rocks is the main source of the world's quicklime and hydrated, or slaked, lime. It is also widely used as a metallurgical flux to scavenge siliceous impurities by forming a slag in smelting furnaces. It provides the essential calcium oxide component in common glasses and cement. Limestones and marbles of lower purity may find uses as dimension stone, soil conditioners, industrial acid neutralizers, and aggregate in concrete and road building. Calcite in transparent well-formed crystals is used in certain optical instruments. *See* CARBONATITE; CRYSTAL OPTICS; LIME (INDUSTRY); LIMESTONE; STONE AND STONE PRODUCTS.

Calcite is the most stable polymorph of $CaCO_3$ under the pressure-temperature conditions existing in most of the Earth's crust. It is by far the most common polymorph to form as a result of carbon dioxide (CO_2) in the atmosphere, hydrosphere, and lithosphere, and in the upper mantle reacting with the calcium oxide (CaO) component in magmas, crustal rocks, and the oceans. Although it is commonly of high purity, calcite itself may contain other cations such as manganese, iron, magnesium, cobalt, barium, and strontium, substituting for the calcium in variable amounts. Well-developed crystals of calcite are common in cavities in limestones and basic igneous rocks. Of special interest are the large crystals of optical quality obtained from near the Eskefiord, Iceland. This material is known as Iceland spar, and single crystals several feet across have been reported. In the United States very large crystals have been mined in Taos County, New Mexico.

Crystallography. Calcite crystals may exhibit a wide variety of external forms in addition to the most common rhombohedra. Tabular, prismatic, and scalenohedral varieties are not unusual, and twinned crystals are often found (**Fig. 1**). Some fibrous varieties of calcite known as satin spar yield specimens of gem quality. Deposits in streams in calcareous regions and forming stalactites and stalagmites in limestone caves occur as masses with overall rounded forms made up of very small crystals. Travertine and onyx often consist of such material. *See* ONYX; STALACTITES AND STALAGMITES; TRAVERTINE.

Calcite has hexagonal (rhombohedral) symmetry and a structure built up of alternate layers of calcium ions (Ca^{2+}) and carbonate groups (CO_3^{2-}) stacked so that the layers lie perpendicular to the main axis of

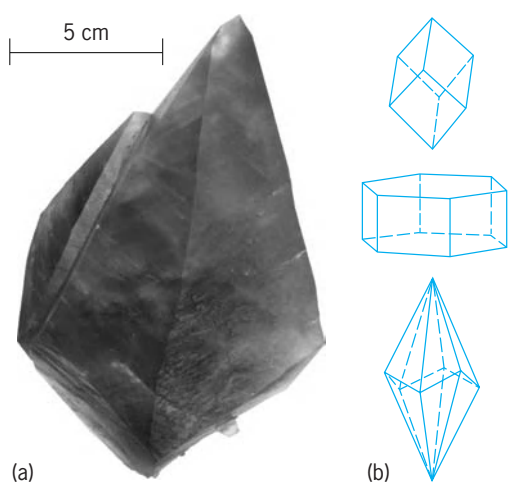

Fig. 1. Calcite. (*a*) Scalenohedral crystal, Joplin, Missouri (*American Museum of Natural History specimens*). (*b*) Common crystal forms (*after C. S. Hurlbut, Jr., Dana's Manual of Mineralogy, 17th ed., John Wiley and Sons, 1959*).

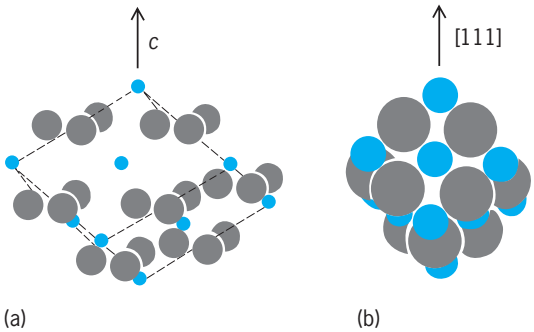

Fig. 2. Structures of (*a*) calcite and (*b*) halite. (*After K. Frye, Modern Mineralogy, Prentice-Hall, 1974*)

symmetry (**Fig. 2**). The structure may be compared with the much simpler sodium chloride structure by aligning the main symmetry axis of calcite with any of the four threefold axes in halite. If the sodium ions (Na^+) and the chloride ions (Cl^-) in the halite model are replaced by Ca^{2+} ions and CO_3^{2-} groups respectively, the selected threefold axis of halite becomes shortened to accommodate the planar CO_3^{2-} groups aligned perpendicular to it. The original halite cube is thus pulled into an obtuse rhomb to give a model of the calcite structure (Fig. 2). *See* CRYSTAL STRUCTURE.

Properties. Pure calcite is either colorless or white, but impurities can introduce a wide variety of colors; blues, pinks, yellow-browns, greens, and grays have all been reported. Hardness is 3 on Mohs scale. The specific gravity of pure calcite is 2.7102 ± 0.0002 at 68°F (20°C).

Calcite has a very low solubility in pure water (less than 0.001% at 77°F or 25°C), but the solubility increases considerably with CO_2 added, as in natural systems from the atmosphere, when more bicarbonate ions and carbonic acid are formed. The solubility is also increased by falling temperature and rising total pressure. Shallow warm seas are supersaturated with calcite, while enormous quantities of calcite are dissolved in the unsaturated deep oceans.

The equilibrium amounts of certain cations that replace calcium in the calcite structure have been determined experimentally for magnesium, iron, and manganese. The amounts of each have been found to increase with temperature. At a fixed temperature the equilibrium amounts increase in the order Mg—Fe—Mn as the size of the substituting cation approaches that of calcium.

The temperature at which calcite dissociates to calcium oxide and carbon dioxide at atmospheric pressure depends on the heating rate, grain size, presence of moisture, and partial pressure of carbon dioxide. Dissociation temperatures range from 1290°F (700°C) to at least 2000°F (1100°C) [at 34 atm or 3.4 megapascals CO_2]. The equilibrium thermal dissociation temperatures have been determined experimentally, and nearly pure calcite has been melted at approximately 2440°F (1340°C) and 1000 atm (100 MPa) of carbon dioxide. Attempts to grow large single crystals of calcite of optical quality by high-temperature and high-pressure techniques have been encouraging, although the dimensions of synthetic crystals do not yet compete with those of natural crystals. *See* HIGH-PRESSURE MINERAL SYNTHESIS.

Partly because it is available in pure, clear, well-developed crystals, calcite has played an important role in the formation of certain fundamental concepts in crystallography and mineralogy. The phenomenon of double refraction was first observed in calcite, and the discovery of the polarization of light resulted from studies of this mineral. Detailed work has shown anisotropism in most of its physical properties. *See* BIREFRINGENCE; POLARIZED LIGHT.

Robert I. Harker

Calcium carbonate relations. Calcium carbonate has a variety of polymorphs, but only calcite, aragonite, and vaterite occur in nature. Vaterite is a rare mineral and occurs as tissues of fractured shells of some gastropods, as gallstones, and as an alteration product after larnite. It probably forms under metastable conditions. The synthetic phase of vaterite has been encountered as high-temperature precipitates of calcium carbonate.

As many as five additional polymorphs not known in nature either have been encountered experimentally or have been presumed to exist. All are apparently closely related to calcite in structure and are not stable at room temperature. **Figure 3** shows the range of pressures and temperatures for which most of these phases occur; the polymorphs $CaCO_3$(II) and $CaCO_3$(III) do not possess pressure-temperature stability fields and are metastable with respect to aragonite. Two additional $CaCO_3$ polymorphs, designated as $CaCO_3$(IV) and $CaCO_3$(V), exist at temperatures above 1470°F (800°C). Phase transformations among these polymorphs appear to be rapid, and the $CaCO_3$(V) decomposes to $CaO + CO_2$ at temperatures above 1800°F (1000°C).

The aragonite–calcite transition has long been of interest to chemists and geologists. The determination of the aragonite stability field at high pressures and temperatures and the positive identification of aragonite in blueschist facies rocks in California have led petrologists to consider the utility of metamorphic aragonite as a geobarometer for blueschist facies

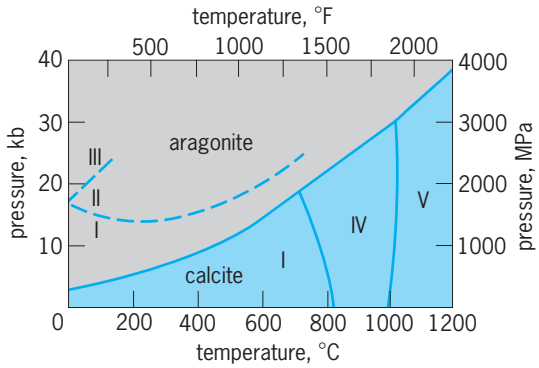

Fig. 3. Phase relations in the unary system $CaCO_3$. Broken lines represent metastable equilibria. Calcite polymorphs are identified as I–V. (*After R. J. Reeder, ed., Reviews in Mineralogy, vol. 11: Carbonates: Mineralogy and Chemistry, Mineralogical Society of America, 1983*)

metamorphism, and much effort has been directed to both phase relations and kinetics for the transition. The pressure-temperature slope of the equilibrium boundary for the aragonite–calcite transition changes between 660 and 930°F (350 and 500°C), as shown in Fig. 3. This change has been attributed to either a phase transition between $CaCO_3$(I) and $CaCO_3$(II) or a gradual disordering of the carbonate (CO_3) groups with increasing temperature. The presence of small amounts of strontium (Sr) in aragonite extends its stability field toward lower pressures; the opposite effect was observed for the presence of magnesium (Mg) in calcite.

Aragonite has been recognized in high-pressure metamorphic rocks such as blueschists. Its rare occurrence is attributed to rapid inversion of aragonite to calcite under shallow crustal conditions. Aragonite cannot survive in rocks in the presence of a fluid phase or in dry rocks subjected to temperatures greater than 570°F (300°C). Partial preservation of aragonite within the calcite stability field requires fast uplift rates (0.004–0.4 in./year or 0.1–10 mm/year) and temperatures of less than 250–370°F (120–190°C) along the uplift pressure-temperature-time path. The occurrence of aragonite may have been overlooked in many blueschist terranes; it does imply depths of 9–19 mi (15–30 km) for its formation. However, because the aragonite can be completely converted to calcite, the absence of aragonite does not limit the rocks to less than these depths. The common aragonite occurrence in Franciscan blueschists in California and its rarity in other blueschist terranes suggests that Franciscan metamorphism reached a lower maximum temperature (for example, 720°F or 380°C), remained dry upon uplifting, or was uplifted much more rapidly than other blueschist terranes. *See* ARAGONITE; BLUESCHIST; CARBONATE MINERALS. J. G. Liou

Bibliography. C. Klein, *Manual of Mineralogy*, revised 22nd ed., 2001; R. J. Reeder (ed.), *Reviews in Mineralogy*, vol. 11: *Carbonates: Mineralogy and Chemistry*, 1983.

Calcium

A chemical element, Ca, of atomic number 20, fifth among elements and third among metals in abundance in the Earth's crust, Calcium compounds make up 3.64% of the Earth's crust. The physical properties of calcium metal are given in the table. The metal is trimorphous and is harder than sodium, but softer than aluminum. Like beryllium and aluminum, but unlike the alkali metals, it will not cause burns on the skin. It is less reactive chemically than the alkali metals and the other alkaline-earth metals. *See* PERIODIC TABLE.

Occurrence of calcium is very widespread; it is found in every major land area of the world. This element is essential to plant and animal life, and is present in bones, teeth, eggshell, coral, and many soils. Calcium chloride is present in sea water to the extent of 0.15%. *See* CARBONATE MINERALS.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Calcium metal is prepared industrially by the electrolysis of molten calcium chloride. Calcium chloride is obtained either by treatment of a carbonate ore with hydrochloric acid or as a waste product from the Solvay carbonate process. The pure metal may be machined in a lathe, threaded, sawed, extruded, drawn into wire, pressed, and hammered into plates. *See* ELECTROMETALLURGY.

In air, calcium forms a thin film of oxide and nitride, which protects it from further attack. At elevated temperatures, it burns in air to form largely the nitride. The commercially produced metal reacts easily with water and acids, yielding hydrogen that contains noticeable amounts of ammonia and hydrocarbons as impurities.

The metal is employed as an alloying agent for aluminum-bearing metal, as an aid in removing bismuth from lead, and as a controller for graphitic carbon in cast iron. It is also used as a deoxidizer in the manufacture of many steels, as a reducing agent in preparation of such metals as chromium, thorium, zirconium, and uranium, and as a separating material for gaseous mixtures of nitrogen and argon.

Calcium oxide, CaO, is made by the thermal decomposition of carbonate minerals in tall kilns using a continuous-feed process. The oxide is used in high-intensity arc lights (lime-lights) because of its unusual spectral features and as an industrial dehydrating agent. The metallurgical industry makes wide use of the oxide during the reduction of ferrous alloys.

Calcium hydroxide, $Ca(OH)_2$, is used in many applications where hydroxide ion is needed. During the slaking process for producing calcium hydroxide, the volume of the slaked lime [$Ca(OH)_2$] produced expands to twice that of quicklime (CaO), and because of this, it can be used for the splitting of rock or wood. Slaked lime is an excellent absorbent for carbon dioxide to produce the very insoluble carbonate. *See* LIME (INDUSTRY).

Calcium silicide, CaSi, an electric-furnace product made from lime, silica, and a carbonaceous reducing agent, is useful as a steel deoxidizer. Calcium carbide, CaC_2, is produced by heating a mixture of lime and carbon to 5432°F (3000°C) in an electric furnace. The compound is an acetylide which yields acetylene upon hydrolysis. Acetylene is the starting material for

Properties of calcium metal

Property	Value
Atomic number	20
Atomic weight	40.08
Isotopes (stable)	40, 42, 43, 44, 46, 48
Atomic volume, cm^3/g-atom	25.9
Crystal form	Face-centered cubic
Valence	2+
Ionic radius, nm	0.099
Electron configuration	2882
Boiling point, °C	1487(?)
Melting point, °C	810(?)
Density, g/cm^3 at 20°C	1.55
Latent heat of vaporization at boiling point, kilojoules/g-atom	399

a great number of chemicals important in the organic chemicals industry.

Pure calcium carbonate exists in two crystalline forms: calcite, the hexagonal form, which possesses the property of birefringence, and aragonite, the rhombohedral form. Naturally occurring carbonates are the most abundant of the calcium minerals. Iceland spar and calcite are essentially pure carbonate forms, whereas marble is a somewhat impure and much more compact variety which, because it may be given a high polish, is much in demand as a construction stone. Although calcium carbonate is quite insoluble in water, it has considerable solubility in water containing dissolved carbon dioxide, because in these solutions it dissolves to form the bicarbonate. This fact accounts for cave formation in which limestone deposits have been leached away by the acidic ground waters. *See* CARBONATE MINERALS.

The halides of calcium include the phosphorescent fluoride, which is the most widely distributed calcium compound and which has important applications in spectroscopy. Calcium chloride has in the anhydrous form important deliquescent properties which make it useful as an industrial drying agent and as a dust quieter on roads. Calcium chloride hypochlorite (bleaching powder) is produced industrially by passing chlorine into slaked lime, and has been used as a bleaching agent and a water purifier. *See* BLEACHING; CHLORINE.

Calcium sulfate dihydrate is the mineral gypsum. It constitutes the major portion of portland cement, and has been used to help reduce soil alkalinity. A hemihydrate of calcium sulfate, produced by heating gypsum at elevated temperatures, is sold under the commercial name plaster of paris. *See* PLASTER OF PARIS.

Calcium is an invariable constituent of all plants because it is essential for their growth. It is contained both as a structural constituent and as a physiological ion. Calcium is found in all animals in the soft tissues, in tissue fluid, and in the skeletal structures. The bones of vertebrates contain calcium as calcium fluoride, as calcium carbonate, and as calcium phosphate. *See* CALCIUM METABOLISM. Reed F. Riley

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; R. Pochet et al. (eds.), *Calcium: The Molecular Basis of Calcium Action in Biology and Medicine*, 2000.

Calcium metabolism

The use and maintenance of calcium levels in the body. Calcium is essential to all forms of life. It plays fundamental roles in cell biology and signaling, in the functioning of organs and organ systems, and in the functioning of organisms. In the most complex organisms, vertebrates, calcium is a key ingredient in the skeleton. In vertebrates, about 99% of total body calcium is found in bones, in the form of a crystalline structure called hydroxylapatite. This calcium not only forms the basis of skeletal rigidity and structure but also serves as a reservoir for calcium that can be mobilized to the blood in response to specific environmental or endocrine signals. When excessive resorption of bone occurs due to inadequate calcium or vitamin D in the diet, diseases result, the symptoms of which may include misshapen skeletal structure and bones which break easily. *See* BONE; CALCIUM.

Concentration and function. The concentration of calcium in the blood is kept within narrow limits, close to 1 *mM* free (unbound) concentration, by a combination of factors including dietary intake, regulated deposition in and resorption from bone, and regulated excretion by the kidneys. This relatively constant extracellular calcium is required for a large variety of functions in different cells of the body. The two major functions of calcium ions are to stabilize membranes and to act as a signal inside cells. The stabilizing functions of extracellular calcium are not well understood, but likely involve interaction with acidic molecules on the cell surface. This surface-bound calcium helps to maintain the relatively low permeability of the plasma membrane to most ions; in nerve and muscle cells, it also helps to maintain control of excitability and the generation of action potentials. *See* BLOOD; CELL PERMEABILITY.

Signaling. Calcium ions also serve as intracellular signals or messengers. To understand the signaling functions of calcium ions, it is necessary first to understand the regulation and metabolism of calcium inside cells. The steady-state concentration of free calcium in the cytoplasm is around 100 *nM*, or 10^{-7} *M*, about one ten-thousandth of the extracellular concentration. This is accomplished by a combination of low plasma membrane permeability to calcium, and active extrusion across the plasma membrane by an adenosine triphosphate (ATP)–consuming calcium pump, the plasma membrane calcium ATPase. In some cells, there is also a sodium-for-calcium exchanger that can facilitate lowering of cytoplasmic calcium. Rapid, transient rises in cytoplasmic calcium can be dampened by a combination of passive and active intracellular calcium buffers. The passive buffers include calcium-binding proteins, and the active buffers are predominantly the endoplasmic reticulum and the mitochondria. The endoplasmic reticulum (and

sarcoplasmic reticulum in muscle cells) accumulates calcium by means of a calcium pump similar in function to the plasma membrane calcium ATPase. This pump is called the sarcoplasmic–endoplasmic reticulum calcium ATPase, or SERCA. Mitochondria take up calcium from the cytoplasm by a process that depends on the mitochondrial membrane potential, which is strongly negative on the inside relative to the cytoplasm. This membrane potential is built from a process coupled to the respiratory chain and is utilized to drive the synthesis of ATP. The uptake of positively charged calcium ions short-circuits this membrane potential to some extent, such that calcium accumulation in mitochondria occurs at the expense of ATP production. However, the uptake of calcium into mitochondria serves as a signal to activate key enzymes involved in ATP production, so the net effect of calcium entering the mitochondria is actually to increase cellular ATP. *See* BIOPOTENTIALS AND IONIC CURRENTS; MITOCHONDRIA.

Endoplasmic reticulum. Calcium stored in the endoplasmic reticulum serves two general functions. First, it is necessary for the proper functioning of chaperone proteins, involved in regulation of protein folding and processing. When endoplasmic reticulum calcium is very low, aberrant protein folding and processing sets in motion a stress response that ultimately leads to apoptosis (cell death). Thus, apoptosis can result from either very high cytoplasmic calcium or very low endoplasmic reticulum calcium. The second major function of calcium in the endoplasmic reticulum is to serve as a rapidly mobilizable signal. When calcium rises in the cytoplasm of cells, it can bind to and regulate a variety of calcium-dependent enzymes, leading to a wide variety of cellular responses (see below). *See* ENDOPLASMIC RETICULUM; PROTEIN FOLDING.

The signal for release of calcium from the endoplasmic reticulum to the cytoplasm is conveyed inward from the cell's plasma membrane by a variety of mechanisms. In skeletal muscle, proteins in the plasma membrane interact with release channels, called ryanodine receptors, in the closely apposed sarcoplasmic reticulum membrane. Depolarization of the plasma membrane causes a conformational change in the plasma membrane proteins that is conveyed by direct protein–protein interaction to the ryanodine receptors, causing their activation and the gated release of stored calcium from the sarcoplasmic reticulum. The released calcium interacts with contractile proteins, causing shortening. This highly specialized system provides an extremely rapid signaling mechanism necessary for rapid muscle movement. In heart, and perhaps in smooth muscles and nerves, calcium ions entering the cell through plasma membrane channels (discussed below) interact with and activate closely associated ryanodine receptors, a process known as calcium-induced calcium release (CICR; see **illustration**).

In nonmuscle cells, other signaling mechanisms can induce release of endoplasmic reticulum calcium. Perhaps the most pervasive mechanism involves inositol 1,4,5-trisphosphate. Neurotransmitters, growth factors, and hormones may interact with a specific class of plasma membrane receptors that are coupled to and activate a phospholipase C. The latter in turn cleaves a relatively minor membrane lipid, phosphatidylinositol 4,5-bisphosphate, into two products, diacylglycerol and inositol 1,4,5-trisphosphate (IP_3). Both of these products act as intracellular messengers; IP_3 binds to a specific

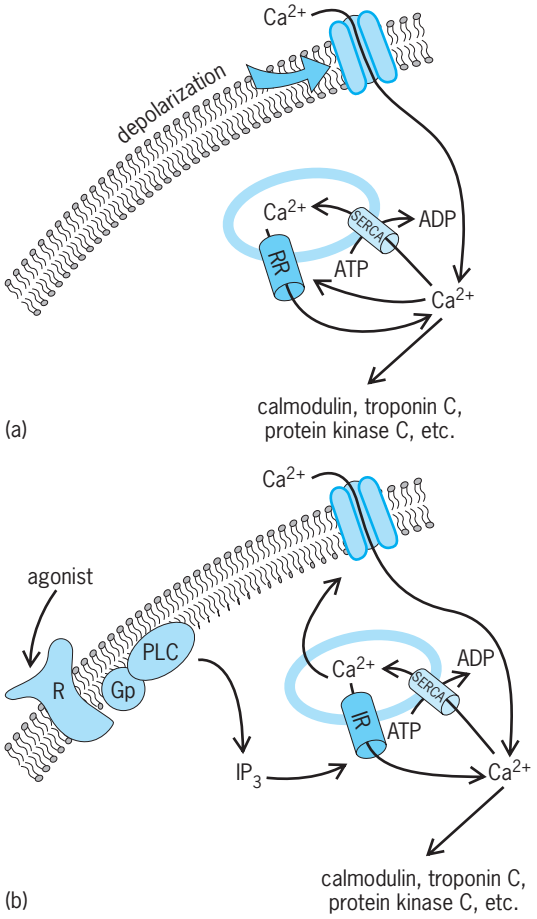

Motifs of $[Ca^{2+}]_i$ signaling. (*a*) In electrically excitable cells such as muscle and nerve cells, extracellular calcium can enter the cells when voltage-dependent calcium channels are activated by the depolarization associated with action potentials. This calcium can cause further release of intracellularly stored calcium by activating a calcium-induced calcium release mechanism associated with the ryanodine receptor calcium channel (RR). (*b*) In electrically nonexcitable cells, signaling is generally initiated when an agonist activates a surface membrane receptor (R) which, usually through a G-protein (guanine nucleotide–dependent regulatory protein, Gp), activates a phospholipase C (PLC) which degrades phosphatidylinositol 4,5-bisphosphate (PIP_2), releasing the soluble messenger, inositol 1,4,5-trisphosphate (IP_3). IP_3 activates an IP_3 receptor (IR) and thus releases calcium from the endoplasmic reticulum to the cytoplasm. The release of calcium from the organelle causes a signal to be generated which activates a plasma membrane calcium entry pathway (capacitative calcium entry). In both *a* and *b* cases, the cytoplasmic calcium signal generated activates various cellular processes by interacting with calcium-binding proteins. SERCA = sarcoplasmic–endoplasmic reticulum calcium ATPase. (*Modified from J. W. Putney, Jr., Science, 262:676–678, 1993*)

receptor on the endoplasmic reticulum. The IP_3 receptor is a calcium-permeable channel, and when IP_3 binds, the channel opens resulting in the release of stored calcium to the cytoplasm (see illustration). The other product of phosphatidylinositol 4,5-bisphosphate breakdown, diacylglycerol, acts in concert with calcium to activate a specific signaling kinase, protein kinase C. *See* CYTOPLASM.

Plasma membrane. In addition to intracellular release of calcium, there are mechanisms for permitting calcium entry to the cytoplasm across the plasma membrane. This entry occurs through calcium channels that are signaled to open in a variety of ways. In electrically excitable cells, such as nerve and muscle, there are calcium channels that open when the membrane is depolarized, as occurs during an action potential. For example, calcium entering heart cells through voltage-gated channels is essential for activation of cardiac muscle contraction (see illustration). In nonexcitable cells, as well as in many excitable cells, there are a variety of calcium-permeable channels activated by either extracellular or intracellular signals. In some cases, the channel is also the receptor for an extracellular ligand, such as acetylcholine or ATP. In other cases, the channels are activated by intracellular messengers, such as diacylglycerol, arachidonic acid, cyclic nucleotides, or even calcium itself. Perhaps the most common mode of regulation of calcium channels in nonexcitable cells involves the capacitative or store-operated channels (see illustration). These channels open when the calcium falls in the endoplasmic reticulum, thus ensuring that persistent depletion of endoplasmic reticulum stores does not occur. Calcium entering through the store-operated channels also activates physiological processes in concert with calcium released from intracellular stores. *See* CELL MEMBRANES; SIGNAL TRANSDUCTION.

The elevation of calcium in the cytoplasm, whether through one of the various mechanisms of intracellular release or through entry across the plasma membrane, signals a variety of cellular processes. These range from very rapid responses, such as muscle contraction and secretion, to more prolonged responses, such as control of glucose metabolism in the liver, to very prolonged responses, such as the control of cell proliferation and differentiation. The response pathways are initiated by calcium binding to and activating specific calcium-binding proteins. In some instances, calcium binds to and directly activates enzymes, for example protein kinase C. In the majority of instances, however, calcium binds to regulatory proteins that in turn regulate the activity of specific enzymes. An example of such a calcium-binding regulatory molecule is troponin, which dissociates from the actin–myosin complex in skeletal muscle when calcium is bound, permitting rapid muscle contraction. Perhaps the most commonly utilized intracellular calcium receptor protein is calmodulin. This protein regulates a wide variety of cellular processes, ranging from smooth muscle contraction to specific gene activation.

Excess cytoplasmic calcium and cell death. In the long term, it is necessary to maintain cellular calcium concentration at low levels because higher concentrations interfere with normal energy metabolism. Prolonged elevation of cytoplasmic calcium is generally deleterious and can eventually lead to cell death. This cell death in some circumstances can occur in a disorderly manner, leading to cell rupture and release of cellular contents. This process is called necrosis, and tends to result in secondary inflammatory responses due to the release of intracellular materials. High calcium can also induce a more ordered, physiological mode of cell death called apoptosis. When apoptosis occurs, the cell structure is dismantled by formation of smaller, membrane-delimited apoptotic bodies which are ingested by lymphocytes.

Overview. The simple chemical nature of divalent calcium ions is deceptive. No other metallic element fulfils such a myriad of significant functions in biological systems. For this reason, the fields of calcium metabolism and calcium signaling continue to be rapidly moving, providing surprises to biologists and benefits to medical science. James W. Putney

Bibliography. E. Carafoli, *Calcium as a Cellular Regulator*, Oxford University Press, 1999; J. W. Putney, Jr., *Calcium Signaling*, CRC Press, Boca Raton, 1999; J. W. Putney, Jr., *Capacitative Calcium Entry*, Landes Biomedical, Austin, 1997.

Calculators

Desktop or, more often, portable electronic devices that are used to perform arithmetic, statistical, or other, more complex processing operations at the step-by-step direction of the user or by execution of a program (a stored sequence of processing operations) selected and initiated by the user. Models are offered with particular stored programs specialized to professional fields such as finance and science, and results are displayed in formats ranging from the still prevalent single line of alphanumeric characters to small viewing screens capable of graphical representations.

Development. Mechanical aides for performing mathematical operations have been available since at least the seventeenth century, when the slide rule was invented. The mechanical adder was introduced in 1887. Calculators in use into the 1960s were large, heavy, noisy desk- or table-top mechanical or electromechanical systems offering columns of numerical keys for selection of an operand (a number to be used in a calculation), and a manually operated lever or a small number of keys for activation of one of the few arithmetic operations of which the system was capable (**Fig. 1**). The operand was set by depression of one of the numeric keys in each of the appropriate columns. Available operand range was therefore established by the number of columns of keys provided. A decimal point location was usually fixed by the calculator, often allowing two columns to the

Fig. 1. Electromechanical calculator manufactured in the 1960s. (*Hewlett-Packard*)

right of the decimal point. Keys remained depressed until execution of the desired arithmetic operation, which was initiated by activation of the lever or crank or, if the unit was motorized, by depression of one of the operation selection keys. Electromechanical calculators offered at most a choice of the four basic arithmetic operations (addition, subtraction, multiplication, and division); many models could only add or subtract.

Improvement in calculator performance was paced by the development of components in the electronics industry. Solenoids and relays were utilized in improved designs, and a transistorized calculator, still capable of only basic arithmetic operations, was introduced in 1964. The invention of the integrated circuit in 1959 and the improvements in the performance offered by integrated circuits and in the technology employed to produce the devices have largely paced the remarkable expansion of calculator capabilities. A prototype calculator constructed of integrated circuits was completed in 1967. The first integrated circuit microprocessor was developed in 1970 for use in a Japanese calculator. The scientific pocket calculator was introduced in 1972, and the programmable scientific calculator in 1974. *See* INTEGRATED CIRCUITS.

Range. The capabilities of the display device employed on a calculator can establish limits to the range of operands and results that the unit can accommodate. Traditionally, electronic calculators have utilized single strings of numeric display devices. If conventional notation is used, and a polarity of the operand or result must be indicated, the allowable range is limited between a positive and a negative number with nines in all digits except the sign position. In a simple example using only whole numbers, eight display positions would provide for display of numbers from −9,999,999 to +9,999,999. Range was expanded with the introduction of scientific notation, which is now prevalent. Scientific notation uses several display positions for what is called an exponent, essentially a power of ten by which the displayed quantity must be multiplied. One position is also allowed for the sign of the exponent, with smaller numbers having less positive or more negative exponents. The quantity −3,256,000 could be represented as −3.256 06 in scientific notation, where the exponent 06 corresponds to ten to the sixth power, or 1,000,000, the number by which −3.2560 must be multiplied. Similarly, the number 0.0007517 could be represented as 7.517 − 04, since ten to the negative fourth power is 0.0001. Depending upon the available display length, two or three digits may be allocated to the exponent. The use of scientific notation greatly expands the range of an operand or calculation result that can be represented by a linear display of a given length.

Advances in components. Current calculator designs use either a customized integrated circuit or a microprocessor and a small number of peripheral integrated circuits to perform their operations. A customized integrated circuit will accept and interpret keystrokes made by the user, perform the requested operation, and generate signals to drive the display device, providing the user an indication of the processing result. Complex mathematical processing abilities are built into some calculator integrated circuits, and some accept plug-in memory modules that extend the available processing sequences. A single line of alphanumeric readout is usually coupled with a customized calculator integrated circuit. Microprocessor-based calculators are generally more complex, more flexible, more capable, and more costly. Some accept memory modules or programs that are selected from a library to meet the particular requirements of an application and downloaded to the unit. The peripheral integrated circuits effect interfacing of the microprocessor with data and program entry and display, readout, and printout devices. *See* MICROPROCESSOR.

Displays have historically consumed the most power in calculator systems and have therefore often limited battery life. To conserve power, many calculators automatically turn their displays off after a fixed period during which no entry has been made by the operator. The large multifilament incandescent readout devices found in early calculator models were supplanted by light-emitting-diode (LED) displays that offered readout of numeric characters, some symbols such as the plus and minus sign, and a limited subset of alphabetic characters. Liquid-crystal-display (LCD) devices, however, have become the predominant calculator display. These devices require only small amounts of power and can provide very fine resolution in a display capable of representing a range of figures and symbols that extends well beyond what could be displayed on the most capable light-emitting-diode readouts. Many calculators utilizing liquid-crystal displays make use of solar cells embedded into the exposed upper surface of the units. The solar cells produce current sufficient for operation of the calculator even at room ambient light levels. *See* ELECTRONIC DISPLAY; LIGHT-EMITTING DIODE; LIQUID CRYSTALS; SOLAR CELL.

Rapid improvement of the density of semiconductor memory has also aided calculator performance.

Large amounts of random-access memory (RAM) can be built into a microprocessor or added to a calculator in the form of a single integrated circuit. Flash memory, which can be programmed and can provide nonvolatile storage between periods of calculator use, yet can be erased and reloaded with new information, facilitates either upgrading of the fixed set of processing functions performed by a model or the loading of the calculator with the specific processing programs needed for a specific application. *See* SEMICONDUCTOR MEMORIES.

Customization. Some calculator models are designed for use in specific professional fields. Most prevalent among these are calculators for business and finance and for science and engineering. A business-finance calculator may provide for determination of true value of money, interest conversions, depreciation, and other customary operations. Selection keys carry labels appropriate to those operations. Customization has more recently taken the form of the downloading of application-specific software, composed of appropriate processing subroutines, into flash memory within the calculator. Individual operations are assigned to specific keys, often function keys, with the display employed to remind the user of function key assignments (**Fig. 2**). The function keys allow the user to select processing programs that have been selected and stored in program memory. Some calculators feature a full qwerty-style keyboard and cursor key (Fig. 2), which represent additional steps away from traditional calculator design and toward the production of a complete pocket computer.

Graphing calculators. The combination of a microprocessor and a liquid-crystal display has made possible the introduction of the graphing calculator. The display is composed of a large rectangular array of liquid-crystal-display pixels (picture elements), each of which can be activated independently. A typical display (Fig. 2) has about one-fourth the pixel display capability of a conventional personal computer monitor screen, packed into a considerably smaller display area. The increased pixel density improves display resolution. This type of array supports the picturelike display of results. For example, the complex voltage waveform representing the calculated response of an electronic circuit to a specific input signal can be displayed for examination. The equation for a complex surface can be entered, and that surface can be displayed. Visual displays of this type are useful to teachers, improving the presentation made available to students, as well as to professionals using the calculator. Engineers and scientists use graphing calculators to enter a sequence of equations in standard form rather than a form dictated by the limitations of the calculator. Matrices can also be represented, and sequences of steps in calculations or in programs stored within the unit can be visually reviewed by recall to the display area.

More capable graphical units can accurately produce representations of complex results such as three-dimensional surfaces (Fig. 2). A number of calculator models support graphing in several coordinate systems (square, rectangular, polar) to improve the user's understanding and interpretation of the result. Graphing calculators are invariably described as being of the science-engineering type, although many of them are, in fact, programmable.

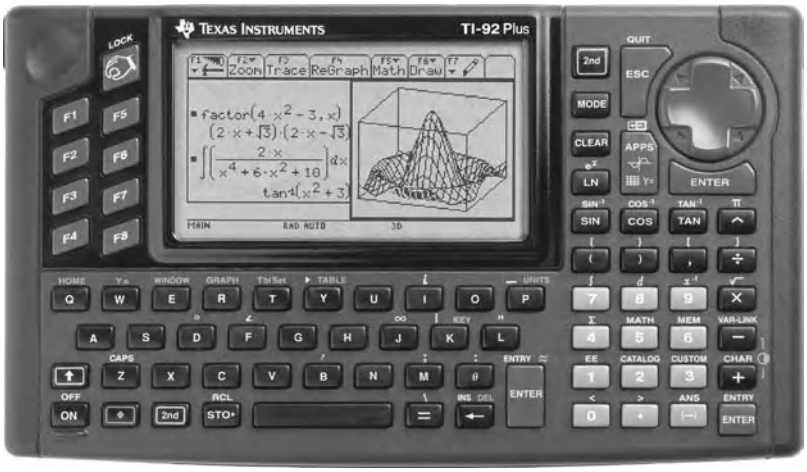

(a)

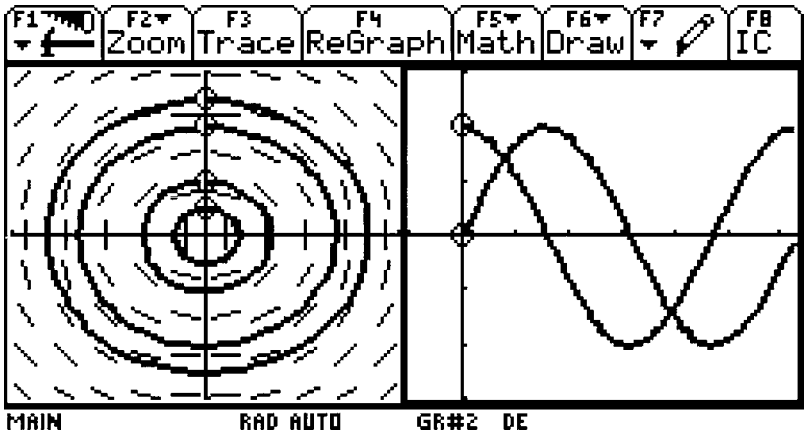

(b)

Fig. 2. Modern programmable electronic calculator. (*a*) Front of calculator, featuring liquid-crystal display, qwerty-style keyboard, cursor key, and function keys. The display is composed of 128 rows of 240 pixels each, and here shows a three-dimensional surface. Function keys allow the user to select processing programs that have been selected and stored in program memory. (*b*) Example of a display, illustrating split-screen capability. (*Texas Instruments Corporation*)

Prospects. The trend in design of the most capable calculators is toward a more general computer-like structure with capabilities set by the processing routines downloaded into program memory rather than by a fixed set committed to permanent memory. The routines employed in the more adaptable approach can be changed as improvements are made or as application requirements change. Connections to peripheral devices such as printers or to computers providing libraries of processing routines are becoming more prevalent. The move away from dedication of a unit to specific processing steps and toward greater dependence on downloaded routines selected to address the problem at hand is the

central development of a broader trend whereby calculators are becoming devices that could more accurately be called pocket computers. *See* COMPUTER; COMPUTER PERIPHERAL DEVICES; DIGITAL COMPUTER; MATHEMATICAL SOFTWARE. William W. Moyer

Bibliography. B. Slozak, *Graphing Calculator Explorations*, Addison-Wesley, 1998; D. E. Varberg and T. D. Varberg, *Algebra and Trigonometry: A Graphing Approach*, Prentice Hall, 1995.

Calculus

The branch of mathematics dealing with two fundamental operations, differentiation and integration, which are carried out on functions. The subject, as traditionally developed in college textbooks, is partly an elementary development of the purely theoretical aspects of these operations and their interrelation, partly a development of rules and formulas for applying calculus to the standard functions which arise in algebra and trigonometry (with exponentials and logarithms included), and partly a collection of applications to problems of geometry, physics, chemistry, engineering, economics, and perhaps a few other subjects.

Derivative. The fundamental concept of differential calculus is that of the derivative of a function of one variable. The classical physical prototype of this concept is that of instantaneous velocity, which is the derivative of distance as a function of time. The derivative also has a highly significant geometrical realization which depends upon the graphical representation of a function in rectangular coordinates (x, y). If y is a differentiable function of x, perhaps as x increases from x_1 to x_2, the graph of the function is a continuous curve with exactly one y for each x, and at each point the curve has a tangent line which is not parallel to the y axis. If ϕ is the angle, measured counterclockwise, from the positive x direction to the tangent (**Fig. 1**), then $\tan \phi$ is equal to the derivative of y with respect to x. (This is on the supposition that the same unit of length is used along the two axes.) This $\tan \phi$ is also called the slope of the curve.

The standard notation for the derivative of y with respect to x is dy/dx. If the functional notation $y = f(x)$ is used, the derivative is often denoted by $f'(x)$. Modern practice is to use f for the function as an abstract entity, while $f(x)$ denotes the value of f at x. Then f' denotes the derivative as a function, and $f'(x)$ is the value of f' at x. For a precise definition of the derivative *see* DIFFERENTIATION

Functions. To say that y is a function of x means that, as a consequence of some rule or definite agreement, there is a designated collection of permissible values of x (called the domain of the function) and an associated set of corresponding values of y (called the range of the function), of such a character that with each permissible value of x is paired a unique well-determined value of y. The function itself is the collection of all the pairs (x, y) which arise in this way. This collection exemplifies the rule or agreement. If f denotes the function, it is customary to write $y = f(x)$ to represent the dependence of y on x.

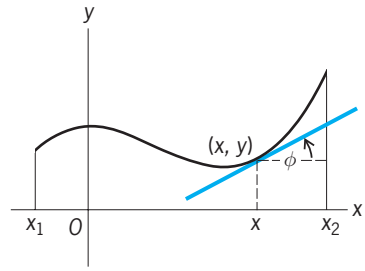

Fig. 1. Graphical representation of the derivative of *f*(*x*).

In calculus the domain of the function is usually composed of one or more intervals of the x axis.

Definite integral. If f is a function defined on the finite interval from x_1 to x_2 inclusive, the definite integral of f from x_1 to x_2, denoted by

$$\int_{x_1}^{x_2} f(x)\, dx$$

is defined by applying to f a rather intricate process which entails the consideration of what are called approximating sums. When the function f is subjected to certain restrictions, this process culminates in the determination of a number as the limit of the approximating sums, and this number is called the definite integral of f from x_1 to x_2. The integral is not defined unless the approximating sums do converge to a well-defined limit. A sufficient condition that this be so is that the function f be continuous.

There is a geometrical representation of the process of defining the definite integral, and it furnishes a plausible argument for the convergence of the approximating sums to a limit. Divide the interval from x_1 to x_2 into a finite number N of not necessarily equal parts. Let the lengths of these parts be $h_1, h_2, \ldots, h_N$, and let t_k be a value of x in the kth part (**Fig. 2**). Then the expression

$$f(t_1)h_1 + f(t_2)h_2 + \cdots + f(t_N)h_N$$

is called an approximating sum. In Fig. 2, where the function is continuous and the function values are all positive, each term $f(t_k)h_k$ in the approximating sum is equal to the area of a certain shaded rectangle, and the whole sum is an approximation of the

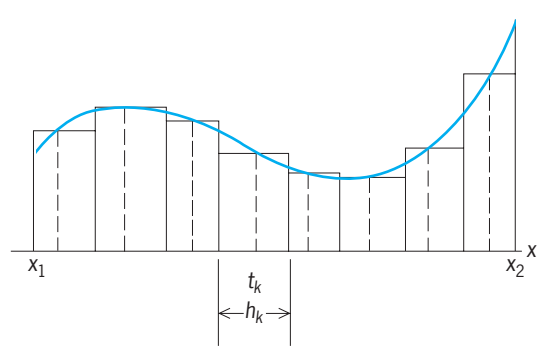

Fig. 2. The definite integral.

area between the graph of the function and the x axis, from x_1 to x_2 inclusive. The limiting process is carried on by increasing N and making the largest of the h_k's approach 0. It is then intuitively clear that the definite integral is the number which represents the exact area between the x axis and the graph. This geometrical interpretation of the integral is the basis of an important application of integral calculus, to the calculation of areas.

It would be tedious and difficult in practice to compute definite integrals by actually working out the limits of approximating sums. It is therefore fortunate that by purely mathematical reasoning it is possible to demonstrate a theorem which links derivatives and integrals and makes it possible, in many important instances, to compute definite integrals by an easier procedure. *See* INTEGRATION.

In the next paragraph are stated the two fundamental theorems of calculus. In these statements the adjective "continuous" appears. For a function $y = f(x)$ to be continuous means, roughly, that as x changes gradually, y must either change gradually, or not at all. Absolutely abrupt jumps are forbidden, and so are many more bizarre modes of irregular behavior. For a precise definition of continuity *see* DIFFERENTIATION

Fundamental theorems. For the calculation of

$$\int_{x_1}^{x_2} f(x)dx$$

find, if possible, a function F with continuous derivative f' such that $f'(x) = f(x)$ when $x_1 \leqq x \leqq x_2$. Then Eq. (1) can be written. This is one of the two central

$$\int_{x_1}^{x_2} f(x)dx = F(x_2) - F(x_1) \tag{1}$$

theorems. The other is stated as follows: Suppose f is continuous, and consider the function F defined by Eq. (2). Then F has a derivative given by $f'(x) = f(x)$.

$$F(x) = \int_{x_1}^{x} f(t)\, dt \tag{2}$$

Law of the mean. The most important theorem about derivatives, exclusive of those which involve integrals, is variously called "the law of the mean" and "the mean-value theorem of differential calculus." It is stated in this form: Suppose f is continuous for x from x_1 to x_2 inclusive, and suppose f has a derivative for each x between x_1 and x_2. Then there is some x of this kind for which Eq. (3) holds true. This theorem

$$f'(x) = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \tag{3}$$

enables one to prove that, if $f'(x)$ is always positive, then $f(x)$ is always zero, then $f(x)$ remains constant as x changes. The first of these results is important for applications to the investigations of graphs. The second is virtually indispensable in the proof of the fundamental theorems of calculus stated above, and in establishing the uniqueness of the solutions for certain problems which involve antidifferentiation (the process of surmising what a function is from a knowledge of its derivative). Such problems abound in application of calculus to problems of motion.

Applications of derivatives. If y is a function of x, the derivative dy/dx is interpretable as the rate of change of y with respect to x. If y is distance and x is time, dy/dx is velocity. If y is work done by a force and x is time, dy/dx is power. Many motion problems, such as those arising in mechanics and formulated by Newton's second law, are expressible by posing an equation which a function and its derivative or derivatives must satisfy. These are called differential equations. The study of such equations is an extension and ramification of calculus. Differential equations are important in all branches of physics. They also arise in many problems of engineering and chemistry, for example, in the deformation of columns and beams, in the study of chemical reactions, and in radioactive decay. *See* DIFFERENTIAL EQUATION.

Problems of maximizing or minimizing a function usually involve derivatives. If a differentiable function f reaches a relative maximum or minimum at a point in the interior of an interval of its domain, then $f'(x)$ must be zero there. This helps in finding extreme values of f.

Applications of integrals. Integrals are used to compute areas bounded by curves in a plane. They are also used to compute other geometrical quantities, such as lengths of curves, areas of surfaces, and volumes of solids. For computing volumes and surface areas the natural tools are multiple integrals. Integrals are also used to compute quantities which occur in physics, such as center of mass, moment of inertia, work done by a variable force, and attraction due to gravitation. The concepts of first, second, and higher moments of a function, as used in statistics, are stated in terms of integrals.

Functions of several variables. To each pair (x_1, x_2) in a specified collection of number pairs, let there correspond a certain definite number y. This defines $y = f(x_1, x_2)$ as a function of two variables. Functions of three or more variables may also be considered. Differential calculus is extended to such functions through the study of partial derivatives with respect to the separate variables $x_1, x_2, \ldots$. A concept of total differential is also relevant. For details *see* PARTIAL DIFFERENTIATION

Integral calculus for functions of several variables is developed through the concept of a multiple integral. For instance, in the case of functions of two variables there is a concept of double integral. The theory of the subject deals with how the double integral is expressible in terms of two successive definite integrals of the sort occurring in calculus of functions of one variable. There is also a concept of a line integral, which is a kind of integral along a curve. There are relations between this concept, the concept of total differential, and the concept of double integral.

New developments. Calculus belongs to the branch of mathematics called analysis. One characteristic

of analysis is its concern with infinite limiting processes. Current research in analysis is mostly on a level of development far beyond the elements of calculus. However, the ideas of calculus persist, though in much more generalized and abstract form, in the theory of functions and in functional analysis.

Angus E. Taylor

Bibliography. R. Courant and F. John, *Introduction to Calculus and Analysis*, vols. 1 and 2, 1965, 1974, reprint 1999; W. Kaplan, *Advanced Calculus*, 4th ed., 1991; S. Lang, *A First Course in Calculus*, 5th ed., 1986, reprint 1993; S. L. Salas and E. Hille, *Calculus*: *Single Variable*, 6th ed., 1990; S. K. Stein and A. Barcellos, *Calculus and Analytical Geometry*, 5th ed., 1992; E. W. Swokowski et al., *Calculus*, 6th ed., 1994; A. E. Taylor and R. W. Mann, *Advanced Calculus*, 5th ed., 1992.

Calculus of variations

An extension of the part of differential calculus which deals with maxima and minima of functions of a single variable. The functions of the calculus of variations depend in an essential way upon infinitely many independent variables. Classically these functions are usually integrals whose integrand depends on a function whose specification by any finite number of parameters is impossible. For example, let C be a smooth bounded region of a space of m variables, $x_1, x_2, \ldots, x_m$, let y be any function of some smooth class on C and its boundary into real numbers or into n-tuples of real numbers and taking specified values on the boundary, and let $f(x, y, p)$ be a smooth function of $2m + 1$ variables $x_1, x_2, \ldots, x_m, y, p_1, p_2, \ldots, p_m$. Then the integral, Eq. (1), is a function

$$J = \int \cdots \int_C f(x, y, y_x) \tag{1}$$

on the space of functions y to the real numbers, and this space of functions is infinite dimensional unless excessive restrictions are placed on it. Here y_x denotes the derivatives $\partial y/\partial x$, and throughout this article subscripts will be used to denote derivatives and occasionally where the context is clear to denote particular values.

The calculus of variations studies such functions and their maxima and minima. The limitation of the competing functions is made realistically, and with sufficient restrictions it is possible to arrive at a rewarding theory; these restrictions do not always include the fixed boundary conditions stated above.

Principal applications may be to physical systems involving flexible components or time-dependent orbits; equilibrium positions or orbits may be determined by minimizing energy or action integrals. The problems are of mathematical interest because of intrinsic difficulties (largely related to lack of topological compactness of bounded regions in spaces of infinitely many dimensions) and possibly because more progress with difficult nonlinear problems has been made here than elsewhere. *See* HAMILTON'S PRINCIPLE; LEAST-ACTION PRINCIPLE.

Theoretical basis. The classical theorems of differential calculus which are used and generalized are the following.

1. Necessary conditions for a function f of a single variable x to attain a local minimum at $x = x_0$ are $f'(x_0) = 0$ if the derivative exists at x_0 and if there are neighboring points of definition of f for arguments on each side of x_0; $f'(x_0) \geqq 0$ if the derivative exists at x_0 and if there are neighboring points of definition of f arguments larger than x_0.

2. If $f'(x_0) = 0$, if f is defined at some points neighboring x_0, and if $f'(x_0)$ exists, then a necessary condition for f to attain a local minimum at $x = x_0$ is $f''(x_0) \geqq 0$; conversely the conditions $f'(x_0) = 0$ and $f''(x_0) > 0$ guarantee a local minimum at x_0. *See* DIFFERENTIATION.

One standard technique of the calculus of variations is to derive necessary conditions for minima by restricting variations to a set of admissible variations depending smoothly on a single parameter and then to apply these theorems 1 and 2 from differential calculus.

For example, in the integral above, suppose that the function $f(x, y, p)$ is defined for all (x, y) in some region of $(m + 1)$-space and for all p, and that f and all its first and second partial derivatives are continuous in this region. Suppose that the admissible functions y are restricted to functions which have continuous second derivatives in C and which take on the required boundary values. Let $g(x)$ be such an admissible function lying in the interior of the region over which $f(x, y, p)$ is defined, and let $z(x)$ be any function with continuous second derivatives on C and taking the value 0 at all points on the boundary of C. Then for e neighboring zero, the function $y = g(x) + ez(x)$ is admissible in the problem (it has continuous second derivatives and takes the prescribed values on the boundary). Then $z(x)$ is an admissible variation. If g does indeed afford a local minimum to J and if it is specified that the functions $y = g(x) + ez(x)$ are close to g for e neighboring zero, then propositions 1 and 2 above may be applied to the function $J(e)$ of the single variable e, where $J(e)$ is the value taken by J for y defined in terms of g and z as above. Under the conditions stated, the derivative $J'(0)$ does exist, and it may be computed by differentiating under the integral sign. The result is shown in Eq. (2).

$$J'(0) = \int \cdots \int_C \left[z(x) f_y(x, g, g_x) + \sum_j z_{x,j} f_{p,j}(x, g, g_x) \right] \tag{2}$$

This expression must be zero for any admissible $z(x)$ according to proposition 1. Under the restrictions set here each term of the expression to be summed may be integrated by parts, the integration of the term with factor $z_{x,j}$ being with respect to the single variable x_j; for the formula

$$\int u\, dv = uv - \int v\, du$$

(with boundary values assigned to the term uv) one sets $dv = z_{x,j}dx_j$ and takes the remainder of the term as u. This yields, after due account is taken of the vanishing of z on the boundary, Eq. (3). A fundamental technique of the calculus of variations lies in exploiting the necessity of this last integral being zero independent of the admissible variation $z(x)$. The integral divided by the measure of C is the average over C of z times the expression in brackets, and the requirement that any function of a sufficiently wide class times a given function averages to zero is met only if the given function is itself zero. Care must be taken to assure that this fundamental condition is met—that is, that the admissible variations are sufficiently general to permit invoking this lemma. In this case the bracketed expression must be zero, and the necessary condition becomes a differential equation, the Euler equation. For the function in Eq. (4), so that Eq. (5) holds, the Euler equation is just the Laplace equation (6).

$$J'(0) = \int \cdots \int_C z(x) \times \left[f_y(x, g, g_x) - \sum_j \frac{\partial}{\partial x_j} f_{p,j}(x, g, g_x) \right] \quad (3)$$

$$f(x, y, p) = p_1^2 + p_2^2 \quad (4)$$

$$J = \int \int \left[\left(\frac{\partial y}{\partial x_1} \right)^2 + \left(\frac{\partial y}{\partial x_2} \right)^2 \right] \quad (5)$$

$$\frac{\partial^2 y}{\partial x_1^2} + \frac{\partial^2 y}{\partial x_2^2} = 0 \quad (6)$$

See DIFFERENTIAL EQUATION; INTEGRATION; LAPLACE'S DIFFERENTIAL EQUATION.

Second variations are studied in a similar way, applying proposition 2 above to $J''(0)$. These give rise to conditions of Jacobi for a minimum. Generally, there are a set of conditions including Euler's condition above, a condition of Weierstrass, a condition of Legendre which may be derived from the Weierstrass condition, and the Jacobi condition, all known as necessary conditions for a minimum. These are generally set in terms of inequalities, and when the conditions are strengthened to demand strong inequalities (excluding equality) the conditions become sufficient for many interesting problems.

Multidimensional derivatives. Much of the work on the calculus of variations is devoted to meticulous detail with regard to the number of derivatives assumed to be available for various functions, particularly the competitive admissible functions $y(x)$. If too many derivatives are assumed, minima may not exist; if too few are assumed, the solution might not be sufficiently smooth to be acceptable in the light of the original statement of the problem. In an attempt to use fewer derivatives, different approaches are used depending on the number of independent variables x. For the multidimensional case, an approach by A. Haar leads to extended complications which are not amenable to description here. Haar introduced an additional function which has the effect of replacing the Laplace equation in the simple Dirichlet problem stated above by the Cauchy-Riemann equations. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

Single-integral problems. For the case of single integrals, however, a lemma of du Bois–Reymond is applicable. This lemma states that a function must be a constant if the average of its product with every function of a sufficiently broad class is zero, where the class may be restricted to have average value zero. The proof takes the given function to be $u(x)$ and its average over the interval to be some constant c, and requires that $u(x) - c$ be one of the admissible functions for comparison. This leads to a requirement that $[u(x) - c]^2$ average zero, and this means that $u(x)$ is essentially c.

Additional illustrative material of general results obtainable will now be presented in terms of single integral problems. Here take Eq. (7) and enlarge the

$$J = \int_a^b f(x, y, y')\, dx \quad (7)$$

class of admissible curves somewhat. The precise nature of this enlargement will not be important for the moment, but assume that among the admissible curves $y = g(x)$, $a(0) \leqq x \leqq b(0)$ affords a minimum to J. Assume also that a family of curves $y = z(x, e)$, $a(e) \leqq x \leqq b(e)$ is admissible and that the value of $z(x, e)$ as e tends to zero is $g(x)$. Most particularly, however, it is desired to avoid the restriction that $z_x(x, e) \to g'(x)$, which was implied above by the variation $y = g(x) + ez(x)$. This allows variations more general than the ones above with regard to the variations of the end points of the curve, the limits of integration, and the behavior of the derivatives.

Curves of the type of $z(x, e)$ introduced here may be admitted as neighboring $y(x)$ for e neighboring zero either by using $\max_x |z(x, e) - y(x)|$ (possibly modified by some measure of disparity between the intervals of definition) or $\int [z(x, e) - y(x)]^2$ as a measure of distance between two points in the function space. In either case, $J(e)$ may be defined as the value of the integral J with f evaluated for x, $z(x, e)$, and $z_x(x, e)$, and it is again possible to study $J'(0)$ if it exists. However, it should be noted that there is no reason for expecting $J'(0)$ to exist in this case, for the integrand involves derivatives z_x which may not vary continuously. Actually a simple example of length shows what may happen; if J is a length integral, the z curves may be taken as zigzagging broken lines approaching a straight line segment but always making an angle of perhaps 45° with it. Then $J(e)$ for any e not zero is $\sqrt{2}$ times $J(0)$. There is illustrated here an important property known as lower semicontinuity. Although $J(e)$ is not continuous, its limit as the argument varies is never less than its value for the limiting value of the argument. This property, along with some compactness properties, suffices for all purposes of the calculus of variations, but it must

be noted that it is reasonable to expect semicontinuity but not continuity in problems of the calculus of variations.

The method used before may be used to derive a formula for $J'(0)$; it is Eq. (8), where Eq. (9)

$$J'(0) = f[b_0, g(b_0), z_x(b_0, 0)]b'(0) - f[a_0, g(a_0), z_x(a_0, 0)]a'(0) + \bar{J} \quad (8)$$

$$\bar{J} = \int_{a_0}^{b_0} \{z_e f_y[x, g, z_x(x, 0)] + z_{xe} f_p[x, g, z_x(x, 0)]\} \quad (9)$$

holds. This may be integrated by parts to give either Eq. (10) or Eq. (11), where H is a function of x such that Eq. (12) holds. In all these formulas $z_x(x,0)$ has

$$\bar{J} = \left\{z_e f_p[x, g, z_x]\right\}_{a_0}^{b_0} + \int_{a_0}^{b_0} z_e \left\{f_y - \frac{d}{dx} f_p\right\} \quad (10)$$

$$\bar{J} = \left\{z_e H\right\}_{a_0}^{b_0} + \int_{a_0}^{b_0} z_{xe} \left\{f_p - H\right\} \quad (11)$$

$$H' = f_y[x, g(x), g'(x)] \quad (12)$$

been retained instead of $g'(x)$ to indicate ambiguity in case of discontinuous derivatives.

End-point problems. If variations with fixed end points are admissible near g, it is possible to choose z with $z_e = 0$ at the end points, and so that a' and b' are zero for $e = 0$ and only $\bar{J}$ is left, so that $J'(0) = 0$ implies $\bar{J} = 0$. In the expression above for $\bar{J}$, Eq. (11), the first term is zero since $z_e = 0$ at the end points, the average of z_{xe} over the interval is zero, since $z_e = 0$ at both end points, and the lemma of du Bois–Reymond mentioned above may be applied. This means that the expression in curly brackets in the integrand is constant whether $f_p(x, g')$ is known to be differentiable or not. Since H is obviously differentiable if g and z have first derivatives which are continuous on each of a finite set of closed intervals covering $[a, b]$, then the whole expression is differentiable, and the Euler equation follows under these weakened conditions. Furthermore, under some conditions of regularity, it is demonstrable that the solutions of this Euler equation have two continuous derivatives, so that minimizing functions in a wide class of functions are comfortably smooth.

The same formula may be used to get conditions which must exist in a corner. It is also applicable to problems with variable end conditions. It may be applied to a special variation to get the Weierstrass condition (**Fig. 1**). These are curves with no general properties except that they vary slowly, approaching straightness at the top and conforming with the continuing curve at the bottom.

Here the lower curve $y = g(x)$ is assumed to be the extremal, and at a point $x = x_1$ on this curve a line with slope q is erected. The variations corresponding to e are taken by proceeding along g to x_1, then up the line to a point with abscissa $x_1 + e$, then along a curve of a smooth family rejoining g at a fixed point x_2, and then along g to b. The family between the line and the point of rejoining g is assumed to

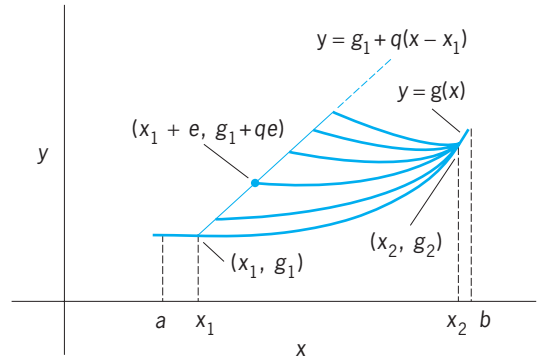

Fig. 1. Weierstrass condition. Symbols explained in text.

be smooth and to approach g with the $z_x(x, e)$ approaching g'. Thus the difficulty about limits of $z_x(x, 0)$ is avoided except at the point $x = x_1$. Note that the variation is defined only for nonnegative values of e, and hence the value $J'(0)$ must be nonnegative if g is a minimum, but not necessarily zero. Intelligent straightforward application of the formulas developed just above yields this condition in the form shown in Eq. (13).

$$E(x, g, g', q) = f(x, g, q) - f(x, g, q') - (q - g')f_p(x, q, g')] \geq 0 \quad (13)$$

In problems with variable end points (which end points may, for example, be constrained to lie on prescribed curves) the end-point conditions which appear in these formulas must be taken into account. Here it is often possible to assume that $\bar{J}$ in the formula is zero, because of admissibility of sufficiently large classes of variations with fixed end points. The remaining conditions are called transversality conditions.

Alternative methods. In classical single-integral problems it is frequently possible to integrate the Euler equations into a family of curves depending on a finite set of parameters (the initial conditions, for example). However, in other cases it is not possible to do this, and a solution must be arrived at or proved to exist in more direct ways. Frequently a minimizing sequence of functions is used; that is, a set of functions y_n is chosen so that the values of J corresponding to y_n approach their greatest lower bound. If there is a convergent subsequence of these y_n and if J itself is lower semicontinuous, then a limit y to the sequence must afford a minimum to J.

Other methods of solution and of proving the existence of solutions depend on general studies of Hilbert spaces. In these studies the integrals appear as operators in Hilbert space, and they are frequently reduced to quadratic forms by some majorizing process.

Problem of Bolza. Here the admissible class of functions is restricted by differential conditions. Problems of Bolza include the classical isoperimetric problems, in which the length of the admissible curves is specified. The problem of Bolza gives rise to multiplier rules which are based on the implicit

function theorem. The theorem is applied by noting that if the rank of a matrix which can be caused to arise is maximal, then the restrictions imposed on admissible curves may be retained unchanged (in particular, or more generally they may be changed arbitrarily) and the function J may still be changed arbitrarily, hence reduced. A submaximal rank for this matrix is a necessary condition for a minimum; using this result the problem of Bolza may be transformed to an ordinary problem whose integral is a linear combination of the integrand of J and functions describing the original restrictions. The coefficients of this linear combination are initially undetermined; they are called Lagrange multipliers.

Critical points. Finally, studies in the spirit of differential calculus but not restricted to local minima should be noted. Here the interest includes not only minima but all critical points (that is, all functions satisfying the Euler condition), and also other studies. It is true that classical problems in mechanics may have stable solutions corresponding to nonminimizing critical points. For sufficiently restricted problems it can be shown that the homology group of the space for which $J \leqq c$ for any c does not change as c varies until a critical level is reached; this is a level of J at which a critical point exists. The change in homology groups at critical points is related closely to the nature of the critical point. For example, in a finite-dimensional space, on a sphere the number of pits minus the number of simple passes plus the number of peaks is two; on a torus the corresponding number is zero; a pass with three grooves running down and three ridges running up from it is to be counted as two simple passes for topological reasons. This is the Morse theory. *See* TOPOLOGY.

Charles B. Tompkins

Multivariable problems. This section deals primarily with the case where the function y is vector-valued. Thus the associated Euler equation is a vector equation, or equivalently a system of real-valued equations.

There are two main cases, one where the approach through the Euler partial differential equation (PDE) has been fruitful, and one where it has been more useful to treat the problem in its original, integral form by rewriting the integral as an integral over the image set rather than the parameter domain. Most of the results discussed are for elliptic problems, but reference is made to some of the work on nonelliptic problems in the calculus of variations.

PDE approach. It will be assumed that the domain C is a bounded open subset of R^m (where R is the real numbers) with a smooth boundary and that the real-valued function f of Eq. (14) is continuous in all

$$f(x, y, p): C \times R^n \times R^{mn} \to R^1 \tag{14}$$

its variables (though somewhat weaker assumptions are possible); as before, $J[y]$ is given by Eq. (15).

$$J[y] = \int_C f(x, y, y_x)dx \tag{15}$$

The key ingredients for the existence of a minimum of J within an admissible class of functions are as follows:

1. The integral J has the property that if a sequence of admissible functions y_i converges to an admissible function y, then $\lim\inf J[y_i]$ is no larger than $J[y]$ (this property is called sequential lower semicontinuity).

2. There exists a minimizing sequence of admissible functions which converges to an admissible function. (Typically this is arranged by requiring the set of admissible functions to have the property that any sequence has some convergent subsequence with limit in the set, which is the property of compactness.)

When these two conditions hold, the direct method can be used to find a minimizer: take any sequence y_i such that $J[y_i]$ converges toward the infimum of the possible values of J, and extract a convergent subsequence.

To have these conditions hold, it is sufficient that f be convex in the last set of variables (so that the Euler equations form a so-called strongly elliptic system of partial differential equations), that either f increases with $|y_x|$ at least as fast as $|y_x|^\alpha$ for some $\alpha > 1$ or there be some known bound given in advance on the slopes $|y_x|$ of a minimizer, and that the class of admissible functions be an appropriate Sobolev space. In particular, the class of admissible functions must allow some kinds of discontinuities, and the first derivatives of y need exist only in a "weak" sense (analogous to Schwartz's generalized functions or distributions).

This leads to the regularity problem: Are the functions obtained by this procedure indeed classical functions with enough derivatives that they are classical solutions to the Euler system of equations?

When $n = 1$ (and under the above assumption of convexity), the answer is yes. But for $n \geqq 2$, the answer is, in general, no. For example, there are integrands f satisfying all the above hypotheses such that the function $y = x/|x|$ (which is discontinuous at the origin) is the unique minimizer among functions with its boundary values.

The problem thus shifts to characterizing the set of points x such that y is not a continuously differentiable function in a neighborhood of x (that is, to characterizing the singular set of y). It has been shown that the singular set has Hausdorff dimension less than $m - 2$; it is an open question in general whether the dimension of the singular set is always less than or equal to $m - 3$. (Fractal sets can have dimensions which are not whole numbers, and the possibility of fractal singular sets has not been ruled out.) *See* FRACTALS.

One particular such case where $y = x/|x|$ is the unique minimizer is with $m = n = 3$, C the unit ball in R^3, $f(x, y, y_x) = y|_x|^2$ (for this f, $J[y]$ is often called the elastic energy of the mapping y), and boundary values $y(x) = x$ for x on the boundary of the unit ball, provided the additional requirement is imposed that the image of y lies on the boundary of the unit ball. Such energy-minimizing mappings where the image is constrained to lie on a unit sphere are typical of the director fields of liquid crystals. The subject of looking for energy-minimizing mappings,

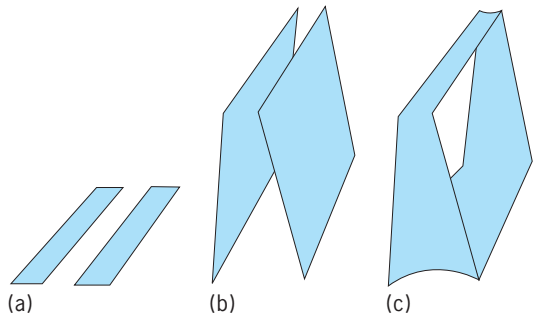

Fig. 2. Example of an area-minimizing surface that cannot be written as the graph of a single function. (*a*) Nonconvex domain *C*. (*b*) Minimal graph over *C* that is not area-minimizing. (*c*) Second surface over *C* that has less area. (*After F. Morgan, Geometric Measure Theory: A Beginner's Guide, Academic Press, 1988*)

or at least for mappings which are critical points of this J (called harmonic maps), is an active field of mathematics research. An example where there is no energy-minimizing mapping arises in the attempt to map a torus (an inner tube) so as to cover a sphere (a beach ball): in the limit of an energy-minimizing sequence, the torus tears apart. *See* LIQUID CRYSTALS.

The major problem with this PDE approach is that many variational problems, when formulated in this way, do not result in a strongly elliptic system of partial differential equations. For example, when a two-dimensional surface in R^3 is written as the graph of a single real-valued function y over a domain C in the plane, its area is given by Eq. (16),

$$J[y] = \int (1 + |y_x|^2)^{1/2}\, dx \tag{16}$$

and if C is convex the above theory can be applied. But if the same surface is written parametrically, as a mapping from a parameter domain in R^2 into three-dimensional space, the area is given by integrating a different integrand, and this integrand is no longer convex in the derivatives. (For mappings of two-dimensional disks, it is possible to chose a particular parametrization so that minimizing area is equivalent to minimizing the elastic energy of the mapping, but for higher-dimensional surfaces, for example, three-dimensional surfaces in R^4, there is no such choice of parametrization or equivalence.) Furthermore, in any dimension a typical area-minimizing surface usually cannot be written as the graph of a single function (**Fig. 2**), and thus (except for special parametrizations of some two-dimensional surfaces) the above theory simply cannot be applied. *See* PARAMETRIC EQUATION.

Geometric variational problems. For functions J such that the value of J depends only on the image surface and not on the particular parametrization chosen (as is the case for area), there is an alternative approach. Integrands F with this property are commonly (and perversely) known as parametric integrands. Since the solution to a variational problem with a parametric integrand is really a surface and not a particular parametrization of it, it has proven fruitful to translate the problem to one of integrating over the image set and not over the parameter domain. This is the point of view adopted in geometric measure theory, a subject whose modern history dates to 1960. Written in this form, the integral has the form given in Eq. (17). Here $\vec{S}(x)$ denotes the tangent plane to the

$$J = \int_{x \in S} F[x, \vec{S}(x)]\, d\sigma \tag{17}$$

surface S at the point x on the surface, and $d\sigma$ denotes integration over the surface (specifically, with respect to Hausdorff m-dimensional measure). For example, F is just constantly 1 for all points and all tangent plane directions when J is the area functional. The integrand F is defined to be elliptic if a piece of a plane always has uniquely (and uniformly, in a certain sense) the least integral, compared to any other surface having the same boundary. If F is elliptic, then when a minimizing surface can be written as the graph of a function, the corresponding system of Euler partial differential equations is strongly elliptic.

If integration over the image set is considered, then it makes sense to speak about minimizing integrals over possible changing surface geometries. For example, surfaces with different numbers of handles or connected components can be compared. A handle is defined topologically by the operation of cutting out two disks from a surface and gluing in a cylinder. The surface in **Fig. 3***a* has no handles and more area than the surface in Fig. 3*b* with one handle, which in turn has more area than the surface in Fig. 3*c* with two handles. In Fig. 3*d*, *e*, and *f*, these three surfaces are deformed to show that they do have zero, one, and two handles, respectively; the cylinders (tubes) are evident in Fig. 3*e* and *f*, compared to Fig. 3*d*. An infinite sequence of handle additions results in a surface with infinitely many handles that is area-minimizing.

In this context, previously intractable problems, such as that of seeking area-minimizing surfaces with prescribed boundaries which are not graphs of functions, become elliptic and amenable to analysis. Some nonelliptic problems can also be handled. More elaborate problems can also be considered,

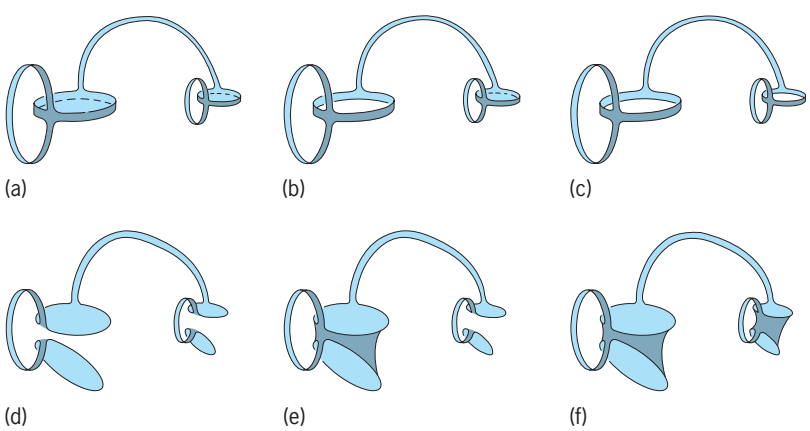

Fig. 3. Comparison of surfaces. (*a*) Surface with no handles. (*b*) Surface with one handle and less area. (*c*) Surface with two handles and still less area. (*d*, *e*, *f*) These three surfaces are deformed to show that they do have zero, one, and two handles, respectively.

such as that of finding the shape of the interface between two substances of prescribed volume filling a given knobby container with tubes running through it. The three types of surface energy per unit area—that between the two fluids as well as that between each fluid and the wall of the container—must be specified, as well as the direction and magnitude of any gravitational attractions. The object here is to minimize the sum of the gravitational energy of the bulk and the integrals of the surface energies over the relevant surfaces.

Many general results on existence, regularity, singularity structure, and computation for geometric variational problems have been developed in this context. The former questions become here questions of the existence of a minimizing surface, the existence of tangent planes to the surface and their smooth variation from point to point, and the possible structures of the surface around points where tangent planes do not exist.

There are four basic classes of surfaces considered in geometric measure theory (playing the role of admissible functions above). These four classes were created in order to solve different types of problems in the calculus of variations.

1. Integral currents are most analogous to mappings; the orientation and multiplicity (number of times a point is hit in the mapping) are included in the definition, and the boundary is defined by Stokes' theorem. An appropriately bounded set of integral currents is compact, and elliptic integrands on integral currents have the lower-semicontinuity property; thus the direct method in the calculus of variations is applicable. The singular set of an m-dimensional area-minimizing integral current is known to have dimension at most $m - 2$ [dimension $m - 7$ if the area-minimizing surface happens to lie in an $(m + 1)$-dimensional space], and there are examples which have singular sets this large. Much less is known about the possible singularities in integral currents minimizing the integral of a general elliptic integrand, except that they can be much larger. *See* STOKES' THEOREM.

2. Flat chains modulo 2 are used for unorientable surfaces like the Möbius band; in other ways, they are similar to integral currents. *See* MANIFOLD (MATHEMATICS).

3. (F,ϵ,δ)-minimal sets are useful in modeling a large variety of complicated physical surfaces, such as soap bubble clusters and soap films or the total grain-boundary structure of polycrystalline materials. Neither orientation nor multiplicity is counted for these surfaces; the defining idea is that of being close to minimizing in small balls. When F is elliptic, singularities can have dimension at most $m - 1$. (The liquid edges in soap bubble clusters form singularities of dimension 1.)

4. Varifolds are quite general (all integrands are continuous on this class of surfaces) and are useful, in particular, in the consideration of infinitesimally corrugated surfaces such as the limit of the sawtooth curves discussed above or certain boundaries of crystalline grains in materials, where the surface energy is so highly anisotropic that the integrand is not elliptic. (When varifoldlike solutions arise in the context of partial differential equations, they are often referred to as Young measures.) *See* MINIMAL SURFACES.

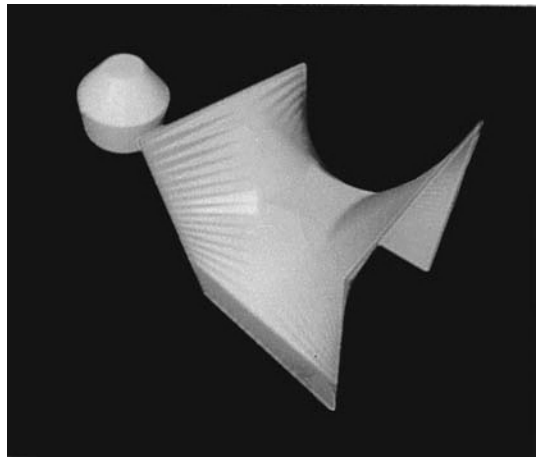

Fig. 4. Examples of surfaces minimizing the integral of a nonelliptic integrand among all bodies having the same volume (the surface at upper left) and among all surfaces having the same boundary (the saddle-shaped surface). Surfaces were computed as part of the Geometry Supercomputer Project sponsored by the National Science Foundation and the University of Minnesota.

Computation. Research in geometric measure theory includes developing and implementing algorithms for the computation and display of surface-energy-minimizing surfaces with a variety of constraints. The integrands (surface energies) that are considered range from the constant 1 (for area-minimizing surfaces) through other elliptic integrands to integrands that are so nonelliptic that the analog of a soap bubble is a polyhedron (**Fig. 4**). A major tool in computing solutions to geometric variational problems is the Surface Evolver, a freeware program. *See* MINIMAL SURFACES.

Additional directions. Optimal design problems, where the objective is to optimize the macroscopic behavior of microscopically heterogeneous materials, can be viewed as a special type of variational problem. Typically, an optimal design problem requires minimization of a sum of energies that is not weakly lower semicontinuous; that is, the total energy of the ordinary limit of a sequence of configurations might well be higher than the limit of the total energies of the configurations themselves. Therefore, special relaxation methods are developed to find and effectively describe new types of limits to highly oscillating configurations; this is analogous to the use of varifolds. The new kinds of limits do have energies that are the limits of the energies of their sequence of oscillatory configurations, and the properties of these new types of limits are called the effective properties of the limit composite material.

Optimal structures are generally unstable. Indeed, they concentrate the resistance capacities in certain directions in order to withstand a given loading. They are thus extremely sensitive to the variation of the loadings. Furthermore, there is a great difference between two-phase and multiphase optimal structures.

Significant advances have been achieved through the use of weak convergence as a rigorous yet general language for the discussion of macroscopic behavior. There is also much interest in novel types of questions, particularly the G-closure problem. (The set of effective properties of all possible composites made from given materials is called the G-closure of the set of original properties, and the problem is to describe the G-closure for different sets of original materials.) Other areas of theoretical progress involve the introduction of new methods for bounding effective properties, and the identification of deep links between the analysis of microstructures and the multidimensional calculus of variations. This work has implications for many physical problems involving optimal design, composite materials, and coherent phase transitions. *See* COMPOSITE MATERIAL; OPTIMIZATION; PHASE TRANSITIONS.

There is at least one variational formulation for all partial differential equations, although some variational formulations are more useful than others. In the past, it has been most useful to pass from calculus of variations problems to partial differential equations, but now it is sometimes better to travel the opposite way. In particular, approximate solutions to partial differential equations involving time can sometimes be constructed via solving minimization problems over discrete time steps, and limits of such approximate solutions as the time step goes to zero can produce solutions to the equations that handle singularities in an appropriate way. Jean E. Taylor

Bibliography. N. I. Akhiezer, *Calculus of Variations*, 1988; U. Brechtken-Menderscheid, *An Introduction to the Calculus of Variations*, 1991; A. Cherkaev and R. Kohn (eds.), *Topics in Mathematical Modelling of Composite Materials*, Birkhauser, 1997; F. Morgan, *Geometric Measure Theory: A Beginner's Guide*, 2d ed., 1995; M. Morse, *Calculus of Variations in the Large*, 8th ed., 1986, reprint 1993; E. R. Pinch, *Optimal Control and the Calculus of Variations*, 1993; F. Y. Wan, *Introduction to the Calculus of Variations and Its Applications*, 1991.

Calculus of vectors

In its simplest form, a vector is a directed line segment. Physical quantities, such as velocity, acceleration, force, and displacement, are vector quantities, or simply vectors, because they can be represented by directed line segments. The algebra of vectors was initiated principally through the works of W. R. Hamilton and H. G. Grassmann in the middle of the nineteenth century, and brought to the form presented here by the efforts of O. Heaviside and J. W. Gibbs in the late nineteenth century. Vector analysis is a tool of the mathematical physicist, because many physical laws can be expressed in vector form.

Addition of vectors. Two vectors **a** and **b** are added according to the parallelogram law (**Fig. 1**). An equivalent definition is as follows: From the end point of **a**, a vector is constructed parallel to **b**, of the same magnitude and direction as **b**. The vector from the origin of **a** to the end point of **b** yields the vector sum $\mathbf{s} = \mathbf{a} + \mathbf{b}$ (**Fig. 2**). Any number of vectors can be added by this rule.

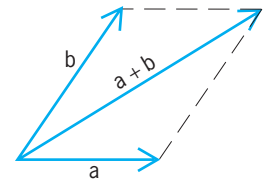

Fig. 1. Addition of two vectors.

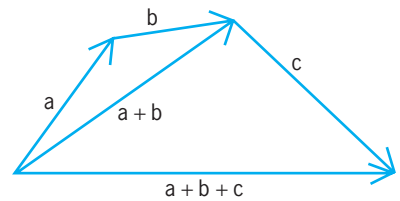

Fig. 2. Addition of three vectors.

Given a vector **a**, a class of vectors can be formed which are parallel to **a** but of different magnitudes. If x is a real number, the vector $x\mathbf{a}$ is defined to be parallel to **a** of magnitude $|x|$ times that of **a**. For $x > 0$, the two vectors **a** and $x\mathbf{a}$ have the same sense of direction, whereas for $x < 0$ the vector $x\mathbf{a}$ is in a reverse direction from that of **a**. The vector $-\mathbf{a}$ is the negative of the vector **a**, such that $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$, with **0** designated as the zero vector (a vector with zero magnitude). Subtraction of two vectors is defined by

$$\mathbf{a} - \mathbf{b} = \mathbf{a} + (-\mathbf{b})$$

The rules shown in notation (1), which conform

$$\begin{aligned}
\mathbf{a} + \mathbf{b} &= \mathbf{b} + \mathbf{a} \\
(\mathbf{a} + \mathbf{b}) + \mathbf{c} &= \mathbf{a} + (\mathbf{b} + \mathbf{c}) \\
x(\mathbf{a} + \mathbf{b}) &= x\mathbf{a} + x\mathbf{b} \\
x(y\mathbf{a}) &= (xy)\mathbf{a} \\
(x + y)\mathbf{a} &= x\mathbf{a} + y\mathbf{a} \\
0 \cdot \mathbf{a} &= \mathbf{0} \\
\mathbf{a} + \mathbf{0} &= \mathbf{a} \\
\mathbf{a} + \mathbf{b} &= \mathbf{a} + \mathbf{c} \text{ implies } \mathbf{b} = \mathbf{c} \\
\mathbf{a} = \mathbf{c}, \mathbf{b} &= \mathbf{d} \text{ implies } \mathbf{a} + \mathbf{b} = \mathbf{c} + \mathbf{d} \\
|\mathbf{a} + \mathbf{b}| \le |\mathbf{a}| &+ |\mathbf{b}|, \text{ with } |\mathbf{a}| = \text{magnitude of } \mathbf{a}, \text{ etc.}
\end{aligned} \tag{1}$$

to the rules of elementary arithmetic, can be readily deduced.

Coordinate systems. The cartesian coordinate frame of analytic geometry is very useful for yielding a description of a vector (**Fig. 3**). The unit vectors **i**, **j**, **k** lie parallel to the positive x, y, and z axes, respectively. Any vector can be written as a linear combination of **i**, **j**, **k**. From Fig. 3, it is noted that Eq. (2) holds. Furthermore, the scalars a_x, a_y, a_z are

$$\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k} \tag{2}$$

simply the projections of **a** on the x, y, and z axes, respectively, and are designated as the components of **a**. Thus, a_x is the x component of **a**, and so on.

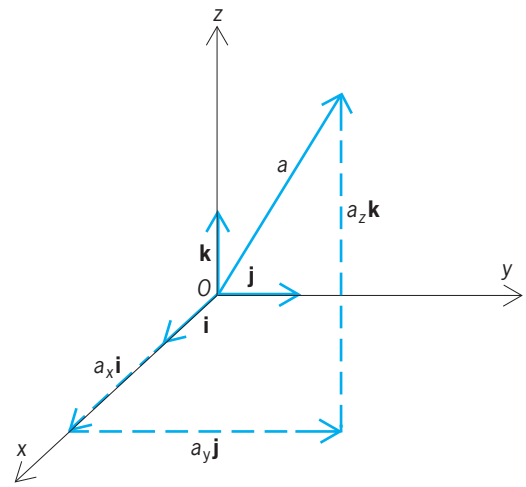

Fig. 3. Vectors in cartesian coordinate system.

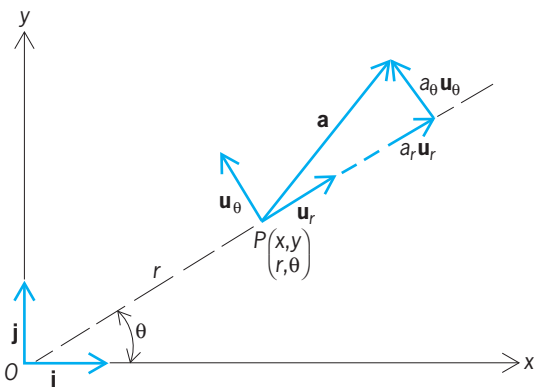

Fig. 4. Vectors in polar coordinate system.

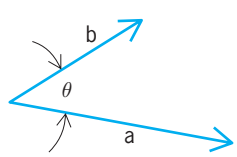

Fig. 5. Scalar product of two vectors.

If the vector **b** is described by $\mathbf{b} = b_x\mathbf{i} + b_y\mathbf{j} + b_z\mathbf{k}$, then Eq. (3) holds.

$$\alpha\mathbf{a} + \beta\mathbf{b} = (\alpha a_x + \beta b_x)\mathbf{i} + (\alpha a_y + \beta y_y)\mathbf{j} + (\alpha a_z + \beta b_z)\mathbf{k} \quad (3)$$

In general, the components of a vector will be functions of the space coordinates and the time. To be more specific, consider a fluid in motion. At any time t the particle which is located at the point $P(x,y,z)$ will have velocity components which depend on the coordinates x,y,z, as well as the time t. Thus, the velocity field **v** of the fluid is represented by Eq. (4). A steady-state vector field exists if the

$$\mathbf{v} = v_x(x, y, z, t)\mathbf{i} + v_y(x, y, z, t)\mathbf{j} + v_z(x, y, z, t)\mathbf{k} \quad (4)$$

components are time-independent. The force field of a fixed gravitating particle is of this type.

It is not necessary to describe a vector in terms of rectangular coordinates (**Fig. 4**). Let **a** be a vector in the xy plane with origin at the point $P(x,y)$. The point P can also be described in terms of polar coordinates (r,ϕ). Let $\mathbf{u}_r$, $\mathbf{u}_\phi$ be unit vectors in the directions of increasing r and θ, respectively. From Fig. 4 it follows that Eq. (5) holds

$$\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j} = a_r\mathbf{u}_r + a_\theta\mathbf{u}_\theta \quad (5)$$

with

$$a_r = a_x\cos\theta + a_y\sin\theta$$
$$a_\theta = -a_x\sin\theta + a_y\cos\theta$$

The components a_r, a_θ yield a description of the same vector **a**. Thus coordinate systems are simply a means of describing a vector. The vector is independent of the description.

Scalar or dot product of two vectors. From two vectors **a** and **b**, a scalar quantity is formed from the definition in Eq. (6), where θ is the angle between

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}|\cdot|\mathbf{b}|\cos\theta \quad (6)$$

the two vectors when drawn from a common origin (**Fig. 5**). It is quickly verified that Eqs. (7) hold. Here

$$\begin{aligned} \mathbf{a}\cdot\mathbf{b} &= \mathbf{b}\cdot\mathbf{a} \\ \mathbf{a}\cdot\mathbf{a} &= |\mathbf{a}|^2 \\ (\mathbf{a}+\mathbf{b})\cdot(\mathbf{c}+\mathbf{d}) &= \mathbf{a}\cdot\mathbf{c} + \mathbf{a}\cdot\mathbf{d} + \mathbf{b}\cdot\mathbf{c} + \mathbf{b}\cdot\mathbf{d} \quad (7) \\ \mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} &= \mathbf{k}\cdot\mathbf{k} = 1 \\ \mathbf{i}\cdot\mathbf{j} = \mathbf{j}\cdot\mathbf{k} &= \mathbf{k}\cdot\mathbf{i} = 0 \end{aligned}$$

$\mathbf{a} \perp \mathbf{b}$ implies $\mathbf{a}\cdot\mathbf{b} = 0$, and conversely, provided $|\mathbf{a}|\cdot|\mathbf{b}| \neq 0$.

For $\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k}$, $\mathbf{b} = b_x\mathbf{i} + b_y\mathbf{j} + b_z\mathbf{k}$, it follows from Eq. (7) that Eq. (8) can be written. If

$$\mathbf{a}\cdot\mathbf{b} = a_xb_x + a_yb_y + a_zb_z \quad (8)$$

a is a force field displaced along the vector **b**, then $\mathbf{a}\cdot\mathbf{b}$ represents the work performed by this force field.

Referring to Eq. (5), let $\mathbf{b} = b_x\mathbf{i} + b_y\mathbf{j} = b_r\mathbf{u}_r + b_\theta\mathbf{u}_\theta$. Then

$$\begin{aligned} \mathbf{a}\cdot\mathbf{b} &= a_xb_x + a_yb_y = a_rb_r + a_\theta b_\theta \\ &\equiv (a_x\cos\theta + a_y\sin\theta)(b_x\cos\theta + b_y\sin\theta) \\ &\quad + (-a_x\sin\theta + a_y\cos\theta) \\ &\quad \times (-b_x\sin\theta + b_y\cos\theta) \end{aligned}$$

Thus the scalar product of two vectors is independent of the descriptive coordinate system as is evident from the definition of the scalar product given by Eq. (6).

Vector or cross product of two vectors. In three-dimensional space, a vector can be formed from two vectors **a** and **b** in the following manner if they are nonparallel. Let **a** and **b** have a common origin defining a plane, and let **c** be that vector perpendicular

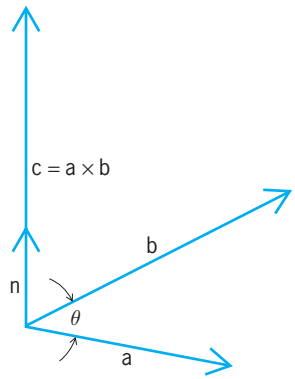

Fig. 6. Vector product of two vectors.

to this plane of magnitude $|\mathbf{c}| = |\mathbf{a}|\ |\mathbf{b}| \sin\theta$. If **a** is rotated into **b** through the angle θ, a right-hand screw will advance in the direction of **c** (**Fig. 6**). Thus Eq. (9) can be written.

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = |\mathbf{a}|\ |\mathbf{b}| \sin\theta\ \mathbf{n} \tag{9}$$

It follows that $\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$, and that if **a** is parallel to **b**, $\mathbf{a} \times \mathbf{b} = \mathbf{0}$. Conversely, if $\mathbf{a} \times \mathbf{b} = \mathbf{0}$, then **a** is parallel to **b** provided $|\mathbf{a}|\ |\mathbf{b}| \neq 0$.

The distributive law can be shown to hold for the vector product so that Eq. (10) holds.

$$(\mathbf{a} + \mathbf{b}) \times (\mathbf{c} + \mathbf{d}) = \mathbf{a} \times \mathbf{c} + \mathbf{a} \times \mathbf{d} + \mathbf{b} \times \mathbf{c} + \mathbf{b} \times \mathbf{d} \tag{10}$$

It follows from

$$\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = 0$$
$$\mathbf{i} \times \mathbf{j} = \mathbf{k}, \mathbf{j} \times \mathbf{k} = \mathbf{i}, \mathbf{k} \times \mathbf{i} = \mathbf{j}$$

that for

$$\mathbf{a} = a_x\mathbf{i} + a_z\mathbf{j} + a_z\mathbf{k} \qquad \mathbf{b} = b_x\mathbf{i} + b_y\mathbf{j} + b_z\mathbf{k}$$

Eq. (11) holds. The expression in Eq. (11) is to be

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix} \tag{11}$$

expanded by the ordinary rules governing determinants.

Multiple products involving vector and scalar products can be generated. The triple scalar product $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ is given by Eq. (12). It can be shown

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix} \tag{12}$$

that Eq. (13) is true. Geometrically, the scalar triple

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} \equiv (\mathbf{abc}) \tag{13}$$

product represents the volume of a parallelepiped formed with **a, b, c** as coterminous sides,

$$\mathbf{V} = |\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})|$$

Of importance in the study of rigid-body motions is the triple vector product shown in Eq. (14). From Eqs. (13) and (14) follow Eqs. (15) and (16).

$$\begin{aligned} \mathbf{a} \times (\mathbf{b} \times \mathbf{c}) &= (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c} \\ (\mathbf{a} \times \mathbf{b}) \times \mathbf{c} &= (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{b} \cdot \mathbf{c})\mathbf{a} \end{aligned} \tag{14}$$

$$\begin{aligned} (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) &= (\mathbf{a} \cdot [\mathbf{b} \times (\mathbf{c} \times \mathbf{d})] \\ &= (\mathbf{a} \cdot \mathbf{c})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d})(\mathbf{b} \cdot \mathbf{c}) \end{aligned} \tag{15}$$

$$\begin{aligned} (\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) &= (\mathbf{acd})\mathbf{b} - (\mathbf{bcd})\mathbf{a} \\ &= (\mathbf{abd})\mathbf{c} - (\mathbf{abc})\mathbf{d} \end{aligned} \tag{16}$$

Pseudovectors. It is easy to verify that a reflection of the space coordinates given by $x' = -x$, $y' = -y$, $z' = -z$, reverses the sign of the components of a vector. Under a space reflection, however, the jcomponents of the vector $\mathbf{a} \times \mathbf{b}$ do not change sign. This is seen from Eq. (11), for if a_x, a_y, a_z are replaced by $-a_x$, $-a_y$, $-a_z$, and if $-b_x$, $-b_y$, $-b_z$ are replaced by $-b_x$, $-b_y$, $-b_z$, the components of $\mathbf{a} \times \mathbf{b}$ remain invariant. Hence, $\mathbf{a} \times \mathbf{b}$ is not a true vector and therefore is given the title pseudovector. In electricity theory the magnetic field vector **B** is a pseudovector, whereas the electric field vector is a true vector provided the electric charge is a true scalar under a reflection of axes. Pseudovectors belong properly to the domain of the tensor calculus. In general, vectors associated with rotations belong to the category of pseudovectors. In particular, the angular velocity vector associated with the motion of a rigid body is a pseudovector. *See* TENSOR ANALYSIS.

Differentiation. There are three differentiation processes that are of conceptual value in the study of vectors; the gradient of a scalar, the divergence of a vector, and the curl of a vector.

Gradient of a scalar. The vector field **v** given by Eq. (17) has the property that the components of

$$\mathbf{v} = v_1(x, y, z)\mathbf{i} + v_2(x, y, z)\mathbf{j} + v_3(x, y, z)\mathbf{k} \tag{17}$$

v differ at individual points $P(x,y,z)$. The differential change in the individual components of v in moving from $P(x,y,z)$ to $Q(x + dx, y + dy, z + dz)$ is given by Eq. (18). It is suggestive to define the differential

$$dv_i = \frac{\partial v_i}{\partial x}dx + \frac{\partial v_i}{\partial y}dy + \frac{\partial v_i}{\partial z}dz \qquad i = 1,\ 2,\ 3 \tag{18}$$

of **v** by Eq. (19), because **i**, **j**, **k** are constant vectors (both in magnitude and direction).

$$d\mathbf{v} = dv_1\mathbf{i} + dv_2\mathbf{j} + dv_3\mathbf{k} \tag{19}$$

If the components of a vector **v** are functions of a single parameter λ, the derivative of v with respect to λ is defined by Eq. (20).

$$\frac{d\mathbf{v}}{d\lambda} = \frac{dv_1}{d\lambda}\mathbf{i} + \frac{dv_2}{d\lambda}\mathbf{j} + \frac{dv_3}{d\lambda}\mathbf{k} \tag{20}$$

The position vector of a particle is designated by $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. If the particle moves along a trajectory, the velocity and acceleration vectors

associated with the motion of the particle are described by Eqs. (21).

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{dx}{dt}\mathbf{i} + \frac{dy}{dt}\mathbf{j} + \frac{dz}{dt}\mathbf{k}$$
$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2} = \frac{d^2x}{dt^2}\mathbf{i} + \frac{d^2y}{dt^2}\mathbf{j} + \frac{d^2z}{dt^2}\mathbf{k} \tag{21}$$

The expression for the derivative of a vector described in curvilinear coordinates becomes more involved simply because the unit vectors in the curvilinear coordinate system need not remain fixed in space. Referring to Eq. (5),

$$\frac{d\mathbf{a}}{dt} = \frac{da_x}{dt}\mathbf{i} + \frac{da_y}{dt}\mathbf{j}$$
$$= \frac{da_r}{dt}\mathbf{u}_r + \frac{da_\theta}{dt}\mathbf{u}_\theta + a_r\frac{d\mathbf{u}_r}{dt} + a_\theta\frac{d\mathbf{u}_\theta}{dt}$$

From $\mathbf{u}_r = \mathbf{i}\cos\theta + \mathbf{j}\sin\theta$ and $\mathbf{u}_\theta = -\mathbf{i}\sin\theta + \mathbf{j}\cos\theta$, it follows that

$$\frac{d\mathbf{a}}{dt} = \left(\frac{da_r}{dt} - a_\theta\frac{d\theta}{dt}\right)\mathbf{u}_r + \left(\frac{da_\theta}{dt} + a_r\frac{d\theta}{dt}\right)\mathbf{u}_\theta$$

The differentiation rules shown in Eqs. (22) exist.

$$\frac{d}{d\lambda}(\mathbf{u}\cdot\mathbf{u}) \cdot \frac{d\mathbf{v}}{d\lambda} + \frac{d\mathbf{u}}{d\lambda}\cdot\mathbf{v}$$
$$\frac{d}{d\lambda}(\mathbf{u}\times\mathbf{v}) = \mathbf{u}\times\frac{d\mathbf{v}}{d\lambda} + \frac{d\mathbf{u}}{d\lambda}\times\mathbf{v} \tag{22}$$
$$\frac{d}{d\lambda}(f\mathbf{v}) = f\frac{d\mathbf{v}}{d\lambda} + \frac{df}{d\lambda}\mathbf{v}$$

From the scalar $\phi(x,y,z)$, one can form the three partial derivatives $\partial\phi/\partial x$, $\partial\phi/\partial y$, $\partial\phi/\partial z$, from which the vector in Eq. (23) can be formed. The vector,

$$\left\{\begin{matrix}\text{gradient of } \phi \\ \text{grad } \phi \\ \text{del } \phi \equiv \nabla\phi\end{matrix}\right\} = \frac{\partial\phi}{\partial x}\mathbf{i} + \frac{\partial\phi}{\partial y}\mathbf{j} + \frac{\partial\phi}{\partial z}\mathbf{k} \tag{23}$$

grad ϕ, has two important properties. Grad ϕ is a vector field normal to the surface $\phi(x,y,z) =$ constant at every point of the surface. Moreover, grad ϕ yields that unique direction such that ϕ increases at its greatest rate. If $T(x,y,z)$ is the temperature at any point in space, then ∇T at any point yields that direction for which the temperature increases most rapidly. If $\phi(x,y,z)$ represents the electrostatic potential, then $\mathbf{E} = -\nabla\phi$ yields the electric field vector.

Divergence and curl of a vector. The del operator defined by Eq. (24) plays an important role in the development of the differential vector calculus. For the vector field of Eq. (17), the divergence of v is defined by Eq. (25). The divergence of a vector is a scalar.

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z} \tag{24}$$

$$\text{div } \mathbf{v} = \nabla\cdot\mathbf{v} = \frac{\partial v_1}{\partial x} + \frac{\partial v_2}{\partial y} + \frac{\partial v_3}{\partial z} \tag{25}$$

If ρ is the density of a fluid and $\mathbf{v}$ the velocity field of the fluid, then div $(\rho\mathbf{v})$ represents the rate of loss of mass of fluid per unit time per unit volume. The total loss of mass of fluid per unit time for a fixed volume V is given by the volume integral of notation (26).

$$\int\int_v\int \text{div } (\rho\mathbf{v})\, dz\, dy\, dx \tag{26}$$

In electricity theory, the divergence of the displacement vector $\mathbf{D}$ is a measure of the charge density, $\nabla\cdot\mathbf{D}$ in the development ρ.

By use of the del operator one obtains quite formally Eq. (27). The vector $\nabla\times\mathbf{v}$ is called the curl

$$\nabla\times\mathbf{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ v_1 & v_2 & v_3 \end{vmatrix}$$
$$= \left(\frac{\partial v_3}{\partial y} - \frac{\partial v_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial v_1}{\partial z} - \frac{\partial v_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y}\right)\mathbf{k} \tag{27}$$

of $\mathbf{v}$ (curl $\mathbf{v}$). Under a reflection of space coordinates, $\partial v_3/\partial y \to \partial(-v_3)/\partial(-y) = \partial v_3/y$ the curl of $\mathbf{v}$ is a pseudovector. As has been noted previously, the class of pseudovectors is associated with rotations in space; thus, it is not strange that the curl of a velocity field is closely associated with an angular velocity vector field. The velocity of a fluid at a point $Q(x + dx, y + dy, z + dz)$ near $P(x,y,z)$ can be characterized as shown in Eq. (28), where $\mathbf{r}$ is the vector from P to Q and Eq. (29) holds. Now in general, $\omega\times\mathbf{r}$ is the

$$\mathbf{v}_Q = \mathbf{v}_P + \frac{1}{2}(\nabla\times\mathbf{v})_P\times\mathbf{r} + \frac{1}{2}\nabla(\mathbf{r}\cdot\mathbf{w}) \tag{28}$$

$$\mathbf{w} = [(\mathbf{r}\cdot\nabla)\mathbf{v}]_P = \left(x\frac{\partial\mathbf{v}}{\partial x} + y\frac{\partial\mathbf{v}}{\partial y} + z\frac{\partial\mathbf{v}}{\partial z}\right)_P \tag{29}$$

velocity of a point due to the angular velocity ω; thus, $\nabla\times\mathbf{v} = 2\omega$. The velocity at Q is simply the sum of a translatory velocity $\mathbf{v}_P$ plus a rigid-body rotational velocity $\omega\times\mathbf{r}$, plus a nonrigid-body deformation, $^1/_2\nabla(\mathbf{r}\cdot\mathbf{w})$.

Formulas involving gradient, divergence, curl.

1. $\nabla(uv) = u\nabla v + v\nabla u$
2. $\nabla\cdot(\phi\mathbf{v}) = \phi\nabla\cdot\mathbf{v} + (\nabla\phi)\cdot\mathbf{v}$
3. $\nabla\times(\nabla\phi) = \mathbf{0}$
4. $\nabla\cdot(\nabla\times\mathbf{v}) = 0$
5. $\nabla\times(\phi\mathbf{v}) = \phi\nabla\times\mathbf{v} + (\nabla\phi)\times\mathbf{v}$
6. $\nabla\cdot(\mathbf{u}\times\mathbf{v}) = (\nabla\times\mathbf{u})\cdot\mathbf{v} - (\nabla\times\mathbf{v})\cdot\mathbf{u}$
7. $\nabla\times(\mathbf{u}\times\mathbf{v}) = (\mathbf{v}\cdot\nabla)\mathbf{u} - \mathbf{v}(\nabla\cdot\mathbf{u}) + \mathbf{u}(\nabla\cdot\mathbf{v}) - (\mathbf{u}\cdot\nabla)\mathbf{v}$
8. $\nabla(\mathbf{u}\cdot\mathbf{v}) = \mathbf{u}\times(\nabla\times\mathbf{v}) + \mathbf{v}\times(\nabla\times\mathbf{u}) + (\mathbf{u}\cdot\nabla)\mathbf{v} + (\mathbf{v}\cdot\nabla)\mathbf{u}$
9. $\nabla\times(\nabla\times\mathbf{v}) = \nabla(\nabla\cdot\mathbf{v}) - \nabla^2\mathbf{v}$
10. $d\phi = d\mathbf{r}\cdot\nabla\phi + \partial\phi/\partial t\, dt$ for $\phi = \phi(x,y,z,t,)$

11. $\nabla \cdot (\nabla\phi) \equiv \nabla^2\phi = \partial^2\phi/\partial k^2 + \partial^2\phi/\partial y^2 = \partial^2\phi/\partial z^2$

12. If $\nabla \cdot \mathbf{f} = 0$, then $\mathbf{f} = \nabla \times \mathbf{A}$ ($\mathbf{A}$ is called the vector potential)

13. If $\nabla \times \mathbf{f} = \mathbf{0}$, then $\mathbf{f} = \nabla\phi$ (ϕ is called the scalar potential)

Integration. If a closed surface is decomposed into a large number of small surfaces, a vector field normal to the surface can be constructed, each normal element being represented by $d\boldsymbol{\sigma}$. The magnitude of $d\boldsymbol{\sigma}$ is the area of the surface element dS, $d\boldsymbol{\sigma} = \mathbf{N}\, dS$.

If $\mathbf{f}$ is a vector field defined at every point of the surface, then notation (30) represents the total flux

$$\int\int_S \mathbf{f} \cdot d\boldsymbol{\sigma} \tag{30}$$

of $\mathbf{f}$ through the surface S. The elements $d\boldsymbol{\sigma}$ point outward from the interior of S.

The divergence theorem of Gauss states that Eq. (31) is true, where R is the region enclosed by S,

$$\int\int_S \mathbf{f} \cdot d\sigma = \int\int\int_R (\nabla \cdot \mathbf{f})\, d\tau \tag{31}$$

$d\tau$ a volume element of R. For $\mathbf{f} = \rho\mathrm{v}$, the divergence theorem states that the net loss of fluid per unit time can be accounted for by measuring the total outward flux of the vector $\rho\mathrm{v}$ through the closed surface S bounding R. *See* GAUSS' THEOREM.

The line integral of a vector field is described as follows: Let Γ be a space curve, and let $\mathbf{t}$ be the unit vector field tangent to Γ at every point of Γ in progressing from A to B, the initial and end points of the trajectory Γ.

The scalar integral in notation (32) is called the

$$\int_A^B (\mathbf{f} \cdot \mathbf{t})\, ds \qquad ((32)$$

line integral of $\mathbf{f}$ along Γ, with arc length s measured along Γ. If $\mathbf{f}$ is a force field, notation (32) represents the work performed by the force field if a unit test particle is taken from A to B along Γ.

The value of the integral of notation (32) will generally depend on the path from A to B. However, if $\mathbf{f} = \nabla\phi$, then

$$\int_A^B (\mathbf{f} \cdot \mathbf{t})\, ds = \int_A^B \nabla\phi \cdot d\mathbf{r} = \int_A^B d\phi = \phi(B) - \phi(A)$$

and the line integral is independent of the path of integration.

The theorem of Stokes states that Eq. (33) holds,

$$\oint_\Gamma \mathbf{f} \cdot d\mathbf{r} = \int\int_S (\nabla \times \mathbf{f}) \cdot d\sigma \tag{33}$$

where Γ is the boundary of the open surface S. If S is a closed surface, then

$$\oiint_S (\nabla \times \mathbf{f}) \cdot d\sigma = 0$$

In electricity theory, a time-changing magnetic flux through an open surface induces a voltage in the boundary of the surface, so that

$$-\frac{\partial}{\partial t}\int\int_S \mathbf{B} \cdot d\sigma = \oint_\Gamma \mathbf{E} \cdot d\mathbf{r}$$

Applying Stokes' theorem yields one of Maxwell's equations,

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

See OPERATOR THEORY; POTENTIALS. Harry Lass

Bibliography. P. Baxandall and H. Liebeck, *Vector Calculus*, 1987; D. Bourne and P. C. Kendall, *Vector Analysis and Cartesian Tensors*, 3d ed., 1992; B. Davis, H. Porta, and J. J. Uhl, *Vector Calculus: Measuring in Two and Three Dimensions*, 1994; H. F. Davis and A. D. Snider, *Introduction to Vector Analysis*, 6th ed., 1991; P. C. DuChateau, *Vector Analysis*, 1993; J. E. Marsden and A. J. Tromba, *Vector Calculus*, 4th ed., 1996; E. C. Young, *Vector and Tensor Analysis*, 1992.

Caldera

A large volcanic collapse depression, typically circular to slightly elongate in shape, the dimensions of which are many times greater than any included vent (**Fig. 1**). Calderas range from a few miles to 45 mi (75 km) in diameter. A caldera may resemble a volcanic crater in form but differs genetically in that it is a collapse rather than a constructional feature. The topographic depression resulting from collapse is commonly widened by slumping of the sides along shallowly rooted faults, so that the topographic caldera wall lies outside the main structural boundary. A caldera may include vents formed by postcollapse volcanism. Its name derives from the Spanish *caldera*, meaning caldron or kettle. As originally defined, caldron referred to volcanic subsidence structures, and caldera referred only to the topographic depression formed by collapse. However, caldera is now common as a synonym for caldron, denoting both topographic and structural features of collapse. *See* PETROLOGY.

Occurrence. Calderas occur primarily in three different volcanic settings, each of which affects their shape and evolution: (1) basaltic shield cones, such as Mauna Loa and Kilauea, Hawaii; (2) stratovolcanoes, for example, Aso and Izu Oshima, Japan; and (3) volcanic centers consisting of preexisting clusters of volcanoes. These last calderas, associated with broad, large-volume andesitic-to-rhyolitic ignimbrite (or ash-flow) sheets, are generally the largest and most impressive. They are the calderas generally implied by the term, and are often referred to as ash-flow calderas. Examples include Toba

(Indonesia) and Yellowstone (Wyoming). Calderas, large and small, have formed throughout much of the Earth's history, and range in age from greater than 1 billion years to historical times. *See* IGNIMBRITE; RHYOLITE; VOLCANO.

In regions of recurrent, shallow magmatic activity, numerous calderas may be partially or entirely superimposed upon each other. For example, Toba caldera is a topographic depression comprising three large, partially superimposed calderas, each of which collapsed during a voluminous ignimbrite eruption (**Fig. 2**). The first (0.8 million years ago) and third (0.074 million years ago) accompanied collapse in the southern part of the composite depression, the second (0.05 million years ago) in the northern part.

In addition to Earth, large calderas occur on Mars, Venus, and Jupiter's moon Io. The presence of calderas on four solar system bodies indicates that the underlying mechanisms of shallow intrusion and caldera collapse are basic processes in planetary geology. *See* MAGMA.

Collapse. Collapse occurs because of withdrawal of magma from an underlying chamber, typically 2.4–3.6 mi (4–6 km) beneath the surface. The result is collapse of the roof into the chamber. Withdrawal of magma may occur either by relatively passive eruption of lavas, as for the small calderas formed on basaltic shield cones, or by catastrophic eruption of pyroclastic (fragmental) material, as accompanies formation of the largest calderas (**Fig. 3**). Eruption of small volumes (no more than 0.24 mi^3 or 1 km^3) of pyroclastic rock, such as has been observed in historical times, does not necessarily involve caldera collapse. Calderas typically accompany pyroclastic eruptions larger than 6–12 mi^3 (25–50 km^3) volume. Eruption of even larger volumes of pyroclastic material appears always to be associated with caldera collapse, because roofs of such large chambers cannot support themselves.

Calderas may form by platelike subsidence of the floor along nearly vertical annular, or ring, faults; by piecemeal collapse whereby different parts subside at different times or rates along multiple faults; by asymmetrical subsidence along incomplete ring faults or simple downsagging with no bounding faults; or by funnel-shaped collapse. From relatively young (<40 million-year-old) calderas, which preserve their broadly circular forms and annular near-surface structures, it has been inferred that large calderas form dominantly by plate subsidence. In contrast, study of older (420–470 million-year-old) and deeply eroded calderas suggests that piecemeal collapse may be a dominant mechanism, but multiple eruption and collapse events complicate their interpretation. Most calderas have transitional features, indicating that multiple mechanisms may apply. Among other factors, the collapse mechanism may depend on the thickness of the roof relative to its diameter, depth to the magma chamber, rate of eruption, and preexisting faults.

Pyroclastic eruption. Caldera-forming eruptions probably last only a few hours or days or, if intermittent, may extend over weeks or longer. Eruption of

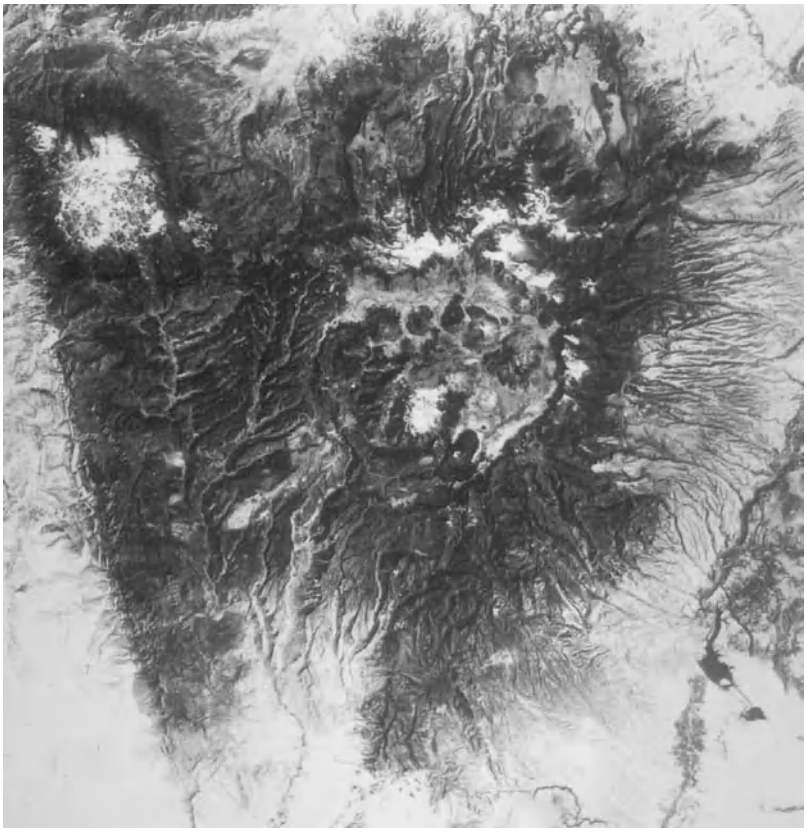

Fig. 1. Landsat Thematic Mapper image of Valles caldera in the Jemez volcanic field of New Mexico. Diameter of the topographic depression is about 13 mi (21 km). (*H. Foote, Battelle—Pacific Northwest Laboratories*)

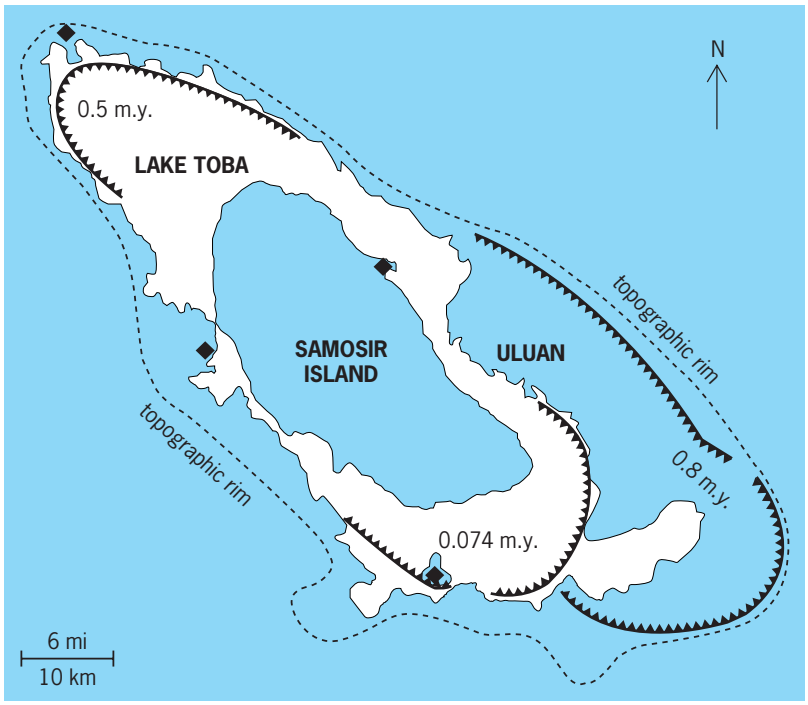

Fig. 2. Toba composite caldera on the Indonesian island of Sumatra comprises several overlapping calderas formed at different times. Spiked lines indicate major faults. Ages of the different caldera segments are indicated by m.y. (million years). Lavas were erupted at four sites (solid boxes) following the eruption 0.074 million years ago.

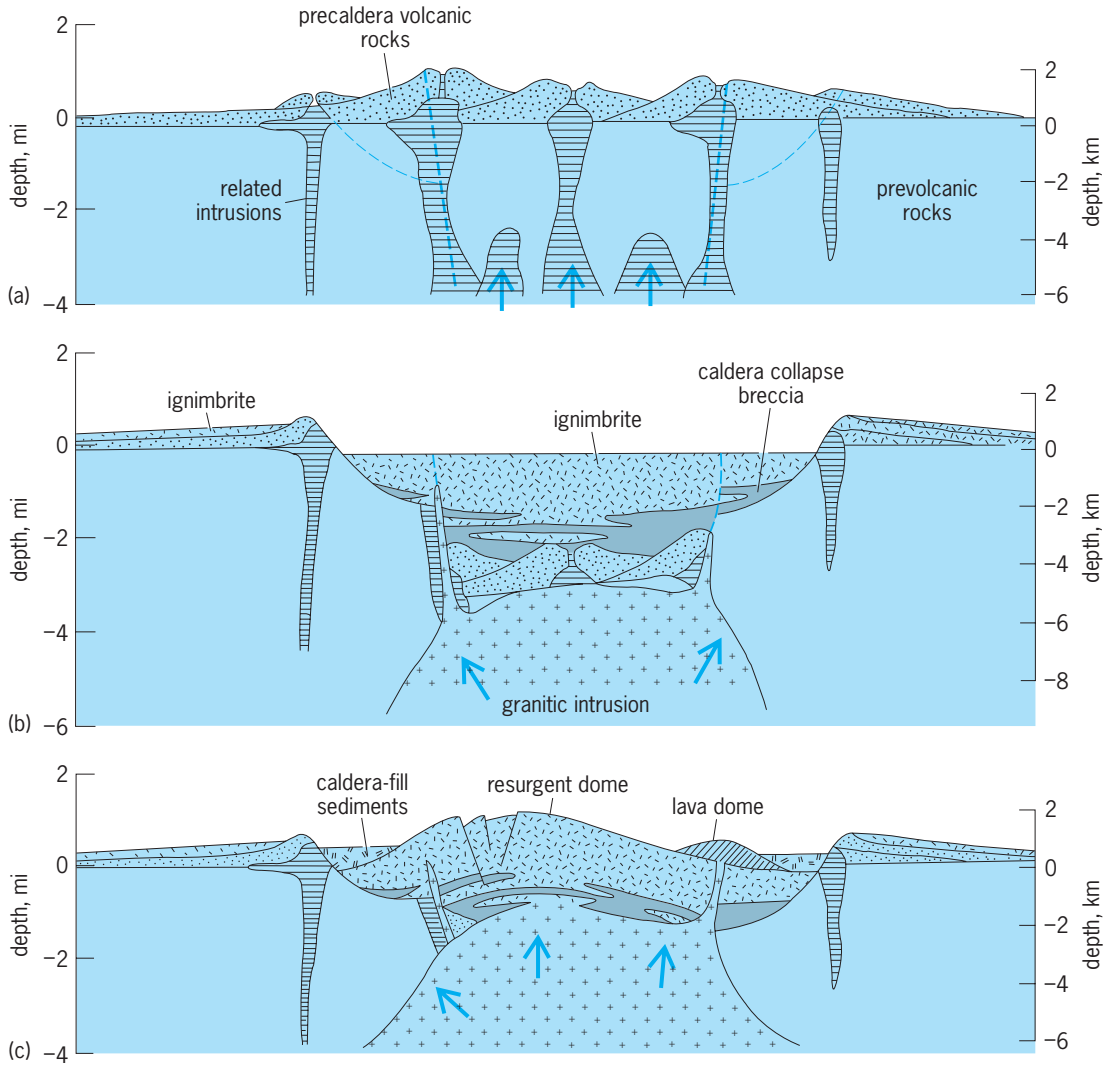

Fig. 3. Generalized evolution of a caldera associated with a pyroclastic eruption. (*a*) Precollapse volcanism. Isolated, small plutons move upward to accumulate into a single large magma chamber. Incipient ring faults are indicated by heavy broken lines, a site of subsequent caldera collapse by light broken lines. (*b*) Caldera geometry just after pyroclastic eruptions and concurrent caldera collapse. Collapse along the ring faults is followed by slumping of the oversteepened walls. (*c*) Resurgence and postcaldera deposition. A magma body has risen and intrudes intracaldera ignimbrites and breccia. Sediments and lava domes partly fill the caldera. (*After P. W. Lipman, The roots of ash flow calderas in western North America: Window into the tops of granitic batholiths, J. Geophys. Res., 89:8801–8841, 1984*)

pyroclastic material begins as gases (predominantly water) dissolved in the magma come out of solution at shallow depths. Magma is explosively fragmented into particles ranging in size from micrometers to meters. An eruption column develops, rising several miles into the atmosphere. This first and most explosive (Plinian) phase of the eruption covers the area around the vent with pumice. The Plinian phase erupts from a single conduit or from a group of closely spaced conduits. Caldera subsidence occurs during eruption, as shown by the thick wedge of ignimbrite (up to 3 mi or 5 km) within calderas (Fig. 3), and probably begins after about 20% of the total volume of associated pyroclastic material has been erupted. The thickness of intracaldera tuff (consolidated pyroclastic material) typically is much greater than that of the corresponding outflow sheets. Evidence for intermittent collapse is provided by landslide deposits, formed by the caving inward of the steep unstable walls, that interfinger with pyroclastic rocks within the caldera, and by the presence of individual cooling units in the tuffs. *See* PYROCLASTIC ROCKS; TUFF.

As caldera subsidence proceeds and eruption becomes less explosive, the Plinian eruption column collapses. This collapse produces pyroclastic flows made of volcanic particles buoyed up and fluidized by hot gases (**Fig. 4**). These hot, ground-hugging flows can travel as far as 93 mi (150 km) outward from the vent at speeds of 330 ft/s (100 m/s). Successive collapses of the column produce multiple flow units with an aggregate thickness that may be several hundreds of feet thick near the caldera. Pyroclastic material may be erupted from multiple vents distributed around the ring fracture; vent geometry may shift during caldera subsidence. The ash and pumice of these ignimbrites are compacted, and individual particles are welded together by the intense heat (932–1112°F or 500–600°C) of the trapped gases.

Fig. 4. Pyroclastic flows from the eruption that accompanied collapse of the Crater Lake caldera in Oregon.

Postcollapse volcanism. Volcanism may resume following caldera collapse and cessation of the massive pyroclastic eruptions. This phase of activity is predominantly passive because the remaining magma is depleted in volatile components. Eruptive activity may begin shortly after caldera collapse and continue intermittently for millions of years. Eruptions may occur from vents within the caldera. Most prominent, however, are the eruptions of lava domes and flows from a series of vents along the ring fractures. These probably mark the intrusion of a ring dike, a feature exposed in some eroded calderas or in subvolcanic igneous centers such as the Mesozoic White Mountains (New Hampshire) and the Permian Oslo graben (Norway). In Valles caldera, New Mexico (Fig. 1), eight domes (up to 2300 ft or 700 m thick) and an obsidian flow were erupted along the main ring fracture from 1 million to about 50,000 years ago. *See* GRABEN; OBSIDIAN.

Resurgence. The floors of many of the largest calderas have been domed upward, resulting in a central massif or resurgent dome. A resurgent dome (Redondo Peak) is visible in the center of the Valles caldera (Fig. 1), and Samosir Island and Uluan are two sides of a single resurgent dome in Toba caldera (Fig. 2). Uplift, which probably occurs intermittently over a period of thousands of years, may exceed 0.6 mi (1 km). During resurgence, the caldera floor is forced upward so that geological units constituting the caldera fill—typically ignimbrite and landslide deposits, and possibly lake beds—have steep dips, often exceeding 45°, away from a central point or axis. Resurgence results from the continued or renewed buoyant rise of magma after collapse. However, resurgence is not always so symmetrical, and instead it may cause broad uplift of adjacent areas of the volcanic field.

Geothermal systems. Calderas typically contain or are associated with extensive hydrothermal systems because of two factors: (1) the shallow magma chambers that underlie them provide a readily available source of heat; and (2) the floors of calderas may be extensively fractured, which condition, along with the main ring faults, allows meteoric water to penetrate deeply into the crust beneath calderas. Hydrothermal activity related to a caldera system can occur any time after magmas rise to shallow crustal levels, but it is dominant late in caldera evolution. In young calderas, such as those of the Yellowstone Plateau and Valles caldera, hydrothermal systems are still active. The spectacular geysers, hot springs, and boiling mud pots of Yellowstone National Park are part of the hydrothermal system associated with a series of three overlapping young calderas.

Magma bodies beneath large calderas cool slowly and may be important heat sources for as long as 2 million years following collapse. The hydrothermal system within Valles caldera (**Fig. 5**) is one of the best-studied systems in the world, largely because of extensive drilling to explore for potentially commercial quantities of geothermal energy. In this system, meteoric water percolates to a depth of 1.2–1.9 mi (2–3 km) in caldera-fill ignimbrites and precaldera volcanic rocks, where it is heated to temperatures of about 572°F (300°C). The heated water then rises convectively to depths of approximately 1640–1968 ft (500–600 m) before flowing laterally southwestward toward the caldera wall. A boiling interface at about 392°F (200°C) marks the upper surface of a convecting liquid-dominated system. Above this interface is a vapor-dominated zone containing steam, carbon dioxide (CO_2), hydrogen sulfide (H_2S), and other volatile components. Acid springs, mud pots, and fumaroles occur in a surface condensation zone a few feet thick. The present hydrothermal system was probably initiated about 1 million years ago, immediately after caldera collapse.

Mineralization. Many metals—including such base and precious metals as molybdenum, copper, lead, zinc, silver, gold, mercury, uranium, tungsten, and antimony—are mobile in hydrothermal circulation systems driven by the shallow intrusions which underlie and give rise to large calderas. Many economically important ore deposits in the western United Sates lie within calderas. However, the relationship of caldera collapse to mineralization is complex. Many mineral deposits are related to precaldera volcanism. Other mineral deposits are related to activity significantly after collapse (in some places by millions of years) and are simply localized by caldera structures. Deep core holes within Valles caldera indicates that mineralization is presently occurring there. *See* ORE AND MINERAL DEPOSITS.

Consequences to humans. Large pyroclastic eruptions associated with such giant (6 mi or 10 km or more in diameter) young calderas as Yellowstone, Long Valley (California), and Valles are, with the

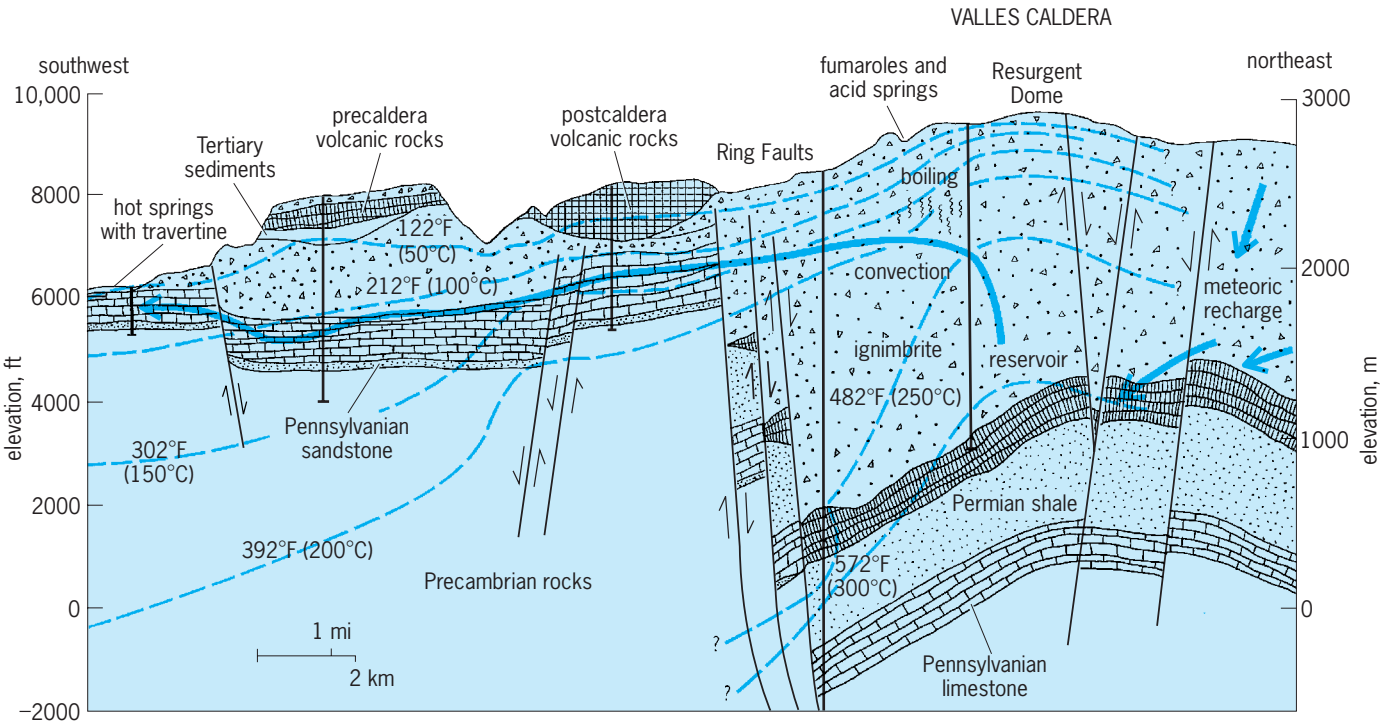

Fig. 5. Hydrothermal system beneath Valles caldera. Heavy vertical lines are drill or core holes. Heavy arrows are possible paths of thermal fluid flow. (*After F. Goff et al., The hydrothermal outflow plume of Valles caldera, New Mexico, and a comparison with other outflow plumes, J. Geophys. Res., 93:6041–6058, 1988*)

exception of impacts of large meteorites or asteroids, the most violent geological phenomena on Earth. In human terms, eruptions of this magnitude are rare, and none are yet documented in historical times.

Formation of several small calderas in the last few thousand years hints at the devastation that such giant caldera-forming eruptions could wreak. The greatest eruption of historical times was that of Tambora (Indonesia) in 1815, which created a caldera about 4 mi (6 km) in diameter. As many as 90,000 people were killed, either directly by the tephra (volcanic rock) fall and pyroclastic flows or by starvation and disease that followed. Ash fallout reached as far as 810 mi (1300 km) from the volcano; ash and sulfurous aerosols injected into the stratosphere are associated with abnormally low summer temperatures and crop failures the following year in the northeastern United States, Britain, and Europe. In approximately A.D. 260, an eruption was accompanied by formation of a caldera 5 × 7 mi (8 × 11 km) at Ilopango, devastating a wide area of central El Salvador. Thick ash flows swept as far as 28 mi (45 km) from the caldera, and ash covered much of El Salvador. Estimates of those killed or permanently displaced range from thousands to hundreds of thousands. The eruption ended Mayan society in the highlands, but migration of refugees stimulated the classic Mayan culture in lowland areas of Central America. In 1883, Krakatau (Indonesia) erupted explosively. The eruption formed a caldera some 5 mi (8 km) in diameter, the floor of which sank over 820 ft (250 m) below sea level. Displacement of the sea as it rushed to fill the void created a series of tsunamis, which killed over 36,000 people in Sumatra and Java. These examples suggest that a truly giant caldera-forming eruption could have catastrophic effect on human life both regionally and globally, with enormous economic and political consequences. *See* TSUNAMI.

W. Scott Baldridge

Bibliography. R. W. Decker and B. B. Decker, *Mountains of Fire: The Nature of Volcanoes*, Cambridge University Press, 1991; P. Francis, Giant volcanic calderas, *Sci. Amer.*, 248(6):60–70, 1983; T. Simkin and R. S. Fiske, *Krakatau 1883: The Volcanic Eruption and Its Effects*, Smithsonian Institution Press, 1983; K. Wohletz and G. Heiken, *Volcanology and Geothermal Energy*, 1989.

Calendar

A system that gives a name to each day. These day names are generally distinct, but not always. For example, the day of the week is a calendar, in a trivial sense, with infinitely many days having the same day name. The design of calendars has been driven by human needs varying over millennia, so the study of calendars is intricately bound to history, astronomy, and religion. Ancient calendars were based on simple observations of phenomena such as the waxing and waning of the Moon, the change in seasons, or the movement of heavenly bodies. Because observational methods suffer from the vagaries of weather and chance, modern calendars tend to be based solely on arithmetical rules, distanced from their motivation in nature.

Most calendars divide a year into an integral number of months and divide months into an integral number of days. However, these astronomical periods—day, month, and year—are incommensurate, so exactly how these time periods are coordinated and the accuracy with which they approximate their astronomical values are what differentiate one calendar from another. *See* DAY; EARTH ROTATION AND ORBITAL MOTION; MONTH; YEAR.

Dozens of calendars are still in use, in addition to the almost universally used Gregorian calendar. Many religious holidays and national events are determined by dates on these calendars. Many tables, algorithms, and rules of thumb for conversion of dates between calendars have been published. Solar calendars—including the Egyptian, Julian, Coptic, Ethiopic, Gregorian, and Persian—are based on the yearly solar cycle, whereas lunar calendars such as the Islamic and lunisolar calendars such as the Hebrew, Hindu, and Chinese take the monthly lunar cycle as the basic building block. Most solar calendars are divided into months, but these months are divorced from the lunar events; they are sometimes related to the movement of the Sun through the twelve signs of the zodiac.

Almost every calendar incorporates a notion of "leap" year to correct the cumulative error caused by approximating a year by an integral number of days and months. Solar calendars add a day every few years to keep up with the astronomical year.

Diurnal calendars. The simplest naming convention would be to assign an integer to each day; fixing day 1 would determine the whole calendar. The Babylonians had such a day count, as did the Maya and the Hindus. Such diurnal calendars are used by astronomers, especially those amateurs and professionals studying variable stars, who use Julian day numbers to specify dates. The Julian period, introduced in 1583 by Joseph Justus Scaliger, was originally a counting of years in a cycle of 7980 years, starting from 4713 B.C.E. (before the common era; or B.C.); nineteenth-century astronomers adapted the system into a strict counting of days backward and forward from JD0 = noon on Monday, January 1, 4713 B.C.E. (Julian) = noon on Monday, November 24, −4713 (Gregorian). A fractional part of a Julian day gives the fraction of a day beyond noon.

Computer scientists often use diurnal calendars as an intermediate device for converting from one calendar to another.

Solar calendars. The Gregorian calendar, now in common use throughout the world, is based on a 12-month year that closely approximates the Earth's solar cycle. This calendar was designed by a commission assembled by Pope Gregory XIII in the sixteenth century; the main author of the new system was the Naples astronomer Aloysius Lilius. This calendar is based on a 365-day common year divided into 12 months of lengths 31, 28, 31, 30, 31, 30, 31, 31, 30, 31, 30, and 31 days, and on 366 days in leap years, the extra day being added to make the second month 29 days long. A year is a leap year if it is divisible by 4 and is not a century year (multiple of 100) or if it is divisible by 400. For example, 1900 is not a leap year; 2000 is a leap year. The Gregorian calendar differs from its predecessor, the Julian calendar, only in that the Julian calendar did not include the century rule for leap years—all century years were leap years.

A simpler solar calendar with a fixed year length of 365 days and no leap year rule was in use in Egypt for millennia prior to the adoption of the Julian calendar in the third century C.E. (common era), and was also used in Babylon and Persia. It served as the canonical calendar for astronomers until the sixteenth century. Each month had 30 days, and at the end of the year was an extra 5-day period, called epagomenae. Though this approximation to the tropical year results in a noticeable shift in the date of onset of the seasons, computationally such a calendar is extremely simple, and it persists in the Armenian calendar.

The Julian calendar was instituted on January 1, 709 *ab urbe condita* (from the traditional founding of Rome; about 45 B.C.E.) by Julius Caesar, with the help of the Alexandrian astronomer Sosigenes. The counting of years according to the Christian era was instituted by the Roman monk Dionysius Exiguus in the sixth century, but only became commonplace a few centuries later; Dionysius erred by a few years in his determination of the year of Jesus's birth. Dionysius did not invent the notion of "B.C."—his system started at 1. The year 1 B.C. is the year before 1 C.E. in the system introduced and popularized by the Venerable Bede around 731; the concept of zero as a number was then unknown in Europe. In this system, the number of years between January 1, x B.C.E., and January 1, y C.E., is $x + y - 1$ (not $x + y$). The use of a year 0 preceding year 1 on the Gregorian calendar is due to Jacques Cassini in 1740. *See* ZERO.

Since every fourth year on the Julian calendar was a leap year, a cycle of 4 years contained $(4 \times 365) + 1 = 1461$ days, giving an average length of year of 365.25 days. This is somewhat more than the mean length of the tropical year, and over the centuries the calendar slipped with respect to the seasons. By the sixteenth century, the date of the spring equinox had shifted from around March 21 to around March 11. If this error were not corrected, eventually Easter, whose date depends on that equinox, would migrate through the whole calendar year. Pope Gregory XIII instituted only a minor change in the calendar: century years not divisible by 400 would no longer be leap years. (He also improved the rules for Easter.) Thus, three out of four century years are common years, giving a cycle of 400 years containing $(400 \times 365) + 97 = 146{,}097$ days and an average year length of $146{,}097/400 = 365.2425$ days. He also corrected the accumulated 10-day error in the calendar by proclaiming that Thursday, October 4, 1582 C.E., the last date in the old-style (Julian) calendar, would be followed by Friday, October 15, 1582, the first day of the new-style (Gregorian) calendar. Catholic countries followed his rule, but Protestant countries resisted. Great Britain and its colonies

(including the future United States) held out until 1752; Russia held out until 1918, after the Bolshevik Revolution. Different parts of what is now the United States changed over at different dates; Alaska, for example, changed only when it was purchased by the United States in 1867. Turkey did not switch to the Gregorian calendar until 1927.

An alternative to arithmetic rules to determine whether a year has 365 or 366 days is to fix the date of the new year according to some astronomical event. The Persian astronomical calendar, for example, uses the vernal (spring) equinox; the short-lived French Revolutionary calendar used the autumnal equinox.

Lunar calendars. The Islamic calendar is an example of a strictly lunar calendar, with no intercalation of months (unlike lunisolar calendars). Its independence of the solar cycle means that its months do not occur in fixed seasons but migrate through the solar year. Virtually all Moslems follow an observation-based calendar in which new moons (and hence months) are proclaimed by officials when that phase is seen. The calendar is computed, by the majority of the Moslem world, starting at sunset of Thursday, July 15, 622 C.E. (Julian), the year of Mohammed's hijra. The introduction of the calendar is often attributed to the Caliph 'Umar, in 639 C.E., but there is evidence that it was in use before his succession. Islamic astronomers developed an arithmetic approximation that is used for estimation. In it, there are 12 Islamic months, which contain, alternately, 30 or 29 days; the twelfth month contains 29 days in an ordinary year and contains 30 days in a leap year—the 2d, 5th, 7th, 10th, 13th, 16th, 18th, 21st, 24th, 26th, and 29th years of a 30-year cycle. This gives an average month of 29.5305555... days and an average year of 354 11/30 days.

Lunisolar calendars. Lunisolar calendars invariably alternate 12- and 13-month years. The so-called Metonic cycle is based on the observation that 19 solar years contain almost exactly 235 lunar months. This correspondence, named after the Athenian astronomer Meton and known much earlier to ancient Babylonian and Chinese astronomers, makes a relatively simple and accurate fixed solar/lunar calendar feasible. The $235 = (12 \times 12) + (7 \times 13)$ months in the cycle are divided into 12 years of 12 months and 7 leap years of 13 months. The 7 leap years are evenly distributed within the 19-year cycle, with gaps of 1 or 2 years between them. The Metonic cycle was used as the basis of lunisolar calendars in Mesopotamia from the fourth century B.C.E, and later in the Seleucid empire. It is (currently) accurate to within 6.5 minutes a year and is still employed in the Hebrew calendar (instituted in 359 C.E.) and for the ecclesiastical calculation of Easter (in both the Gregorian version and the Dionysian formulation of the earlier Nicaean rule). The Hebrew calendar has an average year length of 365.2468 days and month length of $29 + (13{,}753/25{,}920) \sim 29.530594$ days; the Gregorian calculation of Easter results in a mean month length of $29 + (37{,}405{,}943/70{,}499{,}183) \sim 29.530587$ days. Both these calendars may introduce a shift in the start of the year (of a day or two) for technical purposes. (The Hebrew new year, Rosh HaShanah, for example, is not allowed to fall on the alternate days Sunday, Wednesday, and Friday; the calculated Easter full moon is not allowed to fall twice on the same Gregorian date within any one 19-year cycle.) The result is that the Hebrew calendar repeats only after 689,472 years and the ecclesiastical calendar after 5,700,000 years.

The 7-out-of-19 rule of the Metonic cycle is only one of many possibilities. Some of the older Hindu lunisolar calendars evenly distribute 66,389 leap months over 360,000 years.

Alternatively, a lunisolar calendar can follow astronomically determined patterns, as in the Chinese and Hindu, in which each month begins with a new moon, and a thirteenth month is added whenever there are 13 new moons within a solar year. The Chinese calendar has undergone numerous reforms, but since 1645 C.E. it has used astronomical calculations to fix the start of years and months. The current Hindu lunisolar calendar, based on the algorithm of the *Surya Siddhanta* (circa 1000 C.E.), employs Ptolemaic methods. The latter is distinguished by the occasional occurrence of expunged months and days. Nachum Dershowitz; Edward M. Reingold

Bibliography. L. E. Doggett, Calendars, *Explanatory Supplement to the Astronomical Almanac*, ed. by P. K. Seidelmann, University Science Books, Mill Valley, CA, 1992; J. Meeus, *Astronomical Algorithms*, 2d ed., Willmann-Bell, Richmond, VA, 1998; E. M. Reingold and N. Dershowitz, *Calendrical Calculations: The Millennium Edition*, Cambridge University Press, Cambridge, 2000; E. G. Richards, *Mapping Time: The Calendar and Its History*, Oxford University Press, Oxford, 1998.

Calibration

The process of determining the performance parameters of an artifact, instrument, or system by comparing its performance with measurement standards. Adjustment may be part of calibration but not necessarily. It is essential that there be an unbroken traceability chain between the measurement standards being used and the appropriate national measurement standards.

All calibrations involve one or more measurement processes, which may be simple or complex. A calibration can be used to ensure that a device or system will produce results which meet or exceed some defined criteria with a specified degree of confidence. The measurement standards, the process, and the device or system being calibrated must be capable of satisfying these defined criteria.

Two concepts related to calibration are precision and accuracy. Precision refers to the minimum discernible change in the parameter being measured, while accuracy refers to the uncertainty in the measurement process.

All measurement processes used for calibration are subject to various sources of uncertainty. It is now common practice to classify them as Type A

or Type B uncertainties. These approximate to the previous classification into random and systematic uncertainties, respectively. Care must be taken in classifying uncertainties. For example, a "random" component of uncertainty in one measurement may become a "systematic" component of uncertainty in another measurement in which the result of the first measurement is used as an input datum. Type A uncertainties can be estimated as a result of repeated measurements and are amenable to treatment by statistical methods. A Type B uncertainty is obtained by careful consideration of the measurement process, for example, from an estimate of the possible error in the measurement due to the effect of an influence quantity such as temperature. If the error can be calculated reliably, a correction can be made to the result. Otherwise, an estimate is made of the possible uncertainty due to this effect, and it can be incorporated into the uncertainty calculation as a Type B contribution. The two types of uncertainty can be combined by approved methods to give the total uncertainty of the measurement.

A document published by the International Organization for Standardization (ISO), *Guide to the Expression of Uncertainty in Measurement* (popularly known as *GUM*), gives the currently accepted approach to the estimation of uncertainties. Other documents give slightly simplified approaches that are consistent with the ISO standard; most of these can be downloaded from Web sites. *See* ANALYSIS OF VARIANCE; DISTRIBUTION (PROBABILITY); INSTRUMENT SCIENCE; PHYSICAL MEASUREMENT; PROBABILITY; PROBABILITY (PHYSICS); STATISTICS.

R. Gareth Jones

Bibliography. *Guide to the Expression of Uncertainty in Measurement*, 2d ed., International Organization for Standardization, Geneva, 1995.

Caliche

A soil that is mineralogically an impure limestone. Such soils are also known as duricrust, kunkar, nari, kafkalla, Omdurman lime, croute, and race. Many soil profiles in semiarid climates (that is, those characterized by a rainfall of 4–20 in. or 10–50 cm per year) contain concentrations of calcium carbonate ($CaCO_3$). This calcium carbonate is not an original feature of the soils but has been added to the C horizon during soil formation either by direct precipitation in soil pores or by replacement of preexisting material. Fossil analogs of caliche, which are widely reported in ancient sedimentary sequences, are referred to as calcrete or cornstone. *See* LIMESTONE.

Formation. The principal control on the formation of caliche is a hydrologic regime in which there is sufficient moisture to introduce calcium carbonate in solution to the soil but not enough to leach it through the system. As a result, calcium carbonate precipitates in the soil during periods of evaporation, and it will slowly increase in amount as long as the hydrologic setting remains stable. The source of the carbonate may be from the dissolution of adjacent limestones, from the hydrolysis of plagioclase and other silicates, or from carbonate loess.

Within the climatic constraints noted above, most caliche forms in river floodplains and near the surface of alluvial fans. In addition, caliche deposits may form within exposed marine and lacustrine limestones during periods of sea-level fall or lake desiccation. Caliche may also form at inert pediment (eroded rock) surfaces; in the geological record such surfaces will be seen as unconformities. In this context it is interesting that the first unconformity ever recognized as such, by James Hutton in 1787 on the Isle of Arran, western Scotland, is characterized by a development of caliche. *See* UNCONFORMITY.

The mineralogy of the host soil or rock in which a caliche develops may vary considerably; it is not essential for there to be any preexisting carbonate grains within the regolith. The most favorable medium is a clay-rich soil of limited permeability. Low permeability provides the residence time in the soil pores necessary for calcite to precipitate. *See* CALCITE; REGOLITH.

Textures. A variety of textures records the growth of the calcrete. Initially, small nodules of limestone nucleate within the soil; these nodules are usually roughly spherical but may form as elongate tubules and filaments. The latter probably form as casts around roots in the soil. Subsequently, nodules coalesce to form an increasingly solid mass of carbonate, and as a result the soil becomes increasingly rocklike. Eventually the soil profile may become plugged by a massive impermeable limestone. Above the plugged horizon a perched aquifer develops; further calcite precipitation above such a horizon may produce laminated and pisolitic deposits. Continued addition of calcite may result in the buckling and brecciation of laminae. Deoxygenated conditions may develop below the plugged zone; in the reducing environment so formed, elements held in solution in oxygenated waters may precipitate (for example, uranium). *See* OOLITE.

Microscopic fabrics in caliche are initially very fine-grained mosaics of nonferroan low-magnesian calcite crystals. However, more complicated fabrics develop as the paleosol matures. Repeated wetting and drying of the soil results in expansion and contraction of the soil; as a result nodules frequently show signs of multiple cracking. Typically these cracks have been healed because of the precipitation of coarse drusy calcite. This later calcite may be ferroan, that is, may contain some iron. Noncarbonate grains in the caliche usually show signs of intense chemical corrosion and may show signs of physical displacement by precipitating calcite crystals. During the mature stage of development, highly alkaline conditions may develop in caliches. In such conditions quartz in the soil profile may temporarily pass into solution and subsequently reprecipitate as chert, forming a silcrete. *See* CHERT; PALEOSOL.

Mature profile. When the full range of textures referred to above is developed, the caliche is regarded as a mature profile. The vertical profile of such a

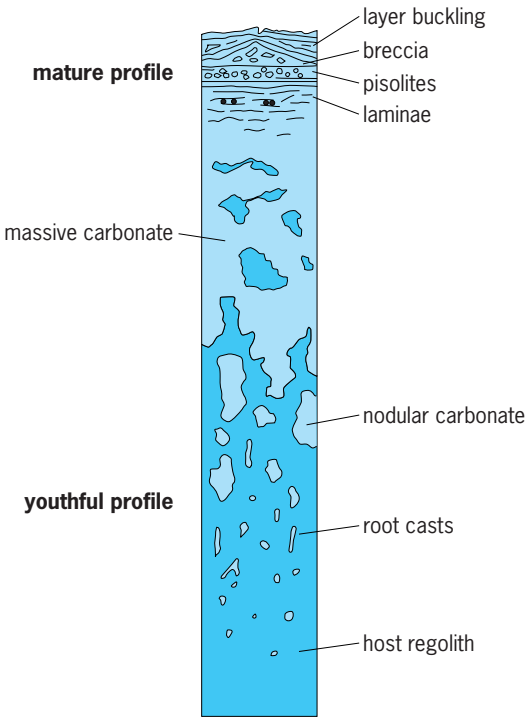

Typical profie through a mature caliche. The lower part of the profile is essentially a mimic of a juvenile caliche.

caliche (see **illus.**) shows a characteristic sequence passing downward from a laminated carbonate into a massive limestone, which in turn passes into increasingly more disconnected nodules. Thus the vertical profile recapitulates the growth of the caliche.

Fully mature caliche profiles may be up to 40 ft (12 m) thick, although 10–20 ft (3–6 m) is more typical. Various workers have tried to estimate the length of time required for the development of a mature caliche. Values of the order of 10,000 years are typical of such estimates. Such lengths of time imply the existence of an environment of great geomorphic stability as continued growth of the caliche requires a constant hydrologic regime. Major fluctuations of climate, variations in base level, and major pulses of tectonic activity are precluded. Environmental stability during caliche formation means that such horizons function as "time lines" in the geological record. In other words, they are important chronostratigraphic markers that permit correlation between adjacent sequences, and thus they help unravel geological history in a more rigorous fashion than would otherwise be the case. *See* SOIL. Nowell Donovan

Bibliography. V. P. Wright (ed.), *Paleosols: Their Recognition and Interpretation*, 1986.

Caliciales

An order of the Ascolichenes. This order is characterized by an unusual apothecium. The hymenial layer originates normally, but by the time the spores are mature, the asci and paraphyses have partially disintegrated into a mass of spores and hymenial tissues known as a mazaedium. In the family Caliciaceae the disk is borne on a short stalk 0.04–0.32 in. (1–8 mm) high, a peculiar structure that is also known in the nonlichenized Roesleriaceae. These two families are so close that some species in either one may or may not lichenize symbiotic algae. The primary thallus of the Caliciaceae is a powdery grayish or lemon-yellow crust on soil or rotten wood. There are six genera separated by spore septation and color. The family Cypheliaceae with seven genera is more typically crustose, with sessile apothecia and a more fully developed thallus. The Sphaerophoraceae are fruticose, much like *Cladonia* sp., but the thallus is solid. The apothecia are open or enclosed in a spherical chamber at the tips of branches. The largest genus, *Sphaerophorus*, is widespread in boreal zones and mountains of both hemispheres. Mason E. Hale

Californium

A chemical element, Cf, atomic number 98, the ninth member of the actinide series of elements. Its discovery and production have been based upon artificial nuclear transmutation of radioactive isotopes of lighter elements. All isotopes of californium are radioactive, with half-lives ranging from a minute to about 1000 years. Because of its nuclear instability, californium does not exist in the Earth's crust. *See* ACTINIDE ELEMENTS; BERKELIUM; PERIODIC TABLE; RADIOACTIVITY.

The chemical properties are similar to those observed for other 3+ actinide elements: a water-soluble nitrate, sulfate, chloride, and perchlorate. Californium is precipitated as the fluoride, oxalate, or hydroxide. Ion-exchange chromatography can be used for the isolation and identification of californium in the presence of other actinide elements. Californium metal is quite volatile and can be distilled at temperatures of the order of 1100–1200°C (2010–2190°F). It is chemically reactive and appears to exist in three different crystalline modifications between room temperature and its melting point, 900°C (1600°F).

The most easily produced isotope for many purposes is ^{252}Cf, which is obtained in gram quantities in nuclear reactors and has a half-life of 2.6 years. It

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

decays partially by spontaneous fission, and has been very useful for the study of fission. It has also had an important influence on the development of counters and electronic systems with applications not only in nuclear physics but in medical research as well. *See* TRANSURANIUM ELEMENTS. Glenn T. Seaborg

Bibliography. B. B. Cunningham, Berkelium and californium, *J. Chem. Educ.*, 36:32–37, 1959; B. B. Cunningham, Chemistry of the actinide elements, *Annu. Rev. Nucl. Sci.*, 14:323–346, 1964; S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; J. J. Katz, G. T. Seaborg, and L. R. Morss (eds.), *The Chemistry of the Actinide Elements*, 2 vols., 2d ed., 1986.

Caliper

An instrument with two legs used for measuring linear dimensions. Calipers may be fixed, adjustable, or movable. Fixed calipers are used in routine inspection of standard products; adjustable calipers are used similarly but can be reset to slightly different dimensions if necessary. Movable calipers can be set to match the distance being measured. The legs may pivot about a rivet or screw in a firm-joint pair of calipers; they may pivot about a pin, being held against the pin by a spring and set in position by a knurled nut on a threaded rod; or the legs may slide either directly (caliper rule) or along a screw (micrometer caliper) relative to each other. *See* MICROMETER.

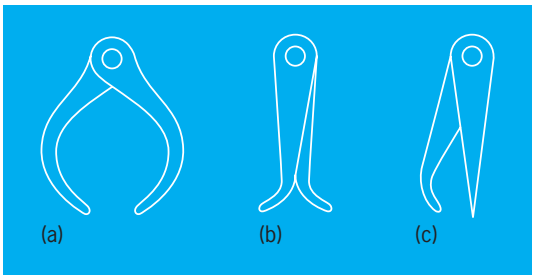

Some typical machinist's calipers. (*a*) Outside. (*b*) Inside. (*c*) Hermaphrodite. (*After R. J. Sweeney, Measurement Techniques in Mechanical Engineering, Wiley, 1953*)

The legs may be shaped to facilitate measuring outside dimensions, inside dimensions, surface dimensions as between points on a plate, or from a surface into a hole as in a keyway (see **illus.**). Other forms are adapted to special needs. *See* GAGES. Frank H. Rockett

Callitrichales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Asteridae of the class Magnoliopsida (dicotyledons). The order consists of three small families with about 50 species, most of which are aquatics or small herbs of wet places, and have much reduced vascular systems. The flowers are small and solitary in the axils of leaves or bracts. The perianth is nearly or completely absent. The pistil consists of two carpels united to form a compound, unilocular or four-chambered ovary, or sometimes the pistil appears to be of a single carpel. The Callitrichales are placed in the Asteridae largely on the basis of their embryology and phytochemistry. The ovules are anatropous and tenuinucellular and have a single integument, and the plants generally produce iridoid substances. *See* ASTERIDAE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM. T. M. Barkley

Calobryales

An order of liverworts. They are characterized by prostrate, simple or branched, leafless stems and erect, leafy branches of a radial organization. The order consists of a single genus, *Calobryum*, and 12 species, most of them occupying restricted ranges in apparently relic areas indicative of an ancient origin and dispersal. The order is considered primitive in comparison with the Jungermanniales, in which the leafy axis tends to be prostrate and the underleaves reduced.

The stems are thick and fleshy, with no differentiated outer layers. Rhizoids are lacking. The leaves may be small or lacking below, larger and more crowded above. They are three-ranked, with those of one rank sometimes more or less reduced. They are broad, unlobed, and entire. The cells, large and thin-walled, are often two- or three-layered at the leaf base. Slime papillae at the base of leaves secrete mucilage. The antheridia occur in axils of upper leaves or in terminal clusters subtended by three leaves; the stalk is long, four-seriate, and enlarged above. The archegonia are clustered in a terminal disk but in the absence of fertilization become lateral; the neck is long and four-seriate. The sporophyte is protected by a cylindric upgrowth of gametophytic tissue. The seta consists of numerous rows of cells, and the long-ellipsoid capsule dehisces by four valves remaining attached at the tip. The wall is unistratose below and bistratose at the apex; its cells have one or two longitudinally oriented bands. The haploid chromosome number is 9. *See* BRYOPHYTA; HEPATICOPSIDA; JUNGERMANNIALES; JUNGERMANNIIDAE.

Howard Crum

Bibliography. E. O. Campbell, The structure and development of *Calobryum gibbsiae* Steph., *Trans. Roy. Soc. N. Z.* 47:243–254; R. M. Schuster, Studies on Hepaticae, XV: Calobryales, *Nova Hedwigia*, 13:1–63, 1966.

Calomel

Mercury(I) chloride, Hg_2Cl_2, a covalent compound which is insoluble in water. The substance sublimes when heated. The formula weight is 472.086 and the

specific gravity is 7.16 at 20°C (68°F). The material is a white, impalpable powder consisting of fine tetragonal crystals. Calomel is manufactured by precipitation when sodium chloride is added to a solution of mercury(I) nitrate or by direct combination of the elements.

Calomel is used in preparing insecticides and medicines. It is well known in the laboratory as the constituent of the calomel reference electrodes which are commonly used in conjunction with a glass electrode to measure pH. *See* MERCURY (ELEMENT); REFERENCE ELECTRODE. E. Eugene Weaver

Calorimetry

The measurement of the quantity of heat energy involved in processes such as chemical reactions, changes of state, and mixing of substances, or in the determination of heat capacities of substances. The unit of energy in the International System of Units is the joule, symbolized J. Another unit still being used, but is strongly discouraged, is the calorie, symbolized cal and defined as 4.184 J. Most calorimetric measurements are made at constant pressure, and the measured change is called the enthalpy change. *See* ENTHALPY; HEAT CAPACITY.

Types of calorimeters. A calorimeter is an apparatus for measuring the quantity of heat energy released or absorbed during a process. Since there are many processes that can be studied over a wide range of temperature and pressure, a large variety of calorimeters have been developed.

Nonisothermal calorimeters. These instruments measure the temperature change that occurs during the process. An aneroid (containing no liquid) nonisothermal (temperature varying during the experiment) calorimeter is normally constructed of a material having a high thermal conductivity, such as copper, so that there is rapid temperature equilibration. It is isolated from its surroundings by a high vacuum to reduce heat leaks. This type of calorimeter can be used for determining the heat capacity of materials at low temperatures. Aneroid nonisothermal calorimeters have also been developed for measuring the energy of combustion for small samples of rare materials.

Another type of nonisothermal calorimeter uses a large quantity of liquid as a heat sink, so that even if there is a large amount of energy released in the process the temperature rise is not excessive. The liquid is kept at a uniform temperature by stirring, and the change in temperature can be measured accurately. In general, the size of the liquid heat sink is designed so that the temperature change is of the order of 5 K (9 °F). For measurements near room temperature, water is commonly used as the working fluid; measurements at higher temperatures use less volatile organic liquids.

With most nonisothermal calorimeters, it is necessary to relate the temperature rise to the quantity of energy released in the process. This is done by determining the calorimeter constant, which is the amount of energy required to increase the temperature of the calorimeter itself by 1 K. This value can be determined by measurement on a well-defined test system or by electrical calibration. In bomb calorimetry, for example, the calorimeter constant is usually determined from the temperature rise that occurs when a known mass of a very pure standard sample of benzoic acid is burned. Electrical calibration, however, is complex and often does not mimic the way that energy is released during the reaction.

Isothermal calorimeters. These instruments make measurements at constant temperature. The simplest example is a calorimeter containing an outer annular space filled with a liquid in equilibrium with a crystalline solid at its melting point, arranged so that any volume change will displace mercury along a capillary tube. The Bunsen ice calorimeter operates at 0 °C (32 °F) with a mixture of ice and water. Changes resulting from the process being studied cause the ice to melt or the water to freeze, and the consequent volume change is determined by measurement of the movement of the mercury meniscus in the capillary tube. While these calorimeters can yield accurate results, their operation is limited to the equilibrium temperature of the two-phase system. Other types of isothermal calorimeters use electrical energy to achieve exact balance of the heat absorption that occurs during an endothermic process. Generally, calorimeters based on this principle are more accurate than similar nonisothermal calorimeters. Many mixing and flow calorimeters operate in this mode.

Calorimeter components. All calorimeters consist of the calorimeter proper and a jacket or a bath, which is used to control the temperature of the calorimeter and the rate of heat leak to the environment. For temperatures not too far removed from room temperature, the jacket or bath contains liquid at a controlled temperature. For measurements at extreme temperatures, the jacket usually consists of a metal block containing a heater to control the temperature. With nonisothermal calorimeters, where

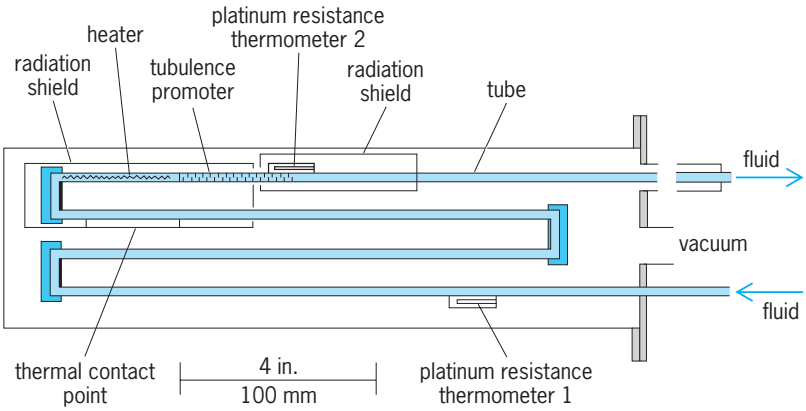

Fig. 1. High-pressure flow calorimeter for measurement of heat capacity. (*After G. Ernst, G. Maurer, and E. Wiederoh, Flow calorimeter for the accurate determination of the isobaric heat capacity at high pressures: Results for carbon dioxide, J. Chem. Thermodyn., 21:53–65, 1989*)

the jacket is kept at a constant temperature, there will be some heat leak to the jacket when the temperature of the calorimeter changes. It is necessary to correct the temperature change observed to the value it would have been if there were no leak. This is achieved by measuring the temperature of the calorimeter before and after the process and applying Newton's law of cooling. This correction can be avoided by using the technique of adiabatic (no energy loss or gain with the surroundings) calorimetry, where the temperature of the jacket is kept equal to the temperature of the calorimeter as a change occurs. This technique requires more elaborate temperature control, and its primary use is for accurate heat capacity measurements at low temperatures.

Thermometry. In calorimetric experiments, it is necessary to measure temperature differences accurately; in some cases, the temperature itself must be accurately known. Modern calorimeters use resistance thermometers to measure both temperatures and temperature differences, while thermocouples or thermistors are used to measure smaller temperature differences. A calibrated standard platinum resistance thermometer can measure temperature differences to 0.00001 °C and temperatures ranging from −260 °C to 630 °C (−436 °F to 1166 °F) with an uncertainty of ±0.001 °C. Multiple-junction thermocouples and thermistors can measure smaller temperature differences, allowing precise measurements of very small heat effects. *See* TEMPERATURE MEASUREMENT; THERMISTOR; THERMOCOUPLE; THERMOMETER.

Calorimetric measurements. Heat capacities of materials and energies of combustion are processes that are routinely measured with calorimeters. Calorimeters are also used to measure the heat involved in phase changes, for example, the change from a liquid to a solid (fusion) or from a liquid to a gas (vaporization). Calorimetry has also been applied to the measurement of the enthalpy of hydrogenation of unsaturated organic compounds, the enthalpy of dissolution of a solid in a liquid, or the enthalpy change on mixing two liquids.

Heat-capacity calorimeter. Heat capacities of gases or liquid can be determined in a flow calorimeter. Some essential features are shown in **Fig. 1**. A known mass of fluid flowing at a constant rate is pumped through the calorimeter. The inlet temperature is measured with thermometer 1. The fluid is heated with a known amount of electrical energy, and the change in temperature is observed at thermometer 2. The radiation shields reduce heat losses to the surroundings.

Enthalpy-of-vaporization calorimeter. A modern enthalpy-of-vaporization calorimeter is shown in **Fig. 2**. The material in the vaporization vessel is heated with a known amount of electrical energy, and the mass of material vaporized is determined from the change in mass of the glass collection vessel. The thermostated shields, which are maintained at the same temperature as the vaporization vessel, ensure that all the electrical energy goes to vaporize the sample and none is lost to the surroundings. *See* CHEMICAL THERMODYNAMICS; HEAT TRANSFER; THERMOCHEMISTRY; THERMODYNAMIC PRINCIPLES.

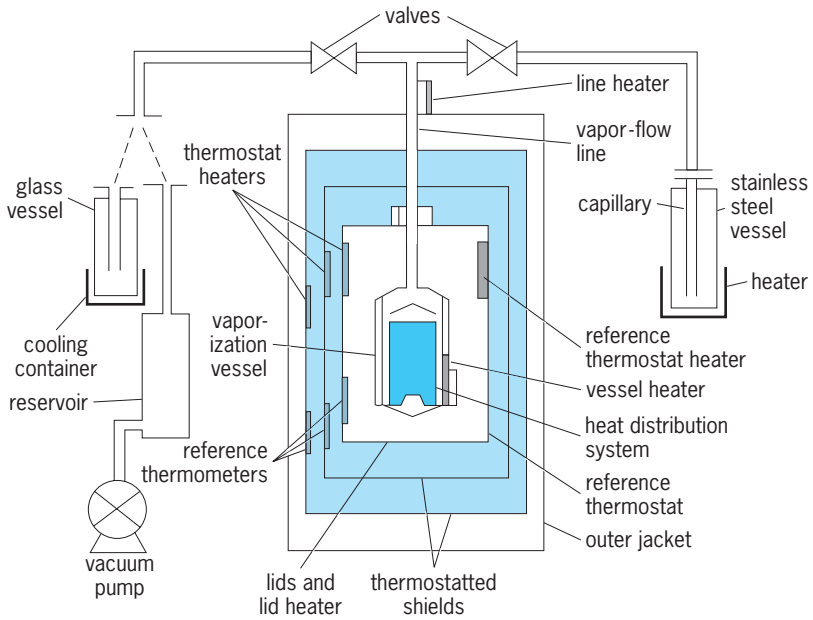

Fig. 2. Heat-of-vaporization calorimeter. (*After L. Ŝváb et al., A calorimeter for the determination of enthalpies of vaporization at high temperatures and pressures, J. Chem. Thermodyn., 20:545–550, 1988*)

Kenneth N. Marsh

Bibliography. K. N. Marsh and P. G. O'Hare (eds.), *IUPAC Experimental Thermodynamics*, vol. IV: *Solution Calorimetry*, 1994; J. P. McCullough and D. W. Scott (eds.), *Experimental Thermodynamics*, vol. 1: *Calorimetry of Non-Reacting Systems*, 1968; M. L. McGlashan, *Chemical Thermodynamics*, 1979; J. H. Moore and N. D. Spencer (eds.), *Encyclopedia of Chemical Physics and Physical Chemistry*, vol. 2, B1.27: *Calorimetry*, 2001; W. R. Parrish, Recent advances in calorimetry, *Fluid Phase Equil.*, 29:172–192, 1986; R. S. Porter and J. F. Johnson (eds.), *Analytical Chemistry*, vols. 2–4, 1970–1977; M. Sorai, *Comprehensive Handbook of Calorimetry and Thermal Analysis*, 2004.

Calycerales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Asteridae of the class Magnoliopsida (dicotyledons). The order consists of a single family with about 60 species native to tropical America. The plants are herbs with alternate, simple leaves that do not have stipules. The flowers are borne in involucrate heads with centripetal flowering sequence. The calyx is reduced to small lobes or teeth, and the corolla consists of (4)5(6) fused lobes and is regular or somewhat irregular. The stamens are attached near the summit of the corolla tube, and the filaments are more or less connate. The pistil consists of two united carpels, which form a compound, inferior ovary wth a single, pendulous

ovule. The order Calycerales is sometimes included within the Dipsacales, and the order has attracted attention because of the overall resemblance of the inflorescence to that of the Asteraceae. *See* ASTERALES; ASTERIDAE; DIPSACALES; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM. T. M. Barkley

Cam mechanism

A mechanical linkage whose purpose is to produce, by means of a contoured cam surface, a prescribed motion of the output link of the linkage, called the follower. Cam and follower are a higher pair. *See* LINKAGE (MECHANISM).

A familiar application of a cam mechanism is in the opening and closing of valves in an automotive engine (**Fig. 1**). The cam rotates with the cam shaft,

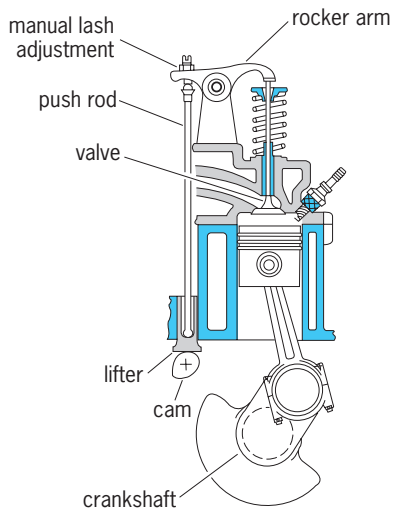

Fig. 1. Cam mechanism for opening and closing valves in automotive engine. (*Texaco, Inc.*)

usually at constant angular velocity, while the follower moves up and down as controlled by the cam surface. A cam is sometimes made in the form of a translating cam (**Fig. 2***a*). Other cam mechanisms, employed in elementary mechanical analog computers, are simple memory devices, in which the position of the cam (input) determines the position of the follower (output or readout).

Although many requisite motions in machinery are accomplished by use of pin-joined mechanisms, such as four-bar linkages, a cam mechanism frequently is the only practical solution to the problem of converting the available input, usually rotating or reciprocating, to a desired output, which may be an exceedingly complex motion. No other mechanism is as versatile and as straightforward in design. However, a cam may be difficult and costly to manufacture, and it is often noisy and susceptible to wear, fatigue, and vibration.

Cams are used in many machines. They are numerous in automatic packaging, shoemaking, typesetting machines, and the like, but are often found as well in machine tools, reciprocating engines, and

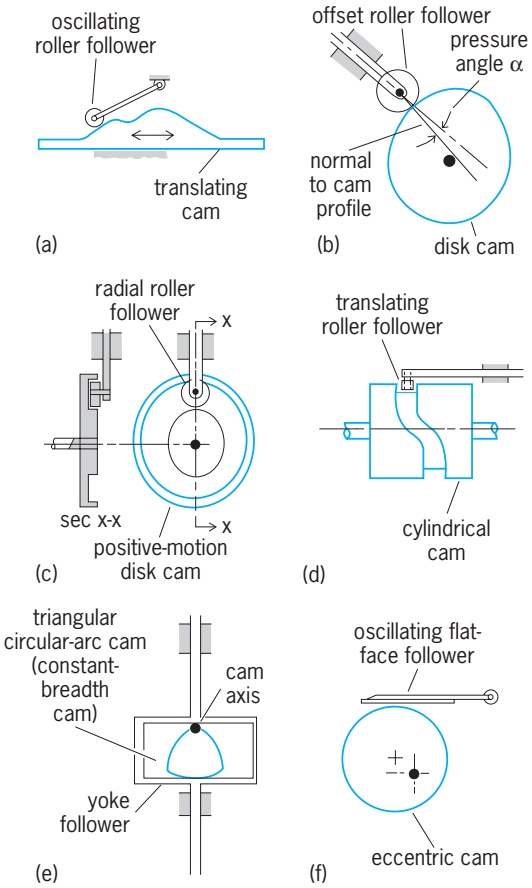

Fig. 2. Classification of cams. (*a*) Translating. (*b*) Disk. (*c*) Positive motion. (*d*) Cylindrical. (*e*) With yoke follower. (*f*) With flat-face follower.

compressors. They are occasionally used in rotating machinery.

Cams are classified as translating, disk, plate, cylindrical, or drum (Fig. 2). The link having the contoured surface that prescribes the motion of the follower is called the cam. Cams are usually made of steel, often hardened to resist wear and, for high-speed application, precisely ground.

The output link, which is maintained in contact with the cam surface, is the follower. Followers are classified by their shape as roller, flat face, and spherical face (**Fig. 3**). The point or knife-edge follower is of academic interest in developing cam profile relationships. Followers are also described by the nature of their constraints, for example, radial, in which motion is reciprocating along a radius from the cam's axis of rotation (Fig. 1); offset, in which motion is reciprocating along a line that does not intersect the axis of rotation (Fig. 2*b*); and oscillating, or pivoted (Fig. 2*a*. Three-dimensional cam-and-follower

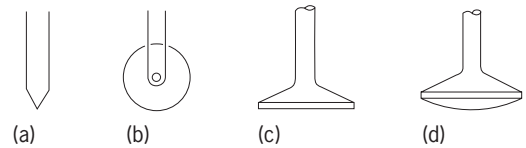

Fig. 3. Cam followers. (*a*) Knife edge. (*b*) Roller. (*c*) Flat face. (*d*) Spherical face.

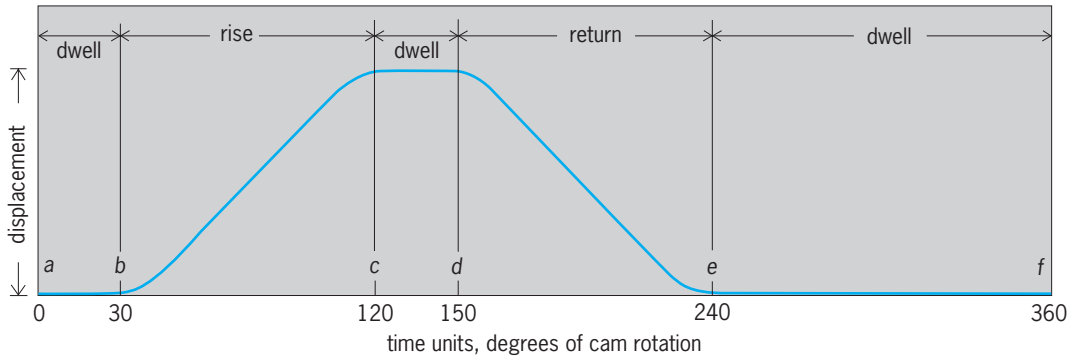

Fig. 4. Displacement-time diagram for a cam, with motion of cam follower indicated.

systems are coming into more frequent use, where the follower may travel over a lumpy surface.

Motion of cam follower. The first step in the design of a cam mechanism is the determination of the motion of the cam follower. In a packaging machine, for example, the ends of a carton may be folded by cam-operated fingers that advance at the proper times, retract as soon as the fold has been made, and then rest or dwell until the next carton is in position. The motion of the cam follower, which in turn moves the folder fingers, can be represented by a displacement-time diagram (**Fig. 4**). The time axis is usually laid off in degrees of cam rotation. The conventional meanings of follower dwell, rise, and return are indicated in the figure.

The maximum displacement of the follower and the periods of dwell are determined, more or less arbitrarily, by the designer, who has the choice of any curve to connect the dwell portions of the complete displacement-time diagram. The practical form for this curve is determined largely by the maximum acceleration that can be tolerated by the follower linkage. In addition, the pressure angle (α in Fig. 2*b*) must be kept fairly small, usually less than 30°, to avoid undue friction and possible jamming of the reciprocating follower in its guides. The space that is available for the cam will affect the maximum pressure angle. Usually, a small cam is preferred; yet the larger the cam can be made (which in effect physically increases the length of the time axis for the same time interval), the smaller the maximum pressure angle will be. The final form may further represent a compromise to make possible economical manufacture of the cam.

If the diagram of Fig. 4 were laid out on and cut out of steel and a knife-edge follower were constrained to move vertically, the translating cam mechanism of **Fig. 5***a* would result. The process of wrapping this translating cam around a disk (Fig. 5*b*), thus producing a disk cam whose follower action would be similar to that of Fig. 5*a*, can be visualized readily. The introduction of a roller or flat-face follower complicates the determination of the actual cam contour that will produce a desired follower displacement-time relationship; but recognition of the similarity between the displacement-time diagram and the final cam contour makes it easier to visualize the conditions that must be met to design a cam that will operate satisfactorily.

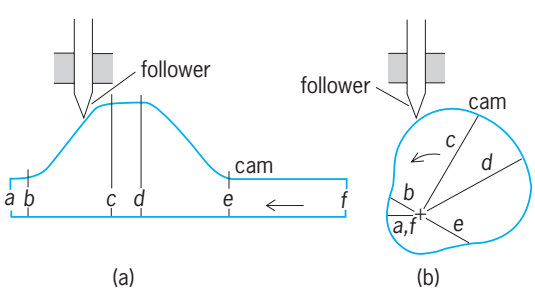

Fig. 5. Converting (*a*) translating cam to (*b*) disk cam.

Consider the dwell-rise-dwell portion of the curve of Fig. 4. Displacement of the follower might be plotted arbitrarily as a straight line (**Fig. 6***a*).

The slope $\Delta s/\Delta t$ of a displacement-time ($s-t$) curve is equal to velocity ($s/t = v$) so that velocity of the follower from A to B (dwell) will be zero, from B to C it will be constant and finite, and from C to D velocity will be zero again (Fig. 6*b*).

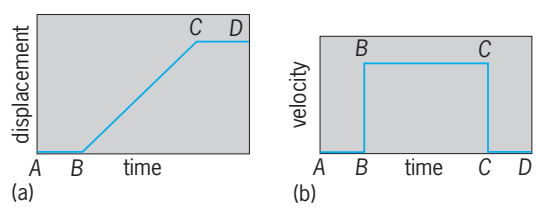

Fig. 6. Effect of (*a*) displacement on (*b*) velocity.

The slope $\Delta v/\Delta t = a$ of the velocity-time curve is correspondingly equal to acceleration. Thus, the acceleration of the follower necessary to increase velocity from zero to a finite value in zero time (B to B in Fig. 6*b*) is infinite. Likewise, the deceleration that occurs at C must also be infinite. The acceleration along the constant velocity line, from B to C, would be zero. Thus the curve chosen in Fig. 6*a* for displacement is unrealistic because of the high inertial forces that would result from abrupt changes of velocity.

Choice of acceleration curve. Therefore, a curve having a gradual transition from dwell to maximum velocity is necessary. Three such curves are plotted in **Fig. 7** and are superimposed in **Fig. 8** for comparison. The derived curves for velocity and acceleration are also plotted so comparisons may be made.

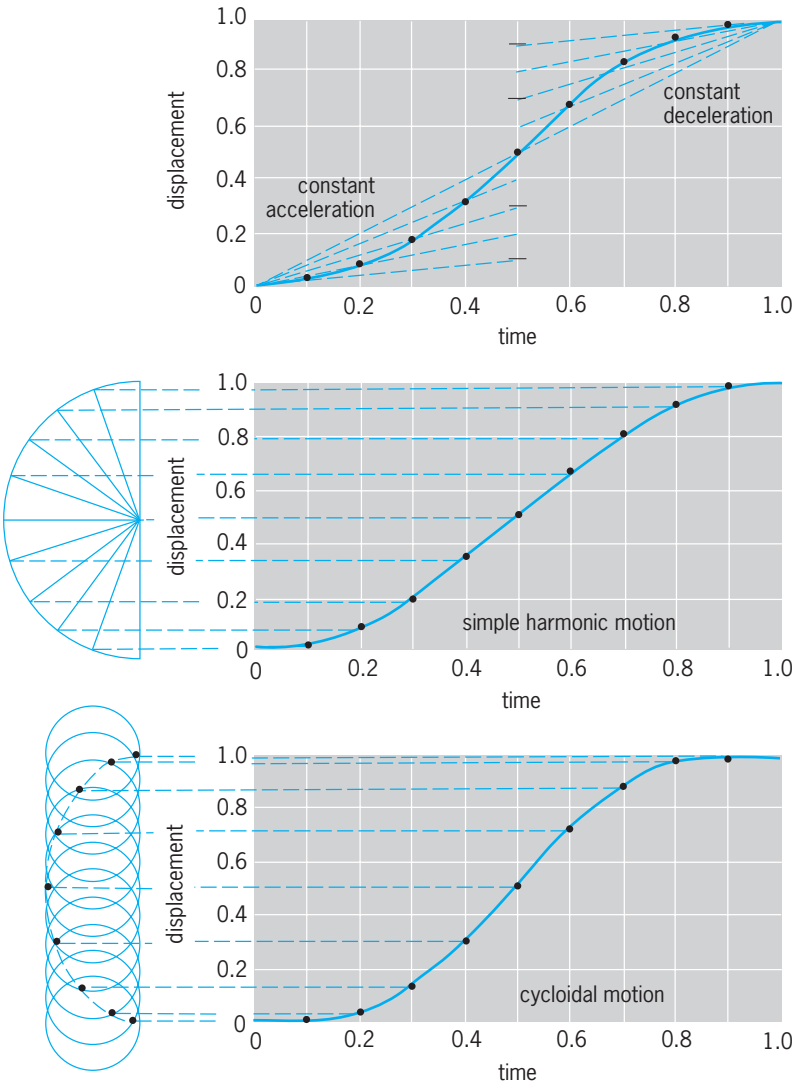

Fig. 7. Displacement-time diagrams for cam contours.

The constant acceleration–constant deceleration curve, in which displacement s is proportional to t^2, is desirable except for the instantaneous reversal of acceleration at the point of maximum velocity; such a reversal would cause high stresses in the mechanism. If the follower were spring-loaded, a heavy spring would be required to prevent the follower's leaving the cam face momentarily, with resulting shock to the linkage as it returned.

The simple harmonic displacement curve is plotted by projecting onto the diameter (equal to follower displacement) of a circle a point moving with constant velocity around the circle's circumference. Although maximum acceleration is higher than in the preceding curve, the abrupt changes of acceleration occur only at the beginning and end of the rise. Both of these curves have been used in cam design; both are satisfactory if speeds are low to moderate and follower mass is not large. However, serious difficulties are encountered when high speeds or heavy followers accentuate the stresses resulting from acceleration.

The cycloidal curve is plotted by projecting points from a cycloid whose generating circle has a diameter equal to follower displacement divided by π, as in Fig. 7. This curve has desirable acceleration characteristics, but requires that the cam face be accurately machined at the beginning and end of rise to accomplish in fact the theoretical performance.

Comparison of the three curves is shown in Fig. 8. The cycloidal curve exhibits a higher pressure angle α for a radial translating follower than the other curves, but its acceleration characteristics are much superior.

Today's extensive use of cams and the extreme demands upon them under increasing speeds of modern manufacturing have yielded still another form of cam profile, which might be called catenoidal because of its association with the catenary [the curve in which a rope or chain (catena) hangs freely]. This curve has an equation expressed primarily in terms of exponential functions ϵ^x, whose slope variation is a curve with the remarkable property that it partakes of the same general form as the original function. Hence the velocity and acceleration curves show

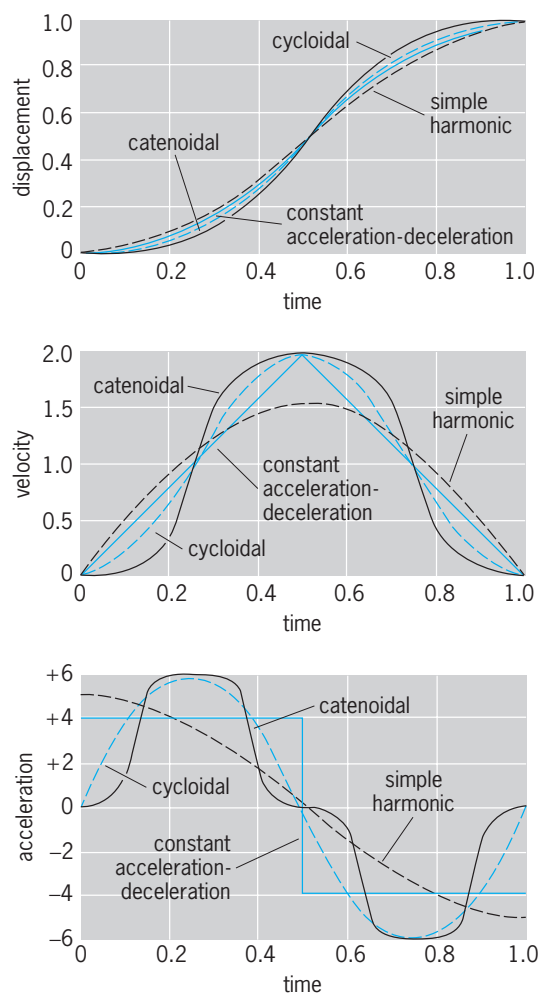

Fig. 8. Comparison of motions for three shapes of rise-return regions. Fourth motion, for catenoidal shape, also superposed. (*After G. L. Guillet and A. H. Church, Kinematics of Machines, 5th ed., Wiley, 1950*)

patterns similar to that of displacement rather than being so different, as was markedly so with the constant acceleration-deceleration profile, and less so with others. This curve (catenoidal) also is shown in the figures; it is alleged to give even smoother performance and less vibration than the cycloidal type. Its slope (jerk) will be zero four times in each cycle.

The manufacture of cam and follower from a master model is dwindling in favor of manufacture using a computer tape, which accurately directs the cutting of the surface. This procedure avoids the expensive manufacture of master cams.

In certain high-speed cam mechanisms, for example, an automotive engine valve gear, the elasticity and vibration characteristics of the follower linkage must be taken into account if faulty operation is to be avoided. The polydyne method derives its name from use of a polynomial displacement curve that suits the dynamic characteristics of the follower linkage.

Construction of cam profile. Empirically, a cam profile can be plotted to as large a scale as desired if the displacement curve and the configuration of the follower linkage are known. The method consists essentially in inverting the mechanism by fixing the cam and rotating the follower linkage about it, plotting only enough of the follower linkage to establish the successive positions of the follower face that will bear on the cam. **Figure 9** shows the method of constructing the profile for a disk cam having a radial translating roller follower whose displacement curve is given. The cam profile is faired in, being at every point tangent to the follower roller. (Accuracy is improved by using a larger scale and by plotting additional positions.)

The method of constructing a cam profile for a flat face, offset, or oscillating follower is similar to that shown in the figure. It is important, however, that the location of the point in the follower linkage whose displacement is described by the displacement-time curve be kept constantly in view while the follower linkage is plotted in various successive positions.

Analytically, a profile can be calculated to any desired accuracy and the cam profile may be shown in tabular form, giving, for example, displacement of the follower for each degree of cam rotation. If a milling cutter or grinding wheel of the same size and shape as the follower is then used to cut the cam contour, the resulting contour will be true except for the small ridges that remain between given positions. These ridges can be removed by hand, using a file or a stone. Douglas P. Adams

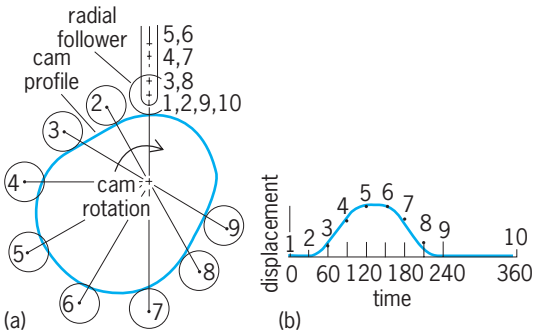

Fig. 9. Method of construction of (*a*) the cam profile for a disk cam having a radial translating roller follower, from (*b*) the displacement curve.

Bibliography. J. Chakraborty and S. G. Dhande, *Kinematics and Geometry of Planer and Spatial Cam Mechanisms*, 1977; F. J. Ogozalek, *Theory of Catenoidal-Pulse Motion and Its Application to High-Speed Cams*, ASME Publ. 66-Mech-45, 1966; J. E. Shigley and J. J. Uiker, *Theory of Machines and Mechanisms*, 2d ed., 1995.

Cambrian

An interval of time in Earth history (Cambrian Period) and its rock record (Cambrian System). The Cambrian Period spanned about 60 million years and began with the first appearance of marine animals with mineralized (calcium carbonate, calcium phosphate) shells. The Cambrian System includes many different kinds of marine sandstones, shales, limestones, dolomites, and volcanics. Apart from the occurrence of an alkaline playa containing deposits of trona (hydrated basic sodium carbonate) in the Officer Basin of South Australia, there is very little provable record of nonmarine Cambrian environments.

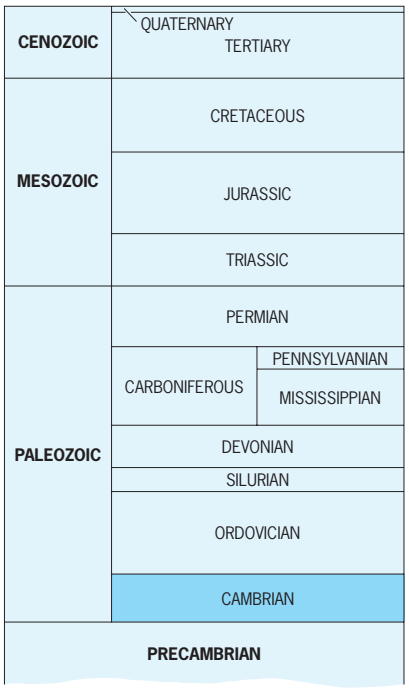

The concept that great systems of rocks recorded successive periods of Earth history was developed in England in the early nineteenth century. The Cambrian, which was one of the first systems to be formally named, was proposed by the Reverend Adam Sedgwick in 1835 for a series of sedimentary rocks in Wales that seemed to constitute the oldest sediments in the British Isles. At that time, there was no real idea

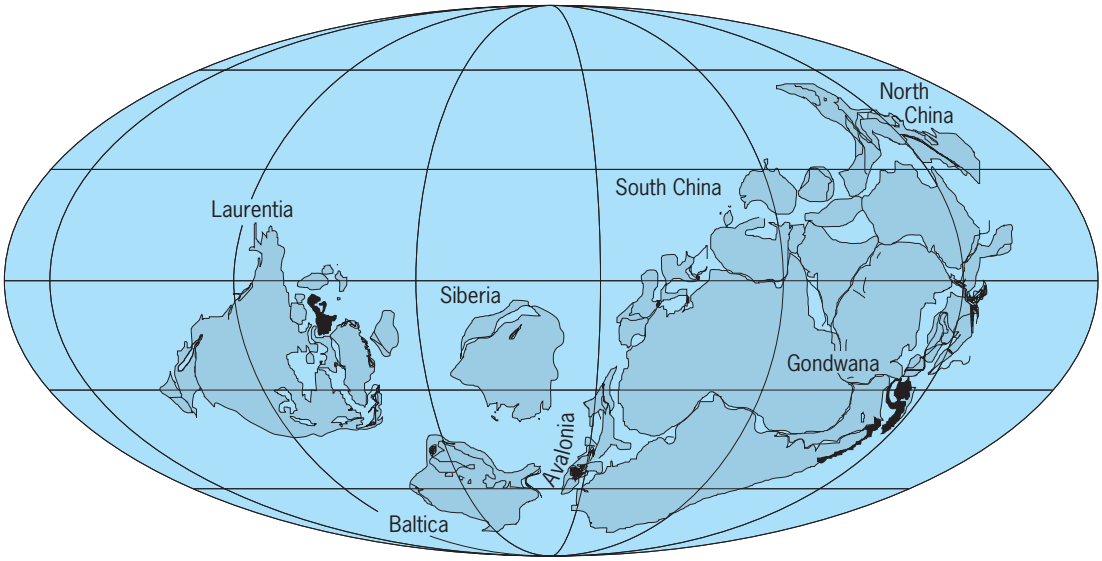

Fig. 1. Reconstruction of the Lower Cambrian world. (*After W. S. McKerrow, C. R. Scotese, and M. D. Brasier, Early Cambrian continental reconstructions. J. Geol. Soc., 149:599–606, 1992*)

of the antiquity of Cambrian rocks. They were recognized by distinctive fossils and by their geologic relations to other systems. In the early part of the twentieth century, radiometric techniques for obtaining the ages of igneous and metamorphic rocks evolved. Because of the difficulty of finding rocks that can be dated radiometrically in association with rocks, usually zircon bearing volcanic ashes, that can be dated empirically by fossils, the age in years of most Cambrian deposits is only approximate. The best present estimates suggest that Cambrian time began about 545 million years ago (Ma; earliest date close to the base of the Cambrian is 543 ± 0.2 Ma) and ended at about 485 Ma (latest date in the Upper Cambrian is 491 ± 1 Ma, but this is not terminal Cambrian). It is the longest of the Paleozoic periods and the fourth long est of the Phanerozoic periods.

Geography. Knowledge of Cambrian geography and of the dynamic aspects of evolution and history in Cambrian time is derived from rocks of this age that have been exposed by present-day erosion or penetrated by borings into the Earth's surface. Despite the antiquity of Cambrian time, a surprisingly good record of marine rocks of Cambrian age has been preserved at many localities throughout the world. Each of the different rock types contains clues about its environment of deposition that have been derived from analogy with modern marine environments. From this information, together with knowledge gained from fossils of about the same age within the Cambrian and information about the present geographic distribution of each Cambrian locality, a general picture of world geography and its changes through Cambrian time is available.

Plate tectonics. The theory of plate tectonics has provided criteria whereby ancient continental margins can be identified. By using these criteria and the spatial information about marine environments derived from study of the rocks, the Cambrian world can be resolved into at least four major continents that were quite different from those of today (**Fig. 1**). These were (1) Laurentia, which is essentially North America, minus a narrow belt along the eastern coast from eastern Newfoundland to southern New England that belonged to a separate microcontinent, Avalonia. This microcontinent, which also included present-day England, and another microcontinent now incorporated in South Carolina were originally marginal to Gondwana; (2) Baltica, consisting of present-day northern Europe north of France and west of the Ural Mountains but excluding most of Scotland and northern Ireland, which are fragments of Laurentia; (3) Gondwana, a giant continent whose present-day fragments are Africa, South America, India, Australia, Antarctica, parts of southern Europe, the Middle East, a nd Southeast Asia; and (4) Siberia, including much of the northeastern quarter of Asia. Unfortunately, there is not enough reliable information to accurately locate these continents relative to one another on the Cambrian globe. Current Cambrian reconstructions rely on similarities of fossil faunas and on studies of magnetic polarity reversals in rock sequences through time (magnetostratigraphy). These have mostly been concentrated across the Precambrian-Cambrian boundary in Australia, Morocco, Siberia, and south China, and across the Cambrian-Ordovician boundary in Australia, North America, Kazakhstan, and north China. *See* CONTINENTAL MARGIN; CONTINENTS, EVOLUTION OF; PLATE TECTONICS.

Time divisions. For most practical purposes, rocks of Cambrian age are recognized by their content of distinctive fossils. On the basis of the successive changes in the evolutionary record of Cambrian life that have been worked out during the past century, the Cambrian System has been divided globally into three or four series, each of which has been further divided on each continent into stages, each stage consisting of several zones (**Fig. 2**). Despite the amount of work already done, precise

SERIES	STAGES	ZONES	
MILLARDAN	Sunwaptan	Saukia	*
		Saratogia	
		Taenicephalus	
		Irvingella major	*
	Steptoean	Elvinia	
		Dunderbergia	
		Prehousia	
		Dicanthopyge	
		Aphelaspis	
		Coosella perplexa	*
LINCOLNIAN	Marjuman	Crepicephalus	
		Cedaria	
		Bolaspidella	
		Ehmaniella	
		Proehmaniella basilica	*
	Delamaran	Glossopleura	
		Albertella	
		Plagiura-Poliella	*
WAUCOBAN	Dyeran	Olenellus	
	Montezuman	"Nevadella"	
		"Fallotaspis"	
BEGADEAN	Unnamed		

Fig. 2. North American divisions of the Cambrian System. Asterisks denote levels of major trilobite extinctions.

intercontinental correlation of series and stage boundaries, and of zones, is still difficult, especially in the Early Cambrian due to marked faunal provinciality. Refinement of intercontinental correlation of these ancient rocks is a topic of research.

Life. The record preserved in rocks indicates that essentially all Cambrian plants and animals lived in the sea. The few places where terrestrial sediments have been preserved suggest that the land was barren of major plant life, and there are no known records of Cambrian insects or of terrestrial vertebrate animals of any kind.

Plants. The plant record consists entirely of algae, preserved either as carbonized impressions in marine black shales or as filamentous or blotchy microstructures within marine buildups of calcium carbonate, called stromatolites, produced by the actions of these organisms. Cambrian algal stromatolites were generally low domal structures, rarely more than a few meters high or wide, which were built up by the trapping or precipitation of calcium carbonate by one or more species of algae. Such structures, often composed of upwardly arched laminae, were common in regions of carbonate sedimentation in the shallow Cambrian seas. *See* STROMATOLITE.

Animals. The animal record is composed almost entirely of invertebrates that had either calcareous or phosphatic shells (**Fig. 3**). The fossils of shell-bearing organisms include representatives of several different classes of arthropods, mollusks, echinoderms, brachiopods, and poriferans. Coelenterates, radiolarians, and agglutinated foraminiferans are extremely rare, and bryozoans are unknown from Cambrian rocks. Rare occurrences of impressions or of carbonized remains of a variety of soft-bodied organisms, including worms and a group of soft-bodied trilobites, indicate that the fossil record, particularly of arthropods, is incomplete and biased in favor of shell-bearing organisms. Some widespread fossil groups, such as Archaeocyatha, are known only from Cambrian rocks, and several extinct groups of Paleozoic organisms, such as hyolithids and conodonts, first appear in Cambrian rocks. Conodonts are thought by some specialists to have affinity with vertebrates, but others prefer to relate them to cephalochordates. Dermal plates recovered from the Late Cambrian of North America and Australia are considered to represent the earliest fish remains. *See* ARTHROPODA; CONODONT; PORIFERA.

Diversity. Although the record of marine life in the Cambrian seems rich, one of the dramatic differences between Cambrian marine rocks and those of younger periods is the low phyletic diversity of

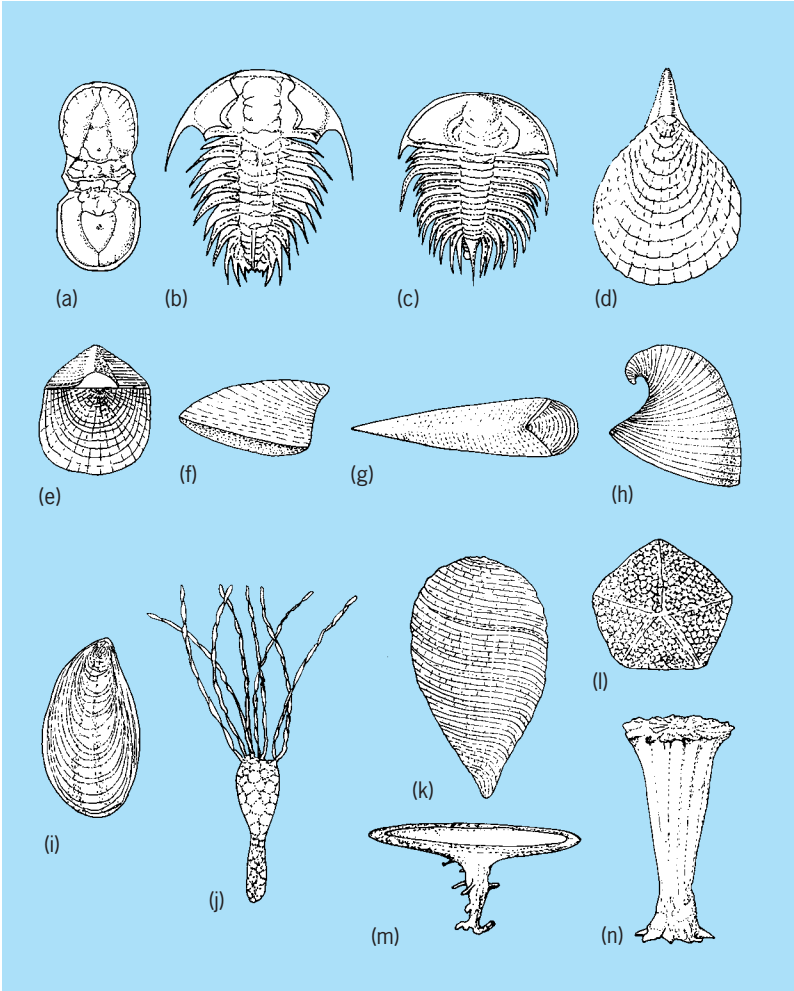

Fig. 3. Representative Cambrian fossils: (*a–c*) trilobites; (*d–f*) brachiopods; (*g*) hyolithid; (*h–i*) mollusks; (*j–l*) echinoderms; and (*m*, *n*) archaeocyathids.

most fossiliferous localities. The most diverse faunas of Cambrian age have been found along the ocean-facing margins of the shallow seas that covered large areas of the Cambrian continents. Because these margins were often involved in later geologic upheavals, their rich record of Cambrian life has been largely destroyed. Only a few localities in the world remain to provide a more accurate picture of the diversity of organisms living in Cambrian time. These are known as Konservat Lagerstätten—conservation deposits containing occurrences of extraordinary preservation, particularly of soft body parts. Globally, they are known from more than 35 localities to date if the "Orsten"-type preservation in the Swedish Alum Shale and elsewhere are considered as Lagerstätten. Orsten is an organic-rich, anthraconitic, concretionary limestone in which phosphatized cuticle-bearing organisms are exquisitely preserved in three dimensions. In Laurentia, the richest localities are in the Kinzers Formati on of southeastern Pennsylvania, the Spence Shale of northern Utah, the Wheeler Shale and Marjum Formation of western Utah, the Buen Formation of northern Greenland, and the Burgess Shale of British Columbia. The last is the largest such deposit, containing about 152 mostly monospecific genera of Middle Cambrian age. Equally spectacular is the Chengjiang fauna found at Maotianshan in Yunnan, southwest China, which contains in excess of 70 arthropod-dominated species of Early Cambrian age. However, for extremely fine morphological detail, Orsten-type preservation in the Lower Cambrian of England, the Middle Cambrian of Russia and Australia, and the Upper Cambrian of Poland and Sweden is unsurpassable.

Trilobites. The most abundant remains of organisms in Cambrian rocks are of trilobites (Fig. 3*a–c*). They are present in almost every fossiliferous Cambrian deposit and are the principal tools used to describe divisions of Cambrian time and to correlate Cambrian rocks. These marine arthropods ranged from a few millimeters to 20 in. (50 cm) in length, but most were less than 4 in. (10 cm) long. Although some groups of trilobites such as the Agnostida (Fig. 3*a*) were predominantly pelagic in habitat, most trilobites seem to have been benthic or nektobenthic and show a reasonably close correlation with bottom environments. For this reason, there are distinct regional differences in the Cambrian trilobite faunas of the shallow seas of different parts of the Cambrian world. *See* TRILOBITA.

Brachiopods. The next most abundant Cambrian fossils are brachiopods (Fig. 3*d–f*). These bivalved animals were often gregarious and lived on the sediment surface or on the surfaces of other organisms. Brachiopods with phosphatic shells, referred to the Acrotretida (Fig. 3*f*), are particularly abundant in many limestones and can be recovered in nearly perfect condition by dissolving these limestones in acetic or formic acids. Upper Cambrian limestones from Texas, Oklahoma, and the Rocky Mountains yield excellent silicified shells of formerly calcareous brachiopods when they are dissolved in dilute hydrochloric acid. *See* BRACHIOPODA.

Archaeocyathids. Limestones of Early Cambrian age may contain large reeflike structures formed by an association of algae and an extinct phylum of invertebrates called Archaeocyatha (Fig. 3*m* and *n*). Typical archeocyathids grew conical or cylindrical shells with two walls separated by elaborate radial partitions. The walls often have characteristic patterns of perforations. *See* ARCHAEOCYATHA.

Mollusks and echinoderms. The Cambrian record of mollusks and echinoderms is characterized by many strange-looking forms (Fig. 3*g–l*). Some lived for only short periods of time and left no clear descendants. Representatives of these phyla, such as cephalopods, clams, and true crinoids, which are abundant in younger rocks, are rare in Cambrian rocks; but rostroconch mollusks are known from the Early, Middle, and Late Cambrian at various times in Laurentia, Australia, Siberia, north China, and Korea. Snails, however, are found throughout the Cambrian. Discoveries of primitive clams have been made in Early Cambrian beds, but they are apparently absent from the later record of life for tens of millions of years until post-Cambrian time. *See* ECHINODERMATA; MOLLUSCA.

Corals. Except for rare jellyfish impressions, the Coelenterata were thought to be unrepresented in Cambrian rocks. Corals have now been discovered in early Middle Cambrian rocks in Australia. However, like clams, they are not seen again as fossils until Middle Ordovician time, many tens of millions of years later. *See* CNIDARIA.

Extinction. The stratigraphic record of Cambrian life in Laurentia (North America) shows perhaps five major extinctions of most of the organisms living in the shallow seas. These extinction events form the boundaries of evolutionary units called biomeres (Fig. 2). Their cause, and their presence in the Cambrian records of other continents, is under investigation. At least one of these extinction events, that at the Marjuman-Steptoean boundary, coincides with a large positive carbon isotope anomaly in Laurentia, Australia, south China, and Kazakhstan. However, perhaps it was these periodic disasters that prevented clear continuity in the evolutionary records of many groups and which led, particularly, to the discontinuous records of the echinoderms, corals, and mollusks. *See* ANIMAL EVOLUTION; EXTINCTION (BIOLOGY).

Faunal origin. One major unsolved problem is the origin of the entire Cambrian fauna. Animal life was already quite diverse before Cambrian time. The earliest Cambrian beds contain representatives of more than 20 distinctly different invertebrate groups. All of these have calcified shells, but none of the Precambrian organisms have any evidence of shells. There is still no clear evidence to determine whether shells evolved in response to predation or to environmental stress, or as the result of some change in oceanic or atmospheric chemistry. *See* PRECAMBRIAN.

History. At the beginning of Cambrian time, the continents were largely exposed, much as they are now. Following some still-unexplained event, the seas were suddenly populated by a rich fauna of shell-bearing invertebrates after 3 billion years of

supporting only simple plants and perhaps 100 million years with shell-less invertebrates. *See* PRECAMBRIAN.

Belts of volcanic islands comparable to those of the western Pacific Ocean today fringed eastern Laurentia, the Australian and western Antarctic margins of Gondwana, and southern Siberia. These belts suggest that crustal plates analogous to those of the present day were in motion at that time. Thick evaporites in Siberia and the Middle Eastern and Indian parts of Gondwana suggest regions of warm temperature and high evaporation rate. Absence of significant development of limestones around Baltica suggest that it was a cool region, probably at high latitudes. Near the continental margins of eastern and western Laurentia, on and around Siberia, and on the western Antarctic, eastern Australian, northwestern African, and southern European margins of Gondwana, archaeocyathid bioherms developed and flourished. By the end of Early Cambrian time, archaeocyathids had become extinct, and shell-bearing organisms capable of building bioherms did not reappear until Middle Ordovician time, at least 45 million years later.

Volcanism and evaporitic conditions continued into the Middle Cambrian in Siberia and parts of Gondwana, and evaporites of this age are also known from northern Canada. However, a dramatic change took place in the southern European and northwestern African parts of Gondwana. Carbonate sedimentation virtually ceased throughout that region as those parts of Gondwana reached areas of cooler water and probably higher latitudes. Sea level was rising over much of the world throughout Middle Cambrian time, flooding the interiors of most continents.

In the Late Cambrian, parts of western Baltica and eastern Laurentia began to show signs of crustal deformation suggesting that Iapetus, the ocean between Laurentia, Gondwana, and Baltica, was beginning to close. Crustal deformation was also taking place in southern Siberia, eastern Australia, and western Antarctica. In the broad, shallow seas over all of the contin ents except Baltica and the southern European and northwestern African parts of Gondwana, extensive areas of carbonate sediments developed. At least five times in the shallow seas covering Laurentia, large parts of the animal populations became extinct and had to be replenished from the oceanic regions. The last of these extinction events marks the end of Cambrian time in Laurentia.

Throughout Cambrian time, terrestrial landscapes were stark and barren. Life in the sea was primitive and struggling for existence. Only in post-Cambrian time did the shallow marine environment stabilize and marine life really flourish. Only then did vertebrates evolve and plants and animals invade the land. Allison R. Palmer; John H. Shergold

Bibliography. C. H. Holland (ed.), *Cambrian of the British Isles, Norden and Spitzbergen*, 1974; C. H. Holland (ed.), *Cambrian of the New World*, 1971; C. H. Holland (ed.), *Lower Paleozoic of the Middle East, Eastern and Southern Africa, and Antarctica*, 1981; W. S. McKerrow, C. R. Scotese, and M. D. Brasier, Early continental reconstructions, *J. Geol. Soc.*, 149:559–606, 1992; M. A. McMenamin and D. L. McMenamin, *The Emergence of Animals: The Cambrian Breakthrough*, 1990; A. R. Palmer, A proposed nomenclature for stages and series for the Cambrian of Laurentia, *Can. J. Earth Sci.*, 35(4):323–328, 1998; A. R. Palmer, Search for the Cambrian world, *Amer. Sci.*, 62:216–224, 1974; R. A. Robison and C. Tiechert (eds.), *Treatise on Invertebrate Paleontology*, pt. A: *Biogeography*, 1979; J. A. Secord, *Controversy in Victorian Geology: The Cambrian-Silurian Dispute*, 1986; H. B. Whittington, *The Burgess Shale*, 1985.

Camel

The name given to two species of mammals which are members of the family Camelidae in the order Artiodactyla. These are the bactrian camel (*Camelus bactrianus*) and the Arabian or dromedary camel (*C. dromedarius*). Both species are domesticated, but a few wild herds of bactrian camels are still in existence in the Gobi desert.

The legs of these animals are long and slender and terminate in two toes. The neck and head are elongate, and there is a cleft upper lip. The upper jaw has both canines and incisors, and the dental formula is I 1/3 C 1/1 Pm 3/3 M 3/3 for a total of 36 teeth. These animals do not have a gall bladder and the stomach has three divisions. The period of gestation is about 1 year and the female breeds every second year, producing one young (colt).

Bactrian camel. The bactrian camel (see **illus.**) is stronger and more heavily built than the dromedary and is more suitable as a pack animal. The long, shaggy hair allows it to withstand the cold climate of its range in central Asia. There are two humps of fatty tissue, one over the shoulders and the other atop the hindquarters. The shoulder height, excluding the hump, is about 6 ft (2 m). This animal is economically important to the region as it provides milk, meat, and leather for the nomads.

Bactrian camel (*Camelus bactrianus*). (*Brent Huffman/ Ultimate Ungulate Images*)

Arabian camel. The Arabian camel is taller than the bactrian and has a single hump of fatty tissue, which can be used as a food reserve. There are two varieties of this species found in the desert. One is the baggage camel, used as a beast of burden, which can average 40 mi (65 km) a day carrying a load of about 400 lb (180 kg). The other type is the more slightly built racing camel, which can travel up to 100 mi (160 km) a day, but with a very light load. This species is well suited to desert life with its broad feet, adapted to walking on sand, its ability to close its nostrils completely, and its double row of interlocking eyelashes.

Adaptations. These two species are able to interbreed. They have a most important physiological adapation in their ability to conserve water. Camels do not store water but conserve it, since the body is well insulated by fur and has a temperature range of over 12°F (7°C) before it perspires sufficiently to prevent a further rise. The body temperature is about 93°F (34°C) in the morning and may rise to 106°F (41°C) during daytime activity before water is expended in any appreciable quantity. The camel can lose over 40% of its body water without fear of dehydration. However, although able to survive for long periods without water, it may drink as much as 15 gal (57 liters) when water is available. *See* ARTIODACTYLA.

Charles B. Curtin

Camel's hair

A fine hair known to the American consumer chiefly in the form of high-quality coat fabrics. This textile fiber is obtained from the two-humped bactrian camel, which is native to all parts of Asia. The protective hair covering of the camel is a nonconductor of both heat and cold, also water repellent. In the spring the year's growth of hair, which hangs from the camel in matted strands and tufts, falls off in clumps. This growth, plus the masses of hair shed throughout the year, is the chief source of supply. The camel is sometimes plucked to obtain the down or underhair.

Camel's-hair fabrics are ideal for comfort, particularly when used for overcoating, as they are especially warm and light in weight. Camel's hair is characterized by strength, luster, and smoothness. The best quality is expensive. It is often mixed with wool to improve the quality of the wool fabric. The price of such a mixed cloth is much less than that of a 100% camel's-hair fabric.

In the textile industry camel's hair is divided into three grades. Grade 1 is the soft and silky light-tan underhair. This is short staple or noil of 1–5 in. (2.5–12.5 cm) but is the choicest quality. This was once the only true camel's hair used in the manufacture of apparel. Grade 2 is the intermediate growth, consisting partly of short hairs and partly of coarse outer hairs. Grade 3 consists entirely of coarse outer hairs measuring up to 15 in. (27.5 cm) in length and varying in color from brownish-black to reddish brown. This grade has no value for apparel manufacture; it is suitable only for cordage and for low-quality rugs. *See* ALPACA; CAMEL; CASHMERE; LLAMA; MOHAIR; NATURAL FIBER; VICUNA; WOOL.

M. David Potter

Cameo

A type of carved gemstone in which the background is cut away to leave the subject in relief. Often cameos are cut from stones in which the coloring is layered, resulting in a figure of one color and a background of another. The term cameo, when used without qualification, is usually reserved for those cut from a gem mineral, although they are known also as stone cameos. The commonly encountered cameo cut from shell is properly called a shell cameo.

Most cameos are cut from onyx or agate, but many other varieties of quartz, such as tiger's-eye, bloodstone, sard, carnelian, and amethyst, are used; other materials used include beryl, malachite, hematite, labrodorite, and moonstone. *See* GEM; INTAGLIO (GEMOLOGY).

Richard T. Liddicoat, Jr.

Camera

A device for forming and recording images; the basic tool of photography. In its simplest form, a camera is a lighttight box in which an image is formed through a pinhole or lens at one end onto a light-sensitive material at the opposite end. Most cameras contain an aperture and shutter for controlling the amount of light reaching the light-sensitive material (exposure). The recording material may be a photosensitive silver halide emulsion film, an electronic device such as a charge-coupled device (CCD) or complementary metal-oxide semiconductor (CMOS) sensor, or some other photosensitive material designed for a specific purpose. *See* CHARGE-COUPLED DEVICES.

Types. The camera was originally a lighttight chamber in which observers saw and sometimes traced the image projected by a pinhole and later by a simple biconvex lens (camera obscura); the first portable model using a lens (camera lucida) was introduced in the mid-1600s. Since the first photographic camera was built by Nicephore Niepce in 1826, many different types have been devised. They can be distinguished by function, application, film size, or format. Distinctions are also based on type of viewfinder and body construction.

Cameras for still photography include box, point-and-shoot, view-and-press, roll-film, 35mm, instant-picture, stereo, underwater, and panoramic. Some categories overlap. Still video and digital cameras use electronic sensors instead of film, and store the image in solid-state memory or on magnetic media or optical disks. Motion picture or cine cameras record movement at regular intervals in a series of frames which are later projected onto a screen to create an illusion of movement. Television and video cameras record movement electronically for broadcast and storage on magnetic media or optical disks. Camcorders are video cameras that contain both the image sensor and recording media in a single unit.

Box and simple cameras. These cameras have no (or very few) adjustments and are generally easy to use. They do not require focusing; called focus-free, universal, or fixed-focus, they have a lens with a small aperture. This design maximizes the range of distances from the subject over which sharp images can be formed. This range is called the depth of field. The focus is preset to produce the greatest depth of field, known as the hyperfocal distance. *See* GEOMETRICAL OPTICS; LENS (OPTICS).

Simple cameras rely on the wide latitude (the range of exposure which yields acceptable negatives) of modern photographic films and in particular films of color-negative emulsions. Most have a single shutter speed of 1/30 to 1/125 s and a single aperture, although a limited selection of settings such as indoors, outdoors, or sunny and cloudy may be provided. A film advance mechanism that prevents double exposure and a basic aiming device are usually present. Many cameras can trigger a flash device (**Fig. 1**).

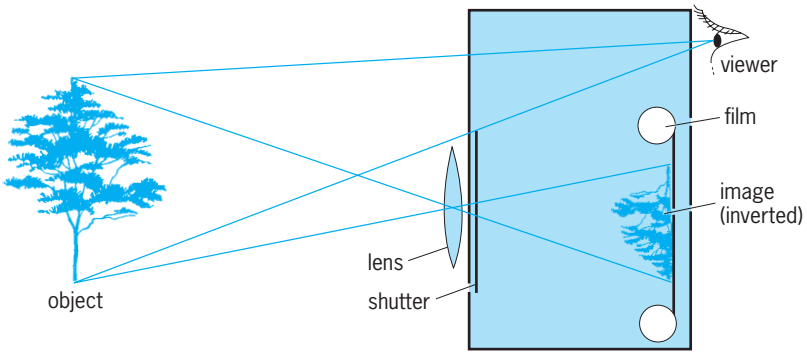

Fig. 1. Simple camera.

Disposable cameras (also known as recyclable or single-use) are simple cameras that come loaded with film and are not reused.

Many cameras have few controls and are as simple to operate as the box cameras, but they have automatic electronic adjustment of the lens opening and shutter speed to produce the correct exposure. These cameras may have a switch to set the film speed, according to the International Standard Organization (ISO), and are often limited to two settings, ISO 100 and ISO 400.

Simple cameras rely on compromises for focus and exposure control. Although they can produce good photographs under average conditions, much better results can be obtained over a wider range of conditions with adjustable focus and exposure.

Zone focusing cameras are similar to the simple cameras described above, except that the focus can be set within limits. Focus settings may exist for close-up, portrait, group portrait, and landscape.

Point-and-shoot cameras. Some cameras operate as box or simple cameras but with automatic electronic adjustments for producing good pictures. These cameras set exposure control automatically. The focus may be fixed or subject to automatic electronic adjustment. These cameras often feature drop-in or automatic film loading and built-in flash. Some have user overrides, which enable the photographer to choose manual settings in extreme situations, such as strong spot or backlighting.

Although the lens is fixed, many models have dual (switchable) focal length or continuously variable focal length (zoom lenses).

Folding cameras. Except for instant-picture cameras, few folding cameras are now being made; a collapsible bellows between the back and the lens permits compact folding for carrying. Folding cameras have been made for most film sizes, especially the larger ones. Some modern cameras use a collapsible or retractable lens which is stored in the camera body when not in use.

View cameras. These large cameras were originally designed for landscapes or views but are equally useful in commercial studio photography. A flexible bellows is attached to front and rear panels that swing on a rigid base. The front panel holds the lens mounted on a board. The rear panel usually contains a ground-glass plate that can be replaced by a film holder. View cameras can be folded for transport; the bellows is primarily functional, connecting the lens board to the film holder, while allowing great freedom of movement between them. These cameras offer a wide range of positions for correcting perspective and controlling the image distortion that results when the subject and film plane are not parallel. Studio or stand cameras used for portraiture are similar in construction, but they may have more limited movements.

Instant-picture cameras. Introduced in 1948 by Edwin Land, these cameras directly produce a print. The development chemicals are contained within the film package in a pod, which is broken by rollers upon ejection of the print from the camera. Instant-picture cameras must be comparatively large, since the full-size image is used. Instant-picture-type backs are made for other cameras, particularly medium- and large-format view cameras, to produce fast proofs to check lighting and composition before making an exposure on conventional film.

Motion picture cameras. Commonly called movie or cine cameras, they photograph moving objects in a series of images for sequential playback to create an illusion of movement. They differ from still cameras principally in the need to advance the film rapidly during the brief time that the shutter closes between frames. Rates of 18 to 24 frames per second are common, with other rates for special purposes such as time-lapse or high-speed photography. The format is usually specified by the width of the film: 8mm and Super 8 (8mm film with slightly larger image and different sprocketing) cameras are amateur formats, popular for home movies; 16mm cameras are used for news, scientific, and some theatrical films; and 35- and 70mm cameras are used primarily for theatrical films. *See* CINEMATOGRAPHY.

Digital or electronic cameras. Photographic technology associated with electronic or digital photography is essentially concerned with using an electronic sensor array instead of film. The optical part of the camera is unchanged as well as the need to keep non-image-forming light away from the photosensitive medium during the making of an exposure. The

fact that photoelectronic recording methods are becoming prevalent also has meant that shuttering can sometimes be accomplished electronically.

The impact of the digital "revolution" is that cameras [point-and-shoot, single-lens reflex (SLR), and medium and large format] are now available with a digital method for capturing and storing images, essentially replacing photographic emulsions.

When records are needed fast, the digital camera can deliver images almost immediately that can be transmitted and distributed to locations worldwide. For the researcher, engineer, and technician, the big advantage is the immediacy of results and the fact that photography can be promptly repeated if not of suitable quality. Digital files, however, can be easily manipulated in a computer, which raises ethical questions.

The resolution of digital systems, some being cameras in their own right while others are simply accessories that can be attached to existing cameras, is close to that of film (especially smaller-size film) and is of suitable quality for many purposes.

Digital or electronic cameras are available in about as many varieties as standard film-type cameras, whether still or motion picture. Large-format cameras are only available as scanning camera backs which are best used with stationary subjects since they capture the whole image by moving one or more linear arrays of photosensors gradually from one side of the image gate to the other. Two-dimensional array cameras permit instantaneous recording of the whole scene and are thus suitable for imaging objects in motion.

A significant problem associated with digital records is the storage capacity requirement both in the camera and for archiving purposes. Film is an excellent and low-cost information storage medium, albeit somewhat susceptible to change over time due to environmental conditions. For the foreseeable future, the digital and film systems will coexist and photographers will be able to decide on the best system for a given application. The future promises development of more hybrid solutions rather than complete elimination of photographic emulsions. The question of sustainable commercial production of specialized silver-halide emulsions is a factor that will determine the ultimate impact of the digital process on traditional processes.

Reflex cameras. Reflex cameras use mirrors to reflect the object scene onto a viewing (and often focusing) screen. The mirror erects the inverted image, making viewing easier, although motion left and right is still reversed in the finder. A camera that uses this mirror with a pair of matched lenses (one for viewing and focusing, the other for picture taking) is called a twin-lens reflex. In such a camera, the two lenses are coupled and the mirror and focusing screen are placed so that a sharp image on the focusing screen corresponds to a sharp image on the film plane (**Fig. 2**).

To eliminate parallax, various methods of viewing through the lens have been used. A beam splitter may be placed behind the lens, channeling part of the image to the viewfinder and part to the film.

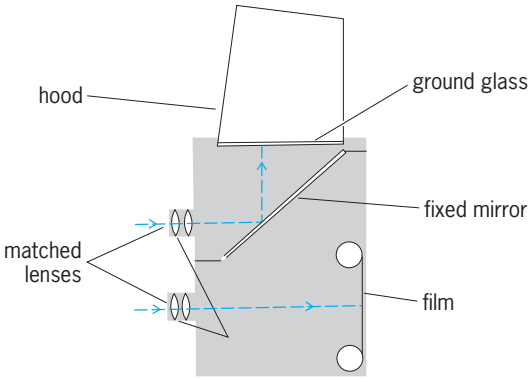

Fig. 2. Twin-lens reflex camera.

This arrangement reduces the efficiency of both the focusing and picture taking. A mirror placed directly behind the lens can reflect the image from it directly to the viewing screen. By removing the mirror from the light path during the exposure, the film receives the full intensity of the image.

The single-lens reflex camera uses the same lens with the aid of a mirror for viewing and photographing (**Fig. 3**). Since picture taking and viewing use the same lens, the single-lens reflex has the advantage that the viewfinder always shows the precise scene to be recorded by the film. In some single-lens reflexes, particularly larger ones, the mirror must be reset manually after exposure, usually by using the film advance lever. Most single-lens reflexes employ an instant-return mirror, which is virtually always in the viewing position but flips out of the way before the exposure. A pentaprism (or a group of mirrors called a Porro prism) may be used between the focusing screen and viewfinder eyepiece to restore proper left-right motion while viewing an image oriented like the subject. This makes it easy to aim the camera and follow moving subjects. Various focusing aids may be present on the built-in focusing screen, including a central split image which is brought into alignment at the point of focus. Some single-lens reflex cameras allow for a change of focusing screen for various purposes.

Autofocusing single-lens reflexes. The semisilvered mirror found in some metering cameras can be used to channel the light to a charge-coupled-device array sensor for detecting focus. The camera electronics can then be used to focus the lens. This type of autofocusing single-lens reflex is widely used by amateurs and some professional photographers and

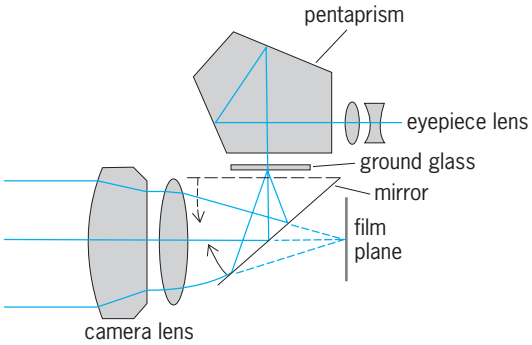

Fig. 3. Single-lens reflex camera.

is a leading feature of current 35mm camera design.

Like the focusing screen, the sensor array must be at the same distance from the mirror as the film plane. The array may detect the intensity maximum or compare two beams to check focus. The result of the measurement may be relayed to the operator via a display in the viewfinder indicating "focus" or "out of focus" or the direction to turn the lens to achieve focus.

The camera itself may focus the lens by using a motor and connecting shaft or by controlling a motor built into the lens.

Other types. Many cameras have been devised for specific purposes. Wide-field or panoramic cameras take extreme wide-angle views, some up to a full 360 degrees. This can be useful in photographing landscapes or large groups of people. In extreme wide-angle photography, some cameras swing the lens during exposure in synchronization with the shutter.

Underwater cameras must be watertight up to a pressure of several atmospheres. Some are specialized in design, although many are conventional cameras placed in a watertight container with linkages to the exterior for control of focus and exposure. *See* UNDERWATER PHOTOGRAPHY.

Aerial cameras may use special mounts to counteract vibration. They may also move the film during exposure to compensate for aircraft motion relative to the ground. These cameras can be used for reconnaissance, mapping, or scientific surveys. *See* AERIAL PHOTOGRAPH.

Stereo cameras have multiple lenses for taking pictures in pairs. When these pictures are viewed simultaneously, each with a different eye, the original three-dimensional depth is created.

Many cameras have been devised for scientific studies. Medical cameras can photograph inside the human body. Streak and slit cameras record images for time-motion studies. Holographic cameras record diffraction patterns instead of real images for reconstruction of three-dimensional subjects. Lenses adapted to ultraviolet or infrared rather than visible light are used with film of corresponding sensitivity. Cameras aboard space satellites gave the first view of the Earth from space and of the dark side of the Moon. *See* HOLOGRAPHY; REMOTE SENSING.

Viewfinders and focusing aids. The image is viewed most accurately on the film plane. The operator places a ground-glass plate at the film plane and adjusts the camera position and lens focus until the desired image appears sharp (and inverted) on the glass, which is replaced by the film in a holder for exposure. This time-consuming process is impractical for many applications, particularly if the subject is moving.

Most cameras, other than large-format view cameras that use a ground glass at the focal plane for composition and focusing, are equipped with an auxiliary framing device called a viewfinder. These viewfinders vary in complexity from open wire-frame devices that, when looked through from a given distance, provide an indication of the approximate field of view reproduced by the camera lens at the film plane, to complex systems that predict more accurately what will be recorded by the camera at the moment of exposure.

In some viewfinders there is also incorporated the ability to focus the lens at some preselected subject detail. In the most complex viewfinder and focusing mechanisms, the focusing of the lens is achieved automatically.

Viewfinders of simple cameras may consist of only a plastic or metal frame that is held up to the eye to approximate the angle of view of the camera lens. In more advanced cameras, one may find a lens or a group of lenses that compensates for the magnification or reduction of the picture-taking lens. Since viewfinders and photographic lenses on a camera are often separate, they do not see the subject exactly from the same position, and this difference in view between the two is called parallax.

Although parallax is not usually a problem when subjects are located at relatively large distances from the camera, when photographing at close range certain parts of a subject that appear in the viewfinder may be cut off in the final photograph.

A major advantage of SLR (single-lens reflex) or DSLR (digital single-lens reflex) cameras, as well as modern "point and shoot" digital cameras that preview the image on an active display such as an LCD screen, is that they do not suffer from parallax because they use the same lens for both viewing and making the photograph. Such cameras also allow the photographer to easily and accurately focus on desired parts of the subject.

Cameras are focused either by changing the distance between lens and film or by moving specific elements within a complex lens, often by turning the outer lens barrel. Some cameras are focused by turning the lens barrel until a pointer is aligned with a point on a scale that represents the distance from the camera to the subject. A rangefinder may be placed on the camera which uses a split image to determine the object's distance. When a lever or dial on the rangefinder is adjusted until the two images coincide, the distance can be read from a scale on the rangefinder. Some cameras connect the rangefinder to the lens via a cam, so that turning the lens adjusts the rangefinder to bring both to a focus at the same point; these are called rangefinder-coupled cameras. *See* RANGEFINDER (OPTICS).

Electronic rangefinding devices built into the camera may be used instead of the optical type. These may measure object distance by timing the reflection of a sound wave emanating from the camera (sonar), or by projecting an infrared beam from a small window on the camera and locating the beam by rotating a mirror behind a second window. By knowing the angle of the mirror and the distance between the mirror and the infrared emitter, the subject distance is electronically calculated by trigonometry. The lens is automatically set to the proper focus or focus zone.

Viewfinders of electronic and video cameras may be simple optical finders like those found in point-and-shoot cameras, but most use electronic viewfinders. Since the electronic sensor at the film plane relays the image information to an electronic recording device, this image can be relayed simultaneously to

a small television monitor. This is usually a miniature liquid-crystal display (LCD) viewed through a magnifying eyepiece.

Exposure control. To produce a photograph or electronic image, the camera must match the image brightness to the sensitivity of the receiving medium or device. When that medium is photographic film, sensitivity is reported as ISO film speed, with each doubling of speed indicating an increase of one photographic stop of sensitivity; that is, one smaller aperture, half the shutter speed, or half the subject illumination would be required to expose the image properly. Electronic sensors may be rated with an equivalent ISO. Image intensity is usually controlled by an iris diaphragm serving as the lens aperture. The ratio of the lens focal length to this aperture is called the *f*/stop. Although technically a function of the lens, in some designs the aperture is located inside the camera. Many cameras can control the aperture in the lens, keeping it at the widest possible setting for bright viewing and focusing of single-lens reflexes and closing the aperture to the proper setting during exposure.

Shutters. The shutter controls exactly when and for how long the light is admitted to the image-forming material. This can be simply a cap that is removed and then replaced over a pinhole; or a metal flap, or a group of such blades (leaf shutter), that swings aside behind the lens to create the exposure; or a slit that moves across the film plane during the exposure (focal-plane shutter). Shutters can provide exposures ranging from several seconds to small fractions of a second. Speeds up to 1/500 s are common for leaf shutters, with focal-plane shutters routinely producing exposures as short as 1/1000 s, and some reaching 1/10,000 s.

Exposure meters. To determine the optimal shutter speed and aperture combination for the film or image sensor, it is necessary to determine the brightness of the scene being recorded. A photosensitive device such as a cadmium sulfide cell or silicon photodiode is often employed to measure illumination. Originally a separate device, the light meter was used to determine *f*/stop and shutter speed combinations for proper exposures. Today this device is often built into the camera, although it only can measure the light reflected from the subject. In some cameras, particularly 35mm point-and-shoot cameras, the cell is placed on the front behind a small lens element designed to approximate the angle of view of the photographic lens. On the most basic cameras, this shows when light is sufficient to take a picture with the recording medium of given sensitivity and the available range of shutter speeds and apertures. Some cameras automatically adjust aperture and shutter speed to obtain the proper exposure as shown by the cell. Other cameras, particularly single-lens reflexes, have an indicator in the viewfinder to show when the shutter speed–aperture combination will give proper exposure. *See* PHOTOCONDUCTIVE CELL; PHOTOELECTRIC DEVICES.

Like viewing, metering through the photographic lens has proved effective and is widely used. Some single-lens reflex cameras have photocells in the mirror box which look forward through the lens and are aimed at the subject. Others use similar cells to measure the intensity of the image (or parts thereof) formed on the focusing screen. Since the image changes with the angle of view provided by the lens, the angle of the metering is proportional to lens coverage and changes with changes of focal length.

Sometimes a semisilvered mirror is used as the instant return mirror. This mirror acts as a beam splitter, allowing a small portion of the light to pass through it to the metering device while reflecting most of the light to the focusing screen. Some cameras have cells facing toward the film plane to monitor the exposure directly from the film. Some cells are intended to measure electronic flash exposures, which are so short that they can be measured only during the actual exposure.

Metering cells may not measure the entire subject evenly. Those that do so, even approximately, are called averaging meters. Meters that concentrate their sensitivity on the center of the frame (where the main subject is often placed) are called center-weighted systems. Some systems measure only a small central portion, in so-called spot metering. Some cameras use more than one metering system. *See* EXPOSURE METER.

Automatic modes. Still cameras control the combination of shutter speed and *f*/stop used for each exposure. (Motion picture and video cameras usually offer less control of shutter speed, which must be higher than the frame rates.) Many combinations of shutter speed and aperture yield the same exposure at the image plane. Different selections produce different photographic results: higher shutter speeds freeze action or reduce blurring caused by camera motion during exposure; smaller apertures yield greater depth of field. The photographer may want the camera to select either setting or both settings automatically.

A camera selecting both shutter speed and aperture is said to be in a programmed autoexposure mode. Some cameras offer a choice of exposure programs, either automatically selected by the camera according to lens focal length, or operator-selected with a preference for greater depth of field or higher shutter speeds.

Some cameras allow the photographer to choose shutter speed or *f*/stop and automatically make the corresponding setting. When the operator sets only the shutter speed, the camera is set to shutter-priority autoexposure. Multimode cameras allow the photographer to select among autoexposure options, often including full manual control, in which the operator sets both *f*/stop and shutter speed.

So-called multimode cameras offer a choice of automated exposure modes. This versatility is especially prominent in 35mm single-lens reflexes. Some cameras offer variations of the basic autoexposure options. Some have autoexposure modes in which the photographer, instead of selecting the actual shutter speed, selects the minimum shutter speed acceptable, and the camera acts as if it were in a special program mode if there is enough light for a higher shutter speed.

Automatic multipattern metering. By dividing the scene into sections, the camera can often automatically set the best exposure in situations where a conventional, single photocell might err, as with a subject lit in silhouette or a strong spotlight. Several variations exist, but all measure the light in the center of the subject and in one or more adjoining areas. By comparing combinations of specific values and checking illumination differences in adjoining sections (and sometimes considering the subject's position and distance from the autofocusing system), the best estimate for proper exposure setting is made electronically.

Auxiliary equipment. Many cameras accommodate a variety of lenses. Those with focal lengths close to the diagonal of the camera format are considered normal lenses, and prints made with them and viewed from the proper viewing distance produce a sensation of perspective that matches well that seen by the unaided eye when examining the original scene. Lenses with shorter focal lengths are wide-angle and may give coverage up to 100°. Telephoto lenses have focal lengths longer than the format diagonal. The narrow angle of view increases the image magnification, allowing photography of small details at greater distances.

A telephoto lens is a photographic lens system specially designed to give a large image of a distant object in a camera of relatively short focal length. A telephoto lens generally consists of a positive lens system and a negative lens system, separated by a considerable distance. If color correction is desired, each of the partial systems must be color-corrected. It is usually not easy to correct distortion in a teleobjective, but occasionally it has been achieved. *See* FOCAL LENGTH; LENS (OPTICS).

Zoom lenses are made of a system of lenses in which two or more parts are moved with respect to each other to obtain a continuously variable focal length and hence magnification, while the image is kept in the same image plane.

If the diaphragm is opened at the same time so that its linear opening increases with the focal length, it is possible to keep the relative aperture of the whole system constant while the focal length varies. In any case, the system must be constructed so that the errors do not vary too much in the shifting. Thus the designer must strive for a design in which the image errors are small, at least for the beginning and end positions of the zooming procedure, and do not become too large for intermediate positions.

In general, a complicated cam is needed to control the motions of the parts of the system. However, it is possible to simplify the mechanism if the focal plane is not kept precisely constant but only required to coincide exactly with a given plane for several focal lengths while approximately coinciding for others. Such a system is called an optically compensated varifocal system.

Some early variable-focal-length lenses contained 15 or more separated elements, but this number has been reduced to as few as four. The zoom ratio (the ratio of maximum to minimum power) is in general 3:1, but it has been possible in some designs to increase the range to 4:1 or even more. Moving sets of achromatic prisms are sometimes used to achieve the zooming effect.

A flash unit produces a brief burst of bright light to allow photography under dim illumination. The camera must open the shutter fully during the brightest portion of the flash. An electronic flash can have various durations. Some cameras, particularly single-lens reflexes, can measure the subject through the lens and then control the flash via a special contact. *See* STROBOSCOPIC PHOTOGRAPHY.

Tripod stands hold cameras steady, giving sharp pictures at longer shutter speeds than are possible with hand-held cameras. Winders and motor drives allow still cameras to take short sequences at rates from one to almost six frames per second.

Special backs used on some cameras imprint digital information on or between the negatives. Accessories can fire the cameras automatically at specified intervals (intervalometers) or for specific events (remote triggers). Some backs can even change the film in midroll. *See* PHOTOGRAPHY.

Andrew Davidhazy; Lawrence R. White

Bibliography. J. Ciaglia, *Introduction to Digital Photography*, 2d ed., 2005; R. Jacobson et al., *Manual of Photography: Photographic and Digital Imaging*, 9th ed., 2000; S. F. Ray, *Applied Photographic Optics*, 3d ed., 2002.

Camerata

An extinct subclass of stalked Crinoidea; a class of the phylum Echinodermata. The Camerata include more than 235 genera confined to the Paleozoic Era, ranging from the Early Ordovician to the Late Permian. Camerate crinoids were one of the four major subclasses of Paleozoic crinoids. Others include the Cladida, Flexibilia, and Disparida. *See* CRINOIDEA; ECHINODERMATA; GEOLOGIC TIME SCALE.

General morphology. Stalked crinoids were elevated above the sea floor by a column with an attachment structure called the holdfast. Camerate crinoids are characterized by a rigid, boxlike calyx composed of many polygonal plates. The lower arm plates (fixed brachials) were solidly incorporated into the upper portion of the theca, and these fixed arm plates were separated by small interradial plates (see **illustration**). The free arms were articulated to the uppermost fixed arm plate, and all except the more primitive Ordovician forms had free arms with biserial brachials, that is, brachials composed of two articulating columns of arm plates. Camerates are also distinctive because the oral part of the calyx is roofed over, covering the mouth and proximal food grooves. Many camerate crinoids had a long, slender, solid anal tube with a terminal anal opening. More derived, younger forms commonly have fewer plates fixed into the calyx, and in extreme instances the free arms may attach directly to radial plates. The morphology of these derived camerates is convergent toward the morphology of cladid crinoids.

Two clades exist among the Camerata: dicyclic forms (order Diplobathrida; more than 60 genera)

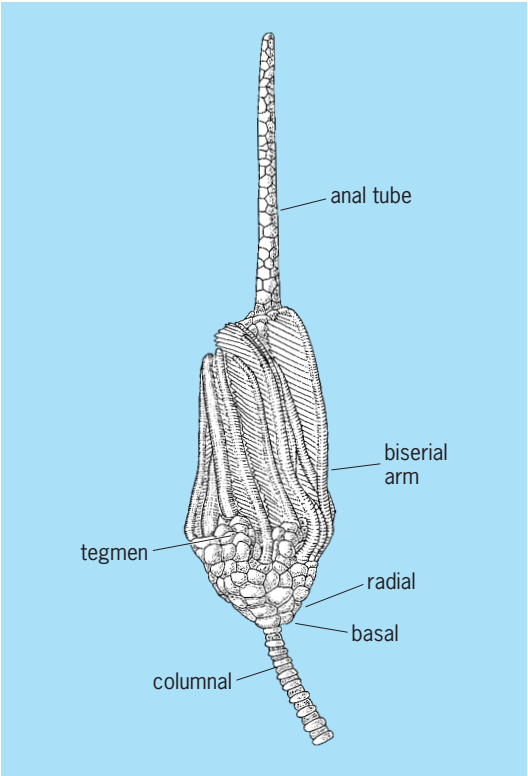

Thylacocrinus, **a Devonian dicyclic camerate crinoid.**

and monocyclic forms (order Monobathrida; more than 175 genera). The lower portion of the dicyclic calyx is composed, from bottom to top, of infrabasal plates, basal plates, and radial plates, whereas the monocyclic calyx has only basal plates and radial plates.

History. The origin of the camerate crinoids is a matter of current debate, with the oldest forms being from the Early Ordovician. Camerate crinoids were especially dominant forms from the Silurian through the Mississippian, where they were very abundant in various shallow, tropical environments (reefs and broad platforms) that deposited limestones. Diplobathrid camerate crinoids became extinct during the middle Mississippian, and monobathrid camerates became extinct during the Late Permian. *See* MISSISSIPPIAN.

N. Gary Lane; William I. Ausich

Bibliography. H. Hess et al., *Fossil Crinoids*, Cambridge University Press, 2002; R. C. Moore and C. Teichert (eds.), *Treatise on Invertebrate Paleontology*, Part T: *Echinodermata 2*, vols. 1 and 2, 1978.

Campanulales

An order of flowering plants, division Magnoliophyta in the subclass Asteridae of the class Magnoliopsida (dicotyledons). It consists of 7 families and about 2500 species. The order is distinguished in this subclass by its chiefly herbaceous habit, alternate leaves, inferior ovary, and stamens which are free from the corolla or attached at the base of the tube. About 2000 of the species belong to the single family Campanulaceae. This family and two others in the order have a specialized pollen presentation mechanism. The anthers converge or unite around the young style, which grows up through the anther tube and pushes out the pollen. Several familiar ornamentals, including the Canterbury bell (*Campanula medium*) and the cardinal flower (*Lobelia cardinalis*), belong to the Campanulaceae. *See* ASTERIDAE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; ORNAMENTAL PLANTS; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Camphor

A bicyclic, saturated terpene ketone. It exists in the optically active dextro and levo forms, and as the racemic mixture of the two forms. All of these melt within a degree of 178°C (352°F). The principal form is *dextro*-camphor, which occurs in the wood and leaves of the camphor tree (*Cinnamomum camphora*). Taiwan is the chief source of natural camphor which is distilled from the wood of this tree. Camphor is also synthesized commercially on a large scale from pinene which yields mainly the racemic variety. The structural formula of the molecule is shown below.

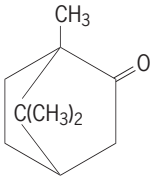

Camphor has a characteristic odor; it crystallizes in thin plates and sublimes readily at ordinary temperatures.

Camphor has use in liniments and as a mild rubefacient, analgesic, and antipruritic. It has a local action on the gastrointestinal tract, producing a feeling of warmth and comfort in the stomach. It is also used in photographic film and as a plasticizer in the manufacture of plastics. *See* KETONE; PINE TERPENE; TERPENE.

Everett L. Saul

Bibliography. T. W. Solomons, *Organic Chemistry,* 5th ed., 1991.

Camphor tree

The plant *Cinnamomum camphora*, a member of the laurel family (Lauraceae) and a native of China, Japan, and Taiwan (see **illus.**). The tree grows to a height of 40 ft (12 m), is dense-topped, and has shiny, dark, evergreen leaves. It is widely planted as an ornamental tree. All parts of the tree contain camphor, as essential oil which is obtained from the finely ground wood and leaves by distillation with steam. The crude camphor crystallizes on the surface of the still. After it is removed and refined, it is the commercial gum camphor, which is used medicinally, in

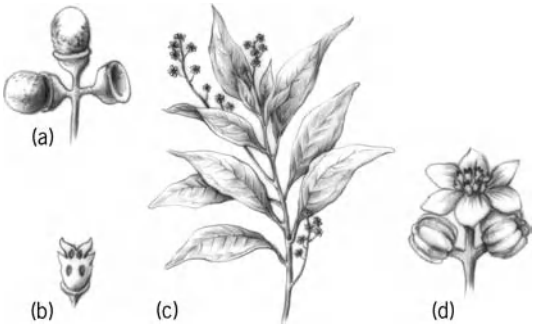

Camphor tree (*Cinnamomum camphora*). (*a*) Fruits. (*b*) Stamen. (*c*) Branch. (*d*) Flower.

perfumes and in the manufacture of celluloid and nitrocellulose compounds. *See* CAMPHOR; MAGNOLIALES. Perry D. Strausbaugh; Earl L. Core

Camptostromatoidea

A small class of primitive echinoderms (subphylum Echinozoa) known from the single species *Camptostroma roddyi* based on about 200 specimens from the Early Cambrian (*Bonnia-Olenellus* Zone) in southeastern Pennsylvania. *Camptostroma* was originally described as a hydrozoan or jellyfish, but it was recognized as an echinoderm and the new class Camptostromatoidea was set up for it. *Camptostroma* has been reinterpreted as an early edrioasteroid, and as a "stem echinoderm" ancestral to several other groups including edrioasteroids. Because of *Camptostroma*'s puzzling morphology, different reconstructions of it have been made (see **illus.**). It has a conical or domal theca or body with a nearly circular outline, divided into a pleated lower theca with larger and smaller overlapping plates, and a domed upper theca made up of large plates with pores on their margins surrounded by numerous smaller plates. The upper theca bears five straight to curved ambulacra protected by numerous cover plates and radiating from the central mouth, and a small anal pyramid at one edge. *Camptostroma* probably had the lower theca either embedded in the soft sediment or attached to objects lying on the sea floor, making it a stationary or attached, low-level suspension feeder using tube feet in the ambulacra to capture small food particles drifting by the theca. The unusual thecal plating is the main feature separating *Camptostroma* from later Edrioasteroidea, which it otherwise resembles, but this genus may also have been ancestral to other classes such as the Eocrinoidea. *See* ECHINODERMATA; ECHINOZOA; EDRIOASTEROIDEA. James Sprinkle

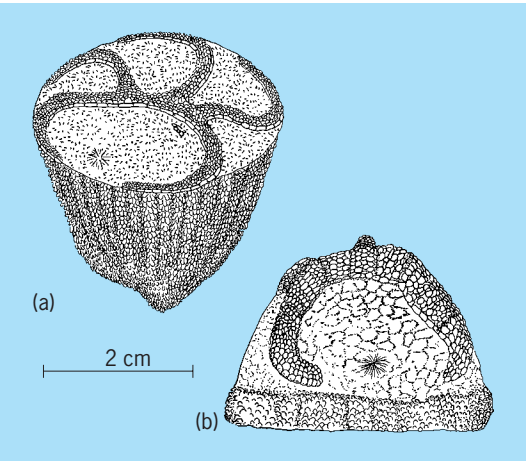

***Camptostroma roddyi* from the Early Cambrian of Pennsylvania. (*a*) Reconstructed oblique side view showing theca with conical lower part inferred to have been partly buried in the sea floor (*after C. R. C. Paul and A. B. Smith, Biol. Rev., 59:451, 1984*). (*b*) Reconstructed side view showing domed theca with flat lower surface inferred to have been attached to the sea floor or to objects on it (*after K. Derstler, Ph.D. dissertation, University of California, Davis, p. 120, 1985*).**

Bibliography. K. Derstler, Ph.D. dissertation, University of California, Davis, 1985; J. W. Durham, *J. Paleontol.*, 40:1216–1220, 1966; C. R. C. Paul and A. B. Smith, *Biol. Rev.*, 59:443–481, 1984.

Canal

An artificial open channel usually used to convey water or vessels from one point to another. Canals are generally classified according to use as irrigation, power, flood-control, drainage, or navigation canals or channels. All but the last type are regarded as water conveyance canals.

Canals may be lined or unlined. Linings may consist of plain or reinforced concrete, cement mortar, asphalt, brick, stone, buried synthetic membranes, or compacted earth materials. Linings serve to reduce water losses by seepage or percolation through pervious foundations or embankments and to lessen the cost of weed control. Concrete and other hard-surface linings also permit higher water velocities and, therefore, steeper gradients and smaller cross sections, which may reduce costs and the amount of right-of-way required.

Water conveyance canals. The character of material along the bottom and sides of a canal must be considered in determining whether lining is required. The velocity in an irrigation canal must be sufficiently low to avoid erosion of the canal banks and bottom, and it must be high enough to prevent deposition of silt: Velocities of 1.5–3.5 ft/s (0.46–1.1 m/s) are normally used for earth canals, and about 7 ft/s (2 m/s) is the usual maximum for concrete-lined irrigation canals. The velocity depends on canal slope, boundary roughness, and proportions of the cross-sectional area. Canal capacity Q in ft^3/s may be expressed by the formula $Q = AV$, where A is the cross-sectional area in square feet, and V is the average velocity in ft/s.

Most irrigation canals are trapezoidal in cross section. Their side slopes are generally either 1.5:1 or 2:1 (horizontal to vertical) depending on the type of soil, but flatter slopes are used in unstable materials. Side slopes of concrete-lined irrigation canals are generally 1.25:1 or 1.5:1. They occasionally may be as high as 2:1, as for example in the San Luis Canal

Fig. 1. Segment of the San Luis irrigation canal in California. A major waterway in the western United States, it is 103 mi (165 km) long, 200 ft (60 m) wide, and 36 ft (11 m) deep. (*U.S. Bureau of Reclamation and State of California*)

(**Fig. 1**). Freeboard, or the vertical distance between maximum water surface and top of canal bank, generally ranges from 1 ft (0.3 m) for small earth laterals to 4 ft (1.2 m) for earth canals. Freeboard for concrete-lined canals ranges from 0.5 to 3.5 ft (0.15 to 1.1 m) to the top of the lining, depending on capacity.

Irrigation canal intakes usually consist of a headworks and a river control structure with some form of intake gates to regulate the quantity of water admitted or stop the flow into the canal. Frequently, when the canal diverts from a stream that transports considerable sediment, the intake may be designed to keep out as much sediment as possible. This is accomplished by (1) locating the canal intake on the stream where the bed load is the smallest; (2) using weirs that divert water from the less silty upper layers of the flowing stream called skimming weirs; or (3) interposing basins in which much of the sand and silt are deposited before the water enters the canal. The deposited material is removed from these settling or desilting basins by occasional sluicing or dredging.

Fig. 2. Walter F. George Lock, Dam and Powerhouse on the Chattahoochee River, Alabama-Georgia boundary. The lock chamber can be seen at right center of the photograph. (*U.S. Army Corps of Engineers*)

Irrigation canals usually require check structures in them to regulate the elevation of the water on the upstream side and wasteways as a safety device to carry off surplus water. The checks hold back the water in case of a break in the canal banks and retard the flow when the canal is being emptied to prevent a sudden evacuation and resulting uplift in the lining or sloughing of saturated banks. Hard-surfaced linings are frequently protected from uplift by automatic relief valves placed in the bottom and sides of the canal. *See* IRRIGATION (AGRICULTURE).

Although generally similar to irrigation canals in design, power, flood-control, and drainage canals or channels have special requirements. Power canals are placed on minimum grade to conserve head for power development; hence velocities are usually low. Rectangular wooden, concrete, or steel flumes

with high ratio of depth to width are often used for this purpose. Because power canals are subject to sudden changes in flow, overflow wasteways are usually provided as well as ample freeboard.

Banks of flood-control channels constructed in earth are lined with grass or rock riprap for protection against erosion by flood flows. Grass protection is used for velocities from about 3 to 8 ft/s (0.9 to 2.4 m/s), and riprap for velocities of 8 to 18 ft/s (2.4 to 5.5 m/s). Either rectangular or trapezoidal concrete-lined channels are used for velocities exceeding 18 ft/s (5.5 m/s), although in some cases concrete lining is used for lower velocities to reduce the size of channel.

Drainage canals are deeply excavated to facilitate the drainage of surrounding land. They usually have a minimum grade and small depths of water flow.

Navigation canals and canalized rivers. Navigation canals are artificial inland waterways for boats, barges, or ships. A canalized river is one that has been made navigable by construction of one or more weirs or overflow dams (**Fig. 2**) to impound river flow, thereby providing navigable depths. Locks may be built in navigation canals and canalized rivers to enable vessels to move to higher or lower water levels.

Navigation canals are often built along portions of canalized rivers or located so as to connect two such rivers. They are adapted to the topography by a series of level reaches connected by locks. Sea-level navigation canals, connecting two tidal bodies of water, are excavated sufficiently deep to preclude the need for locks, if the tidal flow permits.

The dimensions of a navigation canal are determined primarily by the size, and to some extent by the speed, of the vessels that are to use it. Depth must be sufficient to assure bottom clearance under all operating conditions, and ordinarily the width allows vessels to pass each other safely. The canal cross section is usually trapezoidal, with side slopes ranging from 1.5:1 to 3:1 or flatter, depending on the stability of the bank material. Sections cut in rock may have vertical or near-vertical sides. Earth banks of some navigation canals are protected against erosion from wave action by placing rock riprap or similar protection near the water surface. *See* INLAND WATERWAYS TRANSPORTATION.

Locks. A lock (Fig. 2) is a chamber equipped with gates at both upstream and downstream ends. Water impounded in the chamber is used to raise or lower a vessel from one elevation to another. The lock chamber is filled and emptied by means of filling and emptying valves and a culvert system usually located in the walls and bottom of the lock.

After a vessel enters the lock chamber, the afterward gates are closed and the water level is lowered (if the vessel is headed downstream) by operation of the emptying valve, or raised (if the vessel is headed upstream) by operation of the filling valve. When the water level in the chamber reaches the water level forward of the vessel, the forward set of gates is opened and the vessel leaves the lock.

Maximum lift of a lock is the vertical distance from the normal pool upstream of the lock to the low-water surface downstream of the lock. Low lifts simplify design problems, but generally in developing a major waterway it will be more economical to use fewer higher-lift locks. Lock lifts vary from a few feet in tidal canals to over 100 ft (30 m) in major rivers, such as the Columbia and Snake. Lock widths vary from 56 to 110 ft (17 to 33 m) and usable lengths vary from 400 to 1200 ft (120 to 370 m), except that smaller locks are used when all the traffic consists of small craft.

On some canalized rivers having low-lift locks, dams contain movable sections which are lowered during periods of moderate or high flow to permit the unobstructed passage of vessels and barges over the dam. The locks are used in these rivers only during periods of low flow when the movable sections of the dams are closed to create sufficient depth of water for navigation. *See* OPEN CHANNEL.

Corps of Engineers; Bureau of Reclamation

Bibliography. A. J. Chadwick and J. C. Morfett, *Hydraulics in Civil and Environmental Engineering*, 3d ed., 1998; H. W. King and E. F. Brater, *Handbook of Hydraulics*, 7th ed., 1996; J. K. Vennard and R. L. Street, *Elementary Fluid Mechanics*, 7th ed., 1995.

Cancer (constellation)

The Crab, a northern zodiacal constellation, visible in winter evenings (see **illustration**). It is the source of the terrestrial name Tropic of Cancer, for the parallel

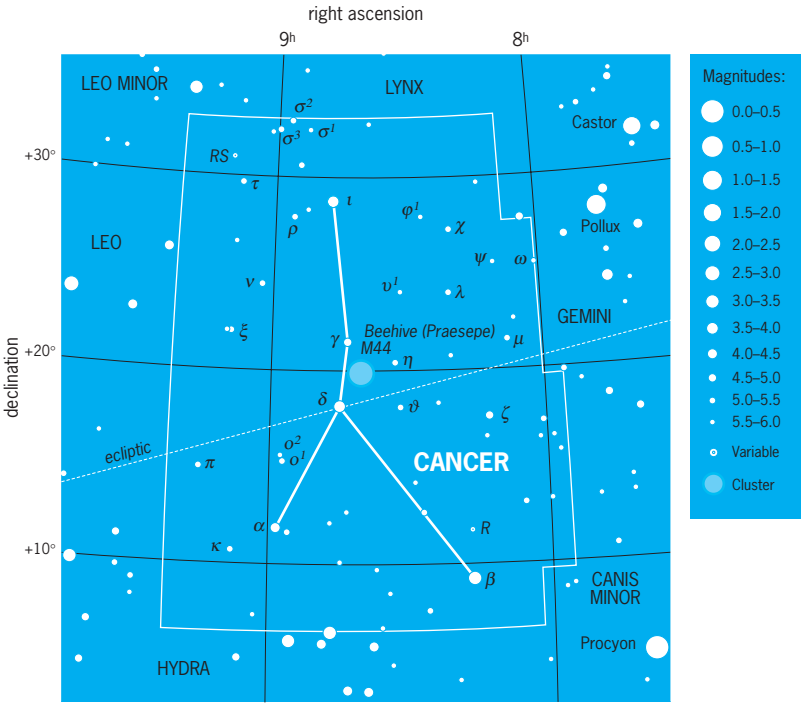

Modern boundaries of the constellation Cancer, the Crab. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

of latitude, about $23^1/_2{}^\circ$ north of the Equator, where the Sun is 90° overhead at noon on the day of the summer solstice in the Northern Hemisphere (about June 21). The summer solstice used to occur when the Sun was in Cancer, but because of the precession of the equinoxes this point has moved to near the modern boundary between Gemini and Taurus (just within Taurus since 1875). *See* GEMINI; PRECESSION OF EQUINOXES; TAURUS; TROPIC OF CANCER; ZODIAC.

Though Cancer has no bright stars, it boasts of an open cluster of stars, Praesepe, the Beehive Cluster, visible to the eye as a hazy spot. In his 1610 book, *Sidereus Nuncius* (The Starry Message), Galileo reported how his newly invented telescope discovered that the haze was really a cluster of stars. *See* STAR CLUSTERS.

The modern boundaries of this and the other 87 constellations were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Cancer (medicine)

The common name for a malignant neoplasm or tumor. Neoplasms are new growths and can be divided into benign and malignant types, although in some instances the distinction is unclear. The most important differentiating feature is that a malignant tumor will invade surrounding structures and metastasize (spread) to distant sites whereas a benign tumor will not. Other distinctions between benign and malignant growth include the following: malignancies but not benign types are composed of highly atypical cells; malignancies tend to show more rapid growth than benign neoplasms, and are composed, in part, of cells showing frequent mitotic activity; and malignant tumors tend to grow progressively without self-limitation. *See* MITOSIS; TUMOR.

Malignant neoplasms that arise from cells of mesenchymal origin (for example, bone muscle, connective tissue) are called sarcomas. Those that develop from epithelial cells and tissues (for example, skin, mucosal membranes, and glandular tissues) are termed carcinomas. Carcinomas usually metastasize initially by way of lymphatic channels, whereas sarcomas spread to distant organs through the bloodstream.

Etiology. The cause of most types of human cancers is unknown. However, a number of factors are thought to be operative in the development of some malignant neoplasms. Genetic factors are thought to be causally related to some human malignancies such as lung cancer in that the incidence of cancer among persons with a positive family history of cancer may be three times as high as in those who do not have a family history. A number of different neoplasms are known to be genetically related and may be due to damage or changes in chromosome structure. These include such neoplasms as retinoblastoma (an eye tumor), pheochromocytoma and medullary thyroid carcinoma (adrenal and thyroid neoplasms), and several types of neoplasms that occur in people who have genetically caused immune deficiency syndromes; in these individuals there is a marked increase in lymphoma (cancer of the lymph glands). *See* IMMUNOLOGICAL DEFICIENCY; LYMPHOMA.

Radiation in various forms is thought to be responsible for up to 3% of all cancers. Persons who receive radiation therapy for some other type of malignant neoplasm, or who are exposed to ionizing radiation from atomic blasts or accidents, have an increase in leukemia. Also, people who receive radiation to the neck have a higher incidence of thyroid carcinoma. Electromagnetic radiation from the Sun (solar radiation) is associated with a higher incidence of many types of skin cancer, including malignant melanoma. There has been a distinct and growing concern about naturally occurring radon gas, from the decay of radioactive rocks in the Earth's crust, concentrated in certain areas of the world such as the Redding Prong in the northeastern United States. Radon gas has been stated to cause up to 10% of all lung cancers, a finding that would be consistent with studies showing that certain types of lung cancer are increased in uranium miners. *See* LEUKEMIA; RADIATION BIOLOGY.

In the United States the carcinogens in tobacco account for up to one-third of all cancer deaths in men and 5–10% in women. The increasing incidence and death rate from cancer of the lung in women is alarming, and is directly related to the increasing prevalence of cigarette smoking by women. There is absolutely no doubt that the single most important action that could be taken to reduce the incidence of cancer in the United States would be to markedly reduce or to eliminate cigarette smoking. Cigarette smoking and the heavy consumption of ethyl alcohol appear to act synergistically in the development of oral, esophageal, and gastric cancers.

There are several carcinogens to which people are exposed occupationally that result in the development of cancer, although the mechanisms by which they cause neoplasms are sometimes poorly understood. For example, arsenic is associated with lung, skin, and liver cancer; asbestos causes mesotheliomas (cancer of the pleural, peritoneal, and pericardial cavities); benzene causes leukemia; and vinyl chloride causes angiosarcomas (tumors of blood vessels) of the liver. *See* MUTAGENS AND CARCINOGENS.

Certain drugs and hormones have been found to cause certain types of neoplasms. Postmenopausal women taking estrogen hormones have a much higher incidence of endometrial cancer (cancer of the lining of the uterine cavity). Alkylating agents that are used in the treatment of a variety of cancers have been shown to cause acute leukemia. Persons who receive organ transplants and are treated with immunosuppressive drugs, such as azathioprine, have an increased incidence of cancer of the lymphoid cells (lymphomas). *See* IMMUNOSUPPRESSION.

The role of diet and nutrition in the development of malignant tumors is controversial and still under investigation. Some epidemiologic studies have shown that certain diets, such as those high in saturated fats, are associated with an increased incidence of certain types of neoplasm, such as colon cancer. Vitamin A–containing compounds have also been shown epidemiologically to prevent some cancers, including lung cancers. In addition, some substances, such as nitrites, used to cure such foods as bacon, are causally associated with gastrointestinal cancers.

The role of viruses in the development of human cancers is being studied. A type of lymphoma that occurs primarily in southern Japan is known to be caused by a retrovirus referred to as human T-cell lymphoma and leukemia virus type I (HTLV-I). Epstein-Barr virus, a deoxyribonucleic acid (DNA) virus of the herpes family, is closely associated with African Burkitt's lymphoma and a tumor of the nasopharynx (nasopharyngeal carcinoma). There is an increased incidence of cancer of liver cells in people who are infected with hepatitis B virus. Whether hepatitis B virus directly causes the tumor or is one of many factors in the development of this type of tumor is uncertain. Certain strains of human papilloma virus (HPV) are thought to cause some cases of uterine cervical carcinoma. *See* EPSTEIN-BARR VIRUS; HERPES; RETROVIRUS; TUMOR VIRUSES.

It is generally accepted that the neoplastic condition is caused by alterations in genetic mechanisms involved in cellular differentiation. In malignant cells, normal cellular processes are bypassed due to the actions of a select group of genes called oncogenes which regulate cellular activities. A group of these highly conserved genes exist in normal cells and are called proto-oncogenes. These genes appear to be important in regulating cellular growth during embryonic development. Although these genes were originally identified in oncogenic retroviruses, proto-oncogenes have been found in normal tissue and also have been identified in a variety of human cancers. Increased amounts of these proto-oncogenes have been seen in small-cell undifferentiated cancer of the lung, colon cancer, breast cancer, and lymphomas. Since proto-oncogenes seem to be important in embryonic growth, it is not surprising that in the process of carcinogenesis these genes can be expressed and can lead to the proliferation of malignant cells. It is thought that in carcinogenesis these proto-oncogenes become unmasked or changed during the breakage or translocation of chromosomes. These genes that were previously suppressed in the cell then become functional, and in some instances lead to the excessive production of growth factors which could be important in the neoplastic state. *See* ONCOGENES.

Incidence. The incidence of malignancies varies with regard to race, geographical location, sex, age, hereditary factors, and socioeconomic status. In certain parts of Africa, for example, liver cancer is the most common malignant neoplasm, whereas in the United States lung cancer in men and breast cancer in women are the most common neoplasms. There has been a marked rise in the incidence of lung cancer in both men and women due to cigarette smoking, whereas most other common tumors, such as cancer of the colon and stomach, have remained stable or even decreased.

Signs, symptoms, and staging. The physical changes that cancer produces in the body vary considerably, depending on the type of tumor, location, rate of growth, and whether it has metastasized. The American Cancer Society has widely publicized cancer's seven warning signals: (1) a change in bowel or bladder habits; (2) a sore that does not heal; (3) unusual bleeding or discharge; (4) a thickening or lump in the breast or elsewhere; (5) indigestion or difficulty in swallowing; (6) an obvious change in a wart or mole; and (7) a nagging cough or hoarseness.

In current medical practice, most cancers are staged according to tumor size, metastases to lymph nodes, and distant metastases. This type of staging is useful in determining the most effective therapy and the prognosis.

Treatment. The progression, or lack thereof, of a given cancer is highly variable and depends on the type of neoplasm and the response to treatment. Treatment modalities include surgery, chemotherapy, radiation therapy, hormonal manipulation, and immunotherapy. In general, each type of cancer is treated very specifically, and often a combination of the various modalities is used, for example, surgery preceded or followed by radiation therapy. The response to treatment depends on the type of tumor, its size, and whether it has spread. *See* CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; IMMUNOTHERAPY; ONCOLOGY; RADIOGRAPHY.

Samuel P. Hammar

Bibliography. V. T. DeVita, Jr., et al., *Cancer: Principles and Practices of Oncology*, 5th ed., 1997; V. T. DeVita, Jr., et al. (eds.), *Important Advances in Oncology*, 1996; J. F. Holland and E. Frei, *Cancer Medicine*, 4th ed., 1999.

Cancrinite

A family of minerals, related to the scapolite family, characteristically occurring in basic rocks such as nepheline syenites and sodalite syenites. A typical specimen is shown in the **illustration**. Cancrinite is hexagonal, $a = 1.28$, $c = 0.52$ nanometers respectively, space group $P6_3$, and is based on a three-dimensional linkage of $[AlO_4]$ and $[SiO_4]$ tetrahedra. Four-, six-, and twelve-membered aluminosilicate rings can be discerned in the structure. Large anions such as $[SO_4]^{2-}$ and $[CO_3]^{2-}$ occur in the hexagonal channels of the structure. The cancrinite member is white, yellow, greenish, or reddish. It has perfect prismatic cleavage, hardness is 5–6 on Mohs scale, and the specific gravity is 2.45.

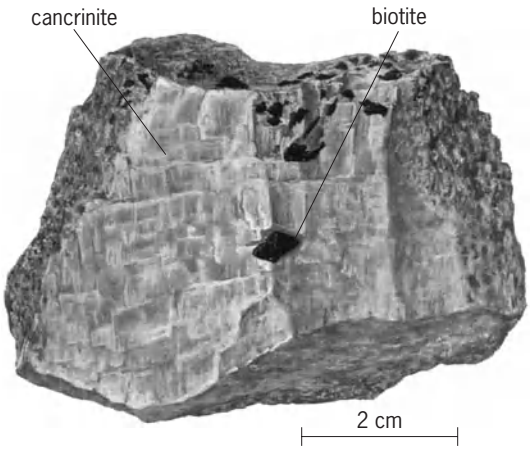

Cancrinite with biotite in nepheline syenite rock, Bigwood Township, Ontario, Canada. (*Specimen from Department of Geology, Bryn Mawr College*)

Compositionally, four members of the cancrinite family are:

Cancrinite	$Na_6Ca_2[AlSiO_4]_6(CO_3)(OH)_2$
Vishnevite	$Na_6Ca_2[AlSiO_4]_6(SO_4)(OH)_2$
Haüyne	$K_2Na_4Ca_2[AlSiO_4]_6(SO_4)(OH)_2$
Afghanite	$Na_6Ca_2[AlSiO_4]_4(SO_4)(OH)Cl$

The cancrinite structure has been proposed as a possible crystalline storage site for nuclear wastes, in particular the gases. Cancrinite can be derived from reaction of far more common nepheline with calcite.

Localities include the Fen area, southern Norway; the Kola Peninsula, Soviet Union; Bancroft, Ontario, Canada; and Litchfield, Maine. *See* FELDSPATHOID; SILICATE MINERALS. Paul B. Moore

Candlepower

Luminous intensity expressed in candelas. The term refers only to the intensity in a particular direction and by itself does not give an indication of the total light emitted. The candlepower in a given direction from a light source is equal to the illumination in footcandles falling on a surface normal to that direction, multiplied by the square of the distance from the light source in feet. The candlepower is also equal to the illumination of metercandles (lux) multiplied by the square of the distance in meters.

The apparent candlepower is the candlepower of a point source which will produce the same illumination at a given distance as produced by a given light source.

The mean horizontal candlepower is the average candlepower of a light source in the horizontal plane passing through the luminous center of the light source.

The mean spherical candlepower is the average candlepower in all directions from a light source as a center. Since there is a total solid angle of 4π (steradians) emanating from a point, the mean spherical candlepower is equal to the total luminous flux (in lumens) of a light source divided by 4π (steradians). *See* LUMINOUS INTENSITY; PHOTOMETRY. Russell C. Putnam

Canine distemper

A fatal viral disease of dogs and other carnivores, with a worldwide distribution. Canine distemper virus has a wide host range; most terrestrial carnivores are susceptible to natural canine distemper virus infection. All animals in the families Canidae (for example, dog, dingo, fox, coyote, wolf, jackal), Mustelidae (for example, weasel, ferret, mink, skunk, badger, stoat, marten, otter), and Procyonidae (for example, kinkajou, coati, bassariscus, raccoon, panda) may succumb to canine distemper virus infection. Members of other Carnivora families, including domestic cats and swine, may become subclinically infected. The virus has also been isolated from large cats (lions, tigers, leopards) that have died in zoological parks in North America, from wild lions in the Serengeti National Park (Tanzania), and from wild javelinas (collared peccaries). *See* CARNIVORA.

Etiologic agent. Canine distemper virus is classified as a morbillivirus within the Paramyxoviridae family, closely related to measles virus and rinderpest virus of cattle and the phocine (seal) and dolphin distemper virus. The virus is enveloped with a negative-sense ribonucleic acid and consists of six structural proteins: the nucleoprotein and two enzymes in the nucleocapsid, the membrane protein on the inside, and the hemagglutinating and fusion proteins on the outside of the lipoprotein envelope. *See* ANIMAL VIRUS; PARAMYXOVIRUS.

Transmission. Canine distemper is enzootic worldwide. Aerosol transmission in respiratory secretions is the main route of transmission. Virus shedding begins approximately 7 days after the initial infection. Acutely infected dogs and other carnivores shed virus in all body excretions, regardless of whether they show clinical signs or not.

Canine distemper virus affects susceptible dogs of all ages, but puppies are most vulnerable when their maternal antibodies are lost. Dogs that completely recover from canine distemper virus infection are immune for years and probably for life. They do not shed the virus and they are not persistently infected. Outside the host, canine distemper virus becomes inactivated rapidly, and all of the available common disinfectants are effective against it.

Disease signs. Great variations occur in the duration and severity of canine distemper, which may range from no visible signs to severe disease, often with central nervous system involvement, with approximately 50% mortality in dogs.

The first fever 3–6 days after infection may pass unnoticed; the second peak (several days later and intermittent thereafter) is usually associated with nasal and ocular discharge, depression, and anorexia. A low lymphocyte count is always present during the early stages of infection. Gastrointestinal and

respiratory signs may follow, often enhanced by secondary infection.

Many infected dogs develop central nervous system disorders following systemic disease. Depending on the virus strain, the signs may be more related to acute gray matter disease or subacute white matter disease. Seizures and muscle spasms accompanied by depression predominate in gray matter disease; lack of muscle coordination, partial paralysis, and muscle tremors occur in white matter disease. Brain and spinal inflammation and cervical rigidity may be seen in both. Inflammation of the optic nerve and retinal lesions in dogs with canine distemper virus are not uncommon. Hardening of the foot pads (hyperkeratosis) and nose is produced by some virus strains. In growing dogs, loss of enamel from the teeth after canine distemper virus infection is common. Old dog encephalitis, a rare disease of middle-aged or older dogs caused by canine distemper virus, is characterized by progressive deterioration and is ultimately fatal. Postvaccinal encephalitis may develop 1–2 weeks after vaccination and usually causes behavioral changes, seizures, and blindness; but this condition is rare.

Pathology. Atrophy of the thymus gland and hardening of lung tissue are the main lesions. Discharges of mucus and pus from the nose and eyes, diarrhea, hardening of nose and foot pads, and skin pustules may be present. Depletion of lymphocytes in lymphoid tissues, interstitial pneumonia, and intracytoplasmic inclusion bodies in epithelium of the respiratory, alimentary, and urinary tracts are characteristic. In the central nervous system, the degeneration of neurons is typical, but not always present. Intracytoplasmic and intranuclear bodies can be found, predominantly in astrocytes and neurons.

Dogs with old dog encephalitis have extensive infiltration of the surrounding vessels by lymphocytes in the central nervous system, degeneration of neurons, and minimal depletion of myelin. Postvaccinal encephalitis causes lesions predominantly in the brainstem.

Treatment. A specific antiviral drug having an effect on canine distemper virus in dogs is not presently available. Treatment of canine distemper, therefore, is nonspecific and supportive. Antibiotic therapy is recommended because of the common occurrence of secondary bacterial infections of the respiratory and alimentary tracts. Administration of fluids and electrolytes may be the most important therapy for canine distemper because diseased dogs with diarrhea are often dehydrated.

Treatment of neurologic forms of canine distemper is usually not very successful. Sedatives and anticonvulsants may temporarily relieve clinical signs, but they do not have a healing effect. When central nervous system disorders are progressive and dogs become recumbent, euthanasia is usually prescribed. Some dogs with central nervous system disorders, however, recover quite well, and residual signs such as muscle spasms or inflammation of the optic nerve can improve with time.

Vaccination and control. Immunization by controlled vaccination is the only effective approach to canine distemper prophylaxis. Active immunization with modified-live virus vaccines induces long-lasting immunity and keeps the disease in dogs under control. With a few exceptions, modified-live canine distemper virus vaccines are derived from either egg and avian cell or canine cell culture adaptations. Both methods of adaptation produce vaccines that are very effective in inducing an immunity that lasts for at least 1 year and probably for several years in most dogs. Any modified-live canine distemper virus vaccine may be fatal for certain wildlife and zoo animals (such as red pandas or black-footed ferrets). Inactivated virus vaccines must be used in some species, but they are not commercially available in the United States.

Heterotypic (measles) virus vaccination has been the best approach to overcome maternal antibody interference with immunization. As with inactivated canine distemper virus vaccines, measles virus induces a limited immunity that can protect dogs against canine distemper virus disease but not against canine distemper virus infection. A combination of attenuated measles virus and canine distemper virus is still commonly used in 6–10-week-old pups. It offers the advantage of complete protection in the absence of, and partial protection in the presence of, maternal antibody. In addition, measles virus antibody titers induced by vaccination usually are below a level that would interfere with measles virus vaccination in the next generation when used in 6–10-week-old pups.

A vaccination schedule for pups against canine distemper should include a combined modified-live measle virus–canine distemper virus vaccination at 6–8 weeks of age. Two additional canine distemper virus vaccinations at 3–4-week intervals should be given. Annual booster inoculations are recommended because some dogs lose antibody titers in that time period. Pups deprived of colostrum (the first milk from the mammary gland) should not be vaccinated with modified-live canine distemper virus before they are 3–4 weeks of age. Modified-live canine distemper virus can be fatal in unprotected younger pups, as it can be in some wild or zoo animals.

Besides immunization, strict isolation of dogs with canine distemper appears to be the most important step in controlling the disease. The virus is shed in all body excretions during the acute systemic disease, and direct dog-to-dog contact appears to be the main route of viral spread. Dogs with subacute canine distemper encephalitis still may infect susceptible contact dogs. Disinfection of the environment can be accomplished with commonly used products because the enveloped virus is rapidly destroyed outside the host.

M. J. G. Appel; B. A. Summers

Bibliography. C. E. Green (ed.), *Canine Distemper in Infectious Diseases of the Dog and Cat*, 1990; M. Horzinek (ed.), *Virus Infections of Vertebrates*, 1987.

Canine parvovirus infection

Severe enteritis caused by a small nonenveloped single-stranded deoxyribonucleic acid (DNA) virus that is resistant to inactivation and remains infectious in the environment for 5–7 months. First observed in dogs in 1976, canine parvovirus may have originated by mutation of a closely related parvovirus of cats or wildlife. The original virus was designated as canine parvovirus, type 2 (CPV-2); however, since its discovery the virus has undergone two minor genetic alterations, designated CPV-2a and CPV-2b. These alterations may have enabled the virus to adapt to its new host, replicate, and spread more effectively.

Epidemiology. Canine parvovirus is transmitted between dogs by the fecal-oral route. The incubation period is 3–7 days. Virus is first shed in the feces on day 3, and shedding continues for an additional 10 days. Chronically infected dogs that shed virus intermittently are rare. Most naturally occurring infections in dogs are subclinical or result in mild signs of the disease. Dogs with subclinical infections play an important role in the spread of the disease by shedding large amounts of virus into the environment. This shedding, along with the ability of the virus to persist in the environment, contributes to the endemicity of the disease. The development of disease following infection ranges from 20 to 90%, and mortality between 0 to 50%.

Pathogenesis. Only mitotically active cells, such as those of the intestine, lymphoid system, bone marrow, and fetal tissues, support replication of the canine parvovirus. After ingestion, the virus replicates in the regional lymphoid tissues of the oropharynx, spreading to other lymphoid tissues, bone marrow, and intestinal epithelial cells by way of the blood. Viral replication in the lymphoid tissue and bone marrow leads to necrosis of the lymphoid cells and a decrease in circulating white blood cells. Viral replication in the germinal epithelium of the intestinal glands of the jejunum and ileum leads to blunting and fusing of the intestinal villi. Diarrhea results from disruption of the normal absorptive and digestive processes. In severe cases, hemorrhaging into the intestine occurs. Secondary bacterial infections occur in severely affected animals.

Diagnosis and treatment. Severe enteritis is commonly observed in dogs 6–20 weeks of age. Initial signs include depression, anorexia, and lethargy, followed by fever, abdominal pain, vomiting, and diarrhea. Blood may be found in both the vomitus and stool. Dehydration, hypothermia, and weakness develop if the vomiting and diarrhea are severe. Signs are more severe in pups that are undernourished, stressed, or concurrently infected with other organisms. Diagnosis is based on history, clinical signs, vaccination status, physical exam, and laboratory tests. The white blood cell count is usually low. Detection of virus in the feces provides a definitive diagnosis. The bowel lumen is often empty, but on occasion there may be watery or hemorrhagic intestinal matter.

The goal of treatment is to support the animal until the infection runs its course. There are no specific antiviral therapies available. The intensity of treatment depends on the severity of signs. Dehydrated pups require intensive intravenous fluid therapy. Antimicrobial drugs are useful because of the risk of secondary bacterial infections, and antiemetic drugs help control vomiting and nausea. Good nursing care is essential. All food and water should be withheld until the pup is no longer vomiting, and the pup should be kept warm, clean, and dry. Because of the infectious nature of the disease, pups should be isolated from other dogs.

Prevention. Canine parvovirus infection in pups is prevented by minimizing exposure and by vaccination. Keeping pups isolated from other dogs until completion of the vaccination series is optimal. If strict isolation is not possible, limiting the exposure of pups to areas where large numbers of dogs congregate is useful. Immunity to canine parvovirus infection can be either passive or active. Passive immunity results from absorption of maternal antibodies following ingestion of colostrum; active immunity develops following natural infection or vaccination. The maternal antibodies protect the pup from natural infection, with the duration of protection dependent on the quantity of immunoglobulins absorbed. Inactivated and modified-live forms of the canine parvovirus vaccine are available. Vaccines that include a modified-live attenuated canine parvovirus strain are preferred because immunity develops more rapidly and is of longer duration. Pups are vaccinated at 2–3-week intervals starting at 6–8 weeks of age and continuing to 16–20 weeks of age. The immune response following vaccination depends on the amount of maternal antibodies present. High levels of antibodies block the immune response to vaccination. This period of nonresponsiveness may extend up to 16–20 weeks of age. *See* ANIMAL VIRUS; IMMUNITY; VACCINATION. Michael J. Coyne

Bibliography. M. S. Leib (ed.), *Gasteroenterology: The 1990's*, 1993; P. Tijssen (ed.), *Handbook of Parvoviruses*, 1989; C. Vella and S. W. Ketteridge, *Canine Parvovirus: A New Pathogen*, 1991.

Canis Major

The Great Dog, a southern constellation at the feet of Orion (see **illustration**). It contains Sirius, at magnitude −1.4 the brightest star in the sky. Sirius is known as the Dog Star. Its heliacal rising—its first visibility at dawn during the year—came at a time that indicated that the annual, agriculturally useful Nile floods would soon begin. Its name led to the phrase "the dog days of summer." Sirius is prominent in the winter evening sky. Its faint white-dwarf companion is difficult to detect in its glare, even using the required telescope. *See* SIRIUS.

In Greek mythology, this Big Dog could be a dog of Orion but could also be a dog of Helen, which allowed Paris to abduct her easily. *See* ORION.

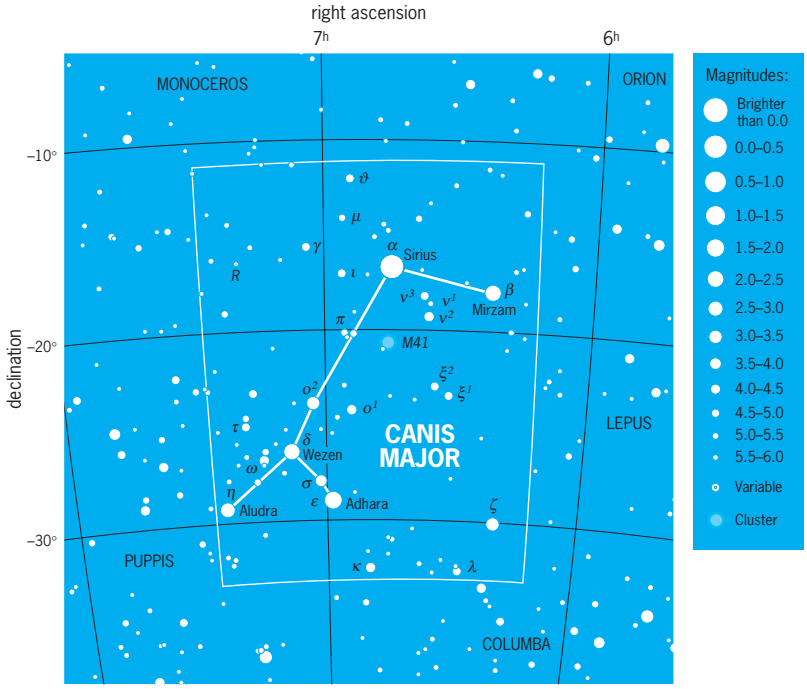

Modern boundaries of the constellation Canis Major, the Great Dog. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Canonical coordinates and transformations

Classical mechanics is the study of the positions and motions of particles or extended bodies. Initially the description used the coordinates and velocities of the objects under consideration and applied Newton's laws to describe their motions. Later the Lagrangian formulation was used to relate the positions, velocities, and resulting accelerations through an action principle. In the Lagrangian method, the velocities are considered as the time derivatives of the positions, and the equations of motion are determined by varying possible paths of the system holding the end points of the variation constant. A more general method is the Hamiltonian formulation, which considers the positions and momenta of the objects in the system under consideration as independent quantities. For the Hamiltonian variation, the momenta and the positions are varied independently subject to a constraint of the relation between the velocities and the momenta.

The Hamiltonian formulation allows the use of any system of coordinates that meet certain conditions. The advantage is that these coordinates can even be a mixture of the original positions and momenta. This freedom can often offer deeper insight into the physics of the system under consideration. The disadvantage is that there are certain restrictions on the possible choice of the new coordinates and momenta. Coordinates that meet these restrictions are named canonical coordinates. The allowed transformations from one set of canonical coordinates to another set are termed canonical transformations. In cases where the energy of the system is conserved, the Hamiltonian is the energy of the system. *See* HAMILTON'S EQUATIONS OF MOTION; LAGRANGE'S EQUATIONS; LAGRANGIAN FUNCTION.

General principles. To be more explicit, we can consider a system of n particles. If we use Newton's laws and the Lagrangian method, the nonrelativistic Lagrangian formulation of mechanics of a system of $3n$ degrees of freedom (x_i) produces $3n$ equations involving, in general, the positions, velocities, and accelerations of the objects in the system under consideration. The momentum, formally called the conjugate momentum (p_i), is defined as equal to the derivative of the Lagrangian with respect to the velocity. In contrast, the Hamiltonian method treats the momenta as independent quantities. The variation is taken of a quantity termed the Hamiltonian, usually denoted by H, which is a function of the independent coordinates q_i and independent momenta p_i. The Hamiltonian is given by Eq. (1), where L is the

$$H = \sum_i p_i \dot{q} - L(q, \dot{q}, t) \qquad (1)$$

Lagrangian and subscripts on the q's in the Lagrangian are suppressed for simplicity. This produces $6n$ equations, which involve only the coordinates and their momenta. The problem has been

reduced from a set of $3n$ second-order equations to a set of $6n$ first-order equations involving the coordinates, the momenta, the first derivatives of the momenta, and possibly time. The set of coordinates and momenta (q_i, p_i) are called canonical coordinates. Canonical transformations are transformations that maintain these definitions. The transformations can mix the original coordinates and momenta into new coordinates and momenta. The methods can be extended to a noncovariant formulation of relativistic mechanics.

Presumably "canonical" is used in the sense of designating a simple general set of standard equations. It appears that the term was introduced by C. G. Jacobi in 1837, but in a slightly different sense. Although the term rapidly gained common usage, the reason for its introduction apparently remained obscure even to contemporaries. Possibly there is an intended or unintended analogy to religious canons.

A straightforward application of the Hamiltonian formulation usually does not materially decrease the effort needed to solve a mechanics problem. The advantages are not in its use as a calculational tool but in the insight it affords to the formal structure of mechanics. Since the momenta and coordinates are treated equally, we may consider the motion of a physical system as a path in a $6n$-dimensional Cartesian space, called the phase space, which has $6n$ independent coordinates that describe the positions and momenta of each of the n particles that constitute the system. We can make invertible coordinate transformations that mix the old p_i's and q_i's into a new set of coordinates and momenta (Q_i, P_i); that is, the new coordinates can be expressed by Eqs. (2).

$$\begin{aligned} Q_i &= Q_i(q_j, p_j, t) \\ P_i &= P_i(q_j, p_j, t) \end{aligned} \tag{2}$$

For Hamiltonian mechanics, only those transformations can be of interest that maintain the canonical nature of the new coordinates. This requirement can be satisfied only if there exists some function, $K(Q, P, t)$, such that the equations of motion in the new set of coordinates are of the Hamiltonian form given by Eqs. (3). That is, Eqs. (4) are satisfied,

$$\dot{q} = \frac{\partial H}{\partial p_i} \qquad \dot{p} = -\frac{\partial H}{\partial q_i} \tag{3}$$

$$\dot{Q} = \frac{\partial K}{\partial P_i} \qquad \dot{P} = -\frac{\partial K}{\partial Q_i} \tag{4}$$

where K is the Hamiltonian in the new coordinate system.

Hamilton's principle is satisfied for both the old coordinates and the new coordinates provided that Eq. (5) holds, where H is the original Hamiltonian.

$$\lambda\Big(\sum_i p_i\dot{q} - H\Big) = \sum_i P_i\dot{Q} - K + \frac{dF}{dt} \tag{5}$$

The quantity F, called the generating function of the transformation, is a function of both the new and old phase-space coordinates and possibly time, with continuous second derivatives. The constant λ corresponds to a scale transformation. (A simple example of a scale transformation would be a change in units from p and q to P and Q such as a change from centimeters to meters. Such a transformation is termed a point transformation of phase space.) It is possible for a canonical transformation to affect only some of the coordinates and momenta, leaving the others unchanged.

Canonical transformations form a group. The generator of two successive transformations is the sum of the generators in the appropriate order. *See* GROUP THEORY.

There are only four basic canonical generating functions. They are given by Eqs. (6), where the subscripts on the q's, p's, Q's, and P's in functions F_1 to F_4 are suppressed for simplicity. The extra terms are needed to preserve the canonical nature of the transformations. Any other generating function can be expressed as a linear combination of the basic ones.

$$\begin{aligned} F &= F_1(q, Q, t) \\ F &= F_2(q, P, t) - \sum_i Q_iP_i \\ F &= F_3(p, Q, t) + \sum_i q_ip_i \\ F &= F_4(p, P, t) + \sum_i (q_ip_i - Q_iP_i) \end{aligned} \tag{6}$$

No other coordinate transformations can preserve the canonical nature of the new coordinates. The main problem in analyzing a physical system is in determining which generating function F_i provides useful insight. Sometimes the transformation is obvious; other times trial and error are used. Some well-known and useful examples of generating functions will be discussed.

Examples. The generating function of Eq. (7) generates the transformations of Eqs. (8). For the

$$F = F_4(p_i, P_i, t) + \sum_i (p_iq_i - Q_iP_i) \tag{7}$$

$$q_i = -\frac{\partial F_{i4}}{\partial p_i} \qquad Q_i = \frac{\partial F_{i4}}{\partial P_i} \tag{8}$$

trivial special case of $F_4 = \sum_i p_iP_i$, the resulting coordinates are $Q_i = p_i$ and $P_i = -q_i$. Therefore, this canonical transformation interchanges the coordinates and the momenta.

The harmonic oscillator in one dimension provides another example. The Hamiltonian can be written as Eq. (9). This Hamiltonian is a sum of squares

$$H = \frac{1}{2m}(p^2 + m^2\omega^2q^2) \tag{9}$$

of the coordinate and the momentum. We can use a generating function to eliminate the coordinate with a canonical transformation of the form (10),

$$p = f(P)\cos Q \qquad q = \frac{f(P)}{m\omega}\sin Q \tag{10}$$

which gives Eq. (11). A generating function of

$$K = H = \frac{f^2(P)}{2m} \tag{11}$$

form (12) produces the transformation of Eqs. (13),

$$F_1 = \frac{m\omega q^2}{2} \cot\,(Q) \tag{12}$$

$$q = \sqrt{\frac{2P}{m\omega}} \sin Q$$

$$p = \sqrt{2Pm\omega} \cos Q \tag{13}$$

so $f(p) = \sqrt{2m\omega P}$ and $H = \omega P$. Hence,

$$\dot{Q} = \frac{\partial H}{\partial P} = \omega \qquad Q = \omega t + \alpha$$

where α is a constant, so the usual solution for the harmonic oscillator is obtained. A coordinate that can be removed from the Hamiltonian by a canonical transformation is said to be a cyclic coordinate since it oscillates in time. This exercise may seem like "cracking a peanut with a sledge hammer," but it shows how cyclic coordinates can be extracted by canonical transformations. *See* ACTION; HAMILTON-JACOBI THEORY; HARMONIC OSCILLATOR.

Some applications. In addition to Hamiltonian-Jacobi theory and action angle variables, infinitesimal canonical transformations were used in the development of the original mathematics of quantum mechanics.

The generating function given by Eq. (14) uses

$$F(q, P, t) = \sum_j q_j P_j + \varepsilon X(q_i, P_i) \tag{14}$$

the form F_2. Here, ε is a small displacement in time whose square can be neglected, and the subscripts in F are suppressed for simplicity. This function generates an infinitesimal canonical transformation given by Eqs. (15). The δq_i and the δp_i are thus small, and

$$\begin{aligned} \delta q_j &= Q_j - q_j = \varepsilon \frac{\partial X}{\partial P_j} \\ \delta p_j &= P_j - p_j = -\varepsilon \frac{\partial X}{\partial q_j} \end{aligned} \tag{15}$$

the P_i can be replaced by p_i in the function X. Thus, X generates an infinitesimal displacement in the system variables q_i and p_i. Such quantities are termed dynamical variables. For example, if X is the z component of linear momentum of a particle, it generates an infinitesimal displacement in the z direction. Likewise, a component of angular momentum generates an infinitesimal rotation, and the Hamiltonian generates a change corresponding to an infinitesimal time displacement, ε.

Poisson brackets. If $Y(q_i, p_i, t)$ is another dynamical variable, distinct from the dynamical variable, $X(q_i, p_i, t)$ introduced above, then we can consider the change in Y, δY, under the transformation X. Equation (16) is obtained where (Y, X) is known as

$$\begin{aligned} \delta Y &= \sum_j \left(\frac{\partial Y}{\partial q_j} \delta q_j + \frac{\partial Y}{\partial p_j} p_j \right) \\ &= \sum_{j=\varepsilon} \left(\frac{\partial Y}{\partial q_j} \frac{\partial X}{\partial p_j} - \frac{\partial X}{\partial q_j} \frac{\partial Y}{\partial p_j} \right)_0 = \varepsilon(Y, X) \end{aligned} \tag{16}$$

the Poisson bracket of two dynamical variables Y and X. Clearly $(X, Y) = -(Y, X)$; that is, the bracket is antisymmetric.

If one of the dynamical variables is chosen to be the Hamiltonian, the Poisson bracket of any other dynamical variable is given by Eq. (17). If a dynamical variable does not explicitly depend upon time, its Poisson bracket with the Hamiltonian is its time derivative. If the Poisson bracket of a dynamical variable with the Hamiltonian vanishes, the dynamical variable is a constant of the motion, and it follows directly that the Hamiltonian is invariant under the infinitesimal transformation generated by that dynamical variable. This relation connects the symmetry of a system directly with the constants of motion of the system. *See* CONSERVATION LAWS (PHYSICS); SYMMETRY LAWS (PHYSICS).

$$\begin{aligned} (X, H) &= \sum_{j=1}^{f} \left(\frac{\partial X}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial H}{\partial q_j} \frac{\partial X}{\partial p_j} \right) \\ &= \frac{dX}{dt} - \frac{\partial X}{\partial t} \end{aligned} \tag{17}$$

Transition to quantum theory. Quantum mechanics is related to classical mechanics through the identification of the quantum commutator of two variables with $ih/(2\pi)$ times the Poisson bracket of the classically analogous dynamical variables, if they exist, where h is Planck's constant. This is one reason for the great importance of Poisson brackets. Thus, for example, classically the angular momentum components j_x, j_y, and j_z satisfy the Poisson bracket relation (18), and so on. The corresponding operators in quantum mechanics obey the commutation relation (19), where the notation [,] is used to distinguish between classical variables and quantum operators. *See* ANGULAR MOMENTUM; NONRELATIVISTIC QUANTUM THEORY.

$$(j_x, j_y) = j_z \tag{18}$$

$$j_x j_y - j_y j_x = \frac{ih}{2\pi} j_z = [j_x, j_y] \tag{19}$$

John L. Safko; Philip Stehle

Bibliography. T. L. Chow, *Classical Mechanics*, 1995; H. Goldstein, C. P. Poole, Jr., and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; I. Percival and D. Richards, *Introduction to Dynamics*, 1982.

Cantaloupe

In the United States the name applied to muskmelon cultivars belonging to *Cucumis melo* var. *reticulatus* of the family Cucurbitaceae. *See* VIOLALES.

Cantaloupes (*Cucumis melo*).

Description. The fruits weigh 2.4 lb (1.1 kg) and are round to oval; the surface is netted and has shallow vein tracts (see **illus.**). At maturity the skin color changes from dark green or gray to light gray or yellow. The flesh is usually salmon-colored, but it may vary from green to deep salmon-orange. When mature the melon is sweet, averages 6–8% sugar, and has a distinct aroma and flavor. The flesh is high in potassium and vitamin C, and when deep orange, rich in vitamin A. The vines usually bear andromonoecious flowers which are pollinated by bees, and the fruit generally separates from the stem when mature. Melons harvested at less than full-slip maturity do not achieve their full potential for sugar content, flavor, texture, and aroma. *See* ASCORBIC ACID; VITAMIN A.

The use of the name cantaloupe to indicate these medium-sized, netted melons with green and yellow-green rinds has become firmly established in the United States. However, this is a misnomer, and the name cantaloupe should be restricted to cultivars of *C. melo* var. *cantalupensis*. The fruits of this group are rough and scaly, with deep vein tracts and a hard rind. Cultivars of the variety *cantalupensis* are grown in Europe and Asia, but seldom in the United States.

Cantaloupe culture in the United States began with the introduction of the Netted Gem cultivar by the Burpee Seed Company in 1881.

Cultivation and production. Cantaloupes require a frost-free season of about 95 days to mature. Average temperatures of 70°F (21°C) are favorable for the production of high-quality melons. Most of the United States production is on irrigated land in the arid and semiarid Southwest because of favorable temperatures and lack of rain during the growing season. California grows approximately 51% of the United States acreage, followed by Texas with 22% and Arizona with 13%. Most of the California production is in the summer in the Sacramento and San Joaquin valleys, while Texas produces in the spring. High harvest and transportation costs tend to make cantaloupes a luxury product on the eastern markets. *See* MUSKMELON.

Oscar A. Lorenz

Diseases. Cantaloupe plants can be infected with bacterial wilt, angular leaf spot, downy mildew, scab, and cucumber mosaic. In addition, there are other fungal and viral diseases which can greatly reduce plant vigor and fruit quantity or quality.

Fungal diseases. Anthracnose may be the most destructive of all diseases affecting cantaloupe. All plant parts, excluding the roots, are susceptible to the pathogen *Colletotrichum lagenarium*. Lesions may coalesce to kill entire leaves, stems and runners. Depressed spots are formed on the fruits and, when wet conditions persist, are covered with a mass of pink spores. Spores easily spread to uninfected plants by splashing rain or are carried by cucumber beetles. The fungus has as many as seven races which are capable of infecting different cultivars. The disease can be controlled by using resistant varieties, although breeding for disease resistance in cantaloupes has lagged behind that in cucumbers and watermelons. Control can be achieved by crop rotation, proper drainage, and fungicides such as the dithiocarbamates.

Powdery mildew is a destructive disease caused by either *Spaerotheca fulginea* or *Erysiphe cichoracearum*. These fungi form a fuzzy white growth on the fruit, stem, or leaf. Tissue covered by the fungus for extended periods of time may be killed. Overwintering structures are tiny black fruiting bodies. Infection is most likely to occur when humidity is high and temperatures are moderately high. Powdery mildew is controlled by dusting plants with Karathane or sulfur. Resistant cultivars derived from P.M.R. 45, Georgia 47, Honey Ball 306, and Homegarden can also be planted.

Fusarium wilt, caused by *Fusarium oxysporum* f. *melonis*, is a particular problem in the northern states from Maine to Minnesota. The fungus either causes damping-off in young seedlings or a root rot and stem blight in older plants. Badly affected plant stems crack open, and the plants desiccate. The pathogen lives in old plant debris and as propagules in the soil. Hence crop rotation is not an effective control measure. The only worthwhile means of control is to plant resistant varieties, such as Honey Dew, Golden Gopher, Iroquois, Delicious 51, and Harvest Queen.

Viral diseases. Muskmelon mosaic and squash mosaic cause serious crop losses from time to time. Muskmelon mosaic virus is seed- and aphid-transmitted. Leaves develop a mottling and may be distorted in shape. Fruit production is greatly decreased. Squash mosaic virus is seed-transmitted and causes symptoms similar to those caused by muskmelon mosaic virus. Both virus diseases can be controlled by using certified disease-free seed and by eliminating wild cucurbit weeds from nearby locations. *See* PLANT VIRUSES AND VIROIDS.

Frank L. Caruso

Bibliography. R. A. Seelig, *Cantaloupes*, United Fresh Fruit and Vegetable Association, 1973; T. W. Whitaker and G. N. Davis, *Cucurbits*, 1961.

Cantilever

A linear structural member supported both transversely and rotationally at one end only; the other end of the member is free to deflect and rotate. Cantilevers are common throughout nature and

engineered structures; examples are a bird's wing, an airplane wing, a roof overhang, and a balcony. *See* WING.

Horizontal axis. If the main axis of the member is horizontal and the member is loaded with a downward gravity load, the member will support the load by forming a structural couple, with the upper portion of the member being in tension and the lower portion being in compression. This is the principle on which the design of a horizontal cantilever truss is based. When a horizontal cantilever truss is under a gravity load, the top member is in tension and the bottom member is in compression. If there are no external horizontal forces, the internal forces in the top and bottom members must be equal. Because these internal forces are separated by a distance, they can resist moments and rotation. The resisting moment is defined as a couple, that is, two equal and opposite forces separated by a distance. The translation force of the cantilever (defined as the shear) is resisted by the diagonals in a truss or the web of a wide flange beam. *See* TRUSS.

For a cantilever to function, it must be counterbalanced at its one support against rotation. This requirement is simply achieved in the design of a playground seesaw, with its double-balanced cantilever, or of a chemical balance. This principle of counterbalancing the cantilever is part of the basic design of a crane, such as a tower crane (see **illus.**). More commonly, horizontal cantilevers are resisted by being continuous with a backup span that is supported at both ends. This design is common for cantilever bridges; the largest of the cantilever bridges is an 1800-ft (540-m) railroad bridge in Quebec, Canada. Cantilever bridges are simpler to erect since they can be built out from their supports. All swing bridges or drawbridges are cantilevers. *See* BRIDGE.

Vertical axis. Cantilevers also occur with the member's main axis being vertical. Vertical cantilevers primarily resist lateral wind loads and horizontal loads created by earthquakes. Common vertical cantilevers are chimneys, stacks, masts, flagpoles, lampposts, and railings or fences. Trees are the largest vertical cantilevers found in nature.

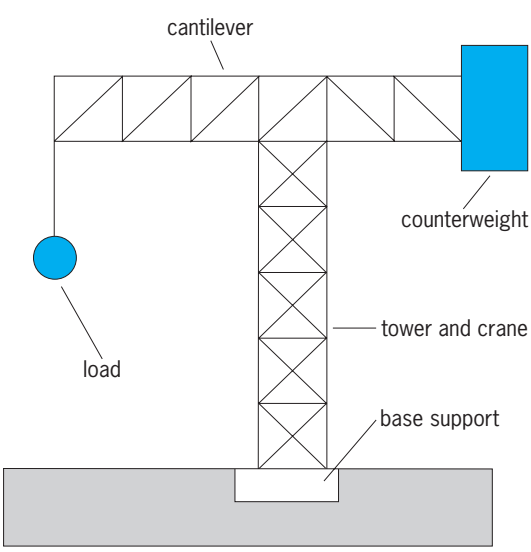

Cantilever configuration in the form of a tower support crane.

All skyscrapers are vertical cantilevers. One common system to provide the strength to resist lateral loads acting on the skyscraper is the use of a truss (known as bracing), which cantilevers out from the foundation in the walls of the core of the building that encloses the elevators and the stairs. In apartment buildings and hotels, vertical walls that separate different living units (demising walls or shear walls) cantilever out of the foundation to resist lateral loads. *See* BUILDINGS; SHEAR.

Applications. Some of the largest cantilevers are used in the roofs of airplane hangars. The reason is that typically at one edge of an airplane hangar there is a sliding door that can be removed; thus it cannot provide support to the roof's edge. Cantilevers of up to 300 ft (90 m) have been used to enclose 747 airplanes. It has become common practice to include cantilevers in the design of theaters and stadiums, where an unobstructed view is desired, balconies and tiers are supported in the back and cantilevered out toward the stage or playing field so that the audience has column-free viewing. *See* BEAM; ROOF CONSTRUCTION. I. Paul Lew

Bibliography. M. Salvadori, *Structure in Architecture*, 1986; M. Salvadori, *Why Buildings Stand Up*, 1994.

Capability maturity modeling

A method of gathering best practices for various work processes into coherent groupings to facilitate quality and process improvement. Capability Maturity Models® (CMM®) represent a relatively recent step in the evolution of process improvement techniques that enable organizations to repeatedly produce high-quality, complex, software-intensive systems faster and within expected costs. The foundations of this work include approaches to improved quality developed by W. Edwards Deming, Joseph Juran, Walter A. Shewart, and Philip B. Crosby, particularly Crosby's quality management maturity grid.

Model structure. Capability Maturity Models are designed to describe discrete levels of improvement within process areas, measured against specific goals and generic goals. A process area is a group of related practices that are performed collectively to achieve a set of objectives. There are typically 15 to 25 process areas represented in a model. Examples of process areas include requirements development, validation, configuration management, and project planning. Each process area has specific goals that describe what must be achieved to satisfy the purpose of the process area. Each specific goal has specific practices that describe the activities expected to result in achievement of the goal. For example, some specific goals of the project planning process area involve

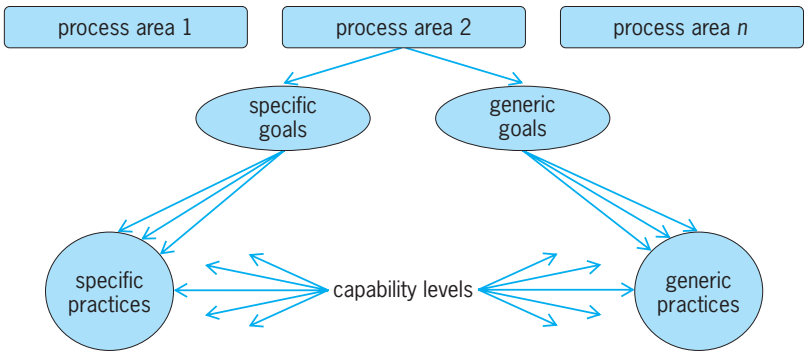

Fig. 1. Capability Maturity Models: the continuous representation.

establishing estimates, developing a project plan, and obtaining commitment to the plan. The specific practices related to the "establish estimates" goal involve estimating the scope of the project, establishing estimates of project attributes, defining the project life cycle, and determining estimates of effort and cost. Although process areas depict behavior that should be exhibited in any organization, practices must be interpreted using in-depth knowledge of the model, the organization, the business environment, and the other specific circumstances involved.

Generic goals apply to all process areas. Achievement of each of these goals in a process area signifies whether the implementation and institutionalization of each process area is effective, repeatable, and lasting. Each generic goal, in turn, has associated generic practices. For example, the generic goal "institutionalize a quantitatively managed process" has two generic practices, "establish quality objectives" and "stabilize subprocess performance."

A Capability Maturity Model does not prejudge which processes are right for any organization or project. Instead, it establishes minimal criteria that processes must meet to be considered capable. A capable process is defined, documented, practiced, supported, maintained, controlled, verified, validated, measured, and able to be improved. Also, the model should allow interpretation and modification as needed to suit an organization's size or business objectives.

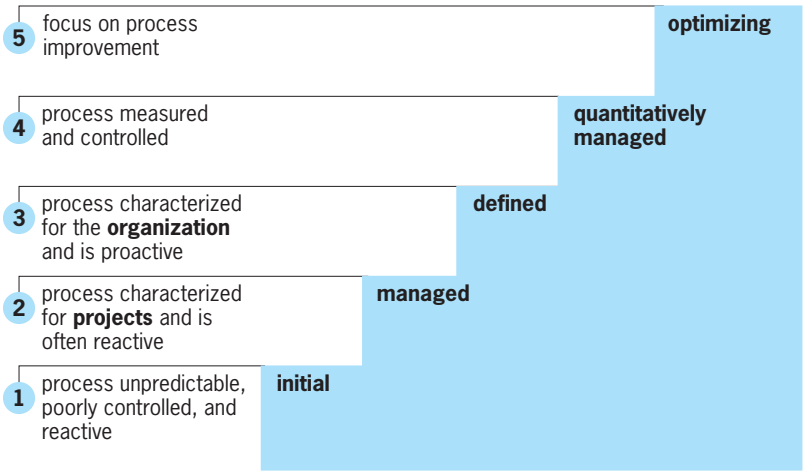

Fig. 2. Organizational maturity levels.

Representations. Capability Maturity Models have been architected in two representations, continuous and staged. These representations provide alternative approaches to process improvement that take advantage of the familiarity of users with either approach. The representations contain the same essential content but are organized differently.

The continuous representation is based on capability within a given process area—the range of expected results that can be achieved by following a process. A low-capability process is improvised and highly dependent on current practitioners; results are difficult to predict, and product functionality and quality are typically compromised to meet schedules. A high-capability process is well controlled, defined, continually improving, supported by measurement, and is a basis for disciplined use of technology.

The continuous representation provides organizations with the flexibility to choose which processes to emphasize for improvement and how much to improve each process. Process improvement is measured using six capability levels: (0) incomplete, (1) performed, (2) managed, (3) defined, (4) quantitatively managed, and (5) optimizing. These capability levels relate to the achievement of specific and generic goals of a process area. For example, an organization can reach capability level 2 of a process area when the specific and generic goals up through capability level 2 are achieved for the process area.

The continuous representation enables selection of the order of process improvement that best meets the organization's business objectives and that most mitigates risk. **Figure 1** illustrates the structure of continuous representation.

The staged representation is based on organizational maturity—the combined capabilities of a set of related processes. Thus, in a highly mature organization, the set of organizational processes, taken as a whole, are of higher capability. Organizational improvement is measured in five maturity levels: (1) initial, (2) managed, (3) defined, (4) quantitatively managed, and (5) optimizing. The maturity levels comprise well-defined evolutionary plateaus on the path to becoming a mature organization. **Figure 2** provides a brief definition of each maturity level.

The staged representation has a recommended order for approaching process improvement, beginning with basic management practices and progressing along a proven path. The Capability Maturity Model for Software (SW-CMM) and People Capability Maturity Model (P-CMM) use this structure. **Figure 3** illustrates the structure of the staged representation.

Generic practices are often grouped by common features in the staged representation. "Commitment to perform" describes practices that relate to creating management policies and securing sponsorship.

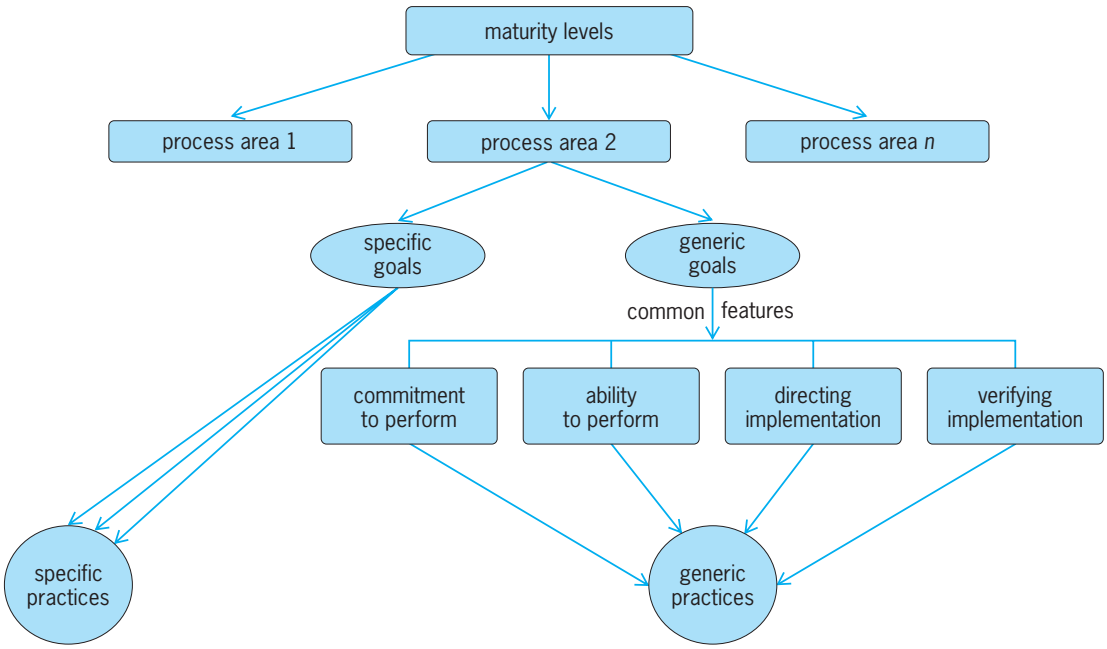

Fig. 3. Capability Maturity Models: the staged representation.

"Ability to perform" characterizes practices related to establishing and maintaining plans, resources, assigned responsibility and authority, and training. "Directing implementation" comprises measurement, control, and performance practices. "Verifying implementation" comprises practices that ensure implementation and compliance. [In the most recent version of Capability Maturity Model Integration (CMMI), discussed below, the grouping of generic practices has been deleted to reduce differences between the staged and continuous representations.]

Model evolution. In 1987, the U.S. Air Force asked Carnegie Mellon University's Software Engineering Institute to identify key practices that a contractor had to perform to deliver software-intensive systems reliably. By 1991, this tracking of practices had matured into the Capability Maturity Model for Software (SW-CMM). The success of this model for one discipline led to similar efforts for other elements of the product development community. Interest in such models for systems engineering process improvement led to two models produced in 1994. These were the Software Engineering CMM (SE-CMM), created by the Enterprise Process Improvement Collaboration (EPIC), and the Systems Engineering Capability and Assessment Method (SECAM), created by the International Council on Systems Engineering (INCOSE). These two models were successfully merged into Electronic Industries Alliance (EIA) Interim Standard 731 (EIA/IS 731) in 1998 as a result of a collaborative effort of EIA, EPIC, and INCOSE. In 1996, a sister model to cover key practices in software acquisition was created, the Software Acquisition Capability Maturity Model (SA-CMM). Concerns about preserving and enhancing the capabilities of developmental engineering staff led to the creation of the People Capability Maturity Model (P-CMM) in 1997.

In 1997, work was underway at the Software Engineering Institute to produce an update to the SW-CMM, and also to produce a model that would capture concurrent engineering practices in an Integrated Product Development CMM (IPD-CMM). The Institute's sponsor, the U.S. Department of Defense, determined that these efforts should be merged into an integrated model, to be called Capability Maturity Model Integration (CMMI®). The feasibility of integrating a diverse set of maturity models had been demonstrated earlier that year by the Federal Aviation Administration (FAA), which had developed an integrated capability maturity model (FAA-iCMM V1.0). Due to the widespread focus on integrated product and process development (IPPD) by the Department of Defense and industry, it was decided that the initial focus of the CMMI effort would be integration of systems engineering, software engineering, and IPPD. A subsequent CMMI version added guidance for supplier sourcing. **Figure 4** depicts the evolution of the CMMI models.

CMMs in use. Several Capability Maturity Models are now in widespread use, with some having achieved the status of de facto or recognized international standard, for the assessment and improvement of such areas as software engineering and acquisition practices, work force development practices, and systems security practices, among others.

SW-CMM. The Capability Maturity Model for Software is used for judging the maturity of the software processes of an organization and for identifying the key practices that are required to increase the maturity of these processes. Used worldwide, the SW-CMM became a de facto standard for assessing and improving software processes. As a result of the

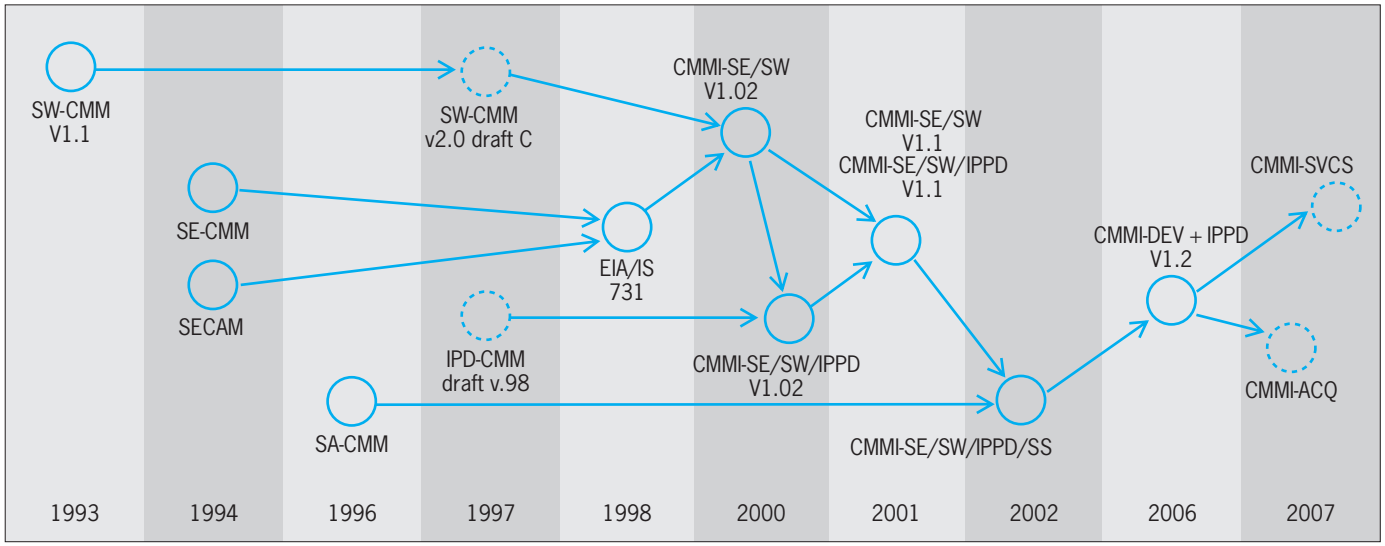

Fig. 4. CMMI model evolution.

creation of CMMI as an evolutionary improvement, appraisals against the SW-CMM are no longer supported by the Software Engineering Institute.

SA-CMM. The Software Acquisition Capability Model (SA-CMM) is used for benchmarking and improving the software acquisition process. The model follows the same architecture as the SW-CMM, but with an emphasis on acquisition issues and the needs of individuals and groups who are planning and managing software acquisition efforts. For example, unique key process areas address contractor tracking and oversight, transition to support, and acquisition risk management. Because of the move to the modernized CMMI approach, a CMMI version for acquisition is currently being developed to modernize and extend this model.

P-CMM. The People Capability Maturity Model is a framework that helps organizations successfully address people issues. Based on the best current practices in fields such as human resources, knowledge management, and organizational development, the P-CMM helps organizations characterize the maturity of their work-force practices, establish a program of continuous work-force development, set priorities for improvement actions, integrate work-force development with process improvement, and establish a culture of excellence. Its five maturity levels establish successive foundations for continuously improving individual competencies, developing effective teams, motivating improved performance, and shaping the work force that the organization needs to accomplish its future business plans. By following the maturity framework, an organization can avoid introducing work-force practices that its employees are unprepared to implement effectively.

SSE-CMM. The Systems Security Engineering Capability Maturity Model (SSE-CMM) describes the essential characteristics of an organization's security engineering process that must exist to ensure good security engineering. It is an emerging international standard (ISO/1EC 21827). The model highlights the relationship between security engineering and systems engineering. It provides a way to measure and improve capability in applying security engineering principles and to address capability-based assurance. Unique process areas include administering security controls; assessing impact, security risk, threat, and vulnerability; and verifying and validating security.

FAA-iCMM. The Federal Aviation Administration developed the FAA-iCMM to increase the efficiency and effectiveness of process improvement by providing a single reference model that integrates engineering, management, and acquisition processes used in developing, managing, acquiring, and maintaining systems. The FAA-iCMM integrates all features and practices of three that were being used separately in the FAA: the SE-CMM (V1.1), the SA-CMM (V1.01), and the SW-CMM (V1.1). A subsequent version (V2.0) has sought to integrate CMMI themes as well as concepts from the quality management standard ISO 9001:2000, which is familiar in the manufacturing domain.

CMMI. CMMI models include a common set of process areas that form the core of a model framework that integrates process improvement guidance for engineering development and integrated product and process development. In addition to the engineering development version, work is underway to produce synergistic variants to cover systems acquisition and the provision of technical services. The resulting Capability Maturity Models may be adapted to an organization's mission and business objectives. The CMMI Product Suite reduces the redundancy and complexity that can result from the use of multiple CMMs, thus improving the efficiency of and return on investment of process improvement. *See* ENGINEERING DESIGN; REENGINEERING; SOFTWARE ENGINEERING; SYSTEM DESIGN EVALUATION; SYSTEMS ENGINEERING; SYSTEMS INTEGRATION.

Mike Phillips

Bibliography. D. Ahern, A. Clouse, and R. Turner, *CMMI® Distilled: A Practical Introduction to*

Integrated Process Improvement, 2d ed., Addison-Wesley, Boston, 2003; M. Chrissis, M. Konrad, and S. Shrum, *CMMI®: Guidelines for Process Integration and Product Improvement*, Addison-Wesley, Boston, 2003; P. B. Crosby, *The Eternally Successful Organization: The Art of Corporate Wellness*, Mentor Books, 1992; P. B. Crosby, *Quality Is Free*, McGraw-Hill, New York, 1979; J. M. Juran and A. B. Godfrey (eds.), *Juran's Quality Handbook*, 5th ed., McGraw-Hill, New York, 1999; Software Engineering Institute, Carnegie Mellon University, *The Capability Maturity Model®: Guidelines for Improving the Software Process*, ed. by M. C. Paulk et al., Addison-Wesley, Boston, 1995.

Capacitance

The ratio of the charge q on one of the plates of a capacitor (there being an equal and opposite charge on the other plate) to the potential difference v between the plates; that is, capacitance (formerly called capacity) is $C = q/v$.

In general, a capacitor, often historically called a condenser, consists of two metal plates insulated from each other by a dielectric. The capacitance of a capacitor depends on the geometry of the plates and the kind of dielectric used since these factors determine the charge which can be put on the plates by a unit potential difference existing between the plates. *See* DIELECTRIC MATERIALS.

For a capacitor of fixed geometry and with constant properties of the dielectric between its plates, C is a constant independent of q or v, since as v changes, q changes with it in the same proportion. This statement assumes that the dielectric strength is not exceeded and thus that dielectric breakdown does not occur. (If it does occur, the device is no longer a capacitor.) If either the geometry or dielectric properties, or both, of a capacitor change with time, C will change with time.

In an ideal capacitor, no conduction current flows between the plates. A real capacitor of good quality is the circuit equivalent of an ideal capacitor with a very high resistance in parallel or, in alternating-current (ac) circuits, of an ideal capacitor with a low resistance in series.

Properties of capacitors. One classification system for capacitors follows from the physical state of their dielectrics. For a discussion based on this classification, *see* CAPACITOR.

Charging and discharging. These processes can occur for capacitors while the potential difference across the capacitor is changing if C is fixed; that is, q increases if v increases and q decreases if v decreases. If C and v both change with time, the rate of change of q with time is given by Eq. (1). Since the current

$$\frac{dq}{dt} = C\frac{dv}{dt} + v\frac{dC}{dt} \tag{1}$$

i flowing in the wires leading to the capacitor plates is equal to dq/dt, Eq. (1) gives i in the wires. In many cases, C is constant so $i = C\,dv/dt$.

Energy of charged capacitor. This energy W_c is given by the formula $W_c = vq/2$, and is equal to the work the source must do in placing the charge on the capacitor. It is, in turn, the work the capacitor will do when it discharges.

Geometrical types. The geometry of a capacitor may take any one of several forms. The most common type is the parallel-plate capacitor whose capacitance C in farads is given in the ideal case by Eq. (2),

$$C = \frac{A\epsilon_r\epsilon_0}{d} \tag{2}$$

where A is the area in square meters of one of the plates, d is the distance in meters between the plates, ϵ_0 is the permittivity of empty space with the numerical value 8.85×10^{-12} farad/m, and ϵ_r is the relative permittivity of the dielectric material between the plates. The value of ϵ_r is unity for empty space and almost unity for gases. For other dielectric materials, ϵ_r ranges in value from one to several hundred. In order for Eq. (2) to give a good value of C for an actual capacitor, d must be very small compared to the linear dimensions of either plate. *See* PERMITTIVITY.

Each plate of a parallel-plate capacitor may be made up of many thin sheets of metal connected electrically with a corresponding number of thin sheets of metal making up the other plate. The sheets of metal and their intervening layers of dielectric are chosen and stacked in such a way that A will be large

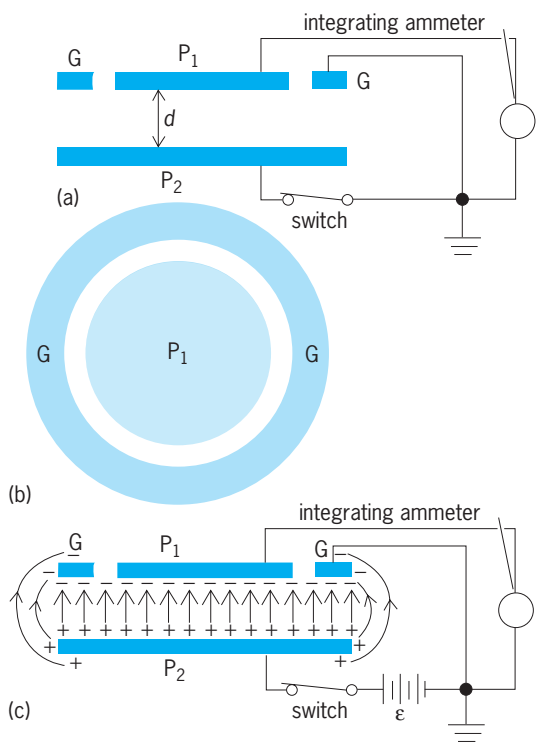

Fig. 1. Parallel-plate capacitor P_1, P_2 with guard ring GG. Spacing between guard ring and plate is exaggerated. (*a*) Cross section with all parts of capacitor at ground potential. (*b*) Top view showing plate P_1 surrounded by guard ring. (*c*) Cross section of capacitor charged from battery whose electromotive force is ⊲ .

and d small without making the whole capacitor too bulky. The result is appreciable capacitance in a reasonable volume.

The cylindrical capacitor, as the name implies, is made of two concentric metal cylinders, each of length l in meters, with a dielectric filling the space between the cylinders. If r_2 and r_1 are the radii in meters of the outer and inner cylinders, respectively, and l is very large compared to $r_2 - r_1$, the capacitance C in farads is given by Eq. (3), where ϵ_r is

$$C = \frac{2\pi\epsilon_r\epsilon_0 l}{ln\,(r_2/r_1)} \tag{3}$$

the relative dielectric constant of the dielectric and ln (r_2/r_1) indicates the natural logarithm of the ratio r_2/r_1.

Guard ring. This is often used with a standard parallel-plate capacitor, as shown in **Fig. 1**, in order that Eq. (2) shall more accurately represent its capacitance. It is the fringing of the electric lines of force which makes Eq. (2) inaccurate for an actual capacitor and, as shown in Fig. 2, the fringing is nearly all at the outside edge of the guard ring, and thus is not associated with the charge Q which is put onto plate P_1 while the capacitor is being charged. It is the charge Q that determines the integrated current flowing during the charging process. Then Eq. (2) gives the correct value of C that is needed to relate Q to the potential difference E across the plates by the equation $Q = CE$. Thus, with C known from Eq. (2) and E, known from a voltmeter measurement, Q may be computed and the integrating ammeter calibrated. An instrument called a ballistic galvanometer, which is now largely of only historical interest, constitutes an integrating ammeter. This illustrates one use of a standard capacitor having a guard ring. *See* GALVANOMETER.

Ralph P. Winch

Types of capacitor. One classification of capacitors comes from the physical state of their dielectrics, which may be gas (or vacuum), liquid, solid, or a combination of these. Each of these classifications may be subdivided according to the specific dielectric used. Capacitors may be further classified by their ability to be used in alternating-current (ac) or direct-current (dc) circuits with various current levels.

Capacitors are also classified as fixed, adjustable, or variable. The capacitance of fixed capacitors remains unchanged, except for small variations caused by temperature fluctuations. The capacitance of adjustable capacitors may be set at any one of several discrete values. The capacitance of variable capacitors may be adjusted continuously and set at any value between minimum and maximum limits fixed by construction. Trimmer capacitors are relatively small variable capacitors used in parallel with larger variable or fixed capacitors to permit exact adjustment of the capacitance of the parallel combination.

Air, gas, and vacuum types. Made in both fixed and variable types, these capacitors are constructed with flat parallel metallic plates (or cylindrical concentric metallic plates) with air, gas, or vacuum as the dielectric between plates. Alternate plates are connected, with one or both sets supported by means of a solid insulating material such as glass, quartz, ceramic, or plastic. **Figure 2** shows a variable air capacitor. Gas capacitors are similarly built but are enclosed in a leakproof case. If the gas capacitor is variable, the shaft supporting the movable plates, or rotor, is brought out through a pressure-tight insulated seal. Vacuum capacitors are of concentric-cylindrical construction and are enclosed in highly evacuated glass envelopes.

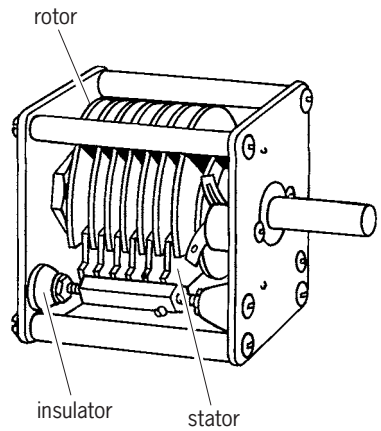

Fig. 2. Variable air capacitor.

The purpose of a high vacuum, or a gas under pressure, is to increase the voltage breakdown value for a given plate spacing. For high-voltage applications, when increasing the spacing between plates is undesirable, the breakdown voltage of air capacitors may be increased by rounding the edges of the plates. Air, gas, and vacuum capacitors are used in high-frequency circuits. Fixed and variable air capacitors incorporating special design are used as standards in electrical measurements. *See* CAPACITANCE MEASUREMENT; ELECTRICAL UNITS AND STANDARDS.

Solid-dielectric types. These capacitors use one of several dieletrics such as a ceramic, mica, glass, or plastic film. Alternate plates of metal, or metallic foil, are stacked with the dielectric, or the dielectric may be metal-plated on both sides. Some of the more popular dielectric materials are listed in **Table 1**, along

TABLE 1. Dielectric materials used in capacitors

Dielectric material	Relative permittivity (ϵ_r)	Power factor, %
Vacuum or air	1	0
Kraft paper	4.0–6.5	3
Polyester	3.0	0.5
Polypropylene	2.2	0.01
Polystyrene	2.5	0.05
Glass	6.7	0.06
Mica	6–8	0.02
Aluminum oxide	10.0	5–10
Tantalum oxide	11.0	5–10
Barium titanate	10.60	0.2–2.5

with their relative permittivities and power factors. *See* DIELECTRIC MATERIALS; POWER FACTOR.

Plastic-film types. These capacitors use dielectrics such as polypropylene, polyester, polycarbonate, or polysulfone with a relative permittivity ranging from 2.2 to 3.2. This plastic film may be used alone or in combination with Kraft paper. The dielectric thicknesses range from 0.06 to over 0.8 mil (1.5 to over 20 micrometers). The most common electrodes are aluminum or zinc vacuum-deposited on the film, although aluminum foil is also used. These types generally utilize wound-roll construction (**Fig. 3**). *See* POLYESTER RESINS; POLYOLEFIN RESINS; POLYSULFONE RESINS.

Plastic-film capacitors may be constructed in a dry roll for dc and low-voltage ac applications. For ac applications above about 250 V, the capacitor winding is usually impregnated with a dielectric fluid to avoid degradation of the dielectric due to partial discharges.

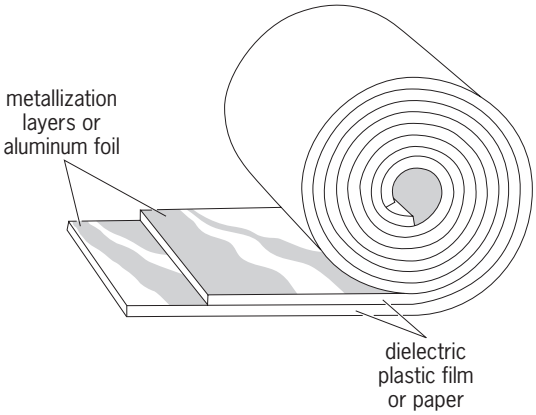

Fig. 3. Wound-roll construction of plastic film and paper capacitors.

Mica types. These capacitors use thin rectangular sheets of mica as the dielectric. The relative permittivity of mica is in the range of 6–8. The electrodes are either thin sheets of metal foil stacked alternately with the mica sheets, or thin deposits of silver applied directly to the surface of the mica sheets. Mica capacitors are used chiefly in radio-frequency applications. They have a low dielectric loss at very high frequencies, good temperature, frequency, and aging characteristics, and low power factor, but have a low ratio of capacitance to volume or to mass. They are made with dc voltage ratings from a few hundred to many thousands of volts and with radio-frequency current ratings up to about 50 A. *See* MICA.

Paper types. A dielectric of Kraft paper usually impregnated with mineral oil, or ester, is used in paper capacitors. Paper and plastic-film capacitors (Fig. 3) are constructed by stacking, or forming into a roll, alternate layers of foil and dielectric. Paper capacitors are gradually being replaced by metallized polypropylene and polyester films with lower cost, smaller size, and lower power factor. *See* PAPER.

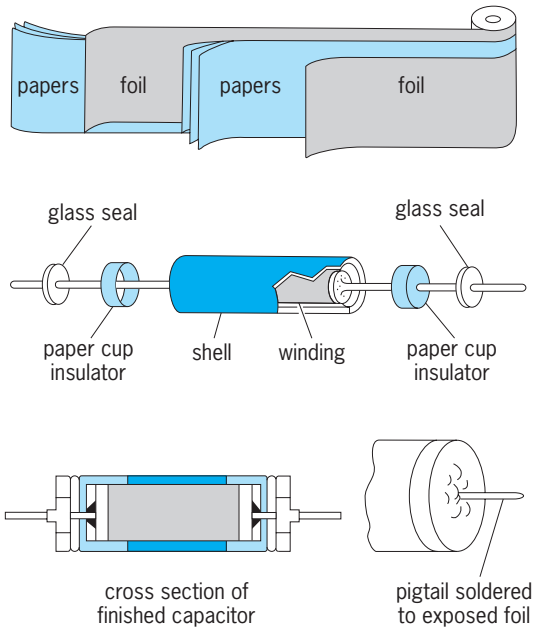

Fig. 4. Ceramic capacitor constructed in chip form, showing cutaway of finished chip.

Ceramic types. The monolithic ceramic capacitor consists of dielectric layers interleaved with electrode layers; the assembly is then compressed and sintered to form a solid monolithic block (**Fig. 4**). The dielectric layer may be as thin as 0.8 mil (20 μm), and as many as 40–50 layers may be interleaved. The most common dielectric material is barium titanate. The electrodes are composed of materials such as a silver-palladium mixture. The end termination is usually composed of several layers of silver-palladium, nickel, and tin. These capacitors may be constructed in radial, axial, or chip form. The chip construction is especially popular for surface mounting on electronic circuits. The chips have a typical size of $^1/_8$ in. × $^1/_6$ in. × $^1/_{16}$ in. (3.2 mm × 1.6 mm × 0.65 mm). These capacitors have the ability to withstand the 450°F (230°C) reflow solder and 540°F (280°C) wave solder mounting temperatures. *See* CERAMICS.

Electrolytic types. A large capacitance-to-volume ratio and a low cost per microfarad of capacitance are chief advantages of electrolytic capacitors. These use aluminum or tantalum plates (**Fig. 5**). A paste electrolyte is placed between the plates, and a dc forming voltage is applied. A current flows and by a process of electrolysis builds up a molecule-thin layer of oxide bubbles on the positive plate. This serves

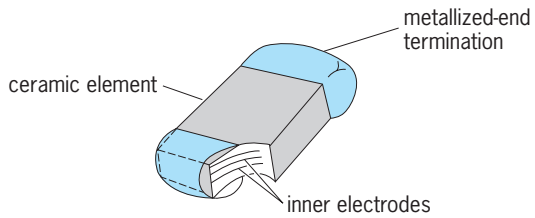

Fig. 5. Tantalum electrolytic capacitor showing cross section. (*General Electric Co.*)

TABLE 2. Major types of capacitors*

Type	Capacitance	Voltage (working voltage, dc), V	Applications
Monolithic ceramics	1 pF–2.2 μF	50–200	Ultrahigh frequency, rf coupling, computers
Disk and tube ceramics	1 pF–1 μF	50–1000	General, very high frequency
Paper	0.001–1 μF	200–1600	Motors, power supplies
Film			
Polypropylene	0.0001–0.47 μF	400–1600	Television vertical circuits, rf circuits
Polyester	0.001–4.7 μF	50–600	Entertainment electronics
Polystyrene	0.001–1 μF	100–200	General, high stability
Polycarbonate	0.01–18 μF	50–200	General
Metallized polypropylene	4–60 μF	400†	Alternating-current motors
Metallized polyester	0.001–22 μF	100–1000	Coupling, rf filtering
Electrolytic			
Aluminum	1 100,00 μF	5–500	Power supplies, filters
Tantalum	0.1–2200 μF	3–150	Small space requirement, low leakage
Gold	0.022–10 F	2.5–5.5	Memory backup
Nonpolarized (either aluminum or tantalum)	0.47–1000 μF	10–200	Loudspeaker crossovers
Mica	330 pF–0.05 μF	50–1000	High frequency
Silver-mica	5–820 pF	50–500	High frequency
Variable			
Ceramic	1–5 to 16–100 pF	200	Radio, television, communications
Film	0.8–5 to 1.2–30 pF	50	Oscillators, antenna, rf circuits
Air	10–365 pF	50	Broadcast receiver
Teflon	0.25–1.5 pF	2000	Very high frequency, ultrahigh frequency

*After A. Mottershead, *Electricity and Electronics,* Prentice Hall, 1990.
†Alternating-current voltage at 60 Hz.

as the dieletric. The rest of the electrolyte and the other plate make up the negative electrode (**Fig. 6***a*). Such a device is said to be polarized and must be connected in a circuit with the proper polarity. To ensure proper connection, the polarization of the capacitor is clearly marked by polarity signs for the axial type (Fig. 6*b*) or by a longer lead for the positive side in the radial type (Fig. 6*c*). Polarized capacitors can be used only in circuits in which the dc component of voltage across the capacitors exceeds the crest value of the ac ripple.

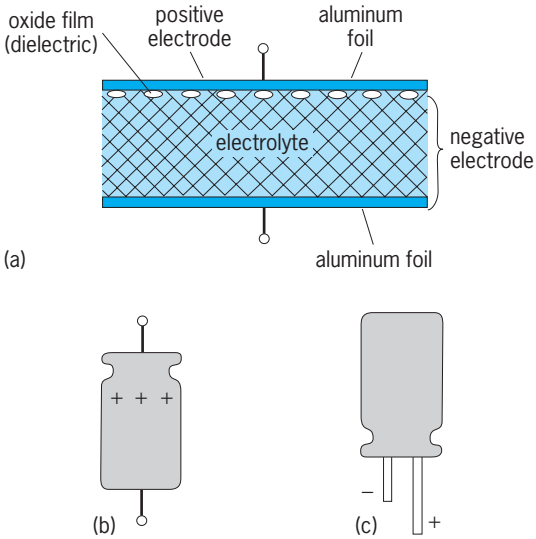

Fig. 6. Construction and polarity markings of aluminum electrolytic capacitor. (*a*) Cross section. (*b* Polarization markings for tabular capacitor with axial leads. (*c*) Polarization markings for capacitor with radial leads. (*After A. Mottershead, Electricity and Electronics, Prentice Hall, 1990*)

Another type of electrolytic capacitor utilizes compressed tantalum powder and the baking of manganese oxide (MnO_2) as an electrolyte. These capacitors may be constructed with either radial, axial, or chip configurations.

Nonpolarized electrolytic capacitors can be constructed for use in ac circuits. In effect, they are two polarized capacitors placed in series with their polarities reversed.

Table 2 provides a summary of the major types of discrete capacitors and some of their applications.

Thick-film types. These capacitors are made by means of successive screen-printing and firing processes in the fabrication of certain types of microcircuits used in electronic computers and other electronic systems. They are formed, together with their connecting conductors and associated thick-film resistors, upon a ceramic substrate. Their characteristics and the materials are similar to those of ceramic capacitors. *See* PRINTED CIRCUIT.

Thin-film types. Thin-film dielectrics are deposited on ceramic and integrated-circuit substrates and then metallized with aluminum to form capacitive components. These are usually single-layer capacitors. The most common dielectrics are silicon nitride and silicon dioxide. Thin-film dieletrics may be deposited by plasma deposition or sputtering. *See* INTEGRATED CIRCUITS; SPUTTERING. Allen Mottershead

Body capacitance. When a part of the human body, say the hand, is brought near a high-impedance network, the body serves as one plate of a capacitor and the adjacent part of the network as the other plate. This situation is the equivalent of a

capacitor of very low capacitance, in parallel between that part of the network and its surroundings, since the human body is usually a conductor at the potential of its surroundings. This capacitance is known as body capacitance and enters as a part of the distributed capacitance of the network. A high-impedance network must be well shielded in order to eliminate the variable and undesirable effects of body capacitance. Ralph P. Winch

Bibliography. W. J. Duffin, *Electricity and Magnetism*, 4th ed., 1990; D. Halliday, R. Resnick, and K. S. Krane, *Physics*, 5th ed., 2002; A. Mottershead, *Introduction to Electricity and Electronics*, 3d ed., 1990; W. F. Mullin, *ABC's of Capacitors*, 3d ed., 1978; E. M. Purcell, *Electricity and Magnetism*, 2d ed., 1985; H. D. Young and R. A. Freedman, *Sears and Zemansky's University Physics with Modern Physics*, 11th ed., 2003.

Capacitance measurement

The measurement of the ratio of the charge induced on a conductor to the difference in potential between it and a neighboring conductor which induces the charge. In a multiconductor system there are capacitances between each pair of conductors. In general, these capacitances are functions of the total geometry, that is, the location of all of the conducting and dielectric bodies. When, as is usually true, only the capacitance between two conductors is of interest, the presence of other conductors is an undesirable complication. It is then customary to distinguish between two-terminal and three-terminal capacitors and capacitance measurements. In a two-terminal capacitor, either one of the conductors of primary interest surrounds the other (in which case the capacitance between them is independent of the location of other bodies except in the vicinity of the terminals); or the somewhat indefinite contributions of other conductors to the capacitance of interest are accepted.

A three-terminal capacitor consists of two active electrodes surrounded by a third, or shield, conductor. The direct capacitance between the two active electrodes is the capacitance of interest, and, when shielded leads are used, it is independent of the location of all other conductors except the shield. Only certain of the measuring methods to be described are suitable for three-terminal capacitors. For greater accuracy or measurement at higher frequencies, four terminal-pair terminations described below are desirable.

Every physically realizable capacitor when being charged, as in ac use, has associated energy loss in the dielectric and in the metal electrodes. At a single frequency these are indistinguishable, and the capacitor may be represented by an equivalent circuit comprising either a parallel or series combination of pure capacitance and pure resistance. The measurement of capacitance, then, in general involves the simultaneous measurement of, or allowance for, an associated resistive element. *See* PERMITTIVITY.

Most capacitance measurements involve simply a comparison of the capacitor to be measured with a capacitor of known value. This should ideally be a reference standard whose capacitance has been established by an unbroken chain of measurements to the standards of a national metrology institute, which have been realized in terms of fundamental constants of nature. Methods that permit comparison of nearly equally valued capacitors by simple substitution of one for the other at the same point in a circuit are frequently possible and almost always preferable. *See* ELECTRICAL UNITS AND STANDARDS.

Properties of standard capacitors. The most stable capacitors available have a fused quartz dielectric and a value of 10 picofarads. The dielectric is a quartz disk about 2.8 in. (70 mm) in diameter and 0.4 in. (10 mm) thick. Gold film electrodes are formed on the two faces of the disk by sputtering gold onto the fused quartz surfaces, and a guard electrode is similarly deposited onto the circumferential surface. Because fused quartz has a temperature coefficient of permittivity of about 10 parts per million per kelvin (5 parts per million per degree Fahrenheit) it is essential to measure its temperature accurately, and to do this a resistance thermometer is wound around the disk. The unit is hermetically sealed in a container which can be put in an oil-bath or other constant-temperature enclosure. *See* SPUTTERING.

Other standard capacitors with values up to 1 nanofarad can be made with parallel-plate or cylindrical electrodes; they are hermetically sealed and evacuated or filled with dry gas. Capacitors of larger value have a solid dielectric, such as mica or polystyrene or other plastic film, and are less constant in value over time. *See* CAPACITOR.

A three-terminal arrangement is usual for standard capacitors; that is, they have a screening case with a separate connection. A three-terminal capacitor can be measured satisfactorily if the screen can be maintained at the potential of one of the capacitor terminals, but at high frequencies—above about 100 kHz—it becomes increasingly difficult to satisfy this condition, and a two-terminal capacitor may be advantageous. The disadvantage of a two-terminal capacitor is the uncertainty arising from the fringing field of its terminals, which may be a few tenths of a picofarad; however, a change in capacitance will be precise if the fringing field is unaltered. For example, the difference in the value of two capacitors with identical terminal arrangements can be measured accurately even if their individual capacitances cannot be determined with high accuracy. This remark applies even to capacitors with precision coaxial connectors for use up to a few gigahertz. Such capacitors have a series inductance of only a few nanohenries, and they, together with transmission lines of accurately known dimensions, provide the standards

for high-frequency measurements. *See* INDUCTANCE; TRANSMISSION LINES.

At the other end of the frequency range, measurements are required on capacitors up to 1 farad in value at frequencies around 100 Hz. The stability of these capacitors is such that the standards used in their measurement need have an accuracy of no better than 0.01% and are often considerably worse. For this range, the standards comprise a capacitor combined with transformers for the input and associated output connections. A single capacitor, which may be a 1-microfarad mica dielectric capacitor, and multitapped input and output transformers can provide values between 1 μF and 1 F. A better performance is obtained if plastic-film capacitors of value larger than 1 μF are used and a transformer designed to give the optimum properties provides each value of the capacitance.

Frequency-dependence of capacitors. The effective value of a capacitor C_e—that is, that which corresponds to the reactance measured at its terminals—is affected by three factors. First, series inductance l apparently increases the actual capacitance C, according to Eq. (1), where ω is the angular frequency. Second, the properties of the dielectric in a capacitor with a solid dielectric separating the electrodes usually cause a reduction in capacitance as the frequency is raised. Third, mechanical displacement of the electrodes caused by the electric forces between them may alter the capacitance in a voltage- or frequency-dependent manner; this effect should be detectable only in air or gas dielectric capacitors, and can be quite appreciable in high-voltage capacitors. In capacitance standards provided with separate current and potential terminals, the mutual inductance between current and potential leads can also alter the effective value of a capacitor, unless the geometrical layout of the leads is arranged to minimize this effect.

$$C_e = \frac{C}{1 - \omega^2 lC} \tag{1}$$

The series inductance can be calculated from a measurement of the self-resonant frequency f_0 of the capacitor with its terminals connected together by a link producing only a small calculable additional inductance. The total inductance of the capacitor and link circuit is $1/[(2\pi f_0)^2 C]$.

The effect of dielectric permittivity variation is measured by ultimate reference to an air or gas dielectric capacitor whose plates have clean metal surfaces.

The electric forces between the electrodes of standard capacitors produce capacitance changes that are usually too small to be significant, except when the highest accuracy is required at more than a single applied voltage or at a frequency near that of a mechanical resonance.

Distributed capacitance of an inductor. An inductor is a particularly impure circuit component. In addition to the series resistance of the winding, distributed capacitance is always present from turn to turn and layer to layer, making the effective inductance a function of frequency. It is customary to assume that the effect of the distributed capacitance in an inductance coil may be represented by a single capacitor connected between the coil terminals. In some cases the value of this equivalent capacitor may be obtained by a determination of the self-resonant frequency of the coil; however, as a result of the distributed nature of both the inductance and the capacitance, the coil may exhibit several resonance modes, and a self-resonance determination is then ambiguous.

A procedure that usually surmounts this difficulty is that of determining the resonant frequency for several settings of a variable capacitor connected in parallel with the coil. A plot of the capacitance of the observed auxiliary capacitor against the reciprocal of the square of the corresponding resonant frequency and extrapolated to infinite frequency then gives a value for the lumped equivalent of the distributed capacitance. These measurements are now made much more conveniently by using a commercial impedance-measuring instrument (a bridge or an impedance analyzer) whose source frequency can be varied either manually or automatically. Connection to the instrument is now most often made by using a four-terminal-pair configuration, via an adapter if necessary.

AC bridge comparison methods. When capacitors must be compared with high accuracy, bridge methods must be adopted. *See* BRIDGE CIRCUIT; WHEATSTONE BRIDGE.

Bridges based on a resistance ratio. These are Wheatstone-bridge configurations in which the potential division of the capacitor being measured and either a parallel combination of a standard loss-free capacitor C_s and a conductance G_s (**Fig. 1***a*) or a series combination of C_s and a resistor R_s (Fig. 1*b*) is equated, when the detector is nulled, to the ratio of potentials across resistors R_1 and R_2. More commonly now, the reference potential division is that of a variable-ratio autotransformer known as an inductive voltage divider (IVD). *See* INDUCTIVE VOLTAGE DIVIDER.

If both the standard and unknown capacitor are represented by their parallel equivalent circuits (Fig. 1*a*), equality of potential division leads to Eqs. (2) and (3) for the unknown capacitance C_x

$$C_x = C_s \frac{R_1}{R_2} \tag{2}$$

$$G_x = G_s \frac{R_1}{R_2} \tag{3}$$

and conductance G_x; while, for their series equivalent circuits (Fig. 1*b*), this equality leads to Eqs. (4) and (5) for C_x and the unknown resistance R_x.

$$C_x = C_s \frac{R_1}{R_2} \tag{4}$$

$$R_x = R_s \frac{R_2}{R_1} \tag{5}$$

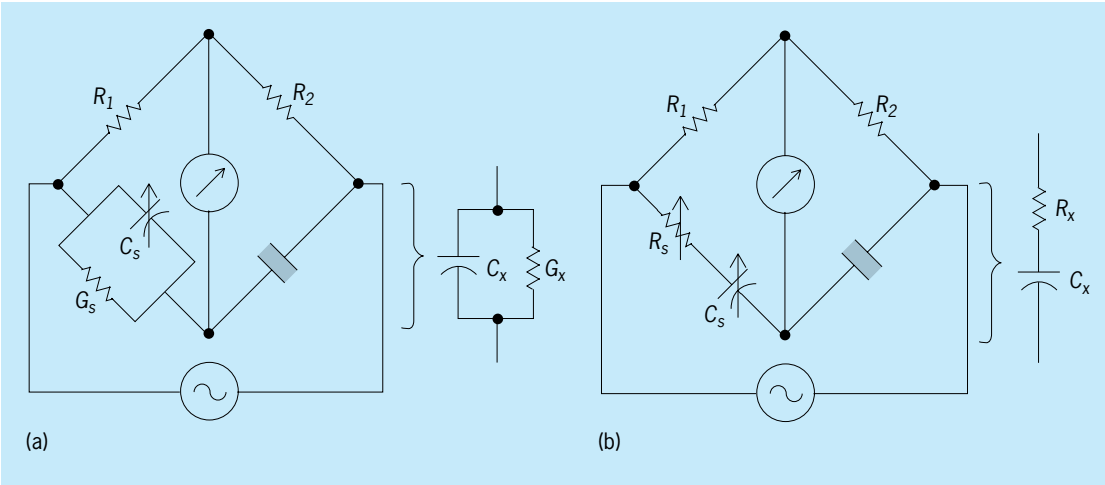

Fig. 1. Resistance-ratio bridges, in which two adjacent arms are resistors. (*a*) Parallel arrangement. (*b*) Series arrangement.

Schering bridge. This bridge yields a measurement of the equivalent series-circuit representation of a capacitor (**Fig. 2**). The equations of balance are written as (6) and (7). As with any balanced bridge network, the positions of source and detector may be interchanged without affecting the balance condition. This change may be done to gain an advantage in, for example, sensitivity to the balance condition or the voltage rating of some component.

$$C_x = C_s \frac{R_1}{R_2} \tag{6}$$

$$R_x = R_2 \frac{C_1}{C_s} \tag{7}$$

Wagner branch. The resistance-ratio and Schering bridges are useful for two-terminal capacitance measurements. Their use may be extended to three-terminal measurements and extended in accuracy and range by the introduction of shielding and the addition of the Wagner branch which maintains a shield at the potential of the terminals connected to the detector (**Fig. 3**).

Balance is effected by the convergent procedure of alternately adjusting the main bridge arms with the switch open and adjusting the Wagner branch elements with the switch closed. The final balance is obtained with the switch open. Capacitances to the shield at two opposite corners (A and C) of the bridge then shunt the Wagner arms and do not affect the balance of the main bridge. Capacitances to the shield at the other two corners (B and D) carry no current at balance and therefore cannot introduce error either.

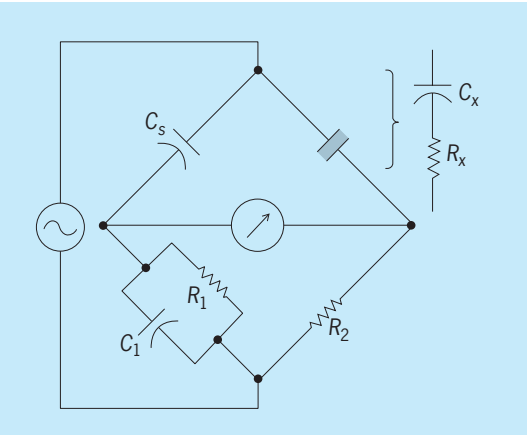

Fig. 2. Schering bridge for measurement of the equivalent series-circuit representation of a capacitor.

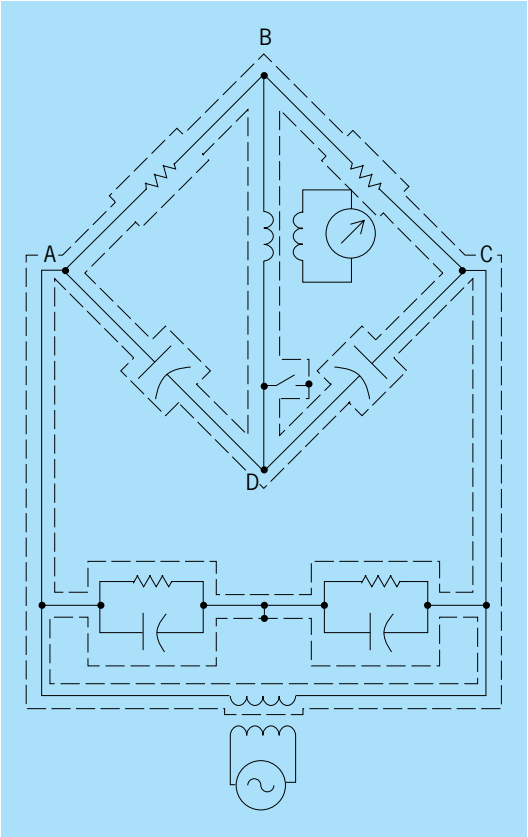

Fig. 3. Shielded resistance ratio-arm bridge with a Wagner branch for three-terminal measurements. With the switch open, capacitances to the shield at A and C shunt the Wagner arms. At balance, capacitances to the shield at B and D carry no current.

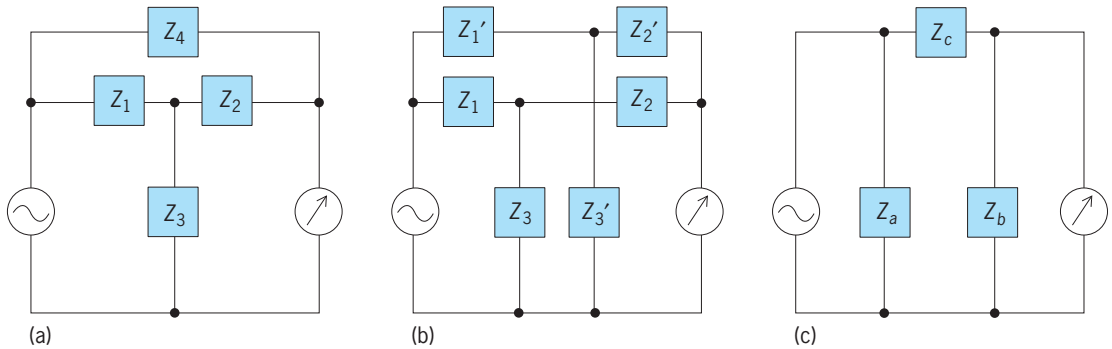

Fig. 4. T networks. (*a*) Bridged-T. (*b*) Parallel-T. (*c*) Equivalent circuit. Balance is achieved by varying network components untill the equivalent impedance Z_c is infinite.

Bridged-T and parallel-T networks. These complex bridges possess a significant advantage over four-arm bridges for medium-accuracy radio-frequency measurements of capacitance because the source and detector have a common point of connection. (**Fig. 4**). Use of the Y-delta network transformation permits reduction of either the bridged-T (Fig. 4*a*) or the parallel-T (Fig. 4*b*) to an equivalent circuit (Fig. 4*c*). Balance, or null indication of the detector, is achieved by variation of the network components until the equivalent impedance Z_c linking the source and the detector is infinite. *See* Y-DELTA TRANSFORMATIONS.

Time-constant methods. If a direct voltage is suddenly applied to the series combination of a resistor and an initially discharged capacitor, the charge and the voltage on the capacitor increase exponentially toward their full magnitudes with a time constant equal in seconds to the product of the resistance in ohms and the capacitance in farads. Similarly, when a charged capacitor is discharged through a resistor, the charge and the voltage decay with the same time constant. Various methods are available for the measurement of capacitance by measurement of the time constant of charge or discharge through a known resistor. *See* TIME CONSTANT.

In one such method the time required for the output voltage of an operational amplifier having a capacitor as a feedback component to increase to a value equal to the step-function input voltage applied through a resistor to its input is determined by an electronic voltage-comparison circuit and timer. With the assumption of ideal characteristics for the amplifier, such as infinite gain without feedback, infinite input impedance, and zero output impedance, the measured time interval is equal to the product of the values of the known resistance and the capacitance being measured. *See* OPERATIONAL AMPLIFIER.

Thompson-Lampard capacitor. The values of capacitors of similar shapes but different sizes are proportional to size; that is, a single length dimension governs their values. A. M. Thompson and D. G. Lampard showed how this principle could be applied to make a precise and practical calculable capacitor. Such capacitors require only a single length to be measured accurately, and this is readily accomplished by optical interferometry. Their design provides the standard for the unit of capacitance defined in the International System of Units (SI).

Thompson and Lampard's theorem relates to a capacitor which, in its simple form, comprises four equal electrodes of parallel round bars arranged with their axes at the corners of a square and with small gaps between adjacent bars (**Fig. 5**). The cross-capacitances (C_A and C_B) between the two diagonally opposite electrodes per unit length of the infinite system in vacuum is given by Eq. (8). Here,

$$\exp\left(\frac{-\pi C_A}{\varepsilon_0}\right) + \exp\left(\frac{-\pi C_B}{\varepsilon_0}\right) = 1 \qquad (8)$$

the permittivity of vacuum $\varepsilon_0 = 8.8541878$ pF/m. If the two cross-capacitances are identical, each cross-

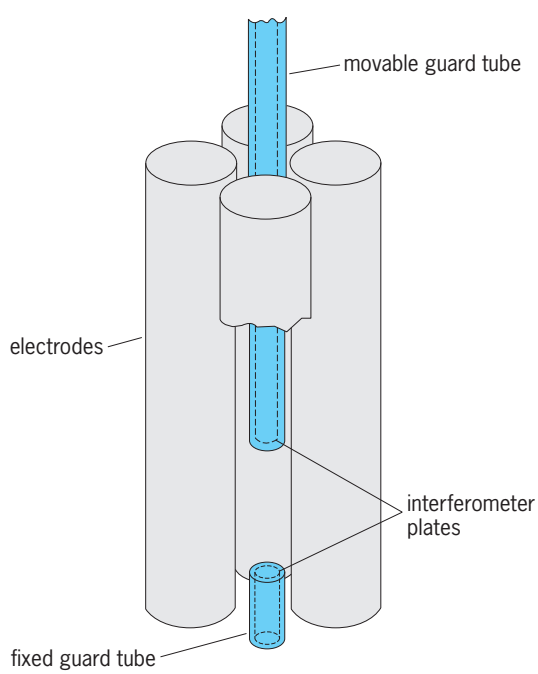

Fig. 5. Calculable capacitor with an interferometer for length measurement. Surrounding shields are omitted.

capacitance is given by Eq. (9). If the two cross-

$$C_0 = 1.95354902 \text{ pF/m} \tag{9}$$

capacitances are not quite equal and differ by ΔC, the mean capacitance is given approximately by Eq. (10).

$$C = C_0[1 + 0.0866(\Delta C/C)^2] \tag{10}$$

Shields around the assembly and at the ends of the capacitor connected to a Wagner branch eliminate from a measurement all capacitances except the required internal ones. The capacitor is in an evacuated enclosure to avoid uncertainties arising from the relative permittivity and the refractive index of air.

If a conducting bar or tube connected to the shield is inserted into the central space between the electrodes, with only a small gap between the bar and the electrodes, the cross-capacitance is zero. Therefore, with a short, fixed tube inserted at one end, a long tube that can be moved axially at the other alters the active length of the capacitor. The change of length is measured interferometrically using reflectors mounted at the inner ends of the fixed and movable tubes. The consequent change in mean capacitance is then calculated from Eq. (10). There is a minimum separation of the tubes—about three times the space between opposing electrodes—below which the capacitance change with length becomes significantly nonlinear.

A practical cross-capacitor departs slightly from the ideal geometry in a number of ways, and consequently the accuracy is limited to about 1×10^{-8} pF; if the working range of the capacitor is 0.5 pF (equivalent to a change of length of 10 in. or 25 cm), the overall accuracy on this account is 2 parts in 10^8. The repeatability is appreciably better.

Four-terminal-pair definition. Four-terminal-pair definition of impedance standards arose from the need to measure them with sub-part-per-million accuracy. All electric and magnetic fields associated with an impedance standard must be completely defined, and the standard must be provided with separate pairs of current and potential terminals, as in the familiar four-terminal dc resistance standards. The solution is to enclose the impedance in a complete shield and to treat this shield also as a four-terminal impedance. When connected by coaxial cables to a coaxial measuring system in which the currents through the outer shield conductors of the cables and the currents through the inner conductors going to the impedance are equal and opposite (both currents being zero in the potential cables), all electric and magnetic fields are reproducibly confined within the shield. The voltage to be measured is that generated between the inner and outer coaxial terminations of the cable on one side (the "high" or H side) of the impedance, and the current is that emerging from the inner conductor of the termination at the other side (the "low" or L side). This current is closely equal to the current

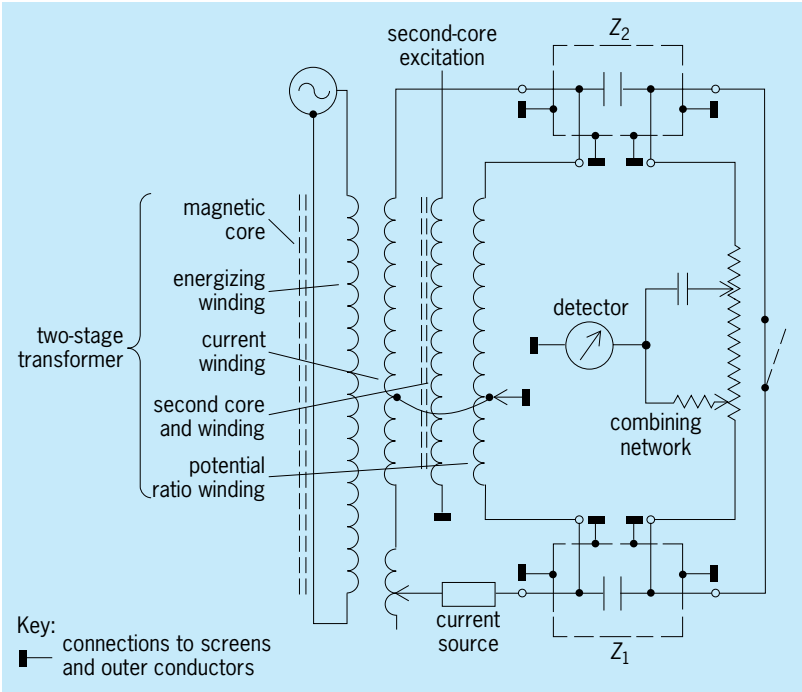

Fig. 6. Transformer bridge for the precise comparison of two similar types of impedances Z_1 and Z_2. Screens and outer conductors of the coaxial connections have been omitted for clarity, but connections to them are shown as black rectangles.

returning via the outer conductor of the same termination. There is one further condition, that the voltage between the inner and outer conductors of the second low coaxial termination should be zero. The second high coaxial termination has a source of current connected across it in order to make the measurement. The magnitude of this current is not relevant to the measured value of an impedance defined in this way since this impedance is the voltage-to-current ratio. One consequence is that the small four-terminal impedance of the outer screen is added to that of the enclosed impedance and thus becomes a part of this seemingly elaborate but complete four-terminal-pair definition.

The defining conditions can be departed from provided that the value measured is demonstrably not thereby altered. Small corrections to account for the finite impedance of the potential- and current-measuring cables are independent of the value of the impedance being measured and are negligible for all but the most accurate measurements or those at the highest frequencies.

Calibration of standard capacitors. The Thompson-Lampard capacitor provides the starting point for the calibration of standard capacitors, which are compared in 1:10 ratios by using an accurately calibrated transformer in a coaxial alternating-current bridge. A four-terminal-pair bridge based on a voltage ratio transformer for comparing capacitance standards is illustrated in **Fig. 6**.

Every mesh of the bridge network has a current equalizer which ensures that equal currents of opposite sign flow in the inner and the outer conductors of every coaxial connection. By using this

coaxial equalized-current technique, the circuitry neither emits significant electromagnetic fields nor responds to external ones. Therefore, the measurements are unaffected by the layout of the conductors or by interference.

The transformer is of the two-stage type; that is, there is an energizing winding on the first core followed by a winding which supplies the current to the impedances to be compared, and finally, a second core and overall winding, which provides a precise potential ratio for comparing the two four-terminal-pair impedances. *See* INSTRUMENT TRANSFORMER.

The bridge is balanced by injecting a small known voltage into the potential circuit of one of the impedances. Adjustment of the combining network for the detector is also required. For this, the switch in the conductor joining the output current connections of the two impedances is opened, the detector is rebalanced by adjusting the combining network, and the switch is closed.

Commercial impedance-measuring instruments. As discussed above, manual bridge systems can make highly accurate impedance measurements, but are time-consuming to operate and have limited frequency range. In contrast, modern automatic inductance-capacitance-resistance (LCR) and impedance analyzer instruments have become commercially available that are claimed to offer moderate accuracies, operate at millisecond speeds, and cover a wider frequency range. In these precision instruments, an attempt is also made to implement the Cutkosky four-terminal-pair (4TP) definition given by Eq. (11). The subscripts refer to the ports of a four-

$$Z_{4\mathrm{TP}} = \left.\frac{V_2}{I_4}\right|_{I_2=I_3=V_3=V_4=0} \tag{11}$$

terminal-pair impedance standard or device under test (DUT), as shown in **Fig. 7**. The definition of $Z_{4\mathrm{TP}}$ given by Eq. (11) is also illustrated in Fig. 7, as well as the notation commonly used on instruments for current (high = H_c and low = L_c) and potential (high = H_p and low = L_p) port identification. The advantage of four-terminal-pair impedance measurements (as opposed to two-terminal or three-

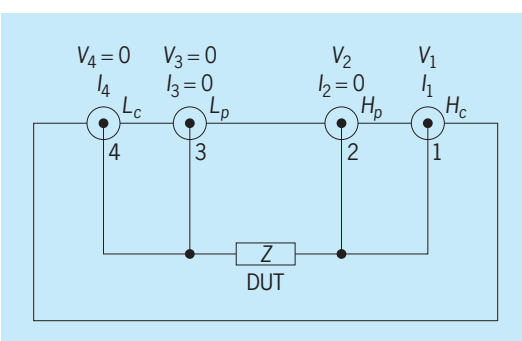

Fig. 7. Four-terminal-pair impedance component or device under test (DUT) and its complete definition, equivalent to that given in Eq. (11).

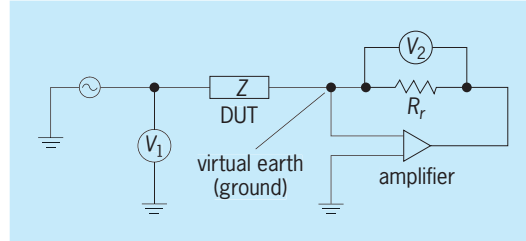

Fig. 8. Autobalancing bridge method employed in modern analyzers for measuring the four-terminal-pair impedance of a device under test (DUT). The outer shield conductors have been omitted for simplicity.

terminal) is that they remove the influence of contact impedances between the cables and the impedance being measured, as well as minimizing the effects of stray capacitance and mutual inductance between cables, assuming current-equalized networks.

The impedance measurement and frequency range of operation of some of the latest precision four-terminal-pair impedance analyzer instruments is typically 1 mΩ to 1 GΩ and few tens of hertz to 100 MHz, respectively. In addition to the alternating-current voltage that these instruments use for impedance measurements (maximum usually 1 V_{rms}), they can provide up to 40 V direct-current bias voltage for various applications. The claimed measurement uncertainty varies depending on the impedance being measured and the test frequency, but typically increases from 0.1 to 10% at the extremes of their impedance capability or at the higher frequencies. The application of such instruments is wide-ranging, from electrical and electronic component and material evaluation to chemical and biological process measurements, in a number of industries (aerospace, biomedical, research and development, and telecommunications).

Most impedance-measuring instruments (which are derivatives of the earlier models) now employ the autobalancing bridge (ABB) method, in which high-input impedance amplifiers and feedback techniques are used to create a virtual Earth (ground) at the low-potential port of the instrument (**Fig. 8**). The instruments usually connect in one of four range resistors (R_r) across which a voltage (V_2) is detected to determine the current flow through the impedance being measured. Often an assumption is made that the direct-current value of these range resistors is the same as their high-frequency value, but this is not entirely true for accurate measurements. In addition, over broad impedance and frequency ranges the defining conditions (Fig. 7) are not implemented to the required level. Several other limitations are also present in the autobalancing bridge method, such as shunt capacitance loading at the high- and low-potential ports, sufficiently high input impedance of the two voltage detectors (V_1 and V_2) over a broad frequency range, and equality of the currents in the inner and shield conductors. Even to achieve the modest accuracy claimed for modern precision impedance-measuring instruments,

significant demand is placed on the very latest electronic passive and active components. For measurements at frequencies higher than 100 MHz, two-port, 50-Ω, matched-impedance or vector network analyzers are normally employed to at least 65 GHz. *See* MICROWAVE IMPEDANCE MEASUREMENTS; MICROWAVE MEASUREMENTS; RADIO-FREQUENCY IMPEDANCE MEASUREMENTS.

For quality-control applications where pass/fail measurements must be performed on passive components relatively quickly, instrument manufacturers supply a number of cable probes and adaptors. One combination of extension cables and connection commonly supplied is known as Kelvin clips, in which the four coaxial cables are connected to the four-terminal-pair impedance analyzer at one end, and at the other end two crocodile clips with independently insulated jaws are provided for connection to a two-terminal component. This configuration essentially converts the four terminal-pair connection at the impedance analyzer to a four-terminal connection scheme at the crocodile clips. The advantage of the Kelvin clips is that the two surfaces of each crocodile clip form the current and potential connections at the passive impedance component, with one clip being the energization side and the other the detection. The Kelvin clip scheme is also employed in various four-terminal-pair to four-terminal adaptors for repeatable measurement of surface-mount (or chip) passive components.

At the other end of the performance range, there are numerous low-accuracy two-terminal handheld LCR meters. These typically operate at one or two fixed frequencies (120 Hz and 1 kHz being the most common) and measure a wide range of inductance-capacitance-resistance values as well as the loss factor, but offer at best a 1% uncertainty. *See* CAPACITANCE; ELECTRICAL MEASUREMENTS.

Bryan P. Kibble; F. Ralph Kotter; G. H. Rayner; Shakil A. Awan

Bibliography. D. Buchla and W. Mclachlan, *Electronic Instrumentation and Measurement*, 1992; W. J. Duffin, *Electricity and Magnetism*, 4th ed., 1990; B. Hague and T. R. Foord (ed.), *Alternating Current Bridge Methods*, 6th ed., 1971; P. Horowitz and W. Hill, *The Art of Electronics*, 2d ed., 1989; B. P. Kibble and G. H. Rayner, *Coaxial A. C. Bridges*, 1984; R. Morrison, *Grounding and Shielding Techniques in Instrumentation*, 4th ed., 1998; L. Schnell, *Technology of Electrical Measurements*, 1993.

Capacitance multiplication

The generation of a capacitance which is some multiple of that of an actual capacitor. Capacitance multiplication circuits have an input impedance which is capacitive and which is proportional to that of an actual capacitor appearing somewhere in the circuit. In most applications, capacitance multiplication circuits are used to generate an equivalent input capacitance which is much larger than that of the actual capacitor. One scenario where capacitance multiplication might prove useful is where a physical capacitor of a required capacitance may be too large, too expensive, or unavailable. A second is where the capacitor must be reasonably large and capable of handling bidirectional signals, thus precluding the direct use of widely available electrolytic capacitors and hence making practical the utilization of a much smaller nonelectrolytic capacitor in a capacitance multiplication circuit. *See* CAPACITANCE.

Capacitance multiplication circuits are often made from operational amplifiers and resistors along with the capacitor that is to be scaled, although transistors and other active devices can also be used. Capacitance multiplication circuits are closely related to classes of circuits termed generalized immittance converters and negative impedance converters. Generalized immittance converters are used to generate an equivalent input impedance that is proportional to products or quotients of specific impedances that appear in the circuit. Generalized immittance converters are often used for capacitance multiplication. Related applications of generalized immittance converters are synthetic inductance simulation and negative-resistance generation. Negative-impedance converters are also used for the generation of negative impedances. *See* ELECTRICAL IMPEDANCE; IMMITTANCE; INDUCTANCE; NEGATIVE-RESISTANCE CIRCUITS; OPERATIONAL AMPLIFIER.

A generalized immittance converter circuit is shown in **illus.** *a*. The input impedance Z_{in} of this circuit is given by Eq. (1). If the impedances

$$Z_{in} = \frac{Z_1 Z_3 Z_5}{Z_2 Z_4} \tag{1}$$

Z_1–Z_4 are resistors and Z_5 is a capacitor, the capacitance multiplication circuit of illus. *b* is obtained. The equivalent impedance of this circuit is that of a capacitor of value C_{eq} given by Eq. (2). The resis-

$$C_{eq} = \frac{C R_2 R_4}{R_1 R_3} \tag{2}$$

tor ratio $(R_2R_4)/(R_1R_3)$ determines the capacitance multiplication factor. A second capacitance multiplication circuit based upon the Miller effect is shown in illus. *c*. The equivalent capacitance of this circuit is given by Eq. (3).

$$C_{eq} = C\left(1 + \frac{R_2}{R_1}\right) \tag{3}$$

Capacitance multiplication circuits have some practical limitations. For example, the input impedance of the circuits of illus. *b* and *c* will remain capacitive only provided the operational amplifiers are operating linearly. This requirement places some restrictions on how capacitance multiplication circuits can be used. For example, the output voltage swing, the output current, and the slew rate of the operational amplifier all have hard limits imposed by the design of the operational amplifier itself, and these limits must not be violated at either internal

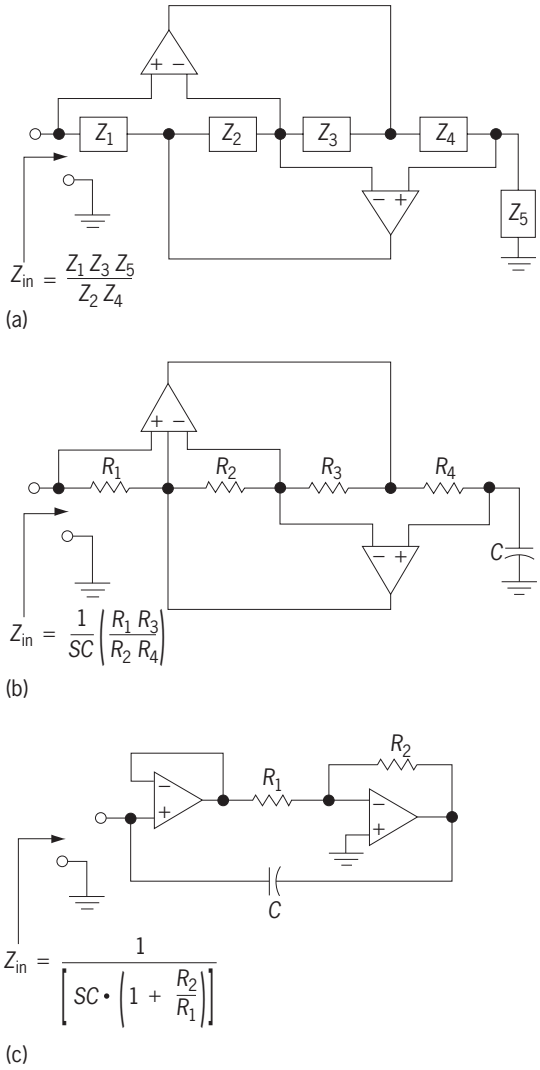

Impedance-generation circuits. (*a*) Generalized immittance converter. (*b*) Capacitance multiplier based on generalized immittance converter. (*c*) Miller capacitance multiplier. Input impedance Z_{in} is given in Laplace transform notation; *S* is the Laplace transform variable.

or external nodes of the capacitance multiplication circuits. The bandwidth of the operational amplifier is also limited. These limits on the operational amplifier restrict the magnitude and frequency of signals that can be placed on capacitance multiplication circuits, as well as how fast these capacitors can be discharged. Although violation of these limits is usually nondestructive (provided the input voltage does not exceed the maximum allowed for the operational amplifier), violations will result in significant signal distortion. These limitations become increasingly problematic when realizing very large capacitors or when using large capacitance multiplication factors. Other capacitance multiplication circuits beyond those shown in the illustration are available which offer varying trade-offs between capacitance multiplication, signal swing, frequency response, and complexity. Randall L. Geiger

Bibliography. L. T. Bruton, *RC-Active Circuits, Theory and Design*, 1980; A. Sedra and K. Smith, *Microelectronic Circuits*, 4th ed., 1997; F. W. Stephenson (ed.), *RC Active Filter Design Handbook*, 1985.

Capparales

An order of flowering plants (Angiospermae) of approximately 15 families of dicotyledons with over 4000 species. In molecular phylogenetic classifications, it is placed near Malvales, Myrtales, and Sapindales. *See* MALVALES; MYRTALES; PLANT KINGDOM; SAPINDALES.

The mustard family, Brassicaceae (Cruciferae), with about 3000 species, and the caper family, Capparaceae, with about 800 species, form the core of the order. An additional 200 or so species are treated in 13 families, including Caricaceae (papaya), Limnanthaceae (meadowfoam), Resedaceae (mignonette), and Tropaeolaceae (garden nasturtium). The plants are diverse in vegetative habit and in floral and fruit morphology. However, they commonly possess myrosin cells, containing the enzyme myrosinase, and produce mustard oil glucosides (glucosinolates), the breakdown products of which are the pungent compounds of radish, wasabi, horseradish, and capers. Primitive members of the order have five-parted flowers, but the core families have four-parted flowers. Common vegetables such as broccoli, brussels sprouts, cabbage, kale, radish, rutabaga, turnip, and water cress belong to the Brassicaceae. Capparales are also important sources of seed oils, from Brassicaceae (*Brassica*) and Moringaceae (*Moringa*), and of several ornamentals, such as candytuft (*Iberis*) and sweet alyssum (*Lobularia*) in Brassicaceae, spider flower (*Cleome*) in Capparaceae, mignonette (*Reseda*) in Resedaceae, and garden nasturtium (*Tropaeolum*) in Tropaeolaceae. *See* BROCCOLI; BRUSSELS SPROUTS; CABBAGE; CAULIFLOWER; COLLARD; CRESS; HORSERADISH; KALE; KOHLRABI; MUSTARD; ORNAMENTAL PLANTS; RADISH; RAPE; RUTABAGA; TURNIP. James E. Rodman

Capricornus

The Sea Goat, a southern zodiacal constellation (see **illustration**). It is the source of the terrestrial name Tropic of Capricorn, for the parallel of latitude, about $23^1/_2{}^\circ$ south of the Equator, where the Sun is overhead at noon on the day of the winter solstice in the Northern Hemisphere (the summer solstice in the Southern Hemisphere, about December 22). The winter solstice used to occur when the Sun was in Capricornus, but because of the precession of the equinoxes this point has moved to Sagittarius. *See* PRECESSION OF EQUINOXES; SAGITTARIUS; TROPIC OF CAPRICORN; ZODIAC.

Capricornus has no bright stars. In the early star atlases, it was drawn as a goat with a fish's tail. In Greek mythology, the god Pan, posing as a goat, dived into a river to escape from a demon. In his panic, he changed only his rear half. He surfaced and saved

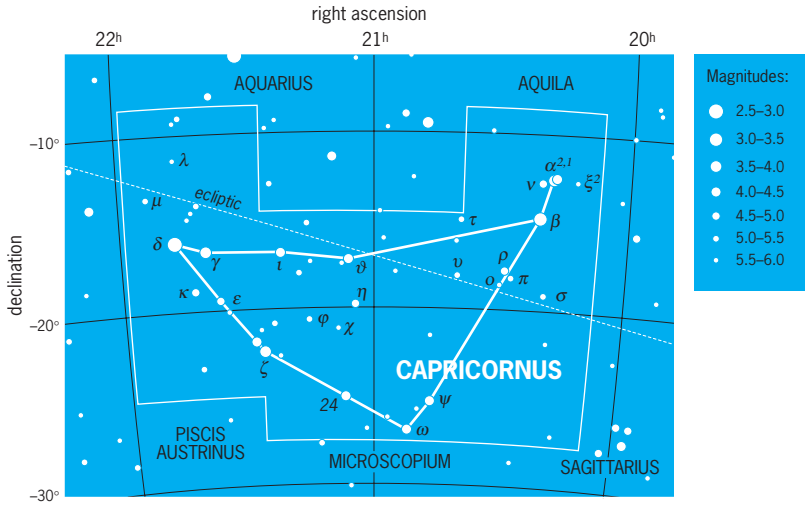

Modern boundaries of the constellation Capricornus, the Sea Goat. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

Zeus, who had been injured and who rewarded the chimera by placing it in the heavens.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Caprimulgiformes

An order of crepuscular (active during predawn or twilight hours) or mainly nocturnal birds collectively known as the goatsuckers. The group is apparently most closely related to the owls (Strigiformes) and is found worldwide, mainly in the tropics and warm temperate regions. Species breeding in the Arctic and cooler temperate regions are migratory. *See* STRIGIFORMES.

Classification. The order Caprimulgiformes is divided into the suborder Steatornithes, containing the single family Steatornithidae (oilbirds; 1 species; northern South America); and the suborder Caprimulgi, including the families Podargidae (frogmouths; 13 species; tropics, India to Australasia), Aegothelidae (owlet-frogmouths; 8 species; tropics, Australasia), Nyctibiidae (potoos; 6 species; Neotropics), and Caprimulgidae (nightjars or goatsuckers; 77 species; worldwide). The largest family, Caprimulgidae is usually divided into two subfamilies, the Caprimulginae (nightjars; worldwide) and the Chordeilinae (nighthawks; New World).

Fossil record. The goatsuckers have a poor fossil record, but some interesting forms are known from the early Tertiary well outside the current range of the different families. Specimens assigned to the oilbirds, the frogmouths, and the owlet-frogmouths have been reported from the Eocene of France.

Characteristics. The goatsuckers are primarily insectivorous, although the large frogmouths also feed on small vertebrates; the oilbird is unique in feeding on fruits, especially of palms. Most goatsuckers have a huge, generally weak mouth with a flexible lower jaw that can expand widely and is surrounded by long, stout bristles; it makes an effective trap for insects caught in flight. Frogmouths and owlet-frogmouths have strong jaws.

The wings are well developed, but the feet are weak and serve mainly for perching. The plumage is soft and fluffy, and is mottled and barred brown and gray, serving as excellent cryptic protective coloration. Tree-dwelling forms, especially frogmouths and potoos, sit quietly on a tree with their body, neck, and head stretched; with their plumage matching the tree bark, the individuals look like a broken branch. White patches may exist on the wings, tail, and throat, which are visible only in flight. Some species have greatly elongated tail feathers, and a few have very elongated primary feathers forming pennants from their wings.

Goatsuckers are generally solitary but may migrate in loose flocks (nighthawks). Most species are highly vocal, using calls to attract mates and defend territories. The English names of a number of species, such as the whippoorwill and chuck-will's-widow, are based on their calls. The Caprimulgidae nest on the ground, but all other members of the Caprimulgi nest in trees; the clutch is from one to five eggs. Young are downy and remain in the nest until they are able to fly.

A few species are known to hibernate. Poorwills in the deserts of the southwestern United States hibernate in rock holes for several months at a time, often returning to the same hole in successive winters. Their body temperature falls from a normal 104–105°F (40–41°C) to 64–66°F (18–19°C).

Oilbirds. The peculiar South American oilbirds are colonial nesters, placing their nest of seeds and droppings on ledges deep in caves. Paired adults remain together throughout the year, roosting on their breeding ledge. The young become very fat, reaching

twice the weight of the adults; they were once harvested in large numbers and boiled down for cooking oil. Oilbirds have excellent night vision, but inside their often totally dark caves they find their way by using echolocation based on pulses of sound of about 7000 Hz, which are audible to the human ear. *See* AVES; ECHOLOCATION. Walter J. Bock

Bibliography. N. Cleere, *Nightjars: A Guide to the Nightjars, Nighthawks and Their Relatives*, Yale University Press, 1998; D. T. Holyoak, *Nightjars and Their Allies*, Oxford University Press, 2001.

Captorhinida

A moderately coherent group of primitive reptiles constituting an order of the subclass Anapsida. Most members are characterized by a closed cheek (temporal) region. The order is divided into four suborders: Captorhinomorpha, Millerosauria, Procolophonia, and Pareiasauria. Except for the Procolophonia, which continue into the Late Triassic, Captorhinida are confined to the Permo-Carboniferous. They lived in lowlands, where they were associated with amphibians and synapsid reptiles. Some of the smaller animals of the Captorhinomorpha and the Millerosauria fed primarily on insects and small vertebrates, but the larger genera were exclusively herbivorous. Along with the caseid pelycosaurs, they were the dominant consumers of vegetation in the middle Permian ecosystems.

At one time the Captorhinida were joined with the Diadectamorpha and Seymouriamorpha to form the order Cotylosauria (stem reptiles). This assemblage, now known to include both reptiles and amphibians, has been dropped because it is polyphyletic and also because it has carried many different meanings.

The Captorhinomorpha are the most diversified of the Captorhinida; the best-known genus is *Captorhinus* of the early Permian (**illus.** *a*). It was preceded, however, by more primitive genera from the upper Carboniferous and very early Permian and succeeded by specialized genera during the later portions of the Permian. *Captorhinus* and its successors were notable in the possession of multiple rows of teeth on the jaws and palates, related to increasingly herbivorous diets. The captorhinomorph evolutionary radiation culminated in a large, massive genus, *Moradisaurus*, from the upper Permian of Niger, Africa; it measured about 6 ft (2 m) in length.

Millerosasaurs are poorly known small reptiles from the middle Permian of South Africa. Unlike most Captorhinida, they developed partially to fully open cheek regions. They have been considered to be close relatives of the lizards, but currently are thought to be merely a somewhat specialized branch of the Captorhinida.

Procolophonia were small to medium-sized animals with generalized skeletons and specialized skulls. Skulls of some of the Late Triassic procolophons such as *Hypsognathus* were set with sharp, bony projections somewhat reminiscent of those in living horned lizards. The teeth of procolophons were arrayed in single rows and were rather blunt, suggesting a herbivorous diet.

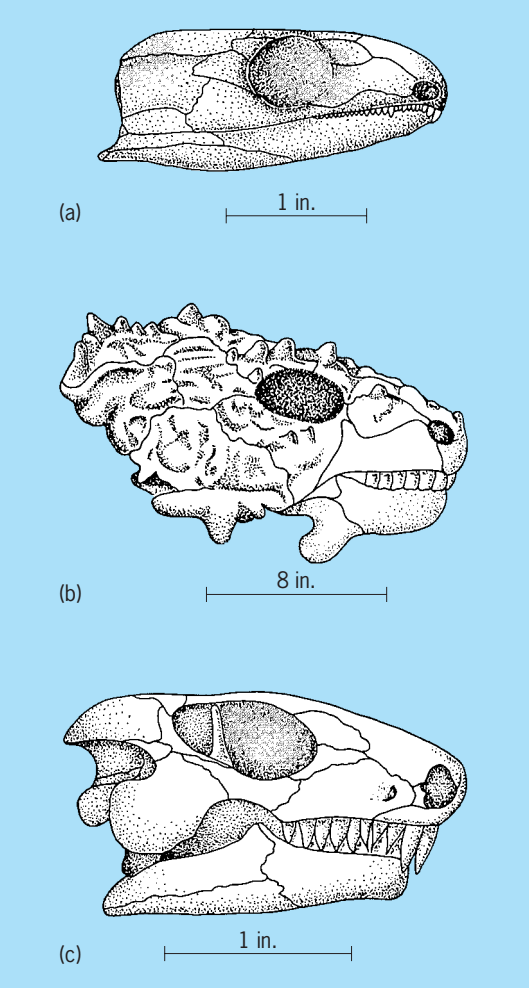

Captorhinid skulls. (*a*) *Captorhinus*. (*b*) *Pareiasaurus*. (*c*) *Procolophon*.

Pareiasaurs were the giants of the Captorhinida, rivaled only by *Moradisaurus*. They ranged up to about 6 ft (2 m) in length and were characterized by broad vertebrae, heavy limbs, and massive feet. Pareiasaurs were prominent herbivores during the middle and late Permian of Eurasia and Africa, where their skeletons occasionally occur in profusion in deposits formed in ancient waterholes. During the late Permian their role as primary consumers was gradually taken over by evolving herbivorous therapsids. *See* ANAPSIDA; REPTILIA; SYNAPSIDA; THERAPSIDA. Everett C. Olson

Bibliography. C. L. Carroll, *Vertebrate Paleontology and Evolution*, 1987; A. S. Romer, *Osteology of the Reptiles*, 1956, reprint 1997.

Carat

The unit of weight now used for all gemstones except pearls. It is also called the metric carat (m.c.). By international agreement, the carat weight is set at 200 milligrams (0.00704 oz). Pearls are weighed in

grains, a unit of weight equal to 50 mg (0.0018 oz), or $^1/_4$ carat.

Despite the great value of many of the gemstones weighed against a unit called the carat, the weight differed from country to country until early in the twentieth century. Although most of the weights to which the name was applied were near 205 mg (0.00722 oz), the range was from less than 190 to over 210 mg (less than 0.00669 to over 0.00739 oz). The 200-mg (0.00704-oz) carat was proposed in 1907 in Paris. This weight slowly gained acceptance, and by 1914 most of the important nations in the gem trade had accepted the metric carat as the legal standard for gemstones.

The word carat comes from the Greek word for the locust tree that is common in the Mediterranean area. This tree produces seeds that are fairly uniform in weight, averaging about 205 mg (0.00722 oz) each. As a result, locust seeds came to be used for gem weight comparisons.

The application of the term carat as a unit of weight must not be confused with the term karat used to indicate fineness or purity of the gold in which gems are mounted. *See* GEM.

Richard T. Liddicoat, Jr.

Caraway

An important spice from the fruits of the perennial herb *Carum carvi*, of the family Apiaceae. A native of Europe and western Asia, it is now cultivated in many temperate areas of both hemispheres. The small, brown, slightly curved fruits are used in perfumery, cookery, confectionery, in medicine, and for flavoring beverages. *See* APIALES; SPICE AND FLAVORING.

Perry D. Strausbaugh; Earl L. Core

Carbohydrate

A term applied to a group of substances which include the sugars, starches, and cellulose, along with many other related substances. This group of compounds is vital in the lives of plants and animals, both as structural elements and in the maintenance of functional activity. Plants are unique in that they have the power to synthesize carbohydrates from carbon dioxide and water in the presence of the green pigment chlorophyll through the energy derived from sunlight, by the process of photosynthesis. This process is responsible not only for the existence of plants but for the maintenance of animal life as well, since animals obtain their entire food supply directly or indirectly from the carbohydrates of plants. *See* PHOTOSYNTHESIS.

The carbohydrates as a group are comparable in importance with the proteins and fats. Cane or beet sugar, D-glucose, D-fructose, starch, and cellulose may be cited as typical representatives. A number of members of this group are of great industrial importance. Among undertakings dependent on carbohydrate materials are the cotton industry, explosives manufacture, brewing, and manufacture of alcohol.

The term carbohydrate originated in the belief that naturally occurring compounds of this class, for example, D-glucose ($C_6H_{12}O_6$), sucrose ($C_{12}H_{22}O_{11}$), and cellulose $(C_6H_{10}O_5)_n$, could be represented formally as hydrates of carbon, that is, $C_x(H_2O)_y$. Later it became evident that this definition for carbohydrates was not a satisfactory one. New substances were discovered whose properties clearly indicated that they had the characteristics of sugars and belonged in the carbohydrate class, but which nevertheless showed a deviation from the required hydrogen-to-oxygen ratio. Examples of these are the important deoxy sugars, D-deoxyribose, L-fucose, and L-rhamnose, the uronic acids, and such compounds as ascorbic acid (vitamin C). The retention of the term carbohydrate is therefore a matter of convenience rather than of exact definition. A carbohydrate is usually defined as either a polyhydroxy aldehyde (aldose) or ketone (ketose), or as a substance which yields one of these compounds on hydrolysis. However, included within this class of compounds are substances also containing nitrogen and sulfur.

The properties of many carbohydrates differ enormously from one substance to another. The sugars, such as D-glucose or sucrose, are easily soluble, sweet-tasting and crystalline; the starches are colloidal and paste-forming; and cellulose is completely insoluble. Yet chemical analysis shows that they have a common basis; the starches and cellulose may be degraded by different methods to the same crystalline sugar, D-glucose.

Classification. The carbohydrates usually are classified into three main groups according to complexity: monosaccharides, oligosaccharides, and polysaccharides.

Monosaccharides. These simple sugars consist of a single carbohydrate unit which cannot be hydrolyzed into simpler substances. These are characterized, according to their length of carbon chain, as trioses ($C_3H_6O_3$), tetroses ($C_4H_8O_4$), pentoses ($C_5H_{10}O_5$), hexoses ($C_6H_{12}O_6$), heptoses ($C_7H_{14}O_7$), and so on. A monosaccharide may be an aldose or a ketose, depending on the type of carbonyl group (C═O) present. This system gives rise to such names as aldotriose, aldotetrose, aldopentose, and aldohexose for the aldose forms and by such names as ketopentoses and ketohexoses when the ketone group is present. More recently the tendency is to indicate the presence of a ketone group by the ending "ulose" in names such as pentuloses and heptuloses. *See* MONOSACCHARIDE.

Oligosaccharides. These compound sugars are condensation products of two to five molecules of simple sugars and are subclassified into disaccharides, trisaccharides, tetrasaccharides, and pentasaccharides, according to the number of monosaccharide molecules yielded upon hydrolysis.

The sugars which include the monosaccharides and oligosaccharides are mostly crystalline compounds. The monosaccharides and disaccharides have a sweet taste. Products obtained from

hydrolysis of higher molecular weight polysaccharides and consisting of compound sugars that contain as many as nine monosaccharide units are also termed oligosaccharides. However, these higher members are not crystalline compounds. The disaccharides and other groups of oligosaccharides are subclassified into reducing and nonreducing sugars, depending on whether or not the sugar has a functional carbonyl group. In a nonreducing oligosaccharide, the carbonyl groups of the constituent monosaccharide units are involved in glycosidic linkage. The compound sugars are in reality sugarlike polysaccharides, but the prefix "poly" is reserved for carbohydrates which consist of large aggregates of monosaccharide units. *See* OLIGOSACCHARIDE.

Polysaccharides. These comprise a heterogeneous group of compounds which represent large aggregates of monosaccharide units, joined through glycosidic bonds. They are tasteless, nonreducing, amorphous substances that yield a large and indefinite number of monosaccharide units on hydrolysis. Their molecular weight is usually very high, and many of them, like starch or glycogen, have molecular weights of several million. They form colloidal solutions, but some polysaccharides, of which cellulose is an example, are completely insoluble in water. On account of their heterogeneity they are difficult to classify. One common system of classification is based primarily upon the class of monosaccharide yielded on hydrolysis. Thus a polysaccharide that yields hexose monosaccharides on hydrolysis is called hexoglycan, $(C_6H_{10}O_5)_n$, where n represents the number of monosaccharide units present. A polysaccharide yielding a pentose sugar is called pentoglycan, $(C_5H_8O_4)_n$. Each major class is further subdivided according to the particular hexose or pentose produced. Thus, a polysaccharide composed of the hexose sugar D-glucose is called glucan; those composed of the pentose sugars D-xylose and L-arabinose are called xylan and araban, respectively.

Some polysaccharides, however, yield both hexose and pentose sugars on hydrolysis and are classified as mixed hexosans and pentosans; others, in addition to hexoses and pentoses, frequently on acid hydrolysis yield the uronic acids, D-glucuronic, D-galacturonic, or D-mannuronic acid. Such polysaccharides are often called polyuronides. *See* POLYSACCHARIDE.

Reducing and nonreducing sugars. The sugars are also classified into two general groups, the reducing and nonreducing. The reducing sugars are distinguished by the fact that because of their free, or potentially free, aldehyde or ketone groups they possess the property of readily reducing alkaline solutions of many metallic salts, such as those of copper, silver, bismuth, mercury, and iron. The most widely used reagent for this purpose is Fehling's solution, in which the oxidizing agent is the cupric tartrate ion, formed in a strongly alkaline solution of copper sulfate and a salt of tartaric acid. The nonreducing sugars do not exhibit this property. The reducing sugars constitute by far the larger group. The monosaccharides and many of their derivatives reduce Fehling's solution. Most of the disaccharides, including maltose, lactose, and the rarer sugars cellobiose, gentiobiose, melibiose, and turanose, are also reducing sugars. The best-known nonreducing sugar is the disaccharide sucrose. Among other nonreducing sugars are the disaccharide trehalose, the trisaccharides raffinose and melezitose, the tetrasaccharide stachyose, and the pentasaccharide verbascose.

The alkali in the Fehling solution, or other such reagents used for the determination of reducing sugars, causes considerable decomposition of the sugar molecule into reactive fragments which may also reduce the metal ions. Thus, while the total reduction for a given sugar may be constant under carefully controlled conditions and can therefore be used for quantitative purposes, it is impossible to write a balanced equation for the reaction in terms of the simple oxidation of the sugar and reduction of the metal ion.

Analysis. This involves separation of the carbohydrate mixtures, identification of the individual carbohydrates, and estimation of their quantities.

Separation and identification. Sugars, in most cases, occur in living tissues as mixtures. Before a sugar can be identified, it must be isolated in pure form, preferably as a crystalline substance. The pure crystalline compound is readily identified by its melting point, optical rotation, and x-ray diffraction pattern. Preparation of certain crystalline derivatives facilitates isolation and identification of reducing sugars. The phenylhydrazone and phenylosazone derivatives of these sugars are especially useful, chiefly because of their ease of preparation. The phenylosatriazoles, benzimidazoles, and diothioacetal acetates are more advantageous for the purpose of identification.

Application of chromatographic technique provides a useful and rapid method for the separation and identification of sugars. The great advantage of this process lies in the fact that it can be applied to minute amounts as well as to relatively large quantities. For micro amounts chromatography on filter paper is used. The mixed sugar solution is applied as a spot on the narrow edge of a strip of filter paper. The sugars are then separated by using a mixture of water and a partially immiscible organic solvent, such as *n*-butanol or phenol. The paper is then dried and sprayed with coloring reagents, such as *p*-anisidine hydrochloride or aniline hydrogen phthalate, which show the location of the individual sugars. For separation of larger quantities of material, column chromatography is used. In this method the proper choice of solvents and absorbents is important. The solvent is passed through a column containing powdered cellulose, silicic acid, alumina, clay, carbon, or other adsorbent by gravity, suction, or pressure. After sufficient development to form the different zones, the products may be eluted from the column in successive batches. *See* CHROMATOGRAPHY.

Color tests. There are a number of color tests which are helpful in the identification of carbohydrates.

These tests are based on the condensation of various aromatic amines or phenolic substances with the degradation products of sugars. The Molisch test is used for the general detection of carbohydrates. In this test a purple color is produced when the solution containing carbohydrate is treated with strong sulfuric acid in the presence of α-naphthol. In the Tauber test pentose sugars produce a cherry-red color when heated with a solution of benzidine in glacial acetic acid. Another test for pentoses and uronic acids is based on the fact that a violet-red color develops when these sugars are treated with hydrochloric acid and phloroglucinol. Seliwanoff's test is used for ketoses, which give a red color with resorcinol in hydrochloric acid. The Tollens naphthoresorcinol and the Dische carbazole tests are used for the detection of the uronic acids, D-glucuronic acid, D-galacturonic acid, and D-mannuronic acid. The polysaccharide, starch, produces a blue color when treated with a dilute solution of iodine in potassium iodide, while glycogen gives a reddish-brown color with this reagent.

Paper chromatography and color tests only indicate the probable presence of a particular carbohydrate. For unequivocal proof it is necessary to resort to isolation procedures.

Estimation. The reducing properties of aldoses and ketoses are most frequently utilized for the quantitative analysis of these sugars. A nonreducing sugar such as sucrose or raffinose, or a polysaccharide such as starch, must be first hydrolyzed with acid or with an appropriate enzyme to its constituent reducing monosaccharide units before it is analyzed. Upon heating the reducing sugars with Fehling's solution, containing copper sulfate, tartrate ion, and sodium hydroxide, a brick-red color develops, a result of the formation and precipitation of cuprous oxide.

The quantity of the cuprous oxide precipitate is a measure of the amount of reducing sugar present. Several modifications of this reaction have been used for the quantitative determination of reducing sugars. The estimation of sugar by the Benedict method involves the determination of the reduced cuprous ions colorimetrically after treatment with phosphomolybdic acid. In the Somogyi-Shaffer-Hartmann method, the sugar is estimated by iodometric titration of reduced copper. Several other quantitative methods for the determination of sugars are based on the reduction of ferricyanide ions in an alkaline solution. These reactions can be applied to micro quantities of sugars. There are also spectrophotometric methods, devised by Dische, which are based on light absorption by the reaction products resulting from treatment of sugars with carbazole or cysteine-carbazole in sulfuric acid. Sugars can be accurately identified and determined by their specific optical rotations, provided they are available in sufficient quantities and no other interfering optically active material is present. *See* ALDEHYDE; TITRATION.

Stereoisomerism. The sugars consist of chains of carbon atoms which are united to one another at a tetrahedral angle of 109°28′. A carbon atom to which are attached four different groups is called asymmetric. A sugar, or any other compound containing one or more asymmetric carbon atoms, possesses optical activity; that is, it rotates the plane of polarized light to the right or left. The specific rotation of a substance possessing optical activity is the rotation expressed in angular degrees which is afforded by 1 gram of substance dissolved in 1 milliliter of water in a tube 1 decimeter in length. It is usually given for the sodium D line at a definite temperature, for example, 20°C (68°F), and is designated by $[\alpha]_D^{20}$. *See* OPTICAL ACTIVITY.

The triose glycerose, or glyceraldehyde, has one asymmetric carbon atom and therefore exists in two optically active forms, one dextrorotatory and the other levorotatory (**Fig. 1**).

The D and L forms of glycerose are related to each other as the right hand is to the left hand, being similar, but not identical. One may not be superposed upon the other. If one model is held before a mirror, the image in the mirror corresponds to the arrangement of the other model. The two compounds whose molecules are mirror images of each other are called optical antipodes or enantiomorphs.

It is difficult to represent stereochemically the more complex sugars having several asymmetric carbon atoms. An examination of the formula for an aldohexose (**Fig. 2**) reveals that it contains four different asymmetric carbon atoms; that is, each of the atoms marked with an asterisk, at positions 2, 3, 4,

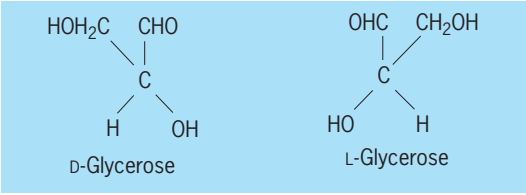

Fig. 1. Spatial arrangement of the groups around the asymmetric carbons of the two isomers of glycerose.

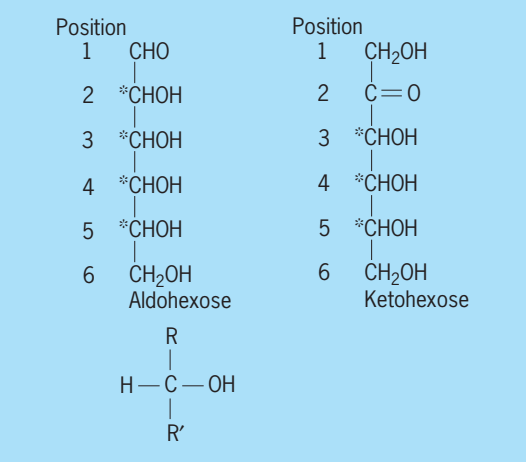

Fig. 2. Structural formulas showing asymmetric carbons (asterisk) for an aldohexose and a ketohexose. Detail of carbon at position 2 in aldohexose is shown, where R = CHO and R′ = everything below carbon atom 2.

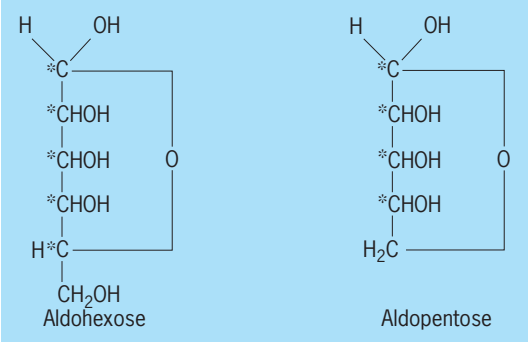

Fig. 3. Structural formulas for ring forms of an aldohexose and an aldopentose. Asymmetric carbons are marked by an asterisk.

and 5, carries four different groups or atoms. The ketohexose contains three asymmetric carbon atoms, at 3, 4, and 5 (Fig. 2).

As the number of asymmetric carbon atoms in the sugar molecule increases, from the trioses to the higher monosaccharides, the number of stereoisomers increases in accordance with the van't Hoff formula 2^n where n represents the number of asymmetric carbon atoms.

Thus the possible number of stereoisomers of an aldohexose, such as glucose or galactose, when written in open-chain form, is 2^4 or 16 (eight enantiomorphous pairs). In the ketose sugars, there is one less asymmetric center than in the aldoses of equal chain length; therefore the number of isomers is reduced by a factor of 2. The possible number of stereoisomers of a ketohexose, exemplified by fructose or sorbose, is 2^3 or 8 (four enantiomorphous pairs). Eight stereoisomers, consisting of four pairs of enantiomorphs, are possible for an aldopentose like xylose or arabinose. If, as seen in **Fig. 3**, the formulas of these sugars are represented in ring form, the first carbon atom in the monosaccharide chain becomes asymmetric because of formation of a cyclic hemiacetal with a new OH group and another center of asymmetry at carbon atom 1. The possible isomers for a cyclic aldohexose therefore become 2^5 or 32, and for an aldopentose 2^4 or 16.

The spatial arrangement of the groups around the asymmetric carbon atom of the dextrorotatory form of glyceraldehyde (glycerose) is arbitrarily called the D configuration, while that of the levorotatory form is called the L configuration. This connotation of the D and L refers to spatial configuration only and is not an indication of the direction of rotation of the plane of polarized light by the sugar. The reference carbon atom regarding D or L configuration is, for sugars containing more than one asymmetric carbon atom, the asymmetric carbon atom farthest removed from the active (that is, the aldehyde or ketone) end of the molecule and adjacent to the terminal group (**Fig. 4**).

Rosanoff, who introduced this criterion for designating the D and L form of sugars, regarded all sugars of the D series as built up, in theory, from D-glycerose, and those of the L series from L-glycerose. Although not accomplished in practice throughout the series, such a process is rendered possible without interfering with the original asymmetric carbon atom by the stepwise addition of the repeating unit CHOH through the Kiliani cyanohydrin synthesis. It must be emphasized that Rosanoff's representation of the D and L forms of sugar is merely a convention and has no relation to their direction of rotation. If it is desired to indicate the direction of rotation of the compound, the symbols (+), meaning dextrorotatory, and (−), meaning levorotatory, are used. Thus a dextrorotatory sugar with the D configuration is indicated by the symbol D(+), while a levorotatory sugar of the same configuration is represented by D(−). **Figure 5** shows the relationship of D-glycerose to the D-trioses, D-tetroses, D-pentoses, and D-hexoses.

Formulation. Thus far the monosaccharides have been considered as open-chain compounds, designated by projection formulas. The projection formula, introduced by Emil Fischer, is based on the convention that tetrahedrons, representing individual asymmetric carbons of a sugar, viewed as though supported by their carbon to carbon bonds with the hydrogen and hydroxyl groups extending outward, are projected upon the depicting plane. This formula is quite convenient, because the relationship between various sugars is easily demonstrated. Its numbering begins at the terminal position associated with the carbonyl group; in ketoses the carbonyl group is in carbon atom 2 position. When Greek letters are used, as in the older literature, the carbon next to the CHO group is called the α-carbon atom.

Although an aldohexose such as D-glucose or an aldopentose such as D-xylose exhibits many of the properties of common aldehydes, its reactivity, when compared with acyclic hydroxyaldehydes such as glycolic aldehyde or glyceraldehyde, is not as great as would be expected if the sugar possessed the simple aldehydic projection formula (**Fig. 6**). Furthermore, the cyanohydrin reaction proceeds with difficulty, and sugars fail to give the Schiff test for aldehydes. Thus, the straight-chain aldehydic projection formula does not represent the essential features of the sugar molecules, particularly their tendency to form five- and six-membered rings. In reality, because of the tetrahedral angles between bonds of the carbon atoms, the chain tends to form a ring. Thus, the ends of a six-carbon chain approach each other as shown in **Fig. 7**, permitting a reaction between the carbonyl group and a hydroxyl on the fourth or fifth carbon atom and forming a hemiacetal.

The reaction is analogous to the formation of hemiacetals from simple aldehydes and alcohols. Admixture of an aldehyde and an alcohol produces first a

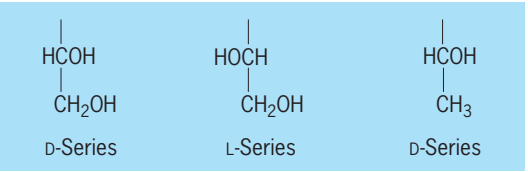

Fig. 4. Spatial configurations of D and L series showing reference carbon attached to the terminal group.

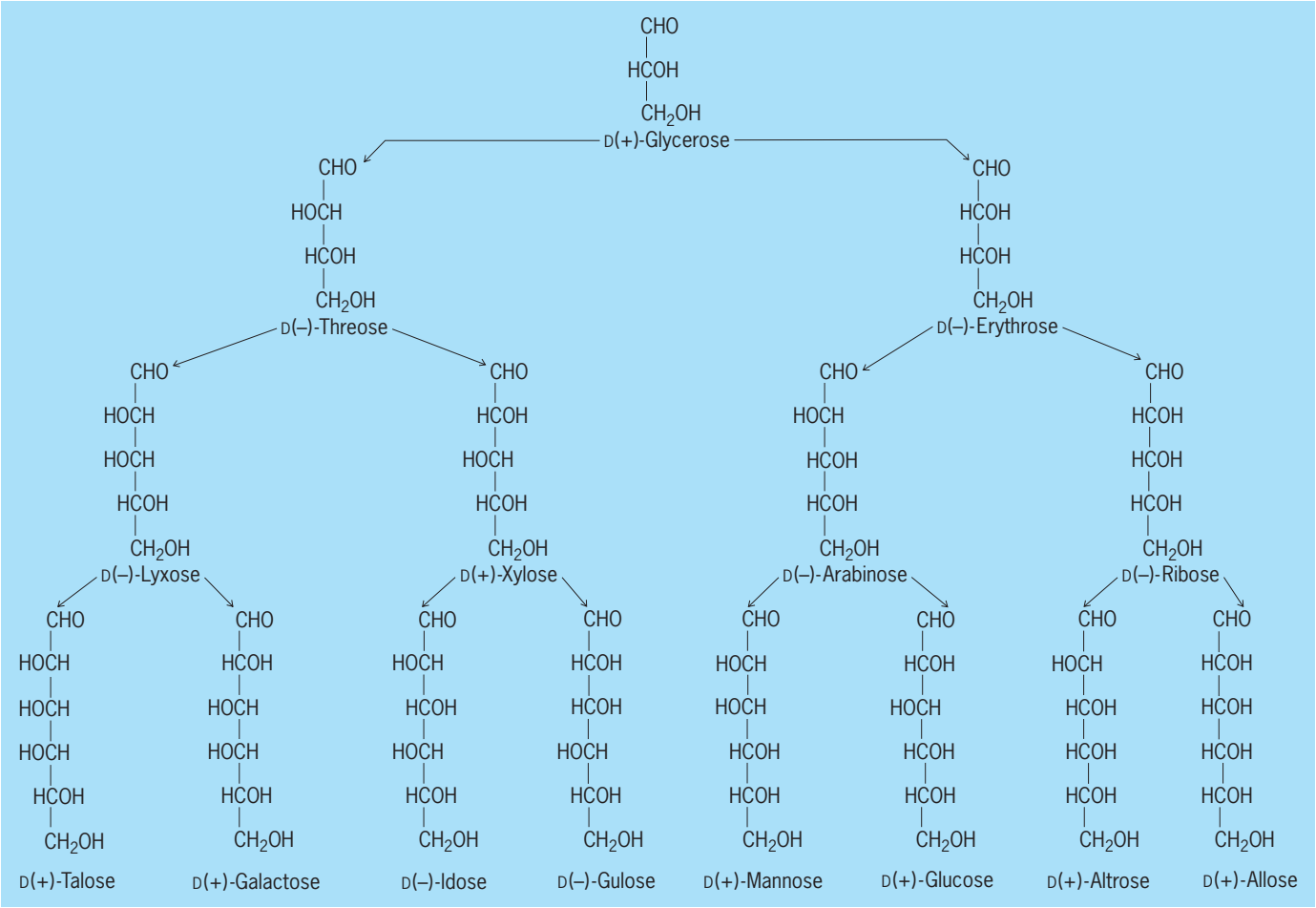

Fig. 5. The D-series of aldose sugars as related to D-glycerose.

hemiacetal. Long standing in an excess of alcohol in the presence of an acid catalyst converts the hemiacetal into a complete acetal, as shown below.

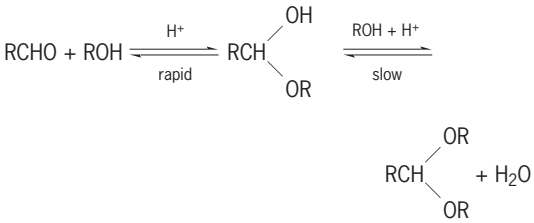

Such a reaction occurs not only with aldose but also with ketose sugars. Formation of a hemiacetal in

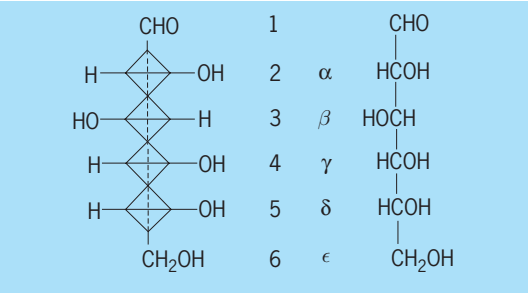

Fig. 6. Projection formula for the aldohexose D-glucose.

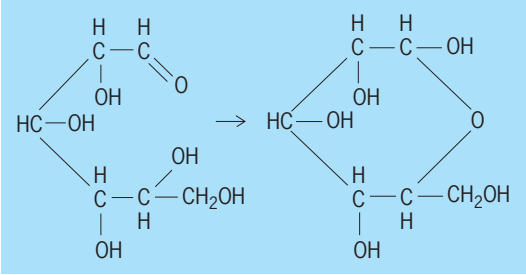

Fig. 7. Six-carbon chain forming hemiacetal ring.

a hexose monosaccharide from a reaction between carbon atom 1 and the OH group of carbon atom 6 is not likely, because this would produce a strained seven-membered ring. Six- and five-membered rings are more easily formed because there is very little distortion of the bond angles. Rings of four carbon atoms or less do not form easily.

There is abundant evidence that sugars actually exist as cyclic compounds. Most decisive proof is furnished by the phenomenon of mutarotation, by the existence of two isomeric methylglycosides, and by Haworth's methylation experiments, in which he was unable to demonstrate open-chain structure.

In the older literature the six-membered ring structure of the aldohexoses in which the first carbon

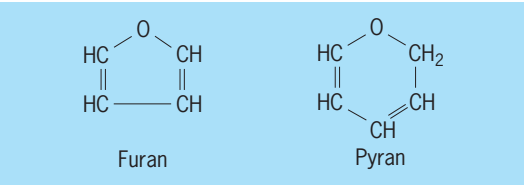

Fig. 8. Structural formulas for furan and pyran.

atom is linked to the fifth carbon atom through an oxygen atom had been named amylene oxide or δ-oxide configuration. The five-membered rings in which carbon atom 1 is linked to carbon atom 4 had been termed butylene oxide or γ-oxide configuration. In 1927 W. N. Haworth suggested that the five-atom ring sugars should be regarded as derived from furan and the six-atom ring configuration from pyran (**Fig. 8**).

Thus a sugar containing a γ-oxide or 1,4-oxide ring is called a furanose sugar, and if it is present as a glycoside, it is described as a furanoside. It must be noted, however, that unsubstituted furanose sugars have never been isolated, although they do exist in small proportion in solution. Although the 1,4 ring is stable and the 1,5 ring is relatively unstable in the sugar lactones, the reverse is true for the sugars and glycosides. Furanose structures are therefore found either in substituted sugars such as 2,3,5,6-*O*-methyl-D-glucose, where there is no free hydroxyl on carbon atom 5, or more commonly, when the monosaccharide (D-fructose or L-arabinose) is a constituent of complex sugars, the oligosaccharides or polysaccharides. Similarly the names pyranose and pyranoside are applied to the sugars and sugar derivatives having the six-membered, δ-oxide, or 1,5 rings. *See* GLYCOSIDE.

In the Haworth perspective formulas, carbon atoms 1–5 and the ring oxygen are represented in a single plane projecting from the plane of the paper on which it is written. The valences of the carbon atoms not involved in ring formation are situated above or below the plane of the ring. The thickened lines at the bottom of the ring formulas represent the sides of the hexagon nearest to the observer.

When the carbon atoms in the ring are numbered clockwise, as in α-D-glycopyranose, the groups that are written on the right-hand side in the projection formula are represented as projecting below the plane of the ring, and the corresponding groups on the left-hand side are projecting above. However, there is a discrepancy in the relative positions of the substituent groups on carbon atom 5 in that the H

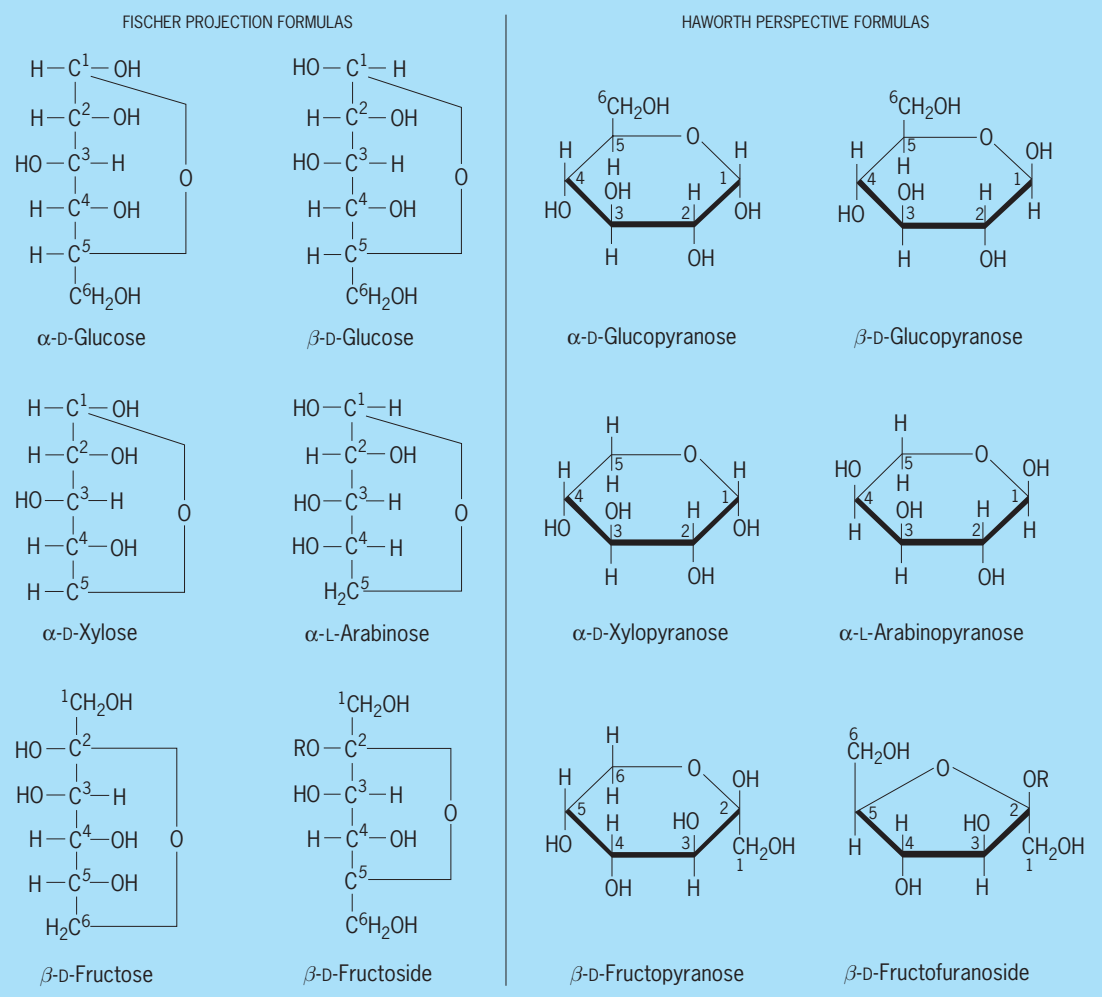

Fig. 9. Fischer projection formulas and Haworth perspective formulas.

atom appears in the perspective formula below instead of above the plane of the ring. This apparent anomaly is the result of torsion required to effect ring closure. In forming a ring from the straight-chain aldehydic form, the fifth carbon atom must be rotated so that the oxygen atom in the OH group on this carbon atom is brought into the plane of the first five carbon atoms. Consequently, the H atom attached to carbon atom 5 is now shifted to the other side of the chain, because this carbon atom has been rotated through more than a 90° angle.

The projection and perspective formulas for the α- and β-isomers of D-glucose and for α-D-xylose, α-L-arabinose, and β-D-fructose are shown in **Fig. 9**.

Since free D-fructose primarily exists in the pyranose configuration, the furanose ring structure of this sugar is shown with the active hydrogen of the potential reducing group substituted. The substituent R may be a D-glucopyranosyl unit as in the disaccharide, sucrose, or a chain of other fructofuranosyl units as in the polysaccharide, inulin.

Mutarotation. Two forms of glucose are known. The α-form is obtained by dissolving the sugar, D-glucose, in water and allowing it to crystallize through evaporation of the solvent. The β form is obtained by crystallizing D-glucose from pyridine or acetic acid. The products are not merely different crystallographic forms of the same substance, because they have different properties when dissolved. At 20°C (68°F) a freshly prepared solution of α-D-glucose has a specific rotation, $[\alpha]_D^{20}$, of +113°, and a freshly prepared aqueous solution of β-D-glucose at the same temperature has a specific rotation of +19°. Upon standing, both solutions gradually change in rotatory power. The optical rotation of the α-D form decreases and that of the β-D form increases until they reach the same equilibrium value of $[\alpha]_D + 52.5°$, which corresponds to a mixture of 37% α and 63% β. This phenomenon of change in its optical property is called mutarotation. It is displayed not only with glucose but with all crystalline sugars that have a free or potentially free aldehyde or ketone group in the molecule. To account for the existence of two structures, it is assumed that the monosaccharides exist in the form of cyclic hemiacetals rather than in open-chain form (**Fig. 10**).

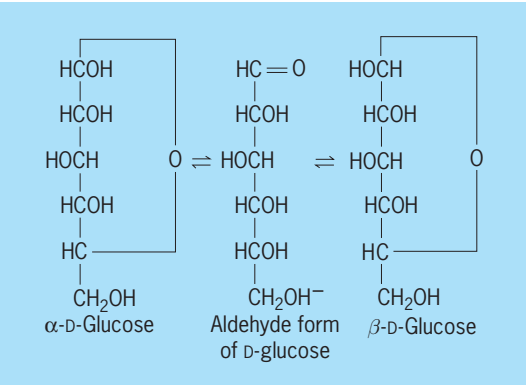

Fig. 10. Possibilities during mutarotation of D-glucose.

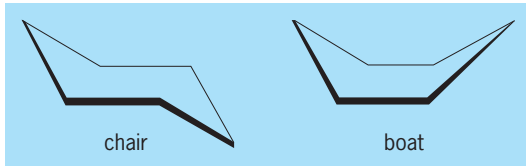

Fig. 11. Conformations of a cyclohexane.

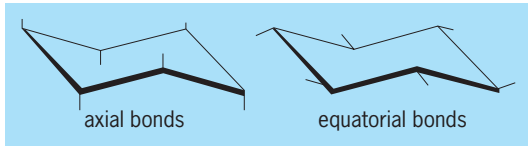

Fig. 12. Geometrical arrangements of carbon-hydrogen bonds on the chair form of a cyclohexane.

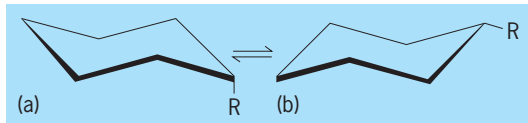

Fig. 13. Interconversion of (*a*) axial and (*b*) equatorial positions for R group of a cyclohexane derivative.

The process of mutarotation is due to the labile nature of the hemiacetal linkage and the interconversion of the α and β forms through the acyclic aldehyde form which exists in solution to a very small extent (about 0.1%). Mutarotation is catalyzed by both acid and base, with the base being by far more effective.

The α- and β-sugars are stereoisomeric, but they are not mirror images. They are known as anomers, differing from each other only in the positions of the atoms or groups attached to the terminal asymmetric carbon atom.

To avoid confusion in the matter of new isomers or anomers, a convention was proposed by C. Hudson which is generally accepted. This states that for pairs of α and β isomers of the D series, the one with the higher rotation in the positive sense shall be called the α form, whereas the isomer with the lower rotation is designated as β. The reverse holds good for the L series, the enantiomorph of α-D-glucose being α-L-glucose with a lower specific rotation of $[\alpha]_D -113°$ than β-L-glucose, $[\alpha]_D - 19°$. In the α-anomer of D-glucose the OH group that is attached to the aldehydic carbon atom and the OH group that is on the adjacent carbon atom are on the same side of the ring (cis position). In the β form these hydroxyl groups are on opposite sides of the hemiacetal ring (trans position).

Sugar conformations. Although the Haworth formulas are an improvement over the Fischer formulas, they still represent an oversimplification of the true configuration of the monosaccharides in space. While the five-membered furanose ring of a sugar is almost planar, a strainless six-membered pyranose ring is not. A nonplanar arrangement of the molecule must be adopted to maintain the normal valency angles in the pyranose ring. The term "conformations" was coined for the various arrangements of

atoms in space that can arise by rotation about single bonds.

Since the only difference between a pyranose ring of a monosaccharide and that of a cyclohexane is the presence of an oxygen in the ring of the former compound, there is a close similarity between the molecular geometries of the two types of rings. The concepts developed for the cyclohexane ring can, therefore, be applied to the sugars containing the pyranose rings.

Results obtained by several investigators show that arrangements of cyclohexane which are free from angle strain are either of the "chair" conformation or the "boat" conformation (**Fig. 11**).

Examination of the chair form of cyclohexane shows that its carbon-hydrogen bonds can be divided into two geometrically different types. Six of them are perpendicular to the plane of the ring and are termed axial (a) bonds, while the other six radiate out from the ring and are termed equatorial (e) bonds (**Fig. 12**). Because of the staggered arrangements of the hydrogen atoms, the chair conformation (which minimizes repulsions between them) has a lower energy content and is thus more stable than the boat form. Consequently, cyclohexane and most of its derivatives tend to assume the chair conformation.

In cyclohexane derivatives, the group R is axial in one conformation but is readily convertible into the equatorial position (**Fig. 13**).

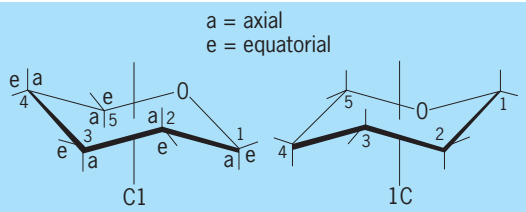

Fig. 14. Interchangeable conformations of a pyranose sugar, designated as C1 and 1C.

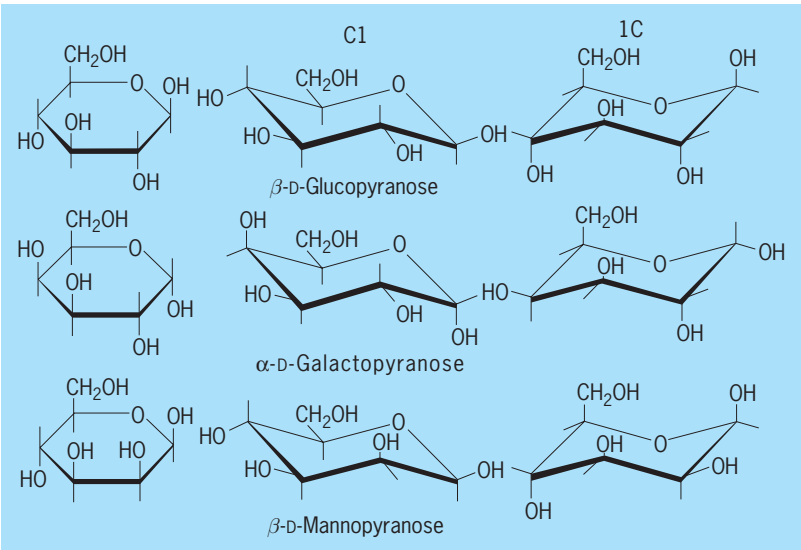

Fig. 15. Relationship between Haworth formula and chair forms for pyranose sugars.

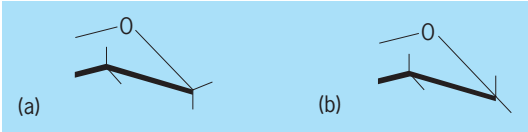

Fig. 16. Glycosidic carbon links in the C1 conformation must be as in *a*, not *b*.

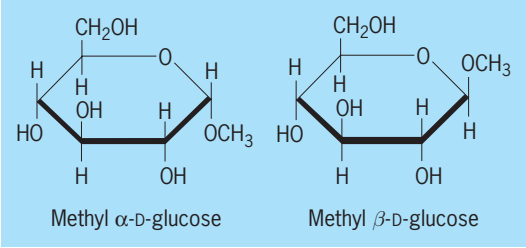

Fig. 17. Formulas for α and β forms of methyl D-glucose.

Similarly, pyranose sugars can exist in two interchangeable conformations, designated by R. E. Reeves as C1 and 1C (**Fig. 14**).

The greatest nonbonded interactions in a sugar are due to the hydroxyl groups. In considering preferred conformational stability for any sugar, conformations with large axial substituents should be avoided. In the absence of large polar interactions in the pyranose series, that chair conformation in which the largest number of the bulky groups are in equatorial position is the most stable one. In this connection it should be noted that the CH_2OH group is bulkier than the hydroxyl.

The relationship between the Haworth formulas and the two chair forms for the pyranose sugars is shown in **Fig. 15**. In each of the pyranose sugars in the illustration there are fewer axial hydroxyl groups and fewer repulsive interactions in the C1 conformation, which is the preferred form.

In conformational formulas, direction of bonds is determined by the tetrahedral valencies of the carbon atoms (**Fig. 16**).

The application of conformational analysis in carbohydrate chemistry is useful in many cases in predicting relative rates of reactions and the extent to which some reactions will proceed. As an example, the greater proportion of β-D-glucose (in which all hydroxyls are equatorial) in the equilibrium mixture compared with α-D-glucose (in which all hydroxyls are equatorial) in the equilibrium mixture compared with α-D-glucose (which has an axial hydroxyl at carbon atom 1) can be explained. Similarly, the equilibrium between α-D-galactose 1-phosphate (axial hydroxyl at carbon atom 4 in C1 conformation) and α-D-glucose 1-phosphate, which has an equatorial hydroxyl at C = 4, resulting in smaller nonbonded interactions, favors the glucose derivative.

Methyl glucosides. When a solution of glucose in cold methanol is treated with dry hydrogen chloride as catalyst, an acetal is formed. Such a compound formed from glucose is called a glucoside (**Fig. 17**); if the sugar galactose were used, the acetal would be called a galactoside. Methyl glucoside

is perhaps the simplest example of a large group of substances in which the hydroxyl of the hemiacetal group at carbon atom 1 of an aldose monosaccharide has been condensed with the hydroxyl group of an alcohol.

In aqueous solution D-glucose exists chiefly as D-glucopyranose. However, a small proportion of this sugar exists in the furanose as well as in the open-chain aldehyde forms, the last in a minute amount. Since the furanose structure reacts more rapidly to form the glucoside than does the pyranose configuration of the sugar, a considerable amount of labile methylglucofuranoside is formed in the first few hours. On prolonged standing or heating, the more stable methyl α-D-glucopyranoside and methyl β-D-glucopyranoside appear. Thus it is evident that when glucose is treated with methanolic hydrochloric acid, the yield of methylglucoside of a particular configuration has little bearing on the structure of the original sugar; that is, the structure of the glucoside may or may not correspond to the structure of the crystalline sugar. However, a relationship in this case can be established through hydrolysis of a methylglucoside. When pure methyl α-D-glucoside is hydrolyzed, the subsequent mutarotation of the hydrolysis product is found to be in the same direction as that of pure α-D-glucopyranose. It is therefore concluded that α-D-glucose and α-D-glucoside have the same configuration about the first carbon atom.

The α- and β-methylglucosides are crystalline compounds. They may be separated by fractional crystallization or through the agency of enzymes. The enzymes maltase and emulsin act selectively upon the two glucosides, the former hydrolyzing the α form only and the latter acting upon the anomeric β compound. The specific rotation of methyl α-D-glucopyranoside is $+159^\circ$, while that of methyl β-D-glucopyranoside is -33°. *See* BIOCHEMISTRY.

William Z. Hassid

Bibliography. G. D. Aspinall, *Carbohydrate Chemistry*, 1976; V. Ginsburg, *Biology of Carbohydrates*, 1981; W. W. Pigman, *The Carbohydrates: Chemistry and Biochemistry*, 2d ed., 1981; J. Staněk et al., *The Monosaccharides*, 1964; J. Staněk, M. Černý, and J. Pacák, *The Oligosaccharides*, 1965.

Carbohydrate metabolism

The fields of biochemistry and physiology deal with the breakdown and synthesis of simple sugars, oligosaccharides, and polysaccharides, and with the transport of sugars across cell membranes and tissues. The breakdown or dissimilation of simple sugars, particularly glucose, is one of the principal sources of energy for living organisms. The dissimilation may be anaerobic, as in fermentations, or aerobic, that is, respiratory. In both types of metabolism, the breakdown is accompanied by the formation of energy-rich bonds, chiefly the pyrophosphate bond of the coenzyme adenosine triphosphate (ATP), which serves as a coupling agent between different metabolic processes. In higher animals, glucose is the carbohydrate constituent of blood, which carries it to the tissues of the body. In higher plants, the disaccharide sucrose is often stored and transported by the tissues. Certain polysaccharides, especially starch and glycogen, are stored as endogenous food reserves in the cells of plants, animals, and microorganisms. Others, such as cellulose, chitin, and bacterial polysaccharides, serve as structural components of cell walls. As constituents of plant and animal tissues, various carbohydrates become available to those organisms which depend on other living or dead organisms for their source of nutrients. Hence, all naturally occurring carbohydrates can be dissimilated by some animals or microorganisms. *See* ADENOSINE TRIPHOSPHATE (ATP); CARBOHYDRATE; CELLULOSE; GLYCOGEN; STARCH.

Dietary carbohydrates. Certain carbohydrates cannot be used as nutrients by humans. For example, the polysaccharide cellulose, which is one of the main constituents of plants, cannot be digested by humans or other mammals and is a useful food only for those, such as the ruminants, that harbor cellulose-decomposing microorganisms in their digestive tracts. The principal dietary carbohydrates available to humans are the simple sugars glucose and fructose, the disaccharides sucrose and lactose, and the polysaccharides glycogen and starch. Lactose is the carbohydrate constituent of milk and hence one of the main sources of food during infancy. The disaccharides and polysaccharides that cannot be absorbed directly from the intestine are first digested and hydrolyzed by enzymes secreted into the alimentary canal. The nature of these enzymes, which are known as glycosidases, will be discussed later. *See* LACTOSE.

Intestinal absorption and transport. The simple sugars, of which glucose is the major component, reach the intestine or are produced there through the digestion of oligosaccharides. They are absorbed by the intestinal mucosa and transported across the tissue into the bloodstream. This process involves the accumulation of sugar against a concentration gradient and requires active metabolism of the mucosal tissue as a source of energy. The sugars are absorbed from the blood by the liver and are stored there as glycogen. The liver glycogen serves as a constant source of glucose in the bloodstream. The mechanisms of transport of sugars across cell membranes and tissues are not yet understood, but they appear to be highly specific for different sugars and to depend on enzymelike components of the cells.

Dissimilation of simple sugars. The degradation of monosaccharides may follow one of several types of metabolic pathways. In the phosphorylative pathways, the sugar is first converted to a phosphate ester (phosphorylated) in a reaction with ATP. The phosphorylated sugar is then split into smaller units, either before or after oxidation. In the nonphosphorylative pathways, the sugar is usually oxidized to the corresponding aldonic acid. This may subsequently be broken down either with or without phosphorylation of the intermediate products. Among the

principal intermediates in carbohydrate metabolism are glyceraldehyde-3-phosphate and pyruvic acid. The end products of metabolism depend on the organism and, to some extent, on the environmental conditions. Besides cell material that is formed through various biosynthetic reactions, the products may include carbon dioxide (CO_2), alcohols, organic acids, and hydrogen gas. In the so-called complete oxidations, CO_2 is the only excreted end product. In incomplete oxidations, characteristic of the vinegar bacteria and of certain fungi, oxidized end products such as gluconic, ketogluconic, citric, or fumaric acids may accumulate. Organic end products are invariably found in fermentations, since fermentative metabolism requires the reduction as well as the oxidation of some of the products of intermediary metabolism. The energy available from the complete oxidation of 1 mole of glucose with molecular oxygen is approximately 688,000 cal (2.98 megajoules), while that which can be derived from fermentation is very much less; for example, alcoholic fermentation yields about 58,000 cal (243,000 joules). In aerobic metabolism, a great deal of the available energy is harnessed for the needs of the organism as energy-rich bonds primarily in ATP through the process known as oxidative phosphorylation. In fermentation, a much smaller amount of ATP is produced. Hence, the amount of biosynthesis and mechanical work that an organism can do at the expense of a given amount of sugar is many times greater in respiration than in fermentation. *See* FERMENTATION.

Glycolysis. The principal phosphorylative pathway involved in fermentations is known as the glycolytic, hexose diphosphate, or Embden-Meyerhof pathway (see **illus.**). This sequence of reactions is the basis of the lactic acid fermentation of mammalian muscle and of the alcoholic fermentation of yeast. Its main features may be summarized as follows:

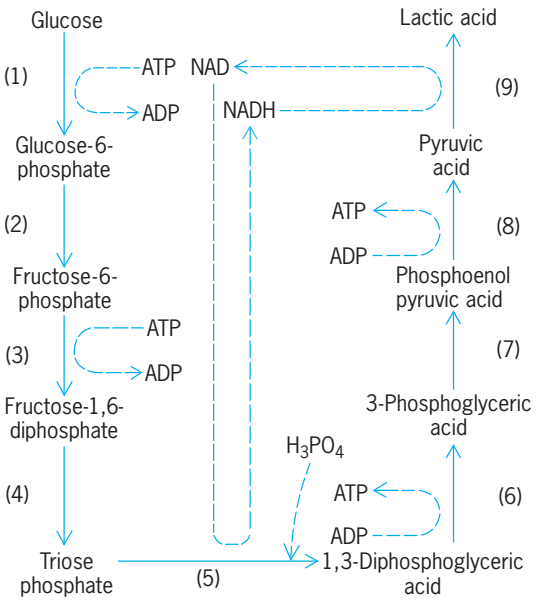

Glycolysis in lactic acid fermentation.

1. Glucose is phosphorylated to yield glucose-6-phosphate by the enzyme hexokinase with the conversion of a molecule of ATP to adenosine diphosphate (ADP).

2. Glucose-6-phosphate is epimerized to fructose-6-phosphate by phosphoglucoisomerase.

3. Fructose-6-phosphate is phosphorylated with ATP to yield fructose-1,6-diphosphate. This reaction is catalyzed by phosphofructokinase.

4. Fructose-1,6-diphosphate is split to two molecules of triose phosphate (dihydroxyacetone phosphate and glyceraldehyde phosphate) by aldolase.

5. Both molecules of triose phosphate, which are interconverted by triose phosphate isomerase, are oxidized to 1,3-diphosphoglyceric acid by triose phosphate dehydrogenase. Inorganic phosphate is esterified in this reaction, and the coenzyme nicotin amide adenine dinucleotide NAD becomes reduced to NADH.

6. 1,3-Diphosphoglyceric acid reacts with ADP to yield ATP and 3-phosphoglyceric acid in the presence of phosphoglycerate kinase.

7. 3-Phosphoglyceric acid is converted to 2-phosphoglyceric acid by phosphoglyceric acid mutase, and the product is dehydrated to phosphoenolypyruvic acid by enolase.

8. In a reaction catalyzed by pyruvic kinase, phosphoenolpyruvic acid reacts with ADP to give pyruvic acid and ATP.

9. Pyruvic acid is converted to the various end products of metabolism. In the lactic acid fermentation, it is reduced to lactic acid with the simultaneous reoxidation of NADH, which had been formed in step 5, to NAD. In alcoholic fermentation, pyruvic acid is decarboxylated to yield CO_2 and acetaldehyde, and the latter compound is reduced to alcohol with NADH.

For every molecule of glucose fermented through the glycolytic sequence, two molecules of ATP are used for phosphorylation (steps 1 and 3), while four are produced (steps 6 and 8). Thus, fermentation results in the net gain of two energy-rich phosphate bonds as ATP at the expense of inorganic phosphate esterified in step 5. The excess ATP is converted back to ADP and inorganic phosphate through coupled reactions useful to the organism, such as the mechanical work done by the contraction of muscle or biosynthetic reactions associated with growth. *See* ADENOSINE DIPHOSPHATE (ADP); NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD).

Oxidative mechanisms. The oxidative or respiratory metabolism of sugars differs in several respects from fermentative dissimilation. First, the oxidative steps, that is, the reoxidation of NADH, are linked to the reduction of molecular oxygen. Second, the pyruvic acid produced through glycolytic or other mechanisms is further oxidized, usually to CO_2 and H_2O. Third, in most aerobic organisms, alternative pathways either supplement or completely replace the glycolytic sequence of reactions for the oxidation of sugars. Where pyruvic acid appears as a metabolic intermediate, it is generally oxidatively

decarboxylated to yield CO_2 and the two-carbon acetyl fragment which combines with coenzyme A. The acetyl group is then further oxidized via the Krebs cycle. *See* CITRIC ACID CYCLE.

The principal alternative pathways by which sugars are dissimilated involve the oxidation of glucose-6-phosphate to the lactone of 6-phosphogluconic acid and are known as the hexose monophosphate pathways. In the best-known hexose monophosphate pathway, 6-phosphogluconic acid is formed from its lactone by hydrolysis and is oxidatively decarboxylated to yield CO_2 and the five-carbon sugar, ribulose-5-phosphate. Ribulose-5-phosphate is then converted to glucose-6-phosphate through a series of reactions in which phosphorylated sugars containing three, four, five, six, and seven carbon atoms are formed as intermediates. The principal enzymes involved in these transformations are transaldolase, transketolase, and epimerases. Since the glucose-6-phosphate that is produced from ribulose-5-phosphate is oxidized by the same sequence of reactions, the cyclic operation of this mechanism results in complete oxidation of glucose. Some of the enzymatic steps involved in this pathway provide the pentoses necessary for the synthesis of nucleotides and participate in the photosynthesis of green plants.

Simple sugars other than glucose. The metabolism of simple sugars other than glucose usually involves the conversion of the sugar to one of the intermediates of the phosphorylative pathways described for glucose metabolism. For example, fructose may be phosphorylated to fructose-6-phosphate, which can then be degraded via the glycolytic pathway or converted to glucose-6-phosphate and oxidized through the hexose monophosphate pathway. Similarly, mannose may be phosphorylated to mannose-6-phosphate, which is then transformed to fructose-6-phosphate by phosphomannose isomerase. Alternatively, in some bacteria, mannose is epimerized to fructose before phosphorylation. The principal mechanism for galactose metabolism requires the phosphorylation of the sugar to galactose-1-phosphate and the conversion of this ester to glucose-1-phosphate through the mediation of uridine diphosphoglucose. Glucose-1-phosphate is then converted to glucose-6-phosphate by phosphoglucomutase. The pentoses may be metabolized via their phosphate esters, which are converted to glucose-6-phosphate as described in the discussion of the hexose monophosphate pathway. *See* URIDINE DIPHOSPHOGLUCOSE (UDPG).

Disaccharides and polysaccharides. The dissimilation and biosynthesis of the oligosaccharides are effected through the enzymatic cleavage or formation of glycosidic bonds between simple monosaccharide constituents of the complex carbohydrates. The principal types of enzyme which split or synthesize glycosidic bonds are the hydrolases or glycosidases, phosphorylases, and transglycosylases. The enzymes are generally highly specific with respect to the glycosidic portion, or moiety, and the type of linkage of the substrates which they attack. For example, the glycosidic moiety may be glucosidic or galactosidic and the linkage may be α or β. The essential function of all three types of enzymes is the transfer of the glycosyl moiety of the substrate, which may be designated as the glycosyl donor, to an appropriate glycosyl acceptor, with the liberation of the aglycone. The different types of enzymes generally require different types of acceptors for the transfer reactions, but some enzymes can use a variety of acceptors and can act in more than one capacity. For example, an enzyme can act as a phosphorylase, transglycosylase, or hydrolase for the same substrate.

Hydrolytic cleavage is an essentially irreversible process and hence is used mostly in dissimilative metabolism. Phosphorylases and transglycosylases, on the other hand, may function in biosynthesis as well as the degradation of oligosaccharides.

The glycosidases or hydrolytic enzymes use preferentially the hydroxyl ion of water as the glycosyl acceptor. The aglycone is liberated after combining with the remaining hydrogen ion.

Lactases, for example, are β-galactosidases, and they combine the β-galactosyl portion of the disaccharide lactose with hydroxyl ion of water to yield galactose, liberating glucose. These enzymes can also hydrolyze methyl- or phenyl-β-galactosides with the formation of galactose and methanol or phenol, respectively.

Maltases are α-glucosidases which can hydrolyze maltose to two molecules of glucose.

Sucrases or invertases are enzymes which hydrolyze sucrose to yield glucose and fructose. The typical yeast invertase is a β-fructosidase, specific for the β-fructosyl portion of sucrose. The sucrase found in the human intestinal mucosa, on the other hand, is an α-glucosidase.

Amylases or diastases are enzymes which hydrolyze (digest) starch. The α-amylases, such as the human salivary and pancreatic amylases, split starch to smaller polysaccharide units, known as dextrins, and the disaccharide maltose. The β-amylase of malt, on the other hand, produces maltose as the main product of hydrolysis.

The phosphorylases catalyze the reversible phosphorolysis of certain disaccharides, polysaccharides, and nucleosides by transferring the glycosyl moieties to inorganic phosphate. The breakdown of glycogen and starch by the enzymes known as amylophosphorylases is an example of biologically important phosphorolytic reactions. α-D-Glycose-1-phosphate, which is the product of this process, can be metabolized after conversion to glucose-6-phosphate. Starch and glycogen can be synthesized from glucose-1-phosphate by the reverse reaction. The polysaccharide chain acts as either donor or acceptor for glucosyl groups, inorganic phosphate being the alternative acceptor. The reversible synthesis of sucrose, sucrose phosphate, and cellulose from uridine diphosphoglucose is similar to the phosphorolytic reactions, with the uridine compound serving as glucosyl donor.

The transglycosylases interconvert various disaccharides and polysaccharides by the transfer of

glycosyl groups. For instance, dextrins can be synthesized from maltose in a reversible reaction catalyzed by the bacterial enzyme, amylomaltase:

$$nC_{12}H_{22}O_{11} \rightleftharpoons (C_6H_{10}O_5)_{11} + nC_6H_{12}O_6$$

Maltose Dextrin Glucose

Many glucosidases have marked transglycosylase activity. *See* BIOCHEMISTRY. Michael Doudoroff

Carbon

A chemical element, C, with an atomic number of 6 and an atomic weight of 12.01115. Carbon is unique in chemistry because it forms a vast number of compounds, larger than the sum total of all other elements combined. By far the largest group of these compounds are those composed of carbon and hydrogen. It has been estimated that there are at least 1,000,000 known organic compounds, and this number is increasing rapidly each year. Although the classification is not rigorous, carbon forms another series of compounds, classified as inorganic, comprising a much smaller number than the organic compounds. *See* ORGANIC CHEMISTRY; PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Elemental carbon exists in two well-defined crystalline allotropic forms, diamond and graphite. Other forms, which are poorly developed in crystallinity, are charcoal, coke, and carbon black. Chemically pure carbon is prepared by the thermal decomposition of sugar (sucrose) in the absence of air. The physical and chemical properties of carbon are dependent on the crystal structure of the element. The density varies from 2.25 g/cm^3 (1.30 oz/in.^{3}) for graphite to 3.51 g/cm^3 (2.03 oz/ in.^{3}) for diamond. For graphite, the melting point is 3500°C (6332°F) and the extrapolated boiling point is 4830°C (8726°F). Elemental carbon is a fairly inert substance. It is insoluble in water, dilute acids and bases, and organic solvents. At elevated temperatures, it combines with oxygen to form carbon monoxide or carbon dioxide. With hot oxidizing agents, such as nitric acid and potassium nitrate, mellitic acid, $C_6(CO_2H)_6$, is obtained. Of the halogens, only fluorine reacts with elemental carbon. A number of metals combine with the element at elevated temperatures to form carbides.

Vaporizing a graphite target with an intense laser creates a nonequilibrium plasma of hot isolated carbon atoms and carbon dimers that, upon rapid cooling, coalesce into pure-carbon cage molecules. The most abundant and most stable of these molecules contain 60 carbon atoms in a highly spherical arrangement. This molecule, C_{60}, was named buckminsterfullerene, or buckyball for short, by its discoverers (in honor of R. Buckminster Fuller). This discovery initiated the field of fullerenes, the name given to this entire class of pure-carbon cage molecules. If one-half of the C_{60} molecule is removed and an extra belt of hexagons is inserted, the molecule will be extended. Successive additions of such belts of hexagons will lead to a tubular structure, a carbon nanotube. Carbon nanotubes were first experimentally observed in high-resolution transmission electron microscopy studies of fullerene by-products. The nonequilibrium synthesis methods previously described for fullerenes are also quite efficient at nanotube production. Depending on the synthesis conditions (such as type of catalyst used), nanotubes with different geometries are produced. Some are multiwalled, where many tubes are arranged coaxially and fit perfectly one inside the other; others are single walled. *See* CARBON NANOTUBES FULLERENE.

Carbon forms three gaseous compounds with oxygen: carbon monoxide, CO; carbon dioxide, CO_2; and carbon suboxide, C_3O_2. The first two oxides are the more important from an industrial standpoint. Carbon forms compounds with the halogens which have the general formula CX_4, where X is fluorine, chlorine, bromine, or iodine. At room temperature, carbon tetrafluoride is a gas, carbon tetrachloride is a liquid, and the other two compounds are solids. Mixed carbon tetrahalides are also known. Perhaps the most important of them is dichlorodifluoromethane, CCl_2F_2, commonly called Freon. *See* CARBON DIOXIDE; HALOGENATED HYDROCARBON.

Carbon and its compounds are found widely distributed in nature. It is estimated that carbon makes up 0.032% of the Earth's crust. Free carbon is found in large deposits as coal, an amorphous form of the element which contains additional complex carbon-hydrogen-nitrogen compounds. Pure crystalline carbon is found as graphite and as diamonds.

Extensive amounts of carbon are found in the form of its compounds. In the atmosphere, carbon is present in amounts of up to 0.03% by volume as carbon dioxide. Various minerals such as limestone, dolomite, marble, and chalk all contain carbon in the form of carbonate. All plant and animal life is composed of complex organic compounds containing carbon combined with hydrogen, oxygen, nitrogen, and other elements. The remains of past plant and animal life are found as deposits of petroleum, asphalt, and bitumen. Deposits of natural gas contain compounds that are composed of carbon and hydrogen. *See* CARBONATE MINERALS.

The free element has many uses, ranging from ornamental applications of the diamond in jewelry to the black-colored pigment of carbon black in

automobile tires and printing inks. Another form of carbon, graphite, is used for high-temperature crucibles, arc-light and dry-cell electrodes, lead pencils, and as a lubricant. Charcoal, an amorphous form of carbon, is used as an absorbent for gases and as a decolorizing agent. *See* CARBON BLACK; CHARCOAL; DIAMOND; GRAPHITE.

The compounds of carbon find many uses. Carbon dioxide is used for the carbonation of beverages, for fire extinguishers, and in the solid state as a refrigerant. Carbon monoxide finds use as a reducing agent for many metallurgical processes. Carbon tetrachloride and carbon disulfide are important solvents for industrial uses. Freon is used in refrigeration devices. Calcium carbide is used to prepare acetylene, which is used for the welding and cutting of metals as well as for the preparation of other organic compounds. Other metal carbides find important uses as refractories and metal cutters. E. Eugene Weaver

Bibliography. M. S. Dresselhaus et al., *Science of Fullerenes and Carbon Nanotubes: Their Properties and Applications*, 1996; *Kirk-Othmer Encyclopedia of Chemical Technology*, vol. 4, 4th ed., 1992; R. T. Morrison et al., *Organic Chemistry*, 6th ed, 1992; P. A. Thrower, *Chemistry and Physics of Carbon*, vol. 24, 1993.

Carbon black

An amorphous form of carbon produced commercially by thermal or oxidative decomposition of hydrocarbons, It is used principally in rubber goods, pigments, and printer's ink.

Because of the irregular arrangement of its carbon atoms, the physical properties of carbon black differ from the graphite and diamond forms of carbon. The surface area of blacks used for rubber reinforcement ranges 3100–46000 ft^2/oz (10–150 m^2/g) when determined by nitrogen adsorption, and the ultimate black particle has an average diameter of 20 to 300 nanometers. Pigment-grade blacks have surface areas of 150,000 ft^2/oz (300–500 m^2/g). Loose carbon black has a density of 2–3 lb/ft^3 (0.032–0.048 g/cm^3), but it is usually compressed for bag shipment or pelleted for bulk shipment in hopper cars.

Principal noncarbon components of carbon black are oxygen, hydrogen, and sulfur. Oxygen and hydrogen are combined chemically with carbon on the surface of the particle. The ratio of hydrogen to oxygen varies with the manufacturing process. In contact (channel) blacks, oxygen content is 2–5% while hydrogen amounts to about 0.5%; in furnace blacks the oxygen content of the black depends on that of the hydrocarbon feed.

Manufacturing processes may be classed as contact, furnace, or thermal. In the channel (contact) process, natural gas is burned with insufficient air for complete combustion. The smoky flame from individual burners impinges on a cool channel iron, and carbon black deposited on the channel is removed by a scraper blade. The yield of channel black is dependent upon the composition of the gas and the grade of black produced. Yield of black decreases as the heating value of the natural gas decreases. Of the three grades of black produced—easy (EPC), medium (MPC), or hard (HPC) processing channel—the yield of EPC is lowest and averages about 1.3 lb/1000 ft^3 (21 kg/1000 m^3) of natural gas.

In the furnace process, the hydrocarbon and air are fed into a reactor. Combustion of part of the hydrocarbon raises the temperature to 2000–3000°F (1100–1700°C), causing decomposition of the unburned hydrocarbon to carbon black. A water spray quickly cools the hot reaction products, and the finely divided black is recovered by cyclones and bag filters. Natural gas was the principal feed to the furnace process until 1943, when a furnace process utilizing heavy petroleum oils was introduced. The oil black produced was superior for reinforcing synthetic rubber, and oil furnace blacks now constitute the major type of carbon black produced. Yield of carbon black from the furnace process depends on feed stock used and type of black produced, ranging from 2 to 16 lb per 1000 ft^3 (32–260 kg/1000 m^3) of natural gas and 2.5 to 6 lb per gallon (0.3–0.7 kg/liter) of oil.

The oil furnace process can produce a variety of blacks specifically designed for use with various types of rubbers. For many years particle size was used as an approximate guide for black quality, the finer-particle blacks generally showing better abrasion resistance in performance tests. However, the physicochemical nature of the black surface, or "structure," became increasingly important. Structure affects the ease of compounding and modulus, stiffness, and extrusion properties of the rubber black mixture. Another development was the production of channel black substitutes from oil furnace blacks.

In the thermal process, natural gas is decomposed to carbon and hydrogen by heated refractories. Because this decomposition cools the refractories rapidly, it must be periodically interrupted to reheat the refractories. In one variant of this process employing acetylene as the feed, heat is liberated during the decomposition of acetylene. Acetylene black is distinguished by unusually high electric conductivity, and it is used in the manufacture of dry cells.

Over 90% of the production of carbon black in the United States is used by the rubber industry; hence, carbon black and rubber technology are interrelated. It is not an inert filler but enhances and reinforces various properties of rubber. Carbon black may be dispersed in rubber latex (masterbatch process) to produce, after coagulation, an intimate solid mixture of black and rubber, ready for fabrication of rubber goods.

Prior to 1940, channel black accounted for nearly 90% of domestic carbon black production; the last plant producing channel black closed in September 1976. This dramatic shift from contact to oil furnace processes was the result of the rapid price

increase of natural gas along with the superior performance of oil blacks in tire manufacture. *See* DIAMOND; GRAPHITE; RUBBER; TIRE. Carl J. Helmers

Carbon dioxide

A colorless, odorless, tasteless gas, formula CO_2, about 1.5 times as heavy as air. The specific volume at atmospheric pressure (100 kilopascals) and 70°F (21°C) is 8.74 ft^3/lb (0.546 m^3/kg). Under normal conditions, it is stable, inert, and nontoxic.

The decay (slow oxidation) of all organic materials produces CO_2. The burning of hydrocarbon fuels also produces CO_2. Fresh air contains approximately 0.033% CO_2 by volume. In the respiratory action (breathing) of all animals and humans, CO_2 is exhaled. *See* RESPIRATORY SYSTEM.

Carbon dioxide gas may be liquefied or solidified. For example, if the gas is compressed to 300 pounds per square inch gage (psig; 2 megapascals) and cooled to 0°F (−18°C) it becomes a liquid; or it may be liquefied at 70°F (21°C) by being compressed to 838 psig (5.78 MPa). Above the critical temperature of 87.9°F (31.1°C), CO_2 exists only as a gas, regardless of the pressure applied.

If liquid CO_2 is cooled to −69.9°F (−56.6°C) the pressure drops to 60.4 psig (417 kPa), and dry ice snow is formed. This condition of pressure and temperature is known as the triple point of CO_2, at which all three phases, solid, liquid, and gas, may exist in equilibrium with one another. Carbon dioxide cannot exist as a liquid below the triple-point temperature and pressure. If the pressure on dry-ice snow is reduced to atmospheric, its temperature drops to −109.3°F (−78.5°C). As solid CO_2 absorbs heat from its surroundings, it transforms directly to a gas (sublimes), hence the name dry ice. In still air, a film of gaseous CO_2 surrounds the dry ice, and the sublimation temperature is −109.3°F (−78.5°C); but in a vacuum or in rapidly moving air in which the CO_2 gas film is stripped away, the sublimation temperature drops to −130°F (−90°C) or lower.

Carbon dioxide is obtained commercially from four sources: gas wells, fermentation, combustion of carbonaceous fuels, and as a by-product of chemical processing.

Applications. Carbon dioxide may be used as a refrigerant, inerting medium, chemical reactant, neutralizing agent for alkalies, and pressurizing agent.

Refrigeration. Solid CO_2 has a greater refrigeration effect than water ice; the latent heat of sublimation of dry ice is 253.8 Btu/lb (595.0 kilojoules/kg), whereas the latent heat of fusion of water ice is 143.4 Btu/lb (336.2 kilojoules/kg). Furthermore, it is much colder than water ice and sublimes to a gas as it absorbs heat. Solid CO_2 may be furnished to the user in pressed blocks weighting 50–60 lb (23–27 kg) each, or as extruded pellets. It may also be made as snow at the point of use by expanding liquid CO_2 to atmospheric pressure; upon release to the atmosphere, liquid CO_2 at a pressure of 285 psig (1.97 MPa) produces about 47% dry ice snow and 53% gas by weight. However, of the total refrigerating effect originally available in the liquid, 85% remains in the ice and 15% in the vapor when the liquid CO_2 is warmed to 0°F (−18°C).

Dry ice blocks and pellets are used to chill, firm, and freeze meats, vegetables, and other perishable foods for preservation during transport. Dry ice snow produced from the expansion of liquid CO_2 may be applied directly to the surface of foods in a freezing chamber or tunnel, or it may be deposited on the contents of a container. Refrigeration may be achieved in trucks and other transport modes by vaporization of either solid or liquid CO_2. *See* DRY ICE.

Liquid CO_2 is injected into pneumatic conveyor systems to cool the contents during transport. Such items as flour, sugar, plastics, and core sand are quickly chilled by the dry ice snow and cold gas produced in the pneumatic line as liquid CO_2 is introduced through a thermostatically controlled orifice.

In low-temperature testing of aircraft and electronic parts to meet military or manufacturing specifications at −65°F (−54°C) and below, liquid CO_2 is expanded through an orifice to form dry-ice snow, which is either blown directly on the part or circulated in a test chamber; solid CO_2 is used in certain kinds of testing.

Carbon dioxide liquid injected into a hollow plastic shape while it is still in the mold reduces the temperature of the internal surface and, in turn, decreases the time required for the plastic to become sufficiently hardened to be self-supporting, allowing the part to be removed from the mold more quickly. Thus blow-molding equipment can operate on a shorter cycle time, and production can be speeded up.

Carbon dioxide liquid or ice may be used to advantage for removing the flash (mold marks) from molded rubber parts. This is done in an insulated tumbling barrel in which the mechanical action easily removes the flash that has been embrittled by contact with solid CO_2.

Carbon dioxide snow is added directly to choppers and mixers used in the preparation of hamburger, sausage, and prepared meat products such as bologna. The quick chilling reduces meat temperature rapidly, retarding bacterial growth.

Carbon dioxide is used to stiffen shortening for homogeneous blending with the dry ingredients used in piecrust mixes; to chill spices, chemicals, sugar, and rubber during high-speed grinding; and to prevent softening of thermoplastic materials during pulverizing.

Inerting. Carbon dioxide does not react with oxygen, nor does it normally function as an oxidizing agent. At high temperatures, it dissociates into carbon monoxide and oxygen (about 1% at 2800°F or 1538°C).

Liquid CO_2 is an effective fire-extinguishing agent because it rapidly reduces the temperature of

burning materials below the ignition point, and the gas, being heavy and inert, displaces air, usually the source of oxygen, and blankets the flames. It is particularly effective in extinguishing fires when water cannot be used, for example, in oil and electrical fires. *See* FIRE EXTINGUISHER.

Oil tankers, barges, storage vessels, and pipelines are quickly and safely inerted with CO_2 to allow welding and other repairs. In the automatic electric–tungsten–inert gas welding of steel using a filler wire, CO_2 is used to blanket the arc, thus preventing oxidation of the molten metal. *See* WELDING AND CUTTING OF MATERIALS.

Carbon dioxide is useful in many other applications in which an inert gas is required to prevent oxidation, as in packaging foodstuffs, spray-drying eggs and other solids, blanketing paint ingredients during manufacture to prevent the formation of "skin," and protecting grain and other bulk foods stored in silos.

Chemical applications. The largest single market for CO_2 for chemical purposes is in the preparation of carbonated beverages, in which the weak carbonic acid formed by the CO_2 acts as a taste enhancer and preservative. Other uses include the hardening of foundry cores in the core box by the reaction of CO_2 with the sand binder, the neutralization of excess alkalinity in water or industrial wastes, the manufacture of salicylic acid for aspirin, the production of pure carbonates and bicarbonates, as an intermediate in the preparation of titanium dioxide, and the stimulation of oil and gas wells.

Pressure medium. Life rafts and life preservers are packaged to include small CO_2 cartridges which permit rapid inflation upon operation of a quick-opening valve. Carbon dioxide in a cartridge is used as the propellent in certain pistols. It is also used as a propellent in pressure packaging, for example, in aerosol cans, and in many instances can replace fluorinated hydrocarbons at a fraction of the cost.

Manufacture. Most CO_2 is obtained as a by-product from steam-hydrocarbon reformers used in the production of ammonia, gasoline, and other chemicals; other sources include fermentation, deep gas wells, and direct production from carbonaceous fuels. Whatever the source, the crude CO_2 (containing at least 90% CO_2) is compressed in either two or three stages, cooled, purified, condensed to the liquid phase, and placed in insulated storage vessels. Commercial liquid CO_2 is usually stored at a pressure of 225–325 psig (1.55–2.2 MPa) and is maintained in this range by refrigeration. When the liquid is placed into high-pressure cylinders and stored at ambient temperature, however, the pressure in the liquid rises to 1071 psig (7.385 MPa) when the temperature is 87.8°F (31.0°C) [the initial point], or higher, if the cylinder is completely filled with liquid.

Distribution. Carbon dioxide is distributed in three ways: in high-pressure uninsulated steel cylinders; as a low-pressure liquid in insulated truck trailers or rail tank cars; and as dry ice in insulated boxes, trucks, or boxcars.

The size of the high-pressure cylinders is limited because of the weight involved. Most commercial cylinders contain either 20 or 50 lb (9 or 23 kg) of CO_2 and are designed to deliver CO_2 gas, which can be reduced to the pressure required by the user through the action of a pressure regulator. However, some cylinders are equipped with a tube that reaches to the bottom of the cylinder and siphons liquid CO_2 to the outlet, provided the temperature of the contents is below the critical temperature (87.9°F or 31.1°C).

Truck trailer capacity is restricted by laws which govern the gross vehicle weight. Up to 20 tons (18 metric tons) of liquid CO_2 may be shipped in an insulated trailer. Suppliers of CO_2 provide their customers with storage vessels varying in capacity from 4 to 150 tons (3.6 to 135 metric tons).

When CO_2 is manufactured and transported as dry ice, losses occur because of sublimation. By storing the dry ice in a well-insulated container, losses may be kept within economical limits.

Solid CO_2 can be converted to the liquid form by placing it in a heavy-walled steel vessel known as a converter. After the vessel is sealed, it is allowed to warm to room temperature. As the temperature of the CO_2 rises past the triple point, it is converted to liquid. J. S. Lindsey

Bibliography. American Society of Heating, Refrigerating, and Air-Conditioning Engineers, *ASHRAE Handbook of Fundamentals*, 1993; W. C. Clark, *Carbon Dioxide Review*, 1982; J. Wisniewski and R. N. Sampson (eds.), *Terrestrial Carbon Fluxes: Quantification of Sinks and Sources of CO_2*, 1993; A. S. Young, *Carbon Dioxide*, 1993.

Carbon nanotubes

Molecular-scale tubes of graphitic carbon with outstanding properties. They are among the stiffest and strongest fibers known, and have remarkable electronic properties and many other unique characteristics. They have attracted huge academic and industrial interest, with thousands of papers on nanotubes being published every year. Commercial applications have been rather slow to develop, primarily because of the high production costs of the best-quality nanotubes.

History. Interest in carbon nanotubes is a direct consequence of the synthesis of buckminsterfullerene, C_{60}, and other fullerenes in 1985. The discovery that carbon could form stable, ordered structures other than graphite and diamond stimulated researchers worldwide to look for other new forms of carbon. The search was given new impetus in 1990 when it was shown that C_{60} could be produced in a simple arc-evaporation apparatus readily available in all laboratories. Using such an evaporator, the Japanese scientist Sumio Iijima discovered fullerene-related carbon nanotubes in 1991. The tubes contained at least two layers, often many more, and ranged in outer diameter from about 3 to 30 nanometers. They were invariably closed at both

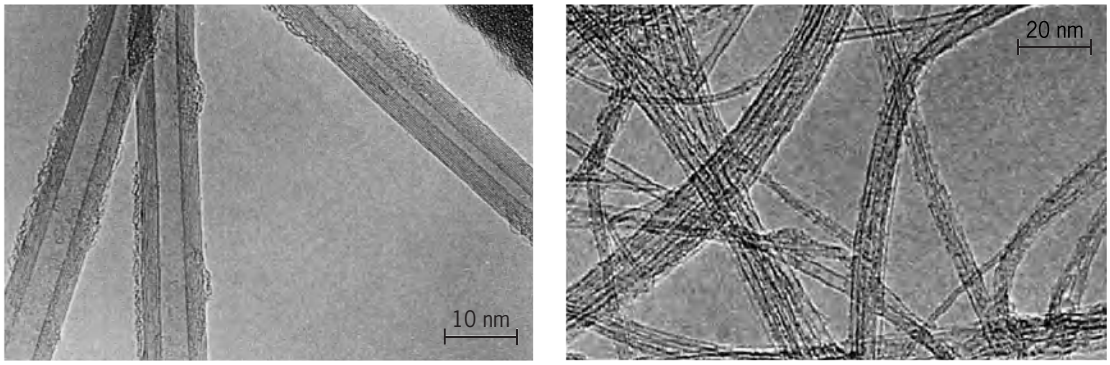

(a) (b)

Fig. 1. Transmission electron micrographs of carbon nanotubes. (*a*) Multiwalled tubes. (*b*) Single-walled tubes.

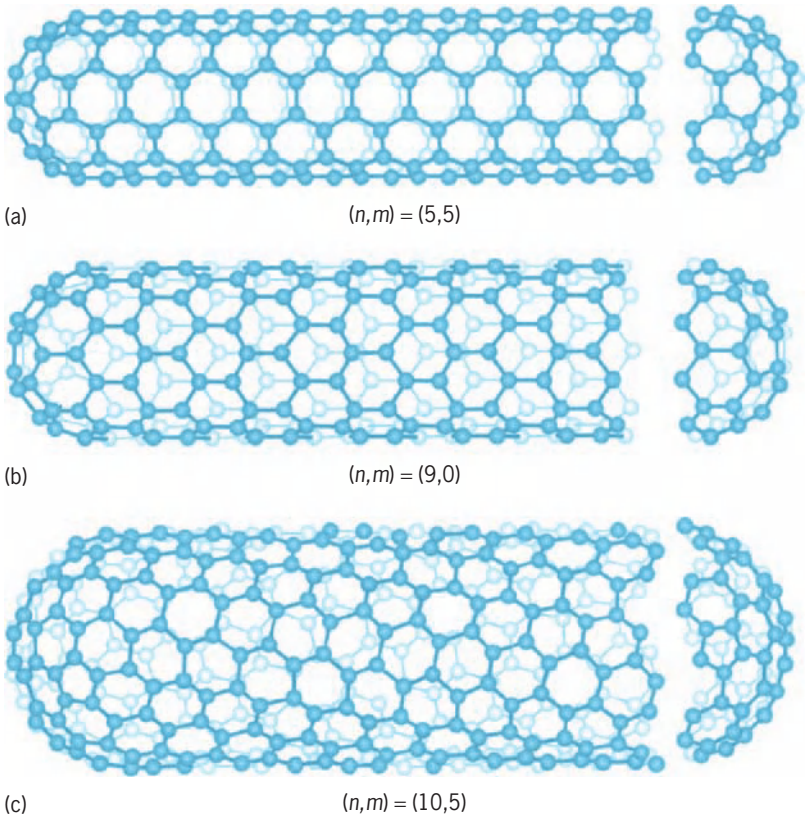

Fig. 2. Nanotube structures of (*a*) armchair, (*b*) zigzag, and (*c*) chiral configurations.

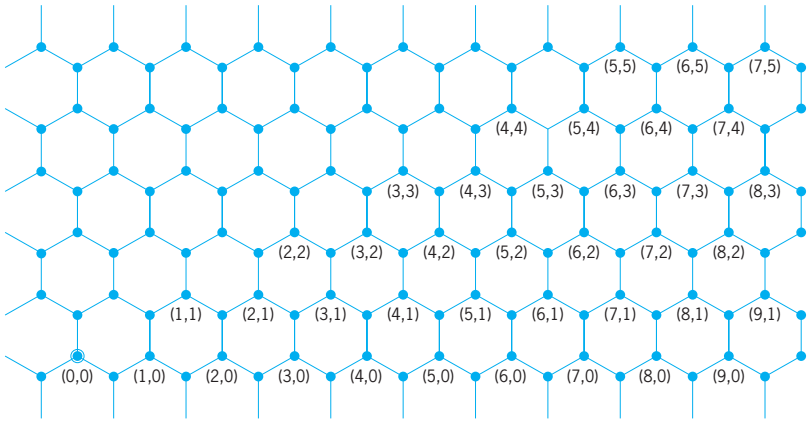

Fig. 3. Graphene sheet with carbon atoms labeled using the (*n*, *m*) notation.

ends. A transmission electron micrograph of some multiwalled nanotubes is shown in **Fig. 1***a*. In 1993, a new class of carbon nanotube was discovered, with just a single layer. These single-walled nanotubes are generally narrower than the multiwalled tubes, with diameters typically in the range 1–2 nm. Figure 1*b* shows some typical single-walled tubes. *See* CARBON; DIAMOND; FULLERENE; GRAPHITE.

It was soon established that these new fibers had a range of exceptional properties, and this sparked an explosion of research into carbon nanotubes. However, nanoscale tubes of carbon, produced catalytically, had been known for many years before Iijima's discovery. The main reason why these early tubes did not excite wide interest is that they were structurally rather imperfect, so they did not have particularly interesting properties.

Structure. The bonding in carbon nanotubes is sp^2, with each atom joined to three neighbors, as in graphite. The tubes can therefore be considered as rolled-up graphene sheets. (Graphene is an individual graphite layer.) There are three distinct ways in which a graphene sheet can be rolled into a tube, as shown in **Fig. 2**. The first two ways, known as armchair and zigzag, have a high degree of symmetry. The third class of tube, which in practice is the most common, is known as chiral, meaning that it can exist in two mirror-related forms. The structure of a nanotube can be specified by a vector, (n,m), which defines how the graphene sheet is rolled up. This can be understood with reference to **Fig. 3**. To produce a nanotube with the indices (6,3), say, the sheet is rolled up so that the atom labeled (0,0) is superimposed on the one labeled (6,3). It can be seen from the figure that $m = 0$ for all zigzag tubes while $n = m$ for all armchair tubes.

Synthesis. The arc-evaporation method, which produces the best-quality nanotubes, involves passing a current of about 50 amps between two graphite electrodes in an atmosphere of helium. This causes the graphite to vaporize, some of it condensing on the walls of the reaction vessel and some of it on the cathode. It is the deposit on the cathode which contains the carbon nanotubes. Single-walled nanotubes are produced when cobalt and nickel, or some other metal, is added to the anode.

It has been known since the 1950s, if not earlier, that carbon nanotubes can also be made by passing a carbon-containing gas, such as a hydrocarbon, over a catalyst. The catalyst consists of nano-sized particles of metal, usually iron, cobalt, or nickel. These particles catalyze the breakdown of the gaseous molecules into carbon, and a tube then begins to grow with a metal particle at the tip. In 1996 it was shown that single-walled nanotubes could also be produced catalytically. The perfection of carbon nanotubes produced in this way has generally been poorer than those made by arc-evaporation, but great improvements in the technique have been made. The big advantage of catalytic synthesis over arc-evaporation is that it can be scaled up for volume production. *See* CATALYSIS.

The third important method for making carbon nanotubes involves using a powerful laser to vaporize a metal-graphite target. This can be used to produce single-walled tubes at high yield.

Properties. The strength of the sp^2 carbon-carbon bonds gives carbon nanotubes amazing mechanical properties. The stiffness of a material is measured in terms of Young's modulus, the rate of change of stress with applied strain. The Young's modulus of the best nanotubes can be as high as 1000 GPa, which is approximately 5× higher than steel. The tensile strength, or breaking strain, of nanotubes can be up to 63 GPa, around 50× higher than steel. These properties, coupled with the lightness of carbon nanotubes, give them great potential in applications, such as in aerospace engineering. It has even been suggested that nanotubes could be used in the "space elevator," an Earth-to-space cable first proposed by Arthur C. Clarke. *See* STRESS AND STRAIN; YOUNG'S MODULUS.

The electronic properties of carbon nanotubes are also extraordinary. A notable fact is that nanotubes can be metallic or semiconducting depending on their structure. Thus, some nanotubes have conductivities higher than that of copper, while others behave more like silicon. There is great interest in the possibility of constructing nanoscale electronic devices from nanotubes, and some progress is being made. However, in order to construct a useful device we would need to arrange many thousands of nanotubes in a defined pattern, and we do not yet have that degree of control. *See* SEMICONDUCTOR.

Carbon nanotubes are already being used in several areas of technology. These include flat-panel displays, scanning probe microscopes, and sensing devices. The unique properties of carbon nanotubes will undoubtedly lead to many more applications. *See* NANOTECHNOLOGY. Peter J. F. Harris

Bibliography. P. M. Ajayan, Nanotubes from carbon, *Chem. Rev.*, 99:1787–1799, 1999; P. G. Collins and P. Avouris, Nanotubes for electronics, *Sci. Amer.*, 283:38–45, December 2000; M. S. Dresselhaus, G. Dresselhaus, and P. Avouris (eds.), *Carbon Nanotubes: Synthesis, Structure, Properties and Applications*, Springer-Verlag, 2001; P. J. F. Harris, *Carbon Nanotubes and Related Structures: New Materials for the Twenty-first Century*, Cambridge University Press, 1999; M. Meyyappan, *Carbon Nanotubes: Science and Applications*, CRC Press, 2004; B. I. Yakobson and R. E. Smalley, Fullerene nanotubes: $C_{1,000,000}$ and beyond, *Amer. Scientist*, 85:324–337, 1997.

Carbon-nitrogen-oxygen cycles

A group of nuclear reactions that involve the interaction of protons (nuclei of hydrogen atoms, designated by ^1H) with carbon, nitrogen, and oxygen nuclei. The cycle involving only isotopes of carbon and nitrogen is well known as the carbon-nitrogen (CN) cycle. These cycles are thought to be the main source of energy in main-sequence stars with mass 40% or more in excess of that of the Sun. Completion of any one of the cycles results in consumption of four protons (4 ^1H) and the production of a helium (^4He) nucleus plus two positrons (e^+) and two neutrinos (ν). The two positrons are annihilated with two electrons (e^-), and the total energy release is 26.73 MeV. Approximately 1.7 MeV is released as neutrino energy and is not available as thermal energy in the star. The energy $E = 26.73$ MeV arises from the mass difference between four hydrogen atoms and the helium atom, and is calculated from the Einstein mass-energy equation $E = \Delta mc^2$, where Δm is the mass difference and c^2 is the square of the velocity of light. Completion of a chain can be thought of as conversion of four hydrogen atoms into a helium atom. Because the nuclear fuel that is consumed in these processes is hydrogen, they are referred to as hydrogen-burning processes by means of the carbon-nitrogen-oxygen (CNO) cycles. *See* SOLAR NEUTRINOS.

Carbon-nitrogen cycle. The original carbon-nitrogen cycle was suggested independently by H. A. Bethe and C. F. von Weiszäcker in 1938 as the source of energy in stars. In the first reaction of the carbon-nitrogen cycle, a carbon nucleus of mass 12 captures a proton, forming a nitrogen nucleus of mass 13 and releasing a photon of energy, 1.943 MeV. This may be written: $^{12}C + {}^1H \rightarrow {}^{13}N + \gamma$, or $^{12}C(p,\gamma)^{13}N$. The nucleus ^{13}N is unstable and decays by emitting a positron (e^+) and a neutrino (ν). In reaction form, $^{13}N \rightarrow {}^{13}C + e^+ + \nu$ or $^{13}N(e^+\nu)$-^{13}C. The cycle continues through ^{14}N, ^{15}O, and ^{15}N by the reactions and decays shown in $^{12}C(p,\gamma)^{13}N(e^+\nu)^{13}$

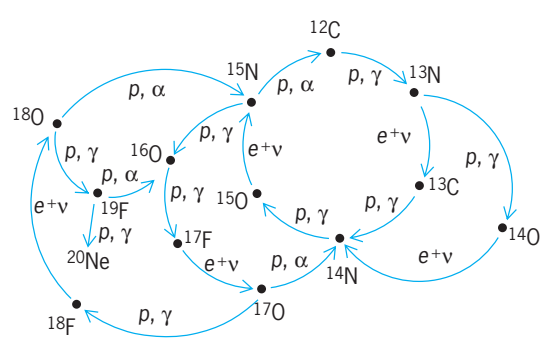

Carbon-nitrogen-oxygen cycles. (*After J. Audouze, ed., CNO Isotopes in Astrophysics, D. Reidel, 1977*)

TABLE 1. Energies involved in the carbon-nitrogen cycle, in MeV

Reaction	Q	Q (thermal)	E_ν^{max}
$^{12}C + {}^1H \rightarrow {}^{13}N + \gamma$	1.943	1.943	
$^{13}N \rightarrow {}^{13}C + e^+ + \nu$	2.221	1.51	1.199
$^{13}C + {}^1H \rightarrow {}^{14}N + \gamma$	7.551	7.551	
$^{14}N + {}^1H \rightarrow {}^{15}O + \gamma$	7.297	7.297	
$^{15}O \rightarrow {}^{15}N + e^+ + \nu$	2.753	1.75	1.731
$^{15}N + {}^1H \rightarrow {}^{12}C + {}^4He$	4.966	4.966	

$C(p,\gamma)^{14}N(p,\gamma)^{15}O(e^+\nu)^{15}N(p, \alpha)^{12}C$. These reactions form the second cycle from the right in the **illustration**. The reaction $^{15}N(p,\alpha)^{12}C$ represents emission of an alpha particle (^4He nucleus) when ^{15}N captures a proton and this cycle returns to ^{12}C. Because of the cycling, the total number of carbon, nitrogen, and oxygen nuclei remains constant, so these nuclei act as catalysts in the production of a helium nucleus plus two positrons and two neutrinos with the release of energy. The positrons that are created annihilate with free electrons rapidly after creation, so the energy used in their creation is returned to the energy fund. Synthesis of some of the rare carbon, nitrogen, and oxygen isotopes is accomplished through hydrogen burning by means of the carbon-nitrogen-oxygen cycles.

Table 1 shows the energy release, Q, and the average release of thermal energy, Q (thermal), in each reaction of the cycle, as well as the maximum neutrino energy E_ν^{max} in the two reactions in which a neutrino is produced.

A standard model for the Sun has been developed by J. N. Bahcall and colleagues to determine the flux of solar neutrinos that should be coming to the Earth. In this model, the site of maximum hydrogen burning is just outside the central core of the Sun. At a distance of 0.0511 solar radius from the center, chosen to lie in the most active part of that site for hydrogen burning in the proton-proton chains, the temperature is 14×10^6 K, the density is $\rho = 112$ g/cm^3 (112 times the density of water), and the mass fraction of hydrogen is $X(^1H) = 0.483$. The reaction rates, R, calculated for the carbon-nitrogen cycle reactions at that temperature, are shown in **Table 2**. Small corrections for screening of the Coulomb field between nuclei by electrons have been neglected. Taking into account the mass fraction of hydrogen and the density at the site, the number of reactions per nucleus per second is given by $\lambda = \rho RX(^1H)/1.0078$. By inverting λ, the mean lifetimes, τ, of the carbon, nitrogen, and oxygen nuclei are obtained; these are shown in the last column of Table 2 in years. The fraction of energy from CNO cycles in the Sun is calculated to be only 1.5%; the remaining 98.5% of the energy is from hydrogen burning in the proton-proton chains.

CNO bicycle. Nuclear research in the 1950s led to the addition of a second cycle to the processes of hydrogen burning by the CNO cycles. Laboratory research has shown that in a very small number of proton captures by ^{15}N a photon is emitted with the formation of an oxygen nucleus of mass 16, leading to reactions which may cycle back to ^{14}N by $^{15}N(p,\gamma)^{16}O(p,\gamma)^{17}F(e^+\nu)^{17}O(p,\alpha)^{14}N$. The pair of cycles consisting of this cycle and the carbon-nitrogen cycle, that forms the main cycle and the first cycle to the left of it in the illustration, is called the carbon-nitrogen-oxygen bicycle. The rate of the $^{12}N(p,\alpha)^{12}C$ reaction is of the order of 10^3 times that of the $^{15}N(p,\gamma)^{16}O$ reaction at most temperatures expected in stellar interiors.

Other CNO cycles. In research since 1960, it has become apparent that many possible branches among the nuclei must be included in any analysis of hydrogen burning by carbon, nitrogen, and oxygen nuclei. For example, if the unstable nucleus ^{13}N to capture a proton before it decays in its mean lifetime of 862 s, a third cycle can occur through $^{13}N(p,\gamma)^{14}O(e^+\nu)^{14}N$. This cycle is displayed in the illustration by the branch on the right-hand side of the main carbon-nitrogen cycle and is known as the fast or the hot carbon-nitrogen cycle. The other possible branches leading to additional cycles shown in the diagram are due to competition between (p,γ) and (p,α) reactions. The added reactions are $^{17}O(p,\alpha)^{18}F(e^+\nu)^{18}O(p,\alpha)^{15}N$ and $^{18}O(p,\alpha)^{19}F(p,\alpha)^{16}O$. The reaction $^{19}F(p,\gamma)^{20}Ne$ shown in the illustration leads out of the carbon, nitrogen, and oxygen nuclei and hence away from CNO cycles.

There are two additional branches that may occur if the unstable fluorine nuclei (^{17}F and ^{18}F) capture protons before they can decay. *See* NUCLEAR FUSION; NUCLEAR REACTION; NUCLEOSYNTHESIS; PROTON-PROTON CHAIN; STELLAR EVOLUTION.

Georgeanne R. Caughlan

TABLE 2. Reaction rates and mean lifetimes for carbon-nitrogen cycle in the Sun

Reaction	R, s^{-1}/(mole)(cm^3)*	λ, s^{-1}	τ, years
$^{12}C + {}^1H \rightarrow {}^{13}N + \gamma$	8.19×10^{-17}	4.40×10^{-15}	7.20×10^6
$^{13}N \rightarrow {}^{13}C + e^+ + \nu$		1.16×10^{-3}	2.73×10^{-5}
$^{13}C + {}^1H \rightarrow {}^{14}N + \gamma$	2.83×10^{-16}	1.52×10^{-14}	2.08×10^6
$^{14}N + {}^1H \rightarrow {}^{15}O + \gamma$	2.94×10^{-19}	1.58×10^{-17}	2.01×10^9
$^{15}O \rightarrow {}^{15}N + e^+ + \nu$		5.68×10^{-3}	5.58×10^{-6}
$^{15}N + {}^1H \rightarrow {}^{12}C + {}^4He$	7.27×10^{-15}	3.91×10^{-13}	8.11×10^4

*From G. R. Caughlan and W. A. Fowler, Thermonuclear reaction rates, V, *Atom. Nucl. Data Tables*, vol. 40, 1988.

Bibliography. J. N. Bahcall et al., Standard solar models and the uncertainties in predicted capture rates of solar neutrinos, *Rev. Mod. Phys.*, 54:767–799, 1982; G. R. Caughlan and W. A. Fowler, Thermonuclear reaction rates, V, *Atom. Nucl. Data Tables*, vol. 40, 1988; M. J. Harris et al., Thermonuclear reaction rates, III, *Annu. Rev. Astron. Astrophys.*, 21:165–176, 1983.

Carbon star

Any star whose spectrum shows a higher abundance of carbon than of oxygen. The carbon enhancements are easily recognizable from strong spectral absorption bands of carbon in molecular forms such as C_2, CN, and CH. Many types of carbon star are known, covering a wide variety of masses, temperatures, abundances, and luminosities. The cool red appearance, high luminosity, and intrinsic variability of bright carbon stars show that they are giant stars.

Most stars with masses between 0.6 and 5 times the mass of the Sun must pass through a carbon-star phase. Stars like the Sun, called dwarf stars, derive their energy from nuclear fusion of hydrogen to produce helium. As a star evolves and its core supply of hydrogen disappears, it can begin to burn helium into carbon, becoming a giant star. Not all giant stars dredge up from the core a sufficient quantity of newly produced carbon to change the appearance of the visible stellar surface. Asymptotic giant branch (AGB) stars are further evolved stars where nuclear burning of helium surrounds a dense carbon-oxygen core. The helium shell burning causes large temperature gradients and convection, which efficiently dredge up the processed, carbon-rich material toward the stellar surface. *See* NUCLEOSYNTHESIS.

Asymptotic giant branch carbon stars expel much of their carbon-rich envelopes as a wind. During a single year of its mass-losing phase, a carbon star may expel up to 10^{-5} of a solar mass, about 3 Earth masses, of carbon-rich material. This material may form cool envelopes of ejected circumstellar dust that dim and redden its appearance. The expelled stellar envelopes eventually mix with the surrounding interstellar environment, so that much of the carbon throughout the Milky Way Galaxy comes from mass-losing carbon stars. Still, carbon stars are not common for two reasons. First, most stars have not yet reached the asymptotic giant branch phase of their evolution. Second, that phase of a star is short compared to the lifetime of a dwarf, so it is rare to see a star in its carbon-rich phase.

Since carbon is produced and brought to the surface only in giant stars, the more recent discovery of dwarf carbon stars at first appeared to contradict basic stellar evolution theory. However, matter expelled from asymptotic giant branch stars in binary star systems may fall onto and enhance the carbon abundance of a dwarf companion star. Since dwarf carbon stars are much fainter, only a few are yet known. The long lifetimes of dwarfs, however, all but ensure that many more currently exist than do carbon giants.

The R and N classification ranks carbon stars by their colors and spectral features. The N stars are in general cooler and show much stronger molecular bands. Spectral subclasses 0 through 8 represent decreasing temperature. CH stars are warm carbon giants with an intense CH molecular band. Both R and CH stars, while giants, also have luminosities too low to be asymptotic giant branch stars. Since CH stars are apparently all binaries, they may be former dwarf carbon stars that have themselves evolved into giants. However, the R stars occur in binary systems only as often as normal (oxygen-rich) red giants, so the origin of their enhanced carbon abundance remains a mystery. *See* BINARY STAR; DWARF STAR; GIANT STAR; STAR; STELLAR EVOLUTION.

Paul J. Green

Bibliography. C. Jaschek and M. Jaschek, *The Classification of Stars*, Cambridge University Press, corrected reprint, 1990; J. B. Kaler, *The Little Book of Stars*, Copernicus Books, New York, 2001; R. Kippenhahn and J. Steinberg (translator), *100 Billion Suns: The Birth, Life, and Death of the Stars*, reprint, Princeton University, 1993; S. J. Shawl, K. M. Ashman, and B. Hufnagle, *Discovering Astronomy*, 5th ed., Kendall/Hunt, Dubique, Iowa, 2006.

Carbonate minerals

The mineral species containing the CO_3^{2-} unit as an essential structural and chemical component. Members are distinguished by chemical composition (type of metal cation and in some cases other anion present) and crystal structure. More than 200 varieties of carbonate minerals are known, but only a small number occur in abundance. Calcite ($CaCO_3$) and dolomite [$CaMg(CO_3)_2$] are the most common species, occurring as the principal mineral components in limestone and Mg-rich limestone sedimentary rocks. Carbonate minerals exhibit limited stability at high temperature unless high carbon dioxide (CO_2) pressure exists, which explains their occurrence dominantly in environments near the Earth's surface and in sedimentary rocks. However, important occurrences in metamorphic rocks include marble and hydrothermal deposits, with only rare occurrences in carbonatite and kimberlite igneous rocks. *See* CARBON DIOXIDE; CARBONATITE; CALCITE; DOLOMITE; IGNEOUS ROCKS; KIMBERLITE; LIMESTONE; MARBLE; METAMORPHIC ROCKS; SEDIMENTARY ROCKS.

Main groups and other varieties. Calcite ($CaCO_3$) is the most common carbonate mineral, notably occurring as the primary mineral constituent of limestone. Other members of the calcite group include magnesite ($MgCO_3$), rhodochrosite ($MnCO_3$), siderite ($FeCO_3$), sphaerocobaltite ($CoCO_3$), gaspeite ($NiCO_3$), smithsonite ($ZnCO_3$), and otavite ($CdCO_3$). These minerals share the same rhombohedral

crystal structure and the general formula ACO_3, where A refers to a divalent metal cation, which balances the charge of the CO_3^{2-} anionic unit. *See* MAGNESITE; RHODOCHROSITE; SIDERITE; SMITHSONITE.

The aragonite group also has the general formula ACO_3 but possesses an orthorhombic crystal structure. In addition to aragonite ($CaCO_3$), which is abundant in modern marine sediments, other members of this group include strontianite ($SrCO_3$), witherite ($BaCO_3$), and cerussite ($PbCO_3$). *See* ARAGONITE; CERUSSITE; STRONTIANITE; WITHERITE.

A second important group of rhombohedral carbonates takes its name from its most abundant member, dolomite [$CaMg(CO_3)_2$], which is the primary constituent of Mg-rich limestone rock known also as dolomite, although some authors favor the term dolostone. With the general formula $AB(CO_3)_2$, members of this group contain two divalent metal cations (A and B) in a structure similar to that of calcite but in which the A and B cations are separated into alternating layers. Other members of this group include ankerite [$Ca(Fe,Mg)(CO_3)_2$], kutnahorite [$CaMn(CO_3)_2$], and the rare minrecordite [$CaZn(CO_3)_2$]. Other carbonate minerals with the $AB(CO_3)_2$ formula are known but their structures differ slightly from that of dolomite, with examples being norsethite [$BaMg(CO_3)_2$] and paralstonite [$CaBa(CO_3)_2$]. The mineral huntite [$Mg_3Ca(CO_3)_4$], known as a weathering product of dolomite or magnesite rocks, is rhombohedral and contains two metal cations but in a different ratio than in dolomite. *See* ANKERITE.

Some carbonate minerals that crystallize from water solutions remain hydrated, such as nesquehonite ($MgCO_3 \cdot 3H_2O$) and natron ($Na_2CO_3 \cdot 10H_2O$). Others may contain hydroxyl ion (OH^-), as do malachite [$Cu_2CO_3(OH)_2$] and hydrozincite [$Zn_5(CO_3)_2(OH)_6$]. Species containing the bicarbonate unit (HCO_3^-) form the group known as acid carbonates, with nahcolite ($NaHCO_3$) being the best-known example. Of the many less common or rare carbonate minerals not described here, some contain other anion units such as sulfate (SO_4^{2-}), phosphate (PO_4^{3-}), borate (BO_3^{3-}), chloride (Cl^-), or fluoride (F^-) in different proportions, combined with a variety of metal cations and existing in numerous complex structures. There are many examples of carbonate minerals containing uranium (U) or members of the rare-earth element group owing to the strong affinity of these metals for the CO_3^{2-} anion when present in water solutions. *See* MALACHITE; RARE-EARTH MINERALS.

Structure and properties. The CO_3^{2-} molecular group resembles an equilateral triangle with an oxygen atom at each corner and the carbon atom at the center, bonding only to the three oxygens. Metal cations form bonds with the oxygen atoms in different configurations, resulting in a variety of crystal structures. In the calcite group (ACO_3), the divalent metal cation is bonded to six oxygen atoms, each belonging to a separate CO_3 group (**Fig. 1***a*). The molar volumes and densities of members of this group differ owing to the different sizes of the metal cations. The crystal structure is rhombohedral (trigonal system) in which layers of CO_3 groups alternate with layers of metal cations perpendicular to the main symmetry axis. Symmetrically arranged planes of weak bonding cause a perfect cleavage (or fracture), giving a distinctive rhombohedral shape often seen in larger specimens of calcite (**Fig. 2**). Clear, colorless crystals of calcite are known as Iceland spar, after the location of an early discovery of high-quality specimens. Optically, rhombohedral carbonates exhibit high birefringence, and Iceland spar played an early role in understanding the polarization of light and remains useful for demonstrating the optical property double refraction (Fig. 2). Localities in Mexico, Brazil, and Russia provide most of the Iceland spar available today. *See* BIREFRINGENCE; CRYSTAL; CRYSTAL OPTICS; CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Members of the aragonite group (ACO_3) have an orthorhombic crystal structure in which the divalent metal cation is bonded to nine oxygen atoms. Layers of divalent metal cations alternate with layers of offset CO_3 groups (Fig. 1*b*). This arrangement results in a more dense packing of atoms than in the calcite structure, as demonstrated by its greater density: 2.94 g/cm^3 for aragonite versus 2.71 g/cm^3 for calcite. A characteristic structural feature commonly observed in aragonite is twinning, where different regions of the crystal are misoriented as mirror images. This may cause crystals to display a

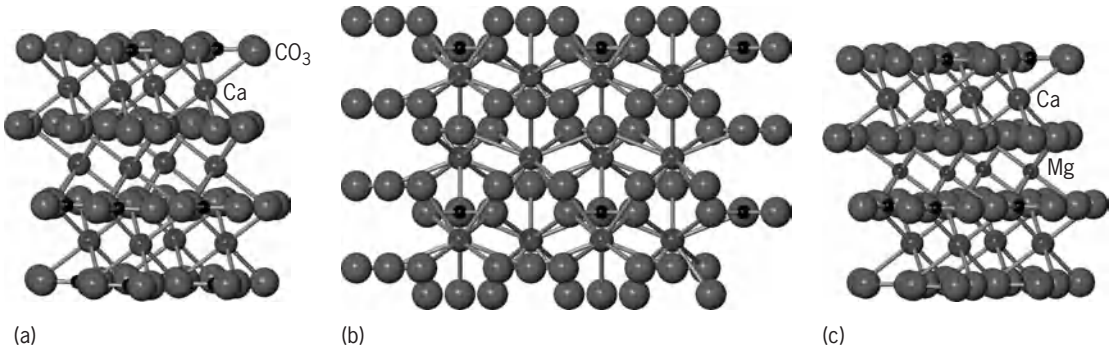

Fig. 1. Crystal structures of (*a*) calcite, (*b*) aragonite, and (*c*) dolomite.

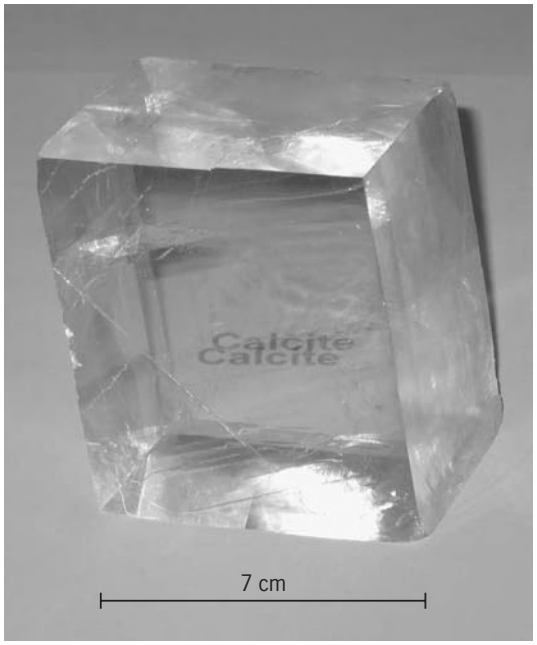

Fig. 2. Rhombohedral cleavage crystal of calcite (Iceland spar), showing the optical property known as double refraction.

pseudohexagonal symmetry that is readily identified in larger specimens (**Fig. 3**). *See* TWINNING (CRYSTALLOGRAPHY).

It is noteworthy that only $CaCO_3$ occurs in both the calcite and aragonite structural forms (as well as a third form called vaterite), a property known as polymorphism. The preference for a metal carbonate to form in either the calcite or the aragonite structure can be attributed to the size of the divalent metal. Mg, Mn, Fe, Co, Ni, Zn, and Cd have radii smaller than Ca, and occur exclusively in the rhombohedral calcite structure. In contrast, Sr, Ba, and Pb have radii larger than Ca, and occur only in the orthorhombic aragonite structure. The larger size of the metal cation in this latter group favors the greater number of bonds to oxygen atoms as found in the aragonite structure. The size of the Ca atom allows $CaCO_3$ to exist in both structure types. *See* POLYMORPHISM (CRYSTALLOGRAPHY).

The crystal structure of the dolomite group is based on that of the calcite group but differs by the segregation of the two divalent metal cations into separate, alternating layers (Fig. 1*c*), referred to as cation ordering. In the mineral dolomite [$CaMg(CO_3)_2$], the layer repeat has the sequence Ca-CO_3-Mg-CO_3-Ca. The dolomite structure is rhombohedral; hence, members can be confused with the calcite group based on appearance, but the true symmetry is lower than that of calcite. The cation ordering necessary to form the dolomite structure is related to the difference in size between the two divalent metal cations. For example, the Ca atom is nearly 40% larger than Mg. This difference favors their ordering into separate layers, as in dolomite. By comparison, divalent iron (Fe^{2+}) is only 8% larger than Mg, and these cations show no tendency to order into separate layers, and there is no $FeMg(CO_3)_2$ member of the dolomite group.

Carbonate minerals may be clear, translucent, or opaque and exhibit a wide variety of colors, usually reflecting the presence of minor impurities or inclusions. Some carbonate minerals display a characteristic color due to the dominance of a particular metal cation, with notable examples including the bright pink color of rhodochrosite ($MnCO_3$), the deep blue color of azurite [$Cu_3(CO_3)_2(OH)_2$], and the green of malachite [$Cu_2CO_3(OH)_2$], all of which are considered semiprecious stones. Calcite is particularly noted for exhibiting a rich variety of crystal morphologies and colors. Certain varieties of dolomite exhibit a distinctive curvature of its crystal faces. Most carbonate minerals have relatively low hardness on the Mohs scale (2.5–4.5) and rhombohedral varieties cleave readily. Many carbonate minerals dissolve in strong acids, releasing carbon dioxide. This serves as the basis for a common diagnostic test of some varieties in which several drops of a strong acid cause effervescence. *See* AZURITE.

Composition and phase relations. Significant compositional variability is observed for many carbonate minerals. Among members of the calcite group, the tendency to form solid solutions depends mostly on the respective sizes of the metal cations. The small difference in size between Fe^{2+} and Mg allows complete substitution for one another on

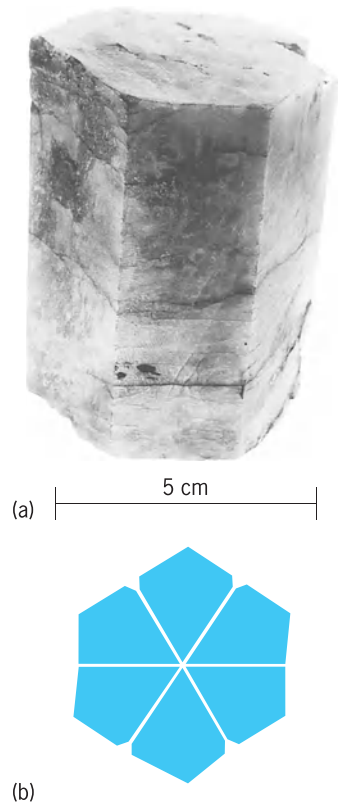

Fig. 3. Aragonite. (*a*) Pseudohexagonal twinning in specimen from Girgenti, Sicily (*American Museum of Natural History specimen*). (*b*) Arrangement of twins (*after C. Klein, Dana's Manual of Mineralogy, 21st ed., Wiley, 1993*).

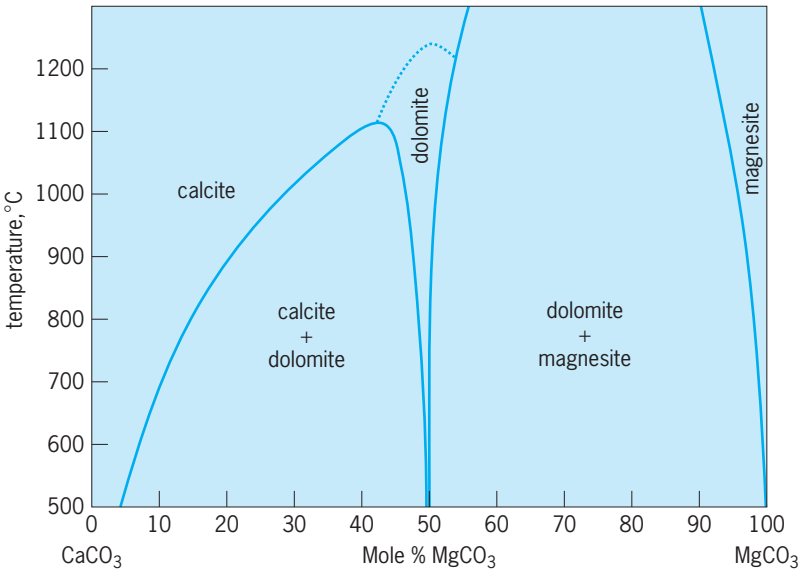

Fig. 4. Temperature-composition phase diagram for the system $CaCO_3$-$MgCO_3$. (*After J. R. Goldsmith and H. C. Heard, Subsolidus phase relations in the system $CaCO_3$-$MgCO_3$, J. Geol., 69:45–74, 1961*)

the same site, as illustrated by the complete solid solution between $FeCO_3$ (siderite) and $MgCO_3$ (magnesite). Extensive solid solutions exist for the end-member pairs $MgCO_3$-$FeCO_3$, $MgCO_3$-$MnCO_3$, and $FeCO_3$-$MnCO_3$. A greater difference in metal cation size results in limited formation of solid solutions between the end-member pairs $CaCO_3$-$FeCO_3$, $CaCO_3$-$MnCO_3$, and $CaCO_3$-$MgCO_3$. As noted above, the large difference in size between Ca and Mg allows the formation of the cation-ordered phase of dolomite at a composition intermediate between $CaCO_3$ and $MgCO_3$. These relationships are best seen in the phase diagram for the $CaCO_3$-$MgCO_3$ system, which also shows that increasing temperature favors solid solubility (**Fig. 4**). Despite unfavorable energetics, Mg substitution is common in calcite that forms in low-temperature environments, including in the shell material of some marine organisms, such as sea urchins. Limited cation substitution occurs among members of the aragonite group, although most aragonite contains minor Sr and strontianite ($SrCO_3$) generally contains some Ca. *See* SOLID SOLUTION.

Substitution among members of the dolomite group is more restricted, with the most important example being Fe^{2+} substitution for Mg in dolomite [$CaMg(CO_3)_2$] to form ankerite [$Ca(Fe,Mg)(CO_3)_2$]. Fe^{2+} is found to substitute up to a maximum of ~70% of the Mg sites, and the pure $CaFe(CO_3)_2$ member of the dolomite group is unknown in nature (**Fig. 5**). A distinction is usually made between species having less than 20% Fe substitution, referred to as ferroan dolomite, and those with greater Fe, called ankerite. Other cations, notably Mn, may substitute in dolomite and ankerite. Some examples of dolomite exhibit imperfect cation order and minor variation in Ca/Mg composition, typically an enrichment of Ca by up to 10%. These examples are less stable than ideal, well-ordered dolomite, and their occurrence is largely restricted to sedimentary conditions.

On heating to sufficiently high temperature, carbonate minerals decompose, giving off CO_2 gas and forming a metal oxide or a mixed solid. For calcite, this can be expressed as reaction (1). The

$$CaCO_3(s) \Leftrightarrow CaO(s) + CO_2(g) \qquad (1)$$

temperature at which decarbonation occurs varies with mineral composition and CO_2 pressure, among other factors. At ambient pressure in air, calcite and magnesite begin to decompose at approximately 700 and 500°C (1300 and 930°F), respectively. Decomposition of these minerals can be prevented up to much higher temperature only when the CO_2 pressure is increased significantly. Consequently, carbonate minerals generally persist under high-temperature conditions only when sufficiently deep in the crust that the high confining pressure prevents decomposition. In the presence of other minerals, carbonates may react at much lower temperature during metamorphism, releasing CO_2. Hydrous carbonate minerals decompose at relatively low temperature, restricting their occurrence to sedimentary deposits. On heating dolomite under sufficient CO_2 pressure to prevent decomposition, the cation ordering begins to decrease at approximately 900°C (1650°F) and vanishes completely by 1200°C (2200°F), at which point the structure is like that of the calcite group. *See* METAMORPHISM.

The $CaCO_3$ polymorphs calcite and aragonite are both common, whereas vaterite is rare, transforming

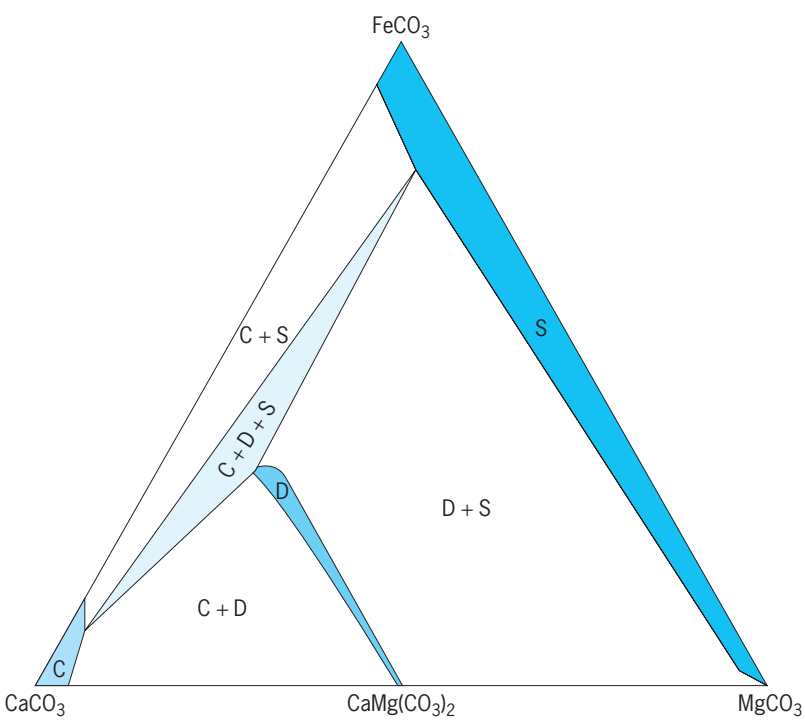

Fig. 5. Ternary phase diagram for the $CaCO_3$-$MgCO_3$-$FeCO_3$ system at 450°C (840°F). Solid solution phases are denoted: C = calcite; D = dolomite; S = siderite-magnesite. (*After P. E. Rosenberg, Subsolidus relations in the system $CaCO_3$-$MgCO_3$-$FeCO_3$ between 350 and 550°C, Amer. Mineral., 52:787–796, 1967*)

to calcite over periods as short as hours. At the temperature and pressure conditions found near the Earth's surface, calcite is the stable form, relative to aragonite. Consequently, aragonite transforms to calcite with time, which explains the occurrence of aragonite predominantly in recent sediments. Under sedimentary conditions, this transformation occurs only in the presence of water by dissolution of aragonite, followed by precipitation of calcite. Aragonite may persist for long periods of geologic time if isolated from water. Aragonite becomes more stable than calcite only at high pressure [approximately 3500 bars (350 MPa) at 25°C (77°F) and 7000 bars (700 MPa) at 300°C (570°F)], consistent with its greater density, and can be found in metamorphic rocks subjected to high pressure (**Fig. 6**). At high temperature and pressure, other structural forms of $CaCO_3$ have been produced experimentally but are not observed in nature (Fig. 6). Dolomite and magnesite have also been observed to undergo transformations to different structural forms at high pressure.

Solubility of carbonate minerals in water varies significantly depending on the acidity, the amount of dissolved carbon dioxide, and the total concentration of other dissolved ions. In general, carbonates tend to have solubilities greater than most silicate and oxide minerals but substantially less than the solubilities of common salts, such as NaCl (halite). Carbonate minerals are unusual in exhibiting a decrease in solubility with increasing temperature, a relationship called retrograde solubility. Under most conditions, this solubility behavior is attributable to a combination of factors, including the heat released during dissolution and the temperature influence on the solubility of carbon dioxide and dissolved carbonate species. *See* HALITE; OXIDE AND HYDROXIDE MINERALS; SILICATE MINERALS.

Occurrences and formation. Calcite and dolomite account for over 95% of the total abundance of carbonate minerals, primarily as constituents of sedimentary rocks (principally limestone and dolomite) but also as minor constituents in sandstone and shale, and in modern sediments. The tendency of carbon dioxide to dissolve in water, hydrate, and dissociate to form bicarbonate and carbonate species provides a source for crystallization of carbonate minerals, as illustrated for calcium carbonate [reaction (2)].

$$\begin{array}{l} CO_2(g) + H_2O \Leftrightarrow H^+ + HCO_3^- \\ \qquad\qquad\qquad\qquad \downarrow \\ \qquad\qquad Ca^{2+} + HCO_3^- \Leftrightarrow CaCO_3(s) + H^+ \end{array} \tag{2}$$

Significant sources of dissolved bicarbonate in water are derived from reactions involving the weathering of rocks. Crystallization of calcium carbonate may occur by purely inorganic reactions or through the biologic activity of organisms, a process known as biomineralization. In seawater, organisms such as bivalves, gastropods, and marine plankton, called foraminifera, crystallize hard exoskeletons or shells composed of calcite or aragonite. This biomineralization accounts for most of the formation of

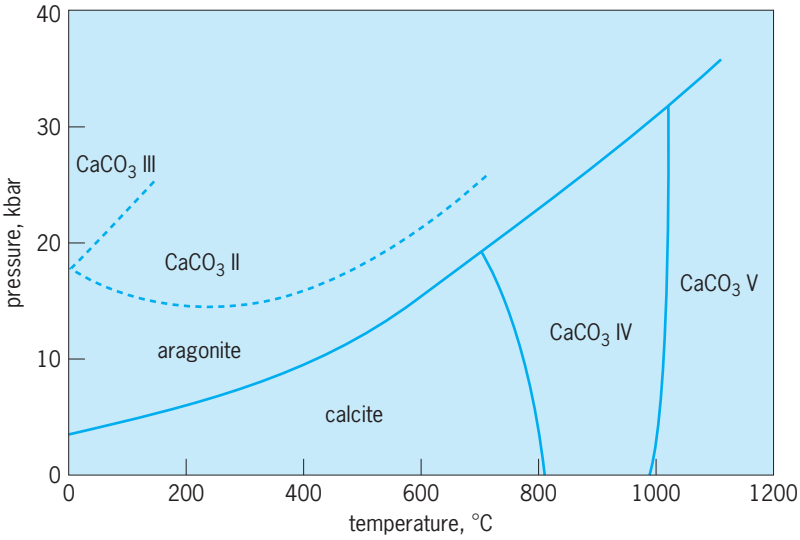

Fig. 6. Pressure-temperature phase diagram for the $CaCO_3$ system. Dashed lines refer to metastable equilibria. (*After W. D. Carlson, The calcite-aragonite equilibrium: effects of Sr substitution and anion orientational disorder, Amer. Mineral., 65:1252–1262, 1980*)

$CaCO_3$ in the oceans, despite the fact that shallow seawater is oversaturated with respect to $CaCO_3$. Accumulation of the hard skeletal remains of dead organisms represents one of the principal sources of calcium carbonate in limestone, which is formed after burial and cementation. The successive transfer of carbon dioxide gas to seawater as dissolved carbonate, its removal through biomineralization, followed by formation of limestone are fundamental processes in the global carbon cycle, which is linked to climate change over time. Carbonate minerals have also played an historic role in climate studies. Much of our present knowledge about global climate change has been derived from the carbon and oxygen isotopic compositions of marine calcium carbonate fossils. The isotopic ratios acquired during formation of the carbonate shells are sensitive to temperature. *See* BIOGEOCHEMISTRY; CLIMATOLOGY; GLOBAL CLIMATE CHANGE; WEATHERING PROCESSES.

Seawater is also oversaturated with respect to dolomite [$CaMg(CO_3)_2$], yet no significant crystallization is known, either inorganically or through biomineralization. The abundance of dolomite in sedimentary rocks, yet the lack of any significant dolomite formation in seawater, has been a source of controversy. The formation of most dolomite in sedimentary rock is thought to result from alteration of calcium carbonate sediments or limestone by Mg-containing water solutions.

Calcite also commonly occurs in soils from arid and semiarid climates, where evaporation favors the accumulation of calcium carbonate, sometimes in dense layers referred to as calcrete or caliche. Where rainfall is abundant, solution of carbonate minerals in limestone or dolomite may produce cave systems. Loss of CO_2 gas from calcium carbonate–rich water solutions in caves causes crystallization of principally calcite and aragonite, often in unusual deposits referred to generally as speleothem

and including forms known as stalagmite and stalactite, soda straws, flowstone, and cave pearls. Calcite may also precipitate from springs, especially hot springs, as layered, massive, formations called travertine. *See* CALICHE; CAVE; STALACTITES AND STALAGMITES; TRAVERTINE.

Calcite and dolomite, and less commonly other carbonates, occur in marble (metamorphosed limestone) and as minor constituents of other metamorphic rocks. Aragonite is found in rare circumstances in metamorphic rocks, where its presence along with other high-pressure minerals is regarded as important evidence for pressures in excess of 5000–7000 bars (500–700 MPa), corresponding to depths greater than 10–15 mi (16–24 km) in the Earth's crust. Well-formed specimens of calcite and dolomite, in which the largest individual crystals may exceed 1 m in size, are generally hydrothermal in origin, typically forming along fractures or in cavities in rocks through which hot-water solutions flow. Carbonate minerals are commonly associated with the formation of ore bodies, often during late stages when circulating hot fluids result in their hydrothermal crystallization. Carbonate minerals also form as alteration products of primary ore minerals, with the resulting carbonate in some instances being an important ore mineral, such as malachite. *See* HYDROTHERMAL ORE DEPOSITS.

Although generally uncommon in igneous rocks, carbonates are the dominant mineral components of the relatively rare carbonatite rocks, and are found in kimberlite rocks, which are host to diamonds. The origin of carbonate magma remains enigmatic, with evidence supporting both a primary mantle source and separation from a parent melt, called liquid immiscibility. *See* DIAMOND; MAGMA.

Iron- and magnesium-rich carbonate minerals have been documented in meteorite fragments thought to have originated on Mars, but their origin remains uncertain.

Uses. Calcite finds widespread application as a filler, binder, or coloring agent in pharmaceuticals, paper, plastic and composite materials, foodstuffs, and paints. It is widely used in chemical processing and in wastewater treatment. Both natural and precipitated forms of calcium carbonate are used as calcium supplements and antacids. Calcite, as limestone, has long been used as an essential reaction ingredient in the production of steel. Calcite and dolomite, as crushed or pulverized limestone, find extensive use in roadway construction, agriculture, and lawn care. Limestone and marble are used as building stone, tile, or facing, and serve as the material in numerous monuments and statues. Calcium carbonate, after heating to produce lime (CaO), is an important raw material in the manufacturing of glass and portland cement. Richard J. Reeder

Bibliography. L. L. Y. Chang, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, vol. 5B: *Non-Silicates: Sulfates, Carbonates, Phosphates, Halides*, 2d ed., Geological Society, London, 1996; J. W. Morse and F. T. Mackenzie, *Geochemistry of Sedimentary Carbonates*. Elsevier, 1990; R. J. Reeder (ed.), *Carbonates: Mineralogy and Chemistry, Reviews in Mineralogy*, vol. 11, Mineralogical Society of America, 1983.

Carbonatite

An igneous rock in which carbonate minerals make up at least half the volume. Individual occurrences of carbonatite are not numerous (about 330 have been recognized) and generally are small, but they are widely distributed. Carbonatites are scientifically important because they reveal clues concerning the composition and thermal history of the Earth's mantle. Economically, carbonatites provide important mines for some mineral commodities and are virtually the only sources of a few. *See* CARBONATE MINERALS.

Mineralogy. The carbonate minerals that dominate the carbonatites are, in order of decreasing abundance, calcite, dolomite, ankerite, and rarely siderite and magnesite. Sodium- and potassium-rich carbonate minerals have been confirmed in igneous rocks at only one locality, the active volcano Oldoinyo Lengai in Tanzania (**Fig. 1**). Noncarbonate minerals that typify carbonatites are apatite, magnetite, phlogopite or biotite, clinopyroxene, amphibole, monticellite, perovskite, and rarely olivine or melilite. Secondary minerals, produced by alteration of primary magmatic minerals, include barite, alkali feldspar, quartz, fluorite, hematite, rutile, pyrite, and chlorite. Minerals that are important in some carbonatites because they carry niobium, rare-earth elements, and other metals in concentrations high enough for profitable extraction are pyrochlore, bastnaesite, monazite, baddeleyite, and bornite.

Styles of occurrence. Carbonatites occur in a variety of forms, both intrusive and volcanic. Lava flows are small and rare, but pyroclastic carbonatites are numerous as thick near-vent accumulations (tuff cones) and thin but widespread airfall and surge deposits. Some pyroclastic carbonatites are especially significant, because they preserve the forms of quenched droplets of carbonate-rich liquid and

Fig. 1. Carbonatite lava from Oldoinyo Lengai volcano, Tanzania; thin section between crossed polarizers. Longest dimension of the field of view is 2.5 mm.

also record the mineral assemblages that are nearly always erased in intrusive rocks by low-temperature alteration. Carbonatite airfall tuffs are paleontologically important, because they tend to be fine grained and rapidly become lithified, preserving delicate organic structures. The hominid footprints at Laetoli, Tanzania, and remarkably detailed molds of plants and caterpillars are examples. *See* TUFF.

Most intrusive carbonatites form arrays of dikes that may be parallel, radiating, or concentric and arcuate. More irregular intrusive bodies are also known, but all are small compared to those of silicate-rich igneous rocks; the largest known carbonatite body, at Sokli in Finland, has a surface area of only 20 km^2 (8 mi^2). Carbonate-rich magmas appear to have had low viscosities, but some were emplaced with explosive violence because of their high gas content. A few carbonatites extend for several kilometers along fault surfaces; they may have been older igneous bodies that were caught up in the fault displacement and acted as cold but weak ductile lubricant between the rock masses; or they may have been injected as magma. Except for the absence of large homogeneous plutons, carbonatites form the full array of igneous rock styles that are shown by silicate-rich igneous rocks. *See* MAGMA; PLUTON.

Associated rocks. Carbonatites usually have been found with low-silica, high-alkali igneous rocks, typically melilitites, nephelinites, tephrites, and phonolites. The silicate and carbonate rocks are intimately mixed in many intrusive complexes, and the carbonatites are among the youngest and generally least abundant of rock types in any complex. Carbonatites that are unaccompanied by other igneous rocks have been recognized. Examples are at Fort Portal, Uganda; Kaluwe, Zaire; and Sarfartoq, Greenland.

Geographic distribution. Carbonatites are widely scattered (**Fig. 2**). About half of all carbonatite occurrences so far recognized are in Africa. The only occurrences known on oceanic crust are in the Canary and Cape Verde islands; however, carbonatite associated with deep-water marine sediments occurs under the Semail ophiolite in the United Arab Emirates. These occurrences suggest that carbonatites might actually be common on and within oceanic crust but are usually buried or destroyed by subduction. *See* OPHIOLITE.

Chemical compositions. Carbonatites, compared to the inferred composition of the Earth's mantle and to other igneous rocks, are greatly enriched in niobium, rare-earth elements, barium, strontium, phosphorus, and fluorine, and they are relatively depleted in silicon, aluminum, iron, magnesium, nickel, titanium, sodium, potassium, and chlorine. These extreme differences are attributed to strong fractionation between carbonate liquid on the one hand and silicate and oxide solid phases on the other during separation of the carbonate liquid from its source. Strontium and neodymium isotope ratios indicate that the sources of carbonatites are geologically old, inhomogeneous, and variably depleted in the radioactive parent elements rubidium and samarium.

Genesis of carbonate-rich magma. For more than two centuries, it has been known that calcite decomposes to lime (calcium oxide; CaO) and carbon dioxide (CO_2) when exposed to high temperature at atmospheric pressure. Furthermore, at high pressure calcium carbonate ($CaCO_3$) liquid is stable only at temperatures far higher than are reasonably expected in the Earth's crust. It was therefore difficult for geologists to imagine liquid carbonate being stable at or near the Earth's surface. In 1960 it was demonstrated experimentally that the addition of water (H_2O) to the system CaO-CO_2 at high pressure will stabilize a carbonate liquid at temperatures expectable in the Earth's crust. Also in 1960, the

Fig. 2. Generalized distribution of carbonatites. Individual spots represent clusters of occurrences.

volcano Oldoinyo Lengai in Tanzania erupted carbonate-rich ash and subsequently carbonatite lava. Geologic mapping has confirmed that most carbonatites occur in small volumes associated with mantle-derived, high-alkali, low-silica igneous rocks. *See* CALCITE.

There are three possible origins (not mutually exclusive) for carbonatite magmas: (1) they are "primary" liquids, arriving directly from their mantle sources without change during ascent; (2) they are products of fractional crystallization from carbonate-bearing silicate magmas (if silicate minerals crystallize early from the magma, carbonate concentration should build up in the remaining liquid); or (3) they are products of immiscible separation of carbonate-rich and silicate-rich liquids (the properties of highly ionic carbonate liquid are so different from those of silicate liquid that, during cooling, carbonate-bearing silicate magma separates into two liquids, one containing nearly all the carbonate, the other nearly all the silicate).

The first origin is favored by evidence from altered mantle xenoliths (rock fragments enclosed by an unrelated igneous rock) carried to the surface by basalts, nephelinites, and kimberlites; these altered fragments demonstrate that carbonatite liquid has percolated through the mantle. A primary liquid origin is also supported by the presence in carbonatites, rarely, of magnesium- and chromium-rich spinels typical of mantle rocks, and by the absence, in some carbonatite occurrences, of associated silicate rocks that could represent parents or siblings of the carbonatites. The lack of mantle xenoliths in carbonatites is evidence against primary liquid, but the low density and low viscosity of carbonatite liquid may permit such dense fragments to settle out quickly. *See* SPINEL; XENOLITH.

The second and third origins are favored by the small volume of carbonatite relative to silicate rocks in most occurrences, and by the lateness of carbonatites in intrusive and eruptive sequences; but they are opposed by differences in isotopic ratios between carbonatite and silicate rocks in some occurrences (neither fractional crystallization nor immiscible liquid separation should change the isotopic ratios). The second origin is also opposed by experimental results indicating that silicates and carbonates crystallize simultaneously through a large temperature interval, preventing buildup of carbonate in the remaining liquid. The third origin is supported by experiments demonstrating that during cooling a carbonate-bearing silicate liquid can split into carbonate-rich and silicate-rich liquids, and by the natural occurrence of "rock emulsions" consisting of blobs of one composition enclosed in the other. On balance, the evidence is still inconclusive, and possibly any of the three mechanisms of origin can operate at different times and places.

Criteria for recognition. It is not easy to discriminate carbonatites from thoroughly recrystallized and metamorphosed marbles or from vein fillings deposited by dilute, low-temperature aqueous solutions. Because calcite has very low strength and easily deforms by plastic flow, carbonate-rich rocks can intrude their more brittle surroundings and thus imitate magmatic rocks. Furthermore, because marbles and recrystallized carbonatites converge toward the same textures and mineral assemblages, carbonatites can be misidentified as metamorphosed sedimentary rocks. The strongest indications of igneous origin for a carbonate-rich rock are considered to be quenched carbonate-rich liquid in the form of droplets or bubble walls, association with nepheline- or melilite- bearing rocks, presence of strontium-rich calcite, presence of pyrochlore and other minerals that are rich in niobium or rare-earth elements, and presence of apatite that has higher concentrations of rare-earth elements and silicon than apatite from other rock types. *See* APATITE; NIOBIUM; RARE-EARTH ELEMENTS.

Economic significance. Carbonatites yield a variety of mineral commodities, including phosphate, lime, niobium, rare-earth elements, anatase, fluorite, and copper. Agricultural phosphate for fertilizer is the most valuable single product from carbonatites; most is obtained from apatite in lateritic soils that have developed by tropical weathering of carbonatites, dissolving the carbonates and thereby concentrating the less soluble apatite. Lime for agriculture and for cement manufacture is obtained from carbonatites in regions where limestones are lacking.

The carbonatites at Bayan Obo, China, and Mountain Pass, California, dominate the world suppliers of rare-earth elements, but many other carbonatites contain unexploited reserves. Tropical weathering at several carbonatites in Brazil has produced economically important concentrations of anatase (TiO_2) from decomposition of perovskite ($CaTiO_3$). Other carbonatites are locally important sources for some commodities; for example, Amba Dongar, India, provides fluorite, and Phalaborwa, South Africa, is a major producer of copper, with zirconium, gold, silver, platinum group elements, and uranium as byproducts.

Tectonic significance. Ultramafic xenoliths from lithospheric mantle commonly show textures and mineral assemblages that indicate modification of the original rock. This alteration typically results in strong enrichment in light rare-earth elements, uranium, thorium, and lead, but much less enrichment in titanium, zirconium, niobium, and strontium. These changes are commonly attributed to interaction of lithospheric mantle with an invading carbonate-rich magma. The wide geographic dispersal of these altered xenoliths suggests that carbonate-rich liquid has been more common in the upper mantle than the low abundance of carbonatites in the upper crust would suggest. According to the testimony of these samples, carbonatite magma, ascending through lithospheric mantle, commonly is trapped before it can invade the crust. In addition to the factors that can stop the rise of any magma (heat loss, increase of solidus temperature with decrease in pressure, decrease in density and increase in strength of wall rock), carbonatite magma can be halted by reaction with wall rock to form calcium

and magnesium silicates plus CO_2, and by less oxidizing conditions to reduce carbonate to elemental carbon (graphite or diamond) or to methane. Both of these changes subtract dissolved CO_2 from the magma, causing crystallization. *See* DIAMOND; EARTH INTERIOR; GRAPHITE.

In some regions carbonatite magmatism has recurred after intervals of 100 million years, suggesting that carbonate-rich liquids can be regenerated by a small degree of partial fusion of locally and repeatedly carbonated mantle. This possibility, in turn, suggests that carbonate in the mantle is replenished by recycling of sedimentary carbonate. *See* SUBDUCTION ZONES.

Carbonatites are not restricted to a single tectonic regime. They occur in oceanic and continental crust and have formed in compressional fold belts and stable cratons as well as regions of crustal extension. Rather than indicating the stress field in the shallow crust in which they were emplaced, carbonatites are useful in modeling the long-term thermal and chemical development of the mantle. *See* IGNEOUS ROCKS.

Daniel S. Barker

Bibliography. K. Bell (ed.), *Carbonatites: Genesis and Evolution*, Unwin-Hyman, Boston, 1989; R. H. Mitchell (ed.), *Undersaturated Alkaline Rocks: Mineralogy, Petrogenesis, and Economic Potential*, Short Course Vol. 24, Mineralogical Association of Canada, Winnipeg, 1996; A. R. Woolley, *Alkaline Rocks and Carbonatites of the World*, pt. 1: *North and South America*, British Museum (Natural History), London, 1987.

Carboniferous

The fifth period of the Paleozoic Era. The Carboniferous Period spanned from about 355 million years to about 295 million years ago. The rocks that formed during this time interval are known as the Carboniferous System; they include a wide variety of sedimentary, igneous, and metamorphic rocks. Sedimentary rocks in the lower portion of the Carboniferous are typically carbonates, such as limestones and dolostones, and locally some evaporites. The upper portions of the system are usually composed of cyclically repeated successions of sandstones, coals, shales, and thin limestones. *See* SEDIMENTARY ROCKS.

The economic importance of the Carboniferous is evident in its name, which refers to coal, the important energy source that fueled the industrialization of northwestern Europe in the early 1800s and led to the Carboniferous being one of the first geologic systems to be studied in detail. Carboniferous coals formed in coastal and fluvial environments in many parts of the world. Petroleum, another important energy resource, accumulated in many Carboniferous marine carbonate sediments, particularly near shelf margins adjacent to basinal black shale source rocks. In many regions the cyclical history of deposition and exposure has enhanced the permeability and porosity of the shelfal rocks to make them excellent petroleum reservoirs. The limestones of the Lower Carboniferous are extensively quarried and used for building stone, especially in northwestern Europe and the central and eastern United States. *See* COAL; PETROLEUM.

The base of the Carboniferous is placed at the first appearance of the conodont *Siphonodella sulcata*, a fossil that marks a widely recognized biozone in most marine sedimentary rocks. The reference locality for this base is an outcrop in Belgium. The top of the Carboniferous is placed at the first appearance of the conodont *Streptognathus isolatus* a few meters below the first appearance of the Permian fusulinacean foraminiferal zone of *Sphaeroschwagerina fusiformis*. The reference locality is in the southern Ural region in Kazakhstan. The equivalent biozone is at the base of *Pseudoschwagerina* in North America. *See* CONODONT; FUSULINACEA.

Subdivisions. The term Carboniferous Order was originally applied by W. D. Conybeare and J. Phillips in 1822 to rocks in the British Isles that included the Old Red Sandstone, Mountain Limestone, Millstone Grit, and Coal Measures (**Fig. 1**). Later, after the Old Red Sandstone was recognized as being a continental facies of the Devonian System, the Mountain Limestone became the Lower Carboniferous and the Millstone Grit and Coal Measures were combined to be the Upper Carboniferous. In North America, the Coal Measures were readily recognized (as Upper Carboniferous), and the underlying, mainly noncoaly beds were initially called Subcarboniferous.

In 1891, the U.S. Geological Survey introduced the terms Pennsylvanian Series for the Coal Measures and Mississippian Series for the Subcarboniferous. In 1906 T. C. Chamberlain and R. D. Salisbury raised the Mississippian and Pennsylvanian to the rank of

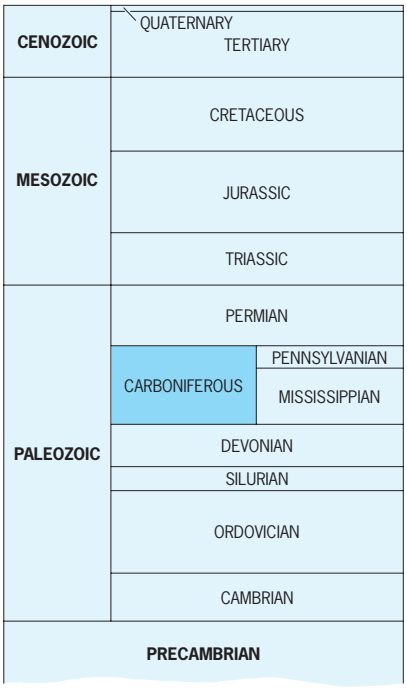

<table>
<tr><th colspan="2">Conybeare and Phillips, 1822 (Great Britain)</th><th colspan="2">Northwestern Europe, mid-1800s</th><th colspan="2">Chamberlain and Salisbury, 1906 (North America)</th><th colspan="2">Soviet Union, 1920s–1970s</th><th colspan="2">Current usage</th></tr>
<tr><td colspan="2">Magnesian Limestone</td><td colspan="2">Permian</td><td colspan="2">Permian</td><td colspan="2">Permian</td><td colspan="2">Permian</td></tr>
<tr><td rowspan="4">CARBONIFEROUS</td><td rowspan="2">Coal Measures</td><td rowspan="4">CARBONIFEROUS</td><td rowspan="3">Upper</td><td rowspan="2">Pennsylvanian System</td><td rowspan="4">CARBONIFEROUS</td><td>Upper</td><td rowspan="4">CARBONIFEROUS</td><td rowspan="2">Pennsylvanian Subsystem</td></tr>
<tr><td rowspan="2">Millstone Grit</td><td>Middle</td></tr>
<tr><td rowspan="2">Mountain Limestone</td><td rowspan="2">Mississippian System</td><td rowspan="2">Lower</td><td rowspan="2">Mississippian Subsystem</td></tr>
<tr><td>Lower</td></tr>
<tr><td colspan="2">Old Red Sandstone</td><td colspan="2">Devonian</td><td colspan="2">Devonian</td><td colspan="2">Devonian</td><td colspan="2">Devonian</td></tr>
</table>

Fig. 1. Development of stratigraphic nomenclature for the Carboniferous.

systems (Fig. 1), noting that they were separated by a major unconformity and were quite different lithologically in both North American and in Europe. This nomenclature was extensively adopted in North America but not in other parts of the world. *See* MISSISSIPPIAN; PENNSYLVANIAN.

A third means of subdividing the Carboniferous was developed during the 1920s and 1930s on the Russian Platform, the western slopes of the Urals, and the greater Donetz Basin areas of Ukrainia where a three part subdivision is recognized. On the platform, the limestone succession of the Lower Carboniferous is widely recognized and is separated by a major unconformity and a hiatus from the overlying Upper Carboniferous. On the platform margins and slopes and toward the center of the Donetz Basin, additional sediments progressively filled the hiatus, so that a complete sedimentary record is available for study in a marine facies. Within the upper portion of the Carboniferous, a second regional unconformity of considerable duration is documented on the Russian Platform between the Moscovian and Kasimovian stages (Fig. 1); and as a result, the Russian Carboniferous was subdivided into three series (Fig. 1). The boundary between the Middle and Upper Carboniferous on the Russian Platform is approximately the same as the boundary between the Middle and Upper Pennsylvanian in the North American midcontinent region.

The International Subcommission on Carboniferous Stratigraphy reached general agreement in the 1970s and 1980s that the Carboniferous would be divided into two parts: a Lower Carboniferous Mississippian Subsystem and an Upper Carboniferous Pennsylvanian Subsystem.

The two Carboniferous subsystems are subdivided into a number of series and stages (**Fig. 2**) that are variously identified in different parts of the world, based on biostratigraphic evidence using evolutionary successions in fossils or overlapping assemblage zones. Many fossil groups have been studied in order to establish consistent worldwide zonations; they have included foraminifers, corals, ammonoid cephalopods, brachiopods, bryozoans, sponges, bivalves, crinoids, radiolarians, conodonts, and plants. To a large extent the distribution of each of these fossil groups was determined by climate, temperature, and other environmental and ecological conditions, evolutionary adaptations, and paleobiogeographic opportunities for dispersal. As the Carboniferous progressed, changes in each of these factors caused biotic redistributions and dispersals and, at other times, restrictions in distribution and strongly provincial biotic associations. These biogeographic differences have made detailed correlations between some regions more difficult, and as a result, there remains strong preferences for regional series and stage names. *See* INDEX FOSSIL; STRATIGRAPHY.

Paleogeography and lithology. Perhaps the strongest of the many ecological factors that controlled biotic distributions were the paleogeographic changes within the Carboniferous that were brought about by the initial assembling of the supercontinent Pangaea and the associated mountain-building activities, which greatly modified climate, ocean currents, and seaways. In the Early Carboniferous (**Fig. 3*a***), a nearly continuous equatorial seaway permitted extensive tropical and subtropical carbonate sedimentation on the shelves and platforms in North America, northern and southern Europe, Kazakhstan, North and South China, and the northern shores of the protocontinent Gondwana (such as northern Africa).

The gradual collision of northern Gondwana

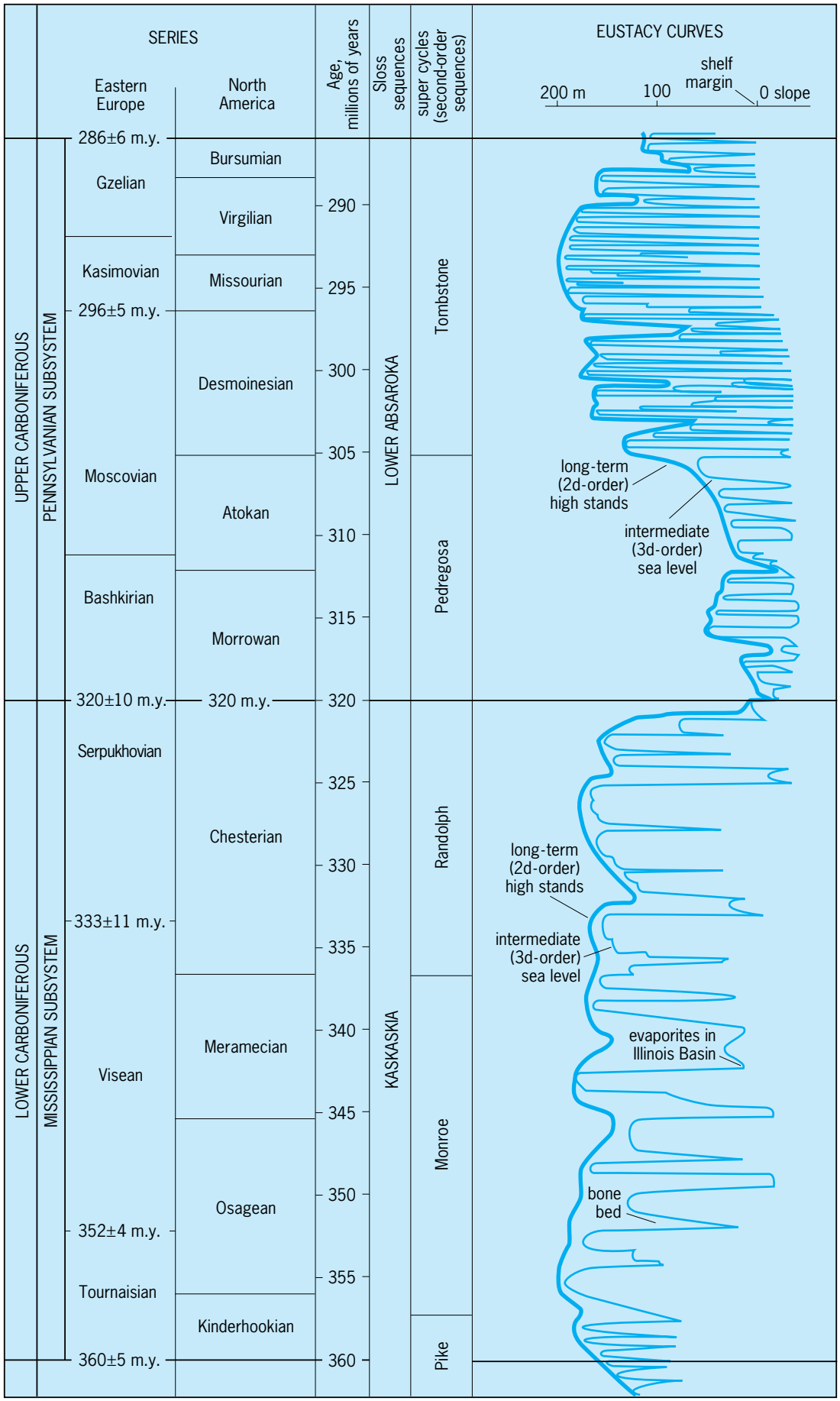

Fig. 2. Curves outlining the fluctuations of sea level during the Carboniferous. The Mississippian sea-level curve is typical of times when glaciation was not a major factor in controlling sea levels, and the Pennsylvanian sea-level curve is characteristic of extensive and protracted glaciation in the Early Pennsylvanian followed by repeated glaciation and nonglaciation events in the Middle and Upper Pennsylvanian. The time scale is estimated based on a few reliable radiometric ages for the interval. 1 m = 3.3 ft. (*After C. A. Ross and J. R. P. Ross, Late Paleozoic sea level and depositional sequences and biostratigraphic zonation of late Paleozoic depostional sequences, Cushman Found. Foraminiferal Res. Spec. Publ., no. 24, 1987*)

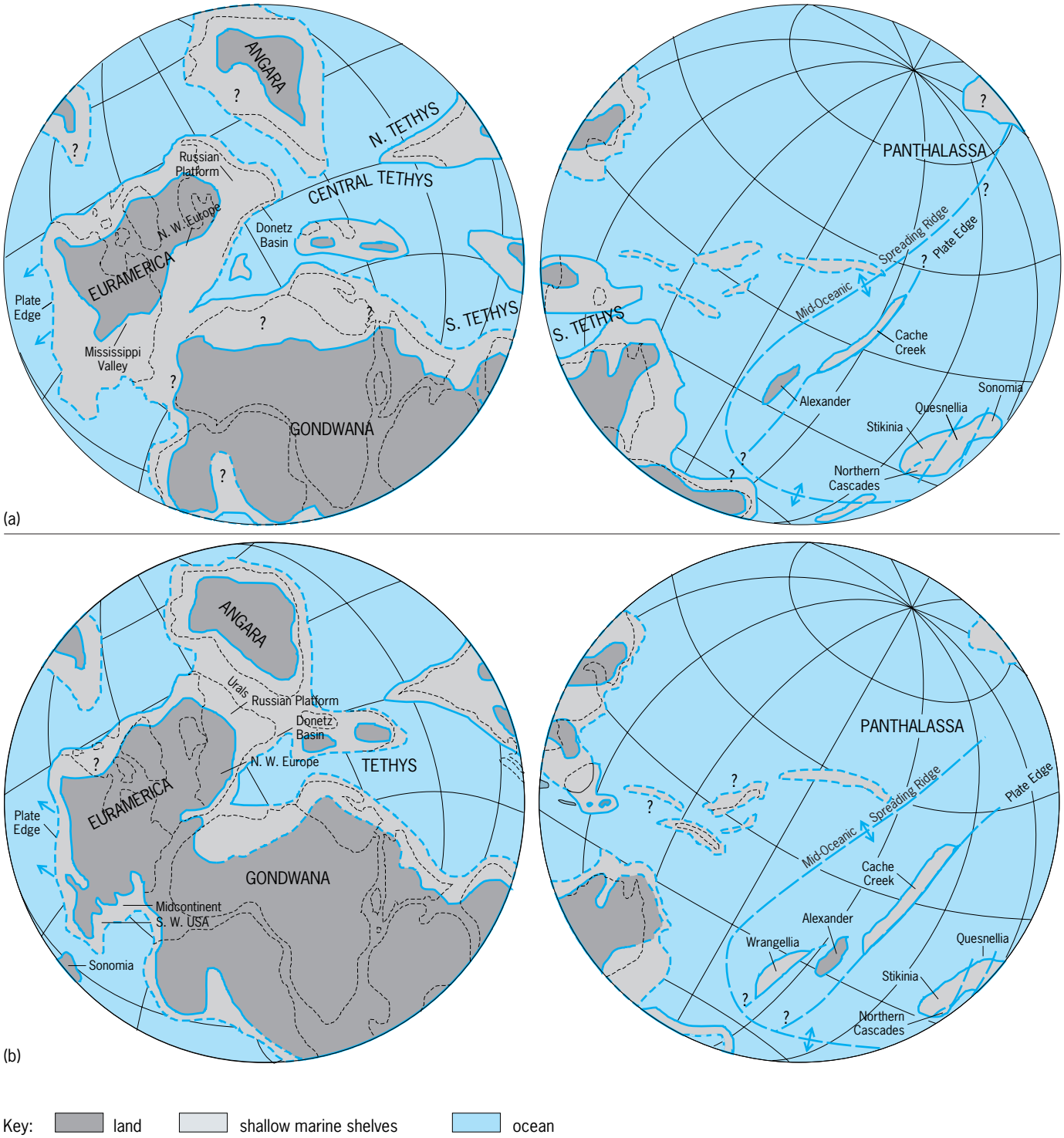

Fig. 3. Paleogeography of the Carboniferous. (*a*) Early Carboniferous (Mississippian) Tournaisian and Visean epochs, with a major east-west equatorial seaway between Euramerica and Gondwana. (*b*) Late Carboniferous (Pennsylvanian) Moscovian (or Desmoinesian) epochs with the east-west seaway closed by the collision of Euramerica and Gondwana and the resulting Ouchita-Appalachian-Hercynian orogeny and land barrier between the western Tethys and eastern Panthalassa oceans.

against northern Europe–North America (also called Euramerica or Laurussia) started the formation of the supercontinent of Pangaea. By the middle of the Carboniferous, this process divided the tropical equatorial seaway into two isolated segments (Fig. 3*b*). One, which represented the beginning of the Paleotethys faunal region, lay to the east of Pangaea as a large arcuate ocean with a number of small island arcs and oceanic plateaus. This warm-water ocean with vast west-flowing equatorial currents resulted in warm-water currents being directed both north and south of their normal latitudes.

The other segment of the former equatorial seaway lay to the west of Pangaea and faced a relatively large, deep Panthalassa Ocean basin, which had considerably fewer island arcs and ocean plateaus. This edge of Pangaea was not bathed by warm-water currents; it was the site of cold-water upwelling and cool-water currents from both the Northern and Southern hemispheres. When these reached equatorial areas they began to warm, and they formed the beginnings of the westward-flowing equatorial currents. As a consequence, the paleolatitudinal extent of carbonate platforms on the western margin of Pangaea was less than in the Paleotethys region. *See* CONTINENTAL DRIFT; CONTINENTS, EVOLUTION OF; PALEOGEOGRAPHY; PLATE TECTONICS.

Orogeny. The collision of Gondwana with North America–northern Europe, in addition to changing the patterns of equatorial ocean currents, created the long chain of mountains that made up the Ouachita-Appalachian-Hercynian orogenic belt. This complex set of mountains also is evident in the sedimentary patterns of clastic rocks, such as conglomerates, river and shoreline sandstones, and coals and other nonmarine clastic deposits, as well as in thick turbidite successions in several foredeep marine basins that formed along the front of the advancing orogenic thrust belts. Although this chain of mountains was mainly within the tropics and subtropics, it became very elevated, perhaps similar to the present-day Himalaya ranges; it disrupted terrestrial climates to a similar degree. *See* OROGENY.

Glaciation. An additional ramification of the formation of Pangaea was the beginning of very extensive glaciation in the Southern Hemisphere polar and high-latitude regions of the supercontinent. Glacial deposits are also known from smaller continental fragments that were at high paleolatitudes in the Northern Hemisphere. The cause for this great increase in ice accumulation is not entirely known. It may be that the diversion of warm-water currents into more southerly (and northerly) latitudes increased precipitation (as snow). Perhaps the mountain heights of the newly formed orogenic belts disrupted general atmospheric circulation from the tropics into high latitudes. Solar radiation possibly was reduced a few percent to initiate these more extensive glacial conditions. Or other, still unknown factors may be the primary causes. In any case, the Earth's climate cooled, tropical carbonate-producing areas became restricted toward the Equator, and eustatic sea-level fluctuations became prominent in the sedimentary record. *See* GLACIAL EPOCH.

Sea-level changes. The sedimentary record of sea-level changes is well documented from the beginning of the latest Mississippian through the Pennsylvanian and into the Permian. These are cyclical sediments in which the succession of depositional facies is regularly repeated by each sea-level fluctuation. As a result, these depositional sequences have great lateral extent and continuity. Because local depositional conditions depended on many other, often unique circumstances, the sedimentary patterns have considerable lateral variation (**Fig. 4**). For example, the

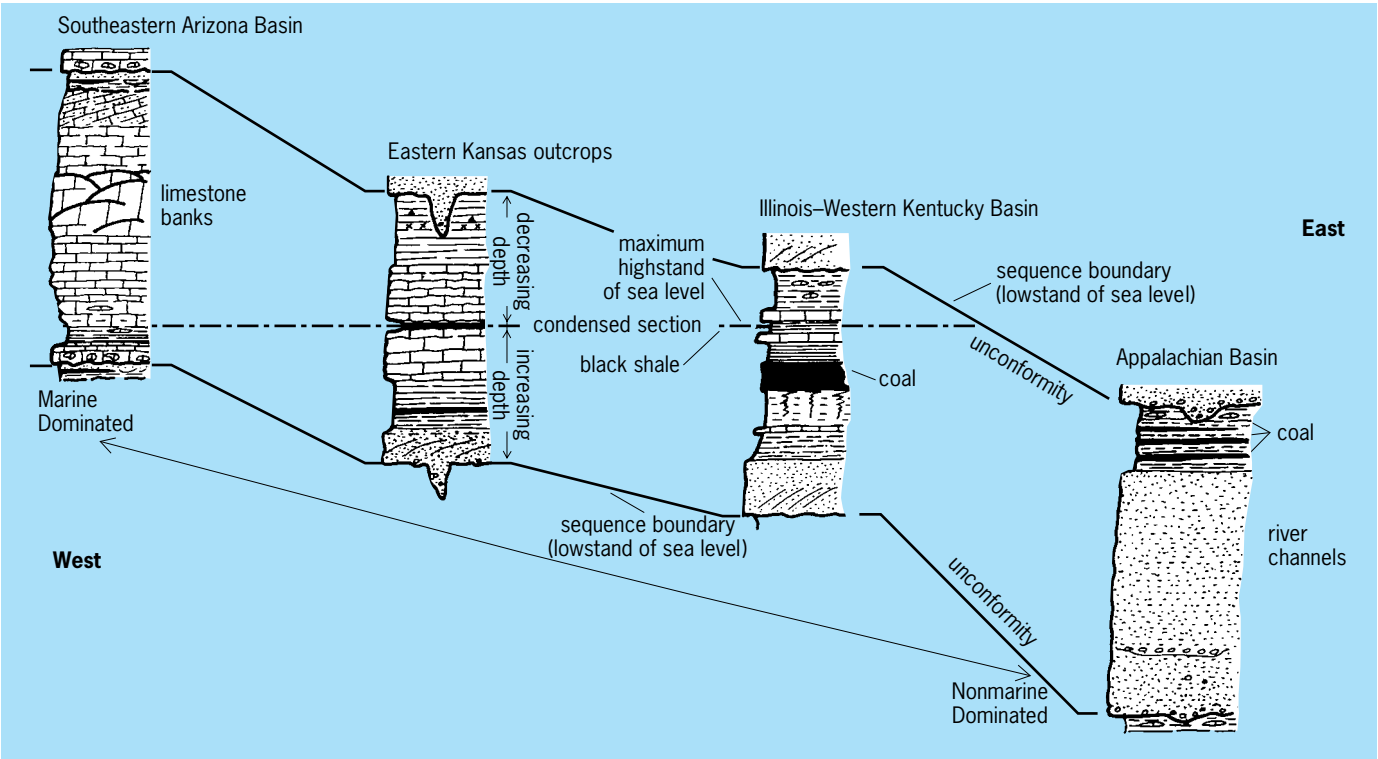

Fig. 4. Typical shelf sedimentary patterns in upper Carboniferous depositional sequences. The broken line represents the condensed section that forms during maximum sea-level highstands when sedimentation rates are low. The prominent unconformities at the base of depositional sequences in Appalachian Basin areas having well-developed fluvial beds become progressively less pronounced westward in marine-dominated successions.

cyclical sediments near the Appalachian orogenic belt contain vast amounts of clastic debris, such as sands, silts, and gravels, which were transported and deposited in rivers, lakes, deltas, and coastal features. Many of these cycles are nonmarine or only marginally marine, and many are viewed as being more strongly influenced by changes in tectonic activity and climate. The suggestion is that the amount of clastic material was supplied mainly by erosion, and sedimentation responded to wet and dry climatic cycles and only indirectly to sea-level fluctuations. In marine-dominated sequences, supratidal carbonates commonly passed upward into thin, fluvial and eolian beds that were usually reworked as debris into the succeeding basal marine transgressive beds. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS.

Farther west (Fig. 4) at greater distances from the Appalachian orogenic belt, as in Ohio, many of these clastic-rich cyclic sediments intertongue with marine sediments, and the influence of sea-level fluctuations becomes increasingly evident. The Illinois-Kentucky basin has about equal representation of marine deposits and fluvial channels and delta deposits. Western Missouri and eastern Kansas were far enough away from the large influx of clastic materials to have predominantly marine sediments and minor amounts of nonmarine, mostly silt and fine sand clastics. Carbonate sedimentation dominated farther from sources of clastic sediments. Oceanic plateaus in the Paleotethys and Panthalassa ocean basins, which lacked influxes of clastic materials, have cyclical sediments that are entirely in various carbonate facies that reflect differences in water depth.

The fluctuations in sea level during the Carboniferous reflect different rates of sea-level change, different durations, and different magnitudes. Minor sea-level fluctuations of a meter or two (3–6 ft) with cyclicities of 20,000–40,000 years and intermediate fluctuations of 4–6 m (13–20 ft) with cyclicities of about 100,000 years may be preserved. More easily recognized are fluctuations of about 400,000 years–1 million years, which have amplitudes of 60–200 m (200–660 ft) [Fig. 2]. Late in the Mississippian and early in the Pennsylvanian, for an interval of approximately 15–20 million years, sea-level high-stands were consistently low and infrequently reached higher than the continental shelf margins (Fig. 4), a condition that was not repeated until quite late in the Permian Period. *See* CONTINENTAL MARGIN.

The sea-level fluctuations resulted in large areas of the cratonic shelves being frequently exposed to long intervals of weathering, erosion, and diagenesis. The effects of these exposure-related events suggest that within the Carboniferous climates were not uniform and different types of paleosols (fossil soil profiles) formed under a variety of pedogenic processes. These paleosols reflect changes in world climate as well as the continued northward motion of continents across latitudinal climatic belts as a result of plate tectonics and sea-floor spreading. *See* PALEOCLIMATOLOGY; SEDIMENTOLOGY; WEATHERING PROCESSES.

Life. During the Carboniferous, life evolved to exploit fully the numerous marine and nonmarine aquatic environments and terrestrial and aerial habitats. Single-cell protozoan foraminifers evolved new abilities to construct layered, calcareous walls and internally complex tests. These single-celled organisms diversified into nearly all shallow-water, carbonate-producing environments from the intertidal, lagoon, shelf, shelf margin, and upper parts (within the photic zone) of the shelf slope. Ammonoid cephalopods, from their first appearance in the later part of the Devonian, diversified rapidly during the Early Carboniferous, and their occurrences were used to establish one of the earliest biostratigraphic zonation schemes for the Lower Carboniferous. Brachiopods, bryozoans, conodonts, crinoids, and corals were also widespread and locally important parts of the marine faunas. The Tournaisian and Visean shallow tropical seas abounded in blastoid echinoderms; however, these creatures became nearly extinct and were geographically restricted to a small area in the eastern Panthalassa Ocean after the Early Pennsylvanian (Morrowan). Conodonts were widespread in deeper-water deposits.

Insects have remarkable evolutionary histories during the Carboniferous. They adapted to flight and dispersed into many terrestrial and fresh-water habitats. Carboniferous insects include many unusual orders, such as one with three pairs of wings (although it is not clear that the foremost pair functioned directly in propulsion). By the Late Pennsylvanian, many large cockroachlike orders were present and also several huge dragonflylike groups, some of which reached wingspans of 90 cm (3 ft). Carboniferous insects included representatives of five Paleoptera orders (mayflies and dragonflylike orders) and at least six orthopterid (cockroachlike) orders. Insects are commonly preserved in coal swamp deposits and display the amazing diversity of life within these swamps. *See* INSECTA.

Vertebrates also evolved rapidly. Although acanthodian fish declined from their Devonian peak, sharklike fishes and primitive bony fishes adapted well to the expanded environments and the new ecological food chains of the Carboniferous. Some sharklike groups invaded fresh-water habitats, where they were associated with coal swamp deposits. For the bony fish (ray-finned fish and air-breathing choanate fishes with lobed fins), the Late Devonian and Carboniferous was a time of considerable evolutionary diversity and ecological expansion, with many lineages independently adapting to both fresh-water and marine conditions. Lungfish are one of the fresh-water choanate lineages which evolved adaptations for survival in temporarily dry lakes and rivers. Near the Devonian-Carboniferous boundary, a lineage from the choanate fish had evolved into the first amphibians. *See* VERTEBRATA.

Carboniferous amphibians evolved rapidly in several directions. The earliest were the labyrinthodont

embolomeres, which had labyrinthodont teeth and were mainly aquatic. Another significant labyrinthodont group was the rhachitomes, which originated in the Early Carboniferous and became abundant, commonly reaching about 1 m (3 ft) or more; they were widespread in terrestrial habitats during the Late Carboniferous and Permian. Some of the Carboniferous amphibians reverted to totally aquatic habitats, such as the lepospondyls, which lost their bony vertebrae and limbs, had large flattened broad heads, and were snakelike in appearance. Ancestors of the present-day anurans (frogs), urodeles (salamanders), and caecilians (apodans) probably date from the later part of the Carboniferous, but their record as fossils is meager. *See* AMPHIBIA.

Primitive reptiles evolved from one of the embolomere amphibian lineages during the Late Carboniferous. They formed the basal stock from which all other reptiles have evolved including the earliest mammallike reptiles in the Late Carboniferous. During the Late Carboniferous, early reptiles coexisted with several advanced amphibian groups which shared at least some, but probably not all, of their reptilelike characters. *See* REPTILIA.

Terrestrial plants also showed major diversification of habitats and the evolution of important new lineages during the Carboniferous. Initially, Early Carboniferous plants were predominantly a continuation of latest Devonian groups; however, they were distinguished in part by their large sizes with many arborescent lycopods and large articulates, and pteridosperms (seed ferns) and ferns became increasingly abundant and varied. By the Late Carboniferous, extensive swamps formed along the broad, nearly flat coastal areas; and these coal-forming environments tended to move laterally across the coastal plain areas as the sea level repeatedly rose. Other coal-forming marshes were common in the floodplains and channel fills of the broad rivers of upper delta distributary systems. During the Late Carboniferous, primitive conifers appeared and included araucarias, which became common in some, probably drier ecological habitats. One of the features of Late Carboniferous plant paleogeographic distributions is the recognition of a southern, high-latitude Gondwanan floral province, the *Glossopteris* province, and a northern high-latitude Anagaran floral province, which were cool adapted. These provinces contrasted with an extensive equatorial belt of much greater plant diversity. *See* PALEOBOTANY; PALEOZOIC.

C. A. Ross; June R. P. Ross

Bibliography. R. H. Dott, Jr., and D. R. Prothero, *Evolution of the Earth*, 5th ed., 1994; H. L. Levin, *The Earth Through Time*, 4th ed., 1994; A. L. Palmer (general ed.), Decade of North American Geology (DNAG) Project, *The Geology of North America*, Geological Society of America, 1986–1994; R. C. Moore et al., The Kansas Rock Column, *State Geol. Surv. Kans. Bull.*, no. 89, 1951; C. A. Ross and J. R. P. Ross, Late Paleozoic Sea Levels and Depositional Sequences and Biostratigraphic Zonation of Late Paleozoic Depositional Sequences, *Cushman Found. Foram. Res. Spec. Publ.*, no. 24, 1987; J. W. Skehan et al., The Mississippian and Pennsylvanian (Carboniferous) Systems in the United States, *USGS Prof. Pap.*, no. 1110-A–DD, pp. A1–DD16, 1979; R. H. Wagner, C. F. Winkler Prins, and L. F. Granados (eds.), *The Carboniferous of the World* (M. C. Diaz, general ed.), Instituto Geologico y Minero de Espana, Madrid, and Nationaal Natuurhistorisch Museum, Leiden, Parts I, II, III, IUGS Publ. 16, 20, 33, 1983, 1985, 1996; H. R. Wanless and C. R. Wright, Paleoenvironmental Maps of Pennsylvanian Rocks, Illinois Basin and Northern Midcontinent Region, *Geol. Soc. Amer. Publ.*, no. MC-23, 1978.

Carbonyl

A functional group found in organic compounds in which a carbon atom is doubly bonded to an oxygen atom

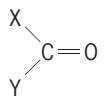

Depending upon the nature of the other groups attached to carbon, the most common compounds containing the carbonyl group are aldehydes (X and Y = H; X = H, Y = alkyl or aryl), ketones (X and Y = alkyl or aryl), carboxylic acids (X = OH, Y = H, alkyl, or aryl), esters (X = O-alkyl or aryl; Y = H, alkyl, or aryl), and amides (X = N—H, N-alkyl, or N-aryl; Y = H, alkyl, or aryl). Other compounds that contain the carbonyl group are acid halides, acid anhydrides, lactones, and lactams. *See* ACID ANHYDRIDE; ACID HALIDE; ALDEHYDE; AMIDE; ESTER; KETONE; LACTONE.

The direct introduction of the carbonyl group (carbonylation) has been accomplished by the reaction of alkenes with a mixture of carbon monoxide and hydrogen (synthesis gas) in the presence of metal carbonyls, of which the cobalt derivatives are the most important. This industrially important process, known as the oxo process, is shown in the reaction below. It is a principal commercial source of straight-chain aldehydes.

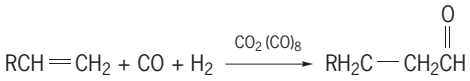

Another mode of direct introduction is by the action of Grignard reagents with carbon dioxide, which produces carboxylic acids. Formally, this reaction is analogous to the fixation of carbon dioxide in plants to produce glucose and fructose and ultimately starch and cellulose. *See* CARBOXYLIC ACID; GRIGNARD REACTION.

Indirectly, the carbonyl group can be introduced by the oxidation of primary and secondary alcohols, the former a source of aldehydes and carboxylic

acids, and the latter of ketones. The hydrolysis of organic cyanides is another source of carboxylic acids.

The presence of a carbonyl group in a molecule often leads to highly desirable properties of taste and odor. Thus a large number of flavoring and perfume components are carbonyl compounds. Familiar examples are oil of spearmint, oil of peppermint, oil of cloves, oil of vanilla, and camphor. *See* ESSENTIAL OILS.

Most of the important materials of nature, including fats and oils, steroids, proteins, many carbohydrates, and other such familiar compounds as caffeine, aspirin, and pencillin, contain the carbonyl group. *See* ASPIRIN; CAFFEINE; CARBOHYDRATE; FAT AND OIL; PROTEIN; STEROID.

Synthetic fibers such as nylon and Dacron, acrylic plastics and latexes, and many adhesives and paper finishes are polymeric materials containing carbonyl groups, principally of the amide and ester type. *See* MANUFACTURED FIBER; POLYMER.

All the compounds containing this functional group are referred to in a general way as carbonyl compounds. It is important, however, to distinguish these compounds from a large group formed from metals and carbon monoxide, which are known as metal carbonyls. In these latter compounds, there is only one group attached to the carbon in addition to the oxygen, and the carbon atom is viewed as triply bonded to the oxygen. *See* METAL CARBONYL.

Jeremiah P. Freeman

Bibliography. R. T. Morrison and R. N. Boyd, *Organic Chemistry*, 7th ed., 2000; S. E. Patai and Z. Rappoport, *Chemistry of Enones and Related Compounds*, 1989; M. B. Smith and J. March, *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*, 5th ed., 2001.

Carborane

A cluster compound containing both carbon (C) and boron (B) atoms as well as hydrogen (H) atoms external to the framework of the cluster. A cluster compound is one with insufficient electrons to allow for classical two-center two-electron bonds between all adjacent atoms. Sometimes the term carborane is used as a synonym for *closo*-1,2-$C_2B_{10}H_{12}$, commonly referred to as *ortho*-carborane. Carboranes are of interest because of their nonclassical bonding, their relatively high thermal stability, and their ability, when containing the ^{10}B isotope, to capture neutrons efficiently. *See* BORANE.

Structure. The structures of carboranes are based upon a series of three-dimensional, cagelike geometric shapes possessing triangulated faces; such shapes are termed delta polyhedra. The structure for any given carborane may be predicted by determining the framework electrons, by determining the number of electrons involved in bonding the boron and carbon atoms of the cluster framework together, and by using Wade's rule. Wade's rule states that a cluster containing n framework electrons will be derived from a delta polyhedron containing $(n-2)/2$ vertices, the parent cluster. Once this parent cluster has been determined, the geometry of the cluster framework may be predicted by clipping off vertices from the parent cluster until a polyhedron whose number of vertices is equivalent to the sum of boron and carbon atoms in the cluster framework is obtained.

Carboranes are placed, according to their structure, into several classifications. The most common classifications are closo (closed), nido (nestlike), and arachno (cobweb). If a carborane's framework structure is that of a closed delta polyhedron, the carborane is said to be a *closo*-carborane. If a carborane's framework structure is that of a closed delta polyhedron minus one or two vertices, the carborane is said to be a *nido*- or *arachno*-carborane, respectively.

The carborane *nido*-$C_2B_4H_8$ can be used as an example for predicting the structure of a carborane. The first step is to sum all the electrons for the compound: the eight valence electrons of the two carbon atoms (four each), the 12 valence electrons of the four boron atoms (three each), and the eight valence electrons of the eight hydrogen atoms (one each). There is then a total of 28 valence electrons contained in *nido*-$C_2B_4H_8$. In the second step, the number of framework electrons is calculated, assuming that each carbon atom and each boron atom contains one bond to a hydrogen atom that is directed away from the center of the cluster, that is, a terminal hydrogen atom. For *nido*-$C_2B_4H_8$, six terminal hydrogen atoms are assumed. Then, since terminal hydrogen atoms are not involved in holding the cluster framework together, for each terminal hydrogen two electrons (the number of electrons involved in the bond between carbon or boron and the terminal hydrogen) are subtracted from the total number of valence electrons to obtain the number of cluster framework electrons—16 framework electrons for *nido*-$C_2B_4H_8$. Third, the number of vertices in the parent cluster can then be calculated by using Wade's formula, where $n = 16$, and the number of vertices $= (16 - 2)/2 = 7$. Therefore, the structure of *nido*-$C_2B_4H_8$ is derived from a seven-vertex delta polyhedron (**Fig. 1**: seven vertices, closo column). However, since the sum of carbon and boron atoms in the cluster framework is only six, one of the vertices of the seven-vertex delta polyhedron must be removed. The result is a framework structure containing a plane of five atoms capped by one atom, a pentagonal pyramid (Fig. 1: six vertices, nido column). The nonterminal hydrogen atoms are then distributed as bridging hydrogen atoms between boron atoms about the open face that resulted from the removal of the vertex from the parent cluster. The structures of other carboranes may be predicted in a similar manner.

Isomers. Since the framework structures of carboranes are based upon multivertex polyhedra, it might be expected that isomers exist in which the arrangement of the carbon and boron atoms within the cluster framework varies. Indeed, this type of isomerism is commonly observed. For carboranes containing two carbon atoms in the framework structure,

dicarbon carboranes, the isomer with adjacent carbon atoms is frequently the one initially synthesized, although it is rarely, if ever, the thermodynamically preferred isomer. Typically, the thermodynamically preferred isomer in dicarbon carboranes is the one in which the carbon atoms have moved as far apart within the cluster as possible. Three isomers of *closo*-$C_2B_{10}H_{12}$ result from moving the carbons farther apart from one another (**Fig. 2**). *See* MOLECULAR ISOMERISM.

Numbering. The rules for numbering carboranes can be quite complicated; however, following is a brief explanation of the numbering employed here. The numbers directly preceding the carbon atoms in a formula indicate the numbers assigned to those carbon atoms. For example, in the formula *closo*-1,2-$C_2B_{10}H_{12}$ the 1 and 2 indicate that the carbon atoms have been assigned as atoms 1 and 2. A functional group immediately following a number in a formula indicates that the functional group is bound to the atom assigned that number. For example, in the formula *closo*-1-R-2-R′-1,2-$C_2B_{10}H_{10}$ the 1 preceding the functional group R indicates that it is bound to carbon 1, while the 2 preceding the functional group R′ indicates that it is bound to carbon 2. *See* STRUCTURAL CHEMISTRY.

Bonding. The bonding within a carborane can be thought of in terms of localized atomic orbitals forming both classical two-center two-electron bonds and nonclassical three-center two-electron bonds. Each vertex boron and carbon atom can be thought of as being sp^3 hybridized, with three of these hybrid orbitals of each vertex atom employed in framework bonding (**Fig. 3**). Employing this simple approach, the bonding within carboranes can be represented by employing resonance structures. For example, *nido*-$C_2B_4H_8$ exhibits three resonance structures (**Fig. 4**). The existence of resonance structures implies a delocalization of electron density throughout the cluster framework. Indeed, carboranes exhibit a high degree of electron density delocalization, and their bonding can be described more accurately by employing molecular orbital theory. *See* CHEMICAL BONDING; DELOCALIZATION; MOLECULAR ORBITAL THEORY; RESONANCE (MOLECULAR STRUCTURE).

Preparation. The typical synthesis of a carborane involves the reaction of a boron hydride cluster, containing only boron and hydrogen, with an alkyne. The resulting carborane contains two carbon atoms in its skeletal structure, a dicarbon carborane. The dicarbon carboranes, because of their relative ease of preparation, have been the most widely studied group of carboranes. In particular, *closo*-1,2-$C_2B_{10}H_{12}$, the most readily available carborane, has been extensively studied. Other common groups of carboranes include the monocarbon and tetracarbon carboranes. In order to illustrate the general synthetic routes to carboranes, the synthesis of several common mono-, di-, and tetracarbon carboranes are described below. *See* ALKYNE.

Monocarbon carboranes. There are two common routes to the monocarbon carboranes. One involves the incorporation of the carbon atom of $[B_{10}H_{13}CN]^{2-}$ into the skeletal structure [reaction (1)], and the other involves the degradation of [*nido*-7,9-$C_2B_{10}H_{13}]^-$ [reactions (2) and (3)].

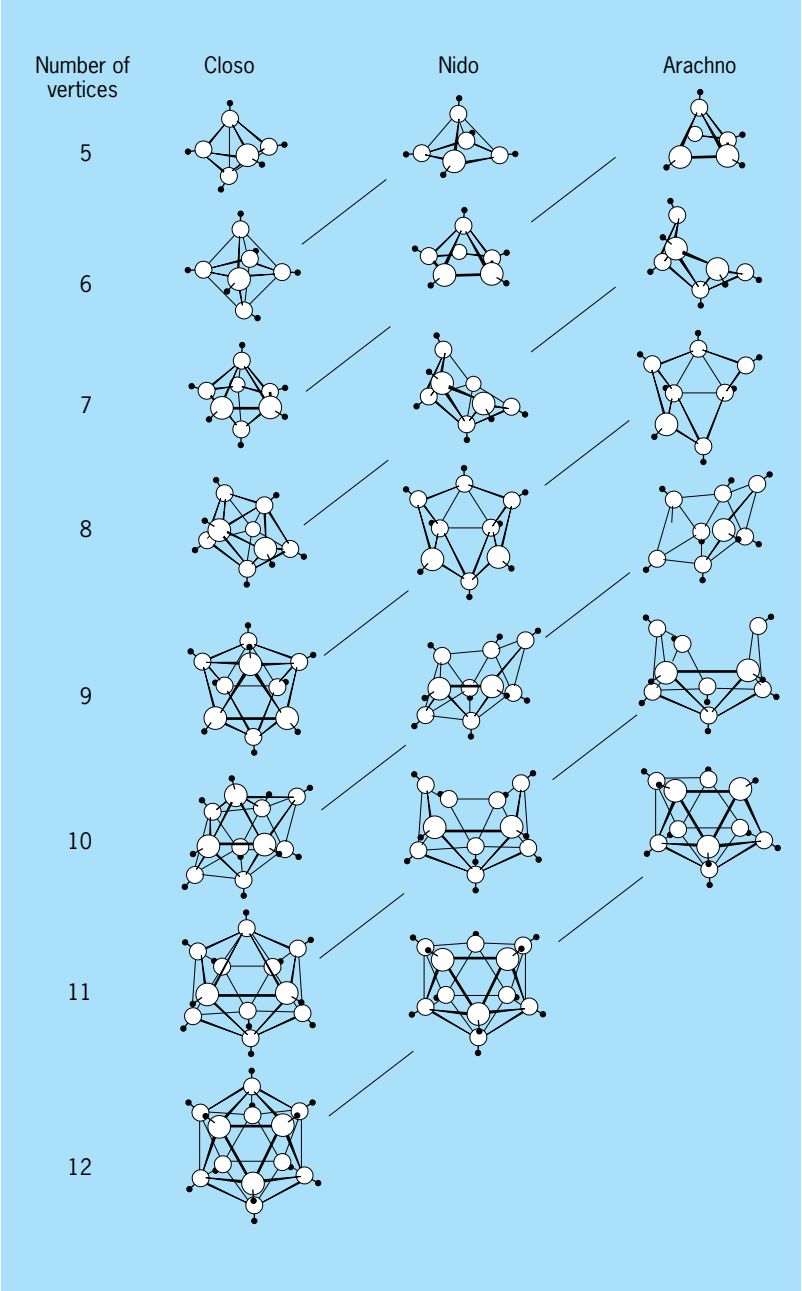

Fig. 1. Parent clusters from which carborane structures are determined. The diagonal lines define series of related closo, nido, and arachno structures.

$$B_{10}H_{14} + 2NaCN \xrightarrow{-HCN} Na_2B_{10}H_{13}CN \xrightarrow{H^+}$$
$$B_{10}C(NH_3)H_{12} \xrightarrow{(CH_3)_2SO_4} \xrightarrow[2.\ H_2O]{1.\ Na} [CB_{10}H_{13}]^- \quad (1)$$

$$[C_2B_{10}H_{13}]^- \xrightarrow[N(CH_3)_3]{C_4H_8O_{(aq)}}$$
$$9\text{-}CH_3\text{-}8\text{-}(CH_3)_3N\text{-}\textit{arachno}\text{-}6\text{-}CB_9H_{12} \xrightarrow[K_2CO_3]{(CH_3)_2CO}$$
$$9\text{-}CH_3\text{-}8\text{-}(CH_3)_3N\text{-}\textit{nido}\text{-}6\text{-}CB_9H_{10} \quad (2)$$

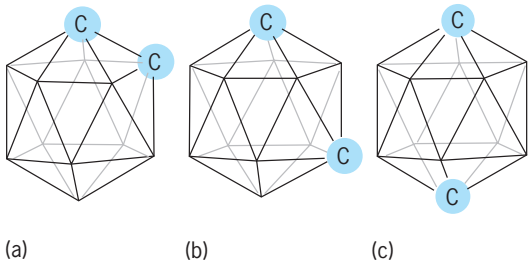

Fig. 2. Skeletal structures of the (*a*) 1,2-, (*b*) 1,7-, and (*c*) 1,12-isomers of $C_2B_{10}H_{12}$.

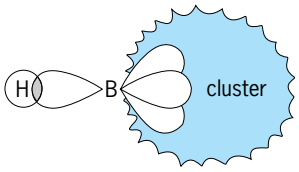

Fig. 3. Vertex boron atom with the three sp^3 orbitals that it employs in cluster bonding. The BH unit contributes three orbitals and only two electrons to the bonding that holds the cluster together (framework bonding).

$$[C_2B_{10}H_{13}]^- \xrightarrow[(CH_3)_2S]{H_3O^+} \xrightarrow{[BH_4]^-} [CB_9H_{12}]^- \qquad (3)$$

Dicarbon carboranes. The most common carboranes are the 12-vertex carboranes, *closo*-1,2-$C_2B_{10}H_{12}$ and its carbon-substituted derivatives. They are readily prepared by the reaction of acteylene or substituted alkynes with compounds of the form $B_{10}H_{12}L_2$ [where L = CH_3CN or $(CH_3CH_2)_2S$]. Such adducts are readily obtained by the reaction of the appropriate Lewis base with decaborane (14), $B_{10}H_{14}$ [reaction (4)]. Yields of the *closo*-12-vertex carboranes

$$B_{10}H_{14} + 2L \xrightarrow{-H_2} B_{10}H_{12}L_2 \xrightarrow{RC{\equiv}CR'} \text{1-R-2-}R'\text{-1,2-}C_2B_{10}H_{10} \qquad (4)$$

depend upon the substituents R and R′ of the alkyne. With R=R′=H, the yield is typically about 60%. The substituents R and R′ can be any number of functional groups, including acetoxy, alkoxy, alkyl, aryl, carboranyl, dimethylaminomethyl, halomethyl, and olefin. In all these syntheses, the 1,2-isomers of the corresponding carboranes are obtained.

Another readily prepared class of dicarbon carboranes are the 10-vertex carboranes, *nido*-$C_2B_8H_{12}$ and *closo*-$C_2B_8H_{10}$ and their carbon-substituted derivatives. A convenient method for the preparation of carbon-substituted derivatives, *nido*-5,6-R,R′-5,6-$C_2B_8H_{10}$, is the reaction of the appropriately substituted alkyne with $B_9H_{13} \cdot O(C_2H_5)_2$ as in reaction (5). The reaction proceeds with the

$$B_9H_{13} \cdot O(C_2H_5)_2 \xrightarrow[-B(CR{=}CR'H)_3]{4\,RC{\equiv}CR'} \textit{nido}\text{-5,6-R,}R'\text{-5,6-}C_2B_8H_{10} \qquad (5)$$

displacement of the $(C_2H_5)_2O$ and the abstraction (removal) of a BH_3 unit, by excess alkyne, from $B_9H_{13} \cdot O(C_2H_5)_2$. The unsubstituted *nido*-5,6-$C_2B_8H_{12}$ is most readily prepared by the aqueous iron(III) chloride oxidation of [*nido*-7,8-$C_2B_9H_{12}]^-$. The *nido*-5,6-$C_2B_8H_{12}$ obtained by this procedure can then, if desired, be converted to *closo*-1,10-$C_2B_8H_{10}$ by vacuum pyrolysis at 525°C (977°F).

The *nido*-six-vertex carboranes, *nido*-2,3-R_2-2,3-$C_2B_4H_6$, further serve to exemplify the preparation of dicarbon carboranes. This class of carboranes is readily prepared by the reaction of pentaborane with triethylamine and the appropriate alkyne. Carboranes with the functional groups CH_3, C_2H_5, C_3H_7, or $Si(CH_3)_3$ are readily prepared via this method.

Tetracarbon carboranes. The tetracarbon carboranes are unusual in that their structures are not that of a regular delta polyhedron. Instead they are 12-vertex clusters that possess two square and 16 triangular faces (**Fig. 5**). The carboranes 2,3,7,8-R_4-2,3,7,8-$C_4B_8H_8$ are prepared by the oxidative fusion of two 2,3-R_2-2,3-$C_2B_4H_4$ fragments from the metallacarboranes $[R_2C_2B_4H_4]_2FeH_2$, where R = CH_3, C_2H_5, C_3H_7. These tetracarbon carboranes have cluster framework structures (Fig. 5).

Chemistry. The chemistry of *closo*-1,2-$C_2B_{10}H_{12}$ and carboranes derived from it provides a reasonable representation of the entire class of carboranes. The most commonly exploited reaction of *closo*-1,2-$C_2B_{10}H_{12}$ is the deprotonation (removal of hydrogen) of its carbon atoms by a lithium alkyl group. Depending on the stoichiometry of the lithium alkyl employed, either one or both of the carbon atoms may

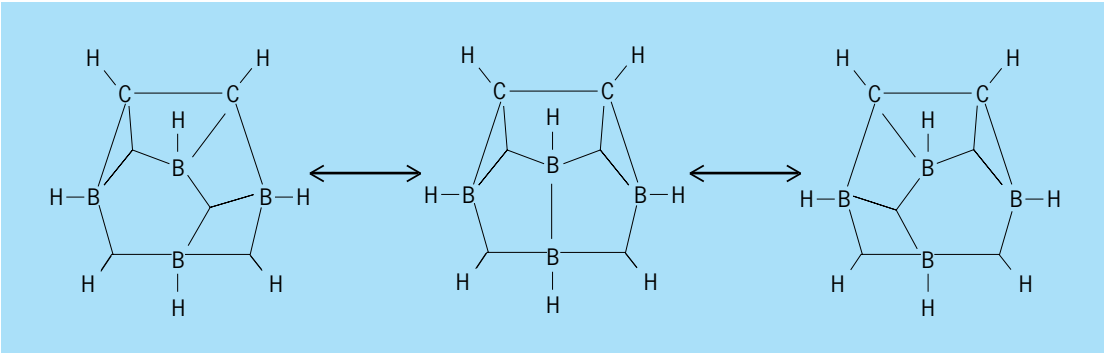

Fig. 4. Three resonance structures of *nido*-$C_2B_4H_8$.

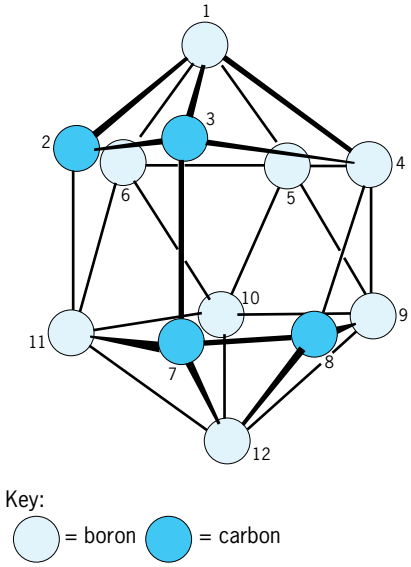

Fig. 5. Skeletal structures of the tetracarbon carboranes, $R_4C_4B_8H_8$.

be deprotonated [reactions (6) and (7), where BuLi is

$$1,2\text{-}C_2B_{10}H_{12} + \text{BuLi} \longrightarrow 1\text{-Li-}1,2\text{-}C_2B_{10}H_{11} + \text{butane} \quad (6)$$

$$1,2\text{-}C_2B_{10}H_{12} + \text{BuLi} \longrightarrow 1,2\text{-Li}_2\text{-}1,2\text{-}C_2B_{10}H_{10} + 2\,\text{butane} \quad (7)$$

the lithium alkyl]. The resulting lithiated carboranes can then be employed as nucleophiles in a variety of organic reactions.

Closo-1,2-$C_2B_{10}H_{12}$ is thermally robust, that is, not easily degraded by heat, and is isomerized to *closo*-1,7-$C_2B_{10}H_{12}$ upon heating to 450°C (842°F) under an atmosphere of nitrogen or argon. Even upon heating to 500°C (932°F) under an atmosphere of nitrogen or argon, no decomposition is observed. However, upon heating to 600°C (1112°F), significant decomposition occurs and rearrangement to *closo*-1,12-$C_2B_{10}H_{12}$ is observed, the *closo*-1,12-$C_2B_{10}H_{12}$ being obtained in low yield. These 12-vertex carboranes are very resistant to degradation by acid. However, both *closo*-1,2-$C_2B_{10}H_{12}$ and *closo*-1,7-$C_2B_{10}H_{12}$ are degraded by strong bases such as methoxide, resulting in the removal of a BH vertex that is adjacent to both carbon atoms. This base degradation results in the formation of the isomeric [*nido*-7,8-$C_2B_9H_{12}]^-$ and [*nido*-7,9-$C_2B_9H_{12}]^-$ anions from the 1,2- and 1,7-isomers, respectively (**Fig. 6***a*,*b*).

The two-electron reduction of a *closo*-carborane results in an opening of the cluster framework and the formation of a *nido*-carborane. This type of reductive opening of a cluster is observed when *closo*-1,2-$C_2B_{10}H_{12}$ reacts with sodium metal in solution to form [*nido*-$C_2B_{10}H_{12}]^{2-}$ [Fig. 6*c*].

One unique feature of carboranes is the ability of those with open faces to incorporate a wide variety of atomic groups into vertex positions. Anions of such carboranes are readily capped by a variety of atomic groups representing many of the elements of the periodic table. Additionally, an atom may be sandwiched between two *nido*-carborane anions; such structures are known as sandwich compounds (Fig. 6*d*). Thomas D. Getman

Bibliography. J. Casanova, *The Borane, Carborane, Carbocation Continuum*, 1998; R. N. Grimes, *Carboranes*, 1970; J. Plesek, Potential applications of the boron cluster compounds, *Chem. Rev.*, 92:269–278, 1992; B. Stibr, Carboranes other than $C_2B_{10}H_{12}$, *Chem. Rev.*, 92:225–250, 1992; R. E. Williams, The polyborane, carborane, carbocation continuum: Architectural patterns, *Chem. Rev.*, 92:177–207, 1992.

Carboxylic acid

One of a large family of organic substances widely distributed in nature, and characterized by the presence of one or more carboxyl groups (—COOH). These groups typically yield protons in aqueous solution. *See* ACID AND BASE.

In the type formula, $R(CXY)_nCOOH$, symbols R, X, and Y can be hydrogen, saturated, or unsaturated groups, carboxyl, alicyclic, or aromatic groups, halogens, or other substituents, and *n* may vary from zero (formic acid, HCOOH) to more than 100, provided that the normal carbon covalence of four is maintained.

Nomenclature and general properties. According to the International Union of Pure and Applied Chemistry, an acid is named by its relation to the parent hydrocarbon, but in common practice trivial names are used often to indicate the origin of the substance. Thus HCOOH, formic acid (from Latin *formica*, ant), is properly methanoic acid, and $CH_3(CH_2)_7COOH$, pelargonic acid, from the leaves of *Pelargonium roseum*, is nonanoic acid. Substituents are located numerically on the carbon chain, counting the carboxylic carbon as 1, but an older, common method uses Greek letters, beginning with the carbon

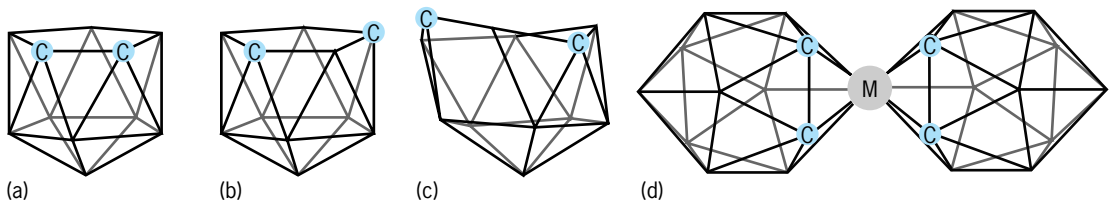

Fig. 6. Skeletal structures. (*a*) [*nido*-7,8-$C_2B_9H_{12}]^-$. (*b*) [*nido*-7,9-$C_2B_9H_{12}]^-$. (*c*) [*nido*-$C_2B_{10}H_{12}]^{2-}$ anions. (*d*) $(C_2B_9H_{11})_2M^{N\text{-}4}$ complexes, where *N* = oxidation state of the metal.

adjacent to the carboxyl. The following exemplifies both systems:

$$\overset{\epsilon}{\underset{6}{CH_2}}\overset{}{O}\overset{\delta}{\underset{5}{H}}\overset{\gamma}{\underset{4}{CH_2}}\overset{\beta}{\underset{3}{CHBr}}\overset{\alpha}{\underset{2}{CHBr}}\underset{1}{CH_2COOH}$$

is 3,4-dibromo-6-hydroxyhexanoic acid, or β,γ-dibromo-ϵ-hydroxycaproic acid. *See* ORGANIC CHEMISTRY.

Physical and chemical properties of carboxylic acids are represented, grossly, by the resultant of the various chemical groupings present in the molecule. A short-chain aliphatic acid, wherein the carboxyl is dominant, is a pungent, corrosive, water-soluble liquid of abnormally high boiling point (because of molecular association), with specific gravity close to 1 (higher for formic and acetic acids). With increasing molecular weight, the hydrocarbon grouping overbalances the carboxyl; sharpness of odor diminishes, boiling and melting points rise, the specific gravity falls toward that of the parent hydrocarbon, and the water-solubility decreases. Thus the typical high-molecular-weight saturated acid is a bland, wax-like solid.

Ionization and acidity. Ionization occurs on exposure to a proton acceptor (base, in the Brønsted-Lowry sense); in water the opposing equations shown in reaction (1) describe the fundamental equi-

$$R(CXY)_nCOOH + H_2O \rightleftharpoons R(CXY)_nCOO^- + H_3O^+ \qquad (1)$$

librium encountered with any feebly ionized, or weak, acid. Hydration of the proton to form H_3O^+ is virtually complete; thus, the process of ionization is analogous to that of neutralization, the degree to which ionization proceeds varying directly with the basicity of the solvent (proton acceptor), and inversely with the O-H bond strength of the carboxylic hydroxyl group.

The above equilibrium leads to the dissociation constant K_a defined as reaction (2), wherein the square brackets signify concentration, measured in

$$K_a = \frac{[H_3O^+][R(CXY)_nCOO^-]}{[R(CXY)_nCOOH]} \qquad (2)$$

moles per liter of solution. With simple, saturated acids, wherein X and Y are hydrogen, the value of K_a is always small (acetic acid, 1.75×10^{-5}), which means that simple organic acids are only slightly ionized in aqueous solution. However, the degree of dissociation is sensitive to electronegative (electron-attracting) substituents, particularly when located on the α- or 2-carbon (for example, K_a for trichloroacetic acid, 1.3×10^{-1}). *See* IONIC EQUILIBRIUM.

Table 1 presents a list of 30 saturated, monocarboxylic acids, together with their important constants and main sources.

Structure and properties. The un-ionized carboxyl group has been shown via many classical studies to have the structure

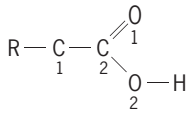

in which carbon 2 is doubly bonded to oxygen 1 and singly bonded to oxygen 2. Geometrical considerations lead to the conclusions that distance

$$\underset{2}{C}{=}\underset{1}{O}$$

should be shorter than

$$\underset{2}{C}{-}\underset{2}{O}$$

and the angle

$$\underset{1}{C}{-}\underset{2}{C}{=}\underset{1}{O}$$

should be greater than angle

$$\underset{1}{C}{-}\underset{2}{C}{-}\underset{2}{O}$$

This is confirmed by measurements on crystalline acids, although the fact that the single bond

$$\underset{2}{C}{-}\underset{2}{O}$$

never reaches the normal length (about 0.144 nm) means that some resonance is involved between the two forms

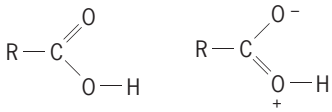

In the ionized state, however, the two resonating forms of the anion are electronically equivalent,

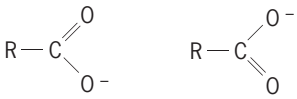

and the actual structure of the carboxylate ion is a stable hybrid of these.. *See* RESONANCE (MOLECULAR STRUCTURE).

In the free state (solid, liquid, or gas), the lower acids tend to exist as dimers, in which two molecules are associated through relatively weak hydrogen bonds (energy about 5 kcal/mole):

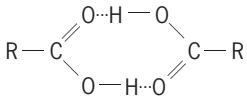

The hydrogen bonds (indicated by dotted valences) account for the abnormally high boiling points of the smaller acids. With longer-chain acids (from C_8) there is a tendency toward polymorphism, ascribable to hydrogen bonding of the following type:

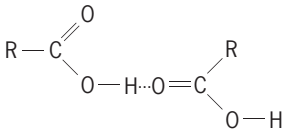

As determined by x-ray diffraction studies, the carboxyl and terminal methyl groups of acids containing an odd number of carbons are on the same side of

TABLE 1. Saturated monocarboxylic acids

Name	Formula	Main source*	Melting point, °C (°F)	Boiling point, °C (°F)
Formic	HCOOH	S	8.4 (47.1)	100.5 (212.9)
Acetic	CH_3COOH	S and fermentation	16.7 (62.1)	118.1 (244.6)
Propionic	CH_3CH_2COOH	S	−22.0 (−7.6)	141.0 (285.8)
Butyric	$CH_3(CH_2)_2COOH$	F, S, and ferm.	−4.7 (23.5)	164.1 (327.4)
Valeric	$CH_3(CH_2)_3COOH$	S and *Valeriana officinalls*	−34.5 (−30.1)	186.4 (367.5)
Isovaleric	$(CH_3)_2CHCH_2COOH$	F and S	−29.3 (−20.7)	176.5 (349.7)
Caproic	$CH_3(CH_2)_4COOH$	F and S and by-product in butyric fermentation	−1.5 (29.3)	205.8 (442.4)
Heptanoic	$CH_3(CH_2)_5COOH$	S from heptaldehyde	−7.5 (18.5)	223.0 (433.4)
Caprylic	$CH_3(CH_2)_6COOH$	F and S; goat's butter	16.3 (61.3)	239.7 (463.5)
Pelargonic	$CH_3(CH_2)_7COOH$	S; oxidation of oleic and *Pelargonium roseum*	12.5 (54.5)	255.6 (492.1)
Capric	$CH_3(CH_2)_8COOH$	F; coconut; palm kernel oils	31.5 (88.7)	268.7 (515.7)
Undecylic	$CH_3(CH_2)_9COOH$	S; reduction of undecenoic acid	29.3 (84.7)	228 (442)/21 kPa
Lauric	$CH_3(CH_2)_{10}COOH$	F; coconut and laurel oils	44.1 (111.4)	225 (437)/13 kPa
Tridecylic	$CH_3(CH_2)_{11}COOH$	S	41.0 (105.8)	236 (457)/13 kPa
Myristic	$CH_3(CH_2)_{12}COOH$	F; nutmeg oil	58.0 (136.4)	250 (482)/13 kPa
Pentadecylic	$CH_3(CH_2)_{13}COOH$	S	54.0 (129.2)	257 (495)/13 kPa
Palmitic	$CH_3(CH_2)_{14}COOH$	F; palm and olive oils	62.8 (145.0)	271.5 (520.7)/13 kPa
Margaric	$CH_3(CH_2)_{15}COOH$	S	59.9 (139.8)	227 (441)/13 kPa
Stearic	$CH_3(CH_2)_{16}COOH$	F; tallow and reduction of oleic acid	69.9 (157.8)	291 (556)/15 kPa
Nonadecylic	$CH_3(CH_2)_{17}COOH$	S	66.5 (151.7)	298 (568)/13 kPa
Arachidic	$CH_3(CH_2)_{18}COOH$	F; peanut oil	77.0 (170.6)	205 (401)/130 Pa
Heneicosanoic	$CH_3(CH_2)_{19}COOH$	S	75.2 (167.4)	
Behenic	$CH_3(CH_2)_{20}COOH$	F; reduction of erucic acid	80.2 (176.3)	306 (583)/8 kPa
Tricosanoic	$CH_3(CH_2)_{21}COOH$	S	79.1 (174.4)	
Lignoceric	$CH_3(CH_2)_{22}COOH$	F; *Adenanthera pavonina* seed; beechwood tar; rotten oak	84.2 (183.6)	
Pentacosanoic	$CH_3(CH_2)_{23}COOH$	S	83.0 (181.4)	
Cerotic	$CH_3(CH_2)_{24}COOH$	Beeswax; carnauba wax	87.7 (189.9)	
Heptacosanoic	$CH_3(CH_2)_{25}COOH$	S	87.6 (189.7)	
Montanic	$CH_3(CH_2)_{26}COOH$	Montan wax from lignite	90.9 (195.6)	
Nonacosanoic	$CH_3(CH_2)_{27}COOH$	S	90.3 (194.5)	
Melissic	$CH_3(CH_2)_{28}COOH$	Beeswax	93.6 (200.5)	

*S stands for synthetic, meaning that the acid is manufactured commercially by one or more methods. F stands for fats; the fatty acids comprise a group of carboxylic acids, both saturated and unsaturated, which may be obtained by hydrolysis of animal or vegetable fats and oils.

the molecule; with acids containing an even number of carbons, these groups are on opposite sides, permitting closer packing in the crystal lattice and consequent increased van der Waals forces. Thus, the melting point of any even-numbered acid lies above that of the preceding and following odd-numbered acid.

Structural variations. These are of two types: those not involving the carboxyl group, and those involving the carboxyl, leading to the acid derivatives.

In the first type are such distinctive classes as branched-chain acids, for example, pivalic acid, $(CH_3)_3CCOOH$; alicyclic acids, for example, cyclopropane carboxylic acid,

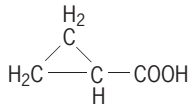

halogenated acids, fluoro, chloro, or bromo, rarely, iodo; hydroxy acids; dicarboxylic and polycarboxylic acids; aromatic acids; β-keto acids or amino acids. *See* AMINO ACIDS; BENZOIC ACID; LACTAM; LACTONE; PROTEIN; TARTARIC ACID.

An important class of acids contains unsaturated groups, for example, acrylic acid ($CH_2{=}CHCOOH$). Acids of this type are unstable and polymerize readily; hence they and their esters are much studied. *See* POLYACRYLATE RESIN.

Among the variations involving the carboxyl group are found many important classes of substances, derived from acids by suitable substitution within the carboxylic group. Thus, replacement of the hydroxyl group with alkoxyl gives esters; replacement of OH by halogen forms acid halides; substitution of OH by amino (NH_2) group gives amides; intermolecular dehydration of two acids generates the anhydride structure

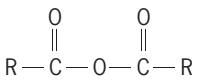

and the hydrolysis of long-chain esters by alkali hydroxides forms soaps. *See* ACID ANHYDRIDE; ACID HALIDE; ACYLATION; AMIDE; DETERGENT; ESTER; FAT AND OIL; SOAP; UREA; WAX, ANIMAL AND VEGETABLE.

Table 2 lists some important dicarboxylic acids, aromatic acids, and unsaturated acids, together with their important physical constants.

Table 2. Dicarboxylic, aromatic, and unsaturated acids

Name	Formula*	Melting point, °C (°F)*	Boiling point, °C (°F)*
Dicarboxylic acids			
Oxalic	HOOC—COOH	189.5 (373.1) (dec)	
Malonic	$CH_2(COOH)_2$	135 (275) (dec)	
Succinic	$(CH_2)_2(COOH)_2$	185–187 (365–369)	235 (455)
Glutaric	$(CH_2)_3(COOH)_2$	98.0 (208)	302–304 (576–579)
Adipic	$(CH_2)_4(COOH)_2$	152 (306)	337.5 (639.5)
Pimelic	$(CH_2)_5(COOH)_2$	105–107 (221–225) (subl)	272 (522) at 13 kPa
Suberic	$(CH_2)_6(COOH)_2$	142 (288)	279 (534) at 13 kPa
Azelaic	$(CH_2)_7(COOH)_2$	106.5 (223.7)	286 (547) at 13 kPa
Sebacic	$(CH_2)_8(COOH)_2$	134 (273)	295 (563) at 13 kPa
Aromatic acids			
Benzoic	C_6H_5COOH	121.7 (251.1)	249 (480)
o-Toluic	*o*-$CH_3C_6H_4COOH$	104 (219)	259 (498) at 100 kPa
m-Toluic	*m*-$CH_3C_6H_4COOH$	111 (232)	263 (505)
p-Toluic	*p*-$CH_3C_6H_4COOH$	180 (356)	275 (527)
o-Ethylbenzoic	*o*-$C_2H_5C_6H_4COOH$	68 (154)	259 (498) at 100 kPa
m-Ethylbenzoic	*m*-$C_2H_5C_6H_4COOH$	47 (117)	
p-Ethylbenzoic	*p*-$C_2H_5C_6H_4COOH$	113 (235)	
o-Fluorobenzoic	*o*-FC_6H_4COOH	122 (252)	
m-Fluorobenzoic	*m*-FC_6H_4COOH	124 (255)	
p-Fluorobenzoic	*p*-FC_6H_4COOH	182–184 (360–363)	
o-Chlorobenzoic	*o*-ClC_6H_4COOH	140.2 (284.4)	
m-Chlorobenzoic	*m*-ClC_6H_4COOH	154.3 (309.7)	
p-Chlorobenzoic	*p*-ClC_6H_4COOH	239.7 (463.5)	Subl
o-Bromobenzoic	*o*-BrC_6H_4COOH	148–150 (298–302)	Subl
m-Bromobenzoic	*m*-BrC_6H_4COOH	154–155 (309–311)	
p-Bromobenzoic	*p*-BrC_6H_4COOH	251–253 (484–487)	
o-Iodobenzoic	*o*-IC_6H_4COOH	162 (324)	
o-Iodobenzoic	*o*-IC_6H_4COOH	162 (324)	
m-Iodobenzoic	*m*-IC_6H_4COOH	187 (367)	Subl
p-Iodobenzoic	*p*-IC_6H_4COOH	270 (518)	Subl
Salicylic	*o*-HOC_6H_4COOH	158.3 (316.9)	211 (412) at 2.7 kPa
m-Hydroxybenzoic	*m*-HOC_6H_4COOH	201 (394)	
p-Hydroxybenzoic	*p*-HOC_6H_4COOH	215 (419)	
Gallic	3,4,5-$(HO)_3C_6H_4COOH$	235 (455) (dec)	
Anthranilic	*o*-$H_2NC_6H_4COOH$	144–145 (291–293)	Subl
m-Anthranilic	*m*-$H_2NC_6H_4COOH$	173–174 (343–345)	
p-Anthranilic	*p*-$H_2NC_6H_4COOH$	187–188 (369–370)	
α-Naphthoic	α-$C_{10}H_7COOH$	160–161 (320–322)	300 (572)
β-Naphthoic	β-$C_{10}H_7COOH$	184 (363)	>300 (572)
Unsaturated acids			
Acrylic	$CH_2{=}CH{-}COOH$	14 (57)	141
Crotonic (trans)	$CH_3CH{=}CHCOOH$	71.6 (161)	185
i-Crotonic (cis)		15.5 (59.9)	
Mesaconic (trans)	$HOOC{-}C(CH_3){=}CHCOOH$	202 (396)	250 (482) (dec)
Citraconic (cis)		91 (196)	
Tiglic (trans)	$CH_3CH{=}C(CH_3)COOH$	64 (147)	198 (388)
Angelic (cis)		45 (133)	185 (365)
Cinnamic (trans)	$C_6H_5CH{=}CHCOOH$	133 (271)	300 (572)
Allocinnamic (cis)		68 (154)	125 (257) at 2.4 kPa
Fumaric (trans)	$HOOCCH{=}CHCOOH$	287 (549)	
Maleic (cis)		130 (266)	
Elaidic (trans)	$CH_3(CH_2)_7CH{=}HOOC(CH_2)_7CH$	44 (111)	288 (550) at 13 kPa
Oleic (cis)		13 (55)	288 (550) at 13 kPa

*dec = decomposes, subl = sublimes.

Reactions and uses. Acids are used in large quantities in the production of esters, acid halides, acid amides, and acid anhydrides. Decarboxylation of acids to form hydrocarbons containing one less carbon is accomplished by pyrolyzing the sodium or barium salt with soda lime. If the potassium salt is electrolyzed at a platinum anode (Kolbe electrolysis), the hydrocarbon RCH_2CH_2R is produced by the

acid $RCH_2COO^-K^+$. Ethylenic and acetylenic acids containing unsaturation α,β to the carboxyl are easily decarboxylated by heat. Aromatic acids, such as benzoic and toluic, are frequently used as sources of hydrocarbons; the decarboxylation is effected by treating the acid in boiling quinoline with a copper powder catalyst.

Many acids obtained by acid hydrolysis of fats or waxes are reduced to the corresponding alcohols, for example, lauric acid to lauryl alcohol. The reaction is carried out in the laboratory by means of lithium aluminum hydride; the industrial procedure utilizes hydrogen at elevated temperatures, over a mixed catalyst of the oxides of copper and chromium.

Acids find use in the manufacture of soaps and detergents, in thickening lubricating greases (stearate soaps), in modifying rigidity in plastics, in compounding buffing bricks and abrasives, and in the manufacture of crayons, dictaphone cylinders, and phonograph records. The solvent action of acids finds use in manufacture of carbon paper, inks, and in the compounding of synthetic and natural rubber. Because of the stability of saturated fatty acids toward oxidation, these are often used as solvents for carrying out oxidation reactions upon the sensitive compounds.

Evans B. Reid

Carburetor

The device that controls the power output and fuel feed of internal combustion spark-ignition engines generally used for automotive, aircraft, and auxiliary services. Its duties include control of the engine power by the air throttle; metering, delivery, and mixing of fuel in the airstream; and graduating the fuel-air ratio according to engine requirements in starting, idling, and load and altitude changes. The fuel is usually gasoline or similar liquid hydrocarbon compounds, although some engines with a carburetor may also operate on a gaseous fuel such as propane or compressed natural gas. A carburetor may be classified as having either a fixed venturi, in which the diameter of the air opening ahead of the throttle valve remains constant, or a variable venturi, which changes area to meet the changing demand. *See* AUTOMOBILE; ENGINE; FUEL SYSTEM; VENTURI TUBE.

Engine air charge. A simple updraft carburetor with a fixed venturi illustrates basic carburetor action (**Fig. 1**). Intake air charge, at full or reduced atmospheric pressure as controlled by the throttle, is drawn into the cylinder by the downward motion of the piston to mix with the unscavenged exhaust remaining in the cylinder from the previous combustion. The total air-charge weight per unit time is approximately proportional to the square root of its pressure drop in the carburetor venturi (with a square-root correction, direct for change in air pressure, and inverse for change in air absolute temperature). Also, each individual cylinder air-charge volume follows the intake pressure minus about one-sixth the exhaust pressure, so that the intake

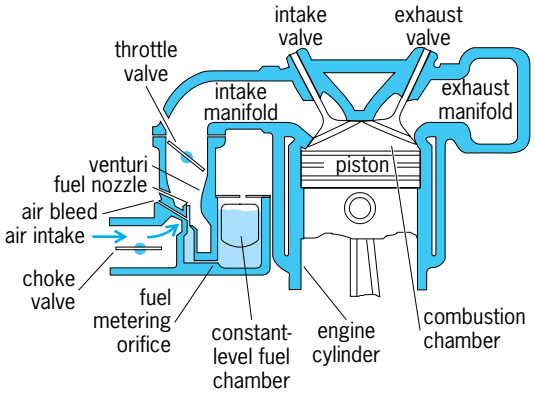

Fig. 1. Elements that basically determine air and fuel charges received by the engine through the carburetor.

manifold pressure and temperature are often taken as an approximate indication of the engine torque; this factor multiplied by the engine speed is similarly taken as a measure of the engine power output. Any given fixed part-throttle opening gives higher intake manifold pressure, increased individual-cylinder air charge, and greater engine torque, as the engine speed is decreased. *See* TORQUE.

A cylinder is most completely filled with the fuel-air mixture when no other cylinder is drawing in through the same intake passage at the same time. For high speed, best power is usually obtained with multiple carburetor and intake passages. However, on an engine so equipped, it is often difficult to throttle all the cylinders equally. *See* VOLUMETRIC EFFICIENCY.

Fuel charge. The fuel is usually metered through a calibrated orifice, or jet, at a differential pressure derived from the pressure drop in a venturi in the intake air passage. Since this fuel flow also follows the square root of its differential pressure, the fuel and air rates tend to be accurately parallel, particularly if a small bleed of air is fed into the fuel delivery or spray nozzle to overcome the retarding effect of surface tension.

Fuel-air requirements. For normal steady speed, the limits of consistent ignition lie between 0.055 and 0.09 fuel-air by weight (**Fig. 2**). Richer mixtures (more fuel) give higher power, apparently cooler combustion, and less detonation. Leaner mixtures (less fuel) give better fuel economy. *See* COMBUSTION; COMBUSTION CHAMBER.

A narrowly controlled rich mixture is required for idling because low charge densities and relatively high contamination with residual exhaust gas make combustion conditions unfavorable. Individual adjustments may be provided both for the idling fuel feed and for the throttle closure stop to keep the engine running at minimum speed. Slight intake or exhaust valve leaks can make smooth idling impossible.

Cold-engine operation. Engine fuels are a mixture of components varying in volatility, usually selected so that they evaporate completely in the intake manifold at ordinary operating temperatures at the reduced pressures of part throttle, but not at full throttle. For starting at low temperatures and for the first

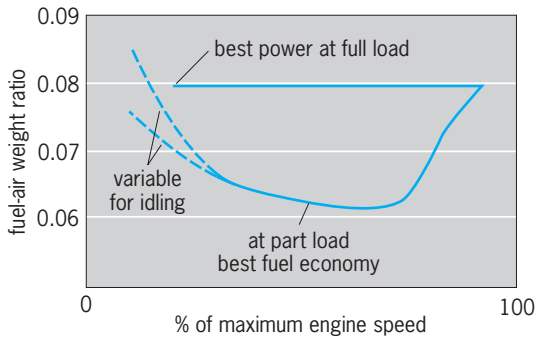

Fig. 2. Fuel-air ratios required for engine operation.

few ignitions with cylinder walls cold, a great increase in fuel feed is momentarily required since only the most volatile part of the fuel can vaporize and burn. *See* GASOLINE.

The higher metering suction necessary to give the excess fuel feed required for cold starting may be created by a manually or automatically operated choke valve. Partial closure of this valve also yields the moderate enrichment needed during warmup. An additional linkage is generally supplied on automotive carburetors so that pressing the accelerator pedal all the way down will open the choke valve and permit return to normal fuel feed.

Automotive carburetion. An automotive engine performs through a wider range of loads and speeds than most prime movers. Furthermore, its transient response to changes of load is highly important. For these reasons, the functions of an automotive carburetor are complex. The multiple main and boost venturi structure is used to increase the fuel metering force for a given air delivery; also, it yields better metering regulation at the higher air velocities.

Idling system. To obtain adequate metering forces at low air speeds and small throttle openings, the fuel passes through the main metering jet and is bypassed to an idle metering jet. Then the fuel mixes with air bled in through the idle air bleed, and this fuel and air are delivered to an idle port in the high suction at the edge of the throttle valve. Some graduation of fuel feed from low to high idling speed is obtained as the throttle valve edge passes across the idle port, which may include multiple delivery holes.

Because of increased piston and other friction with a cold engine, a greater throttle opening as well as more fuel is required for idle at that time. The choke valve is usually interconnected so that, in its partly closed or cold-engine position, the minimum-throttle area is increased.

Power enrichment. The richer fuel feeds necessary for best power are customarily obtained by varying the area of the main metering jet, using a valve responsive to intake manifold pressure beyond the throttle. At light loads the lowered manifold pressure draws the enrichment valve down. A rise in manifold pressure resulting either from a drop in engine speed or added throttle opening will allow the plunger spring to open the metering jet further. An approximation of this function is obtained in some carburetors by mechanical linkage between the enrichment valve and the throttle valve. On most carbureted automotive engines with a catalytic converter, the enrichment valve is computer controlled to help maintain stoichiometry.

Transitional requirements. Proper regulation under change of throttle is a major problem in obtaining good carburetor action. Any lag in evaporation of the fuel when the throttle is opened tends to give a momentary delay in power response, which is usually compensated by an accelerator pump. The pump gives a quick squirt of fuel as the throttle is opened. The apparent willingness of the engine to respond to the accelerator pedal is largely determined by how accurately the accelerator pump discharge is proportioned to the existing engine temperature and to the volatility of the fuel.

A converse transitional problem is associated with sudden closing of the throttle, when wet fuel is present on the intake manifold walls. The sudden drop in pressure causes the liquid to flash into vapor, resulting in a temporarily overrich condition in the cylinder, with subsequent misfiring and incompletely burned combustion products in the exhaust. Better manifold heating helps, as does injection of fuel at cylinder intake ports. *See* FUEL INJECTION.

Aircraft carburetion. Early aviation carburetors followed automobile carburetor practice, but in the 1940s the pressure or injection type came into use. These incorporated the following advantages:

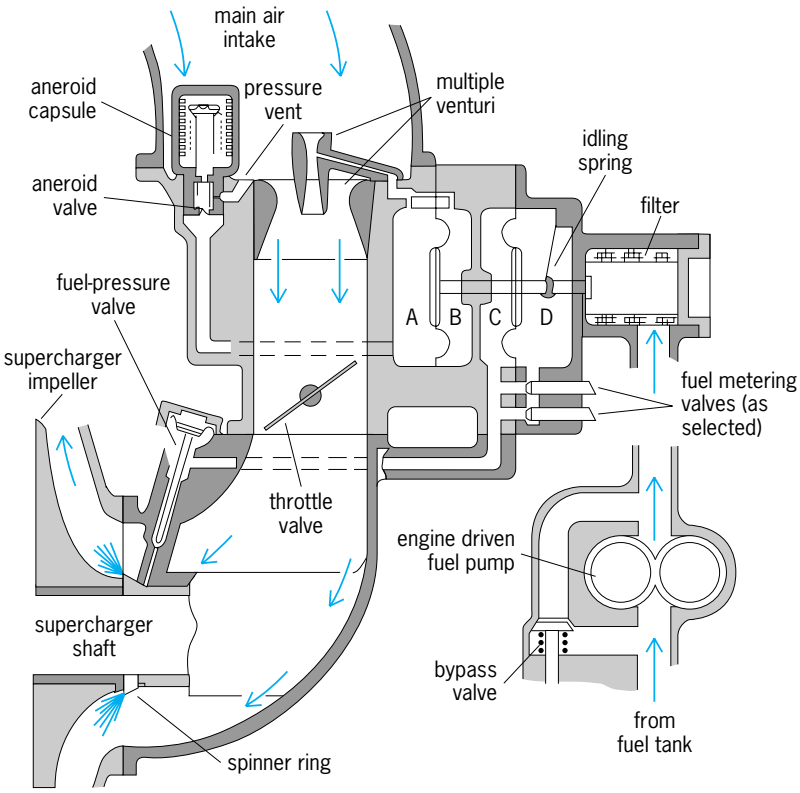

Fig. 3. Schematic diagram of an aircraft injection-type or pressure-type carburetor.

(1) The whole metering system is kept under pressure, to prevent formation of bubbles and disturbances to the metering under the reduced pressures at higher altitudes. (2) The fuel spray can be delivered into a heated part of the intake system so that ice formation resulting from fuel evaporation becomes impossible. (3) Correction of the metering function under changes of altitude and temperature is automatically provided. (4) The closed system avoids disturbance of the fuel metering, or leakage from vents, during vigorous maneuvers of the airplane.

In the most commonly used form, the metering function of the automobile carburetor is retained. The fuel metering differential (**Fig. 3**) is derived from the multiple venturi air differential by the system of opposed synthetic-rubber diaphragms. The idling spring adds a small positive increment to provide the required rich idle fuel flow.

The fuel pressure in the carburetor is held to a desired value above the surrounding atmosphere by the fuel-pressure valve and its regulating diaphragm and spring. Beyond the pressure valve the fuel may be led to a spinner ring which has multiple discharge orifices. The ring is mounted on, and rotates with, the supercharger shaft. On other installations, the fuel, after metering, is taken to an injection pump which divides the charge equally and delivers a portion at much higher pressure to the interior of each engine cylinder on its intake stroke.

Pressure from the engine-driven fuel pump, as regulated by its bypass valve, must be adequate to maintain flow through the system at the highest powers, but beyond this it does not affect the fuel metering. Fuel-air ratio, as desired for light or full load or as selected by the pilot, is controlled by one or more metering valves which collectively determine the flow area between the two fuel-metering chambers. A variety of combinations have been used, but all are directed toward the regulation illustrated in Fig. 2.

Correction of the fuel-air ratio at increased altitude or temperature is accomplished by differentials. A small fixed-size depression vent is provided between the innermost multiple venturi chamber and the venturi communication channel, while a large pressure vent between the main air intake and the chamber is varied by the aneroid valve. This valve is connected to the aneroid capsule, a metallic bellows filled with inert gas and extended by a spring. As the airplane gains altitude, or as the temperature of the entering air rises, the capsule extends to reduce the area of the entering air vent to the innermost chamber. This action reduces the differential between the multiple venturi chambers at a rate determined by the profile and adjustment of the aneroid valve.

Deviations from the form that provides power enrichment by varying the main metering orifice have been used on smaller aircraft engines. Similar construction has also been used to regulate fuel feed by the parameters of engine speed and intake manifold pressure. This operation is commonly called speed-density metering.

Carburetor icing. There are two reasons for condensation of atmospheric moisture and formation of ice in the intake system. First, at partly closed throttle, the adiabatic pressure drop across the throttle orifice generates a temperature drop. When atmospheric humidity is high, ice sometimes partly clogs the throttle orifice, requiring further throttle opening to keep the engine running for a short period after starting and before the engine heat has had time to build up. With carburetors having a fuel discharge at the throttle edge, the addition of water-soluble antifreeze to the fuel may reduce such icing tendency.

The other source of temperature drop in the intake system is the evaporation of fuel. Ice tends to form wherever fuel spray impinges upon an unheated surface. Such icing was formerly encountered frequently in aircraft operation and was dealt with by heating the intake air, which reduced the engine power and sometimes caused detonation. The problem was largely eliminated by use of the fuel-injection system. *See* FUEL INJECTION. Donald L. Anglin

Bibliography. W. H. Crouse and D. L. Anglin, *Automotive Engines*, 1995; J. B. Heywood, *Internal Combustion Engine Fundamentals*, 1988.

Cardamon

A spice from the plant *Elettaria cardamomum* (Zingiberaceae), a perennial herb 6–12 ft (1.8–3.6 m) tall and native to India. The small, light-colored seeds, borne in capsules, have a delicate flavor. An important spice in the Orient for centuries, they are used in curries, cakes, pickles, and in general cooking, as well as in medicine. In India the seeds are a favorite masticatory. *See* SPICE AND FLAVORING; ZINGIBERALES. Perry D. Strausbaugh; Earl L. Core

Cardiac electrophysiology

The science of the electrical activity of the heart. The heartbeat results from the development and organized control of cardiac excitability, including ionic current flow across the cardiac membrane, within and between cells, and throughout the body, which in turn allows the orderly contraction of heart muscle and the efficient pumping of blood. Alterations in these determinants of normal cardiac excitability may lead to abnormalities in the rhythm of the heart (arrhythmias or dysrhythmias) and in the propagation of electrical impulses throughout the heart (conduction defects). Cellular electrophysiologic events lead to the establishment of extracellular potentials on the surface of the body. Electrocardiographic, vectorcardiographic, and other recording systems are used to determine the orientation and magnitude of that extracellular potential. The normal heart produces characteristic sequences of extracellular potentials that may be altered by disease.

Cellular electrophysiology. Cardiac excitability depends on the interaction among the determinants of the resting potential and both the active and passive cellular properties, which can be considered elements in an electrophysiologic matrix. The inside and outside of a cardiac cell are connected by channels that run through a membrane consisting of a lipid bilayer. Electrochemical gradients are established across the membrane in which the ionic activities differ inside and outside the cell. The extracellular activity of sodium ions (Na^+) and calcium ions (Ca^{2+}) exceeds the intracellular activity, which causes an inward leak; the reverse is true for potassium ion (K^+) activity. Energy-requiring pumps and other exchange mechanisms maintain the distribution of the ionic activities across the membrane. The unequal distribution of each ion species charge across the semipermeable membrane is well described by the Nernst equation, which determines the potential required to oppose the activity gradient and is termed the equilibrium potential (E_{ion}). *See* BIOPOTENTIALS AND IONIC CURRENTS; CELL PERMEABILITY.

Active cellular properties relate to the opening and closing of channels that carry ionic currents, which, in turn, are responsible for rapid depolarization and for both the slow and rapid phases of repolarization. The channels are selectively permeable to different ionic species and are controlled by gates that open and close as a function of voltage, time, or, sometimes, binding by a hormone, drug, or other ligand. The current flow of an ionic species (i_{ion}) depends on the driving force (the difference between the transmembrane voltage, V_m, and the ion's equilibrium potential, E_{ion}) and the membrane conductance (g_{ion}), that is, the ease with which ions pass through the channels. In other words, $i_{ion} = g_{ion}(E_{ion} - V_m)$. Conductance ($g$) is the reciprocal of resistance (R).

Some tissues—such as those in the atria, the Purkinje fibers of the specialized infranodal conduction system, and the ventricles—depend on the inrush of sodium ion current (i_{Na^+}) to produce the phase of rapid depolarization, phase 0. The kinetics of the channel, the inrush of sodium ion current, and the maximum rate of rise of voltage ($\dot{V}_{max}$) during phase 0 are rapid. Thus, these tissues are often called fast-response tissues. Much of what is known about cardiac electrophysiology has been determined by using microelectrodes that are inserted through the cell membrane into the interior of the cell, thus allowing the recording and perturbation of transmembrane voltages (**Fig. 1**). The resting potential (V_r) and subthreshold conductance (before excitation) in fast-response tissues are determined primarily by potassium conductance (g_K). As the cell is depolarized from a resting potential of about −90 millivolts to about −60 mV, a point often termed the threshold voltage (V_{th}), the response becomes regenerative and out of proportion to the stimulus. This regenerative response is called the action potential.

The membrane potential in fast-response tissues depends on a strictly controlled balance between depolarizing inward currents, which help maintain the action potential plateau, and repolarizing outward currents, which favor a return to the resting potential. Phase 1 depends on the inactivation of sodium conductance and, in Purkinje fibers, perhaps on the activation of the positive dynamic current, which may be a calcium-activated potassium ion current. Events during phase 2, the plateau phase, are complex. The reduced transmembrane current opens the

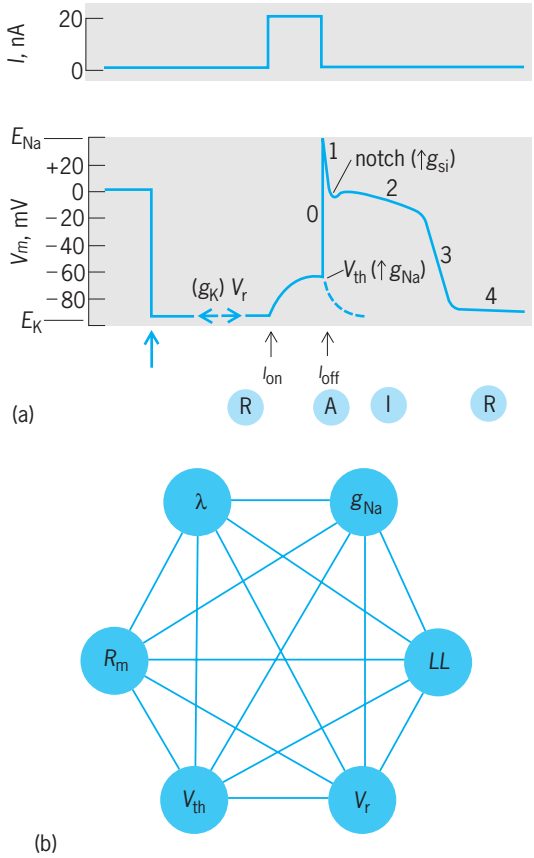

Fig. 1. Cellular electrophysiology. (*a*) Oscilloscopic representations of current (*I*) and the transmembrane voltage V_m. When the microelectrode is in the extracellular space, V_m is 0. With impalement of the tissue, V_m falls from 0 to −95 mV (colored vertical arrow) which is the resting potential (V_r). With physiologic extracellular K^+ activities, V_r approximates the potassium equilibrium potential (E_K). Depolarizing current (I_{on}) causes V_m to become more positive and to approach V_{th} for regenerative phase 0 depolarization. If the liminal length requirements are not met, V_m decays exponentially (broken line following I_{off}) because of passive membrane resistive-capacitative properties. Membrane conductances for both resting membrane and subthreshold voltage range are determined primarily by membrane potassium conductance (g_K). If liminal length requirements are met, an action potential results, and the overshoot approaches the equilibrium potential for sodium (E_{Na}). States of the sodium channel are indicated by R (resting), A (open or activated), and I (inactivated). The phases of the action potential are shown, as are the conductances of particular importance to each phase (g_K, g_{Na}, g_{si}). (*b*) Simple matrix of selected active and passive cellular properties that determine cardiac excitability. The determinants depicted include the resting potential (V_r), threshold voltage (V_{th}), sodium conductance (g_{Na}), membrane resistance (R_m), the length constant (λ), and, as a measure of overall excitability, the liminal length (*LL*). Note that each has its own determinants, but other determinants are also important, so that the scheme is multidimensional and more complex than illustrated.

channel for the mixed calcium ion–sodium ion, or "slow inward," current (i_{si}), which helps maintain the plateau. An inward sodium current, the "window" current, is distinct from that active during phase 0 and helps maintain the plateau. At the transmembrane voltage of the plateau, potassium ion conductance decreases in the outward direction but increases in the inward direction, and thus repolarization is more difficult; this process is called inward rectification. That mechanism helps maintain the plateau. The delayed rectifier potassium current has a slow, rapid, and ultrarapid component termed, respectively, i_{Ks}, i_{Kr}, and i_{Kur}. This rectification mechanism helps maintain the plateau.

Phase 3 repolarization has also been termed the rapid repolarization phase. During phase 3, a rapid return to resting potential results from inactivation of the slow inward current, reversal of the process of inward rectification, and activation of repolarizing outward potassium currents. During this phase, enough of the sodium system is normally reactivated so that another action potential can be electrically induced. Sodium channels may exist in several states: the resting state (R) predominates in normally polarized tissues; the activated state (A) occurs transiently with opening of the sodium channel; and the inactivated state (*I*) dominates at more positive transmembrane conductance, such as during the plateau or in tissues that have been depolarized because of injury.

Other tissues, such as the sinoatrial (SA) and atrioventricular (AV) nodes, depend on a slow inward current of mixed calcium and sodium ions for phase 0 depolarization. They have been termed slow-response tissues because the current flows through a kinetically slow channel distinct from the channel that controls sodium. Because the channel preferentially carries calcium ions, it is often called the calcium current. Phase 0 in the sinoatrial and atrioventricular nodes normally depends on a slow inward current, and depolarized fast-response tissues may switch over to action potentials dependent on slow inward current. Slow-response tissues have a resting potential of about −60 mV, a transmembrane voltage at which the sodium system is virtually inactivated. The slow channel is voltage-dependent and has resting, active, and inactivated phases. The phase of inactivation depends on both voltage and time and extends into electrical diastole, a characteristic that limits the number of impulses that can transverse the atrioventricular node in atrial fibrillation and flutter.

The cardiac cell has a low-resistance myoplasm surrounded by a membrane that has both capacitative elements and a high resistance; it is analogous to a telegraph cable with a low-resistance core and high-resistance insulation. Cells are connected and communicate with each other through gap junctions, which normally have a low resistance to ionic flow. With injury, gap junction resistance may increase and cells may uncouple—actions mediated by changes in pH and intracellular calcium activity. Membrane insulation is imperfect, and so current leaks out, resulting in a drop in available current for the longitudinal flow of ions through the myoplasm and the gap junctions.

Liminal length. Liminal length is the amount of tissue that must be raised above threshold so that the inward depolarizing current from one region is sufficiently greater than the repolarizing influences of the adjacent tissues to permit regenerative depolarization of neighboring tissues. The liminal length is directly proportional to the charge developed by the active cellular properties and inversely proportional to the membrane capacitance and the length constant. If the liminal length requirements of an element are fulfilled, that element will influence its neighboring element. If the electrotonic influence from the first membrane element is sufficient to bring its adjacent membrane element up to threshold (or, more properly, to fulfill its liminal length requirements), regenerative depolarization and an action potential are produced in the second element. If the electrotonic influence is insufficient, the current continues to decay exponentially. Clearly, cardiac excitability depends both on the current strength created by active cellular properties and on the "sink" of the tissue's cable properties. Each membrane element that fulfills its liminal length requirement will bring about regenerative depolarization, and an action potential will propagate one unit further. Impulse propagation, then, depends on both active and passive cellular properties. Conduction velocity is much more rapid in fast-response, as compared to slow-response, tissues.

Excitation-contraction coupling. Excitation-contraction coupling in heart tissue is complex. A rise in intracellular calcium ions is required to activate the proteins responsible for myocardial contraction. The slow inward current seems to trigger and modulate the release of calcium ions from the sarcoplasmic reticulum and other intracellular storage sites and to be involved in replenishing calcium ions in those stores. Sodium current, resultant changes in intracellular sodium activity, and sodium ion–calcium ion exchange, which influences the intracellular milieu, also affect contractility.

Automaticity. True spontaneous or automatic activity arises in the absence of direct external causes. Such automaticity may be physiological, as in the sinoatrial node and the lower escape pacemakers, or pathological, as in states of accelerated, depressed, or abnormal pacemaker function. Yet another type is triggered automaticity, in which nondriven action potentials are initiated by one or more driven action potentials. Normally, the sinoatrial node is the predominant cardiac pacemaker and controls the overall heart rhythm. Pacemaker activity is the product of a changing balance between positive inward currents, which favor depolarization, and positive outward currents, which favor repolarization. The depolarization current of importance is slow inward current, although a second inward current may also play a role. A slowly decaying outward potassium current may also be involved in causing pacemaker activity. *See* HEART (VERTEBRATE); MUSCLE.

Channelopathies. Abnormalities of the function of ion channels may be inherited or acquired. Mutations can occur in the encoding genes of ion channels, which affect the gating (opening and closing) of the channel proteins. The long QT syndrome (LQTS), for example, is the phenotypic description of a clinical disorder characterized by prolongation of the QT interval on the electrocardiogram (see below) associated with a certain type of heart rhythm disturbance termed polymorphic ventricular tachycardia, most commonly torsades de pointes, or "twisting of points." The genetic basis for this syndrome was suggested by the association of LQTS with the Romano-Ward syndrome, an isolated autosomal dominant expression, and by the Jervell and Lang-Nielsen syndromes, an autosomal recessive form associated with congenital deafness. The LQTS can be inherited, but it also can be acquired in response to disease, injury, certain medications, metabolic disorders, or other causes. At the time of writing, mutations in seven genes have been identified and have been designated LQT1 through LQT7, with LQT1, 2, and 3 accounting for some 90% of congenital LQTS. LQT1 involves gene *KVLQT1* (also called *KCNQ1*) located on chromosome 11 and results in suppression of the outward-rectifying potassium current (i_{Ks}). This potassium current normally enhances repolarization, so its suppression will prolong the duration of the action potential, which in turn prolongs the duration of the QT interval. LQT2 involves the *HERG* or *KCNH2* gene on chromosome 7, which affects the rapidly acting component of the outward-rectifying potassium current, i_{Kr}, which also prolongs the action potential and, therefore, the QT interval. LQT3 is caused by a mutation in the sodium channel gene (*SCN5A*) located on chromosome 3, which impairs sodium channel inactivation so that the membrane remains depolarized through leakage of depolarizing sodium current into the cell. This prolongs the QT interval and promotes membrane instability. The other LQT syndromes have other mechanisms, and there are differing mutations within LQT groupings. Other channelopathies include some family groupings of atrial fibrillation, short QT syndrome, familial sick sinus syndrome, catecholaminergic polymorphic ventricular tachycardia, familial Wolff-Parkinson-White syndrome, Brugada syndrome, and progressive cardiac conduction defect of the Lenegre type.

Recording extracellular potentials. A dipole consists of two separated, equal but opposite charges. Positive charges flow along force lines from the positive pole to the negative pole, creating current and, perpendicular to the current flow, an electrical field. The separation of charge across the cell membrane discussed above can be considered to be a dipole layer. The changes in transmembrane voltages during depolarization, repolarization, and impulse propagation alter the polarity of the dipole layer in the activated region. The separation of charge between activated and quiescent tissues results in a boundary between surface dipoles. The dipoles are the source of extracellular current flow across the boundary and throughout the surrounding volume conductor of the human thorax; the distribution of the potential in the electrical field is at right angles to the current flow and is detected electrocardiographically.

The magnitude of the recorded extracellular potential is determined by the size and relationship of the boundary to the recording electrode, by the transmembrane voltage gradient across the boundary, and by a conductivity term. The conductivity term incorporates intercellular conductivity, which is determined largely by gap junctions, and extracellular conductivity, which incorporates the characteristics of the volume conductor of the thorax, skin, fat, muscle, mediastinal structures, electrode paste, and the electrode itself. Dipoles have magnitude and direction and, therefore, can be represented by spatial vectors. Multiple vectors may be summed and expressed as a single equivalent vector, which may be translated to a common origin. In classic physiology, the heart has been considered a single dipole current source, represented by a single equivalent spatial vector that is fixed in position but is allowed to vary in magnitude and direction during the cardiac cycle. More complex multidipole models have also been proposed.

An equivalent spatial vector may be projected onto a plane or, as with the electrocardiograph, onto a lead axis system. Leads detect potential differences between two or more sites. The Einthoven reference system uses bipolar leads that compare left arm and right arm, with the left arm being positive (lead I), as well as the right arm and left leg (lead II) and the left arm and left leg (lead III), with the left leg being positive. Unipolar leads compare an exploring electrode against two or three standard limb leads (Goldberger augmented limb leads and Wilson precordial leads, respectively). The spatial vector, then, can be projected onto a hexaxial lead system (Einthoven and Goldberger leads) that defines the frontal plane (superior, inferior, left, and right) and a precordial lead system (Wilson leads) that approximates a horizontal plane (anterior, posterior, left, and right). Other lead systems provide a more orthogonal, that is, less distorted, three-dimensional axis system. The Frank vectorcardiographic system, for example, uses an X lead (left, right) similar to standard lead I, a Y lead (inferior, superior) similar to the augmented left foot lead, and a Z lead (posterior, anterior) similar to an inverted second unipolar chest lead (V_2).

In electrocardiography the potential in a given lead system is displayed against time, usually with 12 separate leads: I, II, III, augmented right arm (aVR), augmented left arm (aVL), augmented left foot (aVF), and six unipolar chest leads, designated V_1 through V_6. In vectorcardiography the spatial orientation of successive vectors is displayed for the frontal, horizontal, or sagittal plane. The vectorcardiographic loop is formed by the instantaneous plot of one orthogonal lead against another, with the time frame indicated by an interruption in the inscribed loop every 2.0 or 2.5 ms.

Body surface mapping based on recordings made at many sites has been used because each recording

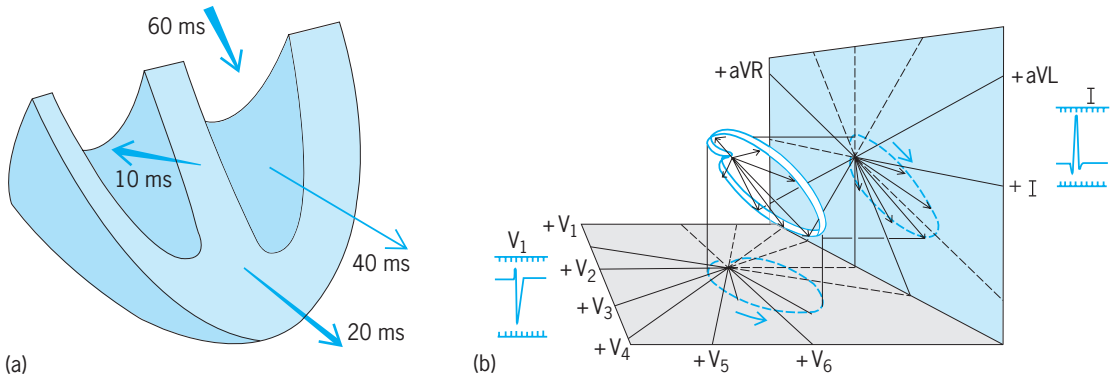

Fig. 2. Spatial vectors after onset of ventricular depolarization. (*a*) Ventricular activation at 10, 20, and 40 milliseconds. At 10 ms the subendocardial Purkinje system is depolarized, and ventricular vectorial forces are balanced except for left septal depolarization; the resultant spatial vector normally has a rightward, anterior, and superior orientation. The 20-ms vector results from activation of the septum and apex and has a leftward, anterior, and inferior orientation. The 40-ms vector arises from depolarization of the inferior and lateral walls of the left ventricle, with lesser contributions from the septum and right ventricle, and has a leftward, posterior, and inferior orientation. At 60 ms the inferior posterolateral and basal walls of the left ventricle are activated, and the spatial vector has a leftward, inferior, and even more posterior orientation. (*b*) The spatial vectors of ventricular depolarization translated in space with a common origin. The spatial QRS loop depicted is projected onto the frontal plane of the hexaxial reference system consisting of Einthoven and Goldberger leads and onto the horizontal plane created by unipolar Wilson leads. The projected loop is interrupted every 2.5 ms. The scalar recordings are shown in leads I and V_1. By convention, when the spatial vector approaches the positive terminal of the lead, the scalar recording is positive, and when directed away from the positive terminal of the lead, negative. (*After M. F. Arnsdorf, Electrocardiography, American Physiological Society, 1978*)

has both redundant and unique information, and computer analysis can extract the unique information. Signal averaging in the time domain and with Fourier analysis yields data that are not available with the standard electrocardiographic and vectorcardiographic approaches.

Extracellular potential measurements. The sinoatrial node is the normal pacemaker of the heart. Its mass is small, and therefore it does not appear on surface electrocardiograms. The sinoatrial nodal potential, however, initiates the activation of the atria. Atrial depolarization appears as the P wave on the electrocardiogram. Conduction from the atria to the ventricles normally takes place through the atrioventricular node and is quite slow. The impulse emerges from the atrioventricular node, penetrates the bundle of His and bundle branches, and enters the terminal Purkinje network. Because the tissue mass of the atrioventricular node, His bundle, bundle branches, and terminal Purkinje network is too small to produce a signal strong enough to be recorded at the body surface, the P wave is followed by an isoelectric interval.

The wavefront enters the ventricular muscle from the Purkinje network. The QRS complex results from depolarization of the large mass of ventricular muscle. It is followed by the T wave, the rapid phase of ventricular repolarization. U waves, which sometimes are recorded, may reflect repolarization of the terminal Purkinje system. The interval from atrial activation to the beginning of ventricular depolarization is known as the PR interval, and the isoelectric segment is the PR segment. The ST segment is inscribed during the plateau phase of ventricular action potential repolarization, which follows QRS depolarization, and it may be abnormal in disease states. The QT interval represents the entire depolarization and repolarization process of the ventricle. Vectorcardiography records P, QRS, and T loops in a frontal, sagittal, and transverse or horizontal plane. Time intervals are represented by interruptions in the recording. The spatial vectors in **Fig. 2** point in the direction of the net potential.

The electrocardiograph and vectorcardiograph are sensitive to abnormal sequences of depolarization such as bundle branch block and preexcitation. They often are useful in detecting increase in myocardial mass, such as atrial and ventricular enlargement. They can detect ischemia and myocardial infarction and often suggest myocardial or pericardial disease, electrolyte disturbance, and drug effects.

The heartbeat is controlled primarily by the sinoatrial node, which is normally the fastest pacemaker. A marked slowing or failure of the sinus node or a conduction defect that prevents an impulse initiated by the sinus node from reaching all of the heart usually leads to the appearance of a lower escape pacemaker to assume control of the heart. Such escape pacemakers, which have intrinsic rates slower than that found in the normal sinoatrial node, may arise in the atrium, the junction (the His bundle and the region of the atrioventricular node adjacent to it), the subjunction (fascicles and bundle branches), and the terminal Purkinje fibers. The so-called idioventricular escape rhythm probably originates in the terminal Purkinje fibers rather than in the ventricle, although pathological pacemakers may arise in the ventricular tissue. Time-dependent inward and outward currents are involved in lower pacemaker function. Abnormalities in automaticity and in impulse propagation may lead to cardiac rhythm disturbances, but electrocardiography is very useful in evaluating the rhythm and conduction of the heart. Normal electrocardiograms and vectorcardiograms are shown in **Fig. 3**.

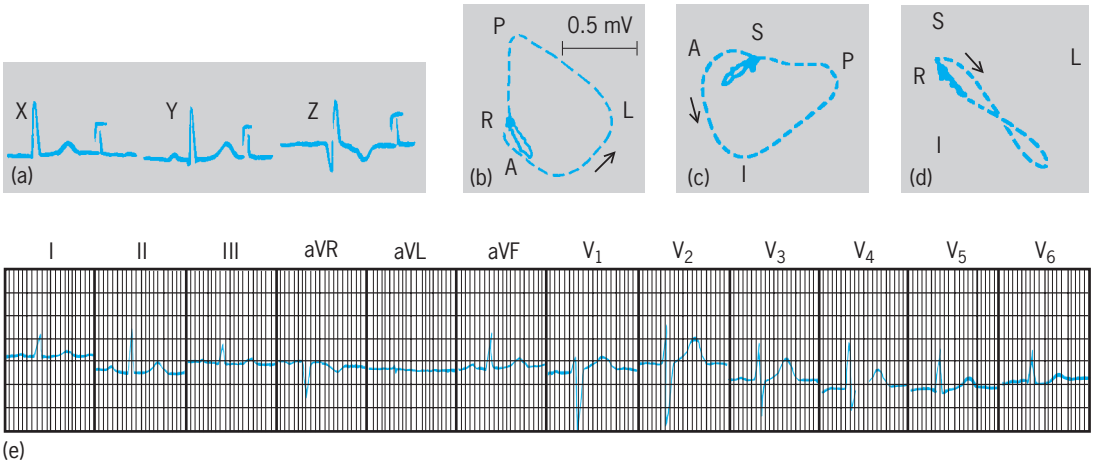

Fig. 3. Normal electrocardiograms and vectorcardiograms. (*a*) Normal orthogonal X, Y, and Z leads that are used to construct the vectorcardiogram (VCG). The X lead defines a left-right lead with the positive terminal on the left; the Y lead defines a superior-inferior lead with the positive terminal inferiority; and the Z lead defines an anterior-posterior lead with the positive terminal posteriorly. (*b–d*) Normal P, QRS, and T vectorcardiographic loops. The beam is interrupted every 2.5 ms, and the direction of inscription is indicated. Orientation in the vectorcardiograph is as follows: A, anterior; P, posterior; L, left, R, right; S, superior; I, inferior. (*b*) The horizontal plane has a left-hand (X lead) and anterior-posterior (Z lead) orientation; (*c*) the sagittal plane has a superior-inferior (Y lead) and anterior-posterior (Z lead) orientation; and (*d*) the frontal plane has a left-right (X lead) and superior-inferior (Y lead) orientation. The P loop is very small, the QRS loop is the large dominant inscription, and the T loop is the intermediate-size inscription. The P and T loops are magnified for analysis. (*e*) Normal 12-lead electrocardiogram. Note the similarity between the electrocardiographic and the orthogonal vectorcardiographic recordings: left-right (I, X), inferior-superior (aVF, Y), and the mirror image of V_1 with anterior positivity and Z with posterior positivity.

At times, pacemaker catheters are inserted into the cavities of the heart to stimulate the myocardium, provoking cardiac arrhythmias, and to record extracellular potentials. Intracavitary recordings can record small potentials from the sinoatrial node, His bundle, bundle branches, and terminal Purkinje fibers. Multiple intracavitary recording catheters permit the identification and mechanisms of reentrant arrhythmias and the localization of anomalous connections between the atria and ventricles, such as those that occur in preexcitation. An example of an intracavitary recording is shown in **Fig. 4**.

Body surface mapping and signal averaging techniques show promise for detecting abnormal potentials that indicate myocardial disease. Some studies suggest that the presence of such abnormal potentials may allow identification of the individual at risk for serious rhythm disturbances and sudden death. *See* BIOPHYSICS; CARDIOVASCULAR SYSTEM; VASCULAR DISORDERS.

Arrhythmogenesis. Heart rhythm disturbances result most commonly from reentry, but may also result from enhanced excitability, increased automaticity, and triggered activity. Reentry can be defined as the reexcitation of previously activated tissue by a continuously propagated electrical impulse. The model, such as that in **Fig. 5**, usually involves an area of slow conduction and unidirectional block. In this model of a Purkinje network, the impulse would normally travel from the central Purkinje fiber (1) down the two branches (2, 3) and activate the ventricular muscle (4) [Fig. 5*a*]. As a result of ischemia or injury, a segment with unidirectional block and slow retrograde conduction is established in one limb of the Purkinje fiber (3) [Fig. 5*b*]. The impulse descending from the central Purkinje is blocked proximally by the area of unidirectional block, travels normally through the other branch (2), enters the ventricular muscle (4), and reenters the depressed limb (3) distally. The impulse conducts slowly through the depressed segment and, if the tissue at the bifurcation and in limb (2) has regained its excitability, it can be reexcited (reentered), once more reaching the ventricular muscle through limb (2). The

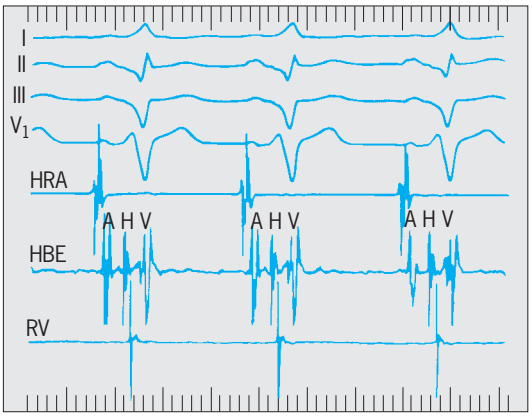

Fig. 4. Body surface and intracavitary electrocardiographic recordings. The electrocardiogram body surface leads I, II, III, and V_1 are shown in the four traces at the top. A catheter with recording electrodes was advanced through the femoral vein, the inferior vena cava, the right atrium, and right ventricle. Recordings were made in the high right atrium (HRA), in the area of the tricuspid valve (HBE, His bundle electrogram), and the right ventricle (RV). Note that the HBE records the atrial (A), His bundle (H), and ventricular (V) depolarization. The ventricular electrogram in the HBE is primarily due to left ventricular depolarization, is delayed because of a conduction block in the left bundle branch, and occurs after the right ventricular depolarization. (*Courtesy of Dr. Thomas Bump, University of Chicago*)

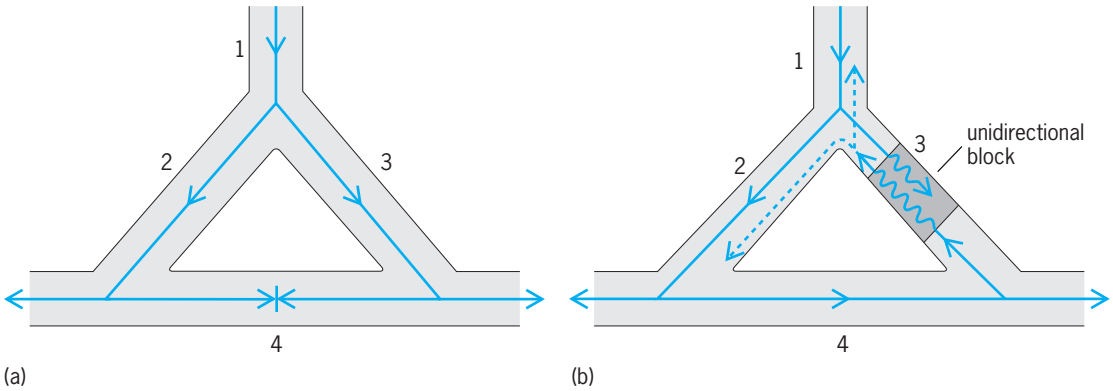

Fig. 5. Model of a Purkinje network. (*a*) Normal propagation. (*b*) Reentrant circuit with a depressed segment demonstrating slow conduction and unidirectional block.

reentry may be a single beat, or it may be sustained. The success or failure of reentry depends on the length of the functionally isolated depressed segment, the integrity of cell-cell communication through gap junctions, the conduction velocity, and the ionic mechanisms of depolarization and repolarization both within the depressed segment and the normal tissue that is to be reactivated, and the refractoriness of the previously excited tissue. These pathways may be anatomic or functional and may exist anywhere in the heart. The reentrant circuits may be large (macroreentry)—as in atrial flutter, atrioventricular nodal reentrant tachycardia, atrioventricular tachycardia that utilizes an accessory pathway, and ventricular—or they may be very small (microreentry) as in atrial and ventricular fibrillation.

Normal automaticity may be enhanced as with sinus tachycardia, or abnormal foci of spontaneously depolarizing tissue may result from ischemia, autonomic surges, electrolyte imbalance, hypoxia, or drugs, resulting in atrial, junctional, and ventricular tachycardias.

Unusual oscillatory activity may be triggered by normal or abnormal activation. As mentioned above, the long QT syndrome is characterized by prolongation of the action potential, which is manifested by QT interval prolongation on the surface electrocardiogram. The prolongation of the action potential allows more time for inward calcium current to flow, which, in turn, may load the cell with calcium and lead to the development of oscillatory afterpotentials. These afterpotentials likely are the cause of the polymorphic ventricular tachycardia called torsades de pointes. At times, tissue may demonstrate increased excitability as a result of disease or external influences.

Clinical cardiac electrophysiology. Patients may receive permanent electronic pacemakers because of slow heart rates, which often cause dizziness, lightheadedness, poor exercise tolerance, and sometimes loss of consciousness. The slow rate may result from a slowing of the normal sinus rhythm with the failure of lower spontaneous pacemakers to appear, or from a conduction defect in which the normal input of the sinus node to the ventricles is interrupted due to block, usually in the AV node, His bundle, or bundle branches. The electronics in these artificial pacemakers has become quite sophisticated in that most can record and pace from the atrium or ventricles, can recognize slow and sometimes fast rates, and can switch from one mode of pacing to another depending on the arrhythmia or conduction abnormality that is encountered. Some pacemakers can synchronize the depolarization of the two ventricles when one is activated before the other as may occur in bundle branch block, and such cardiac resynchronization therapy has proven effective in the treatment of heart failure in some patients.

By permitting the recording of intracavitary electrical potentials, intracardiac catheters can be used to identify the mechanisms of an arrhythmia. Moreover, radio-frequency (RF) energy, which is low-voltage high-frequency electrical energy, can be delivered from the tip of a catheter to the endocardial (inside) surface of the heart to "burn" and scar the site at which an arrhythmia is arising or to interrupt a reentrant pathway. This type of RF ablation has revolutionized the treatment of macroreentrant arrhythmias such as atrial flutter, atrial tachycardias, AV nodal reentrant tachycardia, AV reentrant tachycardia using accessory pathways, conduction of atrial arrhythmias through accessory pathways, ventricular tachycardia, and other rhythms. Progress is also being made in the ablation of atrial fibrillation.

Sudden cardiac death usually results from ventricular fibrillation (VF) that sometimes is heralded by the appearance of ventricular tachycardia (VT). The external application of direct-current (DC) energy has been used since 1962 to cardiovert atrial fibrillation, atrial flutter, ventricular tachycardia, ventricular fibrillation, and other tachyarrhythmias to a normal rhythm. Electronic miniaturization and modern computer technology have allowed the development of the implantable cardioverter-defibrillator (ICD), which consists of sensing electrodes, defibrillation electrodes, and a pulse generator. The modern ICD can recognize an arrhythmia, attempt to convert a reentrant rhythm such as ventricular tachycardia to a normal rhythm by pacing, or, if necessary, defibrillate the heart as may be required with a refractory ventricular tachycardia or with ventricular

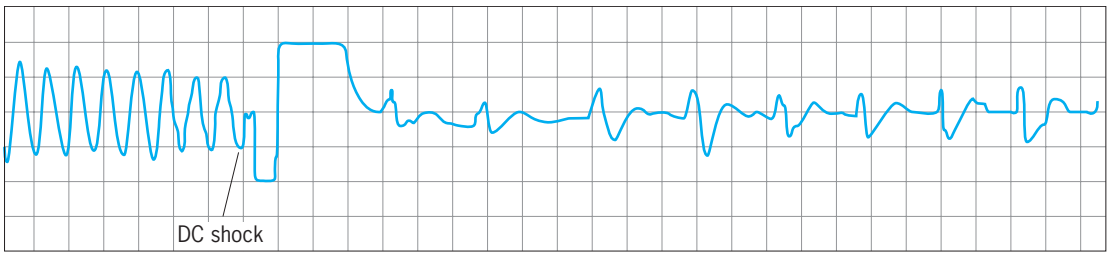

Fig. 6. Appropriate delivery of a DC shock by an implantable defibrillator. (*Courtesy of Dr. Bradley Knight, University of Chicago*)

fibrillation. An example is shown in **Fig. 6** in which a very rapid ventricular rhythm (ventricular flutter) is recognized, a DC shock is delivered, and the rhythm is converted to a normal sinus rhythm with a wide QRS complex. Current ICD lead systems are usually placed transvenously through an axillary, subclavian, or cephalic vein. The ICD has become the first line of defense in protecting the patient with heart disease from sudden cardiac death due to ventricular fibrillation. Morton Arnsdorf

Bibliography. M. F. Arnsdorf, *Electrocardiography: I. Fundamental Theory,* and *II. Applied Theory*, 1978; M. F. Arnsdorf and J. C. Makielski, Excitability and impulse propagation, in *Physiology and Pathophysiology of the Heart*, 4thd ed., ed. by N. Sperelakis, 2000; B. Hille, *Ion Channels of Excitable Membranes*, 3d ed., 2001; M. E. Josephson, *Clinical Cardiac Electrophysiology: Techniques and Interpretations*, 3d ed., 2001; F. Lehmann-Horn and K. Jurkatt-Rott (eds.), *Channelopathies*, 2000; P. W. Macfarlane and T. D. V. Lawrie (eds.), *Comprehensive Electrocardiography*, 3 vols., 1989; D. Zipes and J. Jalife (eds.), *Cardiac Electrophysiology: From Cell to Bedside*, 4th ed., 2004.

Cardinal points

The four intersections of the horizon with the meridian and with the prime vertical circle, or simply prime vertical, the intersections with the meridian being designated north and south, and the intersections with the prime vertical being designated east and west (see **illus.**). The cardinal points are 90° apart; they lie in a plane with each other and correspond to the cardinal regions of the heavens. The four intermediate points, northeast, southeast, northwest, and southwest, are the collateral points. *See* ASTRONOMICAL COORDINATE SYSTEMS.

Frank H. Rockett

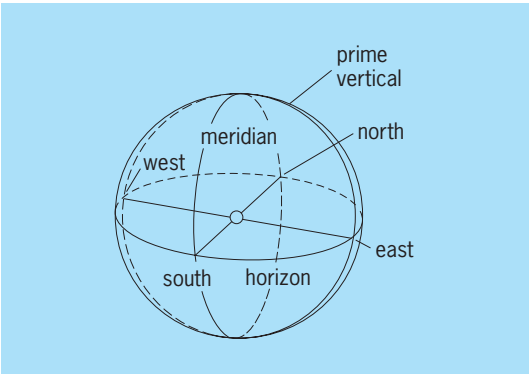

Cardinal points around horizon.

Cardioid

A heart-shaped curve generated by a point of a circle that rolls (without slipping) on a fixed circle of the same diameter. In point-wise construction of the curve, let O be a fixed point of a circle C of diameter a, and Q a variable point of C. Lay off distance a along the secant OQ, in both directions from Q. The locus of the two points thus obtained is a cardioid (see **illus.**). If a rectangular coordinate system

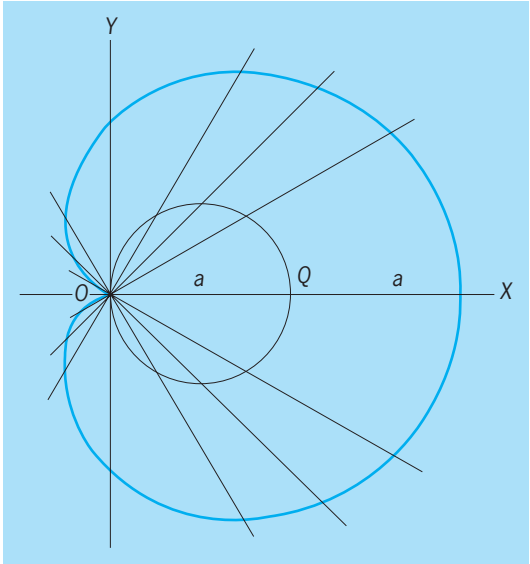

Cardioid. Symbols explained in text.

is chosen with O for origin initially and y axis tangent to C at O, the cardioid has the equation $(x^2 + y^2 - ax)^2 = a^2(x^2 + y^2)$. The equation expressed in polar coordinates is $\rho = a(1 + \cos\theta)$. The cardioid's area is $3/2\pi a^2$, or six times the area of C, and its length is $8a$. *See* ANALYTIC GEOMETRY. Leonard M. Blumenthal

Cardiovascular system

Those structures, such as the heart, or pumping mechanism, and the arteries, veins, and capillaries, which provide channels for the flow of blood. The cardiovascular system is sometimes called the blood-vascular system. The circulatory system includes both the cardiovascular and lymphatic systems; the latter consists of lymph channels (lymphatics), nodes, and fluid lymph which finally empties into the bloodstream. This article discusses the cardiovascular system under the following major headings: comparative anatomy, comparative embryology, functional development of the heart, human fetal circulation at term, postnatal circulation, and comparative physiology. *See* BLOOD; LYMPHATIC SYSTEM.

Comparative Anatomy

All vertebrates feature a closed system of branched vessels, a ventral heart, and a basic pattern of organization that ranges from the single system in most fishes to the double system in land forms.

Heart. The hearts of vertebrates differ from those of animals in the lower phyla in their ventral rather than dorsal location. Blood is pumped anteriorly through arteries and forced to the dorsal side. The greater part then courses posteriorly through arteries which terminate in capillaries in various parts of the body. The blood returns to the heart through veins. *See* HEART (VERTEBRATE).

Phylogeny. In lower vertebrates the heart is located far forward in the body, but there is a gradual backward shifting as the vertebrate scale is ascended. The heart lies in a pericardial cavity which is surrounded by an investing membrane, the pericardium. The pericardial cavity is a portion of the coelom which has been separated from the remainder of the body cavity. In elasmobranchs, the separation is incomplete and the two portions of the coelom are connected by a pericardio-peritoneal canal. A thin serous membrane, the epicardium, covers the surface of the heart. It is continuous with the lining of the pericardial cavity.

The heart is primarily a pulsating tube, the lining (endocardium) of which is derived from fusion of two vitelline veins. Its muscular wall (myocardium) and the epicardium originate from surrounding splanchnic mesoderm. It becomes divided into chambers called atria and ventricles. In addition, two accessory chambers, called the sinus venosus and conus arteriosus, respectively, may be present. Cyclostomes and most fishes have two-chambered hearts with one atrium and one ventricle. A sinus venosus connects with the atrium and a conus arteriosus leads from the ventricle. Dipnoans and amphibians have three-chambered hearts with two atria and one ventricle. The three-chambered heart of most reptiles is similar to that of amphibians, but an incomplete partition appears in the ventricle. In crocodiles and alligators this becomes complete, forming a four-chambered heart with two atria and two ventricles. Birds and mammals have four-chambered hearts. Valves are present to regulate the direction of blood flow.

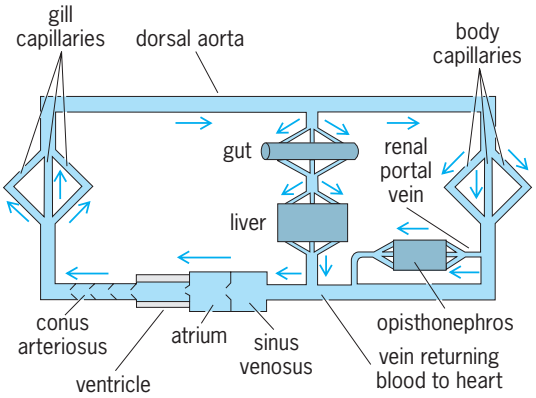

Fig. 1. Single type of circulatory system. (*After C. K. Weichert, Anatomy of the Chordates, 3d ed., McGraw-Hill, 1965*)

Circulation in the heart. Vertebrates with two-chambered hearts have the single type of circulation (**Fig. 1**) in which only unoxygenated blood passes through the heart, which pumps it to the gills for aeration. The double type of circulation (**Fig. 2**) exists in three- and four-chambered hearts through which pass two streams of blood, one oxygenated and the other unoxygenated or partly oxygenated. Even in three-chambered hearts these streams do not mix to any appreciable extent. Partitioning of the heart has been associated in evolution with development of pulmonary circulation accompanying the appearance of lungs.

The conus arteriosus no longer exists as such in adults of higher forms; it is split into trunks leading to the aorta and lungs, respectively.

In those forms having two-chambered hearts unoxygenated blood from all parts of the body is collected by the sinus venosus which joins the atrium. In the three-chambered hearts of dipnoans and amphibians the sinus venosus has shifted its position and joins the right atrium. A large sinus venosus is present in certain reptiles, but in most it is very small or is

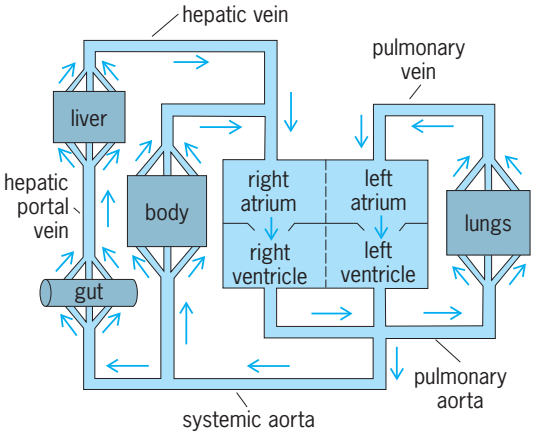

Fig. 2. Double type of circulatory system in a vertebrate having a four-chambered heart (ventral view). (*After C. K. Weichert, Anatomy of the Chordates, 3d ed., McGraw-Hill, 1965*)

lacking. Birds and mammals possess a sinus venosus only during early embryonic development. In vertebrates having three- or four-chambered hearts, unoxygenated blood is returned to the heart through the sinus venosus or directly to the right atrium, as the case may be.

Arterial system. Although the arterial systems of various adult vertebrates appear to be very different in arrangement, a study of development reveals that all systems are built upon the same fundamental plan.

Ventral and dorsal aortas. In lower vertebrates the conus arteriosus leads forward to a ventral aorta. During early development this vessel divides anteriorly into two aortic arches which course dorsally in the mandibular region. These continue posteriorly as the paired dorsal aortas. Additional pairs of aortic arches then appear, forming connections between ventral and dorsal aortas on each side. They appear in sequence in an anteroposterior direction, each coursing through the tissues between adjacent pharyngeal pouches. The typical number of aortic arches to form in vertebrates is six pairs (**Fig. 3***a*), although there are certain discrepancies among lower forms. The paired dorsal aortas unite posterior to the pharyngeal region to form a single vessel. This single dorsal aorta continues posteriorly into the tail as the caudal artery.

Various paired and unpaired arteries arise along the length of the dorsal aorta to supply all structures of the body posterior to the pharyngeal region. Anterior continuations of the ventral aorta and the anterior remnants of the paired dorsal aortas supply the head and anterior branchial regions. Although the vessels arising from the dorsal aorta are fairly uniform throughout the vertebrate series, the aortic arches undergo profound modifications. The changes are similar in members of a given class.

Blood, pumped anteriorly by the heart, passes to the aortic arches. These vessels then carry the blood to the dorsal aorta or aortas from which it goes either anteriorly to the head or posteriorly to the remainder of the body.

Aortic arches. Changes in the aortic arch region constitute the chief differences in the arterial systems of the separate vertebrate classes. A progressive reduction in the number of aortic arches occurs as the evolutionary scale is ascended. In cyclostomes and in most fishes, in which gills are used in respiration, the aortic arches break up into afferent and efferent portions connected by numerous gill capillaries in which blood is oxygenated.

The changes in this region involve primarily a routing of the blood through certain preferred aortic arches with a consequent atrophy or disappearance

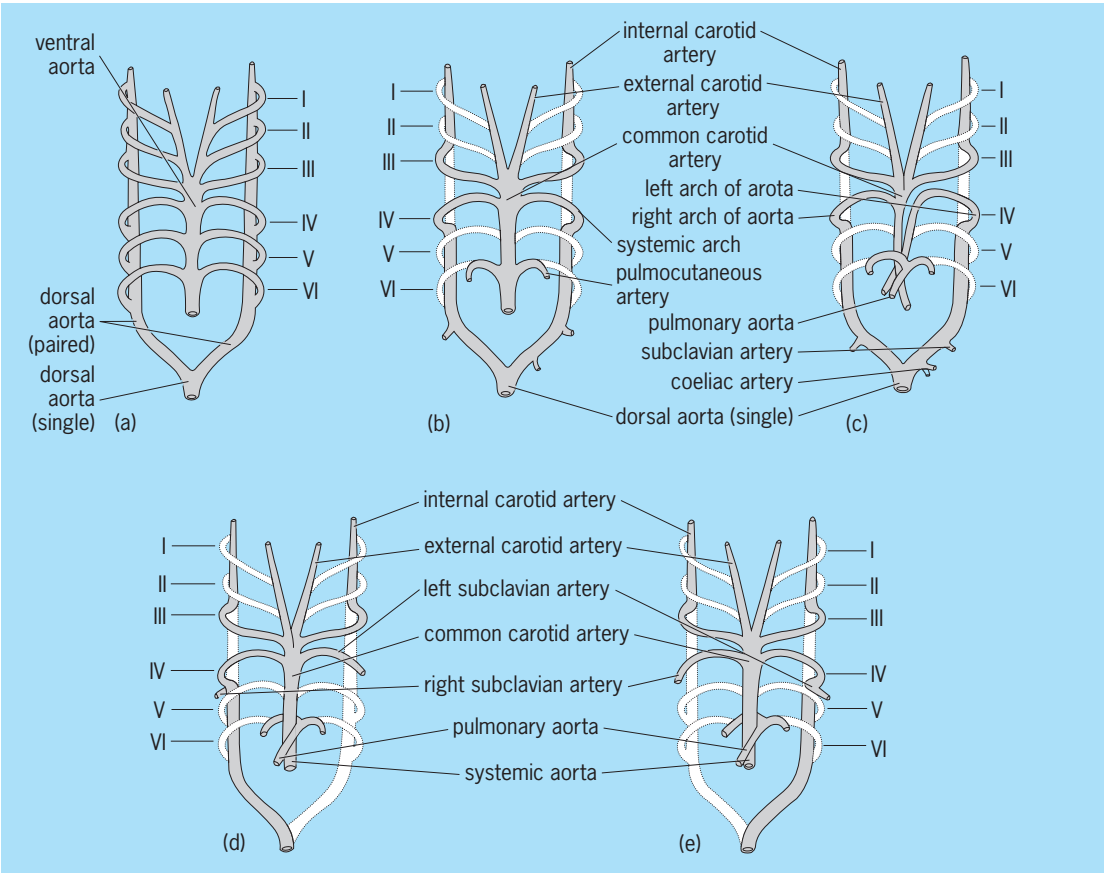

Fig. 3. Modifications of aortic arches in vertebrates, ventral views. (*a*) Typical condition in vertebrate embryos. Six pairs of arches connect dorsal and ventral aortas. (*b*) In anuran amphibians. (*c*) In reptiles. Ventral aorta has split into three vessels, right and left systemic trunks and a pulmonary trunk or aorta. (*d*) In birds. (*e*) In mammals. (*After C. K. Weichert, Anatomy of the Chordates, 3d ed., McGraw-Hill, 1965*)

of those which are no longer used, and a splitting of the conus arteriosus and ventral aorta, particularly in reptiles, birds, and mammals, in such a manner that systemic and pulmonary trunks are established (Fig. 3*b–e*). The systemic trunk, coming from the left ventricle (or left side of the ventricle), distributes oxygenated (or partially oxygenated) blood to the body in general. The pulmonary trunk, coming from the right ventricle (or right side of the ventricle), carries blood to the lungs. Oxygenated blood returns through veins from the lungs to the left atrium. Vessels which carry blood from the heart to the lungs and back constitute the pulmonary circulation. Branches of the sixth pair of aortic arches give rise to the pulmonary arteries.

Coronary arteries, supplying the tissues of the heart itself, arise as branches of certain aortic arches or of the systemic trunk near the point where it emerges from the heart.

Somatic arteries. The arteries supplying the body proper are usually paired structures which clearly show evidences of segmental arrangement (metamerism). They spring from the dorsolateral regions of the dorsal aorta. Fusion of two or more segmental arteries may take place, obscuring the fundamental metameric arrangement. *See* METAMERES.

Visceral arteries. Arteries supplying the viscera are of two kinds, paired and unpaired. The paired arteries, usually restricted to certain regions, are segmentally arranged. They supply structures derived from the part of the embryo from which the urogenital organs and their ducts arise.

Unpaired visceral arteries supply the spleen and digestive tract. There are usually three such vessels in vertebrates. The celiac artery supplies the stomach, spleen, pancreas, liver, and duodenum; the superior mesenteric artery supplies most of the intestine and a portion of the pancreas; and the inferior mesenteric artery goes to the posterior part of the large intestine and rectum. Variations from this condition are accounted for by fusions or separations.

Venous system. A comparison of veins in the various vertebrate groups shows that they too are arranged according to the same fundamental plan and that the variations encountered form a logical sequence as the vertebrate scale is ascended. In its development the venous system of higher forms passes through certain stages common to the embryos of lower forms.

Sinus venosus. The accessory chamber present in the hearts of lower vertebrates through which blood from all over the body is returned to the heart is the sinus venosus.

In cyclostomes and fishes the sinus venosus typically receives a pair of common cardinal veins, or ducts of Cuvier, formed by the union of anterior and posterior cardinals, a pair of small inferior jugular veins from the ventrolateral part of the head, and a pair of hepatic veins coming from the liver. In lower fishes a lateral abdominal vein also joins each duct of Cuvier. Each lateral abdominal vein receives a subclavian vein from the pectoral fin and an iliac vein from the pelvic appendage (**Fig. 4**). In some fishes the inferior jugular veins are lacking; in others they have fused. Teleost fishes lack lateral abdominal veins; rather, the subclavians enter the common cardinals. In the dipnoan *Epiceratodus*, the lateral abdominal veins fuse to form a single anterior abdominal vein which joins the sinus venosus directly.

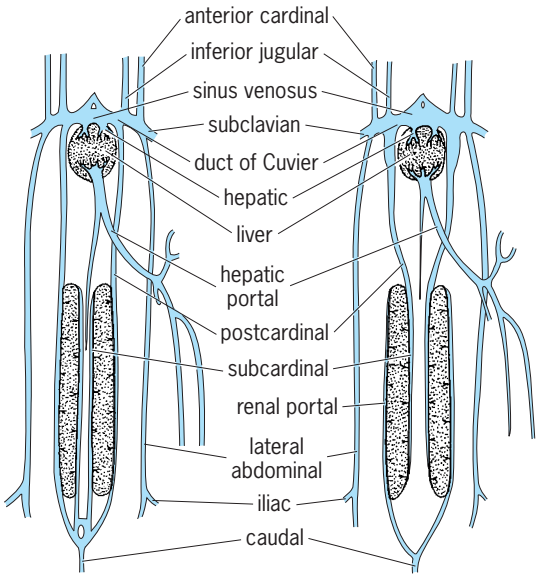

Fig. 4. Changes over the primitive condition which occur in the venous system of lower fishes, ventral view. (*After C. K. Weichert, Anatomy of the Chordates, 3d ed., McGraw-Hill, 1965*)

The sinus venosus in amphibians is conspicuous. The ducts of Cuvier have consolidated to form two precaval veins, each of which, in addition to the usual vessels, receives a large musculocutaneous vein from the skin and body wall before joining the sinus venosus. A postcava, derived partly from the sinus venosus and partly from the vitelline veins, joins the sinus venosus posteriorly.

A large sinus venosus is present in turtles; in most reptiles, however, it has been greatly reduced and much of it may be incorporated within the wall of the right atrium. In such reptiles and in adult birds and mammals which lack a sinus venosus, precaval and postcaval veins enter the right atrium directly. Valves that are present where veins enter the right atrium represent vestiges of the sinus venosus.

Cardinal veins. Anterior cardinal, or jugular, veins are prominent in cyclostomes and fishes. In amphibians, reptiles, birds, and mammals they receive internal and external tributaries before joining the precaval veins. In some mammals a single precava is present, an anastomosis having developed between the two original vessels.

Postcardinal veins, which originally course along the lateral borders of the kidneys, at first carry blood away from the kidneys. In fishes they separate into anterior and posterior portions. The posterior sections terminate in the kidneys to become the renal portal veins. The anterior sections connect with a new pair of vessels, the subcardinals, which appear along the medial borders of the kidneys. These now

carry blood from the kidneys to the heart. In caudate amphibians and a few salientians the condition is similar to that of fishes, but in most salientians the anterior portions of the postcardinals disappear, and the blood from the kidneys reaches the sinus venosus by way of the postcava. In reptiles and birds the renal portal veins have lost most of their significance, and in mammals they disappear entirely. A remnant of the anterior end of the right postcardinal of mammals becomes the azygos vein, a branch of the right precava which drains the intercostal muscles.

Abdominal veins. The lateral abdominal veins of amphibians and reptiles, like those of the lungfish *Epiceratodus*, have fused to form a single anterior abdominal vein. Instead of entering the sinus venosus, however, it joins the hepatic portal vein. In reptiles, birds, and mammals the lateral abdominal veins are represented during embryonic life by allantoic or umbilical veins which lose their direct connections with the sinus venosus when the liver develops. In adult birds the anterior abdominal vein may possibly be represented by the coccygeomesenteric vein or the epigastric vein. It disappears entirely in all mammals, with the exception of the echidnas.

Portal systems. A portal system is a system of veins that breaks up into a capillary network before the blood which courses through it is returned to the heart. All vertebrates have a hepatic portal system in which blood collected from the digestive tract and spleen passes through capillaries (sinusoids) in the liver before reaching the heart. The embryonic vitelline (subintestinal) vessels are represented in adults by the hepatic portal vein and its tributaries. The renal portal system of lower vertebrates has already been mentioned. Another small but important portal system is found associated with the blood vessels draining the pituitary gland. *See* LIVER.

Miscellaneous. Coronary veins draining the tissues of the heart enter the sinus venosus in lower forms or the right atrium in higher vertebrates. Pulmonary veins, found in amphibians, reptiles, birds, and mammals, first appear in the dipnoan fishes, where their function is to drain the swim bladder. They enter the left atrium. *See* RESPIRATORY SYSTEM; SWIM BLADDER.

Charles K. Weichert

Comparative Embryology

The cardiovascular system in vertebrates arises from the splanchnic mesoderm, with the first blood and vessels formed in the wall of the yolk sac.

Heart. The heart of each species attains its characteristic morphology through an orderly series of changes which begin in early embryonic life. Although the details of development differ for different animals, there are a number of basic principles which apply to all species. Emphasis will be placed on the general principles first, and species characteristics will be noted subsequently.

The prospective heart rudiment, that is, the tissue which eventually differentiates into the heart, can be located by experimental procedures (exploration, transplantation, vital dyes, and autoradiographs) before morphological differentiation is observable. The rudiment consists of bilateral areas of splanchnic mesoderm located near the blastopore in the early amphibian gastrulae and near the pharyngeal arches just prior to initial morphological differentiation. *See* AUTORADIOGRAPHY.

The heart field, the mesoderm with potency for heart development, is more extensive than the prospective rudiment which normally forms the organ. For example, a normal heart can develop in amphibians after complete removal of a visibly differentiated primordium. This results from migration of mesoderm from the periphery of the field to the region of extirpation.

The heart field, like the prospective rudiment, is bilateral. The potency for cardiac development is greatest at the center of the field and least at the periphery. The limits of the fields have been determined precisely in chick embryos by studies on the fates of small areas transplanted to the chorioallantoic membrane. *See* FATE MAPS (EMBRYOLOGY).

Induction. The prospective rudiment in amphibian embryos becomes destined (determined) for heart development at the time of gastrulation through an inductive effect exerted by tissue adjacent to the blastopore. Endoderm has a further inductive role on the heart-forming mesoderm at the neurula stage, as evidenced by lack of heart development following extirpation of the endoderm. *See* EMBRYONIC INDUCTION.

Determination and self-differentiation. The prospective rudiment is determined early, as outlined above. This is tested by transplantation to abnormal positions and by explantation as in tissue culture. The rudiment is described as self-differentiating when it attains the ability to develop in a foreign environment.

Polarity. Polarity or axial determination occurs at different times for the different axes. For example, when the anteroposterior axis is changed by 180° in neurula-stage amphibians, the heart develops with reversed morphology and beat. When the early rudiment is rotated around the long axis, altering only the dorsoventral and ventrolateral axes, the heart develops normally.

Totipotency. The ability of part of an organ rudiment to develop the whole organ is called totipotency. Experiments on amphibians, birds, and mammals show that either of the bilateral rudiments is capable of developing a complete organ. Furthermore, a whole organ may form from less than one-half; that is, several hearts can form from one rudiment.

Tubular heart formation. The heart is usually described as tubular during its early morphological differentiation, although it is saccular in shape in most species.

Mesenchymal cells arising from the splanchnic mesoderm of the heart rudiment differentiate into endocardium continuous with the subintestinal blood capillaries. The endocardial channels are bilateral at first; they eventually fuse into a single median tubular endocardium (**Fig. 5***a* and *b*).

The ventral portion of the splanchnic mesoderm of each bilateral heart rudiment thickens and differentiates into epimyocardium around the

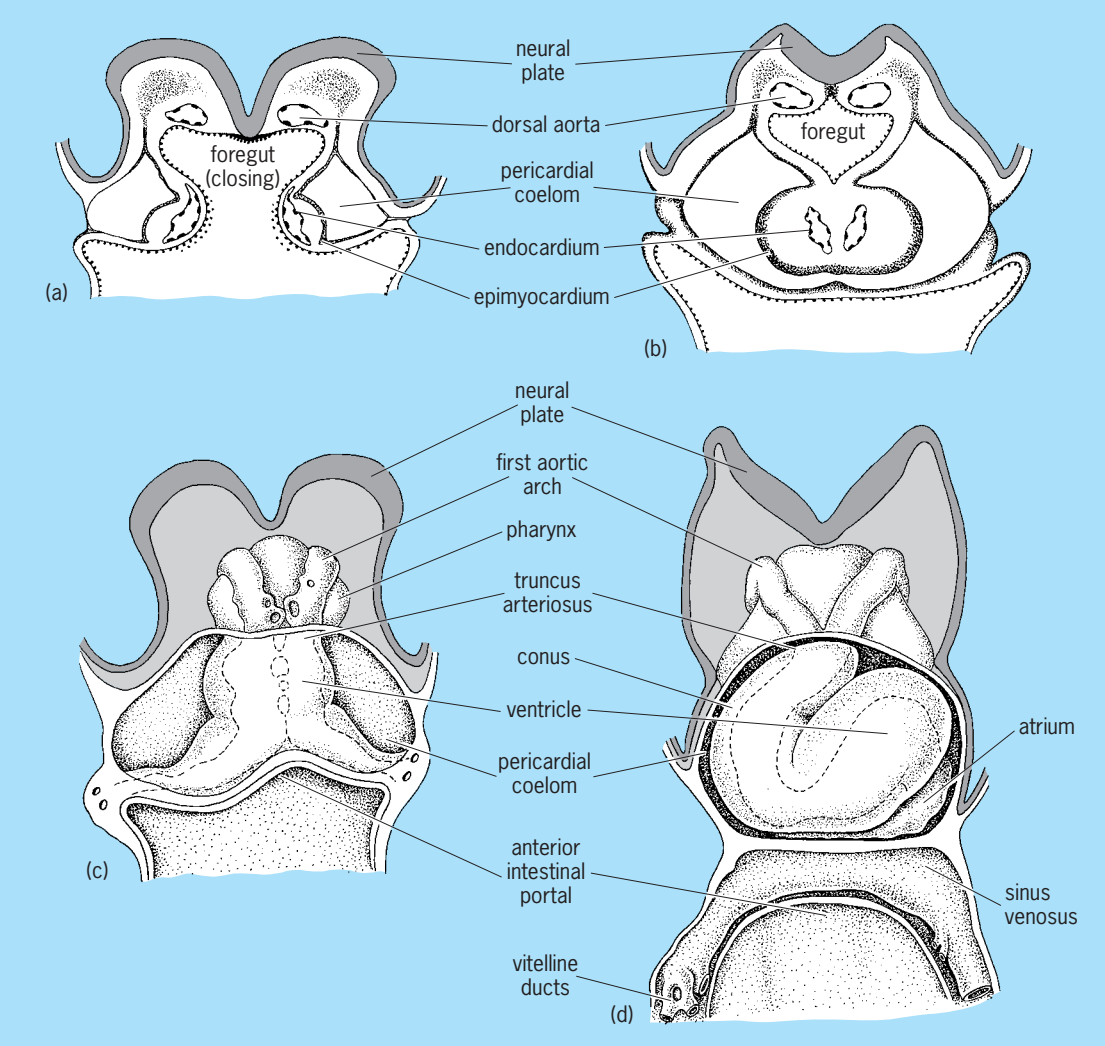

Fig. 5. Stages in development of the human heart. The organ is depicted as translucent, with the endocardium (broken lines) visible through the myocardium. (*a*) Cross section from a four-somite embryo. (*b*) Cross section at seven somites. (*c*) Ventral view at four somites. (*d*) Ventral view at thirteen somites.

endocardium (Fig. 5). A space between endocardium and myocardium during early development is filled with homogeneous material known as cardiac jelly. It is eventually replaced by connective tissue uniting the endocardium and myocardium.

The space between the splanchnic and somatic mesoderm of the heart rudiment becomes the pericardial cavity, and a part of the somatic mesoderm differentiates into parietal pericardium.

The bilateral portions of the tubular heart approach the midline and unite to form a single tubular or saccular heart (Fig. 5). Fusion progresses in a cephalocaudal direction. Likewise, structural and functional differentiation of the chambers progresses cephalocaudally. These chambers, beginning with the most caudal one, are the sinus venosus, atrium, ventricle, and bulbus cordis (conus). The sinus venosus receives the blood from the veins; the bulbus empties into the truncus arteriosus and thence into the aortic arches.

The amount of morphological differentiation seen in the bilateral heart rudiments when they reach the midventral region of the body differs for different species. For example, the bilateral rudiments of the amphibian heart are only slightly differentiated at this time, whereas those of birds and mammals show considerable differentiation prior to union in the midline (Fig. 5*a* and *c*). This is because the heart of birds and mammals develops relatively early, while the mesoderm is spread out over the yolk sac. The two sides cannot come together until the head end of the body grows and folds off from the yolk sac. An early development of the circulatory system is particularly important in mammalian embryos since they have very little yolk and come to depend very early upon metabolic exchange with the mother through a placental circulation. It is significant that the heart is the first organ to function.

Cardiac loop and regional development. The tubular heart is attached by its vascular trunks and by a continuity of visceral pericardium with the serous portion of the parietal pericardium at its venous and arterial ends. It is attached by a dorsal pericardium for only a brief period.

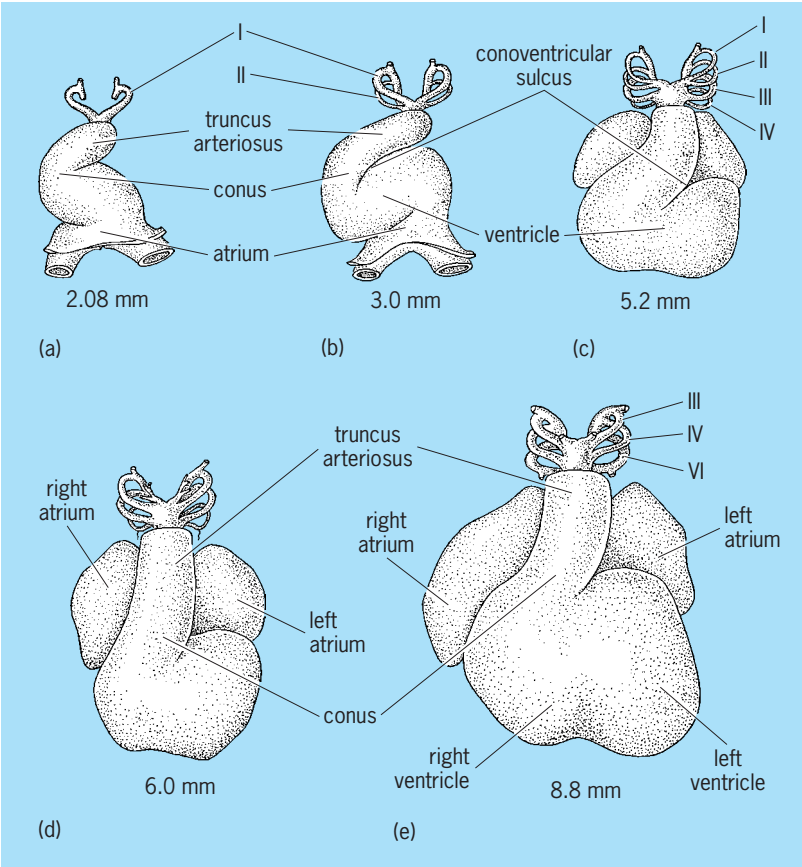

Fig. 6. Ventral views (*a–e*) of the human embryonic heart, showing the bending of the cardiac tube and the establishment of its regional divisions. The roman numerals I to VI indicate the aortic arches of corresponding numbers. (*After T. C. Kramer, Amer. J. Anat., 71(3):343–370, 1942*)

The heart becomes curved primarily due to its rapid growth. The heart increases in length so much faster than the cavity in which it lies that it first makes a U-shaped bend to the right and is then twisted into a loop. This moves the venous end cephalically and the arterial end caudally until the entering and exiting vessels approximate the same level in mammals (**Figs. 6***e* and **7***f*).

With the formation of the cardiac loop during the fifth week of human gestation, the primary regional divisions become clearly recognizable (Fig. 6). This is a result of the curvature and of different intrinsic growth rates in different regions. Blood flow also plays a part in the regional differentiation.

The sinus venosus is the thin-walled most caudal chamber into which the great veins enter (Fig. 7*c*). The atrial region is established by transverse dilation of the heart just cephalic to the sinus venosus, and the ventricle is formed by the bent midportion of the original cardiac tube (Figs. 5, 6, and 7). Between the atrium and ventricle, the heart remains relatively undilated to form the atrioventricular canal. The conus is distinguishable as a separate region prior to its later incorporation into the right ventricle. The most cephalic part of the cardiac tube undergoes the least change in appearance, persisting as the truncus arteriosus connecting the conoventricular region to the aortic arches (Figs. 5 and 6).

Studies on amphibians, birds, and mammals show that initial cardiac contractions begin in the ventricular region. The heart at this time is just slightly less developed than that illustrated in Fig. 6 for a human embryo 0.08 in. (2 mm) long (beginning of fourth week of development). With further cardiac differentiation in a cephalocaudal direction, the atrium and sinus venosus acquire the ability to contract, each with a successively higher intrinsic rate which dominates its predecessor.

Innervation of the heart occurs subsequent to the time when the sinus differentiates and takes over the pacemaker function. The fact that the embryonic heart pulsates rhythmically and carries on circulation prior to innervation indicates the beat is myogenic rather than neurogenic in origin.

The myocardium shows cytological differentiation relatively early. Electron micrographs of developing heart muscle cells show randomly oriented myofilaments (actin and myosin) prior to the time when striations appear due to the alignment of the filaments in register in myofibrils. A compact, well-oriented arrangement of myofibrils, each composed of many myofilaments, occurs earlier in the ventricle than in the atrium. The sinus venosus (shown in **Fig. 8** for the tadpole heart just prior to the time of innervation) retains an irregular arrangement in adult amphibians. In mammals, where the sinus venosus becomes incorporated into the right atrium, the sinoatrial node is composed of cells which have sparse randomly oriented myofibrils, like those which are found in the amphibian sinus venosus. The irregular arrangement is one of the characteristics of pacemaker cells. *See* HEART (INVERTEBRATE); MUSCLE.

Formation of definitive heart. The morphological changes which occur in the heart while it differentiates from a tubular structure to its definitive adult form vary for different classes and species of vertebrates in correlation with different functional requirements. The changes in elasmobranch fishes are few and simple. The heart retains all its primitive chambers in sequence (sinus, atrium, ventricle, and conus) and each chamber remains unpartitioned.

In adult amphibians the heart attains an external form and curvature beyond that illustrated in Fig. 5 for a mammalian embryo, but it retains the chambers in sequence. In urodele amphibians there is a partial division of the atrium into right and left sides; in anuran amphibians the atrial division is complete, but other parts remain undivided.

The heart of reptiles has the atrium divided completely into right and left chambers while the ventricle is divided only partially.

In birds and mammals the sinus venosus is incorporated into the right atrium, which is divided completely from the left atrium; the primitive ventricle is completely divided into right and left chambers; the conus is incorporated into the right ventricle; and the truncus arteriosus is divided into right and left sides to form the roots and proximal portions of the pulmonary artery and aorta.

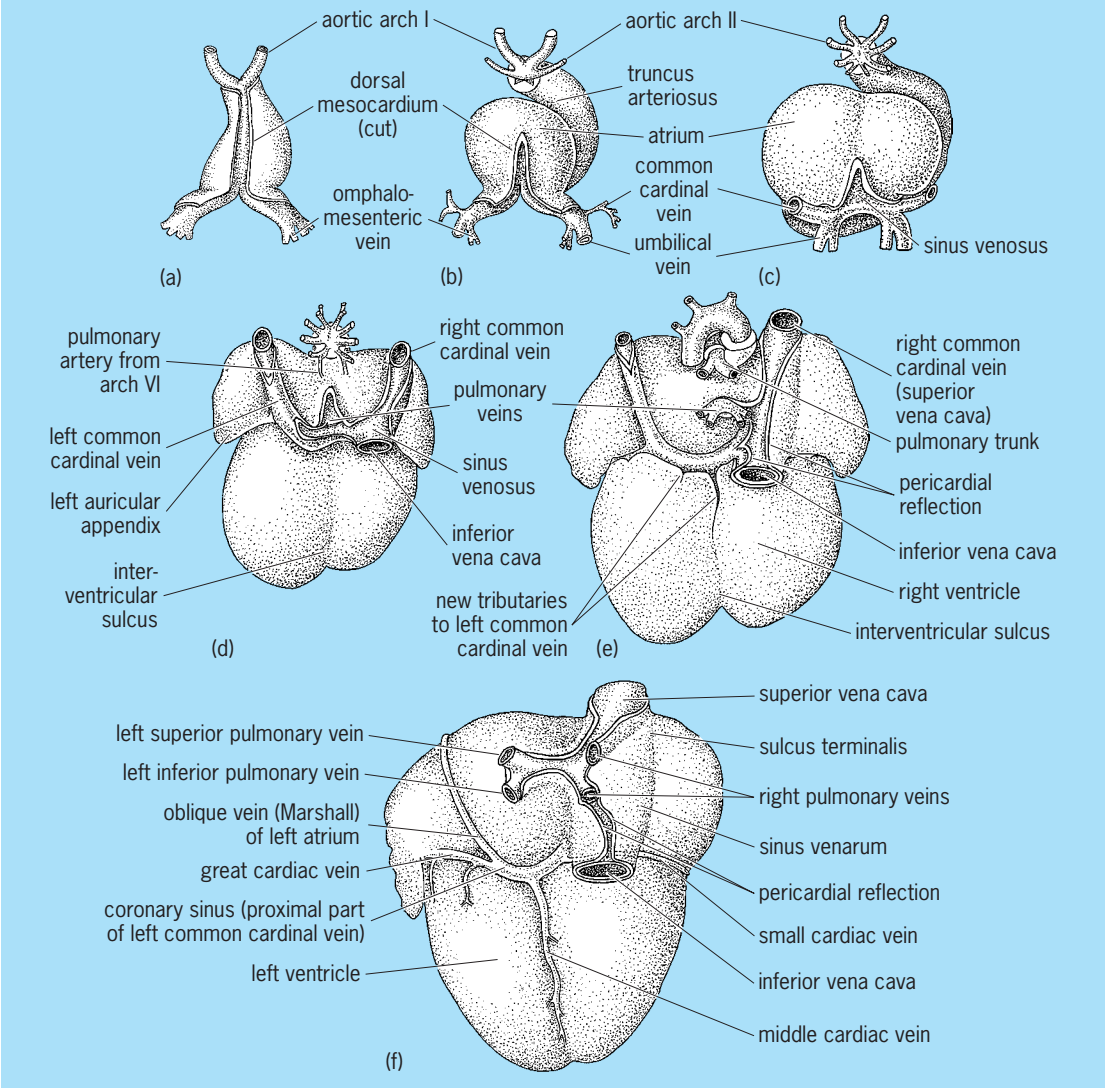

Fig. 7. Six stages in development of the heart, in dorsal aspect to show changing relations of the sinus venosus and great veins entering the heart. (*a*) 2½ weeks (8–10 somites). (*b*) 3 weeks (12–14 somites). (*c*) 3½ weeks (17–19 somites). (*d*) 5 weeks (6–8 mm or 0.24–0.31 in., crown-rump). (*e*) 8 weeks (about 25 mm or 1 in.). (*f*) 11 weeks (about 60 mm or 2 in.). (*After B. M. Patten, Human Embryology, 3d ed., Blakiston–McGraw-Hill, 1968*)

The fusion of the sinus venosus with the right atrium is somewhat more complete in most mammals than in birds. In mammals specialized tissues develop in certain areas of the original sinus region to form the sinoatrial and atrioventricular nodes. The sinoatrial node, located near the entrance of the superior vena cava, serves as the cardiac pacemaker.

Complete division of the heart into right and left sides, that is, a complete separation of respiratory and systemic circulations, is achieved only in birds and mammals. This supplies a higher arterial pressure on the systemic side and distributes oxygenated blood to the tissues more rapidly than in animals with partial separation (reptiles and amphibians) or those with no separation (fishes).

Partitioning of mammalian heart. The division of the heart into right and left sides has been studied in detail in humans and in a number of other mammals. The following account applies primarily to the human heart, but the general plan of partitioning is similar in other mammals. It should be noted that the partitioning of the embryonic heart involves more than its structural division into parts. The process is beset by striking functional exigencies. Cardiac partitioning in the human embryo does not begin until about the fifth week of development, 0.2 in. (6 mm) long, when the heart (Fig. 6*d*) is already maintaining a circulation essential for embryonic life. All of the complex changes in partitioning must be made without interruption of the blood supply to any part of the growing embryo. Starting with a stage when the blood is flowing through it in an undivided stream, the heart, by the time of birth, must become converted into an elaborately valved, four-chambered organ, pumping from its right side a pulmonary stream which is returned to its left side to be pumped out again over the aorta as the systemic bloodstream. Moreover, at the end of gestation, the vascular mechanism must be prepared for air breathing. At the moment of birth the lungs, their blood

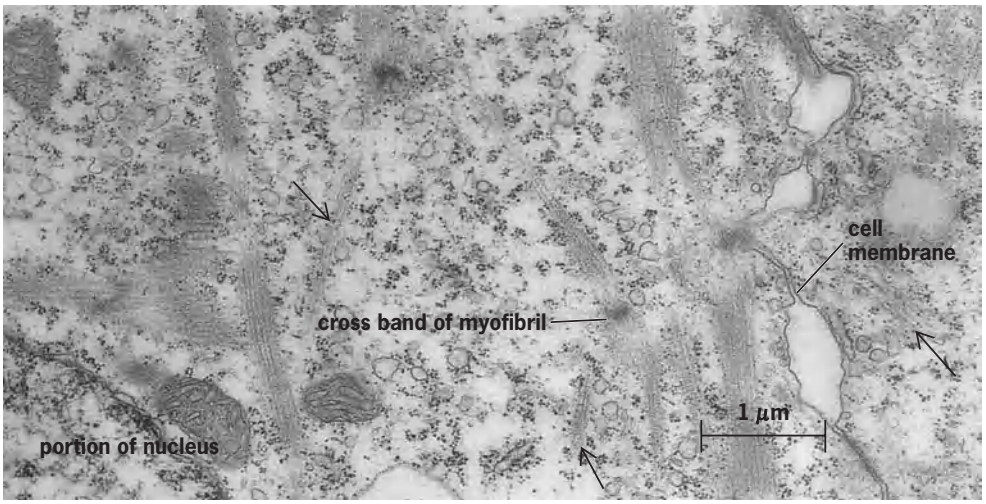

(a)

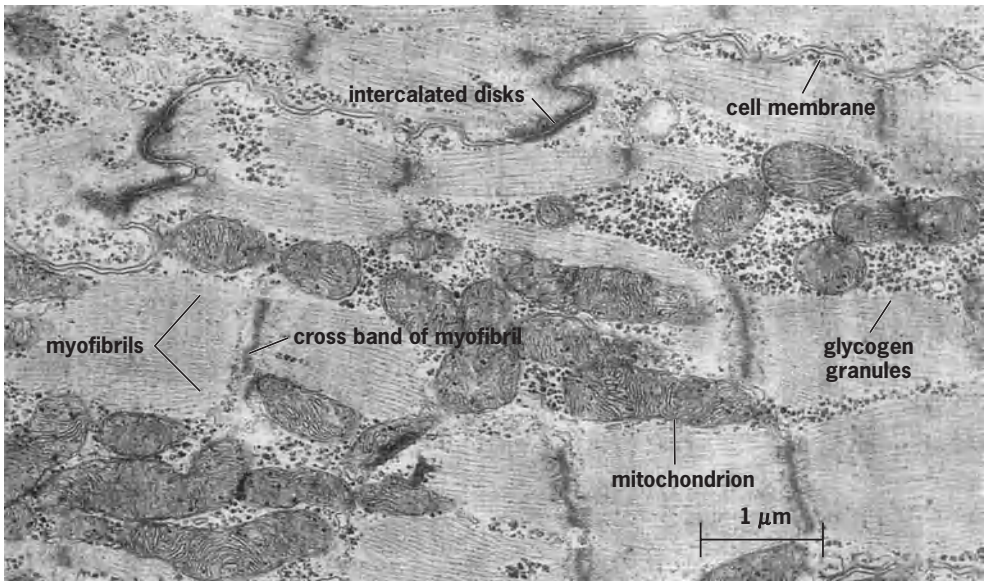

(b)

Fig. 8. Electron micrographs of cardiac muscle at different stages of development. (*a*) Section of sinus venosus from the heart of a tadpole just before the feeding stage. (*b*) Section of ventricle of adult frog. Note that the tadpole sinus has irregularly arranged myofilaments (at arrows) and poor alignment of filaments into myofibrils with only a few cross bands.

vessels, and the right ventricle (which pumps the pulmonary circuit) must be ready to take over from the placenta the entire responsibility of oxygenating the blood.

The systemic part of the circulation must also be prepared. During intrauterine life the left side of the heart receives less blood from the pulmonary veins than the right side receives from the vena cavae. Immediately after birth the left ventricle is called on to do more work than the right ventricle. It must pump sufficient blood through the myriad peripheral vessels of the systemic circulation to care for the metabolism and continued growth of all parts of the body. These are just some of the striking functional situations which the growing heart must encounter before it can attain its adult condition.

Division of atrium and ventricles. Almost from their earliest appearance, the atrium and ventricle show external indication of their impending division into right and left sides. A distinct median furrow appears at the apex of the ventricle (Figs. 7*d* and *e* and 6*c*–*e*). The atrium meanwhile bulges out on either side of the midline (Fig. 7). Its bilobed configuration is emphasized by the manner in which the truncus arteriosus compresses it mid-ventrally (Fig. 6*c*–*e*). These superficial features suggest the more important internal changes.

As the wall of the ventricle increases in thickness, it develops internally a meshwork of interlacing muscular bands, the trabeculae carneae. Opposite the external furrow these muscular bands become consolidated as a partition which projects from the apex of the ventricle toward the atrium. This is the interventricular septum (**Fig. 9**).

Meanwhile, two conspicuous masses of the loosely organized mesenchyme called endocardial

cushion tissue develop in the walls of the narrowed portion of the heart between the atrium and ventricle. One of these endocardial cushions of the atrioventricular canal is formed in its dorsal wall (**Fig. 10***b*). A similar one is formed opposite it on the ventral wall. When these two masses meet, they divide the atrioventricular canal into right and left channels (**Fig. 11**).

Septum primum. Concurrently, a median partition appears in the cephalic wall of the atrium. Because another closely related partition will form adjacent to it later, this is called the first interatrial septum or septum primum. It is composed of a thin layer of young cardiac muscle covered by endothelium. It is crescent-shaped, with its concavity directed toward the ventricle. The apices of the crescent extend to the atrioventricular canal where they merge, respectively, with its dorsal and ventral endocardial cushions (Fig. 10). This leaves the atria separated from each other except for a diminishing opening called the interatrial foramen primum (Fig. 10*b*).

While these changes have been occurring, the sinus venosus has been shifted out of the midline so that it opens into the atrium to the right of the interatrial septum (Fig. 7). The heart is now in a critical stage of development. Its original simple tubular form has been altered so that the four chambers characteristic of the adult heart are clearly suggested. Partitioning of the heart into right and left sides is well under way, but there are still open communications from the right to the left side in both atrium and ventricle. If partitioning were completed now, the left side of the heart would be left almost literally dry, because the sinus venosus, into which systemic, portal, and placental currents all enter, opens on the right of the interatrial septum. Not until much later do the lungs and their vessels develop sufficiently to return any considerable volume of blood to the left atrium. The partitions in the ventricle and in the atrioventricular canal do progress rapidly to completion (Figs. 10, 11, and **Fig. 12**), but an interesting series of events takes place in the interatrial septal complex which assures that an adequate supply of blood will reach the left atrium and thence the left ventricle.

In the sixth week, just when the septum primum is about to fuse with the endocardial cushions of the atrioventricular canal, thus closing the interatrial foramen primum, a new opening is established. The more cephalic part of the septum primum becomes perforated, at first by multiple small holes. These soon coalesce to form the interatrial foramen secundum, thus keeping a route open from the right to the left atrium (Fig. 11).

Septum secundum. During the seventh week the second interatrial partition makes its appearance just to the right of the first. Like the septum primum, the septum secundum is crescent-shaped. The concavity of its crescent is, however, differently oriented. Whereas the open part of the septum primum is directed toward the atrioventricular canal, the open part of the septum secundum is directed toward the lower part of the sinus entrance which later becomes the opening of the inferior vena cava into the right

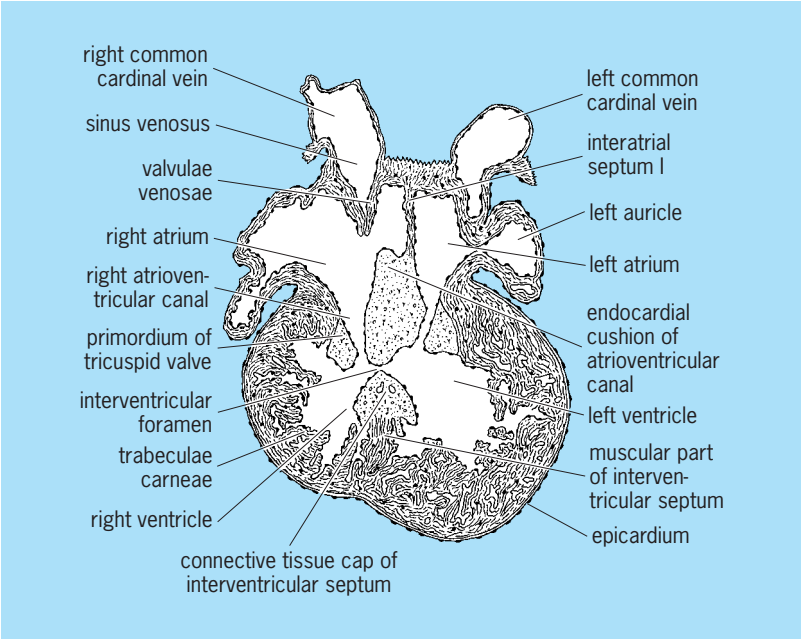

Fig. 9. Section through heart of a 9.4-mm (0.37-in.) pig embryo. Stage of development corresponds to that in a human embryo of the sixth week. (*After B. M. Patten, Foundations of Embryology, 2d ed., McGraw-Hill, 1965*)

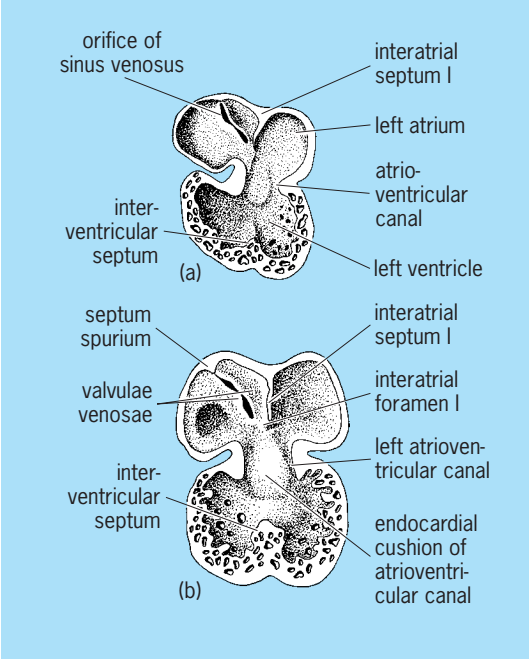

Fig. 10. Interior of heart showing initial steps in its partitioning. (*a*) Cardiac septa in human embryos early in fifth week, showing primary relations of interatrial septum primum. Based on original reconstruction of heart of a 3.7-mm (0.15-in.) pig embryo and on Tandler's reconstructions of corresponding stages of the human heart. (*b*) Cardiac septa in human embryos of sixth week. Note restriction of interatrial foramen primum by growth of interatrial septum primum. Based on original reconstructions of the heart of 6-mm (0.24-in.) pig embryo, on Born's reconstructions of rabbit heart, and Tandler's reconstructions of corresponding stages of the human heart. (*After B. M. Patten, Human Embryology, 3d ed., Blakiston–McGraw-Hill, 1968*)

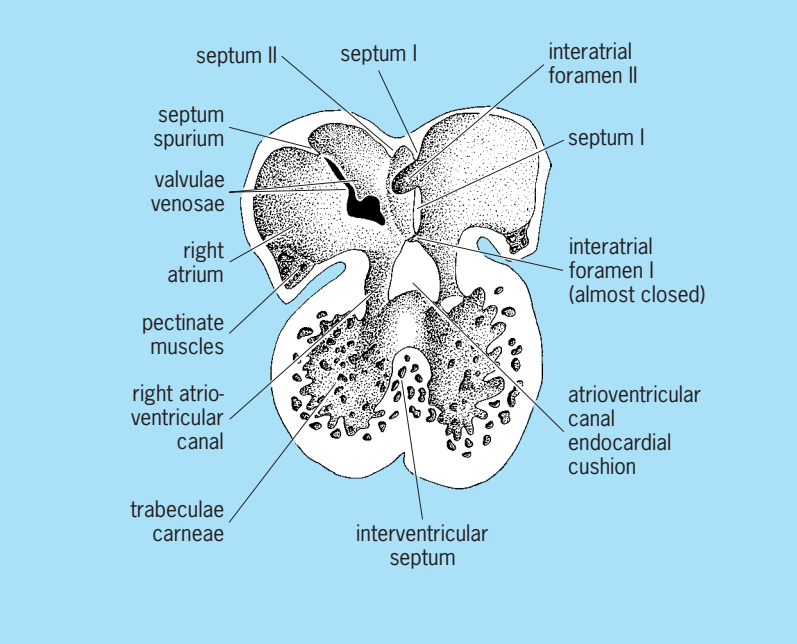

Fig. 11. Interior of heart showing start of interatrial septum secundum and appearance of interatrial foramen secundum in septum primum. Based on original reconstructions of the hearts of pig embryos and on Tandler's reconstructions of the heart of human embryos of the seventh week. (*After B. M. Patten, Human Embryology, 3d ed., Blakiston–McGraw-Hill, 1968*)

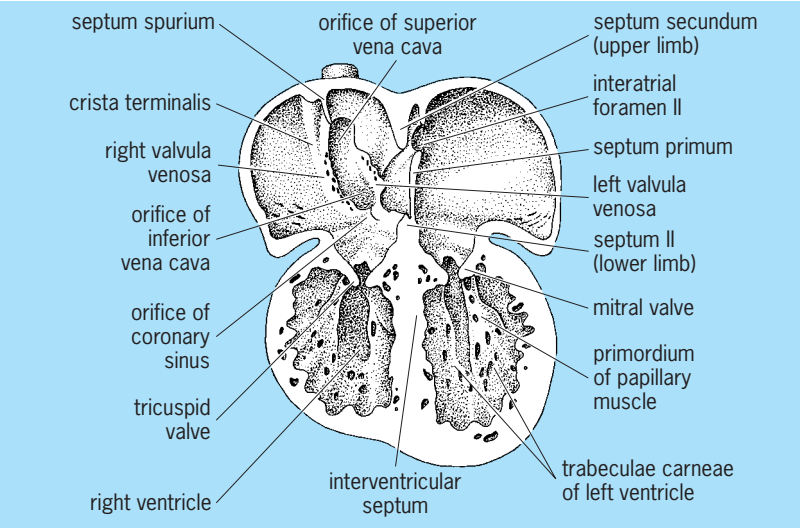

Fig. 12. Heart of third-month human embryo. Resorption has begun in valvulae venosae and septum spurium as indicated by many small perforations in their margins. Left venous valve is coming to lie against and fuse with septum secundum; it usually leaves no recognizable traces in the adult, but occasionally delicate lacelike remains can be seen adhering to septum secundum, and more rarely, extending a short distance onto the valvula foraminis ovalis. (*After B. M. Patten, Foundations of Embryology, 2d ed., McGraw-Hill, 1964*)

atrium (Figs. 11 and 12). This difference in the direction of growth in the two interatrial septa is of vital functional significance because it means that as the septum secundum grows, the opening remaining in it is carried out of line with the interatrial foramen secundum in the septum primum (Fig. 12). The opening in the septum secundum, although it becomes relatively smaller as development progresses, will not be completely closed but will remain as the foramen ovale.

The flaplike persisting portion of the septum primum overlying the foramen ovale constitutes an efficient valvular mechanism between the two atria. When the atria are filling, some of the blood returning by way of the great veins can pass freely through the foramen ovale merely by pushing aside the flap of the septum primum. When the atria start to contract, pressure of the blood within the left atrium forces the septum primum against the septum secundum, effectively closing the foramen ovale against return flow into the right atrium. Without some such mechanism to afford a fair share of blood to the left atrium, the fetal left ventricle would receive a low blood volume, and as a result its muscular wall would not develop adequately to carry its postnatal pumping load. The strength of cardiac muscle and other muscles in the body depends on the work the muscle is called upon to do.

Division of the truncus. While these changes are going on in the main part of the heart, the truncus arteriosus is being divided into two separate channels. This process starts where the truncus joins the ventral roots of the aortic arches. Continuing toward the ventricle, the division is effected by the formation of longitudinal ridges of plastic young connective tissue of the same type as that making up the endocardial cushions of the atrioventricular canal. These ridges, called truncus ridges, bulge progressively further into the lumen of the truncus arteriosus and finally meet to separate it into aortic and pulmonary channels (**Figs. 13** and **14**). The semilunar valves of the aorta and the pulmonary trunk develop as local specializations of these truncus ridges. Toward the ventricles from the site of formation of the semilunar valves, the ridges are continued into the conus of the ventricles (**Fig. 15**). The truncoconal ridges follow a spiral course through the truncus and extend down

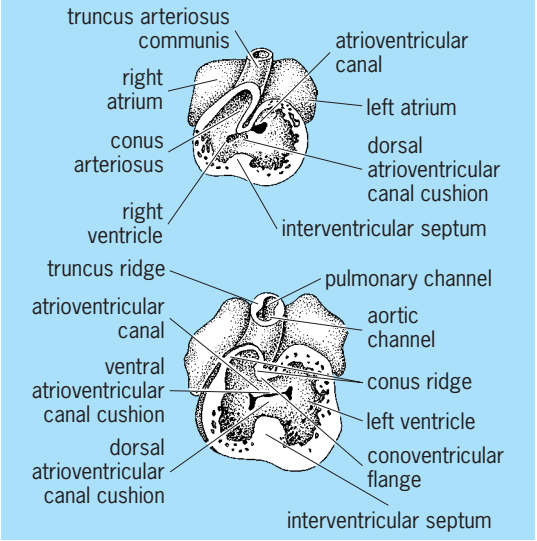

Fig. 13. Dissections of developing hearts in frontal aspect to show relations of importance in establishing aortic and pulmonary outlets. (*After T. C. Kramer, Amer. J. Anat., 71(3):343–370, 1942*)

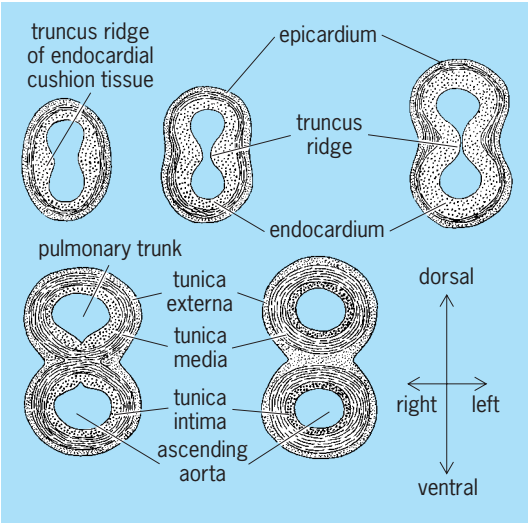

Fig. 14. Partitioning of the truncus arteriosus to form the ascending aorta and the pulmonary trunk. (*After S. E. Gould, Pathology of the Heart, Charles C. Thomas, 1953*)

into the ventricles, where they meet and become continuous with the margins of the interventricular septum. This reduces the relative size of the interventricular foramen but does not close it completely. Its final closure is brought about by a mass of endocardial cushion tissue from three sources. Bordering the interventricular foramen ventrocaudally is the interventricular septum, the crest of which, above its main muscular portion, is made up of endocardial cushion tissue (Fig. 9).

Toward the atrioventricular canal lie the masses of endocardial cushion tissue which were responsible for its partitioning (Fig. 15). In the conus outlet are the ridges that were just considered. From all three of these adjacent areas, endocardial cushion tissue encroaches on the opening. About the end of the seventh week, the interventricular foramen is completely plugged with a loose mass of this young, readily molded connective tissue. Later this mass differentiates into the membranous portion of the interventricular septum and the septal cusps of the atrioventricular valves. When the interventricular septum has thus been completed, the right ventricle leads into the pulmonary trunk and the left leads into the ascending aorta. With this condition established, the heart is completely divided into right and left sides except for the interatrial valve at the foramen ovale which must remain open throughout fetal life until after birth, when the lungs attain their full functional capacity and the entire volume of the pulmonary stream passes through them to be returned to the left atrium.

Ductus arteriosus. This leaves only one of the functional exigencies of heart development still to be accounted for. If, during early fetal life and before the lungs are well developed, the vessels to the lungs were the only exit from the right side of the heart, the right ventricle would not have an outlet adequate to develop its pumping power. It is only late in fetal life that the lungs and their vessels develop to a degree which prepares them for assuming their postnatal activity, and the power of the heart muscle can be built up only gradually by continued functional activity. This situation is met by the ductus arteriosus, leading from the pulmonary trunk to the aorta. Throughout fetal life any blood from the right ventricle that cannot be accepted within the pulmonary circuit is shunted by way of the ductus into the descending aorta. Thus the right ventricle is able to develop its muscular walls by pumping throughout prenatal life its full share of the cardiac load.

The foregoing account of cardiac partitioning applies particularly to the human. There are some notable exceptions in cardiac development in the opossum, a mammal in which the young are born in a primitive stage and continue their development for several weeks in the mother's pouch. This species

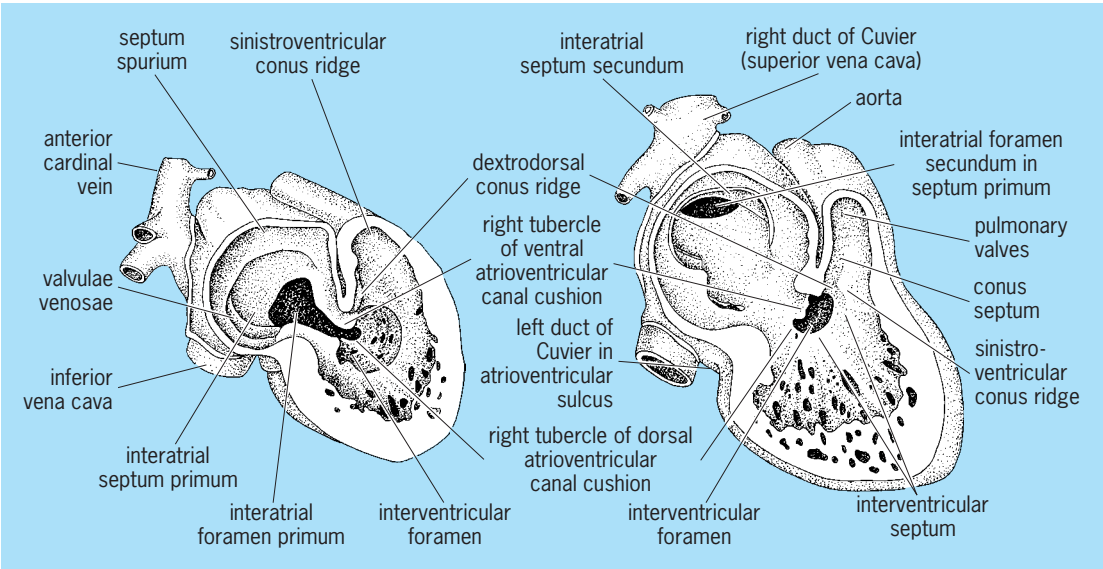

Fig. 15. Lateral dissections showing the relations of the various septa in the developing heart. (*After B. M. Patten, Human Embryology, 3d ed., Blakiston–McGraw-Hill, 1968*)

does not have a secondary interatrial septum or a fossa ovalis. The primary septum has multiple perforations, a few of which persist until birth. The sinus venosus remains in a more primitive stage, clearly defined from the right atrium. The left superior vena cava also persists, opening into the sinus portion of the atrium by way of the coronary sinus.

The rabbit also maintains a more primitive arrangement of the vessels and a more evident sinus region much like the opossum.

Embryogenesis of blood vessels. The endothelial lining of the earliest embryonic vessels arises from mesenchyme, with differentiation occurring first in the wall of the trophoblast and only slightly later in the body stalk and yolk sac. The mesenchymal cells with potency to form endothelium are designated as angioblasts.

According to one view (angioblastic theory), the vessels within the body of the embryo arise by migration of angioblasts from the body stalk and yolk sac. According to the local origin theory, the main intraembryonic vessels, such as the aorta, arise by in situ differentiation from vasoformative mesenchymal cells. Most experimental evidence favors the latter view.

After the main embryonic vessels have arisen, new vessels arise by vascular sprouts from preexisting vessels. Regeneration of endothelium of new vessels following injury in adult life is likewise dependent upon outgrowth from preexisting endothelium.

Angiogenesis. The earliest observable differentiation of endothelium occurs in blood islands of the trophoblast, body stalk, and yolk sac. Each island consists of mesenchymal cell clusters in which the central cells differentiate into blood corpuscles and the peripheral cells elongate and transform into endothelium. Growth and union of the isolated vascular spaces give a plexus of primitive vascular channels. Endothelium of the main vessels of the body of the embryo differentiates also but without relationship to blood islands.

Circulatory system morphogenesis. The earliest vessels are anastomosing, thin-walled tubes (capillaries) lined only by endothelium. In the course of differentiation, some of the capillaries enlarge to form arteries and veins, some remain as capillaries, and some fall into disuse and atrophy. It seems well established that the amount rather than the rate of blood flow through any given portion of a capillary plexus determines whether a vessel merely persists as a capillary enlarges to form an artery or vein. In a similar manner, alterations in vascular pathways occur after the capillary plexus stage of differentiation. For example, arteries atrophy when their blood flow is diverted to newly formed vessels.

Formation of ventral aorta. The aorta arises from bilateral primordia continuous with the bilateral cardiac primordia. The bilateral rudiments quickly fuse to form a common truncus arteriosus continuous with the bulbus cordis. By formation and fusion of two longitudinal ridges along the inner surface of the distal portion of the bulbus, the primitive aortic bulb is split into a pulmonary trunk and an ascending aorta. The latter is relatively short in mammals. It connects with the dorsal aorta by means of aortic arches which vary at different embryonic stages, and the final pattern differs in different species.

Formation of aortic arches. In mammalian embryos there are usually six pairs of arches, but these do not all function simultaneously. Furthermore, the fifth pair is very rudimentary. The arch which follows the questionable fifth one is usually called the pulmonary arch because it becomes a part of the pulmonary arterial system.

The first pair of aortic arches forms in human embryos at the beginning of the fourth week, and the remaining pairs develop in sequence during the fourth week. Transformations (**Fig. 16**) occupy the fifth to seventh weeks.

The first and second pairs become dysfunctional and atrophy as the third and fourth pairs enlarge, and form more direct pathways from the ventral to the dorsal aortas. Following atrophy of portions of the dorsal aortas between levels of the third and fourth arches, the third arches and cephalic portions of the dorsal aortas remain as paired primitive internal carotid arteries. After the external carotid grows out from the third arch, the proximal portion is known as the common carotid.

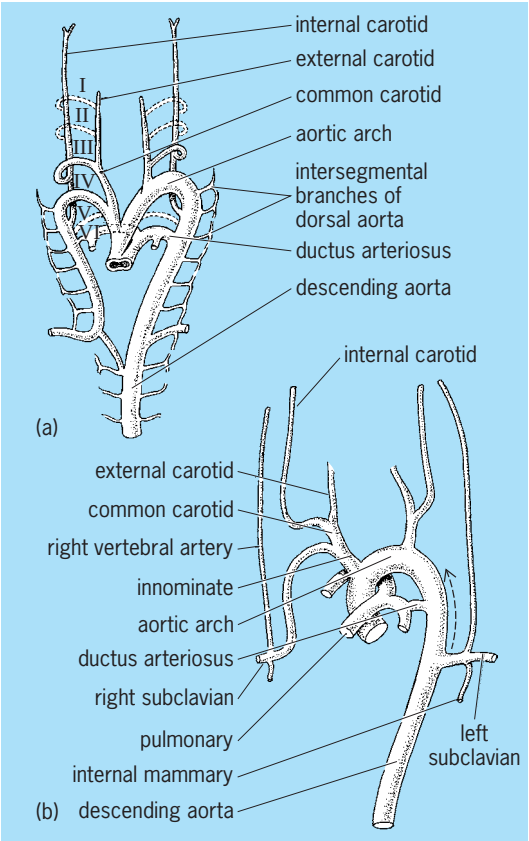

Fig. 16. Transformations of human aortic arches, ventral view. (*a*) Diagram of early stage of transformation when first, second, and fifth pairs of vessels (broken lines) have degenerated. (*b*) Adult derivatives. Diagram, from fetus, shows left subclavian arising from caudal part of arch, but shifts to a higher position at a later stage, indicated by arrow. Ductus arteriosus normally atrophies after birth.

The fate of the fourth arches differs for different classes of vertebrates. In mammals the left fourth arch persists as the arch of the aorta, whereas the right fourth arch forms a portion of the innominate artery and the proximal portion of the subclavian artery (Fig. 16). In birds the condition is the reverse, with the arch of the aorta forming on the left. In reptiles and amphibians both fourth arches retain their connection with the dorsal aorta.

Formation of dorsal aorta. The dorsal aorta arises from bilateral primordia which persist as the paired internal carotids cephalically (Fig. 16) and fuse caudally to form the descending aorta.

The aorta has dorsal, lateral, and ventral branches which are arranged serially. There are about 30 pairs of dorsal branches, which are also known as intersegmental arteries because they occur between successive body segments. Each of the intersegmental arteries divides into dorsal and ventral rami; the dorsal rami give rise to the spinal and vertebral arteries and the ventral rami form the intercostals and lumbars. The lateral aortic branches are not segmentally arranged. They form the renal, suprarenal, inferior phrenic, and internal spermatic or ovarian arteries. The ventral aortic branches develop as paired vitelline arteries to the yolk sac. They fuse when the paired aortas fuse. Later, they are reduced until they occur only at three levels, forming the celiac, superior mesenteric, and inferior mesenteric arteries.

Arteries of the extremities. Vascularization of the extremities illustrates the growth of given pathways within a plexus of vessels. At first, each extremity has a blood supply from several lateral aortic branches. Then, one lateral branch enlarges and the others fall into disuse. In the arm the main stem is the brachial artery which is a continuation from the subclavian. The extension of the brachial into the arm progresses by differential growth of pathways of a capillary plexus. At one stage, the brachial continues by the interosseous to the vessels of the hand. Later, when the median artery arises as a brachial branch and annexes the vessels of the hand, the interosseous becomes less prominent. At a still later stage, the ulnar and radial arteries develop as brachial branches and become the main vessels of the forearm.

The first axial vessel of the leg is known as the sciatic artery. Later, it is superseded by the femoral artery which is a continuation of the external iliac, a branch of the common iliac. After the femoral joins and annexes portions of the sciatic, the proximal portion of the sciatic persists merely as the inferior gluteal, whereas the distal part of the sciatic becomes the popliteal artery in continuity with the femoral.

Primitive venous system. In the early embryo there are three sets of paired veins of particular significance: vitellines carrying blood from the yolk sac to the heart, umbilicals from the placenta, and cardinals from the head and body.

Changes in vitelline veins. Changes in these vessels are correlated with the developing liver in the sense that there is a mutual intergrowth between cords of hepatic cells and endothelial sprouts from the vitelline veins. Thus the vitelline veins are interrupted by liver sinusoids. By enlargement of some of the sinusoids, a direct channel is formed between the left umbilical vein and the proximal portion of the right vitelline (the future hepatic); the channel is known as the ductus venosus.

Caudal to the liver, the right and left vitelline veins become united by three cross anastomoses. By growth of some parts and atrophy of other parts of this system, the S-shaped portal vein arises.

The segment of the right vitelline between the liver and the heart becomes the hepatic vein; the left vitelline of this level disappears.

Changes in umbilical veins. During its growth, the liver encroaches on the umbilical veins until all the umbilical blood enters the liver. Then the entire right umbilical vein atrophies, leaving the left vein which empties into the ductus venosus. The left umbilical vein and the ductus venosus persist as important vessels until birth; then they atrophy and become the ligamentum teres and the ligamentum venosum, respectively.

Plan of the cardinal veins. Paired precardinal veins from the head join paired postcardinals from the body at the level of the heart to form the common cardinals (ducts of Cuvier). During development, parts of the postcardinals are replaced and superseded by subcardinals and supracardinals. The major veins of the body arise through transformation of the cardinal system.

Changes in precardinals, common cardinals. The precardinal of each side is composed of a primary head vein and the precardinal proper extending from the head to the common cardinal. The head vein drains anterior, median, and lateral plexuses over the brain. The rostral portion of the head vein becomes the cavernous sinus; portions of the anterior plexus enlarge to become the superior sagittal sinus; and a dorsal connection between the middle and posterior plexuses becomes the transverse sinus. The main precardinal stem of each side becomes the internal jugular vein as far caudal as the level of a shunt from the left to the right precardinal (**Fig. 17**). The shunt itself becomes the left brachiocephalic (left innominate), and the left precardinal caudal to the shunt remains as the relatively small first intercostal.

The right common cardinal with the right precardinal up to the point of intercardinal anastomosis forms the superior vena cava. The left common cardinal of humans persists as the small oblique vein of the left atrium. As a congenital anomaly, it may remain as a left superior vena cava. In some mammals and in the lower vertebrates, two superior venae cavae (left and right) occur normally.

Postcardinal system transformations. Postcardinals develop primarily as vessels of the mesonephroi and their fate in different species varies with that of the mesonephroi. In *Petromyzon* the postcardinal of each side remains, as in the embryo; in elasmobranchs the plan is modified by a renal portal system; and in humans and some other mammals the postcardinals become altered, as shown in Fig. 17.

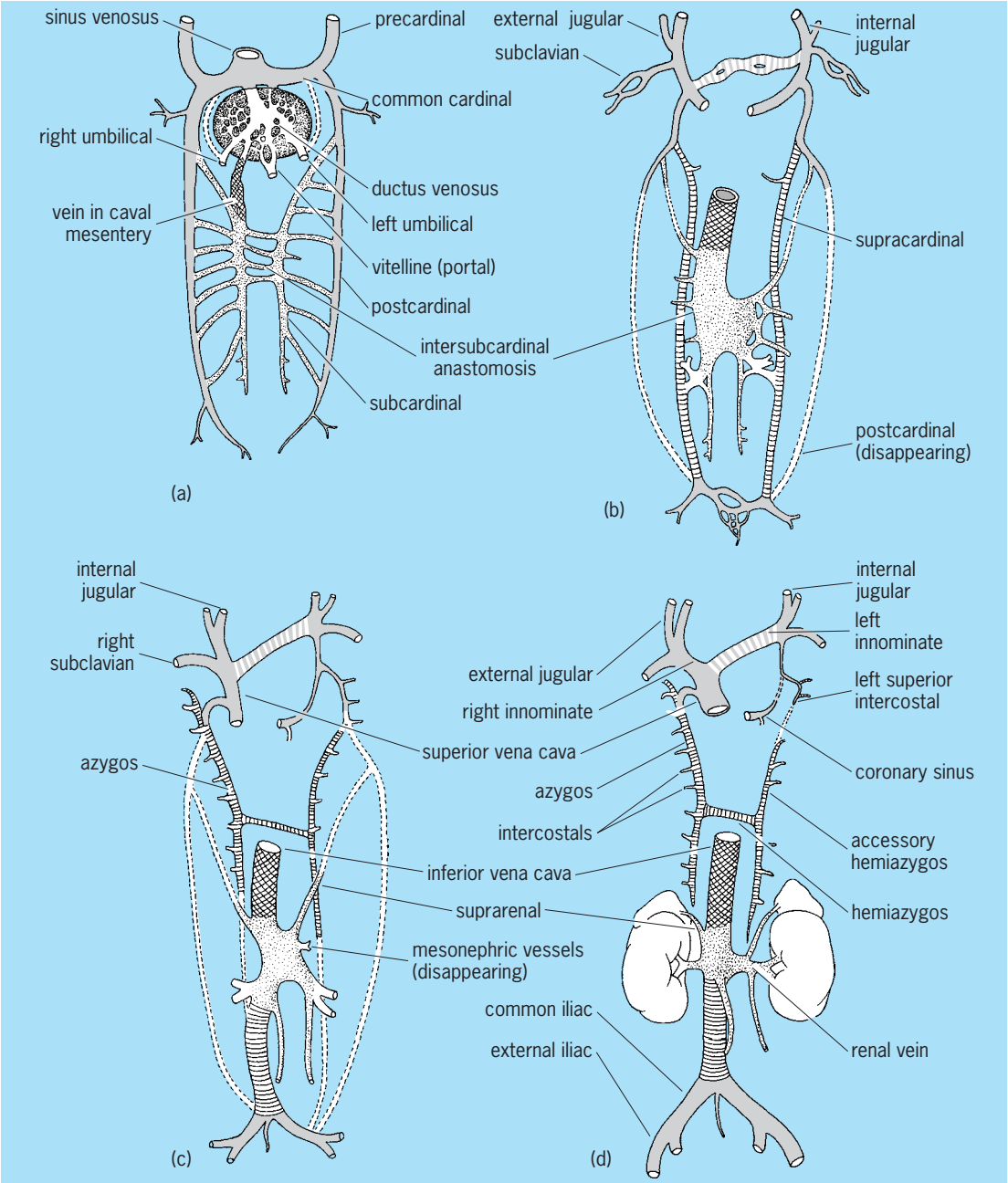

Fig. 17. Transformations of primitive veins of human embryo, ventral view. (*a*) At 5½ weeks. (*b*) At 7 weeks. (*c*) At 8 weeks. (*d*) At term. In *a*, proximal portions of umbilicals (broken lines) have atrophied. Remainder of the right umbilical atrophies later. Vitellines (omphalomesenterics) have been interrupted by liver sinusoids. Right proximal vitelline has enlarged and will become the hepatic. Caudal portions of vitellines are transforming into the portal vein. A connection from hepatic to subcardinals via ventral mesentery will form hepatic segment of inferior vena cava. Vessels associated with liver are omitted from *b*, *c*, and *d*.

This transformation is related to the development of subcardinal veins ventral to the mesonephroi, and supracardinals dorsomedial to the postcardinals. Parts of each of these systems contribute to the formation of the inferior vena cava (Fig. 17).

Veins of the extremities. These develop by channels within capillary plexuses, as already described for the arteries. Each extremity of an early embryo develops a peripheral or border vein. In the upper extremity the border vein persists on the ulnar side as the subclavian, axillary, and basilic veins; in the lower extremity the border vein persists on the fibular side.

Pulmonary veins. These develop from pulmonary plexuses. At one stage of development, all branches combine to open by a single stem into the left atrium. By growth of the atrium, more of the pulmonary vessel is drawn into the atrial wall until there are four pulmonary vein openings into the atrium, two from each lung.

W. M. Copenhaver

Functional Development of Heart

The heart begins to beat before all its parts have been formed. The character of the heartbeat changes with the sequential formation of the regions of the heart.

The heart arises by the fusion of bilateral primordia which converge from positions on opposite sides of the embryo. In the human embryo virtually the entire development of the heart and major blood vessels occurs between the third and eighth weeks of embryonic life. In the chick embryo, one of the principal objects of experimental analysis, the bilateral primordia begin to fuse in the seven-somite stage. The first beats can be recorded in the embryo of nine somites, where the first cross striations also occur.

Contractions of the heart. The first contractions occur along the right side of the developing ventricle, the first part of the heart to form. The beat is at first slow but rhythmical and gradually involves the whole wall of the ventricle. The first contractions may be described as rhythmic but intermittent, that is, rhythmic contractions interrupted by rest periods. Next is added the atrium, which begins to contract at a higher rate than that of the ventricle, stepping up the rate of the entire heart, because the part of the tube with the highest rate of contraction sets the pace for the heart as a whole. At this time the blood is set in motion. The sinus venosus begins to contract somewhat later; again, its inclusion steps up the heartbeat. Contractions begin before any of the nerve fibers (which in their development grow out from the central nervous system), reach the heart and before the specialized system for conducting impulses in the heart is established. The nervous system secondarily assumes the regulation of the rate of the pulsations originating in the myocardium, retarding and accelerating functions being by the vagus and cervical sympathetic fibers, respectively.

Heart-forming areas. The first pulsations of the heart are foreshadowed by the localization of the heart-forming cells to well-demarcated regions of the embryo in the head-process stage, about 10 h earlier. **Figure 18** shows the position of the heart-forming areas as demonstrated by experimental tests of the histogenetic capacity of isolated fragments of the early embryo. Although cardiac muscle arises from the mesoderm, it is impossible to state whether it develops independently in the chick or whether interactions between the mesoderm and endoderm are required. Such an inductive interaction appears to be required for normal development of the heart in the salamander embryo.

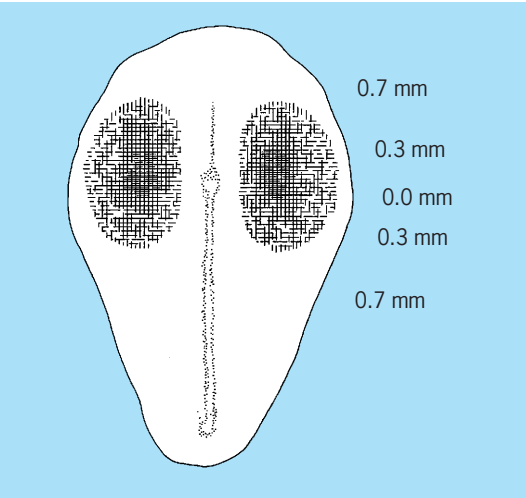

Fig. 18. Heart-forming areas in a head-process stage chick blastoderm. Numerals show distances from level of primitive pit. (*After M. E. Rawles, The heart-forming areas of the early chick blastoderm, Physiol. Zool., 16:22–42, University of Chicago Press, 1943*)

Contractile proteins. The question has been raised as to what extent the early localization of heart-forming cells reflects the early differentiation of specificity in chemical composition or metabolism. It has also been questioned whether the heart-forming areas differ chemically from adjacent regions of the embryo. Immunochemical techniques employing as tools specific antiheart sera reveal that the early embryo contains macromolecules identical with or closely related to those of the adult heart, a conclusion that is supported by the additional finding that when early embryos are cultured in a medium containing antiheart serum the development of the heart is differentially suppressed. Similar techniques show time of origin and pattern of localization of the cardiac contractile proteins, myosin and actin.

Synthesis of contractile proteins. The synthesis of cardiac myosin is initiated during the formation of the mesoderm, the protein being detected first during the movement of prospective mesodermal cells through the primitive streak. As these formative movements are terminated, myosin is distributed widely in the ectoderm and mesoderm, but cannot be detected in the endoderm. At the head-process stage, cardiac myosin is restricted to the heart-forming regions of the embryo where, at the same time, the synthesis of cardiac actin is initiated. It is impossible at the present time to say whether the myosin first detected is adult myosin, or a complete myosin molecule closely related to but not identical with the adult protein, a situation that exists in the sequential formation of fetal and adult hemoglobin molecules. It is postulated that, in the formation of cardiac muscle, the contractile proteins are synthesized and subsequently aggregated to form fibrils. Although this hypothesis of stepwise organization agrees with the limited information available on the development of the fine structure of heart muscle, a note of caution must be expressed, because in skeletal muscle myosin has been detected simultaneously with the formation of the first simple muscle fibrils, and not before. *See* HEMATOPOIESIS.

Action of inhibitors. In further studies of the biochemical differentiation of the heart-forming areas, it has been found that the metabolic pathways operating in the development of the brain and heart differ markedly. When early chick embryos are cultured in vitro on a medium containing traces of the metabolic inhibitor antimycin A, the development of the regions destined to form muscle is inhibited almost completely, whereas the developing brain and spinal cord remain intact. The heart is more sensitive than other mesodermal tissues. Another compound, sodium fluoride, also inhibits development of the heart. The inhibition of the embryo produced by sodium fluoride follows a clear-cut, reproducible

pattern, in which the initial sites of inhibition are the heart-forming regions. At succeeding stages in the establishment of the heart, the locations of the cells destroyed by sodium fluoride reflect the sites of highest ability to form pulsatile heart and the area of greatest capacity for the synthesis of actin and myosin. Thus the primary forces operating in the formation of the heart must be sought at the very outset of development. James D. Ebert

Human Fetal Circulation at Term

The circulation in the mammalian fetus at term must be thoroughly understood as a basis for the consideration of postnatal circulatory changes, because the very mechanisms which ensure intracardiac balance during prenatal life maintain a balanced cardiac load during the changes to the postnatal basis. In this discussion of the fetal circulation at term, the basic conditions presented are essentially applicable to all the higher mammals. Reference to ages and to other specific details are, however, based on the conditions in humans.

By the last trimester of pregnancy all the major blood vessels have been developed in essentially their adult pattern, and all the steps in the partitioning of the embryonic heart are leading progressively closer to the final adult condition, in which the heart is a four-chambered organ completely divided into right and left sides. However, from the nature of its living conditions it is not possible for the fetus in utero to attain fully the adult type of circulation. The plan of the divided circulation of postnatal life is predicated on lung breathing. After birth the right side of the heart receives the blood returning from a circuit of the body and pumps it to the lungs, where it is relieved of carbon dioxide and acquires a fresh supply of oxygen. The left side of the heart then receives the blood that has just been aerated in the lungs and pumps it through ramifying channels to all the tissues of the body. In the last part of intrauterine life the lungs and their blood vessels are fully formed and ready to function, but they cannot actually try out their functional competence until after birth. Nevertheless, in the first minutes of its postnatal life a fetus must successfully change from an existence submerged in the amniotic fluid to air breathing with its hitherto untested lungs. Moreover, this abrupt change must be accomplished without the sudden overloading of any part of the cardiac pump.

It is in the light of these functional exigencies that the fetal circulation at term must be considered. Of primary importance is the fact that at no time during the prenatal life are the atria completely separated from each other. This permits the left atrium to receive a contribution of blood from the inferior vena cava by a transseptal flow which compensates for the relatively small amount of blood entering the left atrium directly by way of the pulmonary circuit. The routes and the relative amounts of this interatrial shunt are different at different ages. Very early in development, before the lungs have grown to any great extent, the pulmonary return is negligible and the flow from the right atrium through the interatrial ostium primum constitutes practically the entire intake of the left atrium (**Fig. 19***b*). After the ostium primum has been closed and while the lungs are but little developed, flow through the interatrial ostium secundum must provide for the major part of the blood entering the left atrium (Fig. 19*d*). During the latter part of fetal life the foramen ovale in the septum secundum becomes the transseptal route for blood (Fig. 19*e* and *f*).

Intracardiac circulatory balance. The precise manner in which this balancing transseptal flow occurs in a fetus at term and where and to what extent the various bloodstreams are mixed has long been a controversial subject. By a synthesis of the most significant of the anatomical evidence with the newer experimental evidence, the course followed by the blood in passing through the heart may be summarized as follows. The inferior caval entrance is so directed with reference to the foramen ovale that a considerable portion of its stream passes directly into the left atrium (**Figs. 20** and **21**). Careful measurements have shown, however, that the interatrial communication through the foramen ovale is considerably smaller than the inferior caval inlet. This implies that the portion of the inferior caval stream which could not pass through this opening into the left atrium must eddy back and mix with the rest of the blood entering the right atrium. Angiocardiographic studies confirm this inference.

Prenatal pulmonary circulation. As the lungs grow and the pulmonary circulation increases in volume relatively less of the left atrial intake comes by way of the foramen ovale and relatively more from the vessels of the growing lungs. However, the circulation of the lungs, although increased as compared with earlier stages, has been shown by angiocardiographic studies to be much less in volume than the caliber of the pulmonary vessels would lead one to expect. Until the lungs are inflated, there are evidently factors, either vasomotor or mechanical or both, which restrict flow through the smaller pulmonary vessels. Therefore, even in the terminal months of pregnancy, a considerable right-left flow must still be maintained through the foramen ovale in order to keep the left atrial intake equal to that of the right.

Balancing ventricular output. The balanced atrial intake thus maintained implies a balanced ventricular intake, and this in turn implies the necessity of a balanced ventricular output. Although not in the heart itself, there is, in the closely associated great vessels, a mechanism which affords an adequate outlet from the right ventricle during the period when the pulmonary circuit is developing. When the pulmonary arteries are formed as downgrowths from the sixth pair of aortic arches, the right sixth arch soon loses its original connection with the dorsal aorta. On the left, however, a portion of the sixth arch persists as the ductus arteriosus connecting the pulmonary artery with the dorsal aorta (Fig. 21). This vessel remains open throughout fetal life and acts as a shunt, carrying over to the aorta whatever blood the pulmonary vessels at any particular phase of their development

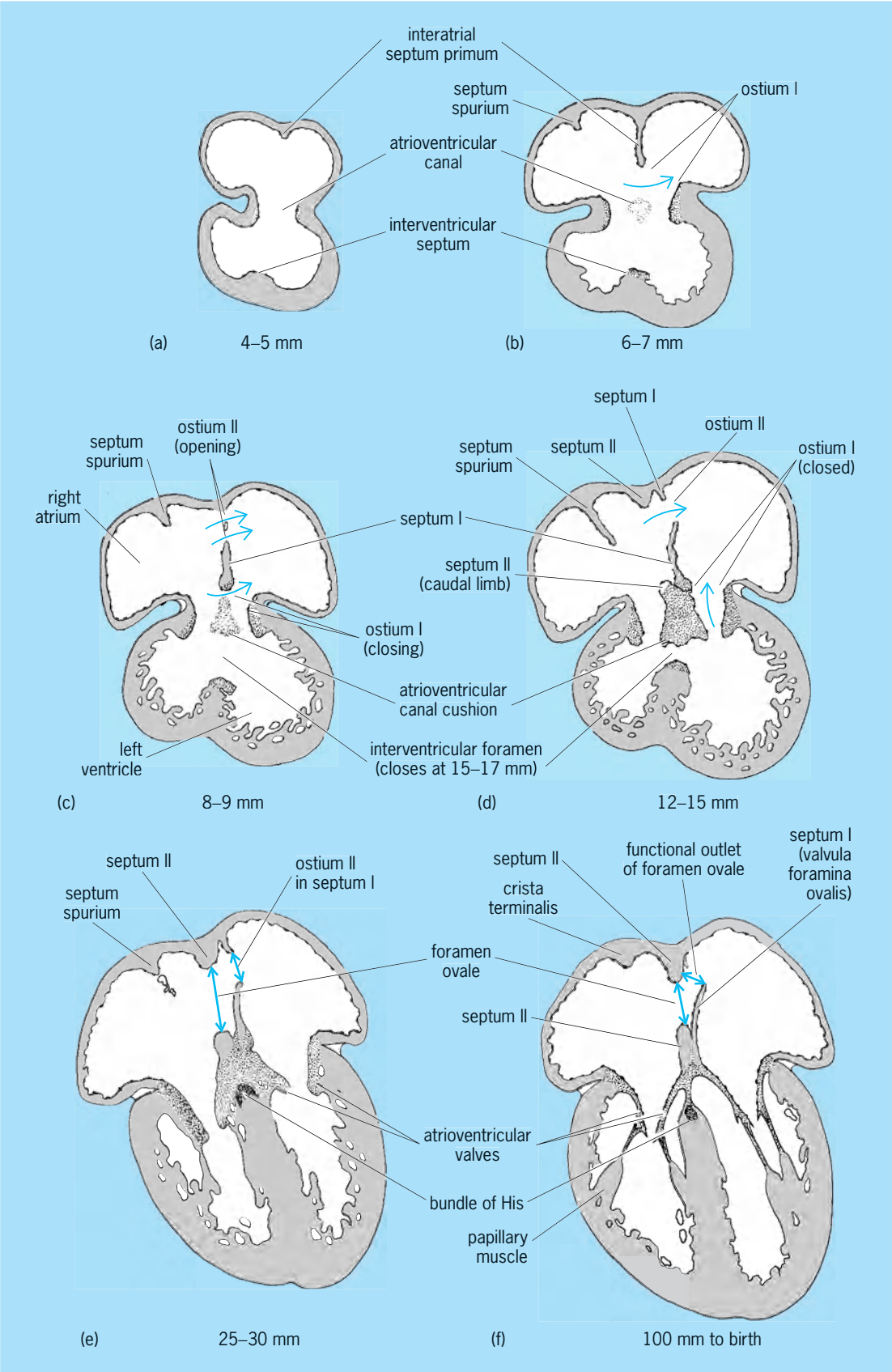

Fig. 19. Sectional plan (*a–f*) of the embryonic heart in frontal plane, showing progress of partitioning in the human embryo. The endocardial cushion tissue is indicated by stippled areas, the muscle by gray areas, and the epicardium by solid black outline. Parts *b* and *c* show the location of endocardial cushions of the atrioventricular canal before they have grown sufficiently to fuse in the plane of the diagram. (*After B. M. Patten, Developmental defects at the foramen ovale, Amer. J. Pathol., 14(2)135–161, 1938*)

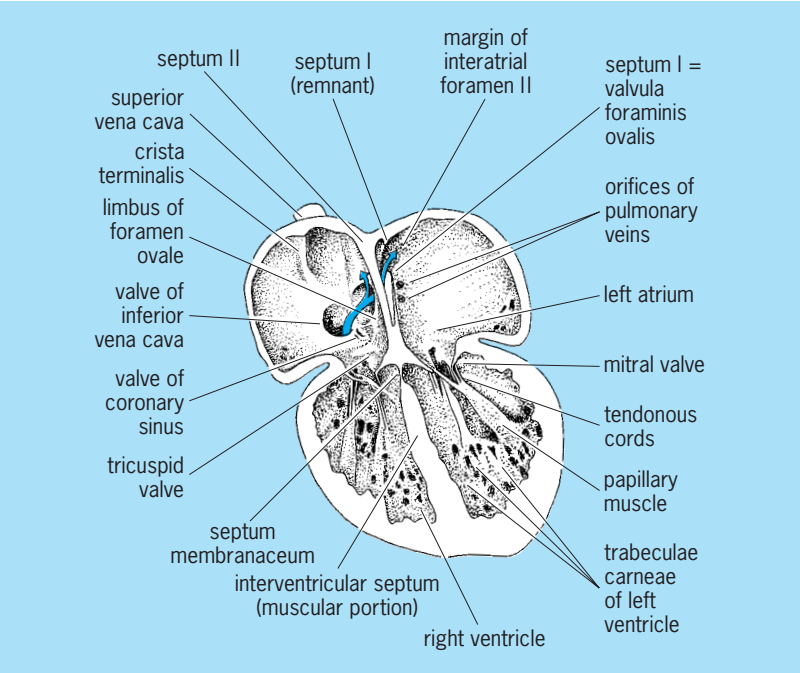

Fig. 20. Interrelations of septum primum and septum secundum during latter part of fetal life. Lower part of septum primum is situated to act as a one-way valve at the foramen ovale in septum secundum. Split arrow indicates the way a considerable part of blood from inferior vena cava passes through foramen ovale to the left atrium while remainder eddies back into the right atrium to mingle with the blood being returned by way of the superior vena cava. (*After B. M. Patten, Foundations of Embryology, 2d ed., McGraw-Hill, 1964*)

are not prepared to receive from the right ventricle. The ductus arteriosus can therefore be called the exercising channel of the right ventricle because it allows the right ventricle to do its full share of work throughout development and thus be prepared to pump its full quota of blood through the lungs at birth.

Thus by means of intake and output shunts there is maintained, throughout prenatal development, an effective right-left balance in the pumping load of the heart. The importance of this for the attainment of normal cardiac structure is forcefully shown by the abnormalities that appear when the balancing mechanisms are in any way disturbed. The importance of these same mechanisms in accomplishing postnatal circulatory changes is discussed in the following section.

Human Postnatal Circulation

The changes in circulation following birth involve some of the most dramatic and fascinating biological processes. One of the most impressive phenomena in embryology is the perfect preparedness for these changes which has been built into the very architecture of the circulatory system during its development. The shunt at the ductus arteriosus, which has been one of the factors in balancing ventricular output throughout intrauterine development, and the valvular mechanism at the foramen ovale, which has been balancing atrial intakes, are perfectly adapted to prevent any abrupt unbalancing of cardiac load as a result of postnatal changes in circulatory routes.

Pulmonary circuit and ductus. In much of the older literature, great emphasis was placed on what happened at the foramen ovale. It is now known that these changes are secondary and that the events of primary significance occur in connection with the pulmonary circuit and the ductus arteriosus. In the section on fetal circulation at term, the importance of the ductus arteriosus as an exercising channel for the right ventricle was emphasized. From the standpoint of postnatal circulatory changes, the reciprocal relation between flow through the ductus and flow through the lungs becomes the center of interest. As the lungs increase in size, relatively more of the blood leaving the right ventricle by way of the pulmonary trunk goes to the lungs, and relatively less goes through the ductus arteriosus to the dorsal aorta.

By the end of gestation the pulmonary vessels must be large enough to handle a blood volume adequate to care for oxygenating the blood. Injection preparations of fetuses clearly show that the vascular channels in the lungs are of generous size for this function well before birth. However, such postmortem material does not show that before birth these large vessels are not carrying the blood volume their size would suggest.

Prenatal pulmonary blood flow. The brilliant work of C. Wegelius and J. Lind, utilizing angiocardiographic methods on living fetuses, has shown that the flow of blood through the pulmonary circuit is actually restricted to a volume much below the potential capacity of the vessels. How much of this is the result of mechanical factors, such as the unexpanded condition of the lungs, and how much it depends upon differential vasoconstriction of the smaller vessels within the lungs remains to be determined. It is clear, however, that pulmonary channels of the requisite size have been formed and are ready to increase their blood flow radically and promptly with the beginning of respiration. Within a short time after birth, under the stimulus of functional activity, the lungs are able to take all the blood from the right side of the heart.

Abandonment and closure of ductus. When the lungs accept all the blood entering the pulmonary trunk, blood flow through the ductus arteriosus ceases. Following its functional abandonment, the ductus arteriosus is gradually occluded by an overgrowth of its intimal tissue. This process in the wall of the ductus is as characteristic and regular a feature of the development of the circulatory system as the formation of the cardiac septa. Its earliest phases begin to be recognizable in the fetus as the time of birth approaches, and postnatally the process continues at an accelerated rate, to terminate in complete anatomical occlusion of the lumen of the ductus some 6–8 weeks after birth.

Postnatal readjustment of the circulation cannot, however, wait on this protracted structural closure. Following birth there appears to be an immediate reduction in the bore of the ductus as a result of its muscular contraction. This is accompanied by a reduced flow of blood through it. This reduction in the

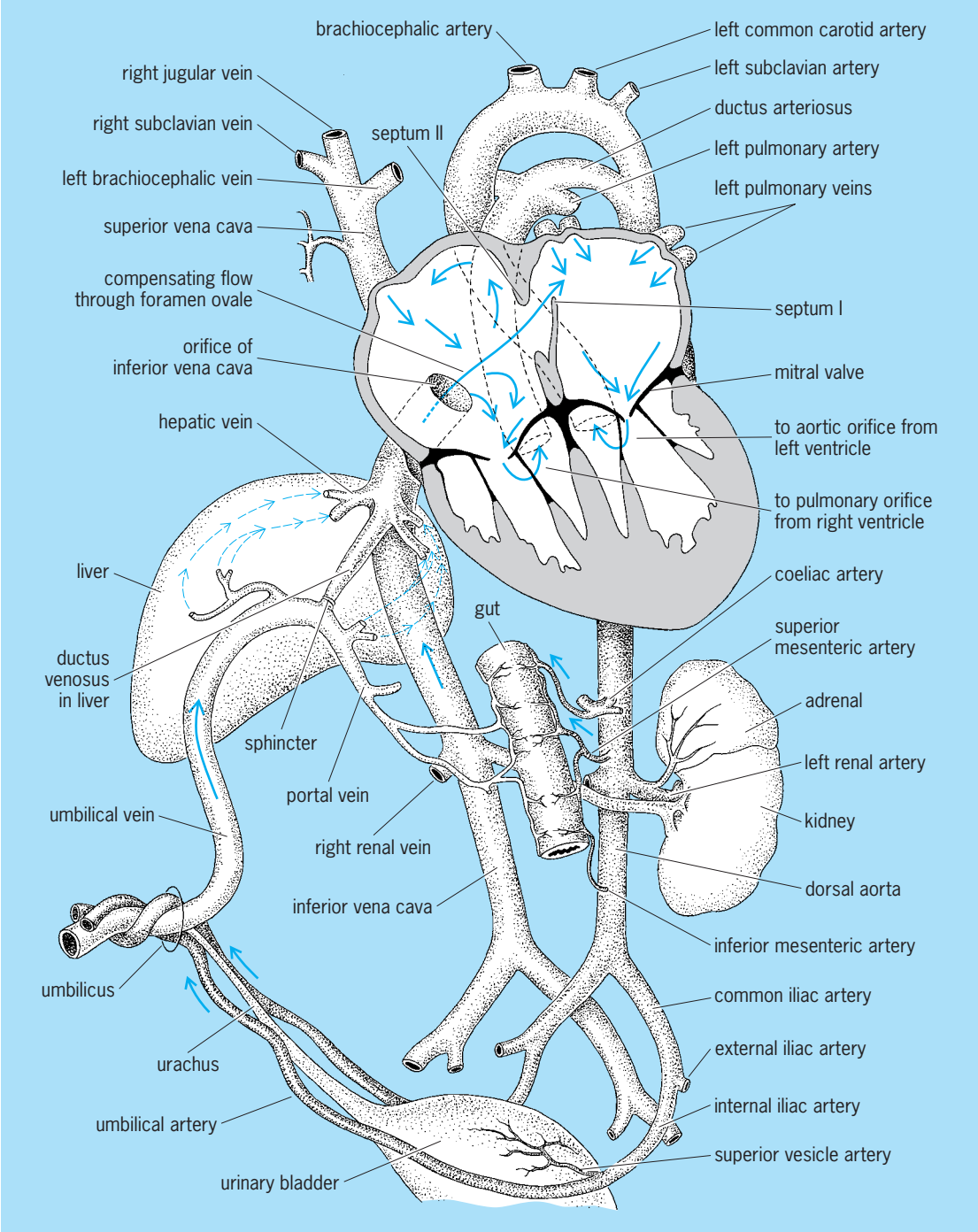

Fig. 21. Diagram of heart and major vessels of fetus just before birth. Small arrows within liver indicate alternative blood routes by way of hepatic sinusoids. (*After B. M. Patten, Human Embryology, 3d ed., McGraw-Hill, 1968*)

shunt from the pulmonary circuit to the aorta, acting together with the lowering of resistance in the vascular bed of the lungs which accompanies their newly assumed respiratory activity, aids in raising the pulmonary circulation promptly to full functional level. At the same time, the functional closure of the ductus arteriosus paves the way for the ultimate anatomical obliteration of its lumen by the active growth of its intimal connective tissue.

Closure of foramen ovale. The results of increased pulmonary circulation with the concomitant increase in the direct intake of the left atrium are manifested secondarily at the foramen ovale. Following birth, as the pulmonary return increases, compensatory blood flow from the right atrium to the left decreases correspondingly, and shortly ceases altogether. In other words, when equalization of atrial intakes has occurred, the compensating one-way valve at the foramen ovale falls into disuse and the foramen may be regarded as functionally closed. The abandonment of the shunt at the foramen ovale is indicated anatomically by a progressive reduction in the looseness

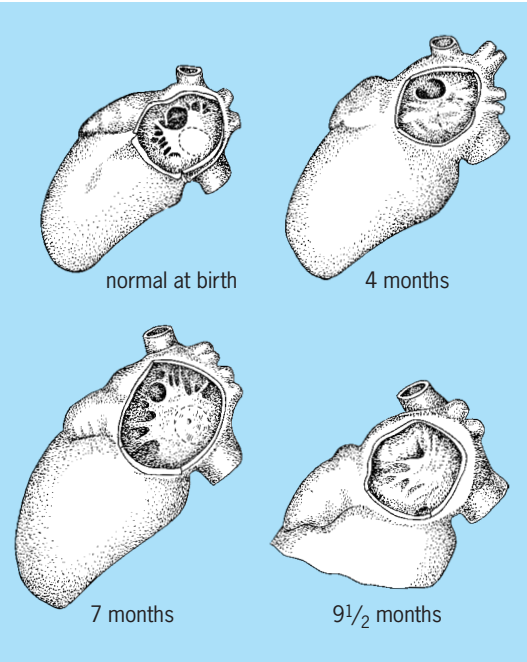

Fig. 22. Hearts with left atrium opened to show gross changes in relations of valvula during period in which foramen ovale closes anatomically. (*After B. M. Patten, The closure of the foramen ovale, Amer. J. Anat., 48:19–44, 1931*)

of the valvula foraminis ovalis and the consequent diminution of the interatrial communication to a progressively narrower slit between the valvula and the septum.

Anatomical obliteration of the foramen ovale slowly follows its functional abandonment (**Fig. 22**). There is a highly variable interval of 3–9 months following birth before the septum primum fuses with the septum secundum to seal the foramen ovale. This delay is, however, of no functional import because as long as the pulmonary circuit is normal and pressure in the left atrium equals or exceeds that in the right, the orifice between them is functionally inoperative. It is not uncommon to find the fusion of these two septa incomplete in the hearts of individuals who have, as far as circulatory disturbances are concerned, lived uneventfully to maturity. Such a condition can be characterized as probe patency of the foramen ovale. When, in such hearts, a probe is inserted under the margin of the fossa ovalis and pushed toward the left atrium, the probe is prying behind the no longer used, but still unfastened, interatrial door.

Physiological aspects of transition. Much yet remains to be learned concerning the more precise physiology of the fetal circulation and concerning the interaction of various factors during the transition from intrauterine to postnatal conditions. Nevertheless, it is quite apparent that the changes in the circulation which occur following birth involve no revolutionary disturbances of the load carried by different parts of the heart. The fact that the pulmonary vessels are already so well developed before birth means that the changes which must occur following birth have been thoroughly prepared for, and the compensatory mechanisms at the foramen ovale and the ductus arteriosus which have been functioning throughout fetal life are entirely competent to effect the final postnatal reroutings of the circulation with a minimum of functional disturbance. The change from living in water to living in air is crowded into a few crucial moments that in phylogeny must have been spread over eons of transitional amphibious existence.

Bradley M. Patten

Comparative Physiology

Comparative physiology describes the structure and operation of the circulation in living animals, and enquires as to how or why the circulatory system may have evolved. Comparing the circulatory system in different animal groups leads to an understanding of general principles and also to various applications of those principles, adapting animals to a wide range of habitats. The circulatory system in all vertebrates has multiple functions, but all functions are involved in regulating the internal environment of the animal (promoting homeostasis).

General physiology of circulation. In all vertebrates the circulatory system consists of a central pump, the heart, which drives a liquid transport medium, the blood, continuously around a closed system of tubes, the vascular system. The arterial portion of this system is divided into larger elastic and smaller resistance vessels (arterioles) which distribute blood to specialized regions or organs (such as muscles, gut, and lungs) where transfer of nutrients, oxygen, or waste products takes place across the walls of a fine network of microscopic capillaries. Blood from the capillaries passes through the venules (small venous vessels) into the main veins and returns to the heart (**Fig. 23**). The arterioles, venules, and capillaries make up the microcirculation, which is arguably the most important role of the vertebrate circulatory system from a functional point of view.

Blood flow in the circulation occurs from regions of high to regions of low fluid energy (or down an energy gradient). The total fluid energy consists of "pressure" energy (created by the heart) and energy contained in the blood due to its motion (kinetic energy). The latter component is small (1 to 2% of the pressure energy in main arteries) and is usually ignored. Hence, the term "pressure" has become synonymous with the total fluid energy. However, this is not always the case as pressures usually recorded in the circulation are transmural pressures; that is, pressure across the wall of the vessel.

The transmural pressure can be very misleading if it is taken as representing the energy for flow. For example, in a recumbent person, transmural pressure in the foot arteries is 1.7 lb in.$^{-2}$ or 12 kilopascals, but in an erect person, the column of fluid between the heart and foot is subjected to the gravitational force, which means the column bears down on the foot arteries with a pressure equivalent to its height (4 ft or 1.2 m), and adds to the pressure created by the pump. Due to this hydrostatic effect, transmural pressure in the foot may be 2.4 lb in.$^{-2}$ or 24 kPa.

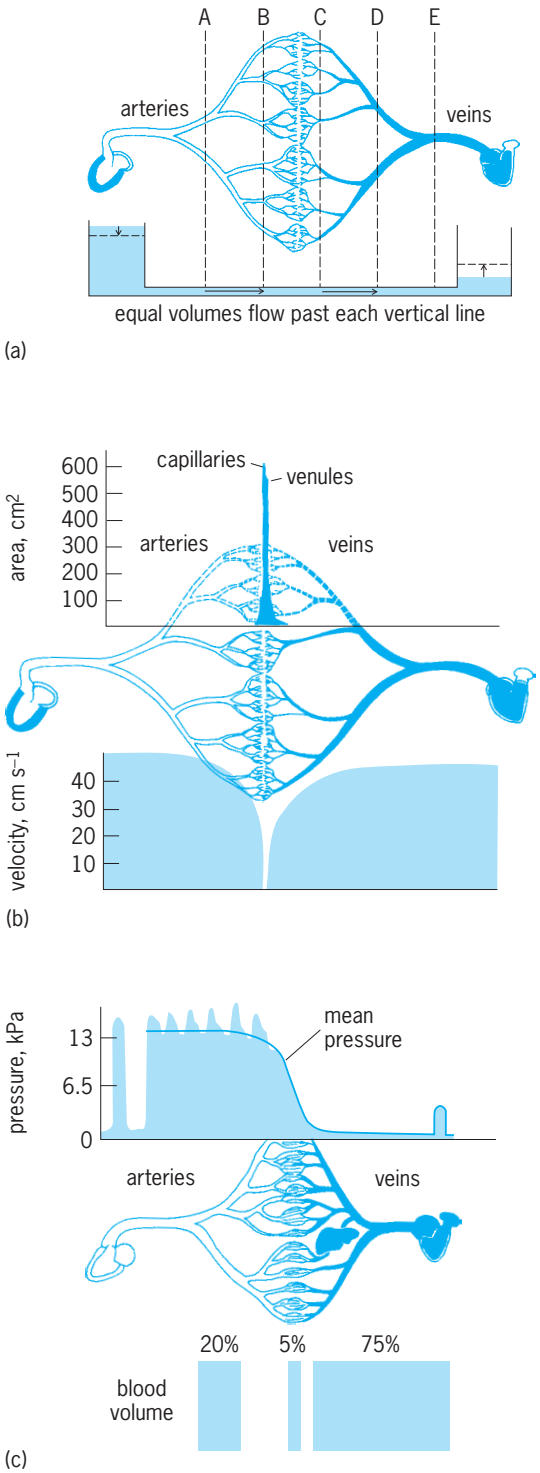

Fig. 23. Systemic circulatory system in a dog (13 kg or 28.7 lb). (*a*) Representation of the aborization of the system. All vessels of the same caliber are arranged vertically (at A, B, C, and so on). (*b*) Increase in cross-sectional area of the various segments. The velocity of blood flow (cm · s^{-1}) is inversely proportional to the cross-sectional area of the tubes through which it flows. (*c*) Pulsatile and mean pressures in the various vascular segments. At the extreme left is a representation of left ventricular pressure, while right ventricular pressure is shown at the extreme right. The boxes show the proportion of blood volume (approximately 85% of total blood volume) in the systemic arteries (including the heart), capillaries, and veins (the pulmonary circulation has about 10–15% of the total blood volume). (*After R. F. Rushmer, Cardiovascular Dynamics, W. B. Saunders, 1961*)

However, this hydrostatic effect is also added to the veins so the energy gradient for flow through the vascular beds remains unchanged. In other words, in a closed circulation, flow cannot be maintained by differences in fluid level (hydrostatic pressure). Flow occurs as a result of difference in total fluid energy between two points in the circulation; if these points are at the same fluid level (central arteries and veins), then the energy difference can be measured in terms of pressure.

Energy is dissipated unevenly around the circulation, the major portion being expended to force blood through the arterial resistance vessels (Fig. 23, energy shown as pressure). Even so, loss of energy is not altogether due to friction between the vessel wall and blood but rather to overcoming the internal friction (or viscosity) of the blood. In the circulation, blood flow is usually streamlined or laminar. If blood, flowing in a vessel, is imagined as made up of a series of thin concentric tubes (or laminae), the layer of fluid next to the wall is stationary while that in the center flows fastest. Therefore, successive layers of blood slide past one another, and friction between these layers of fluid dissipates mechanical energy as heat.

The relation between mean pressures (energy) and flows in the body circulation (equally applicable in general principle to the pulmonary circulation) is shown in Fig. 23. At every division of a blood vessel the cross-sectional area of the branches always exceeds that of the parent vessel (with two branches by a factor of 1.2 to 1.3), so there is a tremendous increase in area as the capillaries are approached (Fig. 23*b*). In a closed circulation what goes in at one end must come out the other, so the total blood flow through each region must be the same, but the rate of flow will be inversely proportional to the cross-sectional area of any given region (Fig. 23*a* and *b*). Since the cross-sectional area of the veins is not much greater than that of the arteries, venous flow rates are high (Fig. 23*b*). On the other hand, mean pressure in the circulation falls from the arteries to the veins. The major pressure drop in the circulation occurs across the arterial resistance vessels which lead to the capillaries (Fig. 23*c*). Consequently, venous pressure is low, despite high blood flow rates, and this explains why blood does not spurt out when one severs a vein. The majority of the total blood volume is stored in the venous vessels (Fig. 23*c*).

Microcirculation. The microcirculation consists of arterial (arterioles) and venous (venules) resistance vessels on either side of the capillaries. The arterioles and venules constitute the pre- and postcapillary resistances respectively. A capillary is a single layer of endothelial cells enclosed by a basement membrane. Blood flow through the capillary is low (at rest, less than 1 mm · s^{-1}) and is controlled by the opening and closing activity of a short muscular region of blood vessel at the capillary entrance (precapillary sphincter). The total number of capillaries, and therefore their surface area for exchange of materials with the cells, is enormous. The systemic (or body) circulation of humans has billion capillaries

with a surface area of 10,800 ft^2 (1000 m^2). In the pulmonary (or lung) circuit there are billion capillaries with a combined surface area of 650 ft^2 (60 m^2). Lipid-soluble substances, including oxygen and carbon dioxide, and some water cross the capillary cell walls, whereas most water and all water-soluble substances are exchanged through pores either within or between the cells.

The structure of capillaries is dictated by their function in the organ supplied and falls into three main types which are basically similar in all vertebrates. Fenestrated capillaries have holes through the cells permitting large and rapid exchange of solvent and solutes; the pores may be closed by a delicate membrane. This type of capillary is found in kidney, intestine, and glands. Discontinuous capillaries have large gaps between the cells of the wall, and are found where macromolecules or even red blood cells are exchanged, as in bone marrow, liver, and spleen. Continuous capillaries have an apparently continuous layer of cells and occur in muscle, lungs, and central nervous system. However, there are small-diameter pores between the cells in all tissues except the brain. The brain capillaries of all vertebrates, except those of lampreys and hagfishes, appear to present an extremely effective barrier to the transport of materials across their walls. In conjunction with the astrocytes, modified nonnervous brain cells which invest the brain capillaries, they form an effective blood-brain barrier. Metabolites must cross the blood-brain barrier by specialized carrier systems or active transport, and this explains the abundance of mitochondria in the brain capillaries. Neurons are more sensitive than most cells to an imbalance in their environment, and homeostasis of the neuronal environment is the function of the blood-brain barrier.

Capillaries are generally about the diameter of the red blood cells (RBCs) which flow through them. Hence, in some amphibians (RBC diameter, 80 micrometers) the capillaries are far larger than, for example, in the mouse deer (RBC diameter, 1.5 micrometers). This relation between size of the RBC and capillary means that blood cells move through capillaries one at a time acting as a moving plug, sweeping unstirred fluid layers away from the capillary wall and facilitating exchange across the wall by diffusion, by reducing the diffusion distance. This is referred to as bolus flow.

Pressure and flow. The pulsatility of pressure and flow allied to the anomalous viscosity of blood greatly complicates any attempt to describe mathematically the relationship between pressure and flow in large arteries. However, in arterial resistance vessels both pressure and flow pulsations are dissipated so that a condition of steady pressure and flow is approached in the microcirculation. In large vessels, blood viscosity is high (some four times that of water) and, due to the presence of RBCs, appears to change with flow velocity (anomalous properties). In small vessels (less than 200 μm diameter) the apparent viscosity falls and, in capillaries, appears constant and similar to that of water (Fähraeus-Lindqvist effect).

Given conditions of steady pressure, flow, and viscosity, the relation between pressure and flow in the microcirculation can be described by the Hagen-Poiseuille law, Eq. (1), where Q = flow per unit time,

$$\dot{Q} = \frac{\pi r^4 \cdot [P_1 - P_2]}{8\eta L} \tag{1}$$

r = radius of the vessel, $P_1 - P_2$ = pressure difference between the upstream and downstream ends of a pipe of length L, and η = blood viscosity. An important relationship emerging from the Hagen-Poiseuille law is that flow varies with the radius of the vessel raised to the fourth power. Thus, if r is halved (by constriction of a blood vessel), then flow through that vessel will be reduced to one-sixteenth [$^1/_2{}^4 = {}^1/_6$] of that existing formerly for the same pressure difference.

Hagen-Poiseuille's law is often rewritten in a simplified form, Eq. (2) [where R = resistance =

$$\dot{Q} = \frac{P_1 - P_2}{R} \tag{2}$$

$8\eta L/\pi r^4$], as a direct analog of Ohm's law describing the relation between electric current, voltage and resistance. This "Ohmic" relation is frequently applied to the whole circulation (not just the microcirculation) to determine total peripheral resistance (TPR) when Q = cardiac output (volume per unit time), P_1 = mean arterial pressure (kPa), and P_2 = mean venous pressure (kPa). TPR is generally expressed in peripheral resistance units (PRU; kPa · ml^{-1} · unit $time^{-1}$). To simplify comparisons between animals of vastly different sizes (or even between individual vascular beds in the same animal), cardiac output is usually expressed on a unit weight basis yielding PRU per unit weight (see **table**).

Since flow per unit time (ml · min^{-1}) through the vascular system is the same at any two points, capillary pressure (P_c) may be calculated from arterial (P_a) and venous (P_v) pressure by application of the Hagen-Poiseuille law, yielding Eq. (3), where R_a and

$$P_c = \frac{P_a \cdot (R_v/R_a) + P_v}{1 + (R_v/R_a)} \tag{3}$$

R_v are the pre- and postcapillary resistances. Since it is the ratio of these resistances which determines capillary pressure, capillary pressure is largely independent of arterial blood pressure over a wide range. For instance, an R_a/R_v ratio of 5:1 gives P_c = 1.875 kPa when P_a = 10 kPa and P_v = 0.25 kPa. Now if P_a rose to 15 kPa, an adjustment of the ratio to 8:1 would leave P_c unchanged.

Measured capillary pressures are in the range of 1.33 to 4 kPa. This raises the question of why the single layer of cells in the capillary is not disrupted by these high internal pressures. The answer is contained in Laplace's law. This law describes how the wall tension (T) which opposes the distending force of the blood pressure (P) is critically dependent on the radius of the vessel, for Eq. (4). Hence, wall

$$P = \frac{T}{r} \tag{4}$$

Some cardiovascular variables in selected vertebrates*

Species name (common name)	Conditions	Body mass and temp.	Heart rate, beats · min^{-1}	Cardiac output, ml · min^{-1} · 100 g^{-1}	Mean arterial pressure, kPa	Mean pulmonary pressure, kPa	Peripheral resistance of body (PRU_{100}), kPa · ml^{-1} · min^{-1} · 100 g^{-1}	Peripheral resistance of gas exchanger (PRU_{100}), kPa · ml^{-1} · min^{-1} · 100 g^{-1}
Salmo gairdneri (rainbow trout)	Unrestrained, rest	1.25 kg, 10°C	38	1.76	4†	5.06‡	2.26	0.6
	Unrestrained, exercise		51	5.26	4.8†	8‡	0.9	0.6
Xenopus laevis (clawed toad)	Restrained, rest, breathing	0.1 kg, 20°C	45	11	3.74	2.8	0.57	0.43§
	Restrained, rest, not breathing		40	6.4	4	3.7	0.56	3.2§
Pseudemys scripta (red-eared turtle)	Unrestrained, rest, breathing	1.25 kg, 21°C	23	5.7	3.74	2.4	1.73	0.65
	Unrestrained, rest, not breathing		11	2.65	2.8	2.0	1.94	1.73
Anas platyrhynchos (white pekin duck)	Restrained, anesthetized	2.5 kg, 41°C	219	22	19	2.0	0.87	0.1
	Diving, unanesthetized		30	2	18	1.67	8.33	0.17
Homo sapiens (human)	Unrestrained, rest	70 kg, 37°C	72	8	12	1.75	1.5	0.2
	Maximum exercise		193	33	14.5	2.4	0.44	0.07

*In fish the gills are in series with the body circulation, so the resistances were calculated by using the pressure difference between the ventral and dorsal aorta and the dorsal aorta and veins, respectively. In fish, amphibians, and reptiles the heart is undivided and cardiac output is the total amount of blood pumped per unit time, whereas in birds and mammals cardiac output is the output of only one ventricle or half the output of the whole heart. In all cases, central venous pressures were assumed to be insignificant, except during diving in ducks when both pulmonary and central venous pressure rise to 1.3 kPa.
†Dorsal aorta.
‡Ventral aorta.
§Includes skin.

tension in a capillary in a human is 1/10,000 of that in the aorta, even though the pressure within is one-fifth of that in the aorta. Not surprisingly the aortic wall is 10,000 times thicker than the capillary wall.

Fluid exchange across capillaries. Water is exchanged both across the wall of the capillary and through pores. It is driven out by the blood pressure in the vessel, which is much higher than the pressure in the fluid between cells. In fact, the extracellular fluid pressure is frequently subatmospheric (negative). The venules (small veins), as well as capillaries, are apparently involved in this exchange because the venules are also highly permeable to water. The pressure which filters water out of the capillary (filtration pressure) is opposed by the colloid osmotic pressure (COP) of the blood; COP tends to pull water back into the blood vessel. The COP difference between the blood and the fluid between the cells (interstitial fluid) is almost entirely due to a higher concentration of relatively impermeable proteins in blood plasma; it is equivalent to a pressure of 1.3 to 4 kPa depending on the animal.

At the arterial end of the capillary there is usually a net loss of water because the blood pressure is higher than COP. As water leaves, the blood is concentrated, and its COP increases. This increase in COP, coupled with the fall in blood pressure along the capillary, causes fluid absorption at the venous end. In a human perhaps 21 quarts (20 liters) of fluid is filtered from the body capillaries each day, excluding kidney filtration. Of this filtrate, 17–19 quarts (16–18 liters) are reabsorbed, and the remainder returns to the blood by the lymphatic system. When filtration and absorption are more or less in balance, a Starling equilibrium is said to exist (named after E. H. Starling). However, the degree of equilibrium in individual tissues may vary; some may be exclusively involved in filtration, and others in absorption due to the independent regulation of capillary blood pressure in different tissues.

Obviously, capillary fluid exchange must be regulated if the volume of the blood is to be controlled. The vital importance of this control is seen when animals, such as humans, assume an erect posture. A column of fluid stretching from the heart to the feet applies an additional pressure of 12 kPa to the capillary pressure. This results in a pressure across the capillary wall (transmural pressure) of perhaps 16 kPa. This pressure is added to both arterial and venous sides of the circulation so the pressure gradient for flow across the vascular bed remains unchanged (except that flow resistance may drop because the vessels expand a bit due to the high transmural pressures). However, capillary pressure now greatly exceeds COP, and fluid should be filtered throughout the microcirculation. In fact, this is prevented

by a marked reduction in the permeability of the microcirculation due, it is thought, to precapillary sphincters closing off capillaries and thereby reducing the surface area for capillary exchange.

In the brain of erect animals, the hydrostatic effect due to standing is reversed. The blood column must be lifted against gravity from the heart. Thus, in the head of an erect human, there is a negative hydrostatic pressure of 5 kPa, whereas in the giraffe it is 20 kPa. Consequently, a positive capillary pressure, even at the arterial end of the microvasculature, can only be achieved by a high arterial blood pressure. This is the explanation for the extremely high arterial blood pressure in the giraffe. However, at the venous end of the microcirculation, pressures could still be negative, causing veins to collapse. Even though the brain is enclosed within a rigid box, which would tend to prevent blood vessel collapse, many veins are tethered to surrounding skeletal structures to ensure that they remain open.

Control. Microvascular activity is coupled to local tissue function through intrinsic mechanisms which are independent of control by the nervous or hormonal systems. The arterioles, metarterioles, and precapillary sphincters are muscular vessels in which muscle fibers contract spontaneously (myogenic activity), keeping these vessels in a state of partial constriction (basal vascular tone). Increases in blood pressure stretch the vessels, causing an increase in myogenic activity in the vascular smooth muscle and a return toward their original diameter. Any tendency for vessels to close completely is prevented by the accumulation of vasodilator metabolites in the tissues as flow is reduced. Decreases in oxygen and increases in carbon dioxide, both in the blood and tissues, cause muscle relaxation and an increase in vessel diameter (vasodilation). However, in the lung circuit, low oxygen causes blood vessel diameter to decrease (vasoconstriction). This reversal of blood vessel response is related to the lung's role as a supplier, rather than user, of oxygen and has important consequences for distributing blood to areas of the lung where oxygen is available.

The arterioles and metarterioles are also subject to remote control by the nervous system. They are innervated by vasodilator and vasoconstrictor nerve fibers which modulate and sometimes dominate local control systems. In contrast, the precapillary sphincters lack innervation and can only be affected by local or blood-borne excitatory or inhibitory influences.

The endothelial (inner) layer of cells of the blood vessels makes an important contribution to regulation of resistance to flow. Physical stimuli such as stretch, flow, and pressure as well as hormonal influences cause the release of contracting and relaxing factors from the endothelial cells, which affect adjacent smooth muscle. These endothelium-dependent regulatory mechanisms integrate blood vessel responses, and may explain regional differences in responses of vascular beds to the same hormone. The most well-known relaxing factor is nitric oxide (NO), produced from L—arginine, which very rapidly loses its potency in the circulation. Contracting factors include peptides, such as endothelin and prostaglandin which remain active for a considerably longer period of time.

The efficacy of the local control system as an unaided regulator of blood supply to the tissues is remarkable. For example, blood flow to the brain can be maintained constant despite induced arterial blood pressure changes from 8 to 18 kPa. Such control, in the absence of nerves or blood-borne factors, is referred to as autoregulation.

Heart. The heart is a pump imparting propulsive energy to the blood. Primitively, the heart is a tubular structure equipped with valves to prevent backflow. In more advanced animals it is differentiated into receiving and storage chambers (sinus venosus and atrium) and pumping chambers (ventricle and conus arteriosus). During evolution, these chambers have become folded upon one another; some were incorporated into others, and some subdivided. Sharks and rays have all four chambers, whereas in birds and mammals, the sinus venosus is incorporated into the right atrium and the conus arteriosus into the outflow tract of the ventricle. In birds and mammals the two remaining chambers (atrium and ventricle) are subdivided to form two parallel circulations, one supplying the body and the other the lungs.

In all vertebrates the heart is enclosed in a double-layered membranous pericardium; a liquid-filled space separates the layers. The inner layer of the pericardium is applied to the ventricular surface. In elasmobranchs, the outer layer of the pericardium is extremely thick and is attached to surrounding skeletal structures. As a result, ventricular contraction creates a subatmospheric pressure within the pericardium which tends to stretch the thin-walled atrium so that it fills with blood by aspiration.

Cardiac cycle. In all vertebrates the period of one heartbeat, or cardiac cycle, can be divided into four phases (**Fig. 24***a* and *b*).

1. *Filling phase:* the inflow valves are open and the outflow valves shut. Ventricular pressure at the start of this phase is low and falling; when it falls below that in the atrium, the atrioventricular valves open and blood flows rapidly into the ventricle. Flow into the ventricle is driven by the energy contained in venous blood. As ventricular pressure rises, flow slows, and atrial contraction "tops up" the ventricle. The end of the ventricular relaxation phase (diastole) is marked by the start of ventricular contraction (systole) which increases ventricular pressure and shuts the atrioventricular valves.

2. *Isovolumetric contraction (contraction without a change in volume):* both inlet and outlet valves are shut. The ventricular muscle contracts, developing tension, and the pressure of the contained blood increases. In hearts with large and rapid pressure generation the free edges of the atrioventricular valves have guy ropes (chordae tendineae), attached to papillary muscles of the ventricular wall; these prevent the valves being turned inside out.

3. *Ejection phase:* the rising ventricular pressure exceeds that in the arteries and the outlet valves

open; the inlet valves remain closed. Blood is ejected rapidly, but pressure continues to rise in both ventricle and outflow tract until peak systolic pressure is reached. About two-thirds through this phase, ventricular contraction stops and pressure falls; when it falls below arterial pressure, the outflow valves shut, causing a disturbance in the arterial pressure pulse (incisura or dichrotic notch). The amount of blood pumped in this phase, by a single ventricle, is the stroke volume (ml), while cardiac output (ml · min^{-1}) is the product of stroke volume and heart rate (beats · min^{-1}). Backflow of blood is not necessary to shut the outflow valves. When the valve opens during blood ejection, vortices are created between the valve cusp (free edges of the valve) and the aortic wall, and the pressure distribution associated with these vortices deflects the cusps toward apposition (closure) as outflow from the ventricle decelerates. Leonardo da Vinci is credited with first describing the role of vortices in valve closure more than 100 years before W. Harvey's discovery of the circulation. Some lower vertebrates have an extra contractile chamber (conus arteriosus) on the ventricular outflow tract; the conus shuts the outflow valves when it contracts.

4. *Isovolumetric relaxation:* inflow and outflow valves shut and pressure falls rapidly as the ventricular muscle relaxes. Subatmospheric pressure can occur in this phase due to "elastic recoil" of the walls of the ventricle.

The plot of ventricular pressure against volume describes a closed loop circling counterclockwise with respect to time (Fig. 24*c*). The area enclosed by the loop is a measure of the work done by the ventricle in ejecting blood. Since pressures in the lung circuit of birds and mammals (see table) are lower than those in the body circulation, the work required to circulate the blood to the lungs is much less.

The ventricles are not divided in amphibians and noncrocodilian reptiles. Since both the lung and body circulations are connected to the same pressure source, flow in each circuit is inversely proportional to the resistance of that circuit. In amphibians such as frogs and toads, during periods of breath holding (apnea) blood can be circulated away from the lung circuit and sent to the body by increasing pulmonary flow resistance (**Fig. 25**). Hence, unlike the situation in avian and mammalian hearts, flows in lung and body circuits can be independent of one another. However, in incompletely divided hearts peak pressures during cardiac contraction have to be the same in both lung and body circuits because they are connected to the same pressure source. The peak pressure must be low; otherwise plasma will filter out of the lung capillaries into the lung and the animal will drown. Only when the heart is completely divided can pressures in lung and body circuits be independent of one another (high pressure in the body, low pressure in the lung), but the price to be paid is that flows in each circuit must now be exquisitely balanced.

Varanid reptiles (for example, monitor lizards), however, show clear pressure separation between

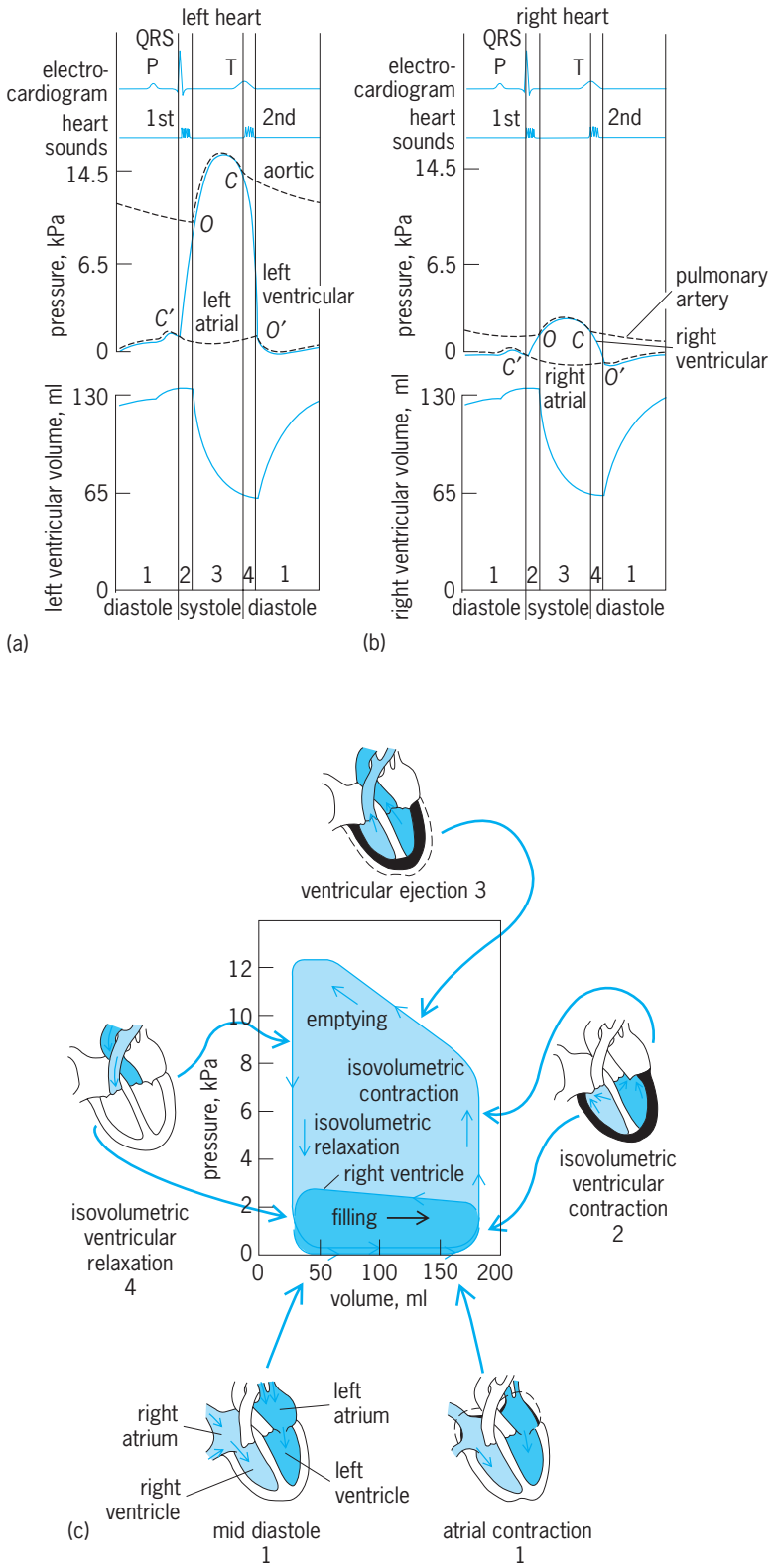

Fig. 24. Mammalian heart. Events in the (*a*) left heart and aorta and (*b*) right heart and pulmonary arteries during a cardiac cycle. In *a*: At *C′* the atrioventricular valve closes; at *O′* it opens. At *O* the aortic valve opens; at *C* it closes. In *b*: At *C′* the atrioventricular valve closes; at *O′* it opens. At *O* the pulmonary valve opens; at *C* it closes (*after A. J. Vander et al., Human Physiology, McGraw-Hill, 1975*). (*c*) Pressure-volume loops for the right and left ventricles. The area enclosed by the loop is a measure of the work done by the heart in ejecting blood. Diagrammatic representations of the heart during one cardiac cycle surround the loops and are linked by arrows with their appropriate position (in time) on the loop. The contracting portions of the heart are shown in black. The numbers (1 to 4) under each heart diagram refer to the phases of the cardiac cycle in *a* and *b* (*after R. Eckert and D. J. Randall, Animal Physiology, W. H. Freeman, 1978*)

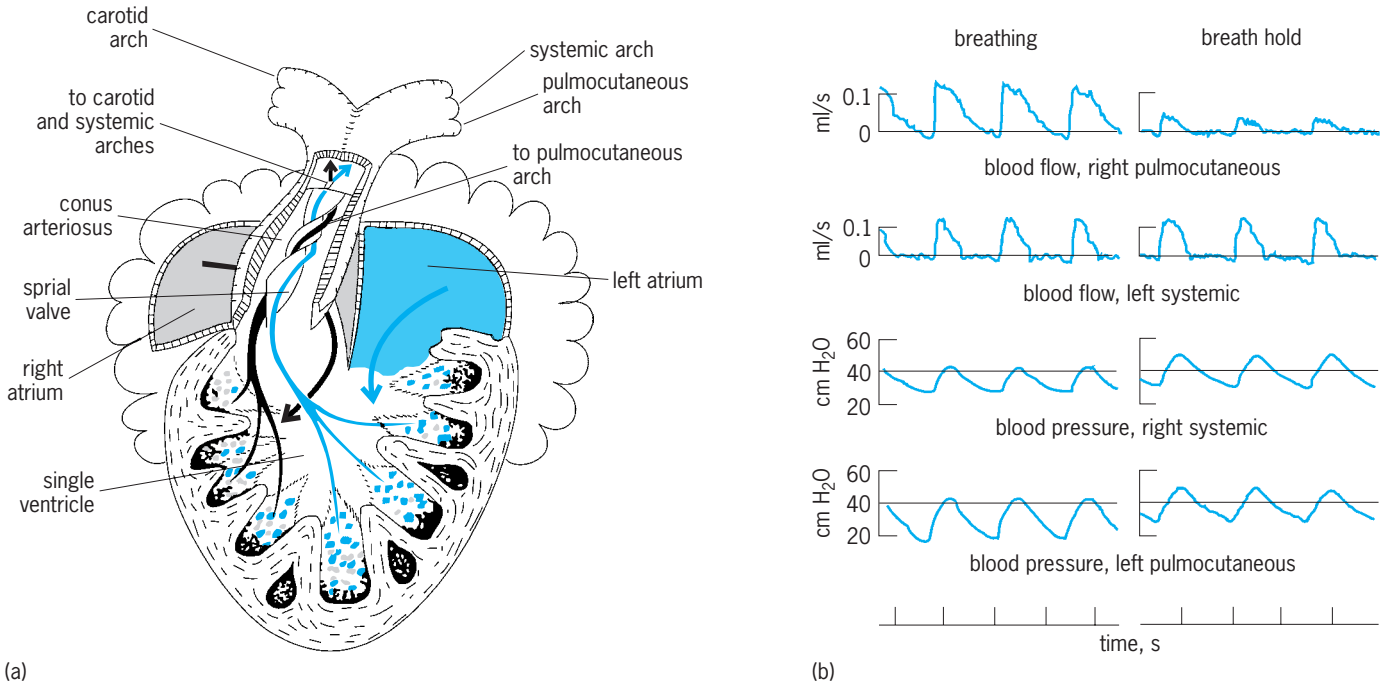

Fig. 25. Frog heart. (*a*) Ventral walls of the atria, ventricles, and conus arteriosus are removed to show flow streamlines that account for selective distribution of oxygenated blood to the head and body via the carotid and systemic arteries, and deoxygenated blood to the lungs via the pulmocutaneous arteries. (*b*) Blood pressures and flows in the blood vessels supplying the lungs (pulmocutaneous) and body (systemic) during a period of breathing and during a breath hold in the anuran amphibian *Xenopus laevis*. (*After P. S. Davies, Perspectives in Experimental Biology, vol. 1, Pergamon Press, 1976*)

lung (low pressure) and body circuits (high pressure), yet the varanid ventricle is morphologically undivided. This extraordinary triumph of physiology over morphology is achieved by partitioning the ventricular cavity during the contraction and relaxation phases. An interventricular partition formed by the atrioventricular valves divides the ventricle during diastole, whereas a musclar ridge within the ventricle divides it at a different site during systole. Crocodilians also have a high-pressure body and low-pressure lung circulation. The heart is completely divided by an interventricular septum, but the left aorta, supplying the body, arises along with the pulmonary artery from the right ventricle. Furthermore, the left and right aortas are joined just outside the heart by a hole in their common wall (foramen of Panizza) and again by a connection behind the heart. Hence, in the crocodilians a unique situation exists, because it is possible for venous blood to be transferred away from the lung circulation and sent back to the body through the left aorta so that both flows and pressures are independent of one another.

In birds and mammals the heart is completely divided, and the flows in both lung and body circuits must be matched. This matching of flows is largely achieved by nervous control although an automatic mechanism, referred to as Starling's law of the heart, could play a role. Starling's law (which also applies to the hearts of lower vertebrates) states that the energy of contraction of ventricular muscle is a function of the length of the muscle fiber. Thus, if in a particular beat a ventricle is filled to a greater extent than the previous one, the next contraction would be more vigorous and a greater volume of blood (stroke volume) would be ejected. So, if there is an increase in pumping by the right heart, a few beats later more blood will return to the left heart and increase its filling, so its output will rise, maintaining a balance between flows in the two circuits without intervention of any nervous or humoral (blood-borne) control mechanisms.

Pacemaker. The cardiac rhythm is myogenic in all vertebrates, the heartbeat being initiated in a specialized group of muscle cells which form the pacemaker. From the pacemaker, a wave of electrical excitation passes across the heart, activating the contractile process in cardiac muscle. This wave of excitation may pass across the muscle tissue itself or along specialized conducting pathways (Purkinje fibers, which are modified muscle cells; **Fig. 26**). Pacemaker cells are different from other cardiac cells in that the electrical charge across their cell membranes spontaneously and repetitively declines (depolarization). These repeated depolarizations are called pacemaker potentials. They reach a threshold, after a depolarization of some 15 mV, and an action potential is generated (a sudden reversal of the membrane electrical charge). The action potential is propagated from cell to cell, initiating muscle contraction (Fig. 26). The cycle then repeats. An unusual feature of the cardiac action potential is its long duration, compared, for instance, with the action potential in skeletal muscle. If a muscle action potential is short, then further action potentials and contractions can be generated before the muscle relaxes. Consequently, in skeletal muscle a sequence

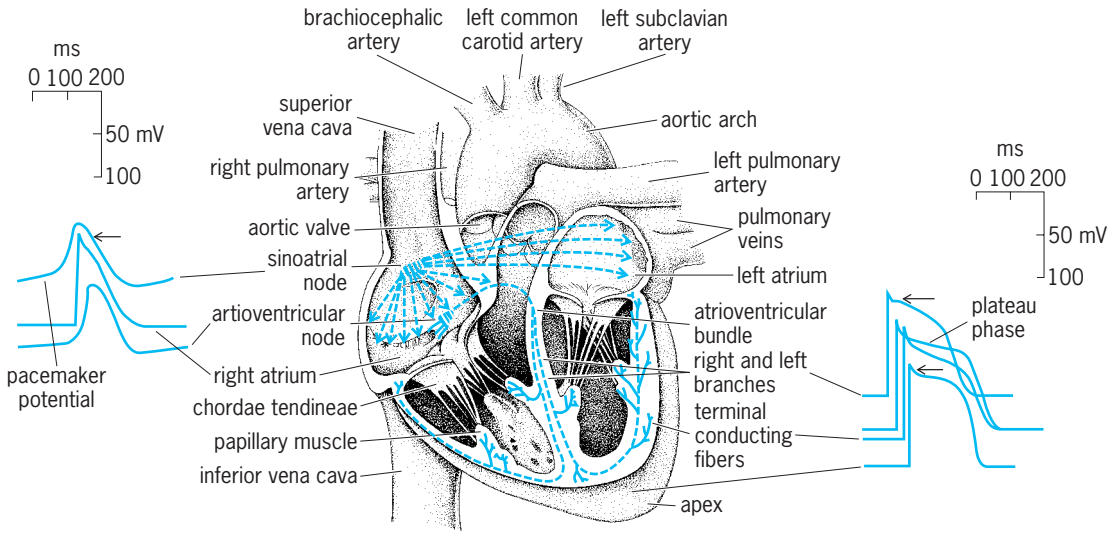

Fig. 26. Anterior view of the opened human heart, showing the pacemaker in the sinus node, the atrioventricular node, and the specialized conduction pathways of the ventricle. Intracellular action potential shapes were recorded at identified sites of the heart during one cardiac cycle. The arrows link the action potentials to the site. Time and voltage calibrations are shown at the left and right. Zero transmembrane voltage is indicated on some of the action potentials by the arrowhead. (*After J. E. Crouch, Functional Human Anatomy, Lea and Febinger, 1978*)

of contractions may sum. However, in the heart the shortest possible interval between two action potentials is still longer than the relaxation time of cardiac muscle. Thus, the sustained contractions (tetanus) characteristic of skeletal muscle, and due to the fusion of individual waves of contraction, cannot occur in cardiac muscle. Tetanus in cardiac muscle would, of course, be maladaptive.

The sum of the electrical events occurring during the synchronous activity of all the cardiac cells can be detected as a series of small voltage changes at points all over the body (Fig. 24*a* and *b*). The record of small voltage changes is the electrocardiogram; it exhibits waves associated with atrial depolarization (P-wave), ventricular depolarization (QRS complex), and repolarization (T-wave) [Fig. 24*a* and *b*].

In all vertebrate hearts there is a hierarchy of pacemakers. This is reflected in embryonic development where the ventricle is formed and begins beating before the other parts of the heart have differentiated. The ventricular pacemaker sets a slow rate and is superseded by pacemakers, first in the atrium, and ultimately in the sinus venosus. In fish and frogs heart rate is set by the pacemaker in the sinus venosus because it beats fastest and drives the other pacemakers. In birds and mammals the sinus venous is incorporated into the right atrium, and the pacemaker zone is called the sinoatrial node (Fig. 26). The wave of excitation crosses the atrium (in some mammals specialized conducting pathways have been described) at a rate of up to 1 $m \cdot s^{-1}$, but it can only cross to the ventricle through the atrioventricular node. In the atrioventricular node, the rate of conduction slows to 0.05–0.1 $m \cdot s^{-1}$, allowing time for the atrium to complete its contraction before the ventricle is activated. From the atrioventricular node the wave of excitation is propagated rapidly (0.05–2.5 $m \cdot s^{-1}$). In birds and mammals propagation takes place through specialized conducting fibers called, in mammals, the right and left bundles of His. Lower vertebrates lack specialized conducting fibers and even a discrete atrioventricular node; in noncrocodilian reptiles there is an almost complete ring of junctional tissue between the atria and ventricle.

In all except the lowest vertebrates the pacemaker region receives innervation both from excitatory (sympathetic) and depressor (parasympathetic) nerves. The sympathetic nerves liberate catecholamines which increase the rate of spontaneous depolarization, and therefore heart rate. The parasympathetic nerves liberate acetylcholine which stabilizes the membrane potential and decreases the rate of spontaneous depolarization, so heart rate falls. The extent to which these nerves innervate cardiac muscle or other structures, such as the atrioventricular node, is variable. However, when muscle innervation is dense, catecholamines increase the force of contraction (positive inotropic effect) while acetylcholine decreases it (negative inotropic effect).

The effects of chemicals released by nerves are due to the activation of receptors on the cardiac cell which are specific for that particular chemical. For noradrenaline and adrenaline the receptors are called α and β receptors, respectively. Adrenaline and noradrenaline increase in the circulation in response to stress and stimulate the heart. In many vertebrates, however, the effects are different from those caused by nerves. For instance, circulating adrenaline is unable to reach the same receptors which are activated neurally because the nerve ending is so closely attached to the cardiac cell. Receptors stimulated by nerves have different effects on pacemaker potentials from those accessed from the circulation.

Arteries. The arteries are the connecting tubes between the heart and the microcirculatory vessels.

They are largest and most distensible just outside the heart. The arteries decrease in diameter and flexibility at every bifurcation, so the arterial tree is said to display both geometric and elastic taper. The vessels within the thorax are extremely distensible due to a much higher proportion of rubbery (elastin) to stiff (collagen) fibers in their walls. In nonthoracic arteries collagen dominates the wall composition, but its proportion remains constant as the vessels get stiffer and stiffer with their approach to the periphery.

During the ejection phase of the cardiac cycle much of the blood is temporarily stored in the central arteries as these become distended by the pressure rise. This blood is fed into the peripheral circulation, by the rebound of the stretched elastic walls, throughout diastole. Consequently, a highly pulsatile input is transformed into a more even outflow (**Fig. 27***a*). However, in a large number of animals (such as humans, dogs, and ducks) this is not the case with the pressure pulse. The pressure pulse is amplified in the peripheral vessels (peaking), and both the size and rate of rise of the pulse wave is greater than in the more central vessels (Fig. 27*a*). On the other hand, small high-frequency components such as the incisura disappear from the peripheral pulse as they are damped out (Fig. 27*a*).

The pressure pulse wave travels through the arteries at a velocity which depends on the viscosity of the fluid and distensibility of the vessel. In humans it picks up speed, from 3 m · s^{-1} centrally, to 5–10 m · s^{-1} in peripheral arteries. At every discontinuity in the arterial tree (regions of geometric and elastic taper and, more importantly, the terminal vascular beds) the incident wave will be reflected toward the heart. At any major reflecting sites, such as the terminal vascular beds, the incident wave and the reflected wave will be almost in phase and will sum so that the pressure pulse increases in size (Fig. 27*a*). If the heart is positioned one-quarter of a wavelength back from the major reflecting sites (wavelength is the pulse velocity divided by the cycle length), the incident wave will have passed here one-quarter wavelength before (with respect to the reflecting sites) and the reflected wave will arrive here one-quarter wavelength later (with respect to the reflecting sites) so the two waves will be out of phase, by half a wavelength, and will cancel one another out. Hence, the pressure pulsations in central arteries will be reduced by wave reflection (Fig. 27*a*). In mammals about 30% of cardiac power is invested in pulsatile pressure and subsequent flow, so by reducing the pressure pulsations just outside the heart, energy can be saved. Peaking occurs only if the transit time is a significant proportion of the cardiac cycle. In humans mean pulse wave velocity is around 5–6 m · s^{-1} at a cardiac interval of 1 s, so the transit time is about a quarter of the cardiac cycle (quarter wavelength) and, as described above, pulse amplification occurs. However, arteriosclerotic individuals (be they ducks, dogs, or humans) do not show peaking of the pressure pulse since the pulse wave velocity in the "hardened" arteries is extremely fast and transit time through the circulation is greatly reduced. This lack of pulse wave amplification may contribute to the deleterious effects of this condition on the heart. In frogs mean pulse wave velocity is around 3 m · s^{-1} at a cardiac interval of 1 s, so the transit time is about 50 ms, and the shape of the peripheral pulse in frogs is, except for damping of fast components (incisura), identical in shape to that in central arteries.

As blood is pumped through the circulation, energy is dissipated and mean pressure drops (Fig. 27*a*). Mean pressure is a steady pressure obtained by integrating the pressure pulse over one or several cardiac cycles and is somewhat lower than the arithmetic mean of systolic and diastolic pressures. When peaking occurs, the pressure pulse gets bigger but much narrower (Fig. 27*a*), and mean pressure is still lower in the peripheral arteries compared with the central arteries. Therefore, the flow occurs down the mean pressure gradient.

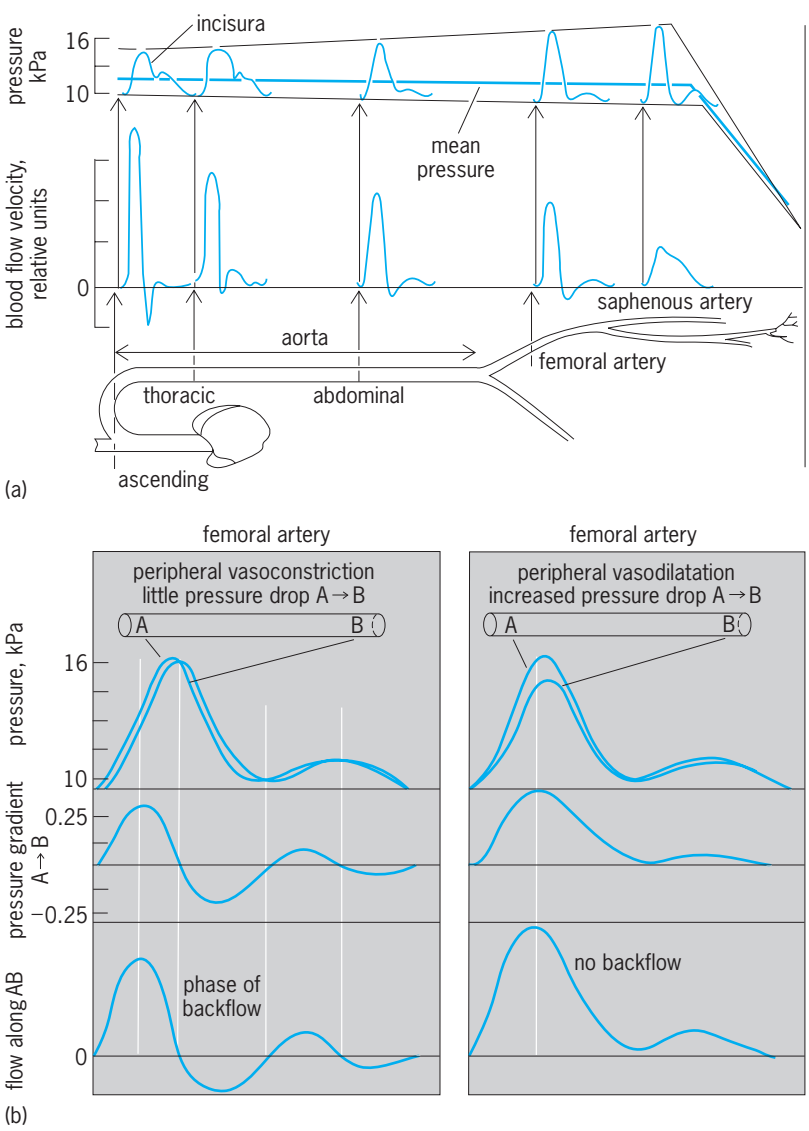

Fig. 27. Arterial circulation. (*a*) Comparison of the behavior of the pressure and flow velocity pulses from the ascending aorta until the saphenous artery. (*b*) The propagation of the pulse wave along an artery, with deduction of the pressure gradient and the phasic flow. (*After B. Folkow and E. Neil, Circulation, Oxford University Press, 1971*)

Unfortunately, an analysis in terms of steady pressures and flows tells little about the shapes of the pressure and flow pulses. Why does flow reverse at certain portions of the cardiac cycle? The reason is that the pulsatile pressure gradient reverses in those phases (Fig. 27*b*). The pressure pulse travels through the arterial tree with a velocity from 3 to 10 $m \cdot s^{-1}$. If pressure is recorded at two sites in a vessel that are a small distance apart, then as the wave passes the upstream point, pressure will exceed that downstream and forward flow will occur (Fig. 27*b*). When the wave moves on and passes the downstream point, pressure there will exceed the pressure upstream and the flow will reverse (Fig. 27*b*). In fact, it is the pressure difference between the upstream and downstream points (the pressure gradient) which oscillates about a mean and is closely linked to the flow velocity (Fig. 27*b*).

Arterial blood pressure is controlled, at least in the short term, through the agency of arterial baroreceptors. Baroreceptors are specialized nerve endings located in the walls of major vessels; they are stimulated when the wall is expanded by the blood pressure. They have been found in the aortic arch and carotid sinus of mammals, the aortic arch of birds, and the carotid labyrinth of amphibians. Claims have been made for their existence in fishes. Blood pressure is controlled using the negative feedback principle. A rise in blood pressure stimulates baroreceptors which send electrical messages (action potentials) to the brain where they are directed to regions involved in control and adjustment of heart rate, stroke volume, and resistance in the peripheral blood vessels. In response to a rise in blood pressure, heart rate, stroke volume, and peripheral resistance are reduced. Input from mechanically stimulated receptors in the heart and lungs, and from chemoreceptors (measuring blood oxygen and carbon dioxide) in the vascular system and brain, are also involved in regulating the circulatory system.

Venous system. From the capillaries the blood passes through venules to successively larger veins and returns to the heart. Aside from being conduits for returning blood to the heart, the major function of the venous system is as a storage site for blood (Fig. 23). Veins are thin-walled, although the larger ones have a muscular coat, and are distended by small changes in pressure across their walls (transmural pressure). A pressure rise causes the veins to change from an elliptical cross-sectional profile to one that is nearly circular, greatly increasing blood storage while reducing flow resistance. Contraction of the muscles in the walls of the larger veins (venomotor activity) reverses the effects of increases in transmural pressure. Venoconstriction is under the control of the sympathetic nervous system.

In larger veins, valves appear as intimal folds and ensure that flow moves only in one direction. In humans some veins such as the vena cava, hepatic, pulmonary, and cerebral veins lack valves, a feature shared with the major veins running the length of the body in fishes. In elasmobranchs, these long veins have a unique structure; they are invested with such a thick connective tissue sheath that they are virtually incompressible tubes. Elasmobranch fishes and some seals also have a muscular venous sphincter located between the liver and heart which regulates venous return when venous pressure rises during exercise (elasmobranch) or diving (seal).

In most animals the majority of the venous reservoir is placed level with or above the heart (even in giraffes). However, in erect animals (primarily humans) the majority of the venous reservoir is below the heart. Consequently, due to the hydrostatic effect (as described above), transmural pressure will increase in veins below the heart and will decrease in those above. Hence, when the animal is upright, leg veins distend and head veins may collapse. This hydrostatic effect has no "direct" effect on flow (aside from the fact that flow resistance will fall as vessels distend or increase if they collapse) since it is added to both arterial and venous sides of the circulation.

Venous return is caused by the forward push of the blood generated by the heart and transmitted in the form of a positive pressure across the capillaries (*vis à tergo*). In the pulmonary circulation and, on the body side in some lower vertebrates, the arterial pressure pulse may be transmitted across the capillaries. Even when these pulsations are damped out, venous pressure pulsations occur in central veins as heart movements are transmitted "backward" to these veins. In some animals venous return is promoted by a suction force (*vis à fronte*) due to subatmospheric pressures in the cardiac cavities caused either by a negative intrapericardial pressure stretching the atrium, as in elasmobranchs, or, in animals lacking a rigid pericardium, to "elastic recoil" of the ventricles, or atrial volume changes associated with ventricular relaxation. When veins with valves run through blocks of skeletal muscle, contraction of these muscles squeezes blood toward the heart (muscle pump).

The muscle pump works in all vertebrates; in some fishes, even the arteries within the muscle mass have valves to ensure unidirectional flow of blood. In animals that have a diaphragm, inspiration aids venous return; since pressure within the thorax is at subatmospheric level, the transmural pressure of intrathoracic veins increases, and they expand. At the same time, abdominal pressure rises, so the transmural pressure in extrathoracic veins will decrease and they will be compressed, thus forcing blood toward the heart. Hence, at the diaphragm a sharp drop in venous pressure occurs which is referred to as a vascular waterfall.

David R. Jones

Bibliography. C. R. Austin and R. V. Short, *Embryonic and Fetal Development*, 1983; H. S. Badeer, *Cardiovascular Physiology*, 1983; R. W. Blake (ed.), *Efficiency and Economy in Animal Physiology*, 1991; W. W. Nichols and M. F. Rourke, *McDonald's Blood Flow in Arteries*, 4th ed., 1997; C. L. Prosser, *Comparative Animal Physiology*, 4th ed., 1991; A. S. Romer and T. S. Parsons, *The Vertebrate Body*, 6th ed., 1986.

Caribou

A ruminant, *Rangifer tarandus*, of the deer family, Cervidae. Reindeer inhabit the Arctic region and have a circumpolar distribution. They have been domesticated for centuries and are economically important to the Laplanders, who use them as draft animals as well as for their skins, flesh, and milk. The Laplanders are nomadic people who migrate during the summer months with their herds from one feeding ground to another.

Reindeer do not display sexual dimorphism; both sexes have antlers and are brown with yellow-white areas on the neck and chest.

Both sexes have antlers (see **illus.**), a characteristic peculiar to this deer species. Although their senses of sight and hearing are not sharp, their sense of smell is quite keen. Reindeer are well adapted for life in the Arctic; the hooves are round and broad, giving the animal stable footing in snow and on ice. In the wild, they are the most migratory of all deer, traveling in large herds in search of lichens as a food source during the winter months. The rutting season begins during the autumn migrations; fawns are born during the spring. *See* ARTIODACTYLA; DEER; MIGRATORY BEHAVIOR; MOOSE.

Charles B. Curtin

Carnauba wax

Product exuded from the leaves of the wax palm, *Copernicia cerifera* (Palmae), a native of Brazil and other regions in tropical South America. The wax accumulates on the surface of the leaves; these are cut from the trees and dried, whereupon the layer of wax becomes loose and is easily removed by flailing. It is the hardest, highest-melting natural wax and is used in making candles, shoe polish, high-luster wax, varnishes, phonograph records, and surface coating of automobiles. *See* ARECALES; WAX, ANIMAL AND VEGETABLE.

Earl L. Core

Carnivora

One of the larger orders of placental mammals, including fossil and living dogs, raccoons, pandas, bears, weasels, skunks, badgers, otters, mongooses, civets, cats, hyenas, seals, walruses, and many extinct groups organized into 12 families, with about 112 living genera and more than twice as many extinct genera. The primary adaptation in this order was for predation on other vertebrates and invertebrates. A few carnivorans (for example, bear and panda) have secondarily become largely or entirely herbivorous, but even then the ancestral adaptations for predation are still clearly evident in the structure of the teeth and jaws. The Carnivora have been highly successful animals since their first appearance in the early Paleocene. During the early Tertiary they shared the predatory mode of life with carnivorous members of some other placental orders, particularly the Deltatheridia, to the extent that few other mammals succeeded in invading this adaptive zone. *See* CREODONTA.

Structural adaptations. Meat-eating adaptations involving the teeth and jaws appeared early in the history of Carnivora. The dentition is sharply divided into three functional units. The incisors act as a tool for nipping and delicate prehension, and the large interlocking upper and lower canines for heavy piercing and tearing during the killing of prey. The cheek teeth are divided into premolars (for heavy prehension) and molars (for slicing and grinding), which may be variously modified depending on the specific adaptation. A diagnostic feature of the Carnivora is the enlargement of the last (fourth) upper premolar and the first lower molar to form longitudinally aligned opposed shearing blades. This important pair of teeth are termed the carnassials. Primitively, the dental formula was that common to all early placentals, but in most carnivores there is often loss of anterior premolars and posterior molars, with the strictly carnivorous forms, such as cats, reducing the number of cheek teeth the most (**Fig. 1**). *See* DENTITION.

In all carnivorans the jaw articulation is arranged in such a manner that movement is limited to vertical

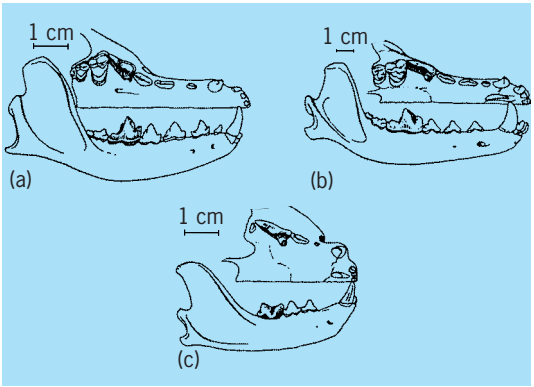

Fig. 1. Upper and lower dentition of the right side representing the superfamilies of Carnivora, carnassials shaded. (*a*) Miacoidea, *Miacis*, Eocene. (*b*) Canoidea, *Canis*, Recent. (*c*) Feloidea, *Felis*, Recent.

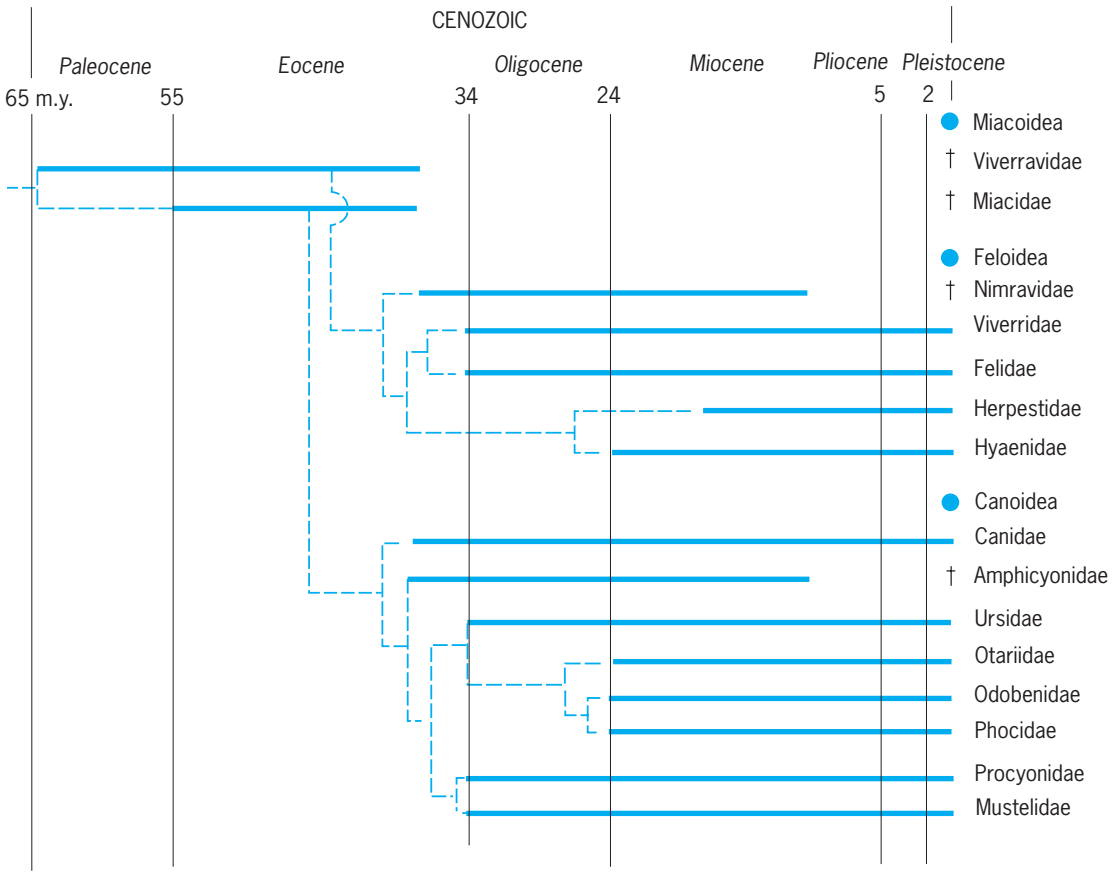

Fig. 2. Phylogenetic relationships and geologic ranges of the carnivoran families. The heavy lines indicate known geologic ranges, and the broken lines indicate predicted ranges inferred from phylogenetic relationships. Daggers indicate extinct families.

hinge motions and transverse sliding. The temporal muscle dominates the jaw musculature, forming at least one-half of the total mass of the jaw muscles. Cursorial adaptations of the limbs appeared early in carnivore history and have always been characteristic of the order. In all modern carnivorans the brain is large and the cerebral hemispheres are highly developed; evolution of the brain seems to have proceeded at an accelerated pace during the latter half of the Cenozoic.

Phylogeny. The history of the Carnivora is relatively well documented despite the fact that fossils are rarely abundant. The earliest records are early Paleocene, but the earliest well-represented material comes from the middle Paleocene of North America. During the Paleocene and Eocene the stem-carnivorans, or miacoids, underwent considerable diversification in both the Old and New worlds. These animals were small- to medium-sized, usually with the full eutherian dental formula, and the fourth upper premolars and first lower molars were developed as carnassials. The feet were plantigrade or subdigitigrade with separate scaphoid and lunar bones in the wrist. The middle-ear region lacked a completely ossified cover or bulla, although the bony tympanic ring was present. The internal carotid artery was primitively three-branched (median, promontory, and stapedial branches).

At the end of Eocene and beginning of Oligocene time throughout the Northern Hemisphere, a dramatic change took place within the Carnivora—the appearance of primitive representatives of modern carnivoran families (**Fig. 2**). These animals had more cursorial limbs than their miacoid predecessors, with fused scaphoid and lunar bones, and completely ossified bullae, and the internal carotid was reduced to a two-branched (median and promontory) or single (median) vessel. Commonly the dentition showed reduction or loss of anterior premolars and posterior molars. There is little doubt that the miacoids are the progenitors of the earliest members of the modern carnivoran families, but analysis of available fossil evidence has allowed only limited determination of the exact lineages.

A second adaptive radiation of the Carnivora was initiated by these primitive members of the living families. Eventually this radiation brought about the extinction of contemporary carnivores, the Creodonta (family Hyaenodontidae did persist into late Miocene in the Old World tropics). During this second adaptive radiation of the Carnivora, a wider variety of ecologic niches was exploited than is shown by the fossil record of the early Cenozoic radiation. Striking among these are the development of crushing dentitions for omnivorous diets, loss of teeth for ant and termite feeding, hypertrophy of

both premolars and molars for bone crushing, and most spectacularly the total modifications of skeleton and soft parts for aquatic existence in the open ocean.

Taxonomy. Historically, the order Carnivora has been divided into suborders Creodonta, Fissipeda, and Pinnipedia. Research on early Cenozoic fossils casts serious doubt on the affinity of the creodonts with the remainder of the Carnivora; they appear to be a convergent group of eutherians, some of which occupied the same adaptive zone as the Fissipeda. They are now placed in their own order, Creodonta. Furthermore, the Fissipeda and Pinnipedia do not denote a primary subdivision of the Carnivora as historically conceived. Although the living pelagic carnivorans and fossil allies (Pinnipedia) do seem to form a natural (that is, monophyletic) group, they originate from carnivorans embedded within the fissiped group. Such a twofold subdivision of the order Carnivora no longer expresses knowledge of the group's phylogeny. A threefold subdivision of the order has long been practiced, and the following superfamilies seem a better expression of the phylogenetic relationships: Miacoidea, Canoidea, and Feloidea. *See* ARCHAIC UNGULATE; EUTHERIA.

Miacoidea. The basic features of these ancestral carnivorans are mentioned above. Two families, the Miacidae and the Viverravidae, are often recognized. These animals occur in early Paleocene through late Eocene deposits in North America. In Europe they are known only from the Eocene. They have not been recognized in Africa. Members of the Viverravidae occur earliest in the fossil record. They have reduced dentitions that suggest special affinity with the Feloidea. The miacids include forms that resemble the Canoidea, and some may eventually be recognized as members of canoid families.

Canoidea. This group of doglike carnivorans first appears in the late Eocene of North America and Eurasia. They possess an auditory bulla, formed from both the entotympanic and ectotympanic (tympanic ring), that primitively lacks an internal septum. The Canoidea consists of seven living families: Canidae (dogs and their allies), Procyonidae (raccoons and their allies), Ursidae (bears and their allies), Mustelidae (weasels and their allies), and the pinniped families, Otariidae (sea lions), Odobenidae (walruses), and Phocidae (seals). All the terrestrial families appear in the fossil record during the Oligocene, and there seems little doubt that they are a closely related group. The pinniped families appear to be related to an extinct group that lies within the Ursidae. The fossil record reveals an eighth canoid family, the Amphicyonidae, that lies near the base of the canoid phyletic tree. This group, which included giant predators in the Miocene, persisted into the late Miocene of Eurasia. *See* BADGER; BEAR; COATI; DOGS; FERRET; FISHER; MARTEN; MINK; OTTER; PANDA; PINNIPEDS; RACCOON; SKUNK; WEASEL; WOLVERINE.

Feloidea. The catlike carnivorans also first appear in the late Eocene or early Oligocene in both the Old and New worlds. Four living families are currently recognized: Viverridae (civets) Herpestidae (mongooses), Hyaenidae (hyenas and aardwolf), and Felidae (cats and saber-toothed cats). Nearly all the fossil and living members of these families possess an auditory bulla divided by a septum formed at the junction of the entotympanic and ectotympanic bones. An exception is the living African palm civet (*Nandinia*) in which only the ectotympanic is fully ossified and the entotympanic is cartilaginous or poorly ossified. In addition, *Nandinia* appears to represent the primitive condition of the bulla in feloids. The herpestids and hyaenids appear in early to medial Miocene time and may be the youngest terrestrial carnivoran families to arise. Members of extinct family of feloids, the Nimravidae, often assigned to the Felidae, retained their primitive cranial features although possessing felid teeth and skeletons. Representatives of this group appeared first in the late Eocene of North America, and by the Oligocene the nimravids were distributed across the continents of the Northern Hemisphere. All nimravid species have bladelike upper canine teeth, and in some, such as the late Cenozoic saber-toothed felids, the canines are greatly lengthened. Saberlike upper canines also evolved in other orders of carnivorous mammals such as the Credonta (family Hyaenodontidae) and the Marsupialia (family Borhyaenidae), and at least twice within the Carnivora (families Nimravidae and Felidae), providing a striking case of convergence in a specialized feeding mechanism. *See* CAT; CIVET; HYENA; MAMMALIA; MONGOOSE. Richard H. Tedford

Carnot cycle

A hypothetical thermodynamic cycle originated by Sadi Carnot and used as a standard of comparison for actual cycles. The Carnot cycle shows that, even under ideal conditions, a heat engine cannot convert all the heat energy supplied to it into mechanical energy; some of the heat energy must be rejected. In a Carnot cycle, an engine accepts heat energy from a high-temperature source, or hot body, converts part of the received energy into mechanical (or electrical) work, and rejects the remainder to a low-temperature sink, or cold body. The greater the temperature difference between the source and sink, the greater the efficiency of the heat engine.

The Carnot cycle (**Fig. 1**) consists first of an isentropic compression, then an isothermal heat addition, followed by an isothermal expansion, and concludes with an isentropic heat rejection process. In short, the processes are compression, addition of heat, expansion, and rejection of heat, all in a qualified and definite manner.

Processes. The air-standard engine, in which air alone constitutes the working medium, illustrates the Carnot cycle. A cylinder of air has perfectly insulated walls and a frictionless piston. The top of the cylinder, called the cylinder head, can either be covered with a thermal insulator, or, if the insulation is removed, can serve as a heat transfer surface for heating or cooling the cylinder contents.

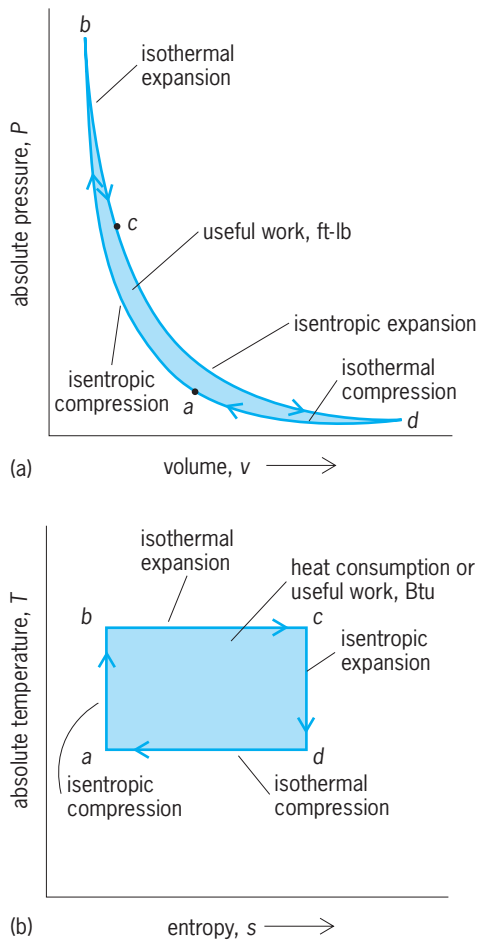

Fig. 1. Carnot cycle for air. (*a*) Pressure-volume diagram. (*b*) Temperature-entropy diagram.

Initially, the piston is somewhere between the top and the bottom of the engine's stroke, and the air is at some corresponding intermediate pressure but at low temperature. Insulation covers the cylinder head. By employing mechanical work from the surroundings, the system undergoes a reversible adiabatic, or an isentropic, compression. With no heat transfer, this compression process raises both the pressure and the temperature of the air, and is shown as the path *a-b* on Fig. 1.

After the isentropic compression carries the piston to the top of its stroke, the piston is ready to reverse its direction and start down. The second process is one of constant-temperature heat addition. The insulation is removed from the cylinder head, and a heat source, or hot body, applied that is so large that any heat flow from it will not affect its temperature. The hot body is at a temperature just barely higher than that of the gas it is to heat. The temperature gradient is so small it is considered reversible; that is, if the temperature changed slightly the heat might flow in the other direction, from the gas into the hot body. In the heat addition process, enough heat flows from the hot body into the gas to maintain the temperature of the gas while it slowly expands and does useful work on the surroundings. All the heat is added to the working substance at this constant top temperature of the cycle. This second process is shown as *b-c* on Fig. 1.

Part way down the cylinder, the piston is stopped; the hot body is removed from the cylinder head, and an insulating cover is put in its place. Then the third process of the cycle begins; it is a frictionless expansion, devoid of heat transfer, and carries the piston to the bottom of its travel. This isentropic expansion reduces both the pressure and the temperature to the bottom values of the cycle. For comparable piston motion, this isentropic expansion drops the pressure to a greater extent than the isothermal process would do. The path *c-d* on Fig. 1 represents this third process, and is steeper on the P-v plane than process *b-c*.

The last process is the return of the piston to the same position in the cylinder as at the start. This last process is an isothermal compression and simultaneous rejection of heat to a cold body which has replaced the insulation on the cylinder head. Again, the cold body is so large that heat flow to it does not change its temperature, and its temperature is only infinitesimally lower than that of the gas in the system. Thus, the heat rejected during the cycle flows from the system at a constant low-temperature level. The path *d-a* on Fig. 1 shows this last process.

The net effect of the cycle is that heat is added at a constant high temperature, somewhat less heat is rejected at a constant low temperature, and the algebraic sum of these heat quantities is equal to the work done by the cycle.

Figure 1 shows a Carnot cycle when air is used as the working substance. The P-v diagram for this cycle changes somewhat when a vapor or a liquid is used, or when a phase change occurs during the cycle, but the T-s diagram always remains a rectangle regardless of phase changes or of working substances employed.

It is significant that this cycle is always a rectangle on the T-s plane, independent of substances used, for Carnot was thus able to show that neither pressure, volume, nor any other factor except temperature could affect the thermal efficiency of his cycle. Raising the hot-body temperature raises the upper boundary of the rectangular figure, increases the area, and thereby increases the work done and the efficiency, because this area represents the net work output of the cycle. Similarly, lowering the cold-body temperature increases the area, the work done, and the efficiency. In practice, nature establishes the temperature of the coldest body available, such as the temperature of ambient air or river water, and the bottom line of the rectangle cannot circumvent this natural limit.

The thermal efficiency of the Carnot cycle is solely a function of the temperature at which heat is added (phase *b-c*) and the temperature at which heat is rejected (phase *d-a*) [Fig. 1]. The rectangular area of the *T-s* diagram represents the work done in the cycle so that thermal efficiency, which is the ratio of work done to the heat added, equals $(T_{hot} - T_{cold})T_{hot}$. For

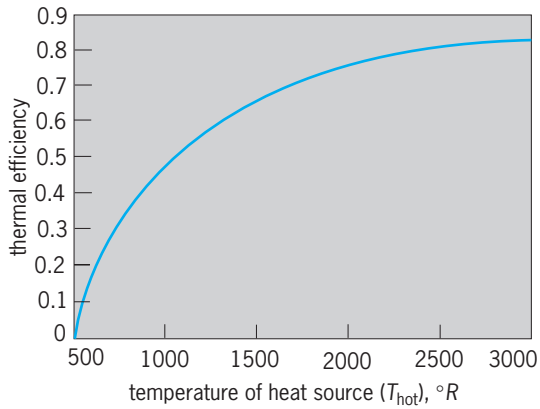

Fig. 2. Thermal efficiency of the Carnot cycle with heat-rejection temperature T_{cold} equal to 500°R (278 K).

the case of atmospheric temperature for the heat sink ($T_{cold} = 500°R$), the thermal efficiency, as a function of the temperature of the heat source, T_{hot}, is shown in **Fig. 2**.

Carnot cycle with steam. If steam is used in a Carnot cycle, it can be handled by the following flow arrangement. Let saturated dry steam at 500°F (260°C) flow to the throttle of a perfect turbine where it expands isentropically down to a pressure corresponding to a saturation temperature of the cold body. The exhausted steam from the turbine, which is no longer dry, but contains several percent moisture, is led to a heat exchanger called a condenser. In this device there is a constant-pressure, constant-temperature, heat-rejection process during which more of the steam with a particular predetermined amount of condensed liquid is then handled by an ideal compression device. The isentropically compressed mixture may emerge from the compressor as completely saturated liquid at the saturation pressure corresponding to the hot-body temperature. The cycle is closed by the hot body's evaporating the liquid to dry saturated vapor ready to flow to the turbine.

The cycle is depicted in **Fig. 3** by *c-d* as the isentropic expansion in the turbine; *d-a* is the constant-temperature condensation and heat rejection to the cold body; *a-b* is the isentropic compression; and *b-c* is the constant-temperature boiling by heat transferred from the hot body. *See* STEAM ENGINE.

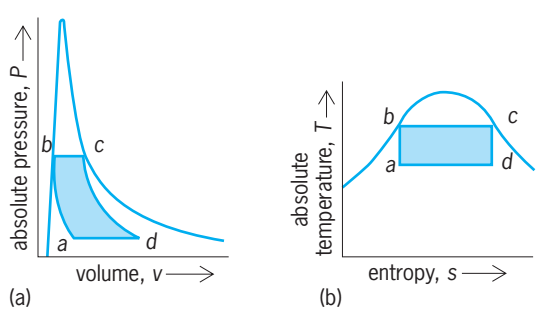

Fig. 3. Carnot cycle with steam. (*a*) Pressure-volume diagram. (*b*) Temperature-entropy diagram.

Carnot cycle with radiant energy. Because a Carnot cycle can be carried out with any arbitrary system, it has been analyzed when the working substance was considered to be a batch of radiant energy in an evacuated cylinder. If the system boundaries are perfectly reflecting thermal insulators, the enclosed radiant energy will be reflected and re-reflected with no loss of radiant energy and no change of wavelength.

The electromagnetic theory of radiation asserts that the radiant energy applies pressure P to the cylinder walls. This radiation pressure is equal to $u/3$, where u represents the radiant energy density, or the amount of radiant energy per unit volume.

The piston moves so that the cylinder boundaries expand, and additional radiant energy is supplied to the system so that the temperature remains constant. Then, cutting off any further supply of radiant energy, the system undergoes a further infinitesimal expansion with its associated pressure drop and temperature drop. The third process is an isothermal compression at the low-temperature level, accompanied by some rejection of radiant energy. One last process closes the cycle with a reversible compression that raises the temperature to the original level. The assumption is made throughout this analysis that energy density is a function of temperature alone. Thus, because this last compression process increases the energy density, it raises the temperature.

A record of all energy quantities in this radiant-energy Carnot cycle indicates that energy density u is proportional to the fourth power of the absolute temperature. Consequently the total rate of emission of radiant energy from the surface of a blackbody is also proportional to the fourth power of its absolute temperature, thereby using a Carnot cycle to provide a relationship by theoretical analysis, the same relationship which had previously been determined experimentally and labeled as Stefan's law.

Conversion of heat to electricity. Several physical phenomena convert heat energy directly into electrical energy. The extent of this direct conversion of heat to electricity is limited by the temperature levels between which the process operates. The ideal efficiency of such direct-conversion thermoelectric cycles equals the efficiency of a Carnot cycle that operates between the same temperature limits of heat source and heat sink. However, the conversion efficiency obtained in practice is only a small fraction of the ideal efficiency at the present stage of development.

The most widely known physical arrangement for direct thermoelectric generation is the thermocouple, which produces an electromotive force, or voltage, when one junction of two dissimilar conductors is heated while the other junction is kept cool. Thermocouples made of metals are inefficient converters of heat energy to electric energy, because metals that have good electrical conductivity unfortunately have equally good thermal conductivity, which permits heat loss by conduction from the hot to the cold junction. In contrast, thermocouples made of

semiconductors offer the prospect of operation at high temperatures and with high temperature gradients, because semiconductors are good electrical conductors but poor heat conductors. Semiconductor thermocouples may have as large a junction potential as the metal couples do. *See* THERMOCOUPLE.

Thermionic emission is another phenomenon that permits the partial conversion of heat energy directly into electrical energy. Externally applied heat imparts kinetic energy to electrons, liberating them from a surface. The density of the emission current is a function of the absolute temperature and work function of the emitter material.

For many years thermionic emission received little attention as a source of power because of its very low conversion efficiency. However, interest was stimulated by development of a contact potential thermionic emission cell. In this device, current flows between the surfaces of two materials which have different work functions. These materials are held at different temperatures, and the gap between the electrode surfaces is filled with gas at low pressure. *See* THERMIONIC EMISSION.

Such direct-conversion techniques show promise of becoming small-scale, if unconventional, power sources. Although these techniques ideally can convert heat to electrical energy with the efficiency of a Carnot cycle operating between the same temperature levels, laboratory devices do not surpass about 8% efficiency, which is just a fraction of the ideal performance.

Reversed Carnot cycle. A Carnot cycle consists entirely of reversible processes; thus it can theoretically operate to withdraw heat from a cold body and to discharge that heat to a hot body. To do so, the cycle requires work input from its surroundings. The heat equivalent of this work input is also discharged to the hot body.

Just as the Carnot cycle provides the highest efficiency for a power cycle operating between two fixed temperatures, so does the reversed Carnot cycle provide the best coefficient of performance for a device pumping heat from a low temperature to a higher one. This coefficient of performance is defined as the ratio of the quantity of heat pumped to the amount of work required, or it equals $T_{hot}(T_{hot} - T_{cold})$ for a warming machine, and $T_{cold}(T_{hot} - T_{cold})$ for a cooling machine, where all temperatures are in degrees absolute. This is one of the few engineering indices with numerical values greater than unity. *See* HEAT PUMP; REFRIGERATION CYCLE.

Practical limitations. Good as the ideal Carnot cycle may be, there are serious difficulties that emerge when one wishes to make an actual Carnot engine. The method of heat transfer through the cylinder head either limits the operation of the engine to very low speeds, or requires an engine with a huge bore and small stroke. Moreover, the material of the cylinder head is subjected to the full top temperature of the cycle, imposing a metallurgical upper limit on the cycle's temperature. *See* HEAT TRANSFER.

A practical solution to the heat transfer difficulties which beset the Carnot cycle is to burn a fuel in the air inside the engine cylinder. The result is an internal combustion engine that consumes and replaces its working substance while undergoing a periodic sequence of processes.

The working substance in such an internal combustion engine can attain very high temperatures, far above the melting point of the metal of the cylinder walls, because succeeding lower-temperature processes will keep the metal parts adequately cool. Thus, as the contents change temperature rapidly between wide extremes, the metal walls hover near a median temperature. The fuel-air mixture can be ignited by a spark, or by the rise in temperature from the compression stroke. *See* DIESEL CYCLE; DIESEL ENGINE; INTERNAL COMBUSTION ENGINE; OTTO CYCLE.

Even so, the necessarily high peak pressures and temperatures limit the practical thermal efficiency that an actual engine can achieve. The same theoretical efficiency can be obtained from a cycle consisting of two isobaric processes interspersed by two isentropic processes. The isobaric process requires that the cycle handle large volumes of low—pressure gas, which can best be done in a rotating turbine. *See* BRAYTON CYCLE; GAS TURBINE.

Although the Carnot cycle is independent of the working substance, and hence is applicable to a vapor cycle, the difficulty of efficiently compressing a vapor-liquid mixture renders the cycle impractical. In a steam power plant the sequence of states assumed by the working substance is (1) condensate and feedwater are compressed and pumped into the boiler, (2) heat is added to the water first at constantly increasing temperature and then at constant pressure and temperature, (3) the steam expands in the engine, and (4) the cycle is completed by a constant-temperature heat-rejection process which condenses the exhausted steam. *See* VAPOR CYCLE.

By comparison to the Carnot cycle, in which heat is added only at the highest temperature, this steam cycle with heat added over a range of temperatures is necessarily less efficient than is theoretically possible. An analysis of engine operations in terms of thermodynamic cycles indicates what efficiencies can be expected and how the operations should be modified to increase engine performance, such as high compression ratios for reciprocating internal combustion engines and high steam temperatures for steam engines. *See* POWER PLANT; THERMODYNAMIC CYCLE; THERMODYNAMIC PRINCIPLES.

Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; K. E. Bett et al., *Thermodynamics for Chemical Engineers*, 1975; T. D. Eastop and A. McConkey, *Applied Thermodynamics for Engineering Technologists*, 5th ed., 1993; D. P. Tassios, *Applied Chemical Engineering Thermodynamics*, 1993; M. W. Zemansky and R. Dittman, *Heat and Thermodynamics*, 7th ed., 1996.

Carotenoid

Any of a class of yellow, orange, red, and purple pigments that are widely distributed in nature, including vegetables, fruits, insects, fishes, and birds. Most of these pigments are carotenoids, which are generally fat-soluble unless they are complexed with proteins. In plants, carotenoids are usually located in quantity in the grana of chloroplasts in the form of carotenoprotein complexes. Carotenoprotein complexes give blue, green, purple, red, or other colors to the outer surfaces or eggs of crustaceans, such as the lobster and crab. Echinoderms, nudibranch mollusks, and other invertebrate animals also contain carotenoproteins. Some coral coelenterates exhibit purple, pink, orange, or other colors due to carotenoids in the calcareous skeletal material. Cooked or denatured lobster, crab, and shrimp show the modified colors of their carotenoproteins.

Colors exhibited by carotenoid pigments are due to the absorption of quanta of visible wavelength light by the conjugated double-bond network of carotenoids. The absorption of a light quantum by the carotenoid molecule promotes an electron (π electron) of the double-bond system to a higher-energy orbital.

Astaxanthin is the carotenoid component of crustacyanin, a carotenoprotein that determines the color of the lobster shells. The native α-crustacyanin of fresh lobster maximally absorbs light of 632-nanometer wavelength whereas the denatured pigment protein of boiled lobster shows an absorption maximum at 540 nanometers, exhibiting dark green and orange-pink colors, respectively. These color changes are apparently due to an alteration of the interactions among the carotenoid molecules or between the carotenoid and protein moieties upon heat denaturation of the crustacyanin.

The total carotenoid production in nature is estimated at about 110,000,000 tons (100,000,000 metric tons) a year. The widespread occurrence of carotenoids in nature suggests that the distribution pattern of carotenoid pigments may be valuable in chemotaxonomic analyses. Patterns of evolution of algal groups have indeed been proposed on the basis of carotenoid pigmentation patterns. *See* PIGMENTATION.

Chemical structure and classification. The general structure of carotenoids is that of aliphatic and aliphatic-alicyclic polyenes, with a few aromatic-type polyenes (**Fig. 1**). Most carotenoid pigments are tetraterpenes with a 40-carbon (C_{40}) skeleton, traditionally regarded as results from the conjugation of eight isoprene units. C_{30}, C_{45}, and C_{50} carotenoids are also known in nature. Carotenes such as α-, β-, γ-, and δ-carotene are C_{40} carotenoids containing 11 or fewer conjugated carbon-carbon double bonds. Oxygenated carotenoids with hydroxyl, epoxy, ether, aldehyde, ketone, or acid groups are generally called xanthophylls. Carotenoids may exist in either cis or trans isomeric forms, with the latter being more stable than the former. The cis-trans isomerization may be catalyzed by iodine, and blue light of 450-nm wavelength may also induce the isomerization (photoisomerization). *See* PHOTOCHEMISTRY.

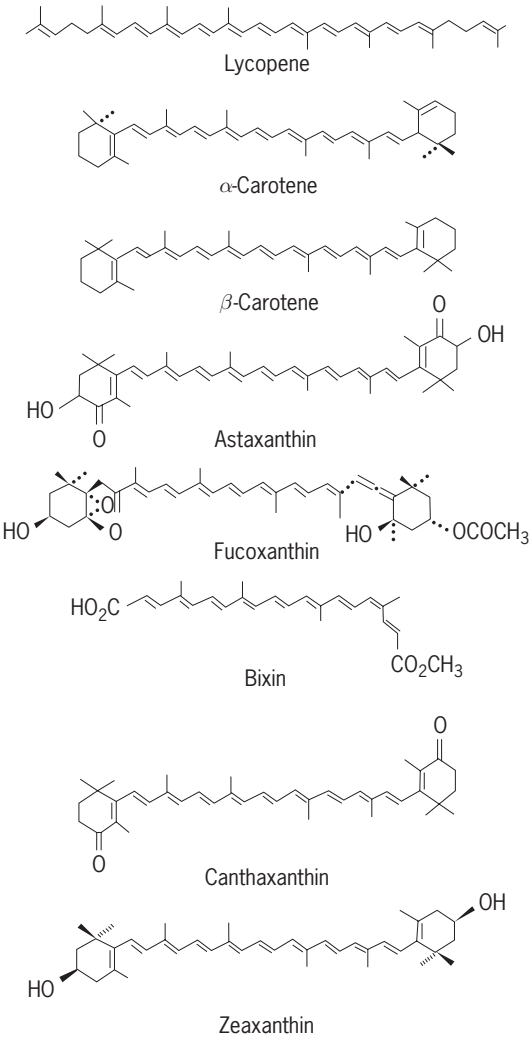

Fig. 1. Chemical structures of typical carotenoids.

More than 400 carotenoids of known structure are recognized, and the number is still on the rise.

Isolation and characterization. A rough separation of carotenes and xanthophylls (oxygen-containing carotenoids) can be achieved by partitioning a crude extract of pigment mixture between two immiscible solvents, petroleum ether or *n*-hexane and aqueous methanol (90%). Petroleum ether or *n*-hexane dissolves carotenes preferentially (epiphasic fraction), whereas aqueous methanol dissolves xanthophylls with two or more hydroxyl groups (hypophasic fraction). Monohydroxy carotenoids tend to be distributed equally between both phases.

Chromatographic methods are most effective and popular for isolation and purification of carotenoids. These include column chromatography, thin-layer chromatography, and high-pressure liquid chromatography, among others. Both column and thin-layer chromatography are based on the differential adsorption of various carotenoids to adsorbents.

Adsorbents chosen for the column method depend on the polarity of the carotenoids to be separated. For example, sucrose powder and magnesium silicate are used for separating strongly polar carotenoids, whereas carotenes are readily separated on an activated alumina column, by eluting them from the column stepwise, with a solvent mixture of increasing polarity. Thin-layer chromatography utilizes adsorbents such as silica gel, alumina, or sucrose, developed with mixtures of organic solvents. Thin-layer chromatography on cellulose layers as the stationary phase can also be used for separating different carotenoids on the basis of their partitioning between the stationary phase and the mobile phase (for example, acetone-methanol-water). Thin-layer chromatographic separation is usually more satisfactory for carotenoids than is column chromatography. **Figure 2** shows an example of the thin-layer chromatographic separation of five representative carotenoids. High-pressure liquid chromatography yields the highest resolution in separating carotenoids for identification and preparative purposes. Stereochemical isomers including cis-trans and enantiomeric isomers can be separated by high-pressure liquid chromatography. For characterization of carotenoids, various chemical and modern physical techniques such as spectroscopic and x-ray crystallographic determinations have been widely used. *See* CHROMATOGRAPHY.

Since carotenoids are blue light–absorbing pigments, absorption spectroscopy is routinely employed for the characterization of carotenoids with regard to the degree of conjugation and geometric isomerism. **Figure 3** illustrates spectral changes of β-carotene upon isomerization of the central carbon-carbon double bond. The near-ultraviolet peak gains its absorptivity, as the electronic system of β-carotene changes from a fully linear all-trans to a bent cis form. In contrast to linear polyenes, cyclic polyenes such as porphyrins exhibit a much stronger near-ultraviolet absorption band (Soret band) than a visible light band. Thus isomerization of fucoxanthin to isofucoxanthin by alkali can be followed spectrophotometrically at the near-ultraviolet band. Alkali also facilitates oxidation of certain carotenoids; for example, astaxanthin is oxidized to astacene. Acids also cause chemical changes in carotenoids; certain acids also form Lewis acid–carotenoid complexes.

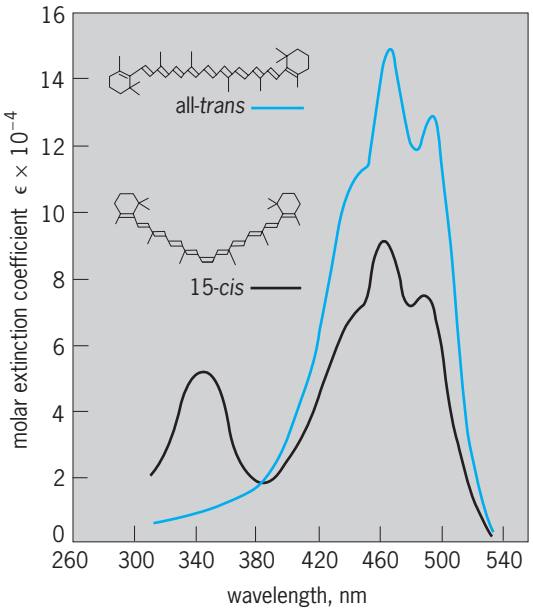

Fig. 3. Absorption spectra of all-trans carotene and 15,15′-mono-*cis*-β-carotene.

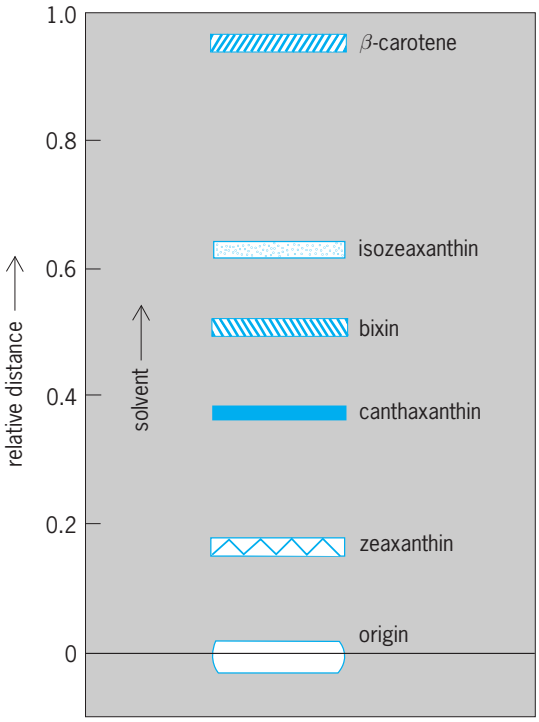

Fig. 2. Thin-layer chromatogram showing the separation of carotenoids. The adsorbent is silica gel, and the developing solvent consists of three parts *n*-hexane and seven parts ether.

Biosynthesis. The carbon skeletons of carotenoids are derived from the starting material, acetic acid, via a biosynthetic route similar to those of rubber and steroid syntheses. The biosynthetic pathway involves enzymatic reactions, such as hydrogenation, condensation, and dehydrogenation, followed finally by conjugation of the carbon-carbon double bonds forming the carotenoid π-electron system. Some xanthophylls are derived from β-carotene. **Figure 4** shows a simplified scheme for the biosynthesis of carotenoids.

The biosynthesis of carotenoids can be regulated by a number of exogenously added compounds. Some amines and compounds containing ammonium groups inhibit carotenoid biosynthesis in fungi and higher plants. On the other hand, many others stimulate carotenoid biosynthesis. A classic example of a compound that stimulates carotenoid biosynthesis is β-ionone, which is the ring structure occurring in β-carotene (Fig. 1). It is known that β-ionone

itself is not incorporated into the carotenoid structure during the biosynthesis of β-ionone-containing carotenoids. It is known too that the regulation of the carotenoid biosynthesis is rather specific with respect to different geometric carotenoid isomers; thus, compounds such as diethyl alkylamines stimulate the formation of all-trans carotenes. Specifically, diethyl octylamine and diethyl nonylamine greatly enhance the biosynthesis of the cyclic carotenes, α-, β-, and δ-carotenes, along with the noncyclic pigment lycopene (Fig. 1).

The substituted benzylamines, such as *N*-benzylphenylamines and *N*-benzyl 2-naphthalene methylamine, induce the accumulation of cis-carotenoids in certain plant fruits, for example, grapefruit.

There are two major regulatory modes of action for these bioregulators in trans-carotenoid biosynthesis. One mode of action includes the inhibition of the enzyme cyclase in the formation of the monocyclic δ-carotene and the bicyclic β-carotene (Fig. 4), as shown in the scheme where X indicates a blocked pathway. The other includes activation or transcriptional induction of a gene regulating the synthesis of specific enzymes required for the biosynthetic pathway of carotenoid pigments.

The mode of action of the bioregulators for the biosynthesis of cis-carotenoids is less clear. The cyclases are not inhibited by the cis-regulators. Thus, the most likely mode of action of these regulators appears to be gene activation that results in biosynthesis of cis-carotenoids at the expense of all-trans carotenoid biosynthesis.

β-Carotene serves as a major precursor to various oxygen-containing carotenoids called xanthophylls. For example, in lobster and crab the biosynthetic pathway from β-carotene to astaxanthin (Fig. 1 and 4) involves several intermediates, as outlined in **Fig. 5**. Zeaxanthin has also been proposed as a precursor to astaxanthin in lobster.

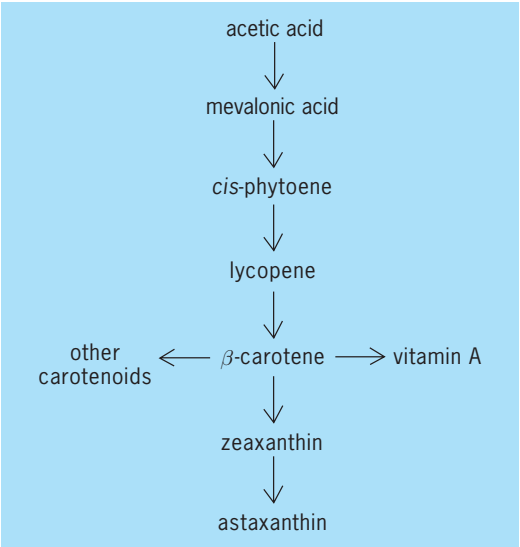

Fig. 4. Biosynthesis of carotenoids.

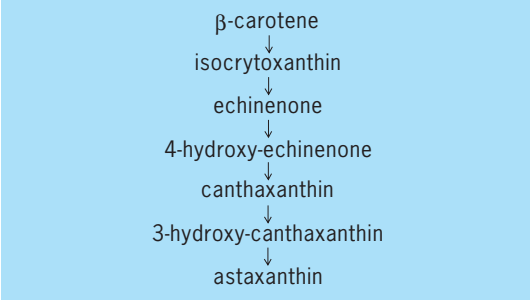

Fig. 5. Biosynthesis of astaxanthin from β-carotene in lobster and crab.

Biosynthesis of carotenoids can be enhanced by light. For example, blue light stimulates the formation of carotenoids in the fungi *Fusarium*, *Phycomyces*, and *Neurospora*. *Fusarium* synthesizes very little quantity of carotenoids in the dark, but blue light markedly stimulates the biosynthesis of carotenoids, including β-carotene and neurosporaxanthin. Blue light does so by regulating the expression of genes for the enzymes or regulatory proteins involved in the biosynthesis of carotenoids, and by photoinducing the production of certain factors used for the biosynthetic pathway of carotenoids. In addition, blue light stimulates the biosynthesis of carotenoids in algae, bacteria, and possibly in certain crustaceans and higher plants, such as angiosperms.

Red light also enhances the biosynthesis of carotenoids in higher plants; the pigment protein phytochrome acts as the red-light receptor. The red-light effect may be coupled to the blue-light effect in enhancing the biosynthetic formation of carotenoids in higher plants.

Function. The biochemical function of carotenoids is still largely unknown, although there are several biochemical functions in which the role of carotenoids is well understood. These include carotenoids in the photosynthetic apparatus of green plants, algae, and photosynthetic bacteria, where carotenoids function as a blue-light–harvesting pigment (antenna or accessory pigment) for photosynthesis. Thus carotenoids make it possible for photosynthetic organisms more fully to utilize the solar energy in the visible spectral region. Although the mechanism for the transfer of absorbed solar energy from carotenoid to chlorophyll is yet to be elucidated, it appears that carotenoid structures involving interactions of carotenoid-carotenoid as well as carotenoid-chlorophyll aggregate pairs are essential for the efficient energy transfer. This feature has been elucidated for the blue-light–harvesting carotenoid-chlorophyll protein complexes of marine dinoflagellates. **Figure 6** shows two sets of peridinin pairs (dimer) which stabilize the excited state of peridinin produced by light absorption, thus enhancing the probability of energy transfer from the carotenoid to chlorophyll. Without such interactions through exciton coupling between the two excited states of the peridinin dimer, each individual peridinin

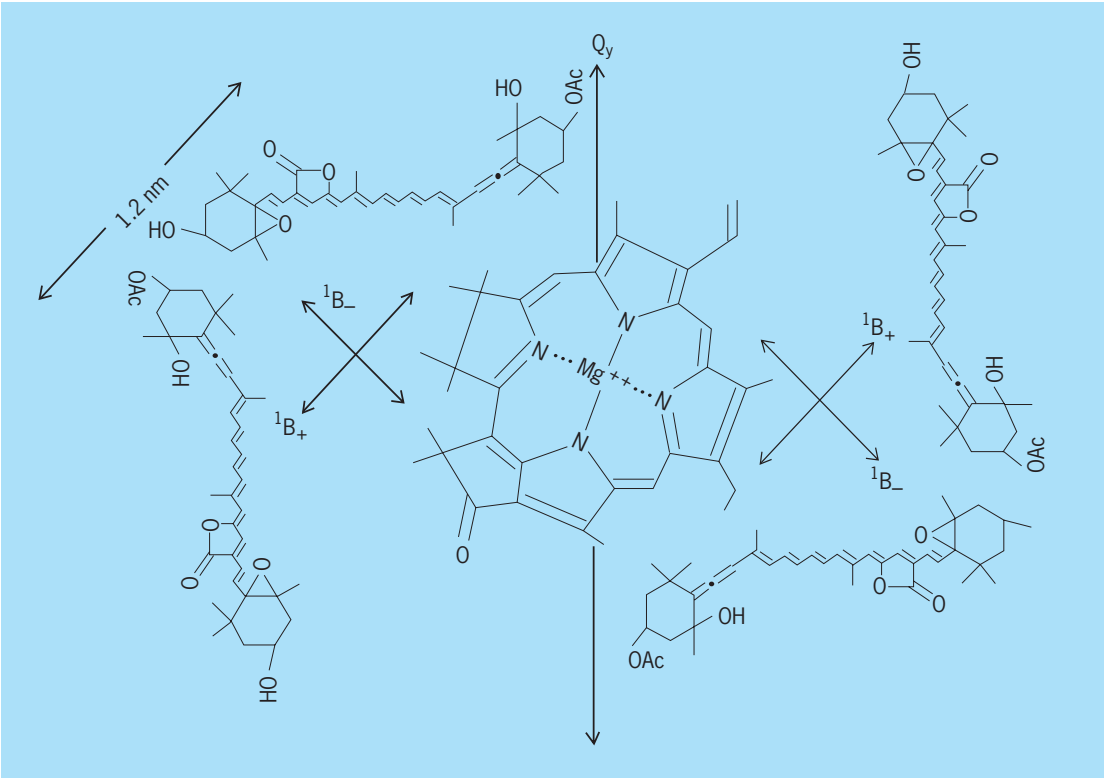

Fig. 6. A possible arrangement of four peridinin molecules (in two sets of dimer) around the chlorophyll *a* molecule. The Q_y axis of chlorophyll *a* represents the direction of fluorescence emission. B_+ and B_- represent two exciton states resulting from dimeric interactions between the peridinin molecules. Energy transfer takes place from the B_+ state to the chlorophyll molecule.

molecule does not possess a long enough lifetime for the energy transfer to chlorophyll to take place. *See* CHLOROPHYLL; PHOTOSYNTHESIS.

In addition to the light-harvesting function of carotenoids for photosynthesis, it is possible that carotenoid acts as an electron donor to the reactive chlorophyll (so-called P680), resulting in the formation of a carotenoid cation radical which may participate as an intermediate in the oxygen-evolving reaction of photosynthesis. However, this possibility remains to be studied.

Many algae, including dinoflagellates, utilize similar carotenoprotein pigments as the photoreceptor for phototaxis (swimming toward or away from blue light) and photophobic response (light-avoiding behavior). Apparently carotenoproteins in these systems serve as the antenna pigment and transfer the absorbed energy to the primary photoreceptor, probably chlorophyll, of the photosynthetic machinery which may control flagellum motion of algae in response to blue light by providing the necessary source of energy for the transduction network.

Another function of carotenoids is to protect biological systems such as the photosynthetic apparatus from photodynamic damage. Thus, plants lacking carotenoids are readily damaged by light. Certain herbicides kill weeds by blocking the biosynthesis of carotenoids so that the weed's photosynthetic apparatus is unprotected from photodynamic damage. Protection from photodynamic damage is achieved by quenching the powerful photodynamic oxidizing agent, singlet oxygen, produced as an undesirable by-product of the exposure of pigmented organisms to light. Carotenoids also protect plants from photodynamic damage by quenching the so-called triplet-state chlorophyll (where two of the π electrons in chlorophyll have their magnetic moments in parallel) in the photosynthetic apparatus.

Carotenoids are synthesized concomitantly with other photosynthetic membrane constituents for photosynthesis in the alga *Euglena*. This simultaneous synthesis appears to ensure that the membrane being assembled for the photosynthetic apparatus is stabilized by carotenoids and protected from photodynamic damage. However, this is not the case in higher plants where assembling of the photosynthetic apparatus takes place before the biosynthesis of carotenoids.

There may be many other functions of carotenoids as yet unknown. For example, it is possible that carotenoids, as important flower-petal pigments, may function in attracting insect pollinators.

Usage. Certain carotenoids are metabolized in biological systems, and some of them possess nutritional value. In particular, β-carotene is widely used as the vitamin A precursor, since it is metabolically oxidized to retinol (vitamin A) in animals, except cats. It is also used as an effective sunscreen agent in erythropoietic protoporphyria. Carotenoids, as well as a derivative, vitamin A, exhibit antitumor activity, for

example, against ultraviolet-radiation–induced skin tumors. *See* VITAMIN A.

Perhaps the most important industrial application of carotenoids, including β-carotene and lycopene, which are rich in carrot and tomato, respectively, is in safe coloration of foods, as exemplified in the coloring and fortification of margarine and poultry feedstuff. *See* FOOD MANUFACTURING. Pill-Soon Song

Bibliography. G. Britton and T. W. Goodwin (eds.), *Carotenoid Chemistry and Biochemistry*, 1982; O. Isler (ed.), *Carotenoids*, 1971; H. Senger (ed.), *Blue Light Responses: Phenomena and Occurrence in Plants and Microorganisms*, 1987; P. S. Song, *Trends Biochem. Sci.*, pp. 25–27, February 1978.

Carotid body

A special sensory organ (glomus caroticum) located in the angle of the branching of the common carotid artery into the external and internal carotid arteries (see **illus.**). The human carotid body is an ovoid organ approximately 0.2 in. (5 mm) long, embedded in the outer connective tissue (adventitia) of the blood vessel. The carotid body is called a chemoreceptor because it is sensitive to oxygen and carbon dioxide in the blood. There is evidence that suggests that the carotid body is also sensitive to changes in blood pressure, blood flow, and blood osmolarity (salt content). *See* CHEMORECEPTION.

Histology. The carotid bodies are highly vascular structures composed of cords or lobules of epithelioid cells. The glomus cells, which constitute the principal cell type in the carotid body, contain numerous cytoplasmic vesicles or granules that are known to contain special transmitter substances called catecholamines. Stored together with these catecholamines are putative transmitter substances called neuropeptides that may modulate the action of the catecholamine neurotransmitter. The carotid body also contains supporting cells, which are less abundant than glomus cells, do not contain vesicles, and partially envelop glomus cells and adjacent nerve fibers in their attenuated processes.

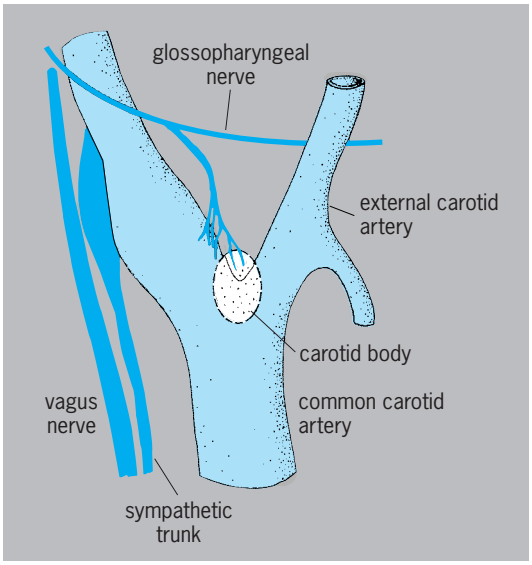

Location of the carotid body in relation to the carotid arteries.

The glomus cells are richly innervated by sensory nerve fibers derived from the ninth cranial (glossopharyngeal) nerve, and to a lesser extent by the ganglioglomerular nerve from the superior cervical ganglion. The nerve fibers enter the carotid body and travel between and around the lobules of glomus cells, forming an interstitial plexus. Small unmyelinated nerve fibers from this plexus ultimately end on the glomus cells in a variety of synaptic configurations.

The embryology of the carotid body is controversial; experiments in birds, however, suggest that all of the cells are derived from a specialized portion of the neuroectoderm (neural crest), and there is little reason to suspect a different origin for the mammalian carotid body. The neural crest also gives rise to the adrenal medulla, whose cells closely resemble those of the carotid body. The adrenal medullary cells, however, secrete the hormones epinephrine (adrenaline) and norepinephrine (noradrenaline), whereas glomus cells release primarily dopamine and norepinephrine as neurotransmitters. *See* ADRENAL GLAND; EPINEPHRINE.

Physiology. The 1930s marked an epoch in carotid body research. It was demonstrated at that time that the carotid body and similar glomus tissue accumulations associated with the aortic arch (aortic bodies) possessed chemoreceptor properties. It was found that blood low in oxygen (hypoxemia) or high in carbon dioxide (hypercapnia) stimulated an increase in respiration and a secondary increase in heart rate and blood pressure. The increase in respiration observed after carotid body stimulation provides for an increase in the oxygenation of the red blood cells (erythrocytes). Likewise, secondary reflexes originating from the lungs cause an increase in heart rate that results in a greater blood flow from the heart to the lungs and, ultimately, the whole body. These important reflexes protect the organism from conditions of hypoxia (low tissue oxygen) such as may be experienced at high altitudes, during exercise, or during various cardiopulmonary disease states. *See* AORTIC BODY.

Although much is known about reflexes initiated by the carotid body, it is still not known how the chemical information derived from the blood is converted into electrical signals that are then transmitted to the cardiopulmonary brain centers. Current hypotheses center on the glomus cell as the immediate chemosensor which responds to stimuli by triggering nerve impulses in the sensory nerve endings. Alternatively, some studies suggest that both glomus cells and sensory nerve endings have chemosensory properties and that these elements coparticipate in the chemoresponse. Experiments to differentiate between these two possibilities are difficult to design. Moreover, different cellular elements within the carotid body may respond to different stimuli, thereby complicating the transduction

mechanisms that normally are operative. Regardless of how chemoreception occurs, the carotid body is uniquely suited for the task of "sampling" the arterial blood. Because of its high vascularity and small size, the carotid body has one of the highest blood flow rates of any tissue in the body.

Pathological and physiological changes. Tumors of the carotid body are termed chemodectomas and are similar histologically to tumors of the adrenal medulla (pheochromocytomas) that secrete catecholamines. Carotid body tumors usually are benign, asymptomatic, and slow-growing. Individuals living at high altitudes have larger carotid bodies, and a higher incidence of carotid body tumors, although compared to all neoplasms the incidence is low. Finally, there is some evidence that the carotid body may be involved in sudden infant death syndrome (SIDS) or "crib death." It is still not known, however, whether SIDS arises from a peripheral (such as the carotid body) or from a central (such as brainstem) defect in respiratory integration. *See* CARDIOVASCULAR SYSTEM; SUDDEN INFANT DEATH SYNDROME (SIDS). John T. Hansen

Bibliography. *Handbook of Physiology*, sec. 3: *The Respiratory System*, vol. 3: *Control of Breathing*, ed. by A. P. Fishman, 1986.

Carp

The common name for a number of cypriniform fishes of the family Cyprinidae. The carp originated in China, where for centuries it was raised for food. It was imported into the United States from Europe, where it also has been raised for years as a source of food.

Two varieties or forms of carp occur in nature (see **illus.**), but are commoner in captivity in breeding ponds. These are the mirror carp, which has a few large scales, and the leather carp, which has no scales. In America the carp has invaded many natural bodies of water, where it has destroyed the weeds which are part of the natural breeding places of native fishes.

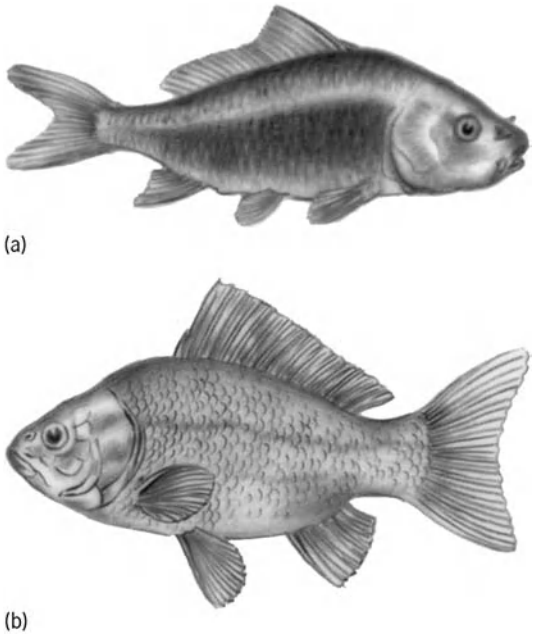

Two varieties of carp: (*a*) the carp (*Cyprinus carpio*) and (*b*) the goldfish (*Crassius auratus*).

The carp (*Cyprinus carpio*) is a member of the Cyprinidae and is closely related to the goldfish (*Crassius auratus*). The silver carp (*Hypopthalmichthys*) is an Asiatic species in which the fourth gill arch is modified to form the helical organ, which is an accessory feeding structure. Fine, particulate food that escapes the gill rakers is entrapped in the mucus of the helical organ and passed into the stomach. The carp has pharyngeal teeth and a suckerlike mouth. A very long dorsal fin is preceded by a strong spine. The swim bladder of the carp is associated with a group of small bones at the anterior end, the Weberian ossicles. This structural modification enables the carp to perceive sound quite well. *See* ACTINOPTERYGII; CYPRINIFORMES; SWIM BLADDER. Charles B. Curtin

Carpal tunnel syndrome

A condition caused by the thickening of ligaments and tendon sheaths at the wrist, with consequent compression of the median nerve at the palm. Affected individuals report numbness, tingling, and pain in the hand; the discomfort often becomes worse at night or after use of the hand. A physical examination of the injured hand during the early stages of the syndrome often reveals no abnormality. With more severe nerve compression, the individual experiences sensory loss over some or all of the digits innervated by the median nerve (thumb, index finger, middle finger, and ring finger) and weakness of thumb movement.

Assessment. Electrodiagnostic testing is important for an accurate diagnosis. The electromyographer uses sensory fibers to measure the nerve conduction velocity from the finger or the palm to the wrist and the motor conduction velocity from the wrist to the thumb muscles. Approximately half the individuals with carpal tunnel syndrome have abnormalities of the opposite median nerve. Electrodiagnostic values of these individuals therefore need to be compared with reference values obtained from normal subjects and with ulnar or radial nerve values. No other test has a higher diagnostic accuracy for individuals who have a final diagnosis of carpal tunnel syndrome, but to make a final diagnosis clinical data, including the individual's response to treatment, must be obtained.

False positive and false negative results are illustrated by the high rate of abnormalities in the opposite, nonsymptomatic hands of individuals with carpal tunnel syndrome. As testing becomes more complex and sophisticated, it becomes increasingly likely that results beyond the normal range (false positive) will be generated. The number of individuals with false negative results should be below 8%.

Thermography shows clear abnormalities in carpal tunnel syndrome, but an abnormal thermogram is found in many other conditions in which the pattern of blood flow to the hand is altered. Computed tomography and magnetic resonance imaging have not been widely used for carpal tunnel syndrome and have no role in management at this time.

Prevalence. Carpal tunnel syndrome is common. One study reported that 125 individuals per 100,000 were affected from 1976 to 1980. Another study estimated that 515 of every 100,000 patients sought medical attention for carpal tunnel syndrome in 1988; the syndrome in half of these patients was thought to be occupational in origin.

The incidence of carpal tunnel syndrome is greater among electronic-parts assemblers, frozen-food processors, musicians, and dental hygienists. Highly repetitive wrist movements, use of vibrating tools, awkward wrist positions, and movements involving great force seem to be correlated with the disorder. Awkward and repetitive wrist motions occur in many office tasks, such as typing and word processing. In some occupations, such as shellfish packing, the incidence is more than 200 times higher than in the baseline data. The highest reported incidence in an industrial setting, based on the numbers of carpal tunnel syndrome–based releases, was 15% among a group of meatpackers.

Carpal tunnel syndrome probably accounts for a minority of the cases of overuse syndrome (cumulative trauma syndrome), which is a common problem in occupational settings. Overuse syndrome symptoms include muscle pain, tendinitis, fibrositis (inflammation of connective tissue in a joint region), and epicondylitis (inflammation of the eminence on the condyle of a bone). Although the causative relationship between the two disorders has not been conclusively proven, the incidence of both carpal tunnel syndrome and overuse syndrome appears to increase in tandem in individuals who are at risk.

Pathophysiology. Under normal circumstances, the pressure within the tissues of a limb is 7–8 mmHg. In carpal tunnel syndrome, the pressure is often 30 mmHg, approaching the level at which nerve dysfunction occurs. With wrist flexion or extension, pressures may increase to 90 mmHg or more.

The increase in pressure within the carpal canal is usually caused by nonspecific inflammation of flexor tendon sheaths. Diabetes, pregnancy, rheumatoid arthritis, and hypothyroidism are the most common medical conditions associated with carpal tunnel syndrome. Amyloidosis, acromegaly, and mycobacterial infections are rare causes. A reduction in the flow of blood to the nerve can account for the intermittent tingling that occurs at night or with wrist flexion. *See* AMYLOIDOSIS; ARTHRITIS; DIABETES; THYROID GLAND DISORDERS.

Treatment. Nonsurgical treatment includes avoidance of the use of the wrist, use of a splint to keep the wrist in a neutral position, and anti-inflammatory medications. These treatments are especially useful in individuals with an acute flare-up and in those with minimal and intermittent symptoms. Conservative nonsurgical treatment will not succeed for individuals over the age of 50, for those who have had the syndrome for more than 10 months, and for those with constant tingling.

Surgical treatment may be used if conservative approaches fail. The procedure is usually done on an outpatient basis with prognoses of good to excellent in 80% of the cases. Although 40% of the individuals regain normal function, the condition of 5% may worsen. Most individuals return to an office job within a week of surgery, but it may be 4–6 months before carpenters, construction workers, or athletes are able to return to work. Many individuals with work-related carpal tunnel syndrome should consider changing jobs.

Surprisingly little information is available regarding the redesign of workstations and its effect on the reversal of symptoms or the prevention of carpal tunnel syndrome. Ergonomic redesign is widely practiced, but rarely described in medical terms. David M. Dawson

Bibliography. D. M. Dawson, M. Hallett, and L. H. Millender, *Entrapment Neuropathies*, 3d ed., 1999; R. B. Rosenbaum and J. L. Ochoa, *Carpal Tunnel Syndrome and Other Disorders of the Median Nerve*, 1993.

Carpoids

The common name for four extinct classes of primitive echinoderms that have a flattened theca or body lacking radial symmetry. These enigmatic fossils were originally classified together in the class Carpoidea, but more recent echinoderm researchers have assigned them to four separate classes in the subphylum Homalozoa: the Stylophora (or Calcichordates), Homoiostelea, Homostelea, and Ctenocystoidea. These four classes include about 50 genera that range from the Early or Middle Cambrian to the Late Carboniferous.

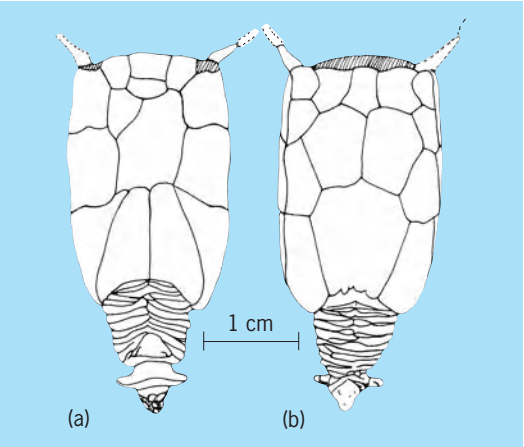

***Enoploura popei*, a stylophoran carpoid from the Late Ordovician of Ohio. (*a*) Concave lower side and (*b*) convex upper side of the plated theca showing part of the attached, tapering appendage used for locomotion and feeding at the bottom.**

Carpoids have a flattened theca that varies from asymmetrical to nearly bilaterally symmetrical (see **illus.**) and is made up of sutured, multiporous, single-crystal, calcite plates like those found in other echinoderms. Three of the classes have a long plated appendage attached to the theca that was used for locomotion and, in the Stylophora, also for feeding. The mouth was either on the anterior margin of the theca with ambulacral grooves leading to it or at the base of the feeding appendage; the anal pyramid was usually on the opposite thecal margin or nearby at a thecal corner. Carpoids were apparently mobile, bottom-living or shallow-burrowing, detritus or suspension feeders, sifting out small food particles from the top layer of soft sediment or from the surrounding seawater. Because of the distinctive skeletons, most researchers consider carpoids as true echinoderms, although they seem only distantly related to other fossil and living echinoderms that have well-developed pentameral symmetry. *See* ECHINODERMATA; HOMALOZOA. James Sprinkle

Bibliography. R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, pt. S, 1968.

Carrageenan

A polysaccharide that is a major constituent of the cell walls of certain red algae (Rhodophyceae), especially members of the families Gigartinaceae, Hypneaceae, Phyllophoraceae, and Solieriaceae. Extracted for its suspending, emulsifying, stabilizing, and gelling properties, it is one of three algal polysaccharides of major economic importance, the others being agar and alginate. *See* AGAR; ALGINATE.

Carrageenans form a family of linear (unbranched) sulfated polysaccharides in which the basic repeating units are exclusively D-galactose, or D-galactose and 3,6-anhydro-D-galactose, linked alternately α-1,3- and β-1,4-. Different types of carrageenans, which are designated by Greek letters, are found in different species and sometimes in different somatic phases of the same species or in different regions of the same plant. The κ- and ι-carrageenans are soluble in hot water but not cold, and gel upon cooling in the presence of potassium ion. The λ-, ϵ-, and π-carrageenans are soluble in hot water or cold, forming a viscous solution but not a gel.

Carrageenan is prepared by boiling the dried algae in slightly alkaline water. The hot solution is filtered and poured into isopropyl alcohol, which causes the carrageenan to precipitate. The product is purified, dried, ground, and packaged as a powder.

The main sources of carrageenan are *Chondrus crispus* (Gigartinaceae) from the Maritime Provinces of Canada and various species of *Eucheuma* (Solieriaceae) from the Philippines. *Chondrus*, popularly called Irish moss, is harvested from naturally occurring intertidal stands, while *Eucheuma* is successfully grown in mariculture on nets and lines.

About 80% of the refined carrageenan is used in food processing, with the dairy industry being the chief consumer. The remainder is used in the cosmetic, pharmaceutical, printing, and textile industries. *See* RHODOPHYCEAE.

Paul C. Silva; Richard L. Moe

Carrier (communications)

A periodic waveform upon which an information-bearing signal is impressed. This process is known as modulation and comprises a variety of forms such as amplitude, phase, and frequency modulation. The most common type of carrier is the sinusoidal carrier (**illus.** *a*), but in reality, any periodic waveform followed by a band-pass filter can serve as a carrier.

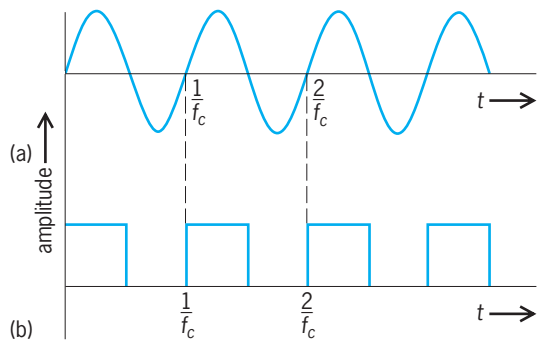

Carriers. (*a*) Sinusoidal carrier. (*b*) Square-wave carrier.

The validity of the last statement can be verified by considering a periodic function $c(t)$, and assuming that it is desired to have a message $m(t)$ double-sideband (DSB)-modulate a carrier. Since the function $c(t)$ is periodic, it can be expanded in a Fourier series of the form given by Eq. (1), where the $\{c_n\}$

$$c(t) = \sum_{n=-\infty}^{\infty} c_n e^{j2\pi n f_c t} \tag{1}$$

are the coefficients of the Fourier series, and f_c is the carrier frequency, chosen so that it exceeds W, the bandwidth of the message $m(t)$. This results in the waveform given by Eq. (2). The Fourier transform

$$m(t)c(t) = \sum_{n=-\infty}^{\infty} c_n m(t) e^{j2\pi n f_c t} \tag{2}$$

of this waveform is given by Eq. (3), where $M(\omega)$

$$F(\omega) = \sum_{n=-\infty}^{\infty} c_n M(\omega - n\omega_c) \tag{3}$$

is the Fourier transform of $m(t)$, and $\omega_c = 2\pi f_c$. If this signal is now put through a band-pass filter with center frequency f_c and bandwidth $2W$, a waveform proportional to $m(t) \cos \omega_c t$ results (that is, a double-sideband signal has been generated). *See* FOURIER SERIES AND TRANSFORMS.

A periodic waveform that is very simple to generate is a square wave composed of ones and zeros of equal duration. This allows the multiplication function inherent in the modulation process represented by the function $m(t)c(t)$ to be replaced by an on-off switch. Indeed, such a modulator is commonly

known as a switching modulator, and results in a periodic square wave (illus. *b*).

At times, it is convenient to use subcarriers in addition to the main carrier. Such a situation arises when frequency-division multiplexing (FDM) is used. A set of, say, *N* information-bearing waveforms individually modulate their own subcarrier, whereby the spacing of the subcarriers is chosen so that it results in (virtually) no spectral overlap of the resulting *N* spectra. This composite baseband waveform then modulates the main carrier in preparation for transmission over the channel. For voice communications, the modulation on the subcarriers is often single-sideband (SSB), whereas typically the modulation on the carrier itself is frequency-modulation. *See* AMPLITUDE MODULATION; FREQUENCY MODULATION; MODULATION; MULTIPLEXING AND MULTIPLE ACCESS; SINGLE SIDEBAND. Laurence B. Milstein

Carrot

A biennial umbellifer (*Daucus carota*) of Asiatic and Mediterranean origin belonging to the plant order Umbellales. The carrot is grown for its edible roots which are eaten raw or cooked (**Fig. 1**). *See* APIALES.

Cultivation and harvesting. Propagation is by seed planted directly in the field. Spacing varies; plants are usually planted $^1/_2$–2 in. (1.2–5 cm) apart in rows of 18–24 in. (45–60 cm). Temperature and soil moisture affect root color and shape; 60–70°F (15–21°C) is optimum for high yields of long, deep-colored roots.

Exposure of the young carrot plants to prolonged cold weather (40–50°F, or 4–10°C, for 15 or more days) favors seed-stalk formation. Misshapen roots commonly result from nematodes, aster yellows disease, or compacted soils. Most commercial plantings are weeded chemically.

Varieties (cultivars) are classified according to length of root (long or short or stump-rooted) and use (fresh market or processing). Popular varieties for fresh market are Imperator and Gold Pak; for processing, Red Cored Chantenay and Royal Chantenay are often chosen.

Harvesting of carrots for fresh market begins when the roots are $^3/_4$–2 in. (1.8–5 cm) in diameter, usually 60 to 90 days after planting. Harvesting by machine is common. Marketing in bunches is declining; most carrots are sold in plastic bags with the foliage removed. Carrots for canning generally have shorter roots and are grown to a larger diameter.

Texas, California, and Arizona are three important carrot-producing states. *See* NEMATA. H. John Carew

Diseases. Carrots are affected in the field by two important leaf spot diseases and by aster yellows. In storage and transit, the roots are attacked by several kinds of rot-inducing organisms.

Carrot leaf spots are caused by the fungi *Cercospora carotae* and *Alternaria dauci*. Cercospora blight occurs on young leaves and Alternaria blight on older leaves. These fungi overwinter in debris of the previous crop. During the growing season, repeated disease cycles are caused by airborne spores.

Fig. 1. Carrots. (*Asgrow Co., Subsidiary of the Upjohn Co.*)

In humid weather, both diseases increase rapidly, and severe defoliation results. Control of these leaf spots is achieved by repeated application of a suitable fungicide.

Aster yellows is caused by a mycoplasmalike organism that can infect many weed and crop hosts. Leafhoppers, principally *Macrosteles fascifrons*, are the insect vectors. Aster yellows causes the development of many adventitious chlorotic shoots to give a witches'-broom condition. These malformed and bushy tops eventually turn bronze or reddish. Roots are also malformed and produce numerous woolly secondary roots (**Fig. 2**) and tough woody tissue in the carrot. Control of aster yellows is by control of

Fig. 2. Carrot yellows. (*Courtesy of Arden Sherf, College of Agriculture, Cornell University*)

the leafhopper vectors. Spraying with an approved insecticide should begin early in the season as vectors begin to migrate from nearby diseased weeds into carrot fields.

Rots of carrots in storage or transit are caused by several different bacteria or fungi. The classic rot is bacterial soft rot caused by *Erwinia carotovora*. This bacterium is a wound parasite and enters the carrot through openings at the crown, insect wounds, or harvest bruises. At temperatures about 40°F (4°C), the rot advances rapidly. Diseased tissues become soft and slimy and have a putrid odor. Control of soft rot requires careful handling during harvest and packing in order to reduce bruising. Storage temperature should be just above freezing, and relative humidity below 90%. *See* PLANT PATHOLOGY.

Thomas H. Barksdale

Bibliography. L. V. Boone et al., *Producing Farm Crops*, 4th ed., 1991; T. W. Whitaker et al., *Carrot Production in the United States*, U.S. Department of Agriculture, Agr. Handb. 375, 1970.

Cartilage

A firm, resilient connective tissue of vertebrates and some invertebrates. Isolated pieces act as a skeleton to provide support and anchor muscles, or cartilage is with bone to contribute its resilience and interstitial growth to bony skeletal functions. Cartilage comprises a firm extracellular matrix synthesized by large, ovoid cells (chondrocytes) located in holes called lacunae. The matrix elements are water bound by the high negative charge of extended proteoglycan (protein-polysaccharide) molecules, and a network of fine collagen fibrils. The elements furnish mechanical stability, give, and tensile strength, but allow the diffusion of nutrients and waste to keep the cells alive. Generally, blood vessels reach only to the perichondrium of fibrous connective tissue, wrapping around the cartilage and attaching it to other tissues. *See* BONE; COLLAGEN.

Cartilage is modified in several ways. In elastic cartilage, the chondrocytes add elastic fibers to the matrix to increase resilience, as in cartilages supporting the Eustachian tube, mammalian external ear, and parts of the larynx. Where cartilage joins bones tightly at certain joints with limited mobility, for example, at the pubic symphysis and between vertebrae, the matrix of fibrocartilage contains prominent collagen fibers and has less proteoglycan than the typical hyaline variety. Hyaline cartilage, named for its glassy translucence, is the major support in the airway; and throughout the embryo, pieces of it develop as a cartilaginous precursor to the bony skeleton, except in the face and upper skull.

The primitive cartilaginous skeleton undergoes another modification, by locally calcifying its matrix. At sites of calcification, invading cells destroy the cartilage and mostly replace it by bone, leaving permanent hyaline cartilage only at the joint or articular surfaces, in some ribs, and, until maturity, at growth plates set back from the joints and perpendicular to the long axis of limb bones. The precarious physiological balance between chondrocytes and matrix materials in the heavily loaded articular cartilage breaks down in old age or in inflamed joints. *See* ARTHRITIS; CONNECTIVE TISSUE; JOINT (ANATOMY); SKELETAL SYSTEM.

William A. Beresford

Bibliography. B. K. Hall (ed.), *Cartilage*, 3 vols., 1983.

Cartography

The techniques concerned with constructing maps from geographic information. Maps are spatial representations of the environment. Typically, maps take graphic form, appearing on computer screens or printed on paper, but they may also take tactile or auditory forms for the visually impaired. Other representations such as digital files of locational coordinates or even mental images of the environment are also sometimes considered to be maps, or virtual maps.

Maps and uses. Because the environment is complex and everchanging, the variety of maps and map uses is unlimited. For instance, maps are indispensable tools for navigating over land, sea, or air. Maps are effective both in exploiting natural resources and in protecting them. They are used to investigate geographic phenomena, including environmental pollution, climate change, even the spread of diseases and the distribution of social phenomena such as poverty and illiteracy. In addition, maps can be used to communicate insights derived from geographic research through publication in periodicals and books and through distribution over computer networks and broadcast media. Every private business, government agency, and academic discipline whose products, services, or objects of study are geographically dispersed benefits from detailed, up-to-date maps. Unfortunately, appropriate maps often are unavailable.

Map scale and geographic detail. Maps often include insufficient or excessive detail for the task at hand. The amount of usable detail on a map varies with its scale, because human visual acuity and the resolution of printing and imaging devices are limited. Maps that depict extensive areas in relatively small spaces are called small-scale maps. For example, on a 1-ft-wide (30-cm) map of the world, on which the ratio of map distance to ground distance is approximately 1:125,000,000, very little perceptible detail can be preserved. As the scale of a map increases, so may the level of geographic detail it represents. Geographic features selected to appear on small-scale maps must be exaggerated in size and simplified in shape so as to be recognizable by the map user. These map generalization operations constitute an intriguing field of research by cartographers attempting to formalize, and ultimately to automate, the map creation process.

Reference maps. Topographic maps record the positions and elevations of physical characteristics of the landscape. They serve as locational dictionaries

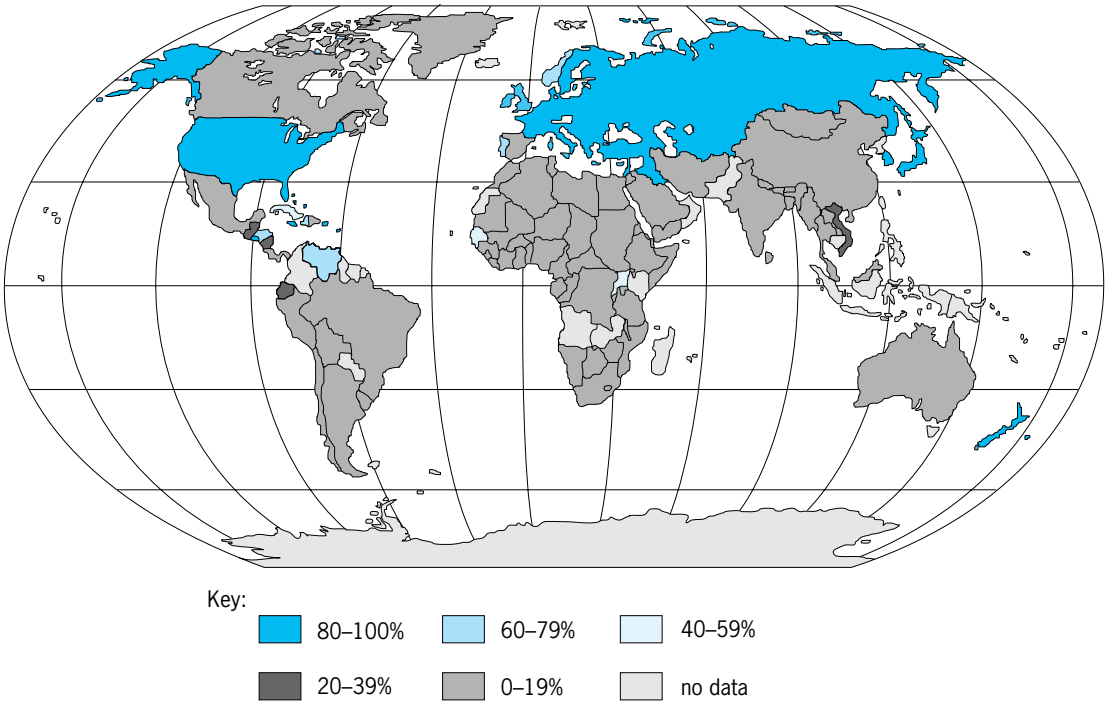

Fig. 1. Large-scale topographic mapping, percent area coverage area at 1:1000 to 1:36,680 showing the uneven spatial distribution of map coverage. (*After United Nations, World Cartography, vol. 20, 1990*)

for many endeavors, including environmental planning, resource management, and recreation. The 1:24,000 scale United States Geological Survey (USGS) series covers the continental United States with approximately 57,000 map sheets and depicts 17 categories of physiographic and cultural features with more than 130 distinct graphic symbols. The enormous costs involved in compiling, producing, and revising a topographic map series account for the fact that only about 15% of Earth's surface is topographically mapped at a scale of 1:25,000 or larger (**Fig. 1**). *See* TOPOGRAPHIC SURVEYING AND MAPPING.

Thematic maps. Another problem with available maps is that they often fail to include a feature of particular interest. Maps that emphasize one or a few related geographic phenomena in the service of a specific purpose are called thematic maps. An example is a thematic map that reveals the uneven distribution of topographic map coverage around the world (Fig. 1). Thematic maps are powerful alternatives to text, tables, and graphs for visualizing potentially meaningful patterns in geographic information. Although the production of large-scale topographic map series requires the resources of large private or government agencies, individuals and small organizations with access to relatively inexpensive personal computers, mapping software, and databases can afford to produce thematic maps in support of business, scientific, political, and creative endeavors.

Constructing geographic information. Maps are composed of two kinds of geographic information: attribute data and locational data. Attribute data are quantitative or qualitative measures of characteristics of the landscape, such as terrain elevation, land use, or population density. Locations of features on the Earth's surface are specified by use of coordinate systems; among these, the most common is the geographical coordinate system of latitudes and longitudes.

Geographical coordinates describe positions on the spherical Earth. These must be transformed to positions on a two-dimensional plane before they can be depicted on a printed sheet or a computer screen. Hundreds of map projections—mathematical transformations between spherical and planar coordinates—have been devised, but no map projection can represent the spherical Earth in two dimensions without distorting spatial relationships among features on Earth's surface in some way. One specialized body of knowledge that cartographers bring to science is the ability to specify map projections that preserve the subset of geometric characteristics that are most important for particular mapping applications.

Prior to World War II, locational data were compiled mainly by field surveys. Aerial surveillance techniques developed for the war effort were then adapted for use in civilian mapmaking. The scale distortions inherent in aerial photographs can be corrected by photogrammetric methods, yielding planimetrically correct projections on which all locations appear to be viewed simultaneously from directly above. Rectified aerial photographs (orthophotos) can be used either as bases for topographic mapping or directly as base maps. *See* AERIAL PHOTOGRAPH; LATITUDE AND LONGITUDE; PHOTOGRAMMETRY.

Influence of computing technology. Periodically, cartographic practice has been transformed by new technologies. Few have had such a profound effect as the development of computer-based mapping techniques. While printed paper maps still constitute the richest store of geographic information, cartography has become as much a digital as a paper-based enterprise. With more and more geographic data available in digital form, the computer has changed the very idea of a map from a static caricature of the environment to a dynamic interface for generating and testing hypotheses about complex environmental and social processes.

Digital geographic data. There are two major approaches to encoding geographic data for computer processing. One, commonly called raster encoding, involves sampling attribute values at some regular interval across the landscape. Imagery scanned from Earth-observing satellites works this way, recording surface reflectance values for grid cells (pixels) from 80 to 30 m (250 to 100 ft) or less in resolution. Digital elevation models are matrices of terrain elevations derived from satellite imagery or sampled from topographic maps (**Fig. 2**).

A second method, known as vector encoding, involves digitizing outlines of landscape features that are homogeneous with regard to some attribute, such as a river, a watershed, a road, or a state boundary. Vector encoding is more expensive than raster encoding, but it is more flexible for many applications. One U.S. government data-base, for example, includes street descriptions and address ranges for 345 metropolitan areas by which U.S. census statistics can be matched and precisely mapped. The U.S. government's Digital Chart of the World vector database includes 16 distinct feature layers, including coastlines, rivers, roads, political boundaries, and 1000-ft (300-m) terrain elevation contours for the entire world at 1:1,000,000 scale.

Although the raster and vector approaches for digital encoding predominate, these have been implemented in dozens of idiosyncratic data formats designed to suit individual mapping agencies and computer vendors. Incompatible formats have impeded data sharing and have resulted in expensive redundancies in database construction. A committee of United States academic and government cartographers specified a Spatial Data Transfer Standard to improve cooperation and data sharing among federal government agencies. The development of international data standards is needed. *See* COMPUTER GRAPHICS.

Constructing geographic understanding. Geographic illiteracy is thought to put afflicted societies at a disadvantage in an increasingly integrated international economy. Access to geographic information is a necessary but insufficient condition for constructing geographic understanding. Access to analytical expertise is required to learn from the available information.

Geographic information systems. Guiding much of the research and development efforts of academic cartographers is the concept of automated geography—an amalgam of computer databases and procedures by which analysts might model, simulate, and ideally predict the behavior of physical and social systems on the landscape. Concurrently, increasing social concern for environmental protection stimulates a market for computerized geographic information systems (GIS) that combine mapping capabilities with techniques in quantitative spatial analysis. Many of the analytical procedures in these

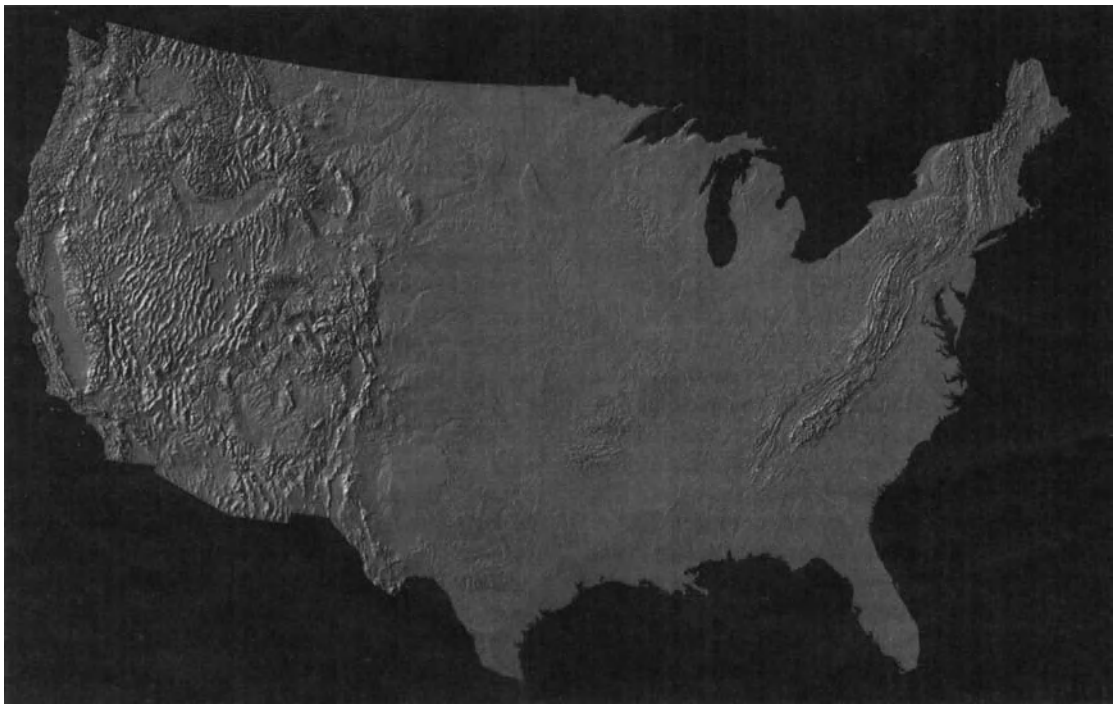

Fig. 2. Computer rendering of the topography of the 48 contiguous United States based on 12 million elevations. (USGS)

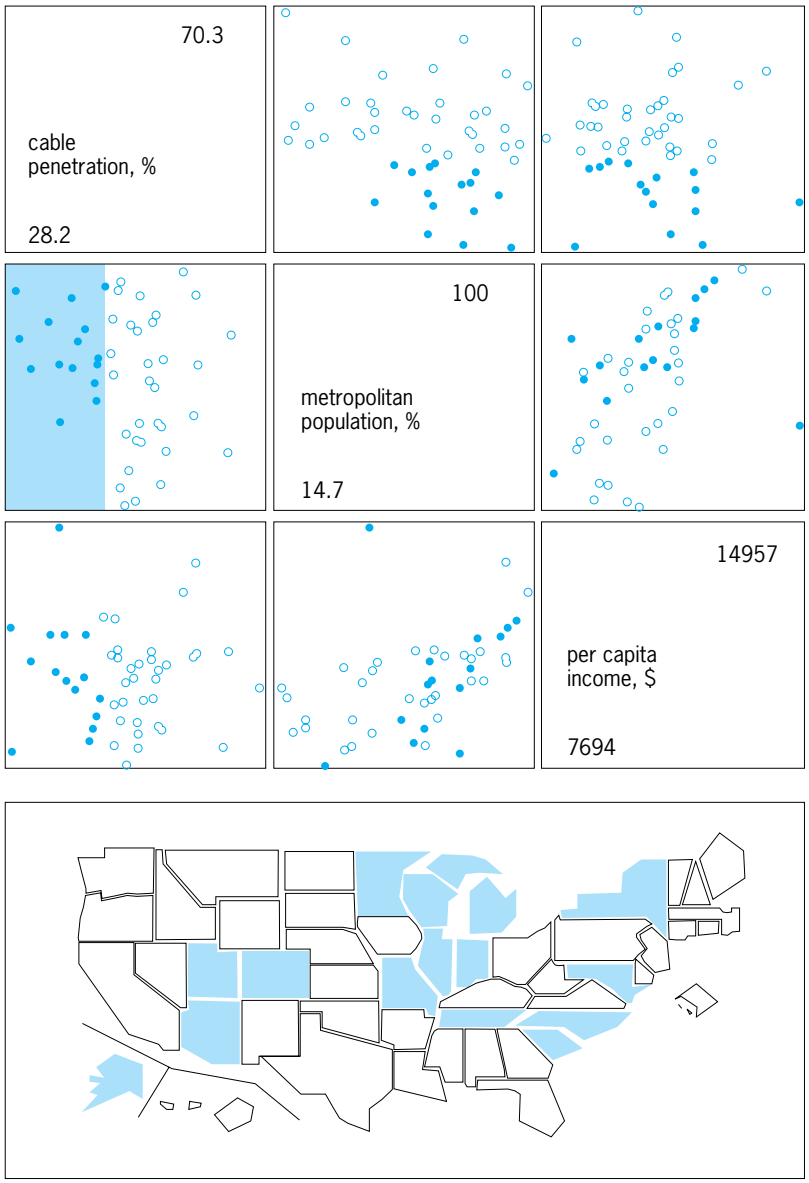

Fig. 3. Three geographic data variables related in a scatterplot matrix and linked to a map. Locations of observations selected in a scatterplot are automatically highlighted on the map. (*From M. Monmonier, Geographic brushing: Enhancing exploratory analysis of the scatterplot matrix, Geog. Anal., 21(1):81–84, 1989*)

systems—such as calculations of distances, areas, and volumes; of terrain surface slope and aspect; defining buffer regions surrounding landscape features; and generating maps of new features formed by the intersection of several related map layers—have been codified by cartographers. Some cartographers have investigated the potential of computerized expert systems that may be used to assist nonspecialists in performing quantitative geographic analyses. *See* EXPERT SYSTEMS; GEOGRAPHIC INFORMATION SYSTEMS.

Cartographic communication. With few exceptions, the outcomes of analyses performed with geographic information systems are maps. Just as careless or biased quantitative analyses result in erroneous conclusions, so can unskilled map designs mislead users, and even the analysts themselves. Concern for the integrity of thematic maps motivated cartography's largest research project—the search for objective guidelines, based on psychological research, to optimize map communication. Map design issues that have garnered the most attention include classification techniques for grouping attribute data into discernible map categories and logical systems for choosing appropriate graphic symbols for representing different types of attribute data.

Interactive cartography. Although many broadly applicable map design principles have been established, the goal of specifying an optimal map for a particular task is less compelling than it once was. Instead, there is interest in the potential of providing map users with multiple, modifiable representations via dynamic media such as CD-ROM, computer networks, and interactive television. Even as computer graphics technologies and numerical models provide ever more realistic environmental simulations, innovative, highly abstract display methods are being developed and tested to help analysts discover meaningful patterns in multivariate geographic data sets (**Fig. 3**). Maps, graphs, diagrams, movies, text, and sound can be incorporated in multimedia software applications that enable users to navigate through vast electronic archives of geographic information. Interactive computer graphics are eliminating the distinction between the mapmaker and the map user. Modern cartography's challenge is to provide access to geographic information and to cartographic expertise through well-designed user interfaces. *See* LAND-USE PLANNING; MAP DESIGN; MAP PROJECTIONS.

David DiBiase

Bibliography. G. L. Gaile and C. J. Willmott, *Geography in America at the Dawn of the 21st Century*, 2004; J. Makower (ed.), *The Map Catalog: Every Kind of Map and Chart on Earth and Even Some Above It*, 3d ed., 1992; P. C. Muehrcke, *Map Use: Reading, Analysis and Interpretation*, 4th ed., 2001; D. Wood, *The Power of Maps*, 1992.

Caryophyllales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the core eudicots. The order consists of 26 families and about 12,500 species. The order has been expanded from the traditional concept of 12 families (core Caryophyllales), which are characterized by P-type sieve-tube plastids, the presence of betalains (except in Caryophyllaceae) instead of anthocyanins, and frequent occurrence of succulent habit. The four largest families of core Caryophyllales are Aizoaceae (about 2500 species), Amaranthaceae (about 2300 species), Cactaceae (about 2000 species), and Caryophyllaceae (about 2000 species). The expanded order is more difficult to define on the basis of morphology, but anomalous secondary growth, multicellular glands (trichomes), ellagic acid, and naphthaquinones occur frequently. Many representatives grow in marginal environments such as saltmarshes and deserts.

Polygonaceae (about 1000 species) are among the additional families in the expanded order, many

of the others being small, little-known groups. Aizoaceae are unarmed leaf succulents, chiefly African, with numerous tepals (divisions of the perianth) and stamens. Cactaceae are American stem succulents, mostly spiny with no leaves, again with numerous tepals and stamens. Amaranthaceae have reduced flowers, usually with a single ovule. Polygonaceae usually have prominent stipules united to form a scarious bract or ocrea, unusual vascular structure, and swollen nodes. Droseraceae are carnivorous, and capture insects by fly-paper or spring-trap mechanisms. Members of Caryophyllales include carnation (*Dianthus caryophyllus*) in Caryophyllaceae, rhubarb (*Rheum* × *cultorum*) in Polygonaceae, sundews (*Drosera* spp.) and Venus flytrap (*Dionaea muscipula*) in Droseraceae, and jojoba (*Simmondsia chinensis*) in Simmondsiaceae. *See* BEET; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM; SPINACH; SUGARBEET.

Michael F. Fay; Mark W. Chase

Caryophyllidae

A relatively small subclass of the class Magnoliopsida (dicotyledons) of the division Magnoliophyta (Angiospermae), the flowering plants, consisting of 3 orders, 14 families, and about 11,000 species. They have mostly trinucleate pollen and ovules with two integuments (bitegmic) and a nucellus that is more than one cell thick (crassinucellate). Also, the ovules are often distorted by unequal growth so that the micropyle is brought near the funiculus and the chalazal end (campylotropous), or half inverted so the funiculus is attached near the middle (amphitropous) and these are usually borne on a free-central or basal placenta.

Most of these plants contain betalain pigments instead of anthocyanins, and the seeds very often have a perisperm. Most of the families and species of the subclass belong to the order Caryophyllales. The other orders (Plumbaginales and Polygonales) have only a single family each. *See* CARYOPHYLLALES; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM; PLUMBAGINALES; POLYGONALES.

Arthur Cronquist; T. M. Barkley

Cascode amplifier

An amplifier stage consisting of a common-emitter transistor cascaded with a common-base transistor (see **illus.**). The common-emitter–common-base (CE-CB) transistor pair constitutes a multiple active device which essentially corresponds to a common-emitter stage with improved high-frequency performance. In monolithic integrated-circuit design the use of such active compound devices is much more economical than in discrete designs. A similar compound device is the common-collector–common-emitter connection (CC-CE), also known as the Darlington pair. *See* INTEGRATED CIRCUITS.

The cascode connection is especially useful in wideband amplifier design as well as the design of high-frequency tuned amplifier stages. The improvement in high-frequency performance is due to the impedance mismatch between the output of the common-emitter stage and the input of the common-base stage. Thus the influence of the Miller effect on the first (common-emitter) transistor is minimal even when the load resistance of the amplifier, R_L, is large, because the load resistance of the first transistor is no longer R_L but the much lower input resistance of the second transistor. Furthermore, since the second transistor is in the common-base configuration, it does not suffer from the Miller effect. This effect limits the frequency response of a common-emitter stage by essentially multiplying the base-to-collector junction capacitance by the voltage gain and making it appear much larger at the input of the stage between base and emitter.

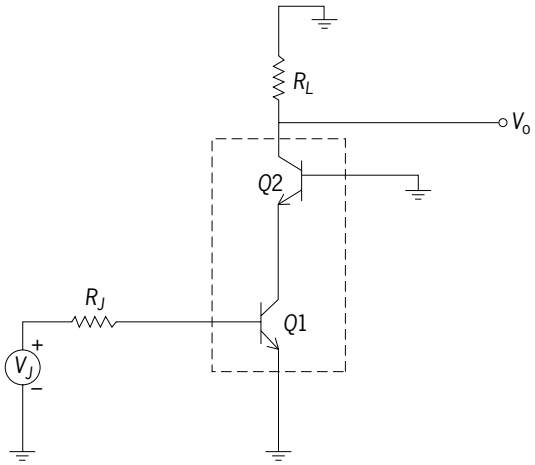

Cascode amplifier. Broken lines enclose a transistor pair consisting of a common-emitter transistor *Q*1 and a common-base transistor *Q*2.

Another important characteristic of the cascode connection is the higher isolation between its input and output than for a single common-emitter stage, because the reverse transmission across the compound device stage is much smaller than for the common-emitter stage. In effect, the second (common-base) transistor acts as an impedance transformer. This isolation effect makes the cascode configuration particularly attractive for the design of high-frequency tuned amplifier stages where the parasitic cross-coupling between the input and the output circuits can make the amplifier alignment very difficult. *See* AMPLIFIER; TRANSISTOR.

Christos C. Halkias

Bibliography. A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Casein

The principal protein fraction of cows' milk. It accounts for about 80% of the protein content and is present in concentrations of 2.5–3.2%. The term casein, derived from the Latin word for cheese, *caseus*,

was introduced into the scientific literature early in the nineteenth century to describe the primary protein of milk. Casein is a mixed complex of phosphoproteins existing in milk as colloidally dispersed micelles 50 to 600 nanometers in diameter. The variable-sized micelles are assembled from spherical subunits of nearly uniform diameter (10 to 20 nm), containing 25–30 casein molecules, and from small amounts of calcium phosphate which, together with serum ionic calcium, plays a significant role in micellar structure. Critical to the stability and size distribution of the micellar state is a calcium-insensitive phosphoprotein component which functions as a so-called protective colloid to the system. When hydrolyzed by rennin (chymosin) as in the cheese-making process, it loses its hydrophilic, carboxy-terminal segment, predisposing the altered micellar system to the aggregation activity of calcium ions and subsequent clot formation. *See* COLLOID; MICELLE.

Caseins can be separated from the whey proteins of cows' milk by gel filtration, high-speed centrifugation, salting-out with appropriate concentrations of neutral salts, acid precipitation at pH 4.3–4.6, and coagulation with rennet (or other proteolytic enzymes), and as a coprecipitate with whey proteins. The first three methods yield preparations in essentially their native micellar state, but are impractical for commercial exploitation. Thus, commercial caseins are produced by methods more amenable to industrial practices (see **illus**).

Acid casein. Fat-free, separated milk (skim milk) is cultured with lactic acid–producing organisms, namely *Streptococcus lactic* or *S. cremoris*, until it clots at pH 4.3–4.6 (isoelectric pH of the casein complex). The clot is broken and heated to about 122°F (50°C) to expel whey, thoroughly washed, pressed, ground, dried, and milled. Alternatively, food-grade HCl or H_2SO_4 is added slowly to agitated skim milk at 86–95°F (30–35°C). The temperature is raised to 111–115°F (44–46°C) to enhance precipitation of the casein, which is collected and processed. Acid caseins are designated according to the acid utilized in their production, such as lactic casein, HCl casein, or H_2SO_4 casein. Acid caseins are insoluble and possess low ash content (see **table**).

Rennet casein. When rennet is added to skim milk, colloidal casein is destabilized and forms a gel. Sufficient rennet and $CaCl_2$ is added to skim milk at about 86°F (30°C) to yield a coagulum in 20 to 30 min. Agitation is initiated following the onset of coagulation to reduce the particle size to optimize processing. After raising the temperature to 131–149°F (55–65°C) to expel whey, the curd is collected and processed. Rennet casein is frequently referred to as *para*-calcium caseinate. It is insoluble and contains a relatively high ash content consisting largely of calcium phosphate (see table). *See* RENNIN.

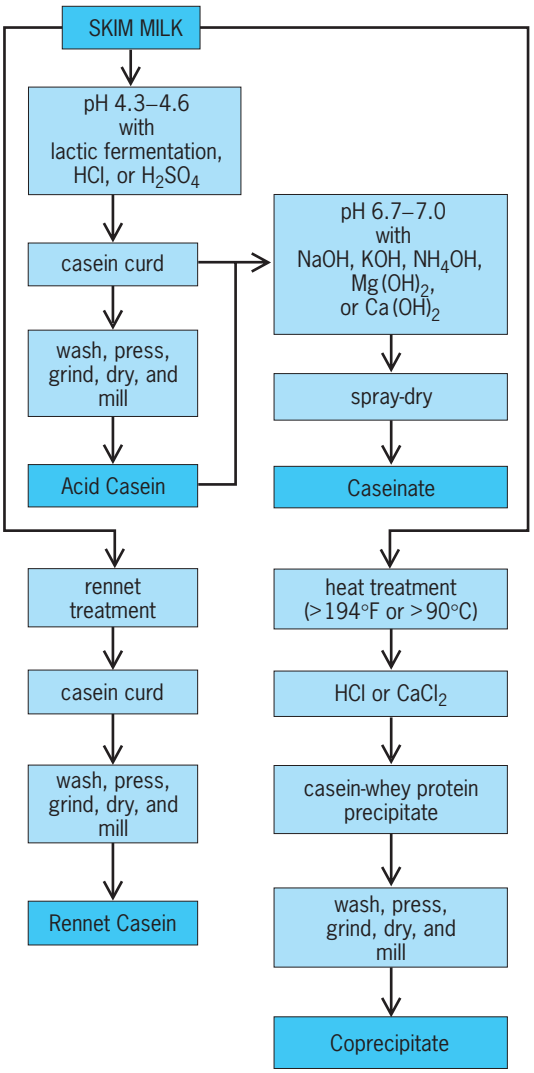

Flow chart of production of commercial casein products from skim milk.

Approximate composition of commercial casein products, in percent

Component	Sodium caseinate	Calcium caseinate	Acid casein	Rennet casein	Coprecipitate
Protein	94	93.5	95	89	89–94
Ash (max.)	4.0	4.5	2.2	7.5	4.5
Sodium	1.3	0.05	0.1	0.02	—
Calcium	0.1	1.5	0.08	3.0	—
Phosphorus	0.8	0.8	0.9	1.5	—
Lactose (max.)	0.2	0.2	0.2	—	1.5
Fat (max.)	1.5	1.5	1.5	1.5	1.5
Moisture (max.)	4.0	4.0	10.0	12.0	5.0
pH	6.6	6.8	—	7.0	6.8

Coprecipitates. In the production of casein isolates, only 80% of the milk protein is recovered, since undenatured whey proteins are soluble at pH 4.6 and are unaffected by rennet. Efforts to increase protein recovery and improve the nutritional quality and functionality of isolates led to the development of coprecipitates. Skim milk is heated to more than 194°F (90°C) to denature the whey proteins and induce protein-protein interactions. Either HCl or $CaCl_2$ is added to induce precipitation of the protein complex, which varies in ash content according to the amount of $CaCl_2$ and HCl employed (see table). $CaCl_2$-induced coprecipitates are usually dispersed with polyphosphate salts prior to drying.

Caseinates. Many applications for casein isolates require that they be water-soluble or dispersable. This is achieved by the slow addition of a suitable alkali to the wet curd of an acid casein, being careful not to exceed a pH of 6.7–7.0. Protein concentrations of about 20% are spray-dried. Although sodium caseinate is most commonly used in the food industry, caseinates containing cations of K, NH_4, Mg, and Ca are produced.

The early production of casein isolates was stimulated by their application in industrial products such as paper, glue, paint, and plastics. These applications have been replaced by petroleum-based polymers. Thus the emphasis has shifted to their utilization in food systems, where they add enhanced nutritional and functional characteristics. They are widely used in the formulation of comminuted meat products, coffee whitener, processed cereal products, bakery products, and cheese analogs. *See* CHEESE; FOOD MANUFACTURING; MILK. J. Robert Brunner

Bibliography. P. F. Fox (ed.), *Developments in Dairy Chemistry*, vol. 1: *Proteins*, 1982; N. P. Wong, *Fundamentals of Dairy Chemistry*, 3d ed., 1988.

Cashew

A medium-sized (20–30 ft or 6–9 m high), spreading evergreen tree (*Anacardium occidentale*) native to Brazil, but now grown widely in the tropics for its edible nuts and the resinous oil contained in the shells. Cashews belong to the family Anacardiaceae in the order Sapindales. The fruit consists of a fleshy, red or yellow, pear-shaped receptacle, termed the apple, 2–3$^1/_2$ in. (5–8.7 cm) long at the distal end of which is borne a hard-shelled, kidney-shaped ovary or nut about 1$^1/_2$ in. (3.7 cm) long. The shell of this nut is about $^1/_8$-in. (3-mm) thick and consists of a smooth, relatively thin outer layer (exocarp) and an inner hard layer (endocarp). Between these layers is a porous layer (mesocarp), which is filled with a caustic black liquid that blisters the skin and makes the processing of the nut difficult. *See* SAPINDALES.

Although cashew trees are spread throughout the tropics, commercial production is centered in India, which handles 90% of the world trade. The trees, mostly seedlings, grow on wastelands with a minimum of care, reach full bearing at 8 years of age, and may continue for 25 years or more. Usually the fruits fall from the trees when mature and are gathered by local labor; the nuts are separated from the apples by hand. Some of the apples are marketed locally but most are wasted.

Processing the nuts is a difficult operation because of the caustic liquid in the shells. The nuts are dried in the sun, cleaned, moistened, and kept for 12 h in heaps or silos to soften the shells. They are then roasted at 350–400°F (176–204°C). In the roasting process the cashew shell liquid is released and collected as a by-product. The kernels are extracted from the shells, heated to remove the enclosing membrane (testa), graded, and packed in tin cans for export.

Shelled cashew nuts, plus cashew shell liquid, are second only to jute in India's exports. More than half of these go to the United States; Europe and Russia are heavy importers also. *See* JUTE.

Cashew nut kernels are eaten as nuts and used extensively in the confectionery and baking trade. The cashew shell liquid is a valuable by-product, containing 90% anacardic acid and 10% cardol, and is used in the varnish and plastic industries.

The cashew apples are too astringent for eating without being processed, but when processed may be used for jams, chutney, pickles, and wine. In India much attention is being given to extending and improving the cashew industry. *See* NUT CROP CULTURE. Laurence H. MacDaniels

Cashmere

The natural fiber obtained from the Cashmere goat, native to the Himalayan region of China and India. The fleece of this goat has long, straight, coarse outer hair of little value; but the small quantity of underhair, or down, is made into luxuriously soft woollike yarns with a characteristic highly napped finish. This fine cashmere fiber is obtained by frequent combings during the shedding season. A microscopic examination reveals that cashmere is a much finer fiber than mohair or wool fiber obtained from sheep. The scales being less distinct and farther apart, the fiber appears to be made of telescoped sections.

Cashmere first became familiar in the beautiful soft, light cashmere shawls for which India has been famous. Today, it is used for such garments as sweaters, sports jackets, and overcoats. Cashmere is soft, lighter in weight than wool, and quite warm; but because of its soft, delicate texture, cashmere is not as durable as wool. *See* ALPACA; CAMEL'S HAIR; LLAMA; MOHAIR; NATURAL FIBER; VICUNA; WOOL.

M. David Potter

Cassava

The plant *Manihot esculenta* (family Euphorbiaceae), also called manioc. It is one of the 10 most important food plants, and the most important

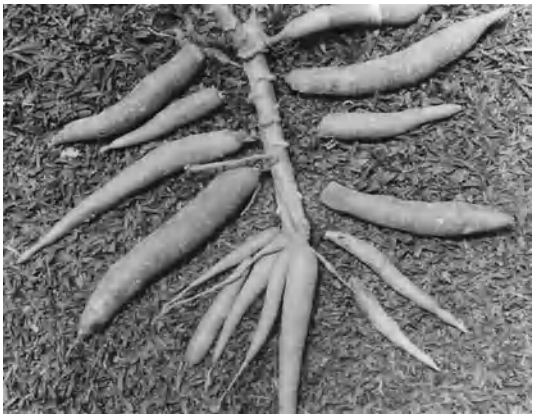

Tuberous roots from a single cassava plant.

starchy root or tuber of the tropics. It originated in Central or South America, possibly Brazil, and was domesticated and widely distributed well before the time of Columbus. Subsequent distribution has established cassava as a major crop in eastern and western Africa, in India, and in Indonesia. Brazil continues to be the largest producer. *See* GERANIALES.

Description and cultivation. The cassava plant is a slightly woody, perennial shrub reaching 10 ft (3 m) in height. The leaves are deeply palmately lobed; the flowers are inconspicuous, and the prominent capsules are three-seeded and explosive at maturity. The roots (see **illus.**) are enlarged by the deposition of starch and constitute the principal source of food from the plant. Normal yields are about 10 lb (4.5 kg) per plant. The leaves are also eaten (after cooking), and are noteworthy for their high protein content. The plant is propagated from mature stems which are planted without special treatment. Tuberization occurs gradually; about 10 or 12 months from planting to harvest is normal. Cassava can be grown for 2 or more years, however, and thus is a food that can be used at any season. Once it is harvested, however, the roots deteriorate within a few days.

Use and food value. The chief use of cassava is as a boiled vegetable. It is also a source of flour, called farinha in Brazil and gari in western Africa, and of toasted starch granules, the familiar tapioca. It can be processed into macaroni and a ricelike food. In the form of dried chips, cassava root is an important animal feed in the international market and is used extensively in Europe. In spite of its popularity, however, cassava root is a poor food. Its protein content is extremely low, and its consumption as a staple food is associated with the protein deficiency disease kwashiorkor. In addition, all parts of the plant contain glucosides of hydrocyanic acid, substances which on decomposition yield the poisonous hydrocyanic acid (HCN, prussic acid). Chronic diseases including goiter are common in regions where cassava is a staple food.

Franklin W. Martin

Diseases. Much of the older literature implies that cassava diseases are of minor importance, but this is not the case. Numerous diseases caused by fungi, bacteria, viruses, and nematodes affect cassava. The most important cassava disease in Africa is African common cassava mosaic, which possibly also occurs in Asia. The causal agent is a virus transmitted by whiteflies. Most plants in Africa are infected, and losses of 30–80% are reported. Some lines with resistance are available, but losses can also be minimized where disease-free planting material can be produced.

Cassava bacterial blight, caused by *Xanthomonas manihotis*, is distributed worldwide and is perhaps the most devastating disease on a worldwide basis. Disease-free planting material and resistant varieties are control measures. The four species of *Cercospora* that induce leaf spots are widespread and cause significant losses. Some varieties resistant to Cercospora leaf spots occur, and wider spacing to reduce humidity can also reduce disease severity. Other important foliar diseases are superelongation (caused by *Sphaceloma manihoticola*), Phoma leaf spot, cassava rust (caused by *Uromyces* spp.), anthracnose (caused by *Colletotrichum* spp.), and cassava ash (caused by *Oidium manihotis*). Stem pathogens (*Glomerella* spp. and *Botryodiplodia* spp.) and root rot pathogens (*Phytophthora* spp. and *Pythium* spp.) may be destructive in some areas. *See* PLANT PATHOLOGY.

H. David Thurston

Bibliography. T. Brekelbaum, A. Bellotti, and J. C. Lozano (eds.), *Proceedings of the Cassava Protection Workshop*, CIAT, Cali, Colombia, 1977; J. H. Cock, *Cassava*, 1984; J. H. Cock, *Cassava: New Potential for a Neglected Crop*, 1985; W. O. Jones, *Manioc in Africa*, 1959; J. C. Lozano and R. H. Booth, Diseases of cassava (*Manihot esculenta* Crantz), *Pest Articles and New Summaries*, 20:30–54, 1974; H. D. Thurston, *Tropical Plant Disease, American Phytopathological Society,* 1984.

Cassiopeia

A constellation named for the queen of Ethiopia, who, in Greek mythology, claimed to be more beautiful than the Nereids, daughters of the sea god, Poseidon. Poseidon sent a sea monster, Cetus, in revenge, and Cassiopeia's daughter Andromeda was sacrificed to appease the monster. (In some versions of the myth, Cassiopeia boasted of Andromeda's beauty.) However, Andromeda was rescued by Perseus. The constellations named after these figures are adjacent in the sky. *See* ANDROMEDA; PERSEUS.

The celestial atlas by Bayer (1603) shows Cassiopeia as a queen in a big chair. To the side of the chair in the atlas is the bright star seen as a supernova by Tycho Brahe in 1572. In the sky, Cassiopeia is a prominent W of bright stars in the northern Milky Way (see **illustration**). It is circumpolar, never setting, from midnorthern latitudes. *See* SUPERNOVA.

The modern boundaries of the 88 constellations,

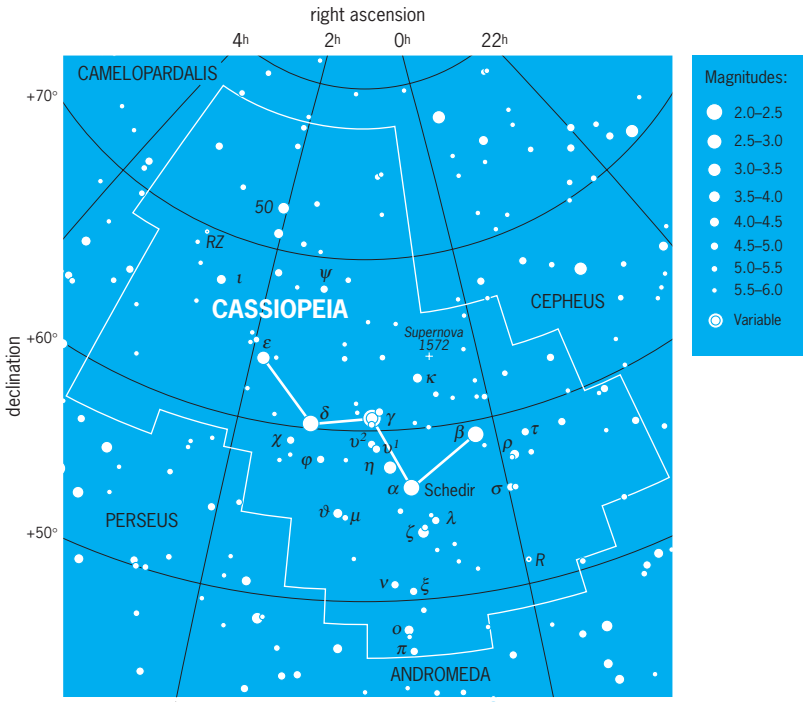

Modern boundaries of the constellation Cassiopeia, mythological queen of Ethiopia. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Cassiterite

A mineral having the composition SnO_2. Cassiterite is the principal ore of tin. Its crystals are of the tetragonal rutile structure type, usually in prisms terminated by dipyramids. Twins with a characteristic notch called visor tin are common. Cassiterite is usually massive granular, but may be in radiating fibrous aggregates with reniform shapes (wood tin). The hardness is 6–7 (Mohs scale), and the specific gravity is 6.8–7.1 (unusually high for a nonmetallic mineral). The luster is adamantine to submetallic. Pure tin oxide is white, but cassiterite is usually yellow, brown, or black because of the presence of iron.

Cassiterite is genetically associated with granitic rocks and pegmatites and may be present as a primary constituent. More commonly it is found in quartz veins in or near such rocks associated with tourmaline, topaz, and wolframite. Cassiterite is most abundantly found as stream tin (rolled pebbles in placer deposits). Early production of tin came from cassiterite mined at Cornwall, England, but at present, the world's supply comes mostly from placer or residual deposits in the Malay States, Indonesia, Zaire, and Nigeria. It is also mined in Bolivia. *See* TIN.

Cornelius S. Hurlbut, Jr.

Cast iron

A generic term describing a family of iron alloys containing 1.8–4.5% carbon. Cast iron usually is made into specified shapes, called castings, for direct use or for processing by machining, heat treating, or assembly. In special cases it may be forged or rolled moderately. Generally, it is unsuitable for drawing into rods or wire, although to a limited extent it has been continuously cast into rods and shapes from a liquid bath or swaged from bars into smaller-dimensional units. Silicon usually is present in amounts up to 3%, but special compositions are made containing up to 6% (Silal) and up to 12% (Duriron).

Cast iron of the above composition range is often made into blocks or rough shapes and called pig iron. It is an intermediate form of cast iron used for remelting into iron castings. Pig iron usually is produced from iron ore and coke smelted into liquid iron in the blast furnace. This liquid iron may be poured into sand molds of specific design which, when cleaned and trimmed, become ingot molds for steel manufacture. In these instances the cast iron may be referred to as direct metal, indicating its production in the blast furnace without the intermediate remelting in the cupola furnace of the foundry, which is common practice in the manufacture of iron castings.

Classification. Cast iron may be purchased in several commercial grades called gray iron, chilled iron, mottled iron, white iron, malleable iron, ductile iron,

spheroidal graphite iron, nodular iron, and austenitic cast iron.

Originally, cast iron was graded by examination of fracture characteristics of the surfaces produced when a sample was broken; its true composition was unknown. This method of grading first was applied to pig iron, which was classified similarly as white, mottled, or gray. The color of the fracture governed the classification of the product until about 1888, when extensive work by W. J. Keep revealed that the silicon content in a large measure determined the character of the fracture. Further work clarified the relationships between composition, microstructure, and fracture characteristics, and gradually the importance of the examination of fractures receded until the present time, when it is used only to classify the products generally. The chemistry of cast iron today is quite well known, and gray, white, and mottled iron can be selected on the basis of composition charts which mark the areas of their occurrence.

Further developments in cast irons have adapted heat-treatment processes to produce special properties that permit classification of special grades of gray, malleable, and white cast irons. Also, metals such as cerium and magnesium have been used to produce the spheroidal graphite structures which are characteristic of ductile irons.

Properties and uses. Gray iron is an iron-base material within the broad composition limits mentioned above. Of its total carbon content, more than one-half is in the form of graphite flakes. It is the presence of the graphite flakes which produce a gray color in a fresh fracture. The remaining carbon is dissolved in the metal and is called combined carbon. It may be present in amounts ranging from 0.05 to 1.20%. Gray iron is the mainstay of iron foundry production. Its application covers practically all engineering fields involving construction, machinery, engines, power, transport, mining, and many daily uses, such as stove parts, pots and pans, and hardware.

Mottled iron solidifies as a mixture of white and gray irons. The fracture is a mottled mixture of white and gray areas. The fracture may contain as few as 10% white areas in a gray background or the reverse, namely, 10% gray areas in a white background. Mottled irons are useful for abrasion- or erosion-resisting services, where their composite structure provides a hard surface with some ability to deform. The high frictional resistance of its surface is occasionally utilized.

Chilled iron has all but disappeared as a descriptive term. It was applied to cast iron which was poured into a metal mold or chill. Frequently, a gray iron would be processed in this manner, and if the cooling rate was rapid enough, a white-iron casting or a surface zone of white iron on the casting would develop, backed by the less rapidly cooled gray iron inside. Since glass molds, for example, are often made by casting one of their surfaces against a chill, such castings would be called chilled iron. Unfortunately, the term spread to the part-white, part-gray types of cast iron and more inaccurately to a solid white iron which had been cast in a chill.

White iron is an iron-base material within the broad composition limits mentioned above whose total carbon content is in the combined form. It therefore contains little or no graphite. A fresh fracture is white (ground-glass white). Its characteristic structure usually is obtained by lowering the carbon content or the silicon content or both, so that the graphitizing power of the composition is subdued.

Malleable cast iron is made as a white-iron casting within the narrower composition range that the term implies. Annealing produces a tough, bendable, and machinable cast iron. During the anneal, all or most of the carbon is precipitated in ragged nodules of graphite. The end product is available in a number of grades that vary in strength, hardness, and toughness. Malleable iron is used widely in the transportation, valve, fitting, and hardware industries for innumerable small parts.

Ductile iron (also called spheroidal graphite iron or nodular iron) is a product of high carbon and high silicon content in which most of the carbon has been coagulated into spheres during processing in the liquid state. During processing, an ingredient such as magnesium or cerium is added, followed by another substance such as silicon, calcium, or combinations of these ingredients. The resulting product acquires the structure of steel peppered with spheres of graphite. Its strength closely approaches that of steel, and it may be heat-treated to develop a ductility exceeding that of malleable cast iron. Its properties make it useful for a great many engineering applications requiring strength and toughness levels which previously were beyond the reach of iron foundry products.

Austenitic cast irons are available in the gray-iron and ductile-iron classifications. When gray or ductile iron is alloyed with substantial amounts of nickel, manganese, silicon, or other elements to alter the basic crystalline structure from magnetic alpha iron to nonmagnetic gamma iron, the resulting product becomes austenitic cast iron. These irons generally offer better corrosion- and heat-resistance than the low-alloy irons. *See* ALLOY; HEAT TREATMENT (METALLURGY); IRON; IRON ALLOYS; METAL CASTING; STEEL; WROUGHT IRON. James S. Vanick

Bibliography. H. T. Angus, *Cast Iron: Physical and Engineering Properties*, 2d ed., 1978; E. Dorazil, *High Strength Austempered Ductile Cast Iron*, 1991; H. Fredriksson and M. Milert (eds.), *Physical Metallurgy of Cast Iron: Materials Research Society Symposia Proceedings*, 1985; Metallurgical Society of the American Institute of Mining, Metallurgical and Petroleum Engineers, *History of Iron and Steelmaking in the United States*, 1961; L. H. Van Vlack, *Elements of Materials Science and Engineering*, 6th ed., 1989.

Castor plant

A plant, *Ricinus communis*, belonging to the spurge family (Euphorbiaceae). Castor varies greatly in height and in color of foliage, stems, and seeds and

also in size and oil content of the seeds. The palmately lobed leaves are borne more or less alternately. The main stem is terminated by a raceme (primary). After appearance of the primary raceme, branches on the main stem are each in turn terminated by a raceme. Growth continues sequentially as long as the plant lives, with branches arising from recently terminated branches. Thus a plant can have racemes in all stages of development from mature to prebloom. Mature racemes mostly have 10 to 70 capsules, each capsule containing three seeds. In frost-free areas castor can attain heights of 9 to 11 m (30 to 36 ft). Castor seeds are poisonous and also contain allergens.

Distribution is in the warmer regions of the world, the plant often growing in waste places. The castor oil plant has been of utilitarian value since antiquity. Oil from the seeds is among the world's oldest nonfood products in commerce. Castor oil contains about 85% ricinoleic acid, used in making industrial products such as alkyde resins for surface coatings, blown oil used in plasticizers, cracked oil for production of synthetic perfumes, nylon, sebacic acid, synthetic detergents, drying oils, and special lubricating oils.

Farmers usually grow castors on less productive land. Culture usually is not mechanized. However, plant breeders in the United States developed types (nonshattering capsules, dwarfs, and F_1 hybrids similar to hybrid corn) suitable for completely mechanized production; concurrently, engineers developed harvesting machinery. Production in the United States, mostly in west Texas, peaked in 1968 but declined to nil because of larger economic return from food and fiber crops. Brazil is the world's largest producer and exporter of castor seed and oil. Leroy H. Zimmerman

Bibliography. E. A. Weiss, *Castor, Sesame and Safflower*, 1971.

Casuarinales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Hamamelidae of the class Magnoliopsida (dicotyledons). The order consists of a single family (Casuarinaceae) and genus (*Casuarina*), with about 50 species. Native to the southwestern Pacific region, especially Australia, they are sometimes called Australian pine. They are trees with much reduced flowers and green twigs that bear whorls of scalelike, much reduced leaves. Some species are grown as street trees in tropical and subtropical regions. *See* HAMAMELIDAE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM. Arthur Cronquist

Cat

The term applied to any member of the mammalian family Felidae. Commonly, the term is restricted to the domestic cat and those felids that resemble it in size, shape, and habits. Those that are larger in size are referred to as the big cats.

All members of the cat family have a round head, are digitigrade (walk on their toes), have retractile claws (with the exception of the cheetah), and have 30 teeth with the dental formula I 3/3 C 1/1 Pm 3/3 M 1/1. All species are carnivorous.

Domestic cats. The origin of domestic cats is unknown, but it is established that they have been associated with humans for many centuries. Remains of the cat have been found with artifacts and remains of cave people, but it is generally conceded that these were not domesticated species. It has been suggested that domestic cats are descended from Old World wildcats such as the caffer cat (*Felis lybica*), an African species. However, there is no evidence for this belief, although it is known that wild and domestic species interbreed. Domestic cats are generally characterized by long, thin tails, straight ears, and short hair of a variety of colorations. These are the so-called alley cats and are mixed breeds. The pure breeds, however, are notable exceptions to this general description. There is a long-haired race, apparently developed in Persia, which is represented by the Persian and the Angora. The Abyssinian breed, in contrast to the common domestic cat, is characterized by being ruddy brown in color and having a longer face and ears. Each short hair is ticked or grizzled in appearance because of its alternate dark and light banding. This is known as an agouti condition.

There are at least two tailless breeds, of which the best known is the Manx from the Isle of Man. It has a boneless tuft of hair instead of a tail. The other tailless variety is found in Japan, and there is a question as to whether the Manx was a mutation or was imported. One of the most popular breeds is the Siamese cat, an animal with great independence and individuality. Its eyes are deep blue, it has a long, kinky tail, and its fur is short and cream to buff in color.

Most cats have two periods of heat each year. After a gestation period of about 8 weeks, a litter of about five kittens is born in April and August. They are helpless, blind, and entirely dependent upon the mother after birth. The kittens are protected by the female since the male will sometimes attack them. Scientific breeding of cats is difficult, since the cat is a vagabond by nature and interbreeding usually occurs in time, making it difficult to keep races pure.

Wildcats. In addition to cats such as the lynx, bobcat, serval, ocelot, puma, leopard, lion, tiger, jaguar, and cheetah, there are a few lesser known forms that are of interest to the biologist and naturalist. Among these are the Scottish wildcat and the caracal. The Scottish wildcat resembles the domestic species but is more heavily built. It has become virtually extinct, except for the small numbers in the Scottish Highlands. It is a nocturnal animal and is both ferocious and untamable. This cat is regarded as a separate subspecies of the European wildcat (*F. sylvestris*). The long-legged caracal (*Caracal caracal*) is built better for running than are other cats. It is found in the deserts of North Africa and western Asia, where it

hunts and kills many types of animals. It resembles a lynx in having ear tufts but is more closely related to the serval.

Lynx. There are a number of species of cats known by this name. The Canadian lynx (*Lynx canadensis*) is the largest North American lynx. It is found in pine-forested areas of Canada and the northern United States, where it feeds upon rabbits, mice, and birds, and during the winter on the remains of animals left by other scavengers. This animal has become relatively rare as a result of land settlement. It is a stout-bodied cat about the size of a cocker spaniel. It is trapped for its fur, and the flesh is reputed to be edible. Usually two cubs are born each year, and the gestation period is about 10 weeks. The maximum weight of the adult is about 30 lb (13 kg), and the life-span is about 10 years.

The European lynx (*L. lynx*) ranges throughout the large forested areas of northern and central Europe. Unlike the Canadian lynx, it will leave the forest to attack sheep and goats when food is scarce. *Lynx pardellus* is known as the Spanish lynx and is the smallest of the species. It is found in the Pyrenees, on the French side of the ranges. *Lynx rufus*, known as the bobcat, bay lynx, or wildcat, is very similar to the Canadian lynx but is smaller, and its ears are not conspicuously tufted. It ranges over the eastern United States from Maine to Georgia and has been divided into a number of subspecies. It is fierce and feeds on mammals and birds.

Lynx differ from other felids in having 28 instead of 30 teeth; an upper premolar is absent.

Serval. Found in the scrub grass regions south of the Sahara desert, this cat (*Leptailurus serval*) usually feeds on birds and small mammals. Being fawn colored with black spots, it blends with the landscape. It is one of the fastest-running felids and can leap up to 6 ft (2 m) to catch a flying bird.

Ocelot. This species (*F. pardalis*) is arboreal and feeds on monkeys and birds in the forested regions of South America, its native habitat. Smaller and more ornamental than the jaguar, it is known as the painted leopard. The head and back are golden and the flanks are silvery, and there are rows of somewhat metallic spots on the body with stripes on the head and neck. Near habitations this cat will feed on domestic and farmyard animals. The adult, which is almost 3 ft (0.9 m) long, can be easily trained when captured young and is apparently harmless to humans.

Puma. This member of the cat family, also known as the cougar or mountain lion, is found in various habitats in the Americas and nowhere else. The puma (*F. concolor*) has been described as resembling a small, short-legged lioness because of its uniform color. It is a good climber and is nocturnally active in its hunting. It has been theorized that the puma aided the extinction of the horse in the Americas before the discovery of these continents. The male may weigh as much as 200 lb (90 kg) and have a length of about 7 ft (2 m), and the female is nearly as large. The male is not adverse to eating the young.

Jaguar. This cat (*F. onca*) is indigenous to Central and South America, where it is found along the banks of forested rivers and in marshy areas. Remaining hidden during the day, it is a nocturnal hunter, a strong swimmer, the largest cat, and a fierce carnivore. The jaguar has been called the leopard of the New World, and its pelts are commercially important.

Cheetah. This is a peculiar animal in that it has some doglike characteristics although it is included in the cat family. Its claws are nonretractile, its legs are long, and it does not climb. A single species, *Acinonyx jubatus*, occurs from northern Africa to Asia as far south as India. This animal is doubtless the fastest animal in running over short distances. It hunts in small groups and attacks such animals as gazelles and antelopes.

Leopard. This species ranges throughout Africa and Asia and is regarded as the most athletic member of the cat family, being equally adept at climbing, swimming, jumping, and running. *Felis pardus* is the only species of true leopard; however, *Uncia uncia*, the snow leopard or ounce, also is called a leopard. This latter species occurs in the Himalayas of central Asia at high altitudes. The leopard has a more extensive distribution than any other species of Felidae. These animals hunt ceaselessly and have been known to eat humans. They can be tamed in captivity when young, but few are tamed since they cannot be trusted.

Lion. Felis leo was a common animal in the Near East, and its range extended into Europe during prehistoric times. In historical times it was recorded as continuing to inhabit the Balkan Peninsula. It has come to be restricted in its distribution as a result of the advances of humans. It is still abundant in the savanna areas between Senegal and East Africa, where there is much herbivorous fauna to prey upon. It is not found in deserts or thickly forested regions. The lion may weigh as much as 550 lb (248 kg) and reach a height of 3 ft (0.9 m) and a length of $6^1/_2$ ft (1.9 m). It is the largest African carnivore. There is one litter of 2–4 young per year after a gestation period of about 15 weeks. The life-span may be as long as 40 years. The lion hunts at night and is silent while approaching its prey in order to surprise it. Various races of this species have been described, such as Cape, Masai, Somali, and Indian lions, but they show few differences in color and size.

Tiger. The tiger is represented by a single species, *F. tigris* (see **illus.**). It inhabits the forests of Asia,

Tiger (*Felis tigris*).

Sumatra, Java, and southern Siberia. As a consequence of the wide climatic and geographical distribution of tigers, they show a wide range of coloration and size, and a number of varieties are recognized. After a gestation period of 3 months, two or three cubs are born helpless and blind. There is one litter every second year, and usually only one cub of each litter survives. Tigers breed in captivity, and crosses between the lion and tigress (the liger) and between the tiger and lioness (the tigron) have been obtained. The tiger takes any food available and, like many members of the cat family, returns to the kill until it is consumed. *See* CARNIVORA. Charles B. Curtin

Cat scratch disease

Cat scratch disease in humans is typically a benign, subacute regional disease of the lymph nodes (lymphadenopathy) resulting from dermal inoculation of the causative agent. This agent has been demonstrated to be *Bartonella* (formerly *Rochalimaea*) *henselae*, a bacterium that has been isolated from immunocompromised patients with bacillary angiomatosis, a distinctive and potentially deadly vascular proliferative host response in skin, bone, or other organs associated with the presence of clumps of bacteria.

Although cat scratch disease was first described in France in 1950, the causative bacterial agent remained obscure until 1992, when *B. henselae* was implicated by serologic studies. In 1993 and 1994, major progress was made in the epidemiology of cat scratch disease. The domestic cat is the major reservoir of *B. henselae*, and the cat flea, *Ctenocephalides felis*, is the main vector of transmission from cat to cat. *Bartonella clarridgeiae* has been isolated from domestic cats, and at least two human cases of cat scratch disease have been associated with this bacterium. Recently, a new bacterium, *B. koehlerae*, has also been isolated from the blood of domestic cats.

Etiology. The cause of cat scratch disease has long been in question. It was initially considered a virus, then a gram-negative bacterium, and only recently was the specific agent confirmed. In 1983 a bacillus was identified by Warthin-Starry (silver deposition) staining of lymph node biopsies from 39 patients with cat scratch disease. In 1988 a pleomorphic, gram-negative bacterium was isolated from the lymph node of an infected individual, and was named *Afipia felis*. Its isolation was difficult, and only a few strains were ever isolated. In addition, serology was not highly specific, and it was difficult to standardize. For some years, *A. felis* was considered to be the most probable agent causing cat scratch disease. However, in the 1990s evidence clearly implicated *B. henselae* as the causative agent.

Bartonella henselae is morphologically very similar to *A. felis* when examined by Warthin-Starry staining, which may explain the previous confusion. Serological studies and isolation of the organism from lymph nodes of probable cat scratch disease cases substantiate the major role played by *B. henselae* in the etiology of the disease. *Bartonella henselae* is a small, curved gram-negative rod which exhibits a twitching motility. Optimal growth on solid media is obtained by use of rabbit blood agar and incubation at 95°F (35°C) in 5% carbon dioxide (CO_2). *Bartonella clarridgeiae* possesses flagella, which are absent in *B. henselae*.

Epidemiology. There were an estimated 22,000 individuals with cat scratch disease in the United States in 1992, and 2000 of these required hospitalization. Cat scratch disease occurs in immunocompetent patients of all ages, with 55–80% being less than 20 years of age. However, cases are more often reported among children and teenagers than adults (45–50% of infected individuals are less than 15 years old). Cat scratch disease is considered the most common cause of chronic benign adenopathy in children and young adults, but it can easily be confused with neoplastic conditions. More than 90% of cases have a history of contact with cats, and 57–83% recall being scratched by a cat. Incidence varies by season; most cases occur in the fall and winter. More cases are observed in males than females.

A 1993 epidemiological study reported that infected individuals were more likely than healthy cat-owning control subjects to have at least one kitten 12 months of age or younger, to have been scratched or bitten by a kitten, and to have at least one kitten with fleas. Of 45 individuals observed, 38 (84%) had antibodies to *B. henselae* compared to 4 of 112 control subjects (3%). Interestingly, 39 of 48 (81%) of the cats of infected individuals also had antibodies to *B. henselae*, as compared with 11 of 29 control cats (38%).

Bartonella henselae antibodies have been found to be common in cats. Geographical variations in seroprevalence were also observed, with 60% of cats seropositive in the Southeast of the United States, but with very low positivity levels in the Midwest, Alaska, and the Rocky Mountains. The presence of bacteria in the blood (bacteremia) is also very common in cats, up to 50% in warm and humid areas of North America, Australia, and the Philippines, and the range is 10–15% in pet cats in most of western Europe and in Australia. Bacteremia can last several months to a few years in cats, but very high bacteremia levels usually do not last more than 4–6 weeks. Cats can be co-infected with *B. henselae* and *B. clarridgeiae* and with various *B. henselae* types.

Clinical signs. In humans, 1–3 weeks may elapse between the scratch (or bite) and the appearance of clinical signs. In 50% of the cases, a small skin lesion, often resembling an insect bite, appears at the inoculation site (usually on the hand or forearm) and evolves from a pimple (papule) to a skin blister to partially healed ulcers. These lesions resolve within a few days to a few weeks. Inflammation of lymph nodes develops approximately 3 weeks after exposure, is generally unilateral, and commonly appears in the epitrochlear, axillary, or cervical lymph nodes.

Swelling of the lymph node is usually painful and persists for several weeks or months. In 25% of the cases, a discharge of pus occurs. A large

majority of the cases show signs of systemic infection, such as fever, chills, malaise, anorexia, or headaches. In general, the disease is benign and heals spontaneously without aftereffects. Atypical manifestations of cat scratch disease occur in 5–9% of the cases. The most common is Parinaud's oculoglandular syndrome (enlargement of the lymph node around the eye and conjunctivitis); but also encephalitis, degenerative bone lesions, and thrombocytopenic purpura (a bleeding disorder due to decreased platelet levels) may occur. Usually, complete recovery occurs with few sequelae. The histological findings characteristically include stellate caseating granulomas, micro abscesses, and follicular hyperplasia. Warthin-Starry staining may reveal bacilli.

No major clinical signs of cat scratch disease have been reported in cats, although enlargement of the lymph nodes caused by a cat scratch disease–like organism has been reported. Fever, lymphadenopathy, moderate neurological symptoms, and reproductive disorders have been reported in experimentally infected cats.

Diagnosis. For years, the diagnosis was based on clinical criteria, exposure to a cat, failure to isolate other bacteria, or histologic examination of biopsies of lymph nodes. A skin test was also developed. However, the skin test antigen, prepared from pasteurized pus from lymph nodes of patients with cat scratch disease, was not standardized, and concerns were raised about the safety of such a product. Since 1992, serological tests and techniques to isolate the organism from human specimens have been developed. Additionally, the polymerase chain reaction is being used to confirm infection by *Bartonella*. At present, diagnosis is mainly based on a serological titer of 1 to 64 ratio or greater (or a fourfold titer increase between an early and late serum samples) using an indirect immunofluorescence test and a history of cat scratch or bite.

In cats, the organism is isolated from blood samples. The samples are collected in pediatric lysis-centrifugation or EDTA tubes. EDTA tubes are frozen at −70°C for a few days to several weeks before plating. After centrifugation, the pellet is resuspended in inoculation media and plated onto 5% rabbit blood agar.

Treatment. Most individuals with cat scratch disease experience mild illness and require minimal treatment. Usually antimicrobial therapy is not necessary since spontaneous resolution is common. In severe forms, antibiotics such as ciprofloxacin, rifampin, or gentamicin have been recommended. Use of oral azithromycin for 5 days has shown significant clinical benefit in typical cat scratch disease. Erythromycin (500 mg four times a day) and doxycycline are effective antibiotics in the treatment of bacillary angiomatosis and may be recommended in immunodeficient persons suffering from cat scratch disease. Rifampin is recommended as a second drug in combination with either erythromycin or doxycycline for severely ill patients. Bruno B. Chomel

Bibliography. B. B. Chomel et al., Experimental transmission of *Bartonella henselae* by the cat flea, *J. Clin. Microbiol.*, 34:1952–1956, 1996; M. J. Dolan et al., Syndrome of *Rochalimaea henselae* adenitis suggesting cat scratch disease, *Ann. Intern. Med.*, 118:331–336, 1993; C. K. English et al., Cat scratch disease: Isolation and culture of the bacterial agent, *J. Amer. Med. Ass.*, 259:1347–1352; J. E. Koehler, C. A. Glaser, and J. W. Tappero, *Rochalimaea henselae* infection: A new zoonosis with the domestic cat as reservoir, *J. Amer. Med. Ass.*, 271:531–535, 1994; D. L. Kordick et al., *Bartonella clarridgeiae*, a newly recognized zoonotic pathogen causing inoculation papules, fever, and lymphadenopathy (cat scratch disease), *J. Clin. Microbiol.*, 35:1813–1818, 1997; R. L. Regnery, M. Martin, and J. G. Olson, Naturally occurring "*Rochalimaea henselae*" infection in domestic cat, *Lancet*, 340:557–558, 1992; K. M. Zangwill et al., Cat scratch disease in Connecticut: Epidemiology, risk factors, and evaluation of a new diagnostic test, *N. Engl. J. Med.*, 329:8–13, 1993.

Cataclysmic variable

A type of close binary star system containing a cool star transferring material to its hotter, high-density, degenerate white dwarf companion. The mass transfer results in a large range of observed variability, including cataclysmic events called outbursts, which can increase the brightness of the systems by 2–10 magnitudes (a logarithmic scale with each magnitude being a factor of 2.5 in brightness) from quiescence, equivalent to a factor of 6–10,000 times in intensity. The specific behavior of each system and the extent and cause of the variability are related to whether the transferred material accumulates in an accretion disk surrounding the white dwarf, or whether it flows in a ballistic stream directly from the cool star to the white dwarf surface. Which of these processes will occur depends on the magnetic field strength of the white dwarf and the separation of the two stars in the binary, properties which are determined during the formation of the system (see **illustration**). *See* MAGNITUDE (ASTRONOMY).

Accretion disk systems. If the white dwarf magnetic field is under 100 tesla (1 megagauss), an accretion disk will form and extend close to the white dwarf surface, while a hot spot will form where the mass stream from the cool star hits the disk (illus. *a*). Systems in which the mass transfer occurs in this way constitute the most common type among the roughly 1500 known cataclysmic variables. They can be further classified as novae, dwarf novae, and novalike systems. These categories probably represent different phases of evolution of similar systems, and each is determined by the current mass-transfer rate.

Novae are the most spectacular cataclysmic variables, with 7–10-magnitude (factor of 630–10,000 times) outbursts caused by thermonuclear runaways triggered when a critical mass of hydrogen builds up in the atmosphere of the white dwarf. Usually this occurs every few thousand years, although a handful of novae with high-mass white dwarfs and mostly

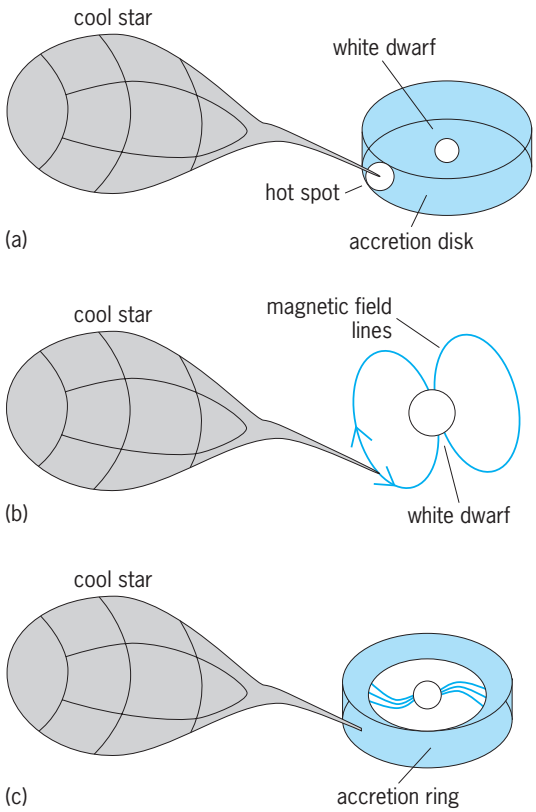

Visualizations of the components of cataclysmic variables. The components are not drawn strictly to scale, for ease of viewing. (*a*) Accretion disk system, the most common type. The shape of the cool star is distorted due to the gravitational influence of its close companion. (*b*) AM Herculis, or polar system. (*c*) DQ Herculis, or intermediate polar system.

giant cool stars recur on time scales of tens of years. There are typically 20 novae per year in a galaxy such as the Milky Way. Large amounts of mass ejection (1/10,000 the mass of the Sun) during the outburst and subsequent expansion of the shell and dust formation, as the nova returns to its quiescent state in months or years, can produce distinctive spectra and light curves for novae. *See* NOVA.

Dwarf novae have much smaller outbursts of 2–5 magnitudes or a factor of 6–100 times (a small number may be as large as 9 magnitudes or a factor of 4000), and they recur much more frequently, on time scales of weeks or months. Their light curves and theoretical models indicate that their outbursts are probably due to accretion disk instabilities, resulting in increased accretion onto the white dwarf when the disk reaches a critical density. The dwarf novae have a wide range of subtypes, ranging from U Gem or SS Cyg types (named after the variable stars U Geminorum and SS Cygni) with orbital periods greater than 3 h, large mass-transfer rates, and long outburst time scales, to the SU UMa types (named after SU Ursae Majoris, the prototype) with orbital periods under 2 h. These short-period systems can be further divided into the ones with medium mass-transfer rates, lowest-outburst amplitudes, and shortest recurrence times (V1159 Ori types, after V1159 Orionis) and those with very low transfer rates, very high outburst amplitudes, and very long recurrence times (WZ Sge types, after WZ Sagittae).

Novalike systems have the highest mass-transfer rates, so that their disks dominate the light output of the system and do not undergo the instability of outbursts. However, for little-understood reasons, the mass transfer may shut down at sporadic times and cause low states of brightness (VY Sc1 types, after VY Sculptoris). Systems with no recorded outbursts are often put into this novalike category until an outburst event occurs.

Magnetic systems. At very high magnetic fields, over 1000 T (10 MG), and small separations (short orbital periods), the mass-transfer stream follows the magnetic field lines of the white dwarf to the white dwarf surface (illus. *b*). This type of magnetic cataclysmic variable is termed an AM Her star (after its prototype, AM Herculis) or as a polar (after its accretion mode). The approximately 80 known polars were primarily discovered by x-ray satellites due to their large x-ray emission, caused by an accretion shock (producing hard x-rays) as the material is channeled to the magnetic pole of the white dwarf and by the subsequent heating of the white dwarf surface (producing soft x-rays). The large field strength locks the spin of the white dwarf to that of the orbit. *See* X-RAY ASTRONOMY; X-RAY TELESCOPE.

For magnetic fields of 100–1000 T (1–10 MG) and larger separations, the material will form an outer accretion disk ring and then flow from its inner edge to the white dwarf following the magnetic field lines (illus. *c*). The 22 known systems that exhibit this behavior are termed DQ Her stars (after DQ Herculis) or intermediate polars (IPs). In this situation, the accretion is more like a curtain of material raining on the magnetic pole, causing primarily hard x-rays, and the white dwarf spin is not locked, so that a signature pulse of light is detected on a time scale (minutes) shorter than the orbital period (hours).

Large-scale ground-based surveys such as the Sloan Digital Sky Survey and the Hamburg Quasar Survey have contributed to the numbers of known polars and intermediate polars. In addition, space surveys of globular clusters accomplished with the *Hubble Space Telescope* have proposed many more candidates that need confirmation. *See* HUBBLE SPACE TELESCOPE; SLOAN DIGITAL SKY SURVEY; STAR CLUSTERS.

Evolution. As most stars are members of multiple systems, the binarity of cataclysmic variables is not unusual. What sets them apart is that at some stage in their life they have a normal low-mass star existing close enough to an evolved white dwarf companion that mass transfer will occur. The evolutionary path is thought to originate with a wide binary with one star having a mass of about 1–10 times that of the Sun. During the course of its normal evolution, this star becomes a giant, engulfing the companion, which then spirals into a closer orbit due to friction, as the giant star becomes a dense, degenerate white dwarf. About 1% of these binaries will eventually evolve into a cataclysmic variable with an orbital period of

several hours. As angular momentum losses continue to shrink the orbit, it evolves to a system with an orbital period near 80 min. At that point, the cool star becomes degenerate, and further loss of mass causes the orbital period to increase again. Evolution calculations show that for the lifetime of the Milky Way Galaxy most cataclysmic variables should be faint systems that have evolved past the orbital period minimum. However, selection effects result in the preferential discovery of the longer-period systems undergoing the highest accretion. In addition, strong magnetic fields may change the time scales and angular momentum losses at different stages of evolution.

During the course of evolution as cataclysmic variables, systems may cycle between nova eruptions (on long time scales), dwarf novae, novalikes, and in-between states of no mass transfer, as nova eruptions and magnetic cycles on the cool star affect the mass-transfer flow. For orbital periods of 2–3 h, the mass transfer may cease, creating a gap in the distribution of orbital periods. About 20 systems are known with extremely short orbital periods (20–50 min), which have helium white dwarfs and degenerate companions.

At late stages of evolution, only the white dwarfs will be visible in systems with faint degenerate companions. These white dwarfs will have had long cooling times, although measured temperatures of white dwarfs in cataclysmic variables show that they are generally hotter than their single counterparts, probably due to accretion heating. Perhaps the most interesting end point of evolution for a cataclysmic variable is the possibility that a system with a white dwarf near 1.4 times the mass of the Sun (the Chandrasekhar limit) will accrete enough matter to undergo a supernova event and collapse to a neutron star. If this does not disrupt the binary, it will then be known as a low-mass x-ray binary. *See* BINARY STAR; NEUTRON STAR; STELLAR EVOLUTION; SUPERNOVA; VARIABLE STAR; WHITE DWARF STAR. Paula Szkody

Bibliography. R. Downes et al., A catalog and atlas of cataclysmic variables, *J. Astron. Data*, vol. 11, 2005; J.-M. Hameury and J.-P. Lasota (eds.), *The Astrophysics of Cataclysmic Variables and Related Objects*, Astron. Soc. Pacific Conf. Ser. no. 330, 2005; B. Warner, *Cataclysmic Variable Stars*, Cambridge University Press, 1995; D. T. Wickramasinghe and L. Ferrario, Magnetism in isolated and binary white dwarfs, *Publ. Astron. Soc. Pacific*, 112:873–924, 2000.

Catalysis

The phenomenon in which a relatively small amount of foreign material, called a catalyst, augments the rate of a chemical reaction without itself being consumed. A catalyst is material, and not light or heat. It increases a reaction rate. *See* ANTIOXIDANT; INHIBITOR (CHEMISTRY).

If the reaction $A + B \rightarrow D$ occurs very slowly but is catalyzed by some catalyst (Cat), the addition of Cat must open new channels for the reaction. In a very simple case [reactions (1)], the two propagation pro-

$$\left.\begin{array}{l} A + Cat \longrightarrow ACat \\ ACat + B \longrightarrow D + Cat \end{array}\right\} \text{Chain propagation}$$
$$A + B \longrightarrow D \qquad \text{Overall reaction} \tag{1}$$

cesses, which are fast compared to the uncatalyzed reactions, $A + B \rightarrow D$, provide the new channel for the reaction. The catalyst reacts in the first step, but is regenerated in the second step to commence a new cycle. A catalytic reaction is thus a kind of chain reaction.

An example that occurs naturally in the stratosphere is the catalytic destruction of odd oxygen (oxone, O_3, and oxygen atoms, O) by nitric oxide (NO) [reactions (2)].

$$\left.\begin{array}{l} O_3 + NO \longrightarrow O_2 + NO_2 \\ O + NO_2 \longrightarrow O_2 + NO \end{array}\right\} \text{Chain propagation}$$
$$O_3 + O \longrightarrow 2O_2 \qquad \text{Overall reaction} \tag{2}$$

See CHAIN REACTION (CHEMISTRY).

If a reaction is in chemical equilibrium under some fixed conditions, the addition of a catalyst cannot change the position of equilibrium without violating the second law of thermodynamics. Therefore, if a catalyst augments the rate of $A + B \rightarrow D$, it must also augment the reverse rate, $D \rightarrow A + B$. *See* CHEMICAL THERMODYNAMICS.

Categories. Catalysis is conventionally divided into three categories: homogeneous, heterogeneous, and enzyme. Heterogeneous catalysis plays a dominant role in chemical processes in the petroleum, petrochemical, and chemical industries. Homogeneous catalysis is important in the petrochemical and chemical industries. Enzyme catalysis plays a key role in all metabolic processes and in some industries, such as the fermentation industry. The mechanisms of these categories involve chemical interaction between the catalyst and one or more reactants. In phase-transfer catalysis the interaction is physical. Electrocatalysis and photocatalysis are more specialized forms of catalysis.

Homogeneous. In homogeneous catalysis, reactants, products, and catalyst are all present molecularly in one phase, usually liquid. Examples are the odd-oxygen reaction [reaction (2)], the hydrogenation of 1-hexene in a hydrocarbon solvent catalyzed by dissolved $[(C_6H_5)_3P]_3RhH$ [reaction (3)] and the hydrolysis of an ester catalyzed by acid [reaction (4)].

$$H_2C{=}CHCH_2CH_2CH_2CH_3 + H_2 \longrightarrow CH_3CH_2CH_2CH_2CH_2CH_3 \tag{3}$$

$$CH_3COOC_2H_5 + H_2O \xrightarrow{H^+} CH_3COOH + HOC_2H_5 \tag{4}$$

See HOMOGENEOUS CATALYSIS.

Heterogeneous. In heterogeneous catalysis, the catalyst is in a separate phase. Usually the reactants and products are in gaseous or liquid phases and the catalyst is a solid. The catalytic reaction occurs on the

surface of the solid. Examples are the dehydration and the dehydrogenation of isopropyl alcohol [reactions (5) and (6)]. Reaction (5) can be affected by

$$CH_3CHOHCH_3 \longrightarrow CH_3HC{=}CH_2 + H_2O \quad \text{Dehydration} \quad (5)$$

$$CH_3CHOHCH_3 \longrightarrow CH_3COCH_3 + H_2 \quad \text{Dehydrogenation} \quad (6)$$

passing the vapors of the alcohol over alumina at about 300°C (570°F), and reaction (6) over copper at 200°C (390°F). It is estimated that one-quarter to one-third of all manufactured goods involve a catalytic antibody. *See* HETEROGENEOUS CATALYSIS.

Enzyme. Transformations of matter in living organisms occur by an elaborate sequence of reactions, most of which are catalyzed by biocatalysts called enzymes. Enzymes are proteins and therefore of colloidal dimensions. Although studies of interrelations between homogeneous catalysis and heterogeneous catalysis have been developing, enzyme catalysis remains a rather separate area in the nature of the catalyst in the type of reactions catalyzed and in the reaction conditions. *See* ENZYME.

Phase transfer. In phase-transfer catalysis the two reactants start in separate phases: a salt as a solid or dissolved in water, and a water-insoluble organic compound as a liquid or dissolved in a nonpolar solvent such as benzene. Mere agitation of the two phases will usually lead to little reaction. However, addition of a suitable phase-transfer catalyst such as an onium salt or a crown ether will considerably augment the rate, ordinarily by carrying the anion into the organic phase, where it is relatively unsolvated and therefore reactive. Such catalytic reactions are run in batch reactors with catalyst at the 1–3% level. Thus, the turnover number (a value related to the number of sets of molecules that react) is about 50 in typical phase-transfer catalysis. For example, an alkyl halide (RX) dissolved in benzene agitated with aqueous potassium cyanide (KCN) forms RCN + KX in the presence of tetrabutyl ammonium chloride as a catalyst. *See* PHASE-TRANSFER CATALYSIS.

Electrocatalysis. Electrochemical processes ordinarily involve an overpotential (overvoltage). Electrocatalysis aims at reducing the overpotential with consequent increase in rate and often of selectivity. The catalytic action may be innate to the electrode (for example, platinum electrodes in a hydrocarbon reaction where dissociative adsorption is involved); it may result from derivatization of the electrode (for example, by a porphyrin); or it may be indirect [for example, the electrochemical hydrogenolysis of halobenzene (RX) with added anthracene (Ar); reactions (7)].

$$\begin{aligned} Ar + e \rightleftharpoons AR^{-}\cdot &\xrightarrow{RX} R\cdot + X^{-} + Ar \\ R\cdot + e \rightleftharpoons R^{-} &\xrightarrow{H_2O} RH + OH^{-} \end{aligned} \quad (7)$$

It is possible to modify the activity of an electrode by the electrochemical deposition of an ion that acts as a promoter or inhibitor of the electrode for a reaction that occurs in the absence of an applied potential. This is not called electrocatalysis but nonfaradaic modification of catalytic activity. *See* OVERVOLTAGE.

Photocatalysis. Photocatalysis is a rather vague term, and in some cases it represents a nonstandard usage of catalysis. Photocatalysis should not be used to describe the use of photons for catalyst preparation; for example, the photoreduction of metal ions to form or deposit colloidal metal particles on a support is not photocatalysis. An appropriate usage is given in the following example. Water is irradiated with light of $\nu = 320$ nanometers. No light is adsorbed so nothing happens. Fine particles of titanium dioxide are now suspended in the water. Light is adsorbed and hydrogen and oxygen are evolved. It is demonstrated that the titanium dioxide is not consumed; that is, it may be recovered and reused in this reaction. This sequence is called photocatalysis. However, some chemists would distinguish this kind of catalysis, where the photon is one of the reactants, from a normal endothermic, catalytic reaction operated isothermally with heat supplied from the surroundings. The distinction is thermodynamic. In the latter case, energy is required in a First Law sense (conservation of energy), but the equilibrium of the reaction is not affected—the normal expectation of catalysis and a corollary of the definition of catalysis. Photocatalysis allows a reaction that would not occur spontaneously, for example, the dissociation of water into hydrogen and oxygen at room temperature and pressure.

Selectivity. In most cases, a given set of reactants could react in two or more ways, as exemplified by reactions (5) and (6). The degree to which just one of the possible reactions is favored over the other is called selectivity. The fraction of reactants that react by a specific path, for example, reaction (5), is called the selectivity by that path, and it will vary from catalyst to catalyst. Selectivity is a key property of a catalyst in any practical application of the catalyst.

Turnover. The number of sets of molecules that react consequent to the presence of a catalytic site (heterogeneous or enzyme catalysis) or molecule of catalyst (homogeneous catalysis), the turnover number is substantially greater than unity and may be very large. The turnover frequency is the number of sets of molecules that react per site or catalyst molecule per second. It must be greater than one or the reaction is stoichiometric, not catalytic.

Catalyst life. In practical applications, catalyst life, that is, the time or number of turnovers before the reaction rate becomes uselessly low, is important. It will be reduced by the presence of molecules that adsorb at (react with) and block the active site (poisons). *See* ADSORPTION.

Robert L. Burwell, Jr.; Gary L. Haller

Bibliography. B. C. Gates, *Catalytic Chemistry*, 1992; H. Gerischer and W. Vielstich, *Electrocatalysis*, in *Handbook of Heterogeneous Catalysis*, vol. 3, ed. by G. Ertl, H. Knözinger, and J. Weitkamp, 1997; E. Pelizzetti and N. Serpone (eds.), *Photocatalysis: Fundamentals and Applications*, 1990.

Catalytic antibody

A type of large protein naturally produced by the immune system with the capability of initiating diverse chemical reactions similarly to enzymes. Catalytic antibodies are elicited against small molecules that are bound to carrier proteins and contain a specific binding site. In their native form, they are constructed of two pairs of polypeptide chains that differ in length and are connected to each other by disulfide bridges. Various antibody molecules share a common structure (**Fig. 1**), but they differ in the N-terminal regions of antibody light and heavy chains which are responsible for antigen recognition. These regions vary greatly in the sequence and number of their constituent amino acids and therefore provide an enormous diversity of antigen-binding domains. *See* ANTIBODY; ANTIGEN; ANTIGEN-ANTIBODY REACTION; ENZYME; IMMUNITY; PROTEIN.

Description. In principle, catalysis is achieved by lowering the free energy of activation for a chemical reaction. Therefore, the catalyst has to bind the transition state more tightly then either the reactant or the product. If an antibody could be elicited against a transition state, it would have the potential to catalyze this chemical reaction. *See* CATALYSIS.

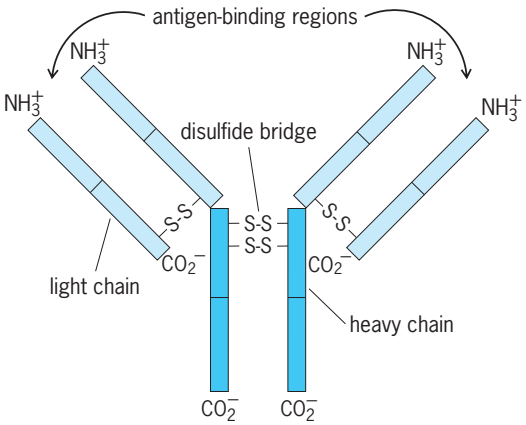

Fig. 1. Schematic structure of an antibody molecule. The protein consists of two heavy chains and two light chains.

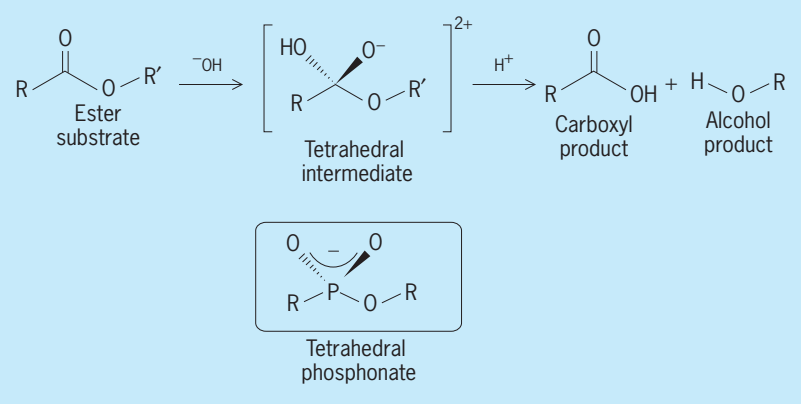

Fig. 2. Early highlights of the first antibody-catalyzed reaction of ester hydrolysis. Stable tetrahedral phosphonate was used as the transition-state analog.

Catalytic antibodies are produced according to standard monoclonal technology. Stable transition-state analogs (haptens) are synthesized and coupled to two carrier proteins. One such conjugate is used for mouse immunization, and the second for screening purposes. The carrier proteins trigger strong immune responses, and specific antibodies which bind only to the hapten are selected and screened for catalysis. *See* MONOCLONAL ANTIBODIES.

Catalytic antibodies were first reported in 1986 independently by Richard Lerner and Peter Schultz. Both used antibodies that were elicited against tetrahedral phosphonate as transition-state analog to catalyze hydrolysis reactions of aryl esters and carbonates (**Fig. 2**).

Since then, several hundred different examples of catalytic antibodies have been reported from various research groups, including highly enantioselective, disfavored, and nonenzymatic reactions. However, the transition-state analog immunization strategy produced antibodies that were highly specific for their substrates, and therefore only a very limited number of compounds could be used for antibody reactions.

A major breakthrough has been achieved with the development of a new immunization concept. Instead of immunizing against transition-state analogs, immunization is done with a highly reactive compound in order to create a chemical reaction during the binding of the antigen to the antibody. The same reaction becomes part of the mechanism of the catalytic event. In other words, the antibodies are elicited against a chemical reaction. This strategy is termed reactive immunization. *See* REACTIVE INTERMEDIATES.

1,3-Diketone was used as a trap for an amino lysine residue in an antibody active site (**Fig. 3**). Two antibodies which contained the desired lysine were found to mimic type I aldolases very efficiently. Type I aldolases are enzymes that catalyze reversible aldol reactions through the intermediacy of enamine formation with an enzyme amino group:

$$R{-}CH{=}O + R'CH_2C{=}OR'' \rightleftharpoons R{-}CHOH{-}CHR'{-}C{=}OR''$$

For the first time, a catalytic antibody was capable of accepting a very broad range of substrates. Furthermore, in most cases, the antibody's reactions were highly enantioselective and could be performed easily on a preparative scale. The synthetic advantages of these unique antibodies were clearly demonstrated, and include organic enantioselective synthesis of natural products and preparation of both enantiomers of a variety of aldol products. The surprising success of this aldolase catalyst made it the first commercially available catalytic antibody. *See* ORGANIC SYNTHESIS; STEREOCHEMISTRY.

Synthetic achievements. The early skepticism of the synthetic capability of catalytic antibodies has rapidly diffused. The first large-scale reaction with a catalytic antibody was reported in 1994, in which the synthesis of 1.4 g of enantiomerically pure

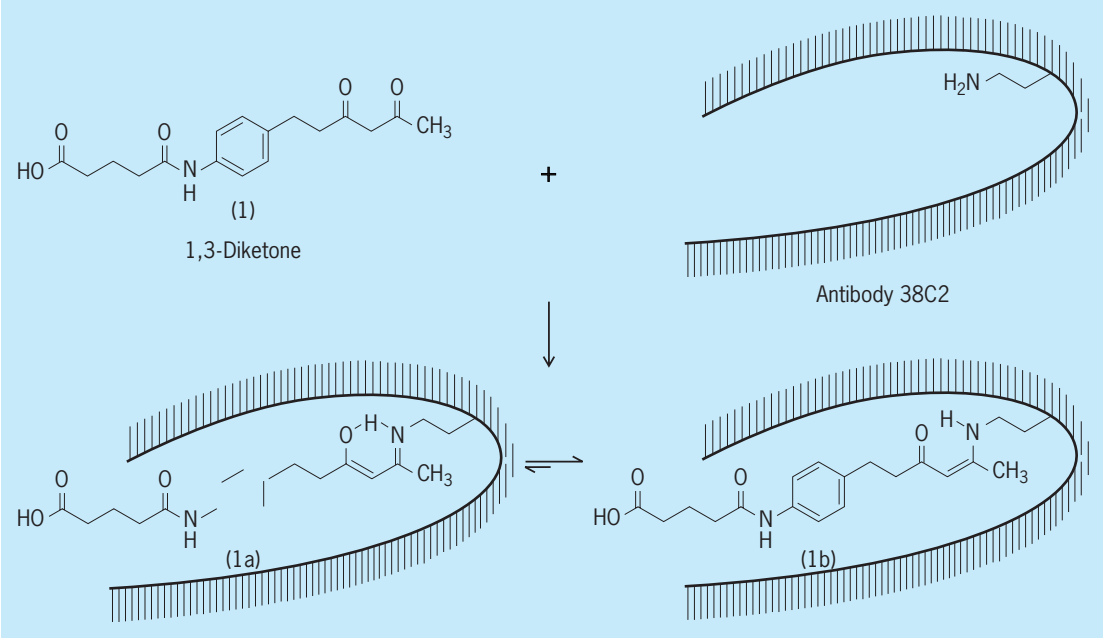

Fig. 3. Mechanism of trapping the essential ε-amino group of a lysine residue in the antibody's binding pocket by using the 1,3-diketone hapten.

enol ether was achieved. Just a year later, the first total synthesis of a natural product using a catalytic antibody was reported. Aldolase antibody 38C2 has accomplished several important achievements in organic synthetic chemistry. It started with enantioselective synthesis of the important synthetic block, the Wieland-Miescher ketone, and followed with the enantioselective total synthesis of 10 different brevicomins (sex pheromones of the bark beetle). A total synthesis of the sex pheromone frontalin was reported shortly after. *See* PHEROMONE.

The most stimulating achievement in the synthesis of a natural product by an antibody catalyst has been accomplished by the total synthesis of Epothilone, in which an important chiral synthon was prepared using antibody 38C2 to resolve 18 g of racemic mixture to 9 g of enantiomerically pure compound. Furthermore, the antibody was shown to perform aldol addition reactions very efficiently with unprotected hydroxyacetone as a donor. This is remarkable in that no other catalyst, chemical or biological, is capable of using hydroxyacetone as a donor substrate for the aldol reaction. The advantage of this reaction was elegantly demonstrated in the short total synthesis of the biologically important sugar 1-deoxy-L-xylulose.

Medical potential. The intriguing concept of using catalytic antibodies as therapeutic agents became more appealing when it was shown that most of the amino acids in a mouse antibody molecule could be replaced with human sequences and thereby make the antibody molecule compatible for in vivo treatment in humans. This strategy has been exploited within the application of prodrug activation by a catalytic antibody.

In principle, a catalytic antibody can be designed to modify a prodrug into an active drug. Several groups reported prodrug activation by antibody catalysis. All of them used phosphonate hapten as tetrahedral transition-state analog for the hydrolytic reaction which released the free drug.

The aldolase antibody 38C2 has a major advantage regarding prodrug activation. Since it has the capability to accept a broad variety of substrates, the antibody may potentially activate any prodrug. A general prodrug chemistry was developed to take advantage of the broad scope and mechanism of catalytic antibody 38C2. The drug masking/activation concept was based on a sequential retro-aldol-retro-Michael reaction catalyzed by antibody 38C2 (**Fig. 4**).

This reaction is not known to be catalyzed by any other enzyme and has a very low background; that is, the reaction is very slow in the absence of the catalyst. This chemistry was applied to the anticancer drugs doxorubicin and camptothecin. Weakly or nontoxic concentrations of their corresponding prodrugs can be activated by therapeutically relevant concentrations of antibody 38C2 to kill colon and prostate cancer cell lines. To further test the therapeutic relevance of this model system, it was shown that antibody 38C2 remained catalytically

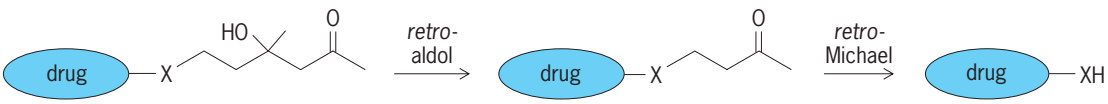

Fig. 4. Prodrug activation via a tandem retro-aldol-retro-Michael reaction. X stands for heteroatoms N, O, or S (nitrogen, oxygen, or sulfur).

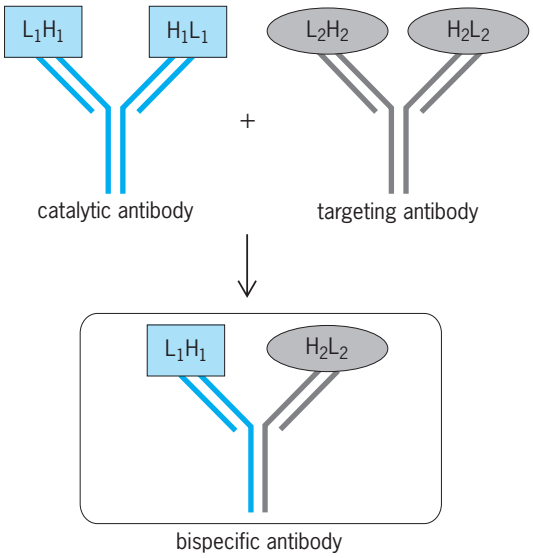

Fig. 5. Generation of bispecific antibody originating from a catalytic antibody and a targeting antibody.

active over weeks after intravenous injection into mice. Based on these findings, it is possible that the system described here has the potential to become a key tool in selective chemotherapy. *See* CHEMOTHERAPY.

The development of strategies that provide for selective chemotherapy presents significant multidisciplinary challenges. Selective chemotherapy, in the case of cancer, might be based on the enzymatic activation of a prodrug at the tumor site. The enzymatic activity must be directed to the site with a targeting molecule, usually an antibody, that recognizes a cell surface molecule selectively expressed at the tumor site. Since a single molecule of enzyme catalyzes the activation of many molecules of prodrug, a localized and high drug concentration may be maintained at the tumor site. This concept of antibody-directed enzyme prodrug therapy (ADEPT) holds promise as a general and selective chemotherapeutic strategy if several criteria can be met.

A number of antigens that are expressed on the surface of tumor cells or in their supporting blood vessels (vasculature) have been shown to be effective targets for antibody-mediated cancer therapy. Thus, for the most part, the targeting antibody component of this strategy is not limiting. By contrast, the requirements of the enzyme component and complementary prodrug chemistries for ADEPT are difficult to achieve. First of all, selective prodrug activation requires the catalysis of a reaction that must not be accomplished by endogenous enzymes in the blood or normal tissue of the patient. Enzymes of nonhuman origin that meet these needs are, however, likely to be highly immunogenic, a fact that makes repeated administration impossible. Finally, the chemistry used to convert a drug into a prodrug should be versatile enough to allow for the modification of many drug classes while not interfering with the operation of the enzyme so that a single enzyme could be used for the activation of a multiplicity of prodrugs.

Outlook. The limitations of the ADEPT complex encourage scientists to suggest that the enzyme component for ADEPT might be replaced by a catalytic antibody. The potential of catalytic antibodies for ADEPT are indeed compelling; both catalysis of reactions not catalyzed by human enzymes and minimal immunogenicity through antibody humanization are feasible. Combining these features, the ADEPT conjugate translates into a bispecific antibody consisting of targeting and catalytic arms (**Fig. 5**). Two parent antibodies are combined in a manner such that each antibody contributes one light chain and one heavy chain. The bispecific construct contains two different binding regions. One originates from the targeting antibody and can bind specifically to antigens which are expressed on tumor cells, and the other originates from a catalytic antibody and is used to activate a prodrug.

A bispecific antibody could be used for selective chemotherapy (**Fig. 6**). This possible future treatment can consist of two steps. First, a dose of the bispecific antibody can be administered, and the targeting arm can locate and attach to specific antigens on tumor cell surface. Excess of the antibody can be cleared after a limited time from the patient's blood, preventing nonspecific prodrug activation. Second, several doses of prodrug can be administered on suitable time gaps. The prodrug can reach the tumor site through the blood circulation and can be activated by the catalytic arm very close to the tumor cell. The damage to noncancer cells should be therefore minimized, and the free drug can specifically target cancer cells.

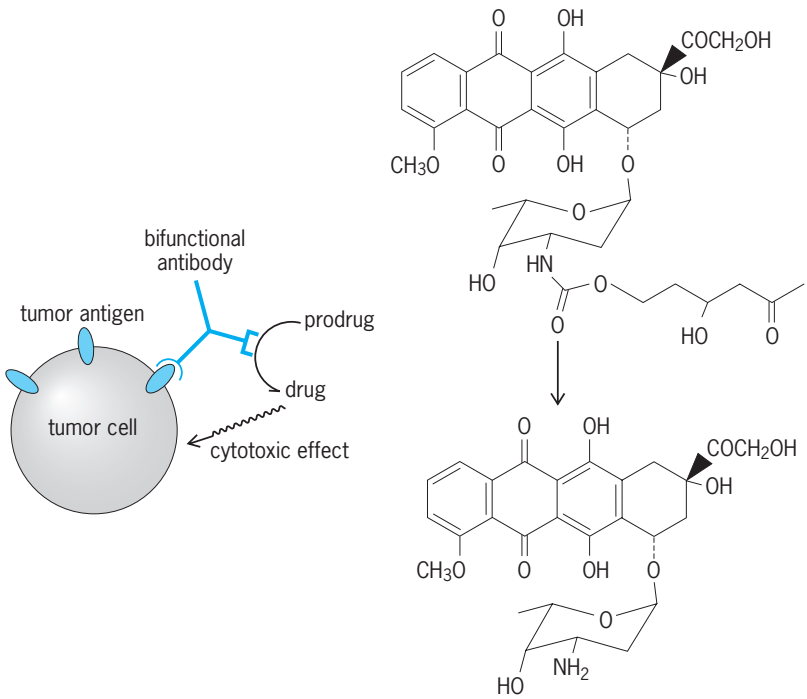

Fig. 6. Chemotherapy targeted specifically to a tumor cell. The prodrug is prodoxorubicin, which is transformed to the active drug doxorubicin by the catalytic arm of the bifunctional antibody. The other arm binds to an antigen on a tumor cell.

Catalytic antibodies are useful synthetic tools in enantioselective synthesis and also play an important role in performing organic reactions with unprotected chemical functionalities. However, the most promising use for them is probably in the medicinal field. The previously described strategy of targeting cancer with selective chemotherapy is highly encouraging. Richard A. Lerner; Doron Shabat

Bibliography. G. M. Blackburn et al., Passive and catalytic antibodies and drug delivery, *Pure Appl. Chem.*, 76(5):983–989, 2004; T. Hoffmann et al., Aldolase antibodies of remarkable scope, *J. Amer. Chem. Soc.*, 120:2768–2779, 1998; B. List et al., A catalytic enantioselective route to hydroxy-substituted quaternary carbon centers: Resolution of tertiary aldols with a catalytic antibody, *J. Amer. Chem. Soc.*, 121:7283–7291, 1999; B. List et al., Enantioselective total synthesis of some brevicomins using aldolase antibody 38C2, *Chem. Eur. J.*, 4(5):881–885, 1998; H. Miyashita et al., Prodrug activation via catalytic antibodies, *Proc. Nat. Acad. Sci. USA*, 90(11):5337–5340, 1993; P. G. Schultz and R. A. Lerner, From molecular diversity to catalysis: Lessons from the immune system, *Science*, 269(5232):1835–1842, 1995; D. Shabat et al., Multiple event activation of a generic prodrug trigger by antibody catalysis, *Proc. Nat. Acad. Sci. USA*, 96:6925–6930, 1999; S. C. Sinha, C. F. Barbas III, and R. A. Lerner, The antibody catalysis route to the total synthesis of epothilones, *Proc. Nat. Acad. Sci. USA*, 95:14603–14608, 1998; J. Wagner, R. A. Lerner, and C. F. Barbas III, Efficient aldolase catalytic antibodies that use the enamine mechanism of natural enzymes, *Science*, 270(5243):1797–1800, 1995.

Catalytic converter

An aftertreatment device used for pollutant removal from automotive exhaust. Since the 1975 model year, increasingly stringent government regulations for the allowable emission levels of carbon monoxide (CO), hydrocarbons (HC), and oxides of nitrogen (NO_x) have resulted in the use of catalytic converters on most passenger vehicles sold in the United States. The task of the catalytic converter is to promote chemical reactions for the conversion of these pollutants to carbon dioxide, water, and nitrogen.

By definition a catalyst is an agent which promotes the rates at which chemical reactions occur, without affecting the final equilibrium as dictated by thermodynamics, but which itself remains unchanged. For automotive exhaust applications, the pollutant removal reactions are the oxidation of carbon monoxide and hydrocarbons and the reduction of nitrogen oxides. Metals, base and noble, are the catalytic agents most often employed for this task. Small quantities of these metals, when present in a highly dispersed form (often as individual atoms), provide sites upon which the reactant molecules may interact and the reaction proceed.

In addition to the active metal, the converter contains a support component whose functions include yielding structural integrity to the device, providing a large surface area for metal dispersion, and promoting intimate contact between the exhaust gas and the catalyst. Two types of supports are used: pellets and monoliths. The pelleted converter consists of a packed bed of small, porous, ceramic spheres whose outer shell is impregnated with the active metal. The monolith is a honeycomb structure consisting of a large number of channels parallel to the direction of exhaust gas flow. The active metals reside in a thin layer of high-surface-area ceramic (usually γ-alumina) placed on the walls of the honeycomb. In either system the support is contained in a stainless steel can installed in the exhaust system ahead of the muffler.

Two types of catalyst systems, oxidation and three-way, are found in automotive applications. Oxidation catalysts remove only CO and HC, leaving NO_x unchanged. An air pump is often used to add air to the engine exhaust upstream of the catalyst, thus ensuring an oxidizing atmosphere. Platinum and palladium are generally used as the active metals in oxidation catalysts. Three-way catalysts are capable of removing all three pollutants simultaneously, provided that the catalyst is maintained in a "chemically correct" environment that is neither overly oxidizing or reducing. To achieve this requires that the engine air-fuel ratio always be at, or very near, stoichiometry under all vehicle operating conditions. Feedback air-fuel ratio control systems are often used to satisfy this requirement. Platinum, palladium, and rhodium are the metals most often used in three-way catalysts. In addition, base metals are frequently added to improve the ability of the catalyst to withstand small, transient perturbations in air-fuel ratio. In both oxidation and three-way catalyst systems, the production of undesirable reaction products, such as sulfates and ammonia, must be avoided.

Maintaining effective catalytic function over long periods of vehicle operation is often a major problem. Catalytic activity will deteriorate due to two causes, poisoning of the active sites by contaminants, such as lead and phosphorus, and exposure to excessively high temperatures. Catalyst overtemperature is often associated with engine malfunctions such as excessively rich operation or a large amount of cylinder misfire. To achieve efficient emission control, it is thus paramount that catalyst-equipped vehicles be operated only with lead-free fuel and that proper engine maintenance procedures be followed. In such cases catalytic converters have proved to be a very effective means for reducing emissions without sacrificing fuel economy. *See* AUTOMOTIVE ENGINE; CATALYSIS. Norman Otto

Cataract

Any clouding or opacity of the crystalline lens of the eye. Cataracts vary markedly in degree of density and may be due to many causes, but the majority are associated with aging. Cataracts are the single leading cause of blindness in the world.

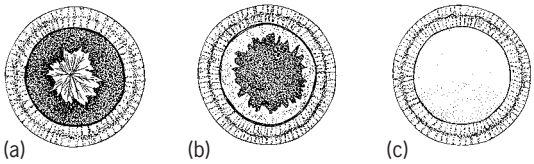

Types of cataracts: (*a*) traumatic "star-shaped," (*b*) senile "cuneiform," and (*c*) senile "morgagnian."

The lens consists of fibers arranged in regular layers, similar to the structure of an onion. It is composed of three distinct parts: an elastic capsule, the epithelium from which the lens fibers originate, and the lens substance which is plastic and capable (below the age of 40) of altering shape according to the condition of the capsule. The average diameter of the lens is 0.36 in. (9 mm), and it has a greater radius of convexity on its posterior surface. This lens is supported by the suspensory ligament and lies between the back of the iris and the anterior base of the vitreous, which is the jelly that fills the space behind the lens and in front of the retina. The lens has no blood supply, so it cannot be inflamed, but it is easily affected by metabolic changes. The essential biochemical change in a cataractous lens is the coagulation of its protein.

Senile cataract occurs with aging and is by far the most common type. Progressively blurred vision is the only symptom. There are a number of other varieties (congenital, metabolic, secondary, or traumatic; see **illus.**) and causes (such as a reaction to certain drugs or irradiation). Cataracts may also be associated with systemic diseases such as hypoparathyroidism, myotonic dystrophy, atopic dermatitis, galactosemia, and Lowe's, Werner's, and Down syndromes.

At present the treatment for cataracts, when they are sufficiently advanced to impair the vision, is surgical removal. When a cataract is surgically removed, the crystalline lens of the eye is removed much as a lens would be removed from a camera. Cataract surgery may be performed in one of three ways: (1) Intracapsular extraction: the lens, enclosed in its complete capsule, is entirely removed through a large 0.48–0.56-in. (12–14-mm) incision in the eye; this remains the most successful and common method of cataract removal in the United States. (2) Extracapsular cataract extraction: the capsule is opened, and the lens contents are removed; the posterior capsule is left intact, while the anterior capsule is partially or totally removed. (3) Ultrasound cataract surgery or phacoemulsification: an ultrasonic needle is introduced into the lens, and the lens is emulsified, fragmented, or broken up and sucked out through a small (0.12-in. or 3-mm) opening in the eye; this is a form of extracapsular cataract extraction since the anterior capsule is opened and removed along with the contents, leaving the posterior capsule intact.

In order to restore normal vision after surgery, any of a number of methods may be followed. Cataract glasses may be prescribed. Contact lenses may be fitted, but these have not been too successful in elderly cataract patients because of their inability to handle lenses. However, since soft contact lenses have been approved for extended wear, they can be left in the eye for weeks or even months, thus obviating the need for handling. Intraocular lenses made of the plastic polymethylmethacrylate can be inserted into the eye to replace the cataractous lens. There are numerous types of intraocular lenses. Their theoretical advantages include permanent placement and less distortion and magnification. Some of the disadvantages include the increased difficulty of the cataract operation, an increased incidence of operative and postoperative complications, and the inability to easily remove a lens of incorrect power. *See* EYE (VERTEBRATE).

Jack Hartstein

Bibliography. O. R. Caldwell (ed.), *Cataracts*, 1988; E. L. Greve (ed.), *Surgical Management of Coexisting Glaucoma and Cataract*, 1987; J. Schulman (ed.), *Cataracts*, 1993; R. W. Young, *Age-Related Cataract*, 1990.

Catastrophe theory

A theory of mathematical structure in which smooth continuous inputs lead to discontinuous responses. Water suddenly boils, ice melts, a building crashes to the ground, or the earth unexpectedly buckles and quakes. The French mathematician René Thom conceived and developed an eclectic collection of ideas into catastrophe theory. His idea was to establish a new basis for a more mathematical approach to biology. Connotations of disaster are misleading, since Thom's intention was to emphasize sudden, abrupt changes.

Advanced areas of modern mathematics, including algebraic geometry, differential topology, and dynamical system theory, contributed to the creation of catastrophe theory. A complete mathematical theory exists for the elementary catastrophes, which can be written as the gradient of an energylike function. The physical, chemical, and engineering applications are less developed, although many are known in optics, laser theory, thermodynamics, elasticity, and chemical reaction theory. The Thom classification theorem gives exactly seven elementary catastrophes. Although a theory of generalized catastrophes exists, which extends the theory beyond gradient systems, it is not nearly as well developed mathematically or physically as that of elementary catastrophes. It does include remarkable examples of chaos (or stochastic behavior) in the solutions to nonlinear deterministic equations. These solutions include strange attractors and omega explosions among the examples of non-elementary catastrophes. *See* GEOMETRY; PERIOD DOUBLING; TOPOLOGY.

Two important aspects of catastrophe theory are frequently overlooked or misconstrued. First, as it is a rigorous mathematical theory, the characteristic catastrophe features can be proved. These features include jumps in the response; hysteresis or a path dependence in the response, representing a storage of energy for some paths; divergence, where a small path change produces a large response change (as if a source or sink were crossed); and type changes in

the response, where a smooth response occurs on one path that becomes discontinuous along a nearby path.

Second, all of these features are topological, so that they are independent of the coordinates used to describe the potential. They are, therefore, qualitative features of the solutions. Some critics have concluded that because these aspects were qualitative, they could not be quantitative. This is contradicted by the solid and growing body of quantitative studies in catastrophe theory. (Problems in quantum optics, thermodynamics, and scattering theory have all been clarified by catastrophe theory.)

Definitions. The object V whose possible behavior is to be classified is a smooth, real function, having the general form given by Eq. (1), the set of real numbers, or, equivalently, Eq. (2). The r real variables

$$V : R^r \times R^n \mapsto R \tag{1}$$

$$V_{a_1,a_2,\ldots,a_r}(x_1,\ldots,x_n) = Va_1,\ldots,a_r(x_1,\ldots,x_n) \tag{2}$$

a_1, a_2, ..., a_r are called the control parameters; the n variables x_1, ..., x_n are the variables; and V is the response function. In mechanics, V represents a potential energy surface, and the control parameters are coupling constants, that is, the strengths of various terms in the expression for the potential energy. For a two-dimensional potential function, $V_{abc}(x_1,x_2)$, the response is the height of hills and valleys in a perpendicular x_3 direction. If a ball (with negligible inertia) is rolled onto this surface, it will roll into the lowest possible position. This lowers its potential energy to a local minimum. This is a simple example of a critical point. *See* ENERGY; PARAMETER; POTENTIALS.

In general, the critical points of $V_{a1},\ldots,_{ar}(x_1,\ldots,x_n)$ are the set of all points $P_i = (x_{1i},x_{2i},\ldots,x_{ni})$ for which Eq. (3) is satisfied, and may be maxima,

$$\left(\frac{\partial V_{a_1,\ldots,a_r}}{\partial x_j}\right)\bigg|_{P_i} = 0 \tag{3}$$

minima, or inflection points. The hessian matrix **H** of the function $V_{a2},\ldots,a_r$ is the $n \times n$ matrix with ijth element given by Eq. (4). A point x_0 is called structurally stable if it is a critical point and satisfies inequality (5).

$$h_{ij} = \frac{\partial^2 V_{a_1,\ldots,a_r}}{\partial x_i \partial xj} \tag{4}$$

$$\det (H)_{x_0} \neq 0 \tag{5}$$

If S is a space of points with boundary ∂S, its codimension is dimension (S) – dimension (∂S). Two subspaces V_1, V_2 of S are transverse if dimension (V_1) + dimension (V_2) = dimension (S). The Morse lemma and splitting lemma allow the number of variables considered to be reduced. The codimension is the number of variables which must be studied.

Thom classification theorem. Consider an r-parameter family of smooth real functions $V : R^r \times R^n \to R$, for any value of n and any value of $r \leq 5$. The Thom classification theorem states that either the family is structurally stable, and is equivalent about any point to noncritical or a nondegenerate critical point (these are not catastrophes); or it is structurally unstable and is equivalent to one of the seven elementary catastrophes: (1) the fold; (2) the cusp; (3) the swallowtail; (4) the butterfly; (5) the elliptic umbecile; (6) the hyperbolic umbecile; or (7) the parabolic umbecile. If the number of parameters is finite and greater than five, complicated combinations of these seven catastrophes can occur. If r is made very large, the elementary class of catastrophes is not complete. Catastrophe geometry becomes increasingly complicated both as n increases and as the numbers of possible catastrophes increase.

One-variable cusp catastrophe. A simple example of a catastrophe is the one-variable cusp catastrophe. The response function for this catastrophe, $V_{ab}(x)$, is given by Eq. (6). A sketch of Eq. (6) for $x = x_0$, a

$$V_{ab}(x) = {}^1/_4x^4 + {}^1/_2ax^2 + bx \tag{6}$$

critical point, is given in **Fig. 1**. Since the value of the critical point x_0 depends on the control parameters, a and b, in a complicated fashion, it follows that V_{ab}, evaluated at $x = x_0$, is a complicated nonlinear function of a and b. This is plotted in Fig. 1. Assuming that V_{ab} is a potential energy surface, a ball placed on the surface in Fig. 1 will roll downhill toward the right. The curves A, B, and C indicate three possible paths and, in **Fig. 2**, points 1 to 10 are labeled on these curves. Suppose further that points (1,3,5) and points (2,4,10) all have the same potential energy. This example will be used to illustrate each of the characteristic features of elementary catastrophes. It is the dependence of the response upon control parameters which is under study.

The curves A, B, and C of Fig. 1 are redrawn in Fig. 2. To see the jumps in response, consider a ball set in motion from point 5 on curve C. It will roll past point 6, and at point 7 it will fall to point 9

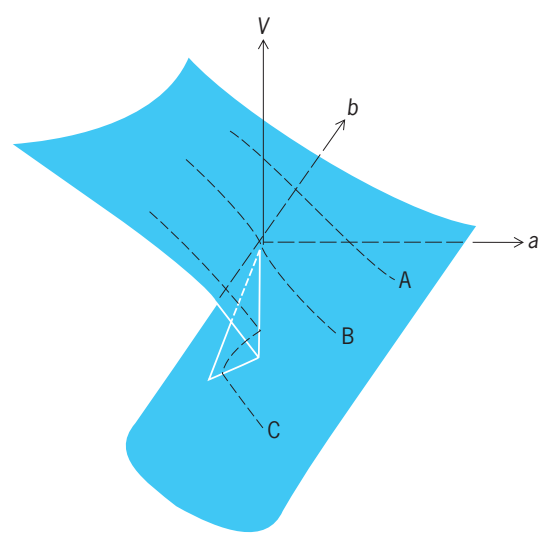

Fig. 1. Cusp catastrophe surface for the response function V_{ab} as a function of (a,b) at a critical point. Paths A, B, and C are labeled.

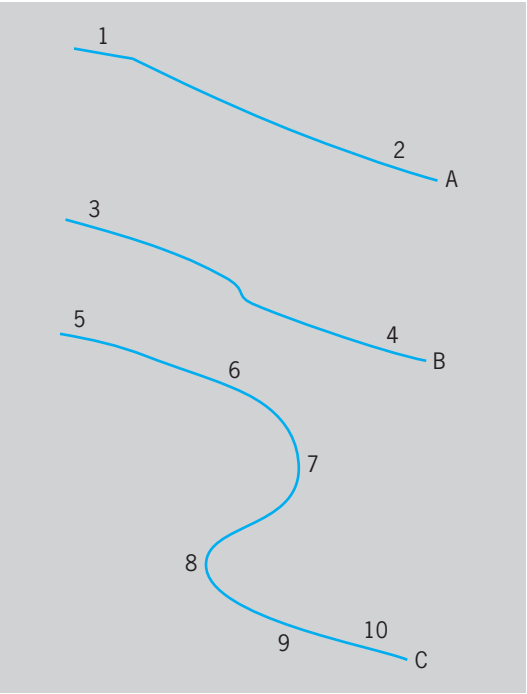

Fig. 2. Paths A, B, and C from Fig. 1, with points 1–10 marked. The response function V_{ab} has the same values at (1,3,5) and at (2,4,10).

(**Fig. 3***a*). In addition, it is impossible to go from point 10 to point 5 along curve C.

To exhibit the hysteresis effect (Fig. 3*b*), first roll a ball from point 2 to point 1 on curve A; from there to point 3 on curve B. Then roll it from point 3 to point 4 and back to the starting point 2. The area enclosed in the figure is a hysteresis effect and is quite similar to hysteresis in magnetism. *See* HYSTERESIS.

The divergence effect can be understood by considering a small counterclockwise circular path about the (*a, b*) origin. A ball which rolls in this path will see the hill at negative *b* as a source of energy. The ball will act as if it had received a divergence of potential (that is, a nonconservative force) on each trip.

To see the change in type of path, simply consider the difference in paths A and C. If the scale of the *b* control parameter is such that these paths are very "near" one another, then a small change in the *bx* term can dramatically change the path from a smooth and continuous path A to a discontinuous path C.

Applications. Many physical problems involve some or all of the four nonlinear effects just described. Phase transitions include many aspects which are surprisingly similar to these four. The jump in the latent heat at a liquid-vapor phase transition is very similar to the jump shown in Fig. 3*a*. However, the cusp catastrophe given in Eq. (6) is quite definite. It contains no cx^3 term, as such effects are transformed away by coordinate transformations in obtaining the generic form. Unfortunately, careful experiments and their analysis show that the liquid-vapor phase transition is not any of the elementary catastrophes. On the other hand, the hysteresis of ferroelectrics is accurately described by the butterfly catastrophe. The conclusion is that catastrophe theory has an important role to play in phase transition physics even if it is not the whole story. *See* FERROELECTRICS; PHASE TRANSITIONS.

In ecology, sharp boundaries are often important. A predator may eliminate some competing species from an area; a parasite may eliminate the food supply of another. An explanation can be obtained from curve C in Fig. 2. Along the segment 5 to 6, the conditions assure the dominance of species A. Point 7 is a catastrophe point. The segment 8-9-10 is a region where another species, species B, dominates. This simple model is, of course, nonunique. Any other model whose solution is approximately such a polynomial is an equally good candidate. Thus, some collaboration from ecology is required to determine the correct response function, but the model can help in finding this. After several successful iterations of this kind, the surviving models have a sound basis. *See* ECOLOGY.

Status. There have been vigorous controversies, as some scientists have argued that catastrophe theory is unmathematical and scientifically empty, and that some applications are foolish. The first two claims are simply false. The first objection was based

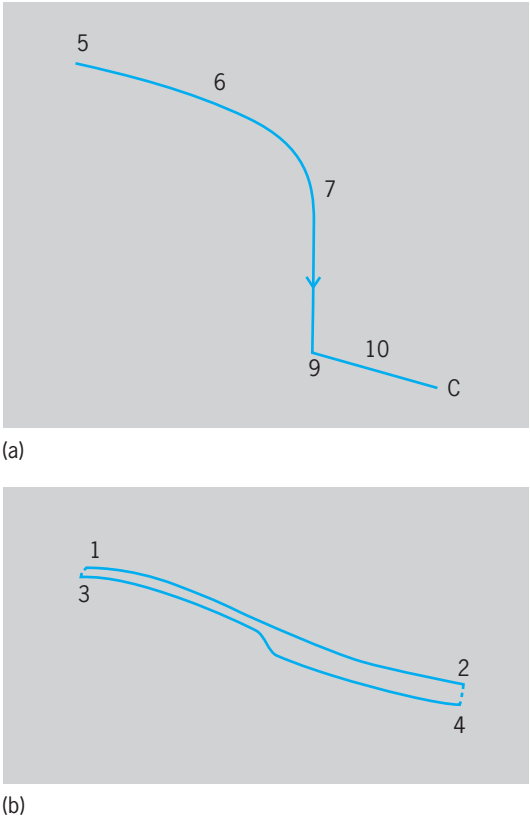

Fig. 3. Paths that exhibit features of catastrophe theory. (*a*) Path on curve C showing a jump in the response function. (*b*) Path using curves A and B which shows hysteresis. No change in potential energy V_{ab} is required to go from 1 to 3 or from 2 to 4. The area enclosed between curves is proportional to an energy stored from the overall process.

on the articles which were explicitly written to popularize the subject, while conveniently ignoring the scholarly articles which contained the mathematical proofs and examples. The second objection was based upon a very careless search of the literature; articles by many mathematicians meet the highest scientific standards. The third criticism is partially correct. Attempts to use simple models on drunkenness, prison disorders, and social problems are questionable. Even in these cases, however, the criticism seems harsh.

Catastrophe theory is an important tool for the physical sciences. If Thom's vision is correct, it may be even more important for biology and the social sciences. Brian DeFacio

Bibliography. M. V. Berry, *Adv. Phys.*, 25:1–26, 1976; R. Gilmore, *Catastrophe Theory for Engineers and Scientists*, 1982, reprint 1999; T. Poston and I. Stewart, *Catastrophe Theory and Its Applications*, 1978; H. T. Sussmann, *Synthesis*, 31:229–270, 1975; R. Thom, *Structural Stability and Morphogenesis*, 1975; E. C. Zeeman, Catastrophe theory, *Sci. Amer.*, 234(4):65–83, April 1976; E. C. Zeeman, *Catastrophe Theory: Selected Papers (1972–1977)*, 1977.

Catenanes

Catenanes are compounds that are made of interlocking macrocycles. The interlocking rings are said to be mechanically rather than chemically bound. Catenanes are named according to the number of interlocking rings. The simplest catenane, containing two interlocking rings, is called [2]-catenane (structure **1**). The most common catenanes con-

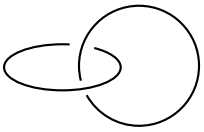

(**1**)

sist of a linear arrangement of interlocking rings, so [2]-catenane is the first member of this series. Next is [3]-catenane (**2**). Linear catenanes containing up to

(**2**)

five rings are known, for example, olympiadane (**3**).

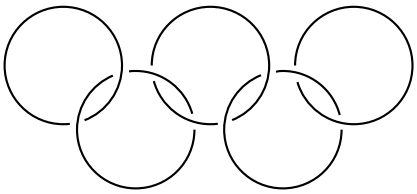

(**3**)

Catenanes involving necklace arrangements of bead-like rings have also been synthesized: [5]-catenane (**4**) represents the topology of this series. [7]-Catenane (**5**) combines linear and necklace topolo-

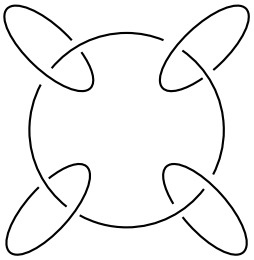

(**4**)

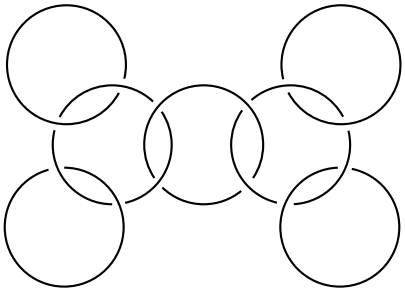

(**5**)

gies; it is the highest-order molecular catenane isolated so far. In fact, catenanes need not be made of simple macrocycles only. Macrobicycles or higher-order macrocycles may be involved too. Examples of catenanes constructed from such molecules are [4]-catenane (**6**), which is based on a macrobicyclic

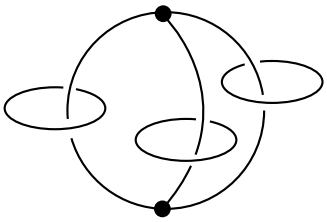

(**6**)

backbone, each lobe being interlocked with a macrocycle, and three- dimensionally interlocked catenane (**7**), where two macrobicycles are mutually inter-

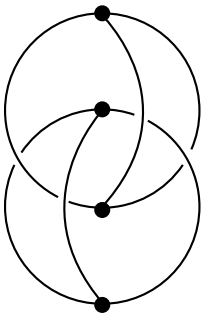

(**7**)

locked. A remarkable class of catenanes, from the

topological viewpoint, are the multiply interlocked catenanes. The first member of this series is the doubly interlocked [2]-catenane (**8**), made of inter-

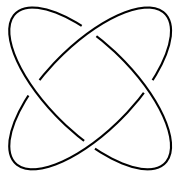

(**8**)

woven macrocycles. It is topologically more complex than [2]-catenane (1), since the projection of the molecule on the plane has four crossing points whereas that of [2]-catenane has only two crossing points.

Many other catenane topologies can be envisaged. An exciting target is the so-called Borromean rings (**9**), a figure used by a famous Italian family as a sym-

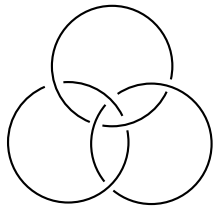

(**9**)

bol of solidarity. There are three interlocked rings, yet any two rings are independent of each other; it is the third ring that fastens the system. Another target is the figure-eight [2]-catenane (**10**).

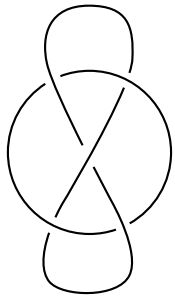

(**10**)

Occurrence. Molecular catenanes are compounds of synthetic origin. However, natural deoxyribonucleic acid (DNA) macromolecules were shown to assemble, in certain conditions, into catenated structures, which were studied by electron microscopy.

Synthesis. The rings of the catenanes may be purely organic macrocycles or metallomacrocycles, that is, macrocycles including transition-metal ions in their bond sequences. Actually, the chemical nature of the ring is dictated by the method of synthesis. Basically, three methods have been developed for the synthesis of catenanes (**Fig. 1**): statistical, directed, and template syntheses.

Statistical synthesis. Statistical synthesis relies on the probability that a macrocycle can be threaded onto a molecular string to afford an intermediate that will undergo the cyclization reaction (Fig. 1*a*). Convincing approaches to the statistical method used the trick of stabilizing the threaded complex by stoppering the extremities of the string with bulky groups: a so-called rotaxane species is obtained (Fig. 1*b*). Subsequently, conventional macrocycle synthesis is used to prepare the catenane, which is obtained after removal of the stoppers. This method produced the first hydrocarbon catenane, made of interlocked $(CH_2)_{28}$ and $(CH_2)_{46}$ macrocycles.

Directed synthesis. Directed synthesis uses a catechol-based acetal incorporated in a macrocycle (Fig. 1*c*). Two pendent arms, as precursors to the second macrocycle, are anchored perpendicularly to the plane of the first macrocycle. Their functionalized extremities are compelled to react with a complementary function localized inside the first macrocycle, so that the construction of the second macrocycle is directed to take place inside the first one. The last key step is the cleavage of the bonds linking the two macrocyclic sequences of atoms. This multistep synthetic method was used to prepare [2]- and [3]-catenanes.

Template syntheses. Template synthesis methods are highly directed and economical in terms of numbers of steps. In these methods, metal cations or molecules gather and preorganize reactive molecular fragments in a spatially controlled manner, so that the desired structure will be obtained preferentially over many others. In the first method, the transition-metal-templated synthesis of catenanes, the metal [generally Cu(I)] gathers a macrocycle incorporating a chelating subunit and a linear fragment made up of the complementary chelate such that both components are more or less at right angles to each other (Fig. 1*d*). Cyclization of the linear fragment affords the metallocatenane, or catenate. Removal of the metal template by competitive complexation provides the catenane as a free ligand, or catenand. In the one-step alternative strategy, two open-chain chelates are assembled orthogonally at a metal center. A double cyclization reaction provides the catenate in one step. The transition-metal-templated synthesis of catenanes is very flexible in terms of chemical functionality; for example, catenands containing pendent metalloporphyrins have been synthesized. The yield of catenane formation depends on the cyclization reaction used; excellent yields were observed for cyclization methods based on the ring-closing metathesis reaction.

Fig. 1. Synthetic routes to catenanes. (*a*) Statistical; f and g are complementary functions that react to close the ring. (*b*) Statistical, using a rotaxane intermediate; f and g are complementary functions that react to anchor the stoppers; i and h are complementary functions that react to form the second, interlocked ring. (*c*) Directed. (*d*) Transition-metal-templated; the thick ring portions represent the coordination sites, and the black disk is the metal; f and g are complementary functions that react to close the second, interlocked ring. (*e*) Templated by aromatic π-stacking interactions; the open rectangles are electron-rich, and the filled rectangles are electron-poor, aromatic groups; f and g are complementary functions that react to form the electron-deficient macrocycle.

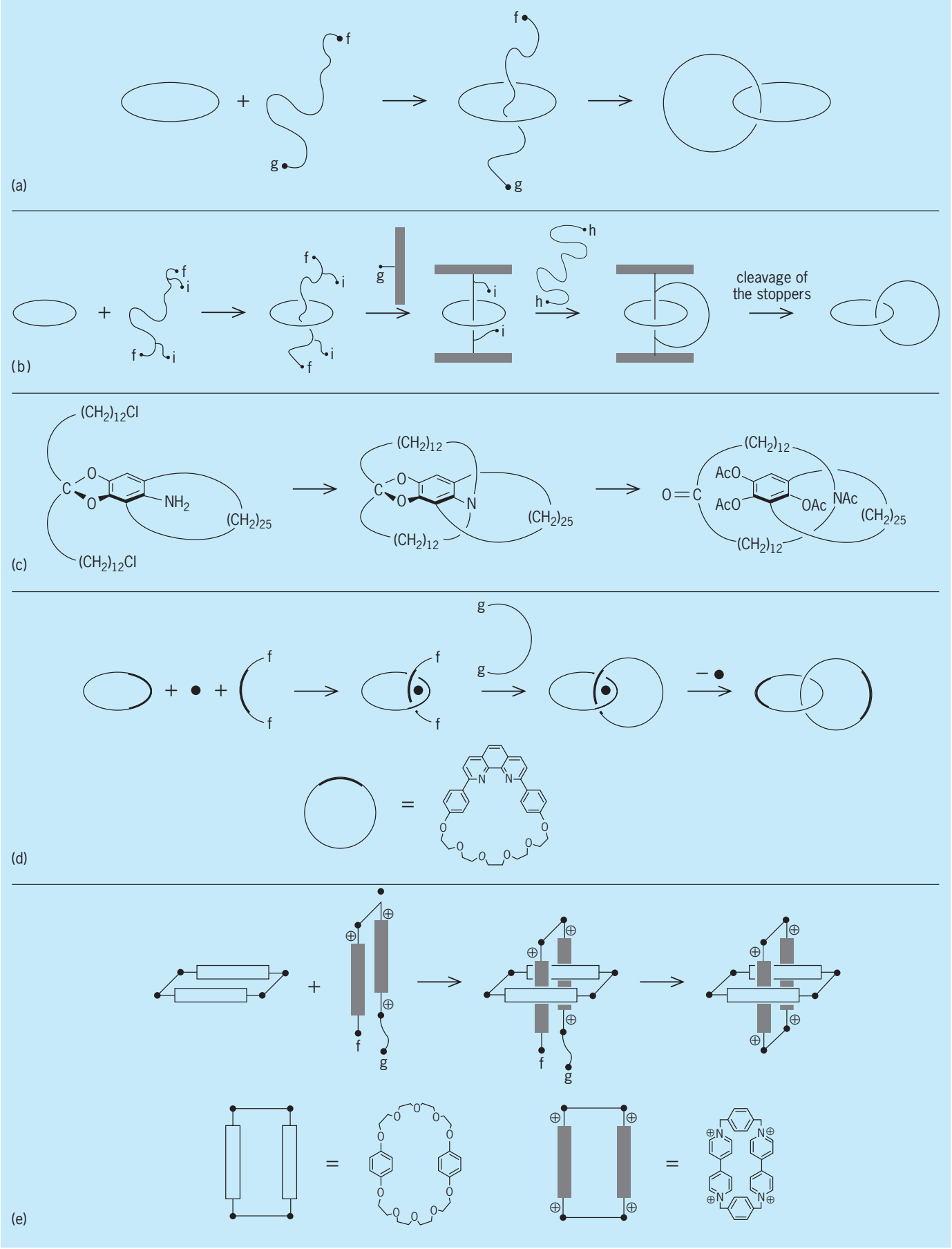
(a)
f
g
(b)
f
i
g
h
cleavage of
the stoppers
(c)
$(CH_2)_{12}Cl$
C
O
NH_2
$(CH_2)_{25}$
N
O=C
AcO
OAc
NAc
(d)
f
g
N
O
(e)
f
g
O
N

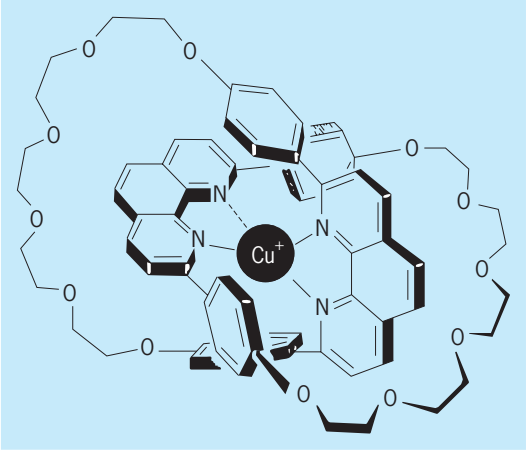

Fig. 2. Molecular structure of the copper(I) [2]-catenate of Fig. 1*d*, from an x-ray crystal structure analysis.

Other template syntheses involve template effects based on noncovalent bonds like van der Waals or aromatic π-stacking interactions, and hydrogen bonds. For example, the formation of a catenane based on π-stacking interactions between an electron-rich macrocycle and the precursor to an electron-deficient macrocycle; these interactions control the threading of the macrocycle onto the linear intermediate (Fig. 1*e*). Then cyclization can occur, affording the catenane. The principle of template methods involving molecular rather than metal templates is that, very schematically, one preformed macrocycle acts as a receptor for a molecular fragment of the other macrocycle. Moreover, in some cases the receptor macrocycle need not be made prior to the final cyclization step; self-assembly effects come into play. This was observed, in particular, in the case of catenanes assembled with the help of hydrogen bonds.

Characterization. Catenanes are characterized and studied like normal molecules, in solution or in the solid state, using routinely employed techniques, including elemental analysis, nuclear magnetic resonance (NMR) spectroscopy, mass spectrometry, osmometry, and single-crystal x-ray analysis. Because of the interactions between the macrocycles, which are compelled to stay in proximity, the NMR spectrum of the catenane species differs from the sum of the spectra of the individual, separated macrocycles. The fast atom bombardment (FAB) mass spectra of the catenanes show typical fragmentation patterns. For example, in the case of a [2]-catenane there is no ionized fragment between that of the mass of the parent catenane and that of the individual macrocycles, which are therefore the first observed fragments. This feature is considered a signature of catenated species. However, an x-ray crystal structure analysis, when possible, is the only technique that allows a direct view of the actual topology of the molecule, such as the molecular structure of the copper(I) [2]-catenate (**Fig. 2**).

Properties. Catenanes are new molecules as compared to the separated macrocycles. Some colligative properties, that is, macroscopic properties resulting from the fact that chemical compounds are molecular assemblies, have been measured. For example, it was demonstrated that the molar volume of a [2]-catenane was slightly larger than the sum of the molar volumes of the separated macrocycles. This unexpected result could be rationalized by considering solvent effects.

In addition to their topology, catenanes may show interesting stereochemical properties; interlocking of rings can lead to novel kinds of stereoisomerism. Topologically chiral catenanes are molecules whose chirality cannot be inverted by continuous deformation, unlike geometrically chiral molecules. Topologically chiral [2]-catenanes can be made by interlocking oriented rings (**11**) or by double interlocking

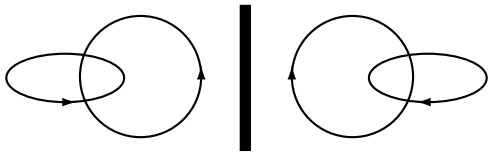

(**11**)

of nonoriented rings (8). They have been resolved by high-pressure liquid chromatography on optically active stationary phases. Isomerism produced by different relative orientations of nonsymmetrical rings (translational isomerism) has been described (**12–14**). Interestingly, these orientations can be

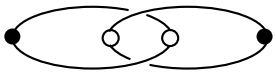

(**12**)

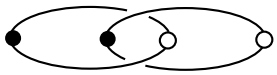

(**13**)

(**14**)

controlled by the redox state of the metal in certain metallocatenates.

Catenands, that is, coordinating catenanes, have very specific properties. As ligands, they complex first-row transition metals with high binding constants and enhanced kinetic stability, as compared to analogous open-chain chelates. This generalized macrocyclic effect has been described as the catenand effect. Most catenands synthesized and studied are tetradentate ligands that complex metals with a tetrahedral coordination sphere, like Li(I), Co(II), Cu(I), Zn(II), Ag(I), and Cd(II), but catenates of Cu(II), Ni(II), and Fe(II) have also been studied. In addition, catenands stabilize low-valent transitionmetal cations like Ni(I), and they are one of the strongest complexants of Cu(I). Catenands also show unusual acid-base properties, being stronger bases than their open-chain chelate analogues.

Applications. Applications of catenanes are a research field in its infancy, and therefore they are more or less speculative. Promising approaches

include incorporation of catenane structures into polymeric species to endow the polymers with peculiar rheological properties, because of the mechanical linking; and use of catenanes as elements of molecular machines—for example, the mechanical link of catenanes could be used for making a primitive rotary motor at the molecular level. *See* ALICYCLIC HYDROCARBON; ANALYTICAL CHEMISTRY; AROMATIC HYDROCARBON; MACROCYCLIC COMPOUND; PHYSICAL ORGANIC CHEMISTRY; STEREOCHEMISTRY.

Jean-Claude Chambron; C. O. Dietrich-Buchecker; J.-P. Sauvage

Bibliography. D. B. Amabilino and J. F. Stoddart, Interlocked and intertwined structures and superstructures, *Chem. Rev.*, 95:2725–2828, 1995; J.-C. Chambron, C. O. Dietrich-Buchecker, and J.-P. Sauvage, Transition metals as assembling and templating species: Synthesis of catenanes and molecular knots, in J.-M. Lehn (ed.), *Comprehensive Supramolecular Chemistry*, vol. 9, pp. 43–83, 1996; C. O. Dietrich-Buchecker and J.-P. Sauvage, Interlocking of molecular threads: From the statistical approach to the templated synthesis of catenands, *Chem. Rev.*, 87:795–810, 1987; M. Fujita, Self-assembled macrocycles, cages and catenanes containing transition metals in their backbone, in J.-M. Lehn (ed.), *Comprehensive Supramolecular Chemistry*, vol. 9, pp. 253–282, 1996; R. Jäger and F. Vögtle, A new synthetic strategy towards molecules with mechanical bonds: Non-ionic template synthesis of amide-linked catenanes and rotaxanes, *Angew. Chem. Int. Ed. Engl.*, 36:930–944, 1997; G. Schill, *Catenanes, Rotaxanes and Knots*, 1971.

Catenary

The curve formed by an ideal heavy uniform string hanging freely from two points of support. The lowest point A (**Fig. 1**) is the vertex. The portion AP is an equilibrium under the horizontal tension H at A, the tension F directed along the tangent at P, and the weight W of AP. If the weight of the string is w per unit length and s is the arc AP, $W = ws$; and from the force triangle, $\tan\psi = ws/H = s/c$, where $c = H/w$, is called the parameter of the catenary. Thus the catenary has the differential equation

$$dy/dx = s/c$$

The horizontal line at a distance c below the vertex A is the directrix of the catenary. With the x axis as directrix and the y axis through the vertex, the integration of the equation above yields

$$y = c\ \cosh\frac{x}{c} \qquad s = c\ \sinh\frac{x}{c}$$

for the ordinate and arc of the catenary. All catenaries are geometrically similar to the hyperbolic cosine curve, $y = \cosh x$. These equations show that $g^2 = s^2 + c^2$, and from the differential equation,

$$s = c\ \tan\psi \qquad y = c\ \sec\psi$$

The radius of curvature of the catenary at P is the

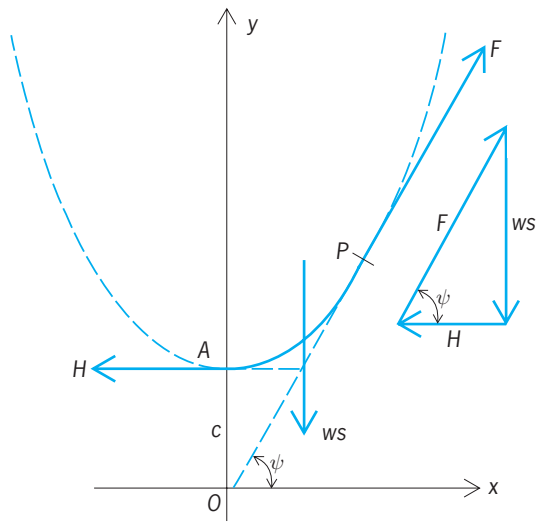

Fig. 1. Catenary and force triangle which keeps a portion of the string in equilibrium.

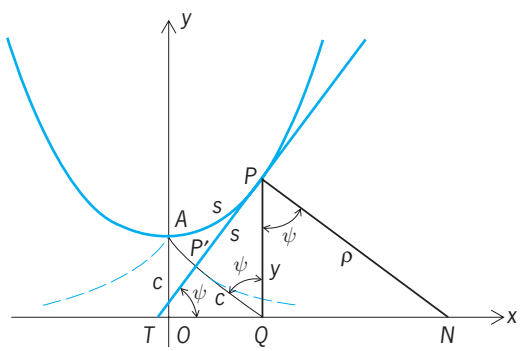

Fig. 2. Relation of catenary to the tractrix which is its involute.

length of the normal cut off by the x axis:

$$\rho = ds/d\psi = c\sec^2\psi = y\sec\psi = PN$$

The tension at P,

$$F = H\sec\psi = wc\sec\psi = wy$$

varies as the height above the directrix.

In **Fig. 2** the tangent PT is tangent to a circle of radius c about Q, the foot of the ordinate; and since $\tan\psi = s/c$, $P'P = s =$ arc AP. Thus arc AP unwraps on the segment $P'P$. The locus of P' is a curve whose tangent $P'Q$ has the constant length c, namely a tractrix; hence the tractrix is an involute of the catenary. *See* TRACTRIX.

The surface generated by revolving a catenary about its directrix is a minimal surface, the catenoid. The catenary is an extremal for the problem of finding a curve joining two given points so that the surface generated by revolving it about a given line has minimum area. *See* HYPERBOLIC FUNCTION.

Louis Brand

Bibliography. P. A. Laurent, A. LeMehaute, and L. L. Schumaker (eds.), *Curves and Surfaces*, 1991; D. H. von Seggern (ed.), *Handbook of Mathematical Curves and Surfaces*, 2d ed. 1992.

Cathode-ray tube

An electron tube in which a beam of electrons can be focused to a small cross section and varied in position and intensity on a display surface. In common usage, the term cathode-ray tube (CRT) is usually reserved for devices in which the display surface is cathodoluminescent under electron bombardment, and the output information is presented in the form of a pattern of light. The character of this pattern is related to, and controlled by, one or more electrical signals applied to the cathode-ray tube as input information. *See* CATHODOLUMINESCENCE; ELECTRON TUBE.

The cathode-ray tube is the archetypal electronic device of the twentieth century, first serving as a valuable research tool in the early experiments in modern physics and electronics, and then during World War II in radar displays on ships, on land, and in the air. In midcentury the cathode-ray tube made possible the development of television and then later the personal computer. In the late 1990s the all-electronic cinema was demonstrated using a motion picture projector incorporating a cathode-ray tube and a liquid-crystal light valve.

Hundreds of millions of cathode-ray tubes were in service at the end of the twentieth century, and tens of thousands more were manufactured daily. These tubes were commonplace in television sets, computers, homes, hospitals, banks, airplanes, and weapons of war. Even so, the cathode-ray tube is being supplanted in many of its traditional uses by flat-panel electronic devices. This trend is expected to continue until, except for perhaps a few specialized applications, the cathode-ray tube will be primarily of historical interest.

These flat-panel devices, for example liquid-crystal, plasma, and field-emission displays, are digital in nature. Their display surfaces are structured into tiny areas, called pixels, which are individually addressed by the supporting electronics. By comparison, the cathode-ray tube is an analog device, often much bulkier and heavier, and a little out of place in what is fast becoming a digital world. *See* ELECTRONIC DISPLAY; FLAT-PANEL DISPLAY DEVICE; LIQUID CRYSTALS.

Development. The technology of the cathode-ray tube was both a product of certain late-nineteenth-century experiments which led to modern physics and an indispensable element in the conduct of such experiments. In particular, the discovery of x-rays in 1895 by William Roentgen and the discovery of the electron in 1897 by J. J. Thompson resulted from experiments in which early cathode-ray devices were involved. *See* ELECTRON; X-RAYS.

While experiments with cathode rays within partial vacuum devices had been conducted by various scientists for at least the previous 200 years, Ferdinand Braun is credited with inventing in 1897 a "cathode-ray indicator tube" from which all modern cathode-ray tubes are derived. Braun's interest in the cathode-ray tube was primarily as a device for the study of phenomena related to electricity. The cathode-ray tube was well suited to this use because of its ability to provide a visual indication of alternating electric currents at low frequencies. Oscillography was thus a natural first application for the new device. To this day, the cathode-ray oscilloscope continues to be an indispensable tool in nearly all fields of electronics.

Early experimenters in television who were attempting to develop electronic scanning realized the potential usefulness of the cathode-ray tube as a television display device. In the 1920s V. K. Zworykin developed a cathode-ray tube which was improved in ways which made it particularly suitable for the display of television images. This tube, the kinescope, first demonstrated in 1929, was the forerunner of the monochrome television picture tubes which were produced commercially in the late 1940s and the 1950s. Modifications to the kinescope, including the addition of a shadow mask, led to the commercial introduction of the color picture tube in 1953. *See* PICTURE TUBE; TELEVISION.

During the 1930s, it had become apparent that the cathode-ray tube would make an excellent display device for use with the new radar technology. Many of the early television pioneers were recruited into the effort to complete the development of radar for use in World War II. *See* RADAR.

In the 1970s the cathode-ray tube became established as a widely used device for the display of computer output information. Modifications have been made to television picture tubes, both monochrome and color, to improve characteristics which are of importance to observers of computer output information. Cathode-ray tubes specialized for this use are called data display tubes. Their use has grown rapidly because of the popularity of personal computers. *See* COMPUTER; MICROCOMPUTER.

Basic elements. In the following description of a typical cathode-ray tube, emphasis is given to those features which are common to most of the tubes now produced. Some variations in construction and function from this basic cathode-ray-tube configuration are described later in the article.

The three elements of the basic cathode-ray tube are the envelope, the electron gun, and the phosphor screen (**Fig. 1**).

Envelope. The envelope is usually made of glass, although ceramic envelopes and metal envelopes have been used. The envelope is typically funnel-shaped.

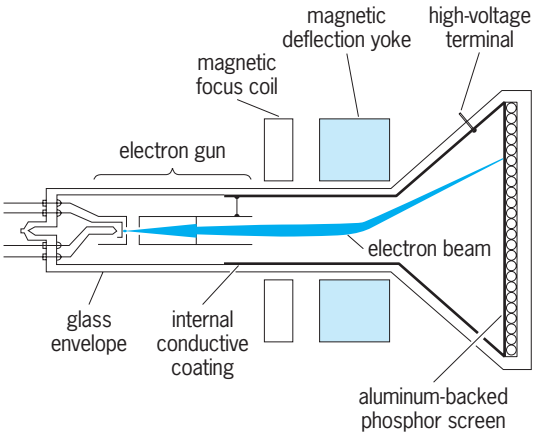

Fig. 1. Elements of a cathode-ray tube.

The small opening is terminated by the stem, a disk of glass through which pass metal leads that apply voltages to the several elements of the electron gun. A tubulation that passes through the stem is used to evacuate the air from the tube during manufacture, and is then sealed off. The electron gun is mounted within the neck portion of the envelope and is connected to the leads coming through the stem. The neck is often made sufficiently narrow to allow positioning of deflection and focusing components outside it.

The large end of the funnel is closed by a faceplate, on the inside of which the phosphor screen is deposited. The faceplate is made of high-quality clear glass in order to provide an undistorted view of the display on the phosphor screen.

Electron gun. The electron gun consists of an electrical element called a heater, a thermionic cathode, and an assemblage of cylinders, caps, and apertures which are all held in the proper orientation by devices such as glass beads, ceramic rods, and spacers.

The cathode is a source of electrons when maintained at about 1750°F (1100 K) by thermal radiation from the heater. The heater requires from 0.5 to 5 W of input power, depending on its design. Electrons emitted by the cathode are formed into a beam, and controlled in intensity by other elements of the electron gun. Depending on the design and application of the cathode-ray tube, the current of the electron beam will have an instantaneous peak value from 0.5 to 5000 microamperes. Means are provided, either within the electron gun itself or externally, to focus the electron beam to a small cross section at its intersection with the phosphor screen and to deflect it to various locations on the screen. *See* CHARGED PARTICLE OPTICS; ELECTRON EMISSION.

A monochrome cathode-ray tube is designed to display information in a single color, usually white, although green, orange, and other colors are used. A color cathode-ray tube, such as a color picture tube, is designed to display information in a full range of colors. In most cases, monochrome cathode-ray tubes employ a single electron gun. Nearly all color picture tubes employ the shadow-mask principle and use three electron guns.

The deflection path of the electron beam on the phosphor screen depends on the intended use of the cathode-ray tube. In oscillography, a horizontal trace is swept across the phosphor screen, with vertical excursions of the beam which coincide with variations in the strength of some electrical signal. In television, a raster of closely spaced horizontal lines is scanned on the phosphor screen by the electron beam, which is intensity-modulated to produce a visible picture. Radar makes use of a variety of specialized electron-beam scanning patterns to present information to an observer. It is also a common practice to apply scan-conversion techniques to the radar signals so that they may be displayed by using a television raster scan.

In the display of computer output information, two general approaches to beam deflection are used: The raster-scan technique may be identical in format to that used for television or may utilize a greater number of scanning lines for increased definition. The random-scan technique involves computer control to direct the electron beam to locations which may be anywhere on the tube face.

Phosphor screen. The phosphor screen consists of a layer of luminescent material coated on the inner surface of the glass faceplate. Typically, the luminescent material is in the form of particles whose size is from 1 to 100 micrometers, depending upon the design and intended use of the cathode-ray tube. Monochrome cathode-ray tubes generally use a single layer of a homogeneous luminescent material. Color cathode-ray tubes typically utilize a composite screen made up of separate red-, green-, and blue-emitting luminous materials. *See* LUMINESCENCE; PHOSPHORESCENCE.

Design considerations. There are certain concerns which must be dealt with in the design of all types of cathode-ray tubes.

Safety. Foremost is the matter of safety. While cathode-ray tubes now in use generally can be considered safe, when employed as intended by their manufacturer, there are potential health hazards which must be considered in the design both of the cathode-ray tubes and of the equipment incorporating them. For example, as discussed below, it is frequently desirable to operate cathode-ray tubes at high voltages and currents, for example, tens of kilovolts and hundreds to thousands of microamperes. Consequently, there are possible hazards of electrical shock, x-radiation, and high temperatures. Designers of cathode-ray tubes and equipment using the tubes are generally quite aware of these possible hazards and are able to effectively utilize designs without undue risk to health and safety.

Vacuum maintenance. Another general concern is the absolute necessity of maintaining a high vacuum within the envelope of the cathode-ray tube so that the electron beam can travel unimpeded from the electron gun to the phosphor screen. Also, it is frequently not possible for the various electrodes within the cathode-ray tube to withstand electrical breakdown except under high vacuum. Cathode-ray tubes are exhausted by vacuum pumps for periods of up to several hours before being tipped off (permanently sealed at the exhaust tubulation). Cathode-ray tubes are typically raised to temperatures of a few hundred degrees Celsius during exhaust in order to outgas the internal parts. Internal pressure in a cathode-ray tube at tip-off is typically in the range of 10^{-4} to 10^{-8} torr (10^{-2} to 10^{-6} pascal), the desired value being determined by the design and intended use of the cathode-ray tube.

With very few exceptions, commercial cathode-ray tubes are never again vacuum-pumped after tip-off, and are expected to maintain an acceptable high level of vacuum. Leaks through the vacuum envelope are extremely rare; those problems which do arise are generally from the outgassing of internal parts. In order to assist in maintaining a high internal vacuum throughout many years of storage and use, it is now a universal practice to include within the cathode-ray tube one or more small devices called getters. Getters, through chemical action, absorb residual gas

molecules within an envelope, and will continue to do so usually for several years.

Voltage control. Another general concern is the need to have several different voltages applied to structures within the cathode-ray-tube envelope. These voltages must be individually passed through the vacuum envelope, either through the stem leads or through metal buttons or pins sealed into the envelope. Frequently, an electrical cable carrying voltages as high as 50 kV will be encapsulated to a button terminal on the outside of the cathode-ray tube. Inside the cathode-ray tube, it is important that all envelope surfaces, as well as all electrode structures, be set at definite, known electric potentials. For this reason, conductive coatings which include materials such as graphite, iron oxide, stannous chloride, or evaporated aluminum are applied to nearly all interior envelope surfaces. This is one of the functions of the aluminum coating on the back of the phosphor screen. By these means, the path of the electron beam is fully established and controlled.

Performance. Two major considerations in cathode-ray-tube design are the choice of the overall operating voltage, that is, the potential difference between the cathode and the phosphor screen, and the electron-beam current to the phosphor screen. The product of the overall operating voltage (called the screen voltage) and the average value of the screen current gives the power input to the phosphor screen. For example, a cathode-ray tube operating at a screen voltage of 20 kV and screen current of 500 μA has 10 W of input power. A small fraction of this power, typically 10 to 20%, is converted into visible light by the phosphor. Less than 1% of the input power is converted into x-radiation, nearly all of which is absorbed within the glass walls of the envelope. The remainder of the input power is dissipated in the faceplate as heat. Most cathode-ray tubes operate at power levels such that this heat is allowed to radiate away or otherwise be removed by conduction and convection. Tubes which operate at very high voltages and currents, such as projection tubes, sometimes require special means, such as cooling fans, to remove heat from the faceplate.

The choice of operating screen voltage and current can have considerable effect on the design and construction of the cathode-ray-tube envelope, electron gun, and phosphor screen. It is important to consider the tradeoffs involved in making these choices. The higher the screen voltage, the greater the level of performance which can be obtained from the cathode-ray tube. The performance characteristics most directly affected by choice of screen voltage and current are screen brightness and resolution. While higher values of screen voltage will improve both brightness and resolution, higher values of screen current will improve brightness but degrade resolution performance. For any particular cathode-ray-tube design, the upper level of screen current which can be utilized is set by the allowable level of resolution degradation. On the other hand, the upper limit of screen voltage that can be utilized is set by (1) the capability of the internal tube elements, as well as the glass wall of the envelope, to withstand high potential gradients without electrical breakdown; (2) the practical problems involved in generating a very high voltage and delivering it to the cathode-ray-tube high-voltage terminal; (3) added power requirements for the deflection drive circuits; and (4) additional precautions necessary to ensure safe operation at high voltages as far as electrical and x-radiation safety are concerned.

A great deal of the design effort for any cathode-ray tube goes into obtaining the maximum attainable brightness and resolution at a selected screen voltage and current. The required brightness and resolution vary greatly, depending on the intended use of the cathode-ray tube (**Table 1**).

Envelopes. Cathode-ray-tube envelopes have been made from metal, ceramic, and glass materials. Glass is now used more than any other material, because it has the most desirable combination of mechanical, electrical, and optical properties. A variety of specialized glasses have been developed for use in cathode-ray tubes. *See* GLASS.

Electrical resistance. An important characteristic of glass for cathode-ray tubes is its electrical resistance. Early oscilloscope cathode-ray tubes, operating at voltages from 500 to 2000 V, were made by using a soda-lime-silica combination similar to that used in windows and bottles. This glass was found to be useless at higher voltages because of low electrical resistance. *See* ELECTRICAL RESISTANCE.

X-ray absorption. A second important characteristic is produced any time that electrons traveling at high speed strike a stationary target. The amount of x-radiation produced is proportional to the square of the energy of the electron and to the first power of the atomic number of the constituents of the target

TABLE 1. Characteristics of cathode-ray tubes

Type of cathode-ray tube	Screen voltage, kV	Screen current, μA	Brightness, footlamberts (candelas/m^2)	Resolution expressed as spot size, in. (mm)	Typical application
Oscilloscope	5	100	40 (137)	0.050 (1.3)	Oscilloscope
Photorecording	15	5	3 (10)	0.001 (0.025)	Phototypesetter
Picture tube (monochrome)	20	300	100 (340)	0.015 (0.4)	Television or data display
Picture tube (color; 3 guns)	30	400	100 (340)	0.020 (0.5)	Television or data display
Projection	45	2000	10,000 (34,000)	0.004 (0.1)	Data display

material. The target material is that of the phosphor screen, the faceplate, and shadow mask, if one is present. The amount of x-radiation generated within the cathode-ray tube can be substantial at higher voltages. Manufacturers and users of cathode-ray tubes depend on the material of the envelope to absorb nearly all of this radiation. Absorption of x-radiation within the envelope depends on the thickness of the envelope and its composition, and is given by Eq. (1),

$$I = I_0 e^{-\mu t} \tag{1}$$

where I is the intensity of x-radiation at the outside surface of the cathode-ray tube; I_0 is the intensity of x-radiation at the inside surface of the cathode-ray tube; μ is the linear x-ray absorption coefficient of the particular type of glass used (usually expressed in units of 1/cm); and t is the thickness of glass (usually expressed in centimeters). *See* X-RAYS.

Optical properties. A third important characteristic for glass used in cathode-ray tubes is the optical properties of the faceplate glass. These properties vary widely, depending on the design and intended use of the cathode-ray tube. The best cathode-ray-tube faceplates have optical characteristics and quality equivalent to those of the finest photographic lenses. Cathode-ray-tube faceplates may be clear or may have intentionally reduced optical transmission in order to improve the contrast of the display. Special-purpose cathode-ray tubes have been built with fiber-optic faceplates which make possible direct-contact exposure of photosensitive film and paper.

Shapes and sizes. **Figure 2** shows a variety of cathode-ray tubes having different shapes and sizes. Cathode-ray-tube envelopes are generally round or are rectangular with a 3×4 product aspect ratio, although numerous other shapes have been built. Extremes in sizes of cathode-ray tubes are exhibited by 1-in.-diameter (2.5-cm) and 36-in.-diameter (91-cm) round cathode-ray tubes.

Another parameter characterizing the shape of a cathode-ray tube envelope is its deflection angle. The deflection angle is related roughly to the angular dimension of the opening of the funnel. Cathode-ray tubes with deflection angles from 15 to 110° have been built.

Electron gun. The electron gun and its associated neck components generate the electron beam and direct it toward the phosphor screen. This assembly may conveniently be broken down into three separate systems: the electron emission system, the focusing system, and the deflection system.

Emission systems. An emission system may consist of the cathode with its associated heater, a control grid, also called grid #1, and an accelerating grid (**Fig. 3**).

The voltage applied to the accelerating grid sets up a field which penetrates the opening in the control grid and determines the magnitude of the space-charge–limited current drawn from the cathode. As the control grid is made more negative in potential, the penetration of the accelerating field and, consequently, the beam current are reduced.

The electrostatic field in the emission system of an electron gun is so shaped by the geometry of the several elements that electrons perpendicularly leaving the cathode enter trajectories which carry them across the axis of the system in the vicinity

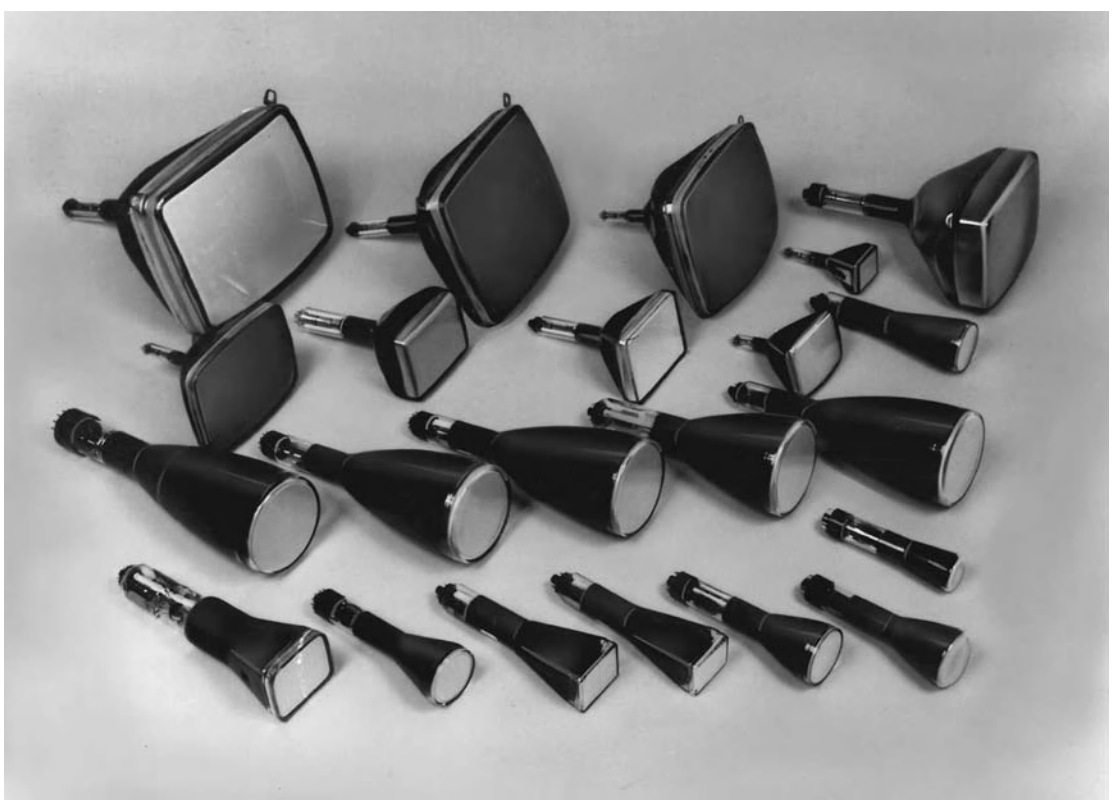

Fig. 2. Typical cathode-ray tubes.

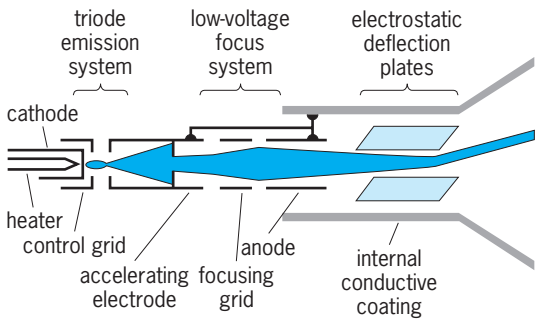

Fig. 3. Simplified electron gun with triode emission system, electrostatic focus, and electrostatic deflection.

of the control-grid–accelerator-grid space. The location at which the electrons cross the axis is called the crossover point and represents the minimum beam cross section in the vicinity of the cathode surface. Beyond this point, beam electrons enter the field-free drift space within the cylindrical electrode attached to the accelerating grid and travel along straight diverging trajectories toward the focusing system. *See* ELECTRON MOTION IN VACUUM.

A key parameter for the emission system of an electron gun is the cutoff voltage, that voltage between the control grid and the cathode which is just necessary to extinguish the flow of electrons from the cathode. Depending on a particular design, this value may lie between −15 and −150 V, the grid being at a negative voltage with respect to the cathode.

The control grid can be modulated by a video signal, or otherwise varied in voltage between cutoff and zero volts, in order to produce a desired level of beam current. The total current from the cathode (known as the Child-Langmuir current) is given by Eq. (2), where I is the actual cathode current in mi-

$$I = KV_d^3 V_c^{-3/2} \tag{2}$$

croamperes, K is a constant whose value typically is in the range of 3 to 5, V_d is the grid drive which is equal to the difference between the actual voltage on the control grid and the cutoff voltage, and V_c is the cutoff voltage. Absolute values are used for both V_d and V_c.

Emission-system operation, as described above, in which the cathode is grounded and the control grid is modulated, is termed grid drive. Another mode of operation is cathode drive, in which grid #1 is grounded and the cathode is modulated. The formula for cathode current in the cathode-drive mode is more complicated than that given above for the grid-drive mode.

Often, a circular baffle termed a limiting aperture is introduced in the path of the electron beam. The limiting aperture intercepts the outer portions of the beam and prevents their reaching the phosphor screen, thus improving the resolution of the display. The portion of the beam intercepted may be from 0 to 98%, depending on the design and intended use of the cathode-ray tube. For this reason, the electron-beam current reaching the phosphor screen may be a small fraction of the current emitted by the cathode.

An emission system consisting of a cathode, a control grid, and an accelerating grid is termed a triode emission system. A tetrode emission system (**Fig. 4**) has an additional electrode, called grid #2, inserted between the control grid and the accelerating grid. A primary purpose of grid #2 is to electrically isolate the emission system from the focusing system and the high-voltage portion of the envelope, so that changes in focusing or anode voltage will not affect cutoff and drive voltages. Grid #2 is typically operated within a range of 150 to 1500 V positive with respect to the cathode, depending on the design of the electron gun. Grid #2 may be adjusted in voltage to select a desired cutoff voltage. Increasing the grid #2 voltage will raise the magnitude of the cutoff voltage, and vice versa.

The element within the cathode-ray tube set at the highest voltage to which the electron beam is subjected prior to its deflection is termed the anode, or sometimes the ultor. Except in the special case of cathode-ray tubes which utilize postdeflection acceleration, the anode is the element at the highest positive direct-current voltage in the cathode-ray tube and is the element to which the aluminum backing of the phosphor screen is connected. In many cathode-ray tubes, the anode will function as the emission-system accelerating electrode.

Focusing systems. The two basic types of focusing used in cathode-ray tubes, magnetic and electrostatic, have been extensively used in many different cathode-ray-tube designs with fully acceptable performance. Electrostatic focus has the advantage of being generally less costly and easier to implement for most applications. In certain applications where resolution at high beam current is required, there is usually a performance advantage in using magnetic focus.

1. *Magnetic focusing*. In magnetic focusing systems, a short focus coil is fitted externally to the neck of the envelope of the cathode-ray tube. The coil is arranged so that magnetic flux lines flow within the neck parallel to the direction of travel of the electron beam. The focus coil is positioned toward the phosphor screen from the anode portion of the electron gun, so that focusing will take place in a region

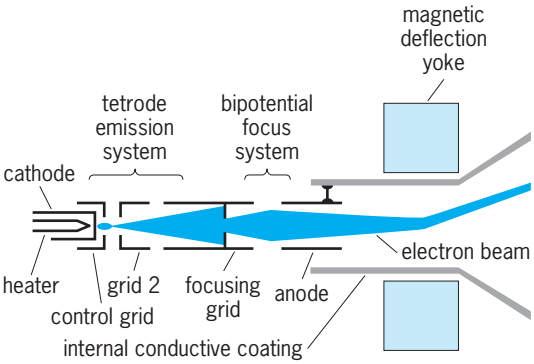

Fig. 4. Simplified electron gun with tetrode emission system, electrostatic focus, and magnetic deflection.

free from electrostatic potential gradients. Current through the focus coil is adjusted to bring the electron beam to a sharp focus on the phosphor screen. Many designs will require dynamic changes in the focus current to maintain sharp focus as the electron beam is deflected to various positions on the tube face. *See* MAGNETIC LENS.

2. *Electrostatic focusing*. In electrostatic focusing systems, an electrostatic lens is formed between two or more electrodes within the neck of the tube. Many different electrode arrangements which will give satisfactory focus are in use. Two frequently encountered arrangements are the bipotential lens (Fig. 4) and the unipotential focus lens (Fig. 3). These lenses are often used with tetrode emission systems, where the presence of grid #2 serves to prevent variations in focus voltage from altering the emission-system cutoff voltage. The bipotential lens requires an additional electrode, the focus electrode, inserted between grid #2 and the anode. The focus electrode operates at an intermediate voltage between that of grid #2 and the anode. Typically, this voltage is in the range of 20 to 40% of the anode voltage, depending on lens design, and is adjusted in value for optimum focus. In the bipotential lens, the main focusing action takes place in the region between the focus electrode and the anode.

In the unipotential lens, the anode is split into two separate cylinders and an additional electrode, the focus electrode, is inserted into the gap between them. This focus electrode operates at a potential which is only about 25% that of the anode. The unipotential lens generally will provide superior focus quality at high electron beam currents.

Some newer designs for electrostatic focusing lens involve the introduction of several focusing electrodes stacked together, each operating at a slightly different voltage. *See* ELECTRON LENS; ELECTROSTATIC LENS.

Deflection systems. Two basic types of deflection are electrostatic deflection and magnetic deflection. In principle, either types of deflection can be used for any application, although each has advantages which have led to its being favored for certain classes of use.

1. *Electrostatic deflection*. With electrostatic deflection, it is possible to very quickly deflect the beam from one location to any other location on the screen. Operation is possible over a wide frequency range. The band pass of an electrostatic deflection system is limited in a practical sense by the amount of power the equipment designer can justifiably expend in achieving adequate display brightness and size. Laboratory oscilloscopes designed around conventional cathode-ray tubes and having useful band pass to the order of 50 MHz are readily obtained.

In the electrostatic-deflection electron gun shown in Fig. 3, the deflecting system consists of a single pair of parallel electrodes mounted symmetrically about the axis of the gun. The electron beam entering and traversing the space between the electrodes does so at a constant axial velocity corresponding to the average potential of the electrodes. This potential is usually adjusted to be essentially that of the anode. Signals to be applied to the deflecting electrodes are first passed through a paraphrase amplifier in which two components equal in voltage but opposite in polarity are generated. These two components can be applied to the opposing deflecting electrodes without affecting the average potential in the space between them. Under the influence of the applied deflection signals, an electric field is set up between the electrodes. This field imparts an acceleration to the electron beam which is perpendicular to the direction of the constant beam velocity. In the space between the electrodes, therefore, the beam travels in a parabolic path. After leaving this transverse field, the beam continues to the phosphor screen on a straight path tangent to the parabola. The assumption is made that no postdeflection acceleration is present. If the signals applied to the electrodes vary linearly with time, a straight line is traced out upon the phosphor screen. In a practical tube, a second pair of deflecting electrodes is mounted farther along the gun axis at right angles to the first. These electrodes provide a second independent axis of deflection, orthogonal to the first.

The deflection produced by a pair of parallel electrodes is given by Eq. (3), where y is the deflection

$$y = \frac{LbV_d}{2aV} \tag{3}$$

observed at the phosphor screen in any linear units, L is the distance from the axial midpoint of the electrodes to the phosphor screen in the same linear units, b is the deflecting-electrode length in the same linear units, a is the deflecting-electrode spacing in the same linear units, V_d is the potential difference between electrodes in volts, and V is the beam acceleration potential in volts.

2. *Magnetic deflection*. Magnetic deflection systems generally require more time, perhaps tens of microseconds, to deflect the electron beam from one location on the screen to another. This is because a change in position requires a change in the value of the current through an inductive coil. Magnetic deflection systems do have an important advantage in that they can deflect the beam through a much wider deflection angle with less distortion in the shape of the cross section of the beam than is possible with electrostatic deflection. Also, with high anode voltages of 20–30 kV, the current required for magnetic deflection yokes can readily be provided; whereas the high dynamic voltages which would be required for electrostatic deflection plates, on the order of 3–4 kV peak to peak, cannot easily and economically be achieved. For these reasons, cathode-ray tubes employing magnetic deflection are widely used, particularly in television, radar, and data display.

An elementary magnetic deflection system takes the form of a pair of coils, each wound on a rectangular form, positioned externally to the cathode-ray tube with major planes parallel to beam direction on opposite sides of the neck. The magnetic field lines are thus perpendicular to the axis of the electron gun mounted therein. On entering this field,

the beam electrons encounter forces which impart to them acceleration perpendicular to their velocity. If the flux density is uniform, the path followed by the electrons is circular and lies in a plane perpendicular to the lines of flux. On leaving the deflection field, the beam travels in a straight-line path tangent to the circle just described. The deflection given an electron beam which traverses a field of uniform flux density limited to the space between two planes perpendicular to the initial direction of the beam motion (that is, the tube axis) is given approximately by Eq. (4), where y is the deflection observed at the

$$y \approx \frac{LbB}{3.37\sqrt{V}} \tag{4}$$

phosphor screen in centimeters (1 in. = 2.54 cm), L is the distance from the axial midpoint of the magnetic field to the phosphor screen in centimeters, b is the axial extent of the magnetic field in centimeters, B is the flux density in gauss (1 tesla = 10^4 gauss), and V is the beam acceleration potential in volts. By an appropriate arrangement of the windings of a magnetic-deflection coil pair, more commonly called a deflecting yoke, it is possible to obtain useful total deflection angles up to 110°.

Phosphor screen. This element converts electrical energy to visible radiation. Materials known as phosphors are said to be luminescent; that is, they are able to emit light at temperatures substantially below those which produce incandescence. A phosphor which is excited to luminescence by electron bombardment is described as cathodoluminescent. In the case of phosphors for cathode-ray tubes, luminescence invariably persists after cessation of excitation. Luminescence which continues for more than 10^{-8} s after excitation is removed is called phosphorescence. Luminescence which is coincident in time with excitation is known as fluorescence.

A wide variety of materials display the property of cathodoluminescence. Phosphors used in commercially fabricated cathode-ray tubes are generally inorganic, nonmetallic, crystalline materials. Many phosphors now make use of materials which incorporate transition elements, such as yttrium, and the lanthanoid metals: lanthanum, cerium, europium, gadolinium, and terbium. Phosphors made from materials using these elements are known for their high luminous output, purity of color, and stability.

Some of these materials will emit radiation in the pure state. Most, however, display practical luminescence only when activated by an impurity (**Table 2**). These impurities, deliberately introduced in amounts ranging from 1 part in 100,000 to 1 part in 100, play a profound role in determining the efficiency, color, and persistence of the emission obtained from a given phosphor. **Figure 5** shows the spectral emission density curves which plot relative radiant energy versus wavelength for certain of these phosphor screens.

One of the criteria by which a phosphor material is judged is its ability to convert electrical energy to useful radiation. In many cases, cathode-ray-tube screens are intended to be viewed by human observers. It is appropriate, therefore, to speak in terms of visible flux output. The luminous efficiencies of practically all commercially significant phosphors are within the range 5–50 lumens per watt of exciting power. Light output increases with increasing bombarding-current density, eventually exhibiting evidence of saturation at current-density levels

TABLE 2. Characteristics of some phosphor screens used in cathode-ray tubes

WTDS* designation	Previous EIA† designation	Base material (activator)	Color of luminous emission‡	Persistence classification§	Typical application
BE	P11	Zinc sulfide (silver)	Blue	Medium short	Photorecording
BH	P47	Yttrium silicate (cerium)	Purplish blue	Very short	Flying-spot scanner
BM	P55	Zinc sulfide (silver)	Blue	Medium short	Projection, large screen
GH	P31	Zinc sulfide (copper)	Green	Medium short	Data display, oscillography
GJ	P1	Zinc orthosilicate (manganese)	Yellowish green	Medium	Projection, oscillography
GM	P7	Zinc sulfide (silver) on top of zinc cadmium sulfide (copper)	Yellowish green	Long	Radar
GY	P43	Gadolinium oxysulfide (terbium)	Yellowish green	Medium	Data display
KA	P20	Zinc cadmium sulfide (silver)	Yellow green	Medium	Storage tubes
KG	P46	Yttrium aluminate (cerium)	Yellow green	Very short	Flying-spot scanner
KH	P48	Yttrium aluminate (cerium) and yttrium silicate (cerium)	Yellow green	Very short	Flying-spot scanner
KJ	P53	Yttrium aluminum garnet (terbium)	Yellowish green	Medium	Data display
RF	P56	Yttrium oxide (europium)	Red	Medium	Projection, large screen
WB	P45	Yttrium oxysulfide (terbium)	White	Medium	Projection, large screen
WW	P4	Zinc sulfide (silver) and zinc cadmium sulfide (silver)	White	Medium short	Monochrome television, data display
		Zinc sulfide (silver)	Blue	Medium short	Color television, data display
X	P22¶	Zinc cadmium sulfide (silver)	Green	Medium short	
		Yttrium oxysulfide (europium)	Red	Medium short	

*World Wide Type Designation System. Designation administered by the Electronic Industries Association.
†Electronic Industries Association.
‡Color shown is that of phosphorescence, that is, the color after cessation of phosphor excitation.
§Persistence categories are based upon time for radiant output to drop to 10% of initial level following interruption of excitation: very long, 1 s or more; long, 100 ms to 1 s; medium, 1 ms to 100 ms; medium short, 10 μs to 1 ms; short, 1 μs to 10 μs; very short, less than 1 μs.
¶Composite structured screen used in conjunction with shadow mask for full color display. The materials shown are one example of several sets of materials used for this screen.

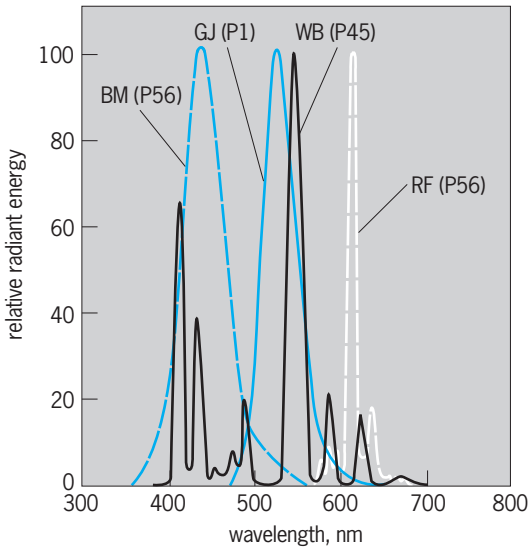

Fig. 5. Spectral emission density curves for phosphor screens.

which are characteristic of particular phosphors. For some phosphors, saturation effects are quite evident at the 10 μA/cm^2 level. Others display relatively undiminished luminous efficiencies at levels of several hundred microamperes per square centimeter.

The color of the emission from a phosphor is governed both by the base material and by the activator. The spectral-energy distributions of most phosphors are fairly broad, although colors of high saturation are readily obtained. Phosphors made from materials incorporating rare-earth and related elements are noted for having spectral-energy distributions in which the radiant energy is concentrated into very narrow emission bands at just a few wavelengths.

As is evident from Table 2, a wide range of colors is available; the tube designer is restricted in this matter only insofar as particular efficiencies and persistences must be achieved simultaneously with a given color.

A great majority of cathode-ray-tube screens now being produced are aluminized. The side of the phosphor screen facing the electron gun is coated with a thin electron-transparent film of aluminum by vacuum-deposition techniques. This process not only stabilizes the phosphor potential at the applied acceleration voltage, but also provides an increased light output because of the mirror effect of the highly reflective layer, which redirects radiation that could otherwise be lost in the interior of the envelope.

The fineness of detail in an image presented on a phosphor screen is limited to some extent by the granular nature of the screen. Both particle size and layer thickness affect performance. Phosphor screens intended for very high-resolution cathode-ray tubes employ thin layers of phosphor with particles as fine as 1 μm in size.

Varieties of tubes. A wide variety of available envelopes, electron guns, and phosphor screens have been combined in different ways to fashion cathode-ray tubes specialized to meet the needs of a host of different applications. For example, cathode-ray tubes intended for use in oscilloscopes used electrostatic deflection electron guns in order to meet the need for high-speed deflection of random waveforms. Cathode-ray tubes for World War II radar displays used a phosphor material which would persist for several seconds (GM/P7 in Table 2) in order to keep the radar image visible until the slowly rotating antenna returned to refresh the image.

When computer data processing with ample memory for image storage became readily available, such highly specialized cathode-ray-tube designs were no longer required. Instead, the basic monochrome and color cathode-ray tubes originally conceived and developed for broadcast television have now adapted for general use for the display of all types of pictorial information.

It is convenient to divide the varieties of cathode-ray tubes into three categories: direct view, projection, and photorecording.

Direct view. A large class of applications for cathode-ray tubes involves either the presentation on the cathode-ray-tube screen of an actual picture with a full black and white halftone range or with full color, such as is required for television, or the presentation of a computer-generated display which may consist of alphanumerics, graphics, or a variety of pictorial subjects. Tubes for the direct viewing of such presentations are required to have large display sizes, high brightness, high resolution, and in many cases a full halftone range and full color capability.

Cathode-ray tubes for these presentations have always employed magnetic deflection and generally electrostatic-focus electron guns operating at high voltages from 15 to 36 kV. Envelopes for television cathode-ray tubes have traditionally been rectangular, of 3×4 aspect ratio for standard television and 9×16 aspect ratio for high-definition digital television. Popular usable screen diagonals have been 10 to 36 in. (25 to 91 cm). Deflection angles have ranged from 10 to 110°.

Cathode-ray tubes for computer-generated data-display applications are very similar to television picture tubes but may use other rectangular or square display formats. The need for uniformity of resolution in the display, particularly from the center of the picture to the edge, is even more acute for tubes intended for data display than it is for tubes intended for television only. Particular attention is paid to this need in the design of cathode-ray tubes for data-display applications.

One special class of direct view applications comprises displays for use in automobiles, airplanes, boats, and military vehicles of all types. While cathode-ray tubes for such applications are similar to those described above, they tend to be smaller and brighter, and they must be quite rugged.

Projection tubes. Projection tubes are not intended to be directly viewed. The display on the phosphor screen is projected by using an optical system, such as a lens, onto large screens. Screen sizes vary widely, the largest being those in theaters and sports arenas that are equipped for projection television.

Cathode-ray tubes for projection applications are usually of the general type described above but generally are optimized for extremely high brightness and resolution capability. Projection cathode-ray tubes generally have a largest phosphor screen dimension of 3 to 8 in. (7.5 to 20 cm). This smaller size is compatible with available high-efficiency optical systems. Projection cathode-ray tubes generally operate at higher voltages than direct-view tubes. Tubes which can be operated at voltages as high as 80 kV have been built.

Projection tubes for very large projection screens, up to 32 ft (9.8 m) wide and more, are used in conjunction with liquid-crystal devices which function as image light amplifiers. The image on the projection tube is projected onto the back surface of the liquid-crystal device, and is then transferred to the front, mirrorlike surface of the device. Light from a powerful arc lamp is reflected from this front surface, projecting the image onto the projection screen. This technique allows the projection tube design to be optimized for best resolution, with a relatively low brightness requirement. Motion picture projectors using cathode-ray tubes and following this approach were used in the 1990s for the first public demonstration of all-electronic cinema.

On a different scale, projection tubes used in avionics helmet-mounted displays make use of infinity optics to project images directly onto the retina of the aviators. Such cathode-ray tubes must be very small, light, low-voltage, low-power devices.

Photorecording. Another class of cathode-ray tubes which are not intended for direct viewing by human observers comprises photorecording tubes. The applications for these tubes require that the phosphor screen display be projected by an optical system, such as a lens, onto a photosensitive medium, such as photographic film. Applications include electronic phototypesetting and the storage of computer output information on microfilm. Photorecording cathode-ray tubes are required to have extremely high resolution capability, to be extremely stable over long periods of time, and to have accurate and precise display geometry. A photorecording display may have 50 times the information content of a typical broadcast television picture. *See* PRINTING.

Photorecording tubes usually employ magnetic deflection and may utilize either magnetic or electrostatic focus. They operate at anode voltages in the range 10 to 20 kV. Many photorecording cathode-ray tubes use envelopes with 3- or 5-in. (7.5- or 12.5-cm) round faceplates of optical quality. Displays on these faceplates are projected through a lens onto the photosensitive medium. Other photorecording tubes have faceplates made from fiber-optic plates which, in use, are in direct contact with the photosensitive medium. *See* OPTICAL FIBERS.

A special cathode-ray tube which is closely related to the photorecording tube is the flying-spot scanner. The flying-spot-scanner display consists of a spot of light moving across the phosphor screen to scan a raster such as that used in television. This raster is blank and includes no video modulation other than retrace blanking. In a typical application, the raster will be projected through a lens onto a slide transparency. The light passing through the transparency is picked up by one or more photomultiplier tubes. The pictorial content of the transparency in this way is made available as a video signal, provided by the photomultiplier, which can then be displayed on a television monitor.

A special group of phosphor screens having extremely short persistence, on the order of 150nanoseconds, has been developed for use with flying-spot scanners. These are identified in Table 2. Short persistence is necessary to avoid severe distortion of the video signal.

Competing technologies. Considering the length of time that the cathode-ray tube has been in use, and the number of uses to which it has been put, it must be considered one of the most successful electronic devices ever conceived. There are, nonetheless, certain drawbacks inherent in the use of the cathode-ray tube: (1) it must have a certain size and weight to provide a vacuum-tight enclosure and sufficient room internally for generation and deflection of the electron beam; (2) it has a long but finite lifetime, limited by degradation processes associated with the high temperatures at which the heater and cathode are operated; and (3) it is essentially an analog device which must be interfaced with other electronic devices, most of which are now exclusively digital.

Display technologies not involving the use of cathode-ray tubes have been under development for many years. The more promising make use of plasma or gas discharge, electroluminescent, and liquid-crystal devices. These devices are increasingly gaining acceptance in data display applications which previously used cathode-ray tubes.

Norman W. Patrick

Bibliography. P. A. Keller, *The Cathode-Ray Tube: Technology, History, and Applications*, 1992; S. Sherr, *Electronic Displays*, 2d ed., 1993; J. Whitaker, *Electronic Displays*: *Technology, Design, and Applications*, 1994.

Cathode rays

The name given to the electrons originating at the cathodes of gaseous discharge devices. The term has now been extended to include low-pressure devices such as cathode-ray tubes. Furthermore, cathode rays are now used to designate electron beams originating from thermionic cathodes, whereas the term was formerly applied only to cold-cathode devices. *See* CATHODE-RAY TUBE.

The basis for the nomenclature is purely historical. The first outward evidence was flourescence from the glass walls of cold-cathode discharge tubes. This fluorescence appeared as the pressure was reduced to the region where the mean free path became greater than the tube dimensions. At these pressures, the gas in the tube no longer emitted an appreciable amount of light. It was ascertained that the wall

fluorescence had its origin in rays of particles coming from the cathode. Furthermore, it was demonstrated that these particles traveled in approximately straight lines. If an object was interposed between the cathode and the wall, the fluorescence disappeared in the optically shadowed region of the wall. In 1897 J. J. Thomson was able to show, using electric and magnetic fields, that the particles were negatively charged. In his experimental arrangement, he eliminated all but a very narrow beam of these rays. By producing an electrostatic deflection of the beam and then counteracting this with a magnetic deflection, he was able to determine the charge-mass ratio e/m. This was found to be the same as that of the electron, and resulted in the identification of these particles. *See* ELECTRON.

Thomson's measurements also form the basis for modern beta-particle spectroscopy. In this case e/m is known, so that either an electric or magnetic field may be used to determine the energy of the electrons which emanate from various radioactive materials, such as beta particles. *See* BETA PARTICLES; ELECTRICAL CONDUCTION IN GASES; ELECTRON EMISSION; PARTICLE ACCELERATOR; X-RAYS. Glenn H. Miller

Bibliography. J. W. Rohlf, *Modern Physics from Alpha to the Z^0*, 1994; T. R. Sandin, *Essentials of Modern Physics*, 1989; J. Taylor and C. Zafaritos, *Modern Physics for Scientists and Engineers*, 1991.

Cathodoluminescence

A luminescence resulting from the bombardment of a substance with an electron (cathode-ray) beam. The principal applications of cathodoluminescence are in television, computer, radar, and oscilloscope displays. In these a thin layer of luminescent powder (phosphor) is evenly deposited on the transparent glass faceplate of a cathode-ray tube. After undergoing acceleration, focusing, and deflection by various electrodes in the tube, the electron beam originating in the cathode impinges on the phosphor. The resulting emission of light is observed through the glass faceplate, that is, from the unbombarded side of the phosphor coating. *See* CATHODE-RAY TUBE.

The luminescence of most phosphors comes from a few sites (activator centers) occupied by selected chemical impurities which have been incorporated into the matrix or host solid. The interaction of cathode rays with the phosphor involves a collective excitation of all the atoms of the host rather than a selective excitation of the luminescent centers, a condition that allows the dissipation of beam energy by competing nonluminescent processes. An appreciable energy loss occurs as soon as the primary cathode-ray beam strikes the phosphor; 25–35% of the electrons are immediately reflected (backscattered) due to coulombic repulsion. The electrons that actually penetrate the phosphor give rise to a combination of several processes that can be described only qualitatively. Some x-rays are produced, but in the main the high-energy electron beam ionizes the solid, producing a plasma of many lower-energy (secondary) electrons. These electrons are multiply scattered, successively losing more and more energy to the solid by various nonradiative paths. Although the bombarded phosphor is in the complex excited condition described above, a small part of the excitation energy is transferred by various mechanisms to the activator centers, causing them to luminesce. Because of the complex mode of interaction of cathode rays with phosphors, the energy efficiency of light production by cathodoluminescence is lower than the best efficiencies obtainable with photoluminescence. Conversion efficiencies of currently used display phosphors are between 2 and 23%.

The brightness B of a phosphor under cathode-ray excitation depends on the accelerating voltage V and the current density j. Many phosphors exhibit a dead voltage V_0 below which they show diminished output, presumably due to radiationless dissipation of the energy by poison centers, which are present only at the surface where this low voltage excitation occurs. Above V_0, the brightness is proportional to $(V - V_0)^q$, where q is between 1 and 3. At a given voltage, the brightness initially varies linearly with the current density, and then may increase more slowly with increasing j (saturation). The blue-emitting zinc sulfide (ZnS) and green-emitting zinc-cadmium sulfide [(Zn,Cd)S] phosphors used in color television exhibit this saturation effect, as well as a color shift and shorter afterglow at high current density, but the red-emitting europium-activated phosphor and other rare earth–activated phosphors do not. Linearity at high current density is an important requirement for phosphors to be used in projection television or aircraft pilot displays.

The blue and green phosphors for television are broadband emitters with dominant wavelengths of 464 and 556 nanometers, respectively; the red europium-activated yttrium oxysulfide phosphor is a line emitter at 605 nm. Their luminous efficiencies are 7.5, 65, and 17 lumens per input watt, respectively, and all have persistences of less than 10^{-4} s to reduce the smearing of fast-moving objects in television pictures. For easy viewing of slowly scanned radar screens, on the other hand, phosphors with persistences of up to 0.55 are used.

The activators in zinc sulfide and zinc-cadmium sulfide are parts-per-million traces of donor-acceptor pairs of impurities, chlorine and silver in the former and aluminum and copper in the latter. Luminescence in these phosphors is produced when electrons trapped at donors recombine with holes trapped at acceptors. The broad emission band is a complex of emissions of slightly different wavelengths from pairs having a variety of separations in the host lattice. This mechanism can also explain phosphor saturation, shortened persistence, and color shifts in the sulfides at high current density. In the rare earth–activated phosphor, the emitting center Eu^{3+} is present at a concentration of 4 mole %. The deep-lying atomic orbits of Eu^{3+} are not appreciably affected by interaction with the other constituent atoms of the phosphor, and the

red light therefore appears in a very narrow range of wavelengths. *See* HOLE STATES IN SOLIDS; SEMICONDUCTOR.

An important requirement for a good cathodoluminescent phosphor is the possession of good secondary electron emission properties; otherwise, it charges up negatively and reduces the effective potential of the bombarding beam. In most cases the secondary electron emission coefficient R is less than 1, and the screen must be coated with a film of aluminum to provide conductance to the power supply in order to prevent charge buildup. The film provides two additional benefits. Its action as an optical mirror nearly doubles the display brightness, and it shields the phosphor from bombardment by residual gas ions that remain or are generated in the tube (ion burn). The phosphor host crystal structure can be disrupted by ion bombardment, and to a lesser extent by prolonged electron bombardment, creating absorbing centers (discoloration), poison centers, and other lattice defects which can reduce the efficiency of the luminescence. *See* LUMINESCENCE; SECONDARY EMISSION.

Herbert N. Hersh; James H. Schulman

Cauliflower

A cool-season biennial crucifer (*Brassica oleracea* var. *botrytis*) of Mediterranean origin. Cauliflower belongs to the plant order Capparales. It is grown for its white head or curd, a tight mass of flower stalks, which terminates the main stem (see **illus.**).

Head of cauliflower. (*Burpee Seeds*)

Cauliflower is commonly cooked fresh as a vegetable; to a lesser extent, it is frozen or pickled and consumed as a relish. Cultural practices are similar to those used for cabbage; however, cauliflower is more sensitive to unfavorable environment. Strains of the variety (cultivar) Snowball are most popular; purple-headed varieties are less common. Cauliflower is slightly tolerant of acid soils and has high requirements for boron and molybdenum. A cool, moist climate favors high quality. Harvest is generally 3–4 months after planting. California and New York are important cauliflower-producing states. *See* CABBAGE; CAPPARALES.

H. John Carew

Causality

In physics, the requirement that interactions in any space-time region can influence the evolution of the system only at subsequent times; that is, past events are causes of future events, and future events can never be the causes of events in the past. Causality thus depends on time orientability, the possibility of distinguishing past from future. Not all spacetimes are orientable.

Causality and determinism. The laws of a deterministic theory (for example, classical mechanics) are such that the state of a closed system (for example, the positions and momenta of particles in the system) at one instant determines the state of that system at any future time. Deterministic causality does not necessarily imply practical predictability. It was long implicitly assumed that slight differences in initial conditions would not lead to rapid divergence of later behavior, so that predictability was a consequence of determinism. Behavior in which two particles starting at slightly different positions and velocities diverge rapidly is called chaotic. Such behavior is ubiquitous in nature, and can lead to the practical impossibility of prediction of future states despite the deterministic character of the physical laws. *See* CHAOS.

Quantum mechanics is deterministic in the sense that, given the state of a system at one instant, it is possible to calculate later states. However, the situation differs from that in classical mechanics in two fundamental respects. First, conjugate variables, for example, position x and momentum p, cannot be simultaneously determined with complete precision, the relation between their indeterminacies being $\Delta x\ \Delta p \geq \hbar$, where $\hbar$ is Planck's constant divided by 2π. Second, the state variable ψ gives only probabilities that a given eigenstate will be found after the performance of a measurement, and such probabilities are also all that is calculable about a later state ψ' by the deterministic prediction. Despite its probabilistic character, the quantum state still evolves deterministically. However, which eigenvalue (say, of position) will actually be found in a measurement is unpredictable. *See* DETERMINISM; EIGENVALUE (QUANTUM MECHANICS); NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; QUANTUM THEORY OF MEASUREMENT; UNCERTAINTY PRINCIPLE.

Causal structure of spacetimes. Nonrelativistic mechanics assumes that causal action can be propagated instantaneously, and thus that an absolute simultaneity is definable. This is not true in special relativity. While the state of a system can still be understood in terms of the positions and momenta of its particles, time order, as well as temporal and spatial length, becomes relative to the observer's frame, and there is no possible choice of simultaneous events in

the universe that is the same in all reference frames. Only space-time intervals in a fused "spacetime" are invariant with respect to choice of reference frame. The theory of special relativity thus rejects the possibility of instantaneous causal action. Instead, the existence of a maximum velocity of signal transmission determines which events can causally influence others and which cannot. *See* SPACE-TIME.

The investigation of a spacetime with regard to which events can causally influence (signal) other regions and which cannot is known as the study of the causal structure of the spacetime. Thus, in the Minkowski spacetime of special relativity, an event E can causally affect another E' if and only if there is a timelike curve which joins E and E', and E' lies in the future of E. Such a curve is contained within the future light cone of E and connects E and E'. The light-cone surface is generated by null geodesics representing the velocity of light. One of the two halves of the light cone intersecting at the "present" is specified as the future, the other as the past. Paths lying within the light cone are timelike, those in the past region (or the present) being capable of influencing later events. Curves joining E with events outside the light cone are called spacelike. Their traversal requires velocities greater than that of light, and thus events which are spacelike-related cannot causally influence one another.

Closed causal curves. Different spacetimes, for example, those allowed by general relativity, are distinguished by their different causal structures. Because general relativity admits distinct spacetime metrics at adjacent points, the future directions of light cones can vary over short distances. (Light cones at adjacent points tilt with respect to one another.) Under these circumstances, it is possible for a continuous sequence of tilted cones to result in a timelike curve intersecting itself, producing a closed curve. An event on such a curve, however, both precedes and succeeds itself, and can be both its own cause and effect. Some solutions of the general relativistic field equations, for example, Gödel and Taub-Nut spacetimes, contain such closed causal curves.

However, many physicists hold that the existence of such curves is unphysical, and seek criteria to exclude them. Among the many such conditions that have been extensively discussed are the following. The causality condition excludes closed nonspacelike (that is, null or timelike) curves. Strong causality further excludes "almost closed" causal curves, wherein a nonspacelike curve returns more than once to the same infinitesimal neighborhood. Need for yet a further condition arises from the possibility that, in a quantum theory of gravity, the uncertainty principle would prevent the metric from having an exact value at every point, leading to the possibility that small variations in the metric would generate closed timelike curves. The condition of causal stability prevents such occurrences by defining a neighborhood of a point in which there are no closed timelike curves. Causal stability is the strongest condition excluding causal anomalies: violations of weaker conditions necessarily violate causal stability. If a spacetime is causally stable, the topology of its manifold follows from the causal structure, as do the differentiable and the conformal structures. Clearly a study of the causal structure gives deep insight into the characteristics of a spacetime.

In contrast, there are also many physicists who do not consider solutions of the equations of general relativity which contain closed causal curves to be unphysical, pointing out that, in the past, possibilities treated as unphysical have frequently turned out to be physically significant. For this reason (among others), they hold that solutions with closed causal curves should be taken seriously, and should be included among, for example, the possible histories of the universe in path-integral calculations in quantum theories of gravity.

Other causality violations. Violations of conventional causality could (hypothetically) arise in ways other than through closed causal curves. For example, the possibility has been considered that tachyons, faster-than-light particles, might exist. This is tantamount to the speculation that the past (or present) could be influenced by future events through the transmission of tachyons. Existence of such particles is generally rejected on both experimental and theoretical grounds. Causality violation would also result if there were more than one dimension of time, as there are of space. Such assignments are therefore usually excluded in theories of quantum gravity. *See* QUANTUM GRAVITATION; RELATIVITY; TACHYON.

Dudley Shapere

Bibliography. R. Geroch and G. Horowitz, Global structure of spacetimes, in S. Hawking and W. Israel, *General Relativity,* 1979; S. Hawking and G. Ellis, *The Large-Scale structure of Space-Time,* 1973; P. S. Joshi, *Global Aspects in Gravitation and Cosmology,* 1993; R. Torretti, *Relativity and Geometry,* 1983; R. Wald, *General Relativity,* 1984.

Cave

A natural cavity located underground or in the side of a hill or cliff, generally of a size to admit a human. Caves occur in all types of rocks and topographic situations. They may be formed by many different erosion processes. The most important are created by ground waters that dissolve the common soluble rocks—limestone, dolomite, gypsum, and salt. Limestone caves are the most frequent, longest, and deepest. Lava-tube caves, sea caves created by wave action, and caves caused by piping in unconsolidated rocks are the other important types. The science of caves is known as speleology. *See* DOLOMITE; GYPSUM; HALITE; LIMESTONE.

Dissolution caves. The principal processes of aqueous dissolution are simple molecular dissociation for salt (NaCl, and other water-soluble salts) and gypsum ($CaSO_4 \cdot 2H_2O$), and bicarbonate solution for limestone ($CaCO_3 + H_2O + CO_2 = Ca^{2+} + 2HCO_3$) and dolomite ($CaMg \cdot 2HCO_3$). Dissociation of salt proceeds until an equilibrium concentration of 360 g/liter is achieved in standard conditions; that

is, there is very rapid creation and destruction, with the consequence that salt caves survive only in arid regions and are rarely longer than 1 km (0.6 mi). Gypsum solutions equilibrate at about 2.4 g/liter, allowing enterable caves to develop in tens to hundreds of years where there is abundant water and groundwater hydraulic gradients are steep. Limestone (calcite) and dolomite solutions normally equilibrate at 0.1–0.3 g/liter; such caves require 10^3–10^6 years to be initiated and then enlarged for human entry, but then may also survive for many millions of years. *See* CALCITE; GROUND-WATER HYDROLOGY.

Limestone and dolomite dissolution is augmented or accelerated at certain locations by addition of carbon dioxide (CO_2) or hydrogen sulfide (H_2S) [oxidizing to sulfuric acid (H_2SO_4)] from deep crustal sources, and along sea coasts where mixing of fresh and salt water enhances the solubility. Microbiological processes (nanobacteria producing CO_2) may also be important. Condensation corrosion by carbonic acid (H_2CO_3) or H_2SO_4 can be significant where periodic high humidity combines with excess release of CO_2 or H_2S in some coastal or arid-zone caves.

Types. Dissolution caves have been classified by their plan morphology (see **table**). In most regions the caves are formed by meteoric water recharging through karst depressions, creating systems that are combinations of the fracture and bedding plane types. Significant karst caves from intergranular dissolution are uncommon, as are caves created by diffusion beneath porous but insoluble sandstones. Spongeworks and rudimentary networks are associated mostly with coastal mixing-zone dissolution in young (late Tertiary–Quaternary) limestones. Hypogene (ascending, deeper crustal) waters sometimes create a single-stem, treelike cave in their rise to the surface. More often they ramify into networks of fissures and rooms; the greatest known examples have invaded filled paleokarst cavities of different kinds. *See* KARST TOPOGRAPHY.

There are four principal varieties of common meteoric water caves (see **illus.**). Where densities of penetrable fissures are low or stratal dips are steep, the cave system may consist largely of a single deep (phreatic zone) loop beneath a water table that was lowered, then stabilized, by expansion of the cave (illus. *a*). Where fissure frequency is high or

Table of dissolution caves

		Type of Recharge				
		via karst depressions		diffuse		hypogenic
		sinkholes (limited discharge fluctuation)	sinking streams (great discharge fluctuation)	through sandstone	into porous soluble rock	dissolution by acids of deep-seated source or by cooling of thermal water
		branchworks (usually several levels) and single passages	single passages and crude branchworks, usually with the following features superimposed:	most caves enlarged further by recharge from other sources	most caves formed by mixing at depth	
Type of Presolutional Porosity	Fractures	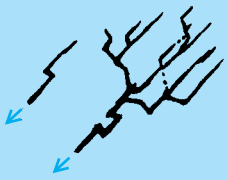 angular passages	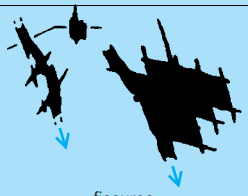 fissures, irregular networks	fissures, networks	isolated fissures and rudimentary networks	networks, single passages, fissures
	Bedding partings	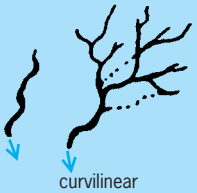 curvilinear passages	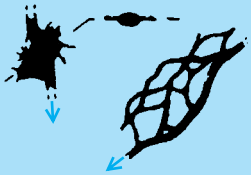 anastomoses, anastomotic mazes	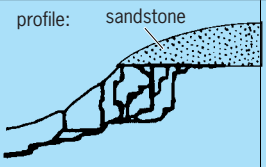 shaft and canyon complexes, interstratal solution	spongework	ramiform caves, rare single-passage and anastomotic caves
	Intergranular	rudimentary branchwork	spongework	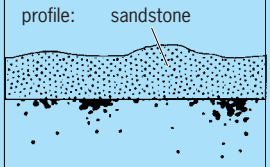 rudimentary spongework	spongework	ramiform and spongework caves

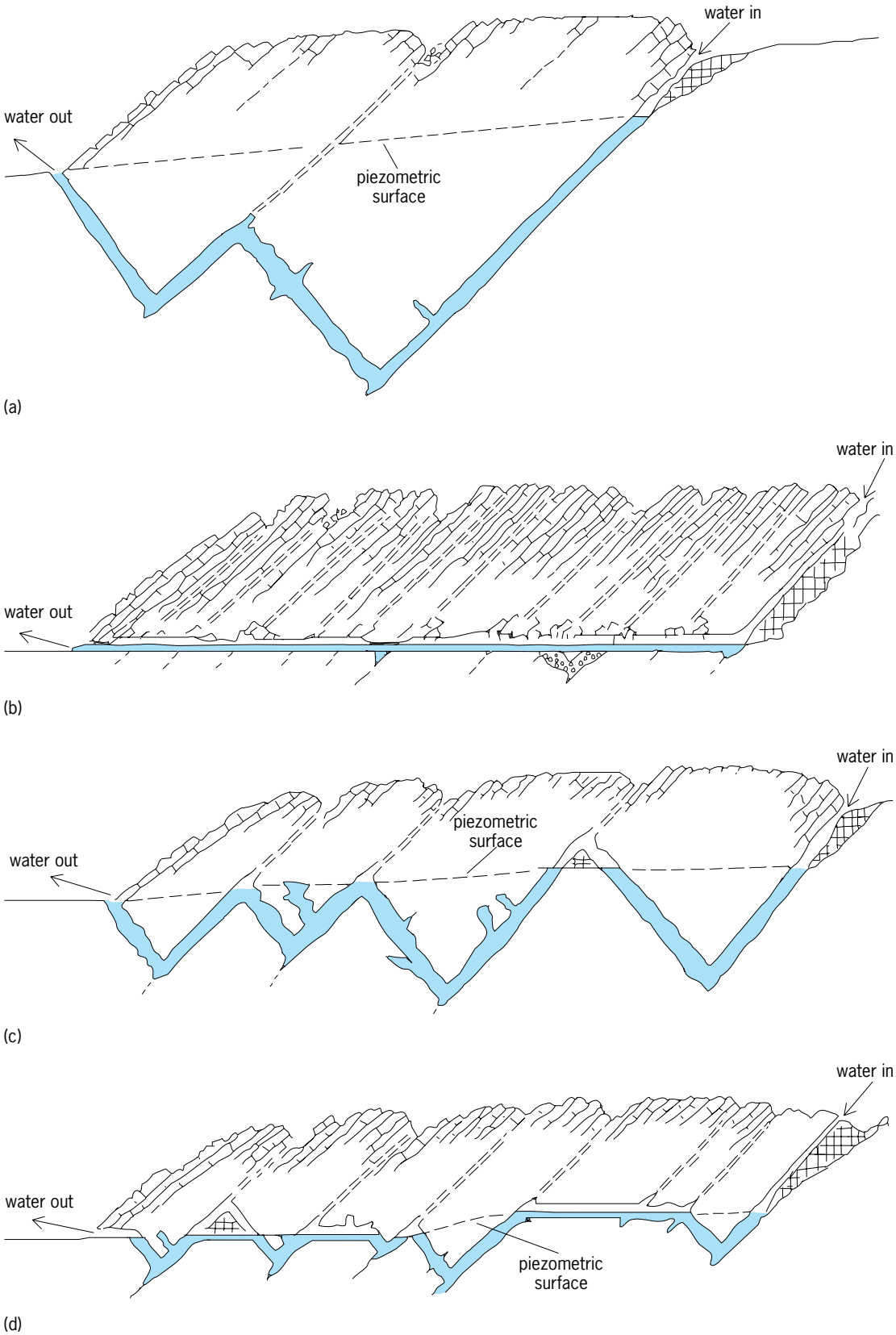

Principal types of meteoric water caves. The piezometric surface is an imaginary surface that represents the static head of ground water and is defined by the level to which the water will rise. (*a*) Bathyphreatic cave. (*b*) Ideal water-table cave. (*c*) Phreatic cave with multiple loops. (*d*) Cave with mixture of phreatic and water-table-leveled components. The letters represent the distribution (density) of major penetrable fissures in which the cave will develop. As fissure density increases, the vertical amplitude of looping in the cave diminishes. (*From D. C. Ford and P. W. Williams, Karst Geomorphology and Hydrology, Routledge, Chapman and Hall, 1989*)

there is significant intergranular porosity, or where strata are nearly flat lying, a subhorizontal, water-table passage may collect waters from higher (vadose zone) feeder galleries that were drained by the early enlargement of the cave (illus. *b*). There are two intermediate types (illus. *c*, *d*). Most of the larger cave systems that are known also display multiple phases (levels) in their development.

In the United States, Mammoth Cave National Park, Kentucky, provides excellent examples of meteoric water caves. Jewel Cave and Wind Cave, South Dakota, are superb examples of thermal water caves. Carlsbad Caverns, New Mexico, are outstanding for sulfur corrosion and magnificent speleothems.

Passages. In cross section (as they are seen by visitors), meteoric water passages display three basic shapes, alone or in combination. Phreatic passages are created below a water table, where pressure flow applies. Their ideal shape is circular (that is, minimum-friction cross section, as in a domestic pipe), but more often there is an ellipse extended along the host fissure. Vadose passages develop in the aerated zone where gravity flow prevails; shapes are varieties of entrenched canyons interspersed with fluted shafts, and there may be widening at the base to undercut the walls. In breakdown passages, parts or all of the walls or ceilings are collapsed into blocks or rubble. Breakdown is caused by loss of buoyant support upon draining, by vadose stream undercutting, by tributary waters descending through roofs, as well as other processes. Most large underground rooms are developed by progressive breakdown.

Shapes in hypogene and mixing-zone caves are phreatic, but with a greater frequency of deep, rounded dissolution alcoves (pockets) than is common in meteoric water caves. Fissure passages taper rapidly in some hydrogen sulfide caves. A swiss-cheese texture is often reported in the walls of mixing-zone caves.

Large systems. The longest known cave system is Mammoth Cave, Kentucky, a meteoric water cave in near-horizontal strata; more than 500 km (300 mi) of interconnected passages have been mapped. Phreatic tubes, vadose, and breakdown rooms are all common in this system, which has evolved through at least four phases at successively lower spring levels. Other caves exceeding 100 km (60 mi) include hypogene mazes such as Jewel Cave and Wind Cave, South Dakota, which are in limestone and dolomite, and Optimists' and Ozernaya, Ukraine, which are contained in a thin gypsum formation. Exploration in the Americas, western Europe, China, and southeastern Asia is revealing many more meteoric caves around or approaching the 100-km (60-mi) mark. Flooded (still phreatic) caves have been explored to lengths of 8 km (5 mi) in Australia, Europe, and Florida. The deepest known caves occur in high mountain areas such as the Alps, Pyrenees, Caucasus, and southern Mexico: Reseau Bernard (France) is 1590 m (5170 ft) deep, and there are at least 20 others exceeding 1000 m (3300 ft). Underwater, great springs at Mante (Mexico) and Vaucluse (France) have been explored by divers to depths of 250 m (825 ft) and 315 m (1040 ft) respectively; the latter was investigated with a remotely controlled submersible. The greatest room is in Lubang Nasib Bagus Cave, Sarawak: it measures 720 m (2400 ft), long, 400 m (1300 ft) wide, and 70 m (230 ft) high and has a volume of approximately 20,000,000 m^3 (700,000,000 ft^3). It is far larger than any space built by humans.

Cave deposits. Caves are important sediment traps, preserving evidences of past erosional, botanic, and other phases that may be obliterated aboveground. Clastic sediments include breakdown, weathering earths, water-laid gravels, sands, silts and clays, windborne dust, colluvium, beach deposits, and even tills injected from glaciers. Organic materials of every kind can be trapped in them. The finer-grained deposits may also contain good paleomagnetic records. *See* EROSION; PALEOMAGNETISM; WEATHERING PROCESSES.

Chemical deposits are very important. More than 100 different minerals are known to precipitate in caves. Most abundant and significant are stalactites, stalagmites, and flowstones of calcite. These may be dated with uranium series methods, thus establishing minimum ages for the host caves. They contain paleomagnetic records. Their oxygen and carbon isotope ratios and trapped organic materials may record long-term changes of climate and vegetation aboveground that can be dated with great precision. As a consequence, cave deposits are proving to be among the most valuable paleoenvironmental records preserved on the continents. *See* STALACTITES AND STALAGMITES.

All kinds of faunal remains are found in cave deposits, including human artifacts and bones. Most known sites of early *Homo sapiens, H. neanderthalensis*, and *H. erectus* are caves. Cave fauna specially adapted to the permanent darkness of underground are termed troglodytes. They include many species of fishes, insects, and crustaceans. Eyeless animals are common. *See* FOSSIL HUMANS.

Derek C. Ford

Bibliography. W. Dreybrodt, *Processes in Karst Systems: Physics, Chemistry and Geology*, 1988; D. C. Ford et al., Uranium series dating of the draining of an aquifer: The example of Wind Cave, Black Hills, S. D., *Bull. Geol. Soc. Amer.*, 105:241–250, 1993; D. C. Ford and P. W. Williams, *Karst Geomorphology and Hydrology*, 1989; C. A. Hill and P. Forti, *Cave Minerals of the World*, 2d ed., 1997; A. N. Palmer, Origin and morphology of limestone caves, *Bull. Geol. Soc. Amer.*, 103:1–21, 1991; W. B. White, *Geomorphology and Hydrology of Karst Terrains*, 1988; I. J. Winograd et al., Continuous 500,000 year climate record from vein calcite in Devils Hole, Nevada, *Science*, 258:255–260, 1992.

Cavies

Rodents comprising the family Caviidae, which includes the guinea pig, rock cavies, mountain cavies, capybara, salt-desert cavy, and mara. All members of the group are indigenous to South America and

The guinea pig (*Cavia aperea*).

comprise 15 species in six genera. Cavies have either rounded bodies with large heads and short ears and limbs, or rabbitlike bodies with long limbs and moderately long ears. They have 20 teeth with short incisors and a dental formula of I 1/1 C 0/0 Pm 1/1 M 3/3.

The guinea pig (*Cavia aperea*) originated in Peru, where there is still a wild stock. The domestic form (*C. porcellus*) [see **illus.**] has been produced by selective breeding and is a valuable laboratory animal with a life-span of 3–5 years. They are stocky animals; adults measure 8–12 in. (20–30 cm) in length and weigh less than 2 lb (0.9 kg). There are two or three litters each year with two to six young born after a gestation period of 2 months. The young cavies can see at birth and are fairly independent of the mother. These animals are docile, are easily reared in captivity, and make excellent pets.

The guinea pig has been used as an experimental animal in numerous studies. Geneticists have studied coat color and hair length and type, and have carried out numerous studies on the effects of inbreeding. Mating behavior in the guinea pig has been thoroughly studied by ethologists; in fact, it is probably the best-known insofar as mating behavior is concerned. Immunological studies using this animal are continually conducted.

A large cavy which is closely related to the guinea pig is the mara or Patagonian cavy (*Dolichotis patagonum*). It resembles a large hare, being 1 ft (0.3 m) high and having a length of 2–3 ft (0.6–0.9 m). It is a burrowing, herbivorous animal that lives in small colonies in drier areas of Argentina and Patagonia. The mara is active during the day, and is quite swift as it runs and leaps. Two litters of usually three offspring each are born in the burrows each year. The salt-desert cavy (*Pediolagus salinicola*) is a smaller species found in the salt deserts of southern Argentina.

The largest of all rodents is the capybara or carpincho (*Hydrochoerus hydrochaeris*). It grows to the size of a small pig, about 4 ft (1.2 m) long and 120 lb (54 kg). It is essentially an aquatic animal which lives in small groups along the banks of lakes and streams in tropical South America. It is a vegetarian, is easily tamed, and breeds well in captivity. The female produces one litter each year after a gestation period of 4 months with from three to eight self-sufficient young; domestication increases fertility. The life-span of the animal is about 10 years. *See* RODENTIA.

Charles B. Curtin

Cavitation

The formation of vapor- or gas-filled cavities in liquids. If understood in this broad sense, cavitation includes the familiar phenomenon of bubble formation when water is brought to a boil under constant pressure and the effervescence of champagne wines and carbonated soft drinks due to the diffusion of dissolved gases. In engineering terminology, the term cavitation is used in a narrower sense, namely, to describe the formation of vapor-filled cavities in the interior or on the solid boundaries created by a localized pressure reduction produced by the dynamic action of a liquid system without change in ambient temperature. Cavitation in the engineering sense is characterized by an explosive growth and occurs at suitable combinations of low pressure and high speed in pipelines; in hydraulic machines such as turbines, pumps, and propellers; on submerged hydrofoils; behind blunt submerged bodies; and in the cores of vortical structures. This type of cavitation has great practical significance because it restricts the speed at which hydraulic machines may be operated and, when severe, lowers efficiency, produces noise and vibrations, and causes rapid erosion of the boundary surfaces, even though these surfaces consist of concrete, cast iron, bronze, or other hard and normally durable material.

Acoustic cavitation occurs whenever a liquid is subjected to sufficiently intense sound or ultrasound (that is, sound with frequencies of roughly 20 kHz to 10 MHz). When sound passes through a liquid, it consists of expansion (negative-pressure) waves and compression (positive-pressure) waves. If the intensity of the sound field is high enough, it can cause the formation, growth, and rapid recompression of vapor bubbles in the liquid. The implosive bubble collapse generates localized heating, a pressure pulse, and associated high-energy chemistry. *See* SOUND; ULTRASONICS.

Bernoulli's principle. As mentioned, cavitation occurs when the pressure in a liquid is reduced to a critical value. For the present, it will be assumed that this critical value is the vapor pressure p_v of the liquid. For clean, fresh water at 70°F (21°C), p_v has a value of about 52 lb/ft^2 (2.5 kilopascals); hence, when a liquid is moving with velocity V over a body at ordinary temperature and the pressure on the body surface is reduced to or near 52 lb/ft^2 (2.5 kPa), cavitation may be expected to occur. The condition for the onset of cavitation is therefore given by relation (1), where

$$p_m \leq p_v \tag{1}$$

p_m is the minimum pressure at any point on the surface of a moving body and p_v is the vapor pressure of the liquid at the prevailing temperature. Inversely, the condition for avoidance of cavitation is given by relation (2).

$$p_m > p_v \tag{2}$$

See VAPOR PRESSURE.

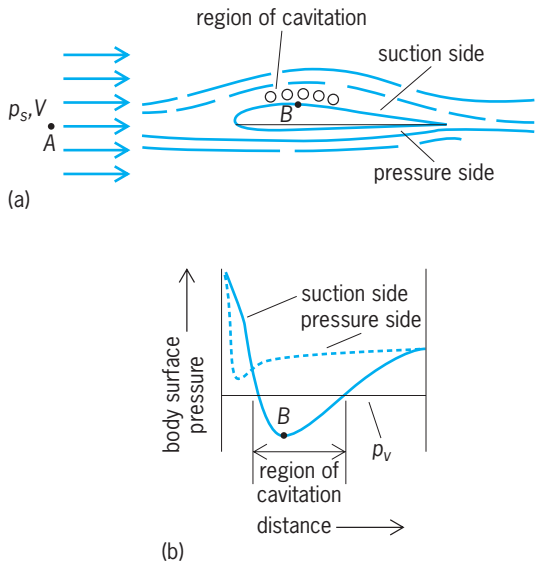

Fig. 1. Flow around a hydrofoil in a frictionless, incompressible fluid. (*a*) Cross section of flow. At point *A*, p_s = absolute static pressure, *V* = velocity of undisturbed flow. (*b*) Variation of body surface pressure along the two sides of the hydrofoil. p_v = vapor pressure of fluid.

A more practical relation than (2) is obtained when the pressure p_m is expressed in terms of easily measurable reference values. This can be done by use of Bernoulli's principle. According to this principle, the sum of pressure head and velocity head in a frictionless incompressible medium remains constant along a streamline. This may be understood by considering a hydrofoil in a stream flowing from left to right with constant velocity V (**Fig. 1**). Writing Bernoulli's equation for two points A and B lying on the same streamline, where the local pressure is minimum at point B, yields Eq. (3), where p_m = absolute local

$$p_m = p_s - \frac{\rho}{2}V^2\left[\left(\frac{v}{V}\right)^2 - 1\right] \tag{3}$$

pressure at point B, v = velocity at point B, V = velocity of undisturbed flow at point A, p_s = absolute static pressure at point A, and ρ = density of liquid. The term in the square brackets is independent of the absolute speed V, because v changes proportionally to V. It follows that for a constant value of p_s the pressure p_m may be reduced to the vapor pressure p_v, to zero, or even to negative values by increasing the speed V. Negative values for p_m have been obtained in carefully conducted laboratory tests, proving that a pure liquid, such as clean air-free water, may sustain tension. However, in most practical cases, negative values are not obtained; instead, the flow is disrupted when cavities are formed. Combining (2) and (3), the condition for avoidance of cavitation becomes relation (4).

$$\frac{p_s - p_v}{(\rho/2)V^2} > \left[\left(\frac{v}{V}\right)^2 - 1\right] \tag{4}$$

See BERNOULLI'S THEOREM.

Cavitation number. The term on the left of inequality (4) contains easily measurable values and is usually denoted by σ, which is called the cavitation number. The magnitude of the term on the right can be calculated for simple bodies, but for more complex configurations it cannot be calculated and must be obtained by experiment. Denoting this value by σ_c, the condition for avoidance of cavitation has the form of relation (5).

$$\sigma = \frac{p_s - p_v}{(\rho/2)V^2} > \sigma_c \tag{5}$$

The cavitation number σ is used for flow through pipes, flow around submerged bodies, and in the design of marine propellers. In pump and hydraulic turbine work, slightly different expressions are used. The simplest one is the expression first introduced by D. Thoma which has form of relation (6), where

$$\sigma_T = \frac{H_{sv}}{H} > (\sigma_T)_c \tag{6}$$

H_{sv} is the net positive suction head at the pump inlet, or just below the turbine runner, and H is the total head under which the turbine or pump operates. The value of $(\sigma_T)_c$, like the value of σ_c, is a fixed number for a given design of pump or turbine which in general must be found by experiment.

Experiments to determine σ_c and $(\sigma_T)_c$ are usually made on models geometrically similar to, but smaller than, the prototype installations. This has the advantage of less cost and more precise control of the experiments and permits correction of undesirable characteristics of a design before the prototype machine is actually constructed. For instance, should it be found in a model test that a given propeller design cavitates heavily at the design operating value, different combinations of diameter and pitch, revolutions, blade width and outline, or section shape may be tried to eliminate or alleviate the observed cavitation. The same procedure is followed in the case of pumps and turbines.

The term scale effect is given to any deviations from the elementary similarity relations linking cavitation number to geometric and kinematic conditions. The many factors producing scale effects can be divided into two types: those which affect the minimum pressure in the liquid flow, and those which cause the cavitation pressure to be different from the equilibrium vapor pressure. *See* DYNAMIC SIMILARITY.

Types. It has been found convenient to differentiate between a type of cavitation in which small bubbles suddenly appear on the solid boundary, grow in extent, and disappear, and another type in which cavities form on the boundary and remain attached as long as the conditions that led to their formation remain unaltered. The former type is known as bubble or traveling cavitation (**Fig. 2***a*) and the latter as sheet or attached cavitation (Fig. 2*b*). Tip vortex cavitation (Fig. 2*b*) occurs in the low-pressure center of a vortex formed at the tip of a lifting blade.

Physical causes. Both experiments and calculations show that with ordinary flowing water

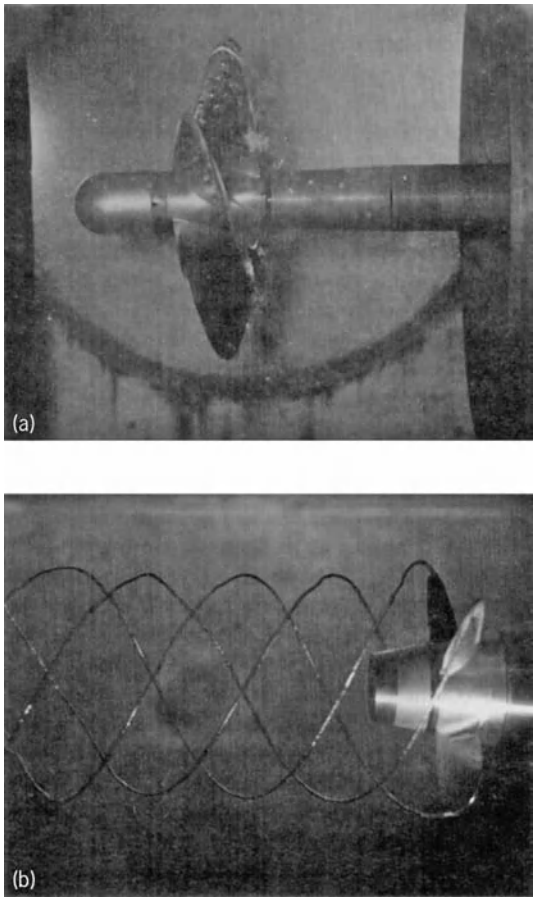

Fig. 2. Examples of cavitation on screw propellers in water tunnels. (*a*) Bubble or traveling cavitation (*U.S. Navy*). (*b*) Sheet and tip vortex cavitation (*Applied Research Laboratory, Pennsylvania State University*)

cavitation commences as the pressure approaches or reaches the vapor pressure, because of impurities in the water. These impurities, called cavitation nuclei, cause weak spots in the liquid and thus prevent it from supporting higher tensions. The various forms of cavitation nuclei are classified in two general groups, stream nuclei and surface nuclei. Stream nuclei exist in the liquid in such forms as inorganic or organic particulates or microbubbles. Surface nuclei originate at the surface of the fluid boundary by means of cracks and crevices in the boundary. Several techniques to measure stream nuclei have been standardized, and two significantly different approaches have developed. One is to measure the particulate or microbubble distribution by utilizing acoustical (wave-velocity, tone-attenuation), electrical (Coulter counter), or optical (photography, holography, laser light-scattering) techniques. The other approach measures a cavitation event rate for a liquid under various tensions (measured by a head-form venturi tube) and establishes a cavitation susceptibility.

The exact mechanism of bubble growth is generally described by mathematical relationships which depend upon the cavitation nuclei. Cavitation commences when these nuclei enter a low-pressure region where the equilibrium between the various forces acting on the nuclei surface cannot be established. As a result, bubbles appear at discrete spots in low-pressure regions, grow quickly to relatively large size, and suddenly collapse as they are swept into regions of higher pressure.

Effects. It was mentioned initially that cavitation produces noise, and, when severe, lowers the efficiency of a machine, and causes rapid erosion of boundary surfaces. The destructive effect of cavitation on solid surfaces may be explained as follows. When cavitation bubbles form in a low-pressure region, the growth is explosive, with high bubble wall velocity. The growth time interval is too short for much air or gas to come out of solution, so that the bubbles are filled primarily with vapor. On subsequent collapse in the high-pressure region, the liquid particles rush toward the center of the bubble virtually unimpeded and form a very high velocity jet. Impingement of this jet and the resulting pressure wave on a solid boundary cause very high impulsive forces. It is estimated that the surface stress caused by the impingement is of the order of 1000 atm (100 megapascals), which is sufficiently high to cause fatigue failure of the material in a relatively short time. Some investigators hold that the explosive formation of the bubbles, intercrystalline electrolytic action, and the collapse of the bubbles all are factors contributing to the observed destruction. Michael L. Billet

Supercavitating propellers. The limitation on ship speed caused by loss of thrust when cavitation is severe has been overcome by a radical departure from conventional propeller designs. In this new design (**Fig. 3**), cavitation on the backs of the blades (forward side) is induced by special blade sections at relatively low forward speed so that, when revolutions and engine power are increased, the whole back of each blade becomes enveloped by a sheet of cavitation. When this is completed, further increase in thrust at still higher engine power and revolutions per minute is obtained by the increase in positive pressure on the blade face (rear side); erosion is avoided because the collapse of the cavitation bubbles occurs some distance behind the trailing edges of the blades. The blade sections are usually wedge-shaped with a sharp leading edge to initiate cavitation at this point, a blunt trailing edge, and a concave face. The supercavitating propeller is no replacement for the conventional propeller,

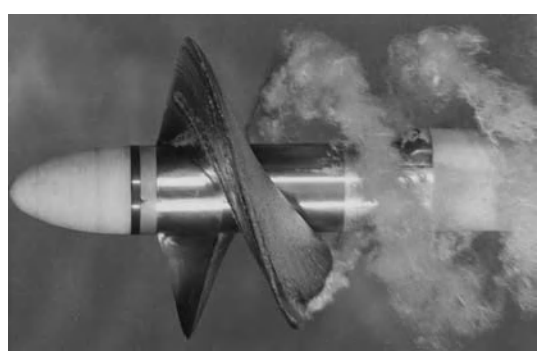

Fig. 3. Supercavitating propeller in water tunnel. (*U.S. Navy*)

being suitable only for very high ship and engine speeds.

In 1968 the supercavitating propeller was applied in two experimental hydrofoil boats: the *Denison* constructed for the U.S. Maritime Administration and another constructed for the Royal Canadian Navy. Probably it has also been used on patrol boats of various navies, but no performance data have been published. In the known installations the propellers functioned as anticipated, but some difficulties have been encountered in finding materials for these propellers strong enough to withstand the very high blade stresses. *See* PROPELLER (MARINE CRAFT). Jacques B. Hadler

Sonochemistry. The chemical and biological effects of ultrasound were first reported by A. L. Loomis in the 1930s. With the advent of inexpensive and reliable sources of ultrasound, there has been a resurgence of interest in the chemical applications of ultrasound. Ultrasound has been found to enhance a very wide range of chemical reactions, and there has been special interest in organometallic reactions. It has been used to increase rates and yields for both stoichiometric and catalytic reactions, both in homogeneous liquids and in liquid–solid heterogeneous systems.

Since ultrasound has acoustic wavelengths of roughly 0.15–75 mm, clearly no direct coupling of the acoustic field with chemical species on a molecular level can account for sonochemistry. Instead, the chemical effects of ultrasound derive from several different physical mechanisms, depending on the nature of the system. The most important of these is cavitation. The dynamics of cavity growth and collapse are strikingly dependent on local environment, and cavitation in a homogeneous liquid should be considered separately from cavitation near a solid interface.

In homogeneous media, the generally accepted sonochemical mechanism involves pyrolysis by a localized hot spot due to the adiabatic heating which is produced by the implosive collapse of a bubble during cavitation. A measurement of the temperature generated during this implosive collapse established that the effective temperature in the gas-phase reaction zone is about 8900°F (5200 K) with pressures of hundreds of atmospheres (tens of megapascals).

When a liquid–solid interface is subjected to ultrasound, cavitation still occurs, but with major changes in the nature of the bubble collapse. No longer do cavities implode spherically. Instead, a markedly asymmetric collapse occurs, which generates a jet of liquid directed at the surface, just as in propeller-induced cavitation. The jet velocities are greater than 300 ft/s (100 m/s). The origin of this jet formation is essentially a shaped-charge effect. The impingement of this jet creates a localized erosion responsible for surface pitting and ultrasonic cleaning. Enhanced chemical reactivity of solid surfaces is associated with these processes. The cavitational erosion generates unpassivated, highly reactive surfaces; causes short-lived high temperatures and pressures at the surface; produces surface defects and deformations; creates high-velocity interparticle collisions; increases the surface area of friable solid supports; and ejects material in unknown form into solution. Finally, the local turbulent flow associated with acoustic streaming improves mass transport between the liquid phase and the surface, thus increasing observed reaction rates. In general, all of these effects are likely to occur simultaneously. *See* SEMICONDUCTOR. Kenneth S. Suslick

Bibliography. C. E. Brennen, *Cavitation and Bubble Dynamics*, 1995; F. G. Hammitt, *Cavitation and Multiphase Flow Phenomena,* 1980; K. S. Suslick (ed.), *Ultrasound: Its Chemical, Physical and Biological Effects*, 1988; J. P. Tullis, *Hydraulics of Pipelines: Pumps, Valves, Cavitation and Transients*, 1989; F. R. Young, *Cavitation*, 2000.

Cavity resonator

An enclosure capable of resounding or resonating and thereby intensifying sound tones or electromagnetic waves. Resonance is the phenomenon which results when the frequency of the impressed driving force is the same as the natural vibration of the cavity. Vibrating rods, the tuning fork, musical instrument strings, radio and television channel tuners, and so forth, constitute resonating systems as well. The cavity resonator enclosure has a volume which stores energy oscillating between one form and another. In the case of sound, the oscillation is between displacement and velocity of particles. In the case of electromagnetic waves, the energy oscillates between the magnetic and the electric fields. *See* MUSICAL INSTRUMENTS; TUNING FORK; VIBRATION.

Sound-resonant pipes. Cavity pipes are used as resonators in musical instruments such as pipe organs and flutes to increase their sonority. The frequency of resonance is determined (to a degree of approximation) by the length of the pipe, by the velocity of sound at the ambient temperature and pressure, by the intensity of the driving force, and by the condition at the ends of the pipe: closed or open. The resulting frequency is related to multiples of quarter-wavelengths or half-wavelengths (depending on the end conditions) contained in the length of pipe. The driving force, if sufficiently strong, can force oscillations to occur at overtone frequencies which are higher multiples of the lowest or fundamental frequency, as well as at the fundamental. **Figure 1** illustrates simple cavity resonators for sound for a pipe open at one end and closed at the other. Other resonators have both ends open or both ends closed, with a small hole through which the driving force is introduced. For the case of both ends open as well as for the case of both ends closed, the resonant frequencies are those frequencies for which the pipe length is a whole number of half-wavelengths. *See* ACOUSTIC RESONATOR; SOUND.

Microwave-resonant cavities. At very high radio frequencies, losses due to radiation can be eliminated and resistive losses can be minimized by using closed resonant cavities instead of lumped-circuit

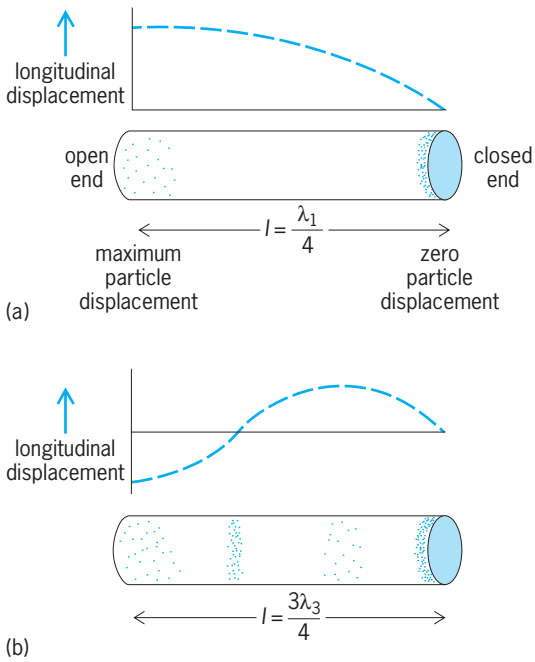

Fig. 1. Closed-end sound pipe resonating (*a*) at the fundamental frequency f_1 (the first harmonic) with wavelength λ_1, and (*b*) at an overtone, $f_3 \approx 3f_1$ (the third harmonic), with wavelength λ_3.

resonators. A cavity resonator stores both magnetic and electric fields, the energy oscillating between the two, losing energy only to the conducting walls if a perfect dielectric fills the space. The resonant frequency of the cavity is determined by the shape of the cavity and the mode, or allowable field distribution, of the electromagnetic energy that the cavity contains. Microwave transmission devices use such cavities. *See* KLYSTRON; MAGNETRON.

The usual cavities consist of closed sections of cylindrical or rectangular waveguides, or of hollow spheres or other symmetrical shapes. The choice of the shape is determined by the ease of fabrication or by the cavity's application. In the case of the cylindrical cavity resonator, the resonant frequency f_0 is computed from Eq. (1):

$$f_0 = \frac{c}{2L}\left[1 + \left(\frac{2L}{K_{nl}a}\right)^2\right]^{1/2} \qquad (1)$$

where a = radius
L = cavity length
c = velocity of light
K_{nl} = lth root of $J_n'(K) = 0$ for transverse electric (TE) modes
K_{nl} = lth root of $J_n(K) = 0$ for transverse magnetic (TM) modes (J_n is the nth order Bessel function and J_n' is its derivative)

If one of the end walls of the cavity is adjustable, the cavity can act as a wavemeter; that is, the resonance frequency can change as L is adjusted. **Figure 2** illustrates how this can be done. *See* MAGNETRON; WAVELENGTH MEASUREMENT.

For the case of a rectangular cavity, the resonance frequency is determined by Eq. (2), where *a, b,* and

$$f_0 = \frac{c}{2}\left[\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2 + \left(\frac{p}{d}\right)^2\right]^{1/2} \qquad (2)$$

d are dimensions of the cavity and *m, n,* and *p* are integers (or zero in certain cases) as determined by the mode of the enclosed fields. *See* MICROWAVE; WAVEGUIDE.

Coupling. Coupling to cavities may be accomplished by (1) introduction of a conducting probe or antenna oriented so it coincides with the direction of the electron field; (2) introduction of a conducting loop with plane normal to the magnetic field; (3) placement of a hole or iris between the cavity and a driving waveguide, the iris being placed so that a field component in the cavity mode has a common direction with one in the waveguide or other cavity; and (4) introduction of a pulsating electron beam passing through a small gap into the cavity and in the direction of the electric field of the cavity. The first three coupling methods are illustrated in Fig. 2*b*.

Quality factor. A good measure of the quality of a typical cavity resonator is expressed by the amount of stored energy as compared with the energy lost to the imperfectly conducting walls (and if present, to the imperfect dielectric of the cavity region).

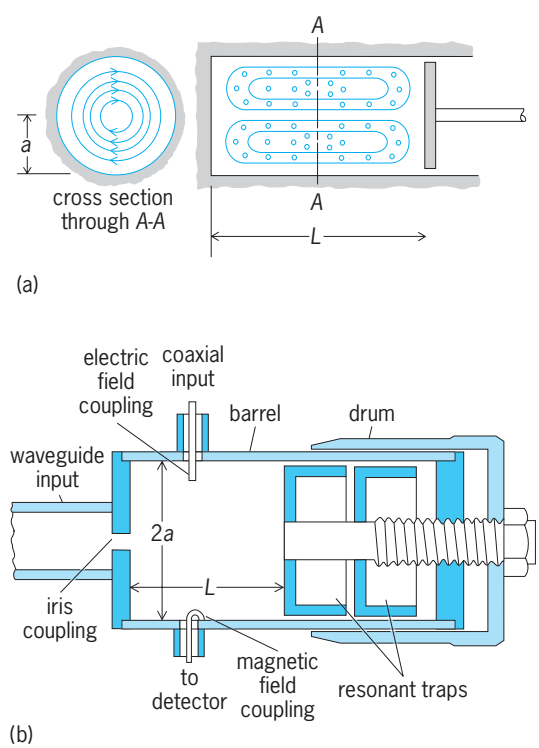

Fig. 2. Wavemeters. (*a*) Wavemeter consisting of cylindrical cavity with an adjustable end-wall, resonating in the TE_{011} mode. The free-space wavelength λ_0 corresponding to the resonant frequency f_0 is given by $\lambda_0 = 2L[1 + (2L/1.64a)]^{-1/2}$. (*b*) Practical wavemeter for the TE_{111} mode.

More precisely defined, the quality factor is represented by Eq. (3). The formula for any particular

$$Q = \frac{2\pi f_0 \times (\text{energy stored})}{\text{average power loss}}$$

$$= \frac{\pi(\text{energy stored})}{\text{energy loss per half cycle}} \qquad (3)$$

cavity configuration operating in a particular mode is derived by calculating the energy stored (either in the magnetic field or in the electric field), determining the losses due to the currents flowing on the imperfect conducting walls, and substituting in Eq. (3) for Q. The formula then contains the dimensions of the cavity relevant to the mode of the field configuration and an expression for the surface resistance of the conducting walls.

If a resonant cavity which has a quality factor Q is coupled to an external load which has a factor denoted by Q_e, the combination of the two leads to a loaded quality factor Q_L, which is related to the others by Eq. (4).

$$1/Q_L = 1/Q_e + 1/Q \qquad (4)$$

The Q can be measured by using two basic techniques. The first is to monitor the field intensity E as the frequency of the driving force is increased and decreased off the resonance of the cavity.

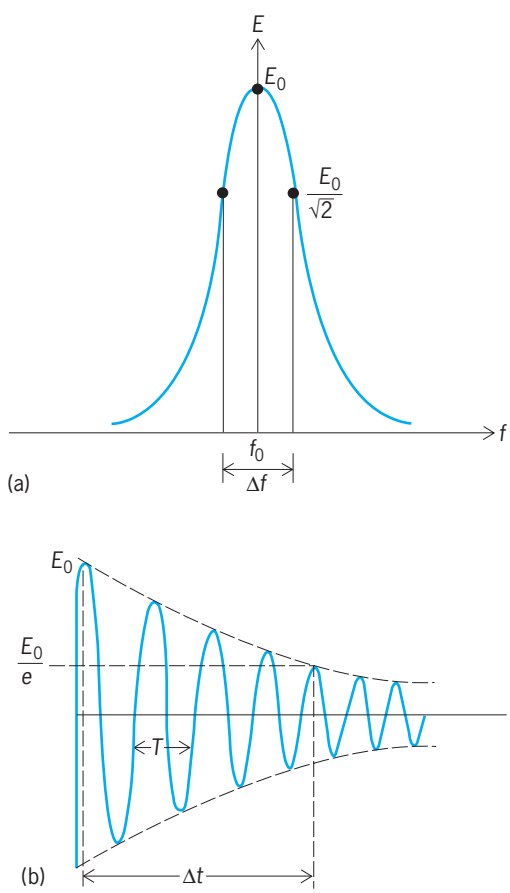

Fig. 3. Determination of the *Q* of a cavity resonator by (*a*) measuring the bandwidth and (*b*) measuring the damping rate of the oscillations.

Figure 3*a* shows the response of such a detuning. The two points representing a drop in field intensity to 0.707 or $1/\sqrt{2}$ of its peak value (half-power points) also represent the bandwidth, or Δf. Under this condition, it can be shown that Q is given by Eq. (5).

$$Q = \frac{f_0}{\Delta f} \qquad (5)$$

Another approach is to make use of the damping properties which indicate the rate at which oscillations would decay if the driving source were removed. For this method the fields E decay as given in Eq. (6), where E_0 is the initial field and $e = 2.718\ldots$.

$$E = E_0 e^{-(\pi f_0/Q)t} \qquad (6)$$

Consequently, if, as shown in Fig. 3*b*, a measure of the time taken for the field to drop to (1/*e*)th of its initial value is denoted as Δt, then Q is given by Eq. (7), and, since $1/f_0$ is the period of the oscillation T_0, then Eq. (8) is valid.

$$Q = \pi f_0 \Delta t \qquad (7)$$

$$Q = \pi \Delta t / T_0 = \pi \times (\text{number of cycles in } \Delta t) \qquad (8)$$

See Q (ELECTRICITY); RESONANCE (ALTERNATING-CURRENT CIRCUITS).

Optical cavity resonators. As the frequency of oscillation of the electromagnetic waves gets higher and well beyond the microwave range, the microwave cavity resonators described above become impractical. For lasers, with operation in the visible range, the infrared, or ultraviolet, the corresponding wavelengths are of the order of micrometers. The resonant cavity would have to be extremely small. If one was constructed with dimensions of millimeters, then the resonances of the possible modes of operation would be extremely close together and it would be impractical to separate them. Removing the sides of a closed cavity eliminates a large number of modes which in the absence of walls would radiate and thus be damped out. It has been shown that proper shaping and placing of the ends does support low-loss modes. Parallel plane mirrors is one system. Another consists of curved mirrors facing each other. Combinations of curved-planar mirrors, confocal mirrors, concentric mirrors, and spherical mirror configurations also constitute optical resonators. The objective in each design is to stabilize (contain) low-loss modes. *See* LASER.

The system of parallel plane reflectors has been used in interferometry and is referred to as a Fabry-Pérot resonator or etalon. *See* INTERFEROMETRY.

Optical resonators have also been made using loops of low-loss optical fibers, and in integrated optical circuits. *See* INTEGRATED OPTICS; OPTICAL FIBERS; RESONANCE (ACOUSTICS AND MECHANICS).

Diogenes J. Angelakos

Bibliography. R. E. Collin, *Foundations for Microwave Engineering*, 2d ed., 2000; J. D. Krause and K. R. Carver, *Electromagnetics*, 4th ed., 1992; M. I. Skolnik, *Radar Handbook*, 2d ed., 1990.

Cayley-Klein parameters

A set of four numbers used to specify the orientation of a body, or equivalently, the rotation R which produces that orientation, starting from some reference orientation. They can be expressed in terms of the Euler angles ψ, θ, and ϕ, as in Eqs. (1).

$$\begin{aligned} \alpha &= \cos\frac{\theta}{2}e^{-i(\psi-\phi)/2} \qquad \beta = -i\sin\frac{\theta}{2}e^{i(\psi-\phi)/2} \\ \gamma &= -i\sin\frac{\theta}{2}e^{-i(\psi-\phi)/2} \qquad \delta = \cos\frac{\theta}{2}e^{i(\psi-\phi)/2} \end{aligned} \tag{1}$$

Often the set δ, $-\beta$, $-\gamma$, α is used with slightly different properties from those given here. The four complex numbers contain eight real numbers and satisfy relations in Eqs. (2), where the bar denotes

$$\begin{aligned} \delta = \overline{\alpha} \qquad \gamma = -\beta \\ \alpha\delta - \beta\gamma = \alpha\overline{\alpha} + \beta\overline{\beta} = 1 \end{aligned} \tag{2}$$

complex conjugate. These constitute five real conditions, leaving three independent parameters as required. *See* EULER ANGLES.

The Cayley-Klein parameters combine in a simple way under compound rotations when they are arranged in the square matrix array shown in (3).

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \tag{3}$$

If α, β, γ, δ correspond to the rotation R_1 and α', β', γ', δ' to R_2, then the paramters α'', β'', γ'', δ'' of the rotation R_2R_1 (R_1 first, then R_2) are found by matrix multiplication. Thus Eq. (4) holds.

$$\begin{pmatrix} \alpha'' & \beta'' \\ \gamma'' & \delta'' \end{pmatrix} = \begin{pmatrix} \alpha' & \beta' \\ \gamma' & \delta' \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \tag{4}$$

A rotation $R(\psi, \theta, \phi)$ can be produced by three successive rotations. The corresponding decomposition of the parameter matrix is given by Eq. (5).

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{pmatrix}\begin{pmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}\begin{pmatrix} e^{-i\psi/2} & 0 \\ 0 & e^{i\psi/2} \end{pmatrix} \tag{5}$$

Although these parameters have been used to simplify somewhat the mathematic of spinning top motion, their main use is in quantum mechanics. There they are related to the Pauli spin matrices and represent the change in the spin state of an electron or other particle of half-integer spin under the space rotation $R(\psi, \theta, \phi)$. *See* MATRIX THEORY; SPIN (QUANTUM MECHANICS).

Bernard Goodman

Bibliography. H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; L. Shepley and R. Matzner, *Classical Mechanics*, 1991.

Caytoniales

A group of Mesozoic plants first recognized in 1925. The remains consist of palmately compound leaves with 3–6 lanceolate leaflets previously known as *Sagenopteris phillipsi*, pinnately branched microsporophylls (named *Caytonanthus arberi*) that bore winged pollen in four-chambered microsporangia, and fruit-bearing inflorescences, *Caytonia* (with two species), which bore a dozen or more globular, short-stalked fruits in subopposite rows. Each fruit was 0.2 in. (5 mm) or less in diameter. Near the attachment stalk at the lower part of the fruit was a small transverse lip beneath which was a minute depression marking the place where a small opening formed at pollination time. This opening admitted the pollen into the interior, where small orthotropous ovules were attached to the surface. The inflorescence thus resembled a simply pinnate frond, of which the individual pinnules were curved so as to enclose the seeds. When first discovered, it was believed that the flap served as a stigmatic surface to receive the pollen in a manner similar to the stigma of a modern flower. Later it was found that the pollen actually entered the young fruit through the small opening and germinated on the nucellus of the ovule after the manner in gymnosperms (Pinopsida). The Caytoniales appear related to the pteridosperms. They range from the Triassic to the Cretaceous. *See* PALEOBOTANY.

Chester A. Arnold

Cedar

Any of a large number of evergreen trees having fragrant wood of great durability. Arborvitae is sometimes called northern white cedar. *Chamaecyparis thyoides* is the botanical name of the southern white cedar, which may grow to 75 ft (22 m). It is characterized by small, spherical cones with peltate, seed-bearing scales fitted together to form a little ball, and by small scalelike evergreen leaves (**illus.** *a*). It grows only in swamps near the eastern coast of North America, where it is also known as Atlantic white cedar. The wood is soft, fragrant, and durable in the soil and is used for boxes, crates, small boats, tanks, woodenware, poles, and shingles. *See* ARBORVITAE; PINALES.

The Port Orford cedar (*C. lawsoniana*), also known as Lawson cypress, grows to 180 ft (54 m), has spherical cones larger than those of the eastern species, and is native to southwestern Oregon and northwestern California. It is the principal wood for storage battery separators, but is also used for venetian blinds and construction purposes. It is an ornamental tree which is sometimes used in shelter belts.

Alaska cedar (*C. nootkatensis*), found from Oregon to Alaska, grows to 120 ft (36 m), and has larger spherical cones which are nearly $^1/_2$ in. (0.6 cm) in diameter. The wood is used for interior finish, cabinetwork, small boats, and furniture. It is also grown as an ornamental tree.

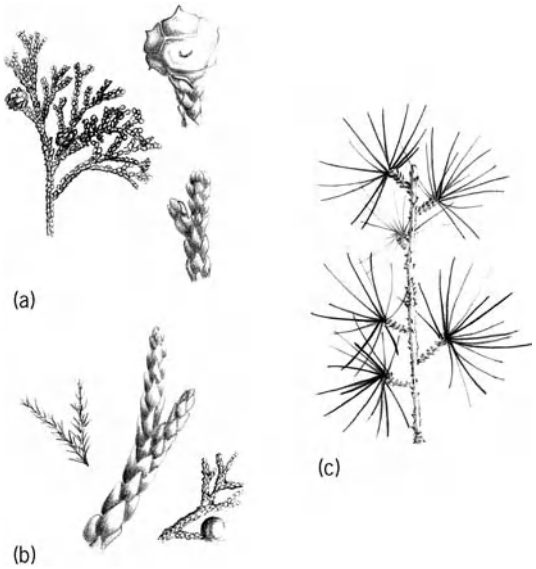

Characteristic leaf arrangements for three cedar species: (*a*) Southern white cedar (*Chamaecyparis thyoides*), (*b*) eastern red cedar (*Juniperus virginiana*), and (*c*) Atlas cedar (*Cedrus atlantica*).

Incense cedar (*Libocedrus decurrens*), a beautifully formed tree which grows to 135 ft (41 m), is found from Oregon to western Nevada and Lower California. The cones, $^3/_4$–1 in. (1.8–2.5 cm) long, resemble those of arborvitae. Incense cedar is one of the chief woods for pencils, and is also used for venetian blinds, rough construction, and fence posts and as an ornamental and shade tree.

Eastern red cedar (*Juniperus virginiana*), which may reach a height of 90 ft (27 m), is distributed over the eastern United States and adjacent Canada. Its leaves, evergreen and scalelike, occur in four rows lengthwise on the branchlets, giving a four-sided appearance (illus. *b*). On young trees or on vigorous shoots the leaves are needlelike, sometimes in threes. The fruit is spherical, berrylike, about $^1/_2$ in. (0.6 cm) in diameter, and borne on the female tree. Male trees bear only stamens. The very fragrant wood is durable in the soil and is used for fence posts, chests, wardrobes, flooring, and pencils. Cedarwood oil is used in medicine and perfumes.

Cedar of Lebanon (*Cedrus libani*) and Atlas cedar (*C. atlantica*) resemble the larch, but the leaves are evergreen and the cones are much larger and erect on the branches. The cedar of Lebanon, a native of Asia Minor, has cones 3–4 in. (7.5–10 cm) long and $1^1/_2$–$2^1/_2$ (4.3–6.8 cm) in diameter, and is sometimes cultivated in the United States. The Atlas cedar is hardier, and the variety *glauca* is popular as an ornamental (illus. *c*). *See* LARCH.

The deodar cedar (*Cedrus deodara*), a beautiful tree with large leaves and drooping branches, is a native of India, and although not hardy in the northern United States, it is widely cultivated in the Southeast and in California.

The cigarbox cedar (*Cedrela odorata*), also known as the West Indian cedar, belongs to the mahogany family, is a broad-leaved tree with pinnate, deciduous leaves and is related to the *Ailanthus* and sumac. The wood is very durable and fragrant and is valued in the West Indies for the manufacture of cabinets, furniture, and canoes. It seldom grows well in the northern United States. *See* FOREST AND FORESTRY; TREE. Arthur H. Graves; Kenneth P. Davis

Celastrales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Rosidae of the class Magnoliopsida (dicotyledons). The order consists of 11 families and more than 2000 species, with the families Celastraceae (about 800 species), Aquifoliaceae (about 400 species), Icacinaceae (about 400 species), and Hippocrateaceae (about 300 species) forming the core of the group. The order is marked by its simple leaves and regular flowers, varying from hypogynous (those with the perianth and stamens attached directly to the receptacle, beneath the ovary) to perigynous (those with the perianth and stamens united at the base into a hypanthium distinct from the ovary) and with a single set of stamens which alternate with the petals. Nearly all of the species are woody plants. Various species of holly (*Ilex,* family Aquifoliaceae; see **illus.**) and *Euonymus* (Celastraceae) are often cultivated. *See* HOLLY; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM; ROSIDAE.

American holly (*Ilex opaca*), a characteristic member of the family Aquifoliaceae in the order Celastrales. The spiny leaves are typical of several species of holly, but not of the whole family or order. (*Photograph by Eric J. Hosking, from National Audubon Society*)

Arthur Cronquist; T. M. Barkley

Celery

A biennial umbellifer (*Apium graveolens* var. *dulce*) of Mediterranean origin and belonging to the plant order Apiales. Celery is grown for its petioles or leafstalks, which are most commonly eaten as a salad but occasionally cooked as a vegetable (see **illus.**). Celeriac or knob celery (*A. graveolens* var. *rapaceum*) is

Celery (*Apium graveolens* var. *dulce*). (*Joseph Harris Co., Inc., Rochester, New York*)

grown for its enlarged rootlike stem and is commonly eaten as a cooked vegetable in Europe.

Propagation of celery is by seed planted in the field or sown in greenhouses or outdoor beds for the production of transplants. Field spacing varies; plants are generally grown 6–10 in. (15–25 cm) apart in 18–36-in. (0.4–0.9-m) rows. Celery requires a long growing season, cool weather, and unusually abundant soil moisture. It has a high requirement for boron; a deficiency results in "cracked stem." Exposure to prolonged cold weather (39–48°F or 4–9°C for 10–30 days) favors "bolting" or seed-stalk formation.

Varieties (cultivars) are classified primarily according to their color, green or yellow (self-blanching). The most popular green variety is Utah, of which a large number of strains are grown, such as Summer Pascal and Utah 52–70H. The acreage of yellow celery has been declining. Pink and red varieties are grown in England.

Harvesting begins generally when the plants are fully grown but before the petioles become pithy, usually 3–4 months after field planting. In periods of high prices, earlier harvesting is often practiced.

California, Florida, and Michigan are important producing states. *See* APIALES. H. John Carew

Celery seedlings are often severely damaged by a variety of soil-inhabiting fungi that cause "damping-off" diseases in seedbeds. Extensive chemical fumigation treatments are commonly used to reduce seedbed losses. Some diseases caused by fungi such as *Fusarium oxysporum* f. sp. *apii* and *Septoria apii* can lead to early plant losses in the seedbed, but they usually damage plants after transplanting in the field. The crown, roots, and petioles are rotted and killed. Basal stalk rot and pink rot are other important fungal diseases in this group.

Foliar diseases caused by both fungi and bacteria reduce quality and affect the appearance of the harvested product. In most celery production areas, pesticides are routinely applied to protect the foliage from infection and damage. Common foliar diseases are caused by the fungi *Cercospora apii* and *Septoria apii,* leading to the early and late blight diseases, respectively. Bacterial leaf blight commonly damages celery foliage in seedbeds and in the field. Virus diseases cause sporadic losses. Cucumber mosaic and western celery mosaic viruses are commonly encountered.

Celery exhibits several physiological disorders traceable to mineral deficiencies or imbalances. Among the most common and important disorders are blackheart, due to calcium deficiency, and cracked stem due to boron deficiency. *See* PLANT PATHOLOGY; PLANT VIRUSES AND VIROIDS.

J. O. Strandberg

Bibliography. G. N. Agrios, *Plant Pathology*, 4th ed., 1997; A. F. Sherf and A. A. MacNab, *Vegetable Diseases and Their Control*, 2d ed., 1986.

Celestial mechanics

The field of dynamics as applied to celestial bodies moving under their mutual gravitational influence in systems with few bodies. It usually describes and predicts motions in the solar system, both of natural bodies such as planets, satellites, asteroids, and comets, and of artificial bodies such as space probes. It can also be applied to small stellar systems.

Newton's laws. Isaac Newton's law of universal gravitation is the foundation of most of the field. It states that the force produced by one particle upon another is attractive along the line connecting the bodies, is proportional to the product of the masses of the bodies, and is inversely proportional to the square of the distance between the bodies. The constant of proportionality is G, the universal constant of gravitation. Newton's second law of motion then says that the acceleration experienced by a body is equal to the force on that body divided by its mass. *See* FORCE; KINETICS (CLASSICAL MECHANICS); NEWTON'S LAWS OF MOTION.

Two-body problem. The simplest and only exactly solvable problem in celestial mechanics is that of one particle moving about another. Since any body with spherical symmetry looks gravitationally like a point mass from the outside, the results from this problem may be used to describe approximately the relative motion of two finite bodies, such as a planet around the Sun or a satellite around a planet. The principal results from this problem had already been recognized empirically by Johannes Kepler and are embodied in his three laws of planetary motion. Usually the motion of the smaller body (the secondary) is described relative to the larger one (the primary). This relative motion is confined to a plane, and the path traced is a conic section such that the primary occupies one focus. If the bodies are gravitationally bound, the conic is an ellipse. The longest segment connecting opposite points on the ellipse is called the major axis, and half this length is called the semimajor axis a (see **illustration**). The departure of the ellipse from a circle is called the eccentricity e, which is usually quite small for planetary orbits. The tilt of the plane from some reference plane is called the inclination, and for the solar system that reference plane is the plane of the Earth's orbit, known as the ecliptic plane. Planetary inclinations are also usually

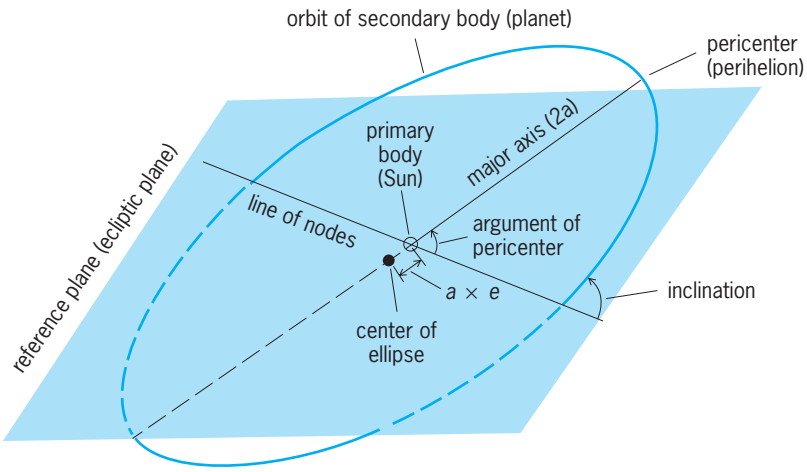

Relative motion of one body about another when the bodies are gravitationally bound. Parameters used to describe the motion are shown. (Terms used to describe the motion of a planet about the Sun are given in parentheses.) e = eccentricity.

quite small. The line of intersection of the plane of motion with the reference plane is called the line of nodes. The point on the orbit closest to the primary, which is at one end of the major axis, is called the pericenter (specifically for planetary orbits, the perihelion), and its angular distance from the node is called the argument of pericenter. The time at which the secondary passes through the pericenter is called the epoch of pericenter. A seventh parameter is the period of revolution, and the cube of the semimajor axis divided by the square of the period is proportional to the sum of the two masses. Since a planetary mass is small compared to that of the Sun, this ratio is essentially constant for the planets; this is Kepler's third law, also known as the harmonic law. *See* ELLIPSE.

A second result applies whenever the forces are directed along the line connecting the two bodies. Angular momentum is conserved, which causes the line connecting the two bodies to sweep out equal areas in equal times, a result stated in Kepler's second law. This results in the relative velocity in the orbit being inversely proportional to the square root of the separation. Ellipses are not the only type of relative motion permitted, and the type of conic depends on the total energy in the orbit. If there is just enough energy for the bodies to escape from each other, the relative orbit is a parabola. If there is more than enough energy for escape, such that some relative velocity would still remain, the orbit is a hyperbola. A hyperbola would also describe the relative motion of two independent bodies encountering each other, as in the case of two stars within the galaxy. *See* CONIC SECTION; ESCAPE VELOCITY; ORBITAL MOTION; PLANET.

Orbit determination. One of the major operational problems of celestial mechanics is that of determining the orbit of a body in the solar system from observations of its position, or distance plus line-of-sight velocity, at various times. The objective is to determine the numerical values of the parameters characterizing the orbit, known as orbital elements. It is usually first assumed that the motion is that of a two-body system and that the mass of the body is negligible. Thus, six elements must be determined. Each observation consists of two values at each time. Since six numbers are needed to solve for six other ones, a minimum of three observations is required. Usually there are more than three observations, which means a unique solution is not possible, but best values must be estimated in some statistical sense.

The problem is made difficult because observations are being made from the surface of the Earth, and hence parameters easily described by the two-body problem are not observed. The location of the point of observation with respect to the Sun must be accurately known, and this is often the limiting factor in orbit determination. In many cases, it is the improvement in knowledge of the orbit of the Earth that is the most useful result of a set of observations. *See* EARTH ROTATION AND ORBITAL MOTION.

Ephemeris generation. Once the orbit is known, the future locations of the object can be predicted. A table of predicted positions is called an ephemeris, and the generation of such a table is relatively straightforward. Usually, further observations are then obtained and checked for discrepancies, which in turn leads to improved values for the elements in a continually repeated cycle. *See* EPHEMERIS.

Space flight and ballistics. Another important problem is determining the proper orbit to get from one point at one time to another point at another time. This may involve getting from one body to another (space flight) or from one point to another on the same body (ballistics). For space flight, the approach is to consider several two-body problems and then patch them together. For instance, for interplanetary flight, there is an initial hyperbolic orbit with respect to one planet, an elliptic orbit with respect to the Sun for the most part, and a final hyperbolic orbit with respect to the other planet. For economy, an orbit with as little energy change as possible is desired. This dictates an elliptic heliocentric orbit that is just tangent to one planetary orbit at one extreme and just tangent to the other orbit at the other extreme. Such an orbit is known as a Hohmann transfer orbit, and it is unique for each pair of planets. The period of the orbit is known, and the tangent points are at opposite ends of the major axis. Therefore, the launch time is dictated by the relative locations of the planets. Since in practice there is a little room for adjustment, there usually is a period of time, known as the launch window, during which a launch is possible. *See* SPACE NAVIGATION AND GUIDANCE; SPACE PROBE.

The ballistic problem requires determining an Earth-centered elliptic orbit that will connect launch and target points. The actual path will be that part of the ellipse above the Earth's surface. Usually the two points are at least approximately at the same distance from the center of the Earth, but the Earth rotates. This requires a cyclic procedure of estimating travel time, allowing for the shift in target location during that time, improving the orbit, and so on. *See* BALLISTICS.

Binary stars. An astronomical application of the two-body problem is the relative motion of

components of a binary star. It was realized in the eighteenth century that this motion obeys the same laws as the solar system, which established the case for the universality of the law of gravitation. In this application, there is the advantage of being on the outside looking in, but there is no information on separation in the line of sight. The same techniques of orbit determination can be applied in usually simplified form, but since the masses are not known, period must be treated as a seventh independent parameter. This leads to the primary astronomical interest in these objects, because if the orbit is known, the masses can be determined, and this is the only presently known way to determine directly these important stellar data. *See* BINARY STAR.

Restricted three-body problem. Only slightly increased in complexity is this problem of the motion of a massless particle moving in the gravitational field of two bodies moving around each other in two-body motion. The simplest version of this problem is called planar circular, which means the particle moves in the plane of revolution of the two bodies which are themselves in circular orbits. This can be applied to objects like a spacecraft going from the Earth to the Moon or asteroids under the influence of the Sun and Jupiter. This problem has no general solution; the analytic and numerical study of the problem is concerned with stability, periodic orbits, and topology of solutions. There are five specific solutions—the fixed points or libration points. If the massless particle is placed at any of these points with zero velocity in the coordinate system rotating with the primaries, it will remain at that point in the rotating system. Three of these points are located along the line connecting the primaries, one between them and one outside of each; these are known as the linear points. Their exact distances from the primaries depend on the ratio of the masses of the primaries.

The other two points form equilateral triangles with the primaries, one ahead and one behind as they revolve. Unlike the linear points, these triangular points can be stable, in that a slight displacement of the massless particle away from the point will not produce unbounded motion but rather an oscillation (called a libration) about the point. The requirement for this is that the primaries' mass ratio must be below about 0.04. This is indeed true for the Sun-Jupiter system, and there are asteroids, known as the Trojans, librating about both triangular points in this system. It is also true for the Earth-Moon systems but the presence of the Sun makes matters more complicated. However, it is possible to keep the particle reasonably close to one of these lagrangian or L points for extended periods of time if initial conditions are chosen properly. *See* TROJAN ASTEROIDS.

Perturbations. If there are three or more bodies, all of which have mass and therefore all of which influence each other, the problem becomes almost hopeless. The degree of complexity is essentially independent of the number of bodies, so the problem is called the n-body problem. This is usually studied by purely numerical means, but in two extreme cases some analytical progress can be made. One is when the number of bodies, n, becomes so large that statistical approaches are possible; this leads into the dynamics of star clusters and galaxies and out of the field of celestial mechanics. The other is when relative geometries or masses are such that the situation becomes a series of two-body problems with small coupling influences, or perturbations. These perturbations can be treated in some approximate way, such as series expansions or iterative solutions. *See* GALAXY, EXTERNAL; MILKY WAY GALAXY; STAR CLUSTERS.

The major efforts in these perturbation studies have been the development of general theories of motion. These are elaborate mathematical approximate solutions which almost invariably involve series expansions in some small parameter of the problem. These can be solutions for slow variations of the elements that are fixed in the two-body problem, or they can be for slow divergence from the coordinates that would result from the two-body problem, or a mixture of the two. The numerical parameters are carried as literal expressions as far as possible, but before application, actual numerical values must be estimated. This usually involves a first estimate from two-body results, then similar ephemeris generation, comparison with observations, improved estimates of the parameters, and so on.

There are two classical areas of general perturbation theory. One is the development of lunar theory, the representation of the motion of the Moon about the Earth, under the influence of rather strong perturbations from the Sun. The small parameter is the period of the Moon (a month) compared to the period of the Earth (a year), and the motion of the Earth-Moon system about the Sun is considered known. This parameter is not really very small, producing slow rates of convergence and thus expressions with thousands of terms. *See* MOON.

The other major development has been planetary theory, the description of the motion of planets (either major or minor) about the Sun, under the influence of (other) major planets. The small parameter is the mass of the perturbing planet or planets compared with the Sun, but series expansions involving powers of ratios of mutual distances are also required. These ratios can become quite large (for Neptune and Pluto, it can become unity), again producing slow convergence. Furthermore, there are many planets, which add more terms to the resulting expressions. However, the biggest complication is resonances, a situation realized when the revolution periods of two planets are related by small integers. (The most famous of these is the great inequality of Jupiter-Saturn, produced by five of Jupiter's almost-12-year periods being almost exactly the same as two of Saturn's almost-30-year periods). Fortunately, these perturbations are long-period, large-amplitude, and sensitive to the perturbing mass, making them useful for planetary mass determinations.

The use of the general theory is now giving way to that of the special or numerical theory. The full equations of motion are solved on computers in

approximated numerical form to simulate the motions of the bodies and produce tables of coordinates as functions of time. Again, ephemerides must be generated and compared to observations to establish the best numerical values of the starting conditions. In addition, the extrapolatory value of these theories is more uncertain than that of analytic ones, because of the approximations inherent in the numerical solutions. *See* PERTURBATION (ASTRONOMY).

Oblateness and tides. Another set of perturbations from two-body theory arises from the finite sizes of the actual bodies. Because of rotation during the time of formation, no body is completely spherically symmetric, but rather possesses at least a small bulge around its equator. This bulge will cause slow changes in the orientation of the orbit of any object circling the oblate body. It will also cause the orbiter to pull asymmetrically on the oblate body and change its orientation in space. These slow changes in orientation are collectively referred to as precession. *See* PRECESSION OF EQUINOXES.

If a body of finite size is under the influence of another one, the parts closer to the perturber will feel a stronger pull and the parts farther away will feel a weaker pull. If the body is not infinitely rigid, this will result in a deformation known as tides. If it has a liquid or gaseous surface and is rotating, the tidal bulges will be carried at least part way around the body, but will reach a point at which the horizontal tidal forces compensate frictional drag and an equilibrium is reached. This drag will change the rotation rate of the body, at the same time releasing energy from the system, and the asymmetric bulge will perturb the orbit of the circling body to change its period and distance. The end result in one extreme is that the system will evolve until rotation and revolution rates are equal. Many planetary satellites are thus locked into synchronous rotation with their planets, and the Earth-Moon system is evolving to a time when both the day and the month will be equal to about 47 of our present days. However, solar tides will carry Earth rotation past the synchronous state, and the lunar distance will start to decrease again. Eventually the opposite extreme situation will arise, when the tidal forces on the Moon will exceed the cohesive ones and the Moon will break up. The limit on how closely a body can approach another one without breaking up is known as Roche's limit. Its exact value depends on the nature of the bodies involved, but it is usually around $2^1/_2$ times the radius of the larger body. *See* ROCHE LIMIT; SATURN; TIDE.

Unseen companions. The most complex situation is where there is a perturbation but no visible perturber. Unexplained variations can exist in observed orbits that can be due only to bodies not yet seen and therefore not yet modeled. Such was the case for Uranus after its discovery, and ultimately Neptune was discovered. Such is the case for both Uranus and Neptune, and while Pluto has been discovered, its mass is much too low to be responsible for the observed perturbations. A comparable situation can arise for a star, in which a wobble is seen in its motion across the sky, due to an unseen companion. Such detections have led to the discovery of many faint, substellar objects and may even result in the location of extrasolar planets. Since 1995, over 100 extrasolar planets have been discovered by using Doppler spectroscopy to observe variations in the radial motions of their parent stars. *See* ASTROMETRY; EXTRASOLAR PLANETS; NEPTUNE.

Post-newtonian theories. The newtonian law of universal gravitation has been remarkably successful in explaining most astronomical dynamical phenomena. However, there have been some discrepancies, the most glaring being a small unexplainable motion in the perihelion of Mercury. For a time, a planet interior to Mercury ("Vulcan") was suspected, but the problem was resolved by Einstein's theory of general relativity. Philosophically, gravitation is quite different in the two theories, but the mathematical description of motion in general relativity shows that Newton's simple relationship is "almost" correct. The errors depend on such quantities as the square of the object's velocity compared to the square of the speed of light, which make them virtually undetectable for astronomical bodies (with the exception of Mercury). These effects are, however, easily detectable in spacecraft trajectories, and thus now have to be routinely considered. Post-Einstein theories of gravitation have also been proposed, but there is no observational support for them and thus no observational need to adopt a more complex theory. *See* GRAVITATION; RELATIVITY.

Robert S. Harrington

Bibliography. R. R. Bate, D. D. Mueller, and J. E. White, *Fundamentals of Astrodynamics*, 1971; G. Beutler, *Methods of Celestial Mechanics*, 2 vols., 2005; J. M. Danby, *Fundamentals of Celestial Mechanics*, 2d ed., 1988; V. Szebehely and H. Mark, *Adventures in Celestial Mechanics*, 2d ed., 1998; M. Valtonen and H. Karttunen, *The Three-Body Problem*, 2006.

Celestial navigation

Navigation with the aid of celestial bodies, primarily for determination of position when landmarks are not available. In celestial navigation, position is not determined relative to the objects observed, as in navigation by piloting, but in relation to the points on the Earth having certain celestial bodies directly overhead.

Celestial bodies are also used for determination of horizontal direction on the Earth, and for regulating time, which is of primary importance in celestial navigation because of the changing positions of celestial bodies in the sky as the Earth rotates daily on its axis.

The navigator is concerned less with the actual motions of celestial bodies than with their apparent motions as viewed from the Earth. The heavens are pictured as a hollow celestial sphere of infinite radius, with the Earth as its center and the various celestial bodies on its inner surface. The navigator visualizes this sphere as rotating on its axis once in

about 23 h 56 min—one sidereal day. The stars are then back where they were when the period started.

Bodies closer to the Earth appear to change position at a different rate than do those at a distance. The Sun appears to make a complete revolution among the stars once a year, as the Earth makes one revolution in its orbit. The apparent motion is along a great circle called the ecliptic, which is inclined nearly 23.5° to the plane of the Equator of the Earth. All of the planets stay within 8° of the ecliptic, in a band called the zodiac. Within this band they appear to move among the stars. The Moon, too, stays within the zodiac as it revolves around the Earth—or more properly as the Earth and Moon revolve around their common center of mass—once each lunar month. *See* EARTH ROTATION AND ORBITAL MOTION.

Body selection and identification. The navigator uses a limited number of celestial bodies—the Sun, Moon, 4 planets, and perhaps 20–30 stars. Although 173 stars are listed in the *Nautical Almanac*, and 57 of these are listed in the *Air Almanac* and on the daily pages of the *Nautical Almanac*, the majority of them are normally not used unless the navigator's favorite ones are unavailable. Some of the navigational stars are not seen at the latitudes traveled by many navigators. *See* ALMANAC.

With relatively few bodies in use and little change in star positions from one evening to the next, identification is seldom a problem. A tentative selection is often made in advance by means of some form of star finder. When set for the latitude of the observer and the time and date of observation, this device provides a graphical indication of the approximate altitude and azimuth of each star shown. The relative positions of other bodies can be plotted by the user. Star finders, tables, and star charts may be used to identify bodies observed before identification.

Celestial equator coordinate system. Several systems of coordinates are available to identify points on the celestial sphere. The celestial equator system of coordinates is an extension of the equatorial system commonly used on the Earth. The intersection of the plane of the terrestrial Equator, extended, with the celestial sphere is a great circle called the celestial equator. The Earth's axis, extended, intersects the celestial sphere at the north and south celestial poles. Small circles parallel to the celestial equator, similar to parallels of latitude on the Earth, are called parallels of declination. Each of these connects points of equal declination, the celestial coordinate similar to latitude on the Earth.

Great circles through the celestial poles, similar to meridians on the Earth, are called celestial meridians if they are considered to remain fixed in relation to terrestrial meridians, and hour circles if considered to remain fixed on the rotating celestial sphere.

Several different quantities on the celestial sphere are analogous to those of longitude on the Earth. Greenwich hour angle is measured westward from the Greenwich celestial meridian, through 360°. Local hour angle is similarly measured from the celestial meridian of the observer. Meridian angle is measured eastward and westward from the local meridian, through 180°. Sidereal hour angle is measured westward from the hour circle of the veneral equinox, the point at which the Sun crosses the celestial equator on its northward travel in spring, through 360°. Right ascension is measured eastward from the hour circle of the vernal equinox, usually in hours, minutes, and seconds, from 0 through 24 h. The various relationships of the celestial equator system are shown in **Fig. 1**.

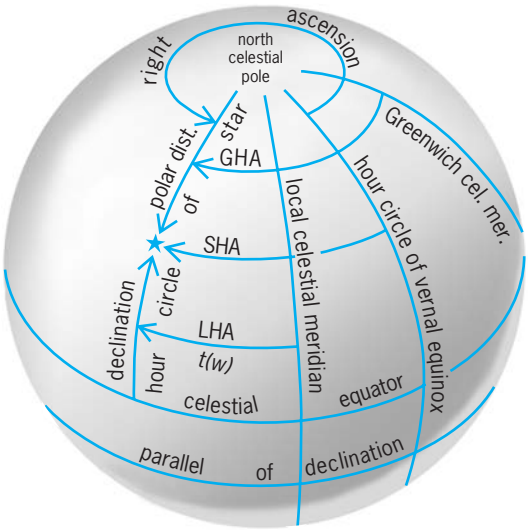

Fig. 1. Celestial equator system of coordinates. GHA is Greenwich hour angle, SHA is sidereal hour angle, LHA is local hour angle, and *t* is meridian angle (shown here as westerly).

With the exception of right ascension, all of these quantities are customarily stated to a precision of one minute of arc by the air navigator, and to one-tenth of a minute of arc by the marine navigator. The celestial equator system is used in the almanacs for indicating positions of celestial bodies at various times. *See* ASTRONOMICAL COORDINATE SYSTEMS.

Horizon system of coordinates. The navigator also uses the horizon system of coordinates, which is similar to the celestial equator system. The primary great circle is the horizon of the observer. The pole vertically overhead is the zenith, and the opposite pole is the nadir. Small circles parallel to the horizon are called parallels of altitude, each connecting all points having the same altitude. Angular distance downward from the zenith is called zenith distance. Great circles through the zenith and nadir are vertical circles. The prime vertical circle passes through the east and west points of the horizon. Azimuth is measured clockwise around the horizon, from 000° (it is generally expressed in three figures) at the north point, to 360°. Azimuth angle is measured eastward and westward to 180°, starting from north in the Northern Hemisphere, and from south in the Southern Hemisphere. Thus, it starts directly below the elevated pole, the celestial pole above the horizon. The relationships of the horizon system are shown in **Fig. 2**. The navigator uses the horizon system because it offers the most practical references for the origin of his or her measurements.

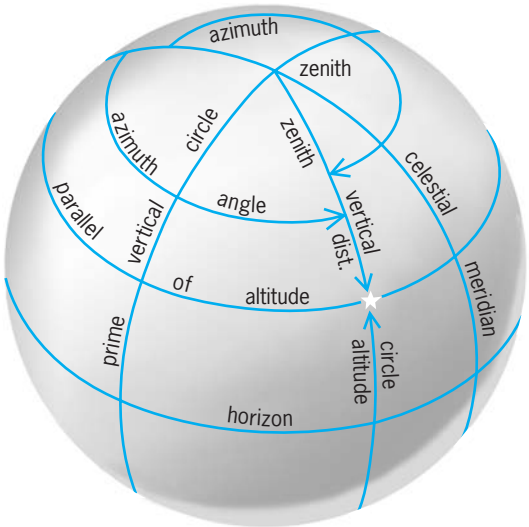

Fig. 2. Horizon system of coordinates.

Altitude and zenith distance are customarily stated by the air navigator to a precision of one minute of arc, and azimuth and azimuth angle to a precision of one degree; the marine navigator states them to one-tenth of a minute and one-tenth of a degree, respectively.

Two similar systems based upon the ecliptic and the galactic equator are used by astronomers, but not by navigators.

Position determination. Position determination in celestial navigation is primarily a matter of converting one set of coordinates to the other. This is done by solution of a spherical triangle called the navigational triangle.

The concept of the spherical navigational triangle is graphically shown in **Fig. 3**, a diagram on the plane of the celestial meridian. The celestial meridian passes through the zenith of the observer, and is therefore a vertical circle of the horizon system. Elements of both systems are shown in Fig. 3, indicating that an approximate solution can be made graphically.

The vertices of the navigational triangle are the elevated pole (P_n), the zenith (Z), and the celestial body (M). The angles at the vertices are, respectively, the meridian angle (t), the azimuth angle (Z), and the parallactic angle (X). The sides of the triangle are the codeclination of the zenith or the colatitude (colat) of the observer, the coaltitude or zenith distance (z) of the body, and the codeclination or polar distance (p) of the body.

A navigational triangle is solved, usually by computation, and compared with an observed altitude to obtain a line of position by a procedure known as sight reduction.

Observed altitude. To establish a celestial line of position, the navigator observes the altitude of a celestial body, noting the time of observation. Observation is made by a sextant, so named because early instruments had an arc of one-sixth of a circle. By means of the double reflecting principle, the altitude of the body is double the amount of arc used. Similar instruments were called octants, quintants, and quadrants, depending upon the length of the arc. Today, all such instruments, regardless of length of the arc, are generally called sextants.

The marine sextant uses the visible horizon as the horizontal reference. An air sextant has an artificial, built-in horizontal reference based upon a bubble or occasionally a pendulum or gyroscope. The sextant altitude, however measured, is subject to certain errors, for which corrections are applied. *See* SEXTANT.

When a marine sextant is used, observations can be made only when both the horizon and one or more celestial bodies are visible. This requirement generally eliminates the period between the end of evening twilight and the beginning of morning twilight. The navigational stars and planets are therefore usually observable only during twilight. Air navigation is not subject to this limitation, star and planet observations being available all night.

In selecting bodies for observation, the navigator considers difference in azimuth, magnitude (brightness of the body), altitude (avoiding both extremes), and sometimes other factors. If speed is of particular concern, the navigator selects a body nearly ahead or astern to provide a speed line. A body near the ship's beam provides a course line. One north or south provides a latitude line, while one east or west provides a longitude line. One perpendicular to a shoreline provides an indication of distance offshore.

Many navigators prefer to observe three bodies differing in azimuth by about 120°, or four bodies differing by 90°, and preferably at about the same altitude. In this way any constant error in the altitudes is eliminated.

Sight reduction. The process of deriving from an observation the information needed for establishing a line of position is called sight reduction. A great variety of methods have been devised. That now in general use is called the Marcq St.-Hilaire method,

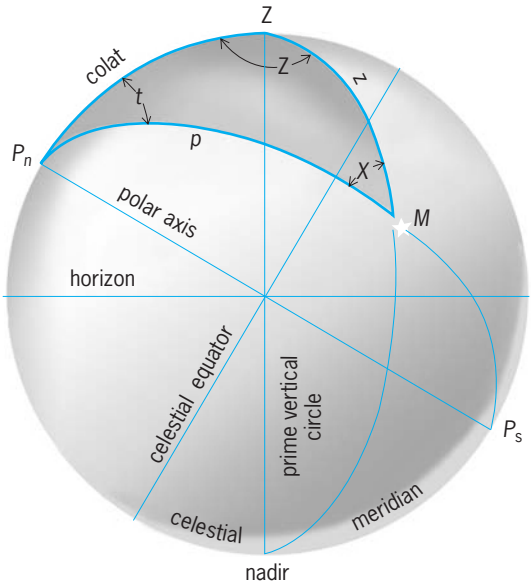

Fig. 3. Navigational triangle.

after the French naval officer who proposed it in 1875, and is described as follows.

Having obtained the body's corrected sextant altitude, called the observed altitude, the navigator uses the almanac to obtain the Greenwich hour angle and declination of the body. The navigator converts the former to local hour angle or meridian angle for an assumed position in the vicinity of actual position. With these two quantities and the latitude of the assumed position, the navigator solves the navigational triangle for altitude and azimuth.

Each altitude of a given celestial body at any one instant defines a circle having the geographical position of the body (the point on the Earth at which the body is momentarily vertically overhead) as the center, and the zenith distance as the radius. The difference between the computed altitude for the assumed position and the observed altitude at the actual position of the observer, called the altitude difference or altitude intercept, is the difference in radii of the circles of equal altitude through the two positions.

This difference is measured along an azimuth line through the assumed position, each minute of altitude difference being considered one nautical mile. A perpendicular through the point so located is considered a part of the circle of equal altitude through the position of the observer. This is the line of position sought. The intersection of two or more nonparallel lines of position adjusted to a common time defines a fix (or running fix if the elapsed time between observations is more than a few minutes) locating the position of the observer at the time of observation. Although two lines of position are sufficient for a fix, most navigators prefer to observe three or more to provide a check and to decrease somewhat the probable error of the fix.

Sometimes, however, the navigational triangle is solved in reverse, starting with altitude, assumed latitude, and declination, and solving for meridian angle. This is then compared with Greenwich hour angle at the time of observation to determine the longitude at which the line of position crosses the assumed latitude. An American, Capt. Charles H. Sumner, used this method in his discovery in 1837 of the celestial line of position. Others used one point, found by this so-called time sight method, and the azimuth of the celestial body to establish a line of position.

Special short-cut methods have been used for a body observed on or near the celestial meridian and for Polaris. Since simple methods of sight reduction have become widely used, the popularity of such special methods has decreased.

A large number of methods have been devised for solution of the navigational triangle. The most widely accepted have been mathematical, but a great many graphical and mechanical solutions have been proposed. Many of these methods are discussed in U.S. Defense Mapping Agency Hydrographic Topographic Center Publication 9, *American Practical Navigator*, originally by N. Bowditch. Most American navigators and many of other nationalities now use one of two methods, both published by the U.S. Defense Mapping Agency Hydrographic Topographic Center. Publication 229, *Sight Reduction Tables for Marine and Air Navigation*, is intended for use with the *Nautical Almanac* for marine navigation and with the *Air Almanac* for air navigation. Editions of both of these sets of tables are also published in certain other countries.

With the emergence of electronic computers and hand-held calculators, sight reduction has been performed increasingly with limited use or elimination of tables. The extent to which a computer replaces tables depends upon the capability of the available computer and the preference of the navigator. The principal use of a nonprogrammable hand-held calculator is to perform simple arithmetical calculations such as addition, subtraction, and interpolation, with respect to both the almanac and slight reduction tables. If a programmable calculator with stored programs is available, the complete sight reduction process can be performed without the aid of tables, and if suitable data are available, even the ephemerides of celestial bodies can be computed, eliminating the need for an almanac in the usual form. An *Almanac for Computers*, providing the essential data, is published by the U.S. Naval Observatory. In the most sophisticated systems, a central computer combines sight reduction data with outputs of electronic positioning systems, perhaps including those associated with navigational satellites, and dead-reckoning data to provide continuous readout of the current most probable position. *See* CALCULATORS.

Time relationships. Time is repeatedly mentioned as an important element of a celestial observation because the Earth rotates at the approximate rate of 1 minute of arc each 4 s of time. An error of 1 s in the timing of an observation might introduce an error in the line of position of as much as one-quarter of a mile. Time directly affects longitude determination, but not latitude. The long search for a method of ascertaining longitude at sea was finally solved two centuries ago by the invention of the marine chronometer, a timepiece with a nearly steady rate.

Several different kinds of time are used by the navigator. A timepiece, which keeps watch time, usually has a small watch error. When this is applied to watch time, the result is usually zone time. This is familiar to most people as standard time (such as Pacific Standard Time) or, when clocks are set an hour ahead, as daylight saving time (such as Central Daylight Saving Time). At sea the zones may be set by each vessel or aircraft, but they are generally 15° wide, centered on the meridians exactly divisible by 15°.

When zone time is increased or decreased by 1 h per 15° longitude (the amount of the zone description, for example, +7 for Mountain Standard Time), Greenwich mean time is obtained. This is the time used in the almanacs.

Local mean time differs from zone time by the difference in longitude between the meridian of the observer and the zone meridian, at the rate of 4 min of time for each degree of longitude, the meridian

to the eastward having the later time. Local mean time is used in tables indicating time of sunrise, sunset, moonrise, moonset, and beginning and ending of twilight.

All forms of mean time are based upon apparent motions of a fictitious mean sun which provides an essentially uniform time. Apparent time, based upon the apparent (visible) Sun, may differ from mean time by a maximum of nearly $16^1/_2$ min. Apparent time, plus or minus 12 h, indicates the actual position of the Sun with respect to the celestial meridian. At local apparent noon the Sun is on the celestial meridian, and the local apparent time is 1200. (Navigators customarily state time in four digits without punctuation, from 0000 at the start of a day to 2400 at the end of a day.)

Sidereal time, based upon motion of the stars, is used (indirectly in many cases) with a star finder or a star chart.

The custom of setting navigational timepieces to Greenwich mean time is growing, particularly among air navigators. This time is used almost invariably in polar regions.

Time signals are broadcast from a number of stations throughout the world to permit checking of standard timepieces. Marine chronometers are not reset by the user, an accurate record being kept of chronometer time, chronometer error, and chronometer rate. *See* TIME.

Day's work. A typical day's work of a marine navigator at sea, when using celestial navigation, is as follows:

1. Plot of dead reckoning.
2. Morning twilight observations for a fix.
3. Report of 0800 position to the commanding officer.
4. Morning sun line and compass check.
5. Winding of chronometers and determination of error and rate.
6. Noon sun line, advance of morning sun line for running fix, and report of 1200 position to the commanding officer.
7. Afternoon sun line and compass check.
8. Determination of time of sunset and preparation of a list of bodies available for observation during evening twilight.
9. Evening twilight observations for a fix.
10. Report of 2000 position to the commanding officer.
11. Determination of time of beginning of morning twilight, time of sunrise, and preparation of a list of bodies available for observation during morning twilight.
12. Time of moonrise and moonset.

Electronic applications, however, have gradually changed the pattern of celestial navigation in at least three important respects: (1) by providing noncelestial position information at sea, (2) by providing devices for automatic observation and sight reduction, and (3) by extending use of celestial navigation to all weather conditions at all times of day or night, by the use of electronic star trackers and of radio astronomy. *See* DEAD RECKONING; NAVIGATION; PILOTING; POLAR NAVIGATION; RADIO ASTRONOMY.

Alton B Moody

Bibliography. T. J. Cutler, *Dutton's Nautical Navigation*, 15th ed., 2003; W. B. Paulk, *Basic and Intermediate Celestial Navigation*, 1989; H. Schlereth, *Progressive Celestial Navigation*, 1996; U.S. Defense Mapping Agency Hydrographic Topographic Center, *American Practical Navigator* (2 vols., Publ. 9), *Sight Reduction Tables for Marine and Air Navigation* (6 vols., Publ. 229), published periodically; U.S. Naval Observatory, *Air Almanac*, *Almanac for Computers*, and *Nautical Almanac*.

Celestial reference system

A system for specifying coordinates for locating celestial objects. The International Celestial Reference System (ICRS) was adopted by the International Astronomical Union in 1994 as the new fundamental reference system. The ICRS is defined from a Barycentric Celestial Reference System, where "barycentric" means that the system is based on the center of mass of the solar system bodies, and a Geocentric Celestial Reference System, where "geocentric" means that the system is based on the center of the Earth. The systems are established based on specified gravitational potentials for the two origins. Their orientations are the same, and a standard set of equations, known as the Lorentz transformation, can be used to change from one to the other. Though the fiducial point (orientation of the x axis) of the new reference system is arbitrary, it has been chosen to agree with the dynamical equinox of the preexisting system of moving (dynamical) solar system objects adopted for the beginning of the year 2000, known as J2000.0, as closely as its accuracy allows. *See* LORENTZ TRANSFORMATIONS.

International Celestial Reference Frame. The International Celestial Reference System is realized by the International Celestial Reference Frame (ICRF), which is determined from the positions of approximately 400 extragalactic radio sources observed with very long baseline interferometry (VLBI). These sources are very distant, so they do not have any apparent proper motions (motions across the sky), but they can have structure that changes with time, so their individual positions have some very small uncertainty. The resulting reference frame is accurate to about 0.2 milliarcsecond or 10^{-9} radian. The International Celestial Reference Frame is a fixed, epochless frame, not changing as the epoch (such as 2000.0) changes. It replaces the historical use of fundamental star catalogs, such as the FK4 and FK5, which were based on optical observations of nearby bright stars. The standard frames, the transformations between them, and all the constants and motions involved are part of the International Celestial Reference System. The time scales, which are specified for the reference frames and origins involved, are also a part of the system. These standard values are subject to ongoing improvements based on an increasing number of highly accurate

observations. *See* RADIO ASTRONOMY; RADIO TELESCOPE; STAR.

Moving reference frames. While the International Celestial Reference Frame is fixed, most observations are made from the Earth, which has many kinematic motions. Thus, a moving reference frame must be established. In the past the determination of this moving reference frame has been based on the dynamical motions of the solar system objects. Thus, the moving equatorial and ecliptic planes determined the equinox, with the celestial equator defined by the Earth's Equator and the ecliptic corresponding to the Earth's orbit around the Sun. With the newly defined fixed reference frame, which is independent of the equinox, it is no longer necessary to use the equinox as the fiducial point for the moving reference frame. Therefore, a new nonrotating origin has been introduced, called the Celestial Intermediate Origin. Also, a new precession-nutation model, IAU 2000A, has been adopted, and from it a Celestial Intermediate Pole is determined. There is a choice of using a moving reference frame that is based on either the equinox or the Celestial Intermediate Origin. *See* ECLIPTIC; EQUINOX.

The new precession-nutation model includes geodesic precession and nutation (which are small corrections related to a relativistic effect), but it does not include free core nutation (that is, nutation coming from unrestrained motions in the Earth's core), which cannot be predicted accurately. With the determinations of subdaily periodic motions, it has been decided to restrict nutation terms to periods greater than 2 days. All shorter-period terms are included in polar motion. For accuracies limited to a milliarcsecond, standard theories can be used, but for better accuracies, values based on current observations available from the International Earth Rotation Service (IERS) must be used. Currently the observational determinations of nutation are more accurate than can be achieved by theoretical models based on a nonrigid Earth. *See* EARTH ROTATION AND ORBITAL MOTION; NUTATION (ASTRONOMY AND MECHANICS); PRECESSION OF EQUINOXES; RELATIVITY.

Optical catalogs. The new reference frame is directly observable at radio frequencies, but optically the sources are generally fainter than 18th magnitude. The *Hipparcos Catalogue*, determined from observations by that astrometric satellite of the early 1990s, provides an optical catalog on the International Celestial Reference Frame with accuracies of about 10–20 milliarcseconds at the present epoch. New catalogs—such as the *Tycho 2* (based on lower-accuracy observations from *Hipparcos*), the *USNO* (*U.S. Naval Observatory*) *CCD Astrographic Catalog* (*UCAC*), where charge-coupled devices (CCDs) are the electronic detectors used, and *USNO B2.0*—increase the density of objects in the optical reference frame and extend it to 21st magnitude on the International Celestial Reference Frame, including fainter stars than had been previously usable. New observations and planetary ephemerides are being made directly in the International Celestial Reference Frame. Observational data determine additional standard frames, such as the solar system dynamical frame and the frames for other wavelengths of the electromagnetic spectrum. *See* ASTRONOMICAL CATALOGS; EPHEMERIS; PARALLAX (ASTRONOMY).

Observations. Continuing observations of the radio sources by VLBI are required to maintain the positional accuracies and to detect possible changes in the structures, and hence the positions, of the sources of radiation. Global Positioning System (GPS) observations can be made continuously by many receivers worldwide, so short-period (less than 1 day) Earth orientation motions can be determined. The positional observations of the sources also determine precession, nutation, and orientation of the Earth. Observations of optical sources are required to maintain the accuracy of optical catalogs, which lose accuracy due to the uncertainties of the stars' proper motions. Lunar laser ranging observations currently can be made at centimeter accuracy to determine the lunar orbit, lunar librations, Earth orientation, precession, and nutation. *See* MOON; SATELLITE NAVIGATION SYSTEMS.

Terrestrial reference frame. An International Terrestrial Reference System (ITRS) has been adopted by the International Geophysical Union. It is realized by the International Terrestrial Reference Frame (ITRF), consisting of coordinates on the surface of the Earth. These positions are specified in terms of the Terrestrial Intermediate Origin and are subject to the variations due to geophysical effects, such as plate tectonics and tidal effects. In order to analyze and use observational data from the surface of the Earth, it is necessary to include all the effects that cause differences between the celestial and terrestrial reference frames. The angle between the Celestial Intermediate Origin and the Terrestrial Intermediate Origin is the Earth Rotation Angle, while the angle between the equinox and the Terrestrial Intermediate Origin is sidereal time. The transformation between the International Celestial Reference Frame and the International Terrestrial Reference Frame includes precession, nutation, polar motion, and the real Earth rotation given by the time system known as UT1. *See* ASTROMETRY; EARTH TIDES; GEODESY; PLATE TECTONICS; TIME.

P. K. Seidelmann

Bibliography. International Earth Rotation Service, *IERS Conventions*, 2003, 2004; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, 1992.

Celestial sphere

An imaginary sphere of infinite radius centered either on an observer or at the center of the Earth. We imagine that the stars and planets are attached to the inside surface of the celestial sphere. Standing outside on a clear moonless night far from city lights, it is easy to imagine that one is at the center of such a sphere and that the stars and planets are attached to its inside surface.

Celestial poles, celestial equator, and ecliptic. Extending the Earth's axis of rotation in both directions until it "touches" the celestial sphere determines two points, the north celestial pole and the south celestial pole. Similarly, projecting the Earth's equator onto the celestial sphere determines the celestial equator. As the Earth revolves around the Sun each year, we see the Sun seeming to travel across the celestial sphere. As it does, it traces out an imaginary great circle, the ecliptic. (It may help to remember that the stars are still in the sky in the daytime; the Sun is so bright that the stars cannot be seen, but they are still there.) The plane determined by the Earth's equator is tilted with respect to the plane determined by the ecliptic, so the Sun is north of the equator for 6 months of each year and south of the equator for the other 6 months. The ecliptic and the equator cross at two points, the vernal equinox and the autumnal equinox. The vernal equinox is the point where the Sun crosses the equator on its way north each year, marking the first day of spring in the Northern Hemisphere (see **illustration**). *See* ECLIPTIC; EQUATOR; EQUINOX; NORTH POLE; SOUTH POLE.

Coordinate systems. Because the distances to the stars and planets are so great, two different observers see the same star in the same direction and thus the star can be thought of as being at a specific "position" on the sphere. Using the two poles, the equator, and the ecliptic, it is straightforward to establish a coordinate system that makes it possible to determine the position of any object in the sky. The system is similar to the system of latitudes and longitudes used for the locations of objects on the surface of the Earth. Right ascension (abbreviated RA or with the Greek letter α) is analogous to longitude. Declination (abbreviated Dec or with the Greek letter δ) is analogous to latitude. *See* LATITUDE AND LONGITUDE.

As with longitude on the Earth, it is necessary to choose a "zero" point for right ascension. On Earth we use a meridian—a great semicircle drawn from one pole to the other, perpendicular to the equator—passing through Greenwich, England. On the celestial sphere we use the meridian that passes through the vernal equinox. Every point on this prime meridian has a right ascension of zero. Right ascension is measured in hours and minutes, ranging from 0 hours, 0 minutes to 23 hours, 59.999. . . minutes. Hours are used rather than degrees because the entire sphere seems to rotate once per day. Each hour of right ascension corresponds to 15°. Right ascension increases toward the east; that is, if one star is slightly east of another, the first star will have a larger right ascension than the second—unless they are on opposite sides of the prime meridian. *See* MERIDIAN.

Declinations are measured north and south of the equator with angles between 0° and 90°, measured in degrees and minutes. Northern declinations are considered positive; southern declinations negative.

The brightest star, Sirius, is at RA 6 hours 45.146′ and Dec −16° 43.049′ (see illustration). The star Betelgeuse is at RA 5 hours 55.172′ and Dec +7° 24.424′. *See* ASTRONOMICAL COORDINATE SYSTEMS.

Motions of celestial objects. The positions of the objects on the celestial sphere are not fixed. The Sun, Moon, and planets are in constant motion relative to the Earth and thus their positions on the celestial sphere change constantly. The same is true of the stars, but to a much lesser extent. The stars are all in motion relative to one another; as a result, their positions on the celestial sphere change all the time. However, the stars are extremely far away so these changes are very small and need to be taken into consideration only when extremely accurate positioning of a telescope is required or to discuss how a star's position has changed over a long period of time. *See* MOON; PLANET; STAR.

Precession of the equinoxes. As the Earth rotates, it also wobbles. It completes one slow wobble every 26,000 years. As a result of this wobble, the vernal equinox slides along the ecliptic, making a full trip around every 26,000 years. This phenomenon is known as the precession of the equinoxes. In addition to the changes mentioned earlier, the right ascension and declination of every object in the sky changes constantly due to precession. The numbers given above for Sirius and Betelgeuse are for the first day of the year 2000. The difference between the right ascension and declination for a star on any nearby date and the values for 2000 is tiny, but that difference does matter when a telescope or other astronomical instrument is being used at high power. *See* PRECESSION OF EQUINOXES.

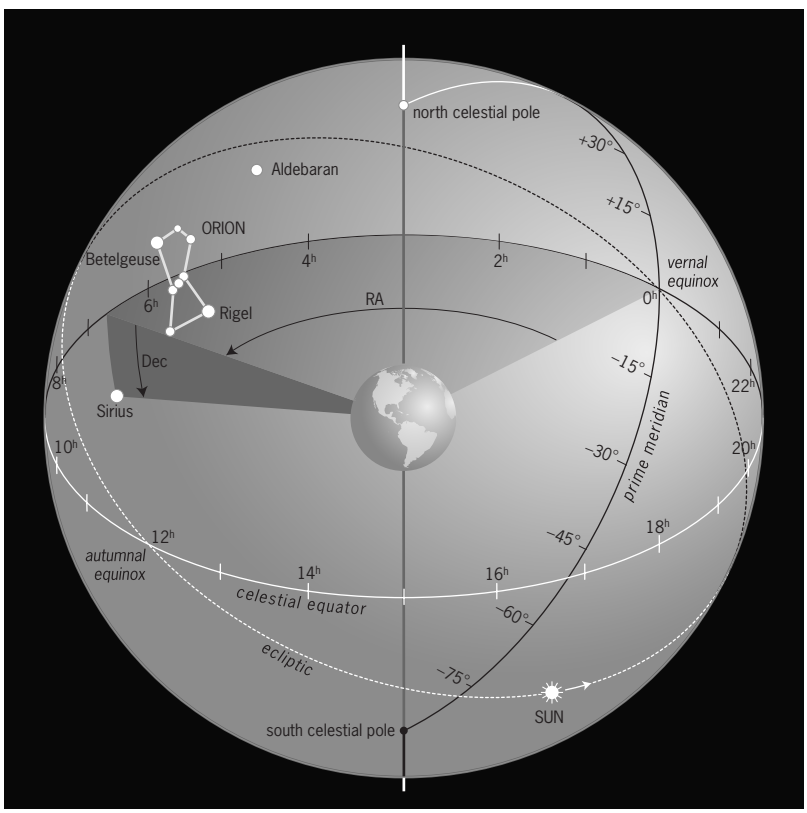

Celestial sphere. The arrows indicate the right ascension (RA) and declination (Dec) of Sirius. (*Wil Tirion*)

Distances. The stars are immensely distant compared to the distances to the Sun, Moon, and planets. The galaxies are immensely distant compared to the distances to the stars we see with the unaided eye. The notion of an imaginary celestial sphere is of little use when we want to talk about the actual structure of our universe. Samuel E. Rhoads

Bibliography. A. Fraknow, D. Morrison, and S. Wolff, *Voyages Through the Universe*, Brooks/Cole, 3d ed., 2006; D. H. Levy, *Skywatching, A Nature Company Guide*, The Nature Company, 1994; G. Lovi and W. Tirion, *Men, Monsters and the Modern Universe*, Willmann-Bell, 1989; J. M. Pasachoff, *Astronomy: From the Earth to the Universe*, Saunders College Publishing, 6th ed., 2002; J. M. Pasachoff, *Field Guide to the Stars and Planets*, 4th ed., Houghton-Mifflin, 2000; J. M. Pasachoff and A. Filippenko, *The Cosmos: Astronomy in the New Millennium*, Brooks/Cole, 3d ed., 2007; H. A. Rey, *The Stars: A New Way To See Them*, Houghton Mifflin, rebound by Sagebrush, 2002; I. Ridpath (ed.), *Norton's Star Atlas and Reference Handbook*, 20th ed., Pi Press, 2004.

Celestite

A mineral with the chemical composition $SrSO_4$. Celestite occurs commonly in colorless to sky-blue, orthorhombic, tabular crystals. Fracture is uneven and luster is vitreous. Hardness is 3–3.5 on Mohs scale and specific gravity is 3.97. It fuses readily to a white pearl. It is only slowly soluble in hot concentrated acids or alkali carbonate solutions. The strontium present in celestite imparts a characteristic crimson color to the flame.

Celestite occurs in association with gypsum, anhydrite, salt beds, limestone, and dolomite. Large crystals are found in vugs or cavities of limestone (see **illus.**). It is deposited directly from seawater, by groundwater, or from hydrothermal solutions. Celestite is the major source of strontium. The principal use of strontium is in tracer bullets and in various red flares used by the armed forces. Minor applications of strontium compounds are in ceramics, depilatories, and medicine. Although celestite deposits occur in Arizona and California, domestic production of celestite has been small and sporadic. Much of the strontium demand is satisfied by imported ores from England and Mexico. *See* STRONTIUM. Edward C. T. Chao

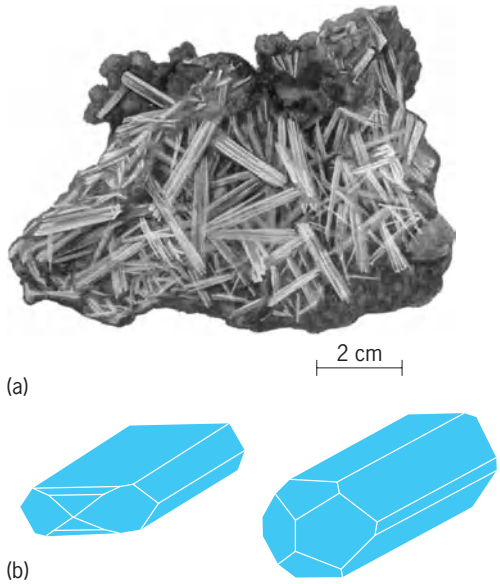

Celestite. (*a*) Bladed crystals in limestone from Clay Center, Ohio (*specimen from Department of Geology, Bryn Mawr College*). (*b*) Crystal habits (*after C. S. Hurlbut, Jr., Dana's Manual of Mineralogy, 17th ed., Wiley, 1959*).

Cell (biology)

Cells can be separated into prokaryotic and eukaryotic categories. Eukaryotic cells contain a nucleus. They comprise protists (single-celled organisms), fungi, plants, and animals, and are generally 5–100 micrometers in linear dimension. Prokaryotic cells contain no nucleus, are relatively small (1–10 μm in diameter), and have a simple internal structure (**Fig. 1**). They include two classes of bacteria: eubacteria (including photosynthetic organisms, or cyanobacteria), which are common bacteria inhabiting soil, water, and larger organisms; and archaebacteria, which grow under unusual conditions. The archaebacteria include methanogens, which live in anaerobic conditions and reduce carbon dioxide (CO_2) to methane (swamp gas); halophiles, which thrive under extremely high salt conditions; and sulfur bacteria, which grow in hot (80°C; 176°F) sulfur springs where the pH is extremely acidic. *See* EUKARYOTAE.

Eubacteria. All eubacteria have an inner (plasma) membrane which serves as a semipermeable barrier allowing small nonpolar and polar molecules such as oxygen, carbon dioxide, and glycerol to diffuse across (down their concentration gradients), but does not allow the diffusion of larger polar molecules (sugars, amino acids, and so on) or inorganic ions such as Na^+, K^+, Cl^-, Ca^{2+} (sodium, potassium, chlorine, calcium) [Fig. 1]. The plasma membrane, which is a lipid bilayer, utilizes transmembrane transporter and channel proteins to facilitate the movement of these molecules. Eubacteria can be further separated into two classes based on their ability to retain the dye crystal violet. Gram-positive cells retain the dye; their cell surface includes the inner plasma membrane and a cell wall composed of multiple layers of peptidoglycan (Fig. 1*b*). Gram-negative bacteria (Fig. 1*a*) are surrounded by two membranes: the inner (plasma) membrane and an outer membrane that allows the passage of molecules of less than 1000 molecular weight through porin protein channels. Between the inner and outer membranes is the peptidoglycan-rich cell wall and the periplasmic space. *See* CELL MEMBRANES.

Eubacteria contain a single circular double-stranded molecule of deoxyribonucleic acid (DNA), or a single chromosome. As prokaryotic cells lack

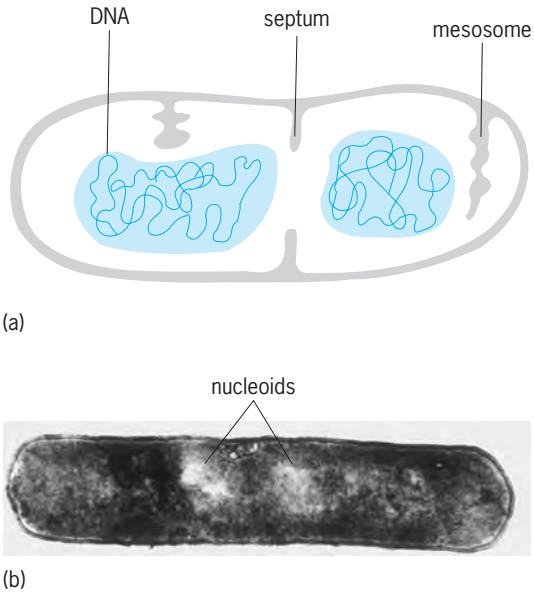

Fig. 1. Bacterial cell structure. (*a*) Artist's rendition of a gram-negative eubacterial cell (*modified from H. Lodish et al., Molecular Cell Biology, 3d ed., Scientific American, New York, 1995*). (*b*) Electron micrograph of a gram-positive eubacterium, *Bacillus licheniformis* (*courtesy of S. R. Goodman*).

a nucleus, this genomic DNA resides in a central region of the cell (Fig. 1*b*) called the nucleoid. *Escherichia coli*, a common gram-negative eubacteria, contains 0.4 picogram of DNA in its genome, or 4×10^6 base pairs. Bacterial replication requires accurate duplication of this chromosomal DNA, with both copies attaching by cross-wall formation to the plasma membrane, allowing the formation of two identical daughter cells by binary fission. The bacterial genome contains all the necessary information to maintain the structure and function of the cell. The processing of this information occurs in two steps. First the DNA is transcribed to form ribonucleic acid (RNA) by the enzyme RNA polymerase. Before this messenger RNA (mRNA) is completely transcribed, there is attachment of a 70S ribosome. The ribosome translates the mRNA into protein by reading three nucleotides (codon) at a time that encode a specific amino acid. The amino acids are added via specific transfer RNA (tRNA) molecules containing anticodon regions which base-pair with the mRNA codon. *Escherichia coli*, for example, contains about 30,000 cytoplasmic ribosomes and produces as much as 2000 distinct mRNAs and, therefore, protein subunits.

Many bacteria are able to move from place to place, or are motile. Their motility is based on a helical flagellum composed of interwoven protein called flagellin. The flagellum is attached to the cell surface through a basal body, and propels the bacteria through an aqueous environment by rotating like the propeller on a motor boat. The motor is reversible, allowing the bacteria to move toward chemoattractants and away from chemorepellants.

Eukaryotic cells. The smallest objects that can be visualized by the modern light microscope are approximately 0.5 μm wide. In a light microscopic view of a eukaryotic cell, a plasma membrane can be seen which defines the outer boundaries of the cell, surrounding the cell's protoplasm or contents. The protoplasm includes the nucleus, where the eukaryotic cell's DNA is compartmentalized away from the remaining contents of the cell (the cytoplasm). Therefore, the protoplasm includes the nucleus and the cytoplasm. In addition to the nucleus, the cytoplasm contains other organelles (Fig. 2*b*). The eukaryotic cell's organelles include the nucleus, mitochondria, endoplasmic reticulum, Golgi apparatus, lysosomes, peroxisomes, cytoskeleton, and plasma membrane (**Fig. 2***a*). The organelles occupy approximately half the total volume of the cytoplasm. The remaining compartment of cytoplasm (minus organelles) is referred to as the cytosol or cytoplasmic ground substance. Eukaryotic cells also differ from prokaryotic cells in having a cytoskeleton that gives the cell its shape, its capacity to move, and its ability to transport organelles and vesicles from one part of the cell cytoplasm to another. The cytoskeleton is composed of microfilaments (7–8 nanometers in diameter), intermediate filaments (10 nm), microtubules (25 nm), and a spectrin-based membrane skeleton. Eukaryotic cells are generally larger than prokaryotic cells and therefore require a cytoskeleton and membrane skeleton to maintain their shape, which is related to their functions.

The development of multicellular organisms required that cells closely related by ancestry become differentiated from one another—some developing certain features, others developing different traits. While different cells within a human tissue (such as nervous tissue) are radically different from each other, they have descended from the same fertilized egg. In most cases (one exception is human erythrocytes), cells retain all of the genetic material contained in the precursor fertilized egg. Differentiation or specialization of cells depends on gene expression. Eukaryotic cells contain a large amount of DNA (about a thousandfold more than bacterial cells), only approximately 1% of which encodes protein. The remaining DNA is structural (involved in DNA packaging) or regulatory (helping to switch on and off genes). Therefore, whether a cell becomes a muscle fiber or a neuron depends on the expression of muscle-specific genes or neuron-specific genes (although both are present).

Plasma membrane. The plasma membrane serves as a selective permeability barrier between a cell's environment and cytoplasm. The fundamental structure of plasma membranes (as well as organelle membranes) is the lipid bilayer, formed due to the tendency of amphipathic phospholipids to bury their hydrophobic fatty acid tails away from water. Human and animal cell plasma membranes contain a varied composition of phospholipids, cholesterol, and glycolipids. Cholesterol intercalates itself perpendicular to the membrane surface between the phospholipids. Glycolipids are primarily found in the noncytoplasmic leaflet of the bilayer. While lipids determine the bilayer structure, membrane proteins are

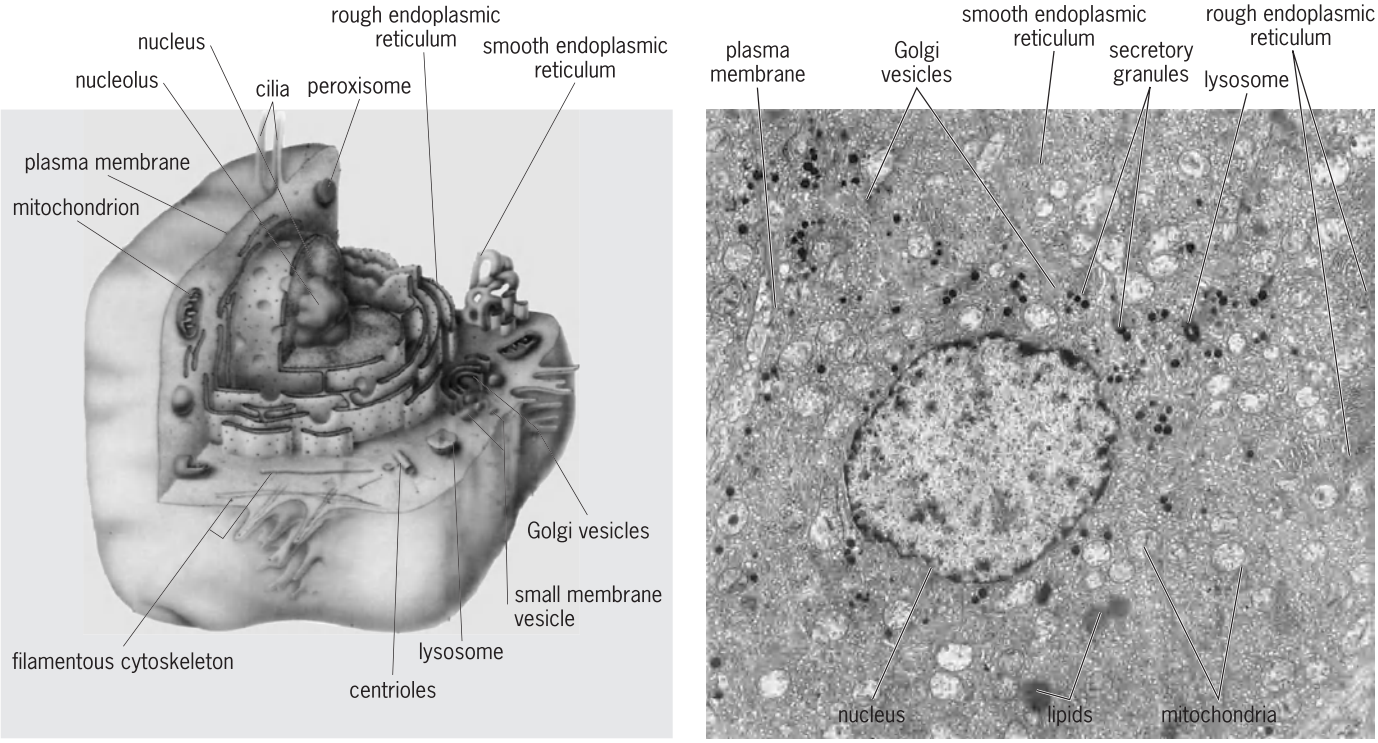

Fig. 2. Animal cell structure. (*a*) Artist's rendition of a eukaryotic animal cell (*modified from H. Lodish et al., Molecular Cell Biology, 3d ed., Scientific American, New York, 1995*). (*b*) Electron micrograph of a bovine large luteal cell (*courtesy of Dr. Phillip Fields*).

primarily responsible for membrane function. Membrane proteins can be divided into two categories: integral proteins that are embedded in the hydrophobic core of the bilayer, and peripheral proteins that are attached noncovalently to the hydrophilic membrane surfaces. Integral membrane proteins can be subdivided into transmembrane proteins that have a single alpha helix transversing the hydrophobic core of the bilayer (single-pass transmembrane proteins); multipass transmembrane proteins (transporters and channels are of this type); or proteins that are covalently linked to a fatty acid chain or phospholipid which is embedded. Both the lipids and proteins of the plasma membrane are asymmetrically distributed across the bilayer. This means that the lipid and protein in the extracellular leaflet (E face) are distinct from the protoplasmic leaflet (P face). Plasma membranes are fluid. The protein and lipids of the plasma membrane are capable of lateral movement and rotation around a central axis, but they are not capable of rapid flip-flop from one leaflet to the other. Therefore, the asymmetry of the plasma membrane protein and lipid, important to the differing functions on either side of the bilayer, is maintained over the lifespan of the cell. The lateral mobility of some transmembrane proteins is limited by their interaction with peripheral proteins on the extracellular or cytoplasmic surface of the plasma membrane. This creates a membrane topography where selected membrane domains can have differing protein and lipid composition and function.

Cytoskeleton. The cytoskeleton is involved in establishing cell shape, polarity, and motility, and in directing the movement of organelles within the cell. The cytoskeleton includes microfilaments, microtubules, intermediate filaments, and the two-dimensional membrane skeleton that lines the cytoplasmic surface of cell membranes. Microfilaments contain filamentous actin (F-actin) which is formed by the adenosine triphosphate (ATP)–dependent polymerization of globular actin (G-actin) monomers. Actin microfilaments have a polarity. They have a plus end where polymerization of additional G-actin monomers occurs at a faster rate than at the minus end of the filament. The plus end of the actin filaments associates with membranes and the Z-disk of muscle sarcomeres. Myosin is a fibrous protein which forms 15-nm-thick filaments in muscle and nonmuscle cells. In the presence of ATP, myosin head groups interact with actin microfilaments, causing contraction in muscle and nonmuscle cells (for example, contractile ring, stress fibers, adhesion belts). In addition to a contractile role, actin microfilaments have a structural role in epithelial cells, where they form the core of the fingerlike projections called microvilli. Each microvillus contains 20–30 bundled actin filaments with their plus ends contacting the tip of the microvillus, and their minus ends down in the terminal web region, attached via nonerythroid spectrin and myosin II to other actin bundles and intermediate filaments. Gel-sol transformation within the cytoplasm of human and animal

cells is controlled by the length of actin filaments and the extent of microfilament crosslinking. Those actin-binding proteins which shorten actin filaments cause solation, while those which crosslink F-actin or lead to lengthening of the actin filament cause gelation. The membrane skeleton on the cytoplasmic surface of all eukaryotic cell membranes contains a fibrous protein called spectrin which crosslinks short actin protofilaments into a two-dimensional meshwork. This spectrin membrane skeleton stabilizes the bilayer, gives shape to the membrane, and controls the lateral mobility of some transmembrane proteins.

Microtubules are hollow tubes (25 nm in diameter) which are composed of alpha, beta tubulin dimers. The tubulin dimers interact head to tail to form protofilaments, and 13 protofilaments form the wall of the microtubule. Microtubules have a polarity, with polymerization occuring faster at the plus end. In vivo and in vitro microtubules undergo dynamic instability, and grow and shrink rapidly. Whether a microtubule is growing or shrinking is determined by the rate of guanosine triphosphate (GTP) hydrolysis versus the rate of tubulin dimer addition. If GTP hydrolysis is occurring faster than tubulin addition, then the microtubule will shrink; if tubulin addition is faster, then the microtubule will grow. Within a nondividing cell, microtubules grow from a single structure near the nucleus, called a centrosome or microtubule organizing center. The microtubule organizing center contains a pair of centrioles surrounded by pericentriolar material, and a microtubule is capped at its minus end by the pericentriolar material. During mitosis, the centrioles within the microtubule organizing center duplicate and form the spindle poles. The mitotic spindle is polar, with the spindle microtubule minus end facing the spindle pole. Microtubules within interphase and mitotic cells are very dynamic. However, other cellular microtubules are stable, such as the microtubules found in neuronal axons and dendrites and those found in axonemes of cilia and flagella. Several factors contribute to the stability of microtubules: posttranslational acetylation of lysine residues or removal of the C-terminal tyrosine of tubulin, microtubule associated proteins (MAPs), or minus-end capping proteins. Microtubules serve as tracks on which organelles and vesicles are transported through the cytoplasm. Two proteins called kinesin and cytoplasmic dynein have the ability to move organelles toward the plus and minus ends of microtubules, respectively. These proteins are translocating adenosine triphosphatases (ATPases), and appear to be responsible for most of the cytoplasmic movement of organelles. Cilia and flagella are extensions of the cell surface which are responsible for movement, and both have a very similar internal structure. Cilia move particulate matter across the apical surface of epithelial sheets; a flagellum forms the tail of the sperm and is responsible for its ability to swim. Both types of movement are based on an internal microtubule-based structure called an axoneme, and require ATP hydrolysis as a source of energy. The axoneme has a characteristic (9+2) pattern, with 9 outer microtubule doublets (A and B subfibers) and 2 inner singlet microtubules. The movement of cilia and flagella is based on an ATPase named dynein which is attached along the length of each A subfiber in the outer doublets. In the presence of ATP, the dynein heads reach out and attach to the B subfiber on an adjacent doublet and pull it toward the tip of the cilia or flagella. Bending rather than sliding results, because the outer doublets are attached to each other by a flexible protein called nexin. The axonemal microtubules are nucleated by basal bodies which have the same structure as centrioles. Basal bodies and centrioles are small cylinders with nine outer triplet microtubules at a 45° angle from the radius of the cylinder.

Intermediate filaments are 10 nm in diameter and play structural roles within cells. Intermediate filaments within muscle cells link the Z-disks of adjacent myofibrils together. Intermediate filaments within neurons (neurofilaments) serve as a structural support for axons and dendrites, and the intermediate filaments of epithelial cells interconnect spot desmosomes, thereby stabilizing epithelial sheets. Intermediate filaments in human and animal cells are made up of cell-type-specific proteins. The intermediate filament subunits are fibrous, and almost all subunits are incorporated into the ropelike intermediate filament in a process that does not require ATP or GTP hydrolysis. The resulting intermediate filaments have no polarity. This heterogeneous group of proteins can all form intermediate filaments of similar dimension because of a common 310-amino-acid alpha-helical region that forms the core of the intermediate filament.

Nucleus. One of the most prominent organelles within a eukaryotic cell is the nucleus. The nuclear compartment is separated from the rest of the cell by a specialized membrane complex built from two distinct lipid bilayers, referred to as the nuclear envelope. However, the interior of the nucleus maintains contact with the cells cytoplasm via nuclear pores. These structures form holes or channels in the nuclear envelope that are responsible for the selective uptake of nuclear components and release of components destined for the cell cytoplasm. Attached to the nuclear pore complex and the inner surface of the nuclear membrane is a fibrous protein network built from the intermediate-filament proteins known as nuclear lamins A, B, and C. This nuclear lamina participates in the breakdown and reformation of the nuclear boundary during mitosis and, in combination with a fibrous complex of proteins referred to as the nuclear matrix, is responsible for maintaining the structure of the nucleus. The primary function of the nucleus is to house the genetic apparatus of the cell; this genetic machinery is composed of DNA (arranged in linear units called chromosomes), RNA, and proteins. Nuclear proteins aid in the performance of nuclear functions and include polypeptides that have a direct role in the regulation of gene function and those that give structure to the genetic material. The complex of DNA and

structural proteins is known as chromatin, which can be either highly compacted (heterochromatin) or dispersed (euchromatin) within the nucleus. A specialized region within the nucleus that is visible in interphase cells is the nucleolus, which functions in the production and assembly of ribosomal subunits. *See* CELL NUCLEUS.

Gene expression. The biochemical and phenotypical properties of a cell are defined by the proteins expressed within the cell. The expression of a protein involves a complex pathway that transfers information from the DNA within the nucleus to the protein-synthetic machinery within the cytoplasm. The first step of this pathway is the synthesis of an RNA molecule from the DNA template by the process known as transcription. All cells in an organism contain a complete complement of genetic material; however, only a selected portion is synthesized into RNA. Thus, the transcription process is highly regulated within each cell. In eukaryotes, RNAs are synthesized from units referred to as genes which contain the information needed to specify a protein sequence stored in three-nucleotide "words" called codons. The informational content of a gene is not necessarily contiguous in that some noncoding DNA may be found interrupting the sequences which form the functional RNA found in the cell's cytoplasm. These noninformational segments are transcribed and removed from the primary transcript in the nucleus, with the informational or exon segments of the RNA joined together by a process called splicing. Once the functional RNA is formed, it is transported to the cytoplasm via the nuclear pores. Upon exiting the nucleus, a mature RNA can be translated into a polypeptide sequence. The amount of protein that is made from an RNA can be regulated by the level of RNA available to the translational machinery or by the ability of the machinery to select specific RNAs for translation. Some proteins are not able to perform their function upon translation and are subjected to modification following translation. In summary, the transfer of information from DNA to a protein capable of performing a function within the cell involves several steps, each of which is subject to regulation. It is the composite of all the individual steps that determines the final level of expression of any given gene product. *See* DEOXYRIBONUCLEIC ACID (DNA); GENE.

Protein synthesis from mRNA occurs within the cytoplasm of eukaryotic cells, catalyzed by RNA and protein containing 80S particles called ribosomes. Proteins to be secreted or inserted into the plasma membrane, endoplasmic reticulum membrane, Golgi membrane, or lysosomal membrane are synthesized on ribosomes attached to the endoplasmic reticulum membrane. Proteins which will reside within the cytosol, some mitochondrial proteins, and peripheral membrane proteins are synthesized on free (not endoplasmic reticulum–associated) ribosomes. There is no difference between endoplasmic reticulum–associated and free ribosomes. Their position is determined by the sequence of the protein being synthesized.

Endoplasmic reticulum. The endoplasmic reticulum is composed of membrane-enclosed flattened sacs or cisternae. The enclosed compartment is called the lumen. The endoplasmic reticulum is morphologically separated into rough (RER) and smooth (SER). RER is studded with ribosomes and SER is not. RER is the site of protein synthesis, while lipids are synthesized in both RER and SER. In the case of proteins which will be synthesized in the RER, an N-terminal signal sequence causes the protein and ribosome to be bound to a signal recognition particle. The signal recognition particle and bound signal sequence and ribosome are attached to the endoplasmic reticulum through a docking protein called the signal recognition particle receptor. Then the signal peptide threads its way through a channel protein into the endoplasmic reticulum lumen, and protein synthesis continues until the nascent protein is complete and located within the lumen. If the protein is soluble, the signal sequence is cleaved, releasing the protein into the lumen. If the protein is a transmembrane integral protein, its topography across the bilayer is determined by carefully placed start and stop transfer signals (hydrophobic stretches of amino acids). Glycosylation of secretory and membrane proteins begins within the lumen of the endoplasmic reticulum and then is completed within the Golgi apparatus. Membrane lipids are also synthesized in the endoplasmic reticulum, and then they are moved to target organelles by transport vesicles or phospholipid exchange proteins. The synthesis of phospholipids occurs on the cytoplasmic leaflet of the endoplasmic reticulum (usually SER); some are moved to the luminal leaflet by specific phospholipid translocating enzymes called flippases. In this manner, membrane asymmetry is established by the specific flipping of phospholipids across the endoplasmic reticulum bilayer. Since the flippases are present at low concentration in the plasma membrane, Golgi membrane, or lysosomal membrane, the asymmetry created in the endoplasmic reticulum remains unaltered. The phospholipids synthesized in the endoplasmic reticulum are transferred to the plasma membrane, Golgi membrane, and lysosomal membrane by the budding-off of endoplasmic reticulum vesicles which travel to and fuse with the membrane of the target organelle. Mitochondria and peroxisomes receive most of their phospholipid via phospholipid exchange proteins which carry phospholipids from the endoplasmic reticulum cytoplasmic leaflet to these organelles. *See* ENDOPLASMIC RETICULUM.

Golgi apparatus. The final posttranslational modifications of proteins and glycolipids occur within a series of flattened membranous sacs called the Golgi apparatus. Vesicles which bud from the endoplasmic reticulum fuse with a specialized region of the cis Golgi compartment called the cis Golgi network. A series of posttranslational modifications occur first in the cis compartment, then in order in the medial, trans, and trans Golgi network compartments. In each case the membrane transfer occurs by the blebbing of coated vesicles from one compartment

and their fusion with the next. In the trans Golgi network, proteins and lipids are sorted into transport vesicles destined for lysosomes, the plasma membrane, or secretion. The targeting of proteins to lysosomes requires mannose phosphorylation within the cis Golgi, and then concentration by a mannose 6-phosphate receptor in the trans Golgi network. Vesicles which bleb from the trans Golgi network carry the targeted proteins to one of several locations, including the lysosomes. *See* GOLGI APPARATUS.

Lysosomes. Lysosomes are membrane-bound organelles with a luminal pH of 5.0, filled with acid hydrolyses. Lysosomes are responsible for degrading materials brought into the cell by endocytosis or phagocytosis, or autophagocytosis of spent cellular material. *See* LYSOSOME.

Endocytosis and exocytosis. Movement of proteins to the plasma membrane or those destined for secretion occurs via secretory vesicles that also bleb off the trans Golgi network. These vesicles either travel directly to the plasma membrane and fuse, releasing their contents (constitutive pathway), or wait for a specific signal (usually elevated Ca^{2+}) before completing the secretion pathway (regulated pathway). The surface area of most plasma membranes does not increase due to the exocytosis that is described above, because membrane is being retrieved by the process of endocytosis. Endocytosis can be separated into pinocytosis (cell drinking) and phagocytosis (cell eating). Pinocytosis can be further subdivided into fluid-phase endocytosis and receptor-mediated endocytosis. In the former there is no concentration of small molecules being internalized, while the latter concentrates the ligand because of the presence of specific receptors within the internalized membrane. Much of the endocytosis occurs at membrane domains called coated pits, which internalize to form coated vesicles. The coat is largely composed of a protein called clathrin. Upon endocytosis or phagocytosis, the internalized endocytic vesicles or phagosomes, respectively, fuse with endosomes or lysosomes, leading to hydrolysis of the ingested molecules or particles. The resulting molecules can be released from the lysosomes into the cytoplasm for utilization. The membrane-associated receptors in receptor-mediated endocytosis are either recycled to the plasma membrane for reutilization or digested within the lysosomes.

Mitochondrion. The mitochondrion contains a double membrane: the outer membrane which contains a channel-forming protein named porin, and an inner membrane which contains multiple infolds called cristae. The porin channel allows molecules of less than 10,000 molecular weight free passage between the cytoplasm and the intermembrane space. The inner membrane, which contains the protein complexes responsible for electron transport and oxidative phosphorylation, is folded into numerous cristae that increase the surface area per volume of this membrane. The transfer of electrons from nicotinamide adenine dinucleotide (NADH) or flavin adenine dinucleotide ($FADH_2$) down the electron transfer chain to oxygen causes protons to be pumped out of the mitochondrial matrix into the intermembrane space. The resulting proton motive force drives the conversion of ADP plus inorganic orthophosphate (P_i) to ATP by the enzyme ATP synthetase. The matrix contains the mitochondrion's double-stranded DNA, which encodes 13 mitochondrial proteins. Therefore, the mitochondrion must import the vast majority of its protein from the cytosol. Proteins synthesized on free ribosomes are targeted to the mitochondria by an N terminal amphipathic, positively charged signal peptide, and they enter into the mitochondria through contact sites between the inner and outer membranes. The signal peptide is cleaved, and the remainder of the protein is unfolded and then enters the matrix. Targeting of proteins to the inner membrane, outer membrane, and intermembrane space is determined by specifically positioned start and stop transfer signals similar to endoplasmic reticulum import.

Peroxisomes. Peroxisomes import all of their proteins from the cytosol. In the case of the peroxisome, the targeting signal is three amino acids at the C terminus of the newly synthesized protein. Within the peroxisome, hydrogen atoms are removed from organic substrates and hydrogen peroxide is formed. The enzyme catalase can then utilize the hydrogen peroxide to oxidize substrates such as alcohols, formaldehydes, and formic acid in detoxifying reactions.

Cell division. Cell growth and division occur in an orderly, highly regulated, progressive series of events called the cell cycle. Advancement of cells through the cell cycle is driven by a small family of intracellular proteins, principal of which are the cyclins and a 34,000-molecular-weight protein, kinase. The cell cycle culminates in mitosis, a process that results in the production of two genetically identical diploid daughter cells. The chromosome separation events of mitosis occur as a result of the workings of the mitotic spindle apparatus, while the process of cytoplasmic division, or cytokinesis, is directed by the contractile ring. Unlike mitosis, meiosis produces nonidentical haploid cells. Meiosis, which occurs only in gonadal tissues, results in the production of either sperm or oocytes. By the process of fertilization, a new diploid is produced. Progression of cells through the cell cycle in multicellular organisms also can be affected by cell-cell interactions. This occurs when one cell secretes a growth factor that either stimulates or suppresses cell division in neighboring cells. Although cellular proliferation is tightly regulated in multicellular organisms, cells occasionally escape from these regulatory controls with resulting formation of cancerous tumors. Molecular analyses of the DNA in tumor cells have resulted in the identification of the mutated genes that allow for the onset of a cancer. The protein products of these genes, called protooncogenes and tumor suppressor genes, appear to play key roles in the regulation of cell growth in humans.

Cell adhesion. Tissues are formed by the interactions of cells with other cells and the extracellular

matrix. Cell-cell interactions can occur either by adherence of neighboring cells at distinct organized regions called junctions or by the binding of randomly scattered cell surface molecules to surface molecules on adjacent cells. Cell junctions can be classified according to function, with the integrity of most types of junctional complexes being dependent upon the presence of extracellular Ca^{2+}. Tight junctions form at the apical surfaces of cells and inhibit the leakage of ions and molecules from one side of an epithelial sheet to the other side, while gap junctions play a crucial role in cell-cell communication. The anchoring junctions share a common molecular organization which allows for the cytoskeleton of one cell to be attached to the cytoskeleton of neighboring cells. This provides tissues, particularly epithelia, with great strength. The desmosome, a type of anchoring junction that acts like a spot weld to hold cells together, interacts with intermediate filaments, while adherens junctions, which are anchoring junctions that allow for cell adhesion and cellular locomotion, utilize microfilaments. The extracellular matrix also is important for tissue formation. Glycosaminoglycans and proteoglycans form a hydrated gel that is the ground substance of the extracellular matrix. The ground substance is embedded with structural proteins such as collagen, which provides tensile strength to the extracellular matrix, and elastin, which is responsible for the resilient nature of the extracellular matrix. The molecules of the extracellular matrix are attached to one another and to the cells which are embedded in the matrix by interactions with adhesive glycoproteins. Adhesive glycoproteins, such as fibronectin, are the glue that holds connective tissues together. A specialized form of extracellular matrix called the basal lamina underlies all epithelia. *See* CELL ADHESION.

Signal transduction. To coordinate cellular activities in multicellular organisms, complex mechanisms of cell-cell communication needed to evolve. A variety of different types of signaling molecules are used by cells, and several different signaling pathways are utilized. In neuronal transmission, electrical signals are converted to chemical signals at synapses. The target cells, either other neurons or muscle cells, then convert the chemical signal back to an electrical signal and respond in an appropriate manner. Cells also can signal target cells at great distances by secreting signaling molecules, or hormones, directly into the bloodstream. Steroid hormones are lipid-soluble molecules that cross the plasma membrane and bind to intracellular receptors. The activated steroid hormone receptors then bind specifically to DNA sequences and directly regulate the transcription of adjacent gene sequences. Most hormones that are secreted by endocrine cells are water-soluble molecules that bind to specific receptors located on the surface of the target cells. The activated receptors are transmembrane proteins that are able to relay the extracellular signal across the plasma membrane to the cytoplasm. This is achieved by the cytoplasmic domain of the activated receptor molecule, which binds to and either activates or represses the activity of a cytoplasmic G-protein. Specific G-proteins then interact with target enzymes located on the inner surface of the plasma membrane to regulate the activity of these proteins. One of these enzymes, adenylate cyclase, is responsible for converting ATP to complementary adenosine monophosphate (cAMP). cAMP then sets off a reaction cascade inside cells by allosterically regulating certain molecules, thereby completing the signal transduction pathway. A second enzyme on the cell surface that is activated by G-proteins is phospholipase C. Phospholipase C cleaves a membrane lipid, phosphatidyl inositol 4,5-bisphosphate (PIP_2), into inositol 1,4,5-trisphosphate (IP_3) and diacylglycerol. IP_3 triggers the release of Ca^{2+} ions from intracellular storage sites, and the Ca^{2+} ions then bind to and activate calcium-binding proteins such as calmodulin. Diacylglycerol, however, activates protein kinase C, completing the signal transduction pathway in the cell.

Cell motility. The ability of eukaryotic cells to move through a substrate requires cytoskeletal proteins, which generate propulsion, and constant remodeling of the membrane by endocytosis and exocytosis. The stages in cell movement are the polarization of the cell and identification of a leading edge, extension of a lamellipodium or ruffled edge in the direction of movement, and breakdown of attachment with the extracellular matrix. Sometimes a cell moves toward or away from a particular chemical signal (chemotaxis). This occurs because the chemical binds to receptors on the surface of the motile cell, increasing intracellular Ca^{2+} at the leading edge. The role of Ca^{2+} in creating a motoring force is not clear, but it may act through Ca^{2+}-dependent severing proteins to cleave actin filaments both within the leading-edge membrane skeleton and in the cell cortex. The result would be a solation of the cytoplasm in the direction of motion, a membrane capable of being extended by exocytosis, and a stabilization of the extension by the rapid polymerization of actin in a direction parallel to the leading-edge lamellipodium. The migratory cell is then steered as a result of the microtubule organizing center and Golgi apparatus being realigned along an axis between the nucleus and leading edge. The Golgi apparatus produces the secretory vesicles which fuse with the leading edge of the plasma membrane, and these vesicles are transported to the leading edge by microtubule-based translocation. Membrane from the old leading edge then ruffles back toward the trailing end of the cell. During the process of cell migration, the motile cell must make and break interactions with the extracellular matrix in a traction-producing process. *See* CELL MOTILITY.

Plant cell. Outside their plasma membrane, plant cells have an extremely rigid cell wall. This cell wall is composed of cellulose and other polymers and is distinct in composition from the cell walls found in fungi or bacterial cells. The plant cell wall expands during cell growth, and a new cell wall partition is created between the two daughter cells during cell

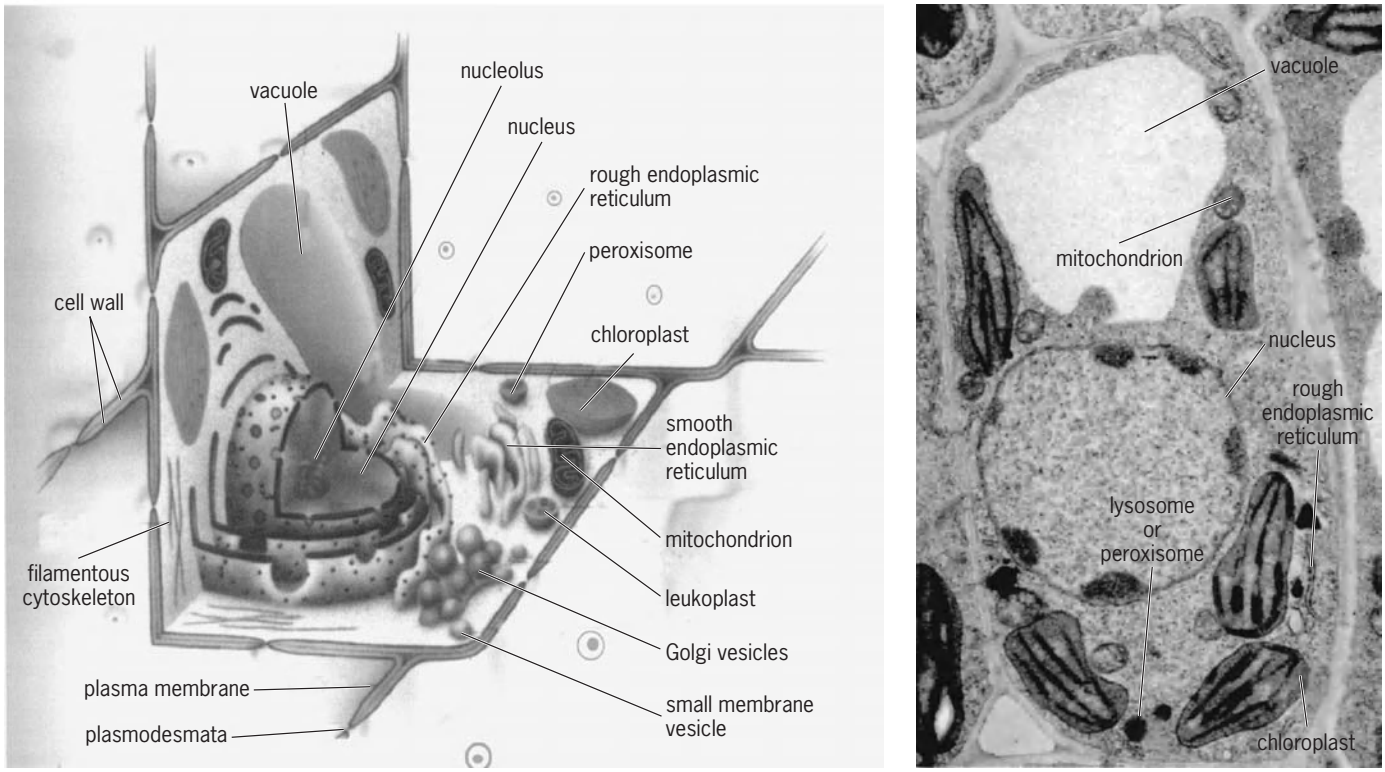

Fig. 3. Plant cell structure. (*a*) Artist's rendition of a typical plant cell (*modified from H. Lodish et al., Molecular Cell Biology, 3d ed., Scientific American, New York, 1999*). (*b*) Electron micrograph of a plant cell (*Electronmicroscopy Core Lab, Interdisciplinary Center for Biological Research, University of Florida*).

division. Similar cell walls are not observed in animal cells (**Fig. 3**).

Most plant cells contain membrane-encapsulated vacuoles as major components of their cytoplasm. These vacuoles contain water, sucrose, ions, nitrogen-containing compounds formed by nitrogen fixation, and waste products. The amount of osmotic particles within the vacuoles is much higher than in the cytosol or extracellular fluid. The encapsulating membrane is permeable to water but impermeable to the molecules within the vacuole. Therefore, water rushes into the vacuole, the vacuole expands, and a hydrostatic pressure or turgor is created within the cell. This turgor is resisted by the cellulose-containing cell wall, but as the cell wall stretches the plant grows. *See* PLANT CELL.

Chloroplasts are the other major organelle in plant cells that is not found in other eukaryotic cells. Like mitochondria, they are constantly in motion within the cytoplasm. One of the pigments found in chloroplasts is chlorophyll, which is the molecule that absorbs light and gives the green coloration to the chloroplast. Chloroplasts, like mitochondria, have an outer and inner membrane. Within the matrix of the chloroplast there is an intricate internal membrane system. The internal membranes are made up of flattened interconnected vesicles that take on a disc-like structure (thylakoid vesicles). The thylakoid vesicles are stacked to form structures called grana, which are separated by a space called the stroma. Within the stroma, carbon dioxide (CO_2) fixation occurs, in which carbon dioxide is converted to various intermediates during the production of sugars. Chlorophyll is found within the thylakoid vesicles; it absorbs light and, with the involvement of other pigments and enzymes, generates ATP during photosynthesis. Like mitochondria, chloroplasts contain their own double-stranded circular DNA sufficient to encode a few essential proteins. Therefore, most chloroplast proteins must be imported from the cytosol following synthesis on free ribosomes.

Steven R. Goodman

Bibliography. B. Alberts et al., *Molecular Biology of the Cell*, 3d ed., 1994; S. R. Goodman, *Medical Cell Biology*, 2d ed., 1998; H. Lodish et al., *Molecular Cell Biology*, 3d ed., 1995; A. G. Moat and J. W. Foster, *Microbial Physiology*, 3d ed., 1995.

Cell, spectral analysis of

Living cells contain various substances, the concentrations and biological activity of which can be investigated by observing the spectrum of light passed through the cells. Such investigations take advantage of the fact that many substances absorb light in an individually characteristic manner. Thus, the spectrum of light passed through a green leaf has two black bands where red and blue light should appear. The absorption of red and blue light is characteristic of the chlorophylls, the photosynthetic pigments. Today, investigations of cells by optical methods go

far beyond the routine analysis of brightly colored pigments that are found widespread and in high concentration, for example, hemoglobin and chlorophyll. Such methods allow the investigation of light-absorbing molecules within the cell whose concentration is 1000 times smaller that that of hemoglobin or of chlorophyll. *See* CHLOROPHYLL; HEMOGLOBIN.

Because the color of biological molecules changes when they undergo chemical reactions, such reactions in the cell can be monitored by spectral analyses. These analyses can be used to monitor reactions occurring in times ranging from 10^{-15} s to minutes. In some cases, the spectral properties can also indicate the environment of the biological molecules in the cell, that is, whether they are rigidly held or free to move and how they react with one another within the cell.

Useful kinds of spectra. Different parts of the spectrum provide different information about molecules in cells.

Infrared spectra (700–5000 nanometers) give information about the structure of molecules; all molecules absorb in the infrared in a characteristic manner.

Visible light spectra (400–700 nm) give information concerning those relatively few biological molecules that absorb light in this region. These molecules thus can be specifically studied in a cell that may contain tens of thousands of other types of molecules.

Ultraviolet spectra (200–400 nm) give information on those molecules that absorb light in this region. Such spectra are not very useful when working with living cells or other light-scattering samples.

Fluorescence and phosphorescence spectra are produced by light emission. When some molecules absorb light in the ultraviolet and visible spectral regions, they can be energized into various electronic excited states. In many instances, the energy of the excited states is dissipated rapidly as heat, and the molecule returns to its original ground state. However, the appropriate excited state can also be dissipated relatively slowly by emitting light at wavelengths slightly longer than that of the excitation light. The emitted light can take one of two forms: fluorescence, which occurs rapidly after excitation and lasts 10^{-9} to 10^{-6} s; and phosphorescence, which has a longer decay time and wavelength range than fluorescence and lasts on the order of 10^{-6} to 10^{2} s. Relatively few molecules emit light after absorption.

Light scattering is of two types. When light is scattered without changing wavelength (elastic scattering), the scattering reveals the size and shape of molecules. Another type of scattering is called Raman scattering. In this case, molecules alter the light by slightly shifting the wavelength in a manner that is very specific for the particular molecule. *See* RAMAN EFFECT; SCATTERING OF ELECTROMAGNETIC RADIATION.

Magnetic resonance spectroscopy techniques frequently can resolve components of cells that cannot be visualized by optical methods. *See* ELECTROMAGNETIC RADIATION; MAGNETIC RESONANCE.

Intracellular substances. Spectral analysis by optical methods is limited to those substances which show characteristic peaks, or maxima, of absorption when light absorption is plotted as a function of light wavelength. Whether or not a given substance can be detected depends on the intensity of absorption relative to other substances present at a given absorption maximum. The molecular extinction coefficient quantitatively describes the light absorption of a given substance at a given wavelength. The dependence of absorption A on extinction coefficient ϵ, concentration c, and optical path l is expressed as Beer's law, $A = 10 \exp(-\epsilon \cdot C \cdot l)$. This dependence imposes limitations on the application of optical methods to the study of living cells; only those substances with a high extinction coefficient can be investigated successfully.

Hemoproteins. There are a variety of substances present within living cells, but the number of those of biological interest which also possess a useful extinction coefficient is limited; often these substances are pigments. A large group of biologically interesting pigments belongs to the class known as hemoproteins. Included in this group are the hemoglobins, myoglobins, catalase, peroxidase, and the cytochromes, a large family of hemoproteins that are universally present in all cells. Reduced hemoproteins are characterized by well-defined absorption maxima in the visible spectrum, while the spectra of oxidized hemoproteins are quite different. The absorption spectrum of each hemoprotein is characteristic and individual to that hemoprotein only, a property that gives a considerable degree of specificity to the optical study of these components in cells. Among the hemoproteins, the cytochromes are of particular interest because of the intimate role they play in various processes, for example, in cellular respiration and in photosynthesis. From the cytochrome spectra obtained from an investigation of cells, the ratio of oxidized to reduced forms (the oxidation state) can be determined. The oxidation state of the cytochromes, in turn, depends on the conditions under which the cell is placed. Given any steady-state condition, the oxidation state of each cytochrome component in the cell will reach a definite value, a value which can be determined by spectral analysis. *See* CYTOCHROME.

Flavoproteins and pyridine nucleotides. Other substances of biological interest that are found in cells and that can be studied optically include flavoproteins, a family of proteins that occur abundantly in nature, and the pyridine nucleotides. As in the hemoproteins, the spectral properties of flavoproteins and pyridine nucleotides depend on their oxidation state. Flavoproteins in their oxidized state absorb light in the yellow portion of the visible spectrum and fluoresce strongly in the green; light absorption and fluoresence is quenched when these proteins are reduced. Reduced pyridine nucleotides, NADH and NADPH, show strong light absorption in the blue portion of the spectrum and also fluoresce strongly at longer wavelengths, also in the blue. In the oxidized forms NAD^+ and $NADP^+$, the molecules do

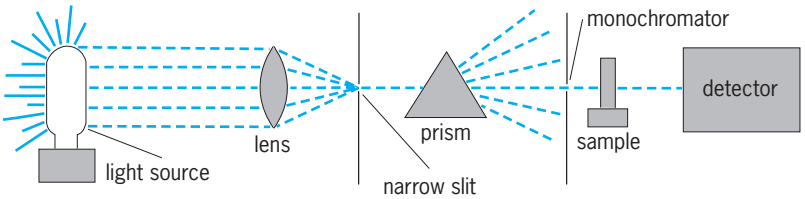

Fig. 1. Basic components used for spectral analysis.

not absorb light or fluoresce. Because of the strong fluorescence, it is often most convenient to study these components by fluorescent techniques.

Iron-sulfur clusters. There is a large group of proteins, the iron-sulfur proteins, which are of considerable biological importance. These substances, referred to individually as iron-sulfur clusters, do not have the properties that can be characterized by optical spectroscopy. However, they can be studied in considerable detail by magnetic resonance spectroscopy, a technique where the spectrum involves magnetic fields and light in the microwave spectral region. This technique is also used to study the position of protein complexes relative to one another in biological membranes.

Subcellular structures. The living cell is a highly organized biological unit. Within the cell there are a variety of structures, each one of which carries out specific functions. The necessary interactions between the subcellular structures that must provide for the normal processes of cell functions and metabolism are not yet well understood. Examples of the well-organized subcellular structures are the mitochondria, where the respiratory processes of the cell are localized, microsomes, peroxisomes, and chloroplasts. In many instances, the study of suspensions of cells, for example, yeast and algal cells, has been reasonably fruitful. However, because of light-scattering problems and the limited numbers of single cells available, it is extraordinarily difficult to study specific processes within a single cell. Rather, it has proven much more profitable to investigate suspensions of subcellular organelles which can be isolated in a reasonably undamaged state, for example, mitochondria and chloroplasts. Such suspensions, prepared by destroying the cell structure, contain a variety of substances suitable for optical, fluorescent, and magnetic resonance analysis and in sufficient concentration to allow precise and quantitative study. Suspensions of mitochondria and chloroplasts have been used extensively in the development of the understanding of cellular respiration and of photosynthesis. *See* CELL (BIOLOGY); CELL PLASTIDS; MITOCHONDRIA.

It is characteristic of biological processes that various proteins, enzymes, and coenzymes react with one another in cycles or sequences. In the processes of cellular respiration, there is a sequence of reactions which involve pyridine nucleotides, flavoproteins, iron-sulfur clusters, and cytochromes in processes involving energy transfer and oxygen consumption. Similarly, photosynthesis involves chlorophylls and accessory pigments, cytochromes, pyridine-nucleotides, and iron-sulfur clusters. It is particularly important to discover which components are included in various biological processes and the order in which they react with one another; spectral analysis provides the methods for such investigations.

Methods. Any spectral analysis, regardless of its application, utilizes the components in **Fig. 1.** These components consist of a light source from which the light is focused by means of a lens on a narrow slit. The narrow beam of light is then resolved into its component wavelengths by means of a prism or a grating. Lasers have been used as a source of monochromatic light. A narrow band of such light is isolated by means of a second narrow slit and passes through the sample onto an appropriate detector. The signal produced on activation of the detector may be amplified and displayed or recorded on a suitable device. Instruments for measuring light absorption at varying wavelengths utilize the above basic components and are suitable for the study of clear solutions—solutions that do not impose any restrictions, other than those of Beer's law, on the amount of light passing through them.

The spectral analysis of living cells and of subcellular organelles is complicated because some suspensions of particulate matter scatter light and thus limit the amount of light that can pass through them. Suspensions of biological material scatter light by refraction and reflection, and the problem is intensified by changes in the state of the particulate matter in the suspension—such as changes of size with time: swelling and shrinking. Problems of light scattering are magnified by monochromatic light, where high light intensities are difficult to obtain. The most important problem is the impossibility of observing spectra of specific light-absorbing species. Problems of light scattering can be avoided in various ways, for example, by using a low-dispersion spectroscope or a dual-wavelength spectrophotometer.

Microspectroscope. Much of early understanding of blood and muscle pigments and of cellular respiration was obtained through the use of the microspectroscope (**Fig. 2**). Basically, this instrument is a hand spectroscope that uses the human eye as the detector. Within a narrow but biologically interesting spectral region, the human eye is capable of sharp delineation between light and dark; hence, absorption bands can be observed with a considerable degree of accuracy. This instrument is now regarded as a museum piece, but many investigators still find that it is useful for quick observations.

Spectrophotometers. Special photoelectric devices have almost completely replaced visual spectroscopy for examination of light-scattering samples.

The split-beam spectrophotometer measures the difference in absorption between two nearly identical biological suspensions. Monochromatic light is passed alternately through the two samples by means of a vibrating mirror and onto a light detector. If the suspensions under study are identical except for their oxidation state, light scattering will be the same in both, and the light detector will measure only the

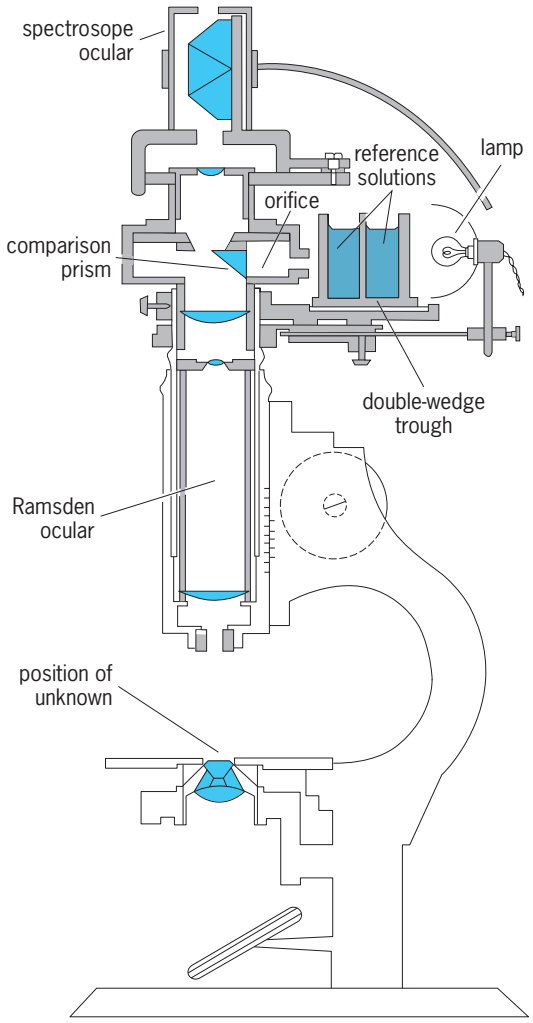

Fig. 2. Microspectroscope used by an observer to view the spectra of the specimen and reference solution simultaneously. (*After D. Keilin and Y. L. Wang, Haemoglobin of Gastrophilus larvae, Biochem. J., 40(5,6):855–866, 1946*)

difference in light adsorption between the two samples. The result is an oxidized-minus-reduced difference spectrum and will give the absorption maxima between the reduced and oxidized redox components present in the sample. Such a spectrophotometer is capable of scanning difference spectra in the visible and near-ultraviolet regions (**Fig. 3**).

The dual-wavelength spectrophotometer uses two sources of monochromatic light that are passed through a single light-scattering sample, again by means of a vibrating mirror. If one light beam is set at a known absorption maximum and the other at wavelength where there is no absorption change between the oxidized and reduced species, the photoelectric circuit will measure the true absorption difference between these two wavelengths and will reject almost completely changes of light transmission that are independent of wavelength. With such a device, one can measure rates of oxidation or reduction of redox components. As with the split-beam spectrophotometer, difference spectra of redox components of biological samples can be scanned.

Temperature studies. The light-absorbing properties of many pigments important to the economy of living cells are altered at low temperatures. Not only is there a marked sharpening of the absorption bands at low temperature, but also a considerably increased absorbancy because of increased path length due to reflections from interfaces on the ice surfaces. Thus, the use of low temperatures, 77 K and below, increases the sensitivity and specificity of detection. The effects of low temperature are illustrated in **Fig. 4** in a suspension of yeast cells, measured at 77°F (25°C) and at 77 K (−196°C). Different path lengths were used at each temperature; hence, the increased absorbancy at low temperature is not shown. The 77 K spectrum clearly shows the sharp delineation of absorbing components that can be realized at that temperature. Note that there is a shift of absorption maxima toward shorter wavelengths at 77 K.

Microspectrophotometer. Spectrophotometers, utilizing the principles discussed above but including a microscope for the localization of the object under study, have also been developed. Such microspectrophotometers are capable of carrying out spectral analyses within the dimensions of a single cell.

Ultraviolet studies. Such studies are difficult to carry out with intact cells and tissues, light-scattering problems increase with decreasing wavelengths, and there are too many substances with low extinction coefficients and overlapping absorbancies in this spectral region. Pertinent information is most easily obtained by using clear extracts of cells and tissues. Extracts can be prepared that contain nucleic acids, amino acids, various metabolites, and quinones, for example, the ubiquinones.

Infrared studies and Raman spectra. Both of these reflect the vibrational structure of the molecule. Water absorbs in the infrared region, so most infrared studies are done on concentrated solutions in order to increase the signal over that of water. Infrared studies are particularly useful for monitoring protein conformation and the structure of lipids in membranes.

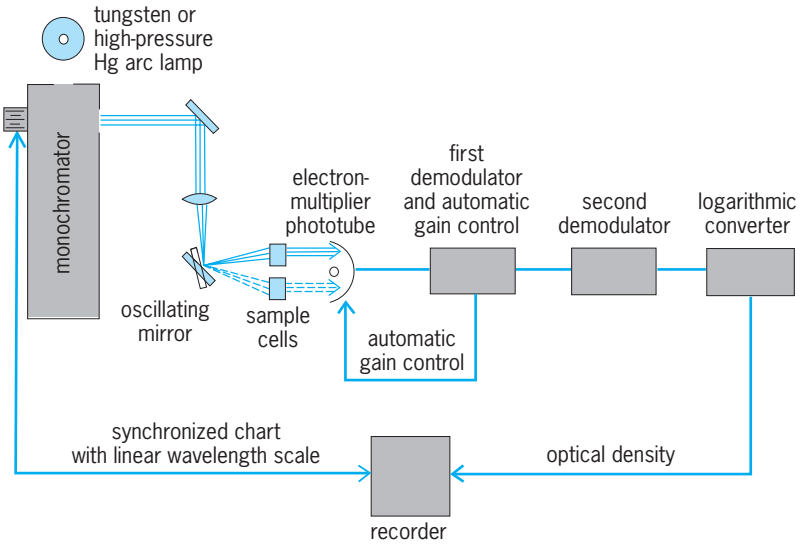

Fig. 3. Split-beam recording spectrophotometer.

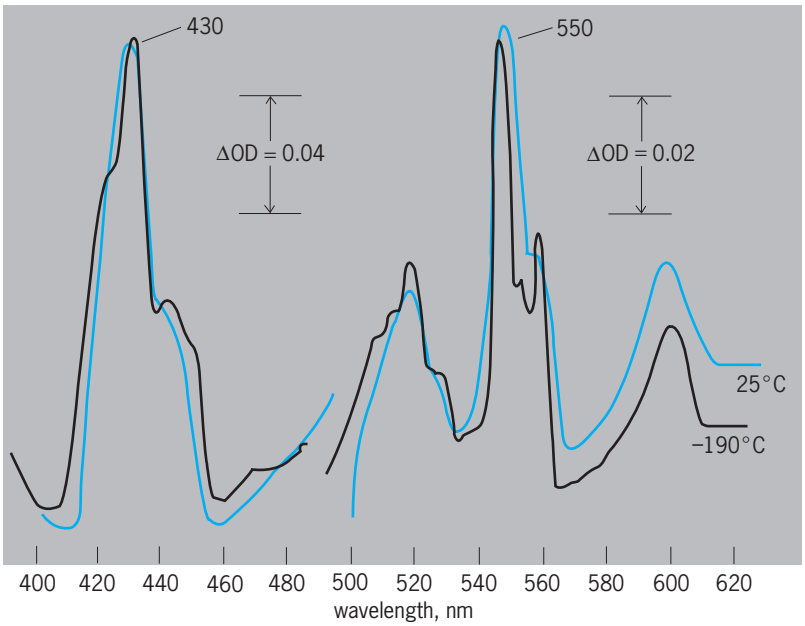

Fig. 4. Spectra, by difference, of the pigments present in suspensions of yeast cells. These spectra show the changes in absorbency plotted as a function of wavelength. ΔOD is optical density increment.

Raman spectra are relatively weak, but very useful for studying the heme proteins, where they can help determine the structure of the heme.

Fluorescence and phosphorescence. Fluorescence is widely used as an analytical tool for the investigation of cells. Similar to the spectrophotometer, a spectrofluorimeter consists of a light source from which the light is made monochromatic by a prism or grating. Instead of measuring how much light is absorbed, however, the light that is emitted from the sample is measured. The light is emitted at higher wavelengths than those of the absorbed light, and it is measured at right angles relative to the exciting light by passing it through a second monochromator. In the absorption instrument the difference in intensity between two light beams is measured. In the fluorescence spectrofluorimeter the absolute value of light intensity is measured. Because single photons of light can be measured with modern equipment, fluorescence can measure very low concentrations of molecules. However, the technique is limited to only those molecules which fluoresce.

Fluorescence measurements are especially adaptable for investigating the environment of molecules. The rotational motion of a molecule depends upon its size and viscosity. These can be measured by fluorescence polarization. In this technique the molecule is excited with polarized light. If the molecule moves in the time scale of the fluorescent excited state (1–100 nanoseconds), the emitted light will be partially depolarized, the amount of depolarization depending upon the molecular rotation. This technique has been used to determine molecular size and the viscosity of the environment around cells.

Phosphorescence can be used in much the same way as fluorescence, but molecular oxygen interferes with the phosphorescence light intensity. By measuring phosphorescence intensity or lifetime, the concentration of oxygen within the cells can be determined. *See* FLUORESCENCE; PHOSPHORESCENCE.

Measurement of cell component interaction. The reactions of the cell occur very specifically, for example, if A→B→C, the scheme A→C→B will never occur. Spectral analysis is very useful for finding the order and rate of reactions of cell components. A common method for carrying out such experiments is to initiate the reactions with a strong light signal, usually from a laser, and then to follow the reactions by measuring the absorption of the components with a weaker light. This method has been used to study photosynthesis, where reactions occur in as fast as 1 picosecond (10^{-12} s).

Walter D. Bonner, Jr.; Jane M. Vanderkooi

Bibliography. C. R. Cantor and P. R. Schimmel, *Biophysical Chemistry II: Techniques for the Study of Biological Structure and Function*, 1980; S. P. Colowick and N. O. Kaplan (eds.), *Special Techniques for the Enzymologist*, vol. 4 of *Methods in Enzymology*, 1957; D. C. Youvan and B. L. Marrs, Molecular mechanisms of photosynthesis, *Sci. Amer.*, 256(6):42–49, 1987.

Cell adhesion

The process whereby cells interact and attach to other cells or to inanimate surfaces. Cell adhesion is mediated by cell surface proteins and associated macromolecules. It forms the physical basis of multicellularity, morphogenesis and embryonic development, tissue integrity, host–pathogen interactions, immune system function, and the ecological integration of microbial communities. In addition to attaching cells to one another, cell adhesion systems act to coordinate a wide range of cellular behaviors. In animal cells, sites of adhesion serve to organize the system of intracellular protein filaments, known as the cytoskeleton, responsible for cell shape and movement. Adhesive interactions are critical for juxtacrine signaling, by which neighboring cells communicate. Juxtacrine signals regulate gene expression and cell behavior and determine whether a cell survives, divides, differentiates, or undergoes programmed cell death (apoptosis).

Cell–Cell Interactions in Microorganisms

Although it is common to think of microorganisms, for example bacteria and yeasts, as isolated individuals, more often these organisms live within complex communities known as biofilms.

Biofilms. Within a biofilm, cells are embedded in an extracellular matrix that they themselves secrete. Within this matrix the organisms interact in complex ways. Their properties, for example their susceptibility to antibiotics as well as their metabolic requirements, are often dramatically different from those of the same organism living in isolation (monoculture). In fact, only a small percentage of known microorganisms can be grown in laboratory monoculture; most are dependent upon interactions with

their neighbors. Within a biofilm, organisms cooperate, compete, and prey upon one another. Many microbially derived antibiotics have their evolutionary origins in such ecological interactions. Biofilms form rapidly on wet surfaces, as in the overnight appearance of plaque on teeth. The formation of biofilms is of great practical significance for any system in which microorganisms can colonize surfaces. *See* ANTIBIOTIC; BIOFILM; CORROSION.

Adhesins. The adhesion of microorganisms to the cells of higher plants and animals is a critical prerequisite for causing disease. Such interactions are mediated by molecules, generally cell-surface proteins, known generically as adhesins (**Fig. 1**). Adhesins interact specifically with host cell surface proteins and glycolipids, which serve as pathogen receptors. The true functions of these "receptors" have nothing to do with pathogenesis: the receptors have simply been exploited by the pathogen. In some cases, the pathogen itself can synthesize and secrete its own receptor, which is taken up by host cells and then displayed on the cell surface (Fig. 1*c*). Once a pathogen adheres to a target cell, either it can modify the behavior of the cell (for example, *Yersinia pseudotuberculosis*) or, in cases such as *Listeria monocytogenes*, it can actually invade the cell and live within the cytoplasm. *See* LISTERIOSIS; PATHOGEN; YERSINIA.

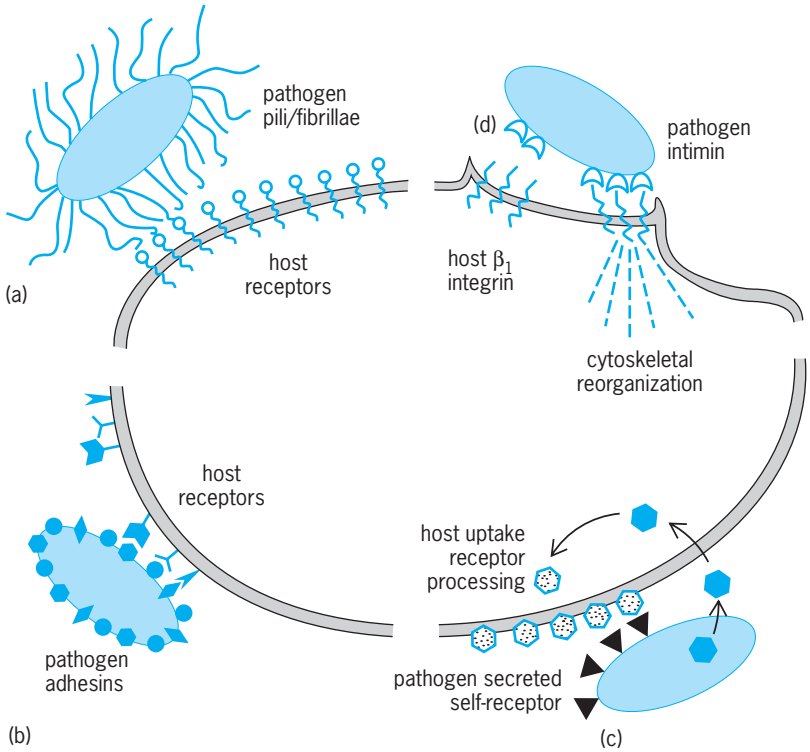

Fig. 1. Adhesive interactions between pathogen and host are central to the process of infection. Were it not for their ability to adhere to host cells, most pathogenic fungi, bacteria, and viruses would be rapidly removed from host-cell surfaces. By specifically adhering to host-cell surface molecules, a pathogen can target itself to a hospitable environment. Such adhesive interactions involve a number of mechanisms, including (*a*) fibrous pili or fibrillae, (*b*) adhesin-type carbohydrate binding proteins, or (*d*) intimins, which bind to surface integrins. In some cases, (*c*) the pathogen itself secretes its own receptor, which is taken up by the target cell; this self-receptor is then placed on the host cell surface, where it mediates pathogen adhesion. (*Adapted from T. M. Wizemann, J. E. Adamou, and S. Langermann, Adhesins as targets for vaccine development, Emerg. Infect. Dis., 5(3), 1999; http://www.cdc.gov/ncidod/eid/vol5no3/wizeman.htm*)

Cell–Cell Interactions in Multicellular Organisms

Cell–cell interactions become more elaborate in true multicellular organisms such as plants and animals. These interactions play a critical role in processes such as embryonic development and the maintenance of tissue integrity, in which they have been studied extensively. In most cases, plant cells are rigidly constrained by a specialized extracellular matrix, the cell wall, which limits cell–cell movement. Dynamic adhesive interactions do occur, however, during the process of fertilization. In animal cells, adhesion is more dynamic and very highly specialized. The regulation of cell adhesion is central to morphogenesis (the process by which cells rearrange to form tissues and organs), the development of the nervous system and the specialized connections that underlie coordinated movement, memory, and cognition, and the functioning of the immune system. Aberrations in cell adhesion mechanisms underlie metastatic processes associated with cancer progression. *See* ANIMAL MORPHOGENESIS; CANCER (MEDICINE); FERTILIZATION; PLANT MORPHOGENESIS.

Origins of cell–cell adhesion. The origins of cell–cell adhesion are closely associated with the origins of true multicellularity. There are a number of protozoa, such as the choanoflagellates, that can exist in either unicellular or multicellular colonial forms. Based on the morphological similarities in cell structure, and now supported by molecular phylogenetic data, it appears that choanoflagellates are close evolutionary relatives to the animals. In true metazoans, such as the sponges (Porifera), hydra, corals, and jellyfish (Cnidaria), the number of distinct cell types increases as does the repertoire of possible cell–cell interactions. These interactions, between cells and between cells and extracellular matrix, are mediated by specific membrane proteins and accessory factors. *See* CELL (BIOLOGY); CELL DIFFERENTIATION; CHOANOFLAGELLIDA.

Cell-surface adhesion proteins. Adhesion is mediated either directly by macromolecules embedded in the plasma membrane or indirectly through molecules deposited on the cell surface. In animal cells, a number of distinct protein families mediate cell–cell and cell–extracellular matrix interactions (see **table**).

Cadherins. Among the best-studied protein families are the cadherins. These proteins share a common structural motif, the cadherin domain (**Fig. 2**). A classical cadherin contains an extracellular domain composed of five linked cadherin domains, a transmembrane domain, and a "cytoplasmic" domain. Cadherins of the same type interact in a homotypic (like-to-like) manner through their extracellular domains. Different cell types synthesize different types of classical cadherins, which enables them to adhere to cells of the same type. The family of membrane proteins that contains cadherin domains is large: over 100 such genes encoding these proteins have been identified in humans. These cadherin

Some major cell adhesion proteins of animals

Protein name	Known function	Details
Cadherins	Cell–cell adhesion	Typically homotypic
Protocadherins	Cell–cell adhesion	Synaptic adhesion
Selectins	Cell–cell adhesion	Interactions between lymphocytes and capillary cells
Ig-family adhesion molecules	Cell–cell adhesion	Diverse developmental events
Integrins	Cell-matrix interactions	Heterodimeric
Notch	Juxtacrine signaling	Transcellular signaling
T-cell receptor/MHC	Juxtacrine signaling	Transcellular signaling
Fas/Fas ligand	Juxtacrine signaling	Regulation of apoptosis

family proteins perform a wide range of adhesive and signaling functions. *See* CELL MEMBRANES; GENE; PROTEIN.

The intracellular domain of a classical cadherin protein interacts with cytoplasmic proteins known as catenins. Catenins mediate the binding of cytoskeletal polymers to the site of cell–cell adhesion, an adherence junction. They also regulate the stability of cadherin proteins. In vertebrate cells there are two generic types of cadherin-mediated cell–cell junctions: adherens junctions, which link to actin filaments, and desmosomes (macula adherens), which anchor intermediate filaments. Variants of these junctions are involved in linking the cells of the heart, skin, and vasculature together. Coordinated changes in adhesion junction–associated cytoskeletal systems are responsible for changes in the morphology of tissue sheets. Defects in either the cytoskeletal system or the adherence junction are associated with cellular fragility diseases, such as epidermolysis bullosum simplex.

Protocadherins. In addition to the classical cadherins, there is a family of cadherin-domain proteins, the protocadherins, that can be distinguished based upon gene organization. Many of these protocadherin genes are organized in a conventional manner, whereas others are more complex. The genes for the mammalian α and γ protocadherins consist of 14–22 regions (exons), each of which encodes a distinct extracellular and transmembrane domain; there is also a set of three exons that encodes a common cytoplasmic domain. Differential RNA splicing (a process by which different messenger RNA molecules are produced from a single gene transcript by splicing together different sets of exon regions) is used to generate a diverse set of protocadherin proteins from a single gene. Currently it is thought that this molecular diversity may be used to establish and maintain specific cell–cell contacts within the developing nervous system.

Selectins. Selectins are another class of membrane proteins associated with the invasion of white blood cells (lymphocytes) into sites of infection. The overall structure of the selectins is similar to that of the cadherins. Their extracellular domain consists of a series of repeating motifs that ends in a "carbohydrate recognition," or lectin, domain. As lymphocytes pass through the bloodstream, they interact weakly with carbohydrates associated with the surface of endothelial cells, which line blood vessels. However, in regions of cell damage, inflammatory signals transmitted by soluble proteins known as cytokines are released by the damaged or infected cell and lead to modifications in the surface of nearby capillary endothelial cells. In regions of cytokine signaling, lymphocytes (which normally move through the capillary) interact with selectins, causing them to adhere to and roll along the capillary's inner surface. Interactions with selectins together with other signaling/adhesion molecules lead lymphocytes to strongly adhere to, and then migrate between, the endothelial cells, moving out of the bloodstream

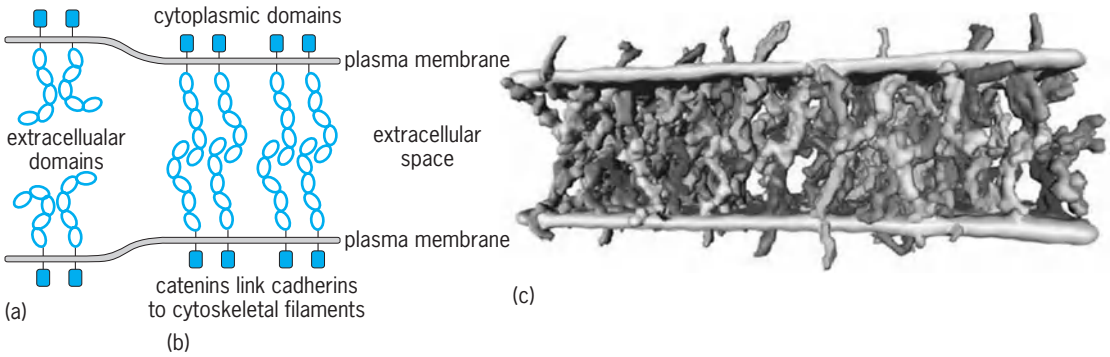

Fig. 2. Cadherins mediate homotypic (like-to-like) adhesion. (*a*) Cadherin proteins project outward from the cell surface. (*b*) When two cells come together, the opposing extracellular domains of the cadherins interact. This interaction is then stabilized through interactions with the cytoskeleton mediated by catenin proteins. Various electron microscopy techniques have been used to generate a more exact description of the adhesive interactions. (*c*) A section of the extracellular portion of a desmosome is modeled; the extracellular domains of individual cadherin molecules are indicated by various shades. (*Reprinted with permission from W. He, P. Cowin, and D. L. Stokes, Untangling desmosomal knots with electron tomography, Science, 302:109–113, 2003. Copyright © American Association for the Advancement of Science*)

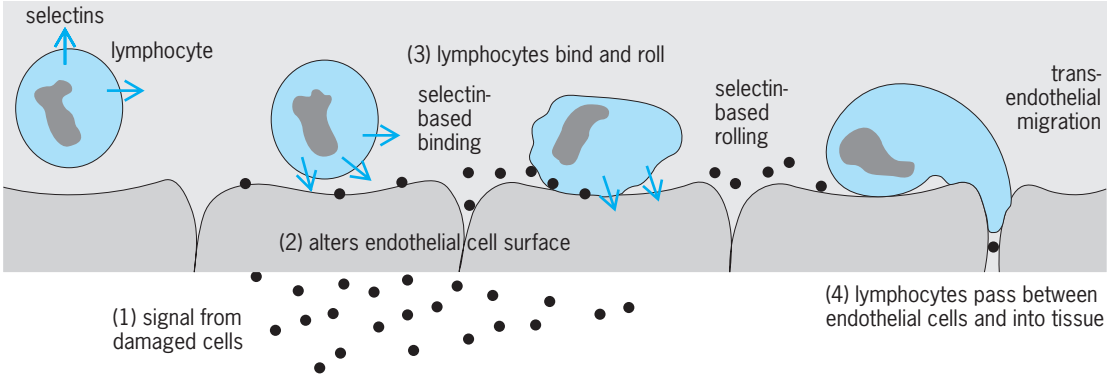

Fig. 3. Regulation of adhesion within the immune system. Normally, white blood cells (lymphocytes) migrate throughout the body via the bloodstream. However, in regions of infection or cell damage, signals from the infected cells lead to changes in the capillary endothelial cells that enable lymphocytes to adhere to the capillary surface via their selectins. The lymphocytes roll along the capillary surface and then migrate between the endothelial cells and into the tissue, where they can protect against the spread of infectious agents such as bacteria and viruses.

and into the damaged tissue (**Fig. 3**). *See* BLOOD; CYTOKINE; INFLAMMATION; LECTINS.

Immunoglobulin superfamily adhesion proteins. A third major group of cell adhesion proteins are built from a structural domain that is also found in immunoglobulin (antibody) molecules. These Ig-domain adhesion proteins can mediate either homotypic or heterotypic adhesive interactions. The genes that encode these proteins are complex, and a large number of variants can be generated from each through a process of differential splicing.

Types of cell adherence junctions. In addition to the adherens junctions and desmosomes (**Fig. 4***a*), there are other distinct types of cell–cell junctions.

Hemidesmosomes. Hemidesmosomes mediate interactions with extracellular matrix fibers (Fig. 4*c*). Like desmosomes, they anchor intermediate filaments to the plasma membrane and act to mechanically strengthen tissues.

Tight junctions. Tight junctions (zonula occludens) are typically formed between epithelial and endothelial cells (Fig. 4*a*). Composed of strands of membrane proteins, tight junctions act to block the movement of extracellular fluids and molecules through the spaces between cells. The disruption of tight junction integrity has been linked to multiple sclerosis and other pathogenic conditions.

Gap junctions. Gap junctions, on the other hand, act to mediate the movement of small molecules directly from the cytoplasm of one cell to that of its neighbor (Fig. 4*b*). Gap junctions form when two cells are so closely opposed to one another that connexin hex-

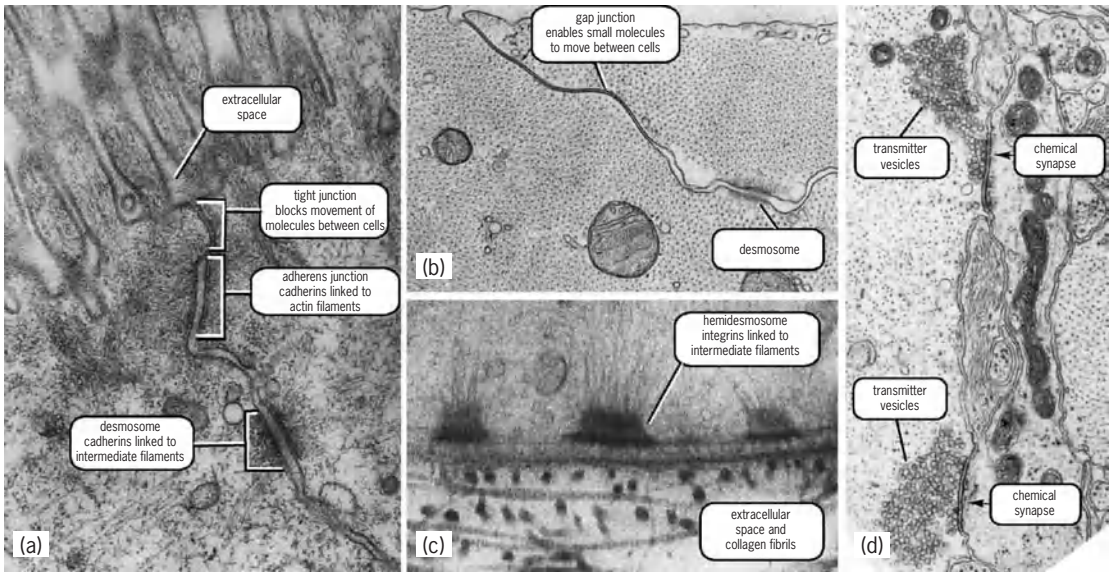

Fig. 4. Gap and tight junctions. Cell are connected by a number of different junctions. (*a*) In a classical electron micrograph of rat intestinal epithelia, the cells are linked by a complex of tight junctions, adherens junctions, desmosomes, and gap junctions (not shown). Tight junctions block molecular diffusion between cells, whereas adherens junctions and desmosomes anchor cytoskeletal systems. (*b*) Gap junctions are particularly prominent in cells that are electrically coupled to one another, such as these heart cells. (*c*) Cell layers are attached to extracellular matrix components through hemidesmosomes which anchor intermediate filaments in this electron micrograph of salamander skin. (*d*) An electron micrograph of lamprey nervous tissue showing typical neuronal synapses, which are modified adherence junctions specialized to facilitate neurotransmitter release by vesicle fusion and receptor. (*Reprinted with permission from D. W. Fawcett, The Cell, 2d ed., W. B. Saunders, 1981. Copyright © Elsevier Inc.*)

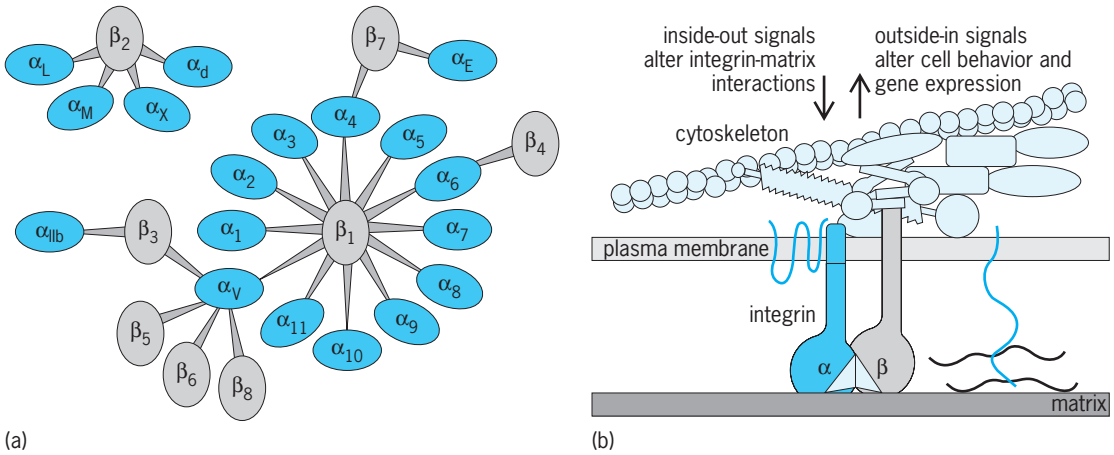

Fig. 5. Integrin family. The integrins are heterodimeric proteins, composed of α and β polypeptides. (*a*) Certain subunits are found in specific pairs; for example, the β_1 polypeptide can form functional heterodimers with many different α-type polypeptides, whereas the β_4 polypeptide is known to pair only with α_6. (*b*) The integrins are membrane proteins—their extracellular domain interacts with components of the extracellular matrix, while their intracellular domain interacts with the components of the cytoskeleton and signaling systems. Different integrin pairs have different binding affinities for matrix and intracellular molecules. The binding of an integrin can lead to a specific signal within the cell (outside-in signaling) that can alter cytoskeletal organization, cellular behavior, and gene expression. Similarly, changes within the cell (inside-out signaling) can alter the binding of the integrin to matrix components. (*Part a courtesy of T. J. Kunicki. Part b adapted with permission from R. O. Hynes and Q. Zhao, The evolution of cell adhesion, J. Cell Biol., 150:F89–F96, 2000. © The Rockefeller University Press.*)

amers on one cell can interact with connexin hexamers on the other. Together they form a transcellular protein, a connexon. Before they interact, the connexin hexamers are closed; but once they bind to one another, they form a pore that allows the direct diffusion of small molecules (less than ~1000 daltons) between cells. This type of junction is used by some neurons to pass electrical information directly from cell to cell, and by cardiac muscle cells to coordinate contractile activity. At an electrical synapse, the depolarization of a presynaptic cell can directly depolarize, and thereby activate or otherwise alter, the postsynaptic cell.

Chemical synapses. At chemical synapses (Fig. 4*d*), there are no gap junctions. Instead, the pre- and postsynaptic cells are tightly adherent, but there is a clear separation between the two cell membranes. The signal between the cell involves the secretion of neurotransmitter from the presynaptic cell and its reception by proteins located in the membrane of the postsynaptic cell. The presynaptic cell can be identified by the concentration of vesicles, which contain neurotransmitter.

Cell–extracellular matrix interactions: integrins. Integrins are heterodimeric membrane proteins composed of an α-integrin and a β-integrin polypeptides (**Fig. 5***a*). Integrins mediate interactions primarily with extracellular molecules, although cell–cell adhesive functions for these proteins have also been reported. Different integrin subunit pairs have different patterns of binding specificity. The cytoplasmic domains of integrins often interact with the cytoskeleton. For example, the α_6,β_4 integrin is associated with hemidesmosomes. Disruption of these junctions is associated with the autoimmune skin disease pemphigus. Most other integrins interact with the actin cytoskeleton.

Outside-in and inside-out signaling. Cell adhesion systems can be regulated in two general ways. The binding of the extracellular domain of an adhesion molecule to a specific target molecule (ligand) can alter the structure of the cytoplasmic domain, leading to changes within the adherent cell. This is a specific form of allosteric regulation known as outside-in signaling (Fig. 5*b*). It enables cells to "know" with which types of cells and extracellular molecules they are associated. Conversely, changes within the adherent cell can lead to changes in the extracellular domain of the adhesion protein, such that the molecule no longer binds effectively to its ligands. This inside-out signaling enables cells to disconnect from their neighbors and plays a critical role in the remodeling of tissues during embryonic development and during cancer progression.

Juxtacrine signaling. There are many ways that cells can communicate with one another. By adhering to one another, signals can be directed to only a signal target cell; this type of signaling is known as juxtacrine signaling. It is most obvious in the synaptic connections that link neurons into networks capable of processing complex inputs. Juxtacrine signaling is, however, widespread within metazoans.

Notch signaling. A phylogenetically conserved juxtacrine signaling system is the Notch system (**Fig. 6***a*). Notch proteins, located on the cell surface, can bind to Notch ligand proteins located on the surface of adjacent cells—something that can occur only if the two cells are very close to one another (that is, juxtaposed). The binding of Notch to its ligand activates a protein complex containing the presenilin and γ-secretase. (Mutations in presenilin have been linked to a specific form of early-onset Alzheimer's disease.) The activated presenilin/secretase complex makes a specific cut in the

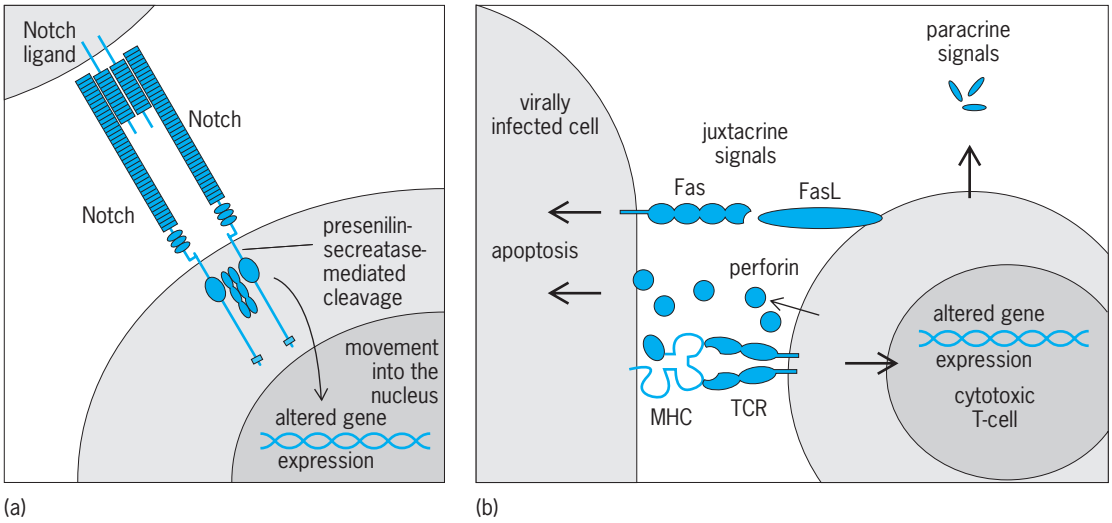

Fig. 6. Juxtacrine signaling. As two cells come close to one another, there is an opportunity for interactions between their surface components. This type of juxtacrine signaling plays critical roles during development and in the adult. (*a*) In Notch signaling, the binding of Notch to its ligand on a adjacent cell activates the proteolytic cleavage of the Notch polypeptide. The released cytoplasmic domain then moves into the nucleus where it alters gene expression and, therefore, cellular behavior and fate. (*b*) An example of immune system juxtacrine signaling. A virally infected cells displays virally derived peptides on its surface, complexed to a major histocompatibility complex (MHC) protein. A T-cell with the appropriate T-cell receptor (TCR) on its surface will recognize the infected cell, and this will initiate changes within the T-cell leading to the release of cytotoxic molecules (perforins in this example) that will lead to the apoptotic death of the infected cell. A second juxtacrine signaling event is facilitated by the MHC–TCR interaction; the surface molecule Fas binds the Fas ligand (Fas L), which induces apoptosis in the virally infected cell.

Notch protein, releasing the protein's cytoplasmic domain, which then moves into the cell nucleus and alters the expression of target genes. *See* ALZHEIMER'S DISEASE.

T-cell receptors/MHC. Juxtacrine signaling is particularly important within the immune system, in which cell surface receptors are used to recognize specific cell types and thereby target signals to these cells (Fig. 6*b*). For example, cytotoxic T-cells recognize virally infected cells through an adhesive interaction between their surface T-cell receptors (TCRs) and major histocompatibility complex (MHC) protein molecules on the infected cell surface. This interaction produces outside-in signaling in the T-cell that leads it to secrete perforins; these proteins generate pores in the target cell and lead to its death. The tight apposition of T-cell and target cell also enables the Fas protein on the target cell to bind to its ligand, FasL, on the T-cell. This interaction triggers an apoptotic (programmed cell death) reaction in the target cell. *See* APOPTOSIS; CYTOLYSIS; HISTOCOMPATIBILITY.

Morphogenesis and the regulation of adhesion. Dramatic changes in cell–cell interactions are associated with developmental events such as gastrulation, which leads to the formation of the three germ layers (ectoderm, mesoderm, and endoderm); neurulation, by which the neural plate folds to form the neural tube, which goes on to become the brain and spinal cord; neural crest formation, by which cells separate from neuroepithelium and migrate away to eventually form much of the craniofacial skeleton and the peripheral nervous system; and organogenesis, by which cells interact to form the various organs, such as the liver, kidney, and heart. In each case, cells must make and break appropriate adhesive interactions.

The signals that instruct cells to maintain or remodel their adhesive properties are complex and depend upon gene and protein activity. Both are influenced by surrounding cells and molecules. The interplay of inside-out and outside-in signaling, together with changes in gene expression, can maintain a cell in stable association with its neighbors or lead it to disconnect. *See* DEVELOPMENTAL BIOLOGY; DEVELOPMENTAL GENETICS; EMBRYOGENESIS; EMBRYONIC INDUCTION.

Study methods. There are several methods used to study cellular adhesion.

Reassociation of dissociated cells. A common approach is to examine the reassociation of dissociated cells (**Fig. 7**). The normal structure of many cell adhesion proteins depends upon the presence of Ca^{2+} ions; removing Ca^{2+} through the use of chelators can lead to the dissociation of tissues into isolated cells. More often, the links between cells are more stable, and the proteins responsible must be digested to release individual cells. This is done using proteases, protein catalysts (enzymes) that degrade proteins. With proteases, it is generally necessary to allow the cells to recover. During this recovery period, the cell replaces membrane proteins damaged by the protease treatment.

To allow reassociation, a suspension of cells is shaken so that only cells that strongly adhere are able to establish stable junctions to one another. This leads to the formation of clumps of cells. The speed at which these cellular aggregates form is a measure of the strength of the adhesion between cells. If cells of different types are present in the suspension, they will form either separate aggregates or aggregates in

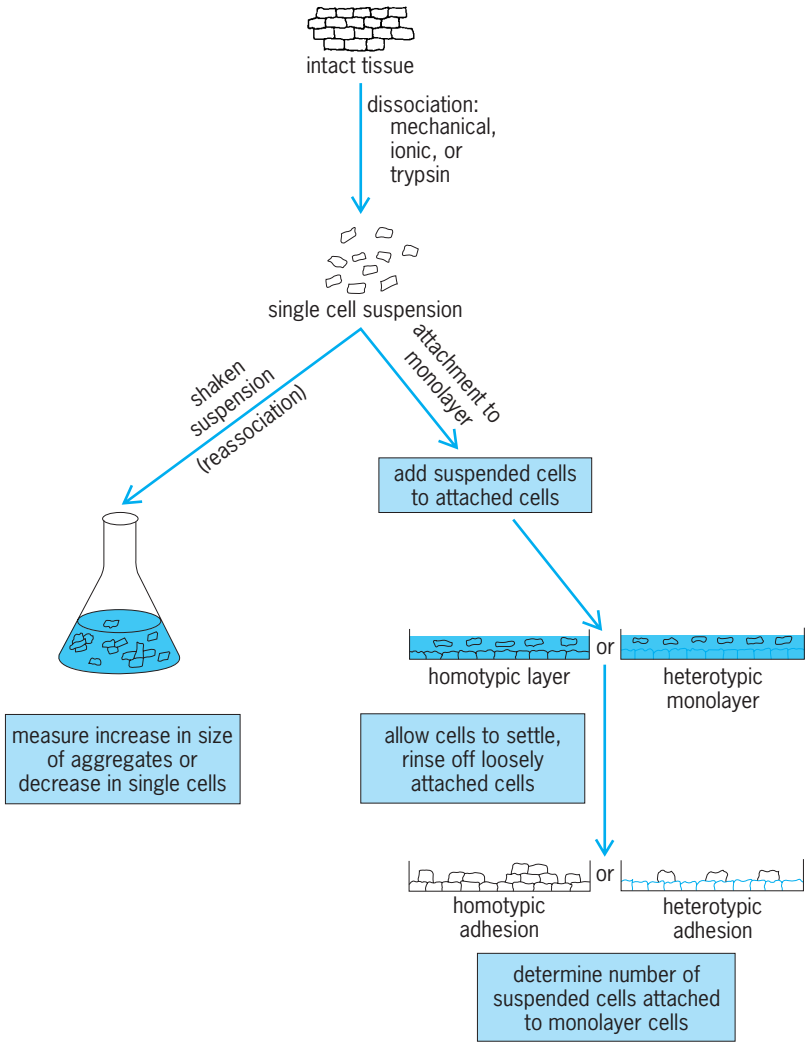

Fig. 7. Methods used to study reassociation of dissociated cells.

which the different cell types are segregated from one another within a single aggregate.

A related technique by which to study cell adhesion involves a preestablished layer of cells, a monolayer, grown on a tissue culture dish. Dissociated cells are added to the culture plate, and after various lengths of time the surface is washed to remove loosely adherent cells. The number of cells that remain attached to the monolayer provides a measure of adherence.

Antibody generation. Another method to study the role of a particular adhesion protein in a specific process is to generate antibodies directed against the extracellular domain of the protein. Some of these antibodies, when applied to a cell, will bind to the protein and interfere with interactions required for adhesion. Most antibody molecules are multivalent, which enables them to mediate adhesive interactions. It is, however, simple to cleave antibody molecules using specific proteases and so produce monovalent antibody fragments. Known as Fab fragments, these molecules retain their binding specificity but do not themselves promote adhesion. If a Fab fragment that binds to a specific cell adhesion molecule blocks cell–cell aggregation, this is taken as direct evidence for the involvement of the target molecule in the adhesion process.

Genetically engineered mutant proteins. More recently, the use of genetically engineered mutant versions of various cell adhesion proteins have been used to study cell–cell and cell–matrix interactions. For example, the DNA sequence encoding a classical cadherin can be manipulated so that a new, modified version is constructed that encodes the extracellular domain and a membrane anchor, but not the cytoplasmic domain. Since the formation of a normal adhesion junction depends upon the cytoplasmic domain, this "dominant-negative"-acting polypeptide will disrupt the normal process of adhesion junction formation. *See* ANTIBODY; GENETIC ENGINEERING; MONOCLONAL ANTIBODIES; MUTATION.

Michael W. Klymkowsky

Bibliography. B. Alberts et al., *Molecular Biology of the Cell*, 4th ed., 2003; N. King, C. T. Hittinger, and S. B. Carroll, Evolution of key cell signaling and adhension protein families predates animal origins, *Science*, 301:361–363, 2003; J. C. McKenzie and R. M. Klein, *Basic Concepts in Cell Biology and Histology*, McGraw-Hill, 2000.

Cell biology

The study of the activities, functions, properties, and structures of cells. Cells were discovered in the middle of the seventeenth century after the microscope was invented. In the following two centuries, with steadily improved microscopes, cells were studied in a wide variety of plants, animals, and microorganisms, leading to the discovery of the cell nucleus and several other major cell parts. By the 1830s biologists recognized that all organisms are composed of cells, a realization that is now known as the Cell Doctrine. The Cell Doctrine constitutes the first major tenet upon which the contemporary science of cell biology is founded. By the late 1800s biologists had established that cells do not arise de novo, but come only by cell division, that is, division of a preexisting cell into two daughter cells. This is the second major tenet upon which the modern study of cells is based. *See* CELL DIVISION; MICROSCOPE.

By the end of the nineteenth century chromosomes had been discovered, and biologists had described mitosis—the distribution at cell division of chromosomes to daughter cells. Subsequent studies showed that the chromosomes contain genes and that mitosis distributes a copy of every chromosome and hence every gene to each daughter cell during cell division. This established the basis of cell heredity and ultimately the basis of heredity in multicellular organisms. *See* CHROMOSOME; MITOSIS.

Microscope studies established that some kinds of organisms are composed of a single cell and some, such as plants and animals, are made up of many cells—usually many billions. Unicellular organisms

are the bacteria, protozoa, some fungi, and some algae. All other organisms are multicellular. An adult human, for example, consists of about 200 cell types that collectively amount to more than 10^{14} cells. Microscope studies revealed that all organisms fall into two major groups, the prokaryotes and the eukaryotes. In prokaryotes, the genetic material is not segregated by a membranous envelope into a distinct nucleus. Prokaryotes are the bacteria, of which there are many thousands of species. Eukaryotes are organisms whose cells possess a nucleus that is clearly defined by an enveloping double membrane. Eukaryotes are the protozoa, fungi, algae, plants, and animals; some are unicellular and some are multicellular. The number of species is unknown, but exceeds 10 million. *See* EUKARYOTAE; PROKARYOTAE.

All modern research recognizes that in both unicellular and multicellular organisms the cell is the fundamental unit, housing the genetic material and the biochemical organization that account for the existence of life. Many millions of different species of cells, showing tremendous diversity in structure and metabolic capabilities, exist on Earth. Cells as different as a bacterium, an ameba, a plant leaf cell, and a human liver cell appear to be so unrelated in structure and life-style that they might seem to have little in common; however, the study of cells has shown that the similarities among these diverse cell types are more profound than the differences. These studies have established a modern set of tenets that bring unity to the study of many diverse cell types. These tenets are: (1) All cells store information in genes made of deoxyribonucleic acid (DNA). (2) The genetic code used in the genes is the same in all species of cells. (3) All cells decode the genes in their DNA by a ribonucleic acid (RNA) system that translates genetic information into proteins. (4) All cells synthesize proteins by using a structure called the ribosome. (5) Proteins govern the activities, functions, and structures in all cells. (6) All cells need energy to operate; they all use the molecule adenosine triphosphate (ATP) as the currency for transfer of energy from energy sources to energy needs. (7) All cells are enclosed by a plasma membrane composed of lipid and protein molecules. *See* CELL MEMBRANES; GENETICS; RIBOSOMES.

In the twentieth century the study of cells, which had been dominated for more than 200 years by microscopy, has been enormously expanded with many other experimental methods. The breaking open of a large mass of cells and the separation of released cell parts into pure fractions led to the discovery of functions contributed by different structures and organelles. For example, mitochondria were shown to be organelles that carry out energy metabolism and supply the entire cell with the energy-rich molecule adenosine triphosphate, and the thousands of small RNA-protein particles in the cell cytoplasm, called ribosomes, were discovered to be factories that carry out protein synthesis. This work was greatly accelerated after World War II by the availability of radioactive isotopes, which were used to identify and trace biochemical activities of various cell parts. *See* MITOCHONDRIA.

Beginning in the 1950s, the application of the electron microscope, with its great power of magnification, led to discoveries about cell structures, especially the endoplasmic reticulum, mitochrondria and chloroplasts, the nuclear envelope, chromosomes, nucleoli, lysosomes, the Golgi complex of membranes, peroxisomes, centrioles, cilia, flagella, the cell membrane, the cytoskeleton, and cell walls. *See* CELL ORGANIZATION; ELECTRON MICROSCOPE.

Contemporary research in cell biology is concerned with many problems of cell operation and behavior. Cell reproduction is of special concern because it is essential for the survival of all unicellular and multicellular forms of life. Cell reproduction is the means by which a single cell, the fertilized egg, can give rise to the trillions of cells in an adult multicellular organism. Disrupted control of cell reproduction, resulting in accumulation of disorganized masses of functionally useless cells, is the essence of cancer. Indeed, all diseases ultimately result from the death or misfunctioning of one or another group of cells in a plant or animal. The study of cells pervades all areas of medical research and medical treatment. Great advances were made during the 1980s in learning how cells of the immune system combat infection, and the nature of their failure to resist the acquired immune deficiency syndrome (AIDS) virus. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); CANCER (MEDICINE); CELL SENESCENCE.

The development of methods to grow plant and animal cells in culture has provided new ways to study cells free of the experimental complications encountered with intact plants and animals. Cell culture has greatly facilitated analysis of abnormal cells, including transformation of normal cells into cancer cells. Cultured cells are also used extensively to study cell differentiation, cell aging, cell movement, and many other cell functions. *See* TISSUE CULTURE.

In parallel with cell biology, the fields of genetics, biochemistry, and molecular biology have likewise greatly expanded. As knowledge and understanding have increased, these fields have overlapped and fused more and more with one another, resulting in a unified multidisciplinary approach to the study of life. This confluence will increasingly facilitate research on the cellular basis of life and provide expanded opportunities for the study of many areas in medicine, genetic engineering, agriculture, evolution, and higher functions of the vertebrate brain. *See* BIOCHEMISTRY; CELL (BIOLOGY); GENETIC ENGINEERING; MOLECULAR BIOLOGY.

David M. Prescott

Bibliography. B. Alberts et al., *Molecular Biology of the Cell*, 3d ed., 1994; W. M. Becker, *The World of the Cell*, 4th ed., 2000; J. Darnell et al., *Molecular Cell Biology,* 1986; D. Fawcett, *The Cell*, 2d ed., 1981; L. J. Kleinsmith and V. M. Kish, *Principles of Cell Biology*, 1988; D. M. Prescott, *Cells*: *Principles of Molecular Structure and Function*, 1988.

Cell constancy

The condition in which the entire body of an adult animal or plant consists of a fixed number of cells that is the same in all members of the species. This phenomenon is also called eutely. The largest group of animals exhibiting eutely are the nematode worms, one of the largest of all animal phyla, and of great medical and agricultural importance as parasites of plants, animals, and humans. A plant that exhibits eutely is usually called a coenobium. Many species of semimicroscopic aquatic green algae exist as coenobia, such as the common *Volvox* and *Pandorina*.

Numerical limitation occurs in certain organs and organ systems, notably the brain and muscles of annelid worms, mollusks, and vertebrates. This is essentially a localized eutely. A related but different phenomenon, observed for many animal cells when cultured, is that normal cells divide some specific number of times and then stop dividing. Thus the life-span, as measured by number of cell cycles, is limited; for many human cell types this is about 50 cell generations.

Investigations on cell constancy and limitation center on three areas. One is the investigation of the factors that limit cell proliferation; this is related to the general problems of growth and differentiation as well as to cancer. A second active area of interest is the use of animals such as mollusks and nematodes in neurophysiological research because they possess nerve centers with a fixed and very small number of cells. Analysis of these simple nervous systems, it is anticipated, will permit better understanding of the organization and functioning of nervous systems, even those of much greater complexity. The third area is muscle development in meat animals such as chickens, beef cattle, and swine. The work on muscle in cattle was stimulated by the discovery of "double-muscled" cattle in which many but not all of the muscles are abnormally enlarged. The muscle hypertrophy in these cattle is due to a single, semidominant gene which has several additional effects, including reduced head size and amount of fatty tissue.

Limiting factors. The factors that sharply limit cell proliferation are unknown. It is clear, however, that in both plants and animals a fundamental distinction exists between growth by cell proliferation and growth by cell enlargement. It is significant in this connection that heavy metals, which will form mercaptides with the sulfhydryl groups of the cell, will block cell division but not cell enlargement. It is also true that the formation of a coenobium in a plant bears a certain resemblance to the cleavage of an egg in an animal. In both cases, a single large cell undergoes a rapid series of mitoses, resulting in a group of many small cells. It is conceivable that in both cases the final slowdown of cell division is due to an exhaustion of the readily available supply of nucleic acids and other constituents required for the rapid formation of new chromosomes.

Although plausible, such a theory gives little insight into why in one case the cells can continue to divide at a slow pace while in the other they must wait until the initial large size is again reached. The work of G. Fankhauser on salamanders has shown that unknown factors correlated with overall size of the animal are somehow important. If these animals are made polyploid so that they have an extra set of chromosomes in every cell, the individual cells are much larger than normal. Nevertheless, the size of the adults remains close to normal. This is because in every tissue, from the brain to the mucous glands of the skin, there has occurred a reduction in cells to compensate exactly for their large size.

The cause or causes of the limit on the number of possible cell divisions in culture for vertebrate cells are unknown. One current theory is that the cumulative effects of thermal, chemical, and background radiation on the cell's deoxyribonucleic acid (DNA) result in a cell which cannot divide. However, why this always occurs after some specific number of cell divisions is hard to explain by this random damage theory. An alternative hypothesis is that cell senescence and the inability to divide are genetically programmed developmental events.

Animals. In rotifers, where cell constancy was first intensively investigated, the number of nuclei in every organ and tissue can be counted exactly. In the common species *Epiphanes* (*Hydatina*) *senta*, which has been studied by several workers, there are 958 nuclei in every individual. The skin and associated structures have 301, the pharynx 167, and the digestive tract 76, of which 15 constitute the esophagus, 35 the stomach, 12 the gastric glands, and 14 the intestine. The urogenital system consists of 43 cells, the general musculature 120, and the nervous system 247, of which 183 are in the brain and 4 in the retrocerebral organ. A. F. Shull, in an investigation of a large number of individuals of this species, found almost no deviations. The presence of 6 nuclei in each gastric gland and of 8 in the vitellarium is common to many species of rotifers.

"An outstanding anatomical feature of nematodes," according to L. H. Hyman, "is their cell constancy." This fact has been confirmed by several workers on various species, including the well-known *Ascaris, Turbatrix aceti* (vinegar eel), and *Rhabditis*. In *Turbatrix* there are 251 nerve cells, 1 excretory cell, 18 cells in the midgut, 64 in the body-wall musculature, 59 in the pharynx, 5 in the esophagus, and so on. Since cell division comes to an end before the birth of the worm, all subsequent growth is by cell enlargement. The gonads are exceptions because they produce new eggs or sperms continuously throughout the reproductive life of the individual.

In annelids and vertebrates, cell proliferation is more or less continuous throughout life only in those tissues that are subject to wear. Thus, in adults, cell division may be found in the germinative zones of the skin, hair, finger and toe nails, the lining of the alimentary canal, and especially in the blood cell–forming tissues. The muscles and nervous system, however, appear to undergo no cell division after early embryonic or fetal stages. In both earthworms

and mammals, including humans, it has been demonstrated that the number of muscle nuclei and muscle fibers, but not fibrils, is fixed early and does not increase with subsequent growth. An earthworm hatches from its egg cocoon with the adult number of muscle fibers and nuclei. A human fetus, aboutbreak 5 in. (13 cm) from crown to rump, has as many muscle fibers and nuclei as an adult. It has been shown that the number of glomeruli in each kidney of a rat or human, and therefore presumably of any mammal, is fixed before birth, and that the subsequent growth of the glomeruli, either normally or resulting from compensatory hypertrophy after unilateral nephrectomy, is due entirely to the enlargement of cells already present. The same holds true for the cells of the ciliated nephrostomes of earthworms.

Plants. Coenobia, that is, adult individuals composed of a fixed and predictable number of cells, are characteristic of the species of green algae belonging to the species of *Volvox* in the order Volvocales, and to the species of *Hydrodictyon* and *Scenedesmus* in the order Chlorococcales. In these forms the number of cells in an adult ranges from 4 to 8 in certain species of *Gonium*, 32 to 64 in *Pandorina*, and from 500 (probably 512) to 60,000 (probably 65,536) in *Volvox*. During asexual reproduction in these species, either every cell or, in the larger forms, certain cells undergo a rapid series of cell divisions. This results in the formation of a miniature replica of the adult, as far as cell number is concerned, within every single cell which has so divided. Thus a very small, young volvox has as many cells as a large, mature one. The minute, new individuals are released from the parent cell by the breakdown of the old cell wall. It will be noted that the number of cell generations is always small. Thirty-two is 2^5, which means only 5 cell generations, 512 is 2^9, and even 65,536 is only 2^{16}, which means only 16 cell generations to make the largest volvox. Gairdner B. Moment

Bibliography. B. C. Edgar, Nematodes, *"Elegans Workshops" Worm Breeder's Gazette*, vols. 1 and 2, University of California, Santa Cruz, 1977; D. F. Goldspink, *Development and Specialization of Skeletal Muscle*, 1981; L. Hayflick, Cell biology of human aging, *Sci. Amer.*, 242(1):58–66, 1980; L. H. Hyman, *The Invertebrates*, vol. 3, 1951; T. L. Lawrence (ed.), *Growth in Animals*, 1980.

Cell cycle

The succession of events that culminates in the asexual reproduction of a cell; also known as cell division cycle. In a typical cell cycle, the parent cell doubles its volume, mass, and complement of chromosomes, then sorts its doubled contents to opposite sides of the cell, and finally divides in half to yield two genetically identical offspring. Implicit in the term cycle is the idea that division brings the double-sized parent cell back to its original size and chromosome number, and ready to begin another cell cycle. This idea fits well with the behavior of many unicellular organisms, but for multicellular organisms the daughter cells may differ from their parent cell and from each other in terms of size, shape, and differentiation state.

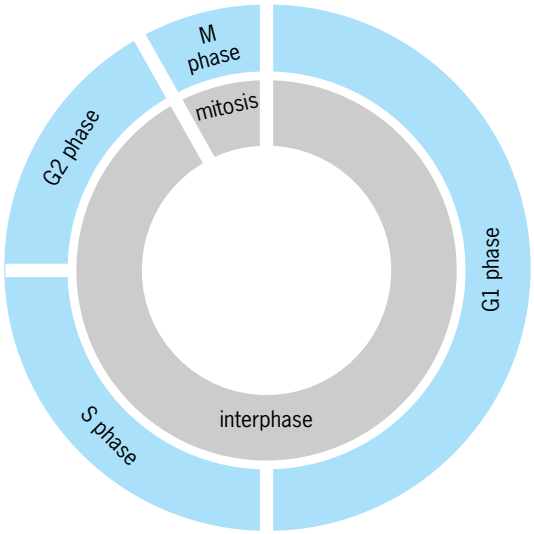

Fig. 1. Phases of the eukaryotic cell cycle.

The time required for completion of a eukaryotic cell cycle varies enormously from cell to cell. Embryonic cells that do not need to grow between divisions can complete a cell cycle in as little as 8 min, whereas cycling times of 10–24 h are typical of the most rapidly dividing somatic cells. Many somatic cells divide much less frequently; liver cells divide about once a year, and mature neurons never divide. Such cells may be thought of as temporarily or permanently withdrawing from the cell cycle.

Eukaryotic phases. The cell cycle is divided into two main parts: interphase and mitosis (**Fig. 1**). During interphase, the cell grows and replicates its chromosomes. Interphase accounts for all but an hour or two of a 24-h cell cycle, and is subdivided into three phases: gap phase 1 (G1), synthesis (S), and gap phase 2 (G2). Interphase is followed by mitosis (nuclear division), and cytokinesis (cell division). This relatively brief part of the cell cycle includes some of the most dramatic events in cell biology. *See* MITOSIS.

G1 phase. Gap phase 1 begins at the completion of mitosis and cytokinesis and lasts until the beginning of S phase. This phase is generally the longest of the four cell cycle phases and is quite variable in length. During this phase, the cell chooses either to replicate its deoxyribonucleic acid (DNA) or to exit the cell cycle and enter a quiescent state (the G0 phase). Late in the G1 phase, the cell becomes committed to replicating its DNA. In mammalian cells, the time at which this commitment occurs is called the restriction point.

S phase. Replication of the chromosomes is restricted to one specific portion of interphase, called S phase (DNA synthesis phase), which typically lasts about 6 h. In mammalian cells, the start of S phase—the actual initiation of DNA synthesis—takes place several hours after the cell has committed to

carrying out DNA synthesis. During S phase, each chromosome replicates exactly once to form a pair of physically linked sister chromatids. In animal cells, a pair of centrioles is also duplicated during S phase. *See* CHROMOSOME; GENETICS.

G2 phase. The portion of interphase that follows S phase is called gap phase 2. Some cells can exit the cell cycle from G2 phase, just as they can from G1 phase.

M phase. M phase includes the overlapping processes of mitosis and cytokinesis. Mitosis is divided into five stages: prophase, prometaphase, metaphase, anaphase, and telophase. During prophase, the chromosomes condense and the football-shaped mitotic spindle begins to form. Prometaphase begins when the nuclear envelope abruptly disappears and the chromosomes begin to migrate toward the spindle's midline. When the chromosomes reach the midline, the cell is said to be in metaphase. Metaphase ends and anaphase begins when the sister chromatids abruptly separate from each other and move toward the spindle poles. During telophase, the nuclear envelope reforms around each set of chromosomes, the chromosomes decondense, and mitosis is completed. Cytokinesis usually begins during anaphase and ends at a point after the completion of mitosis. At the end of cytokinesis, the parent cell has formed its two G1 phase progeny and the cell is ready to repeat the cycle.

Control of cell cycle. The network of proteins that regulate DNA synthesis (G1/S), mitotic entry (G1/M), and mitotic exit (the transition from mitotic metaphase to anaphase and then out of mitosis) appears to be well conserved throughout eukaryotic evolution. At the heart of these cell cycle transitions is the periodic activation and inactivation of cyclin-dependent protein kinases. In addition, in multicellular eukaryotes, pathways regulating entry into and exit from the cell cycle entrain these central cyclin-dependent kinases to extrinsic signals.

Cell cycle entry. In multicellular eukaryotes, most cells spend most of their time in the quiescent G0 state. In response to peptide growth factors, cells initiate a signal transduction cascade that culminates in entry into G1 phase. Components of these mitogenic signaling pathways include receptor tyrosine kinases [such as the epidermal growth factor (EGF) receptor], small G-proteins (such as Ras), and signal-relaying protein kinases (such as MAP kinase). Mitogenic signaling culminates in the activation of at least two waves of gene transcription, leading to the synthesis of a G1 cyclin protein (cyclin D) and neutralization of the retinoblastoma tumor suppressor protein (pRb).

Many components of these mitogenic signaling pathways are capable of causing malignant transformation when overexpressed or inappropiately activated. The cancer-causing forms of these proteins are termed oncoproteins, and their genes are termed oncogenes; the corresponding normal forms from which oncogenes and oncoproteins are derived are termed proto-oncoproteins and proto-oncogenes. Conversely, several tumor suppressor genes (genes that can promote malignant transformation when inactivated) have now been shown to encode proteins that oppose mitogenic signaling proteins (for example, the retinoblastoma protein and the PTEN phosphatase). These discoveries underscore the idea that when the biochemical machinery that controls normal cell growth goes awry, cancer may result. *See* ONCOGENES.

Much less is known about how cells are induced to exit the cell cycle and enter the quiescent G0 state.

Mitotic control. When interphase cells are artificially fused with cells in mitosis, the interphase cells' nuclei rapidly enter mitosis. This discovery indicates that some dominant mitosis-inducing factor is present in M-phase cells. Studies on frog oocytes and eggs, unusually large cells that are naturally arrested in G2 phase (oocytes) or meiotic M phase (eggs), underscore this point: when oocytes are microinjected with cytoplasm from eggs, the oocytes enter M phase, a process termed oocyte maturation. The factor responsible, maturation promoting factor (MPF), is present and active during both meiotic M phase and mitotic M phase, and can be found in M-phase cells from evolutionarily distant organisms. This suggests that MPF is the universal M-phase trigger. The ease of obtaining large quantities of M-phase frog eggs meant that the identification of MPF could be approached biochemically.

Other crucial insights into M-phase control came from studies of growth regulation in the yeast *Schizosaccharomyces pombe*. Mutations in genes that regulate M-phase onset of *S. pombe* give rise to organisms that are too long or too short. Once such a mutant strain is isolated, the gene responsible for its altered size can be identified by molecular genetic methods. Two genes identified through this approach were $cdc2^+$ and $cdc13^+$. Genes related to $cdc2^+$ and $cdc13^+$ were found in a variety of other eukaryotes, suggesting that they encoded universal regulators of M-phase onset. *See* MOLECULAR BIOLOGY; MUTATION.

Ultimately it was realized that the biochemical studies of M-phase regulation in frog oocytes and the genetic studies of M-phase regulation in *S. pombe* had succeeded in identifying the same M-phase regulators. MPF proved to be a complex of the *Xenopus* homologs of *S. pombe* $cdc2^+$ and $cdc13^+$ (the Cdc13 protein is more usually called a B-type cyclin, because it rises and falls in abundance during the cell cycle). The fact that two very different approaches and two evolutionarily distant organisms had converged upon the same M-phase regulators underscored the importance and universality of the regulators.

The Cdc2/cyclin B complex is essential for initiation of all M phases in all organisms. The complex functions as a protein kinase—an enzyme that adds a phosphate group to specific amino acid residues in target proteins. The catalytic subunit of the complex is Cdc2; cyclin B is necessary for the activation of Cdc2 and is responsible for localizing the complex to the nucleus at the onset of M phase. The nuclear lamin proteins are Cdc2 target proteins, which

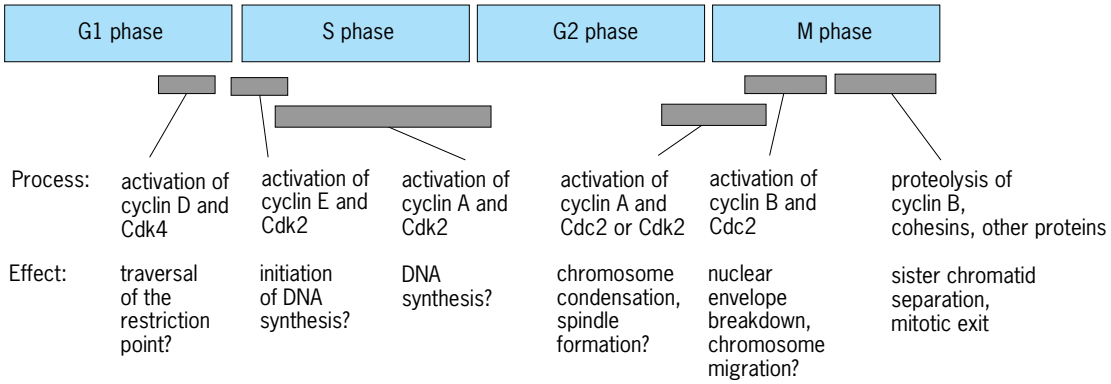

Fig. 2. Behavior of various Cdk/cyclin complexes and their cellular events.

form a scaffolding that supports the nuclear envelope. The Cdc2/cyclin B complex can phosphorylate lamins, causing the lamin network to disassemble. This disassembly allows the nuclear envelope to break down into small vesicles, which will coalesce and reassemble during telophase after cyclin B has been degraded and the lamins have been dephosphorylated. Other likely targets of Cdc2/cyclin B complex include histone proteins, whose phosphorylation may contribute to chromosome condensation, and a number of regulatory proteins.

At the onset of anaphase, a protein complex termed the anaphase-promoting complex (APC) becomes activated, bringing about the proteolytic destruction of several key mitotic regulators. One critical APC target is cyclin B; its destruction is required for mitotic exit. APC also directly or indirectly causes the destruction of other proteins (securins, cohesins) that keep the pairs of sister chromatids attached to each other and keep the cell in metaphase. As was the case with MPF, the rapid progress in the understanding of APC function and the appreciation of its universality were made possible through a combination of genetic studies (in the budding yeast *Saccharomyces cerevisiae*) and biochemical studies (mostly in frog egg extracts).

G1- and S-phase control. The other cell cycle transitions, for example, the commitment to DNA synthesis that occurs in G1 phase, and the initiation of DNA synthesis at the start of S phase, are also triggered by activation of heterodimeric protein kinases consisting of Cdc2-like catalytic subunits (termed Cdks, for cyclin-dependent kinases) and cyclin-like regulatory subunits. At least nine Cdc2-like catalytic subunits (Cdc2/Cdk1 and Cdk2–9) and at least eleven classes of cyclins (cyclins A through J and cyclin T) have now been identified in animal cells. A subset of these proteins have been implicated in cell cycle regulation (Cdc2, Cdk2, Cdk4, Cdk6; cyclins A, B, D, E). Thus, the entire cell cycle may be driven by the sequential activation and inactivation of Cdk/cyclin complexes (**Fig. 2**). James E. Ferrell

Bibliography. A. Murray and T. Hunt, *The Cell Cycle: An Introduction*, 1993; A. W. Murray and M. W. Kirschner, What controls the cell cycle?, *Sci. Amer.*, 264:56–63, 1991.

Cell differentiation

The mechanism by which cells in a multicellular organism become specialized to perform specific functions in a variety of tissues and organs. The life cycle of a higher organism begins with a unicellular stage, the fertilized egg, and becomes more complex as the individual grows and takes on its characteristic form. The life cycle of a unicellular organism permits temporary change in shape and chemical composition. Only a multicellular organism can afford irreversibly differentiated cells.

Differentiated cells. The stable differentiated state is a consequence of multicellularity. A complex organism maintains its characteristic form and identity because populations of specialized cell types remain assembled in a certain pattern. Thus several kinds of cells make up a tissue, and different tissues build organs. The variable assortment of about 200 cell types allows for an almost infinite variety of distinct organisms.

Epithelia, sheets of cells of specific structure and function, cover the outer surface of the vertebrate body and line the lungs, gut, and vascular system. The stable form of a vertebrate is due to its rigid skeleton built from bone and cartilage, forming cells to which the skeletal muscles adhere. All other organs, such as liver and pancreas, are embedded in connective tissue that is derived from fibroblast cells which secrete large amounts of soft matrix material. *See* CONNECTIVE TISSUE; EPITHELIUM.

Some cells, like nerve cells, are so specialized that they need divide no longer in order to maintain a complex network. Their finite number decreases even during embryonic development. Other cell types are constantly worn out and must be replaced; for example, fibroblasts and pancreas cells simply divide as needed, proving that the differentiated state of cells is heritable, as the daughter cells remember and carry out the same special functions. The renewal of terminally differentiated cells that are unable to divide anymore, such as skin and blood cells, is carried out by stem cells. They are immortal and choose, as they double, whether to remain a stem cell or to embark on a path of terminal differentiation. Most stem cells are unipotent because they

give rise to a single differentiated cell type. However, all cell types of the blood are derived from a single blood-forming stem cell, a pluripotent stem cell. A fertilized egg is a totipotent stem cell giving rise to all other cell types that make up an individual organism. *See* BLOOD.

Process of differentiation. Specialized cells are the product of differentiation. The process can be understood only from a historical perspective, and the best place to start is the fertilized egg. Different kinds of cell behavior can be observed during embryogenesis: cells double, change in shape, and attach at and migrate to various sites within the embryo without any obvious signs of differentiation. Cleavage is a rapid series of cell cycles during which the large egg cell is divided into a ball of small cells that line the primitive body cavity as a single layer of embryonic cells. This blastula stage is followed by gastrulation, a complex coordinated cellular migration which not only shapes the embryo but segregates the single-cell layer of the blastula into the three germ layers: endoderm, mesoderm, ectoderm. They give rise to specific cell types; for example, skin and nerves from the ectoderm, the digestive tract from the endoderm, and muscle and connective tissue from the mesoderm. *See* BLASTULATION; CLEAVAGE (DEVELOPMENTAL BIOLOGY); EMBRYOGENESIS; GASTRULATION.

Descriptive embryology has contributed elaborate body plans, but cannot explain the development of various organisms. Experimental embryologists originally concluded that all organisms can be placed in two categories. They found that either the developmental system represented by an egg acts as a fixed mosaic, or it can regulate and, for example, give rise to four complete mice from the cells produced after the first cleavage divisions. In a philosophical context, the terms preformation and epigenesis describe the two opposite kinds of development, the former claiming that an organism is physically preformed inside the egg, the latter proposing that the organism is newly created in each generation. *See* DEVELOPMENTAL BIOLOGY; EMBRYOLOGY.

Cell determination. An explanation of the two egg types by the timing of a single developmental change came from another experiment. A piece from a salamander blastula, which would have become an eye according to the body plan, was transplanted into the region destined to develop into the gut of another blastula. The implant adapted to the new environment and became part of the gut rather than an eye. A short time later, after gastrulation, the transplant resulted in the development of an eye inside the gut. This experiment holds a key to the cell differentiation process. It reveals a critical event, cell determination, which leads to the fixation of a particular developmental pathway. Before that time the embryo regulates; thereafter it has become a mosaic. Once a cell is committed to function in a specific capacity, it will not forget it, irrespective of environment. Thus, the change brought about by determination affects the internal character of a cell and results in a self-propagating change. The egg cell combines all capacities to build organisms as complex as humans, but it can do so only if its potential is first segregated and divided into groups of stem cells with lesser developmental capacities. Two counteracting tendencies describe the developmental process: as embryonic cells lose their developmental potential, they gain specific characteristics which permit sophisticated division of labor in multicellular organisms. *See* EMBRYONIC DIFFERENTIATION.

Cell differentiation is the realization of the determined state, and since different cell types share the same genes, the central concept, if not dogma, of developmental biology is the theory of differential gene expression. This means that a myoblast becomes a muscle cell because genes responsible for muscle-specific structures, such as the contractile apparatus, are selectively activated. Conversely, muscle-specific genes remain inactive in other cell types. Molecular biologists have accumulated good evidence for selective gene expression in differentiated cells. Specific gene products, such as hemoglobin in red blood cells, have been labeled luxury molecules, because they are not essential for the lives of a cell but are required for the diversification of multicellular organisms. However, the processes of cell determination and differentiation are not accompanied by an overall increase in the number of genes that are expressed during embryonic development. To the contrary, many genes that are active in early embryos become inactivated by the time a cell becomes differentiated. It seems that the loss of developmental capacity is accompanied, or may be the consequence of, massive inactivation of genes of unknown functions. *See* GENE.

The determined state is very stable and inherited over many cell generations, but it is not irreversible. Transdetermination occurs occasionally and alters the developmental course of cells. There are some mutations that cause a similar shift of the developmental pathway. Perhaps the respective genes are "master genes" which control a whole set of cell-type-specific genes that evoke the production of the "luxury" molecules in differentiated cells. About 200 of such regulatory genes would be required to determine all existing cell types. If, however, embryonic cells could be determined by a set of regulatory genes in either "on" or "off" position, eight genes would suffice to program all different cell types ($2^8 = 256$). It can be concluded that these control genes are active in determined cells, which makes them different from those that are not determined. Consequently, committed cells must possess some less obvious luxury molecules yet to be discovered. They are distinct from undetermined cells, although not so grossly altered as the differentiated cells that are their progeny. Cell divisions occur during the transition from cell determination to differentiation. Frequently, a special kind of cell cycle, called the quantal cell cycle, is required, following which the differentiated state becomes obvious by the massive production of the cell-type-specific luxury molecules. *See* CELL CYCLE.

Embryonic induction. Another experiment has shed some light on the determination process. Implantation of that piece from the blastula, which would

normally come to lie under the prospective neural tissue after gastrulation, into another blastula turns the host embryo into a monster with two complete nervous systems. This experiment defines embryonic induction. Induction works even with killed inducing cells. This result opened the search for the still elusive inducer substance. Furthermore, the capacity of induction can be traced back through earlier stages into the egg. These observations are consistent with the concept of cytoplasmic localization. It is assumed that cytoplasmic determinants are produced during the process of oogenesis and become deposited in certain areas in the egg. This concept is a modern version of the classical preformation theory. *See* OOGENESIS.

There are two ways to explain embryonic induction: either the inducer instructs the reacting tissue to become determined and, later on, differentiated, or the reacting tissue is multipotential and selects one developmental pathway over one of several others. In the first case, instruction can be viewed as a transfer of information in the form of specific protein or nucleic acid molecules. In the second case, the inducer plays a more permissive role, allowing the reacting tissue to decide what to do next. Consequently, the inducer need not confer information and may just provide a cue, like a change in ion concentration. There is evidence for unspecific cueing in the early embryogenesis of the most complex organisms, the mammals. After a few cleavage divisions, the young embryo forms the blastocyst, which consists of an outer layer of cells with a few cells inside. Any cell that happens to be inside the blastocyst takes part in building the embryo, while outside cells do not. This is germane to the inside-out concept, indicating that the cellular environment can provide cues as to the course of cell determination in the absence of localized cytoplasmic determinants. *See* EMBRYONIC INDUCTION.

Cellular nonequivalence. The shaping of an embryo is an essential routine of the developmental program. While embryonic induction leads to cell determination and differentiation, it is a complex patterning process that eventually puts different cell types into their proper territory. For example, a leg and an arm are quite distinct, although they are built from the very same cell types; but they are arranged in different patterns. More than one cell must collaborate to construct body parts, such as a segment or a wing of an insect. This process is initiated among a few founder cells in response to unknown cytoplasmic controls for which terms such as morphogenetic field, gradient, and centers have been applied.

According to the unifying position information hypothesis, embryonic cells secrete substances, perhaps an activator and an inhibitor, that generate a stable dynamic gradient of either diffusible molecules or a macromolecular extracellular network. These cells not only receive information with respect to their position, but are able to interpret that information. Interpretation is the crucial component of this hypothesis. One can assume that receptor molecules at the cell surface measure the size of the supracellular position information signal, and can postulate that they interpret different increments of that signal according to a preset behavior pattern. In the famous "French flag model," a sheet of cells with the potential to become blue, white, or red turn red at a low value, blue at a high value, and white at a medium value of the position signal. An important implication of the position information concept is that cells remember their position value. Consequently, cells of the same differentiated state can be different: they are nonequivalent. As with determination and differentiation, cellular nonequivalence is a self-perpetuating change in character. It may distinguish each cell from all other cells and allow it to home-in during the animated puzzle of embryogenesis and to stay put in a precise location. There is a class of developmental mutants, in which cells of a certain territory, called a compartment, collectively change their course of development and, for example, form a leg instead of an antenna at the head of the fruit fly *Drosophila melanogaster*. A combinational activity pattern of a few selector genes could provide for a mechanism to interpret different positive values.

It is conceivable that a single position information mechanism operates in all organisms during embryogenesis and later permits determined cells to display overt differentiation at the correct place. Together with genetic information that generates biological materials, position information could arrange these materials inside a cell to maintain a specific subcellular organization, or could operate between cells to create multicellular organisms of distinct forms.

Cellular change. In order to accurately describe a specific organism, one must know which cell types are arranged in what pattern. From a developmental perspective three different components describe the system from which a body plan emerges: cellular determination, differentiation, and nonequivalence. The developmental process proceeds according to the timing and interactions of these three kinds of cellular self-perpetuating change. Therefore, it is impossible to project a body plan into an egg cell and thus reveal the blueprint of a particular organism. The remarkable biological change initiated upon fertilization of an egg is perhaps better described as the product of the developmental program.

Molecular biology has unraveled in almost complete detail how genetic information in nucleic acids programs the linear array of amino acids in proteins. Some model systems (for example, bacteriophages and subcellular particles, such as ribosomes) need little if any additional information to assemble predictably into highly complex and functional structures by self-organization. These particles are almost exclusively programmed in the genome, but they are not alive. Even dissociated specialized cells of an adult organism can self-organize to some extent by the processes of sorting-out and differential adhesion. However, a cell is never reformed from its parts; each cell comes from a cell. Over 90% of its genome does not code for any known product. Even most known genes are made up of a

mosaic of coding stretches of deoxyribonucleic acid (DNA), exons, and introns that are seemingly devoid of genetic sense. The genome is less stable than originally thought. Subtle rearrangements (by transposons) or differential ribonucleic acid (RNA) splicing may be the source for unprogrammed but essential biological change during embryonic development. In addition, the biomembrane, which is never the result of self-assembly, could store information in a two-dimensional network at its inner face to communicate with the cytoplasm, and at its outer face to respond to signals from the environment. To date, the developmental program, if it exists, has remained elusive. *See* BACTERIOPHAGE; DEOXYRIBONUCLEIC ACID (DNA); EXON; INTRON; MOLECULAR BIOLOGY; NUCLEIC ACID; PROTEIN; RIBONUCLEIC ACID (RNA); RIBOSOMES; TRANSPOSONS.

By comparing the differentiation process of cells to evolution, interesting new insights may be gained. The mechanism of cell differentiation and the establishment of nonequivalence are formally equal to speciation. Although it is hard to imagine how random change in embryonic cells will lead to complex organisms, such as humans, the developmental options of the germ cell are limited. The limitations are the linear patterns of genes, which hold the memory for their successful propagation over evolutionary time, and the constraints posed by the protoplasmic environment and its system of chemical interactions. Extinctions are common in the fossil record and, by analogy, also occur during embryogenesis. For example, the big egg cell, the totipotent stem cell, disappears after a few cleavage divisions and must be reorganized during oogenesis in the adult organism; and all neuroblasts, the precursors of terminally differentiated nerve cells, become extinct before birth. Cell death even becomes a formative tool as the human fingers are carved out from a paddlelike appendage, and the tail of a tadpole is destroyed in building the frog. Eventually the whole organism is survived only by its germ cells and the genes contained in them.

Origins. A living unit must reproduce faster than it is destroyed or taken over by other replicators. It seems that one kind of replicator made of nucleic acid and protein survived all other attempts at forming life, because all life on Earth uses the same genetic code. The most primitive cell was formed when a membrane kept the necessary building blocks close to the replicator. Prokaryotic organisms, such as bacteria and blue-green algae, may be descendants of the first cell. The eukaryotic cell achieved a higher degree of division of labor by segregating and enlarging the genome in the nuclear envelope, and by acquiring organelles, perhaps through ingestion and maintaining other primitive cells in a state of symbiosis. Flagellates may resemble that state which permitted a thousandfold increase in size. All unicellular organisms compete with each other for resources. The third step, following replicator and cell formation, is characterized by specialization of all functions assembled in the eukaryotic cell through cellular differentiation and positioning. Thus multicellularity and cell differentiation, and the switch from competitive to altruistic cellular behavior, are closely related. *See* GENETIC CODE; PREBIOTIC ORGANIC SYNTHESIS.

Some specialized functions in extant organisms may provide information as to how cell differentiation originated. Even the simplest bacterium segregates genetic material into one gene set which controls the routine processes of life, and a small additional genome, a plasmid, that permits shuttling of genetic material within and between cells, a prerequisite for sexuality. *See* BACTERIAL GENETICS.

The ciliates are the most advanced single-celled organisms. They have two nuclei, a micronucleus and a macronucleus. The micronucleus functions in gene replication and sexual recombination; its genes cannot be copied into RNA. Growth and division are controlled by the macronucleus, in which a portion of the genome is segregated, amplified, and transcribed. Ciliates display a highly organized cell surface as the result of intracellular patterning. The positioning of arrays of cilia is to some degree self-perpetuating and independent of the genome. *See* CILIOPHORA.

The switch from competitive to altruistic cell behavior happens in the life cycle of cellular slime molds. They live as individual cells while they have sufficient food. Following starvation, about 10,000 individuals aggregate into a tissuelike structure which forms a fruiting body. In this process many individuals die and allow the survival of others to propagate the genes. Formation of these fruiting bodies is a model case of cellular positioning, reversible differentiation, and altruistic cellular behavior. *See* CELL ADHESION.

A stable association of identical cells has occurred in some algae, which form a sphere from identical individuals that resemble a blastula. Plants may have arisen from algal colonies, and the high degree of regeneration in plants could indicate that cells of a plant tissue have never given up the pluripotent state. *See* ALGAE.

The plasmodial slime molds provide a model of developmental organization beyond unicellularity. These organisms can live as minute competing amebas. Under proper conditions a single cell turns into a large mass of protoplasm. Its new developmental feature is the uncoupling of nuclear divisions from cell divisions. A plasmodium is born, a giant cell which can grow to over a foot in diameter and contains billions of nuclei. In some plasmodial slime molds, cellular transformation into a multinucleated system is coupled to sexual cell differentiation and cell fusion. The life cycle continues with a patterning process in which the giant cell is divided into smaller units of about a million nuclei. Each unit undergoes morphogenesis which results in a fruiting body. During this process most nuclei become eliminated. The surviving nuclei end up in the head of the fruiting body, where each is surrounded by a portion of cytoplasm and a proper cell membrane. Thus, this organism demonstrates major events of embryogenesis, which turn a fertilized egg into a complex organism: growth, patterning, determination, cell differentiation, and morphogenesis. It is a simple model system

for two reasons: the transition from the unicellular to the multicellular state occurs in the absence of the still elusive maternal programming during oogenesis, and the giant cell and the building of the fruiting body display, in mutually exclusive fashion, either growth or differentiation. One cannot distinguish whether the plasmodial slime mold represents an evolutionary dead end or whether the animals are its descendants. Once multicellularity is established, all kinds of fungi, plants, and animals can evolve by utilizing the opportunities of specialization to optimize division of labor by differentiated cells, and the constraints posed by the history of their ancestors. *See* CELL (BIOLOGY); CELL SENESCENCE. Helmut Sauer

Bibliography. B. Alberts et al., *Molecular Biology of the Cell*, 3d ed., 1994; J. T. Bonner (ed.), *Evolution and Development* 1982; F. Jacob, *The Logic of Life*, reissued 1982; G. Karp and N. J. Berrill, *Development*, 2d ed. 1981; N. Maclean, *The Differentiation of Cells*: *Genetics—Principles and Perspectives* (a series of texts), 1977; R. A. Raff and T. C. Kaufman, *Embryos, Genes, and Evolution*, 1983.

Cell division

The division of a cell into daughter cells that receive identical copies of its genetic material. The cell cycle comprises the period between the formation of a cell as a progeny of division and its own subsequent division into two daughter cells. The cell cycle falls into two parts. A relatively long interphase represents the time during which the cell engages in synthetic activities and reproduces its components, even though there is no visible change. The relatively short period of mitosis provides an interlude during which the actual process of visible division into two daughter cells is accomplished.

Cell cycle. Prokaryotic and eukaryotic cells differ markedly in the coordination of deoxyribonucleic acid (DNA) synthesis and in the subsequent equal partitioning of DNA during cell division. In prokaryotes the cell cycle consists of successive periods of DNA replication and of cell division where a cell wall forms and divides the cell in two, with no visible condensation and decondensation of DNA. Eukaryotes from yeast to humans have similar cell division phases, termed G1, S, G2, and M phases. After each division there is a time gap (G1 phase) before the synthesis of DNA begins. The cell is metabolically active during this part of the cycle in which proteins and DNA precursors are made. The period of DNA synthesis (S phase) involves the replication of chromosomal DNA, and is followed by another time gap (G2 phase), after which mitosis (M phase) occurs. Interphase is the time that elapses between one M phase and the next. It is composed of successive G1, S, and G2 phases and normally comprises 90% or more of the total cell cycle time. The cell cycle varies from tissue to tissue and during development. Sea urchin blastomeres, for instance, replicate their DNA at the end of mitosis in telophase of the preceding division, thus eliminating the G1 phase entirely.

The products of the series of mitotic divisions that generate the entire organism are called somatic cells. During embryonic development, most of the somatic cells are proceeding through the cell cycle. In the adult organism, however, many cells are terminally differentiated or no longer divide; they remain in a perpetual interphase. These cells are sometimes said to be quiescent or in G0 phase. *See* CELL CYCLE; DEOXYRIBONUCLEIC ACID (DNA).

Mitotic stages. The mitotic or M phase of the cell cycle involves the separation of replicated nuclei into two identical daughter nuclei. Mitosis is divided into six distinct stages: prophase, prometaphase (or late prophase), metaphase, anaphase, telophase, and cytokinesis (see **illus**.). The length of time that a cell takes to complete mitosis varies in different organisms. In rapidly proliferating cells of higher eukaryotes, M phases generally occur only once every 16–24 h, and each M phase itself lasts only 1–2 h.

Prophase. Prior to mitosis, in the S phase of the cell cycle chromosomal DNA replicates, and each chromosome divides into two sister chromatids. At the onset of prophase, these sister chromatids condense while remaining attached at a specific sequence of DNA necessary for chromosome separation called the centromere. At this point, the condensed chromosomes become visible by light microscopy. In interphase, the centrosome also divides, giving rise to two sets of centrioles. The centrioles, which are made from microtubules, also duplicate during interphase and begin to move to opposite sides of the nucleus to form separate poles. As these centrioles move apart, the microtubules disassemble. The new microtubules or asters, which will become intimately involved in the movement of the chromosomes during mitosis, begin to radiate from the centriole pairs in all directions to form the mitotic spindle. The number of microtubules making up the mitotic spindle varies depending on the organism. The microtubules that connect at the center of the chromatid pairs are called kinetochore microtubules; others, called polar microtubules, will extend to the opposite pole. Astral microtubules radiate from the centriole pair outside the mitotic spindle. *See* CENTRIOLE; CENTROSOME.

Prometaphase. Prometaphase, or late prophase, begins with the breakdown of the nuclear membrane. At this point, the kinetochore microtubules interact with the chromosomes to form protein complexes at their centromeres. This association between the microtubules and proteins of the chromatids makes up the structure of the mitotic spindle. The microtubules that extend past the sister chromatids to each pole are stabilized through cross-linking of microtubule binding proteins. These microtubules are also protected from disassembly by the constant formation of new microtubules from tubulin molecules in the fluid portion of the cytoplasm (cytosol). The stabilization of these microtubules is necessary for mitosis, as shown by the effect of the drug colchicine which blocks the assembly of microtubules and makes cells unable to complete mitosis.

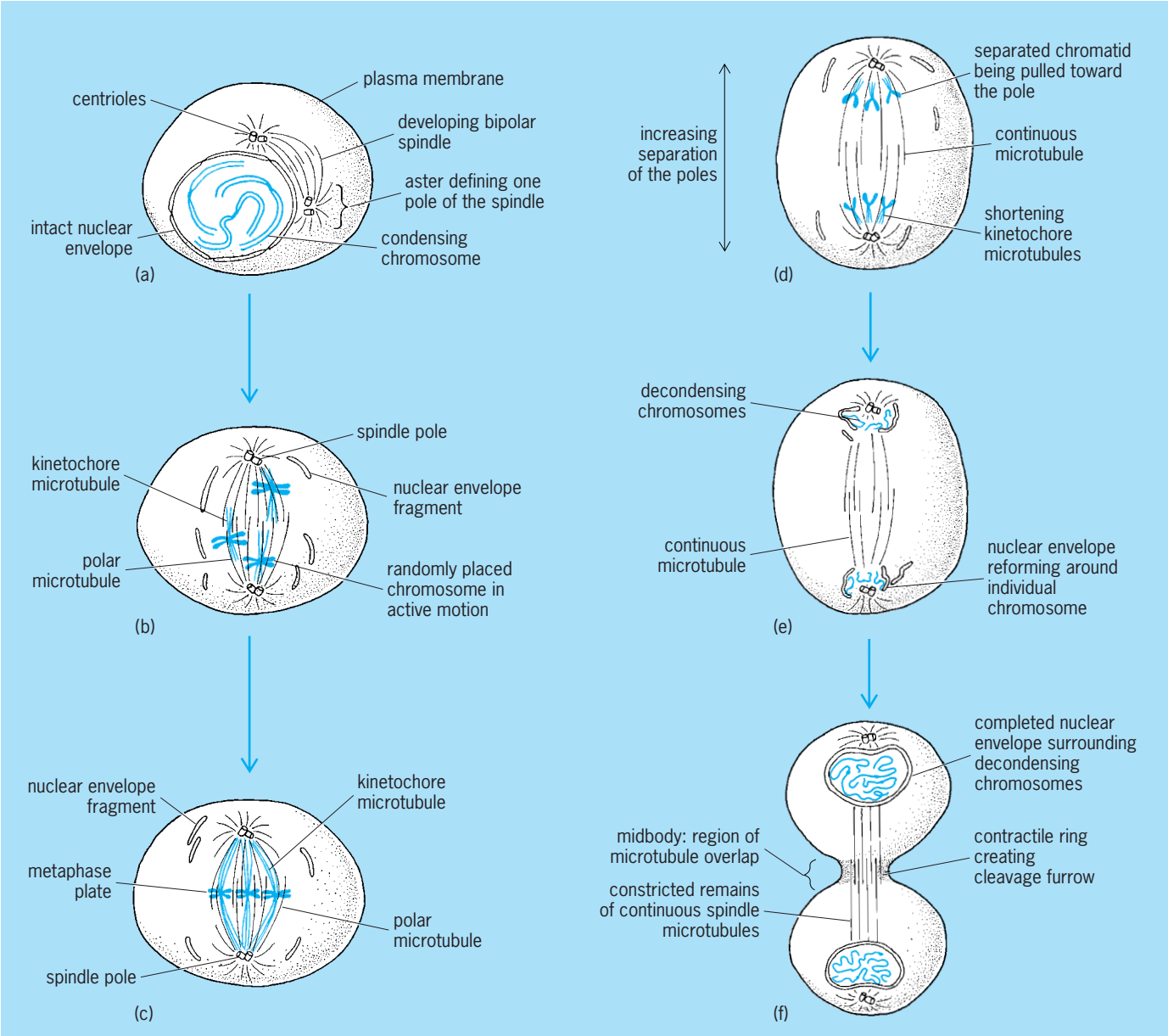

Representation of the stages of mitosis for a typical animal cell: (*a*) prophase; (*b*) prometaphase; (*c*) metaphase; (*d*) anaphase; (*e*) early telophase; (*f*) late telophase.

Metaphase. At metaphase, the kinetochore microtubules align their chromosomes in one plane called the metaphase plate, half way between the spindle poles, with one kinetochore facing each pole. Each sister chromatid is now perpendicular to the opposite poles. The metaphase stage of mitosis is the longest, lasting up to 30 min in mammalian cells. The kinetochore microtubules align the sister chromatids by holding them in tension with the two opposite poles. This tension seems to be necessary for the correct movement of the chromatids in the next stage.

Anaphase. The sister chromatids are triggered to separate and migrate to opposite spindle poles, and assemble there to form the nucleus of a new cell. Anaphase is separated into two stages. During anaphase A, a sister chromatid pair separates at the kinetochore and is guided to an opposite pole as the kinetochore microtubules begin to shorten at a rate of 1 micrometer per minute. This separation occurs simultaneously for all the chromatids. As the chromatids reach the poles, the microtubules begin to disassemble. During anaphase B, the disassembling polar microtubules begin to elongate so that the two poles of the spindle move farther apart.

Telophase. The end stage of mitosis, termed telophase, begins with the arrival of the sister chromatids at each separate pole. As the kinetochore microtubules disappear, the polar microtubules continue to elongate. A new nuclear envelope begins to form around each group of sister chromatids. The chromatin which has been condensed throughout

mitosis expands, and nucleoli appear. After this stage, the cell enters cytokinesis. *See* CELL NUCLEUS.

Cytokinesis. Cytokinesis begins during anaphase and ends the mitotic process. The cytoplasm divides by a process known as cleavage to give rise to two separate daughter cells. The forces required for cleavage are generated through actin-myosin interactions, which pull the plasma membrane down a furrow, a puckering of the plasma membrane. The cleavage furrow forms around the equator of the parent cell so that the two daughter cells produced are of approximately equal size and have similar properties.

Induction of mitosis. The alternation between interphase and M phase is controlled by periodic changes in the activity of maturation promoting factor. This factor in amphibian eggs is present only in M-phase cytoplasm, and its activity is regulated by phosphorylations. Activity of maturation promoting factor depends on two protein species associated in equimolar amounts in purified preparations. The first component is the protein kinase catalytic subunit $p34^{cdc2}$. The second component is a 62-kilodalton protein which must be synthesized at each metaphase-anaphase transition in mitosis. Because of this cyclical behavior, this protein was identified as cyclin B; it is the maturation promoting factor activator protein. There is a delay in an organism between the time after which protein synthesis is no longer required and the time when $p34^{cdc2}$ kinase activation occurs. The suggestion is that cyclin B accumulation required for maturation promoting factor activation is not the only factor involved in the activation process.

Disruption. The order of the cell cycle is ensured by dependent relationships, the initiation of late events depending upon the completion of preceding ones. Such dependence can be circumvented by pharmacological intervention, such as the induction of premature chromosome condensation by caffeine, blockade of mitosis, and inhibition of microtubule formation by colcemid, or inhibition of cytokinesis by cytochalasin B. Mutations, which are more specific than inhibitions, can also detect molecules involved in cell division. Mutants and inhibitors aid in the identification of control mechanisms in cell division, named checkpoints, that govern mitotic entry and exit. *See* MITOSIS; MUTATION.

Khandan Keyormarsi; Nuala O'Leary; Arthur B. Pardee

Bibliography. N. Akkas (ed.), *Biomechanics of Cell Division*, 1987; B. Alberts et al., *Molecular Biology of the Cell*, 3d ed., 1999; R. Baserga, *The Biology of Cell Reproduction*, 1985; J. S. Hyams and B. R. Brinkley (eds.), *Mitosis*, 1989; J. D. Watson et al., *Recombinant DNA*, 1992.

Cell lineage

A type of embryological study in which the history of individual blastomeres (cells formed during division of the zygote) or meristem cells is traced to their ultimate differentiation into tissues and organs.

Animals. The question of how the animal genome can be regulated to produce the various cell types found in the larval and adult organism is a central concern in developmental biology. A possible approach to this problem would involve tracing the structural fates of the descendants of each of a population of progenitor cells, and then trying to determine which gene products are required for particular steps in the process of cell differentiation. The cleaving embryos of a limited number of invertebrates, such as some flatworms, annelids, mollusks, and ascidians, are suitable for such studies because they show stereotyped cleavage patterns and asymmetries in size and arrangement of the blastomeres. Studies of this sort have been of limited interest because the large number of cells produced make it impossible to follow clones beyond the establishment of the germ layers. *See* CLEAVAGE (DEVELOPMENTAL BIOLOGY).

Cell lineages are also studied by marking cells in developing tissues and then observing the pattern of marked descendants in the adult structure. X-ray-induced alteration of cellular phenotype is the most powerful method of studying cell lineages in the imaginal development of insects. The x-rays cause a calculable number of chromosome crossing-over events which mark the affected cells by altered phenotype, such as different color or bristle configuration. The spatial pattern of cells with altered phenotype is then observed in the adult and, in ideal cases, represents the progeny of one founder cell. *See* FATE MAPS (EMBRYOLOGY).

Some of the most promising cell lineage studies are conducted on a nematode worm, *Caenorhabditis elegans*, which is a small (1 mm or 0.04 in. in length), nearly transparent worm that lives in soil. Adults are either males or hermaphrodites; the hermaphrodites contain 959 somatic nuclei. The origin of each somatic cell can be traced back to a single blastomere, and the clonal history of each cell has been determined (see **illus.**). A detailed genetic map for the 80,000 kilobase genome has been worked out. The short (3 days) generation time makes it possible to follow the effects of defined mutations upon the differentiation of specific cells. The relative roles of genetic programs as opposed to extrinsic signals can be assessed, for example, by destroying a particular cell during a defined stage of development and observing the effect on the differentiation of surrounding cells. The techniques of molecular genetics, particularly insertion and deletion of specified stretches of deoxyribonucleic acid (DNA) at specific chromosomal sites in *C. elegans*, should be particularly useful in determining the role of regulatory genes whose product may not be a structural or enzymatic protein. Because each stage of the differentiation of a particular cell can be described, the effects of alteration of the activity of any gene on the differentiation or interaction with surrounding cells can be precisely delineated. *See* DEVELOPMENTAL GENETICS; INVERTEBRATE EMBRYOLOGY; OVUM.

Spencer J. Berry

Plants. Cell lineage analysis in plants, as in animals, involves tracing the origin of particular cells in the adult body back to their progenitor cells. The

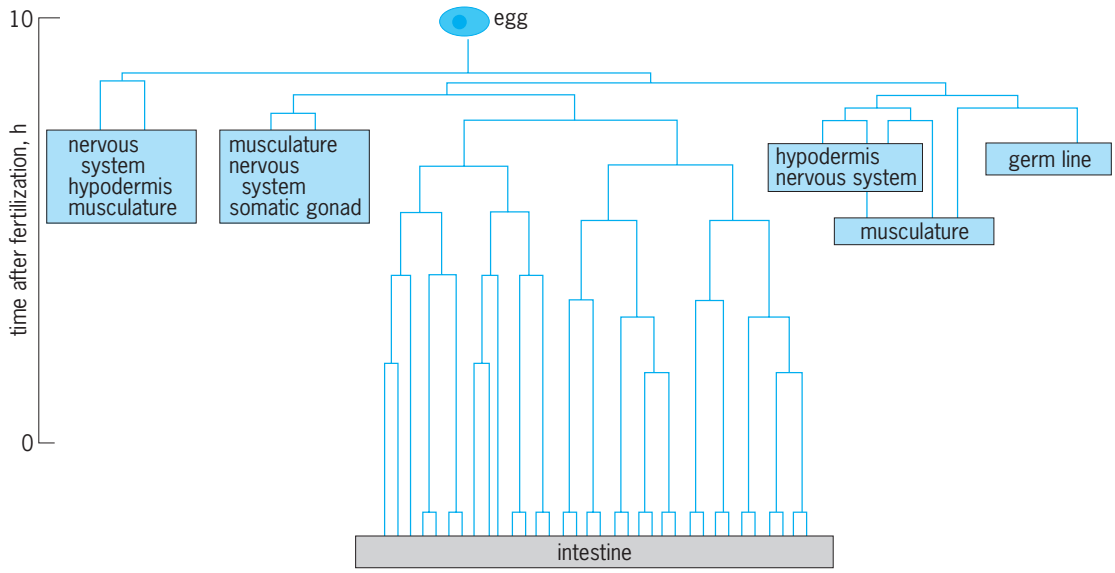

Lineage tree for the cells that form the intestine of *Caenorhabditis elegans*.

adult body of a typical plant consists primarily of leaves, stems, and roots. Cells arise continuously during plant life from specialized dividing cell populations called meristems. A shoot apical meristem produces the leaves and stem, and a root apical meristem produces root tissue. The shoot apical meristem will also produce specialized structures, such as cones, flowers, thorns, tendrils, and so on. Because plant cells do not move during development, and in many cases the plane of cell division is constant, lines of cells, called cell files, all derive from a single meristem cell at the base of the file. Cell files are particularly obvious in root tissue of ferns and higher plants. Although the organization of cells in plants follows regular patterns, the fate of individual cells is not as rigidly determined as cell fates are in animals. In plants, anatomical position determines a cell's typical fate, but if the pattern is disrupted by injury, for example, cells can differentiate into a variety of cell types. *See* APICAL MERISTEM.

Study of chimeras. Cell lineage can be examined in a more sophisticated manner by examining individual organs, such as leaves, to determine how many cells form a leaf and how these cells are oriented to one another as the organ forms. Exact cell lineages are most readily followed by using a color marker, such as chlorophyll-deficient cells which are white. Plant developmental biologists have used naturally occurring green-white chimeric plants, such as variegated ivy, or chimeras have been created by mutation. A common practice is to x-irradiate seeds that are heterozygous for one or more color loci (+ for color, − for albino) to create some meristem cells in which the chromosome arm carrying the (+) allele is lost from the cell, leaving only the (−) allele. This (−) cell and all of its progeny will be albino. *See* CHIMERA.

By using such mutated stocks, the contribution of the albino cell lineage to individual organs can be determined. Two types of observations are made. The first is the apparent cell number. If a leaf is 90% green and 10% albino, the simplest explanation is that the leaf was originally composed of 10 cells: 9 normal green ones and 1 mutant albino one. Thus the apparent cell number of the leaf is 10, indicating that 10 cells in the shoot apical meristem contributed to the leaf's formation. The second observation is the extent of the defect. If the albino sector is found in only a single leaf, this suggests that the meristem is very large and that the albino cell, and by inference all other cells in the meristem, typically contributes to only a single leaf on the plant. More often, albino sectors will be found on a number of leaves, indicating that meristem cells contribute to many different leaves. In some cases in which the plant has opposite leaves, all of the leaves on one side may be completely green while those on the opposite side all have an albino sector. Such an observation suggests that there are two groups of cells in the meristem, each producing leaves exclusively on one side of the stem.

Lineage in growing leaves. To gain further information about the pattern of cell division within an organ such as a leaf, experimentalists have x-irradiated leaves as they are growing. The earlier the mutations are induced, the larger the resulting sectors, because early in development the leaf is composed of few cells, each of which will give rise to a substantial fraction of the final leaf. Mutations induced late in development, when the leaf has thousands of cells, result in sectors that are tiny dots or streaks on the leaf. Such tiny sectors are used to determine where cell divisions are occurring in the leaf late in its development; areas lacking any dots or streaks have their final cell number, so no mutant cell lineages are visible.

Cell lineage analysis is a powerful tool for the study of plant development. By simple rules of observation and analysis, a great deal can be learned about the structure and activity of the meristems and about the

timing and pattern of cell division in the developing organs. *See* PLANT GROWTH; PLANT MORPHOGENESIS.

Virginia Walbot

Bibliography. B. Alberts et al., *The Molecular Biology of the Cell*, 3d ed., 1999; E. H. Davidson, *Gene Activity in Early Development*, 2d ed., 1976; A. Garcia-Bellido, P. A. Lawrence, and G. Morata, Compartments in animal development, *Sci. Amer.*, 1979; P. Grant, *Biology of Developing Systems*, 1978; M. M. Johri and E. H. Coe, Jr., Clonal analysis of corn plant development, *Develop. Biol.*, 97:154–172, 1983; C. Kenyon, The nematode *Caenorhabditis elegans, Science*, 240:1448–1453, 1988; N. Le Douarin (ed.), *Cell Lineage, Stem Cells and Cell Differentiation*, 1979; W. F. Loomis, *Developmental Biology*, 1986; S. Subtelny and I. M. Sussex (eds.), *The Clonal Basis of Development*, 1978; K. M. Wilbur (ed.), *The Mollusca*, vol. 3, 1983.

Cell membranes

Cells maintain their content separate and distinct from the external environment by a semifluid lipid bilayer that prevents the free exchange of most biological molecules and acts as a barrier between the inside and the outside of a cell. Prokaryotic cells contain only one such membrane, the plasma membrane that delineates the border of the cell. Eukaryotic cells, in addition to the plasma membrane, contain a multitude of internal membranes that define distinct intracellular organelles. The role of the plasma membrane and the internal membranes is analogous—all prevent the free exchange of molecules between the inside and the outside of the enclosed compartment, thus allowing the generation and maintenance of specialized microenvironments within each membrane-enclosed space. *See* CELL (BIOLOGY); CELL ORGANIZATION.

Membrane structure and composition. All biological membranes are approximately 7.5–10 nm thick and have a characteristic "railroad tracks" appearance when viewed in thin sections with the electron microscope, with two dark layers of the track, each approximately 2.5 nm thick, separated by a lighter layer, also approximately 2.5 nm (**Fig. 1***a*). This appearance is due to the heavy staining of the charged head groups and the poor staining of the fatty acyl chains (long chains of C-H bonds ending in a —COOH group) of phospholipids, the predominant lipid component of membranes. All membranes are made up of two phospholipid layers in a tail-to-tail arrangement with the charged head groups on the outside and the acyl chains on the inside (Fig. 1*b*). This basic bilayer organization was shown to exist in all examined membranes, indicating that the fundamental structure of all biological membranes is constant, despite differences in their composition and function.

All membranes are made up of lipid and protein, with the ratios varying from approximately 80% lipid (in the myelin membrane) to less than 30% lipid (in the mitochondrial membrane). All lipid constituents of the membrane are amphipathic; that is, each molecule has a hydrophilic (polar) and a hydrophobic (nonpolar) region. The structure of the membrane is intrinsically defined by this property, since amphipathic molecules spontaneously orient themselves with their hydrophobic cores to the interior and the polar groups to the exterior; by attempting to shield their hydrophobic domains, amphipathic lipids form sealed membrane "balloons" in an aqueous environment.

Proteins are the primary determinants of membrane function. They are found embedded within

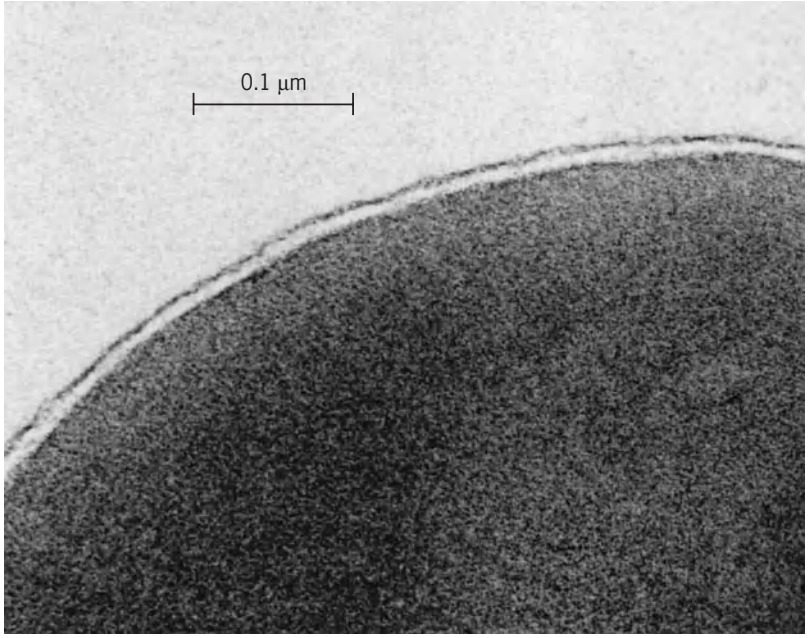

(a)

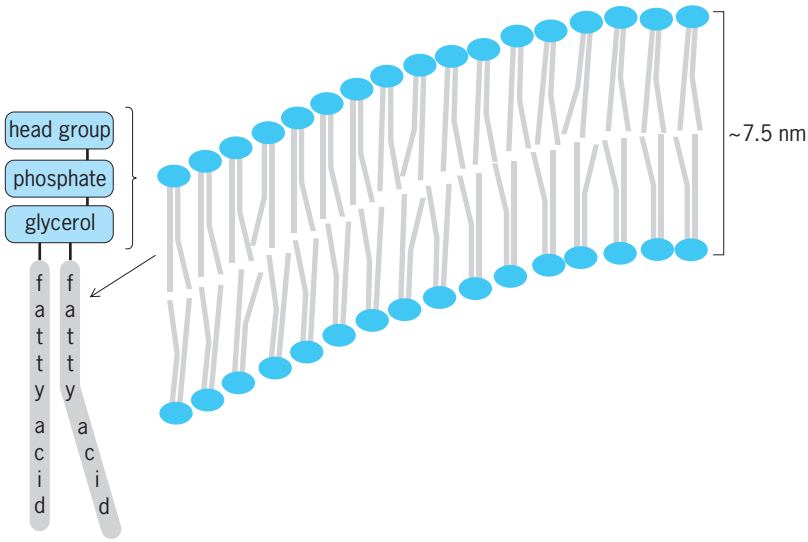

(b)

Fig. 1. Bilayer structure of biological membranes. (*a*) Electron micrograph of a thin section of a human red blood cell, showing the "railroad tracks" appearance of the plasma membrane (*from J. D. Robertson, Cell membranes, McGraw-Hill Encyclopedia of Science & Technology, 9th ed., 2002*). (*b*) Schematic diagram of the bilayer organization of phospholipids within the membrane.

the hydrophobic core of the lipid bilayer or attached to the hydrophilic membrane surface. The protein composition of membranes varies extensively. Many membrane proteins required for basic housekeeping functions are found in all cells, whereas more esoteric proteins are expressed in a restricted manner and are found only in cells specialized to perform specific functions.

Membrane lipids. The major membrane lipids are the phospholipids, the glycolipids, and the sterols (**Fig. 2**). The relative ratios of different lipids vary in different membranes, but usually phospholipids are the major constituent (50–90% of total lipids), followed by glycolipids (5–20%) and cholesterol (0–10%). *See* LIPID.

Phospholipids are based on a glycerol backbone and contain two fatty acyl chains linked to a glycerol esterified to a phosphate that is in turn esterified to a head group such as choline (to make phosphatidylcholine, PC), serine (PS), ethanolamine (PE), or inositol (PI). However, phospholipids can also be based on sphingosine (a single fatty acyl chain linked to a serine). The size of the fatty acyl chains of membrane phospholipids varies but is usually between 14 and 24 carbons in length. The level of saturated (single) to unsaturated (double) bonds also varies depending on the membrane, and acyl chains within a single phospholipid exhibit different levels of unsaturation. *See* PHOSPHOLIPID.

Glycolipids are often based on a glycerol framework and contain two acyl chains attached to a glycerol bound to a sugar chain; these are predominantly found in plants and in bacteria. However, in animal cells the backbone of glycolipids is sphingosine. When the amino group of the serine moiety of sphingosine is linked to a second acyl chain, a ceramide is produced. Ceramide can be further modified by the attachment of a phosphate group esterified to choline to produce sphingomyelin. Alternatively, ceramide can be modified by the addition of a single sugar (galactose) to generate cerebroside, or by the addition of branched sugar chains (composed of glucose, galactose, glucosamine, and sialic acid) to form gangliosides. The complex head groups of gangliosides define their function, and the family has more than 60 different members. *See* GLYCOLIPID.

As for sterols, cholesterol is the main sterol found in animal membranes, whereas various phytosterols predominate in plants. Sterols are rare or absent in prokaryotes. Sterols have the structure of four planar rings and are thus relatively rigid, flattened structures. They orient in the membrane with the small hydrophilic hydroxyl group on the outside and the hydrophobic steroid rings buried within the membrane. *See* CHOLESTEROL.

Membrane proteins. Proteins are responsible for most of the specialized membrane functions by acting, for example, as selective transport portals, signal

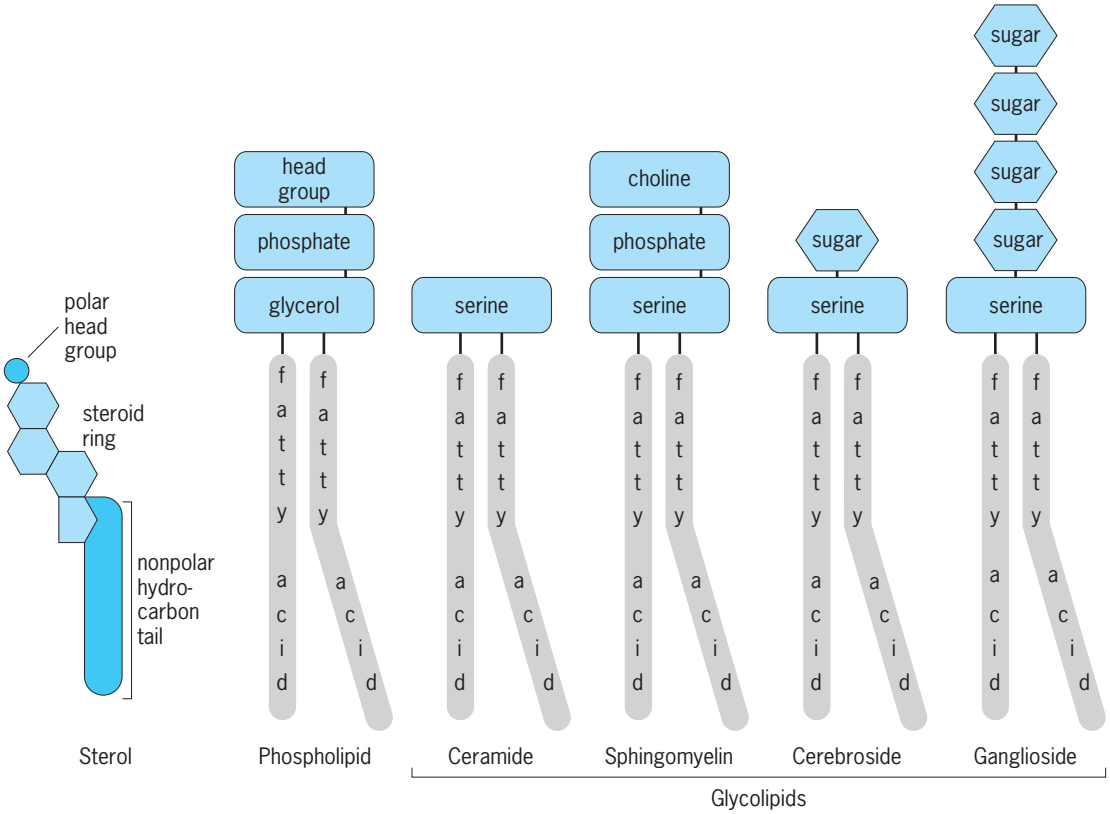

Fig. 2. Structure and arrangement of lipids within membranes. The schematic represents major membrane lipids. The head group of a phospholipid can be choline (phosphatidylcholine, PC), serine (PS), inositol (PI), or ethanolamine (PE). The sugar residue on cerebroside is galactose, and the sugar chains on gangliosides are made up of varying amounts and arrangements of glucose, galactose, *N*-acetylgalactosamine, and sialic acid. The fatty acyl chains can vary in length and in the number and position of double bonds.

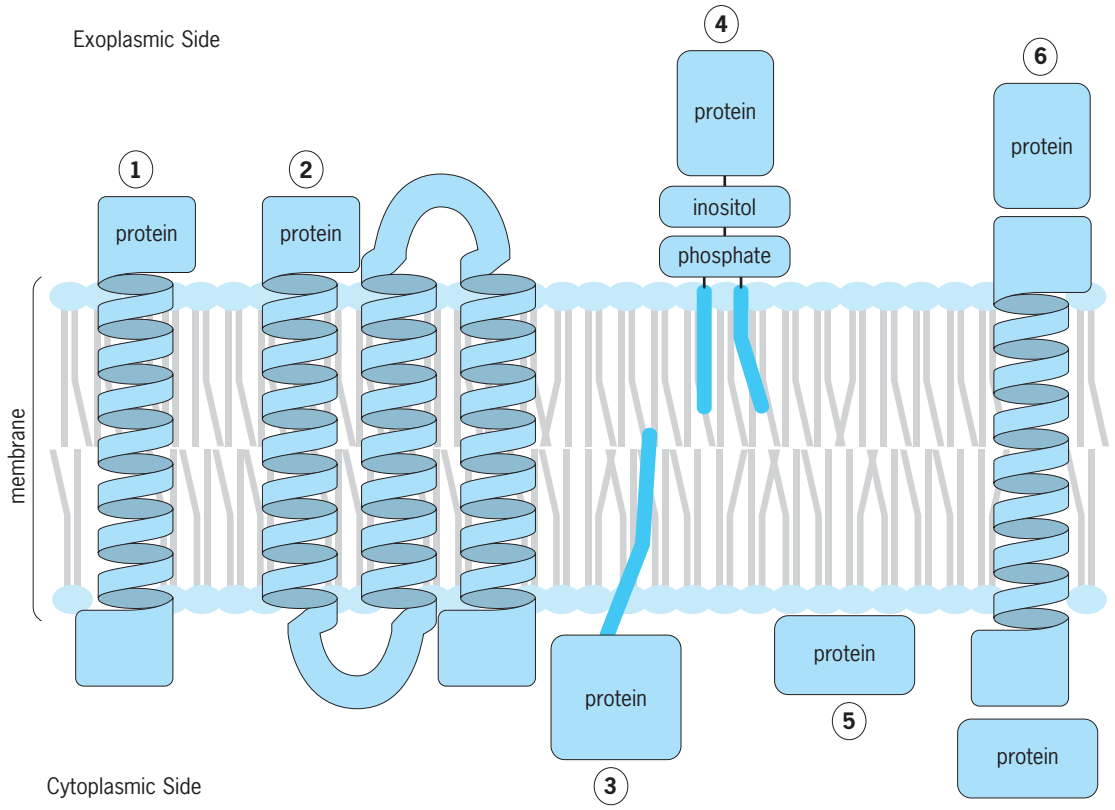

Fig. 3. Structure and arrangement of proteins within membranes. The asymmetric arrangement of various types of membrane proteins is shown in respect to the bilayer: (1) Integral single-span protein: the protein can be arranged with the *N*-terminus facing the exoplasmic side (type I protein) or with the *N*-terminus facing the cytoplasmic side (type II protein). (2) Integral multispan protein: multiple segments of the protein can span the bilayer creating cytoplasmic and exoplasmic loops. (3) Lipid-anchored protein: the protein can be attached to a myristyl, palmityl, prenyl, farnesyl, or geranylgeranyl chain. (4) GPI-linked protein. (5) Peripherally associated protein interacting with lipids. (6) Peripherally associated protein interacting with an integral membrane protein. The protein can be associated with the exoplasmic or cytoplasmic side of the membrane.

receptors and transducers, matrix attachment and cell-cell contact mediators, identity determinants, and enzymes.

Proteins that have one or more segments of their polypeptide chain threaded through the membrane with the remainder of the chain sticking out on either side of the membrane are called integral membrane proteins or transmembrane proteins (**Fig. 3**). In the majority of studied transmembrane proteins, the part that spans the membrane is either α-helical or is arranged in multiple β-strands, forming a barrel-shaped structure. The predicted length of an α-helix necessary to span the bilayer is 3.75 nm, and this usually corresponds to approximately 22 amino acids.

Other proteins are associated with the membrane by being covalently bound to specialized lipid molecules (lipid-anchored proteins), or by interacting noncovalently with transmembrane proteins (peripheral proteins). Lipid-anchored proteins that are found attached to the cytoplasmic aspect of cellular membranes are usually attached to a myristyl-, palmityl-, prenyl-, farnesyl-, or geranyl-carbohydrate moiety buried within the bilayer. In contrast, lipid-anchored proteins that are found within the lumen of secretory organelles are attached to a complex glycosylated phospholipid, glycosylphosphatidylinositol (GPI). The addition of the GPI tail occurs within the lumen of the endoplasmic reticulum after the protein is translocated there, and all GPI-linked proteins face the lumenal or the external leaflet of the membrane. Peripheral membrane proteins are usually associated through direct electrostatic interactions with the polar head groups of lipids, or by binding to integral membrane proteins. *See* PROTEIN.

Membrane fluidity. Membranes are fluid at ambient temperatures. They allow the free axial rotation of all components, and free flow within the plane of the bilayer is observed for lipids and for unrestricted proteins consistent with the fluid mosaic model of membrane structure, in which components are free to diffuse within the two-dimensional plane of the membrane.

Lipid molecules define the fluidity of a membrane. Lipids undergo phase transition, from a relatively rigid gel-like state at low temperature to a fluid state at high temperature. Differences in the length and the saturation of the acyl chains define how tightly the molecules can pack into the bilayer and thereby influence fluidity of the membrane. In general, lipids with short or unsaturated fatty acyl chains do not pack very tightly and undergo phase transition (are more fluid) at lower temperatures than those with long or polyunsaturated chains that can pack more tightly.

A typical lipid molecule changes places with its neighbor approximately 10^7 times per second and diffuses several micrometers per second at 37°C (98.6°F). This means that a single lipid molecule can traverse the length of a typical eukaryotic cell in approximately 5–10 s. However, the movement of lipids within the plane of the bilayer is restricted to only one leaflet, and lipids do not normally flip-flop from one side to the other. Depending on the type of cell, 30–90% of integral proteins in the plasma membrane and intracellular membranes are mobile. The mobility of membrane proteins within the lipid bilayer is significantly (approximately a hundredfold) lower than that of lipids. Protein movement is important since key cellular events, such as receptor clustering and capping and cell-cell and cell-matrix attachment, depend on the mobility of transmembrane proteins. Most transmembrane proteins are restricted by linkage to cytoskeletal components, thus tethering membranes to cytoskeletal systems.

Membrane asymmetry. Since membranes separate two compartments, each leaflet of the bilayer faces a distinct environment. For the plasma membrane, the inner (or cytoplasmic) leaflet is exposed to the cytosol while the outer (or exoplasmic) leaflet faces the extracellular space; for intracellular membranes, the cytoplasmic leaflet is exposed to the cytosol while the exoplasmic leaflet faces the inside (the lumen) of the organelle.

Membranes are asymmetric, due to differences in lipid composition between the cytoplasmic and the exoplasmic leaflet. Since lipids do not ordinarily flip-flop across the bilayer, the differential distribution is maintained throughout the life of the membrane. Lipids with neutral or negative head groups (such as PE, PS, and PI) are found preferentially facing the cytosol, whereas PC (which contains a positively charged head group) is preferentially within the exoplasmic leaflet. The fatty acyl chains in the exoplasmic leaflet tend to have longer, more saturated carbon chains than those in the cytoplasmic leaflet. Glycolipids are found exclusively in the exoplasmic leaflet since sugar addition occurs within the lumen of the endoplasmic reticulum and the Golgi apparatus. Similarly, cholesterol is predominantly located in the exoplasmic leaflet and, due to its planar structure, increases the relative "thickness" of that leaflet. Sterols are most abundant in the plasma membrane, where they increase the rigidity and the mechanical stability of the membrane.

Membrane asymmetry is also evident in the protein arrangement within the bilayer, since only a particular transmembrane orientation, which is determined by the primary sequence of the protein, is possible. Protein asymmetry is essential since it is responsible for the directionality of transport and signaling processes, the sidedness of enzymatic reactions, and the surface exposure of specific protein sequences required for recognition and adhesion of cells; protein asymmetry underlies all other protein-mediated membrane processes as well.

Membrane functions. The fundamental function of all membranes is to serve as barriers that enable cells to maintain a different composition between two aqueous compartments, and this function is due to the lipid components of the membrane. However, lipids can also play additional roles. For example, PI is involved in a signaling cascade mediating the transmission of signals from hormone receptors to influence transcription of specific genes; cholesterol is involved in the synthesis of low-density lipoprotein particles (LDLs), which are utilized in mammals to transport cholesterol from the liver to other tissues; gangliosides have been shown to define the A, B, and O blood groups; and sphingomyelin is required to surround and insulate neuronal axons, allowing neuronal transmission.

A myriad of membrane proteins mediate various additional membrane functions, among them the selective transport of molecules, reception and transduction of signals, and energy conversions. Membranes are impermeant to charged molecules such as ions, amino acids, and adenosine triphosphate (ATP) and to all large molecules such as sugars and peptides. Such compounds are transported across membranes by protein pumps that move compounds against a concentration gradient, channels that transport down concentration gradients, or transporters that move compounds down or against a concentration gradient. The fundamental importance of membrane transport systems is exemplified by the transport of neurotransmitters into synaptic vesicles that subsequently mediate neuronal transmission and the transport of calcium ions into the sarcoplasmic reticulum allowing subsequent muscle contraction. *See* BIOPOTENTIALS AND IONIC CURRENTS; CELL POLARITY (BIOLOGY); SYNAPTIC TRANSMISSION.

One of the most important functions of membranes is in signal reception and relay. Extracellular signaling molecules such as growth factors, antigens, hormones, and neurotransmitters bind to extracellular portions of integral membrane receptors. Binding results in receptor stimulation, such that the cytoplasmic portion of the receptor is activated to mediate signaling events inside the cell. In all cases, the binding of a specific ligand to the outside of the receptor will ultimately result in the activation of particular signaling cascades, resulting in changes in cell morphology, locomotion, differentiation, gene expression, and so on. For example, signaling through the nerve growth factor receptor is required for neuronal survival and development and is essential for restoring innervation in wounded skin or muscle. *See* SIGNAL TRANSDUCTION.

Membranes are also intimately involved in energy generation and conversion reactions, without which eukaryotic life would be impossible. Within mitochondria, proteins that are embedded in the mitochondrial inner membrane and make up the electron-transport chain utilize the energy generated by oxidative reactions to create a transmembrane electrochemical gradient, which ultimately is used to generate ATP. Within chloroplasts, the thylakoid membrane contains protein complexes that capture electrons after chlorophyll is excited by sunlight

and other protein complexes involved in electron transport processes that ultimately result in ATP synthesis. Both ATP-producing processes are fundamentally dependent on the differential in ionic concentrations generated and maintained across the impermeant inner mitochondrial or thylakoid membrane. *See* ENERGY METABOLISM; MITOCHONDRIA; PHOTOSYNTHESIS.

Membrane biogenesis. Nearly all membrane lipid is synthesized by integral endoplasmic reticulum proteins with their catalytic domains protruding from the endoplasmic reticulum membrane into the cytosol. Thus, the synthesis of lipids occurs within an aqueous cytosolic environment. The substrates for the synthesis, for example fatty acids, are maintained in the cytosol by being esterified to acyl carrier proteins (ACPs), thereby sheltering their hydrophobicity. The first step in phospholipid synthesis involves the movement of two fatty acids from the ACP to a soluble glycerol 3-phosphate to generate phosphatidic acid (PA), which is inserted into the cytoplasmic leaflet of the endoplasmic reticulum bilayer. Phosphatidic acid is then modified by the removal of the phosphate to form diacylglycerol, which is subsequently utilized to make various phospholipids by adding specific head groups. The newly made lipids rapidly diffuse within the cytoplasmic leaflet of the endoplasmic reticulum membrane and are apparently rapidly equilibrated between the two membrane leaflets by the action of an enzyme called a scramblase. As a result, the different types of phospholipids are equally distributed within the two endoplasmic reticulum membrane leaflets. The newly synthesized lipids are transferred by vesicle budding and fusion to membranes of distal components of the secretory patway, that is, the Golgi apparatus, lysosomes, secretory granules, and the plasma membrane. The highly asymmetric distribution of phospholipids in the two leaflets of the plasma membrane results from the action of flippases that transfer lipids containing free amino groups—phosphatidylserine and phosphatidylethanolamine—from the leaflet facing the cell exterior to the one facing the cytoplasm. Gangliosides and sphingomyelin are synthesized in the Golgi apparatus from ceramide brought from the endoplasmic reticulum and are located exclusively in the lumenal, or exoplasmic, leaflet as a consequence of their synthesis by enzymes exposed on the surface of that leaflet.

All integral and GPI-linked membrane proteins are synthesized by multicomponent synthetic

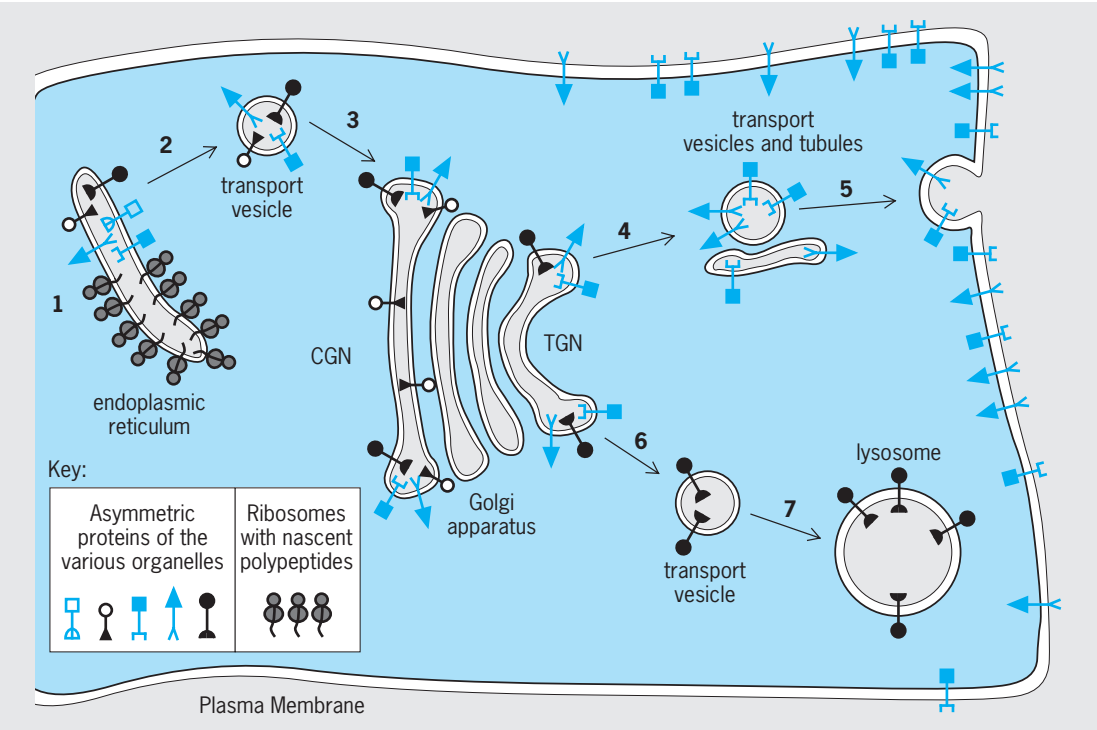

Fig. 4. Biogenesis of membranes. The membranes of the distal organelles of the secretory pathway—the Golgi apparatus, lysosomes, secretory granules, and the plasma membrane—are derived from the endoplasmic reticulum membrane by a process of vesicular transport, in which vesicles containing both lipids and proteins bud off from one compartment and fuse with a receiving compartment. Transmembrane proteins are synthesized in the endoplasmic reticulum (1) by ribosomes attached to the cytoplasmic face of the endoplasmic reticulum membrane and inserted into the membrane during the course of their synthesis. Each protein achieves in the endoplasmic reticulum a characteristic asymmetric disposition relative to the bilayer, which is maintained in recipient membranes as the protein is transported via vesicles between compartments. Some proteins are retained in the endoplasmic reticulum, and others are incorporated into transport vesicles (2) to be delivered (3) to the cis region of the Golgi apparatus (CGN). The proteins move through the Golgi apparatus by mechanisms that are not definitively established, with some remaining as residents in particular cisternae of this polarized organelle. Proteins that reach the farthest region of the Golgi apparatus, known as the trans-Golgi network (TGN), are then sorted (4, 6) into transport vesicles or tubules that bring them to the plasma membrane (5), developing lysosomes (7), or secretory granules (not shown).

machineries associated with the cytosolic side of the endoplasmic reticulum membrane. The membrane insertion of such proteins occurs while the protein is being synthesized (cotranslationally). Depending on its primary sequence, the protein will either span the membrane once or multiple times, or will acquire a GPI anchor. Proteins associated with the cytosolic side of the membrane are synthesized within the cytosol, some are modified by the attachment of lipid anchors, and all associate with the membrane after their synthesis is complete (posttranslationally). *See* ENDOPLASMIC RETICULUM.

New membrane for all secretory and endocytic organelles originates in the endoplasmic reticulum. From the endoplasmic reticulum, the newly synthesized lipid and protein components are transported by being incorporated into membrane vesicles, first to the Golgi apparatus, then from the Golgi to the plasma membrane, endosomes, lysosomes, or secretory granules. In some cases, proteins delivered from the Golgi to the plasma membrane are internalized and then delivered to endosomes or lysosomes (**Fig. 4**). Some lipids and proteins will be "left behind" in each compartment, whereas others will continue to be moved via vesicular traffic to their final destination. *See* GOLGI APPARATUS.

In contrast, organelles such as peroxisomes, mitochondria, and plastids do not receive endoplasmic reticulum–derived membranes. They maintain and enlarge their membranes by directly incorporating lipid molecules that are delivered by phospholipid transfer proteins. The transfer proteins extract a specific type of phospholipid from a membrane, diffuse within the cytosol, and then transfer the phospholipid into another membrane. Exchange appears to be concentration-dependent, and phospholipid is transported from phospholipid-rich (endoplasmic reticulum) membranes to phospholipid-poor (peroxisomal or mitochondrial) membranes. The exchange involves the transfer of phospholipid between the cytoplasmic leaflets of the two membranes, and it has been postulated that peroxisomal and mitochondrial flippases move the newly acquired phospholipid to the lumenal leaflet. Similarly, proteins are imported into peroxisomes, mitochondria, and plastids directly from the cytosol. Specific amino acid sequences in mitochondrial and plastid proteins define whether the protein will be integrated into the outer or the inner membrane and whether it will face the cytoplasmic, intermembrane, or internal compartment of these organelles. *See* CELL PLASTIDS; MITOCHONDRIA; PEROXISOME.

Membrane disease. Most membrane diseases are caused by defects in membrane proteins. One of the most studied diseases is cystic fibrosis, caused by mutations in a plasma membrane channel transporting Na^+ and Cl^- ions in epithelial cells. The most common mutation changes a single amino acid within the protein, and this leads to premature degradation and results in the absence of functional channels at the plasma membrane. As a consequence, the electrolyte homeostasis is altered in affected epithelia, leading to clinical manifestations of cystic fibrosis and ultimately death. *See* PANCREAS DISORDERS.

Elizabeth Sztul

Bibliography. B. Alberts et al., *Molecular Biology of the Cell*, 4th ed., Garland Publishing, 2002; W. R. Bishop and R. M. Bell, Assembly of phospholipids into cellular membranes: Biosynthesis, transmembrane movement, and intracellular translocation, *Annu. Rev. Cell Biol.*, 4:580–611, 1988; R. B. Gennis, *Biomembranes: Molecular Structure and Function*, Springer, New York, 1989; L. E. Henderson and S. Oroszlam, Fatty acylation of proteins, *Annu. Rev. Cell Biol.*, 4:612–648, 1988; H. Lodish, *Molecular Cell Biology*, W. H. Freeman, 2000; D. E. Vance and J. Vance (eds.), *Biochemistry of Lipids, Lipoproteins, and Membranes*, Elsevier, 1996; G. van Meer, Lipid traffic in animal cells, *Annu. Rev. Cell Biol.*, 5:247–276, 1989; K. W. A. Wirtz, Phospholipid transfer proteins, *Annu. Rev. Biochem.*, 60:73–99, 1991.

Cell metabolism

The sum of chemical reactions which transpire within cells. The cell is a self-contained chemical factory which interacts with its environment. This factory is unique in that it is a structure which moves, manufactures products for its own use and for the use of other factories, takes in from the outside that which it needs and discards that which it does not, and has the property of producing factories identical to itself from its own substance. The cell performs chemical, osmotic, mechanical, and electrical work, for which it needs energy. Plant cells obtain energy from sunlight; using light energy, they convert simple compounds such as carbon dioxide and various nitrogen, phosphate, and sulfur compounds into more complex materials. The energy in light is thus "stored" as chemical substances, mostly carbohydrates, within plant cells. Animal cells cannot use sunlight directly, and they obtain their energy by breaking down the stored chemical compounds of plant cells. Bacterial cells obtain their energy in various ways, but again mostly by the degradation of some of the simple compounds in their environment. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; PHOTOSYNTHESIS; PLANT METABOLISM.

Dynamic state. The chemical atoms of the biological world are relatively few in number, being mostly carbon, hydrogen, oxygen, nitrogen, phosphorus, sulfur, iron, magnesium, and minute amounts of others. These atoms are grouped together into biological molecules, the most prevalent being proteins, carbohydrates, fats, and nucleic acids. These latter differ greatly among themselves in their chemistry and their functions. Their characteristics are given elsewhere, but they all have one thing in common: They are being continually broken down or degraded and being resynthesized by the cell. Thus, a feature of the cell is continuous metabolism of its chemical constituents.

Furthermore, cells have definite structures, and even the chemical constituents of these structures are being constantly renewed. This continuing turnover has been called the dynamic state of cellular constituents. For example, an animal cell takes in carbohydrate molecules, breaks down some of them to obtain the energy which is necessary to replace the chemical molecules that are being turned over, while another fraction of these molecules is integrated into the substance of the cell or its extracellular coverings. The energy obtained is, of course, also used for all the purposes mentioned above. Another way of stating it is that the cell is constantly striving to maintain an organized structure, a structure which has many functions, not the least of which is to duplicate itself, in the face of an environment which is continuously striving to degrade that structure into a random mixture of chemical molecules.

Chemical specificity. All the large molecules of the cell have rather specific functions: Carbohydrates, fats, and proteins constitute the structures of the cells; these, particularly the former two, are also used for food, or energy, depots; the nucleic acids are the structures involved in the continuity of cell types from generation to generation. All these large molecules are really variegated polymers of smaller molecules. These smaller molecules interact with one another in chemical reactions which are catalyzed at cellular temperature by biological protein catalysts called enzymes. These chemical reactions involve the breakdown of molecules, the condensing of one molecule to another to form a larger one, and the reforming of a third compound from parts of two constituents.

All these reactions are very specific, being catalyzed by specific enzymes, each enzyme being a specific protein, having a singular structure and performing a singular function, only reacting with its own specified substrate or substrates. At present, about a thousand chemical reactions are known which occur within cells; thus, about a thousand specific enzymes are known, some having been purified to crystalline form. *See* CARBOHYDRATE METABOLISM; ENZYME; LIPID METABOLISM; NUCLEIC ACID; PROTEIN METABOLISM.

Metabolic pathways. By studying how these enzymes operate and what substrates they attack, the biochemist has learned in general, and in many cases in specific, how fats, carbohydrates, proteins, and nucleic acids are synthesized and degraded in cells. Mainly through the use of radioactive tracer atoms, the pathways of many chemical compounds within the cell have been realized. For example, it is known what part of the carbohydrate molecule is used for energy production, what part is used in fat storage, and what parts end up in proteins and nucleic acids. Via the vast array of enzymatic reactions which go on inside cells, the substances which a cell "eats" or brings in are completely changed, becoming transformed into cell substance. This changeover, these syntheses of specific cell substances, need energy for accomplishment.

Energy source. As mentioned, this energy comes from the breakdown of various foodstuffs, a burning up of substrate, an oxidation; but unlike mechanical machines in which energy is given off as heat, the cell has evolved an energy coin. This coin is a chemical compound called adenosine triphosphate (ATP); it is synthesized enzymatically by the cell in a number of reactions in which various compounds coming from foodstuffs are oxidized, and thus the energy gained as a result of this oxidation is not lost as heat but stored in ATP. Subsequently, all cellular reactions which require synthesis of cell-specific substances use this ATP as a source of energy for these syntheses. For example, fats are composed mainly of carbon, oxygen, and hydrogen; when fats are used as foodstuff by the cell, the compounds are degraded, oxygen and hydrogen end up in the form of water and carbon in the form of carbon dioxide, and overall energy is gained in the form of ATP. When fats have to be synthesized by the cell, this ATP is used as an energy source to recombine small chemical molecules containing these atoms into certain kinds of fats. But since the cell needs to synthesize other molecules as well, these same carbon atoms, for example, are also found (via synthetic reactions) in carbohydrates, proteins, and nucleic acids. In this way the molecules, of which these atoms are a part, are being constantly replaced. *See* ADENOSINE TRIPHOSPHATE (ATP).

Control and continuity. Remarkably, even with these constant replacements going on, the cell never loses its own distinctive structure and function. The reason is that the ordering of the cell resides in a code of nucleic acids, which directs the syntheses of specific enzymes designed to do specific tasks; when these enzymes, being proteins and being also replaced like the other molecules of the cell, are degraded and have to be resynthesized, they are made again in exactly the same way as before. In this way continuity is ensured. *See* DEOXYRIBONUCLEIC ACID (DNA); GENETICS; RIBONUCLEIC ACID (RNA).

Metabolic differentiation in cells. To accomplish the many tasks outlined above, the cell is compartmentalized into distinctive functioning structures. The production of usable energy in the form of ATP takes place in plant and animal cells in membranous structures called mitochondria. Even in bacteria, which are about the size of mitochondria, these events occur on the membrane surrounding the bacterium. In plants the capture of light energy and the production of ATP therefrom takes place in structures called chloroplasts. Proteins, including enzyme proteins, are sythesized by tiny particles called ribosomes. The synthesis of carbohydrates in plants occurs within chloroplasts, while the synthesis of fatty acids in animal cells occurs in mitochondria. The structures involved in the continuity of cell structure and function, in the heredity of the cell, are the nucleic acid–containing chromosomes of the nucleus. In addition, cells have a variety of other structures, not as important as those above, but which nevertheless have distinctive functional tasks, such as storage granules, secretory granules, and lytic granules.

See CELL PLASTIDS; CHROMOSOME; MITOCHONDRIA; RIBOSOMES.

Control mechanisms. Finally, although a great deal is known of the metabolism of a large variety of compounds, their degradation and syntheses, and the interaction between chemical substances, little is known of how these multitudinous reactions are regulated within the cell to effect growth, particular size, and division into daughter cells having the same structure and functioning characteristics as those of the parent cells. It is known that enzymatic reaction activities within cells are strictly governed so that in quite a few cases knowledge has been gained of how a cell shuts off the synthesis of a compound of which it has enough, or speeds up the syntheses of those in short supply. It is all done by an enzyme so constructed that the compound which it synthesizes, say, can interact with this enzyme to inhibit further activity of the enzyme; thus not too much of the compound is formed at any one time. In the cells of a multicellular organism, cells interact with cells, their metabolism being geared among themselves, and this gearing, or regulation, probably occurs via the action of intercellular messengers or hormones. *See* HORMONE.

In summary, almost the sole justification for cell metabolism is the functioning of a vehicle whose major task is to reproduce as precise a replica of itself as possible. The efficiency of this metabolism has been maximized with this goal in view. *See* CELL (BIOLOGY). Philip Siekevitz

Bibliography. T. A. Subramanian (ed.), *Cell Metabolism: Growth and Environment*, 2 vols., 1986; G. R. Welch and J. S. Clegg (eds.), *The Organization of Cell Metabolism*, 1987.

Cell motility

The movement of cells, changes in cell shape including cell division, and the movement of materials within cells. Many free-living protozoa are capable of movement, as are sperm and ameboid cells of higher organisms. Coordinated movement of cells occurs during embryogenesis, wound healing, and muscle contraction in higher organisms. Cell division is observed in all organisms and is a requirement for reproduction, growth, and development. Many cells also undergo structural changes as they differentiate, such as the outgrowth of axonal and dendritic processes during nerve cell differentiation. A more subtle form of cell motility involves the active transport of membranous organelles within the cytoplasm. This form of movement is required for proper organization of the cytoplasmic contents, and the redistribution of metabolites, hormones, and other materials within the cell.

Molecules responsible for movement. There are two basic molecular systems responsible for producing a variety of forms of movement in a wide range of cell types: one system involves filamentous polymers of the globular protein actin; the other involves hollow, tube-shaped polymers of the globular protein tubulin, known as microtubules (see **illus.**). Associated with both actin filaments and microtubules are accessory enzymes that convert the chemical energy stored in adenosine triphosphate (ATP) into mechanical energy. Other proteins are responsible for regulating the arrangement, assembly, and organization of actin filaments and microtubules within the cell.

Actin and myosin. Muscle contraction represents one of the most extensively studied forms of cell movement, and it is from muscle that much basic knowledge of actin-based movement has been derived. Striated muscle cells found in skeletal muscle and heart muscle contain highly organized arrays of actin filaments interdigitating with filaments of the protein myosin. Myosin has an enzymatic activity that catalyzes the breakdown of ATP to adenosine diphosphate (ADP) and phosphate. The released energy is used to produce force against the actin filaments, which results in sliding between the actin and myosin filaments. *See* MUSCLE PROTEINS.

These proteins have by now been found in virtually all cell types. The actin and myosin filaments are less highly organized in smooth muscle and nonmuscle cells than in striated muscle; however, they act by the same force-producing mechanism.

Microtubules, dynein, and kinesin. Microtubules were initially described as components of eukaryotic cilia and flagella. Like actin filaments, microtubules have by now been found within the cytoplasm of almost all eukaryotic cells.

Two different molecules, dynein and kinesin, have been identified as enzymes that break down ATP to ADP and phosphate to produce force along microtubules. Dynein is a large enzyme complex that was initially identified in cilia and flagella. It has also been found associated with cytoplasmic microtubules. Kinesin is a force-producing enzyme that was initially found in microtubules prepared from neuronal cells. It is now also known to be widespread.

Despite the basic similarity in how the three force-producing enzymes (myosin, dynein, and kinesin) work, they differ from each other in structure and enzymatic properties and there is no evidence that they are evolutionarily related. Kinesin and dynein differ from each other in another important way: they produce force in opposite directions along microtubules. This suggests that they play complementary roles in the cell. As yet, no enzyme has been identified that produces force along actin filaments in the direction opposite to myosin.

Molecular and genetic analysis of a variety of organisms has revealed a remarkable diversity of kinesin- and myosin-related proteins in cells. The parts of these molecules responsible for producing force are closely related to the corresponding parts of kinesin and myosin. However, the remainder of the kinesin- and myosin-related molecules are structurally diverse. This pattern suggests that the kinesin- and myosin-related proteins attach to different structures within the cell, each individual protein being

responsible for a specific type of movement. Multiple forms of dynein have also been identified. The members of the dynein family of proteins are much more closely interrelated than are the members of the myosin and kinesin families. The dyneins fall into two classes: a cytoplasmic dynein and several ciliary and flagellar dyneins. Novel tubulin- and actin-related proteins have been identified. With the exception of γ-tubulin, the functions of these proteins are as yet poorly understood.

Actin and microtubule assembly and organization. Another factor involved in cellular and intracellular movement is the assembly and rearrangement of actin and tubulin polymers. Actin filaments in muscle and microtubules in cilia and flagella are relatively stable structures. However, in other situations, both polymers are in equilibrium with their subunits. Thus the cell can make rapid and reversible changes in the number and organization of polymers.

Actin and microtubule assembly is often initiated at specific sites by specialized structures in the cell. In typical cultured eukaryotic cells, microtubules assemble outward from a centrosome. The centrosome is a complex structure that contains one or two centrioles (composed of a complex array of short microtubules) surrounded by diffusely organized proteinaceous material containing the novel protein γ-tubulin, which may serve to initiate microtubule growth. During the interphase portion of the cell cycle, the microtubules are seen to emanate radially from a single centrosome. During cell division the centrosome splits, and two overlapping radial arrays of microtubules appear. These arrays are subsequently remodeled to form the mitotic spindle. Differentiated cells exhibit a variety of microtubule arrangements consistent with the specific functional requirements of these cells. For example, neuronal cells are filled with parallel microtubules that are required for process formation and also serve as tracks along which materials are transported. Microtubules span epithelial cells and are involved in transport between the two surfaces of the epithelial cell layer. Microtubules form a circular band around blood platelets and are required for maintaining the discoid shapes of these cells. *See* CELL CYCLE; CENTROSOME.

Actin often assembles from sites associated with membranes. These sites may be diffuse or well defined, such as the cytoplasmic face of intercellular junctions. A variety of proteins that may be involved in organizing the actin filaments are localized at these sites, but the precise arrangement of these proteins and their role in initiating the assembly of actin are not fully understood.

Other factors affect the assembly of actin filaments and microtubules. Energy is required for polymerization. In the case of actin filaments, energy is supplied by the breakdown of ATP to ADP and phosphate as actin subunits add to the end of the growing filament. In the case of microtubules, energy is supplied by the breakdown of guanosine triphosphate (GTP) to guanosine diphosphate (GDP) and phosphate. Microtubule assembly is inhibited by calcium, but it is not certain whether this mechanism is actually used to regulate assembly in the cell.

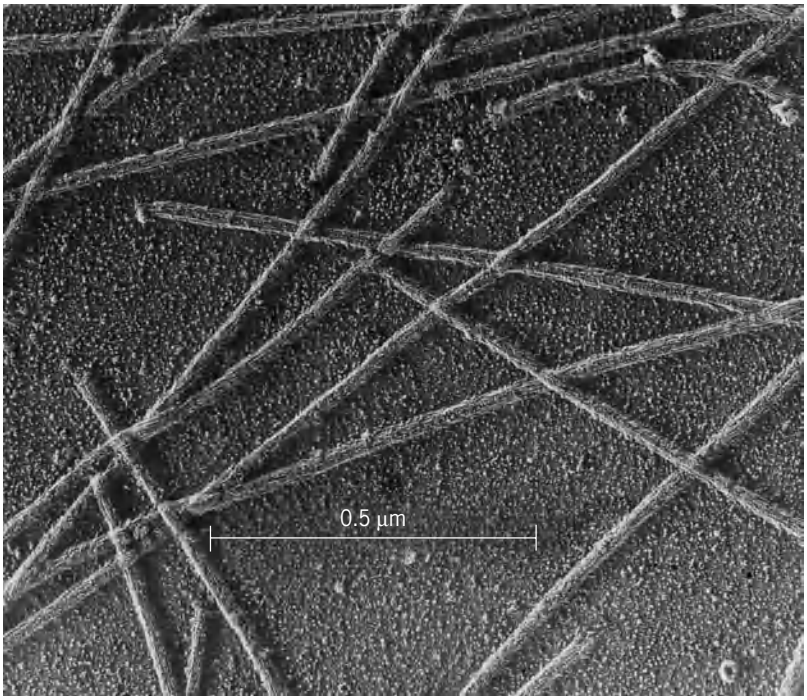

Electron micrograph of microtubules purified from calf brain cytoplasm. Rows of tubulin subunits along the microtubule axis can be discerned by close inspection. (*Courtesy of John Aghajanian*)

Microtubules contain a variety of associated proteins (microtubule associated proteins, or MAPs) that have the general property of enhancing the rate and extent of polymerization. Several of these proteins appear as elongated fibrous projections on the microtubule surface, and may be involved in organizing the microtubules into parallel arrays in the cytoplasm.

A functionally diverse group of proteins affects actin assembly and organization. Profilin and thymosin $\beta 4$ regulate assembly by binding to free actin subunits and controlling their availability for polymerization. Gelsolin and severin are examples of proteins responsible for breaking actin filaments. Other proteins cross-link actin filaments; some, such as filamin and alpha actinin, produce random gels; and others, such as fascin, produce regularly packed bundles.

The organization of myosin may also be controlled in smooth muscle and nonmuscle cells by the activity of a protein kinase, an enzyme that adds a phosphate group to myosin. In the test tube, this reaction switches on the enzymatic activity of the myosin and causes it to form filaments. The protein kinase is under the control of calcium, which also turns on the severing activity of gelsolin and, in nonmuscle cells, inhibits the cross-linking activity of alpha actinin. These combined effects may be important in the solation and mobilization of gelled cytoplasm.

Actin-based cell movement. Actin is involved in a wide variety of movements in many cell types.

Ameboid movement. Ameboid movement occurs in a number of free-living amebas, in slime molds, which are colonial amebas, and in a variety of cells of higher organisms, such as white blood cells. Many of these cells are chemotactic, that is, they move toward or away from food, hormonal stimuli, and other factors. This indicates that, despite the lack of a distinct shape, these cells are capable of coordinated movement. In ameboid cells, the actin and myosin filaments are much less organized than in muscle, and may act as a contractile gel. The actin filaments are randomly arranged, and force production occurs in all directions, unlike muscle where it occurs only along the axis of the muscle fiber. The cytoplasm of the ameboid cell contains relatively rigid as well as fluid regions, which respectively reflect the cross-linked and free states of the actin filaments. Movement of the cell occurs by the extension of pseudopodia (false feet), which appears as the oozing of fluid cytoplasm through the zone of gelled cytoplasm surrounding the cell.

Other forms of locomotion. Other forms of movement also use actin and myosin to provide the motive force. The best characterized of these movements is that shown by mammalian cells, commonly fibroblasts and epithelial cells, grown in culture in plastic petri dishes. Unlike ameboid cells, these cells become very flat, and their cytoplasm shows some evidence of organized actin-containing structures. Most prominent are stress fibers, which are long cables of parallel actin filaments spanning the cells. The stress fibers also contain myosin and other proteins known to interact with actin. This suggests that they may act like myofibrils, the organized arrays of actin and myosin filaments in muscle cells, serving to produce force across the entire cell. However, experiments designed to test this possibility have generally shown that stress fibers do not really serve in force production. They may, instead, act as structural struts or simply represent by-products of the constant reshuffling of actin filaments during movement.

Cultured fibroblasts and epithelial cells, like ameboid cells, show evidence of fluid regions of cytoplasm involved in the forward migration of the cell. At the leading edge of the cell, a lamellipodium of clearer, fluid cytoplasm is usually seen, which extends at relatively rapid rates much like an ameboid pseudopodium. New sites of adhesion to the substratum are produced as the leading edge of the cell moves forward. In time, the trailing portion of the cell is released from the substratum, and contracts to join the leading portion of the cell, resulting in net movement.

Similar lamellipodial extension is observed in other forms of cell movement and may be a very general phenomenon. For example, the migration of sheets of epithelial cells that occurs in wound healing and during embryonic development may occur by a concerted extension of the leading edge of a row of cells. Neuronal cells, which produce extremely elongated axonal and dendritic processes, have a growth cone at their leading edge. The growth cone has the appearance and the actin organization of a leading lamellipodium, and seems to extend by a similar mechanism.

Actin filaments may have a relatively simple organization within lamellipodia. The filaments appear to be attached to the leading edge of the lamellipodium and extend in parallel array into the cytoplasm. These filaments may assemble by intercalation of actin subunits at the membrane. In this case, actin assembly alone, without the need for myosin, could be responsible for movement. Clear evidence that actin assembly can be responsible for movement has come from work with sperm of the sea cucumber. When the sperm contacts an egg, a long acrosomal process composed of actin filaments emerges from the head of the sperm within a few seconds and penetrates the vitelline membrane. There is no myosin in the process. Instead, it is the rapid assembly of actin that acts as the motive force for elongation of the process.

Another form of actin-based intracellular movement is cyclosis or cytoplasmic streaming. It occurs in varying patterns in a variety of plant and algal cells, as well as in the cells of a number of protozoa and the acellular slime molds. In some cases, as in the acellular slime molds, the bulk back-and-forth movement of cytoplasm appears to result from the contraction of gelled actin in the cell cortex. In walled cells, like those of plants and algae, arrays of parallel actin filaments near the cell surface are thought to be involved in propelling the cytoplasm in a continuous, circulating stream. Myosin may be involved in force production, but it is not known how force is exerted on the fluid cytoplasm to induce bulk flow.

Cytokinesis. Cell division consists of two processes: mitosis, which is responsible for chromosome separation, and cytokinesis, which divides the other contents of the cell. Mitosis is a microtubule-based process. Cytokinesis in animal cells involves the formation of a contractile ring of actin filaments. During cell division the ring constricts, as a result of the force-producing action of myosin. Ultimately, the cell is completely separated into two progeny cells. Plant cells contain a rigid cell wall, and cytokinesis occurs by a different mechanism. *See* MITOSIS.

Microtubule-based movement. Microtubules are also involved in a variety of forms of movement.

Ciliary and flagellar movement. Many eukaryotic cells bear cilia or flagella. Unlike the flagella of bacteria, eukaryotic cilia and flagella consist of a bundle of microtubules. The arrangement of microtubules is the same in cilia and flagella, and consists of nine laterally fused pairs of microtubules surrounding a central pair of unfused microtubules. This complex structure, known as the axoneme, whiplashes back and forth at very rapid rates. In the case of free-living cells, including a variety of ciliated and flagellated protozoa and most eukaryotic sperm cells, this movement propels the cell through the surrounding medium. Cilia also coat the surface of many types of epithelial cell layer in eukaryotic organisms (for example, the lining of the lung) and serve to transport the surrounding medium across the cell layer.

The basic mechanism underlying ciliary and flagellar movement is well understood. The enzyme dynein, using ATP as fuel, causes the microtubules to slide against each other. Because other proteins in the cilium or flagellum bind the microtubules together, little net movement of the microtubules relative to each other is possible. Instead, the microtubule bundle bends, much as a bimetallic strip bends in response to changes in temperature. How the elaborate bending movements of the entire structure are coordinated is not yet well understood. *See* CILIA AND FLAGELLA.

Organelle traffic. Many, perhaps all, membrane-bound organelles move about within the cell. This movement occurs in two directions, outward toward the cell membrane and inward toward the nucleus. The outward movement is part of the process of secretion, the export of materials manufactured within the cell. The inward movement represents the processing of materials that have been taken up by endocytosis and pinocytosis, such as peptide hormones that have acted on the cell surface and serum proteins destined for degradation. These movements occur in virtually all cells. In neurons, they are referred to as anterograde (outward) and retrograde (inward) axonal transport, and are responsible for supplying new materials and removing used components from the axon, which has a very limited biosynthetic capability. *See* ENDOCYTOSIS.

The oppositely directed movements of the membrane-bound organelles are thought to be due to the activities of kinesin and dynein, which produce force in opposite directions along microtubules. The tubulin subunits within a microtubule are arranged in a head-to-tail manner, and the microtubule, therefore, has a built-in polarity. The two ends of the microtubule are designated as plus (+) and minus (−). In typical cultured mammalian cells, the plus ends of the microtubules are anchored in the centrosome located on the surface of the nucleus. The plus ends of the microtubules are disposed toward the cell surface. In axons, most of the microtubules are oriented with their minus ends near the nucleus and the plus ends toward the axon terminus where the synapse is located.

Evidence suggests that kinesin and cytoplasmic dynein are attached to the surface of organelles and are responsible for transporting the organelles along microtubules. Based on what is known about the orientation of the microtubules, a simple model for the mechanism of bidirectional organelle traffic is that kinesin is responsible for organelle transport from the nucleus toward the cell surface, or toward the end of the axon, while cytoplasmic dynein is responsible for movement in the opposite, inbound, direction.

Mitosis. All eukaryotic cells reproduce by replication of their deoxyribonucleic acid (DNA) and subsequent separation of the duplicated chromosomes. The process of dividing the genetic material is referred to as karyokinesis or mitosis.

Cell division begins by the condensation of chromatin in the nucleus, followed by the breakdown of the nuclear envelope. The centrosome, the organizing center for microtubules in the interphase cell, duplicates, resulting in two radial arrays of microtubules. The overlapping microtubules interact to form a barrel- or spindle-shaped structure termed the mitotic spindle, also known as the mitotic apparatus. The chromosomes gather at the middle of the spindle, at which time they are found attached to the ends of microtubules. The chromosomal attachment site for microtubules is known as the kinetochore, which is located along a constricted region of the chromosome known as the centromere. *See* CELL DIVISION.

Chromosome movement generally occurs in two stages referred to as anaphase A and B. During anaphase A, the chromosomes move toward the poles of the spindle. During anaphase B, the entire spindle elongates, causing further separation of the chromosomes. During both processes, different classes of microtubules within the spindle shorten or lengthen. It is not known whether this behavior is a cause or a consequence of chromosome movement. Kinesin-related proteins and cytoplasmic dynein are involved in chromosome movement. CHO-1, a kinesin-related protein, has been found associated with microtubules in the middle of the mitotic spindle. It is believed to produce spindle elongation by causing microtubules to slide against one another. Other kinesin-related proteins have been found at the kinetochore, as has cytoplasmic dynein, and may pull on the ends of microtubules at their point of attachment to the kinetochore and draw the chromosome toward the spindle pole. *See* CHROMOSOME.

Other motile proteins. It appears that many forms of movement can be explained by actin and myosin or by microtubules and their associated proteins dynein and kinesin. However, other motile mechanisms certainly exist.

Many bacteria possess flagella. These are very fine helical hairs, unlike the more substantial flagella and cilia of eukaryotic cells. Another important difference is that the bacterial flagella rotate about their axis and propel the bacterium by a corkscrewlike mechanism, unlike the bending and whiplashing movements of eukaryotic cilia and flagella. Bacterial flagella are hollow filamentous polymers, like microtubules, but are composed of the protein flagellin, which has no apparent relationship to tubulin. The flagella are attached at their base to a rotary motor structure in the bacterial cell membrane. *See* BACTERIA.

Other forms of bacteria glide over solid substrata by using an excreted slime for propulsion; the mechanism of gliding is not understood. Gliding motility is also seen in a number of algae and blue-green algae.

The sperm cells of roundworms differ from other types of sperm cells in that they lack flagella and exhibit a form of ameboid movement. These cells, however, contain neither actin, which is involved in ameboid movement in other ameboid cells, nor tubulin. Movement may be produced by insertion of

lipid in the forward region of the plasma membrane and rearward flow of the membrane.

A contractile protein, spasmin, has been identified in *Vorticella* and related ciliated protozoa. Spasmin is organized into a long, thick fiber within the stalk portion of the organism. In response to calcium, the fiber undergoes a rapid, drastic contraction. It is not known whether spasmin exists in other organisms. *See* CELL (BIOLOGY).

Relationship to disease. Understanding how cells move increases the ability to control abnormal cell behavior, such as the increased level of cell division responsible for cancer and the migration of cancer cells from their site of origin in the body. Errors in chromosome segregation are also known to be responsible for Down syndrome, and are prevalent during the progression of neoplastic tumors.

Because the normal functioning of cells is so dependent on proteins that compose and regulate microtubules and actin filaments, defects in these proteins are expected to have severe effects on cell viability. An example of a microtubule defect has been identified in Alzheimer's disease: a microtubule-associated protein (termed tau) is found to be a prominent component of abnormal neurofibrillary tangles seen in affected nerve cells. However, it remains unknown whether the defect involving tau is part of the cause of the disease or represents one of its effects. *See* ALZHEIMER'S DISEASE; CANCER (MEDICINE); DOWN SYNDROME. Richard Vallee

Bibliography. M. De Brabander and J. G. De Mey (eds.), *Microtubules and Microtubule Inhibitors: Nineteen Eighty-Five*, 1987; J. A. De Grado, *Microtubule Proteins*, 1989; P. Dustin, *Microtubules*, 1984.

Cell nucleus

The largest of the membrane-bounded organelles which characterize eukaryotic cells; it is thought of as the control center since it contains the bulk of the cell's genetic information in the form of deoxyribonucleic acid (DNA). The nucleus has two major functions: (1) It is the site of synthesis of ribonucleic acid (RNA), which in turn directs the formation of the protein molecules on which all life depends; and (2) in any cell preparing for division, the nucleus precisely duplicates its DNA for later distribution to cell progeny. The discovery of the nucleus dates back to 1710, when the Dutch microscopist Antonie van Leeuwenhoek noted a centrally located "clear" area in living blood cells of birds and amphibians. However, it was not until 1831 that the British botanist Robert Brown first used the term nucleus and provided a precise morphological description. *See* DEOXYRIBONUCLEIC ACID (DNA); EUKARYOTAE; PROKARYOTAE.

Structure. The diameter of nuclei ranges from 1 micrometer in intracellular parasites and yeast cells to several millimeters in some insect sperm. Spherical or ellipsoidal nuclei are found in most cell types, although occasionally spindle-shaped, lobulated, disc-shaped, or cup-shaped nuclei may be observed. Although nuclear size and shape are somewhat consistent features of a particular cell type, these features are more variable in cancer cells. In addition, tumor cell nuclei are characterized by indentation, furrowing, elongation, and budding.

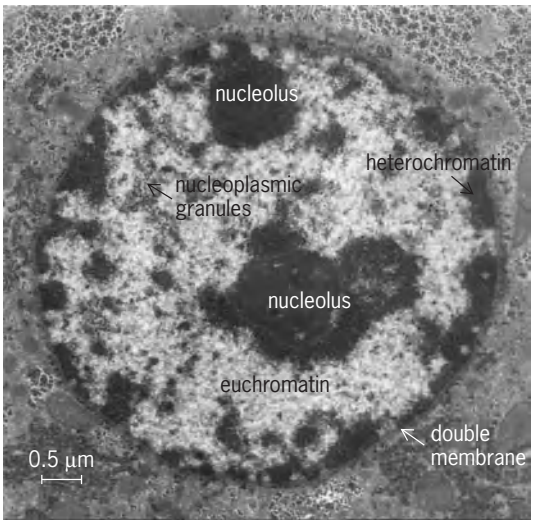

Fig. 1. Transmission electron micrograph of a thin section of a rat liver cell nucleus.

The nucleus is bounded by a double membrane and contains several major components: chromatin, which is composed of DNA and chromosomal proteins; the nucleolus, which is the site of ribosomal RNA (rRNA) synthesis; and nucleoplasmic fibrils and granules, some of which are involved in the processing and transport of messenger RNA out of the nucleus (**Fig. 1**). The constituents of the nucleus are contained within a framework referred to as the nuclear matrix.

Nuclear envelope. The contents of the nucleus are spatially separated from the cytoplasm by a double-membrane-bounded structure called the nuclear envelope. The two membranes are both approximately 10 nanometers thick, and they are separated by a gap of 20–40 nm, known as the perinuclear space. The outer nuclear membrane is continuous with the endoplasmic reticulum and is often studded with ribosomes involved in protein synthesis.

At periodic intervals, the inner and outer membranes appear to be fused together, forming nuclear pores, averaging 120 nm in diameter. A typical mammalian cell contains 3000–4000 pores in its nuclear envelope. Each pore is composed of a nuclear pore complex measuring approximately 80 nm (inside diameter) with an estimated 9-nm open channel. The mass of each mammalian nuclear pore complex has been estimated at about 125×10^6 daltons and is composed of approximately 50 different proteins. The pores allow for the passage of molecules into and out of the nucleus. All nuclear pores appear to be functionally identical in that they can carry out both import and export. The central channel of the nuclear pore complex is thought to be involved in

active transport, and eight peripheral channels, each approximately 10 nm in diameter, are proposed to function in passive exchange of small molecules as well as perhaps serving as an anchoring site for interphase chromatin. Small macromolecules below 50 kDa can diffuse through the nuclear pore complex, whereas larger macromolecules are transported via a signal-mediated process that may occur via a series of association-dissociation reactions. Targeting signals that are encoded in a protein and direct its nuclear import are called nuclear localization signals (NLS), and signals that direct nuclear export are termed nuclear export signals (NES); many different signals have been identified that function in association with different proteins. Recently, a nuclear export signal has been identified that directly interacts with transfer RNA. The largest-diameter complex that can be transported through the nuclear pore complex is estimated to be about 0.26 μm. *See* RIBOSOMES.

Internal and adjacent to the inner membrane of the nuclear envelope is an electron opaque layer termed the nuclear lamina. In most mammalian cells, the lamina is about 10 nm thick and is composed of three major polypeptides, lamins A (70 kilodaltons), B (68 kDa), and C (60 kDa), which are present in equal amounts in the interphase nucleus. Lamins possess all the major structural properties that have been described for intermediate filament polypeptides at the levels both of protomers and of assembled filaments. The lamins are thought to function in regulating nuclear envelope structure and anchoring interphase chromatin at the nuclear periphery. It has also been suggested that the lamins act as a substrate for the formation of DNA replication complexes, or regulate DNA condensation or formation of nuclear matrix elements which are required for efficient replication to occur. *See* CELL MEMBRANES.

Chromatin and chromosomes. Deoxyribonucleic acid is present in the nucleus as a DNA-protein complex referred to as chromatin. The DNA that is present in a typical human cell is composed of 3 billion base pairs of nucleotide sequence. The DNA is highly compacted, approximately 200,000-fold. If it were stretched out in a linear manner, it would extend for 6.6 ft (2 m). This DNA encodes for the estimated 50–100,000 genes which are present in each human cell. However, a significant portion of the genome does not contain any protein coding information, and it is unclear what role this DNA plays in cellular function.

Classes of chromatin. Eukaryotic chromatin has been divided into two classes, heterochromatin and euchromatin, based on its state of condensation during interphase. Heterochromatin is condensed during interphase and is therefore generally considered to be transcriptionally inactive. It is commonly located in an irregular band around the nuclear periphery (Fig. 1) and around the nucleolus as well as in patches throughout the nucleoplasm. The amount of heterochromatin present in the nucleus varies with transcriptional activity; little heterochromatin is present in transcriptionally active cancer cells whereas the nuclei of orthochromatic erythroblasts and mature spermatozoa, both transcriptionally inactive, are practically filled with condensed chromatin. In an average eukaryotic cell, approximately 90% of the chromatin is transcriptionally inactive at any given time, but all of the inactive chromatin may not be in a condensed state. However, the 10% of the chromatin that is transcriptionally active is always in the form of decondensed euchromatin.

Transcription. The process by which a segment of cellular DNA makes a copy of itself in the form of RNA is referred to as transcription. Transcription can be divided into three systems, each using a different set of enzymes and some similar as well as different factors which function to produce different classes of RNA molecules. The three systems are referred to as RNA polymerase I, II, and III. RNA polymerase I is specifically involved in transcribing ribosomal RNAs, which are components of ribosomes and are involved in the process of protein synthesis. RNA polymerase II is primarily involved in transcribing messenger RNAs which act as the template for the production of specific cellular proteins. A majority of the RNAs that are made by RNA polymerase II must be processed prior to their transport from the nucleus into the cytoplasm. This processing, referred to as splicing, involves the removal of noncoding regions (introns) which interrupt the protein coding sequence (exons) present in the transcribed RNA molecule. RNA polymerase III is primarily involved in the production of transfer RNAs, which are involved in directing specific amino acids, the building blocks of proteins, to the ribosomes, where the amino acids become incorporated in a sequence-dependent manner into the newly synthesized protein. The production of RNA molecules may be thought of as the major function of the nucleus, and the process is highly regulated, with some RNAs being transcribed at specific times during the cell cycle. *See* EXON; INTRON; RIBONUCLEIC ACID (RNA).

Replication. A second nuclear function equally important to transcription is the process of DNA replication. Because many cells in the human body are continuously dividing, it is essential to their function for each of the daughter cells to obtain a complete copy of the cell's DNA. Replication of DNA takes place only once during the cell cycle in a period referred to as the synthesis phase or S phase. As is the case for transcription, a specific series of enzymes and factors are responsible for the accurate replication of cellular DNA. *See* CELL CYCLE; CHROMOSOME.

Nucleolus. The nucleolus is a subnuclear organelle within which ribosomal genes and their products are naturally sequestered from the rest of the genome and nucleoplasm. Ribosomal gene transcription, ribosomal RNA processing, and preribosomal particle formation all occur within this highly specialized region of the nucleus. Nuclei may contain one or more nucleoli, and generally cancer cells have multiple nucleoli. Although the nucleolus is not bounded by a membrane, specific proteins have been localized to discrete functional regions within this structure and a nucleolar localization signal has been identified. Ultrastructurally, the nucleolus may be regarded as

being composed of four organizational areas: a dense fibrillar region, fibrillar center or centers, a granular region, and nucleolar vacuoles. The nucleolus is usually surrounded by a shell of perinucleolar chromatin which is connected with regions of intranucleolar chromatin and is continuous with the fibrillar centers.

The cellular DNA that encodes for ribosomal RNA of higher eukaryotes occurs in multiple copies and is organized in tandem repeats which are separated from each other by nontranscribed spacer regions of DNA. Human cells contain approximately 250 copies of ribosomal DNA, and the length of a single ribosomal DNA repeat unit is 44,000 base pairs. The periphery of the nucleolar fibrillar centers are the sites that most likely represent the nucleolar regions where ribosomal DNA transcription occurs. The dense fibrillar component probably represents elongating RNA transcripts and precursor molecules. Finally, the granular region represents the site of assembling and processing mature ribosomal subunits.

Nuclear matrix. The nuclear components are present within specific macromolecular nuclear domains (**Fig. 2**). The large number of molecules present within the nucleoplasm and the high degree of efficiency with which vital nuclear functions take place suggest a high degree of nuclear organization. The concept of a nuclear matrix or nuclear scaffold seems to satisfy this role. By definition, the nuclear matrix is the residual structure left after the majority of DNA, nuclear proteins, and nuclear phospholipid are extracted from the nucleus. Numerous nuclear functions, including transcription, RNA processing, and DNA replication, are thought to take place in association with the nuclear matrix. *See* CELL (BIOLOGY); MOLECULAR BIOLOGY.

David L. Spector

Bibliography. P. S. Agutter, *Between Nucleus and Cytoplasm*, 1990; B. Alberts et al. (eds.), *Molecular Biology of the Cell*, 3d ed., 1994; H. Hillman, *The Living Cell*, 1991; D. L. Spector, Macromolecular domains within the cell nucleus, *Annu. Rev. Cell Biol.*, 9:265–315, 1993.

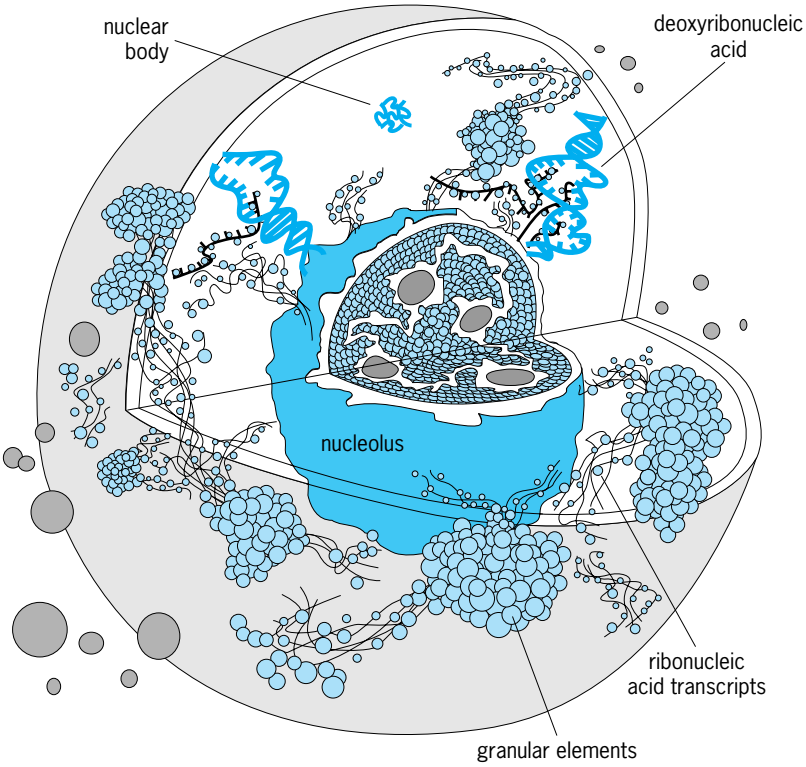

Fig. 2. Three-dimensional organization of the mammalian cell nucleus.

Cell organization

Cells are divided into several compartments, each with a characteristic structure, biochemical composition, and function (**Fig. 1**). These compartments are called organelles. They are delimited by membranes composed of phospholipid bilayers and various proteins specialized for each type of organelle. All eukaryotic cells have a nucleus surrounded by a nuclear envelope, and a plasma membrane that borders the whole cell. Most eukaryotic cells also have endoplasmic reticulum, a Golgi apparatus, lysosomes, mitochondria, and peroxisomes. Plant cells have chloroplasts for photosynthesis in addition to the organelles that they share with animal cells. These organelles are suspended in a gel-like cytoplasmic matrix composed of three types of protein polymers called actin filaments, microtubules, and intermediate filaments. In addition to holding the cell together, the actin filaments and microtubules act as tracks for several different types of motor proteins that are responsible for cell motility and organelle movements within the cytoplasm.

A major challenge in the field of cell biology is to learn how each organelle and the cytoplasmic matrix are assembled and distributed in the cytoplasm. This process is complex, since cells consist of more than 2000 different types of protein molecules together with a large number of lipids, polysaccharides, and nucleic acids, including both deoxyribonucleic acid (DNA) and many different types of ribonucleic acid (RNA). *See* NUCLEIC ACID.

The cell must possess enough information to specify which molecules are to be associated in a specific compartment, to route the appropriate groups of molecules to their compartments, and then to position each type of component appropriately in the cell. As a result of intense research on each of these topics beginning in the 1960s and continuing to the present, a number of specific chemical reactions that contribute to organizing cells are now recognized; even more important, a small number of general principles that explain these complex processes of life can be appreciated.

Principles of assembly. The following principles account for many of the processes that result in the assembly and organization of cells.

Normal cells regulate the production and degradation of all of their constituent molecules so that the right balance of molecules is present at any given time. The genes stored in nuclear DNA are duplicated precisely once per cell cycle. The supply of each of thousands of proteins is usually regulated at the level

of the genes at the time of biosynthesis and by the rate of degradation. These proteins serve as enzymes that determine the rate of synthesis of themselves as well as of other cellular components such as nucleic acids, carbohydrates, and lipids. Each of these processes is regulated by molecular feedback loops to assure the proper levels of each cellular constituent. *See* CELL CYCLE; CELL METABOLISM.

A large majority of cellular components are generated by the self-assembly of their constituent molecules. Self-assembly means that the information required for molecules to bind together in the proper orientation is contained in the molecules themselves. In other words, no outside information is required. Although the amino acid sequence of each protein not only specifies its three dimensional structure and its interactions along assembly pathways, cells often use proteins called chaperones to guide the process, particularly to avoid incorrect folding events that might lead to irreversible aggregation.

Some examples of self-assembly include the binding of histones to DNA, the formation of lipid bilayers from phospholipids, and the polymerization of actin molecules into filaments. The molecules are usually brought together by diffusion. The energy required to hold them together derives from the exclusion of water from their complementary surfaces as well as from the formation of ionic bonds and hydrogen bonds. The variety of molecular structures found in proteins allows each type to self-assemble specifically, only with their correct partner molecules.

After biosynthesis, proteins and nucleic acids are routed to their proper cellular compartment by specific recognition signals consisting of parts of the molecule or, in the case of some proteins, by sugar side chains. These signals are recognized by compartment-specific receptors that guide the molecules to the correct compartment. For example, proteins destined for lysosomes have a specific sugar side chain added in the Golgi apparatus that guides them to lysosomes. Similarly, proteins destined for the nucleus all contain short sequences of amino acids that target the proteins for uptake by the nucleus. These so-called nuclear recognition sequences bind to receptor proteins that move through nuclear pores, channels through the nuclear envelope that connect the nucleus with the cytoplasm.

Most molecules move to their correct compartment by the process of diffusion down concentration gradients, but objects as large as organelles require transport systems composed of microtubules or actin filaments together with specific motor proteins to position them correctly in the cytoplasm. For example, a protein molecule destined to be part of a mitochondrion will diffuse from the site of biosynthesis through the cytoplasm to a mitochondrion, where it will bind to a receptor that guides its incorporation into the mitochondrion. On the other hand, the mitochondrion itself is too large to diffuse through the network of protein fibers in the cytoplasmic matrix, and so it must be pulled through the matrix by a motor protein that moves along microtubules to the correct place in the cell.

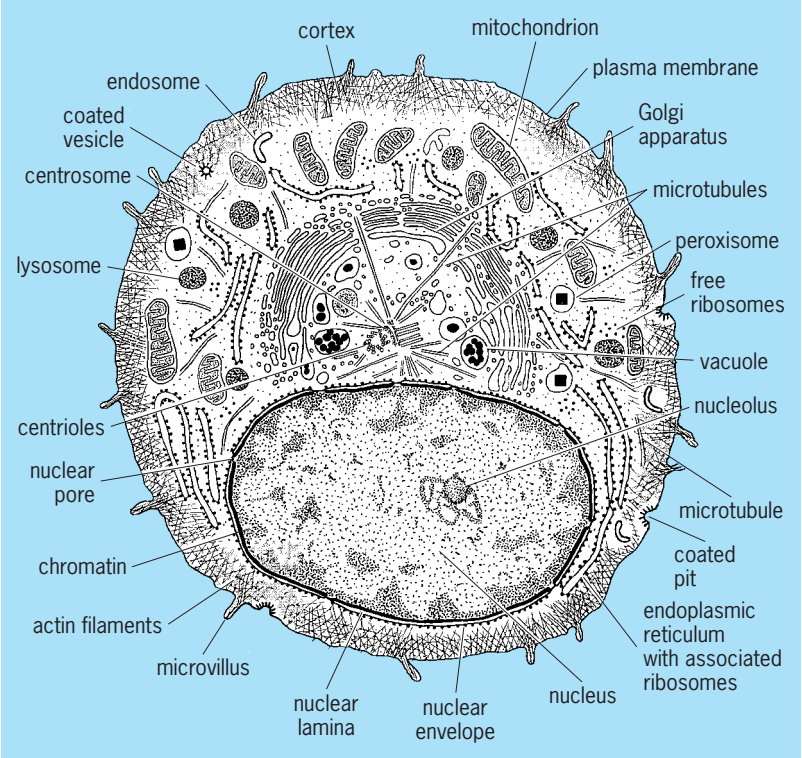

Fig. 1. Section through an animal cell showing the major components visible by electron microscopy. To simplify, a few important components, including intermediate filaments, have been omitted.

Some cellular components, such as ribosomes and the filaments of the cytoplasmic matrix, assemble afresh from their constituent molecules, but all organelles composed of membranes form only by the growth and division of preexisting organelles. The reason that these organelles require precursors is that biological membranes composed of phospholipids can grow only by expansion of preexisting bilayers. As a consequence, organelles such as mitochondria and the endoplasmic reticulum are inherited maternally starting from the egg and expanding into every cell in the body by continued growth and partition into both daughter cells at every cell division. Other membrane-bound organelles, such as lysosomes, form from vesicles that bud off the Golgi apparatus.

Many of the individual self-assembly reactions required for normal cell growth can be reproduced in the test tube, but the cell is the only place known where the entire range of reactions, including biosynthesis, targeting, and assembly, can go to completion. Thus, cells are such a special environment that the chain of life has required an unbroken lineage of cells stretching from all living cells back through their ancestors to the earliest forms of life. This requirement for cellular continuity explains why extinction is an irreversible process. *See* EXTINCTION (BIOLOGY).

Organization and assembly of organelles. All cellular organelles arise by the assembly of their components, either from constituent molecules or from preexisting structures.

Nucleus. The nucleus is a protected environment within the cell where the genes are stored in the form of specific sequences of bases in large DNA molecules. Many genes are linked together end-to-end in one giant DNA molecule. Interaction with proteins called histones coils up each of these linear DNA molecules into more compact units called chromosomes, which can fit inside the nucleus. The DNA is replicated once during each cell cycle and supplied with sufficient histones and other accessory proteins to duplicate each chromosome. Between mitotic events, the chromosomes are not distinct from each other and are called chromatin (Fig. 1). *See* CELL NUCLEUS; CHROMOSOME; GENE.

The nucleus is separated from the cytoplasm by the nuclear envelope, a specialized part of the endoplasmic reticulum consisting of two membranes and penetrated by tiny diaphragms called nuclear pores. All traffic into and out of the nucleus passes through nuclear pores. Traffic moving into the nucleus includes all of the proteins found in the nucleus and the proteins required to form ribosomes. Traffic moving out includes messenger RNA and ribosomal subunits assembled in the nucleolus. The integrity of the nuclear envelope is maintained by a network of intermediate filaments composed of protein subunits called lamins (Fig. 1) that are associated with the inner surface of the envelope. Lamins form links between the nuclear pores and between the membrane and the chromatin.

During mitosis, the replicated chromosomes must partition equally between the two daughter cells, a process that requires the disassembly and reassembly of the nuclear envelope. At the outset of mitosis, phosphate groups are added to the nuclear lamins by an enzyme called a mitotic kinase. This modification causes lamin intermediate filaments to break up into their protein subunits. The nuclear envelope membrane then fragments into small vesicles, releasing the chromosomes into the cytoplasm, where the mitotic apparatus divides them between the daughter cells. At the end of mitosis, the chromosomes cluster so close to each other that they exclude soluble macromolecules. The lamins lose their extra phosphates, and together with associated vesicles of nuclear envelope membranes they repolymerize on the surface of the condensed chromosomes. When these vesicles fuse to form a continuous nuclear envelope, virtually no cytoplasmic macromolecules are included in the nucleus. Remarkably, the entire cycle of chromosome condensation, nuclear envelope breakdown, and nuclear envelope reformation can now be carried out in the test tube. A single process, the phosphorylation and dephosphorylation of the lamins, provides the biochemical cue for this complex process. *See* MITOSIS.

Protein synthesis by ribosomes. The sequence of nucleotides in each gene determines the sequence of nucleotides in messenger RNAs and other RNAs. The nucleotide sequence of a messenger RNA specifies the sequence of amino acids in a protein as it is synthesized by RNA-protein assembly called a ribosome. Ribosomes are giant macromolecular assemblies consisting of one large and one small subunit that are assembled in the nucleolus from more than 50 different proteins and 3 RNA molecules. As the ribosome reads the messenger RNA, transfer RNAs deliver the correct amino acid to the end of the growing polypeptide. The RNA component of the large ribosome subunit catalyzes the formation of the peptide bond between each amino acid in the protein. *See* PROTEIN; RIBOSOMES.

Endoplasmic reticulum. The endoplasmic reticulum is an extensive, membrane-bounded intracellular compartment specialized for the synthesis of lipids and proteins. The membrane forms by expansion of preexisting endoplasmic reticulum. Proteins inserted into the lipid bilayer itself synthesize the phospholipids that constitute the lipid bilayer. The proteins destined for insertion into membranes, or for lysosomes, or for secretion carry signal sequences of amino acids. As ribosomes synthesize these proteins, the signal sequence interacts with an adapter called a signal recognition particle, which links the ribosome producing the protein to a receptor protein of the endoplasmic reticulum. As the protein grows in length, a protein pore guides the protein into or across the lipid bilayer of the endoplasmic reticulum. Other zip codes retain particular proteins in the lipid bilayer or in the lumen of the endoplasmic reticulum; other zip codes direct proteins on to various parts of the cell. *See* ENDOPLASMIC RETICULUM.

Microtubules play an important role in distributing the endoplasmic reticulum within the cell. Motor proteins bind to the membrane and by moving along microtubules can pull the endoplasmic reticulum into a branching network that spreads throughout the cytoplasm.

Golgi apparatus. These stacks of flattened membrane-bounded sacks (Fig. 1) are responsible for adding sugars to proteins and for sorting membranes and secretory proteins. Complexes of cytoplasmic proteins interact with membrane proteins to pinch vesicles from the endoplasmic reticulum. These vesicles fuse with the Golgi apparatus, transferring lipids, proteins in the lumen, and membrane proteins from the endoplasmic reticulum to the Golgi apparatus. Proteins and lipids pass selectively through the compartments of the Golgi apparatus where they are modified by the covalent attachment of sugar polymers. The Golgi apparatus is characteristically located in the middle of the cell near the centrosome by the action of motor proteins that move the membranes along microtubules toward the centrosome. *See* GOLGI APPARATUS.

Plasma membrane. The components of the cell membrane (Fig. 1) are originally formed in the endoplasmic reticulum, where both the phospholipids and proteins are synthesized. The proteins are inserted into the bilayer and then transported from the endoplasmic reticulum through the Golgi apparatus via small vesicles to the plasma membrane. Remarkably the signals and receptors along this pathway are selective enough to allow plasma membrane proteins to move to their target, with the total exclusion of

proteins that belong in the other membrane compartments.

Many of the components of the plasma membrane are not permanent residents there; instead, proteins that serve as receptors for extracellular molecules, including nutrients and some hormones, are taken into the cell by specializations called coated pits (Fig. 1) and can recycle in small vesicles from the plasma membrane to internal compartments called endosomes and back to the cell surface from one to hundreds of times before they are degraded. Much larger particles, such as bacteria, are taken in by the related process of phagocytosis and also delivered to lysosomes for degradation. *See* CELL MEMBRANES; ENDOCYTOSIS; PHAGOCYTOSIS.

Lysosomes. The degradative enzymes found in lysosomes are synthesized by ribosomes attached to the endoplasmic reticulum so that they pass into the lumen of the endoplasmic reticulum. From there they pass to the Golgi apparatus, where an enzyme recognizes surface features and add a special sugar called mannose-6-phosphate. This sugar is the signal for protein receptors to bind lysosomal proteins and to guide them from the Golgi apparatus to the lumen of lysosomes, as cargo in small vesicles that bud from the Golgi apparatus and fuse with forming lysosomes. *See* LYSOSOME.

Mitochondria. Mitochondria arose in an early eukaryote that engulfed a proteobacterium. The two cells established a symbiotic relationship with the proteobacterium providing enzymes to derive energy from the breakdown of lipids and the capacity for synthesis of ATP by oxidative phosphorylation. Over tens of millions of years, most of the proteobacterial genes moved to the nucleus of the host, so now nuclear genes encode most mitochondrial proteins and RNAs. Mitochondria retained a small genome, ribosomes, and messenger RNAs that produce a few special proteins.

Although mitochondria are membrane-bound organelles (Fig. 1), they form in a fundamentally different way from the other organelles since the constituent proteins do not come through the endoplasmic reticulum-Golgi pathway. Most of the proteins are synthesized on free ribosomes located in the cytoplasm. From there they diffuse to the mitochondria. Mitochondrial proteins have an identifying sequence of amino acids that reacts with receptors on the surface of mitochondria. Subsequently they are actively transported across the mitochondrial membranes into the lumen or are inserted into the membranes. A minority of the mitochondrial proteins are synthesized inside the mitochondria themselves. *See* MITOCHONDRIA.

Peroxisomes. As with mitochondria, the constituent proteins are synthesized on free ribosomes and then incorporated into the membranes and lumen of the peroxisome. Enzymes in the lumen of peroxisomes catalyze a variety of oxidation reactions. *See* PEROXISOME.

Chloroplasts. Chloroplasts arose during a singular event that resulted in a cyanobacterium becoming a symbiont in a eukaryotic cell that founded the lineages including algae and plants. The cyanobacterium brought the capacity for photosynthesis by providing the proteins required for both photosystem I and photosystem II. As in the case of mitochondria, over time most bacterial genes moved to the nucleus but chloroplasts retained a small genome and the capacity to make a number of their own proteins. Most of the proteins are made in the cytoplasm and imported as in mitochondria and peroxisomes. *See* CELL PLASTIDS; PHOTOSYNTHESIS.

Cytoskeleton. Assembly of the cytoskeleton represents a somewhat different problem for the cell than any of the examples considered above since the three types of protein polymers making up the cytoskeleton change their form relatively rapidly during the life of the cell. Consequently, information is required not only to guide assembly, but also to remodel these systems on a time scale of seconds to minutes. Since the actin filaments, intermediate filaments, and microtubules are composed of different proteins and each has distinctive functions and dynamics, they will be considered separately. *See* CYTOSKELETON.

1. *Actin filaments.* Actin filaments (**Fig. 2***a*) occur in all eukaryotic cells, being a fundamental structural component as well as linear tracks for movements of myosin motors. In muscle, actin filaments and myosin filaments form a highly specialized contractile system designed for forceful, rapid, one-dimensional contractions. In other cells, actin filaments and myosins are responsible for pinching daughter cells in two during cytokinesis, for some cell shape changes and for movements of some membrane bound organelles (particularly in plants and fungi). *See* CYTOKINESIS; MUSCLE PROTEINS.

Actin is a globular protein that polymerizes into long filaments composed of a double-helical array of actin subunits. Formation of these filaments is a classical example of self-assembly, but the process is tightly controlled in cells by more than 50 regulatory proteins. Some actin-binding proteins bind actin monomers and influence the number available for polymerization. Others bind to an end of actin

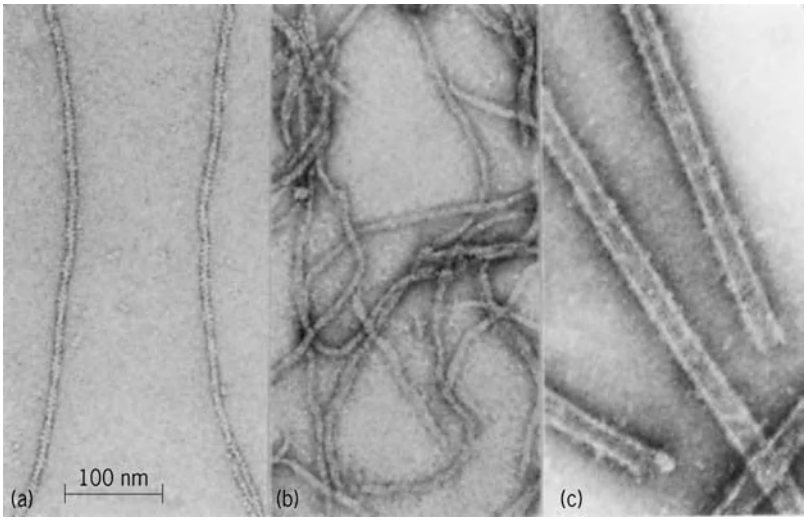

Fig. 2. Electron micrographs of the three cytoskeletal polymers: (*a*) actin filaments, (*b*) intermediate filaments, and (*c*) microtubules.

filaments. These can help initiate polymerization but allow growth in only one direction. Some can cut actin filaments into two fragments. Others cross-link actin filaments into random networks or parallel bundles. These networks form relatively rigid gels in the cortical cytoplasm of cells (Fig. 1). Actin filaments anchored to the plasma membrane reinforce the surface of the cell. Tightly packed bundles of actin filaments with closely applied plasma membrane project from the surface of many cells (like fingers in a glove). These filopodia or microvilli (Fig. 1) increase the surface area of the plasma membrane and participate in sensing the environment. Some cells have bundles of actin filaments attached at their ends to the inside of the plasma membrane. These "stress fibers" reinforce cells that are under mechanical stress such as the cells that line large arteries.

This actin filament system is dynamic, as shown by the way that a cell can change shape as it moves. Such movements require the reorganization of both the cortical network and actin filament bundles. Some of these changes are caused by myosin-mediated contractions such as the contraction of the cleavage furrow during cytokinesis. Other changes, such as the extension of the cortical cytoplasm to form a new pseudopod, depend on the assembly of new actin filaments. External signals such as chemotactic molecules can influence both actin filament organization and the direction of motility. *See* CELL MOTILITY.

About 20 families of myosin molecules are responsible for actin-based movements. In every case the energy for the movements is provided by the hydrolysis of adenosine triphosphate (ATP). Myosin II, the type originally found in muscle, forms filaments that can pull two actin filaments past each other. Other myosins interact with membranes, which they transport along actin filaments. Several types of regulation are known, including phosphorylation of the myosin and calcium-sensitive regulatory proteins bound to the actin filaments. *See* MUSCLE.

2. *Intermediate filaments.* These filaments (Fig. 2*b*) act as intracellular tendons in higher eukaryotic cells that are subjected to substantial physical stresses such as muscle cells and epithelial cells like skin. The various types of intermediate filaments are composed of a family of protein subunits all related to the keratin molecules found in hair. Some intermediate filaments are attached to the plasma membrane at specializations called desmosomes, while others surround the nucleus. Since intermediate filaments are inextensible (try stretching a hair), they help the cell from being over stretched. Intermediate filaments rearrange during mitosis and cell movements. Mutations in the genes for the intermediate filaments cause a variety of human diseases. Defects in keratin cause blistering of skin, while mutations in the proteins of the nuclear lamina can even cause premature aging.

3. *Microtubules.* Microtubules (Fig. 2*c*) are rigid polymers that form a number of structures, including the mitotic apparatus, an interphase cytoplasmic array radiating from the centrosome (Fig. 1), and the supporting axoneme of cilia and flagella. Microtubules get their name from the cylindrical arrangement of tubulin subunits in the polymer. Their rigidity make microtubules the only cytoskeletal polymer that can resist compressive forces. Microtubules provide the internal support for asymmetrical cell processes (such as cilia) and to maintain asymmetrical cell shapes. A simplifying feature of microtubule organization in cells is that virtually all microtubules have the same polarity relative to the centrosome; the rapidly growing end is peripheral. As with actin filament polymerization, microtubule polymerization is a self-assembly process modulated by microtubule-associated proteins. Some of these proteins stabilize or destabilize microtubules. Some associate with the growing end of microtubules. Others cross-link microtubules to intermediate filaments and actin filaments, unifying the cytoskeleton into a single mechanical structure. *See* CILIA AND FLAGELLA.

Microtubules are tracks for rapid movements of membrane-bound vesicles and larger organelles, including mitochondria and the endoplasmic reticulum. In this way the microtubule system is an important determinant of cellular organization. These movements are powered by motor proteins that hydrolyze ATP to provide the energy for the motion. A family of kinesin motors generally move particles toward the peripheral end of the microtubules, while another family of motors called dynein move particles toward the cell center. Together these motors form a two-way transport system in the cell that is particularly well developed in the axons and dendrites of nerve cells, where vesicles can move more than 1 m in a few days' time.

The site and direction of microtubule assembly is determined by the centrosome (or by basal bodies in the case of cilia and flagella). Remarkably, microtubules radiating into the cytoplasm are not particularly stable, but individually undergo spontaneous fluctuations in length by the alternate loss and reassembly of thousands of tubulin subunits at their peripheral ends. These fluctuations in length called dynamic instability are important in the formation of the mitotic spindle and allow microtubules to probe all parts of the cytoplasm. *See* CELL (BIOLOGY); CYTOPLASM.

Thomas Pollard

Bibliography. B. Alberts et al., *Molecular Biology of the Cell*, 4d ed., 2002; H. Lodish et al., *Molecular Cell Biology*, 5th ed., 2004; T. Pollard and W. Earnshaw, *Cell Biology*, 2002.

Cell plastids

Specialized structures found in plant cells, diverse in distribution, size, shape, composition, structure, function, and mode of development.

Types. A number of different types is recognized. Chloroplasts occur in the green parts of plants and are responsible for the green coloration, for they contain the chlorophyll pigments. These pigments, along with certain others, absorb the light

energy that drives the processes of photosynthesis, by which sugars, starch, and other organic materials are synthesized. Amyloplasts, nearly or entirely colorless, are packed with starch grains and occur in cells of storage tissue. Proteoplasts are less common and contain crystalline, fibrillar, or amorphous masses of protein, sometimes along with starch grains. In chromoplasts the green pigment is masked or replaced by others, notably carotenoids, as in the cells of carrot roots and many flowers and fruits. *See* CAROTENOID; CHLOROPHYLL; STARCH.

Occurrence. Plastids lie in the cytoplasm of the cell (**Fig. 1**), normally unattached to other structures but surrounded by a special extension of the nuclear envelope in some algae. There may be from a single plastid to many thousand plastids per cell. A mesophyll cell in a leaf contains up to about 50 chloroplasts. Some specialized cells lose their plastids at maturity. Plastids do not occur in the fungi.

Size and shape. The greatest diversity of chloroplast size and shape is found among genera and species of algae, as exemplified by branched and unbranched ribbons 100 micrometers or more in length, flat plates, and cup-shaped or stellate bodies. In some algae and in most other plants, chloroplasts are spherical or lens- or disk-shaped, with the major axis measuring 4–7 μm. *See* ALGAE.

Composition. A chloroplast in a spinach leaf has a dry weight of about 2×10^{-11} g, of which some 70% is protein (about half of this is soluble and half is membrane-bound), slightly more than 20% is lipid, and up to 7% is nucleic acid. Inorganic constituents include calcium, magnesium (largely in chlorophyll molecules), iron (partly in the proteins ferredoxin and cytochrome), manganese, and copper. Of the lipid approximately 21% is chlorophyll (100), 3% is carotenoid (22), 3% is plastoquinone and other quinonoid compounds (22), 9% is phospholipid (50), 44% is glycolipid (mainly mono- and di-galactosyl glycerides) (23), 2% is sterol (25), 0.5% is prenol (3), and 17% is as yet unidentified. Figures in parentheses are estimates of the numbers of molecules relative to 100 molecules of chlorophyll.

Structure. All types of plastids have one structural feature in common, a double envelope consisting of two concentric sheets of membrane. The outer of these is in contact with the cytoplasmic ground substance; the inner with the plastid matrix, or stroma. They are separated by a narrow space of about 10 nanometers.

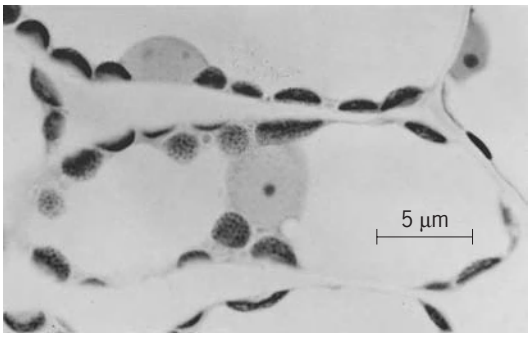

Fig. 1. Light microscope view of part of *Hypochaeris radicata* leaf, showing chloroplasts sectioned in a variety of planes and displaying internal grana.

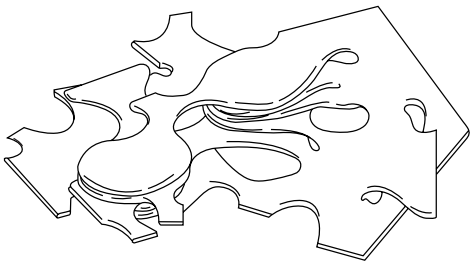

Fig. 2. Three-dimensional model of grana with intergranal connections. (*After W. Wehrmeyer, Planta, 63:13–30, 1964*)

Another system of membranes generally occupies the main body of the plastid. This internal membrane system is especially well developed in chloroplasts, where the unit of construction is known as a thylakoid. In its simplest form this is a sac such as would be obtained if a balloon-shaped, membrane-limited sphere were to be flattened until the internal space was not much thicker than the membrane itself. It is usual, however, for thylakoids to be lobed, branched, or fenestrated. **Figures 2** and **3** show how disk-shaped lobes stack on top of one another to form grana. A granum may have 2–100 layers in it, and there may be more than 50 grana in a chloroplast. It is thought that all the grana are interconnected (**Fig. 4**) and that junctions with the inner layer of the chloroplast envelope are also present. As with their shape and pigmentation, algal chloroplasts are very diverse as regards the architecture of the thylakoids, and often do not possess grana. *See* CELL MEMBRANES.

Functions. The surface area of thylakoids is very large in relation to the volume of the chloroplast. This is functionally significant, for chlorophyll molecules and other components of the light-reaction systems of photosynthesis are associated with these membranes. Techniques such as shadow casting, negative staining, freeze etching, and small-angle x-ray scattering have revealed the presence of subunits within and on the surface of the thylakoid membrane itself. One such unit has been identified as a photophosphorylase enzyme involved in the final stages of the production of adenosine triphosphate by photosynthetic phosphorylation. Others have not yet had specific functions assigned to them, but since the membranes can be fractionated to yield particles that carry out one or the other of the two light reactions, it may well be that photosynthetic units of function are being visualized. Such units might be based on lipoprotein particles, with chlorophyll and other molecules associated with them.

Utilization of the products of light reactions in the actual fixation of carbon dioxide occurs not in the membranes but in the stroma, or ground substance, of the chloroplast. Here reside the enzymes of the Calvin carbon-reduction cycle, including as a major protein component of green tissues the enzyme ribulose diphosphate carboxylase (carboxydismutase). This large molecule, more than 10 nm in diameter, catalyzes the formation of phosphoglyceric

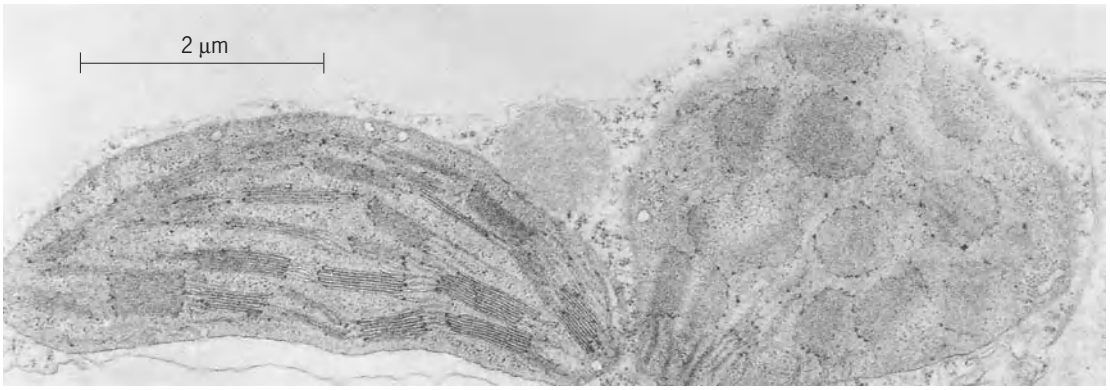

Fig. 3. Electron microscope view of two oat leaf chloroplasts, one showing successive layers of grana in side view, and the other showing grana in face view (disk shape of constituent thylakoid lobes are seen).

acid from ribulose diphosphate (the acceptor molecule) and carbon dioxide. It is, like many other molecular species in the cell, restricted to chloroplasts.

A chloroplast is much more than a device for carrying out photosynthesis. It can use light energy for uptake and exchange of ions and to drive conformational changes. The stroma contains the elements of a protein-synthesizing system—as much deoxyribonucleic acid (DNA) as a small bacterium, various types of ribonucleic acid (RNA), distinctive ribosomes, and polyribosomes. There is evidence to indicate that much of the protein synthesis of a leaf takes place within the chloroplasts. *See* DEOXYRIBONUCLEIC ACID (DNA); PHOTOSYNTHESIS; PROTEIN; RIBONUCLEIC ACID (RNA); RIBOSOMES.

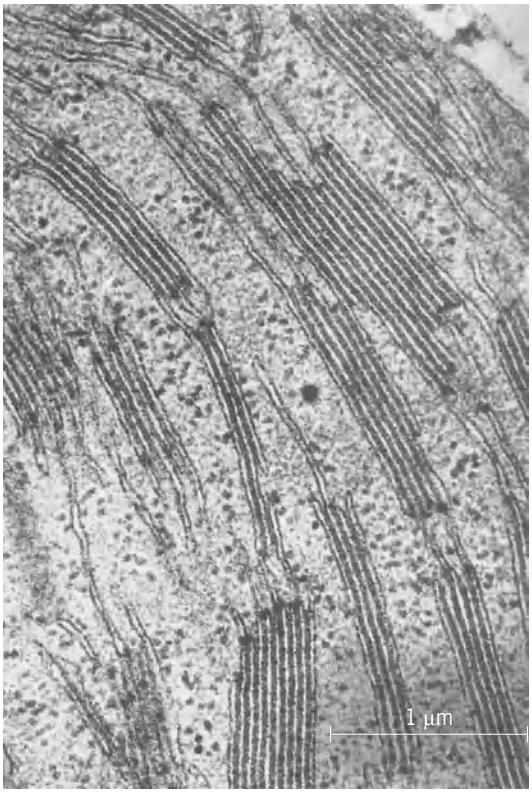

Fig. 4. Part of oat leaf chloroplast at higher magnification. Grana are seen in side view, revealing profiles of membranes. Thylakoids that interconnect grana are shown, with components of stroma, or ground substance, in which thylakoids lie. The dense granules are chloroplast ribosomes.

Mode of development. In algae, mosses, and ferns mature chloroplasts can divide. In cells of flowering plants divisions occur at a much earlier stage of development. Young, undifferentiated cells near the growing points of the plant, the meristems, contain precursor bodies called proplastids (**Fig. 5**), which apparently can divide and multiply. The proplastids only start to develop into one or other of the categories of plastid when the cell that contains them itself starts to differentiate. The ability of amyloplasts to give rise to chloroplasts (for example, when a potato tuber turns green), of chloroplasts to senesce into chromoplasts, and of chromoplasts to "re-green" to give chloroplasts, demonstrates that, although the different types normally develop from proplastids, interconvertibility is not completely lost.

A complex morphogenetic pathway leads from the proplastid to the chloroplast. Precursor molecules are synthesized; some are incorporated into the thylakoids, and the extending thylakoid surface gradually gives rise to interconnected grana. In general, light is required for this process. In darkness a precursor of chlorophyll, protochlorophyll, accumulates and the membranes assume a surprisingly different form, the unit of which is a branched tubule some 20–25 nm in diameter. These tubes interconnect to form semicrystalline lattices based on tetrahedral or cubic arrangements. The lattices, or prolamellar bodies, are up to 3 μm across and contain a relatively large surface area of membrane. When the plant is illuminated, the tubular membranes rearrange and the normal arrangement of interconnected grana develops. Since prolamellar bodies occur in etiolated plants, the plastids that contain them are known as etioplasts.

Morphogenetic autonomy. One of the most challenging problems in cell biology concerns the autonomy of organelles, such as the plastids. Chloroplasts, for instance, have their own DNA, DNA-polymerase,

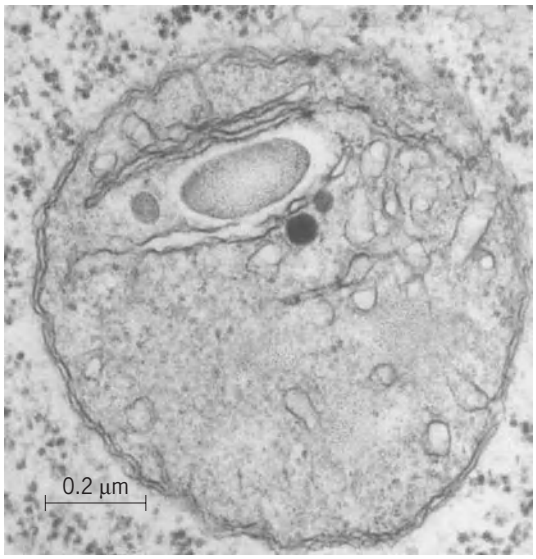

Fig. 5. Proplastid in cell of broad bean root tip. Double membranous envelope is shown and rudimentary nature of internal membrane system is obvious.

and RNA-polymerase; can make proteins; and, significantly, can mutate. All this suggests a measure of independence. It is known, however, that some nuclear genes can influence the production of molecules that are normally found only in chloroplasts, so their autonomy cannot be complete. It remains to be seen whether they control and regulate their own morphogenetic processes. Brian E. S. Gunning

Bibliography. C. J. Avers, *Molecular Cell Biology*, 2d ed., 1986; R. J. Ellis (ed.), *Chloroplast Biogenesis*, 1985; L. Goldstein and D. M. Prescott (eds.), *Cell Biology: A Comprehensive Treatise*, vol. 4, 1980; L. J. Kliensmith and V. M. Kish, *Principles of Cell Biology*, 1987; J. J. Lee and J. F. Fredrick (eds.), *Endocytobiology III*, 1987; W. Waymouth (ed.), *Modern Cell Biology*, vol. 5: *Molecular Mechanisms in the Regulation of Cell Behavior*, 1987.

Cell polarity (biology)

The highly organized condition in most cells that is characterized by a distinct apical-basal axis with an asymmetric distribution of cytoplasmic organelles. This phenomenon of polarization is critical for living organisms to function. For example, cells in secretory organs such as the gall bladder generally secrete only at one end where secretory vesicles are localized. Another example would be a cell in the intestinal wall that must collect nutrients on the side near the lumen and transport them through the cell to the opposite end, where they can be delivered to the blood supply for transport to the rest of the body. Clearly, this cell must exhibit a polarized distribution of membrane proteins so that those necessary for sugar uptake are concentrated in the membrane facing the lumen and those needed to move the sugar out of the cell are located at the opposite side.

Most organisms begin life as a rather symmetrical, spherical, single cell called an egg. During early development this egg divides many times and forms an embryo, which exhibits much more intricate patterns (such as the polarized intestine) than were initially expressed by the egg. Determining how the embryo controls the development of such patterns has been an area of active research, and there is evidence that the plasma (outer) membrane can influence the development of cell polarity by driving ionic currents through the cell which, in turn, can influence the polarization process by generating ion concentration gradients within the cell or voltage gradients between cells. *See* DEVELOPMENTAL BIOLOGY.

Plant eggs. By far the strongest cases for a direct role of membrane-generated ion currents and concentration gradients in cell polarization are found in plant eggs which exhibit no preformed growth axis. Among the most useful have been the eggs of the brown algae *Fucus* and *Pelvetia*, common seaweeds found in the intertidal zone on both the Atlantic and Pacific coasts. The eggs develop by first secreting wall-softening and wall-building enzymes at one end. They then pump in potassium and chloride ions, and water follows these ions into the eggs to generate a large turgor pressure of about 5 atm (500 kilopascals). The eggs bulge (germinate) in the region where the wall has been softened by the secreted enzymes, thereby producing a structure that will develop into the rhizoid or holdfast. This is the first morphologically obvious axis of polarity. The establishment of the axis of secretion, and thus the morphological polarity, can be influenced by a variety of environmental factors, including unilateral light or unilateral chemical influences from nearby eggs or plants. The response to unilateral light is to germinate on the dark side; this is useful since the rhizoid outgrowth will form the holdfast for the plant and should form opposite the Sun so the plant can be anchored to a rock. Since the light receptors are located in the plasma membrane, it is not surprising that this membrane plays a key role in the signal transduction.

Transcellular ionic currents. The earliest indicator of the polarization process is the movement of ions through the egg via pores in the plasma membrane. These ionic currents crossing the plasma membrane can be detected by using a technique called the vibrating probe which measures small voltage gradients in the fluid just outside the egg. As early as 30 min after fertilization, these currents can be found entering the dark hemisphere of a unilaterally illuminated egg. The early spatial current pattern is unstable and shifts position, often with more than one inward current region. However, the current enters mainly on the side where germination will occur, and is usually largest at the prospective cortical clearing region where the rhizoid forms. The current pattern observed during the 2-h period prior to germination is more stable and looks like the pattern on the right side in **Fig. 1**. The site of inward current always predicts the germination site, even when the axis is reversed by light-direction reversal. The most likely hypothesis is that light receptors in the cell's plasma membrane control the distribution of open ion channels so that current enters the dark end.

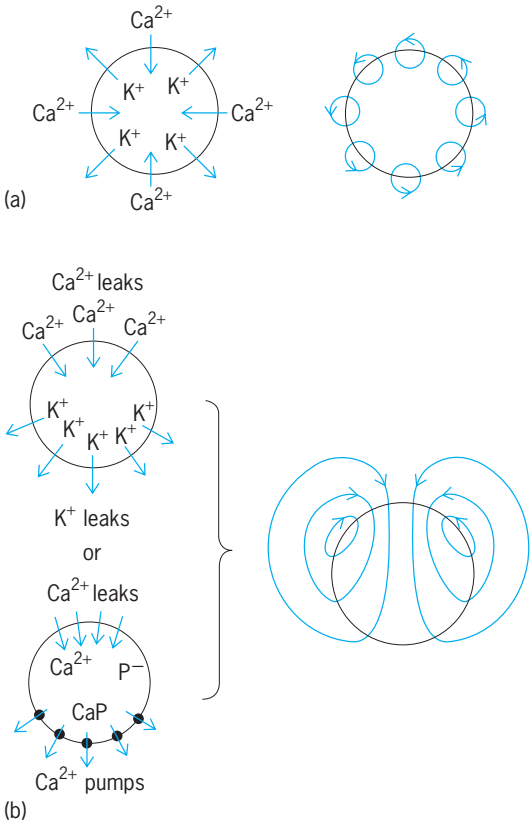

Fig. 1. Possible distribution of ion channels and pumps in a cell's plasma membrane with the resulting current pattern. (*a*) A uniform distribution of calcium (Ca^{2+}) and potassium (K^+) channels will result in many highly localized current loops. (*b*) Separation of the Ca^{2+} and K^+ channels (above) or Ca^{2+} channels and pumps (below) will result in the transcellular ion current pattern shown on the right.

In order to understand how such a transcellular current might be influencing the polarization process, it is important to determine which ions are carrying the current. By varying the composition of the artificial seawaters in which these measurements were made, it was determined that calcium ion (Ca^{2+}) influx was responsible for a small fraction of the current with chloride ion (Cl^-) efflux as the most likely carrier of the bulk of it. This calcium ion component was directly measured by following the movement of radioactive calcium ion through the egg. Therefore, one consequence of the transcellular ionic current in this egg is the generation of a calcium ion flux through the cell, which has been shown to generate a gradient in the intracellular concentration of calcium ion, $[Ca^{2+}]_i$. This gradient appears to play the primary role in polarizing the egg.

Calcium ion gradients. It is known that a central component of the polarization mechanism in the egg of a brown alga involves a cytoplasmic gradient in $[Ca^{2+}]_i$. This hypothesis is supported by the three criteria required for physiological significance: (1) An endogenous $[Ca^{2+}]_i$ gradient exists in the egg (although measured long after polarization); (2) modifying the direction of this gradient during polarization changes the orientation of the axis of polarization; and (3) inhibiting the gradient prevents polarization. The gradient was detected in germinated zygotes by using calcium ion–sensitive microelectrodes, and a 10-fold higher $[Ca^{2+}]_i$ was detected in the tip region than in the subtip region of the rhizoid. A calcium ion gradient was imposed on unpolarized eggs by placing them near a fixed source of calcium ion ionophore in the dark. A majority of the eggs then germinated on the hemisphere facing the ionophore, where the influx of calcium ion is expected to be the greatest. Finally, the magnitude of the gradient can be greatly reduced by injecting the cell with a molecule that can rapidly shuttle calcium ions from regions of high concentration to regions of low concentration. Such a molecule was used to show that polarization could be blocked for weeks while the cell remained viable and even divided. This is the strongest evidence yet that a gradient in $[Ca^{2+}]_i$ is required for cell polarization.

Ionic currents and intracellular ion concentration gradients are not a unique property of brown algae, but have been detected in many other cell types as well. In fact, ionic currents have been detected in nearly every plant and animal cell and tissue investigated with the vibrating probe technique (**Fig. 2**). The transcellular current pattern is usually closely correlated with the axis of polarity. It would therefore appear that most cells do not have a uniform distribution of ion channels and pumps (Fig. 1*a*) but segregate channel types to varying degrees (Fig. 1*b*). *See* BIOPOTENTIALS AND IONIC CURRENTS.

Animal cells. A wide variety of animal systems have been investigated with the vibrating probe, and the best example of the use of ionic currents in maintaining the polarized state comes from the developing egg of the moth *Hyalophora cecropia*. This egg is connected to seven other cells called nurse cells which produce important molecules that are transported into the egg. This transport is polarized since most proteins will move from the nurse cells into the egg but will not move from the egg back to the nurse cells. It was demonstrated that a voltage difference of about 5 millivolts exists between the nurse cells and oocyte in the follicle of the silk moth. The only way that such a voltage gradient can be maintained in the conductive cytoplasm is by a steady current flow. Thus, the plasma membranes of these cells must be driving a steady current along the cytoplasmic bridge between the cells to generate the observed voltage gradient. Moreover, there is strong evidence from the behavior of microinjected proteins that intercellular electrophoresis is involved as a protein segregation mechanism here. The polarity of protein transport across the bridge connecting these two cell types can be reversed by reversing the endogenous electrical field with this bridge.

While it is not too surprising that proteins could be electrophoresed along an intercellular bridge by an imposed field, even more compelling evidence that electrophoresis is involved in this transport has been found. The polarity of movement of given protein across the bridge is dependent only on its net electric charge. Most cellular proteins are negatively charged and so would be driven from the nurse cells into the

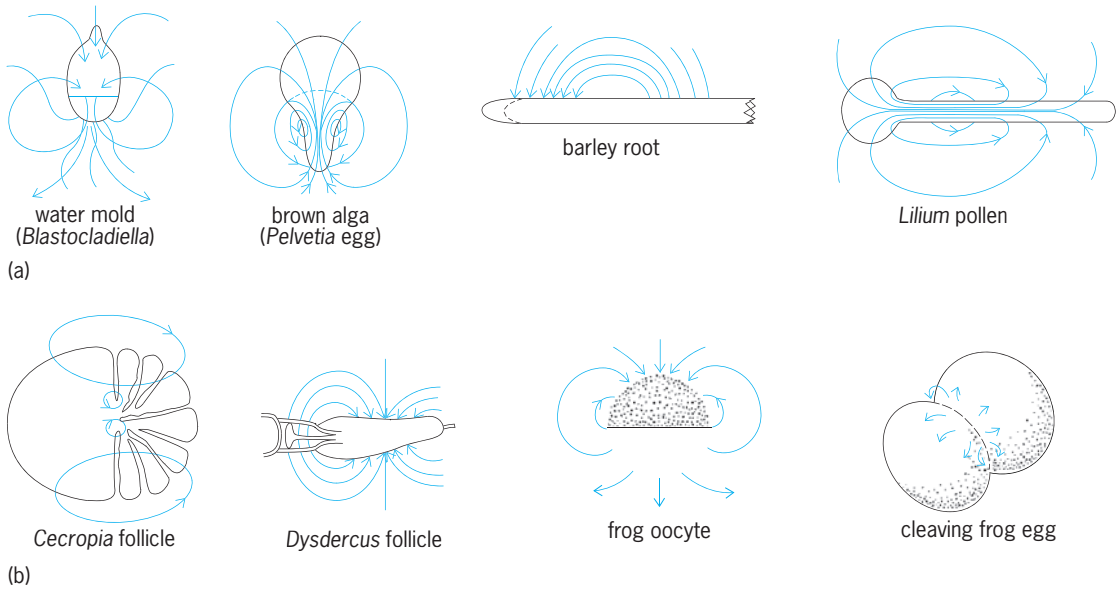

Fig. 2. Transcellular ion current patterns measured in (*a*) plant cells and (*b*) animal cells.

oocyte by the voltage gradient along the cytoplasmic bridge. Lysozyme is a basic protein with a net positive charge, and when it is injected into either the nurse cells or the oocyte, it is found to move only in the opposite direction, from the oocyte to the nurse cells. However, when the net charge on this protein is reversed by methylcarboxylation, its transport direction was reversed, and it was observed only to move from nurse cells to oocyte. In further support for the electrophoresis mechanism, neutral proteins with no net electric charge move in both directions across the bridge. Thus, in the cecropia follicle, the polarity of protein transport, which is a critical component of the oocyte's polarity, is determined by the electrical field across the intercellular bridge. The cecropia follicle is the best example in which transcellular currents play an active, causal role in polarized transport. *See* CELL (BIOLOGY). Richard Nuccitelli

Bibliography. C. Brownlee and J. W. Wood, A gradient of cytoplasmic free calcium in growing rhizoid cells of *Fucus serratus, Nature*, 320:624–626, 1986; L. F. Jaffe and R. Nuccitelli, An ultrasensitive vibrating probe for measuring steady extracellular currents, *J. Cell Biol.*, 63:614–628, 1974; J. R. McIntosh (ed.), *Modern Cell Biology*, vol. 2, 1983; R. Nuccitelli, Ooplasmic segregation and secretion in the *Pelvetia* egg is accompanied by a membrane-generated electrical current, *Dev. Biol.*, 62:13–33, 1978; R. I. Woodruff and W. H. Telfer, Electrophoresis of proteins in intercellular bridges, *Nature*, 286:84–86, 1980.

Cell senescence

The limited capacity of all normal human and other animal cells to reproduce and function. The gradual decline in normal physiological function of the cells is referred to as aging or senescence. The aging process ends with the death of individual cells and then, generally, the whole animal. Aging occurs in all animals, except those that do not reach a fixed body size such as some tortoises and sharks, sturgeon, and several other kinds of fishes. These animals die as the result of accidents or disease, but losses in normal physiological function do not seem to occur. Examples of cells that do not age are those composing the germ plasm (sex cells) and many kinds of cancer cells. These cells are presumed to be immortal.

Normal cells. In the mid-1960s it was found that, contrary to the prevailing belief, normal human cells cultured in the laboratory do have a limited capacity to function and to divide. This phenomenon was interpreted as aging at the cellular level. These observations have been extended to include the cells of all other animals tested. *See* TISSUE CULTURE.

Prior to this finding, it was believed that all cultured cells were potentially immortal. If that were true, aging in animals would be the result of events that occur at organizational levels higher than individual cells. That is, age changes would be attributed to nongenetic events occurring outside cells, such as changes in the substances that bind cells into tissues, externally caused injuries to cells, hormone effects, and other interactions between tissues and organs. Now that it is known that normal human and other animals cells have a limited capacity to divide and function, there is a general belief that cell senescence must be attributed to intracellular events. These events, which would play a major role in senescence or aging, are thought to occur in the deoxyribonucleic acid (DNA) and other information-containing molecules.

Studies have shown that when normal human fibroblasts, derived from human embryos, are grown in laboratory culture, they can undergo only about 50 population doublings (**Fig. 1**). Fibroblasts are found in virtually all tissues, and besides lending structural

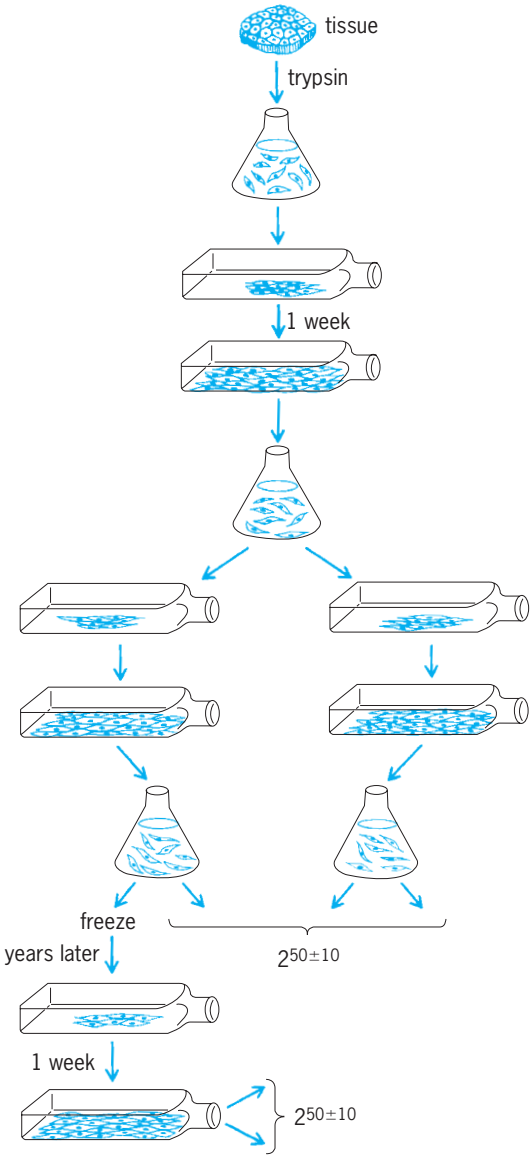

Fig. 1. Culture of normal human cells. The tissue is broken down into individual cells with the digestive enzyme trypsin. The cells are then separated from the trypsin and placed into a flat bottle, where they multiply until they cover the floor of the culture vessel. When they stop multiplying, the cells are divided into equal numbers and placed into two culture vessels. The process is repeated every few days after the cells have covered the available surface area. Each subcultivation results in one population doubling. (*After L. Hayflick, The cell biology of human aging, Sci. Amer., 242(1):58–66, 1980*)

support, they produce collagen and other proteins that bind cells into tissues. The maximum number of doublings attainable was inversely proportional to donor age: fibroblasts from older people were found to undergo fewer doublings than those derived from younger donors.

Cultured fibroblasts derived from different animal species seem to undergo a number of population doublings that are directly proportional to the species lifespan (**Fig. 2**).

Nondividing cells. Fibroblasts have a greater capacity to divide in laboratory cultures than any other normal cell type. Nevertheless, since all normal human and other animal cells have a finite capacity to divide or to function, nondividing cells, such as neurons or muscle cells, also age. The loss of function in brain neurons may indeed play a greater role in age changes than do changes in any other cell type.

When normal human embryonic cells are frozen at a particular population doubling, they "remember" at what doubling level they were frozen and, when thawed, continue to divide until the total of 50 is reached. It has been shown that the frozen cells' memory is accurate over a period of at least 22 years.

Normal human and other animal cells are mortal not only when they are grown in laboratory cultures but even when normal tissues are grafted into laboratory animals. When the laboratory animal grows old, the graft can be removed and placed on a younger host animal. This serial animal-to-animal transfer, which is analogous to cell transfers in laboratory cultures, also reveals the finitude in division capacity and function of normal cells.

Changes in senescent cells. Prior to the death of cultured normal human cells, they incur well over 100 changes that herald their approaching demise. These changes occur in lipid content and synthesis; carbohydrate utilization; protein content, synthesis, and breakdown; ribonucleic acid (RNA) and DNA content, synthesis, and turnover; enzyme activity and synthesis; cell-cycle kinetics; morphology; ultrastructure; cell architecture; and incorporation and stimulation. Of great importance is that many of these age changes seen in cultures at the cell level are also observed in humans as they age.

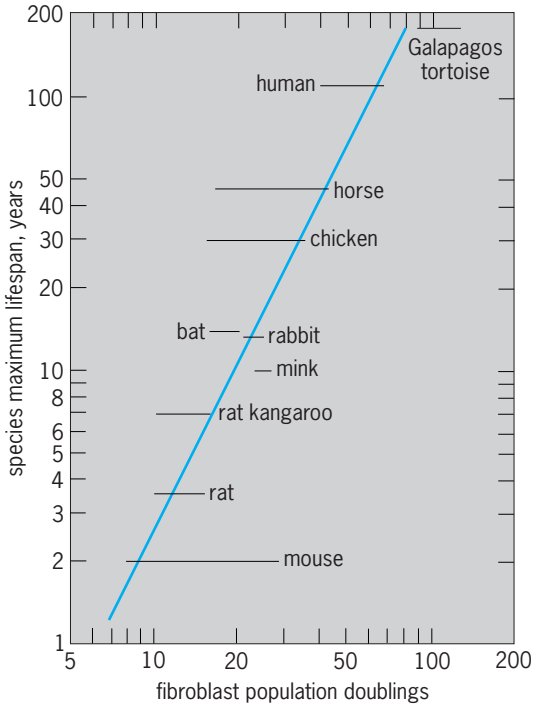

Fig. 2. Fibroblasts from the young of 10 different species multiply in culture roughly proportionate to the maximum lifespan for that species.

Immortality. Although cultured normal human and other animal cells are mortal, they can be converted to a state of immortality. The conversion can be produced in human cells by the SV40 virus and in other animal cells by other viruses, chemicals, and irradiation. This conversion from mortality to immortality is called transformation, and is characterized by the acquisition of many profound abnormal cell properties, including changes in chromosome number and form, ability of the cells to grow unattached to a solid surface, and the property of immortality. These changes, and many more, are characteristic of most cancer cells. *See* CANCER (MEDICINE); CELL (BIOLOGY); MUTATION; ONCOLOGY; TUMOR VIRUSES.

Leonard Hayflick

Bibliography. A. Comfort, *The Biology of Senescence,* 1979; C. Finch and L. Hayflick (eds.), *Handbook of the Biology of Aging,* 1977; L. Hayflick, The cell biology of aging, *Bioscience*, 25:629–636, 1975; L. Hayflick, The cell biology of human aging, *Sci. Amer.*, 242(1):58–66, 1980; L. Hayflick and P. S. Moorhead, The serial cultivation of human diploid cell strains, *Exp. Cell Res.*, 25:585–621, 1961; A. Rosenfeld, *Prolongevity*, 1976.

Cell-surface ionization

All living cells suspended in aqueous salt solutions at neutral pH values possess a negative charge. The charge is due to the dissociation of ionogenic, or charged, groups (carboxyl, amino, and others) in the cell surface. Studies of the charge carried by bacteria give information regarding the nature of the components in their surfaces; these studies have been used to measure the frequency distribution of bacterial variants in a mixed population and can assist in the selection of material for vaccine production. The occurrence and nature of tumor cells and the selection of yeasts for brewing have been studied electrophoretically. *See* BIOLOGICALS.

Electrophoretic mobility. When an electric field is applied between electrodes placed in a cell suspension, the cells migrate, generally toward the positive electrode. This migration can be observed under a microscope in a microelectrophoresis apparatus (see **illus.**). The electrophoretic mobility $\overline{v}$ (velocity per unit field strength) is measured by timing cells across a fixed distance in the observation chamber under a known field strength. In this method the mobility of individual cells of varying shape and size can be measured with minimal cellular disorganization. Migration occurs because charged groups and ions adsorbed on the surface produce a nonuniform distribution of ions in solution at the cell-liquid interface. This distribution is generally expressed in terms of the ζ potential (volts), which may be calculated from the mobility, $\zeta = 4\pi\eta\overline{v}\upsilon/D$, where η and D are the viscosity and dielectric constants, respectively, of the medium. The surface charge density in electrostatic units per square centimeter can be calculated from ζ.

Variation of charge. Factors which may affect the charge of cell surfaces are of three types: (1) biological factors—in bacterial and yeast cells, such things as sex, strain, age, growth conditions, presence of capsules or fimbriae, antibiotic resistance, virulence, and toxicity; in mammalian cells, such things as species, type of tissue, normal or pathological conditions, and the like; (2) nature of suspension medium, that is, ionic strength, pH, presence of dyes, drugs, or surface-active agents in the medium; and (3) chemical or enzymatic treatment of the surface. It is often possible to correlate alteration of charge with changes in the nature or number of ionogenic groups in the surface and hence with a particular biological property. Such a correlation is justified only if the variation of charge is measured in media of constant physical properties. An absolute value of the charge under standard conditions of test is without value; the charge carried by particles with ionogenic and nonionogenic surfaces under the same conditions may be identical. However, variations of charge with changes in the nature of the medium are completely different. With living cells, lowering the pH of the medium at constant ionic strength in general produces a lowering of the negative value of the ζ potential; at low pH value, charge reversal may occur. This change results from the decreased dissociation of acidic groups which may be characterized from the dissociation constant pK of the surface. Simultaneously there is an increased ionization of basic groups. At extreme pH values surface and possibly cellular disorganization must be investigated. When the salt concentration of the suspension medium is increased, biological and nonionogenic surfaces behave similarly. The charge-density concentration curve is a typical Langmuir adsorption isotherm, indicating adsorption of anions from solution. *See* ADSORPTION.

Ionogenic groups can be characterized by the binding properties of a series of cations. Their order of effectiveness in reversing the charge of the

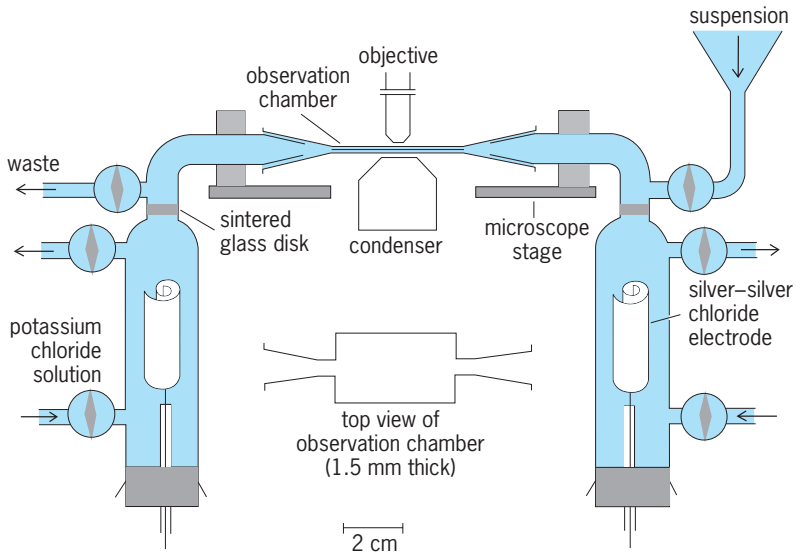

Microelectrophoresis apparatus used to measure electrophoretic mobility.

particles is known as a charge-reversal spectrum. The spectra for various cells can be compared with spectra of model systems with known ionogenic surface groups. Alteration of charge in the presence of anionic surface-active agents can reveal the presence of lipid in the surface. Mild chemical treatment to inactivate specific groups, for example, diazomethane for carboryl, results in marked changes of charge. Enzymatic treatments, for example, using the enzyme trypsin, may produce changes in the charge. *See* ENZYME; LIPID; PROTEIN.

Bacteria. Most strains show slight, if any, variation of charge density during growth. Any drastic alteration of charge always occurs during the period of active cell division, returning to the normal value during the stationary phase, as in hemolytic streptococci. Little alteration of charge generally results from changes in nutrition. Growth in the presence of some drugs, for example, *Aerobacter aerogenes* in crystal violet, results in the formation of a resistant strain with different charge characteristics. The pK values of the surfaces of *Escherichia coli* (2.9) and *A. aerogenes* (3.0), obtained from ζ-pH studies, are the same as those for a large number of polysaccharides, strongly suggesting that carboxyl and not phosphatidic groups (pK = 1.8) are present on the cell surface. Charge-reversal spectra lead to the same general conclusions. Treatment of *E. coli* K12 in aqueous suspension at pH 4.5 with gaseous diazomethane results in an 80% decrease of charge when measured at pH 7.0. Assuming that all the surface carboxyl groups have been esterified, the number of such groups is about 8×10^5 per cell. Spores of *Bacillus subtilis* treated with *p*-toluene sulfonyl chloride in aqueous suspension show a 25% increase in negative charge; this is attributed to the blocking of amino groups.

Teichoic acid, on the surface of cells of *Staphylococcus aureus* in the pH range 4–5, undergoes a change of configuration resulting in a pH-mobility curve which passes through a maximum at pH 4.5. Strains which are resistant to Methicill give a normal curve with no maximum. Cells of *Streptococcus pyogenes* which are resistant to Tetracycline possess larger amounts of surface lipid, as detected by changes of mobility in the presence of surface-active agents, than do sensitive strains.

Yeasts. In suspensions of pH greater than 2.3, yeast cells carry a negative charge. Various strains of *Saccharomyces cerevisiae* and *S. carlsbergensis* have such low charges that ion binding is small, the composition of the wall determining the charge. In some strains the charge is due to combined phosphate in the cell wall and in others to protein. The nature of the growth medium has a pronounced effect on the charge.

Blood cells. The major components on the erythrocyte surface are phospholipid in nature; protein and cholesterol may also be present near the surface. Of sheep blood cells, lymphocytes exhibit an isoelectric point at pH 2.5 (that pH at which there is a zero charge) in contrast to erythrocytes and polymorphonuclear leukocytes, which are damaged at low pH values. Charge-reversal spectra and chemical treatments of aldehyde-treated blood cells indicate that the predominant groups are carboxyl from sialic acid on erythrocytes, carboxyl on polymorphonuclear leukocytes, and both phosphate and amine on lymphocytes and platelets. *See* BACTERIA; BIOPHYSICS; BLOOD; COLLOID; ELECTROKINETIC PHENOMENA; STREAMING POTENTIAL; YEAST.

Arthur M. James

Bibliography. M. Blank, *Surface Chemistry of Biological Systems*, 1970; W. Hoppe, *Biophysics*, 1983.

Cell walls (plant)

The cell wall is the layer of material secreted by the plant cell outside its plasma membrane. All plants have cell walls that are generally very similar in chemical composition, organization, and development. The walls of the Chlorophyta (green algae) show characteristics virtually identical to those of flowering plants, another indication that flowering plants are derived evolutionarily from this division of algae. The wall serves as the first point of entry of materials into cells, functions in the movement of water throughout the plant, and is one of the major mechanical strengthening factors. In addition, the wall must be sufficiently flexible and plastic to withstand mechanical stresses while still permitting the growth of the cell. *See* CELL MEMBRANES.

Cell plate and middle lamella. The plant primary wall is initiated during the process of cell division. After chromosomes line up along the metaphase plate and begin to be pulled apart toward the poles of the cells by the spindle fibers (the anaphase portion of mitosis), a cell plate or phragmoplast can be observed at the equator of the dividing cell. Vesicles apparently derived from the dictyosomes line up on both sides of the equator to form the proteinaceous cell plate. Elements of the endoplasmic reticulum fuse with the cell plate, marking the location of

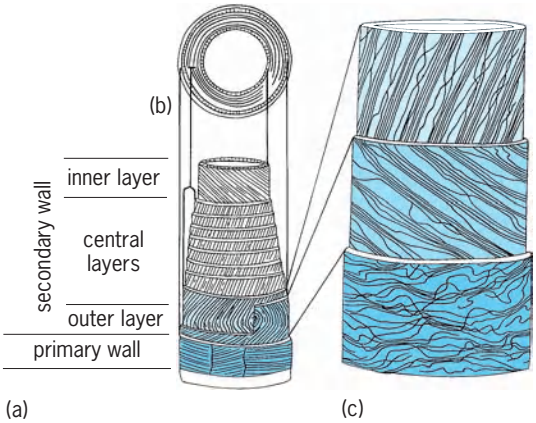

Fig. 1. Wall in a cotton fiber. (*a*) Telescoped segment and (*b*) transverse section showing spatial relation of layers and orientation of microfibrils. (*c*) Primary wall has reticulate microfibrillar structure; the outer layer of the secondary wall combines reticulate and parallel orientation of microfibrils, and the first central layer of the secondary wall has a predominantly parallel microfibrillar structure. (*After K. Esau, Plant Anatomy, 2d ed., Wiley, 1965*)

plasmodesmatal pores and pits which will eventually provide the intercellular connections between adjacent cells. The cell plate forms the matrix within which the middle lamella and primary walls are formed. The middle lamella is composed of pectic substances which are polymers of neutral (arabinogalactans) and acidic (galacturonorhamnans) pectins plus smaller amounts of other sugars, including L-fucose, D-xylose, and D-galactose. The bulk of pectic substances are long chains of α-D-galacturonic acid, with side chains composed of other sugars; the exact chemical composition and organization of pectins are still unclear. Pectins can form reversible gels or viscid solutions in water. As the middle lamella, pectic substances are usually in the water-insoluble gel state, possibly because they are esterified with calcium or magnesium ions. Such gels can be solubilized and extracted from plants. Some tissues, notably those of mature apple and citrus fruits which have high concentrations of pectins, have been extracted and purified, and their pectins have been used as gelling agents in jams, jellies, and some industrial products. The middle lamella provides some of the observed plasticity and extensibility of cell walls during cell growth, and it has also been suggested that pectins are capable of hydrogen-bonding to the cellulose that forms the plant cell primary wall.

Primary wall. Simultaneously with the middle lamella, the primary cell wall begins to form, usually before the end of telophase. During the early stages in cell wall formation, the cellulose wall is isotropic without any ordered orientation, but as cell walls continue to develop in area and in thickness and the cell grows to mature size, the walls become anisotropic or highly ordered (**Fig. 1**).

Biosynthesis and chemistry of cellulose. Cellulose, like starch, is basically a polymer of glucose, a six-carbon monosaccharide, although the biosynthetic route for cellulose synthesis differs from that of starch. The polyglucose chain of starch is derived by enzymatically regulated condensation of uridine diphosphate glucose (UDPG) or adenine diphosphate glucose (ADPG). However, the monomers that form cellulose are guanosine diphosphate glucose (GDPG) with the resulting polymer being a β-(1–4) glucoside rather than the α-(1–4) polymer characteristic of amylose. In both polymers, the terminal glucose residue can unfold into its open-chain configuration, exposing the aldehyde group that has reducing power.

Each chain of cellulose may be as long as 8000 to 12,000 glucose monomers, or up to 4 micrometers long. These are arranged linearly, with no side branching. Cellulose chains are aggregated into bundles of approximately 40 chains each, the cellulose micelles, which are held together by hydrogen bonds between the hydrogen atom of a hydroxyl group of one glucose residue and an oxygen atom of another glucose residue which is adjacent to the first glucose. Although the strength of individual hydrogen bonds is rather weak, there may be so many of them joining adjacent chains that total bonding in a micelle is quite high. There is good, albeit incomplete, evidence that all of the chains in the same micelle bundle are directionally oriented, with the reducing groups at the same end of the micelle; and there is some fairly conclusive evidence from the x-ray diffraction patterns of cellulose that the glucose units of one chain are displaced by one-half the length of a glucose monomer, so that there are two hydrogen bonds connecting adjacent monomers. The micelle thus is a very regular, quasicrystalline structure (**Fig. 2**).

Whether isotropic or anisotropic, the micelles themselves are embedded in a matrix of other polysaccharides, the hemicelluloses. Insofar as is known, hemicelluloses are produced by the dictyosomes and associated vesicles and are moved through the cytoplasm and the plasmalemma to enter into the structure of the primary wall. In addition to glucose, seven monosaccharides are involved in hemicellulose composition. About a quarter of this monosaccharide complex is arabinose, with 10–15% each of galactose, galacturonic acid, and xylose plus smaller percentages of rhamnose, fucose, and mannose. These monosaccharides may form unispecific polymers composed of a single monomer or may be complexes containing two or more monomers. In the primary plant cell wall, these polysaccharides are covalently bonded. One model (**Fig. 3**) suggests

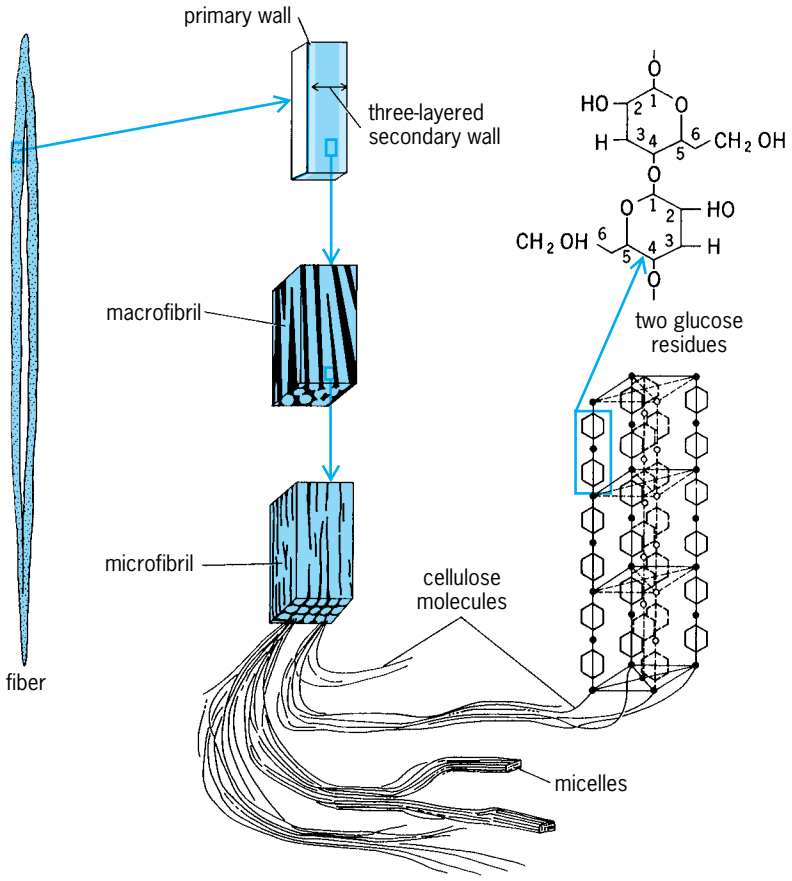

Fig. 2. Interpretation of wall structure. Fiber has a three-layered secondary wall. The macrofibrils consist of numerous microfibrils of cellulose interspersed by microporosities containing noncellulosic wall materials. Microfibrils consist of bundles of cellulose molecules, partly arranged into micelles. Micelles are crystalline because of regular spacing of glucose residues which are connected by β-1,4-glucosidic bonds. (*After K. Esau, Plant Anatomy, 2d ed., Wiley, 1965*)

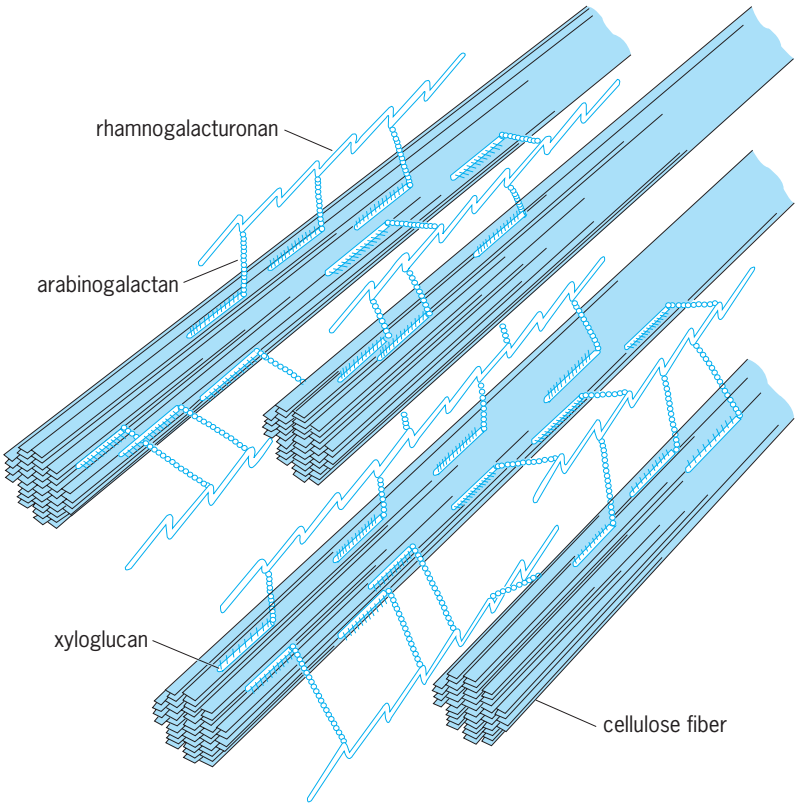

Fig. 3. Model of the cell wall showing cellulosic wall organization. (*After P. Albersheim, The wall growing plant cells, Sci. Amer., 232(4):80–95, April 1975*)

that each of the cellulose micelles is completely coated with a monomolecular layer of xyloglucan, with the glucose units of the xyloglucan lying parallel to the long axis of the micelle and hydrogen-bonded to glucose residues in the cellulose. Radiating at right angles from the coated micelle are arabinogalactan polymers, which are then covalently linked to rhamnogalactan polymers which run parallel to the cellulose micelle. Thus, hemicellulose serves to bind the micelle into a fairly rigid unit which retains a good deal of flexibility.

Micelles, in bundles of variable number, are bound together into the cellulose microfibril, a unit sufficiently large to be seen under the electron microscope; these, in turn, are bound together into macrofibrils which are observable under the light microscope. *See* CELLULOSE.

Organization of primary wall. During the formation of the primary wall, at locations predetermined by attachments of endoplasmic reticulum to the middle lamella, cellulose microfibrillar deposition is minimal, leaving a thin place in the primary wall which forms the plasmodesmatal connections. Running through these pores are fine strands of protoplasm, the plasmodesmata proper, which contain a tube of endoplasmic reticulum–like material. The plasmodesmata provide a cytoplasmic connection between adjacent cells. Such connections are found among all the living cells of a plant, a fact which has led to the concept that all plant cells are so interconnected that the entire plant is a cytosymplast or single unit.

In the formation of the primary cell wall, there may be several discrete depositions of cellulose, with changes in the orientation of the micro- and macrofibrils with successive depositions. There has been some equivocation as to whether the added cellulosic fibrils are inserted among those already part of the wall (intussusception) or whether the new fibrils are laid down inside the older portions of the primary wall as a discrete layer or lamella (apposition). Most of the available evidence suggests that apposition is the only method of increasing the thickness of the wall. This can be readily seen in the formation of both primary and secondary walls in the cotton fiber, the hairs attached to the seed coat which are the starting material for cotton thread (Fig. 1).

Secondary wall. Although there are differences in nomenclature and terminology, secondary walls of plant cells are defined as those laid down after the primary wall has stopped increasing in surface area, essentially at that time when the plant cell has reached mature size. This is particularly true of those cells that, at maturity, have irreversibly differentiated into specialized cells, some of which are destined to lose their cytoplasm and become functional only as dead cells. These include xylem vessels and tracheids, and sclereids (the stone or grit cells of pear fruits). In those xylem cells differentiated from parenchyma cells to form the primary vascular xylem, secondary wall thickenings may not occur uniformly over the surface of the wall but may be in the form of bands, spirals, or reticulate thickenings.

The secondary wall of most plants seems to have the same chemical structure and physical orientation of fibrils and hemicelluloses as do primary walls. While there may be little orientation of fibrils in young primary walls, the secondary walls are composed of fibrils that are highly ordered. In general, there are three layers or lamellae of fibrils. Two of these, usually the inner and outer layers, tend to be thin and birefringent, and the middle layer tends to be thicker and less birefringent. One of the differences between the larger xylem vessels and tracheids of spring wood and the smaller xylem cells of summer wood is the thickness of the middle layer of cellulose fibrils. Part of the difference in birefringence, on the other hand, seems to be due to the reduced anisotropy of the middle secondary wall layer.

Since successive layers of cellulose fibrils are laid down inside one another, with the youngest layer closest to the plasmalemma, the lumen of the cell which contains the protoplast becomes reduced in size and, in some cells, may virtually disappear.

In most secondary walls, and particularly those of the xylem, the fibrillar structure of the primary as well as the secondary walls may become impregnated with more substances, the most prominent of these being lignin. The chemical nature and biological role of lignin is of considerable interest because of the use of wood in the lumber and pulpwood-paper industries. Lignins are polymers of complex phenolic substances (phenylpropanoid C_6-C_3 compounds), and exact composition of the polymer may be species-specific. Certainly the lignins of the

softwoods (conifers) and those of hardwoods (broad-leaved trees) are chemically distinguishable based on both chemical and physical properties and on the ratios of the various phenolic monomers composing the polymer. As the secondary walls develop, lignins may be deposited either from the inside layer toward the outermost or the reverse. The primary roles of the lignins include their ability to render walls mechanically strong, rigid, and—at least to some extent—water-impermeable. It has been suggested that lignins may also serve to make wood less subject to microbially caused decay.

Depending on the species, other compounds may form part of the secondary walls of plant cells. In some of the algae, notably the diatoms and phylogenetically related classes in the division Chrysophyta, mature cell walls may consist almost entirely of silica compounds. Although silica is usually not a major component in the walls of vascular plants, grasses and sedges have walls with significant amounts of silica, and the use of the scouring rush (*Equisetum* spp.) was based in part of the silica content of the walls of stem cells. In some algae and in a few flowering plants (including *Ficus*), deposition of calcium carbonate into both primary and secondary walls has been found. In many plants, an ill-characterized compound called callose is found; it is particularly prominent in the end or sieve plates of phloem cells, where it surrounds the sieve pores. Callose may vary in amount with the developmental stage of the plant, with the seasons, and as a result of trauma.

Although frequently not considered to be a structural component of plant cell walls, water is an important constituent. A good deal of the movement of water through a tissue system occurs not by the movement of water into and out of successive and adjacent cells, but through the intercellular spaces between cells and through the primary and secondary walls. Many plant physiologists consider the walls and intercellular spaces to be the apparent free space that permits the movement of water and dissolved substances at rates considerably in excess of that possible were all water to move from cell to cell.

Other components of cell walls have received relatively little attention from chemists and developmental physiologists, although they play important roles in the economy of the plant and may, in some instances, be of commercial value. Walls of red and brown algae contain polymers such as algin, agar, and carrageenan, which form rigid gels used as substrates for growing microorganisms and as thickening and stabilizing agents in ice cream and industrial products. Cutins, waxes, and suberins are lipoidal materials that serve to waterproof cells in whose walls they are deposited. Cutin, found as the outermost layer of the outside wall of epidermal cells, is the primary barrier of the plant, serving to prevent the leaching of cell constituents and functioning as a protective layer to prevent invasion of the plant by microorganisms. Cutins also serve to lessen water loss from leaves and stems. Suberins are prominent components of the walls of cork and bark cells and of the casparian strip of root endodermal cells. Both are polymers of fatty acids. Waxes may also be found on epidermal layers of leaves and fruit. They are usually small polymers and are responsible for the "bloom" which, in apples, allows the fruit to be polished. Waxes can be removed from plants and used in shoe polish and other applications. *See* WAX, ANIMAL AND VEGETABLE.

Pits. As the primary wall develops, areas of the wall surrounding the plasmodesmatal intercellular connections tend to be thinner than the rest of the wall and are called primary pit fields. In some cells, primary pit fields may be deeply depressed as cellulose microfibrillar layers accumulate and the rest of the primary wall becomes thicker. Pit fields are almost always paired, with the pit field of one cell oriented exactly opposite to that of the adjacent cell. Many cells have large numbers of primary pit fields.

In plant secondary walls, pits are readily distinguishable from the rest of the wall because the layers of secondary wall materials are completely interrupted at the pit with only thin coverings, if any, of primary wall. In mature cells with well-developed secondary walls, the pits may be simple or, as seen in the bordered pits common in xylem vessels, tracheids, and fibers, they may be structurally very complex. *See* XYLEM.

In bordered pits, the developing secondary wall arches over the pit to form a thickened rim (the border), enclosing a pit chamber which may be open and continuous with the cell lumen (**Fig. 4**). In conifers, the primary walls in the center become thickened (the torus), while on the edges (the margo) they become thinner as hemicelluloses are digested away. It has been suggested, primarily but not exclusively on anatomical grounds, that bordered pits in xylem may serve as valves to control the rate of water flow between the adjacent cells bearing the pair of pits. Depending on the species, bordered pits may be very complex.

Cell elongation and enlargement. In vascular plants, cells grow in length, thickness, or both, primarily by

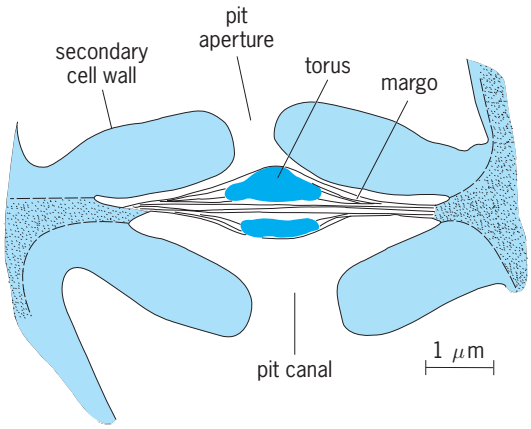

Fig. 4. Cross section through a bordered pit of a transfusion tracheid from a needle of *Pinus sylvestris*, showing the pit membrane. (*After E. G. Cutter, Plant Anatomy: Experiment and Interpretation, pt. 1, Addison-Wesley, 1969*)

the uptake of water into their vacuoles. The initiation and maintenance of growth is dependent not only upon the availability of water and the concentrations of auxins and other growth-regulating compounds that control cell enlargements, but on characteristics of the cell wall. Even in primary walls the cellulosic components and the matrix in which the fibrils are embedded form a fairly rigid structure, and for growth to occur there must be loosening of the fibrillar matrix. Furthermore, the effects of indoleacetic acid (IAA) on cell elongation are expressed within a minute or even less following presentation of auxin, indicating that the necessary loosening must occur quite rapidly. Any comprehensive theory must also account for the fact that, as a cell elongates, its wall does not become thinner, nor is new wall material deposited only at the ends of the cell. The new wall fibrils are deposited fairly uniformly over the entire surface of the cell. *See* AUXIN.

On the manifold questions relating to cell elongation and enlargement, the question of the deposition of new cellulose fibrils appears to be the most amenable to answer. Most, if not all, of the fibrillar material is deposited by apposition on the inner portion of the walls adjacent in the middle lamella. Intussusception, if it occurs at all, is a minor factor in new wall deposition during growth. Little attention has been directed toward understanding the possible changes in the middle lamella. Clearly, if the cell is to continue to grow, with increases in surface area of the wall, the middle lamella must either increase proportionally or be drastically thinned; no unequivocal data are available on this point.

Since cell wall area is accompanied by appositional deposition of fibrils, the fibrils existing prior to the period of growth must slide past one another. It has been suggested that this occurs by the enzymatic breaking of bonds that occur between the cross-linked fibrils with almost immediate reformation of these bonds at new sites, but no information has been provided on the chemical nature of the bonds affected or the polymers involved. Other investigators have postulated that the bond breaking and reforming may be due to changes in the proton concentration caused by application of auxin. This "acid growth theory" suggests that IAA stimulates an adenosine triphosphatase/adenosine triphosphate (ATPase/ATP) reaction, providing the energy for the mutual exchange of potassium ions and protons across the plasmalemma, and that this may be the mechanism for breaking the bonds linking fibrils together. *See* PLANT CELL; PLANT GROWTH.

Richard M. Klein

Bibliography. J. D. Mauseth, *Plant Anatomy*, 1988; R. D. Preston, *The Physical Biology of Plant Cell Walls*, 1974.

Cellophane

A clear, flexible film made from cellulose. It first appeared commercially in the United States in 1924, and it revolutionized the packaging industry, which had been using opaque waxed paper or glassine as wrapping materials. Cellophane was also the first transparent mending tape. By 1960, petrochemical-based polymers (polyolefins) such as polyethylene had surpassed cellophane for use as a packaging film. Nevertheless, 50,000 tons (45,000 metric tons) of cellophane are used in the United States each year for packaging because cellophane gives a material that is stiffer and more easily imprinted than are polyolefin films. *See* CELLULOSE.

Manufacture. Cellophane is manufactured in a process that is very similar to that for rayon. Special wood pulp, known as dissolving pulp, which is white like cotton and contains 92–98% cellulose, is treated with strong alkali in a process known as mercerization, a treatment used in processing cotton for textiles. The mercerized pulp is aged for several days. *See* TEXTILE CHEMISTRY.

The aged, shredded pulp is then treated with carbon disulfide, which reacts with the cellulose and dissolves it to form a viscous, orange solution of cellulose xanthate known as viscose. Rayon fibers are formed by forcing the viscose through a small hole into an acid bath that regenerates the original cellulose while carbon disulfide is given off. To make cellophane, the viscose passes through a long slot into a bath of ammonium sulfate which causes it to coagulate. The coagulated viscose is then put into an acidic bath that returns the cellulose to its original, insoluble form. The cellophane is now clear. *See* MANUFACTURED FIBER.

The cellophane is then treated in a glycerol bath and dried. The glycerol acts like a plasticizer, making the dry cellophane less brittle. The cellophane may be coated with nitrocellulose or wax to make it impermeable to water vapor; it is coated with polyethylene or other materials to make it heat sealable for automated wrapping machines. Cellophane is typically 0.03 mm (0.001 in.) thick, is available in widths to 132 cm (52 in.), and can be made to be heat sealable from 82 to 177°C (180 to 350°F). *See* POLYMER.

Uses. Cellophane is used for packaging snacks, cookies, baked goods, candy, and tobacco products. Cellophane can be cast into a ring-shaped structure to be used as sausage casings (wiener), but paper with hemp fiber is used in large sausages. Cellophane enjoys many specialized uses based on its ability to let small molecules pass through it while retaining larger molecules, although some materials are replacing cellophane for these uses. It has been used in kidney dialysis machines, battery separators for silver-zinc cells, and semipermeable membranes. It is also used in denture compression molding. *See* WOOD CHEMICALS; WOOD PRODUCTS.

Christopher J. Biermann

Bibliography. C. J. Biermann, *Essentials of Pulping and Papermaking,* 1993; *Kirk-Othmer Encyclopedia of Chemical Technology,* 4th ed., vol. 5, 1992; A. P. Peck, Cellophane is born, Sci. Amer., 158: 274–275, May 1938; E. Sjöström, *Wood Chemistry: Fundamentals and Applications,* 2d ed., 1993.

Cellular automata

A cellular automaton is a hypothetical computer constructed of a regular array of cells that interact with neighboring cells according to a set of simple rules. Though conceptually simple, cellular automata can produce complex and surprising behavior. To computer scientists and mathematicians, cellular automata are of great interest from a theoretical perspective, forming one of the fundamental abstractions (along with Turing machines, finite-state automata, and so on) that are a basic model for all of computing. They are also an example of evolutionary computation, a set of computing techniques that are based on nature's laws of evolution and are important in the related field of artificial life, which involves interest in virtual (computer-based) life forms. To cellular automata hobbyists, they are a dynamic art form that often produces interesting results. *See* AUTOMATA THEORY; COMPUTER; EVOLUTIONARY COMPUTATION.

The word automaton comes to us from Greek through Latin, meaning something with the ability to act spontaneously. A cellular automaton is essentially a mathematical abstraction, such as a set or a function, but not a physical machine and not really alive. Nevertheless, when cellular automata are animated, either manually or through the graphical user interface of a computer program, they give a clear sense of complex animated behavior that makes them clearly lifelike. To computer scientists, the word automaton has a more technical meaning, that of a formal mathematical representation of computation.

Conventions and rules. A basic cellular automaton consists of a one- or two-dimensional array of cells, each of which can be in one of two states (black or white). The cells react to each other in a linear (two neighbors), orthogonal (four neighbors), or fully connected (eight neighbors, including the diagonals) fashion, based upon a set of rules that govern the cellular automaton (**Fig. 1**). The cells are synchronous, reacting in unison to their neighbors. During one reaction cycle, the cells assess their neighbors (two, four, or eight) and then reset their respective values in accordance with the universal rules. Over a series of reaction cycles, the cells can collectively exhibit complex, lifelike behavior.

Cellular automata gained prominence among mathematicians, both professional and amateur, with the publication in 1970 of Martin Gardner's *Scientific American* column "The fantastic combinations of John Conway's new solitaire game 'life.'" John Horton Conway, a mathematician at the University of Cambridge at the time, was well known for his contributions to theoretical mathematics and for his interest in recreational mathematics. Conway's "Game of Life" rules established an environment for the evolution of patterns (or populations, from an artificial life perspective) in cellular automata.

A typical set of rules for a cellular automaton comes from the original Game of Life. The rules refer to states of "death" and "life," which are represented by two alternative colors in the cells. These rules assume a cellular array with full connections (eight per cell):

1. If a live cell is surrounded by zero or one live neighbor, it dies from loneliness.
2. If a live cell is surrounded by four or more neighbors, it dies from overcrowding.
3. If a dead cell is surrounded by three neighbors, it comes back to life.

Additional rules are usually included to describe how the edge cells react, since they have fewer connections than the center cells. The cells are "seeded" initially by making a few of them live. Then, over a period of cycles, interesting behavior emerges from the collective. **Figure 2** shows a pattern over four cycles of evolution in a fully connected cellular automaton.

Cells can also be arranged into three-, four-, and higher-dimensional configurations, which can be theoretically interesting, though nearly impossible to visualize and therefore less interesting from a recreational perspective. Other rule sets are possible, including variations on the number of cells associated with loneliness, birth, death, and random variations (mutations). Cells may also have more than two coloring options, giving them the ability to generate colorful graphics.

The classical "Life" rules described above were developed through experimentation with the following goals in mind.

1. There should be no initial pattern for which there is a simple proof that the population can grow without limit.
2. There should be initial patterns that apparently do grow without limit.
3. There should be simple initial patterns that grow and change for a considerable period of time before coming to an end in three possible ways: fading away completely (from overcrowding or becoming too sparse), settling into a stable configuration that remains unchanged thereafter, or entering an oscillating phase in which the patterns repeat an endless cycle of two or more periods.

As a result, populations from the Game of Life exhibit behavior that is surprising and interesting and the underlying mathematics become interesting as well. Conway initially offered a conjecture that unrestricted growth was impossible under the standard Life rules. His supposition was that every possible seed pattern would eventually die, reach a static

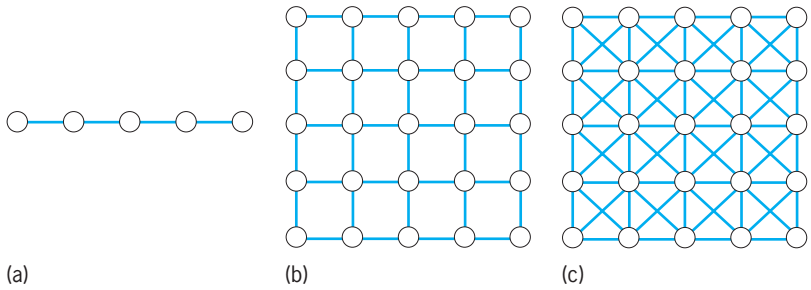

Fig. 1. Cellular automata connections: (*a*) linear, (*b*) orthogonal, and (*c*) full.

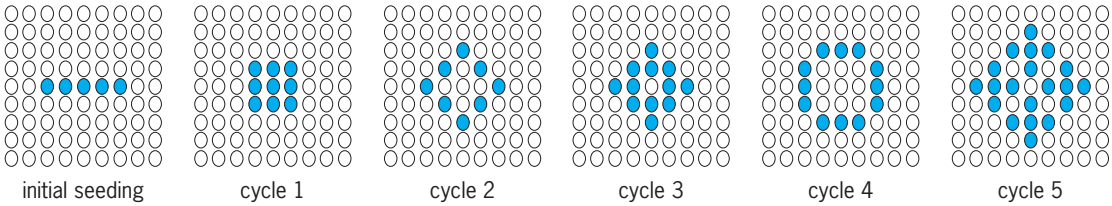

Fig. 2. Cellular automata in action.

state, or enter a permanent oscillation. He posed a reward of $50 for the first challenger to prove otherwise. The first of such patterns was found within a few months, and Conway's conjecture was disproved.

Self-replicating automata. Though Conway is credited as the originator of the Game of Life and is responsible for much of the recreational interest in cellular automata, John von Neumann, famous for his contributions to computer science and mathematics, first proposed cellular automata in 1943 as part of his research into self-replicating machines. As a young mathematician and professor, von Neumann made significant contributions in applied mathematics in a variety of fields, including game theory, fluid dynamics, and the development of the atomic bomb. During World War II, von Neumann became interested in computing and in formal models of computing in particular. He began to develop a comprehensive theory of organisms and machines, including their structure, organization, use of information, control, and language. With his automata theory, he aimed to develop a unified logical model that would explain both the organization of a computer and the operation of the human central nervous system. His theory also incorporated questions of complexity, reproduction, and self-repair. *See* CYBERNETICS; GAME THEORY.

As part of his attempt to model self-replicating automata, von Neumann developed several models of self-reproduction, both continuous and discrete, including one based on cellular automata. His self-reproducing cellular automaton formed a powerful universal computational model that could also reproduce itself within an infinite cellular automata space. In this model, the cells were arranged in an orthogonal lattice and operated on a uniform program of 29 rules. As a result of this development, von Neumann also showed that a computer structured as a cellular automaton was as powerful as any other computer and equivalent in power to the most well known model of computation, the Turing machine. Von Neumann died relatively young and was unable to complete his automata theory, but the work that he did complete was summarized and published by Arthur Burks.

Applications. Scientists in fields other than computer science have also discovered applications of cellular automata to their own research. For example, cellular automata have been used to investigate solutions to problems such as fluid dynamics, text data-compression algorithms, computational approaches to various problems in economics, tracking location of terminals in mobile computing networks, analyzing complex chemical reactions, modeling fundamental physics, and simulating the mechanics of a landslide. *See* ALGORITHM.

The above examples suggest only some of the many different ways in which cellular automata are being used to investigate problems in natural science, social science, mathematics, and computing. There are virtually no fields of modern science in which cellular automata have not been used or proposed to solve outstanding problems. One group even suggests that the best explanation of reality is that the universe is in fact one large collection of cellular automata. In any case, there is no apparent limit to the growing scientific interest in cellular automata and their potential importance to scientific discovery.

Ralph F. Grove, Jr.

Bibliography. M. Gardner, The fantastic combinations of John Conway's new solitaire game "life," *Sci. Amer.*, 223:120–123, 1970; M. Gardner, *Wheels, Life, and Other Mathematical Amusements*, W. H. Freeman, 1983; A. Ilachinski, *Cellular Automata: A Discrete Universe*, World Scientific, 2001; T. Toffoli and N. Margolus, *Cellular Automata Machines*, MIT Press, 1987; S. Wolfram, *A New Kind of Science*, Wolfram Media, 2002.

Cellular immunology

The field concerning the interactions among cells and molecules of the immune system, and how such interactions contribute to the recognition and elimination of pathogens. Humans (and vertebrates in general) possess a range of nonspecific mechanical and biochemical defenses against routinely encountered bacteria, parasites, viruses, and fungi. The skin, for example, is an effective physical barrier to infection. Basic chemical defenses are also present in blood, saliva, and tears, and on mucous membranes. These defenses are nonspecific in that they may be effective against a broad array of organisms. Nonspecific or innate defense mechanisms offer only limited protection against pathogenic organisms, which can proliferate rapidly and overwhelm these first-line defenses. True protection stems from the host's ability to mount responses targeted to specific organisms, and to retain a form of "memory" that results in a rapid, efficient response to a given organism upon a repeat encounter. This more formal sense of immunity, termed adaptive immunity, depends upon the coordinated activities of cells and molecules of the immune system.

Cells involved. Several types of cells play a role in protecting the body from infection, and they are

found primarily in the blood and lymph. Specific immune responses mainly involve the activities of T-lymphocytes and B-lymphocytes, two types of white blood cells. A response is initiated when a pathogen triggers the activity of one or both of the two major types of T cells: $CD4^+$ cells, also known as helper T cells (T_H); and $CD8^+$ cells, also known as cytotoxic T-lymphocytes (CTL) [see **illus.**].

When $CD4^+$ T cells are triggered, they release factors called cytokines, which in turn stimulate B-lymphocytes to make and secrete antibodies. Cytokines, principally the interleukins and interferons, are much like hormones, but they are secreted by individual cells rather than by organized tissues or glands. Cytokines play a significant role in many aspects of immune function and regulation, acting not only on B cells but also on T cells, macrophages, and other cells (thereby helping the immune response). Antibodies are molecules that are capable of binding to specific molecular structures found on microbes, and play a critical role in eliminating pathogens. Pathogen-bound antibodies can initiate a complement cascade (a series of reactions by serum proteins that burst and kill a microbe); they also serve as adapter molecules that facilitate the ingestion and destruction of microbes by phagocytic cells, such as macrophages. *See* ANTIBODY; COMPLEMENT; PHAGOCYTOSIS.

When $CD8^+$ T cells are triggered, they release factors that kill a cell harboring an infectious agent, and they also release cytokines. Virus-infected cells commonly are the targets of cytotoxic T-lymphocytes, since viruses need to get inside a cell in order to reproduce.

Antigen recognition. The immune system must detect a pathogen before a response can be made. This phase of the response is shown in the interaction between a T cell and an antigen-presenting cell (APC) [see illus.]. An antigen is a molecule, or portion of a molecule, that is recognized by a T-cell receptor (TCR) or antibody molecule. The cell surface molecules involved in antigen recognition by T cells are the T-cell receptor and class I or class II molecules of the major histocompatibility complex (MHC). Each T cell expresses a unique T-cell receptor that will interact specifically with a limited set of antigens. The antigens recognized by T-cell receptors are short peptides bound to MHC molecules. In addition to the T-cell receptor, the CD8 and CD4 molecules on the surface of T cells interact with class I and class II molecules, respectively, and define the MHC restriction of a given T cell; a circulating T cell expresses either CD8 or CD4, and it is poised to recognize peptides presented either by class I or by class II molecules. These co-receptor molecules do not engage in antigen recognition per se but help to stabilize the T cell–antigen-presenting cell interaction.

Antigen presentation. When a microbe enters or is ingested by a cell, some of the microbe's proteins are broken down by a process called proteolysis, just as cellular proteins are in the normal course of metabolism. Upon proteolysis, small fragments of protein, termed peptides, are generated. Class I and class II MHC molecules have a

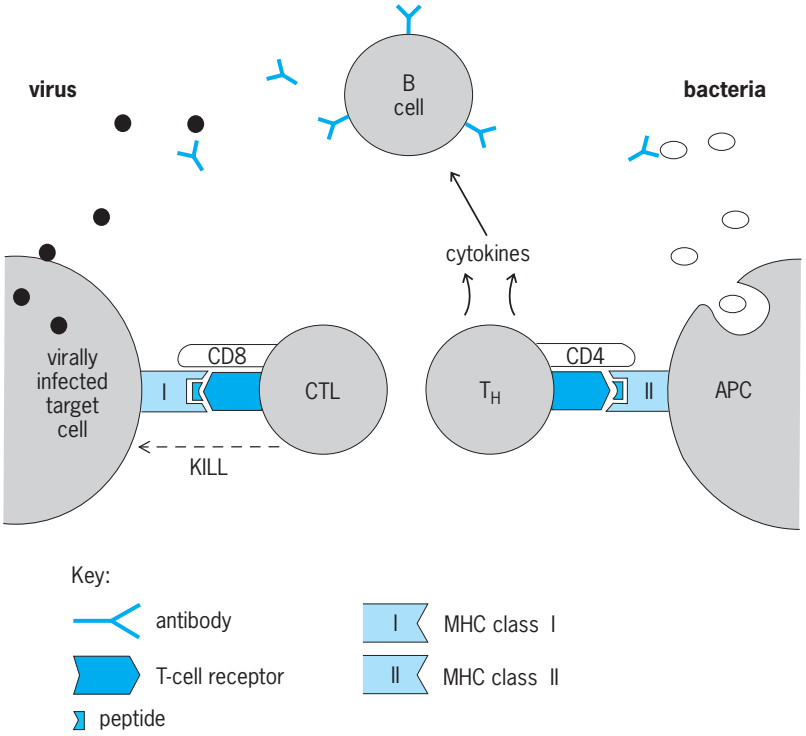

Key processes in a specific immune response.

grooved structure in their membrane-distal portions that can accommodate a short peptide generated by this process. MHC molecules are highly polymorphic (exist as many inherited varieties), and differences in this grooved domain determine the peptide-binding preferences exhibited by class I or class II molecules of a given type. A peptide that is bound to an MHC molecule forms an MHC-peptide complex, which is transported to the cell surface and displayed to roaming T cells. Most of these peptides derive from the body's own proteins, and do not elicit an immune response. T cells with receptors capable of recognizing self-peptides either are eliminated during their development in the thymus or are "trained" to ignore such peptides. However, when an infectious agent is present, some MHC-peptide complexes feature a pathogen-derived peptide. This peptide "message" alerts T cells (only those having a T-cell receptor capable of reading the particular message) to the presence of a microorganism. *See* THYMUS GLAND; TRANSPLANTATION BIOLOGY.

MHC molecules. Although class I and class II MHC molecules are structurally similar, they differ in the means by which they acquire peptides and in their patterns of cellular expression. Class I molecules can be expressed on nearly every cell type in the body, and their acquisition of peptides is optimized for the capture of peptides derived from proteins synthesized within the cell. This scheme is well suited to the presentation of virus-derived peptides, since viruses can enter many different types of cells, where they take over the biosynthetic machinery in order to produce their own proteins and replicate. Class I–peptide complexes alert $CD8^+$ T cells that a given cell is infected, and mark it for destruction.

Class II molecules, in contrast, are expressed on a

limited variety of cell types: lymphoid cells, dendritic cells, macrophages, and endothelium. Their expression on other cell types is inducible by gamma interferon. Peptide acquisition by class II molecules is biased toward the capture of peptides derived from extracellular sources, brought into the cell by phagocytosis and receptor-mediated uptake. Most bacteria thrive extracellularly, and viruses spend a portion of their cycle outside the cell. Cells presenting antigenic class II–peptide complexes generally are not marked for destruction, but stimulate $CD4^+$ cells to promote responses (such as antibody production) geared toward the elimination of extracellular pathogens. *See* HISTOCOMPATIBILITY.

Co-stimulation. Antigen recognition, as mediated by the T-cell receptor–MHC–peptide interaction, is necessary, but generally it is not sufficient for the initiation of an immune response. Several other molecular interactions occur between molecules on the T-cell surface and those on the antigen-presenting cells. Among these is the CD28-B7 interaction, which currently is understood to provide a critical co-stimulatory signal in the initiation of a T-cell response. CD28 and B7 molecules reside on the T-cell and antigen-presenting-cell surfaces, respectively. B7 represents a family of molecules: B7-1 (CD80) and B7-2 (CD86) are two members of this family that are ligands for the CD28 molecule. B7 molecules typically are expressed on cells referred to as professional antigen-presenting cells, such as activated dendritic cells and macrophages.

The co-stimulation provided by these additional interactions drives the production of cytokines by T cells and induces their proliferation. Because B7 molecule expression largely is restricted to professional antigen-presenting cells, this requirement for a second signal (antigen provides the first signal) can ensure that the immune response is not misdirected. Typically, T cells receive both signals from an antigen-presenting cell only in the context of an inflammatory reaction, initiated in part by innate defense mechanisms triggered by a pathogen. This context dependence is an important means for preventing T cells from reacting in a hair-trigger fashion to the presence of harmless foreign proteins (such as dietary proteins) or against self-proteins in cases where tolerance may be incomplete. *See* AUTOIMMUNITY.

Memory. Once an immune response has been initiated, T cells and B cells proliferate and become mature responder cells. These cells do not have the same requirement for co-stimulation once a response is under way; a mature cytotoxic T-lymphocyte, for example, can kill an infected cell if the cell simply presents the appropriate peptide. As pathogen elimination nears completion, many of the T and B cells involved in the response die. However, a subset of cells remain as memory cells, which can be quickly called into action if the same pathogen is encountered on a future occasion. The mechanisms that signal some cells to become memory cells and remain until needed are not well understood at present. However, it is this anamnestic feature of the immune response that gives humans the familiar experience of having a single nonrecurring instance of childhood diseases such as mumps and chicken pox. Although there may be reexposure to the pathogens causing those illnesses, the primed immune system can deal with a repeat threat so efficiently that it goes largely unnoticed. *See* ACQUIRED IMMUNOLOGICAL TOLERANCE; CELLULAR IMMUNOLOGY; CYTOKINE; IMMUNITY; IMMUNOLOGY.

Douglas J. Loftus

Bibliography. E. A. Greenfield, K. H. Nguyen, and V. K. Kuchroo, CD28/B7 costimulation: A review, *Crit. Rev. Immunol.*, 18:389–418, 1998; C. A. Janeway et al., *Immunobiology: The Immune System in Health and Disease*, 4th ed., 1999.

Cellulose

A linear polymer of (1→4)-linked β-D-glucopyranosyl units, and the most abundant of all naturally occurring substances. Cellulose constitutes approximately a third of all vegetable matter and thus it exists in far greater quantity than any other polysaccharide. It occurs as a principal structural component of the cell walls of mosses and seaweeds (25–30%), annual plants (25–35%), and trees (40–50%); cotton fiber contains 98% cellulose along with 1% protein, 0.65% pectic substance, and 0.15% mineral matter. Cellulose also is produced by some microorganisms, in a few cases reaching amounts of 20–30%. *See* CELL WALLS (PLANT).

Source. Cellulose for chemical modification, particularly for derivatization, is often obtained from cotton linters by boiling them with 1% sodium hydroxide solution. High-quality cellulose may be obtained from bast fibers such as flax (80–90% cellulose), hemp (65–75%), jute (60–70%), and ramie (85%). Cellulose fibers used for paper, for cardboard, or for conversion to film and synthetic fiber are obtained from wood pulp.

Structure. While cellulose is a uniform, linear polymer of a β-D-glucopyranosyl units linked (1→4), the β-D-glucosidic bond causes alternate units to be positioned as shown in the **illustration**, so that the molecule is essentially a polymer of the disaccharide, cellobiose, actually the cellobiosyl unit. The molecule has a nonreducing end and an aldehyde end, although at times the latter may be oxidized to a carboxyl group. Chain length varies with previous treatment, but native cellulose molecules have 7000 to 15,000 units, which are termed degrees of polymerization. *See* POLYMER.

Cellulose is the principal structural component of plants. The structural strength and integrity of plant cell walls that support even giant trees is due to the ease with which cellulose molecules fit together and bond intermolecularly, with secondary forces, over long lengths. This crystallization is due, for the most part, to hydrogen bonds such as that between the C3 hydroxyl of one chain and the C6 hydroxyl of a neighboring chain. Crystallinity may account for 60–70% of the cellulose in wood pulp and 70–80% in cotton. Rayon regenerated from pulp may redevelop 45% crystallinity. Cellulose in the cell wall is mainly in threadlike fibers, composed of smaller fibrils. These provide the reinforcing material with non-

crystalline molecules running from one crystallite to another, providing elasticity in conjunction with that of hemicellulose and other filler material of the cell wall. *See* HEMICELLULOSE; HYDROGEN BOND.

Various crystalline forms of cellulose exist with the native form known as cellulose I. The unit cell is monoclinic with three unequal sides and one non-90° angle. Chains are skewed 180° along the chain axis. Strong alkali alters the polymorph to cellulose II, providing greater hydrogen bonding and a more stable, lower-energy, structure. Some covalent bonding occurs between cellulose and with the lignin of the cell wall and of the middle lamella, but mainly these molecules are free and can be extracted with reageants such as a strong alkaline solution.

Pulp. Commercial cellulose is obtained from woods by pulping procedures designed to remove lignin, hemicelluloses, gums, waxes, and other natural noncellulose constituents. Most commercial wood pulp is consumed in paper manufacture, but pulps of high α-cellulose content (90% or greater for rayon and at least 98% for cellulose ester, mainly acetate, manufacture) are considered good chemical cellulose, or dissolving-grade cellulose. α-Cellulose is that which is not soluble in 17.5% sodium hydroxide solution. Though cellulose is resistant to degradation during the pulping operation, it is partially degraded to lower molecular weights. Pulping is accomplished by mechanical or chemical means, or a combination of the two. Mechanical pulps contain the same constituents as the original wood, but fewer water-soluble substances. These pulps have shorter fibers than do chemical pulps and produce weaker paper. If steam is used to produce a thermomechanical pulp, the fibers are less degraded.

The kraft (sulfate) method is the most common chemical pulping method used in the United States, accounting for 70% of pulp production; pulp yield is 40–45%. In this process, wood chips, $^5/_8$–1 in. (1.5–2.5 cm) long by $^1/_4$ in. (0.6 cm) thick, are cooked to a maximum temperature of 334–347°F (168–175°C) for 2–6 h with a solution of sodium hydroxide and sodium sulfide, present in a ratio of 3:1. Other processes used are the sulfite and alkaline. Less than 10% of the pulp made in the United States is produced by the sulfite process. Chemical methods have become focused on improved yield, reduced environmental problems associated with sulfur chemicals, and production of high-strength fibers. Modern processes add anthraquinone to a soda (alkaline) cook, introduce soda-oxygen cooking, or use organic solvent extraction. Combinations of mechanical and chemical pulping tend to produce higher yields of high-strength fibers. *See* PAPER; WOOD CHEMICALS.

Dissolving pulp of 95–98% α-cellulose is used for viscous rayon, cellulose nonwoven materials, and cellulose derivatives. These pulps are made by alkaline washing of sulfite or sulfate pulps or by mild acid hydrolysis of chips followed by kraft pulping and bleaching to give low pulp yields of 29–35%. *See* TEXTILE CHEMISTRY.

Derivatives. Commercial derivatives are ethers that, among other properties, provide water solubility and esters which, in high degrees of substitution, provide organic solubility needed for spinning cellulose into fibers or molding it into shaped objects. One inorganic ester, the nitrate, made by reacting cellulose in a mixture of nitric and sulfuric acids, is used at high derivation as cellulose trinitrate gunpowder. Cellulose trinitrate has a degree of substitution (DS; the average number of hydroxyls substituted per glucose residue) of three. Formerly, it was used at a lower degree of substitution as a lacquer and film former. *See* ESTER; ETHER; MANUFACTURED FIBER.

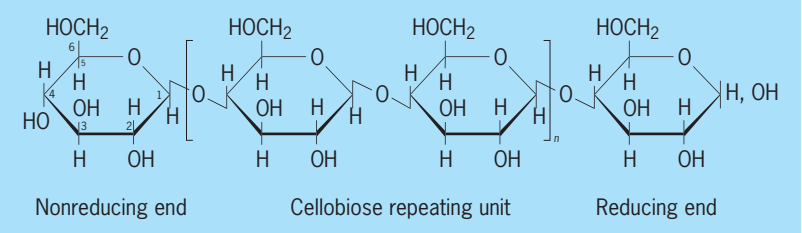

Positioning of units in cellulose structure.

Commercial ethers are carboxymethyl, methyl, hydroxyethyl, or hydroxypropyl. Mixed ethers are also made that contain methyl and hydroxyethyl or hydroxypropyl groups. All are soluble in water or mild alkaline solution. Solubility is conferred on the linear cellulose molecule by the substituent groups, which provide frequent bumps on the regular cellulose chains. This restricts them from associating when sections of chains meet in solution, thus preventing close association of chains that would lead to partial crystallization and insolubilization. Since these rather inflexible chains are still quite long with degrees of polymerization of 150–250, they sweep out large spherical volumes as they gyrate, and they collide frequently to absorb energy, producing highly viscous solutions. Such derivatives have degrees of substitution of 0.1–2.5. They are used to thicken solutions and to act as suspending agents and sometimes as mild emulsifiers.

Carboxymethylcellulose has a negative charge due to ionization of the carboxyl group, giving the molecule some emulsifying character. This derivative is used in washing powders; the molecules bind slightly to cotton fabrics, giving the cloth a negative charge that prevents negative soil particles from redepositing on the cloth.

Cellulose acetate is the principal commercial ester. Reaction is usually conducted in a mixture of acetic anhydride, acetic acid, and sulfuric acid as the catalyst to produce a degree of substitution of about 2.4, resulting in a product soluble in acetone. This has applications as a quick-drying adhesive, for film formation, or for use in cigarette filters. If proprionic acid or its anhydride is used in the esterification reaction, a mixed ester is produced that forms a shock-resistant plastic useful for football helmets, telephones, and other objects that need to withstand shock. A fully acetylated cellulose (degree of substitution = 3) is used for making high-strength products, acetate rayon, and textile tow. This acetate can be made by acetylation of cellulose in dichloromethane

with acetic anhydride and perchloric acid catalyst.

Degradation. Because cellulose molecules bind together so strongly, they are quite resistant to hydrolysis by acids and even cellulytic enzymes, although the free molecules or amorphous regions are hydrolyzed as easily as other polysaccharides. The end product is D-glucose; if it could be produced cheaply from cellulose, it would allow wood to be used as a source of fermentation alcohol. Hot 12% hydrochloric acid converts hemicelluloses mainly to furfural and cellulose to hydroxmethylfurfural. Unless cellulose molecules are oxidized, they are resistant to alkaline solutions, because like other polysaccharides the aldehyde functions of the D-glucopyranosyl units are tied up as acetal linkages that are resistant to alkaline cleavage.

Cellulose undergoes oxidation by numerous oxidants such as peroxides and hypochlorite. Dinitrogen tetraoxide specifically converts the hydroxymethyl group at C6 of each D-glucopyranosyl unit to a carboxyl or uronic acid. Most oxidations are catalyzed by metal ions. Therefore, iron salts allow oxidation of cotton textiles, causing yellow spots and eventually holes.

Cellulose is hydrolyzed slowly by cellulase enzymes produced by bacteria and fungi. *See* POLYSACCHARIDE.

David G. Barkalow; Roy L. Whistler

Bibliography. J. P. Casey (ed.), *Pulp and Paper Chemistry and Chemical Technology*, vol. 1, 3d ed., 1980; J. F. Kennedy et al. (eds.), *Cellulose and Its Derivatives: Chemistry, Biochemistry and Applications*, 1985; T. P. Nevell and S. H. Zeronian (eds.), *Cellulose Chemistry and Its Applications*, 1985; C. E. Shuerch (ed.), *Cellulose and Wood: Chemistry and Technology*, 1989; E. Sjostrom, *Wood Chemistry: Fundamentals and Applications*, 1993; R. A. Young and R. M. Rowell (eds.), *Cellulose Structure, Modification and Hydrolysis*, 1986.

Cement

A material, usually finely divided, that when mixed with water forms a paste, and when molded sets into a solid mass. The term cement is sometimes used to refer to organic compounds used for adhering or for fastening materials, but these are more correctly known as adhesives. *See* ADHESIVE; ADHESIVE BONDING.

In the fields of architecture, engineering, and construction, the term portland cement is applied to most of the hydraulic cements used for concrete, mortars, and grouts. This article discusses portland cement and other special construction cements.

Portland cement sets and hardens by reacting chemically with water. In concrete, it combines with water and aggregates (sand and gravel, crushed stone, or other granular material) to form a stonelike mass. In grouts and mortars, cement is mixed with water and fine aggregates (sand) or fine granular materials. *See* CLAY, COMMERCIAL; CONCRETE; MORTAR.

Production. Portland cement is ground from nodules, which can vary in dimension from dust to fist-sized chunks, known as clinker. Cement clinker is the product of a high-temperature process. Manufacturing entails introduction of finely ground raw materials—mineral sources of calcium oxide, silica, alumina, and iron oxide—into a horizontally sloping rotary kiln. Raw materials are ground to a fineness to pass a 75-micrometer (200-mesh) screen in order to react within the kiln's residence time.

Materials are introduced either in powdered form, known as the dry process, or in slurry form, known as the wet process. Less energy is required to produce cement in the dry process; energy conservation efforts have led to the phasing-out of wet-process systems. Kiln systems differ, but generally materials are funneled through three major zones: drying and preheating, 70–1650°F (20–900°C), in which water is evaporated and calcination occurs; calcining, 1100–1650°F (600–900°C), where carbon dioxide is driven off but lime needed for subsequent reactions remains; and sintering or burning, 2200–2700°F (1200–1480°C), in which calcium oxide reacts with silica to form calcium silicates.

The intense heat yields clinker, whose principal chemical compounds are tricalcium silicate, dicalcium silicate, tricalcium aluminate, and tetracalcium aluminoferrite. Raw-material sources can vary according to local availability, but generally bear lime, iron, silica, alumina, and magnesia. Sources of these materials include limestone, clay, shale, and slag. Cement clinker is ground to extreme fineness in rotary steel-ball mills. Approximately 85–95% of cement particles are smaller than 45 micrometers. Small amounts of gypsum are interground with clinker to control set time.

Calcium silicates constitute about 75% of the weight of cement. When mixed with water, they form calcium hydroxide and calcium silicate hydrate or tobermorite gel. Hydrated cement contains about 25% calcium hydroxide and 50% tobermorite gel. The latter is primarily responsible for the strength and other properties of hydrated cement.

Standards. Under its C 150 classification, the American Society for Testing and Materials (ASTM) designates five basic types of portland cement. Type I is a general-purpose cement suitable for all uses where special properties are not required. Type II is used for resistance to moderate sulfate exposures. Type III is high-early-strength cement. Type IV is used where the rate and amount of heat generated from hydration must be minimized. Type V is used for resistance to severe sulfate exposure.

Types I, II, and III are also produced in air-entraining varieties. Air entrainment (incorporation of tiny air bubbles) is employed to maintain workability and cohesiveness in the fresh concrete mix and provide durability in the hardened product in climates where the material is subject to cycles of freezing and thawing. The microscopic air bubbles in the mix act as pressure-relief valves during freeze cycles. Air entrainment in a concrete mix can be acheived by using the three air-entraining portland cements (types IA, IIA, IIIA) or admixtures, which impart special properties to fresh or hardened concrete.

Adjustments in the physical and chemical compositions allow for tailoring portland cements and other hydraulic cements to special applications. Blended hydraulic cements are produced with portland cements and materials that by themselves might not possess binding characteristics.

Special varieties. Special cements are produced for mortars and architectural or engineering applications: white portland cement, masonry cement, and oil-well cement, expansive cement, and plastic cement.

In addition to acting as the key ingredient in concrete, mortars, and grouts, portland cements are specified for soil-cement and roller-compacted concrete, used in pavements and in dams, and other water resource structures, and as reagents for stabilization and solidification of organic and inorganic wastes. Don Marsh

Bibliography. S. H. Kosmatka and W. C. Panarese, *Design and Control of Concrete Mixtures*, 13th ed., 1988; Portland Cement Association, *A New Stone Age: The Making of Portland Cement and Concrete*, 1992.

Cenozoic

Cenozoic (Cainozoic) is the youngest and the shortest of the three Phanerozoic geological eras. It represents the geological time (and rocks deposited during that time) extending from the end of the Mesozoic Era to the present day.

The geological concept of the Cenozoic, as the youngest era of the Phanerozoic, was introduced by J. Phillips in 1941. He considered it to be a unit equivalent to the Tertiary, a term that was still in use from G. Arduino's classification of Primary, Secondary, and Tertiary rocks, introduced in Italy in 1759. Arduino based his subdivision on physical attributes, such as older magmatic and metamorphic rocks as Primary; limestone, marl, and clay with fossils as Secondary; and youngest, fossil-rich rocks as Tertiary. This classification is now largely obsolete, with the exception of the term Tertiary that is still in use in a modified sense. A. Brogninart first modified the concept of *Tertiaire* in 1810, applying it to the strata deposited above the Cretaceous chalk in the Paris Basin. The term Quaternary (the second and younger period of the Cenozoic) was introduced by M. Morlot in 1854. *See* QUATERNARY; TERTIARY.

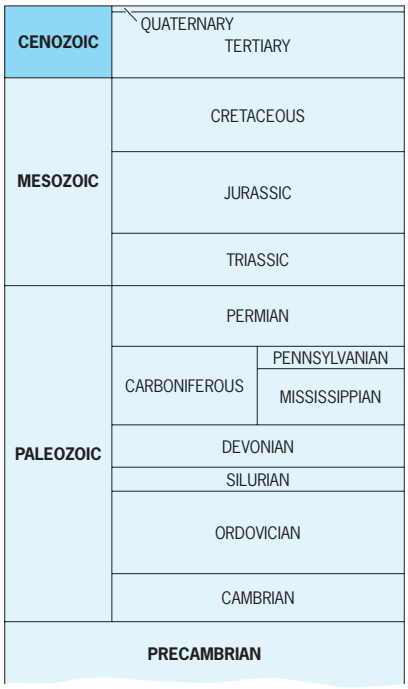

Modern time scales include all of the past 65 million years of geological history in the Cenozoic Era. The distinction of the Cenozoic Era from older eras has been traditionally based on the occurrence of fossils showing affinities to modern organisms, and not on any particular lithostratigraphic criteria. *See* INDEX FOSSIL; PALEONTOLOGY.

Subdivisions. Traditional classifications subdivide the Cenozoic Era into two periods (Tertiary and Quaternary) and seven epochs (from oldest to youngest): Paleocene, Eocene, Oligocene, Miocene, Pliocene, Pleistocene, and Holocene. The older five epochs, which together constitute the Tertiary Period, span the time interval from 65 to 1.8 million years before present. The Tertiary is often separated into two subperiods, the Paleogene (Paleocene through Oligocene epochs, also collectively called the Nummulitic in older European literature) and the Neogene (Miocene and Pliocene epochs). These subperiods were introduced by M. Hornes in 1853. The Quaternary Period, which encompasses only the last 1.8 million years, includes the two youngest epochs (Pleistocene and Holocene). Holocene is also often referred to as the Recent, from the old Lyellian classification. Recent stratigraphic opinions are leaning toward abandoning the use of Tertiary and Quaternary (which are seen as the unnecessary holdovers from obsolete classifications) and in favor of retaining Paleogene and Neogene as the prime subdivisions of Cenozoic. *See* EOCENE; HOLOCENE; MIOCENE; OLIGOCENE; PALEOCENE; PLEISTOCENE; PLIOCENE.

Tectonics. Many of the tectonic events (mountain-building episodes or orogenies, changes in the rates of sea-floor spreading, or tectonic plate convergences) that began in the Mesozoic continued into the Cenozoic. The Laramide orogeny that uplifted the Rocky Mountains in North America, which began as early as Late Jurassic, continued into the Cretaceous and early Cenozoic time. In its post-Cretaceous phase the orogeny comprised a series of diastrophic movements that deformed the crust until some 50 million years ago, when it ended abruptly. The Alpine orogeny, which created much of the Alps, also began in the Mesozoic, but it was most intense in the Cenozoic when European and African plates converged at an increased pace. *See* CRETACEOUS; JURASSIC; MESOZOIC; OROGENY.

In the Pacific Ocean the most significant tectonic event in the Cenozoic may have been the progressive consumption of the East Pacific Rise at the Cordilleran Subduction Zone and the concomitant development of the San Andreas Fault System some

30 million years ago. *See* CORDILLERAN BELT; FAULT AND FAULT STRUCTURES; SUBDUCTION ZONES.

In the intracontinental region of the Tethys between Europe and Africa, the Cenozoic tectonic history is one of successive fragmentation and collision of minor plates and eventual convergence of the African and European plates. Africa's motion was counterclockwise relative to Europe, which began sometime in the early Mesozoic. By mid-Cenozoic, however, the motion between the two plates was largely convergent. The convergence caused a complex series of events that closed the Tethys Seaway between the two continents. In the late Cenozoic (Pliocene) the final collision of the Arabian plate with the Asian plate along Iran produced the Zagros Mountains, and partially or completely isolated the Caspian and Black seas. *See* CONTINENTAL DRIFT.

In the Indian Ocean, perhaps the most significant event during the Cenozoic was rapid movement of the Indian plate northward and its collision with Asia. India had already broken loose from the eastern Gondwana in the late Cretaceous, but around 80 million years ago its motion accelerated, and then increased further in the early Cenozoic. This movement, however, slowed down considerably around 50 million years ago when the Indian plate plowed into the Asian mainland. The first encounter of the two plates caused the initial uplift of the Himalayas. The major phase of the Himalayan orogeny, however, extends from the Miocene to the Pleistocene, when much of the high Himalayas were raised and the Tibetan Plateau was fully uplifted. The encounter also caused major reorganization of the crust both north and south of the collision zone. In the Indian Ocean the plates were reorganized, and spreading was initiated on the Central Indian Ridge system. The collision had important repercussions for the Asian mainland as well. Over 1500 km (900 mi) of crustal shortening has occurred since the collision began. The effects of diastrophism associated with this event extend some 3000 km (1800 mi) northeast of the Himalayas. This includes major strike-slip faults in China and Mongolia, which may account for a major portion of the crustal shortening, which continues to the present day. The present convergence between the Indian and Asian plates is at the rate of about 0.5 cm (0.2 in.) per year.

Another major long-term affect of the tectonic uplift of Tibetan Plateau, which is dated to have been significant by 40 million years ago, may have been the initiation of the general global cooling trend that followed this event. The uplifted plateau may have initiated a stronger deflection of the atmospheric jet stream, strengthening of the summer monsoon, and increased rainfall and weathering in the Himalayas. Increased weathering and dissolution of carbonate rock results in greater carbon dioxide (CO_2) drawdown from the atmosphere. The decreased partial pressure of carbon dioxide (pCO_2) levels may have ultimately led to the Earth entering into a renewed glacial phase.

The convergence between India and Asia and between Africa and Europe in the mid-Cenozoic destroyed the ancestral Tethys Seaway, leaving behind smaller remnants that include the Mediterranean, Black, and Caspian seas. *See* PLATE TECTONICS.

Oceans and climate. The modern circulation and vertical structure of the oceans and the predominantly glacial mode that the Earth is in at present was initiated in the mid-Cenozoic time. The early Cenozoic was a period of transition between the predominantly thermospheric circulation of the Mesozoic and the thermohaline circulation that developed in the mid-Cenozoic. By the mid-Cenozoic the higher latitudes had begun to cool down, especially in the Southern Hemisphere due to the geographic isolation of Antarctica, leading to steeper latitudinal thermal gradients and accentuation of seasonality. The refrigeration of the polar regions gave rise to the cold high-latitude water that sank to form cold bottom water. The development of the psychrosphere (cold deeper layer of the ocean) and the onset of thermohaline circulation are considered to be the most significant events of Cenozoic ocean history, which ushered the Earth into its modern glacial-interglacial cyclic mode.

The overall history of the Cenozoic oceans is marked by a long-term withdrawal of the seas from epicontinental and coastal oceans and the accretion of ice on the polar regions. The ice buildup may have been partly favored by the more poleward position of the landmasses and the eventual thermal isolation of the Antarctic continent.

In the early Paleogene, for the first time a deep connection between the North and South Atlantic was developed to allow deeper water penetration into the southern basin that intermittently led to extensive erosion on the ocean floor. However, the most likely source area for deep waters in the early Cenozoic was still in the temperate and low-latitude shelves. Farther north, in the Norwegian Sea area there is magnetic evidence of initiation of seafloor spreading in the late Paleocene. By the middle Eocene, this area may have become a site for the formation of cold bottom waters. Erosional events on the sea floor indicate that North Atlantic Deep Water may have begun to flow southward at approximately the same time as the initial subsidence of the Greenland-Faeroe Ridge below sea level in the late Eocene. Later on, in the Miocene, the ridge subsided more actively, resulting in greater outflow of higher-salinity water to intermediate and abyssal depths of the Central Atlantic. Evidence also points to vigorous bottom waters in the late Eocene, which increased in intensity through Oligocene and Miocene time.

The most prominent feature of the surface circulation in the early Cenozoic was the westward flowing circumglobal Tethys Current, which dominated the oceanic scene in the tropical latitudes. As the Indian plate approached the Asian mainland in the early Cenozoic, it progressively restricted the flow of Tethys Current to its north. In the middle Eocene, when the general drop in global sea level and the first encounter of the Indian plate with Asia reduced the northern passage, the main flow moved to the west of the Indian plate. By Oligocene time the

westward flow in the Tethys had become intermittent and severely restricted to a narrow western passage. The Tethyan passage had essentially closed by the dawn of the Neogene.

A paleogeographic event of major import for the overall Cenozoic oceanic patterns was the breaching of the straits between Antarctica and South America at the Drake Passage in the Oligocene. This event led to the development of the circum-Antarctic Current, eventual thermal isolation of Antarctica, and further enhancement of the ice cap on the continent by mid-Oligocene. Winter ice accumulation on the Arctic, which lacks a continent in the polar position, may have also begun by the Oligocene. The increased sequestration of water on ice caps may be responsible for a major global sea-level drop in the mid-Oligocene evidenced along most of the world's continental margins. By late Oligocene time the global surface circulation patterns had essentially evolved the major features of the modern oceans. *See* PALEOCEANOGRAPHY; PALEOGEOGRAPHY.

All climatic indicators point to a general warming trend through the Paleocene, culminating in a period of peak global temperatures at the close of Paleocene. The warm climates continued into the early Eocene interval. The latitudinal and vertical thermal gradients in the late Paleocene–early Eocene were low, and mean surface temperature was around 10°C (50°F) in the higher latitudes and 20°C (70°F) in the tropics. Terrestrial flora and fauna also corroborate the peak warming of this interval. For example, the Arctic island of Ellesmere has yielded a rich warm-blooded vertebrate fauna that indicates a range of temperature between 10 and 20°C (50 and 70°F).

Studies have revealed a prominent carbon-isotopic shift in global carbonate reservoir that coincides with the latest Paleocene peak warming. This has been ascribed to the breakdown of deposits of methane hydrates on continental margins and catastrophic release of methane into the water and atmosphere due to rapid warming of the bottom waters. In the latest Paleocene, bottom water temperature increased rapidly (in less than 10,000 years) by as much as 4°C (7.2°F), with a coincident prominent enrichment of ^{12}C isotope of the global carbon reservoir. The isotopic changes are accompanied by important biotic changes in the oceanic microfauna and are synchronous in the oceans and on land. This rapid and prominent isotopic shift cannot be explained by increased volcanic emissions of carbon dioxide, changes in oceanic circulation, or terrestrial and marine productivity alone. Increased flux of methane from gas-hydrate sources into the ocean-atmosphere system and its subsequent oxidation to carbon dioxide is held responsible for this isotopic excursion in the inorganic carbon reservoir. High-resolution data support the gas-hydrate connection to latest Paleocene abrupt climate change. Evidence from two widely separated sites from the low- and high-latitude Atlantic Ocean indicates multiple injections of methane with global consequences during the relatively short interval at the end of the Paleocene.

The Eocene time is also characterized by higher global sea levels and increased oceanic productivity and carbonate deposition on the shelves and banks. The climate became more extreme in the late Eocene and through the Oligocene, when the latitudinal contrast increased due to development of ice in the polar regions. Himalayan uplift that produced the obstruction of the Tibetan Plateau in the path of the jet stream in the late Eocene may have been an important contributory factor to the cooling trend of the mid and late Cenozoic. Vertical thermal gradients also steepened once the cold bottom waters outflow began from the higher latitudes.

The late Cenozoic is characterized by further accentuation of the oceanographic and climatic patterns that were initiated in the early Cenozoic. By Miocene time the Tethyan connection between the Indian Ocean and the Mediterranean Sea had been broken. This event modified the circulation patterns in the North Atlantic and the Mediterranean, which for the first time began to resemble their modern analogs. Neogene climatic proxies (such as stable isotopes in cores and glacial records on land) show evidence of considerable climatic fluctuations. Six major climatic deterioration events have been identified in the Miocene-Pliocene record. These events were also associated with enhanced surface circulation, bottom erosion, and aridity on land over North Africa. The cooler early Miocene climates were followed by a climatic optimum in mid-Miocene, to be in turn followed by a significant deterioration in climate, which has been ascribed to a major enlargement of the ice sheets on Antarctica.

In the late Miocene, the Mediterranean suffered a salinity crisis following a sea-level fall and the isolation of the basin. The growth of ice caps in the Miocene eventually led to the fall of sea level below the depth of Gibraltar Sill, isolating the Mediterranean Basin. The lack of connection to the open Atlantic and excess evaporation led to high salinities and deposition of voluminous quantities of evaporites in a relatively short time. The Mediterranean was reconnected to the Atlantic in the early Pliocene, allowing cold deep waters to suddenly spill over the subsided Gibraltar Sill. The Black Sea was also converted into an alkaline lake during the salinity crisis when the Mediterranean inflow was cut off. The early Pliocene reconnection with the Mediterranean was once again followed by isolation of the Black Sea, this time as a fresh-water lake. These conditions lasted into the Quaternary, when the reconnection to the Mediterranean was established through the Bosporus Straits around 7000 years ago, ushering in the present-day conditions. *See* SALINE EVAPORITES.

Another important threshold event of the late Cenozoic was the closing of the connection between the Central Atlantic and Pacific oceans at the Isthmus of Panama. The connection was operative until the mid-Pliocene, when tectonic events led to its closure some 3 million years ago. The closure most likely led to more vigorous Gulf Stream flow due to deflected energy, displacing the stream northward to its present-day position. The modern circulation

patterns in the Caribbean also date back to this event. Another major climatic-oceanographic event of the Cenozoic was the development of an extensive ice cap on the Arctic. Although there is some evidence of ice cover in the Arctic since the Oligocene, evidence of a significant amount of ice accumulation is only as old as mid-Pliocene, some 3 million years ago, coincident with the closing of the Panama isthmus. The deflection of the Gulf Stream northward due to the latter event may have provided the excess moisture needed for this accumulation.

The Quaternary climatic history is one of repeated alternations between glacial and interglacial periods. At least five major glacial cycles have been identified in the Quaternary of northwestern Europe. The most recent glacial event occurred between 30,000 and 18,000 years ago when much of North America and northern Europe was covered with extensive ice sheets. The late Pliocene and Pleistocene glacial cyclicity led to repeated falls in global sea level as a result of sequestration of water as ice sheets in higher latitudes during the glacial intervals. For example, the sea level is estimated to have risen some 110 m (360 ft) since the end of the last glacial maximum. As a by-product of these repeated drops in sea level and movement of the shorelines toward the basins, large deltas developed at the mouths of the world's major drainage systems during the Quaternary. These bodies of sand and silt constitute ideal reservoirs for hydrocarbon accumulation. *See* DELTA; PALEOCLIMATOLOGY.

Life. At the end of the Cretaceous a major extinction event had decimated marine biota and only a few species survived into the Cenozoic. The recovery, however, was relatively rapid. During the Paleocene through middle Eocene interval, the overall global sea-level rise enlarged the ecospace for marine organisms, and an associated climatic optimum led to increased speciation through the Paleocene, culminating in high marine diversities during the early and middle Eocene. Limestone-building coral reefs were also widespread in the tropical-temperate climatic belt of the early Cenozoic, and the tropical Tethyan margins were typified by expansive distribution of the larger foraminifera known as *Nummulites* (giving the Paleogene its informal name of the Nummulitic period). *See* NUMMULITES.

The late Eocene saw a rapid decline in diversities of marine phyto- and zooplankton due to a global withdrawal of the seas from the continental margins and the ensuing deterioration in climate. Marine diversities reached a new low in the mid-Oligocene, when the sea level was at its lowest, having gone through a major withdrawal of seas from the continental margins. The climates associated with low seas were extreme and much less conducive to biotic diversification. The late Oligocene and Neogene as a whole constitute an interval characterized by increasing partitioning of ecological nichesinto tropical, temperate, and higher-latitude climatic belts,and greater differentiation of marine fauna and flora.

The terminal Cretaceous event had also decimated the terrestrial biota. Dinosaurs, which had dominated the Mesozoic scene, became extinct. Only a few small shrewlike mammalian species survived into the Paleocene. In the absence of dinosaurian competition, mammals evolved and spread rapidly to become dominant in the Cenozoic. The evolution of grasses in the early Eocene and the wide distribution of grasslands thereafter may have been catalytic in the diversification of browsing mammals. Marsupials and insectivores as well as rodents (which first appeared in the Eocene) diversified rapidly, as did primates, carnivores, and ungulates. The ancestral horse first appeared in the early Eocene in North America, where its lineage evolved into the modern genus *Equus*,only to disappear from the continent in the late Pleistocene. A complete evolution of the horse can be followed in North America during the Cenozoic. Increase in overall size, reduction in the number of toes, and increasing complexity of grinding surface of the molars over time are some of the obvious trends. Hominoid evolution began during the Miocene in Africa. Modern hominids are known to have branched off from the hominoids some 5 million years ago. Over the next 4.5 million years the hominids went through several evolutionary stages to finally evolve into archaic *Homo sapiens* about 1 million years ago. Truly modern *Homosapiens* do not enter the scene until around 100 thousand years ago. *See* DINOSAURIA; FOSSIL HUMANS; MAMMALIA; ORGANIC EVOLUTION. Bilal U. Haq

Bibliography. B. U. Haq and F. W. B van Eysinga, *Geological Time Table*, Elsevier, Amsterdam 1998; K. J. Hsü (ed.), *Mesozoic and Cenozoic Oceans*, 1986; C. Pomerol, *The Cenozoic Era: Tertiary and Quaternary*, 1982; M. E. Raymo and W. F. Ruddiman, Tectonic forcing of late Cenozoic climate, *Nature*, 359:117–122, 1992.

Centaurus

The Centaur, a southern constellation in the Milky Way (see **illustration**). It is not visible from mid-northern latitudes, but observers in the Southern Hemisphere see two bright stars, Alpha Centauri (Rigel Kentaurus) and Beta Centauri (Hadar), marking the feet of the Centaur, one of the mythological Greek creatures who are half human and half horse. Centaurus honors the wise centaur Chiron, who educated Achilles and Hercules and who was elevated to the heavens by Zeus when he died. In 800 B.C., in the era when this myth was probably told, the constellation was almost completely visible from latitudes up to 40°N. Because of the precession of the equinoxes, only the upper half of the constellation is ever above the horizon at that latitude today. Other myths identify Chiron with the constellation Sagittarius. *See* PRECESSION OF EQUINOXES; SAGITTARIUS.

Rigel Kent, as it is often called, is the third brightest star in the sky; only Sirius and Canopus are brighter.

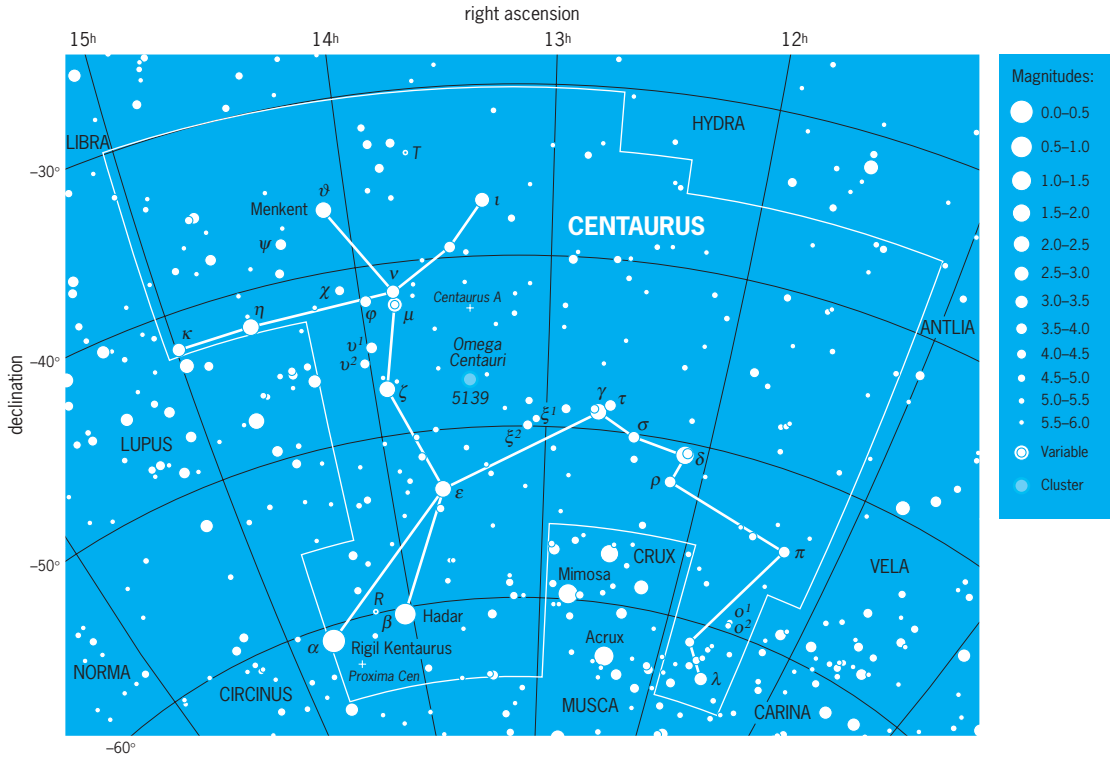

Modern boundaries of the constellation Centaurus, the Centaur. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

Alpha and Beta Centauri together point at the Southern Cross, the constellation Crux. They are thus often called the Southern Pointers, analogously to the Pointers at the end of the Big Dipper's bowl. *See* CRUX; URSA MAJOR.

The Alpha Centauri multiple star system is the closest star system to our own solar system. A faint red-dwarf companion star, known as Proxima Centauri, is marginally closer to us, making it the closest star to the Sun. *See* ALPHA CENTAURI.

A notable globular cluster, Omega Centauri got its name when it was originally mistaken for a star. The radio source Centaurus A turns out to be a giant active galaxy, emitting strongly also in the x-ray part of the spectrum. *See* RADIO ASTRONOMY; STAR CLUSTERS.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Center of gravity

A fixed point in a material body through which the resultant force of gravitational attraction acts. The resultant of all forces or attractions produced by the Earth's gravity on a body constitutes its weight. This weight is considered to be concentrated at the center of gravity in mechanical studies of a rigid body. The location of the center of gravity for a body remains fixed in relation to the body regardless of the orientation of the body. If supported at the center of gravity, a body would remain balanced in its initial position. Coordinates, as for an aircraft, are conveniently chosen with origin at the center of gravity. *See* GRAVITY; RESULTANT OF FORCES.

Nelson S. Fisk

Center of mass

That point of a material object or a system of objects that moves as if all the external forces were acting on the entire mass of the object or the system of objects concentrated at that point. The motion of a material object or a system of objects can be described in terms of the translational motion of the center of mass and the rotational motion about the center of mass. *See* ROTATIONAL MOTION; RIGID-BODY DYNAMICS.

Location. The location of the center of mass x_{CM} for a system of two particles of mass m_1 and m_2 located at positions x_1 and x_2 along the x axis is given by Eq. (1).

$$x_{CM} = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2} \tag{1}$$

If the two masses are equal, the center of mass lies midway between them in accord with Eq. (1). If their

masses are unequal, the center of mass lies closer to the larger mass. When this definition is extended to i particles in three dimensions, the center of mass is given by Eq. (2),

$$\mathbf{r}_{\mathrm{CM}} = \frac{\sum_i m_i \mathbf{r}_i}{M} \tag{2}$$

where M is the total mass of the system, $\mathbf{r}_i$ is the position vector of the ith particle, and m_i is the mass of that particle.

Relation to forces and momenta. The translational motion of an object or a system of particles behaves as if the sum of the external forces acting on the individual particles can be considered a single force $\mathbf{F}$ that acts on the total mass concentrated at $\mathbf{r}_{\mathrm{CM}}$. That is, the motion may be described by Eq. (3).

$$M\frac{d^2\mathbf{r}_{\mathrm{CM}}}{dt^2} = \mathbf{F} \tag{3}$$

Similarly, the product of the total mass times the velocity of the center of mass gives the total linear momentum of the system, $\mathbf{P}$, as in Eq. (4).

$$\mathbf{P} = M\frac{d\mathbf{r}_{\mathrm{CM}}}{dt} \tag{4}$$

Equations (3) and (4) lead to a statement of conservation of linear momentum: If the total external force is zero, the total linear momentum of a system is conserved. In this case, conserved means that a quantity is constant. In addition, the total linear momentum remains constant when the total external force is zero. *See* MOMENTUM.

Description of elastic collisions. The description of the elastic collision of two particles may be simplified by choosing a frame of reference that moves with the center of mass of the two-particle system. In that frame, the center of mass is at rest, the total momentum of the two particles is zero, and the particles are seen to be approaching the center of mass with momenta that are equal in magnitude and opposite in direction. After colliding, to conserve momentum, the particles recoil, moving in opposite directions along a line that makes an angle Θ with the initial line of motion (**Fig. 1**). The angle Θ is the scattering angle in

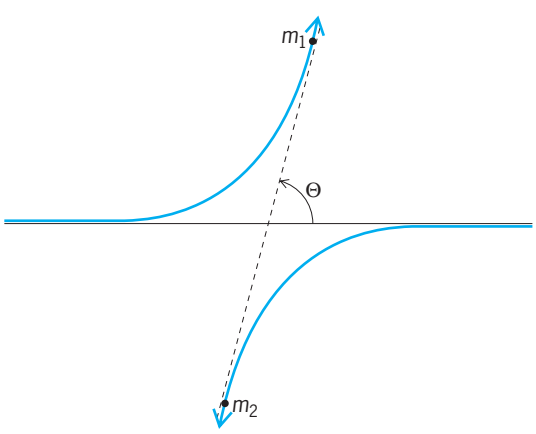

Fig. 1. Collision of two particles as seen from a frame of reference at rest with respect to the center of mass.

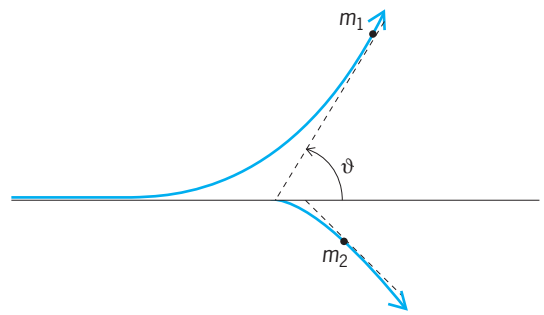

Fig. 2. Collision of two particles as seen from the frame of reference of the laboratory in which one of the particles is initially at rest.

the center-of-mass reference frame. The scattering is assumed to be axially symmetric so that no azimuthal angle need be introduced. The magnitudes of the particles' velocities must be unchanged in order that the kinetic energy is conserved, a requirement for an elastic collision. *See* COLLISION (PHYSICS).

Now consider the situation in which an incoming particle collides elastically with a particle at rest with respect to the laboratory. The collision may still be analyzed using the center-of-mass reference frame, and the results in the laboratory frame of reference may be obtained by a coordinate transformation.

For a collision between an incoming particle m_1 and a target particle m_2 initially at rest, the scattering angle ϑ measured in the laboratory is the angle between the final and incident directions of the scattered particle in laboratory coordinates (**Fig. 2**). The angle ϑ depends on the angle Θ in the center-of-mass frame and on the ratio of the two masses m_1/m_2 through the relationship in Eq. (5).

$$\tan\vartheta = \frac{\sin\Theta}{\cos\Theta + (m_1/m_2)} \tag{5}$$

For the special situation where $m_1 \ll m_2$, the two scattering angles become approximately the same.

Edwin R. Jones

Bibliography. H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3d ed., Addison-Wesley, Boston, 2002; E. R. Jones and R. L. Childers, *Contemporary College Physics 2001 Update*, 3d ed., McGraw-Hill, New York, 2001; J. B. Marion and S. T. Thornton, *Classical Dynamics of Particles and Systems*, 5th ed., Brooks Cole, Belmont, CA, 2004; C. P. Poole, Jr., *The Physics Handbook: Fundamentals and Key Equations*, Wiley, New York, 1998; K. R. Symon, *Mechanics*, 3d ed., Addison-Wesley, Boston, 1971.

Center of pressure

The center of pressure for an object is defined as the point at which the resultant surface pressure force may act and produce the same moment (torque) as the actual distributed pressure force on the object. The center of pressure is a very useful concept for

simplifying the effects of fluid forces acting on a solid object. *See* TORQUE.

In the case of a flat surfaces, the center of pressure will lie on the surface. Examples include the hydrostatic pressure force acting on a dam (**Fig. 1**) or the aerodynamic loading of an aircraft wing section (**Fig. 2**). In neither case is the atmospheric pressure included since it acts equally around all sides and contributes nothing to the net force or moment on the object. The placement of a balancing force equal and opposite to the resultant pressure force at the center of pressure will establish equilibrium of the body. *See* AERODYNAMIC FORCE; EQUILIBRIUM OF FORCES.

For curved plane surfaces the center of pressure may lie off the surface, but it still correctly indicates the point at which a balancing force may be placed. For example, a circular-arc water gate (**Fig. 3**) may be balanced by a single force located at the center of gate curvature.

Axisymmetric bodies moving through a fluid have a center of pressure that lies along the central axis but near or even forward of the nose (**Fig. 4**). However, flight without tumbling requires that the center of pressure lie behind the center of gravity. The feathers on the tail of an arrow increase stability by moving the center pressure rearward. Stability may be further improved by adding weight to the nose, which moves the center of gravity forward.

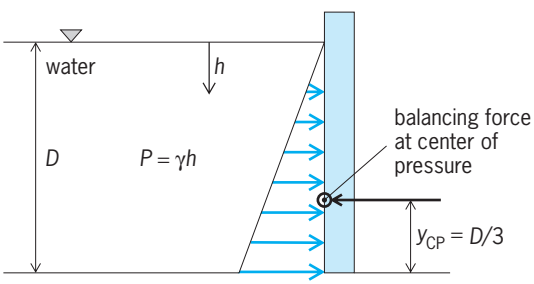

Fig. 1. Hydrostatic pressure force on a dam. The center of pressure is located at the centroid of the triangular pressure distribution indicated by arrows. h = depth beneath surface of water, D = total depth of water, P = pressure, γ = weight density of water ($\gamma = \rho g$, where ρ = mass density of water and g = acceleration of gravity); and y_{CP} = height of center of pressure above base of dam.

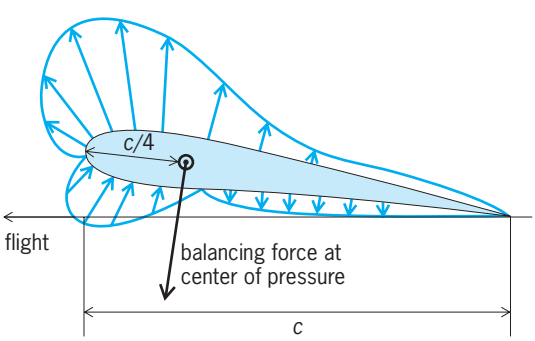

Fig. 2. Aerodynamic pressure force on a wing section. The pressure distribution is indicated by the arrows and their envelope. The center of pressure for the wing section is also referred to as the aerodynamic center of airfoil section. c = chord of airfoil.

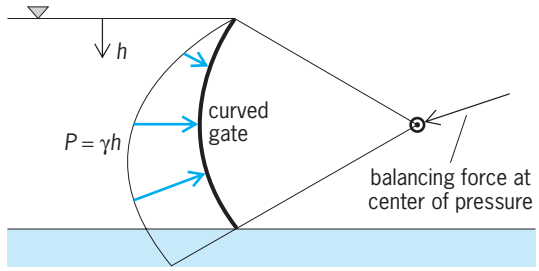

Fig. 3. Hydrostatic force on a circular-arc water gate. The pressure distribution is indicated by the arrows and their envelope. h = depth beneath surface of water, P = pressure, γ = weight density of water.

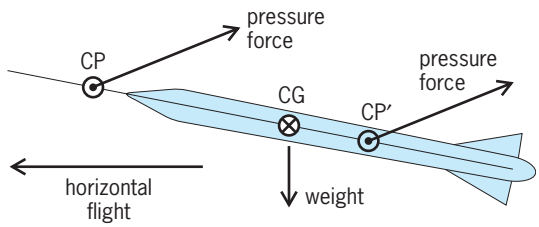

Fig. 4. Projectile in flight. The center of gravity is located at CG. The center of pressure without fins is located at CP and flight is unstable. The center of pressure with fins is located at CP′ and flight is stable.

In the case of objects submerged in a motionless liquid, pressure increases with depth. The center of pressure lies at the centroid of the object's volume and is referred to as the center of buoyancy. Stability is achieved when the center of gravity of the object is below the center of pressure. The center of pressure for an object floating on a liquid surface is located at the centroid of the displaced liquid. *See* BUOYANCY; CENTER OF GRAVITY; CENTROIDS (MATHEMATICS); RESULTANT OF FORCES.

A. Gordon L. Holloway

Central force

A force whose line of action is always directed toward a fixed point. The central force may attract or repel. The point toward or from which the force acts is called the center of force. If the central force attracts a material particle, the path of the particle is a curve concave toward the center of force; if the central force repels the particle, its orbit is convex to the center of force. Undisturbed orbital motion under the influence of a central force satisfies Kepler's law of areas. *See* CELESTIAL MECHANICS.

Raynor L. Duncombe

Central nervous system

That portion of the nervous system composed of the brain and spinal cord. The brain is enclosed in the skull, and the spinal cord within the spinal canal of the vertebral column. The brain and spinal cord are intimately covered by membranes called

meninges and bathed in an extracellular fluid called cerebrospinal fluid. Approximately 90% of the cells of the central nervous sytem are glial cells which support, both physically and metabolically, the other 10% of the cells, which are the nerve cells or neurons. Although the glial cells are much more numerous than neurons, the glial cells are also much smaller and do not directly participate in the propagation and integration of information. *See* MENINGES; NEURON.

Functionally similar groups of neurons are clustered together in so-called nuclei of the central nervous system. When groups of neurons are organized in layers (called laminae) on the outer surface of the brain, the group is called a cortex, such as the cerebral cortex and cerebellar cortex. The long processes (axons) of neurons course in the central nervous sytem in functional groups called tracts. Since many of the axons have a layer of shiny fat (myelin) surrounding them, they appear white and are called the white matter of the central nervous system. The nuclei and cortex of the central nervous system have little myelin in them, appear gray, and are called the gray matter of the central nervous system. The complex of gray and white matter forms the organizational pattern of the central nervous systems of all vertebrates. *See* BRAIN; NERVOUS SYSTEM (VERTEBRATE); SPINAL CORD. Douglas B. Webster

Centrifugal force

A fictitious or pseudo outward force on a particle rotating about an axis which by Newton's third law is equal and opposite to the centripetal force. Like all such action-reaction pairs of forces, they are equal and opposite but do not act on the same body and so do not cancel each other. Consider a mass M tied by a string of length R to a pin at the center of a smooth horizontal table and whirling around the pin with an angular velocity of ω radians per second. The mass rotates in a circular path because of the centripetal force, $F_c = M\omega^2R$, exerted on it by the string. The reaction force exerted by the rotating mass M, the so-called centrifugal force, is $M\omega^2R$ in a direction away from the center of rotation. *See* CENTRIPETAL FORCE.

From another point of view, consider an experimenter in a windowless, circular laboratory that is rotating smoothly about a centrally located vertical axis. No object remains at rest on a smooth surface; all such objects move outward toward the wall of the laboratory as though an outward, centrifugal force were acting. To the experimenter partaking in the rotation, in a rotating frame of reference, the centrifugal force is real. An outside observer would realize that the inward force which the experimenter in the rotating laboratory must exert to keep the object at rest does not keep it at rest, but furnishes the centripetal force required to keep the object moving in a circular path. The concept of an outward, centrifugal force explains the action of a centrifuge. *See* CENTRIFUGATION. C. E. Howe; R. J. Stephenson

Centrifugal pump

A machine for moving fluid by accelerating it radially outward. More fluid is moved by centrifugal pumps than by all other types combined. The smooth, essentially pulsationless flow from centrifugal pumps, their adaptability to large capacities, easy control, and low cost make them preferable for most purposes. Exceptions are those in which a relatively high pressure is required at a small capacity, or in which the viscosity of the fluid is too great for reasonable efficiency. *See* CENTRIFUGAL FORCE.

Centrifugal pumps consist basically of one or more rotating impellers in a stationary casing which guides the fluid to and from the impeller or from one impeller to the next in the case of multistage pumps. Impellers may be single suction or double suction (**Fig. 1**). Additional essential parts of all centrifugal pumps are (1) wearing surfaces or rings, which make a close-clearance running joint between the impeller and the casing to minimize the backflow of fluid from the discharge to the suction; (2) the shaft, which

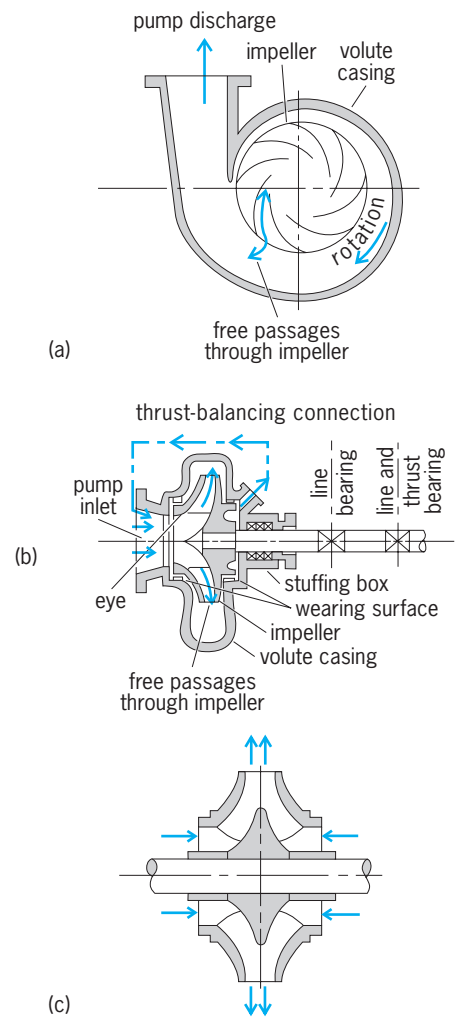

Fig. 1. Views of centrifugal pumps. (*a*) Section across the axis of a single-suction volute pump. (*b*) Section along the axis of a single-suction volute pump. (*c*) Double-suction impeller.

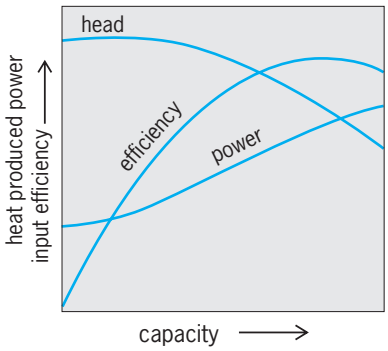

Fig. 2. Some of the typical characteristics of centrifugal pump at constant speed of rotation.

supports and drives the impeller; and (3) the stuffing box or seal, which prevents leakage between shaft and casing.

Characteristics. The rotating impeller imparts pressure and kinetic energy to the fluid pumped. A collection chamber in the casing converts much of the kinetic energy into head or pressure energy before the fluid leaves the pump. A free passage exists at all times through the impeller between the discharge and inlet side of the pump. Rotation of the impeller is required to prevent backflow or draining of fluid from the pump. Because of this, only special forms of centrifugal pumps are self-priming. Most types must be filled with liquid or primed, before they are started.

Every centrifugal pump has its characteristic curve, which is the relation between capacity or rate of flow and pressure or head against which it will pump (**Fig. 2**). At zero pressure difference, maximum capacity is obtained, but without useful work. As resistance to flow external to the pump increases, capacity decreases until, at a high pressure, flow ceases entirely. This is called shut-off head and again no useful work is done. Between these extremes, capacity and head vary in a fixed relationship at constant rpm. Input power generally increases from a minimum at shut-off to a maximum at a capacity considerably greater than that at which best efficiency is realized. The operating design point is set as close as practical to the point of best efficiency. Operation at higher or lower speed results in a change in the characteristic curves, with capacity varying directly with the speed, and head varying as the square of speed. Since the power required is proportional to the product of the head and capacity, it varies as the cube of the speed. These relations are essentially constant as long as viscosity of the fluid is low enough to be negligible.

Centrifugal impeller (and casing) forms vary with the relation of desired head and capacity at a practical rotating speed. Impellers and pumps are commonly classified as centrifugal, mixed-flow, and axial-flow or propeller (**Fig. 3**). There is a continuous change from the centrifugal impeller to the axial-flow impeller. For maximum practical head at small capacity, the impeller has a large diameter and a narrow waterway with vanes curved only in the plane of rotation. As the desired capacity is increased relative to the head, the diameter of the impeller is reduced, the width of the waterway is increased, and the vanes are given a compound curvature. For higher capacities and less head, the mixed-flow impeller is used. It discharges at an angle approximately midway between radial and axial. For maximum capacity and minimum head, the axial-flow or propeller-type impeller is used.

Capacity. Specific speed is used to identify the place occupied by an impeller in this range of form or type. The narrow, purely radial discharge impeller is at the low end of the specific speed scale while the axial-flow impeller is at the high end. The highest pump efficiency is obtainable at a specific speed of approximately 2000 when the capacity is expressed in gallons per minute (gpm), or 94 when the capacity is expressed in cubic feet per second (cfs). Specific speed $N_s = NQ^{1/2}H^{-3/4}$ where N is rpm, H is head in feet, and Q is capacity in gpm or cfs.

Since the highest efficiency is obtained at a medium specific speed, theoretically an optimum speed of rotation may be calculated for any head and capacity. Such a selection must be adjusted for practical considerations. Because most pumps are driven directly by electric motors, available motor speeds govern operating speeds to a large extent. For small capacities, a lower specific speed is usually necessary. The low limit is set by casting limitations for long vanes and extremely narrow impellers, by the hydraulic losses from flow through the small passages, and by the friction of a large disk rotating in the fluid. At the other end of the scale, for large capacities and low heads, high specific speed is necessary to keep the driver speed as high as practical. In large pumps the head per stage may be limited by mechanical considerations. At common motor speeds, single-stage pumps are limited

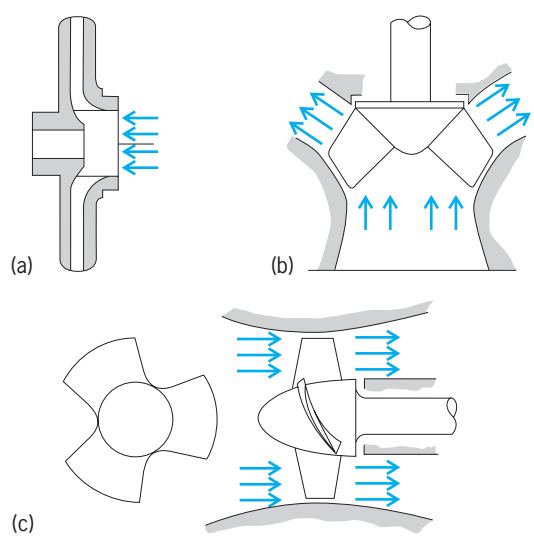

Fig. 3. Common classification of impellers. (*a*) Centrifugal for low speed. (*b*) Mixed-flow for intermediate speed. (*c*) Axial-flow or propeller for high speed.

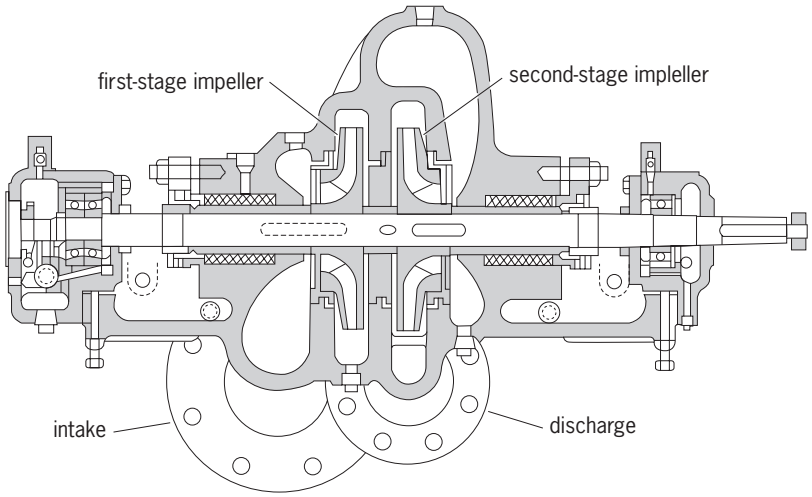

Fig. 4. Two-stage horizontally split casing pump.

to heads of about 600 ft (180 m) per stage. However, at higher speeds heads up to about 2000 ft (600 m) per stage have been attained in multistage boiler feed pumps.

Cavitation limits also play a major part in limiting the operating speed and head per stage for any specified condition. High speeds of rotation require high velocities at the entrance to the impeller. Since the static pressure in a closed stream of fluid drops as velocity is increased, there is a limiting velocity at which the absolute pressure approaches the vapor pressure of the fluid. Beyond this, partial vaporization occurs and the subsequent collapse of these pockets of vapor causes noise, vibration, and ultimately destruction of the surrounding walls. This form of vaporization in a moving stream of fluid is called cavitation. Higher operating speeds require higher suction pressure over the vapor pressure, spoken of as a net positive suction head. *See* CAVITATION.

Multistage pumps. When the required head exceeds that practical for a single-stage pump, several stages are employed. **Figure 4** shows a typical two-stage pump with single-suction impellers mounted back to back to obtain an approximate balance of the hydraulic thrust developed by pressure acting on the greater area of the back side of the impeller. A modern high-pressure multistage centrifugal pump enclosed in a forged steel barrel with all the single-suction impellers facing in the same direction is shown in **Fig. 5**. In this design axial thrust is balanced by a rotating disk mounted at the high-pressure end of the pump. Leak-off from this balancing device, which is piped back to suction pressure, serves also to reduce, practically to suction pressure, the load on the stuffing box packing or seal at this end of the pump. Pumps of this type are built with 4 to 10 or more stages, depending on the speed, pressure, and capacity desired.

Another type of multistage pump is the vertical turbine or deep-well pump. The details of design and the choice of an impeller of relatively high specific speed are the result of a drastic limitation of the outside diameter required to fit inside drilled-well casings. Pumps of this type are built with as many as 20 or 30 stages for high lifts from relatively small-diameter wells. Pumps with this same arrangement are built with a closed-bottom cylindrical housing with the inlet connected near the top of the housing, for applications at moderate head and limited net positive suction head. The cylindrical housing, or can, is then located in a pit or dry well with suction

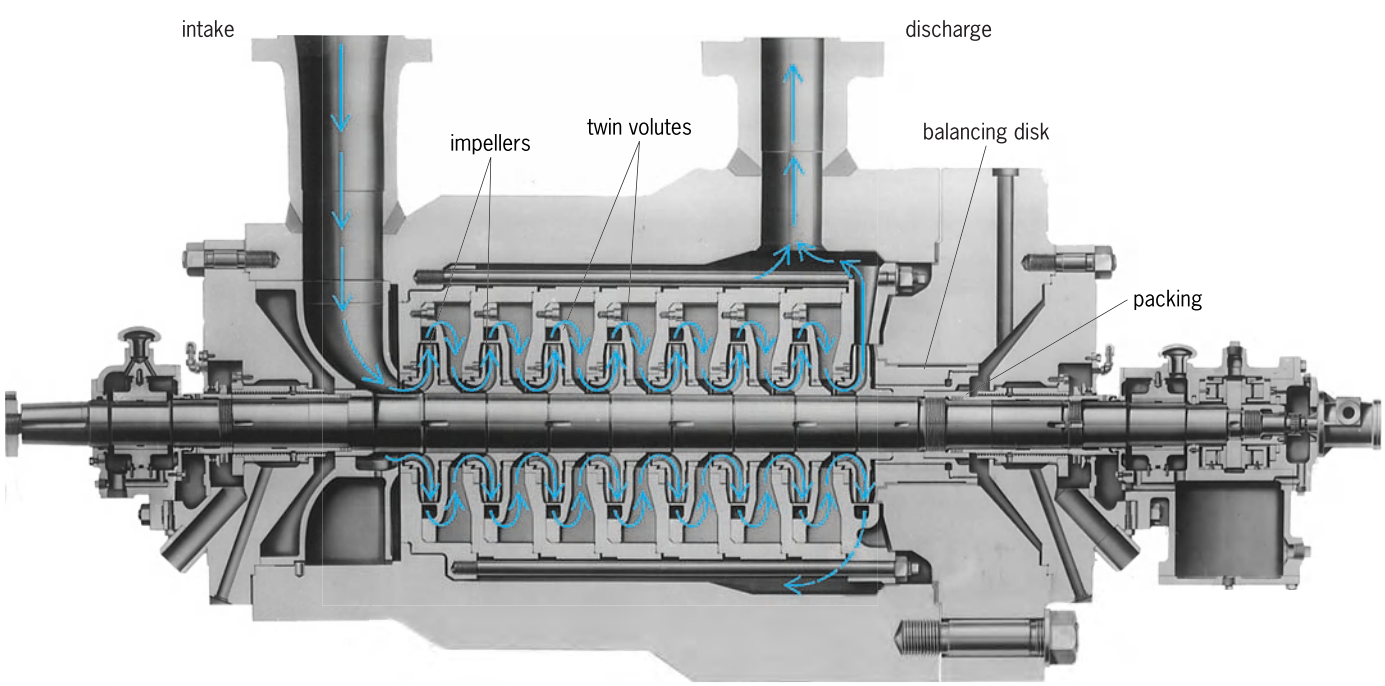

Fig. 5. High-pressure multistage centrifugal boiler-feed pump.

and discharge connections near ground or floor level. *See* PUMP; PUMPING MACHINERY. Elliott F. Wright

Bibliography. J. W. Dufour and W. Nelson (eds.), *Centrifugal Pump Sourcebook*, 1992; I. J. Karassik and W. C. Krutzch, *Pump Handbook*, 3d ed., 2000; V. S. Lobanoff and R. R. Ross, *Centrifugal Pumps*, 1992; A. J. Stepanoff, *Centrifugal and Axial Flow Pumps*: *Theory*, *Design and Application*, 2d ed., 1957, reprint 1993.

Centrifugation

A mechanical method of separating immiscible liquids or solids from liquids by the application of centrifugal force. This force can be very great, and separations which proceed slowly by gravity can be speeded up enormously in centrifugal equipment.

When an object is made to move in a circular path, its direction must constantly be changed. To accomplish this, a force must be applied acting inward toward the center of rotation; otherwise the object would continue on a straight path. An equal and opposite force, the centrifugal force, acts in an outward direction on the rotating elements.

The magnitude of the centrifugal force exerted on an element depends on its angular velocity and its distance from the center of rotation. Centrifugal acceleration is usually compared with the gravitational acceleration at the Earth's surface, and centrifugal force is expressed as a relative force F_R, expressed in multiples of the force of gravity, or G's. Relative centrifugal force is given by Eq. (1), where D is

$$F_R = 5.59 \times 10^{-4} n^2 D = 0.204 V^2/D \qquad (1)$$

the diameter of the circular path, in meters; n is the rotational speed, in rpm; and V is the peripheral speed of an element around the circle of rotation, in meters per second. In a centrifugal machine, the force varies radially from zero at the rotational axis to a maximum where D equals the inside diameter of the bowl or chamber. *See* CENTRIFUGAL FORCE.

Laboratory centrifuges are used for testing, control, and research. Industrial centrifuges find application in such diverse areas as sugar refining, chemical manufacture, dewatering diesel and jet fuel, and municipal waste treatment; they handle large amounts of material, yet require little space.

Stationary equipment: cyclones and nozzles. Centrifugal force is generated inside stationary equipment by introducing a high-velocity fluid stream tangentially into a cylindrical-conical chamber, forming a vortex of considerable intensity. Cyclone separators based on this principle remove liquid drops or solid particles from gases, down to 1 or 2 micrometers in diameter. Smaller units, called liquid cyclones, separate solid particles from liquids. The high velocity required at the inlet of a liquid cyclone is obtained with standard pumps. Units 4 to 12 in. (100 to 300 mm) in diameter make crude separations of large, heavy solid particles from liquids with a pressure drop across the cyclone of about 30–60 lb/in.^{2} (200–400 kilonewtons/m^2). More difficult separations may be done in banks of smaller units called hydroclones, each 0.4 in. (10 mm) in diameter, which can concentrate or remove 2–5-μm particles with a pressure drop of some 100 lb/in.^{2} (700 kN/m^2).

Comparison of bowl diameter and centrifugal force

Type of centrifuge	Bowl or basket diameter, in. (mm)	Rotational speed, rpm	Centrifugal force, multiples of gravity
Ultracentrifuge	0.2 (6)	230,000	177,000
Laboratory tubular	1.8 (45)	32,000	25,800
Industrial tubular	4 (100)	14,300	11,400
Disk	12 (300)	4,800	3,800
	18 (450)	3,200	2,600
Batch	30 (750)	1,900	1,500
	47 (1200)	1,200	970

In the separation nozzle process for uranium enrichment, a dilute mixture of gaseous uranium hexafluoride (UF_6) and hydrogen is caused to flow at high velocity in a curved path of very small radius of curvature (about 0.004 in. or 0.1 mm). The high centrifugal force tends to concentrate the heavier ^{238}U near the outer wall. The issuing stream is divided in two and sent through a long series of similar nozzles to achieve the desired degree of separation.

Rotating equipment: centrifuges. Much higher centrifugal forces than in stationary equipment are generated in a mechanically driven bowl or basket, usually of metal, turning inside a stationary casing. Rotating a cylinder at high speed induces a considerable tensile stress in the cylinder wall; for a thin-walled cylinder, this self-stress is given by Eq. (2),

$$S_s = 2.74 \times 10^{-3} n^2 D^2 \rho_m \qquad (2)$$

where S_s is the stress, in N/m^2, and ρ_m is the density of the material of the wall, in kg/m^3. This limits the centrifugal force which can be generated in a unit of a given size and material of construction. Dividing Eq. (1) by Eq. (2) and rearranging gives Eq. (3). Thus

$$F_R = 0.20 S_s / D \rho_m \qquad (3)$$

if S_s is fixed at the maximum stress allowable in the material, the maximum relative centrifugal force varies inversely with cylinder diameter. Very high forces, therefore, can be developed only in very small centrifuges. The **table** shows forces generated in various types and sizes of centrifuges at a typical peripheral speed of 246 ft/s (75 m/s).

Centrifuges

There are two major types of centrifuges: sedimenters and filters. Either type may be operated batchwise, with intermittent removal of the separated phase; semibatchwise, with essentially continuous removal of the liquid and intermittent removal of the solids; or continuously, with steady uninterrupted removal of both liquid and solids. Special

sedimenting gas centrifuges find application in uranium enrichment.

Sedimenting centrifuges. A sedimenting centrifuge contains a solid-wall cylinder or cone rotating about a horizontal or vertical axis. An annular layer of liquid, of fixed thickness, is held against the wall by centrifugal force; because this force is so large compared with that of gravity, the liquid surface is essentially parallel with the axis of rotation regardless of the orientation of the unit. Heavy phases "sink" outwardly from the center, and less dense phases "rise" inwardly. Heavy solid particles collect on the wall and must be periodically or continuously removed.

The sedimentation velocity of fine particles in a centrifuge is given by the form of Stokes' law shown in Eq. (4), where u is the setting velocity, n the rotational speed, D_p the particle diameter, ρ the particle density, r the radial distance from the axis of rotation, and μ the fluid viscosity. The setting velocity of small low-density particles, to which Stokes' law applies, increases directly with the relative centrifugal force. That of large heavy particles, to which Stokes' law does not apply, increases less rapidly with an increase in centrifugal force.

$$u = \frac{2\pi^2 n^2 D_p^2 \rho r}{9\mu} \tag{4}$$

Sedimenting centrifuges are divided into small, high-speed machines, which generate very high centrifugal forces, and large, comparatively slow-speed machines, which handle large amounts of material. Each group is further subdivided into liquid-liquid separators or decanters, clarifiers, sludge separators, and classifiers. Special machines partially separate gas mixtures for isotope enrichment.

Separators. Bottle centrifuges are batch laboratory centrifuges for research, analysis, or very small-scale production. The mixture to be separated is poured into bottles or test tubes which are placed in a rotor head and spun for a known length of time. An ultracentrifuge is a research unit which generates extremely high centrifugal forces inside a rotor around 0.2 in. (6 mm) in diameter (see table); sedimentation of extremely fine particles or macromolecules may be observed through a microscope and a transparent window in the rotor. *See* ULTRACENTRIFUGE.

Bowls in high-speed centrifuges for industrial use rotate about a vertical axis; liquid feed enters at or near the bottom of the bowl and escapes from the top. Stability considerations dictate one of two preferred bowl shapes: tall narrow cylinders or squat cylinders with a conical top. Tubular centrifuges typify the first design. Industrial models are 4 in. (100 mm) in diameter and 30 in. (750 mm) long and process 1.5 to 10 gal/min (0.3 to 2.5 m^3/h) of liquid. They serve as centrifugal decanters and as clarifiers to remove very small amounts of solids from liquids. A maximum of about 11 lb (5 kg) settled solids may accumulate in the bowl before the machine must be stopped and cleaned.

The bowl in a disk centrifuge (**Fig. 1**) is typically 8 to 16 in. (200 to 400 mm) in diameter and contains a large number of closely spaced "disks" which are actually truncated cones stacked one above the other and perforated to form passages for the liquid. The disks shorten the distance that a particle or drop must travel before being separated, compared with the distance in a tubular unit. Despite the somewhat lower centrifugal force, therefore, disk centrifuges are effective separators, especially for emulsion concentration. As shown in Fig. 1, feed enters the top of the machine and flows outward into the bowl near the bottom; separated liquids are removed through outlets at the top, while solids collect in the bowl. Large units can process 660 gal/min (150 m^3/h) or more and retain 110 to 130 lb (50 to 60 kg) of settled solids before cleaning becomes necessary.

Clarifiers. In a self-discharging separator, the settled solids are removed in more or less concentrated form through openings in the bowl wall. The bowl in a nozzle-discharge centrifuge is usually conical at the bottom as well as at the top. Where the bowl diameter is a maximum, 2 to 12 nozzles, lined with abrasive-resistant material, allow a slurry to issue continuously from the periphery of the bowl. Part of the slurry may be recycled to the bottom of the bowl to build up the solids concentration. Some nozzle machines are three-way separators, discharging two clarified liquids and a solids suspension. In a valve-discharge centrifuge, the peripheral openings are circumferential slits, covered most of the time by a piston. Periodically the piston is lowered by hydraulic or mechanical action, and concentrated solids are forcefully ejected, with some liquid, into the casing.

Sludge separators. In a sludge separator the continuous removal of solids as concentrated sludge requires that settled solids be moved mechanically out of the liquid and allowed to drain while subjected to centrifugal force. This is done in a scroll-conveyor centrifuge (**Fig. 2**), in which a cylindrical bowl with a conical end section rotates about a horizontal axis. Feed enters through a stationary central pipe and sprays outward into a pond or annular layer of liquid inside the bowl. Clarified liquid escapes through overflow ports in the cover plate on the large end

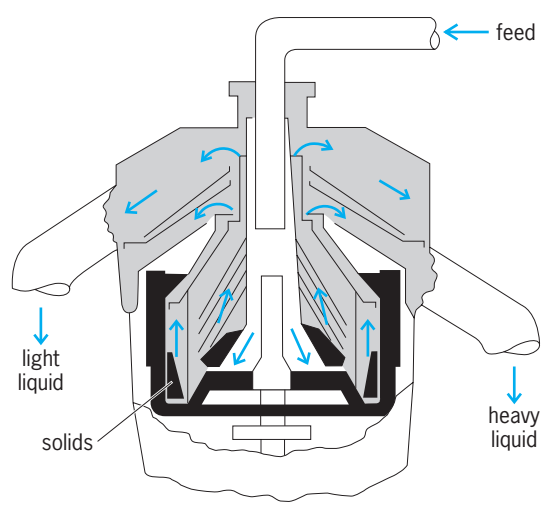

Fig. 1. Cutaway view of disk centrifuge bowl.

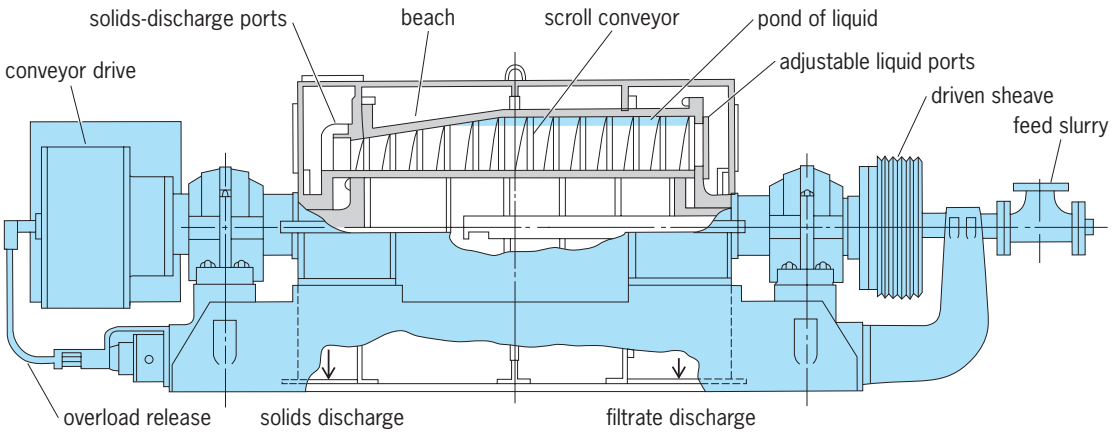

Fig. 2. Cylindrical-conical helical-conveyor centrifuge. (*Bird Machine Co.*)

of the bowl; the position of these ports fixes the depth of the pond. Solids settle under centrifugal force to the inner wall of the bowl, where a helical or scroll conveyor, turning slightly slower than the bowl, moves the solids out of the pond and up the slanting beach to discharge openings in the small end of the cone. Wash liquid may be sprayed on the solids as they travel up the beach. In some designs, feed enters near the large end of the bowl and liquid and solids travel in the same direction, with clarified liquid removed by skimmer pipes; this is effective in concentrating fine low-density solid particles encountered in sewage sludge treatment. Industrial sludge separators range in diameter from 16 to 55 in. (400 to 1400 mm), and separate from 1.1 to 2.2 tons/h (1 to 2 metric tons/h) of solids in a small machine to as much as 110 tons/h (100 metric tons/h) in a large one.

Classifiers. Classification is the sorting of solid particles in a fluid stream into two or more fractions according to particle size. Centrifugal classifiers are used when particles are too fine for separations by gravity. Scroll-conveyor centrifuges find application in large-scale wet grinding operations using ball mills, for example, separating acceptably fine particles from the oversize ones which are returned to the mill for regrinding. High-speed tubular centrifuges remove particles of grit from ultrafine cosmetics, dyes, and pigments.

Gas centrifuges. High-speed centrifuges are extensively used for uranium enrichment, partially separating ^{235}U from ^{238}U. They operate on gaseous uranium hexafluoride at moderate temperatures. The centrifugal force tends to compress the gas into a region near the bowl wall, while thermal motion tends to distribute the molecules uniformly throughout the available volume. Since thermal motion has a greater effect on light molecules than on heavy ones, a concentration gradient is established, with a higher concentration of the heavier molecules near the wall. The theoretical separation factor α is given by Eq. (5), where M_A and M_B are the molecular weights

$$\alpha = \exp \frac{(M_A - M_B)V^2}{2RTg_c} \qquad (5)$$

of the heavy and light components, respectively; V is the peripheral velocity; R is the gas constant; T is the absolute temperature; and g_c is the Newton's-law conversion factor. Since α depends on the difference between the molecular weights of the isotopes and not on their ratio, gas centrifuges have much higher separation factors than do gas diffusion units.

Gas centrifuges operate at very high speeds, for the separative capacity of a given unit depends on the fourth power of the peripheral velocity. In the typical machine shown in **Fig. 3**, a tubular rotor perhaps 8 in. (200 mm) in diameter rotates in an evacuated casing with a peripheral velocity of more

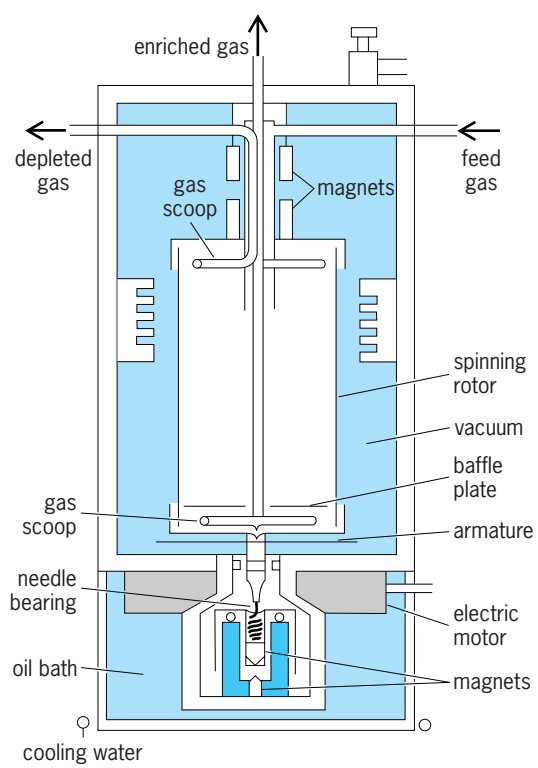

Fig. 3. Gas centrifuge. (*After V. V. Abajian and A. M. Fishman, Supplying enriched uranium, Phys. Today, 26(8):23–29, 1973*)

than 660 ft/s (200 m/s), about three times that in conventional industrial centrifuges. The centrifugal force at the bowl wall is more than 40,000 *G*. Feed gas enters the top near the axis of rotation; enriched gas is collected by a gas scoop near the bottom of the bowl, and depleted gas by a scoop near the bottom. Because of the high separation factor, relatively few centrifuges need be connected in series to achieve a substantial change in the concentration of ^{235}U. Productive capacities are low, however, and thousands of machines operating in parallel are needed for practical rates of throughput.

Filtering centrifugals. Centrifuges which separate by filtration are known as centrifugals or centrifugal filters. Solids and liquid are introduced into a rotating basket; after the liquid has been removed through the basket wall, the retained cake of solids is washed, spun dry, and removed. Centrifugals operate batchwise or continuously; various types differ primarily in the length of the cycle times and in the way in which solids are removed from the basket.

In large batch centrifugals, up to 47 in. (1200 mm) in basket diameter, the basket turns about a vertical axis. A solid cake 2–6 in. (50–150 mm) thick is formed, washed, spun (often at a higher speed), and then cut out of the basket, now turning very slowly, with a traveling knife. Discharged solids fall through openings in the basket floor; discharged liquid, or centrate, is removed from the casing. In some machines the cycle steps are fully automated. Free-draining crystals are separated quasicontinuously in short-cycle automatic horizontal-axis machines with baskets 20–40 in. (500–1000 mm) in diameter; solids are cut out every 45–150 s at full basket speed by a heavy-duty knife. Truly continuous filtering units for the large crystals are typified by the pusher centrifugal (**Fig. 4**), in which the solid cake, 1 to 3 in. (25 to 75 mm) thick, is moved over the inner surface of the perforated basket by a reciprocating piston. Each forward piston stroke moves the crystals an inch or so toward the basket lip; on the return stroke, a space is opened on the filtering surface on which more solids can be deposited. When crystals reach the basket lip, they fly outward into the casing. Filtrate and any wash liquid sprayed on the crystals during their travel leave the casing through separate outlets. Multistage pusher centrifugals are used when the solids cannot be pushed the fairly long distances required in the single-stage unit shown in Fig. 4.

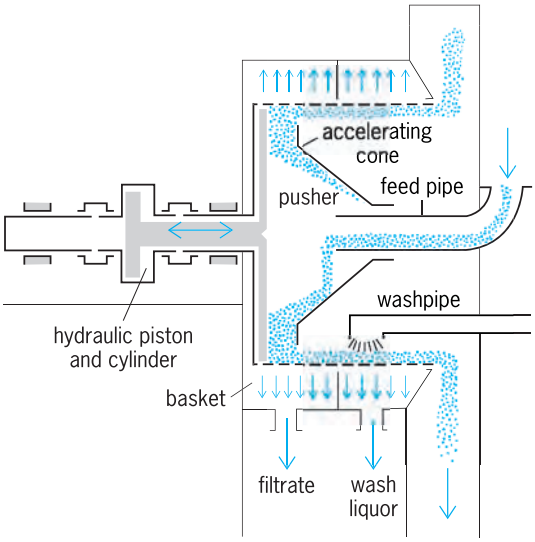

Fig. 4. Pusher-type centrifugal.

A filtering centrifuge operates on the same principle as the spinner in a household washing machine. The basket wall is perforated and lined with a filter medium such as a cloth or a fine screen; liquid passes through the wall, impelled by centrifugal force, leaving behind a cake of solids on the filter medium. The filtration rate increases with the centrifugal force and with the permeability of the solid cake. Some compressible solids do not filter well in a centrifugal because the particles deform under centrifugal force and the permeability of the cake is greatly reduced. The amount of liquid adhering to the solids after they have been spun also depends on the centrifugal force applied; in general, it is substantially less than in the cake from other types of filtration devices. *See* MECHANICAL SEPARATION TECHNIQUES. Julian C. Smith

Biological Applications

Biologists have long been interested in the effects of high centrifugal forces on cells, developing embryos, and protozoa. These studies have yielded valuable information on the properties of cells, such as surface tension, relative viscosity of the cytoplasm, and the spatial and functional interrelationship of cell organelles when redistributed in intact cells. The last finding led to the discovery that in a variety of invertebrate egg cells any disturbance of the intracellular organization of the cytoplasm will severely affect future development. *See* DEVELOPMENTAL BIOLOGY.

Another interest, partly in relation to the space program, has been the physiological effects of centrifugal forces on intact organisms, including humans. These forces are identical with the high gravitational forces astronauts must face on leaving and reentering the Earth's atmosphere. Such studies have been useful to determine the maximum forces that a person can tolerate and to develop equipment to offset the drastic effects. *See* AEROSPACE MEDICINE.

Thus the centrifuge has been applied to the whole gamut of biological systems, from humans to cells to subcellular organelles to biological molecules. It is routinely used in hospitals and clinics to collect blood cells, bacteria, and viruses, but its prime importance in biology is to isolate and determine the biological properties and functions of subcellular organelles and large molecules.

Cellular studies. During the first half of the twentieth century those interested in cell biology were generally segregated into two separate disciplines, cell biology and biochemistry. Cell biologists studied sections of tissues that had been subjected to various fixatives and then stained. From these studies evolved the identification of cell structures, but since these organelles had been rendered inactive by fixation, their function could only be guessed.

The biochemists viewed living systems as an enzymatic and metabolite "soup." They thoroughly ground up tissues and by primarily chemical procedures isolated enzymes and substrates. Again, however, no attempts were made to correlate the discovered metabolic-enzyme pathways with their location in intact cells. *See* CELL BIOLOGY; CELL METABOLISM; CYTOCHEMISTRY.

The modern technique for the isolation of cell parts was first introduced by R. Bensley and N. L. Hoerr in 1934, when they disrupted guinea pig liver cells and isolated mitochondria. A. Claude soon saw the importance of this technique, and from 1938 to 1946 extended the findings of Bensley and Hoerr by isolating two fractions, a "large granule" fraction, which was mitochondria, and a fraction of submicroscopic granules, which he called microsomes. From these beginnings additional pioneering investigators developed the centrifugal techniques of cell fractionation now commonly used.

Differential centrifugation. A tissue, such as the liver, or a population of cells, is broken up in one of several ways. The most common method is mechanical disruption, by grinding the tissue in a test tube with a glass or Teflon pestle (tissue grinder) or with a high-speed mixer such as a Waring blender. To be certain that the cellular particles retain their structure and function, much care is taken not to destroy them by overvigorous cell disruption, and to place the disrupted cell material, or homogenate, into a sucrose and salt solution that is isotonic and as close as possible to the natural conditions inside the intact cell.

The homogenate is then centrifuged in a medium at various centrifugal forces to separate the cellular organelles from each other. Organelles that differ in density, size, or shape sediment at different rates in a field of centrifugal force. Centrifugation simply magnifies the force of gravitational acceleration, which is normally constant, 392 in./s^2 (980 cm/s^2), and accounts for the weight of objects on Earth. Spinning a particle at high speeds produces a force on it proportional to the mass of the particle, the rotational speed or angular velocity in going in circles, and the distance from the axis or point of rotation to the particle. This is stated in Eq. (6), where G is the

$$G = \omega^2 r \tag{6}$$

gravitational acceleration, ω is the angular velocity, and r is the distance from the point of rotation to the particle. The angular velocity can be determined by knowing the rotor speed according to Eq. (7). It

$$\omega = \frac{2\pi(\text{rpm})}{60} \tag{7}$$

is thus possible to calculate the relative centrifugal force G, knowing the revolution per minute (rpm) and the radius r of the centrifuge, using Eq. (8).

$$G = \frac{4\pi^2(\text{rpm})^2 r}{3600} = 1.11 \times 10^{-5}(\text{rpm})^2 r \tag{8}$$

The rate at which a particle sediments in a centrifugal field is proportional to the centrifugal force, but also is dependent on the density and viscosity of the medium through which it is passing. This relationship is shown by Eq. (9), which states that t

$$t = \frac{9}{2}\frac{\eta}{\omega^2 r^2(d_p - d_m)} \cdot 2.3 \log \frac{R_{\max}}{R_{\min}} \tag{9}$$

is the time in which a fairly spherical particle of radius r and density d_p travels from the top of the centrifuge tube, $R_{\min}$, to the bottom of the tube, $R_{\max}$, in a medium having a known viscosity η and density d_m. This makes it possible to determine the time needed to centrifuge to the base of a tube a population of relatively homogenous particles from a larger population of particles of widely varying size and density, such as occurs in a cell homogenate.

This was the basis for the first techniques developed for cell fractionation, and is called differential centrifugation. These procedures have now been standardized. Tissue, such as liver, is homogenized at 32°F (0°C) in a sucrose solution (0.22–45 *M*) containing both a buffer to stabilize the pH and a salt, usually magnesium chloride. The homogenate is then centrifuged at 600 *G* for 10 min, also at 32°F (0°C). The sediment, which is forced to the bottom of the centrifuge tube, forms a solid pellet, and the overlying solution, or supernatant, is now removed and placed in another centrifuge tube. Called the nuclear fraction, the sedimented pellet consists mainly of nuclei which are the largest and densest organelles present in cells. The supernatant is centrifuged at a higher speed to attain a higher centrifugal force, 10,000 *G*, for 15 min. Another group of particles, the mitochondria, now sediment. This process can then be repeated, using higher centrifugal forces, to obtain the lysosome and microsome fractions from each succeeding supernatant medium. Finally, the supernatant, following centrifugation at forces of 100,000 *G* for one hour or longer, is essentially nonparticulate and is called the soluble fraction, containing enzymes and other small-molecular-weight substances that are not associated with any cell organelles. *See* CELL (BIOLOGY).

Density gradient method. Since there was a great deal of cross-contamination of particles among the various fractions, modifications soon developed to improve the separation of the cell organelles into purer fractions. The chief improvement involved the formation of a density gradient in the medium through which the particles migrate by altering the concentration of the sucrose or the salt. The gradient is so formed that the density is least at the top of the tube and is greatest at the bottom. A tissue homogenate is placed on top of a density gradient, and in the presence of centrifugal force each particle sediments to a specific zone of the gradient with an equivalent density to that of the particle. In this way zones, or bands, of the various cell organelles are formed with a minimum of contamination of other particles of differing size and density. A number of variations of the density gradient centrifugation method have been developed to further refine the quality of separation

of particles of similar size and shape but of slightly different density.

Another technique has enabled large volumes of material to undergo separation by a continuous flow system, in which homogenate is continually being added to a centrifugal field while the displaced material, which is essentially nonparticulate supernatant, is being discarded. Large quantities of particulate material may thus be accumulated.

Although biochemists had used centrifuges previously, with the advent of these methods cell biologists began to look at both the chemical nature and the morphology or structures of the isolated particles. In this way function could be correlated with structure and various metabolic processes could be localized in intact cells. It was established that the mitochondria were the centers for oxidative phosphorylation involved in the synthesis of adenosine triphosphate (ATP). The crude "microsome" fraction was found to consist of membranes and associated particles called ribosomes. Biochemical procedures had shown that nearly all protein synthesis occurred in the microsome fraction. By separating the membranes from the ribosomes by mild detergent treatment, the ribosome fraction retained protein synthetic activity. Several years later, ultracareful homogenization of developing red blood cells (reticulocytes) showed the ribosomes usually were aggregated and held together by a strand of ribonucleic acid (RNA). This has led to the polysome concept, in which the strand of RNA is messenger RNA (mRNA), which contains the genetic information that is obtained from the nucleus to make specific proteins. *See* RIBONUCLEIC ACID (RNA); RIBOSOMES.

A new particle was also discovered by centrifugation procedures, the lysosomes. These particles sediment very close to the mitochondria, but by careful manipulation they were separated. Lysosomes perform many of the digestive functions of the cell and contain many hydrolytic enzymes. *See* LYSOSOME; MITOCHONDRIA.

By these centrifugal procedures, then, it has been possible to isolate cell components, determine their composition and their metabolic activity, and examine the components to discover what they look like in order to find the most comparable structures in intact cells.

Molecular studies. Another valuable avenue of research and discovery opened by density gradient centrifugation techniques has been to find and isolate the large macromolecules that are involved in the various metabolic events of the cell, particularly for the nucleic acids. A complete description is not possible, but key examples may be given. There are slight but definite differences in the density of deoxyribonucleic acid (DNA) of various species isolated and centrifuged in cesium chloride gradients. This centrifugation depends of the base composition of DNA (ratio of adenine and thymine to cytosine and guanine). *See* NUCLEIC ACID.

Isotopic labeling combined with density gradient centrifugation has proved extremely useful to define the nature and role of the various RNAs in protein synthesis. A sedimentation "profile" is illustrated in **Fig. 5**, showing the presence of the two species of RNA found in ribosomes (rRNA-I and rRNA-II) and transfer RNA (tRNA). It is known that there is a third type of RNA, messenger mRNA, that is involved with protein synthesis. It is rapidly synthesized and is rather unstable. This synthesis characteristic can be used to advantage by rapidly exposing cells or a tissue to a precursor molecule of RNA labeled with an isotope such as ^{32}P or ^{14}C. After this brief exposure the cells are homogenized and the RNA extracted and centrifuged in a density gradient centrifuge. Fractions are removed drop by drop and analyzed for the presence of RNA by ultraviolet light, which is absorbed at 260 nanometer wavelength, and for

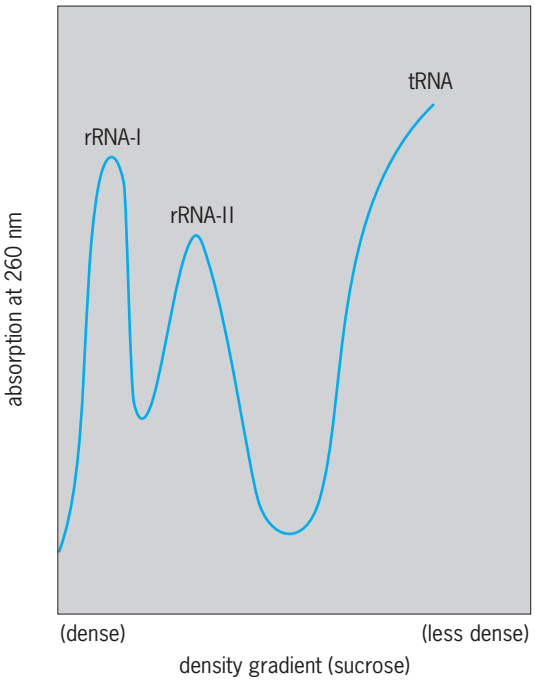

Fig. 5. Sedimentation profile showing ribosomal and transfer ribonucleic acid.

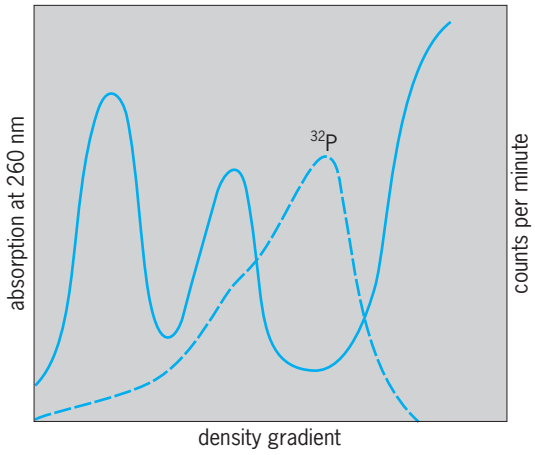

Fig. 6. Localization of radioactivity for RNA labeled with ^{32}P. Sedimentation profile of Fig. 5 is shown.

radioactivity. **Figure 6** illustrates that most of the radioactivity is localized between the tRNA and rRNA-II peaks, and composes a small portion of the total RNA of the cell. *See* PROTEIN.

A final distinction in centrifugation methodology may be made. If the contents of the centrifuge are to be analyzed after the centrifuge stops, as shown in Figs. 5 and 6, the procedure is called preparative; if the sample is examined optically during centrifugation, the procedure is called analytical. This last method is particularly important to determine the sedimentation coefficient and the molecular weight of the molecules under investigation. *See* MOLECULAR BIOLOGY. W. Auclair

Bibliography. D. Green (ed.), *Perry's Chemical Engineer's Handbook*, 7th ed., 1991; W. L. McCabe, J. C. Smith, and P. Harriott, *Unit Operations of Chemical Engineering*, 5th ed., 1993; D. Rickwood (ed.), *Centrifugation*, 2d ed., 1984; D. Rickwood (ed.), *Preparative Centrifugation: A Practical Approach*, 1993; R. W. Rousseau, *Handbook of Separation Process Technology*, 1987.

Centriole

A morphologically complex cellular organelle at the focus of centrosomes in animal cells and some lower plant cells. Prokaryotes, some lower animal cells, higher plant cells, and a few exceptional higher animal cells do not have centrioles in their centrosomes. Centrioles typically are not found singly; the centrosome of higher animal cells contains a pair of centrioles (together called the diplosome), arranged at right angles to each other and separated by a distance ranging from 250 nanometers to several micrometers. *See* CENTROSOME.

Structure and composition. Centrioles are typically 300–700 nm in length and 250 nm in diameter. Although they can be detected by the light microscope, an electron microscope is required to resolve their substructure. At the electron microscopic level, a centriole consists of a hollow cylinder of nine triplet microtubules in a pinwheel arrangement (see **illus.**). Within each triplet, one microtubule (the A tubule) is a complete microtubule, while the others (the B and C tubules) share a portion of their wall with the adjacent tubule. In some cells these nine triplet microtubules are embedded in a densely staining cylindrical matrix that is spatially distinct from the pericentriolar material of the centrosome. Structures found in the lumen or core of the centriole include linkers between the triplets, granules, fibers, a cartwheel structure at one end of the centriole, and sometimes a small vesicle.

Centrioles have a close structural similarity to basal bodies, which organize the axoneme of cilia and flagella. In many types of mammalian somatic cells, the older of the two centrioles in the centrosome can act as a basal body during the interphase portion of the cell cycle. In such cases, tapered projections, called basal feet, are often observed on the external surface

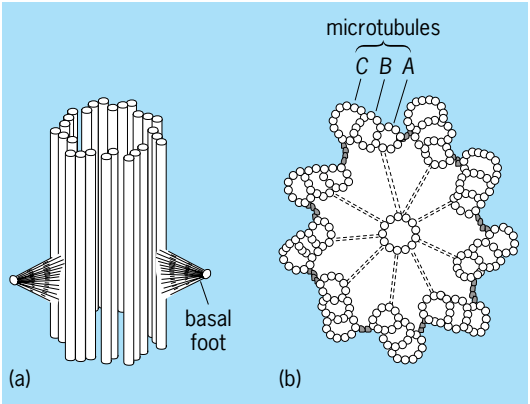

Diagrams of centriole showing (*a*) arrangement of microtubules and (*b*) cross section of proximal end, with nine triplet microtubules (A, B, and C) and central cartwheel structure.

of the centriole that is acting as the basal body. Microtubules are attached to the globular tips of the basal feet and may serve to anchor this centriole in the cell.

The centriole is a polar organelle, having a proximal end and a distal end relative to the cell center. When a centriole acts as a basal body, the ciliary or flagellar axoneme is nucleated from the distal end of the centriole. Daughter centrioles are formed perpendicular to the proximal end of the parent centriole. Within the centriole, the cartwheel structure is located at the proximal end.

The molecular composition of the centriole is poorly understood because it has not been possible to isolate enough centrioles in a sufficiently pure form to conduct biochemical analyses. The microtubule protein, tubulin, is a major constituent of centrioles; however, the identities of the proteins that make up the rest of the centriole are not known. Centrioles probably contain ribonucleic acid (RNA), but it is not known if this RNA is specific to the centriole or only represents cytoplasmic RNAs that happen to reside there. A centriole-specific role for this RNA, if any, has yet to be demonstrated.

Microtubules. During interphase the centrosome nucleates the array of cytoplasmic microtubules; later in the cell cycle the centrosome duplicates, and the daughter centrosomes form the poles of the mitotic (or meiotic) spindle. The terms "centriole" and "centrosome" are sometimes erroneously used interchangeably; centrioles are not the centrosome itself, but a part of it. The centrosome of higher animal cells has at its center a pair of centrioles, arranged at right angles to each other and separated by 250 nm or less. Spindle microtubules are nucleated by gamma tubulin containing ring complexes within the pericentriolar material, an amorphous cloud of densely staining material surrounding the centriole pair. Centrioles do not nucleate spindle microtubules. In some cells the pericentriolar material surrounds primarily the older of the two centrioles; the younger (daughter) centriole acquires pericentriolar material only later in the cell cycle when it separates from its mother and assembles its own daughter centriole.

Distribution and replication. During interphase the centriole pair and associated pericentriolar material lie closely apposed to the nuclear envelope or sometimes lie within a small indentation in the nuclear envelope.

Late in the G1 (first gap) phase of the cell cycle, the two centrioles separate slightly from each other and lose their orthogonal arrangement. When the cell initiates deoxyribonucleic acid (DNA) synthesis, short daughter centrioles (called procentrioles) are assembled at right angles to and separated slightly from the proximal ends of both mature centrioles. It is not known how daughter centrioles are formed in this way. The finding that cells from which centrioles are microsurgically removed cannot reassemble centrioles suggests that daughter centriole formation normally requires a specific organizing activity provided by the mother centriole. The procentrioles later elongate during the S (DNA synthesis) and G2 (second gap) phases of the cell cycle, reaching their mature length during mitosis or the following G1 phase. At a variable time in the G2 phase of the cell cycle, the centrosome as a whole splits, with a mother-daughter pair of centrioles distributed to each daughter centrosome. During mitosis, a pair of centrioles is found at the focus of both spindle poles. *See* CELL CYCLE; MITOSIS.

Functions. The only clearly demonstrated role for the centriole is to organize the axoneme of the primary cilium in cells having this structure, and the flagellar axoneme in sperm cells. Other possible functions for centrioles are a matter of debate. Some authorities assert that when present in the centrosome, centrioles contain activities that serve to organize the centrosome, determine the number of centrosomes in a cell, and control the doubling of the centrosome as a whole before mitosis. Evidence comes from observations that the introduction of extra centrioles into cell leads to the formation of extra centrosomes. Also, the ability of a centrosome to double between divisions is correlated with the number of centrioles that it contains.

Others believe that centrioles have no role in the formation and doubling of the centrosomes but are associated with the centrosomes only to ensure the equal distribution of basal bodies during cell division. This view is supported by the observation that all higher plant cells, most protozoa and algae and fungi, and a few animal cells form functional centrosomes without recognizable centrioles. At a minimum, the nine-triplet-microtubule aspect of the centriole is probably not required for centrosome formation or doubling. In principle, the nonmicrotubule components of the centriole might play important roles in centrosome doubling, but this has not yet been proved directly. *See* CELL (BIOLOGY); CELL DIVISION.

Greenfield Sluder

Bibliography. M. Carroll, *Organelles*, 1989; H. Hermann, *Cell Biology*, 1990; J. L. Hiatt et al., *Cell Biology-Histology*, 1993; K. Miller (ed.), *Advances in Cell Biology*, 1989; C. Widnell, *Essentials of Cell Biology*, 1990.

Centripetal force

The inward force required to keep a particle or an object moving in a circular path. It can be shown that a particle moving in a circular path has an acceleration toward the center of the circle along a radius. *See* ACCELERATION.

This radial acceleration, called the centripetal acceleration, is such that, if a particle has a linear or tangential velocity υ when moving in a circular path of radius R, the centripetal acceleration is υ^2/R. If the particle undergoing the centripetal acceleration has a mass M, then by Newton's second law of motion the centripetal force F_C is in the direction of the acceleration. This is expressed by Eq. (1), where ω

$$F_C = M\upsilon^2/R = MR\omega^2 \qquad (1)$$

is the constant angular velocity and is equal to υ/R. From the laws of motion given by Isaac Newton in his treatise *The Principia*, it follows that the natural motion of an object is one with constant speed in a straight line, and that a force is necessary if the object is to depart from this type of motion. Whenever an object moves in a curve, a centripetal force is necessary. In circular motion the tangential speed is constant but is changing direction at the constant rate of ω, so the centripetal force along the radius is the only force involved.

For example, a small heavy object tied by a cord of length R, with the other end of the cord held in the hand, can be whirled in a horizontal circular path. The person holding the cord has to exert an inward pull, the centripetal force, to maintain this circular motion (**Fig. 1**). When turning a corner in a car, the centripetal force necessary for the turn is provided by the frictional force between the road and the wheels. If the road is banked, then a component of the weight of the car can supply the necessary centripetal force. For a given angle of banking and radius of turn it is necessary to maintain a certain

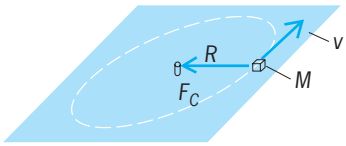

Fig. 1. Diagram of centripetal force in circular motion.

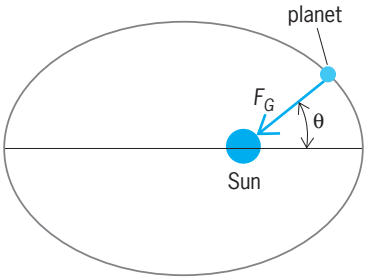

Fig. 2. Planet moving in an ellipse about the Sun.

speed, if the component of weight is to exactly provide the necessary centripetal force. *See* CENTRIFUGAL FORCE.

As another example (**Fig. 2**), consider the motion of the Earth or other planet about the Sun. The planets move in elliptical paths about the Sun. This motion is maintained by the gravitational attraction of the planet and Sun. In this case the radial distance between planet and Sun changes during the motion, so that the gravitational force is not in general equal to the centripetal force. If F_G is the gravitational force of attraction between the planet and Sun and R the distance between them, then Eq. (2) holds, where

$$F_G = M(a_R - R\omega^2) = M[d^2R/dt^2 - R(d\theta/dt)^2] \quad (2)$$

M is the mass of the planet, $a_R = d^2R/dt^2$ is the acceleration along the radial distance R, and $d\theta/dt$ is the instantaneous angular velocity. Since the path of a planet about the Sun is elliptical rather than circular, another force, beside the centripetal one, is required. *See* ORBITAL MOTION. R. J. Stephenson

Bibliography. F. Bueche, *Principles of Physics*, 6th ed., 1994; D. Halliday, R. Resnick, and K. Krane, *Physics*, 5th ed., 2002; F. Miller, Jr., D. Schroeer, and R. W. Stanley, *College Physics*, 6th ed., 1989; H. D. Young, et al., *Sears and Zemansky's University Physics*, 10th ed., 2000.

Centrode

The path traced by the instantaneous center of a plane figure when it undergoes plane motion. If a plane rigid body is constrained to move in its own plane but is otherwise free to undergo an arbitrary translational and rotational motion, it is found that at any instant there exists a point, called the instantaneous center, about which the body is rotating. The path that this instantaneous center traces out in space as the motion unfolds is called the space centrode. The path that it would trace out in a coordinate system which is rigidly attached to the body is called the body centrode. The motion may therefore be specified, when the two centrodes are given, by allowing one curve to roll without slipping along the other. *See* FOUR-BAR LINKAGE; RIGID-BODY DYNAMICS. Herbert C. Corben; Bernard Goodman

Centrohelida

An order of the Heliozoia. There is a central cell mass from which thin stiff arms radiate. The arms are supported internally by arrays of microtubules which terminate internally on a central granule, the centroplast, and the nucleus has an eccentric position. In the majority of species, the body is coated with a layer of siliceous spines or spicules and measures from 20 to 50 micrometers. These organisms are found in fresh-water and marine habitats, feeding on other protozoa, which adhere to the arms after colliding with them. In addition to about five genera

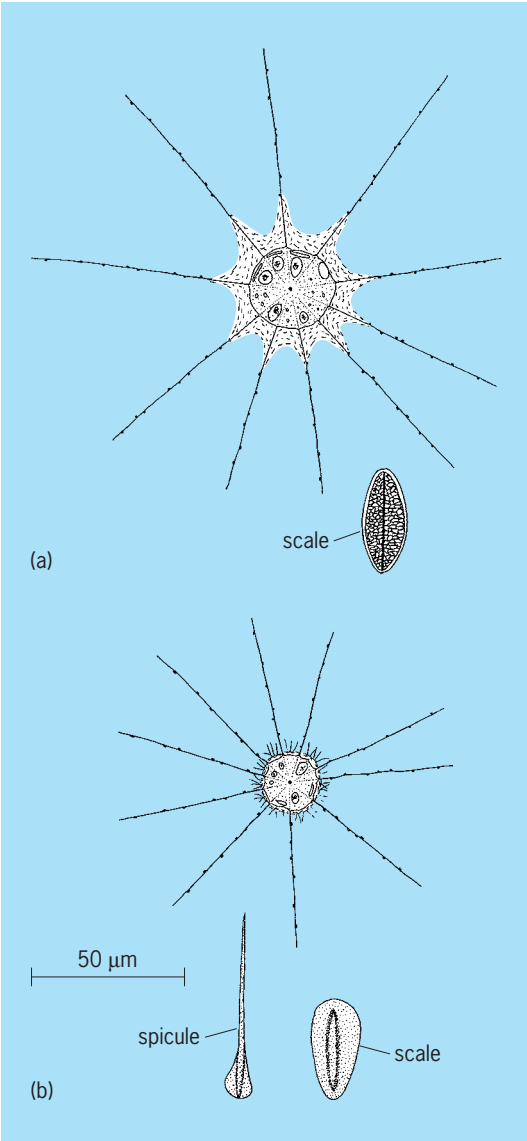

Centrohelida. (*a*) *Raphidiophrys ambigua*. (*b*) *Acanthocystis erinaceoides*.

which certainly are closely related (of which *Acanthocystis* and *Raphidiophrys* are among the most widely represented; see **illus.**), the group contains a variety of other Heliozoia of uncertain affinities. *See* ACTINOPODEA; HELIOZOIA; PROTOZOA; SARCODINA. David J. Patterson

Bibliography. J. J. Lee, S. H. Hutner, and E. C. Bovee, *Illustrated Guide to the Protozoa*, Society of Protozoologists, 1984.

Centroids (mathematics)

Points positioned identically with the centers of gravity of corresponding homogenous thin plates or thin wires. Centroids are involved in the analysis of certain problems of mechanics, for example, the phenomenon of bending.

Centroids by integration. The centroid of plane area A is point C (**Fig. 1**). Coordinates $\bar{x}$ and $\bar{y}$ of C as

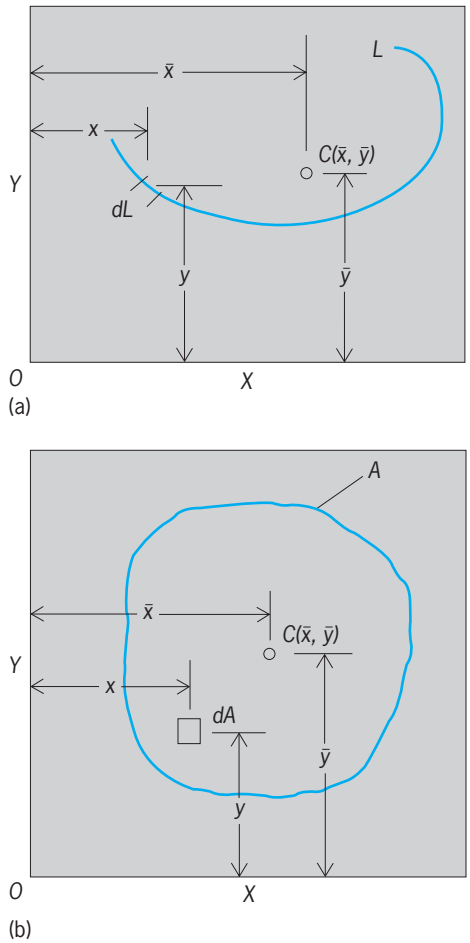

Fig. 1. Notation of integral equations for centroids. (*a*) Centroid of line. (*b*) Centroid of area.

referred to the indicated X and Y coordinate axes are given by Eqs. (1). Similarly, the centroidal coordinates of plane curve L are given by Eqs. (2). In

$$\bar{x} = \frac{\int x\,dA}{\int dA} \qquad \bar{y} = \frac{\int y\,dA}{\int dA} \tag{1}$$

$$\bar{x} = \frac{\int x\,dL}{\int dL} \qquad \bar{y} = \frac{\int y\,dL}{\int dL} \tag{2}$$

these equations x and y are the coordinate locations of infinitesimal area element dA and infinitesimal line element dL, respectively.

Centroids by summation. Whether or not the above integrals can be evaluated, approximate values of $\bar{x}$ and $\bar{y}$ may be determined by summation (**Fig. 2**). To this end, area A is divided as shown into n finite elements whose areas $\Delta A_1, \ldots, \Delta A_n$ contain center points $P_1, \ldots, P_n$. Coordinates $\bar{x}, \bar{y}$ of the centroid of area A are then approximately given by Eqs. (3).

$$\bar{x} = \frac{x_1\Delta A_1 + \cdots + x_n\Delta A_n}{\Delta A_1 + \cdots + \Delta A_n} \qquad \bar{y} = \frac{y_1\Delta A_1 + \cdots + y_n\Delta A_n}{\Delta A_1 + \cdots + \Delta A_n} \tag{3}$$

In analogous fashion, equations for the centroid coordinates of a planar line L are, in summation, given as Eqs. (4). Accuracy of the method increases as the

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i\Delta L_i}{\sum_{i=1}^{n} \Delta L_i} \qquad \bar{y} = \frac{\sum_{i=1}^{n} y_i\Delta L_i}{\sum_{i=1}^{n} \Delta L_i} \tag{4}$$

number of elements increases.

Principle of symmetry. The centroid of a geometrical figure (line, area, or volume) is at a point on a line or in a plane of symmetry of the figure.

A plane figure is symmetrical about a straight line if the figure may be decomposed into pairs of equal elements so that each pair is in reflected position about a point on the line. If every such pair is reflected in the same point, it is the point of symmetry of the figure.

Lines LL are lines of symmetry (and contain the centroids) of both the perimeter and area of the outlined shapes (**Fig. 3**). Point C is the point of symmetry (and the centroid) of the figure (**Fig. 4**).

Surface and volume of revolution. By the following theorems stated by Pappus of Alexandria (Greek mathematician, third century) a knowledge of the

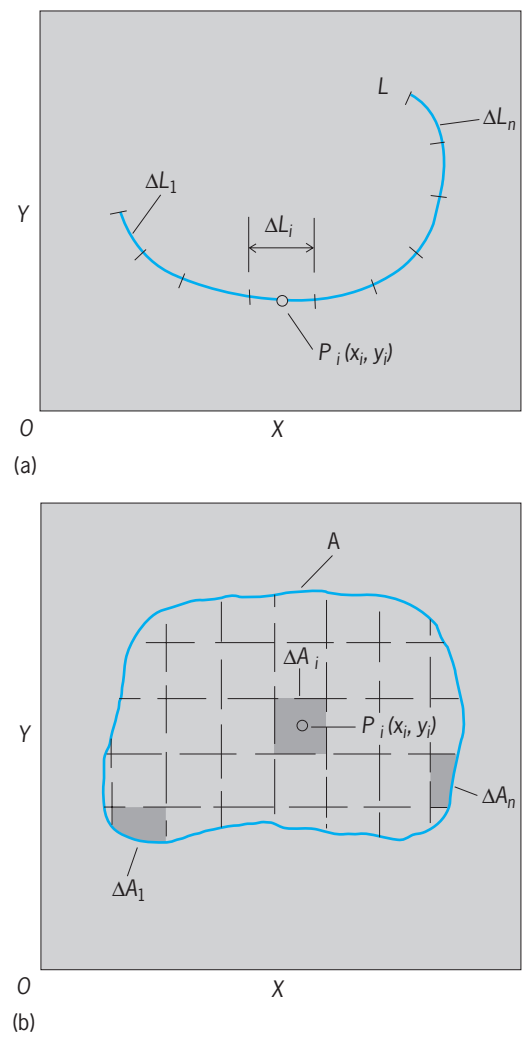

Fig. 2. Determination of centroids by summation. (*a*) Centroid of line. (*b*) Centroid of area.

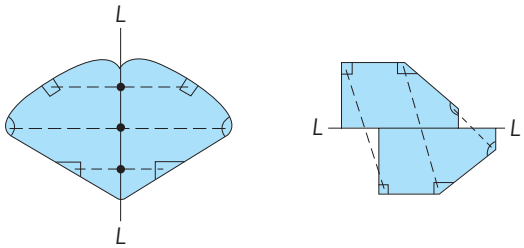

Fig. 3. Plane areas symmetric about a line.

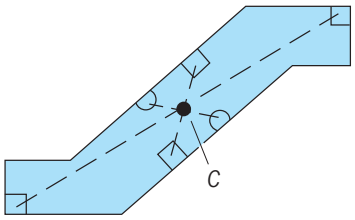

Fig. 4. Plane area symmetric about a point.

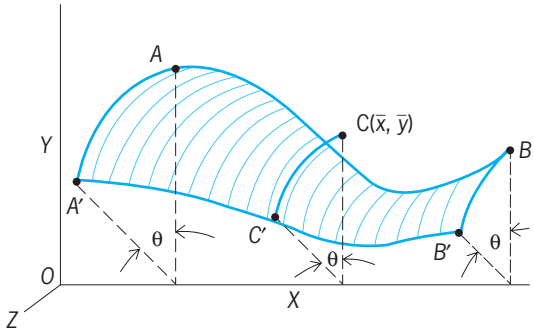

Fig. 5. Use of the centroid of a line to find area of a face of revolution, according to Pappus' theorem.

centroids of plane curves and plane areas may be employed to determine, respectively, surface areas of revolution and volumes of revolution.

Pappus' theorem of surface area is that the area of a surface generated by rotating a plane curve about a nonintersecting line in its plane is the product of the length of the curve and the arc length traveled by its centroid. Area S of the surface (indicated by shaded lines in **Fig. 5**) generated by rotating plane curve AB of length L an amount θ radians about the X axis from position AB in OXY to position $A'B'$ is $S = \bar{y}\theta L$, where $\bar{y}$ is the centroidal coordinate of curve AB referred to the OXY coordinate plane, and $\bar{y}\theta = CC'$ is the arc length traveled by its centroid.

Pappus' theorem of volume is that the volume of a solid generated by rotating a plane area about a nonintersecting line in its plane is the product of the area and the arc length traveled by its centroid. *See* CENTER OF GRAVITY; MOMENT OF INERTIA. Nelson S. Fisk

Centrosome

An organelle located in the cytoplasm of all animal cells and many plants, fungi, and protozoa, that controls the polymerization, position, and polar orientation of many of the cell's microtubules throughout the cell cycle. There is usually one centrosome per cell, located near the cell's center; it doubles during interphase, so there are two when the cell divides. At the onset of mitosis, each centrosome increases the number of microtubules it initiates. These mitotic microtubules are more labile and generally shorter than their interphase counterparts, and as they rapidly grow and shrink they probe the space around the centrosome that initiated them. When the nuclear envelope disperses, the microtubules extend from the centrosome into the former nucleoplasm where the chromosomes have already condensed. Some of these microtubules attach to the chromosomes, while others interact with microtubules produced by the sister centrosome, forming a mitotic spindle that organizes and segregates the chromosomes. During anaphase, sister centrosomes are forced apart as the spindle elongates, allowing each daughter cell to receive one centrosome to organize its microtubules in the next cell generation. *See* CELL CYCLE; CELL MOTILITY.

Morphology and organization. There is no simple morphological description of a centrosome because its shape differs widely between organisms. The centrosomes of animal cells usually contain a pair of perpendicular centrioles (**Fig. 1**), including a parent centriole formed in an earlier cell generation and a daughter centriole formed during the most recent interphase. Centrioles can serve as basal bodies for the initiation of a cilium or flagellum in the cells that make them. They appear to be essential for the formation of these motile appendages, so centriole inheritance by both daughters at cell division is analogous to the transmission of a gene. *See* CILIA AND FLAGELLA.

The organization of a centrosome is partly dependent on centrioles when they are present. The microtubule initiation that defines the centrosome

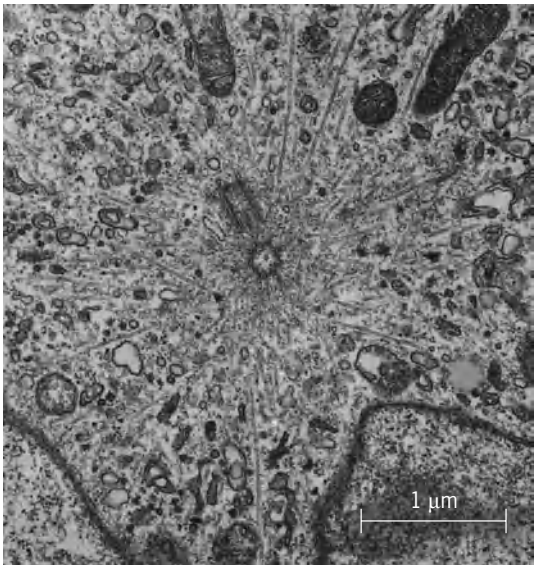

Fig. 1. Centrosome of a mammalian cell. Many microtubules radiate from the cloud of pericentriolar material which surrounds one of the two centrioles. (*Courtesy of Kent McDonald*)

is accomplished by a matrix that surrounds and is attached to the parent centriole; this matrix is often called pericentriolar material. In many cases, the pericentriolar material forms an almost spherical halo around the parent centriole, and the array of microtubules that forms is radially symmetric. In other cases, the pericentriolar material is highly structured and includes cone-shaped projections from the parent centriole (basal feet) or cross-striated fibers that extend from one end of the centriole (striated rootlets). Many of these appendages can initiate microtubules and should be regarded as a part of the centrosome. *See* CENTRIOLE; CYTOSKELETON.

In many lower eukaryotes there are no centrioles, but centrosomes still serve as both the principal organizer of interphase microtubules and as poles for the mitotic spindle (**Fig. 2**). In these centrosomes, the microtubule-initiating material is usually asymmetrically disposed, leading to the initiation of microtubule arrays without radial symmetry. Cells of higher plants lack visible centrosomes, yet they contain microtubules in both interphase and mitosis. During interphase, microtubule initiation is accomplished by cytoplasmic specializations that lie just beneath the plasma membrane at the corners and edges of the cells. During mitosis, microtubules form from two juxtanuclear regions that function similarly to centrosomes but lack a structurally defined microtubule organizing center. The cells of plants that form flagellated sperm also lack organized centrosomes, except in the sperm mother cells. The latter cells contain a more structured organelle associated with the spindle microtubules. The daughters of these cells produce even more structured centrosomes that eventually form many centrioles, each of which serves as a basal body for the formation of a sperm flagellum.

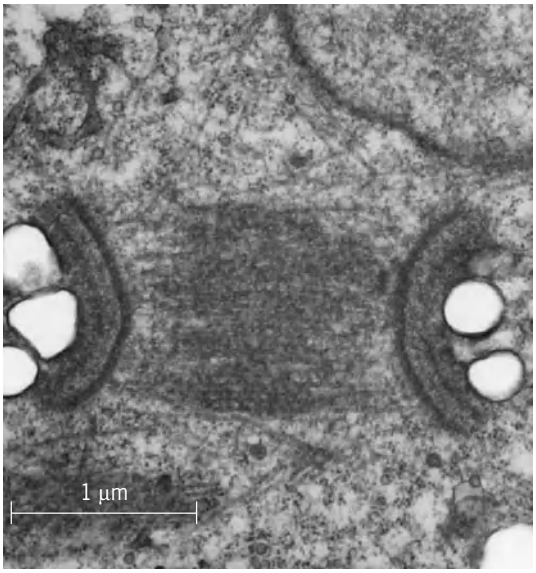

Fig. 2. Developing mitotic spindle in the alga *Stephanopyxis turris*. Microtubules are asymmetrically distributed around each centrosome (spindle pole). (*Courtesy of Kent McDonald*)

Chemistry. Centrosomes have been isolated from several lower eukaryotes, from embryos of *Drosophila*, and from mammalian cells. Biochemical descriptions of these fractions have now identified numerous centrosome components, but most of these components are unique to the cells from which they were isolated. Legitimate generalizations about the functions of centrosome components are still rare. One essentially universal centrosome component is γ-tubulin, a trace isoform of tubulin that is almost entirely confined to the centrosome. This protein may play a key role in initiating the polymerization of α- and β-tubulin. Mutations of or neutralizing antibodies to γ-tubulin and some other centrosome components prevent microtubule initiation or block cell cycle progression, implicating them in centrosome function. There are many candidates for proteins that might help to organize the centrosome and hold γ-tubulin in place, because several antigens have been localized to the pericentriolar material by using either antibodies raised against centrosome components or sera from humans with autoimmune disorders. *See* ANTIBODY; ANTIGEN.

Centrosome action is regulated as a function of time in the cell cycle. The increase in microtubule number that occurs prior to mitosis is correlated with a significant increase in the extent of phosphorylation of several centrosomal proteins. The protein kinase $p34^{cdc2}$, which helps to regulate the cell cycle, is concentrated at the centrosome, together with cyclin-B, a positive regulator of this kinase. Indirect evidence suggests that the changes in centrosome phosphorylation do not result simply from $p34^{cdc2}$, but from other kinases, some of which are regulated by $p34^{cdc2}$. Moreover, one of the regulatory subunits of cyclic adenosine monophosphate (cAMP)–dependent protein kinase is bound to a high-molecular-weight centrosome component, suggesting that additional kinases may also contribute to the regulation of centrosome activity. Phosphatases too have been localized to the centrosome.

Structural regulation. Biological reasons for centrosome regulation can be found not only in the importance of microtubules for interphase cell shape and for mitosis but also in the ways that centrosome behavior is modified in differentiative divisions during the development of metazoa. Developmental distinctions between daughter cells are often made at cell divisions that produce daughters of different size. Such divisions result from the asymmetric position of the cleavage furrow, which correlates with asymmetry in the mitotic asters, which in turn correlates with different activities at the two mitotic centrosomes. For example, when blastomeres cleave asymmetrically, the spindle's asters are of different size and shape. When oocytes go through meiosis and throw off polar bodies, one centrosome is usually attached to the cell cortex, while the other is not. *See* CELL (BIOLOGY); MITOSIS.

J. Richard McIntosh

Bibliography. D. M. Glover and J. W. Raff, The centrosome, *Sci. Amer.*, 268:62–68, 1993; V. I. Kalnins (ed.), *The Centrosome*, 1992.

Cephalaspidomorpha

The subclass of Agnatha that includes the jawless vertebrates with a single median nostril, as opposed to the subclass Pteraspidomorpha, in which there are paired nostrils. The Cephalaspidomorpha includes the orders Osteostraci, Galeaspida, Petromyzontida, and Anaspida. *See* ANASPIDA; JAWLESS VERTEBRATES; OSTEOSTRACI; PETROMYZONTIDA; PTERASPIDOMORPHA. Robert H. Denison

Cephalobaenida

One of two orders in the class Pentastomida of the phylum Arthropoda. This order includes primitive pentastomids with six-legged larvae. The hooks are simple, lacking a fulcrum, and disposed in trapeziform pattern, with the anterior pair internal to the posterior pair. The mouth is anterior to the hooks. In both sexes the genital pores are located anteriorly on the abdomen. There is no cirrus sac; the cirrus is short and united with the gubernaculum. In the female the uterus is straight or saccate.

There are two families, Raillietiellidae and Reighardiidae. The Raillietiellidae contains two genera: *Raillietiella*, with 19 species arranged in five groups; and *Cephalobaena*, with a single species. *Cephalobaena* has equal hooks at the ends of retractile, fingerlike appendages, the parapodia. The posterior end of the body is bifid. In *Raillietiella* the body is elongate with rounded ends, and there are lobes at the bases of the hooks. The anterior pair of hooks is the smaller. These animals are parasitic: adult stages occur in the lungs of lizards and snakes, and larval stages are encysted in reptiles, toads, and possibly insects. The life cycle is not well known and may be direct. The family Reighardiidae contains the single genus *Reighardia*, with one species, which infests the air passages of gulls and terns. The posterior end of this species is rounded, without lobes. The cuticula is covered with minute spines, and the hooks are small and equal. Proboscis and parapodia are lacking. *See* PENTASTOMIDA. Horace W. Stunkard

Cephalocarida

A group of tiny marine crustaceans (body length 2.0–4.0 mm or 0.08–0.16 in.), discovered in 1954. Initially, cephalocarids were thought to be a living representation of what the primordial crustacean looked like. In recent years, that view has changed, with the Branchiopoda currently postulated to represent the most primitive of extant major crustacean taxa. In the classification of J. W. Martin and G. E. Davis (2001), the Branchiopoda are basal, whereas the Cephalocarida are positioned between the Remipedia and Maxillopoda, suggesting a more advanced evolutionary position for the class. Nonetheless, in their discussion of "most primitive," Martin and Davis acknowledge that the question is still unresolved. At present, the Cephalocarida comprise ten species, assigned to five genera in one family. They have been found in flocculent surface deposits of mud or silty sand, from the intertidal zone down to depths of 5000 ft (1500 m), on the shores of all continents except Europe. Population densities up to an average of 16 individuals/ft^2 (177/m^2) have been recorded. *See* BRACHIOPODA; CRUSTACEA.

Species. The best-known species is *Hutchinsoniella macracantha*, from the east coast of North and South America, about 0.12–0.16 in. (3–4 mm) long, with a shovel-shaped head and a slender, very flexible body that is not covered by a carapace (**Fig. 1**). Species in the other four genera are similar, differing principally in the modifications of the

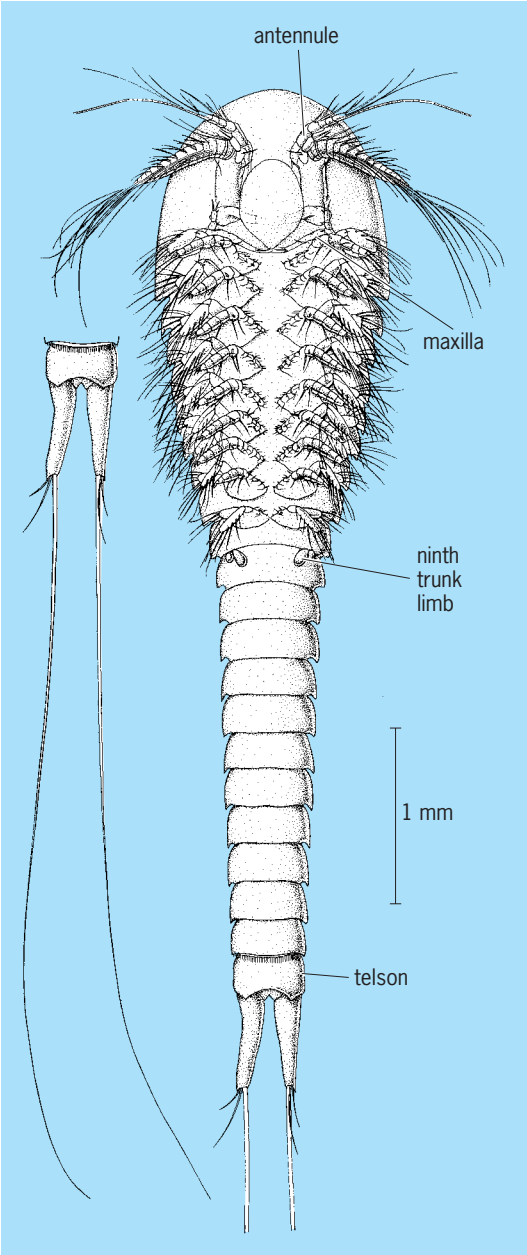

Fig. 1. Ventral view of *Hutchinsoniella macracantha*. (*After H. L. Sanders, The Cephalocarida: Functional morphology: Larval development, comparative external anatomy, Mem. Conn. Acad. Arts Sci., 15:1–80, 1963*)

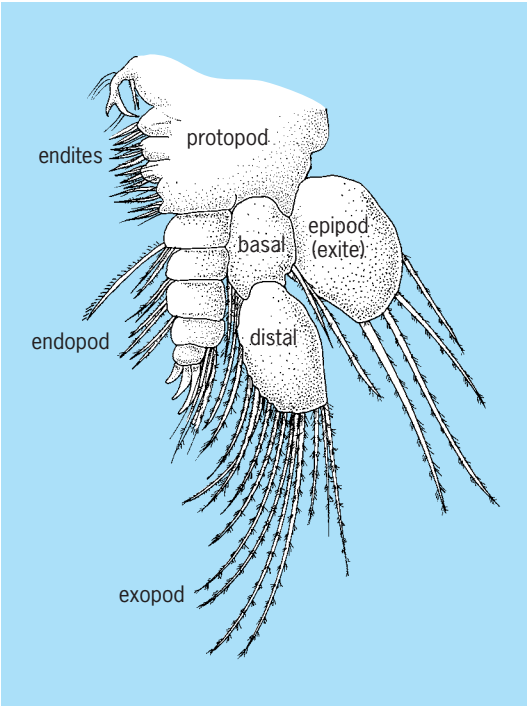

Fig. 2. ***Lightiella*** **fourth thoracopod. (*After P. A. McLaughlin, Comparative Morphology of Recent Crustacea, W. H. Freeman, 1980*)**

posterior four thoracopods, presumably as enhancements to reproductive functions. In *Lightiella*, the eighth thoracopods are entirely absent, whereas they are present but lack endopods in the other genera. In *Sandersiella* and *Hamptonellus*, the exites of the sixth thoracopods are highly modified, while the distal claw of each seventh thoracopod in *Hutchinsoniella* is bluntly rounded. These appendages show little modification in *Chiltonella.* Of the more anterior pairs of thoracopods, all are flexible but stiffened by internal turgor; each has a flattened basal protopod bearing several medial endites, a distal leglike endopod, and a distal paddle-shaped exopod; and an epipodite (or pseudepipodite) projects from the base of the exopod (**Fig. 2**). In the ventral nerve cord there is a partial coalescence of the ganglia in the head, but in the trunk a separate pair of ganglia is present in each somite. The heart extends through the limb-bearing trunk somites, in each of which it has one pair of ostia.

Morphology. Until the discoveries of the crustacean and crustacean-like fossils of the Cambrian Orsten and the Burgess and Chengjiang faunas, a number of attributes of cephalocarid morphology suggested a basic primitiveness. These included the serial homology of trunk appendages, the ventral nervous system, the uniramous, multisegmented antennules, the biramous antennae with both rami multisegmented, the large, flattened pleura on all the limb-bearing somites, and the telson freely articulated with the trunk and bearing a caudal furca. Even cephalocarid development, initially considered primitive in having a nauplius stage (**Fig. 3**) followed by a series of regular additions of body somites, has been shown by developmental stages in Orsten taxa actually to be apomorphic. A relatively slight metamorphosis occurs at the thirteenth molt, when the antennae, mandibles, and maxillules suddenly acquire their adult form.

Cephalocarids are nonswimming, bottom-creeping, nonselective deposit feeders. Distal claws on the endopods of the thoracopods scratch up particles that are passed into the median ventral food groove and moved forward to the mouth by metachronal beating of the thoracopods. The digestive system of cephalocarids has been examined in detail only in *Hutchinsoniella* and has been found in gross morphology to be consistent with other non-malacostracan crustaceans. It consists of a cuticular esophagus and rectum separated by an endodermal midgut with cephalic diverticula. However, *Hutchinsoniella* is unique in possessing many longitudinal muscles inside the circular muscles of the anterior midgut that change to only a few outside the circular muscles in the posterior midgut. Similarly, the gross morphology of the nervous system of *Hutchinsoniella* conforms to other crustaceans, but it also has elements of uniqueness. Most significant is the total absence of eyes and the organ of Bellonci, and with these the absence of brain centers associated with these sensory attributes. The most conspicuous elements of the cephalocarid "brain" are the olfactory lobes and the "mushroom body complex" (corpora pedunculata). The size and complexity of these olfactory lobes are unique for non-malacostracan crustaceans. Likewise,

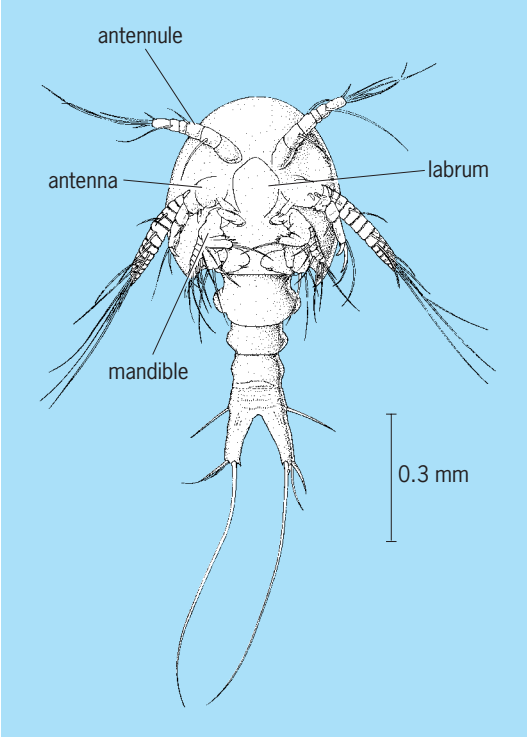

Fig. 3. Stage 1 nauplius larva of *Hutchinsoniella*. (*After H. L. Sanders, The Cephalocarida: Functional morphology, larval development, comparative external anatomy, Mem. Conn. Acad. Arts Sci., 15:1–80, 1963*)

non-malacostracans are thought to lack "mushroom bodies," and those found in *Hutchinsoniella* are distinct from those of malacostracans. *See* FEEDING MECHANISMS (INVERTEBRATE).

Reproduction. Cephalocarids are simultaneous hermaphrodites, with eggs and sperm that ripen at the same time. It is presumed that fertilization must include some form of copulation, but this has never been documented; thus self-fertilization cannot be ruled out. The sperm are not mobile and no indication of spermatophore development has been observed. The paired ovaries are located entirely in the maxillary somite, each adjacent to a tubule of the maxillary gland. The oviducts, beginning at the posteromedial margin of each ovary, extend posteriorly through the thorax and most of the trunk, then loop forward at trunk somite 18 and continue anteriorly to the sixth thoracic somite where they join the vasa deferentia. The sausage-shaped, paired testes occupy space dorsal to the gut from approximately the seventh thoracic somite to the third trunk somite. The testes have an anterior transverse connection from which the paired vasa deferentia descend to join the oviducts. The pairs of oviduct/vas deferens, each as a compact unit, then proceed to the gonopores located on the protopods of the sixth thoracopods. Each gonopore is blocked by a muscular operculum. These opercula, which are hinged laterally, can be pulled aside by the action of a set of muscles that originate in each protopod and insert centrally on the operculum. The extruded egg is attached to the knoblike tip of the very reduced ninth thoracopod by cement formed in a gland in the ninth somite. No more than two eggs are brooded at a time, and the length of time required for embryonic development makes it unlikely that individuals have more than a few broods each season. The actual extrusion and attachment of the egg has not been witnessed, but brooding embryos are always oriented the same way, with the anterior end of the cephalon directed posteriorly and its ventral surface pointed downward.

Phylogeny. The discovery of this presumably primitive class gave rise to speculation that cephalocarids stood close to the ancestral stem of the Crustacea. However, with the more recent discovery of another, and perhaps even more primitive, class, the Remipedia, this hypothesis was challenged. The exceptional morphological details described for some of the three-dimensionally preserved, phosphatized Cambrian Orsten fauna have demonstrated that a number of characters previously thought to be primitive in both the Cephalocarida and Remipedia were misinterpreted. Similarly, as more information has been gathered about the Cephalocarida, a substantial number of their attributes are now acknowledged as being unquestionably advanced. These include the lack of limbs on the last 11 trunk somites. The heart does not extend through these limbless somites. The mandibles of the adult have lost all parts distal to the coxa, and the adult maxillules are also somewhat reduced. The eighth trunk limbs are reduced or absent, and the ninth trunk limbs are greatly reduced. The reproductive system also is unique and unlike those of other Crustacea.

Additionally, although the cephalocarid digestive system and nervous system, as exemplified by *Hutchinsonella macracantha*, are comparable in their gross morphology to that of other crustaceans, they deviate in certain elements. The suggestion made by several carcinologists that cephalocarids are simply highly specialized progenetic paedomorphs may in fact be an accurate assessment.

John H. Lochhead; Patsy McLaughlin

Bibliography. R. Elofsson and R. R. Hessler, Central nervous system of *Hutchinsoniella macracantha* (Cephalocarida), *J. Crust. Biol.*, 10(3):423–439, 1990; R. Elofsson, R. R. Hessler, and A. Y. Hessler, Digestive system of the cephalocarid *Hutchinsoniella macracantha*, *J. Crust. Biol.*, 12(4):571–591, 1992; R. R. Hessler, R. Elofsson, and A. Y. Hessler, Reproductive system of *Hutchinsoniella macracantha* (Cephalocarida), *J. Crust. Biol.*, 15(3):493–552, 1995; R. R. Hessler and Y. Wakabara, *Hampsonellus brasiliensis* n. gen., n. sp., a cephalocarid from Brazil, *J. Crust. Biol.*, 20(3):550–558, 2000; J. W. Martin and G. E. Davis, An updated classification of the Recent Crustacea, *Nat. Hist. Mus. Los Angeles County Sci. Ser.*, 39:1–124, 2001; R. F. Schram and C. H. J. Hof, Fossils and the interrelationships of major crustacean groups, in G. D. Edgecombe (ed.), *Arthropod Fossils and Phylogeny*, pp. 233–302, Columbia University Press, New York, 1998; D. Walossek, The Upper Cambrian *Rehbachiella* and the phylogeny of the Branchiopoda and Crustacea, *Fossils Strata*, 32:1–202, 1993.

Cephalochordata

A subphylum of the phylum Chordata comprising the lancelets, including *Branchiostoma* (amphioxus). They are also known as the Leptocardii. Lancelets are small fishlike animals, not exceeding 3.2 in. (80 mm) in length. They burrow in sand on the ocean bed or in estuaries in tropical and temperate regions throughout the world. Only two genera are recognized, *Branchiostoma* with 23 species and *Asymmetron* with 6 species.

Morphology. The structure of the lancelet is based on the same fundamental plan as all other chordates, but there is neither head nor paired fins (see **illus.**). The skeletal rod of the back, or notochord, extends the entire length of the body. The animal is thus pointed at both ends, a feature to which the name amphioxus refers, and lanceolate in form, hence the name lancelet. The mouth lies at the base of an oral hood fringed with sensory projections or cirri and opens into a capacious pharynx occupying over half the length of the body. The body wall in the pharyngeal region is perforated on each side by a large number of elongated gill slits. This delicate branchial basket is protected by two folds of tissues, the metapleural folds, which grow down on either side of the body and, joining midventrally, form a sac into which

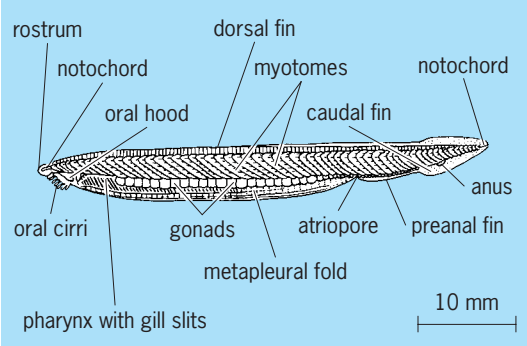

European lancelet (*Branchiostoma lanceolatum*). (*After J. E. Webb, Proc. Zool. Soc. London, 127:131, 1956*)

the gill slits open. The cavity of this sac is called the atrium and opens to the exterior through a posterior atriopore. The circulatory system resembles that of a fish, but there is no heart and the blood is colorless. The excretory system is quite unlike that of higher chordates and consists of nephridia situated above the pharynx with ciliated cells similar to those of the Platyhelminthes and the annelid worms. Lancelets are either male or female. The gonads are segmentally arranged pouches arising from the metapleural folds and discharging into the atrium.

Nutrition. Feeding in lancelets is by a ciliary or filter-feeding mechanism. Water is taken in through the mouth, passed through the gill slits into the atrium, and thence out through the atriopore. The water current is maintained by the beating of cilia on the gill bars. Microorganisms and organic debris on which the animal feeds are strained off by the gill bars onto a mucous string, the endostyle, and transported by cilia back to the intestine for digestion.

Life history. The adults burrow in rather coarse sand or shell gravel and usually lie with the mouth open at the surface of the sand. They select by trial and error sands of high permeability with rounded grains coated with microorganisms and free from organic decay. The selection of the optimum sand deposit available results in the congregation of large numbers of lancelets in small areas for spawning. The egg develops into a larvae not unlike the adult except that it is asymmetrical with a long elliptical mouth on the left side of the body and a single series of gills on the right. At first the larva is bottom-living, but later becomes planktonic. Metamorphosis to the symmetrical adult form takes place after about 11–12 weeks of life. The young adult sinks to the bottom and swims actively until a suitable sand in which it can burrow and remain undisturbed is found. Some species complete the life cycle in a year and death evidently occurs after spawning, but others may live 3 or 4 years and spawn annually. In some species metamorphosis occurs late and comparatively large larvae with 20–30 gills are formed. At one time specimens of this type were placed in a separate genus, *Amphioxides*, but these are now generally regarded as larvae of either *Branchiostoma* or *Asymmetron*.

Importance. The importance of lancelets lies chiefly in their being one of the most primitive chordates. However, lancelets are eaten by the Chinese and there are lancelet fisheries at Amoy. Lancelets are also being used as indicator organisms to assist in tracing the pattern of ocean currents on which their distribution appears to depend. *See* CHORDATA.

J. E. Webb

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; J. E. Webb and M. B. Hill, The ecology of Lagos Lagoon, pts. 3 and 4, *Phil. Trans. Roy. Soc. London B*, 241(683):334–391, 1958; C. K. Weichert, *Anatomy of the Chordates*, 4th ed., 1970; J. Z. Young, *The Life of Vertebrates*, 3d ed., 1981.

Cephalopoda

A class in the phylum Mollusca that is exclusively marine and contains the extant members—squids, cuttlefishes, octopuses, and the chambered nautiluses—and the fossil taxa—ammonoids, nautiloids, and belemnoids. The earliest known cephalopod fossils date to roughly 500 million years ago from the Upper Cambrian of northeast China. Fossilized external shells of the extinct subclass Ammonoidea and the nearly extinct subclass Nautiloidea are interpreted as indicating that these animals were shallow-living and slow-moving. Despite the thousands of species of such shelled cephalopods that have been recognized, all are extinct except for six species in two surviving genera, *Nautilus* and *Allonautilus*. Other living cephalopods, all predators in the subclass Coleoidea with internal shells, number over 725 species, with many more likely to be recognized. *See* MOLLUSCA.

General characteristics. Living cephalopods are bilaterally symmetric predatory mollusks that typically have conspicuous, laterally oriented eyes. Most have 8–10 appendages (8 arms and 2 tentacles) around the mouth; species of nautilus have over 90 arms. The appendages are lined with one to several rows of suckers or hooks, although the arms of nautilus lack conspicuous armature. Paired chitinous jaws, similar to an upside-down parrot's beak, with large muscles form the spherical buccal mass that contains the mouth and a tonguelike, toothed radula (a character that, in part, defines the Mollusca). The highly developed brain is positioned between the eyes where it is the center of the extensive, proliferated nervous system. The shell, which was external in most fossil cephalopods, as it is in modern species of nautilus, has become in all other living forms internal, highly modified, reduced, or absent. Paired fins that emerge from the side of the body in most coleoid cephalopods aid locomotion. The main thrust of primary movement is generated by jet propulsion in which water that is drawn into the sides of the mantle cavity is forcibly expelled through the nozzlelike funnel.

Classification. Among mollusks, cephalopods are thought to be most closely related to the class Gastropoda, although each group has had a very long and distinct evolutionary history. The earliest fossils are

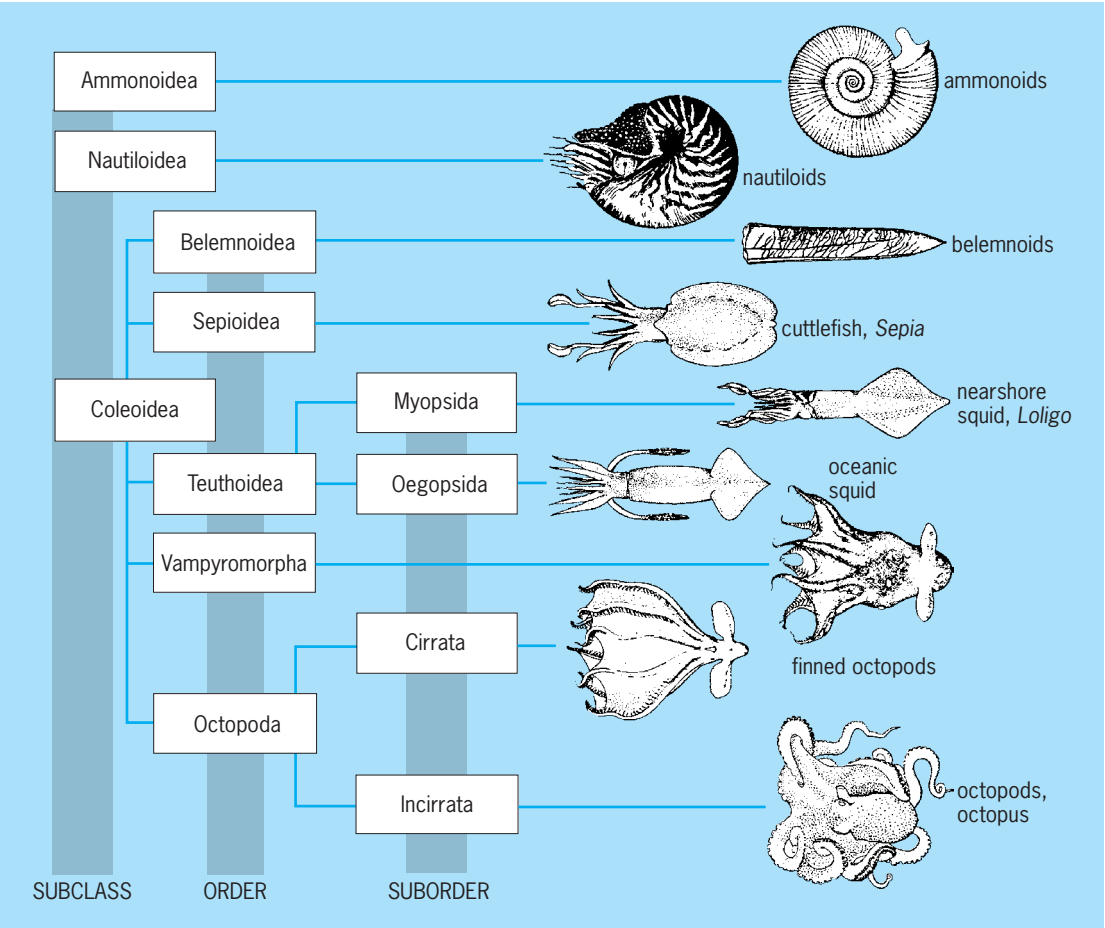

Classification and morphology (body form) of major groups of the class Cephalopoda.

external shells recognized as those of cephalopods by internal septa thought to relate to buoyancy control. The **illustration** shows the taxonomic hierarchy to the level of suborder in the extant coleoid cephalopods, with a representative figure of each taxon. *See* GASTROPODA.

Morphological characters that form the basis of fossil cephalopod classification are those associated with shells. Characters that form the basis of coleoid cephalopod classification, taxa that lack external shells, include the presence or absence of a cornealike membrane over the eye that separates nearshore or myopsid squids from oceanic or oegopsid squids. Although squids share general characters, the more than 20 families are diagnosed by a combination of the kind and arrangement of hooks or suckers on the arms and tentacles, shape of the cartilaginous connections between the mantle and funnel, and the presence or absence of tentacles at maturity. Cuttlefishes are recognized by their characteristic cuttlebone.

Among octopods, the presence or absence of fins separates the two suborders. Among finned or cirrate octopods, fin size and the shape of the supporting structure that underlies the fins provide taxonomic identification. Within-group distinctions among the finless or incirrate octopods include the arrangement of the suckers, relative arm lengths and web sector depths, skin color and texture, and copulatory organ features. Most modern cephalopod diversity lies within the three nearshore genera *Loligo*, *Sepia*, and *Octopus*.

Phylogenetic relationships among cephalopods are poorly resolved. The octopods are recognized as the sister group to the vampyromorphs, as both groups share characters that are otherwise unique in the cephalopods. Relationships within the squids and between them and the cuttlefish remain poorly understood, with several enigmatic groups complicating the issues.

Habitat. Cephalopods live in most marine habitats, from the intertidal zone to depths of over 16,500 ft (5000 m). Each extant cephalopod group tends to share ecological distribution. Chambered nautiluses are restricted to tropical and subtropical waters of the Indo-West Pacific, where they live associated with vertical reef faces from near-surface to 1600 ft (480 m) deep. Squids typically spend their lives swimming; nearshore squids typically occur on continental shelves and oceanic squids in open water. Cuttlefishes tend to occur in shallow water where they rest on the ocean floor, but they actively swim while hunting. Most octopod species are bottom-dwelling animals that occur in virtually all benthic habitats. In shallow water, octopuses tend to associate with shelters, which they either find or

excavate if in heavily sedimented areas. Lesser-known groups of octopods are pelagic at various ocean depths. Cirrate or finned octopus that live at 660 ft (200 m) and deeper were once thought to spend most of their time in the water column; seafloor observations show that they may also occur on the bottom at depths of 660 to 16,500 ft (200 to 5000 m). The only species in the order Vampyromorpha occurs at midwater depths of 1650–3960 ft (500–1200 m) worldwide. Most cephalopods are thought not to tolerate brackish water, although the myopsid squid *Lolliguncula brevis* is known to enter water with salinities as low as 8.5 parts per thousand, much more dilute than the normal ocean salinity of 35 parts per thousand.

Brain and nervous system. Cephalopods are characterized by their ability to change their behavior based on past experience, which requires memory and is consistent with learning. Their eyes are extraordinarily well developed as are their multilobed brains, which by some measures (ratio of brain to body weight) exceed those of most fishes and reptiles. Diffuse ganglionated nerve cords in the arms of octopuses and mantles of squids further indicate that their advanced nervous system approaches that of some birds and mammals. Since cephalopods are direct competitors with and frequent prey of marine fishes, some experts argue that fishes have generated selective pressure resulting in the rapid evolution of skin color and texture, predator detection abilities, and eyes of modern cephalopods. *See* NERVOUS SYSTEM (INVERTEBRATE).

Eye structure. Cephalopods—with the exception of species of nautilus that have pinhole eyes and one species of finned octopod that has reduced eyes—have large eyes that are very similar to those of vertebrates. Two features demonstrate that the structures evolved in convergence. First, to focus the eye on nearby subjects, cephalopods change the shape of the eyeball rather than the thickness of the lens, as do vertebrates. Second, in cephalopods the photosensitive cells point toward the lens, allowing the nerves that connect them to the brain to exit directly out of the back of the eye. In vertebrates, the photosensitive cells point backward, toward the retina, so they detect light that bounces off the retina; their nerves must extend into and their common optic nerve must exit out of the eyeball, creating a blind spot. Because cephalopod eyes are better engineered, they have uninterrupted vision. *See* EYE (INVERTEBRATE); EYE (VERTEBRATE).

Respiration, circulation, and excretion. Cephalopods extract oxygen from the water using gills that are fully contained within the muscular mantle. The gills lie on the side of the mantle cavity between the mantle wall and the viscera, where they are immediately exposed to water as it is drawn inside the mantle. The skin may act as an additional surface for gas exchange. The gills extend most of the length of the mantle in squids. In some cirrate octopods, however, the gills are reduced—rather than two attachments, a single one creates a semicircular gill. *See* GILL (ANATOMY).

Gas exchange, oxygen uptake and release of carbon dioxide from the bloodstream, is somewhat limited in cephalopods because their copper-based oxygen-carrying blood pigment, hemocyanin, is not highly effective in binding oxygen. The nearly closed circulatory system is pressurized by right and left branchial hearts that force deoxygenated blood through their respective gills and a medially situated systematic heart that pumps oxygenated blood through the rest of the body. The elastic nature of the arteries and veins is considered by some to contribute significantly to blood circulation. The same ability to contract also allows the blood vessels to stop blood loss in the event of an injury.

Nitrogenous waste is excreted in the form of ammonia, with several organs involved. The renal sacs, branchial heart appendages, digestive duct appendages, and even the gills contribute, with the branchial heart appendages being the site of ultrafiltration. The intestine and rectum prepare and expel indigestible material eaten with food from the body. The renal sacs are unusual in harboring minute wormlike members of the Dicyemida, which are only known from these organs of cephalopods. Ingested material that resists breakdown is passed through the intestine to be released from the viscera by the rectum, at the base of the nozzlelike funnel. A vigorous exhalation flushes it from the mantle.

Food, feeding, and digestion. Cephalopods are predators that feed on a variety of invertebrates, fishes, and even each other. The relatively sluggish nautiluses feed primarily on slow-moving prey such as reed shrimps, and even scavenge castoff shells of molted spiny lobsters. Cuttlefishes prey on shrimps, crabs, and small fishes, while squids eat fishes, pelagic crustaceans, and other cephalopods. Benthic octopuses prey mostly on clams, snails, and crabs. Prey capture is treated in Coleoidea.

Salivary glands secrete enzymes that begin digestion, and in octopuses contain toxins that subdue their prey. The digestive tract begins with the buccal mass that forms a mobile muscular sphere containing the chitinous beak that holds or bites off chunks of prey. The toothed radula working as a rasp moves these bits into the esophagus. Because the esophagus passes through the brain, prey morsels need to be fairly small and lack hard points. Digestion occurs primarily in the stomach where prey bits mix with digestive juices from the salivary glands, the digestive gland (the analog of the vertebrate liver and a lot more), and its appendages. The cecum is where digestion is completed; absorption of food also occurs in this organ and in the digestive gland and its appendages. Indigestible waste is moved through the system to the rectum, where it is released from the anus at the base of the funnel. It leaves the mantle cavity with water expelled during respiratory or in jetting contractions of the mantle (body).

Size, age, and growth. Cephalopods efficiently incorporate their food into their bodies, allowing very rapid growth rates. Many species of small cuttlefishes, bobtail squids, and octopuses reach reproductive maturity in less than a year, while larger

forms hatch in the summer, overwinter, and spawn at 12–14 months of age. The life spans of giant species of octopuses and squids are not known but are estimated to be 5 years or less. The smallest adult cephalopods are the Indian Ocean sepioid *Idiosepius pygamaea* at 0.2–0.8 in. (6–22 mm) body length, the Caribbean myopsid *Pickfordiateuthis pulchella* at 0.8 in. (22 mm), and the octopuses *Octopus micropyrsus* from California at 0.4–1.4 in. (10–35 mm) and *O. nanus* from the Red Sea at 0.5–1 in. (12–26 mm). The true giant squids of the family Architeuthidae are the largest invertebrates in the world, with a total length (tip of tail to tip of outstretched tentacles) of nearly 64 ft (20 m), a body length of 13–16 ft (4–5 m), and a weight of up to 1 ton. Most *Architeuthis* specimens are 32–38 ft (10–12 m) total length, 6–10 ft (2–3 m) body length, and about 450–660 lb (200–300 kg). The largest octopus, *Enteroctopus dofleini*, can have an arm span of 32 ft (10 m).

Defense. Because most cephalopods are unprotected by hard coverings, they are the prey of many other animals. Cephalopods have a myriad of defenses, most of which rely on escaping detection. To this end, they have become masters of camouflage and escape.

Chromatophore organs allow the skin of cephalopods to change color to match the background or mimic less desirable targets for predation. In addition, musculature in the skin can erect papillae, flaps, and knobs to match the background and break up the outline of the body. Many midwater oceanic squids have photophores on their ventral surface that match the intensity of light from above, thus reducing the conspicuousness of their silhouettes. Most cephalopods have an ink sac, from which blobs of dark ink are ejected to form a "decoy" for a pursuing predator. The two components of the ink, mucus and pigment, protect the cephalopod because when the predator bites the ink decoy, the ink temporarily blinds it and the mucus clouds its chemoreceptors, allowing the cephalopod to escape undetected. Jet propulsion, which despite its energy-intensive nature allows both rapid movement and extraordinarily high maneuverability, also contributes to cephalopod defenses. Benthic octopuses typically take refuge in shelters or dens that provide protection from predators. Some octopuses excavate holes in areas of soft sediment in which to hide. Squids occur in schools that may provide safety in numbers. Cuttlefish rest on the bottom and often cover their dorsal surface with a coat of sediment that effectively camouflages them. *See* CHROMATOPHORE; COLEOIDEA; PHOTOPHORE GLAND; PROTECTIVE COLORATION.

Locomotion. Cephalopods have perfected jet propulsion for many modes of locomotion, from hovering motionless, to normal cruising, to extremely rapid escape swimming. Water is drawn into the mantle cavity through openings on the sides and ventral mantle when the muscular mantle (body) expands. In squids and cuttlefish, the mantle opening can be closed by cartilaginous thickenings to ensure that contraction drives all water out of the mantle through the narrow funnel, propelling the cephalopod tail-first through the water. A complex of nerves controls this action, the most notable of which is the giant axon. Because nerve impulses travel faster in larger nerves, the giant axon allows the impulse that starts the contraction of the mantle to reach the dead-end part of the mantle first. As the wave of contraction moves forward toward the funnel, it moves the largest possible volume of water out to generate the most possible force, moving the squid as efficiently as possible. The giant axon of cephalopods, the largest single nerve fiber in the animal kingdom, was used as a model for neurobiology to increase our understanding of how nerves function. Although the fins of sepioids, teuthoids, and cirrate octopods are conspicuous, they primarily serve as rudders and stabilizers, providing only limited locomotory power.

Benthic octopuses normally move by walking or scuttling on their arms; when hunting they use each arm to explore the substrate. Most species, however, will use jet propulsion when threatened or if they are anxious to get somewhere. The use of buoyancy devices—such as the cuttlebone of cuttlefish (*Sepia*), gas chambers in other shells (species of nautilus, *Spirula*), low-density ions, especially of ammonium in special body reservoirs (Cranchiidae) or even in the body tissues (Histioteuthidae, Architeuthidae), and large amounts of oils in the liver (Bathyteuthidae)—reduces the energy required when these cephalopods are swimming.

Reproduction. Cephalopods have separate sexes and internal fertilization. Copulation in some species (such as nearshore squids) occurs at spawning, that is, when the eggs are released; in other taxa (octopods), sperm can be stored for periods of at least several months. Females produce eggs inside their (single) ovary. When the eggs mature, they are released through the paired oviducts to be inseminated by sperm from the male. Eggs may be encased in individual capsules (most octopods, sepioids, some oceanic squids), or may be enveloped in a gelatinous matrix and shaped into fingerlike clusters (nearshore squids) or formed into sausage- or ball-shaped masses that drift in the open ocean (some oegopsids). The fragile "shell" of the epipelagic argonauts (*Argonauta*) is actually secreted by the female's arms to hold her eggs during incubation. In one midwater pelagic octopod, the eggs develop entirely within the female's long, convoluted oviducts.

Males produce sperm in their single testis; it moves through a series of organs where it is packaged into cylinders (spermatophores) that the males store. The fluid inside the spermatophore is of comparatively high salinity. When the spermatophore is exposed to normal seawater, osmosis drives water inside it. The resultant increase in internal pressure forces the spermatophore to open and eject its sperm. A male transfers spermatophores to females either with a specially modified arm, the hectocotylus, or in those squids that lack a modified arm by a penis, which

is normally contained inside the mantle. Sperm in octopods is stored in glands in the oviduct, but in cephalopods without sperm storage organs, spermatophores may be injected into the female's mantle muscles, around the neck, under the eyes, or around the mouth, depending on the species. In these species, sperm is mixed with eggs as the female prepares the egg cluster.

Although most cephalopods leave their eggs to develop on their own, females of incirrate (finless) octopods care for their eggs during development, as does at least one species of oceanic squid. Egg development proceeds over a few weeks to a few months depending on the egg size and the temperature at which they develop. In deep-sea octopuses that produce very large (over 50 mm or 2 in. long) eggs, development may take years. Most hatchlings emerge from the eggs as planktonic "larvae" that, as they grow, slowly take on the adult form of their species; in some species hatchlings look like miniature adults.

Ecology. Cephalopods are important marine predators that kill and eat most marine animal taxa. They, in turn, are important prey to animals able to capture them. Because cephalopod beaks resist digestion, analysis of gut contents can identify cephalopods eaten by large predators. They are important components in the diets of toothed whales (sperm whales, dolphins), pinnipeds (seals, sea lions), pelagic birds (petrels, albatrosses), and predatory fishes (tunas, billfishes, groupers). In one case, North Atlantic pilot whales were found to prey almost exclusively on one species of squid, *Illex illecebrosus*, that aggregates for spawning in the summer. As cephalopods grow, they become less vulnerable to smaller predators, but they also become more attractive to larger predators.

Squids tend to form schools of individuals of the same size. The similar size of squids within a given school has been suggested to be due to the elimination of any smaller-than-average individuals by cannibalism. Octopuses tend to be solitary, with individuals coming together for copulation or cannibalism (the two are not exclusive, as females can cannibalize their suitors). Observational studies of a limited number of octopus species suggest that in part their high growth rates relate to the amount of time they spend resting, which is typically interrupted only by bouts of feeding. Janet Voight; Clyde F. E. Roper

Fossils. Modern representatives of the Cephalopoda are dominated in numbers by forms without a calcareous skeleton (such as *Octopus*) or by those taxa with only a reduced or rudimentary internal skeleton (most squids). Only a small minority of forms, such as the externally shelled *Nautilus* and *Argonauta*, the internally shelled cuttlefish (*Sepia*), and the squid *Spirula*, possess significant skeletal hard parts.

Fossil record. The fossil record, however, suggests that the typical morphologies of the world's cephalopodan fauna were quite different in the past. During the Paleozoic and Mesozoic eras (from about 560 to 65 million years ago), the majority of cephalopod fossils came from species with external shells, similar to that produced today by *Nautilus*. Although fossils of the subclass Coleoidea (to which squids and octopus are today assigned) are known, and in some deposits can even be extremely common, it appears that they represented only a small fraction of Paleozoic and Mesozoic diversity. Even accounting for their lesser chance of incorporation into the fossil record because of their reduced internal skeletons, it appears that the coleoids were never as abundant as the groups bearing external shells during the Paleozoic and Mesozoic eras. Over 15,000 Paleozoic- and Mesozoic-aged species belonging to the subclasses Nautiloidea and Ammonoidea are now known to science. In contrast, the modern-day cephalopod fauna, dominated by the squids and octopuses, has only several hundred species described to date, and appears to be very different from the typical cephalopod faunas characteristic of the Paleozoic and Mesozoic oceans. Most evidence suggests that the modern fauna is composed of taxa which have evolved only after the end-Mesozoic extinction of most externally shelled forms. *See* MESOZOIC; PALEOZOIC.

Chambered shell. Much of what is known and inferred about the habits of the ancient cephalopods comes from study of the small number of extant forms still possessing the external shells characteristic of the Paleozoic and Mesozoic cephalopod faunas. For this reason the modern-day *Nautilus* has been the subject of much research, for it represents the last living, externally shelled cephalopod. (Although *Argonaut*, a very specialized pelagic octopus, also has an external shell, this shell is produced by females only and serves as a specialized brooding organ for developing eggs.) The *Nautilus* shell has the dual purpose of lending protection to the vulnerable soft parts and serving as a buoyancy device. The rear of the shell contains numerous compartments partitioned by calcareous walls or septa; each chamber contains a central tube within which is a strand of living flesh communicating with the rear of the soft parts. This strand, called the siphuncle, is instrumental in achieving overall neutral buoyancy of the entire animal by removing a seawater-like liquid from each chamber. Although the modern *Nautilus* shows soft-part specializations which may be adaptations to its typical deep-water existence in front of coral reefs (and hence atypical of most extinct species, which appear to have lived in water depths and environments different from those typical of *Nautilus*), the nautilus shell shows close similarity with the shells of all fossil nautiloids and their closely allied descendants, the ammonoids.

In both nautiloids and ammonoids, the soft parts rest in a large space at the front of the shell. The shell enlarges through new calcification along its aperture. New chambers are produced at the back of the soft parts and are originally filled with the seawater-like liquid. This liquid is removed via the siphuncle (a tabular extension of the mantle passing through all chambers to the apex of a shelled cephalopod) by an osmotic process at a rate that balances the density increases brought about by the production of new

shell and tissue growth. After completion of a new chamber, the soft parts of the nautiloid or ammonoid move forward in the shell, creating a space behind the body for another chamber. All fossil cephalopods reached a final mature size and hence ceased shell growth.

With this cessation of shell growth, new chamber formation no longer became necessary. The buoyancy function of the chambered shell portion changed from one primarily dedicated to offsetting density increase due to shell and tissue growth, to one concerned with making small buoyancy compensations in response to slight, nongrowth weight changes, such as weight increases brought about by windfall feeding, or concerned with producing weight reduction, which could occur if shell material was broken off during predatory attacks on the nautiloid or ammonoid. There is reason to believe that buoyancy change was not used to power any sort of movement vertically through the water column. At least in the modern *Nautilus* (and by inference in the fossil nautiloids and ammonoids as well), the process of buoyancy change appears too slow to aid in any type of active locomotion brought about by rapid weight gain or loss.

Origins. The evolution of the chambered cephalopod shell, with its neutral buoyancy capability, dates back to the Cambrian Period; the shell was the central adaptation leading to the origin of the cephalopods from their snail-like molluscan ancestors. The earliest cephalopods are assigned to the subclass Nautiloidea. The first of these, known from Late Cambrian rocks in China, were small (less than 1 ft or 0.3 m long) and bore straight to slightly curved shells. By Ordovician time, however, adaptive radiations among the ancestral nautiloid stocks gave rise to a wide variety of shell shapes and sizes, including giants at least 10 ft (3 m) long. The subclass Ammonoidea arose from nautiloid ancestors in the Devonian Period, and can be differentiated from the nautiloids in details of shell morphology, especially relating to the complexity and position of the septa and siphuncle. It was also during the Devonian Period that the earliest internally shelled forms assignable to the Coleoidea appeared. *See* CAMBRIAN; DEVONIAN; ORDOVICIAN.

Evolutionary radiations and extinctions. The early Paleozoic nautiloids predated the first-known fishes and hence may have been among the earliest large predators in the sea. The subsequent evolutionary history of the nautiloids and their ammonoid descendants was one of successive waves of rapid evolutionary radiations. These externally shelled forms, however, were also very susceptible to extinction, especially during the rare intervals of mass extinction which occasionally occurred during the Paleozoic and Mesozoic eras. Following a period of extinction, a new fauna of nautiloids and ammonoids would rapidly evolve. This characteristic pattern of rapid speciation rates and high extinction rates tends to make most species chronologically short-ranging and hence useful in subdividing strata through biostratigraphy. Ammonoids are among the most useful known fossils for subdividing the stratigraphic record into small, chronological units.

The modern-day cephalopod fauna originated only after one of the most spectacular of the mass extinctions, occurring at the end of the Cretaceous Period. At this time, about 65 million years ago, a still-flourishing ammonoid fauna was totally destroyed; most nautiloids also died out as well. Although the record of Cenozoic cephalopod fossils is spotty, it appears that much of the cephalopod fauna as it is known today evolved only after the demise of the externally shelled forms. Thus, much as in the case of mammals and dinosaurs, the so-called modern fauna (the coleoid cephalopods) inherited the seas not because they were inherently superior to the ancient shelled nautiloids and ammonoids, but because they found themselves in an ocean emptied by mass extinction. *See* CENOZOIC; EXTINCTION (BIOLOGY).

Janet Voight; Peter D. Ward

Bibliography. W. Arkell et al., *Treatise on Invertebrate Paleontology*, pt. L, 1996; M. R. Clarke and E. R. Trueman (eds.), *The Mollusca*, vol. 11: *Form and Function*, 1988; M. R. Clarke and E. R. Trueman (eds.), *The Mollusca*, vol. 12: *Paleontology and Neontology of Cephalopods*, 1988; U. Lehmann, *The Ammonites: Their Life and Their World*, 1981; R. T. Hanlon and J. B. Messenger, *Cephalopod Behaviour*, Cambridge University Press, 1996; K. N. Nesis, *Cephalopods of the World*, T. F. H. Publications, Neptune City, NJ, 1987; M. Norman, *Cephalopods: A World Guide*, Conch Books, Hackenheim, Germany, 2000; P. Ward, *The Natural History of Nautilus*, Allen and Unwin. London, 1987.

Cephalosporins

A group of antibiotics that are effective in eradicating streptococcal, pneumococcal, staphylococcal, *Klebsiella*, *Neisseria*, enteric gram-negative rod bacteria that produce pulmonary, skin and soft tissue, bone and joint, endocardial, surgical, urinary, and bacteremic infections. They have been used most often in a preventive or prophylactic fashion at the time of various surgical procedures. The pharmacology of the early cephalosporins was such that the agents had to be administered by vein in frequent intervals in serious infections. This problem was overcome since most of the newer agents have much longer half-lives. All the third-generation cephalosporins penetrate well into tissues, and antibacterially active levels in various body fluids (pleural, peritoneal, synovial, biliary) and tissues such as bone are excellent. The toxic potential of the agents, considering their broad antibacterial spectrum, has been minor. Toxicities which are seen are those of bleeding due to vitamin K depletion.

There are many areas in which cephalosporins have been used, particularly respiratory, oral, abdominal, urinary tract, bone, gynecological, and neurological infections, and in septicemia. Roles in skin and soft tissue infections have been advocated for the drugs, but the agents still are a second choice

and the investigational agents offer most for the uncommon multiresistant bacterial infection. *See* ANTIBIOTIC. Harold Neu

Bibliography. H. C. Neu, *J. Antimicrob. Chemother.,* 6(Suppl.A):1–11, 1980.

Cepheids

A class of highly luminous yellow stars that vary periodically in brightness. Approximately one out of a thousand of the intrinsically brightest stars in any galaxy, like the Milky Way Galaxy, is a classical Cepheid variable. Among the 6000 apparently brightest stars visible to the unaided human eye in the night sky, more than a dozen are Cepheids (including Delta Cephei, the prototype; Beta Doradus in the south; and Polaris, the Pole Star). Cepheids have the same yellow color as the Sun (and hence, similar surface temperatures), but they have larger radii and therefore are intrinsically 1000 to over 100,000 times more luminous. Named for their naked-eye prototype Delta Cephei, the Cepheids are very well understood, and have repeatedly proved to be immensely important to many aspects of modern astronomy. Not only are Cepheids bright (supergiant) stars, so that they can be seen over large distances, but also they are at a particular stage in their evolution such that they are unstable to periodic oscillations. Motion of their surface manifests itself as regular changes in the apparent luminosity with time and so they can be easily identified. *See* POLARIS; SUPERGIANT STAR.

Distance indicators. The importance of Cepheids to astronomy comes both in their application to practical problems of distance determination (within the Milky Way Galaxy itself, and far beyond) and in their acting as critical tests of both stellar evolution and pulsation theory. From theoretical considerations, the period of oscillation and the intrinsic luminosity of a Cepheid can be understood in terms of the star's internal structure, mass, radius, and temperature. From observations it is known that the luminosity is closely predicted by the period of oscillation, a relation known as the period-luminosity relation, discovered by Henrietta Leavitt in 1912. This relation provides a powerful tool for estimating distances, since the period can be determined without prior knowledge of the distance. Using the period to predict how bright a given star would appear at various distances, it is possible to calculate the distance of a Cepheid from the observer, given its apparent luminosity. Knowledge of the intrinsic luminosity, surface temperature, and period of oscillation enable theorists to place Cepheids into their grid of evolutionary models from which ages and masses of these pulsating stars can be derived. Similarly, computer models can be constructed to match the oscillation behavior of real Cepheids, thereby predicting the mass, temperature, and internal structure needed to maintain such an instability.

In the 1920s, Edwin Hubble conclusively ended the debate as to whether other galaxies existed in addition to the Milky Way when he discovered Cepheids in nebulae now considered to make up the Local Group of galaxies. Among them were M31 and M33, situated at distances now known to be about 700 kiloparsecs (2.3×10^6 light-years). Hubble went on to show that the distances of galaxies correlated with their apparent recession velocities, consistent with an expanding universe. Cepheids have continued to play an extremely important role in measuring distances to galaxies. *See* GALAXY, EXTERNAL; HUBBLE CONSTANT; LOCAL GROUP.

Variability. Cepheids are generally identified by the distinctive and periodic optical variations in their light output. The time period over which a complete cycle is executed ranges from a few days to a few hundred days. Most Cepheids show an abrupt rise in their luminosity followed by a slower decline (**Fig. 1**), although some special varieties have a more symmetric (and generally lower-amplitude) light curve.

Light variability alone is ambiguous to interpret, but the analysis of the spectroscopic properties of a star, including the time variation of the radial velocities of Cepheids, ultimately led to the identification of the mechanism behind the changing light. The total light variation is the result of temperature changes in the stellar atmosphere, induced by, and combined with, an inward and outward motion of the surface of the star. Depending on the particular star, these excursions range from a few percent up to 10% of the mean radius of the Cepheid. The resulting velocities associated with the pulsation are observed to be in the range 10–70 km/s (6–43 mi/s).

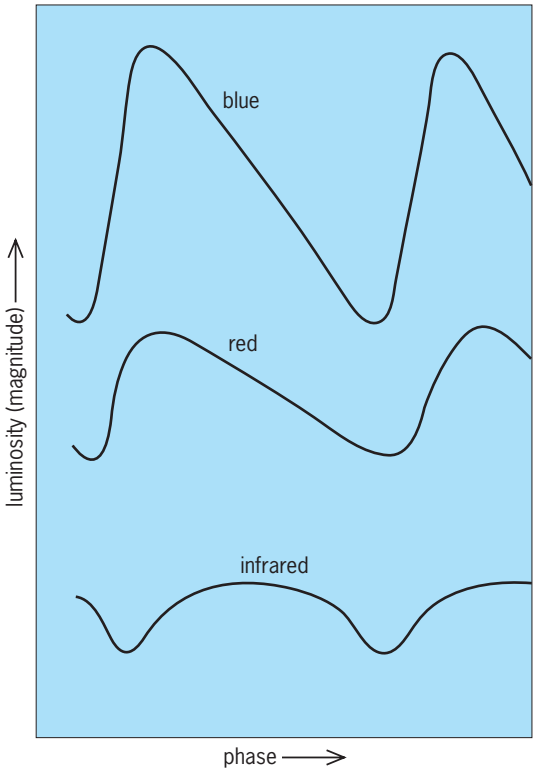

Fig. 1. Changes in Cepheid light curves as a function of wavelength.

Measurements of the color variation, interpreted as surface temperature variations, suggest that the surface temperature is changing by a few hundred kelvins during the pulsation cycle. The surface temperature changes the surface brightness and, especially at visible wavelengths, this is the primary cause of the large observed periodic light variation of a Cepheid. *See* ASTRONOMICAL SPECTROSCOPY; VARIABLE STAR.

The theory of stellar pulsation has been worked out in considerable detail, and confirmed in large measure by both linear and nonlinear computer modeling. The second helium ionization layer is now known to be the main driver of pulsation in Cepheid variables. If slightly perturbed in temperature, this zone can either add considerable energy to the flow of radiation from the center of the star, by recombining and releasing the ionization energy, or it can extract energy from the flow by ionizing new material. This situation is unstable to slight perturbations in temperature because there is an opportunity for a cycling between these two states of ionization to occur. Energy is siphoned from the flow to ionize the material. The increased number of particles (the newly liberated electrons) contribute added pressure, forcing outward the layers above. The outward motion takes the layer to cooler temperatures, which then trigger a recombination of the previously liberated electrons back onto the helium nuclei. The stored energy is released back into the flow. The pressure falls and the layer collapses. The temperature in the layer increases again, and the cycle repeats. This is a so-called heat engine, and it drives the Cepheid variability.

The interesting question remains as to what limits the variability of stars at higher and lower temperatures. If this ionization layer occurs at a point in the atmosphere where the weight of material above it is so large that it dominates the small perturbation caused by the change of state, little effect will be seen at the surface. If the same ionization layer occurs (at the same temperature) very far out in the atmosphere, there may be too little material to make any impact on the remaining atmosphere. However, at a particular placement the ionization layers can in principle impact the stability of a sizable portion of the outer layers of the star. The narrow temperature range over which the ionization layer and the mass of the atmosphere above that layer resonate with each other sets some of the fundamental limits on what observers see as the Cepheid instability strip. Although temperature is the primary cause for setting off the instability from the high-temperature side, it is generally thought that the cool-temperature limit on the strip is primarily set by the onset of convection in the atmosphere. Unfortunately, present understanding of convection is still rudimentary and details have yet to be modeled.

Types. There are several types of Cepheids. The most important distinction is between the young, high-mass classical Cepheids and their old, low-mass counterparts, the W Virginis stars. The latter belong to the population II component of the Milky Way Galaxy, and are now evolving through a completely different stage of stellar evolution than the classical Cepheids. Because their surface temperatures are virtually identical at this stage of evolution, both types of stars are subject to the same instability that allows the pulsations to occur; however, because the stars are of significantly different mass, their radii and intrinsic luminosities are quite different even for the same observed period. This ambiguity led to immense confusion early in the twentieth century, and ultimately accounted for an incorrect estimate by Hubble of the size of the universe.

Evolution. Classical Cepheids are relatively high-mass stars, 3–30 times more massive than the Sun. However, unlike the Sun, they have already exhausted the supply of hydrogen fuel in their cores and have commenced burning helium. Helium ignition occurs at a significantly higher temperature than hydrogen burning and ultimately leads the star to higher luminosities, at cooler surface temperatures and significantly larger radii. During this evolution of the star core, and the subsequent restructuring of the outer envelope and surrounding atmosphere, the supergiant star passes through a region of temperature and radius where stars are unstable to small perturbations of their atmosphere. This instability strip is narrowly defined in temperature but extends over a wide range of luminosities.

Population I stars enter the instability strip as classical Cepheids when their surface temperatures are in the 5000–7000 K range; the lower-luminosity, higher-temperature extension of the instability strip is occupied by stars called Delta Scuti variables, whose temperatures are in the 7000–8000 K range. Population II stars also pass through similar instabilities. In the low-mass, low-luminosity range are found the RR Lyrae variables (once called cluster Cepheids because of their frequent and close association with globular star clusters). RR Lyrae variables have periods in the range of 0.5–1 day. At higher luminosities are the W Virginis stars with periods up to 20 days. *See* STELLAR ROTATION.

Light curves. From an observational point of view Cepheids are known to change their light curve characteristics as a function of the wavelength at which they are observed (Fig. 1). In the green, blue, and ultraviolet wavelengths the light curves are quite asymmetric, showing a rapid rise followed by a slow, linear decline, to give a sawtooth shape. The amplitude of the variation is a declining function of wavelength, with the ultraviolet amplitudes being upward of 2 magnitudes and visual amplitudes for the same variable dropping by a factor of 2 in comparison. In the red and infrared, the amplitudes continue to decline but less rapidly than in the blue. Moreover, the shape of the light curve changes dramatically, converging on a more symmetric shape reminiscent of a cycloid rather than a sawtooth, and the maximum moves to later phases.

This change can be understood as follows. In the green and blue the dominant physics giving rise to the variability is the change in surface brightness (dominated by temperature variations). In the

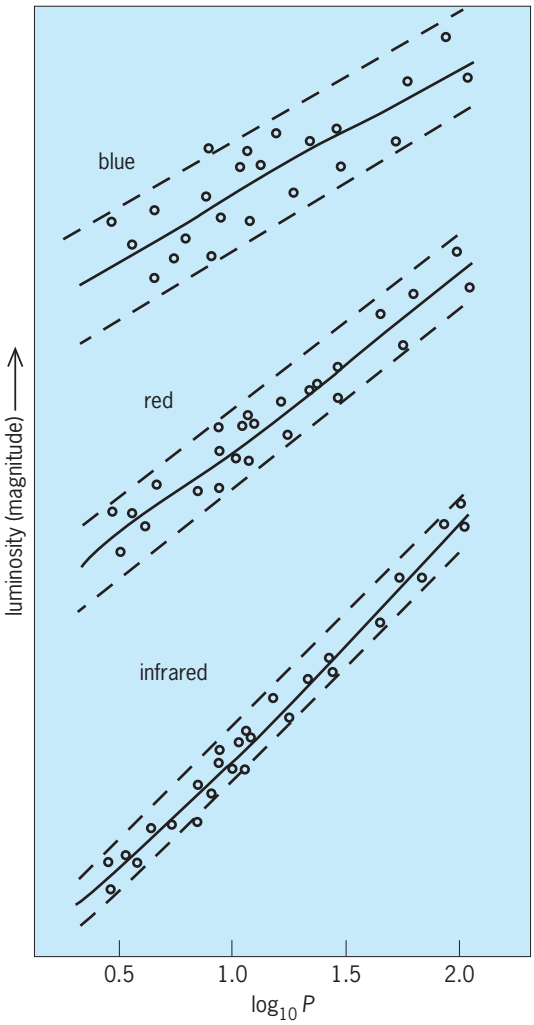

Fig. 2. Changes in the Cepheid period-luminosity relation with wavelength. Luminosity is plotted as a function of the logarithm of the period *P*, in days. In the blue the period-luminosity relation has significant scatter and its rise in luminosity with increasing period is relatively shallow. The scatter decreases with increased wavelength, while the slope of the period-luminosity relation increases toward the red and infrared.

infrared the variability primarily reflects the geometric change in the surface area of the star, being driven by the relatively small radius variations. The physics is such that temperature variations change the surface brightness far more dramatically in the blue than they do in the infrared. Analysis of the radial velocity curves for Cepheids predicts both the shape and amplitude of the radius-area variations; these correspond to what is indeed observed in the infrared. *See* LIGHT CURVES.

Period-luminosity relation. For the same physical reasons described above, the period-luminosity relations change as a function of wavelength. The width of the period-luminosity relation is a reflection of the color width of the instability strip. At a particular period, intrinsically bluer (hotter, therefore higher-surface-brightness) stars are brighter. Cooler, redder stars are intrinsically fainter. At longer wavelengths, where temperature variations play a decreasingly important role in the change of brightness of the Cepheids, the temperature-induced width of the period-luminosity relation at a fixed period is expected to decrease; this is, in fact, precisely what is observed (**Fig. 2**). Observing Cepheids at infrared wavelengths leads to an increase in accuracy in distance determinations since the dispersion (scatter) in the period-luminosity relation is decreased.

There are two other significant advantages to observing at redder wavelengths. First, the effects of extinction due to dust are reduced [by a factor of 10 when going from blue (450 nanometers) to K-band (2.2 micrometers) observations]. Finally, the effects of scattering and absorption due to the presence of metals in the atmospheres of the Cepheids are most severe at bluer wavelengths. *See* INTERSTELLAR EXTINCTION.

Distance measurements. With the aid of the Hubble Space Telescope, Cepheids have been found in galaxies as far away as the Virgo cluster, more than 20 times farther away than the Andromeda galaxy, M31, in the Local Group. At these distances, the general expansion of the universe begins to dominate the radial velocities of the galaxies. (For nearby galaxies, the motions of galaxies can be perturbed by the interaction with neighbors, or motions due to bulk flows can be a significant component of the observed velocity.) Cepheids currently provide the most accurate zero point for the calibration of other (secondary) distance methods (for example, bright supernovae) which extend the range of distance measurements by over a factor of 10. At such distances, the observed velocities are representative of the overall expansion of the universe. *See* COSMOLOGY; STAR; UNIVERSE; VARIABLE STAR. Barry F. Madore; Wendy L. Freedman

Bibliography. A. N. Cox, Cepheid masses from observations and pulsation theory, *Astrophys. J.*, 229: 212–222, 1979; M. W. Feast and A. R. Walker, The Cepheid distance scale, *Annu. Rev. Astron. Astrophys.*, 25:345–375, 1987; W. L. Freedman, The expansion rate and size of the universe, in D. H. Levy (ed.), *Scientific American Book of the Cosmos*, pp. 377–384, Scientific American, 2000; W. L. Freedman et al., Final results from the *Hubble Space Telescope* Key Project to measure the Hubble constant, *Astrophys J.*, 553:47–72, 2001; G. Jacoby et al., A critical review of selected techniques for measuring extragalactic distances, *Publ. Astron. Soc. Pacif.*, 104:599–662, 1992; B. F. Madore and W. L. Freedman, The Cepheid distance scale, *Publ. Astron. Soc. Pacif.*, 103:933–957, 1991.

Ceramics

Inorganic, nonmetallic materials processed or consolidated at high temperature. This definition includes a wide range of materials known as advanced ceramics and is much broader than the common dictionary definition, which includes only pottery, tile, porcelain, and so forth. The classes of materials generally considered to be ceramics are oxides, nitrides,

borides, carbides, silicides, and sulfides. Intermetallic compounds such as aluminides and beryllides are also considered ceramics, as are phosphides, antimonides, and arsenides. *See* INTERMETALLIC COMPOUNDS.

Ceramic materials can be subdivided into traditional and advanced ceramics. Traditional ceramics include clay-base materials such as brick, tile, sanitary ware, dinnerware, clay pipe, and electrical porcelain. Common-usage glass, cement, abrasives, and refractories are also important classes of traditional ceramics.

Advanced materials technology is often cited as an enabling technology, enabling engineers to design and build advanced systems for applications in fields such as aerospace, automotive, and electronics. Advanced ceramics are tailored to have premium properties through application of advanced materials science and technology to control composition and internal structure. Examples of advanced ceramic materials are silicon nitride, silicon carbide, toughened zirconia, zirconia-toughened alumina, aluminum nitride, lead magnesium niobate, lead lanthanum zirconate titanate, silicon-carbide-whisker–reinforced alumina, carbon-fiber–reinforced glass ceramic, silicon-carbide-fiber–reinforced silicon carbide, and high-temperature superconductors. Advanced ceramics can be viewed as a class of the broader field of advanced materials, which can be divided into ceramics, metals, polymers, composites, and electronic materials. There is considerable overlap among these classes of materials. *See* CERMET; COMPOSITE MATERIAL; GLASS; POLYMER.

Advanced ceramics can be subdivided into structural and electronic ceramics based on primary function or application. Optical and magnetic materials are usually included in the electronic classification. Structural applications include engine components, cutting tools, bearings, valves, wear- and corrosion-resistant parts, heat exchangers, fibers and whiskers, and biological implants. The electronic-magnetic-optic functions include electronic substrates, electronic packages, capacitors, transducers, magnets, waveguides, lamp envelopes, displays, sensors, and ceramic superconductors. Thermal insulation, membranes, and filters are important advanced ceramic product areas that do not fit well into either the structural or the electronic class of advanced ceramics.

Fabrication Processes

A wide variety of processes are used to fabricate ceramics. The process chosen for a particular product is based on the material, shape complexity, property requirements, and cost. Ceramic fabrication processes can be divided into four generic categories: powder, vapor, chemical, and melt processes.

Powder processes. Traditional clay-base ceramics and most refractories are fabricated by powder processes, as are the majority of advanced ceramics. Powder processing involves a number of sequential steps. These are preparation of the starting powders, forming the desired shape (green forming), removal of water and organics, heating with or without application of pressure to densify the powder, and finishing.

Powder preparation. The starting powders for traditional ceramics are predominantly clays and other natural minerals which have been processed to achieve the purity and particle size desired. The methods of mining, crushing, and milling are similar to those operations in other industries. The starting powders for advanced ceramics are high-purity specialty chemicals. These powders have many similarities to high-quality inorganic pigments. Powder preparation is a critical step in the fabrication of advanced ceramics, since high-quality advanced ceramics require high-purity, submicrometer powders with controlled physical characteristics such as particle size. The availability of high-quality, low-cost powders with reproducible characteristics is a critical requirement for use of advanced ceramic materials in many emerging applications. *See* CRUSHING AND PULVERIZING; GRINDING MILL.

Advanced ceramic powders are produced by a number of processes, including precipitation, calcination, and vapor-phase processes. The primary process is chemical precipitation, followed by calcination (heating) to remove water and other volatiles and to achieve the desired crystallinity and physical characteristics. After calcination, powders are ground, usually in ball mills, to break the bonds between individual crystallites formed by diffusion during calcination, thus producing a powder with a submicrometer particle size. In addition to conventional precipitation, a number of other precipitation processes can be used, including precipitation under pressure (hydrothermal) and simultaneous precipitation of more than one cation (coprecipitation). Hydrothermal precipitation can produce mixed cation precipitates and can precipitate oxides directly, whereas conventional precipitation usually produces hydroxides or salts which must be calcined to form oxides. Coprecipitation is used for many advanced multication powders to obtain more intimate mixing of the cations than can be achieved by mixing oxides or salts of the individual cations, as discussed below under calcination. Because of the intimate mixing produced by coprecipitation, the calcination temperature required to produce the desired crystalline phase in the powder is usually lower than for calcination of mixtures of individual oxides. *See* PRECIPITATION (CHEMISTRY).

Calcination is involved in the production of most advanced ceramic powders. For precipitated powders, a subsequent calcination step is usually required, as described above. Many multi-cation powders, such as an oxide which contains nickel, zinc, and iron [$(Ni,Zn)Fe_2O_4$], are formed by calcining mixtures of the individual oxides or salts of the cations. Many nonoxide powders can also be produced by some form of calcination. Most silicon carbide (SiC) is produced by high-temperature reaction of a mixture of sand (silicon dioxide; SiO_2) and carbon powders. Silicon nitride (Si_3N_4) powder can be produced by reacting a mixture of SiO_2 and carbon

at high temperature in a nitrogen atmosphere or by reacting silicon metal powder with nitrogen at high temperatures.

Powders for most advanced ceramics can also be produced by vapor-phase reaction. In this process, reactant gases are mixed and heated to nucleate and grow particles from the gas phase. A number of different heat sources for the reaction can be used, such as a direct-current or induction-coupled plasma, a laser, or a furnace.

Green forming. The ceramic fabrication step in which powders are formed into useful shapes is called green forming. A wide variety of processes are used for green forming. The process chosen is based on the material being fabricated, the shape desired, the property requirements, and the production cost. Green-forming processes include casting, extrusion, jiggering, die pressing, isostatic pressing, tape casting, and injection molding. Some form of treatment of the starting powders is usually required for green forming, and the treatment depends on the green-forming process and the material. Green-forming processes for clay-base ceramics usually take advantage of the plastic behavior of clay–water mixtures. Organic plasticizers are often added to achieve the desired plasticity if clay is not one of the starting powders.

Slip casting is a green-forming process in which the starting powders are suspended in water to form a slip. Green shapes are formed by pouring the slip into porous molds made of plaster of paris or another suitable material. The capillary action of the porous mold draws water from the slip and forms a solid layer on the inside of the mold. When the desired part thickness has been obtained, the excess slip is poured out of the mold (**Fig. 1**). The part is partially dried in the mold to cause it to shrink away from the mold and develop adequate rigidity for handling. This process is an economical way to form complex shapes and is used for hollow objects such as artware and sanitary ware. Slip preparation is critical to slip casting and requires optimization among a number of factors, including viscosity, solids contents, slip stability, casting rate, drying rate, and drying shrinkage. One of the critical factors in developing a slip is controlling the degree of deflocculation. This involves an understanding of the surface chemistry of the particles in the slip. *See* PORCELAIN; POTTERY.

Pressure casting has begun to replace conventional slip casting. In pressure casting, a pressure is applied to the slip to supplement or replace the capillary action of the porous mold. The advantages of pressure casting over conventional slip casting are higher casting rate, ability to automate the process, and extension of the useful application of slip casting to colloidal nonclay powders.

Extrusion is a process in which a stiff plastic mass is forced through a die that has the desired cross section of the object to be formed. The extruded column is then cut to the desired length. Secondary shaping processes are often used to further shape the extruded object. Extrusion is the primary process used to fabricate structural clay products such as building brick and clay sewer pipe. The starting materials for these products are primarily clays that have been ground into a powder and blended with 10–15% water to form a stiff mud, which is extruded. Other products which are formed by extrusion include thermocouple insulation tubes and catalyst supports for automotive emission control. In general, any material with a constant cross section can be formed by extrusion. *See* EXTRUSION.

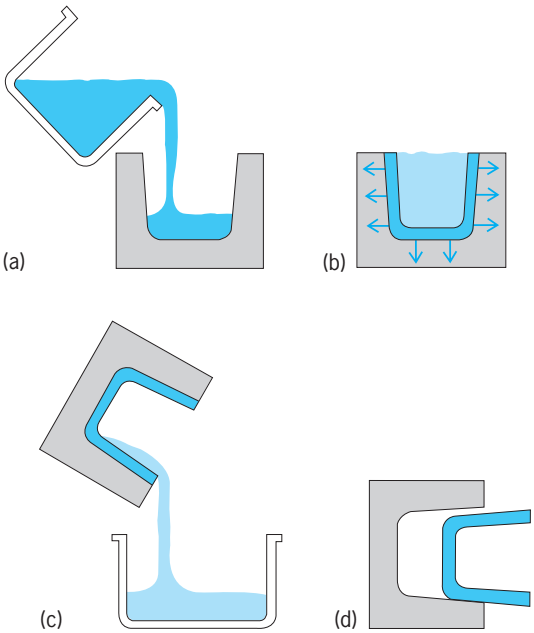

Fig. 1. Drain-casting process. (*a*) Mold being filled with slip. (*b*) Mold extracting liquid, forming compact along mold walls. (*c*) Excess slip being drained. (*d*) Casting removed after partial drying. (*After D. W. Richerson, Modern Ceramic Engineering, Marcel Dekker, 1982*)

Jiggering is a mechanization of the hand-throwing process used by potters. It consists of molding one side of a piece by placing a slab of plastic material (usually cut from an extruded log) on a porous plaster of paris mold, placing the mold and plastic slab on a rotating head, and forming the other side of the object by forcing a template of the desired contour against the plastic slab. Dinner plates are usually formed by this process, and the powder used for this product is a mixture of clays, potter's flint, and feldspar. *See* CLAY; CLAY, COMMERCIAL; FELDSPAR.

Powder pressing is a process in which a powder containing 0–5% water is green-formed by compaction under pressure. This process is used for a wide variety of traditional and advanced ceramic products, including floor and wall tile, some refractories, spark plug insulators, and many electronic ceramics. In general, if the shape permits, powder pressing is used for green forming. There are two basic types of powder pressing, die pressing and isostatic pressing. In die pressing, which is limited to relatively simple shapes, a steel or tungsten carbide die consisting of a die cavity and internal punches is used. Powder is placed in the die cavity, and force is applied through the punches to compact the powder

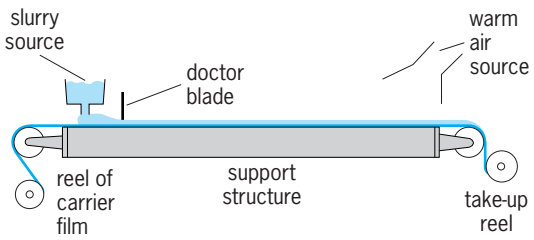

Fig. 2. Schematic diagram of doctor-blade tape casting process. (*After D. W. Richerson, Modern Ceramic Engineering, Marcel Dekker, 1982*)

into the desired shape. Isostatic pressing is used to extend green forming by powder compaction into more complex shapes. For example, spark plug insulators are pressed isostatically by placing the powder around a central mandrel pin in a flexible mold which is compressed by using a pressure vessel to apply fluid pressure to the outside of the mold. The pressed blank is then ground on an automatic grinder to form the green part to the desired dimensions and contours. Most powders used for powder pressing are wet-blended to the desired composition, which includes organic binders and lubricants, and spray-dried so that they flow. *See* ELECTRICAL INSULATION.

In tape casting or the doctor-blade process, a powder slurry is drawn under a dam (doctor blade) to form a thin layer on a moving plastic belt (**Fig. 2**). This is the primary process for forming ceramic substrates for integrated electronic circuits. An organic liquid is used as the medium for the slurry to facilitate rapid removal by subsequent evaporation to form the green tape. *See* INTEGRATED CIRCUITS.

Injection molding is a process in which a plastic blend of ceramic powder and organic materials is extruded into a steel mold to form a green shape. The process is similar to injection molding of filled plastics. However, for ceramics the organic substance must be removed so that the powder shape can be densified. Injection molding is used for forming complex shapes. Thread guides and turbocharger rotors are two ceramic products that are formed by this process.

Removal of water and organics. After green forming, a drying or organic removal step is usually required before proceeding to the densification step. The process required depends on the green-forming process and the ceramic material being fabricated. For injection-molded parts, the organic (binder) removal step is especially critical, while drying is critical for slip-cast or extruded parts.

Densification. In the densification, step the porous powder shape is converted into a dense part. The three primary processes used for final densification are sintering, hot pressing, and hot isostatic pressing. Sintering is often referred to as firing or vitrification, especially for traditional ceramics. Sintering involves heating a porous powder shape to a temperature sufficient to activate various thermal processes that lead to densification of the part. The two general mechanisms of densification during sintering are liquid-phase sintering and solid-state sintering. The driving force for sintering is reduction of the free surface energy. Most traditional ceramics are densified by liquid-phase sintering, as are alumina spark plug insulators. Materials such as dinnerware, sanitary ware, and electrical porcelain contain finely ground feldspar as one of the starting materials. When the part is heated, the feldspar melts and forms a viscous liquid which promotes densification by viscous flow. In many other ceramics, liquid phases are formed by chemical reaction between the individual starting materials. In many advanced ceramics, solid-state sintering is the predominant sintering mechanism. Solid-state sintering occurs by solid-state diffusion processes, and high-surface-area powders, generally below 1 micrometer in diameter, are especially critical to sintering by this process. *See* SINTERING.

Hot pressing is a process in which pressure is applied to a powder or powder compact in a die at elevated temperature (**Fig. 3**). The pressure enhances the thermal processes involved in sintering and often activates new processes. Dies are generally made of graphite. Materials that cannot be densified by sintering without pressure can be densified by using hot pressing. Hot pressing can also be used to produce materials with special properties, such as cutting-tool inserts. The principal disadvantages of hot pressing are cost and a limitation on the shapes that can be fabricated.

Hot isostatic pressing (HIP) is similar to isostatic compaction of powders, except that the compaction process is carried out at elevated temperatures. The equipment consists of a cold-wall autoclave with an internal furnace. Gas, typically argon, is used as the pressurizing fluid in place of the water or oil used for cold isostatic compaction of powders. Hermetically sealed metal or glass enclosures (known as cans) are used to encase the powder compact and act as the barrier to transfer the gas pressure to an isostatic

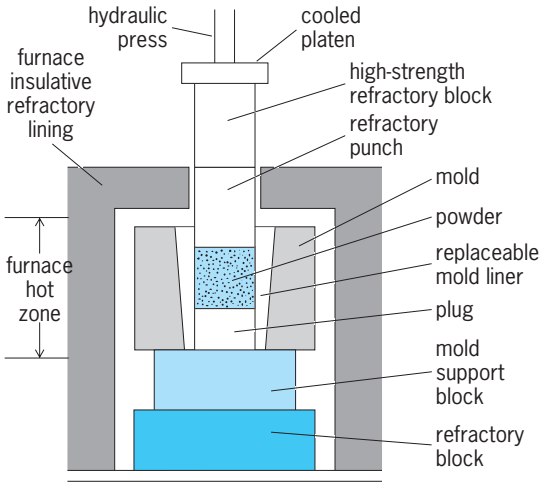

Fig. 3. Schematic diagram showing the essential elements of a hot press. (*After D. W. Richerson, Modern Ceramic Engineering, Marcel Dekker, 1982*)

compaction force on the powder compact. Hot isostatic pressing is also used to further densify materials that have been sintered to eliminate continuous or interconnected porosity. This process is referred to as post-HIP, and a can is not required since the part acts as its own can. The primary advantages of hot isostatic pressing over hot pressing are the ability to densify more complex shapes and the ability to apply higher pressures to densify a wide range of materials. *See* POWDER METALLURGY.

Finishing. After densification, many ceramic products require some finishing to achieve the required dimensional tolerance or surface finish. Since ceramics generally have high hardness, the conventional machining used for finishing metals cannot be used. Diamond cutting, grinding, and core drilling are the most common finishing techniques, but ultrasonic cutting and laser drilling and cutting are frequently used. Electric discharge machining is used if the electrical conductivity of the material is high enough.

Vapor processes. The primary vapor processes used to fabricate ceramics are chemical vapor deposition and sputtering. Vapor processes have been finding an increasing number of applications. Chemical vapor deposition involves bringing gases containing the atoms to make up the ceramic into contact with a heated surface, where the gases react to form a coating. This process is used to apply ceramic coatings to metal and tungsten carbide cutting tools as well as to apply a wide variety of other coatings for wear, electronic, and corrosion applications. Chemical vapor deposition can also be used to form monolithic ceramics by building up thick coatings. A form of chemical vapor deposition known as chemical vapor infiltration has been developed to infiltrate and coat the surfaces of fibers in woven preforms. The coating on individual fibers leads to densification of the preform to produce ceramic-ceramic composites such as silicon carbide-fiber–reinforced silicon carbide that are being developed for applications in high-temperature aerospace engines. *See* HIGH-TEMPERATURE MATERIALS; VAPOR DEPOSITION.

Several variations of sputtering and other vacuum coating processes can be used to form coatings of ceramic materials. The most common process is reactive sputtering, used to form coatings such as titanium nitride on tool steel. This process involves vaporizing titanium from a titanium target by use of a magnetron that accelerates ions such as argon against the target in order to vaporize titanium atoms, which deposit on a substrate maintained at a fixed temperature in a low-pressure plasma containing nitrogen. The titanium and nitrogen react on the substrate (such as a cutting tool) to form a titanium nitride coating. *See* MAGNETRON; SPUTTERING.

Chemical processes. A number of different chemical processes are used to fabricate advanced ceramics. The chemical vapor deposition process described above as a vapor process is also a chemical process. Two other chemical processes finding increasing application in advanced ceramics are polymer pyrolysis and sol-gel technology.

Polymer pyrolysis. This involves the controlled heating of organometallic compounds to break down the organic molecules and remove the organic radicals. The residue is a ceramic material, and the starting organometallics are known as precursors. With appropriate selection of the precursor, a wide variety of oxide and nonoxide ceramics can be formed. The process is similar to the fabrication of carbon fibers or of the matrix in carbon-carbon composites by controlled pyrolysis of pitch. Silicon carbide fibers can be manufactured by using organometallic precursors such as polycarbosilane. Similar precursors can be used to incorporate a ceramic matrix in a woven preform of silicon carbide fibers to produce silicon carbide-fiber–reinforced silicon carbide composites. *See* CERMET.

Sol-gel technology. In addition to its use as a process for manufacturing advanced ceramic powders, sol-gel technology is used to fabricate advanced ceramics. This process involves the formation of a suspension of colloidal particles (sol) which is converted to a gel by chemical treatment. The gel is then dried and sintered to form a ceramic product. The overall process in which dense materials are fabricated directly from the preformed gel without going through a powder stage is known as the sol-gel monolith process. As a result of the fine pore size of the dried gel, the sintering temperatures are lower than those required for similar materials formed from powders. Many ceramic fibers are manufactured by the sol-gel monolith process, and the process is beginning to find many other applications in advanced ceramics in biomedical, electronic, and other fields. *See* COLLOID; GEL.

Melt processes. These are used to manufacture glass, to fuse-cast refractories for use in furnace linings, and to grow single crystals. *See* REFRACTORY.

Thermal spraying can also be classified as a melt process. In this process, a plasma-spray gun is used to apply ceramic (or metal) coatings by melting and spraying powders onto a substrate. Powder particles are entrained in a gas stream in which a plasma is formed to melt (or soften) the particles as they are accelerated onto the substrate surface. For materials with lower melting points, a flame-spray gun can be used. In this process, an oxygen-acetylene flame is used in place of a plasma, and the feed material can be a powder or a rod of the material to be deposited. Ceramic coatings applied by plasma spraying are used for corrosion protection in applications such as aircraft engine parts and for wear resistance in applications such as on rotating shafts in seals and slide bearings.

Applications. The selection of materials for engineering applications is based on performance requirements and cost, as well as other factors.

The general advantages of advanced structural ceramics over metals and polymers are high-temperature strength, wear resistance, and chemical stability, in addition to the enabling functions the ceramics can perform. Typical properties for some engineering ceramics are shown in the **table**.

Typical properties for some ceramic materials

Property	Aluminum oxide	Silicon nitride	Silicon carbide	Partially stabilized zirconia
Density, g/cm^3	3.9	3.2	3.1	5.7
Flexure strength, MPa	350	850	450	790
Modulus of elasticity, GPa	407	310	400	205
Fracture toughness (K_{IC}), MPa·m$^{1/2}$	5	5	4	12
Thermal conductivity, W/mK	34	33	110	3
Mean coefficient of thermal expansion ($\times 10^{-6}$/°C)	7.7	2.6	4.4	10.2

High-temperature strength. This property of advanced ceramics is important in many potential applications, especially in components for the hot sections of engines and heat exchangers. The maximum use temperature for superalloys is 1800°F (1000°C). The only metals usable above this temperature are refractory metals that lack the oxidation resistance required for high-temperature applications in an oxidizing environment. Silicon nitride and silicon carbide have adequate oxidation resistance to be used in air up to 2600 and 2700°F (1400 and 1500°C), respectively. They also have adequate strength for many applications at these temperatures, although many silicon nitride materials are limited to use at lower temperatures because of additives that are used to aid in their fabrication. Many ceramic materials have higher melting points than silicon carbide and silicon nitride but lack adequate oxidation resistance for use at high temperatures. The other factor, which is usually critical for advanced ceramics in high-temperature applications, is thermal-shock fracture resistance. The thermal-shock fracture resistance under steady-state heat transfer is proportional to

$$\frac{\text{Strength}}{\text{Elastic modulus}} \times \frac{\text{thermal conductivity}}{\text{thermal expansion}}$$

Under highly transient conditions, the thermal conductivity is not a factor. Most engineering applications fall between the two extremes of steady-state and highly transient conditions. Comparison of the data in the table shows why silicon nitride and silicon carbide are far superior to alumina and zirconia for high-temperature applications where temperature gradients are expected. Two lower-temperature examples of an application involving thermal stress are ceramic stovetops and cookware. These materials have a thermal expansion near zero; therefore, stresses from thermal gradients are very low.

Wear resistance. Another important requirement for applications of materials is resistance to failure under impact loading. While metals generally exhibit ductible behavior, ceramic materials generally exhibit brittle failure; they are much more susceptible to failure due to high local stresses from impact loading. The toughness values in the table indicate the high toughness of partially stabilized zirconia relative to the other materials shown. In the same units, glass has a toughness of less than 1. The utilitarian fracture resistance of partially stabilized zirconia under impact loading is quite different from that of glass and even of silicon nitride. For example, a hammer made from partially stabilized zirconia can be used to drive nails without chipping, and a partially stabilized zirconia shovel was dropped onto concrete from a height of three stories without breaking. In addition to partially stabilized zirconia, other materials such as alumina can be toughened by adding zirconia particles of the correct size and composition. Ceramic whisker- and fiber-reinforced ceramics possessing high toughness have also been developed. These advances have opened additional engineering applications for advanced ceramics.

In general, ceramic materials have a much higher hardness than metals, a property that leads to high wear resistance. Valve guides and cam followers for advanced engines, sandblast nozzles, grinding media, thread guides, cutting tools, pump parts, and extrusion dies are all applications for ceramics that result from their wear resistance. *See* WEAR.

Chemical stability. Ceramic materials also have better corrosion resistance than metals or polymers, particularly in erosive wear applications. This leads to applications in valves, pumps, and pipe linings for chemical, petrochemical, and other industries. *See* CORROSION.

For electronics, advanced ceramic materials have been developed which have piezoelectric, electro-optic, pyroelectric, and many other first- and second-order effects, and thus many special electronic, optical, or magnetic properties; this allows the ceramics to be tailored for specific advanced systems, a distinct advantage. *See* PIEZOELECTRICITY.

Enabling functions. Many of the applications of advanced ceramics are based on the enabling functions they can perform in advanced systems such as automotive engines, aerospace hardware, and electronics. For example, in advanced automotive engines, silicon nitride rotors enable the design of a turbocharger with a more rapid response to the accelerator, because the density of silicon nitride is lower than that of competitive metals, and its high-temperature strength, oxidation resistance, and resistance to thermal stress are good. The incentive to develop and apply advanced ceramics often lies with the manufacturer of a system, if the value added is sufficient to justify the required investment.

The enabling function of advanced ceramics is most apparent in high-temperature superconductors. In January 1986, J. G. Bednorz and K. A. Müller discovered superconductivity in lanthanum barium copper oxide and, in April 1986, announced their finding that it occurred at temperatures as high as

35 K (−405°F). Also, in February 1986, a group including C.-W. Chu and M. K. Wu discovered superconductivity in yttrium barium copper oxide up to 92 K (−294°F). Subsequent developments have resulted in higher-temperature ceramic superconductors. These events have generated a massive research and development effort around the world. The reason for the interest in this class of superconductors is its strong potential enabling functions, many of which may not yet be conceived. *See* SUPERCONDUCTIVITY.

The primary disadvantages of most advanced ceramics are in the areas of reliability, reproducibility, and cost. Major advances in reliability are being made through development of tougher materials such as partially stabilized zirconia and ceramic whiskers; and reinforced ceramics such as silicon-carbide-whisker-reinforced alumina used for cutting tools, and silicon-carbide-fiber-reinforced silicon carbide for high-temperature engine applications. Reproducibility is being improved by developments in manufacturing technology. Cost is expected to drop as materials science, manufacturing technology, and engineering design with ceramics are further developed.

The major applications for advanced ceramics are in electronics and include substrates for electronic circuits, electronic packages, capacitors, magnetic components, piezoelectric components, electronic sensors, and integrated optic components. Advanced ceramics are also finding many applications where wear resistance and corrosion resistance are critical, such as extrusion dies and pump components. Advanced ceramic components have been introduced into automotive and other engines, and these applications are expected to grow. Automotive engine applications include turbocharger rotors, valves, valve guides, valve seats and other valve-train components, fuel-injector nozzles, exhaust-port inserts, and advanced diesel engine components. Automotive gas turbines whose hot section consists wholly of advanced ceramic components are under development, and advanced ceramic composites are expected to play a major role in the engine for the aerospace plane as well as other aerospace systems. In the machine tool industry, advanced ceramic cutting tools have increased productivity up to tenfold over metal and carbide tools, and advanced ceramic bearings are being applied where high loads and high rigidity are needed. In communications, glass fiber-optic cable has revolutionized the electronic transmission of voice and data. *See* AUTOMOTIVE ENGINE.

Advanced ceramics will continue to play key enabling functions in a wide variety of advanced systems, and traditional ceramics will continue to fill important needs in consumer, building, and industrial products. Dale E. Niesz

Bibliography. M. W. Barsoum, *Fundamentals of Ceramics*, 2002; W. D. Kingery, H. K. Bowen, and D. R. Uhlmann, *Introduction to Ceramics*, 2d ed., 1976; D. W. Richerson, *Modern Ceramic Engineering*, 3d ed., 2005; S. Saito, *Fine Ceramics*, 1987.

Cerargyrite

A mineral with composition AgCl (see **illus.**). Its structure is that of the isometric NaCl type, but well-formed cubic crystals are rare. The hardness is 2.5

Cerargyrite crystals from Leadville, Colorado. (*Specimen from Department of Geology, Bryn Mawr College*)

on Mohs scale and specific gravity 5.5. Cerargyrite is colorless to pearl-gray but darkens to violet-brown on exposure to light. It is perfectly sectile and can be cut with a knife like horn; hence the name horn silver. Bromyrite, AgBr, is physically indistinguishable from cerargyrite and the two minerals form a complete series. Both minerals are secondary ores of silver and occur in the oxidized zone of silver deposits. *See* HALITE; SILVER. Cornelius S. Hurlbut, Jr.

Ceratophyllales

An order of flowering plants (angiosperms) previously thought to be related closely to the waterlilies because one genus of the latter, *Cabomba* (Nymphaeaceae), has similarly highly dissected leaves. Studies of deoxyribonucleic acid (DNA) sequences have revealed that the single genus of the order, *Ceratophyllum* (Ceratophyllaceae), has no close relationship to any other extant group of flowering plants. It is an old, highly specialized plant, modified for a fresh-water aquatic habitat. Fossil fruits attributed to it are more than 120 million years old, which makes it the oldest extant angiosperm genus. The plants are submersed, rootless aquatics with highly dissected, branching leaves; they have reduced, separately sexed, petalless flowers. There are probably only six species distributed throughout fresh-water systems worldwide. *See* FLOWER; FOSSIL SEEDS AND FRUITS; MAGNOLIOPHYTA; PLANT EVOLUTION. Mark W. Chase

Cereal

Any member of the grass family (Gramineae) which produces edible grains usable as food by humans and livestock. Common cereals are rice, wheat, barley, oats, maize (corn), sorghum, rye, and certain millets. Triticale is cereal derived from crossing wheat and rye and then doubling the number of chromosomes in the hybrid. Occasionally, grains from other grasses

(for example, teff) are used for food. Cereals are summer or winter annuals. Corn, rice, and wheat are the most important cereals. Today, as in ancient times, cereals provide more food for human consumption than any other crops. *See* BARLEY; CORN; CYPERALES; OATS; SORGHUM; RICE; RYE; WHEAT.

Cereal grains have many natural advantages as foods. They are nutritious. One kind or another can be grown almost anywhere on Earth. The grains are not bulky and therefore can be shipped inexpensively for long distances and stored for long periods of time. They are readily processed to give highly refined raw foods.

Preference for a cereal depends on the form and flavor of the food made from it, the amount of nourishment and contribution to health, its cost, its availability, and the food habits of a people.

Food products. Four general groups of foods are prepared from the cereal grains. (1) Baked products, made from flour or meal, include pan breads, loaf breads, pastries, pancakes, flatbreads, cookies, and cakes. (2) Milled grain products, made by removing the bran and usually the germ (or embryo of the seed), include polished rice, farina, wheat flour, cornmeal, hominy, corn grits, pearled barley, bulgur (from wheat), semolina (for macaroni products), prepared breakfast cereals, and soup, gravy, and other thickenings. (3) Beverages are made from fermented grain products (distilled or undistilled) and from boiled, roasted grains. Beverages, such as beer and whiskey, made from cereals are as old as recorded history. (4) Whole-grain products include rolled oats, brown rice, popcorn, shredded and puffed grains, breakfast foods, and home-ground meals made from wheat, corn, sorghum, or one of the other cereals. *See* DISTILLED SPIRITS.

Breeding. The long evolutionary development of the cereals has resulted in a wide array of genetic types. These and new mutants have been used to breed varieties that differ in adaptation, productivity, resistance to disease and insect pests, physical properties, composition, and nutrients.

Nutritional value. All cereal grains have high energy value, mainly from the starch fraction but also from the fat and protein. In general, the cereals are low in protein content, although oats and certain millets are exceptions. High-protein genetic types have been found, and, from their use in breeding, new varieties have been developed that possess a higher percentage of protein. Likewise, a better balance has been found among the amino acids in the protein in some lines of corn, barley, and sorghum, and research with wheat is promising. Also, the nature of the starchy fraction can be altered to give products with different physical properties. *See* GRAIN CROPS; GRASS CROPS; NUTRITION.

Louis P. Reitz

Cerebral palsy

A collection of syndromes (not a disease) of nonprogressive motor dysfunction arising from abnormal development of or damage to the brain, either prenatally, at birth, or postnatally. Most cases of cerebral palsy develop in utero. Premature birth is associated with an increased risk of cerebral palsy, with the lowest birth weights carrying the highest risk. A maximum of 15% of cases are related to birth injury or perinatal oxygen deprivation. *See* PREGNANCY DISORDERS.

Although the brain damage in cerebral palsy is nonprogressive and thus deterioration does not occur, the neurological manifestations of cerebral palsy may change with neurological maturation. The precise form of cerebral palsy rarely can be characterized prior to 6 months of age; often it cannot be characterized until the individual is 2 years old. Cerebral palsy is only rarely familial. Its incidence is stable at 0.1–0.3% of live births.

Forms of handicap. Cerebral palsy is classified by the form and distribution of the motor handicap as spasticity, dyskinesia, ataxia, and hypotonia. These four forms have considerable overlap. Spasticity, seen in 75% of cases, presents a clinical picture of muscle stiffness, weakness, and imbalance of muscle tone. Common findings are contractures of joints, resulting in shortened heel cords and thus in toe walking. Tightness of the adductor muscles in the thigh may result in a scissors gait. Painful dislocation of the hip is a common problem associated with severe spasticity. Severe involvement of the brainstem may result in difficulty speaking clearly or impair feeding. Dyskinetic syndromes (athetosis or dystonia) occur in 20% of cases. Both are characterized by a severe lack of voluntary muscle control. Athetosis is characterized by slow, writhing, involuntary movements occasionally combined with quick, distal, involuntary movements (chorea). Dystonia is characterized by more or less constant muscle stiffness. Dyskinetic syndromes show no symptoms during sleep and are aggravated by emotional tension. Unclear speech (dysarthria) can be quite severe. Ataxic syndromes, characterized by impaired coordination without altered motor tone, are uncommon. A fourth syndrome, characterized by severely decreased motor tone, is called atonic cerebral palsy.

Distribution of the altered motor tone is of great importance in predicting the degree of handicap, especially in the spastic forms. (1) Hemiplegia refers to involvement of one side of the body only. (2) Paraplegia refers to involvement of both legs with no involvement of the arms and hands. These two forms constitute about 50% of the spastic cases. Individuals with them can usually walk, and 80% have normal intelligence. Most of these individuals can become independent in society. (3) Spastic quadriplegia implies equal involvement of all four extremities; individuals frequently are severely retarded and have a small head (microcephalic). (4) Diplegia refers to involvement of the legs and lesser involvement of the arms and hands. It should probably be regarded as living a transitional form between paraplegia and quadriplegia and as having an intermediate prognosis.

A specific cerebral palsy syndrome due to severe neonatal jaundice is manifested by athetosis,

deafness, and deficient upward gaze. Advances in obstetric and pediatric care have made this form uncommon. Other problems associated with cerebral palsy include mental retardation, seen in about 50% of cases. Epilepsy occurs in 25% of cases; it is treated with anticonvulsant drugs. *See* SEIZURE DISORDERS.

Treatment. Treatment of cerebral palsy is aimed at maximizing lifetime independence within the limitations of the individual's handicap. Appropriate education, in a normal class when feasible, is critical. Even the most severely affected individual can profit from training in self-care and communication skills. Common treatment modalities include physical and occupational therapy to prevent contractures and facilitate optimal motor control. Speech therapy is used to improve feeding technique and communication skills. Although these therapies are used, their long-term value has not been proven. Surgical procedures to correct contractures and improve muscle balance are valuable in the spastic syndromes. Surgical and pharmacological approaches to reducing spasticity and dyskinesia remain largely experimental.

Hart de Coudres Peterson

Bibliography. E. Blair and F. J. Stanley, Intrapartum asphyxia: A rare cause of cerebral palsy, *J. Pediat.*, 112:515–519, 1988; H. Galjaard et al. (eds.), *Early Detection and Management of Cerebral Palsy*, 1987; K. B. Nelson and J. H. Ellenberg, The asymptomatic newborn and risk of cerebral palsy, *Amer. J. Dis. Child.*, 141:1333–1335, 1987; K. B. Nelson and J. Ellenberg, Obstetric complications as risk factors for cerebral palsy or seizure disorders, *J. Amer. Med. Ass.*, 251:1983–1984, 1985; F. Stanley and E. Alberman, *The Epidemiology of the Cerebral Palsies*, 1991.

Cerenkov radiation

Light emitted by a high-speed charged particle when the particle passes through a transparent, nonconducting, solid material at a speed greater than the speed of light in the material. The blue glow observed in the water of a nuclear reactor, close to the active fuel elements, is radiation of this kind. The emission of Cerenkov radiation is analogous to the emission of a shock wave by a projectile moving faster than sound, since in both cases the velocity of the object passing through the medium exceeds the velocity of the resulting wave disturbance in the medium. This radiation, first predicted by P. A. Cerenkov in 1934 and later substantiated theoretically by I. Frank and I. Tamm, is used as a signal for the indication of high-speed particles and as a means for measuring their energy in devices known as Cerenkov counters. *See* SHOCK WAVE.

Direction of emission. Cerenkov radiation is emitted at a fixed angle θ to the direction of motion of the particle, such that $\cos\theta = c/nv$, where v is the speed of the particle, c is the speed of light in vacuum, and n is the index of refraction of the medium. The light forms a cone of angle θ around the direction of motion. If this angle can be measured, and n is known for the medium, the speed of the particle can be determined. The light consists of all frequencies for which n is large enough to give a real value of $\cos\theta$ in the preceding equation.

Cerenkov counters. Particle detectors which utilize Cerenkov radiation are called Cerenkov counters. They are important in the detection of particles with speeds approaching that of light, such as those produced in large accelerators and in cosmic rays, and are used with photomultiplier tubes to amplify the Cerenkov radiation. These counters can emit pulses with widths of about 10^{-10} s, and are therefore useful in time-of-flight measurements when very short times must be measured. They can also give direct information on the velocity of the passing particle. *See* PARTICLE DETECTOR; PHOTOMULTIPLIER.

Dielectrics such as glass, water, or clear plastic may be used in Cerenkov counters. Choice of the material depends on the velocity of the particles to be measured, since the values of n are different for the materials cited. By using two Cerenkov counters in coincidence, one after the other, with proper choice of dielectric, the combination will be sensitive to a given velocity range of particles. *See* DIELECTRIC MATERIALS.

The counters may be classified as nonfocusing or focusing. In the former type, the dielectric is surrounded by a light-reflecting substance except at the point where the photomultiplier is attached, and no use is made of the directional properties of the light emitted. In a focusing counter, lenses and mirrors may be used to select light emitted at a given angle and thus to give information on the velocity of the particle.

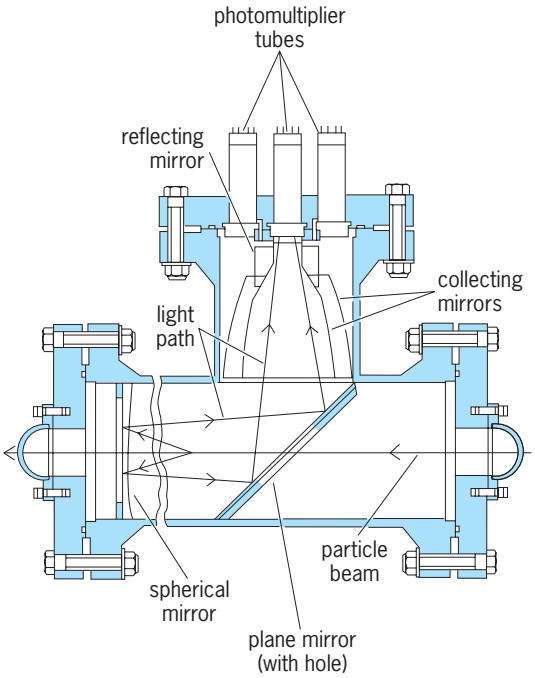

Fig. 1. Differential gas Cerenkov counter.

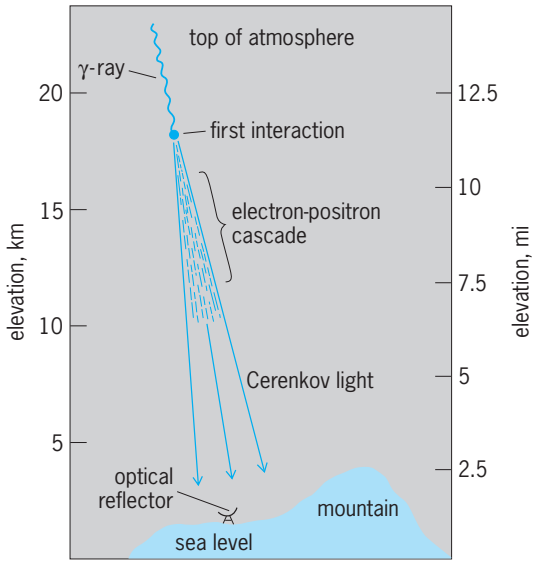

Fig. 2. Schematic view of air shower initiated by a very high energy gamma ray.

Cerenkov counters may be used as proportional counters, since the number of photons emitted in the light beam can be calculated as a function of the properties of the material, the frequency interval of the light measured, and the angle θ. Thus the number of photons which make up a certain size of pulse gives information on the velocity of the particle.

Gas, notably carbon dioxide (CO_2; **Fig. 1**), may also be used as the dielectric in Cerenkov counters. In such counters the intensity of light emitted is much smaller than in solid or liquid dielectric counters, but the velocity required to produce a count is much higher because of the low index of refraction of gas. William B. Fretter

Cerenkov gamma-ray astronomy. The properties of Cerenkov radiation have been exploited in the development of a branch of gamma-ray astronomy that covers the energy range of about 10^5–10^8 MeV. A high-energy gamma ray from a source external to the Earth creates in the atmosphere a cascade of secondary electrons and positrons (**Fig. 2**). This cascade is generated by the interplay of two processes: electron-positron pair production from gamma rays, and gamma-ray emission as the electrons and positrons are accelerated by the electric fields of nuclei in the atmosphere (bremsstrahlung). For a primary gamma ray having an energy of 10^{12} eV (1 teraelectronvolt), as many as 1000 or more electrons and positrons will contribute to the cascade. The combined Cerenkov light of the cascade is beamed to the ground over an area a few hundred meters in diameter and marks the arrival direction of the initiating gamma ray to about 1°. On a clear, dark night this radiation may be detected as a pulse of light lasting a few nanoseconds, by using an optical reflector (**Fig. 3**). *See* BREMSSTRAHLUNG; ELECTRON-POSITRON PAIR PRODUCTION.

This technique offers a means to study regions of the universe where charged particles are accelerated to extreme relativistic energies. Such regions involve highly magnetized, rapidly spinning neutron stars; supernova remnants; and active galactic nuclei. These same motivations drive the satellite observations of the EGRET instrument of the *Compton*

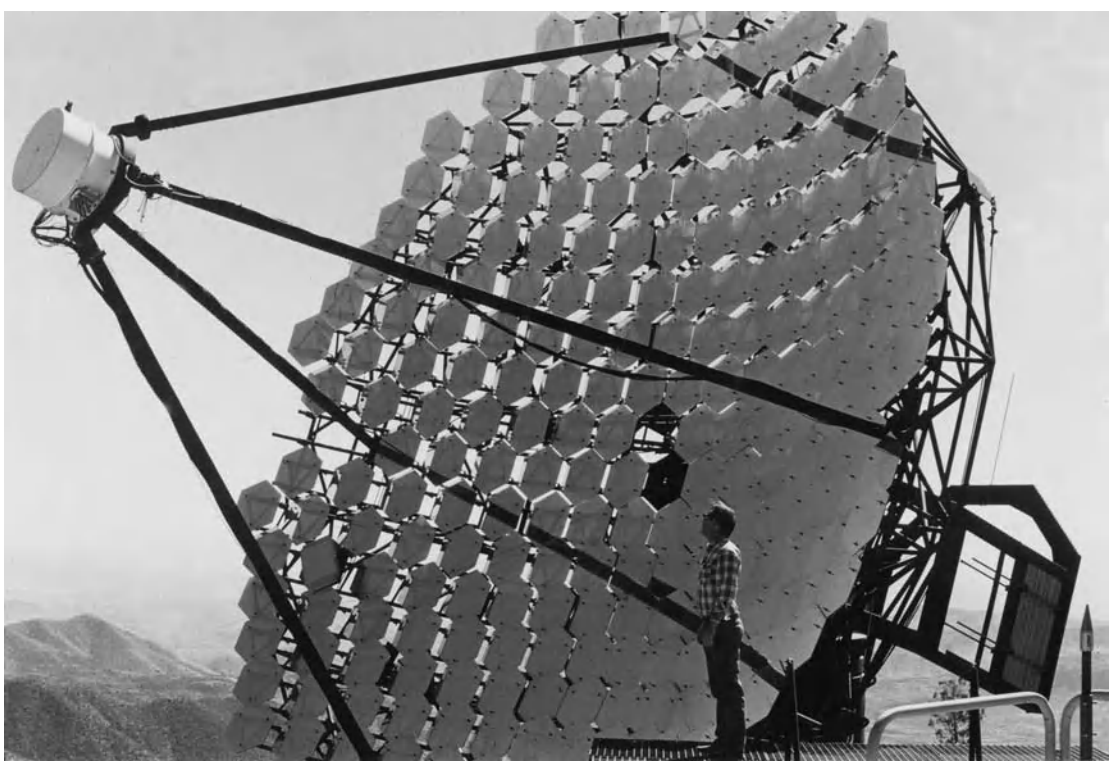

Fig. 3. The 34-ft (10-m) optical reflector at the Harvard-Smithsonian's Whipple Observatory near Tucson, Arizona.

Gamma-Ray Observatory at lower gamma-ray energies (up to about 10^4 MeV). *See* GAMMA-RAY ASTRONOMY.

The idea of using the atmosphere as an essential component of a very high energy gamma-ray detector dates from the 1960s. However, the development of the field was hindered for several decades by the enormous background to any possible gamma-ray signal that arises from cosmic-ray-induced air showers. Such showers are hundreds of times more numerous (within a 1° field of view) than the showers from even the strongest gamma-ray source. However, the technique remains attractive since it appears to be the only viable technique for gamma-ray astronomy in this energy range. *See* COSMIC RAYS.

Rapid development of the field dates from the late 1980s and follows the recognition that cosmic-ray air showers differ substantially from gamma-ray air showers. In order to distinguish between them, it is necessary to image each individual air shower. The imaging is accomplished through the use of an array of fast photomultipliers, operating in the focal plane of the gamma-ray telescope. Comparison of observed image parameters for each shower with the image expected for a gamma-ray shower results in rejection of more than 99% of background cosmic-ray showers, while more than 50% of gamma-ray showers are retained.

A number of imaging Cerenkov telescopes are in operation with, in most cases, focal-plane imaging detectors consisting of more than 100 pixels. By 1998, seven sources had been reliably observed with good signal-to-noise ratios.

Spin-powered neutron stars. Three of these sources are associated with spin-powered neutron stars, the Crab pulsar, Vela pulsar, and PSR B1706-44, all located within the Milky Way Galaxy less than 7000 light-years from the Earth. Unlike earlier detections, however, there is no evidence for pulsations at the spin periods of the neutron stars. (This behavior is in contrast to observations at lower gamma-ray energies from the EGRET instrument, in which the gamma-rays are dominantly pulsed.) For these sources the unpulsed gamma rays are presumably associated with a surrounding nebula. For the Crab, the nebula is clearly observed, but for the other two objects there is some ambiguity concerning the presence of any nebulae.

The observations of the Crab Nebula at teraelectronvolt energies combined with observations at lower gamma-ray energies generally support a theoretical model of the high-energy emission process in which the highest-energy radiation is due to Compton scattering between extremely relativistic electrons, with energies up to about 10^{16} eV, and lower-energy photons in the nebula. In this picture, the lower-energy photons have been created by the relativistic electrons themselves as they spiral in the magnetic field of the nebula. The underlying power for all of this high-energy activity is the rotation of a rapidly spinning neutron star (spin period 33 milliseconds) which itself is the remnant of the relatively recent supernova explosion in the year 1054. *See* COMPTON EFFECT; CRAB NEBULA; NEUTRON STAR; PULSAR; SUPERNOVA.

Shell-type supernova remnants. Another of the very high energy gamma-ray sources is a shell-type supernova remnant associated with a historical supernova of the year 1006. Shell-type supernova remnants are suspected to be the dominant source of cosmic rays with energies less than about 10^{14} eV. The shock waves that occur at shell boundaries accelerate cosmic rays confined by magnetic fields near the shocks to these energies. Although SN 1006 is a shell-type supernova, it does not meet expectations. The teraelectronvolt gamma rays appear to arise principally from electrons rather than the protons and heavier nuclei that make up the bulk of the cosmic rays, since the location of the teraelectronvolt gamma-rays generally coincides with the location of nonthermal x-ray emission attributed to relativistic electrons.

Observations of other shell-type supernova remnants have failed to detect any very high energy gamma-ray emission with, in some cases, upper limits that are slightly below theoretical expectations. If the common view of the origins of cosmic rays is correct, the next generation of imaging Cerenkov detectors should easily detect the nearer shell-type remnants and map their teraelectronvolt photon emission.

Active galactic nuclei. The remaining sources seen by atmospheric Cerenkov telescopes are all extragalactic objects associated with active galactic nuclei: Markarian 421, Markarian 501, and 1ES 2344 + 514. Active galactic nuclei in general are thought to be powered by supermassive black holes (10^6–10^9 solar masses) accreting material from a surrounding disk of gas. In many cases, a jet of relativistic material and radiation is formed perpendicular to the accretion disk. In the case of the objects seen at teraelectronvolt energies, and at 100 MeV–1 GeV energies as well, the jets are apparently directed toward the Earth, so that individual gamma rays are boosted in energy and the apparent luminosity of the sources is enhanced by several orders of magnitude. Active galactic nuclei with jets aligned near to the line of sight to the Earth are called blazars. *See* BLACK HOLE.

Variations at very short time scales have been observed in the emission of the blazars. The gamma-ray flux from Markarian 421, for example, has been observed to rise and fall by a factor of 2 within 15 min. Because of the large collection area associated with Cerenkov telescopes (Fig. 2), they have the ability to probe for flux variations on a time scale unobtainable by any other detector operating above a few megaelectronvolts. The observations of Markarian 421 show the the teraelectronvolt gamma-ray emission is best characterized by succession of rapid flares with a baseline level below the sensitivity limit of the detector, consistent with little steady emission.

When the spectrum obtained at teraelectronvolt energies is combined with lower-energy spectra, a consistent picture emerges in which the teraelectronvolt photons are due to Compton scattering of radiation from a population which is dominantly

relativistic electrons. However, many questions remain, and alternative scenarios in which the dominant particle species are protons are possible. The fact that gamma rays up to energies of 10^{12} eV are observed suggests that this type of source is a copious producer of high-energy cosmic rays. Their contribution to the observed cosmic ray flux at Earth has yet to be assessed. *See* ASTROPHYSICS, HIGH-ENERGY; GALAXY, EXTERNAL. Richard C. Lamb

Bibliography. F. A. Aharonian and C. W. Akerlof, Gamma-ray astronomy with imaging atmospheric Cerenkov telescopes, *Annu. Rev. Nucl. Part. Sci.*, 47:273–314, 1997; R. Ong, Very high energy gamma-ray astronomy, *Phys. Rep.*, 305:93–202, 1998.

Ceres

The largest asteroid in the "main belt" of asteroids between Mars and Jupiter, and the first such object to be discovered.

Discovery. The reason for the apparent gap devoid of planets between Mars and Jupiter was a matter of speculation since Johannes Kepler's time. In the latter part of the eighteenth century, the Titius-Bode law (an empirical law for the spacing of planetary distances from the Sun) suggested that there should be a planet in the gap. The Titius-Bode law seemingly was supported by the location of Uranus discovered by William Herschel in 1781, intensifying the search for the "missing planet" in the region between Mars and Jupiter. On January 1, 1801, Giuseppe Piazzi serendipitously discovered the first minor planet: Ceres. At the time, it was believed to be a planet; only after the discovery in the next few years of more such objects and progress in estimating their sizes was it clear that they were much smaller than planets, and Herschel suggested that these objects be called "asteroids" rather than "planets." Today, the term "minor planets" is used to describe asteroids in the main belt as well as other small solar system bodies such as Kuiper Belt objects and near-Earth objects. *See* KUIPER BELT; PLANET; URANUS.

Properties. Ceres is a G-type asteroid; these asteroids are very dark and have a nonvolatile chemical composition similar to the Sun. Ceres orbits the Sun with a 4.4-year period, a semimajor axis of 2.7 AU (astronomical unit, the mean distance from the Sun to the Earth; 1 AU = 149,597,871 km = 92,955,807 mi), and an eccentricity of 0.097. It has a rotational period of 9.075 hours, and is a prograde rotator (that is, the direction of its rotation is in the same sense as that of the Earth and of the direction of revolution of the planets around the Sun).

Ceres has a visual geometric albedo of about 0.10, and has a red color (meaning that it tends to reflect light more strongly at longer wavelengths) from the visible part of the spectrum through much of the ultraviolet. The brightness of an asteroid is a function of how much light it reflects due to its size and albedo. Although Ceres is the largest asteroid, it is not the brightest. Its absolute magnitude of $H = 3.32$ makes it somewhat fainter than Vesta, which is brighter because it is more reflective. (Vesta's albedo is about four times greater than that of Ceres.) The absolute magnitude of an asteroid (H) is defined differently from that of a star for practicality; H is defined as the magnitude an object would have if it were 1 AU from the Sun and 1 AU from the Earth, at zero phase angle (fully illuminated, as with a full moon). This alignment is a physical impossibility but is useful as a definition.

With a mass of about 9.5×10^{20} kg (2.1×10^{21} lb, or about 0.00016 that of the Earth), Ceres contains about 25% of the combined mass of all the main-belt asteroids (that is, not including objects in the Kuiper Belt). That mass and the measured diameter of roughly 950 km (590 mi) implies a density around 2.1 g/cm^3, roughly twice the density of water.

Ceres, like Vesta, Pallas, and other large asteroids, is a surviving protoplanet from the era of accretion in the inner solar system. It, like Vesta, is thought to have retained its primitive physical state due to the presence of water that, although causing some aqueous alteration, prevented it from melting like Vesta. Consequently, Ceres may well have the oldest intact surface in the inner solar system.

Hubble Space Telescope observations. The *Hubble Space Telescope* has observed Ceres several times, and an image from one observation is shown in the **illustration**. These observations were key in identifying the first large feature on Ceres, informally named "Piazzi" in honor of the discoverer of Ceres.

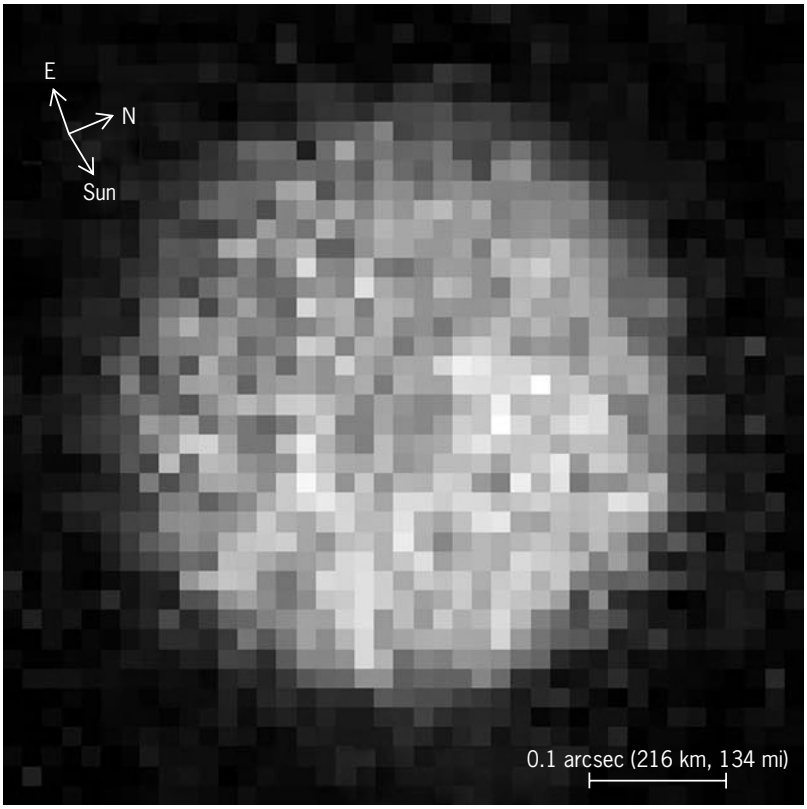

Hubble Space Telescope **ultraviolet image of Ceres. The darkened spot near the center is the Piazzi feature. Arrows show the direction of the Sun, to provide an idea of the direction of illumination, and the North-East orientation of the image. However, the Sun is actually nearly behind the head of the reader looking at the image, so the arrow showing its direction should be pointing out of the page and a little down to the right. (*Southwest Research Institute, NASA*)**

The Piazzi feature is roughly circular with a diameter of about 250 km (150 mi). However, it is not known whether this feature is a crater, an albedo "spot," or another effect.

Hubble observations also were used to analyze the internal structure of Ceres. Those observations showed that Ceres has a shape and smoothness indicative of a gravitationally relaxed object. Its shape is significantly less flattened than that expected for a homogeneous object, but is consistent with a central mass concentration indicative of differentiation. Possible interior configurations include water-ice-rich mantles over a rocky core. *See* ASTEROID; HUBBLE SPACE TELESCOPE.

Dawn mission. Ceres is the target of the National Aeronautics and Space Administration's Discovery mission *Dawn*, which was scheduled to launch in 2007 to orbit and study Vesta and Ceres (in 2010–2012 and 2015, respectively). The mission's goal is to characterize the conditions and processes of the solar system's earliest epoch by investigating in detail these two of the largest protoplanets that have remained intact since their formations. *See* ASTEROID; SPACE PROBE. Joel Parker

Bibliography. J. Cunningham, *The First Asteroid: Ceres 1801–2001*, Historical Studies in Asteroid Research, vol. 1, Star Lab Press, 2001; J. W. Parker et al., Analysis of the first disk-resolved images of Ceres from ultraviolet observations with the *Hubble Space Telescope, Astron. J.*, 123:549–557, 2002; P. C. Thomas et al., Differentiation of the asteroid Ceres as revealed by its shape, *Nature*, 437:224–226, 2005.

Ceriantharia

An order of the Zoantharia, typified by *Cerianthus*, which lives in sandy marine substrata (**illus.** *a* and *b*). The animal is enclosed in a sheath formed by mucus secreted from gland cells of the column ectoderm, in which discharged nematocysts, sand grains, and other foreign objects are embedded.

The polyp is a muscular, skeletonless, elongated, cylindrical body with a smooth wall. Long, slender, unbranching, freely retractile tentacles are arranged in two cycles and consist of smaller labial and larger marginal ones. A siphonoglyph termed the hyposulculus is present and a pedal disk is lacking.

The mesenteric arrangement is most characteristic (illus. *c*). Mesenteries are arranged bilaterally in couples, not in pairs, forming a single cycle. All are complete. The three most dorsal couples or directives appear first, and then the other two appear at both sides of the directives. Additional metacnemic couples are added to the ventral intermesenteric space and are arranged in quartets with regard to length. The longitudinal muscles are weakly developed in the ventral faces of the mesenteries, whose free edges bear filaments and acontialike threads or acontioids that seem to have an adhesive function.

Pelagic larvae (illus. *d–f*) known as *Arachnactis, Cerianthula, Ovactis, Anactinia*, and others always pass through the cerinula stage with three couples of the protocnemes. Life cycles have not been fully investigated. *See* HEXACORALLIA. Kenji Atoda

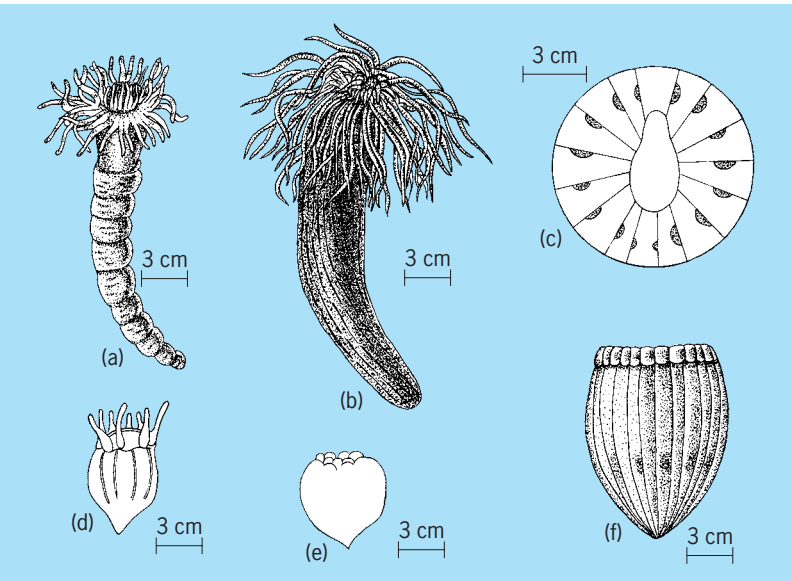

Ceriantharia. (*a*) *Cerianthus solitarius*. (*b*) *Pachycerianthus multiplicatus*. (*c*) Mesenteric arrangement. (*d*) *Cerianthula spinifer* larva. (*e*) *Ovactis aequatorialis* larva. (*f*) *Anactinia pelagica* larva.

Cerium

A chemical element, Ce, atomic number 58, atomic weight 140.12. It is the most abundant metallic element of the rare-earth group in the periodic table. The naturally occurring element is made up of the isotopes ^{136}Ce, ^{138}Ce, ^{140}Ce, and ^{142}Ce. A radioactive α-emitter, ^{142}Ce has a half-life of 5×10^{15} years. Cerium occurs mixed with other rare earths in many minerals, particularly monazite and blastnasite, and is found among the products of the fission of uranium, thorium, and plutonium. *See* PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Although the common valence of cerium is 3, it also forms a series of quadrivalent compounds and is the only rare earth which occurs as a quadrivalent ion in aqueous solution. Although it can be separated from the other rare earths in high purity by

ion-exchange methods, it is usually separated chemically by taking advantage of its quadrivalent state.

Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

Cermet

A group of composite materials consisting of an intimate mixture of ceramic and metallic components.

Fabrication. Cermets can be fabricated by mixing the finely divided components in the form of powders or fibers, compacting the components under pressure, and sintering the compact to produce physical properties not found solely in either of the components. Cermets can also be fabricated by internal oxidation of dilute solutions of a base metal and a more noble metal. When heated under oxidizing conditions, the oxygen diffuses into the alloy to form a base metal oxide in a matrix of the more noble metal. *See* CORROSION; POWDER METALLURGY.

Components. Ceramic components may be metallic oxides, carbides, borides, silicides, nitrides, or mixtures of these compounds; the metallic components include a wide variety of metals whose selection depends on the application of the respective cermet. Some of the component materials are shown in the **table**.

Interactions. The reactions taking place between the metallic and ceramic components during fabrication of cermets may be briefly classified and described as follows: (1) Heterogeneous mixtures with no chemical reaction between the components, characterized by a mechanical interlocking of the components without formation of a new phase, no penetration of the metallic component into the ceramic component, and vice versa, and no alteration of either component (example, MgO–Ni). (2) Surface reaction resulting in the formation of a new phase as an interfacial layer that is not soluble in the component materials. The thickness of this layer depends on the diffusion rate, temperature, and time of the reaction (example, Al_2O_3–Be). (3) Complete reaction between the components, resulting in the formation of a solid solution characterized by a polyatomic structure of the ceramic and the metallic component (example, TiC–Ni). (4) Penetration along grain boundaries without the formation of interfacial layers (example, Al_2O_3–Mo). *See* DIFFUSION; PHASE EQUILIBRIUM.

Representative components of cermets

Class	Ceramic	Metal addition
Oxides	Al_2O_3	Al, Be, Co, Co-Cr, Fe, stainless steel
	Cr_2O_3	Cr
	MgO	Al, Be, Co, Fe, Mg
	SiO_2	Cr, Si
	ZrO_2	Zr
	UO_2	Zr, Al, stainless steel
Carbides	SiC	Ag, Si, Co, Cr
	TiC	Mo, W, Fe, Ni, Co, Inconel, Hastelloy, stainless steel, Vitallium
	WC	Co
	Cr_3C_2	Ni, Si
Borides	Cr_3B_2	Ni
	TiB_2	Fe, Ni, Co
	ZrB_2	Ni
Silicides	$MoSi_2$	Ni, Co, Pt, Fe, Cr
Nitrides	TiN	Ni

Bonding behavior. One important factor in the selection of metallic and ceramic components in cermets is their bonding behavior. Bonding may be by surface interaction or by bulk interaction. In cermets of the oxide-metal type, for example, investigators differentiate among three forms of surface interaction: macrowetting, solid wetting, and wetting assisted by direct lattice fit.

Combinations. One distinguishes basically between four different combinations of metal and ceramic components: (1) the formation of continuous interlocking phases of the metallic and ceramic components, (2) the dispersion of the metallic component in the ceramic matrix, (3) the dispersion of the ceramic component in the metallic matrix, and (4) the interaction between the metallic and ceramic components.

Characteristics. The combination of metallic and ceramic components can result in cermets characterized by increased strength and hardness, higher temperature resistance, improved wear resistance, and better resistance to corrosion, each characteristic depending on the variables involved in composition and processing. In general, these materials should be corrosion resistant, have high-temperature strength, temperature (thermal) shock resistance, and a certain ductility.

The yield strength σ of a cermet depends on the component materials, the volume fraction f of the dispersed phase, the particle diameter d of the dispersed material, and the mean spacing λ between particles. These variables are related by Eqs. (1) and (2), where A and B are material constants.

$$\lambda = \frac{2d}{3f}(1 - f) \tag{1}$$

$$\sigma = -A \log \lambda + B \tag{2}$$

Applications. Friction parts as well as cutting and drilling tools have been successfully made from cermets for many years. Certain nuclear reactor fuel elements, such as dispersion-type elements, are also made as cermets.

Fiber reinforcement. Fiber-reinforced cermets consist either of a metallic matrix reinforced by ceramic fibers (for high-temperature strength) or a ceramic matrix with metallic fibers inserted (for better heat conductivity). The term fiber refers to a multicrystalline material approximately 0.5–2.5 micrometers (0.02–0.1 mil) in diameter. Whiskers which can be used instead of fibers are short single crystals

approximately 1–10 μm in diameter. Combinations of powders and fibers can be made with short or continuous fibers, randomly dispersed or aligned (oriented). Special materials with directional properties can be made by insertion of aligned fibers.

Bulk interactions. Two different forms of bulk interaction between metals and ceramics can be distinguished; solid solution and formation of chemical compounds. Usually solid-solution bonding involves the addition, or formation, within the cermet of a small amount of the appropriate ceramic form of the metal constituent; examples of the type of phase involved are provided by the systems Al_2O_3–Cr_2O_3 and NiO–MgO for oxides, and TaC–Tic and NbC–ZrC for carbides. The systems form continuous series of solid solutions, but there are also many suitable systems in which solid solution occurs over a limited range only.

Formation of a compound in a bonding phase can best be shown by the examples of spinels, having the generalized formula $RO \cdot R'_2O_3$, where R could, for example, stand for Ni^{2+}, Mg^{2+}, Fe^{2+}, or Co^{2+}, and R′ for Al^{3+}, Cr^{3+}, or Fe^{3+}. Much study has been made of the system Al_2O_3–Fe, without achieving the combination of properties required for the high-temperature engineering applications in view. *See* COMPOSITE MATERIAL; METAL MATRIX COMPOSITE. Henry H. Hausner

Bibliography. J. Pask and A. Evans (eds.), *Surfaces and Interfaces in Ceramic and Ceramic-Metal Systems*, 1981; P. Naylor, *Introduction to Metal Ceramic Technology*, 1992; G. S. Upadhyaya (ed.), *Sintered Metal-Ceramic Composites*, 1985.

Cerussite

A common mineral in the upper oxidized zone of ore deposits containing galena (PbS). It is usually associated with anglesite, smithsonite, malachite, and other secondary minerals. Normally cerussite is close in composition to $PbCO_3$ with only minor amounts of zinc (Zn) and strontium (Sr) in place of lead (Pb). *See* ANGLESITE; MALACHITE; SMITHSONITE.

Cerussite has the aragonite-type structure with $a = 5.180$Å, $b = 8.492$Å, $c = 6.134$Å, $Z = 4$, and space group Pmcn. The average C-O bond length is 1.27 Å and O-C-O bond angle is 121.3°, which are distinct in comparison with those in alkaline earth carbonates of the aragonite-type structure. The average Pb-O bond length is 2.69Å. *See* ARAGONITE.

This mineral is biaxial and optically negative with $\alpha = 1.803$, $\beta = 2.074$, $\gamma = 2.076$, and $2V_\alpha = 9°$. Birefringence is extreme (0.273) and is structurally dependent. The mineral is usually white or gray, and often brown or black. It exhibits luminescence of considerable intensity with a broad emission band centered at 470 nanometers. The hardness of cerussite is between 3 and $3^1/_2$ on the Moh's scale, and its specific gravity is 6.55, which is a distinguishing feature of cerussite in comparison with other carbonate minerals of the aragonite-type structure. *See* BIREFRINGENCE; HARDNESS SCALES; LUMINESCENCE; SPECIFIC GRAVITY.

Cerussite decomposes to $PbCO_3 \cdot 2PbO$ at about 200°C (392°F) and CO_2 pressure lower then 1 atm (14.7 lb/in.^{2}). At medium pressures (~1 atm), two intermediates, $PbCO_3 \cdot PbO$ and $PbCO_3 \cdot 2PbO$, were formed at 300°C and 400°C (572°F and 752°F), respectively. Cerussite decomposes to PbO via three intermediates, $2PbCO_3 \cdot PbO$, $PbCO_3 \cdot PbO$, and $PbCO_3 \cdot 2PbO$, at pressures higher than 4 atm (58.8 lb/in.^{2}).

Cerussite generally occurs in crystal aggregates with a varied habit. Twins are common, in reticulated formations (networks), in starlike groups, or in sixling formations. Luke L. Y. Chang

Bibliography. L. Chang, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, vol. 5B: *Non-silicates*, 2d ed., Longman, 1996.

Cesium

A chemical element, Cs, with an atomic number of 55 and an atomic weight of 132.905, the heaviest of the alkali metals in group 1 of the periodic table (except for francium, the radioactive member of the alkali metal family). Cesium is a soft, light, very low-melting metal. It is the most reactive of the alkali metals and indeed is the most electropositive and the most reactive of all the elements.

Cesium reacts vigorously with oxygen to form a mixture of oxides. In moist air, the heat of oxidation may be sufficient to melt and ignite the metal. Cesium does not appear to react with nitrogen to form a nitride, but does react with hydrogen at high temperatures to form a fairly stable hydride. Cesium reacts violently with water and even with ice at temperatures as low as −116°C (−177°F). Cesium reacts with the halogens, ammonia, and carbon monoxide. In general, cesium undergoes some of the same type of reactions with organic compounds as do the other alkali metals, but it is much more reactive. *See* SODIUM.

The physical properties of cesium metal are summarized in the **table**.

Cesium is not very abundant in the Earth's crust, there being only 7 parts per million (ppm) present. Like lithium and rubidium, cesium is found as a constituent of complex minerals and not in relatively pure halide form as are sodium and potassium.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Physical properties of cesium metal

Property	Temp., °C	Valve
Density	20	1.9 g/cm^3
Melting point	28.5	
Boiling point	705	
Heat of fusion	28.5	3.8 cal/g
Heat of vaporization	705	146 cal/g
Viscosity	100	4.75 millipoises
Vapor pressure	278	1 mm
	635	400 mm
Thermal conductivity	28.5	0.044 cal/(s)(cm^2)(°C)
Heat capacity	28.5	0.06 cal/(g)(°C)
Electrical resistivity	30	36.6 microhm–cm

Indeed, lithium, rubidium, and cesium frequently occur together in lepidolite ores, such as those from Rhodesia.

Cesium metal is used in photoelectric cells, spectrographic instruments, scintillation counters, radio tubes, military infrared signaling lamps, and various optical and detecting devices.

Cesium compounds are used in glass and ceramic production, as absorbents in carbon dioxide purification plants, as components of getters in radio tubes, and in microchemistry. Cesium salts have been used medicinally as antishock agents after administration of arsenic drugs. The isotope cesium-137 is supplanting cobalt-60 in the treatment of cancer. *See* ALKALI METALS. Marshall Sittig

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; W. A. Hart et al., *The Chemistry of Lithium Sodium, Potassium, Rubidium, Cesium, and Francium*, 1975; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004.

Cestida

An order of the phylum Ctenophora (comb jellies) containing two monospecific genera, *Cestum* and *Velamen*, which have an unusual morphology. The transparent bodies of cestids are flattened in the tentacular plane and greatly elongated in the stomodeal plane so that they have the shape of a belt or ribbon (hence the specific name *Cestum veneris*, "Venus' girdle"; see **illus.**). The mouth is at the midpoint of the length. Individuals of *Cestum* may be more than 39 in. (1 m) long, and of *Velamen* about 8 in. (20 cm). Brown or yellow pigment spots sometimes occur on the tips of the body. If disturbed, Cestids, like other ctenophores, produce brilliant bioluminescence along their meridional canals. They occur worldwide in tropical and subtropical waters; *C. veneris* is one of the most common species in the oceanic Atlantic. The extreme fragility of their bodies makes cestids difficult to collect and maintain, and living specimens have been seen by few biologists.

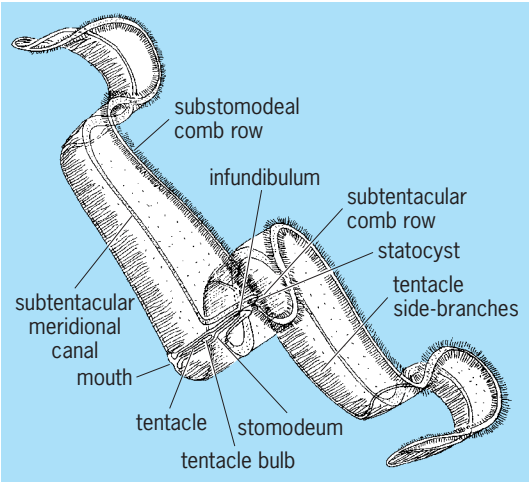

Structure of *Cestum veneris*.

Because of the great lateral elongation of the body, the subtentacular comb rows are reduced to a few ctenes, but the substomodeal rows run along the entire aboral edge of the body. The beating of the comb rows propels the cestid slowly forward like a wing, with the oral edge leading. Two main tentacles arise on either side of the central stomodeum and lie in a groove along the oral edge; numerous fine side branches (tentilla) from these tentacles trail back over the flat sides of the body. The prey, mainly small crustaceans, are captured when they contact the body surface and stick to the tentilla. Contraction of the tentilla moves the prey into the oral groove, which transports it by ciliary action to the mouth. This feeding mechanism allows cestids to catch prey ranging from about 100 micrometers to several millimeters in size. It is energetically efficient, and fishing is continuous, without interruptions to transfer or handle prey. Cestids are also capable of rapid undulatory swimming along their long axis, mainly as an escape response to disturbance. They are frequently parasitized by hyperiid amphipods; predators on cestids include medusae, the ctenophore *Beroë*, and various fishes.

Like other ctenophores, cestids are simultaneous hermaphrodites, with ovaries and testes located along the meridional canals. Very little is known about the timing of their reproduction. Fertilized eggs develop first into typical cydippid larvae but soon begin to elongate and take on the typical cestid form. *See* CTENOPHORA. Laurence P. Madin

Bibliography. G. R. Harbison, L. P. Madin, and N. R. Swanberg, On the natural history and distribution of oceanic ctenophores, *Deep-sea Res.*, 25:233–256, 1978; F. W. Harrison and J. A. Westfall (eds.), *Microscopic Anatomy of Invertebrates*, vol. 2, 2000; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 1982.

Cestoda

A class of the phylum Platyhelminthes commonly referred to as tapeworms. All members are endoparasites, usually in the digestive tract of vertebrates. The class has been subdivided as follows:

Class Cestoda
- Subclass Cestodaria
  - Order: Amphilinidea
    - Gyrocotylidea
- Subclass Eucestoda
  - Order: Proteocephaloidea
    - Tetraphyllidea
    - Lecanicephaloidea
    - Trypanorhyncha
    - Diphyllidea
    - Pseudophyllidea
    - Cyclophyllidea
    - Nippotaeniidea

See separate articles on each group listed.

Morphology and life history. In size the tapeworms range from less than 0.04 in. (1 mm) to several feet in length. The class is differentially characterized by the presence of a cuticle rather than a cellular epidermis and by the total absence of a mouth and digestive tract. External cilia are never present, although the cuticle usually bears minute projections ranging from 0.1 to 1.0 micrometer in diameter. In most species of the class the body is divided into proglottids, each proglottid containing one or two hermaphroditic reproductive systems (**illus.** *a*). The anterior end is usually modified into a holdfast organ, bearing suckers or sucking grooves, and frequently armed with hooks. Early embryonic development occurs in the parental body, usually in a uterus, to a hook-bearing stage, the oncosphere (illus. *b*). The oncosphere leaves the parental body through a uterine pore or by liberation of the terminal segment from the main body of the worm. Further development of the worm always occurs within the body of a host, most often an invertebrate, which commonly ingests the larval form. Growth of the larval worm, frequently with differentiation of the holdfast organ, occurs in this host; in some species asexual reproduction by budding occurs. Further larval growth and development of the holdfast may require a second host. Development of the sexual phase from the larva occurs in another host. Growth with strobilation and development of reproductive systems follow, the worm usually staying in the digestive tract. *See* PSEUDOPHYLLIDEA.

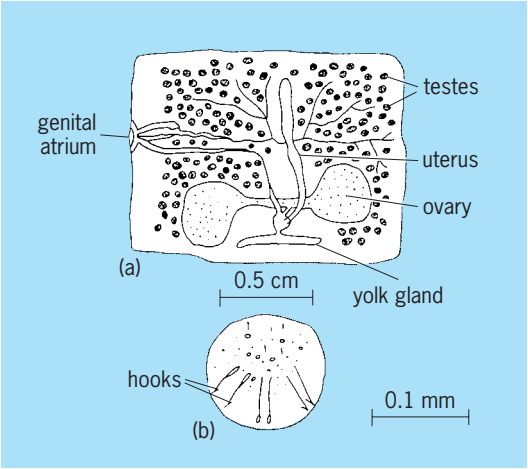

Cestoda. (*a*) Sexually mature segment of *Taenia*. (*b*) Oncosphere of a cyclophyllidean tapeworm.

Nutrition. Food materials are presumed to be absorbed through the external surface. It has been shown that tapeworms have catalytic mechanisms for the active transport into the body of simple sugars and amino acids. Elaboration of digestive enzymes has not been demonstrated. The chemical composition of these worms is unique in that polysaccharide and fat represent a much larger proportion of the dry weight than does protein. *See* PLATYHELMINTHES.

Clark P. Read

Bibliography. A. C. Chandler and C. P. Read, *Introduction to Parasitology*, 10th ed., 1961; L. H. Hyman, *The Invertebrates*, vol. 2, 1951; E. R. Noble and G. A. Noble, *Parasitology: The Biology of Animal Parasites*, 5th ed., 1982; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; R. A. Wardle and J. A. McLeod, *The Zoology of Tapeworms*, 1952.

Cetacea

A mammalian order comprising approximately 90 living species of whales, dolphins, and porpoises and their fossil relatives. Like all other mammals but unlike all fish, cetaceans nurse their young with milk produced by the mother, are endothermic (warm-blooded), breathe air, have a lower jaw that consists of a single bony element (the dentary), and have three small bones (hammer, anvil, and stirrup) subserving sound transmission within the ear.

Unlike most mammals, living cetaceans are aquatic animals that cannot live on land. They have streamlined bodies (**Fig. 1**), with the nasal opening (blowhole) on top of the head. Their forelimbs are modified into flippers, they lack external hindlimbs, and their tail forms a flat horizontal fluke. Modern cetaceans lack hair except for whiskers in the young of some species. Many of these features are common in aquatic vertebrates, and they have evolved convergently as adaptations for life in water.

Brain and sense organs. The brain of cetaceans is large and highly developed. In fact, until humans came on the scene about 2 million years ago, cetaceans had the highest brain size to body size ratio of any mammal. Cetaceans are thought to be very intelligent. The cetacean sense of smell is nearly or totally absent. In most species the nerve that carries olfactory information to the brain is absent, which is very unusual among mammals. The eyes of most species are well developed, although some species of river dolphins are nearly blind. The ear is the most important sense organ. Toothed whales (Odontoceti) echolocate, emitting high-frequency sounds and using the echoes to determine shapes and distances in their surroundings. Odontocetes do not emit sounds with their voice box (larynx) like other mammals, but have modified nasal passages through which bursts of air are forced. These sounds

pass through the tissues of the head and reach a fatty organ on the forehead called the melon. The melon focuses sounds and allows a cetacean to direct its beam of echolocating calls. The sound-receiving organ of odontocetes is also highly specialized. The ear-opening of most cetaceans is minuscule; the head area that is most sound-sensitive in the dolphin is the back of the lower jaw. Sounds are transmitted from this area to the ear by means of a specialized pad of fat connecting the two structures. Mysticetes (baleen whales) do not echolocate; rather, they produce low-frequency sounds with their larynx. These sounds can travel through the ocean for hundreds of miles and are used for communication. *See* ECHOLOCATION.

Feeding. The two extant suborders of cetaceans, odontocetes and mysticetes, have different dietary specializations. Most odontocetes have simple, pronglike teeth which are used to grab and hold, but not chew, large prey items. Prey includes a variety of fish of all sizes, crustaceans, and squid and other mollusks. Some of the larger odontocetes, such as the killer whale (*Orcinus orca*), eat large prey, including sea lions and dolphins. A pod of killer whales will also hunt together, attacking much larger prey such as gray whales. Most odontocetes are active hunters that chase their prey, although some ambush their prey.

Modern mysticetes do not have teeth and are filter feeders, straining water filled with clouds of marine organisms (krill) through a network of baleen. Baleen is a keratinlike substance that hangs down in plates from the upper jaws of the whale. For example, humpback whales (*Megaptera novaeangliae*) take a mouthful of water and squeeze it with their tongue through the baleen sieve. The water passes through, but any suspended food items (such as plankton and small fish) are caught and subsequently swallowed. Mysticetes employ a variety of krill-concentrating behaviors, such as releasing bursts of air bubbles among schools of krill and swimming in formations to encircle the prey.

Locomotion. All modern cetaceans swim by swinging their horizontal tail fluke through the water, while their forelimbs are used for steering and navigating. The flippers of modern cetaceans resemble flat oars, although the bones for five fingers are present internally. The dorsal fin stabilizes the body during swimming and is not supported by a bony skeleton. The tail fluke is supported only by a column of vertebrae. The body of cetaceans is streamlined and contributes to efficient locomotion. The neck of modern cetaceans is short, and individual vertebrae are often fused. Under the skin of cetaceans is a layer of blubber, a fatty tissue that serves to insulate the animal and affects its buoyancy and streamlining. *See* ADIPOSE TISSUE.

Sustained swimming speeds of large whales (gray whale, *Eschrichtius robustus*; sperm whale, *Physeter macrocephalus*) are around 5 mi/h (8 km/h). During flight (such as when a sperm whale is harpooned), speeds of 25 mi/h (40 km/h) can be maintained for an hour. Feeding whales swim slowly, approximately 1–3 mi/h (2–5 km/h). Some species are capable of diving to great depths [commonly more than 5000 ft (1500 m) in the sperm whale], yet all cetaceans must come to the surface to breathe. Whales have a number of adaptations for diving and staying underwater for long periods of time. They exhale before they dive, allowing them to submerge faster. They store oxygen in the muscle (myoglobin) and not in the lungs or blood (hemoglobin) and change circulation patterns of blood to save oxygen. Their chests can easily collapse under increasing pressure (with depth) without causing permanent damage. After surfacing from a dive, whales remain at the surface for an extended period to replenish oxygen stores. *See* DIVING ANIMALS; HEMOGLOBIN.

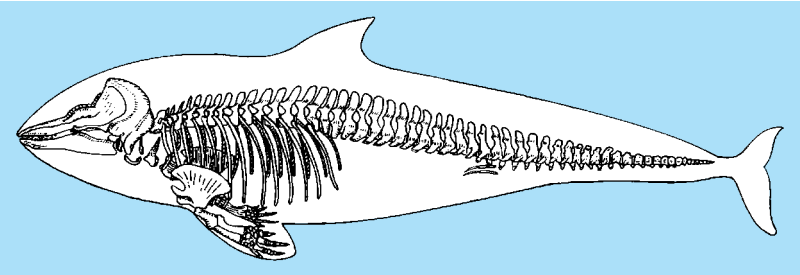

Fig. 1. Skeleton of a porpoise, highly specialized for an aquatic life. (*After Guide to the Hall of Biology of Mammals, Amer. Mus. Nat. Hist. Guide Leafl. Ser. 76, 1933*)

Life history and social structure. A variety of social structures are found among cetaceans. Most marine dolphins live in schools that may contain dozens of animals, sometimes composed of multiple species. These large aggregations may increase the ability to detect large predators, such as killer whales. They also engage in cooperative hunting. Dolphin societies are loose aggregations, males and females mate multiple times with different partners, and the two parents are often not part of the same group. Bottlenosed dolphin (*Tursiops truncatus*) females have an offspring every 3–6 years; the young stays with the mother for several years. Female offspring often remain with the maternal groups, whereas males wander. These dolphins may reach 50 years of age in the wild.

Herds of sperm whales consist of related females and juveniles. Clusters of young males form bachelor groups, and adult males, much larger than the females, are solitary. During the mating season, a male temporarily dominates one of the harems, but may be easily dislodged from this position by another male. Humpback whales (mysticetes) spend the summers feeding at high latitudes but live in temperate or tropical latitudes during the winter. Here, males form large clusters around single females and engage in elaborate courting behaviors, such as special swimming displays and singing. There is also much fighting among males. Right whales (Balaenidae, also mysticetes) ram into each other with their heads, which are studded with hard, sharp layers of barnacles that may cause bloody gashes in their opponent. Many baleen whales form mating pairs that last one summer. Copulation in humpback whales occurs belly to

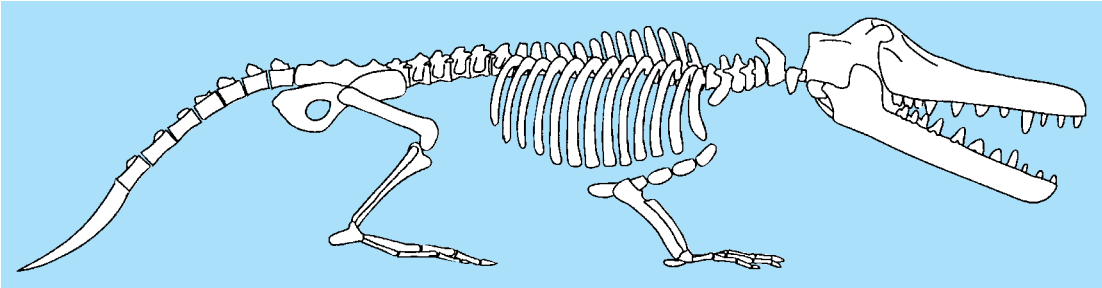

Fig. 2. Reconstruction of *Ambulocetus natans*. (*From J. G. M. Thewissen et al., Courier Forschungs-Institut Senckenberg, 1996*)

belly, while swimming vertically and surfacing. The gestation time is one year, and they may have an offspring every other year. Calving for most whales is done in shallow water; whale calves are born tail-first and are helped to the surface by the mother and other females. Humpback calves can swim in less than an hour after birth. Humpbacks are sexually mature after 8 years and live up to 50 years. Male baleen whales are usually smaller than females.

Distribution. Cetaceans are found in all oceans and seas. Some species are restricted to coastal environments (such as bottlenosed dolphins), whereas others live only in the open sea (such as sperm whales). Some species live in all seas and oceans of the world (such as killer whales). Many mysticetes and sperm whales are migratory. A number of dolphin species have left the sea and live permanently in rivers (such as the Indus, Ganges, Yellow, Amazon, and La Plata). Some river dolphins are among the most unusual modern cetaceans; they are excellent echolocators as demonstrated by their ability to navigate in muddy waters.

Systematics and evolution. Cetaceans originated from a four-footed terrestrial ancestor. This predecessor, a mesonychian, may have resembled a wolf or a hyena and lived approximately 55 million years ago. The closest living relatives of cetaceans are probably the artiodactyls (even-toed ungulates). The most primitive whales, the pakicetids, are known only from India and Pakistan; they were not sea animals, but probably lived in rivers and lakes and were no larger than a wolf. The most primitive marine whale is *Ambulocetus natans* (**Fig. 2**). A nearly complete skeleton is known for this species, and it resembles a crocodile with a long snout and eyes on top of the head. *Ambulocetus* had well-developed forelimbs and large hindlimbs; it did not have a tail fluke or flippers. Large whales first appeared around 40 million years ago; *Basilosaurus* had a long [maybe 50-ft (15-m)] snakelike body, and is known from Africa, North America, and Pakistan. Whales dating back 37 million years or more are often collectively called archaeocetes. They are primitive in that they still had teeth resembling those of land mammals, did

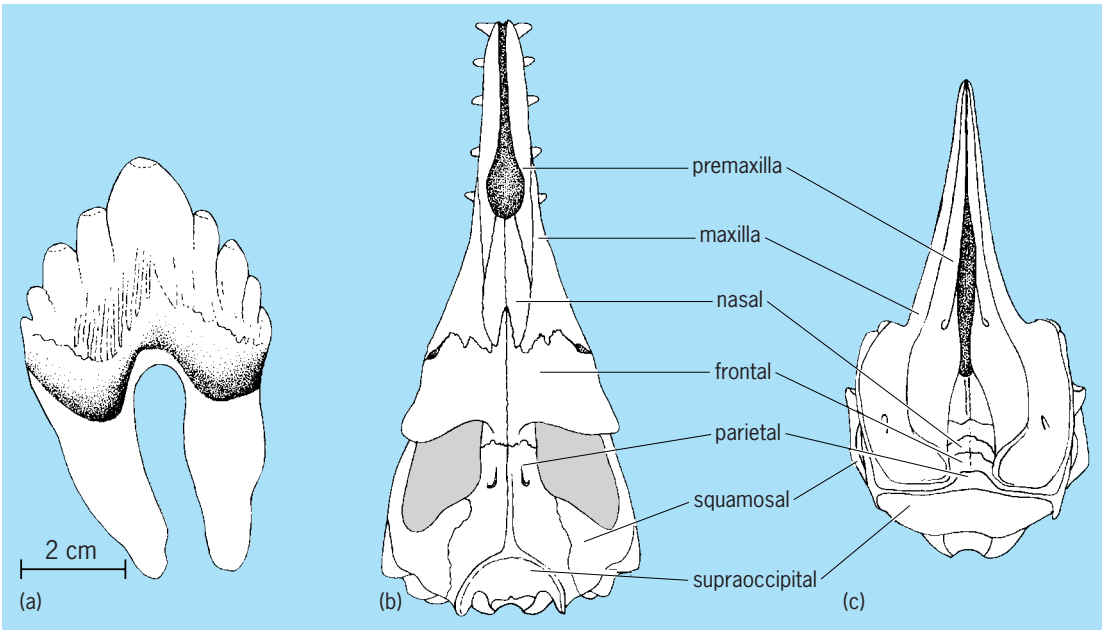

Fig. 3. Cetacean fossils. (*a*) Left upper premolar. (*b*) Skull of *Prozeuglodon stromeri* from early Oligocene of Egypt. (*c*) Skull of *Dephinodon dividum* from the middle Miocene of Maryland.

not echolocate, and had hindlimbs. Cetaceans more recent than 37 million years ago are included in the two modern suborders: Odontoceti (toothed whales, including dolphins and porpoises) and Mysticeti (baleen whales). *See* ARTIODACTYLA.

Primitive odontocetes have teeth that resemble those of sharks to some extent; some of these whales are called squalodontids (*Squalodon* is a genus of sharks; **Fig. 3**). Overall, many early odontocetes resembled dolphins. Even primitive odontocetes appear to have been echolocators. Modern odontocetes are diverse, ranging from the enormous sperm whales [up to 20 m (66 ft) long and 52,000 kg (114,500 lb)] to the tiny porpoises [Phocaenidae, smallest around 9 kg (20 lb), length 1.5 m (5 ft)]. Dolphins (Delphinidae, which includes the killer whale), porpoises (Phocaenidae), and fresh-water dolphins (Iniidae, Pontoporidae, Platanistidae) are the smallest odontocetes. Among the most unusual odontocetes is the extinct *Odobenocetops*, a whale with the face of a walrus, that probably dug for mollusks in the seas of Peru in the Pliocene. *Odobenocetops* had one tooth that was greatly enlarged and tusk-shaped. It is related to the extant narwals, which live in the Arctic. Male narwals have a straight tusk that is longer than the body. The oldest mysticetes (baleen whales) still had teeth, but these were lost early in their evolution. The largest animal ever to live on Earth is the blue whale (*Balaenoptera musculus*), a baleen whale that is 25 m (82 ft) long and weighs 130,000 km (286,340 lb). Other examples of mysticetes are humpback whales, right whales, gray whales, and minke whales. *See* MAMMALIA. J. G. M. Thewissen

Bibliography. R. F. Fordyce and L. G. Barnes, The evolutionary history of whales and dolphins, *Annu. Rev. Earth Planet. Sci.*, 22:419–455, 1994; S. M. Minasian, K. C. Balcomb, and L. Foster, *The World's Whales: The Complete Illustrated Guide*, Smithsonian Books, 1984.

Cetane number

A number, usually between 30 and 60, that indicates the ability of a diesel engine fuel to ignite quickly after being injected into the cylinder. The higher the cetane number, the more easily the fuel can be ignited. In high-speed diesel engines, a fuel with a long ignition delay tends to produce rough operation. *See* COMBUSTION CHAMBER.

Automotive diesel engines use grades 1-D and 2-D of diesel fuel. Both grades have a minimum cetane number of 40. Low- and medium-speed diesel engines use grade 4-D, which has a minimum cetane number of 30.

To determine cetane number of a fuel sample, a specially designed diesel engine is operated under specified conditions with the given fuel. The fuel is injected into the engine cylinder each cycle at 13° before top center. The compression ratio is adjusted until ignition takes place at top center (13° delay). Without changing the compression ratio, the engine is next operated on blends of cetane (*n*-hexadecane), a short-delay fuel, and heptamethylnonane, which has a long delay. When a blend is found that also has a 13° delay under these conditions, the cetane number of the fuel sample may be calculated from the quantity of cetane required in the blend. *See* DIESEL ENGINE; DIESEL FUEL. Donald Anglin

Bibliography. American Society for Testing and Materials, *Test Methods for Rating Motor, Diesel, and Aviation Fuels*, vol. 05.04, annually.

Cetomimiformes

An order of oceanic, mostly deep-water fishes that are structurally diverse and rare; most of the 41 species are known from one or a few specimens. Thus, their scientific study has been hindered, the anatomy is imperfectly known, and the relationships are in dispute. Five of the 10 families and 11 of the 21 genera currently placed in the order have been described since World War II. There is no fossil record.

Morphology and phylogeny. Cetomimiforms are all soft-rayed fishes. In most the mouth is large. Many are naked but a few have scales that are thin and deciduous or form an irregular mosaic; a few have the skin spinulose. Pelvic fins may be abdominal, thoracic, or jugular in position or, commonly, absent. In most forms the single dorsal and anal fins are placed rather well back and are opposed; an adipose fin is present in only one species. Studies now under way suggest that some kinds represent early developmental stages of other fishes, whereas some are the product of degenerative evolution in the deep seas. Cetomimiformes are probably polyphyletic and, when better known, the order may be dismembered and abandoned. The relationships of the constituent groups should be sought in the orders Beryciformes, Lampridiformes, and Salmoniformes.

Taxonomy. Because of the diversity, the five currently recognized suborders may be mentioned separately. Best known are the Cetomimoidei or whalefishes, also known as Cetunculi, a group of 3 families and 15 rare species of small, red or black deep-sea fishes with whale-shaped bodies and enormous mouths; they are bioluminescent. The Ateleopoidei (or Chondrobrachii) consist of 1 family, 3 genera, and 11 species of elongate fishes, the largest over half a meter, in which the long anal fin is continuous with the caudal and there is no dorsal fin. The Mirapinnatoidei (or Miripinnati) are tiny, perhaps larval fishes, all recently described. Three families, 4 genera, and 5 species are included. The Giganturoidei, with 2 families, 3 genera, and 6 species, are small mesopelagic fishes with large mouths and strong teeth; some have telescopic eyes. The Megalomycteroidei, or mosaic-scaled fishes, consist of 1 family, 4 genera, and 4 rare species of small, elongate deep-sea fishes with degenerate eyes and irregularly disposed scales. *See* ACTINOPTERYGII; OSTEICHTHYES; TELEOSTEI. Reeve M. Bailey

Chabazite

A mineral belonging to the zeolite family of silicates. It commonly occurs in well-crystallized groupings of crystals with external rhombohedral symmetry (point group $\bar{3}2/m$) with nearly cubic angles. This well-developed external form suggests an internal structure compatible with space group $R\,\bar{3}m$; however, one structure refinement suggests lower symmetry (triclinic) with space group $P\,\bar{1}$. In addition to recognizing the distinct rhombohedral habit of chabazite crystals, the following physical properties are significant: hardness (on the Mohs hardness scale) in the range 4–5; light colors ranging from white to yellow, pink, and red; vitreous luster; and transparent to translucent in transmitted light. It is distinguished from calcite by a poorer rhombohedral cleavage and lack of effervescent reaction with dilute hydrochloric acid (HCl). *See* CRYSTAL STRUCTURE.

The ideal composition is $Ca_2Al_2Si_4O_{12} \cdot 6H_2O$ (where Ca = calcium, Al = aluminum, Si = silicon, O = oxygen, H_2O = water), but there is considerable chemical substitution of Ca by sodium (Na) and potassium (K), as well as (Na,K)Si for CaAl. The internal structure of chabazite consists of a framework linkage of (AlO_4) and (SiO_4) tetrahedra, with large cagelike openings bounded by rings of tetrahedra. The cages are connected to each other by open structural channels that allow for the diffusion of molecules through the structure of a size comparable to that of the diameter of the channels (about 0.39 nanometer in diameter). For example, argon (0.384 nm in diameter) is quickly absorbed by the chabazite structure, but *iso* -butane (0.56 nm in diameter) cannot enter the structure. In this manner, chabazite can be used as a sieve on a molecular level. *See* MOLECULAR SIEVE.

Chabazite is found, commonly with other zeolites, lining cavities in basalts as a result of low-temperature hydrothermal activity. Most of the fine chabazite crystal specimens that are exhibited in museums were found in large crystal-lined vugs in basalts. Notable localities are the Faroe Islands; the Giant's Causeway, Ireland; Seiser Alpe, Trentino, Italy; Aussig, Bohemia, former Czechoslovakia; Oberstein, Germany; and India. In the United States chabazite has been found at West Paterson, New Jersey, and Goble Station, Oregon. *See* ZEOLITE.

Cornelis Klein

Bibliography. G. Gottardi and E. Galli, *Natural Zeolites*, 1985; C. Klein and C. S. Hurlbut, Jr., *Manual of Mineralogy*, 21st ed., 1993.

Chaetodermomorpha

Burrowing, vermiform (wormlike) mollusks in the group Aplacophora. Chaetodermomorpha (also known as Caudofoveata) are covered by a cuticular integument bearing sclerites (spicules) and are recognizable by the presence of a sensory cuticular oral shield, lack of a foot, and presence of paired gills in a posterior mantle cavity. Chaetoderms (see **illustration**) range in size from less than 0.08 in. (2 mm) to more than 2.8 in. (70 mm) and are found from shelf depths to hadal depths over 23,100 ft (7000 m). There are four families with 14 genera and 128 species worldwide. Chaetoderm species are numerically dominant in certain deep-sea localities. *See* DEEP-SEA FAUNA.

General morphology. Sclerites are aragonitic and sculptured by lengthwise ridges and grooves. They are overlapping, either held flat or slightly-to-perpendicularly raised in relation to the body. The underlying epidermis secretes both cuticle and sclerites, as well as functioning for excretion through specialized papillae. Beneath the epidermis lie circular muscles, two sets of diagonal muscles, and four sets of heavy longitudinal muscle bands, the placement of which defines body shape.

Burrowing is slow, brought about by pushing the anterior end through the sediment by means of hydrostatic pressure; the longitudinal muscles then contract to bring the rest of the body forward.

Digestion. The radula is cuticular and nonrasping with two mirror-image teeth per row on a divided radular membrane, or it is specialized either as a rasping organ with distichous teeth on a unipartite membrane or as grasping pincers with two small distal denticles on a large cone. Food is entirely organic, either detritus or prey organisms, such as foraminiferans. The digestive system consists of a pharynx supplied with salivary glands, sometimes an esophagus, a stomach from which opens a blind digestive gland, and an intestine that runs from the stomach to the mantle cavity. Some families have a gastric shield at the posterior end of the stomach against which turns a mucoid rod.

Respiration, circulation, and nervous system. The paired mantle cavity gills are typical for mollusks, of aspidobranch pattern with afferent and efferent vessels. A ventricle and paired auricles lie in the pericardium. The only blood vessel is the dorsal aorta, which runs anteriorly to the head; otherwise blood moves through open sinuses of the hemocoel. Paired ventral and lateral nerve cords arise from a bilobed cerebral ganglion and run posteriorly, connected by cross-commissures; the lateral cords join above the rectum as a large ganglion, which gives off a nerve to a chemoreceptory dorsal sensory organ.

Reproduction. The gonads are paired or fused and lie in the posterior part of the body; they empty into the pericardium, which in turn empties into paired U-shaped gametoducts. The latter are without elaboration and open separately into the cloaca. Sexes are separate. Embryological development is through nonfeeding larvae.

Phylogeny. The relationship of the Chaetodermomorpha to the other aplacophoran taxon Neomeniomorpha (Solenogastres) remains unresolved, as does their phylogenetic position in the Mollusca.

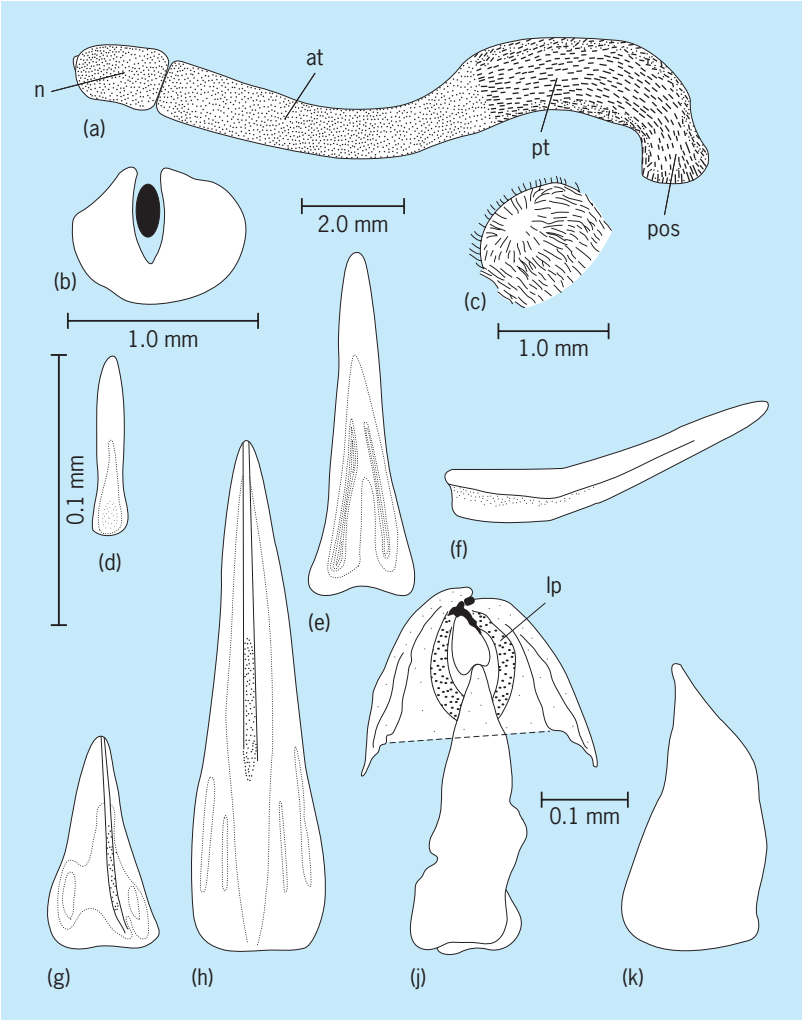

Chaetoderma argenteum **Heath, occurring from Alaska to southern California from 330 to 1980 ft (100 to 600 m) depth. (*a*) Specimen from Santa Maria Basin (USNM). (*b*) Oral shield of presumed syntype of *C. montereyensis* (MCZ). (*c*) Posterium of specimen *a*. (*d–h*) Spicules of specimen *b*, (*d*) from neck, (*e, f*) anterior trunk at constriction, (*g*) posterior region of anterior trunk, and (*h*) midposterior trunk. (*j, k*) Radula of specimen *a*, (*j*) frontal view and (*k*) lateral view of cone, frontal to right. (*From A. H. Scheltema in Taxonomic Atlas of the Benthic Fauna of the Santa Maria Basin and the Western Santa Barbara Channel, vol. 8, pt. a, The Mollusca, Santa Barbara Museum of Natural History, 1998*)**

They are unknown as fossils. *See* APLACOPHORA; MOLLUSCA; NEOMENIOMORPHA. Amelie H. Scheltema

Bibliography. L. v. Salvini-Plawen, Mollusca Caudofoveata, *Marine Invertebrates of Scandinavia*, vol. 4, 1975; A. H. Scheltema, Aplacophora as progenetic aculiferans and the coelomate origin of mollusks as the sister taxon of Sipuncula, *Biol. Bull. (Woods Hole)*, 184:57–78, 1993; A. H. Scheltema, Comparative morphology of the radulae and alimentary tracts in the Aplacophora, *Malacologia*, 20:361–383, 1981; A. H. Scheltema, Reproduction and rapid growth in a deep-sea aplacophoran mollusc, *Prochaetoderma yongei, Mar. Ecol., Prog. Ser.*, 37:171–180, 1987; A. H. Scheltema, in *Taxonomic Atlas of the Benthic Fauna of the Santa Maria Basin and the Western Santa Barbara Channel*, vol. 8, *Mollusca*, pt. 1, Santa Barbara Museum of Natural History, 1998; A. H. Scheltema and D. L. Ivanov, Prochaetodermatidae of the eastern Atlantic Ocean and Mediterranean Sea, *J. Moll. Stud.*, 66:313–362, 2000; A. H. Scheltema and D. L. Ivanov, Use of birefringence to characterize Aplacophora sclerites, *The Veliger*, 47:153–160, 2004; E. R. Trueman and M. R. Clarke (eds.), *The Mollusca*, vol. 10: *Evolution*, 1985.

Chaetognatha

A phylum of abundant planktonic arrowworms. As adults they range in size from 0.2 to 4 in. (5 to 100 mm). Their bodies are tubular and transparent, and divided into three portions: head, trunk, and tail (**illus.** *a*). The head possesses one or two rows of minute teeth anterior to the mouth and usually 7–10 larger chaetae, or seizing jaws, on each side of the head. One or two pairs of lateral fins and a caudal fin are present. In mature individuals a pair of seminal vesicles protrudes from the sides of the tail segment just forward of the caudal fin.

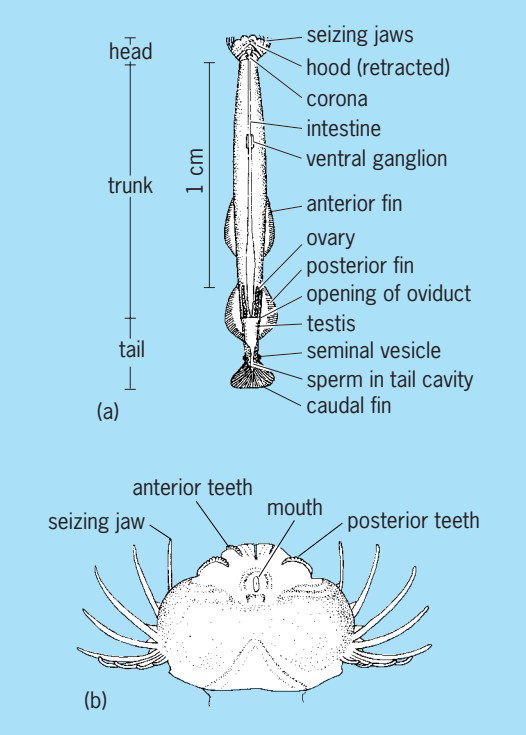

***Sagitta enflata*. (*a*) Dorsal view. (*b*) Head, ventral.**

Nine genera and about 42 species are recognized by some specialists. Most species belong to the genus *Sagitta*, which can be recognized by the presence of two pairs of teeth (illus. *b*) and two pairs of lateral fins. Important diagnostic features for species include the shape and position of the seminal vesicles and the number and arrangement of the anterior teeth.

Morphology. The body is covered by a thin hood of epidermis. When drawn forward, the hood covers the teeth and jaws and streamlines the anterior end of the body. Two small, dark eyes are found on the dorsal side of the head. Behind the eyes, and forming an irregular loop on the epidermis, is a structure called a corona ciliata, which is visible only when stained.

The nervous system consists of a dorsal ganglion, or brain, from which nerves lead to various portions of the head and body. A pair of nerves extends posteriorly from the brain and joins the prominent ventral ganglion which lies just beneath the epidermis.

The head musculature is quite complex, but the trunk has only six longitudinal bands of muscle, which restrict the body to simple lateral movements. Arrowworms are unable to constrict their bodies. Swimming is accomplished by rapid but brief movements of the caudal fin, followed by a quiet period. They do not appear to swim much, but drift passively most of the time.

The digestive system is simple. A mouth opens ventrally in the center of the head. A short esophagus leads to a straight intestine, which begins just behind the head and terminates in an anus opening ventrally at the septum separating the trunk and tail segments. The intestine is supported by dorsal and ventral mesenteries. A large coelom surrounds the intestine.

All chaetognaths are hermaphroditic. The ovaries are tubular and lie in the trunk segment just ahead of the septum dividing the trunk from the tail. The eggs in the ovaries increase in size at maturity, only a small fraction of the eggs maturing and being spawned at one time. Alongside the ovary is an oviduct, and contained within the oviduct is a narrow sperm duct. Fertilization is internal. The fertilized eggs are spawned free in the ocean, and the young chaetognaths hatch in some instances within 48 h. There is no larval stage.

The testes are situated in either side of the tail segment. As the sperm mature, they break away and often circulate in the tail cavity. From time to time some pass through short ducts leading from inside the tail cavity to the seminal vesicles. Small quantities of sperm are stored in the ducts prior to discharge.

Natural history. Chaetognaths are cosmopolitan forms which live not only at the surface but also at great depths; however, no one species is found in all latitudes and at all depths. One of the Arctic species, *Eukrohnia hamata*, may extend to the Antarctic by way of deep water across the tropics. A few species are neritic and are not found normally beyond the continental shelf.

Their food consists principally of copepods and other small planktonic crustaceans; however, they are very predacious and will even eat small fish larvae and other chaetognaths on rare occasions. The suggestion has been made that they feed on plants as well as animals, but this needs more supporting evidence. They seize their food with their pointed jaws and, aided by the teeth, push it into their mouths. Small crustacean larvae eaten one day appear to be digested by the next.

Along the Atlantic Coast of the United States they are most abundant over the continental shelf. Numbers of five or more per cubic meter of water are common along the coast; farther offshore only a fraction of this number is normally encountered. When very abundant, they may exceed 2.8/ft^3 (100/m^3); however, this is unusual and often includes many young specimens.

Studies have shown them to be useful as indicator organisms. Certain species appear to be associated with characteristic types or masses of water, and when this water is displaced into an adjacent water mass, the chaetognaths may be used as temporary evidence for such displacement.

Their abundance and cosmopolitan distribution in marine waters, their little-known life history, their importance in the food chain both as predators and prey, and their value as indicator organisms invite further study. *See* SEAWATER FERTILITY. E. Lowe Pierce

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; E. L. Pierce, The Chaetognatha over the continental shelf of North Carolina with attention to their relation to the hydrography of the area, *J. Mar. Res.*, 12(1):75–92, 1953; T. Tokioka, Chaetognaths of the

Indo-Pacific, *Anat. Zool. Jap.*, 25(1,2):307–316, 1952; R. von Ritter-Zahony, *Revision der Chaetognathen*, Deutsche Südpolar-Expedition 1901–1903, Bd. 13, Zool. 5, Heft 1, 1911.

Chaetonotida

An order of the phylum Gastrotricha. Members have Y-shaped pharyngeal lumina. *Neodasys* is a marine and macrodasyid-like form that reaches 0.8 mm (0.03 in.) in length and has front, side, and rear adhesive tubes. Others seldom exceed 0.3 mm (0.01 in.); they have only two rear adhesive tubes borne on a posterior furca, or none at all. Members of the family Xenotrichulidae have locomotor cilia grouped into cirri; all three genera are marine (**illus.** *a*).

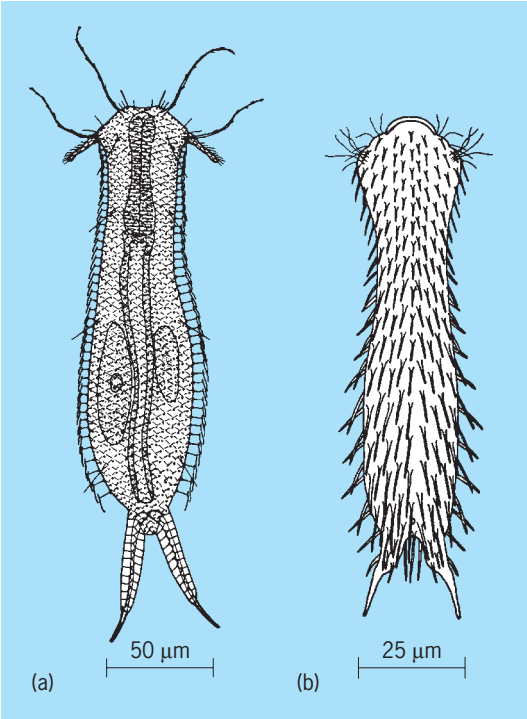

Marine gastrotrichs of the order Chaetonotida; dorsal views. (*a*) *Xenotrichula* (*courtesy of W. D. Hummon*). (*b*) *Chaetonotus* (*after W. D. Hummon, Intertidal marine Gastrotricha from Colombia, Bull. Mar. Sci., 24:396–408, 1974*).

The family Chaetonotidae comprises half of all gastrotrichs: *Musellifer* and *Halichaetonotus* are marine; *Polymerurus* is fresh-water; *Aspidiophorus*, *Chaetonotus* (illus. *b*), *Heterolepidoderma*, *Ichthydium*, and *Lepidodermella* have species in each habitat. *Musellifer* lives in mud; all others inhabit sands of streams, beaches, or offshore banks, or live in the surface detritus of ponds or lake bottoms. The other seven genera, in four families, are all fresh-water and either are rare or tend toward a semiplanktonic life, especially *Neogossea* and *Stylochaeta*. Several chaetonotids are easily kept in aquatic microcosms, and a few, such as *L. squamata*, are readily cultured in pond water on wheat grains. *See* GASTROTRICHA.

William D. Hummon

Chain drive

A flexible device of connected links used to transmit power. A drive consists of an endless chain which meshes with sprockets located on the shaft of a driving source, such as an electric motor/reducer, and a driven source, such as the head shaft of a belt conveyor.

Chains have been used for more than 2000 years. However, the modern contribution of the chain to industrialization began in 1873 with the development of a cast detachable chain. It was a simple cast-metal chain composed of identical links which could be coupled together by hand. This chain so greatly improved the performance of power takeoff drives for farm implements that mechanization became a practical reality. The first development after the success of the cast detachable chain was the cast pintle chain with a closed barrel design and steel pins. Other cast chain designs and variations evolved until an all-steel chain was developed.

Chain types. The new development, now called the roller chain (**Fig. 1**), found uses in the early 1900s on bicycles and other forms of conveyances. Constantly refined and improved, today's roller chain meets the demands of heavy-duty oil well drilling equipment, high-production agricultural machinery (**Fig. 2**), construction machinery, and similar

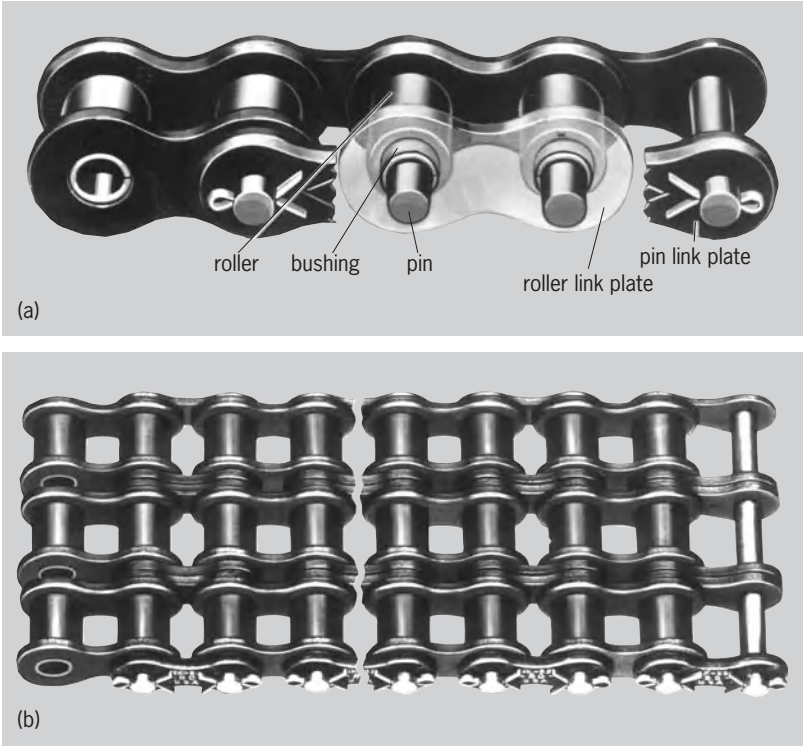

Fig. 1. Roller chains. (*a*) Single-strand. (*b*) Triple-strand.

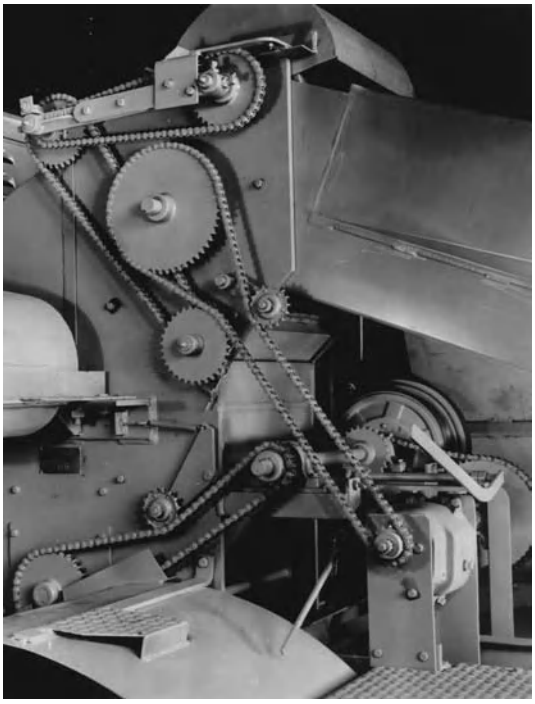

Fig. 2. Roller chain drive. (*FMC Corp.*)

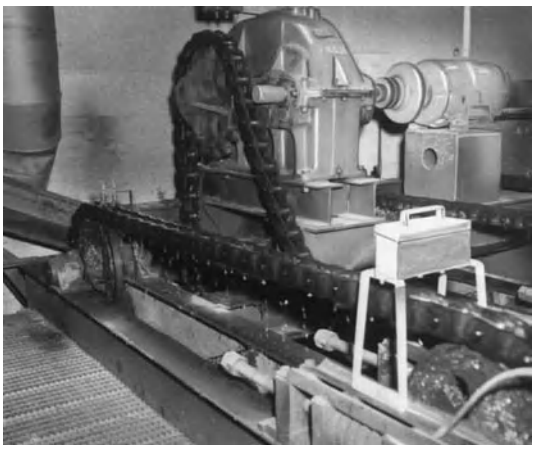

Fig. 3. Engineering steel chain drive. (*Rexnord Inc.*)

equipment. It also meets the precise timing requirements of lighter-duty equipment such as printing, packaging, and vending equipment.

Another type of chain also evolved, called the engineering steel chain (**Fig. 3**). In a broad sense, the early designs of the engineering steel chain were a blend of the other two, the cast and the roller chain. The drive chains in the engineering steel chain category are usually identified by their offset/crank-link sidebar design. Generally, larger pitch sizes as compared to the roller chain and higher-strength chains characterize this group. A third group used for chain drives is the inverted-tooth (silent) chains (**Fig. 4**). A familiar application is their use as automotive timing chains in automobile engines.

Advantage. The use of chains for power transmission rather than another device, such as V-belts or a direct coupling to the power source, is usually based on the cost effectiveness and economy of chains and sprockets. Chains and sprockets offer the following advantages: large speed ratios; sufficient elasticity to absorb reasonable shocks; a constant speed ratio between the driving and driven shaft; long life without excessive maintenance; mechanical understandability regarding installation and functionality; coupling and uncoupling with simple tools; and a simple means to get power from its source to the location where needed. *See* BELT DRIVE.

Drive design. The design of a chain drive consists primarily of the selection of the chain and sprocket sizes. It also includes the determination of chain length, center distance, method of lubrication, and in some cases the arrangement of chain casings and idlers. Chain and sprocket selection is based on the horsepower and type of drive; the speeds and sizes of the shafting; and the surrounding conditions. A properly selected chain, following prescribed chain manufacturers' and the American Chain Association's techniques, is usually based on 15,000 hours of operation without breakage of components, considering chain wear not to exceed that which can be accommodated by the sprockets. Generally, chains are considered worn out when the roller chain exceeds approximately 3% elongation, the engineered steel chain exceeds approximately 5% elongation, and the inverted-tooth (silent) chain exhibits malfunction characteristics. To achieve the rated 15,000-h life, the environment must be clean and the chain lubricated as recommended by the manufacturer for the speed, the horsepower capacity, and the number of teeth on the smaller sprocket.

Chain and sprocket design. The particular design of the sideplate/sidebar/crank-link, pin, bushing, and roller of chains used in chain drives has been standardized to a substantial degree. The dimensional parameters regarding pitch, pin, bushing, roller, and side-plate sizes have been standardized. Interference

Fig. 4. Inverted tooth/silent chain drive.

ANSI and ISO standards for drive chains

Chain description	ANSI standard number	ISO standard number	Title
Roller chain	B29.1	R606	(ANSI) Transmission Roller Chains and Sprocket Teeth (ISO) Short-pitch Transmission Roller Chain
Inverted-tooth (silent) chain	B29.2	—	(ANSI) Inverted-tooth (silent) Chains and Sprocket Teeth (ISO) No standard available
Engineering steel chain	B29.10	ISO3512	(ANSI) Heavy-duty Offset Sidebar Power Transmission Roller Chains and Sprocket Teeth (ISO) Heavy-duty Crank-Link Transmission Chains

between parts is usually controlled by the manufacturer, as is hardness of the chain parts, which determine the life of the chain. The selection of materials and heat treatments to obtain desired hardnesses are selected by manufacturers, but generally certain minimum hardnesses are required to meet minimum ultimate strength or breaking load criteria.

A sprocket is a wheel with teeth shaped to mesh with the chain. The sprocket tooth form, when properly selected, assures the success of the chain drive. Space limitations often determine the chain length and the number of teeth on the sprocket. Usually more than one combination of chain type, chain size, and number of sprocket teeth will satisfy a requirement. The final determination is based on economics and availability.

Lubrication. Lubrication of chains reduces joint wear as the links flex onto and off the sprockets. Lubricated bearing surfaces, that is, the chain joints, can carry high loads without galling. Chain capacity or horsepower ratings are determined for operation in a clean environment with proper lubrication. Lubrication extends the wear life of chains and sprockets operating in any environment, no matter how dirty or abrasive. Chains should not be greased. A nondetergent petroleum-base oil is recommended. For operation at high speeds, oil stream–force feed lubrication is required for cooling. An oil bath is effective at intermediate speeds, and manual lubrication is acceptable at slow speeds. *See* LUBRICANT; WEAR.

ANSI and ISO standards. The American National Standards Institute has established a committee for the standardization of transmission chains and sprockets. Also, the International Standards Organization has established a committee to promulgate international standards for chains and chain wheels for power transmission and conveyors. The **table** identifies those standards which apply to drive chains. Typically, the standards identify those characteristics of a chain which ensure that one manufacturer's chain will couple with another and that the minimum ultimate strength or breaking load characteristics have been established for each chain. The ANSI standard includes a supplemental section with horsepower rating tables and selection information. *See* CONVEYOR.

Victor D. Petershack

Bibliography. American Chain Association, *Chains for Power Transmission and Material Handling*, 1982.

Chain reaction (chemistry)

A chemical reaction in which many molecules undergo chemical reaction after one molecule becomes activated. In ordinary chemical reactions, every molecule that reacts must first become activated by collision with other rapidly moving molecules. The number of these violent collisions per second is so small that the reaction is slow. Once a chain reaction is started, it is not necessary to wait for more collisions with activated molecules to accelerate the reaction, which now proceeds spontaneously.

Photochemical reactions. A typical chain reaction is the photochemical reaction between hydrogen and chlorine as described by reactions (1).

$$\begin{aligned} Cl_2 + \text{light} &\rightarrow Cl + Cl \\ Cl + H_2 &\rightarrow HCl + H \\ H + Cl_2 &\rightarrow HCl + Cl \\ Cl + H_2 &\rightarrow HCl + H \end{aligned} \tag{1}$$

The light absorbed by a chlorine molecule dissociates the molecule into chlorine atoms; these in turn react rapidly with hydrogen molecules to give hydrogen chloride and hydrogen atoms. The hydrogen atoms react with chlorine molecules to give hydrogen chloride and chlorine atoms. The chlorine atoms react further with hydrogen and continue the chain until some other reaction uses up the free atoms of chlorine or hydrogen. The chain-stopping reaction may be the reaction between two chlorine atoms to give chlorine molecules, or between two hydrogen atoms to give hydrogen molecules. Again the atoms may collide with the walls of the containing vessel, or they may react with some impurity which is present in the vessel only as a trace.

The length of the chain, that is, the number of molecules reacting per molecule activated, is determined by the relative rates of the competing reactions, namely, the chain-propagating reaction and the chain-stopping reactions. In the chain reaction just described, 10^6 molecules of hydrogen chloride may be formed by the photodissociation of 1 chlorine molecule.

In photochemical chain reactions, the length of the chain can be determined by measuring the number of photons of light absorbed, that is, the number of molecules activated, and dividing by the number of molecules which react chemically. *See* PHOTOCHEMISTRY.

Thermal reactions. In thermal reactions, the length of the chain may sometimes be estimated from a knowledge of the intermediate steps and the kinetics involved. The presence of a chain reaction can often be proved by adding a trace of an inhibitor, such as nitric oxide. If the reaction is slowed down greatly by a very small amount of a substance which reacts with the chain-propagating units, the reaction involves a chain. While the inhibitor is being consumed in this way, the reaction is slow. After an induction period, the inhibitor is consumed and the rapid chain reaction then takes place.

Chain reactions are erratic and are reproduced with difficulty in different laboratories because they depend so much on the presence and concentration of accidental impurities which act as inhibitors.

In many chemical reactions, particularly organic reactions at elevated temperatures, the chains are carried by free radicals which are very reactive fragments of molecules that have unshared electrons, such as $\cdot CH_3$, $\cdot C_2H_5$, $\cdot H$, and $\cdot OH$. The thermal decomposition of propane is a typical free-radical chain which follows reactions (2). One molecule

$$\begin{aligned} C_3H_8 &\rightarrow \cdot CH_3 + \cdot C_2H_5 \\ \cdot CH_3 + C_3H_8 &\rightarrow CH_4 + C_3H_7 \\ \cdot C_3H_7 &\rightarrow \cdot CH_3 + C_2H_4 \\ \cdot CH_3 + C_3H_8 &\rightarrow CH_4 + \cdot C_3H_7 \end{aligned} \tag{2}$$

of propane is decomposed into free radicals, $\cdot CH_3$ and $\cdot C_2H_5$, which then react with more propane to give the product methane and a free radical, $\cdot C_3H_7$, which decomposes into $\cdot CH_3$ and the product ethylene. The $\cdot CH_3$ reacts with more propane and continues the chain. The chain is terminated by collision of the free radicals with the wall or with each other, in reactions such as (3). Thus it is possible

$$\cdot CH_3 + \cdot C_3H_7 \rightarrow C_4H_{10} \tag{3}$$

to obtain products of higher molecular weight as well as products of lower molecular weight. The finding of these higher-molecular-weight products supports the theory of free-radical formation and chain reactions.

Certain oxidations in the gas phase are known to be chain reactions. The carbon knock which occurs at times in internal combustion engines is caused by a too-rapid combustion rate caused by chain reactions. This chain reaction is reduced by adding tetraethyllead which acts as an inhibitor.

The polymerization of styrene to give polystyrene and the polymerization of other organic materials to give industrial plastics involve chain reactions. The spoilage of foods, the precipitation of insoluble gums in gasoline, and the deterioration of certain plastics in sunlight involve chain reactions, which can be minimized with inhibitors. *See* ANTIOXIDANT; CATALYSIS; CHAIN REACTION (PHYSICS); CHEMICAL DYNAMICS; INHIBITOR (CHEMISTRY). Farrington Daniels

Bibliography. J. H. Espenson, *Chemical Kinetics and Reaction Mechanisms*, 2d ed., 1995; H. Eyring et al., *Basic Chemical Kinetics*, 1980; K. J. Laidler, *Reaction Kinetics*, 2 vols., 1963; J. I. Steinfeld, J. Francisco, and W. Hase, *Chemical Kinetics and Dynamics*, 2d ed., 1998.

Chain reaction (physics)

A succession of generation after generation of acts of division (called fission) of certain heavy nuclei. The fission process releases about 200 MeV (3.2×10^{-4} erg $= 3.2 \times 10^{-11}$ joule) in the form of energetic particles including two or three neutrons. Some of the neutrons from one generation are captured by fissile species (^{233}U, ^{235}U, ^{239}Pu) to cause the fissions of the next generations. The process is employed in nuclear reactors and nuclear explosive devices. *See* NUCLEAR FISSION.

The ratio of the number of fissions in one generation to the number in the previous generation is the multiplication factor k. The value of k can range from less than 1 to less than 2, and depends upon the type and amount of fissile material, the rate of neutron absorption in nonfissile material, the rate at which neutrons leak out of the system, and the average energy of the neutrons in the system. When $k = 1$, the fission rate remains constant and the system is said to be critical. When $k > 1$, the system is supercritical and the fission rate increases. *See* REACTOR PHYSICS.

A typical water-cooled power reactor contains an array of uranium rods (about 3% ^{235}U) surrounded by water. The uranium in the form of UO_2 is sealed into zirconium alloy tubes. The water removes the heat and also slows down (moderates) the neutrons by elastic collision with hydrogen nuclei. The slow neutrons have a much higher probability of causing fission in ^{235}U than faster (more energetic) neutrons do. In a fast reactor, no light nuclei are present in the system and the average neutron velocity is much higher. In such systems it is possible to use the excess neutrons to convert ^{238}U to ^{239}U. Then ^{239}U undergoes radioactive decay into ^{239}Pu, which is a fissile material capable of sustaining the chain reaction. If more than one ^{239}Pu atom is provided for each ^{235}U consumed, the system is said to breed (that is, make more fissile fuel than it consumes). In the breeder reactor, the isotope ^{238}U (which makes up 99.3% of natural uranium) becomes the fuel. This increases the energy yield from uranium deposits by more than a factor of 60 over a typical water-moderated reactor, which mostly employs the isotope ^{235}U as fuel.

A majority of power reactors use water as both the moderator and the coolant. However, a limited number of reactors use heavy water instead of light water. The advantage of this system is that it is possible to use natural uranium as a fuel so that no uranium enrichment is needed. Some other power reactors are gas-cooled by either helium or carbon dioxide and are moderated with graphite. *See* NUCLEAR REACTOR.

Norman C. Rasmussen

Bibliography. H. A. Enge and R. P. Redwine, *Introduction to Nuclear Physics*, 2d ed., 1995; S. Glasstone and A. Sesonske, *Nuclear Reactor*

Engineering, 2 vols., 4th ed., 1993; K. S. Ram, *Basic Nuclear Engineering*, 1990.

Chalcanthite

A mineral with the chemical composition $CuSO_4 \cdot 5H_2O$. Chalcanthite commonly occurs in blue to greenish-blue triclinic crystals or in massive fibrous veins or stalactites (see **illus.**). Fracture is conchoidal and luster is vitreous. Hardness is 2.5 on Mohs scale and specific gravity is 2.28. It has a nauseating taste and is readily soluble in water. It dehydrates in dry air to a greenish-white powder.

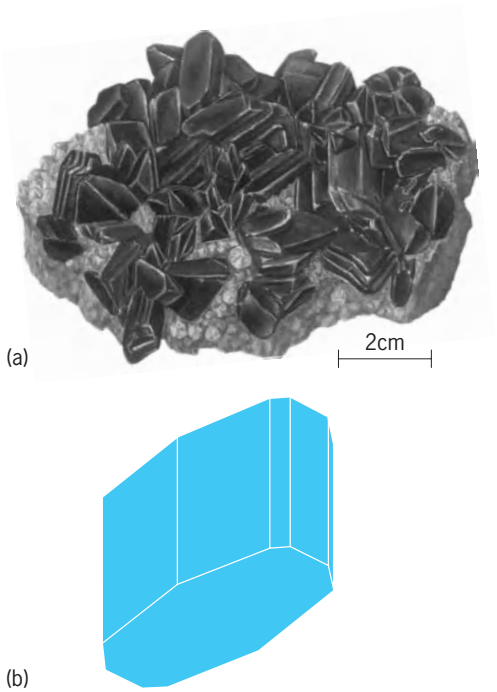

Chalcanthite. (*a*) Crystals associated with quartz, Clifton, Arizona (*specimen from Department of Geology, Bryn Mawr College*). (*b*) Crystal habit (*after L. G. Berry and B. Mason, Mineralogy, W. H. Freeman, 1959*).

Chalcanthite is a secondary mineral associated with gypsum, melanterite, brochantite, and other sulfate minerals found in copper or iron sulfide deposits. It is also found in mine workings.

Although deposits of commercial size occur in arid areas, chalcanthite is generally not an important source of copper ore. Its occurrence is widespread in the western United States. *See* COPPER.

Edward C. T. Chao

Chalcedony

A fine-grained fibrous variety of quartz, silicon dioxide. The individual fibers that compose the mineral aggregate usually are visible only under the microscope. Subvarieties of chalcedony recognized on the basis of color differences (induced by impurities), some valued since ancient times as semiprecious gem materials, include carnelian (translucent, deep flesh red to clear red in color), sard (orange-brown to reddish brown), and chrysoprase (apple green). Chalcedony sometimes contains dendritic enclosures resembling plants or trees. Major kinds of impurities that give color to chalcedony are iron oxides (carnelian and sard), nickel (chrysoprase), and manganese. *See* GEM; QUARTZ.

Chalcedony occurs as crusts with a rounded, mammillary, or botryoidal surface and as a major constituent of nodular and bedded cherts. The hardness is 6.5–7 on Mohs scale. The specific gravity is 2.57–2.64. The ultrafine structure of chalcedony has been deduced from x-ray diffraction and electron microscopy to consist of a network of microcrystalline quartz with many micropores. The amount of amorphous silica, if any, is less than 10%. The yellowish color and anomalously low indices of refraction commonly observed under a transmitted light microscope result from scattering of light from the micropores. Paleozoic and older chalcedony is usually more coarsely crystallized than younger examples, grain growth of microcrystalline quartz being a result of time.

Crusts of chalcedony generally are composed of fairly distinct layers concentric to the surface. Agate is a common and important type of chalcedony in which successive layers differ markedly in color and degree of translucency. In the most common kind of agate the layers are curved and concentric to the shape of the cavity in which the material formed. The successive layers of chalcedony and agate usually differ in permeability to solutions, and much colored agate sold commercially is artificially pigmented by dyes or inorganic chemical compounds. *See* AGATE.

Raymond Siever

Bibliography. C. Frondel, *Dana's System of Mineralogy*, vol. 3: *Silica Minerals*, 1962.

Chalk

The term "chalk" is sometimes used in a broad sense for any soft, friable, or weathered fine-grained limestone; however the term is mostly restricted to pelagic (biogenic) limestones. Chalk is a uniformly fine-grained, typically light-colored marine limestone primarily composed of the remains of calcareous nannofossils and microfossils. These minute pelagic organisms (see **illus.**) live in surface and near-surface oceanic waters and include coccolithophores (algae) and planktic foraminifers (Protozoa). Larger fossil constituents (such as bivalves, pteropods, echinoids, or ammonites) may be present, but only in subordinate amounts. The dominant pelagic skeletal remains are composed of low-magnesium calcite and, after death, settle slowly to the ocean floor, accumulating where the sea floor lies at a depth of less than about 4 km or 13,000 ft (the carbonate is redissolved at greater depths). Typical chalk sedimentation rates are 30 m (100 ft) per million years, so chalk accumulation is also dependent on the exclusion of diluting materials such as reefal detritus or terrigenous debris (clay, silt, or sand) transported from land areas

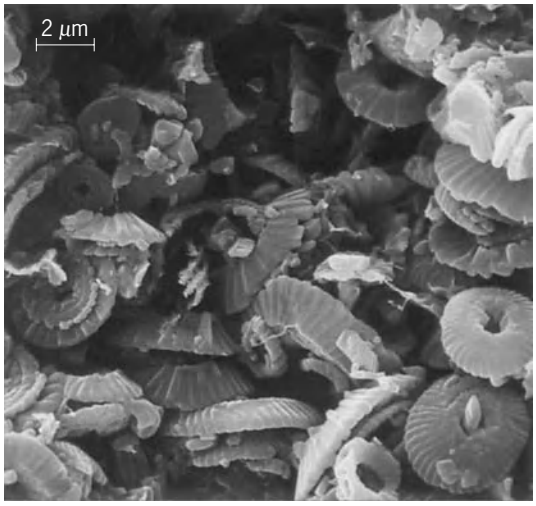

Typical chalk coccoliths, with small grain size and porous fabric. The sample is from late Oligocene chalk from the east side of Halton-Rockall Basin, North Atlantic Ocean.

by rivers. Chalks therefore form mainly in isolated outer shelf or deeper-water settings that are far from land areas. *See* CALCITE; LIMESTONE.

Because coccolithophores and planktic foraminifers did not exist prior to the Jurassic Period (about 150 million years ago), all true chalks are Jurassic or younger. Only 50 million years later (by Late Cretaceous time), chalks had become the single most volumetrically important type of limestone, a status they still retain. Most of those chalks are confined to ocean basins and so are poorly studied. The best-known chalks, such as those from the coastal cliffs of southern England (including the white cliffs of Dover), the Gulf Coast of the United States (Austin or Selma Limestones), and the U.S. Western Interior (Niobrara and Greenhorn Limestones), were deposited outside ocean basins on broad continental shelves during a worldwide sea-level high stand in Late Cretaceous time. *See* CRETACEOUS.

The small size of coccoliths and their constituent crystals means that chalks are extremely fine-grained deposits, typically with average crystal sizes of 1 micrometer or less. Because low-magnesium calcite is the least soluble form of calcium carbonate in near-surface settings, true nannofossil chalks also normally remain porous and friable. The unique combination of light color, compositional purity, softness, and fine texture led to many of the early uses of chalk for writing on blackboards or for marking athletic fields. Chalks are also widely used in the manufacture of portland cement, as lime for fertilizers, and in powders, abrasives, and coatings. More recently, chalks have become major targets for hydrocarbon exploration, especially in the Norwegian and Danish sectors of the North Sea, the southern Persian Gulf region, and U.S. Gulf Coast. Chalks lose porosity mainly through the effects of pressure-related mechanical and chemical dissolution and reprecipitation of calcium carbonate during progressive burial. Where allowed to continue unchecked, such alteration eventually leads to nearly complete porosity loss by about 3.8 km (12,000 ft) of burial. The presence of anomalously high pore-fluid pressures in basins such as the North Sea or the early introduction of hydrocarbons can inhibit or prevent such pressure-related alteration and allow preservation of prolific oil fields in chalk reservoirs, even at depths in excess of 3 km (10,000 ft).

P. A. Scholle

Bibliography. American Geophysical Union, *Short Course in Geology*, no. 4; R. S. Boynton, *Chemistry and Technology of Lime and Limestone*, 2d ed., John Wiley, 1980; H. G. Reading (ed.), *Sedimentary Environments and Facies*, 3d ed., 1994; J. P. Riley and R. Chester (eds.), *Treatise on Chemical Oceanography*, vol. 5, pp. 265–388, 1976; P. A. Scholle, Chalk diagenesis and its relation to petroleum exploration: Oil from chalks, a modern miracle?, *Amer. Ass. Petrol. Geol. Bull.*, 61(7):982–1009, 1977; P. A. Scholle, D. G. Bebout, and C. H. Moore (eds.), Carbonate depositional environments, *Amer. Ass. Petrol. Geo. Mem.*, no. 33, 1983; P. A. Scholle, N. P. James, and J. F. Read (eds.), *Carbonate Sedimentation and Petrology*, 1989.

Chameleon

The name for about 80 species of small-to-medium-sized lizards that comprise the family Chamaeleontidae and occur mainly in Africa and Madagascar. The American chameleons (*Anolis*) belong to a different family of lizards, the Iguanidae. This reptile is insectivorous and beneficial to humans. It commonly is found on vegetation, especially the green parts.

Chamaeleo chamaeleon is the most common species and is a typical example of the group. Its body is flattened from side to side; it has a long, prehensile tail; and both the forelimbs and hindlimbs have two digits that oppose the other three (see **illus.**). These feet and the tail make the chameleon well adapted for its arboreal habitat. The eyes are large and can move independently of each other in all directions. The tongue is also prehensile, being extensible for a great distance, about the length of the animal itself, and is a highly efficient organ for capturing insects. The head of the chameleon is triangular in profile and has a pointed crest.

Chameleons are noted for their ability to change color. Color changes appear to be related to

***Calumma brevicornis*; short-horned chameleon. (*Photo courtesy of Gerald and Buff Corsi, © California Academy of Science*)**

environmental temperatures as well as other external stimuli. When the temperature is high or the animal excited, the chameleon is green in color. When the temperature is cool or the animal unmolested, it is gray in color. There are cells in the skin that contain chromatophores, pigment granules of various colors. These black, red, or yellow pigments are contained in stellate, or star-shaped, color cells. Under nervous stimulation the cell branches expand and the pigments diffuse into these areas and become visible, creating a change in skin color. Contraction of the cells causes the pigment to concentrate and appear as indefinite spots of color. The ability to change color is protective, so that the animal can blend into its surroundings.

The chameleon is oviparous. The female digs out a hollow in the ground for a nest where several dozen eggs are laid and then covered with soil. The period of incubation varies inversely with the temperature and may be as short as 4 months or as long as 10 months. Parental care of the young has not been observed.

The armored chameleon (*Leandria perarmata*), known from one specimen in Madagascar, is a highly developed representative of a group of chameleons having one or more horns on the snout instead of a crest on the head. The body has a row of spines along either side as well as a crest of spines along the back. There is a bony hood on the nape of the neck and a serrated crest down to the eye level. One species, *C. bitaeniatus*, found in the mountainous areas of eastern Africa, is ovoviviparous, and the female produces eggs that are incubated and hatched within her body. After a period of development about six young are born. *See* CHROMATOPHORE; PROTECTIVE COLORATION; REPTILIA; SQUAMATA. Charles B. Curtin

Chamois

One of several species of mammals included in the tribe Rupicaprini of the family Bovidae. The group is heterogeneous in form, but all are intermediate in characteristics between the goats and antelopes. The chamois is the only European species of the group and is indigenous to the mountainous areas, especially the Alps. About nine races are recognized, based on their geographical range. The chamois is, however, becoming rare.

The chamois (*Rupicapra rupicapra*) lives in small herds of both sexes in numbers from 10 to 50. The female leaves the herd after the rutting season for a sheltered area, and after a gestation period of about 20 weeks gives birth to a single young in the spring. The offspring are active almost immediately and begin to forage when 10 days old. They are sexually mature at 3 years of age and the average life-span is about 10 years. The adult is almost 3 ft high and weighs about 90 lb (40 kg) maximum. Both sexes bear horns which are set close together on the forehead, project almost at right angles, and are straight except for the sharp curve backward at the top (see **illus.**). These animals have 32 teeth and the dental formula is I 0/3 C 0/1 Pm 3/3 M 3/3. A soft, pliable leather, known as chamois cloth, is obtained from the skin of this animal. *See* DENTITION.

Alpine chamois (*Rupicapra rupicapra*). (*Brent Huffman/ Ultimate Ungulate Images*)

The goral, serow, and Rocky Mountain goat are included in the same tribe of bovids as the chamois. Both the goral (genus *Naemorhedus*) and serow (genus *Capricornis*) are found in Asia. The goral lives in small herds in the Himalayas, existing on the sparse vegetation found on the rocky slopes. It is about 2 ft (0.6 m) high and has short horns. A single young is born in the early summer after a gestation period of 6 months. The male of the Rocky Mountain goat (*Oreamnos americanus*) is larger than the female, weighing between 200 and 300 lb (90 and 135 kg) and standing over 3 ft (1 m) high. Both sexes are covered with thick, long white hair and have horns and beards. These animals are vegetarians, feeding on moss, lichens, and other vegetation of the area. In the early spring the female bears one or two young, which are quite active soon after birth. *See* ANTELOPE; MAMMALIA. Charles B. Curtin

Chandra X-ray Observatory

A powerful x-ray telescope in orbit around the Earth. Built by the National Aeronautics and Space Administration, it was launched and deployed by space shuttle *Columbia* on July 23, 1999. Named after the Indian-American astronomer Subrahmanyan Chandrasekhar, *Chandra* is unique because of its sensitivity and its extremely high precision mirrors. These features have made possible significant advances in astronomy—especially in relation to the life cycles of stars, the role of supermassive black holes in the

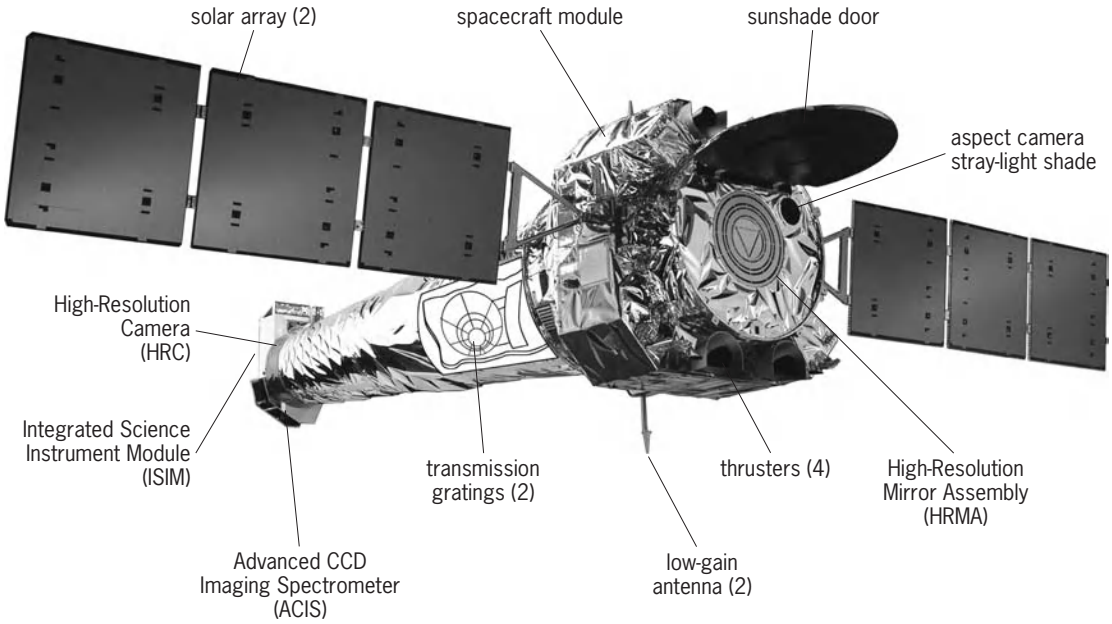

Fig. 1. ***Chandra X-ray Observatory*** **spacecraft.**

evolution of galaxies, and the study of dark matter and dark energy.

X-rays: another form of light. Light is electromagnetic radiation and can take on many forms. Radio waves, microwaves, infrared, visible, ultraviolet, x-ray, and gamma radiation are different forms of light. Each type of light is characterized by a range of wavelengths, with radio wavelengths being the longest and gamma rays the shortest. *See* ELECTROMAGNETIC RADIATION; LIGHT.

When charged particles collide—or undergo sudden changes in their motion—they produce bundles of energy called photons that fly away from the collision at the speed of light. In fact they are light, or electromagnetic radiation. Since electrons have the smallest mass of any known type of charged particle, they are accelerated more easily. According to the laws of electrodynamics, more acceleration means more radiation, so electrons produce most of the photons in the universe. *See* ELECTRON; PHOTON.

The energy of the photon tells what form of light it is. Radio waves are composed of low-energy photons. Optical photons—the only photons perceived by the human eye—are a million times more energetic than the typical radio photon. The energies of x-ray photons range from hundreds to thousands of times higher than that of optical photons. *See* X-RAYS.

The photons collected by x-ray telescopes reveal the hot spots in the universe—regions where particles have been energized or raised to high temperatures by gigantic explosions or intense gravitational fields. X-ray astronomy has revealed that such conditions exist in a remarkable variety of places, ranging from the vast spaces between galaxies to the bizarre, collapsed worlds of neutron stars and black holes.

Description. The *Chandra X-ray Observatory* (**Fig. 1**) has three major parts: (1) the x-ray telescope, whose mirrors focus x-rays from celestial objects; (2) the science instruments which record the x-rays so that x-ray images can be produced and analyzed; and (3) the spacecraft, which provides the environment necessary for the telescope and the instruments to work.

Mirrors and focusing. The mirrors are at the heart of the telescope and are what makes *Chandra* special. They are the smoothest mirrors ever produced—if the continental United States were relatively as smooth as *Chandra's* mirrors, the highest mountain would be less than 13 cm (5 in.) tall.

Although x-rays are a form of light, they do not reflect off mirrors in the same way that visible light does. Because of their high energy, x-ray photons penetrate into a mirror in much the same way that bullets slam into a wall. Likewise, just as bullets ricochet when they hit a wall at a grazing angle, so will x-rays ricochet off mirrors. *See* X-RAY OPTICS.

These properties mean that mirrors built to reflect x-rays look very different from those in optical telescopes. X-ray mirrors have to be precisely shaped and aligned nearly parallel to incoming x-ray photons, so the mirrors resemble barrels in contrast to the dish shape of optical telescopes. They can be nested one inside the other, like Russian dolls, to increase their collecting area. *Chandra* has four nested pairs of mirrors, the largest being about 1.2 m (4 ft) in diameter (**Fig. 2**).

The *Chandra X-ray Observatory* combines the mirrors with four science instruments to record the x-rays from astronomical sources. The incoming x-rays are focused by the mirrors to a tiny spot (about half as wide as a human hair) on the focal plane, about 10 m (30 ft) away. The focal-plane science instruments—or cameras, the Advanced CCD Imaging Spectrometer (ACIS) and the High-Resolution Camera (HRC)—are well matched to capture the

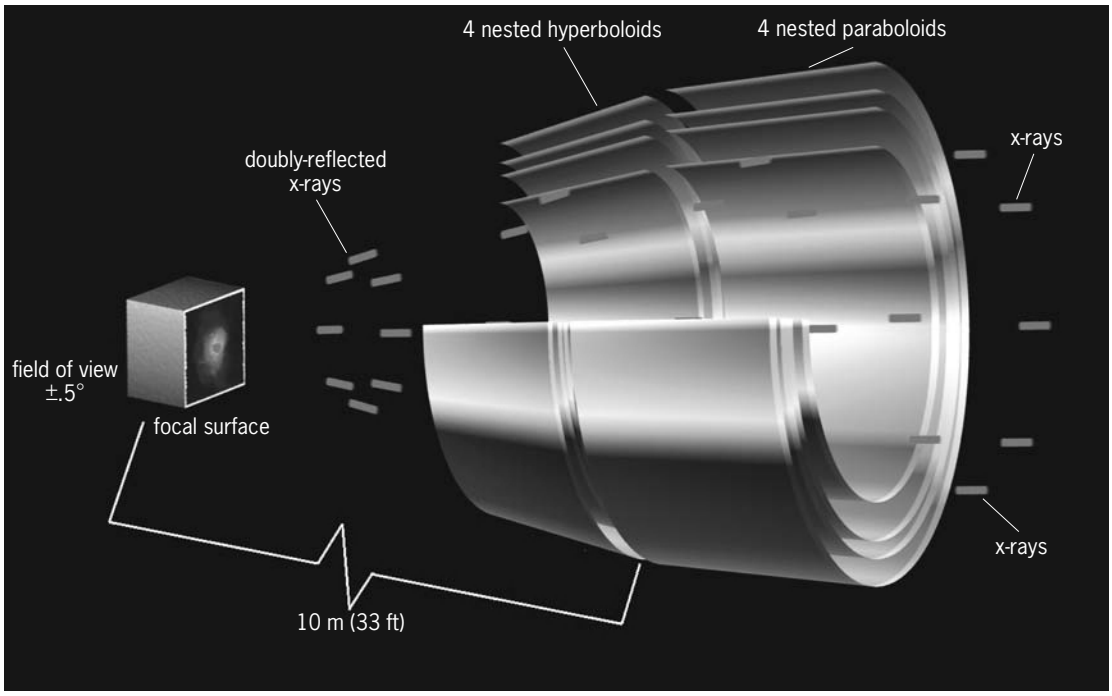

Fig. 2. Schematic diagram of *Chandra*'s grazing-incidence x-ray mirror. Mirror elements are 0.8 m (2.6 ft) long and 0.6–1.2 m (2–4 ft) in diameter. (*NASA/CXC/SAO*)

sharp images formed by the mirrors and to provide information about the incoming x-rays: their number, position, energy, and time of arrival. The close matches of the imaging capability of the instruments and the focusing power of the mirrors enables *Chandra* to make images that reveal detail as small as one-half second of arc. This resolution is equivalent to the ability to read a stop sign at a distance of 20 km (12 mi).

Science instruments. The Advanced CCD Imaging Spectrometer is an array of charge-coupled devices (CCDs), which are sophisticated versions of the crude CCDs used in camcorders. This instrument is especially useful because it can make x-ray images and at the same time measure the energy of each incoming x-ray. It is the instrument of choice for studying temperature variations across x-ray sources such as vast clouds of hot gas in intergalactic space, or chemical variations across clouds left by supernova explosions. The High Resolution Camera is especially useful for making high-quality images and measuring the time variations in x-ray sources. *See* ASTRONOMICAL IMAGING; CHARGE-COUPLED DEVICES.

Two additional science instruments, the transmission grating spectrometers, provide detailed information about the specific energy of the incoming x-rays. These spectrometers are circular gratings that can be swung into position just behind the mirrors, where they redirect (diffract) the x-rays by an amount that depends on their energies. The positions of the x-rays are measured by the focal-plane cameras, so the exact energy of the x-rays can be determined. This enables the temperature, ionization, and chemical composition of the gas around a black hole, for example, to be determined. *See* DIFFRACTION GRATING.

Environment. Because the Earth's atmosphere absorbs cosmic x-rays, x-ray observatories must be placed high above the Earth's surface. This means that the ultraprecise mirrors and detectors, together with the sophisticated electronics that conveys the information back to Earth, must be able to withstand the rigors of a rocket launch and operate in the hostile environment of space.

Orbit. Chandra has an unusual, elliptical orbit that takes the spacecraft more than a third of the way to the Moon before returning to its closest approach to Earth of about 16,000 km (10,000 mi). The time to complete an orbit is 64 hours 18 minutes. The spacecraft spends 85% of its orbit above the belts of charged particles that surround the Earth and interfere with observations. Uninterrupted observations as long as 55 h are possible, and the overall percentage of useful observing time is much greater than for the low Earth orbit of a few hundred miles used by most satellites.

Control and data retrieval. The observatory is controlled by commands transmitted from the Operations Control Center in Cambridge, Massachusetts. A preplanned sequence of observations is uplinked to *Chandra* and stored in the onboard computer for later execution. Data collected by observations with *Chandra* are stored on a recorder for later transmission to the ground every 8 h during regularly scheduled Deep Space Network contacts. The data are transmitted to the Jet Propulsion Laboratory in Pasadena, California, and then to Operations Control at the *Chandra* X-ray Center in Cambridge for processing and analysis by scientists.

See SPACE COMMUNICATIONS; SPACECRAFT GROUND INSTRUMENTATION.

Major scientific achievements. *Chandra*'s unique abilities have made possible significant advances in several areas of astrophysics.

Evolution of stars. *Chandra* gives astronomers a unique window for viewing the drama of stellar evolution, from the formation of stars in dense clouds of dust and gas to their demise, either quietly as white dwarfs or violently as supernovae.

A *Chandra* observation of a rich cluster of young stars in the Orion Nebula showed that young Sun-like stars produce violent x-ray outbursts that are much more frequent and energetic than anything observed today from the 4.6×10^9-year-old Sun. This could have implications for understanding the conditions under which the solar system evolved. *See* ORION NEBULA; SOLAR SYSTEM; STELLAR EVOLUTION.

Stars much more massive than the Sun end their lives with a supernova explosion, one of the most energetic events in the universe. The explosion ejects material rich in the heavy elements necessary for planets and life—carbon, nitrogen, oxygen, silicon, and iron. As this stellar debris plows into surrounding gas and dust, it generates tremendous shock waves that heat gas to millions of degrees. *Chandra* images and spectra have allowed scientists to trace these shock waves and to study the amount and distribution of heavy elements expelled in the explosion (**Colorplate 1**). *See* SHOCK WAVE; SUPERNOVA.

Superbubbles of hot gas thousands of light-years across can be created by the cumulative effect of hundreds of supernovae. *Chandra* observations of the Antennae Galaxies show that superbubbles disperse heavy elements over vast distances (**Colorplate 2**).

Many supernovae leave behind a neutron star—a dense, whirling ball of neutrons that is as massive as the Sun, yet only 20 km (12 mi) in diameter. *Chandra* images have revealed spectacular rings and jets of high-energy particles created by rapidly rotating neutron stars (**Fig. 3**). *See* CRAB NEBULA; NEUTRON STAR; PULSAR.

Black hole discoveries. *Chandra* observations have revealed the presence of many previously undetected stellar black holes in the Milky Way Galaxy and in nearby galaxies, including a possible new class of intermediate-mass black holes (**Colorplate 3**). *See* MILKY WAY GALAXY.

On a larger scale, *Chandra* has shown that there are many more active supermassive black holes in the centers of galaxies than previously thought. At the center of a pair of colliding galaxies, *Chandra* found evidence for two supermassive black holes orbiting each other (**Colorplate 4**). In a few hundred million years, these huge black holes will merge to form an even larger black hole.

Chandra images have revealed startling evidence for the repetitive and far-reaching explosive activity associated with supermassive black holes. Jets several hundred thousand light-years in length have been traced back to central supermassive black holes. A *Chandra* observation of the Perseus galaxy cluster has revealed wavelike features that appear to be sound waves produced by jetlike explosive events occurring in the vicinity of a supermassive black hole in Perseus A, the huge galaxy at the center of the cluster. In another cluster, *Chandra* found evidence for what may be the most energetic explosion ever detected. These massive explosions may be part of a cosmic feedback cycle in which the black hole regulates its own growth and the growth of the galaxy in which it resides. *See* BLACK HOLE; GALAXY, EXTERNAL.

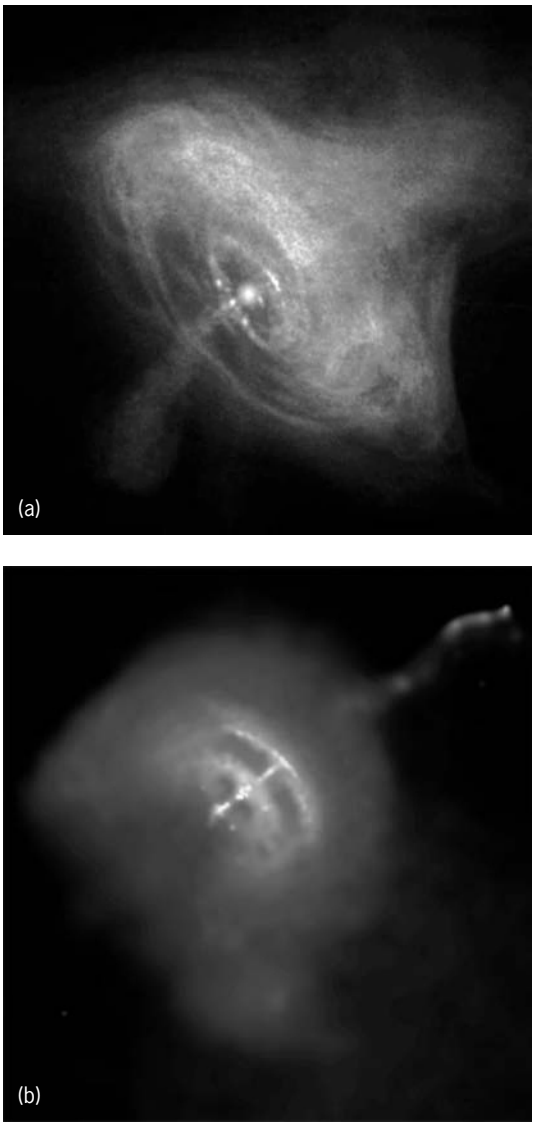

Fig. 3. ***Chandra*** **images of x-rays from the vicinity of rapidly rotating neutron stars. (*a*) Crab Nebula, the remnant of a supernova that was seen in 1054 (*NASA/CXC/SAO*). (*b*) Vela pulsar. A jet of high-energy particles is emerging toward the upper right (*NASA/CXC/PSU/G. Pavlov et al.*).**

Galaxy clusters, dark matter, and the cosmic web. A large cluster of galaxies contains hundreds to thousands of galaxies immersed in an enormous cloud of multimillion-degree hot gas held together by dark matter. Such galaxy clusters are the largest gravitationally bound systems in the universe. The intergalactic hot gas clouds in clusters are detected as x-ray sources and can extend over a region several million light-years in diameter.

The hot gas in a galaxy cluster is in equilibrium with the gravitational pull of all the mass in the cluster. X-ray studies of numerous clusters have shown that 70 to 90% of the mass of a typical cluster consists of dark matter—hypothetical particles left over from the dense early universe that interact with each other and normal matter only through gravity. The favored type of dark matter particle is cold dark matter, which gets its name from the assumption that dark matter particles were moving slowly when galaxies and galaxy clusters began to form. *See* DARK MATTER.

While an explanation for the source of dark matter is still lacking, an effect that is even more enigmatic has been discovered. Astronomers have observed that the visible light from a certain type of supernova of well-known properties is fainter than expected in distant galaxies. The best explanation is that these supernovae are more distant than originally thought, which implies that the expansion of the universe must be accelerating. *Chandra*'s measurements of the dark matter content of clusters of galaxies have provided independent evidence for this remarkable result. The data are consistent with a model in which the expansion of the universe was first decelerating until about 6×10^9 years ago and then began to accelerate. *See* ACCELERATING UNIVERSE.

Cosmic acceleration can be explained if the space between galaxies is filled with a mysterious dark energy. The existence of dark energy and its peculiar properties require either a modification of Albert Einstein's theory of general relativity or a major revision of some other area of fundamental physics. *See* DARK ENERGY.

Assuming that dark energy is responsible for cosmic acceleration, the *Chandra* results, the supernova observations, and the observations of the cosmic microwave background radiation taken together indicate that dark energy makes up about 75% of the energy density of the universe, dark matter about 21%, and visible matter about 4%. *See* COSMIC BACKGROUND RADIATION; COSMOLOGY; UNIVERSE; X-RAY ASTRONOMY; X-RAY TELESCOPE. Wallace Tucker

Bibliography. E. M. Schlegel, *The Restless Universe*, Oxford University Press, 2002; K. S. Thorne, *Black Holes and Time Warps*, W. W. Norton, 1994; W. Tucker and K. Tucker, *Revealing the Universe: The Making of the Chandra X-ray Observatory*, Harvard University Press, 2001; K. Weaver, *The Violent Universe: Joyrides through the X-ray Cosmos*, Johns Hopkins University Press, 2005.

Channel electron multiplier

A single-particle detector which in its basic form (**Fig. 1**) consists of a hollow tube (channel) of either glass or ceramic material with a semiconducting inner surface. The detector responds to one or more primary electron impact events at its entrance (input) by producing, in a cascade multiplication process, a charge pulse of typically 10^4–10^8 electrons at its exit (output). Because particles other than electrons can impact at the entrance of the channel electron multiplier to produce a secondary electron, which is then subsequently multiplied in a cascade, the channel electron multiplier can be used to detect charged particles other than electrons (such as ions or positrons), neutral particles with internal energy (such as metastable excited atoms), and photons as well. As a result, this relatively simple, reliable, and easily applied device is employed in a wide variety of charged-particle and photon spectrometers and related analytical instruments, such as residual gas analyzers, mass spectrometers, and spectrometers used in secondary ion mass spectrometry (SIMS), electron spectroscopy for chemical analysis (ESCA), and Auger electron spectroscopy. *See* AUGER EFFECT; ELECTRON SPECTROSCOPY; MASS SPECTROMETRY; SECONDARY ION MASS SPECTROMETRY (SIMS); SPECTROSCOPY.

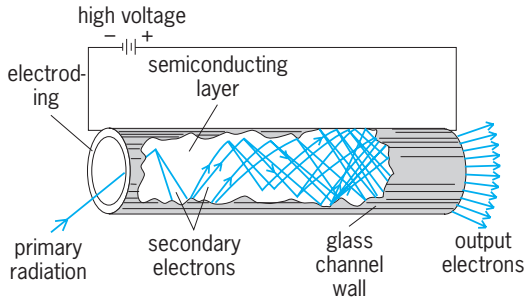

Fig. 1. Cutaway view of a straight, single-channel electron multiplier, showing the cascade of secondary electrons resulting from the initial, primary radiation event, which produces an output charge pulse. (*After J. L. Wiza, Microchannel plate detectors, Nucl. Instrum. Meth., 162:587–601, 1979*)

Photon detection. To directly detect photons, the photon energy must exceed the work function of the multiplier surface. Most channel electron multipliers require energies exceeding 5–10 eV for moderately efficient single-photon detection, limiting direct use to vacuum-ultraviolet wavelengths and shorter. Detection of longer-wavelength, lower-energy photons can be accomplished by the addition of a suitable photoelectron-emitting surface (photocathode) near the channel electron multiplier input, with an electric field to sweep liberated photoelectrons into the channel electron multiplier. *See* OPTICAL DETECTORS; PHOTOEMISSION; PHOTOMULTIPLIER; PHOTOTUBE.

Cascade multiplication. The cascade multiplication process at the heart of the channel electron multiplier's operation results from the secondary emission properties and high resistance (typically $10^9\ \Omega$) of the inner channel surface. For multiplication to occur, there must be a high probability (expressed as a high secondary emission coefficient) that an electron impacting the surface will liberate more than one secondary electron, each of which is then accelerated toward the output to impact the surface and produce even more secondaries. In operation, it is necessary to apply a high voltage of the order of 1–4 kV between the input and output of the channel, thereby producing an electric field within the

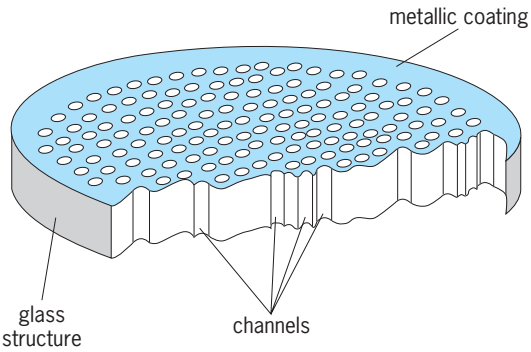

Fig. 2. Cutaway of a microchannel plate, or channel electron multiplier array, not to scale. Metallic coatings on both sides of the plate serve as input and output electrodes. (*After J. L. Wiza, Microchannel plate detectors, Nucl. Instrum. Meth., 162:587–601,1979*)

channel that accelerates secondary electrons to an energy sufficient to guarantee a high secondary emission coefficient. An electron impact energy of the order of 300 eV is generally required. High surface resistivity limits the current required from the high-voltage power supply to a few microamperes. *See* SECONDARY EMISSION.

Ion feedback. The large flux of electrons produced near the exit of a channel electron multiplier leads to a relatively high probability of electron impact ionization of residual gas molecules in the channel. The resulting ions are accelerated by the internal field toward the input, where they can impact the surface to liberate electrons that are then multiplied in the usual fashion. This is a potentially unstable condition called ion feedback and can lead to spurious outputs almost completely unrelated to real inputs. Several techniques are employed to reduce or eliminate ion feedback. Operating the channel electron multiplier in high vacuum (pressure less than 10^{-3} pascal or 10^{-5} torr) controls ion feedback by reducing residual gas density in the channel. Lowering the channel voltage limits the channel gain and hence the electron flux, thereby lowering the ionization rate. Finally, by introducing curvature in the channel, the distance that an ion can travel toward the input can be limited; since most ions are produced near the exit, limiting this distance also limits the multiplication gain and thus the amplitude of spurious ion feedback signals. Virtually all commercially available channel electron multiplier devices employ channel curvature to control ion feedback.

Configuration. A typical single channel electron multiplier will have a channel radius of 0.04 in. (1 mm) and a length of 4 in. (10 cm). The channel is often formed into a spiral or helix to conserve space and to control ion feedback. It is common to find the entrance flared into a funnel shape to increase the surface area sensitive to the primary event.

Microchannel plate. A related device is the channel electron multiplier array, often called a microchannel plate (**Fig. 2**). The channel electron multiplier array is usually a disk-shaped device with a diameter between 1 and 4 in. (2.5 and 10 cm) and a thickness of a fraction of a millimeter, and consists of millions of miniature channel electron multiplier devices arranged with channel axes perpendicular to the face of the disk. In one widely used version, the individual channels have a diameter of about 12 micrometers and are spaced at 15-μm intervals. Channel electron multiplier arrays find application as image intensifiers in night vision devices and are employed to add either large detection area or imaging capabilities, or both, to charged-particle detectors and spectrometers. Ion feedback in channel electron multiplier arrays is often controlled by using two or more arrays in succession and arranging the channel axes at an angle to the axes of the disks, so that the individual channels in successive arrays form a chevron or herringbone pattern with those in previous stages.

Position-sensitive detector. The microscopic channels in a channel electron multiplier array allow the position coordinates of an input primary event (such as a charged particle or photon) to be approximately preserved during the cascade multiplication process. This position sensitivity makes the microchannel plate amplifier a useful imaging device, with position or imaging resolution equal to a few channel spacings in a typical chevronned multiplate device. A microchannel plate with a means of reading out the coordinates of the detected output charge pulse is often called a position-sensitive detector. Several methods of position readout are commonly employed. In the image tube method, the output charge pulses strike a phosphor coating. The "image" of the event distribution at the amplifier input is converted into a visible light image, suitable for processing with conventional video devices [such as a charge-coupled-device (CCD) camera or a vidicon] and techniques. *See* CHARGE-COUPLED DEVICES; TELEVISION CAMERA TUBE.

In the discrete multiple-anode method, the output charge pulse is collected on one or more of

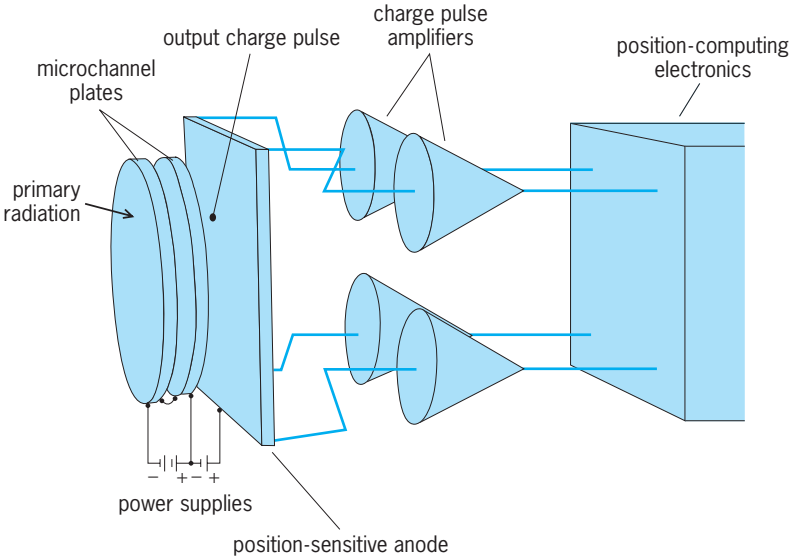

Fig. 3. Position-sensitive detector system. The position-sensitive anode collects a charge "cloud" produced by the microchannel plate amplifier in response to primary radiation. In the charge-division method, shown here, the position of the primary event is calculated from the relative sizes of the charge pulses that subsequently migrate to the corners of the anode.

an array of conducting anodes, with the identity of the active anodes, possibly in a digitally encoded form, then determining the event position. The frequently used charge-division method employs collection of the output pulse on a resistive anode, which is terminated either at two opposite edges (for one-dimensional readout) or at four corners (for two-dimensional readout) [**Fig. 3**]. The relative fractions of the total output charge pulse of a single input event, collected at the two or four terminals, can then be used to determine the event position coordinates using relatively straightforward electronic or computing techniques. A less common time-division method uses the distributed resistance and capacitance of a resistive anode in proximity to the rear surface of the exit microchannel plate (and other possible nearby structures), and the event-position-dependent time delay thereby introduced into the peaking time of the collected charge pulses, to determine event coordinates. *See* IMAGE TUBE (ASTRONOMY); LIGHT AMPLIFIER; PARTICLE DETECTOR.

Stuart B. Elston

Bibliography. M. Lampton, The microchannel image intensifier, *Sci. Amer.*, 245(5):62–71, November 1981; J. H. Moore, C. C. Davis, and M. A. Coplan, *Building Scientific Apparatus: A Practical Guide to Design and Construction,* 2d ed., 1989.

Channeling in solids

The steering of positively charged energetic particles between atomic rows or planes of a crystalline solid. The particles can be positive ions, protons, positrons, or muons. If the angle between the direction of the particle and a particular axis or plane in the crystal is within a small predictable limit (typically a few degrees or less), then the gradually changing electrostatic repulsion between the particle and each successive atomic nucleus of the crystal produces a smooth steering through the crystal lattice (**Fig. 1**). Thus, the trajectory of the channeled particle is restricted to the open spaces between atomic rows and planes of a crystal (**Fig. 2**). *See* CRYSTAL STRUCTURE.

An obvious consequence of this steered motion is that it prevents violent collisions of the particles with atoms on the lattice sites. Hence, as compared with a randomly directed beam of particles, the channeled beam loses energy more slowly, penetrates more deeply, creates much less damage to the crystal along its track, and is prevented from participating in all close-encounter processes (nuclear reactions, Rutherford scattering, and so forth) with lattice atoms. *See* NUCLEAR REACTION; SCATTERING EXPERIMENTS (NUCLEI).

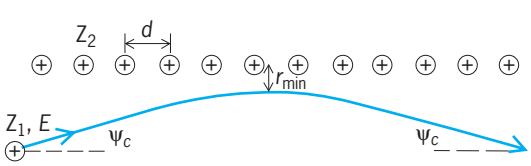

Fig. 1. Typical trajectory of a channeled ion (nuclear charge Z_1, energy E), illustrating the gentle steering caused by electrostatic repulsion from the regularly spaced nuclei of a crystal (nuclear charge Z_2, spacing d), so that the ion approaches no closer to the nuclei than r_{min}. The transverse angle ψ_c with respect to the channel axis remains almost constant as the ion penetrates through the crystal lattice.

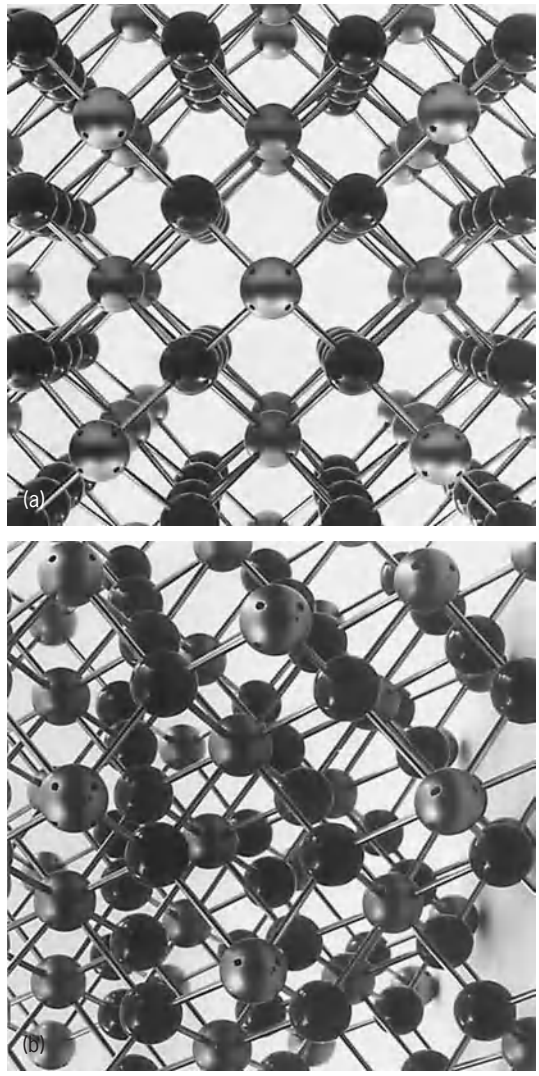

Fig. 2. Model of a cubic crystal lattice. (*a*) Model viewed along a ⟨100⟩ axis. The open (square) spaces are the ⟨100⟩ axial channels. (*b*) The same model viewed along a randomly chosen direction roughly 20° away from a ⟨100⟩ axis.

A related channeling phenomenon is the channeling of energetic electrons or other negative particles. In this case, the particles are attracted to the positively charged atomic nuclei, so that the probability of violent collisions with atoms on lattice sites is enhanced rather than being prevented, and the particles are steered along the rows or planes of nuclei rather than between them.

Blocking. A closely related phenomenon is called blocking. In this case, the energetic positive particles originate from atomic sites within the crystal lattice by means of fission, alpha-particle decay, or by wide-angle scattering of a nonchanneled external beam in a very close encounter with a lattice atom. Those

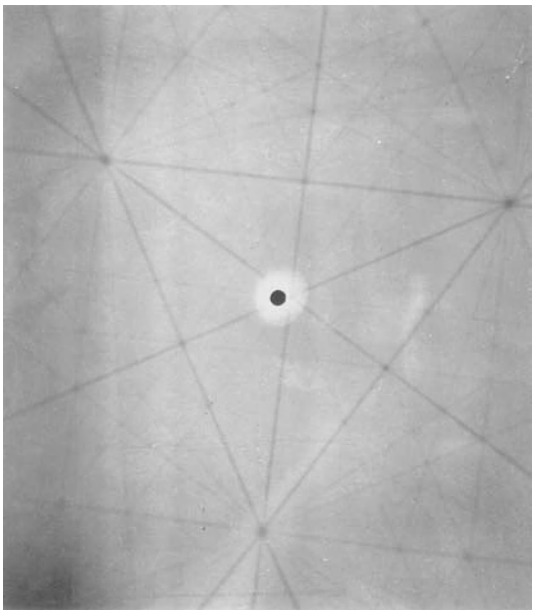

Fig. 3. Blocking pattern produced by a beam of 1-MeV He^+ ions after backscattering from a germanium crystal onto a cellulose nitrate film 6 in. (15 cm) away. The large central spot is the hole in the film through which the incident beam reached the crystal. The intricate pattern of thin dark lines (that is, unbombarded regions of film) corresponds to the network of atomic planes along which backscattered particles are absent due to blocking. The wider spots at the intersection of several planes correspond to blocking along the major crystal axes.

particles emitted almost parallel to an atomic row or plane will be deflected away from the row by a steering process similar to that shown in Fig. 1. Consequently, no particles emerge from the crystal within a certain critical blocking angle of each major crystallographic direction. A piece of film placed some distance from the crystal provides a simple technique for recording blocking patterns (**Fig. 3**). Theoretical considerations show that the same principle is involved in blocking as in channeling; hence, both phenomena exhibit an identical dependence on particle energy, nuclear charge, lattice spacing, and so forth. *See* NUCLEAR FISSION; RADIOACTIVITY.

Development. Channeling effects were predicted by J. Stark in 1912. However, despite the evident simplicity of the channeling phenomenon, firm experimental evidence for its existence was not recognized until the early 1960s, when it was observed that energetic heavy ions sometimes penetrated to unexpectedly large depths in polycrystalline aluminum and tungsten. In 1962, it was shown by computer simulation that these deep penetrations were probably due to particles being channeled along the open directions in some correctly oriented microcrystals. Single-crystal experiments in several laboratories subsequently confirmed that such channeling was indeed an extremely general phenomenon, occurring for all charged nuclear particles (atomic nuclei, positrons, muons, and so forth) over an energy range from a few electronvolts up to hundreds of gigaelectronvolts. At about the same time, a comprehensive theoretical framework for predicting the behavior of channeled trajectories was developed, which led not only to an understanding of the channeling phenomenon but also to its widespread application in many other fields.

Applications. Applications of channeling include the location of foreign atoms in a crystal, the study of crystal surface structure, and the measurement of nuclear lifetimes.

Location of foreign atoms. The location of foreign (solute) atoms in a crystal is one of the simplest channeling applications. It is accomplished by measuring the yields of Rutherford back-scattered particles, characteristic x-rays, or nuclear reaction products produced by the interaction of channeled particles with the solute atoms. Such yields are enhanced for solute atoms that are displaced into channels of the crystal. Thus the solute atom sites can be identified by a triangulation technique, illustrated in **Fig. 4** for a two-dimensional model of a cubic lattice. Numbers in angle brackets ⟨ ⟩ identify the different kinds of crystal axes, whereas numbers in square brackets identify specific directions within a given kind. The shaded areas are shadowed by rows of host atoms (○) in the [10] and [11] channels. As shown by the table in Fig. 4, three possible sites for a foreign atom (●, **x**, and □) are readily distinguished by comparing the channeling behavior along the ⟨10⟩- and ⟨11⟩-type directions. Channeled ions do not have close-encounter interactions with solute atoms in substitutional positions (●), since these are completely shadowed by atomic rows in both ⟨10⟩- and ⟨11⟩-type channels. The **x** interstitial positions (halfway along the sides of the until cells) are 50% shadowed in both the [10] and [01] channels, but lie in the center of the [11] and [1$\bar{1}$] channels. The □ interstitial positions (in the center of the square unit cell) are completely

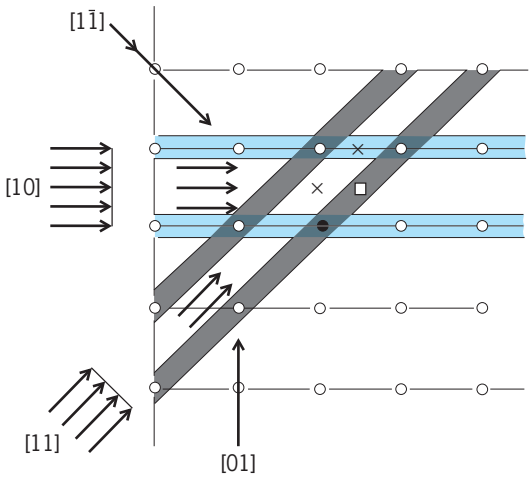

	shadowed fraction		
	●	×	□
⟨10⟩	1.0	0.5	0.0
⟨11⟩	1.0	0.0	1.0

Fig. 4. Two-dimensional model illustrating how the channeling effect may be used to locate foreign atoms in a crystal. (*After J. N. Mundy et al., eds., Methods of Experimental Physics, vol. 21, Solid State: Nuclear Methods, Academic Press, 1983*)

shadowed in both ⟨11⟩-type channels, but lie in the center of ⟨10⟩-type channels. *See* CRYSTAL DEFECTS.

This method has been used to determine the lattice positions of solute atoms that have been introduced into crystals by a variety of means (for example, by melting, diffusion, or ion implantation). For example, channeling measurements show that furnace or laser heating of silicon that has been implanted with energetic boron, arsenic, or other electrically active atoms causes these atoms to move onto normal (substitutional) lattice sites, sometimes to concentrations far in excess of solubility limits. These measurements also are used to determine the amount of lattice damage (created by the ion implantation) which remains after the heating. These and similar applications have proved extremely useful in the development of semiconductor devices. *See* ION IMPLANTATION; LASER-SOLID INTERACTIONS.

When energetic particles strike a crystal, atoms can be ejected from normal lattice sites, leaving vacant lattice sites (vacancies). Solute atoms can trap these vacancies, creating unique geometric configurations, such as the tetravacancy–solute-atom complex (**Fig. 5**). In this complex, the solute atom (for example, a tin atom in aluminum) becomes displaced from a normal lattice site into the tetrahedral interstitial site, surrounded by four lattice vacancies at equal distance. Such a solute atom lies in the center of ⟨100⟩ channels, is halfway to the center of ⟨110⟩ channels, and is completely shadowed in ⟨111⟩ channels. Thus, its position is easily identified by channeling measurements along these three axial channels. These vacancy–solute-atom complexes are very stable thermally, and probably form nucleation centers for the creation of even larger vacancy clusters (voids) in irradiated materials. *See* RADIATION DAMAGE TO MATERIALS.

Study of surface structure. In a similar way, channeling techniques are used to study the structure of crystal surfaces. Because of the asymmetric forces between atoms on the surface of a crystal, the surface plane of atoms is often reconstructed (that is, it has a different atomic structure from that of the bulk); also, it may be relaxed inward or outward. In **Fig. 6**, a method for measuring an outward relaxation of the surface plane of atoms by the channeling technique

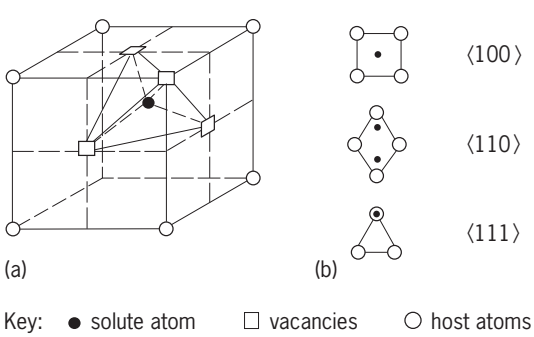

Fig. 5. Tetravacancy–solute-atom complex. (*a*) Perspective view. (*b*) Projection of the solute atom into ⟨100⟩-, ⟨110⟩-, and ⟨111⟩-type axial channels. (*After J. N. Mundy et al., eds., Methods of Experimental Physics, vol. 21, Solid State: Nuclear Methods, Academic Press, 1983*)

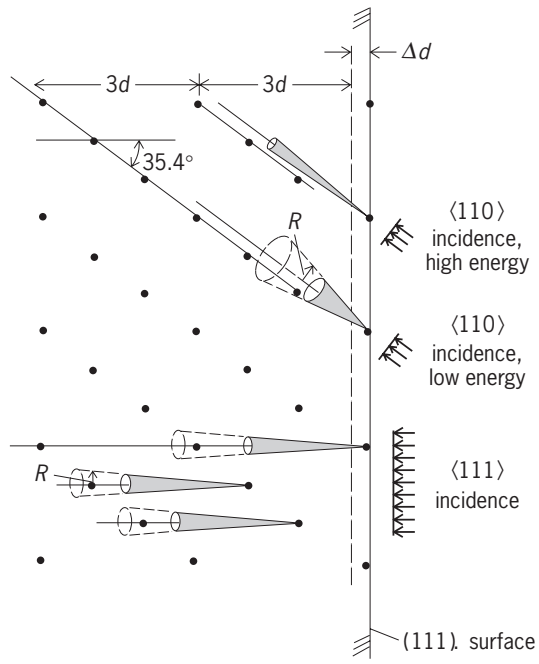

Fig. 6. Atomic configuration near the surface of a Pt(111) crystal, illustrating the use of channeling to investigate Δd, the surface relaxation. R = radius of shadow cone cast by surface atom; d = bulk (111) planar spacing. (*After J. A. Davies et al., Measurement of surface relaxation by MeV ion backscattering and channeling, Phys. Lett., 54A:239–240, North Holland, 1975*)

is shown. A beam of ions directed perpendicular to the (111) surface (⟨111⟩ incidence) "sees" only one atom per row of atoms. However, because of the outward displacement of the surface atoms, a beam of ions incident along a nonperpendicular (for example, ⟨110⟩) axis "sees" two atoms per row at high ion energies (at which each surface atom casts a small shadow cone), and one atom per row at low ion energies (giving a large shadow cone). Information on surface relaxations (Δd) as small as 2×10^{-12} m and also on the vibrational amplitudes of the surface plane of atoms can be obtained by such studies. *See* LATTICE VIBRATIONS; SURFACE PHYSICS.

In addition, premelting at surfaces can be investigated. Channeling studies of the surface of lead near the melting point have shown that a thin surface layer becomes molten at a lower temperature than that of the bulk material. Such studies will contribute to an understanding of melting, which remains one of the more poorly understood physical phenomena.

Measurement of nuclear lifetimes. A different type of-break application is the measurement of extremely short nuclear lifetimes. Nevertheless, the same basic principles of atom location are involved; namely, the ability, by means of channeling or blocking, to pinpoint accurately the location of a nucleus (relative to the crystal lattice) at the instant that it is struck by or emits an energetic positive particle. Suppose that a lattice nucleus under external beam bombardment captures a beam particle to form a compound nucleus, recoiling with known velocity from its lattice site, and that this nucleus subsequently decays by emitting an energetic positive particle (proton,

alpha particle, fission fragment, and so forth). From the resulting blocking pattern of these emitted particles, the recoil distance before emission and hence the nuclear lifetime can be accurately determined. In this way, lifetimes as short as 10^{-18} s are directly observable. John A. Davies; M. L. Swanson

Bibliography. L. C. Feldman, J. W. Mayer, and S. Thomas Picraux, *Materials Analysis by Ion Channeling*, 1982; J. N. Mundy et al. (eds.), *Methods of Experimental Physics*, vol. 21, *Solid State: Nuclear Methods*, 1983; E. Uggerhoj and A. Zucker (eds.), *Channeling and Other Penetration Phenomena*, 1989.

Chaos

System behavior that depends so sensitively on the system's precise initial conditions that it is, in effect, unpredictable and cannot be distinguished from a random process, even though it is deterministic in a mathematical sense. This article begins with a discussion of the notions of order, chaos, and noise as they occur in deterministic dynamical systems, the relation of chaos and periodicity, and the concept of an attractor. Applications of chaos to atmospheric prediction, weather, climate, electronic circuits, astronomy, acoustics, and atoms are then discussed.

Throughout history, sequentially using magic, religion, and science, people have sought to perceive order and meaning in a seemingly chaotic and meaningless world. This quest for order reached its ultimate goal in the seventeenth century when newtonian dynamics provided an ordered, deterministic view of the entire universe epitomized in P. S. de Laplace's statement, "We ought then to regard the present state of the universe as the effect of its preceding state and as the cause of its succeeding state." In everday life, the predictable swing of a long, massive pendulum or the regular motion of the Sun, Moon, and planets provides reassurance of a mechanistic newtonian order in the world.

But if the determinism of Laplace and Newton is totally accepted, it is difficult to explain the unpredictability of a gambling game or, more generally, the unpredictably random behavior observed in many newtonian systems. Commonplace examples of such behavior include smoke that first rises in a smooth, streamlined column from a cigarette, only to abruptly burst into wildly erratic turbulent flow (**Fig. 1**); and the unpredictable phenomena of the weather. *See* FLUID FLOW; TURBULENT FLOW.

At a more technical level, flaws in the newtonian view had become apparent by about 1900, leading J. H. Poincaré to remark, "Determinism is a fantasy due to Laplace," and J. C. Maxwell to assert, "The true logic of this world is the calculus of probabilities." The problem is that many newtonian systems exhibit behavior which is so exquisitely sensitive to the precise initial state or to even the slightest outside perturbation that, humanly speaking, determinism becomes a physically meaningless though mathematically valid concept. But even more is true.

Fig. 1. Transition from order to chaos (turbulence) in a rising column of cigarette smoke. The initial smooth streamline flow represents order, while the erratic flow represents chaos.

Many deterministic newtonian-system orbits are so erratic that they cannot be distinguished from a random process even though they are strictly determinate, mathematically speaking. Indeed, in the totality of newtonian-system orbits, erratic unpredictable randomness is overwhelmingly the most common behavior. *See* CELESTIAL MECHANICS; CLASSICAL MECHANICS; DETERMINISM; STOCHASTIC PROCESS.

Examples. These notions will be illustrated through three simple mapping systems which retain many features of general dynamical systems but are not encumbered by extraneous detail.

Highly ordered motion. An example of highly ordered motion is the discrete mapping of points around a circle of unit circumference in which each point is carried one-fifth way around the circle upon each iteration. After n iterations, a point initially at angular position θ_0 has been carried to position θ_n given by Eq. (1),

$$\theta_n = (0.2)n + \theta_0 \tag{1}$$

where the angle θ has been normalized so that $\theta + 1 = \theta$. This mapping is obviously determinate since each θ_0 uniquely determines an "orbit" of θ_n iterates. The "motion" here is also highly ordered because, from Eq. (1), all orbits are periodic with $\theta_{n+5} = \theta_n$, because the entire circle of points is mapped as a rigid body upon each iteration, and because the angle of rotation is the precisely computable, terminating rational number, $1/5 = 0.2$. Moreover, except for a displacement around the circle, all orbits are identical. This mapping system is thus fully ordered, strictly deterministic, and totally dull. In more general terms,

the monotony of the streamline flow in Fig. 1 may be contrasted with the greater variety of behavior of the turbulent flow. Ordered motion while comfortingly simple is also rigidly constrained motion in which all or almost all capricious options have been discarded.

Ordered motion. A less constrained but still ordered motion is generated by the mapping which rigidly rotates the whole circle of points through the larger angle upon each iteration. Here the result of n iterations is given by Eq. (2), where again

$$\theta_n = (\sqrt{2})n + \theta_0 \tag{2}$$

$\theta + 1 = \theta$. Like the mapping of Eq. (1), except for a displacement around the circle, all orbits are identical. However, here there is one exciting difference. Because the angle of rotation α is now irrational, the θ_n iterates of each θ_0 in Eq. (2) spread out uniformly around the circle, almost as if their positions were selected from a uniform random distribution. In fact, it is possible to rigorously prove for Eq. (2) that Eq. (3) is satisfied, where f is any reasonable

$$\lim_{N\to\infty}\left(\frac{1}{N}\right)\sum_{n=0}^{N} f(\theta_n) = \int_0^1 f(\theta)\,d\theta \tag{3}$$

well-behaved function. Equation (3), a version of the ergodic theorem, is merely a sophisticated way to assert that all θ values occur with equal frequency and that the mapping of Eq. (2) therefore passes one test for randomness. However, this slight randomness in the mathematically deterministic Eq. (2) has appeared only because a sacrifice has been tacitly made for it. Indeed, the introduction of an irrational angle of rotation into Eq. (2) which yields a uniform θ distribution simultaneously also renders Eq. (2) not fully determinate from the physical viewpoint. Specifically, humans can never compute the precise decimal representation for $\sqrt{2}$. Thus, if ϵ is the error in the determination of $\sqrt{2}$, then the error $\Delta(n)$ in θ_n will be at least $\Delta(n) = n\epsilon$. Since this error clearly grows without bound with iteration number n, physical determinism in Eq. (2) is completely lost when the error $\Delta(n)$ is about unity. Nonetheless, physical determinism over human time scales can be maintained, provided the error ϵ in the determination of $\sqrt{2}$ is sufficiently small.

It is now possible to proceed through a sequence of maps which become more "random" and less "physically determinate." For them, the error $\Delta(n)$ in each θ_n would grow as some polynomial in n, but the notions of order and physical determinism could still be maintained, at least over human time scales, by using humanly available accuracy. The great divide occurs when the error growth becomes exponential, yielding mathematically determinate systems which pass every test for randomness. Here the mapping equations require such massive input of data strings in order to maintain accuracy that, informationally speaking, they are not computing or determining the solution but rather merely copying out the solution being given to them. This case will be illustrated with a final example.

Chaotic motion. If the points on the circle of unit circumference move according to the "multiplicative" mapping equation (4), where $\theta + 1 = \theta$ as before,

$$\theta_n = 2^n\theta_0 \tag{4}$$

then orbits starting with slightly differing θ_0 values clearly separate exponentially with iteration number n. In consequence, Eq. (4) exhibits that extreme sensitivity of final state to precise initial state characteristic of all chaotic systems; indeed, Eq. (4) is perhaps the simplest system which can serve as a paradigm for chaos.

However, Eq. (4) is no less mathematically deterministic than are Eqs. (1) and (2), since a θ_n set exists and is unique for each given θ_0. The θ_n iterates of Eq. (4) can nevertheless be unpredictably random and nondeterministic, as can be seen by considering the consequences for determinism of the exponentially separating orbits in Eq. (4). For example, if the spatial angle θ_0 is known to an accuracy of 10^{-31}, far beyond the accuracy available to contemporary science, even then, after only 100 iterations, the error in θ_{100} is of order unity, and the last vestige of determinism in Eq. (4) has been lost.

This issue may be further clarified by computing a full set of θ_n iterates of Eq. (4), each having an accuracy of at least 2^{-1}, that is, at least one-binary-digit accuracy. If 2^{-m} is the accuracy of θ_0 required to compute a given θ_n, then Eq. (5) must be satisfied. Thus

$$2^{-1} = 2^n 2^{-m} \tag{5}$$

$m = n + 1$ or, for large n, $m \cong n$. Now the situation becomes clear. In order to ensure n informational bits of solution output, one per θ_n iterate, it is necessary to include a total of $m = n$ informational bits into the binary representation of θ_0. Equation (4) is thus seen to be merely a "copy" algorithm rather than a "deterministic" one. Alternatively stated, Eq. (4) is merely a fixed "language translation" algorithm which translates a "German" θ_0 into an "English" θ_n sequence.

Even greater insight into these matters can be obtained by writing the θ_0 of Eq. (4) in binary notation as a digit string of zeros and ones; typically it might be given by Eq. (6). Then, each iterate θ_n specified by

$$\theta_0 = 0.11101100000100\ldots \tag{6}$$

Eq. (4) may be obtained by merely moving the "decimal" in Eq. (6) n places to the right and dropping the integer part. Now, the string of zeros and ones in Eq. (6) may be regarded as specifying a coin-toss sequence. Moreover, the set of all binary θ_0 sequences is in a one-to-one correspondence (essentially) with the set of all possible random coin-toss sequences. Consequently, almost all binary θ_0 sequences pass every humanly computable test for randomness. In short, they are random. A straightforward derivation then proves that almost all θ_0 yield random θ_n sequences despite that fact that Eq. (4) is mathematically deterministic.

Equation (4) is mathematically deterministic provided that θ_0 is given or somehow determined

mathematically; however, in Eq. (4), randomness arises precisely because the infinite digit string for θ_0 is, in general, as random as a coin toss and therefore not computable by any finite algorithm. Moreover, and equally important, the exponentially sensitive dependence of final state upon initial state in Eq. (4) transforms randomness of the θ_0 digit string into randomness of the θ_n sequence. Indeed, the θ_n sequences of Eqs. (1) and (2) lack full randomness only because they lack this sensitive dependence upon initial state.

Equation (4), as mentioned above, is the epitome of chaos just as Eq. (1) is the epitome of order. In addition, Eq. (4) exhibits that richness of behavior inherent to all chaotic systems: everything that can happen does happen. In Eq. (4), for example, the full orbit is known once the θ_0 digit string of Eq. (6) is known. But here all digit strings are possible and, in fact, each digit string actually occurs for some θ_0. Thus, chaos results in richness of opportunity rather than meaninglessness. Joseph Ford

Noise. So far, the discussion has focused on the internal dynamics of a hypothetical system. In many applications, external effects also have an impact on behavior of a system. Then, Eq. (7), which summa-

$$\theta_n = F(n, \theta_0) \tag{7}$$

rizes the governing equations for θ_n as a function F of n and θ_0, in the forms of Eqs. (1), (2), and (4), might be written as Eq. (8), in which ε_n represents

$$\theta_n = F(n, \theta_0) + \varepsilon_n \tag{8}$$

the external effect. When it is random and unpredictable, ε_n is usually called noise. Such an external effect causes at least part of θ_n to be random and unpredictable as well.

In summary, chaos is the apparently unpredictable behavior of a nonlinear deterministic system, while noise is an apparently random effect impinging on the system from an external source. The noise may be purely random or the chaotic behavior of another system.

Unpredictable behavior of a deterministic system may thus arise from internally generated chaos, from random forces imposed on the system, or from a combination of both. With real systems, it may be difficult or impossible to distinguish the two phenomena. *See* ACOUSTIC NOISE; ELECTRICAL NOISE. John A. Dutton

Systems in the real world. The use of deterministic, ordered mathematical models to describe behavior of systems in the real world is so familiar that only chaos need be discussed. One example of chaos is the evolution of life on Earth. Were this evolution deterministic, the governing laws of evolution would have had built into them anticipation of every natural crisis which has occurred over the centuries plus anticipation of every possible ecological niche throughout all time. Nature, however, economizes and uses the richness of opportunity available through chaos. Random mutations provide choices sufficient to meet almost any crisis, and natural selection chooses the proper one. *See* ORGANIC EVOLUTION.

Another example concerns the problem that the human body faces in defending against all possible invaders. Again, nature appears to choose chaos as the most economical solution. Loosely speaking, when a hostile bacterium or virus enters the body, defense strategies are generated at random until a feedback loop indicates that the correct strategy has been found. A great challenge is to mimic nature and to find new and useful ways to harness chaos. *See* IMMUNITY.

Another matter for consideration is the problem of predicting the weather or the world economy. Both these systems have much in common with Eq. (4), although their governing equations are certainly more complicated. Nonetheless, like a θ_n sequence of Eq. (4), the weather and the economy are chaotic and can be predicted more or less precisely only on a very short time scale. By recognizing the chaotic nature of the weather and the economy, it may eventually be possible to accurately determine the probability distribution of allowed events in the future given the present. At that point it may be asserted with mathematical precision that, for example, there is a 90% chance of rain 2 months from today. Much work in chaos theory seeks to determine the relevant probability distributions for chaotic systems. *See* WEATHER FORECASTING AND PREDICTION.

Finally, many physical systems exhibit a transition from order to chaos, as exhibited in Fig. 1, and much work studies the various routes to chaos. Examples include fibrillation of the heart and attacks of epilepsy, manic-depression, and schizophrenia. Physiologists are striving to understand chaos in these systems sufficiently well that these human maladies can be eliminated. *See* PERIOD DOUBLING. Joseph Ford

Chaos and periodicity. In systems where the number of possible states, none of which closely resembles any other one, is limited, approximate repetitions of earlier states must eventually occur. If such a system is not chaotic, the history following a near repetition of a state will nearly repeat the history following the original occurrence, so that near repetitions will continue to occur at regular intervals, and the system will vary periodically. If, then, a system of this sort is observed not to vary periodically, it must be chaotic. Absence of periodicity often affords the means by which chaos is most easily recognized, and it has sometimes served in place of sensitive dependence as a definition of chaos. For example, a pendulum in a well-behaved clock does not swing chaotically; a flag in a moderate steady breeze may flap chaotically.

Many real systems exhibit nonperiodic variations superposed on periodic oscillations. The periodic components of these variations are entirely predictable and, like the diurnal and seasonal variations of the Earth's weather, are often externally forced. In such systems, which are also considered chaotic, two rather similar states will eventually evolve into states differing as much as states chosen from a long sequence, at similar phases of the periodic

oscillations but otherwise randomly.

Attractors. Real-world systems generally entail some dissipative processes, and some external forcing that prevents them from coming to a standstill, like the clockwork that keeps a pendulum swinging or the wind that keeps a flag flapping. For such systems the set of states that can occur or be approximated again and again, after transient effects have subsided, is far more restricted than it would be if neither dissipation nor forcing were present; it is most drastically restricted when the system is not chaotic. This set of states is called the attractor of the system. Especially for simple systems, attractors are commonly represented by plane or multidimensional graphs in which each point represents a distinct state, the coordinates of the point equaling the values of some or all of the variables. For a chaotic system the graph is an infinite complex of curves or other manifolds, with finite gaps between any pair of manifolds, and the attractor is called a strange attractor. Edward N. Lorenz

Lorenz butterfly. A celebrated mathematical model exhibiting chaotic behavior was developed by E. Lorenz in the early 1960s. Seeking to understand limitations on the accuracy of weather prediction, Lorenz studied a mathematical model with just 12 differential equations representing atmospheric flow and, by calculating solutions to the equations with a digital computer, discovered that some of them were chaotic—they exhibited sensitive dependence on initial conditions. Failing to find an even simpler model of temporally dependent atmospheric flow that would clarify the mathematical origins of the chaos, Lorenz turned to two-dimensional convection in a vertical plane as modeled by a system of three differential equations that also exhibited chaotic solutions. Lorenz's 1963 article on the mathematical properties of this model became a classic in the numerical study of chaos. It provided an explicit example of chaotic behavior in a system derived from hydrodynamics, and thus seemed to dash hopes that computers would enable accurate large-scale atmospheric prediction for forecast periods of weeks or months. Indeed, Lorenz observed that an implication of his results was that the flap of a butterfly's wings could change the course of the weather a continent away.

Lorenz model and equations. The flow of a fluid heated below and cooled above, known as Rayleigh-Bernard convection, depends on the external temperature difference ΔT maintained on the two bounding surfaces. For small differences, the flow will be a slow regular overturning, while for large temperature gradients it will be intense and turbulent. The ordinary differential equations governing the amplitudes of the representation, X, Y, and Z, are calculated with Fourier techniques from the governing partial differential equations (9). Here, X' is the rate of change of X ($dX(t)/dt$), and the quantities Y' and Z' are similarly defined. All variables are dimensionless, with a Prandtl number σ, an aspect ratio b, and the Rayleigh number r proportional to ΔT. The nonlinear XY and XZ terms are the source of mathematical intractability and chaotic solutions.

$$X' = \sigma(Y - X) \tag{9a}$$

$$Y' = -XZ + rX - Y \tag{9b}$$

$$Z' = XY - bZ \tag{9c}$$

The stationary solutions describing steady flow are readily found by setting the rates of change, X', Y', and Z', all equal to zero. The stationary point at the origin $X = Y = Z = 0$ is always a solution, and is stable only for $r < 1$. A pair of additional stationary points appears for $r > 1$ at $X = Y = \pm$, $Z = r - 1$. For small values of r, solutions spiral rapidly into one of these two points. For $r > 13.96$, solutions with initial points near one of these stable points cross over and spiral into the other point and some isolated chaotic solutions appear. When r exceeds the critical value 24.74, all three stationary points are unstable and chaotic flow is observed for almost all initial conditions.

But even then, the flow is dissipative [as a consequence of the last terms in each of Eqs. (9)] and has the critical property that a volume V of initial points (that is, a three-dimensional set) is mapped asymptotically as time increases to a set of zero volume (a set of dimension less than three). To see this, observe that in Eq. (10) the volume of an initial set of points with volume V_0 is shown in Eq. (11). Thus,

$$\frac{1}{V}\frac{dV}{dt} = \frac{\partial X'}{\partial X} + \frac{\partial Y'}{\partial Y} + \frac{\partial Z'}{\partial Z} = -(\sigma + 1 + b) \tag{10}$$

$$V = V_0 e^{-(\sigma+1+b)t} \to 0 \quad \text{as} \quad t \to \infty \tag{11}$$

the initial volumes are mapped to sets of dimension less than volumes. But from a mathematical theorem, the attractor must be of dimension greater than two if the solutions are to be more complex than periodic orbits.

Chaotic attractor. Solving the nonlinear convection equations (9) with a computer, Lorenz compared solution trajectories in (X, Y, Z) space with those emanating from nearby initial points and discovered that, after a time, they separate from each other when r is large enough. Indeed, for the convection model described here, he found that if the Rayleigh number r exceeds the value 24.74, then the separation occurred eventually, no matter how close the initial points were. Furthermore, through numerical simulation Lorenz discovered that all solution trajectories in (X, Y, Z) space, regardless of their initial point, are attracted to a thin, curved structure where they are then trapped. This Lorenz attractor has two lobes resembling the wings of a butterfly and is called the Lorenz butterfly (**Fig. 2**).

Such attractors are identified as strange because their structure is vastly more complex than the stationary points or periodic orbits that constitute all of the mathematically possible attractors on a two-dimensional surface. The Lorenz attractor may be represented as a collection of nearly parallel sheets connected at the origin. Evolving trajectories can pass by each other on adjacent sheets without actually intersecting. The attractor can be described as resembling a Cantor set, and with advanced

methods its fractal dimension can be determined as 2.06 (where a plane would be of dimension 2).

As a solution trajectory approaches the chaotic attractor, it will first circle one of the paired stationary points a number of times and then pass near the origin on the way to circle the other stationary point in the pair for one or more times. The trajectory thus continues to oscillate between the two stationary points, circling away from each one for a number of times before switching to the other one.

This phenomenon leads to a separation of solutions that forecloses the possibility of extended predictability of chaotic systems. As two solutions with nearby initial points approach the attractor, they circle, and oscillate between, the two paired stationary points for a while, staying fairly close together and tracing out a similar trajectory. But then, one solution will continue to circle one stationary point while the other will move across the origin to orbit the other stationary point. From then on, the time history of the solutions will be different and the limit of numerical predictability has been exceeded (**Fig. 3**).

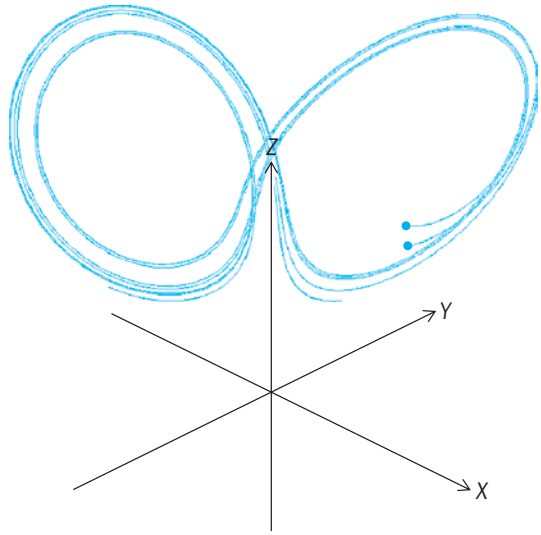

Fig. 3. Separation of solutions on the Lorenz attractor. The two trajectories starting from the nearby initial points indicated by the balls in the right lobe of the attractor stay together for several circuits around both lobes and then separate as they pass over the origin, with one proceeding around the left lobe, the other around the right. (*M. Takatsuka, GeoVISTA Center, Department of Geography, Pennsylvania State University*)

Stretching, tearing, and folding. The mechanism creating chaos in the Lorenz and other chaotic models is a combination of stretching, tearing, and folding mandated by the interaction of nonlinearity and the energy constraints on the flow. In the case of the Lorenz model, it can be shown that a quadratic form in X, Y, Z is constrained and that the flow is confined to a three-dimensional ellipsoid. But because of nonlinearity, the trajectories emanating from a small volume of initial points will diverge from each other and be stretched out into a sheetlike image expanding in two dimensions and contracting in one, thus satisfying the condition that initial volumes map to sets with zero volume as time increases. The images of the initial points are stretched farther and farther apart and are torn when they split to circle different stationary points. Nevertheless, the expanding image surface must eventually fold to stay within the boundaries prescribed by the energy constraint, mixing up the originally nearby trajectories.

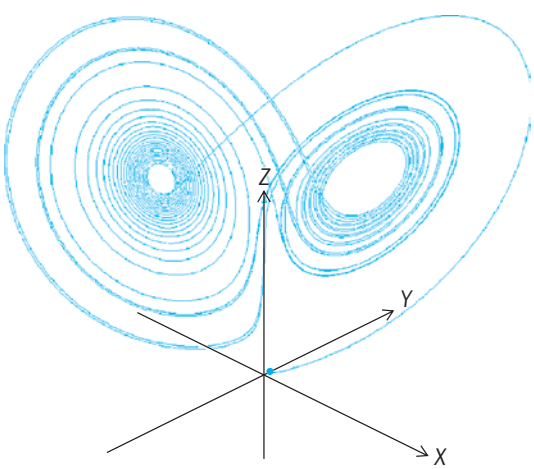

Fig. 2. Three-dimensional visualization of the Lorenz butterfly as traced out by the trajectory of the original simulation reported by Lorenz in 1963 ($\sigma = 10$, $b = 8/3$, $r = 28$) with initial conditions (X, Y, Z) = (0, 1, 0). The integration was performed with a fourth-order Runge-Kutta scheme. (*M. Takatsuka, GeoVISTA Center, Department of Geography, Pennsylvania State University*)

Butterflies and computers. Although the existence of chaotic behavior has been inferred for nearly a century, the detailed exploration of chaos has been made possible by numerical simulation with digital computers. Indeed, even today there is no way other than numerical simulation to determine whether the solutions to a given nonlinear differential system will exhibit chaotic properties for a specific range of its parameters. And thus chaos and computers go together, with computers essential to reveal the properties of chaotic systems.

As exemplified by the Lorenz butterfly, the mathematical structures of chaotic attractors are often elegant and sometimes haunting. But there is the other Lorenz butterfly, too—the one whose wings might change the course of weather events. The manifestations of that hypothetical butterfly appear in reality as differences in observed and actual states of the atmosphere and as inaccuracies inevitably introduced in trying to model an entire atmosphere in a finite computer model. The butterflies are thus in the observations and in the computer models of any chaotic system. Taking account of the effect of these virtual butterfly wings is a key challenge in the development of improved numerical methods of atmospheric prediction.

John A. Dutton

Some mathematical foundations in physics. To the uninitiated a gulf appears to exist between the ways that mathematicians and physicists study chaos. A primary difference between their modeling methods can be shown to result from the choice of size for a time-step (increment, Δt) in the integration of a system's equation of motion. Physicists generally

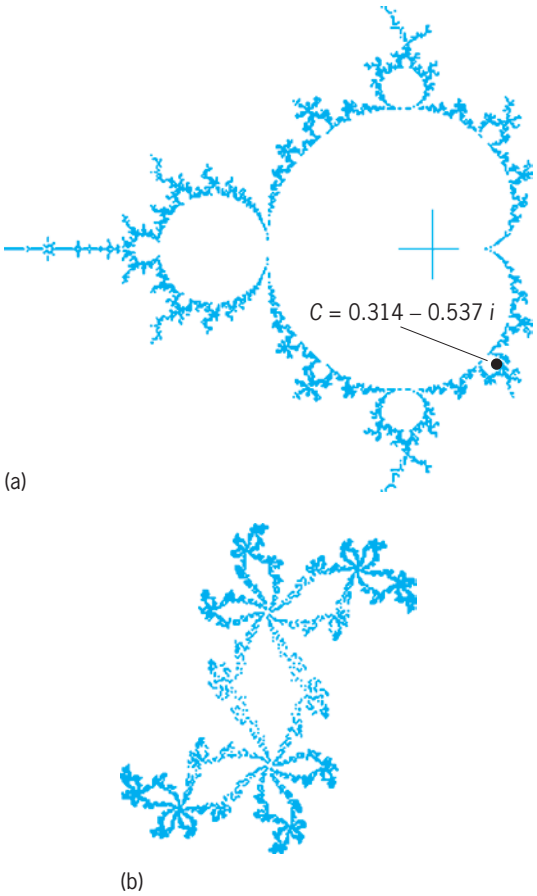

Fig. 4. Examples of fractal (self-similar) sets. (*a*) The Mandelbrot set in the complex plane, with origin indicated by the cross. A single point in the complex plane (in this case $C = 0.314 - 0.537\,i$, where $i = \sqrt{-1}$) generates (*b*) an associated Julia set.

prefer to describe a system in terms of state variables (such as displacement of a pendulum θ and its momentum p) that vary continuously with time t (that is, $\Delta t \to 0$); whereas mathematicians prefer to work with "maps," which involve the repeated application of one or more recursive equations to a set of one or more variables.

The maps known as Julia sets (**Fig. 4**) and studied by Benoit Mandelbrot are among the best-known examples of fractals, which have the property of self-similarity under a change of scale. Maps of this type are generated by iteration of the equation of $Z_{n+1} = Z_n^2 + C$, where C is a complex-number constant, with $Z_0 = C$. The set of points C in the complex plane for which the associated Julia set is bounded is called the Mandelbrot set (Fig. 4), and its boundary is also a fractal. (Details of the methodology are available from Web sites.) *See* COMPLEX NUMBERS AND COMPLEX VARIABLES; FRACTAL.

Hamilton's equations and maps. A connection between phase-space representations and their associated maps, which also exhibit self-similarity for chaotic systems, is realized by applying the methods of Hamiltonian mechanics when generating the differential equations of motion. Consider, for example, the Hamiltonian of Eqs. (12) for the pendulum

$$H(p, \theta) = \frac{p^2}{2} - c\cos\theta$$
$$\dot{p} = -\frac{\partial H}{\partial \theta} = -c\sin\theta \qquad (12)$$
$$\dot{\theta} = \frac{\partial H}{\partial p} = p$$

(obtained from considerations of the total energy of the system), in which some constants that are usually included in this expression have been set equal to unity, to leave the single adjustable parameter c. Here, $\dot{p} = dp/dt$ and $\dot{\theta} = d\theta/dt$, and the two time derivatives in Eqs. (12) are called Hamilton's canonical equations of motion. Taking the time derivative of $\dot{\theta}$ gives the familiar equation of motion for the

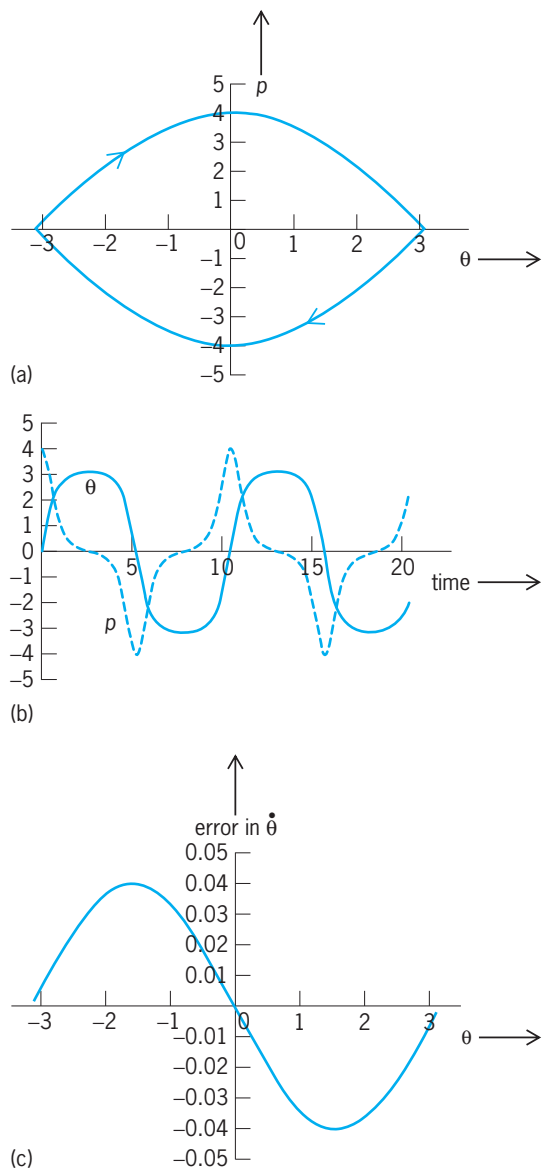

Fig. 5. Representations of pendulum motion corresponding to a periodic variation in angular position θ whose amplitude is just less than π radians (180°). (*a*) Phase-space representation (plot of momentum p versus θ). (*b*) Time plots of θ and of $p = \dot{\theta}$. (*c*) Error in $\dot{\theta}$ from use of the Euler algorithm with $\Delta t = 0.01$.

pendulum; that is, $d\dot{\theta}/dt = \ddot{\theta} = \dot{p} = -c \sin \theta$, yielding $\ddot{\theta} + c \sin \theta = 0$. *See* HAMILTON'S EQUATIONS OF MOTION; PENDULUM.

Transition from a "physics view" to a "math view" involves the use of approximations in finite-difference form, which cause the canonical equations to take on the form of Eqs. (13). Thus θ and

$$\begin{aligned} p_{n+1} &= p_n - c \sin \theta_n \Delta t \\ \theta_{n+1} &= \theta_n + p_{n+1} \Delta t \end{aligned} \tag{13}$$

p can be estimated for any value of n (corresponding to time $n\ \Delta t$) once the initial conditions at $n = 0$ have been specified. To describe the pendulum's behavior according to Eqs. (13), the graphical representation commonly employed is that which is referred to as phase space. For this low-dimensional case, it is simply a plot of p versus θ, an example of which is shown in **Fig. 5***a*. For a more complicated (higher-dimensional) system, such as the Lorenz butterfly, discussed above, a third variable is necessary.

For purpose of comparison, Fig. 5*b* shows the temporal variations of θ and p. Such time plots are less useful than phase plots when the motion is chaotic, as opposed to the highly nonlinear, yet nonchaotic case of Fig. 5. Chaotic motion can be generated by adding to Eq. (13) both a damping term proportional to p and a driving force. This case is characterized by the fractal nature of a plot called the Poincaré section, which is obtained by "strobing" the graph of p versus θ. In other words, instead of continuously graphing p versus θ, one "turns on the graph" at one point per period of the driving force.

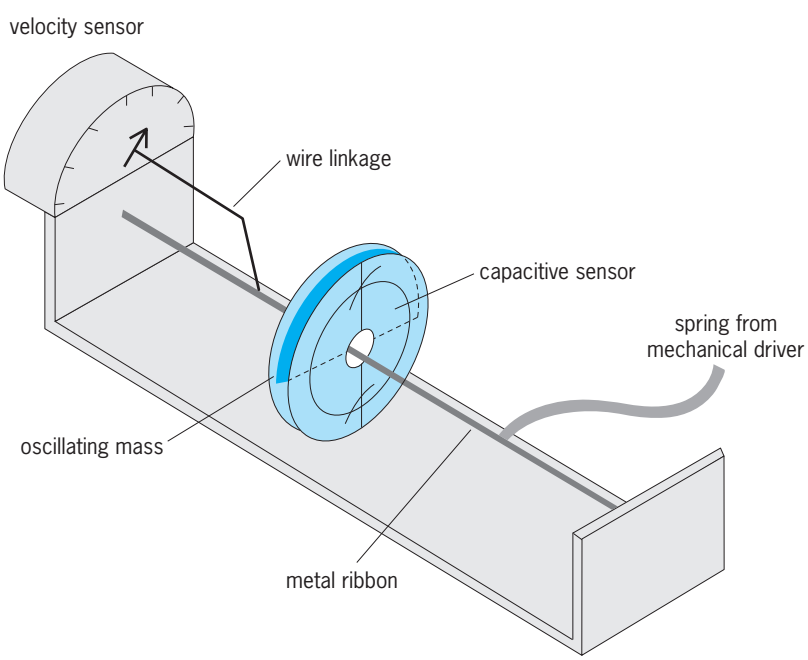

Fig. 6. Example of a pendulum in which gravity and a torsional ribbon act in opposition. The oscillating mass is a thin semicircular metal plate that is rigidly attached to the ribbon at its midpoint. The ribbon is fixed at each end but twists where the mass is attached, and torsional oscillations in the ribbon are excited by the spring. The oscillating mass is sandwiched between two stationary circular metal plates with center holes through which the ribbon passes, and together the three plates form a fully differential capacitive sensor to measure the angular position of the mass. The velocity sensor is a galvanometer operating in "reverse," and the wire linkage drives this galvanometer.

Figure 5*c* shows that the numerical integration algorithm due to Euler, discussed below, is not acceptable for the treatment of these problems. The maximum size of the error increases as Δt is made larger than the indicated value of 0.01.

Iteration and numerical integration subtlety. In Eqs. (13) the update order is important, and for the position update it is also important to observe that the iteration subscript on p is not n but $n + 1$. The importance of this choice becomes clear by noting Eq. (14) for

$$\begin{aligned} \ddot{\theta}(t) &\approx [\theta(2\Delta t + t) - 2\theta(\Delta t + t) + \theta(t)]/\Delta t^2 \\ &= -c \sin \theta(t') \end{aligned} \tag{14}$$

the approximation of $\theta(t)$, in which there are two choices for the time t'. The choice $t' = t$ is equivalent to Euler's method of numerical integration, which is unstable. The instability gives rise to errors such as those discussed above in Fig. 5. On the other hand, the choice $t' = \Delta t + t$, corresponding to p_{n+1} is equivalent to the "leap-frog" method. It is stable because of the symmetry that it brings to Eq. (14), and it is increasingly used instead of more sophisticated integration algorithms, such as fourth-order Runge-Kutta.

In addition to numerical simulation of chaotic pendulum motion, it is possible to remotely control actual hardware by means of an online interactive chaotic pendulum.

Examples in two dimensions. Meaningful characteristics of the pendulum result from setting $\Delta t = 1$ in Eqs. (13), yielding Eqs. (15), a circumstance that is

$$\begin{aligned} p_{n+1} &= p_n - c \sin \theta_n \\ \theta_{n+1} &= \theta_n + p_{n+1} \end{aligned} \tag{15}$$

surprising to some physicists. The resulting map is called the Chirikov or standard map. Its importance derives largely from the status of the pendulum as an archetype of chaos.

A similar map is obtained from consideration of a torsion-gravity pendulum, illustrated in **Fig. 6**. In the absence of torsional restoration, the oscillating mass would rest beneath the ribbon, whereas in the absence of gravity it would rest above the ribbon.

For this system, the Hamiltonian and associated canonical equations are given by Eqs. (16). Depending on the size of its single adjustable parameter b, Eqs. (16) relate to three important oscillator types: (1) the harmonic oscillator for $b \ll 1$, (2) the Duffing oscillator for $1 < b < 1.5$, and (3) the superconducting quantum interference device (SQUID) for large values of b. [Involving a "double-well" potential energy function, the Duffing oscillator was first studied in 1918 by the engineer Georg Duffing. Its chaotic motion is confined in phase space without having to

$$\begin{aligned} H(p, \theta) &= \frac{p^2}{2} + \frac{\theta^2}{2} + b \cos \theta \\ \dot{p} &= -\frac{\partial H}{\partial \theta} = -\theta + b \sin \theta \\ \dot{\theta} &= \frac{\partial H}{\partial p} = p \end{aligned} \tag{16}$$

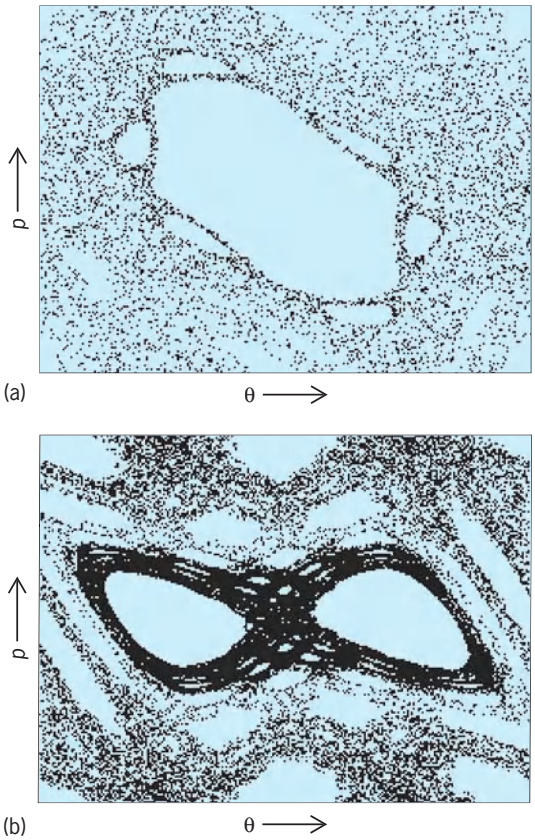

Fig. 7. Fractal geometry attractors generated by computer. (*a*) Attractor generated using the Chirikov map defined by Eqs. (15), with $c = 1.6$ and initial values $p_0 = \theta_0 = 2$. (*b*) Attractor generated using the torsion-gravity pendulum map defined by Eqs. (17), with $b = 1.6$ and $p_0 = \theta_0 = 0.1$. In both plots there are periodic boundary conditions on both p and θ ($\pm\pi$).

use the periodic boundary conditions (modulo 2π) required of the pendulum. Chaotic motion of a pendulum requires motion past the vertical, which is impossible for a bob supported by a string; this in turn involves "winding" through multiple rotations.] *See* SQUID.

Approximation of canonical equations (16) in finite-difference form and setting $\Delta t = 1$ gives Eqs. (17). Examples of self-similar fractal structures

$$\begin{aligned} p_{n+1} &= p_n - \theta_n + b \sin \theta_n \\ \theta_{n+1} &= \theta_n + p_{n+1} \end{aligned} \tag{17}$$

generated by the two-dimensional maps defined by Eqs. (15) and Eqs. (17) are provided in **Fig. 7**.

Examples in one dimension. A one-dimensional map known as the logistic equation and defined by Eq. (18) is better known to mathematicians than the

$$x_{n+1} = cx_n(1 - x_n) \tag{18}$$

previous two-dimensional maps. Its period-doubling route to chaos is illustrated by **Fig. 8***a*. To generate this map, initial values for c and x were chosen at $c = 2.8$ and $x = 0.65$. For each value of c, four iterations were performed to allow convergence to any limit cycle that might exist for that value of c. (When a nonchaotic limit cycle is reached after the settling of transients, points generated by the iterative equation or equations are repetitive. The number of steps n before the sequence repeats designates the order of the limit cycle. In a physical system such as a pendulum, steady-state motion corresponding to a limit cycle manifests itself in the Poincaré section as a set of phase-space points whose total number is n, assuming one point per cycle of the driving force. Denoting the drive frequency by f, the limit-cycle system response is at f/n, that is, a subharmonic of order n.) The value of x following the fourth iteration was subsequently plotted versus the parameter c, and c was then incremented by 0.0001 and the process repeated until the maximum value of 4 was reached. For $c < 3$, there is a single (stable) value for x. For $c = 3.2$ as an example, the previous singlet has branched into two values; that is, bifurcation has occurred. At $c = 3.5$, the consequence of an additional bifurcation can be observed, so that "two begets four," this is part of the "period-doubling route" to chaos that finally results in the vicinity of $c = 3.57$. *See* PERIOD DOUBLING.

Figure 8*b* was generated with an entirely different one-dimensional map that is concerned with solving Johannes Kepler's transcendental equation (19) for

$$E = M + \varepsilon \sin E \tag{19}$$

the position of a planet according to specified time

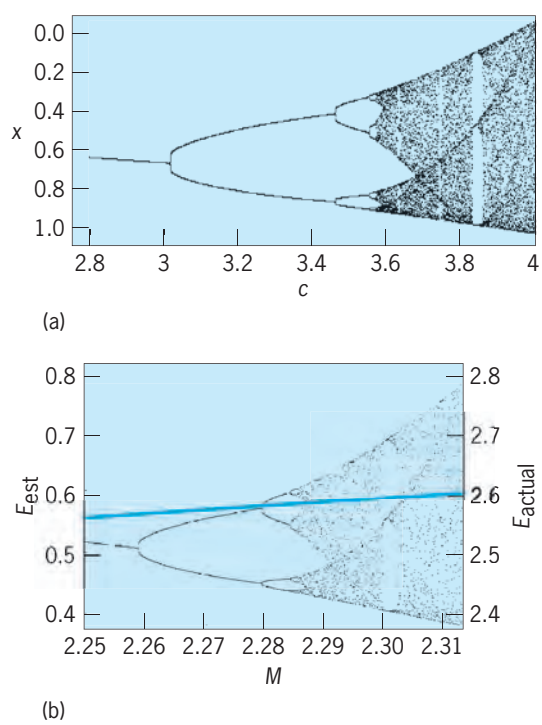

Fig. 8. Examples of the period-doubling route to chaos for two vastly different iterative expressions. (*a*) The logistic map, defined by Eq. (18). (*b*) The Chebyshev root-finding algorithm applied to Kepler's equation, defined by Eqs. (20) and (21) with eccentricity $\varepsilon = 0.55$. The colored line gives the actual value of E (referenced to the right-hand scale, E_{actual}), to which the iteration quickly converges when the starting value of E is located within the known region of convergence. The spurious output from the algorithm is referenced to the left-hand scale (E_{est}).

in orbit. The eccentric anomaly E (giving the angular position of the planet) is obtained for a given value of the mean anomaly M (proportional to the time) for a specified eccentricity ε. The root-finding technique shown in Fig. 8*b* for solving Eq. (19) is that of an "improved" Newton method known as the Chebyshev algorithm, which accounts for the curvature as well as the slope of the function under study. The Chebyshev algorithm is defined by Eq. (20), where,

$$E_{n+1} = E_n - (f/f')[1 + {}^1\!/_2(f/f')(f''/f')] \qquad (20)$$

in the case of its application to Eq. (19), f, f', and f'' are defined by Eqs. (21).

$$\begin{aligned} f &= E_n - M - \varepsilon \sin E_n \\ f' &= 1 - \varepsilon \cos E_n \\ f'' &= \varepsilon \sin E_n \end{aligned} \qquad (21)$$

When the starting value of E is located within the known region of convergence, the iteration converges quickly to the proper single-valued solution shown by the colored line in Fig. 8*b*, referenced to the right-hand scale. The parameter choices shown do not produce this convergence and result instead in the indicated spurious (nonphysical) doubling route to chaos. A noteworthy characteristic of both cases in Fig. 8 is the nonchaotic solution that appears as an "island" in the midst of chaos (period-3 limit cycle, $c = 3.832$ for Fig. 8*a*). *See* KEPLER'S EQUATION.

Feigenbaum constant. As the parameter c is increased in the logistic map given by Eq. (18), there is a progressive series of bifurcations (after transients are permitted to settle following each step in c). A fundamental constant emerges from this progression and was discovered by Mitchell Feigenbaum. Designating by c_n the points at which a period-2^n cycle appears, it can be shown that the Feigenbaum constant is given by Eq. (22).

$$\delta = \lim_{n\to\infty} \frac{c_{n+1} - c_n}{c_{n+2} - c_{n+1}} = 4.6692\ldots \qquad n = 1, 2, \ldots \qquad (22)$$

Claim for the universality of the constant δ in Eq. (22) is supported by the graph of Fig. 8*b*. The points where period-2^n cycles in M_n occur have not been carefully estimated; however, it appears that the Feigenbaum constant could also be generated from this map. The first three values of $M_1 = 2.25896$, $M_2 = 2.27903$, and $M_3 = 2.28368$ yield $\delta_{\text{1st est., Chebyshev}} = 4.32$; whereas the first three values from the logistic map yield $\delta_{\text{1st est., logistic}} = 4.23$.

Randall D. Peters

Atmospheric prediction. Contemporary large-scale weather prediction takes advantage of computer simulation of atmospheric behavior as modeled by a set of nonlinear partial differential equations describing the flow of the atmosphere and its thermodynamic properties. There are two sources of error: (1) the observations of the state of the atmosphere at the start of the prediction cycle may not be representative or accurate; (2) the equations do not represent all aspects of atmospheric behavior accurately. On the basis of experience with simpler nonlinear systems, one might expect that such errors in a nonlinear system would produce chaotic departure of the computer predictions from what actually transpires in the atmosphere.

Indeed, the computer models can be used to study the effect of errors in the initial conditions. When predictions are computed from slightly different initial atmospheric states, they resemble those of the Lorenz model of convection: the solutions remain similar for a time and then begin to depart from each other until eventually they are uncorrelated. Studies show that initial errors in model-scale phenomena double in about two days, but that the models demonstrate some predictive skill for large-scale features for a week or more. The considerable improvements in numerical weather prediction in the past three decades can be attributed to improved atmospheric observations, to increased model resolution made possible by increasing computer power, and to improvements in the formulation of the models and the smaller-scale processes that must be represented statistically. Such improvements push the effect of errors farther into the future, delaying but not eliminating the eventual effects of chaos and the inevitable loss of predictability.

Now a new strategy for improving numerical prediction takes advantage of increasing computer power to predict the accuracy of the forecast itself. Rather than computing a single forecast at increased resolution, the major numerical prediction centers compute a number of predictions, each starting from a set of initial conditions slightly perturbed from those observed. Thus an ensemble of forecasts is computed, suggesting confidence when and where the forecasts in the ensemble agree and suggesting uncertainty when and where they vary significantly.

This approach leads to an even more sophisticated strategy: Determine from the ensemble of forecasts where special observations might be taken to reduce uncertainty and then schedule satellite observations or dispatch piloted or remotely piloted aircraft to obtain special observations to reduce the uncertainty. Such an approach may have favorable benefit to cost ratios in the case of severe weather. Indeed, it has been employed for more than 50 years by sending aircraft to take observations in hurricanes or typhoons that might threaten inhabited regions.

John A. Dutton

Climate. Climate is sometimes identified with the set of all states of the climate system—the system consisting of the atmosphere and the upper portions of the underlying ocean and land surfaces—that may reasonably be expected to occur. It is thus identifiable with the attractor of the climate system. This system is dissipative and is forced by solar heating, and it is almost certainly chaotic; although it is not feasible to examine the actual climate system directly for sensitive dependence, mathematical models of the system, including some used in operational weather forecasting, invariably indicate sensitive dependence, while such features as migratory storms that cross the oceans and continents are not observed to occur in any periodic sequence.

Chaos can enrich the variety of weather patterns that make up the climate. Even though the weather has been observed for centuries, it is reasonable to expect something that is not known to have happened before—a record high or low temperature at some location, or a record rainfall that produces a record flood. Without chaos the weather would vary periodically, and records would simply be tied rather than broken.

Changes in external forcing will change the behavior of the climate system, thereby producing a new attractor, and hence a climatic change. Furthermore, if the state of the system normally fluctuates within one portion of the attractor for a long time and then enters into another portion, the transition will be perceived as an additional climatic change, perhaps indistinguishable from one externally produced. It is therefore sometimes preferable to identify the present climate with the portion of the attractor within which the state is presently fluctuating, thereby allowing for internally produced climatic changes.

Regardless of any chaotic weather variations occurring within the tenure of one climate, it is possible that the variations of the climate itself are chaotic. If this is so, only the periodic components of distant climatic changes can be predicted. Mathematical models of climate have not reached the degree of perfection needed to confirm or reject such a hypothesis, and recourse to observations is indicated. Where weather observations are absent or insufficient, proxy data, such as records of tree rings or oxygen isotope ratios, which are supposed to vary with climatic elements such as temperature, can be useful. There are indications of periodic variations with superposed nonperiodic variations; it is the latter that may be unpredictable. Some of the periodic variations appear to be associated with predictable changes in the Earth's orbital parameters. Some nonperiodic variations may result from irregular external forcing—perhaps nonperiodic volcanic activity. Others may be internally produced, and chaotic. *See* CLIMATIC PREDICTION; CLIMATOLOGY; DENDROCHRONOLOGY.

Since climate variations are unpredictable by deterministic means for periods of decades or centuries in advance, prediction efforts must focus on the statistical characteristics of the climate and its variations as determined by both observations and computer simulations. One possibility lies in identifying a climate attractor that is a multidimensional analog of the attractor for the Lorenz convection model described above. If such an attractor was found, and if the present climate could be located on the attractor, then it might be possible to infer statistical properties expected for the climate over a range of periods into the future. As computer power and capability increase, another possible approach might involve developing statistics from ensembles of computer climate simulations, as described above for weather prediction, that would give an indication of the confidence to be placed in various expectations about climate variations.

To use computer simulation in studying climate variations, it is necessary that the climate variations observed in such models resemble those of the actual climate. To achieve that, the climate simulation models must include realistic representations of atmospheric interactions with the ocean, the land surface, and the biosphere. The longer the period of interest, the more important such interactions presumably are, leading to the conclusion that realistic study of longer-term climate variations requires integrated simulations of the entire Earth system. Edward N. Lorenz; John A. Dutton

Electronic circuits. Chaos is a ubiquitous phenomenon which is a subset of the dynamical behavior of nonlinear systems. In an electronic context, a nonlinear system is one where the output is not simply proportional to the input, and the response to a collection of excitations is not simply the sum of the individual responses to each excitation taken separately. This criterion embraces a large subset of electronic systems, including those in which there is multiple-valued behavior. Many chaotic systems observed in other disciplines are modeled well by electronic circuits. The dynamics are then susceptible to experimental analysis in a way that would be difficult in the source discipline. *See* NONLINEAR PHYSICS.

It is also possible to use electronic chaotic circuits for the direct experimental study of some arcane phenomena suggested by the rigorous mathematical analysis of nonlinear dynamical systems. Many such investigations have been reported. Intermittencies can be observed, trapping and long-lived chaotic transients are easily demonstrated, extensive catalogs of coexisting attractors can be found empirically, and synchronization and entrainment of chaotic systems can be explored, with applications to secure communications.

In analog electronic circuits, chaos can be found in overload situations, where a linear circuit is driven beyond the capacity of its power supply; in switch-mode power supplies, where energy-efficient control is achieved by using dependent switches in association with a linear circuit; and in resonant circuits which are nonlinear. These resonant circuits divide into a group wherein the natural frequency is amplitude dependent (modeled by Duffing-like equations) and also a group which contain amplitude-dependent damping, modeled by Van der Pol–like equations. *See* DIFFERENTIAL EQUATION; ELECTRONIC POWER SUPPLY.

Categories of electrical chaotic system include purely electronic; electromechanical, in which the mechanical element is nonlinear whereas the coupled electronic circuit is linear (for example, ultrasonic cleaners); and control systems generally, in which either or both the control circuit and the system being controlled are nonlinear. There is an important class of chaotic system comprising coupled oscillators; a stable oscillator necessarily has a limit cycle and is therefore nonlinear. Two such oscillators (when coupled) may pull in and out of lock in a chaotic manner.

Analyses of chaos in continuous systems using a digital computer are sometimes grossly in error owing to the extreme sensitivity to initial conditions. However, all real electronic systems are described by

continuous variables plus unavoidable added noise. Real digital systems are modeled as accurately as their analog equivalents when the number of distinguishable states is greater than the number of distinguishable levels in an analog version containing unavoidable noise. Digital filters containing registers which overflow exhibit chaos. Chaos is also found in systems containing phase detectors, owing to the ambiguity in phase angle modulo 360°, for similar reasons. *See* DIGITAL FILTER; PHASE-LOCKED LOOPS.

Chaos is found also in digital systems having more than one autonomous timing element or clock; it is therefore endemic in asynchronous parallel computation. Chaos is suspected to occur in computer networks, at the physical data level. *See* CONCURRENT PROCESSING.

Chaos in circuits is modeled reasonably well by circuit simulation packages that deal with nonlinearities in the device models by solving the dynamical equations by using numerical integration. However, the class of originally linear circuit analysis software packages that handle nonlinear circuits by methods such as harmonic balance fail completely to model chaotic dynamics. David J. Jefferies

Astronomy. The methods of nonlinear dynamics have been applied to a large number of problems in astronomy. In particular, it is clear that chaos has played an important role in determining the dynamical structure and evolution of the solar system.

The Kirkwood gaps in the asteroid belt are now thought to be chaotic regions of space from which asteroids have been removed by the nonlinear gravitational perturbations of Jupiter. An asteroid in such a region can undergo large changes in its orbital elements which can cause it to become planet-crossing; the asteroid will eventually be removed from the region by a close approach to the planet. The existence of prominent gaps at the 1/3, 1/2, and 2/5 Jovian orbit-orbit resonances is consistent with the large chaotic regions which have been discovered at these locations by full numerical integrations and the use of algebraic mappings. The 2/3 resonance is associated with the stable Hilda group of asteroids and a much smaller chaotic zone. *See* ASTEROID.

Chaos has also been used to explain the delivery of one chondrite class of meteorites from the asteroid belt to the Earth. The 1/3 Kirkwood gap at 2.5 astronomical units has been identified as the source region for these meteorites, and studies of the dynamical behavior of asteroid collision fragments at the resonance show that perturbations by the outer planets can cause the orbits to become chaotic and achieve the high eccentricities necessary to become Earth-crossing. The time scales for this process are consistent with the ages of these chondrites determined by the length of time they have been exposed to the cosmic-ray environment of the solar system. *See* METEORITE.

The Saturnian satellite Hyperion exhibits chaotic behavior in its spin motion. Most natural satellites are locked in a 1/1 spin-orbit resonance where, as a result of tidal evolution, the orbital period of the satellite is equal to its spin period. However, Hyperion is distinctly nonspherical, and it has a large orbital eccentricity which is fixed by its 4/3 orbit-orbit resonance with the satellite Titan. As a result, Hyperion is in a chaotic spin state, with its rotational period changing constantly. Additional instabilities imply that its pole position may also have a chaotic behavior and that Hyperion is, in fact, tumbling. *See* SATURN.

The semimajor axis of a natural satellite can change significantly over the age of the solar system because of the tide that it raises on the parent planet. Consequently Kepler's third law implies that pairs of satellites can evolve into and out of resonance. This is an example of a dissipative dynamical system, since energy is lost in the form of tidal friction. Nonlinear dynamical studies have proved useful in determining the orbital histories of the satellite systems of the outer planets.

An 850-million-year numerical integration of the orbit of Pluto has shown some evidence of chaotic motion, although the chaos is unlikely to disrupt Pluto's stable 3/2 resonance with Neptune over the lifetime of the solar system. Long-term numerical integrations have been used to investigate the stability of the solar system over time scales comparable with a significant fraction of its age. Results show that the planets' orbits are chaotic, although there are no dramatic changes over several billion years. In this case the existence of chaos implies that a small error in the measurement of a planet's position will propagate exponentially such that its orbital location cannot be calculated for arbitrary times in the future. The evolution of a planetary orbit is connected with the evolution of the orientation of its spin axis (obliquity). In the case of Mars the obliquity varies chaotically from 0° to 60°, and this has important implications for climate change on the planet. *See* MARS; PLUTO; SOLAR SYSTEM.

Chaos has also helped solve problems in other branches of astronomy, including galactic dynamics and studies of stellar oscillations, and the number of applications of chaos to astronomy is expected to increase. Carl D. Murray

Acoustics. Chaos physics has shed new light on the old problem of noise, a term which has its origin in acoustics. Noise had been thought to be the outcome of interacting processes for which a deterministic formulation and description is no longer feasible because the number of such processes is extremely large. However, chaos physics has shown that noise can be present in systems with only a few interacting processes, for which a deterministic formulation can easily be written down explicitly. An example of such a system is a driven nonlinear oscillator, for example, an electric circuit consisting of only a resistor, an inductance, and a varactor diode (as the necessary nonlinear element) in series, driven by a sinusoidally varying voltage.

A chaotic (irregular, never repeating) sound wave can be produced when the chaotic voltage across this diode is fed to a loudspeaker. But systems of this kind are not considered to be genuine acoustically chaotic systems. Genuine acoustic chaos is found

in ultrasonics, thermoacoustics, musical acoustics, speech, and hearing. In ultrasonics it arises in the breakdown of liquids exposed to high-intensity sound, a phenomenon called acoustic cavitation. At too high an acoustic intensity the liquid ruptures to form bubbles or cavities (almost empty bubbles). They vibrate in the sound field that originally generated them, and emit sound. The total sound output of the cavities at high sound input (of a pure tone) leads to noise with a broadband spectrum. *See* CAVITATION.

A single sound frequency is thus transformed into an almost infinite number of neighboring sound frequencies. To study this transformation, an experiment has been conceived in which the sound output is monitored as the sound intensity is raised by digitizing the output of a hydrophone and storing the samples in a memory. Power spectra calculated from the stored data give the frequency content of the sound output. A period-doubling route to chaos is observed (**Fig. 9**). When the excitation level is raised, spectral lines appear exactly between the old lines by period doubling (frequency halving) until a broadband spectrum (chaos) is reached. Thus, acoustic cavitation has been identified as a chaotic system.

Because the cavities set into oscillation are the source of sound generation, their motion should become chaotic along with the sound output. Holographic observations of acoustically driven bubbles showed that they undergo period doubling to chaos in step with the acoustic output from the liquid. **Figure 10** shows two series of images reconstructed from high-speed holograms taken at 23,100 holograms per second. The black objects on the speckle background are the cavitation bubbles that undergo period-doubling oscillations to chaos.

Sound can be generated by heat—a fact known to glass blowers. When a gas-filled tube, closed at one end and open at the other, is heated at the closed end, the gas column starts to vibrate. Alternatively the open end may be cooled. When this self-excited oscillation is disturbed by another sound wave, the interaction may lead to chaotic sound waves.

In musical acoustics, signs of chaos are observed in string instruments, woodwind instruments (recorder, clarinet), the trumpet, and gongs (vibrated shells). Musicians, however, try to avoid the occurrence of bifurcations when playing their instruments. An example is the so-called wolf note that tends to appear in playing the cello. It is a quasiperiodic oscillation with two incommensurate frequencies that stem from the string and the bridge resonance. Oscillations of this type are also observed in woodwind instruments, there called multiphonics. Subharmonic sounds and chaos have also been found, for example, in the clarinet. *See* MUSICAL ACOUSTICS.

The human speech-production process is intrinsically nonlinear. To produce voice sounds the vocal cords are set into nonlinear vibration through the air flow between them. The fundamental frequency, the pitch, is determined by the length, mass, and tension of the vocal cords. There is no doubt from present knowledge of nonlinear oscillatory systems that the vocal cord system, and thus the human speech-production system, is susceptible to chaotic sound production. *See* SPEECH.

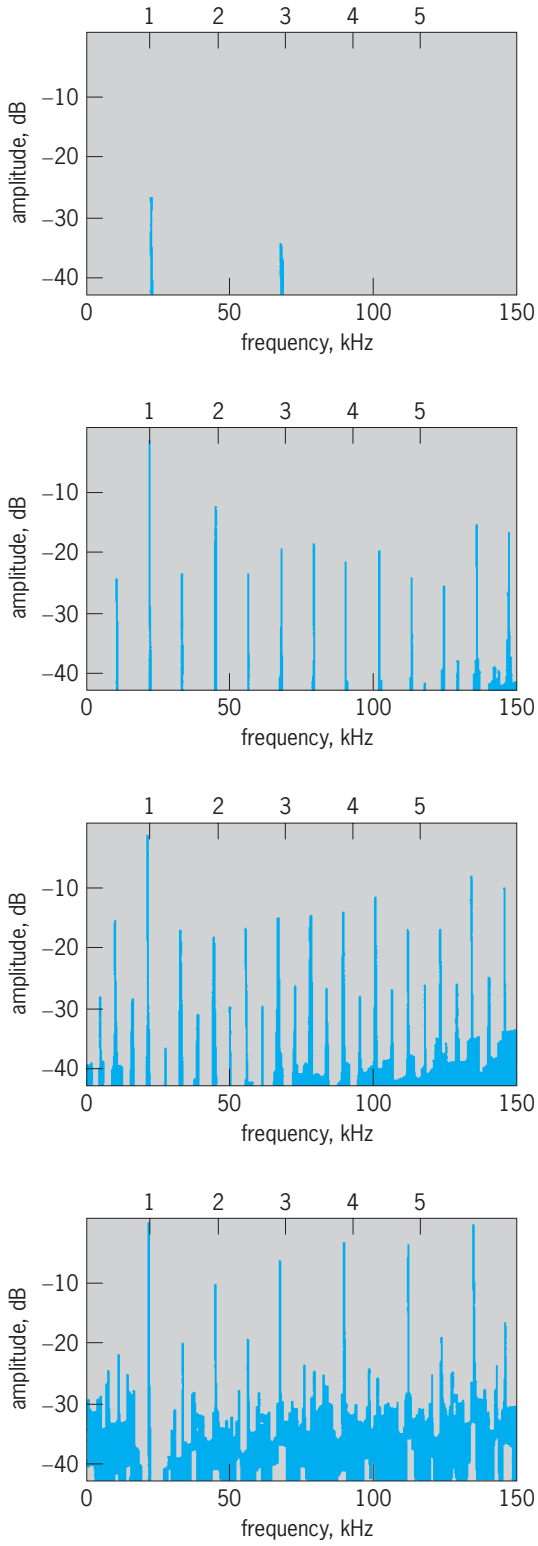

Fig. 9. Power spectra of acoustic cavitation noise at four different excitation levels, showing the successive filling of the spectrum by period doubling (frequency halving) when the sound input intensity is raised. Numbers at tops of the graphs indicate multiples of input frequency (harmonics).

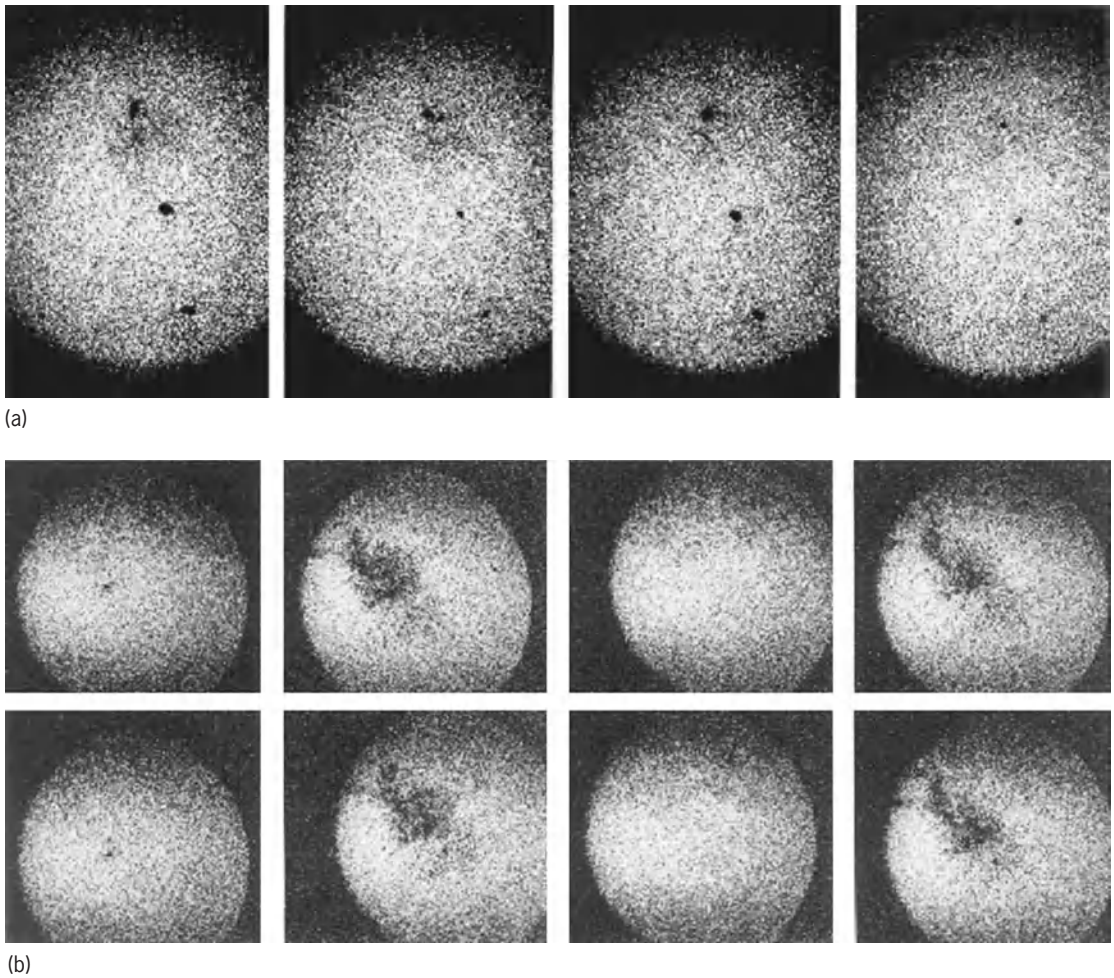

Fig. 10. High-speed holographic images of a bubble field driven at 23.1 kHz. One hologram was taken per period of the driving sound field. The bubble field oscillates as a whole periodically with the indicated period. (*a*) Period 2. (*b*) Period 8, at higher sound intensity.

The ear is a very complex sound receiver with remarkable properties, for instance in frequency selectivity. They stem from nonlinearities and active filtering with feedback, that is, filtering with amplification of sound. Such systems may become unstable and start to oscillate themselves. The ear thus also acts as a sound emitter rather than a pure receiver. This may happen in reponse to a sound wave or spontaneously, a phenomenon called objective tinnitus. Many people suffer from this property of the ear, as they hear the emitted sound. *See* HEARING (HUMAN); HEARING IMPAIRMENT. Werner Lauterborn

Atoms. A challenging problem is the behavior of quantal systems whose nonlinear, nonintegrable classical counterparts involve irregular (chaotic) trajectories, a subject commonly called quantum chaos. Because basic atomic structure is determined by the Coulomb electrostatic force, accurate quantal, semiclassical, or classical numerical calculations are possible. High-resolution atomic spectra, and how they are influenced by externally applied fields that rival the intraatomic Coulomb force binding a weakly bound atomic electron, can be recorded by using lasers or other sources and compared with theory.

In atomic three-body systems such as the helium atom, understanding and classifying the classical orbits, most of which are chaotic, is challenging enough. Quantizing them is a problem that has defied many researchers since the beginning of quantum theory early in the twentieth century. Thus, it is notable that the semiclassical quantization of the helium atom (in states with zero total angular momentum) has been accomplished.

Most atomic studies of quantum chaos have involved a weakly bound electron in the hydrogen atom (principal quantum number $n_0 \gg 1$) exposed to a strong, externally applied force F_{ext} that may be static or time varying. For $F_{\text{ext}} = 0$, the nonrelativistic, classical trajectories are regular, keplerian ellipses with angular frequency w_{at} for bounded motion and hyperbolas for unbounded motion; the corresponding quantal spectrum consists of the discrete (Bohr) levels with total energy E proportional to n_0^{-2} for $E < 0$ and a continuum of levels for $E > 0$. (Rydberg atoms, which have one of their electrons highly excited and approximate hydrogen, are also used.) *See* RYDBERG ATOM.

Static external fields. Hydrogen in a static magnetic field (B) is the most studied example. The linear

(in B) paramagnetic interaction leads to regular classical dynamics and to the Zeeman effect of textbook quantum mechanics. As B increases, however, the nonlinear (quadratic in B) diamagnetic interaction becomes dominant, and irregular (chaotic) electron trajectories begin to appear in classical numerical computations. Experimental studies of the photoabsorption spectrum of hydrogen atoms and of helium, lithium, or other Rydberg atoms (via laser-driven transitions from low-lying quantal states to strongly perturbed states near $E = 0$) revealed a complicated quantal spectrum whose changing statistical properties reflect the onset of classically chaotic trajectories. Series of sharp spectral features observed with lithium Rydberg atoms for magnetic field strengths and energies where the classical (hydrogenic) orbits are all chaotic were well reproduced by quantal numerical calculations for hydrogen. Use of scaled energy spectroscopy, which exploits scaling relations derived for the classical dynamics of the hydrogen atom in the magnetic field and applies whether or not the electron's motion is chaotic, simplifies comparisons with semiclassical theory if the photoabsorption spectrum is recorded with both the laser photon energy and B being varied such that the scaled energy remains constant. Unstable classical periodic orbits (or even just orbits that close on themselves near the nucleus) of ever longer period, which proliferate at bifurcations specified by the theory of hamiltonian dynamics, were shown to correspond in a direct way with important spectral features observed with finite resolution. *See* DIAMAGNETISM; PARAMAGNETISM; ZEEMAN EFFECT.

Examples of even more complicated situations for which classical scaling relationships have also been exploited are highly excited hydrogen or hydrogen-like Rydberg atoms in the presence of either crossed or parallel electric and magnetic fields. Photoabsorption spectra up to and beyond $E = 0$ (well into the domain of classical chaos) have been recorded experimentally and compared with results of classical, semiclassical, and quantal theory.

Time-varying external fields. For an $F_{\text{ext}}(t)$ produced by a microwave electric field with angular frequency ω and amplitude F, exchange of energy between it and the atom can lead to deexcitation to lower bound states or excitation to higher bound states or the $E > 0$ continuum (that is, ionization). (Ionization can require net absorption of many tens or even hundreds of microwave photons, which signals a very high order quantal multiphoton process.) The nature of the dynamics depends strongly on two ratios, the scaled frequency, $\omega/\omega_{\text{at}} = n_0^3\omega$, and the scaled amplitude, $F_{\text{at}} = n_0^4 F$, where F_{at} is the Coulomb electric field at the Bohr (circular) orbit for initial state n_0. If the (scaled) threshold fields $n_0^4 F$ (10%) that produced an ionization probability of 10% after the atoms were exposed to about 300 microwave field oscillations are measured and their dependence on $n_0^3\omega$ is graphed (**Fig. 11**), locally higher thresholds show up as bumps near some rational-fraction values (for example, 1/2, 1/1, and 2/1). These peaks were successfully explained via the stabilizing influence

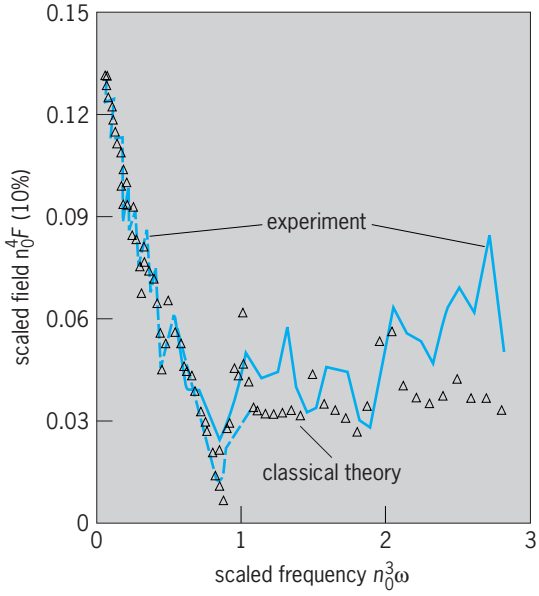

Fig. 11. Dependence of the scaled threshold field n_0^4F (10%) that produces an ionization probability of 10% after weakly bound hydrogen atoms are exposed to about 300 microwave field oscillations, on the scaled frequency $n_0^3\omega$.

of classical, nonlinear trapping resonances. At one extreme, as $n_0^3\omega \to 0$, the measured threshold fields approach, as expected, the value $n_0^4 F_0 \simeq 0.13$ for a static field. As $n_0^3\omega$ rises to near 1, classical, numerical Monte Carlo calculations (Fig. 11) that simulated the experiment reproduce the experimental curve (as do quantal calculations), but there are important, local exceptions. For example, a bump near $n_0^3\omega = 1.3$ is a quantal effect caused by the stabilizing influence of a (scarred) so-called separatrix state. (Broadband noise added to the driving field to decrease its coherence was shown to be a new spectroscopic technique for finding such states experimentally.) When $n_0^3\omega \gtrsim 2$, however, the systematic rise of experimental thresholds above those of classical calculations confirmed a prediction of quantum theory and numerical calculations that quantal interference effects would tend to suppress the effects of classical chaos for high-scaled frequencies. *See* ATOMIC STRUCTURE AND SPECTRA; MONTE CARLO METHOD; NONRELATIVISTIC QUANTUM THEORY.

Peter M. Koch

Bibliography. K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos: An Introduction to Dynamical Systems*, Springer-Verlag, 1996; J. R. Buchler and H. E. Kandrup (eds.), *Stochastic Processes in Astrophysics*, 1993; G. Casati and B. Chirikov (eds.), *Quantum Chaos: Between Order and Disorder*, 1995; L. O. Chua and M. Hasler (eds.), Special issues on chaos in nonlinear electronic circuits, *IEEE Trans. Circ. Sys.*, vol. CAS-40, no. 10 and 11 in CAS-I and no. 10 in CAS-II, 1993; F. Diacu and P. Holmes, *Celestial Encounters: The Origins of Chaos and Stability*, Princeton University Press, 1996; M. C. Gutzwiller, *Chaos in Classical and Quantum Mechanics*, 1990; R. C. Hilborn, *Chaos and Nonlinear Dynamics: An Introduction for Scientists and Engineers*, 2d ed.,

2000; W. Lauterborn and J. Holzfuss, Acoustic chaos, *Int. J. Bifurcations Chaos*, 1:13–26, 1991; E. N. Lorenz, Deterministic nonperiodic flow, *J. Atm. Sci.*, 20:130–141, 1963; E. N. Lorenz, *The Essence of Chaos*, University of Washington Press, 1993; B. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman, New York, 1977; T. Mullin (ed.), *The Nature of Chaos*, 1993; E. Ott, *Chaos in Dynamical Systems*, 2d ed., 2002; H.-D. Peitgen, H. Jürgens, and D. Saupe, *Chaos and Fractals: New Frontiers of Science*, 2d ed., Springer-Verlag, 2004; R. Peters, Chaotic pendulum based on torsion and gravity in opposition, *Amer. J. Phys.*, 63(12):1128–1136, 1995; I. Peterson, *Newton's Clock: Chaos in the Solar System*, W. H. Freeman, 1993; D. Ruelle, *Chance and Chaos*, Princeton University Press, 1991; M. Schroeder, *Fractals, Chaos, and Power Laws*, 1991.

Chaparral

A vegetation formation characterized by woody plants of low stature (3–9 ft or 1–3 m tall), impenetrable because of tough, rigid, interlacing branches, with small, simple, waxy, evergreen, thick leaves. The term refers to evergreen oak, Spanish *chapparo*, and therefore is uniquely southwestern North American. This type of vegetation has its center in California and occurs continuously over wide areas of mountainous to sloping topography. The Old World Mediterranean equivalent is called maquis or macchie, with nomenclatural and ecological variants in the countries ranging from Spain to the Balkans. Physiognomically similar vegetation occurs also in South Africa, Chile, and southwestern Australia in areas of Mediterranean climates, that is, with very warm, dry summers and maximum precipitation during the cool season.

The floras of these five areas with Mediterranean climates are altogether different, however, so only the assumption of close correlation between form and function in plants allows extrapolation of vegetation studies from one area to another.

California types. In western North America a number of kinds of so-called chaparral must be distinguished. In southern California the coastal sage formation consists of shrubs which are gray in color and low in stature, with soft, brittle wood; the formation contains *Artemisia californica*, shrubby species of *Salvia*, *Eriogonum fasciculatum* (typical form), and others. In the humid mountains of California, logged and burned coniferous forests have been replaced by a montane chaparral of the widespread western *Arctostaphylos patula* and *Ceanothus velutinus* plus the West Coast *Prunus emarginata*, the New Mexico–Californian–Cascade *Ceanothus integerrimus*, and the Californian endemics *Chamaebatia foliolosa*, *Castanopsis sempervirens*, *Quercus vacciniifolia*, and *Ceanothus cordulatus* (the last three at higher altitudes). North from Mexico along the Rocky Mountains to Denver and north to Cache Valley in the Wasatch Mountains is a deciduous chaparral characterized by *Q. gambelii*. In northern California and southwestern Oregon and in the Tehachapi and adjacent mountains of southern California is an analogous vegetation type characterized by *Q. garryana* ssp. *breweri*.

The true chaparral of California, however, is different from the above. Its characteristic species include *Adenostema fasciculatum*, *Ceanothus cuneatus*, *Q. dumosa*, *Heteromeles arbutifolia*, *Rhamnus californica*, *R. crocea*, and *Cercocarpus betuloides*, plus a host of endemic species of *Arctostaphylos* and *Ceanothus* and other Californian endemics, both shrubby and herbaceous. These plants determine the formation's physiognomy. It is a dense, uniform-appearing, evergreen, shrubby cover with sclerophyllous leaves and deep-penetrating roots.

Plant life forms. Although the shrub life form is dominant, a variety of life forms is represented in chaparral. A floristic spectrum of Raunkiaer's life forms from an area of serpentine chaparral in Napa County, California, containing 67 species of vascular plants had 4% phanerophytes, 18% nanophanerophytes, 12% chamaephytes, 27% hemicryptophytes, 18% geophytes, and 21% therophytes. Comparable data from other chaparral regions show lower nanophanerophyte and geophyte percentages and a higher therophyte percentage. *See* PLANTS, LIFE FORMS OF.

Geographic variations. The Californian type of chaparral formation runs south from the Rogue River valley of southern Oregon through the lower California mountains, west of the crest of the Sierra Nevada, below the *Pinus ponderosa* belt, to northern lower California. In Oregon *Ceanothus cuneatus* and *Arctostaphylos viscida* are characteristic; *Adenostema fasciculatum*, for example, drops out north of Shasta County, California. In this northern area summer thunderstorms occur. Another variation is evident in Arizona. Again, such Californian plants as *Adenostema fasciculatum* and *Heteromeles arbutifolia* drop out as the climate changes from winter precipitation to a bimodal occurrence of rain with a maximum in summer. The Arizona chaparral is characterized by *Q. turbinella*, *Arctostaphylos pungens*, *Rhus trilobata*, and *Ceanothus greggii*, which are all more or less desert species in California.

Ecology. There is extreme floristic diversity, and therefore presumably also ecological diversity, within the vegetation commonly called chaparral. Since the basic floristic study of the formation has not yet been done, it is extremely difficult to extrapolate ecological or experimental physiological data from one area of chaparral to another.

Climate. Ecologically, chaparral occurs in a climate which is hot and dry in summer, cool but not much below freezing in winter, with little or no snow, and with winter precipitation which is excessive and leaches the soil of nutrients. The need for water and its supply are exactly out of phase. Figures describing chaparral climates are variable. Temperatures may be maritime as at coastal San Diego or continental as at Ash Mountain at 1750 ft (530 m) elevation on the west slope of the Sierra Nevada. Amplitudes of mean

monthly temperatures at these stations are 46.2 and 69.4°F (7.9 and 20.8°C). Means are 60.6 and 63.5°F (15.9 and 17.5°C). Need for water potential evapotranspiration is 31.4 and 39.0 in. (785 and 974 mm) per year. Annual precipitations are 10 and 28 in. (250 and 700 mm) at these stations, with Redding, located at the northern end of the Sacramento Valley, having 38 in. (950 mm). Actual water use varies much less, being 10 in. (250 mm) at San Diego, 15.2 in. (380 mm) at Ash Mountain, and 16 in. (400 mm) at Redding. The differences between these figures and annual precipitation are runoff. In Oregon chaparral areas, need for water is less, about 28 in. (700 mm) per year, corresponding to the lower temperatures. However, actual water use is higher than in California, being 18.8–20.0 in. (470–500 mm) per year. In Arizona chaparral, runoffs are nil, and figures for actual water use are somewhat higher than in California, 14–20 in. (350–500 mm) per year.

A fair conclusion is that chaparral grows under no unique climate. It may, of course, still be differentiated from surrounding vegetation by climatic parameters. Its prevalence on south slopes, whereas some kind of woodland or forest occupies adjacent north slopes, is clear evidence of control by local climates.

Water relations. So far as water relations are concerned, the shrubs of the chaparral presumably transpire continuously during the dry summer since they are evergreen. They photosynthesize below the wilting point, and their roots penetrate very deeply (recorded to 28 ft or 8.5 m) as well as widely (ratio of root to crown spread 2 to 4). At great depths in the fractured rock, water may be perennially available although the soil itself is reduced to the permanent wilting point each summer. Some of the xeromorphic features of the plants may be related to low nitrogen supply, as has been found for ericaceous plants of acid bogs. *See* PLANT-WATER RELATIONS.

Soils. Chaparral soils are generally rocky, often shallow, or of extreme chemistry such as those derived from serpentine, and always low in fertility. In the very precipitous southern Californian mountains, soil erosion rates may be 0.04 in. (1 mm) per year over large watershed areas. The runoff figures mentioned above also contribute to low soil fertility through their leaching effects. The result is low available nitrogen. Quantitative data which would allow a nitrogen balance to be drawn up are lacking, although data do show that a single crop of *Ceanothus leucodermis* can double soil nitrogen contents. *See* SOIL.

Succession. Successionally chaparral is various. *Quercus durata* on serpentine, to which it is limited, may be a continuing type. On deep soils at higher foothill elevations, chaparral is often replaced by *Q. kelloggii*; on the still deeper soils at lower elevations, it may have replaced *Q. douglasii*. *Pinus sabiniana* comes up through dense chaparral stands, but it seems to affect the shrubs little. *Pinus halepensis* occupies a similar ecological position in the Mediterranean maquis. Fire often thickens a stand of chaparral since many of its species are crown sprouters; in other species seed germination is stimulated by fire. Careful management of deer or sheep grazing can keep burned chaparral areas open. Under natural conditions *Eriodictyon californicum* is successional to chaparral after fires, if replacement of the original shrubs does not follow directly by sprouting or seedings. *See* ECOLOGICAL SUCCESSION. Jack Major

Characiformes

An order of teleost fishes sharing the superorder Ostariophysi, series Otophysi, with the Cypriniformes (minnows, carps, and suckers), Siluriformes (catfishes), and Gymnotiformes (knife fishes). *See* CYPRINIFORMES; OSTARIOPHYSI; SILURIFORMES; TELEOSTEI.

Characteristics. The Characiformes are distinguished by the following characters: usually an adipose fin (only one of 18 families totally without an adipose fin; several families and genera with and without); teeth in jaws usually well-developed; pharyngeal teeth usually present but usually not specialized as in cypriniforms; upper jaw usually nonprotractile or, if movable, not protrusible as in most cyprinids; virtually all species have scales, with some being ctenoid or ctenoid-like (with a serrated margin) as opposed to the usual cycloid scales (with annual growth rings); barbels (slender, tactile processes near the mouth) are absent; usually three postcleithra; pelvic fins and pelvic girdle are minute in some species; and adult sizes range from only 13 mm (0.5 in.) to 1.4 m (4.6 ft).

Diversity and habitat. The Characiformes comprise 18 families, about 270 genera, and no less than 1675 species, all of which are limited to freshwaters in Africa (209 species) and in the New World from the southwestern United States (one species, Mexican tetra, *Astyanax mexicanus*), Mexico (eight species), and Central and South America (the remainder of the species).

Classification. The classification of the Characiformes has changed significantly since the mid-1980s. The current consensus classification for the order is as follows.

Suborder Citharinoidei. There are 20 genera and about 98 species. Features distinguishing the Citharinoidei from all other characiforms are bicuspid teeth in the jaws; second and third postcleithra fused; ascending process of premaxilla absent; and virtually all species with ctenoid scales. This suborder is endemic to Africa.

Family Distichodontidae. The family comprises 17 genera and 90 species. There are two distinct forms (see **illustration**). In one, the members have a nonprotractile upper jaw and are micropredators and herbivores, and the body shape varies from deep to moderately deep. The other group has an elongate body and movable jaws, and is carnivorous, feeding on fishes or the fins of fishes. Maximum length is about 83 cm (32.7 in.).

Family Citharinidae. There are 3 genera and 8 species. The body is deep, the dorsal and anal fins are

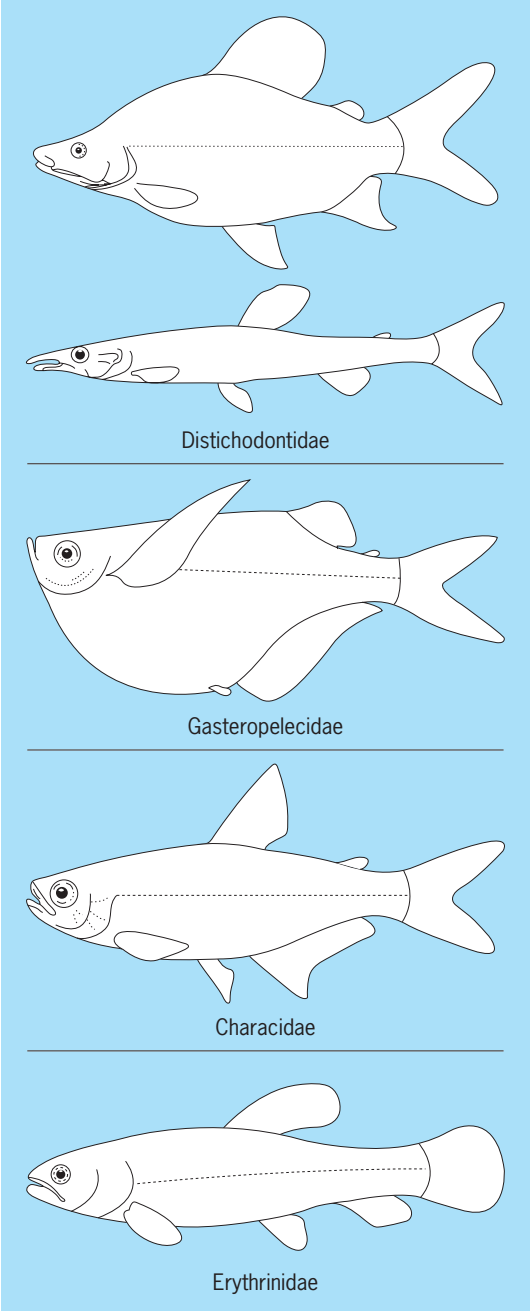

Four families of Characiformes. (*From J. S. Nelson, Fishes of the World, 2006, Wiley, New York*)

relatively long, and the maxilla is quite small and lacks teeth. Maximum length is 84 cm (33 in.).

Suborder Characoidei. The following families are contained in the suborder.

Family Parodontidae. This family comprises 3 genera and about 21 described species. It is characterized by a ventral mouth with greatly enlarged, highly mobile premaxillae and teeth modified for scraping algae off rocks. It inhabits mountain streams of eastern Panama and most of South America. Most species are less than 15 cm (6 in.) in length.

Family Curimatidae. There are 8 genera and 95 species. The absence of jaw teeth distinguishes this family from all other characiforms, except one genus, *Anodus*, in the family Hemiodontidae. Curimatidae occurs from southern Costa Rica to northern Argentina. Maximum length is 32 cm (12.6 in.), usually much less.

Family Prochilodontidae. The family includes 3 genera and 21 species. The mouth is protractile; the lips are enlarged, forming a sucking disc; and the jaws have numerous small teeth. Most species are herbivores or detritovores (organisms that consume dead organic matter), and some swim in an oblique head-down position. This family and the next two families are commonly called headstanders. The range is primarily in the northern half of South America. Maximum length is 74 cm (29 in.).

Family Anostomidae. The Anostomidae have 12 genera and 137 species. The mouth is small and nonprotractile; the premaxilla is enlarged and excludes the maxillae from the gape; the body is usually elongate; and the anal fin is short. The range is Central and South America. Maximum total length is about 88 cm (34.6 in.).

Family Chilodontidae. There are 2 genera and 7 species. The premaxilla is relatively small and the maxillae much enlarged. The range is northern South America. Maximum total length is 18 cm (7 in.).

Family Crenuchidae. The Crenuchidae has 12 genera and 74 species. The frontal bones have paired foramina posterodorsally to the orbits, which are enlarged in the subfamily Crenuchinae. Most of the species (10 genera and 71 species) are in the subfamily Characidiinae; in the largest genus, *Characidium*, are species that have the ability to climb waterfalls by using their paired fins to cling to the underside of rocks. The family inhabits eastern Panama and South America. Maximum total length is under 110 mm (4.3 in.), with the smallest species being only 57 mm (2.2 in.).

Family Hemiodontidae. The family comprises 5 genera and about 28 described species. Members are fast swimmers with a subcylindrical to fusiform body, adipose eyelids, teeth absent on the lower jaw of adults, and typically with a round spot on the side of the body and a stripe on the lower lobe of the caudal fin. Hemiodontids range from northern South America to the Paraná-Paraguay Basin. Maximum total length is about 33 cm (13 in.).

Family Alestiidae. There are 18 genera and 110 species in the family. Many species superficially resemble clupeids, being predominantly silvery. The body varies from deep to moderately slender; the mouth varies from terminal to superior; and many species have a large black spot at the base of the caudal fin. The family, endemic to Africa, is commonly known as African tetras. The maximum length of the small tetras is no more than 21 mm (8.3 in.), whereas the giant tigerfish (*Hydrocynus goliath*), a formidable fish with jaws armed with very large saberlike teeth, attains a length up to 1.4 m (4.6 ft).

Family Gasteropelecidae. Gasteropelicids comprise 3 genera and 9 species. This small family (both in number and size) is known as freshwater hatchetfishes because of the deep and strongly compressed

body in the shape of a half disk (see illustration). An adipose fin is present in the larger species but absent in the smaller species; the pelvic fins and associated girdle are minute; the elongate pectoral fins are provided with greatly enlarged musculature and corresponding skeletal support. These fishes are capable of jumping relatively high out of the water and making short flights by using their heavily muscled pectoral fins. Commercially collected for the aquarium trade, hatchetfishes are found in Panama and all countries of South America except Chile. Maximum total length ranges from about 23 to 76 mm (0.9 to 3 in.).

Family Characidae. With 165 genera and not less than 962 species, this family is by far the largest and most diverse of characiform fishes. It presents uncertainties regarding the taxonomic affinities of numerous genera. Many genera are monotypic, whereas several genera have numerous species (such as *Astyanax* with 86 species and *Hyphessobrycon* with 97). Some characins resemble shads and herrings (Clupeidae), minnows (Cyprinidae), and darters (Percidae) (see illustration). Included in the family are popular aquarium fishes (such as neon tetras), food fishes (such as *Brycon*), and the infamous piranhas (*Serrasalmus*). Characins range from southwestern Texas, through Mexico, and into Central and South America; elsewhere they are introduced. Maximum adult length varies from less than 30 mm (1.2 in.) to over 100 cm (39 in.).

Family Acestrorhynchidae. One genus and 15 species constitute this family. Members have an elongate pikelike body covered with small scales. They are found mostly in the Orinoco and Amazon basins of South America. Maximum total length is about 44 cm (17.3 in.).

Family Cynodontidae. There are 5 genera and 14 species. The mouth is oblique and the jaws bear large canine teeth, which in some species are saberlike. The family is endemic to South America. Maximum length is about 65 cm (25.6 in.).

Family Erythrinidae. There are 3 genera and 14 species. The family is characterized by a cylindrically shaped body (similar to that of the bowfin, Amiidae) (see illustration); relatively large mouth, the gape extending beyond the anterior margin of the eye; numerous palatine teeth, as well as sharp conical teeth in the jaws; no adipose fin; and a round caudal fin. It is endemic to South America, and some species can breathe air and move overland between ponds. Maximum length is about 1.0 m (3.3 ft.).

Family Lebiasinidae. The family comprises 7 genera and 61 species. The mouth is small and superior, the gape not reaching the eye; the adipose fin is present or absent; and in some species the upper lobe of the caudal fin is longer than the lower. The family occurs in Costa Rica, Panama, and South America.

Family Ctenoluciidae. There are 2 genera and 7 species. Members are called pike-characins because they resemble Northern Hemisphere pikes (Esocidae) in having an elongate body, large terminal mouth, jaws equipped with numerous small teeth and a single row of recurved teeth posteriorly in each jaw, and the anal and dorsal fins set far back on the body. Pike-characins are carnivorous, feeding primarily on fishes. They occur in Panama and South America. Maximum length of smallest species is 23 cm (9 in.); largest is 68 cm (27 in.).

Family Hepsetidae. The single species, *Hepsetus odoe*, is called the African pike. It has a pikelike body shape and a large mouth and jaws with small pointed teeth as well as a few large canine teeth; it lays its eggs in a nest of floating foam. These characteristics readily distinguish this monotypic family from other African characiforms. It is endemic to tropical Africa. Maximum length is 30 cm (11.8 in.). Herbert Boschung

Bibliography. P. A. Buckup, Relationships of the Characidiinae (Teleostei, Ostariophysi) and phylogeny of characiform fishes, pp. 123–144 in L. R. Malabarba et al. (eds.), *Phylogeny and Classification of Neotropical Fishes*, Edipucrs, Porto Alegre, Brazil, 1998; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; R. P. Vari, Higher level phylogenetic concepts within Characiformes (Ostariophysi), a historical review, pp. 111–122 in L. R. Malabarba et al. (eds.), *Phylogeny and Classification of Neotropical Fishes*, Edipucrs, Porto Alegre, Brazil, 1998.

Character recognition

The process of converting scanned images of machine-printed or handwritten text (numerals, letters, and symbols) into a computer-processable format; also known as optical character recognition (OCR).

Systems. A typical OCR system (**Fig. 1**) contains three logical components: an image scanner, OCR software and hardware, and an output interface. The image scanner optically captures text images to be recognized. Text images are processed with OCR software and hardware. The process involves three operations: document analysis (extracting individual character images), recognizing these images (based on shape), and contextual processing (either to correct misclassifications made by the recognition algorithm or to limit recognition choices). The output interface is responsible for communication of OCR system results to the outside world.

Image scanner. Four basic building blocks form functional image scanners: a detector (and associated electronics), an illumination source, a scan lens, and a document transport. The document transport places the document in the scanning field, the light source floods the object with illumination, and the lens forms the object's image on the detector. The detector consists of an array of elements, each of which converts incident light into a charge or analog signal. These analog signals are then converted into an image. Scanning is performed by the detector and the motion of the text object with respect to the detector. After an image is captured, the document transport removes the document from the scanning field.

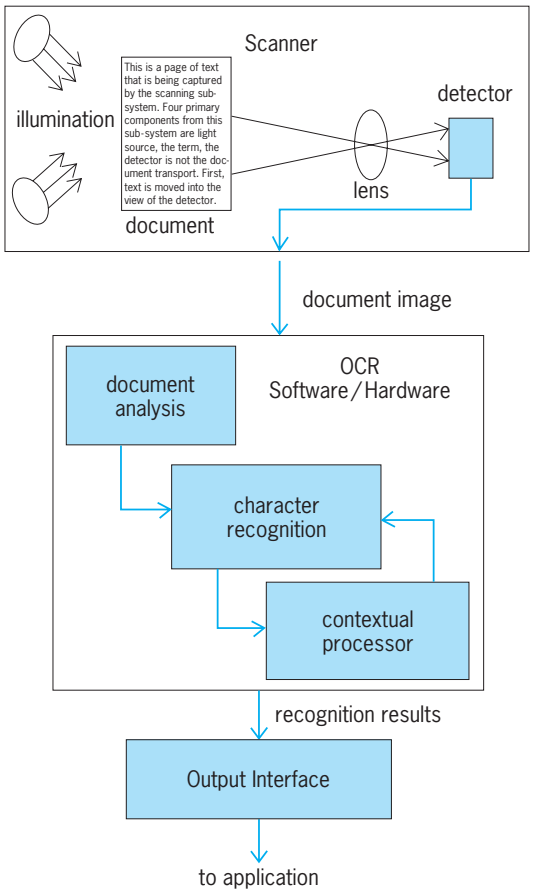

Fig. 1. General structure of an optical character recognition (OCR) system.

Over the years, advances in scanner technology made available higher resolution, often in the range of 300–400 pixels per inch (ppi). Recognition methods that use features (as opposed to template matching) use resolutions of at least 200 ppi and careful consideration of the gray scale. Lower resolutions and simple thresholding (that is, discrimination based only on whether the darkness of a pixel exceeds a certain threshold) tend to break thin lines or fill gaps, thus invalidating features.

Software and hardware. The software-hardware system that recognizes characters from a registered image can be divided into three operational steps: document analysis, character recognition, and contextual processing.

1. *Document analysis.* In this process, text is extracted from the document image. Reliable character segmentation and recognition depend upon both original document quality and registered image quality. Processes that attempt to compensate for poor-quality originals or poor-quality scanning include image enhancement, underline removal, and noise removal. Image enhancement methods emphasize discrimination between characters and objects that are not characters. Underline removal erases printed guidelines and other lines which may touch characters and interfere with character recognition. Noise removal erases portions of the image that are not part of characters. *See* IMAGE PROCESSING.

Prior to character recognition, it is necessary to isolate individual characters from the text image. Many OCR systems use connected components for this process. For those connected components that represent multiple or partial characters, more sophisticated algorithms are used. In low-quality or nonuniform text images, these sophisticated algorithms may not correctly extract characters, and thus recognition errors may occur. Recognition of unconstrained handwritten text can be very difficult because characters cannot be reliably isolated, especially when the text is cursive handwriting.

2. *Character-recognition algorithms.* Two essential components in a character-recognition algorithm are the feature extractor and the classifier. Feature analysis determines the descriptors, or feature set, used to describe all characters. Given a character image, the feature extractor derives the features that the character possesses. The derived features are then used as input to the character classifier.

Template matching, or matrix matching, is one of the most common classification methods. In template matching, individual image pixels are used as features. Classification is performed by comparing an input character image with a set of templates (or prototypes) from each character class. Each comparison results in a similarity measure between the input character and the template. One measure increases the amount of similarity when a pixel in the observed character is identical to the same pixel in the template image. If the pixels differ, the measure of similarity may be decreased. After all templates have been compared with the observed character image, the character's identity is assigned as the identity of the most similar template.

Template matching is a trainable process because template characters may be changed. In many commercial systems, programmable read-only memories (PROMs) store templates containing single fonts. To retrain the algorithm, the current PROMs are replaced with PROMs that contain images of a new font. Thus, if a suitable PROM exists for a font, template matching can be trained to recognize that font. The similarity measure of template matching may also be modified, but commercial OCR systems typically do not allow this modification. *See* COMPUTER STORAGE TECHNOLOGY; SEMICONDUCTOR MEMORIES.

Structural classification methods utilize structural features and decision rules to classify characters. Structural features may be defined in terms of character strokes, character holes, or other character attributes such as concavities. For instance, the letter P may be described as a vertical stroke with a hole attached on the upper right side. For a character image input, the structural features are extracted and a rule-based system is applied to classify the character. Structural methods are also trainable, but construction of a good feature set and a good rule base can be time consuming.

Many character recognizers are based on mathematical formalisms that minimize a measure of misclassification. These recognizers may use pixel-based features or structural features. Some examples are discriminant-function classifiers, bayesian classifiers, artificial neural networks, and template matchers. Discriminant-function classifiers use hypersurfaces to separate the featural descriptions of characters from different semantic classes and, in the process, reduce the mean-squared error. Bayesian methods seek to minimize the loss function associated with misclassification through the use of probability theory. Artificial neural networks, which are closer to theories of human perception, employ mathematical minimization techniques. Both discriminant functions and artificial neural networks are used in commercial OCR systems. *See* BAYESIAN STATISTICS; NEURAL NETWORK.

Character misclassifications stem from two main sources: poor-quality character images and poor discriminatory ability. Poor document quality, image scanning, and preprocessing can all degrade performance by yielding poor-quality characters. However, the character recognition method may not have been trained for a proper response on the character causing the error. This type of error source is difficult to overcome because the recognition method may have limitations and all possible character images cannot possibly be considered in training the classifier. Recognition rates for machine-printed characters can reach over 99%, but handwritten character recognition rates are typically lower because every person writes differently. This random nature often results in misclassifications.

3. *Contextual processing.* Contextual information can be used in recognition. The number of word choices for a given field can be limited by knowing the content of another field. For example, in recognizing the street name in an address, the street name choices can be limited to a lexicon by first correctly recognizing the postal code. Alternatively, the result of recognition can be postprocessed to correct the recognition errors. One method used to postprocess character recognition results is to apply a spelling checker to verify word spelling. Similarly, other postprocessing methods use lexicons to verify word results, or recognition results may be verified interactively with the user.

4. *Nonroman character recognition.* There are 26 different scripts in use that are other than roman. Some have had little work done on their recognition, for example, Kannada, while a significant amount of work has been done on others, for example, Japanese. In addition to alphanumerals, Japanese text uses Kanji characters (Chinese ideographs) and Kana (Japanese syllables). Therefore, it is considerably more difficult to recognize Japanese text because of the size of the character set (usually more than 3300 characters) and the complexity and similarity of the Kanji character structures. Low data quality is an additional problem in all OCRs. A Japanese OCR is usually composed of two individual classifiers (preclassifier and secondary classifier) in a cascade structure. The preclassifier first performs a fast coarse classification to reduce the character set to a short candidate list (usually containing no more than 100 candidates). The secondary classifier then uses more complex features to determine which candidate in the list has the closest match to the test pattern.

Output interface. The output interface allows character recognition results to be electronically transferred into the domain that uses the results. For example, many commercial systems allow recognition results to be placed directly into spread sheets, databases, and word processors. Other commercial systems use recognition results directly in further automated processing and, when the processing is complete, the recognition results are discarded. The output interface, while simple, is vital to the success of OCR systems because it communicates results to the outside world. *See* DATABASE MANAGEMENT SYSTEM; WORD PROCESSING.

Applications. Commercial OCR systems can largely be grouped into two categories: task-specific readers and general-purpose page readers. A task-specific reader handles only specific document types. Some of the most common task-specific readers read bank checks, letter mail, or credit-card slips. These readers usually utilize custom-made image-lift hardware that captures only a few predefined document regions. For example, a bank-check reader may scan just the courtesy-amount field (where the amount of the check is written numerically) and a postal OCR system may scan just the address block on a mail piece. Such systems emphasize high throughput rates and low error rates. Applications such as letter-mail reading have throughput rates of 12 letters per second with error rates less than 2%. The character recognizer in many task-specific readers is able to recognize both handwritten and machine-printed text.

General-purpose page readers are designed to handle a broader range of documents such as business letters, technical writings, and newspapers. These systems capture an image of a document page and separate the page into text regions and nontext regions. Nontext regions such as graphics and line drawings are often saved separately from the text and associated recognition results. Text regions are segmented into lines, words, and characters, and the characters are passed to the recognizer. Recognition results are output in a format that can be postprocessed by application software. Most of these page readers can read machine-written text, but only a few can read hand-printed alphanumerics.

Task-specific readers. Task-specific readers are used primarily for high-volume applications that require high system throughput. Since high throughput rates are desired, handling only the fields of interest helps reduce time constraints. Since similar documents possess similar size and layout structure, it is straightforward for the image scanner to focus on those fields

where the desired information lies. This approach can considerably reduce the image-processing and text-recognition time. Application areas to which task-specific readers have been applied include assigning postal codes to letter mail, reading data entered in forms such as tax forms, automatic accounting procedures used in processing utility bills, verification of account numbers and courtesy amounts on bank checks, automatic accounting of airline passenger tickets, and automatic validation of passports.

1. *Address readers.* The address reader in a postal mail sorter locates the destination address block on a mail piece and reads the postal code in this block. In the United States, additional fields in the address block are read with high confidence, the system may generate a 9-digit ZIP code for the piece. This ZIP code is then used to generate a bar code which is sprayed on the envelope (**Fig. 2**).

The multiline optical character reader (MLOCR) used by the U.S. Postal Service locates the address block on a mail piece, reads the whole address, identifies the 9-digit ZIP code, generates a 9-digit bar code, and sorts the mail to the correct stacker. The character classifier recognizes up to 400 fonts, and the system can process up to 45,000 mail pieces per hour.

2. *Form readers.* A form-reading system needs to discriminate between preprinted form instructions and filled-in data. The system is first trained with a blank form. The system registers those areas on the form where the data should be printed. During the form-recognition phase, the system uses the spatial information obtained from training to scan the regions that should be filled with data. Some readers read hand-printed data as well as various machine-written text. They can read data on a form without being confused with the form instructions. Some systems can process forms at a rate of 5800 per hour.

3. *Bill-processing systems.* In general, a bill-processing system is used to read payment slips, utility bills, and inventory documents. The system focuses on certain regions on a document where the expected information is located, such as the account number and payment value.

4. *Check readers.* A check reader captures check images and recognizes courtesy amounts and account information on the checks. Some readers also recognize the legal amounts (written alphabetically) on checks and use the information in both fields to cross-check the recognition results. An operator can correct misclassified characters by cross-validating the recognition results with the check image that appears on a system console.

5. *Airline ticket readers.* In order to claim revenue from an airline passenger ticket, an airline needs to have three records matched: the reservation record, the travel-agent record, and the passenger ticket. However, it is impossible to match all three records for every ticket sold. Current methods, which use manual random sampling of tickets, are far from accurate in claiming the maximal amount of revenue. Several airlines use a passenger-revenue accounting system to account accurately for passenger revenues. The system reads the ticket number on a passenger ticket and matches it with the one in the airline reservation database. It scans up to 260,000 tickets per day and achieves a sorting rate of 17 tickets per second.

6. *Passport readers.* An automated passport reader is used to speed passengers returning to the United States through custom inspections. The reader reads a traveler's name, date of birth, and passport number on the passport and checks these against the database records that contain information on fugitive felons and smugglers.

General-purpose page readers. The two general categories are high-end page readers and low-end page readers. High-end page readers are more advanced in recognition capability and have higher data throughput. A low-end page reader usually does not come with a scanner, and it is compatible with many flatbed scanners. Low-end page readers are used mostly in office environments with desktop workstations, which are less demanding in system throughput. Since they are designed to handle a broader range of documents, a sacrifice of recognition accuracy is

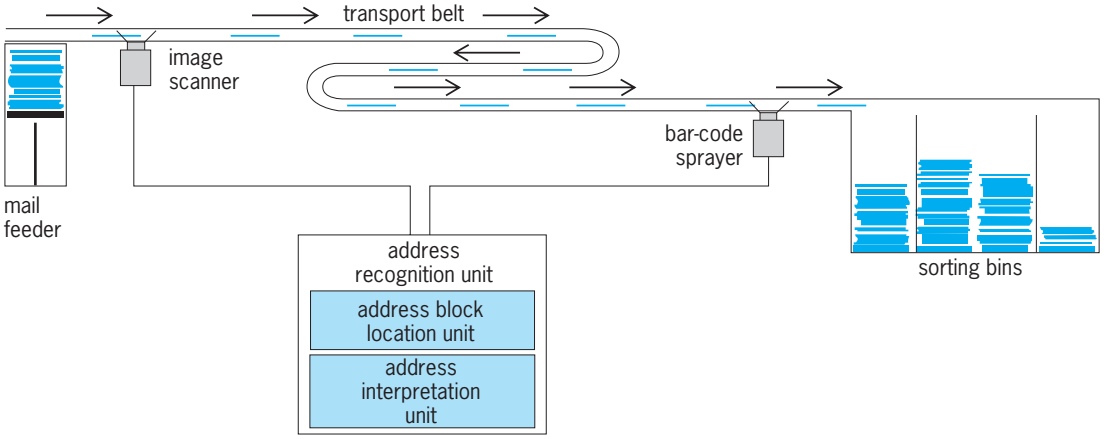

Fig. 2. Architecture of a postal address reading and sorting system.

necessary. Some commercial OCR software allows users to adapt the recognition engine to customer data for improving recognition accuracy. Some high-performance readers can detect type faces (such as boldface and italic) and output the formatted ASCII text in the corresponding style. *See* COMPUTER; COMPUTER VISION. Sargur N. Srihari; Stephen W. Lam

Bibliography. J. Cohen, *Automatic Identification and Data Collection Systems*, 1995; M. Ogg and H. Ogg, *Optical Character Recognition: A Librarian's Guide*, 1992; J. C. Reimel, *Image Processing and Optical Character Recognition: How They Work and How to Implement Them*, 1993; G. Searfoss, *Jis-Kanji Character Recognition Method*, 1994; S. N. Srihari, High-performance reading machines, *Proc. IEEE*, 80:1120–1132, 1992; P. S. Wang (ed.), *Character and Handwriting Recognition: Expanding Frontiers*, 1990.

Characteristic curve

A graphical display depicting complex nonlinear relationships in electronic circuits. A typical use is to show voltage-current relationships in semiconductor devices. Device amplification capabilities, for example, are exhibited by a characteristic plot which traces output current versus output voltage with a third controlling variable as a parameter. This control variable could be the base current of a bipolar junction transistor (BJT) or the gate-to-source voltage of a metal-oxide-semiconductor (MOS) transistor.

The properties of this display can be illustrated by an idealized version of a characteristic curve for an enhancement-mode MOS transistor (see **illus.**). Here, drain current (i_d) is plotted versus drain-to-source voltage (v_{DS}) for each of several values of gate-to-source voltage (v_{GS}) exceeding the device threshold voltage (V_t). The transistor does not conduct if the gate-to-source voltage is below the threshold voltage. The characteristic illustrates how the transistor operating modes depend on its applied voltages, with the transistor gate-to-source voltage having a much larger, highly nonlinear effect on the drain current than does the drain-to-source voltage.

A small increase in drain current with drain-to-source voltage is caused by a phenomenon of semiconductor physics called channel length modulation. To a first-order approximation, this effect can be modeled as a linear resistor (r_a) connected between the device drain and source. The intersection point (V_a) of the projection of the line with the voltage axis is sometimes called the early voltage, a term borrowed from bipolar junction transistor theory. The inverse of V_a is called the channel-length modulation parameter (λ).

Other characteristics often included in transistor data sheets are displays of current gain versus bias current, gain versus frequency, and input and output impedances versus frequency. Less commonly, other graphical nonlinear relationships, such as the variation of thermocouple voltage with temperature or the dependence of electrical motor torque with current, also are known as characteristic curves.

In the past, characteristic curves were used as tools in the graphical solution of nonlinear circuit equations that are followed by relationships of this type. In current practice, this analysis is performed using computer packages for circuit simulation. Designers still use characteristic curves from data sheets, however, to evaluate relative performance capabilities when selecting devices, and to provide the information needed for a preliminary

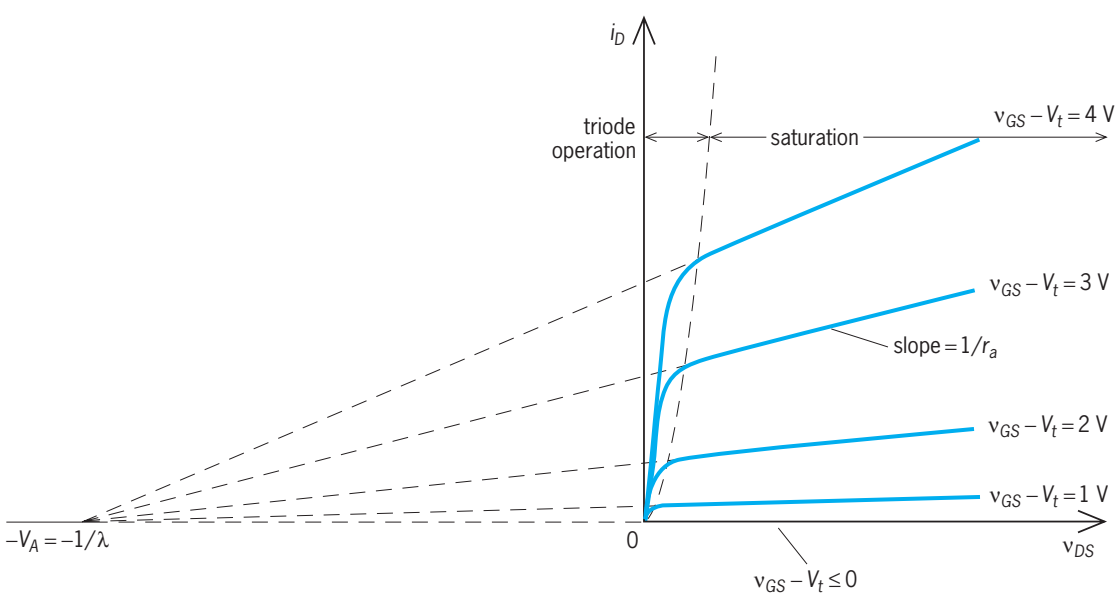

Idealized metal-oxide-semiconductor (MOS) transistor characteristic curve. The increase in drain current with drain-to-source voltage is exaggerated for clarity. (*After A. S. Sedra and K. C. Smith, Microelectronic Circuits, 4th ed., Oxford, 1998*)

pencil-and-paper circuit design. *See* AMPLIFIER; ELECTRICAL MODEL; TRANSISTOR. Philip V. Lopresti

Bibliography. R. T. Howe and C. Sodini, *Microelectronics: An Integrated Approach*, Prentice Hall, 1997; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., Oxford, 1998.

Charadriiformes

A large, diverse, worldwide order of shore and aquatic birds found from dry plains to the open ocean, and from the tropics to the high latitudes. Charadriiformes may be closely related to the pigeons (Columbiformes) on the one side and to the cranes, rails, and their allies (Gruiformes) on the other. The sandgrouse, here included in the Columbiformes, are sometimes placed in the Charadriiformes. Some workers place the flamingos (Phoenicopteriformes) and the ibises (Threskionithidae; Ciconiiformes) in this order, but these decisions are highly controversial. *See* CICONIIFORMES; COLUMBIFORMES; GRUIFORMES; PHOENICOPTERIFORMES.

Classification. The order is arranged into 3 suborders and 17 families as follows:

Order Charadriiformes
- Suborder Charadrii
  - Superfamily Jacanoidea
    - Family: Jacanidae (jacansas or lily-walkers; 8 species; pantropical)
    - Rostratulidae (painted snipe; 2 species; Old World tropics)
  - Superfamily Charadrioidea
    - Family: Graculavidae (Cretaceous of New Jersey)
    - Haematopodidae (oyster catchers; 7 species; worldwide, coastal)
    - Charadriidae (plovers and lapwings; 64 species; worldwide)
    - Scolopacidae (sandpipers, curlews, phalaropes, and snipe; 86 species; worldwide)
    - Recurvirostridae (stilts and avocets; 10 species; worldwide)
    - Dromadidae (crab plovers; 1 species; coast of the Arabian peninsula)
    - Burhinidae (thick-knees; 9 species; worldwide, except North America)
    - Glareolidae (pratincoles; 16 species; Old World)
  - Superfamily Chionidoidea
    - Family: Thinocoridae (seed snipe; 4 species; Andes of South America)
    - Chionidae (sheath-bills; 2 species; southern tip of South America)
- Suborder Lari
  - Family: Stercorariidae (skuas; 5 species; worldwide, oceanic)
  - Laridae (gulls and terns; 88 species; worldwide)
  - Rynchopidae (skimmers; 3 species; pantropical and North America)
- Suborder Alcae
  - Family Alcidae (auks, murres, and puffins; 23 species; northern oceans)

Several of these families, such as the Charadriidae (plovers), Scolopacidae (sandpipers), and Laridae (gulls and terns), are divided into two or more subfamilies. Some workers place the Australian plains-wanderer (Pedionomidae) in this order, but this is still controversial and the group is here included in the Gruiformes. Some workers also place the Presbyornithidae (Eocene of North and South America) in the Charadriiformes, but they are here considered to be Anseriformes. *See* ANSERIFORMES.

Fossil record. Being shore and water birds, the charadriiforms have a long and generally good fossil record, including some of the earliest fossils that are definitely assigned to a living order. The Graculavidae are Cretaceous birds from New Jersey whose affinities to Recent families have not been established, but they may be ancestral to a diversity of orders. A number of other presumably charadriiform fossils are known from the Cretaceous, but their relationships are still uncertain. The larids and alcids have good fossil records, but most of the other families are represented by scattered remains and several are unrepresented as fossils. This lack arises presumably because of the fragile nature of the skeleton of most members of the order.

Suborders. The three suborders of charadiiforms are quite different groups.

Charadrii. The Charadrii are the typical shorebirds, usually found in marshy areas and along shores, but some are in dry areas; a few, the phalaropes, are mainly aquatic. They can run and fly well. Most live in flocks, although most breed solitarily in nests on the ground. Male ruffs (*Philomachus pugnax*; Europe) are much larger than females and court communally, in leks (gathering places for courtship), with females coming to the leks and choosing particular males. The crab plovers nest colonially in tunnels dug into sand banks. Young charadrii are downy and leave the nest on hatching. The phalaropes (Scolopacidae: Phalaropodinae) [see **illustration**], painted snipes, and jacanas show sex reversal, with the females being larger than the males, usually more brightly

***Phalaropus fulicaria*, the red phalarope. (© *California Academy of Sciences; photographers Gerald and Buff Corsi*)**

colored, and assuming the lead in courtship; the males incubate and assume most of the care of the young. Most feed on insects and other small animals; the seed snipes are mainly vegetarian. The bills of many species are highly specialized for food capture, with the most bizarre being that of the wry-billed plover (*Anarhynchus frontalis*) of New Zealand, in which the bill tip bends sharply to the right for probing under stones for insects. Many species breed in the Arctic or cool temperate regions and migrate long distances in large flocks. Some species have long transoceanic migrations, for example, the bristle-thighed curlew (*Numenius tahitiensis*), which flies from Alaska to Hawaii and then to Tahiti and surrounding islands. Many shorebirds were, or still are, of economic importance as game birds. The numbers of many species have been greatly reduced because of overhunting or loss of habitat, mainly migration and wintering grounds. A few species, for example, the Eskimo curlew (*Numenius borealis*), are severely threatened or extinct. *See* ENDANGERED SPECIES; SEXUAL DIMORPHISM.

Lari. The Lari include the skuas, gulls, terns, and skimmers, predominantly aquatic birds that find their food by flying over the water. They are long-winged, excellent fliers with short legs, but they also can walk well. The feet are webbed. Most species breed in large colonies, laying up to four eggs, and the downy young remain in or near the nest until they can fly. Most species can swim on the water surface, although terns do so less often. A few oceanic terns lack waterproof plumage and never alight on the water, sleeping on the wing. Although many species are marine, most are found near the coast and are not truly oceanic. Skimmers have an elongated, knifelike lower jaw with which they skim the surface of the water for fish. Many species breed in the Arctic and cold temperate regions and migrate. The Arctic tern has the longest known migratory flight, up to 22,000 mi (36,000 km) per year. Skuas are highly predatory, as well as robbing other birds of their food; they have the distinction of having been seen closer to both Poles than any other birds.

Alcae. The Alcae include only the alcids, which are true marine, swimming and diving birds found only in the Northern Hemisphere. They are mainly black and white, and have webbed feet and reduced wings. Alcids are excellent divers and swim underwater by using their wings, which are reduced to rather stiff, paddlelike structures. They have a quick, buzzy flight used mainly to reach their nesting sites. Alcids feed on fish and other aquatic animals. They usually breed in large colonies, on rocky ledges, on the ground, or in burrows. However, one species of auklets, the marbled murrelet (*Brachyramphus marmorotus*), breeds in large trees, placing the nest on large branches in trees of old forests. The young are downy, and in some species leave the nest before they are fully grown. The great auk of the North Atlantic, the "penguin" of a number of writers, was the largest known species and was flightless; it is now extinct. *See* AVES.

Walter J. Bock

Bibliography. J. del Hoyo et al. (eds.), Order Charadriiformes, *Handbook of the Birds of the World*, vol. 3, pp. 276–722, Lynx Edicions, 1996; P. Hayman, J. Marchant, and T. Prater, *Shorebirds*, Houghton Mifflin, 1986; E. A. Johnsgard, *The Plovers, Sandpipers and Snipes of the World*, University of Nebraska Press, 1981.

Charcoal

A porous solid product containing 85–98% carbon produced by heating carbonaceous materials such as cellulose, wood, peat, and coals of bituminous or lower rank at 930–1110°F (500–600°C) in the absence of air. Chars or charcoals are rendered more porous and more efficient for sorption by heating in air, carbon dioxide, or steam at 1650°F (900°C) for a brief period. The surface area of activated char produced by this technique is about 800,000 ft^2/oz (1000 m^2/g). Chars from cellulose or wood are soft and friable. They are used chiefly for decolorizing solutions of sugar and other foodstuffs and for removing objectionable tastes and odors from water. Chars from nutshells and coal are dense, hard carbons. They are used in gas masks and in chemical manufacturing for many mixture separations. Another use is for the tertiary treatment of wastewater. Residual organic matter is adsorbed effectively to improve the water quality.

The sorptive properties of chars are the result of their very large internal surface. The selectivity of the sorption can be modified by mixing additives to the char, and by controlling the partial pressure or concentration of components, the temperature of the mixture, and (in liquids) the acidity or polar environment. *See* ACTIVATED CARBON; ADSORPTION; CARBON; COKE; DESTRUCTIVE DISTILLATION; PYROLYSIS; WOOD CHEMICALS.

Joseph H. Field

Charge-coupled devices

Semiconductor devices wherein minority charge is stored in a spatially defined depletion region (potential well) at the surface of a semiconductor, and is moved about the surface by transferring this charge to similar adjacent wells. The formation of the potential well is controlled by the manipulation of voltage applied to surface electrodes. Since a potential well represents a nonequilibrium state, it will fill with minority charge from normal thermal generation. Thus a charge-coupled device (CCD) must be continuously clocked or refreshed to maintain its usefulness. In general, the potential wells are strung together as shift registers. Charge is injected or generated at various input ports and then transferred to an output detector. By appropriate design to minimize the dispersive effects that are associated with the charge-transfer process, well-defined charge packets can be moved over relatively long distances through thousands of transfers.

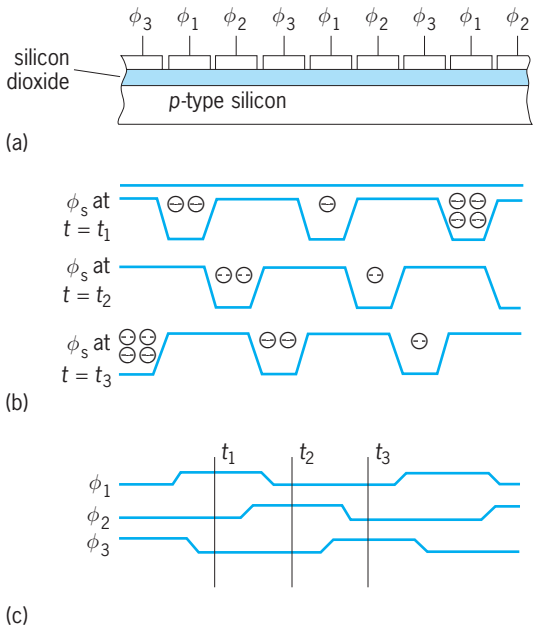

Fig. 1. Operation of three-phase charge-coupled shift register. (*a*) Cross section of register through channel. (*b*) Profile of surface potential ϕ_s for three different clock time intervals t_1, t_2, and t_3. (*c*) Voltage waveforms ϕ_1, ϕ_2, and ϕ_3 for the three-phase clocks.

Control of charge motion. There are several methods of controlling the charge motion, all of which rely upon providing a lower potential for the charge in the desired direction. When an electrode is placed in proximity to a semiconductor surface, the electrode's potential can control the near-surface potential within the semiconductor. The basis for this control is the same as for metal-oxide-semiconductor (MOS) transistor action. If closely spaced electrodes are at different voltages, they will form potential wells of different depths. Free charge will move from the region of higher potential to the one of lower potential. In one arrangement, a charge packet can be moved to the right by alternating the voltage on three electrodes in proper phase **(Fig. 1)**. In another scheme, an asymmetry built into a well can direct the charge in a given direction **(Fig. 2)**. Asymmetries of this type are easily created by using implanted ion layers or varying dielectric thickness. The three-phase structure shown in Fig. 1 has the ability to reverse the charge direction by a change in electrical phase.

Transfer efficiency. An important property of a charge-coupled device is its ability to transfer almost all of the charge from one well to the next. Without this feature, charge packets would be quickly distorted and lose their identity. This ability to transfer charge is measured as transfer efficiency, which must be very good for the structure to be useful in long registers. Values greater than 99.9% per transfer are not uncommon. This means that only 10% of the original charge is lost after 100 transfers.

Several mechanisms influence the transfer of charge from one well to the next **(Fig. 3)**. Initial self-induced drift acts to separate the charge. This repulsion of like charge is the dominant transfer mechanism for large signals, and is effective for the first 99% or so of the charge. Near the edge of the transfer electrode the potential gradient creates a field which sweeps the charge onward. Thermal diffusion accounts for the transfer of the remaining charge.

Electron traps. It would appear that, given sufficient time, almost all of the charge could be transferred. Two other mechanisms are at work to counter this. Within the silicon and at its surface are sites that can act as electron traps. This is especially true at the surface, where numerous surface states exist. These traps collect charge when exposed to a large charge packet and then slowly release it during later cycles of small charge packets. By stringing a large number of empty charge packets together, the traps can be emptied completely. When the first packet containing charge arrives, it could be completely consumed, recharging the traps.

Other than process steps that minimize the trap density, there are two methods to alleviate the severity of this problem. The first is never to allow a series of completely empty charge packets to occur. Instead of an empty charge packet, a minimum charge quantity is always present. This charge, called a fat zero, can be 10–20% of the well capacity. Under some conditions, this can reduce the trapping effects to tolerable levels. The second approach is to use a channel for the charge whose potential minimum does not occur at the surface. This is called a buried channel device, as opposed to a surface channel device. Since the charge is located within the bulk silicon below the surface, it is not exposed to the surface-state traps. Bulk traps remain, but because of their low density they are almost insignificant.

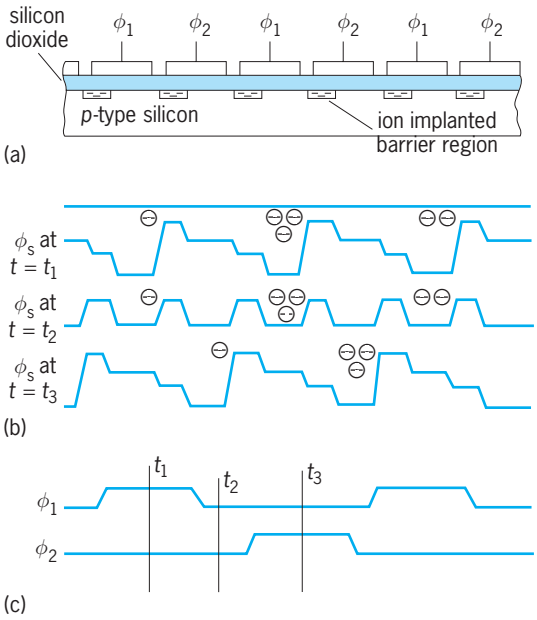

Fig. 2. Operation of two-phase charge-coupled shift register. (*a*) Cross section of register through channel. (*b*) Profile of surface potential ϕ_s for three different clock time intervals t_1, t_2, and t_3. (*c*) Voltage waveforms ϕ_1 and ϕ_2 for the two-phase clocks.

The process to shift the channel potential minimum below the surface can be either an ion-implant layer or an epitaxy layer combined with an implant layer. Because the charge is further removed from the surface control electrode, the maximum charge density has been reduced by a factor of 2 to 3. For the same reason, the fringing fields are greater and charge transfer can be much faster. The charge also moves with bulk mobility rather than the lower surface mobility, further enhancing performance. Fractional charge losses of as low as 5×10^{-5} at clock frequencies greater than 100 MHz have been reported for buried-channel charge-coupled-device structures. This means that such a structure could transfer 10^6 electrons from one well to another in less than 10 ns, leaving fewer than 50 electrons behind. *See* ION IMPLANTATION.

Localized potential minima. In addition to electron-trapping sites, there is another reason that charge can be left behind. Where the two adjacent surface electrodes come together, there is necessarily a gap. This gap is usually quite small, and sometimes may even be covered by one of the electrodes. However, the gap represents a region of poorly controlled surface potential. It is possible, under some conditions, for the transition from one potential well to the next to have perturbations in it. These perturbations, or glitches, can trap charge. They represent a localized minipotential well within the larger one. Charge trapped by the glitch will be left behind after a transfer. If the glitch remained filled, continued clocking of the structure would not represent a problem. However, the size of the glitch may change, depending upon electrode clock amplitude. Also, charge is released over the glitch as thermionic emission over a barrier potential. Thus a glitch can be emptied by a series of empty charge packets to trap the next packet-containing charge. Many times, glitches occur at input or output ports without detriment to the structure's operation, since their effect may be small. However, if a glitch occurs within the repetitive register, the multiplication effect can have a serious impact on overall performance.

Lifetime. A second important property of a charge-coupled-device register is its lifetime. When the surface electrode is clocked high, the potential within the semiconductor also increases. Majority charge is swept away, leaving behind a depletion layer. If the potential is taken sufficiently high, the surface goes into deep depletion until an inversion layer is formed and adequate minority charge collected to satisfy the field requirements. The time it takes for minority charge to fill the well is the measure of well lifetime. The major sources of unwanted charge are the following: thermal diffusion of substrate minority charge to the edge of the depletion region, where it is collected in the well; electron-hole pair generation within the depletion region; and the emission of minority charge by traps. Surface-channel charge-coupled devices usually have a better lifetime, since surface-state trap emission is suppressed and the depletion regions are usually smaller.

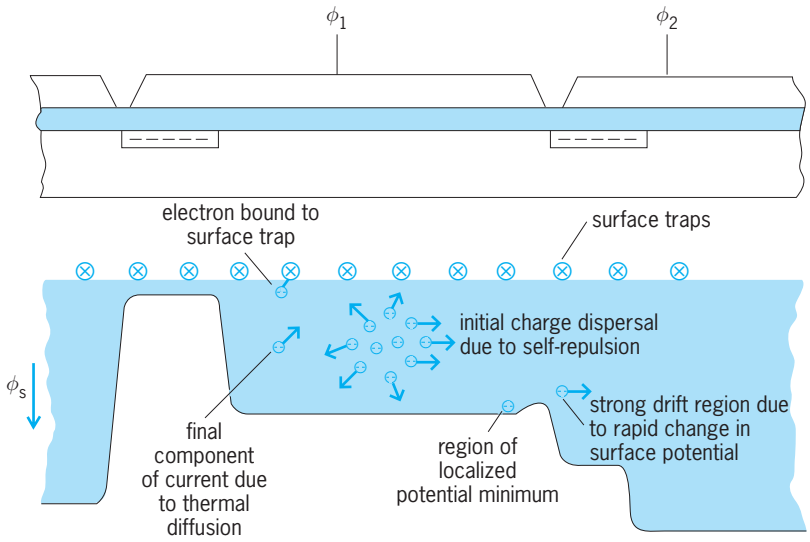

Fig. 3. Mechanisms determining charge-transfer rate. The three terms contributing to charge transfer include self-induced drift, channel drift fields, and thermal diffusion. Charge is left behind due to electron traps and localized potential minima.

Input and output ports. Once adequate transfer efficiency and sufficient lifetime have been achieved, input and output ports must be established. The port structure of a charge-coupled device depends upon its application. Analog registers require linear inputs and outputs with good dynamic range.

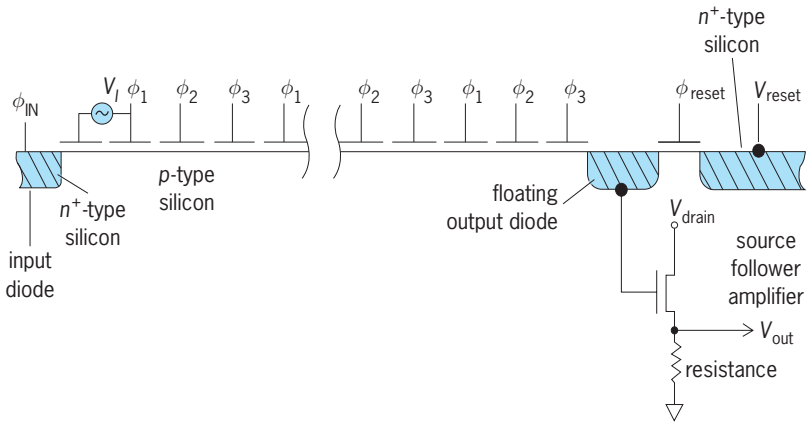

Fig. 4. Charge-coupled register showing analog input and output ports. A source follower output is shown.

Digital registers require precisely metered and detected ones and zeros. Imaging devices use photoemission at the register sites as input with analog-output amplifiers.

The most common input structure relies upon a diode as a source of minority charge **(Fig. 4)**. Whenever the diode potential drops below the threshold potential of the adjacent gate, an inversion layer is formed beneath the gate. By pulsing the diode potential low or the adjacent gate potential high, charge can be injected into the register. To prevent the charge from returning to the diode, the next gate in the register immediately collects it. The amount of charge injected can be controlled by adjusting the channel potential difference between the first two gates.

A common output circuit allows the output diode to be reset at the start of each cycle to the reset level (Fig. 4). During the second part of the cycle, charge from the charge-coupled device register is dumped on the floating diode, causing its voltage to change. The diode is connected to a gate which can be used in an amplifier configuration. For digital applications, the output goes to a regenerative circuit with sufficient gain to produce full-level signals at the output. Mark R. Guidry

Signal processing. Many of the early applications of charge-coupled-device technology were in signal processing. Most of those applications utilized a charge-coupled-device shift register as an analog delay line. In this configuration, an analog input signal was sampled and a packet of charge proportional to its amplitude was created and shifted into the register. After transferring through the register, the charge was converted back into voltage form at the output. The amount of delay was proportional to both the number of shift register elements and the clocking frequency. The charge-coupled-device analog delay line was used in various functions, including analog memory, time base correction, and analog filtering. However, as digital memory and digital signal processing technologies have advanced, the use of charge-coupled devices in signal processing has decreased.

Image sensors. The most significant current application of the charge-coupled-device concept is as an imaging device. Charge-coupled-device image sensors utilize the fact that silicon is sensitive to light. In fact, silicon is sensitive to wavelengths from about 400 to 1100 nanometers (from ultraviolet to near-infrared). When light photons penetrate the silicon surface, hole-electron pairs are created in the silicon. The number of hole-electron pairs created is a function of wavelength (photon energy), intensity (number of photons), and duration (length of time exposed to light).

In a charge-coupled-device image sensor, the light is focused upon an array of picture elements (pixels). These pixels collect the electrons as they are created. The number of electrons collected in each pixel is representative of the light intensity projected onto the sensor at that point. Periodically, the charges from all of the pixels are read out, and the image can then be reconstructed from the intensity and pixel location data.

There are two primary categories of image sensors. Linear image sensors have the pixels aligned along a central axis. Area image sensors have the pixels arranged in a rectangular (rows × columns) array pattern. Linear image sensors require relative motion between the sensor and the object being scanned. The relative motion is precisely known so that, as the object is scanned one line at a time, it can then be reconstructed one line at a time. Area image sensors do not require this motion.

Linear image sensors. The typical linear image sensor consists of a linear array of pixels (to collect the light-generated charge) with two output shift registers (one on each side) to transport the charge to the output circuit. After a set amount of time (the integration time), all of the photogenerated charge is transferred out of the pixels and into the appropriate output shift register. While the pixels begin to collect charge for the next line, the two output shift registers transfer the present line out to the output circuits.

Because the sensor must be moved relative to its target, early applications were in copiers and facsimile machines, in which either the target or the sensor was moved in a direction perpendicular to the axis of the linear array. As the speeds of linear image sensors increased into the 20–60-MHz range, charge-coupled devices were used to scan objects moving at a continuous high speed. Industrial and food inspection systems were developed to improve the quality of product by detection of defects on the line. The U.S. Postal Service now uses linear image sensors to quickly scan addresses on letters. A specialized high-speed application of the charge-coupled device is in telecine machines, which scan movie film and convert the film into a television format, allowing television stations to broadcast movies.

New applications were also created as the resolution of charge-coupled devices increased to 6000 pixels or more. Graphics arts scanners can scan photographs at higher resolution than the film used to create the photograph. High resolution is also critical for remote sensing and aerial reconnaissance. Charge-coupled-device cameras have been placed in orbit aboard satellites for remote sensing of large regions of the Earth to assist in land use management. In aerial reconnaissance applications, charge-coupled-device cameras scan the ground below while the airplane travels at a fixed speed. *See* REMOTE SENSING.

Area image sensors. The area image sensor does not require relative motion between sensor and object because the pixels are arranged in a two-dimensional array. Area image sensors have a variety of architectures, but in general they all have an array of pixels (arranged in rows and columns) and vertical shift registers to transfer the charge in parallel to the output shift register, which transports each line out serially to the output circuit.

Early area image sensors were designed to be used in simple, closed-circuit television and scientific applications. However, as the quality and resolution of area image sensors increased, they began to

be used throughout the television industry. Compared to conventional vidicon-tube-based equipment, chargecoupled-device cameras benefit from reduced size and power consumption, greater ruggedness and reliability, and lower cost. Virtually all camcorders and most television cameras are now based on charge-coupled devices. Charge-coupled-device cameras are also used in surveillance and inspection systems of all kinds and are ideal for robotics applications. Video conferencing and video telephones use lower-resolution versions of these cameras. *See* COMPUTER VISION; TELEVISION CAMERA; TELEVISION CAMERA TUBE; VIDEOTELEPHONY.

The resolution of area image sensors has become equivalent to photographic film, enabling the development of digital photography. Digital imaging provides several benefits. Costly film and toxic film-developing chemicals are eliminated. Images are obtained immediately and can easily be enhanced through software. Image storage and transmission is faster and easier. Cameras with very large, very high resolution area image sensors provide professional photographers better final pictures than are obtainable with conventional film, while lower-resolution, lower-cost, digital charge-coupled-device cameras are available to consumers. *See* CAMERA.

A miniaturized charge-coupled-device camera allows a dentist to see inside a patient's mouth or a physician to see inside a patient's body. Charge-coupled-device area imagers are also used in intraoral dental x-ray systems. These imagers replace the x-ray film that is normally placed within the patient's mouth. The imager is connected to an adjacent computer where the image is immediately displayed. Since silicon-based charge-coupled devices do not image x-rays well, a layer of scintillating material is deposited or placed in front of the sensor. This material has a special phosphor that emits green light when exposed to x-ray energy. The charge-coupled device then "sees" this green image. The charge-coupled device/scintillator combination is more sensitive than standard x-ray film, allowing the dentist to reduce the x-ray exposure of the patient while still producing a superior image.

Charge-coupled-device-based systems with very large area image sensors have been introduced in mammography, to image x-rays of the human breast. Software is being developed to interpret the images and increase the accuracy of diagnosis.

Astronomers have long used charge-coupled device area image sensor cameras mounted on very high power telescopes. By synchronizing the motion of the telescope with the Earth's rotation, the camera can "stare" at one spot in space for hours at a time. The long integration times allow distant objects to be imaged that are otherwise invisible. To keep the sensor from being saturated with thermally generated charge, these cameras typically cool the charge-coupled-device chip down to −50 to −100°C (−58 to −148°F). *See* ASTRONOMICAL PHOTOGRAPHY; INTEGRATED CIRCUITS; SEMICONDUCTOR.

Steven Onishi

Bibliography. G. C. Horst, *CCD Arrays, Cameras and Displays*, JCD Publishing, Winter Park, FL, and SPIE Optical Engineering Press, Bellingham, WA, 1996; A. J. P. Theuwissen, *Solid State Imaging with Charge-Coupled Devices*, Kluwer Academic, 1995.

Charge-density wave

A possible ground state of a metal in which the conduction-electron charge density is sinusoidally modulated in space. The periodicity of this extra modulation is unrelated to the lattice periodicity. Instead, it is determined by the dimension of the conduction-electron Fermi surface in momentum space.

Description. The conduction-electron charge density $\rho_0(\vec{r})$ would ordinarily exhibit a dependence on position $\vec{r}$ having the same spatial periodicity as that of the positive-ion lattice. A metal with a charge-density wave (CDW) has an additional charge modulation described by Eq. (1). The fractional amplitude of the charge-density wave is f and typically has a value of approximately 0.1. The wave vector $\vec{Q}$ of the charge-density wave is determined by the conduction-electron Fermi surface. In a simple metal, having a spherical surface of radius p_F in momentum space, the magnitude of $\vec{Q}$ is approximately the value in Eq. (2), where h is Planck's constant.

$$\rho(\vec{r}) = \rho_0(\vec{r})[1 + f\cos(\vec{Q}\cdot\vec{r} + \phi)] \qquad (1)$$

$$Q \approx \frac{4\pi p_F}{h} \qquad (2)$$

Although the wavelength of a charge-density wave is comparable to the spacing between lattice planes, their ratio is not a rational number. The charge-density wave is then said to be incommensurate. In such a case, the total energy of the metal is independent of the phase ϕ in Eq. (1). *See* FERMI SURFACE.

Origin. In a quasi-one-dimensional metal, for which conduction electrons are mobile in one direction only, a charge-density wave can be caused by a Peierls instability. This mechanism involves interaction between the electrons and a periodic lattice distortion having a wave vector Q parallel to the conduction axis. The linear-chain metal niobium triselenide ($NbSe_3$) is prototypical, and exhibits nonlinear conduction phenomena arising from electrically induced dynamic variations of the phase ϕ in Eq. (1).

For isotropic metals, and quasi-two-dimensional metals, Coulomb interactions between electrons are the cause of a charge-density wave instability. The exchange energy, an effect of the Pauli exclusion principle, and the correlation energy, an effect of electron-electron scattering, both act to stabilize a charge-density wave. However, the electrostatic energy attributable to the charge modulation in Eq. (1) would suppress a charge-density wave were it not for a compensating charge response of the positive-ion lattice. *See* EXCHANGE INTERACTION; EXCLUSION PRINCIPLE.

Lattice distortion. Suppose that $\vec{u}(\vec{r})$ is the displacement of a positive ion from its lattice site at $\vec{r}$. Then

a wavelike displacement given by Eq. (3) will gener-

$$\vec{u}(\vec{r}) = \vec{A} \sin (\vec{Q} \cdot \vec{r} + \phi) \qquad (3)$$

ate a positive-ion charge density that almost cancels the electronic charge modulation of Eq. (1). A typical value of the displacement amplitude A is about 1% of the lattice constant. Ion-ion repulsive interactions must be small in order to permit such a distortion. Consequently, charge-density waves are more likely to occur in metals having small elastic moduli.

Detection. The unambiguous signature of a charge-density wave is the observation of two satellites, on opposite sides of each Bragg reflection, in a single-crystal diffraction experiment, employing either x-rays, neutrons, or electrons. The satellites are caused by the periodic lattice displacement, Eq. (3). Charge-density waves were first seen by electron diffraction in layered metals like tantalum disulfide (TaS_2) and tantalum diselenide ($TaSe_2$), which have three charge-density waves. At reduced temperature, transitions from incommensurate to commensurate $\vec{Q}$'s are observed. The length of a charge-density wave in the latter case is then an integral multiple of some lattice periodicity. The charge-density wave in the elemental metal potassium has been observed by neutron diffraction. The wavelength of the charge-density wave in this case is 1.5% larger than the spacing between close-packed atomic planes. *See* ELECTRON DIFFRACTION; NEUTRON DIFFRACTION; X-RAY DIFFRACTION.

Fermi surface effects. If simple metals like sodium and potassium did not have a charge-density-wave structure, their conduction-electron Fermi surface would be almost a perfect sphere. Many electronic conduction phenomena would then be isotropic, and the low-temperature magnetoresistance would be essentially zero. The presence of a charge-density wave leads to a dramatic contradiction of such expectations. Conduction-electron dynamics will be modified by the presence of a new potential having the periodicity of Eq. (1). The Fermi surface will distort and become multiply connected. In high magnetic fields some electrons will travel in open orbits rather than in closed, cyclotron orbits. The low-temperature magnetoresistance will then exhibit sharp resonances as the magnetic field is rotated relative to the crystal axes. Such phenomena have been observed in sodium and potassium and have been explained with charge-density-wave theory. *See* MAGNETORESISTANCE.

$\vec{Q}$-domains. Generally, the $\vec{Q}$ direction will not be the same throughout an entire sample of, for example, a cubic crystal having a single charge-density wave. There will be a domain structure analogous to magnetic domains in a ferromagnet. The $\vec{Q}$ direction will (of course) prefer some specific axis described by direction cosines α, β, γ. In a cubic crystal there would be 24 equivalent axes and, therefore, 24 $\vec{Q}$-domain types. As a consequence, some physical properties of a sample will depend markedly on the orientation distribution of its $\vec{Q}$-domains. For example, the low-temperature resistivity of a potassium wire might be several times larger if its $\vec{Q}$-domains are oriented parallel to the wire, than the value obtained if the domains were oriented perpendicular. Since $\vec{Q}$-domain distribution can be altered by stress-induced domain regrowth, some physical properties will vary significantly from experiment to experiment, even on the same sample. Such behavior is observed in alkali metals. Techniques for the control of $\vec{Q}$-domain orientation have not yet been developed. *See* DOMAIN (ELECTRICITY AND MAGNETISM); ELECTRICAL RESISTIVITY.

Phasons. The energy decrease caused by the existence of an incommensurate charge-density wave is independent of the phase ϕ, Eq. (1). It follows that there are new low-frequency excitations which can be described by a slowly varying phase modulation of the charge-density wave, as represented approximately in Eq. (4). Quantized excitations of this type

$$\phi \rightarrow \phi(\vec{r}, t) = a(\vec{q}) \sin [\vec{q} \cdot \vec{r} - \omega(\vec{q})t] \qquad (4)$$

are called phasons. They exist only for small $\vec{q}$, $\vec{q}\ \vec{Q}$; and their frequency $\omega(\vec{q})$ approaches zero linearly with q.

A number of physical phenomena caused by phasons have been observed. Phasons give rise to a low-temperature anomaly in the heat capacity. The charge-density waves in lanthanum digerminide ($LaGe_2$) were first suspected from its anomalous heat capacity. The phason spectral density of potassium has been directly observed in point-contact spectroscopy. Phasons reduce the intensity of charge-density-wave diffraction satellites with increasing temperature and contribute a diffuse-scattering cloud which surrounds each satellite. The phason velocity in tantalum disulfide has been determined from measurements of the phason diffuse scattering. Conduction electrons are strongly scattered by phasons, and such processes modify the temperature dependence of the resistivity at low temperature. *See* BAND THEORY OF SOLIDS; CRYSTAL STRUCTURE; SPIN-DENSITY WAVE.

Albert W. Overhauser

Bibliography. F. Bassani et al. (eds.), *Highlights of Condensed-Matter Theory*, Proceedings of the International School of Physics "Enrico Fermi", Course 89, Varenna on Lake Como, 1983; T. Butz (ed.), *Nuclear Spectroscopy on Charge Density Wave Systems*, 1992; L. P. Gor'kov and G. Grüner (eds.), *Charge Density Waves in Solids*, 1990; G. Grüner, The dynamics of charge density waves, *Rev. Mod. Phys.*, 60:1129–1181, 1988.

Charge symmetry

The rule that the interchange of opposite members of a multiplet of subatomic particles in a system or process does not affect any property of the system or the outcome of the process.

Examples. A multiplet is a group of subatomic particles that are (nearly) identical except that each member of the multiplet carries a different amount of electrical charge. The simplest example is the two-member multiplet consisting of the proton and neutron. Both are called nucleons, and both are the

building blocks of the atomic nucleus. They have the same spin ($^1/_2\ \hbar$, where $\hbar$ is Planck's constant divided by 2π) and nearly the same mass, but the proton carries one unit of positive electrical charge while the neutron has no charge. It is the proton's charge that attracts and binds negatively charged electrons to make atoms. Many other multiplets have been found in the particles created in accelerators. These include the pion with three charge states (π^+, π^0, π^-) and the delta with four (Δ^{++}, Δ^+, Δ^0, Δ^-). *See* BARYON; MESON; NEUTRON; NUCLEON; PROTON.

The mathematical framework that best describes the behavior of particles within a multiplet mirrors the framework for spin, with the value of the isospin being $(N - 1)/2$, where the number of multiplet members is N. If a process or property is the same for all members of the multiplet, that property obeys isospin (I-spin) symmetry. Charge symmetry is a special case in which the property is the same when opposite members of the multiplet are swapped. *See* I-SPIN; SPIN (QUANTUM MECHANICS).

For the nucleon, charge symmetry requires that swapping neutrons and protons in a nucleus or a subatomic reaction has no physical effect. Thus, charge symmetry would predict that the triton (nucleus made of one proton and two neutrons) would have the same spin and mass as the helium-3 (^{3}He) nucleus (two protons and one neutron). This rule has proven useful in the classification of subatomic particles and the states of the atomic nucleus. In addition, if there is a reaction among subatomic particles that can proceed through different members of a multiplet, the relative rates of these reactions are related by the mathematical rules that govern isospin. *See* ELEMENTARY PARTICLE; NUCLEAR STRUCTURE; SYMMETRY LAWS (PHYSICS).

Charge symmetry breaking. The predictions of the charge symmetry rule are not exactly obeyed. Neutrons have more mass than protons by 0.14%. Such small deviations are called charge symmetry breaking. They arise from two causes:

1. The change in electric force effects that accompanies the swap of a multiplet member with one electrical charge for another with a different charge. While the strong force between nucleons binds together the nucleus and dominates nuclear properties, the electrical force of repulsion between similar charges makes the force between two protons more repulsive than the force between two neutrons.

2. The heavier mass of the down quark compared to the up quark. Thus, the neutron (composed of two down quarks and one up quark) is heavier than the proton (two up quarks and one down quark).

By measuring the small deviations in nuclear properties and reactions that are a result of charge symmetry breaking, scientists hope to untangle these two causes and measure the size of the quark mass difference effect.

Consequences of charge symmetry breaking. The small mass difference between the neutron and the proton makes it possible for the neutron to decay into a proton, an electron, and an antineutrino while the proton is stable. Left by itself, the neutron will decay with a half-life of about 10 min. The neutron can survive for a long time only if it is bound into a stable nucleus or is a part of a massive object such as a neutron star. In these situations, there is not enough available energy to create the decay products. *See* NEUTRON STAR.

The stability of the proton compared to the neutron led to the abundance of hydrogen (the atom made of a single proton and electron) after the universe was formed in the big bang. This makes possible the process of fusion that powers the Sun and stars. It leads to the formation of water and the many chemicals that are essential for life on Earth. *See* BIG BANG THEORY.

Effects that measure charge symmetry breaking. With the discovery of charmed particles in 1975, it was confirmed that protons and neutrons, as well as many other subatomic particles, were composed of a new kind of particles called quarks. Because quarks are permanently confined inside the nucleons and other particles, their masses cannot be directly measured. For this information, we must rely on experiments that are sensitive to some effect caused by the quark mass in combination with theoretical models that relate the mass to the effect. Measurements of charge symmetry breaking can provide information on the mass difference between the up and down quarks. *See* CHARM; QUARKS.

The mass difference between the neutron and the proton (1.2933317 ± 0.0000005 MeV/c^2, where c is the speed of light) is one such piece of information that is known to great precision. Another piece of information is the scattering length, a measure of the strength of the force between two nucleons when they collide at low speeds. After the effects of the electric force are subtracted using theory, the proton-proton length (-17.3 ± 0.4 femtometer, where the negative value indicates attraction) still differs from the neutron-neutron length (-18.8 ± 0.3 fm), an indication of charge symmetry breaking. The neutron-proton length of -23.77 ± 0.09 fm differs by an even larger amount due to the electric effects of charge independence breaking. If all three lengths had been the same, this property would have obeyed isospin symmetry. *See* SCATTERING EXPERIMENTS (NUCLEI).

At its longest ranges, the strong force that binds nucleons can be described by the exchange between nucleons of short-lived subatomic particles called mesons. Charge symmetry breaking may cause mesons that are otherwise similar but belong to different multiplets with different isospin values to mix. This allows, for example, a ρ meson (isospin = 1) to transform in flight into an ω meson (isospin = 0). Other mixings may occur (pion into η or η' meson) but have not yet been precisely measured.

The scattering of neutrons and protons from each other at speeds generated in accelerators depends on the orientation of their spin axes during the collision. A small difference of 0.6% in the scattering probability appears in the comparison of neutron versus proton spin effects, a difference attributed to charge symmetry breaking.

Effective field theories. The model connecting such experimental results with the underlying quark properties is chiral effective field theory. First developed

by Steven Weinberg, this model tracks nucleons and pions within a framework that obeys the rules of quark interactions. This model points to experiments in which the pion scatters from nucleons as the best place to obtain information on the difference between the up and down quark masses. This has led to two observations of charge symmetry breaking in experiments that produce electrically neutral pions, both of which reported results in 2003. *See* QUANTUM FIELD THEORY.

Neutron-proton fusion with pion production. An experiment conducted by Allena Opper, Elie Korkmaz, and their colleagues measured the rates at which the fusion of a neutron and proton gave rise to a deuteron (the neutron-proton nuclear bound state) and a pion emerging at different angles. A difference in the rates at forward and backward angles, as shown in the **illustration**, is a signature of charge symmetry breaking. *See* DEUTERON.

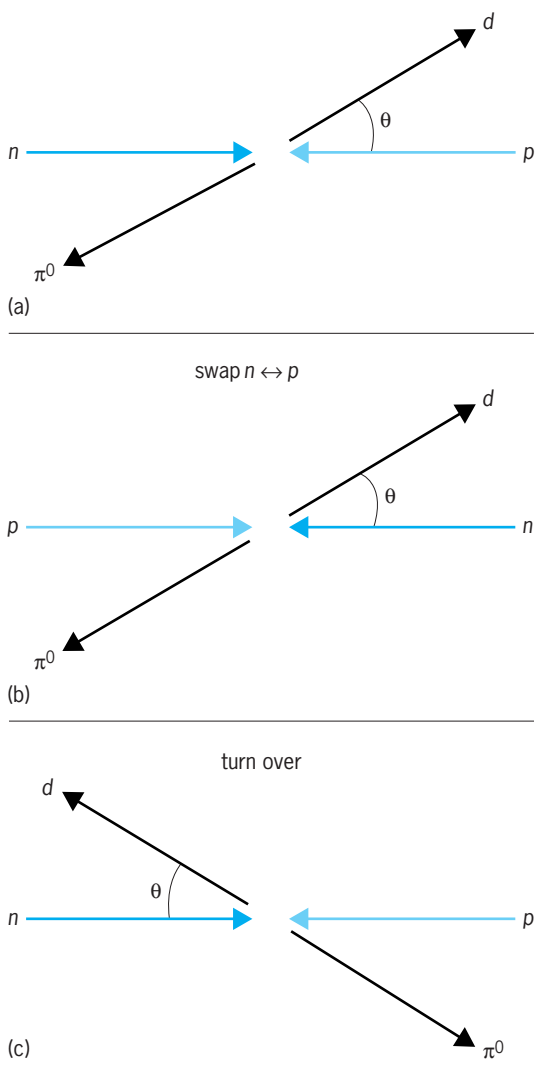

Views of the reaction that brings together a neutron and a proton to form a deuteron and an electrically neutral pion. (*a*) The neutron *n* and proton *p* come together from the left and right, creating a deuteron *d* and an uncharged pion π^0. The deuteron comes out at an angle θ. (*b*) The same process with the neutron and proton swapped. (*c*) The same process as *b* but turned around so that the neutron and proton again come from the left and right as in *a*. Then the deuteron emerges to the left instead of the right.

The illustration shows three views of the reaction. In illus. *a*, the neutron and proton enter from the left and right, producing a deuteron and pion that go out back-to-back; the scattering angle is θ. Charge symmetry requires the same reaction rate if the neutron and proton are swapped (illus. *b*). When this picture is turned around (illus. *c*) to place the proton and neutron as they were in illus. *a*, then θ is on the left instead of the right. So charge symmetry requires that just as many deuterons should go to the left as to the right at the same angle θ. Using a magnetic spectrometer located at the TRIUMF cyclotron laboratory in Vancouver, British Columbia, the experimental team was able to measure a small "fore-aft" asymmetry in this reaction rate equal to $0.17 \pm 0.10\%$, the amount of charge symmetry breaking in this process. *See* PARTICLE ACCELERATOR.

Deuteron-deuteron fusion with pion production. In this experiment, conducted by Edward Stephenson, Andrew Bacher, and their colleagues at the Indiana University Cyclotron Facility, an electron-cooled storage ring was used to bring two deuterons together to make a helium-4 (^{4}He) nucleus and a pion. In this case, swapping neutrons and protons in either the deuteron or ^{4}He nucleus changes nothing as these nuclei are self-conjugate (members of multiplets with only a single member and isospin equal to 0). But the pion is part of a multiplet having three members (π^+, π^0, and π^-) and an isospin value of 1. So this reaction is forbidden to occur to the extent that the amount of isospin before and after the reaction must remain the same. Using a very sensitive magnetic channel for ^{4}He and lead-glass Cerenkov detectors to see the light from the decay of the pion, the team was able to observe this reaction for the first time. The rate is very small; only 1 out of 10^{10} collisions between two deuterons resulted in the production of ^{4}He and a pion. *See* CERENKOV RADIATION.

Interpretation of results. The interpretation of these results is underway. In each case, all indirect mechanisms for breaking charge symmetry must be included along with the quark mass difference and electromagnetic effects. These include meson mixing, as described above. In addition, there may be states in either the deuteron or the ^{4}He nucleus that do not maintain a single value of isospin and would allow these symmetry-breaking effects to appear. The importance of such mechanisms is not well known, and will have to be determined self-consistently among all of the results of charge symmetry breaking. It is likely that such a determination will emerge and yield a better insight into the properties of the quarks inside the proton and neutron.

Edward J. Stephenson

Bibliography. A. G≦rdestig et al., Survey of charge symmetry breaking operators for $dd \rightarrow \alpha\pi^0$, *Phys. Rev. C*, 69:044606, 2004; G. A. Miller, B. M. K. Nefkens, and I. Šlaus, Charge symmetry, quarks and mesons, *Phys. Rep.*, 194:1–116, 1990; G. A. Miller and W. T. H. van Oers, Charge independence and charge symmetry, in W. C. Haxton and E. M. Henley (eds.), *Symmetries and Fundamental Interactions in Nuclei*, pp. 127–168, World Scientific, Singapore,

1995; A. K. Opper et al., Charge symmetry breaking in $np \rightarrow d\pi^0$, *Phys. Rev. Lett.*, 91:212302, 2003; E. J. Stephenson et al., Observation of the charge symmetry breaking $d + d \rightarrow {}^4\text{He} + \pi^0$ reaction near threshold, *Phys. Rev. Lett.*, 91:142302, 2003; U. van Kolck, J. A. Niskanen, and G. A. Miller, Charge symmetry violation in $pn \rightarrow d\pi^0$ and chiral effective field theory, *Phys. Lett. B*, 493:65–72, 2000.

Charged particle beams

Unidirectional streams of charged particles traveling at high velocities. Charged particles can be accelerated to high velocities by electromagnetic fields. They are then able to travel through matter (termed an absorber), interacting with it, losing energy, and causing various effects important in many applications. The velocities under consideration in this article exceed 100,000 m/s (about 60 mi/s or 200,000 mi/h), equivalent to an energy of 100 eV for a proton, and can approach the speed of light c (about 3×10^8 m/s or 6.7×10^8 mi/h). Examples of charged particles are electrons, positrons, protons, antiprotons, alpha particles, and any ions (atoms with one or several electrons removed or added). In addition, some particles are produced artificially and may be short-lived (pions, muons).

Excluded from consideration are particles of, for example, cosmic dust (micrometeorites), which are clumps of thousands or millions of atoms. *See* COSMIC RAYS; ELECTRONVOLT; ELEMENTARY PARTICLE; PARTICLE ACCELERATOR.

Particle properties. Fast charged particles are described in terms of the following properties (values for some particles are given in **Table 1**):

Charge, z, in multiples of the electron charge $e = 1.6022 \times 10^{-19}$ coulomb. At small velocities the charge may be less than the charge of the nucleus because electrons may be present in some of the atomic shells.

Rest mass, M; usually the energy equivalent Mc^2 (in MeV) is given.

Rest mass, m, of electron; $mc^2 = 0.51104$ MeV.

Mass in atomic mass units, u; $A_m = Mc^2/931.481$ MeV.

Kinetic energy, T, in MeV.

Velocity, υ, in cm/s or m/s.

$\beta = \upsilon/c$; $c = 299{,}792{,}458$ m/s = speed of light.

TABLE 1. Properties of charged particles*

Ion	z	Lifetime, ns	Mass 10^{-24} g	Mass u	Mass MeV
Electron†	−1	Stable	0.910956	0.548593	511.004
Muon	1	2198.3	0.188357	0.113432	105.6598
Pion	1	26.04	0.248823	0.149846	139.578
Kaon	1	12.35	0.880322	0.530147	493.82
Sigma+	1	0.081	2.120318	1.276895	1189.40
Sigma−	−1	0.164	2.134436	1.285398	1197.32
		Mass excess,‡ MeV			
1N	0	8.0714	1.674920	1.0086652	939.553
1H	1	7.2890	1.672614	1.0072766	938.259
2H	1	13.1359	3.343569	2.0135536	1875.587
3H	1	14.9500	5.007334	3.0155011	2808.883
3He	2	14.9313	5.006390	3.0149325	2808.353
4He	2	2.4248	6.644626	4.0015059	3727.328
6Li	3	14.0884	9.985570	6.0134789	5601.443
7Li	3	14.9073	11.647561	7.0143581	6533.743
7Be	4	15.7689	11.648186	7.0147345	6534.093
9Be	4	11.3505	14.961372	9.0099911	8392.637
10B	5	12.0552	16.622243	10.0101958	9324.309
11B	5	8.6677	18.276741	11.0065623	10252.406
12C	6	0	19.920910	11.9967084	11174.708
13C	6	3.1246	21.587011	13.0000629	12109.314
14C	6	3.0198	23.247356	13.9999504	13040.691
14N	7	2.8637	23.246166	13.9992342	13040.024
15N	7	.1004	24.901771	14.9962676	13968.741
16O	8	−4.7365	26.552769	15.9905263	14894.875
17O	8	−.8077	28.220304	16.9947441	15830.285
18O	8	−.7824	29.880881	17.9947713	16761.791
19F	9	−1.4860	31.539247	18.9934674	17692.058
20Ne	10	−7.0415	33.188963	19.9869546	18617.472
21Ne	10	−5.7299	34.851833	20.9883627	19550.265
22Ne	10	−8.0249	36.508273	21.9858989	20479.451

*From *American Institute of Physics Handbook*, 3d ed., McGraw-Hill, 1972.
†Electron masses to be divided by 1000.
‡Mass excess given for neutral atoms; it is used to calculate nuclear reaction Q values.

For absorbers, the information needed is:

Atomic number, Z (number of protons in nucleus).
Average atomic weight, A (usually in g/mole, but numerically equal to mass in u).
Absorber thickness, x, in cm or g/cm^2.
Physical state (solid, liquid, gas, plasma).

See ATOMIC MASS UNIT; ENERGY; FUNDAMENTAL CONSTANTS; MASS.

Relation of velocity and energy. For small velocities ($\beta < 0.2$), the approximation $T = \frac{1}{2} Mc^2\beta^2$ is accurate to 3%. The expressions $\beta^2 = \zeta(\zeta + 2)/(\zeta + 1)^2$, and $T = Mc^2[(1/\sqrt{1-\beta^2}), - 1]$, with $\zeta = T/Mc^2$, are correct for all velocities. For $\zeta > 100$, the expression $\beta^2 = 1 - (1/\kappa^2)$, with $\kappa = \zeta + 1$, is more suitable to provide accurate values. *See* RELATIVISTIC MECHANICS.

Range. If a parallel beam of monoenergetic particles (that is, a beam in which all particles have exactly the same kinetic energy T) enters an absorber with a flat and smooth surface, it is found that all the particles (with the exception of electrons) travel along almost straight lines, slow down at approximately the same rate, and stop at approximately the same distance x from the surface (**Fig. 1**). The average distance traveled by the particles is called the mean range $R(T)$.

Energy loss and straggling. If the same beam travels through a thin absorber, it will emerge from it with a reduced average energy $\langle T_1 \rangle$ (**Table 2**). The difference between T and $\langle T_1 \rangle$ is called the average energy loss $\Delta = T - \langle T_1 \rangle$. Owing to the randomness of the number of collisions experienced by each particle, the range and the reduced energy fluctuate around the average values. This fluctuation is called straggling of energy loss or of range. (The shape of the curve in Fig. 1 for $x > 22$ mm is determined by range straggling.)

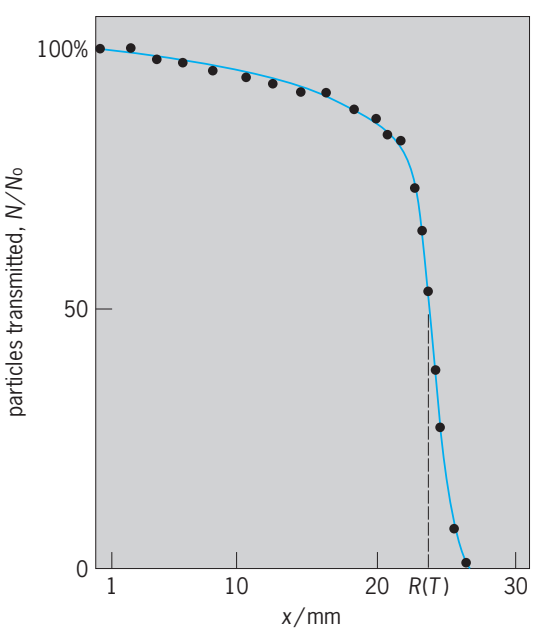

Fig. 1. Transmission curve for protons with kinetic energy $T = 144$ MeV through copper. (*After L. Koschmieder, Zur Energiebestimmung von Protonen aus Reichweitemessungen, Z. Naturforsch., 19a:1414–1416, 1964*)

TABLE 2. Calculated reduced average energy at various depths in copper of protons with T = 144 MeV*

x/mm	T_1/MeV
1	141
5	127
10	107
15	85
20	57

*From *American Institute of Physics Handbook*, 3d ed., McGraw-Hill, 1972.

Charge state. At high velocities, an ion usually has the full charge ze of the nucleus. As soon as υ drops below the velocity $u_K \cong zc/137$ of the K shell electrons, electrons in the absorber will be attracted into K-shell orbits, thus reducing the total charge of the ion to a value $z^* = z - 1$ or $z - 2$. As the velocity drops further, more and more electrons will be attached to the ion; but since some of these electrons will be lost or gained as successive collisions take place, z^* must be considered as an average value which changes with υ (**Fig. 2**). *See* ELECTRON CONFIGURATION.

Interactions. In traveling through matter, charged particles interact with nuclei, producing nuclear reactions and elastic and inelastic collisions with the electrons (electronic collisions) and with entire atoms of the absorber (atomic collisions). Usually, in its travel through matter a charged particle makes few or no nuclear reactions or inelastic nuclear collisions, but many electronic and atomic collisions. The average distance between successive collisions is called the mean free path, λ. In solids, it is of the order of 10 cm (4 in.) for nuclear reactions. It ranges from the diameter of the atoms (about 10^{-10} m) to about 10^{-7} m for electronic collisions. The mean free path, λ, depends on the properties of the particle and, most importantly, on its velocity.

Nuclear interactions. In nuclear reactions [for example, $Be(d,n)B$] the incident particle is removed from the beam. Therefore, a reduction in the fluence (number of particles in the beam per cm^2) will be observed. [The decrease in the number of particles in Fig. 1 for x less than about 22 mm (0.88 in.) is due to nuclear interactions.] Such an attenuation can be described by Eq. (1), where $e = 2.71828\ldots$, $N_0 =$

$$\frac{N}{N_0} = e^{-x\Sigma} = e^{-x/\lambda} \tag{1}$$

initial particle fluence at $x = 0$, $N =$ particle fluence at x, and $\Sigma =$ probability for an interaction to take place per centimeter of absorber. The mean free path is $\lambda = 1/\sigma$; Σ is equal to $n\sigma$, where $\sigma =$ cross section per atom for nuclear reactions in square centimeters, and $n =$ number of atoms per cubic centimeter of absorber material. *See* ATTENUATION; E (MATHEMATICS); NUCLEAR CHEMISTRY; NUCLEAR REACTION; SCATTERING EXPERIMENTS (NUCLEI).

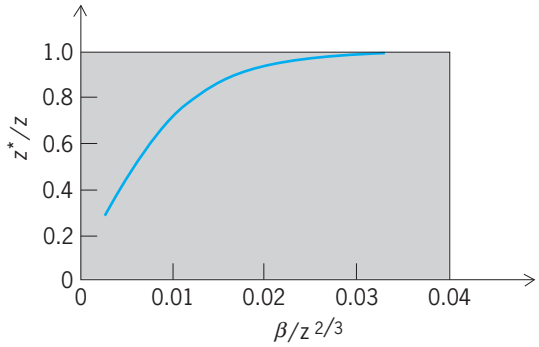

Fig. 2. Average charge z^* of a particle of velocity $v = \beta c$. (*After American Institute of Physics Handbook, 2d ed., McGraw-Hill, 1963*)

A rough estimate of the cross section for particles with $T/A_m > 10$ MeV can be obtained from Eq. (2).

$$\sigma = 5 \times 10^{-26}\ \text{cm}^2\ (A_m^{1/3} + A^{1/3})^2 \qquad (2)$$

For thin absorbers, $N/N_0 \cong 1 - x/\lambda + \cdots$.

An important nuclear interaction is Coulomb (or Rutherford) scattering: because both the incident particles and the nuclei of the absorber atoms have electric charge with values ze and Ze, respectively, a change in the direction of motion of the particles will take place during the passage of the particles near the nuclei. The total cross section for this process is only slightly less than the cross-sectional area of the total atom. Usually, though, the angular deflection is much less than 0.01°, and only multiple scattering, the compounding of collisions with many atoms, will cause noticeable total deflections. If an observation of a very fine beam of particles were made along the direction of motion, the scattering events would be seen as small lateral displacements in random directions, and the final lateral displacement would be their vector sum. Although very few particles experience no deflection, the most probable location of the particles is still on the original line of the beam.

Bremsstrahlung. If a charged particle is accelerated, it can emit photons called bremsstrahlung. This process is of great importance for electrons as well as for heavy ions with $T \gg Mc^2$. It is used extensively for the production of x-rays in radiology. Electrons circulating in storage rings emit large numbers of photons with energies (100–1000 eV) not readily available from other sources. *See* ACCELERATION; BREMSSTRAHLUNG; SYNCHROTRON RADIATION.

Atomic collisions. At low velocities it may be convenient to consider separately collisions in which most of the energy loss is given as kinetic energy to a target atom. Usually, electronic excitation, electron rearrangements, and possibly ionization accompany this process. The term nuclear collision is used by some scientists. No simple quantitative description of atomic collisions is available. *See* SCATTERING EXPERIMENTS (ATOMS AND MOLECULES).

Electronic collisions. The interaction and energy transfer (**Fig. 3**) between the charged particle and the electrons are caused by the Coulomb force. In general, except in tenuous plasmas, electrons are bound. In gases, all electrons are bound to individual atoms or molecules in well-defined orbits. For these isolated molecules (henceforth, atoms will be included with molecules), electrons can be moved into other bound orbits (excitation) requiring a well-defined energy ϵ_2. Another possibility is the complete removal of the electron from the atom (ionization), requiring an energy $\epsilon \geq I$, where I is the ionization energy for the particular electron. The secondary electron, which is called a delta electron, will have kinetic energy $K = \epsilon - I$. In both processes, the charged particle will lose energy; the energy loss is ϵ_e or ϵ, respectively. Also, it will be deflected very slightly. However, the change in direction is so small that it does not show in Fig. 3; the larger deflection caused by Rutherford scattering at point a is visible. *See* ATOMIC STRUCTURE AND SPECTRA; DELTA ELECTRONS; EXCITED STATE.

In liquids and solids, only the inner electrons are associated with a specific nucleus (in aluminum metal the K- and L-shell electrons). Excitation and ionization processes for these electrons are very similar to those in free molecules. The outer electrons are either associated with several neighboring nuclei (nonconducting materials) or, in metals (in Al, the three M-shell electrons), form a plasmalike cloud. Collective or plasma excitations ($\epsilon \cong 20$ eV) take place with high probability, but direct ionization ($\epsilon \gg 20$ eV) also occurs. Because of the requirements of momentum conservation, the maximum energy loss which can occur is given by $\epsilon_M \simeq 2mv^2 = 2mc^2\beta^2$ for particles heavier than electrons. *See* SOLID-STATE CHEMISTRY.

The probability for energy losses ϵ by the incident particle is described by the energy loss spectrum $w(\epsilon)$. Theoretical values have been calculated by H. Bethe. An energy loss spectrum for heavy charged particles in adenine ($C_5N_5H_5$) is shown in **Fig. 4**. The structure between 3 and 30 eV relates to the outer electrons. Similar structures have been observed for many solids (including metals). Excitation and ionization of the K-shell of C (above 280 eV) and N (above 400 eV) cause further structure. The average energy

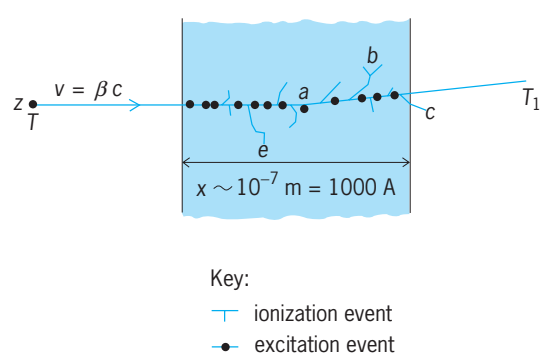

Fig. 3. Energy loss by heavy charged particle. No details shown for energy losses of secondary electrons (delta electrons). Rutherford scattering at point *a*. At point *b*, delta electron experiences collision that results in tertiary delta electron. At *c*, delta electron escapes from absorber.

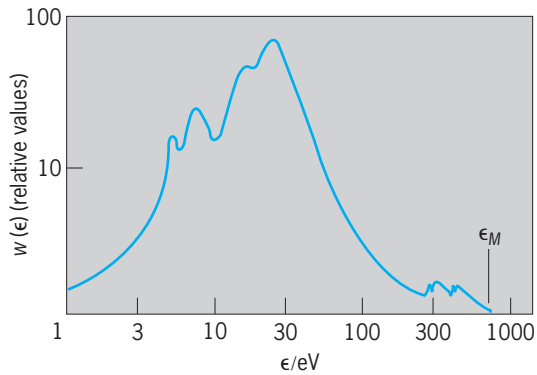

Fig. 4. Schematic single-collision spectrum $w(\epsilon)$ for heavy charged particles in adenine ($C_5N_5H_5$).

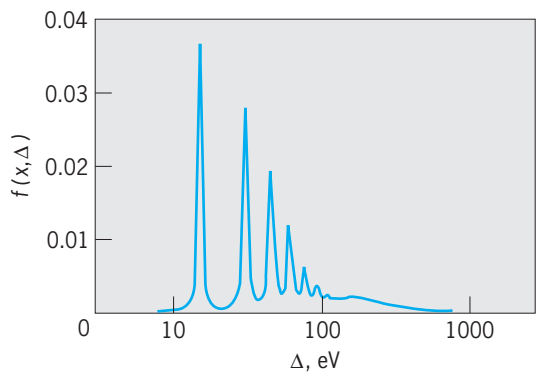

Fig. 5. Calculated straggling curve $f(x, \Delta)$ for 20-MeV protons incident on Al absorber of thickness 5.8×10^{-8} m. Spikes represent multiples of the "plasma loss" at 15 eV. (*After H. Bichsel and R. Saxon, Comparison of calculational methods for straggling in thin absorbers, Phys. Rev., A11:1286–1296, 1975*)

loss per collision is defined by Eq. (3).

$$\langle\epsilon\rangle = \int \epsilon w(\epsilon)\, d\epsilon \Big/ \int w(\epsilon)\, d\epsilon \tag{3}$$

Statistics of energy loss. The total energy loss δ of a particle traveling through matter is the sum of the energy losses ϵ_i in each collision: $\delta = \Sigma\epsilon_i = \epsilon_1 + \epsilon_2 + \epsilon_3 + \epsilon_4 + \cdots \epsilon_v$, where v collisions have occurred, each with a probability given by $w(\epsilon)$ [Fig. 4]. If a large number of particles are observed, they will experience on the average $q = \langle v\rangle$ collisions (q is not an integer), and an average energy loss $\Delta = \langle\delta\rangle = q\langle\epsilon\rangle$, as long as collisions are uncorrelated with another. The number of collisions is distributed according to a Poisson distribution; the fraction $P(v)$ of particles having experienced exactly v collisions is given by Eq. (4). The distribution function for energy losses is called a straggling function $f(\Delta,x)$. Examples are given in **Figs. 5** and **6**. For fairly thick absorbers, $f(\Delta,x)$ is approximately a gaussian distribution of width proportional to $\sqrt{x}$. *See* DISTRIBUTION (PROBABILITY).

$$P(v) = \frac{q^v}{v!}e^{-q} \tag{4}$$

In many applications, the details of energy loss are not important, and a knowledge of the mean or average energy loss Δ is sufficient. If the total collision cross section per atom $w_t = \int w(\epsilon)d\epsilon$ is known (w_t in cm^2), the mean number of collisions is given by $q = xnw_t$, and the mean energy loss by Eq. (5). The quantity stopping power S thus is defined by Eq. (6),

$$\Delta = xn\langle\epsilon\rangle w_t \equiv xS \tag{5}$$

$$S \equiv n\langle\epsilon\rangle w_t \tag{6}$$

in MeV/cm. Since knowledge of w_t and $\langle\epsilon\rangle$ is not extensive, S is frequently determined in experimental measurements in which a beam passes through an absorber, and S is calculated from Eq. (7).

$$S = \lim_{x\to 0}\frac{\Delta}{x} = \lim_{x\to 0}\frac{T - \langle T_1\rangle}{x} \tag{7}$$

Equation (5) is valid only if x is much smaller than the mean range, $R(T)$; otherwise, S varies significantly as the particle loses energy in the absorber. If x is not small, the following procedure can be used to obtain $\langle T_1\rangle$, provided that $R(T)$ has been tabulated: Find $R(T)$ in the range table, calculate $y = R(T) - x$, and then find the energy T_1 corresponding to y in the range table.

Stopping power. The stopping power is a function of the velocity v of the incident particle, its effective charge z^*, and the absorber material. **Figure 7** shows $S(T)$ for heavy ions of various elements in aluminum. S is expressed in MeV cm^2/g. This can be converted to MeV/cm by the formula S(MeV/cm) = ρS(MeV cm^2/g), where ρ is the density of absorber material in g/cm^3. Atomic collisions dominate in region I. For region III, S can be calculated by using the theoretical Bethe expression, Eq. (8), where

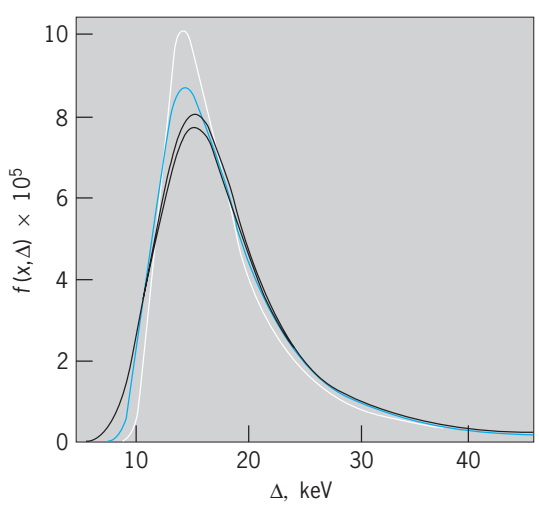

Fig. 6. Calculated straggling curve $f(x, \Delta)$ for 20-MeV protons incident on Al absorber of thickness 3.71×10^{-6} m. Various theoretical calculations are presented. (*After H. Bichsel and R. Saxon, Comparison of calculational methods for straggling in thin absorbers, Phys. Rev., A11:1286–1296, 1975*)

I_A = average excitation energy of absorber (approx-

$$S = \frac{0.30708}{\beta^2}\frac{Z}{A}\left[\ln\frac{2mc^2\beta^2}{I_A(1-\beta^2)} - \beta^2 - \frac{C}{\beta^2} - d(\beta)\right]$$

$$\cdot (z^*)^2[1 + G(z^*, \beta)] \text{ MeV cm}^2/\text{g} \quad (8)$$

imately $Z \times 10^{-5}$ MeV); C = shell correction constant; $d(\beta)$ = density correction, important for $T > Mc^2$; and $G(z^*,\beta)$ = correction due to the second Born approximation, important only for $\beta^2 < 0.01$. *See* PERTURBATION (QUANTUM MECHANICS).

A simpler approximate expression valid to about 10% for $z \leq 10$ in the same region is given by Eq. (9).

$$S = \frac{2.6z^2}{\beta^{1.66}Z^{.25}} \text{ MeV cm}^2/\text{g} \quad (9)$$

$$0.1 < \beta < 0.88$$

Range. A good approximation to the mean range $R(T)$ can be calculated from S using Eq. (10).

$$R(T) = \int_0^T \frac{d\tau}{S(\tau)} \quad (10)$$

Usually, a numerical integration is performed to obtain R from a table of S or the Bethe formula. An approximation formula is Eq. (11). The range of validity

$$R = r \cdot \beta^s \cdot \frac{Z^{.25}}{z^2} \quad (11)$$

$$z \leq 10$$

for this range formula is given in **Table 3**.

Channeling. In absorbers consisting of a single crystal, it has been found that the energy loss will be reduced if the direction of the particle beam coincides with certain preferred alignments of the crystal. It is believed that the particles travel through "open spaces" in the crystal, thus suffering a succession of collisions with relatively small energy losses and angular deflections, and tending to stay in the preferred direction (channel). Thus, even if the total number of collisions q were unchanged, the average energy loss per collision (ϵ) would be reduced, and the total energy loss Δ would be less. *See* CHANNELING IN SOLIDS.

Ionization. The secondary electrons of energies K produced in ionizing electronic collisions will travel through the absorber and will also suffer various collisions, producing further energetic electrons, and so on. This process continues until the electrons have energy $K < I$. It has been found experimentally that the total average number j of ions produced in this

TABLE 3. Range of validity of range formula

Range of kinetic energy T, MeV	Range of β	r, g/cm^2	s
$10 \leq T < 90$	$0.145 < \beta < 0.4$	0.116	1.84
$90 < T < 400$	$0.4 < \beta < 0.7$	0.275	2.33
$400 < T < 1000$	$0.7 < \beta < 0.88$	0.532	3.34

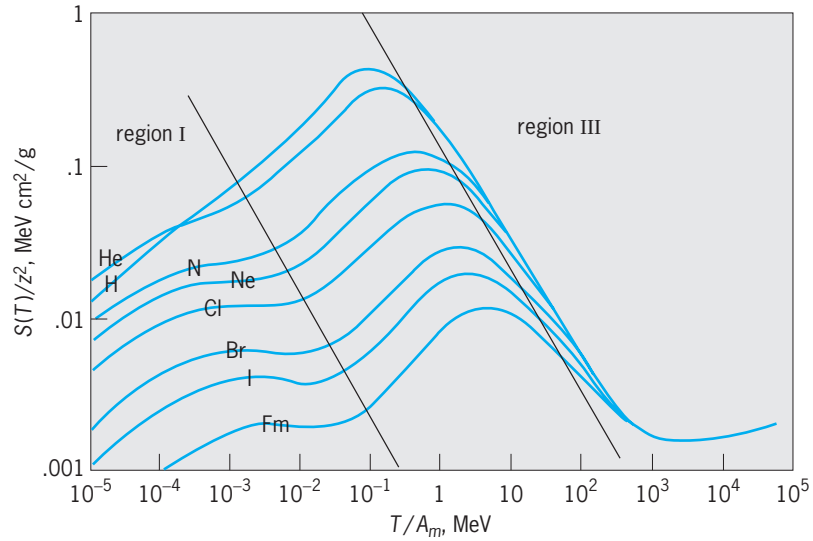

Fig. 7. ***S(T)* for heavy ions in Al. (*After L. C. Northcliffe and R. F. Schilling, Range and stopping power tables for heavy ions, Nucl. Data Tables, A7:233, 1970*)**

way (in the distance x) is proportional to the energy Δ lost by the particle: $j = \Delta/\omega$.

The constant of proportionality ω introduced in this definition has values between 20 and 45 eV for gases, about 3.6 eV for silicon, and 2.96 eV for germanium. It is almost, but not exactly, independent of particle energy and type. If particles lose all their energy in the material, the total ionization J is related to the kinetic energy T: $J = T/W$. The relation between W and ω is $T/W = \int dT/\omega$. $W = \omega$ only if ω is exactly independent of energy.

The ionization j along the path of a single particle increases with the distance traveled, approximately at the same rate as S. For a beam of particles, straggling also influences the total ionization at a given location. The function obtained from the combination (actually the convolution) of both effects is the Bragg curve.

Electrons. Although the interactions discussed earlier all occur for electrons, there are some major differences between electron beams and beams of heavier particles. In general, the path of an electron will be a zigzag. Angular deflections in the collisions will frequently be large. Electron beams therefore tend to spread out laterally, and the number of primary electrons in the beam at a depth x in the absorber decreases rapidly.

Since it is not possible to distinguish individual electrons, it is customary in a collision between two electrons to consider the one emerging with the higher energy as the primary electron. The maximum energy loss in a collision therefore is $mv^2/\sqrt{4}$ (for $T \ll mc^2$). The stopping power expression therefore is somewhat different for electrons. *See* BETA PARTICLES.

Biological effects. In general, for the same dose (the energy deposited per gram along the beam line) heavy charged particles will produce, because of their higher local ionization, larger biological effects than electrons (which frequently are produced by

x-rays). *See* NUCLEAR RADIATION (BIOLOGY); RADIATION BIOLOGY; RADIATION INJURY (BIOLOGY).

Observation. The most direct method of observing a beam of charged particles is to observe the electric current that they form (any flow of electric charges is an electric current). In all accelerators (such as cyclotrons, Van de Graaff generators, linear accelerators, and x-ray tubes) the beam current is measured as a primary monitor of the proper operation of the machine. It is not possible to identify the type of particles with current measurements (except for their electric charge). *See* PARTICLE DETECTOR.

Devices using ionization. If an electric field E is applied to an absorber irradiated with charged particles, the ions and electrons produced will travel in the direction **E,** and the resulting ionization current can be measured (electronic amplification usually is needed). If an oscilloscope is available, the ionization J associated with a single particle can be observed, and the energy (or energy loss) of the particle can be calculated from $T = JW$. Semiconductor detectors (chiefly silicon and germanium) are used extensively for this purpose, but gas-filled ionization chambers have also been used. Cloud chambers and bubble chambers use this principle, but individual ions or clumps of ionization are observed visually. Proportional counters, spark chambers, and Geiger-Müller tubes also operate on the same principle; but in the latter two only the presence of a particle is indicated, and J is not related to T. *See* ELECTRIC FIELD; GEIGER-MÜLLER COUNTER; IONIZATION CHAMBER.

Devices using excitation. The excited state of energy ϵ_e produced in excitation can decay with the emission of light (luminescence or scintillation). Early observations of radioactivity were made with this method (using ZnS screens and visual observation, usually with microscopes), and the method is used extensively with luminescent dials (for example, on wristwatches). The light emitted usually is detected and amplified with photomultipliers. Again, the energy T can be measured. Scintillators used are NaI(Tl), CsI, anthracene, stilbene, and various solids and liquids. *See* LUMINESCENCE; PHOTOMULTIPLIER; SCINTILLATION COUNTER.

A more indirect use of excitations and ionizations is in "chemical" devices (such as photographic emulsions, $FeSO_4$ solutions, and thermoluminescence). *See* PHOTOGRAPHIC MATERIALS; THERMOLUMINESCENCE.

Applications. Electron beams are used in the preservation of food. In medicine, electron beams are used extensively to produce x-rays for both diagnostic and therapeutic (cancer irradiation) purposes. Also, in radiation therapy, deuteron beams incident on Be and ^3H targets are used to produce beams of fast neutrons, which in turn produce fast protons, alpha particles, and carbon, nitrogen, and oxygen ions in the irradiated tissue. Energetic pion (about 100 MeV), proton (about 200 MeV), alpha (about 1000 MeV), and heavier ion beams can possibly be used for cancer therapy. The existence of a Bragg peak for these particles promises improvements in the dose distribution within the human body. *See* RADIOLOGY.

The well-defined range of heavy ions permits their implantation at given depths in solids (this is useful in the production of integrated circuits). Radiation damage studies are performed with charged particles in relation to development work for nuclear fission and fusion reactors. *See* ION IMPLANTATION; RADIATION DAMAGE TO MATERIALS.

Charged particle beams are used in many methods of chemical and solid-state analysis. Nuclear activation analysis can be performed with heavy ions. *See* ACTIVATION ANALYSIS; ELECTRON DIFFRACTION; ELECTRON SPECTROSCOPY; SECONDARY ION MASS SPECTROMETRY (SIMS).

Isotopes can be produced with fast charged ions. *See* BOHRIUM; DUBNIUM; HASSIUM; LAWRENCIUM; MEITNERIUM; MENDELEVIUM; NOBELIUM; NUCLEAR CHEMISTRY; RADIOISOTOPE; RUTHERFORDIUM; SEABORGIUM; TRANSURANIUM ELEMENTS.

Hans Bichsel

Radioactive secondary beams. Beams of nuclei with lifetimes as short as 10^{-6} s are used for studies in nuclear physics, astrophysics, biology, and materials science. Nuclear beams (or heavy-ion beams) are usually produced by accelerating naturally available stable isotopes. However, radioactive nuclei, most of which do not occur naturally on Earth, must be produced as required in nuclear reactions by using various accelerated beams. Because these radioactive nuclei are produced by the nuclear reactions of primary beams, they are called secondary particles and beams of such nuclei are called radioactive secondary beams. *See* RADIOACTIVITY.

Two methods can be used to produce radioactive beams; the methods differ in the way in which the ions of radioactive nuclei are extracted. The so-called isotope separator on line (ISOL) is used to select product nuclei. In this method, nuclei are extracted from a production target by thermal diffusion and then ionized and accelerated to a low energy (approximately 50 keV). The extracted ions are then selected in an isotope separator that consists of magnetic and electrostatic elements; the selected beam is then accelerated to the desired energy. The second method does not use an ion source but rather collects directly the recoil products from the production of nuclear reaction. A recoil mass separator is used to select the desired product. In many cases, these recoil products can be used directly in experiments. *See* ION SOURCES; MASS SPECTROSCOPE.

The recoil separator method has several advantages when applied to the separation of nuclei produced in heavy–ion induced reactions. In such reactions, the velocity of the center-of-mass system is large and product nuclei are emitted in a small-angle forward cone. In particular, nuclei that are produced by projectile fragmentation are emitted in the direction of the primary beam with almost the beam velocity.

For production and separation of radioactive nuclei, a heavy projectile of energy higher than a few

hundred megaelectronvolts per nucleon results in high collection efficiency. For an experiment that requires good momentum resolution, the separated, secondary beam may be injected into a storage-cooler ring in which the radioactive nuclei are cooled to have a relative momentum broadening less than 10^{-3}. Such a cooler ring can also be used to decelerate or to accelerate the radioactive beam for experiments at various energies.

Radioactive secondary beams have made possible the study of the structure of nuclei far from stability. As an example, the sizes of neutron dripline nuclei such as lithium-11 (^{11}Li) and beryllium-14 (^{14}Be) were determined only after beams of these nuclei become available. Furthermore, the internal motion of nucleons inside a radioactive nucleus can be studied in reactions involving the radioactive nucleus as a projectile. Another important application occurs in the study of nuclear reactions of importance in hot stars and in supernovae. Such reactions, involving radioactive nuclei, are crucial for understanding nucleosynthesis in the universe. *See* EXOTIC NUCLEI; NUCLEAR STRUCTURE; NUCLEOSYNTHESIS. Isao Tanihata

Bibliography. *American Institute of Physics Handbook*, 3d ed., 1972; F. H. Attix, *Introduction to Radiological Physics and Radiation Dosimetry*, 1986; R. D. Evans, *The Atomic Nucleus*, 1955, reprint 1982; K. R. Kase et al. (eds.), *The Dosimetry of Ionizing Radiation*, 3 vols., 1985, 1987, 1990; G. F. Knoll, *Radiation Detection and Measurement*, 3d ed., 1999.

Charged particle optics

The branch of physics concerned with the motion of charged particles under the influence of electric and magnetic fields.

Acceleration of charged particles. A positively charged particle that moves in an electric field experiences a force in the direction of this field. If the particle falls in the field from a potential of U volts to a potential zero, its energy gain, measured in electronvolts, is equal to the product of U and the particle's charge. For example, if a singly and a doubly charged particle are accelerated by a potential drop of 100 V, the two particles will gain energies of 100 eV and 200 eV, respectively. If both particles were initially at rest, they would have final velocities proportional to the square root of K/m, where K is the energy increase and m is the mass of the particle. This relation describes the velocities of energetic particles accurately as long as these velocities are small compared to the velocity of light $c \approx 300{,}000$ km/s (186,000 mi/s), a speed that cannot be exceeded by any particle. *See* ELECTRIC FIELD; ELECTROSTATICS.

If an ensemble of ions of equal energies but of different masses is accelerated simultaneously, the ion masses can be determined from their arrival times after a certain flight distance. Such time-of-flight mass spectrometers have successfully been used, for instance, to investigate the masses of large molecular ions, up to and beyond 350,000 atomic mass units. *See* TIME-OF-FLIGHT SPECTROMETERS.

Deflection of charged particles. If a homogeneous electric field is established between two parallel-plate electrodes at different potentials, a charged particle in the space between the electrodes will experience a force in the direction perpendicular to them. If initially the particle moved parallel to the electrodes, it will be deflected by the electric force and move along a parabolic trajectory. Magnetic fields also deflect charged particles. In contrast to electrostatic fields, however, magnetic fields change only the direction of a particle trajectory and not the magnitude of the particle velocity. A charged particle that moves perpendicular to the field lines of a homogeneous magnetic field moves through a circular path (**Fig. 1**) whose radius is determined from the balance between the centrifugal force along this circle and the magnetic force.

Charged particles that enter a magnetic field thus move along circles whose radii increase with the products of their velocities and their mass-to-charge ratios, m/q. If initially all particles start at the same potential U and are accelerated to the potential zero, they will move along radii that are proportional to the square root of $U(m/q)$. Thus, particles of different mass-to-charge ratios can be separated in a magnetic sector field (Fig. 1).

In the plane of deflection, charged particles that diverge from a point in some source are focused back to a point of the image (Fig. 1). In the perpendicular direction, however, the particles remain divergent. A straight line that connects the source and the image also passes through the center of curvature of the trajectory of the reference particle. The images of particles of different mass-to-charge ratios all lie in a plane that is usually inclined with respect to the particle beam.

A sector-field mass analyzer can be used to determine the masses of atomic or molecular ions in a cloud of such ions. Such systems can also be used to purify a beam of ions that are to be implanted in semiconductors in order to fabricate high-performance

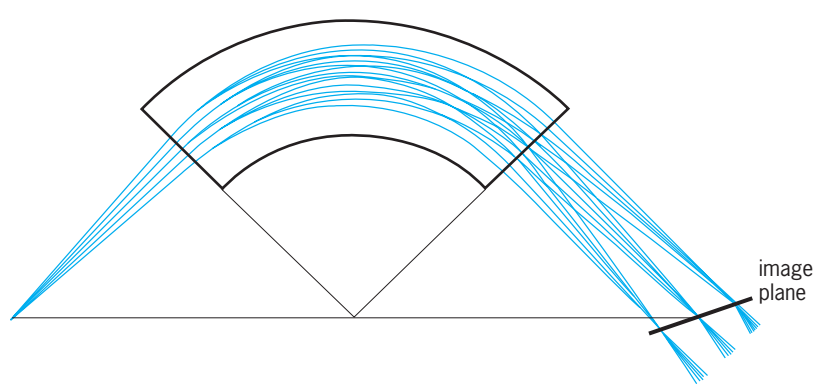

Fig. 1. Sector magnet that separates charged particles according to the products of their velocities and their mass-to-charge ratios. The straight line that connects the particle source and the image of the middle ion bundle also passes through the center of curvature of the corresponding central trajectory.

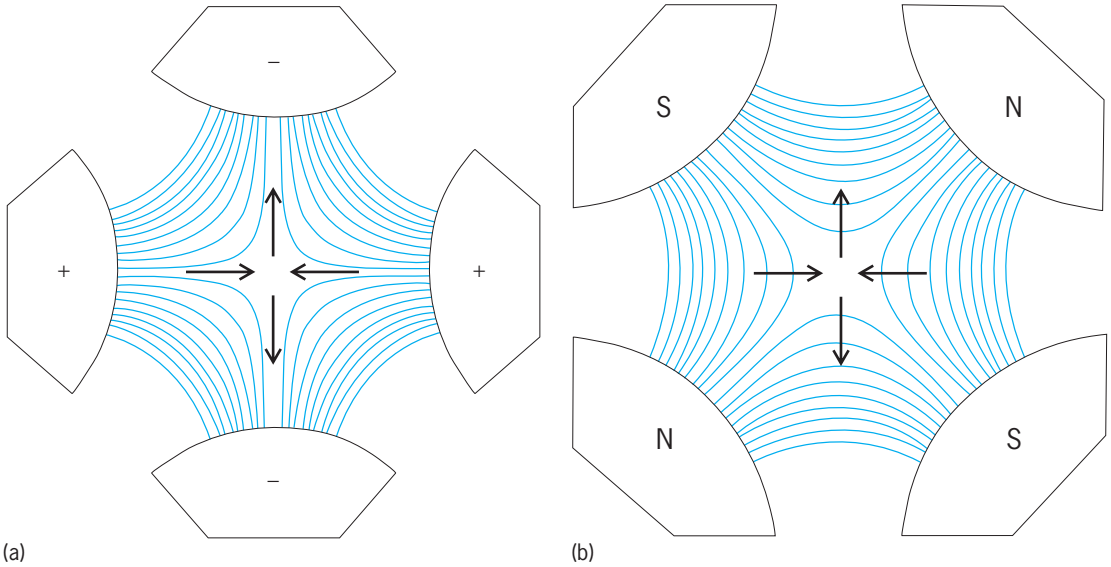

Fig. 2. Cross sections of *(a)* electric and *(b)* magnetic quadrupole lenses. Charged particles move through these lenses perpendicular to plane of drawing. Electric and magnetic field lines are shown. Arrows indicate directions of forces that drive charged particles toward or away from the optic axis. In both lenses, focusing occurs in the horizontal direction, while there is defocusing in the vertical direction.

transistors and diodes. Finally, such magnetic sector fields are found in large numbers in all types of particle accelerators. *See* ION IMPLANTATION; MASS SPECTROSCOPE.

Axially symmetric lenses. An Einzel lens consists of three cylindrical tubes, the middle one of which is at a higher potential than the outer two. Positively charged particles entering such a device are first decelerated and then accelerated back to their initial energies. Since the decelerations and accelerations are perpendicular to the eqipotential surfaces, the particles experience forces that drive them toward the optic axis in some regions and away from it in others. However, the focusing actions are always stronger than the defocusing actions since they occur in regions in which the particles are moving more slowly, that is, in the center regions. Defocusing lenses cannot be built in this fashion unless potential-defining metal grids distort the equipotential surfaces. This statement also holds for axially symmetric lenses in which the potential on the center electrode has been chosen such that the charged particles are not decelerated but accelerated while passing through the lens. *See* ELECTROSTATIC LENS.

Axially symmetric magnetic lenses have also been constructed. Such lenses, also called solenoids, consist mainly of a coil of wire through which an electric current is passed. The charged particles are then constrained to move more or less parallel to the axis of such a coil. Steel tubes with rotationally symmetric geometries can be used to surround the coil to increase the magnetic field strength. *See* MAGNETIC LENS; SOLENOID (ELECTRICITY).

Axially symmetric electric and magnetic lenses are used extensively to focus low-energy particle beams. Particularly important applications are in television tubes and in electron microscopes. *See* CATHODE-RAY TUBE; ELECTRON MICROSCOPE; PICTURE TUBE.

Quadrupole lenses. By passing charged particles through electrode or pole-face arrangements (**Fig. 2**), a particle beam can also be focused toward the optic axis. In such quadrupole lenses the electric or the magnetic field strengths, and therefore the forces that drive the charged particles toward or away from the optic axis, increase linearly with the distance from the axis. This is the main property of the lens. In quadrupoles the focusing occurs in one direction only, while there is a defocusing action in the perpendicular direction. Focusing in both directions thus requires a combination of at least two quadrupole lenses. In quadrupole lenses the particle trajectories are bent toward the optic axis by the main field, while in axially symmetric lenses such a focusing action is achieved by the much weaker fringing fields. Quadrupole lens systems thus have much stronger focusing abilities than other types of lenses, and smaller field strengths can be used to focus more rigid particle beams.

While quadrupole lenses are found in systems in which low-energy particle beams must be focused, for instance, in mass spectrometers, such lenses have become indispensable for high-energy beams. Consequently, quadrupole lenses, especially magnetic ones, are found in many types of particle accelerators used in research in, for example, nuclear and solid-state physics, as well as in cancer irradiation treatment facilities. *See* CHARGED PARTICLE BEAMS; ELECTRON LENS; ELECTRON MOTION IN VACUUM; PARTICLE ACCELERATOR.

Hermann Wollnik

Bibliography. D. A. Dewolf, *Basics of Electron Optics*, 1990; P. W. Hawkes and E. Kasper, *Principles of Electron Optics*, vols. 1 and 2, 1989, vol. 3, 1994; H. Wollnik, *Optics of Charged Particles*, 1987.

Charles' law

A thermodynamic law, also known as Gay-Lussac's law, which states that at constant pressure the volume of a fixed mass or quantity of gas varies directly with the absolute temperature. Conversely, at constant volume the gas pressure varies directly with the absolute temperature. J. A. Charles and J. L. Gay-Lussac independently discovered the relation for an ideal gas. The relation is a useful and close approximation. *See* GAS; KINETIC THEORY OF MATTER.

Frank H. Rockett

Charm

A term used in elementary particle physics to describe a class of elementary particles.

Theory. Ordinary atoms of matter consist of a nucleus composed of neutrons and protons and surrounded by electrons. Over the years, however, a host of other particles with unexpected properties have been found, associated with both electrons (leptons) and protons (hadrons).

Leptons. The electron has as companions the mu meson (μ) and the tau meson (τ), approximately 200 times and 3700 times as heavy as the electron, respectively. These particles are similar to the electron in all respects except mass. In addition, there exist at least two distinctive neutrinos, one associated with the electron, υ_e, and another with the mu meson, υ_μ. In all, there are five or six fundamental, distinct, structureless leptons. *See* LEPTON.

Hadrons. A similar but more complex situation exists with respect to the hadrons. These particles number in the hundreds, and unlike the leptons they cannot be thought of as fundamental. In fact, they can all be explained as composites of more fundamental constitutents, called quarks. It is the quarks which now seem as fundamental as the leptons, and the number of quark types has also increased as new and unexpected particles have been experimentally uncovered. The originally simple situation of having an up quark (*u;* charge $+^2/_3$) and a down quark (*d;* charge $-^1/_3$) has evolved as several more varieties or flavors have had to be added. These are the strange quark (*s;* charge $-^1/_3$), with the additional property or quantum number of strangeness ($S = -1$), to account for the unexpected characteristics of a family of strange particles; the charm quark (*c;* charge $+^2/_3$), possessing charm ($C = +1$) and no strangeness, to explain the discovery of the J/ψ particles, massive states three times heavier than the proton; and a fifth quark (*b;* charge $-^1/_3$) to explain the existence of the even more massive upsilon (Υ) particles. *See* HADRON; *J*/PSI PARTICLE; QUARKS.

The quarks and leptons discovered so far appear to form a symmetric array (see **table**). Both the leptons and quarks come in pairs, although the anticipated partner (*t*) of the *b* quark has not yet been found.

Fundamental constituents of matter

Family of particles	Charge	Particles
Leptons	0	υ_e, υ_μ, υ_τ
	-1	e, μ, τ
Quarks	$+^2/_3$	u, c, $(t)^*$
	$-^1/_3$	d, s, b

Observations. The members of the family of particles associated with charm fall into two classes: those with hidden charm, where the states are a combination of charm and anticharm quarks ($c\bar{c}$), charmonium; and those where the charm property is clearly evident, such as the D^+ ($c\bar{d}$) meson and Λ_c^+ (cud) baryon.

Charmonium. In the charmonium family, six or seven states with various masses and decay modes have been identified. Although a detailed understanding of all these experimentally measured properties has not yet been achieved, everything seems to be in qualitative agreement with theoretical expectations.

Bare charm states. There are several identified bare charm states, including the D ($c\bar{d}$) mesons in both the $J^P = 0^-$ and $J^P = 1^+$ categories (where J is spin, and P is parity), and the Σ_c^{2+} (cuu) and Λ_c^+ (cud) charmed baryons. Information about the lifetimes of these states has been derived from experiments utilizing the high resolution of emulsions to measure the finite distance traveled by charmed particles. The lifetimes of the Λ_c^+ and D^0 have been determined to be on the order of 10^{-13} s with that of the D^+ about a factor of 7 longer. These values are in good agreement with theoretical expectations.

The Λ_c^+ charmed baryon has been observed to be produced in a variety of interactions, including neutrino-proton, proton-proton, electron-positron, and neutrino-neon reactions. A large number of decay modes have been observed, including Kp, $Kp\pi$, $\Lambda\pi$, $\Lambda\pi\pi\pi$, and $Kp\pi\pi$.

Prospects. Although reasonable progress has been made in the study of charmed states, only a handful of states has been observed. Just as the basic SU(3) symmetry arose from a study of the numerous hadron strange and nonstrange resonances, the complete understanding of charm awaits the uncovering of additional states. *See* ELEMENTARY PARTICLE.

Nicholas P. Samios

Bibliography. H. Fritzsch, *Elementary Particles: Building Blocks of Matter*, 2005; I. S. Hughes, *Elementary Particles*, 3d ed., 1991; G. Kane, *Modern Elementary Particle Physics*, 2d ed., 1993; R. F. Schwitters, Fundamental particles with charm, *Sci. Amer.*, 237(4):56–70, 1977; M. Ye and T. Huang, *Charm Physics*, 1988.

Charophyceae

A group of branched, filamentous green algae, commonly known as the stoneworts, brittleworts, or muskgrasses, that occur mostly in fresh- or brackish-water habitats. They are important as significant

components of the aquatic flora in some locales, providing food for waterfowl and protection for fish and other aquatic fauna; as excellent model systems for cell biological research; and as a unique group of green algae thought to be more closely related to the land plants.

Morphology. Charophytes are multicellular, branched, macroscopic filaments from a few inches to several feet in length. Colorless rhizoidal filaments anchor the plants to lake bottoms and other substrates. The main filaments are organized into short nodes forming whorls of branches, and much longer (up to 6 in. or 15 cm) internodal cells. The general morphology varies with environmental conditions such as depth of the water, light levels, and amount of wave action. Reproductive structures occur at the nodes and consist of egg cell–containing structures, the nucules, and sperm cell–containing structures called globules. The biflagellated sperm cells are produced in antheridial filaments within the globules. In many charophytes, calicium carbonate (lime) is secreted on the cell walls, hence the name stoneworts or brittleworts. In some charophytes, the simple structure of nodal and intermodal cells is complicated by corticating elements that cover the cells of the main axis.

Taxonomy and phylogeny. Based on the morphology of the vegetative filaments and the reproductive structures, six extant genera are recognized: *Chara, Nitella, Tolypella, Nitellopsis, Lamprothamnium*, and *Lychnothamnus*. At the species level there is significant disagreement with regard to the importance and interpretation of morphological features, life cycles, and biogeography, especially for *Chara*. Hence, there is no agreement on how many valid species exist, but the maximum number of living species is probably fewer than 100. Among the special features that support the suggestion that the charophytes are closely related to the land plants are the morphologically complicated reproductive structures, the ultrastructure of the sperm cells, and the type of cell division. Some of these features have led some scientists to include other genera of green algae in this group known as the Charophyceae. Deoxyribonucleic acid (DNA) sequencing indicates that the charophytes are a distinct natural group that should be recognized at some taxonomic level (for example, as an order, Charales, or as a class, Charophyceae).

Life history. The charophytes are haploid organisms. The fusion of the sperm and egg cell produces a diploid resistant zygote which, after dormancy, undergoes meiosis to produce a new haploid generation. Thus, there is no alternation of generations. Asexual zoospores are not produced, but plants can reproduce vegetatively as the rhizoidal filaments spread and develop new plants. *See* REPRODUCTION (PLANT).

Fossils. Because the fertilized egg cells produce resistant, calcified zygotes, fossils (known as gyrogonites) provide a record of these organisms extending back to the Devonian Period (approximately 360–408 million years ago). Many genera that once existed are now extinct. *See* ALGAE; CHLOROPHYCOTA; PALEOBOTANY.

Russell L. Chapman

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae: Structure and Reproduction*, 2d ed., 1985; A. Sze, *A Biology of the Algae*, 3d ed., 1997.

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

4 CHE-COS

www.MHEST.com

Cheese

A food product selectively concentrated from major milk components. It is generally rich in flavor and contains high-quality nutrients. There are many varieties of cheese, all produced in the following general manner: Raw or pasteurized milk is clotted by acid, rennet, or both. The curd is cut and shaped into the special form of the cheese with or without pressing. Salt is added, or the cheese is brined after pressing.

Acid is produced during manufacture of cheese by fermentation of the milk sugar, lactose. This fermentation is initiated by the addition of a culture of specially selected acid bacteria (starter culture) to the milk. Acid production in cheese curd retards growth of bacteria that cause undesirable fermentations in cheese. Moreover, it favors the expulsion of the whey and the fusion of the curd particles. Fresh cheese (cottage or cream cheese) does not require any ripening, and it is sold soon after it is made. Other varieties of cheese are cured or ripened to obtain the desirable consistency, flavor, and aroma. The flavor and aroma of cheese are obtained by a partial breakdown of milk proteins and fat by the action of microbial, milk, and rennet enzymes. In hard varieties (Cheddar, Gouda, Edam, Emmentaler or Swiss, and provolone) this is done by the microorganisms in the interior of cheese; in semisoft or soft types (Limburger, Camembert, and Roquefort) by the organisms on, or in contact with, the surface of cheese. *See* MILK.

Cheese Microbiology

Microorganisms are essential to the manufacture of cheese, imparting distinctive flavor, aroma, consistency, and appearance.

Starter cultures. Lactic acid starter cultures are important in making cheese. Different types of lactic acid bacteria (*Lactococcus lactis, L. lactis cremoris, Streptoccus thermophilus*, and homofermentative lactobacilli) are used, depending upon the kind of cheese desired. The starter (both single and mixed strains of bacteria are used) must convert all milk sugar left in the curd into lactic acid within a reasonable time. Several factors may prevent this: the occurrence of natural inhibitors in milk, antibiotics, and bacteriophages.

Some antibiotics (such as nisin) which affect the activity of the starter may be produced by microbes in the milk. Others may be excreted into the milk by cows treated for a particular disease; this happens if penicillin is used to treat mastitis. *See* ANTIBIOTIC.

Bacteriophages, or phages, are viruses which lyse sensitive bacterial cells (**Fig. 1**) and produce new phage particles. Phage particles may get into the milk with infected starter cultures or through contamination with phage-carrying dust particles. Phage

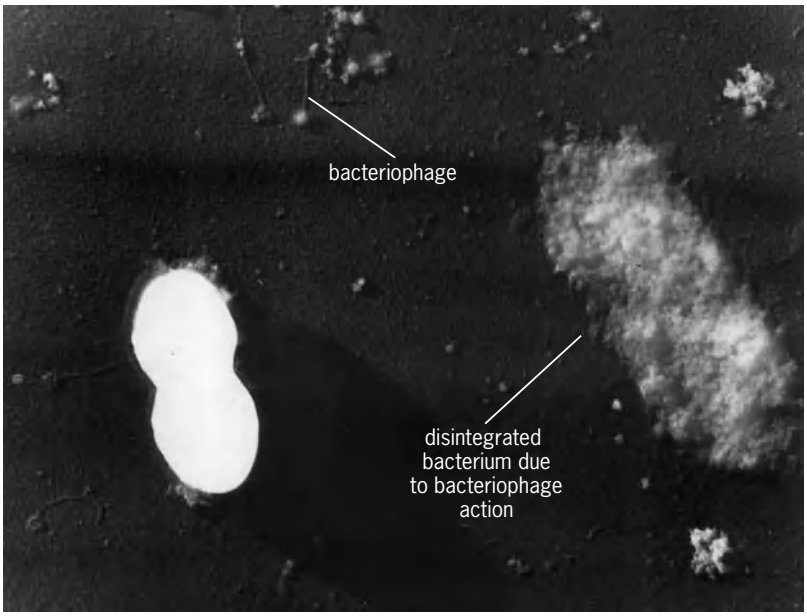

Fig. 1. Electron photomicrograph of *Lactococcus lactis* attacked by phages.

particles slow down or totally inhibit the activity of the starter culture, causing insufficient souring and spoilage of cheese. Phage multiplication is influenced by temperature, pH, and calcium content of the medium. Phage outbreaks cause serious economic losses. They can be controlled by rigorous hygienic handling of starters, by culture rotation, or by culturing starters in calcium-free media. Attempts to isolate phage-resistant strains have not met with success. *See* BACTERIOPHAGE; STREPTOCOCCUS.

Flora of cheese. Cheeses of the hard type contain lactic acid lactococci and other bacteria present in the starter. Moreover, they contain all the microorganisms originally present in the raw milk or pasteurized milk. The types of bacteria most important for ripening are the lactococci, lactobacilli, micrococci, and propionic acid bacteria. They, together with the enzymes of rennet and milk, break down the proteins to peptides and amino acids, hydrolyze the milk fat to fatty acids, and frequently produce carbon dioxide gas which causes the holes in cheese (**Fig. 2**). The structure of cheese is greatly changed during ripening.

In semisoft and soft cheeses the lactic acid originally produced by the starter is broken down afterward by molds, yeasts, and bacteria on the surface. The same flora decomposes the protein and the fat to a much greater extent than in hard cheeses. As a result, strong flavors are produced and the structure of the cheese becomes much softer. For the production of blue-veined cheese, needles are pushed into the interior to bring *Penicillium roqueforti* or the other molds, added during the manufacturing of the cheese, into contact with air they need for growth. The blue color is the color of the mold spores. The red color on the surface of some soft cheeses is caused by *Brevibacterium* (*Bacterium*) *linens* and micrococci. For good surface growth, soft cheeses must be cured in high-humidity air (caves or air-conditioned rooms).

Fig. 2. Swiss cheese with normal eye formation.

Fig. 3. Late blowing in Gouda cheese.

Fig. 4. Evidence of late gas production during the cheese ripening process.

Defects of cheese. In hard cheeses an abnormal gas formation is a problem. It can occur at almost any stage of manufacture or ripening. Early gas production can be caused by coliform bacteria or heterofermentative lactic bacteria. Gas production during cheese ripening (late gas, as in **Figs. 3** and **4**) may be caused by certain sporeforming bacteria, propionic acid bacteria, *Leuconostoc* species, and *Lactococcus lactis lactis* (biovar *diacetylactis*). The use of *L. lactis lactis* (biovar *diacetylactis*) and *L. lactis* as lactic starter cultures can cause fruity flavors, as well as gas formation, in Cheddar cheese. Other defects of hard cheese include insufficient or excessive acidity, various off flavors, and discoloration. Soft cheeses may be spoiled by gas production, excessive acidity, improper development of surface flora, and contamination with atypical molds.

Cheese blending. Cheeses of different ages are blended. The mixture, melted with the aid of emulsifying salts (citrates and phosphates), is packed in sealed containers (tins, paperboard, foil, or plastic). Few bacteria other than sporeformers survive the heat treatment. No substantial growth of the flora takes place in well-preserved processed cheese, but spoilage by anaerobic sporeformers may occur.

Clostridium sporogenes causes putrefaction and slit openness and *C. tyrobutyricum* causes blowing of tins and packages. Acidity, salt content, and temperature of storage are important in controlling spoilage. *See* INDUSTRIAL MICROBIOLOGY. Norman Olson

Cheese Manufacture

More than 150 countries make cheese, but 34 countries produce the major commercial quantities. Primitive cheesemaking exists in much of Africa, Latin America, and southwestern Asia, but advanced techniques are also found in these regions. Cheese manufacturing in developed world areas may be simple too, including that for the sheep-milk Broccio of Corsica, the mountain cheese of the Alps, the queso blanco of Venezuela, and the Pennsylvania pot cheese of the United States, but generally it is highly mechanized. Whether the method is primitive or sophisticated, the quality of the resulting cheese is usually excellent.

The primary objective of cheesemaking is to form a smooth acid curd, to reduce the size of the curd block and remove the whey, with or without prior cooking, and to salt and shape the curds. A different type of cheese emerges when the intensity of approximately eight steps is stressed or minimized as required; when special applications, usually microbial, are introduced; and when the environment is transformed to fit the optimum needs of a specific cheese type.

Materials. Key materials for cheesemaking include fresh or precultured milk, cultures, milk-coagulating enzyme preparations, special microorganisms, salt, and beta carotene or annatto color. The amounts used and the manner in which these materials are applied strongly influence the cheese character. Cheese may be made from the milk of the cow, sheep, goat, water buffalo, and other mammals, but the milk of the cow is most widely used despite some limitations. Sheep milk and water buffalo milk generally give more flavor to the cheeses, and the color is uniformly white because of a lack of carotene in such mammalian milks, but they are more expensive to make into cheese.

Major classes of cheeses. Two major classes of cheeses exist, fresh and ripened. Fresh cheeses are simpler to make than ripened, are more perishable, and do not develop as intense flavors, but give a mild acid, slightly aromatic flavor and soft, smooth texture.

Fresh cheeses. Three basic groups characterize fresh cheese types: group I—ricotta and Broccio; group II—cottage, Neufchâtel, and cream; and group III—mozzarella. Curd formation for these fresh cheese groups results from a combination of acid (pH 6.0) and heat (176°F or 80°C), as in group I; from acid alone to give a pH 4.6, as in group II; or from rennet (an enzyme preparation) at pH 6.3, as in group III. Essentially a dehydration of protein occurs, along with a partial or complete reduction of the negative electrical charges which surround the surfaces of the milk proteins. At a critical point, precipitation occurs, leading to a smooth gelatinous curd of varying strength. The curd may be scooped directly into cloth bags or perforated containers for immediate draining and eventual packaging, or the curd is cut and cooked, followed by drainage of whey, salting, and creaming.

Ricotta. Ricotta, or recooked cheese, in its most acceptable form is made from acidified whole milk heated to 176°F (80°C). By introducing large amounts of lactic starter culture, acid whey powder, food-grade acetic or citric acids, and small amounts of salt to cold milk and heating to the above temperatures in a kettle, smooth white particles appear which rise to the top and collect as a curd bed, which is left undisturbed. In about 30 min the hot curd is scooped into perforated containers for drainage and cooled. Later it is removed and consumer-packaged. The product is utilized directly or in a variety of Italian dishes.

Cottage. Cottage cheese is made from pasteurized skim milk, and in uniform discrete particles classified as small or large curd. A curd forms when the increasing lactic acid of milk during fermentation attains the isoelectric point of casein at pH 4.6. This soft curd additionally contains lactose, salt, and water. Later the curd matrix is cut and cooked to about 126°F (52°C). Separation of whey from the curd is rapid, and is followed by two or three water washings at warm to chill temperatures. Washing removes whey residues and acts as a cooling medium. After drainage of the last wash water, the chilled curd is blended with a viscous, salted creaming dressing to give 4.2% fat and 1% salt, and is packaged. One unsalted, uncreamed, nonfat form of cottage cheese has less flavor, calories, and sodium than standard cottage cheese. Another form is creamed to 9–11% fat and pressed as farmer's pressed cheese. *See* ISOELECTRIC POINT.

Cream and Neufchâtel. Cream and Neufchâtel cheeses resemble cottage cheese in manufacture, but they are made from high-fat mixes instead of skim milk, and no effort is undertaken to attain discrete particles. The minimum fat content of Neufchâtel cheese is 20%, and of cream cheese 33%. The cloth bags traditionally used to separate the whey from the curd have been replaced by centrifugal curd concentrators, which separate the stirred curds and whey at 165°F (74°C). The curd thereafter is pasteurized, treated with stabilizer or gums to prevent water leakage, and salted and homogenized hot. This hot mass is packaged and sealed directly, and tempered at room temperature for a number of hours. The products are used as spreads, in salads, and in cheesecake.

Cottage cheese keeps well at refrigerated temperatures for at least several weeks, but hot-pack cream and Neufchâtel cheeses, because of earlier exposure to high temperature and sealing hot in packages, maintain freshness for up to 2 months or more. Common microbial spoilage agents are molds and yeasts and frequently *Pseudomonas* bacteria which grow at low temperatures.

Mozzarella. Natural mozzarella cheese is normally made from pasteurized whole milk, but milks of 1 or 2% fat may be used if the cheeses are properly labeled. The warm cheese milk is incubated with lactic

acid bacteria (*Lactococcus lactis*, *Streptococcus thermophilus*, or *Lactobacillus delbreuckii bulgaricus*), or may be directly acidified by the addition of food-grade acids such as acetic or citric. Added rennet coagulates the milk to a smooth curd, which is cut with wire knives. The curds, without cooking, may be drained of whey shortly after cutting or, depending upon the type of bacteria, may be cooked to 117°F (47°C) before whey removal. The blocks or patties of curd which result from this action are retained until a pH 5.2–5.5 is achieved, and then are milled into small cubes, placed in hot water, and stretched and molded until smooth. The resulting rectangular blocks are cooled and brined for 2–8 h, depending upon size, and dried and vacuum-packaged.

Acceptable natural mozzarella cheese has a bland, slightly acid flavor. It melts uniformly and smoothly when exposed to hot oven temperatures of 428°F (220°C) for a few minutes and, depending upon type, maintains its quality from several weeks to several months. Natural mozzarella cheese is used in a wide assortment of Italian dishes. Imitations containing vegetable fats, artificial flavors, and a variety of additives are made mainly for institutional and food supplementation use.

Ripened cheese. Ripened cheese is exposed to an optimum environment of temperature and moisture for a period of time to attain characteristic flavor, texture, and appearance. It may be made from raw, heat-treated, or pasteurized milk that is coagulated with a rennet preparation to form a smooth curd. After cutting and cooking to 95–131°F (35–55°C), the curds are salted, before or after pressing, and shaped. The pressed curds, usually salted, may be given a special microbial application, and then are usually held at 41–59°F (5–15°C) in rooms under controlled humidity (85–95%) to develop the desired traits. All standard ripened cheeses are salted, and the milks from which they are made are always coagulated with rennet, usually in the presence of small amounts of lactic acid and sometimes added calcium chloride.

Rennet milk coagulation occurs when rennin, a protease from rennet, hydrolyzes or splits kappa casein from casein into a whey-soluble component, glycomacropeptide. The cleavage of kappa casein by rennin destabilizes the casein molecule, permitting its dominant alpha and beta casein components in micelle form to precipitate. Such coagulated proteins are filamentous and, through strand overlapping under quiescent conditions, appear as a smooth, homogeneous, sweet (pH 6.3) gel or curd. This curd is cut and cooked, leading to a contraction which expresses whey. The dry curds thereafter are salted, pressed, and ripened. *See* MICELLE; RENNIN.

During ripening, many flavor compounds arise in the cheese. They include free fatty acids, amino acids, ketones, diacetyl, and alcohols. Their amounts in a balanced ratio, along with major components of the cheese, largely determine the characteristic flavor. Agents involved through proteolytic, lipolytic, decarboxylation, and deamination reactions include rennin, natural bacteria of milk, added microorganisms, natural milk enzymes, microbial enzymes, and added food-grade enzyme preparations. The microbial enzymes evolve from millions of microbial cells present in cheese; the sources of the added food-grade enzyme preparations usually are fungal.

Most ripened cheeses are contained in one of six basic groups. The characteristic cheeses in each group include: group I—Cheddar and Monterey; group II—Swiss (Emmentaler) and Gruyère; group III—Edam and Gouda; group IV—Muenster, brick, and Limburger; group V—provolone; and group VI—Camembert, Brie, and blue.

Cheddar and Monterey. For these cheeses the curds are formed at 88°F (31°C) by 0.2 qt (200 ml) of rennet in 2200 lb (1000 kg) of underheated whole milk previously inoculated with 1–2% lactic-type starter for 15–30 min. The curds are cut with wire knives, and cooked to 99°F (37°C) for 30–60 min. Then the whey is run off, and the curd bed is cut into rectangular blocks weighing approximately 11 lb (5 kg). These blocks usually are turned over every 15 min for 2 h in a step known as cheddaring, and the stringy nature of the curd evolves as sufficient lactic acid develops to give a pH of 5.3. The blocks, now flatter than at the start, are milled to small cubes, and dry-salted at the rate of 5 lb (2.3 kg) of salt per 220 lb (100 kg) of curd. These salted curds are pressed in single-service paper or cloth-lined stainless steel boxes. The standard 44- to 66-lb (20- to 30-kg) forms which result are removed, and repressed in plastic films of low oxygen permeability. The film-wrapped Cheddar cheese is ripened up to 1 year at 41–50°F (5–10°C). It is generally manufactured into various-size forms prior to reduction into consumer packages.

In modern industrial cheesemaking, another form of Cheddar cheese, known as barrel, is made by the stirred-curd step in contrast to the cheddaring. This avoids the more lengthy period required for the latter. In the stirred-curd step, individual small cut curd particles, after cooking and whey separation, are stirred for about 30 min, salted, and packed into large film-lined metal or fiber barrels of 640-lb (290-kg) capacity. Here, after excess whey is removed, the curds are pressed and the containers sealed and placed in ripening rooms. Later, the cheese is removed and used largely for making processed cheese. Ripening cheese in large blocks appears to improve typical flavor quality, so that some cheese that is cheddared and destined for long ripening, too, may be held in 640-lb (290-kg) capacity film-lined, sealed boxes.

Monterey cheese resembles Cheddar, but has a higher moisture level and ripens more rapidly. Its curds are washed with water in the vat and not cheddared, but are stirred prior to salting and then added to lined standard stainless steel forms for pressing and eventual ripening.

Swiss (Emmentaler) and Gruyère. Swiss and Emmentaler cheeses are the same, but in the United States the term Swiss is widely used, whereas in Europe and elsewhere the name is Emmentaler. The cheeses are made in block or round forms, but in the United States the former shape dominates because it is easier to slice for sandwiches. Also, economics of production are enhanced.

Underheated milk (158°F or 70°C for 16 s), standardized or partly skimmed, is set with rennet to form a curd, accompanied by three special bacterial cultures, *Streptococcus thermophilus, Lactobacillus delbrueckii bulgaricus*, and *Propionibacterium shermanii*. These cultures influence the acid, eye, flavor, and texture development of the cheese. *Propionibacterium shermanii* is mainly responsible for eye and typical flavor formation.

Renneting requires about 30 min by using 0.2 qt (180 ml) of rennet per 2200 lb (1000 kg) of warm (88°F or 31°C) milk. The size of the curds after cutting is very small, about like peas. These are cooked to about 126°F (52°C) in 60 min, and stirred until the pH decreases to 6.3.

To make block Swiss cheese, these stirred, cooked curds and considerable whey are pumped into cloth-lined, perforated stainless steel chambers. Here a large block is formed under the warm whey, the whey is removed, and the bed of curd remaining is cut into about 100-lb (45-kg) sections. These are immersed in cool brine for up to several days. After removal and drying, the blocks are sealed with plastic or rubberized film, placed in wooden boxes, with room for expansion, and covered by lids. Exposure to warm temperature, about 73°F (23°C) for 2–3 weeks, produces the typical eyes and sweet, hazelnut flavor. Further ripening continues at 41°F (5°C) for 4–12 months before the cheese is ready for market.

Gruyère cheese shows eyes like Swiss or Emmental, but they are usually smaller. This 125-lb (57-kg) round wheel cheese is made mostly in France, Switzerland, and Finland. The surfaces of Gruyère cheese, unlike the larger 224-lb (102-kg) round wheel Emmentals, are not washed daily or wiped with cloths. This leads to a special microbial growth on the surface which imparts to the cheese more flavor of a unique nature best typified by the Comté Gruyère of the French Jura area.

Edam and Gouda. Both of these cheese types may contain eyes, but these are small and lack uniform distribution. Edam cheese is produced in its normal size as 5-lb (2.3-kg) round balls waxed in red. Gouda cheese, softer in body because whole milk is invariably used, appears as orange- or yellow-waxed, medium-sized flat wheels of about 11–22 lb (5–11 kg). Manufacturing is essentially similar for both.

Pasteurized milk is set with lactic culture and rennet, but acid development is constrained usually by removing some of the whey and replacing it with water. Before the cheese is pressed, the pH is usually 5.4. This higher pH gives a sweeter flavor to the cheese and permits other species of bacteria to function. In Europe, where most of this cheese is made, sodium nitrate may be added to the cheese milk to suppress spoilage bacteria, but the practice is diminishing, and instead a continuous bacterial spore centrifugal removal process is substituted. Where nitrates are used, only small amounts are introduced, most of which disappear during ripening.

Unique metal forms shape the cheese after light pressing. The cheeses are brined and waxed. Ripening occurs at 50°F (10°C) for 2–3 months.

Muenster, brick, and Limburger. Good-quality pasteurized milk is required for these cheeses. At times, as for brick cheese, the renneted curd may be washed with water to reduce the acidity and to increase moisture. A reddish bacterium, *Bacterium linens*, along with yeasts, grows on the cheese surfaces at 59°F (15°C) in a highly moist room for 11 days. The surfaces take on the brick-reddish color of the bacterium. The three cheeses are differentiated by the intensity of their flavor, softness of body, and sizes and shapes. Muenster has least intensity of flavor, followed by brick cheese.

Provolone. Originating in Italy, provolone cheese also is manufactured in large quantities in Argentina and the United States. It is a pasta filàta, or pulled-curd ripened cheese, and is generally smoked. The texture is smooth and flaky and the flavor highly aromatic or piquante. Initially, the cheese is made almost exactly like that of a fresh, low-moisture mozzarella using *Lactobacillus delbrueckii bulgaricus* as the starter. Following proper acid development in the vat, the curd is stretched and formed into various shapes: cylindrical, pear, or ball. Thereafter, it is bound with ropes or twine and hung in hickory smoke–filled rooms for 3 weeks or exposed to smoke compounds. Ripening occurs at about 44°F (7°C) for up to 12 months, and cheese sizes range from 3 to 500 lb (1.4 to 227 kg). A related Italian type is caciocovàllo, literally meaning horse cheese.

Camembert, Brie, and blue. These are all rennet cheeses, and Camembert and Brie are made with pasteurized milk cultured to about 0.2% titratable acidity. The curds are not cooked and, following whey drainage and hooping in round forms, they are modified by the special application of white or blue-green mold spores. Introduction of the spores to the cheese may result from inoculating the milk, brine, or curds, or inoculating all three media.

Camembert and Brie are white mold cheeses about 1 in. (2.5 cm) thick whose manufacture is similar except that the diameter of the latter is larger. The characteristic feature of these cheeses is the appearance of a snow-white mat of *Penicillium caseicolum* or *P. camemberti* mold on the surfaces. These molds grow well only in an excessive air environment, and hence their appearance on the surface. They ripen the cheese from outside into the center through release of their many active enzymes, which transform the curds to a smooth, soft state. The cheeses are ready for consumption in 12–14 days at 50°F (10°C) and 95% relative humidity. Held thereafter in a refrigerated state for 40 days in supermarkets, the cheeses eventually become overripe. Ammoniacal flavor with the white surfaces becoming brown and the inner sections becoming light tan confirm an overripe state.

Blue cheese requires spores of *Penicillium roqueforti* or *P. glaucum*. These spores, appearing as a black powder, are present throughout the wheel of pressed, highly salted curd, and grow as filamentous blue-green mycelia when minimum air is supplied. In practice, the mold-inoculated wheels of cheese, from 5 to 12 lb (2.3 to 5.4 kg), are penetrated by a

mechanical head containing 50 needles which create the long air channels. The mold grows along these air channels, and moves farther into the cheese if natural openings exist. Here the many lipolytic and proteolytic enzymes originating from the mold mycelia attack the cheese to give characteristic flavor. Molding occurs in 30 days after air channel formation at 50°F (10°C) and 95% relative humidity. Then the cheese is held at 41°F (5°C) during 2–6 months for more complete ripening.

The most famous blue cheese is Roquefort, made in southern France from sheep milk. Other widely known related cheeses are Stilton, Gorgonzola, Danish blue, and American blue.

Processed cheese. Processed cheese is made from natural types. Nearly any natural cheese can be processed, except that for blue cheese technical difficulties cause blackening of the blue mold due to the high heat.

Processed cheese is produced by grinding selected lots of natural cheese and adding cream, color, salt, and emulsifying agents. The mixture is blended and sent to cookers which heat the mass to 165–185°F (74–85°C). During the cooking, fat normally separates from the protein and water, but this is corrected by adding anhydrous citrate and phosphate salts, usually at 3% levels. The salts raise the cheese pH to about 5.6–5.8, making the protein more soluble. Under these circumstances a stable emulsion of protein, fat, and water occurs to give a smooth homogeneous mass. The hot mass is packaged directly. Processed cheese foods and spreads are made similarly, but include more water and less protein, fat, and other solids.

Sliced processed cheese is attained by spreading the hot process cheese mass emerging from the cookers over moving cold steel rolls. The chilled cheese in thin, wide sheets passes over rotary knives and cutting bars to give squares lying on top of each other, which are packaged in film wrappers.

Production developments. The cheese industries of the world have become rapid-growth industries with high consumer acceptance of their product. Cheese-manufacturing plants may be very large, utilizing more than 2.2×10^6 lb (10^6 kg) of cheese milk daily. In the operational areas, cheese vats of 5500-lb (2500-kg) capacity are common, as are giant presses.

Mechanization of cheesemaking too has advanced through the introduction of ingenious cheddaring towers, block formers, and moving web belts. This has led to significantly higher rates of production, with less manual labor, while maintaining quality. True continuous making methods of natural cheese have yet to be realized. However, a unique process incorporating ultrafiltration membrane technology has potential for continuous manufacture of certain cheeses.

Membrane separation is an application of biotechnology to cheesemaking. In this process, skim or whole milk is pumped at low pressure, 48 lb/in.^{2} (3.4 kg/cm^2), through tubes or plates containing many small pores, with a molecular-weight cutoff of 20,000. The milk passes across the membranes, usually at 126–129°F (52–54°C). Much of its water, lactose, soluble salts, and nonprotein nitrogen pass through the pores as permeate. Retained are all the fat, protein, insoluble salts, and some serum, which contains water, lactose, soluble salts, and nonprotein nitrogen. This retentate increases in concentration with time, so that its fat and protein may be five times greater than that of the starting milk. Emerging from the ultrafiltration unit in its final concentrated form as a plastic fluid, this retentate has the same composition as some cheeses. Cheesemaking is completed without vats, as there is no whey to separate. Simple mechanical injection of culture, rennet, salt, color, and fungal spores into the plastic fluid concentrate may follow. A curd develops within 5–10 min, and is placed directly in a ripening room. This is a French process, known as MMV after the names of the inventors, Maubois, Mocquot, and Vassal. Its major advantages are much greater cheese yields due to complete retention of whey protein in the cheese, lower requirements for rennet, a high potential for continuous-process application, and production of a neutral-pH permeate with a reduced biochemical oxygen demand (BOD) value. Much cheese in Europe is made in this manner, particularly fresh, soft acid cheeses—Feta, Camembert, and St. Paulin.

Other variations of ultrafiltration retentate application to cheesemaking include milk retentate supplementation and direct ultrafiltration of cheese milks, but only up to 2:1 protein concentration. The advantages of this alternative ultrafiltration process include improved cheesemaking efficiency, energy savings, and higher quality of marginal cheeses. Use of milk retentates in cheesemaking in this manner has led to the production of natural Cheddar cheese of good quality with significantly reduced sodium levels, 0.015 oz/3.5 oz cheese (420 mg/100 g) or less, compared with approximately 0.02 oz (700 mg) for commercial Cheddar cheeses.

Ultrafiltration has led to whey protein concentrates of 90–95% protein. This nonfat product can be modified to give a texture similar to cheese and is used as a food ingredient to replace butterfat by up to 50% for the making of commercial low-fat cheese. Another technology for the fluid milk and cheese industries is microfiltration. This system follows the same principles as those for ultrafiltration except for membrane pore size. Ultrafiltration retains fat, casein, bacteria, and insoluble salts. Passage through a micromembrane retains fat, bacteria, and insoluble solids, but no protein, making the resulting permeate an almost sterile skim milk. The creamlike retentate obtained is heated to 302°F (150°C) for 2 s, and then it is blended with the almost-sterile skim milk, pasteurized, and cooled to proper temperatures for use in milk and cheese processing. Results from Europe, where much research has been conducted on this subject, indicate an almost 99.5% reduction in total bacteria. Also, quality cheese without gas defects or loss of flavor was produced. *See* ULTRAFILTRATION.

Frank V. Kosikowski

Bibliography. J. G. Davis, *Cheese*, vol. 3: *Manufacturing Methods*, 1976; F. V. Kosikowski and V. V. Mistry, *Cheese and Fermented Milk Foods*, 3d ed., 1997; G. P. Sanders, H. E. Walter, and R. T. Tittsler,

General Procedure for Manufacturing Swiss Cheese, USDA Circ. 851, 1955; A. L. Simon, *Cheese of the World*, 1956; U.S. Department of Agriculture, *Cheeses of the World*, 1972; L. L. Van Slyke and W. V. Price, *Cheese*, 1980.

Cheilostomata

An order of ectoproct bryozoans (in which the anus opens outside the ring of tentacles) in the class Gymnolaemata. Cheilostomes possess delicate colonies composed of loosely grouped, highly complex, short, uncurved, box- or vase-shaped zooecia (chamber in which the animal lives), with solid, porous, chitinous, or calcareous walls, and with apertures closed by lidlike opercula. *See* BRYOZOA; GYMNOLAEMATA.

Morphology. Cheilostome colonies (**Fig. 1***a–e*) may be encrusting threadlike networks (*Pyripora*); thin encrusting sheets (*Membranipora*, *Steginoporella*); tabular, nodular, or globular masses (*Schizoporella*, *Celleporaria*); free-living (unattached) domes (*Cupuladria*); or erect tuftlike (*Bugula*, *Scrupocellaria*), twiglike (*Vincularia*), frondlike (*Coscinopleura*, *Flustra*), or trellislike (*Retepora*, *Sertella*) growths. These colonies are not divisible into distinct endozone (thin-walled, fragile, inner or lower portion of colony) and exozone (thick-walled, sturdy, outer or upper portion of colony) regions. All cheilostome colonies lack stolons (elongated projections of the body wall from which buds are formed, giving rise to new zooids) and monticules (or maculae; raised formations of zooids on the colony surface that control incurrent and excurrent flow). Some have rhizoids (individual rootlike outgrowths that reinforce the colony's attachment to a substrate) or small coelomic chambers, and most cheilostome colonies bear ovicells (or ooecia; brood chambers that protect developing embryonic larvae) developed at the distal ends of the normal autozooecia (feeding zooids; **Fig. 2**). Autozooecia possess several tentacles arranged in a circle (lophophore) around the mouth and held out to form a bell or funnel when protruded for feeding. Many cheilostome colonies are polymorphic, that is, they possess either one or both of two types of specialized zooids (heterozooids): avicularia, which are bird-head-like or mouse-trap-like heterozooids with a movable mandible that can be snapped at small animals attempting to settle on the colony, and vibracula, heterozooids with a bristlelike seta that sweeps animals and debris from the colony surface.

The walls of adjacent zooecia within a cheilostome colony are distinctly separate rather than fused together, and may be perforated by small openings filled with soft tissue. The zooecial walls may be thin chitinous membranes, moderately thick and firm chitinous walls, or thin to thick calcareous walls (in most species, they are made of calcite), displaying a great variety of pores, pits, granules, spines, and lacy fretwork (especially the frontal walls, which form the external surface of the colony). Zooecia lack internal transverse partitions (diaphragms). Round to

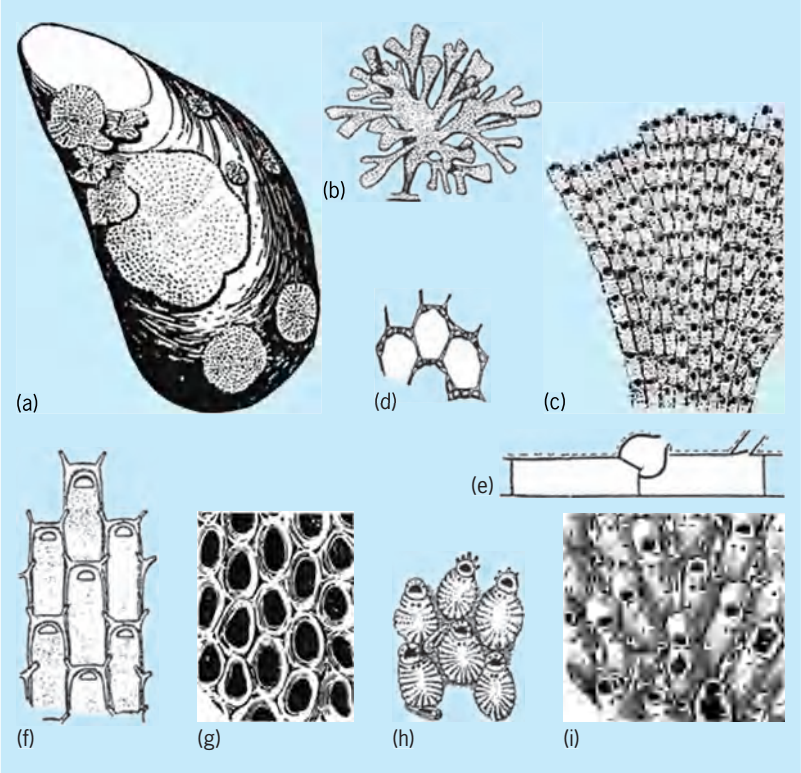

Fig. 1. Cheilostome colonies and zooecia (scale of colonies, ×1). (*a*) Encrusting sheets of *Membranipora* and *Electra* on bivalve shell. (*b*) Erect, flat branches of *Chartella*. (*c*) Upper (frontal) surface section (scale, ×8) of encrusting sheet of *Schizoporella*. (*d*) Tangential section (scale, ×20) of *Escharoides*. (*e*) Longitudinal section (scale, ×25) of an unnamed cheilostome. (*f*) Soft membranous frontal wall of *Membranipora* (scale, ×35). (*g*) Large open opesium of *Conopeum* (scale, ×25). (*h*) Cribrimorph frontal wall of *Cribrilaria* (scale, ×25). (*i*) Bulbous ovicells of *Diplotresis* (scale, ×25). (*Parts a, c, and f reprinted with permission from R. J. Cuffey and J. E. Utgaard, Bryozoans, in R. Singer, ed., Encyclopedia of Paleontology, Fitzroy Dearborn, 1999; all other images reprinted with permission from R. S. Bassler, Treatise on Invertebrate Paleontology, © 1953 Geological Society of America and University of Kansas Press*)

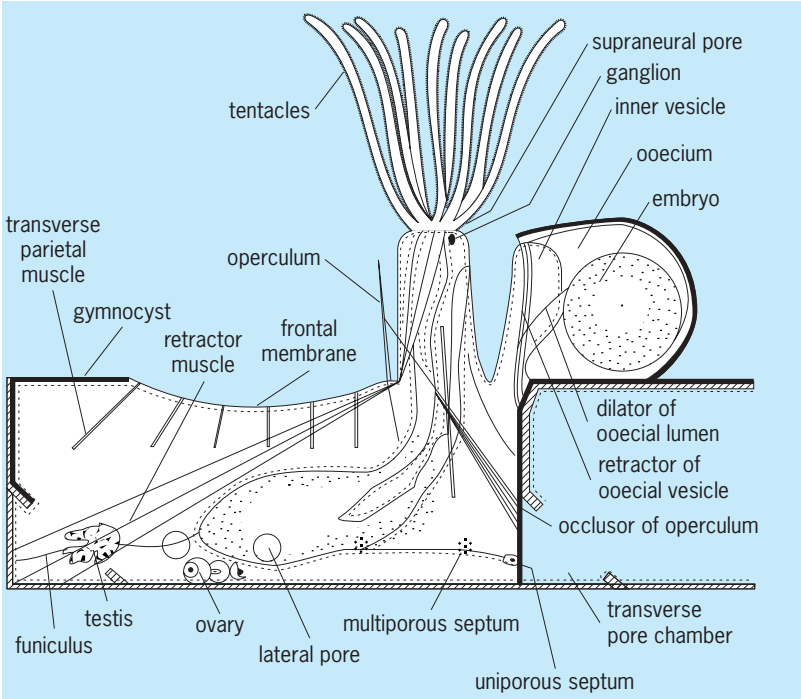

Fig. 2. Anascan cheilostome zooid, showing soft parts and zooecium, with tentacles extended. (*Reprinted with permission from R. J. Cuffey and J. E. Utgaard, Bryozoans, in R. Singer, ed., Encyclopedia of Paleontology, Fitzroy Dearborn, 1999*)

highly complicated in outline, and noticeably smaller in transverse diameter than the zooecium itself, the aperture is closed by a membranous or chitinous lid (operculum) hinged to the proximal edge of, and fitting snugly over, the aperture. The aperture may be flush with the colony surface or surrounded by spines or an elevated rim (peristome).

Life cycle. Usually sheltered in various ways by the parent (female) zooid, each cheilostome zygote develops into one larva. The larva in most species swims freely for less than a day, feeding on its own yolk (that is, it is lecithotrophic), and then attaches to a substrate and metamorphoses (with complete degeneration of its larval organs) into a primary zooid (ancestrula). In a few species, the zygote is shed directly into the water and develops there into a larva (cyphonautes) that swims freely for up to several weeks before settling. Afterward, repeated asexual budding from the ancestrula produces the rest of the cheilostome colony.

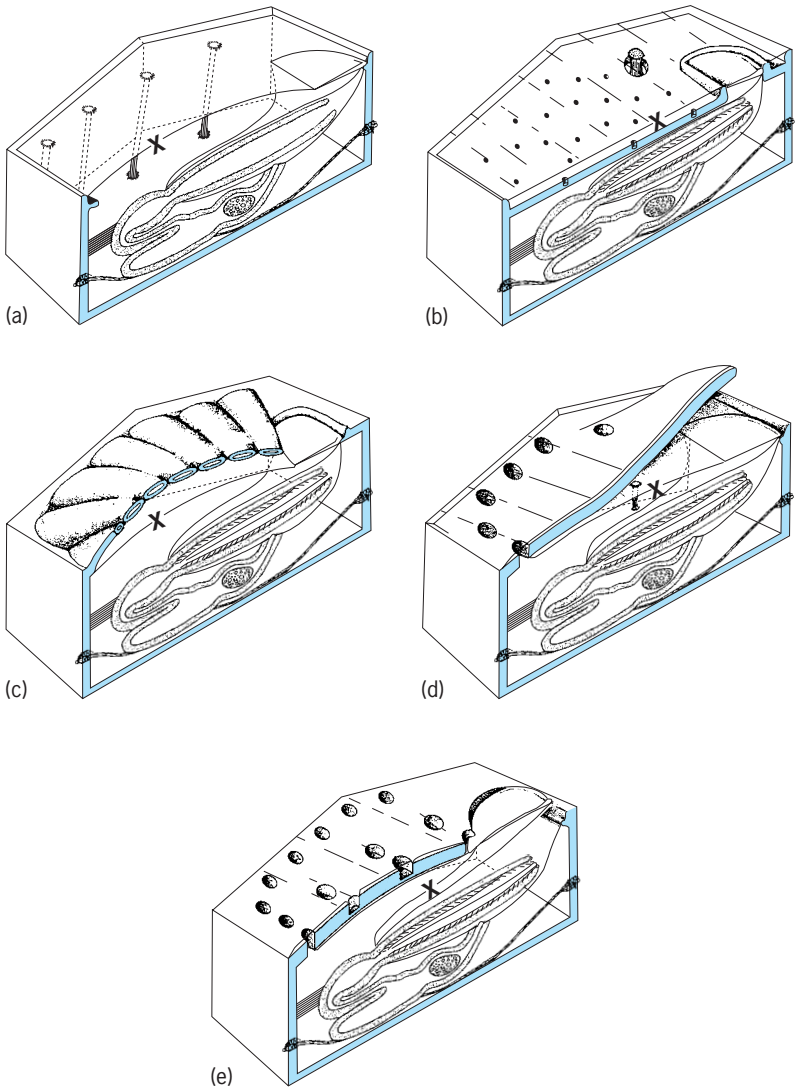

Fig. 3. Arrangements of frontal walls in Cheilostomata. The X indicates the membrane that can be pulled by muscles in order to displace fluid/liquid that pushes the polypide (shaded) out so that the tentacles can feed in the space above the skeleton. (*a*) Anascan with noncalcified frontal wall and retracted polypide unprotected by shield. (*b*) Anascan with perforated skeletal shelf (cryptocyst) below the outer, depressible cuticle. (*c*) Cribrimorph with shield of fused spines arched above the depressible cuticle. (*d*) "Umbonuloid" ascophoran with the frontal cuticle folded back on itself and the upper surface of the middle sheet of the frontal cuticle calcified, above the portion that is depressible. (*e*) Ascophoran as traditionally understood, with a small inflatable sac (ascus) formed by extension of the frontal cuticle lying immediately below the solid frontal wall/skeleton. (*After F. K. McKinney and J. B. C. Jackson, Bryozoan Evolution, Unwin Hyman, Boston, 1989*)

History and classification. Cheilostomes apparently evolved from ctenostomes as a single clade (monophyletic). First appearing in the latest Jurassic, the cheilostomes rapidly expanded during the Cretaceous to become the dominant bryozoan group throughout the Cenozoic, a position which they still maintain today. *See* CTENOSTOMATA.

Although cheilostomes are predominantly marine (ranging from intertidal to abyssal depths), some range into brackish waters, but none are found in freshwater. The oldest cheilostomes occupied nearshore habitats, but they soon invaded deep-water marine environments before the end of the Cretaceous. Maximum taxonomic, life-habit, and growth-form diversities are found near the shelf-slope break. In addition to being abundant in level-bottom communities, cheilostomes extensively but inconspicuously encrust the framework of many Cenozoic and modern coral reefs.

The cheilostomes have diversified into three major groups (anascans, cribrimorphs, and ascophorans) that appeared consecutively in geologic time and can be distinguished on the basis of differences in the frontal or upper wall (Fig. 1*f–i*). They are often classified as suborders, but may be considered as morphological grades of increasing complexity with time.

Anascan cheilostomes, the earliest, have a flexible, membranous, chitinous frontal wall that is completely open to the surrounding water (Fig. 1*f–g*; Fig. 2). Upon fossilization, the cuticle decays, leaving a top opening (opesium) that is much larger than the aperture would have been. Anascans protrude their tentacles via muscles that pull down on the membranous frontal wall, thereby squeezing the liquid body contents out through the aperture or orifice (**Fig. 3***a*). In some anascans, a porous calcareous plate (cryptocyst) may be secreted below, but detached from, the frontal membrane or cuticle (Fig. 3*b*). The primary radiation of anascans occurred in the Late Cretaceous, but they have maintained their high diversity to the present.

Cribrimorph cheilostomes (Fig. 1*b*) possess similar flexible frontal walls but are partially protected externally by overarching, partly fused, calcareous spines that meet above the midline of the zooid or zoeecium; they extend their tentacles in basically the same manner as anascans (Fig. 3*c*). Water passes in and out through the spaces between the spines as the underlying cuticle is flexed to displace polypides (internal contents of the zooid) out or to accommodate displaced fluid as polypides are retracted. The cribrimorphs had evolved by the early Late Cretaceous, reached a moderate diversity during the Late Cretaceous, and then persisted at a reduced or low diversity until the present.

Ascophoran cheilostomes (Fig. 1*i*), the most highly evolved, appeared by the mid-Late

Cretaceous, and display solid, shieldlike, calcareous frontal walls, below which is a special flexible sac (an ascus or compensatrix) used hydrostatically to protrude and retract the lophophore. The ascophoran condition, or grade, evolved several times and involved various ways of maintaining an inflatable structure below the rigid calcified frontal wall. In some (the umbonuloids), the frontal cuticle is deeply folded just behind the orifice (Fig. 3*d*). The reflected portion of the cuticle is calcified on its upper surface, while other portions remain uncalcified. Muscles pull the inner portion of the frontal membrane inward to protrude the polypide. Other ascophorans have developed a calcified plate under the frontal cuticle, and a balloonlike cuticular sac extends down through a pore in the calcified plate just behind or along the proximal edge of the operculum (Fig. 3*e*). By the end of the Cretaceous, ascophorans were diversifying rapidly, and have continued to diversify throughout the Cenozoic.

Many morphological and ecological innovations in cheilostomes originated in the mid-Cretaceous during a period of rapid diversification. Articulated spines, frontal budding, brood chambers, erect rigid and articulated colonies, free-living colonies, avicularia, and cribrimorph and ascophoran conditions originated between the late Early Cretaceous and mid-Late Cretaceous. Another (later) Early Cenozoic phase of rapid diversification primarily involved ascophorans and was accompanied by increased prevalence of frontal budding and more colonially integrated types of zooidal budding along the colony margins. Roger J. Cuffey; Frank K. McKinney

Bibliography. R. J. Cuffey and J. E. Utgaard, Bryozoans, pp. 204–216, in R. Singer (ed.), *Encyclopedia of Paleontology*, vol. 1, Fitroy Dearborn, Chicago, 1999; P. J. Hayward and J. S. Ryland, *British Ascophoran Bryozoans*, Linnean Society and Academic Press, London, 1979; F. K. McKinney and J. B. C. Jackson, *Bryozoan Evolution*, Unwin Hyman, Boston, 1989; J. S. Ryland, *Bryozoans*, Hutchinson University Library, London, 1970; J. S. Ryland and P. J. Hayward, *British Anascan Bryozoans*, Linnean Society and Academic Press, London, 1977.

Chelation

A chemical reaction or process involving chelate ring formation and characterized by multiple coordinate bonding between two or more of the electron-pair-donor groups of a multidentate ligand and an electron-pair-acceptor metal ion. The multidentate ligand is usually called a chelating agent, and the product is known as a metal chelate compound or metal chelate complex. Metal chelate chemistry is a subdivision of coordination chemistry and is characterized by the special properties resulting from the utilization of ligands possessing bridged donor groups, two or more of which coordinate simultaneously to a metal ion. *See* COORDINATION CHEMISTRY.

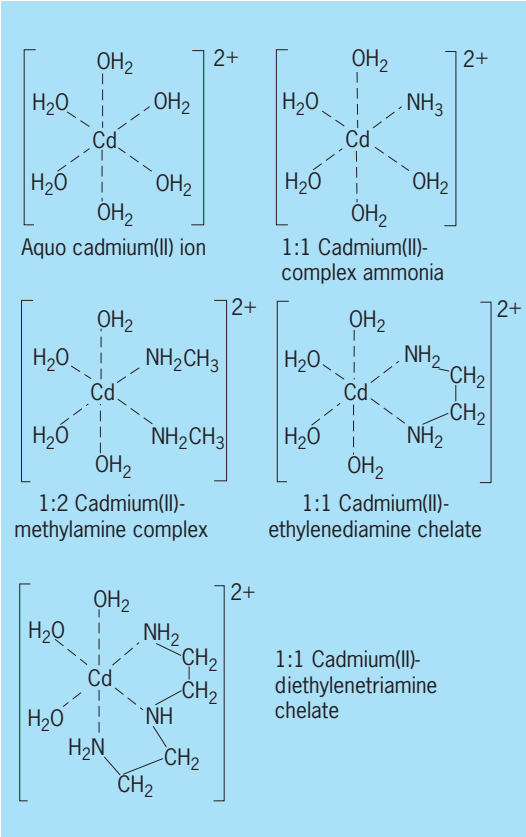

Fig. 1. Formulas of hydrated cadmium(II) ion, two cadmium(II) complexes, and two cadmium(II) chelates.

Ethylenediamine, $H_2NCH_2CH_2NH_2$, is a good example of a bidentate chelating agent consisting of two amino donor groups joined to each other by a two-carbon bridge. The coordination of both nitrogen atoms to the same metal ion would result in the formation of a five-membered chelate ring. If a chelating agent has three groups capable of attaching to a metal ion, as in diethylenetriamine, $NH_2CH_2CH_2NHCH_2CH_2NH_2$, it is tridentate (terdentate); if four, tetradentate (quadridentate); five, pentadentate (quinquidentate); six, hexadentate (sexadentate); and so on. In general, these chelating agents may be designated as multidentate or polydentate ligands. All chelating agents must be at least bidentate; tridentate agents give fused rings. An example of a nonchelated complex and two chelated complexes of the Cd(II) ion are illustrated in **Fig. 1**. In aqueous solution, metal ions are completely solvated, or hydrated, to give an aquo complex, as indicated for the Cd(II) ion in Fig. 1. Thus formation of a complex of a unidentate ligand involves the replacement of a water molecule by the unidentate donor group. Two water molecules are replaced by a single molecule of a bidentate ligand; three water molecules are replaced by a molecule of a terdentate ligand, and so on.

Many of the functional groups of both synthetic and naturally occurring organic compounds can form coordinate bonds to metal ions, producing metal-organic complexes or chelates, many of which

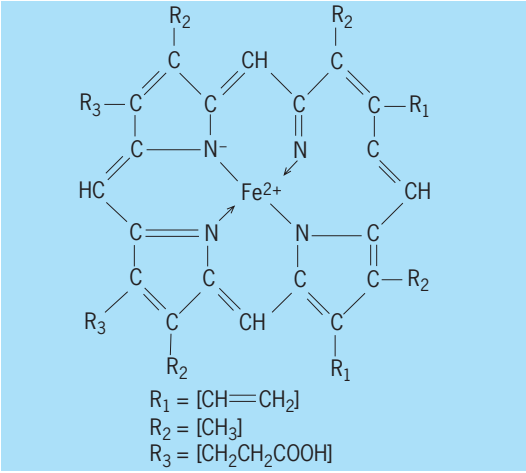

Fig. 2. Heme, a tetradentate chelate of iron(II).

are biologically active. Thus chelate compounds are frequently found in an interdisciplinary field of science called bioinorganic chemistry. The biological significance of chelates can be readily recognized if one notes that a large number of biologically important compounds are either metal chelates or chelating agents. This list includes the alpha amino acids, peptides, proteins, enzymes, porphyrins (such as hemoglobin), corrins (such as vitamin B_{12}), catechols, hydroxypolycarboxylic acids (such as citric acid), ascorbic acid (vitamin C), polyphosphates, nucleosides and other genetic compounds, pyridoxal phosphate (vitamin B_6), and sugars. The ubiquitous green plant pigment, chlorophyll, is a magnesium chelate of a tetradentate ligand formed from a modified porphin compound, and similarly the oxygen transport heme of red blood cells contains an Fe(II) chelate of the type illustrated in **Fig. 2**. *See* BIOINORGANIC CHEMISTRY; ORGANOMETALLIC COMPOUND.

A rapidly growing body of experimental evidence indicates that chelation may be important in the pharmacological action of many drugs. The use of chelating agents to remove certain toxic cations such as lead and plutonium from the body is now widely recognized in medical practice. Intensive research has been undertaken to either develop or discover an effective synthetic or natural chelating agent for the removal of iron deposits in the body that result from certain hereditary metabolic disorders.

Applications. The ability of chelating agents to reduce the chemical activity of metal ions has found extensive application in many areas of science and industry. Ethylenediaminetetraacetic acid (EDTA), a hexadentate chelating agent (**Fig. 3**), has been employed commercially for water softening, boiler scale removal, industrial cleaning, soil metal micronutrient transport, and food preservation. Nitrilotriacetic acid (NTA) is a tetradentate chelating agent (**Fig. 4**) which, because of lower cost, has taken over some of the commercial applications of EDTA. Chelating agents of related type have been used in biological systems to produce metal-ion buffers. By selection of the appropriate chelating agent, the free metal-ion concentration can be maintained at a very low and constant concentration level, just as a relatively constant pH can be maintained through the use of a conventional hydrogen-ion buffer system. Furthermore, the colors of certain chelating agents are sensitive to metal-ion concentration in a manner completely analogous to the pH-dependent color changes observed with acid-base indicators; such chelating agents serve as metal-ion indicators in analytical chemistry. The solubility of many chelating agents and metal chelates in organic solvents permits their use in solvent extraction of aqueous solutions for the separation or analysis of metal ions.

Many commercially important dyes and pigments, such as copper phthalocyanines, are chelate compounds. Humic and fulvic acids are plant degradation products in lake and seawater sediments that have been suggested as important chelating agents which regulate metal-ion balance in natural waters. By virtue of its abundance, low toxicity, low cost, and good chelating tendencies for metal ions that produce water hardness, the tripolyphosphate ion (as its sodium salt) is used in large quantities as a builder in synthetic detergents. Both synthetic ion exchangers and the mineral zeolites are chelating ion-exchange resins which are used in analytical and water-softening applications. As final examples, less conventional chelating agents are the multidentate, cyclic ligands, termed collectively crown ethers, which are particularly suited for the complexation of the alkali and alkaline-earth metals.

Stabilities of metal chelates in solution. One of the most striking properties of chelate ring compounds is their unusual thermodynamic and thermal stability. In this respect, they resemble the aromatic rings of organic chemistry. For example, in reaction (1),

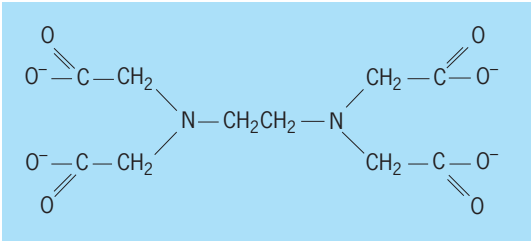

Fig. 3. Structural formula of anion of EDTA, a sexadentate ligand.

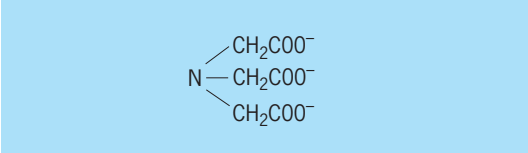

Fig. 4. Structural formula of NTA anion, a quadridentate ligand.

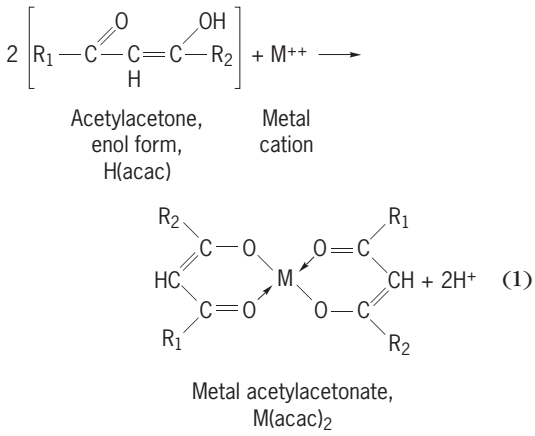

(1)

β-diketones in the enol form can lose a hydrogen ion and coordinate with a metal cation to give a six-membered ring of unusual thermal stability.

Beryllium acetylacetonate boils without decomposition at 270°C (518°F), and G. T. Morgan reported that scandium acetylacetonate, $Sc(acac)_3$, shows very little decomposition at 370°C (698°F). This remarkable stability contrasts sharply with the very low stability of coordination compounds containing simpler monodentate ketones, such as acetone.

Because of enhanced thermodynamic stability in solution, chelating agents may greatly alter the behavior of metal ions. The very insoluble compound, ferric hydroxide, will dissolve in a strongly alkaline solution of triethanol amine, $N(CH_2CH_2OH)_3$. Alternatively, the concentration of the ferric ion can be made vanishingly small at pH 2 by the addition of an equimolar amount of bis(orthohydroxybenzyl)-ethylenediamine-*N,N′*-diacetic acid (HBED; **Fig. 5**).

Since about 1940, both theoretical and practical considerations have focused attention on the factors which contribute to chelate stability. It is convenient to list such factors under three headings: (1) the nature of the metal cation, (2) the nature of the ligand, and (3) the formation of the chelate ring. It should be emphasized that these factors operate together, and their separation is somewhat artificial, but helpful for discussion purposes.

Role of metal in chelate stability. Since nitrogen, oxygen, and sulfur serve as the electron donor atoms in a majority of chelating agents, it is of interest to seek a relationship between the donor atom and the type of metal acceptor atom with which it combines.

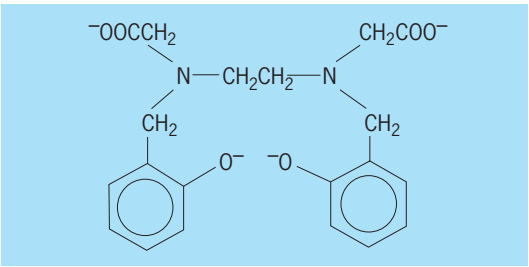

Fig. 5. Structural formula of anion of HBED, a sexadentate ligand.

A large majority of the chelates of Li^+, Na^+, K^+, Rb^+, Cs^+, Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}, Al^{3+}, Ga^{3+}, In^{3+}, Tl^{3+}, Ti^{4+}, Zr^{4+}, Th^{4+}, Si^{4+}, Ge^{4+}, and Sn^{4+} contain oxygen as at least one of the donor atoms. It may be furnished as an acid, alcohol, ether, ketone, or other group. These ions coordinate less frequently through two nitrogen or sulfur atoms. Cations of other metals such as vanadium, niobium, tantalum, molybdenum, and uranium, and the cations Be^{2+}, Al^{3+}, and Fe^{3+} show a preference for oxygen as the donor atom, but they may coordinate through nitrogen, sulfur, or phosphorus under special conditions. Cr^{3+}, Fe^{2+}, and the platinum metals show increasing preference for coordination through nitrogen as opposed to oxygen, while Cu^+, Zn^{2+}, Ag^+, Au^+, Cu^{2+}, Cd^{2+}, Hg^{2+}, V^{3+}, Co^{3+}, and Ni^{2+} show a marked preference for nitrogen and sulfur as the donor atoms. The ions of the last group retain the ability to coordinate with oxygen in even greater degree than do the ions of the first group, but their tendency to form bonds through nitrogen is so great that it exceeds their oxygen-binding tendency.

It must be recognized that broad generalizations such as these have many exceptions, particularly in intermediate regions. On the other hand, such generalizations indicate clearly that attempts to arrange elements in the order of their chelating ability can be of significance only when cations of comparable type are selected. Thus, the stabilities of the alkali-metal, the alkaline-earth, and the rare-earth chelates decrease as the charge on the cation decreases or as the size of the cation increases. For example, the chelates of the alkaline-earth metal ions become less stable as the metal ion becomes larger (if the number of chelate rings remains the same), in the order Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}, Ra^{2+}. The relationship between ion size and chelate stability is of major importance in the separation of the rare-earth and transuranium elements by ion-exchange processes. The selectivity of the ion-exchange column is increased by the use of appropriate chelating agents in the eluting solution.

Stability sequences established for other metal ions are somewhat less satisfactory. Metal chelates of several substituted β-diketones decrease in stability in the order: $Hg^{2+} > (Cu^{2+}, Be^{2+}) > Ni^{2+} > Co^{2+} > Zn^{2+} > Pb^{2+} > Mn^{2+} > Cd^{2+} > Mg^{2+} > Ca^{2+} > Sr^{2+} > Ba^{2+}$. If one restricts stability comparisons to bivalent metals of the first transition series, the following order is obtained: $Zn^{2+} < Cu^{2+} > Ni^{2+} > Co^{2+} > Fe^{2+} > Mn^{2+}$. This latter listing of stabilities appears to be valid for a large variety of chelating ligands.

Lower valence states such as Ir(I) and Rh(I) are effectively coordinated by chelating ligands containing trivalent phosphorus or trivalent arsenic donor groups. For example, recently synthetic coordination chemists have synthesized a series of unusual chelating agents containing trivalent phosphorus, analogous to polyamines, an example of which is illustrated in **Fig. 6**.

General principles of selective coordination of metal ions by various types of donor atoms have now been worked out, and are described through concepts such as the principles of hard and soft acids

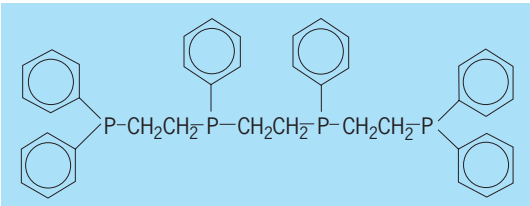

Fig. 6. "Tetraphos," a quadridentate ligand.

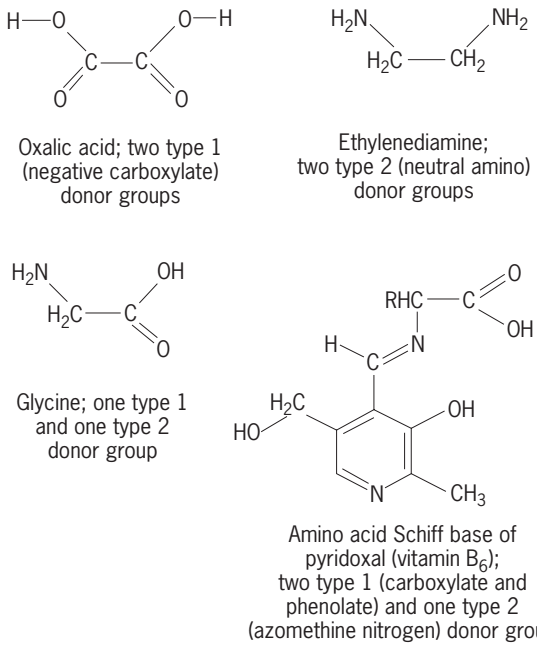

Fig. 7. Chelating ligands indicating the types of donor groups available for coordination.

and bases, and type A and B character of metal ions. These qualitative principles are based on well-known laws of ionic attraction and polarizabilities of atoms, ions, or molecular groups, as well as on general principles of molecular orbital theory.

Role of ligand in chelate stability. Two general types of groups give rise to coordinate bonds between ligands and metal. These are (1) primary acid groups in which the metal ion can replace hydrogen ions and (2) neutral groups which contain an atom with a free electron pair suitable for bond formation. If two groups from either class 1 or 2, or from both classes, are present in the same molecule in such positions that both groups can form bonds with the same metal ion, a chelate ring may be formed. For example, as shown in **Fig. 7**, in oxalic acid two groups of type 1 are present; in ethylenediamine two groups of type 2 are present; glycine possesses one of each type; and pyridoxylidine-glycine (the Schiff base of glycine and vitamin B_6) possesses a carboxylate (type 1), an aromatic phenolate (type 1), and an imino group (type 2) available for chelate formation around a metal ion.

In general, anything which increases the localization of negative charge on the donor atom increases its ability to coordinate to a metal atom. Since a hydrogen ion is bound to a ligand by an electron pair in groups of type 1, an increase in electron density on the donor atom will increase the ability of the donor to bind either hydrogen ion or metal cation. The ability of a ligand to bind hydrogen ion is frequently referred to as its basic strength. It is surprising then that, for ligands of rather similar type, an increase in ligand basic strength implies an increase in its metal-chelating ability. A number of researchers have indicated this relationship by plotting the log of the equilibrium constant for the process represented by Eq. (2) against the log of the equilibrium con-

$$H^+ + L^- \rightleftharpoons HL \qquad K_1 = \frac{[HL]}{[H^+][L^-]} \qquad (2)$$

stant for the process of chelate dissociation shown by Eq. (3).

$$M^{n+} + L^- \rightleftharpoons M^{(n-1)+} \qquad K_2 = \frac{[ML^{(n-1)+}]}{[M^{n+}][L^-]} \qquad (3)$$

L^- is the chelating anion, such as the acetylacetonate anion,

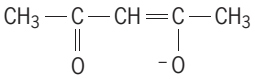

The constant K_1 is the protonation constant (its reciprocal is the acid dissociation constant), and K_2 is designated the stability constant of the metal chelate.

For the organic chemist, the analogy between the hydrogen cation and the metal cation is even more clearly drawn. Rings formed by hydrogen bonding are referred to as chelates; thus, formic acid dimerizes through hydrogen-bond chelation as shown in **Fig. 8**. The high volatility of *o*-nitrophenol compared with the much lower volatility of its meta or para isomers (Fig. 8) can only be explained in terms of

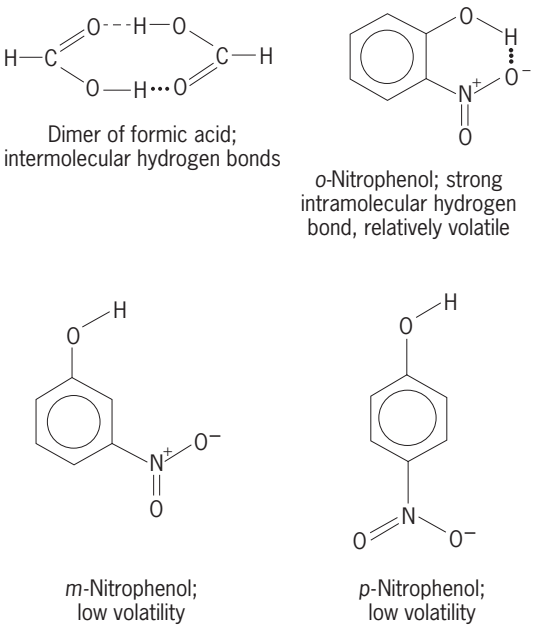

Fig. 8. Compounds forming intramolecular and intermolecular hydrogen bonds.

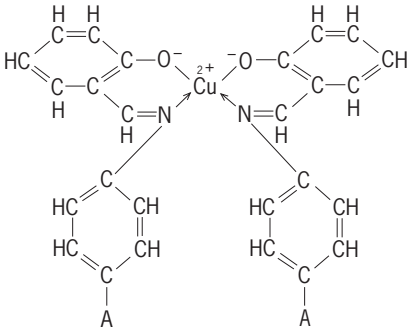

Fig. 9. Structure of a chelated copper compound.

intramolecular versus intermolecular hydrogen bonding. The properties of salicylaldehyde, enhanced enolization in acetoacetic ester, and many other organic compounds are altered by internal chelate-ring formation involving hydrogen bonds. *See* HYDROGEN BOND.

The role of ligand structure on chelate stability is illustrated by studies which have been conducted on compounds of the type shown by **Fig. 9**, where A represents an electron-attracting group. In general, it was found that with an increase in the electron-attracting power of group A, electrons were pulled away from the nitrogen atom, resulting in both lower base strength and lower chelating ability for the ligand. As A is changed successively through the groups shown below, chelate stability increases in the order:

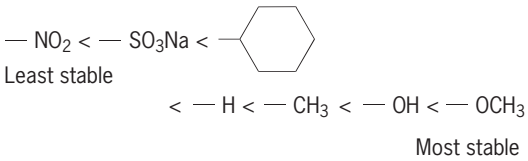

Similar studies on substitute β-diketone chelates of the type

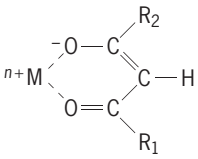

showed that, if R_1 were changed from a methyl group to an electron-withdrawing trifluoromethyl group, the stability of the resulting chelate is greatly decreased. In general, all molecular charge effects which can be invoked to shift charge in an organic molecule, such as inductive and resonance effects, will influence chelate stability. Thus the principles governing charge distribution and bond hybridization in organic chemistry are useful in working out the relationships between ligand structure and chelate stability.

Role of ring closure in chelate stability. The stability factors discussed above are applicable to coordination compounds generally, not to chelates alone. On the other hand, a number of stability factors may be considered to apply uniquely to chelates because of their ring structure. The most obvious variable in this category is ring size, a factor which is uniquely determined by the position of the donor atoms in the chelating ligand. When the groups are present in such a position as to form a five or six-membered ring, the resulting complex is the most stable although four-, seven-, eight-membered, and even larger rings are known. Examples of these would be found among biological ligands as well as in ion exchange resins. The existence of three-membered rings has not been established. Hydrazine, H_2NNH_2, which might in theory form a three-membered chelate ring, appears to be monodentate. There is some evidence for intermediate structures of metal-oxygen complexes

which may be considered three-membered chelate rings. Such compounds may be important reaction intermediates but are generally unstable and present in relatively low concentrations. The four-membered chelate rings are frequently strained. Examples of four-membered chelate rings are copper(II)-carboxylate complexes, and the aluminum chloride dimer (**Fig. 10**). While five-membered rings are very common and are formed preferentially by saturated organic ligands, ligands containing two double bonds tend to form six-membered structures. If only one double bond is present, five- or six-membered rings may form; five-membered saturated rings are illustrated in Fig. 10 by ethylene-diamine-metal chelates, and the conjugated six-membered rings by the metal acetylacetonates.

Rings of seven or more members are comparatively uncommon, but their existence is well established. As the length of the chain between the two donor atoms increases, so does the tendency of the ligand to form polymetallic complexes. Under such circumstances, the two donor atoms on the same chelate molecule coordinate with different

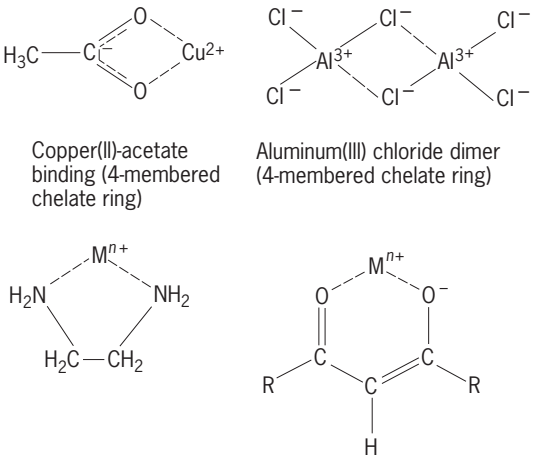

Fig. 10. Metal chelates containing four-, five-, and six-membered rings.

metal atoms rather than with one; thus, a polymeric chain, $M^+—NH_2CH_2CH_2CH_2CH_2NH_2—M^+$, may result instead of a ring structure.

The fact that chelate complexes are usually more stable than comparable nonchelate structures has been called the chelate effect. This effect is partly attributable to the fact that simultaneous rupture of both bonds holding the ligand to the metal is highly improbable, and if only one bond breaks there is a high probability that the broken bond will reform before the second bond is ruptured. As the chain length between the two donor atoms increases, the chance for reformation of the broken bond declines; thus, large rings usually show a decrease in stability. On the other hand, in fused ring systems, formed by polydentate ligands such as ethylenediaminetetraacetic acid, the probability that at least one bond will reform before all bonds are ruptured results in increased complex stability. When the bonding atoms are rigidly positioned around the metal by the organic framework, as in the porphins, the resulting increase in stability is extremely high. It has been reported that the copper phthalocyanine complex, with a completely interlocked ring system, is stable in the vapor phase near 500°C (930°F).

Chelate compounds are differentiated from their nonchelate analogs by several properties besides high stability. Although not all chelates are volatile, the existence of low-boiling metal acetylacetonates and related structures is noteworthy. If the coordination number of the metal cation for the oxygen atoms of the acetylacetone (that is, the number of nearest oxygens around the cation) is equal to twice the ionic charge of the cation, the resulting acetylacetonate is volatile; thus, beryllium with a charge of 2 and a coordination number of 4 forms an acetylacetonate which boils at 270°C (518°F); aluminum with a charge of 3 and a coordination number of 6 forms an acetylacetonate which boils at 314°C (597°F). If the coordination number of the central cation is less than twice the ionic charge, less volatile saltlike complexes are formed.

Isomerism of all types, so important in organic chemistry, is of major concern in chelation, particularly to the biochemist, since desired biological properties are frequently restricted to a particular chelate isomer. Ring formation may result in optical activity where analogous nonchelate structures are inactive. Thus, the ethylenediamine chelate structures shown in **Fig. 11** are optical isomers, whereas the analogous methylamine complexes represented in **Fig. 12** are optically inactive. *See* CHEMICAL BONDING; CHEMICAL EQUILIBRIUM; COORDINATION COMPLEXES; STEREOCHEMISTRY; STERIC EFFECT (CHEMISTRY).

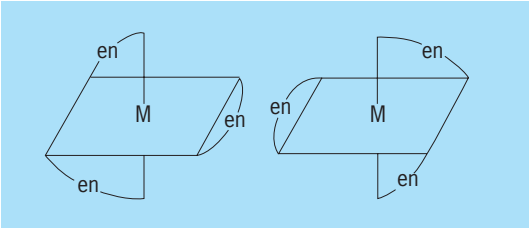

Fig. 11. Optical isomers of chelate complexes.

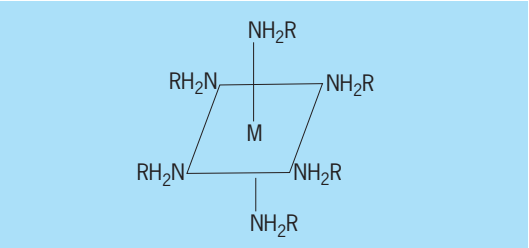

Fig. 12. Formula for an optically inactive nonchelate.

A. E. Martell; R. J. Motekaitis

Bibliography. N. W. Alcock, *Bonding and Structure: Structural Principles in Inorganic and Organic Chemistry*, 1991; I. Bertini, H. Gray, and J. Valentine, *Bioinorganic Chemistry*, 1994; R. Haydock (ed.), *Bonding and Structure of Solids*, 1991; A. E. Martell, The chelate effect, *Advances in Chemistry Series*, no. 62, American Chemical Society, pp. 272–294, 1967; A. E. Martell, *Inorganic Chemistry in Biology and Medicine*, 1980.

Chelicerata

A subphylum of the phylum Arthropoda. The Chelicerata can be defined as those arthropods with the anteriormost appendages as a pair of small pincers (chelicerae) followed usually by pedipalps and four pairs of walking legs, and with the body divided into two parts: the prosoma (corresponding approximately to the cephalothorax of many crustaceans) and the opisthosoma (or abdomen). There are never antennae or mandibles (lateral jaws). The Chelicerata comprise three classes: the enormous group Arachnida (spiders, ticks, mites, scorpions, and related forms); the Pycnogonida (sea spiders or nobodycrabs); and the Merostomata (including the extant Xiphosurida or horseshoe crabs). *See* ARACHNIDA; MEROSTOMATA; PYCNOGONIDA; XIPHOSURIDA.

With a few exceptions, the distinctive body and limb pattern of chelicerates is diagnostic. The Acarina (mites) have the prosoma and opisthosoma fused in a single armored globe. The segments of the prosoma are always fused (except in the minor order Solfugae or sun spiders), and the opisthosoma never bears locomotor appendages (although in the xiphosurids there are paired gills on six segments of the opisthosoma). Chelicerae are always present, but the form of the pedipalps varies greatly from tiny sensory appendages to enormous crablike chelae (in scorpions and certain minor orders). Four pairs of locomotor appendages follow the chelicerae and pedipalp on the prosoma; only in a few species of pycnogonids are there five or six pairs of walking legs. *See* ACARI.

Both Merostomata and Pycnogonida are marine, but the enormous numbers and varied forms of the

Arachnida are almost entirely terrestrial. The respiratory structures of chelicerates include gills, booklungs, and tracheae. Sexes are normally separate, with genital openings at the anterior end of the opisthosoma. Some mites and other small chelicerates are omnivorous scavengers, but the majority of species of larger chelicerates are predaceous carnivores at relatively high trophic levels in their particular ecotopes. *See* ARTHROPODA. W. D. Russell-Hunter

Chelonia

An order of the Reptilia, subclass Anapsida, including the turtles, terrapins, and tortoises. This order is also known as the Testudines. The group first appeared in the Triassic, and its representatives are among the commonest fossils from that time on. Members of the order are most frequently found in freshwater streams, lakes, and ponds or in marshy areas. However, a number of strictly terrestrial species are known, and several are marine. Turtles occur on all the major continents and continental islands in tropic and temperature regions. The marine forms are basically tropic in distribution, but some individuals stray into temperate waters. *See* ANAPSIDA.

The largest carapace length attained by living forms is around 8 ft (2.4 m) in the marine leatherback turtle (*Dermochelys*), which weighs 1500 lb (675 kg). Members of the large land tortoise group (*Geochelone*) reach lengths of 4 ft (1.2 m) and weights of about 560 lb (250 kg). The smallest living species is probably the stinkpot (*Kinosternon odoratus*) of eastern North America, which attains a maximum length of about $4^1/_2$ in. (11 cm). The longevity of turtles is well known, with many captive tortoises having lived 50–60 years and others estimated to be as old as 150 years.

Taxonomy. The living turtles are usually divided into two major groups, the suborders Pleurodira and Cryptodira, based upon the structures of the head and neck. Both lines are thought to be evolved from primitive forms placed in the suborder Amphichelydia (Triassic to Eocene). This latter group differs most markedly from recent species in having elongate neural spines on the neck vertebrae so that the neck is not retractible into the shell (**Fig. 1**). Correlated with the lack of retractibility, the skull is not emarginated to make room for attachment of muscle from the cervical region. The pleurodires, on the other hand, have spines on only the most posterior cervicals and the head is retractile laterally. The skull is emarginate. In the cryptodires the cervical spines are uniformly reduced, the skull is usually emarginate, and the head is folded directly back to within the shell. In several cryptodires, notably in the marine turtles, the neck is secondarily nonretractible because of a reduction in the shell, and the skull is nonmarginate. The **table** shows the principal groups and distribution of the approximately 225 living species of the order.

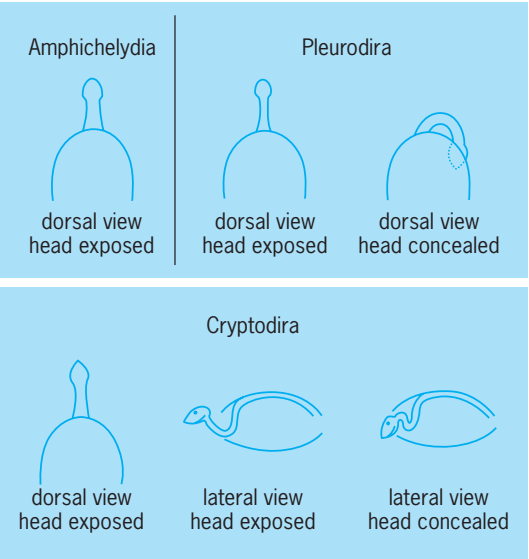

Fig. 1. Characteristics of turtle suborders, showing head in exposed and concealed positions.

Morphology. The Chelonia differ from most other vertebrates in possessing a hard bony shell which encompasses and protects the body (**Fig. 2**). The shell is made up of a dorsal portion, the carapace, and a ventral segment, the plastron, connected by soft ligamentous tissue or a bony bridge. The carapace is composed of the greatly expanded ribs and dorsal vertebrae overlain by a series of enlarged dermal ossifications and an outer covering of tough skin or horny scales. The plastron is similarly arranged

Families of the Chelonia

Group	Common name	Number of species	Distribution
Suborder Pleurodira			
Family Pelomedusidae	Side-necked turtles	14	Southern Africa, Madagascar, South America
Family Chelidae	Side-necked turtles	32	Australia, New Guinea, South America
Suborder Cryptodira			
Family Kinosternidae	Mud and musk turtles	22	North, Middle, and South America
Family Chelydridae	Snapping turtles	3	Middle and northern South America
Family Emydidae	Pond turtles and allies	102	All continents except Australia; only one or two South American species
Family Testudinidae	Tortoises	30	All continents except Australia; two South American species
Family Cheloniidae	Hawksbill, loggerhead, and green sea turtles	6	Tropical and subtropical seas
Family Dermocheylidae	Leatherback turtle	1	Tropical and subtropical seas
Family Trionychidae	Soft-shelled turtles	14	North America, Asia, central and southern Africa
Family Carretochelyidae	Fly River turtle	1	New Guinea

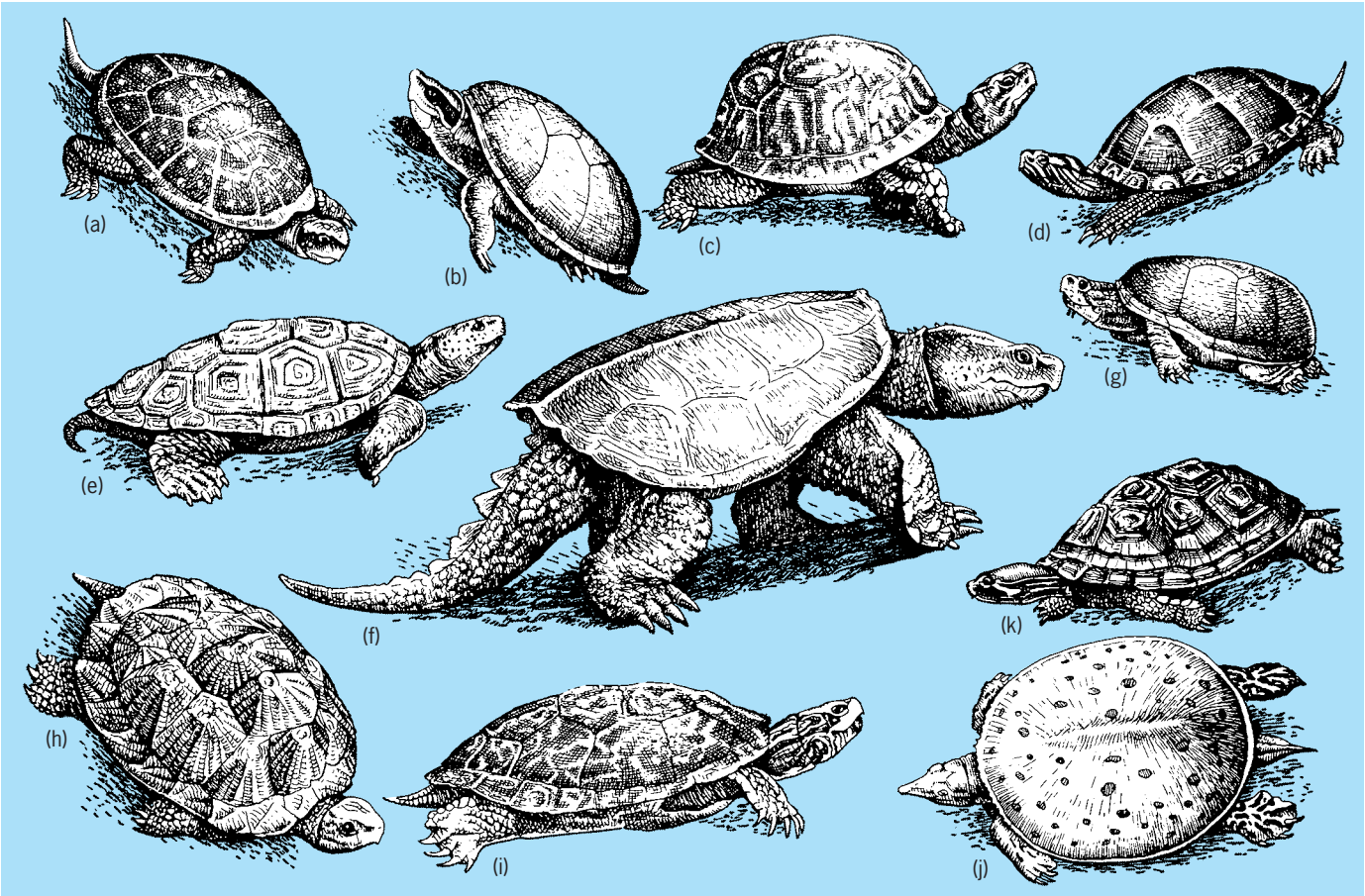

Fig. 2. Common United States turtles. (*a*) Spotted turtle, *Clemmys guttata*. (*b*) Stinkpot, *Kinosternon odoratus*, (*c*) Box turtle, *Terrapene carolina*. (*d*) Painted turtle, *Chrysemys picta*. (*e*) Diamondback terrapin, *Malaclemys terrapin*. (*f*) Common snapping turtle, *Chelydra serpentina*. (*g*) Mud turtle, *Kinosternon subrubrum*. (*h*) Wood turtle, *Clemmys insculpta*. (*i*) Map turtle, *Graptemys geographica*. (*j*) Spiny soft-shelled turtle, *Trionyx spiniferus*. (*k*) River terrapin, *Chrysemys floridana*. (*After J. J. Shomon, ed., Virginia Wildlife, 15*(6)*:27, 1954*)

with remnants of the interclavicle, clavicles, and gastralia fused with dermal ossifications and covered by skin or scales. Other peculiarities associated with the shell include the fusion of the ribs to the vertebrae and the reduction of trunk muscles, the presence of the pectoral girdle completely inside the ribs (found in no other animal), the highly modified short, thick humerus and femur, and the lungs attached dorsally to the shell. In addition, living turtles differ from the tuatara, snakes, lizards, and crocodilians in having an anapsid skull, no true teeth but horny beaks on the jaws, an immovable quadrate, and a single median penis in males. Teeth are present in ancient primitive turtles.

The rigid bony shell of turtles imposes a basic body plan subject to relatively little variation. Turtles are always recognizable as such, and the principal obvious differences between them are in the shell shape, limbs, head, and neck. The general outline of the shell is variable, but in most forms the shell is moderately high-arched and covered with epidermal scales. In some of the tortoises, such as *Geochelone*, the shell is very high, equaling the breadth. In the soft-shelled turtles, Trionychidae, and the saxicolous African tortoise (*Malacochersus*), the shell is greatly depressed to a thin, almost platelike form. No scales are present on the shells of the families Dermochelyidae, Carrettochelyidae, and Trionychidae, which are covered by thick skin.

The majority of turtles have limbs more or less adapted to aquatic or semiaquatic life with moderate to well-developed palmate webbed feet. However, in strictly terrestrial forms such as the tortoises, the limbs are elephantine with the weight of the body being borne on the flattened sole of the feet. The marine turtles and the river turtle (*Carrettochelys*) of southern New Guinea have the anterior limbs modified into paddlelike flippers, used to propel them through the water rapidly.

Physiology. Associated with the tendency of most turtles toward a life in or near water is the auditory apparatus. Even though an eardrum is present, hearing is adapted to picking up sounds transmitted through the water or substratum. However, there is evidence that airborne sound can be heard. Vision is also important, and turtles have color vision. Sounds are produced by many species through expulsion of air through the glottis. Respiration is by means of the rather rigid lungs, which are inflated and deflated by a series of special muscles. Several species with aquatic habits have special vacularized areas in the mouth or cloaca which act as auxiliary respiratory

devices and make possible gaseous exchanges when the turtle is submerged in water. Most forms emit highly pungent oily substances from specialized cloacal glands, and these secretions probably are useful as recognition cues in the water. Most of the species are gregarious and diurnal, and territoriality is unknown in the order.

Reproduction. The courtship patterns of various turtles are distinctive. In general, aquatic forms mate in the water, but tortoises breed on land. The male mounts the female from above, and fertilization is internal. Sperm may be stored in the cloacal region of the female for extended periods before fertilization. Because of the presence of the large median penis in males, the sexes of many species can be distinguished by the longer and broader tail of the male. In addition, in many forms the plastron of the male is concave while in the female it is moderately to strongly convex. No doubt this characteristic of the female is associated with the spatial requirements for egg production and storage. All turtles lay shelled eggs which are buried in sand or soil in areas where females congregate. The shell is calcareous in most forms, but the marine turtles have leathery shells. Eggs are rather numerous, as many as 200 being laid by a single individual in some species, and are highly valued as human food. Incubation takes 60–90 days,

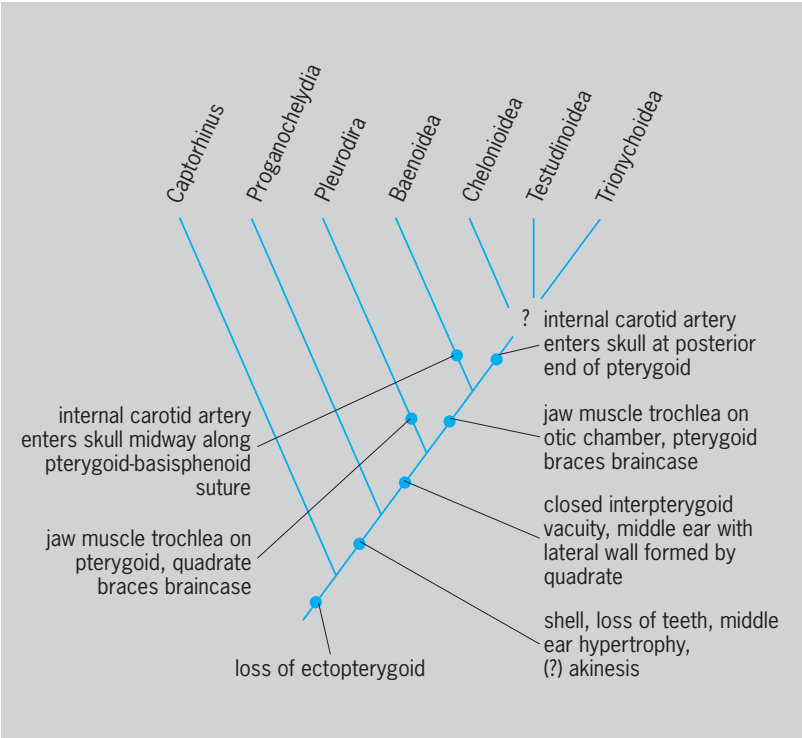

Fig. 3. Relationships of Chelonia derived from analysis of cranial features.

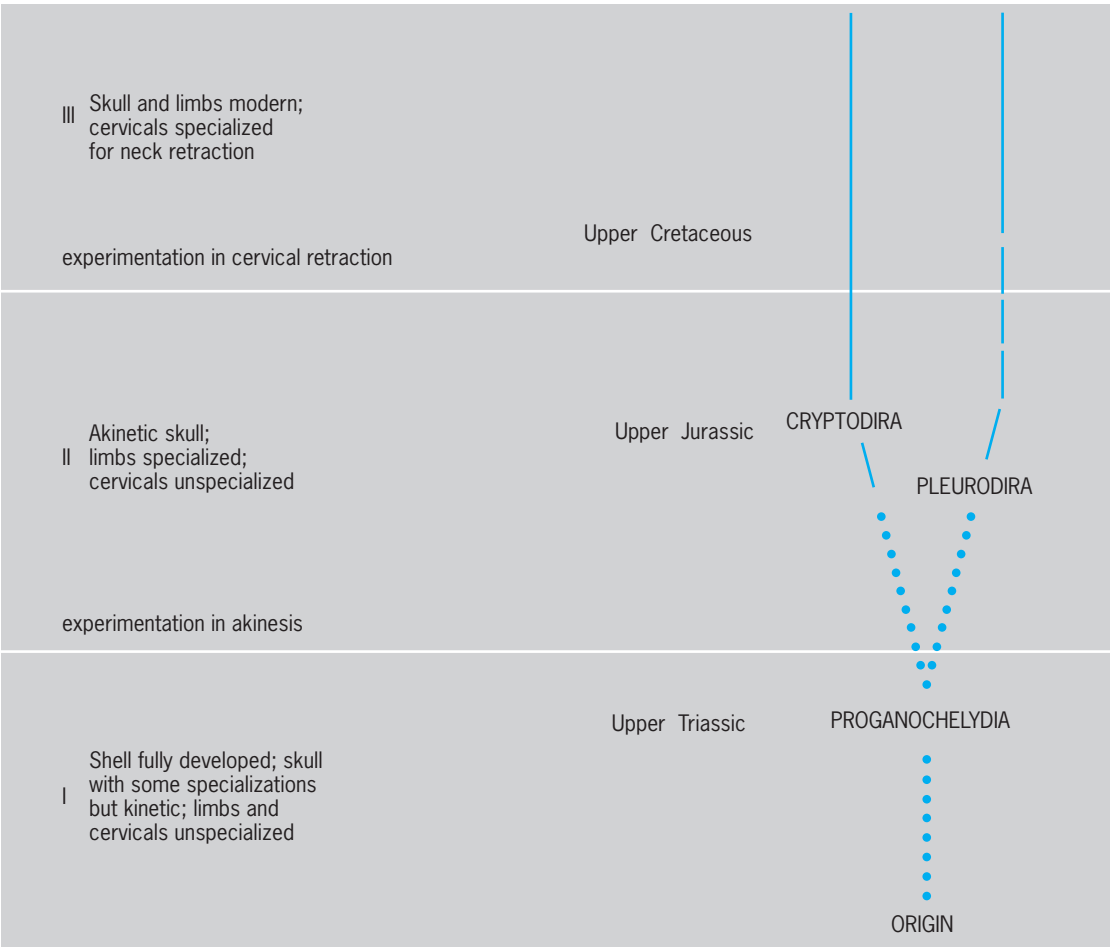

Fig. 4. Levels of organization in the order Chelonia.

and the little turtle cuts its way out of the shell with a small horny egg caruncle at the end of the snout.

Nutrition. Turtles feed on all types of organisms. Aquatic species may eat algae, higher plants, mollusks, crustaceans, insects, or fishes; terrestrial forms are similarly catholic in tastes. Most species are omnivores, but some have very specialized diets. The edges and internal surfaces of the horny beaks of these reptiles are frequently denticulate and modified to form specialized mechanisms adapted to handle particular food items. Jay M. Savage

Phylogeny. Turtles have long been considered a "primitive" or relict reptile group that has maintained a consistent morphologic structure since its origin. Phylogenetic and functional work on the skulls of fossil and modern turtles suggest that this approach is incorrect. The Cryptodira and the Pleurodira probably represent a very early dichotomy which took place soon after the origin of the turtles. These two groups passed through similar stages during their phylogeny and repeatedly evolved parallel adaptations.

The basicranium of turtles has rarely been used in a comparative sense, and the morphology of this region is still poorly known in many families of turtles. Nonetheless, the basicranium is one of the most useful regions for phylogenetic work because of its extreme complexity and relative conservativism compared with other areas. The main arterial and venous circulation of the head, the cranial nerves, the brain, and the neck and jaw musculature are all related to features in the basicranium. Changes in any of these systems can usually be seen in the basicranium. Although turtle skulls are rare compared to shells, the fossil record contains a good series with basicrania dating from the Upper Triassic to the Recent.

Previous phylogenies, however, have been based primarily on shell criteria and vertebral characters. Two major living groups are usually distinguished: the Cryptodira, with vertical neck retraction, and the Pleurodira, with horizontal neck retraction. A few of the shell characters also distinguish the living forms: pelvis sutured to carapace and plastron in pleurodires, pelvis free in cryptodires. These features rather clearly separate living turtles into two groups, but when fossils are examined it becomes difficult to distinguish pleurodires and cryptodires. The neck retractile mechanisms had not evolved until the end of the Mesozoic; thus the earlier phylogeny of the two groups is difficult to determine by neck characters.

Work on the skulls of fossil and recent turtles, however, has resulted in the development of a new hypothesis of chelonian relationships. The use of cranial features in a shared-derived character analysis is expressed in **Fig. 3**. Arterial characters provide the basis for relating the major groups of cryptodires, whereas jaw mechanics and middle ear morphology are used in comparing Proganochelydia, Pleurodira, and Cryptodira. By using these characters, pleurodires and cryptodires can be recognized well into the Mesozoic as distinct lineages, and the divergence of the two groups from a common ancestor probably took place before the Jurassic.

The new phylogenetic information has enabled functional interpretations to be made concerning the biologic significance of the differences between cryptodires and pleurodires. These interpretations show that the two lineages have developed similar structures in parallel during their evolution (**Fig. 4**). A similar but nonhomologous jaw mechanism involving a trochlea to compensate for the expanded otic chamber (**Fig. 5**) and a method of bracing the palatoquadrate against the braincase are complex features evolved independently in both groups. Later in their evolution the neck retractile mechanisms evolved, but by this time the two groups had been separate for many millions of years.

Origin. The origin of turtles is more in doubt than ever. Previous workers have emphasized the peculiar adaptations of a Permian reptile from South Africa, *Eunotosaurus*. This animal, represented by postcranial remains but very little skull material, was suggested as an ancestor for turtles on the basis of its expanded ribs, which form a rudimentary shell. However, T. S. Parsons and E. E. Williams in 1961 and, in particular, C. B. Cox in 1969 showed that *Eunotosaurus* is not related to turtles. Cox stated that *Eunotosaurus* is a captorhinomorph and that its Chelonia-like features are due to structural convergence with turtles. *Eunotosaurus* lacks separate ribs on its last dorsal vertebra, and Cox held that this feature prevents *Eunotosaurus* from being a chelonian ancestor. No reptile known to science

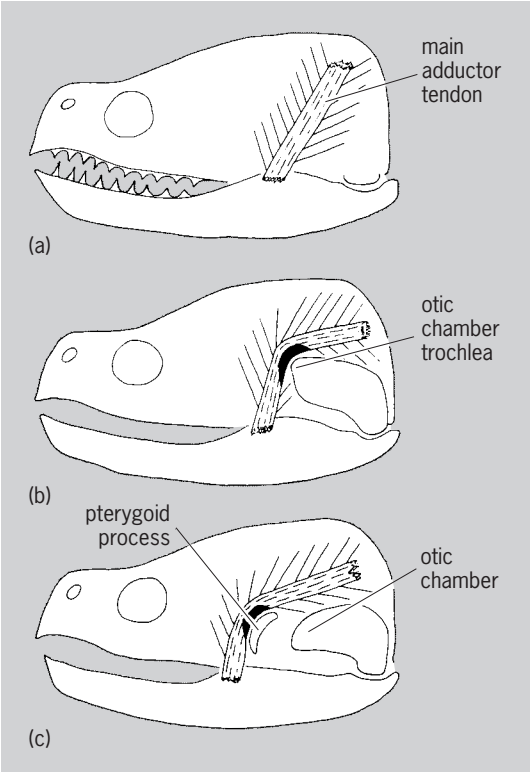

Fig. 5. Position of main jaw muscle tendon in (*a*) early reptiles, (*b*) cryptodires, and (*c*) pleurodires. The trochlea or pulley formed in turtles to get around the expanded otic chamber is independently derived in pleurodires and cryptodires.

definitely possesses specialized characters in common with turtles that would allow it to be a close relative or ancestor of that order. Presumably the stem reptile group Captorhinomorpha includes the eventual forebears of turtles since captorhinomorphs probably were ancestral to all reptiles, but the intermediate forms are unknown. *See* REPTILIA.

Eugene S. Gaffney

Bibliography. P. W. Albrecht, *Tulane Stud. Zool.*, 14:81–99, 1967; R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988; C. B. Cox, *Bull. Brit. Mus.* (*Nat. Hist.*), 18(5):167–196, 1969; E. S. Gaffney, *Bull. Amer. Mus. Nat. Hist.*, 155(5):387–436, 1975; M. Harless and H. Morlock, *Turtles: Perspectives and Research*, 1979, reprint 1988; S. B. McDowell, *Bull. Mus. Comp. Zool.*, 125(2):23–39 1961; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Chemical bonding

The force that holds atoms together in molecules and solids. Chemical bonds are very strong. To break one bond in each molecule in a mole of material typically requires an energy of many tens of kilocalories.

It is convenient to classify chemical bonding into several types, although all real cases are mixtures of these idealized cases. The theory of the various bond types has been well developed and tested by theoretical chemists. *See* COMPUTATIONAL CHEMISTRY; MOLECULAR ORBITAL THEORY; QUANTUM CHEMISTRY.

The simplest chemical bonds to describe are those resulting from direct coulombic attractions between ions of opposite charge, as in most crystalline salts. These are termed ionic bonds. *See* COULOMB'S LAW; IONIC CRYSTALS; STRUCTURAL CHEMISTRY.

Other chemical bonds include a wide variety of types, ranging from the very weak van der Waals attractions, which bind neon atoms together in solid neon, to metallic bonds or metallike bonds, in which very many electrons are spread over a lattice of positively charged atom cores and give rise to a stable configuration for those cores. *See* CRYSTAL STRUCTURE; INTERMOLECULAR FORCES.

Covalent bond. But it is the normal covalent bond, in which two electrons bind two atoms together, as in

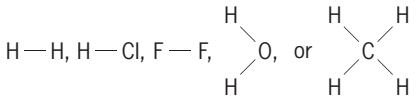

that is the most characteristic link in chemistry. The theory that accounts for it is a cornerstone of chemical science. The physical and chemical properties of any molecule are direct consequences of its particular detailed electronic structure. Yet the theory of any one covalent chemical bond, for example, the H—H bond in the hydrogen molecule, has much in common with the theory of any other covalent bond, for example, the O—H bond in the water molecule. The current theory of covalent bonds both

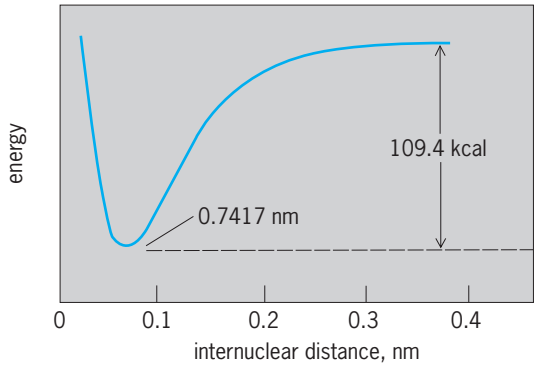

Potential energy of the hydrogen molecule.

treats their qualitative features and quantitatively accounts for the molecular properties which are a consequence of those features. The theory is a branch of quantum theory. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM CHEMISTRY.

Hydrogen molecule. A brief outline of the application of quantum theory to the bond in the hydrogen molecule H—H follows. Here two electrons, each of charge $-e$, bind together two protons, each of charge $+e$, with the electrons much lighter than the protons. What must be explained, above all else, is that these particles form an entity with the protons 0.074 nanometer apart, more stable by $D = 109$ kcal (456 kilojoules) per mole than two separate hydrogen atoms, where D is the binding energy. In more detail, a molecular energy is involved (ignoring nuclear kinetic energy) that depends on internuclear distance, as shown in the **illustration**. This curve can be determined experimentally, and it can be used to interpret the characteristic spectroscopic properties of hydrogen gas. *See* MOLECULAR STRUCTURE AND SPECTRA.

The quantum theory accounts for the properties of isolated atoms by assigning atomic orbitals for individual electrons to move in, not more than two electrons at a time. For the hydrogen atom, the orbitals are labeled $1s$, $2s$., $2p_x$, $2p_y$, $2p_z$, and so on, with $1s$ the one having the lowest energy. For the molecule H_2 one electron, say electron 1, might be assigned to a $1s$ orbital on proton A, with $1s_A(1)$ written to signify this; similarly electron 2 might be assigned to the same kind of orbital on proton B, written as $1s_B(2)$. Since independent probabilities multiply and orbitals represent probability amplitudes, the description for the combined system shown in Eq. (1) is arrived

$$\phi(1, 2) = 1s_A(1)1s_B(2) \qquad (1)$$

at. Unfortunately, this fails to account for the bond properties; it gives a binding energy of only 10 kcal (41.9 kJ) per mole. *See* ATOMIC STRUCTURE AND SPECTRA; EXCLUSION PRINCIPLE.

An essential defect of Eq. (1) is the numbering of the electrons; it puts electron 1 on proton A, electron 2 on proton B. Electrons cannot be distinguished experimentally, so they should not be given unique numbers; the function $1s_A(2)1s_B(1)$ would be just as

good as the foregoing. It is necessary to use a description that is not affected by interchange of electron labels, as in the additive combination of Eq. (2).

$$\phi(1, 2) = 1s_A(1)1s_B(2) + 1s_A(2)1s_B(1) \qquad (2)$$

(The difference combination also is an acceptable description, but it represents an excited state of the molecule.)

Any complete molecular electronic wave function should include electron spin. Symmetric space wave functions like Eq. (2) must be multiplied by antisymmetric spin wave functions to give total wave functions that are antisymmetrical with respect to interchange of electrons. For the ground state of hydrogen, and for the normal covalent bond elsewhere, this requirement means that the electron spins must be paired to give a total electron spin of zero.

The simple relationship described by Eq. (2) qualitatively accounts for the existence of the covalent bond; the predicted binding energy is $D = 74$ kcal (310 kJ) per mole; and the shape of the curve, with the minimum appearing at 0.080 nm, is right.

The description of Eq. (2) can be systematically improved. The charge acting on the electron may be changed from $+1e$ to the larger value, $+Ze$, which is more realistic for the actual molecule. With $Z = 1.17$ this gives $D = 87$ kcal (364 kJ) per mole. Polarization effects may be introduced by taking Eq. (3), where

$$\phi(1, 2) = 1\sigma_A(1)1\sigma_B(2) + 1\sigma_A(2)1\sigma_B(1) \qquad (3)$$

$1\sigma_A = 1s_A + \lambda 2pz_A$ and $1\sigma_B = 1s_B + \lambda 2pz_B$. This gives $D = 93$ kcal (389 kJ). Ionic terms may be introduced, acknowledging the possibility that both electrons may be on one atom, by taking Eq. (4) where C

$$\phi(1, 2) = C_1[1s_A(1)1s_B(2) + 1s_A(2)1s_B(1)] + C_2[1s_A(1)1s_A(2) + 1s_B(1)1s_B(2)] \qquad (4)$$

represents a constant. This also gives (with $Z = 1.19$) $D = 93$ kcal (389 kJ). Another possible approach is to include both ionic terms and polarization effects, and other terms involving $2s$, $3d$, $4f$, and other orbitals. If this is done, eventually one obtains the observed D value and a potential curve that is in complete agreement with experiment.

The linear mixing of terms such as $1s_A(1)1s_B(2)$ is called resonance; the method of mixing covalent and ionic structures is called the valence bond (VB) method. The particular mixing coefficients can be found by the variational principle: The best values for such parameters are those that make the total energy of the molecule, properly computed from quantum mechanics, a minimum. The energy expression contains only terms that have a direct classical interpretation: the kinetic energy of the electrons, their energy of repulsion for one another, their energy of attraction for the nuclei, and the nuclear-nuclear repulsion energy. *See* RESONANCE (MOLECULAR STRUCTURE).

Alternative descriptions of H_2 are possible, of which the most important is provided by the molecular orbital (MO) method. Here electrons are put one at a time into orbitals which are spread over the whole molecule, usually approximating these orbitals by linear combinations of atomic orbitals (LCAO). For H_2 the lowest molecular orbital is $\phi_1 \approx 1s_A + 1s_B$, the next $\phi_2 \approx 1s_A - 1s_B$. The simplest molecular orbital description is displayed in Eq. (5),

$$\phi(1, 2) = \phi_1(1)\phi_2(2) \qquad (5)$$

which represents an equal weighting of covalent and ionic structures; it gives $D = 61$ kcal (255 kJ) for $Z = 1.00$ and $D = 80$ kcal (335 kJ) for $Z = 1.20$. More suitable is a mixture of this function with the function obtained by promoting both electrons from ϕ_1 to ϕ_2. The result of this configuration interaction process has the form of Eq. (6), and it is identical with the valence bond function of Eq. (4). In this manner more terms can be added, using more orbitals, until, again, the accurate potential energy curve is obtained.

$$\phi(1, 2) = D_1\phi_1(2) + D_2\phi_2(1)\phi_2(2) \qquad (6)$$

The most accurate description known for the chemical bond in H_2 is a very complicated electronic wave function. The calculated and observed values of D and other properties agree absolutely.

Complex molecules. The problem of the proper description of chemical bonds in molecules that are more complicated than H_2 has many inherent difficulties. The qualitative theory of chemical bonding in complex molecules preserves the use of many chemical concepts that predate quantum chemistry itself; among these are electrostatic and steric factors, tautomerism, and electronegativity. The quantitative theory is highly computational in nature and involves extensive use of computers and supercomputers.

The number of covalent bonds which an atom can form is called the covalence and is determined by the detailed electron configuration of the atom. An extremely important case is that of carbon. In most of its compounds, carbon forms four bonds. When these connect it to four other atoms, the directions of the bonds to these other atoms normally make angles of about 109° to one another, unless the attached atoms are crowded or constrained by other bonds. That is, covalent bonds have preferred directions. However, in accord with the idea that carbon forms four bonds, it is necessary to introduce the notion of double and triple bonds. Thus in the structural formula of ethylene, C_2H_4 (**1**), all lines denote covalent bonds, the double line connecting the carbon atoms being a double bond. Such double bonds are distinctly shorter, almost twice as stiff, and require considerably more energy to break completely than do single bonds. However, they do not require twice as much energy to break as a single bond. Similarly,

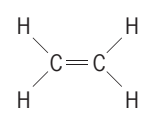

(1)

acetylene (**2**) is written with a triple bond, which

(**2**)

is still shorter than a double bond. A carbon-carbon single bond has a length close to 1.54×10^{-8} cm, whereas the triple bond is about 1.21×10^{-8} cm long. *See* BOND ANGLE AND DISTANCE; VALENCE.

In many compounds the rules for writing bond formulas are not unique. For example, benzene, C_6H_6 (**3**), can be written in two forms. Evidence

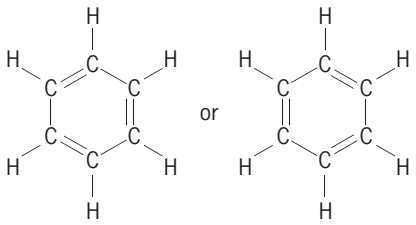

(**3**)

proves that all six C-C bonds are equivalent, so neither formula can be correct. The correct picture is a blend of the two, in which the bonds have many properties intermediate between those of double and single bonds but in which the whole molecule displays an additional stability. This resonance occurs whenever the structure is such that two or more different bond formulas can legitimately be drawn for the same geometry. *See* BENZENE; ELECTRON CONFIGURATION.

Many substances have some bonds which are covalent and others which are ionic. Thus in crystalline ammonium chloride, NH_4Cl, the hydrogens are bound to nitrogen by electron pairs, but the NH_4 group is a positive ion and the chlorine is a negative ion.

Both electrons of a covalent bond may come from one of the atoms. Such a bond is called a coordinate or dative covalent bond or semipolar double bond, and is one example of the combination of ionic and covalent bonding.

The hydrogen bond is a special bond in which a hydrogen atom links a pair of other atoms. The linked atoms are normally oxygen, fluorine, chlorine, or nitrogen. These four elements are all quite electronegative, a fact which favors a partially ionic interpretation of this kind of bonding. *See* ELECTRONEGATIVITY; HYDROGEN BOND.

To illustrate the level of accuracy of contemporary quantum-chemical calculations, the **table** gives observed and calculated values for certain spectroscopic properties of the H_2O molecule.

Spectroscopic properties of H_2O*

Property	Observed value	Calculated value
r_{OH} (Å†)	0.957	0.962
∠(HOH) (degrees)	104.5	104.6
Harmonic vibrational frequencies (cm^{-1})	3832	3912
	1649	1683
	3943	4041

* After B. H. Besler et al., *J. Chem. Phys.*, 89:360, 1988.
† 1 Å = 10^{-10} m.

Robert G. Parr

Bibliography. J. Murell, J. M. Tedder, and S. F. Kettle, *The Chemical Bond*, 2d ed., 1985; L. Pauling, *The Nature of the Chemical Bond and the Structure of Molecules and Crystals*, 3d ed., 1960; L. Pauling and E. B. Wilson, Jr., *Introduction to Quantum Mechanics*, 1935, reprint 1985; M. J. Winter, *Chemical Bonding*, 1994.

Chemical conversion

A chemical manufacturing process in which chemical transformation takes place, that is, the product differs chemically from the starting materials. Most chemical manufacturing processes consist of a sequence of steps, each of which involves making some sort of change in either chemical makeup, concentration, phase state, energy level, or a combination of these, in the materials passing through the particular step. If the changes are of a strictly physical nature (for example, mixing, distillation, drying, filtration, adsorption, condensation), the step is referred to as a unit operation. If the changes are of a chemical nature, where conversion from one chemical species to another takes place (for example, combustion, polymerization, chlorination, fermentation, reduction, hydrolysis), the step is called a unit process. Some steps involve both, for example, gas absorption with an accompanying chemical reaction in the liquid phase. The term chemical conversion is used not only in describing overall processes involving chemical transformation, but in certain contexts as a synonym for the term unit process. The chemical process industry as a whole has tended to favor the former usage, while the petroleum industry has favored the latter. *See* CHEMICAL PROCESS INDUSTRY; UNIT PROCESSES.

Another usage of the term chemical conversion is to define the percentage of reactants converted to products inside a chemical reactor or unit process. This quantitative usage is expressed as percent conversion per pass, in the case of reactors where unconverted reactants are recovered from the product stream and recycled to the reactor inlet. *See* CHEMICAL ENGINEERING.

William F. Furter

Bibliography. N. Basta, *Shreve's Chemical Process Industries Handbook*, 6th ed., 1994.

Chemical dynamics

That branch of physical chemistry which seeks to explain time-dependent phenomena, such as energy transfer and chemical reaction, in terms of the detailed motion of the nuclei and electrons which constitute the system.

Molecular Dynamics

In principle, it is possible to prepare two reagents in specific quantum states and to determine the

quantum-state distribution of the products. In practice, this is very difficult, and experiments have mostly been limited to preparing one reagent or to determining some aspect of the product distribution. This approach yields data concerning the detailed aspects of the dynamics.

Bimolecular reactions. Molecular-beam and luminescence techniques have played a major role in the development of chemical dynamics. Since these techniques largely complement each other, they are illustrated by discussing the results of a single reaction, the formation of deuterium chloride (DCl) from atomic deuterium (D) and molecular chlorine (Cl_2), reaction (1).

$$D + Cl_2 \rightarrow DCl + Cl \qquad (1)$$

Chemiluminescence. Reaction (1) is 46 kilocalories per mole (192 kilojoules per mole) exoergic. If all of the exoergicity were to go into vibration of the newly formed DCl molecule, enough energy would be available to populate the $v' = 9$ vibrational level. In a gas-phase reaction at high pressure, much of this energy would be dissipated as heat at the walls of the container; but at very low pressures, where the frequency of collisions is small, DCl molecules in excited vibrational states, DCl*, will emit infrared emission prior to undergoing a collision.

An apparatus to study this infrared chemiluminescence is shown in **Fig. 1**. Deuterium atoms are made by dissociating D_2 gas in an electrical discharge, and together with Cl_2 are injected into the observation cell at the top. The reagents mix and react inside a vessel with walls maintained at 77 K (−320°F), which freeze out species hitting the wall. The pressure is kept low to minimize vibrationally deactivating collisions. Infrared emission from the products is gathered by mirrors at the end, taken out through the sapphire window, and analyzed with an infrared spectrometer. By analyzing the spectrum, it is possible to determine which vibration-rotation states are emitting and, as a consequence, which vibration-rotation states are formed in the reaction. **Figure 2** shows the relative distribution of vibrational states from the reaction, and also shows, for comparison, the relative distribution of vibrational states calculated from the Boltzmann equation for hot DCl (6000 K or 10,800°F). Thermal distributions at any other temperatures would still show a monotonic decline. (The population for $v' = 0$ is not determined by the chemiluminescence experiments, because that state does not emit.) The DCl formed in the reaction clearly has different properties than hot DCl: the vibrational population displays an inversion, and this system (the hydrogen isotope) was the active medium for the first chemically pumped laser. *See* CHEMILUMINESCENCE.

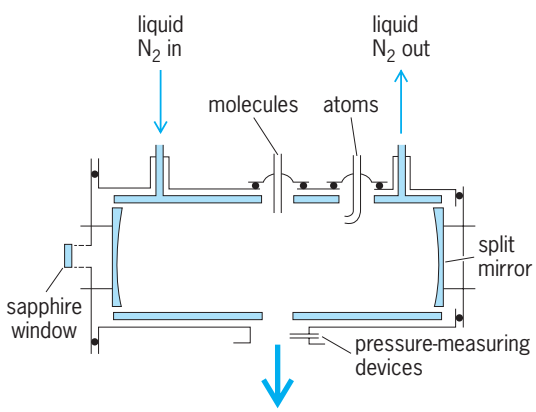

Fig. 1. Reaction vessel for studying infrared chemiluminescence between atoms and molecules at low pressure. (*After D. H. Maylotte, J. C. Polanyi, and K. B. Woodall, J. Chem. Phys., 57:1547–1561, 1972*)

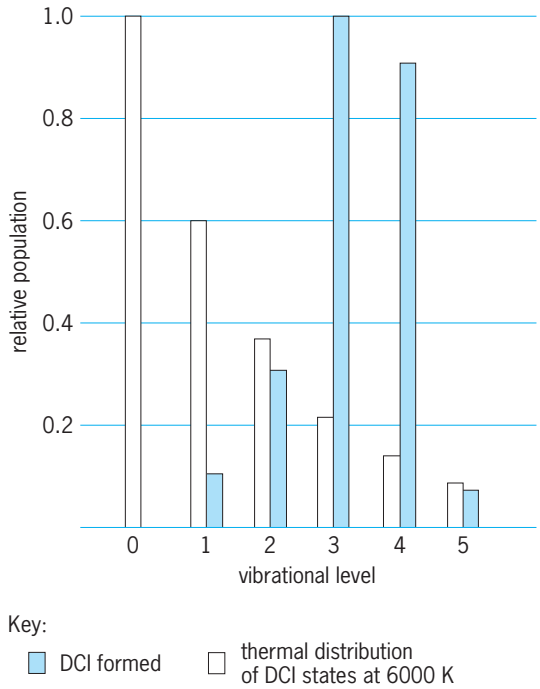

Fig. 2. Relative populations of DCl molecules in different vibrational states formed in the reaction $D + Cl_2 \rightarrow DCl + Cl$. 6000 K = 10,800°F.

Molecular-beam experiments. Molecules can be isolated in molecular beams, and collisions between these isolated molecules can be observed by crossing two tenuous molecular beams in a region of otherwise high vacuum. **Figure 3** shows such an experiment. Gaseous atoms or molecules emerge from the ovens, and collimating slits select the particles which are all going in the same direction. The molecular beams cross at the center of rotation of a large platform which can be rotated under vacuum relative to the two beams. Large vacuum pumps maintain a high vacuum, which ensures that collisions take place only at the intersection of the two beams. Product molecules are ionized by electron bombardment, and detected with a quadrupole mass spectrometer housed within a region of ultrahigh vacuum.

Measurements are made of the scattered product intensity and speed at various scattering angles. For ease of interpretation, these data are transformed into the center-of-mass system in which the two reagents approach each other with equal and opposite momenta. **Figure 4** shows a contour map of the DCl intensity in the center-of-mass coordinate system

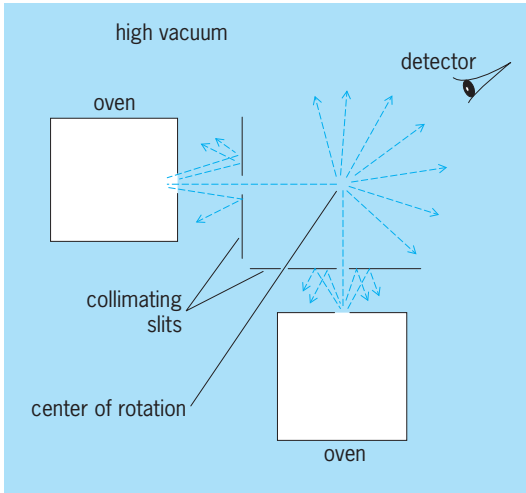

Fig. 3. Schematic of a molecular beam experiment; entire apparatus is under high vacuum.

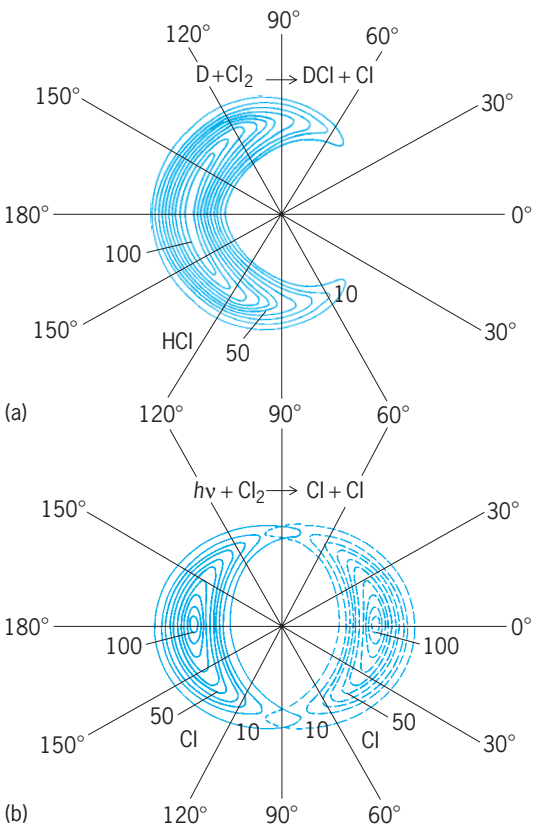

Fig. 4. Contour map of the DCl intensity. (*a*) Center-of-mass angular distribution of DCl from the reaction D + Cl_2. (*b*) Angular distribution of Cl atoms from photodissociation of Cl_2 molecules. Broken lines correspond to the second Cl atom. (*After D. R. Herschbach, Molecular dynamics of chemical reactions, Pure Appl. Chem., 47:61–73, 1976*)

in which the D atom is incident from the left and the Cl_2 molecule is incident from the right. The product DCl recoils backward (in the direction from which the D came) in a broad but nonetheless anisotropic distribution. The speed of the product is high, and corresponds to about half of the reaction exoergicity appearing in translation recoil of the products, with the balance appearing in vibration and rotation of the DCl; this is consistent with the chemiluminescence results. *See* MOLECULAR BEAMS.

The anisotropic product distribution shows that reaction occurs in a time of less than a molecular rotation, about 1 picosecond. The partitioning of energy roughly equally between vibration and translation suggests that the major amount of energy is released in repulsion between the DCl and Cl. This repulsion is similar to that experienced by a Cl_2 molecule in photodissociation, also shown in Fig. 4. Because the deuterium atom is so light, the direction in which the product is expelled is a measure of the orientation of the Cl_2 molecule during reaction. For this reaction the collinear arrangement D-Cl-Cl is preferred, whereas bent arrangements are preferred for D + I_2.

Other investigations. Molecular-beam machines (and to some extent chemiluminescence machines) have been modified in various ways to explore even finer details of specific reactions. For example, electric and magnetic fields have been used to prepare reagents in various orientations or to determine the magnitude and polarization of the product rotational angular momentum; lasers have been used to prepare reagents in initial vibrational states and to determine the final vibration-rotation state of products by inducing fluorescence from the products; and reactions may be conducted at hyperthermal energies to produce dissociation or even ionization.

Advances in subnanosecond laser technology have made it possible to study a few bulk phase reactions under nearly collision-free circumstances. For example, H atoms can be formed by dissociating a hydrogen halide (HX) with a fast laser pulse, and reaction of the H can be monitored by a second pulse delayed enough to probe reaction products after only a few collisions. *See* ULTRAFAST MOLECULAR PROCESSES.

Information about chemical reactions has so far come from observations made on the disappearance of reagents or appearance of products. However, several experiments on systems such as H + H_2, K + NaCl, and ICN (iodine cyanide) have shown that it is possible to use lasers to probe the chemical system during the very short time in which the chemical bonds are rearranging. This forms the basis of a new spectroscopy of the transition region. *See* LASER PHOTOCHEMISTRY.

Unimolecular reactions. Unimolecular reactions are possible when the energy of an isolated molecule lies above the energy of the asymptotically separated products. Molecules were traditionally energized by collisions, and unimolecular reactions have been studied by using collisional excitation in both cross-beam and chemiluminescence experiments. Several laser methods—multiphoton excitation, stimulated emission pumping, and overtone excitation—can also be used to populate a very narrow range of energy levels in isolated molecules to study unimolecular reactions as well as the related process of photodissociation.

Considerable insight into unimolecular processes has been obtained. Energy initially deposited

locally is distributed on a picosecond time scale, and most processes seem to be described by a statistical model. Furthermore, bond-selective photochemistry (in which different reactions would occur for excitation of different molecular vibrations) seems unlikely, because reaction is usually slow compared to the energy redistribution process. *See* PHOTOCHEMISTRY.

Philip R. Brooks

Theoretical Methods

The goal of chemical dynamics is to understand kinetic phenomena from the basic laws of molecular mechanics, and it is thus a field which sees close interplay between experimental and theoretical research.

Energy distribution. An important question regarding the dynamics of chemical reactions has to do with the product energy distribution in exothermic reactions. For example, because the hydrogen fluoride (HF) molecule is more strongly bound than the H_2 molecule, reaction (2) releases a consider-

$$F + H_2 \longrightarrow HF + H \qquad (2)$$

able amount of energy (more than 30 kcal/mol or 126 kJ/mol). The two possible paths for this energy release to follow are into translations, that is, with HF and H speeding away from each other, or into vibrational motion of HF.

In this case it is vibration, and this has rather dramatic consequences; the reaction creates a population inversion among the vibrational energy levels of HF—that is, the higher vibrational levels have more population than the lower levels—and the emission of infrared light from these excited vibrational levels can be made to form a chemical laser. A number of other reactions such as reaction (1) also give a population inversion among the vibrational energy levels, and can thus be used to make lasers.

Most effective energy. The rates of most chemical reactions are increased if they are given more energy. In macroscopic kinetics this corresponds to increasing the temperature, and most reactions are faster at higher temperatures. It seems reasonable, though, that some types of energy will be more effective in accelerating the reaction than others. For example, in reaction (3), where potassium (K) reacts with hy-

$$K + HCl \longrightarrow KCl + H \qquad (3)$$

drogen chloride (HCl) to form potassium chloride (KCl), studies in a molecular beam have shown that if HCl is vibrationally excited (by using a laser), this reaction is found to proceed approximately 100 times faster, while the same amount of energy in translational kinetic energy has a smaller effect. Here, therefore, vibrational energy is much more effective than translational energy in accelerating the reaction.

For reaction (2), however, translational energy is more effective than vibrational energy in accelerating the reaction. The general rule of thumb is that vibrational energy is more effective for endothermic reactions (those for which the new molecule is less stable than the original molecule), while translational energy is most effective for exothermic reactions.

Lasers. As seen from the above, lasers are also an important supplement to molecular-beam techniques for probing the dynamics of chemical reactions. Because they are light sources with a very narrow wavelength, they are able to excite molecules to specific quantum states (and also to detect what states molecules are in), an example of which is reaction (3). For polyatomic molecules—that is, those with more than two atoms—there is the even more interesting question of how the rate of reaction depends on which vibration is excited.

For example, when the molecule allyl isocyanide, $CH_2{=}CH{-}CH_2{-}NC$, is given sufficient vibrational energy, the isocyanide part (—NC) will rearrange to the cyanide (—CN) configuration. A laser can be used to excite a C-H bond vibrationally. An interesting question is whether the rate of the rearrangement process depends on which C-H bond is excited. Only with a laser is it possible to excite different C-H bonds and begin to answer such questions. This question of mode-specific chemistry, that is, the question of whether excitation of specific modes of a molecule causes specific chemistry to result, has been a subject of great interest. (For the example above, the reaction is fastest if the C-H bond closest to the NC group is excited.) Mode-specific chemistry would allow much greater control over the course of chemical reactions, and it would be possible to accelerate the rate of some reactions (or reactions at one part of a molecule) and not others. *See* LASER.

Models and methods. Many different theoretical models and methods have been useful in understanding and analyzing all of the phenomena described above. Probably the single most useful approach has been the calculation of classical trajectories. Assuming that the potential energy function or a reasonable approximation is known for the three atoms in reaction (2), for example, it is possible by use of electronic computers to calculate the classical motion of the three atoms. It is thus an easy matter to give the initial molecule more or less vibrational or translational energy, and then compute the probability of reaction. Similarly, the final molecule and atom can be studied to see where the energy appears, that is, as translation or as vibration.

It is thus a relatively straightforward matter theoretically to answer the questions and to see whether or not mode-specific excitation leads to significantly different chemistry than simply increasing the temperature under bulk conditions.

The most crucial step in carrying out these calculations is obtaining the potential energy surface—that is, the potential energy as a function of the positions of the atoms—for the system. **Figure 5** shows a plot of the contours of the potential energy surface for reaction (2). Even without carrying out classical trajectory calculations, it is possible to deduce some of the dynamical features of this reaction; for example, the motion of the system first surmounts a small potential barrier, and then it slides down a steep hill,

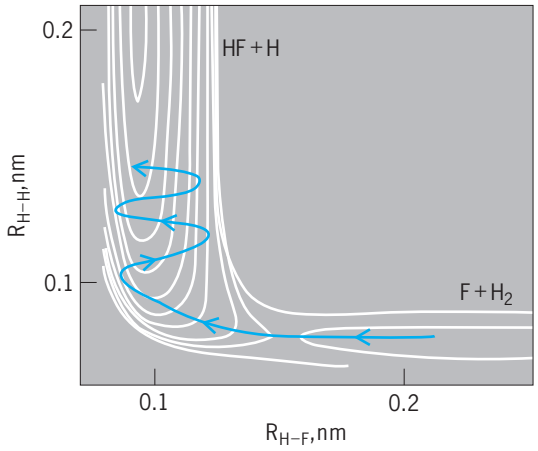

Fig. 5. Contour plot of the potential energy surface for the reaction $F + H_2 = HF + H$, with a typical reactive trajectory indicated.

turning the corner at the bottom of the hill. It is evident that such motion will cause much of the energy released in going down the hill to appear in vibrational motion of HF.

This and other theoretical methods have interacted strongly with experimental research in helping to understand the dynamics of chemical reactions. *See* CHEMICAL KINETICS; INORGANIC PHOTOCHEMISTRY.

William H. Miller

Bibliography. M. Baer (ed.), *Theory of Chemical Reaction Dynamics*, vol. 1, 1985; G. D. Billing and K. V. Mikkelsen, *Advanced Molecular Dynamics and Chemical Kinetics*, John Wiley, 1997; F. A. Gianturco (ed.), *Collision Theory for Atoms and Molecules*, 1989; D. R. Herschbach, Molecular dynamics of elementary chemical reactions, *Angew. Chem. Int. Ed. Engl.*, 26:1221–1243, 1987; K. Kuchitsu (ed.), *Dynamics of Excited Molecules*, 1994; W. H. Miller (ed.), *Dynamics of Molecular Collisions*, 1976; J. I. Steinfeld, J. S. Francisco, and W. L. Hase, *Chemical Kinetics and Dynamics*, 2d ed., Prentice Hall, 1998; A. H. Zewail, *Femtochemistry: Ultrafast Dynamics of the Chemical Bond*, vols. 1 and 2, 1994.

Chemical ecology

The study of ecological interactions mediated by the chemicals that organisms produce. These substances, known as allelochemicals, serve a variety of functions, the growing list of which reflects the rapid discoveries made in this discipline. These compounds influence or regulate interspecific and intraspecific interactions of microorganisms, plants, and animals, and operate within and between all trophic levels—producers, consumers, and decomposers—and in terrestrial, fresh-water, and marine ecosystems.

Function is an important criterion for the classification of allelochemicals. Allelochemicals beneficial to the emitter are called allomones; those beneficial to the recipient are called kairomones. An allomone to one organism can be a kairomone to another. For example, floral scents benefit the plant (allomones) by encouraging pollinators, but also benefit the insect (kairomones) by providing a cue for the location of nectar.

The chemicals involved are diverse in structure and are often of low molecular weight (<10,000). They may be volatile (for example, monoterpene essential oils from peppermint, cyanide from trefoil, musks of mammals, mustard oils from cabbage, aldehyde pheromones in insects) or nonvolatile (such as nicotine alkaloids in tobacco, flavonoids in most plants), and water-soluble (phenols in many insect defenses) or fat-soluble (leaf waxes). Proteins, polypeptides, and amino acids are also found to play an important role.

Plant allelochemicals are often called secondary compounds or metabolites to distinguish them from those chemicals involved in primary metabolism, although this distinction is not always clear. The functions of secondary chemicals were a mystery until 1959, when G. Fraenkel emphasized their potential importance in the interactions of plants and animals. The role of chemicals in the ecology of microorganisms and of animals was recognized earlier: since the discovery of penicillin, scientists have known that fungi can produce substances that destroy bacteria, and the odors of such animals as stink bugs and skunks serve an obvious function.

Chemical defense in plants. Perhaps to compensate for their immobility, plants have made wide use of chemicals for protection against competitors, pathogens, herbivores, and abiotic stresses.

Production of defenses by plants is not cost-free; it requires energy that could be used for other purposes. Although the actual cost is usually not known, many plants do concentrate defensive compounds in those tissues that contain the highest energy reserves or growth potential, such as new leaves or seeds. Some plants may also change the amount of resources used for defense, in response to changes in their environment, and the availability of the raw materials for making defenses, for example, carbon from photosynthesis and nitrogen from soil.

Allelopathy. A chemically mediated competitive interaction between higher plants is referred to as allelopathy. This is one of the few examples in plants where the effects of chemicals can be obvious: observations that areas around the bases of walnut trees are free of plant growth led to the first description of allelopathy. In walnuts, the mechanism involves leaching of the chemical juglone (5-hydroxy napthoquinone) from the leaves to the ground, and this inhibits growth and germination of a variety of plants. Allelopathy appears to occur in many plants, may involve phenolics or terpenoids that are modified in the soil by microorganisms, and is at least partly responsible for the organization of some plant communities. *See* ALLELOPATHY.

Active and passive defenses. Chemicals that are mobilized in response to stress or attack are referred to as active or inducible chemicals, while those that are always present in the plant are referred to as passive or constitutive. In many plants, fungus attack

induces the production of defensive compounds called phytoalexins, a diverse chemical group that includes isoflavonoids, terpenoids, polyacetylenes, and furanocoumarins. The phytoalexin is produced in the area immediately surrounding the fungus and probably acts by interfering with the nutrition or membrane permeability of the fungal cells. Plant susceptibility to fungus attack is partly related to the speed of production of phytoalexins. *See* PHYTOALEXINS.

Defensive chemicals can be induced by herbivore attack. For example, when snowshoe hares feed on paper birch there is an increase in the levels of phenolic resins in new growth. These levels decrease with time, and the activity of these induced chemicals is thought to cause population cycles of hares. There has been increasing evidence that inducible defenses, such as phenolics, are important in plant–insect interactions.

Constitutive defenses include the chemical hydrogen cyanide. Trefoil, clover, and ferns have been found to exist in two genetically different forms, one containing cyanide (cyanogenic) and one lacking it (acyanogenic); acyanogenic forms are often preferred by several herbivores. Acacia trees also have these two forms; the acyanogenic forms are protected from herbivores by ants that live in the tree and feed on special food bodies on the leaves. Other constitutive chemicals include some phenolic acids, flavonoids, pyrethrins, and alkaloids. *See* ALKALOID; FLAVONOID; PHENOL.

Diversity of activities. Chemical defenses frequently occur together with certain structures which act as physical defenses, such as spines and hairs. In tomatoes, glandular trichomes (hairlike structures) on the leaf surface contain the toxic chemical 2-tridecanone that is active against many insects, while other glands contain phenolics that polymerize like fast-setting glue.

While many chemicals protect plants by deterring herbivore feeding or by direct toxic effects, other defenses may act more indirectly. Chemicals that mimic juvenile hormones, the antijuvenile hormone substances found in some plants, either arrest development or cause premature development in certain susceptible insect species.

Plant chemicals potentially affect not only the herbivores that feed directly on the plant, but also the microorganisms, predators, or parasites of the herbivore. For example, the tomato plant contains an alkaloid, tomatine, that is effective against certain insect herbivores. The tomato hornworm, however, is capable of detoxifying this alkaloid and can thus use the plant successfully—but a wasp parasite of the hornworm cannot detoxify tomatine, and its effectiveness in parasitizing the hornworm is reduced. Therefore, one indirect effect of the chemical in the plant may be to reduce the effectiveness of natural enemies of the plant pest, thereby actually working to the disadvantage of the plant.

Types of plant defenses. Most plant chemicals can affect a wide variety of herbivores and microorganisms, because the modes of action of the chemicals they manufacture are based on a similarity of biochemical reaction in most target organisms (for example, cyanide is toxic to most organisms). In addition, many plant chemicals may serve multiple roles: resins in the creosote bush serve to defend against herbivores and pathogens, conserve water, and protect against ultraviolet radiation.

It is argued that there are two different types of defensive chemicals in plants. The first type occurs in relatively small amounts, is often toxic in small doses, and poisons the herbivore. These compounds, for example, some phenolic acids, alkaloids, and monoterpenes, may also change in concentration in response to plant damage; that is, they are inducible. These kinds of qualitative defensive compounds are the most common in short-lived or weedy species that are often referred to as unapparent. They are also characteristic of fast-growing species with short-lived leaves. In contrast, the second type of defensive chemicals often occurs in high concentrations, is not very toxic, but may inhibit digestion by herbivores and is not very inducible. These quantitative defenses, such as tannins, are most common in long-lived, so-called apparent plants such as trees that have slow growth rates and long-lived leaves. Some plants may use both types of defenses.

Aquatic habitats. There is accumulating evidence that marine plants may be protected against grazing by similar classes of chemicals to those found in terrestrial plants. In algae, for example, phenols appear to be an important defense against limpets. One interesting difference in the marine environment is the large number of halogenated organic compounds that are rare in terrestrial and fresh-water systems.

Evolution of defenses and adaptation. Through evolution, as plants accumulate defenses, herbivores that are able to bypass the defense in some way are selected for and leave more offspring than others. This in turn selects for new defenses on the part of the plant in a continuing process called coevolution.

Animals that can exploit many plant taxa are called generalists, while those that are restricted to one or a few taxa are called specialists. Specialists often have particular detoxification mechanisms to deal with specific defenses. For example, certain bruchid beetles that feed on seeds containing the toxic nonprotein amino acid canavanine possess enzyme systems that recognize this "false" amino acid and avoid incorporating it into proteins. In addition, they are able to convert the amino acid to ammonia via urea and generate energy in the process. Some generalists possess powerful, inducible detoxification enzymes, while others exhibit morphological adaptations of the gut which prevent absorption of compounds such as tannins, or provide reservoirs for microorganisms that accomplish the detoxification. Animals may avoid eating plants, or parts of plants, with toxins. For instance, monkeys choose low-toxin, high-nutrient plant tissues and are capable of learning which plants to avoid.

Some herbivores that have completely surmounted the plant toxin barrier use the toxin itself as a cue to aid in locating plants. The common white

butterfly, *Pieris rapae*, for example, uses mustard oil glycosides, which are a deterrent and toxic to many organisms, to find its mustard family hosts. The monarch butterfly and other insects have gone so far as to store plant defensive chemicals in their bodies as protection against predators.

Chemical defense in animals. Many animals make their own defensive chemicals—such as all of the venoms produced by social insects (bees, wasps, ants), as well as snakes and mites. These venoms are usually proteins, acids or bases, alkaloids, or combinations of chemicals. They are generally injected by biting or stinging, while other defenses are produced as sprays, froths, or droplets from glands. For instance, the bombardier beetle expels hot quinones as a spray from its posterior. A response of one predatory mouse to this action is to bury the beetle's dangerous posterior into the sand before biting the head. There are also darkling beetles that mimic the bombardier's actions but do not have a defensive spray. While insects provide the best examples of chemical defense in animals, such defenses are also found in other groups, such as in stinging cells of jellyfish and in poisonous spurs on the duckbill platypus.

Animals frequently make the same types of toxins as plants, presumably because their function as protective agents is similar. Examples include the production of cyanide by millipedes and bracken fern, phenols in the secretions of grasshoppers and in most plant leaves, alkaloids in ant venoms and pepper plants, and cardiac glycosides in milkweed plants and beetles.

Other organisms, particularly insects, use plant chemicals to defend themselves. For example, the monarch butterfly sequesters or stores cardenolides from the milkweed plant, and the lubber grasshopper can sequester a variety of plant chemicals into its defensive secretion by feeding exclusively on certain plants. Sequestration may be a low-cost defense mechanism and probably arises when insects specialize on particular plants.

Microbial defenses. Competitive microbial interactions are regulated by many chemical exchanges involving toxins. They include compounds such as aflatoxin, botulinus toxin, odors of rotting food, hallucinogens, and a variety of antibiotics. *See* ANTIBIOTIC; TOXIN.

Microorganisms also play a role in chemical interaction with plants and animals that range from the production of toxins that kill insects, such as those produced by the common biological pest control agent *Bacillus thuringiensis*, to cooperative biochemical detoxification of plant toxins by animal symbionts.

Information exchange. A large area of chemical ecology concerns the isolation and identification of chemicals used for communication.

Insects. Pheromones, substances produced by an organism that induce a behavioral or physiological response in an individual of the same species, have been studied particularly well in insects. These signals are compounds that are mutually beneficial to the emitter and sender, such as sex attractants, trail markers, and alarm and aggregation signals.

Sex pheromones are volatile substances, usually produced by the female to attract males. Each species has a characteristic compound that may differ from that of other species by as little as a few atoms. For example, the pheromone of the cabbage looper moth is *cis*-7-dodecenyl acetate; the oriental fruit fly's is *cis*-8-dodecenyl acetate. The activity of these compounds can be destroyed by minor changes in their structure. In addition, sex pheromones generally are active at extremely low concentrations—a few molecules impinging on a moth's antenna may be sufficient to stimulate a response. The pheromone blend released by a female moth may be dominated by one or two components but also contains many trace components. The dominant compounds may attract males, but the full repertoire of male behavior that results in successful mating also involves the trace components.

Pheromones are typically synthesized directly by the animal and are usually derived from fatty acids. In a few cases the pheromone or its immediate precursors may be derived from plants, as in danaid butterflies. Male danaids visit certain nonhost plants solely to obtain pyrrolizidine alkaloids, which are used to make pheromones used in courtship.

Trail pheromones are used by social insects and some lepidopterans to recruit others of its species to a food source. These are synthesized by the animal as well as acquired from hosts, and are diverse in structure. The honeybee uses geraniol, which it acquires from nectar. Methyl 4-methylpyrrole 2-carboxylate, the trail pheromone of the leaf-cutting ant, may be synthesized in the gut by bacteria. The ideal trail pheromone loses its potency over time unless reinforced, and is therefore relatively unstable and difficult to analyze chemically. *See* PHEROMONE.

Vertebrates. Very little work has been done in identifying specific pheromones in vertebrates, particularly mammals. It is known, however, that they are important in marking territory, in individual recognition, and in mating and warning signals. Frequently these signals involve sebaceous oils from the skin, urine, or feces. Some mammals have a very complex series of glands; it is possible that microorganisms on these glands convert urine compounds to components of the signal chemical. The complexity of this chemical mix, and the complexity of the animal, means that the strong positive or negative response to single compounds that is found in some insects is not found as often in mammalian systems. In female primates of several species, mixtures of short-chain organic acids, such as acetic and butanoic, lead to sexual excitation in the male. These same chemicals have been found in human females, but in different concentrations. *See* REPRODUCTIVE BEHAVIOR; SCENT GLAND; TERRITORIALITY.

Plants and microorganisms. Chemical communication may also occur among plants and microorganisms, although it is rarer and less obvious than in animals. The release of ethylene by ripening fruit stimulates the ripening of nearby fruit, thus ensuring

simultaneous presentation of an attractive mass for potential dispersers. Certain plants produce high concentrations of volatile isoprenes. The concentrations and composition of these compounds change under stress, but it is not known whether this signal, if it is one, is recognized by other plants.

Bacterial contact and recognition is accomplished by cell wall proteins and by secretions. This type of communication is important in strain recognition, and also takes place in fungi such as *Mucor* prior to mating. Here chemical communication is at the cellular level.

Value. The chemicals involved in chemical interactions between organisms are important in the pharmaceutical, perfume, and many other chemical industries. These substances, together with many antibiotics and pesticides, are based on compounds found in plants. However, information about chemical ecology is only just beginning to be exploited, particularly in the field of plant protection. For example, pheromones which are highly specific and nontoxic are used to trap insect pests for a variety of purposes, including monitoring and population reduction. As the mechanisms of pathogen resistance are understood, the proper chemical attributes may be bred or genetically engineered into plants. For example, the genes for production of a legume seed protein toxin have been transferred to tobacco leaves. The genes were expressed and the leaves became protected from insect attack. The potential for exploiting naturally occurring defenses may thus reduce reliance on costly, potentially hazardous synthetic chemicals. *See* ECOLOGY; GENETIC ENGINEERING; INSECT CONTROL, BIOLOGICAL; INSECTICIDE.

Clive G. Jones; Alcinda C. Lewis

Bibliography. W. G. Abrahamson (ed.), *Plant-Animal Interactions*, 1989; W. J. Bell and R. T. Curde (eds.), *Chemical Ecology of Insects*, vol. 2, 1995; M. S. Blum, *Chemical Defenses of Arthropods,* 1981; R. F. Denno and M. S. McClure (eds.), *Variable Plants and Herbivores in Natural and Managed Systems*, 1983; J. B. Harborne, *Introduction to Ecological Biochemistry*, 4th ed., 1994; D. H. Nordlund, R. L. Jones, and W. J. Lewis (eds.), *Semiochemicals: Their Role in Pest Control*, 1981; G. A. Rosenthal and D. H. Janzen (eds.), *Herbivores: Their Interaction with Secondary Plant Metabolites*, 2d ed., 1991.

Chemical energy

A useful but obsolescent term for the energy available from elements and compounds when they react, as in a combustion reaction. In precise terminology, there is no such thing as chemical energy, since all energy is stored in matter as either kinetic energy or potential energy. *See* COMBUSTION; ENERGY.

When a chemical reaction takes place, the atoms of the reactants change their bonding pattern and become products. The breaking of bonds in the reactants requires energy, and the formation of bonds in the products releases energy. The net change in energy is commonly referred to as chemical energy. To be more precise, when a reaction takes place, there is an overall change in the enthalpy H of the system as bonds are broken and new bonds are formed. This change in enthalpy is denoted ΔH. Under standard conditions [a pressure of 1 bar (100 kilopascals) and all substances pure], the change is noted ΔH° and called the standard enthalpy of reaction. Provided the pressure is constant, the standard enthalpy can be identified with the energy released as heat (when $\Delta H^\circ < 0$) or gained as heat ($\Delta H^\circ > 0$) when the reaction takes place. Reactions for which $\Delta H^\circ < 0$ are classified as exothermic; those for which $\Delta H^\circ > 0$ are classified as endothermic. All combustions are exothermic, the released heat being used either to provide warmth or to raise the temperature of a working fluid in an engine of some kind. There are very few common endothermic reactions; one example is the dissolution of ammonium nitrate in water (a process utilized in medical cold packs). *See* CHEMICAL EQUILIBRIUM; CHEMICAL THERMODYNAMICS; ENTHALPY; THERMOCHEMISTRY.

The "chemical energy" available from a typical fuel (that is, the enthalpy change accompanying the combustion of the fuel, when carbon-hydrogen bonds are replaced by stronger carbon-oxygen and hydrogen-oxygen bonds) is commonly reported as either the specific enthalpy or the enthalpy density. The specific enthalpy is the standard enthalpy of combustion divided by the mass of the reactant. The enthalpy density is the standard enthalpy of combustion divided by the volume of the reactant. The former is of primary concern when mass is an important consideration, as in raising a rocket into orbit. The latter is of primary concern when storage space is a limitation. The specific enthalpy of hydrogen gas is relatively high (142 kilojoules/g), but its enthalpy density is low (13 kJ/L). The values for octane, a compound representative of gasoline, are 48 kJ/g and 38 MJ/L, respectively (note the change in units). The high enthalpy density of octane means that a gasoline tank need not be large to store a lot of "chemical energy." *See* AIRCRAFT FUEL; ENERGY SOURCES; GASOLINE; ROCKET PROPULSION.

Peter W. Atkins

Chemical engineering

The branch of engineering involved in the research, design, improvement, operation, and commercialization of plants, processes, and products for the chemical industry and related fields, such as biotechnology, energy, and materials. The work of chemical engineers has evolved continuously with the changing needs of society. Chemical engineering arose from attempts at inventing ways to make products such as gunpowder and wine. Early on (in the nineteenth century), the large-scale and efficient manufacturing of inorganic chemicals, such as salt, soda ash, sulfuric acid, and nitrogen fixation, were the main occupation of American chemical engineers. However, for the small-scale dyestuff and drug industries in Europe, chemical engineers were less involved. In the twentieth century, large-scale petroleum refining, petrochemicals, and polymers were the dominant industrial problems. In the twenty-first

century, the focus has shifted to smaller-scale processes across many disciplines of science and technology, such as molecular biology, genetics, information technology, and nanotechnology. *See* GENETICS; INFORMATION TECHNOLOGY; INORGANIC CHEMISTRY; NANOTECHNOLOGY; PETROCHEMICAL; POLYMER; PROCESS ENGINEERING.

Professional work and achievements. Chemical engineers work in the development, manufacturing, and service sectors. They are engaged in the design, construction, and operation of equipment and plants. Increasingly, chemical engineers are involved in product engineering, which is the design of new products and the improvement of existing products such as controlled drug-delivery devices, biodegradable polymers, and fuel cells. In addition, chemical engineers play an important role in the commercialization of novel medicines, materials, and environmental remediation and recycling technology. *See* DRUG DELIVERY SYSTEMS; FUEL CELL; RECYCLING TECHNOLOGY.

Chemical engineers often play leading roles in the manufacturing industries of basic chemicals and petroleum refining. They work in pharmaceutical, biochemical, biomedical, food, agricultural chemicals, plastics, rubber, textile fibers, paper, ink, soap and cosmetics, and water purification industries. In the electronics industry, the manufacture of integrated circuits and other microelectronic devices involves chemical processing. Chemical engineers also are employed in services, such as education, government, business and professional services, and, increasingly, health care. *See* BIOCHEMICAL ENGINEERING; BIOMEDICAL CHEMICAL ENGINEERING; CHEMICAL PROCESS INDUSTRY; FOOD ENGINEERING; INK; MANUFACTURED FIBER; MICROLITHOGRAPHY; PAPER; PHARMACEUTICAL CHEMISTRY; RUBBER; TEXTILE.

Some characteristic operations in chemical engineering involve the conversion of raw materials (such as petroleum) into useful products (such as gasoline, jet fuel, and asphalt), as well as the conversion and recycling of environmental pollutants (such as municipal waste). Other conversions involve chemical reactions such as the fixation of nitrogen from air into fertilizers for agriculture, physical modifications such as blowing polymers into films and extruding polymers into fibers, and biochemical reactions such as the fermentation of sugar into alcohol. These conversions also involve separations, such as making freshwater out of seawater, recovering penicillin from very dilute solutions in fermentation tanks, and extracting water and waste products from blood in artificial kidneys. *See* CHEMICAL CONVERSION; CHEMICAL SEPARATION; FERMENTATION; NITROGEN FIXATION; PETROLEUM PROCESSING AND REFINING; PLASTICS PROCESSING; WATER DESALINATION.

Education and professional training. Chemical engineers require knowledge in chemistry, biology, mathematics, physics, and machinery. The education and training of chemical engineers includes a rigorous curriculum of courses, laboratories, design projects, and quite often periods of industrial practice or internships. In common with other engineers, mathematics and physics form the core curriculum. The major distinction of chemical engineers is their understanding of chemistry, and increasingly of molecular and cell biology and materials science. *See* CELL BIOLOGY; MATERIALS SCIENCE; MATHEMATICS; MOLECULAR BIOLOGY; PHYSICS.

The earliest engineering education program began in France, where mathematics played a decisive role. The modern chemical engineering curriculum evolved from a program in 1888 at the Massachusetts Institute of Technology by L. M. Norton, which was centered on the study of the applied chemistry of specific inorganic chemicals. The first paradigm, or systemic knowledge, that could be applied to many problems was the concept of unit operations, introduced in 1923 in the *Principles of Chemical Engineering* by W. H. Walker, W. K. Lewis, and W. H. McAdams. The concept of unit operations became the standard taught everywhere and provided the effective tools for the graduates to engineer new chemical industries. Thermodynamics and chemical reaction engineering were added later, and the accreditation of chemical engineering degrees was based on the mastery of these subjects. *See* CHEMICAL REACTOR; CHEMICAL THERMODYNAMICS; UNIT OPERATIONS.

After World War II, during which scientists with no engineering background made the atomic bomb, engineering educators turned toward science for inspiration. Transport phenomena began as the second paradigm of chemical engineering, with the publication (1960) of *Transport Phenomena* by R. B. Bird, W. E. Stewart, and E. N. Lightfoot. The transport of momentum, heat, and mass is fundamental to the unit operations and chemical reactors. Applied mathematics and computation formed the basis for the rational modeling and design of chemical machinery and operations, as well as the systematic optimization of entire chemical plants. *See* TRANSPORT PROCESSES.

In the United States, the American Institute of Chemical Engineers, founded in 1908, is the principal professional society. Chemical engineering became a worldwide force with the founding of the Institution of Chemical Engineers in 1953 and the World Chemical Engineering Council in 2001.

James Wei

Bibliography. N. R. Amundson, *Frontiers in Chemical Engineering: Research Needs and Opportunities*, National Academies Press, 1988; R. B. Bird, W. E. Stewart, and E. N. Lightfoot, *Transport Phenomena*, Wiley, 1960; E. L. Cussler and J. Wei, Chemical product engineering, *AIChE J.*, 49(5):1072–1075, 2003; R. H. Perry and D. W. Green (eds.), *Perry's Chemical Engineers' Handbook*, 1997; T. S. Reynolds, *75 Years of Progress: A History of the American Institute of Chemical Engineers*, American Institute of Chemical Engineers, 1983; L. E. Scriven, On the emergence and evolution of chemical engineering, in C. K. Colton (ed.), *Advances in Chemical Engineering: Perspectives in Chemical Engineering: Research and Education*, Academic Press, 1991; W. H. Walker, W. K. Lewis, and W. H. McAdams, *Principles of Chemical Engineering*, McGraw-Hill, 1923.

Chemical equilibrium

In a dynamic or kinetic sense, chemical equilibrium is a condition in which a chemical reaction is occurring at equal rates in its forward and reverse directions, so that the concentrations of the reactant and product substances do not change with time. In a thermodynamic sense, it is the condition in which there is no tendency for the composition of the system to change. From the viewpoint of statistical mechanics, the equilibrium state places the system in a condition of maximum freedom (or minimum restraint) compatible with the energy, volume, and composition of the system. The statistical approach has been merged with thermodynamics into a field called statistical thermodynamics; this merger has been of immense value for its intellectual stimulus, as well as for its practical contributions to the study of equilibria. *See* CHEMICAL THERMODYNAMICS; STATISTICAL MECHANICS.

Because thermodynamics is concerned with relationships among observable properties, such as temperature, pressure, concentration, heat, and work, the relationships possess general validity, independent of theories of molecular behavior. Because of the simplicity of the concepts involved, this article utilizes the thermodynamic approach.

Chemical potential. Thermodynamics attributes to each chemical substance a property called the chemical potential, which may be thought of as the tendency of the substance to enter into chemical (or physical) change. Although the absolute chemical potential of a substance cannot be measured directly, differences in chemical potential are measurable. The units of chemical potential are joules per mole.

The importance of the chemical potential lies in its relation to the affinity or driving force of a chemical reaction. It is useful to consider the general reaction (1), and let μ_A be the chemical potential of substance A, μ_B be the chemical potential of B, and so on. Then, according to one of the fundamental principles of thermodynamics (the second law), the reaction will be spontaneous (left to right) when the total chemical potential of the reactants is greater than that of the products. Thus, for spontaneous change (naturally occurring processes), notation (2) applies.

$$aA + bB \rightleftharpoons gG + hH \qquad (1)$$

$$[g\mu_G + h\mu_H] - [a\mu_A + b\mu_B] < 0 \qquad (2)$$

When equilibrium is reached, the total chemical potentials of products and reactants become equal; thus Eq. (3) holds at equilibrium. The difference in chemical potentials in Eqs. (2) and (3) is called the driving force or affinity of the process or reaction; the affinity is zero when the chemical system is in chemical equilibrium.

$$[g\mu_G + h\mu_H] - [a\mu_A + b\mu_B] = 0 \qquad (3)$$

For reactions at constant temperature and pressure (the usual restraints in a chemical laboratory), the difference in chemical potentials becomes equal to the Gibbs energy change $\Delta_r G$ for the process in Eq. (4). The decrease in Gibbs energy represents the maximum net work obtainable from the process. When no more work is obtainable, then the system is at equilibrium. Conversely, if the value of $\Delta_r G$ for a process is positive, some useful work will have to be expended upon the process, or reaction, in order to make it proceed from left to right; the process cannot proceed naturally or spontaneously. The term "spontaneously" as used here implies only that a process can occur. It does not imply that the reaction will actually occur. Thus, the reaction between hydrogen and oxygen is a spontaneous process in the sense of the term as used here, even though a mixture of hydrogen and oxygen can remain unchanged for years unless ignited or exposed to a catalyst.

$$\Delta_r G = [g\mu_G + h\mu_H] - [a\mu_A + b\mu_B] \qquad (4)$$

Because, by definition, a catalyst does not undergo a net chemical change during the reaction, its chemical potential does not appear in Eqs. (2), (3), and (4). A catalyst, therefore, can contribute nothing to the driving force of a reaction, nor can it, in consequence, alter the position of the chemical equilibrium in a system. *See* CATALYSIS.

In addition to furnishing a criterion for the equilibrium state of a chemical system, the thermodynamic method goes much further. In many cases, it yields a relation between the change in chemical potentials (or change in Gibbs energy) and the equilibrium concentrations of the substances involved in the reaction. To do this, the chemical potential must be expressed as a function of concentration and possibly other properties of the substance.

The chemical potential μ is usually represented by Eq. (5), where R is the gas constant, T is the absolute temperature, x is the concentration of the substance, γ is the activity coefficient of the substance, and μ^0 is the chemical potential of the substance in its standard state. *See* CONCENTRATION SCALES.

$$\mu = \mu^0 + RT \ln x + RT \ln \gamma \qquad (5)$$

For substances obeying the laws of ideal solutions (or ideal gases), the last term, $RT\ln \gamma$, is zero ($\gamma = 1$), because it is a measure of the deviation from ideal behavior caused by intermolecular or inter-ionic forces. An ideal solution would then be a solution for which $RT \ln \gamma$ is zero ($\gamma = 1$) over the whole concentration range. For real solutions, the state standard (where γ is unity) is generally chosen as the pure substance [mole fraction = 1 at 1 bar (100 kilopascals) and the temperature of interest] for solids and solvents, and as the hypothetical ideal gas or the unit concentration solute in solution at 1 bar and the temperature of interest. Although the activity coefficient γ is dimensionless, its numerical values depend upon the particular concentration scale x with which it is associated.

Although the choice of concentration scales is a matter of convenience, the following are conventionally used: $x = p$ = partial pressure, for gases; $x = c$

(or m) = molar (or molal) concentrations, respectively, for solutes in electrolytic solutions; x = mole fraction, for solids and solvents. When the choice is not established by convention, the mole fraction scale is preferred.

Activity and standard states. It is often convenient to utilize the product, γx, called the activity of the substance and defined by $a = \gamma x$. The activity may be looked upon as an effective concentration of the substance, measured in the same units as the concentration x with which it is associated. The standard state of the substance is then defined as the state of unit activity (where $a = x = 1$) at 1 bar and the temperature of interest and is characterized by the standard chemical potential μ^0. Clearly, the terms μ^0, γ, and x are not independent; the choice of the activity scale serves to fix the standard state. For example, for hydrochloric acid, the standard state for the solute (HCl) would be an (hypothetical) ideal 1 molar (or 1 molal) solution, and for the solvent (H_2O) the standard state would be pure water (mole fraction = 1), both at 1 bar and the temperature of interest. The reference state would be an infinitely dilute solution; here the activity coefficients would be unity for both solute and solvent. For the vapor of HCl above the solution, the standard state would be the hypothetical gaseous state at 1 bar partial pressure. For gases, the term "fugacity" is often used instead of "activity." *See* FUGACITY.

Equilibrium constant. If the general reaction in Eq. (1) occurs at constant temperature T and pressure P when all of the substances involved are in their standard states of unit activity, Eq. (4) would become Eq. (6). The quantity $\Delta_r G^0$ is known as the standard

$$\Delta_r G^0 = \left[g\mu_G^0 + b\mu_H^0\right] - \left[a\mu_A^0 + b\mu_B^0\right] \qquad (6)$$

Gibbs energy change for the reaction at the temperature and pressure for the chosen standard states. (Standard state properties are commonly designated by a superscript zero, as in $\Delta_r G^0$, μ^0.) Since each of the standard chemical potentials (μ_i^0) is a unique property determined by the temperature, pressure, standard state, and chemical identity of the substance concerned, the standard free energy change ΔG^0 is a constant characteristic of the particular reaction for the chosen temperature, pressure, and standard states.

If, for Eq. (1), at constant temperature and pressure, the chemical potential of each substance is expressed in terms of Eq. (5), then the Gibbs energy change for the reaction, from Eq. (4), becomes Eq. (7) in terms of the activities and the standard

$$\Delta_r G = \Delta_r G^0 + RT \ln \frac{a_G^g a_H^b}{a_A^a a_B^b} \qquad (7)$$

Gibbs energy change, Eq. (6). Equation (7) is often written in the form of Eq. (8), where Q is the ratio

$$\Delta_r G = \Delta_r G^0 + RT \ln Q \qquad (8)$$

of the activities of products to the activities of reactants, each activity bearing as an exponent the corresponding coefficient in the balanced equation for the reaction [as in Eq. (7)]. The standard Gibbs energy change $\Delta_r G^0$ serves as a reference point from which the actual Gibbs energy change $\Delta_r G$ can be calculated in terms of the activities of the reacting substances.

When the system has come to chemical equilibrium at constant temperature and pressure, $\Delta_r G = 0$, from Eq. (3). Equation (7) then leads to the very important relation shown in Eq. (9), where the value of K is shown as Eq. (10), and the activities are

$$\Delta_r G^0 = -RT \ \ln K \qquad (9)$$

$$K = \left[\frac{a_G^g a_H^b}{a_A^a a_B^b}\right] \qquad (10)$$

the equilibrium values. The quantity K is called the equilibrium constant for the reaction. The equilibrium constant is a characteristic property of the reaction system, because it is determined uniquely at a given temperature in terms of the standard Gibbs energy change. The term $-\Delta_r G^0$ represents the maximum net work that the reaction could make available when carried out at constant temperature and pressure with the substances in their standard states. It should be clear from Eq. (6) that the magnitude of $\Delta_r G^0$ is directly proportional to the amount of material represented in the reaction in Eq. (1). Likewise, Eqs. (7) through (10) denote this same proportionality through the exponents a, b, g, and h in the terms K and Q. Naturally, the value of $\Delta_r G^0$, as well as K and Q, will depend upon the particular concentration scales and standard states selected for the system. Therefore it is essential that sufficient information be stated about a system to fully describe the chosen standard states.

Equations (8) and (10) can be combined in the form of Eq. (11), sometimes denoted as the Lewis

$$\Delta_r G = RT \ \ln \frac{Q}{K} \qquad (11)$$

equation after G. N. Lewis, who was the first to formulate this equation. When Q for a specified set of conditions is larger than K (so that $\Delta_r G$ is positive), the proposed reaction cannot occur left to right. On the other hand, when Q is less than K, then the proposed process or reaction can occur. The equilibrium constant thus serves as a measure of the position of chemical equilibrium for a system. For a proposed process for which Q is greater than K, the reaction system would have to move away from its equilibrium state (impossible of its own accord) to proceed from left to right, and for a proposed process for which Q is less than K, the process would bring the system closer to its equilibrium state (as in all naturally occurring processes).

Instead of activities, values of concentrations and activity coefficients at equilibrium may be used to express the form of the equilibrium constant K in Eq. (12). This gives the equilibrium constant as a

$$K = \left[\frac{x_G^g x_H^b}{x_A^a x_B^b}\right] \cdot \left[\frac{\gamma_G^g \gamma_H^b}{\gamma_A^a \gamma_B^b}\right] \qquad (12)$$

product of two terms. The first term, involving concentrations, is directly measurable if the system can be analyzed at equilibrium. On the other hand, the activity coefficient term, as seen in Eq. (13), is frequently difficult to evaluate.

$$K_\gamma = \frac{\gamma_G^g \gamma_H^b}{\gamma_A^a \gamma_B^b} \tag{13}$$

Intensive studies of activity coefficients made upon a wide variety of chemical systems led to a number of simplifying principles and some useful theoretical treatments of the subject. For gases, the activity coefficients differ only slightly from unity for pressures up to a few atmospheres and often can be evaluated from equation-of-state data. For mixtures of nonelectrolytes, the values also are not far from unity in many cases. For solutions of electrolytes, the activity coefficients vary greatly with concentration, and in many cases approach unity only below a useful or even attainable concentration. The theoretical treatments of P. Debye, E. Hückel, and others have systematized the patterns of electrolyte behavior, making possible a reasonable estimate of the activity coefficients of electrolytes in many cases. *See* ACTIVITY (THERMODYNAMICS); SOLUTION.

In general, the function K_γ, Eq. (13), approaches unity as the composition of the system approaches ideal behavior, so in practice most equilibrium constants K are evaluated through some suitable extrapolation procedure involving Eq. (12). See the discussion following Eq. (5).

For many approximate calculations or when data for K_γ are scarce, it is common to express the equilibrium constant as the concentration term only; that is, Eq. (14) is applied. Unless K_γ is a rather insen-

$$K_{\text{app}} = \frac{x_G^g x_H^b}{x_A^a x_B^b} \approx K \tag{14}$$

sitive function of concentration, the so-called constants (K_{app} denotes an apparent or approximate value of K), obtained in this manner are not constant as the composition is varied, and even though approximately constant, may vary considerably from the true value of K. Although the practice of assigning K_γ a value of unity will often give useful results and is frequently the only expedient available, the results must be used with caution.

The kinetic concept of chemical equilibrium introduced by C. M. Guldberg and P. Waage (1864) led to the formulation of the equilibrium constant in terms of concentrations. Although the concept is correct in terms of the dynamic picture of opposing reactions occurring at equal speeds, it was not successful in coping with the problems of activity coefficients and cannot lead to the useful relations, Eqs. (4) and (11), in terms of an energetic criterion for the position of equilibrium. Conversely, the thermodynamic approach yields no relationship between the driving force at the reaction and the rate of approach to equilibrium. *See* CHEMICAL DYNAMICS.

The influence of temperature upon the chemical potentials, and hence upon the equilibrium constant, is given by the Gibbs-Helmholtz equation (15).

$$\frac{d \ln K}{dT} = \frac{\Delta_r H^0}{RT^2} \tag{15}$$

The derivative on the left represents the slope of the curve obtained when values of ln K for a reaction, obtained at different temperatures, are plotted against temperature. The standard enthalpy of reaction $\Delta_r H^0$ for the temperature T at which the slope is measured, is the heat effect which could also be observed by carrying out the reaction involving the standard states in a calorimeter at the corresponding temperature and 1 bar.

For endothermic reactions, which absorb heat ($\Delta_r H^0$ positive), K increases with increasing temperature. For exothermic reactions, which evolve heat ($\Delta_r H^0$ negative), K decreases with increasing temperature and the yield of products is reduced. A more useful arrangement of Eq. (15) is shown in Eq. (16).

$$\frac{d \ln K}{d(1/T)} = \frac{-\Delta_r H^0}{R} \tag{16}$$

Plots of ln K against $1/T$ are nearly linear for reactions where the value of $\Delta_r H^0$ changes slowly with temperature. Hence, over small temperature ranges, Eq. (16) becomes, in integrated form, Eq. (17). This

$$\ln \frac{K_2}{K_1} = \frac{\Delta_r H^0}{R} \left[\frac{T_2 - T_1}{T_1 T_2} \right] \tag{17}$$

relation is much used for calculating heats of reaction from two equilibrium measurements or for determining an equilibrium constant K_2 at temperature T_2 from values of K_1, and $\Delta_r H^0$.

For accurate work, or for extending the calculations over a wide range of temperature, $\Delta_r H^0$ must be known as a function of temperature before Eq. (15) or (16) can be integrated. When sufficient heat capacity data are available, Kirchhoff's equation, involving the difference in heat capacities between the products and reactants, Eq. (18), may be combined with Eq. (15) to yield Eq. (19). In these equa-

$$\Delta_r H^0 = \Delta_r H_0^0 + \alpha T + \frac{\beta T^2}{2} + \frac{\delta T^3}{3} \tag{18}$$

$$R \ln K = -\frac{\Delta_r H_0^0}{T} + \alpha \ln T + \frac{\beta T}{2} + \frac{\delta T^2}{6} + I \tag{19}$$

tions, the coefficients α, β, and δ are determined from heat capacity data; the constants $\Delta_r H_0^0$ and I require knowledge of one value of $\Delta_r H^0$ and one value of K, or values of K at two temperatures. *See* HEAT CAPACITY; THERMOCHEMISTRY.

Homogeneous equilibria. These equilibria involve single-phase systems: gaseous, liquid, or solid solutions.

A representative gas-phase equilibrium is the ammonia synthesis shown in Eq. (20), where (g) de-

$$N_2(g) + 3H_2(g) \rightleftharpoons 2NH_3(g) \tag{20}$$

notes a gas-phase species. The most natural concentration measures are mole fraction or partial pressure. The partial pressures p_i are defined in terms of

mole fractions x_i, and the total pressure P by $p_i = x_i P$. For low pressures, where activity coefficients are practically unity, Eq. (21) holds. Equation (21) is a special case of Eq. (22), where Δn is the change in

$$K_p = \frac{p^2_{NH_3}}{p_{N_2} p^3_{H_2}} = \frac{x^2_{NH_3}}{x_{N_2} x^3_{H_2}} P^{-2} \tag{21}$$

$$K_p = K_x P^{\Delta n} \tag{22}$$

the number of moles of gases [$\Delta r = -2$ for Eq. (20)]. The mole fraction equilibrium constant K_x is pressure dependent, but K_p is independent of pressure. Consequently, an increase in total pressure P must lead to an increase in K_x in this case. If the increase in total pressure is due to a decrease in volume of the system, then the result will be an increased yield of products (NH_3). An increase in pressure brought about by injection of an inert gas into a constant volume system would not affect the partial pressures of the reacting gases nor the ultimate yield of products. The value of K_x, and thus that of $\Delta_r G_x$, will be affected by a change in total pressure owing to a change in the net work of mixing and unmixing the gases. The reaction described by Eq. (20) is exothermic, so better yields will be obtained at lower temperatures. However, the rate of attainment of equilibrium in this case is very slow at low temperatures [Eq. (15)].

A typical liquid-phase equilibrium is the deprotonation of acetic acid in water, reaction (23) and Eq. (24), where (aq) denotes an aqueous-phase

$$HC_2H_3O_2(aq) + H_2O(l) \rightleftharpoons H_3O^+(aq) + C_2H_3O_2^-(aq) \tag{23}$$

$$\frac{K}{K_\gamma} = \frac{(m_{H_3O^+})(m_{C_2H_3O_2^-})}{m_{HC_2H_3O_2} a_{H_2O}} \tag{24}$$

species and (l) denotes a liquid-phase solvent. In this particular case, the activity coefficient ratio K_γ is less than unity in dilute solutions. The activity of water $a_{H_2O} \simeq x_{H_2O}$ at low molalities (≤ 1 m) is close to unity. For this reason, a_{H_2O} is commonly omitted in the formulation of the equilibrium constant expressions of this type, because $a_{H_2O} \simeq 1$. *See* IONIC EQUILIBRIUM.

Heterogeneous equilibria. These are usually studied at constant pressure, because at least one of the phases will be a solid or a liquid. The imposed pressure may be that of an equilibrium gaseous phase, or it may be an externally controlled pressure.

In describing such systems, as for all reaction systems, the nature of each phase must be specified. In the examples to follow, s, l, g, aq, and soln identify solid, liquid, gaseous, aqueous, and nonaqueous phases, respectively. For solutions or mixtures, the composition is needed, in addition to the temperature and pressure, to complete the specification of the system; the nature of any solvent also must be specified.

In the equilibrium shown as (25), the relationship of Eq. (26) holds. Here $K \approx p/x$, the ratio of

$$H_2O(l) \rightleftharpoons H_2O(g) \tag{25}$$

$$\Delta_r G^0 = \mu^0_g - \mu^0_l = -RT \ln \frac{p}{x} \tag{26}$$

the vapor pressure p to the liquid mole fraction x. For pure water, the equilibrium constant is approximately equal to the vapor pressure of water at least at low pressures. When a small amount of solute is added, thereby decreasing the mole fraction of solvent, the vapor pressure p must be lowered to maintain equilibrium (Raoult's law). The effect of the total applied pressure P upon the vapor pressure p of the liquid is given by the Gibbs-Poynting equation (27). Here V_l and V_g are the molar volumes

$$\left[\frac{dp}{dP}\right]_T = \frac{V_l}{V_g} \tag{27}$$

of liquid and vapor, respectively. The equilibrium vapor pressure will increase as the total pressure is increased (both V_l and V_g are positive and the activity of the liquid increases with pressure). If the external pressure is applied to a solution by a semipermeable membrane, then an applied pressure can be found which will restore the vapor pressure (or activity) of the solvent to its standard state value. *See* OSMOSIS.

For a solid, such as barium sulfate, $BaSO_4$, which dissolves in water as shown in Eq. (28), the equilibrium relationship is shown in Eq. (29). When the

$$BaSO_4(s) \rightleftharpoons Ba^{2+}(aq) + SO_4^{2-}(aq) \tag{28}$$

$$K^0 = \frac{a_{Ba^{2+}} + a_{SO_4^{2-}}}{a_{BaSO_4(s)}} \tag{29}$$

solid state is pure, its activity is unity at 1 bar. If the solid is extremely finely divided, then its activity is greater than unity; with this increase in the activity of the solid state, the solubility must increase to maintain equilibrium. On the other hand, inclusion of foreign ions in the crystal lattice (solid solution formation) lowers the activity of the solid state.

When a gas, such as CO_2, is dissolved in a liquid, its equilibrium with the gas phase is as shown in Eq. (30). Equation (31), Henry's law, represents the

$$CO_2(soln) \rightleftharpoons CO_2(g) \tag{30}$$

$$K = \frac{a_{CO_2(g)}}{a_{CO_2(soln)}} \approx \frac{P_{CO_2}}{m_{CO_2}} \tag{31}$$

equilibrium constant for this equation, where m is the molality of dissolved CO_2. When the gas dissociates in the liquid, as in reaction (32), Eq. (33) may

$$H^+(aq) + Cl^-(aq) \rightleftharpoons HCl(g) \tag{32}$$

$$K = \frac{p}{m^2_{HCl}} \cdot K_\gamma \tag{33}$$

be utilized. Here m is the molality of dissolved HCl.

Similarly, when a solute distributes itself between two immiscible phases, that is, nA(soln 1) $\rightleftharpoons$ A_n(soln 2), the equilibrium constant takes the form

of Eq. (34). The equilibrium concentrations c_i in

$$K = \frac{c_{A_n}}{c_A^n} \cdot K_\gamma \tag{34}$$

Eq. (34) reflect the relative solubilities of the solute in the two phases; since solubilities may vary widely, solvent extraction may provide an effective means for concentrating a widely dispersed solute. *See* EXTRACTION.

When components form immiscible phases at equilibrium, each condensed phase will be a saturated solution; complete immiscibility is impossible in principle, because the chemical potential of any component must be the same in all phases. The separation of a liquid system into two liquid phases is a manifestation of the nonideality of the solutions. For practical purposes, however, many solids may be regarded as immiscible because of the stringent requirements associated with formation of solid solutions.

In reactions involving condensed and immiscible phases, for example, reaction (35), there can be no

$$\mathrm{Pb(s) + 2AgCl(s) \rightleftharpoons PbCl_2(s) + 2Ag(s)} \tag{35}$$

change in concentration of any phase during the reaction. In such a case the Q term in Eq. (10) will be constant, and $\Delta_r G$ is fixed at a given pressure and temperature. Although a reaction is possible, there is at most one pressure at each temperature for which $\Delta_r G$ could be zero. Such is the case in transition phenomena, or melting-freezing phenomena. From the phase-rule viewpoint, the system in reaction (35) has only one degree of freedom if reaction is possible, so once a value of P or T is chosen, the remaining variable (T or P) is fixed by nature at equilibrium. *See* PHASE RULE. Peter A. Rock; Cecil E. Vanderzee

Bibliography. G. M. Barrow, *Physical Chemistry*, 6th ed., 1996; I. M. Klotz and R. M. Rosenberg, *Chemical Thermodynamics: Basic Theory Methods*, 5th ed., 1994; I. N. Levine, *Physical Chemistry*, 4th ed., 1995; G. N. Lewis and M. Randall, *Thermodynamics*, 2d ed., rev. by K. S. Pitzer and L. Brewer, 1961; P. A. Rock, *Chemical Thermodynamics*, 1983.

Chemical fuel

The principal fuels used in internal combustion engines (automobiles, diesel, and turbojet) and in the furnaces of stationary power plants are organic fossil fuels. These fuels, and others derived from them by various refining and separation processes, are found in the earth in the solid (coal), liquid (petroleum), and gas (natural gas) phases.

Special fuels to improve the performance of combustion engines are obtained by synthetic chemical procedures. These special fuels serve to increase the fuel specific impulse of the engine (specific impulse is the force produced by the engine multiplied by the time over which it is produced, divided by the mass of the fuel) or to increase the heat of combustion available to the engine per unit mass or per unit volume of the fuel. A special fuel which possesses a very high heat of combustion per unit mass is liquid hydrogen. It has been used along with liquid oxygen in rocket engines. Because of its low liquid density, liquid hydrogen is not too useful in systems requiring high heats of combustion per unit volume of fuel ("volume-limited" systems). In combination with liquid fluorine, liquid hydrogen produces extremely large specific impulses. *See* AIRCRAFT FUEL.

A special fuel which produces high flame temperatures of the order of 9000°F (5000°C) is gaseous cyanogen, C_2N_2. This is used with gaseous oxygen as the oxidizer. The liquid fuel hydrazine, N_2H_4, and other hydrazine-based fuels, with the liquid oxidizer nitrogen tetroxide, N_2O_4, are used in many space-oriented rocket engines. The boron hydrides, such as diborane, B_2H_6, and pentaborane, B_5H_9, are high-energy fuels which are being used in advanced rocket engines. *See* BORANE.

For air-breathing propulsion engines (turbojets and ramjets), hydrocarbon fuels are most often used. For some applications, metal alkyl fuels which are pyrophoric (that is, ignite spontaneously in the presence of air), and even liquid hydrogen, are being used.

A partial list of additional liquid fuels and their associated oxidizers is shown in the **table**.

Fuels which liberate heat in the absence of an oxidizer while decomposing either spontaneously or because of the presence of a catalyst are called monopropellants and have been used in rocket engines. Examples of these monopropellants are hydrogen peroxide, H_2O_2, and nitromethane, CH_3NO_2.

Liquid fuels and oxidizers are used in most large-thrust (large propulsive force) rocket engines. When thrust is not a consideration, solid-propellant fuels and oxidizers are frequently employed because of the lack of moving parts such as valves and pumps, and the consequent simplicity of this type of rocket engine. Solid fuels fall into two broad classes, double-base and composites. Double-base fuels are compounded of nitroglycerin (glycerol trinitrate) and nitrocellulose, with no separate oxidizer required. The nitroglycerin plasticizes and swells the nitrocellulose, leading to a propellant of relatively high strength and low elongation. The double-base propellant is generally formed in a mold into the desired shape (called a grain) required for the rocket case. Composite propellants are made of a fuel and

Liquid fuels and their associated oxidizers

Fuel	Oxidizer
Ammonia	Liquid oxygen
95% Ethyl alcohol	Liquid oxygen
Methyl alcohol	87% Hydrogen peroxide
Aniline	Red fuming nitric acid
Furfural alcohol	Red fuming nitric acid

an oxidizer. The latter could be an inorganic perchlorate such as ammonium perchlorate, NH_4ClO_4, or potassium perchlorate, $KClO_4$, or a nitrate such as ammonium nitrate, NH_4NO_3, potassium nitrate, KNO_3, or sodium nitrate, $NaNO_3$. Fuels for composite propellants are generally the asphalt-oil-type, thermosetting plastics (phenol formaldehyde and phenolfurfural resins have been used) or several types of synthetic rubber and gumlike substances. Recently, metal particles such as boron, aluminum, and beryllium have been added to solid propellants to increase their heats of combustion and to eliminate certain types of combustion instability. *See* HYDROGEN PEROXIDE; ROCKET PROPULSION. Wallace Chinitz

Bibliography. S. S. Penner, *Chemical Rocket Propulsion and Combustion Research*, 1962; G. A. Sutton, *Rocket Propulsion Elements: An Introduction to the Engineering of Rockets*, 6th ed., 1992.

Chemical kinetics

A branch of physical chemistry that seeks to measure the rates of chemical reactions, describe them in terms of elementary steps, and understand them in terms of the fundamental interactions between molecules.

Reaction Kinetics

Although the ultimate state of a chemical system is specified by thermodynamics, the time required to reach that equilibrium state is highly dependent upon the reaction. For example, diamonds are thermodynamically unstable with respect to graphite, but the rate of transformation of diamonds to graphite is negligible. As a consequence, determining the rate of chemical reactions has proved to be important for practical reasons. Rate studies have also yielded fundamental information about the details of the nuclear rearrangements which constitute the chemical reaction.

Traditional chemical kinetic investigations of the reaction between species X and Y to form Z and W, reaction (1), sought a rate of the form given in Eq. (2),

$$\mathrm{X} + \mathrm{Y} \rightarrow \mathrm{Z} + \mathrm{W} \tag{1}$$

$$d[\mathrm{Z}]/dt = kf([X], [Y], [Z], [W]) \tag{2}$$

where $d[\mathrm{Z}]/dt$ is the rate of appearance of product Z, f is some function of concentrations of X, Y, Z, and W which are themselves functions of time, and k is the rate constant. Chemical reactions are incredibly diverse, and often the function f is quite complicated, even for seemingly simple reactions such as that in which hydrogen and bromine combine directly to form hydrogen bromide (HBr). This is an example of a complex reaction which proceeds through a sequence of simpler reactions, called elementary reactions. For reaction (3*d*), the sequence of elementary reactions is a chain mechanism known to involve a series of steps, reactions (3*a*)–(3*c*). This sequence of elementary reactions was formerly known

$$\mathrm{Br_2} \rightarrow 2\mathrm{Br} \tag{3a}$$

$$\mathrm{Br} + \mathrm{H_2} \rightarrow \mathrm{HBr} + \mathrm{Br} \tag{3b}$$

$$\mathrm{H} + \mathrm{Br_2} \rightarrow \mathrm{HBr} + \mathrm{Br} \tag{3c}$$

$$\mathrm{H_2} + \mathrm{Br_2} \rightarrow 2\mathrm{HBr} \tag{3d}$$

as the reaction mechanism, but in the chemical dynamical sense the word mechanism is reserved to mean the detailed motion of the nuclei during a collision.

Bimolecular processes. An elementary reaction is considered to occur exactly as written. Reaction (3*b*) is assumed to occur when a bromine atom hits a hydrogen molecule. The products of the collision are a hydrogen bromide molecule and a hydrogen atom. On the other hand, the overall reaction is a sequence of these elementary steps and on a molecular basis does not occur as reaction (3*d*) is written. With few exceptions, the rate law for an elementary reaction $\mathrm{A} + \mathrm{B} \rightarrow \mathrm{C} + \mathrm{D}$ is given by $d[\mathrm{C}]/dt = k[\mathrm{A}][\mathrm{B}]$. The order (sum of the exponents of the concentrations) is two, which is expected if the reaction is bimolecular (requires only species A to collide with species B). The rate constant k for such a reaction depends very strongly on temperature, and is usually expressed as $k = Z_{\mathrm{AB}}\rho \exp(-E_a/RT)$. Z_{AB} is the frequency of collision between A and B calculated from molecular diameters and temperature; ρ is an empirically determined steric factor which arises because only collisions with the proper orientation of reagents will be effective; and E_a, the experimentally determined activation energy, apparently reflects the need to overcome repulsive forces before the reagents can get close enough to react.

Unimolecular processes. In some instances, especially for decompositions, $\mathrm{AB} \rightarrow \mathrm{A} + \mathrm{B}$, the elementary reaction step is first-order, Eq. (4), which

$$d[\mathrm{A}]/dt = d[\mathrm{B}]/dt = k[\mathrm{AB}] \tag{4}$$

means that the reaction is unimolecular. The species AB does not spontaneously dissociate; it must first be given some critical amount of energy, usually through collisions, to form an excited species AB*. It is the species AB* which decomposes unimolecularly. Philip R. Brooks

Relaxation Methods

Considerable use has been made of perturbation techniques to measure rates and determine mechanisms of rapid chemical reactions. These methods provide measurements of chemical reaction rates by displacing equilibria. In situations where the reaction of interest occurs in a system at equilibrium, perturbation techniques called relaxation methods have been found most effective for determining reaction rate constants.

A chemical system at equilibrium is one in which the rate of a forward reaction is exactly balanced

by the rate of the corresponding back reaction. Examples are chemical reactions occurring in liquid solutions, such as the familiar equilibrium in pure water, shown in reaction (5). The molar equilibrium constant at 25°C (77°F) is given by Eq. (6), where

$$H_2O \underset{k_b}{\rightleftharpoons} H^+(aq) + OH^-(aq) \tag{5}$$

$$K_{eq} = \frac{[H^+][OH^-]}{[H_2O]} = \frac{10^{-14}}{55.5} = 1.8 \times 10^{-16} \tag{6}$$

bracketed quantities indicate molar concentrations. It arises naturally from the equality of forward and backward reaction rates, Eq. (7). Here k_f and k_b are

$$k_f[H_2O] = k_b[H^+][OH^-] \tag{7}$$

the respective rate constants that depend on temperature but not concentrations. Furthermore, the combination of Eqs. (6) and (7) gives rise to Eq. (8).

$$K_{eq} = k_f/k_b = 1.8 \times 10^{-16} \tag{8}$$

Thus a reasonable question might be what the numerical values of k_f in units of s^{-1} and k_b in units of $dm^3\ mol^{-1}\ s^{-1}$ must be to satisfy Eqs. (6) through (8) in water at room temperature. Stated another way, when a liter of 1 *M* hydrochloric acid is poured into a liter of 1 *M* sodium hydroxide (with considerable hazardous sputtering), how rapidly do the hydronium ions, $H^+(aq)$, react with hydroxide ions, $OH^-(aq)$, to produce a warm 0.5 *M* aqueous solution of sodium chloride? In the early 1950s it was asserted that such a reaction is instantaneous. Turbulent mixing techniques were (and still are) insufficiently fast (mixing time of the order of 1 ms) for this particular reaction to occur outside the mixing chamber. The relaxation techniques were conceived by M. Eigen, who accepted the implied challenge of measuring the rates of seemingly immeasurably fast reactions. *See* ULTRAFAST MOLECULAR PROCESSES.

The essence of any of the relaxation methods is the perturbation of a chemical equilibrium (by a small change in temperature, pressure, electric-field intensity, or solvent composition) in so sudden a fashion that the chemical system, in seeking to reachieve equilibrium, is forced by the comparative slowness of the chemical reactions to lag behind the perturbation (**Fig. 1**).

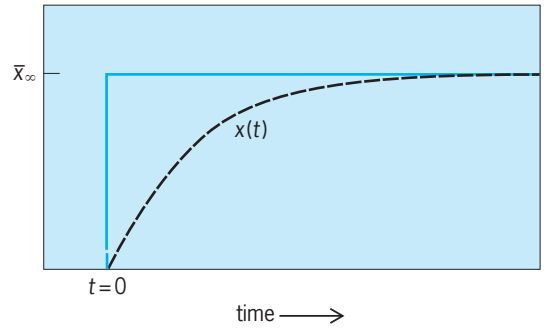

Fig. 1. Relaxational response to a rectangular step function in an external parameter such as temperature or pressure. The broken line represents the time course of the adjustment (relaxation) of the chemical equilibrium to the new temperature or pressure. (*After C. Bernasconi, Relaxation Kinetics, Academic Press, 1976*)

Temperature jump. Reaction (5) has a nonzero standard enthalpy change, ΔH°, associated with it, so that a small increase in the temperature of the water (H_2O) requires the concentrations of hydrogen ions $[H^+]$ and hydroxide ions $[OH^-]$ to increase slightly, and $[H_2O]$ to decrease correspondingly, for chemical equilibrium to be restored at the new higher temperature. Thus a small sample cell containing a very pure sample of water may be made one arm of a Wheatstone conductance bridge, and further configured so that a pulse of energy from a microwave source (or infrared laser of appropriate wavelength) is dissipated in the sample liquid. The resulting rise in temperature of about 2°C (3.6°F) will produce a small increase in conductance that will have an exponential shape and a time constant or relaxation time $\tau \simeq 27$ microseconds; τ is the time required for the signal amplitude to drop to $1/e = 1/2.718$ of its initial value, where e is the base of natural logarithms.

In pure water at 25°C (77°F), $[H^+] = [OH^-] = 10^{-7}\ M$, and for small perturbations, the value for τ is given by Eq. (9), from which it follows that $k_b \simeq 1.8 \times$

$$\begin{aligned}\tau^{-1} &= k_b([H^+] + [OH^-]) + k_f \\ &= k_b([H^+] + [OH^-] + K_{eq})\end{aligned} \tag{9}$$

$10^{11}\ dm^3\ mol^{-1}\ s^{-1}$. This is an exceptionally large rate constant for a bimolecular reaction between oppositely charged ions in aqueous solution and is, in fact, larger than that for any other diffusive encounter between ions in water. Eigen and L. DeMaeyer, who first determined this rate constant (using another relaxation method called the electric-field jump method), attributed the great speed of the back reaction of the equilibrium, reaction (5), to the exceptionally rapid motion of a proton through water, accomplished by the successive rotations of a long string of neighboring water molecules (Grotthuss mechanism). Since sample solutions can be heated by a mode-locked laser on a picosecond time scale or by a bunsen burner on a time scale of minutes, the temperature jump (T-jump) relaxation method just described is very versatile. The choice of the particular means of effecting the temperature perturbation is dictated only by the requirement that the temperature rise somewhat more rapidly than the time constant of the chemical reaction to be explored, so that a tedious deconvolution can be avoided. The discharge of a high-voltage (15–30-kV) capacitor through the sample liquid containing sufficient inert electrolyte to make it a good electrical conductor is the now classic Joule heating T-jump method used by Eigen and coworkers in their pioneering studies. A schematic of such an apparatus is shown in **Fig. 2**. The 30-kV voltage generator charges the 0.1-microfarad condenser to the voltage at which the spark gap breaks down. The condenser then discharges across the spark gap and through the sample cell, containing

an aqueous 0.1 M ionic strength solution, to ground. The sample cell is an approximately 50 ml (3.05 in.^{3}) Plexiglas cell containing two platinum electrodes spaced 1 cm (0.4 in.) apart and immersed in an aqueous 0.1 M ionic strength solution. The surge of current raises the temperature of the 1-ml (0.061-in.^{3}) volume of solution between the electrodes by 10°C (18°F) in a few microseconds.

Electric-field jump. In a situation, such as reaction (8), in which electrically neutral reactant species dissociate into oppositely charged ions, an especially sensitive tool for measuring rate constants of forward and backward reactions is the electric-field jump (E-jump) technique with conductometric detection. In a strong electric field (of the order of 4×10^6 V m^{-1}), a weak acid in solution is caused to dissociate to a greater degree than it would in the absence of the electric field. For weak electrolytes, such as aqueous acetic acid or ammonia, the effect is the order of 10% or less of the total normal dissociation, even at very high electric-field strengths. However, with a sensitive, high-voltage, Wheatstone bridge, the exponential increase with time in the concentration of ions following a precipitous increase in electric-field strength is readily detected. The measured relaxation time (τ) is clearly that corresponding to the high-electric-field environment, but since the rate constants for these reactions differ little in and out of the electric field, no serious problem is posed.

A more serious concern is that the sample solution may have a very high electrical resistance, so that the supposedly square step function in the electric-field strength is distorted by a significant voltage drop with concomitant heating of the sample liquid. Problems of working with high voltages, balancing capacitive and inductive effects in a very sensitive conductance bridge (now often circumvented by spectrophotometric detection), and the comparative difficulty of evaluating amplitudes of relaxations (as opposed to their readily determined time constants) are all factors that have worked against the wide use of the E-jump technique. There are many more ways of achieving a T-jump than an E-jump, and ΔH° values for chemical equilibria are readily available in the thermodynamic literature, whereas the extent to which a chemical equilibrium is displaced by an electric-field increment is rarely already known and is difficult to determine. Thus the commercialization of the T-jump method and the comparative neglect of the E-jump relaxation technique are readily understood.

Notwithstanding these difficulties, the E-jump technique is without peer for the investigation of the kinetics of solvent autoionization or for the exploration of the properties of weak electrolyte solutes in exotic solvents such as acetonitrile or xenon (the latter liquefied under a pressure of about 50 atm or 5 megapascals), so long as the relaxation time to be measured lies in the range 30 nanoseconds $< \tau < 100\ \mu$s.

Ultrasonic absorption. Two other relaxation methods more widely used than the E-jump technique are pressure jump (P-jump) and ultrasonic absorption. Each relies for its effectiveness on a volume change, ΔV°, occurring in an aqueous sample equilibrium undergoing kinetic investigation. (In a nonaqueous solvent it will frequently be more important that ΔH° be large than that ΔV° be so for the equilibrium to be susceptible to study by these two relaxation techniques.) As electrically neutral, weak electrolyte solute species dissociate into ions in aqueous solution, there is an increase in the number of solvent molecules drawn into a highly ordered solvation sheath. The higher the charge density of the ion, the more water molecule dipoles are bound and the greater the change in V° as reactants become products. Thus the dissociation of an aqueous neodymium(III) sulfate complex is particularly susceptible to study by one or more of the four or five ultrasonic absorption methods that cover the $f \simeq$ 100 kHz–1 GHz sound frequency range. Unlike the T-jump and E-jump relaxation methods, which usually employ step function perturbations, the ultrasonic absorption techniques are continuous-wave experiments in which the sample chemical equilibrium absorbs a measurable amount of the sound wave's energy when the frequency of the sound wave (f) and the relaxation time of the chemical equilibrium bear the relation to one another given by Eq. (10).

$$\tau^{-1} = 2\pi f \tag{10}$$

A particularly easy ultrasonic absorption experiment to understand and perform is the laser Debye-Sears technique. A continuously variable frequency sound wave is introduced by a quartz piezoelectric transducer into a 30-ml (1.83-in.^{3}) sample cell that has entrance and exit windows for a visible laser light beam that passes through the cell at about 90° to the direction of travel of the planar sound wave. The regions of compression and rarefaction in the sound wave act as a diffraction grating for the laser light beam. If a chemical equilibrium in the sample strongly absorbs a particular frequency of sound (f), the definition of the "diffraction grating" will deteriorate and the measured intensity of the first-order diffracted laser light will diminish. The frequency of minimum diffracted light intensity will be that of Eq. (10). **Figure 3** shows a diagram of the

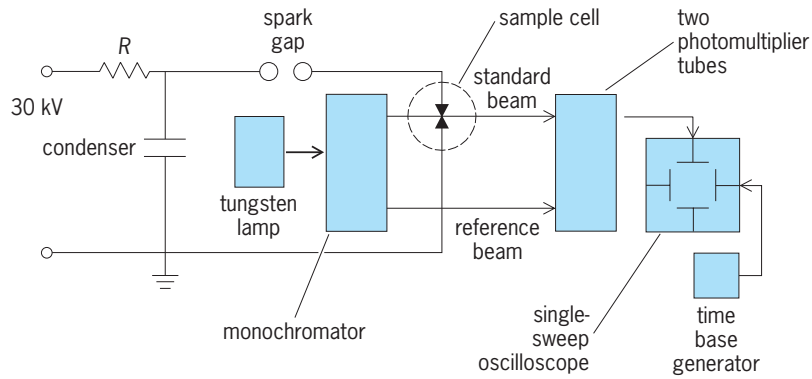

Fig. 2. Schematic of a Joule heating temperature-jump apparatus. (*After H. Eyring and E. M. Eyring, Modern Chemical Kinetics, Reinhold, 1963*)

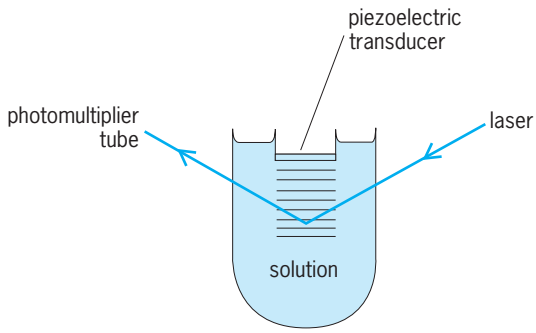

Fig. 3. Schematic of laser Debye-Sears apparatus for measuring ultrasonic absorption (about 15–300 MHz) in a sample liquid. (*After W. J. Gettins and E. Wyn-Jones, Techniques and Applications of Fast Reactions in Solution, D. Reidel, 1979*)

apparatus. The piezoelectric (quartz) tranducer cemented to the bottom of a plastic rod that is driven up and down by a stepping motor is controlled by a mini computer. The angle of diffraction of the laser beam by the alternating regions of compression and rarefaction in the liquid (suggested by the horizontal lines) is exaggerated in the diagram.

Ultrasonic absorption techniques have been used in kinetic investigations of complicated biophysical systems such as the order-disorder transitions that occur in liquid crystalline phospholipid membranes. While the ultrasonic techniques look through a conveniently broad time window at kinetic processes in solution, this picture window is difficult to "see through" in that many equilibrium processes in solution can absorb sound energy and the responsible process is not instantly identified by a characteristic absorption of electromagnetic radiation as in a spectrophotometric T-jump or E-jump experiment. A further disadvantage arises from the great breadth of the ultrasonic absorption "peaks" in a plot of normalized sound absorption versus sound frequency. Unless multiple relaxation times in a chemical system are quite widely separated in time, they are difficult to resolve in an ultrasonic absorption spectrum. *See* ULTRASONICS.

Pressure jump. The typical pressure-jump (P-jump) experiment is one in which a liquid sample under about 200 atm (20 MPa) pressure is suddenly brought to atmospheric pressure by the bursting of a metal membrane in the sample cell autoclave. Relaxation times measured spectrophotometrically or conductometrically are thus accessible if $\tau > 100\ \mu s$. This technique has proven particularly useful in the elucidation of micellar systems of great interest for catalysis and for petroleum recovery from apparently depleted oil fields.

The continuous- and stopped-flow techniques antedate somewhat the relaxation techniques described above, and have the sometimes important advantage of permitting kinetic measurements in chemical systems far from equilibrium. The stopped-flow experiment is one in which two different liquids in separate syringes are mixed rapidly in a tangential jet mixing chamber and then the rapid flow of mixed reactants is almost immediately brought to a halt in a spectrophotometric, conductometric, or calorimetric observation chamber. Reaction half-lives exceeding 2 ms are easily accessible. *See* SHOCK TUBE.

Other relaxation methods. Stopped-flow equipment has been used in concentration-jump and solvent-jump relaxation kinetic studies. An example of an application of the solvent-jump technique to a system insensitive to concentration-jump is a kinetic study of reaction (11) in mixed CCl_4-acetic acid solvents

$$\mathrm{NOCl} + n\mathrm{BuOH} \rightarrow n\mathrm{BuONO} + \mathrm{HCl} \qquad (11)$$

of varying composition (Bu = butyl). The thermodynamic treatment of the solvent jump is just about the only aspect of the presently known relaxation techniques that was not described in exhaustive detail by the earliest publications of Eigen and DeMaeyer. *See* CHEMICAL THERMODYNAMICS.

Edward M. Eyring

Gas-Phase Reactions

The rates of thermal gas-phase chemical reactions are important in understanding processes such as combustion and atmospheric chemistry.

Elementary reactions, mechanisms, and rates. Chemical conversion of one stable, gas-phase molecule into another is an apparently simple process; yet it is highly unlikely to occur in just a single step, but as a web of sequential and parallel reactions involving many species. The oxidation of methane (CH_4) to carbon dioxide (CO_2) and water provides an excellent example. It occurs in combustion (for example, in burning natural gas, which is mostly methane) as well as in the atmosphere. In both cases, the net process may be written down as single reaction (12).

$$\mathrm{CH_4} + 2\mathrm{O_2} \rightarrow 2\mathrm{H_2O} + \mathrm{CO_2} \qquad (12)$$

The reaction does not, however, result from collision of two oxygen (O_2) molecules with one methane molecule. Rather, it involves many separate steps. A simplified list of the steps involved in reaction (12) for both combustion and for the atmosphere is given below.

Combustion:

$$\begin{aligned}
\mathrm{CH_4 + OH} &\rightarrow \mathrm{CH_3 + H_2O}\\
\mathrm{CH_4 + M} &\rightarrow \mathrm{CH_3 + H + M}\\
\mathrm{CH_4 + H} &\rightarrow \mathrm{H_2 + CH_3}\\
\mathrm{CH_4 + O} &\rightarrow \mathrm{CH_3 + OH}\\
\mathrm{CH_4 + CH_2} &\rightarrow \mathrm{2CH_3}\\
\mathrm{CH_3 + M} &\rightarrow \mathrm{CH_2 + H + M}\\
\mathrm{CH_3 + H} &\rightarrow \mathrm{CH_2 + H_2}\\
\mathrm{CH_3 + O} &\rightarrow \mathrm{CH_2O + H}\\
\mathrm{CH_3 + OH} &\rightarrow \mathrm{CH_2 + H_2O}\\
\mathrm{CH_3 + O_2} &\rightarrow \mathrm{CH_2O + OH}\\
\mathrm{CH_3 + O_2} &\rightarrow \mathrm{CH_3O + O}\\
\mathrm{CH_3O + M} &\rightarrow \mathrm{CH_2O + H + M}\\
\mathrm{CH_3O + H} &\rightarrow \mathrm{CH_2O + H_2}\\
\mathrm{CH_3O + O} &\rightarrow \mathrm{CH_2O + OH}\\
\mathrm{CH_3O + OH} &\rightarrow \mathrm{CH_2O + H_2O}
\end{aligned}$$

$$CH_3O + O_2 \rightarrow CH_2O + HO_2$$
$$CH_2O + M \rightarrow HCO + H + M$$
$$CH_2O + H \rightarrow HCO + H_2$$
$$CH_2O + O \rightarrow HCO + OH$$
$$HCO + O_2 \rightarrow CO + HO_2$$
$$CO + OH \rightarrow CO_2 + H$$

Atmospheric oxidation:

$$CH_4 + OH \rightarrow CH_3 + H_2O$$
$$CH_3 + O_2 + M \rightarrow CH_3O_2 + M$$
$$CH_3O_2 + NO \rightarrow CH_3 + NO_2$$
$$CH_3O + O_2 \rightarrow CH_2O + HO_2$$
$$CH_2O + h\nu \rightarrow CHO + H$$
$$CH_2O + h\nu \rightarrow CO + H_2$$
$$CHO + O_2 \rightarrow CO + HO_2$$
$$CO + OH \rightarrow CO_2 + H$$

Even though the overall reaction is the same in both environments, the steps are quite different, as are the rates, temperature dependences, and byproducts. Each is called an elementary reaction, and the sum of the steps that makes up the overall reaction is called the mechanism. Currently, the major thrust of chemical kinetics is to elucidate such mechanisms and to measure (or calculate) the rates of the elementary reactions.

All elementary reactions fundamentally require a collision between two molecules. Even in the case of a unimolecular reaction, in which a single molecule breaks apart or isomerizes to another form, the energy required for the process comes from collision with other molecules. The species involved in many gas-phase elementary reactions are free radicals, molecules that have one or more unpaired electrons. Such species tend to be highly reactive, and they are responsible for carrying out most gas-phase chemistry.

A reaction rate is the rate at which the concentration of one of the reactants or products changes with time. The objective of a kinetics experiment is not to measure the reaction rate itself but to measure the rate coefficient, an intrinsic property of the reaction that relates the reactant concentrations to their time rates of change. For example, the mathematical expression for the rate of a biomolecular reaction, $A + B \rightarrow$ products, is differential equation (13). The

$$-d[A]/dt = -d[B]/dt = k[A][B] \qquad (13)$$

square brackets denote the concentrations of A and B, and k is the rate coefficient described above. The dependence of the rate expression on reactant concentrations is determined experimentally, and it also arises from a fundamental tenet of chemistry known as the law of mass action. Once the rate constant is known, the rate of a reaction can be computed for any given set of concentrations.

Rate constants usually change with temperature because of the change in the mean energy of colliding molecules. The temperature dependence often follows an Arrhenius expression, $k = A \exp(-E_A/RT)$, where A is a preexponential factor that is related to the gas-phase collision rate, R is the universal gas constant, and T is the absolute temperature (in kelvins). The key quantity is the activation energy, E_A, the amount of energy required to induce a reaction. Pressure dependences are usually important only for association reactions, $A + B \rightarrow AB$, since collision of A with B will form an energized complex, AB^*, that will simply redissociate unless a subsequent collision carries away enough energy to stabilize the AB product. The probability of a stabilizing collision increases with the collision frequency and thus the total pressure.

In the simple case of a unimolecular reaction, $A \rightarrow P$, the rate expression is Eq. (14). Equation (14)

$$-d[A]/dt = d[P]/dt = k[A] \qquad (14)$$

is first-order since the rate is proportional to the reactant concentration to the first power, and it leads to an expression for the change in the concentration of A or P with time (called an integrated rate expression), as in Eqs. (15). The subscripts denote

$$[A]_t = [A]_0 \exp(-kt) \qquad (15a)$$

$$[P]_t = [A]_0\{1 - \exp(-kt)\} \qquad (15b)$$

the concentrations at zero time (initial concentration) and an arbitrary time t. The concentration of A decreases (because it is reacting away) as a function of time, while the concentration of P increases with time such that the sum of the concentrations of A and P is always constant and equal to the initial concentration of A (**Fig. 4**). Because first-order reactions are mathematically simple, kineticists try to reduce all studied reactions (if at all possible) to this form. A second-order reaction, for example $A + B \rightarrow$ products, has the rate expression given in Eq. (13). To reduce the second-order expression to the first-order expression, Eq. (14), one chooses one of the concentrations to be in large excess, for example $[B] >> [A]$. The concentration of B is then approximately constant during the course of the reaction, and it may be combined with the rate constant

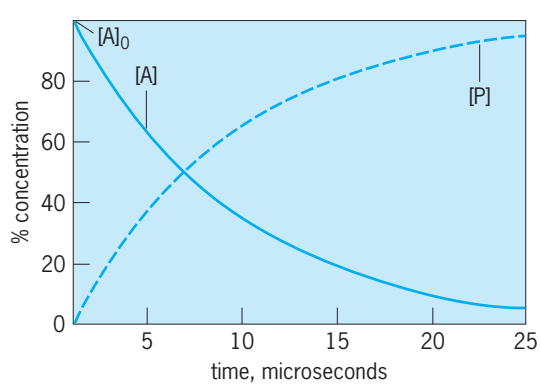

Fig. 4. Concentration of reactant (A) decreases while product (P) increases with time (arbitrarily shown in microseconds).

to give an expression identical to Eq. (14) that depends on the concentration of A alone.

Experimental methods. The experimental challenges to rate constant measurements include generation of reactive species of interest, the measurement of their concentration on a time scale fast enough to follow the course of the reaction, and the measurement of the reaction time itself. The capability for measurement of small concentrations (down to a few molecules in a milliliter) and short times (down to femtoseconds; 1 fs = 10^{-15} s) has improved dramatically since the early 1980s and has enabled kineticists to study extremely fast gas-phase reactions. The following examples illustrate how rate constants are measured.

The rate constant for the bimolecular reaction of OH with methane is very important. **Figure 5*a*** illustrates its measurement via the commonly used pulsed photolysis–laser induced fluorescence (PP–LIF) technique. A pulse of short-wavelength ultraviolet light from a laser irradiates a gas mixture containing a precursor molecule, such as hydrogen peroxide (H_2O_2) or nitric acid (HNO_3), that absorbs the light and fragments instantaneously to produce OH. A second laser produces a pulse of light that is tuned to a color (wavelength) absorbed by the OH radicals, which in turn emit light (fluoresce) in all directions. The fluorescence intensity falling on a detector is proportional to the OH concentration. Both laser pulses have very short duration (typically $<2 \times 10^{-8}$ s or 20 nanoseconds) on the time scale of the reaction. The first pulse creates the OH radicals and defines zero time, and variation of the time delay between the two lasers varies the reaction time. Methane is present in large excess over OH, and thus its concentration, as measured by its pressure, is essentially constant during the course of the reaction. *See* LASER; LASER PHOTOCHEMISTRY.

The same rate coefficient may also be measured using the apparatus shown in Fig. 5*b*. In this case, a flow of gas containing hydrogen (H_2) passes through a microwave discharge, where H_2 breaks down to H atoms. Then it reacts with nitrogen dioxide (NO_2) to make OH and nitric oxide (NO). The NO is unreactive with methane and does not interfere with the OH reaction. Methane flows through a movable injector at the center of a flow tube. The two gas flows mix at the injector and initiate reaction between OH and methane. A laser-induced fluorescence detector, similar to the one described above, at the end of the tube measures the amount of OH present at that point. (Other detectors, such as a mass spectrometer, also can be used.) Since the gas flow velocity down the tube is constant, the distance from the injector to the mass spectrometer can be converted to a reaction time, and movement of the injector to different positions in the tube varies the reaction time. As above, its pressure in the flow tube gives the approximately constant methane concentration.

Both of the above examples are direct methods in which the time rate of concentration change is directly observed. Relative rate constant measurements can be done for the determination of the rate constant for reaction of OH with hydrofluorocarbon 134a (HFC-134a). This compound is used as an automobile refrigerant in place of the banned CFCs

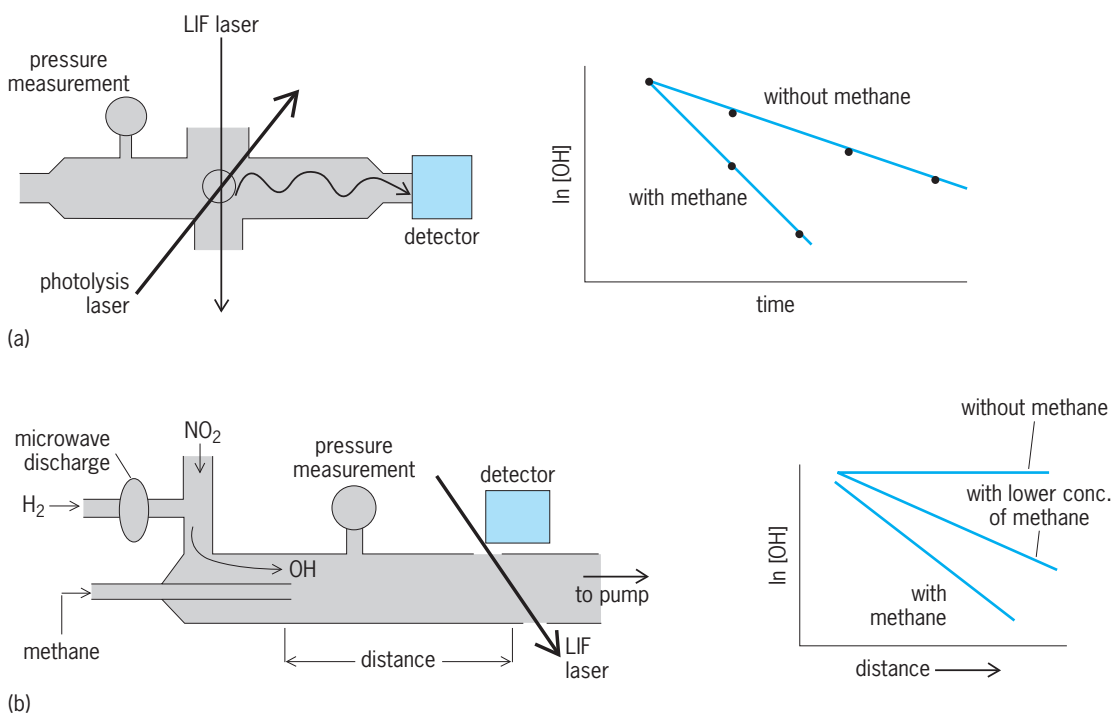

Fig. 5. Measurement of the rate constant for the bimolecular reaction of OH with methane. (*a*) Pulsed photolysis. OH concentrations are measured with methane and without methane. (*b*) Discharge flow. OH concentrations are detected at various distances from a detector. The plot shows OH concentrates measured with methane, with a lower concentration of methane, and without methane.

(chlorofluorocarbons), and, similar to methane, its reaction with OH determines its degradation rate in the atmosphere. The OH radicals are continuously produced in a mixture containing both methane and HFC-134a, and the depletion of methane and HFC-134a is measured. The ratio of the depletions is then proportional to the rate constants for reaction of OH with each compound. Since the rate constant for reaction of OH with methane is known, the measurement provides the previously unknown rate constant for reaction of OH with HFC-134a.

There are many other techniques for measurement of reactant or product concentrations, including ultraviolet, visible, and infrared light absorption, gas chromatography, and a host of additional fluorescence and mass spectrometric methods.

Outlook. Complex gas-phase chemical reactions may be broken down into elementary steps that describe reactions at the level of collisions between individual molecules. The rate constants for such elementary steps allow calculation of the reaction rates from reactant concentrations as well as kinetic modeling of larger chemical processes. Determination of elementary rate constants involves measurement of reactant or product concentrations on the time scale of the reaction. Rate constants may also be made in a relative fashion in cases where two reactions share a common reactant and one rate constant is known. A. R. Ravishankara; S. S. Brown

Bibliography. G. D. Billing and K. V. Mikkelsen, *Advanced Molecular Dynamics and Chemical Kinetics*, John Wiley, 1997; P. R. Brooks and E. F. Hayes, *State-to-State Chemistry*, ACS Symp. Ser. 56, 1977; F. A. Gianturco (ed.), *Collision Theory for Atoms and Molecules*, 1989; Y. T. Lee, Molecular beam studies of elementary chemical processes, *Angew. Chem. Int. Ed. Engl.*, 26:939–951, 1987; J. I. Steinfeld, J. S. Francisco, and W. L. Hase, *Chemical Kinetics and Dynamics*, 2d ed., Prentice Hall, 1998; P. Sykes, *A Guidebook to Mechanism in Organic Chemistry*, 6th ed., 1986.

Chemical microscopy

A scientific discipline in which microscopes are used to solve chemical problems. The unique ability to form a visual image of a specimen, to select a small volume of the specimen, and to perform a chemical or structural analysis on the material in the selected volume makes chemical microscopy indispensable to modern chemical analysis. *See* MICROSCOPE.

Microscopes are instruments which form images of objects too small to be seen with the unaided eye. Microscopes can also be combined with most analytical instruments. For example, a light microscope can be combined with a spectroscope, making it possible to determine the molecular composition of microscopic objects or structures. Similarly, an x-ray spectrometer can be combined with an electron microscope to determine the elemental composition of small objects. *See* ELECTRON MICROSCOPE; SPECTROSCOPY; X-RAY SPECTROMETRY.

Phase analyses can also be made microscopically. The boundaries of amorphous phases can usually be distinguished in the microscope, and an elemental or physical analysis can be used to identify the phase. An example of a physical analysis is the measurement of refractive index. Crystalline phases are even more amenable to microscopical analysis. For example, a polarizing microscope can be used to measure the optical properties of a crystalline phase and thus identify it. Or a transmission electron microscope can be used to select a tiny area of a crystalline phase and identify the crystal structure by means of electron diffraction. *See* ELECTRON DIFFRACTION.

While the idea of selecting a small volume for analysis is straightforward, the minimum volume that can be analyzed varies widely with the instrument used. Light microscopes can be used to identify particles as small as 1 micrometer in diameter and weighing about 1 picogram. A field ion microscope has been combined with a mass spectrometer and used to identify single atoms extracted from the surface of a specimen. *See* FIELD-EMISSION MICROSCOPY; MASS SPECTROMETRY; OPTICAL MICROSCOPE.

After a portion of a specimen has been selected microscopically, it can be analyzed in many ways. An experienced microscopist may learn to recognize various structures by studying known materials, using published atlases, or an atlas that the individual microscopist has constructed. Analyses done in this fashion can be fast and simple. If the object or structure cannot be recognized, many means of analysis are available. For example, a polarizing microscope may be employed to identify the object by using optical crystallographic methods. Other light microscopes useful for chemical analysis include phasecontrast and interference-contrast microscopes, microspectrophotometers, the confocal scanning laser microscope, and the laser Raman microscope. Physicochemical methods may be used to measure melting points, or mixed-melt phenomena and dispersion staining may also be used. *See* INTERFERENCE MICROSCOPE; LASER SPECTROSCOPY; PHASE-CONTRAST MICROSCOPE; RAMAN EFFECT.

Microscopes using other types of image-forming beams serve for chemical analysis. Scanning or transmission electron microscopes are powerful tools for chemical microscopy. Scanning electron microscopes are often fitted with x-ray spectrometers which are capable of both qualitative and quantitative analysis for most of the elements. Other electron microscopes capable of chemical analysis are the Auger electron microscope, field electron microscope, scanning tunneling microscope, and cathodoluminescence microscope. Microscopes which use ion beams, neutron beams, and x-ray beams also have analytical capabilities. *See* AUGER EFFECT; NEUTRON SPECTROMETRY; SCANNING ELECTRON MICROSCOPE; X-RAY MICROSCOPE.

A specific example of a chemical microscopical analysis involves portland cement which is formulated for use in geothermal-energy extraction wells (see **illustration**). In this micrograph, two types of crystals can be recognized by their morphology.

Micrograph of a portland cement formulation cured at 390–450°F (200–230°C) for 72 h; gypsum and xonotlite crystals are growing into a cavity.

If the microscopist does not recognize the crystals from their morphology, the crystals must be analyzed. This can be done in a number of ways. With a light microscope, it is possible to identify the crystals by scraping them from the sample and determining their optical properties through a variety of standard techniques. Another method is to examine each of the crystals by using x-ray or electron diffraction techniques. A third option is to perform an elemental analysis with an electron microscope fitted with an x-ray spectrometer. From the elemental analysis, it should be possible to identify the chemical compound that makes up the crystal. There should be no confusion resulting from the fact that the sample consists of more than one crystalline compound, because with the electron microscope a single crystal or a small portion of a crystal can be analyzed. Each crystal that appears different from the other can be analyzed quantitatively for its elemental composition. Knowledge of the combination of elemental composition and diffraction data nearly always leads to the identification of the crystals in the sample. This does require that the diffraction pattern for the crystals has been published and is accessible to the microscopist. If this information is not available, the microscopist can usually construct a list of data for the materials considered, so that in future work the individual crystals or amorphous materials can be recognized when they are encountered again.

In the example illustrated, only the commonly available microscopes, that is, the polarizing microscope, the scanning electron microscope, and the transmission electron microscope, were considered. Less commonly available microscopes, such as the Auger microscope, the laser Raman microscope, and the microspectrophotometer, can be used to identify crystalline and amorphous phases. *See* ANALYTICAL CHEMISTRY. George Cocks

Bibliography. J. Goldstein et al., *Scanning Electron Microscopy and X-ray Microanalysis*, 3d ed., 2003; T. Lucatorto et al. (eds.), *Experimental Methods in the Physical Science*, vol. 38: *Advances in Surface Analysis*, 2001; T. Mulvey and C. J. R. Sheppard (eds.), *Advances in Optical and Electron Microscopy*, vol. 13, 1994; L. Reimer and P. W. Hawkes (eds.), *Scanning Electron Microscopy: Physics of Image Formation and Microanalysis*, 2d ed., 1998; R. Wiesendanger and H. J. Guntherodt (eds.), *Scanning Tunneling Microscopy III: Theory of STM and Related Scanning Probe Methods*, 2d ed., 1997.

Chemical process industry

An industry, abbreviated CPI, in which the raw materials undergo chemical conversion during their processing into finished products, as well as (or instead of) the physical conversions common to industry in general. In the chemical process industry the products differ chemically from the raw materials as a result of undergoing one or more chemical reactions during the manufacturing process. The chemical process industries broadly include the traditional chemical industries, both organic and inorganic; the petroleum industry; the petrochemical industry, which produces the majority of plastics, synthetic fibers, and synthetic rubber from petroleum and natural-gas raw materials; and a series of allied industries in which chemical processing plays a substantial part. While the chemical process industries are primarily the realm of the chemical engineer and the chemist, they also involve a wide range of other scientific, engineering, and economic specialists.

For a discussion of the more prominent chemical process industries *see* ADHESIVE; BIOCHEMICAL ENGINEERING; BIOMEDICAL CHEMICAL ENGINEERING; CEMENT; CERAMICS; COAL CHEMICALS; COAL GASIFICATION; COAL LIQUEFACTION; DISTILLED SPIRITS; DYEING; ELECTROCHEMICAL PROCESS; EXPLOSIVE; FAT AND OIL; FERMENTATION; FERTILIZER; FOOD MANUFACTURING; FUEL GAS; GLASS; GRAPHITE; HYDROCRACKING; INSECTICIDE; LIME (INDUSTRY); MANUFACTURED FIBER; NUCLEAR CHEMICAL ENGINEERING; NUCLEAR FUELS; PAPER; PETROCHEMICAL; PETROLEUM PROCESSING AND REFINING; PETROLEUM PRODUCTS; PLASTICS PROCESSING; POLYMER; RADIOACTIVE WASTE MANAGEMENT; RUBBER; WATER SOFTENING. William F. Furter

Bibliography. D. Irwin, *Industrial/Chemical Process Design*, 2002; R. M. Murphy, *Introduction to Chemical Processes Principles, Analysis, Synthesis*, 2005; R. H. Perry and D. W. Green (eds.), *Perry's Chemical Engineers' Platinum Edition*, 1999.

Chemical reactor

A vessel in which chemical reactions take place. A combination of vessels is known as a chemical reactor network. Chemical reactors have diverse sizes, shapes, and modes and conditions of operation based on the nature of the reaction system and its behavior as a function of temperature, pressure, catalyst properties, and other factors.

Laboratory chemical reactors are used to obtain reaction characteristics. Therefore, the shape and mode of operation of a reactor on this scale differ markedly from that of the large-scale industrial reactor, which is designed for efficient production rather than for gathering information. Laboratory reactors are best designed to achieve well-defined conditions of concentrations and temperature so that a reaction model can be developed which will prove useful in the design of a large-scale reactor model.

Chemical reactions may occur in the presence of a single phase (liquid or gas), in which case they are

called homogeneous, or they may occur in the presence of more than one phase and are referred to as heterogeneous. In addition, chemical reactions may be catalyzed. Examples of homogeneous reactions are gaseous fuel combustion (gas phase) and acid-base neutralization (liquid phase). Examples of heterogeneous systems are carbon dioxide absorption into alkali (gas-liquid); coal combustion and automobile exhaust purification (gas-solid); water softening (liquid-solid); coal liquefaction and oil hydrogenation (gas-liquid-solid); and coke reduction of iron ore (solid-solid).

Modes of operation. Chemical reactors may be operated in batch, semibatch, or continuous modes.

Batch. When a reactor is operated in a batch mode, the reactants are charged, and the vessel is closed and brought to the desired temperature and pressure. These conditions are maintained for the time needed to achieve the desired conversion and selectivity, that is, the required quantity and quality of product. At the end of the reaction cycle, the entire mass is discharged and another cycle is begun. The complete cycle includes charging, heat-up, reaction (cook) time, cool-down, and discharge. Batch operation is labor-intensive and therefore is commonly used only in industries involved in limited production of fine chemicals, such as pharmaceuticals. *See* AUTOCLAVE.

Semibatch. In this type of reactor operation, one or more reactants are in the batch mode, while the coreactant is fed and withdrawn continuously. For example, a batch of liquid is exposed to a continuously fed and withdrawn gas phase, as in the hydrogenation of unsaturated oils catalyzed by a solid Raney nickel catalyst (a gas-liquid-solid system).

Continuous. In a chemical reactor designed for continuous operation, there is continuous addition to, and withdrawal of reactants and products from, the reactor system. The reaction time is the actual residence time of the molecules within the reactor. There are two extremes of residence or exposure time. If all molecules have the same reaction exposure time, the reactor is a plug- or piston-flow reactor (PFR); that is, reactant-product species move through the reactor as would a piston or plug, with no (zero) mixing in the direction of flow. If, however, the mixing is perfect (infinite) in the direction of flow, the reactor is a continuously fed stirred-tank reactor (CSTR). For this type of reactor, concentrations and temperature are uniform throughout the vessel due to vigor of agitation, and the residence times are distributed around the mean value. In contrast, in the piston-flow reactor, concentrations and temperature are distributed nonuniformly between reactor inlet and outlet, and residence time is unique.

Selection of reactor. The choice between a continuously fed stirred-tank reactor and a piston-flow reactor for continuous operation is determined by reaction characteristics. The reactor environment with respect to temperature is likewise determined, and may be isothermal (heat removal or addition matching heat generation or extraction), adiabatic (no heat removal or addition in the reaction zone), or the intermediate case (partial heat removal or addition).

An example is the sequence of reactions shown below, where k_1 and k_2 represent rate coefficients.

$$\text{Ethylene} \xrightarrow{k_1} \text{ethylene oxide} \xrightarrow{k_2} \text{carbon dioxide} + \text{water}$$

This reaction is exothermic (heat is released), and the rate coefficient k_1 increases with increase in temperature more slowly than does k_2. A decision must be made as to what reactor type, mode of operation, and environment should be employed. If large quantities of ethylene are to be treated annually, the continuous mode is economically justified. The desired product will determine the choice between a continuously stirred tank reactor and a piston-flow reactor.

The adiabatic environment is readily achieved in large-scale industrial reactors by the denial of heat exchange. Such an adiabatic reactor consists of a large-diameter tube packed with catalyst. The adiabatic temperature rise (in exothermic reactions) may require limitation due to safety and equilibrium factors in addition to integrity of the catalyst and reactor materials. These considerations led to the development of the staged adiabatic fixed-bed reactor system for sulfur dioxide oxidation and hydrocracking. In each case, interstage cooling is effected either by means of indirect heat exchange to generate steam or preheat reactants or by means of direct cooling by the addition of "cold" coreactant (cold-shot addition). The adiabatic reactor is the least expensive, since it does not require a costly heat-transfer surface area.

Gas-solid reactors. There are a number of different types of reactors designed for gas-solid heterogeneous reactions. These include fixed beds, tubular catalytic wall reactors, and fluid beds.

The design of these components involves consideration of the properties of solid catalysts. Catalytic agents (for example, platinum, palladium, vanadium pentoxide among other zeolites, metal oxides, sulfides) are usually supported upon porous solids, such as aluminum oxide and silicon dioxide. Thus the components are, typically, support, catalyst, and promoter. Characterization of these components is imperative, since deactivation over time periods of minutes (catalytic cracking), months (reforming), and years (ammonia synthesis) plagues all catalysts. The actual catalytic sites exposed to reactants may be partially inaccessible because of the diffusional limitations of the reactants and the nature of the sites dictated by the fluid-phase environment (oxidizing, reducing). Furthermore, the reactivity and yield selectivity per exposed catalyst site may change with catalyst crystallite size (dispersion), a phenomenon known as fractal behavior. Since the reactions and associated enthalpy changes are forcing functions that affect reactor performance, consideration of the complexities of heterogeneous catalytic phenomena is necessary both in reaction modeling and in reactor modeling. *See* HETEROGENEOUS CATALYSIS.

Fixed bed. A common catalytic reactor that consists of a tube or parallel array of tubes filled with solid catalyst particles is the fixed-bed reactor. It is commonly employed in processes such as partial oxidation, ammonia synthesis, and reforming.

Minimization of tube surface area (large tube diameter) invites adiabaticity. In the general case (nonisothermal or nonadiabatic), temperature varies in axial and radial directions and radial heat-transfer limitations can cause "hot" or "quench" spots with consequent threats to safety and product quality. Radial temperature gradients can be minimized by use of small tube diameter; however, there will be consequent increase in heat-transfer surface and reactor cost. Usually a piston-flow reactor is used. When catalyst decay is rapid (as in catalytic cracking), fresh catalyst is added and spent solid removed continuously. This arrangement is known as a moving bed. *See* CRACKING.

Tubular catalytic wall reactor. Resistance to radial heat transfer can be vastly reduced by using this type of reactor, in which all the catalyst is impregnated upon the alumina wash-coated wall. This is the monolith tubular reactor used in auto-exhaust fume abatement. Catalyst replacement is obviously more difficult in this instance than in the conventional fixed bed; the tubular catalytic wall reactor cannot host catalysts of short life.

Fluid bed. When a fluid is passed upward through an unretained bed of fine solid particles, a velocity can be induced so that the bed moves, with bubbles of gas passing through the agitated solids. The bed behaves as a boiling fluid. Intensity of gas dispersion in the usually baffled bed induces rapid motion and circulation of the solid. Reaction heat is thus nearly uniformly dispersed and, unlike the fixed bed, the fluid bed approaches isothermality. Fine particles can be used with increase in surface area and reaction efficiency. Since the fluid bed is a multiphase system (gas-and-solids emulsion phases), bypassing or short-circuiting of unreacted species can occur, which reduces reactor efficiency. Staging or horizontal baffling can minimize bypassing and deleterious bubble growth.

The fluid bed can be operated in batch or continuous fashion with respect to the solids phase, while the fluiding (gas or liquid) phase is continuous. If catalyst life is short (seconds or minutes) or noncatalytic solid coreactant is converted, the continuous fluid bed must be employed. Residence time is complex.

A variation of the fluidized bed is the spouted-bed reactor, in which a jet of gas is injected at one or more points to disperse even rather large-sized particles with which the gas interacts (for example, drying of wheat, gas-solid reaction, or solid-catalyzed reaction). Solids may be batch or continuous.

At sufficiently high fluid velocities, the fluidized bed becomes entrained; that is, solids are carried out of the bed when fluid velocity exceeds particle terminal velocity. In a transport line or riser reactor, fresh solids are continuously fed at the reactor bottom with the entraining fluid. This is an ideal unit for rapidly decaying catalyst or solid noncatalytic coreactants. Modern catalytic cracking of oils to produce gasoline is largely done in the transport line reactor. Residence time approaches the piston-flow reactor limit as opposed to more complex residence time behavior of the continuous fluidized bed. Reactant exposure times are of the order of seconds, during which deactivating coke is laid down upon the catalyst. Coked catalyst is separated from the product vapors at the top of the reactor and is then sent to a regenerator, where oxygen-bearing gas is contacted with the coked catalyst to burn off the coke. The product of this gas-solid reaction, the regenerated cracking catalyst, is returned to the inlet of the transport line cracker. *See* FLUIDIZATION; FLUIDIZED-BED COMBUSTION.

Gas-liquid-solid reactors. Many different types have been developed for specific reaction conditions. The trickle-bed reactor employs a fixed bed of solid catalyst over which a liquid phase trickles downward in the presence of a cocurrent gas phase. This three-phase reactor, commonly employed for the removal of sulfur or nitrogen from high-boiling oil feedstocks, can pose problems of catalyst wetting, gas-liquid-solid transport, and liquid-phase bypassing. Reasonable catalyst life is required in trickle-bed operation.

An alternative is the slurry reactor, a vessel within which coreactant gas is dispersed into a liquid phase bearing suspended catalyst or coreactant solid particles. Unlike the trickle bed, liquid-solid wetting is complete. With short catalyst life or solid consumption, continuous feed and withdrawal of the liquid-solid phase is feasible, but with attendant problems of solid-liquid separation. Slurry reactors are usually of the continuously stirred tank type.

At high ratios of reactor length to diameter, the gas-liquid-solid reactor is often termed an ebullating-bed (high solids concentration) or bubble column reactor (low solids concentration). A trickle bed operating at high liquid holdup and countercurrent (upward) flow of gas becomes an ebullating bed.

Gas-liquid reactors assume a form virtually identical to the absorbers utilized in physical absorption processes. *See* ADSORPTION OPERATIONS.

Solid-solid reactions are often conducted in rotary kilns which provide the necessary intimacy of contact between the solid coreactants. *See* KILN.

James J. Carberry

Bibliography. L. A. Belfiore, *Transport Phenomena for Chemical Reactor Design*, 2003; O. Levenspiel, *Chemical Reaction Engineering*, 3d ed., 1998; B. Nauman, *Handbook of Chemical Reactor Design, Optimization and Scaleup*, 2001.

Chemical senses

In vertebrates, the senses of smell (olfaction) and taste (gustation) plus the so-called common chemical sense constitute the external chemical senses, as contrasted with such internal chemoreceptors as the carotid sinus detectors of carbon dioxide in blood, glucose receptors of brain and certain visceral organs, and so on. The olfactory cells of vertebrates, usually located in the olfactory mucosa of the upper nasal passages, are specialized neural elements that are responsive to chemicals in the vapor phase. The

floral fragrance of roses and the putrid smell of hydrogen sulfide are two examples of human olfactory sensations. Taste buds of the oral cavity, especially the tongue, are composed of modified epithelial cells responsive to chemicals in solution. Sweet, salty, sour, and bitter are the basic taste qualities experienced by humans. In certain fish (for example, catfish), taste buds occur over the body surface and barbels. The common chemical senses are composed of free nerve endings in the mucous membrane of the eye, nose, mouth, and digestive tract and are responsive to irritants or other chemicals in either the vapor or liquid phase. *See* CAROTID BODY; CHEMORECEPTION.

Chemical sensitivity occurs in single-cell and other lower organisms. Among invertebrates, sense organs occur as specialized hairs and sensilla, or minute cones supplied with sensory nerves and nerve cells. Characteristic of male moths, for example, are their distinctive bushy antennae, by which they detect and locate females by sex pheromones. Rodents, ungulates, carnivores, and other mammals also show sexual attraction to female odors produced by specialized glands. Whether humans in general are susceptible to pheromonal influences from other humans is highly debatable. *See* CHEMICAL ECOLOGY; PHEROMONE.

Taste plays an important role in selection and acceptance of food. Although the "sweet tooth" is widespread throughout the animal kingdom as well as humans, there are great species differences in sweet sensitivity, especially for artificial sweetners. Besides the protective, inborn aversion to bitter (many poisons, but not all, are bitter), a single experience with the particular taste of a toxic substance which caused illness may establish a strong and persistent learned taste aversion. By contrast, a compensatory salt hunger may occur in persons or animals suffering salt deficiency. The intake of salt in civilized people, however, may be determined not only by physiological need but by social custom and habit, and excess intake with high incidences of hypertension has been reported in some groups. *See* HYPERTENSION; SALT (FOOD).

The limbic system of the brain, which modulates appetitive and emotional behavior and hedonic (pleasant vs. unpleasant) experiences, has both taste and olfactory neural pathways to it, providing the neural substrate for the pleasure or displeasure of sensations. *See* NEUROBIOLOGY; OLFACTION; SENSATION; TASTE.

Carl Pfaffmann

Bibliography. C. Pfaffmann, The pleasures of sensation, *Psychol. Rev.*, 67:253–268, 1960; G. M. Shepherd, *Neurobiology*, 3d ed., 1997.

Chemical separation (chemical engineering)

A process used to purify substances or to isolate them from other substances. In industrial applications, the ultimate goal is the isolation of a product of given purity. Most industrial chemical processes involve a separation stage. This may involve a preliminary purification stage before the reaction process, or a final stage involving the separation of the desired product. Two well-known processes are distillation and crystallization.

Distillation and crystallization. The distillation process is based on the relative volatilities (that is, vapor pressures) of the components. It involves the heating of the feed material (which is to be separated) in a boiler. The heated vapors enter a column and partially condense on the many plates or trays. These plates have a large surface area to facilitate the equilibrium between the vapor and the liquid. The vapor in the column is richer in the more volatile (lighter) component. As the vapor travels up the column, this process of enrichment is enhanced. The final stage of the distillation process is the complete condensation of this vapor as it leaves the column. This is done by cooling the vapor to a temperature well below its boiling point. The result is a separation of a light component or components (the condensate) from heavy components left in the boiler. The overall separation achieved depends not only on the relative vapor pressures of the components but also on the number of trays in the column. In spite of the high energy demands of the distillation process, 90% of all separation processes in the chemical and petrochemical industries worldwide are distillation processes. *See* DISTILLATION; VAPOR PRESSURE.

In the crystallization process, separation is based on the relative solubilities of the components. The separation of sea salt from water is usually done by crystallization through the evaporation of the water either by heating or by using solar energy. *See* CRYSTALLIZATION.

Thermodynamics offers a proper understanding of all separation processes. Not only is it useful in quantifying the separation of components, but thermodynamic relationships are also vital in the design of industrial separation plants.

Basic theory. When an equilibrium exists between phases (solids, liquids, or gases), the chemical potential, μ_i, for species i in each phase may be expressed by Eq. (1). For a vapor-liquid equilibrium process

$$\mu_i(\text{solid}) = \mu_i(\text{liquid}) = \mu_i(\text{gas}) \qquad (1)$$

(distillation) this equality results in Eq. (2), the relationship between the vapor composition (y_i) and the liquid composition (x_i) at a temperature T. Here,

$$y_i/x_i = \gamma_i P_i^0/P_{(\text{total})} \qquad (2)$$

γ_i is the activity coefficient of species i, P_i^0 is the vapor pressure of pure i at T, and $P_{(\text{total})}$ is the total vapor pressure of the system. It is this ratio that defines the separating efficiency and is known as the distribution coefficient. Other relationships can be derived for other phase separation processes (such as crystallization), and in every case the results will be based on Eq. (1). *See* CHEMICAL EQUILIBRIUM.

Separation processes. Industrial separation processes involve all the possible phase pairs apart from solid-solid, which is kinetically too slow for useful applications (**Table 1**). *See* ADSORPTION OPERATIONS.

TABLE 1. Some important separation processes

Phase pair	Process	Industrial separation
Gas-liquid	Distillation	(i) Gasoline, paraffin and diesel fuel separated from oil (ii) Acetic acid from water
	Gas adsorption	Ammonia gas and air separated by passing the gas mixture through water
	Gas desorption	Ammonia and water mixture separated by passing air through the solution
	Pervaporation	Alcohols removed from esters and aroma compounds in aroma production
Gas-gas	Membrane	Methane gas or hydrogen gas separated from carbon dioxide
Gas-solid	Sublimation	Removal of water from heat-sensitive foodstuff by freeze drying; the food is cooled below 0°C and vacuum is applied, resulting in the sublimation of water
	Desorption	Drying of clay or wood by exposure to air
	Adsorption	Drying of wet air by passing over silica gel
Liquid-liquid	Solvent extraction	A mixture of acetone and water can be separated by adding carbon tetrachloride (CCl_4); phase separation takes place and the acetone is found in the CCl_4-rich layer
	Reverse osmosis	Seawater and brackish water purified by removal of salts
Liquid-solid	Supercritical fluid extraction	Separation of caffeine from coffee; the caffeine in the ground coffee beans readily dissolves in supercritical carbon dioxide
	Zone refining	Purification of benzene or metals to produce ultrapure compounds
	Fractional crystallization	Salts from seawater are usually separated on a basis of solubility

Supercritical fluid extraction. The process of supercritical fluid extraction is a relatively new separation technique. Interest in it is based on the high solubility of some solids in the supercritical solvent, the high degree of selectivity, the ease of separating the required product, and the ease with which the solvent can be recovered. To understand the process requires an understanding of the critical point. *See* SUPERCRITICAL FLUIDS.

The liquid-vapor phase line (as opposed to the solid-vapor and solid-liquid phase lines) for pressure and temperature discontinues at a certain point (**Fig. 1**), known as the critical point. At temperatures above the critical temperature (T_c), it is not possible to liquefy the gas by pressure alone. In other words, it is impossible to produce a liquid meniscus from the gas phase by pressure alone. At conditions above the critical temperature and pressure (P_c), the gas is known as a supercritical fluid (Fig. 1).

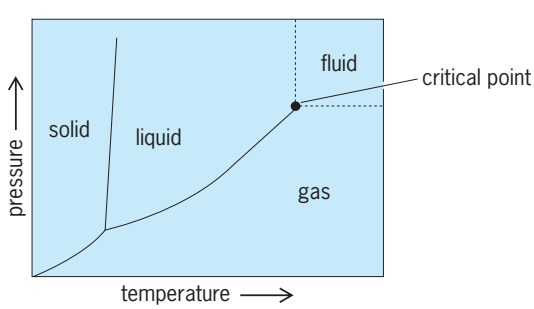

Fig. 1. Typical pressure-temperature (*P-T*) phase diagram showing the critical point discontinuity.

For an understanding of the general process, it is best to look at the most important supercritical fluid extraction process, the decaffeination of coffee using carbon dioxide. Carbon dioxide fluid, at a pressure of 200 atm (20 megapascals) and a temperature of 50°C (122°F) [in excess of P_c (CO_2) = 73 atm and T_c (CO_2) = 31°C (88°F)], is pumped through the extraction cell containing the ground coffee in a mesh basket. The supercritical carbon dioxide has the unique capacity for dissolving only the caffeine and not the compounds that give coffee its taste and flavor. The carbon dioxide fluid, now containing dissolved caffeine, leaves the cell, and its pressure is reduced to below its critical value. As a result, the fluid vaporizes to a gas and the solid caffeine precipitates as it is not soluble in gaseous carbon dioxide. The coffee without caffeine is recovered and the carbon dioxide compressed and recycled. Apart from the simplicity of the process and the fact that it is done at a low temperature (it apparently does not affect the coffee taste or flavor), its main advantage is that the solvent is a natural substance and if any residue is left in the coffee it will be harmless. This is not true for some other decaffeination processes which use organic solvents such as methylene chloride.

TABLE 2. Examples of separations based on membrane processes

Membrane	Mixture	Preferentially selective for	Process
	Gas separation		
Polyimide	O_2/N_2	O_2	Oxygen enrichment
Polyethersulfone	CH_4/CO_2	CO_2	CO_2 recovery
Cellulose acetate	H_2/CO_2 or H_2/N_2	H_2	Hydrogen recovery
	Reverse osmosis		
Cellulose acetate, aromatic polyamides	Seawater	Water	Desalination
	Sugar solutions	Water	Concentration of food ingredients
	Fruit juices	Water	Concentration of food ingredients
	Pervaporation		
Poly(vinyl alcohol)	Water+alcohol	Water	Dehydration
Polyoctylmethylsiloxanes	Water+alcohol	Alcohol	Dealcoholization
Polyoctylmethylsiloxanes	Aroma compounds in water	Aroma compounds	Aroma production

Carbon dioxide is the most popular solvent used in supercritical fluid extraction. This is largely due to its relatively low critical pressure and temperature, and the ease with which the supercritical state can be achieved. It is extensively used in the perfume industry, where high temperatures would destroy the chemicals. It is used, for example, in separating oxygenated compounds (such as geranyl acetate), valuable as perfumes, from lemon oil, which consists largely of terpenes of low value, such as limonene.

Water has received much attention as a possible solvent for supercritical fluid extraction. Its critical properties are, however, too high for convenience [$P_c = 218$ atm (22.1 MPa) and $T_c = 374°C$ (705°F)], but it does have the interesting property of dissolving nonpolar liquids such as hexane. This is not possible at ambient temperature and pressure.

Membrane separation processes. An important new technology uses membranes to effect separation. The membrane is usually a thin plastic film (0.1–5 micrometers) on a thick porous and inert support layer (100–500 μm) [**Fig. 2**]. The driving force pushing the permeate across the membrane is invariably a pressure difference between the feed mixture and the permeate. The membranes are made of synthetic polymers such as poly(vinyl alcohol), for separating water and alcohol liquids; polyurethane, for separating aromatic liquids from cycloalkane liquids; polyimides, for separating gases such as oxygen and nitrogen; and cellulose acetate, for separating pure water from seawater. *See* MEMBRANE SEPARATIONS.

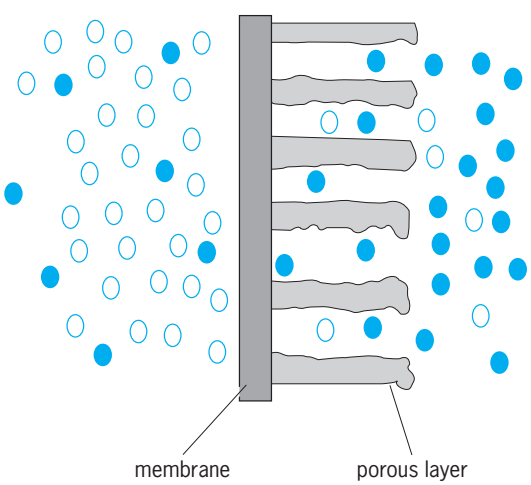

Fig. 2. General description of the membrane separation process. Molecules A (white) and B (black) are separated by a membrane supported on a porous layer and driven by a flux, which is usually pressure.

When the feed and the permeate are both gases, the process is known as gas separation. The pressure difference is usually between 10 and 100 atm (1 and 10 MPa), with the pressure of the permeate being about 1 atm (0.1 MPa). When the feed and permeate are liquids, the process is reverse osmosis, and it is best known for desalination of dilute aqueous solutions such as seawater, brackish water, and industrial wastewater. The pressure applied on the feed side is usually 80–100 atm (8–10 MPa), while the permeate is kept at ambient pressure. If the feed is a liquid under its own saturation pressure and the permeate is a gas or vapor kept at a low pressure ($5–20 \times 10^{-3}$ atm or 0.5–2 kPa) by a suitable vacuum pump, the process is known as pervaporation. The permeate is recovered in a cooling trap. *See* WASTEWATER REUSE; WATER DESALINATION.

Perhaps the main advantage of the membrane technique is that it is possible to tailor-make a membrane for a particular separation. Its limitations include chemical breakdown of the membrane with time and its slowness. The mechanism of the process is complex and involves properties such as the relative solubilities and relative diffusion coefficients of the feed chemicals in the membrane (**Table 2**).

Trevor M. Letcher

Bibliography. C. J. Geankoplis, *Transport Processes and Separation Process Principles: Includes Unit Operations*, 4th ed., 2003; R. D. Noble et al. (eds.), *Principles of Chemical Separations with Environmental Applications*, 2004; R. H. Perry and D. W. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; J. D. Seader and E. J. Henley, *Separation Process Principles*, 1999.

Chemical separation techniques

A method used in chemistry to purify substances or to isolate them from other substances, for either preparative or analytical purposes. In industrial applications the ultimate goal is the isolation of a product of given purity, whereas in analysis the primary goal is the determination of the amount or concentration of that substance in a sample. In principle it is always more convenient to carry out quantitative determinations directly on portions of the original sample. In cases where the analytical methods available are not sufficiently selective to permit this direct approach, it is necessary to employ preliminary separations to reduce the concentration of, or to remove completely, those substances which interfere in the final estimation.

Although special considerations arise in a comparison of separation methods for engineering or laboratory analytical purposes because of differences in the scales of operations, the various separation processes are based on the same principles. There are three factors of importance to be considered in all separations: (1) the completeness of recovery of the substance being isolated, (2) the extent of separation from associated substances, and (3) the efficiency of the separation. The recovery factor or yield R_A of a separation of substance A is defined as Eq. (1) where

$$R_A = \frac{Q_A}{(Q_A)_0} \tag{1}$$

Q_A and $(Q_A)_0$ are the amount of A after and before the separation.

The degree of separation $S_{B/A}$ of two substances A and B is given by the separation factor R_B for B with respect to A, and is defined as Eq. (2). Although

$$S_{B/A} = \frac{(Q_A)_0 Q_B}{(Q_B)_0 Q_A} \cong \frac{Q_B}{(Q_B)_0} \cong R_B \tag{2}$$

complete separations are usually preferred, they are not always necessary in analytical applications. The degree of purity will depend upon the choice of the method of final estimation. Sometimes merely a reduction in the quantity of foreign substance present is enough to simplify the subsequent analytical task.

The third factor, efficiency, is a measure of the amount of work required to obtain a given amount of product with a prescribed purity. This consideration is of much greater consequence in industrial separations in which both the scale and the cost of the operation are important.

There are many types of separations based on a variety of properties of materials. Among the most commonly used properties are those involving solubility, volatility, adsorption, and electrical and magnetic effects, although others have been used to advantage. The most efficient separation will obviously be obtained under conditions for which the differences in properties between two substances undergoing separation are at a maximum.

The common aspect of all separation methods is the need for two phases. The desired substance will partition or distribute between the two phases in a definite manner, and the separation is completed by physically separating the two phases. The ratio of the concentrations of a substance in the two phases is called its partition or distribution coefficient.

In analytical work the original phase is usually a liquid, that is, a solution of the sample, and the separation is brought about through the addition or formation of a solid, a liquid, or a gaseous second phase. Although the actual separation of the phases may be physical in nature, chemical reactions are usually required to convert or modify the substance to a form which permits the formation of the new phase or the partition of the substance to the second phase. In some separation methods this step may also be accomplished by physical means.

If two substances have very similar distribution coefficients, many successive steps may be required for a separation. The resulting process is called a fractionation.

Based on the nature of the second phase, the more commonly used methods of separation are classified as follows: (1) Methods involving a solid second phase include precipitation, electrodeposition, chromatography (adsorption), ion exchange, and crystallization. These methods involve a solid second phase either through the formation of a slightly soluble product, deposition as a metal on the surface of an electrode, or by physical or chemical adsorption on a suitable solid material. (2) The outstanding method involving a liquid second phase is solvent extraction, in which the original solution is placed in contact with another liquid phase immiscible with the first. Separations are achieved as a result of differences in the distribution of solutes between the two phases. Solid materials may also be separated by extraction with organic solvents. (3) Methods involving a gaseous second phase include gas evolution, distillation, sublimation, and gas chromatography. Mixtures of volatile substances can often be separated by fractional distillation. *See* CHROMATOGRAPHY; EXTRACTION.

George H. Morrison

Bibliography. J. C. King, *Separation Processes*, 2d ed., 1980; J. C. King and J. D. Navratil (eds.), *Chemical Separations*, vols. 1 and 2, 1986; E. S. Perry (ed.), *Techniques of Chemistry Separation and Purification*, 1978; R. H. Perry and D. W. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997.

Chemical symbols and formulas

A system of symbols and notation for the chemical elements and the combinations of these elements which form numerous chemical compounds. This system of symbols, devised since about 1800 by chemists, consists of letters, numerals, and marks that are designed to denote the chemical element, formula, or structure of the molecule or compound. These symbols give a concise and instantly recognizable description of the element or compound. In many cases, through the efforts of international

conferences, the symbols are recognized throughout the scientific world, and they greatly simplify the universal languge of chemistry.

Elements. Of the 115 elements known as of mid-2000, 109 have been given symbols, usually derived from the name of the element. Examples of names and symbols are chlorine, Cl; fluorine, F; beryllium, Be; aluminum, Al; oxygen, O; and carbon, C. However, symbols for some elements are derived from Latin or other names for the element. Examples are Au, gold (from *aurum*); Fe, iron (from *ferrum*); Pb, lead (from *plumbum*); Na, sodium (from *natrium*); and K, potassium (from *kalium*). The symbols consist of one or, more commonly, two letters. The first letter is always a capital letter and is followed by a lowercase second letter.

Inorganic molecules and compounds. Simple diatomic molecules of a single element are designated by the symbol for the element with a subscript 2, indicating that it contains two atoms. Thus the hydrogen molecule is H_2; the nitrogen molecule, N_2; and the oxygen molecule, O_2. Polyatomic molecules of a single element are designated by the symbol for the element with a numerical subscript corresponding to the number of atoms in the molecule. Examples are the phosphorus molecule, P_4; the sulfur molecule, S_8; and the arsenic molecule, As_4.

Diatomic covalent molecules containing unlike elements are given a similar designation. The formula for hydrogen chloride is HCl; for iodine monochloride, ICl; and for hydrogen iodide, HI. The more electropositive element is always designated first in the formula.

For polyatomic covalent molecules containing unlike elements, numerical subscripts are used to designate the number of atoms of each element that are present in the molecule. Examples are phosphorus(III) chloride, PCl_3; arsine, AsH_3; ammonia, NH_3; and water, H_2O. Again, as in diatomic molecules, the more electropositive element is placed first in the formula.

Ionic inorganic compounds are designated by a similar notation. The positive ion is given first in the formula, followed by the negative ion; subscripts are again used to denote the number of ions of each element present in the compound. The formulas for several common compounds are sodium chloride, NaCl; ammonium nitrate, NH_4NO_3; aluminum sulfate, $Al_2(SO_4)_3$; and iron(III) oxide, Fe_2O_3.

More complex inorganic compounds are designated in a manner similar to that above. The positive ion is given first, but may contain attached or coordinated groups, and this is followed by the negative ion. Examples are hexammine-cobalt(III) chloride, $[Co(NH_3)_6]Cl_3$; dichlorobis(ethylenediamine)chromium(III) nitrate, $[Cr(en)_2Cl_2](NO_3)_3$, where en is the abbreviation for ethylenediamine; and potassium trioxalatoferrate(III), $K_3[Fe(C_2O_4)_3]$. Hydrates of inorganic compounds, such as copper(II) sulfate pentahydrate, are designated by the formula of the compound followed by the formula for water, the number of water molecules being designated by a numerical prefix. Thus the symbol for the above compound is $CuSO_4 \cdot 5H_2O$.

Organic compounds. Because there are many more organic than inorganic compounds, the designation or notation for the first group becomes exceedingly complex. Many different types of organic compounds are known; in the case of hydrocarbons, there are aromatic and aliphatic, saturated and unsaturated, cyclic and polycyclic, and so on. The system of notation used must distinguish between the various hydrocarbons themselves as well as setting this group of compounds apart from others such as alcohols, ethers, amines, esters, and phenols. *See* CHEMISTRY; COORDINATION COMPLEXES; INORGANIC CHEMISTRY; ORGANIC CHEMISTRY. Wesley Wendlandt

Bibliography. B. P. Block, W. H. Powell, and W. C. Fernelius (eds.), *Inorganic Chemical Nomenclature: Principles and Practice*, 1990; J. C. Richer, R. Panico, and W. H. Powell (eds.), *A Guide to IUPAC Nomenclature of Organic Compounds*, 1993.

Chemical thermodynamics

The application of thermodynamic principles to systems involving physical and chemical transformations is carried out in order to (1) develop quantitative relationships among the identifiable forms of energy and their conjugate variables, (2) establish the criteria for spontaneous change, for equilibrium, and for thermodynamic stability, and (3) provide the macroscopic base for the statistical-mechanical bridge to atomic and molecular properties. The thermodynamic principles applied are the conservation of energy as embodied in the first law of thermodynamics, the principle of entropy production as embodied in the second law, and the principle of absolute entropy and its statistical thermodynamic formulation as embodied in the third law of thermodynamics.

Basic concepts. The basic goal of thermodynamics is to provide a description of a system of interest in order to investigate the nature and extent of changes in the state of that system as it undergoes spontaneous change toward equilibrium and interacts with its surroundings. This goal implicitly carries with it the concept that there are measurable properties of the system which can be used to describe adequately the state of the system, and that the system is enclosed by a boundary or wall which separates the system and its surroundings. Properties that define the state of the system can be classified as extensive and intensive. Extensive properties are proportional to the total mass of the system, whereas intensive properties are not. Typical extensive properties are the energy, volume, and numbers of moles of each component in the system, while typical intensive properties are temperature, pressure, density, and the mole fractions or concentrations of the components.

Extensive properties can be expressed as functions of other extensive properties, for instance, as

in Eq. (1), where the volume V of the system is ex-

$$V = V(U, S, \{n_i\}) \tag{1}$$

pressed in terms of the internal energy U, the entropy S, and $\{n_i\}$, the set of numbers of moles of the various components labeled by the index i. A suitable transformation procedure can be used to replace extensive variables by conjugate intensive variables. For example, the volume can be expressed as in Eq. (2*a*) or (2*b*).

$$V = V(P, S, \{n_i\}) \tag{2a}$$

$$V = V(P, T, \{n_i\}) \tag{2b}$$

Because temperature T and pressure P are particularly convenient variables to control and measure in chemical systems, the form of Eq. (2*b*) is of great utility. All extensive thermodynamic properties X can be rewritten in this form, as Eq. (3). Because all such

$$X = X(T, P, \{n_i\}) \tag{3}$$

properties are linear homogeneous functions of the mass, it can be shown that at a given temperature and pressure Eq. (4) holds, where $\bar{X}_i$ is the partial molar

$$X = \sum_i n_i \bar{X}_i \tag{4}$$

value of the extensive property for the ith component. The value of $\bar{X}_i$ is given by Eq. (5), where the

$$\bar{X}_i = (\partial X/\partial n_i)_{T,P,\{n_i\}'} \tag{5}$$

notation $\{n_i\}'$ means that all amounts are constant, except the ith one involved in the derivative. The variable $\bar{X}_i$ is itself intensive.

Specification of boundaries. The concept of a boundary enclosing the system and separating it from the surroundings requires specification of the nature of the boundary and of any constraints that the boundary places upon the interaction of the system and its surroundings. Boundaries that restrain a system to a particular value of an extensive property are said to be restrictive with respect to that property. A boundary which restrains the system to a given volume V is a fixed, rigid wall. A boundary which is restrictive to one component of a system but not to the other components is called a semipermeable wall or membrane. A system whose boundaries are restrictive to energy and to mass of components is said to be an isolated system. A system whose boundaries are restrictive only to mass of components is called a closed system, whereas an open system has nonrestrictive walls and hence can exchange energy, volume, and mass with its surroundings. Boundaries can be restrictive with respect to specific forms of energy. Two important types are those restrictive to thermal energy but not work (adiabatic walls), and those restrictive to work but not thermal energy (diathermic walls).

First law of thermodynamics. Thermodynamics concerns the conversion of energy. The laws of thermodynamics differ from other scientific laws in that they are stated in terms of the impossibility of achieving certain types of energy transfers. The laws are usually stated in terms of the primary thermodynamic variables of temperature, energy, heat, work, and entropy.

Stated in terms of the primary thermodynamic functions of temperature, energy, and entropy, the laws of thermodynamics for any actual (that is, nonhypothetical) process are the first law: the total energy is conserved (that is, remains unchanged); the second law: the total entropy increases in any real process; and the third law: the absolute zero of temperature always cannot be reached in any real process.

Temperature. A device designed to measure temperature is called a thermometer. Familiar examples are the liquid-in-glass capillary, thermocouples, and resistance thermometers. All these devices have a property that varies monotonically with temperature. The fundamental thermometer is the constant-volume–ideal-gas thermometer, which is based on the fact that gases for which the constituent molecules do not combine or disassociate (that is, $A + A \rightleftharpoons A_2$ or $A_2 \rightleftharpoons 2A$) obey the ideal-gas equation (6) in the limit as $P \to 0$, where P is the

$$PV = nRT \tag{6}$$

pressure of the gas, V is the volume, n is the number of moles of the gas, and R is the gas constant. With a constant-volume gas thermometer, the temperature of a fixed reference point, such as the melting point of lead metal, is measured using Eq. (7).

$$T = (273.16\ \mathrm{K}) \lim_{P_{273.16\ \mathrm{K}} \to 0} \left[\frac{P_T}{P_{273.16\ \mathrm{K}}} \right] \tag{7}$$

The pressure, P_T, of the fixed mass of the gas at temperature T is measured for progressively smaller values of the mass of the gas. The resulting calculated values of temperature are extrapolated to zero pressure to obtain the thermodynamic temperature. The value of $P_{273.16\ \mathrm{K}}$ is the corresponding measured pressure of the gas when the thermometer bulb is in contact with water at its triple point. The triple point of water is the single most important reference point in thermometry. It is arbitrarily assigned a temperature of exactly 273.16 K simply in order to preserve a close numerical agreement with older temperature scales, such as the Celsius scale. *See* TRIPLE POINT.

The temperature scale defined by Eq. (7) is an absolute scale. Zero is the lowest possible temperature, and all measurable temperatures on this scale are positive. This temperature scale is called the absolute Kelvin temperature scale, and temperatures are expressed as kelvins. Celsius temperatures are related to Kelvin temperatures via Eq. (8). The temperatures

$$t(^\circ\mathrm{C}) = T(\mathrm{K}) - 273.15 \tag{8}$$

of fixed reference points obtained in the manner

outlined above are, in turn, used to calibrate other, more convenient thermometers. *See* ABSOLUTE ZERO; GAS THERMOMETRY; TEMPERATURE; THERMOMETER.

Energy, work, and heat. Energy is an abstract mathematical concept that is characterized by an energy function, U. Some other thermodynamic functions also are called energy functions (Gibbs energy, Helmholtz energy) and are chosen specifically as state functions that yield exact differentials in terms of experimentally convenient variables such as temperature and pressure. Energy is not a thing. There are no meters capable of the direct measurement of energy, yet there are instruments that measure electric current, gas flow, volume, mass, pressure, length, time, and temperature. These measurable properties are used to calculate energy changes. Determination of the amount of energy transferred from one system to another always involves measurements of the appropriate physical parameters, followed by calculations involving the appropriate energy formulas. *See* ENERGY.

The principle of energy conservation amounts to the statement that, for any real process, the total energy of the system remains unchanged during the process. Energy can be transferred and transformed, but it cannot be created or destroyed. Work and heat are the two modes of energy transfer between systems. The transfer of energy as work requires the existence of an unbalanced force between a system and its surroundings. The transfer of energy as heat requires the existence of a temperature difference between the system and its surroundings. When a force F acts on a system and displaces the system by a differential amount dx, then the work done on the system is given by Eq. (9).

$$\delta w = -F\,dx \tag{9}$$

See CONSERVATION OF ENERGY.

The work w done on a system depends on the path along which the process is carried out, which must be specified before w can be calculated. When the path must be specified before the integral in Eq. (9) can be evaluated, then the integral is called a line integral. If the work done is independent of the path, then the force F is said to be conservative. A conservative force can always be expressed as the negative derivative of a potential energy, ϕ, with respect to the displacement, as in Eq. (10). Thus, for a conservative force, Eq. (11) holds for the process.

$$F = -\frac{d\phi}{dx} \tag{10}$$

$$w = -\int F\,dx = \int_1^2 \frac{d\phi}{dx}\,dx = \int_1^2 d\phi = \phi_2 - \phi_1 \tag{11}$$

Equation (11) shows that, if the force is conservative, then the work done depends only on the end states (1 and 2) and is thus path-independent. *See* WORK.

If at any stage in the process energy is transferred as heat, then the forces acting on the system cannot be conservative, the value of w depends on the path, and the discipline that describes the process moves from classical mechanics to thermodynamics. For a general process involving both work and heat transfers, the first law of thermodynamics is given by Eq. (12). *See* HEAT.

$$dU = \delta q + \delta w \tag{12}$$

The energy function, U, is called the internal energy, and it is a thermodynamic state function. The differential of a thermodynamic state function is said to be exact because the value of $\Delta U = U_2 - U_1$ for the process is path-independent, even if energy is transferred as heat and work in the process, as in Eq. (13).

$$\begin{aligned}\int_1^2 dU &= U_2 - U_1 = \Delta U \\ &= \int_1^2 \delta q + \int_1^2 \delta w = q + w\end{aligned} \tag{13}$$

See DIFFERENTIAL EQUATION; INTERNAL ENERGY.

A key point is that for a given change in state (that is, $1 \rightarrow 2$) the value of $U_2 - U_1$ is independent of path, whereas the values of both q and w depend on the path. It is only the sum $q + w$ that is independent of path. This statement constitutes the essence of the first law of thermodynamics. The following special cases are notable. If the process is adiabatic (that is, $\delta q = 0$), then δw is exact, because $w = \Delta U$. If no work is done (that is, $\delta w = 0$), then δq is exact (pure heat transfer), because $q = \Delta U$. *See* ADIABATIC PROCESS.

Reversible and irreversible processes. Changes in the state of the system can result from processes taking place within the system and from processes involving exchanges of mass or energy with the surroundings. After a process is carried out, if it is possible to restore both the system and the surroundings completely to their original states, then the process is said to be reversible; otherwise the process is irreversible. All naturally occurring spontaneous processes are more or less irreversible.

The internal energy of the system is given by the fundamental equation of state (14). Processes which

$$U = U(S, V, \{n_i\}, \{X_j\}) \tag{14}$$

give rise to a change in U are then limited to those for which Eq. (15) holds. Because all the extensive

$$\begin{aligned}dU = {} & (\partial U/\partial S)_{V,\{n_i\},\{X_j\}}\,dS + (\partial U/\partial V)_{S,\{n_i\},\{X_j\}}\,dV \\ & + \sum_i (\partial U/\partial n_i)_{S,V,\{n_i\}',\{X_j\}}\,dn_i \\ & + \sum_j (\partial U/\partial X_j)_{X,V,\{n_i\},\{X_j\}'}\,dX_j\end{aligned} \tag{15}$$

state properties are linear homogeneous functions, the coefficients of the differential terms are themselves intensive functions and correspond in each case to the conjugate variable of the respective extensive variable. Their product is thus the differential work associated with the appropriate form of energy transfer. The **table** lists several forms of internal energy transfer, their conjugate pair of variables, and their corresponding work terms.

Internal energy and generalized work

Type of energy	Intensive factor	Extensive factor	Element of work
Mechanical			
Expansion	Pressure (P)	Volume (V)	$-PdV$
Stretching	Surface tension (γ)	Area (A)	γdA
Extension	Tensile stretch (F)	Length (l)	Fdl
Thermal	Temperature (T)	Entropy (S)	TdS
Chemical	Chemical potential (μ)	Amount (n)	μdn
Electrical	Electric potential (E)	Charge (Q)	EdQ
Gravitational	Gravitational field strength (mg)	Height (h)	$mgdh$
Polarization			
Electrostatic	Electric field strength ($\mathcal{E}$)	Total electric polarization (dP)	$\mathcal{E}dP$
Magnetic	Magnetic field strength ($\mathcal{H}$)	Total magnetic polarization (M)	$\mathcal{H}dM$

Equation (15) can be rewritten as Eq. (16), where

$$dU = T\,dS - p\,dV + \sum_i \mu_i\,dn_i + \sum_j I_j\,dX_j \quad (16)$$

μ_i is the chemical potential of the ith component and I_j is the conjugate potential for X_j. The internal energy change given by Eq. (16) is dependent only upon the state properties of the system, and hence is independent of the process causing the change.

Enthalpy and heat capacity. A process involving only pressure-volume work may be described by Eq. (17).

$$\delta w = -F\,dx = -(F/A)A\,dx = -P_{\text{ext}}\,dV \quad (17)$$

Here, P_{ext} denotes the externally applied pressure, the definition of pressure $P = F/A$ = force/area is employed, and the volume element dV = (area of the base) × (vertical displacement) = $A\,dx$. Combination of Eqs. (12) and (17) yields Eq. (18).

$$dU = \delta q - P_{\text{ext}}\,dV \quad (18)$$

The enthalpy H, a thermodynamic state function, is defined by Eq. (19), and differentiation of this equation yields Eq. (20). Combination of Eqs. (18)

$$H = U + PV \quad (19)$$

$$dH = dU + P\,dV + V\,dP \quad (20)$$

and (20) yields Eq. (21), and, at constant pressure ($dP = 0$), Eq. (22) holds. From this equation it fol-

$$dH = \delta q + V\,dP \quad (21)$$

$$dH = \delta q_p \quad (22)$$

lows that the addition or removal of energy as heat from a system at constant pressure is equal to the change in enthalpy of the system ($\Delta H = q_p$). *See* ENTHALPY.

The heat capacity of a system is a quantitative measure of the capacity of a system to take up or supply energy as heat. The greater the heat capacity of a system, the smaller the change in temperature produced by the input of energy as heat. The heat capacity is necessarily positive. The heat capacity C_x of a system at constant x (where x is a thermodynamic state function, such as pressure) is defined by Eq. (23), where $C_x(x, T)$ denotes that C_x is a func-

$$C_x = \lim_{\Delta T \to 0}\left(\frac{q_x}{\Delta T}\right) = C_x(x, T) \quad (23)$$

tion of both x and T. For a constant-pressure process, it follows from Eqs. (22) and (23) that Eq. (24)

$$C_p = \lim_{\Delta T \to 0}\left(\frac{q_p}{\Delta T}\right) = \lim_{\Delta T \to 0}\left(\frac{\Delta H}{\Delta T}\right)_p = \left(\frac{\partial H}{\partial T}\right)_p \quad (24)$$

holds. Also, from Eqs. (24) and (22), it follows that Eq. (25) is valid. A large heat capacity indicates that

$$q_p = \int_1^2 C_p\,dT \quad (25)$$

the system can take up substantial amounts of energy as heat, owing to storage of the energy in the internal modes (translation, rotation, vibration, molecular dissociation) of the constituent particles. For example, the molar heat capacities of solid (s), liquid (l), and gaseous (g) water over the temperature range 250 to 500 K (−23 to 227°C or −10 to 440°F) are as follows:

$$\begin{aligned} H_2O(s):\ C_p &= 37.66\ \text{J}\cdot\text{K}^{-1}\cdot\text{mol}^{-1} \\ H_2O(l):\ C_p &= 75.32\ \text{J}\cdot\text{K}^{-1}\cdot\text{mol}^{-1} \\ H_2O(g):\ C_p &= 30.54\ \text{J}\cdot\text{K}^{-1}\cdot\text{mol}^{-1} \\ &\quad + (1.03\times10^{-2}\ \text{T})\ \text{J}\cdot\text{K}^{-1}\cdot\text{mol}^{-1} \end{aligned}$$

The roughly twofold greater heat capacity of liquid water relative to solid or gaseous water results from the fact that a substantial (roughly half) part of the heat input to liquid water is absorbed in breaking the hydrogen bonds between water molecules in liquid water (about 20 kJ/mol). Ice absorbs energy as heat below its melting point without breaking any hydrogen bonds and thus has a lower value of C_p than liquid water, whereas there are no hydrogen bonds in gaseous water and thus this mode of energy uptake is absent for the gas. *See* HEAT CAPACITY.

Second law of thermodynamics. Most naturally occurring processes involve the transfer of energy as heat. The first law of thermodynamics places no restrictions on energy transfers other than that of

conservation of energy. There are, however, other restrictions on energy transfer, namely:

1. Heat always flows spontaneously from higher- to lower-temperature systems.
2. Heat can be induced to flow from a lower- to a higher-temperature system only through the use of energy as work to drive the transfer.
3. When a system undergoes a change in thermodynamic state, then the system and its surroundings cannot both be restored exactly to their original states. All naturally occurring processes are in this sense irreversible.

A mathematical statement of the second law of thermodynamics is that the change in entropy of the system, dS_{sys}, obeys inequality (26), where T is the

$$dS_{sys} \geq \frac{\delta q_{sys}}{T} \tag{26}$$

absolute thermodynamic temperature [measured in kelvins (K)]. The equality sign holds if the process is reversible, and the inequality sign holds if the process is irreversible. The entropy S is a thermodynamic state function, and it is expressed in units of joules per kelvin (J/K).

It is useful to consider a system and its surroundings between which energy can be transferred. If the system undergoes a change in state, then the total entropy change is given by Eq. (27). If the process is

$$dS_{tot} = dS_{sys} + dS_{sur} \tag{27}$$

adiabatic, then $\delta q_{sys} = \delta q_{sur} = 0$ and from Eq. (26) $dS_{sur} \geq 0$ and $dS_{sys} \geq 0$. Thus $dS_{tot} \geq 0$, or $\Delta S_{tot} \geq 0$.

If the system undergoes an isothermal process, then $T = T_{sys} = T_{sur}$. From the second law, it follows that the system and the surroundings obey inequalities (28), but $q_{sys} = -q_{sur}$ and thus $\Delta S_{tot} \geq 0$. Note

$$dS_{sys} \geq \frac{\delta q_{sys}}{T} \quad \text{and} \quad dS_{sur} \geq \frac{\delta q_{sur}}{T} \tag{28}$$

that in both cases there is a net entropy production if the process is irreversible, but the total entropy remains unchanged if the process is reversible.

All real processes involve some degree of irreversibility and thus lead to an increase in the total entropy. Entropy is not conserved, except in the hypothetical limiting case of a reversible process.

A thermodynamic system is said to be in an equilibrium state when it can no longer undergo any spontaneous (entropy-producing) processes. The increase in total entropy that occurs in an irreversible process is a consequence of the failure to fully exploit the potential of the system for performing work. When less than the maximum possible work is performed in a process, then net entropy production inevitably results. The total entropy can never decrease and, for a given energy and volume for a closed system, a maximum in entropy indicates the most stable configuration of the system, that is, equilibrium.

Because entropy is a state function, Eq. (29) is

$$\int_1^2 dS = S_2 - S_1 = \Delta S \tag{29}$$

satisfied. For a given change in state ($1 \rightarrow 2$) the change in entropy of the system (but not the total entropy change, $\Delta S_{tot} = \Delta S_{sys} + \Delta S_{sur}$) is independent of the path that the system takes from state 1 to state 2.

There are no entropy meters. Rather, entropy changes are calculated from measurable quantities, such as temperature, pressure, volume, and heat capacity. The total entropy of a system is a measure of the spatial and energy disorder (randomness) in a system.

Compared at the same pressure and temperature, liquids are more disordered than solids and gases are more disordered than liquids. This results follows from the fact that the molecules of a gas have more freedom to move around than the molecules of a liquid. The same holds true for the molecules or ions of a liquid as compared to those of a solid. The melting of a solid involves an entropy increase in the range 8–40 $J \cdot K^{-1} \cdot mol^{-1}$. The temperature dependence of the entropy is found from Eqs. (26) and (24). From Eq. (26) it follows that $q_p = T\Delta S$ for an isothermal pure heat transfer (that is, $\delta w = 0$) at constant pressure. Substitution of this result into Eq. (24) yields Eq. (30). Thus, Eqs. (31) and (32) are valid.

$$C_p = \lim_{\Delta T \to 0} \left(\frac{q_p}{\Delta T} \right) = T \lim_{\Delta T \to 0} \left(\frac{\Delta S}{\Delta T} \right)_p = T \left(\frac{\partial S}{\partial T} \right)_p \tag{30}$$

$$\left(\frac{\partial S}{\partial T} \right)_p = \frac{C_p}{T} \tag{31}$$

$$\Delta S = \int_1^2 \frac{C_p}{T}\, dT \quad \text{(constant } P\text{)} \tag{32}$$

Because C_p is greater than 0, the entropy of a substance always increases with increasing temperature at constant pressure. The increase in the entropy of a substance that undergoes an increase in temperature at fixed pressure is a consequence of the increased thermal randomness of the molecules at the higher temperature. A greater amount of energy is more randomly distributed over the same number of particles. *See* ENTROPY.

Gibbs energy and Maxwell relations. Combination of the first law, Eq. (12), with the second law, Eq. (26), yields Eq. (33). If a process involves only pressure-

$$dU \geq T\, dS + \delta w \tag{33}$$

volume work and occurs reversibly, then, applying Eq. (17), Eq. (33) becomes Eq. (34). Because U and

$$dU = T\, dS - P\, dV \tag{34}$$

S are state functions, the values of ΔU and ΔS can be computed along a reversible path between the given initial and final states of the actual process, and the results for ΔU and ΔS for the system can be equated to those for the actual irreversible process. The values of q and w differ for the two paths, but not the values of ΔU_{sys} or ΔS_{sys}.

Equation (34) is the differential equation used to calculate ΔU. It expresses the change in the dependent variable in terms of changes in the independent variables S and V; that is, $U = U(S, V)$. The variables S and V are not convenient independent variables because they are very difficult to control in the laboratory. More convenient independent variables for experimental work are pressure and temperature. Equation (34) can be transformed into a differential equation with P and T as independent variables through Eqs. (35).

$$dU - T\,dS + P\,dV = 0 \tag{35a}$$

$$dU - T\,dS + P\,dV - S\,dT + V\,dP = -S\,dT + V\,dP \tag{35b}$$

$$d(U - TS + PV) = -S\,dT + V\,dP \tag{35c}$$

For compactness, the thermodynamic energy state function G is defined by Eq. (36). It can be used as a

$$G = U + PV - TS = H - TS \tag{36}$$

criterion for system equilibrium in terms of the more convenient variables, T and P. Here, H is defined by Eq. (19).

The function G is called the Gibbs energy after its creator, J. W. Gibbs; it plays a central role in chemical thermodynamics. From Eqs. (34) and (36), it obeys Eq. (37), which gives the differential equation for

$$dG = -S\,dT + V\,dP \tag{37}$$

dG, where $G = G(T, P)$. Because dG is an exact differential, G is a state function, and from mathematics Eq. (38) holds. Comparison of Eqs. (38) and (37) leads to Eqs. (39) and (40). Equation (39) shows that

$$dG = \left(\frac{\partial G}{\partial T}\right)_p dT + \left(\frac{\partial G}{\partial P}\right)_T dP \tag{38}$$

$$\left(\frac{\partial G}{\partial T}\right)_p = -S \tag{39}$$

$$\left(\frac{\partial G}{\partial P}\right)_T = V \tag{40}$$

the temperature dependence of the Gibbs energy at constant pressure is determined by the entropy, and Eq. (40) shows that the pressure dependence of the Gibbs energy at constant temperature is determined by the volume.

Because dG is an exact differential, it follows directly that the second cross partial derivatives of Eq. (37) are equal, as given by Eq. (41). This property is characteristic of an exact differential. Thus, from Eqs. (39), (40), and (41), the Maxwell relation, Eq. (42), is obtained. This relation tells how the entropy changes with pressure at constant temperature.

$$\left[\frac{\partial}{\partial P}\left(\frac{\partial G}{\partial T}\right)_P\right]_T = \left[\frac{\partial}{\partial T}\left(\frac{\partial G}{\partial P}\right)_T\right]_P \tag{41}$$

$$-\left(\frac{\partial S}{\partial P}\right)_T = \left(\frac{\partial V}{\partial T}\right)_P \tag{42}$$

Affinity and chemical equilibrium. Many chemical systems can be considered closed systems in which a single parameter ξ can be defined as a measure of the extent of the reaction or the degree of advancement of a process. If the reaction proceeds or the process advances spontaneously, then entropy must be produced according to the second law, and the uncompensated heat, defined by Eq. (43), must

$$\delta a = T\,dS + \delta w - dU \tag{43}$$

be positive. In terms of the advancement parameter ξ, the uncompensated heat δa is given by Eq. (44),

$$\delta a = \mathcal{A}\,d\xi = T\,d_iS \tag{44}$$

where $\mathcal{A}$ is the affinity of the process or reaction. The affinity is related to internal entropy production d_iS by Eq. (45).

$$\mathcal{A} = T\,d_iS/d\xi \geq 0 \tag{45}$$

The condition that the entropy production is zero represents equilibrium, and hence $\mathcal{A} = 0$ is an equivalent condition for equilibrium in a closed system. For spontaneous processes, because the signs of $\mathcal{A}$ and $d\xi$ must be the same, for positive $\mathcal{A}$ the process must advance or go in a forward direction in the usual sense of chemical reactions or physical processes, while for negative $\mathcal{A}$ the process must proceed in the reverse direction.

For a process that involves only pressure-volume work, Eqs. (17), (36), (43), and (44) can be combined to give Eqs. (46). These equations indicate that the affinity is a state function, given by Eq. (47).

$$dG = -S\,dT + V\,dP - \delta a \tag{46a}$$

$$= -S\,dT + V\,dP - \mathcal{A}\,d\xi \tag{46b}$$

$$\mathcal{A} = -(\partial G/\partial \xi)_{P,T} = \mathcal{A}(T, P, \xi) \tag{47}$$

It is useful to consider a closed system in which a chemical reaction can be characterized by the stoichiometry of reaction (48). This stoichiometry requires that at each time element t in the reaction the number of moles of the ith component n_i be given by Eq. (49), where n_i^0 is the number of moles of i

$$\alpha\mathrm{A} + \beta\mathrm{B} \rightarrow \gamma\mathrm{C} + \delta\mathrm{D} \tag{48}$$

$$n_i = n_i^0 + \upsilon_i\xi \tag{49}$$

in the initial or original state ($t = 0$). The quantity υ_i is the stoichiometric coefficient for the ith component as given in the balanced equation (the convention is that υ_i is positive for products and negative for reactants), and ξ is the degree-of-advancement parameter, whose range is zero to unity. In terms of differential changes in advancement, Eq. (50) holds.

$$dn_i = \upsilon_i\,d\xi \tag{50}$$

For closed systems (constant temperature and pressure) in which only thermal, expansion, and chemical work terms are included, Eqs. (51) and (52)

$$dG = \sum_i \mu_i \, dn_i = -\delta a \tag{51a}$$

$$= \left(\sum_i \upsilon_i \mu_i \right) d\xi = -\mathcal{A} \, d\xi \tag{51b}$$

$$\mathcal{A}(T, P, \xi) = -\sum_i \upsilon_i \mu_i \tag{52}$$

hold. The condition for equilibrium is $\mathcal{A} = 0$, and thus for a chemical reaction, equilibrium is achieved when Eq. (53) holds. If electrical work is included in Eq. (51), Eqs. (54) and (55) hold, where z is the number of units of charge and F is the Faraday constant (the magnitude of the charge per mole of electrons); $F = eN_A$, where e is the electron charge and N_A is Avogadro's number. Because $\mathcal{A} = 0$ is the equilibrium condition, equilibrium in an electrochemical system is given by Eq. (56).

$$\sum_i \upsilon_i \mu_i = 0 \tag{53}$$

$$dG = \sum_i \mu_i \, dn_i + E \, dQ = -\delta a \tag{54a}$$

$$= \left(\sum_i \upsilon_i \mu_i + zFE \right) d\xi = -\mathcal{A} \, d\xi \tag{54b}$$

$$\mathcal{A}(T, P, E, \xi) = -\sum_i \upsilon_i \mu_i - zFE \tag{55}$$

$$\sum_i \upsilon_i \mu_i = -zFE \tag{56}$$

More than one reaction can take place in a chemical system, each characterized by a degree-of-advancement parameter ξ_r, and thus for r independent reactions, Eqs. (57) and (58) hold, where the equilibrium condition is Eq. (59). At equilibrium,

$$dG = \sum_i \mu_i \, dn_i = -\delta a \tag{57a}$$

$$= -\sum_r \mathcal{A}_r \, \delta\xi_r \tag{57b}$$

$$= \sum_r \left(\sum_i \upsilon_{ir} \mu_i \, \delta\xi_r \right) \tag{57c}$$

$$\mathcal{A}_r(T, P, \{\xi_r\}) = -\sum_i \upsilon_{ir} \mu_i \tag{58}$$

$$\sum_r \mathcal{A}_r \, d\xi_r = 0 \tag{59}$$

each of the affinities $\mathcal{A}_r$ must be zero; but for the spontaneous condition, inequality (60) holds. In a two-reaction system inequality (61) holds, but now

$$\sum_r \mathcal{A}_r \, d\xi_r > 0 \tag{60}$$

$$\mathcal{A}_1 \, d\xi_1 + \mathcal{A}_2 \, d\xi_2 > 0 \tag{61}$$

$\mathcal{A}_1$ and $d\xi_1$ do not necessarily have the same sign. If their signs are different, then the first reaction can be driven in the nonspontaneous direction by the second reaction. The reactions are then said to be coupled; reaction coupling is a common situation in biological systems. *See* CHEMICAL EQUILIBRIUM.

Third law of thermodynamics. Although implicit in the concept of the existence of the fundamental equations of state for U or S is an absolute value of these functions, and therefore an extensive quantity which could be calculated on the basis of molecular properties from quantum mechanics and statistical mechanics, neither the first nor second law considers anything but differences in these state functions. The second law does suggest the existence of an absolute zero for an intensive variable, the temperature, but this is not sufficient to bridge the areas of classical and statistical thermodynamics. *See* ABSOLUTE ZERO.

It is found experimentally, for many isothermal processes involving pure phases, that Eq. (62) applies. This equation includes phase transitions between different crystalline modifications, solid-state chemical reactions, and even the solid-liquid and liquid-liquid isotope transitions in helium. This equation along with Eq. (63), obtained from Eq. (31), implies that at zero absolute temperature the entropy of pure crystalline phases of the same substance or different substances are equal. If the entropies of pure phase are equal at $T = 0$, then it is reasonable to take the value of $S(0)$ to be zero. A statement of the third law then is that the entropy of all pure crystalline fully ordered phases at $T = 0$ is zero. This makes it possible, by using Eq. (63), to calculate the absolute or third-law entropy of a substance from experimental measurements of its heat capacities. Comparison of such experimental values with those calculated by statistical thermodynamic methods has provided evidence for the validity of the third law. In some cases, thorough investigation of apparent discrepancies from the third law has led to new conclusions concerning the molecular structure of the substances or new information on the energy level system for the molecules. Calculation of the thermodynamic properties for a gas from the spectroscopic properties of molecules is an important result stemming from the third law. An alternative statement of the third law is that the absolute zero of temperature is unattainable in any real process. *See* STATISTICAL MECHANICS.

$$\lim_{T \to 0} \Delta S = 0 \tag{62}$$

$$S(T, P) - S(0, P) = \int_0^T (C_p/T) \, dT \tag{63}$$

Thermodynamics of irreversible processes. From a mathematical standpoint, if the equations describing a time-dependent physical process change sign when the time variable t is changed to $-t$, then the process is said to be irreversible. If the substitution of t by $-t$ does not change the equations, then the process

described is said to be reversible. An example of an irreversible thermodynamic process is the approach to thermal equilibrium, which is described by the Fourier equation (64), where T is the temperature, t

$$\frac{1}{\alpha}\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \tag{64}$$

is the time, and α is a parameter called the thermal diffusivity. The physicochemical processes of diffusion, heat and electrical conduction, and chemical reactions are irreversible processes. *See* CONDUCTION (ELECTRICITY); CONDUCTION (HEAT); DIFFUSION.

By utilizing Eqs. (44) and (26), the entropy production is found to be given by Eq. (65), where $\mathcal{A}$,

$$dS = \frac{\delta q}{T} + \frac{\mathcal{A}\, d\xi}{T} = d_e S + d_i S \tag{65}$$

the affinity of the chemical reaction, is related to the chemical potentials by Eq. (52). The reaction rate v is defined by Eq. (66), and thus Eq. (67) holds. It is clear

$$v = \frac{d\xi}{dt} \tag{66}$$

$$\frac{d_i S}{dt} = \frac{\mathcal{A} v}{T} > 0 \tag{67}$$

that $\mathcal{A}$ and v always have the same sign and that the affinity $\mathcal{A}$ is directly related to entropy production.

If several reactions are involved simultaneously, then Eq. (68) holds, where the affinities are related

$$d_i S = \frac{1}{T}\sum_k \mathcal{A}_k\, d\xi_k > 0 \tag{68}$$

to the chemical potentials by Eq. (58). At equilibrium all of the affinities are zero. The entropy production per unit time is given by Eq. (69).

$$\frac{d_i S}{dt} = \frac{1}{T}\sum_k \mathcal{A}_k v_k > 0 \tag{69}$$

The net entropy production arising from the simultaneous reactions must be positive. However, it is possible to have the inequalities (70) in a two-reaction system, provided that inequality (71) is sat-

$$\mathcal{A}_1 v_1 < 0, \qquad \mathcal{A}_2 v_2 > 0 \tag{70}$$

$$\mathcal{A}_1 v_1 + \mathcal{A}_2 v_2 > 0 \tag{71}$$

isfied. As noted above, such reactions are said to be coupled in that a reaction can proceed in a direction opposed to its affinity, a phenomenon frequently encountered in biological systems. For example, matter may diffuse against its concentration gradient with negative entropy production driven by positive entropy production arising from heat flow.

Equation (69) can be rewritten as Eq. (72), where J_k and X_k are defined by Eqs. (73). The variable J_k

$$\frac{d_i S}{dt} = \sum_k J_k X_k > 0 \tag{72}$$

$$J_k = v_k, \qquad X_k = \frac{\mathcal{A}_k}{T} \tag{73}$$

is called a flux (rate) and the variable X_k is called a (generalized) force. At thermodynamic equilibrium Eqs. (74) are satisfied for all values of k.

$$J_k = 0, \qquad X_k = 0 \tag{74}$$

Close to equilibrium, various empirical laws, such as Fourier's law for heat flow and Fick's law for diffusion, show a linear relation between the J_k and X_k values. Such linear laws are called phenomenological relations. Although such relations do not necessarily hold far from equilibrium, it will be assumed that they do hold close to equilibrium.

It is useful to consider the case of two simultaneous irreversible processes which, in general, satisfy Eqs. (75). The coefficients L_{ij} are called phe-

$$\begin{aligned} J_1 &= L_{11}X_1 + L_{12}X_2 \\ J_2 &= L_{21}X_1 + L_{22}X_2 \end{aligned} \tag{75}$$

nomenological coupling coefficients. For example, L_{11} and L_{22} may denote heat conductivity and diffusion, while L_{12} describes the interference between these two processes (that is, thermodiffusion). Thus the L_{ij} $(i \neq j)$ are called interference coupling coefficients.

Substitution of Eq. (75) into Eq. (72) yields Eq. (76), where the equality sign prevails when

$$\frac{d_i S}{dt} = L_{11}X_1^2 + (L_{12} + L_{21})X_1X_2 + L_{22}X_2^2 \geq 0 \tag{76}$$

$X_1 = X_2 = 0$. From algebra it follows that the L_{ij} values must satisfy inequalities (77) and (78). Thus

$$L_{11} > 0, \qquad L_{22} > 0 \tag{77}$$

$$(L_{12} + L_{21})^2 < 4L_{11}L_{22} \tag{78}$$

while L_{11} and L_{22} must be positive, L_{12} and L_{21} may be positive or negative.

Onsager's theorem states that the coupling coefficients satisfy Eq. (79), which is called the Onsager

$$L_{ij} = L_{ji} \tag{79}$$

reciprocity relation after its derivation by Lars Onsager. These relations express the fact that if the flux J_i corresponding to the irreversible process i is coupled to the force X_j of the irreversible process j, then the flux J_j also is influenced by the force X_i through the same interference coefficient L_{ij}. *See* OPEN-SYSTEMS THERMODYNAMICS (BIOLOGY); OSMOREGULATORY MECHANISMS; THERMOELECTRICITY; THERMOMAGNETIC EFFECTS.

In the general case, Eq. (80) holds, and when

$$J_i = \sum_i L_{ij}X_j \tag{80}$$

$L_{ij} \neq 0$ there is a coupling between the flux J_i and the force (gradient) X_j. The principle of microscopic reversibility implies that at least close to equilibrium $L_{ij} = L_{ji}$. Far from equilibrium the

situation becomes much more complex, and a detailed molecular-level kinetic (as opposed to a phenomenological) approach is required in the analysis of such cases. *See* THERMODYNAMIC PRINCIPLES; THERMODYNAMIC PROCESSES. Peter A. Rock

Bibliography. S. W. Angrist and L. G. Hepler, *Order and Chaos*, Basic Books, New York, 1967; H. A. Bent, *The Second Law*, Oxford University Press, New York, 1965; J. Keizer, *Statistical Thermodynamics of Nonequilibrium Processes*, Springer-Verlag, New York, 1987; G. N. Lewis and M. Randall, *Thermodynamics*, 2d ed., revised by K. S. Pitzer and L. Brewer, McGraw-Hill, New York, 1961; A. B. Pippard, *Elements of Classical Thermodynamics*, Cambridge University Press, London, 1966; I. Prigogine, *Introduction to Thermodynamics of Irreversible Processes*, 3d ed., Wiley Interscience, New York, 1967; M. W. Zemansky, *Temperatures: Very Low and Very High*, Van Nostrand, Princeton, NJ, 1964.

Chemical vapor deposition

The process of producing a microscopic solid-material structure through chemically reacting vapor- or gas-phase reactants on a heated surface. Many of today's most important and most common technologies make use of very thin films of electrically active materials. Chemical vapor deposition (CVD) is a widely used manufacturing technique for products such as electric circuits and processors. Development of novel CVD processes and of new materials for advanced electronics and sensors produced by CVD are active research areas at universities and laboratories around the world. *See* INTEGRATED CIRCUITS; MICROSENSOR.

Basic principle. The fundamental principle of the CVD process is that the chemical reactants, called precursors, are in the gas or vapor state when they arrive at the base material or substrate. A chemical reaction, which is usually activated by heat or some other energy source, occurs on the substrate surface. The products of the reaction are the desired solid-film material and a gaseous product. Layers of semiconductors, dielectrics, conductors, and insulators are built up on the substrate and further processed to produce electronic devices. Scientists study the way the chemistry, flow dynamics, and deposition rate affect the structure and composition of the films and their electrical properties. *See* DIELECTRIC MATERIALS; FILM (CHEMISTRY); SEMICONDUCTOR; SOLID-STATE CHEMISTRY.

Process sequence. The process is summarized in **Figs. 1** and **2.** Reactants and precursor chemicals either are gases or are vaporized from liquid and solid sources by spraying, bubbling, heating, or sublimation. The precursors are supplied as a gas or as vapors in a carrier gas to the reactor chamber. Once in the reactor, the vapors are transported by convection and gas dynamics to the substrate's surface, which is at the desired deposition temperature. Deposition of the thin film occurs on the substrate's surface through reaction of the vapors and gases, usually on a heated substrate. The exhaust products are removed from the reactor by vacuum and treated.

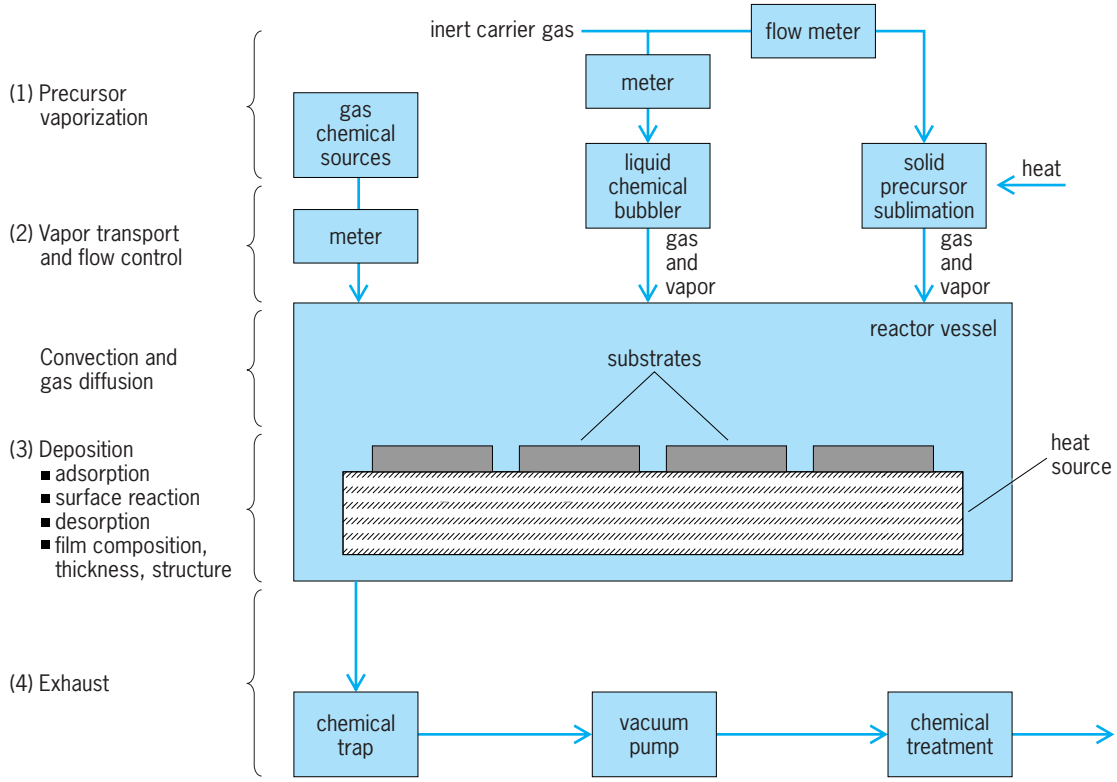

Fig. 1. The components of CVD manufacturing technology reflect the sequence of steps in the CVD process.

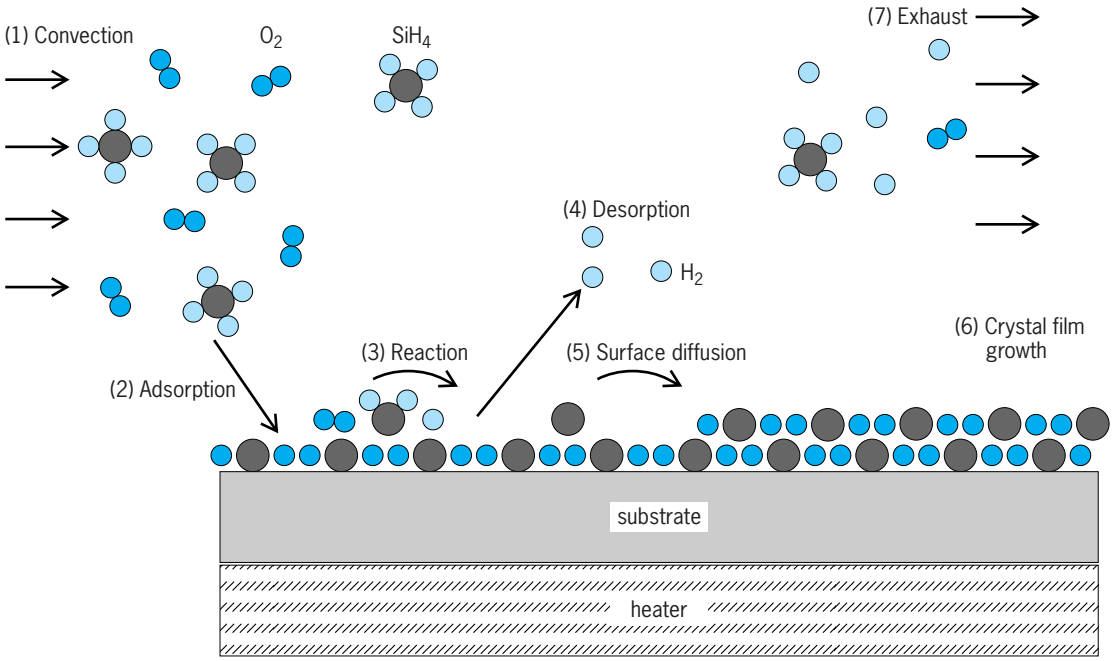

Fig. 2. Sequence of steps involved in the deposition and crystal growth of a solid material (SiO_2) from gas precursors ($SiH_4 + 2O_2$) on the surface of a heated substrate (Si).

Deposition rate. The quality and composition of the thin-film material depends directly on the rates of each step in the process sequence (Fig. 1). The evaporation rate of the precursors is controlled through the flow rate of carrier gases and the evaporation temperature. The feed rate of gas and vapors into the reactor and the pressure at which the reactor is maintained greatly affect the rate at which precursor molecules arrive at the substrate. In addition, the reaction rate at the surface is a strong function of the surface temperature and of the chemical thermodynamics and kinetics. *See* CHEMICAL KINETICS; CHEMICAL THERMODYNAMICS.

Thin-film applications. CVD manufacturing equipment is designed to control each step of the process in order to produce high-quality thin films with the desired properties. The chemistry involved can require the use of high-temperature and high-energy sources, such as plasma, and may require very high vacuum. However, many commercial CVD processes are done at atmospheric pressure, without extreme temperatures. Silicon dioxide (SiO_2) for solid-state electronic devices and optical fibers is produced in a relatively low-temperature (450°C or 840°F) process, using the chemical oxidation reaction of a common precursor, silane (SiH_4) with oxygen (Fig. 2). The CVD system to produce compound semiconductors for solar (photovoltaic) cells (for example, GaAs) must be capable of handling corrosive gases and temperatures over 600°C (1110°F). Batch reactors are used to modify the surface of steel drills and cutting tools by deposition of titanium nitride (TiN) from a precursor gas mixture of titanium tetrachloride ($TiCl_4$), nitrogen (N_2), and hydrogen (H_2) at 1000°C (1840°F). Self-cleaning glass with low reflectivity can be manufactured by deposition of an optically clear film of titanium dioxide (TiO_2) onto glass at less than 450°C (840°F) by using an organometallic precursor [$Ti(OC_3H_7)_4$]. While semiconductor CVD processes have been in practice for well over 20 years, the range of new materials, devices, and applications is expanding rapidly. *See* SEMICONDUCTOR HETEROSTRUCTURES; SOLAR CELL.

Process variations. There are two main categories for CVD processing: atmospheric pressure (APCVD) and low pressure (LPCVD). APCVD is often lower in cost, but it can be more difficult to control the vapor/gas transport processes. LPCVD is most common for electronics, as the low pressure helps ensure the purity necessary for electronically functional materials. When more energy is needed to activate the deposition reaction, plasma-enhanced (PECVD) and laser-enhanced (LECVD) processes are used. Relatively recent developments in a type of precursor chemicals called metal-organic compounds (MOCVD) has led to reduced toxic gas products and reduced processing temperatures.

Developments. CVD researchers investigate the properties of new materials, and the effects that different processing parameters have on these properties. Many scientists also work on reducing toxic-waste gas production through improving the conversion efficiency or through investigating new precursor chemicals. In an effort to reduce the cost of CVD equipment and to improve the control of the evaporation and transport steps in the process, several research groups have begun injecting liquid precursors directly into an evaporation chamber, sometimes using atomization to assist vaporization. For controlling structure, the rate of crystal growth has been found to be enhanced by pulsing, or turning on and off the precursor flow. This gives the crystal time

to grow into more highly ordered structures with improved electrical and other properties. *See* CRYSTAL GROWTH; CRYSTAL STRUCTURE. Susan Krumdieck

Bibliography. F. S. Galasso, *Chemical Vapor Deposited Materials*, CRC Press, 1991; D. A. Glocker and S. I. Shah (eds.), *Handbook of Thin Film Process Technology*, Institute of Physics, Bristol, UK, 1995; M. L. Hitchman and K. F. Jensen (eds.), *Chemical Vapor Deposition Principles and Applications*, Academic Press, London, 1993; M. Ohring, *Materials Science of Thin Films*, Academic Press, San Diego, 2002; H. O. Pierson, *Handbook of Chemical Vapor Deposition (CVD): Principles, Technology, and Applications*, Noyes Publications, Norwich, NY, 1999; H. O. Pierson, *Handbook of Chemical Vapor Deposition: Principles, Technology and Applications*, Noyes, Park Ridge, NJ, 1992; K. K. Schuegraf (ed.), *Handbook of Thin-Film Deposition Processes and Techniques*, Noyes Publications, Park Ridge, NJ, 1988; D. L. Smith, *Thin-Film Deposition Principles and Practice*, McGraw-Hill, Boston, 1995.

Chemiluminescence

The type of luminescence wherein a chemical reaction supplies the energy responsible for the emission of light (ultraviolet, visible, or infrared) in excess of that of a blackbody (thermal radiation) at the same temperature and within the same spectral range. Below 900°F (500°C), the emission of any light during a chemical reaction is a chemiluminescence. The blue inner cone of a bunsen burner or the Coleman gas lamp are examples. *See* BLACKBODY.

Reactions. Many chemical reactions generate energy. Usually this exothermicity appears as heat, that is, translational, rotational, and vibrational energy of the product molecules; whereas, for a visible chemiluminescence to occur, one of the reaction products must be generated in an excited electronic state (designated by an asterisk) from which it can undergo deactivation by emission of a photon. Hence a chemiluminescent reaction, as shown in reactions (1) and (2), can be regarded as the reverse of a photochemical reaction.

$$A + B \rightarrow C^* + D \qquad (1)$$

$$C^* \rightarrow C + h\nu \qquad (2)$$

Energy. The energy of the light quantum $h\nu$ (where h is Planck's constant, and ν is the light frequency) depends on the separation between the ground and the first excited electronic state of C; and the spectrum of the chemiluminescence usually matches the fluorescence spectrum of the emitter. Occasionally, the reaction involves an additional step, the transfer of electronic energy from C* to another molecule, not necessarily otherwise involved in the reaction. Sometimes no discrete excited state can be specified, in which case the chemiluminescence spectrum is a structureless continuum associated with the formation of a molecule, as in the so-called air afterglow: $NO + O \rightarrow NO_2 + h\nu$ (green light). *See* HEAT RADIATION; LUMINESCENCE; PHOTOCHEMISTRY.

Efficiency. The efficiency of a chemiluminescence is expressed as its quantum yield ϕ, that is, the number of photons emitted per reacted molecule. Many reactions have quantum yields much lower (10^{-8} $h\nu$ per molecule) than the maximum of unity. Einsteins of visible light (1 einstein $= Nh\nu$, where N is Avogadro's number), with wavelengths from 400 to 700 nanometers, correspond to energies of about 70 to 40 kcal per mole (300 to 170 kilojoules per mole). Thus only very exothermic, or "exergonic," chemical processes can be expected to be chemiluminescent. Partly for this reason, most familiar examples of chemiluminescence involve oxygen and oxidation processes; the most efficient examples of these are the enzyme-mediated bioluminescences. The glow of phosphorus in air is a historically important case, although the mechanism of this complex reaction is not fully understood. The oxidation of many organic substances, such as aldehydes or alcohols, by oxygen, hydrogen peroxide, ozone, and so on, is chemiluminescent. The reaction of heated ether vapor with air results in a bluish "cold" flame, for example. The efficiency of some chemiluminescences in solution, such as the oxidation of luminol (**1**) and, especially, the reaction of some oxalate esters (**2**) with hydrogen peroxide, can be very high ($\phi = 30\%$).

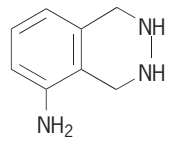

(**1**)

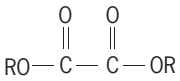

(**2**)

See BIOLUMINESCENCE.

Other aspects. It is believed that the requirements for chemiluminescence are not only that there be sufficient exothermicity and the presence of a suitable emitter, but also that the chemical process be very fast and involve few geometrical changes, in order to minimize energy dissipation through vibrations. For example, the transfer of one electron from a powerful oxidant to a reductant (often two radical ions of opposite charge generated electrochemically) is a type of process which can result, in some cases, in very effective generation of electronic excitation. An example, with 9,10-diphenylanthracene (DPA), is shown in reaction (3). The same is true of the

$$DPA^{-\cdot} + DPA^{+\cdot} \rightarrow DPA^* + DPA \qquad (3)$$

decomposition of four-membered cyclic peroxides (**3**) into carbonyl products, as shown in reaction (4),

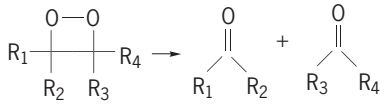

(4)

(**3**)

which may be the prototype of many chemiluminescences. Therese Wilson

Electrogenerated chemiluminescence. Also known as electrochemiluminescence, this is a luminescent chemical reaction in which the reactants are formed electrochemically. Electrochemical reactions are electron-transfer reactions occurring in an electrochemical cell. In such a reaction, light emission may occur as with chemiluminescence; however, the excitation is from the application of a voltage to an electrode. In chemiluminescence, the luminophore is excited to a higher energetic state by means of a chemical reaction initiated by mixing of the reagents. In electrogenerated chemiluminescence, the emitting luminophore is excited to a higher energy state by reactions of species that are generated at an electrode surface by the passage of current through the working electrode. Upon decay to the electronic ground state, light emitted by the luminophore (fluorescent or phosphorescent) can be detected. The luminophore is typically a polycyclic hydrocarbon, an aromatic heterocycle, or certain transition-metal chelates. *See* ELECTRON-TRANSFER REACTION.

Reaction mechanism. There are many reactions which give rise to electrogenerated chemiluminescence. However, they all share some common features. A typical reaction scheme involves the oxidation of a luminophore (Z), as in reaction, and reduction of a second molecule of the luminophore, as in reaction (5), and reduction of a second molecule of the luminophore, as in reaction (6) where e^- is an electron.

$$Z \longrightarrow Z^+ + e^- \tag{5}$$

$$Z + e^- \longrightarrow Z^- \tag{6}$$

The excited state of the luminophore (Z*) is formed by a homogeneous charge-transfer reaction (7), be-

$$Z^+ + Z^- \longrightarrow Z^* + Z \tag{7}$$

tween the oxidized and reduced luminophore. The excited luminophore decays to the ground state and emits a photon at a wavelength characteristic of the luminophore, which is measured by means of a photomultiplier detector [(8)].

$$Z^* \xrightarrow{h\nu} Z \tag{8}$$

An example is the oxidative-reduction reaction using ruthenium tris(2,2′-bipyridine) [$Ru(bpy)_3^{2+}$], and oxalate $C_2O_4^{2-}$. The luminophore $Ru(bpy)_3^{2+}$ and oxalate are oxidized by application of a positive voltage to the electrode [reactions (9) and (10)].

$$Ru(bpy)_3^{2+} \longrightarrow Ru(bpy)_3^{3+} + e^- \tag{9}$$

$$C_2O_4^{2-} \longrightarrow C_2O_4^- \cdot + e^- \tag{10}$$

In both reactions, one electron is removed for each molecule of starting material. The oxidized oxalate rapidly disassociates [reaction (11)]. The products formed are carbon dioxide and, importantly, a strong reducing agent, $CO_{2-}\cdot$ radical anion. The term oxidative-reduction is descriptive of the process of forming a strong reducing agent through an electrochemical oxidation reaction. An exothermic and energetic reaction between the reducing agent and the oxidized luminophore occurs. Provided the enthalpy of the reaction is greater than the excitation energy, the excited state of the luminophore results from reaction (12). After the emission reaction [(8)], the

$$C_2O_4^- \cdot \longrightarrow CO_2 + CO_2^- \cdot \tag{11}$$

$$Ru(bpy)_3^{3+} + CO_2^- \cdot \longrightarrow Ru(bpy)_3^{2+*} + CO_2 \tag{12}$$

luminophore is returned to the initial state and is able to again participate in the reaction sequence, and thus it is able to emit many photons during each measurement cycle.

Applications. Chief among the developments since the early research and discovery of electrogenerated chemiluminescence was the construction of instrumentation for detection of electrogenerated chemiluminescence. These instruments made it possible for the methodology to be used by practitioners other than electrochemists.

Measurement of the light intensity of electrogenerated chemiluminescence is very sensitive and is proportional to the luminophore concentration. Trace amounts of luminophore as low as 10^{-13} mol/liter can be detected, making electrogenerated chemiluminescence very useful in analytical and diagnostic applications.

A commercial application of the phenomenon forms the basis of a highly sensitive technique for detection of biological analytes such as deoxyribonucleic acid (DNA), ribonucleic acid (RNA), proteins, antibodies, haptens, and therapeutic drugs in the clinical laboratory. The technique combines a binding assay method and a system for detecting electrogenerated chemiluminescence. The reaction is an oxidative-reduction type of mechanism using tripropyl amine and $Ru(bpy)_3^{2+}$. The concentration of analyte from a blood or serum sample is proportional to the intensity of the light from the electrogenerated chemiluminescence reaction. The reaction is carried out with a luminophore that is specially bound to the analyte of interest. The process of specifically binding a label (that is, a luminophore) to an analyte and measuring the extent of binding to determine concentration is referred to as a binding assay. An immunoassay uses as the binding reaction an antibody-antigen reaction. By covalently bonding the luminophore to the antibody, an electrogenerated chemiluminescence–active label is produced. The binding reaction may be of many types; antibody-antigen, ligand-receptor, DNA hybridization, and others.

Relating the intensity of the electrogenerated chemiluminescence to concentration requires the integration of several types of systems: a biological binding reaction that is selective for the analyte and attaches the luminophore; an electrochemiluminescent reaction causing the luminophore to emit light;

and an instrument which integrates the electrochemical cell, light measurement system, and means for sample handling and preparation.

Separation scientists requiring more sensitive detection system capabilities for chromatography have used electrogenerated chemiluminescence as a substitute for more traditional methods of detection, such as fluorescence and spectrophotometry.

Jonathan K. Leland; Laurette Nacamulli; Hongjun Yang

Bibliography. A. J. Bard (ed.), *Electroanalytical Chemistry*, vol. 1, 1966, vol. 10, 1977; A. J. Bard and L. R. Faulkner, *Electrochemical Methods*, 1980; A. K. Campbell, *Chemiluminescence: Principles and Applications in Biology and Medicine*, 1988; P. E. Stanley and L. J. Kricka (eds.), *Bioluminescence and Chemiluminescence*, 1991.

Chemiosmosis

The coupling of metabolic and light energy to the performance of transmembrane work through the intermediary of electroosmotic gradients. Processes include synthesis of adenosine triphosphate (ATP) by oxidative phosphorylation or by photosynthesis, production of heat, accumulation of small molecules by active transport, movement of bacterial flagella, uptake of deoxyribonucleic acid (DNA) during bacterial conjugation, genetic transformation and bacteriophage infection, and insertion or secretion of proteins into or through membranes.

Oxidative phosphorylation. During the 1940s and 1950s, mitochondria were shown to be the powerhouses of the eukaryotic cell and the site of synthesis of ATP by oxidative phosphorylation. In the oxidation portion of ATP synthesis, reductants, such as reduced nicotinamide adenine dinucleotide (NADH) and succinate, are generated during metabolism of carbohydrates, lipids, and protein. These compounds are oxidized through the series of redox reactions performed by membrane-bound complexes, called electron transport or respiratory chains. The respiratory chains have common features. All are found associated with the inner mitochondrial membrane, and each chain pumps protons from inside the mitochondrion to the cytosol. Each begins with a membrane-bound dehydrogenase enzyme which oxidizes the soluble reductant. The dehydrogenase complex transfers electrons to a lipid-soluble quinone coenzyme, such as ubiquinone, which in turn passes electrons to a series of cytochromes. Cytochromes are proteins which have a heme moiety. Unlike the oxygen-carrier protein hemoglobin, which maintains its heme iron in the reduced state, the irons of cytochromes undergo a cycle of oxidation-reduction, alternating between the ferrous and ferric states. During reduction a cytochrome accepts an electron from the reduced side of the chain and passes it to the cytochrome on the oxidized side of the chain. The terminal cytochrome interacts with molecular oxygen, forming water. The overall reaction for the mitochondrial NADH oxidase is expressed in (1). The reaction

$$\mathrm{NADH} + \mathrm{H}^+ + {}^1/_2\mathrm{O}_2 \longrightarrow \mathrm{NAD}^+ + \mathrm{H}_2\mathrm{O} \qquad (1)$$

is exergonic, with a standard free-energy change ($\Delta G^{\circ\prime}$) at pH 7 of -52.6 kcal/mol. *See* CYTOCHROME; HEMOGLOBIN; MITOCHONDRIA.

The phosphorylation half of oxidative phosphorylation is catalyzed by a specific H^+-translocating adenosine triphosphatase (ATPase) enzyme. The ATPase is also a membrane-bound enzyme located in the inner mitochrondrial membrane, catalyzing reaction (2), where ADP is adenosine diphosphate and P_i

$$\mathrm{ADP} + \mathrm{P}_i + \mathrm{H}^+ \longrightarrow \mathrm{ATP} + \mathrm{H}_2\mathrm{O} \qquad (2)$$

is inorganic phosphate. The reaction occurs with a standard free-energy change, $\Delta G^{\circ\prime}$, of $+7.3$ kcal/mol. The two processes, oxidation and phosphorylation, are metabolically coupled, so that mitochondria continually synthesize ATP by using energy derived from the respiratory chain. In the case of NADH oxidase, the net reaction is shown as (3). In the overall biological reaction, three ATP's are synthesized for each NADH reduced, even though enough energy was available for six or seven ATP's. This assumes that the actual free-energy change, ΔG, is approximately the same as the standard free-energy change at pH 7, $\Delta G^{\circ\prime}$. Calculations suggest that the free energy of formation of ATP is probably greater than -7.3 kcal/mol, perhaps as much as -10 to -15 kcal/mol, but it is difficult to determine this parameter in a living system. The stoichiometry is a function of the molecular mechanism of coupling.

$$\mathrm{NADH} + {}^1/_2\mathrm{O}_2 + 3\mathrm{ADP} + 3\mathrm{P}_i + 4\mathrm{H}^+ \longrightarrow \mathrm{NAD}^+ + 3\mathrm{ATP} + 4\mathrm{H}_2\mathrm{O} \qquad (3)$$

Most early postulates about the mechanism by which the redox energy of NADH is converted into the energy-rich phosphoric anhydride bond of ATP involved the formation of a high-energy chemical intermediate during the flow of electrons down the respiratory chain. This chemical intermediate could then interact with ADP, forming a high-energy nucleotide which could finally interact with P_i, producing ATP. For two decades this chemical intermediate proved elusive. The search culminated in the formulation of the chemiosmotic hypothesis by Peter Mitchell in 1961. For this contribution Mitchell was awarded the Nobel prize in medicine in 1978.

Mitchell postulated that chemical energy, that is, the energy of a chemical bond, can be reversibly converted into electroosmotic energy. The reactions are catalyzed by enzyme complexes which function as primary ion pumps. From the work of Mitchell and of others, the respiratory chain and the ATPase have both been shown to be proton pumps. The respiratory chain transports protons out of the mitochondrion into the cell cytosol, creating an electrochemical gradient of protons. When the ATPase hydrolyzes ATP, it also pumps protons out of the mitochondrion, hence the name H^+-translocating ATPase. Since the free energy produced by respiration is so

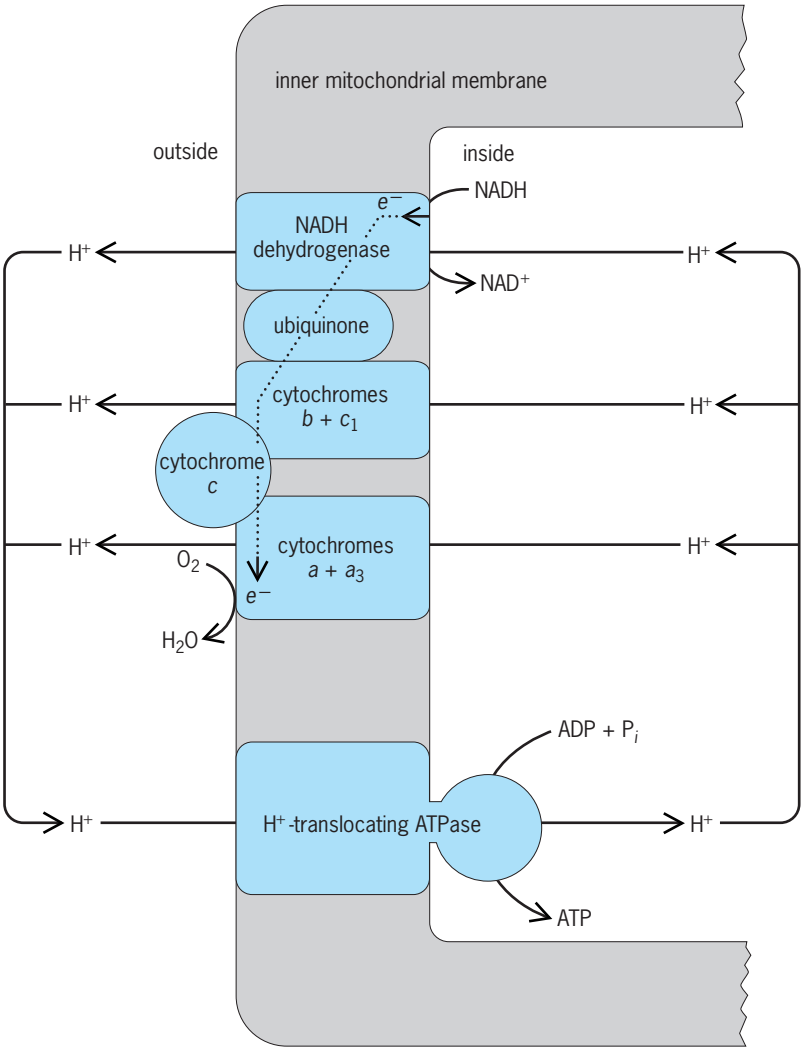

Mitochondrial oxidative phosphorylation. The chemiosmotic proton circuit of the inner mitochondrial membrane is depicted. NADH oxidase is composed of three proton-translocating units connected in series to form an electron transport chain. Electrons flow down a redox potential gradient from NADH to oxygen (broken line). Protons are pumped up an electrochemical gradient from inside to outside the mitochondrion, generating a proton motive force. Uptake of protons by the proton-translocating ATPase couples the phosphorylation of ADP to ATP.

much greater than that derived from ATP hydrolysis, proton translocation by the respiratory chain creates a force which drives the uptake of protons by the ATPase and consequently the synthesis of ATP. A diagram of the coupling of oxidation and phosphorylation is shown in the **illustration**.

Mitchell termed the force derived from proton translocation the proton motive force (pmf), by analogy with the electromotive force (emf) of electrochemical cells. Since the pmf is produced by pumping of H^+ ions across the mitochondrial membrane, it has both electrical and chemical components. The electrical or membrane potential ($\Delta\psi$) is due to the separation of positive and negative charges across the mitochondrial membrane. The chemical component results from the separation of H^+ and OH^-. Since $pH = -\log [H^+]$, $[H^+]$ being the concentration of H^+ ions, the chemical proton gradient is measured as a pH difference, ΔpH. Thus, $pmf = \Delta\psi - (2.3RT/F) \cdot \Delta pH$ and $pmf = \Delta G/F$, where R is the gas constant, T is the temperature, and F is the Faraday constant.

The concept of chemiosmosis is now well accepted, although some still question whether the pmf is the primary source of energy for phosphorylation. Most effort is concentrated on the molecular mechanisms of chemiosmotic coupling. It is clear that respiratory chains pump out protons, but questions such as which components directly transport protons, how many protons are translocated per pair of electrons, and what is the catalytic mechanism of proton translocation are still not adequately answered. It is clear that the ATPase transports protons outward during ATP hydrolysis and inward during synthesis, but the number of protons per ATP is controversial. The protons may directly interact with substrate within an active site, participating directly in catalysis. Conversely, binding of protons may change the shape or conformation of the ATPase enzyme in such a way as to alter its affinity for ATP. These are not trivial distinctions. How these macromolecular complexes work, how chemical energy is transformed into electrochemical and back into chemical energy, cannot be understood without the answers to the more basic questions. *See* ADENOSINE TRIPHOSPHATE (ATP); BIOLOGICAL OXIDATION.

Bacterial oxidative phosphorylation. Bacteria do not contain mitochondria, but many of the functions of the mitochondrial membrane are carried out by the bacterial cytoplasmic membrane. All bacteria contain a H^+-translocating ATPase in their cytoplasmic membrane. This enzyme is nearly identical to that of mitochondria. In fact, the genes for the bacterial and mitochondrial versions are clearly derived from a single evolutionary precursor. Many bacteria also use respiratory chains. Oxygen is not always the terminal electron acceptor of those chains. In some, nitrate or other oxidants can serve this role. But all of the chains are proton pumps. Thus, oxidative phosphorylation in bacteria is also chemiosmotic in nature. This resemblance to mitochondria is more than chance. The evidence, although mostly circumstantial, suggests that mitochondria, chloroplasts, and perhaps other eukaryotic organelles were originally free-living bacteria. These bacteria and larger proto-eukaryotic cells became mutually symbiotic, so that neither was complete or viable without the other. The animal and plant kingdoms arose from these endosymbiotic events. *See* BACTERIAL PHYSIOLOGY AND METABOLISM.

Photosynthesis. Photosynthesis is the conversion of light energy into chemical energy. Overall photosynthetic bacteria and the chloroplasts of eukaryotic plants capture sunlight or other light and use that energy to generate both ATP and a reductant for use in biosynthesis. The mechanism of photophosphorylation, that is, the use of light energy to drive the phosphorylation of ADP to ATP, resembles that of oxidative phosphorylation. Chloroplasts use a H^+-translocating ATPase for phosphorylation. Again, the genes for the chloroplast ATPase are derived from the same evolutionary precursor as the mitochondrial and bacterial genes. Electron transport chains

are also components of membranes found within chloroplasts. These chains have components similar to those of mitochondria, but the direction of electron flow is reversed. Water is oxidized to produce oxygen. Nicotinamide adenine dinucleotide phosphate ($NADP^+$) is reduced to NADPH, which is used for CO_2 fixation and the synthesis of carbohydrates. The energy to drive these endergonic reactions is derived from light through the absorption of photons by chlorophyll-containing photosystems. The chloroplast electron transport chain comprises a proton pump, generating a pmf which drives the phosphorylation of ADP to ATP by the ATPase. *See* PHOTOSYNTHESIS.

Thermogenesis. Not all of the energy of the proton motive force is coupled productively. It was noted above that enough chemiosmotic energy is generated by respiration to allow for phosphorylation of six or seven ATP. Yet, at best only three are produced. If not coupled to the performance of work, the excess energy may be lost as heat. In some cases, heat production or thermogenesis through the dissipation of the pmf is regulated. Some mammals, for example, cold-adapted animals, hibernating animals, and infants of hairless species (including humans), have a specialized form of adipose tissue called brown fat. The inner membrane of brown-fat mitochondria has unique proton channels which open or close depending on the need for heat. When open, the channels allow for the return of protons extruded by respiration. This cycling of protons uncouples the pmf from ATP synthesis. The energy from proton cycling becomes heat energy. By controlling the rate of the cycle, animals can regulate body temperature. *See* ADIPOSE TISSUE; THERMOREGULATION.

Other chemiosmotic systems. Oxidative phosphorylation and photophosphorylation are but specialized examples of chemiosmotic energy coupling. In the broader sense, the pmf is an ion current analogous to an electric current. Electric generators are machines which convert stored energy into a current of electrons. Among the forms of useful energy are chemical energy, such as that derived from fossil fuels, and light energy in the case of solar cells. Electricity is transmitted to motors, which couple electrical energy to the performance of work. Bacterial cells, mitochondria, and chloroplasts have protonic generators and protonic motors. Respiratory and photosynthetic electron transport chains are generators of proton currents pmf's, which then drive the various motors of the cell or organelle. When the H^+-translocating ATPase is "plugged in," the proton current drives phosphorylation. There are other motors present in the cell. Most membranes contain specific transport systems for small molecules, such as ions, sugars, and amino acids. Many of these transport systems are protonic; that is, they use the energy of the pmf to drive the accumulation or extrusion of their substrate. To give one example, in many bacteria the transport system for the sugar lactose consists of a single membrane protein which catalyzes the simultaneous co-transport of both lactose and a proton. Entry of protons through this carrier protein is driven by the pmf. Since a molecule of lactose enters the cell with each proton, entry and accumulation of lactose is thus coupled to the pmf. *See* BIOPOTENTIALS AND IONIC CURRENTS.

Other metabolic events have been found to use the pmf. Flagellar rotation in bacteria is coupled to the pmf. The flagellum is a motor which uses a flow of protons to do mechanical work, just as a paddle wheel uses a flow of water to turn it. Another example of a membrane event which uses the pmf is the uptake of DNA by bacterial cells. Foreign DNA can enter cells by sexual conjugation, by direct uptake (transformation), or by injection of viral DNA by bacteriophage. Similarly, some proteins destined to be secreted by cells require a pmf to be inserted into, or translocated through, the cytoplasmic membrane. Exactly how the pmf functions in the transport of DNA or protein is not clear. In conclusion, chemiosmosis is a universal phenomenon involved in many membrane and transmembrane events. *See* BACTERIAL GENETICS; BACTERIOPHAGE; CELL MEMBRANES.

Barry Rosen

Bibliography. P. Mitchell, Keilin's respiratory chain concept and its chemiosmotic consequences (Nobel Prize lecture), *Science*, 260:1148–1159, 1979; D. G. Nicholls, *Bioenergetics: An Introduction to the Chemiosmotic Theory*, 1982; B. P. Rosen (ed.), *Bacterial Transport*, 1978.

Chemistry

The science that embraces the properties, composition, and structure of matter, the changes in structure and composition that matter undergoes, and the accompanying energy changes. It is important to distinguish chemical change, implicit in this definition, and changes in physical form. An example of the latter is the conversion of liquid water to solid or gas by cooling or heating; the water substance is unchanged. In chemical change, such as the rusting of iron, the metal is consumed as it reacts with air in the presence of water to form the new substance, iron oxide.

Modern chemistry grew out of the alchemy of the Middle Ages, and the attempts to transmute base metals into gold. Seminal observations were made in the early eighteenth century on the changes in volume of air during combustion in a closed vessel, and the French chemist Antoine Lavoisier in the 1770s interpreted these phenomena in essentially modern terms.

Atoms and elements. Underlying all of chemistry is the concept of elementary units of matter which cannot be subdivided. This idea was adumbrated in classical Greek writings, and was clearly expressed by the Englishman John Dalton in 1803, who called these units atoms. Different kinds of atoms were recognized, each corresponding to one of the chemical elements such as oxygen, sulfur, tin, iron, and a few other metals. By the mid-nineteenth century, about 80 elements had been characterized, and these were organized on the basis of regularities in behavior and

properties, into a periodic table. *See* ELEMENT (CHEMISTRY); PERIODIC TABLE.

In the early twentieth century, observations of radiation from various sources and its impact on solid targets led to the recognition of three fundamental particles that are common to all elements; the electron, with negative charge; the proton, with positive charge; and the neutron, with zero charge. An atom consists of a nucleus containing protons and neutrons, and a diffuse cloud of electrons, equal in number to the number of protons and arranged in orbitals of progressively higher energy levels as the distance from the nucleus increases. The atomic number of an element (Z) is defined as the number of protons in the nucleus; this is the sequence of ordering in the periodic table. The mass number corresponds to the total number of protons and neutrons. *See* ATOMIC NUMBER; ELECTRON; NEUTRON; PROTON.

Isotopes. Most elements exist as isotopes, which have differing numbers of neutrons. All isotopes of an element exhibit the same chemical behavior, although isotopes can be separated on the basis of differences in atomic mass. The known elements as of mid-2000 total 115; of these, 97 have been detected in one or more isotopic forms in the Earth's crust. The other elements, including all but one of those with atomic number above 92, are synthetic isotopes produced in nuclear reactions that take place in nuclear piles or particle accelerators. Most of the isotopes of these heavier elements and also some lighter ones are radioactive; that is, the nuclei are unstable and decay, resulting in the emission of radiation. *See* ISOTOPE; PARTICLE ACCELERATOR; RADIOACTIVITY.

Molecules and chemical reactions. Molecules are combinations of two or more atoms, bonded together in definite proportions and specific geometric arrangements. These entities are chemical compounds; a molecule is the smallest unit. The bonding of atoms in compounds involves the distribution of electrons, and is the central concern of chemistry.

Compounds result from chemical reactions of atoms or molecules. The process involves formation and breaking of bonds, and may be either exothermic, in which the net bond charges lead to a more stable (lower-energy) system and heat is evolved, or endothermic, in which energy must be added to overcome a net loss of bonding energy.

A simple case is the reaction of hydrogen and oxygen to give water, which can be expressed as reaction (1).

$$2H_2 + O_2 \rightarrow 2H_2O \qquad (1)$$
$$\Delta H = 572 \text{ kilojoules}$$

The equation is balanced; no atoms are gained or lost in a chemical reaction. The symbols represent the nature of the initial and final materials and also the relative amounts. Thus H_2O represents a molecule of water or a mole, which is the quantity in grams (or other mass units) equivalent to the molecular weight. The symbol ΔH indicates the energy (enthalpy) change for the process. The reaction of hydrogen and oxygen is highly exothermic, and the sign of the energy charge is therefore negative since the system has lost heat to the surroundings. *See* ENTHALPY; MOLE (CHEMISTRY); STOICHIOMETRY.

Bonds. Bonds can be broadly classified as ionic or covalent. An ion is an atom or molecule which has an electric charge. Ionic compounds can be illustrated by salts such as sodium chloride, NaCl, in which a positive sodium ion, Na^+, and negative chloride ion, Cl^-, are associated by electrostatic attractions in regular locations of a crystal lattice. In solution the ions are solvated by water molecules and can conduct an electric current.

In covalent molecules, bonds are formed by the presence of pairs of electrons in overlapping orbitals between two atoms. Thus when two hydrogen atoms (H·) come within bonding distance, a molecule of hydrogen is formed in an exothermic reaction, by formation of a covalent bond. In this case the heat of reaction represents the energy of the H-H bond [reaction (2)].

$$2H\cdot \rightarrow H{-}H \qquad (2)$$
$$\Delta H = -435 \text{ kilojoules}$$

See CHEMICAL BONDING.

Chemical compounds. A compound is specified by the elements it contains, the number of atoms of each element, the bonding arrangement, and the characteristic properties. The number of unique compounds that have been isolated from natural sources or prepared by synthesis is enormous; as of 2000, over 15 million substances were registered in the file maintained by Chemical Abstracts (American Chemical Society). Most of these are organic compounds, containing from a few to many hundred carbon atoms. The element carbon, unlike any others, can form long chains of covalently bonded atoms. Moreover, there can be many compounds, called isomers, with the same atomic composition. Thus a molecular formula such as $C_8H_{16}O$ can represent many thousand different compounds. *See* CARBON; MOLECULAR ISOMERISM.

Branches. Traditionally, five main subdivisions are designated for the activities, professional organizations, and literature of chemistry and chemists.

Analytical chemistry. This subdivision is an overarching discipline dealing with determination of the composition of matter and the amount of each component in mixtures of any kind. Analytical measurements are an integral and indispensable part of all chemical endeavor. Originally, analytical chemistry involved detection, separation, and weighing of the substances present in a mixture. Determination of the atomic ratio and thence the molecular formula of a compound is a prerequisite for any other investigation; the development of balances and techniques for doing this on milligram quantities of material had an enormous impact on organic chemistry. Advances since the 1950s have involved increasingly sophisticated instrumentation; mass spectrometers are a notable example. Other important methods include high-resolution chromatography and various applications of electrochemistry. A constant goal in

analytical chemistry is the development of methods and instruments of greater sensitivity. It is now possible to detect trace compounds such as environmental pollutants at the picogram level. *See* ANALYTICAL CHEMISTRY.

Biochemistry. Biochemistry is the study of living systems from a chemical viewpoint; thus it is concerned with the compounds and reactions that occur in plant and animal cells. Most of the substances in living tissues, including carbohydrates, lipids, proteins, nucleic acids, and hormones, are well-defined organic substances. However, the metabolic and regulatory processes of these compounds and their biological function are the special province of biochemistry. One of the major areas is the characterization of enzymes and their cofactors, and the mechanism of enzyme catalysis. Other topics of interest include the transport of ions and molecules across cell membranes, and the target sites of neurotransmitters and other regulatory molecules. Biochemical methods and thinking have contributed extensively to the fields of endocrinology, genetics, immunology, and virology. *See* BIOCHEMISTRY.

Inorganic chemistry. This discipline is concerned with any material in which metals and metalloid elements are of primary interest. Inorganic chemistry is therefore concerned with the structure, synthesis, and bonding of a very diverse range of compounds. One of the early interests was the composition of minerals and the discovery of new elements; from this has grown the specialized area of geochemistry. Early synthetic work emphasized compounds of the main group elements, and particularly in this century, complex compounds of the transition metals. These studies have led to soluble transition-metal catalysts, and a greatly increased understanding of catalytic processes and the pivotal role of metal atoms in major biochemical processes, such as oxygen transport in blood, photosynthesis, and biological nitrogen fixation. Other contributions of inorganic chemistry are seen in advanced ceramics, high-performance composite materials, and the growing number of high-temperature superconductors. *See* CATALYSIS; CERAMICS; COMPOSITE MATERIAL; INORGANIC CHEMISTRY; PHOTOSYNTHESIS.

Organic chemistry. This subdivision is centered on compounds of carbon. Originally these were the compounds isolated from plant and animal sources, but the term was early broadened to include all compounds in which a linear or cyclic carbon chain is the main feature. Two of the major thrusts have been the elucidation of new structures and their preparation by synthesis; another long-standing interest has been study of the reaction mechanisms and rearrangments of organic compounds. Structure work on naturally occurring compounds progressed over a 150-year period from simple straight-chain compounds with 2–10 carbon atoms, hydrogen, and 1 or 2 oxygen atoms to antibiotics and toxins with many rings and as many as 100 carbon atoms. In modern work, nuclear magnetic resonance spectroscopy and x-ray diffraction have become indispensable tools. Paralleling structural studies has been the synthesis of increasingly complex target molecules. Synthetic work is directed also to the preparation of large numbers of compounds for screening as potential drugs and agricultural chemicals. Plastics, synthetic fibers, and other high polymers are other products of organic chemistry. *See* NUCLEAR MAGNETIC RESONANCE (NMR); ORGANIC CHEMISTRY; X-RAY DIFFRACTION.

Physical chemistry. This discipline deals with the interpretation of chemical phenomena and the underlying physical processes. One of the classical topics of physical chemistry involves the thermodynamic and kinetic principles that govern chemical reactions. Another is a description of the physical states of matter in molecular terms. Experimentation and theoretical analysis have been directed to the understanding of equilibria, solution behavior, electrolysis, and surface phenomena. One of the major contributions has been quantum chemistry, and the applications and insights that it has provided. The methods and instruments of physical chemistry, including such hardware as spectrometers and magnetic resonance and diffraction instruments, are an integral part of every other area. *See* CHEMICAL THERMODYNAMICS; PHYSICAL CHEMISTRY; QUANTUM CHEMISTRY.

Others. Each broad area of chemistry embraces many specialized topics. There are also a number of hybrid areas, such as bioorganic and bioinorganic chemistry, analytical biochemistry, and physical organic chemistry. Each of these areas has borrowed extensively from and contributed to every other one. It is better to view chemistry as a seamless web, rather than a series of compartments. *See* BIOINORGANIC CHEMISTRY.

James A. Moore

Chemometrics

A chemical discipline that uses mathematical and statistical methods to design or select optimal measurement procedures and experiments and to provide maximum chemical information by analyzing chemical data. Chemometrics is actually a collection of procedures, mathematics, and statistics that can help chemists perform well-designed experiments and proceed rapidly from data, to information, to knowledge of chemical systems and processes.

Medicinal chemists use chemometrics to relate measured or calculated properties of candidate drug molecules to their biological function; this subdiscipline is known as quantitative structure activity relations (QSAR). Environmental chemists use chemometrics to find pollution sources or understand the effect of point pollution sources on regional or global ecosystems by analyzing masses of environmental data. Forensic chemists analyze chemical measurements made on evidence (for example, gasoline in an arson case) or contraband to determine its source. Experimental physical chemists use chemometrics to unravel and identify physical or chemical states from spectral data acquired during the course of an experiment. *See* FORENSIC CHEMISTRY; PHYSICAL CHEMISTRY.

In analytical chemistry, chemometrics has seen rapid growth and widespread application, primarily due to the computerization of analytical instrumentation. Automation provides an opportunity to acquire enormous amounts of data on chemical systems. Virtually every branch of analytical chemistry has been impacted significantly by chemometrics; commercial software implementing chemometrics methods has become commonplace in analytical instruments. *See* ANALYTICAL CHEMISTRY.

Whether the analyst is concerned with a single sample or, as in process analytical chemistry, an entire chemical process (for example, the human body, a manufacturing process, or an ecosystem), chemometrics can assist in the experimental design, instrument response, optimization, standardization, and calibration as well as in the various steps involved in going from measurements (data), to chemical information, to knowledge of the chemical system under study.

Tensorial calibration. Chemical processes are usually controlled by acquiring information on physical properties (such as temperature, pressure, and flow rate) and chemical properties (composition) at one or more locations in a chemical reactor. These data are then fed to process control computers that open or close valves and provide heating or cooling of the process to optimize yield and quality of the product. Sensors that measure product in terms of intermediate or starting material composition are calibrated by using pure standards. One problem often encountered is that the sensor may be sensitive to an unsuspected interferent in the process, causing the reported composition data to be erroneous. Simple sensors generating only one measurement or datum can be classified as zero-order sensors and cannot separate the signal generated by the analyte from the interferent. This is a serious problem in process control.

First order. To solve the problem of separating the signals, chemists have turned to multivariate calibration using first-order instruments. These are instruments capable of producing a vector of measurements, a first-order tensor, per sample and include spectrometers and arrays of nonselective zero-order sensors. For the calibration of such instruments, methods such as classical multilinear regression (MLR), principal components regression (PCR), and partial least-squares regression (PLS) have become popular, especially in infrared and near-infrared spectrophotometry. Besides being able to analyze for multiple analytes in one sample while correcting for interferences, the methods have also been used to predict physical properties from spectral data. While these methods correct for interferences, the interferents must be included in the calibration step according to a good statistical experimental design. If an uncalibrated interferent is present in a sample to be analyzed, these methods can detect its presence but cannot, in general, correct for its effect, and therefore the results are biased.

First-order calibration is used in many areas of process control. The physical properties of polymers, such as tensile strength and the ability to incorporate dyes for color, can be predicted from on-line spectral data acquired during a polymerization process before the process is complete. Another important application involves the use of on-line fiber-optic-based near-infrared spectrometers to predict the octane number of gasoline at various stages of refining and blending. Accurate knowledge of fuel properties during production can save millions of dollars per year at a large refinery. A multivariate calibration method is used to relate the spectral data from calibration samples to one or more known performance properties. If the calibration model passes various statistical tests, it can be used in the refinery. *See* PROCESS CONTROL.

Second order. Quite often the relationship between the analytical measurements and the analyte concentration, or physical property, is nonlinear. Methods are available for calibration and analysis in these cases, for example locally weighted regression, nonlinear partial least squares, and neural networks. *See* LEAST-SQUARES METHOD; NEURAL NETWORK.

A common problem in environmental science is determining the concentrations of pollutants in very complex samples containing completely unknown components. For example, in order to determine the migration of toxic chemicals in the ground near a known or suspected dump site, environmental analytical chemists sample ground and subsurface water from monitoring wells located around the site. Environmental analyzers have been developed that can be lowered down the wells for in-place analysis in order to avoid the time, expense, and other problems associated with taking hundreds of samples to the laboratory. The in-place analyzers must be calibrated, and the calibration models must be capable of accurate determination of the toxic chemicals, even though what may be present in the wells is not fully known. When an unknown interferent is present, it must be physically separated from the analyte by chromatography; or the analyst can move up to second-order calibration using an instrument capable of acquiring a matrix of data, second-order tensor, per sample. These instruments involve all so-called hyphenated methods, for example, gas chromatography–mass spectrometry (GC/MS) and liquid chromatography–ultraviolet, and include as well tandem mass spectrometers (MS/MS), two-dimensional nuclear magnetic resonance spectrometers (2D-NMR), and emission-excitation fluorescence spectrometers. Methods to analyze these data include the generalized rank annihilation method (GRAM) and residual bilinearization; they come in a variety of algorithms according to the mathematical characteristics (for example, bilinearity) of the data. The most important benefit associated with second-order calibration (the second-order advantage) is the possibility to analyze for analytes in the presence of unknown interferents. Also, under certain conditions these methods yield qualitative information about sample components as well as quantitative information. This characteristic is particularly important for environmental or medical applications and

assures the analyst that the analyzer is operational. *See* CALIBRATION; CHROMATOGRAPHY.

Standardization. The composition of many agricultural and food products is determined by spectrometers. For example, near-infrared reflectance (NIR) is commonly used to determine the moisture, protein, and fat concentration in wheat or other grains in farm fields. These data are useful for determining harvest time, irrigation, and fertilizer requirements and for setting the value of the product. Either the spectrometers must be calibrated individually or, more desirably, a reference instrument can be calibrated and the calibration model then used with several other similar instruments. Unfortunately, the latter is problematic since no two analytical instruments are exactly alike, even if they come from the same company and have the same model numbers. Also, a single instrument's response to a standard reference sample will change over time because of subtle changes in its components. To circumvent these and similar problems, chemometricians have developed the topic of instrument standardization. This is fundamentally different than calibration, which is used to relate instrument response to the property of interest. The goal of standardization is to correct the response of an instrument to be in accord with another instrument or with itself over the time it is in service.

Standardization can be as simple as using a stable wavelength reference such as a laser line to correct for shifts in wavelength that can invalidate calibration and produce incorrect analyses. There are multivariate instrument standardization methods that use multiple standardization or transfer samples to calculate the parameters of a model that can, for example, correct the response obtained from one instrument so that the response is like that obtained from another instrument. This approach can facilitate the use of a calibration model among several instruments. In process analytical chemistry, this is a major advantage, since a calibration model developed on one instrument, perhaps a high-quality laboratory spectrometer, can be transferred to several process analyzers; thus the need to calibrate each one is avoided, often a very costly process. Also, as the process analyzers change over time (for example, drift), they can be restandardized to generate the same response as when they were first put into service. *See* INSTRUMENT SCIENCE.

Curve resolution. Chromatography is widely used in product testing and quality control, especially in the pharmaceutical industry. As the components of a sample are separated in a chromatogram, it is imperative that the purity of each component be assured in order to avoid having a contaminant in the final product. To help solve this problem, chromatographers employ rapid scanning or diode array spectrometers as detectors so that the spectrum of each component eluting can be checked for purity.

When two analytes coelute near each other during a chromatographic separation of a mixture sample, a single peak in the chromatogram may result. If the analyst has previous knowledge of the shape of a pure analyte elution profile, it may be possible to resolve the mixture peak into the two pure analyte peaks by using, for example, Fourier deconvolution. Fourier deconvolution uses the fast Fourier transform algorithm to mathematically extract a curve with a known shape from a more complex mixture curve, itself composed of a number of curves.

If a spectrometer is used as the detector, multivariate curve resolution methods based on principal components analysis, factor analysis, or singular value decomposition can be used to resolve the mixtures into their component parts, thereby detecting the presence of a product contaminant. *See* CURVE FITTING; FACTOR ANALYSIS.

Pattern recognition. Perhaps the oldest activity in modern chemometrics is the application of data analysis methods developed in computer science and electrical engineering under the heading of pattern recognition to chemical data. As with most of the multivariate methods mentioned above, a pattern recognition application begins when a series of measurements is made on a collection of samples called the training set. For example, the concentrations of a number of trace elements may be measured on several quartzite samples acquired from three known quarries involved in an archeological study. One goal would be to determine if the pattern of trace-element concentrations can be used to distinguish the samples from a given quarry from all of the others. If this training step is successful, the rule or model can be used to classify samples taken from statues or monuments in order to determine their origin. Several studies of this type have been carried out in archeology, where trace-element patterns in antiquities (such as statues or arrowheads) have been traced back to their origins in order to study trade routes or cultural diversity.

Many similar studies have been conducted in forensic chemistry, drug design, screening, taxonomy, and product quality determination. In drug design, structural, chemical, or physical properties are used to classify previously untested drugs as active or inactive in a particular drug screening test. The chemometric methods used are called multivariate classification or supervised learning methods, and they are the same as those used in archeological studies. If the methods are successful, they can be used to predict the activity of untested candidate drug compounds and even to indicate which of the structural, chemical, or physical properties are the most related to biological activity, giving the medicinal chemist information useful for finding more potent drugs.

A medicinal chemist may want to subject a subset of a large set of chemical compounds to a biological screening test. The subset would be selected to represent the structural diversity of the full set as much as possible. Based on structural properties available for each compound, cluster analysis (a type of pattern recognition) might determine that the many compounds can be arranged in, say, 20 groups or clusters, with the compounds belonging to any one group being similar to its class members. The chemist

need only test one compound per group, thereby saving considerable expense but getting a good coverage of available compounds.

Other areas. There are several other areas of chemometrics research. Experimental design is a well-developed topic in statistics that chemists use to get the maximum amount of pertinent information from the smallest investment of time, money, or effort. Using one of several multivariate optimization methods, computers can optimally tune complex instruments to get the best possible results. Wherever experiments generating data, particularly large quantities of data, are being conducted, modern chemometrics methods can be applied to make the best measurements possible and move quickly from data to information to knowledge. *See* STATISTICS.

Bruce Kowalski

Bibliography. S. D. Brown, R. S. Bear, Jr., and T. B. Blank, Chemometrics, *Fundamental Rev.: Anal. Chem.*, 64:22R–49R, 1992; S. J. Haswell (ed.), *Practical Guide to Chemometrics*, 1992; H. Martens and T. Naes, *Multivariate Calibration*, 1992; D. L. Massart et al., *Chemometrics: A Textbook*, 1988; M. A. Sharaf, D. L. Illman, and B. R. Kowalski, *Chemometrics*, 1986.

Chemoreception

The ability of organisms to detect changes in the chemical composition of their exterior or interior environment. It is a characteristic of every living cell, from the single-celled bacteria and protozoa to the most complex multicellular organisms. Chemoreception allows organisms to maintain homeostasis, react to stimuli, and communicate with one another. *See* HOMEOSTASIS.

At the single-cell level, bacteria orient toward or avoid certain chemical stimuli (chemotaxis); algal gametes release attractants which allow sperm to find oocytes in a dilute aqueous environment; and unicellular slime molds are drawn together to form colonial fruiting bodies by use of aggregation pheromones. *See* CELLULAR ADHESION; TAXIS.

In multicellular organisms, both single cells and complex multicellular sense organs are used to homeostatically maintain body fluids (interoreceptors) as well as to monitor the external environment (exteroreceptors). The best-studied interoreceptors are perhaps the carotid body chemoreceptors of higher vertebrates, which monitor the levels of oxygen, carbon dioxide, and hydrogen ions in arterial blood. The best-studied exteroreceptors are those associated with taste (gustation) and smell (olfaction). Such receptors may occur anywhere on the bodies of different organisms, from antennae and tongues to legs and fins, and range from single hair cells in some invertebrates to the complex nasal epithelia of higher vertebrates. Internal communication is also effected by chemical means in multicellular organisms. Thus both hormonal and neural control involve the perception, by cells, of control chemicals (hormones and neurotransmitters, respectively). *See* CAROTID BODY; CHEMICAL SENSES; OLFACTION; SENSE ORGAN; TASTE; TONGUE.

Basic mechanism of action. Although tremendous diversity can be seen in the evolution of complex chemosensory structures, chemoreception is ultimately a property of single cells. The basic mechanism underlying chemoreception is the interaction of a chemical stimulus with receptor molecules in the outer membrane of a cell. These molecules are believed to be proteins which, because of their three-dimensional shapes and chemical properties, will have the right spatial and binding "fit" for interaction with only a select group of chemicals (the same basic mechanism by which enzymes are specific for various substrates). The interaction between a chemical stimulus and a receptor molecule ultimately leads to structural changes in membrane channels by mechanisms that are not totally clear. In some instances, the chemical stimulus–receptor complex activates an enzyme (adenylate cyclase) that synthesizes cyclic adenosine monophosphate (cAMP), which acts as a second messenger interacting with membrane channels and changing their conductance. In other instances, the chemical stimulus interacts with the channel protein directly, altering its conductance. In either case, the net result is usually a change in membrane conductance (permeability) to specific ions which changes both the internal chemical composition of the cell and the charge distribution across the cell membrane. In single-celled organisms, this may be sufficient to establish a membrane current which may elicit responses such as an increase or decrease in ciliary movement. In multicellular organisms, it usually results in changes in the rate of release of hormones or the stimulation of neurons. *See* CELL MEMBRANES.

Characteristics of chemoreceptors. The basic characteristics of all chemoreceptors are specificity (the chemicals that they will respond to); sensitivity (the magnitude of the response for a given chemical stimulus); and range of perception (the smallest or largest level of stimulus that the receptor can discriminate). Specificity is a consequence of the types of proteins found in the membrane of a receptor cell. Each cell will have a mosaic of different receptor molecules, and each receptor molecule will show different combinations of excitatory or inhibitory responses to different molecules. In an excitatory response, there is a net flux of positive ions into the cell (depolarization); for an inhibitory response, there is a net flux of negative ions into the cell (hyperpolarization). In single cells, or in primary sensory neurons, the receptor current will spread along the membrane by passive electrotonic means. The stronger the stimulus—that is, the more of the chemical present—the more receptors affected, the greater the change in conductance, and the larger the membrane current.

In animals with nervous systems, these changes in conductance of primary sensory cells can lead to one of two events. In some receptors, if the current is excitatory and sufficient in magnitude (threshold), an action potential will be generated at a spike-initiating zone on the neuron. Other receptors are not able

to generate action potentials but respond by releasing a neurotransmitter (the depolarizing current causes a change in membrane conductance of the transmitter substance) that acts on a second-order neuron which is excitable and therefore can generate action potentials. *See* BIOPOTENTIALS AND IONIC CURRENTS.

The sensitivity of a chemoreceptor reflects both the amount of chemical substance required to initiate a change in membrane potential or discharge of the receptor cell, and the change in potential or discharge for any given change in the level of the chemical stimulus. There are real limits as to the extent of change in membrane conductance or firing frequency. Thus, for more sensitive cells, there is a smaller range over which they can provide information about the change in concentration of any given chemical before it has reached its maximum conductance or discharge rate and has saturated.

In most chemoreceptors, this problem is reduced by a variety of mechanisms. To begin with, the relationship between stimulus intensity and neuronal discharge frequency is usually logarithmic. This means that the high-intensity end of the scale is compressed, greatly extending the range of discrimination of the receptor. Although this may seem to decrease the sensitivity of the receptor at higher stimulus intensities in absolute terms, the increment in discharge remains the same for any given percentage change in stimulus intensity. In other words, a doubling of the stimulus intensity will produce the same increase in receptor potential anywhere over the full range of sensitivity of the cell. Receptor adaptation, a decrease in receptor discharge despite a constant chemical stimulus, also extends the dynamic range over which receptors operate. Finally, in chemoreceptive organs of multicellular organisms, the interaction of chemoreceptors with different threshold levels (range fractionation) can also extend the dynamic range of sensory perception.

Chemosensory systems can be extremely sensitive. Humans can detect 1 molecule of mercaptan in 5×10^{10} molecules of air, while marine crabs have been shown to respond to as little as 10^{-17} lb (10^{-15} g) of prey tissue in 0.264 gal (1 liter) of seawater. Perhaps most impressive, however, is the ability of the antennal chemoreceptor cells of the male silkworm (*Bombyx mori*) to detect a single molecule of female pheromone. Only about 90 molecules per second impinging on a single receptor are required to alter the firing rate of a cell, and stimulation of only 40 chemoreceptors out of roughly 20,000 per antenna is sufficient to evoke a behavioral response (excited wing flapping). *See* PHEROMONE.

Chemoreceptor tuning. In animals, the responsiveness of some chemoreceptors can be either enhanced or attenuated by other neural input. These influences come in the form of efferent inputs from the central nervous system, from neighboring receptors, or even from recurrent branches of the chemoreceptor's own sensory axons. The net effect is either (1) to increase the acuity of the receptors (excitatory input brings the membrane potential of the receptor cell closer to threshold, requiring less chemical stimulus to elicit a response); or (2) to extend the range of responsiveness of the receptors (inhibitory input lowers the membrane potential of the receptor cell, requiring more chemical stimulus to bring the cell to threshold). For example, chemical sensitivity is greatly heightened in most animals when they are hungry.

Central integration of input. Any given chemoreceptor cell can have any combination of receptor proteins, each of which may respond to different chemical molecules with either an increase or a decrease in the membrane potential. Thus, chemoreceptor cells do not exhibit a unitary specificity to a single chemical substance, but rather an action spectrum to various groups of chemicals. The ability of animals to distinguish such a large number of different, complex, natural chemical stimuli resides in the ability of higher centers in the nervous system to "recognize" the pattern of discharge of large groups of cells. Sensory quality does not depend on the activation of a particular cell or group of cells but on the interaction of cells with overlapping response spectra.

Structural adaptations. Despite the common, basic mechanism underlying chemoreception in all organisms, there is a great diversity in the design of multicellular chemoreceptive organs, particularly in animals. The complex structures of most of these organs reflect adaptations that serve to filter and amplify chemical signals. Thus, the antennae in many insects, and the irrigated protective chambers, such as the olfactory bulb of fishes and nasal passages of mammals, increase the exposure of chemoreceptor cells to the environment. At the same time, they allow the diffusion distances between chemoreceptive cells and the environment to be reduced, thereby increasing acuity. In terms of filtering, they may serve to convert turbulent or dispersed stimuli into temporal patterns that can be more easily interpreted. The extent to which such structural adaptations are seen in various organisms tends to reflect the relative importance of chemoreception to the organism, which, to a large extend, reflects the habitat in which the organism lives. *See* CHEMICAL ECOLOGY.

William Milsom

Bibliography. H. Acker and R. G. O'Regan (eds.), *Physiology of the Peripheral Arterial Chemoreceptors,* 1983; V. G. Dethier, *The Hungry Fly,* 1976; R. Eckert and D. Randall, *Animal Physiology,* 4th ed., 1997; T. V. Getchell and G. M. Shepherd, Responses of olfactory receptor cells to step impulses of odour at different concentrations in the salamander, *J. Physiol. (London),* 282:521–540, 1978; G. M. Shepherd, *Neurobiology,* 3d ed., 1994.

Chemostat

An apparatus (see **illus.**) for the continuous cultivation of microorganisms, such as bacteria, yeasts, molds, and algae, or for the cultivation of plant cells. The nutrients required for cell growth are supplied continuously to the culture vessel by a pump

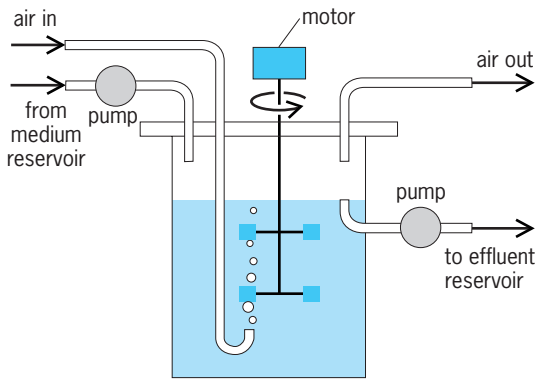

Schematic representation of chemostat apparatus.

connected to a medium reservoir. The cells in the vessel grow continuously on these nutrients. Residual nutrients and cells are removed from the vessel (fermenter) at the same rate by an overflow, thus maintaining the culture in the fermenter at a constant volume.

Parameters. An important feature of chemostat cultivation is the dilution rate, defined as the volume of nutrient medium supplied per hour divided by the volume of the culture. During chemostat cultivation, an equilibrium is established (steady state) at which the growth rate of the cells equals the dilution rate. The higher the dilution rate, the faster the organisms are allowed to grow. Above a given dilution rate the cells will not be able to grow any faster, and the culture will be washed out of the fermenter. The chemostat thus offers the opportunity to study the properties of organisms at selected growth rates. This is particularly important because in their natural environment microorganisms seldom grow at their maximum rate. Under nutrient limitation, cells may display properties which also are of great importance in a number of applications, such as industrial fermentation and wastewater treatment. *See* FERMENTATION; WATER TREATMENT.

The nutrient medium which is fed to the fermenter contains an excess of all growth factors except one, the growth-limiting nutrient. The concentration of the cells (biomass) in the fermenter is dependent on the concentration of the growth-limiting nutrient in the medium feed. Upon entering the fermenter, the growth-limiting nutrient is consumed almost to completion, and only minute amounts of it may be found in the culture and the effluent. Initially, when few cells have been inoculated in the growth vessel, even the growth-limiting nutrient is in excess. Therefore, the microorganisms can grow at a rate exceeding their rate of removal. This growth of cells causes a fall in the level of the growth-limiting nutrient, gradually leading to a lower specific growth rate of the microorganisms. Once the specific rate of growth balances the removal of cells by dilution, a steady state is established in which both the cell density and the concentration of the growth-limiting nutrient remain constant. Thus the chemostat is a tool for the cultivation of microorganisms almost indefinitely in a constant physiological state.

To achieve a steady state, parameters other than the dilution rate and culture volume must be kept constant (for example, temperature and pH). The fermenter is stirred to provide a homogeneous suspension in which all individual cells in the culture come into contact with the growth-limiting nutrient immediately. Furthermore, stirring is also required to achieve optimal distribution of air (oxygen) in the fermenter when aerobic cultures are in use.

Laboratory chemostats usually contain 0.5 to 10.5 quarts (0.5 to 10 liters) of culture, but industrial chemostat cultivation can involve volumes up to 343,000 gal (1300 m^3) for the continuous production of microbial biomass.

Advantages. The chemostat offers a number of advantages for the cultivation of cells as compared with growth in a batch culture. In the latter, a closed system, organisms grow in excess nutrients at their maximum rate, and the nutrient concentration, products, and biomass change continuously. When a nutrient becomes depleted, a rapid fall in the growth rate takes place. This often leads to the death of cells. In the chemostat, however, the constancy of all parameters allows accurate and reproducible experiments. By varying the dilution rate, within limits, the rate of growth of the organisms can be changed at will. The density of cells in the cultures can be chosen by the appropriate concentration of the growth-limiting nutrient in the medium reservoir. Depending on the question to be answered, the type of growth-limiting nutrient can be selected. Of course, possible choices are limited by the nature and capabilities of the organisms studied.

The chemostat can be used to grow microorganisms on very toxic nutrients since, when kept growth-limiting, the nutrient concentration in the culture is very low. The chemostat can be used to select mutants with a higher affinity to the growth-limiting nutrient or, in the case of a mixed population, to select the species that are optimally adapted to the growth limitation and culture conditions. The chemostat is of great use in such fields of study as physiology, ecology, and genetics of microorganisms. *See* BACTERIAL GENETICS; BACTERIAL PHYSIOLOGY AND METABOLISM; MICROBIOLOGY.

Mathematical description. The operation of a chemostat can also be described in mathematical terms. Growth of a microorganism can be described by the empirical formula of J. Monod (1942) as a function of the growth-limiting substrate, Eq. (1), in

$$\mu = \mu_{\max} \frac{C_s}{K_s + C_s} \tag{1}$$

which μ is the specific growth rate (h^{-1}), $\mu_{\max}$ is the maximum specific growth rate, C_s is the concentration of the growth-limiting substrate, and K_s is a Monod saturation constant, numerically equal to the substrate concentration at which $\mu = {}^1/_2\mu_{\max}$. The relationship between C_s and μ is a typical saturation curve. When an organism is grown in a closed system (batch culture) with, initially, excess substrate, it will grow at $\mu_{\max}$. During growth its environment

will constantly change, but if the conditions remain favorable, growth will continue until the growth-limiting compound is depleted. Near the end of growth, μ will fall because of the factors described by Eq. (1), and will finally become zero.

The chemostat can be considered as an open culture system in which fresh (sterilized) medium is introduced at a constant flow rate ϕ, and from which the culture fluid emerges at the same flow rate. At a constant volume υ and an in-flow rate ϕ, the dilution rate D is defined by Eq. (2), in which the dilution rate is expressed in h^{-1}.

$$D = \frac{\phi}{V} \tag{2}$$

Monod demonstrated that over a large range of growth rates a fixed relationship exists between the amount of (growth-limiting) substrate consumed and the amount of biomass produced, Eq. (3), in which C_x

$$\frac{dC_x}{dt} = -Y''_{sx}\frac{dC_s}{dt} \tag{3}$$

is the biomass concentration and t is time. Y''_{sx} is the yield factor and is defined as the amount (weight) of cell material produced per amount (weight, or mole) of substrate consumed. Y''_{sx} is not always a constant.

If the medium in the fermenter is inoculated (for example, with bacteria), the culture will grow at a given rate. At the same time, a quantity of bacteria will be washed out because the culture is continuously fed and diluted with fresh medium. It thus follows that accumulation of biomass in the culture is equal to growth minus washout.

The balance equation is Eq. (4). Hence if μ is

$$\frac{dC_x}{dt} = \mu C_x - DC_x = (\mu - D)C_x \tag{4}$$

greater than D, C_x will increase, while if μ is less than D, C_x will decrease. If μ equals D, Eq. (4) will be zero and an equilibrium will exist.

It can be shown mathematically that, irrespective of the starting conditions, a steady state, with μ equal to D, must inevitably be reached, provided that D does not exceed a critical value.

In order to calculate the steady-state concentrations of biomass and the growth-limiting nutrient in the culture, mass balance equations similar to Eq. (4) for the nutrients entering, being consumed, and leaving the culture must also be made. In Eq. (5), C_s is

$$\frac{dC_s}{dt} = D(C_{si} - C_s) - \frac{\mu C_x}{Y''_{sx}} \tag{5}$$

the substrate concentration in the culture, and C_{si} the concentration of the substrate entering the vessel. The net change in C_{si} per unit of time is equal to the dilution rate multiplied by the difference in substrate concentration entering and leaving the culture, minus the substrate consumed. The substrate consumed is expressed as the growth divided by the yield, Eq. (3). At steady state, both Eqs. (4) and (5) are zero. This, when combined with Eq. (1), will lead to the equilibrium concentrations $\overline{C}_x$ for biomass [Eq. (7)] and $\overline{C}_s$ for the substrate [Eq. (6)]. Here

$$\overline{C}_s = K_s(D/\mu_{\max} - D) \tag{6}$$

$$\begin{aligned}\overline{C}_x &= Y''_{sx}(C_{si} - \overline{C}_s)\\ &= Y''_{sx}\left(C_{si} - \frac{K_s \cdot D}{\mu_{\max} - D}\right)\end{aligned} \tag{7}$$

K_s, $\mu_{\max}$, and Y''_{sx} are constants for a microorganism under the specified conditions of temperature, medium composition, and the nature of the growth-limiting substrate. For C_{si} and D, any realistic constant value can be chosen. From Eq. (6) it appears that $\overline{C}_s$ solely depends on D. Equation (7) shows that $\overline{C}_x$ depends on D and C_{si} and is proportional to C_{si} if $\overline{C}_s \ll C_{si}$, which is usually the case. If K_s, $\mu_{\max}$, and Y''_{sx} are known for a given microgranism, the relationship between C_x or C_s, and D can be predicted at a chosen C_{si}. J. Gijskuenen

Bibliography. R. G. Burns and J. H. Slater (eds.), *Experimental Microbial Ecology*, 1982; D. Herbert, R. Elsworth, and R. C. Telling, The continuous culture of bacteria: A theoretical and experimental study, *J. Gen. Microbiol.*, 14:601–622, 1956; J. R. Norris and D. W. Ribbonds (eds.), *Methods in Microbiology*, vol. 013, 1980; S. J. Pirt, *Principles of Microbe and Cell Cultivation*, 1975; H. Veldkamp, *Continuous Culture in Microbial Physiology and Ecology*, 1976.

Chemostratigraphy

A subdiscipline of stratigraphy and geochemistry that involves correlation and dating of marine sediments and sedimentary rocks through the use of trace-element concentrations, molecular fossils, and certain isotopic ratios that can be measured on components of the rocks. The isotopes used in chemostratigraphy can be divided into three classes: radiogenic (strontium, neodymium, osmium), radioactive (radiocarbon, uranium, thorium, lead), and stable (oxygen, carbon, sulfur). Trace-element concentrations (that is, metals such as nickel, copper, molybdenum, and vanadium) and certain organic molecules (called biological markers or biomarkers) are also employed in chemostratigraphy. *See* DATING METHODS; ROCK AGE DETERMINATION.

Radiogenic isotope stratigraphy. Radiogenic isotopes are formed by the radioactive decay of a parent isotope to a stable daughter isotope. The application of these isotopes in stratigraphy is based on natural cycles of the isotopic composition of elements dissolved in ocean water, cycles which are recorded in the sedimentary rocks. The element commonly used is strontium, but osmium and neodymium isotopes are used as well. *See* ISOTOPE; RADIOISOTOPE.

Strontium (Sr) has four naturally occurring, stable isotopes, with mass numbers (sum of protons and neutrons) 88, 87, 86, and 84. Strontium-87 is radiogenic, formed by the radioactive decay of rubidium-87. Strontium replaces calcium in calcium carbonate and other calcium-bearing minerals (such as aragonite, apatite, and plagioclase). Calcite typically

contains 500–2000 parts per million strontium, and aragonite (another form of calcium carbonate) contains 2000–9000 ppm. Strontium is the major cation in strontianite ($SrCO_3$) and celestite ($SrSO_4$), but its principal occurrence in minerals is as a trace element.

The ocean is a large reservoir of dissolved strontium. The isotopic composition of strontium in the oceans reflects the isotopic composition and relative proportions of strontium entering the ocean. Strontium is delivered to the oceans by dissolution in river water with an average ratio of strontium-87 to strontium-86 ($^{87}Sr/^{86}Sr$) of 0.710 and by dissolution and recrystallization of submarine carbonates. Strontium is removed from the oceans by incorporation into precipitate minerals such as calcium carbonate as well as by inorganic precipitation of evaporites, phosphorites, and ferromanganese oxides. Seawater also exchanges strontium with ocean-floor rocks at mid-ocean ridges. Ocean-floor rocks (basalt) have a relatively low average $^{87}Sr/^{86}Sr$ ratio of 0.703. The variations in the $^{87}Sr/^{86}Sr$ ratio in seawater over millions of years are largely due to the changing inputs of continent-derived strontium delivered by rivers versus strontium derived from exchange of seawater with sea-floor basalt. Comparison of the strontium isotope curve and the geologic record indicate that the $^{87}Sr/^{86}Sr$ ratio in seawater is highest during periods of mountain building and continental collision, when there is a large flux of material from the continents due to increased erosion rates. Conversely, the $^{87}Sr/^{86}Sr$ ratio is lowest in seawater during times when the continents are splitting apart and when sea-floor spreading rates are greatest. *See* SEAWATER.

Measurements of the $^{87}Sr/^{86}Sr$ ratio of calcium-bearing marine precipitates (calcium carbonate, anhydrite, gypsum, and phosphate minerals) show that the $^{87}Sr/^{86}Sr$ ratio of strontium dissolved in seawater has fluctuated over the past 600 million years between the limits of 0.7068 and 0.7092. Because the average strontium atom remains in seawater 3–4 million years before it is removed, which is much longer than the mixing time of the oceans (1000 years), the $^{87}Sr/^{86}Sr$ ratio in seawater is the same throughout the oceans at any given time. Consequently, calcite that forms in the ocean in different places at the same time has the same $^{87}Sr/^{86}Sr$ ratio. When the $^{87}Sr/^{86}Sr$ ratio is changing rapidly with time in seawater, such as between the late Eocene and the present (the past 40 million years), excellent age resolution can be obtained for stratigraphic studies. *See* EOCENE.

An assumption in using strontium isotopes (or any geochemical method) in stratigraphic studies is that the $^{87}Sr/^{86}Sr$ ratio remains unaltered during diagenesis (all the chemical and physical changes that occur in the sediment following its deposition), and hence retains the original strontium isotopic signature. Careful inspection and sample selection prior to analysis are necessary to assess the degree of diagenetic alteration. *See* DIAGENESIS; STRONTIUM.

Stable isotope stratigraphy. The elements hydrogen (H), carbon (C), nitrogen (N), oxygen (O), and sulfur (S) owe their isotopic distributions to physical and biological processes that discriminate between the isotopes because of their different atomic mass. The use of these isotopes in stratigraphy is also facilitated by cycles of the isotopic composition of seawater, but the isotopic ratios in marine minerals are also dependent on water temperature and the mineral-forming processes.

Oxygen isotope stratigraphy. Oxygen isotope stratigraphy has developed since the mid-1950s following the pioneering work of Harold Urey; it has become the most widely used method for dating and correlating marine sediments deposited during the Quaternary (Pleistocene and Holocene). Oxygen isotope stratigraphy is based on the ratio of two isotopes of oxygen (oxygen-18 and oxygen-16) in the calcium carbonate ($CaCO_3$) shells of marine fossils. *See* QUATERNARY.

Detailed measurements of the ratio of oxygen-18 relative to oxygen-16 ($^{18}O/^{16}O$) in planktonic foraminifera from sediments cored in the deep ocean show that the ocean changed periodically over the past 3 million years because of alternating glacial (or ice age) and interglacial conditions. During glacial periods, water depleted in oxygen-18 is transferred to and stored on the continents as glaciers, leaving the oceans enriched in oxygen-18 (higher $^{18}O/^{16}O$ ratio). Although both the water temperature and the $^{18}O/^{16}O$ ratio in the water determine the $^{18}O/^{16}O$ ratio in the carbonate shells, the fluctuations of the $^{18}O/^{16}O$ ratio between interglacial and glacial periods are mainly caused by fluctuating continental ice volume. Each isotopic fluctuation representing a glacial-interglacial cycle has been assigned a stage number, with odd numbers corresponding to interglacial periods and even numbers corresponding to glacial periods (numbered 1 through 23). These stages are globally synchronous and can be intercalibrated with other stratigraphic scales, such as magnetostratigraphy and biostratigraphy. There is much evidence that these climatic fluctuations (called Milankovitch cycles) are a result of changes in the incoming solar radiation reaching the surface of the Earth. The dominant periodicities of these cycles are 100,000, 40,000, and 23,000 years. *See* GLACIAL EPOCH; INSOLATION.

Carbon isotope stratigraphy. Stable carbon (C) isotope stratigraphy involves measurements of the ratio of carbon-13 to carbon-12 ($^{13}C/^{12}C$) in carbonate ($CaCO_3$) fossil shells and organic matter. Organic matter and calcium carbonate are converted to carbon dioxide (CO_2) prior to simultaneous measurement of the ratios $^{13}C/^{12}C$ and $^{18}O/^{16}O$ on a gas-source mass spectrometer.

The carbon isotopic composition in marine organisms (calcite shells or organic tissue) reflects the local and global changes in the cycling of nutrients and CO_2 in the ocean, as well as ocean circulation and ocean-atmosphere interactions. The $^{13}C/^{12}C$ ratio of both organic material and carbonate material can be used as a stratigraphic tool. The $^{13}C/^{12}C$ ratio of a sample depends on the $^{13}C/^{12}C$ ratio of the source [atmospheric CO_2 for land plants, both dissolved CO_2 and bicarbonate (HCO_3^-) for aquatic environments] and the isotopic fractionation produced by equilibrium and kinetic effects (the fractionation of

isotopes of the same element in physical or chemical processes occurs because of their slight mass differences). In general, organic carbon is depleted in carbon-13 relative to carbonate carbon in the marine environment. A major focus of carbon isotope studies in stratigraphy is the assessment of detailed variations and excursions of the $^{13}C/^{12}C$ ratio on a stratigraphic level. For example, during periods of increased productivity in the surface layer of the ocean, the carbon-12 is preferentially removed by phytoplankton during photosynthesis, enriching the surface water in carbon-13. Calcium carbonate and organic tissue that form during these periods have an increased $^{13}C/^{12}C$ ratio. Many organic carbon-rich, laminated black shales of Jurassic and Cretaceous age contain organic carbon that is enriched in carbon-13.

Sulfur isotope stratigraphy. The isotopic composition [the ratio of sulfur-34 to sulfur,-32 ($^{34}S/^{32}S$)] of sulfur in the sulfate molecules (SO_4^{2-}) dissolved in seawater has varied systematically through Phanerozoic time (the past 600 million years). Sulfate is a major constituent of the marine evaporite minerals gypsum and anhydrite, which are the materials commonly used as recorders of the ocean isotopic ratios. The $^{34}S/^{32}S$ ratio in seawater reflects a complex balance of the sulfate delivered to the oceans by rivers; it also reflects the relative importance of sulfur removal as reduced sulfur in pyrite, as opposed to oxidized sulfur in gypsum and anhydrite. *See* ANHYDRITE; GYPSUM; PYRITE.

Radioactive isotope stratigraphy. Radiocarbon [the radioactive isotope of carbon (^{14}C)] and the isotopes uranium-234 (^{234}U) and thorium-230 (^{230}Th) are used in radioactive isotope stratigraphy.

Radiocarbon stratigraphy. Carbon-14 is used in dating and stratigraphy of both marine and terrestrial sediments younger than 50,000 years. Radiocarbon is continually produced in the upper atmosphere by the capture of neutrons by nitrogen (^{14}N), forming carbon-14, which is radioactive with a half-life of about 5730 years. The carbon-14 enters the hydrosphere and biosphere as CO_2, and it is incorporated into plant and animal organic tissues and hard parts (such as calcium carbonate). When the plant or animal dies, exchange with the atmosphere ceases; the carbon-14 begins to decay, starting the radioactive clock.

Uranium-thorium stratigraphy. Marine minerals such as calcite incorporate minor amounts of the radioactive isotope uranium-234 (^{234}U), which has a half-life of 2.4 million years and decays to thorium-230, which is also radioactive with a half-life of 70,000 years. Because marine calcite forms almost free of thorium-230, the accumulation of radiogenic thorium-230 can be used to measure the age of marine calcites up to about 300,000 years old. This technique is valuable for measuring the age of coral reefs. *See* THORIUM; URANIUM.

Molecular stratigraphy. Certain organic molecules that can be linked with a particular source (called biomarkers) have become useful in stratigraphy. The sedimentary distributions of biomarkers reflect the biological sources and inputs of organic matter (such as that from algae, bacteria, and vascular higher plants), and the depositional environment. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS.

One method of molecular stratigraphy uses the composition of long-chain alkenones (lipids whose molecules comprise chains of 37–39 carbon atoms). These compounds are produced by certain photosynthetic algae (coccolithophorids) that regulate the fluidity of their membranes in waters of different temperatures by biosynthetically controlling the extent of unsaturation in their alkenones. The degree of unsaturation of alkenones in marine sediments varies as a function of water temperature, which varies from glacial to interglacial times. A molecular stratigraphic record of sea surface temperatures in Quaternary sequences has been developed and intercalibrated with the oxygen isotope record for the past half million years. Alkenones are found in marine sediments throughout the Quaternary and Tertiary.

Trace elements. Certain trace metals, such as nickel, copper, vanadium, magnesium, iron, uranium, and molybdenum, are concentrated in organic-rich sediments in proportion to the amount of organic carbon. Although the processes controlling their enrichment are complex, they generally form in an oxygen-poor environment (such as the Black Sea) or at the time of global oceanic anoxic events, during which entire ocean basins become oxygen poor, resulting in the death of many organisms; hence large amounts of organic carbon are preserved in marine sediments. The trace-metal composition of individual stratigraphic units may be used as a stratigraphic marker (or "fingerprint"), making it possible to correlate, for example, black shales from location to location within a basin. *See* GEOCHEMISTRY; MARINE SEDIMENTS; STRATIGRAPHY.

Donald J. DePaolo; B. Lynn Ingram

Bibliography. M. E. Brookfield, *Principles of Stratigraphy*, 2004; G. Faure and T. M. Mensing, *Isotopes: Principles and Applications*, 3d ed., 2004; G. Ottonello, *Principles of Geochemistry*, 2000.

Chemotherapy and other antineoplastic drugs

Chemotherapy is the use of chemicals to treat any disease, but the term has come to be applied most commonly to the use of drugs to treat cancer and that exhibit cytotoxic activity through inhibition of cell division. Recently, several other classes of drugs (biologicals) started to be used to treat cancer, often in conjunction with conventional chemotherapy. Cancer is an abnormal growth or proliferation of cells that tends to invade locally or spread to distant parts of the body. Several treatment modalities can be used for cancer. Surgery physically removes the abnormal growth, whereas radiation, immunotherapy, hormonal therapy, and chemotherapy are directed at killing or slowing the growth of cancerous cells. *See* CANCER (MEDICINE).

History. In 1865, in an early, well-documented use of chemotherapy, Heinrich Lissauer gave potassium arsenite (Fowler's solution) to treat leukemia. Despite this work, Paul Ehrlich, a German physician and 1908 Nobel Laureate in Medicine, is generally credited with pioneering the field of chemotherapy and coining the word. However, his efforts were directed toward discovering antibiotics for infections rather than cancer drugs. He was an early advocate of understanding biochemically why certain chemicals effectively kill actively growing bacteria. He also coined the term "magic bullet" for drugs that specifically kill bacteria. *See* ANTIBIOTIC.

The first extensively utilized cancer chemotherapeutic agent was nitrogen mustard, which was used in the mid-1940s by Alfred Gilman and Frederick S. Philips for the treatment of lymphomas and chronic leukemias. During World War I it had been noted that mustard gas, used as a toxic irritant, caused destruction of lymph nodes and bone marrow, which are both sites of active cell proliferation. This observation suggested that mustard gas is toxic to actively growing cells and that it might also kill proliferating cancer cells. Nitrogen mustard is still occasionally used for the treatment of some lymphomas, although it has largely been replaced by other chemotherapeutic regimens. *See* LEUKEMIA.

Folic acid antagonists. The second group of useful anticancer drugs developed were folic acid antagonists. All actively growing cells require the chemical metabolite folic acid to grow and divide. The antagonists were designed to inhibit a cancer cell's ability to produce folic acid, and therefore they are toxic. Aminopterin was the first in this group to be employed clinically. A major breakthrough occurred in 1948, when it was reported that over half of children with acute leukemia could obtain temporary remissions by using aminopterin. This began the era of antimetabolite therapy. Methotrexate, a related drug, replaced aminopterin and is still used in acute lymphoblastic leukemia, choriocarcinoma, carcinomas of the head and neck region, and for other tumors. The first documented chemotherapeutic cure was with methotrexate, which was used in 1961 to treat a gestational choriocarcinoma.

Nucleic acid antagonists. The third group of chemotherapeutic agents was designed to have a toxic effect on the cell's DNA. The development of these drugs was based on the hypothesis, conceived by George Hitchings and Gertrude Elion, that tumor growth might be stopped with chemicals that antagonize nucleic acids, which are necessary for cancer cell growth and division. They eventually were awarded the Nobel Prize in Physiology or Medicine for the development of such compounds. The first clinically useful compound, 6-mercaptopurine, was described in 1952 and is still used for the treatment of acute lymphoblastic leukemia. Another drug, with a different site of action, 5-fluorouracil, was synthesized in 1957 by Charles Heidelberger. It is used today in the treatment of colorectal, head and neck, and breast cancer. Cytarabine (cytosine arabinoside) was reported in 1968 to be useful in adult acute leukemia. It is still used for the primary treatment of most patients with acute myeloid leukemia. Fludarabine is a modern member of this group effective in the treatment of chronic lymphocytic leukemia.

Antitumor antibiotics. Antitumor antibiotics became recognized as potential chemotherapeutic agents in 1954, when actinomycin D (dactinomycin) was studied. They differ from antimicrobial antibiotics in that their spectrum of cytotoxicity includes human cells rather than bacteria and fungi. Doxorubicin (Adriamycin) is one of the most widely used anticancer drugs for both solid tumors and blood or bone marrow tumors.

Since the mid-1950s, many new drugs have been developed. This progress was stimulated by the pharmaceutical industry, by the important preclinical work of Howard E. Skipper and Frank M. Schabel, Jr., at the Southern Research Institute, and by the establishment of a special section of the U.S. National Cancer Institute for drug development and testing, which began work in 1955. Chemotherapeutic drugs now total more than 50, many having novel mechanisms (**Table 1**). The availability of the techniques of molecular biology and a greater understanding of the genetic abnormalities of cancer are resulting in more effective therapies, which tend to be more specific for cancer cells. Drugs which inhibit angiogenesis (blood vessel formation) or which target promoting (oncogenes) or suppressing (tumor suppressor) genes are being tested and incorporated into treatment regimens.

Monoclonal antibodies. A new class of drugs involves the use of monoclonal antibodies that target specific cell receptors. Bevacizumab is a recombinant humanized monoclonal antibody against the vascular endothelial growth factor (VEGF). VEGF is found to be overexpressed in many cancers and is in part responsible for the process of intratumoral angiogenesis. Bevacizumab is currently used in metastatic colon cancer.

Cetuximab, which binds the epidermal growth factor receptor (EGFR), and Trastuzumab, which targets Her2-neu receptor in breast cancer cells, are other examples of antibodies currently used in clinical practice (Table 1). Rituximab is an important therapeutic monoclonal antibody to the B-lymphocyte cell-surface molecule designated CD20. It has important activity in inhibiting the growth of lymphomas that express CD20. It is also possible to deliver a chemotherapy agent to a specific target cell type using a monoclonal antibody. For example, gemtuzumab ozogamicin (Mylotarg) is a drug with the monoclonal antibody anti-CD33 attached to a cytotoxic antitumor antibiotic, calicheamicin. This drug is thereby targeted by the antibody to the leukemia cell. Other than a small incidence of allergic reactions during infusion, side effects of monoclonal antibodies are variable and are closely related to the function of the receptor (VEGF, EGFR, CD20) being targeted by the antibody. Monoclonal antibodies also have been combined with

TABLE 1. Groups of drugs and mechanisms of action

Group	Source	Mechanism of action	Examples	Uses
Alkylating agents	Chemical synthesis	Blocking of DNA synthesis by chemical attachment of a reactive part of the drug (alkyl portion) to DNA and other cell components to form damaging cross links	Ifosfamide, cyclophosphamide, busulfan, chlorambucil, melphalan, dacarbazine, carmustine (BCNU), lomustine (CCNU), streptozotocin, procarbazine, hexamethylmelamine, thiotepa, nitrogen mustard (mechlorethamine), estramustine, temozolomide	Many solid tumors and hematologic (blood and bone marrow) malignancies; brain tumors; often used for high-dose chemotherapy prior to bone marrow transplantation
Antifolates	Chemical synthesis	Inhibition of DNA synthesis by inhibiting the enzyme dihydrofolate reductase	Methotrexate, pemetrexed	Head and neck cancer; lung cancer, leukemia; lymphoma; breast cancer; choriocarcinoma; mesothelioma
Antipurines	Chemical synthesis	Variable, for example, incorporation into DNA as a false building block (6-mercaptopurine, cladribine), inhibition of DNA polymerase which is required for synthesis (fludarabine), and inhibition of adenosine deaminase enzyme with the accumulation of deoxyadenosine which kills lymphocytes	Fludarabine, cladribine, azathioprine, 6-mercaptopurine, 6-thioguanine, pentostatin, clofarabine	Chronic lymphocytic leukemia; acute lymphoblastic leukemia; lymphoma; hairy cell leukemia; immunosuppression after transplantation or for aplastic anemia and autoimmune blood disorders
Antipyrimidines	Chemical synthesis	Variable, for example, incorporation into DNA (cytarabine); inhibition of thymidylate synthase (5-fluorouracil), and inhibition of DNA methylation (5-azacitidine)	Cytarabine, 5-fluorouracil, floxuridine, capecitabine, gemcitabine, 5-azacitidine, nelarabine	Acute leukemia; breast, colorectal, pancreas, and head and neck cancer; myelodysplastic syndrome
Anticancer antibiotics	Fungi or bacteria or chemical synthesis	Variable, for example, oxidation of DNA and RNA (bleomycin), inhibition of RNA synthesis (actinomycin D)	Bleomycin, dactinomycin, mitoxantrone, mitomycin C, streptozotocin	Hodgkin's disease; lymphoma; testis, breast, cervix, penis, islet cell pancreas, and head and neck cancer; acute lymphoblastic leukemia
Anticancer antibiotics of the anthracycline group	Bacteria	Several may be important, including intercalation into DNA, oxygen free-radical formation, and inhibition of topoisomerase type II	Idarubicin, doxorubicin, daunorubicin, epirubicin	Acute leukemia; Hodgkin's disease; lymphomas; multiple myeloma; breast, lung, bone, and thyroid cancer
Drugs that interfere with synthesis of cell microtubules	Semisynthesis or natural occurrence	Binding of the protein tubulin and interference with polymerization which is required to form microtubules	Vincristine, vinblastine, vindesine, vinorelbine	Multiple myeloma; lymphoma; lung and testis cancer; acute lymphoblastic leukemia; Hodgkin's disease
Drugs that stabilize microtubules against depolymerization and induce mitotic block	Tree bark, fungi, or semisynthesis	Enhanced microtubule formation, inhibition of mitosis	Paclitaxel, docetaxel	Colorectal, ovarian, breast, lung, germ cell, and head and neck cancer
Cytokines	Recombinant DNA technology	Interference with intracellular signaling	α-Interferon, aldesleukin (interleukin-2)	Chronic myelocytic leukemia; melanoma; kidney cancer; hairy cell leukemia
Topoisomerase inhibitors	Semisynthesis	Inhibition of topoisomerase enzymes I and II resulting in single- or double-strand DNA cleavage	Etoposide, teniposide, topotecan, irinotecan, anthracyclines	Leukemia; lymphoma; lung, colorectal, and ovarian cancer; myelodysplastic syndrome
Platinum drugs	Chemical synthesis	Binding to and inhibiting function of DNA	Cisplatin, carboplatin, oxaliplatin	Lung, ovary, and bladder cancer

(*cont.*)

TABLE 1. Groups of drugs and mechanisms of action (*cont.*)

Group	Source	Mechanism of action	Examples	Uses
Enzyme therapy	Bacteria	Depletion of asparagine, an essential amino acid	Asparaginase, pegasparaginase	Acute lymphoblastic leukemia
Inhibitors of ribonucleotide reductase	Synthesis	Inhibition of ribonucleotide reduction necessary for DNA synthesis	Hydroxyurea	Myeloproliferative diseases (such as chronic myelocytic leukemia, polycythemia vera, essential thrombocytosis); acute leukemia; melanoma; head and neck cancers; sickle cell anemia
Differentiating agents	Derivative of vitamin A	Binding to receptors on cell nucleus and stimulating differentiation of some cells	Tretinoin (all-*trans*-retinoic acid)	Acute promyelocytic leukemia
Molecular targeted therapy: small-molecule kinase inhibitors	Chemical synthesis	Inhibit dysregulated cancer-related kinases	Imatinib mesylate (targets ABL tyrosine kinases such as BCR/ABL kinase), gefitinib and erlotinib (both target epidermal growth factor receptor kinases), sorafenib (multikinase inhibitor that targets serine/threonine and tyrosine kinases)	Chronic myelocytic leukemia, acute lymphoblastic leukemia with Philadelphia chromosome, gastrointestinal stromal tumors, hypereosinophilic syndrome, non-small-cell lung cancer, kidney cancer
Proteasome inhibitors	Chemical synthesis	Inhibit a large complex that degrades proteins	Bortezomib	Multiple myeloma
Forms of naturally occurring elements	Natural	Not known. Induces apoptosis and differentiation	Arsenic trioxide	Acute promyelocytic leukemia
Monoclonal antibodies	Hybridoma techniques	Monoclonal antibodies against surface proteins	Rituximab (anti-CD20), alemtuzumab (anti-CD52), bevacizumab (antivascular endothelium growth factor), cetuximab (antiepidermal growth factor receptor), trastuzumab	Lymphomas, chronic leukemias, colorectal cancer, breast cancer
Targeted immuno-chemotherapy	Hybridoma techniques and chemical synthesis	Use cancer surface receptors to target chemotherapy	Gemtuzumab ozogamicin	Acute myelocytic leukemia
Targeted immuno-radiotherapy	Hybridoma techniques and chemical synthesis	Use cancer surface receptors to target radioactive compounds	Ibritumomab tiuxetan, tositumomab/iodine I-131	Lymphomas
Immunomodulatory drugs	Chemical synthesis	Multiple (antiangiogenic, T-cell modulation, cytokine secretion modulation)	Thalidomide and lenalidomide	Multiple myeloma, myelodysplastic syndromes
DNA methylation modulators	Chemical synthesis	Reverse abnormal DNA methylation	5-Azacitidine, decitabine	Myelodysplastic syndromes, acute leukemia
Hormonal therapy	Chemical synthesis	Multiple (blocks receptors, inhibits hormones, other)	Goserelin, bicalutamide, ketoconazole, tamoxifen, anastrozole, letrozole, tamoxifen	Prostate and breast cancer

radioisotopes for the treatment of lymphomas.

Tyrosine kinase inhibitors. Another class of anticancer agents that has been recently introduced in oncology clinical practice involves small-molecule tyrosine kinase inhibitors. Among other functions, tyrosine kinase enzymes regulate several of the intracellular processes normally responsible for cell proliferation. In cancers, these processes are often found to be dysregulated. Imatinib mesylate (Gleevec) is one of the most exciting recent drugs. It is a rationally designed ABL (Abelson leukemia oncogene) specific tyrosine kinase inhibitor. It inhibits BCR-ABL, which characterizes chronic myelocytic leukemia cells, and results in hematologic, cytogenetic, and sometimes molecular complete remissions (no abnormality detectable by the most sensitive molecular techniques). Erlotinib inhibits the EGFR-associated tyrosine kinase and has been shown to produce a survival benefit in patients with advanced non-small-cell lung cancer. Special patient populations that may derive greater

benefit from erlotinib are being actively studied, including those with other cancers. These tyrosine kinase inhibitors can frequently be given by mouth. A large number of other inhibitors are currently under clinical investigation.

Hormonal agents. Hormonal agents have been used for several decades to inhibit cancer cells, especially in breast cancer, which frequently expresses hormone receptors on the cell surface. Representative hormonal agents include tamoxifen (a selective estrogen receptor modulator that also has partial estrogen agonist activity) and aromatase inhibitors such as anastrozole, among others. Aromatase is the enzyme responsible for conversion of androgens into estrogens and is the predominant source of estrogens in postmenopausal women. Hormonal therapy is also widely employed in metastatic prostate cancer, such as luteinizing hormone-releasing hormone (LHRH) agonists (goserelin, leuprolide), direct antiandrogen (bicalutamide), and blockers of adrenal androgen production (ketoconazole).

Sources. Chemotherapeutic drugs come from various sources. They may be extracted from natural sources—such as from bacteria, fungi, plants, or trees—or they may be chemically synthesized to target a specific protein—such as tyrosine kinase inhibitors. They may also be semisynthetic, meaning that the process starts with a complex natural compound that is then modified chemically. They may also be produced using modern recombinant DNA technology. Monoclonal antibodies may be murine (-monab); chimeric (-ximab), containing mouse variable regions and human constant regions; humanized (-zumab), containing murine binding regions (about 5% of total antibody) in a human framework; or human (-umab), derived using a variety of techniques from immunoglobulin genes with no murine immunoglobulin sequences.

Clinical trials. After basic laboratory studies suggest that a chemical might be effective in killing cancer cells, animal and then human testing begins. These tests are carried out in steps in order to systematically and safely identify those chemicals with the most usefulness. In phase I tests, the toxicity of the drug (or combination of drugs) and the maximum tolerated dose are determined by gradually increasing the dose given to patients until toxic effects are observed. In phase II studies, efficacy (tumor shrinkage) is determined in a single series of patients with a certain type of cancer. Recently, several new phase II design have been developed (for example, one called randomized discontinuation design) to identify preliminary evidence of activity of drugs that may have a greater effect on preventing cancer growth, rather than actual shrinkage. Typically, results of phase II studies are then used to proceed with larger phase III studies. In phase III tests, the drug (or combination) is randomized against the standard drug (or combination) used for that cancer to determine whether the experimental drug is more effective than the standard therapy. Phase III trials have historically been the gold standard to prove clinical efficacy.

Principles. Chemotherapy or immunotherapy is used when a cancer has spread to multiple sites (has metastasized) and cannot be removed surgically or treated with radiation therapy. It can also be used after all detectable tumor has been eradicated (complete remission) to destroy suspected but undetectable residual tumor. This type of treatment is called adjuvant, intensification, or consolidation chemotherapy. Adjuvant therapy is effective in breast and colon cancer, and consolidation in acute leukemia, for example. Neoadjuvant chemotherapy is used to shrink a tumor prior to surgery or radiation. *See* ONCOLOGY.

Fractional cell kill. For classic chemotherapy agents, the results of treatment depend upon how much chemotherapy is given, how many times it is applied, and how effective it is against the tumor. These facts are derived from a basic principle of chemotherapy that a uniform dose of a drug will destroy a constant fraction rather than a constant number of tumor cells regardless of the size of the tumor or number of cells present (fractional cell kill hypothesis). This result can be seen in the **illustration**, where the tumor

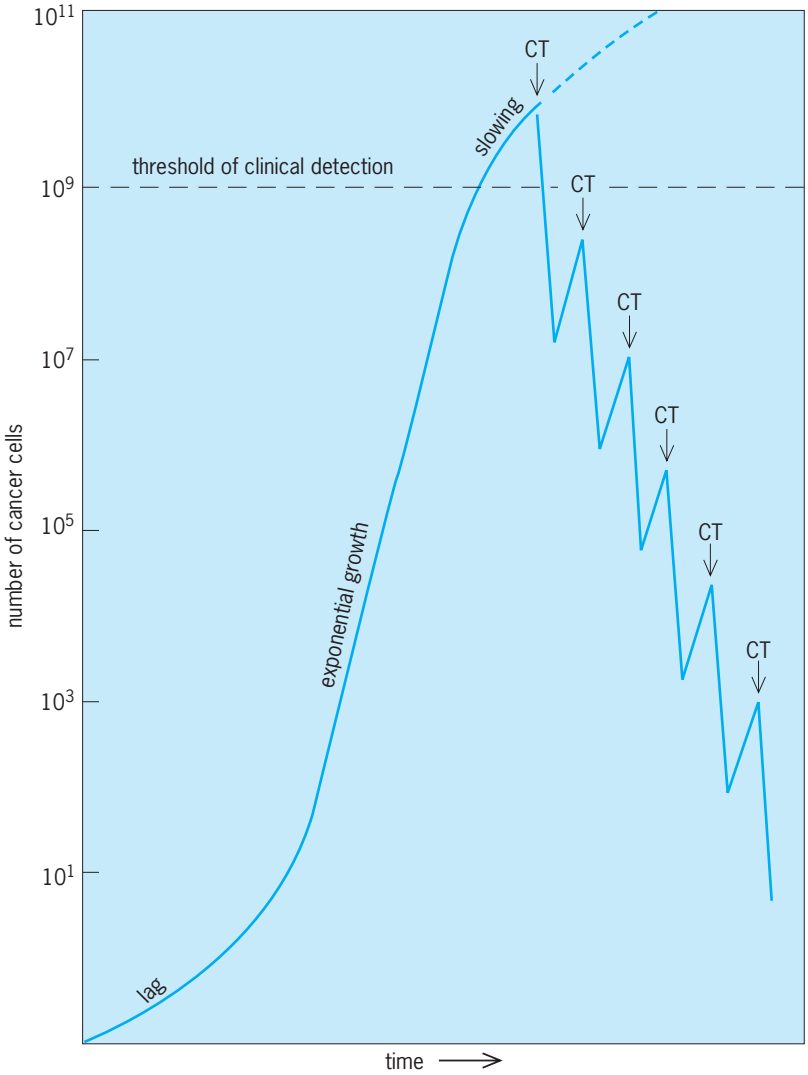

Sigmoid-shaped (Gompertzian) cancer growth (prior to first arrow) and fractional cell kill due to chemotherapy (after first arrow). Arrows indicate times of chemotherapy. Vertical axis is a log scale. The threshold for clinical detection of cancer is 10^8–10^9 total body tumor cells; since cancer probably starts from a single abnormal cell, this requires 27–30 doublings.

TABLE 2. Results of chemotherapy for disseminated disease in various cancers*

Cure in a majority of patients	Testicular cancer, childhood acute lymphoblastic leukemia, Hodgkin's disease, selected non-Hodgkin's lymphomas such as adult large B-cell lymphoma, gestational choriocarcinoma, hairy cell leukemia
Cure in a small subset of patients	Acute adult myeloid and lymphoblastic leukemia, ovarian cancer, some intermediate- and high-grade lymphomas, small-cell lung cancer
Induction of long-lasting remission	Follicular lymphoma, chronic myelocytic leukemia, gastrointestinal stromal tumor
Adjuvant† use with increased cure rates	Breast cancer, colorectal cancer, Ewing's sarcoma, Wilms' tumor, osteogenic sarcoma, rhabdomyosarcoma, glioblastoma multiforme, bladder cancer, gastric cancer
Prolongation of life but rare cures when metastatic‡	Multiple myeloma, follicular lymphoma, chronic lymphocytic leukemia, small cell lung cancer, bladder cancer, gastric cancer, colorectal cancer, pancreatic cancer, lung cancer, breast cancer, renal cell carcinoma
Palliation with unclear prolongation of life when metastatic	Head and neck cancer, hepatoma, melanoma, soft-tissue sarcoma, gallbladder cancer, cholangiocarcinoma

*These categories do not apply to cancers that have not metastasized and are treatable with surgery or radiation therapy alone, or to those that can be treated with bone marrow transplantation.
†Adjuvant refers to the use of chemotherapy in a situation where no tumor is detectable but there is a high chance that undetectable tumor cells are still present; the term for this type of therapy in blood or marrow cancers is consolidation therapy.
‡Some patients in this group can be cured by bone marrow transplantation.

number declines linearly on the log scale after the chemotherapy is applied. Drugs which are more effective in killing cancer cells give a greater slope of decline. Regrowth eventually occurs, but each subsequent dose of chemotherapy brings about a further decline in tumor cell number until eventually, in optimal circumstances, the number of tumor cells is low enough for the body to eliminate them by other mechanisms, resulting in a chemotherapy cure.

Mechanisms of cell death. Although the different groups of chemotherapeutic agents have different mechanisms of toxicity for cells, the cellular death response is often thought to be by apoptosis (programmed cell death). In this process a death signal is generated from within the damaged cells. This initiates a series of energy-requiring, choreographed cellular events leading to cellular suicide. Other mechanisms of death are necrosis and clonogenic (reproductive) death.

Goal. The goal of therapy should be determined in all patients before chemotherapy is started. The possibilities include cure, prolongation of life but not a cure, or palliation (alleviating symptoms) without likely prolongation of life (**Table 2**). The choice of chemotherapy depends not only on the type of tumor and the drug efficacy but, most importantly, on the patient. Also, it must be decided whether the chemotherapy will improve the quality of life of the patient and not just shrink the tumor. Age and the general condition of the patient influence the outcome. Those patients who are more active (better performance scores, which estimate degree of physical incapacitation by the tumor) are more likely to have a favorable outcome to chemotherapy. There are questionnaires to aid in measuring quality of life before and after drugs are given.

Dosage. For conventional chemotherapy, dosing has traditionally been based on the amount of body surface in square meters, or occasionally just on body weight. These dose calculations reflect the fact that toxicity often occurs at doses only slightly higher than those needed for therapy. In clinical trials, often the maximum tolerated dose (MTD) is the one taken forward for further testing and eventual routine use. Tyrosine kinase inhibitors, in contrast, are often prescribed at fixed doses expressed in milligrams. Knowledge of pharmacokinetic principles and drug interactions is very important when administering chemotherapy drugs.

Measurement of effect. The response of a cancer to a drug is determined by measuring the size of the cancer directly or the amount of marker substances produced by some tumors. A complete response is defined as disappearance of the tumor. A partial response means a specified decrease in tumor size that does not qualify for a complete response. In many cancers, the ultimate measure of benefit is whether the patients have a prolongation of life, which in clinical trials may require observation over a long period of time. Time to progression, or the time it takes for the cancer to recur after initial treatment, is another commonly used endpoint in cancer studies. Quality of life is another endpoint often applied in clinical studies in cancer.

Adjuvant therapy. This is a concept that is commonly used in medical oncology and accounts for significant use of chemotherapy at present. It translates into using chemotherapy after surgical removal of the cancer (for example, mastectomy for breast cancer). Chemotherapy is used to kill microscopic amounts of cancer cells that were not removed with the surgery or had spread to distant areas before the surgery. Adjuvant chemotherapy has been shown in numerous phase III trials to increase cure rates in selected common cancers. Typically, large numbers of patients need to be treated to demonstrate a benefit in the population.

Minimal residual disease. The detection of cancer cells that remain in submicroscopic numbers after therapy is referred to as minimal residual disease (MRD). Its importance is being evaluated in several malignancies, and MRD has been shown to correlate with recurrent clinical disease, especially in selected leukemias and lymphomas. The availability of improved molecular biology techniques has allowed increasing ability to detect MRD. These techniques include cytogenetics (chromosome analysis) and fluorescent in-situ hybridization (FISH), which uses DNA probes specific to regions of individual chromosomes, and these targeted areas are then identified by a fluorescent marker. Finally, reverse-transcription polymerase chain reaction (RT-PCR) is the most sensitive method currently available for detection of MRD.

Side effects. Chemotherapeutic drugs have side effects that are specific for the agent and may include transient or permanent damage to almost any organ system of the body. They cause side effects because, in addition to destroying the actively dividing cancer cells, they damage actively growing normal cells. Common sites of damage are the bone marrow, resulting in increased risk for infection, anemia, and bleeding; the gastrointestinal tract, resulting in vomiting and diarrhea; the hair follicles, causing hair loss; and the kidneys, resulting in renal failure. Today, cancer specialists have many ways of preventing or diminishing these toxicities, such as red blood cell transfusions for anemia, drugs to prevent nausea and vomiting, and growth factors which may improve anemia (erythropoietin) or shorten the time of low white blood counts (granulocyte-colony stimulating factor).

Drug resistance. Cells can become resistant to drugs of different groups, including anthracyclines, vincristine, etoposide, and paclitaxel. The general mechanism is referred to as multidrug resistance (MDR). This resistance is due to an acquired ability of the cells to pump these drugs out before they can cause irreversible damage. This pumping action is due to a transmembrane glycoprotein referred to as P-glycoprotein. These tumor cells overexpress the gene *MDR1*, which can be a useful marker of the resistance.

There are other ways by which a cancer cell can become resistant to chemotherapy. The cell may increase the amount of an intracellular drug target (for example, increased dihydrofolate reductase, which produces methotrexate resistance), reduce uptake of drugs (for example, alkylating agents), or decrease expression of a drug target (such as topoisomerase II). In some cases, resistance is due to a malfunctioning tumor suppressor gene called *p53*. The *p53* protein is an important initiator of drug-induced apoptosis; therefore, cells lacking *p53* do not commit cellular suicide in response to chemotherapies in the same way that *p53*-positive cells would.

In the case of tyrosine kinase inhibitors, resistance is often correlated with the emergence of a malignant cell clone with a new mutation in the targeted receptor, making the drug ineffective. Resistance can also be due to the presence of alternative molecular pathways that lead to cell proliferation, other than the pathway targeted by the drug.

Tumor growth. After an early lag phase, tumors grow exponentially, meaning that the time for a tumor to double in size is constant. However, as tumors grow, there is a deceleration in growth rate (see illustration), which may be due to outgrowing their blood supply, lack of sufficient nutrients, or changes in growth signals. This slowing means that there is an increase in the number of cells that are not dividing (G0 cell cycle phase), and these cells may be less susceptible to killing by drugs which act only on dividing cells. This phenomenon may compromise effective anticancer chemotherapy.

Selectivity and toxicity. Some tumors fail to respond to chemotherapy because normal tissues of the bone marrow, gut, or other organs are just as sensitive to the drug as the tumor, and unacceptable side effects preclude increasing the dose to therapeutic levels. These side effects may be early and transient such as vomiting, hair loss, or bone marrow suppression, or late and irreversible such as second malignancies, decline in fertility, lung fibrosis, and heart damage. Fortunately, effective drugs have some selectivity for the tumor, often for unknown reasons. Prednisone is an example of a drug that has the ability to destroy malignant lymphocytes with little damage to normal cells, and it is therefore selective.

Tumor heterogeneity. Cancers are not always uniform in their content of abnormal cells. There may be a mixture of cancerous cells with different features and different sensitivities to chemotherapy. Some cells may be sensitive while others are resistant to a specific drug. The use of a drug that destroys only one type of cell within a heterogeneous tumor will lead to eventual failure of therapy.

Strategies to overcome clinical resistance. Most successful modern chemotherapeutic regimens are combinations of drugs. These are usually named as acronyms, for example, CHOP therapy for lymphomas [cyclophosphamide, hydroxydaunorubicin (another name for doxorubicin), oncovin (another name for vincristine), and prednisone]. Response to combination therapy is usually superior to single-drug therapy. The use of more than one drug increases the chance of a cancer being sensitive to treatment since it is possible that a cancer cell resistant to one drug will be sensitive to another. Combinations of drugs with different mechanisms of tumor cell killing are commonly used. Also, in an effort to decrease side effects, drugs with nonoverlapping side effects are used. Drugs can also be combined with other general types of treatment such as immunotherapy. It is known from preclinical testing that chemotherapy and immunotherapy are more effective when the tumor is small. Therefore, surgery or radiation can be used in combination with chemotherapy to shrink the tumor with one modality and subsequently destroy the remaining smaller tumor with the other. Conversely, chemotherapy may be used to decrease tumor size to make surgery or radiation possible. In the case of tumors sensitive

to tyrosine kinase inhibitors, resistance due to emergence of new mutations can sometimes be partially overcome by increasing the dose of the drug.

Biochemical modulation. Resistance to chemotherapy can be overcome in some instances by giving a second drug either simultaneously or in sequence. An example is the use of leucovorin (folinic acid) that potentiates 5-fluorouracil therapy administered to treat colorectal cancer by increasing reduced folates, which indirectly leads to a more stable complex of 5-fluorouracil with its target enzyme thymidylate synthetase and eventual cell death. Another example is experimental agents that increase the retention of drugs affected by multidrug-resistant cells.

High-dose chemotherapy. A strategy to kill tumors resistant to conventional doses of chemotherapy involves the use of extrahigh (supralethal) doses of chemotherapy. Because this damages the normal tissues, it is necessary to perform transplantation of hematopoietic stem cells (obtained from blood or bone marrow) afterwards to reconstitute vital bone marrow function.

Cell cycle. There has been a good deal of interest in the specific phase of the cell division cycle that is affected by chemotherapy, in the hope that this information could be used to develop more effective cures for cancer. However, the results have been disappointing overall. Some drugs kill cancer cells only in specific phases of the cell cycle (S-phase: cytarabine, methotrexate; M-phase: paclitaxel); others have no specificity (gemcitabine). Some drugs, such as the alkylating agents, cisplatin, and anthracyclines, can also kill nondividing cells. This type of information, which is usually established in tissue culture or animals, has been of limited practical use in planning human therapy. Recently, there has been renewed interest in control of the cell cycle by genes and growth factors. This may eventually lead to effective cell cycle–based drug strategies. Some of the proteins that control the cell cycle (for example, cyclins) might become targets for drug action. *See* CELL CYCLE; CELL DIVISION.

Differentiation. Tumor cells are often immature compared to normal cells and retain the ability to self-replicate. More differentiated mature cells often lose this ability and are end-stage cells. Therapy designed to induce differentiation often results in the loss of the cells' ability to reproduce. In this way tumor cells can be inactivated without administering toxic drugs, thereby avoiding toxicity to normal cells. An active drug in this group is tretinoin (all-*trans*-retinoic acid), which can lead to remissions in a large percentage of patients with acute promyelocytic leukemia. These remissions are usually of short duration unless cytotoxic drugs are added. Unfortunately, the strategy has not been useful in most cancers.

Drug delivery. Techniques are being developed to ameliorate toxicity so that higher doses of chemotherapeutic agents may be delivered directly to the tumor. One approach involves a lipid carrier to deliver more of the toxic drug to the tumor as compared to normal tissues. Thus far, the technique has some usefulness with the drugs doxorubicin in ovarian cancer and cytarabine in central nervous system lymphoma or leukemia.

C. Patrick Burns; Daniel A. Vaena

Bibliography. M. D. Abeloff et al., *Clinical Oncology*, 3d ed., Elsevier Churchill Livingstone, 2004; J. H. Burchenal, From wild fowl to stalking horses: Alchemy in chemotherapy, Fifth Annual David A. Karnofsky Memorial Lecture, *Cancer*, 35:1121–1135, 1975; B. A. Chabner and D. L. Longo (eds.), *Cancer Chemotherapy and Biotherapy: Principles and Practice*, 4th ed., Lippincott Williams & Wilkins, 2005; S. Hellman and E. E. Vokes, Advancing current treatments in cancer, *Sci. Amer.*, 275:118–123, 1996; M. C. Perry (ed.), *The Chemotherapy Source Book*, 3d ed., Lippincott Williams & Wilkins, 2001.

Cherry

Sweet or sour fruit of the *Prunus* genus, having cultivated and wild varieties.

Cultivated varieties. The two principal cherries of commerce are the sweet cherry (*Prunus avium*) and the sour cherry (*P. cerasus*). Both are of ancient origin and seem to have come from the region between the Black and Caspian seas. They were probably carried into Europe by birds before humans were there. The cherry was mentioned in 300 B.C. by the Greek botanist Theophrastus. It was taken to America from Europe.

Cherries of minor importance are the dwarf or western sand cherries (*P. besseyi*) of the plains region of North America; the Duke cherries, which are supposedly natural hybrids between the sweet and sour cherry; and the Padus cherries, which bear their small fruits in long clusters or racemes rather than in short fascicles.

Sweet cherries may be divided into two groups, firm-fleshed types known as Bigarreaus, represented by the Napoleon (also called the Royal Anne), and soft-fleshed types known as Hearts, represented by the Black Tartarian. Sour cherries may also be divided into two groups, clear-juice or Amarelle types, represented by the Montmorency, and colored-juice or Morello types, represented by the English Morello.

In North America the principal sweet commercial varieties are the Napoleon (white), Bing, Lambert, Van, Schmidt, and Windsor (dark). The principal commercial variety of sour cherry is the Montmorency.

Trees are propagated by budding into seedlings of the wild sweet (mazzard) cherry or of *P. mahaleb,* commonly called the perfumed cherry of southern Europe. The former is best suited to the sweet cherry, and the latter to the sour cherry. *See* PLANT PROPAGATION; REPRODUCTION (PLANT).

The sweet cherry is somewhat hardier than the peach, and the sour cherry is hardier than the sweet cherry. Cultivated varieties are usually planted 25–30 ft (7.5–9 m) apart and allowed to grow to heights of only about 15–20 ft (4.5–6 m) for ease of culture and harvest. Sour cherry trees may reach 30–40 ft (9–12 m) but are kept at 15–20 ft (4.5–6 m) and

planted 18–22 ft (5.4–6.6 m) apart. Cherry trees do best in well-drained sandy loam soils. They will not tolerate wet soil. Sweet varieties require cross-pollination; common sour varieties are self-fruitful (self-pollinated). *See* FRUIT; FRUIT, TREE; ROSALES. R. Paul Larsen

Wild varieties. The wild cherry tree, *P. serotina*, grows to a height of 100 ft (30 m) and may reach an age of 200 years. It belongs to the rose family, and is an important timber species of cherry in the United States. Native in the eastern half of the United States and adjacent Canada, it reaches its largest size in the southern Appalachian Mountains. The deciduous leaves are long-pointed, thickish, and shining above (**Fig. 1**). The edible fruit is a shiny black drupe found in long clusters. The hard, strong wood is a red-brown and is used for furniture, often as an imitation mahogany.

The sweet cherry or mazzard (*P. avium*), which grows to 70 ft (21 m), is naturalized over a large part of the United States. It has fruits borne in umbel-like clusters. It also differs from the wild cherry in having more ovate leaves with rounded teeth (**Fig. 2**).

The Japanese flowering cherries are varieties of *P. serrulata*, which attract much attention at blossoming time in early May.

Arthur H. Graves; Kenneth P. Davis

Diseases. A destructive disease of cherry in the orchard and after harvest is brown rot, caused by the fungus *Monilinia fructicola*. Infections occur first in blossoms, and later in the maturing fruit, causing rot (**Fig. 3**). Cherry leaf spot, caused by the fungus *Coccomyces hiemalis*, results in leaf spotting, yellowing, and premature defoliation. Without well-functioning leaves, trees lose vigor and decline in productivity. Powdery mildew, caused by the fungus *Podosphaera oxyacanthae,* impairs the ability of leaves to supply sugars and other products needed by the tree for normal development. These three diseases can be effectively controlled with timely sprays of recommended fungicides.

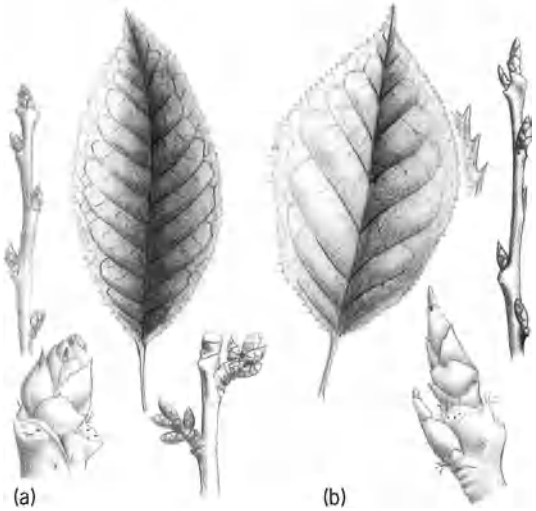

Fig. 1. Leaves, twigs, and buds of (*a*) sweet cherry (*Prunus avium*) and (*b*) choke cherry (*P. virginiana*)

Fig. 2. Leaves, twigs, and buds of (*a*) pin or fire cherry (*Prunus pensylvanica*) and (*b*) wild cherry (*P. serotina*)

Fig. 3. Brown rot disease of sweet cherry.

The major bacterial diseases of cherry are bacterial canker caused by *Pseudomonas syringae* (affecting leaves, branches, and fruit) and crown gall, due to *Agrobacterium tumefaciens*, which causes a tumor-like growth on the crown and roots.

Important virus (or viruslike) diseases are Prunus ring spot, prune dwarf, little cherry, and X disease ("buckskin"). Virus diseases affect certain tissues (wood, bark, leaves, seeds, pollen) of certain branches or even the entire tree. The symptoms vary with each disease, but usually involve small, mottled, yellow, or missing leaves, and small and poorly colored fruit. Tree growth and yield suffer. Preventive control measures include planting resistant varieties and certified virus-free stock and, where applicable, spraying for the control of insect vectors that transmit the pathogens. *See* PLANT PATHOLOGY; PLANT VIRUSES AND VIROIDS. Michael Szkolnik

Bibliography. J.M. Ogawa and H. English, *Diseases of the Temperate Zone Tree Fruit and Nut Crops*, 1991; *Virus Diseases of Small Fruits*, USDA Agr. Handb. 631, 1987; C. Wilson and M.E. Wisniewski (eds.), *Biological Control of Postharvest Diseases: Theory and Practice*, 1994.

Chert

A hard, dense sedimentary rock composed of fine-grained silica (SiO_2). Chert is characterized by a semivitreous to dull luster and a splintery to conchoidal fracture, and is most commonly gray, black, reddish brown, or green. Chert is also used as a field term to describe silica-rich rocks which may be impure; common impurities include carbonates, iron and manganese oxides, and clay minerals. When impurities change the texture of the rock to the extent that it is less dense and hard than chert, and has the appearance of unglazed porcelain, the rock is then called porcellanite or siliceous shale. The term flint is synonymous with chert, but its use has become restricted to archeological artifacts and to nodular chert that occurs in chalk. The term chert, however, is preferred for the nodular deposits. Jasper refers to red or yellow quartz chert associated with iron ore or containing iron oxide. Novaculite is a white chert of great purity and uniform grain size, and is composed chiefly of quartz; the term is mostly restricted to descriptions of Paleozoic cherts in Oklahoma and Arkansas. Chert synonyms that have become obsolete include silexite, petrosilex, phthanite, and hornstone. *See* JASPER.

Composition. The silica that composes chert occurs in four forms: quartz, chalcedony, opal-CT, and opal-A. Chalcedony is a microscopically fibrous form of the mineral quartz with water trapped in void spaces parallel to the fibers. The poorly understood opal-CT phase is probably an interstratified mixture of the silica minerals cristobalite and tridymite, although it has been suggested to be solely disordered tridymite. In chert, the chalcedony, quartz, and opal-CT are secondary forms of silica derived from the dissolution and later precipitation of a primary silica phase, which for most chert is biogenic silica (opal-A). The main rock-forming types of biogenic silica are the tests and frustules of the siliceous plankton diatoms, radiolarians, silicoflagellates, and ebriidians, and the siliceous spicules of sponges. The composition of chert changes with time and temperature, beginning with opal-A, which transforms to opal-CT and finally to quartz or chalcedony. Under special conditions, opal-A can transform directly to quartz. *See* CHALCEDONY; OPAL; QUARTZ.

Diagenesis. Opal-A and opal-CT are metastable phases within the Earth's crust. The diagenetic transformation of opal-A to opal-CT and then to quartz is primarily controlled by temperature and time, but also by surface area, pore-water chemistry, sediment permeability, and host-sediment composition. The mineralogical transformations are commonly, but not invariably, accompanied by textural transformations. For example, diatom or radiolarian ooze (unconsolidated marine sediment composed mostly of diatom frustules or radiolarian tests) or diatomaceous or radiolarian earth (weakly consolidated oozes that have been incorporated into on-land deposits), which are composed of opal-A, transform into porcellanite (most commonly composed of opal-CT), which in turn transforms into chert (most commonly composed of quartz). *See* DIAGENESIS; MARINE SEDIMENTS.

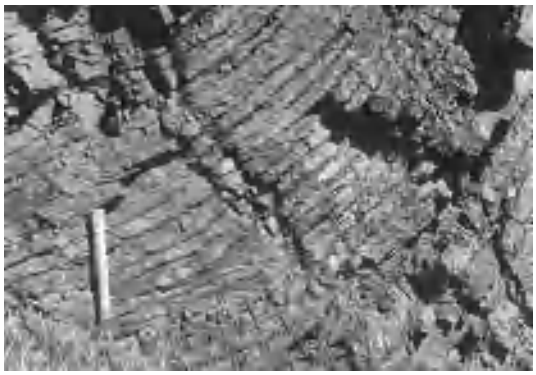

Rhythmically bedded radiolarian chert from the Franciscan Complex, Diablo Range, California; of Late Jurassic age.

Occurrence. Chert occurs mainly in three forms: bedded sequences (see **illus.**), nodular, and massive. Bedded chert (called ribbon chert if beds show pinch-and-swell structure) consists of rhythmically interlayered beds of chert and shale; chert and carbonates; or in some pre-Phanerozoic formations, alternations of chert and siderite or hematite. Bedded sequences can be hundreds of feet thick stratigraphically and cover areas of hundreds of square miles. Individual beds are commonly $^1/_2$–8 in. (1–20 cm) thick. Intercalated shale layers are commonly, but not invariably, thinner than the associated chert layers. In many places individual chert beds are laminated. Chert nodules and lenses occur primarily in chalk, limestone, and dolomite, and are the only types of chert recovered by the Deep Sea Drilling Project from open-ocean sediment. Chert nodules are well known from many places in western Europe, where their dark color contrasts strikingly with the white host chalk. Nodules and lenses vary in size from $^1/_2$ in. to 30 ft (1 cm to 9 m). Fossils and sedimentary structures characteristic of the host rock are preserved within the nodules. Massive cherts occur in the interstices between basalt pillows, and as the basal member of bedded chert that overlies pillow basalts in ophiolites. *See* BASALT; CHALK; DOLOMITE ROCK; LIMESTONE.

Origin. When a supply of silica is available, chert forms in four ways: (1) by replacement of mainly carbonate rock; (2) by deposition from turbidity currents composed primarily of biogenic silica; (3) by increasing the deposition of silica relative to terrigenous input, commonly by increased productivity of biogenic silica; and (4) by precipitation of silica from water under either hydrothermal or low-temperature hypersaline conditions. The replacement of chalk and limestone by silica has been well documented in Deep Sea Drilling Project cores and in some deposits on land. The silica is derived from the dissolution of biogenic silica deposited with the biogenic carbonate. The resultant chert most commonly forms nodules, but the silica may also replace beds of limestone completely, thereby forming bedded chert. Most bedded chert, however, is probably

turbidite sequences, although additional work is necessary to confirm this hypothesis. Transient, sediment-laden density currents (turbidity currents) that flowed down the sides of submarine basins carried mostly biogenic silica which settled in layers. The layers later transformed into cherts through diagenetic processes. Pelagic or hemipelagic clays were deposited in the basin between times of turbidite deposition, thus creating the rhythmically bedded chert-shale sequences found in many mountain belts. In some places, the shales may be the turbidites and the cherts the pelagic deposits. The rhythmic bedding can also be produced by productivity cycles of biogenic silica. Of relatively minor volumetric importance in post-Archean rocks is the primary precipitation of silica. Inorganic precipitation of silica takes place in hypersaline lakes and lagoons. Chert also forms by deposition of silica from hydrothermal solutions into the interstices between basalt pillows, and in other localized areas within ophiolite sequences. *See* TURBIDITE; TURBIDITY CURRENT.

Distribution. The distribution of chert is tied to the distribution of siliceous plankton, which in turn depends on oceanographic conditions. After the plankton die, their shells accumulate in great quantities on the sea floor, especially around the polar regions (diatoms) and the equatorial belt of high biological productivity (radiolarians and diatoms). Shells also accumulate in other areas where cold, nutrient-rich subsurface water upwells, such as off the west coasts of continents and in marginal seas such as the Philippine and Japan seas. Marine siliceous deposits, along with the underlying oceanic crust, are incorporated into mountain belts by uplift of oceanic crust or by obduction and thrust faulting at subduction plate margins. Such on-land cherts of every age are found at many locations around the world. *See* BACILLARIOPHYCEAE.

Radiolarians and sponge spicules occur in Phanerozoic chert. Diatoms first evolved during the Mesozoic Era and came into prominence in the Tertiary, when they became the dominant precursor of chert. The source of silica for pre-Phanerozoic cherts, which are commonly associated with banded iron formations, is equivocal. Direct evidence of a biological source is lacking, and inorganic precipitation of silica is generally postulated. Most Archean (before 2.5 billion years ago) cherts likely formed from hydrothermal or other abiotic processes. Future studies may yield information that would indicate additional sources of silica for pre-Phanerozoic cherts. *See* SEDIMENTARY ROCKS. James R. Hein

Bibliography. S. R. Aston (ed.), *Silicon Geochemistry and Biogeochemistry*, 1983; P. J. Heaney, C. T. Prewitt, and G. V. Gibbs (eds.), Silica: Physical Behavior, Geochemistry and Materials Applications, *Mineralogical Soc. Amer. Rev. Mineralogy*, vol. 29, Washington, DC, 1994; J. R. Hein, Bibliography of Fine-Grained Siliceous Deposits, *USGS Open-File Rep.*, no. 80–391, 1980; J. R. Hein (ed.), *Siliceous Sedimentary Rock-Hosted Ores and Petroleum*, Van Nostrand Reinhold, New York, 1987; J. R. Hein and J. Obradović (eds.), *Siliceous Deposits of the Tethys and Pacific Regions*, Springer-Verlag, New York, 1989; A. Iijima, J. R. Hein, and R. Siever (eds.), *Siliceous Deposits in the Pacific Region*, 1983.

Chestnut

Any of seven species of deciduous, nut-bearing trees of the genus *Castanea* (order Fagales) native to the Northern Hemisphere and introduced throughout the world. The nuts are actually fruits, with the shells enclosing cotyledons. Trees bear both male and female flowers in late spring but must be cross-pollinated for nut production. Nuts are borne in a spiny involucre or bur that opens to release the nuts in late fall. *See* FAGALES.

Varieties. Japanese chestnuts (*C. crenata*) and Chinese chestnuts (*C. mollissima*) are grown in Asia and the United States for their nuts, and many cultivars have been selected. European chestnuts (*C. sativa*) are an important food source, both cooked whole and ground into flour. They are native to the Caucasus mountains, and distributed throughout southern Europe. American chestnuts (*C. dentata*; see **illus.**) have smaller nuts than Asian or European species and are usually sweeter. Only American trees served as an important source of lumber, because of the length of their unbranched trunks; all chestnut species have been used as a source of tannin for the leather-tanning industry. American and Chinese chinquapins (*C. pumila* and *C. henryi*) have very small nuts that are an important source of food for wildlife. All of the species can be crossed, and hybrids have been selected primarily as orchard cultivars.

Diseases. In 1838 a root disease was reported on European chestnut in Portugal, and soon after this "ink disease" was found in several other parts of southern Europe. The pathogen, *Phytophthora cinnamomi*, causes brownish-black lesions on the roots that exude an inky-blue stain. The lesions expand, coalesce, and girdle the root, resulting in tree death.

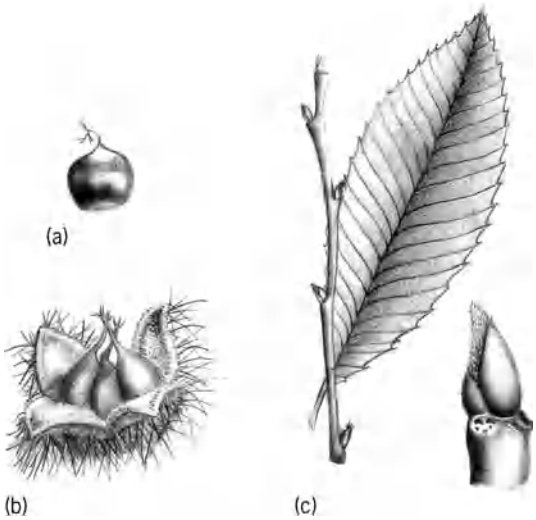

American chestnut (*Castanea dentata*). (*a*) Nut. (*b*) Bur. (*c*) Leaf, twig, and bud.

Breeding with ink disease–resistant Asian species has produced new cultivars, and careful sanitation and site selection allow European chestnuts to still be grown. In the United States, ink disease probably accounted for the recession of chestnut species from large areas in the Gulf and Atlantic states and inland to the foothills and mountains of Mississippi, Alabama, Georgia, Tennessee, Maryland, Virginia, and the Carolinas. The fungus is assumed to have entered the United States in the mid-1800s.

The second disaster for American and European chestnuts was also a fungus disease. *Cryphonectria* (formerly *Endothia*) *parasitica* was first reported in New York City in 1904 and was found to be widespread throughout the northeast region of the United States. This fungus, causing chestnut blight, had probably entered the United States on imported Japanese chestnuts in the late 1800s. The initial spread was on nursery stock (most major mail-order nurseries sold Japanese chestnuts at the turn of the century), but spores were also carried on the fur, feet, and feathers of the birds, animals, and insects that walked up and down the trees. Wounds in the bark allow access and initial fungal growth, and the resulting canker expands to kill the cambium and eventually the twigs, branches, and trunk. The fungus does not kill the root collar, and the dormant embryos there sprout and grow until wounded, infected, and killed, and the cycle is repeated. In the northern part of the natural range of American chestnut (northern Georgia to southern Maine, and west through Ohio, Kentucky, and Tennessee), sprout clumps have survived well, but farther south repeated chestnut blight cycles kill the sprout clumps. American chinquapins are also very susceptible to chestnut blight, and their sprouting system (relying on the old root system, instead of forming new roots as chestnut does) may make them more liable to be killed by blight.

Chestnut blight was first reported in Europe on chestnuts in Italy in 1938. The epidemic there proceeded as in the United States, but natural virus infection of the fungus now provides a biological control. The fungal strains with viruses are less virulent (hypovirulent), and can be inoculated into natural cankers to transmit the viruses and establish biological control in orchards. Hypovirulent strains have been used in the United States for control of blight fungus populations in orchards and forest stands of American chestnut sprouts. Improved control may result from new genetically engineered strains. In addition, trees are bred for resistance by crossing susceptible trees with resistant Chinese and Japanese chestnuts.

The next potential disaster for chestnut trees in the United States and Europe is infestation with a small wasp native to Asia called the Oriental chestnut gall wasp (*Dryocosmus kuriphilus*). Eggs laid in leaf and flower buds develop into galls and heavy infestations can kill the trees. The insect has been in the United States since 1974 (on Chinese chestnut trees in orchards in Georgia), and is now infesting American chestnut trees at the southern part of the range. Japanese chestnut breeders released cultivars with some resistance to the insect, and American chinquapins also seem to have good resistance. *See* FOREST AND FORESTRY; PLANT PATHOLOGY; TREE.

Sandra L. Anagnostakis

Chevrel phases

A series of ternary molybdenum chalcogenide compounds. They were reported by R. Chevrel, M. Sergent, and J. Prigent in 1971. The compounds have the general formula $M_xMo_6X_8$, where M represents any one of a large number (nearly 40) of metallic elements throughout the periodic table; x has values between 1 and 4, depending on the M element; and X is a chalcogen (sulfur, selenium or tellurium). The Chevrel phases are of great interest, largely because of their striking superconducting properties.

Crystal structure. Most of the ternary molybdenum chalcogenides crystallize in a structure in which the unit cell, that is, the repeating unit of the crystal structure, has the overall shape of a rhombohedron with a rhombohedral angle close to 90°. Some of the ternary molybdenum chalcogenides display a slight distortion of the rhombohedral crystal structure at, or below, room temperature to a triclinic structure in which the three axes of the unit cell, and the three angles between them, are unequal.

The building blocks of the Chevrel-phase crystal structure are the M elements and Mo_6X_8 molecular units or clusters. Each Mo_6X_8 unit is a slightly deformed cube with X atoms at the corners, and Mo atoms at the face centers. One of these structures, that of $PbMo_6S_8$, is shown in the **illustration**. *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Superconductivity. Many of the Chevrel-phase compounds exhibit superconductivity, a phenomenon in which a metal, when cooled below its

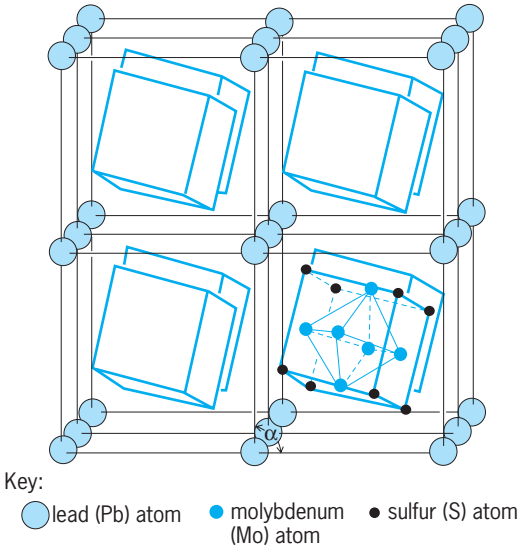

Crystal structure of $PbMo_6S_8$. Each lead atom is surrounded by eight Mo_6S_8 units, the structure of which is shown in the lower right part of the figure. The rhombohedral angle α is indicated also.

superconducting transition temperature T_c, loses all resistance to flow of an electric current. Several of the Chevrel-phase compounds have relatively high values of T_c, the maximum being about 15 K (−433°F) for $PbMo_6S_8$. Prior to 1986, the highest T_c of any known material was 23 K (−418°F) for the A15 compound Nb_3Ge. *See* A15 PHASES.

Since 1987, superconductivity at temperatures above the boiling point of liquid nitrogen (77 K or −321°F) has been discovered in several families of copper oxide compounds. A value of T_c of approximately 122 K (−240°F) has been observed in a compound comprising the elements thallium, barium, calcium, copper, and oxygen. As in conventional superconductors, the superconductivity is associated with electron pairs, called Cooper pairs; but the attractive mechanism responsible for the formation of the Cooper pairs may be different in the high-T_c superconducting oxides than in conventional superconductors, like the Chevrel phases and the A15 phases, where it involves the interaction between the negatively charged electrons and the positively charged ions of the crystal lattice.

The superconductivity of any superconducting compound can be destroyed with a sufficiently high magnetic field, called the upper critical magnetic field $H_{c2}(T)$, which depends on the temperature T and varies from material to material. As the temperature is lowered, $H_{c2}(T)$ generally increases from zero at T_c to a maximum value, $H_{c2}(0)$, as the temperature approaches absolute zero. The Chevrel-phase $PbMo_6S_8$ has a value of $H_{c2}(0)$ of about 60 teslas (600 kilogauss), the largest value observed for any material prior to 1986. The highest $H_{c2}(0)$ that had been previously reported for a non-Chevrel-phase material was 40 teslas (400 kilogauss) for the A15 compound Nb_3Ge. In comparison, the values of $H_{c2}(0)$ for the highest T_c copper oxide superconductors may reach several hundred teslas, values so large that they cannot readily be measured by using the magnetic fields and techniques that are presently available in the laboratory.

A number of Chevrel-phase compounds of the form RMo_6X_8, where R is a rare-earth element with a partially filled *4f* electron shell and X is S or Se, display magnetic order at low temperatures in addition to superconductivity. The superconductivity is primarily associated with the mobile *4d* electrons of Mo, while the magnetic order involves the localized *4f* electrons of the R atoms which occupy regular positions throughout the lattice. Superconductivity has been found to coexist with antiferromagnetic order, but to be destroyed by the onset of ferromagnetic order. *See* ANTIFERROMAGNETISM; FERROMAGNETISM; SUPERCONDUCTIVITY.

M. Brian Maple

Bibliography. G. Burns, *High-Temperature Superconductivity: An Introduction*, 1991; O. Fischer and M. B. Maple (eds.), *Topics in Current Physics*, vol. 32, 1982; T. H. Geballe and J. K. Hulm, Superconductivity: The state that came in from the cold, *Science*, 239:367–374, 1988; H. S. Kwok and D. T. Shaw (eds.), *Superconductivity and Its Applications*, 1988; M. B. Maple and O. Fischer (eds.), *Topics in Current Physics*, vol. 34, 1982.

Chevrotain

The name for four species of mammals which constitute the family Tragulidae in the order Artiodactyla. These animals, also known as mouse deer, are the smallest ruminants, growing to a maximum height of 12 in. (30 cm) at the shoulder. The water chevrotain (*Hyemoschus aquaticus*; see **illustration**) is found in west-central Africa along the banks of rivers in Sierra Leone through the Cameroons to the Congo. It feeds principally on aquatic vegetation. The other species are all members of the genus *Tragulus*, and range through the forested areas of Sumatra into Borneo and Java. These are the Indian chevrotain (*T. meminna*), the larger Malay chevrotain (*T. napu*), and the lesser Malay chevrotain (*T. javanicus*). They are differentiated by the pattern of markings (stripes or spots) on their coats. *Tragulus javanicus* has a coat of uniform color. The chevrotain lacks horns or antlers. There are two well-developed toes on the feet, and the upper canines of the male are elongate and protrude from the mouth as small tusks. The chevrotain is a shy animal which leads a solitary life except during the breeding season. After a gestation period of 120 days, one or two young are born. There is a total of 34 teeth with the dental formula I 0/3, C 1/1, Pm 3/3, M 3/3. The Eocene fossil traguloid, *Archaeomeryx*, which was unearthed in Mongolia, shows many general similarities to the modern chevrotains. The main line of evolutionary development of the traguloids occurred in Eurasia. *See* ARTIODACTYLA; DENTITION.

Charles B. Curtin

Water chevrotain (*Hyemoschus aquaticus*). (*Photo © by Brent Huffman/Ultimate Ungulate Images*)

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Chickenpox and shingles

Chickenpox (varicella) and shingles (herpes zoster) are two different forms of disease caused by the varicella-zoster virus, which is a deoxyribonucleic acid (DNA) virus closely related to herpes simplex and Epstein-Barr viruses. Initial infection causes varicella, a common childhood infection characterized by fever, malaise, and a rash consisting of dozens to hundreds of small fluid-filled lesions (vesicles) that are individually surrounded by reddened skin. Successive crops of lesions appear that eventually ulcerate and crust over during the two-week course of the disease. The virus is spread from person to person by the highly infectious respiratory secretions and lesion drainage. Varicella is rarely a serious disease in normal children but can be severe in immunocompromised individuals or in the rare adult who escaped childhood infection. Primary infection results in immunity to a new varicella-zoster virus, but the original virus lies dormant in nerve ganglia cells. *See* EPSTEIN-BARR VIRUS; HERPES.

At some time in their life, approximately 10% of the population suffers subsequent reactivation of latent virus, which spreads to the skin overlying the affected nerve and causes a localized eruption of vesicles called herpes zoster. The vesicles are similar in appearance and in infectiousness to varicella lesions. This syndrome is usually well tolerated, although elderly persons may develop chronic pain at the site of reactivation. Herpes zoster in immunocompromised individuals may be prolonged or may disseminate to vital organs.

Varicella or herpes zoster in a normal host is self-limited and does not typically require antiviral therapy. In individuals with underlying immune disorders, treatment with the antiviral drug acyclovir decreases the duration and severity of disease. Previously uninfected immunocompromised individuals or pregnant women recently exposed to varicella-zoster virus can be partially protected from subsequent disease through the administration of varicella-zoster immune globulin. However, antibody tests often indicate preexisting immunity from asymptomatic or forgotten childhood varicella. *See* ANIMAL VIRUS; VIRUS INFECTION, LATENT, PERSISTENT, SLOW.

Frederick P. Heinzel

Chicle

A gummy exudate used in the manufacture of chewing gum. It is contained in the bark of a tall evergreen tree, *Manilkara zapota* (see **illus.**), belonging to the sapodilla family (Sapotaceae). The species is a native of Mexico and Central America. The latex (secretion) is collected and carefully boiled to remove excess moisture. When the water content is reduced to 33%, the chicle is poured off and molded into blocks. Crude chicle contains resin, arabin, gutta, sugar, calcium, and different soluble salts. For refining, it is broken up, washed in strong alkali, neutralized with sodium acid phosphate, rewashed, dried, and powdered. The resulting product is an amorphous, pale-pink powder, insoluble in water, and forming a sticky paste when heated. In the manufacture of chewing gum, the chicle is cleaned, filtered, and sterilized, and various flavoring materials and sugar are added. *See* EBENALES; RUBBER.

Fruit of *Manilkara zapota*. (*After L. H. Bailey, The Standard Cyclopedia of Horticulture, vol. 3, Macmillan, 1937*)

Perry D. Strausbaugh; Earl L. Core

Chicory

A perennial herb, *Cichorium intybus* (Asteraceae), with a long taproot, a coarse branching stem, and a basal rosette of numerous leaves. Although the plant is a native of Europe, it has become a common weed in the United States. It is used as a salad plant or for greens. The roasted root is also used as an adulterant of coffee. *See* CAMPANULALES; SPICE AND FLAVORING.

Perry D. Strausbaugh; Earl L. Core

Child-Langmuir law

A law governing space-charge-limited flow of electron current between two plane parallel electrodes in vacuum when the emission velocities of the electrons can be neglected. It is often called the three-halves power law, and is expressed by the formula below.

$$j(\mathrm{A/cm^2}) = \frac{\epsilon}{9\pi}\left(\frac{2e}{m}\right)^{1/2}\frac{V^{3/2}}{d^2}$$

$$= 2.33 \times 10^{-6}\frac{V(\mathrm{volts})^{3/2}}{d(\mathrm{cm})^2}$$

Here ϵ is the dielectric constant of vacuum, $-e$ the charge of the electron, m its mass, V the

potential difference between the two electrodes, d their separation, and j the current density at the collector electrode, or anode. The potential difference V is the applied voltage reduced by the difference in work function of the collector and emitter. The Langmuir-Child law applies, to a close approximation, to other electrode geometries as well. Thus for coaxial cylinders with the inner cylinder the cathode, it leads to a deviation from the true value of the current density of 13% at most. *See* SPACE CHARGE.

Edward G. Ramberg

Chilopoda

A class of the Myriapoda commonly known as the centipedes. As are all myriapods, centipedes are ground dwellers whose permeable cuticle largely restricts their distribution to moist sites and habitats. Centipedes are mostly found under stones, in moss cushions, in leaf litter, and beneath loose bark on trees; several species inhabit caves, especially those of temperate regions. All centipedes are carnivorous; their predatory habits are reflected in their great agility and fleetness and especially in the modification of the first trunk appendages into a pair of usually massive raptorial pincers, the poison claws (also called the forcipules or the prehensors). Each poison claw contains a poison gland whose duct opens near the tip of the claw. Exceptionally, the poison gland may be located in the trunk, in which case the duct connecting it to the apical opening can be as long as twenty body segments. The poison immobilizes or kills arthropod prey and even some small vertebrates. The bite of larger species, particularly of the Scolopendromorpha, may be painful to human beings, but no serious consequences have been reported. *See* ARTHROPODA.

Distribution and taxonomy. The class is divided into two subclasses, the Notostigmophora and the Pleurostigmophora. There are five orders: two are basically tropical (Scutigeromorpha, Scolopendromorpha), two are essentially temperate (Lithobiomorpha, Craterostigmomorpha), and the Geophilomorpha is abundant in both zones. Indeed, there are some Lithobiomorpha and Geophilomorpha that occur as far north as the Arctic Circle, and others at elevations as high as 4000 m (13,123 ft) on the Himalayas.

Fragmentary fossils of Scutigeromorpha are known from the Upper Silurian. Fossil Scolopendromorpha that are very similar to modern types are known from a Carboniferous site. The extinct order Devonobiomorpha, only known for the one species *Devonobius delta* from an upper Middle Devonian site near Gilboa, New York, is probably closely related to the living Craterostigmomorpha.

Morphology. Centipedes have a globular or, more often, flattened head that is distinctly articulated in respect to the long trunk. The head bears a pair of unbranched antennae, which are sometimes very long with up to more than 100 articles. The eyes are of the compound type in the Scutigeromorpha, but are simple ocelli in the other groups; however, there are no eyes in the cryptopid Scolopendromorpha, in the Geophilomorpha, and in many cave-dwelling Lithobiomorpha. The mouthparts include one pair of mandibles and two pairs of maxillae. Excluding the first segment (which bears the poison claws) and the last two or three segments (which are devoid of appendages or bear only short genital appendages), each trunk segment bears one pair of legs. In the adult centipedes, the number of leg-bearing segments is always odd: 15 in the Scutigeromorpha, Lithobiomorpha, and Craterostigmomorpha; 21 or 23 in the Scolopendromorpha; and 27 to 191 (odd numbers only) in the Geophilomorpha. Variation in segment number within the same species is the general rule among the Geophilomorpha (most species in the family Mecistocephalidae being the exception) and has also been reported in one representative of the Scolopendromorpha. Although the trunk segments are uniform ventrally, they are often unequal dorsally. This is most conspicuous in the Lithobiomorpha, in which longer and shorter segments alternate quite regularly except close to midbody.

Gas exchange occurs through lateral spiracles or through dorsal stomata (in Scutigeromorpha only). A circulatory system is present; the copper-containing respiratory pigment hemocyanin has been found in the hemolymph of Scutigeromorpha. Excretion is mainly through a pair of Malpighian tubules similar to those of the insects; in addition, Scutigeromorpha and Lithobiomorpha have a pair of maxillary reins. The genital ducts open at the posterior end of the body. *See* RESPIRATORY PIGMENTS (INVERTEBRATE).

Reproduction and development. All centipedes have separate sexes. The females of Scutigeromorpha and Lithobiomorpha have a pair of pincer-shaped appendages used in egg-laying; otherwise sexual dimorphism is usually limited to the shape of the last pair of legs, which are more elaborate in the male. Males and females of most Scolopendromorpha are virtually indistinguishable without dissection.

Fig. 1. European house centipede, *Scutigera coleoptrata* (order Scutigeromorpha); it is now commonly found in many urban sites in America. (*Photo by Marco Uliana*)

Fig. 2. Scolopender, *Scolopendra cingulata* (order Scolopendromorpha), from the Mediterranean region. (*Photo by Lucio Bonato*)

Although a few instances of parthenogenesis have been described, sexual reproduction among centipedes typically involves the transfer of sperm. Sperm transfer occurs through the production of sperm packets (spermatophores) that the male fixes to the ground (Scutigeromorpha) or, more frequently, places on a small silky web, before inviting the female to collect them. The eggs of centipedes are large. In Scutigeromorpha and Lithobiomorpha, the eggs are manipulated by the female through the posterior pincerlike appendages and finally abandoned, individually, in the soil, covered by soil particles. However, in the Craterostigmomorpha, Scolopendromorpha, and Geophilomorpha, the female not only broods her eggs but also guards the young for a time; without the attention of the female, they would succumb to the fungi and bacteria of the forest floor.

Fig. 3. Geophilomorph centipede, *Clinopodes flavidus*, from northern Italy, illustrating the very elongated, wormlike habit that is characteristic of members of order Geophilomorpha. (*Photo by Lucio Bonato*)

Many centipedes (Scutigeromorpha, Lithobiomorpha, and Craterostigmomorpha) are hemianamorphic, the young hatching with less than the adult number of pairs of legs and continue to molt until the final complement is achieved. However, even once the full complement is attained, a few more molts are required before the animal reaches sexual maturity. Postembryonic development is thus divided into a larval phase, during which the animal achieves the complete number of fully developed segments, and a postlarval phase, during which the segments do not increase further but the animal continues to grow until sexual maturity is reached.

The Scolopendromorpha and Geophilomorpha are epimorphic; that is, upon hatching, the young already possess the adult complement of leg-bearing segments, and successive molts merely permit an increase in size. Centipedes rarely reach sexual maturity in less than 1 year. There is evidence to suggest that all centipedes continue to molt even after they reach sexually maturity and perhaps throughout life; in some species this may entail many years. *See* REPRODUCTION (ANIMAL).

Notostigmophora. The subclass Notostigmophora comprises a single order, Scutigeromorpha. These centipedes are peculiar in embodying primitive as well as highly advanced characters. The group is characterized by the presence of dorsal respiratory openings (stomata); compound-type eyes; long flagellate, multisegmented antennae; and long thin legs with multisegmented tarsi. The trunk has 15 leg-bearing segments but only seven terga. The most familiar member of the order is the common house centipede, *Scutigera coleoptrata*, a species native to southern Europe but accidentally introduced to North America and elsewhere (**Fig. 1**).

Pleurostigmophora. In contrast to the Notostigmophora, the subclass Pleurostigmophora have lateral respiratory openings (or spiracles) and long, branching tracheae similar to those of the insects. Four orders are recognized: Lithobiomorpha (about 1075 species described thus far), Craterostigmomorpha (1 species), Scolopendromorpha (about 810 species), and Geophilomorpha (about 1290 species).

Lithobiomorpha and Craterostigmomorpha. The Lithobiomorpha share some primitive traits with the Scutigeromorpha, with which they were allied in some old classifications: both are anamorphic, have 15 leg-bearing trunk segments, and are rapid runners. However, members of Lithobiomorpha differ in having a flattened rather than globular head, ocelli (if present) rather than compound eyes,

shorter antennae, and, mostly, two-segmented tarsi. The majority of these centipedes are rather small in size, attaining a length of up to only 4 cm (1.6 in.). Although the order is typically temperate in distribution, there are a few mainly tropical or subtropical groups; only a few species are exclusively tropical.

The order Craterostigmomorpha, represented by a single known species inhabiting Tasmania and New Zealand, resembles the Lithobiomorpha in having 15 leg-bearing segments with associated sternites; they differ, notably, in possessing 21 tergites and large, peculiar poison claws.

Scolopendromorpha. The Scolopendromorpha (**Fig. 2**) are the most abundant centipedes in the tropics, where the order exhibits its greatest diversity. They also include the largest centipedes, some of which reach a length close to 30 cm (128 in.). There are also many species, of smaller size, in the temperate zone. The order is divisible into three families: the Scolopendridae, which possess eyes, and the Cryptopidae and Scolopocryptopidae, which are blind.

Geophilomorpha. The Geophilomorpha (**Fig. 3**) represent the most specialized of all centipedes. All are dorsoventrally flattened, and their bodies are long and pliable, permitting them to squeeze through tight openings. All geophilomorphs are blind. The spiracles occur as an unbroken series from the second through the last or next-to-last leg-bearing segment. Unlike most other chilopods, which are voracious hunters that emerge at night to seek their prey in the open, the geophilomorphs are generally found in crevice-cranny habitats; they are adapted to the special requirements of life in the soil, living in the close confines of leaf litter. R. E. Crabill; A. Minelli

Bibliography. J. G. E. Lewis, *The Biology of Centipedes*, Cambridge University Press, 1981; A. Minelli, Chilopoda, in F. W. Harrison and M. E. Rice (eds.), *Microscopic Anatomy of Invertebrates*, vol. 12: *Onychophora, Chilopoda, and Lesser Protostomata*, pp. 57–114, Wiley-Liss, New York, 1993.

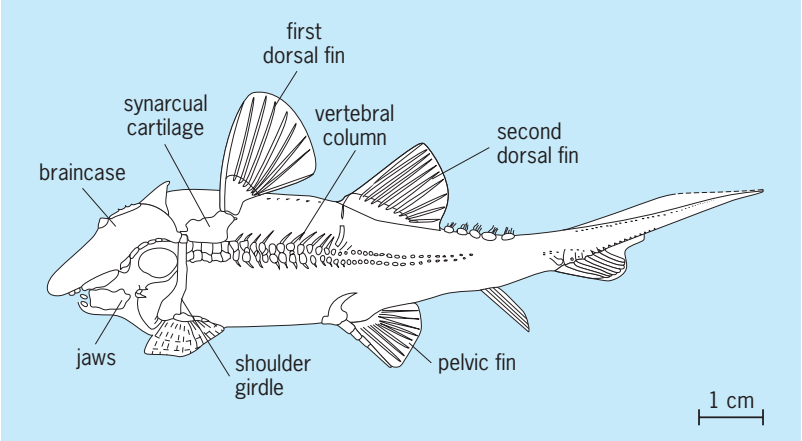

Fig. 1. Lower Carboniferous cochliodontiform *Traquairius*, lateral view. (*After R. Lund, The diversity and relationships of the Holocephali, Proc. 2d Int. Meeting Indo-Pacific Fishes, Ichthyol. Soc. Japan, 1986*)

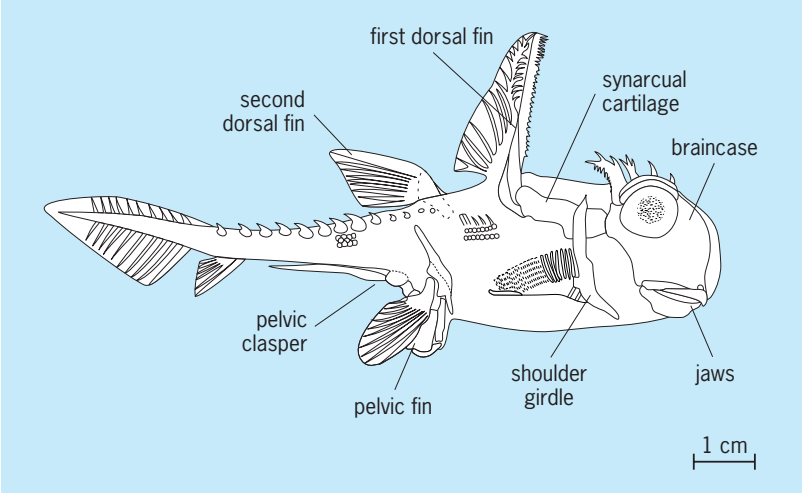

Fig. 2. Lower Carboniferous chimaeriform *Echinochimaera*, lateral view. (*After R. Lund, New Mississippian Holocephalai and the evolution of the Holocephali, Mem. Mus. Nat. Hist. Nat., Paris, 53:195–205, 1988*)

Chimaeriformes

The only extant order of the chondrichthyan subclass Holocephali, and the closest living relatives of the Elasmobranchii (sharks and rays). The chimaeriformes (ratfishes) are a distinctive group of marine fishes that mostly inhabit the deeper parts of the world's oceans, especially in the Atlantic and Pacific. Some species move to coastal waters during the winter months, and a few species may be restricted to these environments. Modern ratfishes are divided into three families, distinguished by the shape of the snout: Chimaeridae, Rhinochimaeridae, and Callorhinchidae. *See* CHONDRICHTHYES.

Morphology. Molecular sequence data and anatomical similarities (such as a cartilaginous and superficially calcified internal skeleton, urea retention mechanism, and clasper organs in the male for internal fertilization) show that chimaeriform fishes and elasmobranchs are more closely related to each other than to any other vertebrate groups. Unlike the elasmobranchs, however, chimaeriforms have the upper jaw fused to the braincase (holostylic condition); a hyoid arch that is not involved in jaw suspension; four gill slits opening into a common outer chamber covered by an opercular skinfold, with a single branchial opening on each side of the head and no spiracle in the adult; massive dental plates instead of individual teeth, with special zones of hard tissue (tritors) separated by softer dentine; a persistent and unconstricted notochord; several ringlike centra per body segment instead of a single one (polyspondylous condition); a narrow, tapering tail; and a dorsal fin with a long, barbed spine just behind the head (**Fig. 1**). A special cartilage (synarcual) connects the back of the braincase to the dorsal fin, and muscles are capable of raising the dorsal fin and its spine (Fig. 1 and **Fig. 2**). The clasper organ (Fig. 2) in males is supplemented by an anterior structure called a tenaculum, which is thought to play a role in copulation, display, or courtship. There is also a large hooklike frontal "clasper" on the head, which may play a similar role. *See* ELASMOBRANCHII. John G. Maisey

Fig. 3. Modern chimaeriform *Hydrolagus colliei* (spotted ratfish). (*Photo © by Joseph Dougherty/www.ecology.org*)

Extant forms. About 31 living species of ratfishes are categorized into six genera and three families, which can be readily distinguished as follows:

Callorhynchidae (plownose chimaeras). The four species in one genus of these chimaeras have an elongate, flexible snout with a hooklike process at the end. They occur in the Southern Hemisphere, off the southern shores of South America, Africa, New Zealand, and Australia.

Chimaeridae (shortnose chimaeras or ratfishes). This family consists of two genera and about 21 species that are known from the Atlantic and Pacific oceans. They have a short, rounded snout and long tapering ratlike tail (**Fig. 3**). The sharp dorsal spine has associated with it a poison gland that produces venom painful to humans.

Rhinochimaeridae. The three genera and six species of this family are similar to the shortnose chimaeras but differ in having a long and pointed snout lacking a hooklike process. They are known from the Atlantic and Pacific oceans, as well as the Indian Ocean off southern Africa.

Most chimaeras feed on mollusks and crustaceans and occasionally on smaller fishes. Herbert Boschung

Fossil record. The fossil record of the chimaeriforms shows that essentially modern forms were present in the Middle Jurassic (175–150 million years ago, ma). Even earlier chimaeriform-like fishes from the Devonian, Carboniferous, and Permian periods (375–250 ma) reveal a surprising variety of morphological diversity, suggesting that they occupied a wide range of exclusively marine habitats (there is no evidence of nonmarine forms). Some had rows of blunt teeth arranged across the jaws (as in sharks) instead of dental plates, while others had continuously growing, curved tooth plates (although these lack the complexity of modern chimaeriform dental plates and do not have tritors). In some, the jaws were not fused to the braincase, as in modern chimaeroids. Many were armored with bonelike plates in the skin, especially around the head. For this reason, it was once thought that chimaeriforms were related to some group of placoderms, although their "armor" is very different in form and structure and the similarities between them are probably either convergent, or represent primitive features of jawed vertebrates generally. Importantly, no placoderms are known to possess the unique features of chimaeriforms. *See* CARBONIFEROUS; DEVONIAN; JURASSIC; PERMIAN; PLACODERMI.

Evolution. The main line of chimaeriform evolution apparently began in the Late Devonian (375 ma). Many spectacular discoveries of chimaeriform-like fossils have been made in the Lower Carboniferous Bear Gulch Limestone of Montana (325 ma). These fossils suggest that several distinctive holocephalan lineages radiated during the late Paleozoic, but only one (Chimaeroidei) survived to the present day. *See* PALEOZOIC.

Two Paleozoic holocephalan orders of particular importance are the Cochliodontiformes and Menaspiformes. These are considered to be successive extinct sister groups to the Chimaeriformes, with which they share one to three pairs of tooth plates. Cochliodontiformes (Late Devonian to Late Permian; worldwide) are mostly known from isolated tooth plates, but complete fossils from Bear Gulch show they had a variety of body forms (Fig. 1). Many cochliodontiforms had a dorsoventrally flattened body; paired mandibular and occipital spines; paired, brushlike frontal spines on either side of the mouth; and "bradyodont" teeth with curved or rounded surfaces, usually with grooves or folds, and multiple small anterior tooth plates in the upper jaw plus two larger posterior tooth pairs in the upper and lower jaws. The branchial arches did not always lie beneath the head as in modern chimaeriforms, but were sometimes behind the head as in elasmobranchs. Some cochliodontiforms lacked a synarcual or dorsal fin, although most had both. Menaspiformes (Early Carboniferous to Permian; North America and Europe) had two pairs of large sickle-shaped frontal or cephalic spines anterior to the orbits, several large dermal "plates" (actually fusions of numerous small tesserae) on the top and sides of the head, holostylic jaws set a distinct angle to the body axis, with the lower jaw rami diverging at about 90°; thin tooth plates, a single pair of lower tooth plates plus anterior teeth at the midline; paired and extremely robust "prepectoral" or "mandibular" spines; a two-part synarcual but no dorsal fin or spine; a complex pattern of dermal scutes, often coalescing to form armor plates; and two rows of enlarged scales on the dorsal part of the trunk (**Fig. 4**).

The order Chimaeriformes includes four suborders: Echinochimaeroidae (Early Carboniferous; North America), Squalorajoidei (Early Jurassic; Europe), Myriacanthoidei (Late Triassic–Late Jurassic; Europe), and Chimaeroidei (Jurassic–Recent, worldwide). Until quite recently, no chimaeriforms were known before the Triassic, but the discovery of *Echinochimaera* in the Bear Gulch Limestone of Montana dramatically extended the stratigraphic

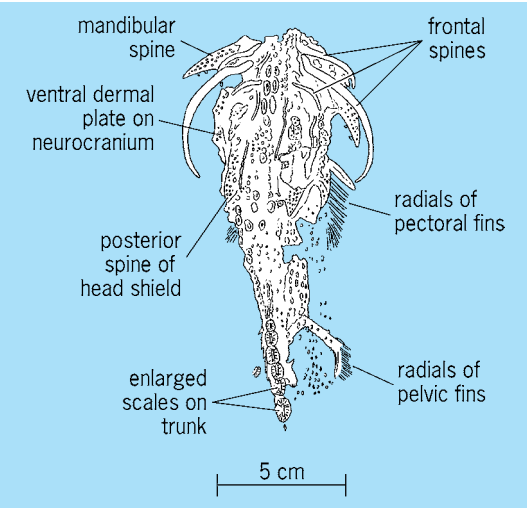

Fig. 4. Permian menaspiform *Menaspis*, dorsal view. (*After C. Patterson, The phylogeny of the chimaeroids, Phil. Trans. Roy. Soc. London, 249:101–219, 1965*)

range of the group (Fig. 2). *Echinochimaera* was remarkably modern in appearance, with a tenaculum and tooth plates (although these lack continuous tritors), but it retained a well-developed second dorsal fin instead of having a narrow, ribbonlike tail. Myriacanthoids resembled modern chimaeriforms in having tritors on at least some tooth plates, polyspondylous ring centra, reduced dermal armor, a single barbed dorsal spine attached to a movable dorsal fin, and a large frontal clasper on the head. Chimaeroids (including modern ratfishes) possess well-developed dental plates, all with specialized tritors; a laterally flattened head; and an ethmoidal canal (an enclosed space in the skull roof above the brain). Apart from the tritors, however, it is unknown whether these characters were also present in other chimaeriforms.

John G. Maisey

Bibliography. D. A. Didier, Phylogenetic systematics of extant chimaeroid fishes (Holocephali, Chimaeroidei), *Amer. Mus. Noviates*, 3119:1–86, 1995; E. D. Grogan and R. Lund, *Debeerius ellefseni* (Fam. Nov., Gen. Nov., Spec. Nov.), an autodiastylic chondrichthyan from the Mississippian Bear Gulch Limestone of Montana (USA), the relationships of the Chondrichthyes, and comments on gnathostome evolution, *J. Morphol.*, 243:219–245, 2000; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994; R. Lund and E. D. Grogan, Relationships of the Chimaeriformes and the basal radiation of the Chondrichthyes, *Rev. Fish Biol. Fisheries*, 7:65–123, 1997; B. J. Stahl, *Chondrichthyes III*, vol. 4: *Holocephali,* in *Handbook of Paleoichthyology*, ed. by H.-P. Schultze, Verlag Dr. F. Pfeil, Munich, 1999.

Chimera

An individual animal or plant made up of cells derived from more than one zygote or otherwise genetically distinct.

Animals

Although some chimeras do arise naturally, most are produced experimentally by mixing cells of very early embryos or by tissue grafting in late embryos or adults. Experimental chimeras have been used to study a number of biological questions, including the origin and fate of cell lineages during embryonic development, immunological self-tolerance, tumor susceptibility, and the nature of malignancy. *See* CELL LINEAGE.

Mixing embryo cells. Two techniques used to form chimeras by mixing embryo cells are aggregation and injection.

Aggregation chimeras. The first experimental mammalian chimeras were produced from mouse embryos in the early 1960s, first by Andrei Tarkowski and later by Beatrice Mintz. The technique involves removing the zonae pellucidae from around 8–16 cell embryos of different strains of mice and pushing the morulae together so that the cells can aggregate. After a short period of laboratory culture, during which the aggregate develops into a single large blastocyst, the embryo is returned to a hormone-primed foster mother (**Fig. 1**). Chimeric offspring are recognized in several ways. If derived from embryos of pigmented and albino strains, they may have strips of pigmented skin and patches of pigment in the eye. Internal chimerism can be detected by use of chromosomal markers or genetically determined enzyme variants. Chimeras accept skin grafts from the two component strains but reject grafts from third-party strains. Aggregation chimeras formed from embryos of strains A and B are denoted A ↔ B, and chimeras developing from two embryos are often known as tetraparental animals. Up to 15 mouse embryos have been aggregated to form a single giant blastocyst. However, the offspring from such experiments are always normal, so that size regulation must occur after implantation. On average, 50% of aggregation

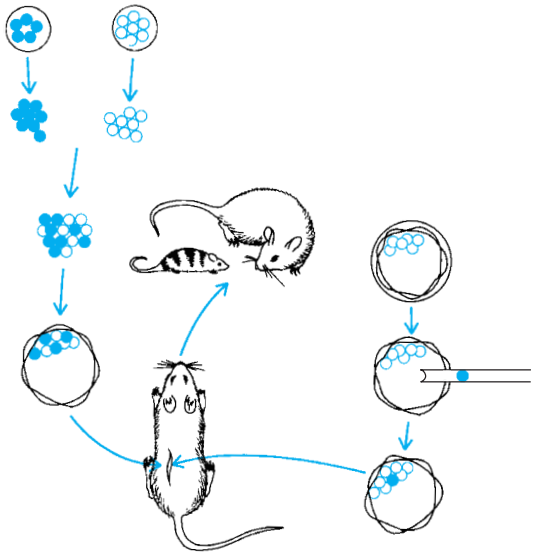

Fig. 1. Techniques for making aggregation chimeras (left) and injection chimeras (right) from early mouse embryos.

chimeras will be XX ↔ XY. Sex determination seems to depend on the relative number of XX and XY cells in the fetal gonads, and in mice most of these animals develop into phenotypic males.

Injection chimeras. This technique was first described for mouse embryos by Richard Gardner in 1968. A blastocyst of the host mouse strain is removed from its zona pellucida and held on a suction pipette. Cells of the donor strain are injected through a fine glass needle, either into the blastocoele cavity or into the center of the inner cell mass (the group of cells from which the fetus is derived). After a short period of culture, the blastocyst is returned to a foster mother. By analogy with aggregation chimeras, injection chimeras are denoted A ↔ B and are recognized by similar, genetically determined markers (pigment formation, and enzyme, serum, and urinary protein variants). A variety of cells have been injected into blastocysts in this way. Cells from adults fail to either grow or become incorporated into the fetus. Early-cleavage-stage embryos injected into the cavity of giant blastocysts (formed by aggregating several embryos in culture as described above) develop inside their zonae into "blastocysts within blastocysts." In contrast, individual cells or groups of cells isolated from embryos of about the same age as the host become incorporated into the host embryo and grow and differentiate along with it. By taking appropriately marked cells from different positions in the donor embryo, placing them in different positions in the host blastocyst, and following their contribution to the late-gestation embryo or adult, important information was obtained about cell interactions and lineages in mouse development.

Another kind of cell—the pluripotent stem cell of mouse teratocarcinomas—was found to give rise to normal tissues in adult chimeras after injection into the mouse blastocyst. Teratocarcinomas are tumors consisting of a disorganized mixture of adult and embryonic tissues. They develop spontaneously from germ cells in the gonads of certain mouse strains, or from cells in early embryos transplanted to ectopic sites. All the differentiated tissues in the tumor arise from pluripotent stem cells known as embryonal carcinoma (EC) cells. When embryonal carcinoma cells are injected into a genetically marked host blastocyst, they continue to divide and participate in normal development, and give rise to fully differentiated cells in all tissues of the adult, including skin, muscle, nerve, kidney, and blood.

Embryonal carcinoma cells from several sources, including spontaneous and embryo-derived tumors and cultured lines selected to carry specific mutations or even human chromosomes, have contributed to normal chimeras. However, embryonal carcinoma cells from some other sources fail to integrate, but produce teratocarcinomas in the newborn animal or adult. The factors controlling successful integration are not yet fully understood, but must include such things as the relative division rate of the embryonal carcinoma cells, their biochemical similarity to cells of the inner cell mass, and their chromosomal constitution. On general grounds, embryonal carcinoma cells should be expected to give rise also to fully functional germ cells, so that EC → blastocyst chimeras might be used to generate nonchimeric mice carrying specific mutations. The fact that certain embryonal carcinoma cells give rise to tumors when injected under the skin or into the body cavity but behave normally in the blastocyst has been used to support the idea that cancers can develop not only as a result of gene mutations but also as a result of disturbances in environmental factors controlling normal cell differentiation (epigenetic theory of cancer). *See* ONCOLOGY.

Interspecies chimeras. A number of interspecies chimeras have been produced by aggregating sheep and goat embryos and reimplanting them in sheep. The offspring have tissues containing a mixture of both sheep and goat cells. The goat, for example, may have patches of both wool and hair. Interspecies chimeras have also been made between rat and mouse, but these do not develop in the uterus, presumably because of immunological rejection by the foster mother. This rejection does not occur in sheep for unknown reasons.

Tissue grafting. Animals that have accepted skin or organ grafts are technically chimeras. Radiation chimeras are produced when an animal is exposed to x-rays, so that blood-forming stem cells in the bone marrow are killed and then replaced by a bone marrow transplant from a genetically different animal. Lymphoid cells in the process of differentiating from stem cells in the donor marrow recognize the recipient as "self" and do not initiate an immune response against the host cells. By removing the thymus from the irradiated host and replacing it with a thymus from another animal prior to the bone marrow transplant, it has been shown that this organ plays an important role in the generation of immunological self-tolerance. *See* TRANSPLANTATION BIOLOGY.

Natural occurrence. Naturally occurring chimeras in humans are rare and are most easily recognized when some cells are XX and others XY. Such individuals are usually hermaphrodite and probably result from fertilization of the egg by one sperm and the second polar body by another, with both diploid cells then contributing to the embryo (the small polar bodies normally degenerate). Blood chimeras are somewhat more common in animals such as cattle where the blood vessels in placentas of twins fuse, so that blood cells can pass from one developing fetus to the other. The female sibling of male and female cattle twins is sterile (a freemartin), because sex hormones from the male fetus interfere with sexual differentiation in the female. Blood chimerism always occurs in marmosets, which produce twins at most pregnancies, but no abnormalities in sexual function are found.

Brigid Hogan

Plants

In modern botanical usage a chimera is a plant consisting of two or more genetically distinct kinds of cells. Chimeras can arise either by a mutation in a cell in some part of the plant where cells divide or by bringing together two different plants so that

their cells multiply side by side to produce a single individual. They are studied not only because they are interesting or ornamental but also because they help in understanding many developmental features of plants that would otherwise be difficult to investigate.

Grafting. The first type of chimera to be used in this way resulted from grafting. Occasionally a bud forms at the junction of the scion and stock, incorporating cells from both; sometimes the cells arrange themselves so that shoots derived from the bud will contain cells from both plants. A plant of this kind is *Laburnocytisus adami*, formed in 1825 by grafting *Cytisus purpureus* onto *Laburnum anagyroides*. The flowers are brown instead of the purple or yellow of the scion and stock. It has the treelike habit of *Laburnum*; its leaves resemble those of *Laburnum* except for the surfaces, which are like those of *Cytisu*; and its seeds always grow into *Laburnum* trees. Occasionally a branch grows out wholly like *Laburnum* or, more rarely, wholly like *Cytisus*. *See* PLANT PROPAGATION.

At first people thought that *Laburnocytisus* and other similar plants derived from grafting must have been nonsexual hybrids, the result of fusing two somatic cells, but there was no evidence that this could occur and result in viable hybrid cells. The occasional reversion to pure *Laburnum* or *Cytisus* branches and the fate of their seeds militated against this theory.

Further investigation showed that the plant consists of a skin of *Cytisus* cells overlying a core of *Laburnum* cells (**Fig. 2**). The skin is only one cell deep, and the flowers appear brown because a purple epidermis covers yellow cells inside the petals. The germ cells which go on to form the embryos in the seeds are always formed from the second layer of the shoot and so always produce *Laburnum* plants. Reversions are due to some buds being formed wholly from the *Cytisus* skin or by the *Laburnum* tissue bursting through the skin in bud formation.

Apical meristem. *Laburnocytisus adami* has been synthesized only once, and all the specimens in existence have been propagated vegetatively from the original. This implies great stability in the constitution of the chimera, and for a long while after the plant's origin, no one knew any mechanism whereby a shoot could maintain such a stable arrangement of genetically diverse cells. In attempts to solve this problem, it was discovered that flowering plants have growing points (apical meristems) where the outer cells are arranged in layers parallel to the surface, and that this periclinal layering is due to the fact that the outer cells divide only anticlinally, that is, by walls perpendicular to the surface of the growing point (**Fig. 3**). In many plants there are two outer (or tunica) layers and, because cell divisions are confined to the anticlinal planes, each layer remains discrete from the other and from the underlying nonlayered tissue called the corpus (Fig. 3). The epidermis of leaves, stems, and petals is derived from the outer layer of the growing point, and the discreteness of this layer accounts for the stability of the *Laburnocytisus* chimera. *See* APICAL MERISTEM.

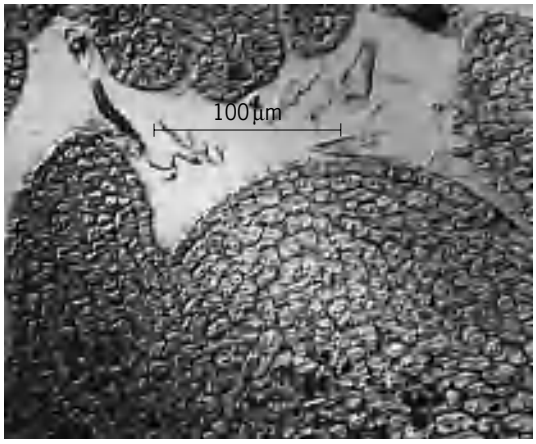

Fig. 2. Section through growing point of *Laburnocytisus adami*.

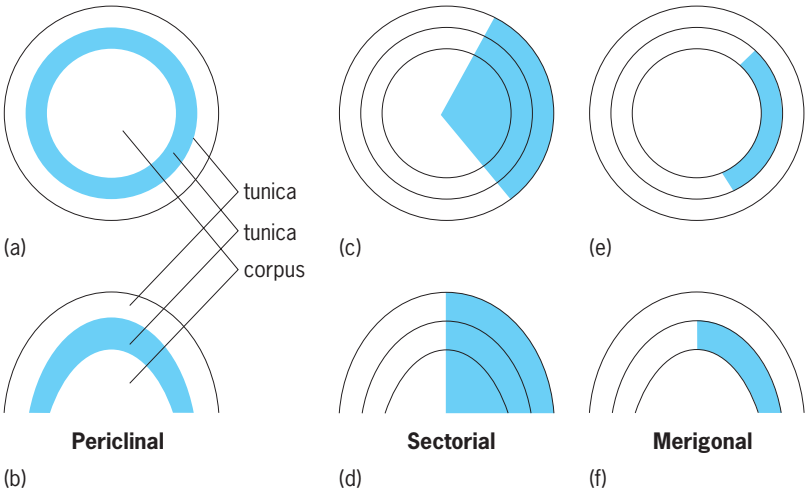

Fig. 3. The two tunica layers and one corpus layer with the three types of chimeras, (*a, b*) periclinal (*c, d*) sectorial, and (*e, f*) merigonal.

An elegant proof of the existence of these self-perpetuating layers in apical meristems came with the synthesis of polyploid chimeras in *Datura* and *Vaccinium*. This was done by treating the embryos or buds with colchicine, a drug that causes nuclei which divide in its presence to double their chromosome number. Since these tetraploid cells and their progeny are larger than ordinary diploid cells, they can be identified in sections cut through the growing point. After the treated meristems had been allowed to grow and produce lateral buds, the buds were sectioned and each tunica layer was seen to contain only cells with the same ploidy; but one layer often consisted of cells of a different ploidy from the other tunica layers and from the corpus. The exclusively anticlinal divisions of the tunica layers clearly prevent the mixing of cells of different ploidy and prevent the chimera from becoming unstable from competition between the cells of different ploidy.

Periclinal chimera. A periclinal chimera has genetically different layers (Fig. 3*a* and *b*). With this type

Fig. 4. Variegated *Pelargonium*, a periclinal chimera whose second tunica layer is genetically white and whose corpus is genetically green.

of chimera, in which the different cells can be identified under the microscope by their relative sizes, it is possible to follow into stems, leaves, and flowers which tissues are derived from each layer in the growing point. For leaves, this can also be done with variegated chimeras, a type of periclinal chimera in which the genetic difference between the cells resides in the plastids due to a mutation that prevents the synthesis of chlorophyll. Tracts of cells whose plastids lack this pigment appear white or yellow instead of green.

A common form of variegated chimera has leaves with white margins and a green center (**Fig. 4**). The white margin is derived from the second layer of the tunica, and the green center from inner cells of the growing point. (The outer tunica layer contributes no color to the leaves of most plants, because it gives rise only to the epidermis, whose cells usually do not develop chloroplasts.) The white leaf tissue overlies the green in the center of the leaf but does not mask the green color. Chimeras with green leaf margins and white centers are usually the result of a genetically green tunica proliferating abnormally at the leaf margin in an otherwise white leaf (**Fig. 5**).

Periclinal chimeras can be used to study the fate of different regions of the growing point, and how the leaves and so forth are constructed; but chimerical plants sometimes do not behave exactly like nonchimerical plants. Nevertheless, they have supplied valuable evidence, sometimes the only evidence, on which to base a view on some developmental problems. Chimeras show, for example, that there is considerable diversity in the number of layers that go to make each kind of organ, and this affects ideas about the homology of different organs.

Sectorial chimera. Since the somatic mutation that initiates most variegated and polyploid chimeras would normally occur in a single cell of a growing point or embryo, it often happens that it is propagated into a tract of mutant cells to form a sector of the plant (Fig. 3*c* and *d*). If the mutation resulted in a failure to form green pigment, the tract would be seen as a white stripe. Such chimeras are called sectorial, but they are normally unstable because there is no mechanism to isolate the mutant sector and, in the flux that occurs in a meristem of growing and dividing cells, one or other of the two sorts of cells takes over its self-perpetuating layer in the growing point. The sectorial chimera therefore becomes nonchimerical or else a periclinal chimera.

Mericlinal chimera. However, in one class of chimeras an isolating mechanism can stabilize the sectorial arrangement (Fig. 3*e* and *f*). This propagates stripes of mutant tissue into the shoot, but because the tunica and corpus are discrete from each other, the plant is not fully sectored and is called a mericlinal chimera.

Many chimeras of this type have a single tunica layer; those with green and white stripes in the leaves have the mutant cells in sectors of the corpus. It was not until C. Thielke tried to find the reason why these chimeras are stable that attention was directed to the fact that they are always plants with leaves in two

Fig. 5. Variegated *Chlorophytum*, a periclinal chimera whose outer cell layer is genetically green and proliferates at the leaf margins to give green tissue in an otherwise genetically white plant.

ranks, and consequently the lateral growth of the growing point occurs by cell expansion only in the plane connecting alternate leaves. This results in the longitudinal divisions of the corpus cells being confined to planes at right angles to the plane containing the leaves. A mutation in one cell therefore can result in a vertical sheet of mutant cells which, in the case of a plastid defect, manifests itself as a white stripe in every future leaf. Transverse sections through the bud confirm that the cells behave as predicted.

Eversports. The eversport is a special type of chimera mainly found in gymnosperms, in which a shoot becomes partially green and partially yellow at a high frequency of about 1 in every 25 nodes. It comes from a periclinal chimera with a single yellow tunica and a green corpus that becomes yellow. This change was explained by H. Derman as coming from a "replacement" periclinal cell division of the topmost cell of the tunica, with one yellow daughter cell remaining in the tunica and the lower yellow daughter cell becoming part of the corpus. This new yellow corpus cell and preexisting green corpus cells proliferate into a partially yellow and partially green corpus, thus seen as a green-yellow chimera (**Fig. 6**).

Recent studies have shown that most eversporting green-yellow chimeras in gymnosperms and dicots eventually become all yellow, leading to the conclusion that the apical meristem has a single apical cell in the top tunica layer similar to what is found in mosses and ferns. Eversports have also been used to understand which regions of a stem give rise to a node and which cells of this nodal region give rise to a leaf.

Roots. The growing points of roots may also become chimerical, but in roots there is no mechanism to isolate genetically different tissues as there is in shoots, and so chimeras are unstable. However, there can be sufficient stability to do experiments with induced root chimeras on the destiny of cells.

Cultivated plants. Since the general acceptance of the existence of organisms with genetically diverse cells, many cultivated plants have been found to be chimeras. This knowledge has explained several phenomena hitherto puzzling. In carnations, for example, white flowers have arisen from red varieties due to the red variety being a periclinal white-red chimera whose normal stability was upset. Flecks of color often indicate the chimerical nature of such plants. Color changes in potato tubers occur similarly because the plants are periclinal chimeras. *See* SOMATIC CELL GENETICS. F. A. L. Clowes; Robert W. Korn

Fig. 6. Eversporting false sawara cypress. Numerous mutant (arrows) green-yellow chimeras are present, most of which become stable mutant shoots.

Bibliography. F. A. L. Clowes, Chimeras and the origin of lateral root primordia in *Zea mays, Ann. Bot.*, 42:801–807, 1978; R. W. Korn, Analysis of shoot apical organization in six species of the Cupressaceae based on chmeric behavior, *Amer. J. Bot.* 88(11):1945–1952, 2001; J. D. Mauseth, *Plant Anatomy*, 1988; R. A. E. Tilney-Bassett, *Plant Chimeras*, 1986.

Chimney

A vertical hollow structure of masonry, steel, or reinforced concrete, built to convey gaseous products of combustion from a building or process facility. A chimney should be high enough to furnish adequate draft and to discharge the products of combustion without causing local air pollution. The height and diameter of a chimney determine the draft.

Construction specifications. For adequate draft, small industrial boilers and home heating systems depend entirely upon the enclosed column of hot gas. A draft is produced because the density of the hot gas inside is lower than that of the air outside. In contrast, stacks, which are chimneys for large power plants and process facilities, usually depend upon force-draft fans and induced-draft fans to produce the draft necessary for operation, and the chimney is used only for removal of the flue gas. Some stacks burn off the gases at the top and are known as flarestacks. *See* FAN.

For fire safety, chimneys for residential construction and for small buildings must extend at least 3 ft (0.9 m) above the level where they pass through the roof and at least 2 ft (0.6 m) higher than any ridge within 10 ft (3 m) of them. Some stacks extend as high as 500 ft (150 m) above ground level, thus providing supplementary natural draft.

A chimney or stack must be designed to withstand lateral loads from wind pressure or seismic forces (earthquakes), as well as vertical loads from its own weight. Wind pressure on a cylindrical surface is about two-thirds that on a plane surface. Thus with a wind pressure of 30 lb/ft^2 (1440 newtons/m^2), the pressure per foot of chimney height is 20 times the diameter, in feet, of the chimney (the pressure in newtons per meter of height is 960 times the diameter in meters). Depending upon geographical location, horizontal inertia loads due to an earthquake can be as high as 35% of the weight of the chimney and can result in greater overturning moments and bending stresses than wind loading. *See* LOADS, DYNAMIC.

Construction materials. Small chimneys used in residential construction are commonly made of brick or unreinforced masonry, while stacks are usually made of steel. In the early 1900s it was generally more economical to construct tall chimneys from masonry or reinforced concrete. However, modern construction costs for tall chimneys favor the use of steel since it allows for more rapid fabrication and erection compared with masonry or reinforced concrete. Where flue gases are particularly corrosive to steel, castable or masonry linings or corrosion-resistant steels (stainless) are often used; or the stack is designed with a corrosion allowance. Sometimes a steel stack has a double steel shell, consisting of concentric cylinders with an air space between the two. The inner shell serves mainly to convect the gases, while the outer shell withstands the wind or seismic loads. *See* REINFORCED CONCRETE; STAINLESS STEEL; STEEL.

Tall steel chimneys of small diameter cannot economically be made self-supporting and must be guyed. Where land costs are at a premium or where land is not available for guying, such as at congested process facilities, it is often more economical to support slender steel stacks with a derrick-type structure of steel latticed or trussed construction as opposed to constructing a self-supported stack. Modern steel stacks are usually shop welded into sections, which are then assembled in the field by flange bolting or welding. *See* JOINT (STRUCTURES); TRUSS.

Concrete chimneys may be plain or reinforced. If the bending moment due to wind or seismic forces on concrete chimneys does not produce excessive tension on the leeward side, reinforcing is unnecessary. To prevent temperature cracking, some plain concrete stacks have a fire-clay lining with an air space between the lining and the outer shell. Masonry or concrete linings are often used in construction of concrete chimneys, resulting in a chimney within a concrete chimney. The outer chimney withstands the lateral loads from wind or earthquake, and the inner one withstands the thermal loads and the acid attacks.

Except for rectangular flues and chimneys commonly used in residential construction, masonry chimneys are usually constructed of perforated radial brick molded to suit the diameter of the chimney. Because the compressive strength of the brick is much greater than that of the mortar, the design strength is determined by the tensile strength of the mortar. Unlike its steel or reinforced counterpart, a brick chimney cannot resist high tensile stresses. *See* BRICK; MASONRY; MORTAR.

Accessory structures. Tall chimneys and stacks, particularly flarestacks, are usually provided with a ladder and rest platforms to allow for inspection and maintenance. Safety regulations require either a cage around the ladder, which increases wind loading, or a safety climb device attached to the ladder to prevent a fall.

Tall stacks are usually fitted with some form of lightning protection and are always grounded at the base. Depending upon the lightning activity in a particular area, a tall stack may be fitted with a lightning dissipator at the top with a special copper line to ground to prevent static buildup and lightning strikes. *See* LIGHTNING AND SURGE PROTECTION.

Cylindrical stacks are subject to a phenomenon known as Kármán vortex oscillations due to the alternating vortices that are shed from opposite sides of the cylinder during gentle winds. Masonry (or gunite-lined stacks) and guyed stacks are generally not susceptible to this phenomenon, because the stack damping ratio is raised in the presence of an internal lining or guying, which tends to abate the buildup of oscillations from the Kármán vortex street. For unlined, self-supported, steel stacks, vortex oscillations are usually controlled by adding strakes or fins to the top third of the stack to spoil the airflow. *See* KÁRMÁN VORTEX STREET.

Foundations. Foundations for tall stacks and chimneys are made from reinforced concrete and are usually octagonal. For very soft soils, the foundation strength is augmented by piles. The design of the foundation is controlled almost entirely by the overturning moment from wind or seismic forces. *See* FOUNDATIONS; PILE FOUNDATION. Joseph Vellozzi

Chinchilla

The name given to two species of rodents which, together with four species of viscachas, make up the family Chinchillidae. These animals are native to several areas of South America and are widely bred on farms in North America and Europe for their fur, which is long, fine, and expensive. The wild population is protected in Chile since the killing of wild stocks in the early 1900s almost led to extinction.

Species. The two species of chinchilla are *Chinchilla brevicaudata* and *C. lanigar*. These animals resemble the squirrel in size and shape and are characterized by long, muscular hindlimbs, with elongate feet bearing four toes, and short forelimbs (see **illus.**). Blunt claws occur on the flexible fingers. These animals have 20 teeth with a dental formula of I 1/1 C 0/0 Pm 1/1 M 3/3. *See* DENTITION.

Chinchillas are gregarious, nocturnally active animals and are found in arid, mountainous regions where they feed principally on vegetation. They often seek shelter in burrows or rock crevices, so

The chinchilla, with more hair per square inch than any other animal, is valued for its fur.

that their capture is difficult. The female, which is larger than the male, bears one to six offspring twice each year after a gestation period of 105–111 days.

Viscacha. There are four species of viscacha belonging to two genera, *Lagidium* (three species) and *Lagostomus*. The mountain viscacha (*Lagidium viscaccia*) lives in burrows near rocks in the foothills of the Andes. It is larger than the chinchilla and has a poor-quality fur, which is woven with wool into cloth. The plains viscacha (*Lagostomus maximus*) inhabits burrows in the pampas region of Argentina. It has coarse fur and is exclusively herbivorous. *See* RODENTIA.

Charles B. Curtin

Chinook

Amild, dry, extremely turbulent westerly wind on the eastern slopes of the Rocky Mountains and closely adjoining plains. The term is an Indian word which means "snow-eater," appropriately applied because of the great effectiveness with which this wind reduces a snow cover by melting or by sublimation. The chinook is a particular instance of a type of wind known as a foehn wind. Foehn winds, initially studied in the Alps, refer to relatively warm, rather dry currents descending the lee slope of any substantial mountain barrier. The dryness is an indirect result of the condensation and precipitation of water from the air during its previous ascent of the windward slope of the mountain range. The warmth is attributable to adiabatic compression, turbulent mixing with potentially warmer air, and the previous release of latent heat of condensation in the air mass and to the turbulent mixing of the surface air with the air of greater heat content aloft.

In winterthe chinook wind sometimes impinges upon much colder stagnant polar air along a sharp front located in the foothills of the Rocky Mountains or on the adjacent plain. Small horizontal oscillations of this front have been known to produce several abrupt temperature rises and falls of as much as 45–54°F (25–30°C) at a given location over a period of a few hours. Damaging winds sometimes occur as gravity waves, which are triggered along the interface between the two air masses.

In the Alpine regions adverse psychological and physiological effects have been noted in humans during prolonged periods of foehn wind. These phenomena have been referred to as foehn sickness. *See* FRONT; ISENTROPIC SURFACES; PRECIPITATION (METEOROLOGY); WIND; WIND STRESS.

Frederick Sanders; Howard B. Bluestein

Chipmunk

The common name for any of 25 species of rodents in the genus *Tamias* of the mammalian family Sciuridae. Chipmunks have bright and dark stripes, large cheek pouches, and live mainly on the ground (see **illustration**). They are found primarily in forested areas in many regions of North America, south to west-central Mexico, as well as in Siberia, Mongolia, and northern and central China. The two species with the widest distribution in North America are the eastern chipmunk (*Tamias striatus*) and the least chipmunk (*T. minimus*). For the most part, the appearance and ecology of all 25 species is similar. *See* MAMMALIA; RODENTIA.

Eastern chipmunk. The eastern chipmunk is a small, fairly stout-bodied, terrestrial squirrel with five conspicuous dark dorsal stripes, a flattened and hairly tail, and large, well-developed cheek pouches. Overall dorsal coloration is tawny or reddish-brown. On each side, a broad whitish longitudinal stripe is bordered by two dark stripes. Two buff stripes are present above and below each eye. The ears are prominent and rounded. Adults are usually 200–255 mm (8.5–10 in.) in total length, including a 75–100-mm (3–4-in.) long tail, and weigh 70–140 g (2.5–5 oz).

Eastern chipmunks live in deciduous forests and brushy areas, especially near broken rocky ground, stone fences, and fallen logs. Although they spend most of their time on the ground, they occasionally climb up stumps or trees to feed or to escape an enemy. (Predators include snakes, weasels, mink, hawks, and domestic cats.) Hearing, sight, and smell are well developed in these solitary and diurnal mammals.

Habitat and hibernation. Chipmunks dig their own burrows, which provide living chambers (including a nest) and a place to store food. They become less active during the hotter summer months and hibernate during the colder winter months. Even though they hibernate, they do not store fat (as do other

Representative species of chipmunk (*Tamias*).

hibernators), so they must awaken periodically to eat. Home ranges are 0.10 to over 1.2 ha (0.04–0.5 acre), depending upon the quality of the habitat and population density. They prefer open-understory and closed-overstory habitats. Home ranges overlap broadly.

Diet. Chipmunks are chiefly vegetarians, feeding mainly on the seeds (including nuts) of woody plants. Corn, wheat, berries, and mushrooms are also eaten along with an occasional earthworm, snail, insect, frog, small snake, bird, and mouse. The large internal cheek pouches greatly assist in transporting food to underground storage chambers or to other cache sites in hollow logs or beneath rocks.

Reproduction and development. Two breeding periods may occur annually in eastern chipmunks, whereas western species have only one breeding season per year. Litters of three to five young are born after a gestation of 31 days. Young chipmunks leave the burrow for the first time at 5 to 6 weeks of age and are probably weaned before they are 40 days old. They are sexually mature at about 3 months, but most do not breed for the first time until they are one year old. Life span in the wild is usually between 2 and 4 years. Donald W. Linzey

Bibliography. D. W. Linzey, *The Mammals of Virginia*, McDonald & Woodward, 1998; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Chiroptera

An order of mammals, commonly known as bats, in which the front limbs are modified as wings, thus making the chiropterans the only truly flying mammals. Bats form the second largest order of living mammals (18 families, 202 genera, 1116 species). They range from the limit of trees in the Northern Hemisphere to the southern tips of Africa, New Zealand, and South America, but most species are confined to the tropics. On many oceanic islands they are the only native land mammals. *See* MAMMALIA.

General characteristics. The wing is formed by webs of skin running from the neck to the wrist (propatagium, or antebrachial membrane), between the greatly elongated second, third, fourth, and fifth fingers (chiropatagium), and from the arm and hand to the body (usually the side) and hindlegs (plagiopatagium). There is also usually a web between the hindlegs (uropatagium, or interfemoral membrane) in which the tail, if present, is usually embedded for at least part of its length (**Figs. 1** and **2**).

Size in bats varies greatly and is well characterized by the length of the forearm (Fig. 1). As familiar reference points, the forearm of *Myotis lucifugus* is about 38 mm long (1.5 in.) and of *Eptesicus fuscus* (**Fig. 3**), about 45 mm (1.8 in.).

Most bats have one young per year, a few have two, and still fewer have three or more. This low reproductive rate is possible because of a long lifespan (up to 34 years in one species) and low predation. Bats have very few natural predators. Most predation is by chance encounters, although some predators may learn to wait at the entrance to a roost and capture bats as they emerge.

Bats are useful as pollinators, as distributors of seeds of tropical plants, and as predators of nocturnal insects. Their prey includes, many agricultural pest insects. Also, they help to maintain the balance of nature, as there are very few other nocturnal insect predators. Although the majority of bats are relatively small and feed on insects, there is great variation in their food habits. Bats can be fruit-eating (there are numerous fruit-eating bats—ranging in size up to the flying fox, with a wingspan reaching nearly 1.5 m or 5 ft); blood-feeding (the vampires); carnivorous (feeding on other small vertebrates—birds, mammals, lizards and frogs); and fish-eating.

Most bats (except pteropids) use echolocation to navigate and to capture insects or other food. Echolocation is similar to radar in that a series of clicks is produced. These clicks bounce off objects, including food items, and return as an echo to the bat, indicating the size, speed, and distance of the object. Most insects are captured by bats in flight using echolocation. Bats catch insects in the air with their mouth, wings, or tail membrane. *See* ECHOLOCATION.

Roost types vary greatly, bats may roost in buildings, in cracks or hollows of trees, under sloughing bark, under bridges, in caves or mines, solitary among foliage, in clusters of leaves, in rolled-up leaves, and in other protected places.

Most temperate bats hibernate during winter; they waken about every 2 weeks for a few hours and then reenter hibernation. About 75% of the fat energy stored for hibernation is used during arousals, leaving only about 25% to support bats throughout hibernation.

The 18 living families of bats may be briefly characterized as follows.

Pteropodidae (Old World fruit bats). These are in general the most primitive of living bats, characterized by more primitive ears and shoulder joints, and are the only family placed in the suborder Megachiroptera. The family, with 42 genera and 186 species, is widely distributed in the tropics and subtropics of the Old World, east to Australia and the Caroline and Cook islands. While some species are quite small, the family includes the largest of all bats, with wingspreads up to 5.5 ft (1.5 m) [Fig. 2]. The external ear of Old World fruit bats is elongate and oval, and there is no tragus. Most still retain a claw on the second digit (absent in all other bats, which constitute the suborder Microchiroptera), as well as the claw on the thumb. The teeth, however, are highly modified for eating fruit or nectar. The tail is short and rudimentary, or absent.

Few have developed the echolocation (radar) mechanism found in all Microchiroptera. One genus emits ultrasonic sounds, but most do not. These bats navigate visually. Many make long flights between feeding and roosting areas. They are slow but strong

flyers. Many of the larger fruit bats congregate in large groups, whereas the smaller ones are usually solitary. Most Old World fruit bats find their food by smell. Most feed on ripe fruit, crushing the pulp in their teeth and spitting out the seeds. These bats may get food by hovering, or they may hold onto a branch with their feet and press the fruit to their chest with a foot to extract it. They sometimes eat flowers. Although some appear to breed at variable times, most breed at well-defined times. Most produce only one young. Some individuals in captivity have lived up to 20 years.

Rhinopomatidae (mouse or long-tailed bats). This insect-eating family, with one genus and four species, is found chiefly in arid regions of northern Africa and southern Asia and is characterized by long wire-like tails and rudimentary nose leaves (dermal outgrowths of unknown function above the nostrils). The tails of rhinopomatids are nearly as long as the head and body combined, and the tail membrane is very short and narrow. Rhinopomatids have large ears connected by skin across the forehead. The tragus (cartilaginous flap at the notch of the ear, Fig. 1) is well developed but rather small. These bats roost in caves, houses, and rock clefts, and are also found in pyramids of Africa. Some roost in very large colonies.

Emballonuridae (sac-winged bats). This insectivorous family includes 13 genera with 51 species and is found in the tropics of both the Old and New worlds. Like the Pteropodidae, but unlike other families of bats, members of Emballonuridae have well-developed bony processes behind the eye sockets. The tail extends only partly across the uropatagium. There is a great deal of variation in this family, with adults varying in weight from 5 to 105 g (0.18 to 3.7 oz). Members lack a nose leaf, but a tragus is present. The tail goes through the tail membrane and appears on the upper surface.

Members of Emballonuridae are commonly known as sac-winged bats because of their glandular wing sacs that open on the upper surface of the wing near the shoulder. These glands secrete a pungent red substance. The sacs are best developed in males and thus may function as a sexual attractant. If present, the sacs are located in the antebranchial membrane (the skin ahead of the upper arm and forearm) and can be seen by partially extending the wing against a light.

Roosting may occur in caves, buildings, ruins, trees, curled leaves, hollow logs, and rock crevices. Although many species form large colonies, some form smaller colonies, and some species are solitary.

Noctilionidae (bulldog or fish-eating bats). This tropical American family, represented by two species in one genus (*Noctilio*), includes a highly specialized fish-eating species. Fish are detected by echolocation and gaffed by the bats' long-clawed feet. The nostrils open forward and down, and the tubular nose extends slightly beyond the lips. There is no nose leaf. The ears are large and pointed, are separate, and a tragus is present. The tail, as in the Emballonuridae, is present but extends only to about the middle of the tail membrane and then perforates the membrane, with the tip being free. The fish-eating bat *Noctilio leporinus* has long hindlegs and large hindfeet with well-developed claws that are useful for catching fish. The other species in the family, *N. albiventris*, is not as well modified for fishing. *Noctilio* is one

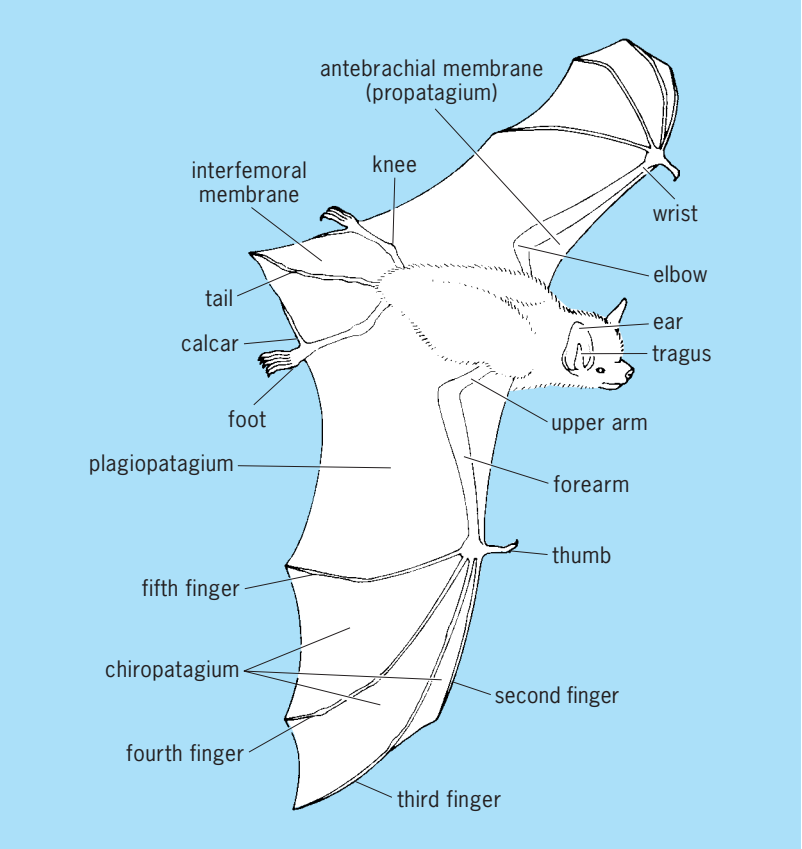

Fig. 1. Anatomical features of bat. (*After R. Peterson, Silently by Night, McGraw-Hill, 1964*)

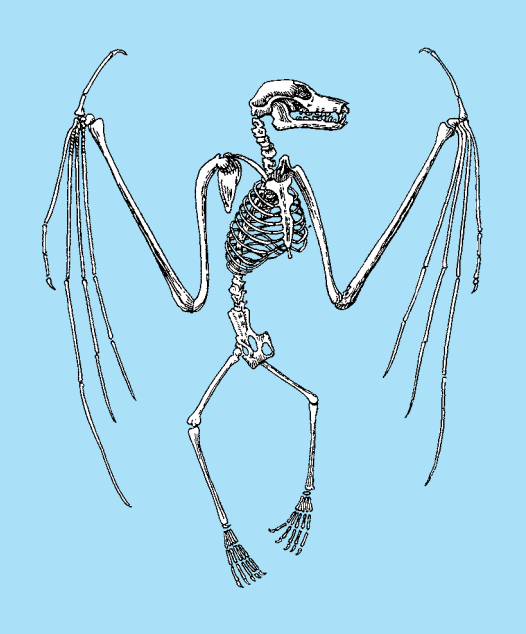

Fig. 2. Skeleton of *Pteropus*, giant fruit bat (lacking tail). (*After R. Peterson, Silently by Night, McGraw-Hill, 1964*)

Fig. 3. ***Eptesicus fuscus*** **(family Vespertilionidae), big brown bat, from Huntington Lake, California. (*Photo by Dr. Lloyd Glenn Ingles; © 1999 California Academy of Sciences*)**

of only three genera of bats that feed on fish. The others are *Myotis* (*M. vivesi* and *M. adversus*) and *Megaderma*. *Noctilio leporinus* apparently uses its echolocation to find fishes and then uses its claws to capture them. Besides fish, it will eat aquatic crustaceans, ants, crickets, scarab beetles, and stinkbugs. *Noctilio albiventris* will also eat fish, although this species has been observed using echolocation to capture insects from the surface of the water.

Nycteridae (slit-faced bats). This family of insect eaters, with one genus and 16 species, is found in Africa and southern Asia. These bats have an extensive basin behind the nose, which is partly bridged over by flaps of skin, leaving a mere slit between them. Slit-faced bats have large ears united at the base. They have a small tragus. They feed mainly on insects, spiders, and sun spiders, but some have been known to eat vertebrates, including fish, frogs, birds, and other bats.

Megadermatidae (false vampire and yellow-winged bats). The four genera (with five species) of this family occur in tropical Africa, southeastern Asia and adjoining islands, and northern Australia. The nose leaf is well developed and simple; the ears are very large and are joined across the forehead. An external tail is lacking. There are no upper incisors. Some species are insect-eating; others feed on small vertebrates, including other bats. The name "false vampire" is from earlier times when these bats were thought to feed on blood.

Rhinolophidae (horseshoe bats). The 77 species (in one genus) of these insect-eating bats are widely distributed in the temperate and tropical areas of the Old World. Horseshoe bats are remarkable for their extremely complex nose leaf expansion of the skin surrounding the nostrils. The lower part is horseshoe-shaped and covers the upper lips, surrounds the nostrils, and has a notch on the lower edge. Above the nostrils is an erect structure (the lancet), connected only at the base and pointed upward. The ears are large but lack a tragus. A tail is present. There are two thoracic teats, but females also have two "dummy teats" on the abdomen that the young grasp as they are being transported. Horseshoe bats roost in many different types of locations, including caves, buildings, trees, and foliage. They feed on insects and spiders. Those in temperate areas hibernate. Some species have delayed fertilization.

Phyllostomidae (New World leaf-nose bats). While most of the 160 species (in 55 genera) of this tropical and subtropical American family possess simple nose leaves (less complex than in Old World leaf-nosed bats), there is tremendous variety in structure, reflecting an equal diversity in food habits. Primitively insect-eating, many have become fruit- or nectar-feeders, and a few are predators on small vertebrates, including other bats. The nose leaf structure is usually present and can be very large, but it may be reduced or absent. The tail may be present or absent. A tragus is present. Phyllostomid bats vary greatly in size, from small species to a false vampire bat, *Vampyrum spectrum*, which is the largest New World bat; its head and body length is about 5 in. (130 cm) and the wingspread can be up to about 3 ft (0.9 m). *Vampyrum spectrum* does not feed on blood. It is carnivorous, feeding on birds, bats, and rodents, and (possibly) on fruit and insects.

The true vampire bats (three species) were previously placed in their own family but are now included among the Phyllostomidae, subfamily Desmodontinae. The vampires are essentially confined to the mainland of tropical America. *Desmodus rotundus* is the best-known species. The vampire bats feed only on fresh blood; *D. rotundus* often feeds on cattle. Vampire bats feed by using the very sharp front incisor teeth (not the canines, as commonly believed) to scratch the surface of the skin so that blood oozes out. There is no sucking of blood; the blood is lapped up with the tongue, but the saliva of the bat contains an anticoagulant which helps keep the blood flowing.

Natalidae (funnel-eared bats). These bats (three genera and eight species) are tropical American insect-eating species with large funnellike ears with glandular papillae on the inside. These are relatively small bats, with the forearm ranging up to about 41 mm (1.7 in.) and the adults weighing 7–10 g (0.25–0.35 oz). The eyes and tragus are small. The tail reaches to the end of the membrane. There is no nose leaf, and the thumb claw is prominent. There is a large "natalid organ" on the face of males which bears a resemblance to sensory cells; it is not always noticeable, and appears to be unique to natalids. These bats often roost in caves, but some have been

found in trees. Natalids are often found in large groups.

Furipteridae (smoky or thumbless bats). This insect-eating family, confined to tropical Central and South America, is represented by two species in two genera. Bats of this family are small (with forearm lengths of only 30–40 mm or 1.2–1.6 in.). The thumb is present but is enclosed in the wing membrane and is so small it appears to be absent. The ears are funnel-shaped and not connected at the base, and the tragus is small. The eyes are close together and point downward. There is no nose leaf, and the tail is long. The third and fourth toes are joined.

Thyropteridae (disk-winged or New World sucker-footed bats). The three species (in one genus) of this insectivorous family are confined to the tropical American mainland. These are small bats, with a forearm length ranging up to 38 mm (1.5 in.). The average weight of *Thyroptera tricolor* is only about 4.2 g (0.15 oz). The large suction disks on short stalks on the thumbs and hindfeet enable them to roost on the smooth inner surfaces of large rolled-up leaves. Unlike the Old World sucker-footed bats, they have a well-developed thumb claw. The tail extends slightly beyond the tail membrane. The third and fourth toes are joined. These bats differ from most others in that they usually hang with the head upward. These bats cling to surfaces, even smooth surfaces, by the suckers.

Myzopodidae (Old World sucker-footed bat). The single species of this family is unique among bats in that it is confined to Madagascar. It eats insects and has suckers on the thumbs and feet rather like those of the Thyropteridae. However, the suckers on this bat are sessile, whereas those of thyropterids are on a short stalk.

Mystacinidae (New Zealand short-tailed bats). The two insect-eating species of this family (both in the genus *Mystacina*) are confined to New Zealand; one (*M. robusta*) is probably extinct, and the other (*M. tuberculata*) is considered vulnerable to extinction. Their body and hindlegs are stout, and they have short tails that do not reach the edge of the uropatagium. The male has a large claw with a projection from it, and the toes also have extra projections. The tail goes through the tail membrane and appears on the upper surface. These bats are very agile, running on the ground and up sloping branches.

Molossidae (free-tailed and mastiff bats). This family includes 100 species in 16 genera and is widely distributed in the tropics and subtropics. They have long tails that extend beyond their uropatagia, and the body and hindlegs are stoutly built. These bats show great size variation, with forearm measuring 27–85 mm (1.1–3.3 in.) in length. The body hair is short and velvetlike in texture.

One genus, *Cheiromeles*, is referred to as the naked or hairless bat. The naked bats are large (with forearm lengths of 70–86 mm or 2.8–3.4 in.), and the hair is so short that the bats appear to be naked. The eyes are small. The ears are leathery and united across the forehead and possess a tragus. There is no nose leaf. The wings are long and thin, and these bats are rapid fliers.

Free-tailed bats feed on insects, and they roost in caves, buildings, tunnels, trees, foliage, and crevices, that is, almost anywhere that bats can roost. They have a strong musky odor and often occur in huge numbers. For example, *Tadarida brasiliensis* number about 20 million in Bracken Cave, Texas, constituting one of the largest vertebrate congregations in the world. They do not hibernate and apparently are generally active throughout the year. They usually have one young per year.

Craseonycteridae (bumblebee bat). Bumblebee bats were described in 1974 as a new species, genus, and family. This family includes only one species. It is from Thailand and is of special interest because it is the smallest bat, and may be the smallest mammal, in the world. Bumblebee bats weigh about 2 g (0.07 oz), or about the same weight as the other two contenders for "world's smallest mammal," the Eurasian pygmy shrew (*Suncus etruscus*) and southern Indiana populations of the American pygmy shrew (*Sorex hoyi*). It apparently feeds on insects and spiders.

Mormoopidae (ghost-face bats). This family includes two genera and 10 species of insectivorous bats, all in the New World. It was formerly placed in the family Phyllostomidae. They occur in tropical lowland habitats from southern Arizona and Texas, south through Mexico, Central America, and northern South America. The largest of these bats, *Pteronotus parnellii*, has a forearm measuring 54–65 mm (2.1–2.6 in.). Mormoopids differ from phyllostomids in lacking the nose leaf. They do, however, have flaps that form a funnel when the mouth is open. The nostrils are in the upper lip, and there are bristles on the outside of the lip flaps that further push air toward the mouth. The eyes of mormoopids are smaller than those of phyllostomids. Members of this family form large colonies in caves.

Hipposideridae (Old World leaf-nosed bats). The hipposiderids include nine genera and 81 species found in the tropics of Asia and Africa, the Philippines, Solomons, and Australia; they were previously considered a subfamily of the Rhinolophidae, but they are currently are recognized as a full family. Old World leaf-nosed bats, like horseshoe bats, have a nose leaf, although the structure differs between the two families. (Members of the genera *Trianops* and *Cloetis* have three leaf points on the nose.) The tail may be absent, but varies greatly in length and may be up to 60 mm (2.4 in.) long. These bats may live in caves, buildings, or hollow trees.

Vespertilionidae (mouse-eared or vespertilionid bats). This nearly cosmopolitan family, with 404 species in 48 genera, occurs almost everywhere that bats occur. All members have a long tail that extends to the edge of the uropatagium. In spite of its many species, the family shows little structural diversity. Most vespertilionids have no special facial modifications. A few have very large ears or small simple nose leaves. Echolocation, present in all families of bats except the Pteropodidae, is perhaps most highly developed in vespertilionids and is used for catching insects as well as avoiding obstacles. Most species are

relatively small, with forearm lengths ranging 22–75 mm (0.9–2.9 in.), and with adults weighing up to 50 g (1.8 oz). Most species have two mammary glands, except for the genus *Lasiurus*, which has four. The many species of this family roost in a very wide variety of locations. Most species appear to have delayed fertilization; however, in a few tropical species, fertilization immediately follows mating. Gestation is usually 40–70 days, but sometimes can last more than 100 days. Almost all are insect-eating, but at least three species of *Myotis* will catch fish. The pallid bat, *Antrozous pallidus*, often comes to the ground to feed on spiders, beetles, and scorpions, and perhaps occasionally even on small vertebrates. This family includes many of the common North American bats, such as *Myotis, Eptesicus* (Fig. 3), *Pipistrellus, Nycticeius*, and *Corynorhinus*.

Fossil record. Bats have a poor fossil record but have been distinct at least since the Eocene, some 50 million years ago. One unusually well-preserved early form, *Icaronycteris*, combined a primitive insect-eating dentition and a retained claw on the second digit with the advanced type of shoulder joint seen in the Vespertilionidae.

John O. Whitaker, Jr.; Karl F. Koopman

Bibliography. R. W. Barbour and W. H. Davis, *Bats of America*, University of Kentucky Press, Lexington, 1969; M. B. Fenton, *Just Bats*, University of Toronto Press, 1983; J. E. Hill and J. D. Smith, *Bats: A Natural History*, University of Texas Press, Austin, 1984; T. H. Kunz (ed.), *Ecology of Bats*, Plenum Press, New York, 1982; T. H. Kunz (ed.), *Ecological and Behavioral Methods for the Study of Bats*, Smithsonian Institution Press, 1987; T. H. Kunz and M. B. Fenton, *Bat Ecology: Advances of the Past 20 Years*, University of Chicago Press, 2003; R. M. Nowak, *Walker's Mammals of the World*, 3 vols., Johns Hopkins University Press, Baltimore, 1999; M. D. Tuttle, *America's Neighborhood Bats*, University of Texas Press, Austin, 1988; J. O. Whitaker, Jr., and W. J. Hamilton, Jr., *Mammals of the Eastern United States*, Cornell University Press, Ithaca, 1998; D. E. Wilson and S. Ruff, *The Smithsonian Book of North American Mammals*, 1999; D. E. Wilson and D. M. Reeder, *Mammal Species of the World: A Taxonomic and Geographic Reference*, 3d ed., 2 vols., Johns Hopkins University Press, 2005.

Chlamydia

A genus of bacteria with a growth cycle differing from that of all other microorganisms. Chlamydiae grow only in living cells and cannot be cultured on artificial media. Although capable of synthesizing macromolecules, they have no system for generating energy; the host cell's energy system fuels the chlamydial metabolic processes. The genome is relatively small; the genomes of *C. pneumoniae* and *C. trachomatis* have been completely sequenced.

Infectious particle. The chlamydial infectious particle, called the elementary body, is round and about 350–450 nanometers in diameter. It enters a susceptible host cell and changes to a metabolically active and larger (approximately 800–1000 nm in diameter) reticulate body that divides by binary fission. The entire growth cycle occurs within a vacuole that segregates the chlamydia from the cytoplasm of the host cell. The reticulate bodies change back to elementary bodies, and then the cell lyses and the infectious particles are released. The growth cycle takes about 48 h.

Chlamydial diseases. Human diseases are caused by three species of *Chlamydia*. *Chlamydia trachomatis*, occurring in humans, is susceptible to sulfonamides and produces inclusions that stain with iodine because they contain glycogen. *Chlamydia psittaci*, occurring in birds, lower animals, and humans, is sulfa-resistant and the inclusions do not stain with iodine. The so-called TWAR strains have been given their own species category, *C. pneumoniae*. The TWAR elementary body is pear-shaped rather than round, and there is little deoxyribonucleic acid (DNA) relatedness. The species *Chlamydia pecorum* is not an important cause of human disease, but infects domestic mammals.

Chlamydia trachomatis. *Chlamydia trachomatis* is almost exclusively a human pathogen, and one of the most common. Those strains infecting humans have no known animal reservoir. Infections occur in two distinct epidemiologic patterns. In many developing countries, *C. trachomatis* causes trachoma, a chronic follicular keratoconjunctivitis. It is the world's leading cause of preventable blindness, affecting approximately 500 million people, with millions losing their sight. In areas where this condition is highly endemic, virtually the entire population is infected within the first few years of life. Most active infections are found in childhood. Scarring can develop as a result of severe inflammation of the conjunctiva. Over a period of years, the scars shrink and distort the upper eyelid, causing an in-turning of the eyelashes which damages the cornea and causes blindness. By age 60, more than 20% of a population can be blinded as a result of trachoma. *See* EYE DISORDERS.

Chlamydia trachomatis is the most common sexually transmitted bacterial pathogen; an estimated 3–4 million cases occur each year in the United States, and there are close to 90 million worldwide. The most common manifestation is nongonococcal urethritis in males. The cervix is the most commonly infected site in women. Ascending infections can occur in either sex, resulting in epididymitis in males or endometritis and salpingitis in females. Chlamydial infection of the fallopian tube can cause late consequences such as infertility and ectopic pregnancy, even though the earlier infection is asymptomatic. For some women, their first knowledge of having had a chlamydial infection occurs when they are being evaluated for infertility or are hospitalized with an ectopic pregnancy. Thus, chlamydial infection in the female can be particularly insidious and dangerous. *See* SEXUALLY TRANSMITTED DISEASES.

The infant passing through the infected birth canal can acquire the infection and may develop either

conjunctivitis (called inclusion conjunctivitis of the newborn, or inclusion blennorrhea) or pneumonia. In the United States, *C. trachomatis* is the leading cause of conjunctivitis in the first month of life, and of pneumonia in the first 6 months.

All of the diseases mentioned above involve mucous membranes. There is a more invasive form of *C. trachomatis* that causes a systemic sexually transmitted disease called lymphogranuloma venereum. Lymphogranuloma venereum has a worldwide distribution, although it is more common in some tropical countries. This disease has a predilection for lymphoid tissue involvement. It occurs in stages: the primary stage involves small, superficial, usually painless genital lesions; in the secondary stage, inguinal buboes (infected lymph nodes) are the common finding. If untreated, the disease may progress to stages involving widespread tissue destruction of genital and intestinal tracts.

Chlamydia psittaci. *Chlamydia psittaci* is virtually ubiquitous among avian species and is a common pathogen among lower mammals. It is economically important in many countries as a cause of abortion in sheep, cattle, and goats. It causes considerable morbidity and mortality in poultry, particularly turkeys and ducks. *Chlamydia psittaci* can infect humans, causing the disease psittacosis. Psittacosis can occur as pneumonia or a febrile toxic disease without respiratory symptoms. Apparently, the chlamydiae that infect turkeys, ducks, and psittacine birds are most infectious for humans. Psittacosis is a potential threat to the health of pet-bird owners and is an occupational hazard in the poultry industry. *Chlamydia psittaci* derived from mammals can also infect humans.

Chlamydia pneumoniae (TWAR). *Chlamydia pneumoniae* appears to be a human pathogen with no animal reservoir. It is of worldwide distribution and may be the most common human chlamydial infection. It appears to be an important cause of respiratory disease, being found in association with sporadic cases of community-acquired pneumonia and epidemics of mild pneumonia. Some evidence has been generated to associate *C. pneumoniae* infections with acute and chronic asthma. In addition, *C. pneumoniae* infection has been linked to coronary artery disease. *Chlamydia pneumoniae* infections start early in life, usually about school age. Infections are likely acquired by the respiratory route. Seroprevalence increases rapidly up to the ages of 30 or 40, where a plateau is approached. Age-specific seroprevalence ultimately reaches 60–70% in many populations. It is believed that asymptomatic, chronic infections are common.

Treatment. Azithromycin is the drug of choice for uncomplicated chlamydial infection of the genital tract. With an antimicrobial spectrum similar to that of erythromycin, it concentrates in cells where the organism grows, and it is effective as a single dose. Two therapeutic agents require longer treatment regimens: doxycycline, a tetracycline antibiotic, is the first alternate treatment; erythromycin may be used for those who are tetracycline-intolerant, as well as for pregnant women or young children. *See* MEDICAL BACTERIOLOGY.

Julius Schachter

Bibliography. J. Boman et al., High prevalence of *Chlamydia pneumoniae* DNA in peripheral blood mononuclear cells in patients with cardiovascular disease and in middle-aged blood donors, *J. Infect. Dis.*, 178:274–277, 1998; L. A. Campbell et al., Detection of *Chlamydia pneumoniae* TWAR in human coronary atherectomy tissues, *J. Infect. Dis.*, 172:585–588, 1995; J. T. Grayston et al., A new respiratory tract pathogen: *Chlamydia pneumoniae* strain TWAR, *J. Infect. Dis.*, 161:618–625, 1990; K. K. Holmes et al. (eds.), *Sexually Transmitted Diseases*, 1999; P. Saikku et al., Serological evidence of an association of a novel *Chlamydia*, TWAR, with chronic coronary heart disease and acute myocardial infarction, *Lancet*, 2:983–986, 1988; J. Schachter and C. R. Dawson, *Human Chlamydial Infections*, 1978; R. S. Stephens et al., Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*, *Science*, 282:754–759, 1998.

Chlorine

A chemical element, Cl, atomic number 17 and atomic weight 35.453. Chlorine exists as a greenish-yellow gas at ordinary temperatures and pressures. It is second in reactivity only to fluorine among the halogen elements, and hence is never found free in nature, except at the elevated temperatures of volcanic gases. It is estimated that 0.045% of the Earth's crust is chlorine. It combines with metals, nonmetals, and organic materials to form hundreds of chlorine compounds, the most important of which are discussed here. *See* PERIODIC TABLE.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Chlorine and its common acid derivative, hydrochloric (or muriatic) acid, were probably noted by experimental investigators as early as the thirteenth century. C. W. Scheele identified chlorine as "dephlogisticated muriatic acid" in 1774, and H. Davy proved that a new element had been found in 1810. Extensive production started 100 years later. During the twentieth century, the amount of chlorine used has been considered a measure of industrial growth.

Physical properties. The atomic weight of naturally occurring chlorine is 35.453 (based on carbon at

12). It is formed of stable isotopes of mass 35 and 37; radioactive isotopes have been made artificially. The diatomic gas has a molecular weight of 70.906. The boiling point of liquid chlorine (golden-yellow in color) is −33.97°C (−29.15°F) at 760 mm Hg (10^2 kilopascals) and the melting point of solid chlorine (tetragonal crystals) is −100.98°C (−149.76°F). The critical temperature is 144°C (292°F); the critical pressure is 78.7 atm; the critical volume is 1.745 ml/g; and density at the critical point is 0.573 g/ml. Thermodynamic properties include heat of sublimation at 7370 ± 10 cal/mole at 0 K, heat of evaporation at 4882 cal/mole at −33.97°C, heat of fusion at 1531 cal/mole, vapor heat capacity at a constant pressure of 1 atm of 8.32 cal/(mole °C) at 0°C (32°F) and 8.46 cal/(mole °C) at 100°C (212°F). Chlorine forms solid hydrates, $Cl_2 \cdot 6H_2O$ (pale-green crystals) and $Cl_2 \cdot 8H_2O$. It hydrolyzes in water as shown in reaction (1).

$$Cl_2 + H_2O \rightarrow HClO + HCl \qquad (1)$$

Chemical properties. Chlorine is one of four closely related chemical elements which have been called the halogen elements. Fluorine is more active chemically, and bromine and iodine are less active. Chlorine replaces iodine and bromine from their salts. It enters into substitution and addition reactions with both organic and inorganic materials. Dry chlorine is somewhat inert, but moist chlorine unites directly with most of the elements. *See* HALOGEN ELEMENTS.

Compounds. Sodium chloride, NaCl, is used directly as mined (rock salt), or as found on the surface, or as brine. It may also be dissolved, purified, and reprecipitated for use in foods or when chemical purity is required. Its main uses are in the production of soda ash and chlorine products. Farm use, refrigeration, dust and ice control, water treatment, food processing, and food preservation are other uses. Calcium chloride, $CaCl_2$, is usually obtained from brines or as a by-product of chemical processing. Its main uses are in road treatment, coal treatment, concrete conditioning, and refrigeration. *See* HALITE.

Wet chlorine reacts with metals to form chlorides, most of which are soluble in water. It also reacts with sulfur and phosphorus and with other halogens as in reactions (2).

$$\begin{aligned} H_2 + Cl_2 &\rightarrow 2HCl \\ 2Fe + 3Cl_2 &\rightarrow 2FeCl_3 \\ 2S + Cl_2 &\rightarrow S_2Cl_2 \\ S + Cl_2 &\rightarrow SCl_2 \\ S + 2Cl_2 &\rightarrow SCl_4 \\ P_4 + 6Cl_2 &\rightarrow 4PCl_3 \\ P_4 + 10Cl_2 &\rightarrow 4PCl_5 \\ Br_2 + Cl_2 &\rightarrow 2BrCl \\ 2F_2 + Cl_2 &\rightarrow ClF + ClF_3 \end{aligned} \qquad (2)$$

The oxides of chlorine, dichlorine monoxide, Cl_2O, chlorine monoxide, ClO, chlorine dioxide, ClO_2, chlorine hexoxide, Cl_2O_6, and chlorine heptoxide, Cl_2O_7, are all made indirectly. Cl_2O is commonly called chlorine monoxide also. Chlorine dioxide, a green gas, has become increasingly important in commercial bleaching of cellulose, water treatment, and waste treatment.

Hydrogen chloride, HCl, is a colorless, pungent, poisonous gas which liquefies at 82 atm at 51°C (124°F). It boils at −85°C (−121°F) at 1 atm (10^2 kPa). Its major production is as the by-product of many organic chlorinations. It can be made by direct reaction of chlorine and hydrogen in an open combustion chamber submerged in cooled, aqueous hydrochloric acid solution. It is used as a strong acid and as a reducing agent.

Aluminum chloride, $AlCl_3$, is an anhydrous, white, deliquescent, hexagonal crystalline substance. Either scrap aluminum or the oxide (bauxite) may be chlorinated. Aluminum chloride is a catalyst for production of cumene, styrene, and isomerized butane. Of the aluminum chloride uses in anhydrous form, ethylbenzene production uses 25%, dyes 30%, detergents 15%, ethyl chloride 10%, drugs 8%, and miscellaneous production 12%. Hydrated and liquid forms are also available, 50% of which are used in drug and cosmetic production.

Ferric chloride, $FeCl_3$, is a solid composed of dark, hexagonal crystals. Much chlorine from chemical processes is converted to ferric chloride, which is then used for the manufacture of salts, pigments, pharmaceuticals, and dyes and for photoengraving, preparation of catalysts, and waste and sewage treatment.

Natural occurrence. Because many inorganic chlorides are quite soluble in water, they are leached out of land areas by rain and ground water to accumulate in the sea or in lakes that have no outlets. Seawater contains 18.97 g of chloride ion per kilogram (3% sodium chloride). Solar evaporation produces large deposits of salts in landlocked areas. Similar evaporation in the past is responsible for vast underground deposits of rock salt and brines in Michigan, central New York, the Gulf Coast of Texas, Stassfurt in Germany, and elsewhere. These deposits are mainly of sodium chloride, the supply of which is unlimited for practical purposes. Other rocks and minerals in the Earth's surface average slightly over 0.03% chloride. *See* HALOGEN MINERALS; SALINE EVAPORITE.

Manufacture. The first electrolytic process was patented in 1851 by Charles Watt in Great Britain. In 1868 Henry Deacon produced chlorine from hydrochloric acid and oxygen at 400°C (750°F) with copper chloride absorbed in pumice stone as a catalyst. The electrolytic cells now used may be classified for the most part as diaphragm and mercury types. Both make caustic (NaOH or KOH), chlorine, and hydrogen. The economics of the chlor-alkali industry mainly involve the balanced marketing or internal use of caustic and chlorine in the same proportions as obtained from the electrolytic cell process.

Uses. Chlorine is an excellent oxidizing agent. Historically, the use of chlorine as a bleaching agent in the paper, pulp, and textile industries and as a germicide for drinking water preparation, swimming pool purification, and hospital sanitation has made

community living possible. Chlorine is used to produce bromine from bromides found in brines and seawater. The automotive age increased the production of bromine tremendously for the manufacture of ethylene dibromide for use in gasoline. Compounds of chlorine are used as bleaching agents, oxidizing agents, solvents, and intermediates in the manufacture of other substances. *See* HERBICIDE.

Joe Downey; Frederick W. Koenker; Robert W. Belfit, Jr.

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. F. Shriver and P. W. Atkins, *Inorganic Chemistry*, 3d ed., 1999.

Chlorite

One of a group of layer silicate minerals, usually green in color, characterized by a perfect cleavage parallel to (001). The cleavage flakes are flexible but inelastic, with a luster varying from pearly or vitreous to dull and earthy. The hardness on the cleavage is about 2.5. The specific gravity varies between 2.6 and 3.3 as a function of composition.

Chlorite is a common accessary mineral in low- to medium-grade regional metamorphic rocks and is the dominant mineral in chlorite schist. It can form by alteration of ferromagnesian minerals in igneous rocks and is found occasionally in pegmatites and vein deposits. It is a common constituent of altered basic rocks and of alteration zones around metallic ore bodies. Chlorite also can form by diagenetic processes in sedimentary rocks. *See* DIAGENESIS.

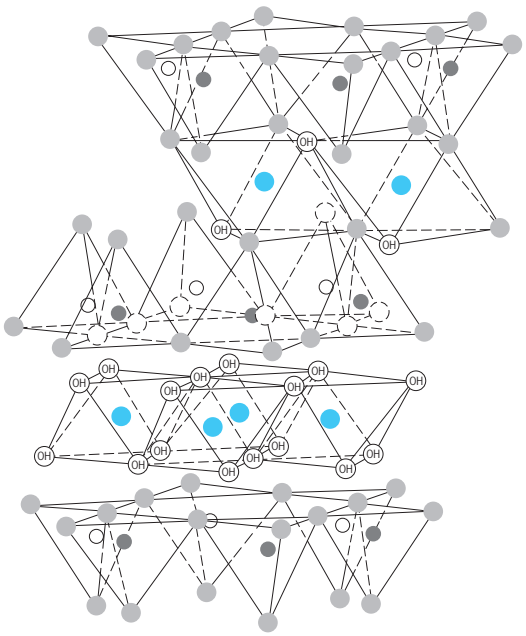

Diagram of the structure of chlorite. (*After R. E. Grim, Clay Mineralogy, McGraw-Hill, 1953*)

The crystal structure of chlorite consists of a 2:1 layer, similar to that in mica, with octahedrally coordinated cations sandwiched between two tetrahedral sheets that are inverted relative to one another. Substitution of Al or Fe^{3+} for Si in the tetrahedral sheets gives a net negative charge to the 2:1 layers, which must be balanced by a positively charged octahedral sheet in the interlayer space. The composition of the 2:1 layer can be expressed as

$$\left(R^{2+}_{3-a-b-c}R^{1+}_{a}R^{3+}_{b}\square_{c}\right)\left(Si_{4-x}R^{3+}_{x}\right)R_{10}(OH)_2$$

and that of the interlayer as

$$\left(R^{2+}_{3-d-e-f}R^{1+}_{d}R^{3+}_{e}\square_{f}\right)(OH)_6$$

with $\square$ = vacancy. Enough R^{3+} cations must be present in the interlayer to create a net positive charge on that sheet. Long hydrogen bonds link basal oxygens of each tetrahedral sheet with surface hydroxyl groups of the interlayer sheet (see **illus.**). Different stacking arrangements of the layers and interlayers are possible. *See* MICA.

Most chlorites are trioctahedral with close to 3.0 octahedral cations present in both the 2:1 layer and the interlayer. The main octahedral cations present in these sheets are Mg, Fe^{2+}, Fe^{3+}, and Al, but with important substitutions of Cr, Mn, Ni, V, Cu, or Li in certain species. Species names are given according to the dominant divalent octahedral cation present: clinochlore for Mg dominant, chamosite for Fe^{2+} dominant, nimite for Ni dominant, and pennatite for Mn^{2+} dominant. A few dioctahedral chlorites are known also. *See* AUTHIGENIC MINERALS; CLAY MINERALS.

S. W. Bailey

Chloritoid

A hydrous iron aluminum silicate mineral with an ideal formula of $Fe_2{}^{2+}Al_4O_2(SiO_4)_2(OH)_4$. Chloritoid occurs as platy, black or dark green crystals, rarely more than a few millimeters in size. Its density ranges from 3.46 to 3.80 g/cm^3, and its hardness on the Mohs scale is 6.5.

Crystal form. Chloritoid most commonly forms monoclinic crystals, having a single twofold rotation axis perpendicular to a mirror plane. This is the symmetry of the prismatic crystal class, symbolized as $2/m$. The atomic structure of chloritoid is characterized by silicon atoms surrounded by four oxygen atoms in a tetrahedral arrangement. These tetrahedral clusters form layers that alternate with layers of ferrous iron (Fe^{+2}) and aluminum (Al; see **illus.**) The iron and aluminum are surrounded by six oxygen (O) and hydroxide (OH) ions in an octahedral arrangement in the ratio $[Fe_2{}^{2+}AlO_2(OH)_4]^-$ and are sandwiched between layers of $(SiO_4)^{4-}$ tetrahedra. The sandwiches of silica tetrahedra–iron and aluminum octahedra–silica tetrahedra are stacked alternately with layers of aluminum. Other structural forms of chloritoid, called polytypes, are created by variations in the stacking of the

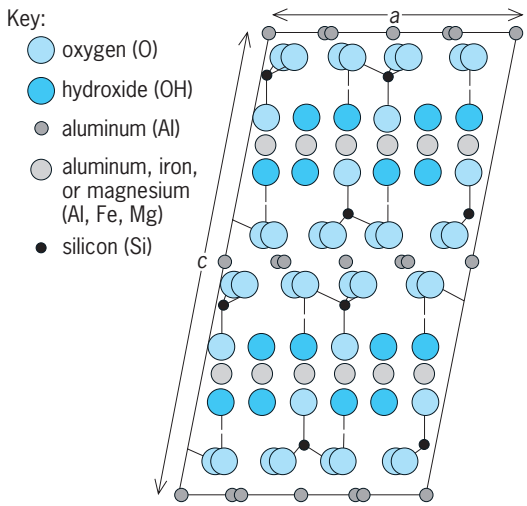

Crystal structure of chloritoid, projected on (010). The unit cell for one structural type of chloritoid, $2M_2$, is shown. The unit cell parameters are $a = 0.9482$ nanometers, $b = 0.5485$ nm, $c = 1.8180$ nm, $\beta = 101.8°$. (*After P. H. Ribbe, ed., Orthosilicates, Reviews in Mineralogy, vol. 5, Mineralogical Society of America, 1980*)

tetrahedral-octahedral-tetrahedral sandwiches. *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

In thin section under the polarizing microscope, chloritoid can appear to be rectangular, with a length two or three times the width. It is observed to be pleochroic, having a color of bluish or greenish gray if the long dimension of the crystal is parallel to the polarizing direction, and pale gray to colorless if the long dimension of the crystal is perpendicular to the polarizing direction. Crystals are commonly twinned across planes parallel to the long dimension. The indices of refraction range from 1.713 to 1.740, with a birefringence (δ) of about 0.010. The mineral has a higher-than-average index of refraction and tends, therefore, to stand out from the other minerals in thin section. Chloritoid exhibits anomalous interference colors in cross-polarized light. *See* BIREFRINGENCE; PLEOCHROISM.

Composition. The composition of chloritoid typically deviates from the ideal formula. Magnesium (Mg) substitutes for as much as 40% of the iron, although the ratio Mg/(Mg + Fe) is generally 0.15 to 0.20. Manganese has been found to substitute for 17% of the iron, but manganese contents are generally less than 1% of the iron. A variety of chloritoid, called ottrelite, contains manganese in substitution for iron in amounts of about 60%. Ferric iron (Fe^{3+}) substitutes for less than 14% of the aluminum, and titanium occurs in trace amounts.

Occurrence. Chloritoid is typically found in low-grade, metamorphosed shales, which are commonly called pelitic schists. It occurs as a comparatively large mineral set in a matrix composed of quartz and fine-grained muscovite. It first appears in most metamorphic terrains near the biotite isograd, that is, near the first appearance of biotite; however, the first appearance of chloritoid depends on many factors, including pressure, oxidation state, and availability of proper rock compositions in the metamorphic terrain. In low-grade rocks, chloritoid is found in association with chlorite, muscovite, and, in some rocks, paragonite. At medium grades of metamorphism, chloritoid can be found with garnet or staurolite in a muscovite and quartz matrix.

The reaction responsible for the synthesis of chloritoid in pelitic schists is not well documented but may occur as in reaction (1), where H_2O may be water or a supercritical fluid phase.

$$12\ \text{pyrophyllite} + 4\ \text{chlorite} \rightarrow 9\ \text{chloritoid} + 49\ \text{quartz} + 10\ H_2O \quad (1)$$

Chloritoid breaks down during metamorphism by either reaction (2) or (3).

$$\text{Chloritoid} + \text{quartz} \rightarrow \text{garnet} + \text{staurolite} + \text{chlorite} + H_2O \quad (2)$$

$$\text{Chloritoid} + \text{muscovite} + \text{quartz} \rightarrow \text{garnet} + \text{staurolite} + \text{biotite} + H_2O \quad (3)$$

The stoichiometric coefficients in these two reactions depend on the Fe/Mg ratios in the minerals, which in turn depend on the pressure and temperature of the reaction. Reaction (2) is found to occur in medium- to high-pressure terrains, whereas reaction (3) occurs in low-pressure terrains.

Experimental studies of the stability of chloritoid have been conducted over a limited range in temperature and pressure and in the possible compositions of chloritoid. The low-temperature stability limit of chloritoid has not been well determined. Calculations indicate that reaction (1) proceeds at a temperature near 400°C (750°F) at 500 megapascals (5 kilobars). The high-temperature stability limit of chloritoid has been studied experimentally under conditions appropriate for reaction (2). The results indicate that chloritoid is stable up to temperatures near 550°C (1020°F) at 500 MPa (5 kbar) and 575°C (1070°F) at 1000 MPa (10 kbar), although the precise temperature and pressure conditions of reaction (2) have not been determined. Reactions at low pressures, including reaction (3), have not been investigated.

Chloritoid is increasingly being recognized as a constituent in rocks that formed under high-pressure conditions. It is found in association with glaucophane in blueschist-facies metamorphic rocks, and with amphibole and pyroxene in eclogite-facies metamorphic rocks. These are rocks that formed under conditions thought to prevail at the base of the Earth's crust or in the mantle. Experimental studies of metamorphism of basalt under these conditions indicate the formation of chloritoid at pressures exceeding 2300 MPa (23 kbar) and at a temperature of 650°C (1200°F). In these high-pressure occurrences of chloritoid, the chloritoid is rich in magnesium, having a value of Mg/(Mg + Fe) ranging from 0.38 to 0.40. *See* METAMORPHIC ROCKS. Theodore C. Labotka

Bibliography. W. A. Deer, R. A. Howie, and J. Zussman, *An Introduction to the Rock-Forming*

Minerals, 1992; C. Klein and C. S. Hurlbut, Jr., *Manual of Mineralogy*, 21st ed., 1993.

Chloromonadida

An order of the class Phytamastigophorea, also known as the Chloromonadina. These poorly known flagellates are grass-green or colorless, somewhat flattened, and have two equal flagella, one anterior, the other trailing. All known genera are free-swimming, although *Reckertia* and *Thaumatomastix* form pseudopodia. They vary in size from 30 to 100 micrometers. Chromatophores are small disks, stigmas are lacking, and fat is the storage product. Trichocysts are found in three genera. The nucleus is large, with nucleoli and chromosomes visible during interphase. One or two anterior vacuoles are present and a reservoir seems

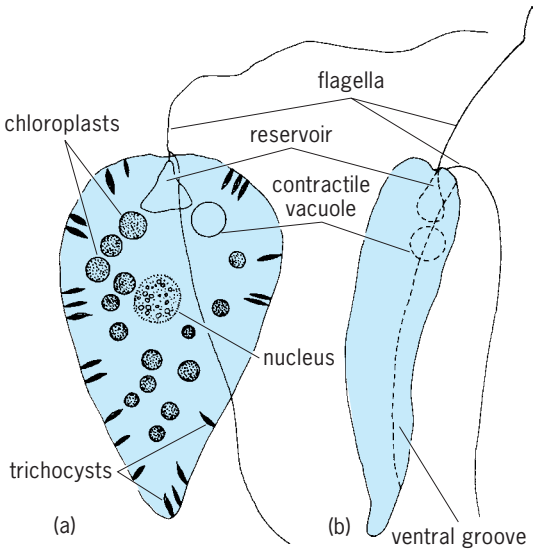

***Gonyostomum semen*. (*a*) Dorsal view. (*b*) Side view.**

to be present in *Trentonia* and *Gonyostomum* (see **illus.**). Life cycles are unknown, but longitudinal division occurs. *Gonyostomum* sometimes occurs as bloom in cedar swamps, and *Vacuolaria* is sometimes common there. The taxonomic position of the class is poorly defined. *See* CILIA AND FLAGELLA; PHYTAMASTIGOPHOREA; PROTOZOA. James B. Lackey

Chlorophyceae

A large and diverse class of plants, commonly called green algae, in the chlorophyll *a-b* phyletic line (Chlorophycota). Estimated number of taxa varies widely; 560 genera and 8600 species are conservative estimates. *See* CHLOROPHYCOTA.

Characteristics. The green algae exhibit great morphological diversity while sharing fundamental biochemical and ultrastructural features. Their photosynthetic pigments are similar to those in higher plants and include chlorophyll *a* and *b*, α-, β-, and γ-carotene, and various xanthophylls, especially lutein, violaxanthin, zeaxanthin, antheraxanthin, and neoxanthin. Chloroplasts vary in number and shape, but always have two membranes, two to five thylakoids per lamella, and usually one or more pyrenoids. A cell wall, which may be calcified, is usually present, and often contains cellulose, hydroxyproline glycosides, xylan, and mannan. The chief food reserve is starch, stored in the chloroplast as granules which often sheathe a pyrenoid. Motile cells usually have two or four apically inserted smooth flagella of approximately equal length. In a few genera, motile cells have more than four apical flagella, one or two pairs of subapical flagella, or a subapical crown of numerous flagella. Flagellar scales and hairs are rare. There is sometimes an eyespot, which lies within the chloroplast. *See* CELL PLASTIDS; CHLOROPHYLL.

Phylogeny. Almost all somatic cell types known for algae occur among the Chlorophyceae, the exceptions being rhizopodial unicells and complex multicellular thalli differentiated into macroscopic organs. In the traditional view of phylogeny within the green algae, the progenitor is a motile unicell (monad) with a wall and a pair of equal, smooth, apical flagella, similar to present-day *Chlamydomonas*. In one line of evolution (volvocine), monads joined to form motile colonies. In a different evolutionary line, monads lost their flagella, forming nonmotile solitary cells and colonies (tetrasporine line), which then underwent specialization and elaboration (chlorococcine line). Multicellular thalli (ulotrichine line) were made possibly by the evolution of desmoschisis, a type of cell division in which the wall of the progeny and that of the parent become an integral structure. Traditionally, the ulotrichine line has been considered to lead from unbranched filaments (Ulotrichales) through forms with both prostrate and erect branched systems (Chaetophorales) to the higher plants, with two very distinctive groups (Zygnematales and Oedogoniales) as early offshoots and a third equally distinctive group (charophytes) diverging from a more advanced position. Siphonous and vesicular green algae (siphonine line) were believed to have evolved prior to the origin of desmoschisis.

Ultrastructural studies, however, have revealed differences in details of nuclear and cell division, cell coverings, plasmodesmata, pyrenoids, and flagellar structure that are correlated among themselves and with certain traditional characters to a degree that strongly suggests that the ulotrichine line is polyphyletic, with desmoschisis having evolved at least four times. Moreover, these studies suggest that the ultimate ancestral green flagellate was prasinophycean (asymmetrical and covered with scales) rather than *Chlamydomonas*-like. Integration of these studies and inferences suggests four phyletic lines: chlorophycean, ulvophycean, charophycean, and pleurastrophycean.

Chlorophycean phyletic line. This phyletic group has motile cells that are usually walled and with two or more apically arising flagella, their basal bodies associated with four or more relatively narrow

bands of microtubules in a cruciate arrangement. The group includes the volvocine, tetrasporine, and chlorococcine lines, as well as the Chlorosarcinales, Microsporales, Sphaeropleales, Oedogoniales, and Chaetophorales. It is believed to be the most derived of the four phyletic lines and the only one with a chlamydomonad ancestor.

Volvocales and Chlorococcales. The order Volvocales includes all Chlorophyceae that normally are flagellate and motile. The cells are solitary or united into colonies of definite structure, called coenobia. Some taxonomists place the unicellular forms in a separate order, Chlamydomonadales, and interpret the coenobial forms as being multicellular. *Chlamydomonas*, presumably similar to the volvocalean archetype, has two smooth flagella, a chloroplast that commonly is cup-shaped or H-shaped, a pyrenoid, a nucleus, an eyespot, and contractile vacuoles. The protoplast is enclosed in a wall (theca). Within the Volvocales there are notable evolutionary trends: isogamy through anisogamy to oogamy; increase in number of cells; and increase in polarity, resulting in morphological and functional differentiation of component cells. *Volvox*, terminating these trends, has a spherical coenobium composed of as many as 50,000 cells. The Tetrasporales, regarded by some taxonomists as a suborder of the Volvocales, are essentially unicellular Volvocales that have lost their flagella and become nonmotile in their vegetative phase. The cells generally have an internal organization similar to that of *Chlamydomonas*. They are usually massed in gelatinous colonies of indefinite shape, but in some genera the thallus is saccate, dendroid, or reticulate. *See* VOLVOCALES.

The order Chlorococcales comprises a large and diverse assemblage of nonmotile algae that appear to follow the Tetrasporales phylogenetically, sharing their inability to form multicellular thalli, but differing cytologically in that vegetative cells lack eyespots and contractile vacuoles and tend to be nonpolar. Most Chlorococcales are solitary cells of various shapes (spherical, ellipsoid, reniform, lunate, acicular, polygonal) and sizes (to 200 micrometers maximum dimension). Cells are free or attached by a gelatinous stalk. The wall is plain or ornamented (spines, bristles, knobs, horns) and uniformly or asymmetrically thickened. In many genera, however, the cells are joined into colonies of indefinite or definite structure. Colonies of indefinite structure result from coalescence of gelatinous envelopes (Coccomyxaceae) or joining of successive progenies by persistent remnants of parental walls (Dictyosphaeriaceae, **Fig. 1***f*). Colonies of definite structure (coenobia) result from coherence of the cells of a progeny, whether the cells are nonmotile (Scenedesmaceae, Fig. 1*i*) or initially motile (Hydrodictyaceae, Fig. 1*d*, *e*, *g*). Asexual reproduction is by cleavage of a protoplast to form autospores or autocolonies (cells or colonies having the same shape as the parent) or zoospores. After being released from the parental wall, autospores and autocolonies enlarge to adult size. Zoospores either metamorphose individually into nonmotile cells or settle collectively to form an autocolony. Sexual reproduction is usually isogamous.

The Chlorococcales constitute an important component of fresh-water plankton and soil microflora, but occupy other aquatic and subaerial habitats, including snow and ice. Some are the algal component (phycobiont) of lichens, while others occur in the tissues of aquatic and terrestrial plants and aquatic animals. *Oophila*, for example, lives in the mucilage of salamander eggs. *Prototheca* is an intestinal pathogen of various domestic animals and humans. *Chlorella*, commonly found in a wide variety of aquatic habitats, is used extensively in physiological research and is grown in mass cultures as a protein source for poultry, livestock, and humans.

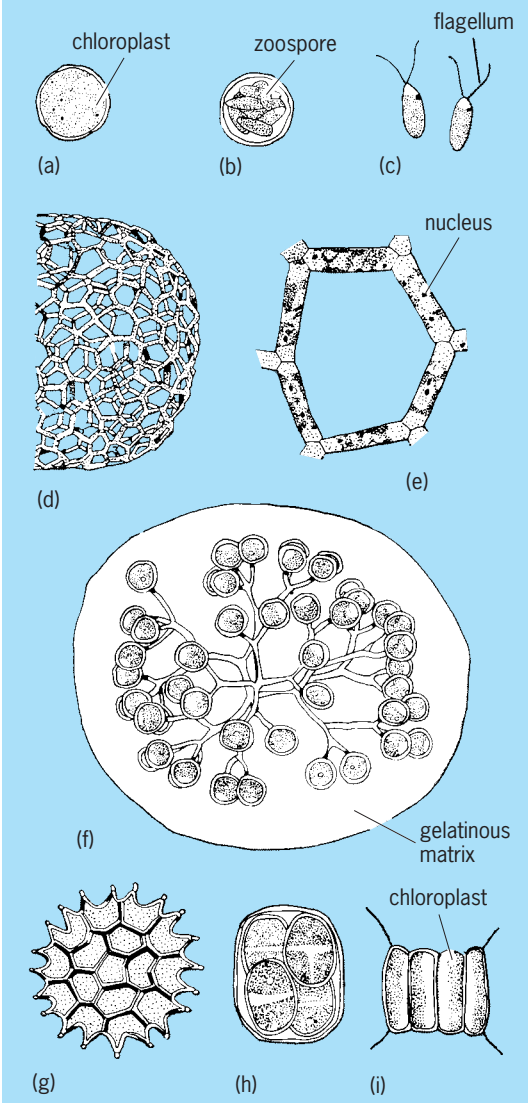

Fig. 1. Chlorococcales. (*a*) *Chlorococcum* (Chlorococcaceae), vegetative cell; (*b*) cell with zoospores; (*c*) zoospores; (*d*) *Hydrodictyon* (Hydrodictyaceae), portion of coenobium; (*e*) portion enlarged; (*f*) *Dictyosphaerium* (Dictyosphaeriaceae), colony in gelatinous matrix; (*g*) *Pediastrum* (Scenedesmaceae), coenobium; (*h*) *Oocystis* (Oocystaceae), colony (*from G. M. Smith, Fresh-water Algae of the United States, 2d ed., McGraw-Hill, 1950*); (*i*) *Scenedesmus* (Scenedesmaceae), coenobium (*from G. M. Smith, Cryptogamic Botany, vol. 1, 2d ed., McGraw-Hill, 1955*).

Nearly 200 genera of Chlorococcales are recognized, distributed among 15 or 16 families. They are almost exclusively freshwater.

In the orders Volvocales, Tetrasporales, and Chlorococcales, cell division is of the type termed eleutheroschisis, in which the daughter protoplasts either remain naked or secrete new walls separate from the parental wall, which ruptures or gelatinizes to release the progeny (zoospores, autospores, or gametes). In theory, multicellularity was made possible by the evolution of desmoschisis, a process by which daughter cells make permanent contributions to the parental wall. In multicellular algae with uninucleate cells, there is a regular sequence of karyokinesis, cytokinesis, and cell-wall deposition. The relationship between daughter and parental walls is not always clear, however; numerous intermediate situations have been demonstrated. A relatively small group of green algae (mainly in soil) exhibit desmoschisis but lack intercellular connections (plasmodesmata) and thus form only cuboidal (sarcinoid) or irregular packets. Both chlorophycean and charophycean phyletic lines are represented, with members of the former assigned to the order Chlorosarcinales and those of the latter to the order Chlorokybales. In a few genera, rudimentary filaments result from end-to-end coherence of daughter cells that do not contribute to the parental wall and thus are equivalent to autospores. Such genera can be assigned to the Chlorococcales. Long but unbranched filaments can also be produced in the absence of desmoschisis of which there are two well-known freshwater examples. In *Microspora* (**Fig. 2***a*), each daughter cell forms a partial wall, so that the filament wall is composed of segments that are H-shaped in optical section, each segment covering the lower half of one cell and the upper half of the subtending cell. The cells are quadrate, each with a reticulate chloroplast lacking pyrenoids. The nucleus lies in a bridge of cytoplasm that crosses the middle of the central vacuole. In *Sphaeroplea* (Fig. 2*b*), daughter cells secrete an entire wall, but the parental wall soon breaks apart. Mature cells are 15–16 times longer than wide and have a curious coenocytic organization, in which a linear series of vacuoles is separated by transverse cytoplasmic bands, each containing several nuclei and a parietal girdle-shaped chloroplast with several pyrenoids or numerous discoid chloroplasts, certain of which contain pyrenoids. Sexual reproduction involves oogamy or an advanced form of anisogamy. *Microspora* and *Sphaeroplea* are traditionally assigned to their own families and, in the present classification, to their own orders.

Oedogoniales. The order Oedogoniales and the coextensive family Oedogoniaceae comprise three genera of filamentous fresh-water algae with several unique features, including (1) an elaborate method of cell division that results in the accumulation of apical caps; (2) zoospores and antherozoids with a subapical crown of numerous flagella; and (3) a highly specialized type of oogamy. *Oedogonium*, with simple filaments, comprises several hundred species that grow attached to various substrates, including stones, wood, and aquatic angiosperms. In *Bulbochaete* and *Oedocladium*, which have fewer species, the filaments are branched. *See* OEDOGONIALES.

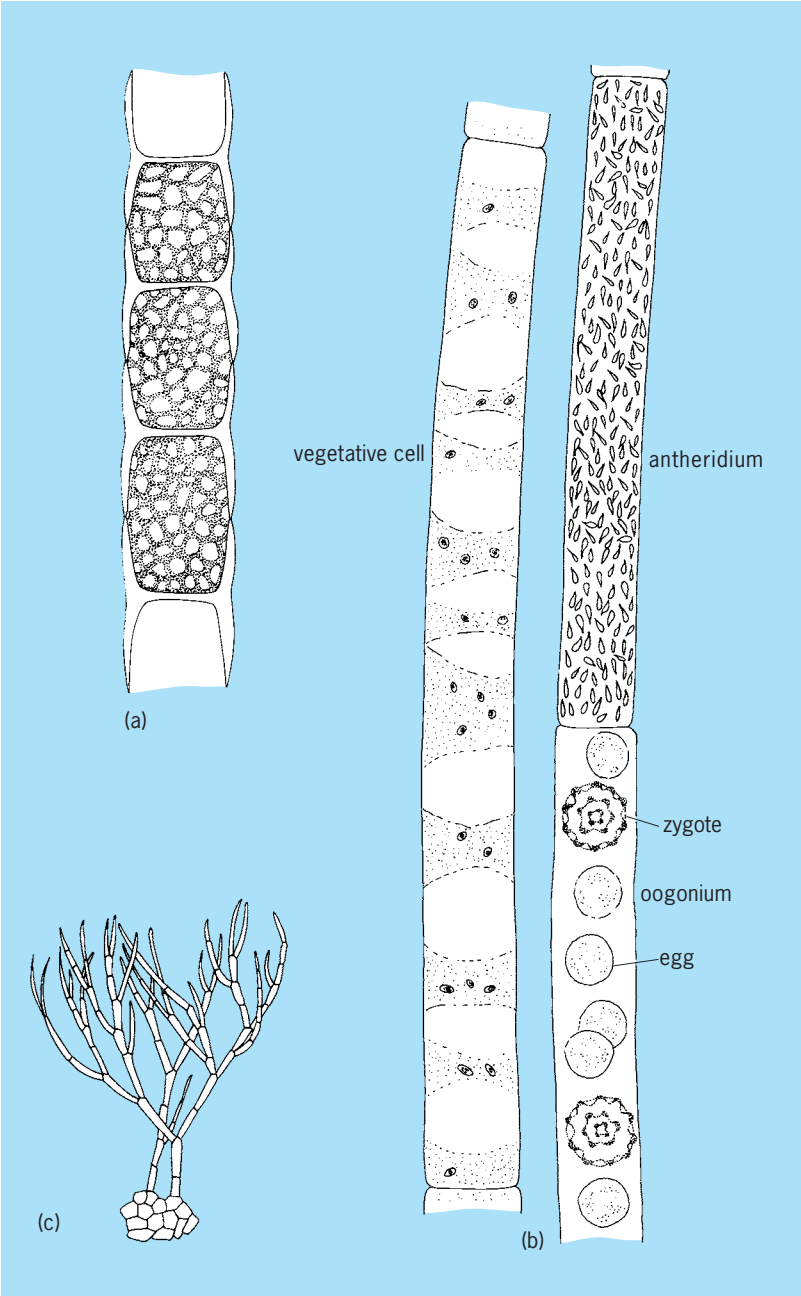

Fig. 2. Filamentous green algae with chlorophycean ultrastructure. (*a*) *Microspora* (Microsporales), showing H-shaped wall segments; (*b*) *Sphaeroplea* (Sphaeropleales); (*c*) *Chaetophora* (Chaetophorales), heterotrichous thallus (erect and prostrate systems).

Chaetophorales. Traditionally, the ulotrichine line was considered to begin with the unbranched filaments of *Ulothrix* and continue through various genera with branched filaments composed of uninucleate cells, each containing a parietal laminate chloroplast. If more than one order (Ulotrichales) was recognized, the branched forms were separated as the Chaetophorales. Modern studies have shown, however, that this morphological spectrum encompasses four phyletic lines as defined by

ultrastructural features. *Ulothrix* is ulvophycean, as are *Trentepohlia* and probably all marine genera previously assigned to the Chaetophorales. *Microthamnion* is pleurastrophycean, while *Coleochaete* is charophycean. Residual chlorophycean forms, which include *Uronema* (unbranched) and *Chaetophora* (branched), constitute the order Chaetophorales *sensu stricto* with about 28 genera distributed among three families. Except for *Uronema*, members of the family Chaetophoraceae typically have a thallus differentiated into a prostrate portion composed of loose or pseudoparenchymatous filaments and an erect, freely branched portion (heterotrichous thallus, Fig. 2*c*). Ultimate branches are usually attenuate, often terminating in a long multicellular hyaline hair. In the Aphanochaetaceae, many cells bear unsheathed unicellular hairs, often with a bulbous base. In *Schizomeris*, the only member of the family Schizomeridaceae, the juvenile is similar to *Ulothrix*, but during maturation all cells in the upper portion of a filament undergo repeated division in various planes, resulting in a parenchymatous cylinder indistinctly marked into bands by the persistent transverse walls of the uniseriate cells.

Ulvophycean phyletic line. Ulvophyceae have motile cells that are scaly or naked, are radially symmetrical, and have two or more flagella arising apically from a cruciate root system, but the distal ends of the basal bodies lie in counterclockwise rotation rather than the clockwise rotation seen in Chlorophyceae. Cytokinesis is effected by a precocious cleavage furrow that impinges on microtubules of the persistent spindle. This phyletic line, which is predominantly marine, includes the Ulotrichales, Ctenocladales, Ulvales, Trentepohliales, Acrosiphoniales, Cladophorales, Siphonocladales, Bryopsidales, Dasycladales, and possibly Cylindrocapsales and Prasiolales.

Ulotrichales. The simplest thallus in the ulvophycean phyletic line is the unbranched filament of *Ulothrix* (**Fig. 3**), the only certain member of the family Ulotrichaceae and the order Ulotrichales. Other genera previously aligned with *Ulothrix* have been shown to be chlorophycean (such as *Uronema*, now in the Chaetophorales) or charophycean (such as *Klebsormidium, Stichococcus*, and *Raphidonema*, now in the Klebsormidiales) or to be reinterpreted as filamentous Chlorococcales (such as *Binuclearia*). Each cell of the *Ulothrix* filament has a nucleus and a parietal laminate or cylindrical chloroplast with one or more pyrenoids. Filaments attach (at least initially) by a modified basal cell to stones or wood, forming extensive stands in various aquatic habitats, such as the shores of lakes and harbors. The life history of fresh-water species involves only one somatic phase, a haploid gametophyte, with meiosis occurring in the production of zoospores during germination of the zygospore. Alternation of a filamentous gametophyte with a filamentous sporophyte has been reported for a marine species.

Ctenocladales. Small green algae with branched filaments are widespread in marine habitats, and have traditionally been assigned to the Chaetophorales, but all prove to have ulvophycean rather than chlorophycean ultrastructure. The order Ctenocladales has been proposed to receive them, but some workers refer them to the Ulvales, an order usually restricted to algae with parenchymatous thalli. In these marine forms, which constitute the family Ulvellaceae, the thallus usually consists solely of prostrate filaments that are free or coherent and pseudoparenchymatous. It is affixed to, or immersed in, a solid substrate, usually an aquatic organism. Free erect filaments, if produced, are short, the cells moniliform or irregularly swollen. Colorless anucleate setalike projections of the cell wall are produced in many genera. Sporangia and gametangia are often

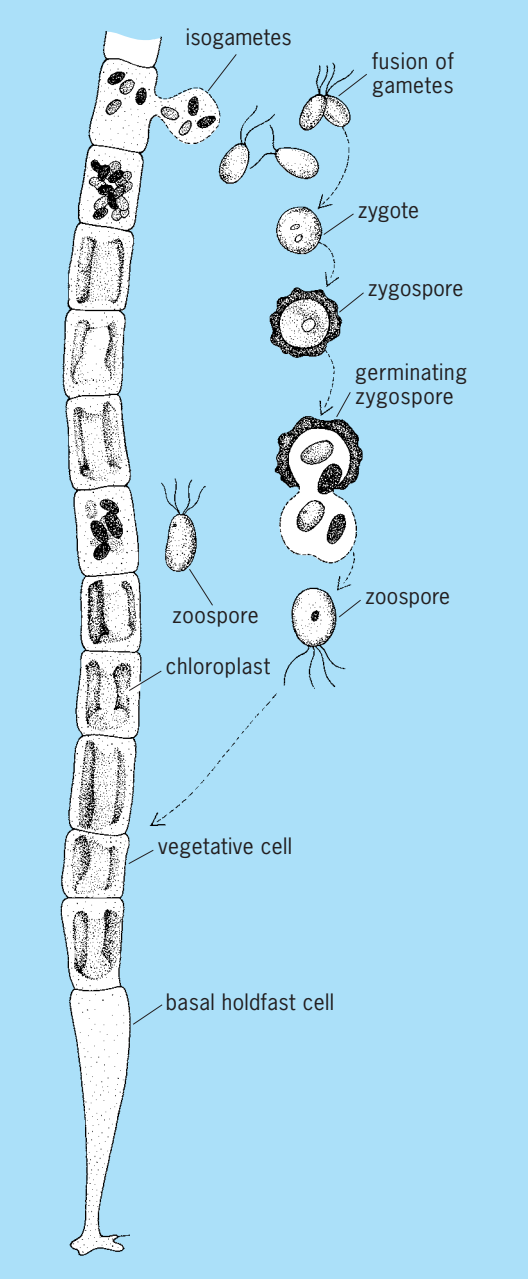

Fig. 3. ***Ulothrix*** **(Ulotrichales).**

conspicuously larger than vegetative cells and undergo sequential cleavage. Swarmers are released by the abrupt rupture of an elongate exit papilla. Various ulvophycean genera that previously were assigned to the Chaetophorales but that lack the special features of the Ulvelaceae are tentatively placed in the Ctenocladaceae.

Ulvales. The order Ulvales comprises green algae with thalli in the form of biseriate filaments, monostromatic blades, tubes, or sacs, and distromatic blades. Cytologically, they are similar to *Ulothrix*, and some workers do not segregate them from the Ulotrichales. Various life histories have been demonstrated. *Ulva* (sea lettuce) and its close relative *Enteromorpha* occur abundantly in the intertidal zone, especially in harbors and estuaries. As treated here, the order comprises five families (Capsosiphonaceae, Percursariaceae, Ulvaceae, Monostromataceae, and Gomontiaceae) with about 12 genera.

Trentepohliales. The order Trentepohliales, with the single family Trentepohliaceae, comprises a small, specialized group of subaerial algae. The thallus is composed of branched uniseriate filaments loosely arranged in an erect tuft or cohering laterally to form a prostrate disk. Sporangia and gametangia are clearly differentiated from vegetative cells. Starch is lacking. Carotenoid pigments are abundant, coloring the thallus yellow-green, yellow, orange, or red, and often imparting the odor of violets. *Trentepohlia*, which has both prostrate and erect systems, grows on rocks, tree trunks, leaves, and wooden structures. *Cephaleuros* grows within the leaves and twigs of various plants in tropical and subtropical regions, including those of commercial plantings of tea and coffee. It is parasitic, inducing a massive wound response in many hosts.

Acrosiphoniales. Thalli in the order Acrosiphoniales are composed of unbranched or branched filaments, each cell usually containing many nuclei and a perforate cylindrical chloroplast with numerous pyrenoids. The life history often includes a vesicular sporophyte derived from the zygote. The vesicle produces haploid quadriflagellate zoospores that give rise to filamentous gametophytes. Gametes and zoospores are liberated through a distinct pore, which is operculate in one family. The order includes two families, Codiolaceae with unbranched filaments and Acrosiphoniaceae with profusely branched filaments, comprising three or four genera restricted to cold marine and brackish waters.

Cladophorales. In the order Cladophorales, the simple or branched filaments have multinucleate cells, as in most Acrosiphoniales, but the characteristic life history is an alternation of isomorphic somatic phases. The cell wall is thick and lamellate. The chloroplast is a parietal network with numerous pyrenoids. *Cladophora* is one of the most common algae in all parts of the world and in all aquatic habitats. Certain species form spherical masses (*Cladophora* balls) that roll to and fro on the bottom of shallow lakes.

Siphonocladales. In the Siphonocladales, an order of tropical marine algae closely related to the Cladophorales, the multinucleate cells are often vesicular rather than cylindrical, and when cylindrical may anastomose in one or more planes. The chief feature of the order supposedly is the occurrence of a unique process of cell division, called segregative division. The protoplast cleaves into multinucleate portions of varying size, which round up, secrete a wall, and expand until they are in complete mutual contact. In some genera the daughter vesicles expand outward and upward, forming branches. The enormous size of the primary vesicle of certain species of *Valonia* has made this genus a favorite subject of biophysical studies, especially in the fields of permeability and solute absorption. In the few species of Siphonocladales in which sexual reproduction has been observed, the thalli are diploid gametophytes, with meiosis occurring during the production of biflagellate isogametes. The failure to observe segregative division in more than a few members of the Siphonocladales and the paucity of information on life histories obscure the border between this order and the Cladophorales. Some authors combine the two orders.

Bryopsidales. Many green algae (almost entirely marine) have thalli composed of nonseptate multinucleate filaments or tubes. These forms, which constitute the order Bryopsidales, are abundant and diverse in most tropical waters. *Halimeda*, which has a calcified thallus, is a major contributor to the buildup of coral reefs and infilling of atoll lagoons. *See* BRYOPSIDALES; HALIMEDA.

Dasycladales. In another group of tropical, essentially marine green algae, the thalli are noncellular, but unlike those in the Bryopsidales, they exhibit radial symmetry, with whorls of loosely arranged or densely aggregated branches arising from the primary axis. These forms constitute the order Dasycladales. Most members have calcified thalli, and there is a rich fossil record extending at least to the Cambrian and including more than a hundred extinct genera. Some genera have a unique life history, involving the fragmentation of an enormous primary nucleus to form numerous small secondary nuclei, which migrate into gametangia. The gametangia cleave into uninucleate units, which develop into operculate cysts containing biflagellate isogametes. Meiosis is believed to occur prior to or during fragmentation of the primary nucleus. *Acetabularia* (mermaid's wine glass) is frequently used in morphogenetic experimentation, since its large and readily accessible nucleus is quite amenable to manipulation. *See* DASYCLADALES.

Charophycean phyletic line. This phyletic group has motile cells that are scaly or naked, asymmetrical, devoid of eyespots, and with two laterally or subapically arising flagella, their basal bodies associated with a broad band of microtubules (multilayered structure) and usually an additional, smaller microtubular root. The distal ends of the basal bodies show counterclockwise rotation. Cytokinesis is effected either by centripetal growth of a furrow that impinges on microtubules of the persistent spindle or by means of a cell plate in the presence of a phragmoplast, the

microtubules of which proliferate from the persistent spindle. This group includes the Chlorokybales, Klebsormidiales, Zygnematales, and Coleochaetales, as well as the charophytes (treated here as the class Charophyceae).

Chlorokybales. Packet-forming algae with charophycean ultrastructure have been segregated from the Chlorosarcinales as the order Chlorokybales. So far only one genus, *Chlorokybus*, is known.

Klebsormidiales. Rudimentary filaments that dissociate easily and lack holdfast cells are formed by members of the order Klebsormidiales. Plasmodesmata are absent. Each cell contains a nucleus and a parietal laminate chloroplast. Zoospores, which lack eyespots, are formed singly in unspecialized cells. Two families (Klebsormidiaceae and Elakatotrichaceae) with at least seven genera are tentatively recognized.

Zygnematales. The order Zygnematales is a large and well-known group of freshwater and subaerial green algae exhibiting two somatic types, unicells (desmids) and unbranched uniseriate filaments. The cell wall is composed of inner fibrillar cellulose layers and an outer slippery pectic layer. Plasmodesmata are absent. Each cell has a nucleus and one or more distinctly shaped chloroplasts. The most distinctive feature of the order is sexual reproduction by conjugation of ameboid or passive gametes (hence the alternate name, Conjugales). Flagellate cells are unknown. The zygote secretes a thick wall and undergoes a period of dormancy. Meiosis occurs at some point (variable, depending upon the species) between conjugation and germination.

Members of the Zygnemataceae (especially *Spirogyra* and *Zygnema*) form the slippery skeins common in stagnant water. Mesotaeniaceae, often called saccoderm desmids, are essentially unicellular Zygnemataceae. The Desmidiaceae and Peniaceae together constitute the placoderm desmids. The cells vary in shape, but are always composed of mirror-image semicells often demarcated by a median constriction. The cell wall has pores and is frequently ornamented. Placoderm desmids are abundant in slightly acidic fresh water. *See* ZYGNEMATALES.

Coleochaetales. Members of the order Coleochaetales are aquatic epiphytes or endophytes. The thallus consists of cells that are clustered or united into filaments. In *Coleochaete*, the only genus in the family Coleochaetaceae, the thallus may be filamentous, pseudoparenchymatous, or rudimentarily parenchymatous. Most or all cells bear one or more cytoplasmic setae, each with a basal gelatinous sheath. Each cell contains a nucleus and usually a parietal laminate chloroplast with a pyrenoid traversed by one or more thylakoids. Asexual reproduction is by scaly biflagellate zoospores, lacking eyespots and formed within ordinary cells. Sexual reproduction is oogamous, with minute colorless antherozoids being formed singly in conical antheridia, and eggs singly in spherical or flask-shaped oogonia. Fertilization occurs in place. The zygote enlarges greatly and secretes a thick wall. Neighboring branches produce a pseudoparenchymatous or parenchymatous sheath around the zygote, suggesting the possibility of the transfer of substances. After a period of dormancy (winter), the zygote germinates meiotically to form 8–32 biflagellate zoospores. Although exhibiting features conducive to the development of a terrestrial life, *Coleochaete* probably is not in the direct line of evolution of any extant embryophyte.

Pleurastrophycean phyletic line. Pleurastrophyceae have motile cells that are thecate or naked and more or less flattened. The flagella have a cruciate microtubular root system showing counterclockwise rotation of the basal bodies. The ancestral flagellate is believed to resemble *Tetraselmis*, a marine genus usually aligned with Prasinophyceae. The Pleurastrophyceae, while apparently a distinct group, is small. The few genera that have been assigned to it have been taken from several orders, including Chlorococcales, Chlorosarcinales, and Chaetophorales.

Micromonadophyceae. In addition to the four phyletic lines described here, each of which includes monads as well as multicellular algae, there is a residual group of primitive green flagellates for which the name Micromonadophyceae has been proposed. These monads are scaly or secondarily naked and have one to four flagella arising from an apical pit or lateral depression. The interzonal mitotic spindle persists during cytokinesis. This group corresponds to the class Prasinophyceae in the present classification.

The ultrastructural classification described in this article could be incorporated in the traditional scheme with minimal difficulty by according the rank of subclass to each major group, a treatment in keeping with the presumed monophyletic origin of green algae. For the sake of consistency, however, Prasinophyceae and Charophyceae are treated elsewhere in this work as separate classes. *See* ALGAE; CHAROPHYCEAE; PRASINOPHYCEAE.

Paul C. Silva; Richard L. Moe

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae: Structure and Reproduction*, 2d ed., 1997; P. Bourrelly, *Les Algues d'Eau Douce . . .*, tome 1: *Les Algues Vertes*, 1990; K. R. Mattox and K. D. Stewart, Classification of the green algae: A concept based on comparative cytology, in D. E. G. Irvine and D. M. John (eds.), *Systematics of the Green Algae*, 1984; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Chlorophycota

A division of the plant kingdom (also known as Chlorophyta or Chlorophycophyta) comprising all algae that have chlorophyll *a* and *b* except the Euglenophyceae, and that in all other respects are so different as to suggest a separate origin of their photosynthetic pigments. Three classes, Charophyceae (charophytes), Chlorophyceae (green algae), and Prasinophyceae, are recognized. *See* ALGAE; CHAROPHYCEAE; CHLOROPHYCEAE; PRASINOPHYCEAE.

Paul C. Silva; Richard L. Moe

Chlorophyll

The generic name for the intensely green, blue, or purple tetrapyrrolic pigments found in higher plants, algae, and photosynthetic bacteria that serve as photoreceptors and transducers of light energy to chemical energy in photosynthesis. *See* PHOTOSYNTHESIS.

Occurrence and Distribution of Major Cholorophylls

Although only about 20 chlorophylls are discussed here, new chlorophylls are found every year, especially in marine microalgae and photosynthetic bacteria of lakes and tidal flats. Chlorophylls *a*, *b*, and *d* are found in oxygenic photosynthetic organisms. All members of the chlorophyll *c* group are found in oxygenic, marine microalgae. Bacteriochlorophylls *c*, *d*, *e*, and *g* are found in anoxygenic (non-oxygen-producing) photosynthetic bacteria. All members of the chlorophyll *c* group are found in oxygenic, marine microalgae.

Chlorophylls a and b. The principal chlorophylls of higher plants and green algae are chlorophylls *a* and *b*. Chlorophyll *a* is found in the light-harvesting pigment-protein complexes (also called antennae) and reaction-center complexes (where the chemistry takes place) of all oxygen-evolving organisms, while chlorophyll *b* is found only in the antennae.

In higher plants, chlorophylls *a* and *b* are contained in the chloroplast (a type of cell plastid). At the highest magnification of the light microscope, one may just see tiny grana in the plastids of higher plants. A granum is made up of 10–100 discs (known as thylakoids), which are the basic photosynthetic apparatus. Chlorophylls *a* and *b* are attached to specific proteins located in the thylakoid membrane. In addition to the outer light-harvesting antennae complexes, two functionally and structurally distinguishable groups of chlorophyll-protein complexes exist: photosystem I and photosystem II. Each photosystem, with the associated antenna, contains about 300 chlorophyll molecules (and a few carotenoid molecules). The majority of these chlorophylls serve as light-harvesting antennae and funnel the energy they absorb to the special reaction-center chlorophylls of photosystem I and II. These reaction-center chlorophylls are special as they are two chlorophyll *a* molecules held in close juxtaposition by a unique protein environment. *See* CAROTENOID; CELL PLASTIDS; PLANT CELL.

Chlorophyll c. The chlorophyll *c* group is a rapidly expanding group of at least eight antenna chlorophylls. All members of the chlorophyll *c* group are found in oxygenic marine microalgae and brown algae. *See* ALGAE.

Chlorophyll d. Chlorophyll *d* has been shown to be the major chlorophyll in the cyanobacterium *Acaryochloris marina* which is a free-living epiphytic bacterium growing in association with red algae. *See* CYANOBACTERIA.

Bacteriochlorophylls. Photoheterotrophic bacteria, which do not evolve O_2, contain bacteriochlorophyll *a* as their reaction-center pigment but may contain bacteriochlorophylls *a, b, c, d, e*, and *g* as antenna pigments. Purple photosynthetic bacteria contain bacteriochlorophylls *a, b*, and *g*, and the green bacteria contain bacteriochlorophylls *a, c, d*, and *e*. *See* BACTERIA; BACTERIAL PHYSIOLOGY AND METABOLISM.

Function

Chlorophyll molecules have three functions: (1) They serve as antennae to absorb light quanta. (2) They transfer this energy from one chlorophyll to another over distances usually of 1.5–2 nm by a process of resonance transfer, so that the energy finally comes to reside in the special chlorophyll molecules P700 or P680 (the number after P represents their long-wavelength absorption maxima) in the reaction centers of photosystem I or II. (3) The special reaction-center chlorophyll molecules P680 and P700, having in their excited states (symbolized as P680* and P700*) the energy of the photon, undergo chemical oxidation by transferring one of their electrons to a neighboring acceptor molecule. Specifically, P680* transfers one electron to a pheophytin molecule (in case of photosystem II) and P700* to another chlorophyll molecule (in the case of photosystem I). In this way the energy of the fleeting light quanta is converted into stable chemical energy.

Light-induced oxidation of reaction-center chlorophyll molecules leads to photobleaching, which is detected spectroscopically by an absorbance decrease. Loss of an electron oxidizes the reaction-center chlorophyll molecules. P680 is transformed to $P680^+$ and P700 to $P700^+$. Analogous bacteriochlorophyll species, P840+ or P890+, is produced in nonoxygenic photosynthetic bacteria. These cationic species are detected both optically as well as by the electron spin resonance technique, because each has an unpaired electron. *See* SPECTROSCOPY.

Chemical Structure

Chlorophylls contain up to four pyrrole rings that are joined by four methine (—CH═) bridges at C5, C10, C15, and C20. However, all chlorophylls also contain a fifth cyclopentanone ring, designated E. Chlorophylls with four pyrrole rings are called phytoporphyrins, and the chlorophylls of the *c* group belong to this family, the members of which are emerald green in color. Chlorophyll *a* contains three pyrrole rings (rings A, B, and C) and one dihydropyrrole ring (ring D) with a hydrogen atom at C17 and C18 (**Fig. 1**). Such chlorophylls are called chlorins, and chlorophylls *a*, *b*, and *d*, and bacteriochlorophylls *c*, *d*, and *e* belong to this family: chlorophyll *a* is blue and the other five are green. Some chlorophylls contain two dihydropyrrole rings (rings B and D) with hydrogen atoms at C7, C8, C17 and C18. These chlorophylls are called bacteriochlorins. Bacteriochlorophylls *a*, *b*, and *g* belong to the bacteriochlorin family and they are purple. These different-colored chlorophylls allow the organisms that possess them to efficiently use (as photosynthetically active radiation) those parts of the visible and near-infrared spectrum of solar radiation that permeate their environment.

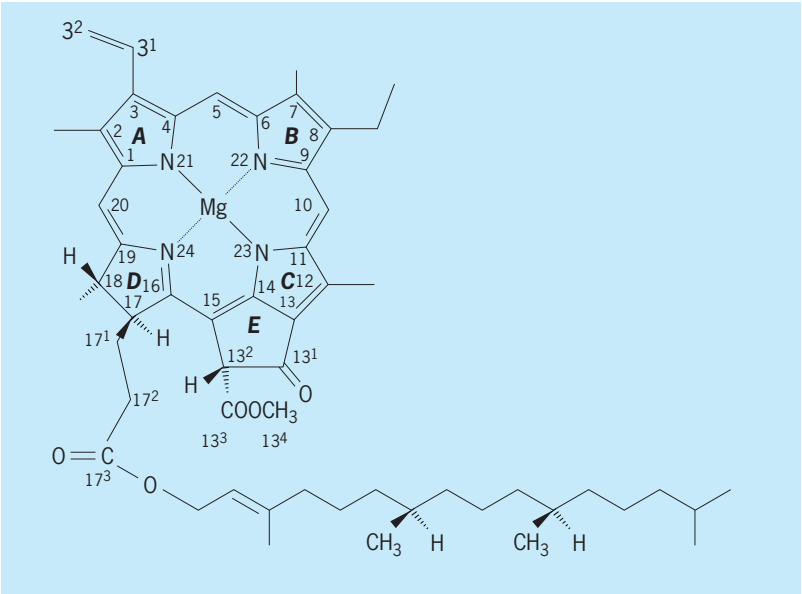

Fig. 1. Structure of chlorophyll *a* ($C_{55}H_{72}O_5N_4Mg$).

Central metal atom. Chlorophyll molcecules contain a central magnesium (Mg) atom that binds extra ligands (besides the four pyrrolic nitrogens) so strongly that it seldom exists only as a four-coordinate metal: one or both of the axial positions are occupied by an electron donor (nucleophile) group so that the Mg is generally a six-coordinate metal, but also sometimes exists as a five-coordinate metal. Axial coordination to the central Mg atom is important for understanding the ligation of chlorophyll to histidine, glutamine, or asparagine residues of chlorophyll apoproteins. Due to the greater than 4 coordination number of Mg, chlorophyll can exist as a monomer, dimer, trimer, tetramer, and as an aggregate of *n* number of chlorophyll molecules. The monomers absorb red light maximally in the 660–670-nm region, whereas the aggregates absorb (darker) red light in the 680–700-nm region. On the other hand, the chlorophyll-H_2O-chlorophyll adduct is suggested to absorb as far as 740 nm; the (chlorophyll·H_2O)$_2$ adduct may absorb around 700 nm.

Chlorins. The structure of chlorophyll *a*, which is a 3-monovinyl compound, is shown in Fig. 1. This molecule has now been synthesized in the laboratory. The characteristic feature of chlorophyll *a* is that it is a Mg-chlorin. Because rings A, B, C, and D are linked through methine bridges at C5, C10, C15, and C20 (Fig. 1), a large resonating 16-membered inner ring of carbon and nitrogen atoms attached to each other through alternating single and double bonds is formed. Closely related derivatives of chlorophyll *a* (the monovinyl pigment shown in Fig. 1) include [3,8-divinyl]-chlorophyll *a* which sometimes has long-chain alcohols other than phytol esterified at C17^3. Some widely dispersed marine phytoplanktonic *Prochlorococcus* strains contain 3,8-divinyl derivatives of chlorophylls *a* and *b*: it has been calculated that these divinyl pigments are responsible for approximately 25% of global photosynthesis.

Other chlorophylls closely related to chlorophyll *a* possess formyl (—CH═O) groups. Chlorophyll *b* is [7-formyl]-chlorophyll *a* and chlorophyll *d* is [3-formyl]-chlorophyll *a*.

Bacteriochlorophylls *c* and *d*, like chlorophyll *a*, belong to the chlorin family. The common features include a hydroxyethyl (—$CHOHCH_3$) group at C3, while the carboxymethyl (—$COOCH_3$) group at C13^2 of chlorophyll *a* is replaced by a H atom and the long-chain alcohol esterified at C17^3 is frequently farnesol ($C_{15}H_{25}OH$) of chlorophyll *a*, which replaces the phytol ($C_{20}H_{39}OH$). Bacteriochlorophylls *c* and *e* in addition have a C20 methyl (—CH_3) substituent.

Phytoporphyrins. The chlorophyll *c* group is a diverse group of Mg-phytoporphyrins including [3,8-divinyl]-protochlorophyllide *a*, which is also shown as a biosynthetic intermediate in the biosynthesis of chlorophyll *a* (**Fig. 2**). Most chlorophylls *c* have a free C17^3 carboxyl group, but two recently discovered nonpolar chlorophylls *c* are esterified at C17^3 to the sugar residue of a monogalactosyldiacylglycerol (MGDG) moiety. Frequently, the propionate group at C17 is replaced by an acrylic acid (—CH═CH—COOH) side chain.

Bacteriochlorins. Bacteriochlorophylls *a, b*, and *g* belong to the bacteriochlorin family of chlorophylls. There are several forms of bacteriochlorophyll *a*: one is esterified at C17^3 with phytol and the other is esterified with geranyl-geranol. Both forms have an acetyl (—CO—CH_3) group at C3 and an ethyl (—C_2H_5) group at C8. Bacteriochlorophylls *b* and *g* have an ethylidene (═CH—CH_3) group at C8 but bacteriochlorophyll *b* has an acetyl group at C3 where bacteriochlorophyll *g* has a vinyl (—CH═CH_2) group. In *Acidiphilium rubrum* the central Mg of bacteriochlorophyll *a* is replaced by a zinc (Zn) atom.

Chlorophyll derivatives. Pheophytins are chlorophyll derivatives lacking the central Mg atom; they occur naturally as electron acceptors in photosystem II. Bacteriopheophytin *a* is a derivative of bacteriochlorophyll *a* and serves as an electron acceptor in bacterial reaction centers. Pheophorbides are chlorophyll derivatives that have lost the central Mg atom and the long-chain alcohol ester moiety.

Biosynthesis

The two major plant tetrapyrrole pigments, chlorophylls and hemes, are both synthesized from ALA (5-aminolevulinic acid). ALA is converted to protoporphyrin in a series of enzymatic steps, which are identical in plants and animals. Thereafter, the pathway branches to form hemes text are porphyrins chelated with iron (Fe^{2+}), which are present in many respiratory hemoproteins including cytochromes. In the other branch, porphyrins are chelated with Mg^{2+} to form precursors of chlorophyll (Fig. 2). *See* HEMOGLOBIN; PORPHYRIN.

ALA is formed by the condensation of succinyl coenzyme A and glycine by ALA-synthase in non-plastid-containing eukaryotes (animals, yeasts and fungi) and in the α subgroup of proteobacteria (which includes the well-known genera of photosynthetic bacteria *Rhodopseudomonas,*

Rhodobacter, and *Rhodospirillum*). However, in plants, algae, Archaea, and most bacteria (excluding the α-proteobacteria), ALA is formed from glutamate, a 5-carbon amino acid, by the C_5 pathway. This is the more ubiquitous pathway of ALA formation in nature. All of the enzymes necessary for converting glutamate to ALA (and further to chlorophyll) are contained in the chloroplasts of higher plants and green algae. The $C13^1$, $C13^2$, and $C13^3$ of ring E (Fig. 1) arise from the propionic acid (—CH_2—CH_2—COOH) group of Mg-protoporphyrin. An oxygen is introduced to form a 13^1-oxo-propionic acid (—CO—CH_2—COOH) moiety and spontaneous bonding then occurs between $C13^2$ and C15, thus forming ring E with a $C13^3$ carboxymethyl ($CHOOCH_3$) substituent. In higher plants and green algae, the oxygen for the 13^1-oxo-group comes from molecular oxygen (O_2), by activity of an oxygenase-type enzyme; in anaerobic photosynthetic bacteria, it comes from water (H_2O), by a hydratase enzyme activity.

In higher plants, a light-dependent and reduced nicotinamide adenine dinucleotide phosphate (NADPH)-dependent protochlorophyllide reductase (LPOR) introduces a hydrogen atom at both C17 and C18, thus converting protochlorophyllide to chlorophyllide. Cyanobacteria, green algae, mosses, ferns, and gymnosperms contain both LPOR and a dark (that is, light-independent) DPOR enzyme so that these organisms can synthesize chlorophyllide in light or dark. The esterification of a geranylgeranyl pyrophosphate forms chlorophyll *a* with a geranylgeranyl ester, which is followed by reduction to a phytyl ester: the esterification and reduction are both catalyzed by chlorophyll synthase, thus completing chlorophyll *a* formation. Among additional cofactors required for chlorophyll formation are Fe^{2+}, Mg^{2+}, adenosine triphosphate (ATP), nicotinamide adenine dinucleotide (NAD), and O_2. Concomitant with chlorophyll synthesis, the disk, or thylakoid, membranes are also synthesized. The possible origins of the phycobilin pigments, chlorophylls *b*, *c*, and *d*, and the bacteriochlorophylls are shown in Fig. 2.

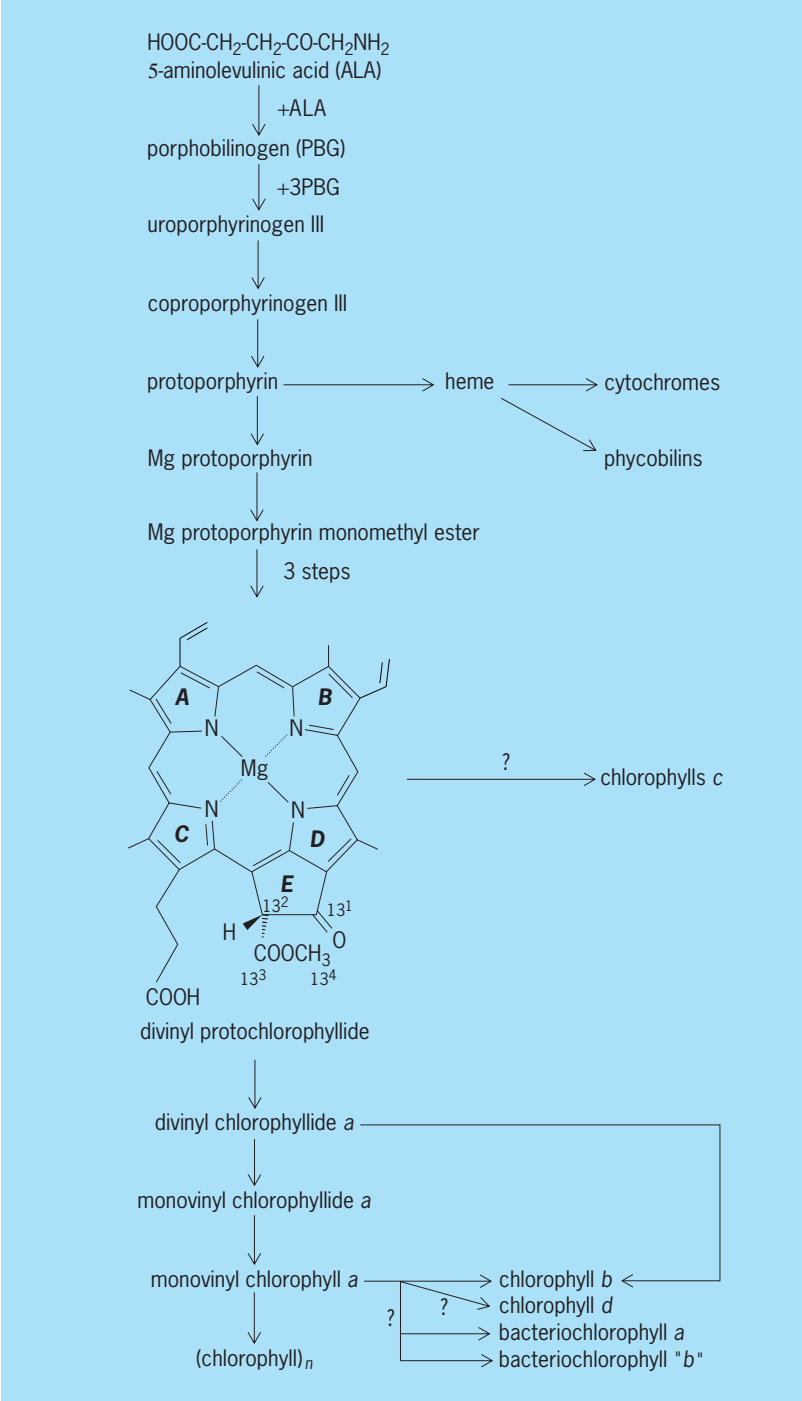

Fig. 2. Simple biosynthetic route for chlorophyll *a* and some related compounds. (In reality most of the pathways are much more complex and are often branched.) The tetrapyrrole structure shown above is Mg 3,8-divinyl protochlorophyllide.

Fluorescence

Chlorophylls reemit a fraction of the light energy they absorb as fluorescence. Regardless of the wavelength of the absorbed light, the emitted fluorescence is always on the long-wavelength side of the lowest energy absorption band, in the red or infrared region of the spectrum. The fluorescent properties of a particular chlorophyll are functions of the structure of the molecule and its immediate environment. Thus, the fluorescence spectrum of chlorophyll in the living plant is always shifted to longer wavelengths (peak at 685 nm) relative to the fluorescence spectrum of a solution of the same pigment (peak at approximately 660 nm). This red shift is characteristic of chlorophyll-protein complexes.

Even in dilute solutions, the capacity of chlorophyll to fluoresce depends on the nature of the solvent. In solvents that can combine with the central Mg atom of chlorophyll by donating a pair of electrons to it, chlorophyll is fluorescent. In solvents that lack this property, chlorophyll is dimeric or polymeric and nonfluorescent at room temperature. The aggregates are formed by combining the carbonyl group of one molecule with the Mg atom of the neighboring molecule.

An excited chlorophyll *a* in diethyl ether has a mean lifetime of 5×10^{-9} s (5 nanoseconds), whereas in the living plant this is reduced to $0.5–2 \times 10^{-9}$ s (0.5–2 ns). A long-lived excited state of chlorophyll *a* (the triplet state) has been observed under

special conditions, such as illumination of concentrated solutions in dry hydrocarbon solvents at low temperatures. Under these conditions, chlorophyll *a* emits phosphorescence at a spectral maximum of 750 nm. Weak phosphorescence of chlorophyll *a* has been recently observed in living cells.

The most widespread chlorophylls in nature, chlorophylls *a* and *b*, fluoresce with a quantum efficiency of 0.33 and 0.16, respectively, in dilute solution in diethyl ether. In the living cell the quantum efficiency drops to 0.03 for chlorophyll *a* and to zero for chlorophyll *b*. This is due to the property of chlorophyll *b* to transfer all its excitation energy to chlorophyll *a*, which, in turn, channels most of its excitation to photosynthesis, allowing only a small fraction to escape as fluorescence.

Measurements of chlorophyll *a* fluorescence are of great practical use as these have been shown to be a probe of the efficiency of photosynthesis. Applications of fluorescence are numerous and include measurements of the photosynthetic efficiency of plants under field conditions; advance estimations of the impact of abiotic and biotic environmental stresses on anticipated crop and biomass and yields; and, the assessment of the photosynthetic efficiency and productivity of plant communities in terrestrial, inland water, and marine ecosystems by remote sensing of chlorophyll *a* fluorescence. *See* FLUORESCENCE.

Govindjee; Robert J. Porra; George C. Papageorgiou

Bibliography. R. Blankenship, *Molecular Mechanisms of Photosynthesis*, Blackwell, Oxford, 2002; S. Granick and S. I. Beale, Hemes, chlorophylls, and related compounds: Biosynthesis and metabolic regulation, *Adv. Enzymol.*, 46:33–203, 1978; B. Grimm et al. (eds.), *Chlorophylls and Bacteriochlorophylls: Chemistry, Biochemistry, Functions and Applications*, Springer, Dordrecht, 2006 [vol. 25 of *Advances in Photosynthesis and Respiration* (series editor: Govindjee)]; S. W. Jeffrey, R. F. C. Mantoura, and S. W. Wright (eds.), *Phytoplankton Pigments in Oceanography: Guidelines to Modern Methods*, UNESCO, Paris, 1997; B. Ke, *Photosynthesis: Photobiochemistry and Photobiophysics*, Springer, Dordrecht, 2001 [vol. 10 of *Advances in Photosynthesis and Respiration* (series editor: Govindjee)]; G. Papageorgiou and Govindjee (eds.), *Chlorophyll a Fluorescence: A Signature of Photosynthesis*, Springer, Dordrecht [vol. 19 of *Advances in Photosynthesis and Respiration* (series editor: Govindjee)]; H. Scheer (ed.), *Chlorophylls*, CRC, Boca Raton, 1991; L. P. Vernon and G. R. Seely (eds.), *The Chlorophylls*, Academic Press, New York, 1966.

Choanoflagellida

An order of the class Zoomastigophorea, superclass Mastigophora, subphylum Sarcomastigophora, in the phylum Protozoa. They are principally small, single-celled or colonial, structurally simple, colorless flagellates distinguished by having the anterior end surrounded by a thin protoplasmic collar, within which is an ingestive area for particulate matter, for example, bacteria. A single flagellum arises from the center of this collar. The genus *Monosiga* is single, sometimes with a stalk, sometimes free in the plankton. *Desmarella* is a linear colony (see **illus.**), and *Astrosiga* and *Sphaerosiga* are colonies radiating from a center. *Proterospongia* cells arise in a common jelly, the zooids being organically connected. *Diploeca* lives in an attached lorica. *Stelexomonas* and *Aulomonas* inhabit tubular loricae. *Diplosiga* has a double colony. *Sportelloeca* and *Pleurasiga* live in complex loricae. *See* PROTOZOA; SARCOMASTIGOPHORA; ZOOMASTIGOPHOREA.

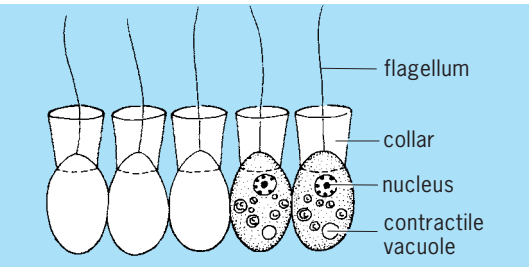

A linear colony of the choanoflagellate, *Desmarella moniliformis*.

The botanist P. Bourrelly puts these organisms in the order Monosigales. It is generally agreed, however, that the group shows relationships to the sponges. They are widespread in fresh and salt water and are common in inshore waters. James B. Lackey

Bibliography. B. M. Honigberg et al., A revised classification of the phylum Protozoa, *J. Protozool.*, 11(1):7–20, 1964; R. E. Norris, Neustonic marine Craspedomonadales (choanoflagellates) from Washington and California, *J. Protozool.*, 12(4):589–602, 1965; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Choke (electricity)

An inductor that is used to prevent electric signals and energy from being transmitted along undesired paths or into inappropriate parts of an electric circuit or system. Power-supply chokes prevent alternating-current components, inherent to a power supply, from entering the electronic equipment. Radio-frequency chokes (RFCs) prevent radio-frequency signals from entering audio-frequency circuits. The printed circuit boards used in virtually all electronic equipment such as computers, television sets, and high-fidelity audio systems typically have one or more chokes. The purposes of these chokes are the (1) attenuation of spurious signals generated in the equipment itself so that these signals will not be transmitted to other parts of the circuit or beyond the overall system to other electronic devices; (2) prevention of undesired signals or electrical noise generated in other parts of the system from adversely affecting circuit performance; and (3) prevention of ripple from the power supply from degrading system behavior. Waveguide chokes keep microwave energy from being transmitted to the wrong part of a waveguide system. *See* ELECTRICAL

Graphic symbols for chokes or inductors in electric circuits. (*a*) Choke or inductor with air core. (*b*) Choke or inductor with magnetic core. (*c*) Waveguide choke.

NOISE; ELECTRONIC POWER SUPPLY; PRINTED CIRCUIT; RIPPLE VOLTAGE; WAVEGUIDE.

In its simplest form, a choke or an inductor is a coil of wire (usually copper) wound around and insulated from a core, which may or may not be ferromagnetic. Ferromagnetic cores tend to increase inductance, reduce physical size, and reduce electromagnetic coupling between circuit elements, and they may increase power loss and resultant heating. Such cores often lead to nonlinear or swinging chokes, commonly found in power supplies, where the nonlinearity may be an advantage. The basic principle of operation derives from the fact that it is impossible to change instantaneously the energy stored in a magnetic field. The various configurations are designed to control the magnetic field in an appropriate manner. The graphic symbols used to represent the various types of chokes or inductors in electric circuits are shown in the **illustration**.

Chokes come in a wide variety of inductance ranges, power-handling capabilities, fabrication technologies, and intended applications. In many cases they are specially designed for a particular situation. Chokes are available in inductance sizes ranging from a fraction of a microhenry to tens of henrys or higher. The frequency range of intended use is a major consideration. *See* INDUCTANCE; INDUCTOR.

Edwin C. Jones, Jr.

Bibliography. D. G. Fink and D. Christiansen (eds.), *Electronics Engineers Handbook*, 3d ed., 1989; Institute of Electrical and Electronics Engineers, *Graphic Symbols for Electrical and Electronics Diagrams*, IEEE 315–1975, 1975; W. T. McLyman, *Transformer and Inductor Design Handbook*, 2d ed., 1988.

Choked flow

Fluid flow through a restricted area whose volume flow rate reaches a maximum when the fluid velocity at the minimum area reaches sonic speed. The phenomenon of choking exists only in compressible flow and can occur in several flow situations. *See* COMPRESSIBLE FLOW.

Flow through varying-area duct. Choked flow can occur through a convergent flow area or nozzle attached to a huge reservoir (**illus.** *a*). Flow exits the reservoir through the nozzle if the back pressure is less than the reservoir pressure. In all cases considered here, the reservoir pressure is assumed to be a constant. When the back pressure is decreased slightly below the reservoir pressure, a signal from beyond the nozzle exit is transmitted at sonic speed to the reservoir. The reservoir responds by sending fluid through the nozzle. Further, the maximum velocity of the fluid exists at the nozzle throat where the area is smallest.

When the back pressure is further decreased, fluid exits the reservoir more rapidly. Eventually, however, the velocity at the throat reaches the sonic velocity. Then the fluid velocity at the throat is sonic, and the velocity of the signal is also sonic. Therefore, further decreases in back pressure are not sensed by the reservoir, and correspondingly will not induce any greater flow to exit the reservoir. The nozzle is thus said to be choked, and the mass flow of fluid is a maximum for a given reservoir pressure. Thus, choking in a converging nozzle occurs when the Mach number (defined as the ratio of fluid velocity to the sonic velocity) at the throat equals 1. *See* MACH NUMBER; NOZZLE; SOUND; SUPERSONIC DIFFUSER.

Flow with friction. Choked flow can also occur through a long constant-area duct attached to a reservoir (illus. *b*). As before, flow from the reservoir through this duct is induced when the back pressure is decreased below the reservoir pressure, and a signal is transmitted through the duct to the reservoir.

As fluid flows through the duct, friction between the fluid and the duct wall reduces the pressure acting on the fluid. As pressure is reduced, other fluid properties are affected, such as sonic velocity, density, and temperature. The maximum Mach number occurs at the duct exit. As the back pressure is decreased, flow exits more rapidly through the duct,

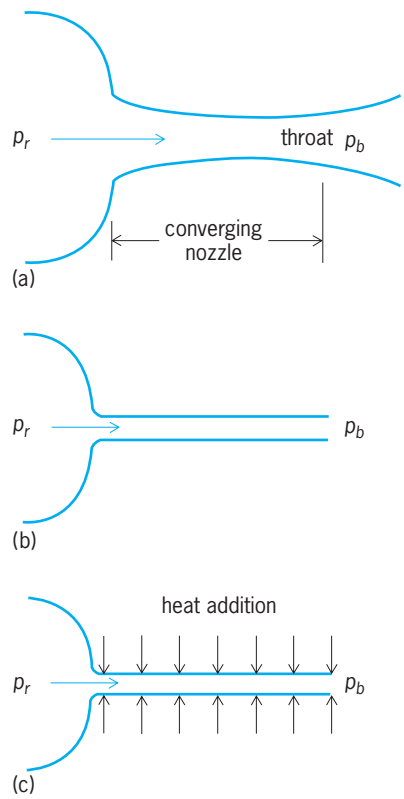

Examples of choked flow. (*a*) Flow through a varying-area duct or nozzle. (*b*) Flow through a constant-area duct. (*c*) Frictionless flow through a constant-area duct with heat addition. P_b = back pressure; P_r = reservoir pressure.

but eventually the Mach number at the exit reaches 1. Any further decrease in back pressure is not sensed by the reservoir and so causes no increase in flow rate. Thus, when the Mach number is 1 at the duct exit, the flow is said to be choked, and the mass flow rate through the duct is a maximum.

Frictionless flow with heat addition. A reservoir with a constant-area duct attached may also be considered in the case that the flow through the duct is assumed to be frictionless, but heat is added to the system along the duct wall (illus. *c*). As fluid receives energy in the form of heat, fluid properties again change. The maximum Mach number occurs at the duct exit. As the total amount of heat increases, the Mach number at the duct exit increases, and eventually reaches 1. At this point, the mass flow through the duct is a maximum and the flow is said to be choked. Any further heat addition results in a decrease in mass flow through the duct. *See* FLUID FLOW; GAS DYNAMICS.

William S. Janna

Bibliography. J. D. Anderson, Jr., *Modern Compressible Flow: With Historical Perspective*, 3d ed., 2003; W. S. Janna, *Introduction to Fluid Mechanics*, 3d ed., 1993; J. E. A. John and T. G. Keith, *Gas Dynamics*, 3d ed., 2001.

Cholera

A severe diarrheal disease caused by infection of the small bowel of humans with *Vibrio cholerae*. This facultatively anaerobic, curved, gram-negative, rod-shaped bacterium is rapidly motile by means of a single sheathed polar flagellum. Cholera is transmitted by the fecal-oral route. From its ancestral home in the Indo-Pakistani subcontinent, through movement of people by ships and other means of transportation, cholera has swept the world in seven pandemic waves. These involved the Western Hemisphere several times in the 1800s, and again in Peru in 1991. Whereas previous cholera outbreaks were associated with high mortality rates, through understanding of its pathophysiology it can now be said that no one should die of cholera who receives appropriate treatment soon enough.

Cholera produces a secretory diarrhea caused by the protein cholera enterotoxin (CTX), which is composed of two kinds of subunits: five B subunits responsible for binding to a glycolipid, the GM1 ganglioside, on host membranes; and a single A subunit, which is an enzyme. The latter, an adenosine ribosyl transferase, transfers adenosine diphosphate P–ribose from nicotinamide adenine dinucleotide to a host cell regulatory protein, G_s. This causes permanent activation of adenylate cyclase and a consequent increase in cyclic adenosine monophosphate. The result is hypersecretion of chloride and bicarbonate and inhibition of sodium absorption, leading to the secretion of the large volumes of isotonic fluid which constitute the diarrhea of severe cholera. The B subunit also activates arachidonic acid metabolism, leading to the production of eicosanoids, which may contribute to the diarrhea as well as to the immune response. The treatment consists of replacing the fluids and electrolytes lost in the voluminous cholera stool. This can be done intravenously, which may be essential initially in comatose patients, or perorally using oral rehydration salt solutions containing a utilizable carbon (energy) source (glucose, rice powder). The oral rehydration solutions are a relatively recent development, can also be used in the treatment of other diarrheal diseases, and have been responsible for saving millions of lives. Appropriate antibiotics can also be used adjunctively to decrease the excretion of the vibrios and the duration of the disease. The incubation period may be less than one day or up to several days; properly treated, the patient should recover in 4 or 5 days. The disease produces immunization, and convalescents rarely get cholera again. *See* DIARRHEA.

Despite the fact that the cholera vibrios were first discovered by Robert Koch in 1883 and a cholera vaccine was introduced 3 years later, there is still no effective, economical, and nonreactogenic vaccine. Use of a killed whole-cell vaccine administered parenterally was eliminated because of expense, reactogenicity, and lack of efficacy. Experimental vaccines currently being evaluated include genetically engineered living attentuated preparations administered perorally (or intranasally), killed whole-cell vaccines administered perorally, and conjugated vaccines (polysaccharide and toxin antigens) administered parenterally. Efforts are also being made to include cholera antigens transgenically in edible plants. A recently developed candidate attenuated vaccine has a deletion in the CTXφ (which carries the genes for CTX) and is also defective in its hemagglutinin/protease. The latter is a zinc metalloprotease which apparently causes detachment of the vibrios (to find another host) and may contribute significantly to the diarrhea which is still produced by CTX^- strains.

A complicating feature is the fact that of approximately 150 recognized serogroups of *V. cholerae*, until 1992 only the two biotypes, classical (first described by Koch) and El Tor (recognized later), of serogroup O1 have been responsible for all epidemic cholera (El Tor first emerged in pandemic form in 1961). In 1992 a recently recognized serogroup, O139, caused epidemic cholera in India and Bangladesh and, for a time, replaced the resident El Tor vibrios. O139 resembles El Tor except for the distinctive O139 (instead of O1) polysaccharide somatic antigen and the possession of a capsule. They are antigenically distinct, so a new vaccine will be required for O139. The emergence of O139 raises the specter that other serogroups of *V. cholerae* may acquire virulence and epidemicity by some means of horizontal genetic transmission. It has recently been recognized that CTX is genetically encoded by a bacteriophage, CTXφ, which has become integrated in the *V. cholerae* chromosome. Virulent cholera vibrios also produce toxin-coregulated pili, surface fibrils, which are essential for colonization and also serve as receptors for CTXφ. Other non-O1/O139 vibrios carry CTXφ but apparently do not have other

essential attributes of virulence, including the ability to adhere and colonize the small bowel and to spread epidemically. Another pilus, the mannose-sensitive hemagglutinin, serves as the receptor for another temperate filamentous choleraphage, 493, which may play a role in horizontal genetic transmission and cholera vibrio territoriality.

The best ways to avoid cholera are by chlorination of water, sanitary disposal of sewage, and avoidance of raw or improperly cooked seafood, which may have become infected by ingesting infected plankton in epidemic areas. Richard A. Finkelstein

Bibliography. D. Barua and W. B. Greenough III, *Cholera*, Plenum, New York, 1992; J. A. Benitez et al., Preliminary assessment of the safety and immunogenicity of a new CTXφ-negative, hemagglutinin/protease defective El Tor strain as a cholera vaccine candidate, *Infect. Immun.*, 67:539–545, 1999; E. A. Jouravleva et al., The *Vibrio cholerae* mannose-sensitive hemagglutinin is the receptor for a filamentous bacteriophage from *V. cholerae* O139, *Infect. Immun.*, 66:2535–2539, 1998; J. B. Kaper, J. G. Morris, Jr., and M. M. Levine, Cholera, *Clin. Microbiol. Rev.*, 8:48–86, 1995; J. W. Peterson et al., Cholera toxin B subunit activates arachidonic acid metabolism, *Infect. Immun.*, 67:794–799, 1999; K. H. Thelin and R. K. Taylor, Toxin-coregulated pilus, but not mannose-sensitive hemagglutinin, is required for colonization by *Vibrio cholerae* O1 El Tor biotype and O139 strains, *Infect. Immun.*, 64:2853–2856, 1996; I. K. Wachsmuth, P. A. Blake, and O. Olsvik, *Vibrio cholerae and Cholera: Molecular to Global Perspectives*, American Society for Microbiology, Washington, DC, 1994; M. K. Waldor and J. J. Mekalanos, Lysogenic conversion by a filamentous phage encoding cholera toxin, *Science*, 272:1910–1914, 1996.

Cholesterol

A cyclic hydrocarbon alcohol commonly classified as a lipid because it is insoluble in water but soluble in a number of organic solvents. It is the major sterol in vertebrate cells and the most common sterol of eukaryotes, but it is absent from most prokaryotes. In vertebrates, the highest concentration of cholesterol is in the myelin sheath that surrounds nerves and in the plasma membrane that surrounds all cells. Insects and most invertebrates cannot synthesize cholesterol; therefore, they rely on dietary cholesterol for use in their membranes and as a precursor of ecdysone, a molting hormone required during growth and development. *See* ECDYSONE; LIPID.

Function. All vertebrate cells require cholesterol in their membranes to function normally. Cholesterol helps maintain normal cell membrane function by associating with the phospholipid acyl chains within the membrane, thereby adjusting up or down the degree of molecular motion or fluidity within the membrane that allows proteins to carry out their tasks. Cholesterol is also the required precursor molecule

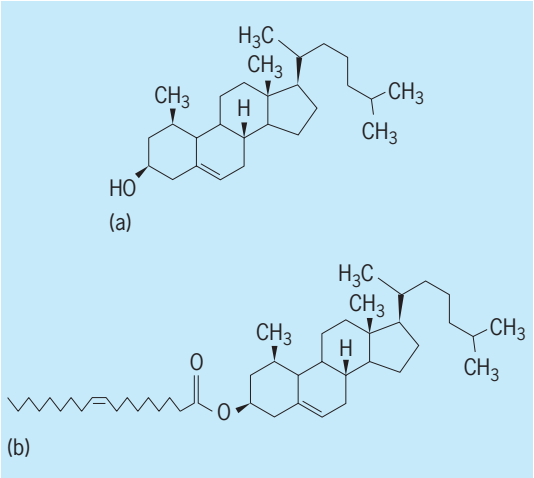

Fig. 1. Cholesterol. (*a*) Structure of unesterified, or free, cholesterol with the hydroxyl group in the 3β, position. (*b*) Structure of a cholesteryl ester molecule with a fatty acid (oleate in this case) attached to the 3β-hydroxy position via an ester bond.

for the formation of steroid hormones, bile acids, and vitamin D, which are required in the execution of various essential bodily processes.

Free and esterified forms. Cholesterol can exist either in the free (unesterified) form, in which the hydroxyl group (OH) at carbon-3 exists in the beta (β) configuration (**Fig. 1***a*), or as a cholesteryl ester to which a fatty acid is attached through an ester bond to the 3β-hydroxyl group (Fig. 1*b*). The free form is found in cell membranes and on the surface of plasma lipoproteins (lipid-protein emulsion-like complexes that transport cholesterol among tissues via the blood), with the hydroxyl group at the membrane aqueous environment interphase and the hydrocarbon tail buried among the fatty acyl chains of the phospholipids forming the membrane or surface layer. The cholesteryl ester molecule is usually found within lipid droplets in the cytoplasm of cells or within the neutral lipid core of plasma lipoproteins.

Cholesterol can be esterified by tissue enzymes termed acyl-coenzyme A (CoA):cholesterol *O*-acyltransferases 1 and 2 (ACAT1 and ACAT2) and by the plasma enzyme lecithin-cholesterol acyltransferase or LCAT. ACAT1 catalyzes the esterification of cholesterol for storage within cells in steroidogenic (steroid-producing) tissues. In the liver, the organ that serves as the central clearinghouse for cholesterol in the body, cholesterol esterification via ACAT2, leads to its storage or its secretion in the form of lipoproteins. Esterification via ACAT1 also occurs in macrophages (phagocytic immune cells) in the arterial wall, forming the lipid-filled foam cells of atherosclerosis in a process apparently designed to protect the cell membrane from becoming overly enriched in free cholesterol. In steroidogenic tissues, cholesteryl ester can be hydrolyzed via catalysis by cholesterol ester hydrolase (CEH), and the resulting cholesterol is then converted into steroid hormones. In macrophage foam cells, hydrolysis via CEH

allows for cholesterol mobilization and removal from the cell; and in the liver, hydrolysis via CEH permits the cholesterol molecule to be used for any of the multiple pathways within the liver, including those resulting in the formation of bile and plasma lipoproteins and those involved in membrane function. *See* CELL MEMBRANE; LIPOPROTEIN; PHOSPHOLIPID.

Sources. Cellular cholesterol is obtained both from the diet, following its absorption in the intestine, and from synthesis within all cells of the body.

Intestinal cholesterol absorption. Much of the cholesterol in the body can be obtained in the diet. Foods that contain cholesterol come from animal sources, and those highest in cholesterol include eggs, dairy products, and organ meats such as liver and brain. The cholesterol from the human diet, typically about 250 mg per day, becomes mixed with endogenous (originating from within the body) cholesterol (about 750 mg per day), secreted into the intestine in bile. Through the processes of digestion, dietary cholesterol is solubilized by bile acids (detergents that facilitate fat and cholesterol absorption) into micelles (water-soluble particles) and transferred through a protein called Niemann-Pick C1-like 1 into enterocytes (cells lining the intestine), whereupon it is esterified by ACAT2, incorporated into absorptive lipoproteins called chylomicrons, and brought into the circulation and deposited in the liver. This process is only about 50% efficient, and about half of the cholesterol in the intestine is excreted each day.

While all cell types have the enzymatic machinery to synthesize cholesterol, only the liver has enzymes to degrade the cholesterol molecule in significant amounts by converting it into bile acids. Assembly of bile acids and cholesterol into bile micelles in the liver with subsequent secretion into the intestine, followed by absorption and return of the absorbed cholesterol to the liver constitutes the process termed enterohepatic recirculation. In a normal day, the bulk of the removal of cholesterol from the body occurs through fecal excretion of approximately equal amounts (~1 g combined) of cholesterol and bile acid. *See* ABSORPTION (BIOLOGY); LIVER.

Biosynthesis and feedback regulation. Cholesterol synthesis occurs in all vertebrate cells, with perhaps the highest rates occurring in the liver, skin, intestine, and in the brain at the time of myelination. The process involves the conversion of two-carbon acetate molecules into isoprenoid units continuing on to produce a 30-carbon precursor known as lanosterol. This molecule is then rapidly converted into the 27-carbon cholesterol molecule. The whole process requires over 40 enzymes and is highly regulated by the cholesterol status of the cell through the process of feedback regulation. The key rate-limiting enzyme in the biosynthetic process is 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMG-CoA reductase). This is the target enzyme for the statin drugs that are used clinically to reduce blood cholesterol concentrations. These drugs work in part by binding to and inactivating HMG-CoA reductase, resulting in a decrease in the rate of cholesterol synthesis. *See* COENZYME; ENZYME.

Although the presence of feedback regulation of cholesterol synthesis was identified some 70 years ago, only recently have we come to understand the complexity of the process. The systematic identification of the many individual steps in this process has provided new fundamental information about the structure and function of cell membranes. A whole series of protein molecules has been evaluated, and the identification of the processes that govern the behavior of these proteins has provided unique insights into how membranes function as molecular sensors responding to changes in the cholesterol availability within the cell.

In the 1950s, the liver was shown to be the site of much of the feedback regulation; nevertheless, many of the details about the feedback mechanism have been identified in skin fibroblasts and in other cells in culture that can be manipulated to induce mutations in individual steps in the regulatory process (**Fig. 2**). Initially, cholesteryl ester-rich low-density lipoproteins (LDL) bind to receptors on the cell surface termed LDL receptors. These receptors localize in regions within the membrane, termed clathrin-coated pits in cell types other than liver cells. These structures bud off the membrane and move into the cell. The bound LDL is delivered to lysosomes (cellular organelles containing digestive enzymes), whereupon the cholesteryl esters are hydrolyzed, making the cholesterol available to the cell. In liver hepatocytes, where most of the LDL clearance from plasma occurs, the LDL receptors do not cluster in coated pits. Rather, the receptor appears to interact with an as yet unidentified protein molecule for internalization of bound LDL particles.

Within cells, the amount of the LDL receptor produced is regulated through transcriptional mechanisms. The cholesterol molecules brought into the cell as a result of receptor activity will eventually transmit the signal for downregulation (a decrease) of receptor-mediated uptake of LDL. This occurs by shutting down the transcriptional activation of the LDL receptor gene in an analogous fashion to the regulation of the genes that participate in cholesterol biosynthesis.

The transcription factor responsible for stimulating transcription of the enzymes of the biosynthetic pathway as well as the LDL receptor is termed the sterol regulatory element-binding protein (SREBP). This protein controls gene transcription by binding to a nucleotide sequence (typically 5′-TCACNCCAC-3′, where N can be any base) within the promoter region of genes. There appear to be a family of three SREBPs (SREBP1a, SREBP1c, and SREBP2), each of which may be involved in regulation of different aspects of lipid metabolism. These proteins are initially integral membrane proteins within the endoplasmic reticulum. As such, they have no access to the nucleus without protein processing. To gain access to the nucleus, the SREBPs must move within vesicles

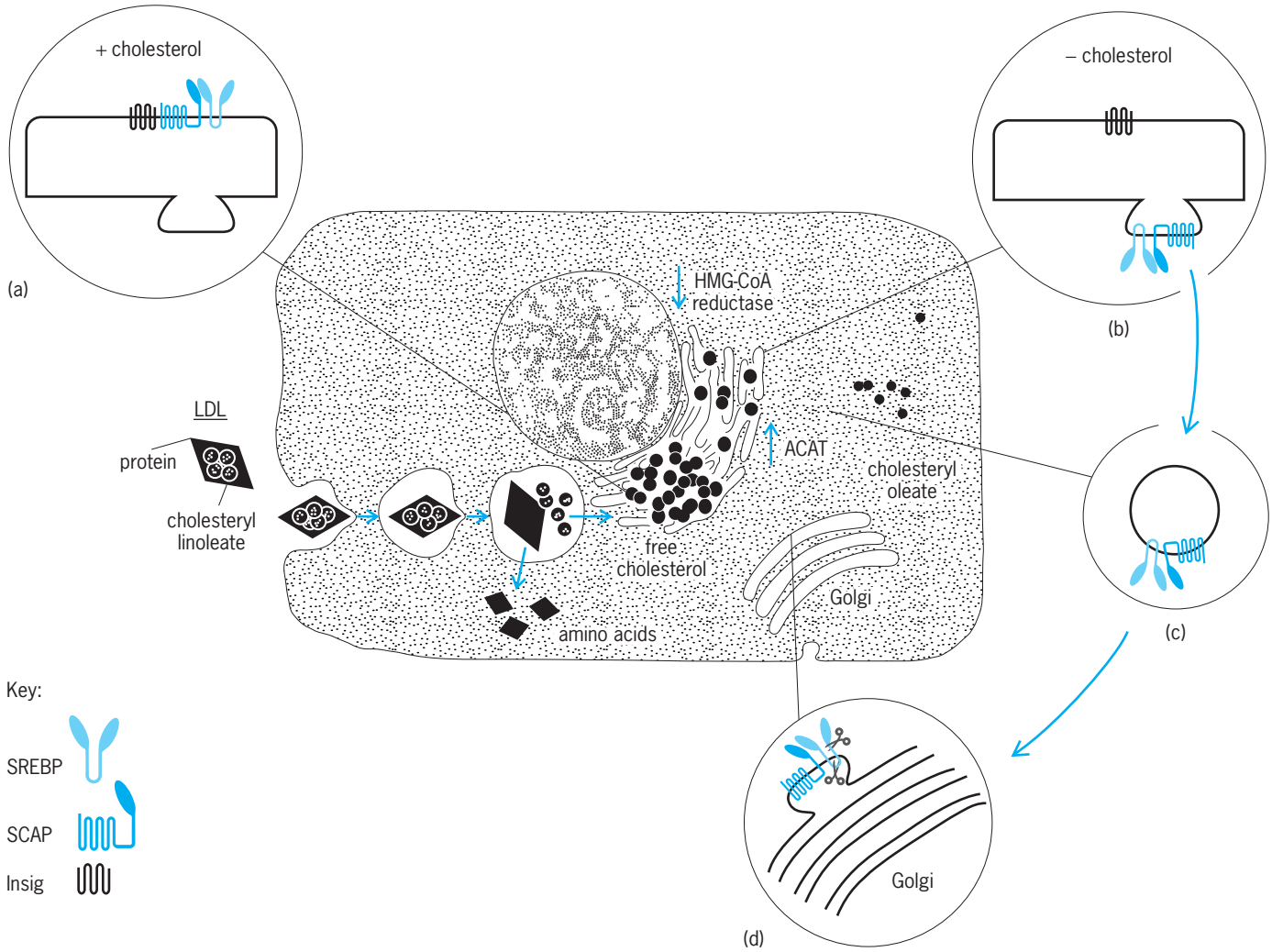

Fig. 2. Insights into the low-density lipoprotein (LDL) receptor internalization pathway, including the uptake of LDL, disassembly in the lysosome, and conversion of the LDL cholesteryl esters into cholesterol, which becomes a part of the regulatory pool of cholesterol inside the cell. (*a*) When cellular cholesterol levels are normal, sterol-regulatory element-binding protein (SREBP), SREBP cleavage-activating protein (SCAP), and Insig-1 form a macromolecular complex in the endoplasmic reticulum (ER). (*b*) Lowering the cholesterol level in the ER stimulates the release of Insig-1 from the complex, which allows SCAP to escort SREBP to vesicles. (*c*) These vesicles bud off and are targeted to carry the SCAP-SREBP complex to the Golgi apparatus. (*d*) In the Golgi, SREBP is cleaved at two sites by proteases, and the sterol-regulatory element-binding fragment is released into the cytoplasm from where it migrates into the nucleus. (*Reprinted from R. G. W. Anderson, J. Goldstein, and M. Brown, From cholesterol homeostasis to new paradigms in membrane biology, Trends Cell Biol., 13:534–539, © 2003, with permission from Elsevier*)

to the Golgi complex inside cells, where the SREBP protein undergoes proteolytic cleavage at two sites by two different proteases, the so-called site 1 and site 2 proteases (Fig. 2*d*). The result is the release of a 480-amino-acid N-terminal portion of the protein that is a transcription factor that travels to the nucleus, where it can bind to the sterol-regulatory element within genes and participate in signaling for their transcription. The SREBP1 transcription factors bind to numerous genes that are involved in fatty acid metabolism, whereas SREBP2 is more involved in promoting the transcription of genes in cholesterol metabolism.

Initially, it appeared that regulation of the release of the processed transcription factor via proteolytic cleavage would be the key regulatory step responsive to cholesterol, but in reality it turned out to be a protein termed SREBP cleavage activating protein, or SCAP (Fig. 2*b*). This protein contains an amino acid sequence termed the sterol-sensing domain (SSD), which is made up of eight membrane-spanning regions that define a signature sequence that is shared with other proteins known to be sensitive to regulation by cholesterol. Several amino acid mutations that abolish cholesterol-mediated regulation occur in the SSD of SCAP. Identification of this fact signaled that it was SCAP that actually detected the level of cholesterol in the membrane. When SCAP binds cholesterol, a conformational change occurs in SCAP that allows the protein to bind to proteins called insigs, which serve as endoplasmic reticulum membrane anchors. When SCAP is bound

to insigs, the SCAP–SREBP complex is retained in the endoplasmic reticulum and cannot participate in vesicular traffic to the Golgi compartment (Fig. 2*a*). The result is that SREBP cannot be processed and the transcriptional component of SREBP cannot be generated and travel to the nucleus and activate cholesterol biosynthetic enzymes. The decrease in nuclear content of processed SREBPs results in a decrease in cholesterol biosynthesis. *See* ENDOPLASMIC RETICULUM; GENE; GENE ACTION; GOLGI APPARATUS.

Metabolism. Steroid hormones are derived from cholesterol and are required for the control of many different cell functions. The conversion of cholesterol to the steroid hormones including cortisol, aldosterone, and androgens occurs in the adrenal glands. Conversion of cholesterol to estrogen, progesterone, or testosterone occurs in the ovary, placenta, or testes. *See* HORMONE; STEROID.

Cholesterol is also a precursor of the bile acids that are synthesized in the liver. Bile acids together with cholesterol and phospholipids are secreted into the bile canaliculi (minute ducts or channels) in the liver. The detergent property of bile acids permits these lipids to become micelles. When the fat in food enters the intestine after a meal, the gallbladder contracts and the liver is stimulated to produce more bile that is secreted into the small intestine, where the bile acids participate in the emulsification (solubilization) of the fat in the diet. The emulsification process allows digestion of the fat to take place efficiently, and the products of digestion, including cholesterol, are absorbed into the intestinal cells. Sometimes, cholesterol in the gall bladder crystallizes and becomes gallstones. The size of the gallstones can vary widely from about the size of a grain of gravel to a lima bean. *See* GALLBLADDER; GALLBLADDER DISORDERS.

A cholesterol biosynthetic precursor, 7-dehydro cholesterol is also the precursor of vitamin D, a hormone that is required for control of normal calcium metabolism. The initial step in the conversion of 7-dehydro cholesterol to vitamin D occurs in the skin in response to ultraviolet light; however, further modifications are required in the liver and kidney in order to form the biologically active form of the vitamin. *See* CALCIUM METABOLISM; LIPID METABOLISM; VITAMIN D.

Plasma lipoproteins and atherosclerosis. Cholesterol and cholesteryl esters are essentially insoluble in water so that, in order to transport these molecules among tissues in the blood, the liver and intestine produce lipid–protein complexes termed lipoproteins. These particles solubilize cholesterol and other lipids, including triglycerides (fats) and phospholipids, so that they can move between tissues to satisfy their metabolic needs. Triglyceride-rich lipoproteins are the largest particles and include the chylomicrons and very low density lipoproteins (VLDL). The LDLs are the major lipoprotein class in human blood, and these particles are enriched in cholesteryl esters, while the smallest particles are the high-density lipoproteins (HDLs).

All of the cholesterol in the blood plasma (after blood cells are removed) is present in the lipoproteins. The relative amounts of the lipoprotein classes are often measured in terms of their cholesterol concentrations as risk factors for the development of atherosclerosis, or hardening of the arteries. The best profile is to have low concentrations of LDL cholesterol (preferably below 100 mg/dl) and to have higher concentrations of HDL cholesterol (above 50 mg/dl). When the concentrations of LDL cholesterol get too high, deposition of cholesterol within the arteries occurs that, over years, can result in accumulations that lead to lesions or plaques in the arterial wall. Higher levels of HDL appear to protect against this deposition. The instability or tendency of plaques to rupture can lead to complications of coronary heart disease, such as myocardial infarction (heart attack). Diets high in cholesterol and saturated fat often result in higher LDL cholesterol concentrations and increased risk for the development of atherosclerosis. *See* ARTERIOSCLEROSIS; HEART DISORDERS.

Lawrence L. Rudel; Peter Edwards

Bibliography. R. G. W. Anderson, J. Goldstein, and M. Brown; From cholesterol homeostasis to new paradigms in membrane biology, *Trends Cell Biol.*, 13:534–539, 2003; J. D. Horton, J. L. Goldstein, and M. S. Brown, *SREBPs: Transcriptional Mediators of Lipid Homeostasis*, vol. 67, pp. 491–498, Cold Spring Harbor Lab Press, 2002; P. A. Mayes and K. M. Botham, Cholesterol synthesis, transport and excretion, in *Harper's Illustrated Biochemistry*, 26th ed., ed. by R. K. Murray et al., pp. 219–230, McGraw-Hill, 2003.

Choline

A compound used by the animal organism as a precursor of acetylcholine and as a source of methyl groups. It is a strongly basic hygroscopic substance with the formula below. It is questionable whether

$$(CH_3)_3{\equiv}N^+{-}CH_2{-}CH_2OH$$

this very necessary metabolite should be considered a vitamin. Choline is determined microbiologically by using a mutant of the mold *Neurospora crassa*.

A choline deficiency has been produced in a number of animal species. The degree of deficiency achieved depends on the amount of choline precursors, particularly methyl donors such as methionine, present in the diet. Choline deficiency is associated with fatty livers, poor growth, and renal lesions. It is a lipotropic agent. There is no direct evidence of disease in humans due to choline deficiency, although there have been suggestions that some of the liver, kidney, or pancreas pathology seen in various nutritional deficiency states may be related to choline insufficiency. Choline is found in acetylcholine, which is necessary for nerve impulse propagation, and in phospholipids.

There is no evidence that choline acts as a coenzyme. Choline functions as a donor of methyl

groups to quaternary nitrogen and sulfonium compounds.

Human beings eat 50–600 mg of choline per day, but only excrete 2–4 mg. Consequently, conventional load tests are of no value in studying choline requirements, and no knowledge of human requirements exists. *See* ACETYLCHOLINE; VITAMIN.

Stanley N. Gershoff

Choline, trimethyl-β-hydroxyethylammonium hydroxide, is a quaternary compound which is made by the reaction of trimethylamine and ethylene oxide in either an aqueous or an alcoholic solvent. If choline salts are desired, the reaction can be carried out with the corresponding trimethylamine salt, or the base can be neutralized with the corresponding acid. The salts in common use are choline chloride, choline bitartrate, choline dihydrogen citrate, and tricholine citrate.

Choline salts are sold as liquids and solids. The liquids, such as a 65% solution of tricholine citrate, are prepared by the aqueous quaternization procedure, decolorized with charcoal, and then evaporated to the desired strength. The solid types are prepared in either aqueous or alcoholic media and are crystallized by cooling concentrated solutions.

Richard J. Turner

Chondrichthyes

A class of fishes characterized by well-developed jaws and bony teeth; pectoral and pelvic girdles, each supporting paired fins; a cartilaginous skeleton that lacks true bone; toothlike scales of ectodermal and mesodermal origin; paired nostrils with blind olfactory sacs (not opening into the mouth); a vertebral column having either a straight or heterocercal posterior end; absence of a swim bladder; cranium without sutures; and internal fertilization.

Chondrichthyes has traditionally included the subclasses Holocephali (chimaeras) and Elasmobranchii (sharks, rays, and skates). A group of Devonian armored fishes, the Placodermi, has been regarded as ancestral to the Chondrichthyes, but this derivation is not certain. Another group of primitive jawed fishes called acanthodians, considered by many as ancestral to the higher bony fishes, exhibit certain primitive elasmobranch-like features. In any case, it is probable that the elasmobranchs and chimaeras arose independently of each other some time during the Silurian or Early Devonian. Because even calcified cartilage is rarely preserved, the fossil record of the Chondrichthyes is represented mainly by teeth and spines, with only occasional associated skeletons. *See* ACANTHODII; PLACODERMI.

Holocephali. The Holocephali differ from the Elasmobranchii in having only four pairs of gill arches and gills that open to the exterior from a single pair of apertures; in the erectile dorsal fin and spine; in the naked skin in adults; and in the absence of a cloaca and ribs. Males of some species are equipped with a frontal clasper on the head. The teeth are consolidated into six pairs of plates, and the upper jaw is immovably fused with the brain case; these adaptations function in grinding mollusks, their chief food. Chimaeras date from the Early Mesozoic. *See* CHIMAERIFORMES.

Elasmobranchii. The elasmobranchs have five to seven pairs of gill clefts, all opening separately; one or two dorsal fins; rigid spines preceding the fins of some taxa; spiracles either present or absent; placoid scales; numerous teeth that are not set in sockets and arranged in several rows; the inner margin of the pelvic fins of males modified to form a copulating organ called claspers; and oviparous, ovoviviparous, or viviparous development. *See* BATOIDEA; ELASMOBRANCHII; SELACHII.

Herbert Boschung

Bibliography. H. B. Bigelow and W. C. Schroeder, Sharks, in *Fishes of the Western North Atlantic*, Mem. Sears Found. Mar. Res. Mem. 1 (pt. 1), pp. 29–556, 1948; E. S. Herald, *Living Fishes of the World*, 1961; J. G. Maisey, Heads and tails: A chordate phylogeny, *Cladistics*, 2(3):201–256, 1986; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; A. S. Romer, *Vertebrate Paleontology*, 3d ed., 1966; B. Schaeffer, Comments on elasmobranch evolution, in P. W. Gilbert et al. (eds.), *Sharks, Skates, and Rays*, 1967.

Chonotrichida

An order of the Holotrichia. This is a small group of curious, vase-shaped ciliates, commonly found as ecotocommensals on marine crustaceans. They attach to their hosts by a short, secreted stalk, and their bodies are practically devoid of ciliature.

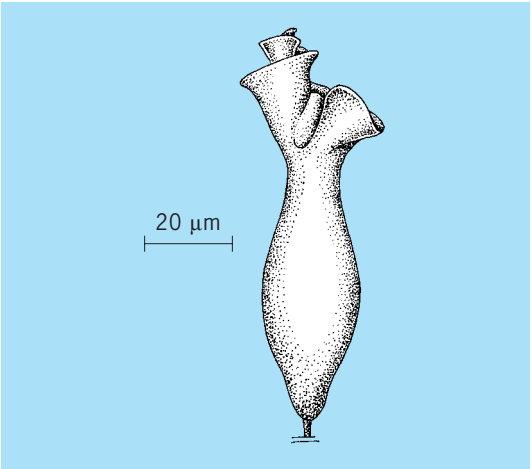

***Spirochona*, an example of a chonotrichid.**

Spirochona (see **illus.**) is a common example of this order. *See* CILIOPHORA; HOLOTRICHIA; PROTOZOA.

John O. Corliss

Chopping

The act of interrupting an electric current, beam of light, or beam of infrared radiation at regular intervals. This can be accomplished mechanically by rotating a vibrating mirror in the path of the beam to deflect it away from its intended source at regular intervals. A current can be chopped with an electromagnetic vibrator having contacts on its moving armature. A current can also be chopped electronically by passing it through a multivibrator or other switching circuit. Chopping is generally used to change a direct-current signal into an alternating-current signal that can more readily be amplified. Control systems for guided missiles make extensive use of chopping. *See* MULTIVIBRATOR.

Chopping has been increasingly used inside analog integrated circuits. Solid-state switches and capacitors are used to chop operational amplifiers, greatly improving their offset voltages. Chopping is also used in analog large-scale-integrated switched-capacitor filters as a means for reducing their undesirable $1/f$ (inverse-frequency) noise. *See* AMPLIFIER; INTEGRATED CIRCUITS; OPERATIONAL AMPLIFIER; SWITCHED CAPACITOR. Eric J. Swanson

Bibliography. R. L. Geiger, P. E. Allen, and N. R. Strader, *VLSI Design Techniques for Analog and Digital Circuits*, 1990.

Chordata

The phylum of animals characterized by possession of a notochord, which includes the Cephalochordata (lancelets), Urochordata (tunicates), and Craniata (vertebrates and hagfishes). The phylum also may include the extinct Yunnanozoons and Vetulicolians. Members of the first two groups are small and strictly marine. The vertebrates include the fishes, amphibians, reptiles, birds, and mammals; they are usually free-living and are found in many environments from marine to freshwater and in almost all terrestrial biomes. Some workers include the acorn worms, pterobranchs, and graptolites (Hemichordata) as a subphylum within Chordata, but here it is considered as a separate phylum, because recent morphological and molecular data indicate these organisms are closely related to echinoderms and do not possess a notochord. However, the phylogenetic position of these animals is still controversial. *See* CEPHALOCHORDATA; HEMICHORDATA; TUNICATA; VERTEBRATA.

Characteristics. The typical chordate characteristics include a notochord, a dorsal hollow nerve cord, pharyngeal slits, and a postanal tail. All chordates are deuterostomes, and most chordates also have bilateral symmetry and a somatic arrangement that includes a head, trunk, and tail region, which they share with their common ancestor. *See* DEUTEROSTOMIA.

Notochord. The notochord arises in the chordate embryo as a slender, flexible rod filled with gelatinous cells and surrounded by a tough fibrous sheath. It contains, at least in some chordates, transverse striated muscle fibers and lies above the developing gut. The notochord persists as the axial support for the body throughout the animal's life in cephalochordates and some urochordates, but it is surrounded and gradually replaced by segmental vertebrae in the Vertebrata, starting with the early fishes.

Nerve cord. The dorsal hollow nerve cord grows from a specialized band of neurectoderm along the middorsal surface of the embryo by folding together two parallel ridges. The anterior end enlarges slightly in larval tunicates and somewhat more in lancelets, but enlarges greatly in the vertebrates to form the brain. Vertebral evolution is characterized by continual enlargement of the brain. Moreover, the amount of neurectoderm in the vertebrates is larger than that in the other two groups, and the lateral rims form the neural crests, the cells of which migrate to all parts of the embryo and develop into a number of diverse structures in the body. In the vertebrates or craniates, the brain is enclosed in the braincase or cranium, and the spinal cord is enclosed by the neural arches. *See* NERVOUS SYSTEM (VERTEBRATE).

Pharyngeal slits. Pharyngeal slits are paired slits that develop as outpocketings from the sides of the embryonic pharynx and connect the lumen of the digestive system to the external environment; they are retained as gills in all aquatic chordates. Pharyngeal slits originated as adaptatations for filter feeding but soon became the primary respiratory organ, as blood vessels line the fine filaments on the margins of each slit. Internal gills were lost with the origin of tetrapods. Larval and some adult amphibians possess external gills that are different in structure from the earliest internal gill slits. In embryonic tetrapods, the pharyngeal slits close early in ontogeny, with the pharyngeal pouches becoming the site for development of glands, such as the parathyroids, thymus, and tonsils in mammals. *See* RESPIRATORY SYSTEM.

Postanal tail. The chordate tail is part of the skeletal, muscular, and nervous system of the animal and continues posterior to the anus. It is a feature not found in any other animal group and serves to increase the force available to the organism for locomotion.

Affinities. Much controversy still exists about the origin and affinities of the chordates, although there is no question that the Hemichordata (acorn worms and pterobranchs) are closely related and part of the pharyngeal-slit filter-feeding radiation. Almost all workers agree that the Echinodermata are allied to the Chordata because of evidence ranging from embryonic development (they are both deuterostomes) to biochemical similarities, but early chordate affinities are just now coming to light. *See* ECHINODERMATA.

Evolutionary origins. With the discovery of *Yunnanozoon lividum* in China, the earliest fossil record of putative chordates now dates back to the Early Cambrian (525 million years ago). *Myllokunmingia*, from the early Cambrian Chengjiang fossil site in southwestern China, and *Pikaia*, from the Middle Cambrian Burgess Shale in Canada, are also

considered to be the earliest chordates. These organisms appear to have chordate characteristics (though they are still being debated), including an impression of a notochord, segmented muscles, and what appear to be pharyngeal slits. *Pikaia* resembles the present-day lancelet or amphioxus (Cephalochordata), which burrows in shifting sands. Presumably all other chordates developed from this ancestral type, with their differing characteristics evolving because of differing habitat conditions.

There are a number of hypotheses for how chordates evolved. One theory postulates that chordates arose from an echinoderm ancestor, because they share similar embryonic developmental patterns (for example, they are both deuterostomes and exhibit similar cleavage and coelomic formation). Another hypothesis suggests that chordates have evolved from tunicate larval form; however, it is far more likely that the vertebrates descended directly from an amphioxus-like form.

What is known is that the first chordate evolved the notochord, which provided a stronger skeleton and permitted the reduction of the transverse and circular muscles. A notochord resulted in a fixed body length and the loss of peristaltic locomotion. Having a notochord for support permitted the appearance of pharyngeal slits through the lateral walls of the anterior parts of the body, which served for increased filter feeding and subsequently for respiration. The presence of the notochord also permitted the appearance of a postanal tail and increased force for locomotion. The dorsal longitudinal muscles enlarged, and with this modification came the evolution of the dorsal hollow nerve cord (dorsal because of closer proximity with the major muscle masses).

Tunicates (Urochordata) are either sessile organisms that live on hard substrates (Class Ascidiacea) or are planktonic (Class Thaliacea and Class Larvacea). Ascidians and thaliaceans have lost the notochord feature in the metamorphosed adult but retain the pharyngeal gill slits and an endostyle (a glandular ciliated groove along the ventral side of the pharynx). Larvae of all tunicate groups (and adult larvacians) have retained the ancestral chordate features (gill slits and notochord).

The most recently evolved group was the Vertebrata, which are presumed to have left the oceans to invade freshwaters and then the land. They retained the full set of chordate features but had to evolve several additional ones. Size and swimming ability increased to swim upstream against the currents. Because of an increased inward flow of water across the body surface, the problem of water balance was solved when the kidneys, with the nephron as the basic unit, evolved to rid the body of excess water. Finally, dermal bone evolved as armored protection against predators. Embryological development of the nerve cord also changed with the appearance of the neural crest, a unique vertebrate feature that is involved in the development of bony scales, teeth, gill bars, pigment cells, and several parts of the nervous system. *See* AMPHIBIANS; ANIMAL KINGDOM; AVES; MAMMALIA; REPTILIA. Walter J. Bock

Bibliography. M. J. Benton, *Vertebrate Paleontology*, 1990; R. L. Carroll, *Vertebrate Paleontology and Evolution*, W. H. Freeman, 1988; W. F. Gutmann, Relationships between invertebrate phyla based on functional-mechanical analysis of the hydrostatic skeleton, *Amer. Zool.*, 21:63–81, 1981; W. F. Gutmann, The hydraulic principle, *Amer. Zool.*, 28:257–66, 1988; K. V. Kardong, *Vertebrates*, Wm. C. Brown, 1995.

Chorion

The outermost of the several extraembryonic membranes in amniotes (reptiles, birds, and mammals) enclosing the embryo and all of its other membranes. The chorion, or serosa, is composed of an outer layer of ectodermal cells and an inner layer of mesodermal cells, collectively the somatopleure. Both layers are continuous with the corresponding tissue of the embryo. The chorion arises in conjunction with the amnion, another membrane that forms the outer limb of the somatopleure which folds up over the embryo in reptiles, birds, and some mammals. The chorion is separated from the amnion and yolk sac by a fluid-filled space, the extraembryonic coelom, or body cavity. In those mammals in which the amnion forms by a process of cavitation in a mass of cells, instead of by folding, the chorion forms directly from the trophoblastic capsule, the extraembryonic ectoderm, which becomes gradually underlain by extraembryonic mesoderm.

In reptiles and birds the chorion fuses with another extraembryonic membrane, the allantois, to form the chorioallantois, which lies directly below the shell membranes. An extensive system of blood vessels develops in the mesoderm of this compound membrane which serves as the primary respiratory and excretory organ for gaseous interchanges. In all mammals above the marsupials, the chorion develops special fingerlike processes (chorionic villi) extending outward from its surface. To a varying degree in different species of mammals, the villous regions of the chorion come into more or less intimate contact with the uterine mucosa, or uterine lining, of the mother, thereby forming the various placental types. *See* ALLANTOIS; AMNIOTA; FETAL MEMBRANE; GERM LAYERS. Nelson T. Spratt, Jr.

Chromadorida

An order of nematodes in which the amphid manifestation is variable but within superfamilies some constancy is apparent. The various amphids are reniform, transverse elongate loops, simple spirals, or multiple spirals not seen in any other orders or subclasses. The cephalic sensilla are in one or two whorls at the extreme anterior. In all taxa the cuticle shows some form of ornamentation, usually punctations that are apparent whether the cuticle is smooth or annulated. When developed, the stoma is primarily esophastome and is usually armed with a dorsal

tooth, jaws, or protrusible rugae. The corpus of the esophagus is cylindrical; the isthmus is not seen; and the postcorpus, in which the heavily cuticularized lumen forms the cresentic valve, is distinctly expanded. The esophagointestinal valve is triradiate or flattened. The females usually have paired reflexed ovaries.

There are four chromadorid superfamilies: Choanolaimoidea, Chromadoroidea, Comesomatoidea, and Cyatholaimoidea. Choanolaimoidea is distinguished by a complex stoma in two parts. The group occupies marine habitats; some species are predaceous, but for many the feeding habits are unknown. Chromadoroidea comprises small to moderate-sized free-living forms that are mainly marine but are also found in fresh water and soil. Known species either are associated with algal substrates or are nonselective deposit feeders in soft-bottom sediments. Comesomatoidea, containing only one family, the Comesomatidae, is found in marine habitats, but the feeding habits are unknown. Cyatholaimoidea is found in marine, terrestrial, and fresh-water environments. *See* NEMATA.

Armand R. Maggenti

Chromatic aberration

The type of error in an optical system in which the formation of a series of colored images occurs, even though only white light enters the system. Chromatic aberrations are caused by the fact that the refraction law determining the path of light through an optical system contains the refractive index n, which is a function of wavelength λ. Thus the image position and the magnification of an optical system are not necessarily the same for all wavelengths, nor are the aberrations the same for all wavelengths.

In this article the chromatic aberrations of a lens system are discussed. For information on other types of aberration *see* ABERRATION (OPTICS); REFRACTION OF WAVES.

Dispersion formula. When the refractive index of glass or other transparent material is plotted as a function of the square of the wavelength, the result is a set of dispersion curves which appear to have an asymptote in the near-ultraviolet region and a straight portion in the near-infrared (up to 1 micrometer). In the glass catalogs the indices are given for selected wavelengths between 0.365 and 1.01 μm, which are the absorption (emission) bands of certain chemical elements. For an extended discussion of dispersion *see* ABSORPTION OF ELECTROMAGNETIC RADIATION; DISPERSION (RADIATION); OPTICAL MATERIALS.

For the visible region, the Hartmann formula, Eq. (1), where a and b are constants for a given glass,

$$n = a/(\lambda - b) \tag{1}$$

has been much used. However, if an attempt is made to apply it to the near-ultraviolet or the near-infrared, it proves to be insufficient.Over this more extended range, n can be given as a function of the wavelength by Eq. (2).

$$n = a + b\lambda^2 + \frac{c}{\lambda^2 - 0.028} + \frac{d}{(\lambda^2 - 0.028)^2} \tag{2}$$

This formula, in which c and d are additional constants, enables one to compute n for any wavelength if it is given for four. An equivalent formula is shown by Eq. (3), where $a_1(\lambda)$, $a_2(\lambda)$, ... are functions of

$$n = n_1 a_1(\lambda) + n_2 a_2(\lambda) + n_3 a_3(\lambda) + n_4 a_4(\lambda) \tag{3}$$

the form of Eq. (2), which assume for λ_1, λ_2,... the values 1, 0, 0, 0; 0, 1, 0, 0; 0, 0, 1, 0; and 0, 0, 0, 1, respectively. Choosing $\lambda_1 = 0.365 = \lambda^{**}$, $\lambda_2 = 0.4861 = \lambda_F$, $\lambda_3 = 0.6563 = \lambda_C$, $\lambda_4 = 1.014 = \lambda^*$, one finds in the **table** the four universal functions $a(\lambda)$ tabulated for different values of λ. Note that for any wavelength Eq. (4) holds.

$$a_1 + a_2 + a_3 + a_4 = 1 \tag{4}$$

Instead of giving the indices of a glass for the wavelengths specified, one can specify the glass by the refractive index n_F at one wavelength F, the dispersion $\delta = n_F - n_C$, and the partial dispersions in the ultraviolet and infrared, defined respectively by Eqs. (5).

$$P^{**} = \frac{n^{**} - n_F}{n_F - n_C} \qquad P^* = \frac{n_F - n^*}{n_F - n_C} \tag{5}$$

The partial dispersion for an arbitrary wavelength λ is then given by Eq. (6).

$$P = P^{**} a_1(\lambda) + a_2(\lambda) + P^* a_4(\lambda) \tag{6}$$

In the literature the Abbe number, defined by Eq. (7), where n_D is the mean index for the D lines of

$$\nu = (n_D - 1)/(n_F - n_C) \tag{7}$$

sodium, is frequently used to designate optical glass.

In **Fig. 1**, $(n_F - 1)$ is shown plotted against $(n_F - n_C)$ for various optical materials. One sees that in this kind of plot glasses of the same type lie on a straight line, while on the ordinary plot of n_D versus ν they lie on hyperbolas.

Values of the four universal functions $a(\lambda)$ for various wavelengths

λ	a_1	a_2	a_3	a_4
1.0140	0.000000	0.000000	0.000000	+1.000000
0.7682	+0.031555	−0.276197	+1.051955	+0.192687
0.6563	0.000000	0.000000	+1.000000	0.000000
0.6438	−0.004774	+0.051075	+0.966511	−0.012813
0.5893	−0.024952	+0.326272	+0.744598	−0.045919
0.5876	−0.025492	+0.336338	+0.735480	−0.046326
0.5461	−0.033080	+0.597795	+0.479049	−0.043764
0.4861	0.000000	+1.000000	0.000000	0.000000
0.4800	+0.009340	+1.036699	−0.052699	+0.006659
0.4358	+0.143980	+1.193553	−0.394569	+0.057036
0.4047	+0.366779	+1.045064	−0.487634	+0.075791
0.3650	+1.000000	0.000000	0.000000	0.000000

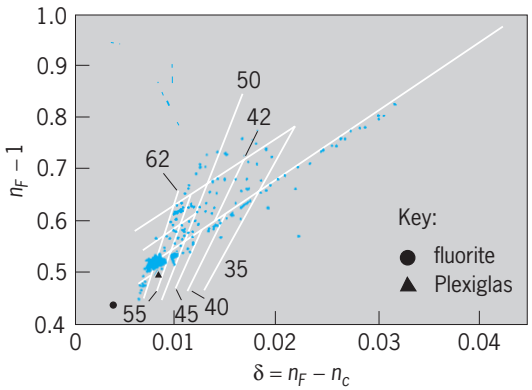

Fig. 1. Plot of ($n_F - 1$) versus δ for selected glasses and for fluorite and Plexiglas. The numbers on the lines dividing the glasses into groups represent ν values.

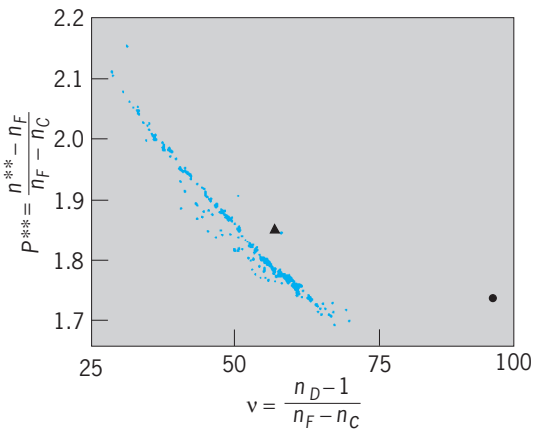

Fig. 2. Plot of P^{**} versus ν for materials of Fig. 1.

Color correction. A system of thin lenses in contact is corrected for two wavelengths λ_A and λ_B if the power of the combination is the same for both wavelengths as in Eq. (8), where the K_κ's are the difference s between the first and last curvatures of the lenses, and the summation is over all the lenses, each with its particular value of n_A, n_B, and K_κ. Two cemented lenses are corrected for wavelengths C and F, for instance, if Eqs. (9) hold.

$$\phi_A = \phi_B = \Sigma(n_A - 1)K_k = \Sigma(n_B - 1)K_k \qquad (8)$$

$$\begin{aligned} \phi_1 + \phi_2 &= \phi \\ \phi_1/\nu_1 + \phi_2/\nu_2 &= \phi_C - \phi_F = 0 \end{aligned} \qquad (9)$$

See LENS (OPTICS).

The two lenses are also corrected for a third wavelength if and only if, in addition, one has $P_1 = P_2$, where P_1 and P_2 are given by Eq. (6).

A system corrected for three colors is called an apochromat (in microscopy this term traditionally demands freedom from asymmetry in addition). An apochromat for the ultraviolet portion of the spectrum is possible only if the two glasses in **Fig. 2** have the same P^{**} value. For the infrared, the glasses must have the same P^* value (**Fig. 3**). The $\nu - 1$ values for the two glasses should lie as far apart as possible to give low values for the powers.

Three lenses can be corrected for four wavelengths and therefore practically for the whole spectrum if the glasses lie on the straight line on the plot of P^{**} against P^* (**Fig. 4**). Such a system may be called a superachromat.

In lenses with finite thicknesses and distances, there are in gaussian optics two errors to be corrected. One is a longitudinal aberration, which means that the gaussian images do not lie in the same plane, and the other is a lateral aberration, which means that the images in different colors have different magnifications.

In the presence of longitudinal aberration, it is best to balance the lateral aberration so that the apparent sizes of the images as seen from the exit pupil coincide.

A system of two uncorrected lenses, such as a simple ocular, with one finite distance, cannot be corrected for both color errors. Two finite distances are needed to balance both color errors at the same time.

In a general system, all image errors are functions of the wavelength of light. However, in a lens system corrected for color, it is easily possible, with a small adjustment, to balance the correction of the aperture rays (spherical aberration) with respect to

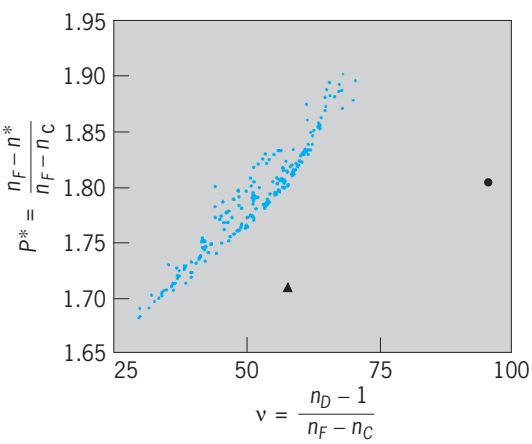

Fig. 3. Plot of P^* versus ν for materials of Fig. 1.

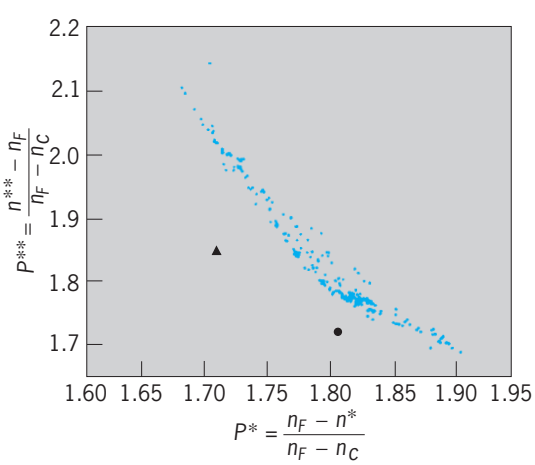

Fig. 4. Plot of P^{**} versus P^* for materials of Fig. 1.

color by introducing a small amount of lateral color aberration.

If only two colors can be corrected, the choice of the colors depends on the wavelength sensitivity of the receiving instrument. For visual correction, the values for *C* and *F* are frequently brought together. Some optical systems contain filters to permit only a narrow spectral band to pass the instrument. This makes correction for color errors easier.

If light of a large band of wavelengths traverses the instrument, it is frequently desirable to use catadioptric systems, such as the Schmidt camera. The mirror or mirrors of these systems are used to obtain the necessary power without introducing color errors, and an afocal lens system can be added to correct monochromatic aberrations. *See* SCHMIDT CAMERA.

Max Herzberger

Bibliography. E. U. Condon and H. Odishaw, *Handbook of Physics*, 2d ed., 1967; M. Herzberger, *Modern Geometrical Optics*, 1958, reprint 1978; F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976; J. Mayer-Arendt, *Introduction to Classical and Modern Optics*, 2d ed., 1984; Optical Society of America, *Handbook of Optics*, 2d ed., 1994.

Chromatography

A physical method of separation in which the components of a mixture of substances are separated from each other by differences in their distribution between two phases, one (stationary phase) is stationary and the other (mobile phase) moves through the stationary phase in a definite direction. The substances must interact with the stationary phase to be retained and separated by it.

Retention and separation. Retention results from a combination of reversible physical interactions that can be characterized as adsorption at a surface, absorption in an immobilized solvent layer, and electrostatic interactions between ions. When the stationary phase is a porous medium, accessibility to its regions may be restricted and a separation can result from size differences between the sample components. More than one interaction may contribute simultaneously to a separation mechanism. The general requirements are that all interactions must be reversible, and that the two phases can be separated (two immiscible liquids, a gas and a solid, and so forth) in such a way that a distribution of sample components between phases and mass transport by one phase can be established. *See* ABSORPTION; ADSORPTION.

Reversibility of the interactions can be achieved by purely physical means, such as by a change in temperature or by competition; the latter condition is achieved by introducing substances into the mobile phase that have suitable properties to ensure reversibility for the interactions responsible for retention of the sample components. Since this competition with the sample components is itself selective, it provides a general approach to adjusting the outcome of a chromatographic experiment to obtain a desired separation. It is an absolute requirement that a difference in the distribution constants for the sample components in the chromatographic system exist for a separation to be possible.

Methods. A distinction between the principal chromatographic methods can be made in terms of the properties of the mobile phase and configuration of the stationary phase. In gas chromatography the mobile phase is an inert gas, in supercritical fluid chromatography the mobile phase is a fluid (dense gas above its critical pressure and temperature), and in liquid chromatography the mobile phase is a low-viscosity liquid. The stationary phase can be a porous, granular powder with a narrow particle-size distribution packed into a tube (called a column) as a dense homogeneous bed. This configuration is referred to as a packed column and is nearly always used in liquid and supercritical fluid chromatography, and is sometimes used in gas chromatography. Alternatively, the stationary phase can be distributed as a thin film or layer on the wall of an open tube of capillary dimensions, leaving an open space through the center of the column. This configuration is referred to as an open tubular column (or incorrectly as a capillary column); it is commonly used in gas chromatography, but rarely used in supercritical fluid or liquid chromatography.

Thin-layer chromatography is a form of liquid chromatography in which the stationary phase is spread as a thin layer over the surface of a glass or plastic supporting structure. The stationary phase must be immobilized on the support by using a binder to impart the desired mechanical strength and stability to the layer. The samples are applied to the layer as spots or bands near the bottom edge of the plate. The separation is achieved by contacting the bottom edge of the plate below the line of samples with the mobile phase, which proceeds to ascend the layer by capillary action. This process is called development and is performed in a chamber, with the lower edge of the layer in contact with the mobile phase and the remaining portion of the layer in contact with solvent vapors from the mobile phase. The chamber may be a simple device such as a covered jar or beaker or a more elaborate device providing control of the mobile-phase velocity and elimination or control of the vapor phase. Thin-layer chromatography is the most popular form of planar chromatography having virtually replaced paper chromatography in laboratory practice. *See* GAS CHROMATOGRAPHY; GEL PERMEATION CHROMATOGRAPHY; LIQUID CHROMATOGRAPHY; SUPERCRITICAL-FLUID CHROMATOGRAPHY.

Uses. Chromatographic methods provide a means of analyzing samples (to determine component identity and relative amount), isolating significant quantities of purified material for further experimentation or commerce, and determining fundamental physical properties of either the samples or the mobile or stationary phases (for example, diffusion coefficients, partition coefficients, or thermodynamic properties). There are virtually no boundaries to the sample types that can be separated. Examples include organic and inorganic compounds in the

form of fixed gases, ions, polymers, as well as other species. Applications are found in all areas of science and technology, making chromatography one of the most widely used laboratory procedures in chemistry. Depending on intent, chromatography can be applied to trace quantities at about the limit of detector response (for example, 10^{-15} g) or to kilogram amounts in preparative separations.

Instrumentation. Modern chromatographic methods are instrumental techniques in which the optimal conditions for the separation are set and varied by electromechanical devices external to the column or layer. Separations are largely automated, with important features of the instrumentation being control of the flow and composition of the mobile phase, introduction of the sample onto the stationary phase, and online detection of the separated components. In column chromatography, the sample components are detected in the presence of the mobile phase after they have exited the stationary phase. In thin-layer chromatography, the sample components are detected in the presence of the stationary phase, resulting in different detection strategies.

Instrument requirements differ by the needs of the method used. Gas chromatography uses a mobile phase of constant composition at a few atmospheres of column inlet pressure and variation in the temperature of the column to effect a separation. Liquid chromatography uses a pump to select or vary the composition of the mobile phase with a high column inlet pressure (typically a few hundred atmospheres) and a constant temperature for the separation. These differences in optimized separation conditions result in different equipment configurations for each chromatographic method.

Interpretation. The results of a chromatographic experiment are summarized in a chromatogram (**Fig. 1**), a two-dimensional record of the detected response to the sample components (y axis) plotted against the residence time of the components in column chromatography or migration distance in planar chromatography (x axis). Individual compounds or mixtures of unseparated compounds appear as peaks in the chromatogram. These peaks are ideally symmetrical and occur at positions in the chromatogram that are characteristic of their identity, with a distribution around the mean position (apex of the peak) that is characteristic of the kinetic properties of the chromatographic system. The area inscribed by the peak is proportional to the amount of substance separated in the chromatographic system.

Information readily extracted from the chromatogram includes an indication of sample complexity (the number of observed peaks), qualitative substance identification (determined by peak position), relative composition of the sample (peak dimensions; area or height), and a summary of the kinetic characteristics of the chromatographic system (peak shapes).

There is a synergistic relationship between the peak profiles and the properties of the detector. The detector may have an approximately equal response to all substances, in which case the peak sizes are roughly equal to their relative concentration in the sample; or the detector may have a selective response to some property of the sample components, such as a particular chromophore or the presence of a particular element, in which case the peak sizes have to be interpreted with respect to prior knowledge of the way a particular compound responds to the detector. This is also a useful feature, because the information in the chromatogram can be edited to suit an individual application. For example, the determination of organophosphorus pesticides in environmental samples by gas chromatography can be made with a detector that responds selectively to phosphorus, so that the information contained in the chromatogram is limited to the number of phosphorus-containing compounds in the sample and is easier to interpret than a chromatogram in which all compounds present in the sample are revealed by the detector.

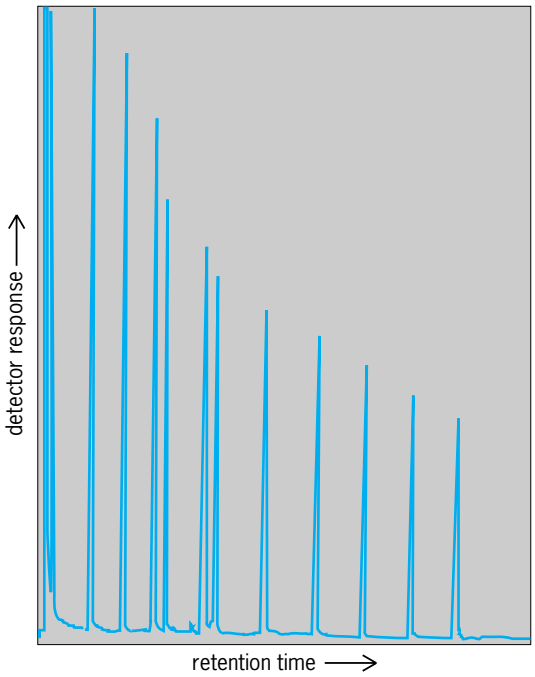

Fig. 1. Typical chromatogram obtained by gas chromatography.

Quality of separation. The object of a chromatographic separation is to obtain a minimum acceptable separation (resolution) between all components of interest in a sample within the shortest possible time. The resolution between two peaks in a chromatogram depends on how well the peak maxima are separated and how wide the two peaks are. This can be expressed numerically, at least for peaks of approximately equal height, by the ratio of the separation of the two peak maxima divided by the average peak widths at their base. When two peaks are separated such that they are just touching at their base, the value calculated for their resolution is 1.5; for difficult-to-separate samples a lower resolution value of 1.0, representing 94% peak separation (6% of the peak areas overlap), is an adequate goal for identification and quantification based on the

characteristic shape of the individual peaks. Resolution is also simply related to the properties of the chromatographic system by the equation below,

$$R_s = \frac{N^{1/2}}{4}\frac{\alpha - 1}{\alpha}\frac{k_2}{1 + k_2}$$

where R_s is the observed resolution, N is the column efficiency measured as the plate number, α is the separation factor (a measure of the relative peak maxima positions; $\alpha = k_2/k_1$), k is the retention factor (the ratio of the time a substance spends in the stationary and mobile phases), and the subscripts 1 and 2 refer to the peak order (peak 2 emerges from the column after peak 1). The equation is derived with the assumption that the average of the two peak widths is identical to the peak width of the second peak, generally a reasonable assumption for peaks that are close to each other in the chromatogram. To a first approximation, the three terms in the equation can be treated as independent variables and optimized separately to obtain a desired separation in the chromatogram.

Efficiency. The efficiency of a chromatographic system is expressed as the plate number. This terminology was adopted from the application of the theory of fractional distillation columns to gas chromatography when this technique was introduced. The term has remained in use, although any analogies between chromatography and fractional distillation are inappropriate and the concept of a theoretical plate has no physical meaning as presently used in chromatography. From a mathematical perspective, efficiency is a characteristic property of the extent of peak dispersion that occurs during the time a compound is resident in the chromatographic system.

Systems with large numerical values for their efficiency are preferred to optimize resolution. Since the resolution is proportional to $N^{1/2}$, unless N can be increased by large increments only small changes in resolution will be obtained. The numerical value for N is governed by kinetic factors that control peak broadening (N affects only peak widths and not peak positions). The chromatographic system should be optimized to minimize peak broadening, leading to better separations. All things being equal, N could never be too large, and chromatographic systems with intrinsically large values of N are desirable. The corollary is also true that systems with intrinsically small values of N are undesirable and diminish the prospects of achieving a desired separation unless other factors in the chromatographic system can be adjusted to compensate for the poor kinetic performance of the column (**Fig. 2***a*).

The rate theory of peak broadening recognizes three processes that contribute to peak widths beyond purely instrumental contributions: (1) resistance to mass transfer between phases prevents instantaneous equilibrium within the chromatographic system; (2) flow anisotropy in packed beds results from variations in the flow characteristics of the bed in regions of unequal packing density; (3) longitudinal diffusion is the natural consequence of molecular diffusion along the axis of migration of the mobile phase. The purpose of column design is to minimize these contributions to peak broadening to achieve the most favorable value of N. The optimum column design is different for each chromatographic method, although it can be easily predicted from the characteristic properties of the chromatographic system and the desired operating conditions (see **table**).

Resistance to mass transfer is often the major cause of peak dispersion in chromatography, and it is related to the absolute magnitude and the relative magnitude of the solute diffusion coefficients in the mobile and stationary phases, and the fact that the relative velocity of the mobile phase is significant on the time scale for these other processes. Resistance to mass transfer, therefore, prevents true equilibrium from being established in the chromatographic system, resulting in additional peak dispersion from that predicted from an equilibrium model.

In gas chromatography, the rate of radial diffusion defines the useful range of column radii for open

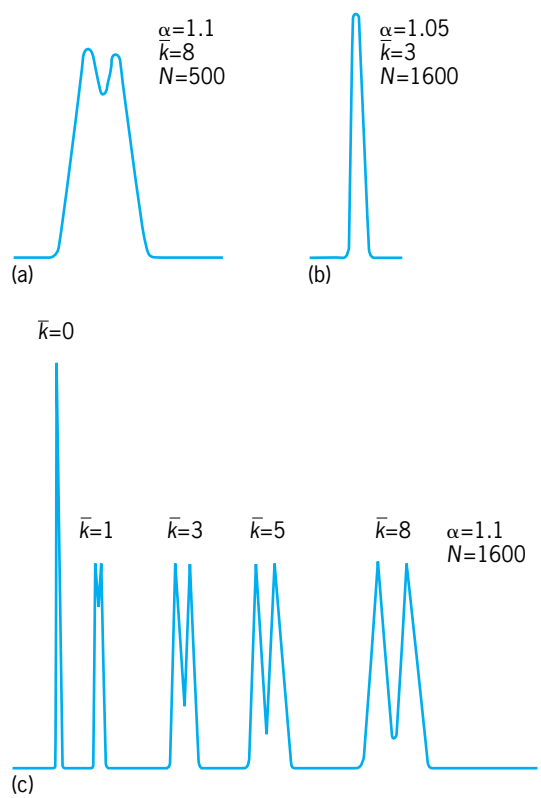

Fig. 2. Observed changes in resolution in a two-component separation for different values of the separation factor, average retention factor, or plate number. (*a*) Viewed vertically, this chromatogram indicates the change in resolution at the same average retention factor and separation factor values when the plate number is varied. (*b*) Viewed vertically, this chromatogram indicates the relationship between resolution and changes in the separation factor with the same plate number and average retention factor value. (*c*) Viewed horizontally, this chromatogram illustrates the change in resolution as the average retention factor is varied with the same value for the separation factor and plate number. (*After P. J. Schoenmakers, Optimization of Chromatographic Selectivity, Elsevier, 1986*)

Approximate values of characteristic parameters in predicting peak broadening

Parameter	Gas	Supercritical fluid	Liquid
Diffusion coefficient (m^2/s)	10^{-1}	10^{-4} to 10^{-3}	10^{-5}
Density (g/cm^3)	10^{-3}	0.3–0.8	1
Viscosity (poise)	10^{-4}	10^{-4} to 10^{-3}	10^{-2}
Column length (m): packed	1–5	0.1–1.0	0.05–1.0
open tubular	10–100	5–25	
Column diameter (mm), open tubular	0.1–0.7	0.02–0.1	
Average particle diameter (μm)	100–200	3–10	3–10
Column inlet pressure (atm)	<10	<600	<400
Efficiency (N): packed	10^3 to 10^4	10^4 to 8×10^4	5×10^3 to 5×10^4
open tubular	10^4–10^6	10^4–10^5	

tubular columns, and the difference in diffusion processes between the mobile and stationary phases dictates that the liquid stationary-phase film should be thin compared to the dimensions of the column radius. High efficiency in gas chromatography can be obtained with open tubular columns of 0.1–0.5 mm diameter, liquid films of 0.1–2.0 micrometer thickness, and column lengths of 5–100 m or more. Open tubular columns of these dimensions offer little flow resistance, and since gases have low viscosities, long columns are practically feasible and compatible with the low operating pressures favored in modern instrument design. Packed columns provide lower efficiencies, since long columns have undesirably high column inlet pressure requirements and the column packing structure results in peak broadening from flow anisotropy, which is unimportant for open tubular columns.

The mobile-phase diffusion coefficients are less favorable in supercritical fluids and liquids than they are for gases. To minimize the contribution of radial mass transfer to peak broadening, the dimensions of an open tube must be reduced proportionately, requiring columns with internal diameters less than 0.10 mm. Small column diameters present significant problems in column preparation and instrument design. These are difficult but possible to achieve for supercritical fluids, but cause exceptional problems in liquid chromatography. Packed columns are invariably used in the practice of supercritical fluid and liquid chromatography. To minimize flow anisotropy and stationary-phase mass-transfer contributions to peak broadening, the columns must have homogeneously packed beds of small particles with a narrow average particle diameter (3–10 μm); the total column length, and therefore the maximum plate number, is limited by the viscosity of the mobile phase and the available column operating pressure (see table). These same restrictions apply to supercritical fluids; however, since their characteristic physical properties are more favorable than those for liquids, supercritical fluids are capable of more efficient separations (larger value for N) than can be obtained in liquid chromatography. Thus, it is generally feasible to optimize resolution through varying N only in gas chromatography and in a more limited way in supercritical fluid chromatography. For liquid chromatography, the largest practical value for N is used for the separation, and the other parameters in the equation are varied to improve resolution. *See* SUPERCRITICAL FLUID.

Selectivity. The separation factor α in the equation is determined by the difference in the distribution coefficients for the two peaks, which is responsible for their relative positions in the chromatogram. Thus, it is the thermodynamic properties of the system that are responsible for the mechanism by which the two components are selectively recognized and therefore retained to different extents. Numerical analysis of the term selectivity reveals its critical importance to obtaining a desired separation. First of all, a separation can result only when α has a value greater than 1.0 ($\alpha = 1$ corresponds to the observed resolution $R_s = 0$), but only a relatively small increase in the value of α close to its minimum value considerably increases the possibility of obtaining a separation (Fig. 2*b*).

In gas chromatography, the mobile phase behaves essentially as an ideal gas and does not contribute to the selectivity of the system. To vary α, a new stationary phase (column) must be selected or the temperature of the system varied. In liquid chromatography, there is a high level of flexibility gained by being able to easily modify the properties of the mobile phase, as well as the possibility of exploiting retention mechanisms that are not available in gas chromatography. These factors account for its widespread use. The properties of the mobile phase are easily varied by changing either the identity or composition of the mobile phase from the wide range of solvents available, or by using additives in the mobile phase to change the properties of individual sample components (for example, buffers, complexing agents, and micelles), as well as by changing the type of stationary phase (column) used for the separation.

In practice, gas chromatography and liquid chromatography should be considered as complementary techniques. Separations by gas chromatography are restricted to thermally stable compounds with significant vapor pressure at the temperatures required for their separation. The upper temperature limit for common stationary phases is in the range 200–400°C (390–750°F). Few compounds with a molecular weight above 1000 have sufficient vapor pressure to be separated in this temperature range, and many low-molecular-weight compounds are known to be labile at temperatures required for their vaporization.

The only requirement for separation by liquid chromatography is that the sample has sufficient solubility in an available solvent whose properties are suitable for the separation. This criterion is much easier to meet in general and results in many of the applications for liquid chromatography as being those for which gas chromatography is unsuitable. Included are applications to high-molecular-weight synthetic polymers, biopolymers, ionic compounds, and many thermally labile polar compounds of chemical interest.

Supercritical fluids have solvating properties that are intermediate between those of gases and liquids. Liquids are virtually incompressible under the conditions used in liquid chromatography, so their capacity for intermolecular interactions is fixed. By contrast, fluids are highly compressible, and their density, which is proportional to their solvating capacity, can be varied over a wide range by simply changing the operating pressure or temperature of the column. This feature is unique to supercritical fluid chromatography, and it is the principal method of varying the separation factor α. The most common mobile phase is carbon dioxide, a relatively nonpolar fluid. More polar fluids, such as water, ammonia, or methanol, tend to have unfavorable critical constants or are highly corrosive to columns and instrument components, limiting their use. Mixed mobile phases, such as carbon dioxide–methanol mixtures, can be used to vary the selectivity but miscibility problems and high critical constants for the mixed mobile phases may restrict the range of properties available for experimentation. Supercritical fluid chromatography can provide faster and more efficient separations than liquid chromatography, but it is more restricted than liquid chromatography in the choice of mobile phases and retention mechanisms to vary the selectivity. However, it finds applications in many areas where the results from gas and liquid chromatography are unsatisfactory. Examples are the separations of low-molecular-weight organic polymers, fats and oils, and enantiomers. *See* CHEMICAL SEPARATION TECHNIQUES.

Separation time. The retention factor k in the equation represents the compromise between the need for retention to achieve resolution and the overall separation time. Numerical evaluation of this term indicates that a minimum value of k is required to obtain close-to-maximum resolution in the chromatographic system, but larger values of k do not increase this resolution significantly and lead to unfavorable separation times (Fig. 2*c*). The most difficult components to separate in a mixture will ideally have k values between 1 and 3, while k values greater than 10–20 have little further influence on resolution. Increasing the residence time of substances in a chromatographic system beyond a certain point does not contribute to their further separation. In practice, larger values than the optimum values may be required to provide sufficient space to fit all the components of a complex mixture into the chromatogram.

Planar chromatography. The separation mechanisms for thin-layer and column liquid chromatography are essentially the same. However, the two processes differ in that in column liquid chromatography all substances migrate the same distance and become separated in time, while in thin-layer chromatography all substances have the same separation time and are separated in space. In column liquid chromatography, the mobile phase moves through the column at a constant velocity, while in thin-layer chromatography its velocity is usually controlled by capillary forces and declines with distance from the solvent entry position. These forces are generally too weak to achieve the optimum velocity required to minimize zone broadening. Compared to a constant mobile-phase velocity, longer separation times and a reduced separation potential result. The useful development distance for a separation is restricted, since eventually the mobile-phase velocity becomes so low that zones begin to broaden faster than their centers are moving apart. The optimum development length depends on properties of the mobile and stationary phases, particularly the average particle size of the layer; and it is typically between 5 and 15 cm for commercially available materials.

A general discussion of zone broadening in thin-layer chromatography is inherently more complex than in column liquid chromatography for reasons which relate to both the practice of the technique and the limitations of subsequent theoretical treatments. Because substances are separated in space, individual zones experience only those theoretical plates through which they travel; and therefore the efficiency N is a function of the migration distance of the substance in the chromatogram. The migration of individual zones occurs across several regions of different local velocity and can be represented only by some average value that is difficult to determine. In practice, when the mobile-phase velocity is controlled by capillary forces, the major contribution to zone broadening is longitudinal diffusion supplemented by a smaller contribution from flow anisotropy. Separations in thin-layer chromatography occur with a plate number of 0–5000, depending on migration distance, and with a capacity for baseline resolved peaks of about 10–14 in the chromatogram.

The general variation of resolution with migration distance is more complex than for column liquid chromatography and is approximately represented by a special curve (**Fig. 3**). The fundamental parameter used to characterize the position of a sample zone in a thin-layer chromatogram is the retardation factor, or R_f value. It represents the ratio of the distance migrated by the sample to the distance traveled by the solvent front in the same time. The boundary conditions for R_f values are $1 \geq R_f \geq 0$ (Fig. 3). There is no separation possible for substances that do not migrate from their origin ($R_f = 0$) or do not interact with the stationary phase and move with the solvent front ($R_f = 1$). A maximum in the resolution of two closely spaced peaks is observed at an R_f of about 0.3, but the resolution does not vary much in the range $R_f = 0.2$–0.5, which is the optimum region in which to obtain a separation. When the difference in R_f values for two substances exceeds 0.1,

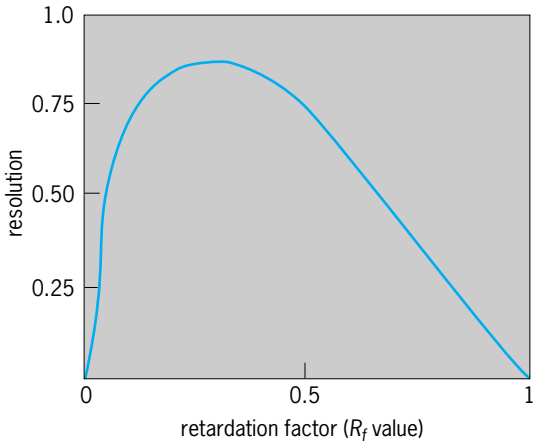

Fig. 3. Variation of resolution of two closely migrating zones as a function of the R_f value of the faster moving zone. (*After C. F. Poole and S. K. Poole, Chromatography Today, Elsevier, 1991*)

separations are easy to achieve; for R_f values less than 0.05, a separation is very difficult or impossible to achieve because of an inadequate range of efficiency values. Therefore, selective separation systems are required.

A common method of improving the separation potential of thin-layer chromatography is to increase the number of development steps used for the separation. This process is known as multiple development. A reduction in zone dimensions is achieved by a refocusing mechanism that is well understood from a phenomenological perspective but is difficult to characterize quantitatively. At the start of each development, the solvent front contacts the bottom portion of the zone first, and this region of the zone starts moving forward before the more distant regions of the zone still ahead of the solvent front. The result is a compression of the sample zone such that its width is reduced in size compared to zones migrated the same distance in a single development. This process can be automated so that a large number of developments can be used in an individual separation, resulting in a peak capacity for baseline resolved peaks of about 25–40, a considerable increase over the separation capacity obtained by using a single development.

The ultimate separation capacity of column liquid chromatography is greater than that of thin-layer chromatography. However, for simple separations where a high sample throughput is required, where crude samples are to be analyzed with minimum sample preparation, and where chemical reactions are to be used for detection, thin-layer chromatography is often preferred to column liquid chromatography. The reasons are that the stationary phase is not reused and multiple samples can be separated in parallel lanes in contrast to the single-sample sequential separations characteristic of column chromatography. After the separation, the components are immobilized on the stationary phase and can be treated with chemical reagents to change their detection properties without time constraints. The types of samples that are suitable for analysis by thin-layer chromatography are the same as those normally separated by column liquid chromatography, and a choice between the two approaches is made on the basis of practical considerations such as those outlined above. *See* ANALYTICAL CHEMISTRY. Colin F. Poole

Bibliography. K. Anton and C. Berger, *Supercritical Fluid Chromatography with Packed Columns: Techniques and Applications*, 1998; G. W. Gehrke, R. L. Wixom, and E. Bayer, *Chromatography: A Century of Discovery 1900-2000*, The Bridge to the Sciences/Technolgy, 2001; J. M. Miller, *Chromatography: Concepts and Contrasts*, 2005; S. Nyiredy, *Planar Chromatography: A Retrospective View for the Third Millennium*, 2001; C. F. Poole, *The Essence of Chromatography*, 2003.

Chromatophore

A pigmented structure found in many animals, generally in the integument. The term is usually restricted to those structures that bring about changes in color or brightness. A majority of chromatophores are single cells that are highly branched and contain pigment granules that can disperse or aggregate within the cell. However, in coleoid cephalopod mollusks (all mollusks except *Nautilus*), the chromatophores function as miniature organs, and changes in the dispersion of pigment are brought about by muscles. Although the mode of action of the two types of chromatophore is completely different, the effect is the same: pigment either is spread out over a large area of the body or is retracted into a small area.

The movement of pigment takes place in many chromatophores simultaneously, so that the effect is a change in the quality of light reflected from the surface of the animal. The color change functions as a camouflage from predator or prey, but it may also serve for regulating temperature, protecting against harmful radiation, and in some animals, signaling. Light stimulates the responses of chromatophores, and although in some species light may affect the chromatophores directly, light generally acts indirectly via the eyes and central nervous system.

Occurrence. Single-cell chromatophores are found in some annelids, in insects (migratory locusts, some grasshoppers, and stick insects), and in echinoderms (some sea urchins). They are much more conspicuous in crustaceans (shrimps and prawns), in fishes (especially in bony fish and teleosts), in anuran amphibians (frogs and toads), and in a few reptiles.

The chromatophores may be uniformly distributed in the skin (chameleons), or they may occur in patches (flounders) or lines (around the abdomen in shrimps). Chromatophores of various colors may be distributed unevenly across the body, and occur at different depths in the skin. This characteristic can be important for enhancing or suppressing colors produced in the skin by reflecting structures.

Colors produced. Chromatophores produce their colors by reflection after absorption of light. Generally, the light comes from above, but it may come

from below after reflection from an underlying structure. The most common type of chromatophore contains melanin (and is, therefore, often called a melanophore), which absorbs all wavelengths so that the chromatophore appears black; other types have red (erythrophores) or yellow (xanthophores) pigments. These pigments generally derive from carotenoids in vertebrates and from ommochromes in invertebrates. It should be noted that although a chromatophore may appear red in the laboratory or in a shallow tidal pool, it will appear black at depths of 165 ft (50 m) or more, and many of the red chromatophores of marine animals may act (like the black ones) as neutral-density filters for matching background brightness rather than for providing color.

Chromatophores do not usually contain green or blue pigments. These colors are often produced structurally, by reflecting or refracting structures (iridophores, also called iridocytes). The skin of many animals contains structures known as leucophores which, being broadband reflectors, can also produce greens and blues. Chromatophores in cephalopods and amphibians often function in conjunction with iridophores and leucophores to produce beautiful and complex color patterns over the whole body.

Crustaceans and vertebrates. Chromatophores contain pigment granules that move within them, giving them an appearance that ranges from spotted to fibrous on the five-stage scale that is widely used to measure the degree of chromatophore expansion. If the pigment within the particular cell is black or brown, the integument takes on a dark appearance when most of the chromatophores are in the last stage of dispersion (stage 5). If the pigment color is yellow or cream, the animal tends to look paler if all the chromatophores are at that stage.

Shrimp or frog chromatophores respond actively to the background illumination (**Fig. 1**). After only an hour or two, changes in the state of the chromatophores become evident with the aid of a microscope. These changes are usually brought about by hormones that are released from a particular site and circulated throughout the body. In shrimp, the hormones acting on the chromatophores originate in organs located in the eyestalk and also in the central nervous system. In the frog, the hormone responsible for color change comes from the pars intermedia of the pituitary organ; it is a peptide, called melanocyte-stimulating hormone, that disperses the pigment granules, thus making the skin darker. In teleosts, the control system is more complicated because there is a second hormone, melanin-concentrating hormone from the hypothalamus, that suppresses melanocyte-stimulating hormone, causing pigment aggregation in all color classes of chromatophores, and hence paling.

In crustaceans, elasmobranch fishes, anurans, and lizards, control of the chromatophores is thought to be exclusively hormonal. Such hormonal control is true also of some teleosts; in others the control is part hormonal and part neural; while in still others control is purely neural, as in the chameleon. Where nerves are involved, the speed of the response is faster, the chromatophores responding in minutes rather than hours. *See* NEUROSECRETION.

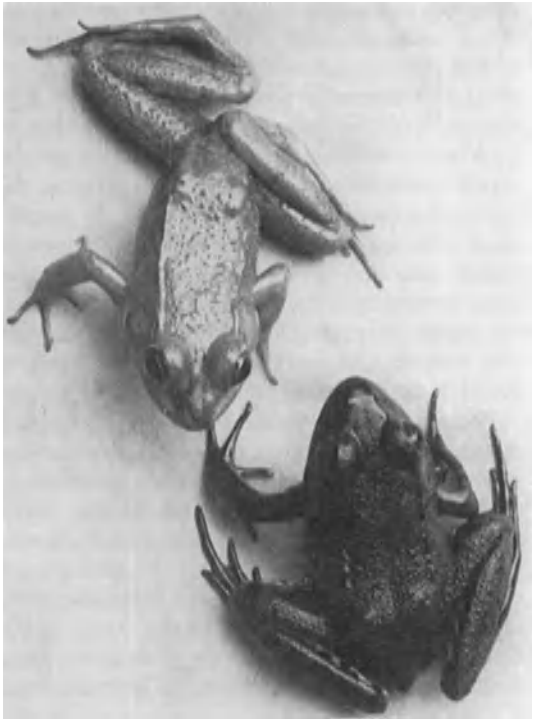

Fig. 1. Light- and dark-adapted frogs. (*After J. Pierce, J. Exp. Zool., 89:293–295, 1942*)

The chromatophores of crustaceans often have pigments of different colors, and they enable the integument to take on the background hue for concealment. When they form patterns, such as lines and bands, they may also serve to conceal the animal disruptively by breaking up the outline of the body. In fish, anurans, and reptiles, camouflage is again a major function, but in terrestrial vertebrates with chromatophores, changes in color may also serve for temperature regulation, at least when the temperature is excessively high or low. In some teleosts, neurally controlled chromatophores permit the rapid color changes that are used for social signaling, for example during sexual or territorial encounters.

Cephalopods. Each cephalopod chromatophore organ comprises an elastic sac containing pigment granules. Attached to the sac is a set of 15–25 radial muscles that are striated and contract rapidly. Associated with the radial muscles are axons from nerve cell bodies that lie within the brain. Active nerve cells cause the radial muscles to contract and the chromatophore sac expands; when the nerves are inactive, energy stored in the elastic sac causes the chromatophore to retract as the muscles relax. The chromatophores receive only nerve impulses, and there is no evidence that they are influenced by hormones.

The chromatophores are ultimately controlled by the optic lobe of the brain under the influence of the eyes. Two consequences follow from the fact that cephalopod chromatophores are under the direct control of the brain. First, color change is

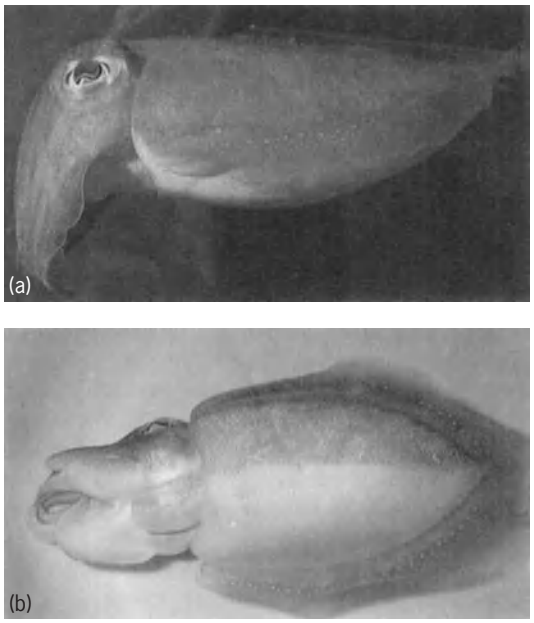

Fig. 2. Orientation of a cuttlefish. (*a*) The animal has a pale underside because all the chromatophores on the ventral mantle are retracted. (*b*) When the cuttlefish is rolled over onto its right side, the chromatophores of the upper (left) half of the ventral mantle expand to maintain the countershading.

instantaneous (an octopus can change from almost white to jet black in less than a second). Such a rapid change may cause a predator to hesitate, permitting escape; also, when the animal reaches the sea floor, it can vanish almost at once. Second, patterns can be generated in the skin in a way impossible in other animals. Thus, cephalopods can use the chromatophores not just to match the background in general color but to break up the body visually (disruptive coloration) so that the predator does not see the whole animal.

Concealment is obviously a major function. It is possible that cephalopod chromatophores evolved for purposes of camouflage after the ancestral cephalopods had abandoned the molluscan shell in the interest of speed and agility. However, because the chromatophores are neurally controlled and patterns can be produced in the skin, they can also be used for signaling. A characteristic of cephalopods is that some species make displays as conspicuous as those of fish or birds. Included are sexual signals destined for members of the same species, and warning signals, such as the large black eyespots used against predators. Signaling is especially well developed in some shallow-water forms, such as the Caribbean reef squid.

Finally, cephalopod chromatophores can be used for countershading and for maintaining countershading after disorientation. Animals that are generally illuminated from above tend to have a dark upper surface and a light lower surface, concealing them from a viewer from the side. This characteristic is true of fish and cephalopods. However, when a fish becomes momentarily disorientated and rolls over on its side, it becomes conspicuous because its countershading is fixed. When a squid rolls over, the chromatophores on the ventral surface of the body expand to maintain the countershading (**Fig. 2**). Such versatility is possible only because the chromatophores are neurally controlled. *See* PIGMENTATION; PROTECTIVE COLORATION. J. B. Messenger

Chromium

A chemical element, Cr, atomic number 24, and atomic weight 51.996, which is the weighted average for several isotopes weighing 50 (4.31%), 52 (83.76%), 53 (9.55%), and 54 (2.38%). The orbital arrangement of the electrons is $1s^2$, $2s^2$, $2p^6$, $3s^2$, $3p^6$, $3d^5$, $4s^1$. The stability of the half-filled *d* shell doubtless accounts for this rather unusual arrangement. In the crust of the Earth, chromium is the twenty-first element in abundance, which ranks it along with vanadium, zinc, nickel, and copper. Traces of chromium are present in the human body; in fact, it is essential to life. *See* PERIODIC TABLE.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

The element was discovered in 1797 and isolated the following year by the French chemist L. N. Vauquelin. It was named chromium because of the many colors of its compounds. It occurs in nature largely as the mineral chromite ($FeO \cdot Cr_2O_3$), which is a spinel, but the ore is usually contaminated with Al^{3+}, Fe^{3+}, Mn^{2+}, and Mg^{2+}. Smaller quantities are found as the yellow mineral crocoite ($PbCrO_4$).

As a transition metal, chromium exists in all oxidation states from 2− to 6+. The chemistry of its aqueous solutions, at least in the 3+ (chromic) state, is complicated by the fact that the compounds exist in many isomeric forms, which have quite different chemical properties.

Pure chromium metal has a bluish-white color, reflects light well, and takes a high polish. When pure, it is ductile, but even small amounts of impurities render it brittle. The metal melts at about 1900°C (3452°F) and boils at 2642°C (4788°F). Chromium shows a wide range of oxidation states; the compounds in which the metal is in a low oxidation state are powerful reducing agents, whereas those in which it shows a high oxidation state are strong oxidizing agents.

The bright color and resistance to corrosion make chromium highly desirable for plating plumbing fixtures, automobile radiators and bumpers, and other decorative pieces. Unfortunately, chrome plating is difficult and expensive. It must be done by electrolytic reduction of dichromate in sulfuric acid solution. This requires the addition of six electrons per chromium ion. This reduction does not take place in one step, but through a series of steps, most of which are not clearly understood. The current efficiency is low (maybe 12%), and the chromium plate contains microscopic cracks and other flaws, and so it does not adequately protect the metal under it from corrosion. It is customary, therefore, to first plate the object with copper, then with nickel, and finally, with chromium.

In alloys with iron, nickel, and other metals, chromium has many desirable properties. Chrome steel is hard and strong and resists corrosion to a marked degree. Stainless steel contains roughly 18% chromium and 8% nickel. Some chrome steels can be hardened by heat treatment and find use in cutlery; still others are used in jet engines. Nichrome and chromel consist largely of nickel and chromium; they have low electrical conductivity and resist corrosion, even at red heat, so they are used for heating coils in space heaters, toasters, and similar devices. Other important alloys are Hastelloy C (Cr, Mo, W, Fe, Ni), used in chemical equipment which is in contact with HCl, oxidizing acids, and hypochlorite. Stellite [Co, Cr, Ni, C, W (or Mo)], noted for its hardness and abrasion resistance at high temperatures, is used for lathes and engine valves, and Inconel (Cr, Fe, Ni) is used in heat treating and in corrosion-resistant equipment in the chemical industry. *See* ALLOY; HEAT TREATMENT (METALLURGY); STAINLESS STEEL.

Several chromium compounds are used as paint pigments—chrome oxide green (Cr_2O_3), chrome yellow ($PbCrO_4$), chrome orange ($PbCrO_4 \cdot PbO$), molybdate orange (a solution of $PbSO_4$, $PbCrO_4$, and $PbMoO_4$), chrome green (a mixture of $PbCrO_4$ and Prussian blue), and zinc yellow (potassium zinc chromate). Several of these, particularly zinc yellow, are used to inhibit corrosion. The gems ruby, emerald, and alexandrite owe their colors to traces of chromium compounds. *See* CORROSION; EMERALD; PAINT; RUBY.

Dichromates are widely used as oxidizing agents, as rust inhibitors on steel, and as wood preservatives. In the last application, they kill fungi, termites, and boring insects. The wood can still be painted and glued, and retains its strength. Other chromium compounds find use as catalysts, as drilling muds, and in photochemical reactions. The last are important in the printing industry. A metal plate is coated with a colloidal material (for example, glue, shellac, or casein) containing a dichromate. On exposure to strong light under a negative image, the dichromate is reduced to Cr^{3+}, which reacts with the colloid, hardening it and making it resistant to removal by washing. The unexposed material is washed off, and the metal plate is etched with acid to give a printing plate. *See* PRINTING.

Chromium is essential to life. A deficiency (in rats and monkeys) has been shown to impair glucose tolerance, decrease glycogen reserve, and inhibit the utilization of amino acids. It has also been found that inclusion of chromium in the diet of humans sometimes, but not always, improves glucose tolerance. Certain chromium(III) compounds enhance the action of insulin.

On the other hand, chromates and dichromates are severe irritants to the skin and mucous membranes, so workers who handle large amounts of these materials must be protected against dusts and mists. Continued breathing of the dusts finally leads to ulceration and perforation of the nasal septum. Contact of cuts or abrasions with chromate may lead to serious ulceration. Even on normal skin, dermatitis frequently results. Cases of lung cancer have been observed in plants where chromates are manufactured.

John C. Bailar, Jr.

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; R. B. Jordan, *Reaction Mechanisms of Inorganic and Organometallic Systems*, 2d ed., Oxford University Press, 1998; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004.

Chromophycota

A division of the plant kingdom (also known as Chromophyta) comprising nine classes of algae: Bacillariophyceae (diatoms), Chrysophyceae (golden or golden-brown algae), Cryptophyceae, Dinophyceae (dinoflagellates), Eustigmatophyceae, Phaeophyceae (brown algae), Prymnesiophyceae, Raphidophyceae, and Xanthophyceae (yellow-green algae). Some of these classes are closely related, while others stand so far apart that they are sometimes assigned to their own divisions. The chief unifying character is the presence of chlorophyll *c* rather than chlorophyll *b* as a complement to chlorophyll *a* (although only chlorophyll *a* is present in Eustigmatophyceae). The chloroplasts are usually brown or yellowish because of large amounts of β-carotene and various xanthophylls, many of which are restricted to one or more classes. Storage products include lipids, starch, and glucans (glucose polymers with β-1,3 linkages). In all classes except Dinophyceae, the chloroplast is surrounded by chloroplast endoplasmic reticulum that is continuous with the nuclear envelope, and the deoxyribonucleic acid of the chloroplast forms an annular genophore. Except in Cryptophyceae, thylakoids are in groups of three. In most classes, motile cells bear two unequal flagella, one of which may be almost completely reduced and at least one of which bears two rows of hairlike appendages. Algae in this division range in size and complexity from unicellular flagellates to gigantic kelps. *See* ALGAE; BACILLARIOPHYCEAE; CHRYSOPHYCEAE; CRYPTOPHYCEAE; DINOPHYCEAE; EUSTIGMATOPHYCEAE; PHAEOPHYCEAE; PRYMNESIOPHYCEAE; RAPHIDOPHYCEAE; XANTHOPHYCEAE.

Paul C. Silva; Richard L. Moe

Chromosome

A DNA–protein complex in the nucleus of eukaryotic cells. Chromosomes were so named because after staining, nineteenth-century light microscopists saw chromosomes as colored bodies in cells. The combination of DNA and proteins in chromosomes is called chromatin. Collectively the DNA from all the chromosomes in a nucleus is the hereditary blueprint for the species. Eukaryotes (plants, animals, fungi, and protists) usually have one nucleus in each cell. A nucleus is a double membrane-bound compartment in which chromosomes are located in interphase between cell divisions. In contrast, prokaryotes, such as bacteria and their allies (including mitochondria and chloroplasts), do not have true chromosomes or nuclei because prokaryotes do not confine their small, single circle of DNA in a membrane-bound compartment. *See* DEOXYRIBONUCLEIC ACID (DNA).

Interphase chromosomes. Chromosomes are usually described when they are easily observed in a condensed state during mitosis (M phase of the cell cycle). However, chromosomes are not visible during interphase, which comprises the largest part of the cell cycle. Interphase consists of G_1 (the first gap in interphase without DNA synthesis from the end of telophase to the beginning of DNA synthesis), followed by S (the synthesis phase, during which the DNA is replicated so each chromosome has two identical chromatids), and finally G_2 (the second gap phase, beginning after S and lasting until M phase begins). Nuclei in cells that stop dividing usually arrest in G_0, which is a G_1 interphase-like state that may persist indefinitely.

During interphase, chromosomes are decondensed to form an apparently homogenous mass of chromatin that consists of long DNA double helices associated with proteins. Chromosomes are still present as individual units in interphase nuclei, but their boundaries are difficult to distinguish, something like looking at a clear bottle full of cotton balls and trying to tell where one cotton ball ends and another begins. *See* CELL CYCLE; MITOSIS.

Structure and composition of mitotic chromosomes. Each chromatid of a chromosome consists of one long, linear DNA double helix that is associated with proteins to form chromatin. A DNA double helix is composed of two parallel polymeric chains of nucleotides that are twisted around each other. Each strand consists of alternating phosphate and sugar molecules with bases extending from each sugar. There are four bases: adenine (A), which forms hydrogen bonds with thymine (T) in the opposite DNA strand, and guanine (G), which forms hydrogen bonds with cytosine (C) in the opposite DNA strand. As a result, the two single DNA strands of polynucleotides in a double helix are complementary. This relationship is used when DNA is replicated during interphase of the cell cycle and when RNA is transcribed using DNA as a template. The DNA double helix of one chromatid is usually millions of base pairs (bp) long.

When chromosomes condense during mitosis, DNA becomes so highly folded that each double helix is approximately 10,000 times as long as the chromatid in which it is packed. (The necessity of this level of folding is apparent when one considers that diploid human cells contain about 2 m of DNA divided between 46 chromosomes in a nucleus that is about 5 μm in diameter.) In a current model for the structure of an average chromosome, the hierarchy of folding begins with formation of nucleosomes. A nucleosome consists of a core of eight basic proteins called histones (two molecules each of histones H2a, H2b, H3, and H4) and DNA. Histones are positively charged so they are attracted to negatively charged DNA. To form a nucleosome, 146 base pairs (bp) of DNA wraps 1.65 times around the core histone particle. There are about 50 bp of coiled DNA between adjacent nucleosomes. A chain of nucleosomes linked by a continuous DNA double helix forms a fiber with a diameter of about 10 nm and a 6:1 packing ratio of DNA length to 10-nm fiber length. Next, the 10-nm fiber associates with other proteins to coil or fold into a fiber with a diameter of 30 nm and a 42:1 ratio of DNA length to 30-nm fiber length. The 30-nm fiber forms ~1-μm-long loops that are attached to a proteinaceous chromosome core (scaffold). The loops condense against the core under the influence of a protein complex called condensin that is found in the core. The condensed loops and core constitute a strand called the chromonema with a diameter of ~0.2 μm. The chromonema coils, possibly due to contraction of the proteinaceous core, to form a chromatid about 1 μm in diameter with a ratio of DNA length to chromatid length of ~10,000:1. Coiling does not occur at the primary constriction (=centromere), resulting in a narrow part of the chromatid (**Fig. 1**). *See* NUCLEOSOME.

Without resorting to special treatments, chromosomes are visible only when they become condensed into sausage-shaped bodies during division of the nucleus (mitosis and meiosis). Although they are always microscopic, mitotic chromosomes in different species and even within a species can vary enormously in size, for example, from 1 to 40 μm in length and 1 to 2 μm in width. A chromosome can consist of either two identical sister chromatids (during G, prophase, prometaphase, and metaphase of mitosis) or of one chromatid (during G_1, anaphase and telophase of mitosis). Having one or two chromatids per chromosome does not change ploidy (the number of complete chromosome sets in a nucleus).

At a gross morphological level, chromosomes often can be distinguished from one another on the basis of their relative lengths and the positions of their primary constriction. The segments of chromosomes to either side of the primary constriction (centromere) are called arms. If the arms are roughly of equal length, the chromosome is metacentric. If the centromere is near one end, the chromosome is acrocentric, and there is a short (p) arm and a long (q) arm. If the centromere is truly at the end (rare), the chromosome is telocentric.

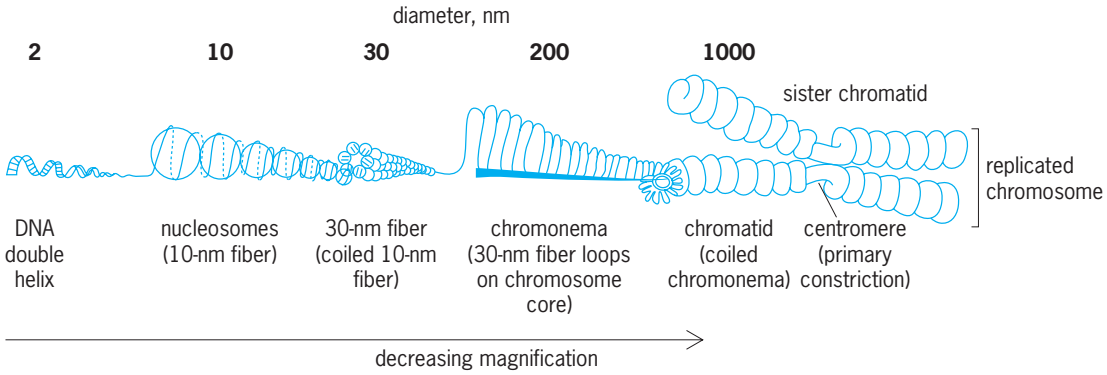

Fig. 1. Model of hierarchy of chromosome structure. Objects on the left are magnified more than those to the right. On the left and progressing to the right, a DNA double helix is associated with nucleosomes to form a 10-nm fiber that is coiled or folded to form a 30-nm fiber that forms loops on the chromosome core to form the chromonema that is coiled to form one chromatid. Note the lack of coiling in the primary constriction (centromere).

Chromosomes with primary constrictions in intermediate positions are described as submetacentric and subacrocentric (**Fig. 2**).

Kinetochores. A kinetochore is a nucleoprotein structure that forms in the centromere on each chromatid. Depending on the species, a kinetochore may appear as an amorphous sphere or a layered disk. The DNA involved in kinetochores often consists of a 150-bp tandem repeat sequence; the exact sequence varies from species to species. [In addition, there are often additional repeated sequences in the centromere and in heterochromatin (chromatin that is tightly coiled even in the nondividing nucleus and stains darkly in interphase). Proteins found in kinetochores are more evolutionarily conserved than the DNA sequences. During mitosis, microtubules grow from the spindle poles and attach to kinetochores. Force is applied through these kinetochore microtubules to move chromosomes to the metaphase plate during prometaphase and then to move sister chromatids (now sister chromosomes) to opposite poles during anaphase.

Telomeres. The ends of chromosomes are called telomeres. Although telomeres usually have no

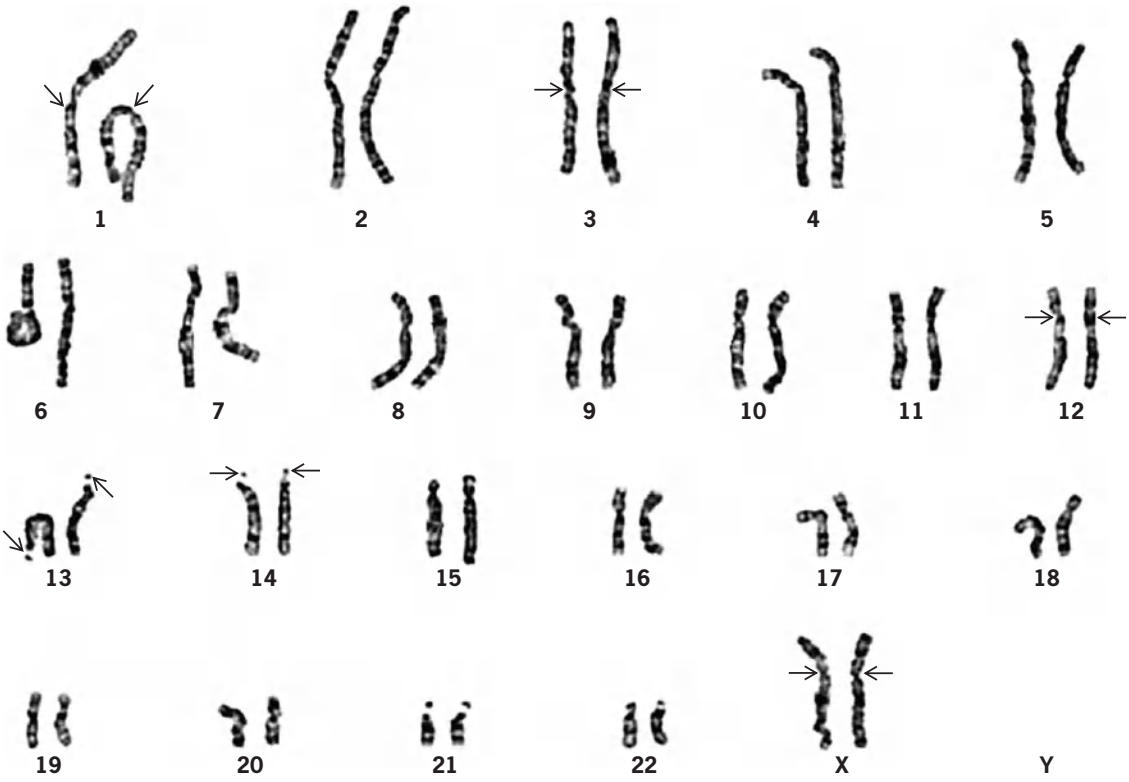

Fig. 2. Diploid set of normal human female chromosomes that have been G-banded for identification. G bands stain more darkly than the lighter R bands, and G bands have more AT-rich DNA than R bands. To construct an idiogram, chromosomes were cut out of a photograph and pasted in a row from longest to shortest with homologs together. Arrows indicate some centromeres. Note metacentric (such as chromosomes 1 and 3), acrocentric (such as chromosomes 13 and 14), submetacentric (such as the X chromosome), and subacrocentric (such as chromosome 12). Since these are chromosomes from a female, there are two X chromosomes and no Y chromosome. (*Courtesy of Genzyme Corporation*)

visibly distinct structure, they are necessary for the survival of chromosomes; without telomeres, chromosomes would be nibbled away slowly from their ends during every round of DNA replication, during which a little of the DNA double helix is lost from both ends of a chromatid. The telomere is the site where an enzyme called telomerase adds expendable DNA, usually in the form of a short repeat sequence such as TTAGGG. Most differentiated animal cells have no telomerase activity, but they inherit telomeres that are thousands of bp long from telomerase activity during meiosis that occurred just before gametes were formed. Telomere lengths are adequate for the normal number of divisions the cells undergo during development. This protects the body from cells that are dividing out of control, because cells that divide too frequently are eventually destroyed when their telomeres are lost. One cost of this protection is aging, and, as one might expect, cancer cells often have telomerase activity that permits them to divide indefinitely.

Euchromatin and heterochromatin. Chromatin is divided into euchromatin and heterochromatin based on its level of condensation during interphase and G_0. Euchromatin is decondensed and available for transcription. However, euchromatin still requires the proper proteins to mark genes (DNA segments) for transcription, that is, genes are not transcribed just because they are in decondensed chromatin, but genes must be in decondensed chromatin to be transcribed. In contrast, heterochromatin generally remains condensed, and its DNA is not used for transcription, if for no other reason than transcription factors and RNA polymerase are sterically hindered from access to the DNA. However, even heterochromatin becomes euchromatic when its DNA is replicated during S phase. There are two types of heterochromatin: constitutive heterochromatin that is always condensed (except during S phase) and facultative heterochromatin that is condensed only in certain cells or in only one sex. Facultative heterochromatin seems to be a way to control gene activity.

During mitosis, euchromatin and heterochromatin are similarly condensed and stained with most dyes. However, heterochromatin in condensed chromosomes can be stained differently from euchromatin after special procedures that typically utilize Giemsa stain. One technique, called C-banding, stains constitutive heterochromatin intensely so that a simple pattern of dark bands is visible on chromosomes. C-bands often occur around centromeres, nucleolus organizers, or at the ends of chromosomes, but they can be anywhere on a chromosome. C-banding has been used successfully on both plant and animal chromosomes. The other technique, called G-banding, stains facultative heterochromatin intensely and results in a relatively complex banding pattern. G-banding works well on mammalian and avian chromosomes and has been important for identifying individual chromosomes in these groups (Fig. 2).

Methyl groups can be added to cytosines in DNA, and this generally causes DNA to become facultatively heterochromatic and unusable for transcription. The pattern of methylation is duplicated during S phase, so methylation is inherited from one cell generation to the next. While spermatozoa and eggs carry genes for the same traits, some genes are differently methylated in spermatozoa compared with eggs. Thus, only the maternal or paternal versions of these genes are active in developing embryos. This is an epigenetic phenomenon (gene activity not determined by DNA sequence, but rather through chemical modification of DNA like methylation) called imprinting.

Chromosome sets, genomes, ploidy, and homologs. Generally, eukaryotes are capable of sexual reproduction that involves fusion of two gametes. Gametes usually carry a haploid set of chromosomes, defined as one copy of each chromosome that is characteristic of the species. An unreplicated haploid set of chromosomes, in which each chromosome consists of one chromatid, contains one genome of DNA called the 1C amount (C standing for "constant" for each species). For example, the human genome (1C amount of DNA) consists of about 3.5 picograms of DNA or 3.4 billion base pairs of DNA or 1 m of DNA. Other species have different genome sizes (1C). Normally each human gamete (spermatozoon or egg) carries a haploid chromosome set (often denoted $1n$) of 23 chromosomes. After fertilization the zygote is usually diploid ($2n$) with 46 chromosomes (Fig. 2 and **Colorplate 1**). If an organism has one or more complete sets of chromosomes with no extra or missing chromosomes, it is euploid. If an organism has more than two sets, it is polyploid. If chromosomes are added or missing from a set, the organism is aneuploid. *See* POLYPLOIDY.

A special class of chromosomes called B chromosomes (or supernumerary chromosomes) occurs sporadically in many organisms. B chromosomes generally have few, if any, active genes, and they are not necessary for the life of the organism. B chromosomes may be thought of as molecular parasites, and their presence or absence does not constitute aneuploidy. Aside from B chromosomes, the numbers of chromosomes in a set vary enormously from one species to another. Generally, haploid numbers above 15 are rare in plants, and haploid numbers above 50 are rare in animals. There are two ways in which chromosomes numbers can change: by breaking one chromosome into two smaller chromosomes or by fusing two chromosomes into one larger chromosome. DNA breaks are usually caused by radiation or chemicals. While breaks usually are repaired without a trace, sometimes repair leads to permanent breaks, fusion of chromosomes, trading segments between chromosomes (translocation), and inversion within a chromosome.

Most eukaryotic organisms are diploid with two sets of chromosomes, so each chromosome is represented twice (one derived from the egg and one from the spermatozoon). These pairs of chromosomes are called homologs or homologous chromosomes. Homologs are the same size and shape, have the same banding pattern, and carry genes for the same traits (but not necessarily the same alleles of those genes) (Fig. 2 and Colorplate 1).

Karyotypes and idiograms. Although the terms karyotype and idiogram have sometimes been used interchangeably, it less confusing to restrict karyotype to a word description of a set of chromosomes (for example, the haploid set of chromosomes is represented by three large metacentric, two large acrocentric, five small acrocentric, and two small metacentric chromosomes), whereas the term idiogram is used for a diagrammatic representation of a chromosome set, in which the chromosomes are arranged in a row from largest to smallest. In general, the longest chromosome is numbered 1, the next longest numbered 2, and so on. The chromosomes can be drawn, but usually they are cut out of a photograph of a chromosome spread and pasted in position. If the cell was diploid, homologs are positioned together in pairs (Fig. 2 and Colorplate 1).

Techniques for spreading and staining chromosomes. To make chromosome preparations suitable for observation by light microscopy, cells are usually fixed (killed and preserved) in a freshly prepared solution of one part acetic acid and three parts ethanol or methanol. The chromosomes are then spread and flattened onto a glass microscope slide, either by squashing cells under a cover glass or by bursting cells that have been hypotonically swollen cells and permitting the chromosomes to dry on the slide. Chromosomes can then be stained for examination by light microscopy.

Many stains have been used for chromosomes. Without special treatments, dyes such as crystal violet, Giemsa stain, orcein, and carmine stain chromosomes homogeneously throughout their length. Historically the Feulgen technique has been important because the intensity of staining is a direct indicator of the amount of DNA present. Some techniques result in the formation of "bands" on chromosomes due to differential staining based on the location of euchromatin and heterochromatin or differences in base composition of DNA along the length of chromosomes. The latter type of banding is observed with fluorescent dyes such as Hoechst 33258, daunomycin, chromomycin A, mithramycin, DAPI, and quinacrine dihydrochloride that either bind better or fluoresce more strongly in AT-rich or GC-rich DNA (Fig. 2).

Fluorescence in-situ hybridization (FISH). Important new tools for cytogenetics have been introduced during the past 30 years. One of these is fluorescence in-situ hybridization (FISH). In this technique, a specific DNA probe is prepared by attaching a small molecule, such as biotin, to a DNA sequence of interest (for example, a gene). A solution containing the DNA probe is placed on a glass microscope slide with chromosome spreads and heated to denature both the probe and the chromosomal DNA, that is, the hydrogen bonds between AT and GC base pairs are broken so that DNA double helices separate into single, complementary DNA strands. When the slide and solution are cooled to a temperature appropriate for renaturation (reformation of double helices by hydrogen bond formation between complementary strands), the probe will renature with any complementary DNA sequence in the chromosomes. Next, an antibody specific for biotin that is itself attached to a small fluorescent dye molecule (such as fluorescein isothiocyanate, FITC) is placed on the slide. This antibody selectively binds to the biotin in the probe and then the slide is examined under a microscope that has an ultraviolet (UV) light source and the appropriate filters for the dye. The dye will fluoresce to show the location of the probe on the chromosome(s) [**Colorplate 2**]. FISH has proven to be a powerful technique for be a powerful technique for locating DNA sequences on chromosomes.

Other forms of FISH include genomic painting, chromosome painting, and multicolor FISH. In genomic painting, the DNA of one species is isolated and labeled with biotin. The labeled DNA is used as a probe to identify chromosomes from that species among the chromosomes from one or more different species in a hybrid. In chromosome painting, one chromosome is isolated and its DNA is amplified by the polymerase chain reaction (a technique for amplifying specific sequences of DNA) involves copying and amplifying the complementary strands of a target DNA) and labeled with biotin; this labeled DNA is then used as a probe to locate a specific chromosome in a group of chromosomes. In multicolor FISH, probes can be prepared for each chromosome in a set, and every chromosome can be identified simultaneously using different dye combinations. *See* GENE AMPLIFICATION.

Immunofluorescence. Immunofluorescence can be used to locate specific proteins on chromosomes. The first step is the preparation of a primary antibody that will specifically bind to the protein of interest in chromosomes on a slide. Next, the primary antibody is bound by a secondary antibody that is labeled with a fluorescent dye that is visualized using fluorescence microscopy. Immunofluorescence is useful for relating proteins to the structure and function of chromosomes (Colorplate 1). *See* IMMUNOFLUORESCENCE.

Function. The nucleus is a molecular library of instructions for building an organism, responding to the environment, and reproducing. Chromosomes are like molecular books, and DNA is the molecular text. In a library, most of the books and the vast majority of the text are stored compactly when not in use, but they must be available when needed. Decondensation of euchromatin in interphase makes the DNA physically available for interaction with proteins called transcription factors that recognize specific base sequences and prepare the DNA for transcription by the enzyme RNA polymerase. During transcription, complementary single-stranded ribonucleic acid is made using one of the DNA strands as a template. Transcribed segments of DNA are called genes. Many RNAs are messages (mRNAs) that are transferred to the cytoplasm, where they associate with ribosomes. There, the base sequence of mRNA is used to determine the amino acid sequence of proteins in a process known as translation of the genetic code. During translation, the linear sequence of bases in mRNA is used to determine the linear

sequence of amino acids in a protein. *See* GENE; GENETIC CODE.

Other RNAs are transcribed that are not messages, but instead function as structural components, enzymes, suppressors of gene activity, or signals in the pathway to forming heterochromatin. Four of these other RNAs are components of ribosomes. Most ribosomal RNA is transcribed from nucleolus organizer regions (NORs), where hundreds or thousands of copies of the ribosomal RNA gene are located. Active NORs appear as secondary constrictions on certain chromosomes, that is, chromosome constrictions in addition to centromeres. *See* RIBONUCLEIC ACID (RNA); RIBOSOME.

During mitosis, when the duplicated DNA is being divided equally between the daughter cells, the DNA is so compactly packaged in condensed chromosomes that it is temporarily unavailable for transcription, like books that are crated for shipping to another library. At the end of mitosis (telophase), chromosomes decondense and return to their interphase condition, that is, the books are uncrated, and the library is open for business.

Special chromosomes. Special chromosomes include sex chromosomes, polytene chromosomes, and the chromosomes of meiosis.

Sex chromosomes. Some animals determine whether embryos develop as males or females by the combination of sex chromosomes they carry. Mammals and birds are prominent examples.

In mammals, males normally have only one large X chromosome and one Y chromosome that is usually small and heterochromatic. The Y carries the *SRY* gene, which determines that testes will be formed rather than ovaries. Females have two X chromosomes and no Y chromosome, so they form ovaries by default (Fig. 2 and Colorplate 1). However, because the number of X chromosomes differs in males and females, there is a problem with the dosage of genes on the X chromosome. Either the male with only one X has half enough genes or the female with two X's has twice too many genes. The problem is solved in mammals by converting one of the two X chromosomes in each cell of females into facultative heterochromatin in a process known as lyonization. This involves transcription of a gene on the mostly inactive X to produce RNA molecules that coat and play a role in inactivating the X chromosome. As a result, females, like males, have only one active (euchromatic) X chromosome in each cell. X inactivation occurs early in embryonic development, appears to be random with respect to which X is inactivated, and is inherited in subsequent cell divisions. Thus, the body of female mammals is a patchwork (mosaic) of tissues in which genes from one or the other X is active. During interphase, the inactive X chromosome in the nuclei of female cells stains darkly and is called a Barr body. Male cells seldom show Barr bodies.

In birds, the system is reversed so females have one Z chromosome and one small W chromosome, while males have two Z chromosomes. Presumably the presence of the W chromosome somehow determines that ovaries rather than testes will form in females. Barr bodies are not visible in male birds, indicating that birds deal with dosage compensation differently than mammals. *See* SEX DETERMINATION.

Fig. 3. *Drosophila melanogaster* polytene chromosomes.

Polytene chromosomes. A few insects, plants, and protozoa have cells that undergo endoreduplication, a process involving round after round of DNA synthesis without nuclear division to produce polytene chromosomes. The best-known example of polytene chromosomes occurs in Dipteran (fly) salivary glands. A polytene chromosome often consists of a thousand or more identical DNA double helices lying parallel in a bundle. Polytene chromosomes show intricate patterns of banding, where in some cases a band corresponds to a specific gene. Often euchromatin is selectively replicated, so heterochromatin

Fig. 4. Electron micrograph of a tomato synaptonemal complex. The kinetochore looks like a fuzzy ball, and a single recombination module is indicated by an arrow.

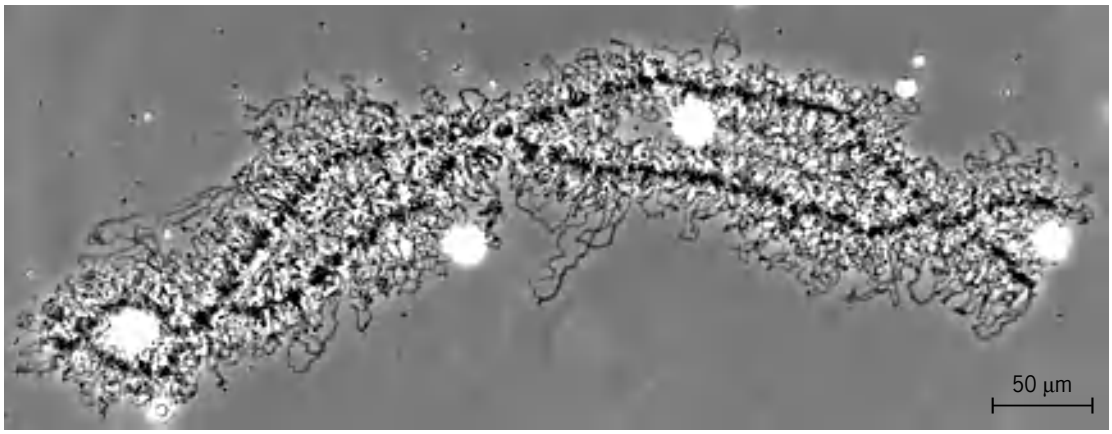

Fig. 5. Lampbrush chromosomes from a salamander (axolotyl) primary oocyte. Chromatin loops extend from the axes of two homologous chromosomes that come together at the chiasmata, where crossing over has occurred. Bright spots are nucleolar material. (*Courtesy of Garry Morgan, University of Nottingham*)

has little representation. Polytene chromosomes are not comparable to mitotic chromosomes; rather they are a special class of interphase chromosomes. Aside from protozoa that produce polytene chromosomes on the way to forming macronuclei, cells with polytene chromosomes are differentiated cells that will never divide again (**Fig. 3**).

Chromosomes of meiosis. Meiosis consists of two special cell divisions referred to as Meiosis I (MI) and Meiosis II (MII). MI is unique among nuclear divisions because crossing over occurs between homologous chromosomes and the chromosome number is reduced from diploid to haploid. For this, elongate homologous chromosomes synapse (come together side-by-side) in a gene-for-gene alignment during prophase I. Synapsed homologs are held together by a proteinaceous structure called the synaptonemal complex (SC) [**Fig. 4**]. The synaptonemal complex looks like a railroad track, with the rails represented by two lateral elements. One lateral element is associated with the two sister chromatids of one chromosome, and the other lateral element is associated with the two sister chromatids of the homologous chromosome. Chromatids in homologs are called nonsister chromatids. Thin transverse filaments span the central region between the lateral elements. It is at this time that new gene combinations are created by crossing over. Pairs of homologous chromosomes held together by synaptonemal complexes or chiasmata are called bivalents. *See* MEIOSIS.

Crossing over involves breakage and repair of DNA in such a way that nonsister chromatids exchange equivalent segments. These events probably occur in recombination nodules, proteinaceous ellipsoids about 100 nm in diameter that reside between the lateral elements of the synaptonemal complex (Fig. 6). The pattern of crossing over (recombination) can be mapped on synaptonemal complexes by examining synaptonemal complexes and recombination nodules using electron microscopy or by locating recombination nodules using immunofluorescent antibodies targeted to a protein in recombination nodules called MLH1 (**Colorplate 3**). Recombination nodule maps show the pattern and frequency of crossing over on chromosomes and thereby link genetic linkage maps to chromosome structure. *See* CROSSING-OVER (GENETICS); RECOMBINATION (GENETICS).

Primary oocytes (immature eggs) of most animals enter a synthetic period right after crossing over is completed and the synaptonemal complex breaks down. During this time, many components that will be needed by the developing embryo are synthesized and stored in the oocyte. Apparently to enhance transcription at this time, the pairs of homologous chromosomes (held together at sites of crossing over called chiasmata) become extremely long to permit rapid RNA synthesis on loops of chromatin extending from the chromosome axes. These bivalents (pairs of homologous chromosomes) are called lampbrush chromosomes because nineteenth-century cytologists thought they looked like the brushes used to clean glass chimneys of oil lamps (**Fig. 5**). Stephen M. Stack; Lorinda K. Anderson

Bibliography. D. W. Cleveland et al., Centromeres and kinetochores: From epigenetics to mitotic checkpoint signaling, *Cell*, 112:407–421, 2003; K. E. Latham, X chromosome imprinting and inactivation in preimplantation mammalian embroys, *Trends Genet.*, 21:120–127, 2005; S. M. Stack and L. K. Anderson, A model for chromosome structure during the mitotic and meiotic cell cycles, *Chromosome Res.*, 9:175–198, 2001; A. T. Sumner, *Chromosomes Organization and Function*, 2003; M. T. Teixeira, Telomere length homeostasis is achieved via a switch between telomerase-extendible and -nonextendible states, *Cell*, 117:323–335, 2004; D. Zickler and N. Kleckner, Meiotic chromosomes: Integrating structure and function, *Annu. Rev. Genet.*, 33:603–754, 1999.

Chromosome aberration

Any numerical or structural change in the usual chromosome complement of a cell or organism.

Heteroploidy. Numerical changes (heteroploidy) are of two types, polyploidy and aneuploidy.

Polyploidy is a change in the number of chromosome sets. The polyploid condition has both advantages and disadvantages. For example, tetraploid ($4n$) plants can be fertile because each chromosome can pair in meiosis with a homologous chromosome and segregate properly to form euploid $2n$ gametes or germ cells. These plants are larger than diploid varieties and produce larger fruit making them valuable commercially. Triploid ($3n$) plants, however, are usually sterile because of the inability of every chromosome to find a homologous chromosome to pair with in meiosis. The result is improper segregation (nondisjunction) of chromosomes and aneuploidy in the progeny. Triploidy rarely occurs in animals. Triploid reptiles and frogs are seen only in populations that reproduce parthenogenetically from unfertilized eggs that have not undergone the meiotic divisions that could produce aneuploid ova. Triploidy occurs in about 1% of human pregnancies, but it is almost always an embryonic lethal condition. *See* MEIOSIS; MITOSIS; POLYPLOIDY.

Aneuploidy is a change in the number of chromosomes from the diploid ($2n$) number (usually found in the somatic cells of sexually reproducing organisms, including animals and plants) or the haploid (n) number (usually found in germ cells and the haplophase of some unicellular organisms.) Aneuploidy involves either an increase or a decrease in the number of chromosomes. It usually involves a single chromosome, and any chromosome in the complement can be involved. Aneuploidy is the result of aberrant segregation of one or more chromosomes during meiosis or mitosis. If malsegregation or nondisjunction occurs during meiosis, one daughter cell receives two copies of the chromosome and the other daughter cell receives none. Fertilization of such an aneuploid germ cell by a euploid gamete will produce a zygote that has either three copies (trisomy) or one copy (monosomy) of the chromosome. Malsegregation of a chromosome can also take place during a mitotic division in a somatic cell, producing trisomic or monosomic cells in an otherwise euploid individual. This outcome is important primarily in the origin and progression of some forms of cancer. In humans, nearly 20% of ova and 7% of sperm may be aneuploid, and at least 3% of recognized pregnancies are trisomic and 1% are XO, that is, having a single sex chromosome per cell. Also, nearly half of all miscarriages and 3% of stillbirths are the result of aneuploidy.

Autosomal trisomy. The most common trisomy in human liveborns is trisomy 21, or Down syndrome, which is a major cause of mental retardation and congenital heart disease. Individuals with trisomy 18 and trisomy 13 also occur but are much less common. Most autosomal trisomies are lethal to embryos, leading to spontaneous abortion. The incidence of trisomy for any autosome increases exponentially with maternal age. For trisomy 21, the risk is about 1 in 100 for women over 35 years of age, and much greater for women over age 40. In about 80% of cases, the extra chromosome 21 is inherited from the mother and in 20% from the father, the latter case showing no paternal age effect on incidence. *See* DOWN SYNDROME.

Sex chromosome aneuploidy. In humans, there are more types of aneuploidy involving the sex chromosomes than the autosomes. The most common is XO, occurring in about 1% of pregnancies, with the risk being slightly higher in very young mothers. Although 99% of XO fetuses die early in pregnancy, the other 1% (about 1 in 10,000 liveborn females) survive. The high frequency of deaths of most XO fetuses is somewhat puzzling because the somatic cells in XX females have only a single active X chromosome. It is believed that the cause of death may be the presence of one or more genes on the tiny shared (homologous) segment of X and Y that remain active on both the Y and the otherwise inactive X chromosome. Adults who are XO tend to be short, with some webbing of the neck. They rarely develop secondary sexual characteristics or have children because the germ cells essential for ovarian development are usually absent. These features are characteristic of Turner syndrome. *See* HUMAN GENETICS.

Trisomy for the human X chromosome, commonly called XXX, is not associated with embryonic death or congenital malformations. The reason is that in all mammals, including humans, only a single X chromosome is active in each somatic cell; additional X chromosomes are inactivated, so almost all of their genes fail to be expressed. These are adult individuals with as many as five X chromosomes, and while they are physically normal they do show some degree of mental retardation.

In contrast to autosomal trisomy, sex chromosome aneuploidy increases only slightly with maternal age, and the extra X chromosome comes from the mother in only about 60% of the cases. About 1 of every 1000 liveborn females is XXX, while 1 of every 1000 liveborn males is XXY and 1 is XYY. An additional X chromosome can be present in either egg or sperm; additional Y chromosomes can be present only in sperm. The XXX and XXY individuals display minimal phenotypic manifestations of their increased number of chromosomes, although XXY males are usually sterile and have an increased risk of mild mental retardation. Individuals who are XYY generally are indistinguishable from XY individuals, although they are questionably at a slightly increased risk of antisocial behavior. The presence of a Y chromosome leads to male sex differentiation no matter how many X chromosomes are present, because of the presence of a single, critical gene, called *SRY*, on the Y chromosome. A mutation of this gene has been found in some XY individuals who developed as females.

Structural abnormalities. Structural abnormalities (chromosome mutations) involve the gain, loss, or rearrangement of chromosome segments after the continuity of the deoxyribonucleic acid (DNA) strand in one or more chromosomes is disrupted (**Fig. 1**).

Deletion. A deletion involves the loss of a chromosome segment and the genes it carries. A terminal deletion involves the loss of a segment extending from the point of disruption (breakpoint) to the end

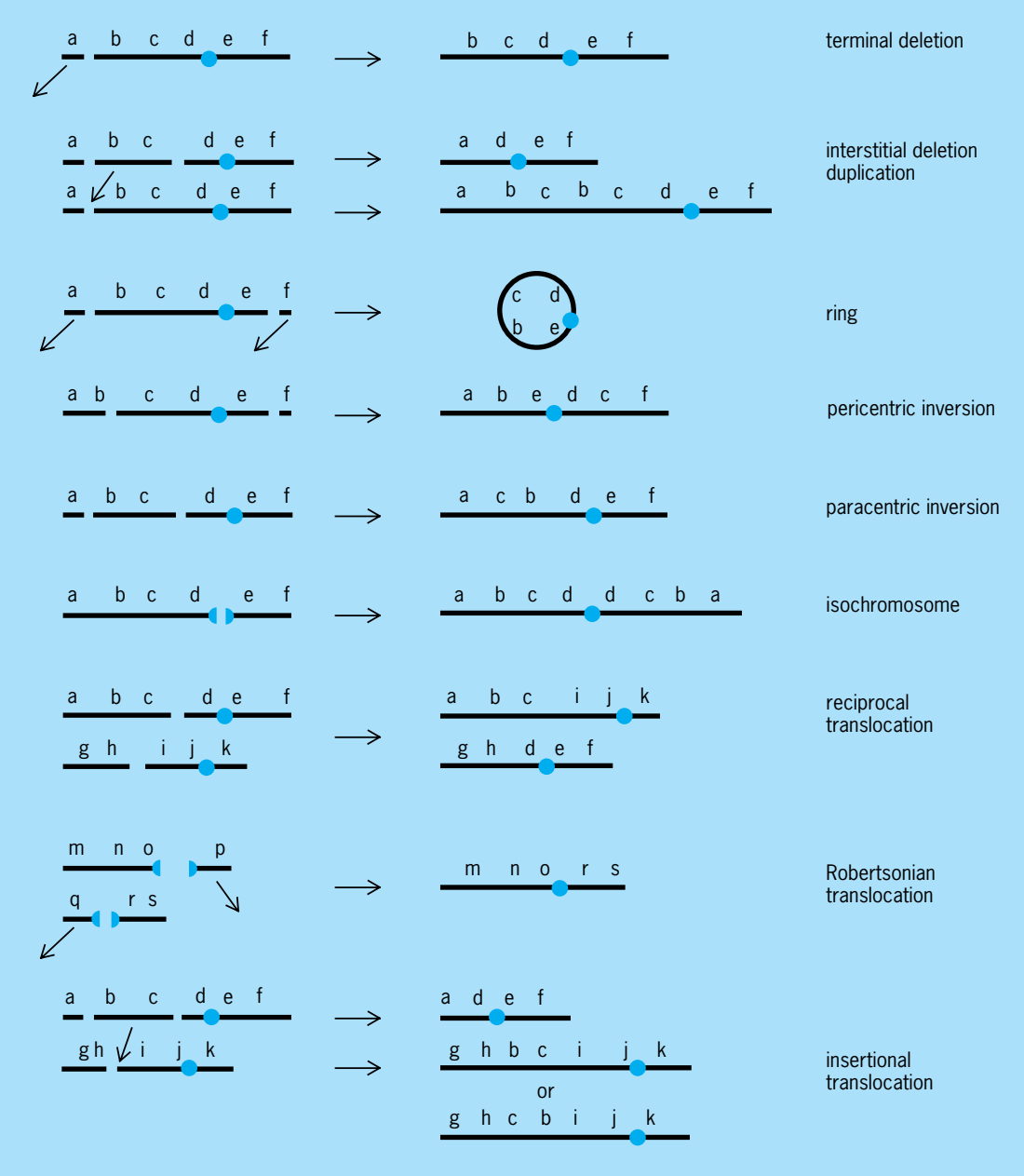

Fig. 1. Diagram of various types of structural abnormalities.

of the same arm of a chromosome. It is relatively uncommon because the broken end of the chromosome is usually repaired during DNA replication, resulting in a dicentric chromosome that is unstable because its two centromeres tend to be pulled toward opposite spindle poles at anaphase of cell division. The result is the breakage and eventual loss of the chromosome. An interstitial deletion involves the loss of the segment between two breakpoints in one arm of a chromosome (**Fig. 2**). The effect of such a loss depends on the genes that are included in the missing segment. If a deletion is associated with a genetic disease, the chromosome breakpoint provides a clue to the location of the gene responsible for the disease. For example, deletions were useful in mapping and isolating the human gene for retinoblastoma.

Fragments and duplication. When one break occurs in each arm of a chromosome, the broken ends of the internal centromeric fragment may join, resulting in the formation of a stable ring chromosome (Fig. 2). Each of the two end segments lacks a centromere, and such acentric fragments are lost during cell division, resulting in deletion of both end segments. Ring chromosomes are subject to reduction in size, as well as doubling. An individual who has a ring chromosome may thus show phenotypic effects not only of deletion but of duplication of part of the chromosome. A duplication more commonly occurs in other ways. For example, a chromosome segment can undergo tandem or inverted duplication at the usual chromosome site, or the second copy of the segment may be carried on another chromosome.

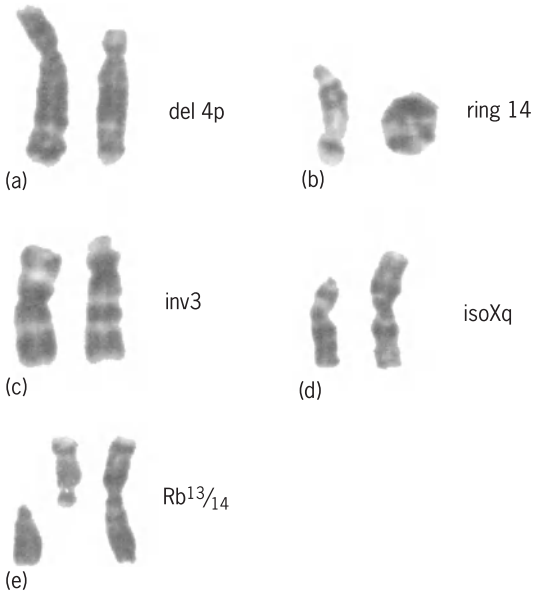

Fig. 2. G-banded human chromosomes illustrating structural chromosome abnormalities present in different individuals. The normal chromosome(s) is on the left in each case. (*a*) Terminal deletion of the short arm of chromosome 4 (4p). (*b*) Ring chromosome 14. (*c*) Pericentric inversion of chromosome 3. (*d*) Isochromosome of the long arm of the X. (*e*) Robertsonian translocation of chromosomes 13 and 14.

Inversion. An inversion is generated by disrupting the DNA strand in a chromosome at two breakpoints and rejoining the broken ends with the interstitial segment in the opposite orientation. This process will invert the order of the genes on the segment (Fig. 2). The centromere is included in the inverted segment in pericentric inversions but not in paracentric inversions. Thus, pericentric inversions can often be identified by an altered position of the centromere, while paracentric inversions can be identified only by an altered banding pattern or altered gene order.

Translocation. A translocation involves the interchange of one or more chromosome segments between two or more chromosomes. A reciprocal translocation involves the reciprocal exchange of the segment produced by a single break in one chromosome for the segment produced by a single break in another chromosome. A Robertsonian translocation (Fig. 2) is produced by the rejoining of the long arms and centromere of two acrocentric or telocentric chromosomes after disruptions in the short arm or at the centromere. It results in the reduction of chromosome number by one, except in the rare case in which the two short arms rejoin and the tiny fragment retains a functional centromere. An insertional translocation is produced when the interstitial segment between two breakpoints in one chromosome is inserted into a breakpoint in another chromosome. A complex translocation involves rearrangement of segments from three or more chromosomes. *See* GENE AMPLIFICATION; MUTATION.

Consequences of translocations and inversions. If a translocation breakpoint disrupts a gene, the gene's function will be blocked or abnormal, and such can have deleterious effects on development or function. Sometimes a normally silent gene is activated by a chromosome rearrangement that places it next to a strong promoter of gene expression, and this change is important as a cause of cancer. If a translocation does not block the function of an essential gene or activate a normally silent gene, the individual carrying the rearrangement will be normal.

In an individual carrying an inversion on one of a pair of homologous chromosomes, pairing leads to the formation of a loop containing the inverted segment. Crossing over within the loop leads either to the formation of gametes with a duplication of one segment of the inversion and a deficiency of the other segment, or to two abnormal chromosomes, one with two centromeres and the other lacking a centromere. *See* CROSSING-OVER (GENETICS).

Causes. Structural aberrations can occur spontaneously or be induced by agents that break chromosomes, such as x-rays, radioactive substances, ultraviolet rays, and certain chemicals. The most frequent cause may be the presence of enormous numbers of a few types of short interspersed elements (SINES), that is, DNA sequences that occur once every few thousand base pairs throughout the genome of most metazoans, including humans. These elements predispose to the occurrence of errors during DNA replication or genetic recombination at meiosis that can lead to the deletion or duplication of the region between two nearby interspersed repeats on one chromosome. They may also play a role in the formation of inversions and, possibly, translocations. *See* MUTAGENS AND CARCINOGENS.

Another cause of structural aberrations is also inherent in the genome: the great abundance of short repeats of a 2-, 3-, or 4-base-pair unit. Some trinucleotide repeats, such as $(CGG)_n$ or $(CAG)_n$, can undergo expansion during meiotic and mitotic cell divisions. This expansion sometimes affects gene function and leads to disease. The most common type of X-linked mental retardation in humans is the result of heritable expansions, in the *FMR-1* gene, of a specific trinucleotide repeat, $(CGG)_n$, where the number of expansions (n) is increased from the normal 8–20 or so to 50–200 or more. For unknown reasons, this expanded region tends to undergo breakage under some conditions, and this particular form of mental retardation is called the fragile X syndrome. There are dozens, if not hundreds, of similar fragile sites in the human and other genomes.

Imprinting. The function of a gene that is inherited from the father may differ from that of the same gene inherited from the mother, a phenomenon known as imprinting. The mechanism for imprinting is not fully understood, but differences in DNA methylation of genes in sperm compared to ova appear to be involved. Imprinting is important in some human diseases, such as Prader-Willi and Angelman syndromes. Although deletions of the same region on chromosome 15 have been observed in some individuals with either of these syndromes, individuals who received the deleted chromosome 15 from the father developed the Prader-Willi syndrome, whereas those

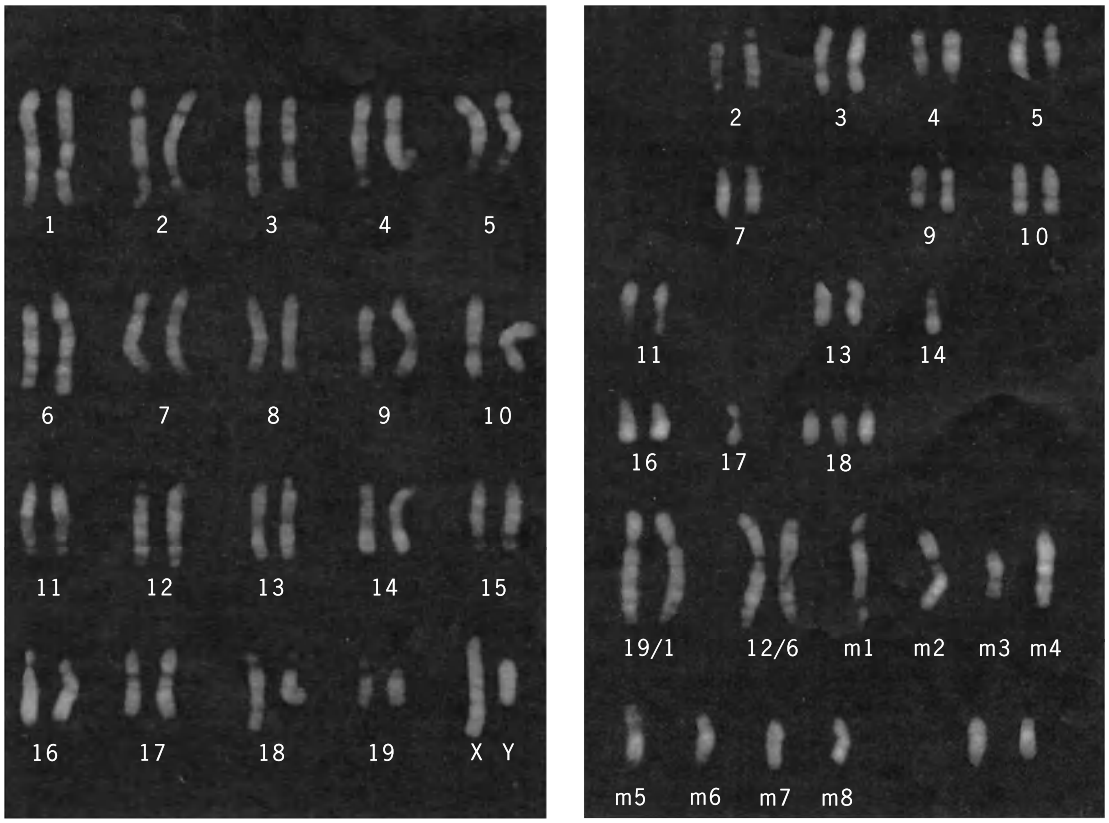

(a) (b)

Fig. 3. Q-banded chromosomes of (*a*) normal mouse metaphase cell, arranged according to the standard karyotype, (*b*) and metaphase karyotype of a mouse cancer cell showing numerical and structural changes. Chromosomes 2–18 are structurally normal, but noteworthy are the monosomy 17 and 14 and trisomy 18.

who received the deleted chromosome 15 from the mother developed Angelman syndrome. Thus, although each syndrome is associated with the presence of a single undeleted chromosome 15, the outcome depended on the parental origin.

Abnormal chromosomes in tumor cells. Specific chromosome aberrations play an important role in the development of many, if not all, cancers. For example, most patients with chronic myelogenous leukemia have a translocation between chromosome 9 and 22 which fuses the normally silent *c-abl* oncogene with the expressed *bcr* gene. The result is the production of an abundant *bcr-abl* fusion gene product and tumor development (**Fig. 3***a* and *b*). Tumor cells tend to be chromosomally unstable, and further chromosome aberrations may enhance tumor malignancy. Translocations and trisomies may lead to additional oncogene products, whereas deletions may remove genes that suppress tumor growth.

Some tumor cells have an elongated region on a chromosome arm that contains an increased number of copies (amplification) of an oncogene. Such regions stain uniformly, rather than being banded, and are called homogeneously staining regions (HSRs). The amplified oncogenes are sometimes extrachromosomal, that is, are present on very tiny collections of chromatin resembling small chromosomes that lack centromeres; these are called double minutes (DMs). *See* ONCOGENES; ONCOLOGY.

Evolution. Chromosome complements are remarkably diverse throughout both plant and animal kingdoms, and chromosome aberrations may have played an important role in the reproductive isolation that fosters speciation. For example, polyploidization creates a reproductive barrier that favors the formation of a new, polyploid species. Some species contain very large chromosome numbers, such as a fern with about 1000 chromosomes, a crayfish with over 350 chromosomes, and a butterfly with over 400 chromosomes.

Structural chromosome changes have also been a common occurrence in evolution. As described above, when a translocation or inversion involves one of a pair of homologous chromosomes, abnormal gametes are produced and fertility is reduced. However, in a small, inbreeding population, both male and female parents could receive a copy of the rearranged chromosome or chromosomes, and each parent could pass them on to an offspring, who would thus have a matched set of rearranged chromosomes. The latter would pair normally at meiosis, leading to the production of normal gametes. In this way, the altered chromosome could spread through a small, inbreeding population and might replace the initial chromosome arrangement, becoming fixed in the population. *See* POPULATION GENETICS.

Chromosome rearrangements do not change the banding patterns of the chromosome segments.

Therefore, it is possible to compare the karyotypes of species to determine what changes have been fixed in their evolution. For example, there is a preponderance of Robertsonian translocations in lemurs compared to other primates, and a preponderance of pericentric inversions in the chromosomes of the great apes and humans. The telocentric chromosomes of the European mouse, *Mus musculus domesticus*, undergo frequent Robertsonian translocations and, indeed, one mouse was found in which 18 of the 19 autosomes had undergone such translocations. At this time there is no evidence that chromosome changes alone have led to the creation of a new species, but it is clear that they can lead to sterility in certain outcrosses and to the reproductive isolation that is critical for speciation. *See* CHROMOSOME; RECOMBINATION (GENETICS); SPECIATION.

Orlando J. Miller; Dorothy A. Miller

Bibliography. C. J. Epstein, *The Consequences of Chromosome Imbalance*, 1986; G. J. Stine, *The New Human Genetics*, 1989; E. Thesman and M. Susman, *Human Chromosomes: Structure, Behavior, Effects*, 3d ed., 1993; B. K. Vig, *Chromosome Segregation and Aneuploidy*, 1993.

Chromosphere

A complex structure of warm gas above the visible surface, or photosphere, of the Sun and most stars. The term "chromosphere" was first applied to the red ring and large prominences seen at the edge of the eclipsed Sun with the unaided eye. Emission in the Balmer-alpha line of hydrogen at 653-nanometer wavelength accounts for the red color. The chromosphere is transparent in visible light, but is opaque and bright in the ultraviolet continuum and in strong lines of abundant elements, including hydrogen, helium (first observed during the 1868 solar eclipse), oxygen, calcium, and magnesium. Gas temperatures range from 3500 to 30,000 K (5800 to 54,000°F) and densities are between 10^9 and 10^{12} particles per cubic centimeter.

Structure of solar chromosphere. The solar chromosphere extends from roughly 500 km (300 mi) above the photosphere to where hydrogen becomes fully ionized and the gas temperature rises rapidly to more than 10^6 K (1.8×10^6 °F) in the corona. High-resolution images of the Sun taken in strong emission lines reveal a filamentary structure consisting of many magnetic loops. In the chromospheric network at the edges of 30,000-km-wide (19,000-mi) supergranule cells, the magnetic field is strong, the gas is dense and hot, and the chomosphere may be only 500 km (300 mi) thick. Here the magnetic heating rate is 10 or more times higher than average. Thin gas jets called spicules are clustered in the network. In the center of supergranule cells, where the magnetic field is weak, the chromospheric gas cools to 3500 K (5800°F) and extends upward several thousand kilometers. Here molecular carbon monoxide (CO) is abundant enough to keep this gas cool. *See* SUPERGRANULATION.

Chromospheres of other stars. Stars with effective temperatures less than about 8000 K (14,000°F; the Sun is 5770 K or 9930°F) show the same emission lines as the solar chromosphere and are thus thought to have analogous regions. Slowly rotating old stars show chromospheric emission at minimal levels comparable to those seen in the centers of solar supergranule cells. This weak emission indicates a basal level of chromospheric heating that is probably produced by the damping of nonmagnetic acoustic waves generated by convection in the lower atmosphere. Young, rapidly rotating stars show chromospheric emission at levels 10 or more times higher than is observed in regions of strong magnetic fields on the Sun. The highest levels of this emission correspond to saturated heating by magnetic-field reconnection events (often called microflaring) or the damping of magnetoacoustic waves.

Wilson-Bappu effect. In 1957 Olin C. Wilson and M. K. V. Bappu discovered that the widths of the near-ultraviolet emission lines of singly ionized calcium (Ca^+) increase with stellar luminosity. The Wilson-Bappu effect is now explained by the increased amount of gas in the chromospheres of lower-gravity stars, which have higher luminosities. Wilson began a long-term program to monitor the strength of the stellar Ca^+ emission, which led to the discovery of magnetic cycles on stars analogous to the 11-year solar cycle. Continued monitoring of this emission feature in many stars revealed that some are going through phases of no magnetic cycle, like the Maunder minimum of sunspots observed between 1645 and 1715. The energy budget and dynamics of stellar chromospheres have been studied with the *Hubble Space Telescope* and the *International Ultraviolet Explorer* satellite. *See* SATELLITE (ASTRONOMY); SUN.

Jeffrey L. Linsky

Bibliography. A. Bhatnagar and W. C. Livingston, *Fundamentals of Solor Astronomy*, World Scientific, 2005; P. Foukal, *Solar Astrophysics*, 2d ed., Wiley, 2004; L. Golub and J. M. Pasachoff, *Nearest Star: The Surprising Science of Our Sun*, Harvard University Press, 2001; K. R. Lang, *Sun, Earth, and Sky*, Springer-Verlag, 1995; K. J. H. Phillips, *Guide to the Sun*, Cambridge University Press, 1992; H. Zirin, *Astrophysics of the Sun*, Cambridge University Press, 1998; J. B. Zirker, *Journey from the Center of the Sun*, Princeton University Press, 2001, paper, 2004.

Chronic fatigue immune dysfunction syndrome

A condition resulting in massive, debilitating fatigue accompanied by diverse symptoms including memory loss, diminished powers of concentration, sleep disorder, headache, low-grade fever, muscle and joint pains, sore throat, tender lymph nodes, and a postexertional malaise lasting more than 24 hours. This condition is also known as postviral fatigue, chronic fatigue syndrome, and myalgic encephalomyelitis. It is often associated with stress.

The severity of chronic fatigue immune dysfunction syndrome (CFIDS) varies considerably. Some patients are bedridden while others suffer only mildly debilitating symptoms. The cause of the disease is unknown, although there are hypotheses relating it to an abnormal immune response due to concurrent stress and a previous infection, perhaps viral. The Epstein-Barr virus, human herpes virus 6, cytomegaloinclusion virus, Coxsackie virus, various stealth viruses, and a retrovirus have been implicated. All of these viruses are common, and the general population is frequently exposed to them. Nevertheless, the general population does not suffer from the syndrome. One possibility is that exposure to the virus occurs during a period of either physical or emotional stress. Genetic and environmental factors may also be implicated. An abnormal immune response to viral or bacterial activity could explain many symptoms of CFIDS. Among possible sources of immune "dysregulation" in CFIDS are decreased natural killer cell (NK) cytotoxity, increased T and B lymphocyte activation, and elevated cytokine production. *See* EPSTEIN-BARR VIRUS; HERPES; INFECTION.

CFIDS is not life-threatening and is not considered communicable. However, it often occurs within families or within a defined geographical area, which, at least in some cases, suggests the possible involvement of an infectious agent. If a viral cause is proven, there would be an infective phase during or immediately after the incubation of the virus.

There is no definitive diagnostic test for CFIDS; thus it is necessary to first eliminate all other potential symptom etiologies. A diagnosis of the syndrome should be made only after alternative medical and psychiatric causes of chronic fatigue have been excluded. Fatigue can occur in the simplest to the most complex diseases. Conditions that must be excluded include simple iron deficiency anemia, familial or genetically related anemias, hematomachrosis, chronic infections (such as tuberculosis and Lyme disease), inflammatory conditions (such as sarcoidosis and thyroid dysfunction), autoimmune diseases (such as lupus erythematosus, Hashimoto's thyroiditis, and hemolytic anemia), and cancer.

While the majority of individuals with confirmed CFIDS have positive antibodies to one or more of the implicated viruses, the viruses are so common that antibody titers may be suggestive but are not diagnostic. *See* ANTIBODY.

There have been suggestions that CFIDS is a psychiatric state manifesting itself as a primary chronic depressive neurosis. This means that the individual is not physically ill but imagines somatic illness and may be labeled as neurotic or lazy. Despite many similarities in the somatic presentation of CFIDS and depression, various psychological, immunologic, and neurological tests indicate clear differences between the two disorders.

There is no standard method of treatment. However, some believe that treatment based on reversal of the abnormal immune response is helpful. Cognitive behavioral therapy and certain approaches to exercise have also demonstrated efficacy as treatments. Some individuals affected by CFIDS improve with time; when untreated, most remain functionally impaired for several years.

Derek Enlander; Christopher Snell

Bibliography. K. Fukuda et al., The chronic fatigue syndrome: A comprehensive approach to its definition and study—International Chronic Fatigue Syndrome Study Group, *Ann. Intern. Med.*, 121:953–959, 1994; C. M. Jorge and P. J. Goodnick, Chronic fatigue syndrome and depression: Biological distinction and treatment, *Psychiat. Ann.*, 27(5):365–371, 1997; N. Kaushik et al., Gene expression in peripheral blood mononuclear cells from patients with chronic fatigue syndrome, *J. Clin. Pathol.*, 58:826–832, 2005; A. L. Komaroff, The biology of chronic fatigue syndrome, *Amer. J. Med.*, 10:169–171, 2000; K. J. Mayer, N. G. Klimas, and M. A. Fletcher, Immunolgy, in L. Jason, P. Fennell, and R. Taylor (eds.), *Handbook of Chronic Fatiue Syndrome*, pp. 124–151, Wiley, New York, 2003; P. Whiting et al., Interventions for the treatment and management of chronic fatigue syndrome, *JAMA*, 286:1360–1368, 2001.

Chronograph

Any device that measures a time interval. The term is usually restricted to measurements that have a small percentage error. Early time measurements used the human pulse rate or counted the number of swings of a pendulum for small time intervals. Astronomers used a regulated rotating drum to measure the time intervals of short changes in objects in the sky. Physicists did the same in the laboratory. The term chronograph seems to have stuck even when the device is digital rather than a graph. The term chronometer is often used interchangeably with chronograph, although again the result may be an electronically produced number rather than a meter reading. *See* CHRONOMETER.

Modern chronographs range from accurate wristwatch chronometers, used to time sporting events, to Doppler or other electronic or atomic means used in the laboratory or everyday living. *See* ATOMIC CLOCK; DOPPLER EFFECT; QUARTZ CLOCK; TIME-INTERVAL MEASUREMENT; WATCH. John L. Safko

Bibliography. L. E. Arushanyan et al., Utilization of the optical chronograph for deep investigation of natural water reservoirs, *Proc. SPIE*, 5149:214–222, 2003; G Lawson et al., In-bore chronograph—A laser radar for interior ballistics measurements, Part 1: System design, *Proc. SPIE*, 1936:44–50, 1993; S. J. Rodano and J. J. D'Amario, Chronograph for measuring projectile velocities, *Amer. J. Phys.*, 44:711–712, 1976; J. J. Sidtis, From chronograph to functional image: What's next, *Brain Cognit.*, 42(1):75–77, 2000.

Chronometer

A large, strongly built watch especially designed for precise timekeeping on ships at sea. The name is sometimes loosely applied to any fine watch. *See* WATCH.

The features that distinguish a chronometer from a watch are (1) a heavy balance wheel, the axis of which is kept always vertical by mounting the entire instrument within two concentric rings, so pivoted as to permit the chronometer to remain undisturbed despite considerable tilting of the box containing it, as the ship rolls and pitches; (2) a balance spring wound in cylindrical shape, instead of a nearly flat helix; (3) a special escapement; and (4) a fusee, by means of which the power of the mainspring is made to work through a lever arm of continuously changing length, being shortest when the spring is tightly wound and longest when it has run down, thus regulating the transmitted power so that it is approximately constant at all times. *See* ESCAPEMENT; SPRING (MACHINES).

These mechanical chronometers are being replaced by quartz digital chronometers, which operate on the same principle as a quartz watch. *See* QUARTZ CLOCK.

Oceangoing ships formerly relied completely on chronometers keeping Greenwich mean time to determine longitude. The broadcast of radio time signals that became widespread in the 1920s, and subsequently the loran, Omega, and satellite navigation systems, have made Greenwich mean time available to mariners at almost any instant, and chronometers are no longer indispensable for determining longitude at sea. However, they are still a necessary backup, and naval vessels will not leave port without one or more chronometers aboard. *See* CLOCK (MECHANICAL); HOROLOGY; LORAN; SATELLITE NAVIGATION SYSTEMS; TIME. Steven J. Dick

Bibliography. R. T. Gould, *The Marine Chronometer*, 1990; M. E. Whitney, *The Ship's Chronometer*, 1985.

Chrysoberyl

A mineral having composition $BeAl_2O_4$ and crystallizing in the orthorhombic system. Chrysoberyl crystals are usually tabular parallel to the front pinacoid and frequently in pseudohexagonal twins (see **illus.**). There is good prismatic cleavage. The hardness is 8.5 (Mohs scale) and the specific gravity is 3.7–3.8. The luster is vitreous and the color various shades of green, yellow, and brown. There are two gem varieties of chrysoberyl. Alexandrite, one of the most prized of gemstones, is an emerald green but in transmitted or in artificial light is red. Cat's eye, or cymophane, is a green chatoyant variety with an opalescent luster. When cut en cabochon, it is crossed by a narrow beam of light. This property results from minute tabular cavities that are arranged in parallel position.

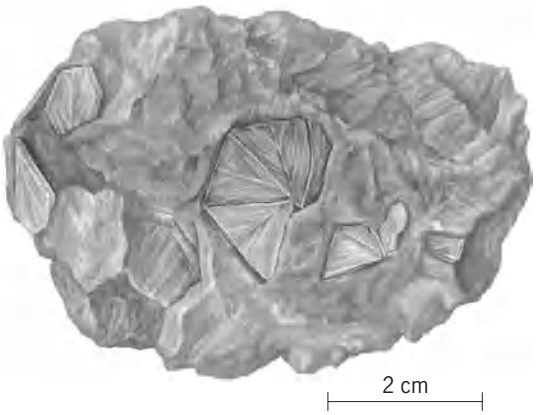

Chrysoberyl crystals in pegmatite, Greenwood, Maine. (*Specimen from Department of Geology, Bryn Mawr College*)

Chrysoberyl is a rare mineral found most commonly in pegmatite dikes and occasionally in granitic rocks and mica schists. Gem material is found in stream gravels in Ceylon and Brazil. The alexandrite variety is found in the Ural Mountains. In the United States chrysoberyl is found in pegmatites in Maine, Connecticut, and Colorado. *See* BERYLLIUM; GEM. Cornelius S. Hurlbut, Jr.

Chrysocolla

A silicate mineral with composition essentially $(Cu,Al)_2H_2Si_2O_5(OH)_4 \cdot nH_2O$; iron, and more rarely manganese, may substitute for copper. Although acicular monoclinic crystals have been reported, the wide variation in the composition of analyzed specimens suggests that chrysocolla is a hydrogel or gelatinous precipitate rather than a well-crystallized mineral. This suggestion is reinforced by the fact that x-ray diffractions are weak and ill defined. The mineral characteristically occurs in impure amorphous-to-cryptocrystalline crusts and masses with conchoidal fracture.

There is a wide variation in physical properties commensurate with the variation in chemical composition. The hardness varies from 2 to 4 on Mohs scale, the specific gravity varies from 2.0 to 4.0, and the refractive index is given variously from 1.40± to 1.60. The luster is vitreous to greasy; and the color is normally green to greenish blue but may be brown or, with manganese present, black.

Chrysocolla is a secondary mineral occurring in the oxidized zones of copper deposits, where it is associated with malachite, azurite, native copper, and cuprite. It is also found in alluvial gravels, indicating that copper is carried in solution far from its primary source. Chrysocolla is a minor ore of copper and a minor gem. Most gem material is essentially chalcedony colored by incorporated chrysocolla. *See* GEM; SILICATE MINERALS. Cornelius S. Hurlbut

Bibliography. C. Klein and C. S. Hurlbut, Jr., *Manual of Mineralogy*, 21st ed., 1993.

Chrysomonadida

An order of the class Phytamastigophorea. Many botanists consider this group a part of the class Chrysophyceae, the golden-brown algae. Chrysomonads, also known as Chrysomonadina, are usually small flagellates, *Mallomonas* being about the largest. They are yellow to brown because of

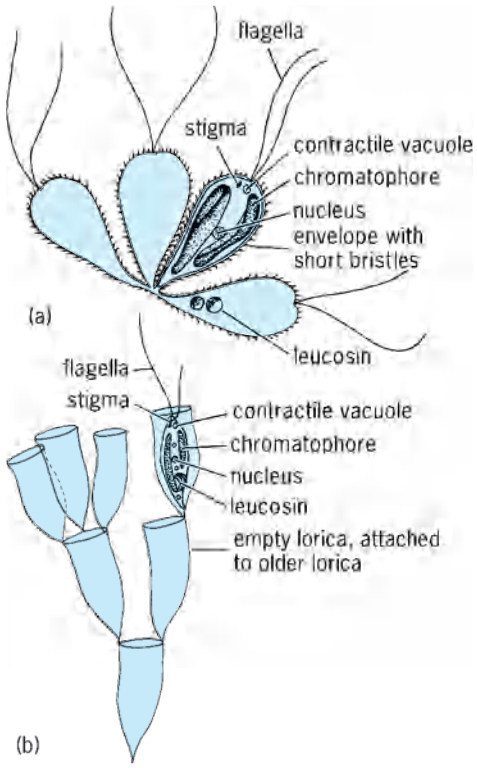

Chrysomonads. (*a*) Portion of colony of *Synura uvella*. Cell size is 20–40 μm, and colonies 100–400 μm. (*b*) *Dinobryon sertularia* colony. Cells measure 30–44 μm.

phycocrysin in the usual one or two chromatophores, but some lack chromatophores. *Uroglena* colonies may be quite large. Many species form diagnostic siliceous cysts. Palmelloid colonies (*Hydrurus*) consist of a tough, gelatinous matrix holding many nonflagellate cells. Starch is not formed, but fats are, and the refractive carbohydrate leucosin is common. The flagella are usually two, rarely three, and are subequal. Nuclei are small. *See* CILIA AND FLAGELLA.

Synura and *Dinobryon* are colonial types which produce enough oil to be a source of tastes and odors in drinking water (see **illus.**). *Dinobryon* often shows large numbers of cysts in the open ends of empty tests of old colonies. Chrysomonads may ingest food through the area near the flagellar base. Chrysomonadida are common in cooler fresh water, and certain types (Coccolithophora and Silicoflagellata) abound in the ocean. *See* CHRYSOPHYCEAE; COCCOLITHOPHORIDA; PHYTAMASTIGOPHOREA; PROTOZOA; SILICOFLAGELLATA. James B. Lackey

Chrysophyceae

A relatively large and diverse class of algae in the chlorophyll *a–c* phyletic line (Chromophycota). In protozoological classification, these organisms constitute an order, Chrysomonadida, of the class Phytomastigophora. Some workers align Chrysophyceae (golden or golden-brown algae) with Bacillariophyceae (diatoms) and Xanthophyceae (yellow-green algae) in the division Chrysophyta.

Chrysophytes typically are flagellate unicells, either free-living or attached, and solitary or colonial. There are, however, ameboid and plasmodial forms, solitary or colonial nonflagellate cells, filaments, and blades. Fresh-water forms, which are usually found in cold clear water, are more common than marine forms.

Although many unicellular chrysophytes have naked protoplasts, others are covered with organic or siliceous scales, or surrounded by a siliceous, cellulosic-pectic, or (rarely) calcareous lorica. Cells that compose blades and filaments have a polysaccharide wall. The statospore, an internally formed resting stage or cyst, occurs in the life history of most chrysophytes. It is urn-shaped, with a siliceous spiny body and a pectic plug. Similar cysts, assigned to the families Archaeomonadaceae and Chrysostomaceae, are found in the fossil record as far back as the Late Cretaceous.

Motile cells bear two subapically inserted, unequal flagella—a hairy (mastigoneme-bearing) flagellum directed forward, and a shorter, smooth flagellum directed backward. In some genera, the smooth flagellum is greatly reduced and may be represented only by its basal region, which may be transformed into a photoreceptive system, often in combination with an eyespot. Contractile vacuoles may be present in fresh-water forms.

Most chrysophytes are photosynthetic, but some are osmotrophic or phagotrophic. Each cell contains one or two chloroplasts (rarely more), which conform to the biochemical and ultrastructural patterns typical of Chromophycota. The chlorophylls are masked by carotenoids (especially β-carotene and fucoxanthin), producing the characteristic golden-brown color. Chloroplast endoplasmic reticulum is present, thylakoids are in groups of three, and chloroplast deoxyribonucleic acid is distributed in a ring. The chief storage product is chrysolaminaran, which is a glucose polymer with β-1,3 linkages, but lipids are also stored.

Asexual reproduction in unicellular forms is by longitudinal binary fission, autospores (nonflagellate cells that are miniatures of the parent), or zoospores, and in multicellular forms by fragmentation or zoospores. Sexual reproduction, observed rarely and in only a few species, is isogamous.

A satisfactory classification of the approximately 1000 species of chrysophytes has not yet been elaborated. One scheme emphasizes somatic organization, by analogy with green algae (Chlorophyceae), while another scheme emphasizes details of flagellation. Studies using electron microscopy and laboratory culture have led to the removal of certain groups to other classes. The choanoflagellates—collared, colorless flagellate unicells—were formerly aligned with chrysophytes or placed in their own class (Craspedophyceae), but similarity to sponge choanocytes supports their placement among animals, as the order Choanoflagellida of the class Zoomastigophora. Members of the order

Isochrysidales, characterized by having two equal flagella, have been transferred to the class Prymnesiophyceae. The bicosoecids, a small group of colorless unicells with a lorica formed of spiral or annular elements and with one of the two flagella serving as a contractile stalk, are sometimes appended to the Chrysophyceae or, alternatively, assigned to the Zoomastigophora. Finally, silicoflagellates, which are characterized by an elaborate siliceous endoskeleton and a unique protoplast, are placed in the subclass Dictyochophycidae or referred to their own class. They are a relictual group common in the fossil record as far back as the Early Cretaceous but represented in modern seas by only a few taxa. *See* ALGAE; BACILLARIOPHYCEAE; BICOSOECIDA; CHOANOFLAGELLIDA; CHRYSOMONADIDA; XANTHOPHYCEAE. Paul C. Silva; Richard L. Moe

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Chrysotile

Chrysotile is a fibrous mineral with a tubular morphology for each fibril (**Fig. 1**). It is a member of the serpentine mineral group, as are antigorite and lizardite. Chrysotile aggregates make up serpentine asbestos, which is the most important type of commercially mined asbestos. Russia and Canada are the main producing countries. Chrysotile displays interesting properties such as being thermally and electrically insulating, sound insulating, chemically inert, fire-resistant, mechanical energy–absorbing, and flexible with enough high tensile strength to be woven. There are hundreds of applications for chrysotile including fire retarder in buildings, roofing tiles, brake pads, weavable material for refractory clothes, filters, and fibers in fibrocement and road surfaces. *See* ASBESTOS; SERPENTINE; SERPENTINITE.

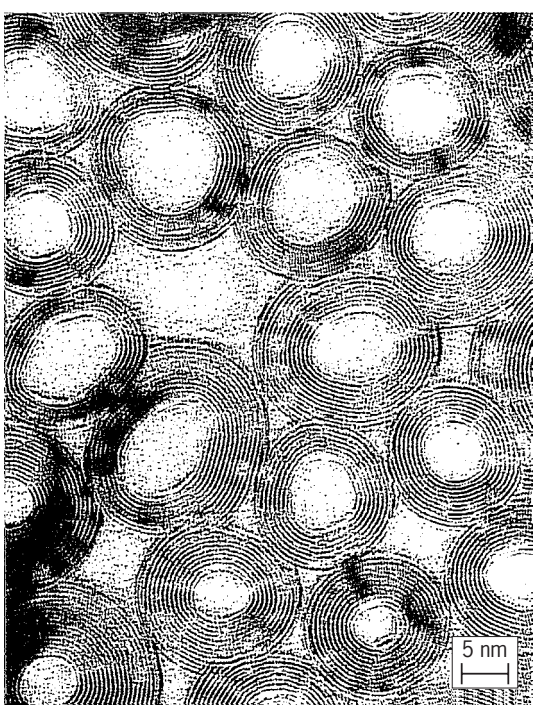

Fig. 1. Cross section of parallel chrysotile fibrils seen with a high-resolution transmission electron microscope (TEM). Black circular fringes image a chrysotile layer containing an octahedral and tetrahedral sheet.

Intensive inhalation of long and thin asbestos fibers over a considerable time period can induce pulmonary diseases such as asbestosis and lung cancers, as well as pleural diseases such as plaques, fibrosis, and mesothelioma. Such health hazards have drastically reduced the use of chrysotile, which is strictly regulated by law in western countries. *See* RESPIRATORY SYSTEM DISORDERS.

Geological occurrence. Chrysotile asbestos shows up mainly as cross-cutting veins in serpentinized green rocks, such as peridotites. It often forms from hydrothermal alteration ($T \leq 350°C$ or 662°F) of magnesium-rich olivines. In general, most serpentines are reacted on and below the sea floor, at slow-spreading mid-ocean ridges, as heated seawater interacts with exposed rocks from the upper mantle of the Earth. Obduction of such rocks accounts for serpentinites in some orogenic belts (for example, the Western Alps). *See* OLIVINE; PERIDOTITE.

Parallel assemblages of chrysotile fibrils in veins are either at a moderate angle (slip fibers) or more commonly normal (cross-fibers) to the vein walls. Progressive shear or opening of the vein by tectonic events seems to be required for massive development of long fibers.

Crystal-chemical properties. The chemical composition of chrysotile expressed as its formula unit is close to $Mg_3Si_2O_5(OH)_4$ [Mg = magnesium, Si = silicon, O = oxygen, OH = hydroxyl], with minor amounts of aluminum (Al) and iron (Fe) and traces of nickel (Ni) and chromium (Cr). This silicate layer is characterized by a tubular shape of each fibril as the basic unit of this material. Each tubule is made of 0.73-nanometer-thick cylindrical layers fitted into each other as telescopic tubes or coils. Most common tubule dimensions are 20 nm outer diameter, 5–7 nm inner diameter, and from 1 micrometer to several centimeters length. Fibrils may aggregate as fibers with a more or less close-packed arrangement of the tubules.

Each curved chrysotile layer is made of two assembled sheet components when viewed at the atomic scale: an outer octahedral sheet (Oc) and an inner tetrahedral sheet (Te). Silicate tetrahedra combine as a honeycomb network to form the tetrahedral sheet. Each silicate tetrahedron consists of a silicon atom surrounded by four close-packed oxygen atoms, and any tetrahedron shares three oxygen atoms with its neighbors whereas the fourth oxygen points toward the octahedral sheet. The octahedral sheet is built up by edge-sharing octahedra all lying on one of the faces, and each is centered by magnesium atoms (**Fig. 2**). All octahedron corners opposed to the tetrahedral sheet are hydroxyl groups. One-third of the inner corners are occupied by hydroxyl groups, whereas the two-thirds left are completed by the above-mentioned nonbridging oxygens of the tetrahedral sheet.

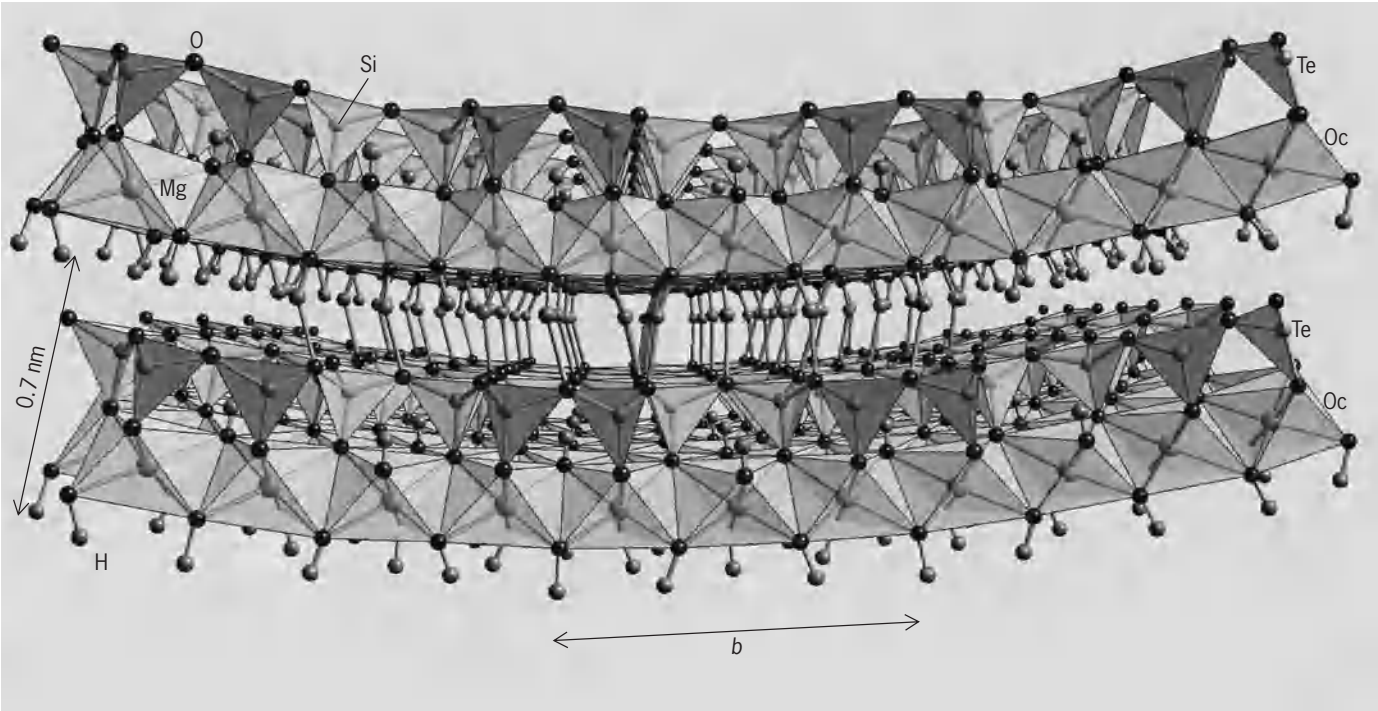

Fig. 2. Atomic structure of a portion of two contiguous chrysotile layers (Oc = octahedral sheet, Te = tetrahedral sheet). Atoms and bonds are seen on and inside translucent coordination polyhedra. Note the progressive deformation and breaking of hydrogen (H) bonds between the layers. This three-dimensional representation was done with CRYSTALMAKER™ software.

The curvature of the serpentine layer of chrysotile is thought to arise from the larger lattice dimensions of the octahedral sheet compared to those of the tetrahedral sheet. Accordingly, curvature is concave toward the tetrahedral sheet, and this partially compensates for any misalignment along with other tetrahedral sheet distortions. Such curvature disturbs the hydrogen (H) bonding between successive layers, that is, the linkage between hydroxyl groups of the octahedral sheet and bridging oxygen of the tetrahedral sheet of the next layer. Due to a progressive structural glide between consecutive layers, a periodic modulation of the hydrogen-bonding strength is observed along the fibril circumference (Fig. 2). This may weaken the average interlayer bonding compared to the case of lizardite, the flat serpentine. *See* HYDROGEN BOND.

Topologies of fibrils. Chrysotile tubules have layers arranged either as pure tube-in-tube cylinders or as spirally wound scrolls. The fibrils elongate either along the x axis (0.53 nm repeat) for normal chrysotile, or along the y axis (0.90 nm repeat) for para-chrysotile. Less frequently, helically wound (chiral) chrysotile tubules are found with curvature axes inside the lattice layer (similar to carbon nanotubes). *See* FULLERENE.

Flat serpentine layers of lizardite are translationally periodic in structure. Once such layers are curved and fitted into each other (as in chrysotile), the tubule loses both radial and circumferential periodicity. Periodicity is maintained along the tubule axis only, thus making chrysotile a kind of a one-dimensional crystal.

Fivefold symmetry. Each time a new cylindrical layer is added on the external surface of a normal chrysotile tubule, five more unit cells of b length are necessary to close up the circumference, because the increment of the perimeter is equal to five times the b length. The uniform distribution of such extra length along a turn implies that any interlayer configuration is recovered regularly five times per turn. Consequently, whatever the stacking sequence of curved unit cells along a tubule radius, the same sequence is recovered at integral multiples of $2\pi/5 = 72^\circ$. Accordingly, fivefold symmetry may be foreseen for the cylindrical lattice of chrysotile tubules as shown by the diffraction pattern of model lattice masks. More direct evidence of fivefold symmetry came from the numerical Fourier transforms of lattice images of tubule cross sections observed by high-resolution transmission electron microscopy (**Fig. 3**). From the above, chrysotile would be the first mineral having a normally forbidden fivefold symmetry axis. This axis would be macroscopic and unique, and in coincidence with the fibril axis. Accordingly, such fivefold symmetry might differ from that of quasicrystals of binary and ternary intermetallic alloys where six distinct fivefold axes are present at both macroscopic and microscopic scales. *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Polygonal serpentines. In serpentinite veins of ultrabasic rocks, long and thin chrysotile fibers coexist frequently with shorter but thicker polygonal fibers of serpentine. When observed with the transmission electron microscope, cross sections of polygonal serpentines exhibit either 15 or 30 radially distributed

sectors of flat serpentine, each akin to lizardite (**Fig. 4**). *See* ELECTRON MICROSCOPE.

There are geological occurrences where either 15-sectored or 30-sectored fibers are present, others where both types coexist. In the latter case, 30-sectored fibers are statistically larger in diameter (up to several micrometers) than 15-sectored ones. Electron diffraction patterns of single polygonal fibers have fivefold symmetry (Fig. 4). Reasons for such symmetry are the same as those invoked for chrysotile. In polygonal serpentines, flat sectors develop where the interlayer hydrogen bonding is satisfied along the radius of the fiber. Such bonding occurs easily 15 times per turn, or less easily 30 times per turn. *See* ELECTRON DIFFRACTION.

Experimental evidences support a transformation of chrysotile into polygonal serpentines, that is, a polygonization process of cylindrical tubules. Such a transition would take place for growing chrysotile when its diameter overshoots 50–100 nm. Further development of polygonal serpentine to micrometer size would occur by nucleation and further growth of serpentine layers inside and outside the polygonal tube.

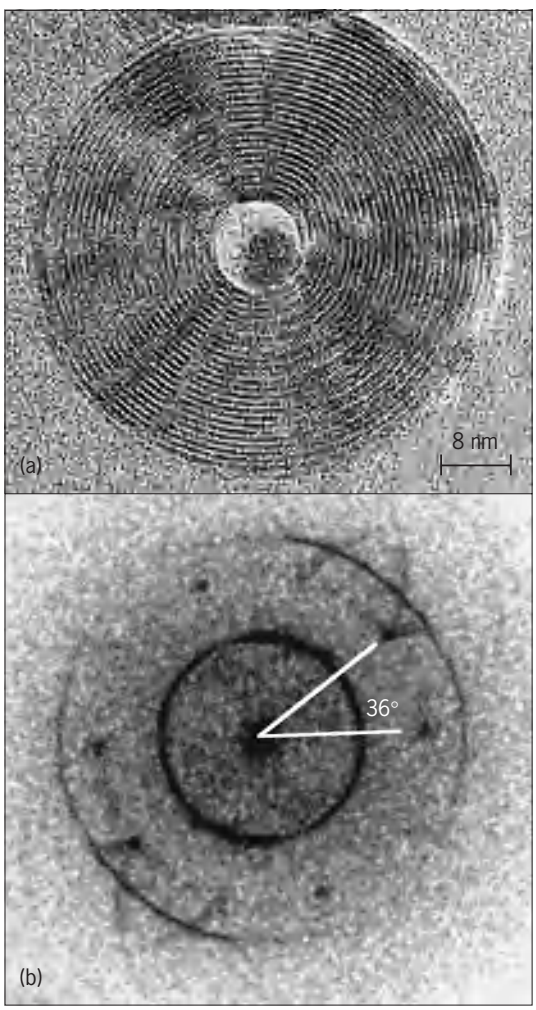

Fig. 3. Chrysotile tube. (*a*) High-resolution TEM image. (*b*) Numerical Fourier transform. The fivefold symmetry of the pattern shows up from the 36° angle between visible spots. (*Courtesy of B. Devouard*)

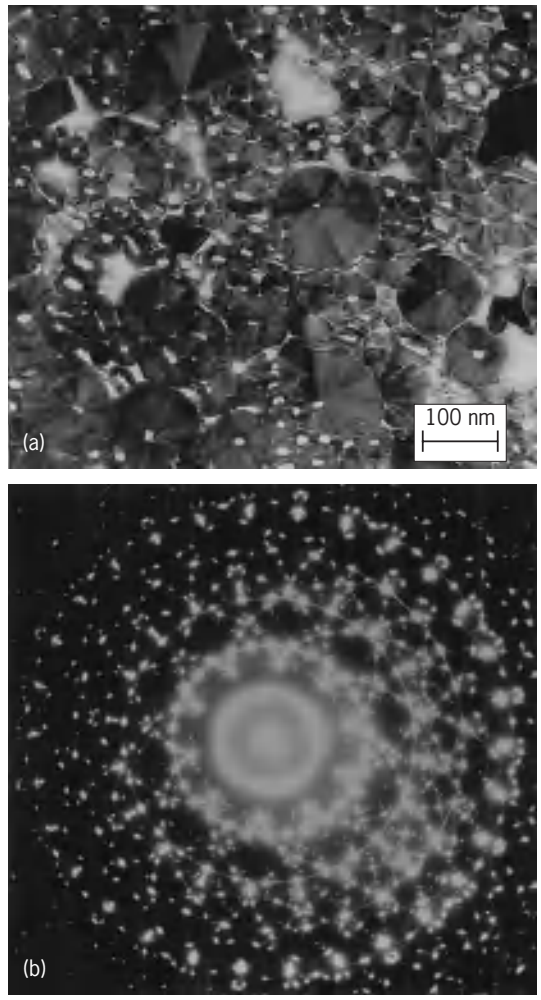

Fig. 4. Polygonal serpentine. (*a*) TEM image of cross-sectioned serpentines, with 15 sectors each. (*b*) Electron diffraction pattern of a serpentine as a rose with fivefold symmetry. (*Courtesy of A. Baronnet and B. Devouard*)

Fiber type and associated hazards. Pure material made of close-packed, poorly coherent chrysotile tubules is finely fibrous and silky. Scratching it with a nail or pouring water droplets on it produces a cottonlike or liquid soap aspect, respectively. Easy separability and thus dispersion of such fibrils in air or in a liquid are likely, so that such material is potentially harmful when inhaled or ingested by humans. Aggregates of polygonal serpentine fibrils that are tightly intergrown and short may be considered less hazardous since they are not able to produce dusts containing long fibers. Alain J. Baronnet

Bibliography. B. A. Cressey and E. J. W. Whittaker, Five-fold symmetry in chrysotile asbestos revealed by transmission electron microscopy, *Mineral. Mag.*, 57:729–732, 1993; B. Devouard and A. Baronnet, Axial diffraction of curved lattices: Geometrical and numerical modeling—Application to chrysotile, *Eur. J. Mineral.*, 7:835–846, 1995; M. Mellini, Chrysotile and polygonal serpentine from the Balangero serpentinite, *Mineral. Mag.*, 50:301–306, 1986; K. Yada, Study of microstructure of chrysotile asbestos by high resolution electron microscopy, *Acta Cryst.*, A27:659–664, 1971.

Chytridiomycota

A phylum of primitive fungi characterized by production of predominantly aquatic fungi characterized by production of uniflagellate zoospores (asexually produced motile spores). Chytrids are basal in the fungal lineage and are thought to have diverged about 1 billion years ago from an ancestor shared with protistian groups such as the Choanoflagellates and Mesomycetozoa. *See* CHOANOFLAGELLIDA; FUNGI.

Phylogeny. Chytridiomycota includes five orders: Blastocladiales, Chytridiales, Monoblepharidales, Spizellomycetales, and Neocallimastigales. Ultrastructural studies reveal distinct zoospore architecture in each order. Molecular analyses indicate that phylum Chytridiomycota is not monophyletic; rather order Blastocladiales is most basal, evolving among members of the Zygomycota and diverging from the core chytrids. Based on thallus structure, 90 genera and over 1500 species of chytrids have been named. Recent molecular and ultrastructural phylogenetic studies, however, have demonstrated unexpected genetic diversity among chytrid fungi, indicating that the actual species richness of this group is untapped. *See* ZYGOMYCOTA.

Morphology and reproduction. Chytrids exist in a variety of morphological forms, enabling them to adapt to life in diverse habitats. All thallus forms ultimately produce zoospore-generating structures called sporangia (**Fig. 1**). Members of the Blastocladiales and Monoblepharidales commonly form filaments, similar to hyphae of higher fungi. Simpler thallus forms consist of sporangia and rootlike rhizoids, and can vary in structure from a holocarpic thallus, in which the entire thallus has developed into a sporangium, to repeating sporangial units in polycentric rhizomycelia. Sexual recombination is achieved with fusion of motile gametes, fusion of contributing thalli, or fusion of motile gametes with nonmotile gametes. Zoospores function in asexual reproduction, in which stored lipids are used as an energy source through the coordinated interactions of organelles in the microbody–lipid complex. Positioned near the flagellum, the microbody–lipid complex can be viewed as the "battery" of the zoospore (**Fig. 2**). The microbody of the complex contains enzymes that help convert stored lipids to energy with the aid of mitochondria, and the membrane cisterna sequesters calcium, suggesting a role in signal transduction and the regulation of flagellar beating.

Fig. 1. Light micrographs of zoospore discharge. The spherical sporangium (diameter 30 μm) of a chytrid thallus growing on sweet gum pollen (the pollen grain is the sphere at the left) progressively releases a globular mass of zoospores into a vesicle. Zoospores swarm within the vesicle for about 5 min before swimming away as uniflagellate zoospores. (*Micrographs by Peter Letcher*)

Fig. 2. Electron micrograph of a longituditudal section through a chytrid zoospore (length 6 μm) with two visible lipid globules (homogenous gray circles at bottom left and upper right of the zoospore). The single flagellum is at the posterior end of the zoospore. The nucleus is surrounded by a dense, granular mass of ribosomes. Mitochondria cluster near microbodies that encircle the anteriorly located lipid globule. A fenestrated cisterna surrounds the posteriorly located lipid globule.

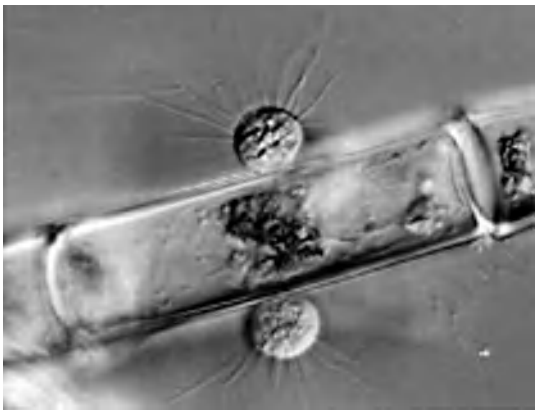

Fig. 3. Light micrograph of two chytrid thalli infecting the filament of a green alga. Hairlike appendages extend from the sporangia (diameter 30 μm) of the chytrid thalli. The contents of the host cell are collapsed.

Ecology. Chytrids are important saprotrophs, feeding on dead and decaying organisms in both aquatic and terrestrial habitats, and have been collected from the Arctic to the tropics. They are biodegraders and nutrient recyclers of resilient materials such as chitin, keratin, cellulose, and pollen. *Rhizophlyctis rosea* is a chytrid resistant to desiccation and a ubiquitous cellulose degrader found in soils. The Neocallimastigales, a group of obligately anaerobic chytrids that live in the digestive tracts of ruminant animals such as cows and sheep, are especially important in improving feed value of cellulosic and hemicellulosic foods. *See* FUNGAL ECOLOGY.

Parasitic activity. Chytrids are also common parasites of algae (**Fig. 3**), other fungi, insects, plants, and animals. Algal parasites are implicated in the natural control of algal blooms. *Coelomomyces* decreases

mosquito populations by parasitizing mosquito larvae and inducing adult female mosquitoes to oviposit its sporangia rather than mosquito eggs. *Synchytrium endobioticum* causes black wart of potatoes and makes potato production difficult in infested soils due to the longevity of chytrid resting spores which remain dormant for long periods before germination, withstanding adverse conditions. *Caulochytrium protostelioides* parasitizes hyphal fungi on plant leaves. *Batrachochytrium dendrobatidis*, discovered in the mid-1990s and considered a primary culprit in the global deaths and mass extinctions of frogs and other amphibians, is the first documented chytrid parasite of vertebrate animals. *See* PARASITOLOGY. Martha J. Powell

Bibliography. C. J. Alexopoulos, C. W. Mims, and M. Blackwell, *Introductory Mycology*, 4th ed., 1996; D. J. S. Barr, Phylum Chytridiomycota, pp. 454–466, in L. Margulis et al. (eds.), *Handbook of Protoctista*, 1990; J. S. Karling, *Chytridiomycetarum Iconographia*, Lubrecht & Cramer, Monticello, 1977; M. J. Powell, Looking at mycology with a Janus face: A glimpse at Chytridiomycetes active in the environment, *Mycologia*, 85:1–20, 1993; F. K. Sparrow, *Aquatic Phycomycetes*, 2d rev. ed., University of Michigan Press, Ann Arbor, 1960.

Great blue heron (*Ardea herodias*), Galápagos Islands, Ecuador. (*Photo by H. Vannoy Davis; © 2001 California Academy of Sciences*)

Ciconiiformes

An order of predominantly long-legged, long-necked wading birds including herons, ibises, spoonbills, storks, and their allies, and also the hawklike New World vultures which were previously placed in the Falconiformes. The flamingos have traditionally been placed in this order; however, they show a number of unique features and differences and so are placed in a separate order, Phoenicopteriformes. Some workers have suggested that the ibises are members of the Charadriiformes and that the herons should be placed in a separate order; however, these suggestions are controversial. *See* CHARADRIIFORMES; FALCONIFORMES; PHOENICOPTERIFORMES.

Classification. The order Ciconiiformes is divided into four suborders and seven families: Ardeae, with the family Ardeidae (herons; 62 species; worldwide); Balaenicipites, with the family Balaenicipitidae (shoe-billed stork; one species; Africa); Ciconiae, with the families Scopidae (hammerhead; one species; Africa); Threskiornithidae (ibises and spoonbills; 33 species; worldwide); Ciconiidae (storks; 17 species; worldwide); and Cathartae, containing the families Teratornithidae (Miocene to Pleistocene of South and North America) and Cathartidae (New World vultures; seven species; New World).

Several of these families are divided into subgroups, the most distinct of which are the ibises (Threskiornithinae) and spoonbills (Plataleinae). The affinities of the New World vultures to the ciconiiforms, and especially to the storks (Ciconiidae), have been demonstrated rather convincingly. For many decades, numerous ornithologists have queried the inclusion of these birds in the Falconiformes, but their true affinities had been obscure.

Fossil record. Except for the Cathartae, the fossil record of the Ciconiiformes is scattered, considering that these are large wading birds, and does not illustrate anything special about the relationships and distribution of these birds. The New World vultures have an extensive fossil record, including in the Old World; the Teratornithidae are found in both North and South America and include the longest winged birds.

Characteristics. The herons (Ardeidae, see **illustration**), shoe-billed stork (Balaenicipitidae), hammerhead (Scopidae), ibises and spoonbills (Threskiornithidae), and storks (Ciconiidae) are mainly wading birds with strong legs, living in marshes and other wet areas. They feed on crustaceans, fish, amphibians, and other animals, which are caught in various ways depending on the structure of the bill. Most are diurnal, but a few are nocturnal. Some of the storks are scavengers. Most species are colonial and nest in mixed, often huge, colonies in trees or on the ground, as well as a few on cliff ledges; the hammerhead builds a huge, domed nest. All are strong fliers, usually having to cover a long distance between the nesting site and feeding areas; some are excellent soarers.

Most species have elaborate courtship displays and form tight pair bonds; many species of herons possess specialized courtship plumes. The young remain in the nest and are cared for by both adults until they are able to fly. They are tropical to cool temperate, with those species breeding in colder areas

migrating to warmer regions for the winter. Many species of herons and egrets had been intensively hunted for their courtship plumes; some were almost extinct before fashions of apparel changed and the birds received protection. The numbers of many species have been drastically reduced, and several are threatened mainly because of lack of the needed wetland habitat. The white stork of Europe, which nests readily on human structures and is a symbol of fertility, has been reduced greatly in much of its range because of accidents with wires and loss of wetlands.

Cathartae. The New World vultures appear to be typical hawks and had been placed in the Falconiformes since the beginnings of avian classification but with ever-increasing doubt. It has been shown rather convincingly that they should be placed in the Ciconiiformes and are closest to the storks. They are placed in a separate suborder to emphasize their specialization as scavengers. These vultures are large soaring birds; the Andean condor has the largest wing span of any living land bird. (However, their wings are shorter than the largest members of the fossil teratornithids.) Vultures locate their food from the air either by sight or by smell; these abilities differ in the several species, with the Turkey vulture (*Cathartes aura*) finding its food in forested areas by smell. Vultures may be solitary or hunt in loose flocks; larger concentrations may exist during the winter. They nest solitarily, with the nest placed on the ground or on a cliff ledge. The one to three young remain in the nest until they can fly. Most species are common, but the California condor (*Gymnogyps californianus*) has decreased steadily in numbers and became extinct in the wild. Survival of this species will depend on captive breeding and reintroduction of birds to the wild, which appears to have been successful following releases in California and in the Grand Canyon of Arizona.

The fossil history of the New World vultures is excellent. It suggests that the Cathartidae had their early evolution in the Old World and invaded the New World only in the Neogene. More interesting are the closely related Teratornithidae known from North and South America, which include the birds with the largest wing spans. The relatively small *Teratornis merriami* had a wing spread of up to 3.8 m (12 ft), and the larger *T. incredibilis* had an estimated wing span of up to 5.9 m (19 ft). However, *Argentavis magnificens* from the Late Miocene of Argentina had an estimated wing span of 7–7.6 m (22.5–24 ft) and a weight of about 120 kg (264 lb), making it the largest flying bird known by far. *See* AVES.

Walter J. Bock

Bibliography. J. del Hoyo et al., Order Ciconiiformes, pp. 375–506, in J. del Hoyo et al. (eds.), *Handbook of the Birds of the World*, vol. 1, Lynx Edicions, 1992; J. A. Hancock, J. A. Kushlan, and M. P. Kahl, *Storks, Ibises and Spoonbills of the World*, Academic Press, 1992; D. C. Houston, Family Cathartidae, in J. del Hoyo et al. (eds.), *Handbook of the Birds of the World*, vol. 2, pp. 24–41, Lynx Edicions, 1994; C. König, Zur systematischen Stellung der Neuweltgeier (Cathartidae), *J. Ornithol.*, 123:259–267, 1982; J. A. Kushland and J. A. Hancock, *The Herons of the World*, Oxford University Press, 2005.

Cilia and flagella

Centriole-based, motile cell extensions. These organelles are usually indistinguishable in fine structure as seen with the electron microscope, but quantitatively there are many (several hundred) cilia and few or fewer (usually one or two) flagella, on one cell. Bacterial or prokaryotic flagella are entirely different organelles that are not considered in this article, which concerns only eukaryotic (higher-cell) flagella. *See* BACTERIA.

Characteristics and occurrence. Flagella move with undulatory motion in which successive bending waves progress along the length of the organelle (**Fig. 1**), whereas cilia move with flexural motion consisting of a planar effective stroke, with the organelle extended perpendicular to the cell body, followed by a nonplanar curving recovery stroke, with the organelle pulled parallel to the cell body. Cilia and flagella are low-Reynolds-number systems, which implies that viscous forces are more important to the motion of the organelles than inertial forces are. On a macroscopic scale, this is equivalent to swimming in molasses, and the fluid can be thought of as being clawed forward by the effective stroke of the cilium or as being pushed away by the propagating undulatory wave of the flagellum. Beat frequencies of cilia and flagella in different organisms range from about 2 to more than 50 per second. During a beat, bends usually originate and grow in restricted regions at the base of the organelle in the form of circular arcs. They then propagate toward the tip at rates of several

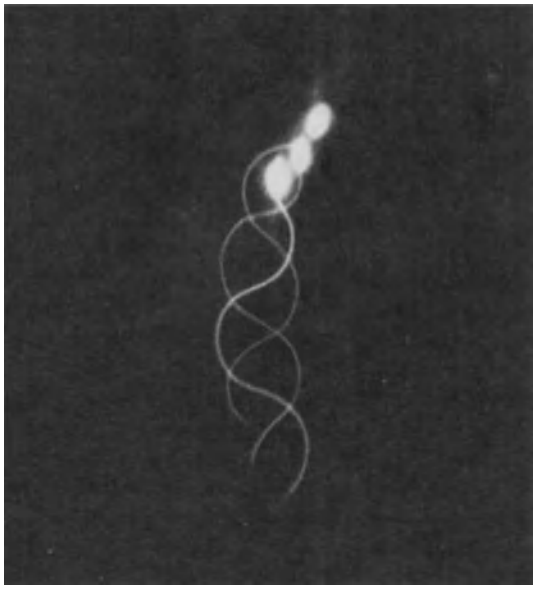

Fig. 1. Sea urchin sperm shown in three positions by dark-field microscopy. Note undulatory waveform as it swims. (*From C. J. Brokaw, Nonsinusoidal bending waves of sperm, J. Exp. Biol., 43:155–169, 1965*)

hundred micrometers per second without changing amplitude or form. On sperm tails, bends of equal magnitude but of opposite direction are separated by straight regions of uniform length to give a symmetrical waveform that is superficially sinusoidal. A similar but asymmetrical waveform on a shorter cilium produces one typical form of flexural motion. Both organelles function to move water past the cell. Their action may bring food and oxygen into an animal, or it may propel the cell to a new environment. Cilia often move sheets of mucus at velocities of about 0.008 in. (0.2 mm) per second, either in water or over surfaces exposed to humid air, as in the human respiratory tract. Particles trapped in such sheets are effectively cleared from the epithelial surface. *See* CELL MOTILITY.

Distribution. Cilia were first described by A. van Leeuwenhoek in 1676. The words cilium (eyelash) and flagellum (whip) are accurate descriptions of the appearance of these cell organelles when they are seen under the light microscope. Cilia are present on protozoa, such as *Paramecium*, and on metazoan cells of many different tissue types. For example, in humans, many cells lining the respiratory and reproductive tracts, the ear and the olfactory epithelium, as well as ependymal cells in the brain, have functioning cilia. A flagellum is present on human sperm; in fact, the sperm of most animals, and the male gametes of many lower plants are flagellated. Even the *Ginkgo* spermatozoon is ciliated. Derivatives that are similar to cilia but nonmotile are found on many sensory structures in a variety of animals and include insect pressure receptors and vertebrate photoreceptors. In addition, many ordinary types of cells of vertebrates, for example, from the thyroid, the kidney, or the pituitary gland, also possess modified nonmotile derivatives that resemble cilia, and other cell lines can be induced to grow them in tissue culture. *See* CILIOPHORA; EPITHELIUM; SPERM CELL; TISSUE CULTURE.

Size. Relative to cell size, both cilia and flagella are very long organelles. Flagella may be more than 50 micrometers long and, in sperm, may be four or more times the cell length. Certain compound cilia, such as the comb plates of ctenophores, are truly macroscopic structures, visible to the naked eye. In these structures, many thousands of individual cilia are joined together laterally, and the length of each cilium runs to a few millimeters. Usually, however, cilia range from 5 to 15 μm in length.

Metachronism. Fields of cilia or flagella such as those in invertebrate flame cells can beat synchronously. Often, however, even on a single cell, cilia beat so that coordinated waves of bending appear to cross the field. Originally, these waves were described as the sort that are seen rippling across a wheat field when the wind is blowing. The phenomenon of beating slightly out of phase is known as metachronism. Several types of metachronism are recognized, since the plane of ciliary beat may be either perpendicular (diaplectic) or parallel (orthoplectic) to the direction of the metachronal wave. Metachronal coordination is usually effected by purely mechanical means; that is, viscoelastic coupling establishes a phase difference between individual cilia so that there is minimal mechanical interference between one beating cilium and its neighbors. In a ciliated epithelium the cells can be functionally linked together so that metachronal waves pass smoothly from cell to cell.

Behavior. Under normal conditions, most cilia and flagella beat continuously, although some cilia (for example, mammalian respiratory cilia and mammalian sperm) require special activation for beating. Many organisms are able to alter properties of ciliary beat, including form, frequency, the direction of the effective stroke, or the direction of bend propagation; sometimes, the ciliary stroke can be arrested. Cells usually alter beat properties by common mechanisms leading to increases in second messengers, including internal calcium ion (Ca^{2+}) or cyclic nucleotide concentrations. For example, in *Paramecium*, depolarization of the cell membrane opens voltage-sensitive calcium channels in the ciliary membrane, which leads to an increase in calcium ion concentration in the cilium above 10^{-7} *M*. This in turn leads to slowed and then to backward swimming at concentrations above 10^{-6} *M*. When calcium ion is pumped out of the cilium, the normal beat resumes. Different second messengers may have different effects. Cyclic adenosine monophosphate (cAMP) increases normal beat frequency in *Paramecium* cilia, leading to faster forward swimming. Such processes allow the cell to respond to various stimuli and form the basis of behavioral responses of free-swimming ciliated cells. The same messenger may have different modes of action in different cilia; for example, calcium ion activates and increases beat frequency in mammalian respiratory cilia.

Ultrastructure. The electron microscope reveals that the cilium or flagellum is really an internal organelle since it is bounded by the cell membrane and enclosed at the tip. The main internal structure of the cilium is the axoneme. Under the electron microscope a single axoneme contains a fixed pattern of microtubules. The microtubules are not simple single units; rather, nine doublet microtubules are found on the periphery of the axoneme surrounding two central elements. This is the so-called 9 + 2 pattern (**Fig. 2**).

Microtubule arrangement. Each peripheral doublet is composed of one complete and one partial microtubule. The microtubules are themselves composed of subunits arranged into microfibers or protofilaments. The complete microtubule, labeled subfiber A, consists of a ring of 13 protofilaments. The incomplete microtubule, subfiber B, is a C-shaped structure of 10 or 11 protofilaments. In a transverse section the doublets appear as figure eights, with two short arms and several linkages extending from one end. The doublets form a ring around the central pair. The diameter of this ring is very constant, about 0.2 μm, even in different types of cilia or in cilia of different species.

Microtubule pattern variation. The 9 + 2 pattern is present in all known motile cilia. Motile flagella, especially sperm tails, sometimes possess altered

Fig. 2. Electron micrograph of mussel gill cilia cross sections showing 9 + 2 pattern of axoneme. At the left are cross sections through tips at various levels. (*From P. Satir, Studies on cilia, II: Examination of the distal region of the ciliary shaft and the role of the filaments in motility, J. Cell Biol., 26:805–834, 1965*)

patterns, although these are usually simple variations upon the basic array. Flatworm sperm, for example, show a 9 + 1 pattern (one central element instead of two), although somatic cilia of the same animal are 9 + 2. Insect sperm show even greater divergences. Mammalian sperm have the 9 + 2 pattern, but in addition have a set of thick fibers alongside the axoneme. The most simply organized motile flagella in eukaryotes are the 6 + 0 and 3 + 0 forms present in male gametes of certain gregarine protozoa. In these flagella, the doublet tubules appear to possess normal arms, but the spokes and central pair are lacking. Nonmotile flagellar derivatives usually have 9 + 0 patterns and lack both the arms and the central pair. All these patterns are constant within any one species of animal or plant, with the exception of certain very rare variants that are probably mistakes in assembly.

Symmetry. If the arms are disregarded, there is a plane of bilateral symmetry in the normal cilium which passes between two peripheral doublets, through the central pair, and through one doublet on the side opposite. This doublet is numbered 1, and the plane of bilateral symmetry is also the plane of the effective stroke. By looking from the cell surface and outward along the axoneme, looking clockwise from subfiber B of doublet number 1, the sequence is subfiber B, subfiber A, arms of doublet number 1, subfiber B, subfiber A, arms of doublet number 2, and so on. The arms always extend from subfiber A of any doublet N toward subfiber B of the adjacent doublet. In this manner the doublets can be numbered unequivocally. The two peripheral doublets between which the plane of bilateral symmetry passes are the numbers 5 and 6, and these may be specially distinguished in some cases by being connected by a distinctive bridge.

Axonemal linkages. The axonemal microtubules are interconnected into a functional whole by groups of projections extending in different directions from each subfiber A (**Fig. 3**). These include the two circumferentially oriented rows of dynein arms. The outer arms repeat at spacings of about 24 nanometers. The outer arms are uniform, but the inner arms are sometimes of two or three types with a more complicated spacing. There are also several types of circumferentially oriented links, some of which form elastic connections, some inelastic connections, between doublets. The periodicity of these links is about 96 nm. Radially oriented spoke groups project toward the central sheath with a major period of 96 nm; the spoke groups often comprise three individual spokes, each of which ends in an expanded spoke head. The spokes are spaced at intervals of 32, 24, and 40 nm, with the first spoke of each group in alignment with a set of dynein arms. There are six central sheath projections per spoke group. Such vernier alignments may prove to be important in the regulation of beat form.

Morphogenesis. At the base of the cilium or flagellum there is a basal body, or kinetosome, that is similar to, and sometimes derived from, a centriole. The basal body may have extensions of various sorts attached to it, notably a basal foot that indicates effective stroke direction and prominent striated rootlet fibers in many cilia. Ordinarily, the ciliary axoneme originates and grows in a membrane protrusion which forms just above the basal body, either at the cell surface or deeper inside the cytoplasm. The basal body remains attached to the cell membrane throughout morphogenesis by a structure that extends from the microtubules to the membrane, where it is seen as a ciliary necklace. The necklace forms a barrier that segregates the ciliary membrane from the general cell membrane. As morphogenesis proceeds, from a pool of subunits, axonemal

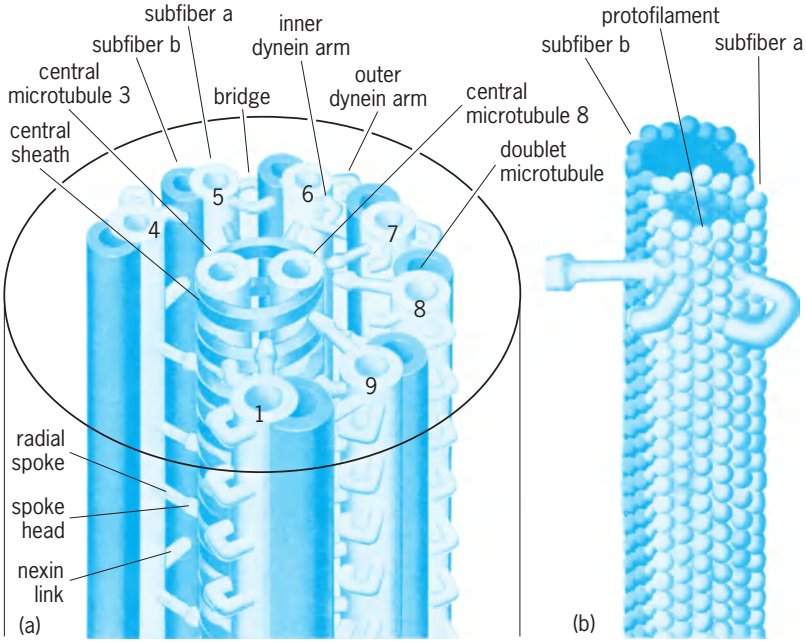

Fig. 3. Ultrastructure of a cilium. (*a*) Three-dimensional reconstruction with doublets 2 and 3 removed to show spoke periodicities more clearly. (*b*) Three-dimensional wall structure of a doublet showing tubulin subunits. (*After P. Satir, How cilia move, Sci. Amer., 231(4):44–52, 1974*)

microtubules polymerize above the basal body. Continuity of the microtubules between axoneme and basal body may be traced through a complex transition zone. Evidently, polymerization of the axonemal microtubules takes place at the growing tip of the cilium. Certain cells such as the ameboflagellate *Naegleria* can grow or resorb the cilia in response to internal programming during the cell cycle, or upon physiological stimulation. *See* CENTRIOLE.

Genetic control. Morphogenesis is under genetic controls, in part by nuclear genes. For example, certain nuclear mutants in *Chlamydomonas* fail to polymerize the central microtubules. Other mutants lack radial spokes or either the outer or the inner row of dynein arms. In these cases, the mutant flagella are partially or completely paralyzed, or the beat form or frequency is altered. Paralyzed and behavioral mutants have also been found in other species, ranging from *Paramecium* to humans. Approximately 1 in 40,000 people suffer from an inherited genetic defect originally known as Kartagener's syndrome, in which the cilia of the bronchial tract and the sperm flagella (in males) are paralyzed, in some instances because dynein arms are completely lacking on the doublet microtubules. People with this syndrome have considerable respiratory difficulty, and males are infertile. These people also have a 50% chance of their heart and other internal organs being on the opposite side of the body compared with normal anatomy, possibly as a result of ciliary immotility during embryonic development.

Growth parameters. The extent of each axonemal component is strictly determined. In mussel gill lateral cilia, all subfiber B's of the peripheral doublets end at the same level. Simultaneously the arms and the spokes also terminate. However, both the central pair and the subfiber A's of the doublets continue, the central pair to the exact end of the ciliary tip, where the elements are topped by a distinctive finial, the central pair cap. Every continuation of subfiber A is unique in length so that subfiber A's of doublets 3 and 8 are overall about 1 μm, or 10%, longer than subfiber A's of doublets 1, 4, and 5. A general reduction in axonemal diameter accompanies microtubule termination and gives the cilium a distinctive conical tip. These growth parameters and the shape of the tip are different for different kinds of cilia and flagella. In certain mammalian cilia the tip is blunt, and the membrane is decorated externally by a cell coat, the ciliary crown.

Chemical composition. Cilia may be isolated en masse from ciliated protozoa, such as *Tetrahymena*, and flagella may be isolated as sperm tails. Such preparations are suitable for chemical analysis, and reasonably pure preparations of axonemes and of membranes can be obtained in quantities adequate for study. The major protein component (70–80%) of the axoneme is tubulin, the structural protein of the microtubules. The second largest component (10–15%) is dynein, the axonemal adenosine triphosphatase protein (ATPase). The remaining 10–15% of the axonemal protein is distributed among approximately 250 distinct polypeptides that can be visualized by two-dimensional electrophoresis. Some of these polypeptides can be assigned to specific axonemal structures, but most have not yet been otherwise characterized. The molecular unit of tubulin is a heterodimer of α and β subunits, each a globular protein about 4 nm long and 5 nm wide. One $\alpha\beta$ tubulin heterodimer (molecular weight 100,000) contains one molecule of bound guanosine triphosphate and one more loosely bound guanosine diphosphate. Axonemal and cytoplasmic microtubules are similar in most properties, but they differ in their stability to depolymerization. In *Chlamydomonas*, axonemal tubulin is acetylated locally after synthesis, while tubulin in most other parts of the cell is not. Another difference is that the midwall of the doublet microtubules is not tubulin, but proteins related to intermediate filament proteins, termed tektins. This may account for the differential stability or other unique properties of the axonemal microtubules.

The axonemal adenosine triphosphatase (ATPase), dynein, consists of a group of related globular proteins of slightly different solubilities with high molecular weights in the range of $1–2 \times 10^6$. Isolated dynein (**Fig. 4**) resembles a bouquet of two or three globular masses attached to a common base by slender stalks. Analysis after brief protease digestion of the molecule shows that each head contains at least one distinctive heavy polypeptide chain (molecular weight $4–5 \times 10^5$) that has ATPase activity, and

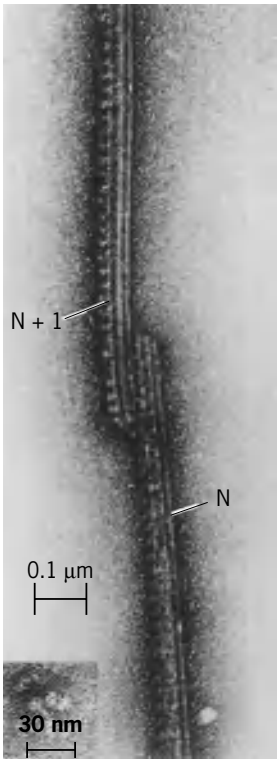

Fig. 4. Sliding of doublet microtubules from an axoneme after destruction of the control system. The dynein arms of the lower doublet (N) have interacted with subfiber B of the upper doublet (N + 1), displacing it toward the ciliary tip. Inset shows an isolated dynein molecule. (*After P. Satir, Structural analysis of the dynein cross bridge cycle, in F. D. Warner et al., eds., Cell Movement, vol. 1, Alan Liss, 1989*; *inset courtesy of K. Barkalow*)

each chain can be cleaved into two distinctive fragments by ultraviolet light in the presence of vanadate ions. In addition, purified dynein contains two or three intermediate-size chains (approximately 1×10^5) and several types of light chains (approximately $1–3 \times 10^4$). Outer and inner arm dyneins of the same axoneme may be different biochemically and structurally. Inner arm dyneins may be closely associated with other axonemal proteins, including small amounts of actin and centrin, in a so-called regulatory complex. The dynein genes of sea urchin and several other organisms have been cloned, and amino acid sequences of the heavy chains have been derived from the deoxyribonucleic acid (DNA) sequences. Each axonemal heavy chain is probably the product of a different gene. The sequences of the ATPase sites of the various dyneins seem highly conserved. Transcription of the axonemal dyneins, of tubulins and, other ciliary proteins, is increased after experimental deciliation of cells.

The membranes of flagella and cilia appear similar in structure, but they differ in composition, from the adjacent cell membranes. Like other membranes, the ciliary membrane is a phospholipid bilayer into which proteins are inserted. Some major membrane components of the cilia and flagella are glycoproteins, including immobilization antigens and other molecules involved in mating or cell recognition. Some ciliary membranes contain specific ion channels, including the voltage-sensitive calcium ion channel responsible for backward swimming in *Paramecium*.

Axonemes sometimes contain target proteins for the second messengers that control their behavior, including calmodulin which binds calcium and loosely or tightly bound protein kinases that may phosphorylate other axonemal proteins in the presence of cyclic nucleotides. A small polypeptide that copurifies with the outer arm dynein is phosphorylated in a cAMP-dependent manner and might act as a regulatory light chain of this dynein.

Mechanism of movement. Cilia and flagella are active structures, and the energy for their motion is provided by the dephosphorylation of adenosine triphosphate (ATP) at the dynein arms distributed along most of the length of the organelle. The membranes of cilia and flagella can be removed selectively by treatment with a nonionic detergent, such as Triton X-100. The motile mechanism in the resulting demembranated cilia and flagella is undamaged, so that when placed in an appropriate saline medium containing added ATP, the organelles regain motility and their pattern of beating closely resembles that of the same organelles in the living organism. An abrupt decrease in ATP concentration causes beating axonemes to set into a rigor state in which the stationary axonemes maintain the bent waveforms that they had at the moment the ATP concentration was lowered. The preservation of these rigor bends results from the dynein arms forming fixed cross-bridges between adjacent doublets in the absence of ATP. Movements between doublets can be regenerated by local addition of ATP; they are the result of a cyclic process in which the binding and hydrolysis of a molecule of ATP at a dynein arm from subfiber A of one doublet (N) causes it to release its cross-bridge attachment to the B subfiber of the adjacent doublet (N+1), change its angle, reattach to a new site closer to the basal end of the B subfiber, and then pull. This paradigm is known as the sliding microtubule model of ciliary motility. Both inner and outer dynein arms operate to produce sliding with the same polarity. Microtubule sliding can be seen in a beating axoneme by attaching gold beads to different parts of the structure and observing their periodic motion in the light microscope.

The outer dynein arms can be quantitatively extracted from certain axonemes. This causes the beat frequency to fall. The beat frequency also decreases in mutants of *Chlamydomonas* where the outer dynein arms or specific portions of them are missing. When corresponding mutants lack a specific type of inner arm, the beat frequency is largely unchanged but the bending pattern of the beat is altered. The suggestion is that sliding produced by the outer dynein arms is a primary determinant of beat frequency, while the inner arms primarily regulate beat form.

Active sliding of the doublet tubules can be demonstrated directly by brief treatment of demembranated axonemes with the protease trypsin. Trypsin rapidly destroys most of the spokes and circumferential links while leaving the arms and doublet tubules relatively intact. When ATP is then added, the treated axonemes no longer generate bending waves, but elongate and grow thinner as the doublets slide apart (Fig. 4). Isolated dynein arms will rebind to axonemal or other microtubules and slide them across glass surfaces in the presence of ATP. Dynein and tubulin are the only proteins necessary to produce sliding. This sliding microtubule mechanism is applicable to other microtubule-based forms of cell motility.

Since all dynein arms operate in the same way, the normal development and propagation of bends in the axoneme requires the presence of a suitable control system to generate appropriately phased localized sliding movements between the axonemal doublets. Trypsin damages this control system that normally limits sliding, leading to an uncoupling of sliding from bending.

The control system permits sliding to be asynchronous; one possibility is that sliding switches back and forth between the two bilaterally symmetrical halves of the axoneme and the other, during each stroke cycle. Inhibitors that act on the switches cause beating to cease abruptly at one particular stroke position. Such arrested cilia can be moved from one position to another by replacing one inhibitor with another, and microtubules from about half of each axoneme prepared from the arrested cilia will slide normally after trypsin treatment. This indicates that arrest is an active state in which the dynein arms producing bending to one side remain potentiated,

while those producing bending to the reverse side are inhibited.

The main elements thought to convert the fundamental ATP-powered sliding interaction into local bending are the spokes. The spoke heads appear to be attached to the central sheath at points where bends develop, so as to provide the needed resistance to sliding, while the central tubules and sheath act as a fulcrum around which bending occurs. The presence of spokes affects the behavior of the inner dynein arms. Flagellar mutants of motile *Chlamydomonas* that possess axonemes lacking spokes appear capable of tubule sliding but not of bending. The 3 + 0 flagella in sperm of certain gregarines mentioned above lack both spokes and central tubules, yet generate helical bending waves of low frequency, which indicates that spokes and central tubules are not absolutely required for bending. However, their interactions may play a critical role in generating the asymmetric bending waves characteristic of normal ciliary motion and in the control of heat parameters. *See* ADENOSINE TRIPHOSPHATE (ATP). Peter Satir

Bibliography. R. A. Bloodgood (ed.), *Ciliary and Flagellar Membranes*, 1990; F. D. Warner, E. R. Gibbons, and P. Satir, *Cell Movement*, vol. 1, 1989.

Ciliatea

The single class of the subphylum Ciliophora. This group has the characteristics of those defined for the subphylum. This protozoan class is divided into the subclasses Holotrichia, Peritrichia, Suctoria, and Spirotrichia. *See* CILIOPHORA; HOLOTRICHIA; PERITRICHIA; SPIROTRICHIA; SUCTORIA. John O. Corliss

Ciliophora

A subphylum of the Protozoa. The ciliates are a fairly homogeneous group of highly differentiated, unicellular organisms. Over 5000 species have been described, and many more surely exist but remain to be discovered. Typically, ciliates are larger than most other protozoa, ranging from 10 to 3000 micrometers or about 0.004 to 0.12 in. Some larger species are easily visible to the naked eye. The majority of them are free-living forms, found abundantly in a variety of fresh- and salt-water habitats, although a few entire groups live in association with other organisms, generally as harmless ecto- or endocommensals. Their principal value to humans is as experimental animals in a host of investigations concerned with fundamental problems of biology.

Systematics. Classification of ciliates followed here is that published in the "Honigberg Report," which appeared in 1964. That classification scheme, of the entire phylum Protozoa, was determined by an international committee of distinguished protozoologists. It is a reasonable and useful classification. Separate articles appear on each group included in the classification listed below. Subordinal divisions of some orders are commonly recognized by protozoologists, but such taxa are not considered here.

Subphylum Ciliophora
- Class Ciliatea
  - Subclass Holotrichia
    - Order: Gymnostomatida
      - Trichostomatida
      - Chonotrichida
      - Apostomatida
      - Astomatida
      - Hymenostomatida
      - Thigmotrichida
  - Subclass Peritrichia
    - Order Peritrichida
  - Subclass Suctoria
    - Order Suctorida
  - Subclass Spirotrichia
    - Order: Heterotrichida
      - Oligotrichida
      - Tintinnida
      - Entodiniomorphida
      - Odontostomatida
      - Hypotrichida

Protociliata or opalinids (**Fig. 1**) are entirely excluded. They are considered members of the subphylum Sarcomastigophora, because they possess only one type of nucleus, a symmetrogenic rather than a homothetogenic type of binary fission, and true syngamy, or sexual reproduction with anisogamous gametes, rather than conjugation. There are four genera of opalinids which are large, mouthless, multiflagellated forms found almost exclusively as saprozoic commensals in the posterior end of the alimentary tract of frogs and toads. *See* SARCOMASTIGOPHORA.

Major subdivisions of the ciliates are reduced to the four subclasses named above. Within the Holotrichia are included the Chonotrichida, a group formerly accorded a high-level, independent taxonomic status. *See* HOLOTRICHIA.

Characteristics. The usual ciliate life cycle is fairly simple. An individual feeds and undergoes binary

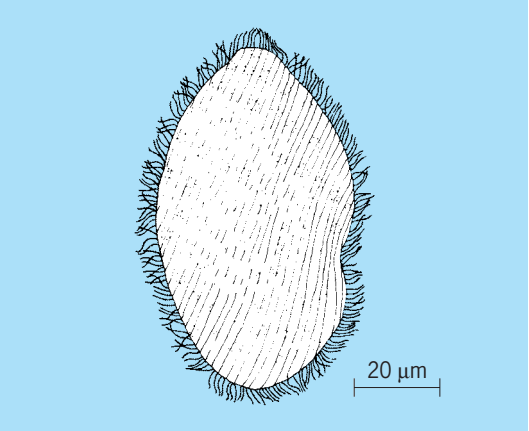

Fig. 1. ***Opaline ranarum*, an opalinid; not considered a ciliate by many protozoologists.**

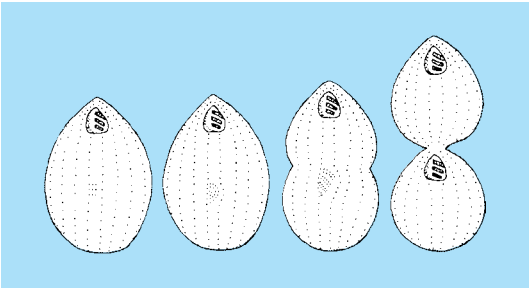

Fig. 2. Binary fission and stomatogenesis in ciliate *Tetrahymena*. Argentophilic basal bodies form part of silverline system and infraciliature.

fission, and the resulting filial products repeat the process. Some commensal or parasitic forms have a more complicated life history, but it is never as involved as the malarial species of the subphylum Sporozoa. Some ciliates, including free-living species, have a cystic stage in their cycle. As in other kinds of Protozoa this stage often serves as a protective phase during adverse environmental conditions, such as desiccation or lack of food. It also may be important in distribution, and thus possibly in preservation, of the species. *See* EUCOCCIDA.

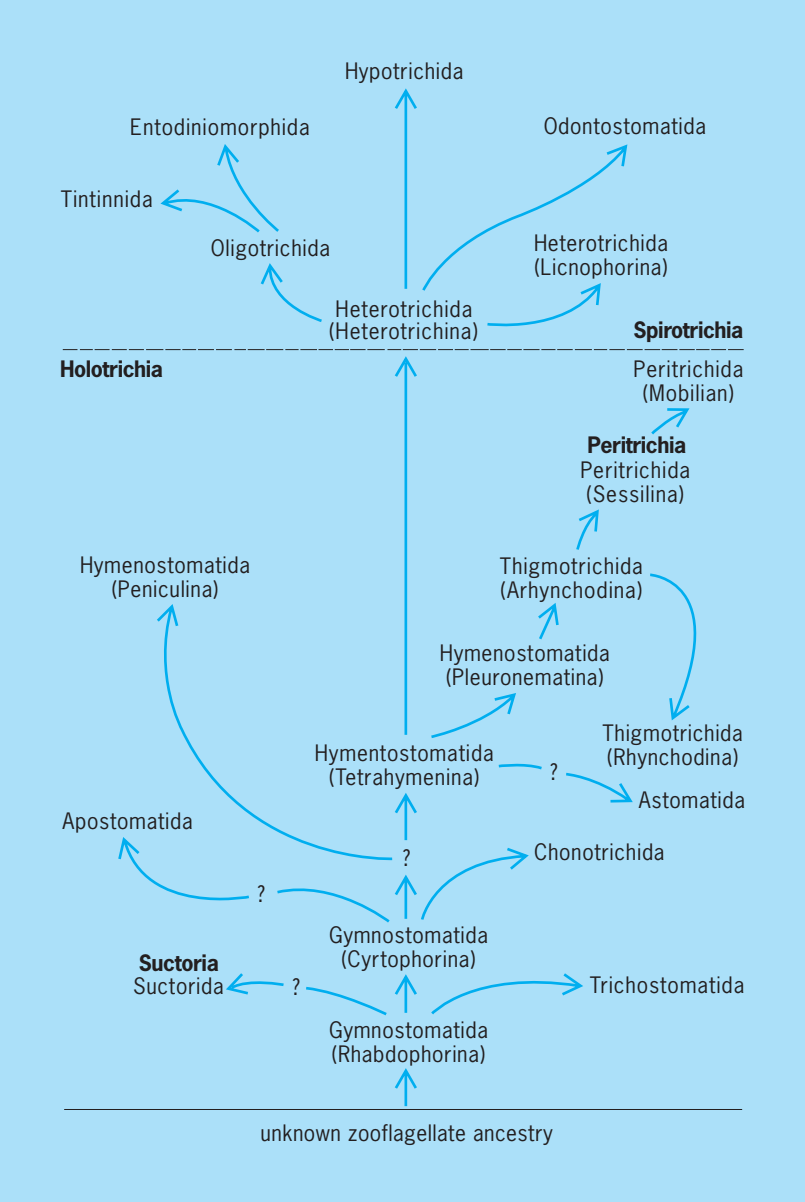

Fig. 3. Phylogenetic tree showing orders of the Ciliophora (names in boldface are of subclasses, and in parentheses of suborders).

Six major characteristics aid in distinguishing the Ciliophora from other protozoan groups. Not all of these are entirely unique, but when taken together they are definitely distinctive of ciliates.

Mouth. Most Ciliophora possess a true mouth or cytostome often associated with a buccal cavity containing compound ciliary organelles. However, some ciliates are completely astomatous, that is, mouthless. Nutrition is heterotrophic in ciliates. Some species are solely or primarily herbivorous since they eat bacteria or algae; some are carnivorous, eating other Protozoa or occasionally small aquatic invertebrate Metazoa; some are omnivorous; and still others live a saprozoic, or saprobic, existence, obtaining all their nutritive needs from materials dissolved in the medium. The process of stomatogenesis, or new mouth formation, which must take place before the completion of binary fission, is a morphogenetic phenomenon of great interest to students of Protozoa. Various modes of stomatogenesis exist in different ciliate groups. Their study helps in postulation of phylogenetic and evolutionary interrelationships within this subphylum of highly differentiated unicellular animals, as well as in throwing light on the important and not yet solved biological process of cell division and its control.

Ciliation. The Ciliophora possess simple cilia or compound ciliary organelles, often in abundance, in at least one stage of their life cycle. Morphologically, cilia are relatively short and slender hairlike structures, whose ultrastructure is known, from electron microscope studies, to be composed of nine peripheral and two central fibrils. Compound ciliary organelles include the undulating membrane, which is essentially a single sheet of fused cilia; the membranelle, a more or less fused block of cilia, typically two to four rows wide and of greater length; and the cirrus, a tuft of closely opposed cilia which is approximately circular in cross section near its base or proximal end but tapered or even fimbriate at its distal end. Functionally, cilia and cirri serve principally as locomotor organelles. Membranes and membranelles are characteristically associated with the mouth or buccal areas and serve to bring food into the oral opening, although they sometimes aid in locomotion as well. Membranelles are typically arranged as an adoral zone. *See* CILIA AND FLAGELLA.

Infraciliature. The presence, without exception, of some sort of infraciliature located at a subpellicular level in the cortex is another characteristic of Ciliophora. The infraciliature consists essentially of basal bodies, or kinetosomes, associated with cilia and ciliary organelles at their bases, plus certain more or less interconnecting fibrils. The basal bodies are argentophilic, that is, they readily stain with silver stains, and are the most important components of

the so-called silverline system present in all ciliates (**Fig. 2**). Various argentophilic fibrils, the contractile vacuole pores, and the cell anus (cytoproct or cytopyge) are included in this system. The controversial neuromotor apparatus of ciliates, best known in hypotrichs, is not well understood. Anatomically, it appears to embrace certain parts of the silverline system, although parts of it are definitely not argentophilic. Much remains to be learned of its function.

Nuclear apparatus. Ciliophora possess two kinds of nuclei, and at least one of each is usually present. The smaller, or micronucleus, contains recognizable chromosomes and behaves much as the single nucleus in cells of metazoan organisms. Its mitotic divisions, however, appear to be acentric (without centrioles). The larger, or macronucleus, does not show mitosis but is considered indispensable in controlling metabolic functions, and is recognized as having genic control over all phenotypic characteristics of ciliates. Micronuclei are typically spherical or ovoid and show little variety in shape. Macronuclei vary in different species from compact bodies to C- or E-shaped forms, or from nodulated chains, which look like a string of tiny linked sausages, to bizarre dendritic configurations. Both nuclei divide in fission, but in sexual phenomena new macronuclei arise from what may be considered micronuclear origin. *See* GENETICS.

Fission. Ciliophora exhibit a type of binary fission commonly known as transverse division. In ciliates the splitting results in two filial organisms, the anterior or proter and the posterior or opisthe which, geometrically speaking, show homothety with respect to identical structures possessed by each (Fig. 2). Thus, homothetogenic is both a broader and more exact descriptive term. Indeed, in a number of groups the fission is more nearly longitudinal than transverse. Some ciliates show a budding type of asexual reproduction; some regularly reproduce within a cystic membrane.

Reproduction. Ciliophora lack true sexual reproduction. Ciliates do not show syngamy, with fusion of free gametes. Processes such as conjugation are considered to be sexual phenomena, since meiosis and chromosome recombination are involved, but not sexual reproduction. In addition to conjugation, certain ciliates exhibit forms of sexual phenomena known as autogamy and cytogamy. *See* REPRODUCTION (ANIMAL).

Evolution of ciliates. Ciliated protozoa, by virtue of their universal possession of an infraciliature and relatable patterns of somatic and buccal ciliature, appear to be the most taxonomically compact group in the phylum Protozoa. They also seem to lend themselves most readily to interpretation of their own phylogenetic interrelationships, that is, their own evolutionary history from the conjectured time of the first recognizable ciliate on the phylogenetic tree of the animal kingdom. Handicaps to study include lack of a usable fossil record, the usually microscopic size, and the very unicellularity which places all complexities of organization of necessity at a subcellular level. However, the student of the ciliated protozoa may still recognize, with a certain amount of reasonable, if indirect, supporting evidence, certain evolutionary lines or trends among the several orders. Most fruitful lines of such investigation have involved comparative study of life cycles, infraciliature, mouthparts, and the morphogenetic phenomenon of stomatogenesis. A possible genealogical tree of Ciliophora is given in **Fig. 3**. The change in somatic and buccal ciliature in the evolution of ciliates is emphasized in three present-day forms which, allegedly, are representative of the most primitive, intermediate, and most highly evolved orders (**Fig. 4**). *See* PROTOZOA.

John O. Corliss

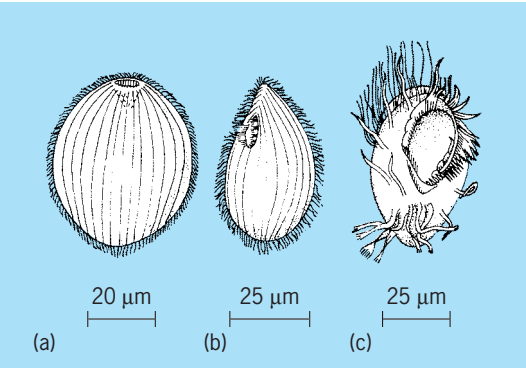

Fig. 4. Ciliates (not drawn to same scale). (*a*) Primitive form. *Holophyra*. (*b*) Intermediate form, *Tetrahymena*. (*c*) Highly complex form, *Diophrys*.

Bibliography. T.-T. Chen (ed.), *Research in Protozoology*, 2 vols., 1967; J. O. Corliss, *The Ciliated Protozoa: Characterization, Classification, and Guide to the Literature*, 1961; J. O. Corliss, The value of ontogenetic data in reconstructing protozoan phylogenies, *Trans. Amer. Microsc. Soc.*, 87:1–20, 1968; B. M. Honigberg et al., A revised classification of the phylum Protozoa, *J. Protozool.*, 11:7–20, 1964; R. R. Kudo, *Protozoology*, 5th ed., 1966; D. L. Mackinnon and R. S. J. Hawes, *An Introduction to the Study of Protozoa*, 1961; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; D. R. Pitelka, *Electron-Microscopic Structure of Protozoa*, 1963.

Cinchona

A genus of trees belonging to the madder family (Rubiaceae), occurring indigenously in the cool, cloud forests of the Andes from Colombia to Peru. Many species have been described, most of which may be variants of *Cinchona pubescens* or *C. officinalis*. The bark contains several alkaloids, the most important of which is quinine. This bitter substance is the most specific drug used in the treatment of malaria. The great demand for quinine and the wasteful methods used in collecting the materials threatened extinction of the plants; therefore cultivation was begun. Now there are extensive cinchona plantations in India, Java, Sri Lanka, Australia, and Jamaica. *See* QUININE; RUBIALES.

Perry D. Strausbaugh; Earl L. Core

Cinematography

The process of producing the illusion of a moving picture. Cinematography includes two phases: taking the picture with a camera and showing the picture with a projector. The camera captures the action by taking a series of still pictures at regular intervals; the projector flashes these pictures on a screen at the same frequency, thus producing an image on the screen that appears to move. This illusion is possible because of the persistence of vision of the human eye. The still pictures appear on the screen many times a second, and although the screen is dark equally as long as it is lighted by the projected image, they do not seem to be a series of pictures but appear to the viewer to be one continuous picture.

Traditional Cinematography

This section discusses motion picture cameras, motion picture projectors, and wide-screen processes. For related information *see* CAMERA; MAGNETIC RECORDING; OPTICAL PROJECTION SYSTEMS; OPTICAL RECORDING; PHOTOGRAPHIC MATERIALS; PHOTOGRAPHY; STEREOPHONIC SOUND.

Cameras. Still photographs are taken by motion picture cameras (movie cameras) at the rate of 24 per second, even faster at times. The photographs are sharp and clear, and if they were superimposed on one another would be found to be extremely uniform with respect to position. This last feature is essential in order to have the image appear steady on the screen. A motion picture camera consists of five basic parts: lens, shutter, gate, film chamber, and pull-down mechanism.

Lens. The lens on a motion picture camera collects the light from the scene being photographed and focuses it on the photographic emulsion in the camera aperture. Lenses vary in focal length. The focal lengths most often used on 35mm cameras are between 30 and 50 mm. Other lenses in common use vary from wide-angle (14.5–25 mm), through long-focal-length (60–250 mm), to telephoto (300–1000 mm). Lenses of shorter focal length have a wider angle of view; those of longer focal length give a larger image size on the film. Lenses for the 16mm camera which give the same angle of view and the same relative image size are approximately one-half the focal length of those used on a 35mm camera. Variable-focal-length lenses (zoom lenses) are useful to cinematographers and directors and are used extensively not only for filming sporting events but also in making feature films, documentaries, and television films. *See* FOCAL LENGTH; TELEPHOTO LENS; ZOOM LENS.

To protect the lens from extraneous light and prevent lens flare, a lens shade must be used. On a motion picture camera this shade is called a matte box. It completely surrounds the lens and acts as a lens shade, but it also has slots to hold various types of filters (such as diffusion, fog, color correcting, and polarizing), and optically flat glass for use when needed. These items are used by cinematographers for the many effects required in making a professional motion picture.

Lenses for motion picture cameras have focusing mounts and can be focused on objects as close as 2 or 3 ft (0.6 to 0.9 m). They are fastened to the camera with threaded or bayonet mounts and can be changed when a lens of a different focal length is needed. Although some older cameras have turrets on the front which hold three or four lenses at one time, making it a simple matter to change from one lens to another, all modern professional cameras have solid fronts because of the size and weight of the zoom lens, and because the faster lenses now used require closer tolerances in the positional accuracy of the lens mount relative to the film plane.

The lens has an iris diaphragm which can be opened and closed to control the amount of light reaching the film, thus controlling the exposure. The *f*-number is a ratio of the focal length of the lens to the diameter of the lens opening, and a lens with a given *f*-number will always allow the same amount of light per unit area to strike the film regardless of the focal length of the lens. The *T*-system used in cinematography takes into account all of the light losses encountered in a lens, such as the number of reflecting surfaces inside the lens and the various types of coating on the surfaces. *T*-stops are determined by the amount of light transmitted through the lens, while *f*-stops are determined by a mathematical formula of the focal length of the lens divided by the diameter of the iris diaphragm opening. Since the *T*-stops system is more accurate as far as light transmission is concerned, this is usually used for exposure, while the *f*-number is used to determine the depth of field, that is, a range of subject distances between which all objects will be in sharp focus on the film. *See* LENS (OPTICS).

Shutter. The shutter in a motion picture camera rotates and exposes the film according to its angle of opening. When the shutter is open to its widest angle, the film receives its maximum exposure. Many cameras have variable shutters; that is, the angle of opening can be changed while the camera is running. Such shutter openings usually vary from 180° when the shutter is wide open to 0° when it is closed. The shutter in a motion picture camera is almost always located between the lens and the film.

Gate. The aperture called the gate is the passageway through which the film is channeled while it is being exposed. It consists of the aperture plate, which is in front of the film and masks the frame, or picture; the pressure plate, which is in back of the film and holds it firmly against the aperture plate; and the edge guides, which keep the film stable laterally.

Film chamber. This holds the unexposed film at one end and collects the exposed film at the other. Although some motion picture cameras have an interior film chamber, it is usually a separate piece of equipment called the film magazine. Magazines come in 400-ft (120-m) and 1000-ft (300-m) capacities. One thousand feet of 35mm film, exposed at the normal speed of 24 frames per second, and with a four-perforation pull-down, lasts slightly more than

11 min. The magazines are loaded with film in a darkroom or in a lightproof changing bag and are fitted on the camera as needed, making camera loading easy and rapid.

Pull-down mechanism. The part of the pull-down mechanism called the intermittent movement pulls the film down through the gate of the camera one frame at a time; in a conventional 35mm camera, each frame is four perforations high. Some widescreen cameras have frames five or six perforations high. The pull-down claw engages the perforations and pulls the film down into place to be exposed. At the bottom of its stroke, the claw remains stationary for a moment to position the film and then disengages itself and returns to the upper portion of the stroke to pull another frame into place. During the time the claw is returning to the top of its stroke, the film is stationary and the shutter opens and exposes the film. Most modern cameras have registration pins which engage the film while the picture is being taken. This ensures that the film is perfectly still and is in exactly the same position as each frame is being exposed. This feature is especially important in special-effects photography where independently photographed images are superimposed.

Mobile cameras. When sound is being recorded, cameras must be absolutely quiet. In the early days of sound, cameras were put in enclosures called blimps. Almost all modern cameras are self-blimped, which enhances their portability. The Mitchell BNC (Blimped News Camera), introduced in 1935, was the first such camera (**Fig. 1**), and remained the worldwide industry standard for 35 years. Although this camera was portable, at approximately 125 lb (55 kg), fully loaded, it was not hand-holdable. In 1968 Arriflex introduced the first hand-held self-blimped camera, the 35 BL. The current model is the 535 (**Fig. 2***a*). In 1972 Panavision introduced the lightweight, silent Panaflex camera, the current version of which is the Millennium (Fig. 2*b*); and in 1979 the silent Moviecam camera was introduced, the current model of which is the lightweight SL (Fig. 2*c*). These later cameras range in weight from 15 to 30 lb (7 to 14 kg), including the lens and a 400-ft (120-m) load of film.

These smaller and lighter cameras have enabled production companies to achieve more mobility and to shoot many pictures on actual locations away from studios, using hotels, cafes, and even private homes in small towns as sets to give the pictures more authenticity. Most feature films, as well as those made for television, are made by independent producers who take advantage of these realistic backgrounds and at the same time save money on stage rentals.

When they go on location, most companies use mobile units which are complete with all the equipment necessary to make a professional motion picture. These vehicles can carry two or three complete camera packages, two or more sound outfits which have been reduced in size even more than the cameras, grip equipment which is used for construction where necessary, camera supports and dollies, and lightweight lamps and cables which feed from a gasoline-powered generator built into the vehicle. Some of these conveyances include space for makeup and wardrobe, although these two departments are usually housed in a separate unit.

Motors. Up until 1970 most sound cameras were driven by 220-VAC (volts alternating current) three-phase synchronous motors, while the sound film, 35mm perforated iron oxide-coated film running over a sprocket, was driven by a similar synchronous motor. After 1970, however, almost all professional

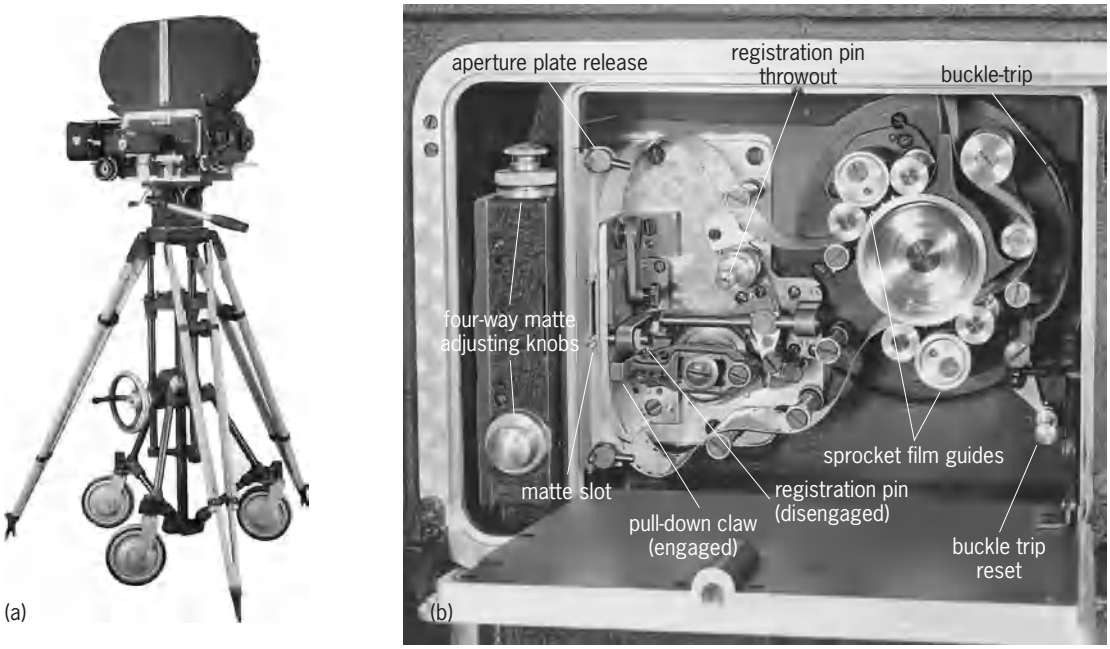

Fig. 1. Mitchell studio camera (BNC). (*a*) Overall view. (*b*) Film-moving mechanism. (*Mitchell Camera Corp.*)

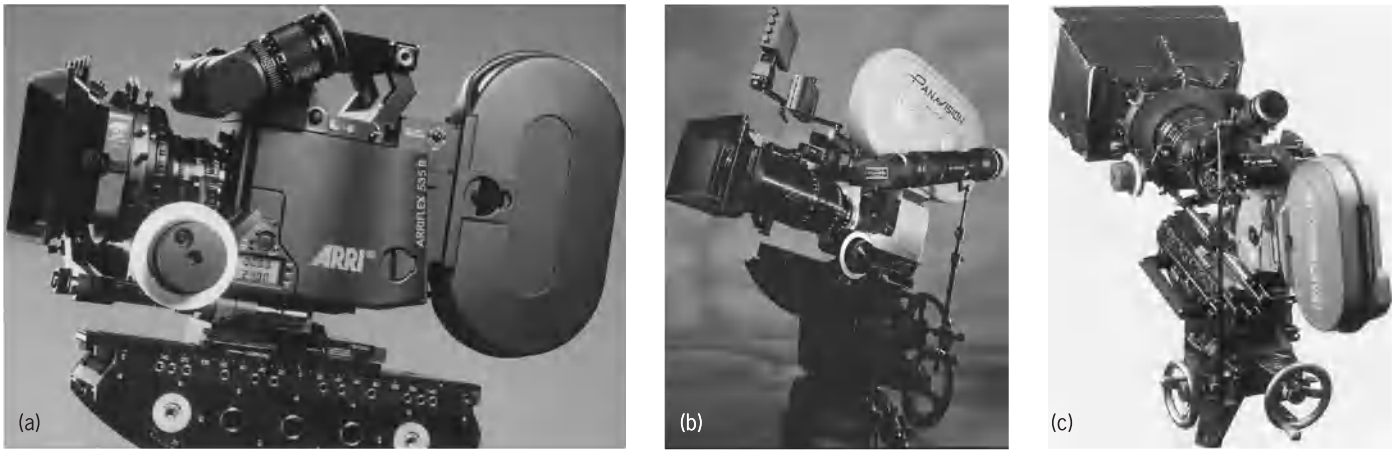

Fig. 2. Hand-held silent cameras. (*a*) Arriflex 535B with 400-ft (120-m) magazine (*Arnold and Richter, Gmbh*). (*b*) Panaflex Millennium camera system with 1000-ft (300-m) magazine and zoom lens (*Panavision*). (*c*) Moviecam SL camera with 400-ft (120-m) magazine and zoom lens (*Moviecam*).

35mm productions in the United States, and many 16mm sound films as well, changed to a new type of motor for driving the camera and a new method of recording synchronous sound, both of which have many advantages over the older system. The motor used on the camera is a direct-current motor, which runs off batteries (12 to 36 V) at a synchronous speed. The speed of the motor is controlled by a crystal oscillator which is so precise that the speed of the motor varies less than one-half a frame of picture in 1000 ft (300 m) of film. All modern cameras (16mm as well as 35mm) use these "crystal motors."

Sound. All sound is recorded on sprocketless 0.25-in. (6-mm) tape. Since tape is an elastic medium, some sort of "sync" pulse must be put on the tape in order to synchronize the recorded sound with the picture when they are put together in a "married" print. The sync pulse is usually an inaudible 1-V, 60-Hz (in Europe, 50-Hz) pulse put on the tape at the time of recording. This pulse can be one generated at the camera and sent to the sound recorder, which is exactly 60 Hz when the camera is running precisely 24 frames/s; or a similar pulse from a crystal oscillator either built into or attached to the tape recorder on which the sound is being recorded. The latter method is the one most commonly used on professional motion picture productions. It is the only system in which there is no physical or electrical connection between sound and camera. Without this connection the camera is completely free; it can be in a separate vehicle, car, boat, or aircraft, while the sound-recording equipment is in another. The only thing required to make a synchronous sound picture with this system is that a sync stick, or clap stick, be used so that sound can have an audible point of reference and the camera can have a visible point of reference.

After the sound is recorded on the 0.25-in. (6-mm) tape, it is put through a resolver, which transfers the sound from the tape to 16mm or 35mm perforated iron oxide–coated film. The resolver corrects for the aforementioned variations in sound-track recording by using the sync pulse as a reference. With the sound on the perforated film it can be precisely synchronized with the picture, using sync-stick marks when the picture is printed for release.

There are two types of sound tracks, optical and magnetic, on release prints, the prints seen by the audience. All 35mm and practically all 16mm release prints have optical sound tracks. A 16mm film may have a magnetic sound track if only one or two prints are to be made, but this is seldom the case. An optical sound track is printed photographically on the film; a magnetic sound track is printed in the same manner as on a tape recorder, after the film is coated with an iron oxide stripe in the area reserved for the sound track. All Super 8 films have magnetic sound tracks.

Until the late 1990s, all 65mm–70mm films, those shot on 65mm film in the camera and printed on 70mm film for release, had magnetic sound tracks. The 70mm prints have six-track sound, one track on each side inside the perforations and two tracks on each side outside the perforations. The space for the tracks outside the perforations is made possible by the additional width of 2.5 mm on each side.

Until the mid-1990s, all sound tracks, both optical and magnetic, were recorded in analog fashion. Developments in the technology, however, have enabled optical sound tracks to be recorded digitally as well, permitting optical sound tracks to provide sound quality comparable to a digital compact disk. For this reason, and for cost savings, the digital-optical sound track has rapidly replaced magnetic-stripe analog sound tracks on 70mm release prints. *See* COMPACT DISK.

Slate. In any motion picture a slate with all pertinent information is of utmost importance (**Fig. 3**). The slate is fastened to the sync stick, and the information is photographed at the beginning of each take. The slate has places for all information relating to the production, such as the name of the picture and its production number for identification at the laboratory, where the film is processed, and the name of the director and the camera operator. In addition, there are spaces for the scene number and the take number, and places to indicate whether the

Fig. 3. Time-code slate. (*Denecke Inc.*)

shot is an interior or exterior, a day or night shot, or sound or MOS (silent); MOS comes from the early European director's pronunciation of "mit-out sound." The term "MOS" is still used extensively throughout the motion picture industry instead of the expression "silent."

Time code. Current practice makes increasing use of some method of coding the film and sound track to facilitate postproduction, especially by permitting transfer to videotape for highly automated editing. Methods have been developed for optically printing time code along the edge of the film beside every frame. The code is machine-readable, and is also human-readable in some systems. A method has also been introduced for printing the human-readable key number, which normally appears as a latent image along the edge of the film, as a machine-readable latent-image bar code. All these schemes facilitate the task of conforming the edit decisions, made in the video domain, back to film for final cutting and assembly. A typical slate (Fig. 3) displays a continuously running time code that is simultaneously recorded as a synchronizing track on the sound recorder. When the clapstick closes, the displayed time stops, providing a visual reference for automatically synchronizing picture and sound in postproduction.

Camera support equipment. A number of devices have been developed to take advantage of the increased mobility of cameras. Their purpose is to give the camera a smooth fluid motion to follow the cinematic action in ever more difficult situations and in ever more ingenious ways. All of these devices are designed to move in such a manner that the film viewer is drawn into the illusion of the scene. A bumpy or jerky camera motion would make the viewer aware of the camera's presence and destroy the illusion.

Crab dolly. One of the first devices developed to permit camera movement was the crab dolly (**Fig. 4**). This dolly can be steered in the usual manner by turning the wheels for curves, and it can be put in a mode where all four wheels turn in the same direction so that it can go forward and backward and also crab across the stage sideways. The camera can be raised or lowered while the dolly is in motion, making it possible to get interesting and creative camera movements.

Cranes and boom arms. A crane (**Fig. 5**) is a larger and more versatile version of the crab dolly. Available in a wide range of sizes, some cranes are slightly larger than a dolly, similar in configuration to a dolly equipped with a boom arm; others extend to 30 ft (9 m) in the air. Cranes are used extensively to achieve large sweeping moves for dramatic effects.

Camera cars. Trucks specially fitted to provide a smooth-moving platform for motion picture cameras (**Fig. 6**) are extremely useful in photographing fast-moving action such as a car chase. These vehicles have become extremely sophisticated. They can support a camera smoothly at highway speeds, in conventional or crab mode, on a boom arm with battery or generator power to satisfy lighting needs.

Steadicam. This is a unique device to support cameras in a mobile situation (**Fig. 7**). It provides a stable platform for a handheld camera supported by the camera operator by means of a vest and supporting spring arm. The mass of the camera and its supporting structure is balanced about a gimbal. The operator guides the camera by gently grasping the system at the gimbal, and is able to walk, run, or go up or down stairs, and still provide totally smooth camera motion.

Film. The film used in the standard motion-picture camera is 35mm (wide) and is perforated along both edges. There are 64 perforations (or sprocket holes) to each foot of film.

Fig. 4. Chapman/PeeWee crab dolly. (*Chapman/Leonard Inc.*)

Fig. 5. Louma crane. (*Panavision Remote Systems*)

Motion-picture film consists of a cellulose acetate base approximately 0.006 in. (0.15 mm) thick and coated with a light-sensitive emulsion. In color film, several layers of emulsion are applied; each of three of the layers is sensitive to one of the three primary colors of light: blue, green, and red. The film is made in large rolls about 54 in. (1.37 m) wide and as long as several thousand feet. It is slit into 35mm strips, perforated, and packed in lightproof bags and cans in rolls 100, 200, 400, and 1000 ft (30, 60, 120, and 300 m) in length.

Smaller widths such as 16 and 8 mm are also made from these wide rolls if they are of the same emulsion type. These sizes are seldom used for feature films but are used extensively by professionals in the audiovisual field and for news and low-budget films for television. The kinds of motion picture film are numerous and varied to meet the needs of a many-faceted industry: black and white, color, high and low speed, fine and coarse grain, high and low contrast, and wide and narrow exposure latitude. These film characteristics describe a particular type of film emulsion.

Fig. 6. Camera car. (*Shotmaker Co.*)

Fig. 7. Steadicam. (*Cinema Products Corp.*)

Film types. There are two different types of film: negative film, from which a print is made in order to see the original subject in its true likeness (for example, Kodacolor, used in amateur photography), and reversal film, in which a negative is first formed in the original film and from this a positive is formed in the same piece of film (for example, Kodachrome, a film familiar to amateurs). In the negative film two pieces of film are necessary to get a picture that can be projected; in the reversal film only one piece of film is required, making it a much less expensive process. In Super 8, as well as regular 8, the reversal film is generally used because of the cost factor. In 16mm and 35mm professional work, the original, that is, the film used in the camera, is never used for projection. A work print is made in order to edit the film and the original negative is assembled to conform to the

edited work print. Master positives are then struck from the original negative, and the master positives are used, in turn, to produce duplicate negatives. The release prints are made from the duplicate negatives, and during this process the sound tracks are added. The original film is all-important and is kept in a vault at the laboratory; it is used only for making prints.

Aspect ratio. The picture taken in standard 35mm motion picture cameras is masked by the aperture plate, which varies according to the camera. The silent aperture, that is, the aperture used on cameras that do not leave room on the film for a sound track, is almost the entire width of the film between perforations, slightly less than 1 in. (25 mm) and approximately 0.75 in. (19 mm) high. This gives an aspect ratio (the ratio of its width to its height) of 1.33:1 (4/3). Sound apertures are masked down to a width of 0.868 in. (22 mm), to allow room for the sound track, and a height of 0.631 in. (16 mm), which gives an aspect ratio of 1.37:1.

Projectors. The projector system in a modern motion picture theater has five main assemblies: the optical sound head, which reproduces optical (photographically recorded) sound; the projector head, which projects the image onto the screen; the lamphouse, which furnishes the illumination for the picture; the shutter; and the platter system, which feeds the film through the projector head. On some projectors, there is a second sound head for the reproduction of magnetic (magnetically recorded) sound.

With the development of digital optical sound, most projectors are now equipped with a charge-coupled-device (CCD) line array to read the digital track. The analog optical track, which is still present as a backup system, is now read by a light-emitting diode (LED) array in lieu of the traditional exciter lamp. An alternate dual-system approach has a time-code track on the film that is read to synchronize the picture with the sound track played back from a compact disk machine. *See* CHARGE-COUPLED DEVICES; LIGHT-EMITTING DIODE.

Optical sound head. This consists of an exciter lamp and a photoelectric cell. The optical sound track is 20 frames ahead of the picture on 35mm film (26 frames ahead on 16mm film). If the film has a magnetic sound track in addition to the optical one, it is first threaded over the magnetic sound head, which is above the projector head (**Fig. 8**). The magnetic sound track is 28 frames behind the picture on 35mm film and 23 frames behind on 70mm film (26 frames ahead on 16mm film, the same as optical sound).

Projector head. This consists of the projection lens and the projector intermittent. The projection lens is located in the front of the projector head between the aperture and the screen. The film is put in the projector aperture upside down and with the sound track on the right. The image is inverted by the lens and appears right side up on the screen. The sound track, if it were visible, would appear on the left side of the screen. The focal length of the lens is determined by the size of the image desired on the screen and the distance between projector and screen. The lens usually has a fixed speed of *f*/2 and has no iris diaphragm.

The projector intermittent uses a Geneva movement to pull the film through the gate. This movement drives a 16-toothed sprocket one-quarter of a revolution at a time. With each movement a new frame is pulled into place. The film is wrapped around this sprocket wheel and its perforations on both sides engage four or five teeth on the wheel at all times. This distributes the stress caused by the acceleration and deceleration over many perforations and results in less wear and tear on the film, substantially increasing its life. A framing device is incorporated in the projector head so that the frame can be moved up or down to position the picture correctly and put the frame lines at the top and bottom of the screen.

Lamphouse. All modern projectors employ xenon arc lamps in place of the previously used carbon arc lamps. The xenon arc lamps form a small plasma arc in a sealed glass bulb. A concave reflector behind the lamp collects the light and concentrates it on the film in the projector aperture. Depending on the size of the screen, the lamps generally range from 1 to 7 kW in power input. *See* ARC LAMP.

Platter system. Most modern theaters use a platter system (**Fig. 9**) rather than the conventional 2000-ft (600-m) reels used in older theaters. Conventional reels are labor-intensive and require two projectors to achieve a smooth transition (changeover) between subsequent reels of a motion picture. In the platter system, only one projector is needed. All reels of a motion picture are spliced together and loaded onto one of the platters with the start of the picture at the center of the reel. The film is then threaded from the center of the reel, through the projector, and onto a second platter. When the show is over, it is

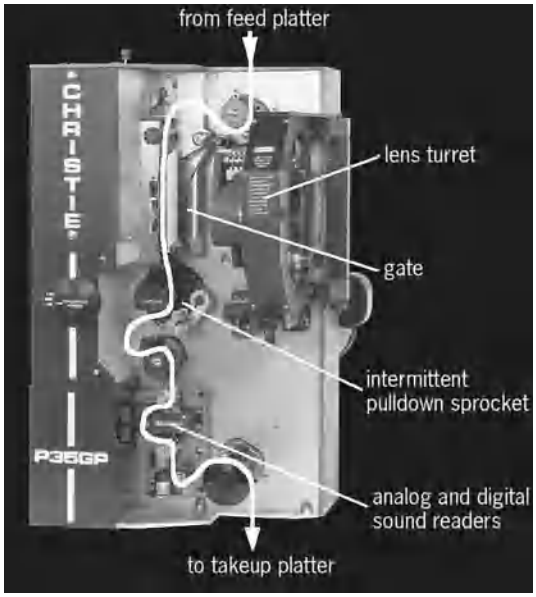

Fig. 8. Motion picture projector. The white ribbon shows the path of film from feed platter down through the projector head gate, over the intermittent pulldown sprocket, over analog and digital sound reader heads, and on the way to the take-up platter. (*Christie Inc.*)

Fig. 9. Nonrewind platter system. (*Kinotone*)

only necessary to rethread the projector, but starting from the second platter.

Shutter. A projector has a two-bladed shutter located between the lamphouse and the gate in the projector head. The shutter cuts off the light to the screen while the film is being moved down from one frame to the next and opens when the new frame has been moved into place. It then closes and opens a second time before another frame has been brought into position. This makes two exposures per frame, or 48 light impulses hitting the screen in 1 s, even though the picture was shot at 24 frames per second. This reduces the flickering effect that would be noticeable if only 24 light impulses per second were flashed on the screen.

Universal types. Universal projectors can show 35mm prints, both normal and anamorphic, with optical sound, four-track magnetic sound, or both. They can also show 70mm normal and anamorphic with six-track magnetic sound. The 35mm sprockets are between and below the 70mm sprockets, and either film can be run by selecting the correct lens and threading the film through the gate and through the appropriate sound head. *See* OPTICAL PROJECTION SYSTEMS.

Wide-screen processes. Until 1952 the motion picture industry as a whole used film and equipment that was well standardized. Film was 35 mm wide, the aspect ratio was 1.33:1, and the camera had a four-perforation pull-down mechanism. These standards had been in effect since early in the twentieth century. With the advent of sound in 1926, standard camera speed became established at 24 frames per second, and most sound tracks were reproduced optically. Beginning in 1952, however, many new processes were introduced which departed from these standards and utilized a multiplicity of film sizes, aspect ratios, and types of sound tracks. Some processes which employ 35mm film use a means of getting more picture on the film. The 1.85:1 ratio has practically become a standard when referring to wide screen. Multitrack digital-optical sound has become fairly common.

Cinerama. The first commercially successful wide-screen process was Cinerama, developed by F. Waller. First shown in 1952, the system employed three separate, matched film strips projected by three separate projectors. The projected image height was equivalent to six perforations, and the width the equivalent of three full 35-mm apertures. The deeply curved screen consisted of 1200 overlapping strips, arranged in such a manner as to minimize lateral reflections. A fourth 35mm film strip carried seven sound tracks for full stereophonic reproduction. The complexities of running these projectors in exact synchronism, in addition to jitter problems between the three panels, caused the system to be simplified some years later to a single 70-mm-wide film with multiple sound tracks and a 1.25:1 anamorphic "squeeze." The first picture made in this manner was *Mad Mad Mad Mad World* (1964).

Cinemascope. The single-film approach was anticipated by Cinemascope, developed by twentieth Century Fox, based on earlier work of the physicist H. Chrétien. Special lens adapters were designed that produced a linear 2:1 anamorphic squeeze when attached to conventional lenses. Integral anamorphic lenses were subsequently developed which were simpler to use. A special projector lens was used to unsqueeze the picture (**Fig. 10**). The first feature

Fig. 10. Action of Cinemascope anamorphic lens. (*a*) Image "squeezed" onto the film by an anamorphic lens. (*b*) "Unsqueezed" image as it appears on the screen. (*Twentieth Century Fox*)

filmed in Cinemascope was *The Robe* (1953). The format continues to be used.

VistaVision. This process, developed by Paramount, utilized an eight-perforation–pull-down, 35mm movement, operating from left to right rather than top to bottom. Conventional spherical lenses were used, and the resulting format was 37.7 mm wide by 25 mm high. *White Christmas* (1954) was the first picture produced by this process. The need for special projection equipment eventually caused VistaVision to be discontinued. It is still used, however, for shooting background plates for process photography since it provides a large format that will not degrade as much as a 35mm format during the duplicating process.

Todd-AO. In the mid-1950s B. O'Brien in cooperation with promoter M. Todd developed the Todd-AO wide-screen process. This process used a 70mm print derived directly from a 65mm negative, five perforations high, with six magnetic sound tracks arrayed as previously noted. A series of special lenses was developed to provide from 37 to 128° of angular coverage. The first picture produced in Todd-AO was *Oklahoma* (1955). As originally presented, Todd-AO used a deeply curved screen, but it was subsequently converted to flat-screen projection. Super Panavision, similar to Todd-AO in format, has been widely used. Mitchell 35mm and 65mm cameras have been used for these processes, as well as cameras specially designed and built for the process.

Screen formats. Films for theatrical release have now been standardized in two basic formats in the United States, as contrasted with the plethora of wide-screen and multiscreen formats that appeared and disappeared during the 1950s and 1960s. With rare exceptions all films for theatrical presentation are now shot either (1) in the 2.35:1 aspect ratio, using the 35mm anamorphic system that was originated by Cinemascope but that now exists largely through the use of Panavision anamorphic lenses and those developed by others, or (2) in the 1.85:1 aspect ratio, which is achieved by cropping the standard 35mm frame top and bottom. In Europe, the popular aspect ratio is 1.66:1.

The 1.85:1 format is extremely wasteful of film, and there have been several attempts to introduce a system involving three-perforation pull-down in the camera as well as in the projector to eliminate this waste.

Techniscope. This was a promising format that has now virtually disappeared in the United States. It was developed in the Technicolor Laboratories in Rome, Italy. The process involves shooting the original camera negative with the camera modified to pull down two perforations, rather than four, in a standard 35mm camera. The taking lenses are the normal complement of spherical lenses. In the laboratory the negative is optically printed through an anamorphic lens which stretches the image vertically by a factor of 2:1 so that the print is identical to one which would have resulted from a negative shot in the Cinemascope anamorphic process. A number of films were shot in this process, but it fell into disfavor in the United States because the increase in granularity and the decrease in resolution due to using a negative area one-half that of the standard 35mm size resulted in a film that was not of acceptable quality when compared with the 1.85:1 format or the anamorphic format. However, in Europe and elsewhere Techniscope is still an acceptable and frequently used process.

Super 16. An alternative means of arriving at current theatrical formats, called Super 16, was pioneered by R. Ericson in Sweden. This process involves photographing with a 16mm camera that is modified so that the full width of the negative is photographed, including that area reserved for sound track in the final release print. The negative is then enlarged, usually through liquid-gate optical printing, to a 35mm format with a 1.65:1 aspect ratio.

70mm presentation. Until the introduction of digital-optical sound, there had been a major move back to theatrical presentation in 70mm film. This was motivated by the ability to provide six magnetic sound tracks on the 70mm print. Original photography was typically on 35mm film and blown up to 70mm. Digital-optical sound technology permits full multichannel surround sound to be recorded on a standard 35mm print, obviating the need to produce very expensive 70mm magnetic-striped prints. However, there has been a growing trend in the use of 70mm prints from 65mm negatives for special-venue presentations such as theme parks, studio tours, dynamic motion simulator rides, and 3D (three-dimensional). To meet this growing need, new 65mm cameras have been developed (**Fig. 11**).

Imax. Epitomizing this trend to larger formats, the Imax process films on a 65mm negative moving

Fig. 11. CP-65 65mm camera system. (*Cinema Products Corp.*)

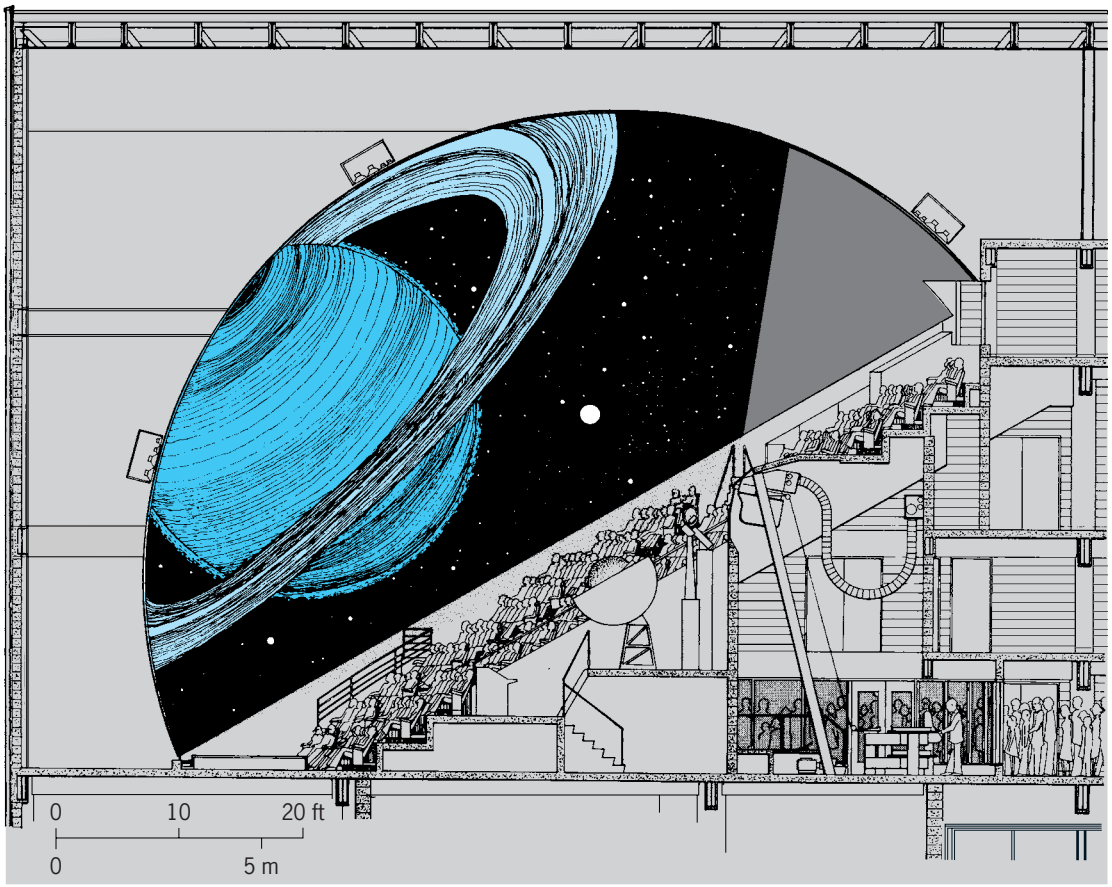

Fig. 12. Cross-sectional view of typical Omnimax theater.

horizontally through the camera, 15 perforations per frame. The resulting aspect ratio is 1.33:1. The 70mm prints are projected on a large flat screen. They can also be photographed and projected with a fisheye lens on a hemispheric dome in a process called Omnimax (**Fig. 12**). The projector is of a unique design that utilizes a rolling loop to rapidly advance the film with no damage and absolute flatness in the film gate. *See* PLANETARIUM.

A less expensive alternative to the Imax format is provided by the Iwerks system of recording eight perforations per frame on 65mm film moving vertically through the camera. This provides essentially the same format, but on a smaller negative area.

Showscan. This process uses the conventional 70mm, five-perforation format, but is filmed and projected at 60 frames per second. This frame rate virtually eliminates all filmic artifacts, such as strobe and flicker, to provide a heightened sense of reality.

Edmund M. DiGiulio

Digital Cinematography

Like other forms of media, cinematography is undergoing a transition from photomechanical to digital processes, which offer greater flexibility, consume fewer resources, offer significant time saving, and allow electronic (versus physical) delivery of products. In cinematography, this transition involves three distinct operations: image capture, image editing, and image display. The complete transition will result in an all-digital workflow for the production of motion pictures. Although the implementation of an all-digital workflow is under way, the use of digital methods in the image editing or "postproduction" processes is far more widespread than in image capture or display.

Image capture. The image-capture process involves filming the scenes that will be edited together to make a movie. In this operation, the switch to digital involves capturing the scenes with a high-definition digital camera instead of traditional cinemagraphic 35mm film cameras.

There are several weaknesses in the traditional approach, such as the high cost of the film and the environmental impact of chemical film processing, which are alleviated by digital image capture. In addition, color negative film cannot be evaluated directly (even after processing), so the creative team cannot get immediate feedback on the success of a filming session; film is easily scratched and susceptible to dust; and in order to use digital editing techniques, film images must be scanned to create digital files. Still, cinematographers have been reluctant to give up film cameras because 35mm film captures a higher-resolution image with greater dynamic range than high-definition digital cameras.

Motion pictures in which digital image capture was used include *Sin City* and *Star Wars: Episode III Revenge of the Sith*. In both of these films, digitally generated backgrounds were extensively used. Thus, the limited resolution of the digital camera did not limit the detail that was evident in the background.

For *Sin City*, Sony HDC-F950 digital cameras were used, capturing red, green, and blue images at 1920 × 1080 pixels of resolution with 10-bit color depth at 24 frames per second. Scenes were viewed on a calibrated monitor immediately after shooting, which benefited the production team by avoiding multiple shoots of some scenes.

Star Wars: Episode III Revenge of the Sith used an all-digital workflow, including image capture. The production cost of $100 million is estimated to be half the amount required for a comparable movie shot on film. The savings came from the increased productivity gained by having immediate feedback during shooting and from doing much of the special-effects work before the shooting took place. This enabled the director to make decisions about which angles and perspectives were needed before (rather than after) the filming of a scene.

Over time, digital image capture should replace film because of the creative flexibility and cost savings it offers. Digital cinema cameras are improving steadily, as are data transport, storage, and compression techniques.

Film digitizing. Scanning rolls of film for digital editing has become a lucrative business because an increasing number of films are doing all the post-production processes digitally. The scanners used for rolls of film were initially developed for converting finished films to DVD format. Ideally, each frame would be scanned at 4K (4100 pixels wide by 3000 pixels high) for optimum quality, but the resulting files were too large and it would take days to scan a feature length film. Scanners now are capable of performing 2K scans in real time and 4K scans at 7.5 frames per second, or about 6 h for a feature-length film, with 4K scanning resolution reserved for critical projects and 2K (2048 pixels wide by 1556 pixels high) scanning is used for most applications.

Image editing. Nearly all special effects are achieved through digital means, making image editing the area with the greatest penetration of digital processes. Traditional film-editing methods involved cutting apart scenes from the developed film reels and taping them back together to form the desired sequence. After the taped-together master was made, it was duplicated with color adjustments to form a new master on a continuous piece of film. This process was extremely time-consuming and not easily reversible.

Digital editing offers significant advantages over film editing. It is faster, more precise, and more readily reedited. It also allows the easy integration of digital special effects into the shot sequences and the individual frames of the film.

Digital editing offers color effects that are not obtainable with photomechanical processes. For example, the film *O Brother Where Art Thou* used a restricted color palette for the background scenery that evoked the Dust Bowl era in the American South. This palette could have been achieved on film with the judicious use of lighting and filters, but it could not have been restricted to background elements only.

Image display. In June 1999, the first two films released for display in digital format were *Star Wars Episode I and An Ideal Husband*. There were two competing projectors in use: the Texas Instruments digital light processor (DLP) projector and Hughes/JVC ILA-12K liquid crystal on silicon (LCOS) projector. The DLP projectors were preferred for their larger color gamut and greater contrast range.

DLP projectors use optical semiconductors, called digital micromirror devices (DMD), to modulate the light and form images. Each DMD is an array of 1.3 million microscopic mirrors, which can be individually pivoted in and out of the projector's light path thousands of times per second. The individual mirrors are spaced 1 micrometer apart, so the device exhibits a high fill factor with good brightness and contrast. Cinema projectors are equipped with three DMD arrays (one for each primary color) to produce an image. Early projectors were limited to 1024 × 768 pixel resolution but current equipment offer 2K pixel resolution, 15-bit color depth, and sufficient brightness for a 75-ft-wide screen. In 2004, Sony released the SXRD 4K projector, which uses a variant of LCD technology. Its improved image resolution (4096 × 2160 pixels) offers a more lifelike image to the viewer.

Standards. The digital cinematography industry is developing standards for file format, compression technology, color display, and encryption needed for interoperability for the rapidly developing projector and hardware industries. Several groups are contributing to this effort, including the Motion Picture Association of America (MPAA) and the Digital Cinema Initiatives (DCI) consortium, a group of seven studios. The Society of Motion Picture and Television Engineering (SMPTE) has a technical committee for digital cinema (DC28) that has recommended specifications for file format, compression, and encryption technologies. The American Society of Cinematographers (ASC) has a technical committee to produce standardized evaluation materials for all aspects of digital cinema.

There has been steady development of encryption and digital rights management strategies for digital cinema. Currently, files are encrypted all the way into the projector, where the display files are extracted. This is of great importance to the industry in thwarting the piracy of intellectual property. It reduces the need for high security on the server since it will only receive and distribute encrypted files.

Anthony Stanton

Bibliography. S. H. Burum (ed.), *American Cinematographers Manual*, 9th ed., American Society of Cinematographers, 2004; J. Careless, The back alleys of Sin City, *Dig. Cinematog.*, 1(1):22–26, 2005; C. Harrison, The evolving digital workflow in cinema, *IS&T Reporter*, vol. 18, no. 6, 2003; M. A.

Hiltzik, Digital cinema take two, *Tech. Rev.*, pp. 36–44, September 2002; M. Hurwitz, Cost effective digital production—*Star Wars: Episode III Revenge of the Sith*, *Dig. Cinematog.*, 1(2):16–19, 2005; T. A. Ohanian and M. E. Phillips, *Digital Filmmaking: The Changing Art and Craft of Motion Pictures*, 2d ed., 2000; D. W. Samuelson, *Hands-on Manual for Cinematographers*, 2d ed., 1998; C. S. Swartz (ed.), *Understanding Digital Cinema: A Professional Handbook*, Focal Press, 2005; D. Viera and M. Viera, *Lighting for Film and Electronic Cinematography*, 2d ed., 2004.

Cinnamon

An evergreen shrub or small tree, *Cinnamomum zeylanicum*, of the laurel family (Lauraceae). A native of Sri Lanka, the plant (see **illus.**) is now in cultivation in southern India, Burma, parts of Malaya, West Indies, and South America. In cultivation the trees are cut back and long, slender suckers grow up from the roots. The bark is removed from these suckers, dried and packaged for shipping. Cinnamon is a very important spice for flavoring foods. It is used in confectionery, gums, incense, dentifrices, and perfumes. Cinnamon oil is used in medicine and as a source of cinnamon extract. *See* MAGNOLIALES; SPICE AND FLAVORING.

Cinnamon (*Cinnamomum zeylanicum*). (*USDA*)

Perry D. Strausbaugh; Earl L. Core

Circle

The curve that is the locus of points in a plane with equal distance (radius) from a fixed point (center) (**Fig. 1**). In elementary mathematics, circle often refers to the finite portion of the plane bounded by a curve (circumference) all points of which are equidistant from a fixed point of the plane, that is, a circular disk. Circles are conic sections and are defined analytically by certain second-degree equations in cartesian coordinates. They were extensively studied by the ancient Greeks, who formulated the famous problem of "squaring the circle," that is, to construct, with compasses and unmarked straightedge only, a square whose area is equal to that of a given circle. It was not until 1882 that this was shown to be impossible, when F. Lindemann proved that the ratio of the length of a circle to its diameter (denoted by π) is not the root of any algebraic equation with integer coefficients. That π is irrational (that is, not the quotient of two integers) was shown by A. M. Legendre in 1794. Numerous approximations of π appeared quite early; for example, 3 (*Book of Kings*); $3^1/_7 > \pi > 3^{10}/_{71}$ (Archimedes); $\sqrt{10}$ (*Ch'ang Hông*, A.D. 78–139); and 3.1416 (Aryabhata, A.D. 476–550). Electronic computers have calculated π to over 10^{12} decimal places. Interesting expressions are

$$\pi = 4(1 - {}^1/_3 + {}^1/_5 - {}^1/_7 + \cdots)$$
$$\pi = 2 \cdot 2(2 \cdot 4/3 \cdot 3)(4 \cdot 6/5 \cdot 5)(6 \cdot 8/7 \cdot 7) \cdots$$

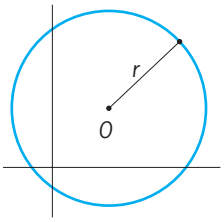

Fig. 1. Circle of radius *r* and center *O*.

The area of a circle (circular disk) with radius r is πr^2; the length (circumference) is $2\pi r$. The area enclosed by a circle is greater than that bounded by any other curve of the same length. If a circular disk can cover each three points of an arbitrary plane set, then the disk can cover the whole set. It is, moreover, the only disk that has this property. *See* ANALYTIC GEOMETRY; CONIC SECTION.

Leonard M. Blumenthal

A great many properties of the circle are in relation to a given triangle. Such, for example, is the nine-point circle theorem. Let the altitudes of any triangle ABC intersect each other in the orthocenter H, and intersect the opposite sides in D, E, and F, respectively. Then the circle DEF passes through the midpoints of each of the six segments AB, BC, CA, AH, BH, CH. It is called the nine-point circle of the triangle ABC, and also of each of the triangles HBC,

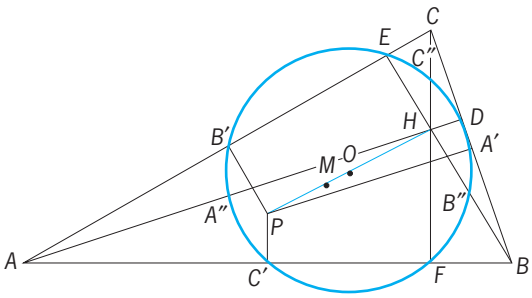

Fig. 2. Nine-point circle and Euler line *PMH*.

HCA, *HAB*, whose orthocenters are *A*, *B*, *C*, respectively. Feuerbach's theorem states that this nine-point circle is tangent both to the inscribed circle and to the three escribed circles of each of these triangles.

The center *O* of the nine-point circle (**Fig. 2**) of any triangle lies midway between the orthocenter *H* and the circumcenter *P* on the so-called Euler line that passes through these points and the centroid *M*. J. Sutherland Frame

Bibliography. J. Arndt and C. Haenel, *Pi Unleashed*, 2001; P. Beckmann, *A History of Pi*, 4th ed., 1977, reprint 1990; D. Blattner, *The Joy of Pi*, 1997; J. L. Coolidge, *Treatise on the Circle and the Sphere*, 1916, reprinted 1971; P. Eynard and J.-P. Lafor, *The Number Pi*, 2004; D. Pedoe, *Circles: A Mathematical View*, rev. ed., 1979, reprint 1995; A. S. Posamentier and I. Lehmann, *Pi: A Biography of the World's Most Mysterious Number*, 2004.

Circuit (electricity)

A general term referring to a system or part of a system of conducting parts and their interconnections through which an electric current is intended to flow. A circuit is made up of active and passive elements or parts and their interconnecting conducting paths. The active elements are the sources of electric energy for the circuit; they may be batteries, direct-current generators, or alternating-current generators. The passive elements are resistors, inductors, and capacitors. The electric circuit is described by a circuit diagram or map showing the active and passive elements and their connecting conducting paths.

Devices with an individual physical identity such as amplifiers, transistors, loudspeakers, and generators, are often represented by equivalent circuits for purposes of analysis. These equivalent circuits are made up of the basic passive and active elements listed above.

Electric circuits are used to transmit power as in high-voltage power lines and transformers or in low-voltage distribution circuits in factories and homes; to convert energy from or to its electrical form as in motors, generators, microphones, loudspeakers, and lamps; to communicate information as in telephone, telegraph, radio, and television systems; to process and store data and make logical decisions as in computers; and to form systems for automatic control of equipment.

Electric circuit theory. This includes the study of all aspects of electric circuits, including analysis, design, and application. In electric circuit theory the fundamental quantities are the potential differences (voltages) in volts between various points, the electric currents in amperes flowing in the several paths, and the parameters in ohms or mhos which describe the passive elements. Other important circuit quantities such as power, energy, and time constants may be calculated from the fundamental variables. For a discussion of these parameters *see* ADMITTANCE; CONDUCTANCE; ELECTRICAL IMPEDANCE; ELECTRICAL RESISTANCE; REACTANCE; SUSCEPTANCE; TIME CONSTANT.

Electric circuit theory is an extensive subject and is often divided into special topics. Division into topics may be made on the basis of how the voltages and currents in the circuit vary with time; examples are direct-current, alternating-current, nonsinusoidal, digital, and transient circuit thoery. Another method of classifying circuits is by the arrangement or configuration of the electric current paths; examples are series circuits, parallel circuits, series-parallel circuits, networks, coupled circuits, open circuits, and short circuits. Circuit theory can also be divided into special topics according to the physical devices forming the circuit, or the application and use of the circuit. Examples are power, communication, electronic, solid-state, integrated, computer, and control circuits. *See* CIRCUIT (ELECTRONICS); NEGATIVE-RESISTANCE CIRCUITS.

Direct-current circuits. In dc circuits the voltages and currents are constant in magnitude and do not vary with time (**Fig. 1**). Sources of direct current are batteries, dc generators, and rectifiers. Resistors are the principal passive element. For a discussion of direct-current circuits *see* DIRECT-CURRENT CIRCUIT THEORY

Magnetic circuits. Magnetic circuits are similar to electric circuits in their analysis and are often included in the general topic of circuit theory. Magnetic circuits are used in electromagnets, relays, magnetic brakes and clutches, computer memory devices, and many other devices. For a detailed treatment. *See* MAGNETIC CIRCUITS.

Fig. 1. Direct current, which is constant in magnitude and does not vary with time.

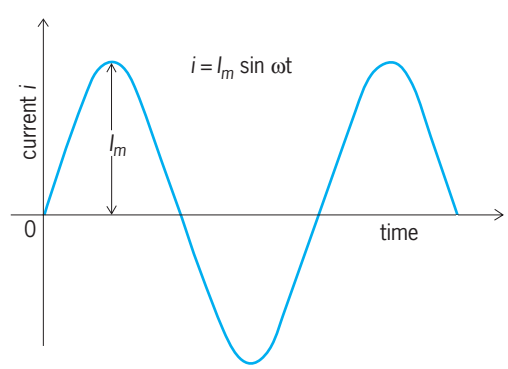

Fig. 2. Alternating current, which periodically reverses direction with time. In this example, the current varies sinusoidally with time.

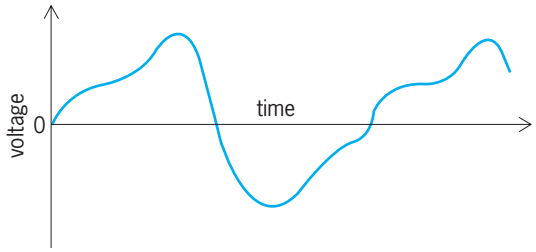

Fig. 3. Nonsinusoidal voltage wave.

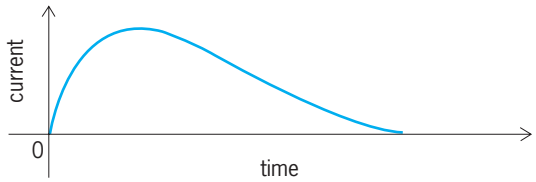

Fig. 4. Transient electric current.

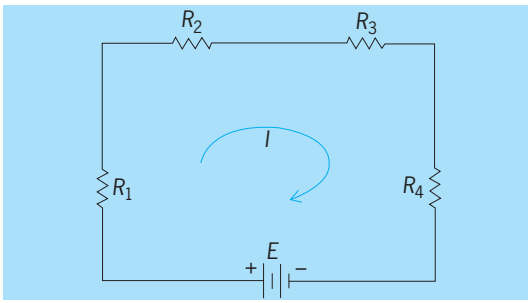

Fig. 5. Series circuit.

Alternating-current circuits. In ac circuits the voltage and current periodically reverse direction with time. The time for one complete variation is known as the period. The number of periods in 1 s is the frequency in cycles per second. A cycle per second was named a hertz (in honor of Heinrich Rudolf Hertz's work on electromagnetic waves).

Most often the term ac circuit refers to sinusoidal variations. For example, the alternating current in **Fig. 2** may be expressed by $i = I_m \sin \omega t$. Sinusoidal sources are ac generators and various types of electronic and solid-state oscillators; passive circuit elements include inductors and capacitors as well as resistors. The analysis of ac circuits requires a study of the phase relations between voltages and currents as well as their magnitudes. Complex numbers are often used for this purpose. For a detailed discussion. *See* ALTERNATING-CURRENT CIRCUIT THEORY.

Nonsinusoidal waveforms. These voltage and current variations vary with time but not sinusoidally (**Fig. 3**). Such nonsinusoidal variations are usually caused by nonlinear devices, such as saturated magnetic circuits, electron tubes, and transistors. Circuits with nonsinusoidal waveforms are analyzed by breaking the waveform into a series ofsinusoidal waves of different frequencies known as a Fourier series. Each frequency component is analyzed by ac circuit techniques. Results are combined by the principle of superposition to give the total response. *See* NONSINUSOIDAL WAVEFORM.

Electric transients. Transient voltage and current variations last for a short length of time and do not repeat continuously (**Fig. 4**). Transients occur when a change is made in the circuit, such as opening or closing a switch, or when a change is made in one of the sources or elements. For a discussion of dc and ac transients. *See* ELECTRIC TRANSIENT.

Series circuits. In a series circuit all the components or elements are connected end to end and carry the same current, as shown in **Fig. 5**. *See* SERIES CIRCUIT.

Parallel circuits. Parallel circuits are connected so that each component of the circuit has the same potential difference (voltage) across its terminals, as shown in **Fig. 6**. *See* PARALLEL CIRCUIT.

Series-parallel circuits. In a series-parallel circuit some of the components or elements are connected in parallel, and one or more of these parallel combinations are in series with other components of the circuit, as shown in **Fig. 7**.

Electric network. This is another term for electric circuit, but it is often reserved for the electric circuit that is more complicated than a simple series or parallel combination. A three-mesh electric network is shown in **Fig. 8**. *See* NETWORK THEORY.

Coupled circuits. A circuit is said to be coupled if two or more parts are related to each other through some common element. The coupling may be by means of a conducting path of resistors or capacitors or by a common magnetic linkage (inductive coupling), as shown in **Fig. 9**. *See* COUPLED CIRCUITS.

Open circuit. An open circuit is a condition in an electric circuit in which there is no path for current flow between two points that are normally connected. *See* OPEN CIRCUIT.

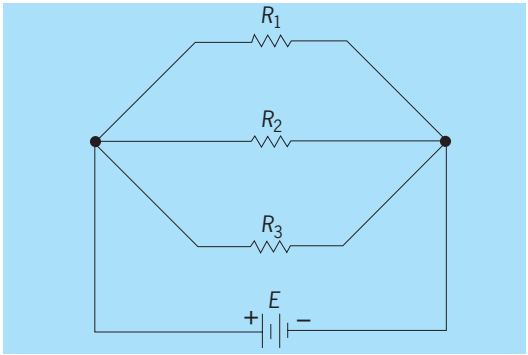

Fig. 6. Parallel circuit.

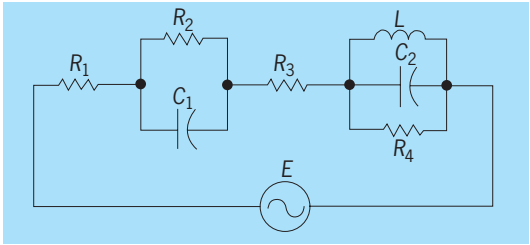

Fig. 7. Series-parallel circuit.

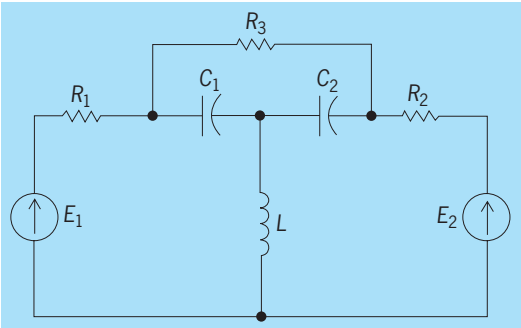

Fig. 8. A three-mesh electric network.

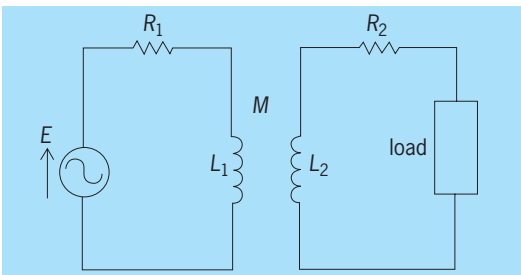

Fig. 9. Inductively coupled circuit.

Short circuit. This term applies to the existence of a zero-impedance or low-impedance path between two points of an electric circuit that are not intended to be connected. *See* SHORT CIRCUIT.

Integrated circuit. In the integrated circuit the entire circuit is contained in a single piece of semiconductor material. Sometimes the term is also applied to circuits made up of deposited thin films on an insulating substrate. *See* INTEGRATED CIRCUITS; SEMICONDUCTOR. Clarence F. Goodheart

Bibliography. N. Balabanian, *Electric Circuits*, 1994; W. Hayt, Jr., and J. E. Kemmerly, *Engineering Circuit Analysis*, 5th ed., 1993; L. P. Huelsman, *Basic Circuit Theory*, 3d ed., 1991; D. E. Johnson, J. L. Milburn, and J. R. Johnson, *Electric Circuit Analysis*, 2d ed., 1991; T. J. Maloney, *Electric Circuits*, 2d ed., 1988; R. J. Smith, *Circuits, Devices, and Systems*, 5th ed., 1995.

Circuit (electronics)

An interconnection of electronic devices, an electronic device being an entity having terminals which is described at its terminals by electromagnetic laws. Most commonly these are voltage–current laws, but others, such as photovoltaic relationships, may occur.

Basic Concepts

Some typical electronic devices are represented as shown in **Fig. 1**, where a resistor, a capacitor, a diode, transistors, an operational amplifier, an inductor, a transformer, voltage and current sources, and a ground are indicated. Other devices (such as vacuum tubes, switches, and logic gates) exist, in some cases as combinations of the ones mentioned. The interconnection laws are (1) the Kirchhoff voltage law, which states that the sum of voltages around a closed loop is zero, and (2) the Kirchhoff current law, which states that the sum of the currents into a closed surface is zero (where often the surface is shrunk to a point, the node, where device terminals join). **Figure 2** represents an electronic circuit which is the interconnection of resistors (R, R_{B1}, R_{B2}, R_E, R_L), capacitors (C), a battery voltage source (V_{CC}), a current source (i_s), a bipolar transistor (T), and a switch (S). Functionally Fig. 2 represents a high-pass filter when S is open, and an oscillator when S is closed and the current source is removed. *See* AMPLIFIER; BATTERY; CAPACITOR; CURRENT SOURCES AND MIRRORS; DIODE; ELECTRIC FILTER; ELECTRONIC SWITCH; INDUCTOR; KIRCHHOFF'S LAWS OF ELECTRIC CIRCUITS; LOGIC CIRCUITS; OPERATIONAL AMPLIFIER; OSCILLATOR; RESISTOR; TRANSFORMER; TRANSISTOR; VACUUM TUBE.

Active and passive devices. The devices in an electronic circuit are classified as being either passive or active. The passive devices change signal energy, as is done dynamically by capacitors and statically by transformers, or absorb signal energy, as occurs in resistors, which also act to convert voltages to currents and vice versa. The active devices, such as batteries, transistors, operational amplifiers, and vacuum tubes, can supply signal energy to the circuit and in many cases amplify signal energy by transforming power supply energy into signal energy. Often, though, they are used for other purposes, such as to route signals in logic circuits. Transistors can be considered the workhorses of modern electronic circuits, and consequently many types of transistors have been developed, among which the most widely used are the bipolar junction transistor (BJT), the junction field-effect transistor (JFET), and the metal-oxide-silicon field-effect transistor (MOSFET). *See* ELECTRONIC POWER SUPPLY.

Fortunately, most of these transistors occur in pairs, such as the *npn* and the *pnp* bipolar junction transistors, or the *n*-channel and the *p*-channel MOSFETs, allowing designers to work symmetrically with positive and negative signals and sources. This statement may be clarified by noting that transistors can be characterized by graphs of output current i versus output voltage v that are parametrized by an input current (in the case of the bipolar junction transistor) or input voltage (in the MOSFET and JFET cases). Typically, the curves for an *npn* bipolar junction transistor or an *n*-channel field-effect transistor are used in the first quadrant of the output i-v plane, while for a *pnp* bipolar junction transistor or a *p*-channel field-effect transistor the same curves show up in the third quadrant. Mathematically, if $i = f(v)$ for an *npn* bipolar junction transistor or *n*-channel field-effect device, then $i = -f(-v)$ for a *pnp* bipolar junction transistor or *p*-channel field-effect device when the controlling parameters are also changed in sign.

Transistor properties. Because the transistors are so basic to the operation of electronic circuits, their performance will be briefly discussed.

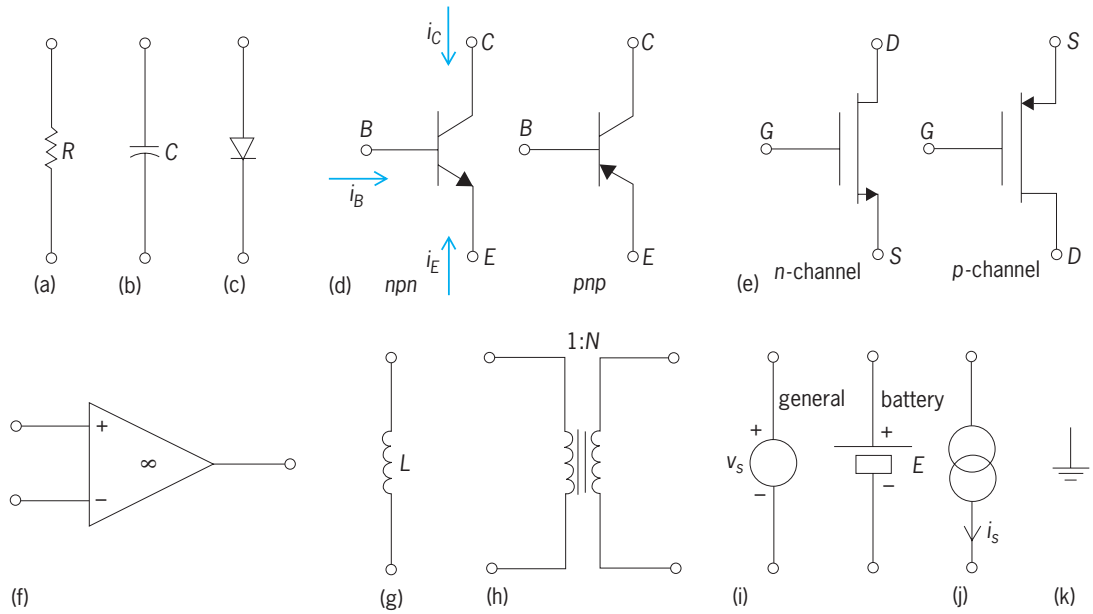

Fig. 1. Representation of some typical electronic devices. (*a*) Resistor. (*b*) Capacitor. (*c*) Diode. (*d*) Bipolar junction transistors (BJTs). (*e*) Metal-oxide-semiconductor field-effect transistors (MOSFETs). (*f*) Operational amplifier. (*g*) Inductor. (*h*) Transformer. (*i*) Voltage sources. (*j*) Current source. (*k*) Ground.

Bipolar transistors. These have three terminals, designated as the base B, the collector C, and the emitter E. These terminals connect to two diode junctions, B-C and B-E, these forming back-to-back diodes. The B-E junction is often forward-biased, in which case its voltage is about 0.7 V, while the B-C junction is reverse-biased for linear operation. In the case of linear operation, the bipolar junction transistor is much more than two back-to-back diodes because it has a very thin base which allows almost all of the current i_E entering the emitter to be collected by the collector. Thus, the collector current is nearly equal to the emitter current. This is expressed in Eq. (1), where

$$i_C = -\alpha i_E \tag{1}$$

the transistor parameter α is very close to 1. (The negative sign arises because the positive direction of current is assumed into the terminals, while one current must come out when the other goes in.) By Kirchhoff's current law for a sphere around the transistor ($i_B + i_C + i_E = 0$), the collector current can be expressed in terms of the base current to obtain Eq. (2), where the parameter β, given by Eq. (3), is

$$i_C = \beta i_B \tag{2}$$

$$\beta = \frac{\alpha}{1-\alpha} \tag{3}$$

very large. From this it follows that in linear operation the (bipolar junction) transistor acts like a current-controlled current source which, when used with base current as input and collector current as output, can have a high current gain β (typically 20 to 600). Because a current can be changed into a voltage via resistors, the transistor in its linear mode of operation also can be used to obtain a high voltage gain. For example, if the current source, the three capacitors, the two resistors labeled R, and the switch are replaced with a capacitor-coupled voltage source v_s between the point B and ground in Fig. 2 the voltage gain is given by Eq. (4), and can be chosen large

$$\frac{v_L}{v_s} = -\alpha \frac{R_L}{R_E} \tag{4}$$

in magnitude by appropriately choosing the ratio of the load resistance R_L to the emitter resistance R_E to be large.

Besides biasing of the junctions for linear operation, any state of the two junctions can occur. For example, both junctions might be forward-biased, in which case the transistor is said to be in saturation and acts nearly as a short circuit between E-C, while if the junctions are simultaneously back-biased the transistor is said to be cut off and acts as an open circuit between all terminals. The transistor can be controlled between saturation and cutoff to make it act as an electronically controlled switch. This mode

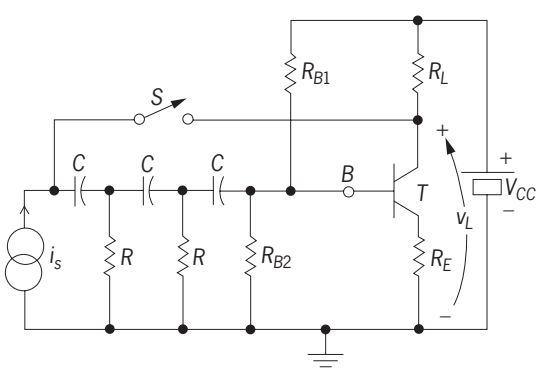

Fig. 2. Diagram of electronic circuit.

of operation is especially useful for binary arithmetic, as used by almost all digital computers, where 0 and 1 logic levels are represented by the saturation and cutoff transistor states.

MOSFET. These also have three regions of operation: cutoff, saturated, and resistive. The MOSFET also has three terminals, the gate G, the drain D, and the source S. A key parameter characterizing the MOSFET is a threshold voltage V_{tb}. When the G-S voltage is below the threshold voltage, no drain current flows and the transistor is cut off. When the G-S voltage is above threshold, then for D-S voltages large in magnitude the MOSFET acts as a voltage-controlled current source, with the drain current given by Eq. (5), where K can conveniently be con-

$$i_D = \frac{K(v_{GS} - V_{tb})^2}{2} \tag{5}$$

trolled by process parameters. Finally, for D-S voltages of small magnitude, the drain curent law is given by Eq. (6), in which case the MOSFET acts as a resis-

$$i_D = K\left[(v_{GS} - V_{tb})v_{DS} - \frac{v_{DS}^2}{2}\right] \tag{6}$$

tor between drain and source which can be varied by the G-S voltage v_{GS}. This resistor is nonlinear, because of the square term, but for very small values of v_{DS} can be considered linear [of conductance $K(v_{GS} - V_{tb})$].

The MOSFET is a versatile device, acting as a voltage-controlled current source in the saturation region and approximately as a voltage-controlled resistor in the resistive region. It can also be electronically controlled between cutoff and the resistive region to make it act as a switch, while for small signals around an operating point in the saturation region it acts as a linear amplifier. Another feature of the MOSFET is that, besides the categories of n-channel and p-channel devices, there are also enhancement- and depletion-mode devices of each category. In practice, for electronic circuit considerations, an n-channel device has $V_{tb} > 0$ for enhancement-mode devices and $V_{tb} < 0$ for depletion-mode devices, while the signs are reversed for p-channel devices.

Biasing of circuits. Since active devices usually supply signal energy to an electronic circuit, and since energy can only be transformed and not created, a source of energy is needed when active devices are present. This energy is usually obtained from batteries or through rectification of sinusoidal voltages supplied by power companies. When inserted into an electronic circuit, such a source of energy fixes the quiescent operation of the circuit; that is, it allows the circuit to be biased to a given operating point with no signal applied, so that when a signal is present it will be processed properly. To be useful, an electronic circuit produces one or more outputs; often inputs are applied to produce the outputs. These inputs and outputs are called the signals and, consequently, generally differ fromthe bias quantities, though often it is hard to separate signal and bias variables. For example, i_s in Fig. 2 would be an input signal, while V_{CC} would be a bias supply voltage. The voltage v_L would contain an output signal but would also include a bias component. The resistors R_{B1}, R_{B2}, R_E, and R_L are used to set the bias point of the transistor, but since they also affect the signal-handling capabilities, their choice is important in the design of the circuit. Biasing of electronic circuits is an important, nontrivial, and often overlooked aspect of their operation. *See* BIAS (ELECTRONICS).

Analog versus digital circuits. Electronic circuits are also classified as analog or digital. Analog circuits work with signals that span a full range of values of voltages and currents, while digital circuits work with signals that are at prescribed levels to represent numerical digits. Analog signals generally are used for continuous-time processes, while digital ones most frequently occur where transistions are synchronized via a clock. However, there are situations where it is desirable to transfer between these two classes of signals, that is, where analog signals are needed to excite a digital circuit or where a digital signal is needed to excite an analog circuit. For example, it may be desired to feed a biomedically recorded signal, such as an electrocardiogram into a digital computer, or it may be desired to feed a digital computer output into an analog circuit, such as a temperature controller. For such cases, there are special electronic circuits, called analog-to-digital and digital-to-analog converters. *See* ANALOG-TO-DIGITAL CONVERTER; DIGITAL-TO-ANALOG CONVERTER.

Feedback. An important concept in electronic circuits is that of feedback. Feedback occurs when an output signal is fed around a device to contribute to the input of the device. For example, in Fig. 2, when the switch S is closed, the voltage v_L is fed around the transistor T back to the input point; removal of i_s then allows v_L to contribute to its own generation, resulting in oscillations if the proper conditions are met. Consequently, when positive feedback occurs, that is, when the output signal returns to reinforce itself upon being fed back, it can lead to the generation of signals which may or may not be wanted. Circuit designers need to be conscious of all possible feedback paths that are present in their circuits so that they can ensure that unwanted oscillations do not occur. In the case of negative feedback, that is, when the output signal returns to weaken itself, then a number of improvements in circuit performance often ensue; for example, the circuit can be made less sensitive to changes in the environment or element variations, and deleterious nonlinear effects can be minimized. *See* CONTROL SYSTEMS; FEEDBACK CIRCUIT.

Specific Circuits

In designing integrated electronic circuits, one of the most useful tools is the current mirror. Two basic current mirrors are that shown in **Fig. 3***a* for taking current out of a terminal attached to point A, and that shown in Fig. 3*b* for sending current into a terminal attached to point B. In Fig. 3*a*, the base-to-emitter

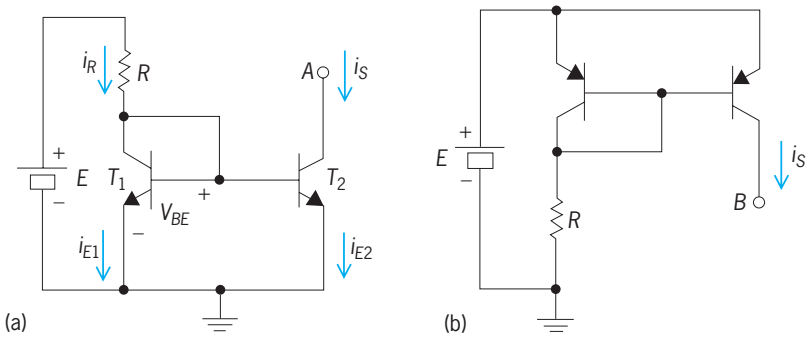

Fig. 3. Basic current mirrors. (*a*) Sink. (*b*) Source.

voltages of both transistors are identical and given by V_{BE}, with typically $V_{BE} = 0.7$ V (since the base-to-emitter junction of T_1 is a forward-biased diode). The current in the resistor i_R is then given by Ohm's law as Eq. (7) since the short circuit between collector

$$i_R = \frac{E - V_{BE}}{R} \tag{7}$$

and base makes T_1 a diode-connected transistor that forces the voltage $E - V_{BE}$ to be across R. From Kirchhoff's current law applied to a sphere encircling T_1, it follows that this current is about equal to the emitter current i_{E1} of T_1 (since the current into the base of transistor T_2 is very small). If the two transistors are identical, then $i_{E2} = i_{E1}$ (because the base-to-emitter voltages are the same). Finally, the collector current i_s of T_2 is about equal to its emitter current [by the construction and characteristics of transistors so that this current is given by Eq. (8)], mirroring the current i_R.

$$i_s = i_R \tag{8}$$

One of the most versatile circuits is the differential pair, since it finds increasing use in both analog and digital circuits. A bipolar-junction-transistor differential pair is shown in **Fig. 4**, where the dc bias current source I_E can be constructed from the output of a current mirror of the type of Fig. 3*a*. If the differential input voltage v_{id} is defined by Eq. (9), then it

$$v_{id} = v_{i1} - v_{i2} \tag{9}$$

can be shown that the differential output voltage v_{od} is given by Eq. (10), where V_T is the thermal voltage

$$v_{od} = \alpha R I_E \tanh \frac{v_{id}}{2V_T} \tag{10}$$

($V_T = 25$ mV at room temperature), α is the parameter alpha (defined above) for the two transistors (which are assumed identical), and tanh[·] is the hyperbolic tangent function. As Eq. (10) shows, the differential pair performs as a differential voltage-controlled voltage source. An MOS differential pair is obtained simply by replacing the bipolar junction transistors in Fig. 4 by MOSFETs, although a somewhat more complicated input-output law results. *See* DIFFERENTIAL AMPLIFIER.

Analog circuits. By replacing the resistive loads of the differential amplifier of Fig. 4 with the current mirror of Fig. 3*b*, a differential voltage-controlled current source (DVCCS) is generated, as shown in **Fig. 5**, where from the above Eq. (11) follows. Since

$$i_o = i_{c2} - i_{c1} = aI_E \tanh \frac{v_{id}}{2V_T} \tag{11}$$

$\tanh(x) \simeq x$ when x is very small in magnitude, for small differential inputs the linear law (12) holds.

$$i_0 = \frac{\alpha I_E v_{id}}{2V_T} \tag{12}$$

An application of the differential voltage-controlled current source is seen by connecting its output to a capacitor, in which case an integrator results. Connecting the output of a differential voltage-controlled current source to its input can result in a resistor.

Operational amplifiers. The differential pair forms the input stage to the operational amplifier (op amp), an

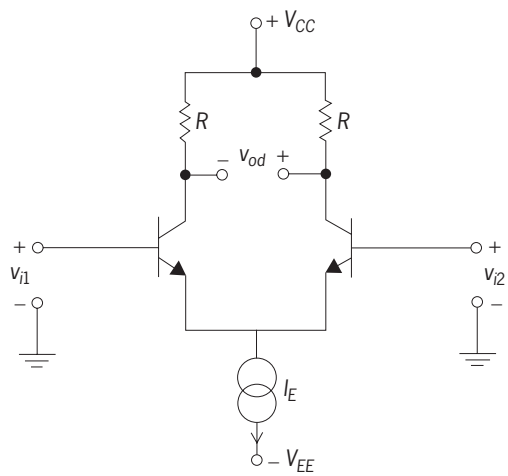

Fig. 4. Differential pair based on bipolar junction transistors.

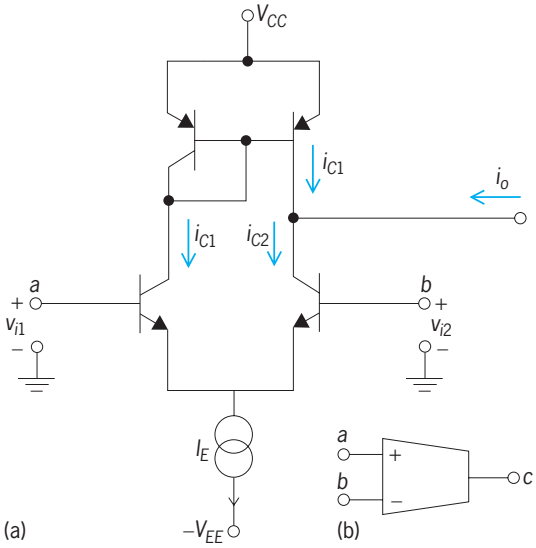

Fig. 5. Differential voltage-to-current converter. (*a*) Circuit. (*b*) Symbol.

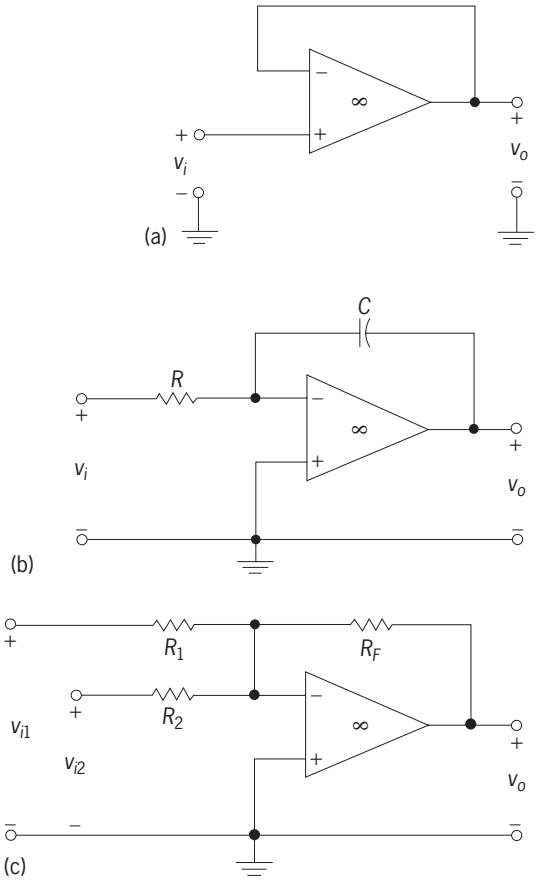

Fig. 6. Basic operational amplifier circuits. (*a*) Unity-gain (isolation) amplifier. (*b*) Integrator. (*c*) Summer.

extremely important electronic circuit device. The operational amplifier is typically realized by taking v_{od} of Fig. 4, greatly amplifying it, terminating with an output stage that allows for loading, and improving the input terminals so that miniscule input current is drawn. From this description, it follows that an operational amplifier in the linear region is characterized by the input-output voltage relationship given by Eq. (13), with K tending to infinity. Since v_o is

$$v_o = Kv_{id} \tag{13}$$

finite, this means that either $v_{id} \simeq 0$ or the output is saturated at a bias voltage of the operational amplifier. In either case, for most practical purposes the input current can be assumed to be zero. Therefore, in the linear region the operational amplifier has the very interesting description given by Eq. (14),

$$v_{id} = 0, \quad i_{\text{input}} = 0 \tag{14}$$

in which case its input looks like a simultaneous short and open circuit (while its output can be "anything").

By applying feedback through a direct connection of the output to the negative input, the unity-gain buffer amplifier satisfying Eq. (15) results. The buffer

$$v_o = v_i \tag{15}$$

amplifier is shown in **Fig. 6***a*. If the positive input is grounded and negative feedback is applied, then the negative input must be also at ground potential and, since the operational amplifier draws no current, the negative input terminal is said to be a virtual ground. Using the concept of the virtual ground, it is readily seen that the circuit of Fig. 6*b* functions as an integrator, given by Eq. (16), where s is the derivative

$$v_o = \frac{-v_i}{RCs} \tag{16}$$

operator (and hence its inverse, $1/s$, is the integral operator). The circuit of Fig. 6*c* functions as a summer, given by Eq. (17). Using the three circuits of Fig. 6,

$$v_o = -R_F\left(\frac{v_{i1}}{R_1} + \frac{v_{i2}}{R_2}\right) \tag{17}$$

any set of linear differential equations with constant coefficients can be simulated, and arbitrary transfer functions realized. The operational amplifier is one of the most useful of the electronic circuit devices. *See* ANALOG COMPUTER.

Applying positive feedback in operational amplifier circuits, some interesting results can also be obtained. For example, the circuit of **Fig. 7***a* is a practical way to obtain a negative resistance; its operation depends on Eq. (18). When the operational

$$i = \frac{v - v_o}{R} \tag{18}$$

amplifier is in its linear region, that is, when notation (19) is valid, Eq. (20) holds, giving the negative resistor equation (21). When the operational amplifier

$$-V_{EE} < v_o < V_{CC} \tag{19}$$

$$v_o = \left(1 + \frac{R_F}{R_1}\right)v \tag{20}$$

$$i = -\left(\frac{R_F}{R_1 R}\right)v \tag{21}$$

is forced into saturation the behavior of the circuit

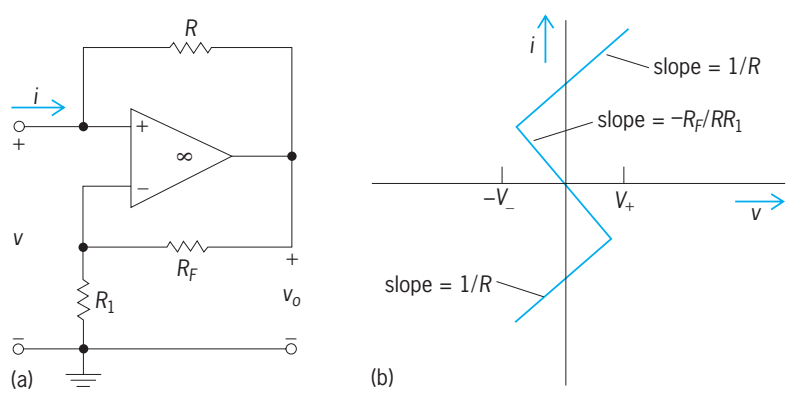

Fig. 7. Negative resistor circuit. (*a*) Diagram. (*b*) Voltage-current characteristic.

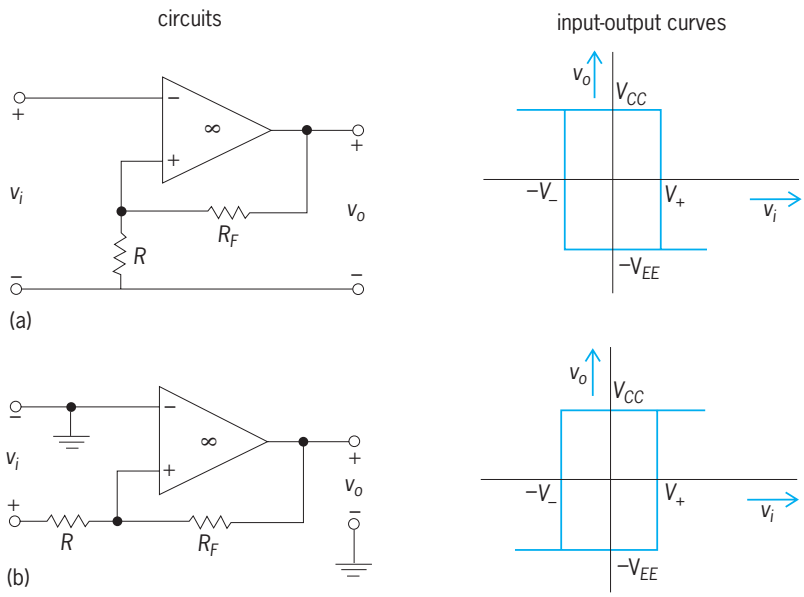

Fig. 8. Schmitt trigger circuits and their input-output curves. (*a*) Circuit with isolated input (negative gain). $V_+ = V_{CC}/[1 + (R_F R)]$; $V_- = V_{EE}/[1 + (R_F R)]$. (*b*) Circuit with coupled input (positive gain). $V_+ = RV_{CC}R_F$; $V_- = RV_{EE}/R_F$.

is given by Eqs. (22) and (23) [Fig. 5]. The resulting

$$i = \frac{v - V_{CC}}{R} \quad \text{when } v < V_+ = \frac{V_{CC}}{1 + (R_F/R_1)} \tag{22}$$

$$i = \frac{v + V_{EE}}{R} \quad \text{when } v > -V_- = -\frac{V_{EE}}{1 + (R_F/R_1)} \tag{23}$$

transfer characteristic is shown in Fig. 7*b*. This negative resistor is multiple-valued when the input voltage v is used as the independent (that is, control) variable. Hence, an attempt to use it as a voltage-controlled device will result in its exhibiting hysteresis, since a switch onto the upper curve occurs at $v = V_+$ whereas a switch off the upper curve occurs at $-V_-$. *See* HYSTERESIS; NEGATIVE-RESISTANCE CIRCUITS; OPERATIONAL AMPLIFIER.

Schmitt triggers. Other circuits exhibiting hysteresis are Schmitt trigger circuits (**Fig. 8**), where positive feedback is again used. In these circuits the output is taken as the operational amplifier output voltage, which saturates at the bias voltages. As shown in the input-output curves of Fig. 8, the Schmitt triggers give binary-valued hysteresis, which makes them useful for storage of binary information or for maintaining a constant output in the presence of noise.

Role of nonlinearity. Since the characteristics of the circuits of Figs. 7 and 8 are very nonlinear, they serve as examples to illustrate that nonlinear theory plays an important role in the area of electronic circuits. There are many other instances where nonlinearities are important, for example, in the construction of analog multipliers where the output is the product of two inputs. (Such a multiplier can be obtained via the differential pair of Fig. 5 if I_E is made to vary with an input. This is achievable by using the current sink of Fig. 3*a* to generate I_E. With this circuit the voltage E becomes the second input.) Likewise, nonlinearities are required in the modern theories of neural networks, where so-called sigmoid-type nonlinearities are used to obtain convergence toward set patterns. (The hyperbolic tangent function realized by the differential pair is a sigmoid-type of nonlinearity.) *See* NEURAL NETWORK.

Filters. One of the most common uses of analog active circuits is in the construction of electrical filters. Electrical filters are circuits whose input-output characteristics allow signals at certain frequencies (called the passband) to pass through unattenuated while signals at unwanted frequencies (called the stopband) are eliminated. By appropriately positioning the passband and the stopband, various filter types such as low-pass, band-pass, and high-pass can be attained. Filter design often concentrates on low-pass filters, since through relatively elementary transformations the band-pass and high-pass responses can be obtained from the low-pass transfer function. The transfer function, $H(s)$, is the ratio of the output v_o to the input v_i as a function of the complex frequency variable s.

Maximally flat filters are common with these resulting from using the Butterworth polynomials, $B_n(s)$, to obtain the low-pass transfer function given by Eq. (24). Here, $B_n(s)$ is given by Eq. (25), where

$$H(s) = 1/B_n(s) \tag{24}$$

$$B_n(s) = (s - s_1) \cdot \ldots \cdot (s - s_n) \tag{25}$$

$s_1, \ldots, s_n$ are the $2n$th roots of $(-1)^{n+1}$ with negative real parts. As a function of frequency, f, the magnitude of the transfer function satisfies Eq. (26), so

$$|H(j2\pi f)| = \frac{1}{[1 + (2\pi f)^{2n}]^{1/2}} \tag{26}$$

that it is equal to 1 at dc and to $1/\sqrt{2}$ at $f = 1/2\pi$ [termed the 3-decibel attenuation point]. The substitution of Eq. (27) shifts the 3-dB point to the frequency $\Omega/(2\pi)$; the substitution of Eq. (28) changes

$$s = \frac{S}{\Omega} \tag{27}$$

$$s = \frac{\Omega}{S} \tag{28}$$

the low-pass characteristic to high-pass. The substitution of Eq. (29) leads to a band-pass characteristic

$$s = \Omega_o \left(\frac{S}{\Omega} + \frac{\Omega}{S} \right) \tag{29}$$

(with Ω_o adjusting the bandwidth and Ω adjusting the center frequency).

Since the complex frequency variable s also serves as the derivative operator, a description of the filter in terms of differential equations can be obtained from the transfer function. One way these differential

equations can be realized is with the electronic circuits of Fig. 6. Alternatively, the transfer function can be realized by specialized circuits created for the purpose of filtering. Indeed, because of the importance of filters to technology, their theory is extensive, and vast tables of filter transfer functions and circuits to realize them have been compiled. *See* ELECTRIC FILTER.

Digital circuits. The digital computer is based on digital electronic circuits. Although some of the circuits are quite sophisticated, such as the microprocessors integrated on a single chip, the concepts behind most of the circuits involved in digital computers are quite simple compared to the circuits used for analog signal processing. The most basic circuit is the inverter; a simple realization based upon the MOS transistor is shown in **Fig. 9***a*. The upper (depletion-mode) transistor acts as a load "resistor" for the lower (enhancement-mode) transistor, which acts as a switch, turning on (into its resistive region) when the voltage at point *A* is above threshold to lower the voltage at point *B*. Adding the output currents of several of these together into the same load resistor gives a NOR gate, a two-input version of which is shown in Fig. 9*b*; that is, the output is high, with voltage at V_{DD}, if and only if the two inputs are low. Placing the drains of several of the enhancement-mode switches in series yields the NAND gate, a two-input version of which is shown in Fig. 9*c*; that is, the output is low if and only if both inputs are high. From the circuits of Fig. 9, the most commonly used digital logic circuits can be constructed. Because these circuits are so simple, digital circuits and digital computers are usually designed on the basis of negation logic, that is, with NOR and NAND rather than OR and AND circuits. *See* DIGITAL COMPUTER; INTEGRATED CIRCUITS; MICROPROCESSOR.

Conversion. Because most signals in the real world are analog but digital computers work on discretizations, it is necessary to convert between digital and analog signals. As mentioned above, this is done through digital-to-analog and analog-to-digital converters. Most approaches to digital-to-analog conversion use summers of the type shown in Fig. 6*c*, where the voltages representing the digital bits are applied to input resistors, either directly or indirectly through switches gated on by the digital bits which change the input resistance fed by a dc source (as is the case in the R-to-2R type of analog-to-digital converters).

One means of doing analog-to-digital conversion is to use a clocked counter that feeds a digital-to-analog converter, whose output is compared with the analog signal to stop the count when the digital-to-analog output exceeds the analog signal. The counter output is then the analog-to-digital output. The comparator for such an analog-to-digital converter is similar to an open-loop operational amplifier (which changes saturation level when one of the differential input levels crosses the other). Other types of analog-to-digital converters, called flash converters, can do the conversion in a shorter time by use of parallel operations, but they are more expensive.

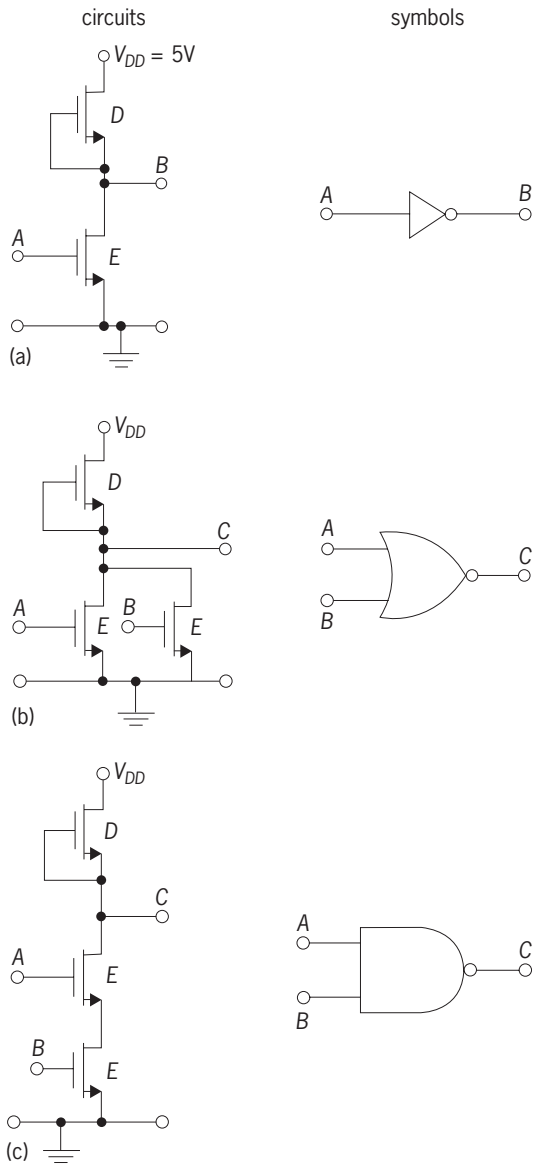

Fig. 9. Digital logic gates and their symbols. *D* = depletion-mode transistor; *E* = enhancement-mode transistor. (*a*) Inverter. (*b*) NOR gate. (*c*) NAND gate.

Other circuits. The field of electronic circuits is very broad and there are a very large number of other circuits besides those discussed above. For example, the differential is a key element in operational amplifier design and in biomedical data acquisition devices which must also be interfaced with specialized electronic sensors. Light-emitting and -detecting diodes allow for signals to be transmitted and received at optical frequencies. Liquid crystals are controlled by electronic circuits and are useful in digital watches, flat-panel color television displays, and electronic shutters. *See* BIOMEDICAL ENGINEERING; ELECTRONIC DISPLAY; LIGHT-EMITTING DIODE; LIQUID CRYSTALS; OPTICAL DETECTORS; TRANSDUCER.

Design

Because some circuits can be very complicated, and since even the simplest circuits may have complicated behavior, the area of computer-aided design

(CAD) of electronic circuits has been extensively developed. A number of circuit simulation programs are available, some of which can be run on personal computers with good results. These programs rely heavily upon good mathematical models of the electronic devices. Fortunately, the area of modeling of electronic devices is well developed, and for many devices there are models that are adequate for most purposes. But new devices are constantly being conceived and fabricated, and in some cases no adequate models for them exist. Thus, many of the commerical programs allow the designer to read in experimentally obtained data for a device from which curve fitting techniques are used to allow an engineer to proceed with the design of circuits incorporating the device. Reproducibility and acceptability of parts with tolerances are required for the commerical use of electronic circuits. Consequently, theories of the reliability of electronic circuits have been developed, and most of the computer-aided design programs allow the designer to specify component tolerances to check out designs over wide ranges of values of the elements. Finally, when electronic circuits are manufactured they can be automatically tested with computer-controlled test equipment. Indeed, an area that will be of increasing importance is design for testability, in which decisions on what to test are made by a computer using knowledge-based routines, including expert systems. Such tests can be carried out automatically with computer-controlled data-acquisition and display systems. *See* CIRCUIT (ELECTRICITY); COMPUTER-AIDED DESIGN AND MANUFACTURING; EXPERT SYSTEMS; RELIABILITY, AVAILABILITY, AND MAINTAINABILITY; ROBOTICS. Robert W. Newcomb

Bibliography. P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3d ed., 1992; D. A. Hodges and H. G. Jackson, *Analysis and Design of Digital Integrated Circuits*, 2d ed., 1988; J. Millman and A. Grabel, *Microelectronics*, 2d ed., 1987; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997; G. C. Temes and J. W. LaPatra, *Introduction to Circuit Synthesis and Design*, 1977.

Circuit breaker

A device to open or close an electric power circuit either during normal power system operation or during abnormal conditions. A circuit breaker serves in the course of normal system operation to energize or deenergize loads. During abnormal conditions, when excessive current develops, a circuit breaker opens to protect equipment and surroundings from possible damage due to excess current. These currents are usually the result of short circuits created by lightning, accidents, deterioration of equipment, or sustained overloads. *See* CIRCUIT (ELECTRICITY); LIGHTNING AND SURGE PROTECTION.

Formerly, all circuit breakers were electromechanical devices. In these breakers a mechanism operates one or more pairs of contacts to make or break the circuit. The mechanism is powered either electromagnetically, pneumatically, or hydraulically. The contacts are located in a part termed the interrupter. When the contacts are parted, opening the metallic conductive circuit, an electric arc is created between the contacts. This arc is a high-temperature ionized gas with an electrical conductivity comparable to graphite. Thus the current continues to flow through the arc. The function of the interrupter is to extinguish the arc, completing circuit-breaking action.

Current interruption. In alternating-current circuits, arcs are usually extinguished at a natural current zero, when the ac voltage applied across the arcing contacts reverses polarity. Within a short period around a natural current zero, the power input to the arc, equal to the product of the instantaneous current and voltage, is quite low. There is an opportunity to remove more energy from the arc than is applied to it, thus allowing the gas to cool and change from a conductor into an insulator. *See* ALTERNATING CURRENT.

In direct-current circuits, absence of natural current zero necessitates the interrupter to convert the initial arc into one that could only be maintained by an arc voltage higher than the system voltage, thus forcing the current to zero. To accomplish this, the interrupter must also be able to remove energy from the arc at a rapid rate. *See* DIRECT CURRENT.

Different interrupting mediums are used for the purpose of extinguishing an arc. In low- and medium-voltage circuits (1110–15,000 V) the arc is driven by the magnetic field produced by the arc current into an arc chute (**Fig. 1**). In the arc chute the arc is either split into several small arcs between metal plates or driven tightly against a solid insulating material. In the former case the splitting of the arc increases the total arc voltage, thus increasing the rate of energy dissipation. In the latter case the insulating material is heated to boiling temperatures, and the evaporated material flows through the arc, carrying a great deal of energy with it.

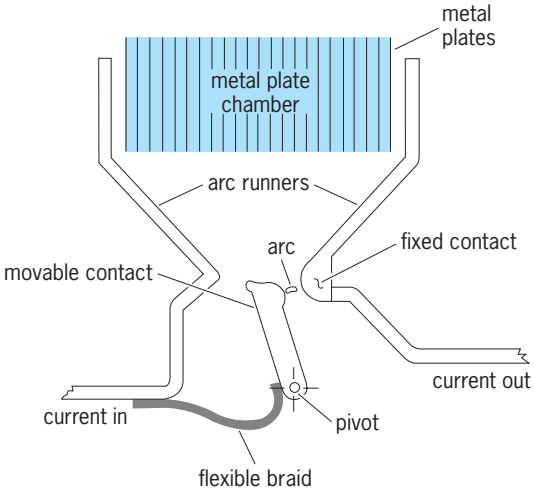

Fig. 1. Cross section of interrupter for a typical medium-voltage circuit breaker.

Oil and gas-blast circuits. For outdoor applications at distribution and subtransmission voltage (10 kV and above), oil breakers are widely used. In the United States, bulk oil breakers are used (**Fig. 2**), and in European and many other countries, "low-oil-content" breakers are quite popular. The principles of operation of both kinds of oil breakers are basically the same. Only the amount of oil used and the detailed engineering design differ. In oil breakers, the arc is drawn in oil. The intense heat of the arc decomposes the oil, generating high pressure that produces a fluid flow through the arc to carry energy away. At transmission voltages below 345 kV, oil breakers used to be popular. They were, however, increasingly replaced by gas-blast circuit breakers such as air-blast breakers (**Fig. 3**) and sulfur hexafluoride (SF_6) circuit breakers (**Fig. 4**).

In air-blast circuit breakers, air is compressed to high pressures (approximately 50 atm = 5×10^6 pascals). When the contacts part, a blast valve is opened to discharge the high-pressure air to ambient, thus creating a very high-velocity flow near the arc to dissipate the energy. In SF_6 circuit breakers, the same principle is employed, with SF_6 as the medium instead of air. In the "puffer" SF_6 breaker, the motion of the contacts compresses the gas and forces it to flow through an orifice into the neighborhood of the arc. Both types of SF_6 breakershave been developed

Fig. 2. Bulk oil circuit breaker for 138-kV application.

Fig. 3. Air blast circuit breaker rated for 500 kV.

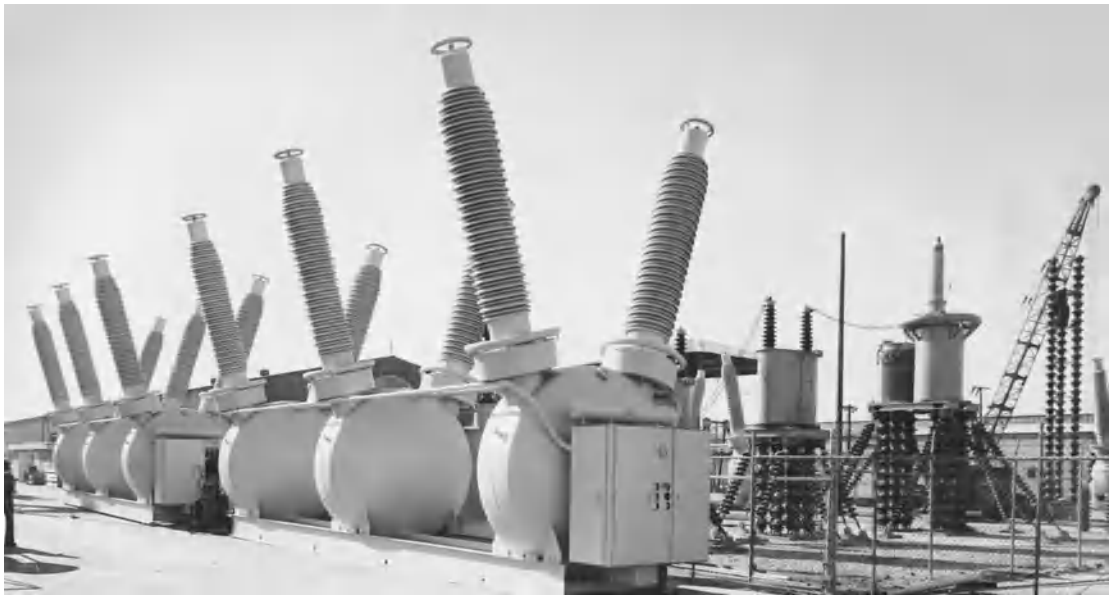

Fig. 4. Sulfur hexafluoride circuit breakers, 500-kV, 3-kA.

for EHV (extra-high-voltage) transmission systems.

For EHV systems, it was discovered that closing of a circuit breaker may cause a switching surge which may be excessive for the insulation of the system. The basic principle is easy to understand. If a breaker is closed at the peak of the voltage wave to energize a single-phase transmission line which is open-circuited at the far end, reflection can cause the transient voltage on the line to reach twice the peak of the system voltage. If there are trapped charges on the line, and if it happens that at the moment of breaker-closing the system voltage is at its peak, equal in magnitude to but opposite to the polarity of the voltage due to the trapped charges left on the line, the switching surge on the line can reach a theoretical maximum of three times the system voltage peak.

One way to reduce the switching surge is to insert a resistor in series with the line (**Fig. 5**) for a short time. When switch A in Fig. 5 is closed, the voltage is divided between the resistance and the surge impedance Z of the line by the simple relationship $V_L = V[Z/(R + Z)]$, where V is voltage impressed, V_L is voltage across the transmission line, R is resistance of the resistor, and Z is surge impedance of the line. Only V_L will travel down the line and be reflected at the other end. The magnitude of the switching surge is thus considerably reduced.

Vacuum breakers. The vacuum breaker, another electromechanical device, uses the rapid dielectric recovery and high dielectric strength of vacuum. A pair of contacts is hermetically sealed in a vacuum envelope (**Fig. 6**). Actuating motion is transmitted through bellows to the movable contact. When the contacts are parted, an arc is produced and supported by metallic vapor boiled from the electrodes. Vapor particles expand into the vacuum and condense on solid surfaces. At a natural current zero the vapor particles disappear, and the arc is extinguished. Vacuum breakers of up to 242 kV have been built (**Fig. 7**).

Solid-state breakers. The other type of breaker uses a thyristor, a semiconductor device which in the off state prevents current from flowing but

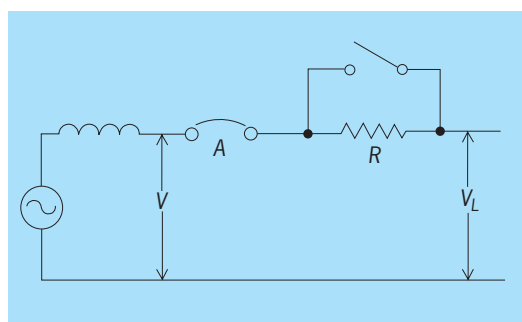

Fig. 5. Resistor insertion to reduce switching surge.

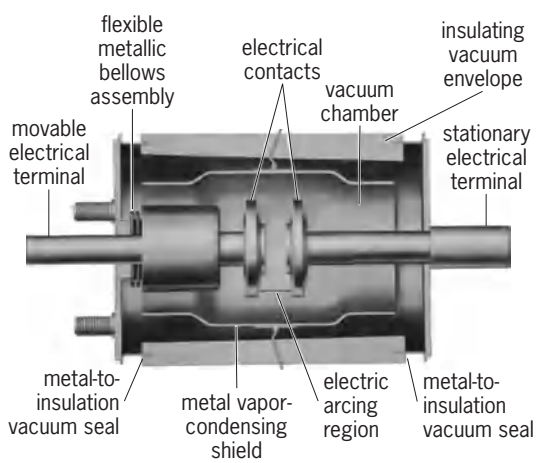

Fig. 6. Cutaway view of vacuum interrupter.

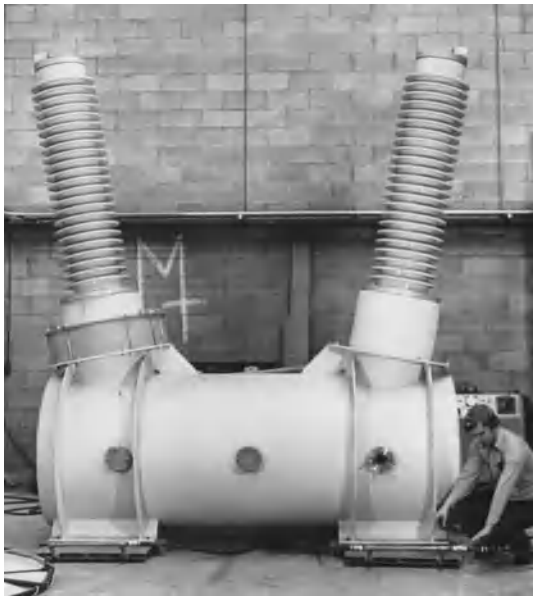

Fig. 7. Vacuum circuit breaker, 242-kV, 40-kA.

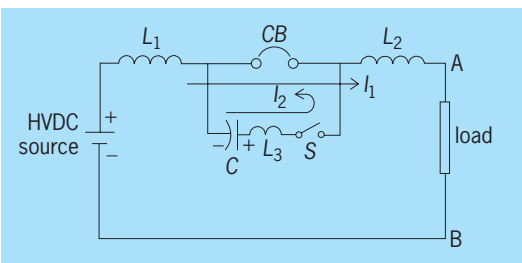

Fig. 8. Commutation principles of circuit interruption.

which can be turned on with a small electric current through a third electrode, the gate. At the natural current zero, conduction ceases, as it does in arc interrupters. This type of breaker does not require a mechanism. Semiconductor breakers have been built to carry continuous currents up to 10,000 A.

Semiconductor circuit breakers can be made to operate in microseconds if the commutation principle is applied. **Figure 8** illustrates the commutation principle for an hvdc (high-voltage direct-current) circuit, but it can easily be extended to ac circuits. During normal operation, the circuit breaker (CB in the diagram) would be closed from the hvdc source. In this diagram, inductances L_1 and L_2 represent the circuit inductance on either side of the breaker. Suppose that a fault occurs which applies a short circuit between points A and B. The current will commence to increase, its rate being determined by L_1 and L_2. When the increased current is detected, the contacts of the circuit breakers are opened, drawing an arc, and the switch (S in the diagram) is closed, causing the precharged capacitors C to discharge through the circuit breaker. The current I_2 so produced is traveling in such a direction as to oppose I_1 and drive it to zero, thereby giving the circuit breaker an opportunity to interrupt.

Thyristors can be used for both circuit breakers and switches. The closing and opening operations are, of course, not mechanical but are controlled by the gates. Such a circuit breaker was built for very special switching applications, such as thermonuclear fusion research. Thomas H. Lee

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; K. Nakanishi (ed.), *Switching Phenomena in High-Voltage Circuit Breakers*, 1991; M. H. Rashid, *Power Electronics*, 2d ed., 1992.

Circulation

Those processes by which metabolic materials are transported from one region of an organism to another. Ultimately, the essential gases, nutrients, and waste products of metabolism are exchanged across cell membranes by diffusion. Diffusion is the movement of material, by random motion of molecules, from a region of high concentration to one of low concentration. The amount of material moved from one place to another depends on the difference in concentrations and on the distance between the two points. The greater the distance, the less movement of material per unit time for a given difference in concentration. Consequently, in all but the smallest animals, convection (or bulk circulation) of materials to the cell must be employed to supplement diffusion.

Protoplasmic movement aids diffusion at the intracellular level. In multicellular animals, however, either the external medium or extracellular body fluids, or both, are circulated. In sponges and coelenterates (**Fig. 1***a* and *b*), water is pumped through definite body channels by muscular activity or, more often, by cilia or flagella on the cells lining the channels. Even external water currents can be tapped for this purpose. It is generally assumed that flows driven by these mechanisms will be slow. However, this may not be the case.

In sponges, flow driven by flagella in flagellated chambers enters through pores in the body wall and exits through the large central osculum. The cross-sectional area of all the flagellated chambers is approximately 6000 times that of the osculum. Hence, the exceptionally low flow velocity in the flagellated chambers increases markedly as the total volume of water is forced out through the osculum. In fact, water may flow through the osculum at velocities of $5\ \text{cm} \cdot \text{s}^{-1}$, approaching those attained in major arterial vessels in vertebrates. This is an example of the principle of continuity, which states that the total volume of fluid moving past any point in a circulation (providing it has no major leaks) must be the same as that flowing past any other point. In regions of large cross-sectional area, the velocity of flow is low, whereas when cross-sectional area is small the flow velocity will be high.

Coelenterates have a body wall derived from two cell layers; an outer ectoderm is separated from an inner endoderm by a noncellular gelatinous

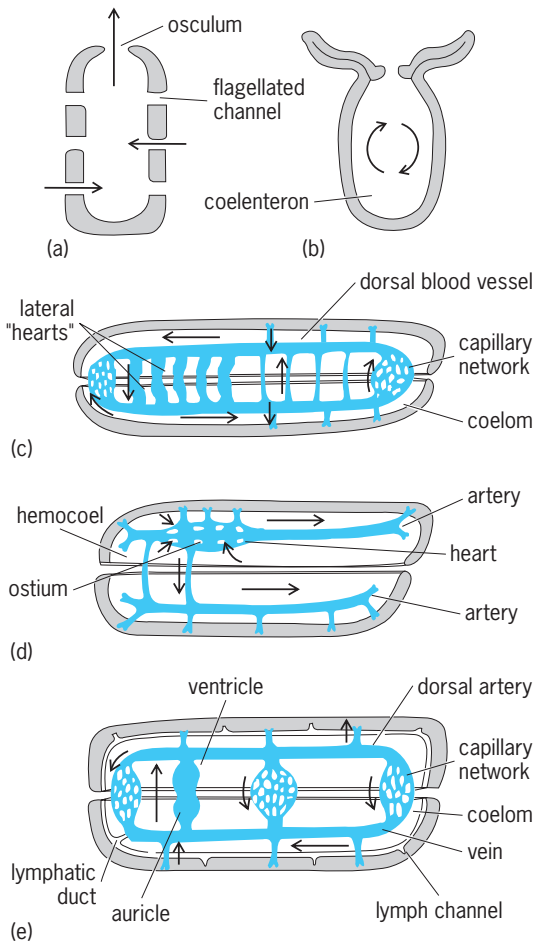

Fig. 1. Systems of internal transport. (*a*) Sponge type, with flagellated channels. (*b*) Coelenterate type. (*c*) Oligochaete type, with coelom and closed blood system. (*d*) Crustacean-molluscan type, with blood system open to hemocoel which is derived from primary cavity or blastocoel. (*e*) Vertebrate type, with coelom, closed vascular system, and lymphatic channels. (*After C. L. Prosser, Comparative Animal Physiology, 3d ed., W. B. Saunders, 1973*)

material (mesoglea). All higher animals have bodies consisting of three cell layers, with the ectoderm being separated from the endoderm by a cellular layer of mesoderm. The mesoderm proliferates and separates to develop a fluid-filled body cavity or coelom. The coelom separates the ectoderm (together with an outer layer of mesoderm) from the endoderm (which has an inner layer of mesoderm). Coelomic fluid is moved around by body movements or ciliary activity, but in larger animals this movement is usually inadequate to supply the metabolic requirements of the organs contained within the coelom. These needs are provided for by pumping a fluid, blood, to them through vessels, the blood vascular system (Fig. 1*c*–*e*).

Open and Closed Circulatory Systems

When the blood is in a separate compartment from the rest of the extracellular fluid, the vascular system is described as closed. In such a system, the blood is circulated by a pump, the heart, through special channels, blood vessels; it comes into close association with the tissues only in the capillaries, fine vessels with walls only one cell thick. In some tissues or regions, larger blood spaces may exist, called sinuses. A closed vascular system is found in most annelids (segmented worms and leeches), cephalopod mollusks (squids and octopods), holothurian echinoderms (sea cucumbers), and vertebrates (Fig. 1*c* and *e*).

In most vertebrates, a functional but anatomically closed connection exists between the extracellular spaces (between the cells) and the blood vascular system in the form of lymph channels. Lymph is derived from the noncellular component of blood (plasma), modified in its passage through the tissues, and is picked up by blind-ending lymphatic vessels, which conduct the lymph to the veins (Fig. 1*e*). The presence of a lymphatic system in fish is controversial. In many fishes, a secondary circulation exists which parallels the primary circulation in extent and connects to it by narrow channels, which are small enough to exclude most red blood cells from the secondary circulation. It is unlikely that the secondary circulation is the evolutionary forerunner of a true lymphatic system, which is first observed in lungfishes, or that the secondary circulation subserves the same function as a true lymphatic system.

In most arthropods (crustaceans, insects), most mollusks (shellfish), and many ascidians (sea squirts), the extracellular spaces are confluent with the blood system. In these animals, blood is pumped through a limited network of vessels into a body cavity called a hemocoel (Fig. 1*d*). After bathing the tissues, blood (called hemolymph in these organisms) collects in sinuses and returns to the heart. This is the open vascular system. In animals with open circulatory systems, the coelom is much reduced.

Dynamics

A pump is required to circulate the blood in animals. The pump (usually called a heart) imparts propulsive energy to the blood, which flows from regions of high fluid energy to regions of low fluid energy (or down an energy gradient). The total energy of flowing blood consists of pressure energy, imparted by the pump, and energy contained in blood due to its motion (kinetic energy). Even in mammals, where blood flows very rapidly, the kinetic energy component is small compared with pressure energy and is usually ignored. Consequently, it is usual to refer to blood flow as taking place down a pressure gradient rather than energy gradient.

As the blood courses around the circulatory system, pressure energy is lost as heat. This loss of energy is described as resistance to flow. In closed circulations, flow resistance is concentrated in short narrow regions of blood vessels, known as resistance vessels. These are fine vessels (diameter of 0.02 mm), and since they have a muscular coat they are able to change their diameter to regulate flow. However, it is not the friction between the wall and blood that causes energy dissipation but rather friction between the various layers of the flowing blood (internal

friction or viscosity). Blood flow in tubes is streamlined; the layer of blood at the wall is stationary while that in the middle flows fastest. Hence, slower-moving layers close to the wall tend to slow down and stop faster-moving layers, while they, in turn, will try to speed up slower-flowing layers. It is this internal friction which must be overcome by flowing blood, and the higher the viscosity of the fluid, the more energy will be dissipated in doing this. For instance, it takes much more effort to force treacle (a high-viscosity fluid) than water, at the same flow rate, through the same-diameter pipe. Blood is also much thicker (or more viscous) than water.

In a closed circulation, the total peripheral resistance (R) can be expressed by the equation below,

$$R = \frac{P_1 - P_2}{Q}$$

in which P_1 = pressure in the outflow vessels of the heart, P_2 = pressure in the inflow vessels, and Q = the output of a single propulsive chamber (ventricle). This equation is a simplified form of a formula used to describe flow in tubes (Poiseuille's law) and is a direct analog of Ohm's law for the relationship between electrical potential difference (pressure gradient), current (flow), and resistance. The resistance of the vascular system will depend on its geometry, as well as the viscosity of the fluid. Of the geometric factors, the radius of the tubes is far more important than length because only the radius is a variable in the circulation. In fact, flow per unit time changes in proportion to the fourth power of the radius for a given fall in pressure along the vessel. Hence, in the resistance vessels, halving the radius by muscular contraction will reduce flow to one-sixteenth of what it was before for the same pressure gradient.

In open circulations, the concept of total peripheral resistance cannot apply. Certainly, the pump imparts propulsive energy to the blood which is dissipated, due to viscous losses, as the blood courses through the hemocoel. But a pressure difference between the outflow and inflow vessels of the heart can, at best, indicate only the resistance of the major shunts in the circulation. A shunt is a short circuit between the outflow and inflow vessels. In the simplest possible case, when the pump sits in the hemocoel and there are no blood vessels, the heart activity only stirs the blood, and pressure differences across the heart are mostly a reflection of its filling and emptying mechanisms.

Physiology of Open and Closed Systems

In open blood systems, there is no separation between extracellular fluid and blood, so the total volume of hemolymph in these animals is large (up to 40% of the body weight). In closed circulations, blood is separated from fluid around the cells, and blood volume is much smaller, about 5–10% of the body weight. In both open and closed systems, blood is conveyed directly to the organs, but only in closed systems can its distribution be regulated with any precision. The best that can be done to facilitate the normal functioning of open circulations is to regulate hemolymph flow to a particular region by opening or closing valves of the junction of the major arteries and the heart.

On the outflow side of the pump (arterial side), closed vascular systems are effectively overfilled with blood, and the elastic blood vessel walls bear down on their contained volume and maintain a high pressure, even when the pump is in the filling phase (diastole). The arterial system is "topped up" again with each ejection of the pump (systole). Obviously, blood pressure can be maintained only if rapid escape of blood from the arteries is prevented by having a high resistance to flow in the periphery of the circulation. As a consequence, almost continuous blood flow is provided to the tissues by the elastic recoil of the stretched arterial walls during diastole. In open systems, peripheral flow resistance is low, as are the pressures generated in these systems. Furthermore, pressures in the outflow vessels of the pump will be highly pulsatile, reaching a peak in systole (the pump-ejection phase) and perhaps falling to zero in diastole (the pump-filling phase). Likewise, flow will be extremely pulsatile, often stopping between heartbeats.

In both open and closed circulatory systems, the volume of blood pumped by the heart must be the same volume as returns to the heart in any given time period. However, only in closed circulations does this have important consequences with regard to the rate at which blood flows through the vessels in the various regions of the vascular system. The velocity of blood flow at any part of a closed system depends on the cross-sectional area of the vessels in that part. Since the vessels that arise from or join to the heart must necessarily be about the same size, flow velocity in the veins is almost as rapid as in the arteries. Furthermore, since the cross-sectional area of the vessels within all the tissues is many times greater than vessels leaving the heart, the flow in the tissue vessels is very slow (usually less than $1\ \text{mm} \cdot \text{s}^{-1}$), allowing plenty of time for exchange across cell walls.

Components of Circulatory Systems

The two principal components of all circulatory systems are hearts and blood vessels.

Hearts. The circulatory pumps are the hearts, all of which operate by causing a wave of muscular contraction followed by a wave of relaxation (peristaltic wave) to travel over the walls of a tube. When the muscle contracts, the volume is reduced and fluid contents are expelled.

Tubular heart. If the heart persists in adult animals as a tube, such as the dorsal vessel of annelids or the dorsal heart of insects (**Fig. 2*b***), it is called a tubular heart. The peristaltic wave passes along the tube in an anterior direction, pushing the blood before it in the same way the air in a balloon is squeezed down to one end by pulling the balloon through a cupped hand. Therfore, valves are unnecessary to ensure unidirectional flow of blood. In earthworms, the peristaltic waves travel up the dorsal vessel at

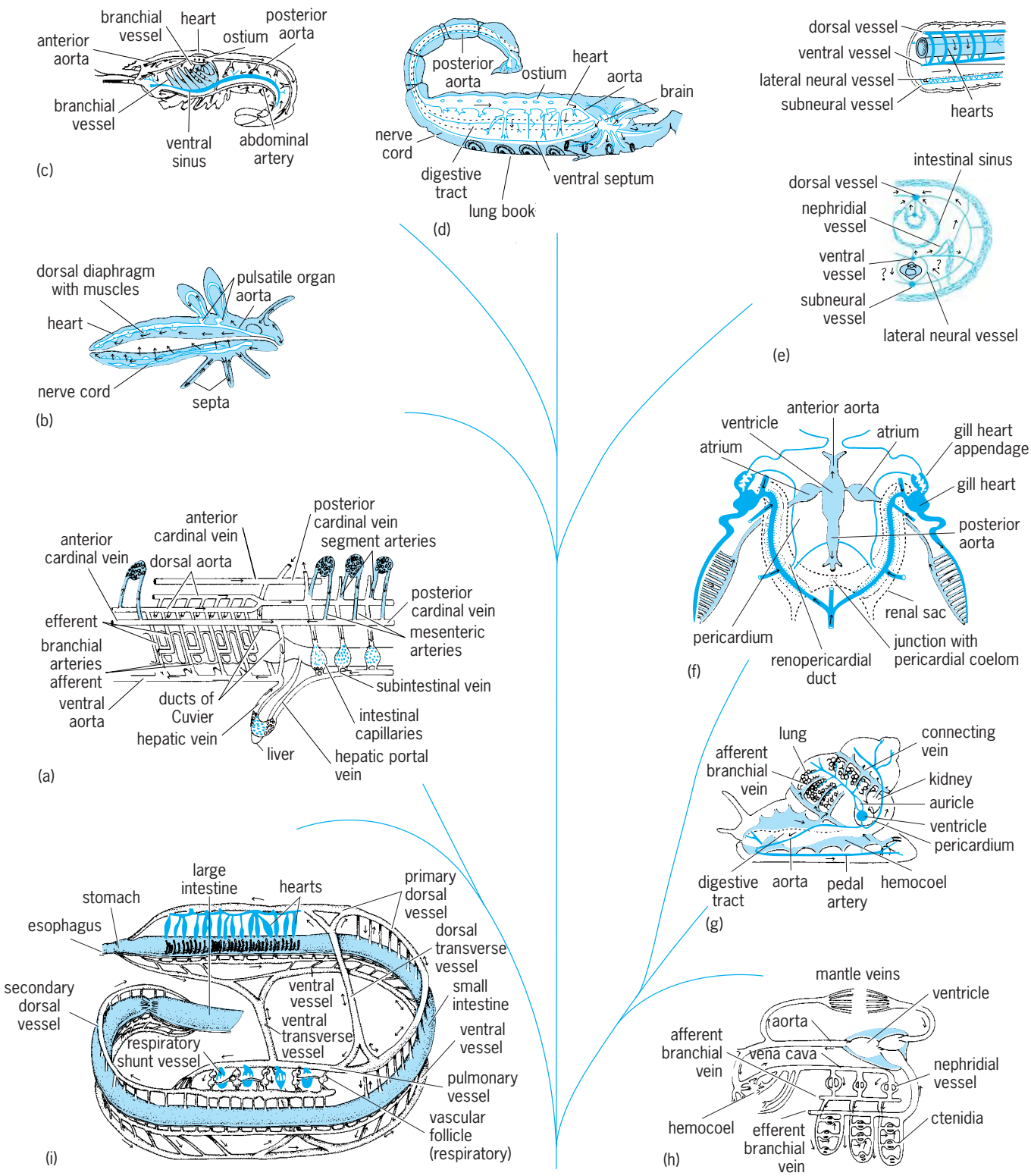

Fig. 2. Representative invertebrate and ancestral vertebrate circulatory systems on a schematic evolutionary tree. In all diagrams the arrows indicate the direction of blood flow. (*a*) Ancestral vertebrate, represented by *Amphioxus*. (*b*) Insect: a main dorsal vessel carries blood from the heart. (*c*) Crustacean: circulation in the lobster. (*d*) Arachnid, represented by the scorpion *Bathus*. (*e*) Annelid: cross section and longitudinal section. (*f*) Cephalopod mollusk, represented by the closed circulation in the octopus. (*g*) Gastropod mollusk: an open circulation represented by the snail *Helix*. (*h*) Bivalve mollusk: an open circulation represented by the fresh-water mussel *Anodonta*. (*i*) Echinoderm represented by the sea cucumber, *Stichopus*. (*Part a after A. J. Waterman et al.*, *Chordate Structure and Function*, *MacMillan*, *1971; b*, *c*, *e after K. Schmidt-Nielsen*, *Animal Physiology*, *2d ed.*, *Cambridge University Press*, *1979; d*, *g*, *h after W. F. Hamilton and P. Dow*, *eds.*, *Handbook of Physiology*, *Sec. 2*, *Circulation*, *vol. 3*, *American Physiological Society*, *1965; f after E. Florey*, *An Introduction to General and Comparative Animal Physiology*, *W. B. Saunders*, *1966; i after C. F. Herreid et al.*, *Blood vascular system of the sea cucumber*, *Stichopus moebi*, *J. Morphol.*, *150:423*, *1976*)

rates of 2–4 cm · s^{-1} about 10 times each minute. Vessels in the gills of hagfishes and octopods also contract rhythmically and propel the blood.

Multiple and chambered hearts. In the majority of animals, the zone of propulsive activity is restricted to a localized region. Some animals have many hearts: earthworms have five pairs of lateral hearts connecting the dorsal and ventral vessels (Fig. 2*e*); octopods and squids have hearts which boost blood flow through the gas exchanger (gills) in addition to a main or central heart (Fig. 2*f*); sea cucumbers may have over 150 hearts between dorsal and ventral vessels (Fig. 2*i*). However, many animals have only one heart, and in order to pump a volume equivalent to that pumped by a long contractile vessel, the lumen is greatly enlarged and usually divided into a number of chambers (chambered heart). To ensure undirectional flow of blood, chambered hearts require one-way valves. Valves are flaps of tissue situated around the edges of orifices between heart chambers or in vessels.

Contractile chambers. The more powerful the heart muscle becomes, the faster the blood can be driven into the outflow vessels, but filling the heart then requires a greater force. In vertebrates, there is often enough energy in the fast-flowing venous blood to do this, but other chambers may be necessary to fill the main propulsive chamber (or ventricle) in stages. These chambers are auricles and must have sufficiently thin walls to be expanded by the energy of the inflowing blood, yet must be powerful enough to fill the ventricle when they contract.

Most animals with chambered hearts have a single auricle, but in mollusks there is one for each pair of gills, so there may be from one to four (Fig. 2*f* and *h*). Fishes have a single auricle, but in amphibians the auricle is divided into two atria. In birds and mammals, the ventricle is divided, giving two separate circulations: one to the body and the other to the lungs. Fishes and amphibians have an extra contractile chamber, the sinus venosus, which fills the auricle when it contracts. Elasmobranchs and amphibians also have an extra contractile chamber on the ventricular outflow pathway, the conus arteriosus. Its role has not been established, but it could be in actively closing the outflow valves of the heart.

All vertebrate hearts, and many in invertebrates (such as mollusks and arthropods; Fig. 2*f* and *g*), are enclosed in a small portion of the coelom called the pericardial cavity. This cavity is fluid-filled and serves to prevent the heart from being buffeted by other body organs, but it can also aid in returning blood to the heart. If the pericardial wall is rigid, as in elasmobranchs and most mollusks and crustaceans, one chamber of the heart cannot contract unless another is expanding. If the ventricle starts to contract, the reduction in volume will tend to increase the volume of the pericardial cavity, and since it has rigid walls, the pressure inside it will be reduced. This suction pressure is transmitted through the thin auricular wall and draws blood in from outside the pericardium. This mechanism ensures successful filling of the heart even in open blood systems.

In tubular hearts, blood will tend to be sucked into the lumen when the muscle relaxes behind the wave of contraction. It has even been suggested that the major function of the dorsal contractile vessel in annelids is to prime the lateral hearts with blood. In insects, the tubular heart is suspended by alary or aliform muscles attached to the external skeleton, and when these contract the expansion of the lumen sucks blood through valved openings (ostia) into the heart (Fig. 2*b*).

Electrophysiology. The wave of contraction which passes across the heart is driven by an electrical impulse (action potential) which arises spontaneously, either in modified heart muscle cells, or in special nerve cells located on or near the heart. These cells form the pacemaker, which sets the heart rhythm. When nerve cells initiate the heartbeat, the pacemaker is called neurogenic. A pacemaker of muscle cells is called myogenic. Myogenic pacemakers are found in vertebrates, mollusks, echinoderms, and perhaps in some annelids. Neurogenic pacemakers are found in most arthropods. Many hearts have more than one pacemaker. In tunicates, the tubular heart pumps into the open circulation in two directions. After several hundred beats the rhythm slows and stops, and the direction of the wave of contraction reverses when the heart starts up again. Consequently, the tunicate heart must have at least two pacemakers. In frogs, heart contraction can be driven by electrical impulses arising in the sinus venosus, atria, or ventricle. Usually the sinus pacemaker beats fastest and drives all the others.

In birds and mammals, the heart is large and generates high pressures. To prevent the pressure set up by the contraction of one part of the heart from stretching, and perhaps rupturing, the relaxed part, all of the ventricular muscle must be activated simultaneously. This is achieved by conveying the electrical impulse to all parts of the ventricle through specialized conduction pathways composed of modified muscle cells. *See* CARDIAC ELECTROPHYSIOLOGY.

Chemical and mechanical excitation. Many hearts are innervated by nerves which regulate the beat. When these nerves are active, they liberate chemicals such as noradrenalin or acetylcholine from their terminals, and these neurotransmitters affect the cardiac muscles directly. In general, the pacemaker of myogenic hearts is inhibited by acetylcholine; excitation is produced by noradrenaline (vertebrates) or serotonin (mollusks). In neurogenic hearts, gamma aminobutyric acid (GABA) is inhibitory, whereas glutamic acid is excitatory. Neurotransmitters affect not only the rate but also the strength of cardiac contraction. For example, noradrenaline increases the amount of blood pumped by a single ventricle (stroke volume) while acetylcholine decreases stroke volume. *See* ACETYLCHOLINE; ENDOCRINE MECHANISMS.

Mechanical factors may also affect rate and strength of heart contraction. In earthworms, mollusks, and some fishes, stretching the pacemaker region of the myocardium increases its rhythm. Moreover, in all vertebrates, and probably in all mollusks as well, the length of the muscle fibers in the heart

at the end of diastole is directly proportional to the strength of the next muscle contraction (Starling's law of the heart).

Many animals, both vertebrate and invertebrate, take advantage of muscular activity in the body, such as during exercise, as an aid to circulation of the blood. In vertebrates, valved thin-walled veins run through blocks of skeletal muscle and are compressed when the muscle contracts, squeezing blood toward the heart (muscle pump). In other animals, the blood functions directly in locomotion. Spiders have no leg extensor muscles, so leg extension is brought about by forcing hemolymph into the leg at high pressure (50 kilopascals), which is generated by the contraction of lateral cephalothoracic muscles depressing the carapace. Backflow into the abdominal heart (which can generate pressures of only 13 kPa) is prevented by valves on the outflow vessels at the junction of the cephalothorax and abdomen. But there are no valves in the veins, and blood returns rapidly to the abdomen.

Caudal heart. Hagfishes have a unique caudal heart. A longitudinal rod of cartilage separates two chambers, and when body muscles on one side contract, the rod is bent to that side. The bend in the rod causes the volume of the chamber on the actively contracting side to expand so that it fills with blood, while the volume of the chamber on the other side is decreased, expelling blood. Valves at the inflow and outflow ends of the chambers ensure unidirectional flow of blood. *See* HEART (INVERTEBRATE); HEART (VERTEBRATE).

Blood vessels. Blood ejected from the heart is conveyed by vessels which are usually called arteries or aortas. In some animals, structural differences between outflow and inflow vessels are not obvious, and they are usually referred to simply as blood vessels. In closed systems, the arteries (outflow vessels) divide and subdivide so that ultimately tubes only a few micrometers thick (capillaries) run between the cells of each organ.

Capillaries. In vertebrates, each capillary consists of a single layer of endothelial cells surrounded by a basement membrane. In some echinoderms with well-developed circulatory systems (such as holothurians; Fig. 2*i*) capillary structure is similar to that in vertebrates, while in cephalopods (such as the octopus) capillaries are extremely small (1 micrometer in diameter) but are usually more than one cell thick. The total number of capillaries in a vertebrate is large, and in humans it is enormous: over 50 billion in the whole body. Therefore, although each capillary is small the total cross-sectional area of all capillaries is perhaps a thousand times greater than that of the main arterial vessels. Consequently, blood flows very slowly in the capillaries, about one-thousandth of the velocity in the aorta, allowing lots of time for exchange of materials with the cells.

In open circulations, after a greater or lesser number of branchings blood vessels ultimately open to the spaces between the cells. Even so, networks of fine blood pathways may occur, especially in gas-exchange organs, nephridia, and masses of nervous tissue (such as brains or ganglia). In animals with closed circulatory systems, a plexus of fine capillaries brings blood close to the external environment in skin (annelids, frogs), respiratory trees (holothurians), gills (fishes), and lungs. Similar close associations between blood and external media are seen in animals with open circulatory systems (Fig. 2*c* and *d*). For instance, gills of shrimp look similar to those of fish in all but the finest morphological details. Book lungs of spiders are unique structures in which blood is moved across rows of fine air tubes, stacked one on the other, in channels only a few micrometers in width. In insects, air is distributed directly to the body cells by way of the trachea, a system of tubes independent of the circulation. Fine blood channels are lacking in these animals. Nevertheless, the circulation in insects reduces the diffusion distances for metabolites, nutrients, and hormones. Furthermore, in large flying insects it plays an important role in controlling body temperature, particularly in the thoracic region where the large flight muscles are located.

Veins and portal systems. After traversing the capillaries, blood is collected into large vessels which return it to the heart. In cephalopods and vertebrates, these collecting vessels are called veins. In open circulations, channels returning blood to the heart exist, but the pattern of blood flow is less precise than in closed systems. In insects and arachnids, longitudinal membranes impose some direction on blood movement (Fig. 2*b* and *d*).

In many vertebrates, venous blood is transported to another organ (such as the liver or kidney) on its way back to the heart. In these organs the veins divide into a capillary network supplying blood to the tissues at low pressure. This is a portal blood supply, and there are a large number of these in vertebrates. The hepatic portal system occurs in all vertebrates and serves to transport materials from the intestine, where they are absorbed, to the liver, where they are stored (Fig. 2*a*). Another prominent portal system is the renal portal system. It occurs in all vertebrates except mammals, and transports blood from the caudal parts of the body to the kidneys. The function of the renal portal system is not known. *See* KIDNEY; LIVER.

Contraction. Many blood vessels have muscular coats, the muscle cells being rhythmically contractile (myogenic rhythm). Contraction of blood vessels is found in the dorsal and ventral vessels and some capillaries of annelids, in the gill vessels of cephalopods and cyclostomes, and also in the fine vessels leading to the capillary beds in higher vertebrates. In these small vertebrate vessels (called arterioles, with a diameter of 50–200 μm), the muscles are in a continual state of activity, modulating the radius of the vessel and therefore the blood flow. This activity can be altered by nerves, by blood-borne agents such as adrenaline, and also at the local level by changes in the level of tissue metabolites (a fall in oxygen tension in the body tissues causes the muscles to relax).

Aortic elasticity. The aortas (or arteries leaving the heart) are fairly elastic in nearly all animals. In closed

circulations, the stroke volume is stored by the elastic distension of the arteries and is fed to the periphery between cardiac ejections. This elastic reservoir (or Windkessel effect) is exemplified by the arterial vessels of vertebrates. In humans, the whole arterial tree participates in producing a Windkessel effect. In contrast, fishes have a very short ventral aorta, and only a very elastic bulb (bulbus cordis) just outside the ventricle provides a pressure reservoir, which smooths blood flow through the gills. Many mollusks with open circulatory systems have a similar elastic expansion of the aortic wall just beyond the ventricle. *See* BLOOD VESSELS.

Evolution of Vertebrate Circulation

The evolution of an organ system which is not preserved in fossils can only be inferred from investigations of extant species. Thus, understanding must always be clouded by uncertainty. For example, a major pattern of increasing complexity, which emerges from comparative studies of circulatory systems, is related in vertebrates to life on land. More promising are investigations of the early stages of development of the heart and circulation. These show that all types of vertebrate circulations can be related to a common ancestral pattern (Fig. 2*a*). The ancestral pattern of circulation in vertebrates is for a ventral heart to pump blood anteriorly in a ventral vessel, connecting to the dorsal vessel through the gas exchanger (gills; Fig. 1*e*). The gills are situated on the high-pressure side of the circulation, and blood flows posteriorly in the dorsal vessel.

Annelid worms exhibit a superficial resemblance to the circulation in the ancestral vertebrate in the pattern of dorsal and ventral vessels (Fig. 1*c*), except that the heart is dorsal, flow in the dorsal vessels is anteriorly directed, and flow in the ventral vessel is posteriorly directed. More important, in most invertebrates the gas exchanger, when it exists as a discrete entity, is on the low-pressure side of the circulation. This arrangement limits metabolic scope, a problem which is solved in large, active invertebrates (such as cephalopods) by the development of high-pressure, prebranchial hearts.

Apparently, the chordate stock arose from ancestors in common with primitive echinoderms, a group with a very poorly developed circulation. Hence, the apparent similarities between the circulations in vertebrates and invertebrates are superficial and presumably arose as an evolutionary response to similar physiological demands. This is an example of convergent evolution due to common selection pressures, the circulatory systems in vertebrate and invertebrate groups being analogous rather than homologous. *See* CARDIOVASCULAR SYSTEM. David R. Jones

Circulation disorders

Circulation disorders include both localized disturbances of circulation and diseases of the circulatory system.

Circulatory system. The function of the circulatory system is to transport and distribute substances either used or produced by cells or both. Excluded are those materials that are discharged directly from sweat glands, digestive glands, and renal tubule cells. Included, however, are nutritive and metabolic substances, hormones, waste products, water, and heat.

The circulating fluids are distributed between the heart, blood vessels, capillaries, intercellular spaces and lymphatics and within the cells. The blood and body fluids are in a constant state of dynamic equilibrium (see **illus.**). Blood flows through the vessels to the capillaries where some fluid passes through the endothelium to the intercellular spaces to bathe the cells. Fluid with contained electrolytes and metabolic substances can then pass into and out of the cells. At the distal end of the capillaries most of this fluid passes back into the vascular compartment. A portion of this fluid is returned to the circulation by way of the lymphatics. *See* CARDIOVASCULAR SYSTEM; LYMPHATIC SYSTEM.

Vascular disturbances. Disturbances in this pattern can either result in or from disease conditions. An example is edema, which is an abnormal accumulation of fluid in the cells, tissue spaces, or cavities of the body. There are three main factors in the formation of generalized edema and a fourth which plays a role in the formation of local edema. They are the permeability of the capillary wall, the colloid osmotic pressure of the plasma proteins, and the hydrostatic pressure in the capillaries. The fourth factor, which is of importance in local edema formation, is lymphatic obstruction.

An example of edema formation secondary to a decrease in the colloid osmotic pressure of the plasma proteins is nephrotic edema. In chronic Bright's disease there is a marked loss of urine albumin and the plasma protein level drops as a result.

Cardiac edema follows the generalized venous congestion of cardiac failure and is an example of edema formation resulting from increased hydrostatic pressure in the capillaries.

Increased permeability of the capillary walls plays an important role in the formation of inflammatory edema, the edema of severe infections, metabolic intoxications, asphyxia, anaphylactic reactions, secondary shock, and acute nephritis. *See* EDEMA; HYPERTENSION.

Deficiencies. A deficiency of circulating blood volume, both cellular elements and fluid, is called oligemia. This may be the result of an acute blood loss or it may be of a chronic nature, such as an anemia combined with dehydration. Anemia, or oligochromemia, is a deficiency of circulating red cell volume or, more specifically, hemoglobin content. This is the oxygen-carrying component of the circulating red blood cell. The normal value for the adult human is 14–16 g/100 ml of blood. The norm for most domestic animals is 10–12 g/100 ml of blood, except for the dog with a norm of 13 g. *See* ANEMIA; HEMOGLOBIN.

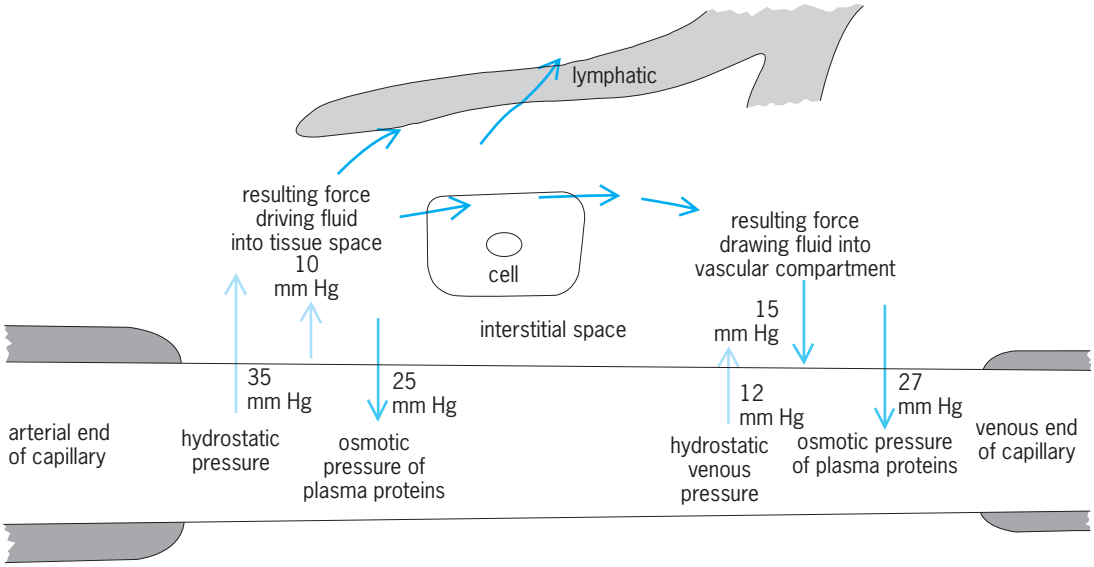

arterial end of capillary		venous end of capillary	
hydrostatic pressure (force tending to drive fluid out)	= 35 mm Hg	hydrostatic pressure	= 12 mm Hg
osmotic pressure of proteins (force tending to draw fluid into vascular compartment)	= 25 mm Hg	osmotic pressure of proteins	= 27 mm Hg
resulting force pushing fluid out of vascular compartment into interstitial space	= 10 mm Hg	resulting force drawing fluid back into vascular compartment	= 15 mm Hg

Fluid movements and resulting pressures at the capillary level of circulating blood.

Anemia can result from a variety of causes. It may be secondary to a chronic nutritional deficiency with an inadequate intake of iron. Iron is a necessary substance used by the blood-forming tissue to make hemoglobin. Anemia can also result from a defect in the blood-forming tissue (the bone marrow) so that it is unable to use the materials available to make red blood cells. It can also be the result of a displacement of blood-forming tissue, as from a tumor. It may also follow periods of acute blood loss. These are just a few examples of the causes of anemia. *See* BLOOD; HEMATOPOIESIS.

Pancytopenia, or oligocythemia, is a deficiency of all circulating blood cellular elements. This is usually the result of a deficiency of the blood-forming tissue, the bone marrow.

Ischemia. The decrease of blood flow to an organ or tissue is known as ischemia. This can be sudden as when a vessel is ligated, when a thrombus or blood clot forms, or when an embolus comes to lodge in the vessel. A gradual occlusion can follow arteriosclerotic changes in the vessel wall with resulting diminution of the lumen and blood-carrying capacity of the vessel. The effect of a sudden occlusion depends to a great extent on the collateral circulation to the organ or tissue involved. If an adequate collateral circulation comes to be established, the tissue survives; it not, it dies and an infarct results.

An infarct is a region of necrosis resulting from a vascular obstruction. The usual cause is occlusion of an artery or vein by a thrombus or embolus. If the vessel is the sole source of blood supply, the tissue dies. *See* ARTERIOSCLEROSIS; EMBOLISM; INFARCTION; THROMBOSIS.

Plethora. An excess of blood is referred to as plethora. This increase may be the result of an increase in the size or the number of red blood cells. The increase in red cell volume may be a polycythemia vera, or true polycythemia, which is a primary increase in the number of red blood cells with no regard to the needs of the organism. Polycythemia, or erythrocytosis, is usually a secondary increase in red cells following conditions of chronic hypoxia, such as congenital heart disease, emphysema, or residence at high altitudes. Serous plethora is an increase in the watery part of the blood.

Hyperemia. Hyperemia, or congestion, refers to an excess of blood within an organ or tissue. This condition may be localized or generalized. Active hyperemia is congestion which is caused by an active dilatation of arterioles and capillaries. This occurs under certain physiological conditions, such as in the muscle when there is an increased need for blood during exercise. It also occurs in pathological states such as inflammation. *See* INFLAMMATION.

Passive hyperemia is a condition which results in an accumulation of blood in the venous system. It may be generalized or localized and can result from any obstruction or hindrance to the outflow of blood from the venous circuit. The generalized form can result from valvular diseases of the heart, such as mitral or aortic stenosis or insufficiency, or myocardial

failure from any cause. Tricuspid valvular disease is slightly more frequent in animals. *See* HEART DISORDERS.

Diseases of the lungs such as emphysema, fibrosis, or pulmonary hypertension of any origin can result in right ventricular failure and generalized venous congestion. The late stages of these conditions can result in cor pulmonale, which is a hypertrophy and dilatation of the right ventricle secondary to an obstruction to the pulmonary blood flow associated with a generalized venous congestion.

Localized venous congestion results when a main vein from a region is occluded either by a thrombus or some extrinsic pressure such as a tumor or enlarged lymph nodes.

In congestion the organ is usually wet and bloody. Microscopically there is an enlargement of the veins and capillaries associated with edema formation.

Hemorrhage. The escape of blood from within the vascular system is hemorrhage. This process can be the result of trauma to, or disease of, the vessel wall. Diapedesis is the passage of the corpuscular elements of the blood through an intact vessel wall.

The trauma which results in hemorrhage may be direct, such as a cut from a knife, or it may be indirect, such as a blow to the head, resulting in a tearing of cerebral or meningeal vessels.

The causes of hemorrhage other than trauma can be divided into three main groups. In the first group are those conditions in which there is a disease process affecting the vessel wall, such as arteriosclerosis or aneurysm formation. An infarct, or tissue death from any case, can result in hemorrhage.

In the second group are conditions in which there is an acute process affecting the vessel wall, as in septicemia, poisoning by heavy metals, or even anoxia.

The third and last group consists of those conditions in which there is a defect in the blood itself, which results in hemorrhage. Under this heading are included the blood-clotting disorders, thrombocytopenia, leukemia, and pernicious anemia.

Apoplexy, or stroke, is an acute vascular lesion of the brain. This can be the result of hemorrhage of, thrombosis in, or embolism to a cerebral vessel. *See* HEMORRHAGE.

Thrombosis. Thrombosis is the formation of a thrombus, which is a solid body formed during life and composed of the elements of the blood: platelets, fibrin, red cells, and leukocytes.

Thrombosis is essentially platelet deposition and may occur on a vessel wall anywhere that the endothelium is damaged. However, because the platelets release thromboplastinogen, which activates the clotting mechanism, thrombosis and blood coagulation may occur together. Platelets normally exhibit a tendency to stick together. This tendency is increased if the number of platelets is increased, as in thrombocytosis, if the velocity of the stream is decreased below a certain speed, or if the endothelium is roughened. These are the same factors which promote thrombus formation.

Thrombi may form in the veins, arteries, capillaries, or within the heart. The most common sites for the formation of thrombi are in the veins.

The venous thromboses can be divided into two groups: venous thrombosis (phlebothrombosis) and thrombophlebitis. Phlebothrombosis is the more important and is often a sequela of generalized circulatory failure, trauma, or prolonged bed rest from any cause.

Thrombophlebitis is the formation of a thrombus secondary to inflammation in the wall of the vein. Arterial thrombosis is usually the result of local causes, for example, arteriosclerosis of the vessel wall.

Embolism. The sudden blocking of an artery or vein by a clot or other substance which has been brought to its place by the blood current is an embolism. The material carried in the circulation in this process is an embolus. Emboli may be composed of thrombi, fat, air, tumor cells, masses of bacteria or parasites, bone marrow, amniotic fluid, or atheromatous material from the vessel wall.

Emboli originating in the right side of the heart or the great venous system of the body come to lodge in the lungs, those from the portal system in the liver, and those from the pulmonary veins or left side of the heart in some segment of the peripheral arterial tree.

One of the dire consequences of an embolus is infarction or death of the tissue supplied by the occluded vessel. This occurs if the collateral circulation is inadequate to supply the tissue with blood. *See* DEATH; VASCULAR DISORDERS. Romeo A. Vidone

Bibliography. E. Chung, *Quick Reference to Cardiovascular Disease*, 3d ed., 1987; W. R. Felix, Jr., *Noninvasive Diagnosis of Peripheral Vascular Disease*, 1987; S. B. McKenzie, *Textbook of Hematology*, 1988; N. C. Staub and A. E. Taylor (eds.), *Edema*, 1984; J. H. Wood, *Cerebral Blood Flow: Physiologic and Clinical Aspects*, 1987.

Cirrhosis

Severe scarring of the liver due to a marked increase in fibrous connective tissue, resulting in a firm, nodular, distorted liver. Cirrhosis is the eighth most common cause of death in the United States and the fourth leading cause of death among Americans between the ages of 30 and 60.

Etiology. Cirrhosis is the end product of any one of a number of different liver diseases. Regardless of the cause, the potential consequences of cirrhosis are the same. Whereas a healthy liver typically repairs and regenerates itself when injured, once cirrhosis has occurred, the damage may never be undone. Not all liver diseases cause cirrhosis—only those that cause chronic, ongoing damage to the liver. Alcoholic liver disease and chronic hepatitis C are the two most common causes of cirrhosis in the United States. However, there are numerous causes of cirrhosis as listed below.

Viral hepatitis—hepatitis B, C, and D
Autoimmune hepatitis (a disease in which the body's immune system attacks liver cells)

- Primary biliary cirrhosis (a liver disease that slowly destroys the bile ducts in the liver)
- Nonalcoholic fatty liver disease [a range of liver diseases that commonly affect people who have type 2 diabetes (adult onset diabetes), with hypertriglyceridemia (high level of fats, or triglycerides, in the blood), and who are obese but drink little or no alcohol)]
- Excessive alcohol consumption
- Hemochromatosis (excessive deposition of iron in the body that may be genetic or acquired)
- Excessive intake of vitamins, such as vitamin A
- Certain herbal remedies, such as comfrey
- Certain medications, such as methotrexate, isoniazid, methyldopa (hypertension)
- Primary sclerosing cholangitis (a liver disease characterized by inflammation, narrowing, and scarring of the bile ducts inside and outside the liver)
- Vascular anomalies (for example, Budd-Chiari syndrome)
- Congestive heart failure
- Wilson's disease (a genetic disorder of copper overload)

See ALCOHOLISM; HEPATITIS; LIVER DISORDERS.

Symptoms and complications. The symptoms of cirrhosis are essentially uniform regardless of its cause. Some people with cirrhosis feel perfectly normal or have nonspecific symptoms, such as fatigue, decreased appetite, and loss of libido. These individuals are said to have compensated cirrhosis. People with compensated cirrhosis often live a normal life span with relatively few health-related consequences. However, they remain at risk for developing many life-threatening complications such as portal hypertension, ascites, encephalopathy, jaundice, hepatorenal syndrome, and hepatocellular carcinoma. When one of these complications has occurred, individuals are said to have decompensated cirrhosis.

Portal hypertension. Due to extensive scarring of the liver that occurs in cirrhosis, the vessels associated with blood flow to and from the liver may become obstructed. This, in combination with increased blood flow to the liver, leads to elevated pressure in the portal circulation, a condition known as portal hypertension. The liver attempts to adapt to this situation by creating alternative routes that bypass this obstruction. These alternative passageways for blood flow are known as collateral shunts or simply "collaterals." These shunts enable blood to be rerouted to and circulated within the rest of the body. Unfortunately, the formation of these shunts has its drawbacks, as they can give rise to serious and even life-threatening complications. When any one of the complications associated with portal hypertension (for example, ascites, esophageal varices, and/or encephalopathy) has occurred, the body can no longer "compensate" for the extensive scarring that has occurred in the liver, and the patient has decompensated cirrhosis. People with any form of portal hypertension should be evaluated for liver transplantation.

Portal hypertension can result in congestion of the spleen, leading to a markedly enlarged spleen (splenomegaly).

Esophageal varices portal hypertension may result in the formation of new collateral venous channels that bypass the liver. The most important collateral channels are those in the region of the lower esophagus, called esophageal varices. When varices become dilated, they can erode through the lining of the esophagus, resulting in life-threatening bleeding.

Ascites is the most common complication of portal hypertension. It is characterized by massive accumulation of fluid in the peritoneal cavity (the space between the abdominal organs and the skin). A fever in a person with ascites may indicate that this fluid is infected and constitutes a serious condition known as spontaneous bacterial peritonitis.

Encephalopathy. Encephalopathy is a disorder of brain function, typically leading to a coma that can occur in people with cirrhosis. It is mainly associated with poor coordination, fetor hepaticus (foul-smelling breath), and asterixis (uncontrollable flapping of the hands). The exact cause of encephalopathy is unknown; most researchers believe that it is mainly associated with the ailing liver's inability to clear toxins, primarily ammonia, from the body.

Jaundice. In addition to obstruction of blood flow, the bile ducts within the liver may become distorted and partially obstructed. This results in jaundice, a yellow discoloration of the skin and eyes. *See* JAUNDICE.

Hepatorenal syndrome. Some individuals with cirrhosis develop hepatorenal syndrome, a progressive deterioration of kidney function associated with lack of urination. This may lead to renal failure, and kidney dialysis may become necessary. Hepatorenal syndrome can result in death unless a liver transplantation is performed. *See* KIDNEY DISORDERS.

Hepatocellular carcinoma. Anyone who has cirrhosis, whether compensated or decompensated, is at risk for developing liver cancer—also known as hepatocellular carcinoma or hepatoma. Risk varies with the cause of liver disease. *See* CANCER (MEDICINE).

Morphological changes. In general, cirrhosis results in similar physical damage to the liver, regardless of which type of liver disease caused it. The liver may be slightly enlarged, but as the disease progresses it usually becomes smaller due to progressive loss of liver cells. In the final stage the liver may weigh as little as 14–18 oz (400–500 g; normal is 50–60 oz or 1400–1700 g). Microscopic examination reveals that liver cells of varying size and shape are surrounded by dense fibrous connective tissue that contains inflammatory cells and distorted bile ducts. Fibrosis, or scar tissue, is the liver's effort to keep contained the damage done by alcohol or hepatitis C virus, for example. But scar tissue can block the blood flow through the liver, resulting in an inability to perform its normal duties. *See* LIVER.

Diagnosis. A liver biopsy, the removal of a tiny piece of liver tissue using a special needle,

continues to be the "the gold standard" to determine the presence or absence of cirrhosis. However, signs of cirrhosis such as muscle wasting, a hard and bumpy liver, and splenomegaly may be detected on a physical exam. Some laboratory test results (such as a low albumin level, an abnormally low cholesterol level, an elevated or prolonged prothrombin time, a decreased platelet count, and/or an elevated alpha-fetoprotein level) may also suggest cirrhosis. Imaging studies such as a sonogram or computerized tomography (CT) scan of the liver may suggest, but cannot confirm, cirrhosis. Once cirrhosis has occurred, a physician will not be able to definitively identify which liver disorder caused it, even from a liver biopsy specimen.

Treatment. Recent studies have shown that with removal of the cause of the underlying liver disease (for example, alcohol or the hepatitis C virus), and with effective treatment (such as interferon and ribavirin for hepatitis C virus), cirrhosis can be reversed—at least in its very early stages, when scarring is minimal. The point at which cirrhosis becomes irreversible is not clear. However, there is a broad consensus among liver disease experts that once decompensated cirrhosis has occurred, it can never be reversed. Much research is being conducted on drugs (known as antifibrotic agents) and drug combinations that can potentially reverse cirrhosis. Examples of promising antifibrotic agents include interferon, pentoxifylline, prostaglandins, antioxidants, pioglitazone, and angiogenesis inhibitors. Currently, therapy of cirrhosis is aimed primarily at preventing or reducing complications. However, while the complications of decompensated cirrhosis can be controlled, the liver is beyond the point of any possible repair, and a liver transplant must be considered. *See* CARDIOVASCULAR SYSTEM; LIVER DISORDERS.

Melissa Palmer

Bibliography. P. Bonis, Is liver fibrosis reversible?, *NEJM*, vol. 344, no. 6, February 2001; *Liver Disease: From Bench To Bedside*, AASLD 55th Annual Meeting, 2004; M. Palmer, *Dr. Melissa Palmer's Guide to Hepatitis and Liver Disease*, rev. ed., Avery Penguin Group, May 2004; *Schiff's Diseases of the Liver*, 9th ed., Lippincott Williams and Wilkins, January 2003.

Cirripedia

An order of the crustacean class Maxillopoda. The cirripedes, particularly the organisms known as barnacles, are the most well-known and easily recognized members of the subclass Thecostraca. The thecostracans are characterized by two features. First, all exhibit a distinctive and unique larva, called the cypris (**Fig. 1***b*). Second, when they have a carapace, it generally envelops the entire body, which typically is composed of a five-segment head, a seven-segment thorax, and a short or rudimentary limbless abdomen. In addition to the cirripedes, the thecostracans include the enigmatic Facetotecta, or "y-larvae," as well as Ascothoracica, a group of parasites of echinoderms and anthozoan cnidarians. *See* ASCOTHORACICA; CRUSTACEA; MAXILLOPODA.

The Cirripedia sensu stricto are composed of three infraorders: Thoracica (the goose and acorn barnacles), Acrothoracica (the burrowing barnacles), and Rhizocephala (internal parasites of shrimp and crabs) [see **table**].

Thoracica and Acrothoracica

Generally, the two free-living cirripede infraorders bear long, setose, filtering thoracic limbs called cirri. In addition, cirripedes have nauplius larvae that bear peculiar structures called frontolateral horns (Fig. 1*a*), whose function is still not certain. As adults, the free-living cirripedes are permanently attached and commonly referred to as barnacles.

Morphology. The carapace forms a complete cover of the body, and it is usually strengthened by calcareous shell plates (**Fig. 2**). In principle, the head bears the usual five pairs of limbs. However, the antennules and antennae are seen only in the larvae (Fig. 1). The antennules form the points of attachment of the adult barnacle but persist only in adult goose barnacles as the peduncular stalk (Fig. 1*d*). The mouth appendages include the paired mandibles with palps, maxillules, and maxillae. The thorax bears up to six pairs of biramous appendages (cirri) composed of numerous segments, each with a considerable number of setae (Fig. 2*b*). A rostral vessel and/or sinus (Fig. 2*b*) serve as a heart that pumps blood through a virtually closed circulatory system. Naupliar and compound eyes occur only in the larvae (Fig. 1*b*).

Reproduction and development. Thoracican adults are typically hermaphrodites, whereas acrothoracicans have separate sexes. The female genital pore is on the first thoracic segment at the base of cirrus I,

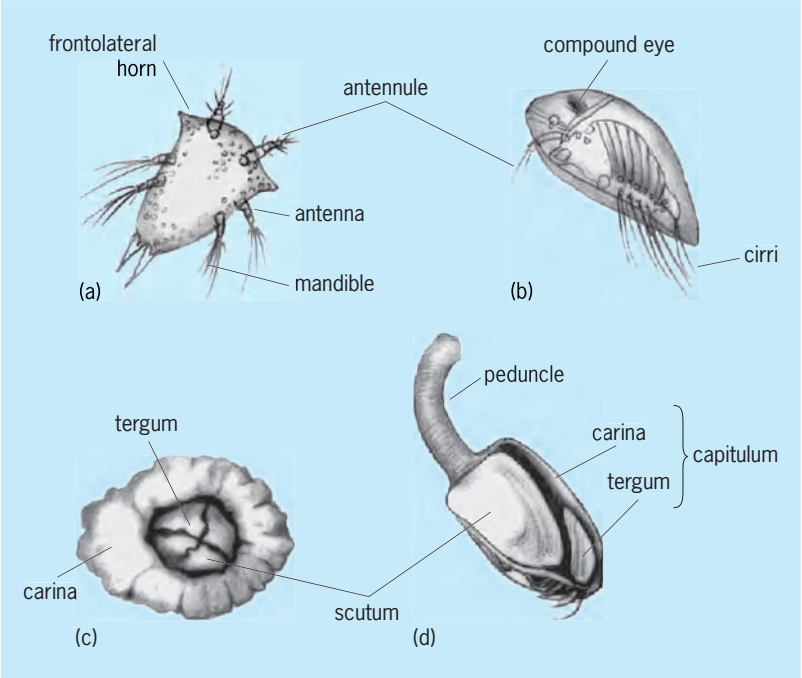

Fig. 1. Larval and adult stages of thoracican barnacles. (*a*) Nauplius larva (note diagnostic frontolateral horns). (*b*) Cypris larva. (*c*) *Balanus*, a typical acorn barnacle (top view). (*d*) *Lepas*, the goose barnacle (lateral view).

Characteristics and some genera of Cirripedia

Infraorder	Characteristics	Some genera
Thoracica	Permanently attached, stalked or sessile, independent or commensal barnacles; with filtering thoracic limbs as long setose cirri; typically with carapace developed as calcareous plates; hermaphroditic	*Lepas, Scalpellum, Conchoderma, Heteralepas, Balanus, Coronula, Verruca*
Acrothoracica	Burrowing barnacles living in shells, coral, or limestone; cirri modified or often reduced in number; carapace without calcareous plates; bisexual with dwarf males	*Trypetesa, Lithoglyptes, Kochlorine*
Rhizocephala	Internal parasites of decapod crustaceans; adult body is a formless ramifying mass among the host tissues; no gut; eggs and larvae produced in an external sac on host's body; bisexual with dwarf males	*Sacculina, Loxothylacus, Thompsonia*

while the male genital pore is on an elongated penis lying medially on the segment posterior to cirrus VI. Cross-fertilization of the hermaphrodites is normal. The extremely long penis is extended out of the shell to grope around for a receptive neighbor. The eggs are incubated in the shell cavity and hatch into free-swimming nauplius larvae (Fig. 1*a*). Thoracicans pass through six successive naupliar stages, whereas acrothoracicans have four. The final nauplius molts into an entirely different form with a bivalved carapace, the cypris larva (Fig. 1*b*), so named because it resembles the ostracode genus *Cypris*. The cypris has prominent antennules and well-developed, though short, cirri. The cypris seeks a suitable substrate or host and attaches by the antennules. Thereafter, an extremely strong cement is secreted from glands in the antennules, and the animal is permanently anchored.

The attached cyprid then undergoes metamorphosis, and a minute replica of the adult (only without calcified shell plates) is produced. By differential growth and elongation of the preoral region, the free body of the animal within the body cavity comes to lie parallel to the substrate with the ventral side uppermost. The thoracic appendages, or cirri, face away from the substrate and are immediately below the opening of the shell, which is uppermost (Fig. 2*b*). Through this aperture, the cirri are extended and unfolded into a fanlike array. The cirri sweep through the water in the manner of a cast net and withdrawn back into the shell with whatever has been netted. Edible particles are then removed from the cirri by the mouthparts and passed to the mouth to be swallowed.

Stalked thoracicans. In the stalked thoracican, or goose barnacles, the antennular region elongates and forms a stalk, or peduncle. At the distal end is the capitulum with the rest of the animal's body inside. In *Lepas* (Fig. 1*d*) the penduncle may be up to 1 ft (30 cm) long and the capitulum only 2–3 in. (5–8 cm). The capitulum in *Lepas* is strengthened by five calcareous plates: thecarina, or keel, extends

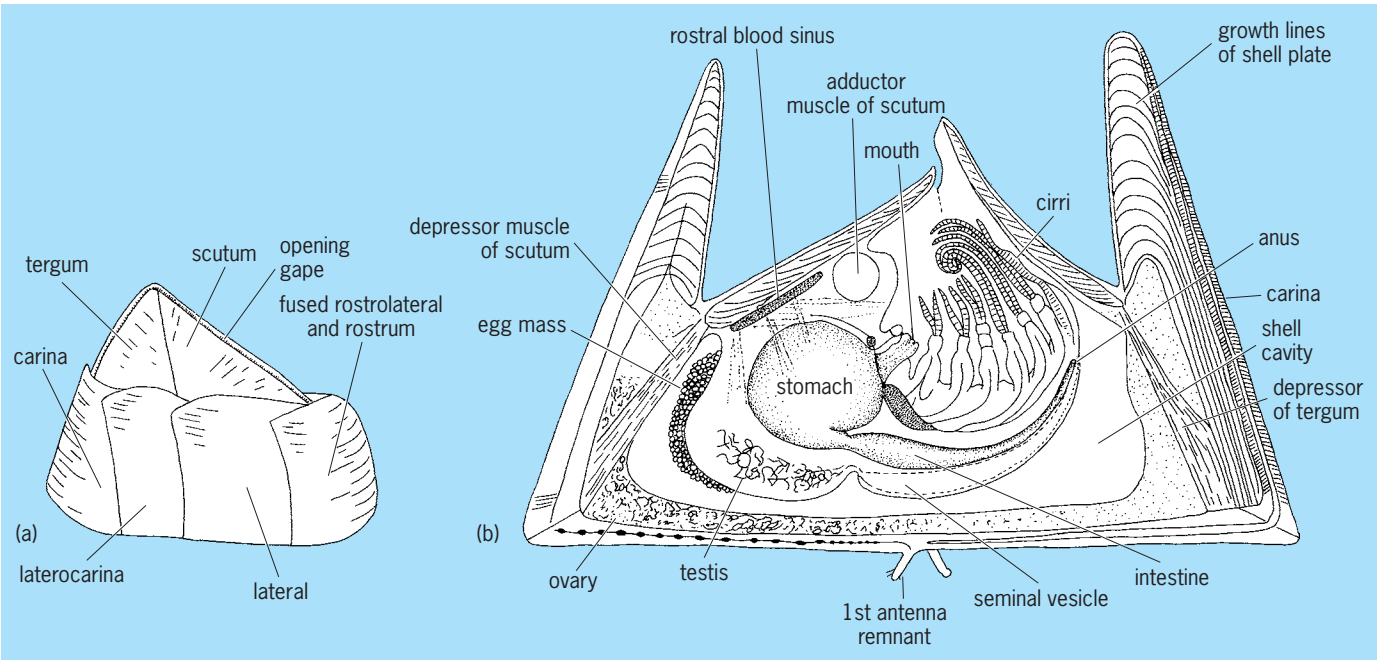

Fig. 2. Diagrams of the anatomy of *Balanus*, a sessile acorn barnacle. (Note the lack of a stalk.) (*a*) Lateral view. (*b*) Vertical cross section. (*After E. R. Ruppert, R. S. Fox, and R. D. Barnes, Invertebrate Zoology: A Functional Evolutionary Approach, 7th ed., Brooks Cole, 2003*)

down the posterior or morphologically dorsal side of the capitulum; and paired scuta and terga lie on either side of the capitulum. Calcification occurs from centers all over the capitulum and peduncle, giving rise to small oval or triangular scales. In many species of *Scalpellum* (sensu lato), in which these peduncular scales are present, there is a ring of larger scales around the base of the capitulum.

The number of capitular plates can be variable both within and between species. However, forms such as the fouling barnacle *Conchoderma* and the commensal of sea urchins *Heteralepas* are protected by the spines of its host and have both lost most of the capitular plates. *Lepas* and *Conchoderma* are found on floating debris in the ocean. *Lepas fascicularis* actually secretes a horny float that supports a cluster of individuals. The more heavily armored species of scalpellids are found on the bottom of the ocean and extend into very deep waters.

The peduculate goose barnacles are not a single monophyletic group. Rather, they form a transition series of distinct lineages leading to the sessile acorn barnacles.

Sessile thoracicans. Most of the sessile thoracicans, or acorn barnacles, live in shallow waters or between the tide marks, where they can be present in immense numbers. They lack any stalk and can be thought of as composed only of a capitulum (Figs. 1*c* and 2). These intertidal forms have economic importance. They settle on any suitable immersed surface, including the hulls of ships and submerged equipment. Their presence interferes with efficient operation and increases running costs. Large sums have been spent devising antifouling paints, principally using tributyl tin, and other techniques to discourage settlement of barnacles. Typically, the shell is formed of six to eight plates and a basal membrane, which may be calcified. The opening of the shell is closed by strong depressor muscles of a pair of movable lids (Fig. 2*a*), each composed of two articulated plates homologous with the scutum and tergum seen in the *Lepas* capitulum. Apart from ship bottoms, rocks, docks, pilings, and shells, acorn barnacles also cling to turtles (*Chelonibia*) and whales (*Cornonula*, *Tubincinella*). *See* THORACICA.

Acrothoracicans. The burrowing acrothoracicans are cryptic forms, millimeters in size. The settled cypris does not develop shell plates on the carapace during metamorphosis. Rather, the carapace bears chitinous teeth that assist in excavating the adult's burrow in the calcarous substrates on which they settle. Although some chemical action in this process has been suspected, apparently most of the burrow excavation is achieved by mechanical abrasion. Once properly housed, virtually little of the adult body ever extends beyond the burrow entrance, with all life processes taking place within the adults snug little home. *See* ACROTHORACICA.

Rhizocephala

The parasitic rhizocephalans can be identified as cirripedes only because of their nauplius larva, which bears frontolateral horns, and the possession of the thecostracan cypris larva. The cypris injects a worm-like stage into a decapod crustacean host, and this grows into an entangling network of female parasitic tissue that effectively takes over the host body. When the time comes to reproduce, an external gonad sac is protruded under the pleon (limbed abdomen) of the host. This attracts a male cypris that settles on the externa and becomes a hyperparasite on the female. Since they do not have an adult body form that can be compared with that of any other crustacean, let alone a thecostracan, the relationship of the rhizocephalans to the barnacles was unclear for a century or more, and until recently most researchers thought they were a group quite separate from the barnacles. However, analysis of DNA sequence data has shown that the rhizocephalans are closely related as a sister group to the thoracicans within Cirripedia. *See* RHIZOCEPHALA.

Fossil Record

Unquestioned fossil remains of Cirripedia are known from the Silurian, but there are also possible plateless forms that occurred earlier in the Cambrian Burgess Shale of Canada. Proper stalked barnacles are known from the Late Paleozoic onward. In Mesozoic and Cenozoic deposits, disarticulated shell plates occur, but only a few intact capitular shells are recognized. *See* FOSSIL; GEOLOGIC TIME SCALE.

Frederick R. Schram

Bibliography. D. T. Anderson, *Barnacles: Structure, Function, Development, and Evolution*, Chapman & Hall, London, 1994; J. T. Høeg et al., New evidence on the basic phylogeny of the Cirripedia Thoracica, in *Crustaceans and the Biodiversity Crisis*, ed. by F. R. Schram and J. C. von Vaupel Klein, pp. 101–114, Koninklijke Brill, Leiden, 1999; F. R. Schram, *Crustacea*, Oxford University Press, New York, 1986; J. T. Tomlinson, The burrowing barnacles, *Bull. U.S. Nat. Mus.*, 296:1–162, 1969.

Citric acid

A hydroxytricarboxylic acid, general formula $C_6H_8O_7$, with the structure shown below. It is

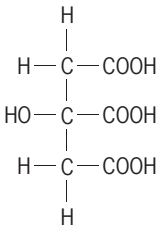

also available primarily as anhydrous material but as the monohydrate. The major commercial salts are sodium and potassium, with calcium, diammonium, and ferric ammonium (complex) also available. *See* ACID AND BASE.

Citric acid and its salts are widely used because they are nontoxic, safe to handle, and easily biodegraded. They are universally accepted, without

TABLE 1. Citrate content of natural products

Product	Percent citric acid
Lemon juice	4.0–8.0
Orange juice	0.6–1.0
Tangerine juice	0.9–1.2
Grapefruit juice	1.2–2.1
Strawberry	0.65–0.82
Raspberry	1.1–1.3
Black currant	1.5–3.0
Red currant	0.7–1.4
Gooseberry	0.98–0.99
Tomato	0.25
Cabbage	0.05–0.07
Corn	0.02
Eggplant	0.01
Peas	0.05
Cow milk	0.08–0.23
Human milk	0.35–1.25
Chicken organs	0.01–0.07

restriction on level, as food ingredients, including acceptance by the Joint FAO/WHO Expert Committee on Food Additives and the U.S. Food and Drug Administration. Because of their food use, specifications for purity are included in the *Food Chemicals Codex* (except diammonium citrate). In addition, specifications for citric acid, sodium citrate, and potassium citrate are included in the *U.S Pharmacopeia* and many pharmacopeia throughout the world.

Occurrence. Citric acid occurs in relatively large quantities in citrus fruits. It also occurs in other fruits, in vegetables, and in animal tissues and fluids either as the free acid or as citrate ion (**Table 1**). It is an integral part of the Krebs (citric acid) cycle involving the metabolic conversion of carbohydrates, fats, and proteins in most living organisms. *See* CITRIC ACID CYCLE.

Manufacture. Prior to World War I, all of the world's citric acid was produced by extraction from cull citrus fruit. Today, essentially all of the commercial citric acid is produced by fermentation. Processes employed are surface or submerged fermentation by mold (*Aspergillus niger*) and submerged fermentation by yeast (*Candida guilliermondii, C. lipolytica*), using a variety of substrates including sucrose, molasses, corn syrup, enzyme-treated starch, and normal paraffins. Citric acid is recovered from the fermentation broth by solvent extraction or more commonly by precipitation as calcium citrate, followed by treatment with sulfuric acid to convert the calcium citrate to calcium sulfate and citric acid. The calcium sulfate is removed by filtration, and the citric acid solution is further purified. Crystallization of citric acid from a hot aqueous solution (above the transition temperature of 36.6°C or 97.9°F) yields anhydrous citric acid; crystallization from a cold solution yields the monohydrate. Although total chemical syntheses for citric acid have been published, they have never achieved commercial success. *See* FERMENTATION; INDUSTRIAL MICROBIOLOGY.

Properties. Key properties of anhydrous citric acid are summarized in **Table 2**. Citric acid is a relatively strong organic acid (see dissociation constants), and is very soluble in water.

Uses. Citric acid is widely used in the food and pharmaceutical industries. In foods it is used primarily to produce a tart taste and to complement fruit flavors in carbonated beverages, beverage powders, fruit-flavored drinks, jams and jellies, candy, sherbets, water ices, and wine. It is also used to reduce pH in certain canned foods to make heat treatment more effective, and in conjunction with antioxidants to chelate trace metals and retard enzymatic activity. *See* FOOD MANUFACTURING.

In pharmaceuticals, citric acid provides the acid source in effervescent tablets in addition to being used to adjust pH, impart a tart taste, and chelate trace metals. It is also used as a blood anticoagulant. *See* PHARMACEUTICAL CHEMISTRY.

Citric acid, because of its low toxicity, relative noncorrosiveness, and biodegradability, is also being used for applications normally reserved for the strong mineral acids. These include preoperational and operational cleaning of iron and copper oxides from boilers, nuclear reactors, and heat exchangers; passivation of stainless steel tanks and equipment; and etching of concrete floors prior to coating. It is also used as a dispersant to retard settling of titanium dioxide slurries and as a sequestering and pH control agent in the textile industry. *See* TEXTILE CHEMISTRY.

Salts. Since citric acid is a tribasic acid, it can form acid and neutral salts, and will buffer over a broad pH range. The salts which are readily available commercially are trisodium citrate (dihydrate and anhydrous), tripotassium citrate (monohydrate), calcium citrate (tetrahydrate), diammonium citrate (anhydrous), and ferric ammonium citrate. *See* BUFFERS (CHEMISTRY); PH.

Trisodium citrate dihydrate, the salt which is most widely used, is a stable white crystal or granule. It is used as an emulsifier in processed cheese products where it prevents fat separation, imparts good meltdown properties, and produces slices with proper flexibility. In other dairy products, such as ice cream, whipping cream, and evaporated milk, it acts as a stabilizer and viscosity control agent. The sharpness of high-acid beverages is mellowed with sodium citrate, and it is used as a buffer to control pH in jams and jellies, gelatin desserts, and many pharmaceutical preparations. It is also used as a blood anticoagulant. Its largest single industrial use is as a detergent builder since it is rapidly biodegradable, environmentally acceptable, and can chelate

TABLE 2. Properties of anhydrous citric acid

Property	Value
Formula	$H_3C_6H_5O_7$
Molecular weight	192.13
Dissociation constants	K_1 8.2×10^{-4}
	K_2 1.8×10^{-5}
	K_3 3.9×10^{-7}
Melting point	153°C (307°F)
Solubility, g/100 ml at 25°C (77°F)	
Water	181
Ethanol	59
Ether	0.75

calcium, magnesium, and other metal ions. Sodium citrate is also used as a water conditioner, set-retardant for cement, buffer, and scrubbing agent to remove sulfur dioxide from stack gases or process tail gases. In most applications the dihydrate is used since it is less expensive. Occasionally, certain dry formulations require the use of the anhydrous salt. *See* CHEESE; DETERGENT.

Potassium citrate is a white crystalline granule or powder which readily absorbs moisture from the air. It is often used as a substitute for sodium citrate in special food products where sodium ion is undesirable. Calcium citrate is a stable, white, free-flowing powder used as a source of calcium either for nutritional purposes or for functional purposes (that is, as a firming agent). Ferric ammonium citrate is a complex salt, of undetermined structure, composed of iron, ammonia, and citric acid. Product containing 14.5–16% iron is green in color, and product containing 14.5–18.5% iron is brown. Both absorb moisture readily and are affected by light. They are used as a nutrient source of iron in food and in pharmaceutical syrups and elixirs. Diammonium citrate is a white granule or powder which is stable in air and used in dry formulations for metal cleaning. Fred Sauer

Bibliography. National Research Council, *Food Chemicals Codex*, 4th ed., 1996; Pfizer Chemical Division, Technical Literature: *Citrate Salts in Foods*, 1983, *Citric Acid and Citrates*, 1978, *Food Acidulants*, 1981, *Organic Chelating Agents*, 1984; *Scientific Literature Review on GRAS Food Ingredients: Citrates*, PB 223 850, National Technical Information Service, Springfield, VA, April 1973; *Scientific Literature Review on GRAS Food Ingredients: Citric Acid*, PB 241 967, National Technical Information Service, Springfield, VA, October 1974.

Citric acid cycle

In aerobic cells from animal and certain other species, the major pathway for the complete oxidation of acetyl coenzyme A (the thioester of acetic acid with coenzyme A); also known as the Krebs cycle or tricarboxylic acid cycle. Reduced electron carriers generated in the cycle are reoxidized by oxygen via the electron transport system; water is formed, and the energy liberated is conserved by the phosphorylation of adenosine diphosphate (ADP) to adenosine triphosphate (ATP). Reactions of the cycle also function in metabolic processes other than energy generation. The role of the cycle in mammalian tissues will be emphasized in this article. *See* ADENOSINE DIPHOSPHATE (ADP); ADENOSINE TRIPHOSPHATE (ATP); COENZYME; ENZYME.

Reactions. The first step in the cycle involves the condensation of the acetyl portion of acetyl coenzyme A (CoA) with the four-carbon compound oxaloacetate to form citrate, a tricarboxylate containing six carbons (see **illus.**). A shift of the hydroxyl group of citrate to an adjacent carbon results in the formation of D-threo-isocitrate, which in turn is oxidized to the five-carbon compound α-ketoglutarate and carbon dioxide (CO_2). In a second oxidative decarboxylation reaction, α-ketoglutarate, in the presence of CoA, is converted to succinyl CoA and another molecule of CO_2. In the subsequent formation of the four-carbon compound succinate and CoA, the energy in the thioester bond of succinyl CoA is conserved by the formation of guanosine triphosphate (GTP) from guanosine diphosphate (GDP) and inorganic phosphate. Fumarate is formed from succinate by the removal of two atoms of hydrogen, and the unsaturated compound is then hydrated to L-malate. The dehydrogenation of malate forms oxaloacetate, the starting four-carbon compound of the metabolic cycle. Thus, beginning with the two-carbon acetyl group, one completion of the cycle results in the formation of two molecules of carbon dioxide. The obligatory role of the intact cycle for the complete oxidation of acetyl CoA has been demonstrated by inhibiting enzymes required for specific reactions (for example, inhibition of succinate dehydrogenase by the substrate analog malonate) or by altering the regeneration or depletion of an intermediate of the cycle.

Electron transport and oxidative phosphorylation. The oxidation of acetyl CoA to CO_2 in the cycle occurs without direct reaction with molecular oxygen. The oxidations occur at dehydrogenation reactions in which hydrogen atoms and electrons are transferred from intermediates of the cycle to the electron carriers nicotinamide adenine dinucleotide (NAD^+) and flavin adenine dinucleotide (FAD). The NAD-specific dehydrogenase reactions are isocitrate dehydrogenase [reaction (1)], α-ketoglutarate dehydrogenase (2), and malate dehydrogenase (3). FAD is the coenzyme for succinate dehydrogenase, (4). The electrons from NADH and $FADH_2$ are transferred to molecular oxygen via a series of electron transport carriers, with regeneration of NAD^+ and FAD. The energy liberated in the electron transport chain is partially conserved by the formation of ATP from ADP and inorganic phosphate, by a process called oxidative phosphorylation. The energy generated as oxygen accepts electrons from the reduced coenzymes generated in one turn of the cycle results in the maximal formation of 11 molecules of ATP. Because GTP obtained by phosphorylation of GDP at the succinyl CoA to succinate step of the cycle is readily converted to ATP by nucleotide diphosphokinase, as in reaction (5), the yield is 12 molecules of ATP per molecule of acetyl CoA metabolized. *See* NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD).

$$\text{Isocitrate} + NAD^+ \longrightarrow \alpha\text{-ketoglutarate} + CO_2 + NADH + H^+ \quad (1)$$

$$\alpha\text{-Ketoglutarate} + CoA + NAD^+ \longrightarrow \text{succinyl CoA} + CO_2 + NADH + H^+ \quad (2)$$

$$\text{Malate} + NAD^+ \longleftrightarrow \text{oxaloacetate} + NADH + H^+ \quad (3)$$

$$\text{Succinate} + FAD \longleftrightarrow \text{fumarate} + FADH_2 \quad (4)$$

$$GTP + ADP \longleftrightarrow GDP + ATP \quad (5)$$

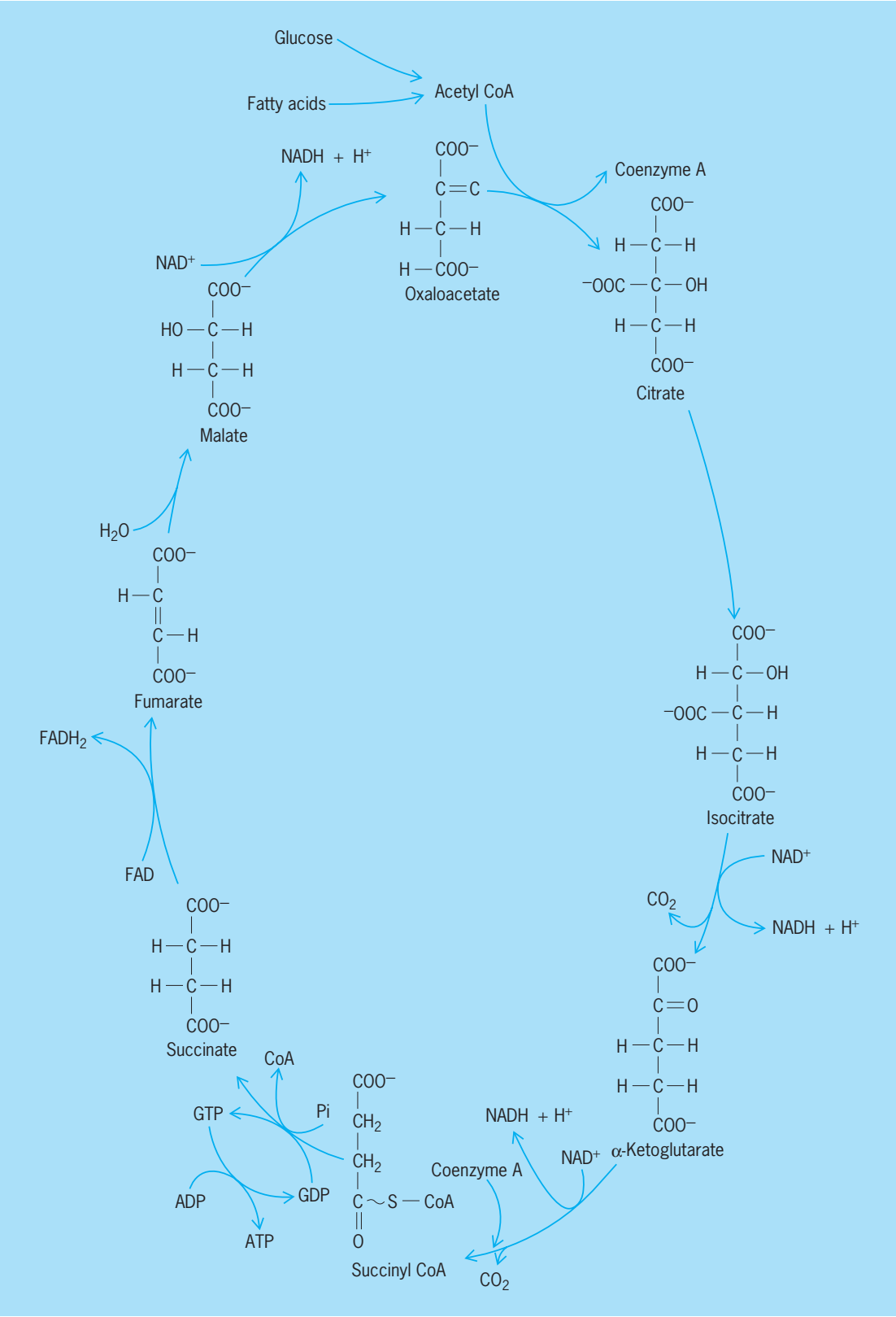

Citric acid cycle.

The electron transport and oxidative phosphorylation systems and the enzymes required for the citric acid cycle are located in the mitochondria of cells. These mitochondrial systems are the major source of ATP for energy-consuming reactions in most tissues. The citric acid cycle does not occur in all cells. For example, mature human red blood cells do not contain mitochondria and the cycle is absent. In these

cells, ATP is formed by the anaerobic conversion of glucose to lactate (anaerobic glycolysis). *See* MITOCHONDRIA; PHOSPHATE METABOLISM.

Regulation. A major rate-limiting step of the citric acid cycle in aerobic mammalian tissues (for example, heart) is at NAD-specific isocitrate dehydrogenase [reaction (1)]. The activity of the enzyme is dependent on the concentration of the substrate (magnesium isocitrate): it is activated by ADP, calcium (Ca^{2+}), and citrate, and it is inhibited by NADH and reduced nicotinamide adenine dinucleotide phosphate (NADPH). Differences in energy demand affect the rate of citric acid flux and concentrations of modulators of NAD-dependent isocitrate dehydrogenase. At rest, the flux through the cycle is slowed, the cellular concentration of the isocitrate dehydrogenase inhibitor NADH is raised; and the activator ADP is lowered; the opposite occurs during high energy demand. The rate of cycle oxidation in liver mitochondria increases with decreased NADPH concentration; increased mitochondrial free Ca^{2+} enhances cycle flux. *See* NICOTINAMIDE ADENINE DINUCLEOTIDE PHOSPHATE (NADP).

Formation of acetyl CoA and intermediates. Acetyl CoA is formed from carbohydrates, fats, and the carbon skeleton of amino acids. The origin of a precursor and the extent of its utilization depend on the metabolic capability of a specific tissue and on the physiological state of the organism. For example, most mammalian tissues have the capacity to convert glucose to pyruvate in a reaction called glycolysis. Pyruvate is then taken up from cellular cytosol by mitochondria and oxidatively decarboxylated to acetyl CoA and carbon dioxide by pyruvate dehydrogenase, as in reaction (6). Acetyl CoA is also the end product

$$\text{Pyruvate} + NAD^+ + \text{CoA} \longrightarrow \text{acetyl CoA} + CO_2 + \text{NADH} + H^+ \quad (6)$$

of fatty acid oxidation in mitochondria. However, the fatty acid oxidation pathway occurs in fewer tissues than does glycolysis or the citric acid cycle. For example, nervous tissue utilizes the complete oxidation of glucose but not of fatty acids to maintain energy needs. In prolonged starvation, the level of blood glucose declines to a concentration inadequate to support the optimal energy needs of brain, whereas the level of blood ketone bodies (formed from fatty acids in liver) rises. In this case, the oxidation of ketone bodies (acetoacetate and 3-hydroxybutyrate) via acetyl CoA supplements the diminished oxidation of glucose to fulfill the energy requirements of the brain.

The amino acids follow varied pathways for forming compounds that can enter the citric acid cycle. For example, the paths of degradation of leucine and lysine lead directly to formation of acetyl CoA; only part of the carbons of aromatic amino acids and of isoleucine are converted to acetyl CoA. Pyruvate, formed bytransamination of alanine with the citric acid cycle intermediate α-ketoglutarate [reaction (7)], can be oxidatively decarboxylated to acetyl

$$\text{Alanine} + \alpha\text{-ketoglutarate} \longleftrightarrow \text{pyruvate} + \text{glutamate} \quad (7)$$

CoA by pyruvate dehydrogenase [reaction (6)] or carboxylated to the cycle intermediate oxaloacetate by pyruvate carboxylase [reaction (8)]. Oxaloacetate

$$\text{Pyruvate} + CO_2 + \text{ATP} \longrightarrow \text{oxaloacetate} + \text{ADP} + P_1 \quad (8)$$

can also be formed by aspartate aminotransferase [reaction (9)]. Portions of the carbon skeletons of va-

$$\text{Aspartate} + \alpha\text{-ketoglutarate} \longleftrightarrow \text{oxaloacetate} + \text{glutamate} \quad (9)$$

line, isoleucine, methionine, and threonine are converted to methylmalonyl CoA, which rearranges to the citric acid cycle intermediate succinyl CoA. *See* AMINO ACIDS.

Role in lipogenesis and gluconeogenesis. In addition to the cycle's role in yielding catabolic energy, portions of it can supply intermediates for synthetic processes, such as the synthesis of the fatty acid moiety of triglycerides from glucose (lipogenesis), and formation of glucose from the carbon skeletons of certain amino acids, lactate, or glycerol (gluconeogenesis).

Lipogenesis. When dietary intake exceeds the energy needs of the body, the excess calories are deposited as body fat. Under these conditions in humans, the synthesis of fatty acids from glucose occurs mainly in the liver. It is favored by metabolic changes resulting principally from a rise in blood insulin: (1) The rate of pyruvate formation increases because of the rise in tissue glucose and enhancement of glycolysis. (2) Increases in enzyme activities catalyzing the conversion of pyruvate to acetyl CoA at pyruvate dehydrogenase [reaction (6)] and to oxaloacetate at pyruvate carboxylase [reaction (8)] raise the concentrations of both of these substrates for citrate synthesis. When this occurs at a rate exceeding that of the citric acid cycle, the citrate concentration rises and citrate exits from the mitochondria to the cytosol. (3) In the cytosol, citrate is cleaved to oxaloacetate and acetyl CoA in a reaction requiring CoA and ATP. The acetyl CoA is converted to palmitate and other long-chain fatty acids in a series of condensation steps requiring CO_2, ATP as energy source, and NADPH as reductant. *See* LIPID METABOLISM.

Gluconeogenesis. Dietary deprivation of an adult human for more than 2–3 days results in the depletion of liver glycogen stores supporting blood glucose, and the oxidation of fatty acids derived from body triglyceride depots becomes the principal body energy source. However, certain tissues, such as brain and red and white blood cells, remain partially or wholly dependent on glucose as the energy source. In the starved state, where blood insulin decreases and glucagon increases, increased mobilization of amino acids from skeletal muscle proteins occurs, and carbons from some of these amino acids

become available for glucose synthesis. The interactions of degradation pathways of individual amino acids in various tissues result in markedly increased blood alanine, an important substrate for gluconeogenesis.

Some of the steps in the conversion of alanine to glucose are as follows: (1) Alanine is taken up from the bloodstream by the liver and converted to pyruvate by transamination [reaction (7)]. (2) The further metabolism of pyruvate in liver mitochondria is affected by the increase in the rate of fatty acid oxidation in the starved state, which raises the levels of acetyl CoA and NADH. Since NADH inhibits pyruvate dehydrogenase [reaction (6)], and acetyl CoA activates pyruvate carboxylase [reaction (8)], the conversion of pyruvate to oxaloacetate rather than acetyl CoA is favored. The increased NADH displaces the malate dehydrogenase reaction [reaction (3)] from oxaloacetate toward malate. Malate is transported from the mitochondria to the cytosol, where it is reoxidized by NAD^+ to oxaloacetate by cytosolic malate dehydrogenase. (3) Oxaloacetate is converted to phosphoenolpyruvate (PEP) by the PEP-carboxykinase reaction [reaction (10)].

$$\text{Oxaloacetate} + \text{GTP} \longrightarrow \text{PEP} + CO_2 + \text{GDP} \qquad (10)$$

Phosphoenolpyruvate is an intermediate of both gluconeogenesis and glycolysis. In the starved state, the levels of modulators, metabolites, and so forth activate certain rate-limiting steps favoring the formation of glucose from PEP, and also inhibit key enzyme activities facilitating glycolysis. The net effect of these opposing trends is to funnel the flux of carbons from amino acids toward glucose synthesis. *See* BIOLOGICAL OXIDATION; CARBOHYDRATE METABOLISM; CELL (BIOLOGY); GLUCOSE; GLYCOGEN; METABOLISM.

Gerhard W. E. Plaut

Citron

Citrus medica, a species of true citrus. The plant has a long and interesting history, being the first contact between citrus, which came from the Far East, and western Judaeo-Christian civilization. A citron, the Etrog, was the "hadar" or "goodly fruit" offering at the ancient Jewish Feast of the Tabernacles. At the time of the first Jewish revolt against Rome in A.D. 66–70, the rebels minted their own coins with the Etrog substituting for Nero's imperial visage.

Commercially, citrons are still grown almost exclusively in the Mediterranean area, principally in Italy, Sicily, Corsica, Greece, and Israel. The tree is evergreen, as are all citrus, and frost-tender. It is thorny, straggly, shrubby, and tends to be short-lived.

The fruit is scarcely edible fresh, having a very thick skin with little flesh and that lacking in juice (see **illus.**). However, it is very fragrant and was valued in ancient times for its aroma and its fragrant peel oil, used in perfumes and as a moth repellent. Later, when the citron was the only known citrus fruit, it was very highly valued. Today, it is grown commercially only as a source of candied peel for use in cakes and confections. The actual candying is usually done in the importing country, the citron peel being exported in brine. Traditionally the fruit was halved and held in large unsealed casks in seawater, a mild fermentation taking place for a week or more, after which additional coarse salt was added and the casks sealed. Today brining is more typically done in a 15% salt solution to which about 2000 parts per million of SO_2 has been added as a preservative and 1% $CaCl_2$ to firm the tissues. The candying process involves first leaching out the brine with boiling water and then transfers through successive syrups of increasing strength up to about 75%.

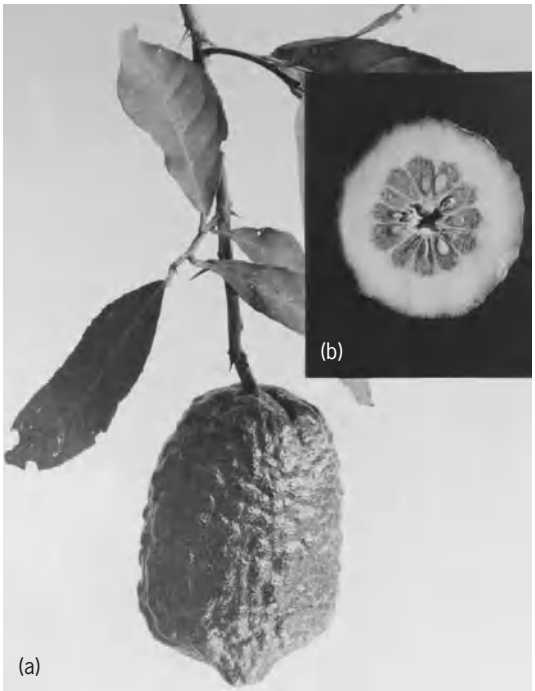

Commercial citron (*Citrus medica*), approximately one-half natural size. (*a*) Foliage and fruit. (*b*) Cross section of fruit. (*From J. Horace McFarland Co.*)

Confusion sometimes arises due to the name "citron" also being used for a small, wild, inedible melon in the United States and for lemons (*C. limon*) in France. *See* FRUIT; FRUIT, TREE.

William Grierson

Citronella

A tropical grass, *Cymbopogon nardus,* from the leaves of which oil of citronella is distilled. This essential oil is pale yellow, inexpensive, and much used in cheap perfumes and soaps. It is perhaps best known as an insect repellent. A large acreage is devoted to the cultivation of this grass in Java and Ceylon. *See* CYPERALES; ESSENTIAL OILS.

Perry D. Strausbaugh; Earl L. Core

Civet

The common name for any of various carnivorous catlike mammals of the family Viverridae of Africa and Asia. These small and medium-sized carnivores are often called "weasel cats" because of their long, slender weasel-like body with short limbs and their long, bushy tail and catlike disposition. They inhabit the warmer parts of the Old World, including southern Asia, the East Indies, portions of Africa, and Madagascar.

Classification. The family Viverridae includes 35 species divided among the following genera: genets (*Genetta*, *Osbornictis*), linsangs (*Poiana*, *Prionodon*), binturongs (*Arctictis*), Oriental civets (*Viverra*), the African civet (*Civettictis*), lesser Oriental civet (*Viverricula*), African palm civet (*Nandinia*), small-toothed palm civet (*Arctogalidia*), masked palm civet (*Paradoxurus*, *Paguma*), Sulawesian palm civet (*Macrogalidia*), banded palm civet (*Hemigalus*), Hose's palm civet (*Diplogale*), Owston's palm civet (*Chrotogale*; see **illustration**), otter civets (*Cynogale*), Malagasy civet (*Fossa*), and falanoucs (*Eupleres*).

Morphology. The civet's head is elongate and the muzzle is pointed. Civets have five toes on each foot, with each toe ending in a sharp claw. Claws may be retractile, semiretractile, or nonretractile. Like cats, civets walk softly and spring swiftly upon their prey. Like skunks, civet cats are strikingly marked so that they are easily recognized, remembered and avoided by potential enemies. They may be striped, spotted, or have a uniform color pattern. In some species the tail is banded or ringed. Females usually have two or three pairs of abdominal mammae. There are between 32 and 40 teeth, with the carnassial teeth being well developed. The senses of slight, hearing, and smell are well developed. Head-body length is 350–950 mm (13.5–37 in.), tail length is 130–900 mm (5–35 in.), and adult weight is 0.6–20 kg (1.3–44 lb).

Habitat. Viverrids are primarily forest dwellers, although they also inhabit dense brush and thick grass. Most species are terrestrial, although many are skillful climbers and spend most of their lives in trees. Shelters include burrows, caves, crevices, tree cavities, or thick brush. Some civets are solitary, whereas others live in pairs or groups. They may be diurnal or nocturnal.

Otter civets (*Cynogale*) are semiaquatic. They have the flat head, broad toes and feet, and thick, dense fur of an aquatic animal. They swim and fish like otters and are as much at home in the water as on dry land.

Diet. Civets are omnivorous. Food consists of small vertebrates and invertebrates and occasionally some vegetable matter such as fruits, bulbs, and nuts. Carrion is taken by some species. Since civets do not digest the seeds of fruits and berries, they distribute the seeds over a considerable distance where they germinate and eventually grow into trees and shrubs.

Owston's palm civet (*Chrotogale owston*). (*Photo © by Mark Kostich*)

Breeding and longevity. Breeding may occur seasonally or throughout the year. Some species produce two litters of one to six young annually. Newborn civets are blind but covered with hair. Potential longevity ranges from 5 to 15 years.

Economic importance. Members of the family Viverridae are valued for their musk, fur, and meat.

Civet musk. Most members of this family possess scent, or musk, glands in the anal region that secrete a fluid with a musky odor, known as civet, as a defensive measure. This thick yellowish substance is composed of ammonia, resin, fat, and volatile oil. When rubbed on tree trunks, the ground, and similar objects, the secretion is probably used as a means of communication between individuals of the same species. Civet is obtained from several genera (*Civettictis*, *Viverra*, and *Viverricula*) for use in perfumes and medicine, primarily in the Orient. Some civets are tamed and kept in captivity specifically for their musk, which is extracted every 14 to 20 days. Many perfume manufacturers now use civetone, a synthetic alternative.

One excellent producer of civet is the Oriental civet (*Viverra*), as its musk glands are greatly enlarged. It typically has a head-body length of 60 cm (2 ft), a 45-cm (18-in.) tail, and a weight of 11 kg (25 lb). The Oriental civet has a full, black-and-white ringed tail, many ill-defined markings on its grizzled body, and a broad black band set off by areas of white across the throat. A prominent crest of long erectile hairs extends down the middle of the back. The feet are black, and the fur is soft and thick, making it of considerable importance to the fur trade. *See* SCENT GLAND.

Fur and meat. Besides their importance to the perfume industry, civets are prized for their fur and meat. The civet cat is considered a delicacy in southern China. It has been suggested that the practice of eating civets may have resulted in the severe acute respiratory disease (SARS) virus outbreak of 2003. SARS, a type of atypical pneumonia, first broke out in Guangdong province in southern China in November 2002. In January 2004, Guangdong province

banned sales of civet cats and ordered the slaughter of all captive civets. In January 2004, the United States announced an embargo on the importation of civets.

Status. Several species are listed as either vulnerable or endangered by the International Union for the Conservation of Nature and Natural Resources (IUCN) and/or the U.S. Department of the Interior. The Malabar civet (*Viverra civettina*) is one of the world's rarest mammals and is listed as critically endangered. It is known only from an isolated belt of rainforests in the Western Ghats of southwestern India. Extensive deforestation has reduced the forests it inhabits to a series of isolated patches. Habitat loss continues. Cashew plantations, which probably hold most of the surviving populations of this species, are now threatened by large-scale clearing for planting rubber. *See* CARNIVORA; ENDANGERED SPECIES; MAMMALIA. Donald W. Linzey

Bibliography. D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Civil engineering

A branch of engineering that encompasses the conception, design, construction, and management of residential and commercial buildings and structures, water supply facilities, transportation systems for goods and people, as well as control of the environment for the maintenance and improvement of the quality of life. Civil engineering is a people-serving profession; it includes planning and design professionals in both the public and private sectors, contractors, builders, educators, and researchers who strive to meet the needs and desires of the community.

The civil engineer holds the safety, health, and welfare of the public paramount. Civil engineering projects and systems should conform to governmental regulations and statutes; should be built economically to function properly for a reasonable period of time with a minimum of maintenance and repair while withstanding anticipated usage and weather; and should conserve energy and allow hazard-free construction while providing healthful, safe, and environmentally sound utilization by society.

Because the desired objectives are so broad and encompass an orderly progression of interrelated components and information to arrive at the visually pleasing, environmentally satisfactory, and energy-frugal end point, civil engineering projects are actually systems requiring the skills and inputs of many diverse technical specialties, all of which are subsets of the overall civil engineering profession.

The infrastructure is the foundation of the modern industrialized urban society. The infrastructure includes roads; mass transit; railroads; bridges; airports; storage buildings; terminals; communication and control towers; water supply and treatment systems; storm water control systems; wastewater collection, treatment and disposal systems; as well as living and working areas, recreational buildings, and ancillary structures for civil and civic needs. Without a well-maintained and functioning infrastructure, the urban area, and society, cannot stay healthy, grow, and prosper. Civil engineers, with their various technical specialties, play a major role in developing workable solutions to construct, renovate, repair, maintain, and upgrade this infrastructure. The civil engineer works with other professionals in developing functional projects, arriving at realistic cost estimates, and advocating reasonable investment strategies.

Civil engineers who move into management combine their technical training and skills with the ability to organize and properly employ the four basics of any project—personnel, time, materials, and money. Civil engineers in management may work for private engineering firms, for industries, for the military, for governmental agencies, or for regional authorities. The civil engineer may also choose a career in education. The teaching civil engineer educates students in technical specialties, inculcates engineering ethics and standards of safety, and engages in research and scholarly endeavors. Civil engineering faculty members are frequently involved as consultants or specialists on boards or commissions for standards, planning, or educational certification.

Photogrammetry, surveying, and mapping. The civil engineer in this specialty precisely measures the Earth's surface to obtain reliable information for locating, planning, and designing engineering projects. This specialty involves everything from traditional land surveying to high-technology methods using aerial surveying, satellites, and computer-processing of photographic imagery. The information obtained by optical sighting, radio signal, laser scan, or sonic beam is converted to maps to provide accurate locations for designing planned developments, boring tunnels, building highways or pipelines, plotting flood control, irrigation or hydroelectric projects with dams and other structures, delimiting subsurface geologic formations that may affect construction projects, and aiding many other building uses. *See* AERIAL PHOTOGRAPH; PHOTOMETRY; REMOTE SENSING; SURVEYING; TOPOGRAPHIC SURVEYING AND MAPPING.

Community and urban planning. Those engaged in this segment of civil engineering evaluate environmental, social, and economic factors in affecting the use and development of land and natural resources as they plan to establish planned developments, recreational park areas, or industrial complexes for regions, cities, or portions of a community. They work with other professionals to coordinate required infrastructure to ensure social, economic, and environmental well-being. *See* LAND-USE PLANNING.

Geotechnical engineering. Civil engineers in this specialty analyze and evaluate the behavior of earth materials and the ability of these materials to support structures on or below the ground. They design and establish procedures to prevent the pressure or weight of structures from compressing the

supporting earth and causing settlement. They are also concerned with methods to stabilize slopes and fills, to protect structures and infrastructure from earthquakes, and to mitigate the unwanted effects of ground water. *See* ENGINEERING GEOLOGY; FOUNDATIONS; PERMAFROST; PILE FOUNDATION; ROCK MECHANICS; SOIL MECHANICS.

Construction engineering. Civil engineers in this area manage and direct the physical construction of a project from start to finish; this field is also known as construction management. Construction engineers apply the knowledge of construction methods and equipment along with principles of financing, scheduling, planning, organization, and coordination to convert the paper designs into completed usable facilities. They maintain a continuous record of the personnel, time, materials, and costs expended and prepare periodic reports depicting the project's progress to completion. *See* CONSTRUCTION ENGINEERING; CONSTRUCTION METHODS; ENGINEERING AND ARCHITECTURAL CONTRACTS.

Structural engineering. In this specialty, civil engineers plan, design, and evaluate a wide variety of structures, including buildings, bridges, offshore platforms, space platforms, amusement park rides, towers, and any other construction that must support its own weight and the loads it must carry. These structures must be designed for, and capable of withstanding, the loads due to rain, snow, or ice; the forces due to wind, hurricanes, and earthquakes; the effects of temperature changes and the like. The structural engineer must determine which construction materials or combinations of them such as steel, concrete, stone, brick, aluminum, plastics, and glass are most economical, safe, and appropriate for the project under the site-specific conditions. *See* BRIDGE; BUILDINGS; STRUCTURAL ANALYSIS; STRUCTURAL DESIGN; STRUCTURAL MATERIALS; STRUCTURE (ENGINEERING).

Environmental engineering. In this branch of the profession, civil engineers plan, design, construct, operate, and supervise systems to protect human health and the ecological balances necessary for environmental quality in both natural and humanmade environments. These systems include development and treatment of drinking water; collection and treatment of sanitary and industrial wastes to prevent contamination of the land and water; containment of hazardous and toxic materials; safe disposal of garbage, refuse, and other solid wastes; minimization or elimination of air pollution caused by industrial manufacturing or combustion processes; and design for containment of radioactive discharges. The environmental engineer is involved, either personally or through a team effort, to reduce, mitigate, or eliminate the hazards and risks associated with air, food, and water contaminants, radiation, toxic chemicals, solid wastes, disease vectors, safety hazards, and habitat alterations to preserve or conserve ecological biomes. *See* AIR POLLUTION; HAZARDOUS WASTE; INDUSTRIAL WASTEWATER TREATMENT; LAND RECLAMATION; SEWAGE TREATMENT; WATER POLLUTION; WETLANDS.

Water resources engineering. Civil engineers who specialize in this area deal with all aspects of the physical control of water. They plan, evaluate, design, construct, supervise, and operate systems to control floods; to supply water for cities; to develop irrigation; to manage and control river navigation, locks, and banks; to prevent beach erosion; and to maintain waterfront facilities. They are also involved in the planning and design of harbors, docks, wharfs, port facilities, canals, and offshore platforms and systems. They are involved with other civil engineers in planning and designing flood control dikes, dams, and hydroelectric power impoundments. *See* CANAL; COASTAL ENGINEERING; DAM; HARBORS AND PORTS; RESERVOIR; REVETMENT; RIVER ENGINEERING; WATER SUPPLY ENGINEERING.

Waste management and risk assessment. Civil engineers involved in this area must function as a part of a team. The improper disposal of hazardous or toxic wastes, as well as the indiscriminate dumping of wastes in landfills, has created a situation that impacts on the land and subsurface water supplies. A subset of environmental engineering has evolved that is specifically concerned with the management of these sites. At the same time it is imperative that the available time, money, and personnel be utilized to treat these existing sites in a systematic and cost-effective manner. One method of establishing the hierarchy for treating the existing sites is to perform a risk assessment and establish which areas pose the greatest threat to people and natural resources. The civil engineers in this area of work must develop treatability procedures and estimate costs needed to reduce the risk. The data must then be presented to the public to explain the trade-offs of potential risk and the costs necessary to reduce them, since the locally affected community must be willing to expend time, effort, and money to clean up and manage the site. *See* RISK ASSESSMENT AND MANAGEMENT.

Transportation engineering. Civil engineers working in this specialty plan, design, construct, and maintain facilities to safely and efficiently move people and goods. Accomplishing this task involves modal system planning for pedestrians, bikeways, streets, roads, highways, mass transit, railroads, airports, ports, and harbors. The ancillary structures to the vehicles and the travel way, which includes parking areas and terminals as the transfer facilities between the various modes of traffic, must be considered. Transportation engineers must apply technological acumen and consider the economic, social, and political factors that impact on each system. It is essential that transportation planning be done in conjunction with community and urban planners. The goal is an acceptable means of enhancing the movement of goods and people from residence to place of employment or entertainment, to shopping areas, and to health care facilities while always permitting quick response of emergency vehicles. The prosperity and quality of life of a community are intimately tied up with the transportation system. *See* AIRPORT ENGINEERING; HIGHWAY ENGINEERING; RAILROAD ENGINEERING; TRANSPORTATION ENGINEERING.

Pipeline engineering. Civil engineers in this highly specialized task of transporting liquids, gases, or slurries plan, design, maintain, and operate the pipelines in which these various types of noncombustible to highly combustible materials are moved from one location to another. The engineers determine pipeline design, location, and the environmental and economic impacts on the region that the line traverses. They select the material for constructing the line based on the qualities of durability and safety. Further, they specify installation techniques, methods of testing, and controls for maintaining proper pressure and rate of flow of the material within the line. When a hazardous or toxic material is to be carried, safety is a major concern. *See* ENGINEERING; PIPELINE.

Kenneth Lazuruk; Gerald Palevsky

Bibliography. American Society of Civil Engineers, *The Civil Engineering Technologist*, 1993; American Society of Civil Engineers, *Our Past, The Present, Your Future . . . in Civil Engineering*, 1993; N. S. Grigg, Infrastructure and economic development: Civil engineering perspective, *J. Prof. Iss. Eng. Educ. Prac.*, 119:51–61, January 1993; W. J. Hall et al., *Civil Engineering in the 21st Century: A Collection of Papers*, 1987; W. J. Hall et al., *Civil Engineering in the 21st Century: A Vision and a Challenge for the Profession*, 1988; F. S. Merritt, J. T. Ricketts, and M. K. Loftin, *Merritt's Civil Engineering Handbook*, 4th ed., 1995.

Cladding

An old jewelry art, now employed on an industrial scale to add the desirable surface properties of an expensive metal to a low-cost or strong base metal. In the process a clad metal sheet is made by bonding or welding a thick facing to a slab of base metal; the composite plate is then rolled to the desired thickness. The relative thickness of the layers does not change during rolling. Cladding thickness is usually specified as a percentage of the total thickness, commonly 10%.

Uses. Gold-filled jewelry has long been made by this process; the surface is gold, the base metal bronze or brass with the cladding thickness usually 5%. The process is used to add corrosion resistance to steel and to add electrical or thermal conductivity, or good bearing properties, to strong metals. One of the first industrial applications, about 1930, was the use of a nickel-clad steel plate for a railroad tank car to transport caustic soda; by 1947 stainless-clad steels were being used for food and pharmaceutical equipment. Corrosion-resistant pure aluminum is clad to a strong duralumin base, and many other combinations of metals are widely used in cladding; a development includes a technique for cladding titanium to steel for jet-engine parts.

Today's coinage uses clad metals as a replacement for rare silver. Since 1965 dimes and quarters have been minted from composite sheet consisting of a copper core with copper-nickel facing. The proportion of core and facing used duplicates the weight and electrical conductivity of silver so the composite coins are acceptable in vending machines (**Fig. 1**).

Fig. 1. Rolls of copper clad with copper-nickel facing for minting composite coins.

A three-metal composite sole plate for domestic steam irons provides a thin layer of stainless steel on the outside to resist wear and corrosion. A thick core of aluminum contributes thermal conductivity and reduces weight, and a thin zinc layer on the inside aids in bonding the sole plate to the body of the iron during casting.

Cladding supplies a combination of desired properties not found in any one metal. A base metal can be selected for cost or structural properties, and another metal added for surface protection or some special property such as electrical conductivity. Thickness of the cladding can be made much heavier and more durable than obtainable by electroplating.

Combinations. The following clad materials are in common use.

Stainless steel on steel. Provides corrosion resistance and attractive surface at low cost for food display cases, chemical processing equipment, sterilizers, and decorative trim.

Stainless steel on copper. Combines surface protection and high thermal conductivity for pots and pans, and for heat exchangers for chemical processes.

Copper on aluminum. Reduces cost of electrical conductors and saves copper on appliance wiring.

Copper on steel. Adds electrical conductivity and corrosion resistance needed in immersion heaters and electrical switch parts; facilitates soldering.

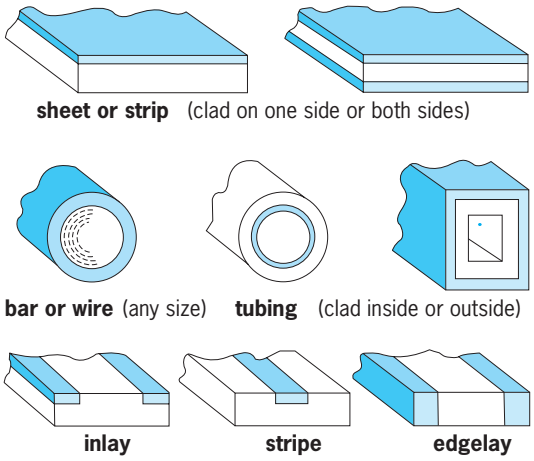

Fig. 2. Types of cladding.

Nickel or monel on steel. Provides resistance to corrosion and erosion for furnace parts, blowers, chemical equipment, toys, brush ferrules, and many mechanical parts in industrial and business machines; more durable than electroplating.

Titanium on steel. Supplies high-temperature corrosion resistance. Bonding requires a thin sheet of vanadium between titanium and steel.

Bronze on copper. Usually clad on both sides, for current-carrying springs and switch blades; combines good electrical conductivity and good spring properties.

Silver on copper. Provides oxidation resistance to surface of conductors, for high-frequency electrical coils, conductors, and braiding.

Silver on bronze or nickel. Adds current-carrying capacity to low-conductivity spring material; cladding sometimes is in form of stripes or inlays with silver areas serving as built-in electrical contacts.

Gold on copper. Supplies chemical resistance to a low-cost base metal for chemical processing equipment.

Gold on nickel or brass. Adds chemical resistance to a stronger base metal than copper; also used for jewelry, wristbands, and watchcases.

Types. Cladding can be added to both sides of a sheet or strip of base metal. Tubing can be supplied with a clad surface on inside or outside; round and rectangular wire can be clad similarly (**Fig. 2**).

For some forms of electrical contact, the composite materials are bonded side by side, or silver is inset as a stripe on one side or along the edges. This construction can place solid silver just where it is needed to form an electrical contact with no waste of costly metal.

Another related form of cladding is found in thermostatic bimetals in which equal thicknesses of low- and high-expansion metals are bonded together. With a change in temperature, differing expansion rates of the two metals cause the composite material to bend and thus operate valves in automobile cooling systems, or electrical contacts in room thermostats.

Clad wires with properly chosen proportions of materials of different thermal-expansion rates can match the thermal expansion of types of glass used for vacuum-tight seals for conductors in lamp bulbs and hermetically sealed enclosures.

In making parts from clad metal, the composite material can be bent, drawn, spun, or otherwise formed just the same as the base metal without breaking the bond. The maximum service temperature is limited by the melting point of the material at the juncture of the two metals; annealing temperatures during manufacturing are similarly limited.

Cut edges of clad sheets expose the base metal and thus may require special protection. Welding of thick clad sheets sometimes offers special problems in maintaining the integrity of the protective surface, but joining methods are generally the same as for the base metal. *See* ELECTROPLATING OF METALS; METAL COATINGS.

Robert W. Carson

Cladoselache

A genus of Paleozoic chondrichthyan "sharks," characterized by broad-based paired fins, a short stout bony spine preceding the first of the two dorsal fins, and a large lunate tail-fin. Often cited as among the best known, earliest, and/or most primitive of the early chondrichthyans (cartilaginous fishes including modern chimaeroids, sharks and rays), *Cladoselache* now appears ill defined and poorly diagnosed. Most of the best fossil specimens of *Cladoselache* are from the Upper Devonian Cleveland Shales, and include complete individuals of around 1 m (3.3 ft) in length. Several of these have mineralized traces of soft tissues, including much of the superficial jaw and gill musculature, and even apparent traces of internal organs such as the kidneys. The dentition consists of small tricuspid teeth in which the central cusp is largest; the body is mostly naked except for compound scales on the head and fins. No specimens of *Cladoselache* are known to have claspers—the sex organs of male chondrichthyans used for internal fertilization (otherwise preserved as skeletal rods extending from the rear of the hind set of paired fins). *See* CHONDRICHTHYES.

In the context of more recent discoveries of early sharklike fishes, *Cladoselache* is emerging as unusual and specialized in several respects. The high-aspect-ratio (tall and narrow) tail is like that of only a subset

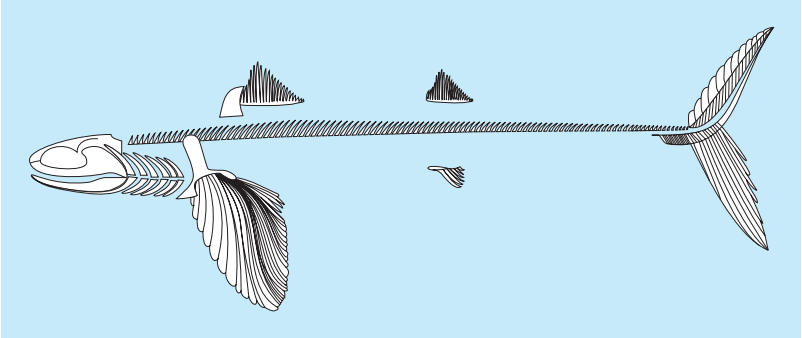

***Cladoselache fyleri*, from the Upper Devonian of Cleveland. (*After S. E. Bendix-Algreen, 1975*)**

of Paleozoic sharks, all of which have a transformed stretch of the vertebral column supporting the upper lobe, but this is far from primitive and completely absent from most contemporaneous chondrichthyan and nonchondrichthyan fishes. The pectoral fins are remarkably large for the total body size, and are supported by a series of broad straplike cartilages—most of which extend directly from the shoulder girdle. This pattern was once considered primitive because it resembled a hypothesized archetypical fin condition (the continuous lateral fin-fold model) for which there is now little fossil and embryological support. Comparisons with paired fins among other early fishes now highlight the cladoselachian pattern as unusual. The braincase (neurocranium) is known in outline only; it is peculiarly wide, although this might result from postmortem compression.

Few other aspects of the paleobiology of *Cladoselache* are known. The Cleveland Shale was deposited in a shallow inland sea covering what is now northeastern Ohio. *Cladoselache* was probably a medium-sized pelagic member of the fauna, dwarfed by large arthrodire placoderms (such as *Dunkleosteus*) known from the same fossil localities. Stomach contents of *Cladoselache* include the intact remains of small ray-finned (actinopterygian) fishes (*Kentuckia*), preserved with heads directed toward the front of the enclosing gut-trace. This rare detail from a fossil fish suggests that they (the ray-finned prey) were engulfed from the rear, tail-first, implying that in this instance *Cladoselache* behaved as an active, pursuit predator. *See* ACTINOPTERYGII; PLACODERMI. Michael Coates

Bibliography. J. A. Long, *The Rise of Fishes*, Johns Hopkins University Press, Baltimore, 1995; J. G. Maisey, *Discovering Fossil Fishes*, Westview Press, New York, 1996.

Clarification

The removal of small amounts of fine, particulate solids from liquids. The purpose is almost invariably to improve the quality of the liquid, and the removed solids often are discarded. The particles removed by a clarifier may be as large as 100 micrometers or as small as 2 micrometers. Clarification is used in the manufacture of pharmaceuticals, beverages, rayon, nylon, and other fiber and film polymers; in the reconditioning of electroplating solutions; in the recovery of dry-cleaning solvent; and for the purification of drinking water and wastewater. The filters in the feed line and lubricating oil system of an internal combustion engine are clarifiers.

The methods of clarification include gravity sedimentation, centrifugal sedimentation, filtration, and magnetic separation. Clarification differs from other applications of these mechanical separation techniques by the low solid content of the suspension to be clarified (usually less than 0.2%) and the substantial completion of the particle removal.

Gravity sedimentation is employed, usually in continuous settling tanks of simple construction, to concentrate the solid fraction and to clarify the liquid fraction of suspensions amenable to separation by this method. Gravity clarifiers differ from thickeners only in the concentration of solids in the feed and in the demand that the overflow be practically particle free. *See* SEDIMENTATION (INDUSTRY); THICKENING.

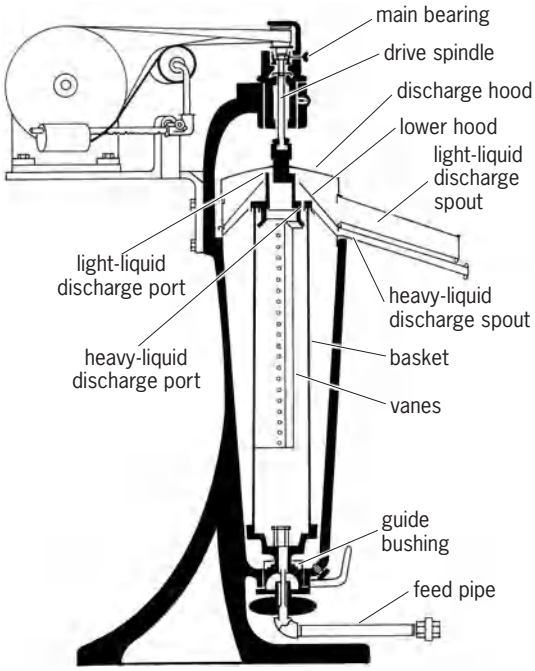

Fig. 1. Continuous high-speed clarifying centrifuge. (*After W. L. Bader and J. T. Banchero, Introduction to Chemical Engineering, McGraw-Hill, 1955*)

When the settling rate is too slow for gravity sedimentation to be feasible, centrifugal force up to approximately 100,000 times the force of gravity may be employed to clarify a suspension. In a separator called a centrifuge the suspension is passed through a rapidly rotating tube or basket (**Fig. 1**). The relatively dense solids are forced to the walls of the rotor, and the clarified liquid overflows a dam in the end of the centrifuge. *See* CENTRIFUGATION; ULTRACENTRIFUGE.

Clarifying filters comprise a versatile class of mechanical separators that are effective over the entire range of particle sizes encountered in clarification (from 0.1 mm down to molecules with a molecular weight on the order of 1000), regardless of relative densities and in spite of high viscosities. Sand or coal filters are used extensively for potable and industrial water clarification. Precoat filters are used for water and suspensions carrying slimy, amorphous impurities. Disk and plate filters (**Fig. 2**) service a wide variety of clarifying demands, removing particles over a range of 0.1 to 5 μm. Cartridge filters are in-line units with cleanable or replaceable elements that can remove particles in the range of 0.5 to 50 μm. Ultrafilters (also called microfilters and hyperfilters) contain fibrous sheets or porous membranes with carefully sized pores. These can remove particles as small as 0.002 μm. *See* FILTRATION; STERILIZATION; ULTRAFILTRATION.

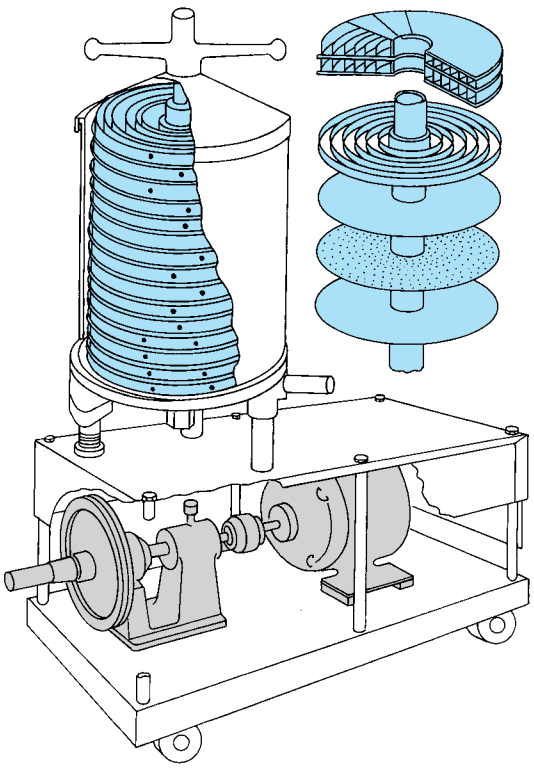

Fig. 2. Disk-and-plate clarifying filter. (*After R. H. Perry and D. Green, eds., Perry's Chemical Engineers' Handbook, 6th ed., McGraw-Hill, 1984*)

Magnetic separators are effective for the removal of ferrous particles. *See* MAGNETIC SEPARATION METHODS.

Clarification processes often depend on flocculating agents for their success. These are additives that cause the suspended solids to agglomerate into flocs that settle more rapidly or are more readily filtered than their smaller components. Shelby A. Miller

Bibliography. C. O. Bennett and J. E. Myers, *Momentum, Heat and Mass Transfer*, 3d ed., 1993; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineer's Handbook*, 7th ed., 1991; W. L. McCabe, J. C. Smith, and P. Harriot, *Unit Operations of Chemical Engineering*, 5th ed., 1993.

Classical field theory

The mathematical discipline that studies the behavior of distributions of matter and energy when their discrete nature can be ignored; also known as continuum physics or continuum mechanics. The discrete nature of matter refers to its molecular nature, and that of energy to the quantum nature of force fields and of the mechanical vibrations that exist in any sample of matter. The theory is normally valid when the sample is of laboratory size or larger, and when the number of quanta present is also very large. *See* PHONON; PHOTON; QUANTUM MECHANICS.

Classical field theories can be formulated by the molecular approach, which seeks to derive the macroscopic (bulk) properties by taking local averages of microscopic quantities, or by the phenomenological approach, which ignores the microscopic nature of the sample and uses properties directly measurable with laboratory equipment. This article follows the phenomenological course, although the microscopic treatment sometimes yields profounder insights. The problems considered below illustrate a selection of classical field theories without long mathematical discussions. Partial differential equations are used because the neglect of microscopic structure allows quantities like density and pressure to be expressed by continuously varying numbers.

Bending of a beam. A first example, leading to an ordinary differential equation, is the bending of a solid beam. To calculate the bending, isolate a segment of length dx (see **illus.**), and take account of the conditions that produce a mechanical equilibrium. The forces exerted up and down on the beam must balance, as well as the moments or torques tending to turn it. These conditions, combined with some geometry, lead to an equation from which the curvature of the beam can be calculated.

The quantity $M(x)$ is defined to be the moment or torque exerted by the part of the beam to the right of point x on the part to the left, and similarly $S(x)$ is the shear force exerted from right to left. If $\rho(x)$ is the beam's linear density, it follows from definitions that S and M change as in Eqs. (1) and (2), where g is

$$dS = -\rho g dx \qquad \frac{dS}{dx} = -\rho g \tag{1}$$

$$dM = S dx \qquad \frac{dM}{dx} = S \tag{2}$$

the acceleration of gravity. To find the bending, the beam is assumed to consist of parallel fibers; those above the center line are in tension and those below in compression. It is easy to see that at any point x, M is as given in Eq. (3), where R is the beam's

$$M = \frac{E}{R} I \qquad I = \int y^2 dA \tag{3}$$

radius of curvature, E is Young's modulus of elasticity for the material, dA is an element of area, and I is the moment of inertia of the cross section. If the beam is nearly horizontal, Eq. (4) relates R to the second derivative of y. For the rectangular cross section

$$\frac{d^2y}{dx^2} = \frac{1}{R} \tag{4}$$

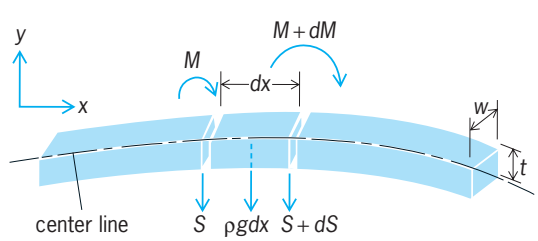

Portion of a rectangular solid beam, showing quantities used in calculating the bending of the beam.

shown, I is evaluated around the center line and is given by Eq. (5).

$$I = {}^{1}/_{12}\omega t^3 \tag{5}$$

See MOMENT OF INERTIA; SHEAR; STATICS.

Combining Eqs. (1)–(4) gives Eq. (6) for the shape

$$\frac{d^4y}{dx^4} = -\frac{pg}{EI} \tag{6}$$

of the beam. It can be integrated, requiring four boundary conditions, to find the curve under different conditions of loading; for example, if the beam is of length l and is supported horizontally at one end, the sag Y at the other end, given in Eq. (7), varies as

$$Y = -\frac{pg}{EI}\frac{l^4}{8t^3} \tag{7}$$

l^4/t^3. Historically, the analysis in terms of the moment of inertia I was begun by G. W. Leibniz in 1694, and results equivalent to the foregoing were obtained by C. Coulomb in 1773. The representation of the beam as a sheaf of fibers of infinitesimal thickness leads to the field equation (6), which is extremely simple in form and involves only the bulk properties of the material. *See* LOADS, TRANSVERSE.

Diffusion. In a vessel of fluid (liquid or gas) a small quantity of a different fluid is released and spreads slowly (in the absence of convection) through the vessel. To calculate the rate at which it spreads, let $n(x)$ be the number of molecules of the new fluid per unit volume, and $\mathbf{j}(x)$ be the flow (molecules per unit area per unit time) produced by any nonuniform distribution of n. This current has the direction in which n changes most rapidly, and the simplest assumption is that it is given approximately by Eq. (8),

$$\mathbf{j}(x) = -D\nabla n(x) \tag{8}$$

proposed by A. Fick in 1855, where D is the diffusivity, a coefficient which for simplicity is here assumed independent of x. If the molecules of the new fluid are neither created nor destroyed, $\mathbf{j}$ satisfies the conservation equation (9), and Fick's law becomes the

$$\nabla \cdot \mathbf{j} + \frac{\partial n}{\partial t} = 0 \tag{9}$$

field equation (10) giving the variation of n with position and time. *See* CALCULUS OF VECTORS.

$$\frac{\partial n}{\partial t} = D\nabla^2 n \tag{10}$$

If the initial distribution of n is $n_0(x)$ and the region is so large that its boundaries are far away, this equation is easy to solve. Let $\tilde{n}(k)$ be the Fourier transform of $n_0(x)$, so that Eq. (9) is solved by Eq. (11) where

$$n(x, t) = \int \tilde{n}(k)\exp(i\mathbf{k}\cdot\mathbf{x} - k^2Dt)d^3k \tag{11}$$

the integration is over all values of k; this expression satisfies Eq. (9) whatever the function $\tilde{n}$ may be. If $t = 0$, the value of the integral is n_0, and if $t > 0$, Eq. (11) shows what happens as time goes on. For any $\tilde{n}$, as t becomes large, the distribution spreads and becomes uniform. (If $t < 0$, the integral shows correctly only in exceptional cases what happened earlier.)

The diffusion equation contains a single parameter, the diffusivity D. It can be determined from experiment and also in other ways, for diffusion can be analyzed as a random molecular process. The drag β that a diffusing particle experiences as it moves through the fluid is defined by Eq. (12), where v is

$$F = -\beta v \tag{12}$$

the particle velocity. Then the diffusivity can be calculated as in Eq. (13), where k is Boltzmann's constant and T is the absolute temperature. Evidently the warmer the environment and the less the resistance, the faster the diffusion. Equation (13) was first given by A. Einstein in 1905. *See* DIFFUSION.

$$D = \frac{kT}{\beta} \tag{13}$$

Temperature, T, is another field quantity in classical physics, and the argument used for diffusion makes it plausible that heat diffusing through a solid follows the same law. Equation (14) duplicates the

$$\frac{\partial T}{\partial t} = \alpha^2\nabla^2 T \tag{14}$$

diffusion equation and has the same solutions. The coefficient α is calculated in the theory of solids by methods that differ according to the physical properties of the solid. *See* CONDUCTION (HEAT).

Hydrodynamics. The calculation of the flow of a fluid presents great mathematical difficulties; in fact, the essentially chaotic nature of turbulence often makes it impossible to represent flow in terms of a continuously varying field. The following discussion is therefore restricted to nonturbulent motion. *See* TURBULENT FLOW.

The basic equation of fluid flow was given by L. Euler in 1755. In modern notation, it can be derived as follows: A small volume of the fluid moves with velocity $\mathbf{v}$. The force exerted on it by the pressure p of the surrounding fluid is

$$-\oint p d\mathbf{A}$$

where dA is an outward-directed element of area, and this is transformed by a vector identity into

$$-\int \nabla p dV$$

that is to say, the force on the element of fluid is $-\nabla p$ per unit volume. If p is the density of the fluid, Newton's law gives the equation of motion, Eq. (15), where the gravitational force on the element is included. It is more convenient mathematically to study changes in velocity at a given point in space than to follow an element of fluid as it moves.

$$\frac{\rho d\mathbf{v}}{dt} = -\nabla p + p\mathbf{g} \tag{15}$$

It is easy to show that Eq. (16) is valid, which gives Euler's equation (17).

$$\frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} \tag{16}$$

$$\frac{\partial \mathbf{v}}{\partial t} = -(\mathbf{v} \cdot \nabla)\mathbf{v} - \frac{1}{\rho \nabla} p + \mathbf{g} \tag{17}$$

Solutions of Euler's equation are as varied as the conditions under which fluids flow; to see what it does and does not contain let us look at a very simple case, the propagation of sound through a fluid. Let $p = p_0 + p'$ and $\rho = \rho_0 + \rho'$, where p_0 and ρ_0 are the ambient values, and assume that the amplitude is so small that second-order terms in p, ρ, and $\mathbf{v}$ can be neglected, as well as $\mathbf{g}$. In addition, $\mathbf{v}$ can be represented as $\nabla\phi$, where ϕ is called the velocity potential. Then Euler's equation gives Eqs. (18). Equations (19) compare the time derivative of Eqs. (18) with the conservation equation (9) expressed in terms of ϕ. An equation for ϕ is possible if p' can be expressed in terms of ρ'. To do this requires a hypothesis that is not included in either Euler's equation or the conservation law, namely, an equation of state for the fluid. If $(\partial p/\partial \rho)_0$ is denoted by c^2, comparison of Eqs. (18) and (19) gives Eq. (20),

$$\begin{aligned} \nabla\left(\frac{\partial \phi}{\partial t} + \frac{p'}{\rho_0}\right) &= 0 \\ \rho_0 \frac{\partial \phi}{\partial t} + p' &= 0 \end{aligned} \tag{18}$$

$$\begin{aligned} \rho_0 \frac{\partial^2 \phi}{\partial t^2} + \frac{\partial p'}{\partial t} &= 0 \\ \rho_0 \nabla^2 \phi + \frac{\partial \rho'}{\partial t} &= 0 \end{aligned} \tag{19}$$

$$\nabla^2 \phi - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} = 0 \tag{20}$$

which is d'Alembert's equation for a wave traveling with velocity c.

The foregoing example shows how the equation of motion for the acoustical field arises from three hypotheses: Newton's second law, the conservation of the fluid, and its equation of state. Whereas the previous examples of field equations could have been derived by reasoning from the microscopic properties of matter, Eq. (20) would have been very difficult to derive in this way. *See* FLUID-FLOW PRINCIPLES; HYDRODYNAMICS; SOUND; WAVE EQUATION; WAVE MOTION.

Electromagnetic field. The theory of the electromagnetic field, which J. C. Maxwell developed during the 1860s, was intended as a phenomenological theory of the ether analogous to the theories outlined above. There is of course no ether, and therefore no phenomenological properties, but this was not known at the time, and to discover the properties of ether was generally considered the most important task of physics. Maxwell based his work on a microscopic model of the ether involving tiny whirling vortices, but when it was finished the model's phenomenological constants had almost entirely disappeared from the equations; only one remained. The electric current could be calculated from the theory in two ways, and the constant, now called c, is determined by the ratio of the two resulting numbers. During the succeeding 40 years Maxwell's equations were regarded as a clever device that gave correct answers while avoiding the important questions.

In terms ofthe electrodynamic potentials, using the modern International System of Units (SI), Maxwell's equations can be written in the form of Eqs. (21); here $\mathbf{A}(x)$ and $\varphi(x)$ are vector and scalar potentials, respectively, and $\mathbf{j}(x)$ and $\rho(x)$ are the current and charge densities, respectively, that produce them. They occur on the right sides multiplied by μ and $1/\epsilon$, respectively, and since $\mathbf{j}$ is simply the result of moving ρ, the units are effectively different in the two equations. Their ratio is $c^2 = 1/(\epsilon\mu)$; in Maxwell's day its value was approximately known. In empty space, $\mathbf{j} = \rho = 0$, and the equations are d'Alembert's for waves traveling at a speed c. The correspondence between this c and the measured speed of light was convincing evidence that light is an electromagnetic phenomenon.

$$\begin{aligned} \nabla^2 \mathbf{A} - \epsilon\mu \frac{\partial^2 \mathbf{A}}{\partial t^2} &= -\mu \mathbf{j} \\ \nabla^2 \varphi - \epsilon\mu \frac{\partial^2 \varphi}{\partial t^2} &= -\frac{1}{\epsilon}\rho \end{aligned} \tag{21}$$

Maxwell's equations are still phenomenological in content, but they are now considered to describe the electromagnetic fields and potentials and not properties of the ether. They are fully relativistic and survive in form, though not exactly in content, in the modern theory of quantum electrodynamics. *See* ELECTROMAGNETIC RADIATION; MAXWELL'S EQUATIONS; QUANTUM ELECTRODYNAMICS; RELATIVISTIC ELECTRODYNAMICS.

Gravitational field. During the nineteenth century, tentative efforts were made to create a theory of gravity along the same lines as Maxwell's theory of electromagnetism. The results, however, were trivial, and the development of special relativity showed that they could not possibly be correct. Only with a reformulation of the entire problem in terms of curved space-time was Einstein able in 1915 to create a classical theory of the gravitational field. Like the classical theories mentioned above, it gives its results in terms of phenomenological constants: the speed of light and the force of attraction between two unit masses one unit distance apart. It is very successful in explaining not only the phenomena of gravitation known in 1915 but also various new facts that have been learned since. Unlike Maxwell's equations, however, Einstein's do not give reasonable results in the context of a quantum theory of gravity. They must represent some sort of approximation, but so far no successful quantum version has been developed with which they can be compared. *See* GRAVITATION; QUANTUM GRAVITATION; RELATIVITY.

Mathematical development. The foregoing examples illustrate some of the most important classical field theories. Solving the equations has produced a vast body of mathematics. Computers have aided in special calculations, but many mathematicians are working on the analytical theory of partial differential equations, and new results continue to be produced. *See* DIFFERENTIAL EQUATION. David Park

Bibliography. G. Carrier, *Partial Differential Equations*, 2d ed., 1988; G. Joos and I. M. Freeman, *Theoretical Physics*, 3d ed., 1986; L. Landau, *Course of Theoretical Physics*: *The Classical Theory of Fields*, 4th ed., 1980.

Classical mechanics

The science dealing with the description of the positions of objects in space under the action of forces as a function of time. Some of the laws of mechanics were recognized at least as early as the time of Archimedes (287?–212 B.C.). In 1638, Galileo stated some of the fundamental concepts of mechanics, and in 1687, Isaac Newton published his *Principia*, which presents the basic laws of motion, the law of gravitation, the theory of tides, and the theory of the solar system. This monumental work and the writings of J. d'Alembert, J. L. Lagrange, P. S. Laplace, and others in the eighteenth century are recognized as classic works in the field of mechanics. Jointly they serve as the base of the broad field of study known as classical mechanics, or Newtonian mechanics. This field does not encompass the more recent developments in mechanics, such as statistical, relativistic, or quantum mechanics.

The general principles of classical mechanics are stated in mathematical form. With mathematical logic, one can deduce countless possible motions of bodies and then compare the predictions with experimental observations. Classical mechanics illustrates the essential nature of a physical theory, and it is usually an important ingredient in or a starting point for the various branches of modern physics. Its study offers one the opportunity to become acquainted with mathematical techniques and procedures which are useful in other fields.

In the broad sense, classical mechanics includes the study of motions of gases, liquids, and solids, but more commonly it is taken to refer only to solids. In the restricted reference to solids, classical mechanics is subdivided into statics, kinematics, and dynamics. Statics considers the action of forces that produce equilibrium or rest; kinematics deals with the description of motion without concern for the causes of motion; and dynamics involves the study of the motions of bodies under the actions of forces upon them. An important example of a force whose effect on bodies is often studied is the Earth's gravitational force. For some of the more important areas of classical mechanics *see* BALLISTICS; CANONICAL COORDINATES AND TRANSFORMATIONS; CELESTIAL MECHANICS; COLLISION (PHYSICS); DYNAMICS; ENERGY; FORCE; GRAVITATION; KINEMATICS; LAGRANGE'S EQUATIONS; MASS; MOTION; PRECESSION; RIGID-BODY DYNAMICS; STATICS; WORK.

The mathematical structure of classical mechanics is essential to understand both the historical development and the current status of relativity theory and of quantum mechanics. *See* QUANTUM MECHANICS; RELATIVITY. Newell S. Gingrich; John L. Safko

Bibliography. T. L. Chow, *Classical Mechanics*, 1995; H. Goldstein, C. P. Poole, Jr., and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; J. V. José and E. J. Saletan, *Classical Dynamics: A Contemporary Approach*, 1998; Theme issue on classical mechanics and non-linear dynamics, *Amer. J. Phys.*, 72(4):423–559, April 2004.

Classical swine fever

A highly contagious epizootic disease of pigs, also known as hog cholera. The causative agent is a virus in the family Flaviviridae, genus *Pestivirus*. Due to its significant economic impact, this disease is the subject of statutory controls in a majority of countries. It has been eradicated from many areas, including the United States, Canada, Australasia, and most parts of Europe. Clinically and pathologically, it closely resembles African swine fever, which is caused by an unrelated virus.

Transmission. Classical swine fever affects domestic and feral pigs as well as European wild boar. Humans are not susceptible. The primary mode of transmission is by contact or proximity. Infected animals shed virus in all bodily secretions. The virus may also be spread by contact with contaminated equipment and vehicles. It can survive for many months in frozen or refrigerated meat from infected pigs, and is not inactivated by mild forms of curing. Most primary outbreaks of classical swine fever are attributable to the feeding of pigs with waste human foodstuffs containing uncooked pig meat. Pregnant sows that recover from classical swine fever infection may give birth to congenitally infected carrier piglets with or without clinical signs that are a source of infection for other pigs.

Clinical signs. Pigs of any age may be affected; however, clinical signs are more severe in young pigs. High fever, loss of appetite, and dullness are the most typical symptoms of the acute form of the disease. Other symptoms may include blotchy discoloration of the skin (particularly the extremities), incoordination and weakness of the hindquarters, constipation followed by diarrhea, conjunctivitis, and coughing. Death may occur within 4 to 7 days after onset of symptoms. Mortality is usually high among young animals. Older pigs often develop immunity and recover completely. Recovered sows may abort or may give birth to poor litters that experience congenital tremors and convulsions.

Some pigs develop chronic disease that is characterized by undulating fever, dullness, poor

performance (such as decreased daily weight gain), capricious appetite, and variable degrees of coughing, diarrhea, and emaciation. There may be joint swellings and ulceration of the skin. These animals will die after 2 or 3 months.

Piglets that have been infected during pregnancy will develop the late-onset form of the disease, and death will occur within a few months after birth.

Strains vary in virulence, and classical swine fever may still be suspected when milder signs occur in epizootic form. Low-virulence strains may produce few signs apart from reproductive failure in sows or congenital tremors in their offspring.

Pathology. A few days after the infection, white blood cells and thrombocytes (platelets) are greatly reduced. Petechial (small, round) to confluent hemorrhages occur in the body, but particularly in the lymph nodes, kidneys, laryngeal mucosa, and urinary bladder. Infarcts are sometimes seen in the spleen. In chronic cases, a characteristic lesion is button ulceration of the large intestine. In cases of congenital infection, a proportion of piglets may show incomplete development of the cerebellum or other developmental abnormalities.

Microscopic and ultrastructural lesions include signs of microthrombosis (blood clot formation in the capillaries and other small blood vessels) and apoptosis (programmed cell death) as well as activation of macrophages and platelets. Damaged capillary endothelial cells lead to either obstruction or dilation of the vascular lumen and affect permeability. In addition, there is depletion of lymphocytes in the lymphoid tissues.

Diagnosis. The majority of outbreaks of classical swine fever are diagnosed on clinical grounds, especially in the severe acute form of disease; however, it is essential to carry out laboratory tests to confirm a presumptive diagnosis. Differential diagnoses include African swine fever, pseudorabies, porcine epidemic abortion and respiratory syndrome, porcine dermatitis and nephropathy syndrome, parvovirosis, bacterial septicemias (for example, salmonellosis, leptospirosis, swine erysipelas), thrombocytopenic purpura (a bleeding disorder), and poisoning with the rodenticide warfarin. Infections of pigs with the related pestivirus from cattle (bovine viral diarrhea virus) can occasionally cause similar symptoms and cross-reactions in diagnostic laboratory tests. *See* LEPTOSPIROSIS; SALMONELLOSES.

Control and vaccination. There is no specific treatment for classical swine fever. In classical swine fever–free countries, disease outbreaks are controlled by a stamping-out policy; that is, once the disease has been confirmed all pigs on affected and suspected farms are slaughtered to prevent further spread. Vaccination is usually prohibited, as it interferes with international trade. In areas with dense pig populations this policy might be very costly. Vaccination programs are widely practiced in regions with enzootic infection. Vaccination can control the severe clinical effects of the disease but cannot completely prevent infection and spread of the virus. Recently marker vaccines have been licensed. These vaccines are promising tools in the control of the infection, since they allow the distinction between vaccinated and infected animals. (Serological tests can be performed that detect the presence of antibodies to viral proteins that are produced only by the infected animals.) Thus, they can prevent the unnecessary slaughter of healthy animals. Epidemics of classical swine fever in wild boar populations may be long lasting and difficult to control. They are a threat to any local domestic pig holding. Specific hunting measures and oral vaccination campaigns are suitable tools to shorten the duration of these epidemics. *See* ANIMAL VIRUS; SWINE PRODUCTION; VACCINATION.

Volker Moennig; Steven Edwards

Bibliography. V. Moennig, G. Floegel-Niesmann, and I. Greiser-Wilke, Clinical signs and epidemiology of classical swine fever: A review of new knowledge, *Vet. J.*, 165:11–20, 2003; M. B. Pensaert (ed.), *Virus Infections of Porcines*, 1989.

Classification, biological

Arrangements of the names of species and groups of species meant to convey how these organisms are related to one another. The purposes of biological classifications are to organize the known biodiversity, provide a framework that will incorporate newly discovered biodiversity, and provide a predictive framework for studying biodiversity. Systematics focuses on studies of biodiversity and relationships among organisms. Taxonomy is a subdiscipline of systematics that focuses on formal classifications of organisms. *See* ANIMAL SYSTEMATICS; BIODIVERSITY; TAXONOMY.

Hierarchy of Ranks

Most, but not all, biological classifications are hierarchical, group-within-group arrangements in which a larger group will include one to many smaller groups within it. A group, or taxon, is given a formal name. The name refers to the organisms that are classified in the group. For example, *Homo sapiens* is the taxon name for all of the organisms classified as living humans, whereas Eukaryota is the taxon name for all plants, animals (including humans), and fungi. The place within a hierarchy is traditionally indicated using ranks. *See* TAXON; TAXONOMIC CATEGORIES.

Domain Eukaryota (organisms with organelles)
- Kingdom Metazoa (multicellular animals)
  - Phylum Chordata (vertebrate animals and kin)
    - Class Mammalia (mammals)
      - Order Primates (primate mammals)
        - Family Hominidae (fossil and living humans)
          - Genus *Homo* (living humans and close relatives)
            - Species *Homo sapiens* (living humans)

Purposes and Applications of Modern Biological Classifications

Most systematists now agree that modern biological classifications must serve two purposes. First, they must serve the primary biological function of expressing the common ancestry relationships (phylogenies) of organisms when possible, that is, when we have reliable information from which we can infer these relationships. One of the major activities of systematists is discovering these phylogenies using the characters of organisms (from DNA to behavior) and building from their studies the Tree of Life. Second, biological classifications must serve as systems from which information about the relationships of organisms can be reliably obtained. This requires that the classifications be consistent with the Tree of Life for groups of organisms with reliable phylogenies. It also requires that names be used consistently so that all biologists understand their meaning.

Classifications have very important applications. For example, the manner in which bacteria are classified forms the basis for predicting the pathogenesis of newly discovered or unknown bacteria. The entire genomes of organisms, such as the mouse and roundworm, are used as model systems to understand the development and causes of disease. Where these organisms fit into a classification has great value in predicting how generalizable the results of this research might be and to which other organisms the findings might apply.

Medical researchers frequently use rats and mice as the first subjects of studies, but it is understood that research using monkeys and great apes is critical if the findings are to be applied to humans. This is because our biological classifications show a close evolutionary relationship between humans and chimps, and we know that the more a medical breakthrough applies to species closely related to humans, the better the chances that they are applicable to humans.

Development of Modern Biological Classifications

The history of the development of the purposes and uses of modern biological classifications has two major themes: (1) how biologists have agreed to name taxa in a way that the use of names is consistent across international borders and over generations and (2) how the function of classifications to serve the needs of modern evolutionary biology emerged with increasing understanding of the nature of taxa.

Use of consistent names. Our modern biological classifications date to classifications published by the eighteenth-century Swedish botanist Carol Linnaeus (1707–1778). Linnaeus gathered both his work and the works of others into volumes outlining his system for organizing plant (*Species Plantarum*, 1753) and animal diversity (*Systema Naturae*, 1758). Linnaeus was by no means the first person to classify plants and animals, but his system of binomial nomenclature (that is, giving each species a two-part name, such as *Homo sapiens*) was widely adopted.

Three International Codes of Nomenclature. By the midnineteenth century, the chaos caused by individual taxonomists proposing different names for the same groups resulted in international efforts to propose rules that govern both the way that names were formed and what names should be used for any particular group. These rules have evolved into the three International Codes of Nomenclature, one each for plants, animals, and prokaryotes (bacteria). Each code specified when the official nomenclature of a particular group begins. For example, animal names proposed before 1758 (that is, before publication of the 10th edition of *Systema Naturae*) cannot be used; only names proposed in the *Systema Naturae* or after the publication of this work are permissible. Other works were selected for other groups as the starting point for names. Further, when possible, the oldest name that was proposed for a group is the official name of the group (Principle of Priority). Finally, each code includes a principle of typification.

Typification signals the intention of the person who names a group, usually by pointing to an actual specimen (for species names) or species (for group names) that the original author meant to include in the taxon. For example, the commercially important Indo-Pacific tuna commonly known as the kawakawa tuna was first described in 1849 by T. E. Cantor. The current name of this species is *Euthynnus affinis* (Cantor, 1849) and a dry skin of the fish used by cantor in his description can be found in the Natural History Museum, London. We understand that Cantor's intention was to apply this name to this specimen because he made reference to the specimen and it still exists and can be studied. We infer that this is the same fish that Cantor named by comparing the skin to specimens that we continue to collect and harvest for commercial purposes, taking into consideration modern concepts of how species in nature can vary and how the specimens we collect might differ from the type specimen as a result of this variation.

Naming is also regulated by the various Codes. For example, all Codes require that species have two names, the name of a genus and a specific epitaph. *Homo sapiens* is the name for living humans, not "*sapiens*." *Pinus ponderosa* is the name of the ponderosa pine tree, not "*ponderosa*." The genus name is always capitalized and the specific epitaph is never capitalized. And, the name is always set apart from the rest of the text, usually by using italics. The names of higher taxa (those that have ranks higher than species) may have specific endings and are frequently formed using the names of lower taxa. For example, in zoological nomenclature, priority and typification extends to the level of the rank family. Family names take on the root of a genus name and end in "idea." Subfamilies take on the name of a type genus and carry the ending "inae." The names of Tribes take on a genus name and end with "ini." Using this knowledge, we can actually "decode" the classification of the kawakawa tuna:

Kingdom Animalia (animals)
  Phylum Chordata (animals with notochords)
    Subphylum Vertebrata (animals with backbones)
      Superclass Gnathostomata (vertebrates with jaws)
        Class Osteichthyes (ray-fin fishes and tetrapods)
          Subclass Actinopterygii (ray-fin fishes)
            Order Perciformes (fishes related to perches)
              Suborder Scombroidei (mackerels, billfishes, and tunas)
                Family Scombridae (mackerels and tunas, but not billfishes)
                Subfamily Scombrinae (tunas, mackerels)
                Tribe Thunni (tunas but not mackerels)
                Genus *Euthynnus* (lesser tunas, but not the larger tunas)
                Species *Euthynnus affinis* (Cantor, 1849)

These efforts to regulate the names of organisms were designed to result in an international language of classification; biologists in China would use the same name for the same organisms as biologists in Sweden. Since Latin was the unofficial language of western scientists when Linnaeus formed his systems of classification, Latin continued to be used. Although cumbersome for the layperson, this has effect of producing a "neutral" language for international communication, setting aside national or regional preferences.

Zoological Code. From Kingdom to Suborder, the rules of priority and typification do not apply in the Zoological Code, so zoologists are free to use the names they wish, subject to general community consensus. However, some names are traditionally derived from lower group names. For example, Perciformes is a name formed around the genus name *Perca*, to which the common yellow perches of Europe and North American belong. At the rank level Family, the rules begin to apply. Scombridae and Scombrinae refer to all fishes of the family that includes the genus *Scomber* (mackerels), Thunni refers to a tribe that includes the large tunas of the genus *Thunnus* and relatives, including *Euthynnus; E. affinis* is one of three species of that genus. The form of the name is subject to the grammatical rules of Latin, thus "Scombridae" rather than "Scomberidae." *See* ZOOLOGICAL NOMENCLATURE.

Botanical and Bacteriological Codes. The rules for plants and prokaryotes (bacteria) work a bit differently and are independent of zoological rules. The names of prokaryotes names must be registered in the approved Lists of Prokaryotic Names, but no such requirement exists for zoological or botanical names. Name endings of plant and prokaryote groups are mostly similar, but they differ from the endings for animal groups. In botany, typification is applied to all ranks, whereas in prokaryotes, it is applied only to ranks below class, and thus naming is more rigid. For example, the classification of the ponderosa pine tree is as follows:

Kingdom Plantae
  Division Pinophyta (the conifers)
    Class Pinopsida (the conifers again)
      Order Pinales (living conifers)
        Family Pinaceae (pines but not yews or cypresses, etc.)
          Genus *Pinus* (pines but not spruces, etc.)
            Species *Pinus ponderosa*

This classification is very simple to decode because each name after Plantae refers to conifers, and *Pinus* is the type name of the entire Division, so all names leading to *Pinus* are formed around this name. Note that some names refer to the same group. For example, Pinophyta contains only one class, so Pinophyta and Pinopsida refer to exactly the same group of organisms. This redundancy is one of the quaint features of nomenclature when the rules state that each category rank be associated with a name. The first name, Pinophyta, serves a purpose, which is to classify conifers relative to other plants. The second name is redundant (and thus biologically unnecessary) but was retained to present a complete hierarchy of ranks. *See* BACTERIAL TAXONOMY; PLANT TAXONOMY.

Changing ideas of taxa. The International Codes have largely achieved their goal of promoting the consistent and regular use of taxon names. However, from the eighteenth century to the middle of the twentieth century, there was no general agreement on what constituted a "natural" classification. There were many ideas. Some of them, such as grouping in fives, seem quite quaint to modern readers. As late as the 1960s, there were three general "schools" competing for the attention of systematists (the others having faded into history). The evolutionary taxonomists, who emerged during the 1930s with the rise of the Neo-Darwinian Revolution, asserted that classifications should be based on phylogeny and similarity. The pheneticists, who emerged with the Computer Revolution of the 1950s, asserted that ideas of phylogeny were speculative and that taxonomic classification should be based on similarity alone. The phylogeneticists (cladists), led by the German entomologist Willi Hennig, emerged in the 1960s and asserted that there was a rigorous method for building phylogenies and that classifications should be based on phylogeny (common ancestry) alone, letting similarity fall where it may.

The pheneticists lost, in spite of sophisticated computer algorithms for calculating similarity, because there were too many ways to calculate similarity. Similarity, it seems is in the eyes of the beholder, and each beholder has a different idea. In addition, classification based purely on similarity mixes up the different types of similarities. That is, some similarities may be due to descent from a

common ancestor (the evolutionary concept of homology), whereas other similarities may be gained independently, perhaps due to similar environmental selection pressures (the evolutionary concept of convergence). Classifications that mix the two kinds of similarity might miss some important phenomena. The evolutionary taxonomists also lost because their idea of basing classifications on a combination of phylogeny and similarity would make it difficult to determine whether a particular classification was based on phylogeny or on similarity (the use of which had already been rejected).

The phylogeneticists won for two reasons. First, they demonstrated rigorous methods for the reliable reconstruction of common ancestry relationships. Second, they demonstrated that their classifications had predictive power. However, there are costs; for example, many traditional taxonomic groups (such as Reptilia) are no longer recognized, and some groups traditionally ranked at very high levels (such as Aves) have been demoted. The group Reptilia has been abandoned because birds are more closely related to crocodiles than crocodiles are to lizards. Reptilia was what Hennig called a "paraphyletic group," a group that did not include all the descendants of a common ancestor. Only "monophyletic groups," groups that include all the descendants of a common ancestral species, are permitted. Otherwise, the classification becomes inconsistent with the phylogeny, and efficient retrieval systems cannot be based on inconsistency. Thus, groups such as Reptilia no longer appear in the most modern classifications even if they persist in introductory and general-interest publications. Further, there is no longer a "Class" Aves. Birds and crocodiles are classified along with dinosaurs in the taxon Archosauria. If ranks were applied (they are not in the National Center for Biotechnology Information, NCBI, taxonomy system), Aves would probably be a suborder (as would Mammalia), not a class. Further, certain questions simply disappear. For example, the question of whether *Archaeopteryx lithographica*, the famous "missing link" between birds and reptiles, is a bird or a reptile disappears if "Reptilia" is not a natural group and is thus unrecognized in taxonomy. Rather, two new questions appear: What is the relationship of *A. lithographica* to Aves and other archosaurs? Is this species more closely related to birds than to, for example, dromaeosaur dinosaurs (the raptors made famous in the movie "Jurassic Park")?

Biological meaning of classifications. The Codes of Nomenclature are designed to stabilize the names applied to organisms and regulate how names are formed. However, the modern rules are not meant to impose a particular view of what the names mean biologically. The debate over the biological meaning of classifications is significant because classifications form the framework for comparing organisms.

From the time of Linnaeus to the middle of the twentieth century, taxonomists generally thought that the rank of a taxon had some biological meaning. Exactly what was meant varied from taxonomist to taxonomist and from system to system, but generally the fact that a taxon was classified at the rank "Family" as opposed to the rank "Subfamily" was thought to confer some significance on the taxon that made it comparable to other taxa of the same rank. In one system this might take on the significance of some perceived perfection, echoing "the great chain of being," with classifications arranged from "lower" forms of life to "higher" forms of life, with humans at the pinnacle, an idea quite popular with medieval philosophers. In other cases, it might have signified the uniqueness of form or adaptation; this approach treated the arrangement of organisms as a form of strict classification in which the ranks formed classes and the classes had biological significance quite apart from the taxa. Thus, a group of fishes classified at the level in the hierarchy of Family (for example, mackerels and tunas in Scomberidae) was considered comparable to a group of frogs placed at the same Family level (for example, toads in Bufonidae).

Classification to systematization. Today, systematists are moving away from this old idea of classification. Linnaean ranks are meant only to place taxon names in the hierarchy. A family of frogs and a family of fishes are no longer considered equivalent in terms of some biological attributes; they are ranked as families only because that is where they happen to fit in the classification hierarchy relative to other fishes or other frogs. The exception is the species category. Species are hypothesized units of process (species give rise to other species through various evolutionary processes) and thus all recognized species are thought to be comparable. Monophyletic groups of species are not units of process, but they are units of history.

Using ranks to show relationships rather than to express a biological level of organization is referred to as systematization rather than classification. As biological classifications change to reflect strictly evolutionary/common ancestry relationships rather than grades or levels of organization, they no longer reflect ancient and outmoded ideas reflected in older notions of "classification." This is because we have come to understand that taxa, groups of organisms in nature, are not the static classes of Aristotle but the dynamic evolutionary units of Darwin. We can express the evolutionary relationships of these groups through the use of Linnaean ranks (such as family, order, genus) because the Linnaean system is a hierarchical system easily adapted to express the relationships that result from evolutionary descent. But, the ranks no longer have the same meanings; they are simply placeholders for which we could substitute any number of alternative systems, such as numerical codes or even simply the phylogeny. For example, the Tree of Life Project uses taxon names and tree graphs to express the relationships among taxa, dispensing with Linnaean ranks (but not Linnaean names) entirely. The NCBI, a repository for all the genetic information known for most organisms, is very consistent in following the various Codes in naming taxa but dispenses with ranking in favor of a linked indented list that can be used to navigate within and between taxa.

In the modern era, classifications are meant to convey specific information as to the evolutionary

relationships of organisms to form predictive systems for comparison. Given this, is there any use for the seemingly outmoded, eighteenth-century systems of nomenclature represented by the various Linnaean Codes of Nomenclature? Interestingly, the answer is "yes." The major functions of the codes are to promote stability and correct use of the names of taxa. That these names now refer to groups of evolutionary significance rather than to static classes of preevolutionary taxonomists does not mean that the need for stable and universally understood names has disappeared. In fact, it means just the opposite: Natural taxa are the result of evolutionary descent and thus have biological significance and biological predictability. The need to have international codes that govern the use of these names is as important now as in the past. E. O. Wiley

Bibliography. J. Cracraft and M. J. Donoghue (eds.), *Assembling the Tree of Life*, Oxford University Press, New York, 2004 (an overview of the phylogeny of life, from bacteria to humans); W. Hennig, *Phylogenetic Systematics*, University of Illinois Press, Urbana, 1996 (phylogenetics); D. L. Hull, *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*, University of Chicago Press, 1988 (an account of the three schools by a philosopher of science); *International Code of Botanical Nomenclature (St Louis Code), Regnum Vegetabile 138*, Koeltz Scientific Books, Königstein, 1999; *International Code of Nomenclature of Bacteria* (1990 revision), *Bacteriological Code*, American Society for Microbiology, Washington, DC, 1999; *International Code of Zoological Nomenclature*, 4th ed., International Trust for Zoological Nomenclature, London, 1999; E. Mayr and P. D. Ashlock, *Principles of Systematic Zoology*, 2d ed., McGraw-Hill, New York, 1991 (evolutionary taxonomy); P. H. A. Sneath and R. R. Sokal, *Numerical Taxonomy: The Principles and Practice of Numerical Classification*, W. H. Freeman, San Francisco, 1973 (phenetics); E. O. Wiley, *Phylogenetics: The Theory and Practice of Phylogenetic Systematics*, Wiley, New York, 1981 (phylogenetics).

Clathrate compounds

Well-defined addition compounds formed by inclusion of molecules in cavities existing in crystal lattices or present in large molecules. The constituents are bound in definite ratios, but these are not necessarily integral. The components are not held together by primary valence forces, but are the consequence of a tight fit preventing the smaller partner, the guest, from escaping from the cavity of the host. Thus, the geometry of the molecules is the decisive factor.

Clathrate compounds are of theoretical interest because they represent a special type of bonding. In practice they can be used for separation of hydrocarbons, for stablization of drugs or pesticides, and as enzyme models.

Inclusion compounds can be subdivided into (1) lattice inclusion compounds (inclusion within a lattice which, as such, is built up from smaller single molecules); (2) molecular inclusion compounds (inclusion into larger ring molecules with holes); and (3) inclusion compounds of macromolecules (see **table**).

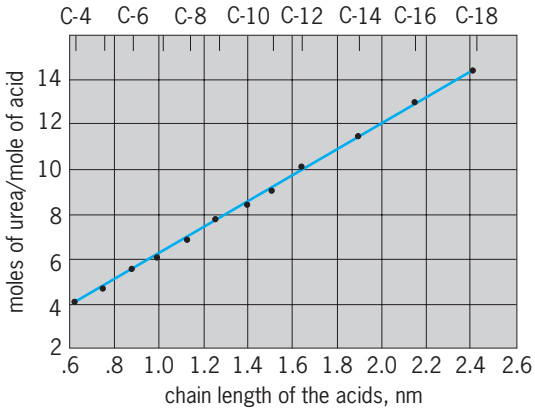

Fig. 1. Dependence of the urea/fatty acid composition of the inclusion compound on the chain length of the fatty acid. (*After W. Schlenk, Jr., Organic occlusion compounds, Fortschr. Chem. Forschg., 2:92–145, 1951*)

The best-known lattice inclusion compounds are the urea and thiourea channel inclusion compounds, which are formed by mixing hydrocarbons, carboxylic acids, or long-chain fatty alcohols with solutions of urea. Urea, in this case the host lattice, crystallizes in the presence of the other molecules in a particular crystal structure which contains long channels or tubes 0.4–0.5 nanometer in diameter. The hydrocarbons are accommodated in these channels, but only unbranched molecules find enough space to fit into the urea channel. The molar ratio of urea to hydrocarbon is simply determined by the lengths of the hydrocarbon molecule, not by any functional group. The longer the paraffin chain, the

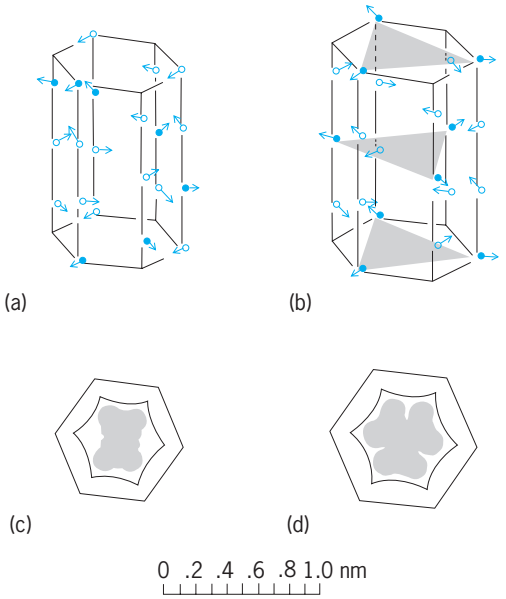

Fig. 2. Schematic drawing of the lattices of (*a*, *c*) urea and (*b*, *d*) thiourea; *c* and *d* show accommodation of unbranched and branched paraffin chains, respectively.

Inclusion compounds

Host substance	Shape of cavity	Guest substance examples
Lattice inclusion compounds		
Urea	Channel	Derivatives of straight-chain hydrocarbons
Thiourea	Channel	Branched-chain hydrocarbons
Deoxycholic acid	Channel	Paraffins, fatty acids, aromatic hydrocarbons
Dinitrodiphenyl	Channel	Diphenyl derivatives
Hydroquinone, phenol	Cage	Hydrogen chloride, sulfur dioxide, acetylene
Gas hydrates	Cage	Halogens, inert gases, hydrocarbons
Trithymotide	Cage	Benzene, chloroform
Nickel dicyanobenzene	Cage	Benzene, chloroform
Cyclodextrins	Channel or cage	Hydrocarbons, iodine, aromats
Molecular inclusion compounds		
Cyclodextrins	Cage	Hydrocarbons, iodine, aromats
Bis(*N,N′*-alkyl benzidines)	Cage	Benzene, dioxane
Crown ether compounds	Cage	Inorganic cations
Ionophore antibiotics	Cage	Inorganic cations
Inclusion compounds of macromolecular substances		
Clay minerals	Channel or sheet	Hydrophilic substances
Graphite	Sheet	Oxygen, bisulfate ion, alkali metals
Cellulose, starch	Channel	Hydrocarbons, dyes, iodine

more urea is necessary in order to envelop the paraffin (**Fig. 1**). Thiourea has a similar lattice structure, but a larger diameter of the channel. Therefore, only branched paraffins are included; the unbranched ones are too small, and cannot be held tightly (principle of closest packing), as shown in **Fig. 2**.

Other representatives of lattice inclusion compounds are the choleic acids, which are adducts of deoxycholic acid with fatty acids, and other lipoic substances. Most likely, these compounds play a role in the digestion process by emulsifying or dissolving fat and fat-soluble substances. Hydroquinone and phenol, cyclic anhydrides of phenol carboxylic acids, such as *o*-trithymotide and tetrasalicylide, as well as some other aromatic compounds, form an open crystal lattice which can accommodate smaller gas and solvent molecules (clathrates in the stricter sense of the word). The gas hydrates, which have been known since about 1810 (for example, $Cl_2 \cdot 6H_2O$), are inclusion compounds of gases in a somewhat expanded ice lattice. The gas or solvent molecules are inserted into definite places within the ice lattice and are surrounded by water molecules on all sides.

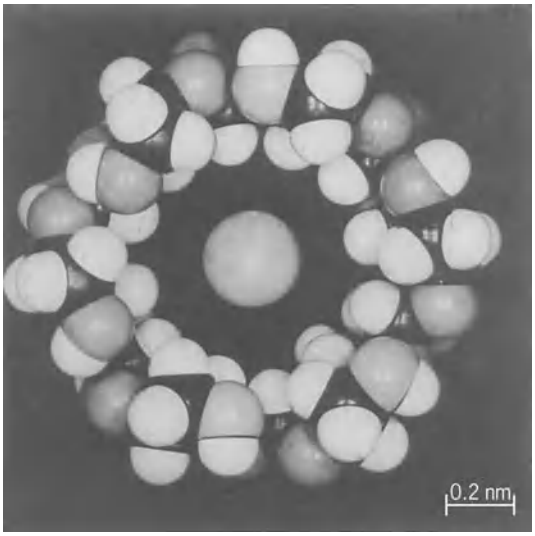

Fig. 3. Molecular model of α-dextrin–iodine compound seen in the direction of the *c* axis, that is, in the direction of the iodine chains, which lie in channels.

The cyclodextrins (cycloamyloses) are cyclic degradation products of starch which contain six, seven, or eight glucose residues (α-, β-, γ-cyclodextrin) and have the shape of large-ring molecules, the cavity diameters being 0.6–1.0 nm. Here, a great number of inclusion compounds in solution or in the crystalline state can be prepared with various kinds of molecules, the only requirement being the fit of the guest within the cyclodextrin cavity. α-Cyclodextrin forms blue iodine inclusion compounds in which the host molecules are stacked to form a cylinder reminiscent of the starch helix (**Fig. 3**). Polyiodide, located within the thus formed channellike cavity, consists of I_2, I_3^-, or I_5^- units in α-cyclodextrin, as well as in starch. A similar structure is likely for the blue iodine poly(vinyl alcohol) inclusion compound which exhibits a strong dichroism when stretched. This is used technically to prepare (light) polarizing sheets and spectacles.

Crown ether compounds are cyclic or polycyclic polyether compounds (for example, with the repeating unit . . . O—CH_2—CH_2. . .) capable of including another atom in the center of the ring. In this way, sodium or potassium compounds can be solubilized in organic solvents. Similarly, a series of ionophore antibiotics can complex inorganic cations. *See* MACROCYCLIC COMPOUND.

Some clay minerals are made up of distinct silicate layers. Between these layers some free space may exist in the shape of channels. Smaller hydrocarbon molecules can be accommodated reversibly within these channels. This phenomenon is used in some technical separation processes for separating hydrocarbons (molecular sieves). Furthermore, ion-exchange processes used for water deionization

are based on similar minerals. *See* CLAY MINERALS; MOLECULAR SIEVE; ZEOLITE.

Enzymes are believed to accommodate their substrates in active sites, pockets, or clefts prior to the chemical reaction which then changes the chemical structures of the substrates. These binding processes are identical to those of low-molecular-weight inclusion compounds. On the other hand, some low-molecular-weight inclusion compounds can mimic enzyme action, such as ester hydrolysis, but only with modest efficiency.

Friedrich Cramer; Wolfram Saenge; D. Gauss

Bibliography. J. L. Atwood, *Inclusion Phenomena and Molecular Recognition*, 1990; P. Bernier (ed.), *Chemical Physics of Intercalation II*, 1993; D. J. Cram and J. M. Cram, Design of complexes between synthetic hosts and organic guests, *Accounts Chem. Res.*, 11:8–14, 1978; S. A. Jenekhe (ed.), *Macromolecular Host-Guest Complexes*, 1992.

Clay

A term used to refer to the finest-grain particles in a sediment, soil, or rock. According to the Wentworth scale (see **table**), clay is finer than silt, characterized by a grain size of less than approximately 4 micrometers. However, the term clay can also refer to a rock or a deposit containing a large component of clay-size material. Thus clay can be composed of any inorganic materials, such as clay minerals, allophane, quartz, feldspar, zeolites, and iron hydroxides, that possess a sufficiently fine grain size. Most clays, however, are composed primarily of clay minerals. *See* CLAY MINERALS; FELDSPAR; QUARTZ; ZEOLITE.

Although the composition of clays can vary, clays can share several properties that result from their fine particle size. These properties include plasticity when wet, the ability to form colloidal suspensions when dispersed in water, and the tendency to flocculate (clump together) and settle out in saline water. *See* COLLOID.

Origin and occurrence. Clays occur most abundantly in nature in soils, sediments, sedimentary rocks, and hydrothermal deposits.

Clays, together with organic matter, water, and air, are one of the four main components of soil. Clays can form directly in a soil by precipitation from solution (neoformed clays); they can form from the partial alteration of clays already present in the soil (transformed clays); or they can be inherited from the underlying bedrock or from sediments transported into the soil by wind, water, or ice (inherited clays). *See* SOIL.

Wentworth scale for size classification of sediment particles

Particle	Grain size, mm
Boulder	>256
Cobble	64–256
Pebble	4–64
Granule	2–4
Sand	1/16–2
Silt	1/256–1/16
Clay	<1/256

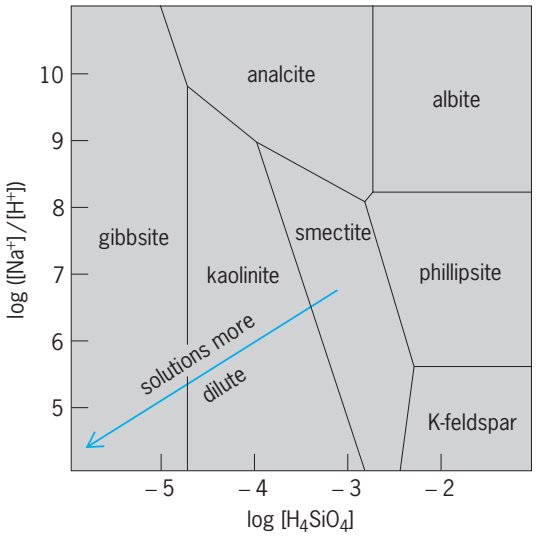

Example of a stability diagram for the system Na_2O-K_2O-Al_2O_3-SiO_2-H_2O drawn for log ($[K^+]/[H^+]$) = 4. The soluble species in brackets are in activity units. (*After J. I. Drever, The Geochemistry of Natural Waters, Prentice-Hall, 1982*)

The type of clays neoformed in a soil depends on the composition of the soil solution, which in turn is a function of climate, drainage, original rock type, vegetation, and time. Generally, neoformed clays that have undergone intense leaching, such as soils formed under wet, tropical climates, are composed of the least soluble elements, such as ferric iron, aluminum, and silicon. These soils contain clays such as gibbsite, kaolinite, goethite, and amorphous oxides and hydroxides of aluminum and iron. Clays formed in soils that are found in dry climates or in soils that are poorly drained can contain more soluble elements, such as sodium, potassium, calcium, and magnesium, in addition to the least soluble elements. These soils contain clays such as smectite, chlorite, and illite, and generally are more fertile than those formed under intense leaching conditions. The importance of the composition of soil solution to the type of clay neoformed at equilibrium can be shown quantitatively in mineral stability diagrams, such as that given in the **illustration**. *See* CHLORITE; GOETHITE; ILLITE; KAOLINITE.

Examples of clays formed by the transformation of other clays in a soil include soil chlorite and soil vermiculite, the first formed by the precipitation of aluminum hydroxide in smectite interlayers, and the second formed by the leaching of interlayer potassium from illite. Examples of inherited clays in a soil are illite and chlorite-containing soils formed on shales composed of these minerals. *See* SHALE.

Clays also occur abundantly in sediments and sedimentary rocks. For example, clays are a major component of many marine sediments. These clays generally are inherited from adjacent continents, and are carried to the ocean by rivers and wind, although some clays (such as smectite and glauconite) are

neoformed abundantly in the ocean. Clays also are a chief constituent of shale, are found in pore spaces in sandstones, and are found dispersed in limestones, dolomites, and coals. The ashes that remain after coal is burned, for example, are the residue of clays and silts washed into coal swamps and later incorporated into the coal. Hydrothermal clays can form abundantly where rock has been in contact with hot water or steam. Illite and chlorite, for example, form during the deep burial of sediments, and smectite and chlorite form by the reaction of hot, circulating waters at ocean ridges. Kaolinite can form hydrothermally where hot springs invade volcanic ash, given proper solution chemistries. *See* COAL; DOLOMITE; GLAUCONITE; LIMESTONE; MARINE SEDIMENTS; SANDSTONE; SEDIMENTARY ROCKS.

Properties and uses. Various clays possess special properties which make them important industrially. Bentonite, a smectite formed primarily from the alteration of volcanic ash, swells; is readily dispersible in water; and possesses strong absorptive powers, including a high cation exchange capacity. These properties lead to uses in drilling muds, as catalysts and ion exchangers, as fillers and absorbents in food and cosmetics, and as binders for taxonite and fertilizers. Swelling clays also must be taken into account in building design. In agriculture, swelling clays in a soil may lead to undesirable tilth and water runoff characteristics, and to unexpected reactions with soil additives, such as the catalytic conversion of pesticides into other compounds, or the fixation of ammonia fertilizer. *See* BENTONITE; CATALYSIS; ION EXCHANGE; SOIL CHEMISTRY.

Kaolinite-rich clays are used as paper-coating materials because they are white and show desirable rheological properties; as additives to rubber because they can bind with organic compounds; and as refractory materials due to their high melting point. Other important uses for clays include the manufacture of brick, ceramics, molding sands, decolorizers, detergents and soaps, medicines, adhesives, liners for ponds and landfills, lightweight aggregate, desiccants, molecular sieves, pigments, greases, paints, plasticizing agents, emulsifying, suspending, and stabilizing agents, and many other products. Ore deposits sometimes are composed of clays (for example, lateritic nickel and bauxites), and clay in shale may play an important role in the generation of petroleum. *See* CLAY, COMMERCIAL; REFRACTORY.

Dennis D. Eberl

Bibliography. R. Bennett and M. H. Hulbert, *Clay Microstructure*, 1986; A. Meunier, *Clays*, 2005; B. Velde (ed.), *Origin and Mineralogy of Clays: Clays and the Environment*, vol. 1, 1995.

Clay, commercial

Commercial clays, or clays utilized as raw material in manufacturing, are among the most important nonmetallic mineral resources. The value of clays is related to their mineralogical and chemical composition, particularly the clay mineral constituents kaolinite, montmorillonite, illite, chlorite, and attapulgite. The presence of minor amounts of mineral or soluble salt impurities in clays can restrict their use. The more common mineral impurities are quartz, mica, carbonates, iron oxides and sulfides, and feldspar. In addition, many clays contain some organic material. *See* CLAY MINERALS.

Clay is an abundant raw material; however, certain high-grade types of clay deposits are limited in geographic occurrence and extent. Examples are the white kaolin clays found in abundance only in Georgia and South Carolina; the high-quality bentonites found in Wyoming, South Dakota, California, Texas, and Mississippi; and the attapulgite clays found in Georgia and Florida.

Kaolinitic clays. Clays containing a preponderance of the clay mineral kaolinite are known as kaolinitic clays. Several commercial clays are composed predominantly of kaolinite. These are china clays, kaolins, ball clays, fireclays, and flint clays. The terms china clay and kaolin are used interchangeably in industry. *See* KAOLINITE.

China clay. China clays are high-grade white kaolins found in the southeastern United States, England, Brazil, and many other countries. In Georgia and South Carolina, the largest producing states, the deposits are composed almost wholly of the mineral kaolinite and occur as lenticular-shaped bodies generally 6–25 ft (1.8–7.5 m) in thickness. Many grades of kaolin are used in the manufacture of ceramics (whitewares, refractories, and porcelain), paper, rubber, paint, plastics, insecticides, adhesives, catalysts, and ink. They are also used for many other industrial purposes.

Kaolin has the desirable properties of being white in color, very fine in particle size, nonabrasive (hardness 2–2.5 on Mohs scale), and chemically inert in most uses. The individual kaolin particle is a thin, flat, pseudohexagonal platelet, so tiny that approximately 10,000,000 spread on a postage stamp would form a layer thinner than the thickness of a human hair. This thin, flat particle shape is a distinct advantage in many uses. The grades of kaolin that are commercially available generally are based on particle size, color, and brightness.

By far the largest consumer of white kaolins is the paper industry, which uses them to make paper products smoother, whiter, and more printable. The kaolin is used both as a filler in the sheet to enhance opacity and receptivity to ink and as a thin coating (about 0.005 in. or 0.013 cm) on the surface of the sheet to make it smoother and whiter for printing. The clay coating makes possible sharp photographic illustrations and brilliant color printing. Other important properties of kaolin for paper manufacture are softness of texture, ease of dispersion in the coating formulation, and ability to produce a high-gloss finish.

Because of the trend toward whiter coatings and lighter-weight paper (to minimize postal charges), the kaolin producers have developed new products. The colorant in kaolin is generally iron and titanium minerals; through flotation procedures, selective

flocculation, and high intensity magnetic separation, a portion of the iron and titanium can be removed, making a significantly whiter product. To aid in the development of lighter-weight coatings, a delaminated kaolin is produced. The thicker kaolin particles are separated by force to produce thin, white particles which give excellent coated-sheet properties with lighter coat weights.

In ceramics the kaolin clays are used in whitewares, sanitary ware, special refractories, and porcelain insulators. In whitewares and sanitary ware the kaolin provides a white body, gives easy molding properties, and adds strength, dimensional stability, and smooth surfaces to the finished product. Refractoriness, dimensional stability, and chemical inertness make the kaolins uniquely suitable for special refractories. In addition, the excellent dielectric properties of kaolin are very suitable for porcelain electric insulators. *See* CERAMICS; PORCELAIN.

In the manufacture of both natural and synthetic rubber, kaolin clays are used as a functional filler. The kaolin adds strength, abrasion resistance, rigidity, and whiteness. It costs less than most other fillers used in rubber.

In paint and ink kaolins are used as extenders because of their chemical inertness, smooth flow properties, ease of dispersion, nonabrasiveness, relatively low cost, and the high pigment volume concentrations that can be attained.

In plastics kaolin is used to provide smooth surfaces, dimensional stability, and better resistance to chemical attack, and to improve physical and electrical properties, to minimize water absorption, to reduce brittleness, and to lower costs.

In insecticides the inertness, low abrasiveness, and good adsorption and flow properties of kaolin make it an effective carrying agent in dust and spray forms.

A relatively new field is the chemical modification of kaolinite to make it an effective catalyst for use in petroleum refining. The kaolin is treated and calcined to produce zeolites, which are used as molecular sieves. Molecular sieves are widely used in petroleum refining.

Because the United States has very small reserves of bauxite, kaolin, which contains 38% Al_2O_3, has been studied as a source of alumina. Many patents utilizing either bases or acids to extract the alumina have been issued. Some companies which are basic producers of aluminum metal have acquired large reserves of kaolin, but to date no processing plants have been constructed.

Calcined clay is another commercial product of the kaolins. It is used in paint, ceramics, plastics, paper, and other materials. The calcined kaolins have the advantage of increased opacity and brightness. The kaolinite is changed during the calcination to the minerals mullite and cristobalite. The hardness after calcination is about 6.5 to 7 on Mohs scale, and therefore the calcined kaolin makes an exceptional, very fine polishing agent. This product is used in toothpaste, automobile polishes, and silver and brass polishes. High-temperature calcined kaolins (1803–2703°F or 984–1484°C) are used as grog by the refractory industry.

Ball clay. Ball clays are composed mainly of the mineral kaolinite but usually are much darker in color than kaolin. The term ball clay is used for a fine-grained, very plastic, refractory bond clay. Most ball clays contain minor amounts of organic material and the clay mineral montmorillonite. Usually ball clays are finer grained than china clays. This fineness, together with the montmorillonite and organic material, gives ball clays excellent plasticity and strength. For these reasons and because they fire to a light-cream color, ball clays are commonly used in whitewares and sanitary ware. Commercial deposits of ball clay are located in western Kentucky, western Tennessee, northern Mississippi, and northeastern Texas.

Fireclay. This term is used for clays that will withstand temperatures of 2730°F (1500°C) or higher. Such clays are composed primarily of the mineral kaolinite. The most common fireclays are the under-clays, which occur directly under coal seams. Not all under-clays are fireclays because many contain materials that lower the fusion point to below 2730°F (1500°C). Fireclays are generally light to dark gray in color, contain minor amounts of mineral impurities such as illite and quartz, and fire to a cream or buff color. Most fireclays are plastic, but some are nonplastic and very hard; the latter are known as flint clays.

Fireclays are used primarily by the refractories industry. The refractoriness of fireclays is proportional to the amount of alumina present in the clay and is reduced by impurities such as calcium carbonate and iron compounds. The foundry industry uses fireclay to bind sands into shapes in which metals can be cast. From 10 to 15% fireclay is necessary to bind the sand grains together. Fireclay-bonded sands give the metal a smoother surface than when other clays, such as bentonite, are used, but equivalent or higher mold strengths can be attained with approximately one-third as much bentonite. In many cases even though the strength properties of bentonite-bonded sands are higher, fireclays are used if the cost of clay and transportation is lower. Large quantities of fireclay are mined in Pennsylvania, West Virginia, Kentucky, Ohio, and Missouri.

Flint clay. This term is used for a special type of fireclay. Flint clays are composed essentially of kaolinite and are also refractory but are nonplastic and very hard.

Diaspore clay. This clay is composed of the minerals diaspore and kaolinite. Diaspore is a hydrated aluminum oxide with an Al_2O_3 content of 85% and a water content of 15%. This mineral is quite hard (6.5–7 on Mohs scale) and very refractory. Diaspore clay is used almost exclusively by the refractories industry in making refractory brick. However, after calcination, it is sometimes used as an abrasive material. Diaspore clay is produced commercially in Missouri and Pennsylvania. It is found in nature mixed with flint and plastic fireclays. High-grade diaspore clay contains 65–70% Al_2O_3. When its Al_2O_3 content is 50–65%, this material is called burley clay.

Mullite. Mullite is a high-temperature conversion product of many aluminum silicate minerals, including kaolinite, pinite, topaz, dumortierite, pyrophyllite, sericite, andalusite, kyanite, and sillimanite. Mullite is used in refractories to produce materials of high strength and great refractoriness. Mullite does not spall, withstands the shock of heating and cooling exceptionally well, and is resistant to slag erosion. It is used in making spark plugs, laboratory crucibles, kiln furniture, saggars, and other special refractories. Mullite crystals occur as long needlelike or lathlike forms. The intermeshing of these crystals produces high-strength bodies. The only natural occurrence of mullite is on the island of Mull in the Western Isles of Scotland, and here it is found in such small quantities that it has no commercial significance.

Bentonites. Those clays that are composed mainly of the clay mineral montmorillonite and are formed by the alteration of volcanic ash are known as bentonites. They are used in many industries; the most important uses are as drilling muds, as bonding clays in foundries, as bonding agents for taconite pellets, and as adsorbents in many industries. *See* MONTMORILLONITE.

The largest producing areas of bentonite are Wyoming, western South Dakota, central and eastern Texas, and northeastern Mississippi. The Wyoming and South Dakota bentonites are known as sodium bentonites, because the montmorillonite has a relatively high sodium content. The Texas and Mississippi bentonites are known as calcium bentonites, because the montmorillonite has a relatively high calcium content. Sodium montmorillonite and calcium montmorillonite have significantly different physical properties which regulate their use and acceptance by many industries.

The term bentonite is a rock term, but in industrial usage it has become almost synonymous with swelling clay. Bentonite contains minor amounts of other clay minerals, particularly illite, chlorite, and kaolinite, and up to approximately 30% opal C-T, a partially amorphous form of silicon dioxide. The properties of the bentonites are largely dependent on the type and amount of montmorillonite present.

Drilling mud. Drilling mud consists of bentonite, water, a weighting material such as barite, and various electrolytes. This mud is circulated in the hole at all times while the drilling is progressing. The functions of the drilling mud are to remove the drill cuttings from the drill hole, to form a filter cake on the wall of the drill hole to keep the drilling fluid in the hole and prevent water in the rock formations from entering, to cool and lubricate the bit, and to prevent blowouts. Bentonite, especially the sodium variety, imparts high viscosity and shear strength to the mud and also imparts a property called thixotropy. A thixotropic body acts as a fluid when agitated and as a thick gel when standing. Because the properties imparted by the bentonites to the drilling fluid are so important to the success of the oil industry, these clays are an extremely important mineral commodity.

Bonding clay. Bentonites are used in the foundry industry to bind sands into desired shapes in which metals can be cast. Only 3–5% bentonite is needed to bond the sand grains together. Because of their fine particle size and their nature of their water adsorption, bentonites give the mold a higher green, dry, and hot strength than does any other type of clay.

Another use of bentonite as a bonding agent is in the taconite industry. Taconite, low-grade iron ore, is crushed into a fine powder and the iron particles are separated from silica and other rock particles magnetically or by flotation. The fine iron particles are pelletized with the aid of bentonite, of which 12–18 lb (5.4–7.2 kg) is used to pelletize a ton of ore.

Other uses. Bentonites are also used as bleaching clays to remove coloring matter from oils, and as clarifying agents for products such as wines and beers. Bentonite is used as an adsorbent for oils, insecticides, alkaloids, vitamins, proteins, and many other materials. Calcium montmorillonites generally have better decolorizing and adsorbent properties than the sodium variety. Bentonite suspensions in water are widely used in industry, in addition to that suspension used as a drilling mud by the petroleum industry. Some applications are as fire retardants, as media for suspending materials for medicinal purposes, and as media for suspending lumps of coal to separate them from heavier impurities. Sodium bentonites are the better suspending agents because of their higher viscosities. Large quantities of sodium bentonite are used to line irrigation ditches and to seal leaks in dams, because it forms an impermeable cake or film by swelling and filling the pores. Sodium montmorillonite swells in water to many times its original size. Other uses are as catalysts, emulsion stabilizers, deinking agents, suspending agents in enamels, plasticizers in ceramics and pastes, and ingredients in cosmetics. *See* BENTONITE.

Attapulgite clays. Attapulgite is a hydrated magnesium aluminum silicate with a needlelike shape. It is commercially produced from a small area in southwestern Georgia and northwestern Florida. The needlelike shape gives attapulgite its unusual properties; each individual needle is exceedingly small, about 1 micrometer in length and approximately 0.01 micrometer across. Attapulgite is used as a suspending agent, and gives high viscosity because of the interaction of the needles. Some commercial uses are as an oil well drilling fluid, in adhesives as a viscosity control, in oil base foundry sand binders, as thickeners in latex paints, in liquid suspension fertilizers, and as a suspending agent and thickener in pharmaceuticals. Other uses are as adsorbent beds for pets, floor adsorbents, carriers for agricultural chemicals, catalyst, flatting agent in paint, and anticaking agent. An unusual use is as a coating on paper used in NCR (no carbon required) multicopy business forms. The attapulgite acts as a catalyst to convert the dye intermediates to the colored dyes when the encapsulated intermediates are released by pressure. The names attapulgite and palygurskite are synonymous.

Miscellaneous clays. Most of these clays and shales contain mixtures of differing proportions of illite, chlorite, kaolinite, and montmorillonite, plus a variety of nonclay minerals. By far the largest user of these miscellaneous clays is the structural clay products industry. This industry manufactures brick, drain tile, sewer pipe, conduit tile, glazed tile, terra cotta, and other items.

The clays that can be used for the different structural clay products necessarily have wide variations in properties. For example, a clay that can be used for conduit tile must be very plastic and have high green and dry strength and uniform shrinkage; for drain tile or common brick, these properties do not have to be controlled so closely. In general, however, the properties of clays that are important in the manufacture of structural clay products are plasticity, green strength, dry strength, drying and firing shrinkage, vitrification range, and fired color. These properties are largely dependent on the composition of the clay and the particle size of the constituents.

Mining and processing. Almost all the commercial clays are mined by open-pit methods, with overburden-to-clay ratios ranging as high as 10 to 1. The overburden is removed by motorized scrapers, bulldozers, shovels, or draglines. The clay is won with draglines, shovels, or bucket loaders, and transported to the processing plants by truck, rail, aerial tramways, or belt conveyors, or as slurry in pipelines.

The clay is processed dry or in some cases, wet. The dry process usually consists of crushing, drying, and pulverizing. The clay is crushed to egg or fist size or smaller and dried usually in rotary driers. After drying, it is pulverized to a specified mesh size, such as 90% retained on a 200-mesh screen with the largest particle passing a 30-mesh screen. In other cases, the material may have to be pulverized to 99.9% finer than 325 mesh. The material is shipped in bulk or in bags. All clays are produced by this method.

Where more exacting products must be produced, as in the case of kaolins for the paper industry and bentonites for certain pharmaceutical uses, wet processing is used. In the case of kaolins, the material is slurried in water at the mine face at about 35–40% solids with the aid of a dispersant. The slurry is pumped to a degritting station, where the very coarse impurities are removed, and then through pipelines to the processing plant. The general processing steps are particle size fractionation, that is, separation into coarse and fine particles; magnetic separation and/or flotation to remove titanium and iron minerals, leaching out colloidal iron impurities; dewatering by filtration; and drying and pulverization. Fractionation is done by centrifuging or by gravity settling; dewatering by centrifuges, filter presses, or rotary vacuum drum filters; and drying by rotary, apron, drum, or spray driers. Most of the kaolin is shipped dry bulk by rail, but about 20% is bagged and shipped by rail or truck. Some kaolin is shipped in slurry form in tank cars at 70% solids.

Properties. Most clays become plastic when mixed with varying proportions of water. Plasticity of a materal can be defined as the ability of the material to undergo permanent deformation in any direction without rupture under a stress beyond that of elastic yielding. Clays range from those which are very plastic, called fat clay, to those which are barely plastic, called lean clay. The type of clay mineral, particle size and shape, organic matter, soluble salts, adsorbed ions, and the amount and type of nonclay minerals all affect the plastic properties of a clay.

Strength. Green strength and dry strength properties are very important because most structural clay products are handled at least once and must be strong enough to maintain shape. Green strength is the strength of the clay material in the wet, plastic state. Dry strength is the strength of the clay after it has been dried. Plasticity and green strength are closely related and are affected largely by the same variables. Dry strength is dependent on the proportion of fine particles present, the degree of hydration of the colloidal fraction, the method of forming the ware, and the extent of drying. The presence of a small amount of montmorillonite, which is of very fine particle size and highly hydrated, generally increases the dry strength.

Shrinkage. Both drying and firing shrinkages are important properties of clay used for structural clay products. Shrinkage is the loss in volume of a clay when it dries or when it is fired. Drying shrinkage is dependent on the water content, the character of the clay minerals, and the particle size of the constituents. Drying shrinkage is high in most very plastic clays and tends to produce cracking and warping. It is low in sandy clays or clays of low plasticity and tends to produce a weak, porous body. Montmorillonite in relatively large amounts (10–25%) causes excessive shrinkage, cracking, and slow drying. Firing shrinkage depends on the volatile materials present, the types of crystalline phase changes that take place during firing, and the dehydration characteristics of the clay minerals.

Vitrification range. The temperature range of vitrification, or glass formation, is a very important property in structural products. Vitrification is due to a process of gradual fusion in which some of the more easily melted constituents begin to produce an increasing amount of liquid which makes up the glassy bonding material in the final fired product. Some clays have a short vitrification range so that the temperature of the kiln must be very closely regulated. Illites, montmorillonites, and chlorites all have lower vitrification temperatures than kaolinite. Some nonclay minerals such as calcite and feldspar lower the vitrification temperature of the body by acting as fluxes. The degree of vitrification developed depends on the duration of firing as well as on the temperature attained. Usually the degree of vitrification is regulated by the amount of shrinkage and porosity that is needed in the final product. There is, however, a practical limit as to the amount of shrinkage and reduction in porosity than can be attained at a given temperature.

Color. Color is important in most structural clay products, particularly the maintenance of uniform

color. The color of a product is influenced by the state of oxidation of iron, the state of division of the iron minerals, the firing temperature and degree of vitrification, the proportion of alumina, lime, and magnesia in the clay material, and the composition of the fire gases during the burning operation. The best white-burning clays contain less than 1% Fe_2O_3. Buff-burning clays contain 1–5% Fe_2O_3 and red-burning clays contain 5% or more Fe_2O_3. Other constituents also affect the color, but finely divided iron-bearing minerals are considered to be the principal colorants.

Uses. All types of clay and shale are used in the structural products industry but, in general, the clays that are used are considered to be relatively low grade. Clays that are used for conduit tile, glazed tile, and sewer pipe are underclays and shales that contain large proportions of kaolinite and illite. The semirefractory plastic clays found directly beneath the coal seams make the best raw material for the above mentioned uses. Brick and drain tile can be made from a wide variety of clays depending on their location and the quality of product desired. Clays used for brick and drain tile must be plastic enough to be shaped. In addition, color and vitrification range are very important. For common brick, drain tile, and terra-cotta, shales and surface clays are usually suitable, but for high-quality face bricks, shales and underclays are used. Geographic location is a prime factor in the type of clay used for structural clay products because, in general, these products cannot be shipped great distances without excessive transportation costs. Many raw materials of questionable quality are utilized in certain areas because no better raw material is available nearby. *See* BRICK; TILE.

Lightweight aggregate is produced from many types of clays and shales. The raw material is heated very rapidly in rotary kilns or sintering machines to the temperature range between incipient and complete fusion. After fusion has developed a molten jacket around the particle, certain constituents, if present, release gas to produce bloating and vesiculation. The molten jacket must be viscous enough to retain the expanding gas. Shales containing illites, montmorillonites, or chlorites, plus nonclay materials such as carbonates, sulfates, and organic matter, are the best raw material for lightweight aggregate. Adequate testing of a clay or shale is the only way to determine whether or not it will make a suitable lightweight aggregate. Mineral and chemical composition, particle size distribution, fusion temperature, and final strength of the aggregate particles are all important factors which control the production of lightweight aggregate from a clay or shale. The presence of a high percentage of kaolinite causes the material to be too refractory and the presence of excess fluxing agents causes the material to melt much too rapidly to control the vesiculation necessary to make the material lightweight.

The cement industry uses large quantities of impure clays and shales. Clays are used to provide alumina and silica to the charge for the cement kiln. Generally, a suitable clay can be found in the area in which the cement is being manufactured. *See* CEMENT; POTTERY; REFRACTORY.

Haydn H. Murray

Bibliography. W. E. Brownell, *Structural Clay Products*, 1976; H. H. Murray, *Clays For Our Future*, 1999.; H. Van Olphen, *An Introduction to Clay Colloid Chemistry*, 2d ed., 1977, reprint 1991; B. Velde, *Introduction to Clay Minerals: Chemistry, Origins, Uses, and Environmental Significance*, 1992; W. E. Worrall, *Clays and Ceramic Raw Materials*, 1986.

Clay minerals

Fine-grained, hydrous, layer silicates that belong to the larger class of sheet silicates known as phyllosilicates. Their structure is composed of two basic units. (1) The tetrahedral sheet is composed of silicon-oxygen tetrahedra linked to neighboring tetrahedra by sharing three corners to form a hexagonal network (**Fig. 1**). The fourth corner of each tetrahedron (the apical oxygen) points into and forms a part of the adjacent octahedral sheet. (2) The octahedral sheet is usually composed of aluminum or magnesium in sixfold coordination with oxygen from the tetrahedral sheet and with hydroxyl. Individual octahedra are linked laterally by sharing edges (**Fig. 2**). Tetrahedral and octahedral sheets taken together form a layer, and individual layers may be joined to each other in a clay crystallite by interlayer cations, by van der Waals and electrostatic forces, or by hydrogen bonding.

Because clay minerals are nearly ubiquitous in the Earth's upper crust, they offer a unique record of earth processes and earth history. Thus, the study of clay minerals forms an important branch of the science of geology. It has been suggested also that clay minerals may have been a necessary precursor for the origin of life by providing protection for primitive organic molecules, and by catalyzing their transformation into more complex substances. *See* PREBIOTIC ORGANIC SYNTHESIS.

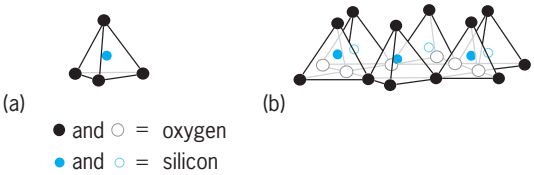

Fig. 1. Diagrammatic sketch showing (*a*) single silica tetrahedron and (*b*) sheet structure of silica tetrahedrons arranged in a hexagonal network. (*After R. E. Grim, Clay Mineralogy, McGraw-Hill, 1953*)

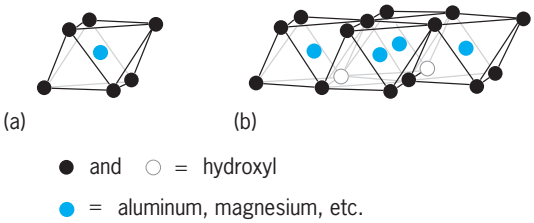

Fig. 2. Diagrammatic sketch showing (*a*) single octahedral unit and (*b*) sheet structure of octahedral units. (*After R. E. Grim, Clay Mineralogy, McGraw-Hill, 1953*)

Classification. Clay minerals are classified by their arrangement of tetrahedral and octahedral sheets. Thus, 1:1 clay minerals contain one tetrahedral and one octahedral sheet per clay layer; 2:1 clay minerals contain two tetrahedral sheets with an octahedral sheet between them; and 2:1:1 clay minerals contain an octahedral sheet that is adjacent to a 2:1 layer (see **table**).

Ionic substitutions may occur in any of these sheets, thereby giving rise to a complex chemistry for many clay minerals. For example, cations small enough to enter into tetrahedral coordination with oxygen, cations such as Fe^{3+} and Al^{3+}, can substitute for Si^{4+} in the tetrahedral sheet. Cations such as Mg^{2+}, Fe^{2+}, Fe^{3+}, Li^{+}, Ni^{2+}, Cu^{2+}, and other medium-sized cations can substitute for Al^{3+} in the octahedral sheet. Still larger cations such as K^{+}, Na^{+}, and Cs^{+} can be located between layers and are called interlayer cations. F^{-} may substitute for $(OH)^{-}$ in some clay minerals. *See* COORDINATION CHEMISTRY.

Clay minerals and related phyllosilicates are classified further according to whether the octahedral sheet is dioctahedral or trioctahedral. In dioctahedral clays, two out of three cation positions in the octahedral sheet are filled, every third position being vacant. This type of octahedral sheet is sometimes known as the gibbsite sheet, with the ideal composition $Al_2(OH)_6$. In trioctahedral clay minerals, all three octahedral positions are occupied, and this sheet is called a brucite sheet, composed ideally of $Mg_3(OH)_6$. Some dioctahedral and trioctahedral clay minerals and related phyllosilicates are listed in the table.

Clay minerals can be classified further according to their polytype, that is, by the way in which adjacent 1:1, 2:1, or 2:1:1 layers are stacked on top of each other in a clay crystallite. For example, kaolinite shows at least four polytypes: *b*-axis ordered kaolinite, *b*-axis disordered kaolinite, nacrite, and dickite. Serpentine shows many polytypes, the best-known of which is chrysotile, a mineral that is used to manufacture asbestos products. *See* KAOLINITE; SERPENTINE.

Finally, clays are named on the basis of chemical composition. For example, two types of swelling clay minerals are the 2:1, dioctahedral smectites termed beidellite and montmorillonite. The important difference between them is in the location of ionic substitutions. In beidellite, charge-building substitutions are located in the tetrahedral sheet; in montmorillonite, the majority of these substitutions are located in the octahedral sheet. Other examples of chemistry used in classification are nontronite, an iron-rich beidellite, and sauconite, a zinc-containing beidellite.

Another family of clay minerals is the chain clays, which have a structural resemblance to the chainlike arrangement of silica tetrahedra in pyroxene. In sepiolite and palygorskite, 2:1 layers are joined at their corners to form long channels that can contain water, a few exchangeable cations, and other substances.

Because clay minerals are composed of only two types of structural units (octahedral and tetrahedral sheets), different types of clay minerals can articulate with each other, thereby giving rise to mixed-layer clays. The most common type of mixed-layer clay is mixed-layer illite/smectite, which is composed of an interstratification of various proportions of illite and smectite layers. The interstratification may be random or ordered. The ordered mixed-layer clays may be given separate names. For example, a dioctahedral mixed-layer clay containing equal proportions of illite and smectite layers that are regularly interstratified is termed potassium rectorite. A regularly interstratified trioctahedral mixed-layer clay mineral containing approximately equal proportions of chlorite and smectite layers is termed corrensite.

Properties. Many of the properties of clay minerals are related to their crystal structure. Some of these properties are discussed below.

Kaolinite-serpentine group. These 1:1 layer silicates possess a *c* dimension of approximately 0.7 nanometer. The dioctahedral clays include the polytypes of kaolinite mentioned previously. The trioctahedral minerals include varieties of serpentine such as chrysotile, antigorite, lizardite, and amesite. These clays are nonswelling in water, with the exception of halloysite, a variety of kaolinite which can swell to about 1.0 nm. Kaolinites, however, can be made to swell by using intercalation compounds. Kaolinites can be distinguished from serpentines in x-ray diffraction analysis by their smaller *b* dimension, and by heat treatment: the kaolinite structure will decompose at 550°C (1020°F), whereas the serpentine structure will not. Also, serpentines are more susceptible to acid attack and will not intercalate. The 1:1 clay minerals possess a cation exchange capacity of 3–15 milliequivalents per 100 g, this arising mainly from broken bonds on crystal edges. Kaolinites are used in the manufacture of ceramics, paper, rubber, and medicine. A variety of serpentine (chrysotile) is used to manufacture asbestos products such as fireproof cloth and brake linings. *See* HALLOYSITE.

Pyrophyllite-talc group. These 2:1 layer silicates possess a *c* dimension of slightly more than 0.9 nm and exist in several polytypes and chemical varieties. Talc generally shows a triclinic structure (lTc), but a disordered variety is called kerolite. Minnesotaite is an iron-rich variety of talc, and willemseite is a nickel-rich variety. Polytypes of pyrophyllite are monoclinic, triclinic, and disordered pyrophyllite. Generally, these 2:1 minerals possess a high thermal stability, have a low cation-exchange capacity, and do not swell in water or in intercalation agents. *See* PYROPHYLLITE; TALC.

Smectite group. These 2:1 clays possess a variable *c* dimension because of their ability to swell. It ranges generally from 1.2 to 1.5 nm in air, depending on the type of interlayer cation, on the layer charge, and on the relative humidity. When saturated with ethylene glycol, however, smectites will swell to approximately 1.7 nm regardless of these factors, and therefore this treatment is a commonly used test for smectite. The ability to swell and other unusual properties associated with this group arise from isomorphous

Some dioctahedral and trioctahedral clay minerals and related phyllosilicates

Type	Layer structure*	Mineral	Idealized formula
		DIOCTAHEDRAL	
1:1		Kaolinite	$Al_2Si_2O_5(OH)_4$
2:1		Pyrophyllite	$Al_2Si_4O_{10}(OH)_2$
	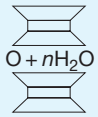	Dioctahedral smectite Beidellite Montmorillonite Dioctahedral vermiculite	 $Al_2(Si_{3.7}Al_{.3})O_{10}(OH)_2Na_{.3} \cdot nH_2O$ $(Al_{1.7}Mg_{.3})Si_4O_{10}(OH)_2Na_{.3} \cdot nH_2O$ $Al_2(S_{3.2}Al_{.8})O_{10}(OH)_2Mg_{.4} \cdot nH_2O$
	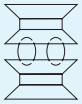	Illite Muscovite Celadonite Leucophyllite	 $Al_2(Si_{3.2}Al_{.8})O_{10}(OH)_2K_{.8}$ $Al_2(Si_3Al)O_{10}(OH)_2K$ $(Fe,Al,Mg)_2Si_4O_{10}(OH)_2K$
	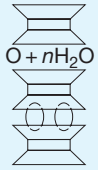	Mixed-layer illite/smectite	
2:1:1		Dioctahedral chlorite	$Al_4Si_4O_{10}(OH)_8$
		TRIOCTAHEDRAL	
1:1		Serpentine	$Mg_3Si_2O_5(OH)_4$
2:1		Talc	$Mg_3Si_4O_{10}(OH)_2$
	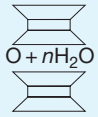	Trioctahedral smectite Saponite Hectorite Trioctahedral vermiculite	 $Mg_3(Si_{3.7}Al_{.3})O_{10}(OH)_2Na_{.3} \cdot nH_2O$ $(Mg_{2.7}Li_{.3})Si_4O_{10}(OH)_2Na_{.3} \cdot nH_2O$ $Mg_3(Si_3Al)O_{10}(OH)_2Mg_{.5} \cdot nH_2O$
	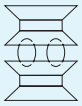	Phlogopite	$Mg_3(Si_3Al)O_{10}(OH)_2K$
2:1:1		Trioctahedral chlorite	$Mg_6Si_4O_{10}(OH)_8$
Chain	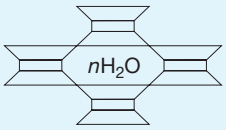	Sepiolite Palygorskite (attapulgite)	$Mg_8Si_{12}O_{30}(OH)_4(OHH)_4 \cdot 8H_2O$ $Mg_5Si_8O_{20}(OH)_2(OHH)_4 \cdot 4H_2O$

*Tetrahedral sheets are drawn as trapezoids, octahedral sheets as rectangles, and interlayer cations as circles. The *c* axis is approximately parallel to the table's columns.

substitutions in the tetrahedral and octahedral sheets which give rise to a negative charge on the 2:1 layers. This charge is balanced by interlayer cations which usually are hydrated.

The association of water with interlayer cations gives rise to crystalline (or first-stage) swelling for smectites. Smectities may undergo further osmotic (second-stage) swelling in wet environments if the interlayer cation is Na^+ or Li^+. In third-stage swelling, 2:1 layers are dispersed completely in water. An osmotically swelling smectite may be transformed into a smectite that undergoes only the more limited form of crystalline swelling by exchanging interlayer Na^+ for a divalent cation such as Ca^{2+}. Other properties related to the expanded interlayer region are a large cation-exchange capacity, a large surface area, and a Brönsted acidity related to a greater degree of dissociation for interlayer water. Many of these properties may be lost, either reversibly or irreversibly, by heating smectite in air to drive off interlayer water, thereby causing the 2:1 layers to collapse around interlayer cations. Water also can be driven out of the interlayer by increasing negative charge on 2:1 layers. This process happens spontaneously in nature during burial metamorphism, wherein layer charge is increased, leading to water expulsion and layer collapse around interlayer K^+, thereby transforming smectite into illite. The many industrial uses for smectite include drilling muds, binders, cation exchangers, and catalysts.

Vermiculite group. Clays in this group are similar to smectite, except that a larger layer charge leads to a decreased ability to swell, although some can swell osmotically under certain conditions. Their *c* dimension is about 1.4 nm, and generally, unlike smectite, they will not swell further in ethylene glycol. Many vermiculites will collapse to 1.0 nm on potassium exchange and on heating in air. Like smectite, these clays also possess special properties related to the expanding interlayer region, including a large cation-exchange capacity and surface area. Macroscopic vermiculite is exploded by heating in air for use in potting soil and insulation. Some micas and illites can be transformed into vermiculite by replacing interlayer K^+ with cations of greater hydration energy. *See* VERMICULITE.

Illite group. These clays possess a *c* dimension of about 1.0 nm. They are nonswelling and possess a small cation-exchange capacity, unless they are interlayered with smectite to form mixed-layer illite/smectite. Illite/smectites show properties intermediate between pure illite and smectite. Illite's layer charge and potassium content are less than those of a true mica. *See* ILLITE.

Chlorite group. Chlorites, like vermiculites, show a *c* dimension of about 1.4 nm, but unlike vermiculite, they will not collapse on heating or potassium saturation. Chlorites share many polytypes and chemical varieties. *See* CHLORITE.

Sepiolite and palygorskite. The structure of these minerals are similar, except that sepiolite shows a *b* dimension of about 1.21 nm, whereas that for palygorskite is about 1.04 nm. This difference arises from the fact that palygorskite incorporates only two linked pyroxenelike chains in each 2:1 layer rather than three. Both clays possess a fibrous morphology and channels which can accommodate water and a few cations. The ability of these channels to absorb organic substances gives rise to their use in clarification. Sepiolite also is used to manufacture meerschaum tobacco pipes. *See* SEPIOLITE; ZEOLITE.

Origin. A primary requirement for the formation of clay minerals is the presence of water, because clay minerals are hydrous silicates. Hence clay minerals are not expected to form on the Moon because there is no water on the Moon. Evidence for the presence of water on Mars, and the unusual catalytic properties of Martian soil, suggest that clay minerals may be present on this planet. Likewise, the presence of clay minerals in some meteorites suggests that these materials have been exposed to water during some part of their long and complex history.

Clay minerals form on Earth in many different environments, including the weathering environment, the sedimentary environment, and the diagenetic-hydrothermal environment. Clay minerals composed of the more soluble elements (for example, smectite and sepiolite) are formed in environments in which these ions can accumulate (for example, in a dry climate, in a poorly drained soil, in the ocean, or in saline lakes), whereas clay minerals composed of less soluble elements (for example, kaolinite and halloysite) form in more dilute water such as that found in environments that undergo severe leaching (for example, a hilltop in the wet tropics), where only sparingly soluble elements such as aluminum and silicon can remain. Illite and chlorite are known to form abundantly in the diagenetic-hydrothermal environment by reaction from smectite. *See* CLAY; CLAY, COMMERCIAL; LITHOSPHERE; SILICATE MINERALS.

Dennis D. Eberl

Bibliography. A. Meunier, *Clays*, 2005; A. C. Newman, *Chemistry of Clays and Clay Minerals*, 1986; B. Velde, *Introduction to Clay Minerals: Chemistry, Origins, Uses, and Environmental Significances*, 1992; B. Velde (ed.), *Origin and Mineralogy of Clays: Clays and the Environment*, 1995.

Clear-air turbulence

Turbulence above the boundary layer but not associated with cumulus convection. The atmosphere is a fluid in turbulent motion. That turbulence of a scale sensed by humans in aircraft is primarily associated with the boundary layer within a kilometer or so of the Earth, where it is induced by the surface roughness, or in regions of deep convection such as cumulus cloud development or thunderstorms. However, aircraft occasionally encounter turbulence when flying at altitudes well above the surface and far from convective clouds. This phenomenon has been given the rather unsatisfactory name of clear-air turbulence (CAT).

What is a primarily sensed in clear-air turbulence by the human is vertical acceleration. This acceleration will depend on the person's location in the plane, the speed of flight relative to the air, and the response characteristics of the airframe. A plane with a wing that generates aerodynamic lift more efficiently or an airframe with less weight per unit wing area will respond more strongly to a given gust magnitude. When allowances are made for different tolerance thresholds in different humans, a rough scale for subsonic aircraft has become accepted, with one-tenth the acceleration of gravity (g) as the threshold for light clear-air turbulence, 0.5 g for moderate, and 1 g for severe. *See* AEROELASTICITY.

Occurrence. Clear-air turbulence is encountered in the atmosphere with probability depending on flight altitude, geographical location, season of the year, and meteorological conditions. Given this variability and the small scale of the phenomenon, it is difficult to establish reliable statistics on the frequency of its occurrence. A number of extensive investigations have been carried out with this aim. One of these was a series of 90 flights by the National Aeronautics and Space Administration (NASA) supersonic aircraft YF-12A. At commercial-jet cruising levels (around 7.5 mi or 12 km), clear-air turbulence was encountered during approximately 12% of flight time in winter and 4% in summer. Above 11 mi (18 km) altitude (in the stratosphere) the frequency falls off rapidly. Other NASA studies have estimated clear-air turbulence frequency below 7.5 mi (12 km) to be somewhat increased, especially over mountainous terrain. The turbulent patches varied from 0.5 or 1.5 mi (1 or 2 km) in length to as large as 30 mi (50 km), although such large patches were very infrequent. The depth of the turbulent air was found to be almost always less than 0.6 mi (1 km).

Although clear-air turbulence may be encountered in unexpected meteorological contexts, there are highly favored locations for its occurrence. One is in the vicinity of the jet stream, particularly in ridges and troughs where the wind direction is turning sharply. A second and even more common location of occurrence is in the lee of a mountain range when a strong air flow is distorted by being forced over the range. In this situation a gravity lee wave is generated, which propagates to stratospheric heights. At various altitudes and distances from the mountain, this wave may break, and as many as a dozen or more patches of clear-air turbulence, light to severe, may be formed. *See* JET STREAM.

Forecasting and detection. Despite knowledge of these favored meteorological areas, it is not possible to forecast with confidence the precise location of a patch of clear-air turbulence. Warning forecasts for substantial portions of routes are typically given to pilots when conditions of clear-air turbulence prevail.

In view of this situation, effort has been expended to develop remote-sensing systems to give warning of a patch of clear-air turbulence with sufficient lead time to be avoided by minor diversions of flight path or at least to assure that all passengers are seated with safety belts fastened. Some ground-based techniques have the capability to produce a unique clear-air turbulence signature; of these, ultrasensitive radar probably has had the most demonstrable success. However, it is clearly impracticable to attempt to monitor any significant portion of commercial flight paths with costly ground-based instrumentation. Airborne infrared and microwave sensors, some using multiple frequencies, detect sharp temperature gradients or water-vapor concentration irregularities. The usefulness of these as clear-air turbulence warning systems is a subject of vigorous debate and investigation. *See* AERONAUTICAL METEOROLOGY; METEOROLOGY.

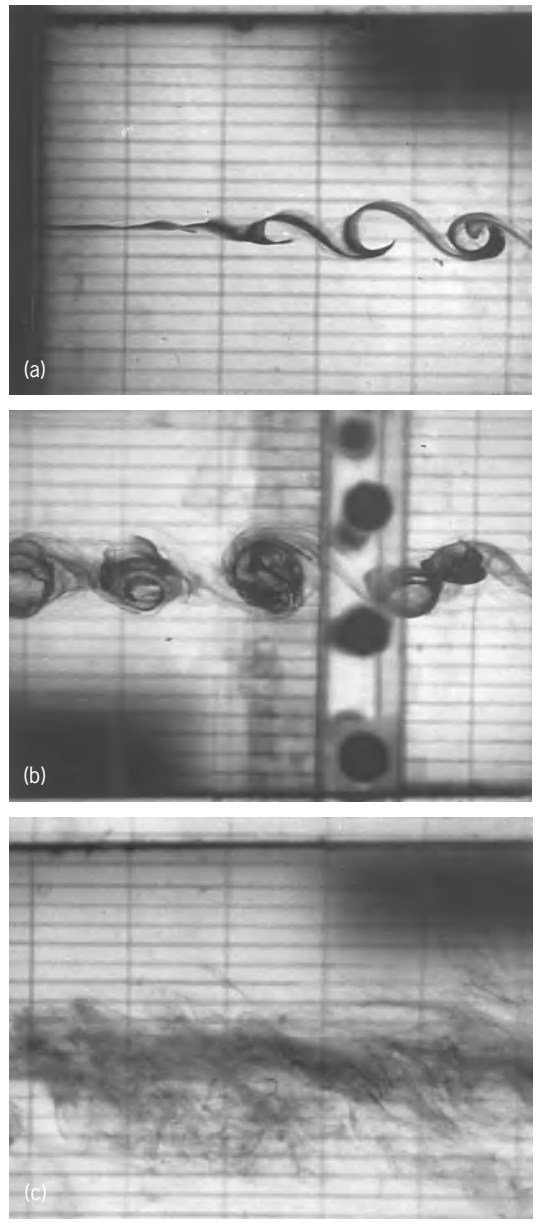

Laboratory realization of (*a*) Kelvin-Helmholtz instability, (*b*) the subsequent generation of turbulent eddies, and (*c*) mixing and dissipation. Fluid flow is left to right. (*Courtesy of Drs. G. G. Koop and F. K. Browand*)

Flight safety. Despite sophisticated design techniques applied to airframe response and despite pilot avoidance procedures, clear-air turbulence is still a

factor in 10% of air transport accidents and safety incidents in the United States. Of all weather-related accidents, at least 20% may be associated with turbulence in flight.

Origin and explanation. The early observational documentations of clear-air turbulence excited much theoretical interest and speculation as to the fluid-dynamic explanation of this phenomenon. The consensus of researchers is that clear-air turbulence is the result of Kelvin-Helmholtz instability, a breakdown of smooth flow that occurs when the destabilizing velocity gradient of the fluid becomes large relative to the stabilizing density gradient.

It was initially obvious that this hypothesis was consistent with the occurrence of clear-air turbulence in the jet stream and in distorted flow downwind of mountain ranges where velocity gradients are large. Further quantitative experiments in the field identified a number of clear-air turbulence events precisely under the theoretically predicted conditions. In the laboratory it has been possible to reproduce the instability mechanism and follow its breakdown into turbulence (see **illus.**). This is the dynamic mechanism widely thought to be the fundamental one in clear-air turbulence.

One way of viewing clear-air turbulence is as an increase in energy of the smaller scales (30 ft to 0.6 mi, or 10 m to 1 km) of atmospheric motion. This is best presented in a spectral diagram, that is, a graph of energy versus wavelength.

Aside from the practical implications of clear-air turbulence for air transport, this phenomenon plays a role of undetermined magnitude in the dissipation of the kinetic energy of the atmosphere. *See* UPPER-ATMOSPHERE DYNAMICS.

M. G. Wurtele; L. J. Ehernberger

Bibliography. F. K. Lutgens and E. J . Tarbuck, *The Atmosphere: An Introduction to Meteorology*, 7th ed., 1997.

Cleavage (developmental biology)

The subdivision of fertilized eggs (zygotes) into progressively smaller cells called blastomeres. **Figure 1** shows the development of the frog *Xenopus laevis* from the fertilized egg to the stage of the hatched larva (tadpole). At an intermediate stage called blastula, the *Xenopus* embryo is a hollow sphere of blastomeres surrounding a fluid-filled cavity called the blastocoel. Cleavage covers the period from egg to midblastula, when the *Xenopus* embryos consists of about 1200 cells. *See* BLASTULATION.

Cleavage shares the basic characteristics of most cell divisions. First, the cell nucleus is duplicated in a process known as mitosis. Then the cytoplasmic body of the cell divides so that each daughter cell inherits part of the cytoplasm along with one nucleus. The division of cytoplasm, called cytokinesis, may be executed by a contractile ring of proteins that pinches the daughter cells apart. This mode of cytokinesis is typical of eggs that are cleaved in their entirety, as in *Xenopus* and many other species. Alternatively, cytokinesis may occur by folds growing out from the outer cell membrane. This mode is characteristic of large eggs that contain major amounts of yolk, as do the eggs of birds, reptiles, and fishes. Here, only a small disc of yolk-free cytoplasm is cleaved while the yolky part of the egg remains undivided. *See* CELL DIVISION; CYTOKINESIS; MITOSIS.

Nucleocytoplasmic ratio and midblastula transition. Cleavage differs from later cell divisions by its fast pace, which leaves little or no time for cell growth. Thus, blastomeres become smaller and smaller while the overall size of the embryo remains nearly constant. One advantage of this fast pace is the rapid increase of the nucleocytoplasmic ratio, that is, the number of gene copies (typically two) relative to the volume of cytoplasm. Most cells need nuclear DNA to maintain a healthy turnover of their RNA and protein molecules. Embryos can synthesize new proteins using an endowment of maternal RNA (stored while the egg was growing in the ovary) but only for a limited time. Another advantage of rapid cleavage is the shortening of embryonic development, a period during which organisms are especially vulnerable to parasites and predators.

Typically, cleavage divisions slow down during the blastula stage. The longer cell cycles leave time for blastomeres to transcribe their nuclear DNA into RNA, which then replaces the maternal RNA. Along with making their own RNA, blastomeres may synthesize more protein and begin to grow between cell divisions. As blastomeres use their own genetic information, they may drive their cell cycles at different speeds, causing them to fall out of synchrony. Around the same time, blastomeres may also start moving relative to one another in different ways. In *Xenopus*, all these changes occur simultaneously at the blastoderm stage and are known as the midblastula transition, or MBT. In other species, MBT events may occur at successive stages. It is thought that key MBT events are triggered when the nucleocytoplasmic ratio reaches a threshold level, possibly as some cytoplasmic component is titrated out by nuclear DNA. *See* CELL CYCLE; DEOXYRIBONUCLEIC ACID (DNA); RIBONUCLEIC ACID (RNA).

Role of cleavage in cell determination. Cleavage is the beginning of a developmental process that will lead to the formation of specialized cells, such as epidermal or blood cells. What becomes of a blastomere in the course of normal development is called its fate. However, in most species early blastomeres are not restricted to their fate but are pluripotent, meaning they are capable of forming multiple embryonic parts. For instance, when the blastomeres of a frog embryo are teased apart at the two-cell stage, each may give rise to a complete larva. Conversely, two complete mouse embryos, when stuck together experimentally and reinserted into the uterus of a foster mother, will give rise to one mouse consisting of a mosaic of cells derived from the two constituent embryos.

The stepwise process that restricts pluripotent blastomeres to their fate is known as cell

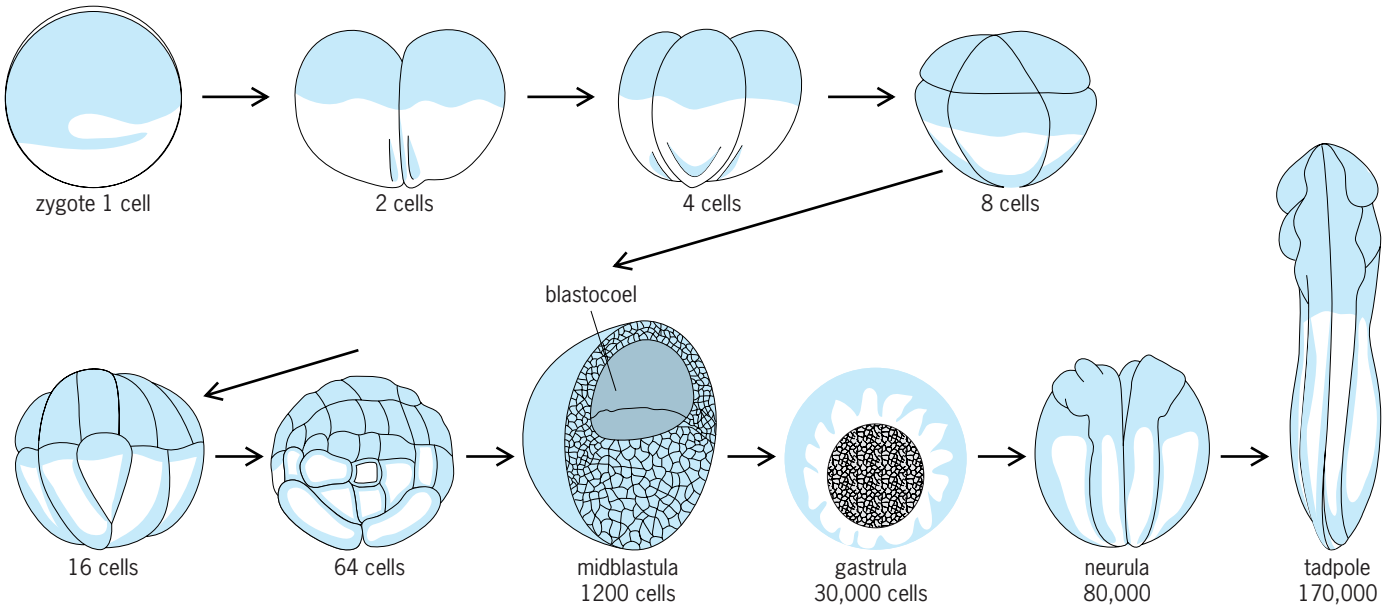

Fig. 1. Development of the African clawed frog, *Xenopus laevis*, from the fertilized egg to the tadpole stage. The midblastula stage is shown in section to reveal the blastocoel. (*Modified from K. Kalthoff, Analysis of Biological Development, 2d ed., McGraw-Hill, 2001*)

determination. In many embryos, such as amphibians, cell determination relies in part on asymmetrically distributed molecules in the egg cytoplasm, which are segregated into different blastomeres as the egg is cleaved. These cytoplasmic determinants elicit different patterns of gene activity in the enclosed cell nuclei. Blastomeres that have inherited different cytoplasmic determinants may later exchange signals that lead to increasingly complex patterns of cells with different determined states. In some species, the process of cell determination occurs early; their embryos cleave in stereotypical patterns that are virtually identical from one embryo to the next. In other species, cleavage patterns are more variable, and blastomere determination relies more on later adjustments by intercellular signals.

Cleavage of mammalian embryos. Cleavage of mammalian embryos is peculiar in several respects. The mammalian egg is unusually small (about 0.1 mm in diameter) and cleaves at a very leisurely pace. Embryonic RNA synthesis starts at the two-cell stage, and cell growth begins early on. All these traits are possible because the embryo develops in the uterus, which provides nutrients and protects the embryo from parasites and predators. The mammalian cleavage pattern produces a ball of cells, which secrete fluid to generate a blastocoel in an off-center position. This leads to the hallmark stage of mammalian development known as the blastocyst (**Fig. 2**). The blastocyst consists, in part, of an outer cell layer known as the trophoblast, which contributes to the formation of the placenta. Positioned excentrically within the trophoblast is the inner cell mass, or ICM, which gives rise to the embryo proper.

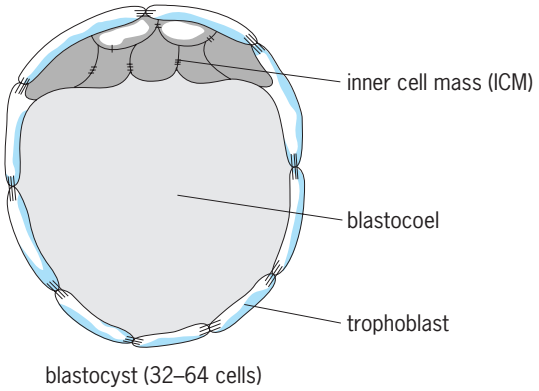

Fig. 2. Mammalian embryo at the blastocyst stage (about 50 cells). (*Reprinted with permission from K. Kalthoff, Analysis of Biological Development, 2d ed., McGraw-Hill, 2001*)

ICM cells are of medical interest because, if cultured under favorable conditions, they may give rise to embryonic stem cells, or ES cells. ES cells divide and grow indefinitely in an undifferentiated state, and they can divide asymmetrically, with one daughter being again an ES cell while its sister cell becomes a committed progenitor cell. The latter, depending on signal molecules in the culture medium, may give rise to all kinds of specialized cells, such as neurons or heart muscle cells. ES cells may, therefore, be used for cell replacement therapy for patients with diseases that leave certain cell lineages dysfunctional, such as Parkinson's disease or heart muscle failure. Cell replacement would be especially useful in conjunction with another technique known as nuclear transfer. Here, the nucleus of a cell from the patient is transferred into an enucleated human egg, which is grown in culture to the blastocyst stage. ICM cells are then harvested to raise ES cells that carry the patient's genetic information, so that their descendants will not be rejected by the patient's immune

system. *See* EMBRYONIC DIFFERENTIATION; STEM CELLS.

Klaus Kalthoff

Bibliography. B. I. Balinsky, *Introduction to Embryology*, 4th ed., Saunders, Philadelphia, 1975; K. Kalthoff, *Analysis of Biological Development*, 2d ed., McGraw-Hill, Boston, 2001.

Cleft lip and cleft palate

Two of the most common congenital anomalies in humans, resulting from incomplete closure of the lip and palate during early embryonic life. During the first trimester of pregnancy, the face and mouth are formed by the fusion of several different parts. The lip is fully fused 6 or 7 weeks after fertilization. The palate, which forms the roof of the mouth and separates the oral cavity from the nasal cavity, is fully fused 10 weeks after fertilization. If fusion fails to occur or breaks down, an opening, or cleft, occurs in the lip, the palate, or both. Although cleft lip and cleft palate are associated with separate embryologic events, that is, are separated by a long time on the embryonic time scale, both often occur in an affected individual. Clefts vary in frequency among racial subgroups, with incidence ranging from 1 per 250 live births to 1 per 750 live births. *See* MOUTH.

Clefts of the lip may occur on one side (unilateral, **Fig. 1**) or both sides (bilateral). When they occur in the exact midline of the lip, they are often associated with severe anomalies of the brain. The cleft usually extends through the lip and nostril floor. The underlying dental arch is also typically involved.

The palate consists of an anterior bony portion (hard palate) and a posterior muscular portion (soft palate). The soft palate alone may be cleft, or both the soft and hard palate may be cleft (**Fig. 2**), but rarely is the hard palate cleft with the soft palate intact. A variant of cleft palate is the submucous cleft palate, in which the muscles of the palate are cleft but the overlying mucous membrane is intact.

Fig. 1. Unilateral, incomplete cleft lip in which the notch does not extend to the nose.

Fig. 2. Cleft palate, resulting in a fissure in the palate connecting the nasal and oral cavities.

Causes. Clefts are the only major abnormality in approximately half of newborns who have them. In the other half, they occur as a part of a pattern of multiple anomalies (a syndrome). It has been determined that most clefts are associated with genetic factors; the genetic patterns range from autosomal dominant inheritance (50% risk of inheritance) in many syndromes to a genetic predisposition for clefting (about 4% risk of inheritance). Clefting is often associated with chromosome abnormalities, intrauterine disturbances (such as tears in the amnion), and teratogenic (malformation-causing) substances, such as alcohol, anticonvulsant drugs, and excessive levels of retinoic acid (vitamin A). *See* HUMAN GENETICS.

Secondary complications. Cleft lip is both an esthetic and a functional problem because of the disruption of the lip musculature. The defect in the floor of the nose causes asymmetry of the nose. The cleft of the dental arch causes orthodontic and occlusal (bite) problems, including congenitally missing teeth. The palatal cleft may result in some minor difficulties with feeding in early life, but a more significant problem is abnormal speech, including excessive nasal resonance (hypernasality) and articulation disorders. Clefting is also associated with chronic middle-ear disease (ear infections) and mild hearing loss. Additional structural or functional disorders can be associated with other features of a syndrome that includes cleft lip or cleft palate. *See* HEARING (HUMAN); SPEECH.

Management. Because of the combination of esthetic, structural, and functional requirements for total repair of clefts, that is, surgical, dental, ear, hearing, and speech, the care of children with clefts has usually been managed by comprehensive teams of specialists, including plastic surgeons, speech pathologists, audiologists, geneticists, otolaryngologists, oral surgeons, orthodontists, and pediatricians.

Cleft lip is typically repaired within the first few months of life, and the palate is most often surgically closed before speech development.

Extensive dental work is normally deferred until some of the permanent teeth have erupted. Middle-ear disease is usually treated aggressively through surgery in order to prevent long-term hearing loss. In general, the outlook for children with these problems is very favorable, and in many cases the problems associated with clefts are completely remediable. Genetic counseling is extremely important. *See* CONGENITAL ANOMALIES. Robert Shprintzen

Bibliography. J. Hobbs (ed.), *Cleft Lip and Cleft Palate*, Cleft Palate Foundation, 1989; D. R. Millard, *Cleft Craft: The Evolution of Its Surgery*, 1980; R. J. Shprintzen et al., Anomalies associated with cleft lip, cleft palate, or both, *Amer. J. Med. Genet.*, 20:585–596, 1985; R. J. Shprintzen and C. B. Croft, Abnormalities of the eustachian tube orifice in individuals with cleft palate, *Int. J. Pediatr. Otorhinolaryngol.*, 3:15–23, 1981; J. E. Trost, Articulatory additions to the classical description of the speech of persons with cleft palate, *Cleft Palate J.*, 18:193–203, 1981; K. Vargervik, Growth characteristics of the premaxilla and orthodontic treatment principles in bilateral cleft lip and palate, *Cleft Palate J.*, 20:289–302, 1983.

Clevis pin

A fastener with a head at one end and a hole at the other used to join a clevis to a rod. A clevis is a yoke with a hole formed or attached at one end of a rod (see **illus**.). When an eye or hole of a second rod is aligned with the hole in the yoke, a clevis pin can be inserted to join the two. A cotter pin can then be inserted in the hole of the clevis pin to hold it in, yet the fastening is readily detachable. This joint is used for rods in tension where some flexibility is required. Paul H. Black

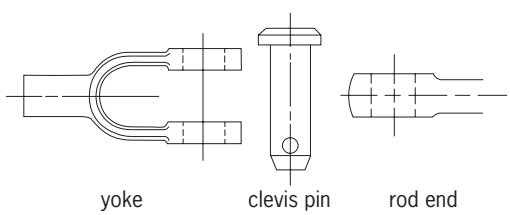

Clevis pin which joins yoke to rod end.

Click chemistry

A simplified chemical reaction method recognized in 2001 by a group led by Nobel Laureate Prof. K. Barry Sharpless at the Scripps Research Institute. "Click" is meant to signify that joining molecular pieces is as easy as clicking together the two pieces of a buckle. The buckle works no matter what is attached to it, as long as its two pieces can reach each other and the components of the buckle can make a connection only with each other. Click reactions are easy to perform, give rise to their intended products in very high yields with little or no by-products, work well under many conditions (usually especially well in water), and are unaffected by the nature of the groups being connected to each other.

The central question posed by Sharpless and colleagues is whether desired functions can be obtained from molecules made using only click chemical reactions. In this deliberate eschewing of more sophisticated synthetic techniques is an implicit challenge to organic chemists that make complex structures by complex methods (can complex structures be made by simple methods?), and an explicit tip of the hat to the synthesis of functional polymers, which cannot occur without the use of reactions that meet the click chemistry standard. *See* ORGANIC SYNTHESIS.

Reactions. Click reactions share the following attributes. (1) Many click components are derived from alkenes and alkynes, and thus ultimately from the cracking of petroleum. Carbon-carbon multiple bonds provide both energy and mechanistic pathways to be made into reactive structures for click connections. (2) Most click reactions involve the formation of carbon-heteroatom (mostly N, O, and S) bonds. This stands in contrast to much of synthetic organic chemistry of recent years, which has emphasized the formation of carbon-carbon bonds. (3) Click reactions are strongly exothermic by virtue of highly energetic reactants or strongly stabilized products. (4) Click reactions are usually fusion processes (leaving no by-products) or condensation processes (producing water as a by-product). (5) Many click reactions are highly tolerant of and often accelerated by the presence of water. *See* ALKENE; ALKYNE.

These features are illustrated by the case of epoxide ring opening by a wide range of nucleophiles (**Fig. 1**). Since the epoxide is a strained three-membered ring, its ring opening is a strongly favored process. Yet that ring opening is required to occur in a particular way; that is, the nucleophile can attack the epoxide carbon atom only along the C-O bond axis, so by-products are avoided and yields are high. In addition, epoxides are relatively unreactive with water, while the hydrogen bonding ability and polar nature of water both serve to support the epoxide ring-opening reaction by other nucleophiles. *See* ELECTROPHILIC AND NUCLEOPHILIC REAGENTS; EPOXIDE; HYDROGEN BOND.

Most click reactions were discovered many years ago and have been widely appreciated, but have not been tapped for their full potential. They include the following classes of reactions, some of which are shown in **Fig. 2**.

Nucleophilic opening of spring-loaded electrophiles. The opening of the three-membered epoxide, aziridine, aziridinium, and episulfonium rings are all simple and versatile reactions. Included in this category are Michael addition reactions to α,β-unsaturated

carbonyl compounds, which are also strongly favored processes.

Mild condensation reactions of carbonyl compounds. The formation of 1,3-dioxolane rings from aldehydes or ketones and of 1,3-diols, hydrazones, and oximes from aldehydes and hydrazines or hydroxylamine ethers, and of heterocycles from α- and β-carbonyl aldehydes, ketones, and esters are highly reliable and widely used reactions. *See* ALDEHYDE; CARBONYL; HETEROCYCLIC COMPOUNDS; KETONE; OXIME.

Cycloaddition reactions. Distinguished by the relatively nonpolar nature of their reactive groups, these fusion reactions comprise a wide range of processes, including the Diels-Alder reaction. Most useful are 1,3-dipolar cycloaddition reactions, with the cream of the crop being the reaction of organic azides with alkynes. *See* AZIDE; DIELS-ALDER REACTION.

Azide-alkyne cycloaddition. The goals and philosophy of click chemistry have been most extensively explored with the uncatalyzed (**Fig. 3***a*) and copper-catalyzed (Fig. 3*b*) versions of the cycloaddition reaction between azides and alkynes to form triazines. The former was first reported by A. Michael in 1893 and was later studied and popularized by the German chemist Rolf Huisgen in 1960s through the 1980s. The latter reaction was discovered independently in 2002 by the research groups of M. Meldal in Holland and K. B. Sharpless in the United States, and requires the use of a terminal alkyne. The abbreviation AAC (azide-alkyne cycloaddition) is used for the uncatalyzed process and CuAAC for the copper-mediated version.

These processes are uniquely useful because of the properties of the reactive substituents. Both azides and alkynes are high in chemical potential energy, and their fusion to make triazoles is exothermic by more than 45 kcal/mol. However, the rate of this reaction is quite slow, normally requiring prolonged heating for unactivated (not strongly electron-deficient or strained) alkynes. The azide and alkyne groups are stable in the presence of the nucleophiles, electrophiles, and solvents common to standard reaction conditions, with the azide being the only 1,3-dipolar reagent to have this quality. Most importantly, azides and alkynes are almost completely unreactive toward biological molecules. They are small and incapable of significant hydrogen bonding, and thus are unlikely to significantly change the properties of structures to which they are attached. Both can be easily introduced into organic compounds.

Click chemistry in situ. Because of their special reactivity—inert toward everything else and only sluggishly reactive with each other—azides and alkynes have been used to assemble molecules that bind very tightly to enzymes by using the enzyme itself as the reaction vessel (**Fig. 4**). The technique, known as click chemistry in situ, involves the tagging of molecules that bind to adjacent locations on an enzyme with azide and alkyne groups. If the derivatized molecules can interact simultaneously with the target in the proper orientation to hold the azide and alkyne units in proximity, a triazole may form to link the enzyme-binding components to each other. Since attachments to any target by two arms is better than either arm alone, the result is invariably a molecule of much higher binding affinity. Prior knowledge of the structure of the target is not required, nor is an assay for enzyme activity necessary. Since no reaction occurs between azides and alkynes at the concentrations used in these experiments without templating by the target enzyme, the formation of a new product is easy to detect by mass spectrometry

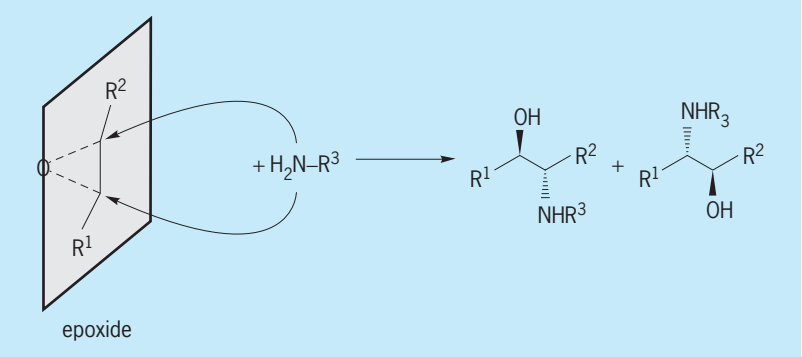

Fig. 1. Ring opening of epoxides with amines, a prototypical click reaction.

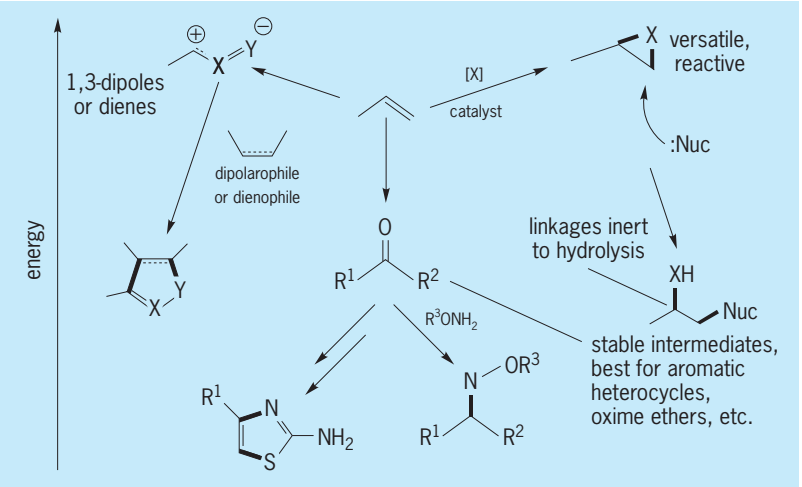

Fig. 2. Click chemistry employs highly energetic but selectively reactive groups.

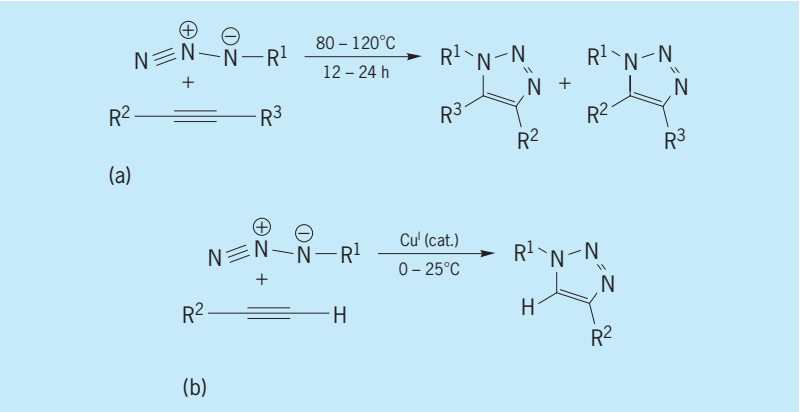

Fig. 3. Azide-alkyne cycloaddition (AAC) reaction to form triazoles: (*a*) Uncatalyzed and (*b*) copper-catalyzed version.

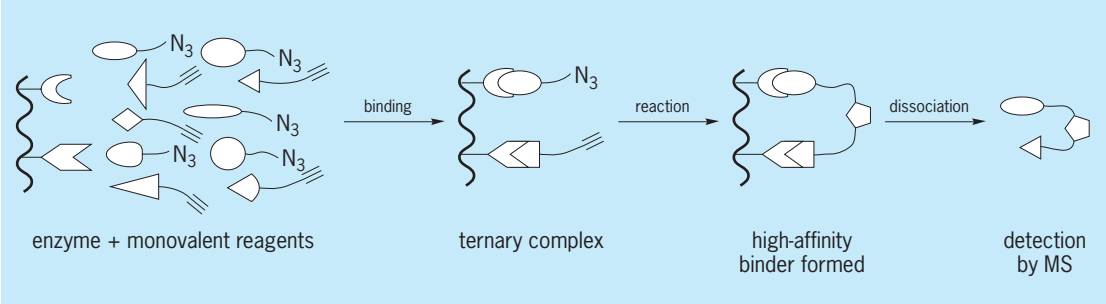

Fig. 4. Click chemistry in-situ technique.

and is proof of the creation of an excellent enzyme inhibitor. *See* ENZYME.

Click chemistry in situ has been used to create novel high-affinity binders to the important neurotransmitter enzyme acetylcholinesterase, the metabolic enzyme carbonic anhydrase, and HIV protease. It has become apparent in these and other studies that the triazole linkage has advantageous properties for drug discovery. Having a strong dipole moment, the capacity for significant hydrogen bonding, and the ability to engage in π-stacking interactions, triazoles can engage in productive interactions with proteins. The creation of this unit from two "invisible" pieces in a pocket of an enzyme is an event that influences the selectivity of bond formation during in-situ drug discovery efforts. The click chemistry in-situ technique is used in many laboratories and pharmaceutical companies around the world as a complement to traditional methods of synthesis and screening. *See* DIPOLE MOMENT.

While not yet used directly in living cells because of the cytotoxicity and attendant bioregulation of copper, the copper-catalyzed reaction has achieved extraordinarily wide use in organic chemistry and materials science. Applications include the synthesis of biologically active compounds, the preparation of conjugates to proteins and polynucleotides, the synthesis of dyes, the elaboration of known polymers and the synthesis of new ones, the creation of responsive materials, and the covalent attachment of desired structures to surfaces.

Conclusion. Click chemistry expands the scope of structures that can be made by expert chemists and nonchemists alike. The rationale is simply that the more tolerant is the connection reaction between chemical pieces, the more diverse the pieces that can be brought to bear on any problem. Chemists possess neither the ability to control reactions as well as living cells do, nor the ability to manufacture large structures such as proteins for every purpose. Therefore, while nature can inspire us, making such complex molecules comes at a fearful cost of time and expense. If indeed it is chemical function that is the goal, keeping it simple is a useful rule.

M. G. Finn

Bibliography. R. S. Bohacek, C. McMartin, and W. C. Guida, The art and practice of structure-based drug design: A molecular modeling perspective, *Med. Res. Rev.*, 16(1):3–50, 1996; Y. Bourne et al., Freeze-frame inhibitor captures acetylcholinesterase in a unique conformation, *Proc. Nat. Acad. Sci.*, 101(6):1449–1454, 2004; F. Himo et al., Copper(I)-catalyzed synthesis of azoles: DFT study predicts unprecedented reactivity and intermediates, *J. Amer. Chem. Soc.*, 127(1):210–216, 2005; R. Huisgen, in A. Padwa (ed.), *1,3-Dipolar Cycloaddition Chemistry*; vol. 1, pp. 1–176, Wiley, New York, 1984; R. Huisgen, Kinetics and reaction mechanisms: Selected examples from the experience of forty years, *Pure Appl. Chem.*, 61(4):613–628, 1989; H. C. Kolb and K. B. Sharpless, The growing impact of click chemistry on drug discovery, *Drug Disc. Today*, 8(24):1128–1137, 2003; W. G. Lewis et al., Click chemistry in situ: Acetylcholinesterase as a reaction vessel for the selective assembly of a femtomolar inhibitor from an array of building blocks, *Angew. Chem. Int. Ed.*, 41(6):1053–1057, 2002; W. G. Lewis et al., Discovery and characterization of catalysts for azide-alkyne cycloaddition by fluorescence quenching, *J. Amer. Chem. Soc.*, 126(30):9152–9153, 2004; R. Manetsch et al., In situ click chemistry: Enzyme inhibitors made to their own specifications, *J. Amer. Chem. Soc.*, 126(40):12809–12818, 2004; V. P. Mocharla et al., In situ click chemistry: Enzyme-generated inhibitors of carbonic anhydrase II, *Angew. Chem. Int. Ed.*, 44(1):116–120, 2005; V. O. Rodionov, V. V. Fokin, and M. G. Finn, Mechanism of the ligand-free Cu(I)-catalyzed azide-alkyne cycloaddition reaction, *Angew. Chem. Int. Ed.*, 44(15):2210–2215, 2005; V. V. Rostovtsev et al., A stepwise Huisgen cycloaddition process: Copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes, *Angew. Chem. Int. Ed.*, 41(14):2596–2599, 2002; C. W. Tornøe, C. Christensen, and M. J. Meldal, Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides, *J. Org. Chem.*, 67(9)3057–3062, 2002.

Client-server system

A computing system that is composed of two logical parts: a server, which provides services, and a client, which requests them. The two parts can run on separate machines on a network, allowing users to access powerful server resources from their personal computers (**Fig. 1**). *See* LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

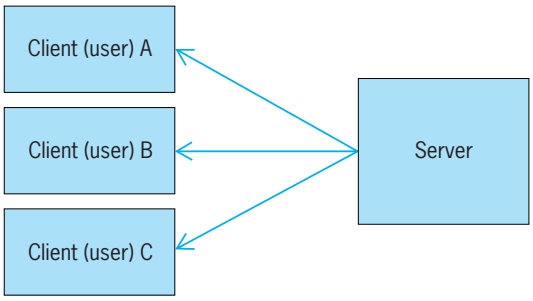

Fig. 1. Client-server system. Clients request services of the server independently but use the same interface.

Client-server systems are not limited to traditional computers. An example is an automated teller machine (ATM) network. Customers typically use ATMs as clients to interface to a server that manages all of the accounts for a bank. This server may in turn work with servers of other banks (such as when withdrawing money at a bank at which the user does not have an account). The ATMs provide a user interface and the servers provide services, such as checking on account balances and transferring money between accounts.

To provide access to servers not running on the same machine as the client, middleware is usually used. Middleware serves as the networking between the components of a client-server system; it must be run on both the client and the server. It provides everything required to get a request from a client to a server and to get the server's response back to the client. Middleware often facilitates communication between different types of computer systems. This communication provides cross-platform client-server computing and allows many types of clients to access the same data.

The server portion almost always holds the data, and the client is nearly always responsible for the user interface. The application logic, which determines how the data should be acted on, can be distributed between the client and the server (**Fig. 2**). The part of a system with a disproportionately large amount of application logic is termed "fat"; a "thin" portion of a system is a part with less responsibility delegated to it. Fat server systems, such as groupware systems and web servers, delegate more responsibility for the application logic to the server, whereas fat client systems, such as most database systems, place more responsibility on the client. *See* HUMAN-COMPUTER INTERACTION.

Multitiered systems. The canonical client-server model assumes two participants in the system. This is called a two-tiered system; the application logic must be in the client or the server, or shared between the two. It is also possible to have the application logic reside in a third layer separate from the user interface and the data, turning the system into a three-tier system. Complete separation is rare in actual systems; usually the bulk of the application logic is in the middle tier, but select portions of it are the responsibility of the client or the server.

The three-tier model is more flexible than the two-tier model because the separation of the application logic from the client and the server gives application logic processes a new level of autonomy. The processes become more robust since they can operate independently of the clients and servers. Furthermore, decoupling the application logic from the data allows data from multiple sources to be used in a single transaction without a breakdown in the client-server model. This advancement in client-server architecture is largely responsible for the notion of distributed data. *See* DISTRIBUTED SYSTEMS (COMPUTERS).

Standard web applications are the most common examples of three-tier systems. The first tier is the user interface, provided via interpretation of Hyper Text Markup Language (HTML) by a web browser. The embedded components being displayed by the browser reside in the middle tier, and provide the application logic pertinent to the system. The final tier is the data from a web server. Quite often this is a database-style system, but it could be a data-warehousing or groupware system. *See* WORLD WIDE WEB.

Software components. The desire to distribute servers across a network and have clients find them on demand has led to the development of component technologies. Components are the smallest self-managing, independent, and useful parts of a system that can work with other components to provide useful services. Each component provides an interface, which is a contract to do certain things asked of it. Therefore, systems can be quickly assembled by reusing existing components. For example, a program that needs to sort a list may ask a sorting component to do this so the programmer writing the new application need not write the sorting code. Components that provide the same interfaces can be interchanged for one another; for example, a user could purchase and install a faster sorting component, and programs using the old sorting component would automatically use the improved version.

Middleware packages. The use of components requires additions to the traditional middleware package that can coordinate components and help components locate one another. The two primary middleware packages providing these services are the Common Object Request Broker Architecture (CORBA) and the Distributed Component Object Model (DCOM).

CORBA. An industry-wide consortium created and advocates the standard known as CORBA. CORBA

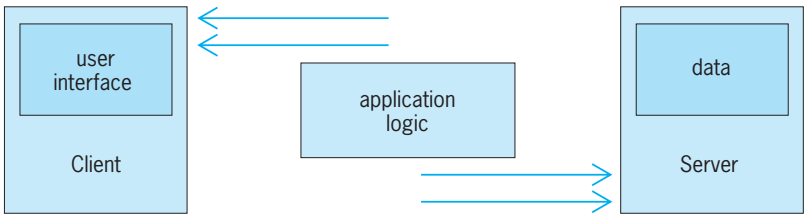

Fig. 2. Distribution of elements of a client-server system. Since the distribution of the user interface and data is fixed, the distribution of the application logic distinguishes fat client systems from fat server systems.

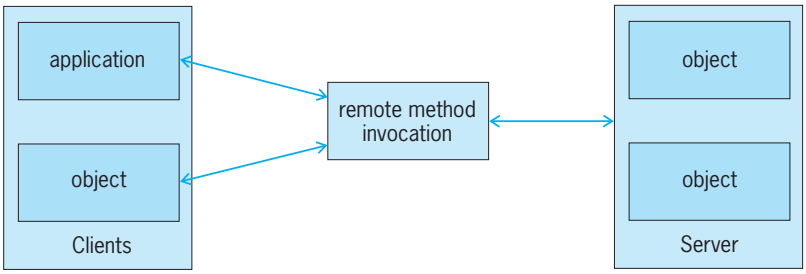

Fig. 3. Client-server computing using distributed objects. Communication between components (denoted by arrows) is facilitated by object request brokers (ORBs), which have been omitted for clarity. The remote method invocation system allows an object on one host to cause a method of an object that resides on another host to be executed on that remote host.

objects can exist anywhere on a network; their location is completely transparent (**Fig. 3**). Details such as an object's language or its operating system are also hidden from clients. Thus, the implementation of an object is of no concern to an invoking object; the interface is the only entity that a client needs to consider when selecting a serving object (Fig. 2).

The most important part of a CORBA system is the object request broker (ORB). CORBA uses the object request broker to establish a client-server relationship between components. The object request broker intercepts service requests from the client objects and routes them to an appropriate server. CORBA uses an interface file written in an interface definition language (IDL) to specify the services that an object can provide, but programmers write code in their language of choice.

When clients request services, the requests are routed to the remote server using a local "stub" object and the object request broker. The stub relieves the application programmer from the burden of routing the message.

CORBA specifies two means by which an object can locate another object in a system. The first is the Naming Service. This service is analogous to the white pages in a phone book: An object looks up another object by the name under which the object registered itself with the object request broker when it was initialized. The second service is the Trading Service, which is like the yellow pages: Objects can ask the Trading Service which objects have registered with certain service characteristics. The trading repository then returns references to salient objects and gives the client information regarding the properties of the services. *See* OBJECT-ORIENTED PROGRAMMING.

DCOM. Like CORBA, DCOM separates interface from functionalty by using an interface definition language. The client-server contract is fulfilled by using object interfaces that make the implementation language and the location of objects transparent. When a DCOM client needs a server, it issues a request for an object supporting the required interface instantiated so that the interface becomes available.

Using interface pointers to mask stublike objects, DCOM supports transparent communication between clients and remote server objects. DCOM uses a model quite similar to the mechanism employed by CORBA to provide location-independent access to objects. Since DCOM provides access to interfaces only (and not objects), there is no means by which to locate specific objects.

Java. Java is a programming language that allows programs to run on a variety of computing platforms without being modified; such a system is referred to as a mobile code system. Mobile code is distinguished from traditional application code in that it is dynamically executed by an application that provides the run-time environment. For code to be considered mobile, it must be portable, meaning the code can be run on a variety of platforms. Java uses bytecodes to provide portability and security. The bytecode system calls for programs to be "compiled" to the point where they can be run on a Java virtual machine, which translates the bytecodes into actual machine instructions in real time. Java virtual machines must adhere to a standard; this ensures that a Java program executes in the same manner on any machine. To provide security, the Java Verifier checks all code scheduled to be executed for malicious code.

Java's promise of "write once, run anywhere" makes it an appealing language in which to author pieces of a client-server system. Java also provides core classes that are a part of standard Java implementations. These core classes provide basic services that many programs require, and can decrease application development times by reducing the amount of code a programmer needs to write and debug.

When coupled with distributed object technology, Java forms a strong basis for the development of a robust system that supports client-server computing. A platform for universal network computing can be created by using Java as a programming language and mobile code system and CORBA or DCOM as an integration technology. Using DCOM or CORBA allows Java components to be split into client- and server-side components in many cases, providing the system with enhanced flexibility in distributing web-based applications. *See* DIGITAL COMPUTER; PROGRAMMING LANGUAGES. Scott M. Lewandowski

Bibliography. K. Arnold and J. Gosling, *The Java Programming Language*, Addison-Wesley, Reading, MA, 1996; D. Chappell, *Understanding ActiveX and OLE*, Microsoft Press, Redmond, WA, 1996; R. Orfali, D. Harkey, and J. Edwards, *The Essential Client/Server Survival Guide*, Wiley, New York, 1996; R. Orfali, D. Harkey, and J. Edwards, *The Essential Distributed Objects Survival Guide*, Wiley, New York, 1996.

Climate history

The long-term records of precipitation, temperature, wind, and all other aspects of the Earth's climate. The climate, like the Earth itself, has a history extending over several billion years. Climatic changes have occurred at time scales ranging from hundreds of millions of years to centuries and decades. Processes in the atmosphere, oceans, cryosphere (snow cover,

sea ice, continental ice sheets), biosphere, and lithosphere (such as plate tectonics and volcanic activity) and certain extraterrestrial factors (such as the Sun) have caused these changes of climate.

The present climate can be described as an ice age climate, since large land surfaces are covered with ice sheets (for example, Antarctica and Greenland). The origins of the present ice age may be traced, at least in part, to the movement of the continental plates. With the gradual movement of Antarctica toward its present isolated polar position, ice sheets began to develop there about 30 million years ago. For the past several million years, the Antarctic ice sheet reached approximately its present size, and ice sheets appeared on the lands bordering the northern Atlantic Ocean. During the past million years of the current ice age, about 10 glacial-interglacial cycles have been documented. Changes in the Earth's orbital parameters, eccentricity, obliquity, and longitude of perihelion are thought to have initiated, or paced, these cycles through the associated small changes in the seasonal and latitudinal distribution of solar radiation. The most recent glacial period ended between about 15,000 and 6000 years ago with the rapid melting of the North American and European ice sheets and an associated rise in sea level, and the atmospheric concentration of carbon dioxide.

The present interglacial climate reached its maximum warmth about 6000 years ago when, compared to the present, the high northern latitudes were slightly warmer, the northern continental interiors were drier, and the monsoon lands of Africa and southern Asia were wetter. These differences appear also to have been caused, at least in part, by differences in Earth's orbital parameters, compared to the present.

Changes of climate spanning centuries to decades occur within the more gradual changes just described. For example, the climate of northern Europe was slightly warmer about 1000 years ago than at present, but colder than at present 100 to 300 hundred years ago, when mountain glaciers advanced beyond present-day limits in Scandinavia and the Alps. The causes of these shorter-term changes of climate are not well understood, although volcanic activity, solar variability, and the internal dynamics of the atmosphere and ocean may all be involved. There is concern that human activities, such as massive deforestation and the burning of fossil fuels, may become causes of climatic change.

The climates of the distant geologic past were strongly influenced by the size and location of continents and by large changes in the composition of the atmosphere. For example, around 250 million years ago the continents were assembled into one supercontinent, Pangaea, producing significantly different climatic patterns than are seen today with widely distributed continents. In addition, based upon models of stellar evolution, it is hypothesized that the Sun's radiation has gradually increased by 10–20% over the past several billion years and, if so, this has contributed to a significant warming of the Earth. *See* CONTINENTS, EVOLUTION OF; GLACIAL EPOCH; PALEOCLIMATOLOGY.

Evidence. Instrumental records of climatic variables such as temperature and precipitation exist for the past 100 years in many locations and for as long as 200 years in a few locations. These records provide evidence of year-to-year and decade-to-decade variability, but they are completely inadequate for the study of century-to-century and longer-term variability. Even for the study of short-term climatic fluctuations, instrumental records are incomplete, because most observations are made from the continents (covering only 29% of the Earth's surface area). Aerological observations, which permit the study of atmospheric mass, momentum and energy budgets, and the statistical structure of the large-scale circulation, are available only since about the mid-1960s. Again there is a bias toward observations over the continents. It is only with the advent of satellites that global monitoring of the components of the Earth's radiation budget (clouds; planetary albedo, from which the net incoming solar radiation can be estimated; and the outgoing terrestrial radiation) became possible. *See* METEOROLOGICAL SATELLITES.

There remain important gaps in the ability of climatologists to describe the present state of the climate. For example, precipitation estimates, especially over the oceans, are poor. Oceanic circulation, heat transport, and heat storage are only crudely estimated. Thus, although climatic changes are discernible from instrumental records, the task of defining the scope of these changes and diagnosing potential causes is difficult.

In spite of the inadequacy of the instrumental records for assessing global climate, considerable evidence of regional climatic variations exists. For example, there is evidence of climatic warming in the polar regions of the Northern Hemisphere during the twentieth century. *See* GREENHOUSE EFFECT.

Evidence of climatic changes prior to instrumental records comes from a wide variety of sources. Tree rings, banded corals, and pollen and trace minerals retrieved from laminated lake sediments and ice sheets yield environmental records for past centuries and millennia. Advanced drilling techniques have made it possible to obtain long cores from ocean sediments that provide geologic records of climatic conditions going back hundreds of millions of years. *See* DENDROCHRONOLOGY; PALYNOLOGY.

Causes. Many extraterrestrial and terrestrial processes have been hypothesized to be possible causes of climatic fluctuations. These include solar irradiance, variations in orbital parameters, motions of the lithosphere, volcanic activity, internal variations of the climate system, and human activities. It is likely that all of the natural processes have played a role in past climatic changes. Also, the climatic response to some particular causal process may depend on the initial climatic state, which in turn depends upon previous climatic states because of the long time constants of lithosphere, oceans, and cryosphere. True equilibrium climates

may not exist, and the climate system may be in a continual state of adjustment.

Solar irradiance. Variations in solar irradiance would almost certainly have an influence on climate. In addition to the possible long-term evolution of the Sun, the changing appearance of the Sun, the so-called sunspots, hint at shorter-term variability on the scale of years to centuries. Satellite observations have detected decadal-scale fluctuations in solar radiation on the order of 0.1%. Radiance variability in limited portions of the solar spectrum has also been observed but has not been linked clearly to climate variability. *See* SOLAR RADIATION.

Orbital parameters. Variations of the Earth's orbital parameters (eccentricity of orbit about the Sun, longitude of perihelion, and inclination or tilt of the rotational axis) lead to small but significant variations in incoming solar radiation with regard to seasonal partitioning and latitudinal distribution. These variations occur at times scales of 20,000 to 100,000 years and have been linked to glacial-interglacial cycles and sedimentation cycles in geologic records. *See* EARTH ROTATION AND ORBITAL MOTION.

Lithosphere motions. Sea-floor spreading and continental drift, continental uplift, and mountain building operate over long time scales and are almost certainly important factors in long-term climate. *See* CONTINENTAL DRIFT; MARINE GEOLOGY; OROGENY; PLATE TECTONICS.

Volcanic activity. Volcanic activity produces gaseous and particulate emissions that lead to the formation of persistent stratospheric aerosol layers. It may be a factor in climatic variations at all time scales. *See* VOLCANO.

Internal variability of climate system. Components of the climate system (atmosphere, ocean, cryosphere, biosphere, land surface) are interrelated through a variety of feedback processes operating over a wide range of time scales. These processes could, in principle, produce fluctuations of sufficient magnitude and variability to explain any observed climatic change. For example, atmosphere-ocean interactions may operate over time scales up to centuries, the basic period of overturning of the ocean thermohaline circulation; and atmosphere-ocean-cryosphere interactions may operate over time scales up to tens of millennia, the time required to build a major ice sheet. Several hypotheses have been proposed to explain glacial-interglacial fluctuations as complex internal feedbacks among atmosphere, ocean, and cryosphere. (Periodic buildup and surges of the Antarctic ice sheet and periodic fluctuation in sea ice extent and deep-ocean circulation provide examples.) Atmosphere-ocean interaction is being studied intensively as a possible cause of short-term climatic variations. It has been observed that anomalous ocean temperature patterns (both equatorial and midlatitude) are often associated with anomalous atmospheric circulation patterns. Although atmospheric circulation plays a dominant role in establishing a particular ocean temperature pattern (by means of changes in wind-driven currents, upwelling, radiation exchange, evaporation, and so on), the anomalous ocean temperature distribution may then persist for months, seasons, or longer intervals of time because of the large heat capacity of the oceans. These anomalous oceanic heat sources and sinks may, in turn, produce anomalous atmospheric motions. The Earth's biosphere may also be actively involved in climatic changes. For example, the atmospheric concentration of carbon dioxide is linked to the carbon reservoirs of the continents and oceans via photosynthesis-respiration cycles, and gas exchange at the air-ocean interface. *See* ATMOSPHERIC GENERAL CIRCULATION; BIOSPHERE; OCEAN CIRCULATION.

Human activities. Forest clearing and other large-scale changes in land use, changes in aerosol loading, and the changing concentration of carbon dioxide in the atmosphere are examples of possible mechanisms through which human activities may influence the large-scale climate. Because of the large observational uncertainties in defining the state of the climate, it has not been possible to establish the relative importance of human activities (as compared to natural processes) in recent climatic fluctuations. There is, however, considerable concern that future human activities may lead to large climatic variations (for example, continued increases in atmospheric carbon dioxide concentration due to burning of fossil fuels) by the early twenty-first century. *See* AIR POLLUTION; CLIMATE MODIFICATION.

Modeling. Because of the complexity of the real climate system, simplified numerical models of climate are being used to study particular processes and interactions. Some models treat only the global-average conditions, whereas others, particularly the dynamical atmosphere and ocean models, simulate detailed patterns of climate. These models will undoubtedly be of great importance in attempts to understand climatic processes and to assess the possible effects of human activities on climate. *See* CLIMATE MODELING; CLIMATIC PREDICTION; CLIMATOLOGY.

John E. Kutzbach

Bibliography. T. J. Crowley and G. R. North, *Paleoclimatology*, 1991; J. T. Houghton, G. J. Jenkins, and J. J. Ephraums (eds.), *Climate Change: The IPCC Scientific Assessment*, 1990; J. Imbrie and K. P. Imbrie, *Ice Ages: Solving the Mystery*, 1986; H. H. Lamb, *Climate History and the Modern World*, 1982; H. E. Wright, Jr., et al. (eds.), *Global Climates since the Last Glacial Maximum*, 1993.

Climate modeling

Construction of a mathematical model of the climate system of the Earth capable of simulating its behavior under present and altered conditions. The Earth's climate is continually changing over time scales ranging from hundreds of years to a few years. Since the climate is determined by the laws of classical physics, it should be possible in principle to construct such a model. The advent of a worldwide weather-observing system capable of gathering data for validation and the development and widespread

routine use of digital computers have made this undertaking possible, starting in the mid-1970s.

The first attempts at modeling the planetary climate showed that the Earth's average temperature is determined mainly by the balance of radiant energy absorbed from sunlight and that emitted by the Earth system. About 30% of the incoming radiation is reflected directly to space, and 72% of the remainder is absorbed at the surface (**Fig. 1**). The incoming solar radiation is divided among reflection, absorption by the atmospheric constituents, and absorption by the surface of the planet. The outgoing infrared radiation comes from the surface, atmospheric gases, and clouds. In addition, the atmosphere radiates down to the surface, and the surface gives energy to the atmosphere in the forms of latent and sensible heat. The radiation is absorbed unevenly over the Earth, which sets up thermal contrasts that in turn induce convective circulations in the atmosphere and oceans. Climate models attempt to calculate from mathematical algorithms the effects of these contrasts and the resulting motions in order to understand better and perhaps predict future climates in some probabilistic sense. *See* SOLAR RADIATION; TERRESTRIAL RADIATION.

Climate models differ in complexity, depending upon the application. The simplest models are intended for describing only the surface thermal field at a fairly coarse resolution. These mainly thermodynamical formulations are successful at describing the seasonal cycle of the present climate, and have been used in some simulations of past climates, for example, for different continental arrangements that occurred millions of years ago. At the other end of the spectrum are the most complex climate models, which are extensions of the models used in weather forecasts. These models aim at simulating seasonal and even monthly averages just shortly into the future, based upon conditions such as the temperatures of the tropical-sea surfaces. Intermediate to these extremes are models that attempt to model climate on a decadal basis, and these are used mainly in studies of the impact of hypothesized anthropogenically induced climate change. *See* CLIMATE MODIFICATION.

Climate models are used in a large variety of applications that aid in the understanding of Earth history. Many simple climate model simulations have been used to sort out the mechanisms responsible for climate change in the past. For example, although not yet fully understood, the astronomical theory of the ice ages states that the waxing and waning of the great continental ice sheets has been forced by the periodic changes in the Earth's ellipitical orbit parameters in the past. Similarly, the onset of glaciation in Antarctica and Greenland has been studied by such means. *See* CLIMATE HISTORY.

A problem that has received considerable attention is the greenhouse effect. Models are studied to achieve better understanding of how the increase of atmospheric carbon dioxide and other trace gases from anthropogenic sources is likely to change the climate in the coming decades. The models are compared to past climates ranging from the ice ages to the records of the last hundred years, for which an instrumental record exists. *See* GREENHOUSE EFFECT.

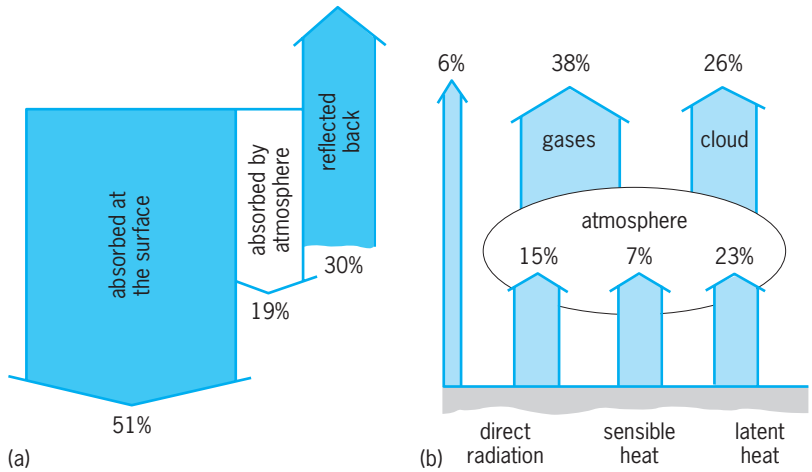

Fig. 1. Schematic diagram of the budget of incoming and outgoing radiation affecting the Earth's climate. (*a*) Solar radiation. (*b*) Terrestrial radiation.

Anatomy of models. Since the main interest of climate modelers is in computing the thermal field over the Earth, a primary objective is to represent the conservation of energy at each location in the system. One requirement is accurate formulations of the absorption and reflection of solar radiation as it passes through the atmosphere and strikes surfaces. Another is the radiation emitted from each mass element in the Earth–atmosphere system. Conversions of heat from latent to sensible must be taken into account as water changes its phase in the system. Clouds must be included since they participate in the radiation transfer and in the changes of water phase. Similarly, snow and ice cover enter both energy disposal accounts. The thermodynamic expression of the conservation of energy is not complete until allowance is made for the flow of matter of a different temperature into a given region. To include this effect, a model of the circulation must be considered, and this in turn is governed by the same thermal contrasts given by the thermodynamic equation.

The circulation of atmospheric and oceanic material is governed by Newton's second law: local acceleration of a fluid element is proportional to total forces on it. In fluid mechanics this is known as the Navier-Stokes equation. It is a nonlinear partial differential equation that is exceedingly complex. Climate models that include a detailed attempt to solve the fluid dynamics equations must be approached by computer simulation. Even here the solutions are known to differ in detail from nature after only a few weeks at most. However, it is thought that statistics, for example, long-term means for the climate of the numerical model and those of nature, should agree; and this has been borne out in numerous tests. *See* NAVIER-STOKES EQUATIONS.

In order to simulate the climate system, the problem must be cast onto a three-dimensional grid in the global ocean–atmosphere volume. The intermediate-sized models typically being used for decadal

simulations have a horizontal resolution of about 300 to 600 mi (500 to 1000 km) and vertical resolutions of about 0.6 to 1.8 mi (1 to 3 km). This implies about 25,000 grid points, at each of which the model is keeping track of about six meteorological variables. The system is solved by numerically advancing in time at each grid point, updating at each time step (typically about intervals of 1 h of so-called model time). Simulation of 15 years of model time, which is typical of models that include only simple formulations of the oceanic interaction, may take tens of hours on the fastest computers. About a half dozen models of this type are being investigated around the world. The models as a group simulate the present seasonal cycle of different geographical regions with remarkable fidelity, considering the short time that has passed since the inception of this field. The models are best at simulating the thermal surface field and weakest in modeling such secondary features as precipitation. *See* SIMULATION.

Simulation models which include ocean circulation are important in understanding the adjustment processes in going from one climate state to another. Inclusion of the ocean dynamics adds significantly to the computational burden, since a finer grid is required owing to the importance of small-scale motions. The irregularity of the bottom topography also complicates the problem. Gerald R. North

Feedback mechanisms and sensitivity. Attempts at modeling climate have demonstrated the extreme complexity and subtlety of the problem. This is due largely to the many feedbacks in the system. One of the simplest and yet most important feedbacks is that due to water vapor. If the Earth is perturbed by an increase in the solar radiation, for example, the first-order response of the system is to increase its temperature. But an increase in air temperature leads to more water vapor evaporating into the air; this in turn leads to increased absorption of space-bound long-wave radiation from the ground (greenhouse effect), which leads to an increased equilibrium temperature. This effect, known as a positive-feedback mechanism, is illustrated in **Fig. 2**. Water vapor feedback is not the only amplifier in the system. Another important one is snowcover: a cooler planet leads to more snow and hence more solar radiation reflected to space, since snow is more reflecting of sunlight than soil or vegetation. Other, more subtle mechanisms that are not yet well understood include those involving clouds, oceans, and the biosphere.

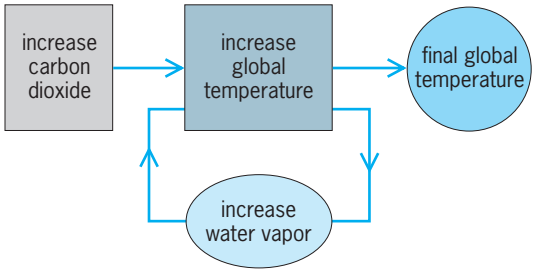

Fig. 2. Water vapor feedback in amplifying the sensitivity of the Earth's climate to increases in atmospheric carbon dioxide concentration. Increases in the Earth's temperature cause an increase in water vapor, causing a warming of the surface because of the greenhouse effect. The net effect is to amplify the response to increases of carbon dioxide.

While water vapor and snowcover feedback are fairly straightforward to model, the less understood feedbacks differ in their implementations from one climate model to another. These differences as well as the details of their different numerical formulations have led to slight differences in the sensitivity of the various models to such standard experimental perturbations as doubling carbon dioxide in the atmosphere. All models agree that the planetary average temperature should increase if carbon dioxide concentrations are doubled. However, the predicted response in planetary temperatures ranges 4.5–9°F (2.5–5.0°C). Regional predictions of temperature or precipitation are not reliable enough for detailed response policy formulation. Many of the discrepancies are expected to decrease as model resolution increases (more grid points), since it is easier to include such complicated phenomena as clouds in finer-scale formulations and coupling with dynamic models of the ocean. Similarly, it is anticipated that some observational data (such as rainfall over the oceans) that are needed for validation of the models will soon be available from satellite sensors. *See* METEOROLOGICAL SATELLITES.

Atmosphere–ocean–land models. Simulating the naturally varying present-day climate and its future evolution can be accomplished only with coupled atmosphere–ocean–land models.

Atmosphere models. Numerical models of the atmosphere are used to reveal its underlying dynamics and to predict the weather. The climatic application of models requires a knowledge of the change in lower-boundary conditions as the model evolves in time: over the oceans, sea surface temperature; and over land, more complex descriptors of the land surface.

Both the ocean and land interact with the atmosphere on multiple time scales. Ocean interactions depend on the thermal capacities of water columns that undergo rapid vertical exchanges, usually about 165–3300 ft (50–1000 m) of water. An unchanging field of sea surface temperature may be used as a model boundary to predict weather a few days in advance (at most) because the huge thermal capacities of the oceans cause hardly any change in the bottom boundary. Dry land, however, interacts with the atmosphere with effective heat capacities of a fraction of a meter of water, so that large thermal variations in interaction with the atmospheric boundary occur over time scales as short as the diurnal cycle. Longer time scales occur for moist soils: for example, soil moisture stores may require several weeks to be depleted. Also, the atmosphere above the boundary layer may require several weeks to be influenced by the underlying boundaries. For simulations or predictions beyond a week, the correct time evolution of the bottom boundary conditions for sea surface temperature and land soil moisture must be calculated.

Coupled models. To consistently simulate both atmospheric and boundary quantities, it is necessary to couple the media and allow them to mutually and consistently evolve. Even to predict evolution of the atmosphere alone, climate models must be coupled atmosphere–ocean–land models.

Numerical models of the ocean are used to understand its underlying dynamics and to predict the sea surface temperature. For these purposes, the fluxes of heat and momentum from the atmosphere must be specified. On sufficiently long time scales, the atmosphere requires sea surface temperatures from the ocean, and the ocean requires momentum and heat fluxes from the atmosphere. Numerical models of the land are used to study basic hydrological processes and to predict temperature, stored moisture, and runoff. Fluxes of heat and water from the atmosphere must be specified because the atmosphere requires land properties and the land requires atmospheric fluxes. Many atmospheric processes must be adequately simulated in order to correctly supply the fluxes of heat, momentum, and moisture to the ocean and to the land surfaces, for example, those processes involving rainfall, cloud optical properties, surface winds, radiation, and atmospheric temperature. These variables, in turn, are affected by the properties of the underlying surface.

Sensitivity studies. Significant properties of the ocean and land components of coupled climate models can be tested and demonstrated by the use of sensitivity studies. These studies assume simple changes in model parameters and assess the significance of the responses. They have demonstrated that changes in surface boundary conditions can have a profound influence on regional rainfall patterns and intensities and can also determine the specifics of how the surface conditions influence atmospheric precipitation.

Surface–precipitation interactions are affected by the heat and moisture supplied by the lower boundary to the planetary boundary layer. Moist convection occurs when an atmospheric column is unstable with respect to a parcel of boundary layer air. Wetter and warmer surfaces promote this instability, as does colder air in the overlying atmosphere. Convection occurring in one location acts to warm and stabilize the atmosphere elsewhere. Therefore, convective rainfall in one location can suppress convection elsewhere. Over land, the diurnal cycle of solar heating promotes the greatest convective instability in the afternoon. Over oceans, warmest sea surface temperatures usually, but not always, produce the greatest convective instabilities. Climate model simulations can test their treatment of surface coupling through comparisons of the simulated and observed diurnal variations of rainfall. These diurnal variations are especially pronounced over land, but they are also notable over the tropical oceans. For example, a model with initially dry soil predicts substantially lower summer convective rainfall in subsequent days, or even weeks, than does a model with initially wet soil. *See* CONVECTIVE INSTABILITY.

Uses of coupled models. Coupled atmosphere–ocean–land models are used to simulate climate over time scales of longer than a few weeks. These models are used to predict future natural variations (that is, variations that occur without human-induced changes of the gas and aerosol composition of the atmosphere) or to project responses to human-induced changes such as the addition of greenhouse gases.

One natural variation is El Niño, the complex of phenomena connected with warm sea surface temperature in the eastern and central tropical Pacific that seems to occur on an irregular 4-year time scale. The simulation of such warm and cold sea surface temperature variations is difficult for the present generation of global coupled climate models. The resolution required for such simulation is approximately 20 km (12 mi), in the tropical Pacific, in order to resolve the thin strip of equatorial upwelling of cold water from below, a crucial process in changing sea surface temperature. Questions such as the dependence of the El Niño cycle on global warming cannot be answered until global coupled models successfully incorporate the full range of processes known to be important for this cycle. By limiting the active extent of the coupled models to the tropical Pacific, such high resolution can easily be attained and predictions of El Niño using such models become possible. Data in the tropical Pacific are used to initialize the forecast, that is, to specify the initial internal state of the upper ocean. The coupled model is then allowed to run freely, and the future evolution of the entire system is predicted. Used in this way, coupled models have provided correct predictions up to a year in advance that have been useful in making economic decisions in nations bordering the tropical Pacific. Extending this approach to the United States is a far more difficult problem for coupled models, but remains a future possibility. *See* EL NIÑO; TROPICAL METEOROLOGY.

In terms of human-induced changes, coupled atmosphere–ocean–land models are used to project a climatic response to the increase of greenhouse gases. Difficulty arises because models are limited in their resolution (that is, in the fineness with which it is possible to subdivide the atmosphere and ocean) and in the parametrization of those processes that cannot be resolved by the present resolution (for example, clouds and turbulence in the atmosphere, small eddies in the ocean).

To project what the climatic response would be to an increase of greenhouse gases, the present climate must first be simulated. Next, an assumption must be made that the composition of the atmosphere will change in a specified way; for example, it may be assumed that carbon dioxide will increase by 1% per year. Current atmosphere–ocean–land models of long-term response to the anthropogenic addition of atmospheric gases have indicated that carbonaceous and sulfate aerosols have a major cooling affect (by direct blocking of the Sun's rays) on regions of the Earth's surface local to the emissions of such constituents. They may also have an indirect cooling effect by acting as cloud condensation nuclei to increase the amount of cloud cover. Future projections of the climatic response to additions of greenhouse

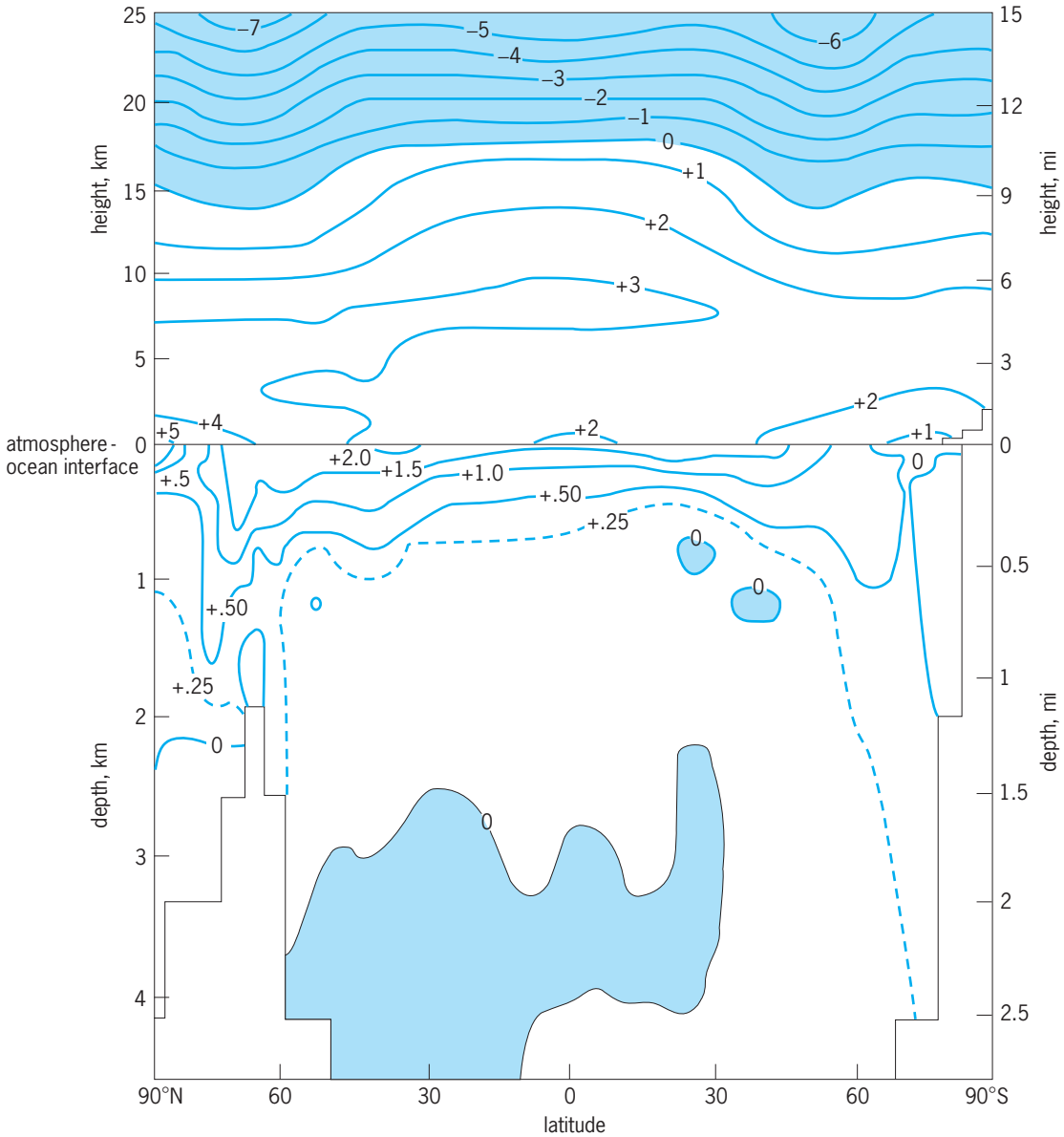

Fig. 3. Model showing the change of temperature (°C) in the atmosphere and ocean averaged over the 60th to 80th year of integration in response to an increase of 1% per year in the concentration of carbon dioxide. The temperature changes are averaged around circles of longitudes. The shaded areas are regions where cooling occurred. 1°C = 1.8°F. (*After S. Manabe, M. J. Spelman, and K. Bryan, Transient response of a coupled atmosphere–ocean model to gradual changes of atmospheric CO_2, pt. I: Annual mean response, J. Climate, 4:785–818, 1991*)

gases must include both the direct and poorly understood indirect effects of aerosols. With this information, a 70-year projection of a climate response, using a state-of-the-art coupled model (**Fig. 3**), shows that the higher latitudes warm more than the tropics, the ocean surface temperatures increase less than the surface atmosphere temperatures, the ocean warms very little below 0.6 mi (1 km) in depth, and the stratosphere cools. The model also shows a substantial summer drying of soils in midlatitudes of the summer hemisphere. This drying could have serious consequences if it were to occur. *See* CLIMATIC PREDICTION; CLIMATOLOGY; WEATHER FORECASTING AND PREDICTION. Robert E. Dickinson; E. S. Sarachik

Uncertainties in climate models. Because the Earth's climate is far too complex to reproduce in a laboratory, scientists have developed numerical climate models, based on the conservation of energy, momentum, and mass, which can be used to simulate past, present, and future climate conditions. These models incorporate the key physical parameters and processes that govern the climate behavior. Climate models that simulate the entire global atmosphere are called general circulation models (GCMs). Those models applied to a limited domain are called regional climate models (RCMs). The horizontal resolution of most general circulation models is 100 km (60 mi) or larger. Regional climate models have a horizontal resolution of a few tens of kilometers, but the initial and lateral boundary conditions of the RCMs depend upon the GCM output or reanalysis, which is a combination of the GCMs' output and

observations. Consequently, RCMs may have more detailed physics and finer resolution, but the uncertainties of the GCM may pass to the RCMs through initial and/or boundary conditions when the RCM is used to predict the future climate.

Climate models (referring to atmospheric models coupled with ocean and land-soil models, hereafter) can be used to simulate changes in temperature, rainfall, snow cover, winds, soil moisture, sea ice, and ocean circulations for periods from seasons to decades. However, the mathematical models used are obviously simplified versions of the real Earth and cannot capture its full complexity. The nonlinear partial differential equations of the climate systems (that is, Navier-Stokes equations) cannot be solved analytically. In climate models, these equations are converted into the discrete difference equations in both time and space and solved numerically. Many numerical schemes and grid structures have been developed and applied in climate models to obtain the approximate numerical solutions. Theoretically, the numerical solutions should converge to the analytical (or true) solutions, but the accuracy can quickly deteriorate as the number of integrations increases in a model using coarse resolution, inappropriate numerical schemes, or improper grid structure. The discrete mode of the solutions also imposes a constraint that the smallest element that can be resolved by the model is twice the resolution (2Δ). The nonlinear terms of the Navier-Stokes equations can generate waves shorter than 2Δ. Those short waves can be misrepresented as the resolvable waves (that is, aliasing), causing nonlinear instability. Artificial filters or smoothing (which can also be generated by truncation errors of the numerical schemes) have been applied to control these instabilities.

Because of the resolution, climate models cannot resolve individual clouds, small-scale turbulence, detailed features of complex terrain, soil properties, or land distributions. They are called the subgrid-scale properties. The effect of the subgrid-scale properties should be formulated by parametrizations. For example, the effect of the subgrid-scale clouds is assumed as a function of the moisture convergence and other fields, which can be resolved by the models. Meteorologists have proposed various cumulus parametrizations to represent the effect of the subgrid-scale clouds; however, it is very difficult to formulate the cumulus parametrizations or other subgrid-scale features based on rigorous mathematics and physics. Models with a fine resolution can resolve the majority of the flow explicitly and reduce the spurious contributions from the subgrid-scale features. But the subgrid-scale parametrizations are still important, even in J. W. Deardorff's large-eddy simulation of a convective boundary layer with a model resolution of 50 m (160 ft). Subgrid-scale features and short-wave disturbances may affect the large-scale weather/climate pattern through the nonlinear terms of the equations.

Uncertainties also exist in formulating cloud physics, radiation, ocean current, and the effects of aerosol, as well as snow–land-surface, air–sea, sea–ice interactions. For example, using the same observed data as the initial conditions, a comparison of the snow–land-surface models shows considerable variation among the results, and most were unable to reproduce the observations due to deficiencies in numerical schemes, physics, or model resolution. For this reason, land models need to be refined before they can appropriately represent the effect of the soil and snow in a climate system. Better parametrizations in cloud physics, turbulence, radiation, aerosols, sea-ice and air-sea interactions, as well as finer model resolution are required in most models. *See* CLOUD PHYSICS.

The accuracy of the initial conditions is very important for numerical weather forecasting. The initial conditions for a GCM come from the reanalysis, which combines observations and model results, both of which may contain errors. The errors of observations may come from instruments or interpretation errors in both time and space. The sources of errors from modeling may be due to numerical schemes, physics, resolution, as well as initial and boundary conditions used in the model. These errors may be amplified or suppressed during the climate model integration. The result of the numerical simulations can become quite different from the observation after just 1 or 2 weeks.

Since uncertainty always exists in a numerical weather prediction model, scientists have applied ensemble-forecasting techniques. The ensemble method may include running the same model with different initial and boundary conditions (for example, to disturb the initial conditions or to start to integrate the model at different times), running several different models with the same initial/boundary conditions, or running different models with different initial/boundary conditions to estimate the probability of the model forecasting. If the results diverge, the confidence of the results is low; otherwise, the confidence is high. There is no guarantee that the real climates are completely covered by the model results. For example, if all of the models are deficient in describing specific physical process, the ensemble average of these models will also be deficient. The model also may not be able to predict extreme cases, which are needed most.

If a model can reasonably simulate the current weather and produce good short-term forecast (7–14 days), it has a better chance to predict future climate change or simulate the past climate because the climate is an accumulation of daily weather processes. A model with more detailed physics, finer resolutions, and more accurate numerical schemes and initial/boundary conditions usually performs better. Although an RCM can simulate the observations reasonably well even after a few years integration when reanalysis is used as the initial/lateral boundary conditions, it cannot be used alone to predict the climate in the future without coupling with a GCM. On the other hand, it is unlikely that GCM-simulated weather systems will compare favorably with observations after a few weeks of integration. Climate studies usually emphasize statistics instead of

describing any single weather event. The climate simulations also depend more on the boundary conditions, which include solar radiation and properties of the soil and ocean, than the initial conditions. Currently, many climate models are being used to simulate the climate in the future or the past. For example, the climate models show that the mean atmosphere becomes warmer under the doubled atmospheric carbon dioxide (CO_2) scenario, with quite different regional responses. When climate models are applied to forecast the future climate, they also encounter uncertainties in volcanic activity, forest fires, the change of trace gases/aerosols and emission, and the change of vegetation and land surface due to human activities. Wen-Yih Sun

Bibliography. T. Crowley and G. North, *Paleoclimatology*, 1996; J. W. Deardorff, Three-dimensional numerical study of turbulence in an entraining mixing layer, *Bound. Layer Meteorol.*, 7:199–226, 1974; G. J. Haltiner and R. T. Williams, *Numerical Prediction and Dynamic Meteorology*, 2d ed., Wiley, New York, 1980; A. Henderson-Sellers and K. McGuffie, *A Climate Modeling Primer*, 3d ed., 2005; J. T. Houghton et al. (eds.), *Climate Change 1995: The Science of Climate Change*, 1996; N. A. Phillips, An example of non-linear computational instability, *The Atmosphere and the Sea in Motion*, Rossby Memorial Volume, pp. 501–504, Rockefeller Institute Press, 1959; S. H. Schneider, Climate modeling, *Sci. Amer.*, 256(5):72–80, 1987; K. Trenberth (ed.), *Climate System Modeling*, 1992; W. Washington and C. L. Parkinson, *An Introduction to Three-Dimensional Climate Modeling*, 2d ed., 2005.

Climate modification

Alteration of the Earth's climate by human activities; humans have the capacity to modify the Earth's climate in several important ways.

Local and regional scale. Conventional agriculture alters the microclimate in the lowest few meters of air, causing changes in the evapotranspiration and local heating characteristics of the air-surface interface. These changes lead to different degrees of air turbulence over the plants and to different moisture and temperature distributions in the local air.

Another example of human influence on climate at a larger scale is that the innermost parts of cities are several degrees warmer than the surrounding countryside, and they have slightly more rainfall as well. These changes are brought about by the differing surface features of urban land versus natural countryside and the unique ways that cities dispose of water (for example, storm sewers). The altered urban environment prevents evaporation cooling of surfaces in the city. The modified surface texture of cities (that is, horizontal and vertical planes of buildings and streets versus gently rolling surfaces over natural forest or grassland) leads to a more efficient trapping of solar heating of the near-surface air. The characteristic scales of buildings and other structures also lead to a different pattern of atmospheric boundary-layer turbulence modifying the stirring efficiency of the atmosphere. *See* MICROMETEOROLOGY; URBAN CLIMATOLOGY.

At the next larger scale, human alteration of regional climates is caused by changes in the Earth's average reflectivity to sunlight. For example, the activities of building roads and highways and deforestation change the reflectivity characteristics of the Earth's surface and alter the amount of sunshine that is reflected to space, as opposed to its being absorbed by the surface and thereby heating the air through contact. Such contact heating leads to temperature increases and evaporation of liquid water at the surface. Vapor wakes from jet airplanes are known to block direct solar radiation near busy airports by up to 20%.

Human activities also inject dust, smoke, and other aerosols into the air, causing sunlight to be scattered back to space. Dust particles screen out sunlight before it can enter the lower atmosphere and warm the near-surface air. An extreme case is popularly referred to as the nuclear winter scenario, where a massive injection of smoke particles into the upper atmosphere occurs during a hypothetical exchange of nuclear blasts. The resulting smoke veil theoretically remains in the atmosphere for up to 6 months and leads to a shading of the world and a resulting cooling of continental interiors by as much as 90°F (50°C) for several months. Ocean and coastal regions would experience less than about 18°F (10°C) of cooling. *See* AIR POLLUTION; NUCLEAR EXPLOSION; SMOG.

Global scale. Humans are inadvertently altering the atmospheric chemical composition on a global scale, and this is likely to lead to an unprecedented warming of the global atmosphere during the next generation. It comes about by anthropogenic injection into the atmosphere of relatively inert trace gases that perturb the radiation balance of the globe as a whole. Most of this gaseous waste comes from burning fuels that contain carbon and nitrogen. Other sources include inert gases used in aerosol spray cans and cooling devices. *See* ATMOSPHERIC CHEMISTRY.

Since alteration of the atmospheric composition by human activities has taken place for more than 150 years, it is important to ask whether the climate system has started to respond to the forcing. While no definitive answer exists, globally averaged temperature data suggest that on the whole the Earth has been warming over the last 100 years. *See* ATMOSPHERE; CLIMATIC PREDICTION; WEATHER MODIFICATION. Gerald R. North

Global warming. One of the most important and best understood features of the atmosphere is the process that keeps the Earth's surface much warmer than it would be with no atmosphere. This process involves several gases in the air that trap infrared radiation, or heat, emitted by the surface and reradiate it in all directions, including back to the surface. This recycling of heat maintains the surface and lower atmosphere of the Earth at a temperature higher than it would otherwise be.

Heat-trapping gases. The heat-trapping gases include water vapor, CO_2, CH_4, and N_2O. These gases

constitute only a small fraction of the atmosphere, but their heat-trapping properties raise the surface temperature of the Earth by a large amount, estimated to be more than 55°F (30°C).

Human activities, however, are increasing the concentrations of CO_2, CH_4, and N_2O in the atmosphere, and in addition industrially synthesized chemicals—chlorofluorocarbons (CFCs) and related compounds—are being released to the air, where they add to the trapping of infrared radiation. Carbon dioxide is released to the air mainly from fossil fuel use, which also contributes to the emissions of CH_4 and N_2O. Agricultural and industrial processes add to the emissions of these gases. These concentration increases add to the already powerful heat-trapping capability of the atmosphere, raising the possibility that the surface will warm above its past temperatures, which have remained roughly constant, within about 3°F (1.7°C) for the past 10,000 years. The atmospheric concentration of CO_2 has increased to about 363 parts per million by volume (ppmv) from its preindustrial value of about 280 ppmv (see **illus.**); CH_4 has increased at a greater rate, from 0.75 ppmv to more than 1.7 ppmv. Nitrous oxide is growing slowly, at about 0.2% per year, and has a present concentration of 0.32 ppbv. The CFCs and related halocarbons have had the fastest rates of growth, up to 15% per year, but their concentrations are very small—300 parts per trillion by volume (pptv) for CFC-11, 535 pptv for CFC-12, and 3 pptv for halon-1301—the fastest growing species. Treaties are in existence that control internationally the production and use of many of the CFCs and related compounds, so their rate of growth has slowed, and for some a small decrease in atmospheric concentration has been observed. Large emissions of sulfur dioxide in industrial regions are thought to result in airborne sulfate particles that reflect sunlight and decrease the amount of heating in the Northern Hemisphere. *See* GREENHOUSE EFFECT.

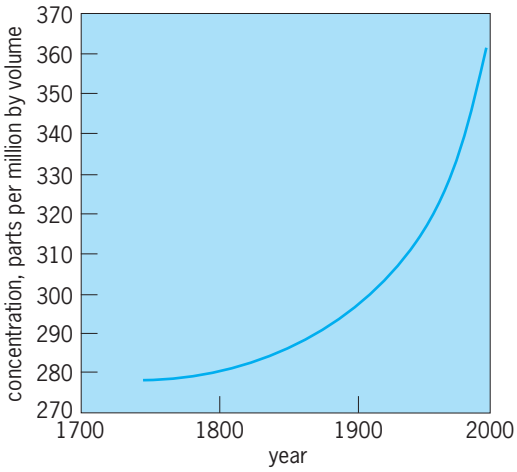

History of atmospheric CO_2 concentration as inferred from an analysis of bubbles taken from ice cores in the Antarctic. (*After E. M. Neftel, H. Oeschger, and B. Stauffer, Evidence from polar ice cores for the increase in atmospheric CO_2 in the past two centuries, Nature, 315:45–47, 1985; and for measurements since 1970, data from Carbon Dioxide Information Analysis Center home page*)

Climate warming impacts. Estimates indicate that if human-induced emissions of heat-trapping gases continue according to the current pattern the globally averaged surface temperature could increase at a rate of about 0.4°F (0.22°C) per decade in the twenty-first century. This rate of global temperature change is much faster than any sustained global average temperature change documented in the last 10,000 years. Lifeforms have evolved with and are adapted to current climate conditions, including temperature. For this reason it is anticipated that should the warming occur as estimated, there will be serious damage to ecological systems.

Even slow, small climate changes can have large regional effects. There is evidence that about 6000 years ago parts of the Sahara Desert had sufficient rainfall to support vegetation and large mammals, while prairies advanced eastward and replaced forests, in what is now the United States, because of lack of rain. These changes were part of a global change characterized by an average temperature increase of only 2–4°F (1–2°C).

Rapid regional changes in climate occurring more than 10,000 years ago have been deduced from evidence in ice cores and from pollen found in layers in bogs and ponds; these changes were accompanied by major ecological changes in the region of the climate change. For example, about 11,000 years ago in what is now Europe trees were suddenly replaced by shrubs at the time of a large cooling in areas around the North Atlantic. Such events add to the evidence that a rapid climate change can damage ecosystems. Besides ecological damage, higher temperatures will melt glaciers and expand seawater, thereby raising sea level. Some regions will have climate changes considerably larger or smaller than the global average.

Even though the process is called global warming, the focus on temperature is partly for convenience; all climate features are expected to change by varying amounts as part of an overall climate change. For example, the temperature of the stratosphere is expected to decrease, and total precipitation is expected to increase while its annual distribution shifts in time and place. The role of CFCs in climate warming will be to increase the amount of heat trapped in the lower part of the atmosphere but, by destroying ozone in the stratosphere, to increase the ability of heat to leave the atmosphere. The net heating by CFCs will therefore be small; they will produce instead a change in the vertical temperature distribution of the atmosphere.

Climate change theory. There are many difficulties in producing a completely convincing calculation of the rate of average climate warming to be expected, and a specification of the detailed change in each climate parameter that will be experienced in each major region. These difficulties stem from the fact that the climate system consists of many components—atmosphere, oceans and ocean biota, ice caps and sea ice, land surface and ecosystems, and

the human societies that are forcing rapid changes in the composition of the air—making the calculation very complex. In addition, small changes in the amount of sunlight reaching the Earth and large volcanic eruptions that place sulfur dioxide in the stratosphere complicate studies of human-induced climate change.

Attempts to simulate the interactions of all of these climate components, though using the largest computers available, require great simplification in the description both of the components and of the interactions among them. These simplifications lead to sizable uncertainties, especially when they involve processes that tend to amplify or diminish the degree of warming. Increased evaporation of water from the oceans as the surface warms and an increased water-holding capacity of the warmer air will lead to a higher water vapor concentration in the atmosphere. Water vapor is a heat-trapping gas; an increase in its average concentration in the air will add its heat-trapping effect to that of the human-emitted gases and amplify the warming. The size of the warming will be decreased by the fact that a warmer Earth will emit more infrared radiation to space, increasing the Earth's heat loss.

Some of the most complicated processes that are important in the calculation describe the role of clouds and the changes that clouds would undergo on a warmer Earth. Clouds both reflect sunlight away from the Earth and trap infrared radiation in the lower atmosphere, so changes not only in the amount of cloud but also in their optical characteristics must be correctly estimated. However, recent progress in climate calculations has contributed to improved agreement between observed climate heating since 1870 and model results, suggesting that the simple treatment of clouds in models is not seriously deficient. The comparison has also been aided by the rapid surface heating since 1975 that provides a larger heating signal compared to random temperature fluctuations. *See* CLOUD PHYSICS.

In addition to the known model deficiencies, there is reason to believe that models in general do not adequately represent sudden changes in complex systems. Thus a gradual change in surface temperature might lead to a sudden reversal of a major ocean current, or a gradual change in rainfall might lead to a collapse of a tropical ecosystem, without either event being foreseen in the model calculations.

All of these difficulties combine to produce a wide range of estimates for the rate of future climate warming. The Intergovernmental Panel on Climate Change (IPCC) of the United Nations estimates that if the current pattern of emissions of heat-trapping gases continues the average surface temperature should increase 0.31–0.43°F (0.17–0.24°C) per decade. But the IPCC asserts that there are enough uncertainties that a faster or slower change cannot be completely ruled out, and the threat of surprises must also be kept in mind. *See* CLIMATE HISTORY; CLIMATE MODELING; HEAT BALANCE, TERRESTRIAL ATMOSPHERIC.

Response to climate warming. The presence of numerical uncertainties and the possibility of surprises increase the difficulty for those who must develop national and international responses to the threat implicit in the human-induced climate change. Their task of selecting a response is made simpler, however, by the fact that the principal driving force for such a climate change—the use of fossil fuels—is also the driving force for a number of other atmospheric problems. Urban air pollution, regional pollution by low-altitude ozone, visibility problems over scenic vistas, and acid precipitation all involve fossil fuel emissions, so that an uncertainty in specific aspects of any one problem does not shift the focus away from reduced fossil fuel use as the core strategy for slowing harmful human impacts on the atmosphere. *See* ACID RAIN.

John Firor

Climate Change Impacts

Substantial uncertainties exist as to the severity, timing, and rate of human-induced global climate change. Nevertheless, there are particular attributes of Earth's physical processes, natural ecosystems, and human socioeconomic activities that would be affected by a changing climate.

Climatic change is often broadly discussed in global terms, but the impacts of a warmer Earth would be experienced at the local and regional levels. Many potential impacts may be considered undesirable from an environmental or socioeconomic perspective, whereas others may be seen as beneficial. Although it is difficult to predict with certainty what the impacts of a possible global climate change may be in particular regions, a range of possible consequences can be described.

Physical processes. Global warming is projected to be greater in the higher latitudes and polar regions than in the tropics, causing a decrease in the difference in temperatures between the poles and the tropics as well as altering global atmospheric circulation patterns. This in turn would change the timing, distribution and amount of regional precipitation, affect the occurrence and location of extreme weather events such as hurricanes and tornadoes, alter the timing of mountain snowmelt and river basin runoff, reduce high-latitude snow and ice cover, and eventually lead to rising sea levels around the world. *See* ATMOSPHERIC GENERAL CIRCULATION; HURRICANE; PRECIPITATION (METEOROLOGY); TORNADO.

Natural ecosystems. The projected climate change may occur at a faster rate than any natural climatic warming has occurred in the past 10,000 years—a period covering the entire history of human civilizations. A rapid warming, or abrupt shifts in regional climate, could harm biological communities that have developed under relatively stable climatic conditions over the past several centuries and exist within relatively narrow ranges of average climatic conditions.

Terrestrial natural ecosystems would be the first to be affected by global warming, followed by marine ecosystems. The terrestrial environment is sensitive not only to average climatic conditions but also to extreme conditions such as extended periods of high or low temperatures, droughts, floods, and

other weather hazards. Some species of vegetation can survive and reproduce under a broad range of climate conditions. Other species require not only a stable climate but a narrow range of meteorological extremes such as high and low temperatures, precipitation, and soil moisture. The Intergovernmental Panel on Climate Change, created by the United Nations in 1988, has concluded that over the next 50 years projected changes in regional temperatures and precipitation could cause climatic zones to shift hundreds of kilometers toward the polar regions. Some biotic communities may be capable of adapting to a climatic shift of such magnitude, while others would experience substantial loss that could include some species extinction. *See* EXTINCTION (BIOLOGY).

In the Northern Hemisphere's midlatitude temperate zone spanning North America and the Eurasian landmass, such a large climatic shift would have a tremendous impact on forest ecosystems. Individual tree species might not be able to migrate into climatic zones similar to those that they presently occupy. If climate change were to proceed more rapidly than temperate forests could successfully migrate poleward, large-scale demise of trees could be anticipated. Climate change could cause weather-related stress for individual tree species, exposing them to pest damage. Changes in forestry would also affect associated flora and fauna that are adapted to existing forest ecosystems and climatic conditions. *See* FOREST ECOSYSTEM.

According to the IPCC, the permafrost that presently underlies the 20–25% of the Northern Hemisphere that is perenially frigid may begin to thaw during the next 40–50 years. Degradation of the permafrost cover would damage high-latitude surface ecosystems, destabilize terrain, and cause soil erosion. *See* PERMAFROST.

Rising sea levels would inundate low-lying river estuaries and coastal wetlands (including tropical mangrove forests), which support established biotic communities. Again, the degree of environmental impact would be a function of both the rate and extent of sea-level rise, as rapid change would not permit adaptation by biotic communities. In addition, some low-elevation islands could be partially inundated, with substantial impacts to coastal ecology and freshwater supplies as well as to terrestrial biota.

The biological productivity and diversity of marine ecosystems would be affected by changes in water temperature and, in the long term, changes in ocean currents and ocean circulation. Inundation of estuaries and coastal wetlands would adversely affect the reproductive cycles of valuable fish and shellfish species. *See* BIOMETEOROLOGY; CLIMATOLOGY; ECOSYSTEM; MARINE ECOLOGY. Steven L. Rhodes

Bibliography. D. E. Abrahamson (ed.), *The Challenge of Global Warming*, 1989; W. S. Broecker, Unpleasant surprises in the greenhouse, *Nature*, 328:123–124, 1987; T. J. Crowley, Causes of climate change over the past 1000 years, *Science*, 289:270–277, 2000; J. Houghton, *Global Warming: The Complete Briefing*, 3d ed., 2004; M. E. Mann, Climate during the past millennium, *Weather*, 56:91–101, 2001; E. M. Neftel, H. Oeschger, and B. Stauffer, Evidence from polar ice cores for the increase in atmospheric CO_2 in the past two centuries, *Nature,* 315:45–47, 1985; Panel on Policy Implications of Greenhouse Warming, Committee on Science, Engineering and Public Policy (U.S. National Research Council), *Policy Implications of Greenhouse Warming: Mitigation, Adaptation, and the Science Base*, 1992; V. Ramanathan, The greenhouse theory of climate change: A test by an inadvertent global experiment, *Science*, 240:293–299, 1988; S. H. Schneider, Climate modeling, *Sci. Amer.*, 256(5):72–80, 1987; R. T. Watson (ed.), *Climate Change 2001: Synthesis Report*, Intergovernmental Panel on Climate Change 3d Assessment Rep., 2002.

Climatic prediction

The description of the future state of the climate, that is, the average or expected atmospheric and Earth-surface conditions, for example, temperature, precipitation, humidity, winds, and their range of variability. *See* CLIMATOLOGY.

Seasonal and interannual climate predictions, made many months in advance, provide useful information for planners and policy makers. Various countries are working on operational (systematic and regular) climate prediction products, and by the beginning of the twenty-first century climate prediction will be added to weather prediction as an operational mission requirement of the world's major meteorological services. *See* METEOROLOGY. Jay S. Fein

Seasonal to Interannual Prediction

It is well known that the Earth revolves around the Sun in a periodic orbit giving rise to a well-defined annual cycle in the Earth's climate, and that the annual cycle in the incident solar radiation is nearly identical from one year to the next. The question as to why the winter climate of one year is so different from the winter climate of another year has long interested climate scientists, because an understanding of the causes of interannual variations of the Earth's climate might provide some clues about the predictability of these variations. *See* SOLAR CONSTANT; SOLAR RADIATION.

The fundamental reason for departures from perfectly periodic variations in the atmosphere and oceans—in spite of the fact that the solar forcing is nearly perfectly periodic—is that the physical laws that determine the motions of air and water in the Earth-atmosphere-ocean system are inherently nonlinear and chaotic. The internal instabilities of this nonlinear dynamical system are capable of producing a nonperiodic behavior even if the external forcing (the Sun) is perfectly periodic. Because of the nonlinear nature of the mathematical equations describing motions in the Earth-atmosphere-ocean system, it is impossible to predict these motions for a long time in the future. One major endeavor of researchers in the field of meteorology and oceanography is to determine the range at

which different properties of the nonperiodic components of atmospheric and oceanic motions can be predicted.

Weather prediction. The day-to-day changes of weather cannot be predicted for a long time because even minute errors in observations amplify with time and become quite large within a few days or a few weeks, destroying the value of the forecast. The track of a storm seen on a weather map for a given day, for example, can be predicted for several days; but beyond a week or two the storm intensity, structure, and position are unpredictable. Improvements in weather-observing systems and the use of more complete equations in the weather prediction models will improve the quality of weather forecasts and the time for which they are valid, but the limitations arising from the amplification of imperceptible errors in the initial conditions of the forecast calculation will always restrict detailed weather forecasts to periods of a week or two in the future.

Despite this fundamental limitation in the ability of meteorologists to predict day-to-day changes in weather, it seems possible to predict seasonal averages of weather with sufficient accuracy to be of benefit to society. For example, even if it is not possible to predict the life cycle of individual weather disturbances that provide the monsoon rainfall over Asia, it may be possible to predict the seasonal mean rainfall over the monsoon region. *See* MONSOON METEOROLOGY; WEATHER FORECASTING AND PREDICTION.

Boundary conditions at Earth's surface. The conditions at the Earth's surface lower boundary of the atmosphere (for example, surface temperature, soil moisture and vegetation, or snow and sea ice cover) can have large variations from one year to another. It has been proven by statistical analysis of observations and by experiments using mathematical climate models that interannual changes in these boundary conditions are directly responsible for interannual changes in the seasonal mean climatic conditions in many parts of the globe. Therefore, if changes in these boundary conditions could be predicted, it would be possible to predict changes in seasonal mean climatic conditions in certain regions of the globe. Since changes in the boundary conditions are caused by interactions with the atmosphere, there is no guarantee that they are predictable. However, boundary conditions change rather slowly, and therefore they can be assumed to persist for periods up to a season. Thus it is possible to predict seasonal mean climatic conditions for the coming season (to the extent that they are influenced by the anomalous boundary conditions) by using the current boundary conditions. For longer periods, the boundary conditions cannot be assumed to persist, so that prediction of climatic conditions must necessarily involve coupled models of the land-atmosphere-ocean system, which enable prediction of future boundary conditions. *See* CLIMATE MODELING.

Seasonal predictions. For seasonal climate predictions to be useful, the seasonal climate variability caused by interactions with the atmosphere's lower boundary (for example, sea surface temperature) must be significantly larger than the changes in the atmosphere that can occur in the absence of any boundary forcing (normally referred to as the natural variability). The natural variability of seasonal mean climate is different in different seasons and in different regions. Therefore, the extent to which seasonal mean climate variations can be usefully predicted is also different for different regions. For example, the weather fluctuations over North America are far more vigorous during the winter season than during the summer season. The weather fluctuations over large parts of the tropics are far less energetic (day-to-day changes are generally small) compared to those in the higher latitudes. Therefore, the seasonal mean climate in the tropics is far more predictable than that in the higher latitudes. Likewise, seasonal mean climate in high latitudes is more predictable in summer than in winter.

Tropical droughts. Most of the tropical regions, and especially the regions that experience monsoons, have well-defined rainy seasons. For example, most of the annual rainfall over India and sub-Saharan Africa occurs during the summer monsoon season (June–September), and over northeast Brazil during March–May. Although the individual rainfall episodes occur in association with individual clouds and weather-scale disturbances, it is the large-scale convergence of moisture, generally associated with the intertropical convergence zone, that is responsible for the seasonal mean rainfall. Most of the tropical droughts and floods can be understood as space-time shifts in the annual cycle. For example, severe droughts in northeast Brazil occur whenever the intertropical convergence zone remains farther north of its mean position during the rainy season. Likewise, droughts over India and Australia are often associated with the eastward shift of the mean climatological rainfall maximum, giving rise to large increases in rainfall in the central Pacific. An important discovery was the recognition that these space-time shifts of rainfall maxima in the tropics are due to changes in the boundary conditions of sea surface temperature and land surface conditions. The tropical atmosphere has a unique property, distinct from that of mid and high latitudes, that seasonal and interannual variations are largely determined by changes in the slowly varying boundary conditions at the Earth's surface. Therefore, the prospects for long-range forecasting in the tropics are good. *See* DROUGHT; PRECIPITATION (METEOROLOGY); TROPICAL METEOROLOGY.

Extratropical variations. Seasonal climate variations in the extratropics are not as strongly controlled by the boundary conditions, but are strongly influenced by the internal dynamical processes in the atmosphere. For example, a major blocking event over North America can significantly change the monthly or even seasonal mean climate for that year. Prediction of seasonal and interannual variations during the winter season is particularly difficult because the day-to-day fluctuations are highly energetic and so give rise to large uncertainty in the prediction of the seasonal mean climate. Some evidence,

however, indicates that large changes in the boundary conditions in the tropics, causing large changes in tropical rainfall and therefore atmospheric heating, can produce predictable effects even in the extratropics. This result is possible because the tropical influences can propagate away from the Equator and influence the extratropical climate variations. This mechanism of tropical-extratropical interaction has given some hope for seasonal and interannual prediction in the extratropics.

The day-to-day variations in the extratropics in the late spring and summer season bear some similarity to the tropical weather variations because they are not as energetic or intense as during the winter season. This reduction in the uncertainty due to the internal dynamical processes should make it possible to predict the influences of slowly varying boundary conditions in the spring and summer season. Relationships between spring rainfall and summer droughts over North America are good examples of predictability due to the effects of the land surface boundary conditions. *See* ATMOSPHERIC GENERAL CIRCULATION; STORM.

El Niño Southern Oscillation (ENSO). The outstanding example of interannual climate variability is the occurrence of multiyear episodes of warm and cold sea surface temperature in the central and eastern tropical Pacific; these episodes are generally referred to as El Niño. These variations are a consequence of the interactions between the upper layers of the ocean and the overlying atmosphere. Changes in the prevailing southeasterly winds over the equatorial Pacific can produce changes in the wave activity in the upper layers of the ocean and the upwelling of cold water, giving rise to changes in sea surface temperature, which in turn can produce changes in the surface winds. This mutual enforcement of oceanic and atmospheric changes (a form of instability of the coupled ocean-atmosphere system) is responsible for multiyear variations in sea surface temperature in the equatorial Pacific and the associated changes in sea-level pressure. In particular, the simultaneous occurrence of higher-than-average pressure in the Pacific Ocean and lower-than-average pressure in the Indian Ocean, and vice versa, is noteworthy. This recurrent seesaw pattern of pressure changes is referred to as the Southern Oscillation.

Associated with these changes in sea surface temperature, surface pressure, and surface wind are large changes in rainfall along the equatorial belt. During the episodes of colder-than-average sea surface temperature in the central and eastern Pacific and stronger-than-average southeasterly trades, the maximum seasonal rainfall is found at the extreme western position near the warmest surface waters of western Pacific. During these years, monsoon rainfall over India and Australia is generally above average. During the episodes of warmer-than-average sea surface temperature in the central and eastern Pacific and weaker-than-average surface winds, the rainfall maximum shifts eastward, producing severe droughts over India and Australia and heavy floods over the central Pacific and the west coast of South America. Actually, there are global effects of these changes in sea surface temperature, rainfall, and circulation in and over the tropical Pacific Ocean.

The discovery of these interrelationships among sea surface temperature, rainfall, and circulation, and the ability to predict these variations by using coupled models of the ocean-atmosphere system, has established a scientific basis for prediction of seasonal and interannual climate variations by using dynamical models of the climate system.

Forecasts. Several countries, including the United States, routinely issue forecasts for the coming season. These forecasts, generally issued in terms of probabilities of occurrence of above- or below-average temperature and rainfall, are based on empirical methods in which certain properties of atmospheric circulation, rainfall, and surface boundary conditions (generally sea surface temperature and snow cover) in the preceding months are used to forecast the climate for the coming season. It is generally accepted that there is some skill in these forecasts. Efforts have been made to develop and implement dynamical models of varying complexity of the coupled land-atmosphere-ocean system to predict seasonal and interannual variations. J. Shukla

Decadal-to-Century Climate Prediction

While the potential benefits of being able to predict climate on the decadal-to-century time scale are very large, so too are the problems associated with trying to make such predictions. To understand the nature of these problems, it is useful to consider first the decadal-to-century prediction of the change in the surface temperature averaged over the Earth, and then to consider the even more difficult problem of predicting the geographical distributions of surface temperature and the other climatic quantities of importance to humanity, especially precipitation.

Global average temperature. To facilitate thinking about predicting future global average surface temperature, it is necessary to consider the past behavior of this quantity. The observed record of decadal average temperature, T, relative to its value in 1860–1869 (**Fig. 1**) shows that during the first six decades of the

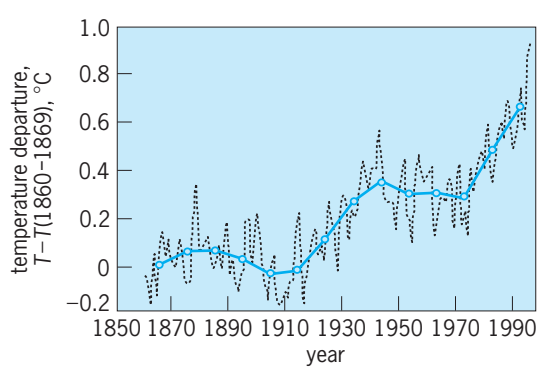

Fig. 1. Observed decadal-average global-average surface temperature (solid line) and observed annual-average global-average surface temperature (broken line). Temperature change in °F equals 1.8 times the temperature change in °C.

record, 1860–1869 to 1910–1919, T varied very little, with no discernible trend. From 1910–1919 to 1940–1949, T increased each decade. This warming trend then ceased. Then T increased from 1970–1979 to 1990–1998 such that it was 0.66°C (1.19°F) warmer in 1990–1998 than in 1860–1869. In order to understand why the decadal average temperature varied in this manner, to determine if predictions of such variations can be made and by what means, it is necessary to consider the transfer of energy across the top of the Earth's atmosphere.

When the amount of infrared thermal radiation emitted by the Earth to space equals the amount of solar radiation absorbed by the Earth from the Sun (240 W m^{-2}), the net input of energy into the atmosphere-Earth system is zero. If somehow the net input energy becomes positive, energy will enter the system, the decadal average temperature will increase, and the infrared radiation emitted to space (which depends on the decadal average temperature) will increase. Conversely, if the net energy input becomes negative, energy will leave the system, the decadal average temperature will decrease, and the infrared radiation emitted to space will decrease. In either case the net input of energy into the atmosphere-Earth system will be restored to zero. Consequently, any factor that causes the net energy input to become nonzero can cause the decadal average temperature to vary. To predict the change in this temperature value, the change in the net energy input from zero (the radiative forcing) must be predicted.

One factor that has changed the net energy input since the mideighteenth century, and will likely change it in the future, is the rising concentrations of atmospheric greenhouse gases: carbon dioxide, methane, nitrous oxide, and chloroflurocarbons (whose emission is now restricted by international convention). Acting alone, the increaseing greenhouse gases would have made the net energy input positive by about 2.6 W m^{-2} in 1998 (**Fig. 2**). Another factor influencing the net energy input is the input of sulfur dioxide gas into the atmosphere near the Earth's surface by the anthropogenic burning of sulfur-containing coal and oil. This gas is chemically converted into sulfate aerosol particles, which enhance the reflection of solar radiation to space in cloudless air. It is thought that these particles also act as nuclei for the formation of cloud droplets, smaller than would be the case in their absence, which therefore enhance the reflection of solar radiation in cloudy air. The enhanced reflection in clear and cloudy air reduces the absorption of solar radiation by the Earth, which, acting alone, would make the net energy input negative. The size of this anthropogenic sulfate aerosol radiative forcing is quite uncertain, particularly that for cloudy air. One estimate gives a value of about −1.1 W m^{-2} in 1998 (Fig. 2). *See* GREENHOUSE EFFECT; TERRESTRIAL RADIATION.

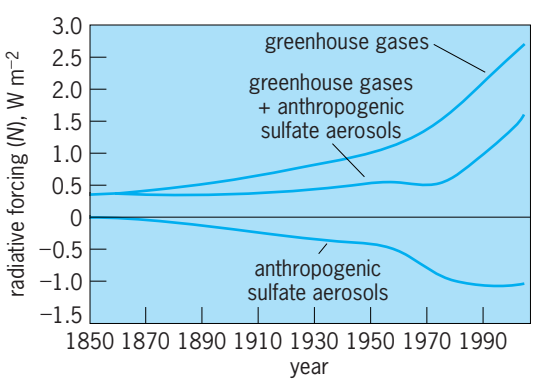

Fig. 2. Radiative forcing of climate by greenhouse gases and anthropogenic sulfate aerosols, together with their sum. The greenhouse gas forcing is taken to be zero in 1765, and the anthropogenic sulfate aerosol forcing is taken to be zero in 1850.

If the future radiative forcing of climate change could be predicted as a function of time, to make a prediction of the resulting climate change a climate model would be needed. A climate model is the mathematical expression of the fundamental laws of nature, for the present example, conservation of energy. The resulting equations express the time rate of change of the decadal average temperature at any time in terms of the value of the decadal average temperature at that time. Starting from the known value of this temperature at some initial time, the time rate of change of the decadal average temperature is calculated from the equations and multiplied by a small time interval; and the result is added to the initial temperature value to give a value at a future time. This procedure is then repeated to predict temperatures at further time increments. The climate model must include the ocean as well as the atmosphere, for the rate of heat removal from the upper ocean by transport into the deeper ocean, together with the net energy input into the atmosphere-Earth system at a given time, determines the rate at which the decadal average temperature changes in time.

In one simple climate model, the ocean is divided vertically into many layers, with the uppermost being the mixed layer wherein temperature is uniform with depth. Also, the ocean in each hemisphere is divided horizontally into a polar region where bottomwater is formed, and a nonpolar region where there is vertical upwelling. In the nonpolar region of each hemisphere, heat is transported upward toward the surface by the water upwelling there, and downward by physical processes whose effects are treated as an equivalent diffusion. Heat is also removed from the mixed layer in the nonpolar region of each hemisphere by transport to the polar region and downwelling toward the bottom. This heat is ultimately transported upward from the ocean floor in the nonpolar region.

Several quantities must be prescribed in this simple climate model. One of the most important is the sensitivity of the climate system, a parameter that represents the change in the global average surface temperature resulting from a radiative forcing equal to that caused by doubling the concentration of carbon dioxide (the net energy input equals 4.4 W m^{-2}) after sufficient time has passed for the net energy

input to be restored to zero. The value of the sensitivity is quite uncertain. The most comprehensive climate models that predict this value—the general circulation models—give values that range from about 1.5 to about 4.5°C (2.7 to 8.1°F). The simple climate model can be used to estimate the sensitivity of the climate system, together with the radiative forcing of the anthropometric sulfate aerosol for a specific year, by minimizing the discrepancy between the simulated and observed decadal average temperature, together with the discrepancy between the simulated and observed interhemispheric temperature difference. The resulting decadal average temperature simulated by the model does reproduce the overall increase in the observed decadal average temperature, but not the detailed variations.

One possible reason for the discrepancy between the observed and simulated temperatures is that one or more factors influencing the net energy input may not have been taken into account. For example, some scientists believe that the energy emitted by the Sun has varied over time, thereby causing a variation in the solar radiation entering the top of the Earth's atmosphere—the solar irradiance. Although there are no useful observations of solar irradiance prior to 1978, when satellite observations began, and although these satellite observations show irradiance variations that are too small to influence the decadal average temperature by more than a few hundredths of a degree Celsius, observations of sunspots since the seventeenth century and of solar magnetic activity during the twentieth century, plus astronomical observations of Sun-like stars since 1970, suggest that variations of solar irradiance sufficiently large to influence the decadal average temperature cannot be ruled out. Other scientists believe that variations in the occurrence of large volcanic eruptions may influence the decadal average temperature by varying the amount of solar radiation reflected to space. Although the sulfate aerosols created in the stratosphere by the injection of sulfur dioxide gas there by large volcanic eruptions do tend to reflect more solar radiation to space and decrease the decadal average temperature, the episodic nature of volcanic eruptions limits the temporal influence of a single volcanic eruption to less than a decade. It is argued, nevertheless, that the variation in the number of large volcanic eruptions per decade can significantly influence the decadal average temperature. Yet another factor which can influence the net energy input is the aerosol particles released into the lower atmosphere by biomass burning, such as by tropical deforestation and slash-and-burn agriculture. Furthermore, changes in the fraction of solar radiation reflected by the Earth's surface—the surface albedo—due to anthropogenic modifications of the surface can also change the net energy input. *See* ALBEDO.

A different possible reason for the discrepancy is that, even if there were no changes in the net energy input due to factors such as those described above, the decadal average temperature might vary because of interactions between the atmosphere and ocean, or by variations within the ocean itself. A well-known example of an atmosphere-ocean interaction is the El Niño Southern Oscillation, which occurs irregularly over the tropical Pacific Ocean with a time scale of 3–7 years. One analysis of the discrepancy suggests that it is likely the result of a temperature oscillation within the North Atlantic Ocean, with a time scale of about 65 years.

Distributions of temperature and precipitation. Although the future course of global average surface temperature can, in principle, be predicted with a simple climate model, the prediction of the geographical distributions of temperature and other climatic quantities requires a much more complex climate model, one at the other terminus of the climate model hierarchy.

Only the general circulation model determines the geographical and vertical distributions of an ensemble of climatic quantities, including temperature, wind, water vapor, clouds and precipitation in the atmosphere; soil moisture, soil temperature, and evaporation on the land; and temperature, currents, salinity, and sea ice in the ocean. The mathematical equations in general circulation models that express the fundamental laws of nature are so complex, however, that they can be solved only at discrete geographical and vertical locations, and only over discrete time intervals. A typical general circulation model divides the atmosphere into thousands of three-dimensional volumes, each having dimensions of about 250 km (155 mi) in both the north-south and east-west directions, and 1 km (0.6 mi) in the vertical direction. Reducing these dimensions is severely constrained by the speed of available supercomputers. For example, increasing the horizontal resolution of a general circulation model from 250 km (155 mi) to 25 km (16 mi) would increase the required computer time a thousandfold, from about 2 weeks to compute a greenhouse-gas induced change in equilibrium climate—the climate change when the net energy input is restored to zero—to more than 30 years.

The constraint on horizontal resolution imposed by available supercomputers also has profound significance for the physical processes that can be included in a general circulation model. In particular, the Earth's climate system encompasses physical processes ranging in horizontal size from 40,055 km (24,889 mi) for disturbances that span the entire circumference of the Earth, to about one-millionth of a meter (0.00004 in.) for the condensation of water vapor onto aerosol particles and dust to form clouds. General circulation models using supercomputers explicitly include physical processes having horizontal sizes of about 250 km (155 mi) and larger. Worse yet, the physical processes smaller than 250 km (155 mi) cannot be ignored, because their effects significantly determine climate and climate change. Thus, climate modelers face the dilemma that their models cannot resolve the small-scale physical processes, and they cannot ignore the effects. This is the major difficulty in modeling the Earth's climate.

The approach taken to overcome it is to determine the effects of the unresolved physical processes on the scales resolved by a general circulation model by using information available only on the resolved scales. This approach is called parametrization. The principal differences among general circulation models lie in their different parametrizations, particularly of cloud and precipitation processes. These parametrization differences have a significant influence on the climate simulated by the general circulation models.

The geographical distributions of the change in surface air temperature simulated by general circulation models for a doubling of the carbon dioxide concentration show that although there are similarities in the regional temperature changes simulated by the global circulation models, there are also marked differences. This fact is even more apparent in the changes in precipitation simulated by the general circulation models. This fact is not surprising because many of the physical processes that determine precipitation have horizontal scales smaller than the resolutions of the general circulation models, and it is these unresolved processes that have had to be parametrized in the models. To improve this situation will require that at least mesoscale meteorological processes, with horizontal sizes down to 50 km (31 mi), be explicitly resolved by the models. A fivefold increase in the horizontal resolution of the general circulation models and an attendant 125-fold increase in computer speed will be required. Such high-speed supercomputers, capable of performing several trillion mathematical operations per second, are projected to be available about the year 2000.

However, to be able to predict the regional climate on a decadal-to-century time scale requires the knowledge of the three-dimensional distributions of temperature, salinity, and velocity in the ocean at the beginning of the prediction period. The panoceanic measurement system required to obtain this knowledge does not yet exist. Michael E. Schlesinger

Bibliography. S. Baliunas and R. Jastrow, Evidence for long-term brightness changes of solar-type stars, *Nature*, 348:520–523, 1990; J. T. Houghton et al. (eds.), *Climate Change 1995: The Science of Climate Change*, Cambridge University Press, Cambridge, 1996; P. P. Jones, Hemispheric surface air temperature variability: A reanalysis and update to 1993, *J. Climate*, 7:1794–1802, 1994; D. E. Parker, C. K. Folland, and M. Jackson, Marine surface temperature: Observed variations and data requirements, *Climatic Change*, 31:559–600, 1995; M. E. Schlesinger and N. Ramankutty, An oscillation in the global climate system of period 65–70 years, *Nature*, 367:723–726, 1994.

Climatology

The scientific study of climate. Climate is the expected mean and variability of the weather conditions for a particular location, season, and time of day. The climate is often described in terms of the mean values of meteorological variables such as temperature, precipitation, wind, humidity, and cloud cover. A complete description also includes the variability of these quantities, and their extreme values. The climate of a region often has regular seasonal and diurnal variations, with the climate for January being very different from that for July at most locations. Climate also exhibits significant year-to-year variability and longer-term changes on both a regional and global basis.

Climate has a central influence on many human needs and activities, such as agriculture, housing, human health, water resources, and energy use. The influence of climate on vegetation and soil type is so strong that the earliest climate classification schemes were often based more on these factors than on the meteorological variables. While technology can be used to mitigate the effects of unfavorable climatic conditions, climate fluctuations that result in significant departures from normal cause serious problems even for modern industrialized societies and more so for the less developed ones. The goals of climatology are to provide a comprehensive description of the Earth's climate over the range of geographic scales, to understand its features in terms of fundamental physical principles, and to develop models of the Earth's climate for sensitivity studies and for the prediction of future changes that may result from natural and human causes. *See* CLIMATE MODELING; CLIMATIC PREDICTION.

Physical basis of climate. The global mean climate and its regional variations can be explained in terms of physical processes. For example, the temperature is warmer near the Equator than near the poles (see **table** and **Figs. 1** and **2**). This is because the source of heat for the Earth is the radiant energy coming from the Sun (insolation), and the Sun's rays fall more directly on the Equator than on the poles. The circulations in the atmosphere and in the oceans transport heat poleward and thereby reduce the Equator-to-pole temperature difference that is continually forced by insolation. The energy released by the rising of warm air in the tropics and the sinking of cold air in high latitudes drives the great wind systems of the atmosphere, such as the trade winds in the tropics and the westerlies of middle latitudes. *See* ATMOSPHERIC GENERAL CIRCULATION; OCEAN CIRCULATION; TROPICAL METEOROLOGY; WIND.

The flux of solar energy at the mean distance of the Earth from the Sun is about 1367 $W \cdot m^{-2}$. The supply of energy per unit of the Earth's surface area is controlled by geometric and astronomical factors. Because the axis of rotation of the Earth is inclined at an oblique angle to the plane of the Earth's orbit, the declination angle of the Sun undergoes a seasonal variation as the Earth makes its annual circuit around it. The declination angle is equivalent to the latitude at which the Sun is directly overhead at noon. The declination angle varies between 23.5° N at northern summer solstice (June 21) and 23.5° S at northern winter solstice (December 22), for the current

Temperature (T) and precipitation (P) for selected stations in North America, South America, and Antarctica*

Station	Latitude, degrees	T_{annual}	T_{Jan}	T_{July}	P_{annual}	P_{Jan}	P_{July}
Alert	82.50 N	−18.0	−32.1	3.9	156	8	18
Barrow	71.30 N	−12.5	−26.8	3.9	109	5	20
Fairbanks	64.80 N	−3.4	−23.9	15.4	287	23	47
Baker Lake	64.30 N	−12.3	−33.6	10.7	213	7	36
Anchorage	61.17 N	1.8	−10.9	13.9	373	20	47
Juneau	58.37 N	4.5	−3.8	12.9	1288	102	114
Edmonton	53.57 N	2.8	−14.7	17.5	447	25	83
Seattle	47.45 N	10.8	3.9	18.2	980	153	19
Montreal	45.50 N	7.2	−8.9	21.6	999	80	93
Des Moines	41.53 N	10.5	−5.2	25.2	789	32	77
Salt Lake City	40.78 N	10.7	−2.1	24.7	353	34	15
Washington, D.C.	38.85 N	13.9	2.7	25.7	1087	82	107
San Francisco	37.62 N	13.8	9.2	17.1	475	102	t
Nashville	36.12 N	15.6	4.4	26.8	1146	139	94
Los Angeles	33.93 N	18.0	13.2	22.8	373	78	t
Birmingham	33.57 N	17.8	8.1	27.6	1349	128	131
Phoenix	33.43 N	21.4	10.4	32.9	184	19	20
New Orleans	29.95 N	20.0	12.3	27.3	1369	98	171
Havana	23.17 N	24.6	21.8	27.0	1126	54	108
Acapulco	16.83 N	27.6	26.1	28.7	1401	8	230
Caracas	10.60 N	26.1	24.4	26.4	545	42	72
Guayaquil	02.18 S	25.5	26.5	24.2	811	199	0.3
Manaus	03.13 S	27.5	26.7	27.6	2294	279	65
Brazilia	15.78 S	20.4	21.2	18.0	1643	248	6
Rio de Janeiro	22.90 S	23.7	26.1	20.9	1218	211	52
Antofagasta	23.47 S	16.2	19.8	13.2	1.9	0.0	0.3
Santiago	33.45 S	14.4	21.2	8.1	264	0.1	69
Buenos Aires	34.58 S	16.9	23.7	10.6	1029	104	61
Puerto Aisen	45.50 S	9.2	13.9	3.9	3001	203	331
Comodoro Rivadavia	45.78 S	12.6	18.6	6.9	216	16	21
Punta Arenas	53.00 S	6.0	10.4	1.3	362	40	24
Melchior	64.32 S	−3.6	1.0	−9.3	1116	42	90
Byrd Station	80.02 S	−27.9	−14.6	−35.1	39	7	2.5
Amundsen Scott	90.00 S	−49.4	−28.7	−60.3	1.5	t	t

*Temperature is in degrees Celsius; $^\circ F = (^\circ C \times 1.8) + 32$. Precipitation is in millimeters; 1 mm = 0.04 in. t indicates trace of precipitation.

alignment of the Earth's orbit. The approximate sphericity of the Earth and the annual variation of the declination angle of the Sun cause the incoming solar radiation to be a function of latitude and season (**Fig. 3**). The annual mean insolation is largest at the Equator and decreases toward the poles. The seasonal variation is largest at high latitudes. At the poles 6 months of daylight alternate with 6 months of darkness. The Earth is closer to the Sun during summer in the Southern Hemisphere, so this region receives about 7% more insolation at this time than the Northern Hemisphere. *See* INSOLATION.

About half of the energy from the Sun that is incident at the top of the atmosphere is transmitted through the atmosphere and absorbed at the Earth's surface (**Fig. 4**). About 30% is reflected back directly to space, and another 20% is absorbed in the atmosphere. The fraction of the incoming solar radiation that is reflected to space is called the albedo. The albedo increases where clouds or surface ice and snow are present. It increases toward the poles because the amount of cloud cover and surface ice increases and because the Sun is closer to the horizon. The albedo of desert areas is generally higher than that of heavily vegetated areas or oceans. The solar energy that reaches the surface may be used to raise the surface temperature, or the energy can be used to evaporate water. The energy that is used to evaporate water (latent heat) is later released into the atmosphere when the water vapor condenses to form clouds and returns to the surface in the form of precipitation (**Fig. 5**). *See* ALBEDO; PRECIPITATION (METEOROLOGY).

Greenhouse effect. In order to achieve an energy balance, the solar energy that is absorbed by the Earth must be balanced by an equal amount of energy that is returned to space. The Earth emits radiative energy at frequencies that are substantially different from those of the Sun because of the Earth's lower temperature. The Earth emits primarily thermal infrared radiation [wavelengths from 4 to 200 micrometers], whereas most of the energy from the Sun arrives in the form of visible and near-infrared radiation (0.4–4 μm). The atmosphere is much less transparent to Earth's thermal radiation than to solar radiation, because water vapor, clouds, carbon dioxide, and other gases in the atmosphere absorb thermal radiation. Because the atmosphere prevents thermal radiation emitted from the surface from escaping to space, the surface temperature is warmer than it would be in the absence of the atmosphere. The combination of the relative transparency of the atmosphere to solar radiation and the blanketing effect of the gases in the atmosphere that absorb thermal radiation is often referred to as the greenhouse effect. The greenhouse effect on Earth is quite important. For example, the net emission of 21 units of

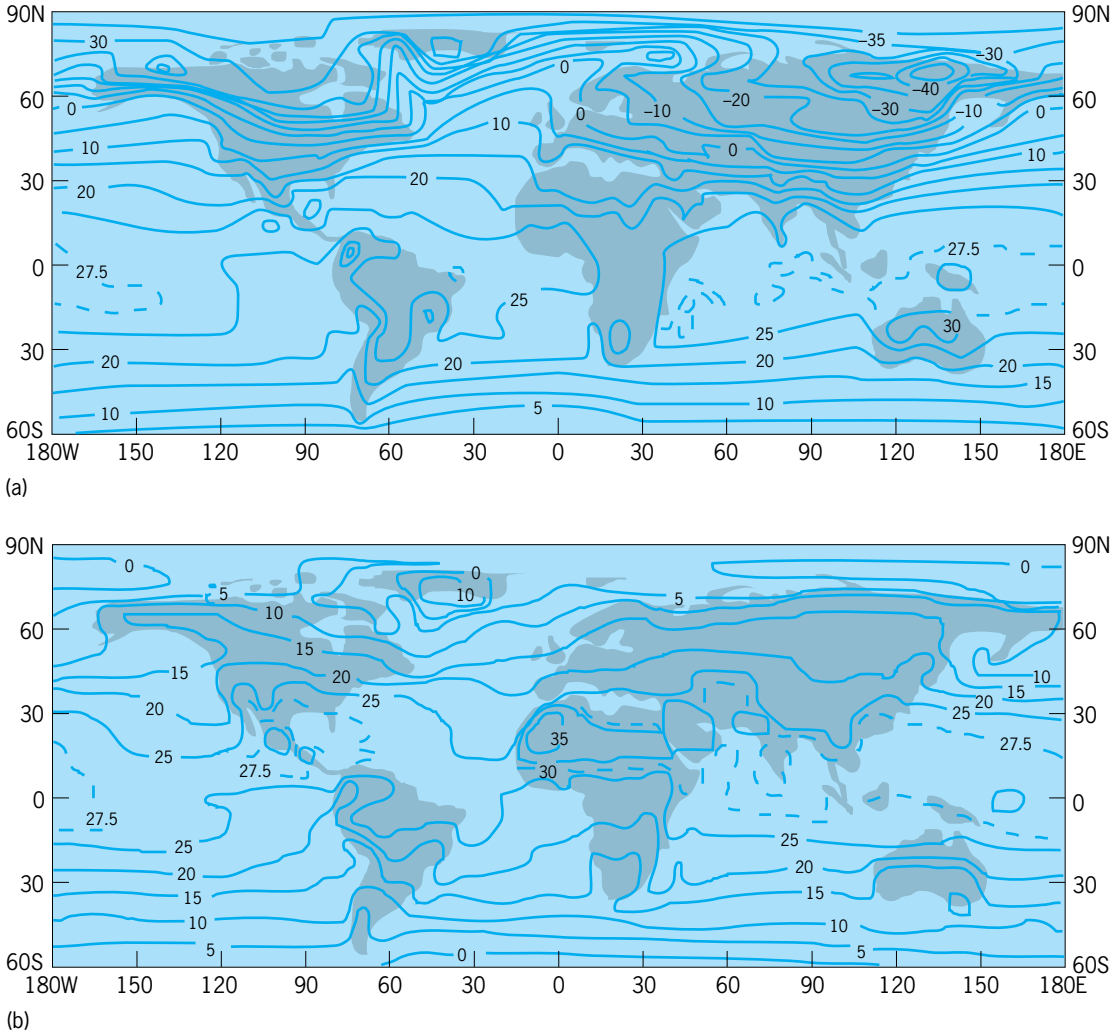

Fig. 1. Mean surface temperature (°C) for (*a*) January and (*b*) July. Note the cold temperatures in the centers of the large northern landmasses in January. The broken contour line (27.5) is included to give more information concerning the tropics. °F = (°C × 1.8) + 32. (*After D. L. Hartmann, Global Physical Climatology, Academic Press, 1994*)

infrared thermal radiation from the surface (Fig. 4) is the difference between a downward flux from the atmosphere of 89 units and a surface emission of 110 units. The downward flux of infrared radiation from the atmosphere is nearly double the solar radiation directly absorbed by the surface. *See* ATMOSPHERE; GREENHOUSE EFFECT.

Clouds affect the energy emission of the Earth.

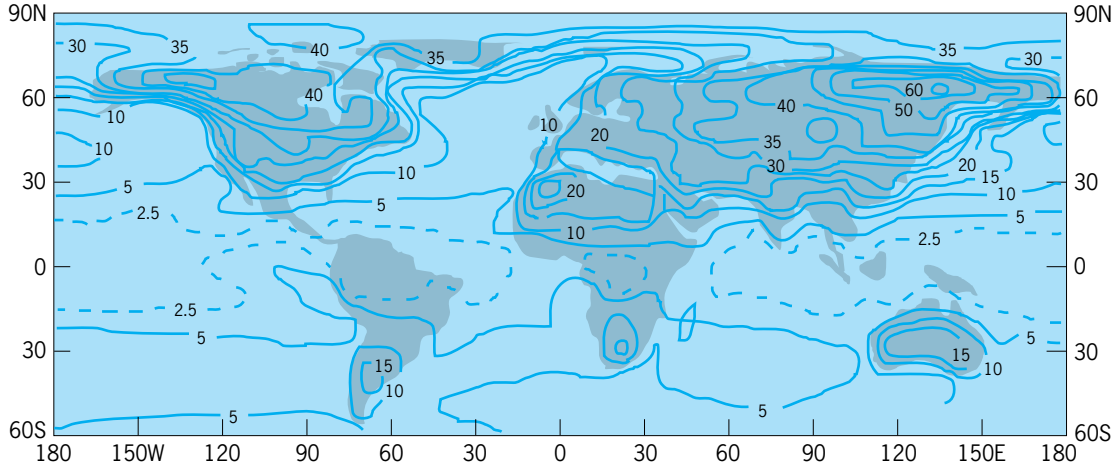

Fig. 2. Annual range of temperature (°C), the difference between the warmest and the coldest months. The broken contour line (27.5) is included to give more information concerning the tropics. °F = (°C × 1.8) + 32. (*After D. L. Hartmann, Global Physical Climatology, Academic Press, 1994*)

The thermal emission is low over tropical regions where high, cold clouds are present, such as the major precipitation regions of equatorial Africa, South America, southern Asia, Indonesia, and the band of clouds along the Equator. Because the albedo is higher, less solar energy is absorbed when clouds are present. Since clouds decrease both the solar energy absorbed and the terrestrial energy emitted to space, the net effect of tropical convective clouds on the energy balance of the Earth is less than their individual effects on solar and terrestrial energy. As a result, the distribution of the net radiative energy exchange at the top of the atmosphere does not clearly reflect the influence of the tall cloud so clearly. In general, the net radiation shows a strong influx of energy at those latitudes where the insolation is strongest. *See* TERRESTRIAL RADIATION.

A striking feature of the distribution of net radiation during July is the low values over the Sahara and Arabian deserts compared with the surrounding ocean and moist land areas. This is because desert sand has a relatively high albedo, so that less solar radiation is absorbed there. In addition, few clouds and little water vapor are present in the atmosphere over the deserts to absorb the thermal radiation emitted by the very hot surface. Since the emitted thermal radiation is high and the absorbed solar radiation is low, desert areas often show a net loss of radiative energy. This fact plays a central role in the maintenance of desert dryness. The loss of radiative energy is often balanced by convective transport of heat energy into the desert regions by the atmosphere. This results in downward motion of dry air, which flows outward near the ground and prevents the moist surface air of surrounding regions from reaching the desert interior. *See* DESERT.

Continentality. Annual temperature ranges are much greater near the center of continents in middle and high latitudes than they are over oceans. Land and ocean areas have very different seasonal variations because of their different capacities to store heat. Because the ocean is a fluid, its upper 150–300 ft (50–100 m) is generally in direct thermal communication with the surface. Therefore solar energy incident on the surface can be absorbed in this large heat reservoir without raising the surface temperature very much. In addition, over the oceans it is possible for solar energy to evaporate water and thus never be realized as heat at the surface. In contrast,

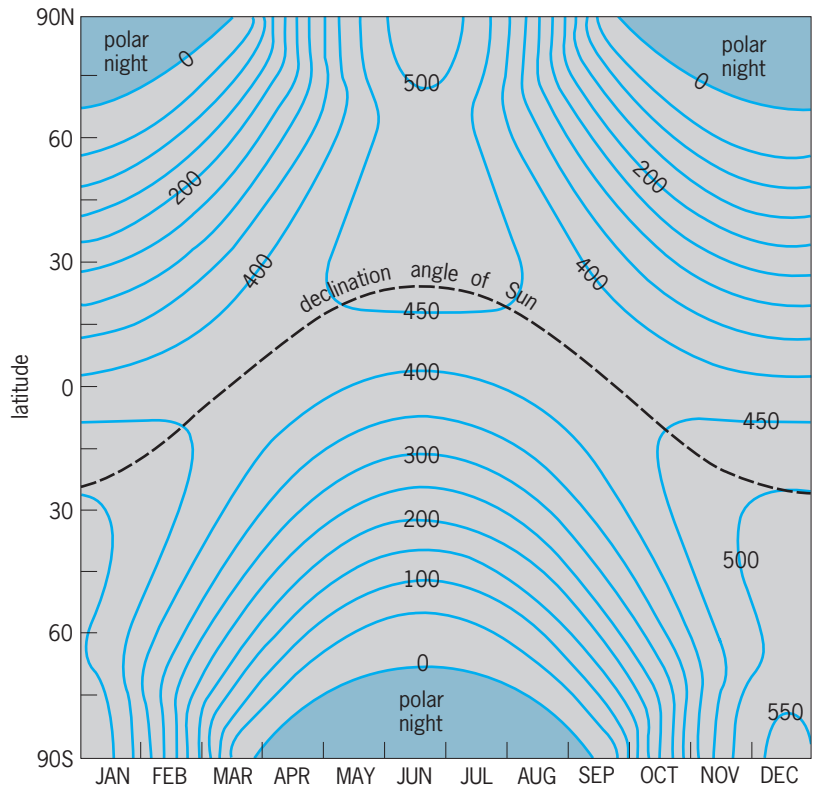

Fig. 3. Daily-average solar energy incident at the top of the atmosphere as a function of latitude and season in watts per square meter of surface area.

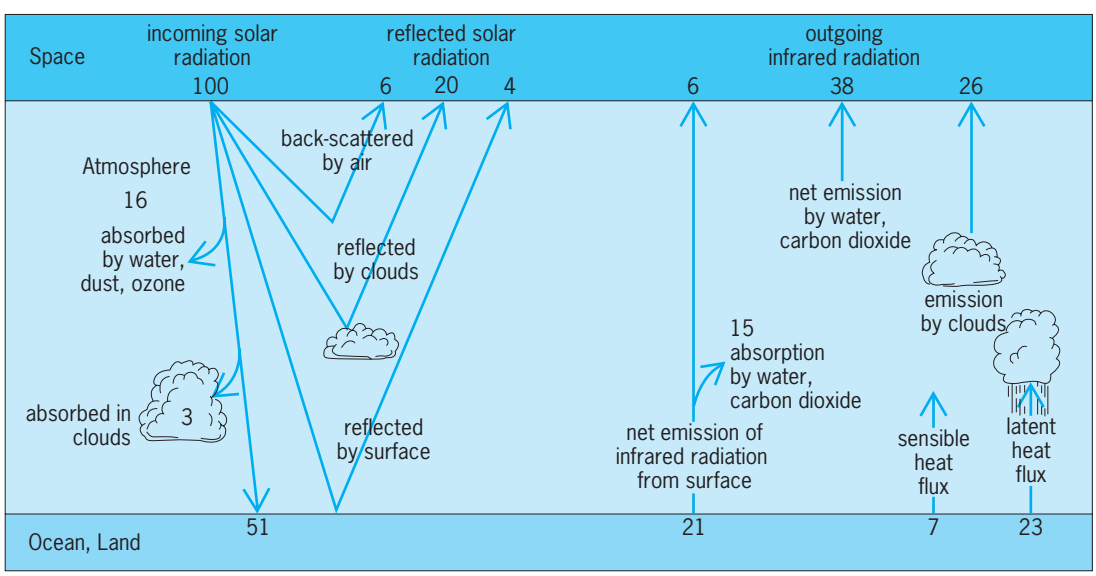

Fig. 4. Annual-mean global energy balance for the Earth–atmosphere system. Numbers given are percentages of the globally averaged solar energy incident upon the top of the atmosphere. (*After J. M. Wallace and P. V. Hobbs, Atmospheric Science: An Introductory Survey, Academic Press, 1977*)

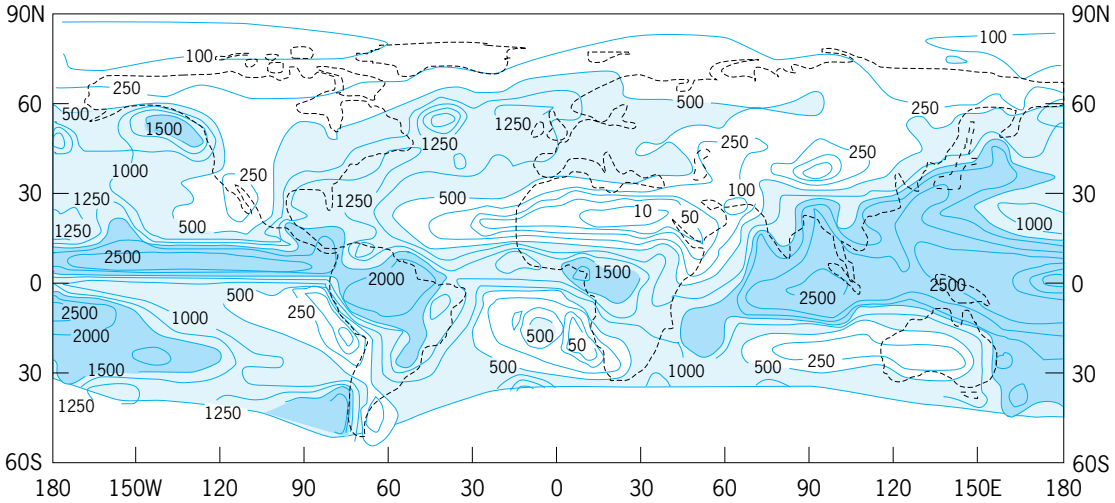

Fig. 5. Geographic distribution of annual mean precipitation in millimeters. (*After D. L. Hartmann, Global Physical Climatology, Academic Press, 1994*)

over land only about the first meter of soil is in thermal contact with the surface, and much less water is available for evaporation. During the winter, the ocean can return heat stored during the summer to the surface, keeping the surface relatively warm. The land cools off very quickly in winter, warms up rapidly in summer, and experiences large day-night differences. Continental climates in middle latitudes are characterized by hot summers and cold winters, whereas climates that are over or near the oceans are more equable with milder seasons (Figs. 1 and 2).

The distinction between maritime and continental climates can be seen by comparing the annual variations of temperature and precipitation for Tatoosh Island, Washington, with those for Minneapolis, Minnesota (**Fig. 6**). The temperature at Tatoosh Island

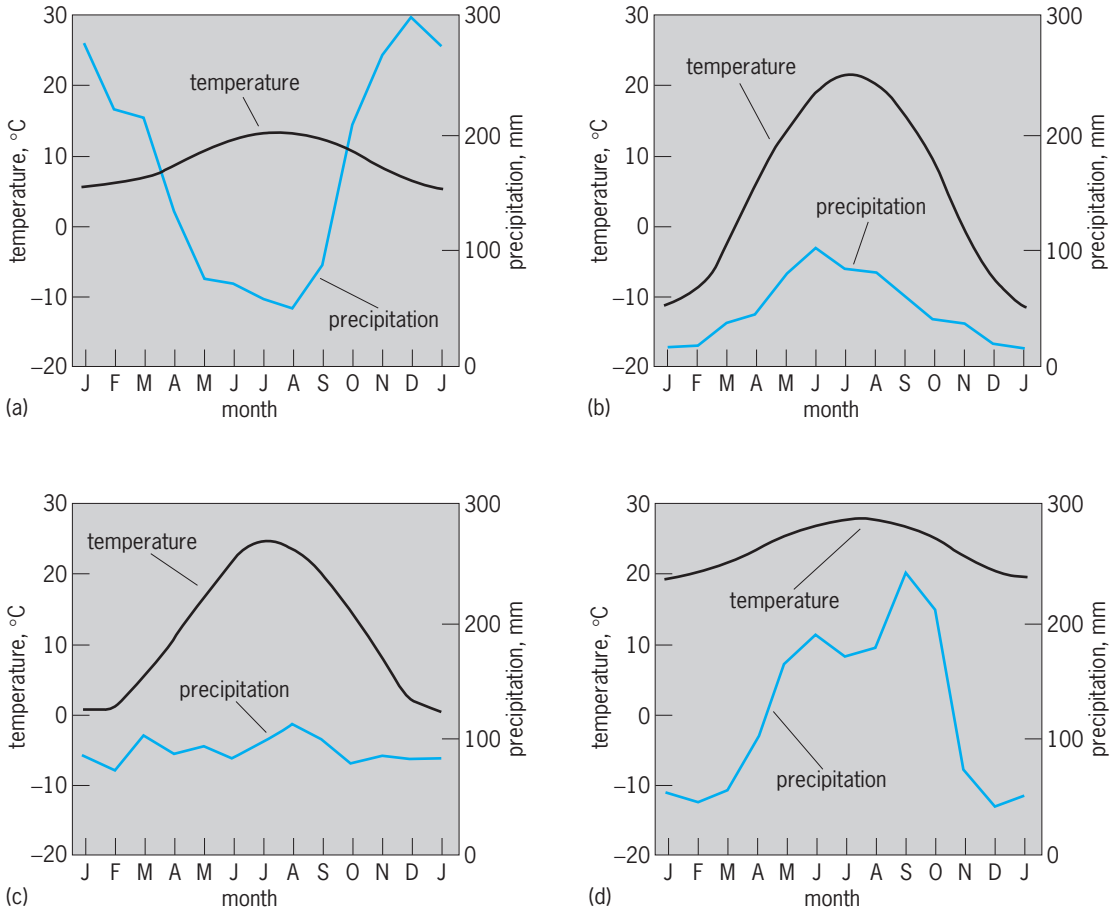

Fig. 6. Annual variation of monthly mean temperature and precipitation at (*a*) Tatoosh Island, Washington; (*b*) Minneapolis, Minnesota; (*c*) New York, New York; and (*d*) Miami, Florida. °F = (°C × 1.8) + 32; 1 mm = 0.04 in.

varies between a mean of 42°F (5.6°C) for January and a mean of 56°F (13.3°C) for August. This narrow temperature range results from the strong influence of the ocean, which is adjacent to and upwind of the Pacific coast, and has a small seasonal variation in temperature because of its large capacity to store and release heat. Although it is at nearly the same latitude, Minneapolis has a much larger annual variation of monthly mean temperatures, with readings of 12°F (−10.9°C) for January and 72°F (22.4°C) for July. Figures 1 and 2 show the large annual variation of temperature near the centers of the continents. *See* MARITIME METEOROLOGY.

The annual variation of precipitation is also very different in the two climate regimes (Fig. 6). At Tatoosh Island the precipitation peaks in the winter season in association with rainfall produced by the cyclones and fronts of wintertime weather. In Minneapolis the precipitation peaks in the summer season. Most of this precipitation is associated with thunderstorms. Adequate precipitation during the warm summer season is an essential ingredient of the agricultural productivity of the American Midwest.

New York City shows an annual variation of temperature and precipitation that is a combination of maritime and continental. It is near the Atlantic Ocean, but because the prevailing winds are out of the west, it also comes under the influence of air that has been over the continent. It has a fairly large annual variation of temperature, but monthly mean precipitation is almost constant through the year (Fig. 6*c*). It receives precipitation both from winter storms and summer thunderstorms. Miami, Florida is in the subtropics at 26°N. The annual variation of insolation is small, so that the seasonal variation of temperature is also modest. The precipitation shows a strong seasonal variation, with maximum precipitation during the summer half-year associated primarily with thunderstorm activity. *See* THUNDERSTORM.

General circulation of atmosphere and climate. Many aspects of the Earth's climate are influenced by the nature of the circulation that results from the radiative heating of the tropics and cooling of the polar regions. In the belt between the Equator and 30° latitude the bulk of the poleward atmospheric energy transport is carried by a large circulation cell, in which air rises in a narrow band near the Equator and sinks at tropical and subtropical latitudes. The upward motion near the Equator is associated with intense rainfall and wet climates, while the downward motion away from the Equator results in the suppression of rainfall and very dry climates (Fig. 5). Most of the world's great deserts, including the Sahara, Australian, Arabian, Kalahari, and Atacama, are in the belts between 10 and 30° latitude.

In middle latitudes the poleward flow of energy is produced by extratropical cyclones rather than by a mean circulation cell. These storms are thousands of miles wide and are characterized by poleward motion of warm, moist air and equatorward motion of cold, dry air. Most of the wintertime precipitation in middle latitudes is associated with weather disturbances of this type.

Surface features are also of importance in determining local climate. Mountain ranges can block the flow of moist air from the oceans, resulting in very low rainfall beyond the mountain ranges. The dryness of the Great Basin of North America and the Gobi Desert of Asia is maintained in this way. On the upwind side of mountain ranges, forced ascent of moist air can result in very moist climates; this occurs, for example, on the west slope of the coastal mountains of western America and the south slope of the Himalayas during the summer monsoon. The downwind sides of such mountains are often very dry, since the moisture precipitates out on the upwind side. Two locations that show the effect of topography on local climates are Puerto Aisén, Chile, and Comodoro Rivadavia, Argentina, which are both located near a latitude of 45°S (see table). Puerto Aisén is on the westward and upwind side of the Andes mountains in a deep valley that is exposed to the midlatitude westerly winds coming off the Pacific Ocean. It receives more than 10 ft (3 m) of precipitated water annually. Only a few hundred miles downwind of the Andes on the Atlantic seaboard, Comodoro Rivadavia receives only about 8 in. (0.2 m) of precipitation each year.

Complete climate system. The climate of the Earth results from complex interactions among externally applied parameters, like the distribution of insolation, and internal interactions among the atmosphere, the oceans, the ice, and the land (**Fig. 7**). The composition of the atmosphere, which plays a key role in determining the surface temperature through the greenhouse effect, has been radically changed by the life-forms that have developed, and continues to be modified and maintained by them. The atmosphere and the oceans exchange heat, momentum, water, and important constituent gases such as oxygen and carbon dioxide. The exchange of constituent gases is strongly influenced by life in the ocean. The hydrologic cycle of evaporation, cloud formation, and precipitation as rain or snow is intimately connected to the climate through the effects of water vapor, clouds, and surface ice on the radiation balance of the planet. Vegetation interacts strongly with the hydrologic cycle over land to determine the soil moisture, surface albedo, evaporation, precipitation, and surface water runoff. *See* HYDROLOGY.

Evolution of Earth's climate. The climate of the Earth is unique among the planets in the solar system. All of them evolved out of material in the rotating cloud from which the solar system was formed. The subsequent evolution of the planets' atmospheres depended critically on the mass of each planet and its distance from the Sun.

The mass of the Earth and its distance from the Sun are such that water can exist in liquid form rather than being frozen or escaping to space. The liquid water formed the oceans and led to the development of photosynthetic life, which reduced the carbon dioxide content and increased the molecular oxygen content of the atmosphere. Planets closer to the Sun receive more solar energy per unit area and are thus much hotter than the Earth. The surface of Venus is sufficiently hot that water vapor cannot

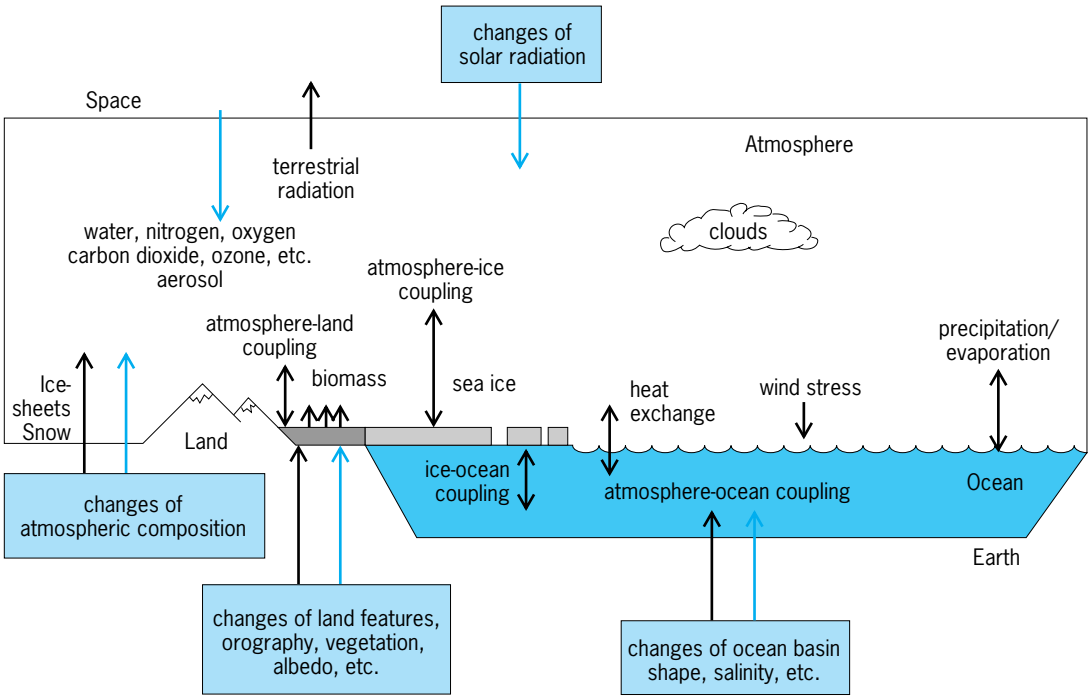

Fig. 7. Schematic illustration of the processes and interactions in the climate system. The colored arrows indicate externally applied conditions, and black arrows show internal processes that act to influence climate change. (*After J. G. Lockwood, Causes of Climate, Halstead Press, 1979*)

condense. Venus has a very thick atmosphere composed mostly of carbon dioxide. On Mars, which is farther from the Sun than the Earth and therefore colder, water freezes, leaving a very thin atmosphere of carbon dioxide. Thus the basic climatic conditions of a relatively circular orbit at a favorable distance for the Sun led to a drastically different evolutionary history for the Earth than for the neighboring planets. *See* ATMOSPHERE, EVOLUTION OF; SOLAR SYSTEM.

History of Earth's climate. Direct measurements allow the estimation of global mean surface temperature of the Earth for only about the last 100 years (**Fig. 8**). Global surface air temperature rose by about 0.9°F (0.5°C) between 1880 and 1940 and then declined slightly in the following three decades. The decade of the 1980s contained some of the warmest years on record.

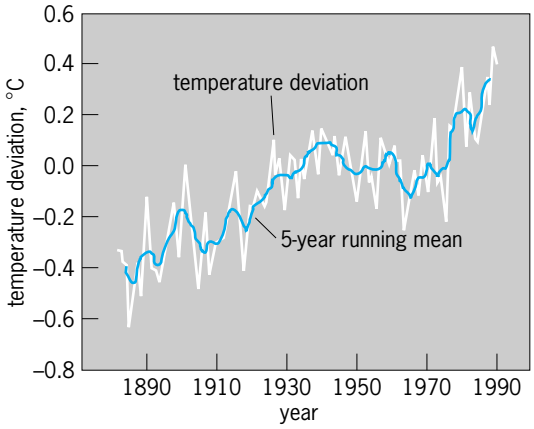

Fig. 8. Global surface air temperature change (°C) over the last century inferred from measurements. Temperature deviation is from 1951–1980 mean. °F = (°C × 1.8) + 32. (*After D. L. Hartmann, Global Physical Climatology, Academic Press, 1994*)

Evidence of climate variations in prehistoric times must be obtained from proxy indicators in dated sediments, such as pollen spores, the shells of small animals, or isotopic abundances, or from geologic features such as terminal moraines of glaciers or dry lake beds. A wealth of geological evidence exists to indicate that the Earth underwent a great glaciation as recently as 20,000 years ago. During this period, ice sheets nearly 2 mi (3 km) thick covered parts of North America and western Europe. Variations in the relative abundance of oxygen isotopes in deep-sea cores indicate that there has been a succession of periods of major glaciations separated by relatively warm, ice-free periods called interglacials during the last million years of Earth history (**Fig. 9**). *See* GLACIAL EPOCH; PALEOCLIMATOLOGY.

Analyses of time series show a relationship between global ice volume and known variations in the Earth's orbit. The parameters of the Earth's orbit include the eccentricity, which measures the departure from a circular orbit to a more elliptical one; the obliquity, which measures the tilt of the axis and therefore the magnitude of the annual variation in the declination angle of the Sun; and the longitude of perihelion, which measures the season at which the Earth makes its closest approach to the Sun. These parameters control the seasonal and latitudinal distribution of insolation. Because ice sheets form primarily on land and much of the land area is in northern latitudes, ice sheets form more

readily when the summertime insolation is relatively low, which allows some of the winter snow accumulation to survive the summer season. The global climate responds to the presence of the ice sheets through a process called ice-albedo feedback. Ice sheets are more reflective to solar radiation than other surfaces, so that their presence tends to reduce the amount of solar heat absorbed by the Earth. This in turn leads to a cooling of the planet and a greater tendency for the ice sheets to grow. *See* EARTH ROTATION AND ORBITAL MOTION.

Evidence from air bubbles trapped in glacial ice indicates that there were lower atmospheric concentrations of greenhouse gases such as carbon dioxide and methane during past glacial ages than there is at present. The concentration of carbon dioxide 20,000 years ago during the last glacial age was only about 190 parts per million by volume (ppmv), compared with the preindustrial level of 280 ppmv and the current concentration of 355 ppmv. During past glacial periods, the rather low values of carbon dioxide are thought to have been produced by changes in ocean chemistry and biology. Such low greenhouse gas concentrations contribute significantly to the cooling associated with glacial advances.

Future climate changes. Past relationships between the Earth's orbital parameters and global ice volume indicate that the Earth will undergo another major glacial age within about 25,000 years from now. Currently, however, the composition of the Earth's atmosphere is changing rapidly as a result of human influences. Of climatic interest are industrial gases that are transparent to solar radiation entering the Earth's atmosphere, but are opaque to the thermal radiation emitted by the Earth. Gases with these characteristics are called greenhouse gases. In order to achieve an energy balance in the presence of increased concentration of greenhouse gases, the Earth's surface must warm, so that its thermal emission will increase to offset increasing downward emission of thermal energy from an atmosphere that is becoming more effective at intercepting outgoing thermal radiation, if other factors remain the same. The principal greenhouse gases are carbon dioxide, which is increasing principally because of the burning of coal and oil; methane or natural gas, whose increase is related to agriculture and coal and gas mining; nitrous oxide, which is a product of the decay of fixed nitrogen in plants or artificial fertilizer; and chlorofluorocarbons, which are gases used in industry for refrigeration, foam-blowing, and cleaning, and as aerosol propellants. Carbon dioxide is expected to contribute most of the increase in atmospheric thermal opacity due by the year 2030.

Within the next 100 years, if current trends continue, climate models indicate that human activities will result in a climate that is 4–8°F (2.2–4.4°C) warmer than it is today. The magnitude of this increase is similar to that between the present climate and the last glacial age, and would represent an extremely rapid warming by the standards of natural climate variability and change. Studies also suggest increased probability of drought during the summer growing season in midlatitude agricultural areas and rising sea levels that may be 27 in. (70 cm) higher than present levels by the year 2080. *See* CLIMATE HISTORY; CLIMATE MODIFICATION; DROUGHT; MESOMETEOROLOGY; METEOROLOGY; MICROMETEOROLOGY.

Dennis L. Hartmann

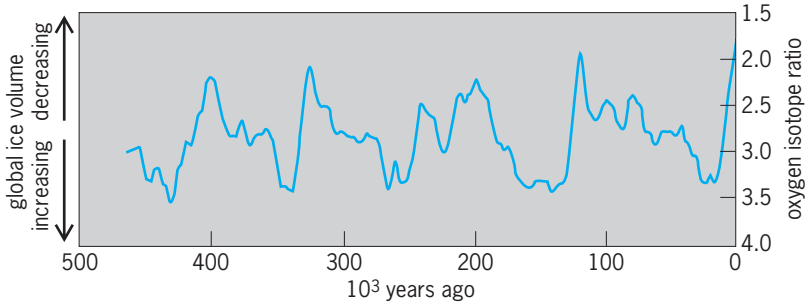

Fig. 9. Climate of the past half-million years. Oxygen isotope ratios are from ocean sediments, which reflect variations in the global volume of ice in glaclers and ice sheets (*After J. Imbrie, Ice Ages: Solving the Mystery, Enslow Publishers, 1979*)

Bibliography. R. S. Bradley, *Paleoclimatology*, International Geophysics Series, vol. 64, 2d ed., 1999; W. J. Burroughs, *The Climate Revealed*, 1999; M. H. Glantz, *Currents of Change: El Nino's Impact on Climate and Society*, 1993; T. E. Graedel and P. Crutzen, *Atmosphere, Climate, and Change*, Scientific American Library Paperback, no. 55, 1997; D. L. Hartmann, *Global Physical Climatology*, 1994; K. McGuffie and A. Henderson-Sellers, *A Climate Modelling Primer: Research and Developments in Climate and Climatology*, 2d ed., 1997; P. J. Robinson and A. Henderson-Sellers, *Contemporary Climatology*, 2d ed., 1999; K. E. Trenberth (ed.), *Climate System Modeling*, 1993.

Clinical immunology

A branch of clinical pathology concerned with the role of the immune defense system in disease. The subject encompasses diseases where a malfunction of the immune system itself is the basic cause, together with diseases where some external agent is the initiating factor but an excessive response by the immune system produces the actual tissue damage. It also extends to the monitoring of the normal immune response in infectious diseases and to the use of immunological techniques in disease diagnosis. *See* CLINICAL PATHOLOGY.

Clinical immunology is a specialization born of the rapid expansion of knowledge in basic immunology together with the recognition that many diseases have an immunological component. Some areas of medicine which might logically be classified as clinical immunology were developed before immunology emerged as a separate discipline, and these have become the province of other medical departments. For example, vaccination against infectious disease is part of the practice of pediatrics, and asthma is treated by respiratory physicians.

Immune defense system. Many features of the immune system make it prone to shift from protecting the body to damaging it. This complex system not only must distinguish between the body's own cells and a foreign invader but must also recognize and eliminate the body's own cells if they are damaged or infected with a virus. The recognition receptors used to make this fine distinction between "self" and "not self" are not encoded in the genes. Rather, they are assembled following random rearrangement of information carried in small gene segments. During their development, immune system cells are subjected to a selective process, those bearing potentially useful receptors being preserved while those bearing dangerous, self-reactive receptors are eliminated. This process is closely balanced, and some potentially self-reactive cells often persist.

The cells that carry these highly discriminating surface receptors are the lymphocytes. They are migratory cells, concentrated in lymph nodes and also found among the white cells of blood. Lymphocytes are of two types: B cells (which are formed in bone marrow) and T cells (which are formed in the thymus). CD8 T cells on activation directly kill the body cells presenting the foreign antigens. CD4 T cells on activation produce short-range factors (cytokines) that recruit and activate macrophages and natural killer cells, which in turn release their own cytokines. These less specific cells are also the first line of defense early in infection. If this complex immune response system is not well controlled, it can cause substantial damage to normal cells besides eliminating abnormal cells and foreign invaders. *See* CELLULAR IMMUNOLOGY; CYTOKINE.

Immunodeficiency diseases. Immune deficiencies may be acquired as a consequence of an external factor such as an infection or may be inherited as a genetic defect. Immunodeficient individuals are subject to recurrent infections by organisms that are not normally a hazard. Some immunodeficiencies affect only one part of the immune system. In Di George's syndrome, T lymphocytes are deficient because of a developmental failure of the thymus. Hypogammaglobulinemia results from a developmental failure of B lymphocytes. Chronic granulomatosis disease results from an enzyme deficiency in phagocytic cells. *See* IMMUNOLOGICAL DEFICIENCY.

AIDS. Acquired immunodeficiency syndrome (AIDS) is a worldwide epidemic. It is caused by infection with the human immunodeficiency virus (HIV), transmitted sexually or through contaminated blood transfusion or by use of contaminated needles for injections. In the late stages of the disease, there is a massive loss of CD4 T cells, resulting in susceptibility to multiple infections. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS).

Autoimmune diseases. Many diseases have an autoimmune component, that is, the central cause of the disease, or much of the pathology, is due to the action of the immune system on the body's own tissues. The factor triggering self-reactivity is usually unknown; it may be a viral infection. In some diseases there is a genetic predisposition according to the human cell surface antigen (HLA) type of the individual.

In some autoimmune diseases one particular organ is targeted, as in thyroiditis, where antibodies react with thyroid hormones; or in juvenile onset diabetes, where the insulin-producing cells of the pancreas are attacked. In others, a widespread but particular cell type is the target. One example is autoimmune hemolytic anemia, where antibodies initiate complement lysis of red cells. Finally, some complex diseases with multiple targets are the result of autoimmune phenomena. For example, systemic lupus erythematosus can affect most organs of the body, especially kidney, skin, and joints, and is associated with antibodies against many intracellular components, including deoxyribonucleic acid (DNA). Rheumatoid arthritis is another multisystem disease, but synovial joints bear the brunt of the inflammatory response, and an antibody against the body's own immunoglobulin (rheumatoid factor) is often present. *See* ANTIBODY; AUTOIMMUNITY; IMMUNOGLOBULIN.

Hypersensitivity or allergic states. These represent an extreme response of particular individuals to environmental antigens that are not intrinsically harmful. Examples are allergies to drugs, food components, pollen, or house-dust mites. The nature of the response varies, but the most common response involves immunoglobulin E antibodies. When complexed with the antigen, these bind to the surface of mast cells, triggering the release of histamine and other pharmacological agents. This is the basis of atopic diseases, including dermatitis, allergic rhinitis (hay fever), bronchial asthma, and even extreme anaphalactic shock. *See* ALLERGY.

Cancer. The body mounts an immune response against at least some forms of cancer. It is not yet clear whether it is possible to generate an effective immune response against all types of cancer cells. Some forms of cancer represent the uncontrolled proliferation of component cells of the immune system itself. These include myelogenous leukemia (phagocytic cells), lymphatic leukemia (B lymphocytes), and multiple myeloma (antibody-producing cells). *See* CANCER (MEDICINE).

Transplantation. The response of the immune system to cells of another individual was first apparent in blood transfusion, where a match for the A, B, O, and other blood groups was needed to avoid a reaction. The transplantation of skin or internal organs presents the more difficult task of matching the complex HLA antigens. These major histocompatibility complex proteins, whose normal function is to present antigenic fragments to T cells, can trigger an especially vigorous T-cell response. The best available matching for HLA type, as well as careful immunosuppression, is needed for graft survival. *See* IMMUNOLOGY; TRANSPLANTATION BIOLOGY.

Therapy. There are several approaches to suppressing excessive immune reactivity. Drugs with a general toxicity for dividing cells have been used, although now more selective and better-tolerated drugs such as cyclosporin may be employed. The most widely used immunosuppressive agents are

corticosteroids, which affect many body systems but have a particularly strong action on lymphoid cells. Desensitization, or modifying the nature of the response by injecting small amounts of the foreign antigen, is sometimes used to treat allergic states. *See* IMMUNOSUPPRESSION.

There are few therapies for enhancing immune responses. Bone marrow transplantation is used to restore the immune system in some immunodeficiency diseases. Passive transfer of preformed antibody protects against some infections, and transfusion of immunoglobulin is used to treat immunoglobulin deficiencies. However, vaccination or immunization, the introduction of modified forms of an infectious agent in order to develop specific protection against infectious diseases, is one of the most effective of all medical procedures. *See* IMMUNOLOGICAL DEFICIENCY; VACCINATION.

The clinic. Many large hospitals have separate clinical immunology departments. However, clinical immunology also is part of the practice of other specialists, such as endocrinologists who treat individuals with thyroiditis or diabetes, or hematologists who treat individuals with pernicious anemia. In addition, common immunological diseases such as allergies or rheumatoid arthritis are often handled in specialized clinics. Individuals in a pediatric clinical immunology department are typically children with immunodeficiency diseases or leukemias. Hematologists and oncologists may be involved in their treatment. The adult patients of a clinical immunology department are usually a diverse group, ranging from hospitalized patients requiring continuous care (for severe allergic responses, late-stage AIDS, infection following immunosuppression) to outpatients whose progress needs regular but not continuous monitoring (for immunoglobulin deficiencies being treated by immunoglobulin transfusion, autoimmune diseases being treated with immunosuppressive drugs).

Clinical immunologists take blood samples as a convenient indicator of the status of the immune system. These samples are analyzed in the laboratory. Other immunological tests use the skin as an indicator of the immune response, most often to test for hypersensitive or allergic states. Test materials are placed in contact with the skin or are injected subcutaneously. A local red inflammatory reaction indicates a positive response. A rapid response, that is, within hours, is referred to as immediate hypersensitivity; this indicates the presence of preformed antibody reactive with the test material. A slow response, within days (delayed hypersensitivity), indicates the development of an immune response involving recruitment of lymphocytes. *See* HYPERSENSITIVITY.

Laboratory tests. The specialized clinical immunology laboratory within a hospital is often directed by the same clinician who sees the patients in the clinic, with trained technicians performing most of the actual tests. Most tests utilize blood samples, although other body fluids and tissue samples are occasionally analyzed. Tests on blood may be divided into those that analyze the serum and those that analyze blood cells. The concentration, type, and specificity of antibodies in serum may be measured by using tests that range from simple agglutination assays, through immunoelectrophoresis, to more sophisticated enzyme-linked immunosorbent assays (ELISA). The number of lymphocytes in the blood may simply be counted, or they may be analyzed or separated to test their function in culture. A preliminary separation of lymphocytes from other blood cells may be made by centrifugation procedures. Further separation into subclasses of lymphocytes is usually effected by a combination of immunofluorescence and flow cytometry. The essential reagents are fluorescent-dye conjugated antibodies specific for individual cell surface molecules; that is, after interaction with these reagents, only particular cells acquire a fluorescent label. The cells are then passed across a laser beam, one at a time in a stream, and their fluorescence intensity is measured. In fluorescence-activated cell sorting, the stream of cells is broken into droplets, and droplets containing particular types of cells are deflected and collected. Immunofluorescence may also be used to examine tissue sections under a microscope. The results of these tests are used to help diagnose a disease or to evaluate the effectiveness of a treatment. *See* IMMUNOASSAY; IMMUNOFLUORESCENCE.

Ken Shortman

Bibliography. P. J. Lachman and D. K. Peters (eds.), *Clinical Aspects of Immunology*, 1982; I. Roitt, *Essential Immunology*, 1991.

Clinical microbiology

The adaptation of microbiological techniques to the study of the etiological agents of infectious disease. Some of the most significant scientific advances during the latter half of the nineteenth century and the first few decades of the twentieth century were in medical bacteriology. From the work of Louis Pasteur, Robert Koch, Theobald Smith, and others, it was shown that bacteria were responsible for many diseases of humans. During the early part of the twentieth century, the technology for the growth of bacteria was advanced to the extent that pure cultures of bacteria could be studied from defined environments. Not only did bacteria become models for the study of their own biochemical characteristics, but the same technology enabled bacteriologists to isolate and identify bacteria from patients with infectious disease. This spurred the discovery of most of the bacterial diseases that are currently known.

The role of fungi, protozoa, rickettsias, and viruses in human disease was also elucidated during the first half of the twentieth century. The inhibitory power of materials from a fungus called *Penicillium* was discovered in 1928 by Alexander Fleming. His discovery of penicillin and the contributions of many others marked a great era in the treatment of bacterial diseases. The discovery of penicillin was soon followed by that of streptomycin and tetracycline.

For the first time, bacteriologists could detect and identify bacteria and then test in the laboratory the ability of various antibiotics to kill these isolated bacteria. Such was the birth of contemporary clinical microbiology. *See* ANTIBIOTIC.

Clinical microbiologists determine the nature of infectious disease and test the ability of various antibiotics to inhibit or kill the isolated microorganisms. A contemporary clinical microbiologist is also responsible for a wide range of microscopic and cultural studies in mycology, parasitology, and virology. The consultative skill of the clinical microbiologist is sought by many clinicians. The clinical microbiologist is often the most competent person available to determine the nature and extent of hospital-acquired infections, as well as public-health problems that affect both the hospital and the community. The clinical microbiologist is a valuable resource to the hospital and community. As clinical microbiology encompasses all aspects of infectious disease, the various specialties will be discussed individually. *See* EPIDEMIOLOGY; HOSPITAL INFECTIONS.

Bacteriology. Historically, the diagnosis of bacterial disease has been the primary job of clinical microbiology laboratories. Many of the common ailments of humans are bacterial in nature, such as streptococcal sore throat, diphtheria, and pneumococcal pneumonia. Only rarely, however, in recent decades have new bacterial diseases been described. A prime example is the disease legionellosis caused by a hitherto unknown bacterium, *Legionella pneumophila*. *See* LEGIONNAIRES' DISEASE.

The bacteriology laboratory plays a major role in the diagnosis of bacterial diseases. It accepts specimens of body fluids, such as sputum, urine, blood, and respiratory or genital secretions, and inoculates the specimens onto various solid and liquid growth media. Following incubation at body temperature, the microbiologist examines these agar plates and tubes and makes a determination as to the relative numbers of organisms growing from the specimen and their importance in the disease process. The microbiologist then identifies these alleged causes of disease and determines their pattern of antibiotic susceptibility to a few chosen agents.

Clinical microbiologists also microscopically examine these body fluids. They report on the presence of bacteria in body fluids and the cellular response to infection, such as the numbers of white blood cells observed in the specimen. Sometimes the kinds of white blood cells observed can provide hints for the identity of the disease. Such microscopic evaluations can provide a rapid aid in the diagnosis of infectious disease, as well as an assessment of the quality of the specimen. For example, a specimen called "pus" which does not contain white blood cells may not be pus. *See* MEDICAL BACTERIOLOGY.

Mycology. Mycology is the study of fungi. The applications of mycology in the clinical microbiology laboratory include the isolation of fungi which are known to cause disease. In general, diseases caused by fungi are relatively harmless but annoying, such as ringworm and athlete's foot. However, they can be severe, systemic disorders such as blastomycosis, coccidioidomycosis, and histoplasmosis. Fungi can also cause disease in those people whose immunologic processes have been suppressed by drugs or radiation. The diagnosis of fungal infection differs from bacterial infection because of the relatively slower rate of fungal growth. In general, diagnosis of fungal disease is a function of not only isolating and identifying the causative organism and determining which of three or four antifungal antibiotics might be effective, but also determining the patient's antibody response to the disease. *See* MEDICAL MYCOLOGY.

Parasitology. Parasitic diseases are widespread and of worldwide importance to humans. Such diseases include amebic dysentery, malaria, and Chagas' disease. The clinical microbiologist working in parasitology examines specimens of feces and blood for the presence of parasitic agents. These examinations are primarily microscopic, and the agents are identified on the basis of their characteristic structure. Although antiparasitic agents are available for the treatment of such infection, the parasitologist is usually not involved in the determination of the relative effectiveness of these agents on parasites. *See* MEDICAL PARASITOLOGY.

Virology. Virology is a specialized part of clinical microbiology. Because of the nature of viruses and their proclivity for growth only in living cell cultures or in eggs, few clinical microbiology laboratories have facilities for the isolation and identification of viruses. Most of these procedures are best done in large hospitals or in state departments of health. There is a growing trend, however, for smaller hospitals to develop screening virology laboratories in which specimens are brought to the laboratory, placed in tissue culture, and preliminary identification steps taken in order to provide a rapid diagnosis of viral disease. Although only a few drugs exist that can be used to treat virus infections, the early recognition of viral disease may allow the clinician to discontinue the use of powerful antibiotics that are effective only for bacterial diseases. Common virus diseases include influenza, herpes, chicken pox, the common cold, and many nonspecific respiratory infections. Much of the viral diagnosis that is done in the laboratory is by the determination of a specific antibody response to viral disease. The routine clinical microbiology laboratory usually does not become involved in such procedures. A special procedure for the detection of viruses and other organisms that affect infants is called the TORCH screen. It is a procedure for the detection of antibodies to the parasitic disease toxoplasmosis and the viral diseases rubella, cytomegalovirus, and herpes. These diseases may cause devastating health problems in babies. *See* ANIMAL VIRUS; DISEASE; TISSUE CULTURE; VIRUS.

Richard C. Tilton

Technology. While direct microscopy and culture continue to be methodological mainstays in diagnostic microbiology laboratories, the nonculture methods are growing in the variety of applications and the sophistication of the technology.

Antibodies. Polyclonal antibodies raised in animals such as mice, sheep, goats, and rabbits, and monoclonal antibodies produced by hybridization technology are used to detect bacteria, fungi, parasites, or virus-infected cells by using direct or indirect fluorescent techniques. Additional methods include latex agglutination tests to detect particulate antigens and enzyme immunoassays to detect soluble antigens. Some of these technologies can also be used to detect antibodies rather than antigens, usually using the antigen to capture antibody that is in serum. *See* IMMUNOASSAY; MONOCLONAL ANTIBODIES.

Instrumentation. Instrumentation for identification of bacteria and yeastlike fungi and for antimicrobial susceptibility testing includes a variety of technologies. The systems use either microtiter plates or plastic cards containing dried or frozen substrates or antimicrobial agents. A known concentration of the organism to be tested is introduced into the wells, and the plate is incubated until the instrument reader detects sufficient growth. The pattern of growth and reactivity is compared to patterns of known organisms. The likelihood that the test organism is one of them is reported from information maintained in a database. Because urinary tract infections in outpatients are usually caused by a single organism, some of these systems can be directly inoculated with urine. Generally, however, a pure culture must first be obtained by overnight incubation of inoculated agar plates.

Nonamplified hybridization. Probes for deoxyribonucleic acid (DNA) or messenger ribonucleic acid (mRNA) are available for various applications. There are two advantages in testing for mRNA. (1) It is present only if the gene it represents is expressed (contributing to the phenotypic characteristics of the organism). (2) If the gene is expressed, there will be multiple copies of the mRNA, thus increasing the analytical sensitivity of the test. A drawback is that ubiquitous ribonucleases can render a false negative test. Probes are used for direct detection of organisms in clinical material and for culture confirmation.

Amplification. Further increases in analytical sensitivity have been achieved by nucleic acid amplification techniques. Several methods are currently available to clinical laboratories.

In polymerase chain reaction (PCR), double-stranded DNA is denatured; oligonucleotide probes bind to homologous strands of single-stranded DNA, and the enzyme polymerase extends the probes using deoxyribonucleotides in the milieu. Over 30 cycles, the number of copies increases a millionfold. This product is then detected.

In ligase chain reaction (LCR), the enzyme ligase can fill the 3-nucleotide gap between two probes that attach to homologous, target, single-stranded DNA. By cycling in a manner similar to PCR, a million joined copies can be generated in a short period of time, after which they are detected.

In nucleic acid sequence–based amplification (NASBA), reverse transcriptase is used to make double-stranded complementary DNA (cDNA) and the target RNA is digested by ribonuclease H. Another enzyme, T_7 polymerase, is used to make more RNA target from which more cDNA is made.

Molecular typing. Analysis of lipopolysaccharides and proteins by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) and cellular fatty acid analysis by gas-liquid chromatography have given way to nucleic acid–based methods. Restriction enzymes, which cut DNA at a constant position within a specific recognition site usually composed of four to six base pairs, are used to cut chromosomal DNA; the resulting fragments are compared by pulse field gel electrophoresis (PFGE) or ribotyping. Electrophoresis of isolated plasmid DNA is another method for comparing organisms. DNA sequencing can also compare segments of the DNA of organisms from the same genus and species.

Chip arrays. DNA chip or microarray technology is expected to have a greater effect on medicine than either DNA sequencing or PCR. Over 30,000 small cDNA clones of expressed fragments of individual genes (expressed sequence tags, or ESTs) are spotted onto a thumbnail-sized glass chip. Fluorescein-labeled genomic or cDNA from the sample being evaluated is passed over the chip to allow hybridization. A laser measures the fluorescent emissions and a computer analyzes the data. As an example, relevant cDNA fragments from the bacteria, viruses, and protozoa that are known to cause diarrhea or dysentery could be placed on one chip and a stool specimen could be analyzed for all known pathogens at one time. Genes coding for known antimicrobial resistance mechanisms could also be represented so that an appropriate therapeutic agent could be prescribed. *See* CLINICAL PATHOLOGY. Carol A. Spiegel

Bibliography. B. J. Howard et al. (eds.), *Clinical and Pathogenic Microbiology*, 1987; H. D. Isenberg, The future of clinical microbiology, *Lab. Med.*, 19:321–323, 1988; E. W. Koneman, *Color Atlas and Textbook of Diagnostic Microbiology*, 5th ed., 1997; K. M. Kurian, C. J. Watson, and A. H. Wyllie, DNA chip technology, *J. Pathol.*, 187:267–271, 1999; P. R. Murray et al. (eds.), *Manual of Clinical Microbiology*, 7th ed., 1999; D. H. Persing et al. (eds.), *Diagnostic Molecular Microbiology: Principles and Applications*, 1993; F. C. Tenoven, Deoxyribonucleic acid probes for infectious diseases, *Clin. Microbiol. Rev.*, 1:82–101, 1988; W. A. Volk et al., *Essentials of Medical Microbiology*, 3d ed., 1986.

Clinical pathology

The discipline in which disease processes are assessed and monitored by the examination of blood, body fluids, secretions, and aspirated specimens. The detection of chemical, morphological, microbiological, and immunological abnormalities in this material is reported as laboratory results. In contrast, anatomical pathology is that branch of pathology dealing with the nature of a disease process through studying the gross and microscopic features of tissues and fluids removed either at postmortem examination (autopsy) or during surgery. However,

divisions between anatomical and clinical pathology are somewhat artificial and largely embedded in how pathology evolved in academic practice. Information derived from autopsy examinations, surgical specimens, or clinical laboratory testing provides physicians with the means to confirm a final diagnosis and to evaluate nonsurgical conditions. Such information is used by physicians in diagnosis and in therapy management.

History and Laboratory Administration

Early clinical pathology laboratories were simple, often confined to a single room (**Fig. 1**). A limited number of tests were performed by hand, with visual observation of color changes or turbidity. Microscopes, often monocular, were used primarily to identify abnormal cells in stained smears or blood films, inflammatory cells and atypical crystals in urinary sediment, or foreign particles such as bacteria in direct mounts of clinical specimens. A group of biochemists, working primarily in hospitals in Boston and New York before the outbreak of World War I, soon recognized that quantitative results must be derived from clinical laboratory tests to detect and monitor underlying diseases. Crude instruments such as manometers to measure blood gases, colorimeters to quantify chemical reactions, and fluorometers on which levels of blood electrolytes could be quantified were introduced between the two world wars. During this time, all reagent samples and reagents were dispersed using mouth pipetting, and it was not uncommon to observe lit cigarettes and open caches of food on the workbench, practices that later were strictly forbidden.

Clinical pathology laboratory testing experienced its first transition following WWII. From the late 1950s into the early 1960s, the number of new tests introduced into clinical laboratories multiplied logarithmically, accompanied by ever-increasing sophistication of instrumentation and advances in laboratory organization and administration. A number of new self-operating analytical systems were introduced to replace the laborious and time-consuming hand-operated processes characteristic of early laboratories. Through the advancement of technology, new tests, new procedures, and new reporting systems were made accessible to medical practitioners and their patients. On the other hand, the massive numbers and complexity of tests also placed burdens on those who use this information in patient care. How to determine which tests meet the standards for evidence-based medicine is an ongoing challenge. Testing was made more creditable by instrument testing. The subjective testing by multiple technologists could be standardized to better inform practicing physicians.

In keeping with the continuing increase in the number of laboratory tests and the complexity resulting from the introduction of automation and molecular techniques, nontechnical matters of the organization, administration, and staffing of clinical pathology laboratories has changed dramatically as well. Typically, a laboratory director must oversee the overall operation of the laboratory and relate to physicians and other medical personnel in choosing appropriate diagnostic tests, interpreting laboratory data, and monitoring therapeutic responses. The director also assumes many administrative duties relating to hospital settings and financial affairs. The day-to-day analysis and quality control of test procedures, the review of the accuracy and precision of test results, and the supervision of laboratory personnel are the several activities commonly carried out by clinical scientists and medical technologists with advanced degrees. Materials obtained from patients are commonly assayed by highly trained laboratory scientists and medical technologists. Laboratory assistants and clerical personnel assist in processing specimens and in transcribing, recording, and logging in laboratory test results. Although computer-based

Fig. 1. Clinical pathology laboratory at Christus Santa Rosa Hospital, San Antonio, Texas, 1923.

laboratory information systems are now indispensable, intervention by laboratory personnel is critical to ensure that samples are procured from the correct patient at the correct time and delivered to the correct section in the laboratory.

Role in patient care. Clinical pathology laboratories function as an integral part of the diagnostic cycle beginning as an individual with signs and symptoms suggestive of disease consults a physician.

Specimen collection and processing. After obtaining a clinical history and examining the patient, a physician may order laboratory studies on blood, urine, other body fluids and secretions, or tissue biopsies to establish or confirm a working diagnosis. Specific protocols are followed in the preparation of the patient and in the techniques in collecting a specimen. Specimen collection must be made into a proper container with preservatives or anticoagulants appropriate to the test being requested. Particular attention to proper packaging of the specimen is critical, especially if the delivery time is delayed by transport through long distances.

A visual and/or microscopic examination may be the first evaluation of specimen quality. If unsatisfactory (not meeting a variety of predetermined criteria), particularly misidentifications, the specimen may be rejected with a request for repeat collection. Patient information and test requests are bar-coded into the laboratory information system. The processing and analysis stages follow.

Processing may require centrifugation to separate the cells if the test requested requires an analysis of plasma or serum. Commercially prepared reagent systems and automated or nonautomated electronic instruments have replaced manual testing. Using direct-read instruments and/or special chemical and immunological reagent packages, test results may be obtained within a few minutes. In highly automated laboratories, test results may be transferred directly from the instrument into the hospital computer system, providing clinicians with direct access to results from terminals present on the wards.

Point-of-care testing. On-site or point-of-care laboratory testing has evolved as an alternative testing arrangement to meet the increasing demand for rapid turnaround time of certain test results. Emergency rooms, outpatient clinics, and physician offices, as well as select special-care hospital settings, are particularly served. Rapid-response read-out devices, packaged reaction cassettes that include positive and negative controls, and reagents to perform a variety of "spot" tests are currently available. The license to perform an increasing variety of simple or "waived" tests has been afforded through regulations provided under the Clinical Laboratory Improvement Amendment of 1988 (CLIA-1988).

The menu of waived point-of-care laboratory procedures provides for the immediate assessment of a number of diseases and conditions. Methods for enumerating blood cells, performing a urinalysis, measuring abnormal levels of chemical constituents such as glucose and cholesterol, determining early pregnancy, and detecting several bacterial and viral infections are currently available. Many of these tests can be performed on a small quantity of blood, 25–50 microliters obtained by a simple finger stick precluding the need for a vena puncture. Again, each of these areas of point-of-care testing must be carefully monitored for accuracy and reproducibility of test results under strict administrative supervision and careful quality control. When all such parameters are in place, the availability of test results within 5 min or less provides for the immediate implementation of an appropriate therapeutic regimen.

Quality control. The clinical laboratory generates data that are converted into information that the physician may use to supplement his clinical knowledge of the patient. Laboratory directors and personnel are obligated to have in place systems of quality maintenance for all of its practices in supplying laboratory test information. Established testing standards must be applied to the assay. In practice, a control sample containing a known amount of the substance being tested is treated in exactly the same way as the patient sample.

This known sample interjection concept of quality control was developed from industrial models used in the 1950s. Included within the run of each batch of specimens is a pair of control samples that contain a known concentration of analyte, commonly set at both high- and a low-level readings. The control results must come within a fixed standard deviation before the specimen results can be approved. Thus, observing a drift in the results of the control samples provides a means to detect reagent deterioration, operator errors, or technical problems that may have occurred during analysis. All specimen test results must be checked against the quality control results before a report is issued. Each analytical instrument must be calibrated and paced through test runs, and all storage compartments and heat-controlled devices must be checked daily for temperature maintenance. All quality control data must be posted on daily log charts that are clearly visible and in usable formats.

Laboratory inspection and accreditation. The College of American Pathologists (CAP) in 1961 launched a laboratory inspection program focused initially on hospital laboratories. Standards were developed for technical performance. Meeting these standards, directed toward continuous improvement of laboratory testing in health care, is rewarded by accreditation. The thrust of the quality-improvement efforts remains educational, and the inspections currently are carried out by volunteer pathologists, Ph.D. scientists, and medical technologists. Through collaborative efforts of the CAP with regulatory agencies, such as the state public heath departments, the FDA, the American Association of Blood Banks, the American Association of Clinical Chemists, and the Joint Commission on Hospital Accreditation, among others, a highly sophisticated and integrated system of ongoing laboratory surveillance and regulation has evolved, leading to high-quality performance.

Proficiency standards are monitored by requiring participating laboratories to analyze, in 3- or 4-month cycles, a set of proficiency test samples that are indistinguishable from those obtained from a patient. The analytical results derived from these proficiency samples are then surveyed by independent observers. When aberrations are spotted, corrective actions must be immediately put in place.

Evidence-based medicine. Resource constraints combined with demands by physicians and the public to have the best evidence about medical testing have led to new approaches to medicine in the clinical laboratory. Clinical epidemiology, also known as evidence-based medicine, has grown in importance as an approach to medical treatment as well as testing and diagnosis. Evidence-based medicine focuses on objective measurements of diagnostic and prognostic accuracy and therapeutic effectiveness, and the implementation of best practices based upon those measurements.

Clinical Laboratory Subsections

Evolution of laboratory testing and its complexities has directed the laboratory's functioning into separate sections and disciplines. Although this division varies from laboratory to laboratory, the following sections and subsections are common:

- Chemical pathology
  - General chemistry
  - Spirometry (measurement of lung breathing capacity and blood gases)
  - Toxicology, in particular the monitoring of drugs of abuse
  - Radiometry (measuring minute amounts of radiant energy) and enzyme immunoassay (determination of binding between antigens and antibodies)
  - Medical microscopy/urinalysis
  - Electrophoresis
- Hematology
  - Morphologic hematology, involving glass-slide preparation and microscopy
  - Chemical hematology for the detection of hemoglobinopathies
  - Blood clotting and dissolution (fibrinolysis)
  - Flow cytometry (fluorescence microscopics) for the determination of cell markers
- Immunohematology
  - Transfusion medicine (blood bank)
  - Donor recruitment and service
  - Transfusion preparations (cross matching)
  - Component separation (red cells, white cells, platelets, plasma)
  - Plasmapheresis (separating out directly from the patient certain blood components)
- Microbiology
  - General aerobic bacteriology
  - Anaerobic bacteriology: bacteria that grow only in the absence of oxygen
  - Mycobacteriology: the agents of tuberculosis
  - Mycology: the agents of fungal infections
  - Parasitology: the animal parasites causing disease in humans
  - Virology: minute microorganisms having only a core of nucleic acid and a protein envelope
  - Immunology and serology: the diagnosis of disease by detection of antigens and antibodies
  - Antimicrobial susceptibility testing: determination of the best drug to use in the treatment of infections

Chemical pathology. Chemical pathology, also known as clinical chemistry or clinical biochemistry, encompasses determination of the presence or quantity of chemicals present in blood, serum, plasma, urine, cerebrospinal fluid, or other body fluids.

The age of automation, which has become so much a part of the modern chemical pathology laboratory, began in the late 1950s with the Technicon Autoanalyzer, possessing the following features:

1. A pipetting robot that procured small volumes of sample from tiny cups placed in a rotating carousel.
2. Streams of liquid passed through a series of plastic tubes connected to reagent baths.
3. A peristaltic proportioning pump under which the reagent tubes passed, propelling a measured amount of sample and reagents in air-segmented aliquots, coming together in a mixing coil housed in a constant-temperature water bath.
4. A photometric flow-through cell through which a light beam passed, producing proportional electrical currents from a photovoltaic cell that activated a recording needle that traced quantitative peak-and-valley line grafts on a roll of moving paper.

Upon this foundation has evolved ever-increasing innovations in robotic, electronic, and computer-based technology. Automated instruments now find service in several subsections of the chemical laboratory: blood gas analysis, toxicology and therapeutic drug monitoring, core chemistry, immunoassay, electrophoresis, and urinalysis. The ultimate extension of this technology is the sequential "streaming" of instruments currently utilized in larger reference laboratories, where the specimen is advanced along an automatic hands-off treadmill assembly, passing through a series of stations including centrifugation, sampling, analysis, and computer-based reporting (**Fig. 2**).

Core chemistry. In the core chemistry section of the laboratory are evaluated many different chemicals in blood samples, including electrolytes such as sodium, potassium, calcium, and chloride ions. Quick measurement of these ions often permits doctors to avoid delays in making accurate diagnostic decisions. Other important molecules are also quantified in this laboratory section; for example, creatinine and urea nitrogen are widely employed to evaluate kidney function.

Blood gas analysis. The blood gas analysis section is found in almost all chemical pathology laboratories. Included among the chemicals analyzed are gases that are dissolved or bound to other molecules, ions,

Fig. 2. Overview of a 2005 automated chemical analysis system. Specimen racks are placed on a movable track that moves sequentially through sampling, read-out, and reporting stations in the course of multitest analyses.

small metabolic molecules, and proteins. Establishing the blood pH as well as the quantity of oxygen and carbon dioxide permits physicians to determine rapidly the acid–base status of patients. From these test results a quick diagnosis may be made on patients with drug poisoning, kidney malfunction, and diabetic ketoacidosis—an acute illness in some diabetics.

Toxicology. In the toxicology and therapeutic drug monitoring sections of the laboratory, suspected patient overdose and poisoning are investigated through the analysis of blood, urine, or other samples for the presence of drugs or drug metabolites. Early identification of the cause of patient overdose permits physicians to begin therapy quickly and often makes the difference in patient outcome. Medication drug levels are also evaluated in this section of the laboratory. Physicians are thus able to tailor the medication and dosage to provide the best therapy for patients. *See* TOXICOLOGY.

Immunochemistry. Larger molecules such as proteins are often measured using immunoassays in the immunochemistry section. Immunoassays use antibodies to detect and quantify the specific antibody target. For example, ferritin, a molecule used to evaluate iron deficiency, is typically measured using antiferritin antibodies. Patient antibodies directed against viruses or bacteria such as hepatitis C, human immunodeficiency virus (HIV), cytomegalovirus, and *Helicobacter pylori* are measured in this section of the laboratory. *See* IMMUNOCHEMISTRY.

Electrophoresis. Electrophoresis is the process through which large molecules are separated by exposing them to an electric field through a solid porous medium, such as agarose or polyacrylamide. Such separation can be visualized with certain dyes and evaluated either visually or by instrument densitometry. Serum protein electrophoresis (SPEP) allows pathologists to identify patterns to detect or monitor patients with multiple myeloma. Electrophoresis is also commonly used to assay urine (urine protein electrophoresis, UPEP) and cerebrospinal fluid. The diagnosis of multiple sclerosis is supported by certain findings seen after the electrophoresis of cerebrospinal fluid. *See* ELECTROPHORESIS.

Urinalysis. The urinalysis section of the laboratory is dedicated to measuring chemicals present in urine as well as the microscopic analysis of solid-formed elements in the urine. Such formed elements include red blood cells, white blood cells, crystals, casts, and bacteria. For example, abnormalities detected in the urinalysis laboratory are often some of the earliest signs of kidney failure. *See* URINALYSIS.

Hematology. Blood is an easily accessible source of human material for analysis. Several blood cell types circulate in the serum and become available for study after blood is drawn (phlebotomy) from a patient's vein or artery.

Microscopy. Blood cells may be scrutinized directly by smearing a drop of blood on a glass slide, applying a stain, and examining the cells under the

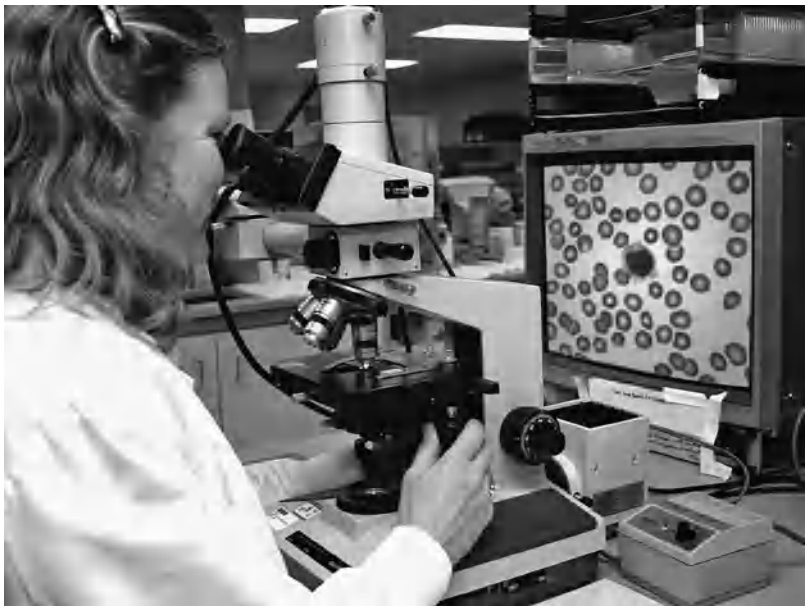

Fig. 3. Microscopic examination of a stained peripheral blood smear. The field of view used for cell identification observed by the microscopist is projected on the monitor screen located behind the microscope.

microscope. Valuable information is there before the eyes of the microscopist, and key observations may point immediately to a diagnosis or direct the physician to additional testing (**Fig. 3**). *See* MICROSCOPE.

Automatic counting. Examinations of the blood are facilitated by the use of automated counting instruments. Among these is the Coulter Counter, introduced in the 1960s. The major features of its operation include:

1. The aspiration of a small sample of blood by negative vacuum pressure into a fine-bore tube that channels the test specimen into the reading chamber of the instrument.
2. The tube is of such small diameter that the individual blood cells are partitioned out in single file.
3. Each cell in turn passes through a reading device, by which its diameter, color, nuclear density, and light transmission properties are measured.
4. This information is electronically captured to produce a multidot graph called a histogram by which the several cell types can be separated and enumerated.
5. Either red cell or differential white cell counts can be automatically generated depending on how the sample was processed before aspiration into the instrument.

Thus, cellular components are separated into red cells (blood), white cells (granulocytes, lymphocytes, and monocytes), and platelets. These different cell types are categorized and sorted electronically, resulting in computer-generated reports. Problems in relationship of cells or cells out of the range of size, shape, and content are flagged for review by the technologist or hematologist. An extension of this technology is the three-part and five-part automated differential counters by which the percentage of these white cells is displayed and printed out on a multicolored histogram, eliminating manual cell counting.

Molecular targeting. As with most testing procedures in the clinical laboratory, technology in hematology has advanced far beyond the counting and visual categorization of cells. Rapid developments in genomics and proteomics, as applied to hematologic disorders, have supplanted the simple hands-on, glass-slide examination of blood cells and test-tube shaking for the observation of clotting, practices that now have limited usefulness in defining a clinical problem. Basic knowledge in molecular and cellular biology, molecular immunology, and genetics have been translated into specific diagnostic and therapeutic testing methods. In further similar developments, clinicians look forward to these same reactions of specific molecular targeting for therapy of hematologic and coagulation disorders. *See* BLOOD; HEMATOLOGIC DISORDERS.

Flow cytometry. In a sense, flow cytometers are specialized fluorescence microscopes attached to a computer that translates signals to data. Specific bands of fluorescence are measured through a series of optic beam splitters and filters, as a flow of single cells passes through a light source. Through this method cells, such as blood cells, are analyzed and sorted. Physical characteristics of cells such as size, shape, and internal complexity can be measured. Any component or fraction that can be detectable by fluorescence is opened for examination. Fluorescence technology with high speeds (1000 cells per second or more) combined with the capability of correlating quantitative dots from simultaneous measurements on each cell has led to widespread use in the biological and medical fields.

In medicine, flow cytometry is used to identify lymphocyte subsets in conditions of immunodeficiency, and in immunophenotyping for leukemia and lymphoma. One common use is immunophenotyping of peripheral blood, lymph node, and bone marrow samples to establish the presence of clonality or leukemia. A wide range of cell surface antigens are present on hematopoetic cells. Flow cytometry is geared to their identification and allows the sorting of these cells into normal or abnormal expression patterns. *See* LEUKEMIA.

Microbiology. The implementation of automation in the clinical microbiology laboratory progressed more slowly than in other sections of the clinical laboratory. The tried and true conventional culture recovery methods of Robert Koch and making bacterial identifications by visually observing color reactions in reaction tubes are still used to varying degrees. As a prelude to automation, several laboratory supply companies introduced packaged kits in the 1960s, in which a series of dehydrated biochemical substrates were contained within miniaturized cupules embedded in a plastic strip. First was the Roche Enterotube, a pencil-shaped plastic tube containing 12 chambers with slants of differential culture media. Then followed the API 20E strip, containing 20 cupules containing dehydrated substrates, which is still used today. The color reactions within

the reagent chambers of each of these systems is read visually, from which a biocode number linked to a bacterial identification is derived. *See* MEDICAL BACTERIOLOGY.

Automated bacterial identification. The impetus for the automation of bacterial identifications came in the mid-1970s as part of the NASA space program. The challenge was to design a small instrument that would be capable of identifying bacterial species encountered in space flights, particularly from environmental samples collected from the Moon. The outcome was the AMS bacterial identification system, since marketed under the name VITEK® (bioMerieux, Inc). The key feature of this system is a multiwell reagent card that, after inoculation of a bacterial suspension, is inserted into the incubation chamber of the instrument. The results of 30 biochemical reactions are provided by the instrument, available as soon as 4 h for biochemically active bacterial species. Competitive systems are also available for use in microbiology laboratories. Each of these widely used systems also has cards for the determination of antimicrobial susceptibility test results, replacing in many laboratories the time-honored method of measuring the zones of inhibition surrounding antibiotic-impregnated filter paper disks (**Fig. 4**).

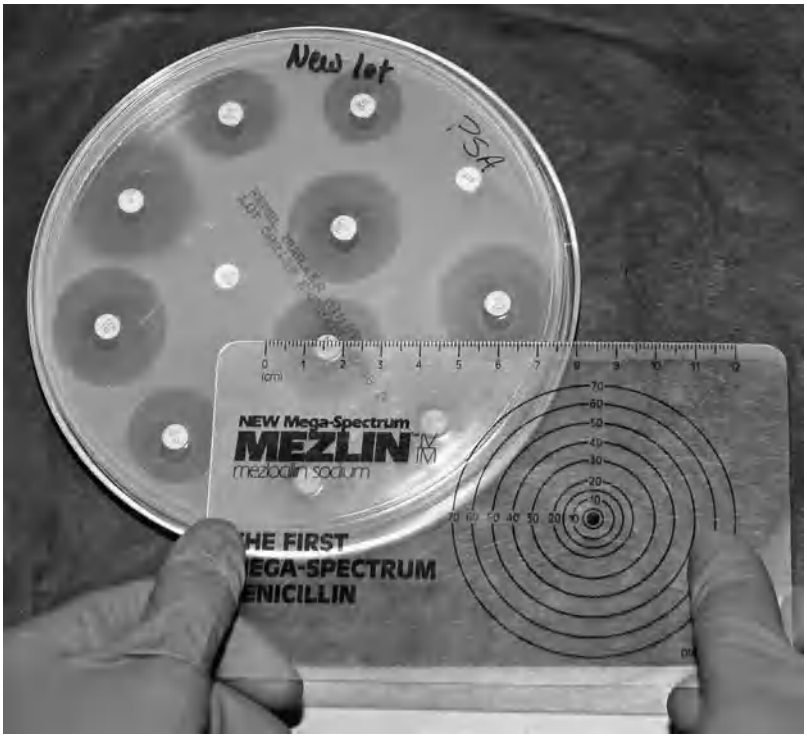

Fig. 4. Disk diffusion antimicrobial susceptibility plate illustrating assessment of antibiotic action by measuring zones of inhibition surrounding antibiotic-impregnated filter paper disks.

Continuous-read blood culture systems. The recent introduction of automated continuous-read blood culture systems has been a major breakthrough in providing the more timely detection of bloodborne infections. Previously, blood cultures were prepared by injecting a given volume of venous blood into bottles containing an enriched culture medium. The inoculated bottles were placed into a 35°C incubator. After 24–48 h of incubation, each bottle was visually observed against a strong transmitted light source for the detection of turbidity as a signal for bacterial growth. Subcultures were made of any turbid bottle to identify the organism.

In continuous-read systems, culture bottles are placed in a holding chamber within the instrument. Located in the bottom of each bottle is a pH-sensitive plastic disk that undergoes a color change when acid is produced by growing bacteria in the culture broth. As a threshold color change is achieved, the system automatically produces both a visual and an audio "alert." The distinct advantage of this system, in addition to being more labor-saving, is the detection of positive blood cultures in as short a time as 4 h after collection.

Immunoassay. The introduction of immunoassays in the late 1960s has impacted in microbiology as well. Using a technology developed by the 1975 Nobel Prize winners George Köhler and César Milstein, monoclonal antibodies are now produced by hybridizing "immortal" antibody-secreting cancer cells (myeloma cells) from mice with a genetically equivalent cell producing a specific antibody. These specific antibodies can be used to detect the complementary antigens on bacterial cell walls, by which species identifications can often be directly made without the need for culture. Immunoassay can also be used to detect cell-surface antigens using monoclonal antibodies labeled with fluorescent dyes. Using this technology, more than 140 different antigens and adhesion molecules have been demonstrated on the surfaces of bacteria, white blood cells, platelets, and lymphocytes. *See* IMMUNOASSAY.

Blood banking and transfusion medicine. The blood bank and the transfusion service are sections of the laboratory dedicated to preparing, storing, and dispensing blood products as well as testing those who donate blood and those who will receive blood products. One hundred years ago, transfusions posed a significant risk to both the donor and recipient. Advances in laboratory testing for blood antigens and infectious disease has made the blood supply and the transfusion process relatively safe. Blood products are critical for treating trauma patients in the emergency department, surgical patients in the operating room, and patients with bleeding or clotting disorders.

Processing of blood for transfusion begins with the blood donor. Most blood donors are individuals who voluntarily give about $^1/_2$ liter of blood in a procurement procedure lasting about 20–30 min. Other donors selectively donate platelets, plasma, or red blood cells, using apheresis technology, a process that typically requires between 45 and 90 min. Patient donors are screened for many infectious diseases to maximize the safety of the blood supply. After a donation, patient blood is processed into components such as platelet concentrate, fresh frozen plasma, red blood cell concentrate, or cryoprecipitate. These components are then stored at the appropriate temperature until used.

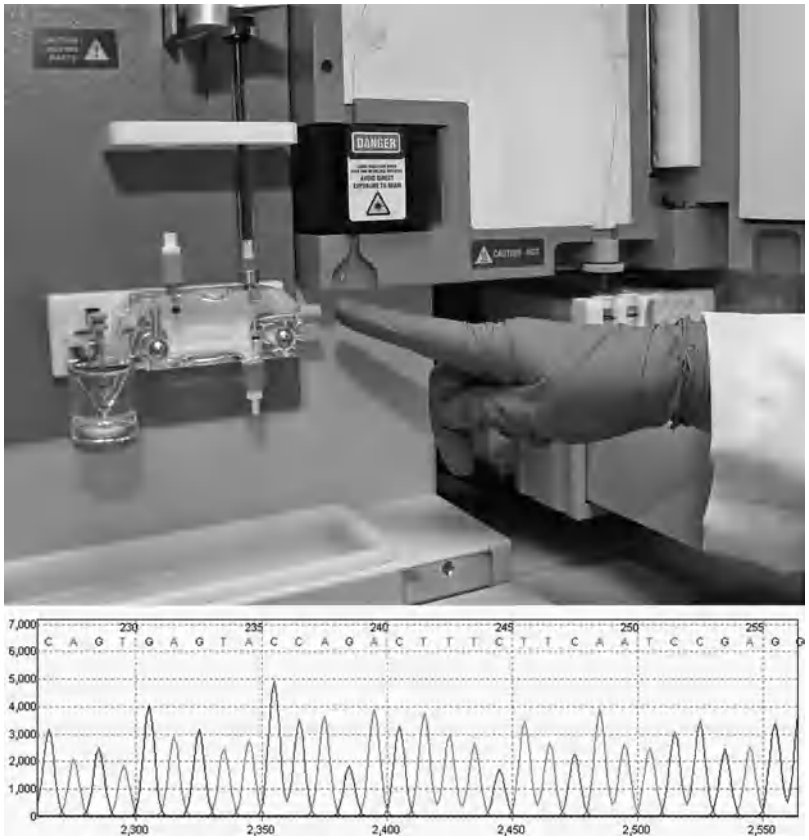

Fig. 5. Capillary electrophoresis is used for DNA sequencing. Data are processed by a computer which then converts the findings into a graphic display. This information is converted to identification of mutations, deletions, and insertions relating to human disease.

Prospective blood recipients are tested to determine their blood type and to detect the presence of antibodies and blood antigens. Using blood type and antibody screening information, blood products are selected that minimize the probability of any adverse transfusion reaction between donor and recipient blood. Many uses for blood products have evolved. For example, red blood cell concentrates are used to treat severe bleeding and anemia. Platelet concentrates are used to treat patients who are bleeding or at risk of bleeding due to lack of platelets. Fresh frozen plasma is used to replenish clotting factors in those with deficiencies or to avert severe bleeding.

The clinical pathologist is trained in blood banking and transfusion medicine and provides advice to other physicians about the best products to use in treating individual patients. In addition, the clinical pathologist oversees the collection of donor blood and the testing of recipients. The clinical pathologist also uses apheresis technology to treat patients. For example, apheresis is often employed to remove harmful antibodies or cells from patients with a wide variety of diseases such as Guillain-Barre syndrome or thrombotic thrombocytopenic purpura. *See* TRANSFUSION.

Recent Advances

A transition in the clinical pathology laboratory began in the early 1990s and is still progressing—the introduction of molecular biology and genetics.

Molecular and genetic biology. This technology, based on the direct amplification and analysis of DNA from a variety of biological cells, facilitates the diagnosis of diseases associated with defects and alterations in genes, including many hereditary diseases and cancers known to have a genetic association. These techniques have also allowed the direct and rapid identification of a variety of microorganisms that cause infectious diseases. For example, the detection of certain DNA sequences in *Mycobacterium tuberculosis*, the cause of tuberculosis, extracted directly either from infected body secretions or from bacterial cells recovered in culture, can lead to a diagnosis within hours, in contrast to the 2 to 3 weeks previously required. As part of forensic examinations, human beings, both alive and postmortem, can be identified based on direct molecular analysis of shed cells, aspirated body fluids, and/or autopsy specimens. The working parts of a nucleic acid sequencing instrument with the resultant printout of the graph depicting the peaks of the individual nucleotides are illustrated in **Fig. 5**. An analysis of the individual presence and sequence of these nucleotides provides information about gene defects and hereditary diseases. *See* DEOXYRIBONUCLEIC ACID (DNA); GENE.

Molecular pathology-laboratory. The molecular pathology laboratory is one of the newest additions to clinical laboratories. Its creation follows advances in molecular biology over the past 30 years such as the polymerase chain reaction, nucleic acid hybridization, and DNA sequencing. The term "molecular pathology" may be a bit misleading since all other laboratory areas also measure molecules in some form. The term, however, is reserved at present for the measurement of molecules using the most advanced technologies. Molecular pathology laboratories include tests for the detection of DNA fragments related to infectious disease, certain types of cancer, and genetic disease (**Fig. 6**).

Most cases of chlamydia and gonorrhea are diagnosed using nucleic acid testing for the organism responsible for each disease. Other organisms such as

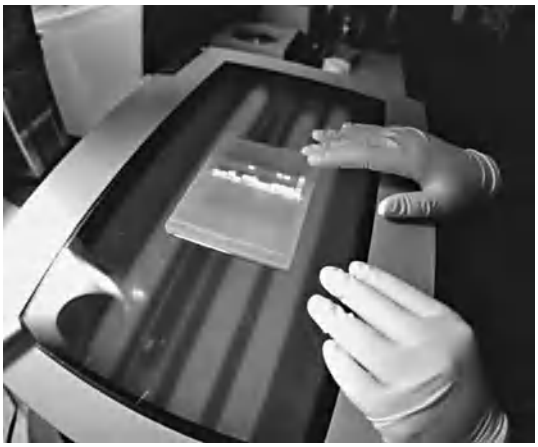

Fig. 6. Examination of an agarose gel with an ultraviolet light source to identify diagnostic fragments of DNA (orange-yellow fluorescent bands in the gel).

human papilloma virus (HPV) are also detected using DNA probes or polymerase chain reaction technology. In addition to detection of infectious diseases, treatment for patients with hepatitis C is monitored increasingly with viral load testing, in which the RNA from the virus is measured as treatment progresses. Patients with HIV infections are monitored the same way. In the workup of hematologic malignancies, the type of malignant cells can be detected by molecular analysis.

Molecular pathology techniques such as immunoglobulin and T-cell receptor gene rearrangements are often employed to help in the diagnosis of a specific type of leukemia or lymphoma. In addition, several types of cancer have characteristic chromosomal translocations. These translocations are usually detected using cytogenetics; however, in some cases DNA-based techniques are used to detect translocations of chromosomal material.

In addition to diagnosis and monitoring the therapy of infectious diseases and cancer, molecular pathology is often the only method by which genetic diseases may be detected prior to the development of symptoms. It is also the most powerful method for identifying patients who are carriers of genetic mutations. Several tests for genetic mutation analysis are commonly used for identifying Huntington's disease, factor V Leiden mutation, muscular dystrophy, and many other conditions. *See* HUMAN GENETICS; MUTATION.

The molecular pathology laboratory is likely to see expansion in both scale and scope. For example, methods to quickly measure large numbers of genes, messenger RNAs, proteins, and small metabolic molecules simultaneously have been developed. These advancements have led to the formation of several new disciplines: genomics, transcriptomics, proteomics, and metabolomics. The clinical pathologist as a laboratory physician becomes responsible for developing these sophisticated testing methods and interpreting results for other physicians.

Elmer W. Koneman; C. William Reiquam; Bradley B. Brimhall

Bibliography. M. L. Bishop, E. P. Fody, and L. E. Schoeff, *Clinical Chemistry: Principles, Procedures, and Correlations*, 5th ed., Lippincott Williams & Wilkins, Philadelphia, 2004; S. H. Goodnight, Jr., and W. E. Hathaway, *Disorders of Hemostasis and Thrombosis*, McGraw-Hill, New York, 2001; J. B. Henry et al., *Clinical Diagnosis and Management by Laboratory Methods*, 20th ed., W. B. Saunders, Philadelphia, 2001; C. Kjelsberg et al., *Practical Diagnosis of Hematological Disorders*, 3d ed., ASCP Press, Chicago, 2000; M. Lichtman et al., *Williams Hematology*, 7th ed., McGraw-Hill, New York, 2006; J. H. Nichols, *Point of Care Testing*, Lippincott Williams & Wilkins, Philadelphia, 2005; T. Sun, *Text of Flow Cytometric Analysis of Hematologic Neoplasms*, 2d ed., Williams & Wilkins, Philadelphia, 2002; W. C. Winn, Jr., et al., *Koneman's Color Atlas and Textbook of Diagnostic Microbiology*, 6th ed., Lippincott Williams & Wilkins, Philadelphia, 2005.

Clock (mechanical)

A mechanical device for indicating the passage of time. Most clocks contain a means for producing a regularly recurring action, such as the swing of a pendulum, the oscillation of a spiral spring and balance wheel, the vibration of a tuning fork, the oscillation of a piezoelectric crystal, or the comparison of a high-frequency signal with the radiation from the hyperfine structure of the ground state of atoms. This article describes the refinements of mechanical clocks. *See* ATOMIC CLOCK; QUARTZ CLOCK; WATCH.

Basic movement. The recurring action of a mechanical clock depends on the swing of a pendulum, or the oscillation of a balance wheel and balance spring or hairspring, or the vibration of a tuning fork, mechanisms capable of repeating their cyclic movements with great regularity. A counting mechanism, consisting of a gear train with calibrated dial and indicating hands, sometimes with a striking mechanism, marks the number of oscillations that have occurred, although the graduations are in seconds, minutes, and hours. A weight or spring ordinarily supplies power to operate the oscillating and the counting mechanisms. However, temperature changes, accelerations, automatic windings, or electricity may provide the power. (The conventional electric clock is not an independent timepiece but is a repeater whose accuracy depends on a remote primary clock.) Usually an escapement transmits power from the counting mechanism to the oscillating mechanism. The accuracy of a clock depends primarily on the escapement.

Escapements. Figure 1 depicts an anchor or recoil escapement. Anchor *A* is connected loosely to the pendulum and swings about *C*. At some time after midswing, a tooth of the wheel escapes from pallet P_1 or P_2, giving the pallet a push to maintain oscillation. The other pallet then checks another tooth, the curve of the pallet forcing the wheel slightly backward. The reaction helps to reverse the pendulum swing and to correct for circular error.

Figure 2 shows a deadbeat escapement. Faces *F* are circles about *C* so that the wheel is stationary

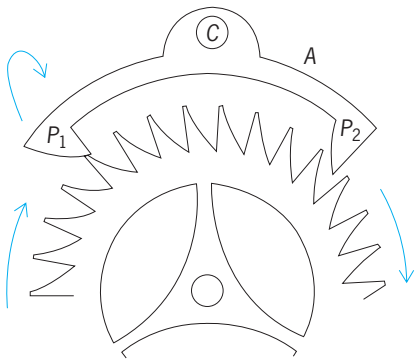

Fig. 1. Anchor escapement. Common in domestic pendulum clocks, this escapement tends to compensate for changes in amplitude of the swing because of irregularly cut gears and varying lubrication.

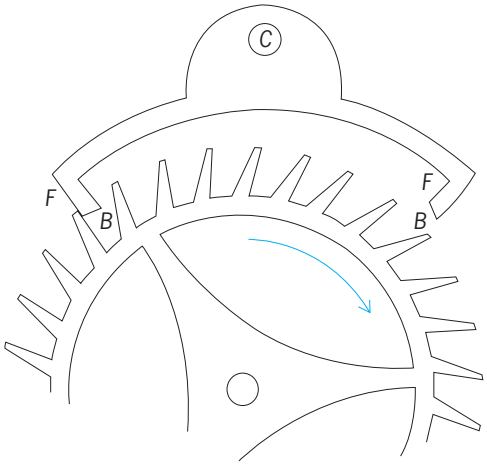

Fig. 2. Deadbeat escapement. Although lacking the corrective feature of the recoil type, this escapement gives better performance if well made.

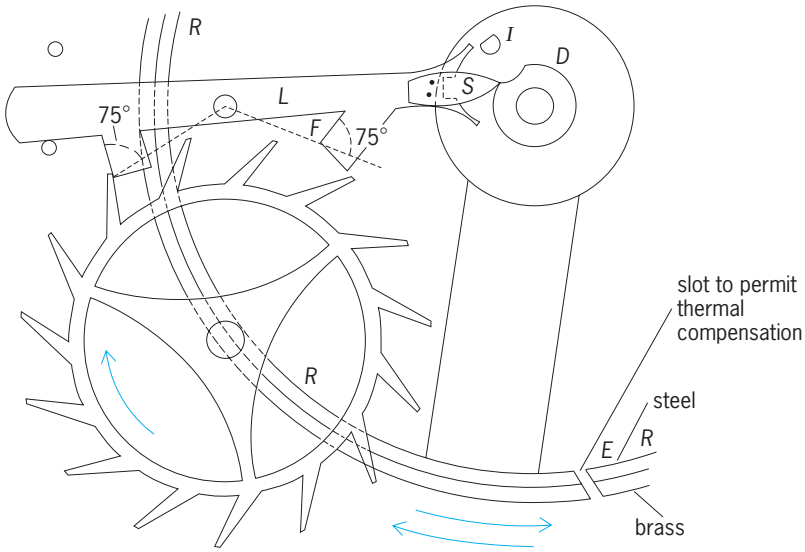

Fig. 3. Lever escapement. Used in good spring clocks and watches, it is detached from the balance wheel during most of a cycle, thus promoting accurate timekeeping.

except near midswing when a thrust on *B* allows a tooth to escape. **Figure 3** shows a lever escapement in which a balance wheel swings through a much larger angle than does a pendulum. Balance wheel and escapement are connected only while impulse pallet *I* is in a notch of lever *L*. As *I* enters a notch, it pushes the lever and then receives an impulse from it. Disk *D* and safety pallet *S* prevent the lever from moving accidentally before *I* returns; thus the mechanism tolerates appreciable vibration. The escapement pallets are similar to those of Fig. 2, but faces *F* are drawn or inclined so that the escape wheel moves slightly forward after each tooth locks; this motion pulls *S* clear of *D*, leaving the balance wheel completely detached. *See* ESCAPEMENT.

In the escapement of **Fig. 4**, the balance wheel is completely detached except for three brief action periods: (1) when pallet *P* presses detent *D* to unlock catch *L*, (2) when an impulse is being given to pallet *I*, and (3) when *P* lifts gold spring *G* on its return half swing. *See* CHRONOMETER.

Power supply. A weight gives uniform power, but a spring develops less force as it runs down. A fuzee by which a cord or chain is unwrapped from an increasing spiral on the first gear provides uniform force from a spring. Alternatively, there may be a remontoir with a small secondary spring driving the clock and being rewound frequently by the mainspring or by an electric motor.

Wind and snow apply large force to the hands of outdoor turret clocks. A gravity escapement isolates the pendulum from such forces. In this design, first used on the Great Westminster Clock (Big Ben), which has run since the mid-1800s with nearly chronometric accuracy, the escape wheel alternately lifts two weighted arms away from the pendulum as it swings out and releases them alternately against the pendulum as it swings back. The maintaining pushes come from the weighted arms rather than directly from the driving power.

Temperature compensation. The rate of a pendulum or balance wheel decreases if it expands with rising temperature. Expansion is counteracted by using two metals of different expansion coefficients. In a compensated balance wheel, rim *R* is of brass and steel fused together and cut (Fig. 3). Small screw weights added near free end *E* increase the effect; near the spoke they counteract it. Alternatively, the hairspring can be of an alloy that changes in stiffness with temperature.

Shortt clock. For still greater accuracy, in the Shortt or free-pendulum clock, a master pendulum is enclosed in an airtight and nearly evacuated case in a constant-temperature cellar. This pendulum swings freely for about 30 s until wheel *W* is released by electromagnet *E* and rolls along arm *A* attached to the pendulum rod (**Fig. 5**). Near the end of the swing, wheel *W* rides down slope *B*, giving the pendulum the slight impulse needed to maintain oscillation for another 30 s. In falling off, wheel *W* triggers the circuit and is reset by electromagnet *F*. The trigger

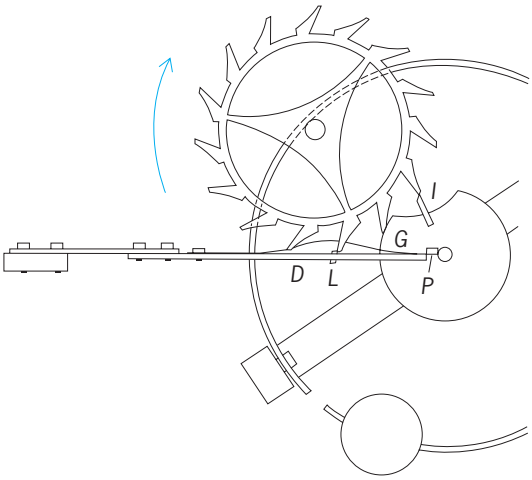

Fig. 4. Escapement used in ships' chronometers. It is sensitive to sudden movements but otherwise can keep within 1 s/day; pallets need no oil.

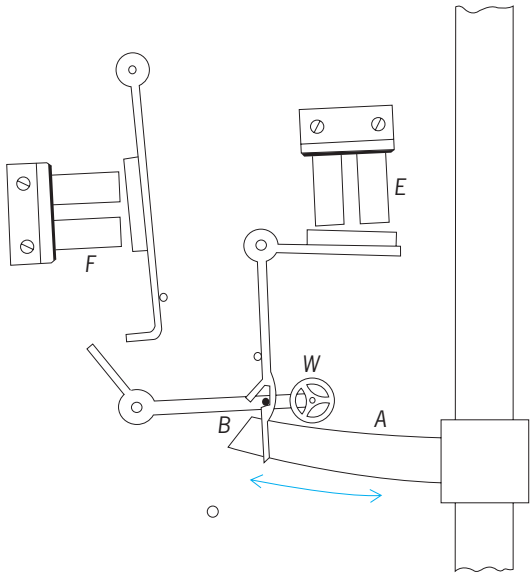

Fig. 5. Shortt free-pendulum clock. The master has only a pendulum and impulse giver; escapement and counting gears are in a separate slave clock (not shown).

pulse also serves as a synchronizing signal to a slave clock which is a complete clock with pendulum and indicating mechanism. If the slave is slow, the pulse accelerates one swing of its pendulum. *See* HOROLOGY; TIME.

Robert D'E. Atkinson

Tuning-fork mechanism. In the Accutron design the balance wheel and hairspring are replaced by a precision tuning fork which is driven with energy from a tiny battery. A permanent magnet and a surrounding magnetic cup are mounted on each tine of the tuning fork, which vibrates 360 times per second. Fixed coils of wire extend into the area between each magnet and cup, lying inside the magnetic field but not in contact with the moving parts of the tuning fork. Passing a current through one of these coils causes it to become an electromagnet which, depending on the polarities involved, will either repel or attract the corresponding magnet and cup assembly. Conversely, motion of the magnet and cup assembly within its electromagnetic coil causes a voltage to be induced, the polarity being determined by the direction of motion. This electromagnetic system serves three purposes: (1) It drives the tuning fork by converting electrical energy into mechanical impulses, (2) allows a determination of the tuning-fork amplitude by measuring the induced voltage, and (3) permits the accurate determination of the point within each tuning-fork cycle when the driving impulse should be applied. Conversion of the vibrational motion of the tuning fork to rotary motion is achieved by a simple ratchet and pawl mechanism.

Raynor L. Duncombe

Bibliography. F. J. Britten et al., *Britten's Old Clocks and Watches and Their Makers*, 9th ed., 1990; J. E. Haswell, *Horology*, 1928, reprint 1976; R. W. Swedberg and H. Swedberg, *American Clocks and Clockmakers*, 1989; H. Tait, *Clocks and Watches*, 1983.

Clock paradox

The phenomenon occurring in the special theory of relativity wherein two observers who start together with identical clocks and then undergo different motions can have different total elapsed time on their clocks when they rejoin later. This effect is a well-defined, mathematically consistent prediction of special relativity which has been verified by experiment but, historically, it has been referred to as a paradox because of erroneous reasoning in the manner in which the effect is commonly analyzed. The clock-paradox phenomenon arises because there is no notion of absolute simultaneity in the theory of special relativity.

The clock-paradox effect is illustrated by the following hypothetical example. (In the context of the example the effect usually is referred to as the twin paradox.) Two identical twins are separated as young adults. One of the twins remains on Earth and, for the purposes of the discussion, is assumed to undergo inertial motion. (In fact, in the context of general relativity, the twin on Earth would be viewed as accelerating—only freely falling observers would be viewed as inertial—but in this example the corrections made to the effect by performing a proper, general-relativistic analysis would be negligible.) The other twin is placed in a rocket ship which accelerates very rapidly away from Earth until it is receding from Earth at nearly the speed of light. After coasting for a while at this speed, the rocket ship turns around and accelerates very rapidly back toward Earth, so that it soon is traveling at nearly the speed of light. The rocket ship then lands on Earth and the twins rejoin. The twin on Earth is old (and everything else on Earth has aged considerably), but the twin in the rocket ship (and everything else therein) has barely aged at all.

Comparison of path lengths. To explain the nature of this prediction of special relativity and why it has been thought to be paradoxical, it is useful to consider the following elementary phenomenon in ordinary euclidean plane geometry, which is mathematically very closely related to the clock paradox. In particular, two points A and B (**Fig. 1***a*) and the straight line segment AB connecting these points are considered, as well as a third point C and the path ACB comprising the straight line segments AC and CB. Then, although both paths AB and ACB start at A and end at B, they have different total lengths. Indeed, in the simple case where C is equidistant from A and B, by the pythagorean theorem, the lengths of the two paths are related by Eq. (1), where L' denotes

$$L' = L\sqrt{1 + \left(\frac{2d}{L}\right)^2} \qquad (1)$$

the length of the path ACB, L denotes the length of AB, and d denotes the perpendicular distance of C from AB.

It is possible to use the following method to calculate the increased length of the path ACB given by Eq. (1). A series of closely spaced lines is drawn

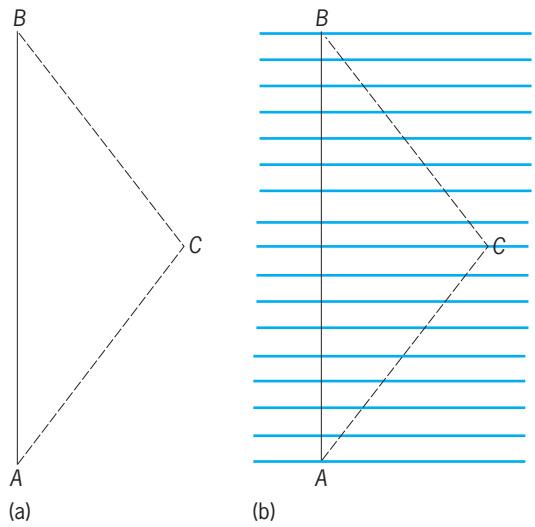

Fig. 1. Comparison of path lengths and elapsed times. (*a*) Straight line *AB* and broken straight line *ACB* connecting points *A* and *B* in euclidean plane geometry. This figure can also represent the world lines of an inertial observer and a "twin." (*b*) Method for comparing the lengths of *AB* and *ACB* in euclidean plane geometry. It can also be used to compare the elapsed time on clocks carried by an inertial observer and a "twin."

perpendicular to AB (Fig. 1*b*). An adjacent pair of such lines is considered, both of which intersect AC. Then it can be seen that the length of the portion of AC lying between these two lines is greater than the corresponding portion of AB by the factor $\sqrt{1+(2d/L)^2}$. The same result holds if two adjacent lines are considered which intersect CB rather than AC. However, the total length of AB can be obtained by summing the lengths of all of the portions of AB lying between adjacent perpendicular lines. By doing the sum, it is possible to correctly rederive the result that ACB is longer than AB by the factor $\sqrt{1+(2d/L)^2}$.

However, an apparent distance paradox arises from an attempt to redo this calculation as follows. A series of lines is drawn perpendicular to AC (**Fig. 2***a*).

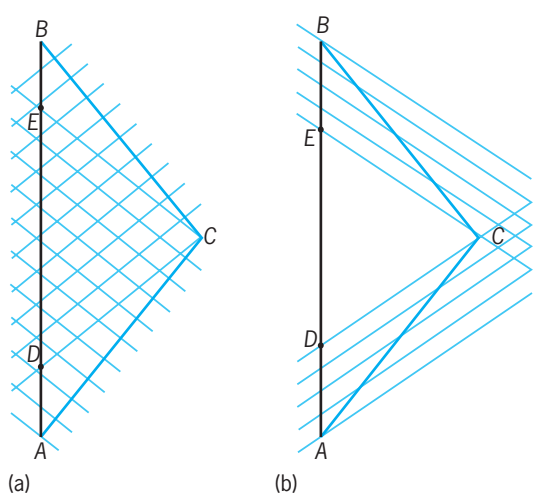

Fig. 2. Alternate methods for comparing (*a*) the lengths of *AB* and *ACB* in euclidean plane geometry, and (*b*) the elapsed times on the clocks carried by an inertial observer (world line *AB*) and a "twin" (world line *ACB*).

The portion of AC lying between adjacent lines now will be shorter than the corresponding portion of AB by the factor $\sqrt{1+(2d/L)^2}$. The same will be true for a similar calculation for the segment CB. Therefore, it might also seem possible to derive the manifestly false conclusion that the path ACB is shorter than the straight line AB by the factor $\sqrt{1+(2d/L)^2}$.

It is not difficult to spot the fallacy in this latter argument. When the series of lines is drawn perpendicular to CB, they will intersect the lines perpendicular to AC (Fig. 2*a*). As a consequence, this argument implicitly double counts a portion, DE, of AB. When a correction is made for this double counting, the correct conclusion that ACB is longer than AB by the factor $\sqrt{1+(2d/L)^2}$ is restored.

Comparison of elapsed times. The original diagram (Fig. 1*a*) may also be used to represent the twin-paradox phenomenon if it is now viewed as a space-time diagram with the vertical axis corresponding to time t and the horizontal axis corresponding to the spatial dimension x, so that "forward in time" corresponds to "upward" in the diagram. At event A, the twins separate, and the first twin undergoes inertial motion in space-time represented by the straight line AB. The segment AC represents the portion of the world line of the second twin while rapidly receding from the first twin, whereas CB represents the portion of the second twin's world line during the return. The comparison of the elapsed time T along the world line AB with the elapsed time T' along the world line ACB differs from the previous distance calculation only in that lorentzean geometry rather than euclidean geometry must be used in the analysis. Taking account of the sign change occurring in the lorentzean analog of the pythagorean theorem, the elapsed times are related by Eq. (2), where

$$T' = T\sqrt{1-\left(\frac{2d}{cT}\right)^2} = T\sqrt{1-\left(\frac{v}{c}\right)^2} \qquad (2)$$

$v = 2d/T$ is the magnitude of the relative velocity of the twins, and c is the velocity of light. In special relativity, the "straight line" (that is, inertial motion) AB has longer elapsed time than the motion represented by the world line ACB or, indeed, any other noninertial motion between A and B. *See* LORENTZ TRANSFORMATIONS.

The ratio of T' to T also can be calculated by the analog of the method of drawing lines perpendicular to AB (Fig. 1*b*), and this method correctly reproduces Eq. (2). There is also a calculation analogous to that of drawing lines perpendicular to AC and CB (Fig. 2*a*), where the lines are perpendicular to AC and CB in the lorentzean geometry of space-time (Fig. 2*b*). A careless use of this method yields the conclusion that T' is longer than T, resulting in the clock paradox. This paradox is easily resolved by noting that this method fails to count the portion DE of AB.

Experimental verification. The clock-paradox phenomenon has been observed directly in an experiment performed in 1971 by J. C. Hafele and R. E. Keating, who observed differences in elapsed times of atomically stabilized clocks flown

in airplanes as compared with ones on the ground. In this experiment the special relativistic effect is so small, since the velocities achieved by airplanes are much smaller than *c*, that the tiny corrections due to general relativity cannot be neglected, so the experiment actually must be viewed as verifying the analogous clock effect in general relativity rather than purely the clock effect of special relativity described here. *See* RELATIVITY; SPACE-TIME.

Robert M. Wald

Bibliography. H. Bondi, *Relativity and Common Sense,* 1964, reprint 1980; H. Chang, A misunderstood rebellion: The twin paradox controversy and Herbert Dingle's vision of science, *Stud. Hist. Phil. Mod. Sci.*, 24:741–790, 1993; J. C. Hafele and R. E. Keating, Around-the-world atomic clocks, *Science,* 1787:166–170, 1972; R. M. Wald, *Space, Time, and Gravity,* 2d ed., 1992.

Cloning

The asexual creation of a genetic copy, a capability possessed by plants but not by most animals. Thus, plants generate genetic copies spontaneously, and rooting "cuttings" is widely used by horticulturists to propagate millions of clones annually. In animals, only some lower invertebrates can be cloned by "cutting"; for example, earthworms when bisected will regenerate the missing half, resulting in two whole, genetically identical individuals. However, asexual reproduction and cloning do not normally occur in vertebrates except for the special case of identical twinning. This is despite the fact that individual cells, called blastomeres, within the very early embryo are totipotent; that is, each is capable, if evaluated on its own, of developing into a viable term pregnancy and infant.

A major scientific interest in cloning revolves around the question of whether the hereditary material in the nucleus of each cell remains intact throughout development, regardless of the cell's fate. On a more practical level, the production of genetic copies of mammals could support the rapid improvement of livestock herds by propagation of valuable founder animals, the creation and production of disease models or transgenic animals for biomedical research, and the preservation of the genetic contribution of a particularly valuable animal, even after death. Therapeutic cloning, a variation that involves the isolation of embryonic stem cells, may provide new cell-based medical approaches to the treatment of human diseases or degenerative conditions.

Reproductive cloning. Scientific inquiry into reproductive cloning in animals began with a "fantastical experiment" suggested by Hans Spemann in 1938 that involved the insertion of a nucleus into an ovum bereft of its own genetic material. This experiment was eventually conducted in 1952 by Robert Briggs and Thomas King in an amphibian, the northern leopard frog, and the technology was quickly extended to a number of other lower vertebrates and invertebrates, and eventually to mammals. The first step in mammalian reproductive cloning is removal of the genetic material from an egg by micromanipulation to create an enucleated egg called a cytoplast. Then genetic material from a donor cell is added, in the form of an intact cell or an isolated nucleus, to produce a diploid, reconstructed embryo. The cell cycle of the nuclear donor cell may be temporarily slowed or stopped in advance of nuclear transfer. Development of the nuclear transfer embryo is triggered chemically, and the cloned embryo is subsequently transferred into a host mother in order to establish a pregnancy (see **illustration**).

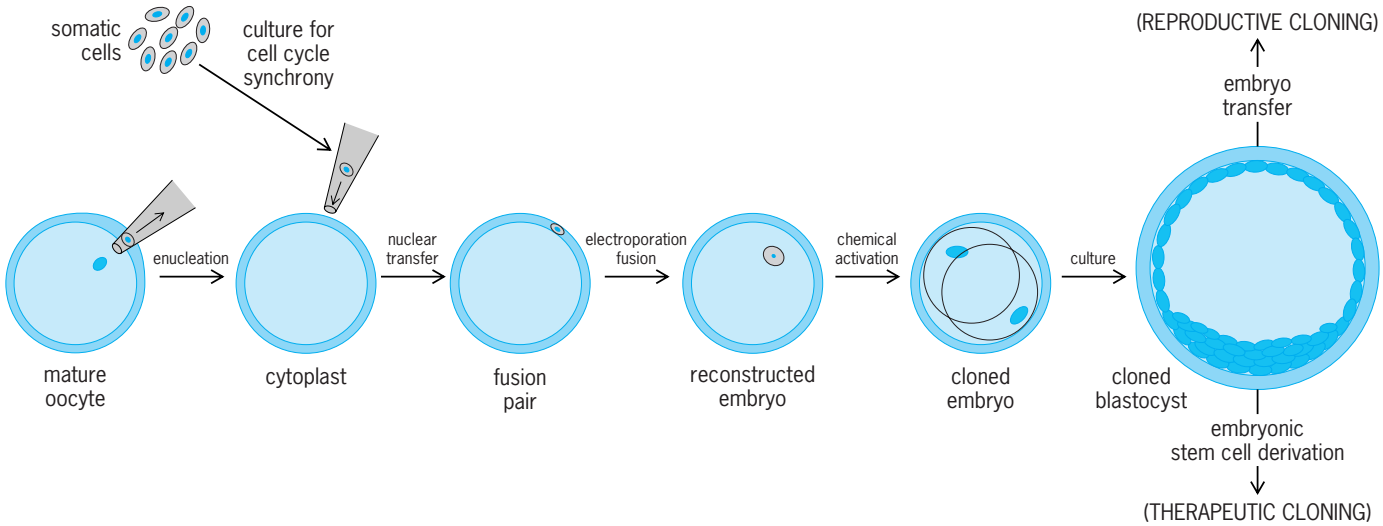

Mammalian somatic cell cloning. The starting requirements include a cohort of mature fertile eggs and a nuclear donor source. Reproductive cloning in mammals involves the transfer of a nucleus or an entire donor cell into an enucleated egg or cytoplast, followed by the subsequent reprogramming of that nucleus by the egg, to the degree that the events of development are recapitulated in the cloned embryo. Therapeutic cloning is accomplished by conducting nuclear transfer with a nuclear donor cell derived from the patient, a skin fibroblast perhaps, transferred into a cytoplast. The resulting nuclear transfer embryo, after culture to the blastocyst stage, would then provide a source of inner-cell-mass cells from which an embryonic stem cell line could be derived.

When the nuclear donor cell originates from an embryo, the process is called embryonic cloning, first reported in cattle and sheep in 1986 and in primates in 1997. When the donor nucleus is derived from a fetus or a juvenile or adult animal and is not a germ cell, the process is called somatic cell cloning. This form of cloning was thought to be impossible in mammals even as recently as 1995, when scientists in Scotland cloned sheep from differentiating embryonic cells. In 1996, the birth of Dolly, a cloned sheep, was the first successful generation of a viable mammal derived from the transfer of an adult cell (a mammary gland cell).

Successful somatic cell cloning in other species (including the cow, mouse, goat, pig, rabbit, cat, horse, mule, and rat) quickly followed. A wide variety of somatic cells—even those that are unequivocally, terminally differentiated—have been used as nuclear donors. It now appears logical that cloning will be possible in all mammals, given an adequate opportunity to define the unique conditions required by each species.

While dramatic progress has been realized in somatic cell cloning, virtually nothing is known about how the process works, and significant limitations continue to undermine efforts to produce viable offspring. Thus, somatic cell cloning is characterized by high fetal and neonatal wastage, along with age-onset health problems (including obesity, hepatic and immune impairment, and premature death) in surviving clones. It is in fact amazing that somatic cell cloning works at all, given that over the course of normal development as a cell specializes, some genes are silenced temporarily or in some cases permanently. In order for the events of development to be accurately recapitulated following nuclear transfer of a differentiated cell, the donor nucleus must be reprogrammed so that it ceases to express genes of importance to a differentiated or somatic cell in favor of those important to a young embryonic cell; reprogramming is a responsibility of the cytoplast that must occur in a timely manner following nuclear transfer, perhaps along with dramatic remodeling of chromosomal architecture. In addition, nuclear reprogramming must ensure retention of the pattern of expression of imprinted genes, that is, those genes that are differentially expressed depending on their maternal or paternal origin. The inappropriate or abnormal expression of imprinted genes, as regulated by epigenetic changes (factors that alter the expression of genes rather than the genes themselves) such as DNA methylation, can impact placental development and fetal wastage.

Other possible explanations for the losses associated with somatic cell cloning are that clones age prematurely because the nuclear donor cells have shortened telomeres (chromosome tips) and because of heteroplasmy, the presence of more than one type of mitochondrial DNA in a cell. Since mitochondrial DNA is contributed by both the cytoplast and by the injection of a donor cell or nucleus preparation, clones are not exact genetic copies; indeed, they can be phenotypically quite different from one another. Fetal exposure to unique environments is another factor that results in phenotypic differences. *See* CELL DIFFERENTIATION; CELL DIVISION; GENE; SOMATIC CELL GENETICS.

Therapeutic cloning. In therapeutic cloning, embryos are produced by nuclear transfer, as in reproductive cloning, but with the intent of isolating embryonic stem cells rather than transferring cloned embryos into a host to produce a pregnancy. Embryonic stem cells are undifferentiated cells present in the embryo that are pluripotent; that is, they are capable of giving rise to all of the major cell lineages in the adult body but unable to generate a functioning organism. Adult stem cells exist and can be recovered in limited numbers from a host of different tissues or organs. They can, in theory, be used in cell-based therapies to treat human disease, but their low numbers and limited developmental potential are drawbacks. It is now clear that the pre-implantation-stage embryo contains a cohort of stem cells confined to the inner cell mass at the blastocyst stage, and these embryonic stem cells can be harvested, propagated, and established as permanent cell lines. These embryonic stem cells could be stored frozen until needed and then thawed and used directly for transplantation into the patient, or after their fate was directed into the desirable cell type in vitro by exposure to the appropriate chemical signals (see illustration).

Embryonic stem cells are the most primitive, most mitotically capable stem cells, and they hold virtually unlimited promise in the treatment of diseases attributable to the loss or dysfunction of known cell types. This cell-based therapy, as envisioned, would involve the transplantation of embryonic stem-cell-derived progeny that had been induced to differentiate into the missing or dysfunctional cells. Transplantation of allogeneic stem cells (stem cells from another individual of the same species) or their progeny, however, is complicated by the likelihood of host-graft rejection. As an alternative to overcome this problem, the patient could generate his or her own embryonic stem cells by contributing the donor nucleus or cell in somatic cell cloning, thereby creating self-cells transparent to the immune system such that their transplantation would not elicit an immune response. Proof-of-principle experimentation with adult cells in therapeutic cloning has now been reported in both mice and humans. *See* STEM CELLS.

Future research. An immediate focus of cloning research is to overcome the low success rates with somatic cell cloning in order to extend this technology to other nonhuman species. Undoubtedly, improvements will be realized with additional experience and increased surveillance at birth, with cloned pregnancies treated as high risk. This is already happening with cattle as substantial financial incentive exists in the cattle industry to produce cloned animals, and with mice due to their amenability to laboratory experimentation. Scientists are asking what kinds of cells make the best donors, what environments

are most conducive to early development of the cloned embryo, what cell cycle stages are optimal for reprogramming, and how reprogramming can be initiated in the test tube. It seems reckless to extend reproductive cloning to humans at present because of the unacceptable risks imposed on the embryo, fetus, and neonate. Additional ethical considerations have been addressed extensively, for instance, in a United States National Bioethics Advisory Commission report issued in 1997. Therapeutic cloning in humans is less controversial, since it is not associated with efforts to produce cloned children and may represent a viable approach to treat human disease. Clearly, though, ethical concerns do persist with therapeutic cloning, based on the requirement to recruit egg donors in order to create a cloned embryo and then the need to destroy the embryo in the process of deriving embryonic stem cell lines. An alternative to creating individual embryonic stem cell lines by therapeutic cloning may ultimately be the development of stem cell banks that can be used to match the needs of all patients.

While many challenges lie ahead for somatic cell cloning, the mere demonstration that this process is possible in mammals is a highly significant beginning. Ongoing research efforts will lead to a better understanding of the molecular correlates to reprogramming, information that in turn will allow the development of in vitro conditions for reprogramming (for instance, the use of readily available cytoplasm as might be harvested from amphibian eggs). Ultimately, reprogramming will be accomplished with chemically defined reagents and, in the process, we will have developed profound new insights into the events underlying development in mammals, humans included. *See* GENETIC ENGINEERING.

Don P. Wolf

Bibliography. M. A. DiBerardino, R. G. McKinnell, and D. P. Wolf, The golden anniversary of cloning: A celebratory essay, *Differentiation*, 71:398–401, 2003; K. Hechedlinger and R. Jaenisch, Monoclonal mice generated by nuclear transfer from mature B and T donor cells, *Nature*, 415:1035–1038, 2002; W. S. Hwang et al., Evidence of a pluripotent human embryonic stem cell line derived from a cloned blastocyst, *Science*, 303:1669–1674, 2004; A. McLaren, Cloning: Pathways to a pluripotent future, *Science*, 288:1775–1780, 2000; National Bioethics Advisory Commission, *Cloning Human Beings: Report and Recommendations of the National Advisory Commission*, National Bioethics Advisory Commission, Rockville, MD, 1997; E. Pennisi and G. Vogel, Clones: A hard act to follow, *Science*, 288:1722–1727, 2000; I. Wilmut et al., Somatic cell nuclear transfer, *Nature*, 419:583–586, 2002.

Closed-caption television

A service, primarily for deaf people, that allows individual viewers to display the dialog of a television program in a readable form on the television screen. In the United Kingdom the service is called subtitling. Closed-caption television was introduced in the United States in 1980.

Caption preparation. Captions for television programs can be prepared in three ways. (1) Off-line captioning is used when the television program has been prerecorded. This is the preferred method for creating captions in that they can be caused to appear on different parts of the home television screen, for example, under the individual speaking, and they can be displayed in close synchronism with the dialog. Also, the caption display can be edited if necessary to a comfortable reading rate of about 125 words per minute. (2) Live television broadcasts that are scripted in advance, such as presidential State of the Union speeches, employ a live display captioning method. Here the script is stored in a computer in advance of the broadcast. During the broadcast the script appears as roll-up captions at the bottom of the screen. (3) In live broadcasts, where a script is not available, real-time captions have to be generated during the broadcast. This is by far the most difficult way to caption a television program. Highly skilled court reporters using a stenograph keyboard and an associated computer create the captions, often at instantaneous rates of 200–250 words per minute. In this method, inevitably, typographical errors are often generated.

Caption transmission. Prior to being broadcast, the captions are digitally encoded and inserted into the vertical blanking interval (VBI) of the television picture. In the United States and Canada, a single VBI line is allocated exclusively for the captioning service. In other countries such as the United Kingdom and Australia, the captions are integrated with other text material on a series of VBI lines that are used to deliver Teletext service to a more general audience. As the Teletext transmission data rate is about ten times the caption-only service data rate, it cannot be used in videocassette recorder (VCR) applications. This led to the caption-only system being introduced in home-television movies in the United Kingdom in the early 1990s.

Digital television. The world's terrestrial television systems are in the process of converting from analog to digital broadcasting. In the United States, the Federal Communication Commission (FCC) has adopted the Advanced Television Standards Committee (ATSC) system. Europe has adopted the Digital Video Broadcasting–Terrestrial (DVB-T) system. Elsewhere in the world, those countries that previously adopted the United States or European analog television standards are typically adopting the respective digital television standards. However, the International Telecommunications Union (ITU) recommendation on a digital video, audio, and ancillary data multiplex model has been adopted worldwide. Captions have been classified as an ancillary service and will be multiplexed with other ancillary services such as text services.

Captioned program service growth. Legislative and regulatory actions are spurring the worldwide

introduction and growth of captioned television service. In many countries, specific goals have been set as to the percentage of television programs that must be captioned by the broadcasters. In the United States, it is illegal to sell a television set with a screen size 12 in. (30 cm) or larger that does not contain caption decoder circuitry built in by the manufacturer. *See* TELEVISION. John E. D. Ball

Bibliography. J. E. D. Ball, Reaching Millions of Viewers Through Closed Captioning, *Public Telecom. Rev.*, vol. 8, no. 6, November/December 1980; J. E. D. Ball and D. R. Wells, *Encoded Captioning of Television for the Deaf: International Television Symposium, Montreaux, Switzerland*, PBS Engineering and Technical Operations Report E 7207A, May 24, 1973; V. Murphy-Berman, The impact of captions on hearing-impaired children's affective reactions to television, *J. Spec. Educ.*, 17:47–62, 1983; *Requirements That Video Programming Be Closed Captioned: Federal Communication Commission Report and Order* implementing Section 713 of the Communications Act of 1934, as amended by Section 305 of the Telecommunications Act of 1996, released Aug. 22, 1997.

Closed-circuit television

Television transmitted to a particular audience at specific locations via coaxial cables, telephone wires, fiber-optic strands, microwave radio systems, or communications satellites, as compared to open-circuit (broadcast) television intended for the general public. *See* COAXIAL CABLE; COMMUNICATIONS CABLE; COMMUNICATIONS SATELLITE; MICROWAVE; OPTICAL COMMUNICATIONS; TELEPHONE SERVICE.

The usual closed-circuit television picture display device is a television receiver. Most receivers use a cathode-ray tube to produce the visible image, although other devices such as flat-panel displays have become common. *See* CATHODE-RAY TUBE; ELECTRONIC DISPLAY; FLAT-PANEL DISPLAY DEVICE; LIQUID CRYSTALS; TELEVISION RECEIVER.

Applications. Closed-circuit television applications include information display, remote monitoring, instruction, cable television programming, and teleconferences and special events.

Information display. Flight times in airports are displayed by closed-circuit television, typically on large cathode-ray-tube video screens. A video screen may be used to display an image of the sales floor to a cashier at the front of a store; this setup is frequently installed for security reasons. In both examples, closed-circuit television provides information, but different methods are used for generating the video. The airport display usually shows text produced by a computer. The store screen shows the image of the sales floor produced by a video camera. These applications exemplify simple forms of closed-circuit television used to connect a source of information to a restricted group of viewing locations.

Remote monitoring. Many types of closed-circuit television systems produce pictures intended for distant viewing. These images may include those taken from a space probe passing close to a distant planet, or pictures of blast furnaces or other industrial operations that could be hazardous to human observers. At the opposite extreme, the low cost of modern video equipment makes practical the use of small cameras to observe sleeping children at home and to enjoy video images of people far away while chatting with them on the Internet.

Parking lots at shopping centers, store cashier locations, bank and hotel lobbies, and other public places are often equipped with closed-circuit television surveillance cameras and displays. These systems may be designed to function in nearly total darkness to detect suspicious activities or in brilliant surroundings to watch for hot spots in factory processes. Medical and scientific equipment and industrial robots often incorporate specialized closed-circuit television systems. Modern closed-circuit television cameras used for remote monitoring are sensitive, rugged, relatively inexpensive, and durable, with long service lives. Most modern solid-state designs use charge-coupled imaging devices. Video signals are often recorded on magnetic tape for future use. *See* CHARGE-COUPLED DEVICES.

Instruction. Closed-circuit television has been available for many years for educational uses. Many universities and school systems employ television for teaching. Classrooms may be equipped with closed-circuit television cameras, recorders, and receivers. Lesson material is often presented via video tape recordings, CD-ROMs, or DVD recordings. In other cases, classroom receivers are supplied with signals from specially reserved broadcast channels (called ITFS channels, for Instructional Television Fixed Service), cable television channels, or satellite channels. Audio circuits installed at viewing locations sometimes allow two-way conversations between lecturer and students. *See* COMPACT DISK; MAGNETIC RECORDING.

Cable television. Cable television was originally developed to improve broadcast reception in difficult receiving situations—for example, to bring distant signals from broadcast television stations to residents of mountain valleys or to serve viewers in cities where large buildings prevent satisfactory reception of broadcast signals. Contemporary cable systems continue to distribute signals from local and distant television stations, but they now offer additional television services not available from broadcast stations, such as satellite-distributed news, music, movies, weather information, and locally originated community-based programming. Cable systems offer services to an estimated 65 million households in the United States (75% of the total) using either coaxial cables or fiber-optic strands for signal delivery. *See* CABLE TELEVISION SYSTEM.

Wider availability has reduced the cost of using fiber-optic strands for high-speed signal transmission. High-speed interconnections enable

closed-circuit television systems to utilize important new kinds of digitally based services. For example, many telephone and cable television companies provide fiber-optic connections to schools and classrooms. These permit relatively inexpensive two-way video and audio connections between schools and among school systems, enabling great flexibility in system design and utilization.

Teleconferences and special events. Closed-circuit television is used to conduct conferences among geographically separated groups of people. Signals from television-equipped meeting rooms or studios are distributed by fiber-optic circuits, communications satellites, or the Internet to convention centers, school or university classrooms, motels, hotels, or corporate locations. Two-way audio is typically provided, and conferences may involve many hundreds or thousands of people. Sports or musical events may also be televised and distributed in this fashion for viewing in suitably equipped theaters or auditoriums. *See* TELECONFERENCING.

Technology. Closed-circuit television systems frequently utilize equipment that produces signals based on broadcasting industry specifications. However, in contrast to the many possible designs for closed-circuit television systems, broadcast television is always limited to the use of standardized signals for its transmissions. The standardized television broadcast system used at present in the United States, Canada, Mexico, Japan, and some other countries was devised in the 1950s by the National Television Systems Committee (NTSC). NTSC systems display 30 complete television pictures (frames) per second, and each frame is composed of 525 horizontal scanning lines arranged from top to bottom. *See* TELEVISION STANDARDS.

Video cameras, television receivers, videocassette recorders, and accessories are produced in large quantities for NTSC broadcast applications. The NTSC system is well suited for home entertainment, but its signal structure is technically limited by the requirement for simple, inexpensive home receivers and video tape recorders. Inauguration of improved kinds of advanced television systems, most of them digital, is under way in many countries. Such technologies will eventually permit higher picture quality and the mass production of affordable receivers with new features. These features will eventually also become available for closed-circuit applications.

Special requirements. Closed-circuit television systems are not required to use NTSC signals, but many do, mainly for economic reasons. In many applications, normal NTSC-based closed-circuit television equipment is satisfactory. If the transmission path to the viewer is of low quality, however, special steps must be taken. In such cases, the signal can be modified to meet more closely the characteristics of the transmission system. For example, channel bandwidth can be reduced without loss of image resolution if the user accepts reduced quality in the representation of image motion. In that case, a rate lower than the normal 30 frames per second can be chosen. This is called slow-scan closed-circuit television.

Still-frame video images sometimes suffice. A high-quality still frame can be transmitted with picture elements being stored for seconds or even minutes in a memory system and then assembled for display after receipt of all data. This allows use of transmission channels having limited bandwidth. Most video pictures from space are transmitted at low frame rates over narrow-bandwidth communications channels, and complete reception requires several seconds. Narrow-band channels are also found on the Internet, where both slow-scan and still images are often used. *See* INTERNET; REMOTE SENSING.

NTSC broadcast signal characteristics may also be poorly suited to other kinds of specialized closed-circuit television applications. Some applications require high image resolution, great sensitivity to light, the ability to respond to infrared or ultraviolet light, or very rugged system components. Closed-circuit television requirements for high-resolution display can be satisfied by increasing the number of image scanning lines and the overall video channel bandwidth. Industrial standards exist for high-resolution closed-circuit systems. Cameras and monitors are available to meet other specialized requirements.

Advanced developments. The digital revolution has affected the design of all kinds of television equipment. Video signals can be converted into digital form and processed by specialized computers, called digital signal processors. Visual special effects are often seen on entertainment television, including slow motion, still frames, rotations, inversions, and picture shape changes. They are accomplished by use of digital signal processing systems. In closed-circuit television applications, digital processing can make transmission via the Internet practical and can also provide image enhancement. The excellent pictures from weather and surveillance satellite imaging systems are obtained with digital signal processing and enhancement. High-performance solid-state camera sensors and the digital processing of video signals, taken together, have extended the applications of closed-circuit television systems. The advanced television systems now entering service make use of the advantages of digital signal processing to provide high-quality images to the home, to the class-room, and to industrial locations. *See* DATA COMPRESSION; IMAGE PROCESSING; TELEVISION. Frederick M. Remley

Bibliography. H. Benoit, *Digital Television: MPEG-1, MPEG-2 and Principles of the DVB System*, John Wiley, New York, 1997; V. Damjanovski, *CCTV*, Butterworth-Heinemann, Boston, 1999; H. Kruegle, *CCTV Surveillance: Video Practice and Technology*, Butterworth-Heinemann, Boston, 1996; A. C. Luther, *Principles of Digital Audio and Video*, Artech House, Norwood, MA, 1997; A. C. Luther, *Video Engineering*, McGraw-Hill, New York, 1999; M. Robin, *Digital Television Fundamentals: Design and Installation of Video and Audio Systems*, McGraw-Hill, New York, 1997.

Clostridium

A genus of bacteria comprising large anaerobic spore-forming rods that usually stain gram-positive. Most species are anaerobes, but a few will grow minimally in air at atmospheric pressure.

The clostridia are widely distributed in nature, and are present in the soil and in the intestinal tracts of humans and animals. They usually live a saprophytic existence, and play a major role in the degradation of organic material in the soil and other natural environments. A number of clostridia release potent exotoxins and are pathogenic for humans and animals. Among the human pathogens are the causative agents of botulism (*Clostridium botulinum*), tetanus (*C. tetani*), gas gangrene (*C. perfringens*), and an antibiotic-associated enterocolitis (*C. difficile*). *See* ANAEROBIC INFECTION; TOXIN.

Characteristics. Clostridial cells are straight or slightly curved rods, 0.3–1.6 micrometers wide and 1–14 μm long. They may occur singly, in pairs, in short or long chains, or in helical coils. The length of the cells of the individual species varies according to the stage of growth and growth conditions. Young cultures of most species of *Clostridium* are gram-positive and stain readily with the usual aniline dyes, but in older cultures and clinical specimens the organisms frequently stain gram-negative. Most clostridia are motile with a uniform arrangement of flagella. *See* CILIA AND FLAGELLA.

The endospores produced by clostridia are dormant structures capable of surviving for prolonged periods of time, and have the ability to reestablish vegetative growth when appropriate environmental conditions are provided. The spores of most clostridia are resistant to heat, chemical agents, and radiation and thus provide the organism with a selective advantage in its native habitat, enabling it to survive under the most adverse environmental conditions. The spores of clostridia are oval or spherical and are wider than the vegetative bacterial cell. Among the distinctive forms are spindle-shaped organisms, club-shaped forms, and tennis racket–shaped structures.

Clostridia are obligate anaerobes: they are unable to use molecular oxygen as a final electron acceptor and generate their energy solely by fermentation. Clostridia exhibit varying degrees of intolerance to oxygen. Some species are sensitive to oxygen concentrations as low as 0.5%, but most species can tolerate concentrations of 3–5%. The sensitivity of clostridia to oxygen restricts their habitat to anaerobic environments; habitats that contain large amounts of organic matter provide optimal conditions for their growth and survival.

A primary property of all species of *Clostridium* is their inability to carry out a dissimilatory reduction of sulfate. Most species are chemoorganotrophic. The substrate spectrum for the genus as a whole is very broad and includes a wide range of naturally occurring compounds. Extracellular enzymes are secreted by many species, enabling the organism to utilize a wide variety of complex natural substrates in the environment. Four nutritional groups are recognized: saccharolytic species (species that break down sugars), proteolytic species (species that hydrolyze proteins and peptides), species that are both saccharolytic and proteolytic, and species that are neither but are specialized on one or a few substrates. A number of saccharolytic species ferment carbohydrates, with the production of such compounds as butyric acid, butanol, isopropanol, and acetone. Proteolytic species metabolize proteins, often with the production of foul-smelling amines.

Pathogens. Species of *Clostridium* associated with human disease are divided into four major groups on the basis of the types of diseases: histotoxic clostridia cause a wide variety of tissue infections; enterotoxigenic clostridia produce gastrointestinal disorders; *C. tetani* is the etiologic agent of tetanus; and *C. botulinum* is the etiologic agent of botulism. In all clostridial diseases, pathogenesis is attributable to potent exotoxins released by the organism.

Histotoxic forms. *Clostridium perfringens* is the most important of the histotoxic clostridia that cause tissue infections in humans, especially of the muscle tissue (clostridial myonecrosis, or gas gangrene). This short, plump, nonmotile rod (2–4 μm wide) is more aerotolorant than most other anaerobes. Its strains produce at least 12 different toxins and enzymes, many of which have lethal, cell-destroying, and hemolytic (red blood cell–destroying) properties. A number of nonlethal enzymes are also produced, and apparently contribute to the invasiveness of the organism. These include collagenase, deoxyribonuclease, and hyaluronidase. *See* GANGRENE.

Other histotoxic clostridia include *C. novyi* and *C. septicum*. More than 30 additional species have been isolated from human clinical specimens, but most of these are considered to be nonpathogenic opportunistic organisms.

In individuals with malignancies of the cecum or with leukemia, invasion of the bloodstream by *C. septicum* or *C. perfringens* may lead to a life-threatening fulminating septicemia. *Clostridium perfringens* can also cause food poisoning if large numbers of an enterotoxin-producing strain are ingested. A more severe form of food poisoning, called enteritis necroticans, is caused by type C strains of *C. perfringens*. It is seen most often in malnourished children in the less developed countries. *See* FOOD POISONING.

Enterotoxigenic forms. *Clostridium difficile* causes pseudomembranous colitis, a severe disease associated with the administration of antibiotics. The antibiotics most often involved include clindamycin, ampicillin, and the cephalosporins. *Clostridium difficile* is widespread in the environment and is part of the normal intestinal flora of about 3% of healthy adults. When antibiotic-sensitive organisms are eliminated from the intestine, there is an overgrowth of *C. difficile*, severe intestinal mucosa damage, and resultant clinical disease.

Clostridium tetani. This species is the causative agent of tetanus. It is a long, thin rod that forms a round

terminal spore that gives the sporulating organism a characteristic drumstick appearance. The organisms are widely distributed in the soil and in the gastrointestinal tracts of humans and other animals. The spores remain viable in the soil for many years and provide the primary source of infection. Often after a simple puncture wound from a nail or splinter contaminated with soil containing the organisms, they multiply locally at the site and elaborate a potent neurotoxin (tetanospasmin) which is responsible for all of the symptoms of tetanus. *See* TETANUS.

Clostridium botulinum. The most common form of botulism is an intoxication that usually follows the ingestion of a preformed toxin produced by *C. botulinum*. There are seven types (A to G) of *C. botulinum* based on the immunologically distinct type of toxin produced. The production of a specific type of botulinum toxin is determined by the specific bacteriophage carried by the organism. Types A, B, and E are responsible for most cases of human botulism. The toxins are neurotoxins and are among the most toxic compounds known. They act on neuromuscular junctions by inhibiting the release of acetylcholine. *See* ACETYLCHOLINE.

The spores of *C. botulinum* are widely distributed in soils, lake sediments, and decaying vegetation in many parts of the world. They are very resistant to heat and can withstand 212°F (100°C) for 3–5 h; moist heat at 250°F (121°C) will kill them in 15 min. In the canning of food (especially vegetables), if the spores on contaminated food are not killed, they will germinate and release toxins when conditions become sufficiently anaerobic. *See* BOTULISM; FOOD PRESERVATION.

Disease in animals. A number of diseases in mammals and birds are caused by the ingestion of toxins elaborated by toxinogenic species of *Clostridium. Clostridium botulinum* causes grass or fodder sickness of horses, silage disease in cattle, limberneck in chickens, lamziekte in cattle, and dust sickness in wild birds. Type C botulinum toxin causes large epidemics of botulism in aquatic birds. In the western United States, the disease is the major natural cause of death of wild ducks. Other species of clostridia associated with major diseases in animals are *C. chauvoei*, *C. haemolyticum*, *C. novyi* type B, *C. septicum*, and *C. tetani*. *See* MEDICAL BACTERIOLOGY.

Hilda P. Willett

Bibliography. J. G. Holt et al. (ed.), *Bergey's Manual of Systematic Bacteriology,* vol. 2, 1986; L. S. McClung, *The Anaerobic Bacteria: Their Activities in Nature and Disease,* 7 vols., 1982; M. T. Parker and L. H. Collier, *Topley and Wilson's Principles of Bacteriology, Virology and Immunity,* 2 vols., 8th ed., 1990; M. P. Starr et al., *The Prokaryotes,* vol. 2, 1981.

Cloud

A suspension of minute droplets or ice crystals produced by the condensation of water vapor. This article presents an outline of cloud formation upon which to base an understanding of cloud classifications. For a more technical consideration of the physical character of atmospheric clouds, including the condensation and precipitation of water vapor, *see* CLOUD PHYSICS; FOG.

Rudiments of cloud formation. A grasp of a few physical and meteorological relationships aids in an understanding of clouds. First, if water vapor is cooled sufficiently, it becomes saturated and is in equilibrium with a plane surface of liquid water (or ice) at the same temperature. Further cooling in the presence of such a surface causes condensation upon it; in the absence of any surfaces, no condensation occurs until a substantial further cooling provokes condensation upon ions or random large aggregates of water molecules. In the atmosphere, even in the apparent absence of any surfaces, nuclei always exist upon which condensation proceeds at barely appreciable cooling beyond the state of saturation. Consequently, when atmospheric water vapor is cooled sufficiently, condensation nuclei swell into minute water droplets and form a visible cloud. The total concentration of liquid in the cloud is controlled by its temperature and the degree of cooling beyond the state in which saturation occurred, and in most clouds approximates to 1 g/m^3 (0.001 oz/ft^3) of air. The concentration of droplets is controlled by the concentrations and properties of the nuclei and the speed of the cooling at the beginning of the condensation. In the atmosphere these are such that there are usually about 100,000,000 droplets/m^3 (2,800,000 droplets/ft^3). Because the cloud water is at first fairly evenly shared among them, these droplets are necessarily of microscopic size, and an important part of the study of clouds concerns the ways in which they become aggregated into drops large enough to fall as rain.

The cooling which produces clouds is almost always associated with the upward movements of air which carry heat from the Earth's surface and restore to the atmosphere that heat lost by radiation into space. These movements are most pronounced in storms, which are accompanied by thick, dense clouds, but also take place on a smaller scale in fair weather, producing scattered clouds or dappled skies. *See* STORM.

Rising air cools by as much as 10°C for each kilometer of ascent, so that even over equatorial regions temperatures below 0°C (32°F) are encountered a few kilometers above the ground, and clouds of frozen particles prevail at higher levels. Of the abundant nuclei which facilitate droplet condensation, very few cause direct condensation into ice crystals or stimulate the freezing of droplets, and especially at temperatures near 0°C (32°F) their numbers may be vanishingly small. Consequently, at these temperatures, clouds of unfrozen droplets are not infrequently encountered (supercooled clouds). In general, however, ice crystals occur in very much smaller concentrations than the droplets of liquid clouds, and may by condensation alone become large enough to fall from their parent cloud. Such ice nuclei are particles of clay, or sometimes bacteria,

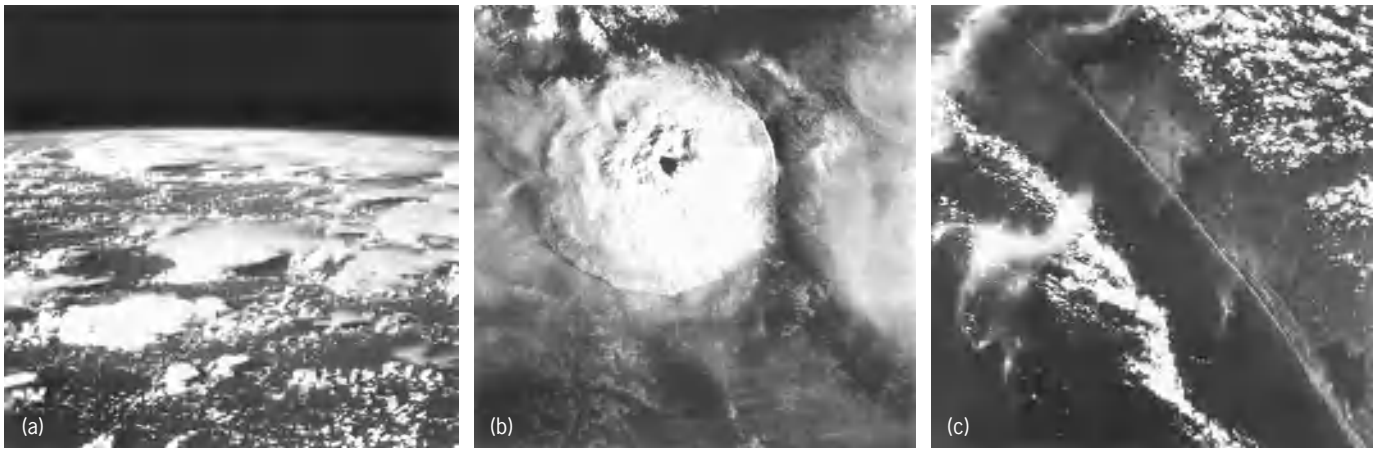

Fig. 1. Clouds as photographed by meteorological satellites. (*a*) A garden of thunderstorms, showing the anvils, forming at the tropopause inversion. (*b*) A hurricane, showing the eye. (*c*) Convective cloud over land and sea-breeze clouds.

mixed upward from the Earth's surface. All cloud droplets freeze by self-nucleation (ice crystal formation by random motion of the water molecules) at temperatures below −35 to −40°C (−31 to −40°F). Even small high clouds may produce or become trails of snow crystals, whereas droplet clouds are characteristically compact in appearance with well-defined edges, and produce rain only when well developed vertically (2 km or 1.5 mi, or more, thick).

Clouds play a major role in the heat balance of the Earth by scattering back sunlight to space. The heat balance is also affected by clouds' absorbing and scattering thermal radiation from regions partly above the cloud. Changes in cloud particle size and concentration by anthropogenic effects—air pollution from various sources—may influence the heat balance, thus offsetting the effects of carbon dioxide warming. *See* CLIMATE MODIFICATION.

Nacreous and noctilucent clouds. Clouds sometimes occur at much higher levels in the atmosphere, particularly when the temperature falls below −90 to −95°C (−130 to −139°F). Nacreous clouds (or polar stratospheric clouds) occur at heights near 25 km (17 mi) in the Arctic and Antarctic winter or early spring. They are visible when their level is illuminated by the Sun from some 20° below the horizon, so that the observer is on the surface in darkness. Such low temperatures occur preferentially in wave disturbances propagating upward and downstream of surface mountain ranges. Particles are composed of sulfuric acid (possibly injected by extreme volcanic events) and nitric acid, water ice, and combinations thereof as hydrates, usually some 1–5 micrometers in diameter. Such clouds are sometimes colored (particularly when particles are unfrozen supercooled or saturated droplets), a phenomenon

Fig. 2. Cirrus, with trails of slowly falling ice crystals at a high level. (*F. Ellerman*, *U.S. Weather Bureau*)

Fig. 3. Small cumulus. (*U.S. Weather Bureau*)

Fig. 4. Overcast of stratus, with some fragments below the hilltops. (*U.S. Weather Bureau*)

Fig. 5. View from Mount Wilson, California. High above is a veil of cirrostratus, and below is the top of a low-level layer cloud. (*F. Ellerman, U.S. Weather Bureau*)

Fig. 6. Cirrocumulus, high clouds with a delicate pattern. (*A. A. Lothman, U.S. Weather Bureau*)

which also occurs in lower-level clouds because of sunlight diffraction from regions containing particles of near-uniform size; hence the alternative name, mother of pearl clouds. At higher levels, near 80 km (50 mi), at the temperature minimum near −80°C (−112°F), noctilucent clouds form and can be viewed under similar circumstances, with viewing time extended for an even lower Sun because of the additional height. The particles tend to be smaller (by about one-tenth); however, these clouds show a much more fibrous structure than nacreous clouds and are not colored. Observation of these clouds is associated with north-south-oriented mountains which give rise to gravity waves in easterly circumpolar flow, being more frequent in Scandinavia, Alaska, and northern Canada. Particles making up these clouds may form the substrate for chemical reactions leading to destruction of ozone.

Classification of clouds. The contrast in cloud forms mentioned above was recognized in the first widely accepted classification, as well as in several succeeding classifications. The first was that of L. Howard in 1803, recognizing three fundamental types: the stratiform (layer), cumuliform (heap), and cirriform (fibrous). The first two are indeed fundamental, representing clouds formed respectively in stable and in convectively unstable atmospheres, whereas the clouds of the third type are the ice clouds which are in general higher and more tenuous and less clearly reveal the kind of air motion which led to their formation. Succeeding classifications continued to be based upon the visual appearance or form of the clouds, differentiating relatively minor features, but later in the nineteenth century increasing importance was attached to cloud height, because direct measurements of winds above the ground were then very difficult, and it was hoped to obtain wind data on a great scale by combining observations of apparent cloud motion with reasonably accurate estimates of cloud height, based solely on their form. With the advent of satellite and rocket observations in the 1970s, a broader perspective became available for examination of clouds on a global basis. Such pictures readily show the distribution of clouds from systems as a midlatitude cyclone, hurricane, or the intertropical convergence zone; they show the influence of land and ocean in initiating convective motion; examples are given in **Fig. 1**. *See* METEOROLOGICAL ROCKET; METEOROLOGICAL SATELLITES; SATELLITE METEOROLOGY.

WMO cloud classification. The World Meteorological Organization (WMO) uses a classification which, with minor modifications, dates from 1894 and represents a choice made at that time from a number of competing classifications. It divides clouds into low-level (base below about 2 km or 1.5 mi), middle-level (about 2 to 7 km or approximately 1 to 4.5 mi), and high-level (between roughly 7 and 14 km or approximately 4 to 8.5 mi) forms within

Fig. 7. Altocumulus, which occurs at intermediate levels. (*G. A. Lott, U.S. Weather Bureau*)

Fig. 8. Altostratus, a middle-level layer cloud. Thick layers of such clouds, with bases extending down to low levels, produce prolonged rain or snow, and are then called nimbostratus. (*C. F. Brooks, U.S. Weather Bureau*)

Fig. 9. Cumulonimbus clouds photographed over the upland adjoining the upper Colorado River valley. (*Lt. B. H. Wyatt, U.S.N., U.S. Weather Bureau*)

the middle latitudes. The names of the three basic forms of clouds are used in combination to define 10 main characteristic forms, or genera.

1. Cirrus are high white clouds with a silken or fibrous appearance (**Fig. 2**).
2. Cumulus are detached dense clouds which rise in domes or towers from a level low base (**Fig. 3**).
3. Stratus are extensive layers or flat patches of low clouds without detail (**Fig. 4**).
4. Cirrostratus forms from an abundance of cirrus as a layer, well above and distinct from the lower layer cloud (**Fig. 5**).
5. Cirrocumulus is formed of high clouds broken into a delicate wavy or dappled pattern (**Fig. 6**).
6. Stratocumulus is a low-level layer cloud having a dappled, lumpy, or wavy structure. See the foreground of Fig. 5.
7. Altocumulus is similar to stratocumulus but lies at intermediate levels (**Fig. 7**).
8. Altostratus is a thick, extensive, layer cloud at intermediate levels (**Fig. 8**).
9. Nimbostratus is a dark, widespread cloud with a low base from which prolonged rain or snow falls.
10. Cumulonimbus is a large cumulus which produces a rain or snow shower (Fig. 1*a*; **Fig. 9**).

Classification by air motion. Modern studies of clouds have been stimulated by the need to know their composition from the viewpoint of aircraft and rocket penetration, which gives rise to icing, turbulence, and lightning hazards, and by the discovery that seeding of supercooled clouds could, on occasion, give rise to enhanced precipitation. *See* WEATHER MODIFICATION.

These studies show that the external form of clouds gives only indirect and incomplete clues to the physical properties which determine their evolution. Throughout this evolution the most important properties appear to be the air motion and the size-distribution spectrum of all the cloud particles, including the condensation and ice-forming nuclei. These properties vary significantly with time and position within the cloud, so that cloud studies demand the intensive examination of individual clouds with aircraft, radar, and satellites. Nevertheless, the overall cloud shape does give information on the formation process. Lenticular clouds with smooth tops may form from smooth air flow with little environmental mixing (**Fig. 10**); clouds with irregular tops (Figs. 3

Cloud classification based on air motion and associated physical characteristics

Kind of motion	Typical vertical speeds, cm/s*	Kind of cloud	Name	Characteristic dimensions, km†		Characteristic precipitation
				Horizontal	Vertical	
Widespread slow ascent, associated with cyclones (stable atmosphere)	1–10	Thick layers	Cirrus, later becoming:			
			cirrostratus, altostratus, altocumulus	10^3	1–2	Snow trails
			nimbostratus	10^3	10	Prolonged moderate rain or snow
Convection, due to passage over warm surface (unstable atmosphere)	10^2	Small heap cloud	Cumulus	1	1	None
	10^3	Shower- and thundercloud	Cumulonimbus	10	10	Intense showers of rain or hail
Irregular stirring causing cooling during passage over cold surface (stable atmosphere)	10–10^2	Shallow low layer clouds, fogs	Stratus	$<10^2$	<1	None, or slight drizzle or snow
			Stratocumulus	$<10^3$		

*1 cm = 0.4 in. †1 km = 0.62 mi.

Fig. 10. A cloud formed by moist air flowing over a mountain. Note the smooth upper profile, indicating low-turbulence flow. (*Courtesy of John Hallet*)

and 9) show that mixing is strong. The former has nonturbulent, laminar air motions and give a smooth flight; the mixing clouds give modest to strong turbulence and a bumpy flight. An observer of clouds can readily see motions of convective clouds; motions in other systems can be viewed by time-lapse photography at 5–50-s intervals, or sequential satellite pictures at intervals of several hours. From a general meteorological point of view, a classification can be based upon the kind of air motion associated with the cloud, as shown in the **table**.

Frank H. Ludlam; John Hallett

Cloud physics

The study of the physical and dynamical processes governing the structure and development of clouds and the release from them of snow, rain, and hail (collectively known as precipitation). *See* PRECIPITATION (METEOROLOGY).

The factors of prime importance are the motion of the air, its water-vapor content, and the number and properties of the particles in the air which act as centers of condensation and freezing. Because of the complexity of atmospheric motions and the enormous variability in vapor and particle content of the air, it seems impossible to construct a detailed, general theory of the manner in which clouds and precipitation develop. However, calculations based on the present conception of laws governing the growth and aggregation of cloud particles and on simple models of air motion provide reasonable explanations for the observed formation of precipitation in different kinds of clouds.

Cloud formation. Clouds are formed by the lifting of damp air which cools by expansion under continuously falling pressure. The relative humidity increases until the air approaches saturation. Then condensation occurs (**Fig. 1**) on some of the wide variety of aerosol particles present; these exist in concentrations ranging from less than 2000 particles/in.^{3} (100/cm^3) in clean, maritime air to perhaps 10^7/in.^{3} (10^6/cm^3) in the highly polluted air of an industrial city. A portion of these particles are hygroscopic and promote condensation at relative humidities below 100%; but for continued condensation leading to the formation of cloud droplets, the air must be slightly supersaturated. Among the highly efficient condensation nuclei are the salt particles produced by the evaporation of sea spray, but it appears that particles produced by human-made fires and by natural combustion (for example, forest fires) also make a major contribution. Condensation onto the nuclei continues as rapidly as the water vapor is made available by cooling of the air and gives rise to droplets of the order of 0.0004 in. (0.01 mm) in diameter. These droplets, usually present in concentrations of several thousand per cubic inch, constitute a nonprecipitating water cloud.

Mechanisms of precipitation release. Growing clouds are sustained by upward air currents, which may vary in strength from about an inch per second to several yards per second. Considerable growth of the cloud droplets (with falling speeds of only about 0.4 in./s or 1 cm/s) is therefore necessary if they are to fall through the cloud, survive evaporation in the unsaturated air beneath, and reach the ground as drizzle or rain. Drizzle drops have radii exceeding 0.004 in. (0.1 mm), while the largest raindrops are about 0.24 in. (6 mm) across and fall at nearly 30 ft/s (10 m/s). The production of a relatively few large particles from a large population of much smaller ones may be achieved in one of two ways.

Coalescence process. Cloud droplets are seldom of uniform size for several reasons. Droplets arise on nuclei of various sizes and grow under slightly different conditions of temperature and supersaturation in different parts of the cloud. Some small drops may remain inside the cloud for longer than others before being carried into the drier air outside.

A droplet appreciably larger than average will fall faster than the smaller ones, and so will collide and fuse (coalesce) with some of those which it overtakes (**Fig. 2**). Calculations show that, in a deep cloud containing strong upward air currents and

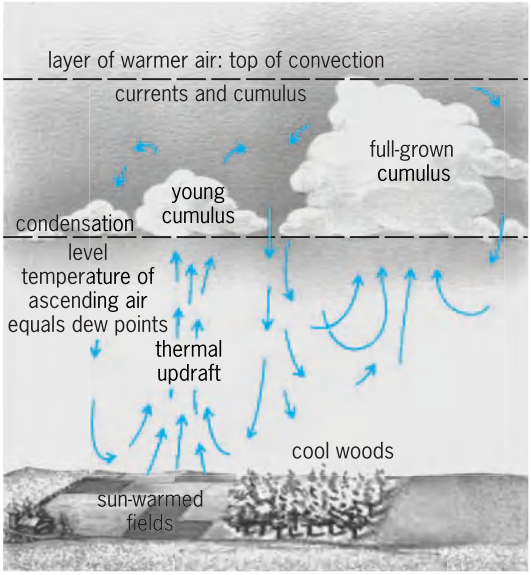

Fig. 1. Conditions leading to birth of a cumulus cloud.

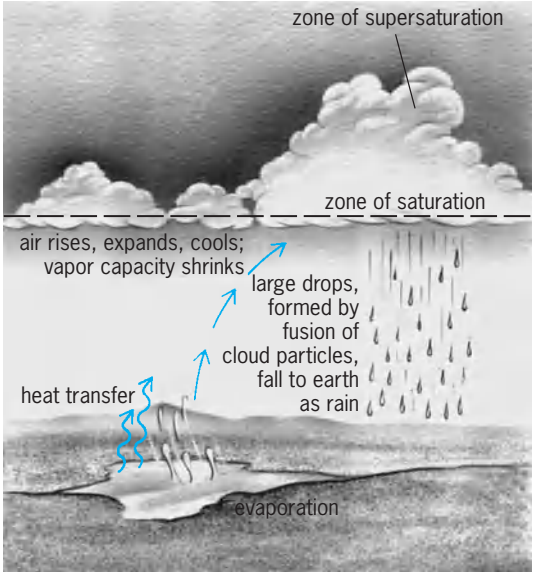

Fig. 2. Diagram of the steps in the formation of rain.

high concentrations of liquid water, such a droplet will have a sufficiently long journey among its smaller neighbors to grow to raindrop size. This coalescence mechanism is responsible for the showers that fall in tropical and subtropical regions from clouds whose tops do not reach the 32°F (0°C) level and therefore cannot contain ice crystals which are responsible for most precipitation. Radar evidence also suggests that showers in temperate latitudes may sometimes be initiated by the coalescence of waterdrops, although the clouds may later reach to heights at which ice crystals may form in their upper parts.

Initiation of the coalescence mechanism requires the presence of some droplets exceeding 20 micrometers in diameter. Over the oceans and in adjacent land areas they may well be supplied as droplets of sea spray, but in the interiors of continents, where so-called giant salt particles of marine origin are probably scarce, it may be harder for the coalescence mechanism to begin.

Ice crystal process. The second method of releasing precipitation can operate only if the cloud top reaches elevations where temperatures are below 32°F (0°C) and the droplets in the upper cloud regions become supercooled. At temperatures below −40°F (−40°C) the droplets freeze automatically or spontaneously; at higher temperatures they can freeze only if they are infected with special, minute particles called ice nuclei. As the temperature falls below 32°F (0°C), more and more ice nuclei become active, and ice crystals appear in increasing numbers among the supercooled droplets. But such a mixture of supercooled droplets and ice crystals is unstable. The cloudy air, being usually only slightly supersaturated with water vapor as far as the droplets are concerned, is strongly oversaturated for the ice crystals, which therefore grow more rapidly than the droplets. After several minutes the growing crystals will acquire definite falling speeds, and several of them may become joined together to form a snowflake. In falling into the warmer regions of the cloud, however, the snowflake may melt and reach the ground as a raindrop.

Precipitation from layer-cloud systems. The deep, extensive, multilayer-cloud systems, from which precipitation of a usually widespread, persistent character falls, are generally formed in cyclonic depressions (lows) and near fronts. Such cloud systems are associated with feeble upcurrents of only a few centimeters per second, which last for at least several hours. Although the structure of these great raincloud systems, which are being explored by aircraft and radar, is not yet well understood, it appears that they rarely produce rain as distinct from drizzle, unless their tops are colder than about 10°F (−12°C). This suggests that ice crystals may be responsible. Such a view is supported by the fact that the radar signals from these clouds usually take a characteristic form which has been clearly identified with the melting of snowflakes.

Production of showers. Precipitation from shower clouds and thunderstorms, whether in the form of raindrops, pellets of soft hail, or true hailstones, is generally of greater intensity and shorter duration than that from layer clouds and is usually composed of larger particles. The clouds themselves are characterized by their large vertical depth, strong vertical air currents, and high concentrations of liquid water, all these factors favoring the rapid growth of precipitation elements by accretion.

In a cloud composed wholly of liquid water, raindrops may grow by coalescence with small droplets. For example, a droplet being carried up from the cloud base would grow as it ascends by sweeping up smaller droplets. When it becomes too heavy to be supported by the vertical upcurrents, the droplet will then fall, continuing to grow by the same process on its downward journey. Finally, if the cloud is sufficiently deep, the droplet will emerge from its base as a raindrop.

In a dense, vigorous cloud several kilometers deep, the drop may attain its limiting stable diameter (about 0.2 in. or 5 mm) before reaching the cloud base and thus will break up into several large fragments. Each of these may continue to grow and attain breakup size. The number of raindrops may increase so rapidly in this manner that after a few minutes the accumulated mass of water can no longer be supported by the upcurrents and falls out as a heavy shower. The conditions which favor this rapid multiplication of raindrops occur more readily in tropical regions.

The ice crystals grow initially by sublimation of vapor in much the same way as in layer clouds, but when their diameters exceed about 0.004 in. (0.1 mm), growth by collision with supercooled droplets will usually predominate. At low temperatures the impacting droplets tend to freeze individually and quickly to produce pellets of soft hail. The air spaces between the frozen droplets give the ice a relatively low density; the frozen droplets contain large numbers of tiny air bubbles, which give the pellets an opaque, white appearance. However, when the

growing pellet traverses a region of relatively high air temperature or high concentration of liquid water or both, the transfer of latent heat of fusion from the hailstone to the air cannot occur sufficiently rapidly to allow all the deposited water to freeze immediately. There then forms a wet coating of slushy ice, which may later freeze to form a layer of compact, relatively transparent ice. Alternate layers of opaque and clear ice are characteristic of large hailstones, but their formation and detailed structure are determined by many factors such as the number concentration, size and impact velocity of the supercooled cloud droplets, the temperature of the air and hailstone surface, and the size, shape, and aerodynamic behavior of the hailstone. Giant hailstones, up to 4 in. (10 cm) in diameter, which cause enormous damage to crops, buildings, and livestock, most frequently fall not from the large tropical thunderstorms, but from storms in the continental interiors of temperate latitudes. An example is the Nebraska-Wyoming area of the United States, where the organization of larger-scale wind patterns is particularly favorable for the growth of severe storms. *See* HAIL.

The development of precipitation in convective clouds is accompanied by electrical effects culminating in lightning. The mechanism by which the electric charge dissipated in lightning flashes is generated and separated within the thunderstorm has been debated for more than 200 years, but there is still no universally accepted theory. However, the majority opinion holds that lightning is closely associated with the appearance of the ice phase, and the most promising theory suggests that the charge is produced by the rebound of a small fraction of the supercooled cloud droplets that collide with the falling hail pellets. *See* LIGHTNING.

Basic aspects of cloud physics. The various stages of the precipitation mechanisms raise a number of interesting and fundamental problems in classical physics. Worthy of mention are the supercooling and freezing of water; the nature, origin, and mode of action of the ice nuclei; and the mechanism of ice-crystal growth which produces the various snow crystal forms.

It has been established how the maximum degree to which a sample of water may be supercooled depends on its purity, volume, and rate of cooling. The freezing temperatures of waterdrops containing foreign particles vary linearly as the logarithm of the droplet volumes for a constant rate of cooling. This relationship, which has been established for drops varying between 0.0004 and 0.4 in. (10 μm and 1 cm) in diameter, characterizes the heterogeneous nucleation of waterdrops and is probably a consequence of the fact that the ice-nucleating ability of atmospheric aerosol increases logarithmically with decreasing temperature.

When extreme precautions are taken to purify the water and to exclude all solid particles, small droplets, about 0.00004 in. (1 μm) in diameter, may be supercooled to -40°F (-40°C) and drops of 0.04 in. (1 mm) diameter to -31°F (-35°C). Under these conditions freezing occurs spontaneously without the aid of foreign nuclei.

The nature and origin of the ice nuclei, which are necessary to induce freezing of cloud droplets at temperatures about -40°F (-40°C), are still not clear. Measurements made with large cloud chambers on aircraft indicate that the most efficient nuclei, active at temperatures above 14°F (-10°C), are present in concentrations of only about 30/ft^3 (10/m^3) of air, but as the temperature is lowered, the number of ice crystals increases logarithmically to reach concentrations of about 25/ft^3 (1/liter) at -4°F (-20°C) and 2500/ft^3 (100/liter) at -22°F (-30°C). Since these measured concentrations of nuclei are less than one-hundredth of the number that apparently is consumed in the production of snow, it seems that there must exist processes by which the original number of ice crystals are rapidly multiplied. Laboratory experiments suggest the fragmentation of the delicate snow crystals and the ejection of ice splinters from freezing droplets as probable mechanisms.

The most likely source of atmospheric ice nuclei is provided by the soil and mineral-dust particles carried aloft by the wind. Laboratory tests have shown that, although most common minerals are relatively inactive, a number of silicate minerals of the clay family produce ice crystals in a supercooled cloud at temperatures above -0.4°F (-18°C). A major constituent of some clays, kaolinite, which is active below 16°F (-9°C), is probably the main source of highly efficient nuclei.

The fact that there may often be a deficiency of efficient ice nuclei in the atmosphere has led to a search for artificial nuclei which might be introduced into supercooled clouds in large numbers. Silver iodide is a most effective substance, being active at 25°F (-4°C), while lead iodide and cupric sulfide have threshold temperatures of 21°F (-6°C) for freezing nuclei.

In general, the most effective ice-nucleating substances, both natural and artificial, are hexagonal crystals in which spacings between adjacent rows of atoms differ from those of ice by less than 16%. The detailed surface structure of the nucleus, which is determined only in part by the crystal geometry, is of even greater importance. This is strongly indicated by the discovery that several complex organic substances, notably steroid compounds, which have apparently little structural resemblance to ice, may act as nucleators for ice at temperatures as high as 30°F (-1°C).

The collection of snow crystals from clouds at different temperatures has revealed their great variety of shape and form. By growing the ice crystals on a fine fiber in a cloud chamber, it has been possible to reproduce all the naturally occurring forms and to show how these are correlated with the temperature and supersaturation of the environment. With the air temperature along the length of a fiber ranging from 32 to -13°F (0 to -25°C), the following clear-cut changes of crystal habit are observed (**Fig. 3**): hexagonal plates–needles–hollow prisms–plates–stellar dendrites–plates–prisms. This multiple change of habit over such a small

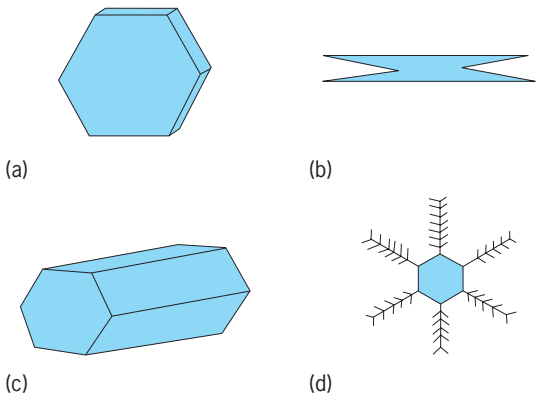

Fig. 3. Ice crystal types formed in various temperature ranges: (*a*) Thin hexagonal plate at 32 to 27°F (0 to −3°C), 18 to 10°F (−8 to −12°C), 3 to −13°F (−16 to −25°C). (*b*) Needles at 27 to 23°F (−3 to −5°C). (*c*) Hexagonal prismatic column at 23 to 18°F (−5 to −8°C), below −13°F (−25°C). (*d*) Dendritic star-shaped crystal at 10 to 3°F (−12 to −16°C).

temperature range is remarkable and is thought to be associated with the fact that water molecules apparently migrate between neighboring faces on an ice crystal in a manner which is very sensitive to the temperature. Certainly the temperature rather than the supersaturation of the environment is primarily responsible for determining the basic shape of the crystal, though the supersaturation governs the growth rates of the crystals, the ratio of their linear dimensions, and the development of dendritic forms. *See* SNOW.

Artificial stimulation of rain. The presence of either ice crystals or some comparatively large waterdroplets (to initiate the coalescence mechanism) appears essential to the natural release of precipitation. Rainmaking experiments are conducted on the assumption that some clouds precipitate inefficiently, or not at all, because they are deficient in natural nuclei; and that this deficiency can be remedied by seeding the clouds artificially with dry ice or silver iodide to produce ice crystals, or by introducing waterdroplets or large hygroscopic nuclei. In the dry-ice method, pellets of about 0.4-in. (1-cm) diameter are dropped from an aircraft into the top of a supercooled cloud. Each pellet chills a thin sheath of air near its surface to well below −40°F (−40°C) and produces perhaps 10^{12} minute ice crystals, which subsequently spread through the cloud, grow, and aggregate into snowflakes. Only a few pounds of dry ice are required to seed a large cumulus cloud. Some hundreds of experiments, carried out mainly in Australia, Canada, South Africa, and the United States, have shown that cumulus clouds in a suitable state of development may be induced to rain by seeding them with dry ice on occasions when neighboring clouds, untreated, do not precipitate. However, the amounts of rain produced have usually been rather small.

For large-scale trials designed to modify the rainfall from widespread cloud systems over large areas, the cost of aircraft is usually prohibitive. The technique in this case is to release a silver iodide smoke from the ground and rely on the air currents to carry it up into the supercooled regions of the cloud. In this method, with no control over the subsequent transport of the smoke, it is not possible to make a reliable estimate of the concentrations of ice nuclei reaching cloud level, nor is it known for how long silver iodide retains its nucleating ability in the atmosphere. It is usually these unknown factors which, together with the impossibility of estimating accurately what would have been the natural rainfall in the absence of seeding activities, make the design and evaluation of a large-scale operation so difficult.

Little convincing evidence can be found that large increases in rainfall have been produced consistently over large areas. Indeed, in temperate latitudes most rain falls from deep layer-cloud systems whose tops usually reach to levels at which there are abundant natural ice nuclei and in which the natural precipitation processes have plenty of time to operate. It is therefore not obvious that seeding of these clouds would produce a significant increase in rainfall, although it is possible that by forestalling natural processes some redistribution might be effected.

Perhaps more promising as additional sources of rain or snow are the persistent supercooled clouds produced by the ascent of damp air over large mountain barriers. The continuous generation of an appropriate concentration of ice crystals near the windward edge might well produce a persistent light snowfall to the leeward, since water vapor is continually being made available for crystal growth by lifting of the air. The condensed water, once converted into snow crystals, has a much greater opportunity of reaching the mountain surface without evaporating, and might accumulate in appreciable amounts if seeding were maintained for many hours.

Trials carried out in favorable locations in the United States and Australia suggest that in some cases seeding has been followed by seasonal precipitation increases of about 10%, but rarely have the effects been reproduced from one season to the next, and overall the evidence for consistent and statistically significant increases of rainfall is not impressive. Indeed, as the experiments have been subjected to stricter statistical design and evaluation, the claims have steadily became more modest, and the difficulty of improving on natural processes has become increasingly apparent.

During the 1960s and 1970s, remaining trials were carried out in some 75 countries, and at one time about 25% of the land area of the United States was being seeded. However, by the late 1980s the activity was much reduced and had virtually ceased in Australia and many other countries. Nevertheless, China and Israel persevered. The Israeli experiments were unique in that they appeared to have produced rainfall increases on the order of 15% year after year for about 15 years. No convincing explanation for this has yet emerged. *See* CLOUD; WEATHER MODIFICATION.

Basil J. Mason

Bibliography. E. M. Agee and T. Assai (eds.), *Cloud Dynamics*, 1982; L. J. Battan, *Cloud Physics and Cloud Seeding*, 1962, reprint 1979; B. J. Mason,

Clouds, *Rain and Rainmaking*, 1975; H. R. Pruppacher and J. D. Klett, *Microphysics of Atmospheric Clouds and Precipitation*, 2d ed., 1997.

Clove

The unopened flower bud (see **illus.**) of a small, conical, symmetrical, evergreen tree, *Eugenia caryophyllata*, of the myrtle family (Myrtaceae). The cloves are picked by hand and dried in the sun or by artificial means. The crop is uncertain and difficult to grow. Cloves, one of the most important and useful spices, are strongly aromatic and have a pungent flavor. They are used as a culinary spice for flavoring pickles, ketchup, and sauces, in medicine, and for perfuming the breath and the air in rooms. The essential oil distilled from cloves by water or steam has even more uses. The chief clove-producing countries are Tanzania with 90% of the total output, Indonesia, Mauritius, and the West Indies. *See* MYRTALES; SPICE AND FLAVORING.

Closed (cloves) and open flower buds on a branch of the evergreen tree *Eugenia caryophyllata*.

Perry D. Strausbaugh; Earl L. Core

Clover

A common name used loosely to designate the true clovers, sweet clovers, and other members of the plant family Leguminosa. This article discusses true clovers, sweet clover, and clover diseases.

True Clovers

The true clovers are plants of the genus *Trifolium*, order Rosales. There are approximately 250 species in the world. Collectively they represent the most important genus of forage legumes in agriculture. Different species constitute one or more crops on every continent. As indicated by the number of species and their ecotypes (subspecies) that have been collected and described, the center of origin appears to be southeastern Europe and southwestern Asia Minor. Many species are found in countries that border the Mediterranean Sea. About 80 species are native to the United States, most of them occurring in the general regions of the Rocky and Cascade mountains and the Sierra Nevada. Although these native clovers are not used as crop plants, they contribute to the range for grazing and wild hay and also supply nitrogen to associated grasses. Of those that are now named, it is possible that many are variants of other species. Only a few are native to the humid Eastern states. Most clovers are highly palatable and nutritious to livestock. The name clover is often applied to members of legume genera other than *Trifolium*. *See* ROSALES.

Characteristics. The true clovers are herbaceous annual or perennial plants. However, many perennial species behave as biennials or annuals under attacks of disease or insects, unfavorable climatic or soil conditions, or improper management. In general, clovers thrive under cool, moist conditions although one native species, desert clover (*T. gymnocarpum*), tolerates semiarid conditions. Annual species usually behave as winter annuals where winter conditions are not severe, and as summer annuals at northern latitudes and high altitudes.

Clovers grow best in areas where adequate supplies of calcium, phosphorus, and potassium are naturally present in the soil or where these elements are applied in limestone and fertilizers. They grow from a few inches to several feet tall, depending on the species and the environmental conditions. The leaves have three to five roundish to spearlike leaflets. The florets, ranging from 5 to 200 in number, are borne in heads. Colors include white, pink, red, purple, and yellow, and various mixtures. The basic haploid chromosome numbers of the species are 6, 7, and 8. Somatic chromosomes range from 12 to about 130, forming diploids, tetraploids, and polyploids. The flowers of some species are self-sterile and must be cross-pollinated before seed will form; the flowers of other species may be self-fertile and capable of self-pollination; others are self-fertile but do not set seed unless the flowers are tripped and the pollen scattered onto the stigma. Cross-pollination and tripping of flowers are effected principally by bees that visit the flowers for nectar and pollen. Seed color ranges from yellow to deep purple, some being bicolored. There are approximately 60,000–700,000 seeds per pound.

Uses. Clovers are used for hay, pasture, silage, and soil improvement. Certain kinds may be used for all purposes whereas others, because of their low growth, are best suited for grazing. All kinds, when well grown in thick stands, are good for soil improvement. Thoroughly inoculated plants add 50–200 lb of nitrogen per acre (56–222 kg per hectare) when plowed under for soil improvement, the amount added depending on growth, thickness of stand, and length of growing season. Clovers may be grown alone, in combination with grasses and other legumes, or with small grains. In the humid states, or where irrigated, clovers are sown most

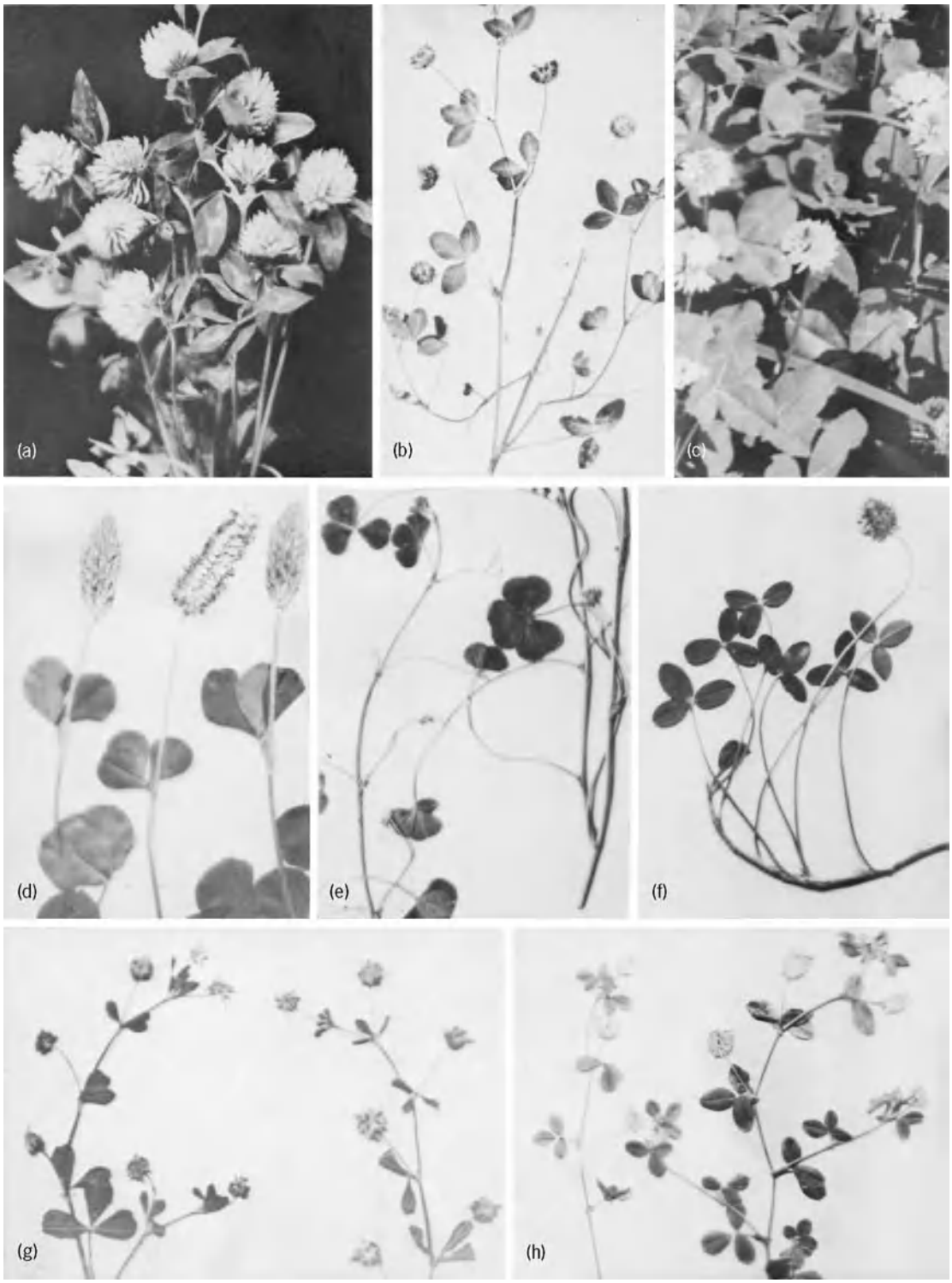

Fig. 1. Important clovers. (*a*) Red clover (*Trifolium pratense*). (*b*) Alsike clover (*T. hybridum*). (*c*) White clover (*T. repens*). (*d*) Crimson clover (*T. incarnatum*). (*e*) Subclover (*T. subterraneum*). (*f*) Strawberry clover (*T. fragiferum*). (*g*) Persian clover (*T. resupinatum*). (*h*) Large hop clover (*T. campestre* or *procumbens*). (*USDA*)

frequently with small-grain companion crops. In the Corn Belt they are generally spring-seeded, whereas in the Southern states and in the West Coast region they should be seeded in the fall for best results.

Important species. All the clover species of agricultural importance in the United States are introduced (exotic) plants. **Figure 1** shows some of the species most widely used: red clover (*T. pratense*), alsike clover (*T. hybridum*), white clover (*T. repens*), crimson clover (*T. incarnatum*), subclover (*T. subterraneum*), strawberry clover (*T. fragiferum*), persian clover (*T. resupinatum*), and large hop clover (*T. campestre*, or *procumbens*). Other clovers of

regional importance, mostly adapted to specific environmental conditions, are rose clover (*T. hirtum*), berseem clover (*T. alexandrinum*), ball clover (*T. nigrescens*), lappa clover (*T. lappaceum*), big-flower clover (*T. michelianum*), and arrowleaf clover (*T. vesiculosum*).

Species characteristics. The following sections discuss the characteristics of the most important species.

Red clover. Red clover is composed of two forms, medium and mammoth, producing two and one hay cuts, respectively. The large purplish-red flower heads are round. An upright-growing perennial, red clover generally persists for 2 years in the northern United States, but behaves as a winter annual in the South. When planted alone or with small grain, seeding rates are 10–15 lb/acre (11–17 kg/hectare). When seeded with grass and other mixtures, 4–8 lb/acre (4.9–9.0 kg/ha) is sufficient. Under favorable growth conditions, seed yields average 70 lb/acre (79 kg/ha); in the West under irrigation they may reach 600–800 lb/acre (675–900 kg/ha). There are several varieties and strains of red clover such as Kenland, Pennscott, Lakeland, Dollard, and Chesapeake. These produce higher yields of forage and are more persistent than common red clover.

Alsike clover. Alsike clover is an upright-growing species that behaves like a biennial. The growth pattern, seeding methods, mixtures, and uses are similar to those of red clover. The flower heads are much like those of white clover in shape and size, but are slightly more pinkish. Alsike clover is more tolerant of wet, poorly drained soils than red clover and occurs widely in mountain meadows of the West.

White clover. White clover, an inhabitant of lawns and closely grazed pastures, is the most important pasture legume in the humid states. The flowers are generally white, but sometimes they are tinged with pink. The stems grow on the soil surface and root at the joints (nodes) as the plant spreads. There are three main types, large, intermediate, and small, with all gradations between. Ladino is the best-known and most widely seeded variety of the large type, and Louisiana Sl represents the intermediate type. The small type is found principally in pastures that are continuously and closely grazed. All types are nutritious and are relished by all classes of livestock and poultry. The protein content ranges from 15 to 30%, depending on plant age and the succulence of growth. Ladino white clover is sometimes used for silage and hay. White clover is mainly grown with low-growing grasses, not being tolerant of the tall-growing kinds. Seeding rates vary from $^1/_2$ to 2 lb/acre (0.6 to 2.3 kg/ha) depending on the seed mixture. For successful growth, it is important to use the variety or strain best adapted to the particular climatic and soil conditions. Limestone and fertilizer must usually be used to get a thick stand and good growth. Spring seeding is recommended for the Northern states and fall seeding for regions having mild winters. For best growth of the clover, grass-clover mixtures should be grazed or cut frequently.

Crimson clover. Crimson clover is used principally as a winter annual for pasture and as a soil-improving crop from the latitude of the Ohio River southward, and along the West Coast. The yellow seed, approximately 120,000 per pound (260,000 per kilogram), are planted at the rate of 15–20 lb/acre (17–22 kg/ha) during the late summer and fall months in a prepared seedbed or on grass which is closely grazed or cut. The seedlings make a rosette of leafy growth during the winter months. With the advent of spring, growth is rapid; the plant reaches $1^1/_2$–3 ft (0.4–0.9 m) in height and the flower stems elongate, terminating with pointed crimson flower heads in late spring or early summer. The plant dies when the seed matures. Seed yields average about 250 lb/acre (280 kg/ha) even though large quantities are lost by shattering. Crimson clover is seeded alone, with small grains and grasses, or on grass turf. During the winter it may be grazed when growth reaches 4 in. (10 cm), although if it is too heavily grazed, regrowth is slow. The greatest return for soil improvement is obtained when the largest growth is plowed under. There are several varieties of crimson clover, including Dixie, Auburn, Autauga, Chief, Talladega, and many local strains. When conditions are favorable, all of these will reseed, forming volunteer stands in the fall from seed shattered the previous spring. Common crimson clover does not reseed. When used with Bermuda or other perennial summer-growing grasses, the tall growth of the grass must be closely grazed or clipped.

Subclover. Subclover, a winter annual extensively used for grazing in the coastal sections of the Western states, is the basic pasture crop of the sheep and cattle industry of Australia. The seed are blue-black and the largest of any clover species (about 60,000 per pound or 132,000 per kilogram). The flower heads on decumbent stems are inconspicuous. As the seed develop, fishhook-shaped appendages are produced which, by their twisting action, pull them into the soil, and hence the name subclover. Seed production is large, but the seed are difficult to harvest. When once established, fall stands develop from seed produced the preceding months. The Australian varieties Mount Barker, Tallarook, and Nangeela have proved to be best adapted to most conditions in the United States. Subclover appears to have considerable promise as a pasture legume under many conditions in the southern United States, but better adapted varieties are needed.

Eugene A. Hollowell

Berseem clover. Berseem, or Egyptian clover (*T. alexandrinum*), is used as a cool-season annual forage crop in the states of Washington, Oregon, California, and Arizona and in the southeastern United States, and as a summer annual in northern areas. It is an upright-growing legume with oblong leaflets and hollow stems. It produces yellowish-white florets in heads and has short tap roots. In southern Florida, plants may obtain a height of 18–24 in. (45–60 cm) during a growing season from October to May and heights of 34–42 in.

(19–105 cm) in Arizona. Berseem clover's greatest potential is probably as green-chopped forage or pasture. It provides high-quality forage with a crude protein concentration of 20–24% on a dry-weight basis. Berseem clover produces an abundance of seed with yields exceeding 990 lb/acre (1100 kg/ha). Berseem flowers are essentially self-sterile, with cross pollination being accomplished by honeybees.

Cultivars of berseem clover are classified according to their branching behavior; branching influences the number of cuttings that can be made. The cultivars Miscawi and Kahdrawi, which exhibit basal branching, can yield four cuttings per growing season, whereas the Saidi produces both basal and apical branching and can be cut twice. The Fahl cultivar exhibits apical branching and can be cut only once; Nile and Hustler cultivars were developed in the United States and are equal to or superior to Miscawi. A somewhat hardier cultivar Multicut, which tolerates temperatures down to 23°F (−5°C), was released in 1988 for use in parts of California and Mexico; it produces five or six cuttings under irrigation. The release in 1984 of the winter-hardy, reseeding cultivar Bigbee, extended the range of adaptation of berseem clover to Tennessee and Oklahoma and enabled producers to obtain reseeding stands. Prior to the release of Bigbee, berseem clover was considered to be a nonwinter-hardy, nonreseeding annual forage legume. W. E. Knight

Other clovers. Strawberry clover, a perennial, is a nutritious pasture plant similar in growth habit to white clover, for which it is frequently mistaken. The pinkish flower head looks like a strawberry, and hence the common name. It is highly tolerant of the wet, salty soil common to the Western states.

Persian clover is a winter annual used mostly for pasture and soil improvement but is also an excellent silage plant. It is particularly adapted to the wet, heavy soils of the lower southern part of the United States. The flowers are light purple, and the mature seed are enclosed in balloonlike capsules that shatter easily, float on water, and are readily blown about by the wind. Persian clover is relished by all classes of livestock.

Large hop clover and small hop clover, widely distributed as winter annuals in the Southern states, are used in pastures. They are more tolerant of low soil fertility than the other species and less productive, but are highly palatable to livestock. The flower heads are bright yellow and, when mature, look like the flowers of the hop plant. The two species are similar.

Rose clover is a relatively new winter annual and appears to be best adapted to the foothill rangelands of California.

Other species are ball clover, lappa clover, big-flower clover, and arrowleaf clover, all winter annuals. Lappa clover grows best in the heavy dark marl soils of the South, whereas ball, big-flower, and arrowleaf clovers appear to be more widely adapted. All are used for pasture, although big-flower and arrowleaf clovers may also be used for hay.

Seaside clover (*T. willdenovii*), white-tipped clover (*T. variegatum*), and long-stalked clover (*T. longipes*) are the most widely distributed of the native western species. All produce thick stands and good growth under varying conditions. *See* GENETICS; LEGUME FORAGES; REPRODUCTION (PLANT).

Sweet Clover

Sweet clover is the common name applied to all but one species of legumes of the genus *Melilotus*, order Rosales. The exception is sour clover (*M. indica*).

Origin and distribution. There are approximately 20 species of sweet clover. Some of the biennial species have an annual form. The roots of the biennials develop crown buds from which the second-year growth arises; the roots of annuals do not develop crown buds. Sweet clovers are native to the Mediterranean region and adjacent countries, but several are widely scattered throughout the world, generally by chance introduction. None is indigenous to the United States. White sweet clover (*M. alba*) and yellow sweet clover (*M. officinalis*) are important forage and soil-improvement plants in the United States and Canada and are found growing along roadsides and in waste places in every state (**Fig. 2**). Sour clover, a yellow-flowered winter annual, is of some value for soil improvement only along the Gulf Coast and in southern New Mexico, Arizona, and California. Improved varieties of white and yellow sweet clovers are available for farm use.

Uses. Sweet clover is used as a field crop in regions of the United States and Canada where the rainfall is 17 in. (42 cm) or more during the growing season, where the soil is neutral, or where

Fig. 2. White sweet clover (*Melilotus alba*). (*USDA*)

limestone and other needed minerals are applied. It is most extensively grown in the Great Plains and Corn Belt, either alone or in rotations with small grains and corn, and is used for grazing, soil improvement, and hay.

Except for those of certain improved varieties, the plants are somewhat bitter because of the presence of coumarin. External and internal bleeding of animals may result from feeding spoiled sweet-clover hay or improperly preserved silage containing sweet clover, a decomposition product of coumarin (4-hydroxycoumarin, commonly called dicumarol), which develops during spoilage, being the toxic principle. Research has led to its use in medicine and for making warfarin, a rodenticide.

Eugene A. Hollowell

Diseases

Diseases of clovers are caused by bacteria, fungi, nematodes, viruses, a mycoplasma, and air pollutants. These diseases reduce forage yield, quality, and stand. Frequently, perennial clovers are productive for only 1 or 2 years because of diseases. Crop losses are best minimized by growing adapted cultivars that are tolerant of or resistant to one or more of the major diseases. Fungi and viruses are the most important pathogens.

Fig. 3. Red clover roots cut longitudinally to show type of injury caused by crown and root rot organisms.

Fig. 4. Typical petiole and stem lesions of northern anthracnose disease of red clover, caused by the fungus *Kabatiella caulivora*.

Fungal diseases. Fungi, including *Pythium, Fusarium*, and *Sclerotinia*, cause root and crown rots that reduce stand establishment and persistence (**Fig. 3**). These pathogens can attack clover plants at any developmental stage. Plants not killed are weakened, and are thus predisposed to further injury from physical and biological stress.

Fungi, including *Stemphylium, Colletotrichum*, and *Kabatiella*, cause leaf and stem diseases, which usually do not kill the plants but reduce their productive and competitive capabilities (**Fig. 4**).

Viral diseases. Most clovers are susceptible to destructive viruses that cause leaf mottling and distortion and reduced plant vigor. Viruses may be seed-borne or spread by insects and mowing. *See* PLANT PATHOLOGY; PLANT VIRUSES AND VIROIDS.

Kenneth T. Leath

Bibliography. N. C. Brady (ed.), *Adv. Agron.*, 35:165–191, 1982; W. L. Graves et al., Registration of Multicut berseem clover, *Crop Sci.*, 29:235–236, 1989; M. E. Heath, D. S. Metcalfe, and R. F. Barnes (eds.), *Forages: The Science of Agriculture*, 4th ed., 1985; W. E. Knight, Registration of Bigbee berseem clover, *Crop Sci.*, 25:571–572, 1985; N. L. Taylor (ed.), *Agronomy Monograph*, no. 24, ASA-CSSA-SSSA, 1984; J. L. Wheeler and R. D. Mochrie, *Forage Evaluation*: *Concepts and Techniques*, 1982.

Clupeiformes

The single order of the teleost fishes in the subdivision Clupeomorpha. Clupeiforms, including anchovies, herrings, sardines, shads, and allies, are classified in two suborders, five families, 83 genera, and about 357 species. A unique feature of clupeiforms is the ear–swim bladder connection (otophysic connection) composed of a pair of anterior extensions of the swim bladder that enter the skull to connect with the utriculus of the inner ear. They are further

distinguished by a series of median scutes (keel-like scales) along the abdomen before and behind the pelvic fin (some species have predorsal scutes as well as ventral scutes), and a single scute is present at the insertion of the pelvic fins. The body is compressed in most species; fin spines are lacking, as are an adipose fin and gular plate; the jaws are not protractile; the pelvic fins are abdominal in position and free from the shoulder girdle; the pectoral fin is placed low on the side; and a mesocoracoid arch is present; the upper jaw is bordered by premaxillae and maxillae, but the detention is usually feeble; the cycloid scales are usually thin and loosely attached, especially in anchovies; the air bladder is connected to the gut (physostomous); and there is no leptocephalus larva.

Clupeiforms are mostly silvery and typically occur in large schools, feeding on plankton that is strained from the water by numerous long gill rakers. A notable exception is the wolf herring (*Chirocentrus*) of the Indo-Pacific, which is a predator with strong teeth.

Some of the great fisheries of the world are based on the conversion of plankton by clupeiform fishes. These include the tremendous anchovy fishery off western South America, the California fisheries for anchovy and (until depleted) Pacific sardine, the menhaden fisheries of the western Atlantic, and the herring and sardine fisheries of northern seas. The catch is processed variously into oil, fertilizer, or fish meal or is prepared directly for human consumption. Indirectly, clupeiforms are of tremendous importance in the food cycles of piscivorous fishes and of some sea birds which depend on them as dietary staples.

Clupeiform fishes have left a rich fossil record, especially from the Upper Cretaceous to the early Tertiary, but appeared first in the Upper Jurassic.

Denticipitidae. This family, in the suborder Denticipitoidei, comprises one fossil species, probably from the Miocene, and one extant species, *Denticeps clupeoides* (denticle herring), known from the coastal rivers of Nigeria and Cameroon. It differs from other clupeiforms in having a complete lateral line, denticles variously placed on the skull and branchiostegal rays, and 16 principal caudal fin rays.

Commerson's anchovy (*Stolephorus commersonnii*). Maximum size approximately 10 cm (4 in.). (*Photo © by John E. Randall*)

Clupeoidei. This suborder comprises the remainder of the clupeiforms and is distinguished by the lateral line not extending onto the body, but essentially limited to a cephalic canal system extending onto the operculum; and 19 principal caudal fin rays.

Engraulidae. Anchovies, which make up this family, are for the most part small, delicate fishes that have a conical snout overhanging a large mouth (see **illustration**). The maxilla is extremely long, reaching beyond the large eye and often to the margin of the operculum. Gill rakers are long and numerous; teeth are absent in jaws of most species—however, the large piscivorous species have large jaw teeth. Scales are thin and cycloid and usually deciduous, even when carefully handled. There are 16 genera and 139 species, with about 17 species occurring in freshwater and the rest in the Atlantic, Indian, and Pacific oceans. Thirty species occur in North American waters, 23 of which are in the genus *Anchoa*. Most of the freshwater species are from South America.

Pristigasteridae. In this family the mouth is superior in most species, otherwise it is terminal; jaw teeth are small (one genus has canine teeth); abdominal scutes are present; and the anal fin is long, with 30 to 92 rays. This family comprises nine genera and 34 species. Four species are freshwater and the rest occur in tropical and subtropical Atlantic, Indian, and Pacific oceans. Five species occur on the Pacific coast of Mexico.

Chirocentridae (wolf herring). This family consists of only one genus and two species that are characterized by an elongate, highly compressed body and fanglike jaw teeth. It has a spiral valve in the intestine, an uncommon feature for a teleost fish. The pelvic fin is small, the scales are small; the pelvic scute is reduced; and abdominal scutes are absent. These fishes occur in the western Pacific and the western Indian oceans to South Africa and the Red Sea.

Clupeidae. This family includes herrings, shads, sardines, sprats, pilchards, and menhadens. Two long, rodlike postcleithra distinguish most members of this family from other clupeiforms. In addition, the mouth is terminal and abdominal scutes are present, except in the subfamily Dussumieriinae (round herrings). The family is composed of five subfamilies, 50 genera, and 181 species, 50 of which occur in freshwater, some of which are anadromous. The rest are marine species that occur worldwide, mostly in tropical waters. Eleven genera and 33 species occur in North American waters.

Distribution and ecology. Some herrings and a few anchovies live in lowland rivers and lakes, and others such as the shad and alewife enter rivers to reproduce; but the majority of clupeiforms occur in bays or shore waters of tropical, temperate, or even northern seas, where they commonly make up enormous schools. None inhabits deep water. Most clupeiforms feed on plankton or other minute organisms, thus they are efficient converters from the base of the food chain to fish flesh.

Reeve M. Bailey; Herbert Boschung

Bibliography. L. Grande, Recent and fossil clupeomorph fishes with materials for revision of the

subgroups of clupeoids, *Bull. Amer. Mus. Nat. Hist.*, 181(2):231–372, 1985; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994; J. E. Nelson et al., *Common and Scientific Names of Fishes from the United States, Canada, and Mexico*, American Fisheries Society Spec. Publ. 29, Bethesda, 2004; P. J. P. Whitehead, *FAO Species Catalogue*, vol. 7: *Clupeoid Fishes of the World (Suborder Clupeoidei) Part 1: Chirocentridae, Clupeidae and Pristigasteridae*, FAO Fish. Synop. (125), 1985; P. J. P. Whitehead, G. J. Nelson, and T. Wongratana, *FAO Species Catalogue*, vol. 7: *Clupeoid Fishes of the World (Suborder Clupeoidei). Part 2: Engraulidae*, FAO Fish. Synop. (125), 1988.

Clutch

A coupling device which permits the engagement and disengagement of coupled shafts during rotation. There are four major types: positive, friction, hydraulic, and electromagnetic.

Positive clutch. This type of clutch is designed to transmit torque without slip. It is the simplest of all shaft connectors, sliding on a keyed shaft section or a splined portion and operating with a shift lever on a collar element. The jaw clutch is the most common type of positive clutch. This is made with square jaws (**Fig. 1***a*) for driving in both directions, or spiral jaws (Fig. 1*b*) for unidirectional drive. Engagement speed should be limited to 10 revolutions per minute (rpm) for a square-jaw clutch and 150 rpm for a spiral-jaw clutch. If desengagement under load is required, the jaws should be finish-machined and lubricated. *See* MACHINE KEY; SPLINES.

Friction clutch. This type of clutch is designed to reduce coupling shock by slipping during the engagement period. It also serves as a safety device by slipping when the torque exceeds its maximum rating. The three common designs for friction clutches are cone, disk, and rim, according to the direction of contact pressure. *See* FRICTION; TORQUE.

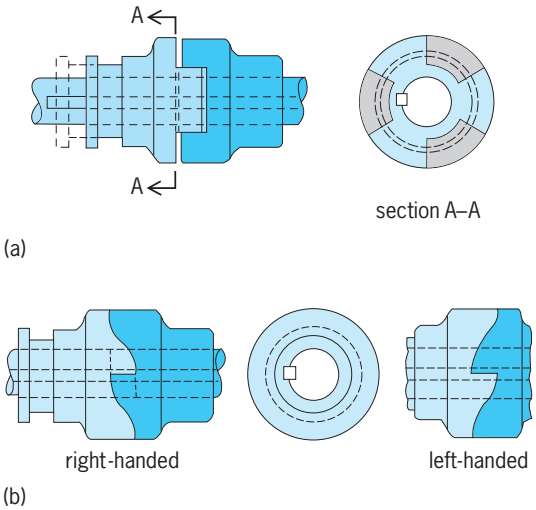

Fig. 1. Positive clutches. (*a*) Square-jaw clutch. (*b*) Spiral-jaw clutch. (*After E. A. Avallone and T. Baumeister III, eds., Marks' Standard Handbook for Mechanical Engineers, 9th ed., McGraw-Hill, 1987*)

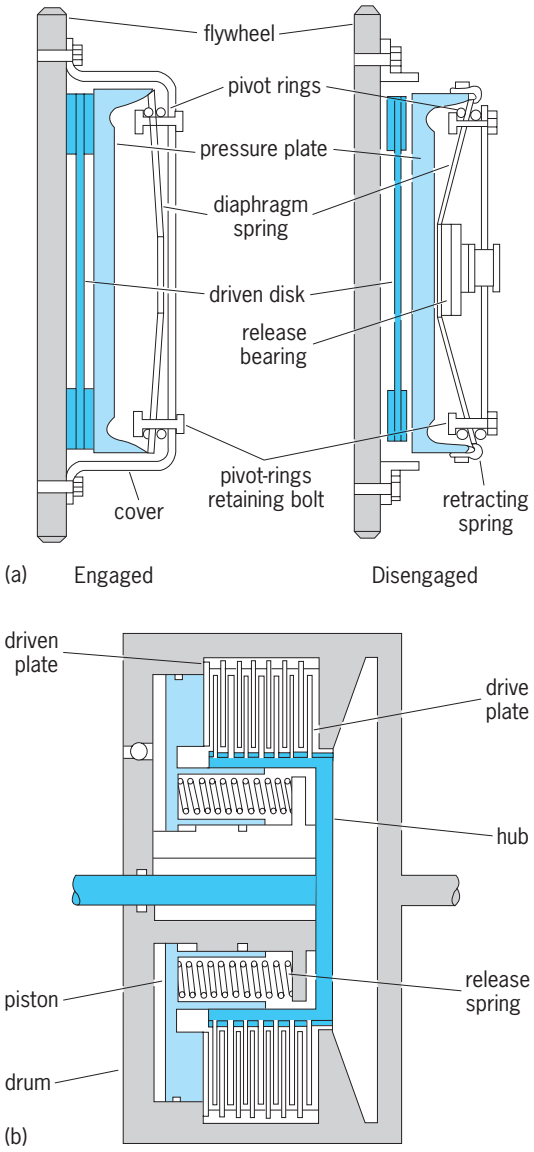

Fig. 2. Friction clutches. (*a*) Plate clutch. (*b*) Multidisk clutch. (*After E. A. Avallone and T. Baumeister III, eds., Marks' Standard Handbook for Mechanical Engineers, 9th ed., McGraw-Hill, 1987*)

Cone clutch. This is an axial type clutch. The surfaces of the cone clutch are sections of a pair of cones. This shape uses the wedging action of the mating surfaces under relatively small axial forces to transmit the friction torque. These forces may be established by the compression of axial springs or by the outward displacement of bell-crank levers to apply axial thrust to the conical surfaces.

Disk clutch. This is also an axial-type clutch. The disk clutch may consist of a single plate or multiple disks (**Fig. 2**). With the development of improved friction material, disk clutches have become more common than cone clutches, with wide applications in the industrial and automotive fields. Disk clutches are not subjected to centrifugal effects, present a large

Friction materials for clutches and design data*

Material	Friction coefficient		Maximum temperature		Maximum pressure	
	Wet	Dry	°F	°C	lb/in.^{2}	kPa
Cast iron on cast iron	0.05	0.15–0.20	600	320	150–250	1000–1750
Powdered metal† on cast iron	0.05–0.1	0.1–0.4	1000	540	150	1000
Powdered metal† on hard steel	0.05–0.1	0.1–0.3	1000	540	300	2100
Wood on steel or cast iron	0.16	0.2–0.35	300	150	60–90	400–620
Leather on steel or cast iron	0.12	0.3–0.5	200	100	10–40	70–280
Cork on steel or cast iron	0.15–0.25	0.3–0.5	200	100	8–14	50–100
Felt on steel or cast iron	0.18	0.22	280	140	5–10	35–70
Woven asbestos† on steel or cast iron	0.1–0.2	0.3–0.6	350–500	175–260	50–100	350–700
Molded asbestos† on steel or cast iron	0.08–0.12	0.2–0.5	500	260	50–150	350–1000
Impregnated asbestos† on steel or cast iron	0.12	0.32	500–750	260–400	150	1000
Carbon graphite on steel	0.05–0.1	0.25	700–1000	370–540	300	2100

*From J. E. Shigley and L. D. Mitchell, *Mechanical Engineering Design,* p. 742, McGraw-Hill, 1983.
†The friction coefficient can be maintained within ± 5% for specific materials in this group.

friction area in a small space, establish uniform pressure distribution for effective torque transmission, and are capable of effectively dissipating the generated heat to the external housing. The disk clutch may be operated dry, as in most automobile drives, or wet by flooding it with a liquid, as in heavier automotive power equipment and in industrial engines. The advantage of wet operation is the ability to remove heat by circulating the liquid enclosed in the clutch housing. Typical friction materials for clutches and design data are shown in the **table**.

Rim clutch. This is another form of frictional contact clutch that has surface elements applying pressure to the rim either externally or internally. Rim clutches may be subdivided into two groups: those employing either a frictional band or block (**Fig. 3**) to make contact with the rim; and the overrunning clutch (**Fig. 4**), employing the wedging action of a roller or ball. The clutches of the former type may have hinged shoes connected by an expanding or contracting kinematic mechanism on a hub fixed to one shaft and riding of a drum or rim attached to a hub on a second shaft. They have expansion rings or external bands that, when displaced, are capable of transmitting torque to the clutch rim. The rim may be grooved to increase the surface area; the clutch may have a double grip, internal and external; or pneumatically expanding flexible tubes may set the friction surfaces against the rim (**Fig. 5**). Because the clutch rotates with the shaft when it is engaged, its balancing must be considered.

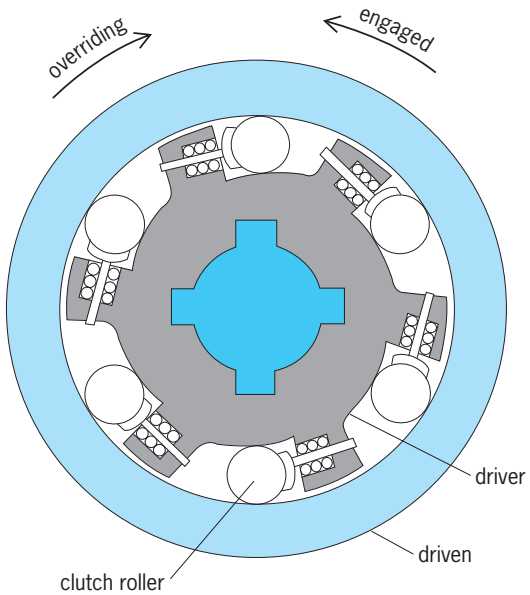

Fig. 4. Overrunning clutch. (*After E. A. Avallone and T. Baumeister III, eds., Marks' Standard Handbook for Mechanical Engineers, 9th ed., McGraw-Hill, 1987*)

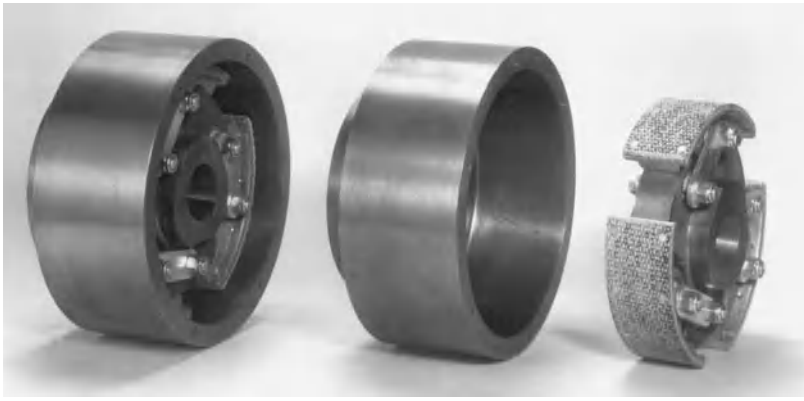

Fig. 3. Internal-expanding centrifugal-acting rim clutch. (*Hilliard Corp.*)

For overrunning clutches, the driven shaft can run faster than the driving shaft. This action permits freewheeling as the driving shaft slows down or another source of power is applied. Effectively this is a friction pawl-and-ratchet drive, wherein balls or rollers become wedged between the sleeve and recessed pockets machined in the hub. The clutch does not slip when the second shaft is driven, and is released automatically when the second shaft runs faster than the driver. Specially shaped struts or sprags may replace the balls; springs may be used to hold the pawl elements in position. Mechanisms have been devised to reverse the direction of operation. *See* PAWL; RATCHET.

Hydraulic clutch. Clutch action is also produced by hydraulic couplings, with a smoothness not possible

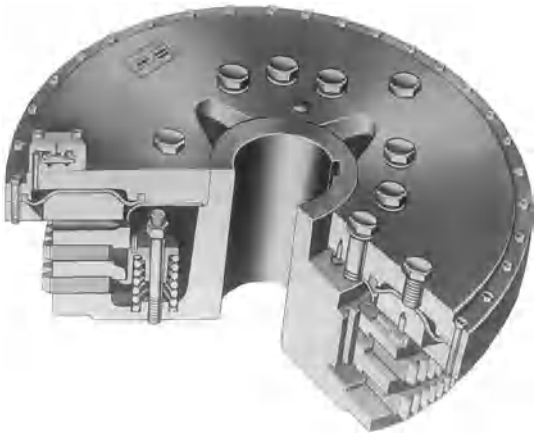

Fig. 5. External-contracting clutch that is engaged by expanding the flexible tube with compressed air. (*Twin Disc, Inc.*)

when using a mechanical clutch. Automatic transmissions in automobiles represent a fundamental use of hydraulic clutches. *See* AUTOMOTIVE TRANSMISSION; FLUID COUPLING; HYDRAULICS.

Electromagnetic clutch. Magnetic coupling between conductors provides a basis for several types of clutches. The magnetic attraction between a current-carrying coil and a ferromagnetic clutch plate serves to actuate a disk-type clutch. Slippage in such a clutch produces heat that must be dissipated, and wear that reduces the life of the clutch plate. Thus the electromagnetically controlled disk clutch is used to engage a load to its driving source. A typical unit 24 in. (61 cm) in diameter and weighing about 600 lb (270 kg) develops 2400 ft-lb (3300 joules) of torque when excited at 2 amperes and 115 dc volts; at a maximum safe speed of 1200 rpm it transmits 540 hp (400 kW). Multiple interleaved disks alternately splined to the driving and driven shafts provide a compact structure (**Fig. 6**). There are three basic types of electromagnetic clutches: magnetic fluid and powder, eddy-current, and hysteresis. *See* ELECTROMAGNET.

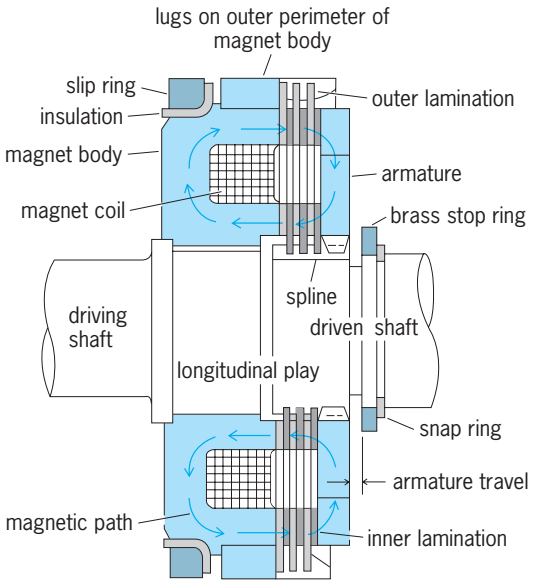

Fig. 6. Electromagnetically activated multiple-disk clutch.

Magnetic fluid and powder clutch. These may use finely divided magnetic particles in an oil carrier or be of the dry can type. Upon the application of a magnetic field, the particles tend to coalesce, thereby creating viscous or friction forces between the clutch members. The wet type provides the smoother operation; the dry type is freer from loss of particles. Torque at a fixed moderate excitation is independent of slip.

Eddy-current clutch. The current that is induced by a magnetic field can serve as the torque-transmitting means. In the eddy-current clutch, a coil-carrying direct current in one rotating member establishes poles that generate currents in the adjacent smooth conductive ring of the mating member. The coupling acts as an untuned damper and therefore does not transmit torsional vibration. An eddy-current clutch consists of a separately excited motor and generator that are combined into a single machine. Clutch action can be obtained by electrically connecting a generator on the driving shaft to a motor on the driver. *See* DYNAMOMETER; EDDY CURRENT; ELECTRIC ROTATING MACHINERY.

Hysteresis clutch. For low-power applications where continuous control with slip is needed, as in instrument servos, a hysteresis clutch may be used. Direct-current excitation of a coil in one part generates a steady flux field that magnetizes an iron ring on the other part, the ring having high magnetic retentivity. The action induces poles in the iron ring. Attraction between the induced poles in the ring and the control field produces a torque that opposes rotation between the two parts of the clutch. Within its load limit the hysteresis clutch transmits rotation without slip. As with the eddy-current clutch, there is no contact surface to wear so that the characteristics are stable and the life is long. *See* HYSTERESIS; SHAFTING; TORQUE CONVERTER. Y. S. Shin

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers,* 10th ed., 1996; J. E. Shigley and L. D. Mitchell, *Mechanical Engineering Design,* 4th ed., 1993.

Cnidaria

A diverse phylum of aquatic animals that are characterized by the possession of (1) intrinsic nematocysts, (2) a single body space that opens to the external medium, (3) radial or biradial symmetry, (4) tissue level of organization, and (5) diploblastic structure. With few exceptions, a cnidarian also has tentacles surrounding the mouth (the opening of the body cavity to the exterior). Most definitions of the phylum include the feature metagenesis; however, although the life cycle of many cnidarians involves this alternation between an asexually reproducing polyp and a sexually reproducing medusa, this is not true of two entire classes (Anthozoa and Staurozoa) and of some members of class Hydrozoa. The name Coelenterata,

which has been used for this phylum, is being supplanted because, as sometimes used, it included the sponges (Porifera) and comb jellies (Ctenophora), as well as the animals now called cnidarians. *See* CTENOPHORA; PORIFERA.

Taxonomy

Cnidarians include jellyfish, corals, sea anemones, the Portuguese man-of-war, hydroids, and the freshwater hydras. For many years, the phylum was divided into three classes, but four are now accepted and a fifth has been proposed: (1) Anthozoa (sea anemones, corals, sea fans, sea pens, and sea pansies), (2) Hydrozoa (hydroids, hydras, and hydromedusae), (3) Scyphozoa ("true" jellyfish), (4) Cubozoa (box jellies; formerly included in Scyphozoa), (5) Staurozoa (the Stauromedusae; formerly included in Scyphozoa) [see **table**]. All except a few hydrozoans are marine (**Fig. 1**). The Hydrozoa, Scyphozoa, Cubozoa, and Staurozoa are collectively referred to as

A classification of the phylum Cnidaria

Subdivision	Characteristics	Representative genera
Class Anthozoa	All marine	
Subclass Octocorallia	Almost exclusively colonial, each polyp with eight tentacles and mesenteries; also called Alcyonaria	
Order: Alcyonacea	Soft corals, sea fans, precious coral; sclerites in most species; thin or thick coenenchyme or stolonic network connects polyps; some with axial skeleton	*Alcyonium, Anthomastus, Clavularia, Corallium, Muricea, Tubipora*
Helioporacea	Blue coral; calcareous skeleton of aragonite; also called Coenothecalia	*Heliopora*
Pennatulacea	Sea pens and pansies; colony retractile, with bilateral symmetry	*Pennatula, Renilla, Virgularia*
Subclass Hexacorallia	Tentacles and mesenteries more or less hexamerously arranged; also called Zoantharia	
Order: Actiniaria	Sea anemones; solitary (may be clonal); live at all depths in all seas	*Actinia, Dactylanthus, Edwardsia, Metridium*
Antipatharia	Thorny or black corals; axial skeleton	*Antipathes, Dendrobranchia*
Ceriantharia	Tube anemones; solitary; lack pedal disc	*Cerianthus*
Corallimorpharia	Resemble corals; lack skeleton; tentacles radially arranged	*Corallimorphus, Corynactis*
Rugosa	Solitary or compound corals; calcareous skeleton; all fossil	*Lambeophyllum, Petraia, Syringaxon*
Scleractinia	True or stony corals; solitary and colonial; half of species reef formers	*Astrangia, Fungia, Porites*
Tabulata	"Honeycomb corals"; colonial, calcareous, tabulate corallites; all fossil	*Favosites*
Zoanthidea	Mat anemones; no skeleton; solitary or colonial, many commensal	*Palythoa, Parazoanthus, Zoanthus*
Class Cubozoa	Velarium present; largely tropical, subtropical	*Carybdea, Tripedalia*
Class Hydrozoa	Most marine; few species fresh-water	
Order: Actinulida	Hydroid generation represented by actinula-like form	*Halamohydra, Otohydra*
Anthoathecata	Athecate hydroids, hydrocorals, chondrophores	*Clava, Hydra, Millepora, Pennaria, Tubularia, Velella*
Limnomedusae	Some species commensal; hydroids naked	*Craspedacusta, Gonionemus*
Leptothecata	Hydroids with a hydrotheca	*Obelia, Plumularia, Sertularia*
Trachymedusae	Polypoid generation absent	*Aglantha, Liriope*
Narcomedusae	Holopelagic; most without polypoid generation; some parasitic	*Aegina, Cunina, Tetraplatia*
Siphonophora	Polymorphic colonies; planktonic	*Abyla, Forskalia, Physalia*
Class Scyphozoa	All marine	
Subclass Scyphomedusae	Tetramerous medusae that are produced through strobilation and grow through ephyra stage	
Order: Coronatae	Coronal furrow; lack ocelli	*Atolla, Linuche, Nausitho, Periphylla*
Semaeostomeae	Pedalia and coronal furrow lacking	*Aurelia, Cyanea, Pelagia*
Rhizostomeae	Numerous small mouths; no marginal tentacles	*Cassiopeia, Cephea, Rhizostoma*
Subclass Conulata	Tetramerous cone-shaped to elongate pyramidal; unattached; tentacles on oral margin; all fossil	
Order Conularida	Characters of subclass	*Conularia, Conulariella, Paraconularia*
Class Staurozoa	Sessile medusae; largely temperate	*Haliclystus, Lucernaria*

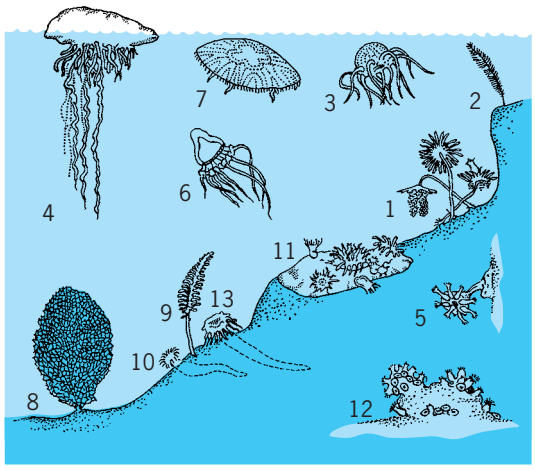

Fig. 1. Marine cnidarians in typical habitats (not to scale): 1, *Tubularia* (solitary hydroid); 2, *Plumularia* (colonial hydroid); 3, *Gonionemus* (hydromedusa); 4, *Physalia* (Portuguese man-of-war); 5, *Haliclystus* (stauromedusa); 6, *Periphylla* (scyphomedusa); 7, *Aurelia* (scyphomedusa); 8, *Gorgonia* (sea fan); 9, *Pennatula* (sea pen); 10, *Edwardsia* (sea anemone); 11, *Epiactis* (sea anemone); 12, *Astrangia* (stony coral); and 13, *Cerianthus* (tube anemone). 1–4 are members of Hydrozoa, 5 is a member of Staurozoa, 6 and 7 are members of Scyphozoa, and 8–13 are members of Anthozoa. (*After T. I. Storer and R. L. Usinger, General Zoology, 3d ed., McGraw-Hill, 1957*)

the Medusozoa because at least some species in each class have a medusa in the life cycle, whereas no anthozoans do. *See* ANTHOZOA; CUBOZOA; HYDROZOA; SCYPHOZOA; STAUROMEDUSAE.

General Morphology

The two basic body forms of cnidarians—the polyp and the medusa—typically live in different habitats (the former benthic, the latter pelagic), and although they appear superficially rather different, they share many features (**Fig. 2**). Of the two epithelial tissue layers, the endoderm (also termed the gastrodermis) lines the single body space, the coelenteron (also termed the gastrovascular cavity), and the ectoderm (also termed the epidermis) forms the external surface. [The name Coelenterata is derived from the term coelenteron (meaning "hollow gut")] Between ectoderm and endoderm is the mesoglea, which varies from a thin, acellular cementing layer

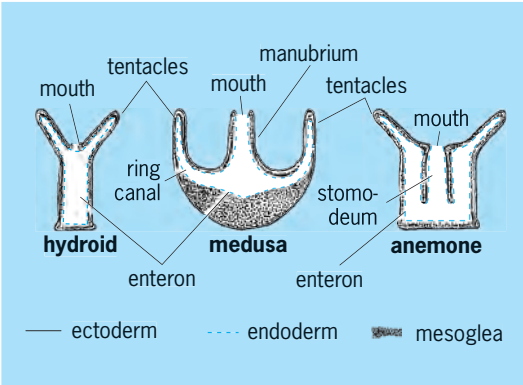

Fig. 2. Comparison of hydrozoan polyp, medusa (inverted), and anthozoan polyp. (*After T. I. Storer and R. L. Usinger, General Zoology, 3d ed., McGraw-Hill, 1957*)

to the thick, cellular and fibrous gelatinous layer that is the "jelly" of jellyfishes.

The ectoderm is, in general, a single layer of cells, varying in form from flat and squamous to tall and columnar. It may be ciliated; in its food-gathering mode, the ciliary beat can move material toward the mouth from the surface of the animal and its tentacles, but in its cleaning mode, the direction is reversed. The endoderm may also be ciliated: cilia circulate water through the coelenteron and move food around within it.

Being at the "tissue grade" of construction, cnidarians lack organs and, therefore, organ systems. Their diploblastic structure means that they essentially consist of folded sheets of epithelia—and so have been termed "origami animals." Types of cells in the epithelia (both ectoderm and endoderm) include gland cells (that produce mucus and digestive enzymes), sensory cells, undifferentiated interstitial cells, ciliated cells, and cnidocytes (cells that produce cnidae, including nematocysts). Epitheliomuscular cells, in which the elongated cell base contains contractile fibrils, may constitute the entire musculature of simple polyps, such as most hydroids, but sea anemones and medusae have distinct muscle cells (contrary to many textbooks).

Polyp. A polyp is typically benthic (lives associated with the bottom of an ocean, lake, or stream), sessile, cylindrical, and elongated along the oral-aboral axis. The mouth, at the polyp's free (distal) end, is almost invariably surrounded by few to many hollow or solid tentacles (depending on the taxonomic group) arrayed in whorls or radially (or both). The aboral (proximal) end of a polyp (commonly referred to as the base or pedal disk may be adapted for attachment to a solid substratum or for burrowing into a soft substratum; but in colonial species, it arises from another polyp or the common tissue mass. With strong bilaterality in some taxa, especially Class Anthozoa, the polyp's symmetry is typically tetraradial, hexameral, octameral, or decameral, but it may be otherwise. An anthozoan polyp has sheets of tissue extending radially from the column wall into the coelenteron; some or all of these mesenteries connect with the actinopharynx (or stomodeum), a tube of tissue extending partway into the coelenteron from the mouth (Fig. 2, right). The mesoglea of a polyp may be little more than an adhesive, holding the two cell layers together (as in many hydroids), or it may be thick, spongy, and with many cells (as in some sea anemones). Polyps range in length from 1 mm to 2 m.

Medusa. A medusa is typically pelagic (lives free in the water), with thick mesoglea; its "body" is termed a bell due to its form. The outer (aboral) surface of the bell is termed the exumbrella and the underside (oral part), the subumbrella. The mouth commonly opens at the end of a tube of tissue (the manubrium) that extends from the underside of the bell (Fig. 2, center); there may be some tentaclelike structures immediately around the mouth, but typically the tentacles are located at the margin of the bell. Conventionally portrayed with the tentacles

hanging downward and the exumbrella pointing toward the surface of the sea, most medusae spend much of their time in postures other than this. The mouth, commonly via the manubrium, leads to the central coelenteron. Radial canals may run between the cavity and the margin on the subumbrellar side, and they commonly communicate with a ring canal that is continuous around the margin; canals may also run from the ring canal along the tentacles. The mesoglea serves as a type of internal skeleton, against which muscles act, and its elasticity returns the bell to its original shape after each contractile pulse.

Cnidae

The name of phylum Cnidaria is derived from "cnidae," microscopic structures that are the phylum's most distinctive feature, and the most complex secretory product made by any animal. Each is formed within a cell (a cnidocyte) and consists of a capsule and a tubule that is coiled or folded within the capsule. When discharge occurs, the tubule emerges from one end of the capsule; typically, the tubule is many times longer than the capsule. There are three types of cnidae: nematocysts are formed by all cnidarians (and only cnidarians); spirocysts are restricted to the Subclass Hexacorallia of Class Anthozoa; and ptychocysts are restricted to Order Ceriantharia of Subclass Hexacorallia.

Nematocysts. When a nematocyst (a "stinging capsule," not a "stinging cell" as it is sometimes called) is appropriately stimulated, its tubule rapidly everts (turns inside out, like a coat sleeve). Some types are used offensively (they are the means by which the exclusively carnivorous cnidarians capture their food) and some types are used defensively. Some types of hydrozoan nematocysts have a cnidocil (a trigger) that must be stimulated for discharge to occur, but in other cnidarians, it appears that the sensation triggering discharge is received at some distance from the nematocyst, and nerves transfer the impulse that stimulates discharge. Thus, a cnidarian has some control over the use of its nematocysts, although previously nematocysts were thought to discharge upon appropriate stimulation, without any control by the animal. Such control is reasonable; each nematocyst can be used only once, so producing them represents a large portion of the energy budget of a cnidarian.

Approximately 30 morphological types of nematocysts have been identified, differing in capsule shape, tubule length, diameter, and spination. However, all seem to fall into one of three functional categories (**Fig. 3**): a volvent has an unarmed tubule that wraps around and entangles prey; a penetrant has a tubule armed with barbs that penetrates prey and may deliver a toxin; and a glutinant has a sticky tubule that hangs on to prey so that it can be carried to the mouth, or helps to anchor the cnidarian to an object

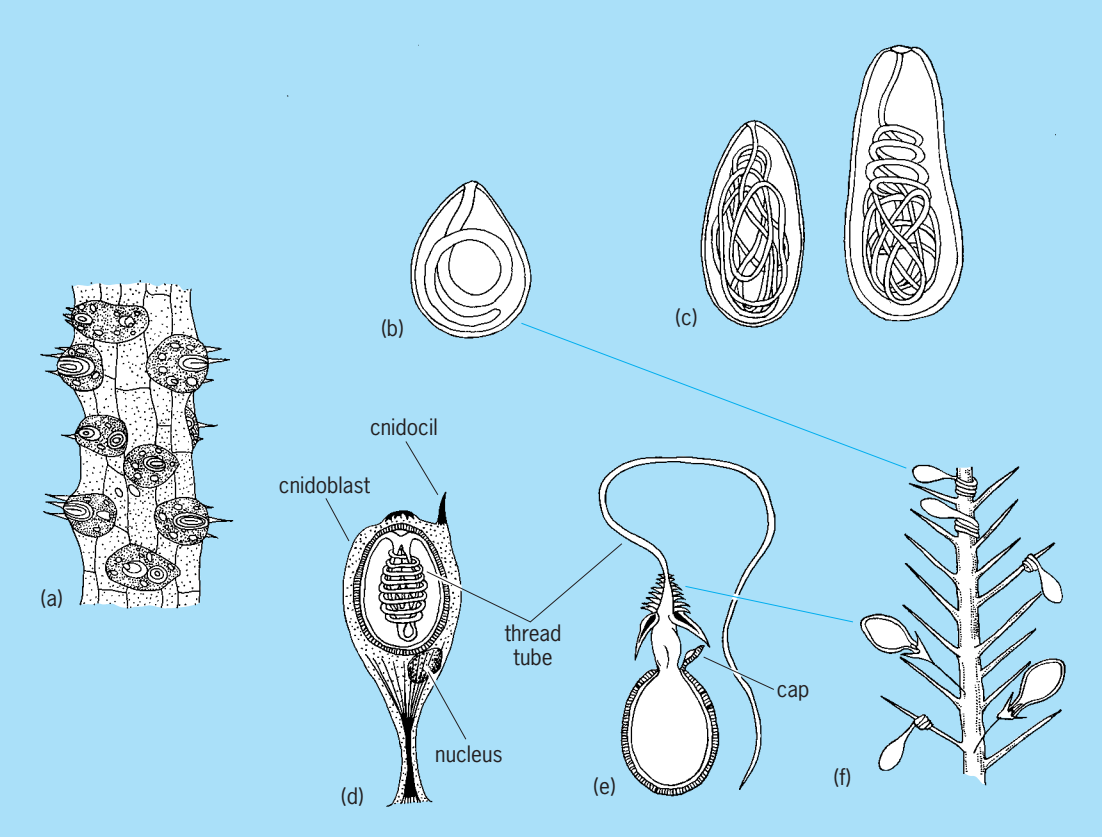

Fig. 3. Nematocysts of Hydra. (*a*) Tentacle with nematocyst batteries. (*b*) Volvent. (*c*) Glutinant. (*d*) Penetrant before discharge. (*e*) Penetrant discharged. (*f*) Discharged nematocysts on bristle of a crustacean. (*After T. I. Storer and R. L. Usinger, General Zoology, 3d ed., McGraw-Hill, 1957*)

(for example, glutinants appear to be important in attaching its tentacles to the substratum when *Hydra* somersaults).

The double-walled nematocyst capsule is made of collagenous proteins. Nematocysts are typically densest in the tentacles, where types important in prey capture are particularly abundant. Nematocysts are limited to the ectoderm of hydrozoans, whereas in Scyphozoa and Anthozoa, they occur in both ectoderm and endoderm. In sea anemones, certain types are restricted to the filaments on the mesenteries, and they seem to contain digestive enzymes. Some types of nematocysts can be painful to humans, and cubozoan stings may cause vasomotor dysfunction and even death. *See* TOXIN.

Although formed only by cnidarians, nematocysts may occur in other kinds of animals. Some flatworms, comb jellies, and nudibranchs that feed on cnidarians can store the capsules in their own bodies without causing them to discharge. Oriented so the eversible end of the capsule is at the surface of the animal's body, these "kleptocnidae" can be used to defend the cnidarian predator.

Spirocysts and ptychocysts. Spirocysts are especially common in tentacles, where they seem to function in entangling prey. The single-walled spirocyst capsule is thin, and the tubule is long, unarmed, spirally coiled, and of uniform diameter. Spirocysts stain with acid dyes, whereas nematocysts stain with bases. Ptychocysts are larger than either nematocysts or spirocysts (although they, too, are intracellular and therefore microscopic), and the tubule of each is folded within the capsule. Ptychocysts occur largely on the exterior surface of the body of a cerianthid, where their discharged tubules entangle, forming a feltwork that incorporates mud and organic matter to create the tube that is typical of these "tube anemones."

Functional Morphology

Cnidarians lack organs and therefore organ systems. Because the endoderm and ectoderm are each a single cell thick, all their cells are in contact with either the external medium or the fluid in the coelenteron, allowing gas exchange and excretion to be carried out by individual cells. Other physiological functions are described below.

Nerves and sensory structures. The nerves form a meshwork, with no centralized ganglia. However, the individual nerve cells and the synapses are polarized, like those of other metazoans (previously the system was considered unpolarized). Nerves are situated at the base of the ectoderm and endoderm, with a variable number of connections across the mesoglea, in some animals particularly around the mouth and in the tissue connecting members of a colony. Some medusae have sensory structures such as eyespots or statocysts (which function in balance) along the ring canal, on the tentacle bases, or on marginal structures. Cubozoan eyes are the most highly developed of these sensory structures: they can form images. In other cnidarians that lack discrete sensory structures, there is evidence that at least some nerves are sensitive to particular stimuli. For example, scleractinian corals typically expand at night and contract by day; thus they appear to be able to detect light despite lacking any structure that is recognizably light sensitive.

Functionally, in at least some species, there appear to be nervous pathways specialized for rapid conduction and others for slower, more general spread of impulses. Colonies of some species are functionally integrated by means of nerves, whereas in other species, the members of the colony function more independently. In the former case, stimulation of a single polyp may provoke contraction of many other members of the colony, whereas in the latter case, it will cause contraction only in the area of stimulation. Potentiation is demonstrable in some cnidarians, with a single stimulus causing only a local contraction, and succeeding stimuli at the same location causing a more widespread response. Many cnidarians show simple reflexes, such as righting movements and swallowing.

Swimming contractions of medusae are coordinated by nerves. In species that have them, marginal sensory structures generally regulate these actions. When the structures are removed, swimming movements stop unless the medusa is artificially stimulated; when some of the structures are removed, movement of the medusa is uncoordinated and erratic. In the laboratory, if a preparation constituting a ring from the margin of a medusa is stimulated, a wave of contraction is initiated. *See* NERVOUS SYSTEM (INVERTEBRATE).

Muscles. The column and tentacles of a polyp have an ectodermal longitudinal muscle layer, and an endodermal, circular muscle layer; the contractile elements are located subepithelially. Most small organisms have epitheliomuscular cells (**Fig. 4***c*), but

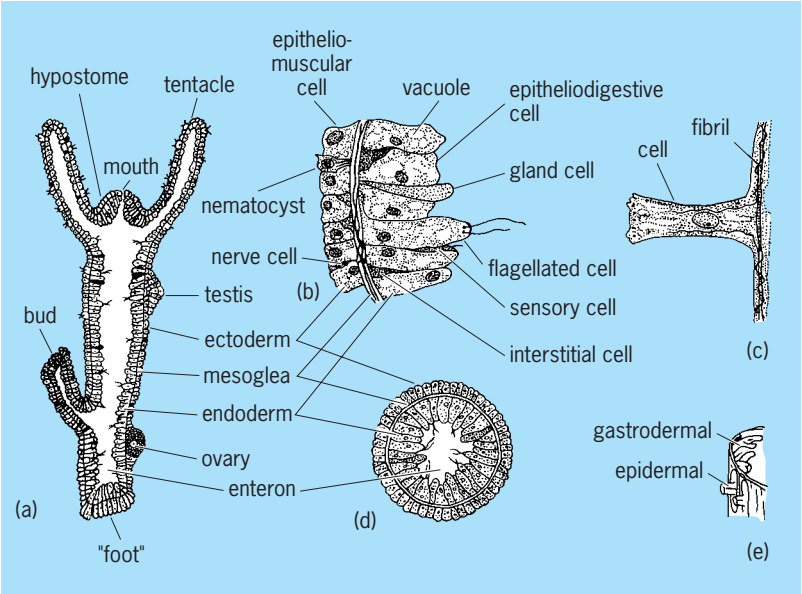

Fig. 4. Histological sections of Hydra. (*a*) Longitudinal section. (*b*) Enlarged cross section through column. (*c*) Epitheliomuscular cell. (*d*) Cross section of tentacle. (*e*) Fibril arrangement of the epitheliomuscular cells. (*After T. I. Storer and R. L. Usinger, General Zoology, 3d ed., McGraw-Hill, 1957*)

many larger ones have discrete muscle cells, which in some medusae are striated. The primary swimming muscles of medusae are large subumbrellar bands, oriented both circularly and radially. In many anthozoans, hypertrophy of the circular muscle at the margin of the oral disk forms a sphincter muscle that allows the animal to close up like a drawstring purse; in some taxa, this muscle is embedded in the mesoglea. In addition to column and tentacle musculature, anthozoan polyps have longitudinal and radial muscles in the mesenteries; the former are the major means by which anthozoan polyps shorten. The muscles work against the hydrostatic skeleton or the mesoglea, depending on the taxon and the part of the body.

Skeleton. The body form of most polyps and some medusae is maintained at least in part by a hydrostatic skeleton, which constitutes water in the coelenteron (and tentacles in the groups with hollow tentacles) maintained under a slight positive pressure through the action of muscles surrounding the cavity and closure of the mouth. By contracting its circular muscles, the polyp or medusa can narrow and extend in length, and by contracting its longitudinal muscles, it can broaden and shorten. When contraction is excessive, some water is expelled through the mouth and, in polyps and some species, through pores at the tentacle tips or in the body wall; expansion can occur again only by pumping water in to reinflate the animal. The mesoglea imparts rigidity and elasticity to the system. Hydrostatic pressure can also be important in supporting colonies (such as those of sea pens), although mechanical support by skeletons made of calcium carbonate and organic materials such as chitin is widespread in the phylum.

The ectoderm of most hydrozoan polyps, some scyphozoan polyps, and all scleractinian corals secretes an exoskeleton, which supports and protects the organism. The hydrozoan skeleton is tubular, surrounding the living tissue. Although typically made of chitin or sclerified protein, and therefore rather flexible, in some taxa it is calcified and rigid. Some hydrozoan skeletons are elaborate, with an ornamented cup into which each polyp may withdraw; others are simple, covering only the column of the polyp. In some colonial pelagic hydrozoans, such as the by-the-wind sailor *Velella*, part of the skeleton may be specialized for flotation. All scleractinians form calcareous exoskeletons, which vary from delicate, branched structures to massive encrusting formations. The living animal tissue overlies the skeletal material, with the polyps occurring in cups or depressions in the skeleton; in colonies, the polyps abut one another or are connected by sheets of living tissue.

Members of several anthozoan groups have an axial skeleton, which is also secreted by ectodermal cells and hidden in life by the polyps and common colony tissue, so it can be seen only after the death of the animals that made it. Composed of sclerified proteins, this tough skeleton is flexible; it is calcified to a greater or lesser extent in some taxa, and flexibility diminishes with calcification. An axial skeleton is characteristic of the hexacoral antipatharians (black corals) and octocoral gorgonians (sea fans and sea whips) and pennatulaceans (sea pens).

Cells of some hydrozoans and anthozoans secrete calcareous sclerites (also known as spicules). Varying in size from microscopic to several millimeters in length, many sclerites are needlelike in form, but others are more globular. In most octocorals, they function to support and strengthen the living tissue in which they are scattered. Some sclerites in some species surround individual polyps or groups of polyps, clearly functioning to protect them. In two groups of octocorals, the sclerites are fused to form solid skeletons. Their morphology is akin to that of hydrozoan corals, with the polyps living in tubes surrounded by protective and supportive skeleton, but, as in scleractinians, the skeleton is covered by tissue in life.

Locomotion. Solitary polyps of many species can creep slowly along the substratum; some can lose contact and be carried by currents. Whether this has functional significance is unclear. A *Hydra* polyp can move in a somersault fashion by alternately attaching the base and tentacles to the substrate. The polyps of a few species of anemones can "swim": some do so by paddling their tentacles—whereas others do so by flexing the column. A few others routinely roll along the sea floor; some of them have suctorial structures on the column that allow them to remain stationary as well. Polyps of *Hydra* and small sea anemones may hang from the surface film of water bodies. One species of sea anemone secretes from its base a chitinous mass that traps air, forming a float from which the anemone hangs as it drifts across the open ocean.

Colonies of polyps can also move about. A sea pansy peduncle anchors the colony in soft substrata and can be released so that the colony can be passively moved by water motion; the peduncle may also function to pull the colony along. The most common examples of motile colonies are the siphonophores and chondrophores, pelagic hydrozoan "superorganisms" (consisting of polyps and/or medusae) such as the Portuguese man-of-war (*Physalia*) and the by-the-wind sailors (*Velella*) that are mistakenly thought by some people to be jellyfish.

A medusa swims by contracting muscles that encircle the edge of the bell, thereby reducing the volume of the subumbrellar space. The water that is forced outward jets the medusa in an aboral direction. In medusae of Class Hydrozoa, a thin sheet of tissue (the velum) extends inward from the margin of the bell, reducing the effective diameter of the subumbrellar aperture, and helping the animal to steer. Cubomedusae have an analogous structure, known as the velarium. Once the circular muscles relax, the resilient and elastic mesoglea returns the umbrella to its original form.

Feeding, digestion, and nutrition. All cnidarians are exclusively carnivorous. (Although there have been claims of phytophagous behavior in some cnidarians

such as soft corals, there is no evidence that these animals possess enzymes capable of digesting plant material.) Larger prey items are secured by the tentacles: the tubules of some nematocysts hold onto the prey while the tubules of other types of cnidae deliver toxins to kill it. The tentacles then pass the prey to the mouth, where it is swallowed into the coelenteron. Tiny prey items may be trapped in mucus on the tentacles and oral disk, and carried by ciliary action to the mouth and into the coelenteron.

Gland cells that produce digestive enzymes are abundant in the endoderm. It is likely that in cnidarians that have endodermal nematocysts, the nematocysts also deliver digestive enzymes to prey items. Although some enzymes may be released into the coelenteron, most or all digestion takes place when the prey comes into contact with the endoderm so that the enzymes can be brought directly to the prey. In scyphozoans, gland cells tend to be most abundant on the gastric tentacles, and in anthozoans they are most abundant in the filaments along the free edges of the mesenteries: these structures wrap around the prey, making close contact between it and the gland-rich endoderm. Once the prey is broken into small particles, the particles are engulfed by endodermal cells and digestion is completed in food vacuoles. Thus, digestion in cnidarians is intracellular, although extracellular proteases effect the first stages of food processing.

Distribution of the compounds, in the absence of a circulatory system, depends at least in part on diffusion from the endodermal cells in which digestion occurs. Wandering ameboid cells may also play a role in distributing these materials, particularly to ectodermal cells in animals with a thick mesoglea, because they are more efficient than diffusion at crossing the mesoglea.

Some cnidarians appear to feed seldom or never. Two alternatives to feeding are uptake of dissolved organic matter and acquisition of small organic molecules from photosymbionts. At least some cnidarians can acquire dissolved organic matter directly from seawater, but not all their ectodermal cells have this capability, so this seems to be a specialized type of nutrition. Many cnidarians, particularly polyps, that live in the photic zone contain intracellular photosymbionts, but this is far from universal. In the marine environment, most photosymbionts are dinoflagellates known as zooxanthellae; the photosymbionts of all freshwater and some marine species are green algae known as zoochlorellae. Photosymbionts are generally contained in endodermal cells, but in a small number of species they are ectodermal; a single animal cell may contain several photosynthetic cells. A large proportion of the carbon fixed by the zooxanthellae (and a smaller proportion of that fixed by zoochlorellae) is released to the animal host; reciprocally, the photosymbiont obtains from its animal host carbon dioxide and nitrogenous wastes. Nearly all reef-forming corals contain zooxanthellae; some of these corals can acquire 100% of their required energy from zooxanthellae, but it appears these animals must ingest animal prey to obtain other compounds such as amino acids and vitamins. *See* FEEDING MECHANISMS (INVERTEBRATE).

Reproduction and Development

Cnidarians of all species reproduce sexually, although not all individuals may be capable of sexual reproduction. Cnidarians of many, but not all, species can reproduce asexually through a variety of mechanisms. The capacity to regenerate is widespread, but seems especially characteristic of animals that can propagate asexually. Thus, pieces cut from most medusae or many sea anemones will not survive, and the part that is excised may not be replaced. *See* REGENERATION (BIOLOGY).

Metagenesis. Metagenesis, the alteration between medusa and polyp generations, is commonly considered a characteristic feature of the phylum as a whole (**Fig. 5**). In metagenesis, a medusa is produced asexually by a polyp (which may also asexually produce polyps), and a medusa produces gametes that, through the processes of fertilization and development, result in polyps. Clearly, anthozoans, which are exclusively polypoid, do not undergo metagenesis; this is also true of the groups of hydrozoans that exist only as polyps or medusae. Staurozoa, too, has only one generation (a sedentary medusa or medusa-like polyp).

Many hydrozoans produce free-swimming medusae that have relatively long lives and are carried some distance from their hydroid before forming and releasing gametes. However, in some taxa the medusa is reduced so that it essentially constitutes a package of gametes that lives only long enough to spawn shortly after release from the hydroid; in others, it is not released at all, spawning while still attached to the hydroid. The ultimate reduction involves the disappearance of a morphologically distinct medusa, with gametes formed in structures on the polyps that are remnants of the medusa and essentially constitute a gonad of the hydroid.

Sexual reproduction. Since cnidarians lack organs, they do not have true gonads; rather, gametes are formed in localized areas, which are often referred

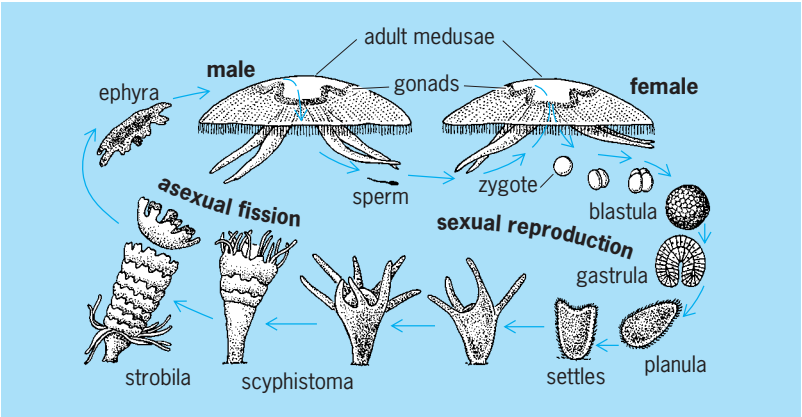

Fig. 5. Life cycle of jellyfish *Aurelia aurita*, class Scyphozoa. (*After T. I. Storer and R. L. Usinger, General Zoology, 3d ed., McGraw-Hill, 1957*)

to as gonads. Thus, there are no accessory sexual structures or ducts to permit the passage of gametes. Gametogenic tissues typically produce gametes only during certain times of the year; at other times, there is no evidence of their existence. Cnidarians lack a germinal cell line, so it is unclear if gametes develop from undifferentiated cells (such as interstitial cells) or if they can develop from cells of any type. In hydrozoans, gametes are produced in the ectoderm, on the manubrium or beneath the radial canals of medusae, and almost anywhere on a hydroid polyp (although in any group or species, gametogenic tissue forms in a characteristic position). Gametes in groups other than Hydrozoa are formed from endodermal cells: in scyphomedusae they develop on the floor of the coelenteron, and in anthozoans they develop in the mesenteries between the filaments and the retractor muscles.

Most cubozoans, hydrozoans, scyphozoans, and staurozoans are gonochoric (have separate sexes); hermaphroditism is widespread in Anthozoa, although most anthozoans are gonochoric. Hermaphroditic species may be self-fertile, but most cross-fertilize. Typically, a cnidarian sheds its gametes into the water surrounding the animal, where fertilization occurs and the embryo develops. However, in some cnidarians, the eggs are fertilized within the body; the developing embryo may be retained in the body or on its surface for its entire duration or for only a part of it.

The first stage of embryogenesis results in the formation of a ciliated blastula (Fig. 5). Gastrulation produces a gastrula that elongates into the typical larva of the phylum, the planula, which is free-swimming, ciliated, mouthless, and pear-shaped (the anterior is the broad end) and has an elongate anterior tuft of sensory cilia. At metamorphosis, if the planula is destined to become a polyp, it attaches by its anterior end, developing a mouth and tentacles at its posterior end. If the the planula is to become a medusa, the anterior end becomes the exumbrellar surface of the bell, and the mouth develops from the posterior end. In some hydrozoans, the planula develops a mouth and tentacles while still swimming; the animal at this stage is termed an actinula. The actinula of some species may attach and develop into a polyp, whereas in others it metamorphoses into a medusa. *See* INVERTEBRATE EMBRYOLOGY; REPRODUCTION (ANIMAL).

Asexual reproduction. If the individuals produced by asexual means separate, they are collectively termed a clone. If the individuals remain attached to one other, they collectively constitute a colony (although it is true that the elements, being genetically identical, are clonal).

One mode of asexual reproduction is budding (a term that is incorrectly applied to any mode of vegetative propagation). Budding involves the creation of a new individual from an outgrowth of an existing organism or the living tissue between individuals (as in Fig. 4*a*). This is the typical mode of formation of hydroid and many scleractinian colonies. Budding is also the means by which hydromedusae are produced by a hydroid colony. Budding is very rare in sea anemones.

Some scleractinian colonies are created by longitudinal fission of existing polyps. In the few medusae that normally propagate asexually, most (or all) do so by binary fission. Some individual sea anemones propagate asexually through longitudinal fission, and others through transverse fission. Transverse fission is the means by which a scyphozoan polyp (at the stage known as a scyphistoma) forms medusae: the distal end develops into an incipient medusa (the ephyra). In some species, these form one at a time, whereas in others, several form simultaneously like a stack of dinner plates of diminishing maturity proximally. When mature, the ephyra pops off the top of the polyp and swims away (Fig. 5).

Small pieces of hydroids may detach and wander about as undifferentiated, elongated entities called frustules, each of which can attach and develop into a new polyp and, ultimately, a colony. Basal bits of sea anemones can also detach and move off and are each capable of developing into a whole anemone. A few species of anemones can even regenerate a whole individual from a cast-off tentacle.

Asexual progeny are genetically identical, at least initially, but may differ morphologically. The most dramatic example of this is a polyp giving rise to a medusa. Even among polyps, in a hydrozoan or octocoral colony, individuals can be very different morphologically. Polymorphic colonies may contain individuals specialized to feed; others specialized to reproduce; and others to protect, propel, or support the colony. In hydrozoan Class Siphonophora, which constitutes the acme of polymorphism, the colony is also highly integrated neurologically and behaviorally. Polymorphism does not occur in scleractinian coral colonies.

Somatic mutation will cause members of a clone or colony to diverge genetically. In colonies that live decades or centuries (like those of some scleractinians), considerable genetic diversity can accumulate.

Ecology

Hydra, the animal used as an exemplar of the phylum, could hardly be less typical—a solitary hydrozoan that lacks a medusa, makes no skeleton, and lives in freshwater. Only a few hydrozoans besides it occur in freshwater. The other 10,000 or so species of cnidarians occur in all oceans and from the highest intertidal to the greatest hadal depths. Medusae are important predators in the pelagic realm, as sea anemones can be in some benthic environments; scleractinians are the primary constructors of one of the most important ecosystems in shallow tropical waters—coral reefs. Calcifying cnidarians do not occur below about 5 km (3 mi) depth, and those at high latitudes do not form large colonies. Both limitations seem due to the saturation state of calcium carbonate in the sea, which is greatest in shallow tropical waters and declines with latitude and depth. Moreover, for reasons that are not entirely understood, the robust skeletons that are important in reef formation are deposited only by corals that

possess zooxanthellae (mostly scleractinians, but also a few species of hydrozoans and octocorals).

Only about half the approximately 1500 species of scleractinians participate in forming coral reefs; the others live at high latitudes and below the photic zone. The so-called deep-sea coral reefs are species-rich areas physically structured by skeletons of scleractinian corals, like their shallow counterparts. The diversity of corals, however, is low, and their skeletons are relatively delicate—they would be easily broken in many of the habitats where tropical shallow reefs thrive. *See* ECOLOGY; OCTOCORALLIA (ALCYONARIA); REEF; SCLERACTINIA. Daphne G. Fautin

Fossils and Phylogeny

Cnidaria has a long and impressive fossil record stretching back to the Precambrian, about 700 million years ago. Thus, the known duration of this phylum equals or exceeds that of other animal phyla. Forms with skeletons (primarily the Conulata, Tabulata, Rugosa, and Scleractinia) have left fairly complete fossil records, whereas primarily soft-bodied forms, such as most hydrozoans, scyphozoans, alcyonarians, and anemone-like animals, are represented by few fossils.

Fossils of the earliest-known cnidarians have been found in Late Precambrian rocks on several continents; the Ediacaran is an especially well-preserved fauna known from southern Australia. These early soft-bodied cnidarians were both pelagic medusoid forms and benthic polypoid forms, and seem to represent all the major classes, indicating that cnidarian origins are truly ancient. Appearing in the succeeding Cambrian strata is the extinct group Conulata (**Fig. 6***a*), which possibly belongs in Scyphozoa. These forms are represented in the fossil record by a thin, steep-sided, pyramidal, chitino-phosphatic external skeleton.

Fossils of two of the three major orders of anthozoans with calcium carbonate skeletons, the Rugosa (tetracorals) and the Tabulata (Figs. 6*b*, *c*, and *d*), appear in rocks of the succeeding Ordovician system. Both groups quickly diversified, and include colonial forms that helped form carbonate banks and reefs. In addition, the rugosans are represented by many solitary forms known as horn corals. Rugose corals, especially, were numerous and widespread in the Early Permian, but in the later part of that period their diversity dwindled, the more complex, bank-forming groups disappearing first. Both tabulate and rugose corals disappeared during the great wave of Late Permian extinctions, along with many other marine invertebrates. *See* PERMIAN; RUGOSA; TABULATA.

There are no corals in the Early Triassic, although the last Conulata have been reported from beds of this age. Scleractinians first appear in Middle Triassic strata (Figs. 6*e*, *f*). Scleractinians cannot have evolved from rugosans because they differ in fundamental morphology and there are no transitional forms in the Early Triassic. Thus, it is inferred that all the coral

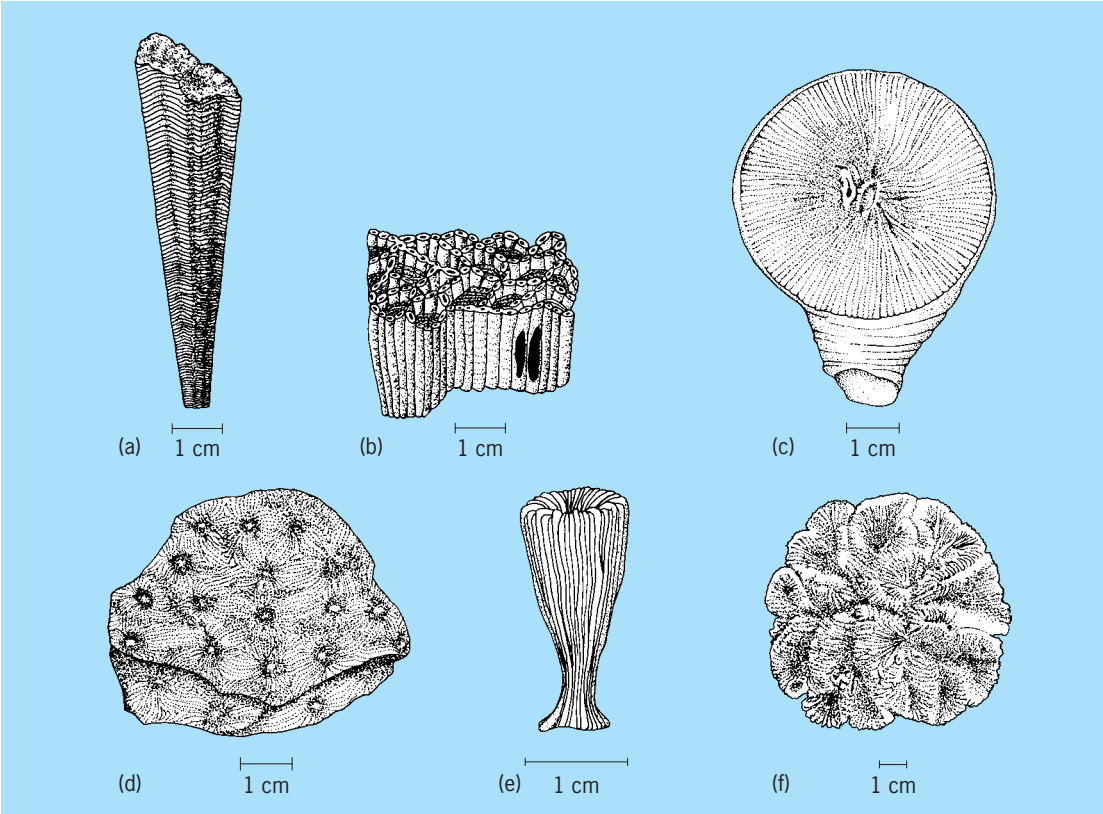

Fig. 6. Cnidarian fossils. (*a*) *Paraconularia*, Middle Paleozoic Conulata. (*b*) *Halysites*, Middle Paleozoic tabulate coral. (*c*) *Bethanyphyllum*, Devonian solitary rugose coral. (*d*) *Billingsastraea*, Devonian colonial rugose coral. (*e*) *Parasmilia*, Late Mesozoic and Cenozoic solitary scleractinian coral. (*f*) *Isophyllia*, Recent colonial scleractinian coral. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. F, Geological Society of America and University of Kansas Press, 1956*)

groups arose independently from soft-bodied anthozoans. These coral ancestors are often referred to as "anemones," and sometimes even as Actiniaria, but there is no evidence that those animals had the features characterizing modern sea anemones—or any other currently extant group of naked polyps. The origin of scleractinians is more mysterious because the first fossils of them are remarkably diverse, with many of the modern taxa represented. By the Jurassic, scleractinians had begun to form coral reefs similar to modern reefs of the tropics; it has been hypothesized that the acquisition of zooxanthellae is what allowed this development. However, many died out in the great Cretaceous-Tertiary (K-T) extinction, so that early Tertiary reefs were formed by bivalve mollusks, not corals. It was not until the Oligocene that scleractinians once again became dominant reef formers. *See* CRETACEOUS; TERTIARY.

Molecular evidence is likely to play a large part in understanding cnidarian evolution, since few members of the phylum leave phylogenetically informative fossils (because they are soft-bodied). A longstanding conundrum was whether Anthozoa is the basal group and development of the medusa generation followed or whether Anthozoa is derived from some ancestral cnidarian by loss of the medusa generation. Data that could be used to resolve this debate were problematic until the advent of gene analysis. Molecular analysis revealed that the mitochondrial chromosome in members of the Medusozoa (Cubozoa, Hydrozoa, Scyphozoa, and Staurozoa) is linear, whereas that of Anthozoa is circular, as in all other animals. This is clear evidence that Medusozoa is derived ("advanced"), relative to Anthozoa.

The phylogenetic relationships within each group are beginning to be understood, but many questions remain. *See* ANIMAL EVOLUTION; PHYLOGENY.

Daphne G. Fautin; Calvin H. Stevens

Bibliography. D. Doumenc, *Traité de Zoologie, Tome III: Cnidaires: Anthozoaires*, Fascicule 3, Masson, Paris, 1987; D. Doumenc, *Traité de Zoologie, Tome III: Cnidaires: Cténaires*, Fascicule 2, Masson, Paris, 1993; D. F. Dunn, Cnidaria, pp. 669–706 in S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, vol. 1, McGraw-Hill, New York, 1982; L. H. Hyman, Metazoa of tissue grid construction—the radiate phyla—Phylum Cnidaria, chap. 7, pp. 365–661 in *The Invertebrates: Protozoa through Ctenophora*, vol. 1, McGraw-Hill, 1940.

Cnidospora

A subphylum of spore-producing Protozoa containing the two classes Myxosporidea and Microsporidea. Myxosporidea contains the three orders Myxosporida, Actinomyxida, and Helicosporida. Microsporida is the only order in the class Microsporidea. Cnidospora are parasites of cells and tissues of invertebrates, fishes, a few amphibians, and turtles. *See* MICROSPORIDEA; MYXOSPORIDEA.

A distinctive structure of this subphylum is the spore, which contains one or more protoplasmic masses called sporoplasms. The spore also contains one or more filaments which lie coiled within the spore proper or within one or more polar capsules. The spore membrane may consist of a single chitinous piece or two or more parts called valves.

When the spore is ingested by a new host, the filaments are extruded and at the same time the sporoplasm is released. The sporoplasm, now called an amebula, reaches the specific site of infection directly through the intestinal wall or by way of the bloodstream. The amebula becomes a trophozoite, a stage in the life cycle, which feeds and grows at the expense of the host. Asexual reproduction by repeated binary fission, multiple fission, internal and external budding, or plasmotomy results in cells which develop into sporonts and eventually into sporoblasts, producing one or more spores. *See* SPOROZOA.

Because of similarity in multinucleated trophozoites, some protozoologists suggest that the Cnidospora have evolved from "slime mold–like" (Mycetozoia) ancestors. *See* PROTOZOA.

Ross F. Nigrelli

Coal

A brown to black combustible rock that originated by accumulation and subsequent physical and chemical alteration of plant material over long periods of time, and that on a moisture-free basis contains no more than 50% mineral matter. The plant debris accumulated in various wet environments, commonly called peat swamps, where dead plants were largely protected from decay by a high water table and oxygen-deficient water. The accumulating spongy, water-saturated, plant-derived organic material known as peat is the precursor of coal. Over time, many changes of the original vegetable matter are brought about by bacteria, fungi, and chemical agents. The process progressively transforms peat into lignite or brown coal, subbituminous coal, bituminous coal, and anthracite. This progression is known as the coalification series. Increasingly deeper burial under hundreds to thousands of feet of younger sediments is required to advance coalification to the bituminous coal and anthracite stages. The pressure exerted by the weight of the overlying sediment and the heat that increases with depth, as well as the length of exposure to them, determine the degree of coalification reached.

The types of plants contributing to peat accumulations vary greatly, depending on when in geologic time the peat was formed, that is, how far plants had evolved at the time, and on the regional and local environments of deposition (for example, water depth and its fluctuations, climate, and flooding by nearby streams or the sea). *See* FOSSIL FUEL; KEROGEN; LIGNITE; PEAT.

Formation

Economic coal deposits originate from accumulations of plant material (peat) in wetlands commonly

referred to as moor, swamp, mire, bog, fen, muskeg, or peatland. Most peat-forming plant material accumulates at or close to the place of growth (autochthonous origin), but some organic material may be transported before deposition, for instance, along beaches or in lakes (allochthonous origin). Autochthonous peats often form in various kinds of regionally coexisting depositional environments; low-lying swamps with forested, marshy, and open water environments (low moors) and raised peat bogs (high moors) are common peat-forming environments. Allochthonous plant material may contribute significantly to otherwise autochthonous peat deposits as open water areas develop and disappear in the evolutionary course of a peat swamp. Coal beds commonly are underlain by seat earths, remnants of ancient soils in which the initial swamp plants were rooted (hence the common occurrence of stigmaria in seat earths). *See* BOG; MUSKEG; SWAMP, MARSH, AND BOG.

Peat. Examples of present-day peat accumulation abound both in humid temperate climates and in subtropical to tropical climates. In North America, peat-forming environments are found in the Everglades of Florida, the Okefenokee Swamp of Georgia and Florida, the Dismal Swamp of Virginia, and in numerous depressions with poor drainage that were created during the ice ages in the plains between the Rocky and Appalachian mountains. In the tropics, thick peat deposits are known in the coastal areas of Sumatra and Borneo, for instance. Ultimate preservation of deposits of peat depends on long-term subsidence of the area and on burial under younger sediments. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; SEDIMENTOLOGY.

The nature of a peat deposit depends on the kind of plants composing the accumulation: sedges, reeds, and other herbaceous vegetation; forest trees and bushes; mosses; and so forth. Such differences create different coal facies that can still be identified in highly coalified coal seams. *See* COAL PALEOBOTANY.

Freshly deposited plant material is attacked by microorganisms, fungi, and chemical agents in air and water, enzymes and oxygen in particular. This commonly leads to degradation of the plant material to the point where the original plant structure is no longer recognizable, and new organic compounds are formed. Cellulose, a major constituent of cell walls, and proteins are much less resistant to decomposition than other chemical plant constituents such as lignin, tannins, fats, waxes, and resins. Humic acids and other humic substances are important decomposition products resulting from the peatification process. Fungi and aerobic bacteria depend on the presence of oxygen to degrade the organic material. During burial, as oxygen is consumed, anaerobic bacteria take over degradation. Woody material passes through a plastic, gelatinous stage before it solidifies to huminite and eventually vitrinite. As burial depth increases, overburden pressure causes increasing compaction that reduces the porosity of the mass and expels moisture, while increasing temperature with greater burial promotes chemical changes in the organic components of the coal. Peat is transformed into coal through the process of coalification. Important by-products of the coalification process are water, carbon dioxide, and, especially in the more advanced stages, methane. *See* CELLULOSE; DIAGENESIS; HUMUS; LIGNIN.

Coal seams. Minable coal seams occur in many different shapes and compositions. Some coal seams can be traced over tens, even hundreds, of miles in relatively uniform thickness and structure. The extensively mined Herrin coal bed of the Illinois Basin and the Pittsburgh coal bed of the northern Appalachian Basin are examples. They are 6–8 ft (2–2.5 m) thick over thousands of square miles. These coals originated in peat swamps that developed on vast coastal plains during the Pennsylvanian Period. The German brown coal deposits near Cologne are characterized by very thick coal deposits (300 ft or 100 m). However, their lateral extent is much more limited than are the two examples from the United States. These peat deposits formed in a gradually subsiding structural graben bounded by major faults. Land lay to the south and the sea to the north. Only a relatively small portion of the subsiding graben block provided optimal conditions for peat accumulation over a long period of time. Thus each coal bed has its own depositional history that determined many of its characteristics. *See* CARBONIFEROUS; CRETACEOUS; GRABEN; PENNSYLVANIAN; TERTIARY.

Coal seams are commonly composed of a number of benches of alternating coal and more or less carbonaceous shale. The shale represents periods when the peat accumulation was interrupted by flooding from a river or the sea, or, more rarely, interrupted by volcanic ash deposition (tonsteins). The individual benches of a coal bed vary laterally in thickness and composition, sometimes quite rapidly. The degree of variability is related to the stability of conditions during accumulation. Fluvial and lacustrine depositional environments produce greater lateral variability than deltaic or coastal plain environments. *See* SHALE.

Analyses

Many kinds of chemical, physical, petrological, and technological analyses are commonly performed on coal samples to ascertain characteristics and origin. The selection of analyses depends on the information needed. Many of these analyses have been standardized at both national and international levels [for example, American Society for Testing and Materials (ASTM) and International Standardization Organization (ISO)].

Megascopic coal description. Megascopic descriptions, often with the help of a hand lens, are routinely performed in the field, in mines, and on diamond drill cores. The detail of description varies between merely distinguishing coal from impure coal, carbonaceous shale, and other rocks, and quite detailed descriptions in terms of coal types or lithotypes.

Megascopic descriptions are used primarily during geologic investigations to characterize lateral

variations in thickness and structure. Their value in predicting specific coal properties such as ash content or coal rank is limited, and they are generally considered to be only the first step in the characterization of a coal deposit. However, megascopic descriptions are the principal basis for coal resource and reserve determinations.

Proximate analysis. This is probably the most commonly performed type of chemical analysis on coal samples. It involves determination of the ash, moisture, volatile matter, and fixed-carbon contents, which total 100%.

High-temperature ash. Combustion of the organic material at 1290–1380°F (700–750°C) leaves behind the high-temperature ash. At these high temperatures, certain minerals disintegrate, leading to both loss of volatile products and reaction with oxygen from the air to form oxides. Some volatile decomposition products react with the ash before they can escape. Therefore, the weight of the high-temperature ash may differ significantly from the weight of the coal's original mineral matter. High-temperature ash is composed primarily of unaltered, altered, and newly formed oxides and silicates (in particular, unaltered quartz and feldspars, altered clay minerals, and newly formed mullite, hematite, and magnetite).

Moisture content. This term has different meanings depending on when, how, and on what sample type it is determined, for example, as received, surface, air-dried, bed, inherent, and equilibrium moistures. The determination may be made on raw or cleaned coal, on selected grain sizes, and so forth. The moisture content in the fully saturated stage, but without presence of surface moisture, is used for determining coal rank. Bed, inherent, and equilibrium moistures are representative of this type of moisture content.

Volatile matter and fixed carbon. These are derived from a small-scale retorting test which is performed under rigidly controlled conditions at a temperature of 1740 ± 36°F (950 ± 20°C). The process results in the disintegration of coal into volatile matter, a complex mixture of gases (exclusive of water), and a cokelike residue rich in carbon (fixed carbon).

Mineral matter. This term refers to both inorganic constituents that occur as discrete mineral grains and so-called inorganic elements that are more or less homogeneously distributed through coal. Minerals that were present before combustion can be approximated from a chemical analysis of the ash (through a normative analysis). A more accurate way to analyze for mineral matter in coal is by oxidation of the organic matter in a low-temperature asher in which an oxygen plasma, activated by a radio-frequency field, is passed over the coal sample. Because the oxidation occurs at the relatively low temperature of 250–300°F (120–150°C), most minerals remain essentially unaltered and the resulting ash can be subjected to a wide range of analyses (in particular, x-ray diffraction) to determine the minerals present. The Parr formula provides a simple method to estimate the original total amount (but not the composition) of the mineral matter in a given coal; its main purpose is to adjust the ash weight by correcting for the disintegration of iron sulfides and other reactions during combustion (that is, loss of sulfur dioxide, and reaction of iron with oxygen in air). *See* MINERAL; X-RAY DIFFRACTION.

Ultimate (elemental) analysis. The elements carbon, hydrogen, oxygen, sulfur, and nitrogen make up essentially all the organic substance of coal and are determined, along with moisture and ash, by the ultimate analysis. The first four of these so-called organic elements are also present in common inorganic components of coal such as the moisture, and the carbonate, sulfide, sulfate, and oxide minerals, making attribution to either the organic or the inorganic constituents of coal potentially difficult. Ultimate analyses are significantly more involved than proximate analyses. Both analyses are now commonly done on sophisticated computer-controlled instruments.

Calorific or heating value. This parameter is determined by observing temperature changes that result from combustion of a known quantity of coal in an adiabatic bomb calorimeter. In the United States, the calorific value is reported in British thermal units given off per pound of coal burned (Btu/lb); the equivalent SI unit is megajoules per kilogram (MJ/kg) or, until relatively recently, cal/g or kcal/kg (1000 Btu/lb = 2.326 MJ/kg = 560 kcal/kg). The calorific value of a coal is influenced not only by the nature of the organic material of which it is made but also by the proportion of noncombustible material present, primarily mineral matter and moisture. *See* CALORIMETRY; THERMOCHEMISTRY.

Major, minor, and trace elements. Coal contains most elements in the periodic table, at least in trace amounts. Elements not yet identified may be present below detection limits. The total concentrations of the ash-forming elements in whole coals ranges from more than 0.5% for major inorganic elements such as aluminum, calcium, iron, and silicon; to 0.02–0.5% for minor elements such as potassium, magnesium, sodium, titanium, chlorine, and others; to less than 0.02% (200 parts per million) for trace elements such as arsenic, beryllium, cadmium, and mercury. Knowledge of the concentration of these elements is required when evaluating coals for combustion and conversion (for example, relative to corrosion, boiler fouling, and poisoning of catalysts) and to assess potential negative environmental and health impacts resulting from coal use. *See* PERIODIC TABLE.

In the United States, 16 minor and trace elements in coal are included in the list of hazardous air pollutants (HAPs) of the 1990 Amendments to the Clean Air Act (**Table 1**).

Analytical methods that are frequently used to determine minor and trace elements in coal include flame and graphite furnace atomic absorption spectrometry, x-ray fluorescence (XRF) spectrometry, optical emission photographic spectrometry, and neutron activation analysis. Inductively-coupled-plasma (ICP) atomic emission spectrometry and ICP mass spectrometry have been used because of their higher sensitivity for certain elements. To study spatial distribution of trace elements, techniques such as laser ablation ICP mass spectrometry, secondary-ion

TABLE 1. Minor- and trace-element content of cleaned coal shipped by mines in Illinois, in parts per million

Element	Arithmetic mean	Standard deviation
Antimony	0.9	0.7
Arsenic	7.5	8.1
Beryllium	1.2	0.7
Cadmium	0.5	0.9
Chlorine	1671.0	1189.0
Chromium	14.0	6.1
Cobalt	3.5	1.3
Fluorine	93.0	36.0
Lead	24.0	21.0
Manganese	38.0	32.0
Mercury	0.09	0.06
Nickel	14.0	5.0
Phosphorus	87.0	83.0
Selenium	1.9	0.9
Thorium	1.5	0.4
Uranium	2.2	1.9

mass spectrometry, and synchrotron-radiation XRF microprobe have been developed. *See* ACTIVATION ANALYSIS; ATOMIC SPECTROMETRY; MASS SPECTROMETRY; X-RAY FLUORESCENCE ANALYSIS.

Macerals. Macerals are the microscopically recognizable constituents of coal. Microscopic analyses are normally performed on representative crushed coal samples embedded in epoxy and polished for observation under incident light, generally with oil immersion objectives, at a magnification of 400–750×. A sufficiently large number (500–1000) of points evenly distributed over the sample are selected for maceral analysis to determine the percent of macerals in the sample by volume. The measuring spots are only a few micrometers in diameter.

The reflected-light method permits identification of a few nonclay minerals during maceral analysis. However, much more accurate determinations of the amount and composition of mineral matter are accomplished through other analyses, in particular, low-temperature ashing followed by x-ray diffraction analysis of the ash.

Reflectance of vitrinite. The percentage of normally incident light that is reflected from the polished surface of vitrinite under oil immersion has become a commonly used measure of coal rank. A reflected-light microscope equipped with a photomultiplier is used. The small size of the measuring spot, about 5 micrometers in diameter, and the large magnification used (500–1000×) permit selection of highly comparable material for this analysis, in contrast to chemical analyses that test bulk coal with variable mixtures of macerals and minerals. Vitrinite is the preferred material on which to measure reflectance, because of its common occurrence in most coals, its homogeneous appearance, its easy recognition under the microscope, and the considerable range in reflectance exhibited by the maceral, depending on a coal's rank.

Density, porosity, and internal surface area. Because of the large proportion of very small pores (a few angstroms to a few tens of angstroms in diameter) in coal, true values for these parameters are difficult to obtain with standard methods. Only very small molecules or atoms (for example, helium) are able to penetrate these micropores. Determinations with gases or liquids whose molecules are of the same order of magnitude as or larger than the micropores and their entrances yield only apparent values. For example, in one experiment the internal surface area of a sample of bituminous coal was measured separately with nitrogen and then carbon dioxide; the values obtained were 9 and 205 m^2/g, respectively. Another method to estimate the internal surface area is by measuring the heat of wetting with various liquids; but with this method, too, the molecules of the liquids do not necessarily penetrate all pores and thus do not measure the entire internal surface area.

Washability. A large proportion of mined bituminous and higher-rank coal is cleaned before it is used. Washability tests are performed in the laboratory on drill core and test pit samples to predict the cleanability of a given coal when mined. Coal samples are crushed to sizes comparable to those anticipated to be produced by the mine and are subjected to float-sink tests in heavy liquids. The amounts floating and sinking in each gravity fraction are measured, and their ash and sulfur contents as well as other parameters (for example, heating value) are determined. The data are tabulated and plotted to produce washability curves that characterize the coal's cleanability.

The fine fraction of mined coal [<0.3–0.5-mm (0.01–0.02-in.) grain size] is difficult or impossible to clean by the float-sink method. Instead of utilizing the differences in specific gravity of particles of the crushed raw coal, differences in the surface properties of coal and minerals (hydrophobic versus hydrophilic) are employed in the flotation method used to clean coal fines. However, since this method is relatively expensive, fine coal has commonly been discarded (in slurry ponds) without cleaning.

Plastic and coking properties. A significant use of coal is in making coke, which involves heating finely crushed coal in coke ovens in the absence of air. Bituminous coals generally become plastic or fluid during heating as they disintegrate into gases, liquids, and solids (char, coke); certain bituminous coals, commonly blends of several coals of different rank, can form a coke that is strong enough for use in blast furnaces. Several empirical tests have been used widely for many years to characterize the behavior of coal during heating and to predict its coke-forming ability.

Free Swelling Index (FSI). This is probably the most widely used test. It requires rapid heating of 1 g (0.35 oz) of finely ground coal to a temperature of 1508 ± 9°F (820 ± 5°C) in a lid-covered crucible. A coke button is formed, and its shape is compared to standard shapes 1–9 and reported in half-step increments. The FSI test provides a rapid and inexpensive first assessment of the coke-forming behavior of a coal.

Plastometer test. In a Gieseler plastometer, 5 g of finely crushed coal is heated in a retort; a small stirrer submerged in the sample applies constant torque while the temperature is raised at a rate of 5.4°F/min

(3°C/min). As the coal becomes plastic and fluid during heating, the stirrer's rate of rotation is recorded. The temperatures of initial softening, maximum fluidity, and resolidification, as well as the maximum dial movement in dial divisions per minute (ddpm), are recorded. These data differ characteristically for coals of different rank and composition and provide good indicators of a coal's coking properties.

Dilatometer test. This test is widely used outside the United States to characterize the behavior of coal during heating in the range 570–1020°F (300–550°C). A finely ground sample of coal is heated inside a crucible with a piston resting on the sample. The displacement of the piston is recorded as a percentage of the original sample volume. While the sample is heated at a rate of 5.4°F/min (3°C/min), the coal becomes plastic and fluid and releases gas, expands, then solidifies to coke. Volume change is plotted against temperature; the resulting curve reveals much about the coking characteristics of a given coal.

Other tests. The Gray-King and ROGA tests, commonly performed in several European countries, are essentially standardized laboratory-scale coking tests which produce useful index numbers. *See* COKE.

Strength and hardness. Strength and hardness are important coal properties in mine stability, coal winning, and comminution for preparation and utilization. A number of tests are commonly performed in the laboratory to obtain information on a coal's compressive and shearing strengths and hardness; an example is the Hardgrove grindability test. The values thus obtained are only index numbers because of the small size of sample used. They must be translated with caution when full-scale industrial operations are involved.

Ash fusion. During combustion of coal in boilers, the noncombustible portion, such as minerals and inorganic elements, either is volatilized or becomes ash, which forms deposits within the boiler or is removed as fly ash or bottom ash. The ash-fusion test identifies the temperatures at which the ash undergoes major phase changes. In general, for combustion uses of coal, ash that fuses at high temperatures (above 2400°F or 1538°C) is preferred to ash that fuses at low temperatures (below 2000°F or 1093°C). The range of ash-fusion temperatures is about 1600–2800°F (871–1538°C). The properties and relative abundances of minerals present may be critical in the formation of ash deposits in a boiler, as well as in the corrosion of boiler tubes. High alkali and alkali chloride contents are known to cause such problems in boilers. Boilers are designed for certain ash fusion properties.

Classification Schemes

The parameters obtained from coal analyses are used to classify coal for general as well as specific purposes. Both national and international standardization organizations have adopted classification schemes that are used in the coal trade, for statistical purposes, and to compile coal resources and reserves by classification category. One of the most widely used classification systems is the American Society for Testing and Materials (ASTM) Standard Classification of Coals by Rank (**Table 2**).

The European Commission for Europe—UN (UN/ECE) Working Party on Coal developed a new

TABLE 2. American Society for Testing and Materials classification of coals by rank (in box, ASTM standard D 388) and other related coal properties

ASTM class	ASTM group	1000 Btu/lb[1]	Agglomerating[2]	Volatile matter,[3] %	MJ/kg[1]	Maximum reflectance,[4] %	Moisture,[1] %	C,[3] %	O,[3] %	H,[3] %
	Peat	1.0–6.0	No	72–62	2.3–14.0	0.2–0.4	95–50	50–65	42–30	7–5
Lignite	Lignite B	>6.3[5]	No	65–40	<14.7[5]	0.2–0.4	60–40[5]	55–73	35–23	7–5
	Lignite A	6.3–8.3[5]	No	65–40	14.7–19.3[5]	0.2–0.4	50–31[5]	55–73	35–23	7–5
Subbituminous	Subbituminous C	8.3–9.5[5]	No	55–35	19.3–22.1[5]	0.3–0.6	38–25[5]	60–80[6]	28–15[6]	6.0–4.5
	Subbituminous B	9.5–10.5[5]	No	55–35	22.1–24.4[5]	0.3–0.6	30–20[5]	60–80[6]	28–15[6]	6.0–4.5
	Subbituminous A	10.5–11.5[5]	No	55–35	24.4–26.7[5]	0.3–0.7	25–18[5]	60–80[6]	28–15[6]	6.0–4.5
Bituminous	High volatile C	10.5–13.0[5]	Yes	55–35	24.4–30.2[5]	0.4–0.7	25–10[5]	76–83[6]	18–8[6]	6.0–4.5
	High volatile B	13.0–14.0[5]	Yes	50–35	30.2–32.6[5]	0.5–0.8[6]	12–5[6]	77–84[6]	12–7[6]	6.0–4.5
	High volatile A	≥14.0	Yes	45–31	≥32.6	0.6–1.2[6]	7–1[5]	78–88[6]	10–6[6]	6.0–4.5
	Medium volatile	>14.0	Yes	31–22[5]	>32.6	1.0–1.7[5]	<1.5	84–91	9–4	6.0–4.5
	Low volatile	>14.0	Yes	22–14[5]	>32.6	1.4–2.2[5]	<1.5	87–92	5–3	6.0–4.5
Anthracite	Semianthracite	>14.0	No	14–8[5]	>32.6	2.0–3.0[5]	<1.5	89–93[6]	5–3[6]	5–3[5]
	Anthracite	>14.0	No	8–2[5]	>32.6	2.6–6.0[5]	0.5–2	90–97[6]	4–2[6]	4–2[5]
	Meta-anthracite	>14.0	No	≤2	>32.6	>5.5[5]	1–3	>94[6]	2–1[6]	2–1[5]

[1] Moist, mineral-matter-free.
[2] Agglomerating coals form a button of cokelike residue in the standard volatile-matter determination that shows swelling or cell structure or supports a 500-g weight without pulverizing. The residue of nonagglomerating coal lacks these characteristics.
[3] Dry, mineral-matter-free.
[4] Reflectance of vitrinite under oil immersion.
[5] Well suited for rank discrimination in range indicated.
[6] Moderately well suited for rank discrimination.
SOURCE: Modified from H. H. Damberger et al., in B. R. Cooper and W. A. Ellingson (eds.), *The Science and Technology of Coal and Coal Utilization*, 1984.

standard that attempts to incorporate both coal rank and coal facies parameters; this standard is intended for the characterization of coal resources in the ground. Another group of experts of ECE developed a codification system for coals traded. This standard replaces the UN/ECE's 1956 classification scheme for hard coal, which combines rank (class), caking (group), and coking (subgroup) parameters into a three-digit code system; this scheme has not been widely used.

For the purpose of classification, coal analyses must be standardized, through computation, to a common basis, such as ash-free (af), moisture-free (mf) or dry (d), moisture-and-ash-free (maf) or dry ash-free (daf), mineral-matter-free (mmf), dry mineral-matter-free (dmmf), and so forth. Generally, chemical laboratories determine values on an air-dried basis; from these all other bases can be easily computed.

Coal Facies

The coal facies reflects the environments of deposition during accumulation of plant material. The types of plants growing, their preservation, and the kinds and amounts of mineral matter introduced vary with the depositional environments. Coal facies can be recognized both megascopically and microscopically through all or most ranks.

Nonbanded coal. Nonbanded coals are homogeneously fine-granular and devoid of distinct megascopic layers. They have a greasy luster and conchoidal fracture. They formed from subaqueous deposition of finely comminuted plant detritus, commonly with significant amounts of spore or pollen grains, algal remains, wax and resin granules, and various other small plant fragments. Well-known nonbanded coals are cannel coal (composition dominated by spores) and boghead coal (predominance of algae). Transitions exist to oil shales. Their proper identification requires microscopic analysis. *See* OIL SHALE; SAPROPEL; TORBANITE.

Nonbanded coals are relatively rare. They may form entire coal seams or layers within banded coal.

Banded coal. This is the most common coal type. In bituminous coal, recognizable bands are composed of lithotypes such as vitrain, attrital coal, and fusain, and of impure coal. The banding is less pronounced in lower rank (subbituminous, lignitic) coals and may be difficult to discern in anthracites. Most bands are between a fraction of a centimeter to several centimeters thick.

Vitrain. This lithotype, also referred to as bright coal, is highly lustrous in bituminous coals and enhances their banded appearance. Vitrain represents the coalified remains of relatively large fragments of wood and the bark of stems, branches, and roots of trees and bushes. In bituminous coal, it forms bands and lenses of very bright coal, generally a few millimeters to 10 mm thick. In lignite the remains of woody material lack the shiny luster.

Attrital coal. This lithotype represents the fine-grained ground mass of banded coal in which layers of vitrain and fusain are embedded. Attrital coal is highly variable in appearance and composition. It exhibits a finely striated, granulose, or rough texture, much like nonbanded coal. Relatively lustrous attrital coal is often referred to as clarain, dull attrital coal as durain. Dull attrital coal commonly contains relatively high amounts of finely dispersed mineral matter, clays and quartz in particular.

Fusain. In most banded coals fusain represents only a subordinate proportion of the total volume, and it generally occurs in lens-shaped laminae and thin bands that rarely exceed a few centimeters in thickness. Fusain resembles charcoal; it is fibrous, very friable, and has a silky luster. It tends to be concentrated on selected bedding surfaces, forming planes of weakness in coal beds. Its many pores are often filled with secondary mineral matter that strengthens it and reduces its friability.

Impure coal. This type of coal is defined as containing 25–50% of ash by weight on the dry basis. The mineral impurities were introduced either during deposition, mostly as clay or as fine-grained detrital mineral matter (bone coal), or later as secondary mineralization (mineralized coal). Bone coal has a dull appearance; it is commonly finely striated. When scratched with a knife or nail, it forms a brown powder; the scratch mark appears shiny. Pyritized coal is the most common mineralized coal.

Maceral groups. Macerals are the microscopically recognizable organic constituents of coal, equivalent to minerals of other rocks. However, unlike minerals, their chemical and physical properties are not fixed but vary over a wide range depending upon coal rank and facies. Submicroscopic inorganic impurities are considered part of macerals.

Three major groups of macerals are generally recognized: vitrinites, exinites (or liptinites), and inertinites. These maceral groups are divided into maceral subgroups and, depending on the degree of detail desired, even further (**Table 3** and **illus.**). In polished sections, the macerals are distinguished on the basis of their reflectance and morphology, and to some extent on their size and maceral association.

Vitrinite. This is the predominant group of macerals in most coals (commonly 70–85% by volume). The reflectance of vitrinite is intermediate between the darker exinite and brighter inertinite (see illus.).

TABLE 3. Classification of macerals*

Maceral group	Maceral
Vitrinite	Vitrinite
Exinite (or liptinite)	Sporinite
	Cutinite
	Resinite
	Alginite
Inertinite	Fusinite
	Macrinite
	Micrinite
	Sclerotinite
	Semifusinite
	Inertodetrinite

*After American Society for Testing and Materials, standard D 2796.

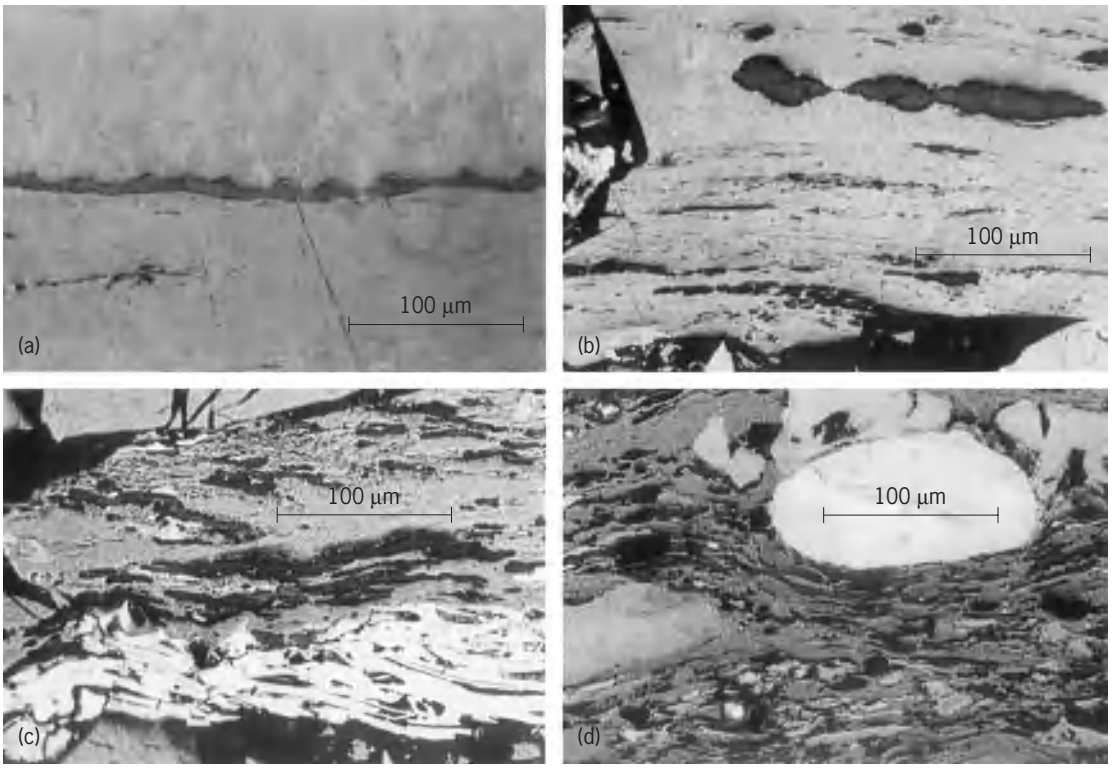

Macerals in polished surfaces of bituminous coals in reflected light, under oil immersion. (*a*) Thin, "saw-toothed" cutinite (dark) enclosed in vitrinite. (*b*) Elongate globs of resinite (dark) and other, smaller exinite macerals (dark) enclosed in vitrinite. (*c*) Sporinite (dark lenses) in vitrinite (gray) with fusinite (white) on lower part of photo. (*d*) Large grains of macrinite (white), smaller inertodetrinite (light gray), and sporinite and other exinites (dark gray) interlayered with vitrinite (gray).

Vitrinite generally originated from woody plant parts (tree trunks, branches, roots, and bark). The original cellular plant structure is still more or less visible in many vitrinites. The identifiable cell walls are known as telinite, while the structureless vitrinite is known as collinite. These two are sometimes further subdivided into submacerals for very detailed investigations.

The reflectance of vitrinite is a good indicator of coal rank, especially in higher-rank coals. Viewed under oil immersion, reflectance values range from 0.2–0.3% in low-rank lignites to over 4% in anthracites (Table 2).

Exinite (or liptinite). The macerals of the exinite group reflect light distinctly less than vitrinite (see illus.). Their volatile matter and H contents are significantly higher than those of associated vitrinite. These differences in reflectance, volatile matter, and H contents disappear quickly with increasing rank in coals above about 1% R_{oil}, or about 30% volatile matter (dry ash-free), and converge at about 1.4% R_{oil}, or about 20% volatile matter. The origin of the macerals of the exinite group is clearly implied by their names (Table 3); they are derived from algae, spores, cuticles, resins, waxes, fats, and oils of plants.

Inertinite. Inertinites reflect light significantly more than associated vitrinites (see illus.), except in high-rank anthracites. Inertinites do not become plastic during heating, such as in coke making; they react as inert materials. Fusinite is characterized by well-preserved cell wall structure (sometimes broken by compaction, a characteristic known as bogenstructure). Cell cavities commonly are filled with minerals. Fusinite's general appearance suggests origin by rapid charring in the peat stage (for example, in a forest fire). On the other hand, macrinite and micrinite lack visible cell structure. Particle size is greater than 10 μm for macrinite and less than 10 μm, commonly about 1 μm, for micrinite. Micrinite appears to be a solid by-product of the coalification of liptinite macerals in the bituminous rank range. Sclerotinite occurs as high-reflecting round to oval cellular bodies of varying size (20–300 μm) or as interlaced tissues derived from fungal remains. Semifusinite is transitional between fusinite and associated vitrinite in both reflectance and cell structure. Inertodetrinites are small, angular clastic fragments of other inertinites commonly enclosed in vitrinite or associated with mineral matter.

The presence of inertinites indicates oxidative conditions during the peat stage of coal formation. Inertinite contents of 5–15% are the rule for many coals of the eastern United States and Europe. But some coals have inertinite contents in the 30–40% range (for example, the High Splint and Winefrede coal seams of eastern Kentucky), indicating strong oxidizing conditions during peat formation. The mostly Permian-age Gondwana coals of Australia, India, South Africa, and South America are well known for their high inertinite contents, up to 85%. These coals were formed in a relatively cold climate with alternating dry and rainy seasons, providing

TABLE 4. Minerals frequently occurring in coals and their stoichiometric compositions, relative abundances, and modes of occurrences*

Mineral	Composition	Common minor- and trace-element associations	Frequency of occurrence in coal seams	Concentration in mineral matter	Chief occurrences	
					Physical†	Genetic‡
Clay minerals						
Illite (sericite)	$KAl_2(AlSi_3O_{10})(OH)_2$	Sodium, calcium, iron, lithium, titanium, manganese, fluorine, and other lithophile elements	Common	Abundant	D,L	d,s(?)
Smectite (including mixed layered)	$Al_2Si_4O_{10}(OH)_2 \cdot H_2O$		Common	Abundant	D,L	d,s(?)
Kaolinite group	$Al_2Si_2O_5(OH)_4$		Common	Abundant	L,F	d,e,s(?)
Chlorite	$Mg_5Al(AlSi_3O_{10})(OH)_8$		Rare	Moderate	L	d
Sulfides						
Pyrite	FeS_2 (isometric)	Arsenic, cobalt, copper, and other chalcophile elements	Rare to common	Variable	D,N,F	s,e
Marcasite	FeS_2 (orthorhombic)		Rare to moderate	Trace	D(?)	s(?)
Pyrrhotite	$Fe_{1-x}S$		Rare	Trace	D	s(?)
Sphalerite	ZnS		Rare	Minor to trace	F	e
Galena/chalcopyrite	$PbS/CuFeS_2$		Rare	Trace	F	e
Carbonates						
Calcite	$CaCO_3$	Manganese, zinc, strontium	Rare to common	Abundant	N,F	e,s
Dolomite (ankerite)	$CaMg(CO_3)_2$		Moderate	Trace	N,L	s,e
Siderite	$FeCO_3$		Rare	Minor	N	s,e
Oxides						
Quartz	SiO_2	–	Common	Abundant	D,L,N	d,s(?)
Magnetite/hematite	Fe_3O_4/Fe_2O_3	Manganese, titanium	Common	Minor to trace	N	s
Rutile and anatase	TiO_2	–	Common	Trace	D	d(?)
Others						
Goethite/limonite	FeO(OH)	Manganese, titanium	Common	Trace	N	w
Feldspar	$K(Na)AlSi_3O_8$	Calcium	Moderate	Trace	D,L	d
Zircon	$ZrSiO_4$	–	Moderate	Trace	D	d
Sulfates						
Gypsum	$CaSO_4 \cdot 2H_2O$	–	Moderate	Minor	D,F	w
Barite	$BaSO_4$	Sodium, strontium, lead	Rare	Minor	F	e
Szomolnokite	$FeSO_4 \cdot H_2O$	–	Rare	Trace	D	w
Apatite	$Ca_5(PO_4)_3(F,Cl,OH)$	Manganese, cerium, strontium, uranium	Moderate	Trace	D	d,s(?)
Halite	NaCl	Potassium, magnesium	Rare	Trace	D	e

*After R. D. Harvey and R. R. Ruch, Mineral matter in Illinois and other U.S. coals, in K. S. Vorres (ed.), *Mineral Matter and Ash in Coal*, Amer. Chem. Soc. Symp. Ser. 301, pp. 10–40, 1986.
†D, disseminated; L, layers (partings); N, nodules; F, fissures (cleat). First listed is the most common occurrence.
‡d, detrital; s, syngenetic; e, epigenetic; w, weathering. First listed is the most common occurrence. (?) indicates there is divergence of opinion about the genetic occurrence.

ample opportunity for widespread oxidation of the accumulating plant material. In contrast, the Pennsylvanian-age coals of North America and Europe originated in the humid equatorial zone.

Mineral matter. Whether they occur as discrete mineral grains or are disseminated through coal macerals, the inorganic components are referred to collectively as mineral matter. The low-temperature oxygen-plasma ashing technique developed in the mid-1960s has permitted detailed analyses of the minerals found finely dispersed in coal (**Table 4**). Mineral matter occurs in coal seams as discrete layers or partings, as nodules, as fissure fillings, as rock fragments, and as small particles finely disseminated throughout. Genetically, minerals can be classified as inherent (plant-derived), detrital (water- or wind-borne during sedimentation), syngenetic (formed early in the peat stage), and epigenetic (formed during and after lithification and during weathering). There is no sharp boundary between syngenetic and epigenetic origins. Minerals, just as macerals, indicate the environments that prevailed during their incorporation into peat or later during coalification. Similarly, the inorganic major, minor, and trace elements of coal were introduced at various stages of coal formation. Multivariate statistical methods permit attribution of most minor and trace elements to specific mineral phases.

Coal Rank

During burial to greater and greater depth, peat is transformed into brown coal or lignite, subbituminous coal, bituminous coal, and anthracite (coalification series) as a result of increasing overburden pressure and rising temperature. Both physical (pressure, heat) and chemical (biochemical, thermochemical) factors are influential in the transformation of peat into the other members of the coalification series (**Table 5**). The boundaries between the members of the series are transitional and must be chosen somewhat arbitrarily. The term rank is used to identify the stage of coalification reached in the course of coal metamorphism. Rank is a fundamental property of coal, and its determination is essential in the characterization of a coal. Classification of coal by rank generally is based upon the chemical composition of the coal's ash-free or mineral-matter-free organic substance (Table 2), but parameters derived from empirical tests indicative of technological properties, such as agglomerating characteristics (Table 2) or the Free

TABLE 5. Examples of typical proximate and ultimate analyses of samples of each rank of common banded coal in the United States*

		Proximate analysis, % (as-received basis)				Ultimate analysis, % (dry basis)					
Rank	State	Moisture	Volatile matter	Fixed carbon	Ash	Sulfur	Hydrogen	Carbon	Nitrogen	Oxygen	Heating value, Btu/lb† (as-received basis)
Anthracite	Pa.	4.4	4.8	81.8	9.0	0.6	3.4	79.8	1.0	6.2	13,130
Semianthracite	Ark.	2.8	11.9	75.2	10.1	2.2	3.7	78.3	1.7	4.0	13,360
Bituminous coal											
Low volatile	Md.	2.3	19.6	65.8	12.3	3.1	4.5	74.5	1.4	4.2	13,220
Medium-volatile	Ala.	3.1	23.4	63.6	9.9	0.8	4.9	76.7	1.5	6.2	13,530
High-volatile A	Ky.	3.2	36.8	56.4	3.6	0.6	5.6	79.4	1.6	9.2	14,090
High-volatile B	Ohio	5.9	43.8	46.5	3.8	3.0	5.7	72.2	1.3	14.0	13,150
High-volatile C	Ill.	14.8	33.3	39.9	12.0	2.5	5.8	58.8	1.0	19.9	10,550
Subbituminous coal											
Rank A	Wash.	13.9	34.2	41.0	10.9	0.6	6.2	57.5	1.4	23.4	10,330
Rank B	Wyo.	22.2	32.2	40.3	4.3	0.5	6.9	53.9	1.0	33.4	9,610
Rank C	Colo.	25.8	31.1	38.4	4.7	0.3	6.3	50.0	0.6	38.1	8,580
Lignite	N. Dak.	36.8	27.8	30.2	5.2	0.4	6.9	41.2	0.7	45.6	6,960

*After *Technology of Lignitic Coals,* U.S. Bur. Mines Inform. Circ. 769, 1954. Sources of information omitted.
†1000 Btu/lb = 2.326 MJ/kg.

Swelling Index, as well as the ROGA, Gieseler, and Gray-King tests, are commonly used in several countries outside the United States in addition to coal rank parameters. *See* METAMORPHISM.

Coal rank increases with depth at differing rates from place to place, depending primarily on the rate of temperature increase with depth (geothermal gradient) at the time of coalification. Coal rank also changes laterally, even in the same coal seam, as former depth of burial and thus exposure to different pressure and temperature vary. Originally established vertical and regional coalification patterns can be significantly altered by various kinds of geologic events, such as the intrusion of large magma bodies at depth (plutonism), or renewed subsidence of a region. Volcanic activity may cause significant local anomalies in coal rank, but rarely leads to regional changes in coalification pattern. *See* MAGMA.

The chemical parameters customarily used to rank coal may yield misleading results in coals that are not primarily composed of vitrinite. Coals rich in exinite may have significantly higher contents of volatile matter and coals rich in inertinite may have significantly lower contents of volatile matter than associated so-called normal coals. This problem can be minimized by the use of vitrinite reflectance to indicate rank. Reflectance data have the additional advantage of being applicable to small coaly particles that occur dispersed throughout most sediments, permitting rank measurements in strata devoid of coal beds.

As rank increases during maturation, most coal properties change in tandem. This is reflected in the values given in Tables 2 and 5, but can also be represented graphically in correlation charts.

Occurrence and Resources

During geologic history, coal could not form in economic quantities until vascular land plants evolved during the Devonian Period some 400 million years ago and until they conquered the continents during the Carboniferous Period. The Carboniferous (Mississippian and Pennsylvanian of North America) witnessed an explosive development of vast swamp forests along the Euramerican equatorial coal belt, which extended from the North American midcontinent to Russia. The Atlantic Ocean did not exist then, and the North American and European continents were fused together as part of a continent called Laurasia. Another important coal-forming belt stretched over what was then Gondwana of the Southern Hemisphere. Gondwana subsequently drifted north and broke up into South America, Africa, India, and Australia. Most major coal deposits on Gondwana are of Permian age. They formed in a cool, humid climate. Most of the vast coal deposits of the western United States and Canada, along the eastern margin of the Rocky Mountains, originated during the Cretaceous and Tertiary periods. *See* CARBONIFEROUS.

Not all coal deposits are of economic interest. Governments and private industry need to know how much coal exists in the ground (resources) and what portion may be available for development under current economic and legal conditions (reserves). Various classification schemes have been adopted by national and international agencies for determination of coal resources and reserves on comparable bases. However, no universally accepted standard has yet emerged. An international framework for the classification of reserves/resources for solid fuels and mineral commodities has been developed under the auspices of the United Nations to fill the gap.

Two independent factors generally are considered in these assessments: the degree of economic minability, and the degree of geologic assurance of the existence of a coal deposit of given properties. The World Energy Conference (WEC) has attempted to overcome the lack of standardization between countries by adopting carefully worded definitions of resource and reserve categories, and has asked national organizations to interpret their databases in these terms. **Table 6** is a summary of the worldwide compilation by the World Energy Conference.

TABLE 6. World coal resources and reserves, in 10^9 metric tons*

Continent or country	Proved amount in place	Estimated recoverable reserves	Estimated additional reserves in place	Additional recoverable reserves
Africa	134	62	114	0.1
South Africa	121	55	5	0
North America	455	253	1199	2
United States	431	241	1139	0
South and Central America	21	10	46	7
Brazil	10	3	22	6
Asia (excl. former Soviet Union)	562	228	767	198
China	287	115	668	150
India	223	70	90	47
Former Soviet Union	287	241	5200	0
Australia	117	91	704	433
Europe (excl. former Soviet Union)	819	150	552	39
Former CSSR	571	5	5	2
Germany	122	67	186	0
Poland	79	42	144	27
United Kingdom	1.0	2.5	190	0
New Zealand, New Caledonia	0.6	0.1	8	<0.1
World total	2396	1034	8590	680

*After survey by World Energy Council, 1995.

Uses

Coal is used primarily for producing steam in electric power plants [87% of coal consumed in the United States in 1997, 67% in countries of the Organization for Economic Cooperation and Development (OECD) in 1996, and 52% in the world in 1996]. Other important uses are by industry for producing steam and heat (9% United States, 12% OECD), and by the steel industry for coke making (3% United States, 7% OECD). Conversion of coal to synthetic liquid or gaseous fuels does not constitute a major use of coal worldwide or in most countries. However, coal conversion has been important under special circumstances in some countries. During World War II, Germany produced a significant amount of liquid fuels from coal, and South Africa, when it became isolated because of its apartheid policy, established a major coal liquefaction industry to reduce its dependence on oil imports. One large coal gasification plant in the United States produces synthetic natural gas from North Dakota lignite; it was built with government support. It will take a significant rise in the price of oil and natural gas to make conversion into synthetic liquid and gaseous hydrocarbons competitive in the future. *See* SYNTHETIC FUEL.

The massive Clean Coal Technology (CCT) cooperative program between the United States government and private industry that began in 1988 included several large-scale projects where coal is gasified, the gas cleaned and then burned on site, first to drive high-temperature gas turbines, and then to generate steam for steam turbines; this makes for very clean and high-efficiency electric power generation. The CCT program has also demonstrated a wide range of other technologies to produce energy from coal cleanly and efficiently. *See* COAL CHEMICALS; COAL GASIFICATION. Heinz H. Damberger

Bibliography. American Society for Testing and Materials, *Annual Book of ASTM Standards*, vol. 05.05, sec. 5: *Petroleum Products, Lubricants and Fossil Fuels*; N. Berkowitz, *The Chemistry of Coal: Science and Technology*, 1985; B. R. Cooper and W. A. Ellingson (eds.), *The Science and Technology of Coal and Coal Utilization*, 1984; I. Demir et al., Environmentally critical elements in channel and cleaned samples of Illinois coals, *Fuel*, 77:95–107, 1998; J. A. Luppens, S. E. Wilson, and R. W. Stanton, *Manual on Drilling, Sampling, and Analysis of Coal*, ASTM Manual Series, MNL 11, 1992; J. G. Speight, *The Chemistry and Technology of Coal*, 1994; G. H. Taylor et al., *Organic Petrology*, 1998; UN/ECE, *International Classification of In. Seam Coals*, 1998; UN/ECE, *International Codification System for Medium and High Rank Coals*, 1998; D. W. van Krevelen, *Coal: Typology–Physics–Chemistry–Constitution*, 3d ed., 1993; C. R. Ward (ed.), *Coal Geology and Coal Technology*, 1984; World Energy Council, *1995 Survey of Energy Resources*, 17th ed., ed. by M. G. Schomberg.

Coal balls

Variously shaped nodules consisting of fossilized peat in which the individual cells and tissue systems of the plant parts are infiltrated by minerals, principally calcium carbonate, along with pyrite, dolomite, and occasionally silica. This type of fossilization, in which the cell walls are filled with minerals, is termed permineralization. Coal balls occur principally in Pennsylvanian (upper Carboniferous) bituminous and anthracite coals, often below black marine shales in North America, but permineralized plants have also been reported in coals from as early as the Devonian and extending well into the Paleocene.

How coal balls formed is not well understood, but some are believed to represent accumulations

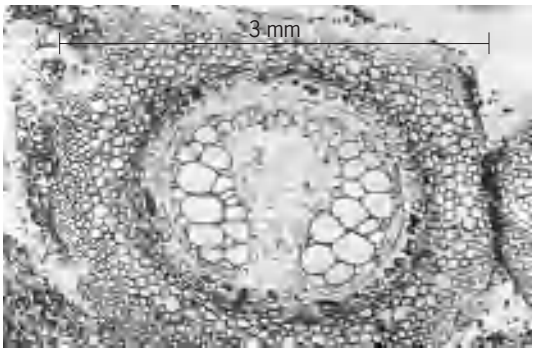

Photomicrograph of the peel made from a coal ball containing the petiole of a fern. 1 mm = 0.04 in.

of peat in which the plants were growing in low-lying swampy areas close to the sea. According to this model, seawater provided the high source of calcium carbonate in the permineralization processes. It has also been suggested that the permineralizing fluids were derived from percolating carbonate-rich ground water.

Because the individual plant cells in coal balls have not been crushed, they offer a wealth of information about the structure, morphology, and biology of the plants. Plants are studied by cutting a coal ball with a lapidary saw and then preparing an acetate peel of the surface. That is, the surface is etched with dilute acid that removes mineral material from the cell lumens, and then a thin sheet of acetate is added to the etched surface that has been flooded in acetone. The acetate-acetone mixture flows in and around the cell walls that are standing in relief. When the acetate sheet dries, a thin layer of plant cells can be peeled from the coal ball surface. After lightly grinding the coal ball surface with fine silicon powder, another peel can be prepared. Thus, the paleobotanist can prepare a set of serial sections of a fossil plant organ just as is done in the preparation of living plant and animal tissues. Peels can be observed with a dissecting microscope, or mounted on microscope slides under a cover glass to be examined with a compound microscope (see **illus.**).

Coal balls also contain the reproductive parts of plants. These can be observed on peels, or in the case of pollen grains and spores they can be macerated and studied with scanning and transmission electron microscopy. The preservation is sometimes so good that delicate structures such as nuclei, chromosomes, and membranes can be observed in great detail. *See* COAL PALEOBOTANY; PALEOBOTANY.

Thomas N. Taylor

Coal chemicals

For about 100 years, chemicals obtained as by-products in the primary processing of coal to metallurgical coke were the main source of aromatic compounds used as intermediates in the synthesis of dyes, drugs, antiseptics, and solvents. However, over 97% of the aromatic hydrocarbons, such as benzene, toluene, and xylenes, are now obtained largely from petroleum. Naphthalene, phenanthrene, heterocyclic hydrocarbons such as pyridines and quinolines, and some phenols are still obtained from coal tar.

Coke oven by-products, as a percentage of the coal used, are gas 18.5, light oil 1.0, and tar 3.5. Coke oven gas is a mixture of methane, carbon monoxide, hydrogen, small amounts of higher hydrocarbons, ammonia, and hydrogen sulfide. Most of the ammonia is recovered as about 20 lb (9 kg) of ammonium sulfate per ton (0.9 metric ton) of coal, but only about 10% of the hydrogen sulfide is recovered as elemental sulfur. Most coke oven gas is used as fuel. The composition of a typical coke oven crude light oil is shown in **Table 1**. The unidentified fractions contain very small amounts of a large number of hydrocarbons, and organic compounds containing oxygen and nitrogen.

Although several hundred chemical compounds have been isolated from coal tar, a relatively small number are present in appreciable amounts. These may be grouped as in **Table 2**.

All the compounds in Table 2 except the monomethylnaphthalenes are of some commercial importance. The amounts recovered and sold, however, are only 5–25% of the totals present in the coal tar.

The direct utilization of coal as a source of bulk organic chemicals has been the objective of much research and development. Oxidation of aqueous alkaline slurries of coal with oxygen under pressure yields a mixture of aromatic carboxylic acids. Because of the presence of nitrogen compounds and hydroxy acids, this mixture is difficult to refine. Hydrogenation of coal at elevated temperatures and pressures yields much larger amounts of tar acids and

TABLE 1. Analysis of a typical coke oven crude light oil

Component	% by vol
Forerunnings	
Cyclopentadiene	0.5
Carbon disulfide	0.5
Amylenes, unidentified substances	1.0
Crude benzol	
Benzene	57.0
Thiophene	0.2
Saturated nonaromatic hydrocarbons	3.0
Unsaturates, unidentified substances	3.0
Crude toluol	
Toluene	13.0
Saturated nonaromatic hydrocarbons	0.1
Unsaturates, unidentified substances	1.0
Crude light solvent	
Xylenes	5.0
Ethylbenzene	0.4
Styrene	0.8
Saturated nonaromatic hydrocarbons	0.3
Unsaturates, unidentified substances	1.0
Crude heavy solvent	
Coumarone, indene, dicyclopentadiene	5.0
Polyalkylbenzenes, hydrindene, etc.	4.0
Naphthalene	1.0
Unidentified heavy oils	1.0
Wash oil	
(Used to separate light oil from coke oven gas)	5.0

TABLE 2. Coal tar chemicals

Compound	Fraction of whole tar, %	Use
Naphthalene	10.9	Phthalic acid
Monomethylnaphthalenes	2.5	Organic syntheses
Acenaphthenes	1.4	Dye intermediates
Fluorene	1.6	Organic syntheses
Phenanthrene	4.0	Dyes, explosives
Anthracene	1.0	Dye intermediates
Carbazole (and other similar compounds)	2.3	Dye intermediates
Phenol	0.7	Plastics
Cresols and xylenols	1.5	Antiseptics, organic syntheses
Pyridine, picolines, lutidines, quinolines, acridine, and other tar bases	2.3	Drugs, dyes, antioxidants

aromatic hydrocarbons of commercial importance than are obtained by carbonization.

Synthetic fuels may be produced from coal by either direct or indirect methods. In direct liquefaction processes, pulverized coal is suspended in a solvent, the resultant slurry is heated and exposed to gaseous hydrogen under pressure, and the liquids produced are separated from the ash and distilled to obtain the fuel fractions. In the indirect process, the coal is decomposed thermally to yield combustible gases, which are synthesized to yield gasoline, alcohols, and waxes. *See* COAL LIQUEFACTION; COKE; DESTRUCTIVE DISTILLATION; ORGANIC CHEMISTRY; PYROLYSIS.

Howard W. Wainwright

Bibliography. L. L. Anderson and D. A. Tillman, *Synthetic Fuel from Coal*, 1979; B. D. Blaustein et al. (eds.), *New Approaches in Coal Chemistry*, 1981; K. R. Payne, *Chemicals from Coal: New Processes*, 1987; F. Probstein and R. Hicks (eds.), *Synthetic Fuels*, 1982, reprint 1990; J. C. Speight, *The Chemistry and Technology of Coal*, 1994.

Coal gasification

The conversion of coal or coal char to gaseous products by reaction with steam, oxygen, air, hydrogen, carbon dioxide, or a mixture of these. Products consist of carbon monoxide, carbon dioxide, hydrogen, methane, and some other gases in proportions dependent upon the specific reactants and conditions (temperatures and pressures) employed within the reactors, and the treatment steps which the gases undergo subsequent to leaving the gasifier. Similar chemistry can also be applied to the gasification of coke derived from petroleum and other sources. The reaction of coal or coal char with air or oxygen to produce heat and carbon dioxide could be called gasification, but it is more properly classified as combustion. The principal purposes of such conversion are the production of synthetic natural gas as a substitute gaseous fuel and synthesis gases for production of chemicals and plastics. *See* COMBUSTION.

Industrial uses. In all cases of commercial interest, gasification with steam, which is endothermic, is an important chemical reaction. The necessary heat input is typically supplied to the gasifier by combusting a portion of the coal with oxygen added along with the steam. From the industrial viewpoint, the final product is either chemical synthesis gas (CSG), medium-Btu gas (MBG), or a substitute natural gas (SNG). Heating values, compositions, and end uses for these gases are compared in the **table**.

Each of the gas types in the table has potential industrial applications. In the chemical industry, synthesis gas from coal is a potential alternative source of hydrogen and carbon monoxide. This mixture is obtained primarily from the steam reforming of natural gas, natural gas liquids, or other petroleum liquids. Fuel users in the industrial sector have studied the feasibility of using medium-Btu gas instead of natural gas or oil for fuel applications. Finally, the natural gas industry is interested in substitute natural gas, which can be distributed in existing pipeline networks. *See* NATURAL GAS.

There has also been some interest by the electric power industry in gasifying coal by using air to provide the necessary heat input. This could produce low-Btu gas (because of the nitrogen present), which can be burned in a combined-cycle power generation system. *See* ELECTRIC POWER GENERATION.

Gasification processes. In nearly all of the processes, the general process-flow diagram is the same (**Fig. 1**). Coal is prepared by crushing and drying, pretreated if necessary to prevent caking, and then gasified with a mixture of air or oxygen and steam. The resulting gas is cooled and cleaned of char fines, hydrogen sulfide, and CO_2 before entering optional processing steps to adjust its composition for the intended end use.

Thermodynamics. In discussions of the thermodynamics of coal gasification, at least one simplifying assumption is usually made, namely, that coal and coal char can be treated as pure carbon. Coal and coal char are really nonhomogeneous solids containing hydrogen, oxygen, sulfur, nitrogen, and mineral matter, but errors associated with this assumption are not likely to be significant.

The basic chemical reactions common to all coal gasification processes are coal and char reactions (1)–(3) and gaseous reactions (4) and (5). Reaction (2) is highly endothermic, requiring 32 kilocalories/mole (134 kilojoules/mole). The principal thermodynamic issue in most practical gasification processes is how to supply the heat for this reaction. Several methods have been considered, but generally

$$\text{Coal} \xrightarrow{\text{heat}} \text{gases } (CO, CO_2, H_2, CH_4) + \text{char} \quad (1)$$

$$C\ (\text{char}) + H_2O \rightarrow CO + H_2 \quad (\text{endothermic}) \quad (2)$$

$$2C\ (\text{char}) + {}^{3}/_{2}\ O_2 \rightarrow CO_2 + CO \quad (\text{exothermic}) \quad (3)$$

$$CO + H_2O \rightarrow H_2 + CO_2 \quad (\text{mildly exothermic}) \quad (4)$$

$$CO + 3H_2 \rightarrow CH_4 + H_2O \quad (\text{exothermic}) \quad (5)$$

Coal gasification products

Product	Higher heating value*	Major constituents	Use
Synthesis gas (CSG)	250–400 Btu/SCF† (9–15 MJ/m^3)	CO, H_2	Chemical feedstock for synthesis of H_2, NH_3, methanol, hydrocarbons, and so forth
Medium–Btu gas (MBG)	300–500 Btu/SCF (11–19 MJ/m^3)	CO, H_2, CH_4	General-purpose fuel for utilities and industries
Substitute natural gas (SNG)	900–1050 Btu/SCF	CH_4	Gaseous fuel

* Based on producing water as a liquid.
† SCF = Standard cubic feet.

heat is supplied by allowing exothermic reactions (3) and (5) to occur in the same vessel as reaction (2). The degree to which one reaction or the other is used depends on the process configuration and the desired product. Reactions (4) and (5) are often carried out in downstream processing blocks in order to tailor the gas composition for particular end uses. In these arrangements it is generally not feasible to return the heat to the gasifier, as the downstream reactions are carried out at much lower temperatures. The heat input methods used in the principal commercial and developing processes are given in the process descriptions below.

Processes. The differences in coal gasification processes principally involve the method of heat input, gasifier type, and ash recovery system. First-generation (commercial) gasifiers include the Lurgi (dry bottom), Koppers-Totzek (K-T), and Winkler gasifiers.

Lurgi. The Lurgi gasifier utilizes a bed of crushed coal moving downward through the reactor with countercurrent flow of gas, and operates at pressures up to 450 lb/in.^{2} gage (3100 kilopascals). Crushed coal, screened to remove fines, is fed to the top of the gasifier through a coal lockhopper, and passes downward through the drying pyrolysis, gasification, and combustion zones. Steam and oxygen are admitted through a revolving grate at the bottom of the gasifier which also removes the ash produced. After burning a portion of the coal to provide the heat required, the hot combustion gas passes upward through the zones of the coal bed. Steam is used as a reactant to gasify the coal, and is also added to keep the temperature below the ash melting point. Although the predominant source of methane (CH_4) is devolatilization, some methane comes from coal hydrogenation, releasing heat which is transferred to the coal and minimizing the oxygen requirements. As the coal moves down through the gasifier, the temperature initially rises slowly. Further down in the gasifier, it rises dramatically as the coal approaches the combustion zone where the temperature exceeds 2200°F (1200°C). The Lurgi reactors commonly used are of the general type shown in **Fig. 2**. Because of the countercurrent flow in the gasifier, high carbon utilization and good heat recovery are obtained.

The dry-bottom Lurgi process is generally applicable to subbituminous coals and lignites because the gasifiers operate best with noncaking, nonfriable coals with high reactivity and high ash melting points. The internal design and operating conditions of these reactors may be altered for coals of different caking characteristics and reactivities. Dry-bottom Lurgi gasifiers are used in the Great Plains coal gasification project in North Dakota and also in the world's largest gasification facilities at the SASOL coal-to-oil plants in South Africa.

Koppers-Totzek. The K-T gasifier is an atmospheric, entrained reactor. Dried and pulverized coal is mixed with steam and oxygen and fed to coaxial burners. The pulverized coal is gasified as it is carried in the flow of gas and as a portion of the coal is combusted. Practically every type of coal can be gasified by this technique. The operability of the process is not affected by caking properties of the coal.

The complete entrainment of the feed particles requires high gas velocities to achieve reaction times of only a few seconds; consequently the reaction temperature required is 3300–3500°F (1825–1925°C). At these temperatures, most of the carbon reacts very rapidly with oxygen and steam to produce carbon oxides, hydrogen, and molten slag. Less than 0.1% CH_4 is present in the product gas, and no tars or phenols are produced. More than 50% of the ash flows down the gasifier walls as molten slag into a

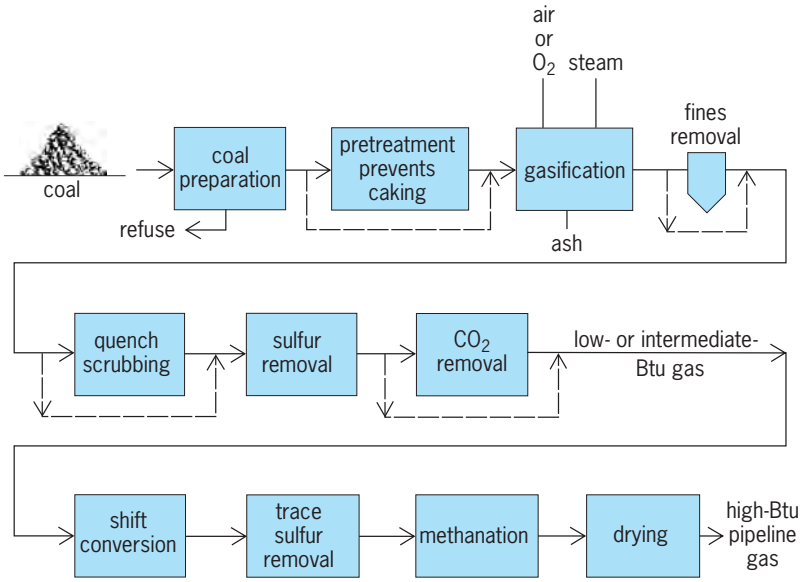

Fig. 1. Schematic representation of the processing steps in coal gasification.

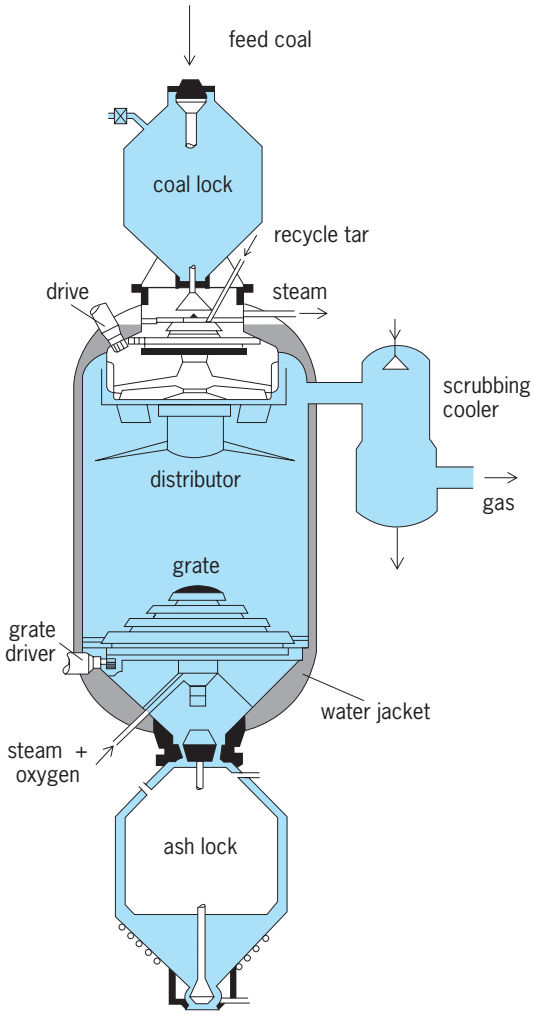

Fig. 2. Diagram of a Lurgi pressure gasifier.

slag quench tank. The rest leaves as entrained fly ash in the exit gas. The gasifier has a double-walled shell and a thin refractory lining. Water is circulated in the vessel jacket to protect the refractory. Most of the operating plants utilizing K-T gasifiers have been for the production of a H_2-rich gas for NH_3 manufacture.

Winkler. The Winkler gasifier is an atmospheric, fluidized-bed reactor that was developed in Germany in the 1920s. The vessel is cylindrical and refractory-lined, with nozzles for injecting a mixture of steam and oxygen at the bottom of the gasifier and above the top of the char bed to control fines. Crushed, dried coal is fed to the gasifier by means of a screw conveyor. The gasification temperature is controlled around 1800–2000°F (1000–1100°C) by adjusting the ratio of oxygen and steam to coal, and thus the portion of the coal being combusted. Gas is removed overhead, cooled, and dedusted. Product gas is low in CH_4 and contains no tars. Heavy char particles, which contain ash, settle to the bottom of the fluidized bed and are withdrawn from the gasifier by a screw conveyor. Most of the Winkler gasifiers that have been built are in eastern Europe and Asia.

Research and development. Most process research and development is directed toward increasing the efficiency of certain process steps, gasifying a wider range of coals or lowering plant costs. Important programs in the United States include work on partial oxidation, gasification with controlled ash agglomeration, and catalytic gasification. Major European efforts include work on the slagging gasifier and partial oxidation.

Slagging gasifier. The British Gas/Lurgi slagging gasifier is a modification of the commercial dry-bottom Lurgi design. It uses less steam and is less coal-sensitive than the dry-bottom Lurgi. These modifications are achieved by incorporating a high-temperature slagging zone at the bottom of a gasifier. Coal is screened and crushed before being introduced via lock-hoppers into the top of the reactor. As the coal moves downward through the reactor as a slowly moving bed, it passes through the drying, pyrolysis, gasification, and combustion zones. The coal ash is melted in the combustion zone and collects in a pool, from which it is tapped and quenched with water before being withdrawn from the gasifier. The oxygen and steam for the reaction are supplied through tuyeres in the bottom of the reactor. Tar, oil, and some of the coal fines produced in the crushing step can also be injected through the tuyeres and gasified. Some quantity of dusty tar can also be distributed at the top of the bed. The gas moving upward countercurrently through the moving bed of coal supplies the heat for the various gasification reactions and for drying the coal. The gas exits the reactor at approximately 1000°F (540°C), and is quenched in a scrubber to remove heavy tars and fines before being cooled in a waste heat boiler. This process is considered to be more attractive than the dry-bottom Lurgi for caking coals, such as bituminous coals of the eastern United States.

Partial oxidation. Both Texaco and Shell have pursued partial-oxidation process development. The Texaco gasifier is a pressurized (up to about 1000 lb/in.^{2} absolute or 6800 kPa), entrained reactor. Coal is pulverized in a wet mill and slurried (50–70 wt % solids) in water. The slurry is then pumped to pressure and injected with oxygen and steam into the downflow entrained gasifier. The hot gas and slag exit the bottom of the gasifier and pass through a radiant cooler. Most of the ash is collected at the bottom of the radiant section. The gas that contains some fly ash passes through a convective cooler and a scrubber. Both the radiant and convective coolers generate high-pressure, saturated steam. An alternate design involves cooling the hot gas from the reactor by water quench. In this case, no steam is generated, but the gas produced contains sufficient water for subsequent conversion of the carbon monoxide to hydrogen.

The Shell process is similar, but utilizes a dry feed system rather than a water slurry. In this process, ground coal is metered into an entrained gasifier where it reacts with oxygen and steam.

Gasification with controlled agglomeration. Both Westinghouse and the Institute of Gas Technology have pursued gasification under conditions which result in controlled ash agglomeration. The Westinghouse

gasifier is a pressurized (about 250 lb/in.^{2} absolute or 1600 kPa), fluidized-bed reactor. Coal is crushed before pressurizing via lockhoppers. The coal is injected with steam and oxygen into the fluid bed through an axial lift pipe. The higher temperature at the outlet of the lift pipe gives controlled ash agglomeration in the fluid bed. Larger particles of ash concentrate at the bottom of the gasifier and are removed and rejected. Steam and recycled gas are also injected in the lower portion of the reactor in order to fluidize the bed and control ash particle size. The resulting gas product is removed overhead, dedusted by cyclones, and cooled in waste heat boilers. The cooled gas is scrubbed to remove fines, and a portion is recycled to provide fluidization gas.

Catalytic coal gasification. The Exxon catalytic coal gasification process is intended for substitute natural gas production, and differs significantly from other gasification methods in that it uses a potassium catalyst to accomplish steam gasification [reaction (2)] and methanation [reaction (5)] in the same vessel. The catalyst enables the gasifier to operate at mild process conditions of 1300°F (700°C) and 500 lb/in.^{2} gage (3550 kPa). A major advantage of this "one-step" approach is that the overall reaction is almost thermally neutral. The gasification heat requirement is essentially supplied by the methanation heat release. Therefore, the catalytic coal gasifier does not require a large heat input by means of oxygen or other methods. The gasifier off-gas is treated by commercially available acid-gas–removal technology, and the methane product is cryogenically separated, with H_2 and CO recycled to the gasifier. The catalyst-char-ash material which is removed from the gasifier is washed countercurrently with water to recover most of the potassium catalyst for reuse. *See* HETEROGENEOUS CATALYSIS.

Underground gasification. Only about one-fifth of the total coal in the ground can be recovered by mining. The remainder is either too deep or in seams that are too thin to be mined. Long-term underground gasification is a promising possibility for utilizing this unminable coal. This process involves drilling wells into coal deposits, preparing a rubbled bed of coal through which gas can flow, and initiating gasification in the bed by using steam and air or oxygen. Such underground field tests have been conducted in the United States, but problems with underground gasification include process control, land subsidence, and ground-water pollution. *See* COAL MINING.

Several organizations have been active in this field, emphasizing work on the nature and direction of subsurface fracture systems, the means to calculate fluid movement underground, and environmental impacts. Directional drilling techniques and fracturing of the underground formations with chemical explosives have been studied. *See* COAL. W. R. Epperly

Bibliography. M. Elliott (ed.), *Chemistry of Coal Utilization*, 2d suppl. vol., 1981; J. L. Figuerido and J. A. Moulijin (eds.), *Carbon and Coal Gasification: Science and Technology*, 1986; R. L. Hirsch et al., Catalytic coal gasification: An emerging technology, *Science*, 215:121–127, January 8, 1982; Proceedings of the 1st, 2d, and 3d International Gas Research Conferences GRI-AGA-DOE-IGU, 1980, 1981, 1983; R. H. Schlosberg (ed.), *Chemistry of Coal Conversion*, 1985; M. K. Shcad and C. F. Hafke, Recent developments in coal gasification, *Chem. Eng. Prog.*, 79(5):45–51, May 1983.

Coal liquefaction

The conversion of most types of coal (with the exception of anthracite) primarily to petroleumlike hydrocarbon liquids which can be substituted for the standard liquid or solid fuels used to meet transportation, residential, commercial, and industrial fuel requirements. Coal liquids contain less sulfur, nitrogen, and ash, and are easier to transport and use than the parent (solid) coal. These liquids are suitable refinery feedstocks for the manufacture of gasoline, heating oil, diesel fuel, jet fuel, turbine fuel, fuel oil, and petrochemicals.

Liquefying coal involves increasing the ratio of hydrogen to carbon atoms (H:C) considerably—from about 0.8 to 1.5–2.0. This can be done in two ways: (1) indirectly, by first gasifying the coal to produce a synthesis gas (carbon monoxide and hydrogen) and then reconstructing liquid molecules by Fischer-Tropsch or methanol synthesis reactions; or (2) directly, by chemically adding hydrogen to the coal matrix under conditions of high pressure and temperature. In either case (with the exception of methanol synthesis), a wide range of products is obtained, from light hydrocarbon gases to heavy liquids. Even waxes, which are solid at room temperature, may be produced, depending on the specific conditions employed. **Table 1** lists potential uses for coal liquefaction products and indicates the additional processing which would likely be required.

Indirect liquefaction. Indirect coal liquefaction consists of three important steps: first, the coal is gasified; second, the composition of the resulting gas is adjusted (shifted, if necessary, to increase H_2 content; and H_2S and CO_2 are removed); and third, the CO and H_2 in the resulting synthesis gas are catalytically reacted to form the liquids. Commercial processes for catalytically reacting CO and H_2 to form the indirect coal liquids include Fischer-Tropsch synthesis and methanol synthesis. In addition, a Mobil process for conversion of methanol to gasoline is expected to be an important commercial operation.

Fischer-Tropsch synthesis. The synthesis processes based on the Fischer-Tropsch chemistry are represented by reaction (1). This reaction represents a

$$2nH_2 + nCO \rightarrow (\text{—}CH_2\text{—})_n + nH_2O \qquad (1)$$

polymeric addition of methylene (—CH_2—) groups to form a distribution of linear paraffins. In practice, a broad range of molecular weights is produced, and the product includes branched, olefinic, and oxygenated compounds, depending on reactor type,

TABLE 1. Potential uses for coal liquefaction products

Product	Commercial uses	Additional processing required
Light hydrocarbons		
Methane and ethane	Pipeline gas	None
C_3–C_4 hydrocarbons	Liquefied petroleum gas (LPG)	None
C_2–C_4 hydrocarbons	Ethylene and propylene manufacture	Steam cracking
Naphtha		
C_4–160°F (71°C) naphtha	Motor gasoline	Hydrotreating
160–350°F (71–180°C) naphtha	Motor gasoline	Hydrotreating and reforming
	Chemicals (benzene, toluene, xylene)	Extraction
350–430°F (180–220°C) naphtha	Motor gasoline	Hydrotreating
Middle distillate	Turbine fuel	Direct use or hydrotreating
350–650°F (180–340°C) or 430–650°F (220–340°C)	Low- or medium-speed diesel fuel	Direct use
	No. 6 fuel oil blendstock	Direct use
	Home heating oil	Direct use (with burner adjustment) or hydrotreating
	Auto diesel blendstock	Moderate hydrotreating
	Jet, auto diesel, No. 2 fuel oil	Severe hydrotreating or hydroconversion
Heavy distillate		
650–1000°F (340–540°C) vacuum	Fuel oil blendstock	Direct use or hydrotreating
Gas oil	Fluid catalytic cracking or hydroconversion feed	Hydrotreating
Other products		
Solvent refined coal (solid)	Boiler fuel	Direct use
Methanol	Motor fuel, distillate fuel	Direct use (with special distribution system)
Fischer-Tropsch liquids	Gasoline, diesel oil, jet fuel, wax	Hydrotreating, reforming, isomerization, polymerization

operating conditions, and catalyst type. The average molecular weight of the product mixture can be adjusted to some extent by the selection of process conditions.

Two synthesis processes (ARGE and Synthol) are used in the commercial coal liquefaction plants of the South African Coal, Oil, and Gas Company (Sasol) in the Republic of South Africa. The ARGE process utilizes a fixed-bed reactor and a precipitated-iron catalyst to produce a high-boiling, linear paraffin product, including a range of waxes. The Synthol process uses an entrained-bed reactor and a fused-iron catalyst to produce a somewhat lower-boiling hydrocarbon product which contains more isoparaffins and olefins.

The Sasol I commercial coal liquefaction plant started operations in 1955 and utilizes both the ARGE fixed-bed and the Synthol entrained-bed processes. Sasol II operation was initiated in 1980, when a site was also prepared for Sasol III. Both Sasol II and Sasol III are modified and enlarged versions of Sasol I. Each is designed to include seven of the Synthol entrained-bed reactors. Production of nonfuel products will be minimized in Sasol III.

Modified process techniques and conditions can have a significant impact on the Fischer-Tropsch product distribution. New catalyst formulations can improve activity and tailor the product distribution to meet fuel and chemical product demands. Engineering research to define better reactor designs and optimum operating conditions for more economic processes have also been undertaken. *See* FISCHER-TROPSCH PROCESS.

Methanol synthesis. The methanol synthesis processes are based on the chemistry represented by reactions (2) and (3).

$$CO + 2H_2 \rightarrow CH_3OH \quad (2)$$

$$CO_2 + 3H_2 \rightarrow CH_3OH + H_2O \quad (3)$$

Commercial processes use copper- and zinc-based catalysts and fixed-bed reactors operating at high pressure (100 atm or 10 megapascals) and low conversion per pass with high recycle ratios. Crude methanol can be utilized in three ways: as a neat fuel in fleet applications, as a gasoline blendstock, or as a feed to the Mobil methanol-to-gasoline process. *See* METHANOL.

Mobil methanol-to-gasoline process. In the Mobil process, a complex sequence of reactions occurs over a zeolite catalyst. Methanol is the first dehydrated to an equilibrium mixture of dimethyl ether, methanol, and water. This mixture is then fed to catalytic reactors where the remaining methanol is dehydrated to dimethyl ether, which reacts to form olefins that polymerize to form aromatic gasoline components. A fixed-bed version of this process is commercially available and has been applied in a New Zealand natural gas-to-gasoline project. A fluid-bed version of the methanol-to-gasoline process is used in a pilot plant in Germany, which started up in late 1982. *See* ZEOLITE.

TABLE 2. Leading direct coal liquefaction technologies*

	EDS	H-Coal	SCR II	Modified Bergius-Pier
Demonstration plant	250 tons/day (230 Mg/day), Baytown, Texas	200 tons/day (180 Mg/day), Catlettsburg, Kentucky	35 tons/day (32 Mg/day), Fort Lewis, Washington	220 tons/day (200 Mg/day), Bottrop, Germany
Recycle solvent	Hydrogenated distillate	Distillate	Product slurry	Distillate
Pressure, atm (MPa)	140–170 (14–17)	210 (21)	140 (14)	300 (30)
Bottoms recycle	Optional	No	Essential	No
Catalyst	Ni-Mo in solvent hydrotreater (none in liquefaction)	Co-Mo (in ebullating bed reactor)	None (relies on catalytic activity in recycled coal ash)	$FeSO_4$ as red mud
Types of coal	Bituminous Subbituminous Lignite	Bituminous Subbituminous	Bituminous	Bituminous

*Mg = megagrams.

Direct liquefaction. Direct liquefaction or hydrogenation of coal was developed in Germany in the 1920s by F. Bergius, who received the Nobel Prize in chemistry for his work. This technology was later modified by M. Pier and commercially proved during World War II. However, this technology operated at extremely high pressure (almost 700 atm or 70 MPa) and in small reactors (3-ft or 1-m diameter) and was too costly to be economically justified in peacetime.

Since the war, interest in direct coal liquefaction has periodically been revived, generally in response to projected shortages of petroleum, but no commercial plants have been built. However, four second-generation liquefaction technologies have been demonstrated in large pilot plants, and these are compared in **Table 2**. These technologies were an outgrowth of the Bergius-Pier process, but operate at much lower pressure (140–300 atm or 14–30 MPa). Other technologies have also benefited from major advances in reactor design since the 1940s, which will allow much larger (and therefore more economical) plants to be built. However, notwithstanding the technical improvements which have been made, the basic processing steps in coal liquefaction (see **illus.**) have not changed conceptually.

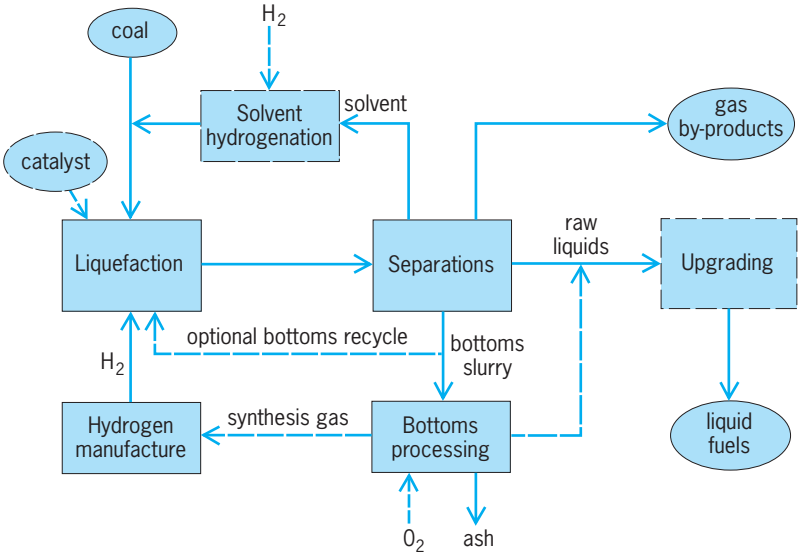

Processing steps in direct coal liquefaction. Broken lines represent optional steps or flow streams.

Coal is first crushed and slurried in a process-derived solvent and then heated and pumped to typical reaction conditions, in the range of 800–930°F (425–500°C) and 140–300 atm (14–30 MPa). Depending on the process, a catalyst may be added to the liquefaction step, or the solvent may be hydrogenated to improve hydrogen transfer to the coal. As the coal is heated to reaction temperatures and held there for 40–80 min nominal residence time, thermal cracking (thermolysis) occurs. At liquefaction temperature, chemical bonds are broken, that is, cracked, generating free-radical fragments which can participate in secondary reactions—combining with hydrogen to produce stable liquids, reacting further to form lighter liquids, or polymerizing to form high-molecular weight, high-boiling-point bottoms. Control of these secondary reactions is important in achieving high conversions and good selectivity to distillable liquid products. During liquefaction, heteroatoms in the coal (oxygen, nitrogen, and sulfur) are also liberated, principally as carbon oxides, water, ammonia, and hydrogen sulfide.

After liquefaction, products must be separated from bottoms (ash, heavy coal liquids, and unconverted coal) and solvent recovered. The most common technique for separation is distillation, although hydroclones, which separate by centrifugal force, have been used in Hydrocarbon Research's H-Coal and extraction in Gulf solvent-refined coal (SRC) for removal of ash and coal residues. Gulf's SRC II process, which evolved from SRC, recycles bottoms to increase the ash level in the liquifaction reactor and improve conversion. The EDS process has operated with and without bottoms recycling, and employs recycling to increase conversion and produce a lower-molecular-weight product slate. Product slate refers to the end products that are actually sent to the market from the plant. It has been shown that the bottoms contain high-boiling asphaltenes and preasphaltenes which can be readily converted to lighter products through recycling.

The net bottoms product is sent to a bottoms processing step, where the carbon is utilized and coal ash is separated. Carbon utilization can involve recovery of additional liquids by coking or extraction, partial oxidation to produce plant fuel or hydrogen, or combustion to provide plant heat. Hydrogen generation is frequently the preferred bottoms-processing step. In this case, bottoms would be sent to a partial oxidation unit and gasified with oxygen to produce a synthesis gas for hydrogen production. Alternatively, hydrogen could be produced by partial oxidation of coal or by conventional steam reforming of light gases as in present-day petroleum refineries and chemical plants. *See* COAL GASIFICATION; PETROLEUM PROCESSING AND REFINING; SOLVENT EXTRACTION.

Upgrading of coal liquids. In most cases, coal liquids require upgrading prior to use. The type and severity of upgrading depends on the initial characteristics of the liquids and the desired product properties (Table 1). In general, however, the upgrading processes are similar to those employed for petroleum liquids. Hydrotreating (mild hydrogenation) is commonly selected for removal of nitrogen, sulfur, or oxygen in coal liquids. Direct coal liquids are rich in cycloparaffins and aromatics. In the gasoline fraction, the cycloparaffins can be dehydrogenated by catalytic reforming to aromatics, which have a high octane rating. In the middle distillate fraction, the aromatics can be hydrogenated to cycloparaffins, which have improved performance in diesel and jet engines. In addition, catalytic cracking or hydroconversion of heavy distillates can be used to increase the yield of gasoline and middle distillate.

Indirect coal liquids are primarily paraffins and olefins. Polymerization of the propane-butane fraction can be used to increase the yield of gasoline. Reforming of the gasoline fraction increases octane rating by producing aromatics via cyclization of paraffins and subsequent dehydrogenation. Isomerization of linear paraffins also increases octane. Heavy, waxy streams can be cracked to gasoline and middle distillate. *See* HYDROCRACKING.

The severity of product upgrading typically decreases with increasing severity of direct liquefaction. The reason for this lies in the basic chemistry of converting high-carbon coal (H:C atomic ratio of about 0.8) into naphtha and light distillate (H:C of 1.2–1.9). Since the overall reaction involves hydrogen addition, it can be seen that a low level of hydrogenation in the liquefaction section would yield a high percentage of high-boiling components which require extensive hydrotreating to produce salable products. On the other hand, a liquefaction process of relatively high severity would yield liquids which require less extensive treatment to achieve the same salable products. As a result, yield comparisons of different coal liquefaction processes are usually closer on a finished product basis than on an intermediate product basis. It is generally agreed that a commercial coal liquefaction plant will produce approximately 3 barrels (0.5 m^3) of liquid product for every ton (0.9 metric ton) of dry, ash-free coal feed.

Continuing research. Research on improving coal liquefaction processes is concentrated in the general areas of staged liquefaction and catalysis. The principal goal of this work is to increase the yield of distillate products while reducing the amount of gas produced. In staged processes, the conditions in each stage are selected to optimize production of the desired liquid products. In catalytic processes, a catalyst which promotes formation of the desired molecular species is selected. Most emerging technologies incorporate elements of both staging and catalysis to obtain the advantages of each approach.

Staged liquefaction processes utilize two or more distinct stages to promote coal conversion or to improve the product slate. Though processes with many stages are theoretically possible, all of the staged liquefaction concepts use only two stages. Their essential features include a coal dissolver section (first stage), a deashing section to remove solids, and a catalytic hydroconversion section (second stage) to upgrade the heavy liquids. The dissolver section could range between a plug-flow reactor and a short-contact-time preheater. Deashing may be accomplished with hydroclones or by using proprietary antisolvent or critical-solvent deashing techniques. The hydroconversion section is typically operated at a lower temperature than the dissolver section (to prevent thermal cracking of liquids into gases) and contains catalyst in either a fixed bed, expanded bed, or ebullating bed.

In addition to these process differences, several variations in processing sequence are possible improvements to the two-stage technology. For example, the degree to which process streams from the hydroconversion section are recycled to the dissolver section may be adjusted to yield an integrated, partially integrated, or nonintegrated operation. Another option is placement of the second stage; since some hydroconversion processes are less sensitive to solids concentration, it may be possible to locate the deashing section downstream of both liquefaction stages. Finally, within the framework of any given processing sequence, the number of recycle streams, the cut points of these streams, and the severity of hydrotreating can also be changed.

The major two-stage liquefaction processes employ catalysts for additional conversion of heavy liquids. The two-stage process of the Electric Power Research Institute and the U.S. Department of Energy and the integrated two-stage liquefaction (ITSL) process of Lummus, both have high-temperature, plug-flow first stages and solvent-extraction deashing followed by lower-temperature, catalytic hydroconversion. In addition, ITSL uses a short-contact time preheater. The Chevron Coal Liquefaction process uses the same process sections but employs solids separation after both of the liquefaction stages. Overall, direct liquefaction is a fertile area for continuing research. *See* CATALYSIS; COAL; HETEROGENEOUS CATALYSIS.

W. R. Epperly

Bibliography. M. Crow et al., *Synthetic Fuel Development Technology in the United States: A Retrospective Assessment*, 1988; M. A. Elliott (ed.), *Chemistry of Coal Utilization*, 2d suppl. vol., 1981; R. F. Probstein, *Synthetic Fuels*, 1982, reprint 1990.

Coal mining

The technical and mechanical activities involved in removing coal from the earth and preparing it for market. Coal mining in the industrialized countries is characterized by the integration of a number of complex systems into a production methodology that varies for surface versus underground mining. *See* COAL.

The basic systems of the production methodology are the following. (1) Extraction systems: the methods and techniques used to break out or "win" the coal. (2) Materials-handling systems: the transport of coal and waste products away from the active production area, and the transport of the necessary materials, equipment, supplies, and workers to service the extraction system. (3) Ventilation: the development and operation of an air distribution system to provide the quantity, quality, and velocity of air where and when needed, to meet health and safety requirements. (4) Ground control: the control of the behavior of underground and surface openings developed by the extraction of coal. (5) Reclamation: the restoration of the mined area to its approximate original state or to an approved state.

To properly plan, design, and engineer a production system, knowledge of the geology of the deposit and the chemical and physical properties of the coal must be assembled and assessed. Basic information on the geology of the deposit is obtained from surface prospecting and mapping, and borehole drilling. This information is used to determine the size and shape of the coal area, the geologic column above and below all minable seams, the continuity and persistence of geologic features throughout the deposit, the presence of water or methane gas, and other special conditions. Proximate chemical analyses are made to determine coal characteristics, which affect its utilization. Tests are made to determine the cleaning, grinding, and handling properties of the coal. Ultimate chemical analyses are made to determine the fundamental chemical constituents of the coal. Maps are drawn to summarize this information, and are used for scheduling and sequencing production. *See* ANALYTICAL CHEMISTRY; ENGINEERING GEOLOGY; PROSPECTING; ROCK MECHANICS; SOIL MECHANICS; SPECTROSCOPY.

Surface Mining

When a coal seam lies sufficiently close to the ground surface, a surface mining method may be employed. The geographic location and geologic column—in particular, the ratio of the thickness and quality of the overburden and interseam waste to be removed, to the thickness and quality of the coal to be produced—are the initial considerations. Site conditions, which affect the production method,

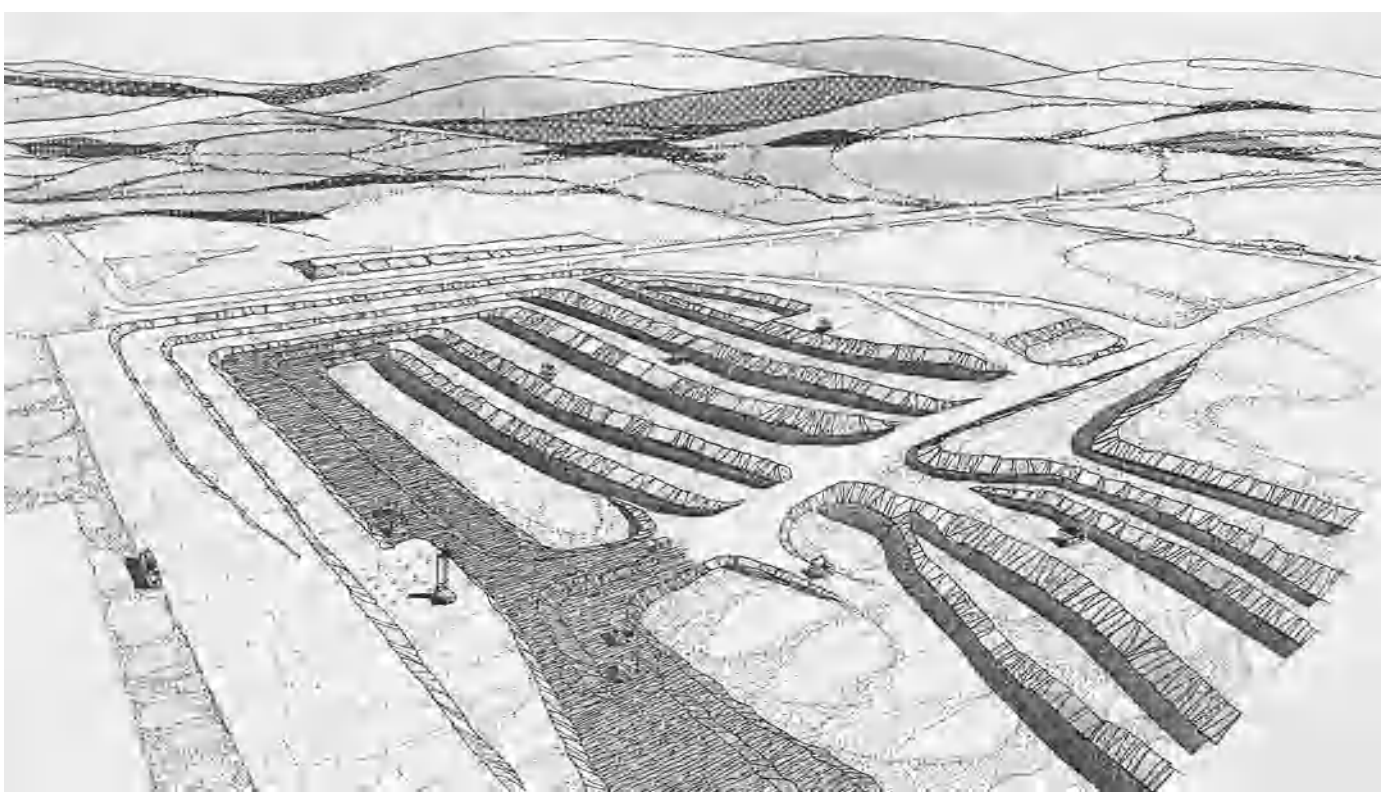

Fig. 1. Diagram of a terrace pit coal mine, showing terrace pit stripping and reclamation. (*After U.S. Department of Energy, The Development of Optimal Terrace Pit Coal Mine Systems, DOE Rep. FE/10023-1, January 1971*)

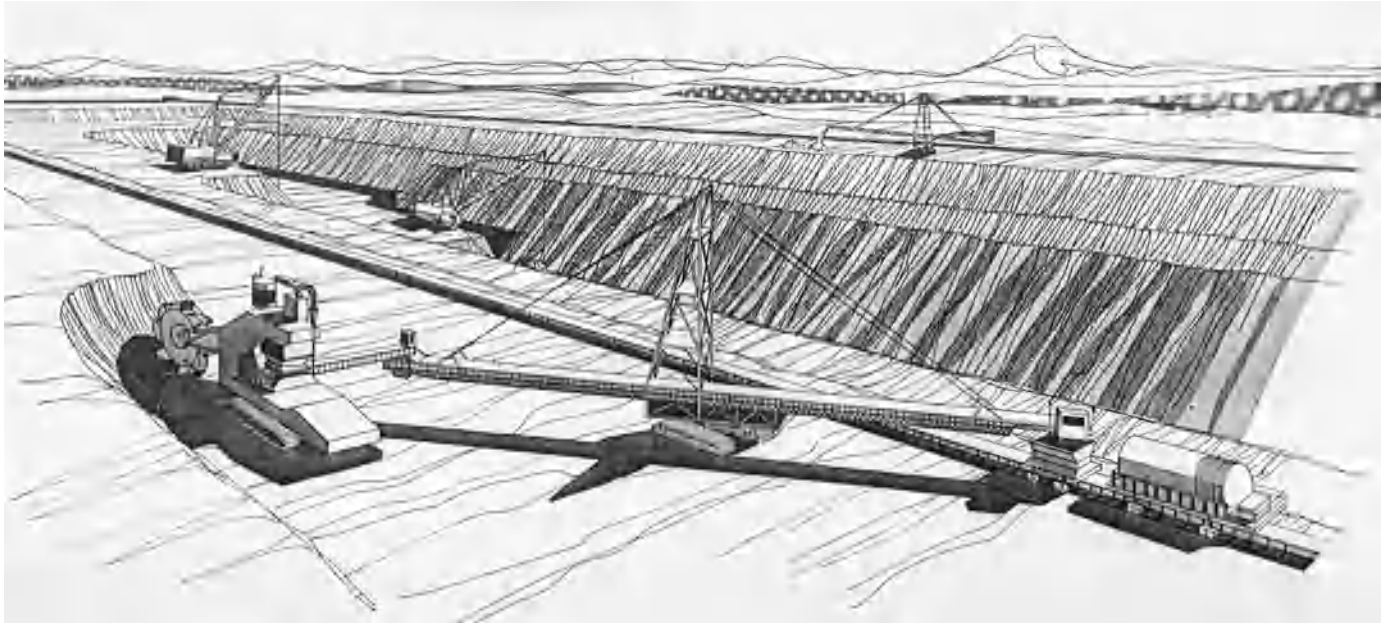

Fig. 2. Diagram of a strip coal mine, showing bucket-wheel excavator prestripping for draglines. (*After U.S. Department of Energy, The Development of Optimal Terrace Pit Coal Mine Systems, DOE Rep. FE/10023-1, January 1971*)

include types of terrain, temperature range, altitude, and rainfall and snowfall. Overburden and interseam waste characteristics affecting the production methods are material weight, stickiness, and swell; ability to form a stable highwall and spoil pile; ratio of rock and other material requiring blasting, to soil in the overburden; and need for special handling of waste materials for reclamation purposes. The presence of thicker and good-quality coal that commands higher prices can justify the removal of thick overburden.

Surface coal mines involve area mining or modified open pit mining, contour mining and mountaintop removal, and auger mining. Removal of overburden is called stripping, and hence the term "strip mining" is often applied to surface coal operations (**Figs. 1** and **2**). Area mining is applicable in relatively flat to gently undulating terrain where coal seams are of considerable area and may be at various dips. Contour mining and mountaintop removal are used in hilly and mountainous country and can be modified to handle coal seams at any dip. Auger mining follows the other surface methods when overburden removal becomes uneconomic, and is generally limited to more or less horizontal coal seams. *See* SURFACE MINING.

Area or modified open pit. In a typical area mine the layout is roughly rectangular with adjustments made to fit the boundaries of the property. The length of the pit varies depending on the rate of planned production, but may range to about 2 mi (3 km) or more. Extraction operations involve topsoil and subsoil removal, overburden removal, and coal loading. Drilling and blasting or ripping of the overburden and coal may be required.

The equipment used in topsoil and subsoil removal and in stripping overburden is selected based on the characteristics of the soil and rock to be removed. In soft digging, bulldozers, wheel loaders, and elevating scrapers may be used to excavate the soil and overburden and to load it into trucks (**Fig. 3**). In many parts of Europe, where the geologic column consists of soft-digging materials, bucket wheel excavators, used singly and in tandem, excavate and convey the soil and overburden to storage areas. Pan scrapers are found to be useful both in soil and overburden removal and in reclamation. When rock is present, drilling and blasting or ripping are required to loosen and break out the overburden. Draglines (**Fig. 4**) and power shovels (**Fig. 5**) are used to dig the harder and more consolidated materials.

Since outcrops rarely occur where area surface mining is practiced, mining begins with an opening or box cut, with the spoil from the cut being

Fig. 3. Front-end loader.

Fig. 4. Small dragline.

cast to one side or removed to a stockpile, and the uncovered coal is mined. Then an adjoining parallel cut of overburden is removed and deposited in the previous excavation, and the exposed coal is mined. Mining proceeds in this fashion until an economic, technologic, or property limit is reached.

The box cut is usually taken where the coal has the least overburden and in an orientation which allows the pit to advance toward coal with the greatest overburden. In a single-bench operation the box cut is designed to allow for a pit floor width which will accommodate the in-pit equipment and the initial overburden spoil used in backfilling, which must lie at a safe angle of repose. In a dragline or stripping shovel operation the toe of the spoil is adjacent to the coal seam in order to minimize the required operating radius for the equipment. Where multiple benches or terraces are developed for overburden and coal removal, the box cut width varies at each bench or terrace level.

In a single-bench, simple sidecasting dragline operation, there is no rehandling of material, and it is therefore a widely practiced and economical method for overburden excavation. Coal is mined by loading shovels or wheel loaders and is hauled by truck. The dragline may be operated so as to make a rough separation between soil and overburden and hence aid in reclamation of the mined area. Large power shovels are also used for stripping overburden in this type of operation.

Draglines have been used exclusively since the 1960s. The superiority of the dragline (**Fig. 6**) is its long reach and dumping, which permits thicker overburden to be mined and the spoil to be placed at a greater distance and in higher piles; and the large capacity of its buckets, which determines the productivity of the machine.

Multiple bench or terraced pits may be developed by using different combinations of equipment, but a shovel-truck combination or a shovel-truck-dragline combination is usually high on the list of feasible options. In a typical shovel-truck combination, overburden is removed with power shovels and loaded into trucks. A bulldozer works with each power shovel for cleanup, leveling, and bench preparation. The overburden benches are about 50 ft (15 m) in height and are mined in 200-ft (60-m) pushbacks. Shovels advance the pushback on the top level first and then proceed to the lower benches. Trucks haul the overburden up ramps on the highwall side of the

Fig. 5. Power shovel loading coal trucks.

Fig. 6. Dragline working from the surface of the ground. (*Morrison Knudsen Co.*)

pit, then around the pit on the spoil side, advancing the spoil dump at the same rate as the highwall. Spoil from the box cut is stockpiled for later use in reclamation.

Overburden drilling is done by rotary drills with bits of 6.78–12.12-in. (17.5–31.8-cm) diameter. Coal drilling is done with truck-mounted rotary drills using 6-in.-diameter (15-cm) bits. Blastholes are loaded with ANFO (ammonium nitrate–fuel oil) explosive, usually from bulk mixing trucks, which mix and meter the amount of explosive for each hole. After blasting, coal is loaded, with either power shovels or wheel loaders, onto bottom dump trucks of 100–150-ton (90–135-metric ton) capacity which are favored for coal hauling because of their low body height. Coal benches are about 30 ft (9 m) high and 200 ft (60 m) wide, and are advanced in 100-ft (30-m) pushbacks. *See* DRILLING AND BORING, GEOTECHNICAL.

A dragline may be added to the shovel-truck combination after the box cut is excavated and backfilling with spoil has begun. A bench is excavated for the dragline at the appropriate level above the seam. Operating from this bench, the dragline excavates the overburden above the seam and sidecasts it into the pit floor area from which coal has been extracted.

A number of variations on the shovel-truck combination are possible. For example, an around-the-pit conveyor transport system may be used to reduce or eliminate truck haulage. Excavated overburden may be loaded directly or may be transported by truck to a hopper equipped with a fixed grizzly screen to reject oversize material. From the hopper, the spoil is fed to a shiftable conveyor system, which carries the spoil out of the pit to a stacker-spreader. This in turn spreads the material along a spoil bench in the pit. The conveyor system is shifted forward in 100-ft (30-m) increments as the benches advance.

Bucket wheel excavators are useful in certain mining conditions. Essentially a bucket wheel excavator consists of (1) a large wheel with toothed buckets pinned to the circumference and mounted on the end of a maneuverable boom; (2) a discharge boom supporting a belt conveyor; (3) a revolving frame mounted on crawler supports; and (4) electric motors and control devices. Material is excavated by moving the rotating wheel into the overburden or coal and swinging it on the maneuverable boom through an arc. Material is cut, chopped, or scraped into the buckets, and as the wheel continues to rotate, the buckets are emptied. The excavated material is transported away from the wheel by a system of conveyors, and finally moves up the discharge conveyor to the discharge point. The conveyor system may consist of an around-the-pit conveyor incorporating a shiftable conveyor section. The shiftable conveyor must be periodically moved to keep it within the discharge range of the bucket wheel excavator. A mobile transfer conveyor may be added to the system to improve the continuity of material flow. In all cases, however, material flow is more or less continuous.

Bucket wheel excavators are capable of high production under favorable conditions with outputs in the range of 1600–12,000 yd^3 (1200–9200 m^3) per hour. However, they have difficulty in handling hard, consolidated material even when blasted; alternating the hard and soft formations slows production. Also, large boulders, buried vegetation, and sticky materials clog the buckets, transfer points, and machinery. However, bucket wheel excavators are widely used in Europe for the mining of brown coals. The rock strata, consisting of unconsolidated earth, soft shale, and siltstone overlying thick beds of brown coal, is ideally suited to the capabilities of bucket wheel excavators. The brown coals extend over large areas and may be found in thicknesses of 600 ft (180 m) or more, interbedded with sand and clay zones. Area pit dimensions may be a number of miles, and coal and spoil transport to delivery points is by belt conveyor or rail.

Bucket wheel excavators have found limited application in the United States. *See* OPEN PIT MINING.

Contour mining and mountaintop removal. This method of surface mining is used in the hilly and mountainous areas of the Appalachians, where coal seams are numerous and are generally flat-lying or subhorizontal, outcropping on the steeply sloping valleys and ridges. The geologic column consists of shales, sandstones, and limestones of varying thickness and hardness.

Now illegal, conventional contour mining started with an initial cut along the outcrop line on the mountainside. The overburden was indiscriminately cast over the side of the hill and the coal extracted. This procedure was continued into the mountainside and along the outcrop line until an economic, technical, or property limit was reached. The resulting

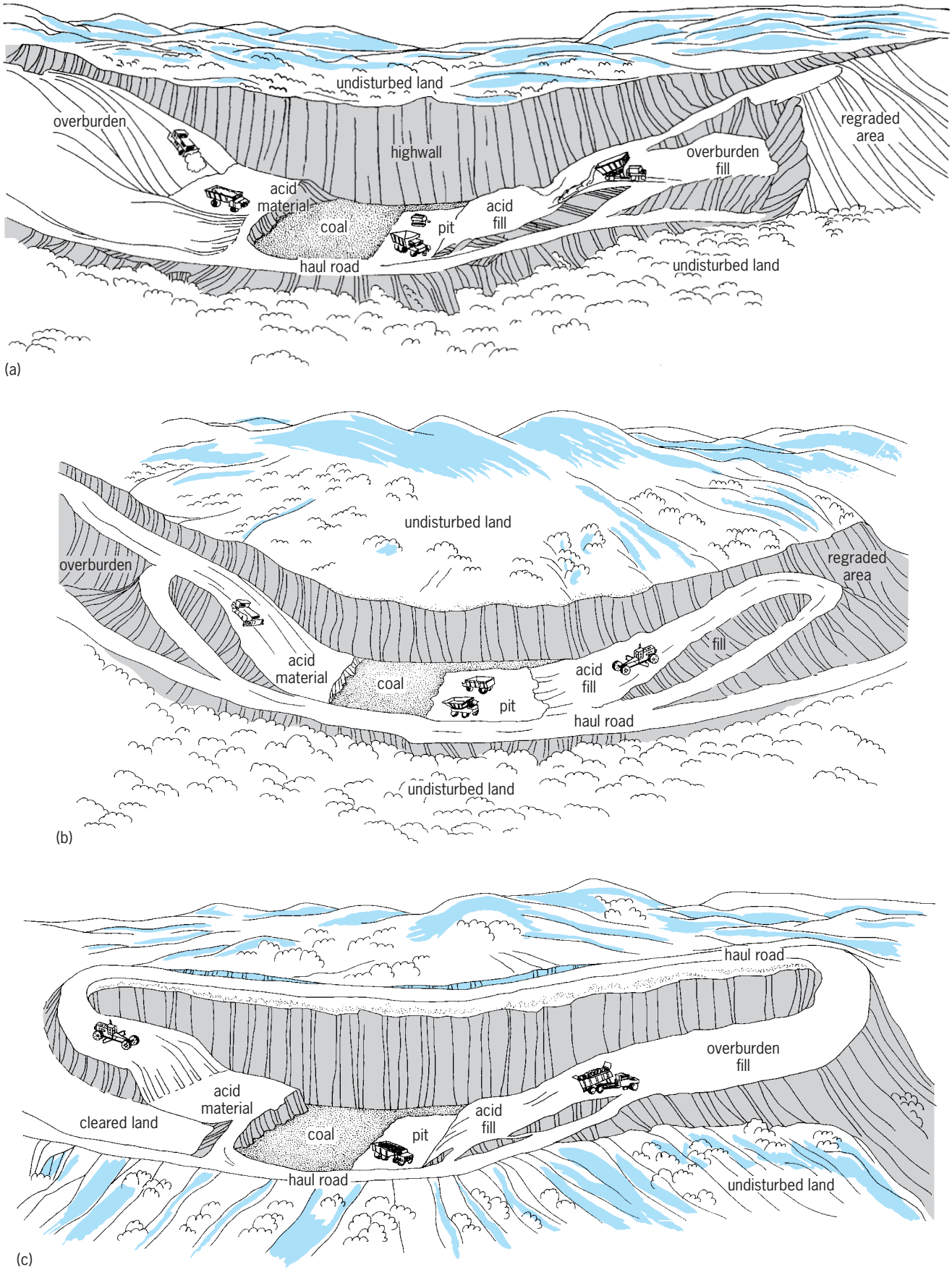

Fig. 7. Typical haulback or lateral movement methods. (*a*) Truck. (*b*) Scraper. (*c*) Truck and scraper. (*After R. Stefanko, Coal Mining Technology, Theory and Practice, Society of Mining Engineers, AIME, 1983*)

excavation left a collar around the mountain consisting of a highwall, bench, and castover spoil bank.

The haulback method of contour mining is environmentally sound because it eliminates the casting-over of spoil and provides for concurrent reclamation (**Fig. 7**). The initial cut is made into the mountainside to the highwall limit, and the excavated overburden is hauled to a prepared storage area. The initial cut serves the same purpose as the box cut in area mining. Mining proceeds along the outcrop line, with overburden, subsoil, and topsoil being trucked and placed as coal is removed. The method leads to minimum disturbance of the land and allows for contemporaneous reclamation.

In the haulback method the layout of haul roads and the design and construction of the spoil storage area are critical. Haul roads are designed and situated so as to assist in the control of water flow from the pit and the spillage of spoil on the downslope. Spoil storage areas are generally located in hollow valleys or low areas, and provisions must be made for prevention of contamination of streams by the spoil material and for stability of the material.

The equipment used in the haulback method is small in comparison to that found in an area mine. Wheel loaders, trucks, dozers, and scrapers are used in various combinations, and sometimes power shovels are employed. Drilling and blasting are done with rotary rigs and ANFO. Overburden is drilled and blasted, then wheel-loaded into trucks for haulage to the mined-out pit. Subsoil and topsoil that have been removed by dozers are then placed on top of the overburden, and the spoil materials are graded, seeded, and revegetated.

Mountaintop removal might be considered as an extension of area mining techniques to the mountainous and hilly terrains. A box cut is made across the ridge or mountain, and the spoil from this cut is placed in hollow or valley spoil-storage areas. The pit is advanced from the box cut by excavating parallel slices, with the spoil from the new cut being placed in the adjacent excavated cut. If multiple seams are to be extracted, a bench or terrace pit is developed. The equipment is similar to that in contour mining. The result of mountaintop removal is the transformation of a rugged topography into more level land, but in the process total coal recovery is obtained. Future use of this method will depend upon its perceived advantages in comparison with other techniques.

Auger mining. This method consists of boring large-diameter holes as far as possible into more or less horizontal coal beds, with hole diameter and spacing depending on seam thickness and geologic conditions. The level bench left from a conventional contour mine provides an ideal platform for equipment operation; otherwise a bench must be developed along the coal outcrop line. The auger consists of a cutting head and a set of twisted flights, resembling a carpenter's wood bit or drill. It can be extended by additional flight sections: the drive unit is uncoupled and withdrawn, allowing another flight to be lowered into place for making another advance. As boring proceeds, the coal is moved along the flights and discharged onto an elevating conveyor and then into trucks. Auger mining is attractive because it recovers coal from the highwall side of the contour mine that might otherwise be lost,

Fig. 8. Highwall miner. (*Superior Highwall Miners*)

and it does this with minimal capital and labor cost. Reclamation and environmental problems can significantly affect the cost of auger mining.

A variation in auger mining, called highwall mining, is a system that can be adapted for varying coal seam thickness (**Fig. 8**). Instead of the typical auger cutting head, a continuous miner or a similar arrangement cuts the coal and loads it onto a conveyor. The mined coal is loaded directly on a truck behind the highwall miner, much the same as in the auger mining system.

Future trends. Surface mining has steadily increased its share of the United States market, and this trend is expected to continue into the foreseeable future. The fundamental extraction and materials-handling systems that have evolved present a range of methods that appear capable of coping with future mining problems. Coal lying under deeper and more difficult geologic columns may require some modification in extraction techniques and methods, but it does not appear that a new generation of large equipment needs to be developed. Emphasis will be placed on better planning and managing of the existing methodologies and automation of cyclic operations, where feasible, to obtain highly efficient mines. Moreover, the development of highly productive and efficient underground mines will provide an economic limit to surface mines. *See* LAND RECLAMATION.

Underground Mining

When a coal seam does not lie close to the surface, it must be extracted by underground methods. The methods may be classified as room-and-pillar, longwall, and others. In each method, modifications to the basic techniques are needed to cope with varying geologic conditions and seam factors. Irregular seam thickness, steep dips, changing rock quality, seam partings, and other factors have a marked influence on the mine geometry and equipment specifications. Seams dipping between 0 and 10° and with regular seam thickness are assumed in the following discussion of room-and-pillar and longwall mining.

Mine development. In developing a particular extraction system, development openings and production openings must be driven. Development openings provide the primary access to various parts of the coal deposit and are called mains. In room-and-pillar mining, the production openings are the rooms driven in the panel and the extraction cuts made during pillar retrieval. In longwall mining, the production openings are the longwall faces. In general, the term "development" includes all openings and other work which precedes production.

Main access to a horizontal or near-horizontal coal seam is usually by drift entry when the coal seam outcrops on a mountainside; otherwise, slope and shaft entries are used. Slopes are preferred over shafts when the seam is not too deep because conveyor belt haulage can be used. In addition, mining machinery can usually be brought underground with less disassembling. From the foot of these access openings, a set of parallel main entries (usually 7–10 individual headings) are driven, and when these have advanced 3000–4000 ft (915–1220 m) a set of submain entries are driven to the right and left of the main entries. This entire set of entries divides the coal property into large blocks, which are further divided by panel entries driven off the submains.

Because of ventilation restrictions, mining machinery movements, and ground conditions, connections between entries are made at intervals. These are called crosscuts. The crosscuts divide the coal block between entries into pillars, often called chain pillars, which provide the primary support for the ground overlying the coal deposit. Since main and submain entries must remain open and stable for long periods of time, the design of the pillars and the roof support in the entries is a critical item in mine planning.

Panel entries, up to 17,000 ft (5180 m) or longer, divide the large coal blocks into rectangular blocks (panels) for production mining. Barrier pillars are left between the mains, the submains, and the rectangular panels in order to minimize the effect of the coal extraction associated with production on the stability of the mains.

Room-and-pillar mining. In the room-and-pillar system (**Fig. 9**) the mining methods may be by either continuous or conventional. The essential differences between the two methods are in the equipment spectrum and the cycle of face operations that the equipment dictates.

In continuous mining, the cycle of operations begins with a continuous miner sumping into the coal face. A shuttle car is positioned behind the miner to receive and transport the cut coal to the belt feeder. When the shuttle car is filled, it moves out and is replaced by the next shuttle car. This operation continues until the continuous miner operator reaches the last roof support. The continuous miner is then withdrawn, and safety posts or hydraulic jacks are placed in the excavated opening. The miner then widens the room to its final size (usually 18–20 ft or 5.5–6 m) by cutting the adjacent strip. The miner moves onto the next working place—a room or a crosscut—and roof-bolt support is installed. Continuous miners equipped with roof-bolting capabilities eliminate this intermediate step and continue to widen the room without the installation of hydraulic jacks. The final roof-bolting pattern is completed after the continuous miner moves to the next room. The face ventilation system is adjusted to compensate for the freshly cut room and the new working location of the continuous miner. Electric cables to the equipment, the power center, the belt feeder, and other items must be adjusted and moved forward as the overall mining section advances.

In conventional mining, the breaking out of the coal from the face is done by undercutting, drilling, and blasting operations. First, the cutting machine saws a 10–12-ft-deep (3–4-m) slice horizontally across the room width at the bottom and then moves out to the next place to be cut. Then the drilling machine (**Fig. 10**) moves in and drills holes into the

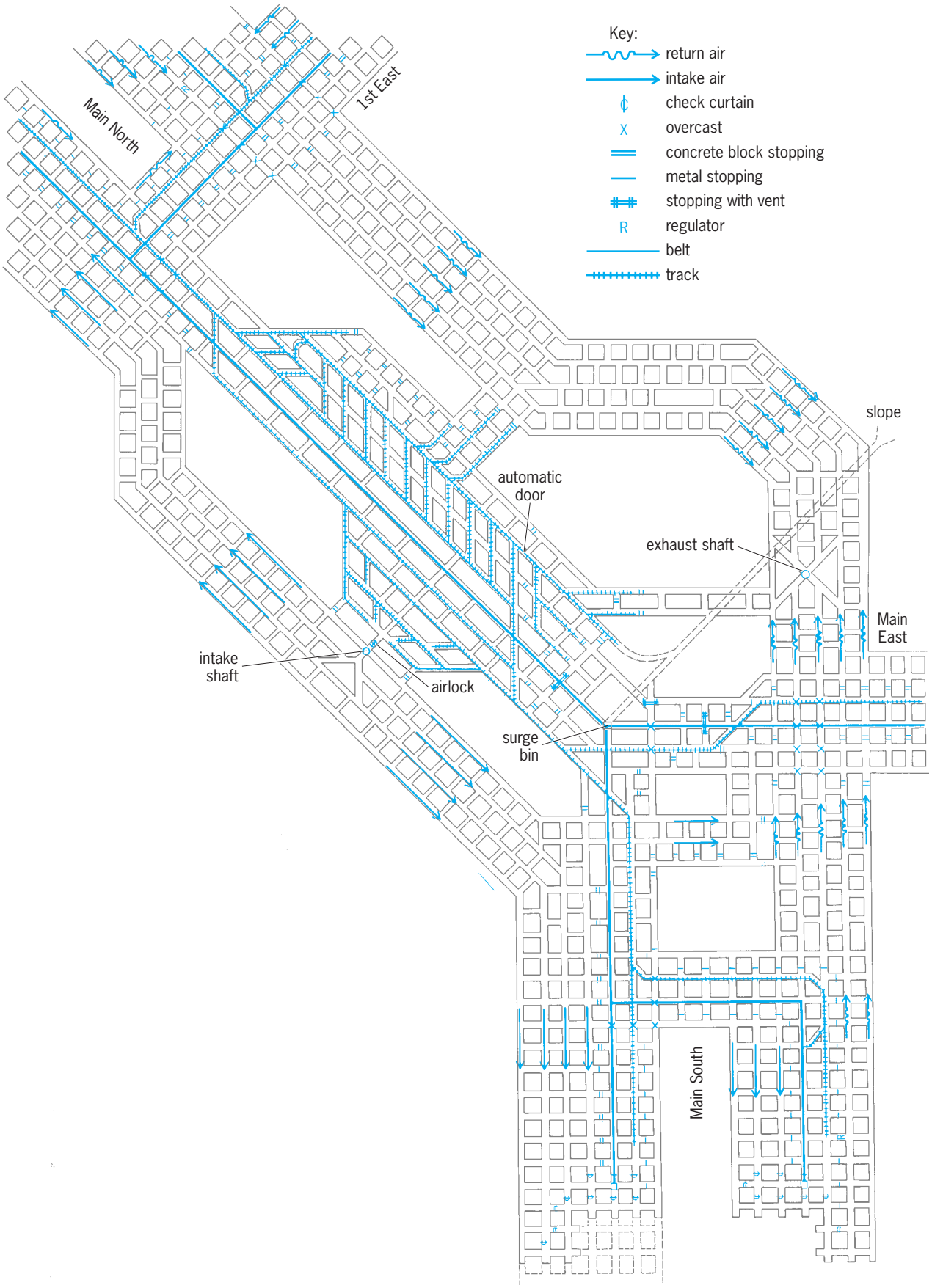

Fig. 9. Plan of a room-and-pillar mining operation. (*After R. Stefanko, Coal Mining Technology, Theory and Practice, Society of Mining Engineers, AIME, 1983*)

Fig. 10. Single-boom face drill operating in low-seam coal. (*Long-Airdox Co.*)

cut face, and this is followed by the blasting operation. After the area is examined for safety, the loading machine and shuttle cars move in and load out the coal. Sufficient working places should be available to minimize the idle time of equipment. Ventilation, electric cables, power center, belt feeder, and other items must be adjusted and moved forward as the overall mining section advances.

Retreat mining. The system of rooms and crosscuts driven in the production panels divides the panel into a series of coal pillars of approximately equal size. These coal pillars are extracted by methods that allow the mining operation to retreat toward the panel entries. As the pillars continue to be extracted, the bridging span of the immediate roof and then of the main roof is exceeded and the ground caves into the opening. This caved rock is called the gob. Since methane gas may accumulate in the gob, a bleeder system of entries is established at the top of the panel to allow a portion of the ventilating air to circulate through the gob.

Fig. 11. Remote-controlled continuous miner.

There are four basic pillar extraction techniques: split-and-fender, pocket-and-wing, outside lift, and open end. The first two are suited for continuous mining methods. Open end is utilized with conventional mining equipment, and outside lift is primarily used to extract long, thin pillars.

In the split-and-fender method, the pillar is split into two fenders of coal by driving a room parallel to the long side of the pillar. Because of roof control restrictions, which do not allow the splitting of the pillar by a single pass of the continuous miner, two pillars are mined simultaneously. The continuous mining unit, operating in an alternating fashion, mines a cut in each pillar. Rock bolts are used for roof support in the splitting operation, but breaker posts and cribs are used for roof control when the fenders of coal are removed. Check curtains are used to direct the ventilating current to the working faces.

The cuts used in the outside lift method are similar to those used in removing a fender of coal.

In the pocket-and-wing method, two working places can be sequenced in the same pillar. The pillar is alternately split transversely, then longitudinally, producing three fenders. These are then removed by using breaker posts for roof control. Ventilation is controlled by check curtains.

In the open-end method of pillar extraction, the sequence of extraction cuts is made adjacent to the gob without a protecting fender of coal.

Equipment. Continuous miners (**Fig. 11**) consist of a cutting head, a coal-gathering system, and a scraper conveyor system; the whole piece of equipment is electrically powered and mounted on a crawler-supported carrier. The cutting heads of various designs rip, bore, or auger the coal from the face. The ripper type of miner is most popular. Rippers have drum cutting heads mounted horizontally and fitted with steel bits. The cutting head, which is supported by arms that allow it to be positioned in a vertical plane, is crowded into the face by moving the carrier forward. Coal is ripped from the face, drops to the floor, and is gathered onto the scraper conveyor by a set of rotating disks or gathering

Fig. 12. Shuttle car for transporting coal within the mine.

Fig. 13. Diesel scoop in operation. (*WVA Mining Equipment Co.*)

arms. The scraper conveyor transports the coal into the positioned shuttle car. Geologic conditions, especially the seam height and the presence of in-seam rock, influence the operation and design of the miner.

A shuttle car is a low-bodied diesel or electrically powered unit (**Figs. 12** and **13**) that carries the coal to a nearby belt conveyor system that takes the coal out of the mine.

Roof bolts are an essential component of both continuous and conventional mining. They are used to provide ground control for the immediate roof strata overlying the mined room, and a means of support that does not obstruct the mine opening. Roof bolters are used to drill the holes and install the bolts. Roof bolting is a prime bottleneck in room-and-pillar mining and a dangerous operation; hence roof-bolting equipment must be fast and efficient and provide temporary support to the mine roof while the operator is installing the bolt.

A cutting machine is a large chain saw mounted on a rubber-tired carrier. A universal cutter can be maneuvered to cut a slice at any angle to the horizontal. Blasthole drilling in coal is done with electrically driven augers. Blasting may be done with special coal mine explosives, compressed air, or carbon dioxide cartridges. The loading machine consists of a gathering system, which sweeps the coal off the mine floor onto an inclined scraper conveyor for transport to the shuttle car. A belt feeder has a toothed roll crusher for sizing the coal, which is then fed at a uniform rate by a scraper conveyor onto the section belt conveyor.

Ventilation. The primary function of mine ventilation is to dilute, render harmless, and carry away the gas and dust produced in the underground mining of coal. Federal law requires the following conditions to be met: (1) Minimum air velocity in the working place must be at least 60 ft (18 m) per minute. (2) Minimum air volume at the last open crosscut must by 9000 ft^3 (255 m^3) per minute. (3) Minimum air volume reaching each working face must be 3000 ft^3 (85 m^3) per minute. (4) Respirable dust levels in the mine air cannot exceed an average of 2 mg/m^3. (5) Air used to ventilate belt conveyors cannot be used to ventilate working faces and must be vented to return airways. (6) Methane accumulations cannot be allowed to exceed 1% in the working area.

The mine ventilation system is composed of three subsystems: (1) the primary distribution system, which consists of main fans and the main intake and return entries; (2) the face ventilation system (**Fig. 14**), which controls airflow throughout the active production area; and (3) control devices to direct the air through the mine. Main fans are required to be installed on the surface, and usually pull air through the mine. They do this by developing a pressure gradient between the intake air entries and the return air entries in the mine. Hence, crosscuts between these two sets of main entries must be closed off by permanent stoppings. Splitting the intake air at each set of submains allows different areas of the mine to have their own independent air, and also serves to minimize power costs. Return air from these splits must be carried over the intake mains by air bridges or overcasts. The flow of air within a split can be controlled by increasing the resistance to flow with regulators placed at strategic locations in various entries and rooms.

The production area of the mine generates the most gas and dust, and hence requires constant

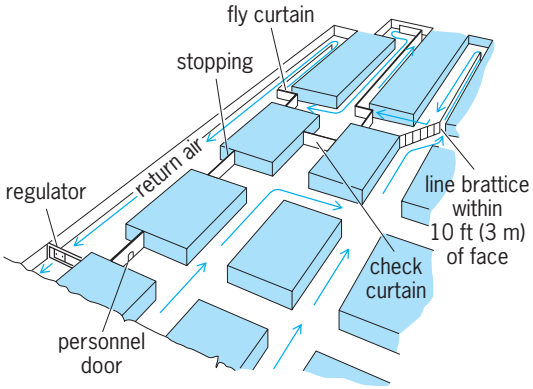

Fig. 14. Diagram of a blower line-brattice face ventilation. Arrows indicate direction of flow of air. (*After R. Stefanko, Coal Mining Technology, Theory and Practice, Society of Mining Engineers, AIME, 1983*)

examination and modification of the airflow pattern to cope with the changing situation developed by the advancing faces. The blind room headings ahead of the last open crosscut may be ventilated by line brattices (nylon-plastic curtains which are installed so as to split the blind room heading longitudinally) or auxiliary fans. Air is then directed up one side and down the other. The line brattices are also used to direct the movement of air throughout the production area by acting as check curtains in certain rooms. In order to do an effective job of diluting and sweeping dust and gas from the working face, line brattices must be kept close to the face. Auxiliary fans and tubing in a push-pull arrangement are alternatives to line brattices, and better diffuse air across the working face but introduce other problems.

Ground control. The rock strata above the coal seam may be divided into an immediate roof and a main roof. The immediate roof responds quickly to the removal of the coal support and collapses unless some artificial support is installed to prevent and control the deformation. The main roof shows stability for long periods of time. In general, the boundary between these two roofs is formed at the contact zone defined by fairly thick competent rocks, such as sandstone and limestone, and thinly bedded incompetent rocks, such as shale and coal. The function of the pillars of coal developed by driving rooms and crosscuts is to provide support for the total roof structure over both the pillars and the mined openings. The ground load immediately over the pillars is transferred directly to the pillars, but the ground load over the rooms and crosscuts must transfer its load to the pillars. This mechanism of load transfer is called the bridging action of the rocks; competent rocks have good bridging action and incompetent rocks have poor bridging action. Artificial support is used to assist, enhance, or develop the bridging action of the rocks in the immediate roof.

Simple posts with cap pieces were the first roof supports used. The concept is to decrease the overall bridging span between posts. More posts with larger cap pieces and three-piece sets are added to control less competent roof strata. Posts for roof control are still used in pillar recovery work or retreat mining. However, for the highly mechanized continuous and conventional mining of entries, rooms, and crosscuts, timber posts and three-piece sets cannot be used because they restrict the movement of equipment and may be knocked down.

Prop-free openings with the immediate roof spanning the full entry width can be developed with the aid of roof bolts. The most commonly used roof bolt is an expansion-shell-anchored, 5/8-in.-diameter (1.6-cm), extrastrength, square-head steel bolt. The law specifies that the bolt must be a minimum of 30 in. (76 cm) in length. The function of the bolt is to bind together the various laminae in the immediate roof and make them act as a unit in developing bridging action. The bolts, when anchored in competent rock, also prevent sagging of the immediate roof. The binding and suspension action of the bolt is developed when tension is produced in the bolt by tightening it into the anchoring shell. A number of variations of this basic bolt are in use.

Ground control in main entries or in areas where roof bolts cannot develop the desired support is done by other means. Room and entry intersections often exceed mandated limits on roof bolts, so steel sets or steel beams resting on wood cribs or concrete blocks are used for the support of the intersections. In some main entries, steel arches or yielding steel arches are used. Where rapid temporary support is required, screw jacks or hydraulic props are easy to install and make excellent supports.

In the United States, federal law mandates that each mine operator have an approved roof control plan, a history of all unintentional roof falls, and a systematic evaluation of the roof control system in use. The roof control plan must show the geologic column, and identify the rock type and thickness above and below the coal seam. Further, the plan must describe the sequence of mining and installation of supports, including temporary supports, and must designate the area of the mine covered by the roof control plan.

Longwall mining. Longwall mining (**Fig. 15**) is the basic underground coal mining method practiced in Europe, and it enjoys wide acceptance throughout the world. It has a greater production potential than room-and-pillar systems, and is safer since mining takes place beneath a complete overhead steel canopy (**Fig. 16**) that moves as the face is mined. Longwall methods were tried in the United States for a number of years, but could not compete with the room-and-pillar system until modern longwall mining was introduced in 1960. Since that time, longwall mining has gained wide acceptance by coal producers in the United States. More coal is being produced now by longwall mining than any other underground mining method in the United States. Longwall systems are classified as retreat or advancing systems based on whether the panel is completely developed by panel entries before panel extraction commences.

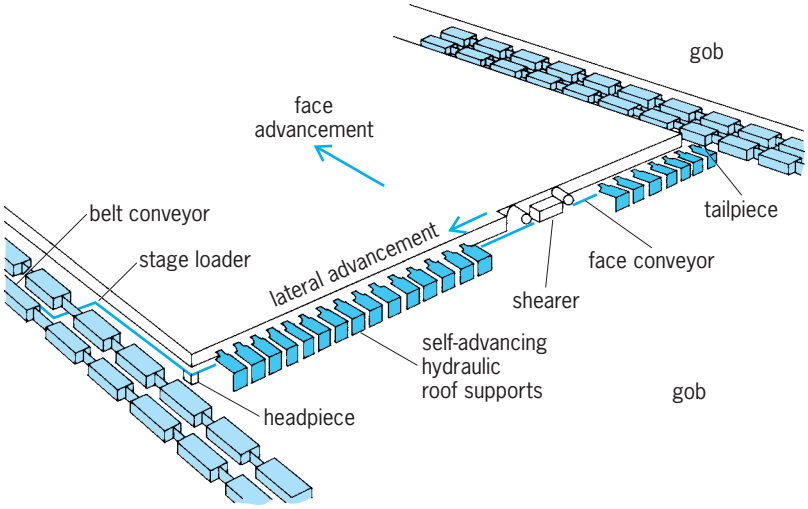

Fig. 15. Diagram showing the basic equipment components on a longwall face. (*After R. Stefanko, Coal Mining Technology, Theory and Practice, Society of Mining Engineers, AIME, 1983*)

Fig. 16. Longwall shearer and hydraulic roof support.

In the modern retreat system, entries (gates or roadways) are driven to block out a coal panel. Coal is then retreat-mined from the far end of the panel back to the submains. In the advancing system, development entries are kept only a short distance ahead of the working face which advances to the far end of the panel. In the United State, panels are usually 600–1100 ft (180–330 m) wide and 3000–17,000 ft (915–5180 m) long.

Regulatory constraints in the United States require longwall panels to be developed by multiple entries similar to those used in room-and-pillar mining. These entries are usually developed by continuous mining methods. Bleeder entries as well as a longwall equipment setup entry are driven across the top of the panel and connect panel entries. The panel entry containing the beltline, electrical and hydraulic equipment, and other items is called the head gate. Its complement on the other side of the panel is called the tail gate. Panels are arranged so that the previous panel's head gate is the new panel's tail gate.

Single head and tail gate entries are common in Europe. When retreat mining is practiced, full-face tunneling machines are used to drive the entries and rock bolts are used for support. These will be supplemented with steel arches or some other method of support as conditions change due to mining. When advance mining is practiced, an in-team miner or stable driving equipment is used. One aspect of longwall mining on the advance is that the head and tail gate must be established and maintained through the gob. This usually means that the roof or floor of these entries must be ripped and the entry cross section reestablished and supported.

Three pieces of equipment are fundamental to modern longwall mining: armored face conveyors, powered supports, and the coal cutting machine. The cycle of face operations is based on the movement of this equipment. The armored face conveyor is erected along the coal face, and is connected to powered supports by means of double-acting hydraulic jacks. The cutter-loader usually slides along the top of the armored conveyor and breaks out a strip of coal 20–30 in. (50–75 cm) wide along the face and then pushes the broken coal onto the conveyor. As the cutter-loader progresses, it leaves a web 20–30 in. (50–75 cm) wide into which the armored conveyor is pushed by the hydraulic jacks on the powered supports. The hydraulic jacks supporting the overhead steel canopy are retracted, freeing the powered supports from roof pressure. These supports are advanced up to the armored conveyor by using the double-acting hydraulic jacks connecting the two, and are then reset. When the cutter-loader completes a pass across the face, it is reset to take a cut as it returns across the face, and the cycle is repeated. As coal is extracted and the powered supports are moved forward, more and more of the roof is left unsupported. The immediate roof collapses and is followed by the main roof collapsing or settling on the broken rock in the gob.

Equipment. An essential element of the modern longwall system is the armored face conveyor, sometimes called a snaking conveyor because in plan it can be curved around behind the cutter-loader into the web that is created as the cutter-loader advances along the face. The face conveyor is a steel trough running the length of the coal face and strong enough to carry the weight of the cutter-loader. One or more endless chains within the trough pull steel flights (paddles) that drag the coal to the head end, where it discharges into the head gate transportation system. The face conveyor may be driven by single or multiple drive units situated at the head and tail gate ends and on either side or both sides of the conveyor. Each drive unit consists of a motor, fluid coupling, and reduction gearing. The overall bulk of these units has an influence on the width of the head and tail gates. The preferred alternative of having drive units at both ends usually cannot be met because the excessive ground pressure on the tail gate makes it difficult to maintain the necessary room for tail drives. Face lengths are limited by the drive arrangements, the chain pull needed to move the coal, and other factors. The face conveyor does not dump directly onto the belt conveyor taking coal out of the panel. Instead it dumps onto a stage loader which serves as a feeder-breaker system for changing the direction and velocity of coal travel and breaking large lumps of coal and rock.

The good roof control developed by powered supports is a key to successful longwall mining. Powered supports have evolved, with a number of different features being introduced as the need arose. A very popular version in the United States is called a shield support. The unit consists of a set of hydraulic props placed between a steel floor beam and a roof beam with the beams pin-connected to a steel caving shield so that the individual members may rotate with respect to each other. Control valves, hydraulic hoses, and a double-acting jack complete the assembly. The caving shield linking the two beams is arranged so that adjacent supports provide skin-to-skin protection against broken rock in the gob flushing into the traveling way and working area of the face. The roof beam cantilevers over the traveling way and working

area of the face, completing the steel canopy protecting workers and giving a prop-free front face. The design principles of the shield allow it to cope with both vertical and lateral movements of the immediate roof as it responds to coal extraction. Powered supports have not been designed to work in thin coal seams.

Cutter-loaders are of two types: shearers and plows. A shearer (Fig. 16) consists of a drum fixed to the end of the ranging arm, and the combined unit is connected through gearing to a power unit riding on the conveyor. The drum is a spiral arrangement fitted with picks for breaking out the coal. A cowl is attached to the machine at the back end of the drum, and the combination of spiral drum and cowl acts as a screw conveyor moving the newly mined coal into the face conveyor. A shearer may have two drums mounted at each end of the power unit. Double-drum shearers with cowls that can be repositioned make bidirectional mining feasible as well as allowing the shearer to accommodate to variable seam height. The shearer moves across the face by a rack-and-pinion mechanism between the shearer and the face conveyor.

The coal plow (planer) is pulled back and forth across the face by an endless chain system. It is designed to travel in the space between the face conveyor and the coal, and cuts a layer of coal 2–6 in. (5–15 cm) thick. The coal is deflected or plowed onto the face conveyor. The plow is held against the face by hydraulic jacks located between the conveyor frame and the plow. An operator does not follow the plow along the face, and the plow does not cut as thick a web as the shearer, so planers can probably work under more difficult roof conditions than shearers.

Ventilation and ground control. The speed of coal production and the length of the longwall face give rise to difficulties in face ventilation. Shearers not only produce a fine product and create dust, but also develop a deep web. The net effect of these factors is an increase in methane liberation. Part of the input face ventilation escapes into the gob, and hence is not used to sweep away and dilute these products. Methane that escapes and collects in the gob area may be drained away by a system of boreholes.

Ground control problems usually result from seam factors. Soft floors may have insufficient bearing capacity for support loads. Massive roof strata with long bridging spans will not cave and overload face support. Variations in seam height and uneven floor conditions can unbalance face support systems. Thick seams must be mined by successive lifts. Tail entries are subjected to severe ground pressures and require heavy timbering and hydraulic prop support.

Other methods. Two of the numerous variations on the basic room-and-pillar and longwall methods will be described. Shortwall mining, using equipment from both room-and-pillar and longwall methods, is an attempt to handle the technical and economic alternatives in choosing an extraction system. The shortwall layout is similar to the longwall panels except that the panel width is 150–200 ft (46–60 m) wide. A continuous miner, loading shuttle cars, substitutes for the cutter-loader-face conveyor system, and the roof beams of the powered supports need to give cantilever action over a wider working area. Shortwall mining is considered when the coal seam is relatively shallow but is overlain by massive roof strata; when a large, relatively homogeneous and undisturbed area of coal is not available for longwall panels; or when ventilation difficulties from dust preclude face personnel from working downwind from the cutter-loader.

Hydraulic mining uses large quantities of water at high pressure to break and convey the coal from the working area. It is most effectively applied to a steeply dipping, relatively thick coal seam that is bounded by an immediate roof and floor and breaks into large pieces. Jets of water are directed tangentially along the coal face to break off lumps of coal. The coal is washed into a system of troughs for conveying away from the working area. Subsequently the coal may be dewatered and moved by conveyor belt, or it may remain in troughs or be pumped to its final location. Hydraulic mining is successfully applied in China, Russia, the United States, and Canada.

Remote control and teleoperation. Remote control enables the operation of mining machines at some distance from the operator. In underground coal mining, remote control of continuous mining machines is a common practice. In all cases, the operator keeps visual contact with the machine, particularly with the cutting head. In cases where visual contact would be limited or impossible, this practice would have little or no advantage. In coal mining, other remotely controlled equipment includes loaders, longwall shearers, belt conveyors, highwall mining machines, roof bolters, and mobile roof support systems. Although health and safety improvements have been made using remotely controlled equipment, efficiency improvements are limited since operation of these systems still requires manual control and visual contact.

In teleoperation, mining machinery is controlled from a remote location without direct visual contact. Ideally, the operator is at the surface or at a relatively safe area and does not enter the work area at all. Some examples have been demonstrated for metal mining, but no applications in coal mining have been implemented. *See* REMOTE-CONTROL SYSTEM.

Future trends. New equipment and equipment modifications to meet the challenges of difficult underground mining conditions will dominate the future. Remotely controlled longwall systems and continuous miners are expected to proliferate for their improved safety and higher productivity. How to mine thin seams, thick seams, steeply dipping seams, seams at greater depths, extremely gassy seams, and seams under bodies of water, or combinations of these needs to be solved. Integration of extraction, haulage, ventilation, ground control, and reclamation systems into an economic and ecologically sound plan is the challenge for future mining engineers. *See* MINING; UNDERGROUND MINING.

Malcolm T. Wane/Tuncel M. Yegulalp

Bibliography. G. Brauner, *Rockbursts in Coal Mines and Their Prevention*, 1994; Code of Federal Regulations, Title 30: *Mineral Resources*, 1997; D. F. Crickmer and D. A. Zegeer, *Elements of Practical Coal Mining*, 2d ed., 1981; H. L. Hartman, *SME Mining Engineering Handbook*, 2 vols., 1992; H. L. Hartman and J. M. Mutmansky, *Introductory Mining Engineering*, 2d ed., 2002; T. Larry, *Handbook of Practical Coal Geology*, 1992; R. D. Merritt, *Coal Exploration, Mine Planning, and Development*, 1986; R. Stefanko, *Coal Mining Technology, Theory and Practice*, 1983.

Coal paleobotany

A special branch of the paleobotanical sciences concerned with the origin, composition, mode of occurrence, and significance of the fossil plant materials that occur in, or are associated with, coal seams. Information developed in this field of science provides knowledge useful to the biologist in efforts to describe the development of the plant world, aids the geologist in unraveling the complexities of coal measure stratigraphy in order to reconstruct the geography of past ages and to describe ancient climates, and has practical application in the coal, coke, and coal chemical industries. *See* COAL CHEMICALS.

Nature of coal seams. All coal seams consist of countless fragments of fossilized plant material admixed with varying percentages of mineral matter. The organic and inorganic materials initially accumulate in some type of swamp environment. Any chemical or physical alteration experienced by the organic fragments during the course of transportation and deposition is followed by another series of changes effected by the chemical, physical, and microbiological agents characterizing the environment in which the particle comes to rest. Subsequent burial beneath a thick cover of sediment induces further physical and chemical alteration of the particles comprising the coal seam. Usually a consolidated layer of well-bonded fragments results. In some instances, crustal deformation and even volcanism add their modifying effects. Thus, the chemical composition, size, shape, and orientation of the fossilized plant remains are influenced both before and after death by biological processes and by the chemical and physical processes attending their postdepositional history. Even in coal seams that have been metamorphosed to anthracitic rank, certain of the constituents remain recognizable as portions of particular plant organs, tissues, or cells. Such entities are classed as phyterals, and in paleobotanical descriptions of coal seams these are identified as specifically as possible as megaspores, cuticles, periderm, and so on. In seams of peat, lignite, and high-volatile bituminous coal, entities are more readily recognized as particular phyterals than they are in higher-rank deposits. This is partly because fewer are destroyed by the metamorphic processes and partly because of the distinctiveness of the substances composing the entity.

In some instances, the fossilized plant fragments can be recognized as remnants of a plant of some particular family, genus, or species. When this is possible, information can be obtained on the vegetation extant at the time the source peat was formed and such data aid greatly in reconstructing paleogeographies and paleoclimatic patterns. Perhaps the most extensive studies of this type are those conducted on the Tertiary brown coals of the Rhine Valley in Germany. From these, detailed reconstructions have been prepared, describing and illustrating the three major swamp environments that gave rise to the sedimentary layers comprising these spectacular coal seams.

Plant fossils in coal seams. Pollen grains and spores are more adequately preserved in coals than are most other plant parts. Recognition of this fact, coupled with an appreciation of the high degree to which these fossils are diagnostic of floral composition, has led to the rapid development of the paleobotanical subscience of palynology. Fruits, seeds, and identifiable woods also occur as coalified fossils and are deserving of more attention than has been accorded them in the past. Often minute structural details are preserved (see **illus.**). Identifiable leaf fossils are comparatively uncommon, although the coalified cuticles of leaves are frequently encountered and their botanical affinities determined. Occasionally, coal seams contain fossil-rich coal balls. Essentially all types of plant fossil, including entire leaves, cones, and seeds, are encountered in these discrete nodular masses. Within the coal ball, the altered but undistorted plant tissues are thoroughly impregnated with mineral matter, usually with calcite or dolomite. Comparatively few coal seams have been encountered with large concentrations of coal balls, but these seams have provided a wealth of detailed information on the nature of the plants which gave rise to the coals concerned. *See* COAL BALLS; FOSSIL SEEDS AND FRUITS; PALYNOLOGY.

Sedimentary units of coal seams. As implied previously, coal seams generally are composed of several superposed sedimentary layers, each having formed under somewhat different environmental conditions. The coal petrologist and paleobotanist recognize these layers because of their distinctive textural appearance and because each consists of a particular association of organic and inorganic materials. Accordingly, each coal seam usually contains several types of coal. These coal types, or lithotypes, possess characteristic suites of physical properties, and knowledge of these properties is very profitably employed in manipulating coal composition in coal preparation and beneficiation plants.

A given lithotype may form several of the constituent layers of lithobodies of a coal seam. Each lithobody often possesses a characteristic assemblage of botanically identifiable fossil plant fragments rendering the unit recognizable to the coal paleobotanist without reference to its textural or compositional features. Paleobotanical descriptions of coal beds may relate to the entire thickness of the seam

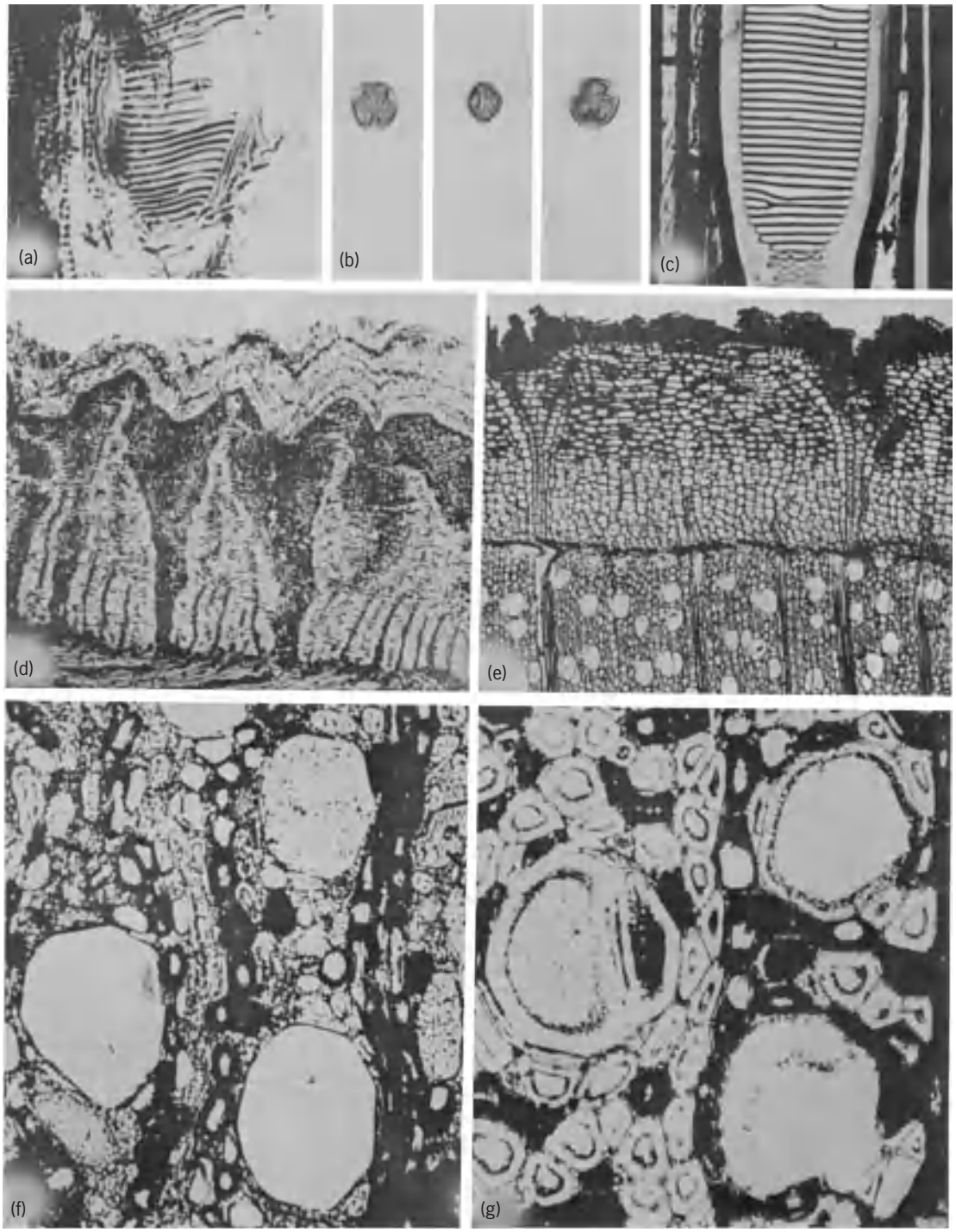

Plant fossils in coal. (*a*) Coalified remnants of scalariform perforation plate in vessel element of piece of lignitized *Cyrilla* wood from Oligocene lignite deposit near Brandon, Vermont. (*b*) Coalified *Cyrilla* pollen grains extracted from Brandon lignite. (*c*) Scalariform perforation plate of vessel in wood of modern *C. racemiflora*. (*d*) Transverse section of lignitized bark of *Cyrilla* from Brandon lignite. (*e*) Transverse section of extant *C. racemiflora* bark with portion of secondary xylem shown in lower half of photo. (*f*) Transverse section of coalified secondary xylem of *Cyrilla* wood from Brandon lignite showing preservation of fine structural detail. (*g*) Transverse section of secondary xylem of *C. racemiflora* showing the structure of vessels, fiber tracheids, wood parenchyma, and rays. Note the coalification effects by making a comparison with *f*.

without regard for the seam's composite nature, or the fossils may be described in relation to the particular lithobody sequence. The more detailed type of description is required if paleoecological interpretations are to be made or if detailed stratigraphic work is involved.

Interrelations with coal petrology. Another facet of coal paleobotany concerns the nature of the substances which compose the coalified plant fragments. When focusing attention upon such matters, the coal paleobotanist is indistinguishable from the coal petrologist and both begin to encroach on the

province of the coal chemist. It is noteworthy, however, that the paleobotanist tends to concentrate attention on the genetic relationships of the various coal substances and hence upon the steps in the derivation and subsequent evolution of maceral materials. The initial stages of coalification tend to be ignored by the coal chemist and petrologist, and the study of these has, quite properly, been thought of as an integral part of coal paleobotany. Thus, through the study of the botanical character of coalified fossils and by means of investigating the substances produced by the coalification process, coal paleobotanists extend knowledge of the organically derived sediments of the Earth's crust. *See* COAL.

William Spackman

Coalbed methane

Coalbed methane production, sometimes referred to as coal degasification, is the process of extracting natural gas from underground coal formations (coalbeds or coal seams). This extraction is accomplished with production technology similar to that used in the conventional natural gas industry. The objectives of producing the natural gas from a coal seam are to improve mine safety, to sell the natural gas, or both. *See* COAL; NATURAL GAS.

Coalbed methane production is a leading form of unconventional production in terms of annual gas production and reserves. Other unconventional gas resources include low-permeability sandstones, Devonian shales, geopressured aquifers, and gas-hydrate reservoirs.

Historical development. The behavior of explosive gases in coal seams has been investigated since the inception of the underground coal mining industry. In 1845, Michael Faraday identified the chain of events leading to devastating mine explosions. First, a localized methane outburst generates a cloud of fine coal particles in the mine. This coal dust may then be ignited through the ignition of the original methane outburst, the ignition of a subsequent methane outburst, or continued mine activity. Clouds of fine coal particles can have an explosive force comparable to black powder. Most of the early coalbed methane production was conducted with the objective of improving mine safety.

With the onset of the energy crisis in the 1970s, a second objective of coalbed methane production was identified: the production of a valuable natural resource, gaseous hydrocarbons.

In 1980, nonconventional fuel tax credits were enacted in the United States to provide economic incentives to stimulate the development of new energy resources.

In 1985, approximately 1×10^9 standard cubic feet of gas were produced from United States coal seams. A standard cubic foot is that volume of gas measured at standardized conditions, for example, 14.7 lb/in^2. (1.5 megapascals) at 60°F (16°C). Due to the increased number of coalbed methane wells and to improvements in production technology, approximately 1.6×10^{11} standard cubic feet of gas were produced from domestic coal seams in 2003. With the success of the United States coalbed methane industry in the 1980s, other governments and international energy companies began investigating the viability of producing natural gas from coal seams outside the United States.

Production and reserves. In 2003, coalbed methane production accounted for approximately 8.0% of the United States' annual natural gas production and 10% of its total proven natural gas reserves. In addition, natural gas production from all of the unconventional gas resources accounted for approximately 33% of the United States' total annual production. **Table 1** provides the annual growth of coalbed

TABLE 1. Coalbed methane proved reserves for 1989–2003, in billion cubic feet at 14.7 psia and 60°F

Year	Alabama	Colorado	New Mexico	Utah	Wyoming	Eastern states*	Western states†	Others‡	United States
1989	537	1,117	2,022	NA§	NA	NA	NA	0	3,676
1990	1,224	1,320	2,510	NA	NA	NA	NA	33	5,087
1991	1,714	2,076	4,206	NA	NA	NA	NA	167	8,163
1992	1,968	2,716	4,724	NA	NA	NA	NA	626	10,034
1993	1,237	3,107	4,775	NA	NA	NA	NA	1,065	10,184
1994	976	2,913	4,137	NA	NA	NA	NA	1,686	9,712
1995	972	3,461	4,299	NA	NA	NA	NA	1,767	10,499
1996	823	3,711	4,180	NA	NA	NA	NA	1,852	10,566
1997	1,077	3,890	4,351	NA	NA	NA	NA	2,144	11,462
1998	1,029	4,211	4,232	NA	NA	NA	NA	2,707	12,179
1999	1,060	4,826	4,080	NA	NA	NA	NA	3,263	13,229
2000	1,241	5,617	4,278	1,592	1,540	1,399	41	—	15,708
2001	1,162	6,252	4,324	1,685	2,297	1,453	358	—	17,531
2002	1,283	6,691	4,380	1,725	2,371	1,488	553	—	18,491
2003	1,665	6,473	4,396	1,224	2,759	1,528	698	—	18,743

* Includes Pennsylvania, Virginia, and West Virginia.
† Includes Kansas, Montana, and Oklahoma.
‡ Includes Oklahoma, Pennsylvania, Utah, Virginia, West Virginia, and Wyoming; these states are individually listed or grouped in Eastern states and Western states after 1999.
§ N/A, not available.
SOURCE: *U.S. Crude Oil, Natural Gas, and Natural Gas Liquids Reserves: 2003 Annual Report, Advance Summary*, Energy Information Administration, Office of Oil and Gas, U.S. DOE, September 2004.

TABLE 2. Coalbed methane production for 1989–2003, in billion cubic feet at 14.7 psia and 60°F

Year	Alabama	Colorado	New Mexico	Utah	Wyoming	Eastern states*	Western states†	Others‡	United States
1989	23	12	56	NA§	NA	NA	NA	0	91
1990	36	26	133	NA	NA	NA	NA	1	196
1991	68	48	229	NA	NA	NA	NA	3	348
1992	89	82	358	NA	NA	NA	NA	10	539
1993	103	125	486	NA	NA	NA	NA	18	752
1994	108	179	530	NA	NA	NA	NA	34	851
1995	109	226	574	NA	NA	NA	NA	47	956
1996	98	274	575	NA	NA	NA	NA	56	1,003
1997	111	312	597	NA	NA	NA	NA	70	1,090
1998	123	401	571	NA	NA	NA	NA	99	1,194
1999	108	432	582	NA	NA	NA	NA	130	1,252
2000	109	451	550	74	133	58	4	NA	1,379
2001	111	490	517	83	278	69	14	NA	1,562
2002	117	520	471	103	302	68	33	—	1,614
2003	98	488	451	97	344	71	51	—	1,600

*Includes Pennsylvania, Virginia, and West Virginia.
†Includes Kansas, Montana, and Oklahoma.
‡Includes Oklahoma, Pennsylvania, Utah, Virginia, West Virginia, and Wyoming; these states are individually listed or grouped in Eastern states and Western states after 1999.
§N/A, not available.
SOURCE: *U.S. Crude Oil, Natural Gas, and Natural Gas Liquids Reserves: 2003 Annual Report, Advance Summary*, Energy Information Administration, Office of Oil and Gas, U.S. DOE, September 2004.

methane proven reserves from 1989 to 2003, while **Table 2** provides the annual coalbed methane production during this period.

Coal seam reservoirs. Coal seam reservoirs are generally characterized as a dual-porosity system (**Fig. 1**). That is, most coal seams contain a primary porosity system (micropores) and a secondary porosity system (macropores, cleat system, or natural fracture system). The primary porosity system consists of the void space within the solid coal matrix. This porosity system is a result of the processes that occur during the deposition of the original organic matter which, over the geological time scale, is eventually converted to coal. Due to the dimensions of the pores (diameters range from 0.8 to 2.0 nanometers), the primary porosity system is inaccessible to ground water and, consequently, is fully saturated with gas. *See* GROUND-WATER HYDROLOGY.

The secondary porosity system of the coal seam consists of the void space in the fractures created after deposition. This fracture system is a result of shrinkage during the coalification process (by which organic matter is converted to coal). Natural fracturing may also be induced by localized tectonic forces. In general, the secondary porosity system in coal seams is composed of two types of fractures: the face cleat and butt cleat. The face cleat is continuous throughout the coal seam, while the butt cleat is discontinuous, generally ending at an intersection with the face cleat. Typically, the face cleat and butt cleat intersect at approximately 60–90° angles with a fairly uniform cleat spacing. Both ground water and gas may exist in the secondary porosity system of coal.

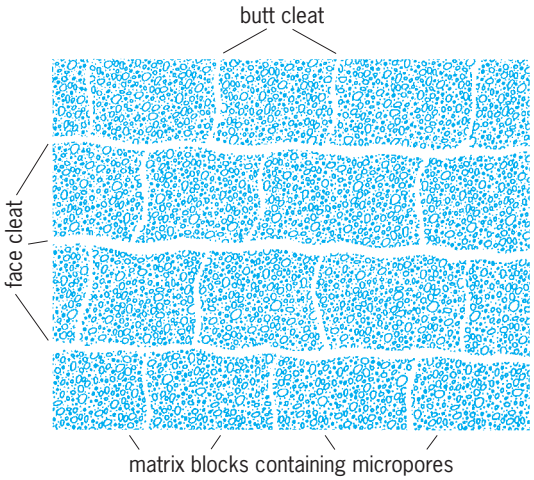

Fig. 1. Plan view of the dual porosity characterization of coal.

Natural gas in coal seams. The natural gas associated with coal is a by-product of the coalification process. Therefore, coal-seam reservoirs are different from conventional natural gas reservoirs; the coal seam is both the source rock and reservoir rock for the gas. Although no universal correlation exists, it has been observed that the original gas content of a coal seam tends to increase with increasing rank (a measure of the maturity and chemical structure of the coal) and increasing depth.

The natural gas associated with coal is composed of low-molecular-weight hydrocarbon gases, predominantly methane, and inorganic gases such as carbon monoxide, carbon dioxide, hydrogen sulfide, and oxygen. The compositions of gases from several coal seams in the United States are listed in **Table 3**. Samples of gas from virgin coal seams have a heat content greater than 950 British thermal units per standard cubic foot, making them of commercial quality with little processing.

Gas storage mechanisms. Gas is stored in coal seams in one of three states: as an adsorbed gas on the pore walls of the coal matrix, as a free gas within the pore structures of the coal, and as gas in solution in resident brine.

TABLE 3. Characteristics of gas from United States coal seams

	Pittsburgh	Pocahontas #3	Mary Lee	Sunnyside	Typical natural gas
Rank:	High-volatile bituminous	Low-volatile bituminous	High-volatile bituminous	High-volatile bituminous	
Geologic era:	Carboniferous (Pennsylvanian)	Carboniferous (Pennsylvanian)	Carboniferous (Pennsylvanian)	Cretaceous	
Gas:					
Methane	90.68*	96.32*	95.97*	99.01*	95.00*
Ethane	0.29	1.39	0.01	0.18	3.50
Propane	—	0.02	—	0.03	0.60
Butane	—	Trace	—	—	0.30
Pentane	—	—	—	—	0.20
Oxygen	0.20	0.17	0.15	0.30	—
Nitrogen	0.59	1.70	3.50	0.48	0.40
Carbon dioxide	8.24	0.36	0.10	—	—
Hydrogen	—	0.01	—	—	—
Helium	—	0.03	0.27	—	—
Heat content, Btu/SCF†:	973	1059	1024	1007	1068

* Chemical composition in mole percent.
† British thermal units/standard cubic foot.

Due to the dimensions of the pores in the primary porosity system and the organic, carbon-rich nature of coal, a large internal surface area containing many potential adsorption sites is present in the coal matrix. The majority of gas stored in coal exists as a physically adsorbed, molecular monolayer on the walls of the primary porosity system. The gas adsorption phenomenon in coal conforms to Langmuir theory and is commonly modeled with a Langmuir adsorption isotherm. *See* ADSORPTION.

In addition to adsorbed gas, free gas may exist in the secondary porosity system of the coal seam. Free-gas storage results from the compressible nature of gas in the presence of the hydrostatic pressure exerted by local ground water. This is the gas storage mechanism that typically dominates in conventional natural gas reservoirs. Free-gas storage is generally modeled with the real gas law.

To a lesser extent, gas is stored in solution in the ground water within the secondary porosity system. The amount of gas dissolved in connate water is dependent on the water salinity and on the temperature and pressure of the coal seam.

Gas transport mechanisms. With gas and water production and the resulting pressure depletion, gas is desorbed from the pore walls of the coal matrix and migrates through the primary porosity system to the natural fractures. The gas is then transported, along with ground water, through the natural fracture system to the production wells. Thus, the natural fracture system acts both as a low-pressure sink to the primary porosity system and as a conduit to the production wells. Gas transport through coal seams is a two-stage process (**Fig. 2**).

Due to low permeability and small pore dimensions, the principal gas-transport mechanism in the primary porosity system is diffusion. Within the primary porosity system, gas can be transported by bulk diffusion (where molecule–molecule interactions dominate), Knudsen diffusion (where molecule–surface interactions dominate), or surface diffusion of the adsorbed gas layer (where molecular movement from adsorption site to adsorption site dominates). Regardless of the dominant diffusion mechanism, gas transport through the primary porosity system obeys Fick's law. The potential that drives the diffusion process is the partial pressure (concentration) gradient. *See* DIFFUSION; RAREFIED GAS FLOW.

Within the secondary porosity system of the coal seam, gas and water are transported by Poiseuille, or forced, flow. For porous media such as the secondary

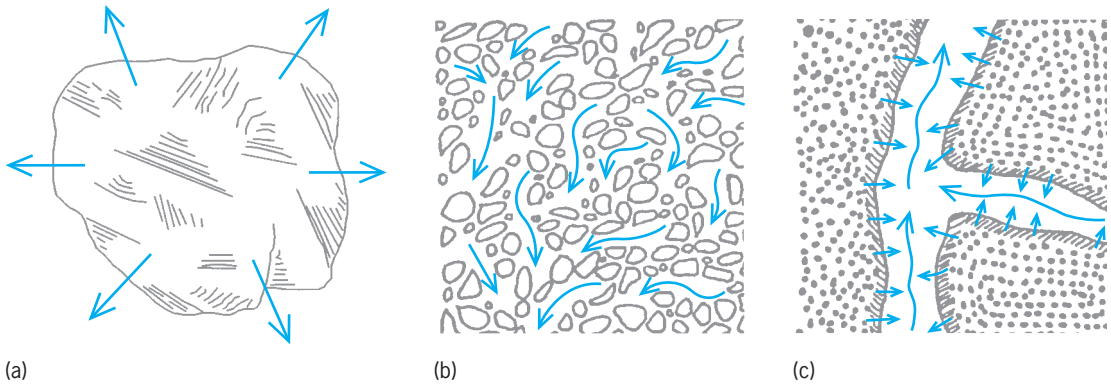

Fig. 2. Transport of gas through coal. (*a*) Gas desorption from coal grains in the coal matrix. (*b*) Step 1: diffusion through the primary porosity system to the cleat system. (*c*) Step 2: Poiseuille flow through the cleat system to production wells.

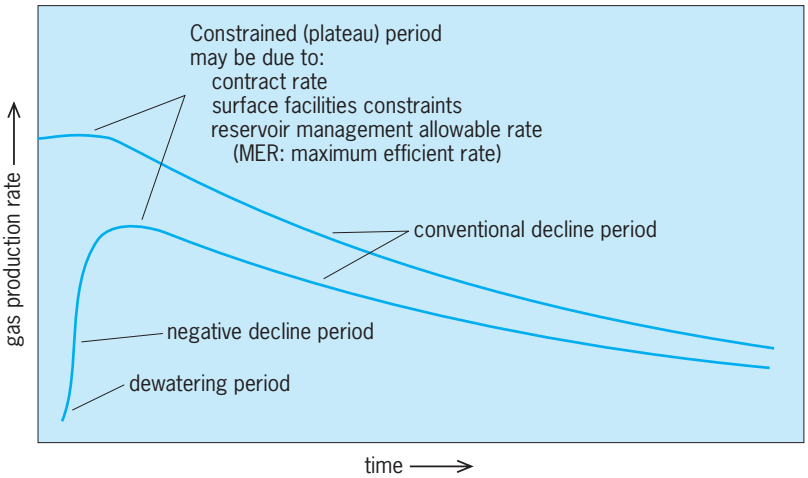

Fig. 3. Comparision of typical gas production rates of a coalbed methane well and a conventional natural gas well.

porosity system in a coal seam, Poiseuille flow is typically modeled with Darcy's law. The potential that drives Poiseuille flow is the total pressure gradient. This is the transport mechanism that normally dominates in conventional natural gas reservoirs. *See* SOIL MECHANICS; VISCOSITY.

Production technology. The drilling technology used for coalbed methane wells is identical to the technology used for conventional natural gas wells. The production technologies, however, do differ in well completions. The well completion is the portion of the well that connects the wellbore to the reservoir. Specialized well completions are required for coalbed methane wells because of the unique mechanical properties of coal.

There are two well completions commonly used in coalbed methane wells: open-hole cavity completions and hydraulically stimulated completions. In the open-hole cavity completion, the well is injected with high-pressure gas and water and is then rapidly depressurized. This process is repeated several times to break down the coal in order to create a large effective well radius, or cavity, in the coal seam.

In the hydraulic stimulation process, a high-pressure fluid (usually a gelled liquid, foam, or gas) is injected into the well along with a proppant (generally sand). The injection of the fracturing fluid either creates and propagates a fracture from the well into the coal seam, or opens preexisting, near-wellbore fractures. The proppant is used to keep the fractures open once the well is put into production. The hydraulic stimulation process used for coalbed methane wells differs from the hydraulic fracturing process used for conventional natural gas wells in that stimulation treatment must be optimized for the low-bulk modulus of coal. Most of the conventional hydraulic fracture technology has been developed for high-bulk modulus clastic and carbonate reservoirs.

Two characteristics distinguish gas production from coalbed methane wells: (1) a dewatering period, which may take up to several months, is often required to initiate the desorption process and, eventually, gas production; and (2) a "negative" gas-decline period, during which the gas production rate increases with time. This is in contrast to the conventional behavior in which gas production normally decreases with time. The dewatering and negative gas-decline periods are a direct result of the unique mechanisms of gas storage and transport in coal seams (**Fig. 3**). Gregory R. King

Bibliography. *A Guide to Coalbed Methane Reservoir Engineering*, GRI-94-0397, Gas Research Institute, Chicago, 1996; *Coalbed Methane*, SPE Reprint 35, Society of Petroleum Engineers, Richardson, TX, 1992; C. T. Rightmire et al. (eds.), *Coalbed Methane Resources of the United States*, AAPG Stud. Geol. 17, American Association of Petroleum Geologists, Tulsa, 1984.

Coastal engineering

A branch of civil engineering concerned with the planning, design, construction, and maintenance of works in the coastal zone. The purposes of these works include control of shoreline erosion; development of navigation channels and harbors; defense against flooding caused by storms, tides, and seismically generated waves (tsunamis); development of coastal recreation; and control of pollution in nearshore waters. Coastal engineering usually involves the construction of structures or the transport and possible stabilization of sand and other coastal sediments.

The successful coastal engineer must have a working knowlege of oceanography and meteorology, wave mechanics, geomorphology, soil mechanics, statistics, and structural mechanics. Tools that support coastal engineering design include analytical theories of wave motion, wave-structure interaction, diffusion in a turbulent flow field, and so on; numerical and physical hydraulic models; basic experiments in wave and current flumes; and field measurements of basic processes such as beach profile response to wave attack, and the construction of works. Postconstruction monitoring efforts at coastal projects have also contributed greatly to improved design practice.

Environmental forces. The most dominant agent controlling coastal processes and the design of coastal works is usually the waves generated by the wind. Wind waves produce large forces on coastal structures, they generate nearshore currents and the alongshore transport of sediment, and they mold beach profiles. Thus, a primary concern of coastal engineers is to determine the wave climate (statistical distribution of heights, periods, and directions) to be expected at a particular site. This includes the annual average distribution as well as long-term extreme characteristics. In addition, the nearshore effects of wave refraction, diffraction, reflection, breaking, and runup on structures and beaches must be predicted for adequate design. *See* OCEAN WAVES.

Other classes of waves that are of practical importance include the astronomical tide, tsunamis, and waves generated by moving ships. The tide raises

and lowers the nearshore water level and thus establishes the range of shoreline over which coastal processes act. It also generates reversing currents in inlets, harbor entrances, and other locations where water motion is constricted. Tidal currents which often achieve a velocity of 3–6 ft/s (1–2 m/s) can strongly affect navigation, assist with the maintenance of channels by scouring sediments, and dilute polluted waters. *See* TIDE.

Tsunamis are quite localized in time and space but can produce devastating effects. Often the only solution is to evacuate tsunami-prone areas or suffer the consequences of a surge that can reach elevations in excess of 33 ft (10 m) above sea level. Some attempts have been made to design structures to withstand tsunami surge or to plant trees and construct offshore works to reduce surge velocities and runup elevations. *See* TSUNAMI.

The waves generated by ships can be of greater importance at some locations than are wind-generated waves. Ship waves can cause extensive bank erosion in navigation channels and undesirable disturbance of moored vessels in unprotected marinas.

On coasts having a relatively broad shallow offshore region (such as the Atlantic and Gulf coasts of the United States), the wind and lower pressures in a storm will cause the water level to rise at the shoreline. Hurricanes have been known to cause storm surge elevations of as much as 16 ft (5 m) for periods of several hours to a day or more. Damage is caused primarily by flooding, wave attack at the raised water levels, and high wind speeds. Defense against storm surge usually involves raising the crest elevation of natural dune systems or the construction of a barrier-dike system. *See* STORM SURGE.

Besides the cyclic variations in the mean sea level along the coast that are caused by the tide, tsunamis, and storm surge, there are longer-term sea-level variations owing to land uplift or subsidence with a superimposed rise in water levels that has been associated with atmospheric warming trends. The result in most coastal locations is a net sea-level rise relative to the land. In turn, there is a slow recession of the shoreline owing to the raised water level and to the beach profile's response to this raised mean sea level.

Other environmental forces that impact on coastal works include earthquake disturbances of the sea floor and static and dynamic ice forces. Direct shaking of the ground will cause major structural excitations over a region that can be tens of kilometers wide surrounding the epicenter of a major earthquake. Net dislocation of the ground will modify the effect of active coastal processes and environmental forces on structures. *See* EARTHQUAKE.

Ice that is moved by flowing water and wind or raised and lowered by the tide can cause large and often controlling forces on coastal structures. However, shore ice can prevent coastal erosion by keeping wave action from reaching the shore.

Coastal processes. Wind-generated waves are the dominant factor that causes the movement of sand parallel and normal to the shoreline as well as the resulting changes in beach morphology. Thus, structures that modify coastal zone wave activity can strongly influence beach processes and geometry.

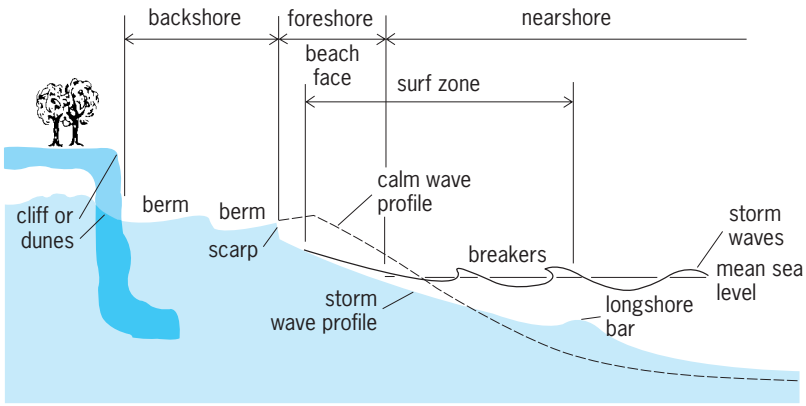

Fig. 1. Typical beach profiles (vertical scale is exaggerated).

Active beach profile zone. In typical beach profiles found at a sandy shoreline (which may be backed by a cliff or a dune field), the backshore often has one or two berms with a crest elevation equal to the height of wave runup during high tide (**Fig. 1**). When low swell, common during calm conditions, acts on the beach profile, the beach face is built up by the onshore transport of sand. This accretion of sand adds to the seaward berm. On the other hand, storm waves will attack the beach face, cut back the berm, and carry sand offshore. This active zone of shifting beach profiles occurs primarily landward of the 33-ft (10-m) depth contour. If the storm tide and waves are sufficiently high, the berms may be eroded away to expose the dunes or cliff to erosion. The beach profile changes (Fig. 1) are superimposed on any longer term advance or retreat of the shoreline caused by a net gain or loss of sand at that location.

Any structure constructed along the shore in the active beach profile zone may retain the sand behind it and thus reduce or prevent erosion. However, wave attack on the seaward face of the structure causes increased turbulence at the base of the structure, and usually increased scour which must be allowed for or prevented, if possible, by the placement of stone or some other protective material.

It is desirable to keep all construction of dwellings, recreational facilities, and other such structures landward of the active beach profile zone, which usually means landward of the frontal dunes or a good distance back from retreating cliffs. It is also desirable to maintain and encourage the growth of the frontal dune system by planting grass or installing sand fencing.

In addition to constructing protective structures and stabilizing the dune system, it is common practice to nourish a beach by placing sand on the beach face and nearshore area. This involves an initial placement of sand to develop the desired profile and periodic replenishment to make up for losses to the profile. Often shore stabilization structures are constructed along with beach nourishment to increase

the stability of the nourished beach section. A common source for sand, which should be clean and at least as coarse as the native sand, is the offshore area near the nourishment site. Sand for beach nourishment has also been obtained from bays and estuaries and from land quarries.

Alongshore current and transport. Waves arriving with their crest oriented at an angle to the shoreline will generate a shore-parallel alongshore current in the nearshore zone. The current flows in the direction of the alongshore component of wave advance and has the highest velocity just inside the breaker line. It may be assisted or hindered by the wind and by tidal action, particularly along sections of the shore adjacent to tidal inlets and harbor entrances. There is a continuous accumulation of flow in the downcoast direction which may be relieved by seaward-flowing jets of water known as rip currents.

The alongshore current transports sand in suspension and as bed load, and is assisted by the breaking waves, which place additional sand in suspension. Also, wave runup and the return flow transports sand particles in a zigzag fashion along the beach face. Coastal structures that obstruct these alongshore transport processes can cause the deposition of sand. They do this by blocking wave action from a section of the shore and thus removing the wave energy required to maintain the transport system; by interfering with the transport process itself; or by directly shutting off a source of sand that feeds the transport system (such as a structure that protects an eroding shoreline).

The design of most coastal works requires a determination of the volumetric rate of alongshore sand transport at the site—both the gross rate (upcoast plus downcoast transport) and the net rate (upcoast minus downcoast transport). The most reliable method of estimating transport rates is by measuring the rate of erosion or deposition at an artificial or geomorphic structure that interrupts the transport. Also, field studies have developed an approximate relationship between the alongshore transport rate and the alongshore component of incident wave energy per unit time. With this, net and gross transport rates can be estimated if sufficient information is available on the annual wave climate. Typical gross transport rates on exposed ocean shorelines often exceed 500,000 yd^3 (382,000 m^3) per year.

Primary sources of beach sediment include rivers discharging directly to the coast, beach and cliff erosion, and artificial beach nourishment. Sediment transported alongshore from its sources will eventually be deposited at some semipermanent location or sink. Common sinks include harbors and tidal inlets; dune fields; offshore deposition; spits, tombolos, and other geomorphic formations; artificial structures that trap sand; and areas where beach sand is mined.

By evaluating the volumetric transports into and out of a segment of the coast, a sediment budget can be developed for the coastal segment. If the supply exceeds the loss, shoreline accretion will occur, and vice versa. When a coastal project modifies the supply or loss to the segment, geomorphic changes can be expected. For example, when a structure that traps sediment is constructed upcoast of a point of interest, the shoreline at the point of interest can be expected to erode as it resupplies the longshore transport capacity of the waves.

Harbor entrance and tidal inlet control structures built to improve navigation conditions, stabilize navigation channel geometry, and assist with the relief of flood waters will often trap a large portion of the alongshore sand transport. This can result in undesirable deposition at the harbor or inlet entrance and subsequent downcoast erosion. The solution usually involves designing the entrance structures to trap the sediment at a fixed and acceptable location and to provide protection from wave attack at this location so a dredge can periodically pump the sand to the downcoast beach. *See* NEARSHORE PROCESSES.

Computer models. Computer modeling of wave and coastal processes has greatly enhanced the ability of coastal engineers to more realistically and accurately define the effects of these processes. Specific examples of computer models in common use are models for the forecasting of wind-generated wave action caused by hurricanes and extratropical storms; prediction of wave transformation as waves propagate from deep water to the shore; prediction of related surge levels during storms; and prediction of beach morphology changes over a coastal segment caused by incident wave action. Computer models are also available that will predict the changing wave characteristics as the tide and tsunami waves enter the coastal zone, as well as the resulting resonant response of harbors and estuaries to these waves.

Coastal structures. Coastal structures can be classified by the function they serve and by their structural features. Primary functional classes include seawalls, revetments, and bulkheads; groins; jetties; breakwaters; and a group of miscellaneous structures including piers, submerged pipelines, and various harbor and marina structures.

Seawalls, revetments, and bulkheads. These structures are constructed parallel or nearly parallel to the shoreline at the land-sea interface for the purpose of maintaining the shoreline in an advanced position and preventing further shoreline recession. Seawalls are usually massive and rigid, while a revetment is an armoring of the beach face with stone rip-rap or artificial units. A bulkhead acts primarily as a land-retaining structure and is found in a more protected environment such as a navigation channel or marina.

A key factor in the design of these structures is that erosion can continue on adjacent shores and flank the structure if it is not tied in at the ends. Erosion on adjacent shores also increases the exposure of the main structure to wave attack. Structures of this class are prone to damage and possible failure caused by wave-induced scour at the toe. In order to prevent this, the toe must be stabilized by driving vertical sheet piling into the beach, laying stone on the beach seaward of the toe, or maintaining a protective beach by artificial nourishment. Revetments that are sufficiently porous will allow leaching of sand from behind the structure. This can lead to

structure slumping and failure. A proper stone or cloth filter system must be developed to prevent damage to the revetment. In an attempt to develop low-cost shore protection, a number of novel materials have been used for shoreline revetments, including cinder blocks, tires, sand-filled rubber tubes, woven-fiber mattresses, and soil-cement paving. *See* REVETMENT.

Groins. A groin is a structure built perpendicular to the shore and usually extending out through the surf zone under normal wave and surge-level conditions. It functions by trapping sand from the alongshore transport system to widen and protect a beach or by retaining artificially placed sand (**Fig. 2**). Typical groin alongshore spacing-to-length ratios vary from 1.5:1 up to 4:1.

There may be erosion downcoast of the groin field, the volume of erosion being approximately equal to the volume of sand removed by the groins from the alongshore transport system. Groins must be sufficiently tied into the beach so that downcoast erosion superimposed on seasonal beach profile fluctuations does not flank the landward end of a groin. Even the best-designed groin system will not prevent the loss of sand offshore in time of storms.

Jetties. Jetties are structures built at the entrance to a river or tidal inlet to stabilize the entrance as well as to protect vessels navigating the entrance channel. Stabilization is achieved by eliminating or reducing the deposition of sediment coming from adjacent shores and by confining the river or tidal flow to develop a more uniform and hydraulically efficient channel. Jetties improve navigation conditions by eliminating bothersome crosscurrents and by reducing wave action in the entrance.

At many entrances there are two parallel (or nearly parallel) jetties that extend approximately to the seaward end of the dredged portion of the channel. However, at some locations a single updrift or downdrift jetty has been used, as have other arrangements such as arrowhead jetties (a pair of straight or curved jetties that converge in the seaward direction). Jetty layouts may also be modified to assist sediment-bypassing operations. A unique arrangement is the weir-jetty system in which the updrift jetty has a low section or weir (crest elevation about mean sea level) across the surf zone. This allows sand to move over the weir section and into a deposition basin for subsequent transport to the downcoast shore by dredge and pipeline.

Breakwaters. The primary purpose of a breakwater is to protect a shoreline or harbor anchorage area from wave attack (**Fig. 3**). Breakwaters may be located completely offshore and oriented approximately parallel to shore, or they may be oblique and connected to the shore where they often take on some of the functions of a jetty. At locations where a natural inland site is not available, harbors have been developed by the construction of shore-connected breakwaters that cover two or three sides of the harbor.

The offshore breakwater intercepts incident waves, thus trapping the predominantly southeastern longshore sand transport; it provides a protected area where a dredge can operate to bypass sediment; and it provides protection to the harbor entrance. A series of shore-parallel offshore breakwaters (with or without artificial nourishment) has been used for shore protection at a number of locations. If there is sufficient fill or trapped material, the tombolo formed in the lee of each breakwater (Fig. 3) may grow until it reaches the breakwater.

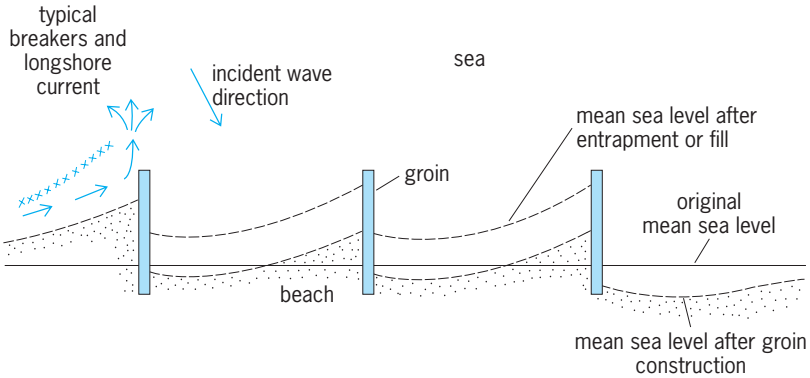

Fig. 2. Groin system and beach response.

Breakwaters are designed to intercept waves, and often extend into relatively deep water, so they tend to be more massive structures than are jetties or groins. Breakwaters constructed to provide a calm anchorage area for ships may have a high crown elevation to prevent overtopping by incident waves and subsequent regeneration of waves in the lee of the breakwater.

Rubble mound structures. Groins, jetties, and breakwaters are most commonly constructed as rubble mound structures (**Fig. 4**). The breakwater has an outer armor layer consisting of the largest stones or, if sufficiently large stones are not available, the armor units may be molded of concrete with a special shape. Stone sizes decrease toward the core and base in order to develop a filter system so that the fine core stone and base sand are not removed by wave and current action. The core made of fine stone sizes is provided to diminish wave transmission through

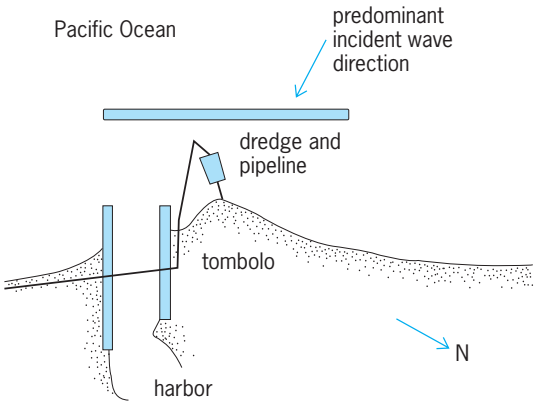

Fig. 3. Overhead view of breakwater and jetty system at the entrance to Channel Islands Harbor, California.

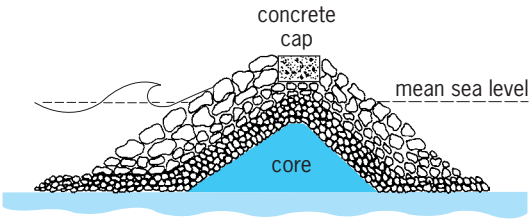

Fig. 4. Cross section of a typical rubble mound breakwater placed on a sand foundation.

the structure. Jetties and groins have a simpler cross section, consisting typically of only armor and core layers. Breakwaters, groins, and jetties have also been constructed of steel or concrete caissons with sand and gravel fill; wood, steel, and concrete sheet piles; and sand-filled bags.

A different type of breakwater that can be effective where incident wave periods are short and large water-level fluctuations occur (as in a reservoir marina) is the moored floating breakwater. This type has been constructed of hollow concrete prisms, scrap tires, logs, and a variety of other materials.

Economic, environmental, and esthetic considerations have resulted in some changes in the types of rubble mound structures being constructed along the coast. Rather than structures with several defined layers (Fig. 4) including an armor layer of large stone designed to remain stable during wave attack, berm breakwaters and revetments have been constructed in regions where large stones are not readily available. The armor layer in these structures consists of smaller stones with a larger size variation that is expected to remold itself to a stable profile in the same manner that a beach profile responds to wave attack. This armor layer must be thicker to allow for this remolding action.

Also, there has been a trend in offshore shore-parallel breakwaters to build the structure with its crest below sea level. This approach saves on stone costs and is visually more acceptable. During calm wave conditions, most of the wave action reaches the shore, a positive condition for shore recreation. However, the larger storm waves are broken and reflected by the structure so that storm-induced beach erosion is reduced. Robert M. Sorensen

Bibliography. P. Bruun, *Port Engineering*, 1990; J. B. Herbich, *Handbook of Coastal and Ocean Engineering*, 1990; R. M. Sorensen, *Basic Coastal Engineering*, 1978; R. M. Sorensen, *Basic Wave Mechanics: For Coastal and Ocean Engineers*, 1993; U.S. Army Coastal Engineering Research Center, *Shore Protection Manual*, 1984.

Coastal landforms

The characteristic features and morphology of the land in the coastal zone. They are subject to processes of erosion and deposition as produced by winds, waves, tides, and river discharge. The interactions of these processes and the coastal environments produce a wide variety of landforms. Processes directed seaward from the land are dominated by the transport of sediment by rivers, but also include gravity processes such as landslides, rockfalls, and slumping. The dominant processes on the seaward side are wind, waves, and wave-generated currents. Mixed among these locations are tidal currents which also carry large volumes of sediment.

All of these processes are relatively continuous or rapid. Some processes that influence coastal development are much slower. Changes in sea level are important over centuries and millennia with rates of a few millimeters per year. Over thousands of years these changes have resulted in an increase or decrease in sea level of as much as 100 m (330 ft). Another process that influences coastal development is plate tectonics, which has a time frame of millions of years. Although not a process, the geologic framework and geomorphology both seaward and landward of the shoreline are also factors in the development of coastal landforms. *See* GEOMORPHOLOGY; PLATE TECTONICS.

Subcontinental- to continental-scale coastal landform patterns are related to plate tectonics. A landmark paper by D. Inman and C. Nordstrom relating coastal development to plate tectonics provided a simple but useful approach to classification of coasts on this scale. It showed that there are three general geologic framework types into which all coasts can be classified. Some coasts are associated with plate boundaries and others are not. The three major tectonic coastal types are leading-edge, trailing-edge, and marginal sea coasts. Leading-edge coasts are associated with colliding plate boundaries where there is considerable tectonic activity. Trailing-edge coasts are on stable continental margins, and marginal seas have fairly stable coasts with plate margins, commonly characterized by island arcs and volcanoes, that form their seaward boundaries. *See* CONTINENTAL MARGIN.

Leading-edge coasts. The primary characteristic of these coasts is the rugged and irregular topography, commonly displaying cliffs or bluffs right up to the shoreline (**Fig. 1**). Seaward, the topography reflects this with an irregular bottom and deep water near the shoreline. The geology is generally complex with

Fig. 1. Rugged, high-relief coast in central Oregon, an example of a leading-edge coast.

Fig. 2. Central Atlantic coastal plain and coast of the United States, showing large estuaries bounded on the seaward side by a system of long barrier islands. (*NASA*)

Fig. 3. Mississippi Delta, an example of a large muddy coast landform in a marginal sea.

numerous faults and folds in the strata of the coastal zone. These coasts tend to be dominated by erosion with only local areas of deposition, typically in the form of small beaches or spits between headlands. Waves tend to be large because of the deep nearshore water, and form wave-cut platforms, terraces, notches, sea stacks, and caves.

Trailing-edge coasts. The most diverse suite of coastal landforms develops along trailing-edge coasts. These coasts are generally developed along the margin of a coastal plain, they are fed by well-developed and large river systems, and they are subjected to a low-to-modest wave energy because of the gently sloping adjacent continental shelf. The overall appearance is little topographic relief dominated by deposition of mud and sand. The spectrum of environments and their associated landforms includes deltas, estuaries, barrier islands (**Fig. 2**), tidal inlets, tidal flats, and salt marshes. *See* BARRIER ISLANDS; COASTAL PLAIN; DELTA; FLOODPLAIN; SALT MARSH.

Although there is typically only low-to-moderate wave energy along such coasts, there is potential for considerable change in landforms through the passage of intense storms, including hurricanes. It is common for such events to cause more change in the coast in one or two days than might take place in a century or more under so-called normal conditions.

These trailing-edge coasts are subdivided into three categories based upon the maturity of the continental margin: neo-trailing-edge coasts, Afro-trailing-edge coasts, and Amero-trailing-edge coasts. The neo-trailing-edge coast develops only a short time after the splitting and spreading apart of a landmass. There has been insufficient time for development of a coastal plain, and the continental margin is narrow and steep but not tectonically very active. The Red Sea that separates Africa from the Middle East is an excellent example of this type. The Afro-trailing-edge coast has a significant level of maturity and a well-developed continental margin and is characterized by both the west coast and southeast coast of Africa. The Amero-trailing-edge coast is one where a very long geological time has passed since the plates separated. Both the continental margin and the coastal plain are wide and well developed. The east coasts of North and South America are excellent examples.

Marginal sea coasts. Some coastal settings are along a stable continental mass and are protected from open ocean processes, commonly an island arc system or other form of a plate boundary. The consequence is a coastal zone that tends to be subjected to small waves and where considerable mud is allowed to accumulate in the coastal zone. Many coastal landforms are present in much the same fashion as on the trailing-edge coasts. Marginal sea coasts tend to have large river deltas (**Fig. 3**). Examples are the eastern margin of Asia along the Gulf of Korea and the China Sea, the entire Gulf of Mexico, and the Mediterranean Sea. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; NEARSHORE PROCESSES.

Richard A. Davis, Jr.

Bibliography. E. C. F. Bird, *Coasts: An Introduction to Systematic Geomorphology*; R. W. G. Carter, *Coastal Environments*, 1988; R. A. Davis, *The Evolving Coast*, 1993; D. L. Inman and C. E. Nordstrom, On the tectonic and morphologic classification of coasts, *J. Geol.*, 79:1–21, 1971; J. S. Pethick, *Introduction to Coastal Geomorphology*, 1983.

Coastal plain

An extensive, low-relief area that is bounded by the sea on one side and by some type of relatively high-relief province on the landward side. The geologic province of the coastal plain actually extends beyond

Fig. 1. Coastal plains of the United States.

the shoreline across the continental shelf. It is only during times of glacial melting and high sea level that much of the coastal plain is drowned. *See* CONTINENTAL MARGIN.

The coastal plain is a geologic province that is linked to the stable part of a continent on the trailing edge of a plate. The extent and nature of the coastal plains of the world range widely. Some are very large and old, whereas others are small and geologically young. For example, the Atlantic and Gulf coasts of the United States are among the largest in the world. They extend continuously from the New Jersey area beyond the border with Mexico (**Fig. 1**). In some areas the coastal plain is hundreds of kilometers wide. In the United States it extends back about 100 million years. By contrast, local coastal plains in places like the east coast of Australia and New Zealand are only 1 or 2 million years old and extend only tens of kilometers landward from the shoreline.

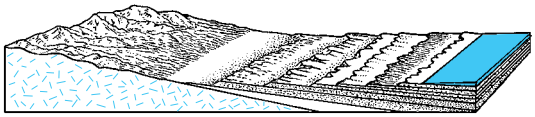

Fig. 2. Block diagram of a coastal plain showing the gentle dip of the strata and the erosional ridges.

The typical character of a coastal plain is one of strata that dip gently and uniformly toward the sea. There may be low ridges that are essentially parallel to the coast (**Fig. 2**) and that have developed from erosion of alternating resistant and nonresistant strata. These strata are commonly a combination of mudstone, sandstone, and limestone, although the latter is typically a subordinate amount of the total. These strata resulted from deposition in fluvial, deltaic, and shelf environments as sea level advanced and retreated over this area. Coastal plain strata have been a source of considerable oil and gas as well as various economic minerals. Although the coastal plain province is typically stable tectonically, there may be numerous normal faults and salt dome intrusions. *See* COASTAL LANDFORMS; DELTA; GEOMORPHOLOGY; PLAINS. Richard A. Davis, Jr.

Bibliography. J. S. Pethick, *Introduction to Coastal Geomorphology*, 1983.

Coati

Any of three species of carnivorous mammals assigned to the raccoon family Procyonidae. The common or ring-tailed coati (*Nasua nasua*) ranges through the forests of Central and South America (see **illustration**), the mountain coati (*Nasuella olivacea*) is found only in South America, while *Nasua nelsoni* is confined to Cozumel Island in Central America.

The coatis are characterized by their elongated snouts, body, and tail, which is held erect when they walk. They are adept at climbing trees in search of birds and lizards, and their diet also includes fruit,

Ring-tailed coati (*Nasua nasua*). (*Photo © by Lisa Kostich*)

insects, and larvae. *Nasua nasua* roams in bands of females and young during the day in search of food. The males are excluded from the band except during the brief breeding season, when a single male is allowed to join. He is aggressive toward any other male but subordinate to the females. After mating, the male leaves and the female builds an isolated nest in a tree. Here, following a gestation period of 11 weeks, a litter of about five young are born and cared for by the female away from the group. After about 1 month the female and her young rejoin the band. *See* CARNIVORA. Charles B. Curtin

Coaxial cable

An electrical transmission line comprising an inner, central conductor surrounded by a tubular outer conductor. The two conductors are separated by an electrically insulating medium which supports the inner conductor and keeps it concentric with the outer conductor. One version of coaxial cable has periodically spaced polyethylene disks supporting the inner conductor (**Fig. 1**). This coaxial is a building block of multicoaxial cables used in L-carrier systems. *See* TRANSMISSION LINES.

The symmetry of the coaxial cable and the fact that the outer conductor surrounds the inner conductor make it a shielded structure. At high frequencies, signal currents concentrate near the inside surface of the outer conductor and the outer surface of the inner conductor. This is called skin effect. The depth to which currents penetrate decreases with increasing frequency. Decreased skin depth improves the cable's self-shielding and increases transmission loss. This loss (expressed in decibels per kilometer) increases approximately as the square root of

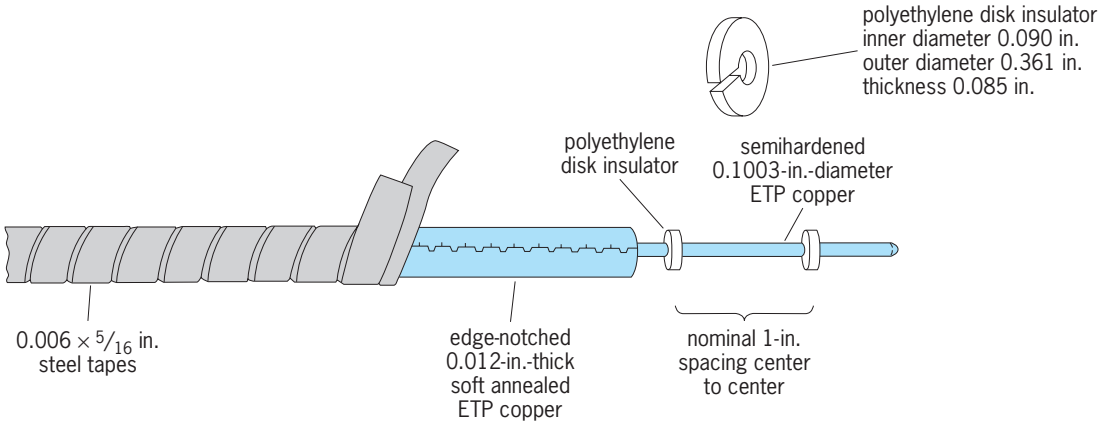

Fig. 1. Air dielectric disk-insulated coaxial cable. 1 in. = 25 mm.

frequency because of the skin effect. *See* ELECTRICAL SHIELDING; SKIN EFFECT (ELECTRICITY).

Coaxial cables can carry high power without radiating significant electromagnetic energy. In other applications, coaxial cables carry very weak signals and are largely immune to interference from external electromagnetic fields.

A coaxial cable's self-shielding property is vital to successful use in broadband carrier systems, undersea cable systems, radio and TV antenna feeders, and community antenna television (CATV) applications.

Coaxial units are designed for different mechanical behavior depending upon the application. Widely used coaxials are classified as flexible or semirigid.

Materials. Conductors for coaxial cable are generally made of copper or aluminum. The low resistivity of these metals results in low cable loss. A typical dielectric material is polyethylene; special grades are made with very low dielectric loss. Where mechanical factors allow, much of the space between inner and outer conductor is occupied by air. This may be done by using spaced insulating disks (Fig. 1) or expanded or spiraled insulation. *See* ALUMINUM; CONDUCTOR (ELECTRICITY); COPPER; POLYOLEFIN RESINS.

Flexible coaxials. These generally have a braided outer conductor and a stranded inner conductor. Insulation is made of solid or expanded plastic of a low-loss electrical grade. Since a single layer of braid has some air gaps between intersecting wires, a second layer of braid is placed over the first when shielding requirements are severe.

Flexible coaxials are used for high-frequency patch cords and intra- and interbay cabling. Very small-diameter coaxials are used to interconnect components on printed wiring boards. Since runs are short in these applications, loss is not generally important.

Semirigid coaxials. This class of coaxials finds the most extensive use in communications systems, such as cables used in long-distance terrestrial and undersea communications. Also, CATV and closed-circuit television often use semirigid coaxials. To minimize cable loss, the insulation may be of expanded plastic, effectively using air bubbles for up to 80% of the dielectric.

Terrestrial coaxials. The coaxial unit in the L3, L4, L5, and L5E carrier systems uses polyethylene disks spaced 1 in. (2.5 cm) apart to keep the inner conductor centered (Fig. 1). The outer conductor is then precision-formed around the disk-insulated inner conductor. Two layers of steel tape are overlapped to complete the unit. The tapes give the assembly mechanical stability and provide added shielding at low frequencies.

Such a coaxial unit has low loss and excellent impedance characteristics. In the L5E system a pair of such coaxial units can carry 13,200 telephone calls. Cables carrying 20 of these coaxial units are used in L-carrier systems (**Fig. 2**), and a 22-unit coaxial

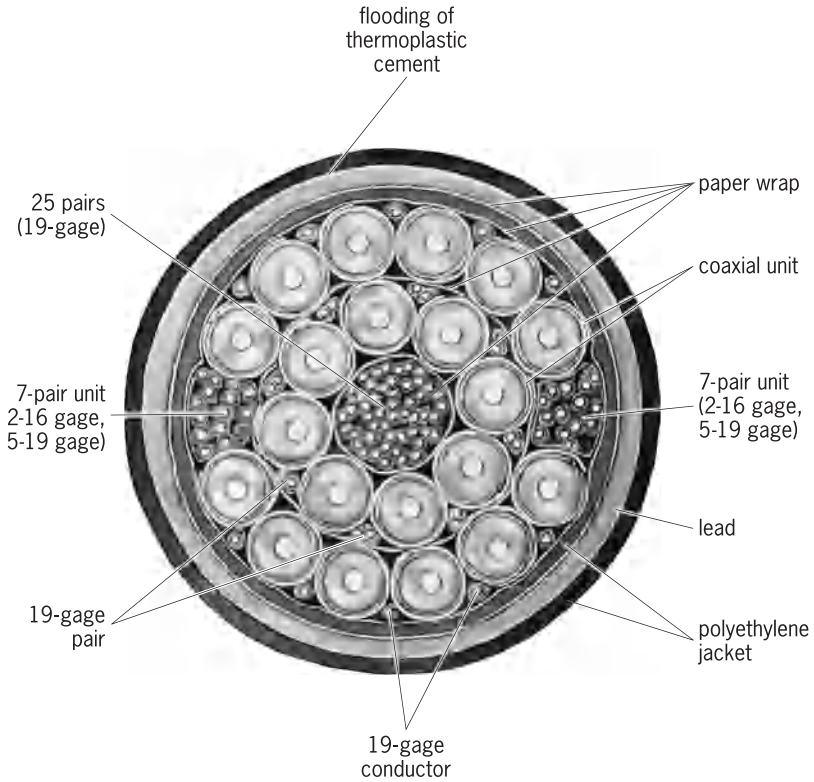

Fig. 2. Construction of multicoaxial transmission line with twenty 0.375-in. (9.5-mm) coaxial units.

version of this cable was used in later L5E installations. Fully equipped, this cable provides 132,000-telephone-channel capacity over the 10 working coaxial pairs. The eleventh pair provides service protection for the 10 working pairs.

Two coaxial cables span the United States from east to west, with many intermediate and shorter installations. For current and future use, optical fiber cables are playing the key role earlier played by coaxial cables.

Undersea coaxials. To function reliably in a high-performance system in the ocean environment, undersea coaxials must meet severe requirements. Some of these are (1) stable dimensions and predictable transmission characteristics under pressures of up to 11,000 lb/in.^{2} (76 megapascals); (2) 25-year service lifetime; and (3) ability to withstand handling and scuffing. A solid polyethylene core withstands ocean pressure which would crush a cable with any air fill. The use of inert materials that resist attack by marine organisms assures the desired lifetime. The high-strength (300,000-lb/in.^{2} or 2-gigapascal) steel strand and high-density jacket assure survival and continuity despite severe handling.

In coaxial cables used in fourth-generation analog undersea cable systems, the high-tensile steel strand provides longitudinal strength needed to lay or recover the cable in ocean depths of up to 4 nautical miles (7 km). The inner conductor is a high-conductivity copper tube, 0.48 in. (12 mm) in diameter, which is formed and swaged over the steel strand. Low-loss polyethylene is extruded over the inner conductor assembly to form the 1.7-in.-diameter (43.2-mm) cable core. To achieve an accurate (±0.001 in. or 0.025 mm) core diameter, the core is extruded slightly oversize and mechanically shaved.

In a further manufacturing operation, the 10-mil-thick (0.010-in. or 0.25-mm) copper outer conductor is formed over the cable core with an overlapped seam. A high-density polyethylene jacket, 2 in. (50 mm) in diameter, is extruded over the outer conductor to complete the cable. *See* COMMUNICATIONS CABLE; SUBMARINE CABLE.

S. Theodore Brewer

Cobalt

A lustrous, silvery-blue metallic chemical element, Co, with an atomic number of 27 and an atomic weight of 58.93. Metallic cobalt was isolated in 1735 by the Swedish scientist G. Brandt, who called the impure metal cobalt rex, after the ore from which it was extracted. The metal was shown to be a previously unknown element by T. O. Bergman in 1780.

Cobalt is a transition element in the same group as rhodium and iridium. In the periodic table it occupies a position between iron and nickel in the third period. Cobalt resembles iron and nickel in both its free and combined states, possessing similar tensile strength, machinability, thermal properties, and electrochemical behavior. Constituting 0.0029% of the

1												13	14	15	16	17	18
1 H	2																2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Earth's crust, cobalt is widely distributed in nature, occurring in meteorites, stars, lunar rocks, seawater, fresh water, soils, plants, and animals. *See* PERIODIC TABLE; TRANSITION ELEMENTS.

Cobalt and its alloys resist wear and corrosion even at high temperatures. The most important commercial uses are in making alloys for heavy-wear, high-temperature, and magnetic applications. Small amounts of the element are required by plants and animals. The artificially produced radioactive isotope of cobalt, ^{60}Co, has many medical and industrial applications.

Cobalt, with a melting point of 1495°C (2723°F) and a boiling point of 3100°C (5612°F), has a density (20°C; 68°F) of 8.90 g·cm^{-3}, an electrical resistivity (20°C) of 6.24 microhm·cm, and a hardness (diamond pyramid, Vickers; 20°C) of 225. It is harder than iron and, although brittle, it can be machined. The latent heat of fusion is 259.4 joules/g, and the latent heat of vaporization is 6276 J/g; the specific heat (15–100°C; 59–212°F) is 0.442 J/g · °C. Cobalt is ferromagnetic, with the very high Curie temperature of 1121°C (2050°F). The electronic configuration is $1s^2 2s^2 2p^6 3s^2 3p^6 3d^7 4s^2$. At normal temperatures the stable crystal form of cobalt is hexagonal close-packed, but above 417°C (783°F) face-centered cubic is the stable structure. Although the finely divided metal is pyrophoric in air, cobalt is relatively unreactive and stable to oxygen in the air, unless heated. It is attacked by sulfuric, hydrochloric, and nitric acids, and more slowly by hydrofluoric and phosphoric acids, ammonium hydroxide, and sodium hydroxide. Cobalt reacts when heated with the halogens and other nonmetals such as boron, carbon, phosphorus, arsenic, antimony, and sulfur. Dinitrogen, superoxo, peroxo, and mixed hydride complexes also exist. In its compounds, cobalt exhibits all the oxidation states from −I to IV, the most common being II and III. The highest oxidation state is found in cesium hexafluorocobaltate(IV), Cs_2CoF_6, and a few other compounds.

There are over 200 ores known to contain cobalt; traces of the metal are found in many ores of iron, nickel, copper, silver, manganese, and zinc. However, the commercially important cobalt minerals are the arsenides, oxides, and sulfides. Zaire is the chief producer, followed by Zambia. Russia, Canada, Cuba, Australia, and New Caledonia produce most

of the rest. Zaire and Zambia together account for just over 50% of the world's cobalt reserves. Nickel-containing laterites (hydrated iron oxides) found in the soils of the Celebes, Cuba, New Caledonia, and many other tropical areas are being developed as sources of cobalt. The manganese nodules found on the ocean floor are another large potential reserve of cobalt. They are estimated to contain at least 400 times as much cobalt as land-based deposits. *See* MANGANESE NODULES.

Since cobalt production is usually subsidiary to that of copper, nickel, or lead, extraction procedures vary according to which of these metals is associated with the cobalt. In general, the ore is roasted to remove stony gangue material as a slag, leaving a speiss of mixed metal and oxides, which is then reduced electrolytically, reduced thermally with aluminum, or leached with sulfuric acid to dissolve iron, cobalt, and nickel, leaving metallic copper behind. Lime is used to precipitate iron, and sodium hypochlorite is used to precipitate cobalt as the hydroxide. The cobalt hydroxide can be heated to give the oxide, which in turn is reduced to the metal by heating with charcoal.

Cobalt ores have long been used to produce a blue color in pottery, glass, enamels, and glazes. Cobalt is contained in Egyptian pottery dated as early as 2600 B.C. and in the blue and white porcelain ware of the Ming Dynasty in China (1368–1644).

An important modern industrial use involves the addition of small quantities of cobalt oxide during manufacture of ceramic materials to achieve a white color. The cobalt oxide counteracts yellow tints resulting from iron impurities. Cobalt oxide is also used in enamel coatings on steel to improve the adherence of the enamel to the metal. Cobalt arsenates, phosphates, and aluminates are used in artists' pigments, and various cobalt compounds are used in inks for full-color jet printing and in reactive dyes for cotton. Cobalt blue (Thenard's blue), one of the most durable of all blue pigments, is essentially cobalt aluminate. Cobalt linoleates, naphthenates, oleates, and ethylhexoates are used to speed up the drying of paints, lacquers, varnishes, and inks by promoting oxidation. In all, about a third of the world's cobalt production is used to make chemicals for the ceramic and paint industries. *See* CERAMICS; DYE; INK; METAL COATINGS.

Cobalt catalysts are used throughout the chemical industry for various processes. These include hydrogenations and dehydrogenations, halogenations, aminations, polymerizations (for example, butadiene), oxidation of xylenes to toluic acid, production of hydrogen sulfide and carbon disulfide, carbonylation of methanol to acetic acid, olefin synthesis, denitrogenation and desulfurization of coal tars, reductions with borohydrides, and nitrile syntheses, and such important reactions as the Fisher-Tropsch method for synthesizing liquid fuels and the hydroformylation process. Cobalt catalysts have also been used in the oxidation of poisonous hydrogen cyanide in gas masks and in the oxidation of carbon monoxide in automobile exhausts. *See* CATALYTIC CONVERTER.

Although cobalt was not used in its metallic state until the twentieth century, the principal use of cobalt is as a metal in the production of alloys, chiefly high-temperature and magnetic types. Superalloys needed to stand high stress at high temperatures, as in jet engines and gas turbines, typically contain 20–65% cobalt along with nickel, chromium, molybdenum, tungsten, and other elements.

In parts of the world where soil and plants are deficient in cobalt, trace amounts of cobalt salts [for example, the chloride and nitrate of Co(II)] are added to livestock feeds and fertilizers to prevent serious wasting diseases of cattle and sheep, such as pining, a debilitating disease especially common in sheep. Symptoms of cobalt deprivation in animals include retarded growth, anemia, loss of appetite, and decreased lactation.

The principal biological role of cobalt involves corrin compounds (porphyrin-like macrocycles). The active forms contain an alkyl group (5′-deoxyadenosine or methyl) attached to the cobalt as well as four nitrogens from the corrin and a nitrogen from a heterocycle, usually 5,6-dimethylbenzimidazole. These active forms act in concert with enzymes to catalyze essential reactions in humans. However, the corrin compounds are not synthesized in the body; they must be ingested in very small quantities. Vitamin B_{12}, with cyanide in place of the alkyl, prevents pernicious anemia but is itself inactive. The body metabolizes the vitamin into the active forms. Although the cobalt in corrins is usually Co(III), both Co(II) and Co(I) are involved in enzymic processes. Roughly one-third of all enzymes are metalloenzymes. Cobalt(II) substitutes for zinc in many of these to yield active forms. Such substitution of zinc may account, in part, for the toxicity of cobalt. *See* ENZYME; VITAMIN B_{12}.

L. Marzilli; Patricia A. Marzilli

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; A. Earnshaw and N. Greenwood, *Chemistry of the Elements*, 2d ed., Butterworth-Heinemann, 1997; J. A. McCleverty and T. J. Meyer (eds.), *Comprehensive Coordination Chemistry II*, 2d ed., Elsevier Science, 2003.

Coca

Shrubs of the genus *Erythroxylum* in the coca family (Erythroxylaceae). The genus contains over 200 species, most in the New World tropics. Two species, *E. coca* (*see* **illus.**) and *E. novogranatense*, are cultivated for their content of cocaine and other alkaloids. Leaves of cultivated coca contain 14 alkaloids, chiefly cocaine, and significant amounts of vitamins and minerals. The cocaine content averages 0.5%, rarely exceeding 1.0%.

Millions of South American Indians "chew" coca every day as a mild stimulant by placing dried leaves in the mouth and masticating with various alkalies such as lime or plant ashes to promote the release of the alkaloids, which are swallowed. Coca is

Flowering branch of coca (*Erythroxylum coca*).

important in the folk medicine of Andean Indians, especially as a treatment for gastrointestinal disorders. It may regulate carbohydrate metabolism in a useful way and supplement an otherwise deficient diet. Coca also has great magical and religious significance in Andean societies. Although it has long been the focus of political debate, especially in Peru, ritualized use of the leaf in traditional cultures is not inconsistent with good physical health and social productivity. *See* COCAINE; COLA; LINALES.

Andrew T. Weil

Bibliography. R. T. Martin, The role of coca in the history, religion, and medicine of South American Indians, *Econ. Bot.*, 24(4):422–438, 1978; A. T. Weil, Coca leaf as a therapeutic agent, *Amer. J. Drug Alcohol Abuse*, 5(1):75–86, 1978.

Cocaine

The principal alkaloid of coca leaves, a topical anesthetic and stimulant, and popular illicit drug. The molecular formula is given below. Cocaine was first

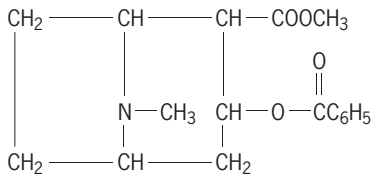

isolated by A. Niemann in 1860. In 1884 C. Koller demonstrated its efficacy as an anesthetic in eye surgery, introducing the age of local anesthesia. For the next decade cocaine enjoyed the status of a wonder drug and panacea. It fell into disfavor with increasing reports of acute toxicity and long-term dependence. Today it is used as a topical anesthetic in the eye, nose, mouth, and throat; for injection anesthesia it has been replaced by synthetic drugs with fewer central nervous system effects. *See* ANESTHESIA; COCA.

Cocaine increases heart rate and blood pressure and causes feelings of alertness and euphoria. It stimulates noradrenergic pathways by blocking reuptake of norepinephrine at the synapse, and also stimulates dopaminergic pathways. In animals it can be shown to be strongly reinforcing. It does not produce physical dependence, as alcohol and opiates do, but many people find it hard to use in a stable and moderate fashion if they have access to it in quantity. Although it is quite active orally, most users of illicit cocaine take it intranasally by snuffing; few inject it intravenously. Aside from local irritation of the nasal membranes, occasional users suffer few adverse effects. The psychological consequences of frequent use can be devastating. The soluble hydrochloride salt is the common form. Insoluble cocaine free base, that is, the free alkaloid form of the drug, may be smoked, a practice that increases the toxicity of the drug and is more likely to be associated with dependence. *See* ADDICTIVE DISORDERS; ALKALOID; NORADRENERGIC SYSTEM.

Andrew T. Weil

Bibliography. M. S. Gold, *Cocaine*, 1993; A. T. Weil and W. Rosen, *From Chocolate to Morphine: Everything You Need to Know about Mind-Altering Drugs*, revised, 1998; R. D. Weiss, S. M. Mirin, and R. L. Bartel, *Cocaine*, 2d ed., 1993.

Coccidia

A subclass of the class Telosporea. These protozoa are typically intracellular parasites of epithelial tissues in both vertebrates and invertebrates. The group is divided on the basis of life cycles into two orders, the Protococcida (in which there is only sexual reproduction) and the Eucoccida (in which there is both sexual and asexual reproduction). There are only a few species in Protococcida, and all are parasites in marine invertebrates. There are hundreds of species in Eucoccida; these are parasites in both invertebrates and vertebrates. The Eucoccida contains three suborders: Adeleina, which occur mostly in invertebrates and in which there is usually one host for the sexual stages and a different host for the asexual stages (an example is *Hepatozoon muris*, found in rats and transmitted by mites); Haemosporina, which occur mainly in vertebrates and which have their sexual stages in invertebrates (an example is *Plasmodium falciparum* which causes malaria in humans and is transmitted by mosquitoes); and Eimerina, which occur mostly in vertebrates and in which the sexual and asexual stages occur in the same host (an example is *Eimeria*, species of which cause coccidiosis in many domestic animals and wildlife, and which are transmitted by filth). The malaria parasites and coccidia are of direct importance to humans, and also cause important losses in chickens, cattle, sheep, and other domesticated animals. *See* EUCOCCIDA; PROTOCOCCIDA; PROTOZOA; SPOROZOA; TELOSPOREA.

Elery R. Becker; Norman D. Levine

Coccolithophorida

A group of unicellular, biflagellate, golden-brown algae characterized by a covering of extremely small (1–35 micrometers) interlocking calcite (the hexagonal form of calcium carbonate) plates called coccoliths. The plates show extinction crosses in a polarizing light microscope; however, detailed study requires the use of electron transmission or electron scanning microscopes. The Coccolithophorida are usually considered plants but possess also some animal characteristics. Botanists assign them to the class Haptophyceae (based on the possession of a haptonema, a threadlike organ of attachment) of the phylum Chrysophyta, and zoologists to the class Phytamastigophorea, superclass Mastigophora, subphylum Sarcomastigophora of the phylum Protozoa. *See* CILIA AND FLAGELLA; PHYTAMASTIGOPHOREA; PROTOZOA.

Classification. The Coccolithophorida are mainly classified by the shape of their coccoliths into two groups: the holococcoliths, with simple rhombic or hexagonal crystals arranged like a mosaic; and the heterococcoliths, with complex crystals arranged into boat, trumpet, basket, or collar-button shapes. Classification is complicated by the presence of polymorphism (more than one type of coccolith per cell) and complex life cycles involving alternation of generations either with different types of coccoliths or with none at all (see **illus.**).

Life cycle. Reproduction occurs mostly by fission. Haploid and diploid phases, motile versus nonmotile phases, resting spores, and ameboid and filamentous stages have been noted but their interrelationships require further study.

Ecology. Chiefly photosynthetic, although epiphytism, phagotrophism, and saprophytism have also been shown, the Coccolithophorida form a significant percentage of the nannoplankton (plankton of less than 74-μm size) of the tropic through subarctic-subantarctic waters of all oceans; a few brackish and fresh-water species also exist. Together with the diatoms they constitute the primary producers of the open ocean food chain. *See* BACILLARIOPHYCEAE; PHYTOPLANKTON.

Coccoliths. Ease of culture makes these organisms a prime choice for studying biologic calcification. Research shows that coccoliths are extruded to the cell surface after being formed intracellularly on an organic matrix located between the nucleus and a "reticular body," which plays some yet undiscovered but essential role, and that the process is light-dependent and linked to photosynthesis. Coccolith formation has been introduced in naked cells by lowering the nitrogen content of the culture medium, and coccolith-bearing cells have lost their covering with the addition of carbon dioxide to their medium.

The function of coccoliths is unknown, but one worker suggests that they serve as a shield against excessive sunlight.

Fossil record. Coccoliths preserve well and have a fossil record dating back into the Jurassic, 180,000,000 years ago. They make up limestone and chalk deposits such as the White Cliffs of Dover on land and up to 30% of the sediments underlying tropical to arctic-antarctic ocean waters.

Their long and involved evolutionary record makes them useful for dating ancient sediments. The onset of the Pleistocene, or glacial, period is marked by the appearance of new coccolith species and

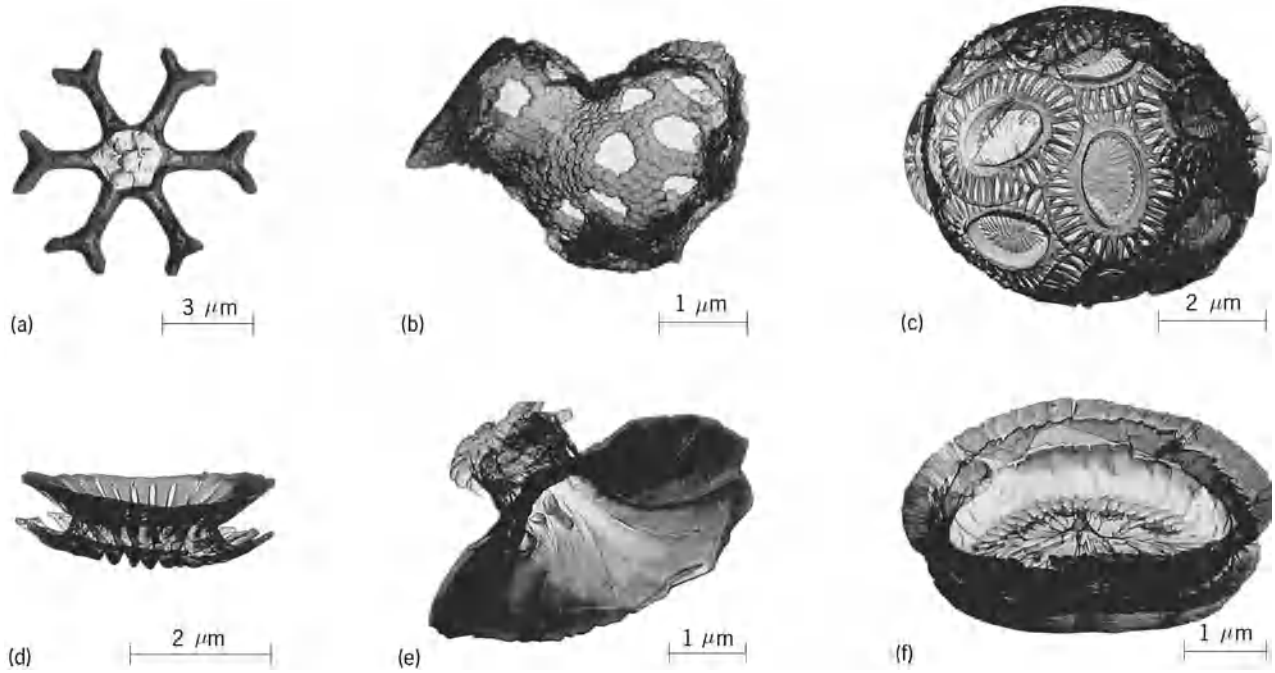

Examples of Coccolithophorida. (*a*) *Discoaster*; (*b*) a holococcolith; (*c*) a coccosphere; (*d*) a coccolith; (*e, f*) heterococcoliths. (*A. McIntyre, Lamont-Doherty Geological Observatory of Columbia University*)

the disappearance of older forms, including the discoasters, a group of star-shaped members prominent during the Tertiary. Extinction of the latter in a sedimentary sequence is a classic marker of the Pliocene-Pleistocene boundary.

Because they live in surface waters, coccoliths are under direct climatic control. Thus many modern species with relatively long fossil records are excellent temperature indicators. Research indicates that these can be used to unravel the climatic variations of the Pleistocene period that evolved into present weather patterns. *See* MICROPALEONTOLOGY.

Andrew McIntyre

Bibliography. M. Black, *Endeavor*, 24:131–137, 1965; W. Fairbridge (ed.), *The Encyclopedia of Oceanography*, 1966; A. McIntyre, *Science*, 158: 1314–1317, 1967; A. McIntyre and A. W. H. Be, *Deep-Sea Res.*, 14:561–597, 1967; R. E. Round, *The Biology of the Algae*, 1965; E. J. F. Wood, *Marine Microbial Ecology*, 1965.

Cocoa powder and chocolate

Products derived from the seeds of the tropical tree *Theobroma cacao*, which grows within a narrow belt along both sides of the Equator. There are three basic varieties: Criollo, which was native to Central and South America; Forastero, which constitutes the bulk of the world supply and is mostly cultivated in West Africa and Brazil; and Trinitario, consisting of various hybrids. *See* CACAO.

The seeds, contained in pods which grow on the trunks and lower branches of the trees, are surrounded by a mucilaginous pulp. They are whitish in color and have no normal chocolate flavor. Each mature bean is about $^3/_4$–1 in. (2–2.5 cm) in length with an average weight not less than 0.04 oz (1 g). The shell (testa) is about 12% of the total bean weight, the nib (cotyledon) 87%, and germ 1%—the last being the point at which germination would have commenced in the live bean. The main constituents of the dried nib are approximately the following: fat (cocoa butter) 54%; protein 12–14%; starch, other carbohydrates, and mucilage 20%; tannins 6%; theobromine 1.3%; caffeine 0.2%; and moisture 3–5%. Good-quality beans show a brown or purplish-brown color and striations in the sliced cotyledon. Unfermented beans have a dense structure and slaty color. *See* SEED.

Processing. To develop a chocolate flavor, the beans are first subjected to a fermentation process. The pods are split open and the seeds and pulp placed in heaps surrounded by plantain leaves or in "sweat boxes." Under these conditions microbiological and enzymatic action occurs, with a rise in temperature to about 120°F (50°C).

After 4–6 days, during which the beans are turned over, they develop a purplish-brown color, and the outer shell and nib become more distinct. The pulp liquefies and drains away, and precursors of the true chocolate flavor are formed. After fermentation the beans are sun-dried in shallow trays that are protected from rainstorms by sliding canopies; in very wet climates, such as that found in Cameroon, various forms of artificial drying are used. After drying, the product is described commercially as cocoa or cocoa beans and should have a moisture content of 6–7%.

A series of well-defined operations are used to produce good-quality cocoa or chocolate.

Cleaning. Machines, by a combination of sieving and air elutriation, remove stones, sand, bag fiber, stalks, and immature beans. Without this treatment, grit would pass to subsequent processes, resulting in machine damage and poor-quality products.

Roasting. In this essential process the chocolate flavor is fully developed from the precursors, the shell loosened from the nib, and the moisture content reduced.

There are many roaster designs. One type is a rotating drum with external heating, which may handle single batches or have a continuous throughput. Other roasters use high-temperature air passing through a continuous stream of beans. Roasting temperatures vary greatly, ranging from 200 to 300°F (95 to 150°C). The temperature used is related to the time of roasting and the type of product required. The shell should never be burned.

In two-stage roasting, the beans are heated at a low temperature in a closed, moist atmosphere, and then a higher temperature and drying follow.

Winnowing. After roasting, the winnowing process separates the shell from the nib. The ideal is complete separation with full recovery of the nib, but in practice this is not possible. Winnowing machines have been greatly improved. A modern system uses a bank of vibrating sieves with an air lift to remove shell pieces at the discharge end of each sieve (**Fig. 1**). At the machine entry the beans are broken by hexagonal impact rollers which minimize the production of fine dust. The shell is a waste product that is used as a boiler fuel and an agricultural mulch.

Liquor milling. The roasted nibs are ground at a temperature above the melting point of the cocoa butter constituent (93–95°F or 34–35°C) to produce a dark brown liquid called liquor, mass, or unsweetened chocolate. Finely ground liquor is the basis for manufacture of cocoa powder and chocolate.

Different machines are used for this purpose. Pin and hammer mills in conjunction with roller refiners are frequently used. Other machines include horizontal disk mills with three stages and the vertical ball mill. In the ball mill partially ground liquor is fed through a vertical cylinder containing hard steel or ceramic balls rotated at high speed by means of a central spindle and attached plates.

Cocoa powder manufacture. To manufacture cocoa powder, liquor is subjected to hydraulic pressing, which separates some of the cocoa butter from the solid cocoa matter. The pressing machine consists of a number of pots arranged horizontally, each provided with a filter and linked to a hydraulic ram capable of providing pressures of over 6000 lb/in.^{2} (10^8 pascals). The pots are filled automatically with hot liquor, and when the ram operates liquid cocoa

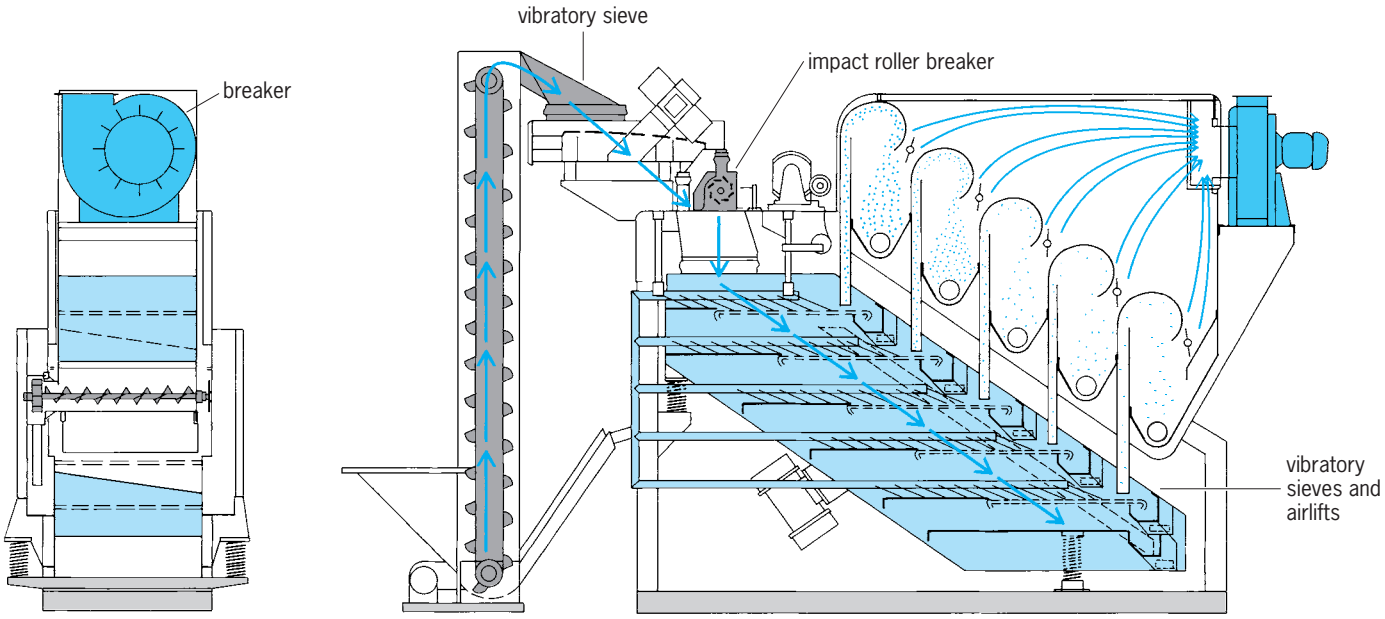

Fig. 1. Two views of winnowing machine. (*Bauermeister Maschinen Fabrik, Hamburg, Germany*)

butter is forced through the filters and into a common channel at the base of the press. The amount of cocoa butter pressed out is determined by control of time and pressure.

At the end of the pressing cycle the ram is reversed, thus opening the pots and releasing compressed cocoa cakes. To make cocoa powder, these cakes are first mechanically broken into small pieces called kibbled cake and then put through pulverizers which, in conjunction with air separators, produce a very fine powder. Such cocoa powder is called natural cocoa.

Additionally, if dark brown or reddish-brown cocoas are required, the liquor, kibbled cake, or preferably the nibs are subjected to a process called alkalization or dutching. This consists of saturating the nibs (or other intermediate) with a solution of potassium or sodium carbonate followed by drying and roasting. In most countries, the amount of added alkali permitted is limited to 3% of the nib.

Color is controlled by the degree of roast of the nib and the quantity and concentration of the alkali used. Alkalization also serves to neutralize the natural acidity of the nib.

Commercial cocoa powders may have a residual cocoa butter content of 10–22%. They are used extensively as ingredients in cookies, biscuits, chocolate syrups and spreads, and vegetable fat coatings.

Chocolate manufacture. Dark, bitter, or sweet chocolate is manufactured from liquor (or nibs), sugar, and cocoa butter—the cocoa butter being obtained from cocoa powder manufacture. Cocoa butter is a very stable, natural fat; it melts just below body temperature and has a narrow melting range. It displays good contraction on solidifying from the liquid state, which enables chocolate to be molded into various shapes. Its melting properties impart good texture in the mouth and flavor release.

Milk chocolate is made from liquor, sugar, milk solids, and cocoa butter. The milk solids are derived from liquid milk, usually by a spray-drying process. Another milk product often used, called milk crumb, is prepared by concentrating and drying liquid milk in the presence of sugar and liquor. This gives a caramelized milk flavor which has proved very popular. *See* MILK.

Refining. In the manufacture of chocolate, the solids are mixed mechanically with liquor and some cocoa butter to give a plastic paste. This paste is fed to refiners, which consist of five steel rolls mounted vertically with the bottom feed roll offset to enable the paste to be delivered continuously from the mixers. The rolls work on a system in which the speed of rotation increases from the bottom to the top roll, thereby obtaining a grinding shear between the roll surfaces and enabling the paste to travel upward from one roll to the next. The refined paste is scraped off the top roll by a sharp blade. In some instances refining is done in two stages. **Figure 2** shows a mixing and double refining system.

Conching. The final process in chocolate making is called conching, which removes moisture and volatile substances that would produce acid and harsh flavors. Conching also reduces viscosity and contributes a smoothness to the eating qualities of the chocolate.

At one time conching was carried out in machines which consisted of a heated tank with a flat bottom over which a roller traveled backward and forward through the liquid chocolate. While these were effective, much time was necessary to produce good chocolate; the process also took much energy, and extra cocoa butter was needed to give fluidity.

Later it was shown that with more powerful machinery it was possible to directly conche the friable

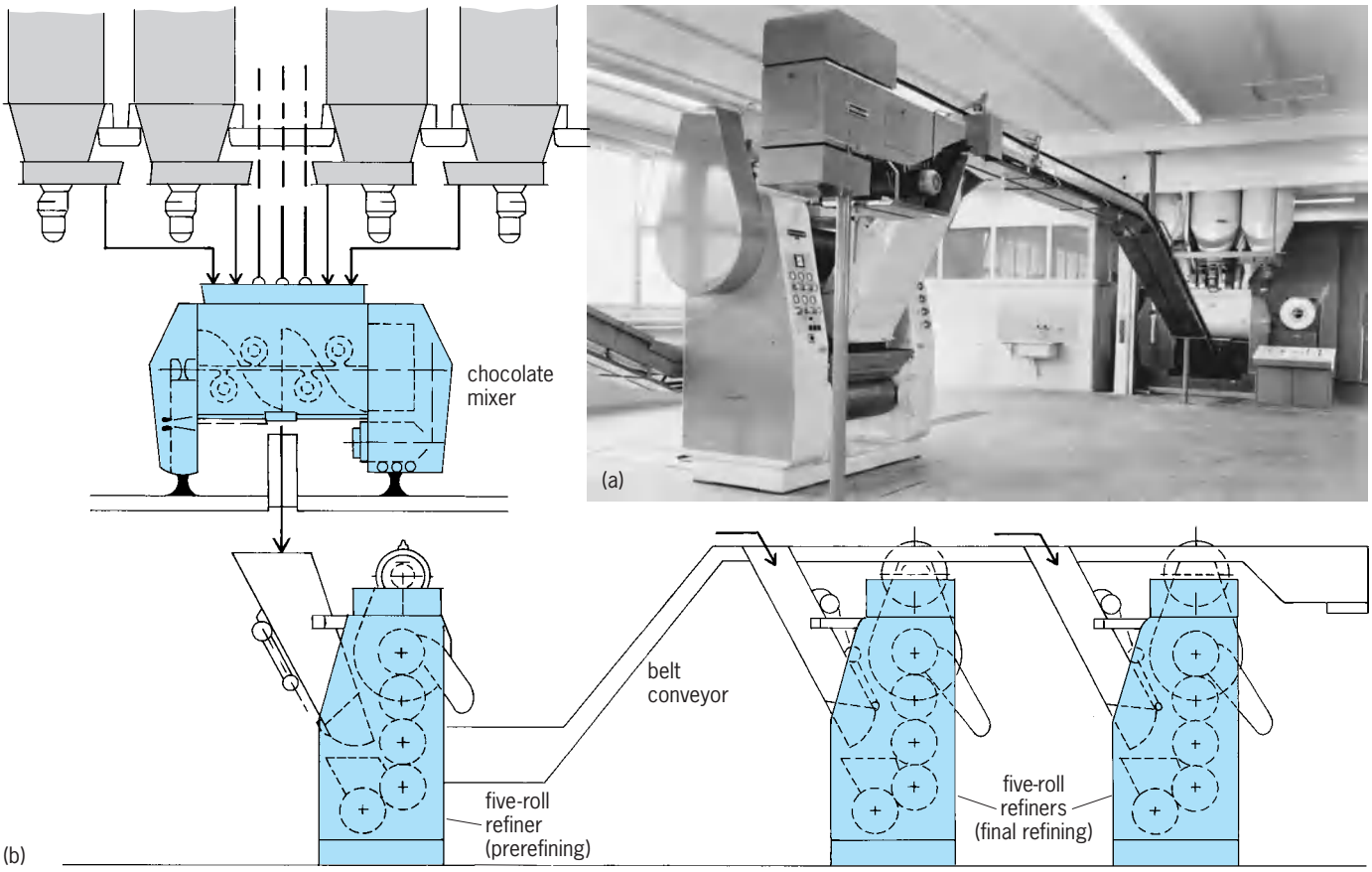

Fig. 2. Modern chocolate-making plant. (*a*) Mixer, conveyor, and refining rolls. (*b*) Diagram of operation of this equipment. (*Buhler-Miag*, *Uzwil*, *Switzerland*)

flake from the refining rolls. This is known as dry conching, a process which removes the volatiles in a much shorter time. Extra fluidity is obtained by mechanical mixing alone. Dry conching is now mostly done in rotary conches (**Fig. 3**), taking about 4 h. Then extra cocoa butter is added and the speed of rotation increased for a further period of 6–12 h depending on the type of chocolate desired.

Fig. 3. Modern rotary conche. (*Petzholdt*, *Frankfurt*, *Germany*)

Besides mechanical improvements, the use of emulsifiers has reduced costs by obtaining the correct fluidity with less cocoa butter. The universally used emulsifier is natural lecithin obtained from the soya bean; in most countries the amount permitted by law is limited to 0.5% of the final chocolate product. About 1 h before the end of the conching period, the emulsifier is added, and the viscosity is checked and adjusted by the addition of more cocoa butter. The final viscosity depends on whether the chocolate is to be used for molding bars or coating confectionery centers.

Tempering. In both refining and conching, the chocolate must be subjected to tempering before it is applied to confectionery centers or deposited into molds. Cocoa butter is polymorphic, which means it crystallizes in several forms on solidifying. Some forms are unstable, and these will transform to the stable form slowly if the chocolate is not tempered, resulting in an unsightly grayish surface film called bloom.

To temper chocolate it is necessary to seed the liquid chocolate with the stable form of crystal. There are several ways of doing this. The most common way is to reduce the temperature of the liquid chocolate from about 100 to 84°F (38 to 29°C), with constant stirring, until cocoa butter seed is formed, and then raising the temperature again to 90–91°F (32–33°C), which will melt out the remaining unstable

forms and render the chocolate sufficiently fluid for application. Modern machines called enrobers or coaters carry out this process. *See* FOOD MANUFACTURING. Bernard W. Minifie

Bibliography. S. T. Beckett (ed.), *Industrial Chocolate Manufacture and Use*, 3d ed., 1999; B. W. Minifie, *Chocolate, Cocoa, and Confectioner Science and Technology*, 3d ed., updated 1999.

Coconut

A large palm, *Cocos nucifera* (in the family Areacaceae), widely grown throughout the tropics and valuable for its fruit and fiber. Usually found near the seacoast, it requires high humidity, abundant rainfall (60 in. or 1.5 m), and mean annual temperature of about 85°F (31°C). Southern Florida, with mean temperature of 77°F (25°C), is at the limit of successful growth. The origin of the coconut has been in dispute, but strong evidence points to southern Asia with wide dispersal by ocean currents and human migrations. *See* ARECALES.

The fruit, 10 in. (25 cm) or more in length, is ovoid and obtusely triangular in cross section (**Fig. 1**). The tough, fibrous outer husk (exocarp) encloses a spherical nut consisting of a hard, bony shell (endocarp) within which is a $^1/_2$-in. (1.25-cm) layer of fleshy meat or kernel (endosperm). The meat is high in oil and protein and, when dried, is the copra of commerce.

Production and culture. Principal commercial producers are the Philippines, Indonesia, Malaya, Ceylon, and Oceania. Although many trees grow along the seashore and in native villages without special care, the crop lends itself to plantation culture with control of weeds, fertilization, and protection from diseases, insects, and animal pests.

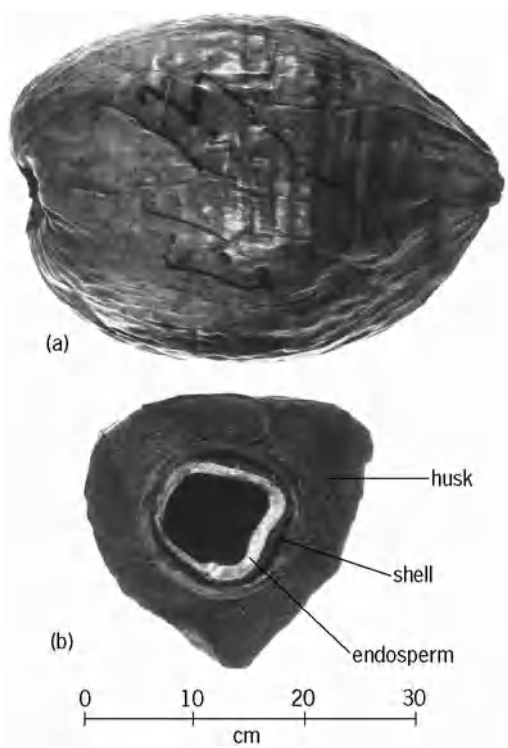

Fig. 1. Coconut. (*a*) Fruit. (*b*) Cross section of fruit showing husk, shell, and endosperm (meat). 2.5 cm = 1 in.

Palms begin to bear nuts the sixth year after planting and reach full bearing about the eighth year. Individual nuts mature about a year after blossoming and normally fall to the ground. In plantation culture, clusters of mature nuts are cut from the trees with knives on long poles just before maturity. Picking individual nuts by climbing the trees is sometimes practiced.

Processing and products. The thick husks are removed from the nuts, which are then split open with a heavy knife and partially dried to loosen the meat. This is pried out of the shell and dried to about 7% moisture either in the sun, if conditions are favorable, or in kilns with additional heat often furnished by burning the husks and shells.

The oil from the dried coconut meats, or copra, is widely used for margarine, soap, and industrial purposes. High-quality copra may be shredded for confectionery and the baking trade. The residue, after oil removal, is used for animal feed.

Coconut husks are an important source of fiber called coir. In coir production, mostly limited to India, the nuts are harvested about a month before maturity and the husks retted in brackish water for 8–10 days to rot away the soft tissues between the fibers. The fiber is cleaned and washed by hand and dried. Recently the industry has become partially mechanized. The various grades of coir are used for ropes, mats and matting, and upholstery filling. *See* COIR.

In the tropics of both the Orient and the Occident the coco palm is the most useful of all plants to the native population. An important source of food and drink, it also furnishes building material, thatch, hats, dishes, baskets, and many other useful items. *See* NUT CROP CULTURE. Laurence H. MacDaniels

Diseases. Cadang-cadang (**Fig. 2**), a disease confined to certain islands of the Philippines, is caused by a viroid. This same agent is also present in the African oil palm and in the Buri palm. No vector is yet known. Since the disease affects mature palms, control is based on replanting with early maturing palms.

Red ring is caused by the nematode *Rhadinaphelenchus cocophilus* and is confined to tropical America. Transverse sections of stems from diseased palms show a typical ring. The disease is always lethal, causing death shortly after infection. The nematode is spread by a palm weevil (*Rhyncophorus palmarum*) which in turn is attracted by damaged palm tissues. Control is effected through sanitation and by the application of insecticides.

Heart rot is confined to tropical America and is caused by a trypanosomatid flagellate protozoon of the genus *Phytomonas*. This organism also attacks the African oil palm. Many of the symptoms resemble lethal yellowing. No effective control is known.

Bud rot is worldwide and is caused by the fungus *Phytophthora palmivora*, which attacks the

Fig. 2. Midstage symptoms of cadang-cadang. Fronds assume an erect position in the crown, and no nuts are being produced.

growing point and causes it to rot. There is no effective control.

Kerala wilt is a debilitating disease of coconut palms in southern India. It results in production of fewer buttons and nuts, reduction of nut size, and thinner husks. The cause is unknown.

Blast and dry bud rot affect young palms in the nurseries and right after planting. Both diseases seem most serious in newly released hybrids. Mycoplasma organisms have been suggested as the cause of blast in the Ivory Coast, where the insect *Ricelia mica* is the vector. *See* PLANT PATHOLOGY.

Luigi Chiarappa

Bibliography. T. O. Diener, Viroids: The smallest known agents of infectious diseases, *Annu. Rev. Microbiol.*, 28:23–29, 1974; K. Maramorosch, *A Survey of Coconut Diseases of Unknown Etiology*, FAO, Rome, 1964; University of Florida, *Proceeding of the the 4th Meeting of the International Council on Lethal Yellowing*, 1979.

Coelacanthiformes

An order of lobefin fishes placed in the subclass Coelacanthimorpha and well known as fossils. Coelacanthiformes and the fossil suborders Rhizodontimorphi and Osteolepimorpha (rhipidistians) were once grouped as the subclass Crossopterygii. However, since it is now believed this assemblage is artificial, the coelacanth fraction was elevated to the subclass Coelacanthimorpha, with one order (Coelacanthiformes), four fossil families, and one extant family (Coelacanthidae).

Members are easily recognized by the two dorsal fins, by the paired pectoral and pelvic fins, and the anal fin, and by their symmetrical caudal fin with small central prolongation (**Fig. 1**). The only living fish with such features is *Latimeria chalumnae*; it is also the only extant fish with an intracranial joint, which otherwise occurs only in fossil rhipidistians. In 1952 its habitat was discovered to be the deep waters around the Comoro Islands. A new population of *Latimeria* was discovered in 1998 north of Sulawesi in the Indonesian archipelago.

Anatomy. The Coelacanthiformes (or Actinistia) are the only fishes with a special rostral organ, a deep postcoronoid in the lower jaw, a tandem double articulation of the lower jaw, a postspiracular bone, and an additional bone, the extracleithrum, in the shoulder girdle. Like lungfish, coelacanths lack the marginal upper jaw bone, the maxilla, and possess a short dentary. Except for two Devonian (*Miguashaia* and *Gavinia* with heterocercal tail) and a Carboniferous (*Allenypterus* with diphycercal tail) genus, coelacanths have a caudal fin with equal-sized upper and lower lobe of unbranched fin rays separated by an axial notochordal lobe (triphycercal tail).

On the snout are the three paired openings of the rostral organ in addition to the two paired openings of the nasal sacs (**Fig. 2**). The rostral organ has a central chamber from which three tubes diverge to both sides. It is an electroreceptive organ, unique in its shape and concentration of electroreceptors among fishes.

The lateral side of the lower jaw is formed by angular and splenial bones; a short toothed dentary lies above the splenial bone. The coronoid bones on the dorsal margin of the lower jaw and the prearticular on the inner side of the jaw carry small pointed teeth. The deep postcoronoid reaches high above the dorsal margin of the lower jaw and does not carry teeth. The posterior end of the lower jaw forms a double groove for the articulation (joint) of the double condyle (the rounded process at the end of a bone) of the quadrate and another groove for the symplectic bone (**Fig. 3**).

The posterior moiety (otico-occipital region) of the endocranium carries the brain with labyrinth (the hollow in the temporal bone); a long extension of the pituitary gland (pituitary foot) surpasses the articulation between posterior and anterior moieties and reaches to the base of the anterior. Below the brain cavity the notochord extends in a tube to the posterior end of the anterior moiety. A paired basicranial muscle connects ventrolaterally the anterior with the posterior moiety. The basicranial muscle together with jaw and gill arch muscles acts in the movement of the anterior moiety against the posterior (lifting the snout) and in the opening of the mouth (**Fig. 4**).

The shoulder girdle is composed of anocleithrum, cleithrum, clavicle, and extracleithrum (**Fig. 5**). An unpaired interclavicle unknown in *Latimeria* is

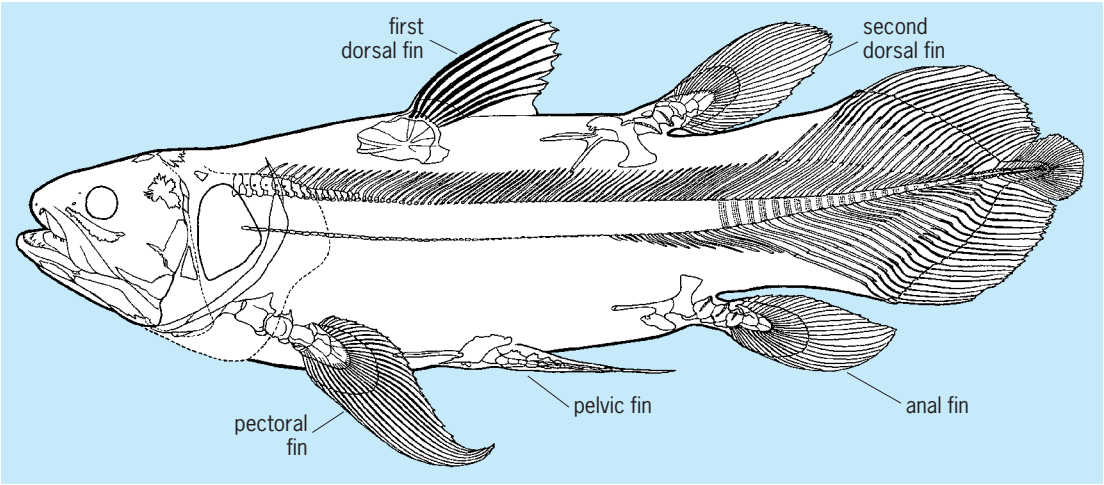

Fig. 1. Overall shape and internal structures of *Latimeria chalumnae*.

known from at least two fossil forms. The endoskeletal scapulocoracoid forms a knob on which the first mesomer of the axial endoskeletal elements of the pectoral fin articulates. Superficial muscles reach from the scapulocoracoid to the distal end of the axial skeleton and the bases of the fin rays; muscles connect internally the axial elements. This fleshy lobe is covered by small scales. The pelvic, second dorsal, and anal fins show the same structure. The first dorsal fin lacks axial elements and the fleshy lobe. The fin rays are articulated but not branched except for one Devonian genus. The exposed surface of the round scales is covered by ridges or tubercles converging toward the median posterior tip of the scales. The vertebral column is formed by an unrestricted notochord with neural arches and spines dorsally and with hemal arches and spines ventrally in the caudal region; ribs occur in some fossil forms. The notochordal fluid is considered by some Far Eastern people to be a life-prolonging elixir. The lung of *Latimeria* is an elongated oil organ well adapted to a deep-water existence. In contrast, fossil coelacanths were shallow-water dwellers with a much larger ossified lung. The intestine is formed as a scroll or longitudinal valve. The rectal gland is a salt-excreting organ which developed in parallel to that of cartilaginous fishes and serves for osmoregulation with high retention of urea. Coelacanths are carnivorous; fish are known to be stomach content in the extant *Latimeria* and in fossil forms.

Reproduction. *Latimeria* shows sexual dimorphism; the female is longer (measuring up to 1.8 m or 6 ft) than the male (up to 1.4 m or 5 ft). Reproduction is by internal fertilization even though the male has no intermittent organ. *Latimeria* has the largest eggs (9 cm or 3.6 in. in diameter) of any bony fish. It is ovoviviparous; up to 30 developing juveniles have been found in one female. Ovovivipary occurs also in Mesozoic coelacanths. Yolk-sac-carrying juveniles of comparatively small Paleozoic coelacanths have been found fossilized; these finds, in addition to the occurrence of unhatched specimens, indicate egg-laying early coelacanths.

Fossil record. Coelacanths are known since the Middle Devonian. They reach high diversity in the Early Triassic and Late Jurassic. Coelacanths acquired their common structure in the Carboniferous. The number of morphological changes is minor thereafter. There is no fossil record for the last 80 million years. The last known fossil coelacanths from the Upper Cretaceous of North America are also the largest known (estimated length of 3.5 m or 11 ft for *Megalocoelacanthus*). Fossil coelacanths are distributed worldwide; Paleozoic forms are known exclusively from the Northern Hemisphere, whereas Mesozoic forms have been found also in southern continents. Most fossil records refer to marine forms;

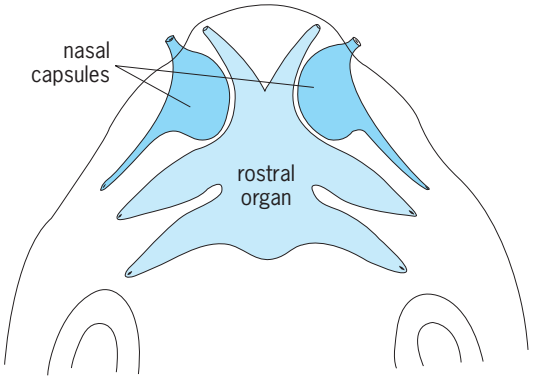

Fig. 2. Snout of *Latimeria*.

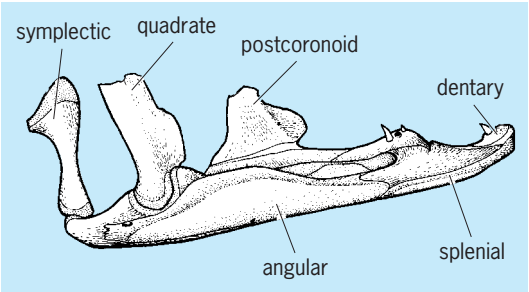

Fig. 3. Lower jaw of *Latimeria* with external bones.

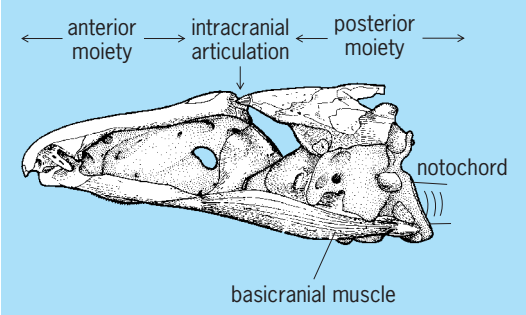

Fig. 4. Endocranium of *Latimeria*.

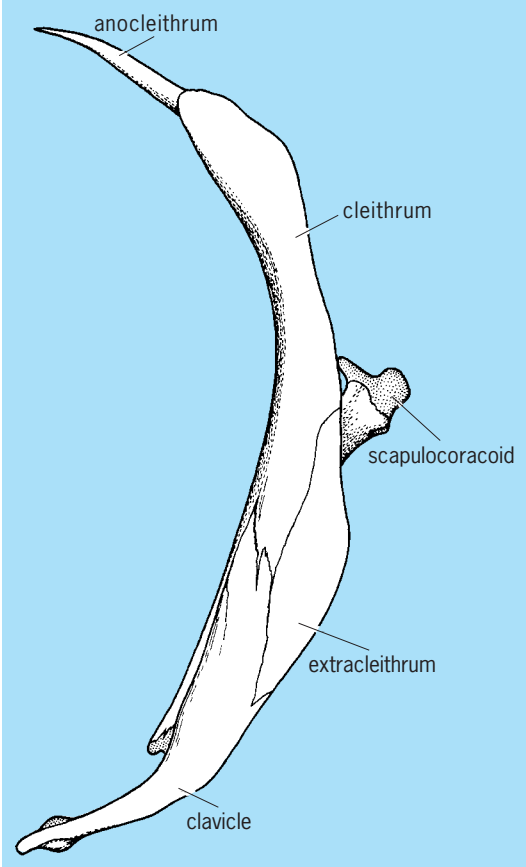

Fig. 5. Shoulder girdle of *Latimeria*.

nevertheless, some coelacanths were able to enter fresh water. *See* FOSSIL.

Relationship. The phylogenetic position of the coelacanthiforms within sarcopterygian fishes is debated. Within extant fishes, they may be closer to tetrapods than to the lungfish, or vice versa. Even the possibility that both coelacanthiforms and lungfish are more closely related to each other than each is to tetrapods is postulated. In any case, both are not close relatives of tetrapods; the latter are related to rhipidistian fishes, especially Elpistostegalia. The living coelacanth is not, as often written, a survivor of human ancestry. *See* OSTEICHTHYES; SARCOPTERYGII.

Hans-Peter Schultze

Bibliography. P. L. Forey, *History of the Coelacanth Fishes*, Chapman & Hall, London, 1998; J. A. Musick, M. N. Bruton, and E. K. Balon (eds.), *The Biology of Latimeria chalumnae and Evolution of Coelacanths,* vol. 32 of *Environmental Biology of Fishes*, Kluwer Academic, Dordrecht, 1991; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994.

Coelom

The coelom is the mesodermally lined body cavity of most animals above the flatworms and nonsegmented roundworms. Its manner of origin provides one basis for classifying the major higher groups.

Annelida. Annelids, arthropods, and mollusks have a coelom which develops from solid mesodermal bands. Within the trochophore larva of annelids, a single pole cell proliferates two strips of mesoblast lying on either side of the ventral midline. These bands subdivide transversely into bilateral solid blocks, the somites. Each somite then splits internally to form a hollow vesicle, the cavity of which is the coelom. The somitic vesicles of each side expand until their surfaces become apposed above and below the gut, thus forming its supporting mesenteries. The anteroposterior surfaces of contact between somites form the septa between segments of the body. In some of the annelids which have developed a parasitic mode of life, as the leeches, septa disappear between segments, and the coelom is much reduced by encroachment of mesenchymatous tissue. *See* ANNELIDA.

Mollusca. The mollusks also form bands of mesoderm from a single pole cell, but these bands do not segment. They split internally to form single right and left coelomic sacs, but the cavities are soon reduced and the surrounding mesoblast disperses as separate cells, many of which become muscle. The only remnants of the coelom in the adult are the pericardial cavity and the cavities of the gonads and their ducts. *See* MOLLUSCA.

Arthropoda. In arthropods paired bands of mesoblast may proliferate from a posterior growth center or may separate inward from a blastoderm, a superficial layer of cells, on the ventral surface of the egg. These bands divide into linear series of somites which then hollow out. Their cavities represent the coelom. The upper wall of each somite moves dorsally to meet its partner from either side and fuse into a longitudinal tube, the heart. The outer and inner somitic walls give rise to muscles of the body wall and gut, respectively. The lower wall may form fat body or move into the appendages to form excretory organs or glands. In adult arthropods the coelom persists only as small cavities within the segmental excretory organs or glands and as the cavities of the gonads and gonoducts. *See* ARTHROPODA.

Echinodermata and Chordata. Echinoderms and chordates constitute a second major group, characterized by the origin of the coelom from outpocketings of the primitive gut wall. In echinoderms one pair of bilateral pouches evaginates and separates

from the archenteron or primitive digestive cavity. Each pouch constricts into three portions, not homologous to the metameres of other animals. Of these the left middle coelom assumes major importance, giving rise to the water-vascular system and primordia of the radii of the adult. Contributions to the stone canal and axial sinus are made by the left anterior coelom. The rearmost cavities are converted into the main perivisceral coelom of the adult. *See* CHORDATA; ECHINODERMATA.

Protochordata. The protochordates of the groups Hemichordata and Cephalochordata exemplified by *Balanoglossus* and *Branchiostoma*, respectively, have three coelomic pouches formed by separate evaginations of the archenteral roof. The first or head cavity is single and median in origin, the other bilateral and paired. In hemichordates the head cavity remains single as the cavity of the proboscis and has a pore to the exterior on each side. The second pouches form cavities within the collar and also acquire external pores. The third pair is contained within the trunk and forms the major perivisceral cavity.

Cephalochordata. In cephalochordates the head cavity divides into lateral halves. The left side communicates, by a pore, to an ectodermal pit called the wheel organ. The second pair of pouches forms the pair of mesoblastic somites, and the third pouches subdivide transversely to give rise to the remainder of the linear series of somites. The upper or myotomic portion of each somite remains metameric and forms the segmental muscles. As it enlarges, the coelomic space is displaced ventrally and expands above and below the gut to form the perivisceral cavities and mesenteries, as described for annelids. However, the transverse septa perforate and disappear between adjacent somites, so that a continuous anteroposterior body cavity is formed on either side of the gut.

Vertebrata. In vertebrates the mesoderm arises as a solid sheet from surface cells that have been involuted through the blastopore. Lateral to the notochord, beginning at about the level of the ear, the mesoderm subdivides into three parts: (1) the somites; (2) the nephrotomic cord, temporarily segmented in lower vertebrates, which will form excretory organs and ducts; and (3) the unsegmented lateral plate. The coelom arises as a split within the lateral plate. In lower vertebrates the split may continue through the nephrotome into the center of the somite, thus resembling the condition in *Branchiostoma*. Such dorsal extensions are soon lost. The lateral plate expands to meet its partner of the other side, thus forming the body cavity and mesenteries of the gut. *See* GASTRULATION; INVERTEBRATE EMBRYOLOGY.

Howard L. Hamilton

Coelomycetes

Filamentous fungi that reproduce asexually by making spores inside microscopic fruiting bodies. Estimates of the number of genera range up to 1000, with the number of species between 7000 and 10,000. They are especially numerous on living or decomposing plants, but they can occur on several other substrata. Many are plant pathogens. A few have been isolated in clinical or veterinary settings. Although all members of the group reproduce asexually by making their spores (conidia) via mitosis, some species also have known sexual states (teleomorphs) which produce spores via mating and meiosis.

Taxonomy and identification. The group Coelomycetes was accorded formal status in traditional fungal taxonomy, but it has long been recognized that this group is artificial (nonphylogenetic, that is, encompassing multiple evolutionary lineages). Most members are asexual states of fungi traditionally classed as ascomycetes (Ascomycota), and a very few are asexual states of basidiomycetes (Basidiomycota). Some mycologists now prefer to write the name in lowercase (coelomycetes) to accentuate the nonphylogenetic nature of the group. The coelomycetes are one of three groups of fungi (coelomycetes, hyphomycetes, and mycelia sterilia) reproducing asexually and collectively referred to as imperfect fungi or deuteromycetes. Other names for this collective group are anamorphic fungi or, for those producing conidia, mitosporic fungi. By convention, asexual states of ascomycetes and basidiomycetes are included in deuteromycetes, and asexual states of Zygomycetes are usually excluded. *See* ASCOMYCOTA; BASIDIOMYCOTA; ZYGOMYCETES.

The traditional orders are Melanconiales, Sphaeropsidales, and Pycnothyriales, which correspond to coelomycetes producing fruiting bodies that are disklike with wide apical openings (acervuli, **Fig. 1**), globose with restricted apical openings (pycnidia, **Fig. 2**), or flattened with small apical openings (pycnothyria). Members of the first two orders are much more common and more important as plant pathogens than members of the third order. Each order is also nonphylogenetic, so

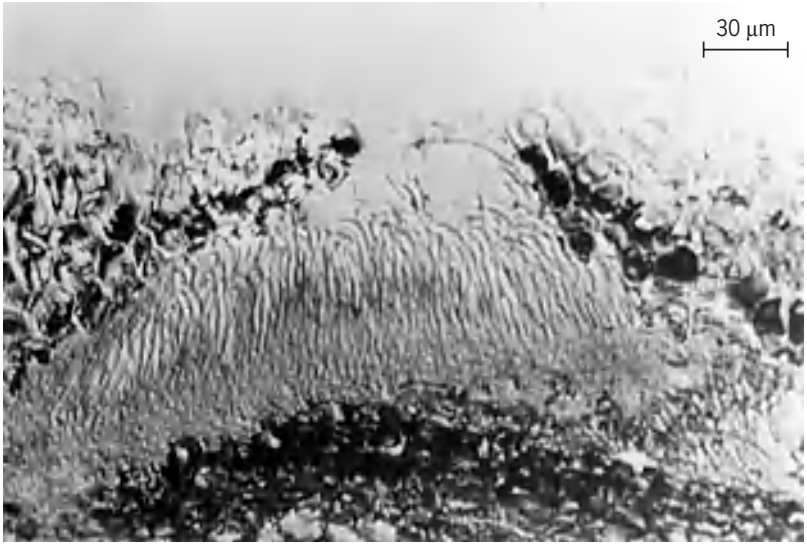

Fig. 1. Section of a typical acervulus, produced by the fungus *Cryptosporiopsis curvispora*, a cause of branch canker and fruit rot of apple.

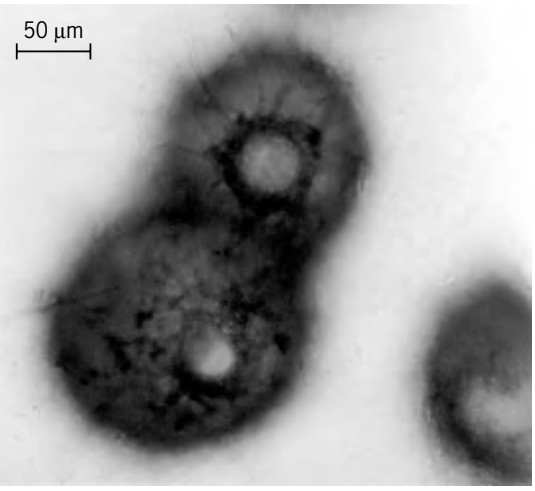

Fig. 2. Surface view of typical pycnidia, produced by the fungus *Ascochyta pinodella*, a cause of stem and leaf spot of pea.

contemporary mycologists often abandon the formal names of the orders and simply refer to acervular, pycnidial, or pycnothyrial fungi. For some fungi, the morphology of the fruiting bodies (conidiomata) is variable or otherwise nonconforming to the above terminology. In such instances, all the above terms may be discarded, and the conidiomata are described in terms of the number of cavities (locules), shape and dimension, wall construction, color, and so on. Of equal or greater importance in modern classification is the structure of the cells that produce the conidia (conidiogenous cells). This approach, first formulated for hyphomycetes by S. J. Hughes, was applied to the coelomycetes by B.C. Sutton, whose name became synonymous with coelomycete taxonomy. Contemporary identification relies on structure and color of the conidia, conidiogenous cells, conidiomata, and proper identification of the host plant. Routine diagnosis still depends on taxonomic keys and descriptions, especially those of Sutton, or the taxonomically more conservative keys of J. A. von Arx, plus literature specific to a given host. Increasingly, DNA sequence data are used in research. Commercial identification kits utilizing molecular-genetic technology are available for a small number of coelomycetes. *See* HYPHOMYCETES.

Cultivation and preservation. Most common coelomycetes can be grown on a variety of artificial, agar-based media. However, the ability to produce conidia or other reproductive structures (such as thick-walled spores called chlamydospores or multicellular structures called sclerotia) may be enhanced by the proper choice of medium. Some coelomycetes may not produce in culture the structures that characterize their growth on plants, and this is especially true of fungi that are acervular *in planta*. Prolonged growth and repeated transfers on artificial media may impair formation of structures or pathogenic abilities. Coelomycetes producing abundant conidia on artificial media usually can be preserved for many years by lyophilization, in glycerol in ultralow-temperature freezers, or in liquid nitrogen vapor. Alternatively, some species will produce conidia on sterilized filter paper squares placed on top of an agar medium. The squares can be removed, desiccated, placed in sterile envelops, and stored in a conventional freezer (approximately −15 to −20°C).

Biology, economic importance, and control. Plant pathogenic genera from temperate and tropical climates have received the greatest attention. Among those with highest impact are *Ascochyta, Colletotrichum*, and *Phoma* species on fruits and vegetables, and *Septoria* and *Stagonospora* species on cereal grains. A given species may parasitize leaves, stems, or fruits of the growing plant, or may be a postharvest pathogen on stored produce. Several genera (for example, *Cryptosporiopsis, Marssonina, Phomopsis*) cause stem cankers on berry bushes or on trees producing fruits or nuts. *Diplodia, Macrophoma*, and *Phomopsis* cause diseases on grape vines or grapes. Species in some genera (such as *Phomopsis, Sphaeropsis*) attack forest trees. *Colletotrichum, Cryptocline, Phomopsis*, and *Sphaceloma* present disease problems for the ornamental flower industry. *Ascochyta, Selenophoma, Septoria,* and *Stagonospora* afflict turf grasses. Fiber crops are attacked by *Colletotrichum* (on cotton, flax, and jute) and *Ascochyta* and *Phomopsis* (on hemp). For important species in which the teleomorph is abundant, much research has focused on the implications that sexual recombination of genes has for generation of new races, fungicide resistance, or enhanced aggressiveness. Sometimes epidemics in field crops are initiated by wind-blown sexual spores (ascospores), in which case control of the sexual stage becomes as important as control of the conidial state. Conidia are usually produced later in the season, are more abundant than ascopores, and tend to be produced in sticky masses dispersed by rain splash or irrigation. Control may be via breeding of resistant plant cultivars, fungicide applications, or cultural practices. In clinical situations, infection by coelomycetes usually occurs through wounds. Itraconazole, an antifungal compound, has been used successfully for therapy. A few coelomycetes are parasitic on lower animals; for example, *Aschersonia* is known as a pathogen of scale insects and whiteflies. Species with saprophytic capabilities can become destructive if colonizing manufactured materials, such as occurs with some *Phoma* species on carpet, paper, or plastics. However, coelomycetes are also recovered in studies of decaying plant litter, where they probably play a beneficial role.

Frank Dugan

Bibliography. H. L. Barnett and B. B. Hunter, *Illustrated Genera of Imperfect Fungi*, APS Press, St. Paul, 1998; E. Kiffer and M. Morelet, *The Deuteromycetes*, Science Publishers, Enfield, 2000; R. T. R. Nag, *Coelomycetous Anamorphs with Appendage-Bearing Conidia*, Mycologue Publications, Waterloo, 1993; B. C. Sutton, *The Coelomycetes*, CMI, Kew, 1980; J. A. von Arx, *Plant Pathogenic Fungi*, J. Cramer, Berlin, 1987.

Coenopteridales

True ferns which span the Late Devonian through Permian time between the recognizable beginnings of fernlike morphology and the earliest-appearing extant filicalean families (Gleicheniaceae, Osmundaceae). A true fern is a relatively advanced type of vascular land plant with distinct stem, fronds, roots, and foliarborne annulate sporangia. Coenopterid ferns are mostly small and simple in contrast to late Paleozoic tree ferns of the Marattiales. All well-known genera of the Coenopteridales in the Pennsylvanian Period exhibit fronds with circinate vernation (croziers), laminate vegetative foliage, and annulate sporangia. Their stems are mostly protostelic with a few siphonostelic species, and one genus (*Zygopteris*) has secondary growth. These ferns are better from anatomically preserved specimens, especially from peat beds, than from coalified compressions, but they occurred in a variety of habitats mostly known on Northern Hemisphere continents. Their fronds have distinct vascular strands, useful in identifying genera (**Fig. 1**).

There are two major distinct groups. Zygopterid ferns are the most ancient and diverse, and apparently a dead-end evolutionary line; they differ the most from other ferns. *Zygopteris* (Zygopteridaceae: Etapteroideae) had dichotomous rhizomes with distantly spaced erect fronds exhibiting three-dimensionally arranged primary pinnae in four rows. Sessile sori of elongate multiseriate sporangia with paired annuli and with sclerotic nests in their interannular walls were borne abaxially on laminate foliage in *Corynepteris* or trailed from pinnae on slender stalks in *Biscalitheca* (**Fig. 2**). Coenopterid ferns are usually placed in the Anachoropteridaceae and Botryopteridaceae, or in one of several extinct families assigned to the Filicales. Coenopterid ferns *sensu stricto* are probably ancestral to, and consequently form an imperceptible transition with, the Filicales. The best-known genera, *Anachoropteris* and *Botryopteris*, commonly have semierect (like *Osmunda*), protostelic rhizomes with closely spaced fronds bearing two rows of branched pinnae, occasional to frequent buds or vegetative shoots, laminate foliage except on some fertile pinnae, and small to large clusters of uniannulate sporangia. Some of the Pennsylvanian-age fern genera (*Sermaya, Doneggia*) with *Anachoropteris*-like anatomy have been assigned to the filicales as members of an extinct family, Sermayaceae.

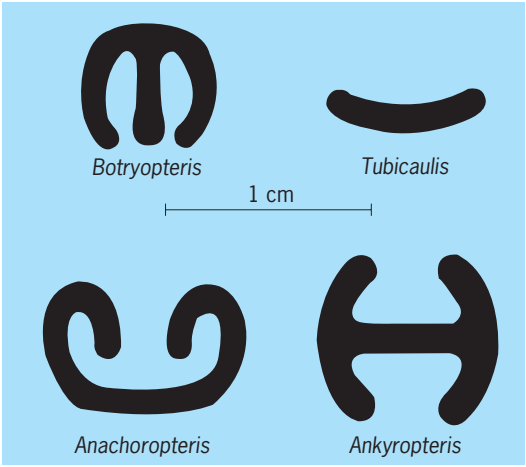

Fig. 1. Diagrammatic cross sections of petiolar vascular strands of representative coenopterid genera, oriented so the parent stem would appear below each diagram.

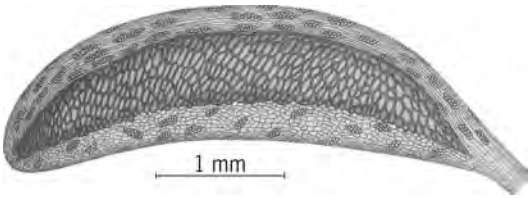

Fig. 2. Side-view reconstruction of *Biscalitheca musata*, a zygopterid sporangium with dorsiventral symmetry, two lateral multiseriate annuli, and small nests of thickened cells in the sporangial wall. (*After S. H. Mamay, Amer. J. Bot., 44(3):236, 1957*)

Evolution of stelar morphology from protostelic to siphonostelic, with or without leaf gaps, is demonstrated in the Anachoropteridaceae and Botryopteridaceae in the Pennsylvanian. Sporangia of various coenopterid genera are osmundoid, gleichenioid, or schizaeoid, generalized and simple morphological types.

The Zygopteridaceae appear in Late Devonian time, the Botryopteridaceae in Visean (Mississippian), the Anachoropteridaceae in the lower Westphalian A (Pennsylvanian), and all families extend into the Permian. *See* PALEOBOTANY.

Tom L. Phillips

Bibliography. R. L. Dennis, Studies of Paleozoic ferns: *Zygopteris* from the middle and late Pennsylvanian of the United States, *Palaeontographica*, 148B:95–136, 1974; D. A. Eggert and T. Delevoryas, Studies of Paleozoic ferns: *Sermaya*, gen. nov., and its bearing on filicalean evolution in the Paleozoic, *Palaeontographica*, 120B:169–180, 1967; J. Galtier and A. C. Scott, Studies of Paleozoic ferns: On the genus *Corynepteris*, a redescription of the type and some other European species, *Palaeontographica*, 170B:81–125, 1979; T. L. Phillips, Evolution of vegetative morphology in coenopterid ferns, *Ann. Mo. Bot. Garden*, 61:427–461, 1974.

Coenurosis

The infection by the larval stage of several species of *Taenia* from canids, such as *T. multiceps, T. serialis*, and *T. brauni*. It is rare in humans. A bladder as large as a hen's egg and containing scolices (protoscolices) forms and daughter bladders bud off either internally or externally, attached by stalks. In subcutaneous or ocular locations the bladders are unilocular, but in the central nervous system they are frequently multilocular (racemose). Symptoms are the result of toxic and allergenic metabolites and pressure effects. Surgery may be attempted; chemotherapy may be

effective, but secondary effects still render a poor prognosis.

Herbivores feeding on grass contaminated by the parasite's eggs also develop coenurosis. If the central nervous system is involved (as is frequent in sheep) neurological signs appear and the animal staggers and circles. Such an animal is said to be giddy (thus the origin of the term "gid tapeworm"). Economic losses in sheep-raising regions due to coenurosis are considerable. The disease could be controlled by treatment of shepherd dogs. *See* CYCLOPHYLLIDEA.

José F. Maldonado-Moll

Coenzyme

An organic cofactor or prosthetic group (nonprotein portion of the enzyme) whose presence is required for the activity of many enzymes. In addition, many enzymes need metal ions, such as copper, manganese, and magnesium, for activation. The prosthetic groups attached to the protein of the enzyme (the apoenzyme) may be regarded as dissociable portions of conjugated proteins. The conenzymes usually contain vitamins as part of their structure. Neither the apoenzyme nor the coenzyme moieties can function singly, since dissociation of the two results in inactivation. In general, the coenzymes function as acceptors of electrons or functional groupings, such as the carboxyl groups in α-keto acids, which are removed from the substrate. Some of the well-known coenzymes are mentioned below. *See* PROTEIN; VITAMIN.

The pyridine nucleotides, nicotinamide adenine dinucleotide (NAD) and nicotinamide adenine dinucleotide phosphate (NADP), function as hydrogen acceptors for a large number of dehydrogenase enzymes. Thiamine pyrophosphate (TPP), or cocarboxylase, the prosthetic group of carboxylase, is responsible for the decarboxylation of α-keto acids in the body. Flavin mononucleotide (FMN) and flavinadenine dinucleotide (FAD) act as hydrogen acceptors in aerobic dehydrogenases. Pyridoxal phosphate, also known as codecarboxylase and cotransaminase, serves as coenzyme for various racemases in reactions involving the decarboxylation of α-amino acids, and also functions in transamination.

Iron protoporphyrin (hemin) acts as coenzymes of catalase, peroxidases, and cytochromes. The uridine phosphates are involved in carbohydrate metabolism. Uridine diphosphate (UDP) is the chief transferring coenzyme for carbohydrates. UDP-glucose is responsible for the interconversion of glucose and galactose, is used in the synthesis of glycogen, and can be oxidized to UDP-glucuronate for the synthesis of glucuronides and chrondroitin or converted to UDP-galactose for the synthesis of lactose and chrondroitin.

Adenosine triphosphate (ATP), adenosine diphosphate (ADP), and adenosine monophosphate (AMP), the phosphorylated derivatives of adenylic acid, can be active in phosphate transport or transphosphorylation or both. Lipoic acid functions in acyl-generating and in acyl-transfer reactions. In combination with TPP and NAD, it functions in oxidative decarboxylation of α-keto acids. Coenzyme A (CoA), a coenzyme in certain condensing enzymes, acts in acetyl or other acyl group transfer and in fatty acid synthesis and oxidation. Folic acid coenzymes are involved in the metabolism of one carbon unit. Biotin is the coenzyme in a number of carboxylation reactions, where it functions as the actual carrier of carbon dioxide. *S*-Adenosylmethionine, active methionine, functions is transmethylation reactions. *See* ADENOSINE DIPHOSPHATE (ADP); ADENOSINE TRIPHOSPHATE (ATP); BIOCHEMISTRY; BIOLOGICAL OXIDATION; BIOTIN; CARBOHYDRATE METABOLISM; CYTOCHROME; ENZYME; HEMOGLOBIN; NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD); NICOTINAMIDE ADENINE DINUCLEOTIDE PHOSPHATE (NADP); URIDINE DIPHOSPHOGLUCOSE (UDPG).

Mary B. McCann

Coesite

A naturally occurring mineral of wide interest, the high-pressure polymorph of SiO_2. Coesite was first discovered and identified in various meteorite impact craters and in some tektites. Because it requires a unique physical condition, extremely high pressure, for its formation, its occurrence is diagnostic of a special natural phenomenon, in this case, the hypervelocity impact of a meteorite.

Coesite occurs in grains that are usually less than 5 micrometers in size and are generally present in small amounts. The properties of the mineral are known mainly from studies of synthesized crystals. It is colorless with vitreous luster and has no cleavage. It has a specific gravity of 2.915 ± 0.015 and a hardness of about 8 on Mohs scale. It is biaxial positive with $2V$ about 64°. Its indices of refraction are α 1.5940, β 1.5955, and γ 1.5970 ± 0.0005. Its dispersion is horizontal, with r less than v (weak). The optical orientation is $X = b$, $Z \angle C = 4\text{–}6^\circ$, and $\beta = 120^\circ$. Synthetic crystals occur as euhedral to subhedral hexagonal platelets, and laths with positive elongation. Simple contact twins occur with (021) as twin and composition plane.

Natural coesite can be identified from thin sections of rocks by its peculiar habit and high index of refraction. However, its positive identification must rest on the x-ray powder diffraction technique.

Coesite has as yet no evident commercial use and therefore has no obvious economic value. As a stepping-stone in scientific research, it serves in at least two ways: where it occurs naturally, coesite is diagnostic of a past history of high pressure; the natural occurrence of coesite suggests that shock as a process can transform a low-density ordinary substance to one of high density and unique properties.

Coesite has been found in materials ejected from craters formed by the explosion of 500,000 tons (450,000 metric tons) of TNT. The study of coesite and craters is important to the understanding of the

impact craters on the Earth as well as those on the Moon. Edward C. T. Chao

Bibliography. F. R. Boyd and J. L. England, The quartz-coesite transition, *J. Geophys. Res.*, 65(2): 749–756, 1960; E. C. T. Chao, E. M. Shoemaker, and B. M. Madsen, The first natural occurrence of coesite from Meteor Crater, Arizona, *Science*, 132(3421):220–222, 1960; L. S. Walter, Coesite discovered in tektites, *Science*, 147:1029–1032, 1965; T. Zoltai and M. J. Buerger, The crystal structure of coesite, the dense, high pressure form of silica, *Z. Krist.*, 111:129–141, 1959.

Coffee

A tropical evergreen shrub or small tree of the genus *Coffea* (Rubiaceae), a native of northeast Africa and adjacent southwest Asia. The beverage known as coffee is made by the hot-water extraction of solubles from the ground roasted beans (seeds) of the shrub. *Coffea* grows mainly between the Tropic of Cancer and Tropic of Capricorn at elevations of 2000–6000 ft (600–1800 m) above sea level, at temperatures near 70°F (21°C), and with annual rains near 60 in. (150 cm).

Annual international commerce is about 75 million bags (standard of 130 lb or 60 kg per bag) of green coffee beans, that is, coffee beans that have been cultivated, harvested, and are ready for commercial processing. The supply generally is a third each from Brazil, Africa, and all other countries. The United States and Europe consume 85% of all world exports. Coffee beverage has no nutritive value and is consumed for its flavor and stimulating effects. The stimulating ingredient in coffee is caffeine. The two main coffee varieties are Arabica with about 1.1% caffeine and Robusta with 2.2%. *See* CAFFEINE.

History. The *C. arabica* shrub (**Fig. 1**) is native to Ethiopia, where it grows wild in rainforests, and its seeds have been used by the native populations since before recorded time. The *C. robusta* shrub grows taller (to 20 ft or 6 m) than the Arabica shrub and is more productive and hardy. It grows in warmer climates and is native to the Congo area south of Ethiopia. About A.D. 900, Arabs planted the Ethiopian coffee shrub in Yemen, and for 700 years coffee propagation was limited to this area.

Fig. 1. *Coffea arabica* **branch.**

Commercial coffee cultivation was started by the Dutch in Java and Ceylon in the late 1600s; by the French, English, and Spanish in the West Indies in the early 1700s; in Mexico, Central America, and Hawaii in the early 1800s; in Colombia about 1900; and in Africa in the late 1940s. Brazil had been the dominant supplier since the early 1800s, but its dominance receded in favor of African coffee. However, Brazil has become a large producer of soluble coffee powder, producing over 60 million pounds (27 million kilograms) per year.

Green coffee cultivation. The coffee plant is a shrub or relatively small tree, often controlled to a height of 15–18 ft (4.5–5.5 m). *Coffea arabica* (milds) accounts for 69% of world production; *C. canephora* (robustas), 30%; and all others, 1%. Each species includes several varieties.

After the spring rains, the plant produces white flowers. About 6 months later the flowers are replaced by fruit the size of a small cherry (Fig. 1). The cherries can be underripe, ripe (red, yellow, and purple color), and overripe; and they can be selectively picked to harvest only ripe ones, or strip-picked to yield mostly ripe but also some over- or underripe fruit. The green coffee beans are the two halves of the seed derived from the processed cherries. Franklin Kramer

Diseases. There are more than 350 known diseases of coffee, and new ones are still being discovered on *C. arabica*, *C. robusta*, *C. liberica*, and *C. excelsa* in widely separated tropical lands. Probably no crop grows under more varied conditions or is subjected to so many unknown diseases.

Leaf rust (caused by *Hemileia vastatrix*), one of the most important diseases of the crop worldwide, was described as early as 1869 in Ceylon, where there were once flourishing coffee plantations. Rust became so damaging that, in the absence of control measures, coffee production had to be abandoned. In Eastern and African coffee-producing countries, rust restricted cultivation of coffee to particular altitudes and districts, rendering cultivation more expensive and difficult. As a consequence, the center of the world's potential coffee production shifted to Brazil and other South and Central American countries, where leaf rust was unknown until 1970. Control of the disease can be achieved by fungicidal sprays, but this inevitably increases production costs. Development of resistant varieties is complicated because of the existence of many different races of the coffee rust fungus, but some new introductions show promise.

Other diseases, which are less widespread than leaf rust, have also been extremely damaging. In East Africa, bark disease (caused by *Fusarium stilboides*) destroyed the once-flourishing coffee

Fig. 2. Coffee berry disease (caused by *Colletotrichum coffeanum*). Berries, which may be infected at any stage of development, are completely destroyed.

industry of Malawi and, in the late 1960s, berry disease (caused by *Colletotrichum coffeanum*; **Fig. 2**) devastated the crop in Kenya and threatened the viability of the coffee production. Fungicidal sprays now give control. Berry disease spread from Kenya to all other African coffee-producing countries, but is unknown outside the African continent.

American leaf spot (caused by *Mycena citricolor*) is very serious but is found only in the Americas, whereas Cercospora leaf spot and the leaf rot caused by *Pellicularia koleroga* occur in both hemispheres. The last two diseases are controlled by spraying.

Fruit, stem, and branch infections are injurious. Wilts and trunk cankers, caused by *Gibberella, Nectria*, and *Ceratocystis*, are troublesome. Probably six *Rosellinias*, a few *Fusaria*, two *Armillarias*, and fungi like *Fomes, Ganoderma*, and *Polyporus* attack roots. Root and trunk infections are reduced by good cultural practices. Coffee is also affected by some half dozen viruses, a few bacteria, and numerous parasitic mistletoes. *See* FUNGISTAT AND FUNGICIDE; PLANT PATHOLOGY; PLANT VIRUSES AND VIROIDS.

Ellis Griffiths

Coffee cherry processing. Coffee cherries are processed by either dry or wet methods. Dry-method products are so-called natural coffees, while wet-method products are usually the more uniform and higher-quality washed coffees.

The dry method is used in Brazil and in other countries where water is scarce in the harvesting season. The cherries from strip picking are spread on open drying ground and turned frequently to permit thorough drying by the sun and wind. Sun drying usually takes 2–3 weeks, depending on weather conditions. Some producing areas use hot air, indirect steam, and other machine-drying devices. When the coffee cherries are thoroughly dry, they are transferred to hulling machines that remove the skin, pulp, parchment shell, and silver skin (a thin cellulosic layer that envelops the coffee bean) in a single operation. In Colombia, freshly picked ripe coffee cherries are fed to a tank for initial washing and removal of stones and other foreign material. The cherries are then transferred to depulping machines that remove the outer skin and most of the pulp, although some pulp mucilage clings to the parchment shells that encase the coffee beans. This is removed in fermentation tanks, usually containing water, a process that may last from 12 h to several days. Because prolonged fermentation may cause undesirable flavors and odors in the beans, enzymes may be used to speed up the process.

The beans are then dried, either in the sun or in mechanical dryers. Machine drying has become more popular in spite of higher costs, because it is faster and independent of weather conditions. After drying, the coffee is further processed in machines that remove silver skins, producing green coffee beans. The coffee beans are then machine graded by screens, oscillating tables, and airveying (pneumatic conveying) into large, medium, and small beans. Damaged beans or foreign materials are removed by hand or by mechanical sorting devices. The beans are then further graded according to appearance, color, and cupping of roasted samples. After processing, the green beans have 10–12% moisture and maintain acceptable quality for about a year.

Coffee chemistry. Chemical analysis of green coffee reveals a complex chemical composition (**Table 1**). The main differences between milds and robustas are caffeine content (higher for robusta coffee), lipids (higher for milds), chlorogenic acid (higher for robusta), trigonelline (higher for milds), and sucrose (higher for milds).

Green coffee has no desirable taste or aroma—these qualities result from roasting. Many complex physical and chemical changes occur during roasting, including the change in color from green to brown. In the first stage of roasting, free water is removed. In the second stage, chemical reactions lead to the formation of compounds associated with aroma and flavor.

Some reactions during roasting are exothermic. These produce a rapid rise in temperature, usually accompanied by a sudden expansion or puffing of the beans, with a volume increase from 50 to 100% depending on the bean variety and roasting conditions. The relatively high value of caffeine in brewed and instant coffee solids (**Table 2**) is due to the very high solubility of caffeine in hot water compared to the other constituents in the roasted beans. The total solids content in a brewed cup of coffee is 0.7–1.2%; espresso coffee has about 3–5% solids. Most aromatic compounds are lost or changed in the instant coffee

TABLE 1. Typical analyses of green coffee, %

Constituent*	Type†	
	C. canephora (robustas)	*C. arabica* (milds)
Moisture	11 (10–13)	12.5 (10–13)
Lipids	10 (7–11)	15 (14–17)
Ash	4.2 (3.9–4.5)	4.0 (3.5–4.5)
Caffeine	2.0 (1.5–2.6)	1.3 (1.1–1.4)
Chlorogenic acid	9 (7–10)	7 (5–8)
Carboxylic acids	2 (1–3)	2.5 (1.5–3.5)
Trigonelline	0.7 (0.3–0.9)	1.1 (0.9–1.2)
Protein	11 (9–13)	11 (9–13)
Free amino acids	0.8	0.5
Sucrose	4 (3–6)	8 (5–9)
Reducing sugars	0.5 (0.4–0.6)	0.1 (0.1–0.2)
Others‡	8.8 (5–10)	5.5 (3–8)
Polymeric carbohydrate		
Mannan	22	22
Arabinogalactan	17 (16–18)	15 (14–16)
Cellulose	8 (7–9)	7 (7–8)

*Dry basis (ex moisture; as-is basis).
†Typical value (range).
‡Others by difference.

processing steps. Instant coffee producers are particularly interested in the nature of these compounds and how to capture and include them in the final product.

Green-bean processing. A series of basic bean processing steps is required to produce the major types of commercial coffee products (**Fig. 3**). First, the green beans are weighed and cleaned, and then they are stored in silos, each containing a single bean type. Green coffee beans to be used for decaffeinated coffee products are processed separately to remove the required amount of caffeine (usually 97% in the United States) and then stored in special silos before further processing.

Roasting and grinding technology. Roasting and grinding are key steps in producing all coffee products.

Roasting. The roasting step, including the type of beans, roasting machinery, and operating parameters (particularly temperature and time), is the most important operation in defining the final product characteristics. Roasting is performed with individual bean types, blended beans, decaffeinated beans, and beans with additional ingredients—for example, sugar and chicory may be added to form specialized flavored coffees.

Roasting is usually conducted by use of hot combustion gases in rotating cylinders or fluidized-bed systems. Infrared and microwave energy has been used to a limited extent in the Far East. The hot-gas roasters are designed for batch and continuous operations. Batch cylinder roasters, which continue to be used only by smaller producers and for specialty products, traditionally have had roasting times of 10–20 min. Fluidized-bed batch roasters roast beans for 2–5 min. Continuous roasters, which are available in cylinder and fluidized-bed form and can achieve very short roasting times, are best suited for large producers that are roasting a consistent blend.

After moisture has been eliminated and the bean

TABLE 2. Approximate analyses of roasted, brewed, and instant coffee, %*

Constituent†	Roasted	Brewed	Instant
Moisture	4 (1–5)		3.5 (2–5)
Oil	17 (16–20)	0.8 (0.2–1.0)	(0.1–0.6)
Ash	4.5 (4–5)	14	9 (7–11)
Caffeine	1.2 (1.0–1.6)	4.8	3.5 (2–5)
Chlorogenic acid	1.5 (1–3.5)	14.8	(3–9)
Carboxylic acids	3 (2–4)	3.0	5.5 (4–8)
Trigonelline	0.5–1.0	1.6	(0.5–2)
Protein	8–10	6	(1–6)
Reducing sugar	0.2	0.4	(1–5)
Sucrose	0.2 (0–0.5)	0.8	(0.6–0)
Aroma compounds	0.1	0.1–2	0.05
Browning products and others	(22–11)	29.4	(20–35)
Polymeric carbohydrate	38	24	(30–50)
Mannan	20		
Arabinogalactan	12		
Cellulose	6		

*Typical value (range).
†Dry basis (ex moisture; as-is basis).

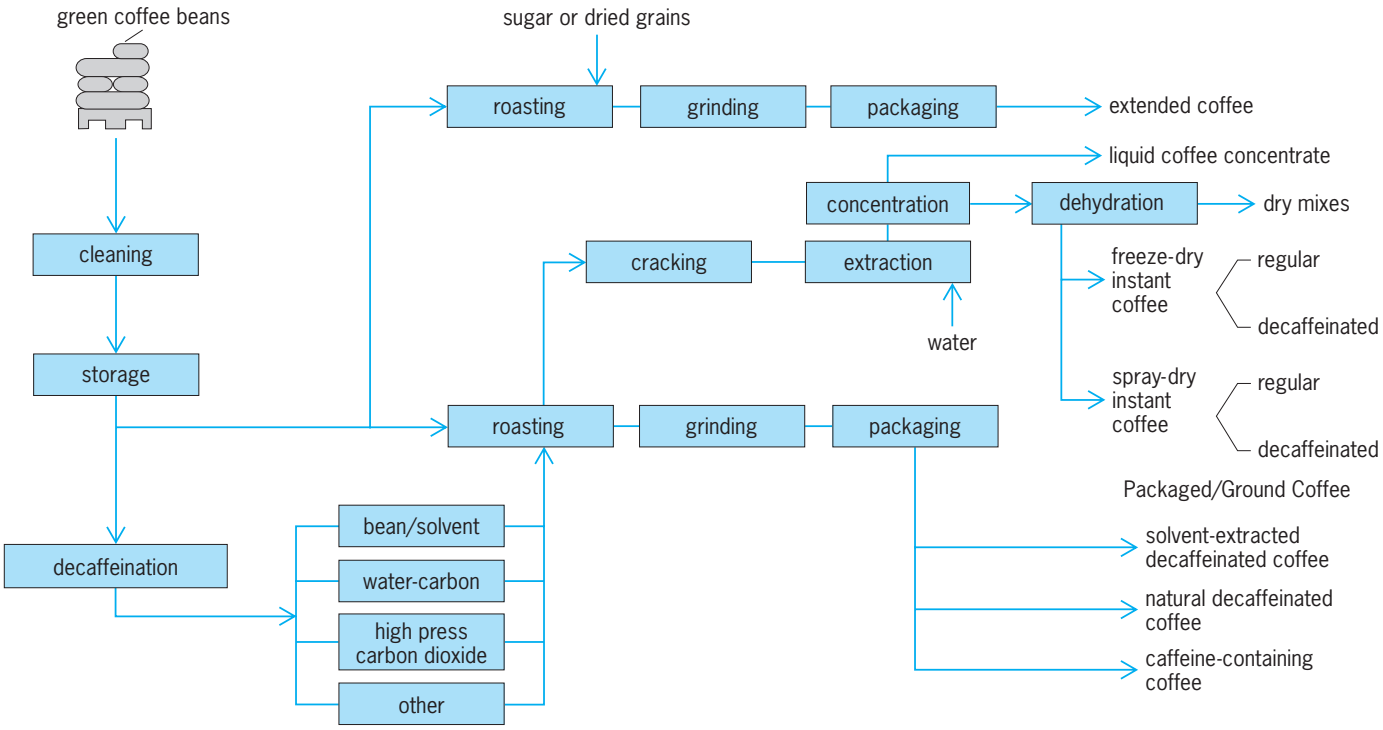

Fig. 3. Flow chart showing the basic processing steps for producing coffee products from green coffee beans.

temperature has reached 330°F (166°C), the reaction changes from endothermic to exothermic, and the required features of the roasted product begin to develop while the beans swell to as much as twice their original size. When the bean temperature reaches 375–475°F (190–246°C), roasting is terminated by rapidly adding a water quench. The final roast temperature is the critical control point that fixes the flavor and color of the product, and it must be consistent for good quality control. Periodic photometric reflectance tests on a ground sample are used for fine-tuning the required bean temperature just before quenching. The water quench adds moisture (usually 2–4%) to the roasted beans.

The faster fluidized-bed roasters are used for producing high-yield coffee. This method provides faster roasting with significantly lower combustion gas temperatures, resulting in more uniform temperature within the beans and hence a larger percentage of soluble solids. Such coffees lend themselves to the low-weight packaged roast and ground coffee, and offer higher yield opportunities for instant coffee processes. Roasting gas temperatures range from 450 to 750°F (230 to 400°C). The exhausted roaster gas normally requires afterburners to remove particulates to conform with environmental regulations. *See* FLUIDIZATION.

The quenched beans are air cooled and conveyed to storage bins for moisture and temperature equilibration prior to grinding. Residual foreign material (mostly stones) that has passed through the initial cleaning step is removed in transit to the storage bins by means of a high-velocity air lift, which leaves the heavier debris behind. The separation is much more efficient at this point because bean density has been reduced by the roasting step.

Grinding. Grinding of the roasted coffee beans is tailored to the requirements of the intended beverage preparation. For example, only coarse cracking of the beans is needed for instant coffee, whereas finer grinds are required for packaged roast and ground coffee.

The most commonly used grinder is a multistep steel roll mill. After passing through a pair of coarse cracking rolls, the broken beans pass through two successive sets of paired longitudinally scored rolls that are operated at speeds designed to cut, rather than crush, the particles. The second set of rolls are more finely scored and operate at higher speeds. The final part of the roll grinder is the normalizer mixing section, to provide a uniform distribution. Particle size is controlled by regulation of roll separation.

Instant coffee processing. Instant coffee comprises the dried solids made from water extraction of the soluble solids in roasted and partially ground (cracked) coffee. The extraction process follows the same principle as home brewing using a filter-drip percolator, except that it is much more efficient and results in very high solid-extraction yields. The countercurrent staged extraction system consists of a battery of 5–10 vertical columns connected by appropriate manifolds of pipes and valves. The water, heated under pressure to 300–350°F (150–180°C), circulates countercurrent to the cracked roasted coffee by intermittent column stage position changes. The columns do not move, but by valve changes the flow of water is stepwise from the final fresh column to the fully extracted first position, each step lasting 30–45 min. The flowing extract is cooled just prior to the final fresh column so that the finished extract is

collected at 150–180°F (66–85°C). Yields of 40–60% by weight of roasted coffee are achieved, depending on the flow rates, type of coffee used, and heat input. *See* EXTRACTION.

The coffee extract is concentrated from about 20–35% to 50% by using multiple-effect vacuum evaporators or freeze concentration, or a combination of both. The flavor of the final dry product can be enhanced by recovering the aroma and flavor compounds prior to, or as part of, the concentration process and then introducing these to the concentrated extract just before drying. The method of drying defines the instant coffee as either spray dried or freeze dried. Spray drying involves spraying the concentrated aroma-enhanced extract into a concurrently flowing stream of hot combustion gases. The amount of flavor retained is greater as the concentration of sprayed droplets increases. The dry product is collected at the bottom of the dryer as a powder of spherical particles. Its appearance and solubility are improved by forming agglomerates of the particles, usually by mixing them with a flow of steam that condenses on them to provide adhesive. The excess moisture is then removed by air drying.

Freeze drying uses special high-vacuum equipment to remove water from the frozen extract by sublimation. The vacuum has to be maintained below 0.5 mmHg (67 pascals). The very low temperature maintained during the drying period minimizes flavor changes and aroma losses. Freeze-dried coffee is more expensive to produce and requires a much higher investment than the spray-dried product. *See* DRYING.

Decaffeinated coffees. Decaffeinated coffees represent about 18% of the total coffee consumption in the United States. Decaffeinated coffee was commercially developed in Europe about 1900. The early process involved direct contact of water-moistened beans with a solvent. The principle of moisturizing green coffee before decaffeination remains a key process step for all decaffeination processes, because it facilitates transport of caffeine through the cell walls. The older European process requires removing the solvent from beans by steam distillation, followed by air drying to remove the water. The caffeine that is removed, after refining to United States Phamacopeia (USP) standards, is a valuable by-product for pharmaceuticals and soft drinks. Though solvent decaffeination has been considered undesirable, solvents not found harmful by various government laboratories continue to be used. In fact, much of the commercial decaffeinated coffee is made from processes closely resembling the original European method. Other processes have been developed to improve coffee quality as well as to minimize solvent residues.

A water extraction process was developed in the United States in 1943 which provided a decaffeinated coffee with improved flavor and with very low solvent residues, in the parts-per-million range. This process used water to extract solids from green beans in a countercurrent staged column system. The green extract was then decaffeinated by extraction with a non-water-miscible solvent, such as methylene chloride, which had solubility characteristics that favored caffeine over the other soluble solids. The decaffeinated green extract was steam stripped to remove residual solvent and then returned to the extraction system. The caffeine-rich solvent extract was concentrated and recovered for reuse.

Natural decaffeination processes were invented and commercialized from the mid-1970s through the 1980s in response to increasing demand for

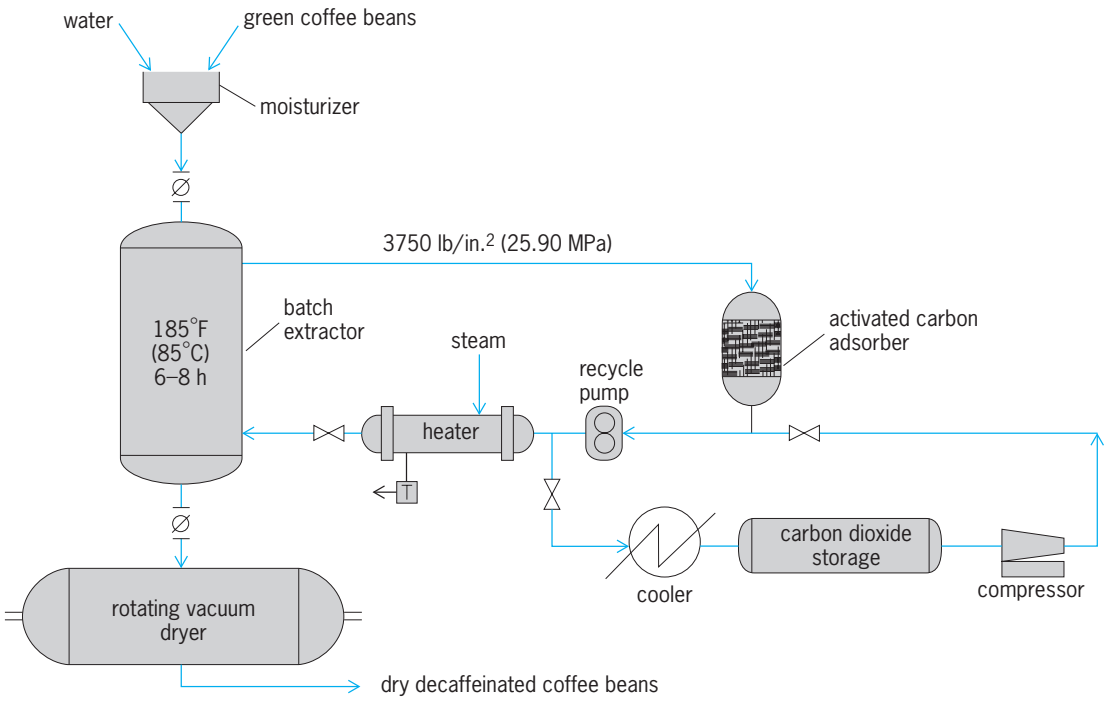

Fig. 4. Flow chart for coffee decaffeination with supercritical carbon dioxide.

decaffeinated coffee that had the same characteristics and quality as regular brewed coffee. Also, decaffeinating solvents, even in trace amounts, were considered to be harmful. Three processes for natural decaffeinated products are commercially available.

Water-carbon process. Referred to as the water process, this method is in many ways the same as the earlier water extraction process. However, instead of using a solvent to remove caffeine from a green coffee extract, the extract is decaffeinated by adsorption using activated carbon. The attraction by the carbon for many of the noncaffeine solids is partially reduced by sugar-coating the carbon. One disadvantage of this process is the inability to recover the caffeine. *See* ACTIVATED CARBON; ADSORPTION.

Supercritical CO_2 processes. Carbon dioxide (CO_2) above its critical temperature (supercritical CO_2) can be used to remove caffeine from green coffee beans (**Fig. 4**). It is a batch operation requiring pressures of about 3500 lb/in. (25.9 megapascals). Supercritical CO_2 is a very specific solvent for caffeine in moistened green beans, but few noncaffeine compounds are removed. Activated carbon adsorbs all the caffeine in the recirculating gas. Even at moderate extraction temperatures, times of 5–8 h are sufficient to obtain the desired degree of caffeine removal. After saturation, the carbon can be regenerated, but the caffeine cannot be recovered. An improved version of the supercritical CO_2 process features two major differences: (1) It simulates a continuous process by intermittently (15-min intervals) removing decaffeinated coffee from the bottom of an extraction tower and then filling space formed at the top of the tower with fresh green coffee. (2) Water is used instead of carbon to remove caffeine from the recirculating supercritical CO_2. The caffeine is then recovered, purified, and dried for reuse.

Other processes. Other so-called natural processes are similar to older solvent processes but use solvents (such as ethyl acetate) considered to be natural, based on the designation "generally regarded as safe" (GRAS) by the U.S. Food and Drug Administration.

Standards. Regulations for decaffeinated coffee differ. The European Economic Community requires that roasted decaffeinated coffee contain no more than 0.1% caffeine (dry basis). This differs from the United States standard, which states "97% caffeine removed," meaning less than 0.03% caffeine in milds and 0.06% in robustas. The United States standard is actually more rigorous, and more difficult to achieve and control.

Product stability and quality. All types of coffee products get stale rapidly when exposed to air, specifically oxygen. Therefore, raw materials, process materials, and the final product must have minimal exposure to oxygen. Although green beans are stable for long periods of time, once roasted, they grow stale in about 7–10 days. Ground coffee becomes stale very rapidly in air and therefore requires vacuum or modified atmospheric packaging (less than 1% oxygen). Once consumer coffee packages are opened, they should be kept refrigerated or in a freezer for extended shelf life.

The instant coffees are less subject to staling but require some of the same packaging and use concerns. The main concern is moisture adsorption due to their hygroscopic nature. Because so-called jar aroma is frequently added, it is desirable to minimize exposure to oxygen. Aromatized instant coffees for consumer sales are normally packaged in nitrogen-flushed jars. *See* FOOD ENGINEERING; FOOD MANUFACTURING.

Franklin Kramer

Bibliography. R. J. Clarke and R. Macrae (eds.), *Coffee*, vol. 1: *Chemistry*, 1985, vol. 2: *Technology*, 1987; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., vol. 6, 1993; P. H. Marse et al., *Volatile Compounds in Food: Qualitative and Quantitative Data*, vol. 11, TNO-CIVO, Food Analyses Institute, 1989; M. Sivetz, *Encyclopedia of Food Technology*, vol. 2, 1974; M. Sivetz and N. Desrosier, *Coffee Technology*, 1979; *World Coffee Situation*, U.S. Department of Agriculture, 1991.

Cofferdam

A temporary wall-like structure which keeps water and soil out of a building site so that construction can proceed under dry conditions. Cofferdams are used in the construction of foundations, bridge piers, dams, dry docks, and similar structures which require dewatering and excavation of an area. The dewatered area can be completely surrounded by a cofferdam structure or a combination of earth slopes and cofferdam structure. The type of construction is dependent on the site depth, soil conditions, fluctuations in water level, availability of materials, working conditions inside the cofferdam, and location either on land or in water. In most situations, since safety, ease of construction, and convenience are the most important considerations, and economy of materials is less important, refined methods of analysis may not be justified. However, an understanding of soil and water pressure against the cofferdam walls is necessary for even the most approximate design. Other important considerations in cofferdam design are the hydraulic analysis of seepage conditions (including filtering), and erosion of the bottom (scour) in streams and rivers.

Where the cofferdam structure can be built on a layer of low-permeability soil (which minimizes the passage of water), the area within the cofferdam can be essentially sealed off. Where the foundation soils are pervious, flow of water into the cofferdam cannot be stopped economically, and the water must be pumped out to maintain dry working conditions.

Types. The **illustration** shows the types of cofferdam construction: *a*, *b*, and *e* are used to enclose small areas, such as individual bridge piers where cross bracing, consisting of steel, wood, or concrete beams and struts, to the opposite wall is practical; the other types are for large areas such as dams,

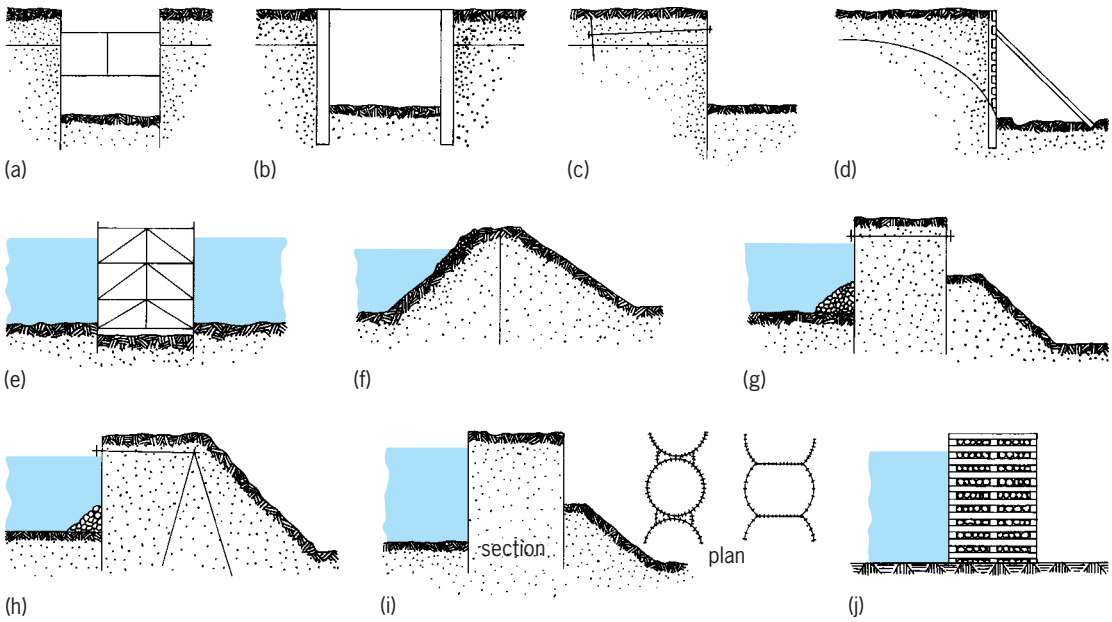

Types of cofferdams. For use on land: (*a*) cross-braced sheet piles; (*b*) cast-in-place concrete cylinder; (*c*) anchored sheet piles; (*d*) braced vertical piles with horizontal sheeting. For use in water: (*e*) cross-braced sheet piles; (*f*) earth dam; (*g*) tied sheet piles; (*h*) anchored sheet piles with earth berm; (*i*) steel sheet-pile cellular cofferdam; (*j*) rock-filled crib.

river locks, or large buildings, where cross bracing is impractical or undesirable. Cross bracing can be replaced with steel or concrete rings acting in compression when a circular form is used for *a* and *e*.

Sheeted cofferdams are made of sheets of steel, timber, or concrete, driven vertically or placed horizontally. The sheeting may be made reasonably watertight by using interlocking, overlapping, or grouted joints. In gravel or boulders, where sheeting cannot be driven to the required depth, steel H piles are driven vertically into the ground and timber sheeting is placed horizontally between or against the vertical piles (illus. *d*) as the excavation progresses. Watertight cofferdams are designed for a full hydrostatic head of water, plus the pressure of earth when they are on land. Nonwatertight cofferdams have a reduced hydrostatic pressure because of seepage.

Another type of cofferdam suitable for use in gravel or boulders is constructed of cast-in-place concrete. Illustration *b* shows this type in a cylindrical form.

Most sheet-pile cofferdams, with the exception of the cellular type (illus. *i*), require a foundation soil into which the sheets are embedded to give the walls stability. A cellular cofferdam, consisting of interlocking steel sheet piling driven as a series of interconnecting cells, is independently stable when filled and may be constructed on either rock or soil. The cells of the circular type may be filled individually. The cells of the straight-wall diaphragm type must be filled systematically to avoid a differential height of fill that would deform the diaphragms.

The filled crib (illus. *j*) is an open, boxlike structure of timber, precast concrete, or fabricated steel sections, filled with a suitable material. A rock-filled crib is suitable for construction on a rock bottom in swiftly flowing waters. Sheeting or an impervious core provides the necessary watertightness. Earth dams (illus. *f*) can be constructed of soil or rock, but if pervious, a cutoff is required to control seepage.

Another type of cofferdam for a large open building area is a slurry wall anchored back to the earth (illus. *c*). The slurry wall is constructed by pouring a concrete mix into a trench dug along the proposed wall before excavating the cofferdam area.

Uses. When cofferdams are required for year-round construction in waters subject to flooding, they are made sufficiently high to exclude the flood waters. To guard against instability, scour must be prevented along the stream side of cofferdams in swiftly flowing waters by protecting the stream bed. A rock fill, shown in illus. *g* and *h* on the stream side, may be used to provide such protection. In addition, the berms shown in illus. *g*, *h*, and *i* on the dewatered side furnish additional stability and reduce seepage of water under the cofferdam.

The cofferdam in illus. *b* has been used to form a mine shaft 220 ft (66 m) deep. Type *d* was used extensively in subway construction in New York City.

A nautical application of the term cofferdam is a watertight structure used for making repairs below the waterline of a vessel. The name also is applied to void tanks which protect the buoyancy of a vessel. *See* BRIDGE; DAM; FOUNDATIONS.

Edward J. Quirin
Patrick J. Fox

Bibliography. T. D. Dismuke, Retaining structures and excavations, in H.-Y. Fang (ed.), *Foundation Engineering Handbook*, 2d ed., 1991; M. Reimbert and A. Reimbert, *Retaining Walls, Anchorages, and Sheet Piling*, pt. 1, 1974; L. White and E. A. Prentis (eds.), *Cofferdams*, 2d rev. ed., 1956.

Cogeneration systems

Systems in which combined production of electrical power and useful thermal energy is achieved by the sequential use of a fuel or fuels. The electrical power is produced by a generator that is most often powered by a prime mover such as a steam turbine, gas turbine, or reciprocating (piston) engine. Examples of useful thermal energy include hot exhaust gases, hot water, steam, or chilled water. Other names for cogeneration systems are combined heating and power (CHP), district heating, and total energy. *See* DISTRICT HEATING; ELECTRIC POWER GENERATION; GENERATOR; PRIME MOVER; STEAM TURBINE; TURBINE.

The use of cogeneration systems has a number of benefits, including energy efficiency, monetary savings, and reduced environmental emissions. Cogeneration systems may have overall efficiencies of up to about 90%, depending on the application. Because of these high efficiencies, less fuel is consumed and less pollution is generated relative to conventional power plants. Any facility that uses electrical power and has thermal energy needs is a candidate for cogeneration. The basic consideration in determining the feasibility of cogeneration for a particular facility is whether the savings on thermal energy costs is sufficient to justify the capital expenditure. Facilities that may be considered for cogeneration include those in the industrial, commercial, and institutional sectors.

The technology for cogeneration exists for sizes ranging from less than 100 kilowatts to over 100 megawatts. The major equipment requirements include a prime mover, electrical generator, electrical controls, heat-recovery systems, and other typical power-plant equipment. These components are well developed, and the procedures to integrate them into cogeneration systems are well established.

Cogeneration is not new. In the early twentieth century, many industrial facilities generated their own electrical power and often supplied power to nearby communities. They used the thermal energy that was available during the electrical power production to provide or supplement process or building heat. These industrial facilities were the first "cogenerators." As electricity became widely available and less expensive, facilities often stopped generating their own electrical power. It is estimated that from 1954 to 1976 industrial cogenerated electrical power decreased from about 25% to 9% of the total electrical power generated in the United States. Since about the mid-1980s, cogenerated electrical power has been fairly constant at about 5%.

Although arguments for cogeneration technology are persuasive, a number of obstacles constrain the implementation of these systems, including the relatively low cost of electricity, high cost of fuels, high cost of capital investment, environmental regulations, and difficulties working with local utilities. In many cases, these obstacles can be minimized or eliminated, and cogeneration facilities can be developed.

Topping and bottoming cycles. A cogeneration system may be classified as either a topping cycle system (**Fig. 1**) or a bottoming cycle system (**Fig. 2**). For topping-cycle systems, a prime mover uses fuel to power an electrical generator to produce electricity. In one configuration, the hot exhaust gases are directed to a heat-recovery steam generator (HRSG) to produce steam or hot water. This cogeneration system is classified as a topping cycle because the electrical power is generated first at the higher (top) temperature associated with the fuel combustion process, and then the rejected or exhausted energy is used to produce useful thermal energy (such as the steam or hot water in this example). The majority of cogeneration applications are based on topping-cycle systems.

In bottoming-cycle systems (Fig. 2), the high-temperature combustion gases are used first in a high-temperature thermal process (such as high-temperature metal treatment), and then the lower-temperature gases are used in a special low-temperature cycle to produce electrical power. Bottoming-cycle cogeneration systems have fewer applications than topping-cycle systems, and must compete with waste-heat-recovery systems, such as feedwater heaters, recuperators, and process heat exchangers. One of the difficulties with bottoming-cycle systems is the low-temperature electrical power–producing cycle. These cycles are generally much less efficient than conventional-power cycles, often involve special equipment, and use working fluids that are more expensive.

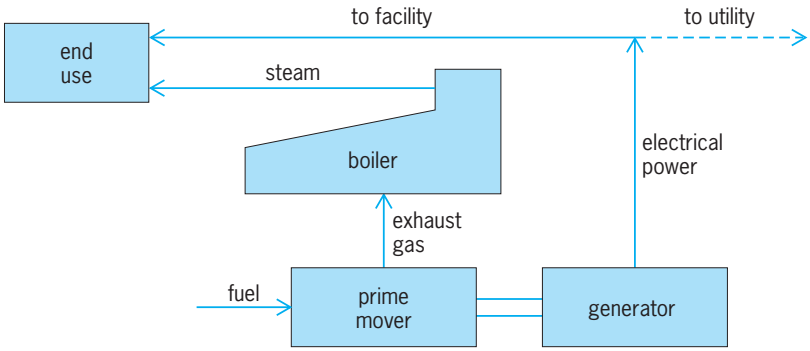

Fig. 1. Cogeneration topping-cycle system.

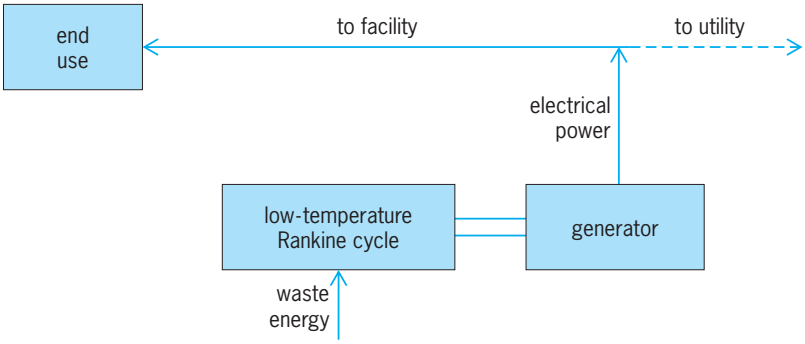

Fig. 2. Cogeneration bottoming-cycle system.

Combined cycles. One power-plant configuration, known as a combined cycle, is based on a form of a topping cycle and is widely used in industry and by electrical utilities. Typically, a gas turbine generates electrical power, and the exhaust gas is ducted to a partially fired heat-recovery boiler. The produced steam drives a steam turbine that produces additional electrical power. In this example, the majority of the steam is then condensed and pumped back to the boiler. To classify as a cogeneration application, some steam would need to be used to satisfy a thermal requirement.

Heat-recovery equipment. Several configurations of heat-recovery devices, or heat-recovery steam generators (HRSGs), are available. HRSGs are often divided into unfired, partially fired, and fully fired. An unfired HRSG is essentially a convective heat exchanger. A partially fired HRSG may include a "duct burner," which often uses a natural-gas burner upstream of the HRSG to increase the exhaust gas temperature. A fully fired HRSG is a boiler that simply uses the exhaust gas as preheated air. These devices can produce hot water, saturated steam, or superheated steam at a range of pressures and temperatures, depending on the need. *See* CONVECTION (HEAT); HEAT EXCHANGER; HEAT TRANSFER.

For chilled water (such as for building cooling), absorption chillers may be used. Absorption chillers use special fluids and a unique thermodynamic cycle to produce low temperatures without the requirement of a vapor compressor that is used in mechanical chillers. Instead of the vapor compressor, an absorption chiller uses liquid pumps and energy from low-temperature sources, such as hot water, steam, or exhaust gas.

Governmental regulations. In 1978 the United States government passed the National Energy Act (NEA), which included the Public Utility Regulatory Policies Act (PURPA). PURPA defined cogeneration systems to include those power plants that supplied a specified fraction of their input energy as useful thermal output in addition to a mechanical or electrical output. For a cogenerating facility to qualify for the privileges and exclusions specified in PURPA, the facility must meet or exceed certain legislated standards. A qualified facility is eligible to use the provisions outlined in PURPA regarding nonutility electrical power generation. In particular, these provisions enhance the interactions between cogenerators and utilities.

Other regulatory legislation directed at managing air and water quality significantly affects the installation of cogeneration systems. Based on the type of fuel and generation process incorporated, the air pollutants regulated for a specific project include nitrogen oxides (NOx), sulfur oxides (SOx), hydrocarbons (HC), carbon monoxide (CO), and particulates. The regulation of emissions from power generation is determined by several different agencies at the federal, state, and local levels. The regulations in some geographical areas have become complex and numerous. For example, certain regions may need to meet limits imposed by new source review or by prevention of significant deterioration.

In addition to air quality standards, most facilities must meet water-quality and solid waste–disposal standards. Generally for these standards, compliance is routine. *See* AIR POLLUTION; WATER POLLUTION.

Pollution control. The reduction of emissions for power-generating equipment may be accomplished by a number of methods depending on the level of reductions required, the pollutant of concern, and the specific equipment. Generally, emissions control equipment either reduces the initial amount of the pollutants generated, referred to as combustion modification control, or reduces the amount of pollutants in the combustion exhaust gas, known as postcombustion control. For gas turbines, combustion modification methods to reduce nitric oxides include steam or water injection, lean premixed combustion, and combustion designs known as "low NOx burners." An example of postcombustion control is selective catalytic reduction, which uses a catalyst and the injection of an ammonia species in the exhaust gas. At the appropriate temperature and conditions, the catalyst reduces nitric oxide to molecular nitrogen (N_2).

Project development/financing. Even with a reliable technical design and favorable economics, the successful completion of a cogeneration project often depends on acceptable financial arrangements. Financing is critical to the success of a cogeneration project, and it is best to determine early on the financial arrangement to be used. The selected financial arrangement will be intimately linked to the ownership structure. The ownership may be structured in a variety of ways. By increasing the number of participants in the project, the individual risk decreases, but the venture is more difficult to organize and the individual potential gains decrease. A cogeneration system may be owned by the facility, by a third-party entity, or by a partnership. In this context, the facility refers to the entity with the electrical and thermal loads. This is sometimes called the thermal consumer, heat consumer, or thermal host. A third party refers to an entity separate from the thermal load owner (first party) and from the local utility (second party).

The thermal and electrical loads of a facility are the items that might motivate others to be involved in developing a cogeneration system. In one respect, these loads represent the opportunity for financial gains. There are a number of potential project participants that could conceivably have an interest in the development of the cogeneration system. These potential participants include equipment manufacturers, power-plant operators, investors (such as banks, insurance companies and pension funds), electrical utilities, fuel suppliers, engineering firms, and governmental agencies. In general, the ownership structure and the financial arrangement are selected to provide the highest profits, lowest risks, and greatest tax benefits.

External participants, partners, and investors will examine a number of issues before deciding to be involved in the project. These issues include the overall

economics, the revenue from thermal and electricity sales, the accuracy of the capital costs and operating expenses, the experience of the participants, the projected availability of the plant, the previous success of the proposed technology, the assurance that permits, contracts, and agreements will be obtained on a timely basis, and the availability and cost of fuel. A net positive assessment of these issues will be necessary for any external participation. Jerald A. Caton

Bibliography. G. R. Baxter and J. A. Caton, Technical and economic assessments of cogeneration systems: An overview of the development and application of a generalized computer simulation, *Proceedings of the 1997 Power-Gen International Conference*, Dallas, Dec. 9–11, 1997; M. P. Boyce, *Handbook for Cogeneration and Combined Cycle Power Plants*, American Society of Mechanical Engineers, New York, 2002; J. H. Horlock, *Cogeneration: Combined Heat and Power (Thermodynamics and Economics)*, Krieger Publishing, 1997; J. A. Orlando, *Cogeneration Planner's Handbook*, Fairmont Press, Lilburn, GA, 1991; S. A. Spiewak and L. Weiss, *Cogeneration and Small Power Production Manual*, 5th ed., Fairmont Press, Lilburn, GA, 1997.

Cognition

The internal structures and processes that are involved in the acquisition and use of knowledge, including sensation, perception, attention, learning, memory, language, thinking, and reasoning. These behaviors are studied by cognitive scientists, a group that includes, but is not limited to, researchers in the areas of cognitive psychology, philosophy, linguistics, computer science, and cognitive neuroscience. Cognitive scientists propose and test theories about the functional components of cognition based on observations of an organism's external behavior in specific situations.

Sensation, perception and attention. Cognition throughout life can be broadly described as an interaction between top-down, knowledge-driven processes and bottom-up, sensory processes, and between controlled processes and automatic processes. When environmental input impinges on an organism's sensors, such as eyes or ears, that sensory information is used to construct an internal representation of the object or event causing the input. This representation undergoes several transformations as more knowledge is activated to make sense of an event. Over time, there is a trade-off between the amount of surface information that is retained in the representation (bottom-up processing) and the amount of meaning that is incorporated (top-down processing).

Following exposure to a stimulus, a sensory representation (sometimes called an image, icon, or echo) is constructed that encodes nearly all the surface characteristics of the stimulus (for example, color, shape, location, pitch, and loudness). The information is short lived, lasting less than a second. Much evidence suggests that extraction of information from this veridical representation takes place in two stages, a feature analysis stage and an object recognition stage. *See* MEMORY; PERCEPTION.

The first stage, sometimes called a preattentive stage, involves the segregation of pattern subcomponents into figures and background by feature analysis. This process is sometimes called parsing the input. A small number of primitive features are extracted from the sensory patterns, including color, size, contrast, tilt, curvature, line ends, movement, stereoscopic depth, and closure. These features are extracted independently by different parts of the visual system, and only in the second stage are they combined to form complete images of three-dimensional objects. For example, if a matrix of T's is juxtaposed next to a matrix of tilted T's, the boundary between the vertical and tilted T's stands out for the viewer. The boundary between juxtaposed matrices of T's and L's, however, does not. These different effects indicate that line orientation (vertical or tilted) is an important feature in early visual processing but the conjunctions of lines (that is, the way vertical and horizontal lines are combined to make up T's and L's) are not. A more surprising feature of object recognition and feature analysis is closure. When jumbles of tilted lines and angles are presented very briefly (200 ms) to viewers, they seldom report seeing triangles. But if small circles (that is, closed geometric figures) are distributed among the lines and angles, the viewers frequently report seeing triangles. When parsing normal triangles, the visual system apparently detects tilted lines, angles, and the abstract feature "closed" (a triangle is a closed geometric figure). The circles in the latter display provide this feature to the viewers' visual systems, along with the lines and angles, allowing them to see illusory triangles. These primitive features, once extracted from the visual input, are combined in the second stage to form objects. It is during this stage that attention (controlled processing) and previous knowledge come into play. For example, the time it takes to indicate whether a slide of O's also contains a Q increases linearly with the number of O's in the display. This relationship suggests that viewers scan the slide serially, focusing attention on each object or location in turn in order to identify each object. Viewers can complete these tasks faster and more accurately when familiar rather than unfamiliar objects are used, demonstrating the influence of top-down processes. Selecting particular objects from among others requires selective attention, that is, the ability to focus on certain inputs and ignore or inhibit processing of others.

Learning and memory. Information from the second-stage image rapidly decays. Typically, no more than five to nine pieces of information can be identified and maintained from the stimulus presented. Aspects of this short-term memory representation that become permanent constitute long–term memory. These aspects include items that mark the beginning and end of events, repeated events, and events that are particularly meaningful to the individual.

There are two independent memory systems. The first is called explicit or declarative memory. It comprises memories that can be consciously recollected, such as facts, episodes, and images. Explicit memories have the self as the agent or experiencer of the event in question and require controlled processing in order to be formed and retrieved. Implicit or nondeclarative memories, in contrast, constitute information acquired during skill learning, habit formation, classical conditioning, emotional learning, and priming. Implicit memories are expressed through performance rather than through conscious recollection. They do not necessarily depend on awareness or attention for their formation or retrieval (automatic processing). While learning to drive a car, a person focuses attentional resources on each subcomponent, such as depressing the clutch, shifting gears, and checking mirrors and blind spots. Once each skill is mastered, however, driving proceeds almost effortlessly, requiring few attentional resources. Such automatic skills or procedures are easier to do than to describe because they consist of memorized sequences of interconnected motor commands that are coordinated for action and executed as a single, fluid movement. *See* LEARNING MECHANISMS.

Conceptual knowledge. Conceptual knowledge is needed to classify objects and events in the world. Some aspects of conceptual knowledge are innate or emerge very early in development, while others are acquired through learning and inference. Advances in the understanding of innate knowledge have come about through an experimental procedure called the habituation technique. In this technique, a preverbal infant or animal is presented with a novel stimulus, which usually elicits a change in some observable behavior, such as directed looking, electrical skin conductance, brain wave activity, or changes in heart rate and blood pressure. The stimulus is presented repeatedly until the response in question declines to half of its initial value. At this point, the subject is said to have habituated to the stimulus. A new, similar stimulus is then presented. If the subject dishabituates and response levels of the subject return to their prehabituation levels, the subject is said to have detected the difference between the old and new stimuli.

By using this technique, human infants have been found to be able to detect differences in small numerosities shortly after birth, indicating innate sensitivity to the concept of number. Six-month-olds have been shown to recognize numerical equivalence across sense modalities, and infants as young as 5 months are capable of adding and subtracting small numbers. By 3 months, infants realize that solid objects cannot occupy the same location at the same time, cannot pass through each other, continue to exist even when out of sight, and cannot influence the trajectory of other objects without direct contact (no action at a distance). In contrast, they do not seem to appreciate the properties of gravity or inertia until about 9 months of age.

Also early to emerge is a sensitivity to ontological category. Studies that explore this type of knowledge typically employ verbal tasks, so it is not clear how early such knowledge develops. By the age of 2 years, however, infants make clear distinctions in their inductive generalizations between objects and substances; by 3 or 4 years, sensitivity is evidenced to distinctions among plants, animals, nonliving natural objects, and human-made artifacts. For example, 4-year-olds believe that a coffee pot can become a bird feeder with a little surgery, but they do not believe that a porcupine can become a cactus. Similarly, if they are told that the human body contains an omentum, they believe that dogs are likely to have one, too, but not human-looking dolls, flowers, or rocks.

Categorization learning is augmented by bottom-up processes, such as grouping together objects that share detectable common features or uses. These groupings typically show graded structure, with objects containing many typical features forming the core, or prototype, of the category. For example, when subjects are asked to list features that typify birds, sparrows, and penguins, the lists generated for sparrow and bird are found to have more features in common than the lists generated for penguin and bird, indicating that sparrows are more typical of the category "bird" than penguins. Because sparrows are closer to the prototype for a bird, people typically classify sparrows as birds more quickly and with fewer errors than they classify penguins as birds. Early on, individuals become sensitive to the ontological distinctions among living things, but learn to form precise categories for members within these distinctions by detecting common features.

Communicating, reasoning, and problem solving. A primary cognitive function of all social species is communication, which can be accomplished by a combination of vocal, gestural, and even hormonal signals. Of all species on Earth, only humans have developed a communication system based on abstract signs. This evolutionary development is closely tied to the greater reasoning capacity found among members of the species as well. All reasoning can be broadly described as pattern recognition and search. People search their conceptual knowledge for relevant information in order to draw a conclusion, solve a problem, or guide behavior. Thinking often takes the form of a chain of associations among concepts in long-term memory, with one thought retrieving others to which it is related. The most common reasoning strategies include direct retrieval, imaging, means-ends analysis, analogy, classification, deduction, and formal procedures.

Reasoning by direct retrieval involves retrieving a known fact from memory to solve a problem, such as determining how much three times three is or where something is located. Reasoning imagistically involves constructing or retrieving images from conceptual memory and examining or manipulating them to solve a problem. For example, individuals reason imagistically when they determine how many windows there are in their living rooms by retrieving an image of the room and counting the windows in the image.

Means-ends analysis is typically employed when solving problems in unfamiliar domains. When a solution is not immediately apparent, reasoners typically compare the goal to the current situation and select means with which to reduce the differences between the two situations. A simple example is getting into a car. The current situation involves an individual standing outside a car with the door locked. The goal situation involves sitting inside the car. The primary difference between the two is the individual's position, inside versus outside the car. To reduce this difference requires opening the car door. If the door is locked, a subgoal is generated to unlock the door. The goal state is an unlocked door, and the current state is a locked door. A search is made for an operator, that is, a means to achieve the end of opening the door. Possible operators includes keys, coat hangers, and bricks. Applying the operator (such as the car key) accomplishes the subgoal of opening the door.

Insight involves the restructuring of a problem representation that allows an available means to be used in a novel way or a seemingly unrelated bit of knowledge to be accessed to solve the problem. For example, scissors can be used not just to cut a string in half but also to serve as a weight to set one of the pieces swinging like a pendulum—a handy solution if a person needs to pass the string or scissors to someone standing out of arm's reach.

Reasoning by analogy is used when a current situation allows an individual to recall another, similar situation that has a known solution or other information relevant to the task at hand. It is a technique that is powerful but error prone. Without proper constraints on the analogy process, irrelevant or incorrect information can be transferred between the target and source analogs. For example, if told "The structure of the atom is like the solar system," the reasoner might also conclude that the nucleus of the atom is hot, just like the Sun.

Reasoning by classification involves making inferences about an object or event based on its category membership. For example, when asked to estimate the likelihood that each of a series of descriptions refers to an engineer or a lawyer, people have been found to base their estimates on the similarity of the descriptions to stereotypes for these professions—regardless of the base rates for the two types of professionals in the pool. Persons described as being interested in mathematical puzzles and reserved in manner were assigned a high likelihood of being engineers, regardless of whether the percentage of engineer descriptions in the pool was reported to be 30% or 70%. Neutral descriptions, such as successful or average weight and height, were judged to have a 50% probability of being an engineer, again regardless of whether subjects were told the percentage of engineers in the pool was 30% or 70%. Judgments like these are called reasoning fallacies because they violate formal rules of reasoning. Ignoring base rates violates Bayes' theorem, a formal model for computing conditional probabilities.

Deductive reasoning involves drawing a conclusion based on its logical relation to one or more premises. A common example is Modus Ponens (If p, then q/p/Therefore, q). Sufficient evidence exists that the Modus Ponens reasoning rule is part of cognitive architecture. Errors on this type of argument rarely occur, judgments are extremely rapid even when other tasks are performed concurrently, and evidence of use of Modus Ponens emerges early in development and independently of formal training. Such unequivocal and converging evidence exists for a scant few other formal rules of deduction.

A second common use for deduction is testing hypotheses. Naive reasoners typically commit a reasoning fallacy termed confirmation bias when testing their hypotheses. For example, when asked to determine the rule used to generate the number series 2, 4, 6, most subjects hypothesized that the rule was even ascending numbers. Importantly, they tested this hypothesis by seeking to confirm their hypothesis rather than trying to disprove it. They did so by generating only instances that were consistent with their hypothesis (8, 10, 12; 22, 24, 26). The correct rule was any ascending sequence. Discovering this required testing a sequence that disconfirmed their hypothesis.

Formal procedures for reasoning and for solving problems have been developed over the centuries. These include logic, mathematics, probability theory and statistics, and scientific investigation. Understanding of the behavior and properties of physical, biological, and cognitive systems has been greatly enhanced through the use of these techniques. *See* PSYCHOLINGUISTICS.

Cognition and the brain. By using noninvasive techniques such as positron emission tomography (PET scan), magnetic resonance imaging, electrical skin conductance, invasive surgical and chemical investigations of animal brains, and data from clinically observed syndromes associated with brain injury, cognitive neuroscientists have pieced together a good deal of information concerning the role that specific brain regions play in the processing of emotional and cognitive events. High-level visual processing, such as object recognition, takes place in the occipital lobes of the cortex (back of the head), although recognition of certain highly complex visual stimuli, such as faces, is handled by the right cerebral hemisphere. Auditory stimuli in general are processed by the temporal lobes of the cortex (near the ears), and written and spoken word recognition and syntactical components of language processing are handled by certain regions of the left hemisphere of the cerebral cortex, notably Broca's and Wernicke's areas; while emotional, idiomatic, and prosodic aspects of language are handled by corresponding regions in the right hemisphere. Higher cognition, such as reasoning and problem solving, involves the frontal lobes of the cortex. Memory and the processing of emotional stimuli are handled by the combined effort

of the cortex (notably the anterior and frontal regions) and subcortical structures (notably the limbic system).

One particular subcortical structure—the hippocampus—plays a major role in the formation of new explicit memories. By using positron emission tomography scanning, conscious recollection of events can be seen to involve increased blood flow to the hippocampal region, while implicit memory tasks produce increased blood flow to posterior, extrastriate areas of the cerebral cortex. Damage to the hippocampal region and the medial temporal lobe produces profound amnesic syndrome in which the patient displays normal perceptual linguistic and intellectual functioning but cannot consciously remember recent events (such as having lunch or playing a round of golf). Such patients do, however, show implicit memory for recent events, including the acquisition of new motor skills, habits, and emotional preferences. It is believed that an intact hippocampus is needed to temporarily bind together distributed sites of activation in the cortex that together make up a whole, explicit memory for an event. *See* BRAIN; COMPUTERIZED TOMOGRAPHY.

Computer models. Theories of cognition are often tested by building computer models that embody the theories and then comparing the model's performance with human performance on selected tasks. These models tend to be of two types. Rule-based models simulate cognition as the execution of condition-action rules. The models consist of a long-term memory containing rules which specify actions to take in the presence of particular input patterns, a short-term memory that encodes input patterns and temporarily stores data structures constructed by the rules, and a control structure that guides the process and resolves conflicts when more than one rule applies to the current input. Rule-based models have successfully simulated expert performance in a variety of domains. Neural network models simulate cognition as a strengthening and weakening of associations among cognitive events. They are called neural networks because they learn in a way that is analogous to the way in which brains learn (by modifying synaptic connections among neurons) and have proven to be excellent models of perception and memory. Neural networks consist of a network of interconnected nodes, a mathematical formula for modifying the connections, and a mathematical formula for propagating activation through the network. *See* EXPERT SYSTEMS; INFORMATION PROCESSING (PSYCHOLOGY); INTELLIGENCE; NEURAL NETWORK; PROBLEM SOLVING (PSYCHOLOGY).

Denise Dellarosa Cummins

Bibliography. J. R. Anderson, *Cognitive Psychology and Its Implications*, 5th ed., 1999; D. D. Cummins, *The Other Psychology*, 1994; Y. C. Lee, *Evolution, Learning and Cognition*, 1988; N. Stillings et al., *Cognitive Science: An Introduction*, 2d ed., 1995.

Coherence

The attribute of two or more waves, or parts of a wave, whose relative phase is nearly constant during the resolving time of the observer. The concept has been developed most extensively in optics, but is applicable to all wave phenomena.

Coherence of two beams. Consider two waves, with the same mean angular frequency ω, given by Eqs. (1) and (2). These expressions as they stand

$$\Psi_A(x,t) = A \exp\{i[k(\omega)x - \omega t - \delta_A(t)]\} \qquad (1)$$

$$\Psi_B(x,t) = B \exp\{i[k(\omega)x - \omega t - \delta_B(t)]\} \qquad (2)$$

could describe de Broglie waves in quantum mechanics. For real waves, such as components of the electric field in light or radio beams, or the pressure oscillations in sound, it is necessary to retain only the real parts of these and subsequent expressions. Assume that the frequency distribution is narrow, in the sense that a Fourier analysis of Eqs. (1) and (2) gives appreciable contributions only for angular frequencies close to ω. This assumption means that, on the average, $\delta_A(t)$ and $\delta_B(t)$ do not change much per period. For any actual wave, however, they must undergo some change; only waves that have existed forever and that fill all of space can have absolutely fixed frequency and phase. *See* ELECTROMAGNETIC RADIATION; QUANTUM MECHANICS; SOUND.

Suppose that the waves are detected by an apparatus with resolving time T; that is, T is the shortest interval between two events for which the events do not seem to be simultaneous. For the human eye and ear, T is about 0.1 s, while a fast electronic device might have a T of 10^{-10} s. If the relative phase $\delta(t)$, given by Eq. (3), does not, on the average, change

$$\delta(t) = \delta_B(t) - \delta_A(t) \qquad (3)$$

noticeably during T, then the waves are coherent. If during T there are sufficient random fluctuations for all values of $\delta(t)$, modulus 2π, to be equally probable, then the waves are incoherent. If during T the fluctuations in $\delta(t)$ are noticeable, but not enough to make the waves completely incoherent, then the waves are partially coherent. These distinctions are not useful unless T is specified. Two waves for which the relative phase changes appreciably in 0.01 s would seem incoherent to unaided human perception, but would seem highly coherent to a fast electronic device.

The degree of coherence is related to the interference patterns that can be observed when the two beams are combined. The following variations of Young's two-slit interference experiment illustrate several possibilities. *See* INTERFERENCE OF WAVES.

Young's two-slit experiment. In the usual Young's experiment (**Fig. 1**) beams A and B come from the same point source and pass through separated slits in a screen. Near a point P equidistant from the slits, for linear wave phenomena, the resultant wave $\Psi(x,t)$ is the sum of $\Psi_A(x,t)$ and $\Psi_B(x,t)$. The observable

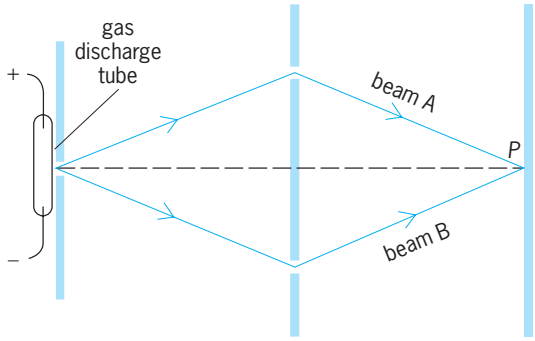

Fig. 1. Young's experiment. If beams *A* and *B* come from the same point source and traverse the same distance, they are coherent around *P* and produce there an interference pattern that has high visibility.

intensity is proportional to $|\Psi|^2$, the square of the magnitude of Ψ. This observable intensity is the energy density or the mean photon density for electromagnetic waves, the energy density for acoustic waves, and the mean particle density for the wave functions of quantum mechanics. With the wave forms of Eqs. (1) and (2), for real A and B, it is given by Eq. (4). The first term gives the intensity of beam A

$$\begin{aligned} |\Psi|^2 &= |\Psi_A + \Psi_B|^2 \\ &= A^2 + B^2 + 2AB \cos \delta(t) \end{aligned} \tag{4}$$

alone, the second gives the intensity of B alone, and the third term depends on the relative phase, given by Eq. (3).

The mean life of a typical excited atomic state is about 10^{-8} s. Collisions and thermal motion reduce the effective time of undisturbed emission to around 10^{-10} or 10^{-11} s in standard discharge tubes. The phase $\delta_A(t)$ of beam A therefore dances about erratically, with substantial changes occurring perhaps 10^{11} times per second. However beam B comes from the same atoms and travels nearly the same distance, and the changes in $\delta_B(t)$ are the same as those in $\delta_A(t)$. The relative phase is always zero at the exact midpoint P and takes on other time-independent values in the neighborhood of P. For each point in that neighborhood, $|\Psi|^2$ has a constant value that satisfies inequality (5). A clear interference pattern

$$|A - B|^2 \le |\Psi|^2 \le |A + B|^2 \tag{5}$$

is observed even with a large-T detector such as a photographic plate.

Independent sources. Instead of a single source, two independent sources can be used to produce the two beams (**Fig. 2**). If the two sources are standard discharge tubes, $\delta_A(t)$ and $\delta_B(t)$ independently change erratically around 10^{11} times per second. Equation (4) is still valid, but the observed quantity is $|\Psi|^2$ averaged over the resolving time T, denoted by $\langle|\Psi|^2\rangle_T$. When the average of Eq. (4) is formed, the last term on the right contributes nothing because $\cos \delta(t)$ randomly takes on values between $+1$ and -1 and thus obeys Eq. (6). The two beams are incoherent, the observed intensity is simply the sum of the separate intensities, and the superposition does not produce a visible interference pattern. In general, if incoherent waves $\Psi_A, \Psi_B, \Psi_C, \ldots$, are combined, much the same argument applies because the average over T of each cross term is zero, so that the observed intensity is given by Eq. (7).

$$\langle\cos \delta(t)\rangle_T = 0 \tag{6}$$

$$\langle|\Psi|^2\rangle_T + \langle|\Psi_A|^2\rangle_T + \langle|\Psi_B|^2\rangle_T + \langle|\Psi_C|^2\rangle_T + \cdots \tag{7}$$

The argument that results in Eq. (6) depends on there being random independent changes in the separate phases during times much shorter than T. If the discharge tubes are replaced by very well stabilized lasers, or by loudspeakers or radio transmitters, then the separate phases $\delta_A(t)$ and $\delta_B(t)$ can be made nearly constant during 0.1 s. The relative phase $\delta(t)$ is then also nearly constant, an interference pattern is visible, and the two waves can be called coherent. *See* LASER.

Effect of path length difference. In another arrangement (**Fig. 3**), the two beams are again emitted by the same

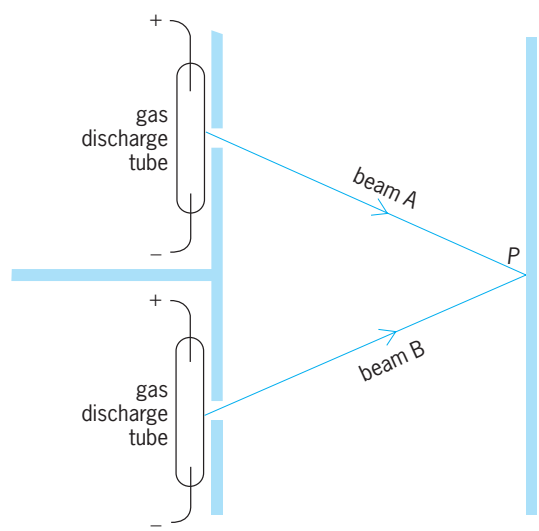

Fig. 2. Independent sources that change their phases frequently and randomly during the resolving time *T* are incoherent, and their superposition does not produce a visible interference pattern.

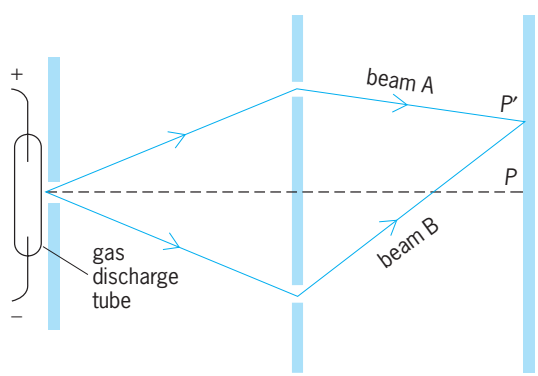

Fig. 3. If the difference between the distances traversed by beams *A* and *B* is of the order of the coherence length, then the beams are partially coherent at *P′*.

point source and are therefore coherent at their origin. Here, however, the distance from the source to the point P' of wave addition is larger along beam B than along beam A by an amount l. In this arrangement, the path length difference is due to the observation point P' not being near the central axis, but it could equally well be caused by the insertion of different optical devices in the two beams.

The waves that arrive at P' via path B at any instant must have left the source earlier by a time $\tau = l/c$, where c is the wave speed, than those that arrive via path A. Suppose again that the source is a standard discharge tube, so that the phase changes randomly about 10^{11} times per second. For $\tau < 10^{-11}$ s, the coherence between A and B at P' is not spoiled by the path length difference, and there is a visible interference pattern around P'. For $\tau > 10^{-11}$ s, the light that arrives via A left the source so much later than that which arrives via B at the same instant that the phase changed many times in the interim. The two beams are then incoherent, and their superposition does not produce a visible interference pattern around P'. For $\tau \simeq 10^{-11}$ s, perfect coherence is spoiled but there is less than complete incoherence; there is partial coherence and an interference pattern that is visible but not very sharp. The time difference that gives partial coherence is the coherence time Δt. The coherence time multiplied by the speed c is the coherence length Δl. In this example, Δl is given by Eq. (8).

$$\Delta l \simeq 10^{-11}\text{s} \times 3.0 \times 10^{10}\ \text{cm/s} = 0.3\ \text{cm} \qquad (8)$$

Effect of extended source. The original arrangement can also be modified by making a large hole in the collimator in front of the discharge tube, so that there is an extended source rather than a point source (**Fig. 4**). Consider two regions of the source, S near the center and S' near one end, separated by a distance y. During times longer than the coherence time, the difference between the phases of the radiation from the two regions changes randomly. Suppose that, at one moment, the radiation from S happens to have a large amplitude while that from S' happens to be negligible. The waves at the two slits are then in phase because S is equidistant from them, and interference at P gives a maximum. Suppose that a few coherence times later the radiation from S happens to be negligible while that from S' happens to have a large amplitude. If the path length along A' differs from that along B' by half of the wavelength λ, then the waves at the two slits are out of phase. Now interference at P gives a minimum, and the pattern has shifted in a few coherence times. If the resolving time T is much larger than the coherence time Δt, the consequence is that beams A and B are not coherent at P.

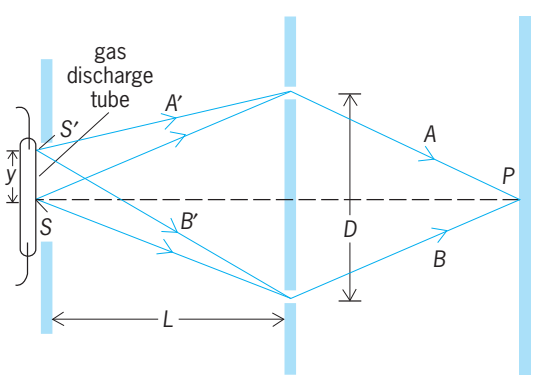

Fig. 4. If the lateral extent y of the source is of the order of $L\lambda/D$, then the beams are partially coherent at P.

The condition that there be a path length difference of $\lambda/2$ between B' and A' is given by Eq. (9),

$$\sqrt{L^2 + (y + D/2)^2} - \sqrt{L^2 + (y - D/2)^2} = \lambda/2 \qquad (9)$$

and for D and y both much smaller than L, expansion of the square roots gives Eq. (10). This argument is

$$2Dy = L\lambda \qquad (10)$$

rather rough, but it is clear that a source that satisfies Eq. (11) is too large to give coherence and too

$$y \simeq \frac{L\lambda}{D} \qquad (11)$$

small to give complete incoherence; it gives partial coherence.

Coherence of a single beam. Coherence is also used to describe relations between phases within the same beam. Suppose that a wave represented by Eq. (1) is passing a fixed observer characterized by a resolving time T. The phase δ_A may fluctuate, perhaps because the source of the wave contains many independent radiators. The coherence time Δt_W of the wave is defined to be the average time required for $\delta_A(t)$ to fluctuate appreciably at the position of the observer. If Δt_W is much greater than T, the wave is coherent; if Δt_W is of the order of T, the wave is partially coherent; and if Δt_W is much less than T, the wave is incoherent. These concepts are very close to those developed above. The two beams in the arrangement with independent sources are incoherent with respect to each other if one or both have $\Delta t_W < T$, but are coherent with respect to each other if both have $\Delta t_W > T$ (Fig. 2).

The degree of coherence of a beam is of course determined by its source. Discharge tubes that emit beams with Δt_W small compared to any usual resolving time are therefore called incoherent sources, while well-stabilized lasers that give long coherence times are called coherent sources.

Consider two observers that measure the phase of a single wave at the same time. If the observers are close to each other, they will usually measure the same phase. If they are far from each other, the phase difference between them may be entirely random. The observer separation that shows the onset of randomness is defined to be the coherence length of the wave, Δx_W. If this separation is in the direction of propagation, then Δx_W is given by Eq. (12), where

$$\Delta x_W = c\Delta t_W \qquad (12)$$

c is the speed of the wave. Of course, the definition

of Δx_W is also applicable for observer separations perpendicular to the propagation direction.

The coherence time Δt_W is related to the spectral purity of the beam, as is shown by the following argument. The wave can be viewed as a sequence of packets with spatial lengths around Δx_W. A typical packet requires Δt_W to pass a fixed point. Suppose that n periods of duration $2\pi/\omega$ occur during Δt_W, as expressed in Eq. (13). The packet can be

$$\Delta t_W \frac{\omega}{2\pi} = n \tag{13}$$

viewed as a superposition of plane waves arranged to cancel each other outside the boundaries of the packet. In order to produce cancellation, waves must be mixed in which are in phase with the wave of angular frequency ω in the middle of the packet and which are half a period out of phase at both ends. In other words, waves must be mixed in with periods around $2\pi/\omega'$, where Eq. (14) holds. The difference

$$\Delta t_W \frac{\omega'}{2\pi} = n \pm 1 \tag{14}$$

of Eqs. (13) and (14) gives Eq. (15), where the spread

$$\Delta t_W \Delta\omega = 2\pi \tag{15}$$

in angular frequencies $|\omega - \omega'|$ is called $\Delta\omega$. This result should be stated as an approximate inequality (16). A large Δt_W permits a small $\Delta\omega$ and therefore

$$\Delta t_W \Delta\omega \gtrsim 1 \tag{16}$$

a well-defined frequency, while a small Δt_W implies a large $\Delta\omega$ and a poorly defined frequency.

Quantitative definitions. To go beyond qualitative descriptions and order-of-magnitude relations, it is useful to define the fundamental quantities in terms of correlation functions.

Self-coherence function. As discussed above, there are two ways to view the random fluctuations in the phase of a wave. One can discuss splitting the wave into two beams that interfere after a difference τ in traversal time, or one can discuss one wave that is passing over a fixed observer who determines the phase fluctuations. Both approaches concern the correlation between $\Psi(x,t)$ and $\Psi(x, t + \tau)$. This correlation is described by the normalized self-coherence function given by Eq. (17). The brackets

$$\gamma(\tau) \equiv \frac{\langle \Psi(x, t+\tau)\Psi^*(x, t)\rangle}{\langle \Psi(x, t)\Psi^*(x, t)\rangle} \tag{17}$$

$\langle\ \rangle$ indicate the average of the enclosed quantity over a time which is long compared to the resolving time of the observer. It is assumed here that the statistical character of the wave does not change during the time of interest, so that $\gamma(\tau)$ is not a function of t.

With the aid of the self-coherence function, the coherence time of a wave can be defined quantitatively by Eq. (18). That is, Δt_W is the root-mean-squared

$$(\Delta t_W)^2 \equiv \frac{\int_{-\infty}^{+\infty} \tau^2 |\gamma(\tau)|^2 \, d\tau}{\int_{-\infty}^{+\infty} |\gamma(\tau)|^2 \, d\tau} \tag{18}$$

width of $|\gamma(\tau)|^2$. The approximate inequality (16) can also be made precise. The quantity $|g(\omega')|^2$ is proportional to the contribution to Ψ at the angular frequency ω', where $g(\omega')$ is the Fourier transform of Ψ, defined by Eq. (19). The root-mean-squared width

$$g(\omega') \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \Psi(0, t) e^{-i\omega' t} dt \tag{19}$$

$\Delta\omega$ of this distribution is given by Eq. (20), where ω

$$(\Delta\omega)^2 = \frac{\int_0^{\infty} (\omega' - \omega)^2 \left|g(\omega')\right|^2 d\omega'}{\int_0^{\infty} \left|g(\omega')\right|^2 d\omega'} \tag{20}$$

is again the mean angular frequency. With these definitions, one can show that inequality (21) is satisfied.

$$\Delta t_W \Delta\omega \geq {}^1/_2 \tag{21}$$

The symbol $\gtrsim$ has become $\geq$; no situation can lead to a value of $\Delta t_W \Delta\omega$ that is less than $^1/_2$. Relation (21) multiplied by Planck's constant is the Heisenberg uncertainty principle of quantum physics, and can in fact be proved in virtually the same way. *See* NONRELATIVISTIC QUANTUM THEORY; UNCERTAINTY PRINCIPLE.

Complex degree of coherence. To describe the coherence between two different waves Ψ_A and Ψ_B, one can use the complex degree of coherence $\gamma_{AB}(\tau)$, defined by Eq. (22), where the averaging process $\langle\ \rangle$

$$\gamma_{AB}(\tau) \equiv \frac{\langle \Psi_A(t+\tau)\Psi_B{}^*(t)\rangle}{\left[\langle \Psi_A{}^*(t)\Psi_A(t)\rangle \langle \Psi_B{}^*(t)\Psi_B(t)\rangle\right]^{1/2}} \tag{22}$$

and the assumptions are the same as for the definition of $\gamma(\tau)$. If beams A and B are brought together, interference fringes may be formed where the time-averaged resultant intensity $\langle|\Psi|^2\rangle$ has maxima and minima as a function of position. The visibility of these fringes, defined by Eq. (23), can be shown, for

$$V = \frac{\langle|\Psi|^2\rangle_{\max} - \langle|\Psi|^2\rangle_{\min}}{\langle|\Psi|^2\rangle_{\max} + \langle|\Psi|^2\rangle_{\min}} \tag{23}$$

the usual interference studies, to be proportional to $|\gamma_{AB}(\tau)|$.

Examples. The concept of coherence occurs in a great variety of areas. The following are some application and illustrations.

Astronomical applications. As is shown in the above discussion, extended sources give partial coherence and produce interference fringes with visibility V [Eq. (23)] less than unity. A. A. Michelson exploited this fact with his stellar interferometer (**Fig. 5**), a modified double-slit arrangement with movable mirrors that permit adjustment of the effective separation D' of the slits. According to Eq. (11), there is

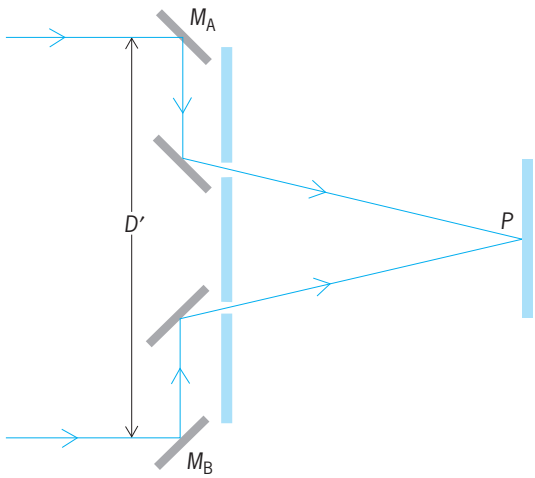

Fig. 5. Michelson's stellar interferometer. The effective slit separation D' can be varied with the movable mirrors M_A and M_B. The angular size of the source can be calculated from the resulting change in the visibility of the interference fringes around P.

partial coherence if the angular size of the source is about λ/D'. It can be shown that if the source is a uniform disk of angular diameter Θ, then the smallest value of D' that gives zero V is $1.22\lambda/\Theta$. The interferometer can therefore be used in favorable cases to measure stellar diameters. The same approach has also been applied in radio astronomy. A different technique, developed by R. Hanbury Brown and R. Q. Twiss, measures the correlation between the intensities received by separated detectors with fast electronics. *See* INTERFEROMETRY; RADIO ASTRONOMY.

Lasers and masers. Because they are highly coherent sources, lasers and masers provide very large intensities per unit frequency, and double-photon absorption experiments have become possible. As a consequence, atomic and molecular spectroscopy were revolutionized. One can selectively excite molecules that have negligible velocity components in the beam direction and thus improve resolution by reduction of Doppler broadening. States that are virtually impossible to excite with single-photon absorption can now be reached. *See* LASER SPECTROSCOPY; MASER.

Musical acoustics. Fortunately, for most musical instruments the coherence time is much less than human resolving time, so that the separate contributions in an orchestra are perceived incoherently. Equation (7) applies, and 20 violins seem to provide 20 times the intensity of 1 violin. If human resolving time were much shorter, the perceived intensity would drift between 0 and 400 times the intensity of a single violin. *See* MUSICAL ACOUSTICS.

Thermodynamic treatment of radiation. Coherence must be considered in thermodynamic treatments of radiation. As one should expect from the connection between entropy and order, entropy is reduced by an increase in coherence. The total entropy of two beams that are coherent with respect to each other is less than the sum of the entropies of the separate beams, because coherence means that the beams are not statistically independent. *See* ENTROPY; THERMODYNAMIC PRINCIPLES.

Quantum physics. In quantum physics, a beam of particles can for some purposes be described by Eq. (1) without the phase $\delta A(t)$. If the wavelength is well defined, measurables such as barrier transmission coefficients and scattering cross sections of sufficiently small targets do not depend on phases and coherence. If, however, the scattering object is large, these aspects may have to be considered.

Electrons emitted by hot filaments show coherence times of roughly 10^{-14} s. In low-energy electron diffraction, where speeds on the order of 10^8 to 10^9 cm/s are used, the coherence length is then on the order of 10^{-5} cm, and the structure can be seen clearly only if the region being studied does not introduce a path length difference greater than this distance. *See* ELECTRON DIFFRACTION.

Most nuclear and particle physics experiments deal with energies so high that the wavelengths of the projectiles are less than 10^{-12} cm, far less than the amplitudes of motions of the scattering nuclei. The many nuclei therefore contribute to the scattered beam incoherently. Equation (7) applies, and the intensity of the scattered beam is proportional to the scattering cross section per nucleus times the number of nuclei. Rolf G. Winter

Photon statistics. Photon statistics is concerned with the probability distribution describing the number of photons incident on a detector, or present in a cavity. By extension, it deals with the correlation properties of beams of light.

According to the quantum theory of electromagnetism, quantum electrodynamics, light is made up of particles called photons, each of which possesses an energy E of $\hbar\omega$, where $\hbar$ is Planck's constant divided by 2π and ω is the angular frequency of the light (the frequency multiplied by 2π). In general, however, the photon number is an intrinsically uncertain quantity. It is impossible to precisely specify both the phase $\phi = \omega t$ of a wave and the number of photons $n \approx E/(\hbar\omega)$ that it contains; the uncertainties of these two conjugate variables must satisfy $\Delta n \Delta\phi \geq {}^1/_2$. For a beam to be coherent in the sense of having a well-defined phase, it must not be describable in terms of a fixed number of particles. (Lacking a fixed phase, a single photon may interfere only with itself, not with other photons.) *See* PHOTON; QUANTUM ELECTRODYNAMICS.

The most familiar example of this uncertainty is shot noise, the randomness of the arrival times of individual photons. There is no correlation between photons in the coherent state emitted by a classical source such as an ideal laser or radio transmitter, so the number of photons detected obeys Poisson statistics, displaying an uncertainty equal to the square root of the mean. The shot noise constitutes the dominant source of noise at low light levels, and may become an important factor in optical communications as well as in high-precision optical devices (notably those that search for gravitational radiation). *See* DISTRIBUTION (PROBABILITY); ELECTRICAL NOISE.

Instead of considering light in terms of phase and photon number, one may equivalently discuss the amplitudes of the two quadrature components, cos ωt and sin ωt. Another uncertainty relation applies between these conjugate variables, implying that even in the so-called vacuum state (zero photons) these two field amplitudes cannot both be precisely zero. Shot noise may be seen as a relic of these ever-present vacuum fluctuations.

Coherent states (in which the two quadratures are determined to equal accuracy) and states in which the photon number is known precisely may both seem somehow ideal. Until the invention of the laser, essentially all light sources fell far from this ideal: In stars, candles, and light bulbs, the randomness of thermal fluctuations leads to more noise than the minimum required by the uncertainty principle. This noise may be described in terms of bunching. To every process in which a photon can be emitted spontaneously, there corresponds a process in which an already-present photon can stimulate the emission of a second photon, with probability equal to that of the spontaneous process. Once a photon is observed in a thermal beam, the probability of detecting another photon is thus twice the mean probability of photon detection in the beam as a whole. *See* BOSE-EINSTEIN STATISTICS.

Antibunching effects, in which detection of one photon makes detection of subsequent photons less likely, cannot be explained by any classical wave theory. Such behavior is precisely what would be expected for a beam of a fixed number of particles. If a single photon is incident on a beam splitter, it may be detected at either output, but never both. Single-photon states have been created in the laboratory, and their indivisibility has been observed.

So-called squeezed states are those in which the uncertainty of a given variable (for example, the number, phase, or a given quadrature) is smaller than it would be for a coherent state, while the uncertainty for the conjugate variable is larger. The noise in certain measurements is thus reduced below the so-called standard quantum limit, while the uncertainty principle is still obeyed. *See* SQUEEZED QUANTUM STATES.

Aephraim M. Steinberg

Bibliography. M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; R. P. Feynman, *QED: The Strange Theory of Light and Matter*, 1985; F. Haake, L. M. Narducci, and D. Walls (eds.), *Coherence, Cooperation and Fluctuations*, 1986; P. Hariharan, *Optical Interferometry*, 1985; R. Loudon, *The Quantum Theory of Light*, 2d ed., 1983; E. R. Pike and H. Walther (eds.), *Photons and Quantum Fluctuations*, 1989; R. G. Winter, *Quantum Physics*, 2d ed., 1986, reprint 1993.

Cohesion (physics)

The tendency of atoms or molecules to coalesce into extended condensed states. This tendency is practically universal. In all but exceptional cases, condensation occurs if the temperature is sufficiently low; at higher temperatures, the thermal motions of the constituents increase, and eventually the solid assumes gaseous form. The cohesive energy is the work required to separate the condensed phase into its constituents or, equivalently, the amount by which the energy of the condensed state is lower than that of the isolated constituents. The science of cohesion is the study of the physical origins and manifestations of the forces causing cohesion, as well as those opposing it. It is thus closely related to the science of chemical bonding in molecules, which treats small collections of atoms rather than extended systems. *See* CHEMICAL BONDING; INTERMOLECULAR FORCES.

When the constituent atoms or molecules are far separated, an attractive force is present. This force is due, directly or indirectly, to the electrostatic interactions of the negatively charged electrons and the positively charged ions. However, when the distances between the constituents become very short, a strong repulsive force results, which is due to the quantum-mechanical kinetic energy of the electrons. The competition between the attractive and repulsive forces determines the equilibrium density for the condensed state (see **illus.**). A positive pressure at a given volume means that pressure must be applied to the solid to keep it at this volume; thus the forces between the atoms are predominantly repulsive. Similarly, the pressure is negative when the forces are predominantly attractive. Both molybdenum and xenon have attractive interactions at large volumes. As the volume drops, the attractive forces reach a maximum and then begin to be canceled by the repulsive forces. The equilibrium volume of the solid is that for which the cancellation is complete, or the pressure vanishes. The equilibrium point actually occurs when the pressure equals the atmospheric pressure, which is approximately 1 bar (100 kilopascals). However, this pressure is effectively zero on the pressure scale of the illustration. *See* COULOMB'S LAW; QUANTUM MECHANICS.

Attractive forces. The origin and magnitude of the attractive forces depend on the chemical nature of the constituent atoms or molecules. Strong attractive

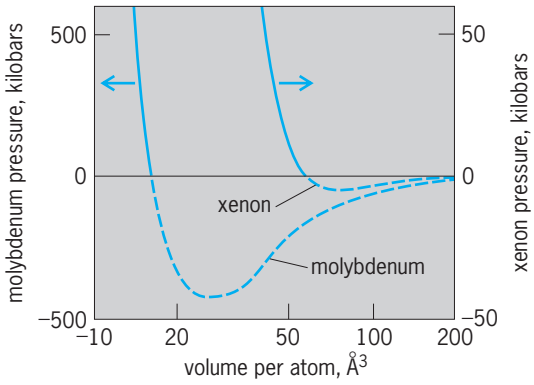

Calculated pressure-volume curves for molybdenum and xenon at absolute zero. Negative-pressure portions of curves are represented by broken lines because they cannot be directly measured. 1 kilobar = 0.1 GPa; 1 Å^3 = 0.001 nm^3.

interactions are usually associated with constituents having valence electron shells which are partly filled or open; if the valence electron shells are completely filled or closed, the interactions are weaker. *See* VALENCE.

Open-shell constituents. This is most easily understood for atomic systems. As the atoms approach, the electron energy levels on different atoms begin to interact, forming a complex of energy levels in the solid. Some of these are below the atomic energy levels and some above. Since the atomic shells are partly filled, the lower energy levels in the solid are filled, but at least some of the higher levels are empty. Thus the average energy of the occupied levels in the solid is lower than that in the isolated atoms, resulting in an attractive force. Bonding in open-shell systems can be approximately divided into three categories, although most cases involve a combination *See* BAND THEORY OF SOLIDS; FERMI-DIRAC STATISTICS; SOLID-STATE PHYSICS; VALENCE BAND.

1. *Covalent bonding*. This type of bonding is most similar to the molecular bond. The electron energy levels in the solid are split into a lower and a higher portion, with the states in the lower one filled and the higher one empty. Covalent bonds are strongly directional, with electron charge accumulating around the bond centers. Materials bonded in this fashion typically form structures with low coordination numbers, prototypical materials elements in group 14 of the periodic table, the insulator carbon, and the semiconductors silicon and germanium. *See* PERIODIC TABLE; SEMICONDUCTOR.

2. *Metallic bonding*. In this case, there is no split between the lower and higher states of the electrons in the solid; rather, they occupy levels from the bottom up to a cutoff point known as the Fermi level. For example, in transition metals, the electron states in the solid derived from the atomic *d* orbitals form a complex which is gradually filled with increasing atomic number. The bulk of the cohesive energy is due to this complex. The metallic bond is less directional than the covalent bond, with a more uniform distribution of electronic charge. Metals usually form closely packed structures. *See* FERMI SURFACE; FREE-ELECTRON THEORY OF METALS.

3. *Ionic bonding*. This occurs in compounds having at least two distinct types of atoms. One or more of the species of atoms (the cations) have only a small number of electrons in their valence shells, whereas at least one species (the anions) has a nearly filled valence shell. As the atoms approach each other, electrons drop from the cation valence states into holes in the anion valence shell, forming a closed-shell configuration in the solid. The different types of atoms in the solid have net charges; a strong attractive force results from the interaction between unlike charges. For example, in sodium chloride (NaCl), the sodium atoms acquire positive charges, and the chlorine atoms acquire negative charges. The closest interatomic separations in the solid are between sodium and chlorine, so that the attractive electrostatic interactions outweigh the repulsive ones. *See* IONIC CRYSTALS; SOLID-STATE CHEMISTRY.

Closed-shell constituents. In these systems the above effects are greatly reduced because the atomic or molecular shells are basically inert. The constituents retain their separate identities in the solid environment. If the constituents are atomic, as in rare-gas solids, the cohesion is due to the van der Waals forces. The positions of the electrons in an atom fluctuate over time, and at any given time their distribution is far from spherical. This gives rise to fluctuating long-ranged electric fields, which average zero over time, but can still have appreciable effects on neighboring atoms. The electrons on these atoms move in the direction of the force exerted by the electric field. The net result is that the interactions between unlike charges (electrons and nuclei) are increased in the solid, whereas the interactions between like charges are reduced. Thus the solid has a lower energy than the isolated atoms.

In solids made up of molecules, there are additional electrostatic interactions due to the nonspherical components of the molecular charge density. These interactions are strongest if the molecules are polar. This means that the center of the positive charge on the molecule is at a different point in space from that of the negative charge. Polar molecules, such as water (H_2O), form structures in which the positive charge on a molecule is close to the negative charges of its neighbors. For nonpolar molecules, the electrostatic interactions are usually weaker than the van der Waals forces. The nonspherical interactions in such cases are often so weak that the molecules can rotate freely at elevated temperatures, while the solid is still held together by the van der Waals forces.

Repulsive forces. The repulsive forces in the condensed phase are a dramatic illustration of the combined action of two quantum-mechanical principles, the exclusion principle and the uncertainty principle.

The exclusion principle states that the quantum-mechanical wave function for the electrons in the solid must be antisymmetric under the interchange of the coordinates of any two electrons. Consequently, two electrons of the same spin are forbidden from being very close to each other. *See* EXCLUSION PRINCIPLE.

The uncertainty principle states that if the motion of an electron is confined, its kinetic energy must rise, resulting in a repulsive force opposing the confinement. The kinetic energy due to the confinement is roughly inversely proportional to the square of the radius of the region of confinement. According to the exclusion principle, the motion of an electron in a solid is partially confined because it is forbidden from closely approaching other electrons of the same spin. Thus the uncertainty principle in turn implies a repulsive force. *See* UNCERTAINTY PRINCIPLE.

The magnitude of the repulsive force always exceeds that of the attractive force when the volume per atom becomes sufficiently small. This is most readily understood in the simple metals, which are, roughly, metals in groups 1 through 14. In these metals the electrons can be approximately treated as if

they were spread uniformly throughout the solid. If the metal is compressed, the repulsive kinetic energy due to the electron confinement is approximately proportional to d^{-2}, where d is the interatomic separation. The energy due to the attractive force, in contrast, is proportional only to d^{-1}. Thus the repulsive force due to the kinetic energy dominates if d is sufficiently small.

Variations between materials. There are enormous variations between materials in the strength of the cohesive forces. For example, the maximal attractive force in molybdenum, an open-shell solid, corresponds to a negative pressure of nearly 500,000 bars (50 gigapascals); that in xenon, a closed-shell solid, is roughly 100 times smaller (see illus.) This is reflected in a much larger equilibrium volume per atom for xenon, even though the size of an isolated molybdenum atom exceeds that of a xenon atom. The variations in the cohesive forces are also reflected in the boiling point, which is the highest temperature at which a condensed state exists. This can be over 5000°C (9000°F) in open-shell systems, and is usually higher than 200°C (400°F). In elemental closed-shell systems, in contrast, the boiling point is usually below 0°C (32°F); in helium it is −269°C (−452°F), or 4°C (7°F) above absolute zero. Many other properties correlate with the strength of the cohesive forces, including the melting point, elastic moduli, and the energies required to form defects such as vacancies and surfaces. The variations in these properties from material to material have a large impact on their technological applications. *See* CRYSTAL DEFECTS.

Weakening of attractive forces. In analyzing the mechanical properties of solids, it is often conceptually useful to think of the force on a particular atom as being the sum of independent forces due to neighboring atoms. Such a picture is justified for solids consisting of closed-shell atoms or molecules, but is often an inaccurate representation of the forces in open-shell systems, particularly in metals. The attractive force in a solid metal can be several times weaker than that in a dimer molecule, because the atoms interact too strongly with each other to maintain their separate identities; the electrons, instead of being localized on individual atoms, are in states that extend throughout the solid. The weakening of the attractive forces in the solid manifests itself in the atomic geometries around defects. For example, the attractive forces in the vicinity of vacancies or surfaces are intermediate in strength between those in the solid and those in the dimer molecule. This usually results in shorter interatomic separations around these defects, if it is possible for the bonds to contract. In addition, a major contribution to the surface tensions of solids results from changes in the attractive forces near the surface. Anders E. Carlsson

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; P. A. Cox, *The Electronic Structure and Chemistry of Solids* 1987; W. A. Harrison, *Electronic Structure and the Properties of Solids*: *The Physics of the Chemical Bond*, 1980; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; A. Sutton, *Electronic Structure of Materials*, 1993.

Coining

A cold metalworking process in a press-type die. Coining is used to produce embossed parts, such as badges and medals, and for minting of coins. It is also used on portions of a blank or workpiece to form corners, indentations, or raised sections, frequently as part of a progressive die operation. The work is subjected to high pressure within the die cavity and thereby forced to flow plastically into the die details. The absence of overflow of excess metal from between the dies is characteristic of coining and is responsible for the fine detail achieved. However, because this action requires highly accurate dies and higher-than-usual die pressures, only essential surfaces are coined. *See* METAL FORMING.

Ralph L. Freeman

Coir

A natural fiber, also known as coco fiber, obtained from the husks of the coconut (*Cocos nucifera*). Although coconuts are grown for copra throughout the tropics and into the subtropics, more than 95% of the coir exports are from Sri Lanka and India. Small amounts are exported from Thailand, Tanzania, Mexico, the Philippines, Malaysia, Kenya, Trinidad, and Tobago. A few other countries produce small amounts of fiber for their own use.

The cells of coir average about 0.028 in. (0.7 mm) in length and about 20 micrometers in width. The cells may have either blunt or pointed ends, and the cell wall has numerous pits. The lumen width and the cell wall thickness are about equal. Coir is a highly lignified fiber and becomes more lignified as it matures.

The outstanding characteristic of coir fiber is its resistance to rot. For example, a coir fiber doormat can stay out in the weather for years. Coir is only intermediate in strength among the vegetable fibers, but its elongation at break is greater than any of the bast or hard fibers.

Types. There are three main types of coir—yarn fiber, bristle fiber, and mattress fiber. Nearly all the yarn fiber comes from southern India; much of the mattress fiber and essentially all the bristle fiber come from Sri Lanka.

Only the finest and longest fiber is suitable for spinning into yarn. It is obtained from the husks of unripe nuts and is the main cash crop rather than a by-product. The bristle fiber and most of the mattress fiber come from mature nuts and are by-products of copra production.

Production. The yarn and bristle fiber and much of the mattress fiber extraction starts with a natural bacteriological process called retting. The loosened

fibers are later separated from the nonfibrous material by mechanical action. Immature husks are used for yarn fiber, while mature husks are used to make bristle and mattress fiber.

Yarn fiber. Within a week after harvest the husks are separated from the shell containing the copra and split into four or five pieces which are put into water for retting, to decompose the nonfibrous materials. Retting usually takes 5 to 6 months but sometimes it requires as long as 10 months. The favored retting centers are in the brackish waters of river estuaries, but many husks are retted in tanks, ponds, lakes, and streams. The circulation of water through the husks—either as a result of a tide, the flow of a stream, or changing the water in a tank—improves the quality of the fiber.

After retting, the husks are washed, air-dried, and then beaten with mallets, stones, or sticks to loosen the fiber and remove the nonfiber materials that have decomposed. Some combing is usually done, either by steel combs mounted on wooden wheels or by stationary combs. Then the fiber is spun by hand into single- or two-ply yarn.

Bristle and mattress fiber. For these fibers, the period of retting is much shorter. Before retting, the nuts are frequently allowed to mature further after harvesting so as to even out the quality of the copra produced from them.

The bristle and mattress fiber is usually separated from retted husks by machines, although the separation still requires extensive hand labor. An increasing number of establishments, especially in new areas, are extracting the fiber from green or dry, unretted husks by automatic machines. However, fiber obtained in such a manner is shorter than fiber obtained by traditional processes and is irretrievably mixed so it must be sold as mattress fiber.

Sorting and grading. Coir fiber ranges in color from light tan to dark brown. Bristles are sorted according to length, which ranges from about 6 to 14 in. (15 to 35 cm), and are packed into bundles of about 17.5 oz (500 g) each. Mattress fiber is packed in small bales of about 11 lb (5 kg) each, and yarn is put up in skeins. Before these packages are put into large bales for shipment, the fiber is graded and sometimes resorted. The fiber is graded on the basis of length, stiffness, color, and cleanliness. Long, clean fibers and light color are desirable for any purpose, but stiffness is desirable in bristles and undesirable for yarn.

Uses. In developed countries, yarn fiber is used chiefly in mats and mattings; in the United States, for instance, its best-known use is in light-brown tufted doormats. In Asia it is used extensively for ropes and twines, and locally for hand-made bags. The principal outlet for bristle fiber has been in brush making, but the market has declined and most bristle fiber is now used in upholstery padding. Mattress fiber is used chiefly in innerspring mattresses, though it has found uses as an insulating material. *See* COCONUT; NATURAL FIBER.

Elton G. Nelson

Coke

A coherent, cellular, carbonaceous residue remaining from the dry (destructive) distillation of a coking coal. It contains carbon as its principal constituent, together with mineral matter and residual volatile matter. The residue obtained from the carbonization of a noncoking coal, such as subbituminous coal, lignite, or anthracite, is normally called a char. Coke is produced chiefly in chemical-recovery coke ovens (see **illus.**), but a small amount is also produced in beehive or other types of nonrecovery ovens.

Uses and types. Coke is used predominantly as a fuel reductant in the blast furnace, in which it also serves to support the burden. As the fuel, it supplies the heat as well as the gases required for the reduction of the iron ore. It also finds use in other reduction processes, the foundry cupola, and household heating. About 91% of the coke made is used in the blast furnace, 4% in the foundry, 1% for water gas, 1% for household heating, and 3% for other industries, such as calcium carbide, nonferrous metals, and phosphates. Approximately 1300 lb (560 kg) of coke is consumed per ton of pig iron produced in the modern blast furnace, and about 200 lb (90 kg) of coke is required to melt a ton of pig iron in the cupola.

Coke is classified not only by the oven in which it is made, chemical-recovery or beehive, but also by the temperature at which it is made. High-temperature coke, used mainly for metallurgical purposes, is produced at temperatures of 1650–2100°F (900–1150°C). Medium-temperature coke is produced at 1380–1650°F (750–900°C), and low-temperature coke or char is made at 930–1380°F (500–750°C). The latter cokes are used chiefly for household heating, particularly in England. The production of these is rather small as compared to high-temperature coke and usually requires special equipment. Coke is also classified according to its intended use, such as blast-furnace coke, foundry coke, water-gas coke, and domestic coke.

Production. Coke is formed when coal is heated in the absence of air. During the heating in the range of 660–930°F (350–500°C), the coal softens and then fuses into a solid mass. The coal is partially devolatilized in this temperature range, and further heating at temperatures up to 1830–2000°F (1000–1100°C) reduces the volatile matter to less than 1%. The degree of softening attained during heating determines to a large extent the character of the coke produced.

In order to produce coke having desired properties, two or more coals are blended before charging into the coke oven. Although there are some exceptions, high-volatile coals of 32–38% volatile-matter content are generally blended with low-volatile coals of 15–20% volatile matter in blends containing 20–40% low-volatile coal. In some cases, charges consist of blends of high-, medium-, and low-volatile coal. In this way, the desirable properties of each of the coals, whether it be in impurity content or in its contribution to the character of the coke, are utilized.

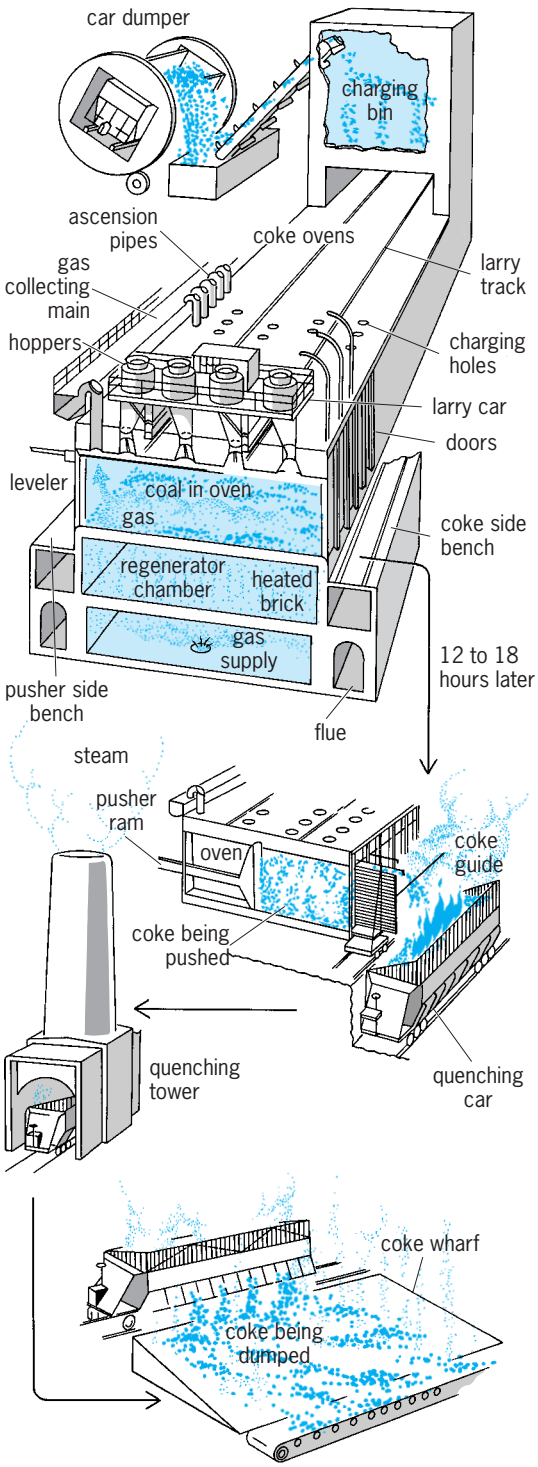

Steps in the production of coke from coal. After the coal has been in the oven for 12–18 h, the doors are removed and a ram mounted on the same machine that operates the leveling bar shoves the coke into a quenching car for cooling. (*American Iron and Steel Institute*)

The low-volatile coal is usually added in order to improve the physical properties of the coke, especially its strength and yield. In localities where low-volatile coals are lacking, the high-volatile coal or blends of high-volatile coals are used without the benefit of the low-volatile coal. Insome cases, particularly in the manufacture of foundry coke, a small percentage of so-called inert, such as fine anthracite, coke fines, or petroleum coke, is added.

In addition to the types of coals blended, the carbonizing conditions in the coke oven influence the characteristics of the coke produced. Oven temperature is the most important of these and has a significant effect on the size and the strength of the coke. In general, for a given coal, the size and shatter strength of the coke increase with decrease in carbonization temperature. This principle is utilized in the manufacture of foundry coke, where large-sized coke of high shatter strength is required.

Properties. The important properties of coke that are of concern in metallurgical operations are its chemical composition, such as moisture, volatile-matter, ash, and sulfur contents, and its physical character, such as size, strength, and density. For the blast furnace and the foundry cupola, coke of low moisture, volatile-matter, ash, and sulfur content is desired. The moisture and the volatile-matter contents are a function of manner of oven operation and quenching, whereas ash and sulfur contents depend upon the composition of the coal charged. Blast-furnace coke used in the United States normally contains less than 1% volatile matter, 85–90% fixed carbon, 7–12% ash, and 0.5–1.5% sulfur. If the coke is intended for use in the production of Bessemer or acid open-hearth iron, the phosphorus content becomes important and should contain less than 0.01%.

The requirements for foundry coke in analysis are somewhat more exacting than for the blast furnace. The coke should have more than 92% fixed carbon, less than 8% ash content, and less than 0.60% sulfur.

Blast-furnace coke should be uniform in size, about $2^1/_2$–5 in. (6.4–13 cm), but in order to utilize more of the coke produced in the plant, the practice has been, in some cases, to charge separately into the furnace the smaller sizes after they were closely screened into sizes such as $2^1/_2 \times 1^3/_4$ in. (6.4 × 4.5 cm), $1^3/_4 \times 1$ in. (4.5 × 2.5 cm), down to about $^3/_4$ in. (1.9 cm) in size. There is now a trend, when pelletized ore is used, to crush and screen the coke to a more closely sized material, such as $2^1/_2 \times {}^3/_4$ in. (6.4 × 1.9 cm).

The blast-furnace coke should also be uniformly strong so that it will support the column of layers of iron ore, coke, and stone above it in the furnace without degradation. Standard test methods, such as the tumbler test and the shatter test, have been developed by the American Society for Testing and Materials for measuring strength of coke. Good blast-furnace cokes have tumbler test stability factors in the range of 45 to 65 and shatter indices (2-in. or 5-cm sieve) in the range of 70 to 80%. These are not absolute requirements since good blast-furnace performance is obtained in some plants with weaker coke—but with the ore prepared so as to overcome the weakness.

Foundrymen prefer coke that is large and strong. The coke for this purpose ranges in size from 3 to 10 in. (7.6 to 25 cm) and larger. The size used usually depends upon the size of the cupola and, in general,

the maximum size of the coke is about one-twelfth of the diameter of the cupola. High shatter strength is demanded by the cupola operator; shatter indices of at least 97% on 2-in. (5-cm) and 80% on 3-in. (7.6-cm) sieves have been specified.

Other important requirements of foundry coke are high reactivity toward oxygen so that carbon dioxide is produced with high heat evolution, and low reactivity toward carbon dioxide so that the amount of carbon monoxide formed is minimized.

The requirements for coke intended for house-heating vary in different localities. Usually the coke is of a narrow size range, and the size used depends upon the size and type of appliance. In general, a high ash-softening temperature is desirable (more than 2500°F or 1370°C).

Coke intended for water-gas generators should be over 2 in. (5 cm) in size, and should have an ash content below 10% and a moderately high ash-softening temperature. *See* CHARCOAL; COAL; COAL CHEMICALS; DESTRUCTIVE DISTILLATION; FOSSIL FUEL; PYROLYSIS. Michael Perch

Bibliography. American Society for Testing and Materials, *Petroleum Products, Lubricants, and Fossil Fuels*, vol. 0.505: *Gaseous Fuels: Coal and Coke*, 1993; Commission of the European Communities, *Coke Oven Techniques*, 1983; R. Loison, *Coke: Quality and Production*, 2d ed., 1989.

Coking (petroleum)

In the petroleum industry, a process for converting nondistillable fractions (residua) of crude oil to lower-boiling-point products and coke. Coking is often used in preference to catalytic cracking because of the presence of metals and nitrogen components that poison catalysts. There are various coking processes: delayed coking, fluid coking, and flexicoking as well as several others. *See* CRACKING.

Delayed coking. The oldest and most widely used coking process, delayed coking has changed very little in the five or more decades that it has been on-stream in refineries. In the semicontinuous process, the residuum or other heavy feedstock is heated to the cracking/coking temperature (>350°C or 660°F, but usually at temperatures on the order of 480°C or 895°F), and the hot liquid is charged, usually by upflow, to the coke drum, where the coking reactions occur. Liquid and gaseous products pass to the fractionator for separation, and the coke deposits in the drum. The coke drums are arranged in pairs, one on-stream and the other off-stream; and they are used alternately to allow continuous processing. Thus they can be operated on a cycle, typically 24–48 h.

The overhead oil is fractionated into fuel gas (ethane and lower-molecular-weight gases), propane-propylene, butane-butene, naphtha, light gas oil, and heavy gas oil. Yields and product quality vary widely because of the broad range of feedstock types charged to the delayed coking process.

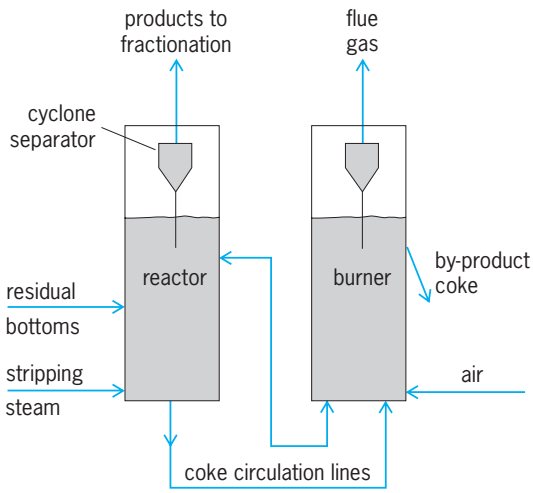

Fluid coking process.

The function of the coke drum is to provide the residence time required for the coking reactions and to accumulate the coke. Hydraulic cutters are used to remove coke from the drum.

Fluid coking. Fluid coking is a continuous fluidized solids process that cracks the feedstock thermally over heated coke particles in a reactor vessel to gas, liquid products, and coke. Heat for the process (see **illus.**) is supplied by partial combustion of the coke, with the remaining coke being drawn as product. The new coke is deposited in a thin fresh layer on the outside surfaces of the circulating coke particles.

Small particles of coke made in the process circulate in a fluidized state between the vessels and serve as the heat transfer medium. Thus the process requires no high-temperature preheat furnace. Fluid coking is carried out at essentially atmospheric pressure and temperatures in excess of 485°C (900°F) with residence times of the order of 15–30 s. The shorter residence time in the fluid coker is in contrast to the residence time of the delayed coker, where the coking reactions are allowed to proceed to completion. This difference is evident from the somewhat higher liquid yields observed in many fluid coking processes. However, the products from a fluid coker may be somewhat more olefinic and slightly less desirable for downstream processing.

Flexicoking. A modification of the fluid coking process, flexicoking also includes a gasifier adjoining the burner/regenerator to convert excess coke to a clean fuel gas with a heating value of about 90 Btu/ft^3 (335 joules/m^3). The coke gasification can be controlled to burn about 95% of the coke to maximize production of coke gas, or at a reduced level to produce both gas and a coke. This flexibility permits adjustment for coke market conditions over a considerable range of feedstock properties and coking processes.

Suitable feedstocks include vacuum residua of all types, asphalt, tar-sand bitumen, and visbreaker bottoms (volatile residues remaining from the viscosity-breaking process).

Uses. The liquid products from the coker, after cleanup via commercially available hydrodesulfurization technology, can provide large quantities of low-sulfur liquid fuels (less than 0.2 wt % sulfur). Another major application for the processes is upgrading heavy low-value crude oils into lighter products.

Petroleum coke is used principally as a fuel or, after calcining, for carbon electrodes. The feedstock from which the coke is produced controls the coke properties, especially in terms of sulfur, nitrogen, and metals content. A concentration effect tends to deposit the majority of the sulfur, nitrogen, and metals in the coke. Cokes exceeding around 2.5% sulfur content and 200 parts per million vanadium are mainly used, environmental regulations permitting, for fuel or fuel additives. The properties of coke for nonfuel use include low sulfur, metals, and ash as well as a definable physical structure. *See* PETROLEUM PROCESSING AND REFINING; PETROLEUM PRODUCTS. James G. Speight

Bibliography. J. G. Gary and G. E. Handwerk, *Petroleum Refining: Technology and Economics*, 3d ed., 1984; J. G. Speight, *The Chemistry and Technology of Petroleum*, 2d ed., 1991; J. G. Speight, *Fuel Science and Technology Handbook*, 1990.

Cola

A tree, *Cola acuminata*, of the family Sterculiaceae and a native of tropical Africa. Its fruit is a star-shaped follicle containing eight hard seeds, the cola nuts of commerce (see **illus.**). These nuts are an important masticatory in many parts of tropical Africa. They have a caffeine content twice that of coffee. The nuts also contain an essential oil and a glucoside, kolanin, which is a heart stimulant. Cola nuts, in combination with an extract from coca, are used in the manufacture of the beverage Coca-Cola. Cola is now cultivated in West Africa, Jamaica, Brazil, India, and other parts of tropical Asia. *See* COCA; MALVALES.

Cola (*Cola acuminata*) branch.

Perry D. Strausbaugh; Earl L. Core

Colchicine

The major alkaloid obtained from the seed capsules, corms, and bulbs of the meadow saffron (autumn crocus; *Colchicum autumnale*). Colchicine is a fairly abundant alkaloid in dried corms (0.03–0.08%) and seeds (0.6–0.9%). It has been isolated also from other species of *Colchicum* and a few related genera, but they are not important sources of alkaloid. *See* ALKALOID; LILIALES.

Colchicine has an unusual structure (shown below) containing two seven-membered rings. Ring

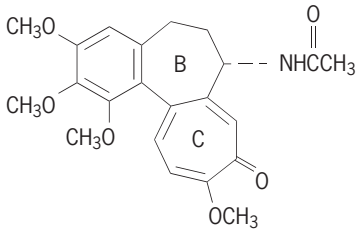

C is a tropolone system, present also in some fungal metabolites. The molecule has been synthesized in several ways, by constructing the ring system in either the sequence A → AB → ABC or A → AC → ABC.

The use of the colchicum plant to relieve the pain of gout has been known since medieval times, and colchicine is still the standard treatment for gout, although it is an extremely toxic substance.

An important property of colchicine is its ability to interrupt the mitotic cycle before cell division occurs. This effect leads to cells with multiple chromosomes, which are of value in plant breeding. The antimitotic effect of colchicine is due to strong binding of the alkaloid to the protein tubulin, thus preventing the assembly of microtubules and formation of the mitotic spindle. *See* MITOSIS; POLYPLOIDY.

James A. Moore

Cold hardiness (plant)

The ability of temperate zone plants to survive subzero temperatures. This characteristic is a predominant factor that determines the geographical distribution of native plant species and the northern limits of cultivation of many important agronomic and horticultural crops. For example, plant hardiness zone maps (**Fig. 1**) are frequently used in recommending the selection of various horticultural crops, especially woody ornamentals (see **table**) and fruit trees, for various regions. Further, freezing injury is a major cause of crop loss resulting from early fall frosts, low midwinter temperatures, or late spring frosts. Problems of cold hardiness are of concern to farmers in diverse areas of agriculture—from coffee growers in Brazil to wheat farmers in Canada; from peach growers in Georgia to apple growers in New York; from tropical foliage plant growers in Florida to producers of woody ornamentals in the northeast; from viticulturists (grape growers) in France to citrus growers

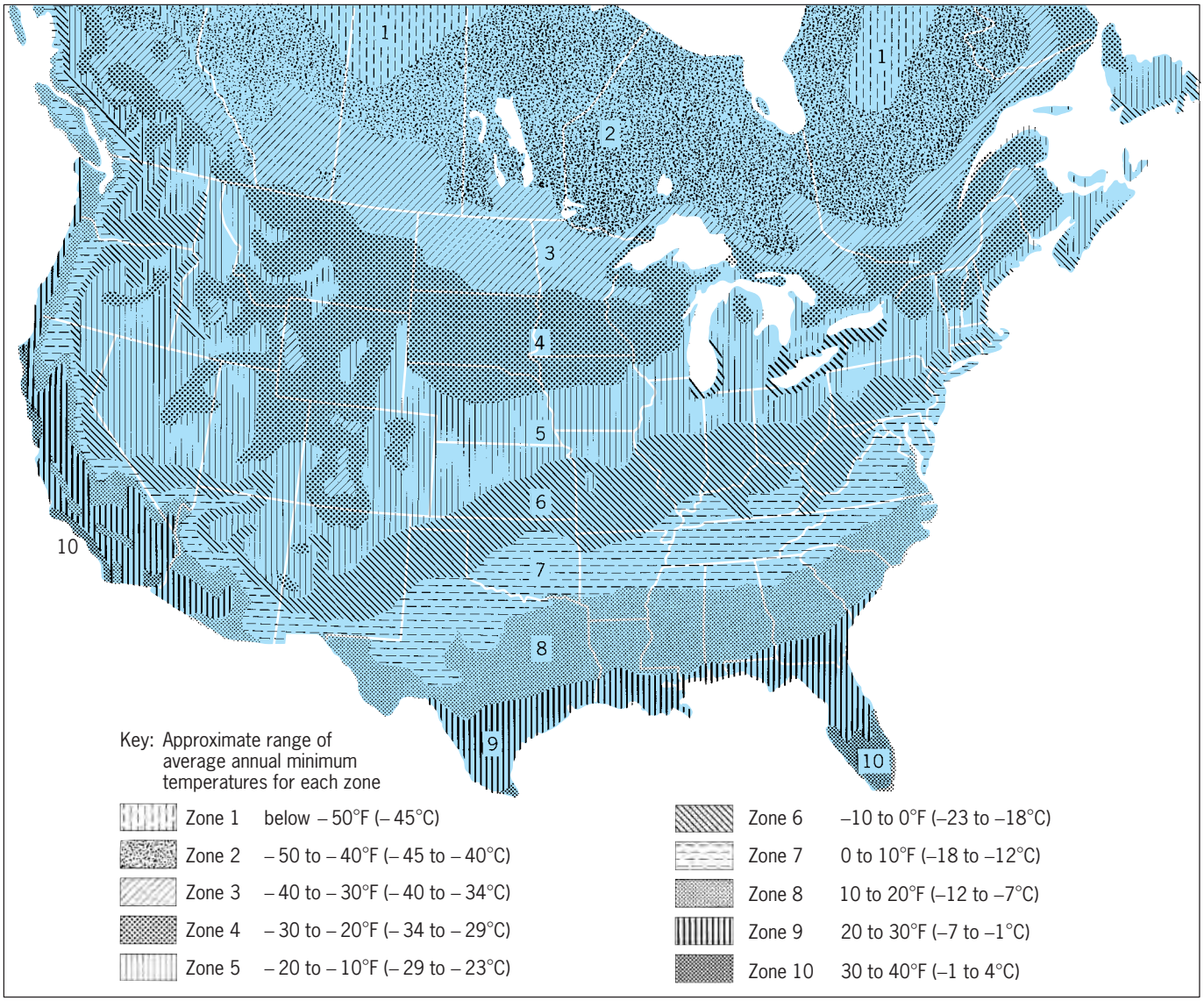

Fig. 1. Plant hardiness zone map of the United States based on isotherms of average minimum winter temperatures.

in Israel. As a result, the development of varieties of cultivated plants with improved cold hardiness is of long-standing concern.

Within the plant kingdom there is a wide range of diversity in low-temperature tolerance—from low levels of hardiness in herbaceous species such as potatoes (27 to 21°F or −3 to −6°C), to intermediate levels of hardiness for winter annuals such as wheat and rye (−4 to −22°F or −20 to −30°C), to extremely hardy deciduous trees and shrubs such as black locust and red osier dogwood that can withstand temperatures of liquid nitrogen (−321°F or −196°C). Within a given species, the range in hardiness can be substantial. For example, the cold hardiness of different varieties of deciduous fruit trees may vary by 18 to 36°F (10 to 20°C). Within a given plant there is a wide range in the cold hardiness of different tissues and organs. For example, roots are much less tolerant of subzero temperatures than shoots; flower buds are more sensitive than vegetative buds.

Environmental effects. The cold hardiness of a given species is an inherent genetic trait that requires certain environmental cues for its expression. With the shorter days and cooler nights of autumn, temperate zone plants become dormant and increase their cold hardiness. This process is referred to as cold acclimation. In the spring, increasing daylength and warmer temperatures result in the resumed growth and development of the plant and a corresponding decrease in cold hardiness. Thus, hardy species such as winter annuals, biennials, and perennials exhibit an annual periodicity in their low-temperature tolerance (**Fig. 2**). In the spring and summer these species are as susceptible to freezing temperatures as are nonhardy species.

Cold hardiness may be influenced by radiation, temperature, photoperiod, precipitation, and stage of development of the plant, with different optimum conditions for different species or cultivars and ecotypes within a species. The various environmental

Representative ornamental plants that will survive in the various hardiness zones

	Scientific name	Common name
Zone 1 below −50°F (−45°C)	*Betula glandulosa*	Dwarf birch
	Empetrum nigrum	Crowberry
	Populus tremuloides	Quaking aspen
	Potentilla pennsylvanica	Pennsylvania cinquefoil
	Rhododendron lapponicum	Lapland rhododendron
	Salix reticulata	Netleaf willow
Zone 2 −50° to −40°F (−45 to −40°C)	*Betula papyrifera*	Paper birch
	Cornus canadensis	Bunchberry dogwood
	Elaeagnus commutata	Silverberry
	Juniperus communis	Common juniper
	Picea glauca	White spruce
	Potentilla fruticosa	Bush cinquefoil
Zone 3 −40 to −30°F (−40 to −34°C)	*Elaeagnus angustifolia*	Russian olive
	Lonicera tatarica	Tatarian honeysuckle
	Malus baccata	Siberian crabapple
	Parthenocissus quinquefolia	Virginia creeper
	Syringa vulgaris	Common lilac
	Thuja occidentails	American arborvitae
Zone 4 −30 to −20°C (−34 to −29°C)	*Berberis thunbergi*	Japanese barberry
	Hydrangea paniculata	Panicle hydrangea
	Juniperus chinensis	Chinese juniper
	Ligustrum amurense	Amur River privet
	Spiraea vanhouttei	Vanhoutte spirea
	Taxus cuspidata	Japanese yew
Zone 5 −20 to −10°F (−29 to −23°C)	*Cornus florida*	Flowering dogwood
	Deutzia gracilis	Slender deutzia
	Forsythia ovata	Early forsythia
	Parthenocissus tricuspidata	Boston ivy
	Rosa multiflora	Japanese rose
Zone 6 −10 to 0°F (−23 to −18°C)	*Acer palmatum*	Japanese maple
	Buxus sempervirens	Common boxwood
	Forsythia suspensa	Weeping forsythia
	Hedera helix	English ivy
	Ilex opaca	American holly
	Ligustrum ovalifolium	California privet
Zone 7 0 to 10°F (−18 to −12°C)	*Azalea Kurume hyb. Hinodegiri*	Red Hussar azalea
	Cedrus atlantica	Atlas cedar
	Cercis chinensis	Chinese redbud
	Chamaecyparis lawsoniana	Lawson cypress
	Cotoneaster salicifolia	Willowleaf cotoneaster
	Ilex aquifolium	English holly
Zone 8 10 to 20°F (−12 to −7°C)	*Arbutus menziesi*	Pacific madrone
	Choisya ternata	Mexican orange
	Melia azedarach	Chinaberry
	Olearia haasti	New Zealand daisybush
	Prunus laurocerasus	Cherry laurel
	Viburnum tinus	Laurestinus
Zone 9 20 to 30°F (−7 to −1°C)	*Arbutus unedo*	Strawberry tree
	Eucalyptus globulus	Tasmanian blue gum
	Grevillea robusta	Silk oak
	Myrtus communis	Myrtle
	Pittosporum tobira	Japanese pittosporum
	Quercus virginiana	Live oak
Zone 10 30° to 40°F (−1 to 4°C)	*Acacia baileyana*	Cootamundra wattle
	Arecastrum romanozoffianum	Queen palm
	Bougainvillea spectabilis	Bougainvillea
	Casuarina equisetifolia	Horsetail beefwood (Australian pine)
	Eucalyptus citriodora	Lemon eucalyptus
	Ficus macrophylla	Moreton Bay fig

cues serve to synchronize plant development with the environment. This synchronization has taken centuries to evolve, and freezing injury in cultivated species can result from any factor that disrupts this synchrony. Some varieties may not be responsive to the prevailing environmental cues or may not respond rapidly enough; some may not develop a sufficient degree of hardiness; and some may deacclimate too rapidly. These factors may arise when individual varieties are introduced into areas that are vastly or even slightly different from their natural habitat, where centuries of selection pressures have evolved those individuals most closely synchronized with the prevailing environment.

Temperature is the key environmental parameter for increasing a plant's capacity to withstand freezing temperatures. Low, above-freezing temperatures are conducive to an increase in hardiness in the fall, and warm temperatures are responsible for the decrease in the spring. Generally, it is considered that most plants will acclimate as temperatures are gradually lowered below 50°F (10°C). However, during acclimation, the progressive decline in temperatures (from the relatively warm temperatures in early fall, followed by low, above-freezing temperatures in late fall and early winter, followed by freezing temperatures in winter) is extremely important. The development of cold hardiness may take 4 to 6 weeks.

Photoperiod is the second major factor influencing cold acclimation, but only in those species that are photoperiodically responsive in relation to growth cessation or induction of dormancy (a true physiological rest period). In other species, light is important only in providing sufficient photosynthetic reserves required for the cold acclimation process. In some cases (for example, germinating seeds), sufficient energy reserves are already present and acclimation can occur in the dark. *See* PHOTOPERIODISM.

There are conflicting reports on the role of moisture in relation to cold hardiness. High soil moisture may reduce the degree of cold acclimation; however, severe winter injury of evergreens will occur if soil moisture levels are too low. Most often tissue moisture levels will influence the survival to a given freeze-thaw cycle rather than directly influencing the process of cold acclimation. Thus, whereas temperature and light effects on hardiness are probably mediated through the development of hardiness (cold acclimation), tissue moisture content directly affects the stresses that are incurred during a freeze-thaw cycle. In addition, various cultural practices can influence the cold hardiness of a given plant. For example, late fall applications of fertilizer or improper pruning practices may stimulate flushes of growth that do not have sufficient time to acclimate. Conversely, insufficient mineral nutrition can also impair the development of maximum cold hardiness.

Freezing of plant tissues. Because plants are poikilotherms and assume the temperature of their surroundings, cold hardiness requires that they tolerate ice formation in tissues. When plant tissues freeze, the water they contain changes both its physical state and location. Ice formation may occur within cells (intracellular ice formation), in the intercellular spaces (extracellular ice formation), or in regions that separate different tissues or organs (extraorgan ice formation). In some organs and cells, such as floral buds or xylem ray parenchyma cells of deciduous forest species, ice formation may be precluded entirely. In these cases, the cytoplasm of the cell remains unfrozen below its freezing point (that is, supercooled) down to temperatures as low as −40°F (−40°C). The location of ice formation (intracellular versus extracellular) is strongly influenced by the rate of cooling, the minimum temperature attained, and the degree of hardiness of the plant.

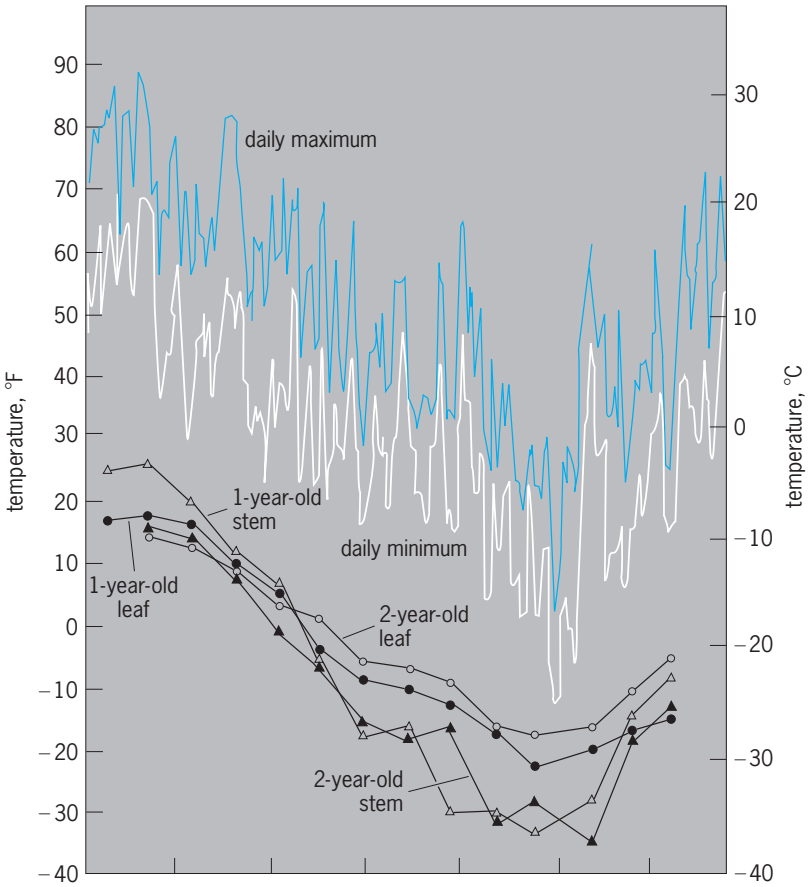

Fig. 2. Development of cold hardiness of *Hedera helix* (English ivy). The curves show killing points for structures at different ages. (*From P. L. Steponkus and F. O. Lanphear, Factors influencing artificial cold acclimation and artificial freezing of Hedera helix "Thorndale," Proc. Amer. Soc. Hort. Sci., 91:735–741, 1967*)

During cooling, plant tissues initially supercool; that is, ice formation does not occur at the freezing point of the cytoplasm. Instead, ice nucleation occurs on the external plant surfaces because of the presence of foreign nucleating agents, such as ice crystals in the air or, in some cases, bacteria. Ice then propagates through the extracellular spaces, but the plasma membrane (the outer membrane of the cell) prevents seeding of the cytoplasm by the external ice. Because ice has a lower chemical potential than water at the same temperature, a gradient in water potential exists between the unfrozen cytoplasm and the extracellular ice. Water moves in response to

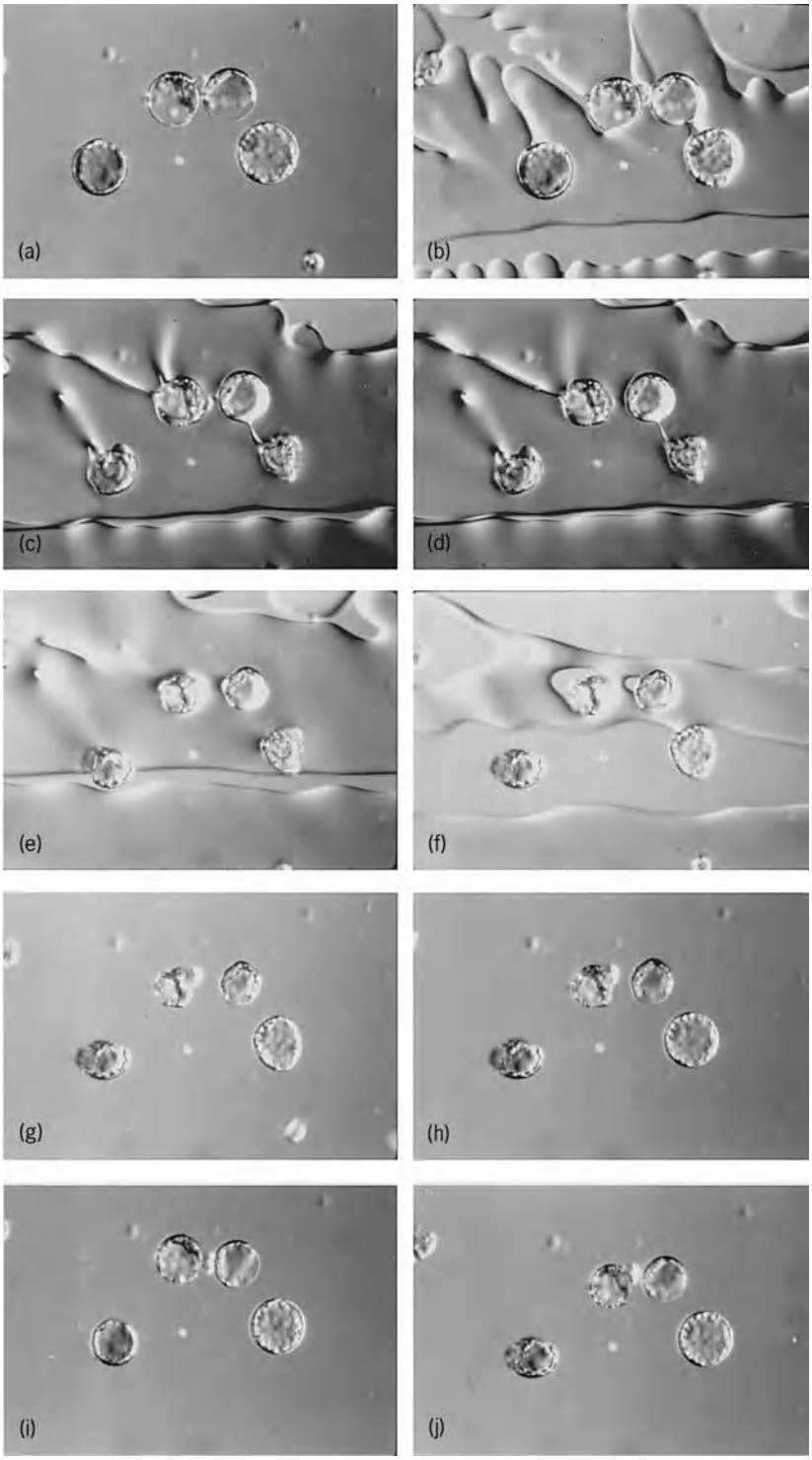

Fig. 3. Cryomicroscopic observations of isolated rye (*Secale cereale*) protoplasts during a freeze-thaw cycle. Cooling: (*a*) 32°F (0°C), (*b*) 28°F (−2°C), (*c*) 23°F (−5°C), (*d*) 14°F (−10°C). Warming: (*e*) 17°F (−8°C), (*f*) 25°F (−4°C), (*g*) 32°F (0°C). (*h–j*) Isothermal at 32°F (0°C). Note the extensive dehydration of the cell that occurs at subzero temperatures. (*Courtesy of Peter L. Steponkus*)

gradients in water potential; hence water will move from the cytoplasm to the extracellular ice and the cell will dehydrate. Because the plasma membrane is a semipermeable membrane that allows for the rapid efflux of water but does not permit the ready efflux of solutes, the cytoplasmbecomes more concentrated. With increasing solute concentration, the freezing point of the cytoplasm is progressively lowered and the intracellular solution remains unfrozen. During warming, the sequence of events is reversed. The extracellular ice melts and the extracellular solution becomes less concentrated than the intracellular solution. As a result, water will flow back into the cell. Thus, during a freeze-thaw cycle, the cells behave as osmometers in response to ice formation in the extracellular spaces. *See* CELL MEMBRANES; CELL WALLS (PLANT).

The osmometric behavior of plant cells depends on the semipermeable characteristics of the plasma membrane that allow for the rapid movement of water while retaining solutes within the cell and excluding extracellular ice crystals. This behavior is readily observed microscopically. Although microscopic observations of biological specimens at subzero temperatures have been reported for over a century, there has been a renaissance in cryomicroscopy since the early 1970s. This has occurred because of the development of sophisticated temperature-control systems and advances in video recording and computer-enhanced video image analysis. As a result, cryomicroscopy has emerged as a very powerful technique for quantifying and understanding cellular behavior during a freeze-thaw cycle. Although it has been nearly 100 years since the cooling-rate dependency of intracellular ice formation was observed in higher plants, quantitative studies of the volumetric behavior of higher plant protoplasts are only now being reported. *See* CRYOBIOLOGY.

Isolated protoplasts are obtained from plant tissues by the enzymatic digestion of the cell walls. The liberated protoplasts behave as ideal osmometers. Their spherical shape over a wide range of solute concentrations facilitates quantitative cryomicroscopic studies of their osmometric behavior. Direct measurements of volumetric behavior during a freeze-thaw cycle (**Fig. 3**) allow for the calculation of water efflux, the extent of supercooling, and water permeability of the plasma membrane. Coupled with a precise time and temperature history, such studies provide considerable insight into the causes and manifestations of freezing injury. Extensive dehydration of the cell occurs at subzero temperatures; following thawing of the suspending medium, the protoplasts rehydrate and osmotically increase in volume. Because the protoplasts in Fig. 3 were isolated from acclimated tissues, they survived the freeze-thaw cycle to a temperature of 14°F (−10°C).

Intracellular ice formation. Whether ice formation is restricted to the extracellular spaces or also occurs within the cells is of utmost importance because intracellular ice formation is invariably associated with lethal injury. In isolated rye protoplasts subjected to rapid cooling rates (9°F or 5°C per minute; **Fig. 4**), intracellular ice formation is manifested as a darkening of the cells (Fig. 4*c* and *d*). This optical darkening or flashing is commonly observed, and was long thought to be the result of the diffraction of light by small ice crystals. High-resolution cryomicroscopy of isolated protoplasts reveals that the

optical darkening is due to the formation of gas bubbles following intracellular ice formation. Upon melting of the ice, some of the gas bubbles coalesce (Fig. 4*e*). More importantly, however, intracellular ice formation is associated with lethal injury as manifested by the ruptured protoplasts (Fig. 4*f*). Intracellular ice formation is associated with early frost injury of herbaceous plants, and may also occur when there are sudden temperature changes in deciduous fruit tree buds.

The incidence of intracellular ice formation is strongly dependent on the rate at which the tissue is cooled and the minimum temperature attained. At rapid cooling rates, efflux of water from the cell is limited by the water permeability of the plasma membrane or other impediments to water flux. As a result, concentration of the intracellular solutes and depression of the freezing point will not keep pace with the decreasing temperature, and the intracellular solution will become increasingly supercooled during cooling. With suspensions of isolated plant cells or protoplasts where the cells are surrounded by ice and the only barrier to water efflux is the plasma membrane, excessive supercooling will occur at cooling rates greater than 5°F (3°C) per minute. In intact tissues, rates greater than a few degrees per hour are sufficiently rapid to result in supercooling of the intracellular solution. This is because the ice crystals may be far removed from the intracellular solution and water flow is slowed by impediments resulting from tissue organization, contact of cell walls, or other barriers. Supercooling per se does not result in intracellular ice formation; it only predisposes the cell to this possibility. Nucleation or seeding of the supercooled solution is required for intracellular ice formation. This may occur by either homogeneous or heterogeneous nucleation or by seeding by the extracellular ice.

Homogeneous nucleation is a result of the spontaneous aggregation of water molecules to form ice nuclei and occurs only at temperatures below −38°F (−39°C). Thus, floral buds of some species or xylem ray parenchyma of some deciduous species will tolerate temperatures down to −38°F (−39°C) because they remain supercooled. At lower temperatures, however, intracellular ice formation occurs because of homogeneous nucleation. The sudden and rapid formation of ice is invariably lethal. There is a close relationship between the northern limits of many deciduous forest species that exhibit supercooling of the xylem ray parenchyma cells and the occurrence of temperatures below −40°F (−40°C).

Heterogeneous nucleation may occur at higher temperatures, but it requires the presence of foreign materials that are effective ice-nucleating agents. However, there is little evidence that such compounds are present in the cytoplasm. Instead, there is considerable evidence that intracellular ice formation is a consequence of seeding by extracellular ice as a result of perturbation or penetration of the plasma membrane. Direct cryomicroscopic observations of isolated protoplasts reveal that mechanical failure of the plasma membrane occurs immediately

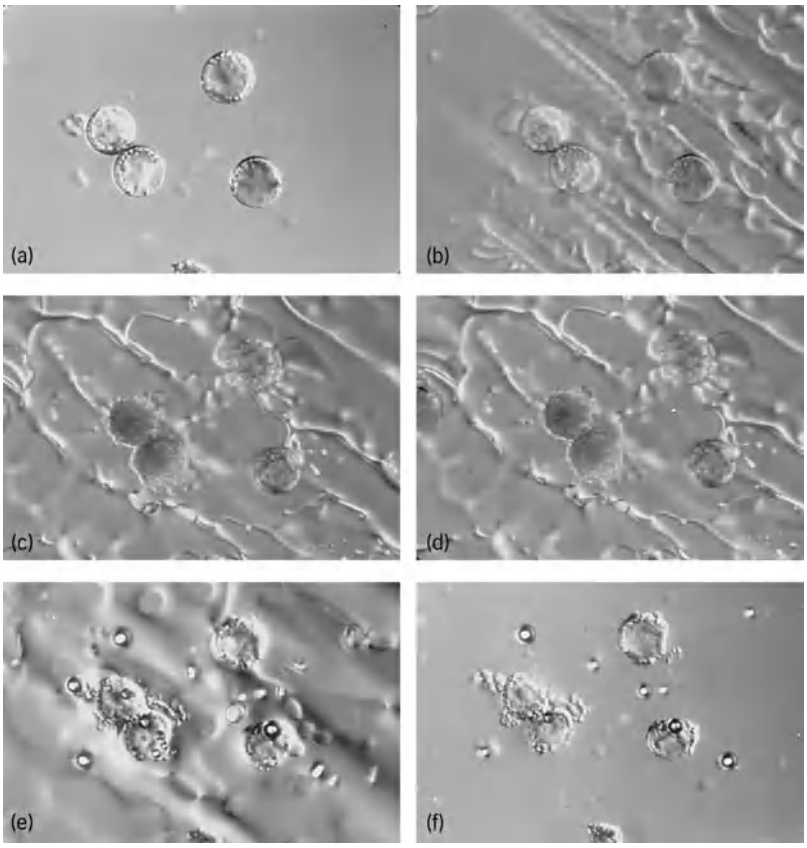

Fig. 4. Intracellular ice formation in isolated rye (*Secale cereale*) protoplasts. Cooling: (*a*) 32°F (0°C), (*b*) 28°F (−2°C), (*c*) 14°F (−10°C), (*d*)14°F (−10°C). Warming: (*e*) 28°F (−2°C), (*f*) 32°F (0°C). (*Courtesy of Peter L. Steponkus*)

before intracellular ice formation. This occurs in the range of 23 to 5°F (−5 to −15°C) in nonacclimated plants. Thus, intracellular ice formation is a consequence of injury to the plasma membrane rather than a cause of injury.

Following cold acclimation, plants are less prone to intracellular ice formation. For many years it was assumed that this was because of an increase in the water permeability of the plasma membrane, which minimized the extent of supercooling. Recently, however, this was shown to be erroneous. Instead, cold acclimation decreases the incidence of intracellular ice formation because the stability of the plasma membrane at lower temperatures is increased and defers seeding by extracellular ice to temperatures as low as −40°F (−40°C).

Freezing injury due to cell dehydration. The visual symptoms of freezing injury—a darkened, water-soaked, flaccid appearance—are apparent immediately following thawing. The rapidity in which injury is manifested indicates that injury is not the result of metabolic dysfunction. Instead, the primary cause of freezing injury is the disruption of the plasma membrane. Manifestations of injury may range from alterations in the semipermeable characteristics to mechanical rupture, and injury can occur at various times during a freeze-thaw cycle (**Fig. 5**). The various forms of injury observed are (1) expansion-induced lysis during warming andthawing of the suspending

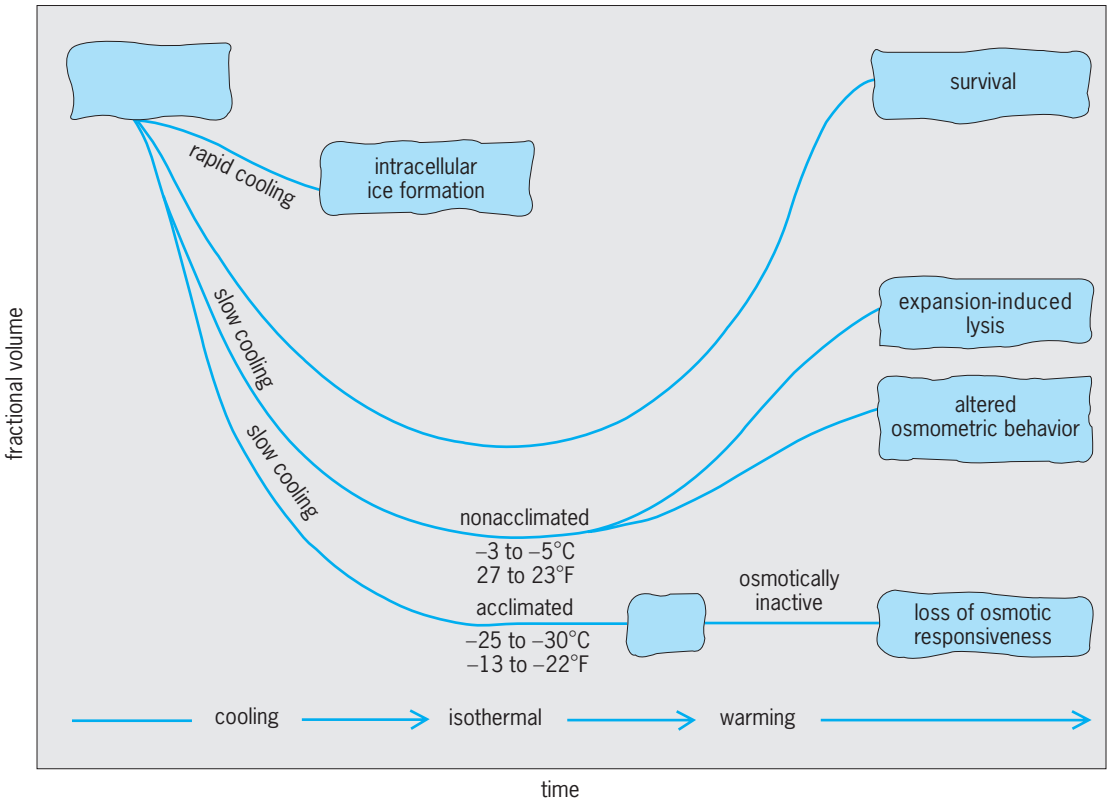

Fig. 5. Manifestations of freezing injury to the plasma membrane.

medium when the decreasing concentration of the suspending medium results in osmotic expansion of the protoplasts; (2) loss of osmotic responsiveness following cooling at slow rates so that the protoplast is osmotically inactive during warming; and (3) altered osmometric behavior during warming, suggestive of altered semipermeable characteristics or leakiness of the plasma membrane.

These symptoms of injury are not consequences of low temperatures per se because low temperatures in the absence of freezing are not injurious to hardy plants. Thus, thermal stresses that may result in membrane phase transitions are unlikely to contribute to freezing injury. Similarly, there is little evidence that injury occurs because of the presence of ice crystals. Instead, injury is primarily a consequence of cell dehydration, and there is a very good correlation between cold hardiness and the ability to withstand dehydration in the absence of freezing. There are, however, several consequences of cell dehydration that can result in injury to the plasma membrane.

In nonacclimated tissues, cell dehydration results in the deletion of membrane material from the plasma membrane in the form of membrane vesicles (**Fig. 6***a*). During contraction of the protoplast, these vesicles are detached from the membrane and become visible in the cytoplasm. Upon thawing, the tissue increases in volume, but the previously deleted membrane material is not reincorporated into the membrane, and rupture of the plasma membrane (lysis) occurs before the cells regain their original size. Hence, this form of injury is referred to as expansion-induced lysis or deplasmolysis injury, and is the predominant form of injury following cell dehydration in nonacclimated herbaceous tissues such as winter cereals (wheat or rye). Injury is largely a consequence of the mechanical stresses imposed on the plasma membrane during cell dehydration and rehydration.

Expansion-induced lysis is not observed in acclimated tissues because membrane deletion does not occur during contraction. Instead, extrusions of the plasma membrane that remain contiguous with the plasma membrane are formed (Fig. 6*b*). These extrusions are readily reincorporated into the plane of the membrane during expansion. Hence, mechanical stresses of the plasma membrane are not responsible for injury in acclimated tissues. Instead, injury is largely the result of chemical stresses associated with cell dehydration.

Loss of osmotic responsiveness occurs in protoplasts isolated from nonacclimated tissues when the protoplasts are cooled to 14°F (−10°C) or lower. At these temperatures severe freeze-induced dehydration occurs with a loss of nearly 90% of the intracellular water, and with a tenfold increase in the concentration of the intracellular solutes. The semipermeable properties of the plasma membrane are disrupted so that following thawing the protoplasts are osmotically unresponsive. In tissues, the cells appear to be plasmolyzed, and therefore this form of injury has been referred to as frost plasmolysis. However, this term is misleading in that the cells cannot be considered plasmolyzed if they are not osmotically responsive.

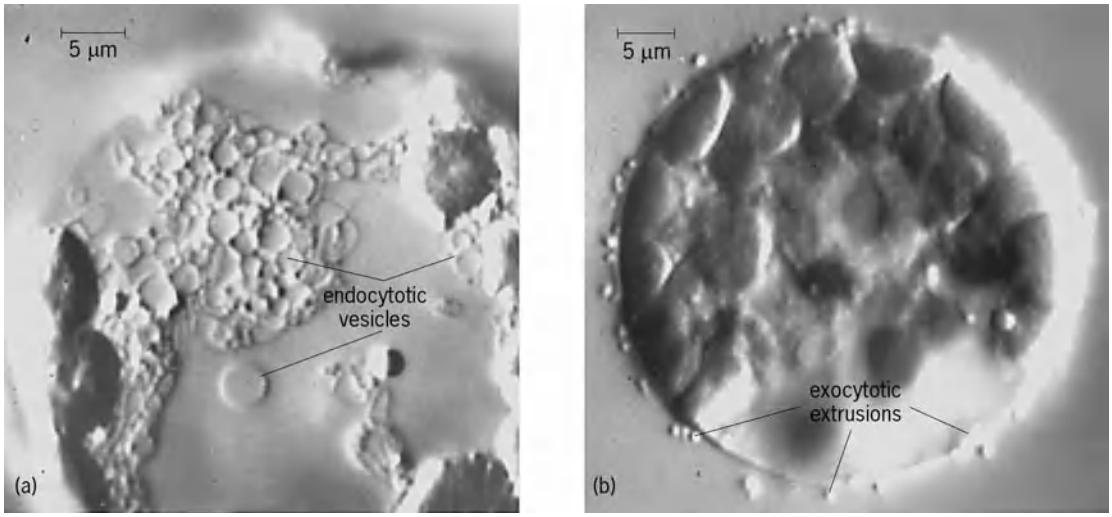

Fig. 6. High-resolution computer-enhanced differential interference contrast micrographs of (*a*) nonacclimated and (*b*) acclimated protoplasts subjected to hypertonic solutions. (*From P. L. Steponkus et al., Destabilization of the plasma membrane of isolated plant protoplasts during a freeze-thaw cycle: The influence of cold acclimation, Cryobiology, 20:448–465, 1983*)

Loss of osmotic responsiveness is associated with several changes in the ultrastructure of the plasma membrane, including lateral phase separations and bilayer-to-nonbilayer transitions, such as the formation of inverted cylindrical micelles or hexagonal$_{II}$ phase. These changes occur as a consequence of the close approach (less than 2 to 3 nanometers) of membranes during cell dehydration. Normally, strongly repulsive hydration forces preclude the close approach of two bilayers. However, the high osmotic pressures that occur during freeze-induced dehydration overcome the hydration forces. Dehydration-induced lateral phase separations and the hexagonal$_{II}$ phase are not observed in protoplasts isolated from cold-acclimated tissues. The increased stability may be a consequence of alterations in the plasma membrane lipid composition that preclude dehydration-induced demixing of the lipids, and the accumulation of cytoplasmic solutes such as sucrose or proline.

Cold acclimation. The process of cold acclimation results in numerous biochemical changes within the plant. These include increases in growth inhibitors and decreases in growth promoters; changes in nucleic acid metabolism; alterations in cellular pigments such as carotenoids and anthocyanins; the accumulation of carbohydrates, amino acids, and water-soluble proteins; increases in fatty acid unsaturation; changes in lipid composition; and the proliferation of cellular membrane systems. Not all of these changes are involved in increased cold hardiness; some are merely changes in response to slower growth rates and decreased photosynthate utilization; others are changes associated with growth at low, above-zero temperatures; and still others are associated with other developmental phenomena, such as vernalization or the induction of dormancy, that also occur during the period of cold acclimation.

As cellular membranes are the primary site of freezing injury, it follows that cold acclimation must involve cellular alterations that allow the membranes to survive lower freezing temperatures. Such alterations may be in the cellular environment so that either the freezing stresses are altered or there is direct protection of the membranes. In addition, cold acclimation may involve changes in the membrane itself, so that its susceptibility to the freezing stresses is decreased.

Large increases in cellular solute concentrations are one of the most universal manifestations of cold acclimation. A doubling of the intracellular solute concentration, most notably sugars, is not uncommon. Such increases have several beneficial effects. First, they serve to depress the freezing point of the intracellular solution. This, however, is of limited advantage because such increases only lower the freezing point to the extent of 1.03°F/osmolal (1.86°C/osmolal).

More important is their effect on the extent of cell dehydration that occurs during freezing. A doubling of the initial intracellular solute concentration will decrease the extent of cell dehydration at any subzero temperature by 50%. An increase in intracellular solutes will also decrease the concentration of toxic solutes at temperatures below 32°F (0°C), because less water will be removed. During dehydration protective compounds, such as sugars, are concentrated along with toxic compounds, such as electrolytes. Because the total concentration of the intracellular solution is a function of the subzero temperature, the toxic compounds will account for only a portion of the total. If the ratio of protective to toxic compounds is initially greater, the final concentration of the toxic compounds at any subzero temperature will be lowered accordingly. In addition, there are some compounds, such as sucrose or proline, that may interact directly with cellular membranes to increase their stability either at low water contents or at high electrolyte concentrations.

Following cold acclimation there are substantial changes in the lipid composition of the plasma membrane. Although there are no lipid molecular species that are unique to the plasma membrane of either nonacclimated or acclimated rye leaves, cold acclimation alters the proportion of virtually every lipid component. This includes an increase in free sterols with corresponding decreases in steryl glucosides and acylated steryl glucosides, a decrease in the glucocerebroside content, and an increase in the phospholipid content. Although the relative proportions of the individual phospholipids (for example, phosphatidylcholine and phosphatidylethanolamine) do not change appreciably, there are substantial differences in the individual lipid molecular species. These include increases in species of phosphatidylcholine and phosphatidylethanolamine that contain two unsaturated fatty acids, whereas species with a single unsaturated fatty acid remain relatively unchanged. The complexity of the plasma membrane lipid composition and the numerous changes that occur during cold acclimation preclude the possibility that any simple correlative analysis of the changes will establish their role in the cold acclimation process; instead, a mechanistic analysis is required.

One approach has been to demonstrate that the differential behavior of the plasma membrane observed in protoplasts isolated from nonacclimated and cold-acclimated leaves (that is, the formation of endocytotic vesicles versus exocytotic extrusions during osmotic contraction and the differential propensity for dehydration-induced bilayer-to-nonbilayer phase transitions) is also observed in liposomes formed from plasma membrane lipid extracts of nonacclimated and acclimated leaves. Thus, the increased cryostability of the plasma membrane following cold acclimation is a consequence of alterations in the lipid composition. However, there are over 100 different lipid molecular species present in the plasma membrane, so that the problem is to determine which of the changes are responsible for the differential cryobehavior. In studies that have used a protoplast–liposome fusion technique to enrich the plasma membrane with specific lipids, it has been established that increases in the proportion of phosphatidylcholine species composed of one or two unsaturated fatty acids increase the freezing tolerance of protoplasts isolated from nonacclimated rye leaves. The increase in freezing tolerance is the result of a transformation in the cryobehavior of the plasma membrane during osmotic contraction so that exocytotic extrusions are formed rather than endocytotic vesicles. As a result, expansion-induced lysis is precluded. Thus, these studies demonstrate that although there are over 100 lipid molecular species in the plasma membrane, altering the proportion of a single species dramatically alters the cryobehavior of the plasma membrane. Similar membrane engineering studies have been undertaken for the other forms of freezing injury (for example, loss of osmotic responsiveness) and should ultimately provide a molecular understanding of the role of plasma membrane lipid alterations in increasing the cryostability of the plasma membrane. *See* ALTITUDINAL VEGETATION ZONES; PLANT PHYSIOLOGY; PLANT-WATER RELATIONS.

Peter L. Steponkus

Bibliography. B. W. W. Grout and G. J. Morris (eds.), *The Effect of Low Temperatures on Biological Membranes*, 1987; O. L. Lange et al. (eds.), *Encyclopedia of Plant Physiology*, 1981; J. Levitt, *Responses of Plants to Environmental Stresses*, 2d ed., 1980; P. H. Li and A. Sakai (eds.), *Plant Cold Hardiness and Freezing Stress*, vol. 1, 1978, vol. 2, 1982; D. V. Lynch and P. L. Steponkus, Plasma membrane lipid alterations associated with cold acclimation of winter rye seedlings (*Secale cereale* L. cv Puma), *Plant Physiol.*, 83:761–767, 1987; P. L. Steponkus, Cold hardiness and freezing injury of agronomic crops, *Adv. Agron.*, 30:51–98, 1978; P. L. Steponkus, The role of the plasma membrane in freezing injury and cold acclimation, *Annu. Rev. Plant Physiol.*, vol. 35, 1984; P. L. Steponkus and D. V. Lynch, Freeze/thaw-induced destabilization of the plasma membrane and the effects of cold acclimation, *J. Bioenergetics Biomembranes*, 21:21–24, 1989.

Cold storage

Keeping perishable products at low temperatures in order to extend storage life. Cold storage vastly retards the processes responsible for the natural deterioration of the quality of such products at higher temperatures. Time and temperature are the key factors that determine how well foods, pharmaceuticals, and many manufactured commodities, such as photographic film, can retain properties similar to those they possess at the time of harvest or manufacture.

Food that is placed in cold storage is protected from the degradation that is caused by microorganisms. At 80°F (27°C), bacteria will multiply 3000 times in 12–24 h; at 70°F (21°C), the rate of multiplication is reduced to 15 times; and at 40°F (4°C), it is reduced to 2 times. The lower limit of microbial growth is reached at 14°F (−10°C); microorganisms cannot multiply at or below this temperature. *See* BACTERIAL GROWTH; FOOD MICROBIOLOGY.

Cold storage through refrigeration or freezing makes it possible to extend both the seasons of harvest and the geographic area in which a product is available. In the past, food products were grown locally and had to be marketed within a short period of time. Modern cold storage technology makes virtually any product available year-round on a global basis. Other technologies have been combined with refrigeration to further improve this availability. For example, in controlled-atmosphere storage of apples, controlled temperatures above freezing are maintained in a sealed room where the air is also modified to increase its nitrogen content (20.9% oxygen, 0.03% carbon dioxide, and 78.1% nitrogen with other gases) to keep apples orchard-fresh from one fall harvest through the next.

The cold storage food chain begins at the farm or packing plant where the product is chilled by

three principal methods: hydrocooling (immersion in chilled water), forced-air cooling, and vacuum cooling (placing the product in a sealed chamber and creating a vacuum, causing evaporation of some of the water in the product, and subsequent cooling).

If the product is to be frozen, one of five methods is used: (1) air blast freezing (cold air at high velocity is passed over the product); (2) contact freezing (the product is placed in contact with metal plates and heat is drawn off by conduction); (3) immersion freezing (the product is immersed in low-temperature brine); (4) cryogenic freezing [the product is exposed in a chamber to temperatures below −76°F (−60°C) by using liquid nitrogen or liquid carbon dioxide]; and (5) liquid refrigerant freezing (the product is immersed and sprayed with a liquid freezant at atmospheric pressure).

The next step in the cold storage food chain is transport by railroad cars, trucks, airplanes, or boats fitted with refrigeration units that maintain temperatures to critical specifications. Electronic technology permits monitoring of temperature and location by satellite transmission from the transport vehicle to a land-based monitoring station.

Refrigerated warehouses and distribution centers maintain the temperatures required to assure continued maintenance of quality before the product makes its last commercial move to supermarkets or to food-service-industry outlets. Typically, a refrigerated warehouse is a fully insulated structure fitted with refrigeration equipment capable of precise maintenance of specific temperatures in rooms holding up to several million pounds of product. For example, individual rooms may be set to maintain refrigerated storage temperatures of 34°F (1°C) or freezer-room temperatures of 0°F (−18°C). Some food products are being marketed that may require storage temperatures as low as −20°F (−29°C) to maintain texture and to preclude separation of ingredients. *See* REFRIGERATION.

Anhydrous ammonia is the principal refrigerant used in refrigeration systems for cold storage. A number of systems had been introduced that used chlorofluorocarbon refrigerants. However, they are being phased out because of concern about the ozone layer. It is anticipated that such systems will be converted to ammonia or other refrigerants. Most engine rooms are computer-controlled in order to allow optimum use of energy and to transmit alerts if temperatures stray out of the programmed range. *See* GREENHOUSE EFFECT; HALOGENATED HYDROCARBON.

Handling of foods through distribution and cold storage is done by using pallets loaded with 1400–3000 lb (630–1350 kg) of product. Modern freezer rooms are 28–32 ft (8.5–10 m) high for conventional forklift operation, and up to 60 ft (18 m) high when automated or specialized materials-handling equipment is used. *See* FOOD MANUFACTURING; FOOD PRESERVATION; MATERIALS HANDLING.

Michael Shaw

Bibliography. American Society of Heating, Refrigerating, and Air-Conditioning Engineers, *ASHRAE Applications Handbook*, 1992; D. Colin and M. Stringer (eds.), *Chilled Foods: A Comprehensive Guide*, 1992; A. M. Pearson and T. R. Dutson (eds.), *AVI Advances in Meat Research*, vol. 3, 1987; R. C. Wiley (ed.), *Minimally Processed Refrigerated Fruits and Vegetables*, 1994.

Colemanite

One of the more common minerals of the borate group (in which boron is chemically bonded to oxygen).

Physical properties. Colemanite is white to gray with a white streak, is transparent to translucent, and has a vitreous luster. It occurs either as short highly modified prismatic crystals or as massive to granular aggregates. It has perfect cleavage in one direction, parallel to the sheet nature of its crystal structure. Colemanite has hardness of 4–4$^1/_2$ on Mohs scale and a specific gravity of 2.42. *See* HARDNESS SCALES.

Crystal structure. Colemanite is a hydroxy-hydrated calcium borate with the chemical formula $Ca[B_3O_4(OH)_3](H_2O)$. [Ca = calcium, B = boron, O = oxygen, OH = hydroxyl, H_2O = water]. The fundamental building block of the structure is a ring of one triangular ($B\phi_3$) group ($\phi = O^{2-}, OH^-$) and two tetrahedral ($B\phi_4$) groups that is conventionally symbolized as $1\Delta 2\square:\langle\Delta 2\square\rangle$. Here $1\Delta 2\square$ denotes that there is one ($B\phi_3$) group (Δ) and two ($B\phi_4$) groups ($2\square$), the delimiters $\langle\rangle$ indicate the presence of a ring of linked polyhedra, specified between the delimiters (**Fig. 1**). This fundamental building block also occurs in many other borate minerals. In colemanite, the $\langle\Delta 2\square\rangle$ rings share two vertices between triangles and tetrahedra of adjacent rings to form a chain extending along one axis [100]. There is one unique calcium atom coordinated by eight anions in an irregular dodecahedral arrangement.

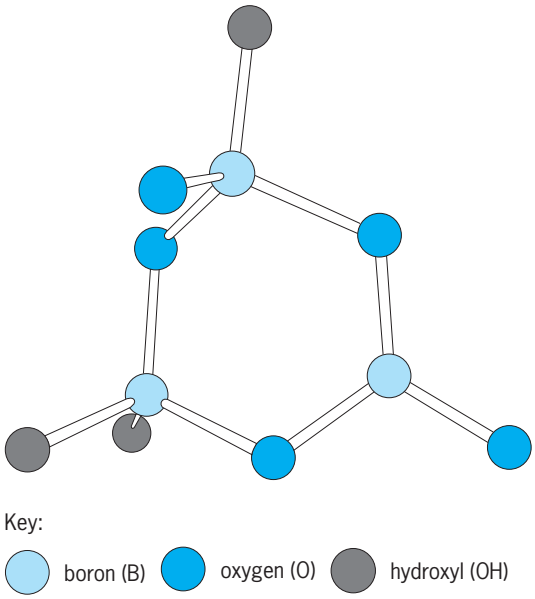

Fig. 1. Fundamental building block in the crystal structure of colemanite.

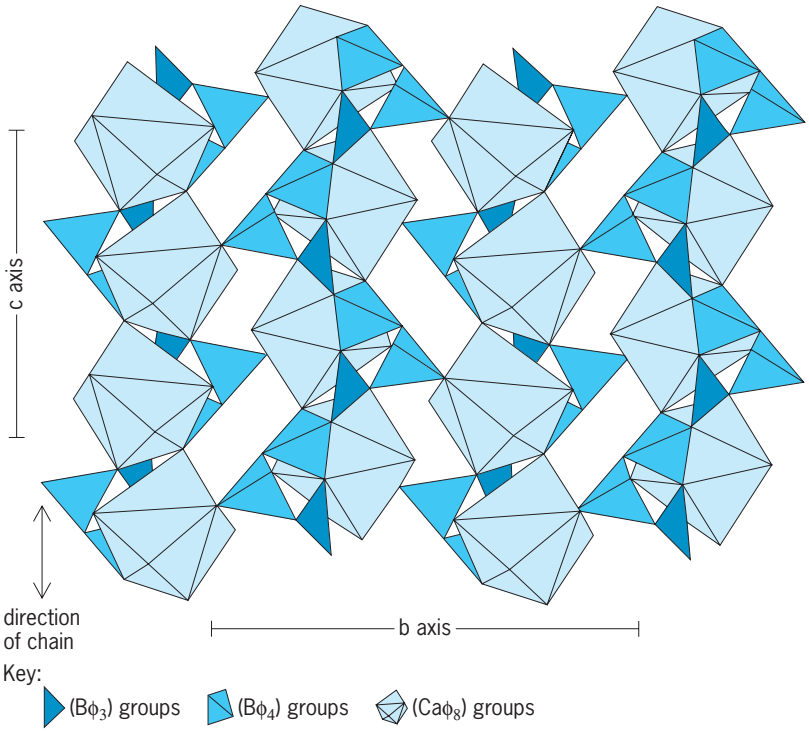

Fig. 2. Crystal structure of colemanite viewed down the *a* axis. The ⟨Δ2□⟩ rings link to form chains of borate polyhedra parallel to the *c* axis.

The ($Ca\phi_8$) polyhedra share corners to form chains parallel to {100} (**Fig. 2**). The [$Ca\phi_7$] chains cross-link the borate chains into a heteropolyhedral sheet of composition [$Ca_2B_3O(OH)(H_2O)$] parallel to {010}. These sheets, seen edge-on in Fig. 2, link via corner-sharing between (BO_4) tetrahedra and ($Ca\phi_8$) dodecahedra, and through an extensive network of hydrogen bonds.

Occurrence. Colemanite typically occurs in lacustrine evaporite deposits of Tertiary age. It forms by thermal diagenesis of primary borate minerals, such as ulexite, $NaCa[B_5O_6(OH)_6](H_2O)_6$, and borax, $Na_2[B_4O_5(OH)_4](H_2O)_8$ (Na = sodium). Both ulexite and borax contain fundamental building blocks that are finite borate clusters consisting of two ⟨Δ2□⟩ rings linked together. During diagenesis, the chemical bonds between these finite borate clusters and the interstitial cations break, and the ⟨Δ2□⟩ rings then polymerize to form chains, as in colemanite. *See* BORATE MINERALS; DIAGENESIS; SALINE EVAPORITES.

Industrial uses. Colemanite is a principal source of boron, together with borax and ulexite, and is mined extensively in California and Nevada, Turkey, and Argentina. Boron is used in a wide variety of industrial commodities: soaps and washing powders, glasses and ceramics, specialty alloys, and fillers in many products. *See* BORON. Frank C. Hawthorne

Bibliography. E. S. Grew and L. M. Anovitz (eds.), *Reviews in Mineralogy*, vol. 33: *Boron: Mineralogy, Petrology and Geochemistry*, 1996; J. D. Grice et al., Borate minerals, II. A hierarchy of structures based on the borate fundamental building block, *Can. Mineral.*, 37:731–762, 1999.

Coleoidea

The cephalopod subclass in the phylum Mollusca that includes living octopods, squids, cuttlefish, and sepiolids (bob-tail squids), and the extinct belemnites. *See* CEPHALOPODA; MOLLUSCA.

General characteristics. The group is defined by the possession of five arm pairs (which are reduced to four in octopods; see **illustration**); arm suckers or hooks with which prey are subdued; chromatophore organs in the skin; a muscular mantle that typically provides jet propulsion; and an internal shell that, although lost in most extant cephalopods, provided buoyancy to the common ancestor.

The shell of belemnites, thought to have been internal because of its unusual microstructure, has a clear buoyancy-regulating structure known as a phragmocone; its discrete chambers are sealed off as the animal grows. As they are emptied of fluid, gas fills them to allow the animal to become neutrally buoyant. The cuttlebone of cuttlefish and the internal shell of the Rams' horn squid, *Spirula spirula*, may be modified from the primitive phragmocone. Most extant coleoids, however, have further reduced the internal support; squids have a thin gladius, and among octopods, only the deep-sea finned taxa have a sizable internal fin cartilage.

Taxonomy and phylogeny. The fossil record of Coleoidea dates back to the Carboniferous. Orders of the extinct belemnite coleoids include the Aulacocerida, Belemnoteuthina, Belemnitida, Diplobelida, and Phragmoteuthida. Phylogenetic relationships among extant coleioids and between living taxa and the extinct belemnites remain unresolved. Although the greatest taxonomic and morphologic diversity of squids and octopods occur in the deep sea, these taxa remain largely unknown. Living coleoids number over 725 species, with many more suspected to remain undiscovered. Squids constitute the order Teuthida (also known as Teuthoidea), which is sometimes recognized as two orders, Myopsida (nearshore squids) and Oegopsida (oceanic squids). The three orders Spirulida (Ram's horn squid), Sepiida (cuttlefish), and Sepiolida (bob-tail

Octopus *Benthoctopus* photographed by the deep-submergence vehicle *ALVIN* at a depth of 2770 m (9140 ft), Gorda Ridge, Northeast Pacific Ocean.

squid) are sometimes treated as the single order Sepioidea. The Vampire squid constitutes the order Vampyromorpha, and order Octopoda contains deep-sea finned octopods, argonauts, midwater pelagic octopus, and the familiar benthic octopuses. *See* BELEMNOIDEA; OCTOPODA; OCTOPUS; SEPIOIDEA; TEUTHOIDEA; VAMPYROMORPHA.

Ecology. Coleoids occupy virtually every habitat in the world's oceans; they are important predators to which many animal taxa potentially fall prey, at least during part of their lives. Several factors, including well-developed brains and large eyes (which are prime examples of evolutionary convergence with vertebrates), contribute to coleiods' effectiveness as predators as well as their defense against predation.

The skin of coleoids is unique in that it contains chromatophore organs, which allow these animals to change color in milliseconds. Coleoids use camouflage to hide from potential predators as well as to sneak up on unsuspecting prey. Chromatophore organs are composed of pigments held in an enervated muscular sac. A section of the coleoid brain controls the animal's coloration by expanding or relaxing the chromatophore organs. *See* CHROMATOPHORE.

The hooks on belemnite arms, the suckers with thorny rings in squids and cuttlefish (secondarily evolved into hooks in some groups), and the muscular gripping suckers in octopods provide excellent prey-handling ability. The ancestral coleoid likely had five arm pairs of equal length. Octopods, having lost one arm pair, likely use stealth (assisted by their elaborate chromatophore-enabled camouflage) and their numerous suckers to secure their prey before injecting it with potent venom. Most other extant coleoids snatch unsuspecting prey from a distance and then draw it to their mouth by using their tentacles, a modified fourth arm pair that are highly extensible, are capable of extremely rapid movement, and have specialized hooks and/or suckers at the tips.

Reproduction and development. Intensive study of a relatively few species indicates that extant coleoids generally live short lives, typified by extremely rapid growth early in the life cycle, which slows with full sexual maturity. In most taxa, egg production occurs only near the end of the life cycle, after which both males and females begin to senesce. Although squids generally leave their eggs to survive unattended, benthic and epipelagic octopuses, such as argonauts, exhibit maternal care of their eggs through development.

Janet R. Voight

Bibliography. R. T. Hanlon and J. B. Messenger, *Cephalopod Behaviour*, Cambridge University Press, 1996; K. N. Nesis, *Cephalopods of the World*, T. F. H. Publications, Neptune City, NJ, 1987; M. Norman, *Cephalopods: A World Guide*, Conch Books, Hackenheim, Germany, 2000.

Coleoptera

An order of Insecta (generally known as beetles) with a complete metamorphosis (with larva and pupa stages), the forewings forming hard protective elytra under which the hindwings are normally folded in repose, and the indirect flight muscles of the mesothorax having been lost. The head capsule is characterized by a firm ventral (gular) closure, the antennae are basically 11-segmented, and mouthparts are of the biting type with 4-segmented maxillary palpi. The prothorax is large and free. The abdomen has sternite I normally absent, sternite II usually membranous and hidden, and segment X vestigial, with cerci absent. The female has one pair of gonapophyses, on segment IX. There are four or six malpighian tubules, which are often cryptonephric. True labial glands are nearly always absent (**Fig. 1**).

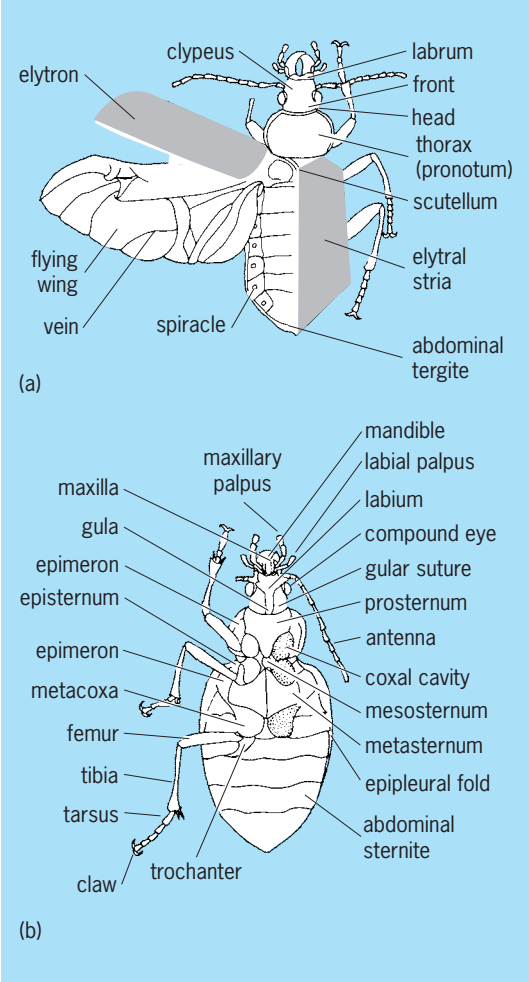

Fig. 1. Diagrammatic sketch of Carabidae member. (*a*) Dorsal view. (*b*) Ventral view.

General Biological Features

The order has well over 250,000 described species, more than any comparable group, showing very great diversity in size, form (**Fig. 2**), color, habits, and physiology. Some Ptiliidae have body lengths as small as 0.01 in. (0.25 mm), and some Cerambycidae may exceed 4 in. (100 mm). Beetles have been found almost everywhere on Earth (except as yet for the Antarctic mainland) where any insects are known, and species of the order exploit almost every habitat and type of food which is used by insects. Many species are economically important.

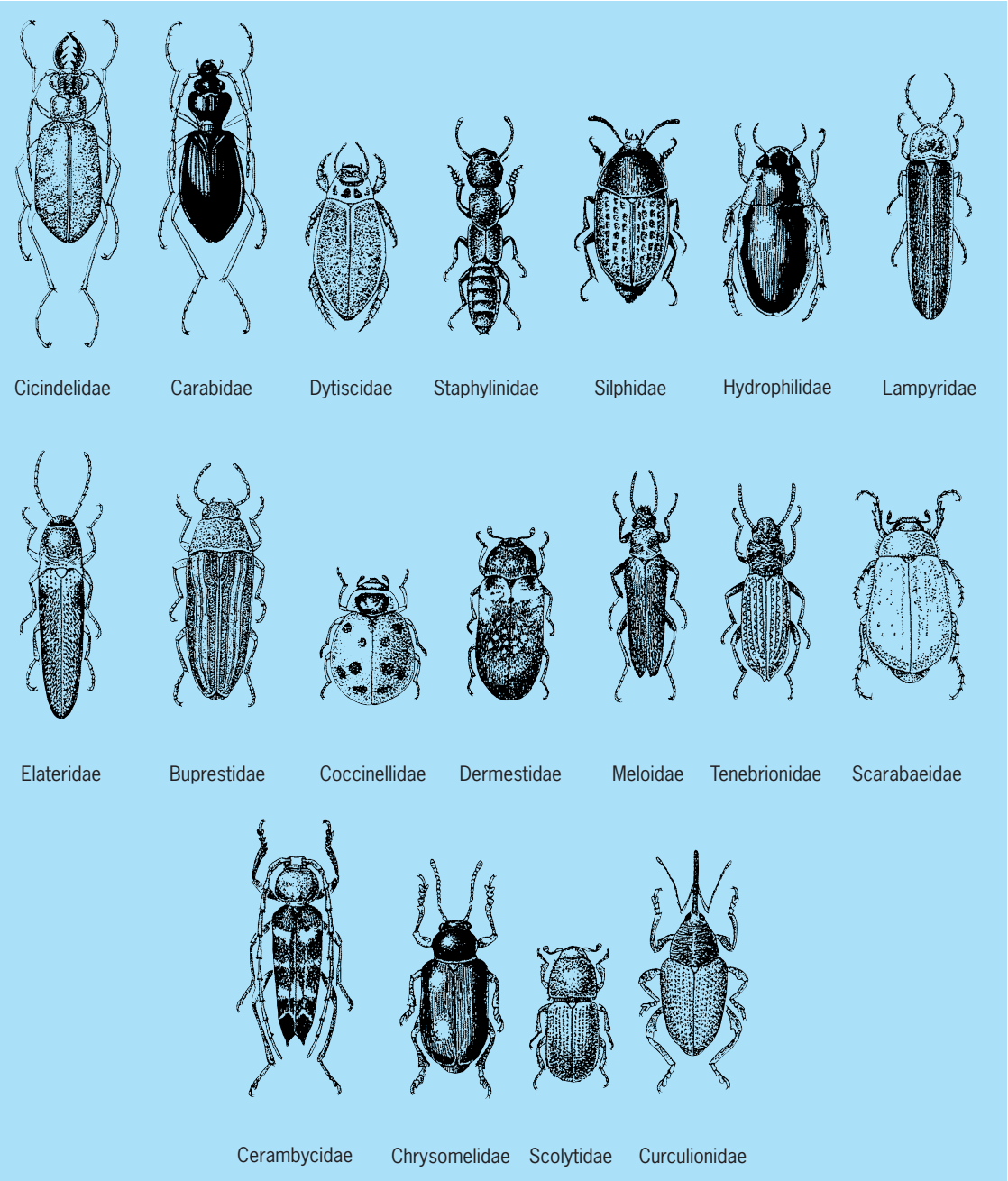

Fig. 2. Representatives of well-known families of Coleoptera. (*After T. I. Storer and R. L. Usinger, General Zoology, 3d ed., McGraw-Hill, 1957*)

A common feature is for the exoskeleton to form an unusually hard and close-fitting suit of armor, correlated with a tendency for the adults to be less mobile but longer-lived than those of other orders of higher insects. It seems to be common in a number of groups for adults to survive and breed in 2 or more successive years, a rare occurrence in other Endopterygota. Most beetles, at least in markedly seasonal climates, have a 1-year life cycle, with developmental stages adapted to specific seasons, but many larger species (particularly when larvae burrow in wood or the soil) may take 2 or more years per generation, and in suitable climates some (usually smaller) species may have more than one generation per year. Another feature unusual in Endopterygota is that many beetles have adults and larvae living in the same habitats and feeding on similar food, though many others have adult foods and habitats quite different from the larval ones. Adult beetles tend to be better runners than most other Endopterygota, and some (such as Cicindelidae) are among the fastest of running insects.

Wings and flight. The flight of most beetles is not very ready or frequent, often requiring prior climbing up onto some eminence, and in many beetles occurs only once or twice in an average life. The elytra may be held in a rigid lateral projection, as in Adephaga, or vibrate through some (usually small) angle

in phase with the hindwings, as in many Polyphaga, or be held vertically upward or over the abdomen, almost in the position of repose. In a number of groups, there is evidence that the elytra contribute significantly to lift or to stability in flight. In males of Stylopidae (Strepsiptera) they are reduced to structures resembling (and probably functioning like) the halteres of Diptera. The hindwings have their insertions shifted very far forward, making the pleura of the metathorax very oblique, and their area is expanded in the anal region. Wing loading and wing-beat frequency are usually higher than in most other Pterygota, with the indirect flight muscles of the metathorax of the asynchronous type, contracting several times in response to a single nerve impulse. Wing venation is adapted to the needs of folding as well as of flight. In some groups, the wings unfold automatically if the elytra are lifted (as in Staphylinidae), but folding after flight is a more complex process, involving active participation of the flexible abdomen. In others unfolding requires considerable abdominal activity, but folding is quick and quasi-automatic, resulting from muscle pulls on the axillary sclerites of the wing base. Most beetles seem to be inherently stable (hence not very maneuverable) in flight. In a number of groups (such as Pleocomidae, Cebrionidae, Drilidae, Phengodidae, and Stylopidae), flight may be confined to the males, indicating that its prime function is to promote gene dispersal.

Cuticle. Apart from a tendency for the sclerotized layers to be thicker, the cuticle of beetles differs little from that of other insects in constitution or ultrastructure. The epicuticular wax layer shows diversity, being particularly well developed in some of the desert Tenebrionidae. Macrochaetae on the body surface may develop into scales (a trait of many Curculionidae) but are rarely plumose; in larval Dermestidae, very complex defensive hairs are developed.

The black coloring which is prevalent in many groups of beetles is produced by deposition of melanin in the process of cuticle hardening. In many groups, with adults active by day, the black is masked by a metallic structural color produced by interference in the surface cuticle layers. Pigmentary colors, mainly ommochromes, giving reddish to yellowish colors, occur in the cuticle of many species, whereas some groups, notably the tortoise beetles (Cassidinae), may have more unstable greenish pigments in the hypodermal cells beneath the cuticle. Pigmentary colors are often adapted to produce aposematic or pseudaposematic appearances, or cryptic ones. *See* PROTECTIVE COLORATION.

Defense adaptations. Special defenses against predation are common in species with long-lived adults, particularly the ground-living species of Carabidae, Staphylinidae, and Tenebrionidae. These groups commonly possess defensive glands, with reservoirs opening on or near the tip of the abdomen, secreting quinones, unsaturated acids, and similar toxic substances. In some cases, the secretion merely oozes onto the body surface, but in others may be expelled in a jet. In the bombardier beetles (Brachinini), the jet contains quinones and is expelled forcibly at nearly 212°F (100°C) through the agency of hydrogen peroxide in an explosion chamber. In other groups, toxins may be dispersed in the body, as in species of Coccinellidae, Lycidae, and Meloidae, and warning (aposematic) coloration is frequently developed.

Other common defensive adaptations include cryptic and mimetic appearances. The most frequent behavioral adaptations are the drop-off reflex and appendage retraction. Numerous Chrysomelidae, Curculionidae, and other adults will react to visual or tactile stimuli by dropping off the plant foliage on which they occur, falling to the ground with retracted appendages, and lying there for some time before resuming activity. In death feigning, the appendages are tightly retracted, often into grooves of the cuticle, and held thus for some time. In such conditions, beetles often resemble seeds and are difficult for predators to perforate or grasp.

Stridulation. Sound production (stridulation) is widespread in beetles and may be produced by friction between almost any counterposed movable parts of the cuticle. Stridulatory organs may show sex dimorphism, and in some species (such as among Scolytinae) may play an important role in the interrelations of the sexes. Elsewhere, as in species of *Cychrus, Hygrobia,* and *Necrophorus,* stridulation is a common response to molestation. Specialized sound-detecting organs, other than Johnston's organ in the antennal pedicel, are not well documented in Coleoptera.

Cryptonephrism. Cryptonephrism, in which the apical parts of the malpighian tubules are closely applied to the hindgut in the region of the rectal glands, is characteristic of a large section of Polyphaga, including all the "stored product" pests and the dominant desert beetles. This condition, not known in other adult insects, has been shown in *Tenebrio* to economize water by extracting it from the rectal contents to produce very dry feces. In cryptonephric groups, certain more or less aquatic members seem to have lost the condition secondarily.

Symbioses. Of the many types of symbiotic relations into which beetles enter with other organisms, perhaps the most ecologically important are those with fungi, which can be divided into ecto- and endosymbioses. In ectosymbiosis, the beetles carry spores or propagules of the fungi and introduce them into suitable new habitats, and are themselves nutritionally dependent (at least as larvae) on the presence of the right type of fungus in the habitat. Classic examples are in numerous bark beetles (Scolytinae) and Lymexylidae. In endosymbiosis, fungi (or sometimes bacteria) are maintained in the gut (or in mycetomes attached to it), usually of the larvae, and contribute either essential nutrients (such as steroids) or essential enzymes. Endosymbionts are usually passed externally or internally to the eggs by the ovipositing female. In endosymbiosis, it seems that as a rule the symbionts are specific to the beetles concerned, to a rather greater degree than in ectosymbionts.

Genetics. The large majority of beetles have an X-y type of chromosomal sex-determining system, with

the y often small (or lost altogether) and forming a "parachute" figure with the X in meiosis. Sex ratios in natural populations are usually 50:50, and sex dimorphism, usually slight, is sometimes very marked. Thelytokous parthenogenesis, often accompanied by polyploidy, is developed in several groups. A haploid male system, like that of Hymenoptera, has been found in Micromalthidae and a few *Xyleborus* ambrosia beetles. A basic complement of nine pairs of autosomes, plus the sex pair, is maintained in many groups of beetles. *See* GENETICS.

Mating and life cycle. Mating in beetles is often a simple process, the male climbing onto the back of the female after a minimum of courtship, but in some groups (such as many Meloidae, Cerambycidae, and Scarabaeidae) more elaborate courtship is the rule. In long-lived species, individuals may copulate several times, often with different partners. Care for the young is commonly limited to the deposition of the eggs in suitable habitats under some sort of cover, but nest construction and provisioning, after the manner of solitary Aculeata Hymenoptera, develop particularly in the dung-feeding Scarabaeinae and Geotrupidae, and in a few Carabidae. Other beetles showing unusual provisions for their offspring include the ambrosia beetles or Attelabidae, some Staphylinidae-Oxytelinae, and some Lamiinae.

Beetle eggs rarely have heavily sculptured chorions or specialized caps for eclosion. Embryonic development is of typically insect types but shows considerable variation within the order. Eggs are often adapted to absorb moisture during embryonic development.

Larvae are very diverse, but usually have a well-developed head capsule with biting mouthparts, often have thoracic legs but very rarely abdominal prolegs (other than a pygopod developed from segment X), and never have true appendages on segment X. The tergite of segment IX often bears paired (sometimes articulated) posterior outgrowths known as urogomphi. First instars often differ considerably from later ones and may have toothlike egg bursters on parts of the body. The number of larval instars usually lies between 3 and 10. Pupae are usually white and exarate, resting on their backs on dorsal spines in some kind of pupal cell. They are never decticous, nor does the adult become active inside the pupal skin.

Hypermetamorphosis, in which two or more sharply different larval forms succeed during individual development, is known in several parasitic or wood-boring groups of beetles. The most common type has an active long-legged first-instar larva which seeks out a suitable feeding site and molts into a grublike type with reduced legs, as in Stylopidae, Rhipiphoridae, and Micromalthidae. The bee-parasitic Meloidae are notable in having two or even three distinct later larval stages.

Adult Anatomy

The head of beetles rarely has a marked posterior neck, and the antennal insertions are usually lateral, rather than dorsal as in most other Endopterygota, a feature possibly related to an original habit of creeping under bark. In most Curculionoidea and a few other groups, the head may be drawn out in front of the eyes to form a rostrum. Paired ventral gular sutures (sometimes partially confluent) are usually present, giving rise to a fairly normal tentorium; a frontoclypeal suture (representing an internal ridge) is often present. Two pairs of ventral cervical sclerites are found in Polyphaga only. Dorsal ocelli are rarely present, and never more than two; compound eyes range from occupying most of the head surface (as in male Phengodidae) to total absence (as in many cavernicolous or subterranean forms). Antennal forms are very various, with the number of segments often reduced, but rarely increased, from the basic 11. Mandible forms are very diverse and related to types of food.

Thorax. The prothorax commonly has lateral edges which separate the dorsal part (notum) from the ventral part (hypomeron) of the tergum. The propleuron may be fully exposed (Adephaga, Archostemata) or reduced to an internal sclerite attached to the trochantin (Polyphaga). The front coxal cavities may or may not be closed behind by inward processes from the hypomera or pleura to meet the prosternal process.

The mesothorax is the smallest thoracic segment. Dorsally, its tergum is usually exposed as a small triangular sclerite (scutellum) between the elytral bases, and ventrally, its sternite may have a median pit receiving the tip of the prosternum. The elytra of beetles develop in a different way from those of earwigs, cockroaches, and such. The first stage is the modification of wing venation to form a series of parallel longitudinal ridges connected by a latticework of crossveins (as seen in many modern Archostemata). The veins then expand to obliterate the remaining membrane, often leaving 9 or 10 longitudinal grooves or rows of punctures representing the original wing membrane and having internal pillars (trabeculae) connecting the upper and lower cuticle. There are various types of dovetailing along the sutural margins of the elytra, and their outer margins are often inflexed to form epipleura.

The form of the metathorax is affected by far-forward insertions of the hindwings and far-posterior ones of the hindlegs, so that the pleura are extremely oblique and the sternum usually long. Usually, the episterna are exposed ventrally, but the epimera are covered by the elytral edges. The metasternum commonly has a median longitudinal suture (with internal keel), and primitively, there may be a transverse suture near its hind edge. Internally, the endosternite takes diverse and characteristic forms.

Tarsi. The tarsi of beetles show all gradations from the primitive five-segmented condition to total disappearance. The number of segments may differ between the legs of an individual or between the sexes of a species. A very common feature in the order is the development of ventral lobes, set with specialized adhesive setae, on some tarsal segments. Often, strong lobes of this kind occur on the antepenultimate segments, and the penultimate one is very small

and more or less fused to the claw joint. Lobed tarsi are particularly characteristic of species frequenting plant foliage.

Abdomen. Most beetles have the entire abdomen covered dorsally by the elytra in repose, and ventrally only five or six sternites (of segments III to VII or VIII) exposed; segment IX is retracted inside VIII in repose. Tergites I–VI are usually soft and flexible, those of VII and VIII more or less sclerotized. In some groups, the elytra may be truncate, leaving parts of the abdomen exposed dorsally, in which case the exposed tergites are sclerotized. Sternite IX is usually a single (internal) sclerite in males, but divided in females to form a pair of valvifers. Male gonapophyses are typically represented by a median basal piece bearing a pair of parameres (lateral lobes), and those of the female by a pair of coxites and styli attached to the valvifers. In females of many groups, an elongate ovipositor is produced by elongation into rods of sclerites of abdominal segment IX.

Elytra. A feature developed in several groups of Coleoptera, but particularly in Staphylinidae, is the abbreviation of the elytra to leave part of the abdomen uncovered in repose, but usually still covering the folded wings. One advantage of this may be to give greater overall flexibility to the body. This is probably of particular advantage to species living in the litter layer in forests, a habitat in which Staphylinidae are strikingly dominant. Perhaps surprisingly, many beetles with short elytra fly readily.

Internal anatomy. Internally, the gut varies greatly in length and detailed structure. The foregut commonly ends in a distended crop whose posterior part usually has some internal setae and may be developed into a complex proventriculus (Adephaga, many Curculionoidea); the midgut may be partly or wholly covered with small papilae (regenerative crypts) and sometimes has anterior ceca. The four or six malpighian tubules open in various ways into the beginning of the hindgut and in many Polyphaga have their apices attached to it (cryptonephric). The central nervous system shows all degrees of concentration of the ventral chain, from having three thoracic and eight abdominal ganglia distinct (as in some Cantharoidea) to having all of these fused into a single mass (as in some higher Scarabaeidae). There are two main types of testes in the males: Adephaga have each testis as a single, long, tightly coiled tube wrapped around in a membrane; Polyphaga have testes usually of several follicles. In the female, Adephaga have polytrophic ovarioles, while Polyphaga have acrotrophic ones, unlike any others known in Endopterygota.

Larval Anatomy

The head capsule in beetle larvae is usually well developed with fairly typical biting mouthparts, with or without a ventral closure behind the labium. Antennae are basically three-segmented, segment II (pedicel) usually bearing a characteristic sensory appendage beside segment III. In several groups, segment I may be divided to give four-segmented antennae (Carabidae, many Staphylinidae, Scarabaeidae), but division of segment III seems to be peculiar to Helodidae. There are up to six ocelli on each side, variously grouped. The frontal sutures (ecdysial lines) often have a characteristic inverted horseshoe form; a median coronal suture may be present. The mandibles have very diverse forms, related to types of food, and the labrum is often replaced by a toothed nasale, particularly in carnivorous types practicing extraoral digestion. The maxillae also vary considerably. In the more specialized predatory types, the cardo tends to be reduced. In some groups, maxillae and labium are more or less fused together to form a complex resembling the gnathochilarium of Millipedes. In some groups with burrowing or internal-feeding larvae, the head capsule may be largely retracted into the prothorax, with corresponding reductions of sclerotization.

Thorax. In the thorax, the prothorax is usually markedly different from the next two segments. The legs are often well developed, particularly in ground-living predatory types, and in Adephaga and Archostemata may have a separate tarsus with two claws. Reduction or loss of legs is characteristic of many burrowing or internal-feeding larvae. The first pair of spiracles is usually situated close to the front border of the mesothorax, lateroventrally; very rarely, a second thoracic spiracle is distinct.

Abdomen. Ten abdominal segments are usually visible, prolegs are rarely developed on any of the first eight segments, but segment X often forms a proleglike pygopod. The tergite of segment IX is often drawn out into posterior processes known as urogomphi, which may become articulated and movable (as in Hydradephaga, Staphylinoidea). In some aquatic larvae, segments IX and X are reduced, and the spiracles of segment VIII become effectively terminal, as in most Diptera larvae. Larval spiracles take diverse forms, which are systematically important.

Body types. The more specialized burrowers among beetle larvae may be divided into two broad types, the straight- and the curved-bodied (**Fig. 3**). Straight burrowing larvae are normal in Archostemata, Carabidae, Elateroidea, various Heteromera, Cerambycidae, and halticine Chrysomelidae, while curved larvae are the norm in Scarabaeoidea, Bostrychoidea, Chrysomelidae-Eumolpinae, and Curculionoidea. The basic difference is probably in the way in which the larva anchors itself in its burrow, using its mandibles to burrow into a resistant medium. In straight larvae, purchase may be secured by short stout legs (as in Elateridae), by terminal abdominal structures, or by protrusible asperate tuberosities (ampullae) on trunk segments (as is the case with Archostemata and Cerambycidae). In both groups, ocelli rarely number more than one and are often lost, and legs are liable to be reduced or lost. The antennae, which are always reduced in straight burrowers, may be quite long in some curved ones, such as Scarabaeoidea.

Physiology

In the basic life functions, such as digestion, respiration, excretion, circulation, muscle contraction,

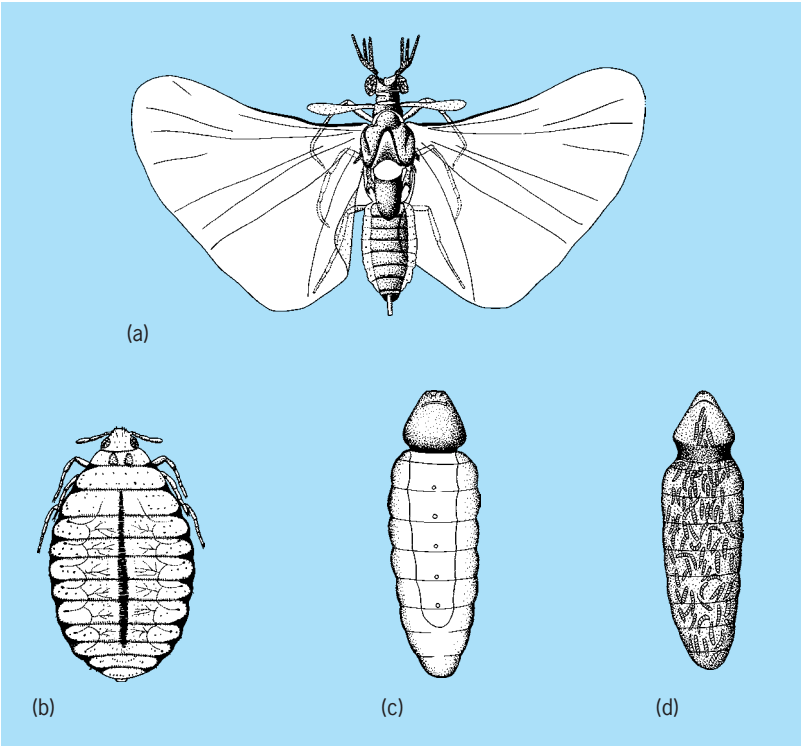

Fig. 3. Stylopidae. (*a*) *Eoxenos laboulbeni*, adult male and (*b*) adult female. (*c*) *Stylops melittae*, adult female (ventral view) and (*d*) gravid female with larvae. (*After C. P. Clausen, Entomophagous Insects, McGraw-Hill, 1940*)

nerve conduction, sense reception, and growth, most beetles show normal insect features, but certain groups are unusual in one or another of these respects. In digestion, some dermestid larvae can digest keratin, a faculty otherwise known only in some tineoid moth larvae. Some wood-boring larvae, notably in Anobiidae and Cerambycidae, have some ability to digest celluloselike substances, apparently through the aid of microorganisms carried in mycetomes associated with the midgut, and some scarabaeid larvae make similar use of microorganisms in a fermentation chamber in the hindgut. Some beetle larvae are able to live and develop normally in fungus-infested decaying wood that is very low in oxygen and high in CO_2. Some aquatic beetles and their larvae may be found living normally in hot springs at temperatures up to 111°F (44°C). In some adult Scolytinae, it has been found that the indirect flight muscles of the metathorax may be greatly reduced and absorbed during gonad development after an initial nuptial flight, yet redeveloped to become functional again with post-reproductive adult feeding. Effective color change in adults, mainly through changes in the hydration of cuticular layers, has been found in certain Scarabaeidae-Dynastinae and Chrysomelidae-Cassidinae. Something approaching homothermy during periods of adult activity has been reported in some of the larger scarabaeoid beetles, utilizing heat generation by muscles in a preliminary warm-up stage, and air-cooling of the thorax by large tracheae during high activity.

Digestion. The digestive organs and capacities of beetles, and their food requirements, are generally like those of other insects, but some groups have unusual features. In many predatory larvae, also in some adults, extraoral digestion occurs, commonly marked in larvae by the development of a rigid toothed nasale in place of the labrum. In some such larvae, particularly when aquatic, the mandibles may become tubular-channeled. In at least some staphylinine larvae, there appear to be "venom" glands developed in the mandibles.

Among herbivores and wood-borers, capacities to digest celluloses and other "difficult" carbohydrates may be developed, usually and perhaps always by the aid of symbiotic microorganisms producing suitable enzymes. Dermestid larvae have developed the ability to digest keratin, breaking its disulfide bonds in an unusually alkaline midgut and producing free hydrogen sulfide as in some tineoid larvae.

Sense organs. The sense organs of beetles, adult or larval, are generally similar in type and function to those of other insects. For vision, the adult compound eyes show great diversity in form and ommatidial structure. Some types, such as *Cicindela*, are adapted for high visual acuity in bright light, others to high sensitivity in poor light, and many show light-dark adaptations of screening pigment. Color vision seems to be restricted to certain groups in the order, notably some of the floricolous Cetoniinae. In certain beetles, specialized infrared sensitivity is evidenced, notably in the region of the vertex of the head in many Curculionidae, and in complex metasternal organs in certain Buprestidae which are attracted to burning trees in coniferous forests. *See* EYE (INVERTEBRATE).

Antennae. Antennae in bettles are mainly receptors for smell and touch, but detection of aerial vibrations (by Johnston's organ in the pedicel) may be important in some. Some of the long hairlike sensilla may detect movements in the air and serve as "wind socks" when beetles are about to fly. Many of the more actively flying groups have the antennal insertions shifted dorsally from their primitive lateral position. The olfactory sensilla, which commonly are thin-walled multiperforate setae, are often concentrated on expanded apical segments, forming a club. This arrangement probably helps in the detection of odor gradients in still air by crawling beetles. Sex dimorphism in the antennae is not uncommon in groups with short-lived adults, the males being specialized to detect female pheromones. Fairly high olfactory acuities for specific substances have been found, especially in some carrion beetles.

Taste. Taste, or contact chemoreception, is usually mediated by sensilla with a single large apical opening. These sensilla are concentrated on the palpi and other mouthparts and are often found on tibiae or tarsi of the front legs (also probably on the ovipositor in many). In water beetles, the smell-taste distinction becomes obscured.

Sound receptors. Specialized sound receptors, other than Johnston's organ, have not been much studied in Coleoptera, though many species produce sounds and in some cases (in various Scolytinae and other Curculionidae, for example) these play a significant

part in courtship. The tibial spurs, in at least some species, serve to detect vibrations in the substratum. *See* ANIMAL COMMUNICATION.

Trichobothria. Very long slender hairs arising from deep cuticular pits occur in many blind cave beetles, along the outer edges of the elytra in most Carabidae, and in larvae of Staphylinidae-Paederinae. These organs are believed to register slight movements in the air, resulting from nearby activity of potential prey or predators.

Gravitational sense. A gravitational sense is evident in many burrowing bettles, which are able to sink accurately vertical burrows despite inclinations of the surface started from. Indications are that such a sense is located in the legs rather than in the antennae, though no beetles are known to have subgenual organs, which may serve this function in other insects.

Temperature sense. A temperature sense is clearly present in some beetles and larvae, but little is known of the sensilla concerned. A type of coeloconic sensillum has been shown to serve this function in palpi of Culicidae, and very similar sensilla have been found in Coleoptera (as on the head capsule of elaterid larvae), which may be sensitive to differences in soil temperature.

Biological clock. What might be called a time sense (otherwise known as a biological clock) clearly operates in many beetles that react specifically to changing daylight lengths marking the seasons in nontropical latitudes. It also manifests itself as a circadian rhythm in "wild" beetles maintained artificially in conditions of uniform temperature, illumination, and humidity. *See* INSECT PHYSIOLOGY; PHOTOPERIODISM.

Food Specialization

The biting mouthparts of many adult and larval beetles are adaptable to many types of food, and fairly widely polyphagous habits are not uncommon in the order, though most species can be assigned to one or another of a few main food categories: fungivores, carnivores, herbivores, or detritivores. Included in the last category are species feeding on decaying animal or vegetable matter and on dung. Most parasites could be included with carnivores, but those myrmecophiles and termitophiles which are fed by their hosts, and those Meloidae which develop on the food stores of bees, form special categories.

There are reasons for thinking that ancestral Coleoptera fed on fungi on dead wood, and beetles of many groups today are largely or wholly fungus eaters. Many are specialists of particular types of fungi, among whose hosts most major groups of fungi are represented. It could well be that the habit of producing fruit bodies and spores underground (as in truffles, and *Endogone*) originally coevolved in relation to the activities of burrowing beetles (such as Liodini and some Scarabaeoidea), which may still be important spore-dispersal agents for such fungi. The long-lasting fruit bodies of various Polyporaceae on dead wood are important breeding grounds for beetles of a number of families, but few beetles can complete their larval development fast enough to exploit the short-lived types of toadstool. Beetles of several families (including the Scolytinae) may serve as effective vectors, inoculating new habitats with fungi which are significant for the beetles' nutrition.

Predators. Predation, mainly on other insects, is probably the fundamental mode of life in adult and larval Adephaga, and has developed in a number of different lines of Polyphaga. Many carnivorous beetles (such as Carabidae and Staphylinidae) are fairly widely polyphagous and liable to supplement their diets with nonanimal matter to some extent, but more specialized predation also occurs (for example, among Histeridae and various Cleridae). Predacious larvae may give rise to largely herbivorous adults (such as in many Hydrophilidae and Melyridae); the converse relation, though rarer, also occurs (in some Cerambycidae, for example). Predatory beetles often do not require movement of the prey to stimulate attack and feeding behavior, and may feed extensively on eggs or pupae of other insects which escape many other predators. Predation on pupae by beetle larvae may lead to parasitic-type habits, as in Lebiini and Brachinini among Carabidae, or Aleocharini among Staphylinidae.

Herbivores. Herbivorous beetles may eat green plants of all groups, from algae to angiosperms, may attack any part of the plant, and display a great range in host plant specialization. Some, like the Colorado beetle, will feed and develop only on a few closely related species of a single genus. Others, like many of the short-nosed weevils, are highly polyphagous, and are able to live on angiosperms of many different families. Feeding on roots, or on the enclosed seeds of Angiospermae, are mainly larval habits, whereas adults usually eat stems, leaves, or flowers. Many species are injurious to crops (**Fig. 4**), and a few have been used successfully in the biological control of weeds (such as *Chrysolina* spp. against *Hypericum*). A number of beetles are effective transmitters of viruses of field crops (such as *Phyllotreta* spp. on Cruciferae); some may be useful pollinators (such as *Meligethes* spp. on Cruciferae). Many botanists believe that beetle pollination was a feature of the first

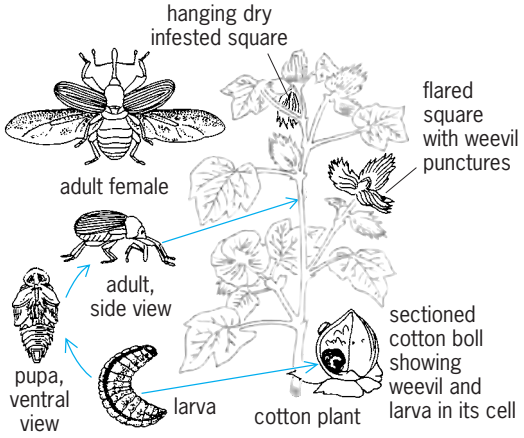

Fig. 4. Cotton boll weevil (*Anthonomus grandis* Boheman). (*After C. H. Metcalf, W. F. Flint, and R. L. Metcalf, Destructive and Useful Insects, 3d ed., McGraw-Hill, 1951*)

Angiospermae and that the enclosed carpels were developed to protect ovules against floricolous beetles.

Adults of a number of families frequent flowers, often eating pollen as well as licking nectar. Botanists distinguish cantharophilous flowers, which depend particularly on beetles for their pollination, and which include what are believed to be some of the most primitive surviving Angiospermae. Characteristic features are white flowers, either large or massed in heads, abundant pollen production, and open shapes, with nectar production often absent.

Larvae developing in flowers or buds are characteristic of Meligethinae and Cateretinae in Nitidulidae. In the latter group, larvae tend to feed on the gynoecium; in the former, on the anthers. Other groups, such as Bruchidae and various weevils, have larvae developing later, in the ripening fruits, and less typically floricolous adults. The tough exines of pollen grains are not easily permeable. For assimilation of the contents, it may be necessary either to crack the outer wall by the mandibles or to cause the grains to begin germination in the crop, as may happen in Oedemeridae and Mordellidae, for example.

Water Beetles

Beetles are unusual among the higher insects in that aquatic habits, where present, usually affect adults as well as larvae. The elytra (like the hemielytra of Heteroptera) may have been preadaptive to the invasion of water. Almost any type of fresh or brackish water body is liable to contain some types of water beetles, though very few species can live permanently in full marine salinities. Almost all aquatic beetles maintain an air reservoir, into which the second thoracic and abdominal spiracles open, under their elytra; and some, mainly small, species may have plastrons or physical gills. Aquatic larvae tend to have a last pair of abdominal spiracles that is large and effectively terminal, or tend to develop tracheal gills. Pupae are terrestrial in the large majority of water beetles, and eggs, when deposited underwater, are commonly laid in contact with airspaces in the stems of water plants, or in an air-filled egg cocoon (Hydrophilidae). Most water beetles are confined to shallow waters, and many of them are ready colonists of temporary pools, through adult flight.

Many of the adults, but fewer of the larvae, are active swimmers by means of the legs, while others rely mainly on crawling for locomotion. The whirligig beetles (Gyrinidae) are exceptional in that the adults swim very fast on the water surface, with only the ventral side submerged, by means of highly modified middle and hind legs, and are able to dive when alarmed, while the larvae are fully aquatic with long tracheal gills. The special modifications of the adults include divided compound eyes, one part for aerial, one for underwater vision; short antennae with a highly developed Johnston's organ, serving to pick up vibrations from potential prey or from the movements of other Gyrinidae in the surface film; long raptorial front legs; and a rudderlike end of the abdomen.

Aquatic larvae and more or less terrestrial adults occur in Psephenidae, some Ptilodactylidae, and the Donaciine leaf beetles; the converse relation is known in at least some Hydraenidae.

The foods of water beetles include algae (Myxophaga, Hydraenidae, many Dryopoidea), higher plants (larval Donaciinae, some Curculionidae, adults of many Hydrophilidae), other small animals (most Hydradephaga, Hydrophilid larvae), and mixed diets.

Special Habitats and Adaptations

More or less unusual adaptations and modes of life are manifest in dung and carrion beetles, ambrosia beetles, cave and subterranean beetles, desert beetles, and luminous beetles. Dung and carrion beetles exploit sporadic and very temporary food resources in which competition (particularly with Diptera larvae) tends to be fierce. Many dung beetles avoid the difficulties by burying dung stores in underground cells where their larvae can develop safely, and some beetles (Necrophorinae) adopt a similar strategy for carrion. They may also develop phoretic and symbiotic relations with specific mites preying on fly eggs.

In ambrosia beetles, the adults usually excavate burrows in wood, carrying with them spores of special fungi, which proceed to develop along the walls of the burrows and are fed on by larvae developing from eggs laid there. This type of relation between beetles, fungi, and trees exists in various forms in a number of families and may be of ecological importance if the fungi concerned are liable to kill trees.

Marginal habitats. More than one family of beetles contains highly adapted types living exclusively in deeper cave systems (cavernicolous) or the deeper layers of the soil (hypogeous). Such species are usually flightless and eyeless, poorly pigmented, and slow-moving, and of very restricted geographical distribution. In the cavernicoles, long trichobothrium-type setae are often developed on the body, and the adults may produce very large eggs singly, with a very abbreviated larval development, approaching the pupiparous condition of some Diptera.

Another marginal habitat in which beetles are the principal insect group represented is the desert, where Tenebrionidae are the dominant group. Here, too, wings are usually lost, but the cuticle tends to be unusually thick, firm, and black. Most desert beetles have burrowing larvae, and the adults also usually take refuge in the soil during the heat of the day. Elaborate adaptations for water economy and dew collection are frequent.

Another marginal habitat in which beetles may be found is hot springs. *Hydroscapha* and certain Hydrophilidae have been found breeding apparently normally in water of temperatures up to 111°F (44°C), which would be lethal for most other animals. A few water beetles are adapted to underground (phreatic) waters, showing parallel features with the cavernicolous and hypogeous terrestrial ones.

Dung and carrion beetles. Dung and carrion provide important but temporary food resources whose

insect users are almost exclusively beetles and Diptera. Dung and carrion beetles are almost always good fliers, with strong antennal clubs bearing heavy concentrations of sensilla. Many of the adults are burrowers and, particularly among dung beetles, frequently make and provision underground nests, thereby removing the food supply for their larvae from the twin dangers of Diptera and drying up. Necrophorini among carrion beetles shows similar features, with the added feature of actual feeding of the larvae by adults. Another special feature of *Necrophorus* and some dung beetles is an apparently symbiotic relation with *Poecilochirus* and allied mites, which are carried on the bodies of adult beetles and which specifically attack the eggs of competing Diptera in the breeding habitat. A well-known feature of many Scarabaeine dung beetles is the formation of a ball of dung which is rolled to a nest site some distance from the dung patch.

Wood borers. One mode of life in which beetles are undoubtedly the dominant group among insects is wood boring, mainly pursued by larvae, but also by adults in some Bostrychidae, Scolytinae, and Platypodinae. Rather few beetles bore normally into the wood of healthy living trees, which are liable to drown borers in sap, but moribund, dead, or decaying timber is liable to heavy attack. Some borers can develop only in unseasoned wood, requiring remains of cell contents; others develop in seasoned wood; and many require the presence of some type of fungus for successful development. Hardwoods and softwoods have their beetle specialists, and females of some Cerambycidae restrict their oviposition to branches or trunks of some particular girth.

Parasites. More or less parasitic relations to animals of other groups have developed in a number of lines of Coleoptera. The closest parallels to Hymenoptera-Parasitica are to be seen in the families Stylopidae (Strepsiptera of many authors; Fig. 3) and Rhipiphoridae. Stylopid larvae are exclusively endoparasites of other insects, while the adult females are apterous, often legless, and remain in the host's body, while males have large fanlike hindwings and the elytra reduced to haltere-like structures. Adult mouthparts are vestigial in both sexes. The more modified of Rhipiphoridae, parasitizing cockroaches, show close parallels with the more primitive Stylopidae, while more primitive ones, probably parasitizing wood-boring beetles, resemble the allied nonparasitic Mordellidae. Larvae ectoparasitic on lignicolous beetle larvae or pupae are found in Passandridae and some Colydiidae.

A considerable variety of Coleoptera develop normally in the nests of termites and social or solitary Hymenoptera-Aculeata, some being essentially detritivores or scavengers, but most feed either on the young or the food stores of their hosts. Many of the myrmecophiles and termitophiles are highly modified structurally and behaviorally and highly host-specific. Some of these beetles appear to have successfully "broken" the chemical or tactile communication codes of their hosts.

Ectoparasitism on birds or mammals is a rare development in beetles, the best-known examples being in Liodidae-Leptininae, with adults and sometimes larvae living on the bodies of small mammals, and Staphylinidae-Amblyopinini, with apparently similar habits. A few Cryptophagidae and Languiriidae-Loberinae have been recorded as adults on bodies of rodents. In the little-known Cavognathidae of the Southern Hemisphere, in at least some species and probably in all of them, the breeding habitat is birds' nests and the larvae feed on nestling birds. *See* POPULATION ECOLOGY.

Luminescence. The beetles are notable in that some of them manifest the highest developments of bioluminescence known in nonmarine animals, the main groups concerned being Phengodidae, Lampyridae, and Elateridae-Pyrophorini, the glowworms and fireflies. In luminescent beetles, the phenomenon always seems to be manifest in the larvae and often in the pupae, but not always in the adults. *See* BIOLUMINESCENCE.

In luminous adult beetles, there is often marked sex dimorphism, and a major function of the lights seems to be the mutual recognition of the sexes of a species. In such cases, there is often a rapid nervous control of the luminescence, and a development of species-specific flash codes. In some cases, notably in adult Pyrophorini and some larval Phengodidae, an individual may carry lights of two different colors.

Another possible function of adult luminosity, and the only one seriously suggested for larvae, is as an aposematic signal. There is definite evidence that some adult fireflies are distasteful to some predators, and the luminous larvae of Phengodidae have dorsal glandular openings on the trunk segments that probably have a defensive function.

Phylogenetic History

Coleoptera are older than the other major endopterygote orders. The earliest fossils showing distinctively beetle features were found rather before the middle of the Permian Period and generally resemble modern Archostemata. By the later Permian, fossils indicate that beetles had become numerous and diverse, and during the Mesozoic Era they appear as a dominant group among insect fossils. Fossils in Triassic deposits have shown features indicative of all four modern suborders, and the Jurassic probably saw the establishment of all modern superfamilies. By early Cretaceous times, it is likely that all "good" modern families had been established as separate lines. In the Baltic Amber fauna, of later Paleogene age (about 40,000,000 years ago), about half the fossil beetle genera appear to be extinct, and the other half still-living representatives (often in remote parts of the world). Beetle fossils in Quaternary (Pleistocene) deposits are very largely of still-living species.

Beetles appear to have been the first insect wood borers, with insect borings in fossil wood of Triassic age (for example, in the Petrified Forest of Arizona) being probably the work of larval

Archostemata. Some fossil woods of slightly younger age have shown the oldest indications of a resin-secreting system, probably a defensive reaction against the attacks of wood-boring beetles. Later still, it seems that beetles may have played a major part in the evolution of the angiosperm-type of flower. The Adephaga probably evolved as ground-living predators, and their pygidial defensive glands may well have been a response to amphibian predation, much like the metathoracic glands of the roughly coeval Hemiptera-Heteroptera. Fossils clearly of adephagan water beetles are known from Lower Jurassic deposits, making a further parallel with Heteroptera-Cryptocerata.

Classification

A modern classification of beetles is given below. This system is adapted from that of Crowson (1955) and is intended to be phylogenetic in the sense of Hennig. An asterisk indicates that the family is not known in the fauna of the United States; a dagger, that the family is not known from the New World.

Suborder Archostemata
- Superfamily Cupedoidea
  - Family: Ommadidae†
    - Tetraphaleridae*
    - Cupedidae
    - Micromalthidae

Suborder Adephaga
- Superfamily Caraboidea
  - Family: Paussidae (including Ozaeninae, Metriinae, Sicindisinae)
    - Cicindelidae
    - Carabidae (including Omophroninae, Rhysodinae)
    - Trachypachidae
    - Amphizoidae
    - Haliplidae
    - Hygrobiidae†
    - Noteridae
    - Dytiscidae
    - Gyrinidae

Suborder Myxophaga
- Superfamily Sphaeroidea
  - Family: Lepiceridae (Cyathoceridae)*
    - Torridincolidae*
    - Hydroscaphidae
    - Sphaeridae (Sphaeriidae)

Suborder Polyphaga
- Series Staphyliniformia
  - Superfamily Hydrophiloidea
    - Family: Hydraenidae
      - Spercheidae
      - Hydrochidae
      - Georyssidae
      - Hydrophilidae
  - Superfamily Histeroidea
    - Family: Sphaeritidae
      - Synteliidae*
      - Histeridae (including Niponiidae)
  - Superfamily Staphylinoidea
    - Family: Ptiliidae (including Limulodidae)
      - Empelidae
      - Liodidae (including Leptininae, Catopinae, Coloninae, Catopocerinae)
      - Scydmaenidae
      - Silphidae
      - Dasyceridae
      - Micropeplidae
      - Staphylinidae (including Scaphidiinae)
      - Pselaphidae
- Series Eucinetiformia
  - Superfamily Eucinetoidea
    - Family: Clambidae
      - Eucinetidae
      - Helodidae
- Series Scarabaeiformia
  - Superfamily Scarabaeoidea
    - Family: Pleocomidae
      - Geotrupidae
      - Passalidae
      - Lucanidae
      - Trogidae
      - Acanthoceridae
      - Hybosoridae
      - Glaphyridae
      - Scarabaeidae
  - Superfamily Dascilloidea
    - Family: Dascillidae
      - Karumiidae*
      - Rhipiceridae (Sandalidae)
- Series Elateriformia
  - Superfamily Byrrhoidea
    - Family Byrrhidae
  - Superfamily Dryopoidea
    - Family: Eulichadidae
      - Ptilodactylidae
      - Chelonariidae
      - Psephenidae
      - Elmidae (Elminthidae)
      - Lutrochidae
      - Dryopidae
      - Limnichidae
      - Heteroceridae
  - Superfamily Buprestoidea
    - Family Buprestidae (including Schizopidae)
  - Superfamily Artematopoidea
    - Family: Artematopidae
      - Callirhipidae
  - Superfamily Elateroidea
    - Family: Cebrionidae
      - Elateridae
      - Throscidae
      - Cerophytidae
      - Perothopidae
      - Phylloceridae†
      - Eucnemidae
  - Superfamily Cantharoidea
    - Family: Brachypsectridae
      - Cneoglossidae*

Family: Plastoceridae†
Homalisidae†
Lycidae
Drilidae†
Phengodidae
Telegeusidae
Lampyridae
Omethidae
Cantharidae
Series Bostrychiformia
Superfamily Dermestoidea
Family: Derodontidae
Nosodendridae
Dermestidae
Thorictidae†
Jacobsoniidae (Sarothriidae)*
Superfamily Bostrychoidea
Family: Bostrychidae (including Lyctinae)
Anobiidae
Ptinidae (including Ectrephini, Gnostini)
Series Cucujiformia
Superfamily Cleroidea
Family: Phloiophilidae†
Series: Peltidae
Family: Lophocateridae
Trogossitidae
Chaetosomatidae†
Cleridae
Acanthocnemidae
Phycosecidae†
Melyridae (including Malachiinae)
Superfamily Lymexyloidea
Family: Lymexylidae
Stylopidae (Strepsiptera, Stylopoidea)
Superfamily Cucujoidea
Family: Nitidulidae
Rhizophagidae (including Monotominae)
Protocucujidae*
Sphindidae
Boganiidae†
Cucujidae
Laemophloeidae
Passandridae
Phalacridae
Phloeostichidae*
Silvanidae
Cavognathidae*
Cryptophagidae (including Hypocoprinae)
Lamingtoniidae†
Helotidae†
Languriidae
Erotylidae (including Pharaxonothinae)
Biphyllidae
Cryptophilidae (including Propalticinae)†
Cerylonidae
Corylophidae
Alexiidae (Sphaerosomatidae)†
Endomychidae
Coccinellidae
Discolomidae
Merophysiidae
Lathridiidae
Byturidae
Mycetophagidae
Cisidae
Pterogeniidae†
Tetratomidae
Melandryidae
Mordellidae
Rhipiphoridae
Merycidae†
Colydiidae
Synchroidae
Cephaloidae
Pythidae
Pyrochroidae
Anthicidae
Meloidae
Aderidae
Scraptiidae (including Anaspidinae)
Oedemeridae
Cononotidae
Othniidae
Salpingidae
Inopeplidae
Mycteridae (including Hemipeplinae)
Monommidae
Zopheridae
Tenebrionidae (including Lagriinae, Nilioninae, Alleculinae)
Superfamily Chrysomeloidea
Family: Disteniidae (including Oxypeltinae, Philinae,† Vesperinae)
Cerambycidae
Megalopodidae (including Zeugophorinae)
Bruchidae (including Sagrinae)
Chrysomelidae
Superfamily Curculionoidea
Family: Nemonychidae
Anthribidae (including Bruchelinae)
Belidae*
Oxypeltidae*
Aglycyderidae (Proterhinidae)†
Allocorynidae
Attelabidae
Apionidae (including Antliarrhininae,† Ithycerinae, Nanophyinae)
Brenthidae
Curculionidae (including Scolytinae and Platypodinae)

This classification is based on a great variety of characteristics, including internal ones and those of immature stages, and takes account of the fossil evidence. Important characteristics include the

structure of the propleura, the cervical sclerites, the ovarioles, larval leg segments, form of testes, wing venation and folding, number and arrangement of malpighian tubules, forms of the female and male external genitalia, presence or absence of a free larval labrum, larval spiracular structures, mode of formation of adult midgut in the pupa, form of the metendosternite, number of female accessory glands, form of adult antennae, tarsal segmentation, presence of pupal "gin traps," number of pupal spiracles, number of adult abdominal spiracles, presence of defensive glands in adults or larvae, egg bursters of the first-instar larva, form of adult tentorium and gular sutures, forms of adult coxal cavities, and internal structure of adult ommatidia.

Geographical Distribution

Almost every type of continuous or discontinuous distribution pattern which is known in any animal group could be matched in some taxon of Coleoptera, and every significant zoogeographical region or area could be characterized by endemic taxa of beetles. In flightless taxa, distributional areas are generally more limited than those of comparable winged taxa. Distinct distributional categories can be seen in those small, readily flying groups (in Staphylinidae and Nitidulidae, for example) which are liable to form part of "aerial plankton," and in those wood borers which may survive for extended periods in sea-drifted logs, both of which are liable to occur in oceanic islands beyond the ranges of most other beetle taxa. Climatic factors often seem to impose limits on the spread of beetle species (and sometimes of genera or families), and beetle remains in peats have been found to be sensitive indicators of climatic changes in glacial and postglacial times. *See* INSECTA; SOCIAL INSECTS. Roy A. Crowson

Bibliography. R. A. Crowson, *The Natural Classification of the Families of Coleoptera,* 1955; R. A. Crowson, The natural classification of the families of Coleoptera: Addenda and corrigenda, *Entomol. Mon. Mag.,* 103:209–214, 1967; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Coliiformes

A small order of birds containing only the family Coliidae with six species (the mousebirds) restricted to Africa. Mousebirds, or colies, are small, grayish to brownish, with a crest on the head and long tail (see **illustration**). The legs are short and the feet strong, with the four toes movable into many positions, from all four pointing forward to two reversed backward. Mousebirds perch, climb, crawl, and scramble agilely in bushes and trees. They are largely vegetarian but eat some insects. They are nonmigratory and gregarious, living in flocks and sleeping in clusters, but they are monogamous in loose colonies as breeders. The nest is an open cup in a tree or bush, and the two to four young remain in the nest, cared for by both adults until they can fly.

White-headed mousebirds (*Colius leucocephalus*). (© *2004 DonGettyPhoto.com*)

The relationships of the mousebirds to other birds are obscure. Mousebirds have a surprisingly good fossil record from the Miocene of France and Germany, suggesting a wider distribution of this group. *See* AVES. Walter J. Bock

Bibliography. E. de Juana, Order Coliidae (Mousebirds), in J. del Hoyo et al. (eds.), *Handbook of the Birds of the World*, Lynx Edicions, vol. 6, pp. 60–77, 2001; H. Schifter, Systematics and distribution of the mousebirds (Coliidae), in *Proceedings of the International Symposium of African Vertebrates*, 1985.

Coliphage

Any bacteriophage able to infect the bacterium *Escherichia coli*. Many are phages able to attack more than one strain of this organism. The T series of phage (T1–T7), propagated on a special culture of *E. coli*, strain B, have been used in extensive studies from which most of the knowledge of phages is derived.

A typical phage particle is shaped something like a spermatozoon. The outer skin, or shell, consists of protein and encloses contents of deoxyribonucleic acid (DNA). This DNA can be separated from the protein by osmotic shock, and it is then found that the protein shell alone can cause death of a bacterium when adsorbed to the cell surface. This does not constitute a true infection, however, since no new phage is formed.

The usual lytic infection of a *coli* B cell by a T phage proceeds as follows (see **illus.**). The phage is adsorbed, by its tail, to the cell surface. Following this adsorption, the phage DNA gains access to the cell. The protein skin has now served its purpose and, if desired, can be removed from the cell surface without affecting subsequent events. Phage in

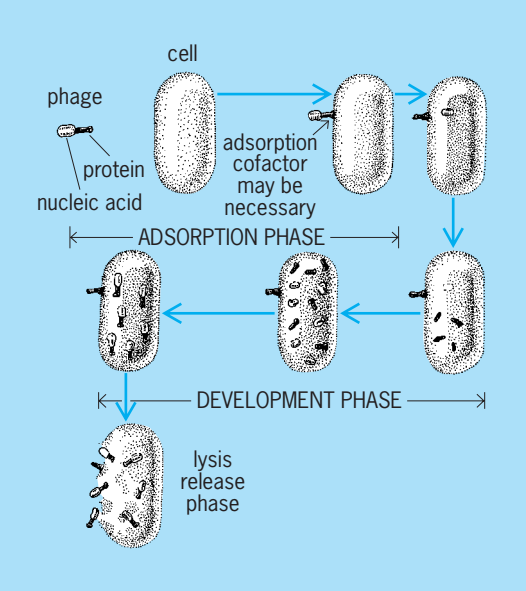

Steps in the two phases of a lytic infection by phage. (*After E. L. Oginsky and W. W. Umbreit, An Introduction to Bacterial Physiology, Freeman, 1954*).

an infectious state no longer exists; it is now spoken of as vegetative phage and will remain in this state during its life cycle until the production of new infective particles occurs. Immediately following the injection of the DNA, the metabolic activities of the cell are profoundly altered. Normal growth ceases abruptly, although respiration continues at a steady rate. Cell enzymes still function, but their activity is now redirected and controlled by "information" supplied by the phage DNA. New protein synthesis can be detected at once; this is not bacterial protein, however, but phage protein which is to serve later as part of the new phage particles. Synthesis of bacterial ribonucleic acid (RNA) ceases. Bacterial DNA breaks down, and after about 10 min, phage DNA begins to make its appearance, formed in part from the bacterial DNA and in part from new synthesis. This new DNA is incorporated with the phage protein to make a new phage which can now be detected as an infectious unit. In another 10 min or so, lysis occurs after more phage has accumulated within the cell. The time will vary with different phages. The infectious phage, released in numbers of 100 or more particles, is able to infect susceptible bacteria. *See* BACTERIOPHAGE; LYSOGENY; LYTIC INFECTION; VIRUS.

Philip B. Cowles

Collagen

The major fibrous protein in animals. Collagen is present in all types of multicellular animals and is probably the most abundant animal protein in nature. It is estimated that collagen accounts for about 30% of the total human body protein. Collagen is located in the extracellular matrix of connective tissues. It is part of the interacting network of proteoglycans and proteins that provides a structural framework for both soft and calcified connective tissues. By self-associating into fibrils and by binding to proteoglycans and other matrix components, collagen contributes to tissue integrity and mechanical properties. Collagen interacts with cells through the integrin cell receptors and mediates cellular adhesion and migration. Important roles for collagen have been identified in development, wound healing, platelet aggregation, and aging. Its commercial importance in leather and the production of gelatin and glue have long been recognized. More recently, it is being used as a basis for biomaterials. Examples of its biomedical applications include injectable collagen to lessen facial wrinkles and defects; surgical collagen sponges to increase blood clotting; and artificial skin for the treatment of burns. *See* GELATIN; LEATHER AND FUR PROCESSING.

Structure and properties. The classification of an extracellular matrix protein as a collagen is based on the presence of a domain with the distinctive triple-helical conformation. The collagen triple helix consists of three polypeptide chains, each of which adopts an extended polyproline II–like helix. The three chains are supercoiled about a common axis and linked by hydrogen bonds. Only recently were the detailed features of this conformation determined by x-ray crystallography, confirming the general structure and showing an extensive network of ordered water hydrogen bonded to the triple helix.

The triple-helical conformation requires unique amino acid sequence features. Glycine, which is the smallest amino acid, must be present as every third residue. A high content of the sterically restricted imino acids, proline and hydroxyproline, is also necessary to stabilize the extended helix. Hydroxyproline provides additional stability through water-mediated hydrogen bonds. Collagen is the only animal protein, other than elastin, with significant amounts of hydroxyproline and hydroxylysine residues. It is easy to identify a collagen triple helix from its amino acid sequence pattern of repeating glycine-X-Y sequences, and glycine-proline-hydroxyproline is the most frequent tripeptide sequence (see **illus.**).

Types. At least 19 distinct molecules have been classified as collagens, and specific types are associated with particular tissues. The most prevalent and well-studied collagens belong to the fibril-forming or interstitial collagen family. These collagens form the characteristic fibrils seen by electron microscopy with a 67-nanometer repeating banding pattern. The banding pattern arises from the axial staggering of the collagen molecules. The molecules in a fibril are covalently cross-linked by an enzymatic mechanism to strengthen and stabilize them. Inhibition of the enzyme involved in cross-linking results in a dramatic decrease in the tensile strength of tissues, a condition known as lathyrism.

Type I is the most common fibril-forming collagen. Its fibrils make up the mineralized matrix in bone, the strong parallel bundles of fibers in tendon, and the plywoodlike alternating layers in the transparent cornea. Type II is the major fibril-forming collagen

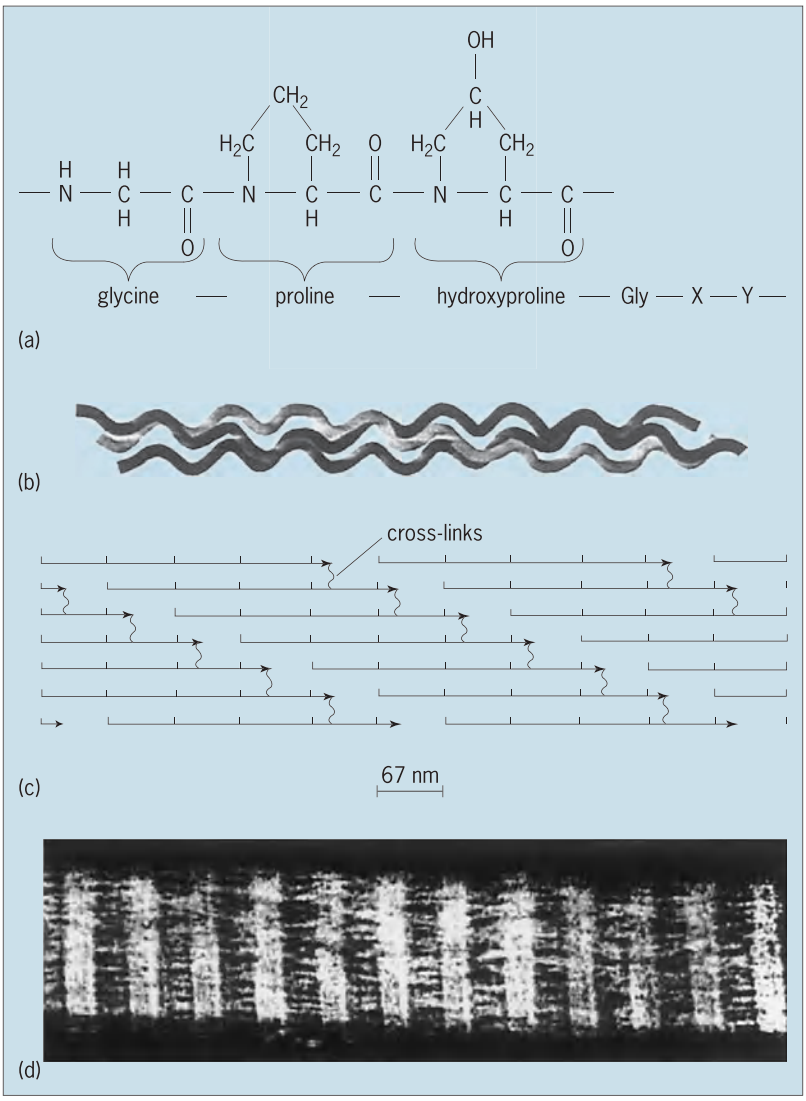

Levels of collagen structure. (*a*) Typical amino acid repeating sequence. (*b*) Triple-helical molecular structure. (*c*) Staggered arrangement of molecules which generates the 67-nm periodic fibrils. (*d*) Electron micrograph of collagen fibril.

in cartilage, while type III is found in blood vessels and skin, together with type I. The family known as fibril-associated collagens with interrupted triple helices, or FACIT collagens, decorates the outside of the 67-nm periodic fibrils. The FACIT collagens appear to link fibrils and bind to other matrix components.

Basement membranes, which serve to separate cell layers and act as filtration barriers, contain a distinctive group of collagens, denoted as type IV collagens. Type IV collagens have many breaks in the repeating glycine-X-Y pattern of their triple helix. These collagens are organized into a network or meshlike sheet structure. In the kidney glomerulus, the network based on type IV collagen acts as a filter to determine which molecules will pass from the blood into the urine.

A number of proteins involved in host defense functions also contain collagenous amino acid sequences that adopt a triple-helical conformation. These include the complement protein C1q and mannose-binding protein. The collagenous domains of these proteins associate into hexamers or other oligomers to form the functional molecule, and are also involved in receptor and enzyme binding. *See* BONE; CONNECTIVE TISSUE; FIBROUS PROTEIN.

Biosynthesis and degradation. The genes coding for different collagen types have many exons, and the biosynthesis of collagen is a complex multistage process. The synthesis of collagen takes place in the endoplasmic reticulum. Hydroxyproline and hydroxylysine are formed by enzymes after proline and lysine are incorporated into polypeptide chains, but before the chains are folded into a triple helix. Vitamin C (ascorbic acid) is required for the hydroxylating enzymes; the failure to make hydroxyproline residues and stable collagen leads to the fragile blood vessels and skin lesions found in scurvy. After three chains associate to form the triple helix, the molecule is secreted by the cell and, following cleavage of precursor regions, fibril formation or other molecular association occurs in the matrix. Enzymes required for the posttranslational modifications of collagen include hydroxylases, enzymes involved in the attachment of sugars, proteases which cleave precursor regions, and lysyl oxidase for cross-linking.

An orderly breakdown of collagen is necessary during development and tissue remodeling. For instance, following childbirth, the uterus reduces in size, which involves a massive degradation of collagen. Collagen is resistant to degradation by most vertebrate proteases, but it is cleaved by several specific matrix metalloproteinases (collagenases). These zinc-requiring enzymes cleave fibril-forming collagens at one specific site. An abnormal increase in the degradation of cartilage collagen is seen in osteoarthritis. Collagen breakdown also appears to be essential for tumor metastases. With the recent determination of the three-dimensional structures of several matrix metalloproteinases, there are active attempts by pharmaceutical companies to develop enzyme inhibitors for use in arthritis and cancer therapies.

Diseases. Collagen is impaired in both inherited and acquired diseases involving connective tissues. A number of hereditary diseases have been shown to be due to mutations in specific collagen genes. Osteogenesis imperfecta (brittle bone) disease is characterized by fragile bones and is due to mutations in type I collagen, the major collagen in bone. Some cartilage disorders are caused by mutations in type II collagen. Ruptured arteries are found in Ehlers-Danlos syndrome type IV, which arises from mutations in type III collagen, important in blood vessels. The most common type of mutations in collagen are single-base substitutions that result in a substitution of glycine by another amino acid. Such a substitution breaks the repeating glycine-X-Y triplet pattern which is essential to form a good triple helix. Transgenic mice are being used as models for human collagen diseases.

Barbara Brodsky

Bibliography. B. Brodsky and J. A. M. Ramshaw, The collagen triple-helix structure, *Matrix Biol.*,

15:545–554, 1997; M. E. Nimni (ed.), *Collagen*, 3 vols, CRC Press, Boca Raton, 1988; D. J. Prockop and K. I. Kivirikko, Collagen: Molecular biology, diseases and potential for therapy, *Annu. Rev. Biochem.*, 64:403–434, 1995; P. M. Royce and B. Steinmann (eds.), *Connective Tissue and Its Heritable Disorders*, Wiley-Liss, New York, 1993.

Collard

A cool-season biennial crucifer, *Brassica oleracea* var. *acephala*, similar to nonheading cabbage. Collard is of Mediterranean origin and is grown for its rosette of leaves which are cooked fresh as a vegetable (see **illus.**). Kale and collard differ only in the form of their leaves; both have been referred to as coleworts, a name taken from the Anglo-Saxon term meaning cabbage plants.

Collard (*Brassica oleracea* var. *acephala*).

Propagation is by seed. Cultural practices are similar to those used for cabbage; however, collards are more tolerant of high temperatures. Georgia and Vates are popular varieties (cultivars). Collard is moderately tolerant of acid soils. Harvesting is usually 75 days after planting. Important production centers are in the southern United States, where collards are an important nutritious green, especially during winter months. *See* CABBAGE; CAPPARALES; KALE.

H. John Carew

Collembola

An order of primitive insects, commonly called springtails, belonging to the subclass Apterygota. These tiny insects do not undergo a metamorphosis. They have six abdominal segments, some of which may be ankylosed to give an apparently smaller number (see **illus.**).

The tarsi are united with the tibiae, and a pretarsus is present, which bears a dorsal median claw

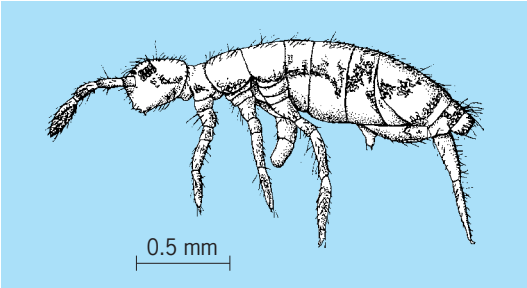

A collembolan, *Entomobrya cubensis*. (*After J. W. Folsom, Proc. U.S. Nat. Mus., 72*(6), *plate 6, 1927*).

and usually a ventral clawlike appendage. The ventral tube always occurs on the first abdominal segment and connects with the mouthparts through a ventral canal consisting of two integumentary folds which overlap. Appended to the fourth abdominal segment in most of the species is a spring, or furcula. This consists of a basal piece and two apical parallel structures which end in hooks. Another pair of structures, united at the base, occur ventrally on the third abdominal segment. These bear teeth which engage with the furcula and hold it beneath the body. When released, the spring flies backward, catapulting the insect. Two families, Sminthuridae and Actaletidae, have simple tracheal systems, while the others breathe through the integument. Collembola live in humid places, often in leaf mold. Some species are active at cool temperatures, hence another common name, snowflea, is used. The earliest known insect fossil, *Rhyniella praecursor*, is thought to belong to this order.

Harlow B. Mills

Bibliography. H. B. Mills, *A Monograph of the Collembola of Iowa*, Iowa State Coll. Div. Ind. Sci. Monogr. 3, 1934.

Collenchyma

A primary, or early differentiated, supporting tissue of young shoot parts appearing while these parts are still elongating. It is located near the surface, usually just under the epidermis (**illus.** *a*). When observed in transverse sections, it is characterized structurally by cell walls that are intermittently thickened, generally in the corners or places of juncture of three or more cells (illus. *b* and *c*). Collenchyma is typically formed in the petioles and vein ribs of leaves, the elongating zone of young stems, and the pedicels of flowers. *See* CELL WALLS (PLANT).

As in parenchyma, the cells in collenchyma are living and may contain chloroplasts and starch grains. The cell wall of a collenchyma cell is its most striking feature structurally and functionally. It is composed of cellulose and pectic compounds plus a very high proportion of water. In some studies collenchyma cell walls were found to contain 50–60% water by fresh weight in a variety of species examined. The cellulose and pectic compounds are present in alternating layers. A striking feature of collenchyma cell walls is their plasticity. They are capable of great

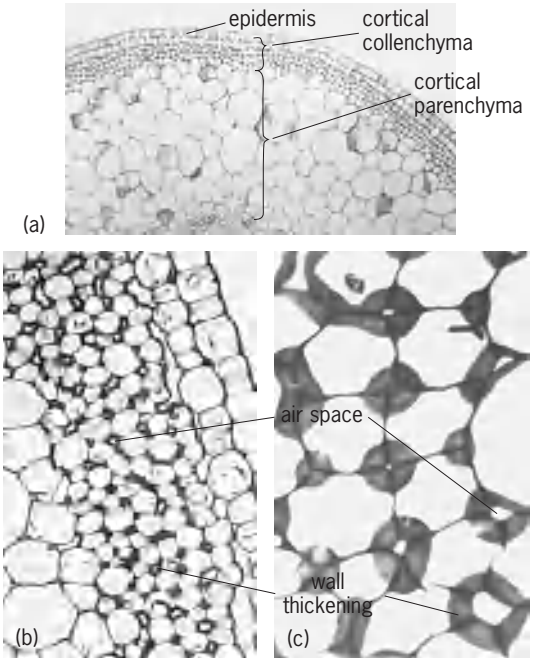

Transverse sections of Jimsonweed (*Datura stramonium*), progressively magnified. (*a*) Stem section. (*b*) Stem section, showing the corner thickenings of collenchyma cell walls and air spaces surrounded by wall thickenings. (*c*) Collenchyma section, showing lamellation of the cell wall forming the corner thickenings. (*Courtesy of W. S. Walker*)

elongation during the period of growth in length of the plant. Despite this elongation, the characteristic thickening of the wall continues to increase because new wall material is added during growth. The plasticity of collenchyma is associated with a tensile strength comparable to that shown by fibers of sclerenchyma. The combination of strength and plasticity makes the collenchyma effective as a strengthening tissue in developing stems and leaves having no other supporting tissue at that time. *See* CELLULOSE; CHLOROPHYLL; EPIDERMIS (PLANT); FLOWER; LEAF; PARENCHYMA; PECTIN; STEM. Robert L. Hulbary

Bibliography. J. O. Mauseth, *Plant Anatomy*, 1988.

Collision (physics)

Any interaction between particles, aggregates of particles, or rigid bodies in which they come near enough to exert a mutual influence, generally with exchange of energy. The term collision, as used in physics, does not necessarily imply actual contact.

In classical mechanics, collision problems are concerned with the relation of the magnitudes and direction of the velocities of colliding bodies after collision to the velocity vectors of the bodies before collision. When the only forces on the colliding bodies are those exerted by the bodies themselves, the principle of conservation of momentum states that the total momentum of the system is unchanged in the collision process. This result is particularly useful when the forces between the colliding bodies act only during the instant of collision. The velocities can then change only during the collision process, which takes place in a short time interval. Under these conditions the forces can be treated as impulsive forces, the effects of which can be expressed in terms of an experimental parameter known as the coefficient of restitution, which is discussed later. *See* CONSERVATION OF MOMENTUM; IMPACT.

The study of collisions of molecules, atoms, and nuclear particles is an important field of physics. Here the object is usually to obtain information about the forces acting between the particles. The velocities of the particles are measured before and after collision. Although quantum mechanics instead of classical mechanics should be used to describe the motion of the particles, many of the conclusions of classical collision theory are valid. *See* SCATTERING EXPERIMENTS (ATOMS AND MOLECULES); SCATTERING EXPERIMENTS (NUCLEI).

Classification. Collisions can be classed as elastic and inelastic. In an elastic collision, mechanical energy is conserved; that is, the total kinetic energy of the system of particles after collision equals the total kinetic energy before collision. For inelastic collisions, however, the total kinetic energy after collision is different from the initial total kinetic energy.

In classical mechanics the total mechanical energy after an inelastic collision is ordinarily less than the initial total mechanical energy, and the mechanical energy which is lost is converted into heat. However, an inelastic collision in which the total energy after collision is greater than the initial total energy sometimes can occur in classical mechanics. For example, a collision can cause an explosion which converts chemical energy into mechanical energy. In molecular, atomic, and nuclear systems, which are governed by quantum mechanics, the energy levels of the particles can be changed during collisions. Thus these inelastic collisions can involve either a gain or loss in mechanical energy.

Consider a one-dimensional collision of two particles in which the particles have masses m_1 and m_2 and initial velocities u_1 and u_2. If they interact only during the collision, an application of the principle of conservation of momentum yields Eq. (1), where

$$m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2 \tag{1}$$

v_1 and v_2 are the velocities of m_1 and m_2, respectively, after collision.

Coefficient of restitution. It has been found experimentally that in collision processes Eq. (2) holds,

$$e = \frac{v_2 - v_1}{u_1 - u_2} \tag{2}$$

where e is a constant known as the coefficient of restitution, the value of which depends on the properties of the colliding bodies. The magnitude of e varies from 0 to 1. A coefficient of restitution equal to 1 can be shown to be equivalent to an elastic collision, while a coefficient of restitution of zero is equivalent to what is sometimes called a perfectly inelastic collision. From the definition of e one can show that in a perfectly inelastic collision the

colliding bodies stick together after collision, as two colliding balls of putty or a bullet fired into a wooden block would do. Equations (1) and (2) can be solved for the unknown velocities v_2 and v_1 in the one-dimensional collision of two particles.

The concept of coefficient of restitution can be generalized to treat collisions involving the plane motion of smooth bodies—both of particles and larger bodies for which rotation effects must be considered. For these collisions, experiments show that the velocity components to be used in Eq. (2) for e are the components along the common normal to the surfaces of the bodies at the point where they make contact in the collision. For smooth bodies the velocity components perpendicular to this direction are unchanged. Use of this result and the principle of conservation of momentum is sufficient to solve two-dimensional collision problems of smooth bodies. For collisions of smooth spheres the velocity components to be used in Eq. (2) for e are those on the line joining the centers of the spheres. Velocity components perpendicular to this direction are unchanged.

Center-of-mass coordinates. A simplification of the description of both classical and quantum mechanical collisions can be obtained by using a coordinate system which moves with the velocity of the center of mass before collision. (Since for an isolated system the center of mass of the system can be shown to be unaccelerated at all times, the velocity of the center of mass of the system of particles does not change during collision.) The coordinate system which moves with the center of mass is called the center-of-mass system, while the stationary system is the laboratory system.

The description of a collision in the center-of-mass system is simplified because in this coordinate system the total momentum is equal to zero, both before and after collision. In the case of a two-particle collision the particles therefore must be oppositely directed after collision, and the magnitude of one of the velocities in the center-of-mass system can be determined if the other magnitude is known. *See* CENTER OF MASS.

Paul W. Schmidt

Bibliography. D. Halliday, R. Resnick, and K. Krane, *Physics*, 4th ed., 1992; F. W. Sears, M. W. Zemanski, and H. D. Young, *University Physics*, 7th ed., 1987.

Colloid

A state of matter characterized by large specific surface areas, that is, large surfaces per unit volume or unit mass. The term colloid refers to any matter, regardless of chemical composition, structure (crystalline or amorphous), geometric form, or degree of condensation (solid, liquid, or gas), as long as at least one of the dimensions is less than approximately 1 micrometer but larger than about 1 nanometer. Thus, it is possible to distinguish films (for example, oil slick), fibers (spider web), or colloidal particles (fog) if one, two, or three dimensions, respectively, are within the submicrometer range.

Types of colloid dispersions

Medium	Dispersed matter	Technical name	Examples
Gas	Liquid	Aerosol	Fog, sprays
	Solid	Aerosol	Smoke, atmospheric or interstellar dust
Liquid	Gas	Foam	Head on beer, lather
	Liquid	Emulsion	Milk, cosmetic lotions
	Solid	Sol	Paints, muddy water
Solid	Gas	Solid foam	Foam rubber
	Liquid	Solid emulsion	Opal
	Solid	Solid sol	Steel

A colloid consists of dispersed matter in a given medium. In the case of finely subdivided particles, classification of a number of systems is possible, as given in the **table**. In addition to the colloids in the table, there are systems that do not fit into any of the listed categories. Among these are gels, which consist of a network-type internal structure loaded with larger or smaller amounts of fluid. Some gels may have the consistency of a solid, while others are truly elastic bodies that can reversibly deform. Another colloid system that may occur is termed coacervate, and is identified as a liquid phase separated on coagulation of hydrophilic colloids, such as proteins. *See* GEL.

It is customary to distinguish between hydrophobic and hydrophilic colloids. The former are assumed to be solvent-repellent, while the latter are solvent-attractant (dispersed matter is said to be solvated). In reality there are various degrees of hydrophilicity for which the degree of solvation cannot be determined quantitatively.

Properties. Certain properties of matter are greatly enhanced in the colloidal state due to the large specific surface area. Thus, finely dispersed particles are excellent adsorbents; that is, they can bind various molecules or ions on their surfaces. This property may be used for removal of toxic gases from the atmosphere (in gas masks), for elimination of soluble contaminants in purification of water, or decolorization of sugar, to give just a few examples. *See* ADSORPTION.

Colloids are too small to be seen by the naked eye or in optical microscopes. However, they can be observed and photographed in transmission or scanning electron microscopes. Owing to their small size, they cannot be separated from the medium (liquid or gas) by simple filtration or normal centrifugation. Special membranes with exceedingly small pores, known as ultrafilters, can be used for collection of such finely dispersed particles. The ultracentrifuge, which spins at very high velocities, can also be employed to promote colloid settling. *See* ELECTRON MICROSCOPE; SCANNING ELECTRON MICROSCOPE; ULTRACENTRIFUGE; ULTRAFILTRATION.

Colloids show characteristic optical properties. They strongly scatter light, causing turbidity such

as in fog, milk, or muddy water. Scattering of light (recognized by the Tyndall beam) can be used for the observation of tiny particles in the ultramicroscope. Colloidal state of silica is also responsible for iridescence, the beautiful effect observed with opals. *See* SCATTERING OF ELECTROMAGNETIC RADIATION; TYNDALL EFFECT.

Preparation. Since the characteristic dimensions of colloids fall between those of simple ions or molecules and those of coarse systems, there are in principle two sets of techniques available for their preparation: dispersion and condensation. In dispersion methods the starting materials consist of coarse units which are broken down into finely dispersed particles, drawn into fibers, or flattened into films. For example, colloid mills grind solids to colloid sizes, nebulizers can produce finely dispersed droplets from bulk liquids, and blenders are used to prepare emulsions from two immiscible liquids (such as oil and water). *See* EMULSION.

In condensation methods, ions or molecules are aggregated to give colloidal particles, fibers, or films. Thus, insoluble monolayer films can be developed by spreading onto the surface of water a long-chain fatty acid (for example, stearic acid) from a solution in an organic liquid (such as benzene or ethyl ether). Colloidal aggregates of detergents (micelles) form by dissolving the surface-active material in an aqueous solution in amounts that exceed the critical micelle concentration. *See* MICELLE; MONOMOLECULAR FILM.

The most common procedure to prepare sols is by homogeneous precipitation of electrolytes. Thus, if aqueous silver nitrate and potassium bromide solutions are mixed in proper concentrations, colloidal dispersions of silver bromide will form, which may remain stable for a long time. Major efforts have focused on preparation of monodispersed sols, which consist of colloidal particles that are uniform in size, shape, and composition (see **illus.**). *See* PRECIPITATION (CHEMISTRY).

Dispersion stability. Dispersions of colloids (sols) are inherently unstable because they represent systems of high free energy. Particles may remain in suspended state as long as they are small enough for gravity to be compensated by the kinetic energy (brownian movement). Consequently, particle aggregation on collision must be prevented, which can be achieved by various means. *See* BROWNIAN MOVEMENT.

The most common cause for sol stability is electrostatic repulsion, which results from the charge on colloidal particles. The latter can be due to adsorption of excess ions (such as of bromide ions on silver bromide particles), surface acid or base reactions (for example, removal or addition of protons on the surface of metal hydroxides), or other interfacial chemical reactions. Alternately, sol stability can be induced by adsorption of polymers or by solvation. Stability of colloid dispersions that occur naturally may represent an undesirable state, as is the case with muddy waters whose turbidity is due to the presence of tiny particles (clays, iron oxides, and so forth). When needed, stability can be artificially induced, as in the production of paints. *See* SOLVATION.

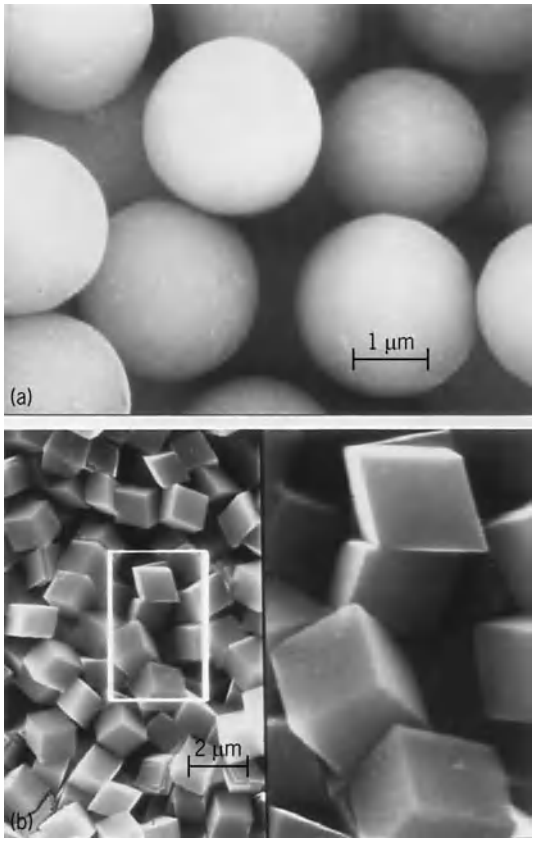

Examples of colloid particles. (*a*) Zinc sulfide (sphalerite). (*b*) Cadmium carbonate, with an enlargement of the boxed section (*courtesy of Egon Matijević*).

Coagulation or flocculation. Processes by which a large number of finely dispersed particles is aggregated into larger units, known as coagula or flocs, are known as coagulation or flocculation; the distinction between the two is rather subtle. The concepts most commonly refer to colloidal solids in liquids (sols), although analogous principles can be employed in explaining aggregation of solid particles or liquid droplets in gases (aerosols).

Coagulation of stable sols can be accomplished by various additives (electrolytes, surfactants, polyelectrolytes, or other colloids) whose efficiency depends on their charge and adsorptivity, as well as on the nature of suspended particles to be treated. A common example of applications of coagulation is in water purification in which particulate colloid contaminants must be aggregated in order to be removed. Coagulation is also used to describe solidification of polymers (for example, proteins), which can be achieved by heating, as exemplified by boiling an egg. *See* EMULSION; FOAM; FOG; POLYMER; SMOKE. Egon Matijević

Bibliography. K. S. Birdi (ed.), *Handbook of Surface and Colloid Chemistry*, 2d ed., 2000; P. C. Hiemenz and R. Rajagopalam, *Principles of Colloid and Surface Chemistry*, 3d ed., 1997; D. Shaw, *Introduction to Colloid and Surface Chemistry*, 4th ed., 1992; E. D. Shchukin et al., *Colloid and Surface Chemistry*, 2001.

Colloidal crystals

Periodic arrays of suspended colloidal particles. Common colloidal suspensions (colloids) such as milk, blood, or latex are polydisperse; that is, the suspended particles have a distribution of sizes and shapes. However, suspensions of particles of identical size, shape, and interaction, the so-called monodisperse colloids, do occur. In such suspensions, a new phenomenon that is not found in polydisperse systems, colloidal crystallization, appears: under appropriate conditions, the particles can spontaneously arrange themselves into spatially periodic structures. This ordering is analogous to that of identical atoms or molecules into periodic arrays to form atomic or molecular crystals. However, colloidal crystals are distinguished from molecular crystals, such as those formed by very large protein molecules, in that the individual particles do not have precisely identical internal atomic or molecular arrangements. On the other hand, they are distinguished from periodic stackings of macroscopic objects like cannonballs in that the periodic ordering is spontaneously adopted by the system through the thermal agitation (brownian motion) of the particles. These conditions limit the sizes of particles which can form colloidal crystals to the range from about 0.01 to about 5 micrometers. *See* BROWNIAN MOVEMENT; KINETIC THEORY OF MATTER.

The most spectacular evidence for colloidal crystallization is the existence of naturally occurring opals. The ideal opal structure is a periodic close-packed three-dimensional array of silica microspheres with hydrated silica filling the spaces not occupied by particles. Opals are the fossilized remains of an earlier colloidal crystal suspension. Synthetic opals can be made by allowing suspensions of silica spheres to sediment and then letting the suspending medium evaporate. Colloidal crystals can also be made from suspensions of synthetic plastic (organic polymer) microspheres, formed by emulsion polymerization. Such suspensions have become important systems for the study of colloidal crystals, by virtue of the controllability of the particle size and interaction. Colloidal crystals have also been found in aqueous emulsions of oil drops in water, in which all the drops are made the same size by fractionation. Another important class of naturally occurring colloidal crystals are found in concentrated suspensions of nearly spherical virus particles, such as *Tipula* iridescent virus and tomato bushy stunt virus. *See* OPAL; VIRUS.

Crystal structure and properties. In order for colloidal crystals to form, the forces between the particles must be repulsive, since if the forces are attractive the particles will aggregate into noncrystalline clusters. The repulsion can be electrostatic, if the particles are charged (charge-stabilized), or can come from interparticle collisions (steric-stabilized). In the latter case the particle surface must be treated with an amphiphilic or polymer layer to prevent the particles from sticking together.

Similarities to atomic crystals. In colloidal crystals of spherical or nearly spherical particles, the crystal structures adopted are identical to those found in crystals formed by spherical atoms, such as the noble gases and alkali metals. The phase behavior is also quite similar.

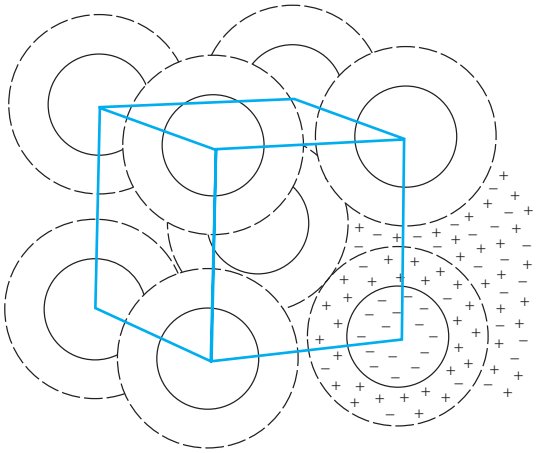

Fig. 1. Arrangement of particles and distribution of electrical charges in a charged stabilized body-centered-cubic colloidal crystal. Spheres indicated by broken lines show extent of positive counterion clouds.

For example, in suspensions of polymer spheres in organic solvents the interaction between the spheres is short-ranged, consisting only of repulsion on contact (the hard sphere limit). As with the noble gas crystals where the interatomic repulsion is short-ranged, the crystals formed are close-packed, either hexagonal or face-centered cubic, first appearing when the spheres' volume fraction is about 55%.

In aqueous suspension the particles can interact electrically, without ever actually touching. Polymer particles are readily prepared with acid surface groups which dissociate in water, leaving the particle with a large negative charge (as large as several thousand electron charges). This charge attracts positively charged small ions, which form clouds around the particles (**Fig. 1**). The spheres will repel each other if they come close enough so that their ion clouds overlap. At low concentration, the particles rarely encounter one another, executing independent brownian motion and forming the colloidal gas phase. As the concentration is increased to where the average interparticle separation is comparable to the overlap separation, the particles become strongly confined by their neighbors and form the colloidal liquid phase. If the concentration is further increased, an abrupt freezing transition to the ordered colloidal crystal phase will occur. The colloidal crystal structure found near the freezing transition is the body-centered cubic, as is found in the alkali metals, where the atomic repulsion is more long-ranged. *See* GAS; LIQUID.

Differences from atomic crystals. Despite these structural similarities, atomic and colloidal crystals are characterized by vastly different length, stiffness, and time scales.

First, particles in atomic crystals are separated by less than a nanometer compared to micrometers in colloidal crystals. This means that colloidal crystal structure can be determined by visible-light

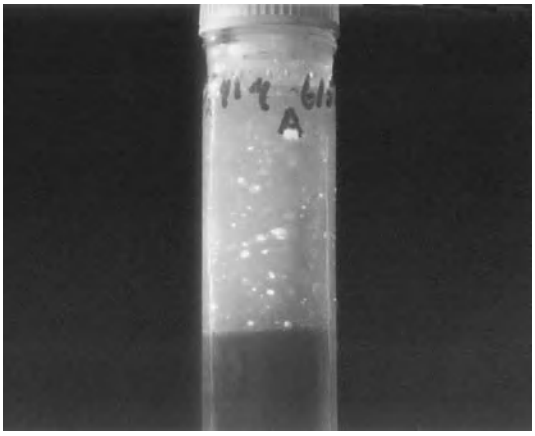

Fig. 2. Reflection of white light from a polycrystalline colloidal crystal suspension. The crystals are body-centered cubic, with a particle diameter of 0.1 μm and interparticle spacing of 1.5 μm.

techniques, as opposed to x-rays, which are required in the atomic crystal case. The spectacular iridescent appearance of opal and other colloidal crystals (**Fig. 2**) is a direct result of constructive interference in the reflection of light by the particles' periodic lattice structure; that is, the colloidal crystals Bragg-reflect visible light in the same way as atomic crystals Bragg-reflect x-rays. *See* OPALESCENCE; X-RAY CRYSTALLOGRAPHY; X-RAY DIFFRACTION.

This optical property is being developed for photonic bandgap applications, in which light of frequency in a particular range is reflected for all possible propagation directions. The search for useful photonic crystals has led to the development of inverse colloidal crystals and opals, in which a hard-sphere colloidal crystal is made (of silica particles, for example), the void space (74% of the volume for a close-packed structure) is filled with another material (a polymer, for example), and then the original particles are chemically removed (by dipping in hydrofluoric acid in the case of silica). The result is a crystal structure in which the solid material occupies only about 25% of the volume (**Fig. 3**).

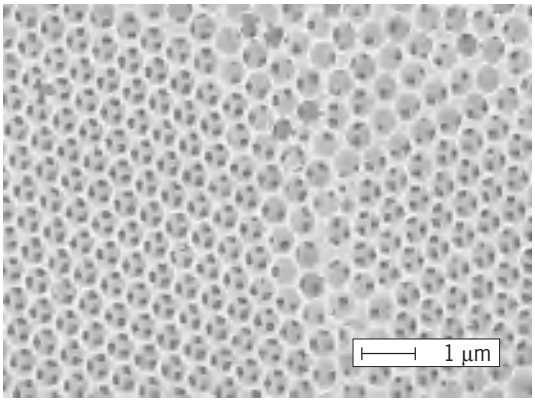

Fig. 3. Inverse opal structure formed by filling the void space of a face-centered cubic colloidal crystal emulsion of oil drops in water and then condensing silica in the aqueous void spaces. (*D.J. Pine, University of California at Santa Barbara*)

Second, colloidal crystals are about 10^8 times weaker than atomic crystals; that is, their elastic constants are 10^8 times smaller. Consequently, colloidal crystals can be melted very easily, even by a gentle shearing of the suspension.

Finally, significant atomic motions, such as the relative vibration of adjacent particles, occur in 10^{-12} s, whereas the equivalent time in a colloidal system is 10^{-2} s. Hence, it is possible to watch in detail through a microscope the motions of particles in a colloidal lattice. The distinctive properties of colloidal crystals thus provide a new avenue to the understanding of many phenomena found in atomic systems, such as lattice vibrations (sound waves), crystal dislocation and other defects, the response of crystals to large stresses (plastic flow), and crystallization itself. *See* CRYSTAL DEFECTS; CRYSTAL GROWTH; LATTICE VIBRATIONS; PLASTICITY.

Colloidal crystallization has been studied in space, taking advantage of the microgravity environment to avoid the sedimentation arising from the density mismatch between particles and suspending medium. Crystals form in space under conditions where the suspensions remain liquid on Earth.

Complex systems. In addition to the simple colloidal crystals found in bulk (three-dimensional) suspensions of spherical particles, there are a number of more exotic colloidal systems. Colloidal alloys can be formed in suspensions containing particles of several different sizes. Binary mixtures form mixed crystals, sometimes with approximately 100 particles per unit cell, analogous to the alloy structures found in binary atomic metal mixtures. Other systems include colloidal glasses, which are binary or multicomponent suspensions that exhibit shear rigidity without periodic structure; two-dimensional colloidal crystals, which are obtained by trapping a monolayer or several layers of particles; and colloidal liquid crystals, which are found in suspensions of interacting non-spherical (rod- or disk-shaped) particles such as tobacco mosaic virus. *See* ALLOY STRUCTURES; AMORPHOUS SOLID; COLLOID; CRYSTAL STRUCTURE; LIQUID CRYSTALS.

Noel A. Clark

Bibliography. N. A. Clark, A. J. Hurd, and B. J. Ackerson, Single colloidal crystals, *Nature*, 281:57–60, 1979; D. J. Darragh, A. J. Gaskin, and J. V. Sanders, Opals, *Sci. Amer.*, 234(4):84–95, April 1976; R. M. Fitch (ed.), *Polymer Colloids*, vols. 1 and 2, 1978; A. Imhoff and D. J. Pine, Uniform macroporous ceramics and plastics by emulsion tampering, *Adv. Mater.*, 10:697–700, 1998; A. Zakhidov et al., Carbon structures with three-dimensional periodicity at optical wavelengths, *Science*, 282:897–901, 1999; J. Zhu et al., Crystallization of hard sphere colloids in microgravity, *Nature*, 387:883–885, 1997.

Colon

The portion of the intestine that runs from the cecum to the rectum; in some mammals, it may be separated from the small intestine by an ileocecal valve. It is also known as the large intestine. The colon is

usually divided into ascending, transverse, and descending portions. In the humans a fourth section, the sigmoid, is found. The colon is longer in herbivores and shorter in carnivores, and is about 4 to 6 ft (1 to 2 m) long in humans. Although no digestive enzymes are secreted in the colon, digestion is assisted by an alkaline fluid. Much digestion (for example, all breakdown of cellulose) occurs by bacteria, of which *Escherichia coli is* the most common. Most of the fluid added to the food during digestion is reabsorbed into the body in the colon. All digestive action, water absorption, and so on, is completed before the food materials pass out of the colon into the rectum. *See* DIGESTIVE SYSTEM. Walter J. Bock

Color

That aspect of visual sensation enabling a human observer to distinguish differences between two structure-free fields of light having the same size, shape, and duration. Although luminance differences alone permit such discriminations to be made, the term color is usually restricted to a class of differences still perceived at equal luminance. These depend upon physical differences in the spectral compositions of the two fields, usually revealed to the observer as differences of hue or saturation.

Photoreceptors. Color discriminations are possible because the human eye contains three classes of cone photoreceptors that differ in the photopigments they contain and in their neural connections. Two of these, the L and M cones, are sensitive to all wavelengths of the visible spectrum from 380 to 700 nanometers. (Even longer or shorter wavelengths may be effective if sufficient energy is available.) L cones are maximally sensitive at about 570 nm, M cones at about 540 nm. The ratio L/M of cone sensitivities is minimal at 465 nm and increases monotonically for wavelengths both shorter and longer than this. This ratio is independent of intensity, and the red-green dimension of color variation is encoded in terms of it. The S cones, whose sensitivity peaks at about 440 nm, are not appreciably excited by wavelengths longer than 540 nm. The perception of blueness and yellowness depends upon the level of excitation of S cones in relation to that of L and M cones. No two wavelengths of light can produce equal excitations in all three kinds of cones. It follows that, provided they are sufficiently different to be discriminable, no two wavelengths can give rise to identical sensations.

The foregoing is not true for the comparison of two different complex spectral distributions. These usually, but not always, look different. Suitable amounts of short-, middle-, and long-wavelength lights, if additively mixed, can for example excite the L, M, and S cones exactly as does a light containing equal energy at all wavelengths. As a result, both stimuli look the same. This is an extreme example of the subjective identity of physically different stimuli known as chromatic metamerism. Additive mixture is achievable by optical superposition, rapid alternation at frequencies too high for the visual system to follow, or (as in color television) by the juxtaposition of very small elements which make up a field structure so fine as to exceed the limits of visual acuity. The integration of light takes place within each receptor, where photons are individually absorbed by single photopigment molecules, leading to receptor potentials that carry no information about the wavelength of the absorbed photons. *See* EYE (VERTEBRATE); LIGHT.

Colorimetry. Although colors are often defined by appeal to standard samples, the trivariant nature of color vision permits their specification in terms of three values. Ideally these might be the relative excitations of the L, M, and B cones. Because too little was known about cone action spectra in 1931, the International Commission on Illumination (CIE) adopted at that time a different but related system for the prediction of metamers (the CIE system of colorimetry). This widely used system permits the specification of tristimulus values X, Y, and Z, which make almost the same predictions about color matches as do calculations based upon cone action spectra. If, for fields 1 and 2, $X_1 = X_2$, $Y_1 = Y_2$, and $Z_1 = Z_2$, then the two stimuli are said to match (and therefore have the same color) whether they are physically the same (isometric) or different (metameric).

The use of the CIE system may be illustrated by a sample problem. Suppose it is necessary to describe quantitatively the color of a certain paint when viewed under illumination by a tungsten lamp of known color temperature. The first step is to measure the reflectance of the paint continuously across the visible spectrum with a spectrophotometer. The reflectance at a given wavelength is symbolized as R_λ. The next step is to multiply R_λ by the relative amount of light E_λ emitted by the lamp at the same wavelength. The product $E_\lambda R_\lambda$ describes the amount of light reflected from the paint at wavelength λ. Next, $E_\lambda R_\lambda$ is multiplied by a value $\bar{x}_\lambda$, which is taken from a table of X tristimulus values for an equal-energy spectrum. The integral $\int E_\lambda R_\lambda \bar{x}_\lambda d_\lambda$ gives the tristimulus value X for all of the light reflected from the paint. Similar computations using $\bar{y}_\lambda$ and $\bar{z}_\lambda$ yield tristimulus values Y and Z.

Tables of $\bar{x}_\lambda$, $\bar{y}_\lambda$, and $\bar{z}_\lambda$ are by convention carried to more decimal places than are warranted by the precision of the color matching data upon which they are based. As a result, colorimetric calculations of the type just described will almost never yield identical values, even for two physically different fields that are identical in appearance. For this and other reasons it is necessary to specify tolerances for color differences. Such differential colorimetry is primarily based upon experiments in which observers attempted color matches repeatedly, with the standard deviations of many matches being taken as the discrimination unit.

Chromaticity diagram. Colors are often specified in a two-dimensional chart known as the CIE chromaticity diagram, which shows the relations among tristimulus values independently of luminance. In this plane, y is by convention plotted as a function of

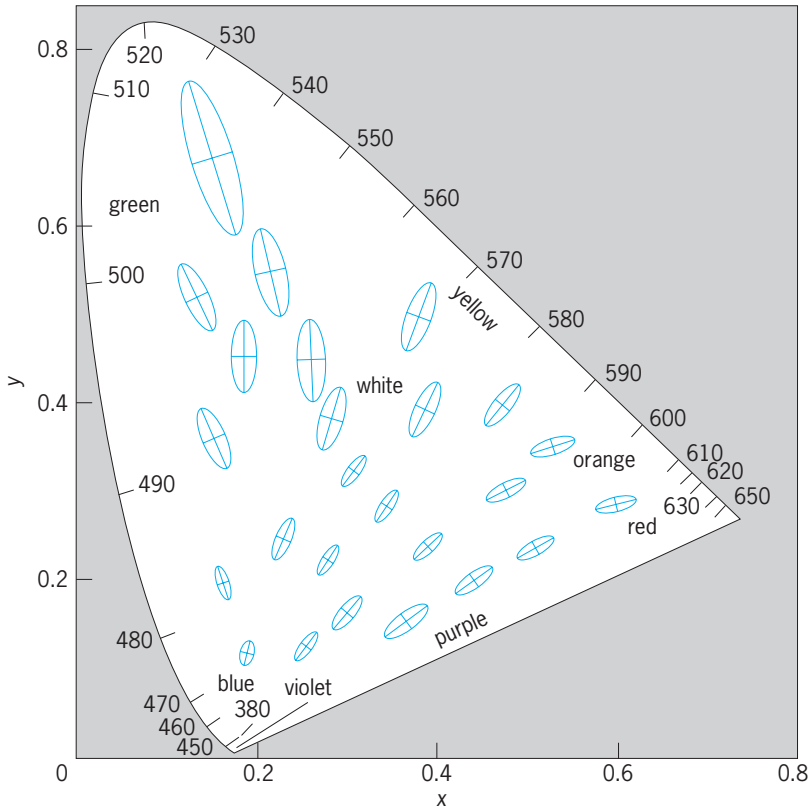

The 1931 CIE chromaticity diagram showing MacAdam's ellipses 10 times enlarged. (*After G. W. Wyszecki and W. S. Stiles, Color Science, Wiley, 1967*)

x, where $y = Y/(X + Y + Z)$ and $x = x/(x + Y + Z)$. [The value $z = Z/(X + Y + Z)$ also equals $1 - (x + y)$ and therefore carries no additional information.] Such a diagram is shown in the **illustration**, in which the continuous locus of spectrum colors is represented by the outermost contour. All nonspectral colors are contained within an area defined by this boundary and a straight line running from red to violet. The diagram also shows discrimination data for 25 regions, which plot as ellipses represented at 10 times their actual size. A discrimination unit is one-tenth the distance from the ellipse's center to its perimeter. Predictive schemes for interpolation to other regions of the CIE diagram have been worked out.

If discrimination ellipses were all circles of equal size, then a discrimination unit would be represented by the same distance in any direction anywhere in the chart. Because this is dramatically untrue, other chromaticity diagrams have been developed as linear projections of the CIE chart. These represent discrimination units in a relatively more uniform way, but never perfectly so.

A chromaticity diagram has some very convenient properties. Chief among them is the fact that additive mixtures of colors plot along straight lines connecting the chromaticities of the colors being mixed. Although it is sometimes convenient to visualize colors in terms of the chromaticity chart, it is important to realize that this is not a psychological color diagram. Rather, the chromaticity diagram makes a statement about the results of metameric color matches, in the sense that a given point on the diagram represents the locus of all possible metamers plotting at chromaticity coordinates x, y. However, this does not specify the appearance of the color, which can be dramatically altered by preexposing the eye to colored lights (chromatic adaptation) or, in the complex scenes of real life, by other colors present in nearby or remote areas (color contrast and assimilation). Nevertheless, within limits, metamers whose color appearance is thereby changed continue to match.

For simple, directly fixated, and unstructured fields presented in an otherwise dark environment, there are consistent relations between the chromaticity coordinates of a color and the color sensations that are elicited. Therefore, regions of the chromaticity diagram are often denoted by color names, as shown in the illustration.

Although the CIE system works rather well in practice, there are important limitations. Normal human observers do not agree exactly about their color matches, chiefly because of the differential absorption of light by inert pigments in front of the photoreceptors. Much larger individual differences exist for differential colorimetry, and the system is overall inappropriate for the 8% of the population (mostly males) whose color vision is abnormal. The system works only for an intermediate range of luminances, below which rods (the receptors of night vision) intrude, and above which the bleaching of visual photopigments significantly alters the absorption spectra of the cones. *See* COLOR VISION.

Robert M. Boynton

Bibliography. F. W. Billmeyer and M. Saltzman, *Principles of Color Technology,* 2d ed., 1981; R. W. Burnham, R. M. Hanes, and C. J. Bartleson, *Color: A Guide to Basic Facts and Concepts,* 1963; D. B. Judd and G. W. Wyszecki, *Color in Business, Science, and Industry,* 3d ed., 1975; P. K. Kaiser and R. M. Boynton, *Human Color Vision,* 2d ed., 1996; K. Nassau, *The Physics and Chemistry of Color,* 1983; G. W. Wyszecki and W. S. Stiles, *Color Science,* 2d ed., 1982.

Color (quantum mechanics)

A term used to describe a quantum number carried by the quarks which make up the strongly interacting elementary particles, such as the proton and neutron of the atomic nucleus. It has nothing to do with the ordinary, visual use of the word color.

The quarks which are thought to make up the strongly interacting particles have a spin angular momentum of one-half unit of $\hbar$ (Planck's constant). According to a fundamental theorem of relativity combined with quantum mechanics, they must therefore obey Fermi-Dirac statistics and be subject to the Pauli exclusion principle. No two quarks within a particular system can have exactly the same quantum numbers. *See* EXCLUSION PRINCIPLE; FERMI-DIRAC STATISTICS.

In making up a baryon, however, it often seemed necessary to violate this principle. The omega particle, for example, is made of three strange quarks, and all three had to be in exactly the same state. O. W. Greenberg was responsible for the essential idea for the solution to this paradox. In 1964 he suggested that each quark type (*u, d,* and *s*) comes in three varieties identical in all measurable qualities but different in an additional property, which has come to be known as color. The exclusion principle could then be satisfied and quarks could remain fermions, because the quarks in the baryon would not all have the same quantum numbers. They would differ in color even if they were the same in all other respects.

The color hypothesis triples the number of quarks but does not increase the number of baryons and mesons. The rules for assembling them ensures this. Tripling the number of quarks does, however, have at least two experimental consequences. It triples the rate at which the neutral π meson decays into two photons and brings the predicted rate into agreement with the observed rate.

The total production cross section for baryons and mesons in electron-positron annihilation is also tripled. The experimental results at energies up to 60 GeV are in good agreement with the color hypothesis and completely incompatible with the simple quark model without color. *See* BARYON; ELEMENTARY PARTICLE; GLUONS; MESON; QUARKS.

Thomas Appelquist

Bibliography. H. Georgi, A unified theory of elementary particles and forces, *Sci. Amer.*, 244(4):48–63, 1981; S. L. Glashow, Quarks with color and flavor, *Sci. Amer.*, 233(4):38–50, 1975; I. S. Hughes, *Elementary Particles*, 3d ed., 1991; G. Kane, *Modern Elementary Particle Physics*, 2d ed., 1993.

Color centers

Atomic and electronic defects of various types in solids which produce optical absorption bands in otherwise transparent crystals such as the alkali halides, alkaline earth fluorides, or metal oxides. They are general phenomena found in a wide range of materials. Color centers are produced by gamma radiation or x-radiation, by addition of impurities or excess constituents, and sometimes through electrolysis. A well-known example is that of the *F*-center in alkali halides such as sodium chloride (NaCl). The designation *F*-center comes from the German word *Farbe*, which means color. *F*-centers in sodium chloride produce a band of optical absorption toward the blue end of the visible spectrum; thus the colored crystal appears yellow under transmitted light. On the other hand, potassium chloride (KCl) with *F*-centers appears magenta, and potassium bromide (KBr) appears blue. *See* CRYSTAL DEFECTS.

Theoretical studies guided by detailed experimental work have yielded a deep understanding of specific color centers. The crystals in which color centers appear tend to be transparent to light and to other forms of electromagnetic radiation, such as microwaves. Consequently, experiments which can be carried out include optical spectroscopy, luminescence and Raman scattering, magnetic circular dichroism, magnetic resonance, and electromodulation. Color centers find practical application in radiation dosimeters, x-ray image storage phosphors, and tunable lasers. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION.

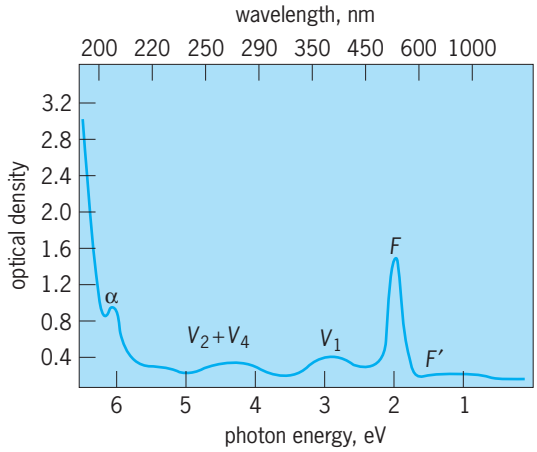

Fig. 1. Absorption bands produced in KBr crystal by exposure to x-rays at 81 K. Optical density is equal to $\log_{10}(I_0/I)$, where I_0 is the intensity of incident light and I the intensity of transmitted light. Steep rise in optical density at far left is due to intrinsic absorption in crystal (modified slightly by existence of *F*-centers).

Origin. **Figure 1** shows the absorption bands due to color centers produced in potassium bromide by exposure of the crystal at the temperature of liquid nitrogen (81 K) to intense penetrating x-rays. Several prominent bands appear as a result of the irradiation. The *F*-band appears at 600 nanometers and the so-called *V*-bands appear in the ultraviolet. Uncolored alkali halide crystals may be grown readily from a melt with as few imperfections as one part in 10^5. In the unirradiated and therefore uncolored state, they show no appreciable absorption from the far-infrared through the visible, to the region of characteristic absorption in the far-ultraviolet (beginning at about 200 nm in Fig. 1).

Color bands such as the *F*-band and the *V*-band arise because of light absorption at defects dispersed throughout the lattice. This absorption is caused by electronic transitions at the centers. On the other hand, colloidal particles, each consisting of many atoms, dispersed through an optical medium also produce color bands. In this case, if the particles are large enough, the extinction of light is due to both light scattering and light absorption. Colloidal gold is responsible for the color of some types of ruby glass. Colloids may also form in alkali halide crystals—for example, during heat treatment of an additively colored crystal which contains an excess of alkali metal.

Atomically dispersed centers such as *F*-centers are part of the general phenomena of trapped electrons and holes in solids. The accepted model of the *F*-center is an electron trapped at a negative ion vacancy. Many other combinations of electrons, holes,

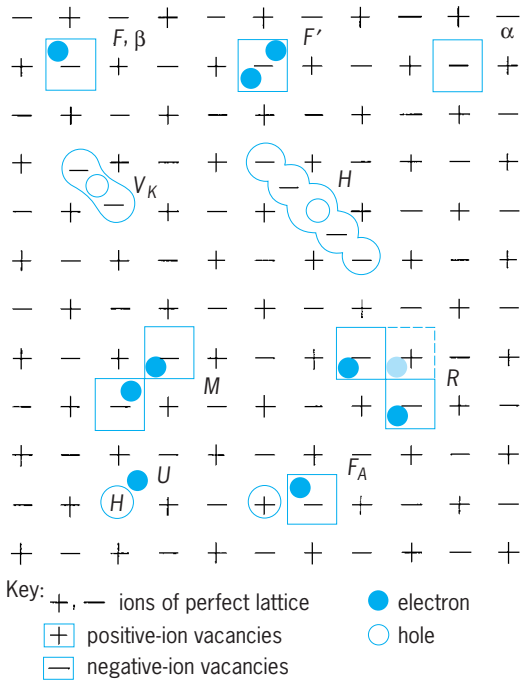

Fig. 2. Well-established models for several color centers in ionic crystal of NaCl structure. Designation β for *F*-center signifies disturbing influence of *F*-centers on tail of fundamental absorption band; characteristic absorption is also influenced by presence of vacancies, or α-centers (see Fig. 1). The V_K-center can be described as a self-trapped hole forming a Cl_2-molecule-ion, and the *H*-center can be described as a hole trapped at an interstitial halide ion. Both are stable at low temperature only. The *R*-center contains three negative-ion vacancies, one of which is in an adjacent lattice plane. *H* in circle at lower left represents hydrogen impurity, which, together with an electron, forms a *U*-center. The F_a-center is an *F*-center adjacent to a monovalent positive-ion impurity.

and clusters of lattice vacancies have been used to explain the various absorption bands observed in ionic crystals. The centers and models shown in **Fig. 2** have been positively identified either by electron spin resonance or by a combination of other types of experimental evidence. *See* HOLE STATES IN SOLIDS.

Impurities can play an important role in color-center phenomena. Certain impurities in ionic crystals produce color bands characteristic of the foreign ion. For example, hydrogen can be incorporated into the alkali halides as H^- ions substituting for halogen ions with resultant appearance of an absorption band (the *U*-band) in the ultraviolet (Fig. 2). The rate at which *F*-centers are produced by x-irradiation is greatly increased by the incorporation of hydrogen, the *U*-centers being converted into *F*-centers with high efficiency.

F-centers. *F*-centers may be produced in uncolored crystals by irradiation with ultraviolet, x-rays, gamma rays, and high-speed particles, and also by electrolysis. However, one of the most convenient methods is that of additive coloration. Alkali halide crystals may be additively colored by heating to several hundred degrees Celsius in the presence of the alkali metal vapor, then cooling rapidly to room temperature. The *F*-band that results is the same as that produced by irradiation and is dependent upon the particular alkali halide used, not upon the alkali metal vapor. For example, the coloration produced in potassium chloride is the same whether the crystal is heated in potassium or in sodium vapor. This and other evidence indicate that excess alkali atoms enter the crystal as normal lattice ions by donating an electron to accompanying negative-ion vacancies to form *F*-centers. **Figure 3** shows the position of the *F*-band in several alkali halides.

Both the width and the exact position of the *F*-band change with temperature in the manner shown in **Fig. 4**. This behavior can be explained in terms of the thermal motion of the ions surrounding the center. Thermal motion is most important at high temperatures, whereas at low temperatures the width of the absorption band becomes less and approaches a constant value which prevails down to the very lowest temperatures.

Concentration of centers. The height of the absorption maximum and the width at half maximum may be used to calculate the concentration N of absorbing centers. For potassium chloride, classical theory gives the concentration as $N = 1.3 \times 10^{16}\, AH\ \mathrm{cm}^{-3}$,

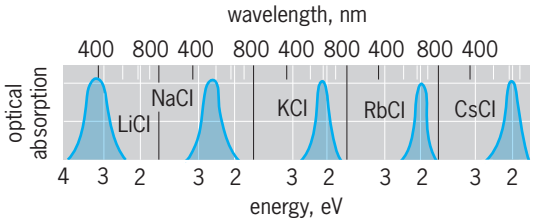

Fig. 3. *F*-bands in different alkali halide crystals.

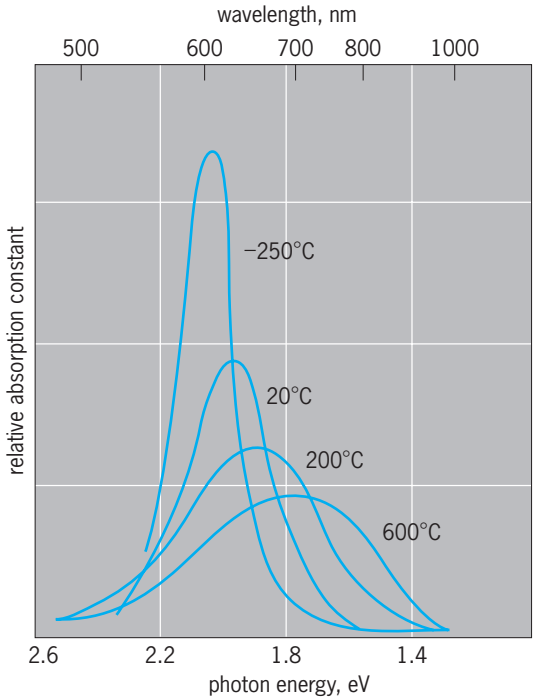

Fig. 4. Variation with temperature of the width of the *F*-band in a crystal of potassium bromide. °F = (°C × 1.8) + 32°.

where A is the maximum absorption in cm^{-1}, and H is the half width of the absorption band in electron-volts. The numerical factor in this equation depends upon the material under consideration and is given here for potassium chloride. The general relation between absorption characteristics and defect concentration is slightly more complicated but is nevertheless useful to determine concentrations below the limit of detection by chemical means.

Shape of F-band. Refined measurements of the *F*-band at low temperatures show that it is not a simple bell-shaped curve indicated in Figs. 3 and 4. In fact, it is found to contain a shoulder on the short-wavelength side, which is referred to as the *K*-band. It has been proposed that the *K*-band is in agreement with a model for the *F*-center in which an electron is trapped in a vacancy surrounded by point ions which produce the cubic crystal field of the lattice. An electron in the vicinity of a negative-ion vacancy should have a ground state and one or more excited states below the bottom of the conduction band. The main absorption peak corresponds to excitation of the electron from the 1*s*-like ground state to an excited 2*p*-like state. The *K*-band is thought to be due to transitions to the 3*p* and higher states which finally merge with the conduction band itself. *F*-center photoconductivity measurements tend to confirm the idea of discrete excited states, such as the 2*p*, which lie below the conduction band or continuum. *See* PHOTOCONDUCTIVITY.

Magnetic experiments. Some of the most detailed information on *F*-centers comes from electron spin resonance and from electron-nuclear double resonance. From these experiments the extent and character of the wave function of the electron trapped at the halogen-ion vacancy can be obtained. The electron even in the ground state is not highly localized but is spread out over the six neighboring alkali ions and to some extent over more distant neighbors. Faraday rotation studies indicated fine structure in the excited state of the *F*-center due to spin-orbit interaction. *See* MAGNETIC RESONANCE.

Lifetime of excited F-center. At low temperature an excited *F*-center decays spontaneously to the ground state with the emission of an infrared luminescent quantum. The radiative lifetime has been measured and found to be unexpectedly long, 0.57 microsecond in potassium chloride and as long as 2.0 μs in potassium iodide (KI). There are indications that this is due to a drastic change in the wave function following excitation and lattice relaxation. Lattice polarization must be taken into account in order to calculate the effect. The relaxed excited state is very spread out in the lattice, whereas the ground state is somewhat more localized.

F′-centers. When a crystal containing *F*-centers is irradiated at low temperature with light in the *F*-band itself, the *F*-absorption decreases and a new band, known as the *F′* band, appears (**Fig. 5**). Experiments have shown that for every *F′*-center created, two *F*-centers are destroyed. The *F′*-center consists of an *F*-center which has captured an extra electron. Important photoelectric effects occur during conversion of *F*- to *F′*-centers and during irradiation of alkali halides with *F*-band light.

V-centers. As Fig. 1 shows, absorption bands arise in the ultraviolet as well as in the visible portion of the spectrum when potassium bromide is irradiated while cold. These *V*-bands are apparently not present in crystals additively colored by the alkali metal vapor. Irradiation of a crystal produces both electrons and holes so that bands associated with trapped positive charges are to be expected in addition to electrons trapped at negative-ion vacancies (*F*-centers). *V*-bands can also be produced, at least in the case of potassium bromide and potassium iodide, by introduction of an excess of halogen into the crystal.

Crystals containing *V*-centers have been studied by spin resonance techniques with the result that a center not previously observed optically has been identified. This center, sometimes referred to as the V_K-center (Fig. 2), is due to a hole trapped in the vicinity of two halogen ions which have become displaced from their normal lattice position to form a Cl_2-molecule-ion (in the case of potassium chloride). It has been found in several different alkali halides, and its optical absorption and energy level structure have been correlated. The V_K-center is known to play an important role in the intrinsic luminescence of ionic crystals.

Mechanism of radiation damage. When an alkali halide crystal is exposed to ionizing radiation, such as ultraviolet light, x-rays, or beta particles, radiation damage occurs in the form of point defects. *F*-centers and *H*-centers are now known to be the primary defects formed. They are produced mainly by recombination of electron-hole pairs associated with ionizing radiation. Luminescence may also occur, depending upon temperature and upon the presence of impurities. Practical applications of these phenomena include dosimeters for detecting ionizing radiation and image plate detectors for x-rays. Analogous

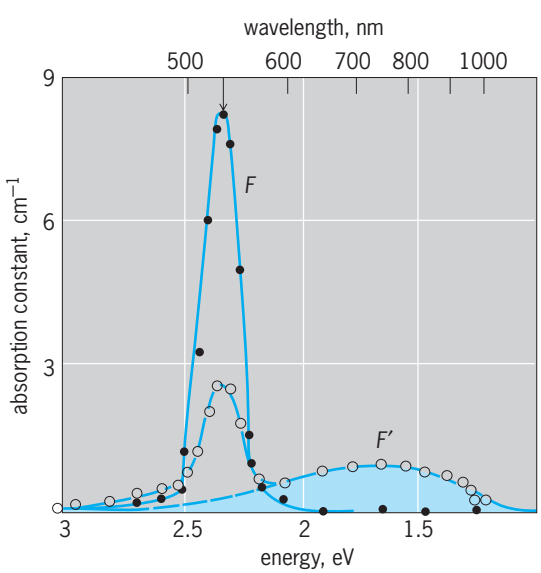

Fig. 5. ***F*- and *F′*-bands in KCl at −235°C (−391°F). *F′*-band was produced by irradiating crystal with light of wavelength shown by arrow at center of *F*-band. Such light causes the *F*-band to bleach as *F′*-band grows.**

phenomena occur in the case of crystals of the fluorite structure (CaF_2) and also in oxides such as calcium oxide (CaO) and magnesium oxide (MgO). In these cases the photolytic processes are somewhat different and apparently not as efficient as in the alkali halides. *See* DOSIMETER; ELECTRON-HOLE RECOMBINATION; PHOTOLYSIS; RADIATION DAMAGE TO MATERIALS.

Frederick C. Brown

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; F. C. Brown, *The Physics of Solids*, 1967; J. Crawford and L. Slifkin (eds.), *Point Defects in Solids*, vol. 1, 1972; W. B. Fowler, *The Physics of Color Centers*, 1968; W. Gellermann, Color center lasers, *J. Phys. Chem. Solids*, 52:249, 1991; N. Itoh and K. Tanimura, Recombination-induced defect formation, *J. Phys. Chem. Solids*, 51:717, 1990.

Color filter

An optical element that partially absorbs incident radiation, often called an absorption filter. The absorption is selective with respect to wavelength, or color, limiting the colors that are transmitted by limiting those that are absorbed. Color filters absorb all the colors not transmitted. They are used in photography, optical instruments, and illuminating devices to control the amount and spectral composition of the light.

Color filters are made of glass for maximum permanence, of liquid solutions in cells with transparent faces for flexibility of control, and of dyed gelatin or plastic (usually cellulose acetate) for economy, convenience, and flexibility. The plastic filters are often of satisfactory permanence, but they are sometimes cemented between glass plates for greater toughness and scratch resistance. They do not have as good quality as gelatin filters.

Color filters are sometimes classified according to their type of spectral absorption: short-wavelength pass, long-wavelength pass or band-pass; diffuse or sharp-cutting; monochromatic or conversion. The short-wavelength pass transmits all wavelengths up to the specified one and then absorbs. The long-wavelength pass is the opposite. Every filter is a band-pass filter when considered generally. Even an ordinary piece of glass does not transmit in the ultraviolet or infrared parts of the spectrum. Color filters, however, are usually discussed in terms of the portion of the visible part of the spectrum. Sharp and diffuse denote the sharpness of the edges of the filter band pass. Monochromatic filters are very narrow band-pass filters. Conversion filters alter the spectral response or distribution of one selective detector or source to that of another, for example, from that of a light bulb to that of the Sun (see **illus.**).

Neutral-density filters transmit a constant fraction of the light across the spectrum considered. They are made either by including an appropriate absorbent like carbon that is spectrally nonselective, or by a thin, partially transparent reflective layer, often of aluminum.

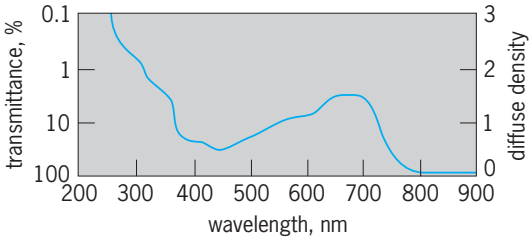

Transmission of a conversion filter, used to convert the spectral distribution of a light bulb (color temperature of 2360 K or 3790°F) to that of the Sun (color temperature of 5500 K or 9440°F). (*After Eastman Kodak Co., Kodak Filters for Scientific and Technical Uses, 1972*)

The transmittance of a filter is the ratio of the transmitted flux to the incident flux, expressed as either a ratio or a percentage. The density of a filter is the logarithm to base 10 of the reciprocal of the transmittance. For example, a filter with a 1% transmittance has a density of 2.

Other, closely related filters are the Christiansen, Lyot or birefringent, residual-ray, and interference filters. They use the phenomena of refraction, polarization, reflection, and interference. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; COLOR; INTERFERENCE FILTER; REFLECTION OF ELECTROMAGNETIC RADIATION; SUN.

William L. Wolfe

Color index

A quantitative measure of a star's color. Even a casual look at the night sky reveals the stars to be colored. From spectral analysis and the laws of radiation, it has been known since the 1920s that such color is closely related to temperature. Stars are reddish at about 3000 K (4940°F), orange at about 4500 K (7640°F), yellowish-white at about 6000 K (10,340°F), white at about 10,000 K (17,540°F), and bluish-white above 10,000 K. With proper calibration, color provides a means by which the temperature and spectral class can be evaluated. Visual examination is too crude, however. Instead, color must be properly defined and quantified. *See* SPECTRAL TYPE; STAR.

Color is defined through a star's various magnitude measures (m), which indicate brightness at different wavelengths. The unaided-eye stars are historically grouped into six magnitude classes, 1st magnitude being the brightest and 6th magnitude the faintest. The system was quantified in the nineteenth century so that five magnitude divisions were made to correspond to a factor of 100 in brightness. One magnitude division is then a factor of 2.512 . . . . The star Vega, with $m = 0.03$, now provides a fundamental calibration point. Telescopes extend the range to magnitude 30. *See* MAGNITUDE (ASTRONOMY).

The original magnitudes were visual, m being estimated by eye, and depended on the eye's response, which is maximized for yellow light. Photography provided a way of permanently recording starlight and of measuring magnitudes. Early photographic emulsions, however, were sensitive primarily to blue

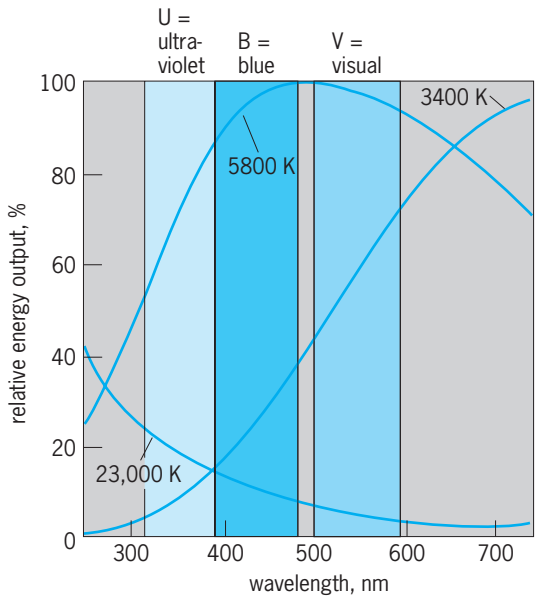

Fig. 1. Responses of the U, B (close to photographic), and V (close to visual) filters are superimposed onto blackbody curves for 3400, 5800, and 23,000 K (5660, 9980, and 40,940°F). The higher-temperature body is brighter in the blue and has a lower magnitude. (*After S. P. Wyatt and J. B. Kaler, Principles of Astronomy: A Short Version, Allyn and Bacon, 1981*)

light. As a result, blue stars look relatively brighter to the photographic plate than they do the eye, and red stars considerably fainter (**Fig. 1**). It was then necessary to establish a separate magnitude system for photography, called m_{ptg}. The original visual magnitudes were distinguished by calling them m_{vis}. The two magnitudes were set equal for white class A0-A1 stars with temperatures of 9600 K (16,820°F). Color can then be quantified by taking the difference between the two magnitudes, the color index becoming m_{ptg}-m_{vis}. Blue stars have slightly negative color indices, red stars positive ones (**Fig. 2**). *See* ASTRONOMICAL IMAGING.

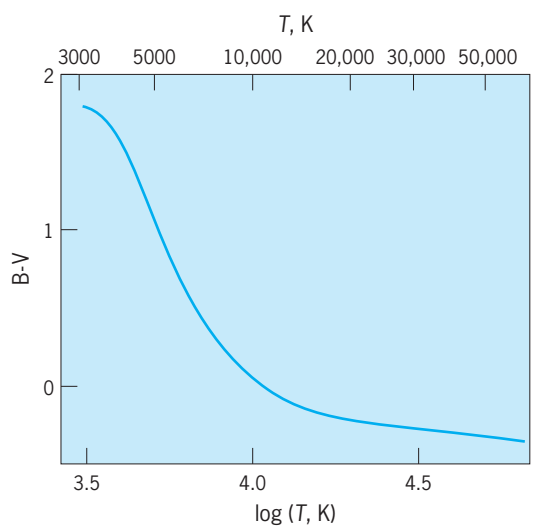

Fig. 2. Graph of the color index B-V versus the logarithm of the temperature *T* (in kelvins). As temperature climbs, the B-V color index decreases. (*After P. J. Flower, Transformations from theoretical Hertzsprung-Russell diagrams to color-magnitude diagrams: Effective temperatures, B-V colors, and bolometric corrections, Astrophys. J., 469:355–365, 1996*)

Photoelectric photometry allowed the establishment of the more precise UBV system, wherein V (yellow) and B (blue) filters respectively mimic the response of the human eye and the untreated photographic plate, and a U filter is added in the ultraviolet (Fig. 1). The traditional color index is then replaced by B-V, and the addition of U allows a second color index, U-B. Main-sequence stars trace a locus on the color-color diagram in which U-B is plotted against B-V. Deviations from the plot allow the evaluation of the degree of interstellar extinction. *See* HERTZSPRUNG-RUSSELL DIAGRAM; INTERSTELLAR EXTINCTION.

The standard system has been expanded into the red with an R filter, into the near-infrared with an I filter, and even deeper into the infrared with J, K, H, L, and M filters that go all the way to 4.75 μm. Numerous color indices are then available to examine stars that radiate primarily in the infrared. Other photometric schemes, such as the Strömgren four-color system (uvby, for ultraviolet, violet, blue, and yellow) and that used by the Sloan Survey (u′g′r′i′z′, which run from the ultraviolet to the near-infrared), allow more sophisticated color indices (for example g′-r′ instead of B-V) that are responsive not only to temperature but also to a variety of other parameters. *See* INFRARED ASTRONOMY; SLOAN DIGITAL SKY SURVEY.

James B. Kaler

Bibliography. J. B. Kaler, *The Cambridge Encyclopedia of Stars*, 2005; J. B. Kaler, *Stars and Their Spectra: An Introduction to the Spectral Sequence*, 1989.

Color vision

The ability to discriminate light on the basis of wavelength composition. It is found in humans, in other primates, and in certain species of birds, fishes, reptiles, and insects. These animals have visual receptors that respond differentially to the various wavelengths of visible light. Each type of receptor is especially sensitive to light of a particular wavelength composition. Evidence indicates that primates, including humans, possess three types of cone receptor, and that the cones of each type possess a pigment that selectively absorbs light from a particular region of the visible spectrum.

If the wavelength composition of the light is known, its color can be specified. However, the reverse statement cannot be made. A given color may usually be produced by any one of an infinite number of combinations of wavelength. This supports the conclusion that there are not many different types of color receptor. Each type is capable of being stimulated by light from a considerable region of the spectrum, rather than being narrowly tuned to a single wavelength of light. The trichromatic system of colorimetry, using only three primary colors, is based

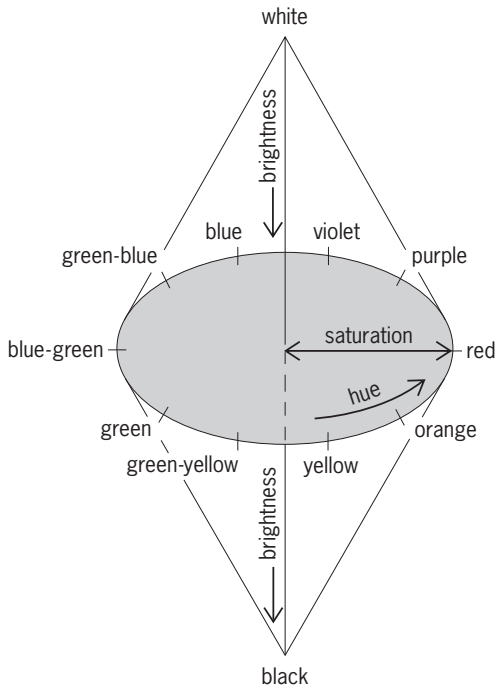

Fig. 1. Diagram of a color solid for illustrating the appearance of surface colors. (*After N. L. Munn, Psychology, 2d ed., Houghton Mifflin, 1951*)

on the concept of cone receptors with sensitivities having their peaks, respectively, in the long, middle, and short wavelengths of the spectrum. The number of such curves and their possible shapes have long been subjects for study, but not until the 1960s were direct measurements made of the spectral sensitivities of individual cone receptors in humans and various animals. *See* COLOR.

Color recognition. Color is usually presented to the individual by the surfaces of objects on which a more or less white light is falling. A red surface, for example, is one that absorbs most of the short-wave light and reflects the long-wave light to the eye. The surface colors are easily described by reference to the color solid shown in **Fig. 1**. The central axis defines the lightness or darkness of the surface as determined by its overall reflectance of white light, the lowest reflectance being called black and the highest, white. The circumference denotes hue, related primarily to the selective reflectance of the surface for particular wavelengths of light. The color solid is pointed at its top and bottom to represent the fact that as colors become whiter or blacker they lose hue. The distance from central axis to periphery indicates saturation, a characteristic that depends chiefly on the narrowness, or purity, of the band of wavelengths reflected by the surface. At the center of the figure is a medium-gray surface, one which exhibits a moderate amount of reflectance for all wavelengths of light. Colors that are opposite one another are called complementaries, for example, yellow and blue, red and blue-green, purple and green, and white and black.

Color mixture and contrast. Two complementaries, when added together in the proper proportions, can oppose or neutralize one another to produce a colorless white or gray. Various contrast effects also attest to the opposition of the complementaries. Staring at a bright-red surface against a gray background results in the appearance of a blue-green border around the red; this is an example of simultaneous contrast. Similarly, when a red light is turned off, there is frequently a negative afterimage that appears to have the opposite color, blue-green.

A set of primary colors can be chosen so that any other color can be produced from additive mixtures of the primaries in the proper proportions. Thus, red, green, and blue lights can be added together in various proportions to produce white, purple, yellow, or any of the various intermediate colors. Three-color printing, color photography, and color television are examples of the use of primaries to produce plausible imitations of colors of the original objects. *See* PHOTOGRAPHY; PRINTING.

Achromatic colors. Colors lying along a continuum from white to black are known as the gray, or achromatic, colors. They have no particular hue, and are therefore represented by the central axis of the color diagram in Fig. 1. White is shown at the top of the diagram since it represents the high-brightness extreme of the series of achromatic colors. With respect to the surface of an object, the diffuse, uniform reflectance of all wavelengths characterizes this series from black through the grays to white in order of increasing reflectance. Whiteness is a relative term; white paper, paint, and snow reflect some 80% or more of the light of all visible wavelengths, while black surfaces typically reflect less than 10% of the light. The term white is also applied to a luminous object, such as a gas or solid, at a temperature high enough to emit fairly uniformly light of all visible wavelengths. In the same connotation, a sound is described as white noise if its energy is nearly the same at all audible frequencies.

Gray. Gray is the term applied to all the intermediate colors in the series of achromatic colors. Gray may result from a mixture of two complementary colors, from a mixture of all primary colors, or from a fairly uniform mixture of lights of all wavelengths throughout the visible spectrum. Grayness is relative; a light gray is an achromatic color that is lighter than its surroundings, while a dark gray is so called because it is darker than its surroundings. Thus, the same surface may be called light or dark gray when carried from one situation to another. A gray color is one of minimum saturation; it corresponds to zero on a scale of excitation purity. At the other extreme on this scale is the color evoked by pure monochromatic light.

Black. The opposite extreme from white in the series of achromatic colors is black. Blackness is a relative term applied to surfaces that uniformly absorb large percentages of light of all visible wavelengths. A black object in sunlight absorbs a large percentage of the light, but it may reflect a larger absolute quantity of light to the eye than does a white object in the shade. Black may also be used to refer to invisible light; ultraviolet rays, for example, may be called

black light if they fall on fluorescent materials that thereby emit visible light.

Color blindness. Color blindness is a condition of faulty color vision. It appears to be the normal state of animals that are active only at night. It is also characteristic of human vision when the level of illumination is quite low or when objects are seen only at the periphery of the retina. Under these conditions, vision is mediated not by cone receptors but by rods, which respond to low intensities of light. In rare individuals, known as monochromats, there is total color blindness even at high light levels. Such persons are typically deficient or lacking in cone receptors, so that their form vision is also poor.

Dichromats are partially color-blind individuals whose vision appears to be based on two primaries rather than the normal three. Dichromatism occurs more often in men than in women because it is a sex-linked, recessive hereditary condition. One form of dichromatism is protanopia, in which there appears to be a lack of normal red-sensitive receptors. Red lights appear dim to protanopes and cannot be distinguished from dim yellow or green lights. A second form is deuteranopia, in which there is no marked reduction in the brightness of any color, but again there is a confusion of the colors normally described as red, yellow, and green. A third and much rarer form is tritanopia, which involves a confusion among the greens and blues. *See* HUMAN GENETICS.

Many so-called color-blind individuals might better be called color-weak. They are classified as anomalous trichromats because they have trichromatic vision of a sort, but fail to agree with normal subjects with respect to color matching or discrimination tests. Protanomaly is a case of this type, in which there is subnormal discrimination of red from green, with some darkening of the red end of the spectrum. Deuteranomaly is a mild form of red-green confusion with no marked brightness loss. Nearly 8% of human males have some degree of either anomalous trichromatism or dichromatism as a result of hereditary factors; less than 1% of females are color-defective. A few forms of color defect result from abnormal conditions of the visual system brought on by poisoning, drugs, or disease.

Color blindness is most commonly tested by the use of color plates in which various dots of color define a figure against a background of other dots. The normal eye readily distinguishes the figure, but the colors are so chosen that even the milder forms of color anomaly cause the figure to be indistinguishable from its background. Other tests involve the ability to mix or distinguish colored lights, or the ability to sort colored objects according to hue.

Theories. Theories of color vision are faced with the task of accounting for the facts of color mixture, contrast, and color blindness. The schema shown in **Fig. 2** may be useful in considering the various theories of color vision. It has no necessary resemblance to the structures or functions that are actually present in the visual system.

Young-Helmholtz theory. Thomas Young, early in the nineteenth century, realized that color is not merely a property of objects or surfaces. Rather, it involves a sensory experience, the characteristics of which are determined by the nature of the visual receptors and the way in which the signals that they generate are processed within the nervous system. Young was thus led to a trichromatic theory of color vision, because he found that only three types of excitation were needed to account for the range of perceived colors. Hermann von Helmholtz clarified the theory and provided it with the idea of differential absorption of light of various wavelengths by each of three types of receptor. Hence, the theory, commonly known as the Young-Helmholtz theory, is concerned with the reception phase in the color vision schema shown. This theory assumes the existence of three primary types of color receptor that respond to short, medium, and long waves of visible light, respectively. Primary colors are those that stimulate most successfully the three types of receptor; a mixture of all three primaries is seen as white or gray. A mixture of two complementary colors also stimulates all three. Blue light, for example, stimulates the blue-sensitive receptors; hence, its complementary color is yellow, which stimulates both the red receptors and the green. Similarly, the complement of red is blue-green and the complement of green is purple (blue and red). Protanopia is explained as a condition in which no long-wave sensitive receptors are present. Deuteranopia results from loss of the green receptor.

Hering theory. The theory formulated by Ewald Hering is an opponent-colors theory. In its simplest form, it states that three qualitatively different processes are present in the visual system, and that each of the three is capable of responding in two opposite ways. Thus, this theory is concerned with the processing phase (Fig. 2), in which there is a yellow-blue process, a red-green process, and a black-white process. The theory obviously was inspired by the opposition of colors that is exhibited during simultaneous contrast, afterimages, and complementary colors. It is a theory best related to the phenomenal

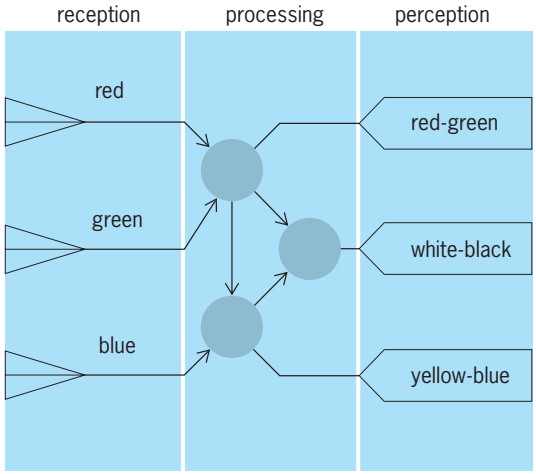

Fig. 2. Schematic representation of the three phases of color vision which can be generally used to describe the various theories of color vision.

and the neural-processing aspects of color rather than to the action of light on photopigments. This theory gains credibility when it is realized that opponent responses are common enough in the central nervous system. Well-known examples are polarization and depolarization of cell membranes, excitation and inhibition of nerve impulses, and reciprocal innervation of antagonistic response mechanism. The theory also maintains that there are four primary colors, since yellow appears to be just as unique perceptually, as red, green, or blue. However, it does not necessarily assume the existence of a separate yellow receptor.

Other theories. Various forms of stage theory or zone theory of color vision have been proposed. One form combines the principles of the Young-Helmholtz and Hering theories. It proposes that (1) there are three types of cone receptor; (2) the responses of red-sensitive, green-sensitive, and blue-sensitive receptors (and possibly others) are conducted to the higher visual centers; (3) at the same time, interactions are occurring at some stage along these separate conducting paths so that strong activity in a red response path inhibits, for example, the activity of the other response paths; and (4) this inhibiting effect is specific to the time and place of the strong red activity. The last point means that in neighboring regions and at subsequent times the blue and green mechanisms are less inhibited, so that blue-green afterimages and borders are commonly experienced as an opponent reaction to strong red stimulation.

Experimental evidence. Techniques of microspectrophotometry have been used to measure the absorption of light by single cone receptors from the eyes of primates, including humans. The results confirm the major assumption of the Young-Helmholtz theory, namely, that three types of visual receptor are specialized to absorb light over characteristic ranges of wavelength, with a maximum absorption at particular regions of the spectrum. Wide differences are found, from one receptor to another, in the wavelength at which this maximum absorption of light occurs. The three types of human cone receptors absorb maximally at about 420, 530, and 560 nanometers. In addition there are rod receptors sensitive to low intensities of light over a broad range of wavelengths peaking at about 500 nm. In each of the four types of receptor there is a photosensitive pigment that is distinguished by a particular protein molecule. This determines the range and spectral location of the light which it absorbs.

In other experiments, a microelectrode has been used to penetrate individual cone receptors in living retinal tissue of certain species of fish. Electrical responses are thus recorded as the receptor cell is stimulated by a beam of light. The wavelength of this test beam is varied from one end of the spectrum to the other. The result is that each cone shows an electrical response to light which presumably initiates impulses in the nerve cells of the visual system and is qualitatively the same regardless of the wavelength of the stimulating light. Quantitatively, however, the cones are found to differ widely in the size of response potential that they generate at each wavelength. In fact, they fall into three classes, responding maximally at about 460, 530, and 610 nm. This confirms, at least in the case of certain species of fish, the trichromatic aspect of the Young-Helmholtz theory.

The Hering theory has also received powerful support from electrophysiological experiments. A microelectrode may be used to record from retinal nerve cells that are close to the cone receptors described above in the fish. Electrical responses from some of these cells are again found to depend on the wavelength of the stimulating light. But, unlike the receptors themselves, these cells often have qualitatively different responses depending on wavelength of stimulation. Thus, a given cell may exhibit a positive response potential over one region of the spectrum and a negative one over another region. A retinal basis thus exists in these cells for the opponent process aspect of the Hering theory.

In monkeys having visual systems similar to that of humans, experiments have also shown opponent-color modes of responding. Cells in certain lower brain centers, for example, exhibit nerve impulses more or less continuously. Some of these cells are found to increase their activity in response to red light, for example, but decrease it in response to green. Other cells have the opposite reaction, and still others show blue-yellow opposition of the sort that is also demanded by the Hering theory.

Central nervous system factors are evident from the fact that a red light in one eye and a green light in the other can sometimes be fused to produce an appearance of yellow. Further evidence of central effects is found in the phenomenon of color constancy. The greenness of trees or the redness of bricks is only slightly less vivid when both are viewed through dark or yellow goggles; a black cat in the sunshine reflects more light to the eye than a white cat seen by indoor illumination. Color vision, like other forms of perception, is highly dependent on the experience of the observer and on the context in which the object is perceived. *See* EYE (INVERTEBRATE); EYE (VERTEBRATE); NERVOUS SYSTEM (VERTEBRATE); PERCEPTION; PHOTORECEPTION; VISION. Lorrin A. Riggs

Bibliography. R. Fletcher and J. Voke, *Defective Color Vision*, 1985; C. H. Graham (ed.), *Vision and Visual Perception*, 1965; L. M. Hurvich, *Color Vision*, 1981; P. K. Kaiser and R. M. Boynton, *Human Color Vision*, 2d ed., 1996; J. Mollon and T. Sharpe (eds.), *Color Vision: Physiology and Psychophysics*, 1983; W. S. Stiles, *Mechanisms of Color Vision*, 1978.

Columbiformes

An order of birds containing the family Columbidae (worldwide; true pigeons and doves) and two smaller related families. The members of this order are characterized by an ability to drink water by "sucking" (actually lapping using the tongue) instead of the sip-and-tilt method of most birds. However, some other

groups of birds are able to "suck" water by various methods. *See* CHARADRIIFORMES.

Classification. Columbiformes is here divided into the suborder Pterocletes, with the single family Pteroclidae (sandgrouse; 16 species; Old World except Australasia), and the suborder Columbae with the families Raphidae (dodos; 3 species; Mascarene Islands, western Indian Ocean) and Columbidae (pigeons and doves; 303 species; worldwide, with most groups found in Australasia to Southeast Asia). However, the relationships of the sandgrouse (Pteroclidae) are disputed; many workers argue that these birds are members of the Charadriiformes. The pigeons are divided into several subfamilies: the worldwide true pigeons (Columbinae), the pheasant-pigeon of New Guinea (Otidiphapinae), the large, crowned pigeons of New Guinea (Gourinae), the tooth-billed pigeon of Samoa (Didunculinae), and the fruit pigeons of the Old World tropics (Treroninae). Relationships of Columbiformes appear to be to the Charadriiformes in one direction and to the Psittaciformes in another. Possibly, the Columbiformes are a central stock in the evolution of birds. *See* PSITTACIFORMES.

Fossil record. The fossil record of the Columbiformes is meager, as would be expected for these terrestrial, mainly woodland and grassland birds. The extinct dodos (Raphidae), from the Mascarene Islands, are known mainly from subfossil remains.

Columbidae. Pigeons are found mainly in the tropics, but a number of species are common in temperate regions. They have a sleek plumage ranging from browns and grays to the brilliant greens, yellows, and reds of the tropical fruit pigeons (see **illustration**). They feed on seeds, fruit, and other vegetarian food. Most pigeons live in flocks but breed solitarily. Almost all are nonmigratory. The one or two young remain in the nest where they are fed by both parents, first on milk produced in the crop and then on seeds mixed with this milk. The extinct North American passenger pigeon (*Ectopistes migratorius*) lived and bred obligatorily in vast flocks, feeding largely on acorns and other mast. Although the passenger pigeon was once an abundant bird, its numbers dwindled because of overhunting and the reduction of extensive eastern forests on which it depended for food and breeding; the last individual died in captivity in 1914.

Pigeons are important game birds throughout the world. Two species of Old World origin, the rock dove (*Columba livia*) and the ring dove (*Streptopelia risoria*), have been domesticated; both have escaped and exist as feral birds throughout the world. Numerous breeds of the rock dove have been developed, which served as the primary example used by Charles Darwin in introducing his idea of evolution by natural selection. One breed of rock doves, the carrier pigeons, have served for centuries to transport messages back to their home loft. Most species of pigeons are kept and bred by aviculturists. Both domesticated species have important roles in biological research, such as studies of orientation and navigation, including work on the "magnetic compass" possessed by birds. In spite of hunting pressures, most species of pigeons are still in good numbers. A few, however, are endangered. *See* MIGRATORY BEHAVIOR.

Triangular-spotted pigeon (*Columba guinea guinea*), Fresno Zoo, California. (*Photo by Dr. Lloyd Glenn Ingles, © 2001 California Academy of Sciences*)

Raphidae. The three species of dodos and solitaires were found only in the several Mascarene Islands and were completely exterminated by the seventeenth century by sailors and released pigs. They are the symbol of extinction, with the slogan "dead as a dodo." One of the common forest trees of these islands has almost become extinct as a result; its fruits were eaten by dodos, and the hard seeds were adapted to the grinding action of the bird's stomach. The hard seed coat had to be cracked in this way so that water could reach the ovary of the excreted seed and start germination. Seeds can now be planted after cracking the seed coat artificially.

Pteroclidae. The dull-colored, ground-dwelling sandgrouse live in flocks in dry grasslands and deserts of the Old World, but depend on ponds of water to which flocks come in large numbers. The nest is generally a long distance from water, which is transported to the young by the adults by soaking their specialized belly feathers. The young drink only from these moistened feathers, and will refuse water in pans placed before them. *See* AVES. Walter J. Bock

Bibliography. D. Gibbs and E. Barnes, *Pigeons and Doves*, Yale University Press, 2001; D. Goodwin, *Pigeons and Doves of the World*, 3d ed., Cornell University Press, 1983; E. de Juana, Family Pteroclidae (Sandgrouse), in J. del Hoyo et al. (eds.), *Handbook of the Birds of the World*, vol. 4, pp. 30–57, Lynx Edicions, 1997.

Column

A structural member that carries its load in compression along its length. Most frequently, as in a building, the column is in a vertical position transmitting gravity loads from its top down to its base. Columns are present in other structures as well, such as in bridges, towers, cranes, airplanes, machinery, and furniture. Other terms used by both engineers and lay persons to identify a column are pillar, post, and strut. Columns of timber, stone, and masonry have been constructed since the dawn of civilization; modern materials also include steel, aluminum, concrete, plastic, and composite material. *See* COMPOSITE MATERIAL; LOADS, TRANSVERSE; STRUCTURAL MATERIALS; STRUCTURAL STEEL.

Fabrication. In the past, steel columns were built up of angle and plate elements welded or riveted together to form the desired cross-sectional shape (**illus**. *a*). Modern steel columns are made by rolling, extruding, or forming hot steel into predetermined cross-sectional shapes in the manufacturing facility (illus. *b–d*). Reinforced concrete columns are fabricated either in their final locations (cast-in-place concrete) or in a precast plant (precast concrete) with steel reinforcing rods embedded in the concrete (illus. *e*). Masonry columns are usually built in their final locations; they are made of brick or concrete masonry blocks; sometimes steel reinforcing rods are embedded within the masonry. *See* BRICK; CONCRETE; MASONRY; PRECAST CONCRETE; REINFORCED CONCRETE.

Classification. According to their behavior under load, columns are classified as short, slender, or intermediate. Definitions of these ranges are qualitative rather than quantitative.

A short column is one whose length is relatively short in comparison to its cross-sectional dimensions and, when loaded to its extreme, fails by reaching the compressive strength of its material. This is called failure in axial compression. A slender column is one whose length is large in comparison to its cross-sectional dimensions and, when loaded to its extreme, fails by buckling (abruptly bending) out of its straight-line shape and suddenly collapsing before reaching the compressive strength of its material. This is called a condition of instability, and it is reached at a load that depends not on the compressive strength of the material but on the stiffness of the material and the geometric properties of the column. (Stiffness is a property of materials expressed in terms of the magnitude of the load that produces a unit deformation. The greater the load required to produce a unit deformation, the stiffer the material. The geometric properties that influence the buckling of a column are the area and shape of the cross section, and the length of the column). An intermediate column falls between the classifications of short and slender. When loaded to its extreme, the intermediate column fails by a combination of compression and instability.

In addition to its length and cross section, the methods of support or fixity at the two ends of the column also influence its strength: the more effective the fixity of support, the more stable the column. (End fixity is the amount of restraint against displacement and rotation of the end of a column.)

Equilibrium, instability, and loads. As the axial load on a perfectly straight slender column with elastic material properties is increased in magnitude, this ideal column passes through three states: stable equilibrium, neutral equilibrium, and instability. The straight column under load is in stable equilibrium if a lateral force, applied between the two ends of the column, produces a small lateral deflection which disappears and the column returns to its straight form when the lateral force is removed. If the column load is gradually increased, a condition is reached in which the straight form of equilibrium becomes so-called neutral equilibrium, and a small lateral force will produce a deflection that does not disappear and the column remains in this slightly bent form when the lateral force is removed. The load at which neutral equilibrium of a column is reached is called the critical or buckling load. The state of instability is reached when a slight increase of the column load causes uncontrollably growing lateral deflections leading to complete collapse. *See* STRESS AND STRAIN; STRUCTURAL DEFLECTIONS.

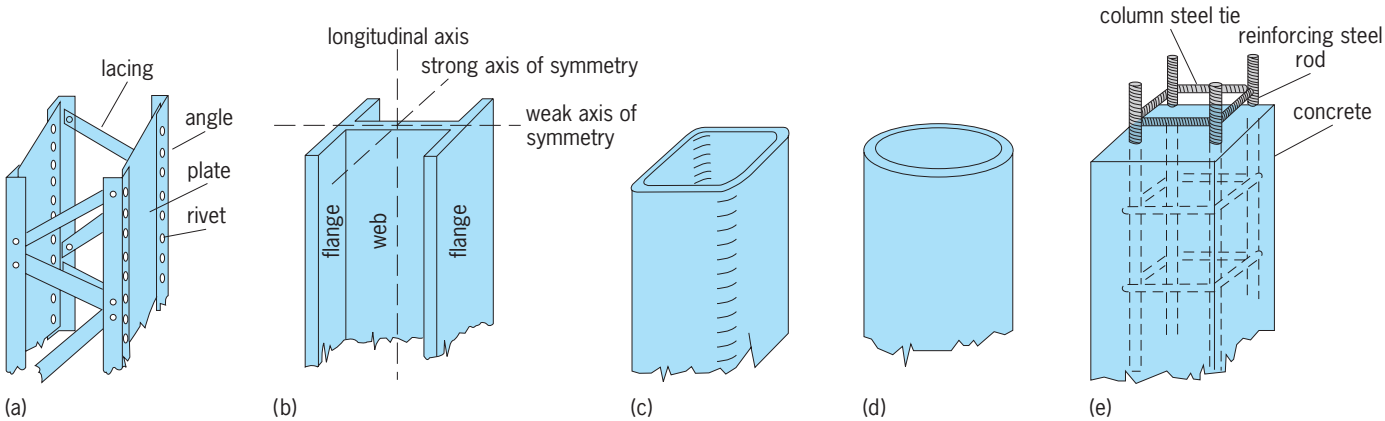

Typical column configurations. (*a*) Built-up steel column. (*b*) Wide-flange steel column. (*c*) Tubular steel column. (*d*) Pipe column. (*e*) Reinforced concrete column.

For an axially loaded straight column with any end support conditions, the equation of static equilibrium, in the form of a differential equation, can be solved for the deflected shape and critical load of the column. With hinged, fixed or free end support conditions the deflected shape in neutral equilibrium of an initially straight column with uniform cross section throughout its length always follows a partial or composite sinusoidal curve shape, and the critical load is given by Eq. (1), where E = modulus of elas-

$$P_{cr} = \frac{\pi^2 EI}{l^2} \tag{1}$$

ticity of the material, I = moment of inertia of the cross section, and l = length of the longest half sine wave along the deflected shape. A variant of (1) is given by Eq. (2), where L = actual length of the col-

$$P_{cr} = \frac{\pi^2 EI}{(KL)^2} \tag{2}$$

umn between its two end supports, K = ratio of the longest half sine wave to the actual column length, and KL = effective length (length of an equivalent hinged-hinged column). From Eq. (2) it can be noted that the buckling strength of a column is inversely proportional to the square of its length.

When the critical stress, F_{cr} ($F_{cr} = P_{cr}/A$, where A = cross-sectional area of the column), is greater than the proportional limit of the material, the column is experiencing inelastic buckling. Since at this stress the slope of the material's stress-strain curve, E_t (called the tangent modulus), is smaller than that below the proportional limit, the critical load at inelastic buckling is reduced. More complex formulas and procedures apply for such cases, but in its simplest form the critical buckling load formula is given as Eq. (3), where E_t = tangent modulus at the stress F_{cr}.

$$P_{cr} = \frac{\pi^2 E_t I}{(KL)^2} \tag{3}$$

A column with a cross section that lacks symmetry may suffer torsional buckling (sudden twisting) before, or in combination with, lateral buckling. The presence of the twisting deformations renders both theoretical analyses and practical designs rather complex. *See* TORSION.

Eccentricity of the load, or imperfections such as initial crookedness, decreases column strength. If the axial load on the column is not concentric, that is, its line of action is not precisely coincident with the centroidal axis of the column, the column is characterized as eccentrically loaded. The eccentricity of the load, or an initial curvature, subjects the column to immediate bending. The increased stresses due to the combined axial-plus-flexural stresses result in a reduced load-carrying ability. *See* BEAM COLUMN.

Design standards. Since it is difficult, if not impossible, to construct a perfectly frictionless pin and an infinitely rigid fixity as practical column support, the theoretical formulas are only approximations for real structures. The various structural design codes and standards recognize this, and they specify adjustment factors in the form of modified effective lengths or load eccentricities. In addition, because the geometries of actual columns and the materials of which they are constructed are not perfect and since actual loadings cannot be predicted precisely, factors of safety are prescribed by design codes and standards. These are applied to the calculated capacities of ideal columns to provide margins of safety against failure.

A column that contains relatively thin plate elements, such as columns of steel, aluminum, and other metals, may fail at a load less than the critical load for the column as a whole, because one or more of the plate elements buckle locally into one or a series of short waves or wrinkles. Outstanding legs of thin angles, flanges of channels, and other such plates are elements that are susceptible to local buckling, thereby reducing the stiffness, and hence the stability, of the entire column. The likelihood of local buckling can be predicted by the mathematical theory of flat plates, and its strength reduction influence is addressed by the design codes and standards for steel, aluminum, and other metals. *See* PLATE GIRDER.

Extensions. When a column is too long to be built or transported in one piece, it has to be extended or spliced at the construction site. A reinforced concrete column is extended by having the steel reinforcing bars protrude a few inches or feet above the top of the concrete, then placing the next level of reinforcing bars to overlap, and pouring the concrete of the next level. A steel column is extended by welding or bolting splice plates on the flanges and webs or walls of the columns to provide a few inches or feet of load transfer from the upper to the lower column section. A timber column is usually extended by the use of a steel tube or wrapped-around sheet-metal plate bolted onto the two connecting timber sections.

Foundations. A column that carries the load down to a foundation must have means to transfer the load without overstressing the foundation material. Reinforced concrete and masonry columns are generally built directly on top of concrete foundations. A steel column, when seated on a concrete foundation, must have a base plate to spread the load over a larger area and thereby reduce the bearing pressure. The base plate is a thick rectangular steel plate usually welded to the bottom end of the column. *See* BUILDINGS; FOUNDATIONS.

Robert T. Ratay

Bibliography. E. H. Gaylord and J. E. Stallmeyer, *Design of Steel Structures*, 3d ed., 2000; A. H. Nilson and G. Winter, *Design of Concrete Structures*, 12th ed., 1997; S. P. Timoshenko and J. M. Gere, *Theory of Elastic Stability*, 2d ed., 1961.

Combinatorial chemistry

A method in which very large numbers of chemical entities are synthesized by condensing a small number of reagents togther in all combinations defined by a small set of reactions. The main

objective of combinatorial chemistry is synthesis of arrays of chemical or biological compounds called libraries. These libraries are screened to identify useful components, such as drug candidates. Synthesis and screening are often treated as separate tasks because they require different conditions, instrumentation, and scientific expertise. Synthesis involves the development of new chemical reactions to produce the compounds, while screening aims to identify the biological effect of these compounds, such as strong binding to proteins and other biomolecular targets.

Chemical Libraries

Combinatorial chemistry is sometimes referred to as matrix chemistry. If a chemical synthesis consists of three steps, each employing one class of reagent to accomplish the conversion, then employing one type of each reagent class will yield $1 \times 1 \times 1 = 1$ product as the result of $1 + 1 + 1 = 3$ total reactions. Combining 10 types of each reagent class will yield $10 \times 10 \times 10 = 1000$ products as the result of as few as $10 + 10 + 10 = 30$ total reactions; 100 types of each reagent will yield 1,000,000 products as the result of as few as 300 total reactions. While the concept is simple, considerable strategy is required to identify 1,000,000 products worth making and to carry out their synthesis in a manner that minimizes labor and maximizes the value of the resulting organized collection, called a chemical library. *See* MATRIX THEORY.

The earliest work was motivated by a desire to discover novel ligands (that is, compounds that associate without the formation of covalent bonds) for biological macromolecules, such as proteins. Such ligands can be useful tools in understanding the structure and function of proteins; and if the ligand meets certain physiochemical constraints, it may be useful as a drug. For this reason, pharmaceutical applications provided early and strong motivation for the development of combinatorial chemistry. *See* LIGAND.

Peptide libraries. In 1984 H. M. Geysen reported on the synthesis of a library of peptides and the screening of that library to probe how changes in a single amino acid in a peptide would change its association strength with an antibody. Biochemists assume that the detailed chemical structure of any ligand will affect that ligand's binding strength to a macromolecular receptor, but even with powerful modern computer modeling methods it is normally not possible to predict either the magnitude or even the direction of the effect. By synthesizing a large number of peptides, each varying from another by only one amino acid, it was possible to determine empirically which amino acid substitutions made binding stronger and which made it weaker. The synthesis of this library benefited from the fact that chemical methods for linking amino acids together to make peptides had been optimized such that little optimization of the reaction conditions was required. Instead, the challenge was how to conveniently make thousands of peptides in a format that facilitated their use in the subsequent binding studies. The solution was to synthesize the peptides on a rack of plastic pins and to test their binding ability while the peptides remained chemically attached to the pin. While solving technical challenges, perhaps more importantly this approach highlighted a philosophy in studying ligand-receptor binding problems: start with a source of molecular diversity (that is, lots of compounds) organized in a way that makes their empirical testing straightforward; then test them all and analyze the results at the end.

Because there are 20 naturally occurring amino acids, the synthesis of a linear peptide that is n amino acids long can be done in $20n$ different ways. Thus, there are 64,000,000 possible hexapeptides ($n = 6$). It would not be convenient, perhaps not even possible, to synthesize that many peptides on individual pins. In 1985 a synthesis method was reported in which the plastic used for synthesis was encased in an inert mesh resembling a tea bag. Collections of such tea-bag reactors were subjected to the chemical addition of an amino acid at the same time. After the addition, the bags were washed thoroughly and the bags (not the bag contents) mixed. The bags were redistributed to new beakers, and another amino acid was chemically added. In this method, each bag contains only one peptide. By chemically cleaving the peptide from the polymeric support, the peptide itself can be obtained. In a variant of this procedure, a mixture of reagents can be used in any given beaker such that each bag contains a mixture of peptides bound to the polymer. While the production of such mixtures complicates the subsequent assay step, it does provide for a much greater number of assays to be accomplished. Still unresolved is the question of whether the time economy afforded by using mixtures compensates for the pitfalls inherent in their testing for binding activity.

To make very large numbers of individual peptides would require very small bags for purely practical reasons. In 1988 work involving the use of polymer resin beads was reported; approximately one-tenth the width of a pinhead, these beads served as a kind of bag. By utilizing the "split-pool" approach, it was possible to synthesize extremely large peptide libraries in which each bead possessed a single peptide. The amount of peptide on one bead is only around 200 picomoles; however, this is enough both for a simple ligand-receptor binding assay and for the analytical techniques required to establish the exact chemical structure of that peptide. *See* PEPTIDE.

Organic libraries. Because much of the practical application in discovering tightly binding ligands derives from the pharmaceutical industry, the combinatorial synthesis of druglike compound libraries is of great interest. Two practical considerations make this a greater experimental challenge than the synthesis of peptide libraries.

The synthetic methods required to make druglike molecules (that is, low-molecular-weight, organic molecules) on a polymer support have not been optimized. While solid-supported peptide synthesis originated in the early 1960s and has been extensively developed since, initial experiments with organic

solid-phase synthesis in the early 1970s were not followed up widely. In addition, while there are only 20 naturally occurring amino acids and therefore a finite number of reactions required to use them efficiently, an almost infinite number of organic chemical reagents exists and a very large number of reaction types. However, by the early 1990s several groups had reported the synthesis of moderately sized organic libraries by the solid-phase synthesis method. Each approach utilized a strategy like that of the Geysen pin method and thus was amenable to the parallel synthesis of hundreds to thousands of compounds.

The synthesis of much larger-membered libraries using the polymer-resin-bead approach was inhibited by the second practical issue: while 200 picomoles of an organic compound is enough for the ligand-receptor study, it is not enough to identify the structure of the ligand. The reason is that analytical methods for structure determination are both easier and more advanced for use with biological macromolecules such as polypeptides and polynucleotides. Beginning in the mid-1990s, the solution of bead tagging, or encoding, solved this issue. The strategy is simple: if the result of a chemical synthesis step cannot be easily read at low concentration, something should be added to the bead that conveniently encodes the reaction history of that bead for later analysis. The first reported methods of bead encoding involved the use of biological macromolecules themselves. After each step in the organic synthesis, either an amino acid or a nucleotide was added to a growing oligopeptide or oligonucleotide on the same bead so that the specific sequence could be read later, and from that sequence the reaction history could be elucidated. However, neither oligopeptides nor oligonucleotides are chemically inert enough to survive the conditions required for organic synthesis. Later, tags were introduced that were much more inert chemically. *See* ORGANIC SYNTHESIS.

More recent methods of encoding include the use of radio-frequency memory microchips and optical bar-coding strategies. **Figure 1** shows a scheme that uses optical-pattern encoding for the synthesis of a combinatorial library. Each of the nine chips has a different pattern written on it with a laser (a 3×3 matrix pattern as shown could encode for 2^9, or 512, unique chips). With the use of pattern-recognition software, the chips are sorted into appropriate reaction vessels so that chemical groups A, B, and C can be covalently added to the core organic molecule on them. After pooling the chips, optical scanning again permits sorting into appropriate reaction vessels so that a new set of three chemical groups can be added. At the end of this combinatorial synthesis involving two steps with three reagents at each, a total of 3^2, or 9, different products is obtained. Larger libraries are obtained by employing additional reaction steps, additional reagents per step, and an appropriately larger matrix pattern to enable the unique encoding of each product in the library. *See* OPTICAL INFORMATION SYSTEMS.

Major challenges in combinatorial chemistry focus on both the characterization of compounds and the screening of very large compound libraries. In addition, the use of combinatorial chemistry methods is under active study for the discovery of new catalysts, chemosensors, and other chemical substances in which prior binding of the cognate ligand (the molecule that fits into another, usually larger molecule) is necessary. *See* MOLECULAR RECOGNITION.

Anthony W. Czarnik

Assaying Problems

In combinatorial chemistry, attention has been focused on the problem of how to identify the set of molecules that possess the desired combination of properties. In a drug-discovery effort, the library members that strongly bind to a particular biological receptor are of interest. In a search for new materials that behave as superconductors at relatively high temperatures, the special combination of elements yielding the best electrical properties is a goal. In each case, the library might consist of up to a million members, while the subset of target molecules might consist of several thousand contenders or just a single highly selective binder. This subset could then be studied in more detail by conventional means.

To take advantage of the parallel nature of the combinatorial chemistry strategy, rapid screening and assaying protocols are necessary. For example, gas chromatography/mass spectrometry of bead-encoded libraries is currently limited to several analyses per hour, a difficult proposition when the researcher is faced with analyzing several thousand samples. The problem is to find ways of keeping track of each library member so that parallel screening can be coupled with parallel assaying.

Spatial array. Several emerging strategies promise to address this problem. In the first case, a library is constructed in a spatial array such that the chemical composition of each location in the array is noted during the construction. The binding molecules, usually labeled with a fluorescent tag, are exposed to the entire assay. The locations that light up can then be immediately identified from their spatial location. This approach is being actively developed for libraries of proteins and nucleotides. A problem is that the chemistry required to attach various molecules to the solid surface, usually silicon, is quite tricky and difficult to generalize. The assaying strategy is intertwined with the available procedures for synthesizing the libraries themselves.

Polystyrene-bead support. A conceptually straightforward approach is to first synthesize the library by using polystyrene beads as the solid support. The product molecules are then stripped from the support and pooled together into a master solution. This complex mixture consisting of a potentially large selection of ligand molecules could then be exposed to an excess of a target receptor. The next step is to devise a method for identifying the ligand-receptor pairs that point to molecularly specific binding. One approach is to examine a part of the mixture en masse by using affinity capillary electrophoresis. With this technique, the migration times of the ligand-receptor pair are significantly longer than the

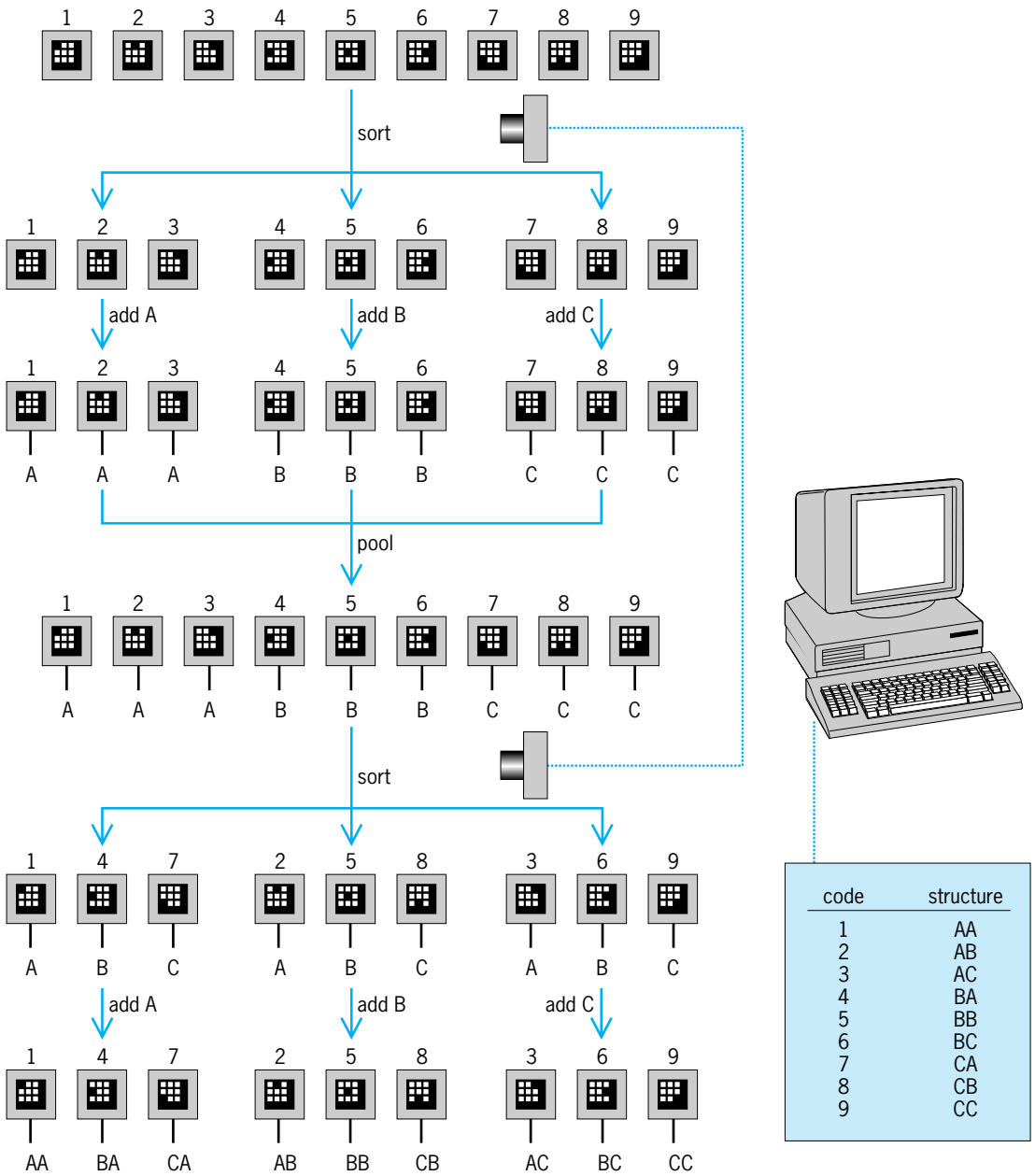

Fig. 1. Scheme for using optical pattern encoding for the synthesis of a combinatorial library. (*Courtesy of X. Xiao*)

unreactive ligands, and can be interrogated by electrospray mass spectrometry.

Mass spectrometry. The mass spectrometric method often provides a direct structural identification of the ligand, either by determination of its molecular weight or by collision-induced dissociation experiments. In the latter case, the molecular ion is selected by a primary mass spectrometer and is driven into a region of high-pressure inert gas for fragmentation. The fragment ions are then used to reconstruct the original molecular structure. This direct approach to screening and assaying has the advantage that the screening is carried out in solution rather than on a solid support, and it avoids steric problems associated with resin-bound molecules. At present the approach seems limited to libraries of about 1000 compounds because of interference from unbound ligands, and limited by sensitivity issues. New strategies using mass spectrometry may eliminate this limit.

A different tack involves assaying the polystyrene beads one by one after the resin-bound molecules are exposed to a receptor. With this approach, active beads may be identified by color or by fluorescence associated with the receptor, and are subsequently indexed in standard 96-well titer plates. Identification is then possible by using a variety of spectroscopic techniques; at present, the most popular methods are electrospray mass spectrometry and matrix-assisted laser desorption ionization mass spectrometry.

Successful structure determination using mass spectrometry is often routine, particularly for protein libraries where sequencing strategies have been

worked out; however, ambiguities often arise for nonpeptide libraries. Encoding strategies promise to resolve this issue. For example, it is now possible to construct the library such that the first molecule attached to the resin can be identified by its ratio of carbon-13 to carbon-12. By using isotope enrichment protocols, a particular isotope ratio can be associated with a specific molecule. Then, for a library synthesized with three steps, the molecule in the first step can be identified from its isotope ratio, the molecule in the third step is known from the synthesis, and the molecular weight is measured by mass spectrometry. Hence, the molecular weight of the molecule used in the second step can be determined by the difference. This approach has been successfully employed to assay a 1000-member library. There are many ways that this approach could be extended to larger systems.

Another advantage of the isotope-encoding strategy is that the code is actually carried with the molecule after it is cleaved from the bead. This property may have important implications when performing bioassays. This scheme appears to be functional in a practical sense. The mass spectrometry methods are generally sensitive enough ($\approx$100 femtomole) to provide a reliable assay, and robotic techniques can be developed to increase the number of assays to perhaps a few dozen per hour.

Another level of sophistication involves the assay of a single polystyrene bead by using a method such as infrared spectroscopy. The advantages of optical techniques are that they are nondestructive and can be employed without removing the target molecules from the linker. There are, however, several obvious drawbacks, including lack of specificity and sensitivity for single-bead studies. Recently, however, experiments using infrared microspectroscopy have partially overcome the sensitivity limitations. The selectivity issue has been addressed by using a carbon-deuterium stretching frequency to determine the deuterium content of a given compound. This signature can lead to identification under suitable circumstances. *See* MASS SPECTROMETRY.

Improving assaying power. It would be desirable to speed up the whole process. To fully capitalize on the elegant concept implicit in combinatorial chemistry, parallel screening methods capable of identifying large numbers of library members need to be developed. When the libraries are synthesized on beads, it would also be preferable to assay the compound without cleaving it from the polymeric resin support. This process inherently destroys the library and opens the possibility of incomplete cleaving reactions. One approach is to take advantage of a special type of mass spectrometry where it is possible to record spectra from very small areas of a solid surface. With this technique, the sample is bombarded by a focused energetic ion beam having a kinetic energy of several thousand electronvolts. The energetic particle loses some of its momentum in the top layers of the solid, causing desorption of molecular ions near the point of impact. If the energetic ion beam is formed in a short (nanosecond) pulse, the secondary ions may be measured by using time-of-flight detection. The resulting mass spectra can be recorded from an area that is much less than 1 square micrometer. Hence, it is feasible to spatially resolve the chemical components on a single resin particle that is typically in the 20–300-μm size range.

Experiments using this idea were attempted in the late 1990s. The first attempts were successful only if the covalent bonds attaching the molecules to the resin were first clipped. This was accomplished by exposing the bead to a vapor of trifluoracetic acid, a standard release agent for acid-sensitive linking moieties. After clipping, the molecules were found to remain on their respective beads, even if prior to their treatment they were essentially touching. For example, in this method, two distinct types of coated beads can be used—one with phenylalanine and one with leucine. It might be possible to extend this method to a collection of thousands of beads arrayed onto a plate.

None of the above approaches provides assaying power that satisfies the need to characterize massive combinatorial libraries. Many schemes appear to have enough sensitivity or selectivity to perform the job, but whether any will be truly practical remains to be seen. It is likely, given the high level of activity in this field, that one of the above methods (or perhaps a completely new one) will become practical. *See* ANALYTICAL CHEMISTRY; ELECTROPHORESIS.

Nicholas Winograd

Dynamic Combinatorial Chemistry

Dynamic combinatorial chemistry is in its early developmental stage. It integrates library synthesis and screening in one process, potentially accelerating the discovery of useful compounds. In the dynamic approach the libraries are not created as arrays of individual compounds, but are generated as mixtures of components, similar to natural pools of antibodies. One important requirement is that the mixture components (**Fig. 2**) exist in dynamic equilibrium with each other. According to basic laws of thermodynamics (Le Chatelier's principle), if one of the

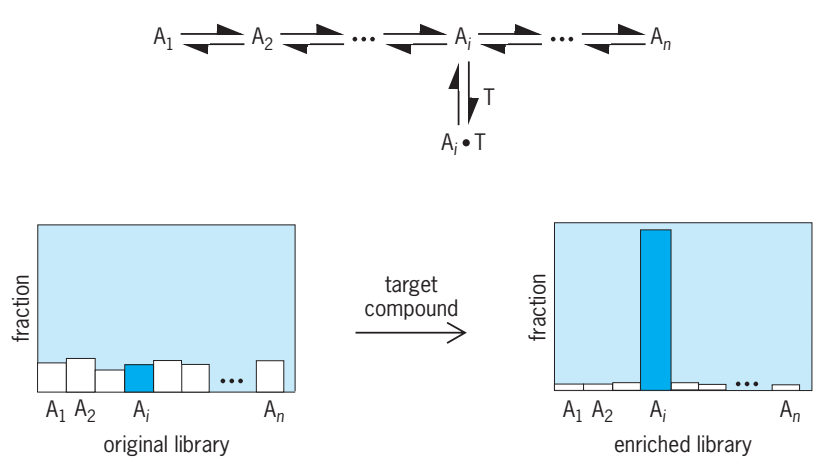

Fig. 2. Dynamic equilibrium for mixture components ($A_1 \ldots A_n$), following Le Chatelier's principle. T indicates the target.

components (A_i) is removed from the equilibrated mixture, the system will respond by producing more of the removed component to maintain the equilibrium balance in the mixture. *See* CHEMICAL EQUILIBRIUM; LE CHATELIER'S PRINCIPLE.

The dynamic mixture, as any other combinatorial library, is so designed that some of the components have potentially high affinity to a biomolecular target. These high-affinity (effective) components can form strong complexes with the target. If the target is added to the equilibrated mixture, when the effective components form complexes with the target they are removed from the equilibrium. This forces the system to make more of these components at the expense of other ones that bind to the target with less strength. As a result of such an equilibrium shift, the combinatorial library reorganizes to increase the amount of strong binders and decrease the amount of the weaker ones (Fig. 2). This reorganization leads to enrichment of the library with the effective components and simplifies their identification.

The general idea of the equilibrium shift is not new, and its combinatorial application was eventually demonstrated by several groups. As an example, researchers described a dynamic library that produced inhibitors for the enzyme carbonic anhydrase. The components of this library were formed by a reversible reaction between two series of compounds: amines and aldehydes (**Fig. 3***a*). The products of this reaction, Schiff bases, existed in the dynamic equilibrium with the starting amines and aldehydes (building blocks) and with each other. Functional groups in the building blocks (R^x and R^y) could recognize binding pockets of the target enzyme, and the best combination of those groups could make a particularly strong binder. To render the library components stable to hydrolysis, they were converted to similar amines (Fig. 3*a*) with retention of their basic target-binding properties. The researchers showed that the addition of carbonic anhydrase to the library led to the preferred formation of a sulfonamide compound (Fig. 3*b*), as compared to other library components. Structures similar to the sulfonamide compound are known to inhibit carbonic anhydrase via strong binding to its active site.

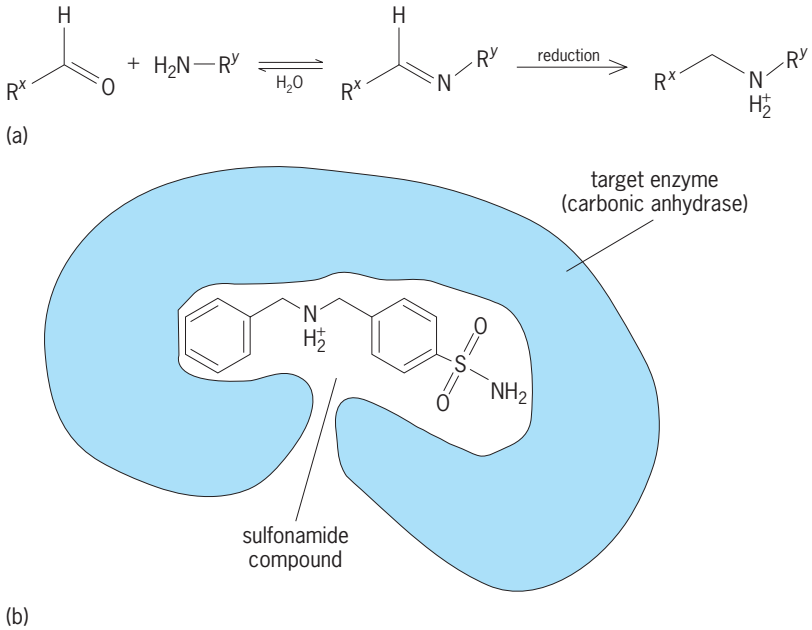

Fig. 3. Equilibrium shift utilized to produce a dynamic library. (*a*) Reversible reaction between amines and aldehydes produces a Schiff base. (*b*) Addition of the enzyme carbonic anhydrase leads to the formation of the preferred sulfonamide compound.

This example shows that the dynamic combinatorial approach requires a reversible reaction that can keep multiple components in equilibrium. A number of such reactions have been used, leading to various types of dynamic libraries. These reactions include peptide bond formation and cleavage to form libraries of peptides that bind to an antibody, photoisomerization of alkenes that changes their shape to fit a target amino acid, transesterification resulting in dynamic libraries of macrocycles that can selectively recognize metal ions, thiol-disulfide exchange that organizes a library of artificial peptide receptors, and scrambling of ligands in zinc complexes to find an effective deoxyribonucleic acid (DNA) binder.

All the above examples, dealing with relatively small libraries, were designed to prove the concept of the dynamic approach. Remarkably, however, all the systems described so far use different base reactions, target compounds, reaction media, and instrumentation. This fact indicates that numerous variations of the dynamic approach can be expected to develop in the future.

Problems. The dynamic combinatorial approach has some inherent problems. One is that it combines two intrinsically different processes: equilibrium between the library components and molecular recognition (selective binding) of the target, which generally require different conditions. One way to overcome this problem is to physically separate the equilibration and binding sites. Such separation has been accomplished by a stepwise approach to the organization of the dynamic libraries (chemical evolution). The target compound is immobilized by attachment to solid beads placed in a continuous-flow column, similar to a chromatography column (selection site). The immobilization does not allow the target to leave the column, and retains the target's properties. When the combinatorial mixture passes through the selection site, the effective components remain, bound to the target. Next, the depleted mixture enters another column (equilibration site), where special conditions, such as heating, catalysis, and irradiation, restore the equilibrium in the combinatorial mixture (**Fig. 4**). During this restoration, part of the mixture converts back into the effective components. The combinatorial mixture then returns to the selection site. After a number of such cycles, the original mixture evolves into a high-affinity subset accumulated on the target-containing beads. Later, this subset can be washed off the target and analyzed to identify the effective components.

Dynamic combinatorial chemistry is an emerging method that can potentially grow into a powerful tool for the discovery of new and useful compounds

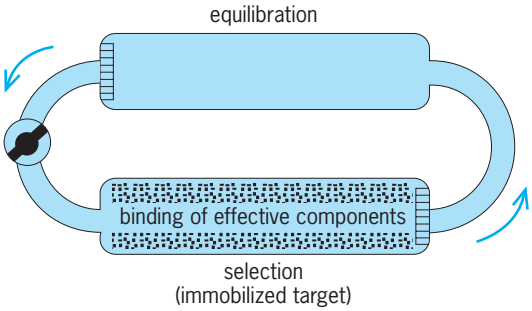

Fig. 4. Separation of equilibration and binding sites in a dynamic combinatorial experiment.

in molecular libraries. This approach may find its niche, for example, in the drug discovery process. The success of this method will depend upon the development of new reversible reactions that can be used to generate the dynamic libraries, and of analytical methods that help identify the structure of components in complex mixtures. Alexey V. Eliseev

Bibliography. P. A. Brady and J. K. M. Sanders, Thermodynamically-controlled cyclisation and interconversion of oligocholates: Metal ion templated living macrolactonisation, *J. Chem. Soc. Perkin Trans.*, 1:3237–3253, 1997; C. L. Brummel et al., A mass spectrometric solution to the address problem of combinatorial libraries, *Science*, 264:399–402, 1994; Y. Chu et al., Affinity capillary electrophoresis: Mass spectrometry for screening combinatorial libraries, *J. Amer. Chem. Soc.*, 118:7827–7835, 1996; S. H. DeWitt and A. W. Czarnik (eds.), *A Practical Guide to Combinatorial Chemistry*, 1997; A. V. Eliseev and M. I. Nelen, Use of molecular recognition to drive chemical evolution, 1. Controlling the composition of an equilibrating mixture of simple arginine receptors, *J. Amer. Chem. Soc.*, 119:1147–1148, 1997; P. A. Fodor et al., Light directed spatially addressed parallel chemical synthesis, *Science*, 251:767–773, 1991; H. M. Geysen et al., Isotopes or mass in coating of combinatorial libraries, *Chem. Biol.*, 3:679–688, 1996; H. Hioki and W. C. Still, Chemical evolution: A model system that selects and amplifies a receptor for the tripeptide (D)Pro(L)Val(D)Val, *J. Org. Chem.*, 63:904–905, 1998; I. Huc and J. M. Lehn, Virtual combinatorial libraries: Dynamic generation of molecular and supramolecular diversity by self-assembly, *Proc. Nat. Acad. Sci. USA*, 94:8272, 1997; B. Klekota, M. H. Hammond, and B. L. Miller, Generation of novel DNA-binding compounds by selection and amplification from self-assembled combinatorial libraries, *Tetrahed. Lett.*, 38:8639–8642, 1997; K. Russell et al., Analytical techniques for combinatorial chemistry: Quantitative infrared spectroscopic measurements of deuterium-labeled protecting groups, *J. Amer. Chem. Soc.*, 118:7941–7945, 1996; P. G. Swann et al., Nonspecific protease-catalyzed hydrolysis synthesis of a mixture of peptides: Product diversity and ligand amplification by a molecular trap, *Biopolymers*, 40:617–625, 1996; S. H. Wilson and A. W. Czarnik (eds.), *Combinatorial Chemistry: Synthesis and Application*, 1997.

Combinatorial synthesis

A method for preparing a large number of chemical compounds, commonly known as a combinatorial library, which are then screened to identify compounds having a desired function, such as a particular biological or catalytic activity. Combinatorial synthesis is an aspect of combinatorial chemistry, which allows for the simultaneous generation and rapid testing for a desired property of large numbers of chemically related compounds. One could regard combinatorial chemistry as the scientist's attempt to mimic the natural principles of random mutation and selection of the fittest. Combinatorial chemistry has already become an invaluable tool in the areas of molecular recognition, materials science, drug discovery and optimization, and catalyst development. *See* COMBINATORIAL CHEMISTRY; ORGANIC SYNTHESIS.

Combinatorial synthesis was developed to prepare libraries of organic compounds containing from a few dozen to several million members simultaneously, in contrast to traditional organic synthesis where one target compound is prepared at a time in one reaction (**Fig. 1**). Thus, a target compound AB, for example, would be prepared by coupling of the substrates A and B in a traditional (orthodox) synthesis and would be isolated after reaction processing (workup) and purification (such as crystallization, chromatography, or distillation). Combinatorial synthesis offers the potential to prepare every combination of substrates, type A_{1-m} and type B_{1-n}, providing a set of compounds, $A_{(1-m)}B_{(1-n)}$. There are two main approaches to combinatorial synthesis: parallel synthesis and split synthesis (also known as split-and-pool or mix-and-split).

Parallel synthesis. In parallel synthesis, each compound is prepared separately in an individual reaction vessel. In order to accomplish parallel combinatorial synthesis, large numbers of reactions must be carried out simultaneously and the resulting products must be isolated and purified very quickly. Recently, various automated devices, such as weighing, reaction, filtration, and chromatography machines, were developed for high-throughput parallel combinatorial synthesis. Another approach employs insoluble reagents that can be removed by filtration from a solution-phase reaction mixture. Acids, bases, reducing reagents, oxidation reagents, or various catalysts are attached to insoluble supports (such as

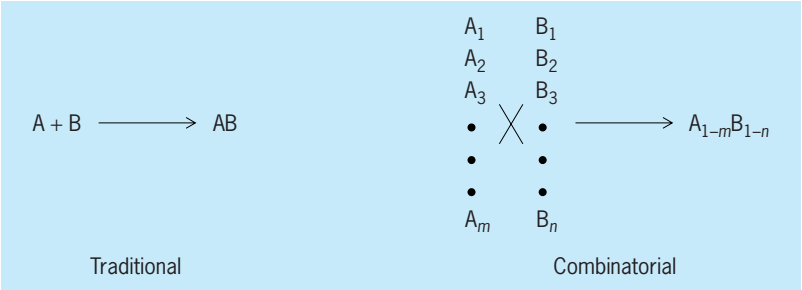

Fig. 1. Traditional synthesis versus combinatorial synthesis.

silica-gel or polymers) in this approach. Functional groups bound to insoluble supports are also applied to scavenge unreacted starting materials in solution.

Transformation of a substrate anchored to an insoluble support by treatment with a solution of reagents, generally called solid-phase synthesis, is a frequently used protocol for combinatorial synthesis. A starting substrate is first attached to a solid support by an appropriate linker group and undergoes various synthetic transformations to furnish a product on the solid support which is readily separated from the unreacted reagents and coproducts in the solution by filtration. The resulting solid-supported product can then be used in subsequent transformation steps. Solid-phase peptide synthesis is a typical example. The final products, as well as the synthetic intermediates, are detached from the solid support on demand. An example of the solid-phase parallel synthesis giving a $3 \times 3 = 9$-membered library is shown in **Fig. 2**. This library can be screened for optimum "hit" candidates before or after release of the supported products; however, the preferred method for biological screening is to test the compounds in solution after cleavage from the solid support. $A_{1\text{-}3}$ and $B_{1\text{-}3}$ could be amino acids to give dipeptides, nucleotides to give dinucleotides, or sugars to give disaccharides. *See* NUCLEOTIDE; PEPTIDE.

In both protocols—solution-phase synthesis with solid-supported reagents and solid-phase synthesis with solution-phase reagents—various polymers, glass surfaces, and silica gel are typical solid supports; cross-linked polystyrene beads having appropriate functional groups (such as NH_2, Cl, OH, and COOH) are the most frequently used.

Split synthesis. Polystyrene beads bearing a reactive functional group are also used as a starting support in split synthesis. The beads range from 10 to 200 micrometers in diameter. For example, 1 g of beads having a 100-μm diameter contains about 1 million beads. The split synthesis is illustrated in **Fig. 3** for a library of $3 \times 3 \times 3 = 27$ possible combinations of trimeric products. Three steps of solid-phase synthesis to attach A, B, and C to the starting beads (step 1) are carried out individually to give the supported A, B, and C. The resulting three kinds of beads are suspended in a solvent, mixed by shaking, and filtered to give one portion of a mixture. The pooled mixture of beads is partitioned to three portions where each portion of beads contains A, B, and C in a statistical ratio (1:1:1). Similar manipulations (step 2) afford three portions of beads where each portion contains every combination of nine compounds (from AA to CC) in a statistical ratio. The three portions of a mixture consisting of nine compounds are subjected to coupling with X, Y, and Z (step 3) to give a nine-membered mixture having a terminal group X, a nine-membered mixture having a terminal group Y, and a nine-membered mixture having a terminal group Z, respectively. Finally, the resulting beads are all mixed together to

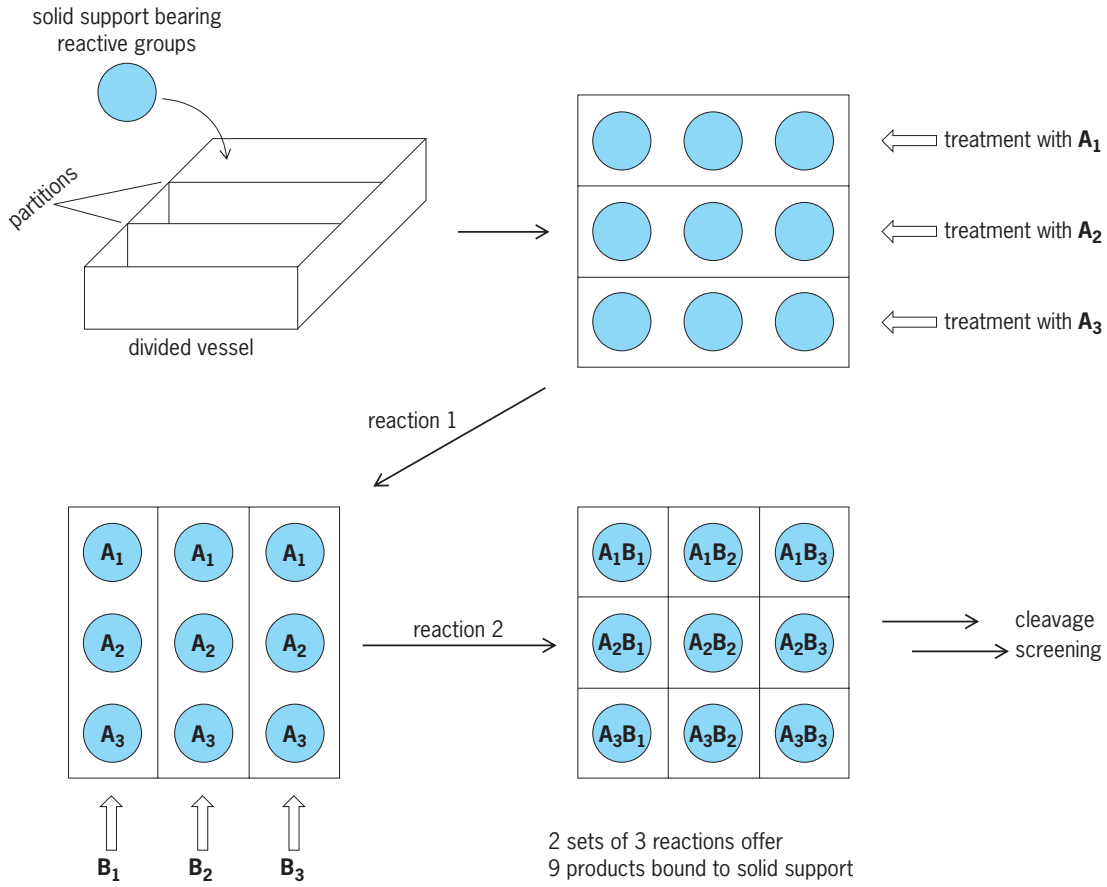

Fig. 2. Solid-phase parallel synthesis of a nine-membered library.

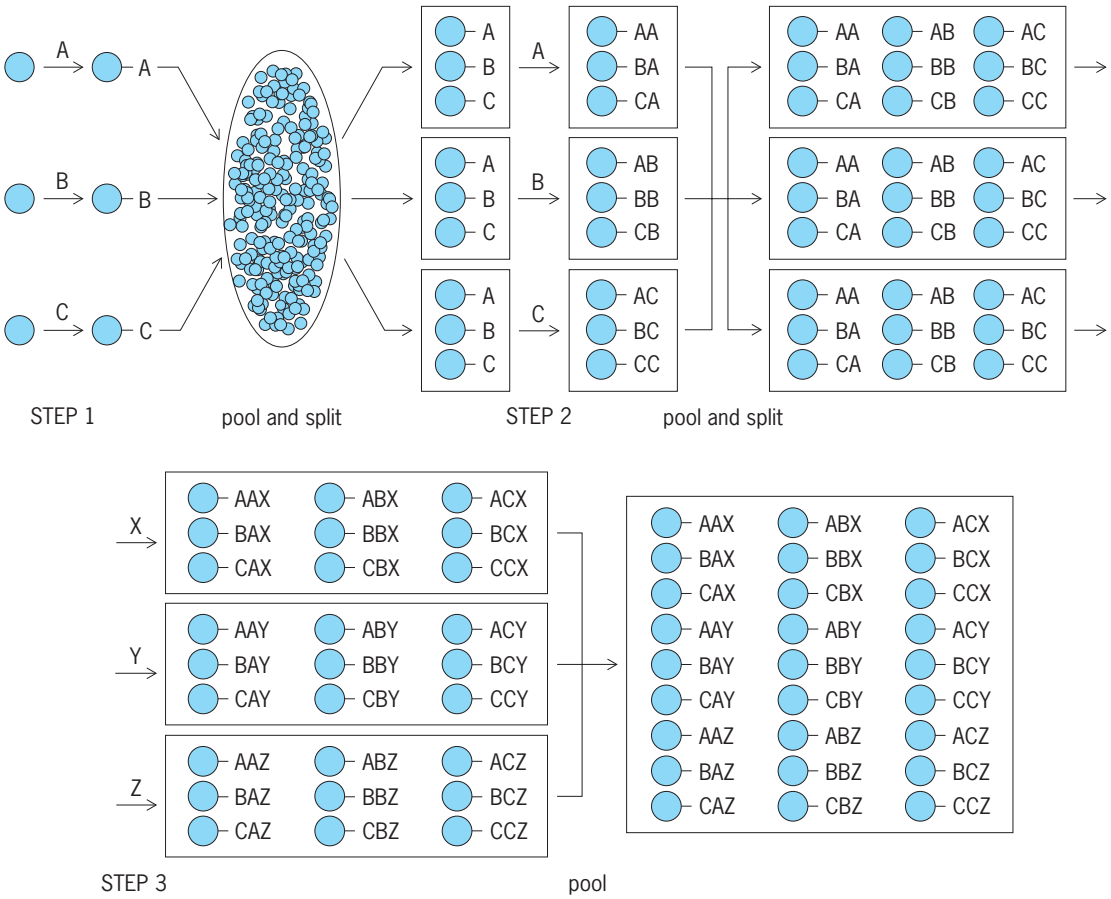

Fig. 3. Solid-phase split synthesis of a library consisting of 27 compounds.

give a library consisting of 27 compounds bound to the solid support, where three sets of three coupling reactions ($3 \times 3 = 9$ manipulations) afford the cube of three ($3^3 = 27$) compounds. In this manner, split synthesis can offer an enormous number of combinatorial products. For example, linear coupling of 10 amino acids five times ($10 \times 5 = 50$ manipulations) can furnish 10^5 pentapeptides. Since one bead of a library prepared by split synthesis must have one compound, the library is also referred to as the one-bead–one-compound library. Identification of the most active compound from the one-bead–one-compound library has gained in popularity with the advent of more efficient methods for one-bead spectroscopic analysis. Yasuhiro Uozumi

Bibliography. B. A. Bunin, *The Combinatorial Index*, Academic Press, 1998; G. Jung (ed.), *Combinatorial Chemistry: Synthesis, Analysis, Screening*, Wiley-VCH, 1999; K. C. Nicolaou, R. Hanko, and W. Hartwig (eds.), *Handbook of Combinatorial Chemistry*, Wiley-VCH, 2002; N. K. Terrett, *Combinatorial Chemistry*, Oxford, 1998.

Combinatorial theory

The branch of mathematics which studies arrangements of elements (usually a finite number) into sets under certain prescribed constraints. Problems combinatorialists attempt to solve include the enumeration problem (how many such arrangements are there?), the structure problem (what are the properties of these arrangements and how efficiently can associated calculations be made?), and, when the constraints become more subtle, the existence problem (is there such an arrangement?).

Other names for the field include combinatorial mathematics, combinatorics, and combinatorial analysis, the last term having been coined by G. W. Leibniz in his *Dissertatio de arte combinatoria* (1666) wherein he hoped for application of this field "in the entire sphere of science." Indeed, the availability today of high-speed computers which can deal with large but finite configurations seems to be bringing his prophecy near fulfillment. On the other hand, the study of combinatorial phenomena dates back to the very beginnings of mathematics. According to a Chinese legend recorded in a divinatory book of the lesser Taoists (about 2200 B.C.), the magic square

$$\begin{matrix} 4 & 9 & 2 \\ 3 & 5 & 7 \\ 8 & 1 & 6 \end{matrix}$$

was observed on the back of a divine tortoise. Apparently, permutations were considered as early as 1100 B.C., and Rabbi Ben Ezra, in the twelfth century, probably knew the formula for the number of k-element subsets of an n-element set.

Enumeration

From the seventeenth well into the twentieth century, combinatorial theory was used primarily to solve certain problems in probability. For example, to compute the probability that exactly three heads appear in five flips of a coin, one need only count the total number of sequences of heads and tails which may occur (32) and divide this number into the number of sequences which result in three heads and two tails (10). *See* PROBABILITY.

Permutations and combinations. In enumeration of arrangements it is always essential to know which arrangements are considered the same and which different. When k objects are to be selected from a set S and different orderings on the selected objects are to be counted separately, the arrangements are called permutations or relabelings of the set of k objects. But if order is to be disregarded, so that attention is focused only on the subset consisting of the k selected elements, then the choice is termed a combination. If elements of S are allowed to be selected more than once, then the arrangements are called permutations with repetition or combinations with repetition. If S has n different elements, Eqs. (1)–(4) enumerate these

$$P(n,k) = n(n-1)(n-2)\cdots(n-k+1) = \frac{n!}{(n-k)!} \tag{1}$$

$$C(n,k) = \binom{n}{k} = \frac{P(n,k)}{k!} = C(n,n-k) \tag{2}$$

$$\overline{P}(n,k) = n^k \tag{3}$$

$$\overline{C}(n,k) = \binom{n+k-1}{k} \tag{4}$$

four basic cases. The number of k-permutations is shown in Eq. (1). In particular, $P(n, n) = n! = n(n-1)\cdots(2)\,2\cdot 1(1)$ is the number of ways to order S. The number of k-combinations (or k-element subsets of S) is shown in Eq. (2). When repetition is allowed, Eq. (3) gives the number of permutations with repetition (or k-tuples) and Eq. (4) gives the number of combinations with repetition. Formula (4), although not as widely known as formula (2), recurs throughout mathematics: for instance, it gives the number of mixed partial derivatives of order k with n variables, or the number of solutions in nonnegative integers of the equation $x_1 + \cdots + x_n = k$. *See* CALCULUS.

Generating functions. The symbol

$$\binom{n}{k}$$

in Eq. (2) is sometimes called a binomial coefficient, because it is the coefficient of x^k in the expansion of $(1+x)^n$ given in Eq. (5). When a finite or infinite

$$C(n,x) = (1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k \tag{5}$$

sequence $(a_0, a_1, \ldots)$ is exhibited in this way as the coefficients of a polynomial or series $p(x) = \Sigma a_i x^i$, then $p(x)$ is called a generating function for the sequence. The general method of generating functions was developed by P. S. Laplace in 1812 but had already been used extensively in the eighteenth century by L. Euler. For example, substituting $x = -1$ in Eq. (5) shows that there are exactly as many even subsets of a set as odd subsets (and thus the probability of an even number of heads in n flips of a coin is $^1/_2$). *See* BINOMIAL THEOREM; SERIES.

Recursions. Algebraic identities involving generating functions are often proved from recursions (or difference equations) satisfied by the enumerating sequence. Thus the numbers $C(n,k)$ are uniquely determined by the boundary conditions $C(n,0)=1$ for all n and $C(0,k) = 0$ for all $k > 0$, and by recursion (6). The generating function defined by Eq. (5) sat-

$$C(n,k) = C(n-1,\ k-1) + C(n-1,\ k) \tag{6}$$

isfies the algebraic identity $C(n,x) = (1+x)C(n-1,x)$.

Catalan numbers. The Catalan numbers are given in Eq. (7). They count the ways to insert parentheses

$$c_n = \frac{1}{n}\binom{2n-2}{n-1} \tag{7}$$

in a string of n terms so that their product may be unambiguously carried out by multiplying two quantities at a time. Thus $c_4 = 5$, since $abcde$ may be computed by $(ab)(cd)$, $((ab)c)d$, $(a(bc))d$, $a((bc)d)$, or $a(b(cd))$. The Catalan numbers also enumerate the ways to place nonintersecting diameters in a convex polygon with $n+1$ vertices so that only triangular regions result. Further, c_n gives the number of ways to draw in succession all the balls from a container which originally contains $n-1$ white balls and $n-1$ black balls, so that at every stage there are at least as many black balls as white balls remaining in the container. Catalan numbers also appear in the theoretical measurement of preferences. Let A be a finite set, and let f be a function which assigns a real number to each element of A. Then x is said to be preferred to y if $fx \geq f(y) + 1$, and this induces a partial order on the elements of A. Thus, if $A=\{x, y, z\}$ and $f(x) = 0, f(y) = {}^1/_2$, and $f(z) = 1$, the only preference is z to x. Assignments $\{1,2,3\}$, $\{0,1,1\}$, $\{0,0,1\}$, and $\{0,0,0\}$ give the four other essentially distinct preferential orders on A, and, in general, there are $c_n + 1$ such orders on an n-element set.

Catalan numbers obey the recursion $c_n = c_1c_{n-1} + c_{n-2} + \cdots + c_{n-1}c_1$ and have the generating function $c(x) = {}^1/_2\,(1 - \sqrt{1-4x})$, where the root is to be expanded by means of the series form of the binomial expansion.

Stirling numbers. If instead of selecting elements from the set S, it is desired to distribute the elements of S into r indistinguishable cells so that no cell remains empty, there results a set-partition or partition of S into r blocks, and these are counted by the Stirling number of the second kind $S(n, r)$. The Stirling

numbers satisfy the boundary conditions $S(n, 1) = S(n, n) = 1$, and the recursion $S(n, r) = S(n-1, r-1) + rS(n-1, r)$ [which may be established by distinguishing within the family of partitions between the distributions for which the first $n-1$ elements of S leave a cell unoccupied which must be filled by the last element and those for which the first $n-1$ elements occupy every cell]. In addition, they satisfy formula (8), where $(x)_r$ stands for falling factorial

$$x^n = \sum_{r=1}^{n} S(n, r)(x)_r \tag{8}$$

$x(x-1)(x-2)\cdots(x-r+1)$. The Bell number, given by Eq. (9), is the total number of partitions

$$B(n) = \sum_{r=1}^{n} S(n, r) \tag{9}$$

of an n-element set. These numbers also give the number of rhyme schemes for a poem of n lines. There are $B(4) = 15$ ways to rhyme a quatrain: *abcd*, *aabc*, *abac*, ..., *aabb*, ..., *abbb*, *aaaa*. The Stirling numbers of the first kind, $s(n, r)$, are defined as the coefficients of the generating function of $(x)_n$.

Fibonacci numbers. The Fibonacci numbers are defined by the elementary recursion $f_{n+2} = f_n + f_{n+1}$; $f_0 = 0, f_1 = 1$. Each f_n counts, for example, the number of sequences of $n-2$ zeros and ones with no consecutive zeros. The generating function $f(x)$ of the Fibonacci numbers satisfies the identity $(1 - x - x^2)f(x) = x$, from which the closed form $f_n = (\tau^n - (1-\tau)^n)/\sqrt{5}$; where $\tau = (1+\sqrt{5})/2$, may readily be deduced.

Asymptotic formulas. Occasionally, only a rough estimate of the number of arrangements of a certain kind is desired (for example, to estimate how long a computer will take to list all such arrangements), and then an asymptotic enumeration formula is sought. Thus, $n!$, which is the number of (complete) permutations $P(n,n)$, obeys for large n Stirling's formula (10), where $e = 2.71828\ldots$ [here $f(n) \approx$

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \tag{10}$$

$g(n)$ means that as n grows without bound the ratio $f(n)/g(n)$ approaches 1]. *See* E (MATHEMATICS).

Integer partitions. An integer partition of n is a nonincreasing sequence of positive integers whose sum is n; the number of these partitions is denoted by $p(n)$. Thus $p(4) = 5$, since $4, 3+1, 2+2, 2+1+1$, and $1+1+1+1$ all partition 4. No usable formula for $p(n)$ is known, but Euler discovered an elegant expression for the generating function: $p(x) = (1-x)^{-1}(1-x^2)^{-1}\cdots(1-x^{m})^{-1}\ldots$. Asymptotically, $p(n)$ obeys formula (11), where e^x is the exponential

$$p(n) \approx \left(\frac{1}{4n\sqrt{3}}\right) e^{\pi\sqrt{2n/3}} \tag{11}$$

function which uses the base e. Thus $p(n)$ is eventually greater than any polynomial function of n but is eventually less than a^n for any $a > 1$.

If the requirement that the sequence be nondecreasing is removed, so that, for example, the partition $1+3$ is considered distinct from the partition $3+1$, then the number of such partitions of n is much easier to count and is 2^{n-1}. If $\overline{p}(n)$ denotes the number of partitions of n into distinct parts, then $\overline{p}(x) = (1+x)(1+x^2)(1+x^3)\cdots$, and if $\overline{p}_k(n)$ is the number of partitions of n into distinct parts whose largest part is at most k, then $\overline{p}_k(x) = (1+x)(1+x^2)\cdots(1+x^k)$.

Partially ordered sets. Results in asymptotic enumeration include a formula for the number of partially ordered sets with n elements. The asymptotic estimate of approximately $2^{n^2/4}$ for the number of partially ordered sets with n elements results from the observation that as n becomes large the partially ordered sets which contain four elements w, x, y, z such that $w < x < y < z$ become asymptotically inconsequential (in the sense that the probability that a random partially ordered set contains a four-element chain approaches zero). *See* LATTICE (MATHEMATICS).

Probabilistic method. A nonconstructive proof technique termed the probabilistic method guarantees the existence of certain configurations by showing that the probability of their occurrence is positive. For any i and j, the existence of a graph all of whose circuits have at least i edges but whose chromatic number is at least j can be demonstrated by this technique. *See* GRAPH THEORY.

Ferrer's diagrams. Associated with an integer partition $n = a_1 + \cdots + a_k$ is an array of dots called its Ferrer's diagram, whose ith row contains a_i dots. Thus $4 = 3 + 1$ gives rise to the diagram $\begin{smallmatrix}\bullet&\bullet&\bullet\\ \bullet&&\end{smallmatrix}$.

A reflection of the diagram through its main diagonal interchanges rows and columns, converting $4 = 3 + 1$ into its dual partition $4 = 2 + 1 + 1$. This correspondence proves a number of identities, such as the number of partitions of n into k parts equals the number of partitions of n which have greatest part k.

Möbius inversion. Another useful counting technique exploits the relationship between a pair of functions f and g defined on a lattice L. Suppose that g satisfies Eq. (12), and that the values of g are known;

$$g(s) = \sum_{s \le t} f(t) \tag{12}$$

then the values of f can be found if the calculation in Eq. (12) can be "inverted" to define f in terms of g. The principle of Möbius inversion states that there is a function $\mu(s, t)$, called the Möbius function of the lattice, defined for pairs of elements (s,t) in L and which depends only on the lattice, such that Eq. (13) holds.

$$f(s) = \sum_{s \le t} \mu(s,t)g(t) \tag{13}$$

For example, assume that the lattice in question is the boolean algebra on the two points H and P. Let the subset $\{H\}$ stand for the property that a specific card in a deck is a heart, $\{P\}$ for the property that it is a picture card, $\{P,H\}$ for the property that it is

both a heart and a picture card, and ϕ, the zero of the lattice, for the property that it is neither. As t varies over these four subsets, let $f(t)$ denote the number of cards in the deck which have precisely the property (or properties) t (and no others), and let $g(t)$ denote the number which have at least the property (or properties) t. Then clearly $g(\phi) = 52$, $g(\{H\}) = 13$, $g(\{P\}) = 12$, and $g(\{P,H\}) = 3$, and g and f stand in the relation given by Eq. (12). For boolean algebras the Möbius function is given as follows: $\mu(s,t) = 0$ if s is not a subset of t, and $\mu(s,t) = (-1)^p$ otherwise, where p is the number of elements in t but not in s. [Thus, for all $s \subseteq t$, $\mu(s,t) = 1$ if the difference in sizes of s and t is even and $\mu(s,t) = -4$ if it is odd.] In our example, $f(\phi) = g(\phi) - g(\{H\}) - g(\{P\}) + g(\{P,H\}) = 30$ cards which are neither hearts nor picture cards. *See* BOOLEAN ALGEBRA.

For a general boolean algebra, the Möbius inversion formula for $f(\phi)$ is more often called the principle of inclusion-exclusion, or sieve formula, because at various stages of the summation the correct number is alternately overcounted and then undercounted. Assume S has $n = g(0)$ elements each of which may have some of the properties $P_1, \ldots, P_r$. Further, let $g(i_1,i_2,\ldots,i_k)$ be the number of elements with (at least) properties $P_{i1}, P_{i2}, \ldots, P_{ik}$. Again Eq. (12) is satisfied, so by Eq. (13), one arrives at formula (14) for $f(0) = f(\phi)$.

$$f(0) = n - \sum_{1 \le i_1 \le n} g(i_1) + \cdots + (-1)^p \sum_{1 \le i_1 < \cdots < i_p \le n} g(i_1, \ldots, i_p) + \cdots \quad (14)$$

Derangements. The derangement numbers D_n count complete permutations (without repetition) subject to the additional constraint that every element be "wrongly labeled." That is, if $S = \{1, \ldots, n\}$, then D_n is the number of ways to order S into sequences $a_1, \ldots, a_n$ such that $a_i \neq i$ for $i = 1, \ldots, n$. If the property P_i of a permutation is that $a_i = i$, then the sieve formula may be applied to prove Eq. (15). In probabilistic terms, Eq. (15) implies that no matter what the size of the decks is, the probability is approximately $1/e$ that two identical and randomly shuffled decks of cards will never have the same card in the same place.

$$\begin{aligned} D_n &= P(n, n) - C(n, 1)P(n-1, n-1) \\ &\quad + C(n, 2)P(n-2, n-2) - \cdots \\ &= n!\left(1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n}\right) \approx \frac{n!}{e} \end{aligned} \quad (15)$$

Gaussian coefficients. The Möbius function has been computed for other lattices which arise in enumeration, such as the lattice of all positive divisors of a given integer n (with elements ordered by divisibility), a lattice of integer partitions, the lattice of partitions of an n-element set (ordered by refinement: $s \le t$ if every block of s is contained in a block of t), and the lattice of all subspaces of a vector space of dimension n over a finite field with q elements. The number of subspaces of dimension k in such a vector space is expressed in formula (16) for gaussian

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{(q^n - 1)(q^{n-1} - 1)\cdots(q^{n-k+1} - 1)}{(q^k - 1)(q^{k-1} - 1)\cdots(q - 1)} \quad (16)$$

coefficients. This apparent rational function of q is, in fact, a polynomial which is the generating function for the numbers $p_{m,k}(n)$ of integer partitions of n into at most m parts whose largest part is at most k, Eq. (17). An easy fact, established from the Ferrer's

$$\begin{bmatrix} m+k \\ k \end{bmatrix}_q = \sum_{q=0}^{mk} p_{m,k}(n) q^n \quad (17)$$

diagram, is that $p_{m,k}(N) = p_{m,k}(mk - n)$. A much harder result is that the sequence of coefficients is unimodal: $p_{m,k}(n + 1) \ge p_{m,k}(n)$ for all $n < mk/2$. *See* LINEAR ALGEBRA.

The lattices for boolean algebras, subspaces, and set-partitions show striking similarities [such as the similarity between Eqs. (2) and (16)], which are studied in the general theory of geometric lattices or combinatorial geometries. For example, the methods applied to set partitions to study the four-color theorem of graph theory, when applied to the lattice of subspaces, provide a way to attack the problem of finding efficient error-correcting codes. *See* INFORMATION THEORY.

Magic squares. A (generalized) magic square is an n by n matrix of (not necessarily distinct) nonnegative integers all of whose row and column sums are equal to a prescribed number x. Thus $x = 15$ in the 3×3 magic square given earlier. The number of such magic squares is 1 when $n = 1$, $x + 1$ when $n = 2$, and in general is a polynomial of the form $a_m x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + 1$, where $m = (n-1)^2$. Although the coefficients a_i are not given explicitly in this formula, for each n they may be computed by the method of undetermined coefficients, so that a computer search of magic squares for small values of x will also determine the number of such squares for all values of x. *See* MATRIX THEORY.

Pólya counting formula. If the number of ways to paint the faces of a cube with red and blue is to be computed, formula (3) would seem to give $2^6 = 64$; however, it is reasonable to assume that all six arrangements in which exactly one face is blue are identical, since for all such arrangements a rotation of the cube will make the top face blue. Under the condition that two coloring patterns are to be assumed equivalent if one can be transformed into the other by a rotation, the number of configurations is 10. This number can be derived by exhausting all cases or by the Pólya counting formula. The Pólya formula counts the number of functions from a set D (in the example, the faces of the cube) to a set R (the colors red and blue), with two functions f and g assumed to be the same if some element of a fixed group G of (complete) permutations of D (in this case, the 24 rotations of a cube) takes f into g. Actually, the

Pólya formula provides a way to compute the generating function, with variables $x_1, \ldots, x_k$ corresponding to the elements of R, for which the coefficient of $x_1^{a_1} \cdots x_k^{a_k}$ is the number of functions which use each x_i value exactly a_i times. (For the painted cube, if x_i = red and x_2 = blue, then the generating function is given by $x_1^6 + x_1^5x_2 + 2x_1^4x_2^2 + 2x_1^3x_2^3 + 2x_1^2x_2^4 + x_1x_2^5 + x_2^6$.) The proof of the Pólya theorem depends on a special case proved earlier by W. Burnside: if G is a group consisting of m (complete) permutations of S, then the number of inequivalent permutations of S is

$$\frac{1}{m}\sum_{\sigma \in G} f(\sigma)$$

Here two permutations π_1 and π_2 are regarded as equivalent if there is some σ in G such that the relabeling that σ effects on the sequence π_1 gives the sequence π_2; and for each σ in G, $f(\sigma)$ denotes the number of permutations of S which do not change when relabeled by σ. *See* GROUP THEORY.

Application of Lefschetz theorem. A wide range of techniques is employed in enumeration theory. For example, the hard Lefschetz theorem from algebraic geometry was used to show that if A is any set of n distinct positive real numbers, and if for any subset B of A, $s(B)$ is the sum of its members, then $s(B)$ is constant on at most $g(n)$ subsets B, where $g(n) = \overline{p}_n([(n+1)(n)/4])$, the number of partitions of $[(n+1)(n)/4]$ into distinct parts whose largest part is at most n. Thus $g(n)$ is the middle coefficient of the polynomial $(1+x)(1+x^2)(1+x^3)\cdots(1+x^n)$ and $g(n)$ is achieved by letting A be the arithmetic progression $\{1,2,\ldots,n\}$ and setting the sum equal to $[(n+1)(n)/4]$.

Properties of Arrangements

An arrangement is primarily a family of subsets of a set S. An alternate way of formulating this concept is as an incidence system, also called a relation, which specifies when a particular element s of S is in a subset S_i of the family, for example, by means of the incidence matrix $M = [m_{ij}]$ of the family: the rows of the incidence matrix are indexed by the elements $s_1, \ldots, s_n$ of the set S, the columns of the matrix are indexed by the subsets $S_1, \ldots, S_m$ in the family, $m_{ij} = 1$ if s_i is in S_j, and $m_{ij} = 0$ otherwise. For example, if $S = \{s_1,s_2,s_3,s_4\}$, $S_1 = \{s_1, s_2\}$, $S_2 = S_4 = \{s_1,s_3,s_4\}$, and $S_3 = \{s_2\}$, the incidence of matrix of this arrangement is

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

Systems of distinct representatives. If $n = m$ in an incidence system, a system of distinct representatives for the sets is a permutation $a_1, \ldots, a_n$ of S such that a_i is in S_i for $i = 1, 2, \ldots, n$. (In a congressional committee S, where each S_j might denote a particular cause or coalition, it might be desirable to find a system of distinct representatives so that each cause or coalition would have a different proponent.) In the above example, s_1,s_2,s_3,s_4 is such a system.

Marriage theorem. A system of distinct representations certainly cannot exist unless any k of the subsets S_i together contain at least k distinct elements, and in fact it can be shown that this condition is also sufficient. This theorem is whimsically called the marriage theorem, since if S consists of a collection of n boys and each subset S_i consists of the boys acquainted with a particular girl g_i, then the conditions of the theorem guarantee that each girl could marry a boy of her acquaintance.

Algorithms for finding systems. Algorithms for computers exist which find systems of distinct representatives with reasonable efficiency: as n increases, the time a computer needs to find a system increases as the square of n. *See* DIGITAL COMPUTER.

Permanent. To enumerate the systems of distinct representatives, it is enough to compute the permanent of the incidence matrix defined by $\Sigma m_{1,a_1} m_{2,a_2} \cdots m_{n,a_n}$, where the sum is taken over all $n!$ permutations $(a_1, a_2, \ldots, a_n)$ of the set of column indices.

Maximal systems. When a complete system of distinct representatives does not exist, the size of a largest possible system is equal to the smallest possible value of the function $n - k + s_{i_1 \ldots, i_k}$ for all possible combinations $S_{i_1}, \ldots, S_{i_k}$ of subsets in the family, where $s_{i_1, \ldots, i_k}$ is the number of elements in $S_{i_1} \cup \cdots \cup S_{i_k}$. A related theorem asserts that the minimum number of blocks into which a partially ordered set may be partitioned subject to the constraint that any pair of elements x and y in the same block be related ($x \le y$ or $y \le x$) is equal to the maximum possible size of a set S of pairwise unrelated elements. It is always possible to find such a set S with the further property that any (complete) permutation π which preserves the partial order [if $x \le y$, then $\pi(x) \le \pi(y)$] must take members of S to other members of S. This fact can be used to show that in a boolean algebra the maximum size of S is $C(n, [n/2])$. This number is the middle (largest) binomial coefficient and corresponds to the set S of all subsets of size $[n/2]$.

Assignment problem. Suppose it is desired to assign n workers to do n jobs, and entry a_{ij} of a matrix A measures how well the ith worker does the jth job. An assignment would be a permutation $j_1, \ldots, j_n$ of the jobs so that the ith worker performs the job j_i. The assignment problem requires that one find an assignment maximizing the total utility T, given by Eq. (18).

$$T = \sum_{i=1}^{n} a_{ij_i} \tag{18}$$

The problem may be solved by introducing auxiliary row and column numbers $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ subject to the condition $u_i + v_j \ge a_{ij}$ for all i and j. Then Eq. (19) is valid for any assignment,

$$T = \sum_{i=1}^{n} a_{ij_i} \le \sum_{i=1}^{n} u_i + \sum_{j=1}^{n} v_j = R \tag{19}$$

since in an assignment each row and column is used exactly once. The marriage theorem shows that the maximum utility M of T is equal to the minimum m of R. Further, an algorithm is known which systematically changes the u's and v's to reduce R, arriving at numbers u_i and v_j for which the set of a_{ij}'s such that $a_{ij} = u_i + v_j$ admits a complete system of distinct representatives. For an assignment derived from such a system, $T = R$ and so M is achieved for this T. This method is practical in that the algorithm leads directly to a solution; and since the number of possible assignments is usually very large (for example if $n = 25$, this number is $25! > 10^{25}$), without some guidance there is no practical way to search for the optimal assignment or to recognize it when it is found. Another application of the marriage theorem is to prove topological fixed-point theorems. *See* TOPOLOGY.

Doubly stochastic matrices. A doubly stochastic matrix, a matrix of nonnegative real numbers such that every row sum and every column sum is equal to one, is a real-number analog of a magic square. The set of all n by n doubly stochastic matrices is convex in the sense that if $M_1, \ldots, M_k$ are all doubly stochastic and $a_1, \ldots, a_k$ are nonnegative numbers whose sum is one, then the convex combination

$$\sum_{i=1}^{k} a_i M_i$$

is also doubly stochastic. These matrices in fact form a convex polytope (a generalization of the notion of convex polygon) in euclidean space $\mathbf{R}^{n^2}$, in that their entries are bounded and there are a finite number of vertices (a minimal set of matrices such that every doubly stochastic matrix is a convex combination of elements of this set). The marriage theorem can be used to prove that the $n!$ permutation matrices with exactly one 1 in each row and column are the vertices of the convex polytope of doubly stochastic matrices.

By a theorem in linear programming, linear functions are optimized on the vertices of a convex polytope. For the assignment problem this means that for a more general assignment in which the ith worker spends a fraction m_{ij} of the work time on the jth job and each job is always being worked on (so that the matrix $M = [m_{ij}]$ is doubly stochastic), a maximum utility is obtained when each worker spends all the work time at one job. *See* LINEAR PROGRAMMING.

Upper-bound problem. In a typical problem in linear programming, polytopes in $\mathbf{R}^m$ occur which are defined by n linear constraints, such as those given by (20), where the x_j's are the variables (corresponding to vectors in $\mathbf{R}^m$. Algorithms to maximize linear functions on these polytopes usually involve proceeding from one vertex to another at which the function is greater. [An algorithm was given in 1979 using different (nonlinear) techniques, which guarantees that the number of steps needed to solve any such linear program is no more than a polynomial function of mn, but for most practical applications it has been found that a vertex-to-vertex method such as the simplex algorithm gives a faster answer.] It is important when allocating computer time to place an upper bound on the number of vertices. In 1970 it was shown that the number of vertices of a polytope in $\mathbf{R}^m$ defined by n linear constraints is at most $M(m,n)$, given by Eq. (21).

$$\sum_{j=1}^{m} a_{ij} x_j \le b_i \qquad i = 1, \ldots, n \tag{20}$$

$$M(m, n) = \binom{n - \left[\dfrac{m+1}{2}\right]}{n-m} + \binom{n - \left[\dfrac{m+2}{2}\right]}{n-m} \tag{21}$$

Existence and Construction

In the 1920s R. A. Fisher noted that "the design and analysis of statistical experiments requires the construction of orthogonal Latin squares and balanced incomplete block design." *See* STATISTICS.

Orthogonal Latin squares. A Latin square is a square n by n matrix with entries from the set $N = \{0, 1, \ldots, n-1\}$ so that each number occurs exactly once in each row and exactly once in each column. Two Latin squares are said to be orthogonal if when they are superposed the n^2 cells contain each of the n^2 pairs of numbers from N exactly once. For $n = 3$ one has the superposed orthogonal squares

00	11	22
12	20	01
21	02	10

where the first digits form the first square, and the second digits the second square. This square may be used to design an agricultural experiment which tests the interaction of three varieties of grain with three types of fertilizers. A field is divided into nine plots in each of which the choice of grain is made according to the first digit and the choice of fertilizer according to the second digit. The Latin squares assure the even distribution of the varieties of grain and fertilizer in both directions, so that effects such as the variation of the soil are mimimized, and the orthogonality allows the experimenter to try each fertilizer with each variety of grain.

For $n \le 10$ each of these arrays can be viewed as a particular type of magic square whose entries are distinct two-digit numbers and such that each row and column sum is $11n(n-1)/2$ (33 for $n = 3$ and 495 for $n = 10$).

If n is odd, it is easy to construct two orthogonal squares A and B. One takes $A = [a_{ij}]$ and and $B = [b_{ij}]$ where $a_{ij} = i + j$, $b_{ij} = i + 2j$, reducing these sums by n or $2n$ if necessary to put them in the range $0, \ldots, n-1$. If n is a multiple of 4, it is not much more difficult to make a construction. But for n of the form $4m + 2$ it is considerably harder. It is clearly impossible when $n = 2$, and, by an exhaustive trial, G. Tarry showed it to be also impossible for $n = 6$. Euler had

Fig. 1. Two orthogonal Latin squares for $n = 10$. (*Courtesy of Thomas Brylawski; Karl Petersen, photographer*)

conjectured in 1782 that no pair of orthogonal Latin squares existed for any n of the form $4m + 2$, but in 1959 orthogonal Latin squares were found for every $4m + 2$ greater than or equal to 10 (**Fig. 1**).

Block designs. T. Kirkman posed the following puzzle in the 1850s: Is it possible for 15 schoolchildren to go for walks in 5 groups of 3 every afternoon, so that in seven afternoons every child shall have walked with every other child? This arrangement is indeed possible and the following is a solution, the girls being represented by letters $a, \ldots, o$:

Sun.	Mon.	Tue.	Wed.	Thur.	Fri.	Sat.
abi	*acj*	*adk*	*ael*	*afm*	*agn*	*aho*
cdf	*deg*	*efh*	*fgb*	*ghc*	*hbd*	*bce*
gjo	*hki*	*blj*	*cmk*	*dnl*	*eom*	*fin*
ekn	*flo*	*gmi*	*hnj*	*bok*	*cil*	*djm*
hlm	*bmn*	*cno*	*doi*	*eij*	*fjk*	*glk*

The cited design is a special case of a balanced incomplete block design or (b, v, r, k, λ)-design: an arrangement of v elements or treatments x_i, $i = 1, \ldots, v$, into b subsets or blocks B_j, $j = 1, \ldots, b$, so that each B_j contains exactly k distinct elements, each element occurs in r blocks, and every combination of two elements $x_u x_v$ occurs together in exactly λ blocks.

Counting the total number of plots (incidences of x_i in B_j or symbols in the above arrangement) in two ways—first summing over the blocks and then summing over the treatments, the constraint (22) is obtained. By considering all two-element subsets which contain x, constraint (23) is obtained. The two technical constraints (24) imply that the blocks are incomplete in that none of the blocks contains all or all but one of the treatments and none consists of a single treatment.

$$bk = vr \tag{22}$$

$$r(k - 1) = \lambda(v - 1) \tag{23}$$

$$2 \rangle k \rangle v - 1, \lambda \rangle 0 \tag{24}$$

Steiner triple systems. When $\lambda = 1$ and $k = 3$, such (b, v, r, k, λ)-designs are known as Steiner triple systems, and in this case constraints (22) and (23) imply that v is of the form $6m + 1$ or $6m + 3$. Kirkman himself proved that triple systems exist for all such v.

The schoolchild arrangement has the additional property that the blocks themselves are partitioned into r families called parallel classes of v/k blocks each such that every element occurs in exactly one block of each parallel class. Such designs are termed resolvable. A resolvable Steiner triple system is called a Kirkman triple system, and clearly in this case v must be divisible by three. In 1968 Kirkman triple systems (resolvable balanced incomplete block designs with block size three) were shown to exist for all $v = 6m + 3$.

An application of Steiner triple systems might arise in testing taste preferences among v products if every pair of products is to be compared, but each taster can efficiently taste only three products per day. A Kirkman system will, in addition, give an efficient schedule in which each product is used exactly once each day.

Fisher's inequality. Fisher's inequality states that $b \geq v$ in every design. If $b = v$, then necessarily $k = r$ and the design is called a symmetric (v, k, λ)-design. It is a property of these designs that every two blocks intersect in λ elements, and hence the transpose of the incidence matrix of a (v, k, λ)-design is itself a (v, k, λ)-design. Elementary matrix theory may be used to show that if v is even, then a symmetric design with parameters v, k, λ exists only if $k - \lambda$ is a perfect square; and deep results in number theory have been used to show that when v is odd, there must exist integers $c, y,$ and z not all zero such that Eq. (25) holds.

$$x^2 = (k - \lambda)y^2 + (-1)^{(v-1)/2}\lambda z^2 \tag{25}$$

See NUMBER THEORY.

Existence for fixed k. If k is fixed, then, for all but finitely many sets of parameters b, v, r, k, λ satisfying Eqs. (22), (23), and (24), there exists a block design having those parameters. When $k \leq 5$, all sets of parameters for which a block design exists have been determined. (A design is simple when $\lambda = 1$.)

Projective planes. When a symmetric design has $\lambda = 1$, condition (23) shows that if $k = n + 1$, then $v = n^2 + n + 1$. In analogy with the points (treatments) and lines (blocks) of a plane in projective geometry, such configurations are called projective planes. In this case, condition (25) states that if the parameter n, called the order of the plane, is of the form $4m + 1$ or $4m + 2$, then integers u and v exist with $n = u^2 + v^2$. Thus an infinite number of possible values of n are excluded, beginning with 6, 14, and 21. Whenever n is a power of a prime (and thus the size of a finite field), a plane exists, namely, the projective plane coordinatized over the field (but these are not the only planes which exist for

prime-power orders). In the early part of the twentieth century, planes were constructed over many algebraic systems in which linear equations could be solved, but which were not associative. Planes coordinatized over such systems are different from those coordinatized over fields. The smallest of these planes has order 9 and thus has 91 points. *See* ALGEBRA; PROJECTIVE GEOMETRY.

Non-Desarguesian planes. The theorem of G. Desargues (which states that if corresponding vertices of two triangles lie on three concurrent lines, then the three points of intersection of respective sides of those triangles will lie on a line) holds in general only for those planes which arise from fields, and hence the term non-Desarguesian plane is used for the others. No (necessarily non-Desarguesian) projective plane has been found for any n that is not a prime power [and no n has been excluded except those not satisfying condition (19)]. Whether a projective plane exists for an n that is not a power of a prime is an important unsolved problem in existence theory and projective geometry. Orthogonal Latin squares and projective planes are related in that the existence of a projective plane of order n is equivalent to the existence of a set of $(n-1)$ n by n Latin squares every pair of which is orthogonal. This would mean that for the first unresolved case, $n = 10$, nine pairwise orthogonal Latin squares must be found to produce a projective plane of order 10. No one has found more than two such squares.

Friendship theorem. Only for trivial projective planes can all the points be paired with all the lines in such a way that the incidence system is dualized, and each line is paired with a point not on that line. Stated differently, this is the friendship theorem: Among n people, if every pair of people has exactly one common friend, then there is someone who knows everyone else.

t-Designs. The concept of a block design in which every pair of elements is in the same number of blocks can be generalized to t-designs with parameters $t-(v, k, \lambda)$, in which each t-element subset of a v-element set is in a fixed number λ of k-element blocks. Every t-design is an s-design for all s such that $0 \le s \le t$ where the λ_s are given by Eq. (26), and

$$\lambda_s = \frac{\lambda P(v-s, t-s)}{P(k-s, t-s)} \tag{26}$$

these numbers must all be integers. Formulas (22) and (23) are special cases of Eq. (26), with $\lambda_0 = b$, the number of blocks, and $\lambda_1 = r$, the replication number.

Designs are only known for relatively small values of t. For $t = 3$, there are the inversive planes which, like finite projective planes, combine classical geometry and finite fields. Let P be the points of a finite affine plane coordinatized by a field F with q elements. These points are in correspondence with the elements of an extension field F' with q^2 elements, where each point (r,s) is paired with the element $r + sx$, with x the root of an irreducible quadratic equation (like $\sqrt{-1}$ over the real numbers). An ideal point at infinity, ∞, is added giving $v = q^2 + 1$ points. The blocks are then the images of the set B_0 consisting of the line $\{(s,0)$: s element of $F\}$ along with ∞ under all linear fractional transformations, given by Eq. (27). Here ∞ acts like 1/0 so that Eq. (28)

$$y \to \frac{ay+b}{cy+d} \tag{27}$$
$$(a,b,c,d \text{ elements of } F',\ ad = bc)$$

$$\frac{a^\infty + b}{c^\infty + d} = \frac{a}{c} \tag{28}$$

holds. The blocks (all with $q + 1$ elements) are then called circles (with the original lines in the affine plane corresponding to circles through ∞). Circles have the property that every triple of points determines a unique circle ($\lambda = 1$), and further that if C is a circle which contains p but not q, there is a unique circle which contains q and is "tangent" to C at p.

The construction of t-designs has been closely related to the study of multiply transitive groups. The Mathieu groups give two designs for $t = 5$. One has parameters $5-(12,6,1)$, and the other has parameters $5-(24,8,1)$. In the first case, the blocks which contain a fixed three-element subset of points form a design isomorphic to the affine plane of order three, and in the second the blocks through three fixed points form a projective plane of order four.

At present no simple designs are known for $t > 5$. Steiner quadruple systems [designs with parameters $3-(v,4,1)$] have been constructed for all v congruent to 2 or 4 modulo 6.

Error-correcting codes. The following matrix is the incidence system of a projective plane of order 2 (and symmetric Steiner triple system):

$$\begin{matrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{matrix}$$

Since $k = 3$ and $\lambda = 1$, each row has exactly three 1's and any two rows overlap in only one of these columns. Thus, any two rows differ in four of their entries. This matrix is an example of a binary block code of distance $m = 4$ and length $n = 7$: a set of codewords, each consisting of a sequence of n binary digits (zeros and ones), such that any two codewords differ in at least m places. If seven messages are coded by rows of the above matrix and one is transmitted to a receiver, the original message can be decoded—even if during transmission one of the 1's were changed to a 0 or conversely—by choosing the codeword closest to the received sequence. For example, if 1001011 were received, the last codeword was transmitted. On the other hand, if two digits were changed, resulting for example in a received message 1000110, the receiver could not tell whether the intended message corresponded to the first codeword or the second. Such a code will then

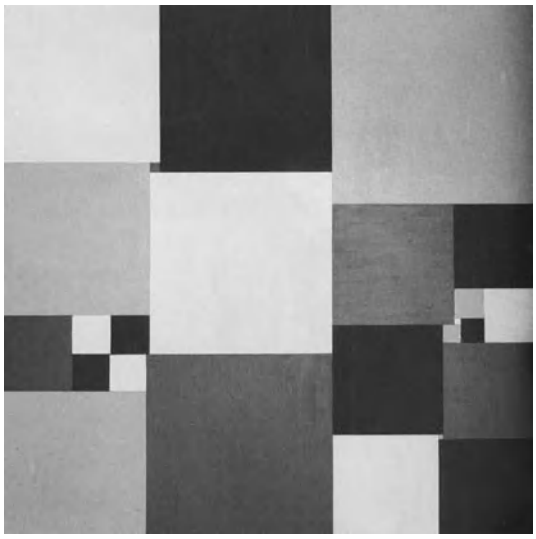

Fig. 2. Subdivision of a square into 26 smaller squares; no two of which are the same size. (*Courtesy of Thomas Brylawski*; *Karl Petersen*, *photographer*)

"correct" one transmission error and decide if an intended codeword with up to three such errors has been correctly transmitted.

In general, a block code will correct up to $[(m-1)/2]$ errors. The fundamental problem in coding theory is to construct codes with a large number of words and large distance but small length. In the example, the size of the code may be increased without decreasing the distance by adding a 1 to the end of all seven codewords, adjoining a new word of eight 1's, and then forming the eight additional words constructed from each of the others by interchanging all 0's and 1's.

Subdivision of square. An elegant combinatorial application of the theories of electrical networks and three dimensional convex polytopes led in the 1940s to the construction of a square subdivided into n smaller squares no two of which are the same size (**Fig. 2**, where $n = 26$). The least number of unequal squares into which a square can be divided is 21, and that, up to symmetry, there is only the following subdivision. A square of length 112 is partitioned into squares of length 50, 35, 27, 8, 19, 15, 17, 11, 6, 24, 29, 25, 9, 2, 7, 18, 16, 42, 4, 37, and 33, respectively, where each square enters the big square as far north as possible and then as far west as possible.

Pigeonhole principle. The pigeonhole (or Dirichlet drawer) principle states that: if a very large set of elements is partitioned into a small number of blocks, then at least one block contains a rather large number of elements. From this result it follows, for example, that any permutation of the numbers from 1 to $mn + 1$ must contain either an increasing subsequence of length $m + 1$ or a decreasing subsequence of length $n + 1$. To see this, associate with each number in the permutation P the pair of numbers (i,d), where i (respectively d) is the length of the longest increasing (respectively decreasing) subsequence in which it is the first element. Then no two numbers in P are associated with the same pair (since if p precedes q, P initiates a longer increasing subsequence if $p < q$ and a longer decreasing subsequence if $p > q$), so that not every number can be associated with one of the mn pairs (i, d), $i = 1, \ldots, m$; $d = 1, \ldots, n$. One can do no better than $mn + 1$, since if the numbers from 1 to mn are arranged sequentially in an n by m array and read off from left to right starting at the bottom, then the longest increasing subsequence has length m and the longest decreasing subsequence has length n. This construction belongs to the field of extremal combinatorial theory, a branch of combinatorics which constructs counterexamples of size n for theorems which hold for all $k \geq n + 1$.

Ramsey's theorem. Ramsey's theorem, proved in the 1930s, generalizes the pigeonhole principle as follows: For any parameters $(r; r_1, \ldots, r_k)$ with each $r_i \geq r$, there exists a number N such that for all $n \geq N$, if S is an n-element set all

$$\binom{n}{r}$$

whose r-element subsets are partitioned into blocks $A_1, \ldots, A_k$, then there is some r_i-element subset S' of S such that all

$$\binom{r_i}{r}$$

of the r-element subsets of S' are in the block A_i. For a given set of parameters, one then asks for the Ramsey number, or smallest N for which the theorem is true. When $r = 1$, by the pigeonhole principle $N = r_1 + \cdots + r_k - k + 1$. For the parameters (2;3,3), the theorem can be interpreted as follows: When the edges and diagonals of a regular polygon with a sufficiently large number N of vertices are all colored red or blue, there are three vertices of the polygon which are also the vertices of a triangle all of whose edges are the same color. One can always find a monochromatic triangle in a hexagon, and the extremal configuration consisting of a pentagon whose edges are red and whose five diagonals are blue shows that the Ramsey number for these parameters is 6.

Analogs of Ramsey's theorem have been proved for graphs, vector spaces, configurations in the Euclidean plane, and (with infinite cardinals as parameters) in set theory. However, very few Ramsey numbers have been computed. For example, when the edges and diagonals of a polygon are colored with four colors, the smallest N which guarantees a monochromatic triangle [that is, for (2;3,3,3,3)] is known only to be between 51 and 63.

Many asymptotic results are known and have interpretations in graph theory. For example, the Ramsey number $(2;3,t)$ is known to be at most $ct^2/\log t$, where c is an absolute constant not depending on t. This means that in any graph with $ct^2/\log t$ vertices and no triangles, there is a subset of t vertices, no two of which are connected by an edge. Thomas Brylawski

Bibliography. I. Anderson, *A First Course in Combinatorial Mathematics*, 2d ed., 1989; K. P. Bogart, *Introductory Combinatorics*, 3d ed., 1999; R. A. Brualdi, *Introductory Combinatorics*, 3d ed.,

1998; V. W. Bryant, *Aspects of Combinatorics*, 1993; P. J. Cameron, *Combinatorics*, 1994; C. D. Godsil, *Algebraic Combinatorics*, 1993; A. Tucker, *Applied Combinatorics*, 3d ed., 1994; J. H. Van Lint and R. M. Wilson, *A Course of Combinatorics*, 1994.

Combining volumes, law of

The principle that when gases take part in chemical reactions the volumes of the reacting gases and those of the products, if gaseous, are in the ratio of small whole numbers, provided that all measurements are made at the same temperature and pressure. The law is illustrated by the following reactions:

1. One volume of chlorine and one volume of hydrogen combines to give two volumes of hydrogen chloride.
2. Two volumes of hydrogen and one volume of oxygen combine to give two volumes of steam.
3. One volume of ammonia and one volume of hydrogen chloride combine to give solid ammonium chloride.
4. One volume of oxygen when heated with solid carbon gives one volume of carbon dioxide.

It should be noted that the law applies to all reactions in which gases take part, even though solids or liquids are also reactants or products.

The law of combining volumes was put forward on the basis of experimental evidence, and was first explained by Avogadro's hypothesis that equal volumes of all gases and vapors under the same conditions of temperature and pressure contain identical numbers of molecules. *See* AVOGADRO'S NUMBER.

The law of combining volumes is similar to the other gas laws in that it is strictly true only for an ideal gas, though most gases obey it closely at room temperatures and atmospheric pressure. Under high pressures used in many large-scale industrial operations, such as the manufacture of ammonia from hydrogen and nitrogen, the law ceases to be even approximately true. *See* GAS. Thomas C. Waddington

Combustion

The burning of any substance, in gaseous, liquid, or solid form. In its broad definition, combustion includes fast exothermic chemical reactions, generally in the gas phase but not excluding the reaction of solid carbon with a gaseous oxidant. Flames represent combustion reactions that can propagate through space at subsonic velocity and are accompanied by the emission of light. The flame is the result of complex interactions of chemical and physical processes whose quantitative description must draw on a wide range of disciplines, such as chemistry, thermodynamics, fluid dynamics, and molecular physics. In the course of the chemical reaction, energy is released in the form of heat, and atoms and free radicals, all highly reactive intermediates of the combustion reactions, are generated. *See* FLAME; FREE RADICAL; REACTIVE INTERMEDIATES.

Processes. The physical processes involved in combustion are primarily transport processes: transport of mass and energy and, in systems with flow of the reactants, transport of momentum. The reactants in the chemical reaction are normally a fuel and an oxidant. In practical combustion systems the chemical reactions of the major chemical species, carbon and hydrogen in the fuel and oxygen in the air, are fast at the prevailing high temperatures (greater than 1200 K or 1700°F) because the reaction rates increase exponentially with temperature. In contrast, the rates of the transport processes exhibit much smaller dependence on temperature and are, therefore, lower than those of the chemical reactions. Thus, in most practical flames the rate of evolution of the main combustion products, carbon dioxide and water, and the accompanying heat release depends on the rates at which the reactants are mixed and heat is being transferred from the flame to the fresh fuel-oxidant mixture injected into the flame. However, this generalization cannot be extended to the production and destruction of minor species in the flame, including those of trace concentrations of air pollutants such as nitrogen oxides, polycyclic aromatic hydrocarbons, soot, carbon monoxide, and submicrometer-size inorganic particulate matter. *See* TRANSPORT PROCESSES.

Nitrogen oxides are formed by the capture of atmospheric nitrogen and the oxidation of organically bound nitrogen in fossil fuels in the flame. Both their oxidative formation, and their destruction in high-temperature fuel-rich flame zones are controlled by the chemistry of the interconversions of the nitrogen species in the flame. The oxidation reaction of the trace combustible species may slow down in the tail end of the flame because of the reduced temperature, or it may get quenched in the flow close to cooled walls of the combustion chamber. Hence, the emission of these trace species strongly depends on the rate of their oxidative destruction.

The fine inorganic particulate matter originates from the vaporization of a small fraction of fuel ash, mainly metals in coal and oil ash, and its subsequent condensation to form an aerosol as the combustion products cool down. The chemical transformation of the ash in the course of combustion (for example, the formation of suboxides of reduced melting temperature) has a strong influence on the aerosol formation and emission. *See* AEROSOL.

Types. When the fuel and oxidant are both gaseous, the combustion is called homogeneous. The combustion reactions will commence as the reactants are heated beyond a critical (ignition) temperature. Such a temperature is dependent on the types of the fuel and the oxidant and their mixture ratio, and also on the physical characteristics of the surroundings.

Liquid fuels are usually prepared for combustion by atomization; they are injected into the flame in the form of a spray of fine droplets. The drops vaporize, and the fuel vapor burns in a manner similar to a fuel gas. However, there are two phases present, liquid

and gas, and the combustion is called heterogeneous. *See* ATOMIZATION.

The heavy residual fuel oils used in large industrial and utility boilers are not fully vaporizable, and a small fraction (about 10%) of the fuel oil burns in the form of small coke particles in the tail of the flame. Because of the higher carbon-to-hydrogen ratio of petroleum fuels compared to natural gas (the latter being mainly methane), oil flames, with higher concentrations of soot, tend to be more luminous than gas flames. The soot particles, although performing the useful function of increasing the radiating ability of the flame, ought to be completely burned within the combustion chamber to ensure a clean exhaust.

Solid fuels (coal, wood, and agricultural and municipal waste) are often burned on grates, with the fuel lying on the grate as a bed of solids and the air admitted through the openings in the grate. The grate can be stationary or moving, the feed being correspondingly periodic or continuous.

Coal, the most abundant fossil fuel, is burned predominantly in large industrial and utility boilers in pulverized form. After fine grinding, the pulverized coal is pneumatically transported by a fraction of the combustion air or an air–flue-gas mixture to the burner through which it is injected into the combustion chamber. Boiler combustion chambers equipped for burning pulverized coal are large so as to provide ample time for burning even the largest coal particles completely and for cooling the flyash particles below their softening temperature before they leave the combustion chamber. *See* COMBUSTION CHAMBER; FOSSIL FUEL.

Technology. In the early years of the development of modern combustion technology, after World War II, the objectives were to complete combustion in the smallest possible combustion space and with a minimum of air in excess of that which was stoichiometrically required. Environmental concerns changed this agenda, emphasizing the reduction of combustion-generated pollution. This includes emissions of the acid rain precursors, that is, nitrogen oxides and sulfur oxides, and air toxics such as polycyclic aromatic hydrocarbons, soot, and submicrometer-size inorganic particulates. Methods for reducing such emissions include fuel treatment (for example, desulfurization), combustion process modification, and postcombustion fuel-gas cleanup. Of these, the suitable modification of the combustion process is the most cost effective but also the most challenging. Notable examples of the achievements of this method are the significant reduction (up to 95%) of nitrogen oxide emissions from boilers and industrial furnaces obtained by the combination of staged admission of the fuel and the air into the combustion process, the use of flue-gas recirculation, and the use of very fuel lean premixed mixtures in industrial gas turbine combustors. In order to allay environmental concerns and to satisfy tight regulations, it is necessary to combine combustion modifications with fuel treatment and post-combustion cleanup to reduce pollutant emissions further. *See* ACID RAIN; AIR POLLUTION.

Applications. Combustion applications are wide ranging with respect to the fields in which they are used and to their thermal input, extending from a few watts for a candle to hundreds of megawatts for a utility boiler. Combustion is the major mode of fuel utilization in domestic and industrial heating, in production of steam for industrial processes and for electric power generation, in waste incineration, and in propulsion in internal combustion engines, gas turbines, or rocket engines.

Although the specific conditions of the individual applications impose different constraints on technologically acceptable solutions for the common goal of high thermal efficiency with minimal pollutant emission, basic studies of combustion science are a unifying source relevant to the development of clean combustion technologies in diverse fields of practical applications.

János M. Beér

Combustion chamber

The space at the head end of an internal combustion engine cylinder where most of the combustion takes place. *See* COMBUSTION.

Spark-ignition engine. In the spark-ignition engine, combustion is initiated in the mixture of fuel and air by an electrical discharge. The resulting reaction moves radially across the combustion space as a zone of active burning, known as the flame front. The velocity of the flame increases nearly in proportion to engine speed so that the distance the engine shaft turns during the burning process is not seriously affected by changes in speed. *See* INTERNAL COMBUSTION ENGINE; SPARK PLUG.

For high efficiency the mixture should burn as quickly as possible while the piston is near top center position (constant-volume combustion). Short burning time is achieved by locating the spark plug in a central position to minimize the distance the flame front must travel. The chamber itself is made as compact as possible for the required volume, to keep the flame paths short (**Fig. 1**). A compact chamber presents less wall area to the enclosed volume, thereby reducing the heat loss from the flame front. Because motion of the flame depends upon transferring heat from the flame to the adjacent layers of unburned mixture, the reduced heat loss increases the flame velocity. One method for increasing flame velocity is to provide small-scale turbulence in the cylinder charge, often by designing the chamber so that part of the piston head comes close to the cylinder head at top center position (Fig. 1*a*). The squish that results forces the mixture in this region into the rest of the combustion chamber at high velocity. The turbulence so produced increases the rate of heat transfer from the flame to the unburned mixture, greatly increasing the flame velocity. Other engines may use high swirl in the intake port and combustion chamber or an auxiliary or precombustion chamber to accomplish the same result (Fig. 1*e*). When a prechamber is used, combustion begins in the richer mixture around the spark plug in the prechamber.

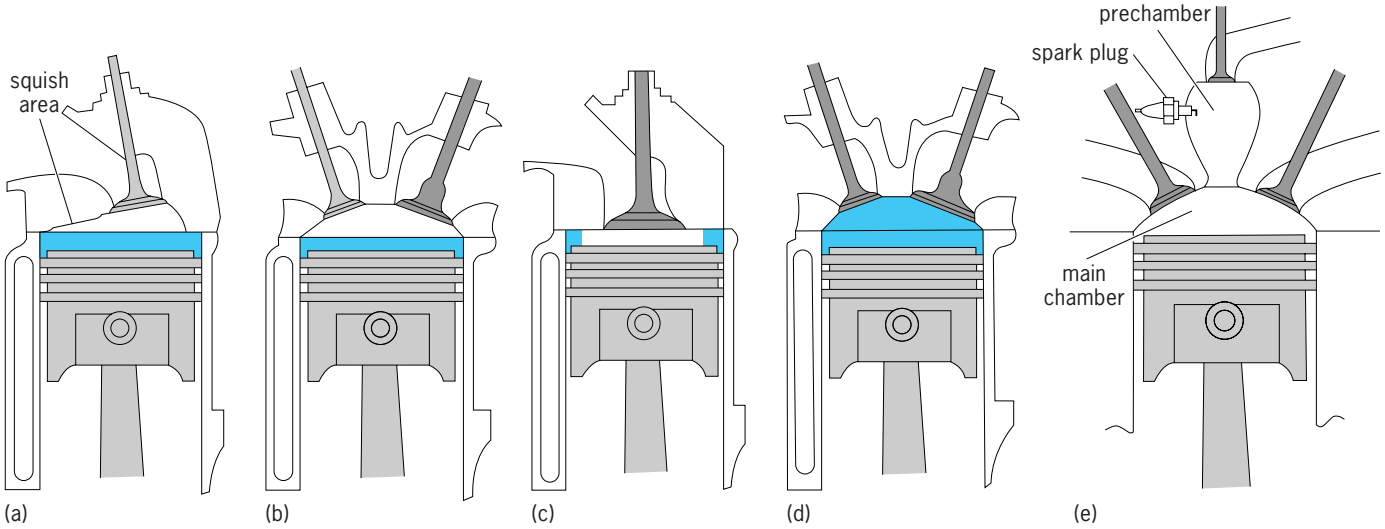

Fig. 1. Various combustion chambers used in spark-ignition engines. (*a*) Wedge. (*b*) Hemispheric (open). (*c*) Bowl-in-piston (cup). (*d*) Crescent (pent-roof). (*e*) Prechamber. (*Automotive and Technical Writing*, *Charlottesville*)

The burning mixture then enters the main chamber, which is filled with a leaner mixture, where combustion is completed. This process is known as stratified charge.

Occasionally a high burning rate, or too rapid change in burning rate, gives rise to unusual noise and vibration called engine roughness. Roughness may be reduced by using less squish or by shaping the combustion chamber to control the area of the flame front. A short burning time is helpful in eliminating knock because the last part of the charge is burned by the flame before it has time to ignite spontaneously. *See* SPARK KNOCK.

For high power output, the shape of the combustion chamber must permit the use of large valves with space around them for unrestricted gas flow. In many engine designs, three and four valves per cylinder are used to provide higher volumetric efficiency. *See* VALVE TRAIN.

Pockets or narrowed sections in the combustion chamber that trap thin layers of combustible mixture between adjacent surfaces interfere with combustion and increase the unburned fuel hydrocarbons normally present in the exhaust gases. Because unburned hydrocarbons are one of the constituents of smog-forming air pollution, the newer combustion chambers have less squish and less surface area per unit of volume. *See* AIR POLLUTION; SMOG.

Compression-ignition engine. In compression-ignition (diesel) engines, the fuel is injected late in the compression stroke into highly compressed air. Mixing must take place quickly, especially in smaller high-speed engines, if the fuel is to find oxygen and burn while the piston remains near top center. After a short delay, the injected fuel ignites from contact with the hot air in the cylinder. There is no flame front travel to limit the combustion rate. Injection may be into either an undivided combustion chamber (direct injection) or a divided combustion chamber (indirect injection) [**Fig. 2**].

During direct injection into an undivided chamber, mixing can be improved by producing a swirl in the intake air so that it is moving relative to the fuel spray (Fig. 2*a*). Swirl velocity increases with higher inlet velocities present at high engine speeds. Therefore, mixing becomes more rapid as the time available for mixing becomes less. Another type of undivided combustion chamber is formed by a narrow cavity in the piston head (Fig. 2*b*). For rapid

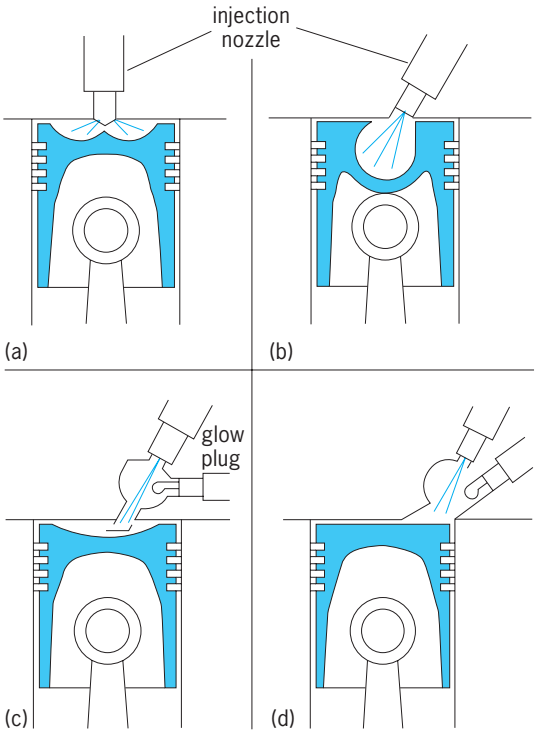

Fig. 2. Various combustion chambers used in compression-ignition (diesel) engines. (*a*) Direct injection. (*b*) MAN M system (German manufacturer). (*c*) Prechamber. (*d*) Swirl chamber. (*Robert Bosch Corp.*)

vaporization, the fuel is sprayed against the hot wall of the combustion chamber in the direction of the swirling air. Smooth combustion with relatively little exhaust smoke results.

Most high-speed diesel engines have indirect injection, in which the combustion chamber is divided into a main and a connected auxiliary chamber. Fuel is sprayed into the auxiliary chamber, where rapid mixing and combustion takes place (Fig. 2*c* and *d*). The burning mixture then issues into the main chamber, where any remaining fuel is burned.

If mixing of fuel and air is too thorough by the end of the delay period, high rates of pressure rise result, and the operation of the engine is rough and noisy. To avoid this condition, the auxiliary chamber in most compression-ignition engines operates at high temperature so that the fuel ignites soon after injection begins. This reduces the amount of fuel present and the degree of mixing at the time that ignition takes place.

High rates of pressure rise can also be reduced by keeping most of the fuel separated from the chamber air until the end of the delay period. Rapid mixing must then take place to ensure efficient burning of the fuel while the piston is near top center. *See* DIESEL ENGINE. Augustus R. Rogowski; Donald L. Anglin

Bibliography. *Bosch Automotive Handbook*, 1986; W. H. Crouse and D. L. Anglin, *Automotive Engines*, 8th ed., 1994; R. Stone, *Introduction to Internal Combustion Engines*, 3d ed., 1999.

Comet

One of the major types of objects that move in closed orbits around the Sun. Compared to the orbits of planets and asteroids, comet orbits are more eccentric and have a much greater range of inclinations to the ecliptic (the plane of the Earth's orbit). Physically, a comet is a small, solid body which is roughly 2 mi (3 km) in diameter, contains a high fraction of icy substances, and shows a complex morphology, often including the production of an extensive atmosphere and tail, as it approaches the Sun. *See* ASTEROID; PLANET.

About 10 reasonably bright comets are discovered or rediscovered each year. On the average, one per year is bright enough to be visible to the unaided eye and to generate much interest among the public as well as among comet workers.

Astronomers consider comets to be worthy of detailed study for several reasons: (1) They are intrinsically interesting, involving a large range of physical and chemical processes. (2) They are valuable tools for probing the solar wind. (3) They are considered to be remnants of the solar system's original material and, hence, prime objects to be studied for clues about the nature of the solar system in the distant past. (4) Comets may be required to explain other solar system phenomena.

Present knowledge of comets is heavily influenced by the Halley's Comet campaign in 1986, which included images of the nucleus obtained by spacecraft sent to intercept the comet, and by the extensive studies of the very large Comet Hale-Bopp in 1997. While the data obtained from these comets are invaluable, comets are believed to be highly individualistic. *See* HALLEY'S COMET.

Fig. 1. Comet West as photographed on March 9, 1976, showing the general appearance of a bright comet. The fan-shaped structure emanating from the head or coma is the dust tail, while the single straight structure is the plasma tail. The tails are some 30° in length. (*S. M. Larson, Lunar and Planetary Laboratory, University of Arizona*)

Appearance. As seen from Earth, comets are nebulous in appearance, and the tail is usually the most visually striking feature. This tail can in some cases stretch along a substantial arc in the sky (**Fig. 1**). Some fainter comets, however, have little or no tail.

The coma or head of a comet is seen as the ball of light from which the tail or tails emanate. Within the coma is the nucleus, the origin of the material in the tail and coma.

Discovery and designation. Comets are discovered by both amateur and professional astronomers. The fainter ones are often discovered by professionals on wide-field images taken for other purposes. In particular, searches for near-Earth objects and other asteroids have dramatically increased the discovery rate of fainter comets. Comets are also being discovered by space-borne solar coronagraphs. Amateurs usually carry out systematic searches of the sky using wide-field binoculars or telescopes. Discoveries are communicated to the Bureau for Astronomical Telegrams, Smithsonian Astrophysical Observatory, Cambridge, Massachusetts, and are then announced by the International Astronomical Union.

In the system for comet designations in use since January 1, 1995, objects are recorded by a letter for the half-month in which they are found; note that the letters I and Z are not used. Thus, the second comet discovered during the second half of February 1995 would be designated C/1995 D2; the C (for comet) is replaced by P if the comet is periodic. The

naming of comets carries on the tradition of naming them for their discoverers while striving for fairness and simplicity. For example, Comet C/1995 O1 (Hale-Bopp) was independently discovered by A. Hale and T. Bopp on July 23, 1995, and was the first comet discovered in that half-month period. Comets discovered prior to 1995 retain their old system designation, where the comets in a given year were assigned a letter in order of their discovery. The letter was later replaced by a roman numeral indicating order of perihelion passage. Halley's Comet at its last appearance was first called Comet 1982i, and later became Comet 1986III. Halley's Comet also illustrates the exceptional case of a comet being named after the person who computed its orbit.

Oort Cloud and comet evolution. A major step in understanding the origin of comets was taken in 1950 by J. Oort. He developed the idea that comets are in effect "stored" in a Sun-centered spherical cloud of radius approximately 50,000 AU from the Sun (1 AU is the distance from the Sun to the Earth, 9.3×10^7 mi or 1.50×10^8 km). Occasionally, gravitational perturbations by passing stars or clouds of interstellar gas send comets into the inner solar system, where they are discovered and their phenomena are observed. This process, of course, leads to the eventual destruction of the comet, because passages near the Sun cause a loss of material which is not recovered by the comet. Very rough estimates indicate that a comet loses 0.1% to 1% of its mass at each perihelion passage. Thus, a comet with radius of 0.6 mi (1 km) would lose a layer approximately in the range 1–10 ft (0.3–3 m) thick on each passage. The origin and evolution of comets are discussed in more detail below.

The current view considers the Oort Cloud as a steady-state reservoir which loses comets as just described and gains them from an inner cloud of comets located between the orbit of Neptune and the traditional Oort Cloud. Some comets can enter the inner solar system directly from the innermost part of the inner cloud, called the Kuiper Belt, discussed below. These discoveries mark a trend to expand the study of comets to bodies both larger and smaller than the traditional range. Ideas that periodic comet showers could result from a tenth planet in the inner cloud or a faint solar companion, called Nemesis, are novel but not generally accepted.

Observations. Observations of comets run the gamut of modern observing techniques. These include images in visual and ultraviolet wavelengths, photometry (accurate brightness measurements) in visual and infrared wavelengths, spectral scans in many wavelength regions, radio observations, and observations in extreme ultraviolet wavelengths and x-rays from rockets and orbiting spacecraft above the Earth's atmosphere. Visual wavelength spectra of comets show considerable variation. Ultraviolet spectra, obtained regularly with the *International Ultraviolet Explorer* satellite during its lifetime, have been found to be rather similar. Observations of meteors also contribute to the understanding of comets. *See* ASTRONOMICAL SPECTROSCOPY.

Measurements. The advent of space missions to comets has added an entirely new dimension to obtaining cometary information. A wide variety of direct measurement techniques have already been applied. Mass spectrometers have determined the properties of neutral gases, plasmas, and dust particles (via impact ionization). Dust counters have recorded the fluxes of dust and determined the mass distribution. Magnetometers have measured the strength and orientation of the magnetic field. Further plasma properties have been determined by plasma wave detectors. In addition, detailed imaging of the nucleus was made possible by the space missions. *See* HALLEY'S COMET.

Near-Sun and near-Earth comets. Most comets are observed under so-called normal circumstances; that is, the comet is between 0.5 and 1.5 AU from the Sun and a few tenths of an AU from Earth. Two specific exceptions are discussed here.

While comets have been observed at distances from the Sun of less than 0.1 AU for decades, only in 1979 was it learned that comets occasionally impact the Sun (**Fig. 2**). This fact was determined from a satellite (*SOLRAD*) used to monitor the Sun's corona; hence, the cometary discovery was not planned. The instrument carries an occulting disk which blocks out the bright glare of the Sun's surface, producing

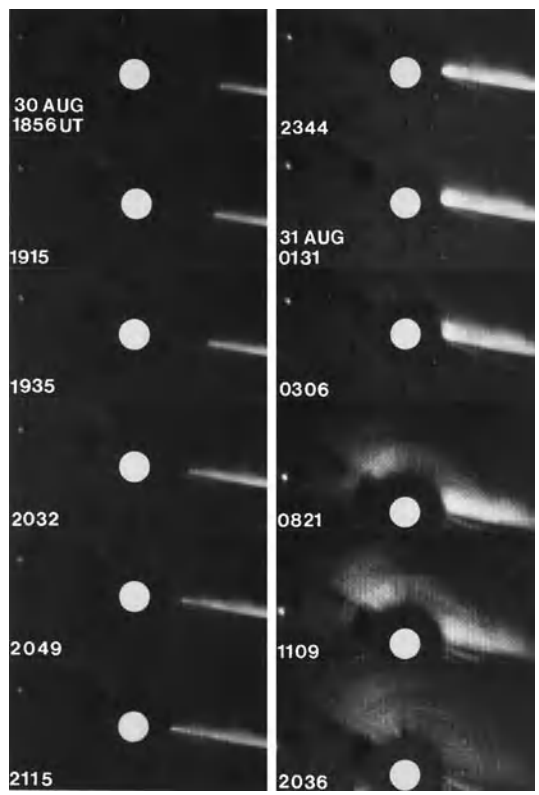

(a) (b)

Fig. 2. Sequence of photographs obtained from a *SOLRAD* satellite showing Comet Howard-Koomen-Michels hitting the Sun in August 1979. Captions give date and universal time (Greenwich mean time) of each photograph. (*a*) Before impact. (*b*) After impact. (*Naval Research Laboratory, Washington, DC*)

a dark circle with an apparent diameter of 5 solar radii. Comet Howard-Koomen-Michels hit the Sun on August 30, 1979; the estimated time of impact was 22^h29^m Greenwich mean time. The comet was clearly photographed (Fig. 2) approaching the Sun, but its reappearance, which would have produced a cometary image on the *SOLRAD* photographs, did not occur. Instead, a brightening of a major portion of the corona was observed which then slowly faded away. The brightening was presumably caused by cometary debris ultimately blown away by the Sun's radiation pressure following the disintegration of the nucleus.

Analysis of the data from this solar monitoring satellite and from other satellites, such as the *Solar Maximum Mission* (*SMM*) and the *Solar and Heliospheric Observatory* (*SOHO*), has turned up many more cometary impacts with the Sun. These results imply that solar impacts by relatively small comets are quite frequent. Observations by *SOHO* led to discovery of many comets and are providing a database for studying cometary phenomena near the Sun.

Comets very near the Earth are rare, but provide unique opportunities for new discoveries. Comet IRAS-Araki-Alcock was another example of a comet being first discovered by a satellite, in this case the *Infrared Astronomical Satellite* (*IRAS*), which also discovered additional comets. This comet passed within 0.031 AU of Earth on May 11, 1983. This distance was only about 12 times the average distance from the Earth to the Moon. Comet Hyakutake passed relatively close to Earth in late March and early April of 1996, thus providing spectacular naked-eye views.

The small distance to Comet IRAS-Araki-Alcock also provided a perfect opportunity to bounce radar waves off the nucleus. The experiment was successful, and the echoes provided evidence for the existence of a solid cometary nucleus with diameter of 0.6 mi (1 km).

The discovery of small comets hitting the Sun and the comets discovered by *IRAS* (including Comet IRAS-Araki-Alcock) imply that small comets are more numerous than had been generally believed. As the details for the origin of comets are worked out, it will be necessary to account for this fact. There has been increasing interest in small comets or comet-like objects, and the number of these objects in the inner solar system may be much larger than previously thought. In the past, the total population of comets might have been an important source of atmospheres for the terrestrial planets as well as a possible source of moderately complex organic molecules necessary for the initial development of life on Earth.

The proposal that the Earth is constantly being bombarded by very large numbers of house-size, icy objects has received considerable public attention. These objects have been called comets, but they are not comets because they do not sublimate. Their origin is unclear, and if they existed, they would have to have extraordinary properties. In sum, their existence is highly unlikely.

Kuiper Belt objects. Observations since 1992 have revealed a major population of large comets in a vast region beyond the orbit of Neptune. This discovery was anticipated about 1950 when K. E. Edgeworth and Gerard Kuiper, working independently, noted that the distribution of mass in the solar system dropped precipitously beyond the orbit of Neptune. They postulated that the missing mass was in a ring of undiscovered bodies in the trans-Neptunian region. These are commonly called Kuiper Belt objects.

Many of these objects, which continue to be discovered, are not small by cometary standards, and they are numerous. Credible estimates place roughly 100,000 objects with diameters larger than 60 mi (100 km) in a ring extending from 30 to 50 AU. Clearly, this is a major, largely unexplored population of cometary objects in this region. *See* KUIPER BELT.

Orbits. The first closed comet orbit to be calculated was Edmond Halley's elliptical orbit for the comet of 1680. This work indicated that comet orbits were ellipses with the Sun at one focus. In subsequent work, Halley noticed the striking similarity of the orbits of what were thought to be three different comets observed in 1531, 1607, and 1682. In 1705 he concluded that these were the same comet with a period of 75 or 76 years, and predicted its return in 1758. This comet is the one bearing Halley's name. It made an appearance in 1910 (**Fig. 3**), and again in 1985 and 1986.

The second comet to have its return successfully predicted was named after J. F. Encke. This comet has the shortest known period, 3.3 years. At its 1838 return, Encke's Comet showed another common property of comets, that is, a steadily changing period. This phenomenon is now known to result from the so-called nongravitational forces which must be explained by any successful comet model.

Six parameters are necessary to describe completely the orbit of a comet. They specify the orientation of the orbital plane in space, the orientation of the orbit in this plane, the size and shape of the ellipse, and the position of the comet along the orbit. In principle, three observations of position on the celestial sphere are sufficient, because each observation consists of independent measures of two coordinates. In practice, definitive orbits are derived from many observations, often hundreds.

While comet orbits are represented by ellipses to a good approximation, there are departures caused by the nongravitational forces and by the gravitational perturbations of the planets. When orbital parameters are listed for a comet, they refer to an ellipse which exactly matches the comet's position and motion at a specific time. Such an osculating orbit, as it is called, forms the starting point for studies of orbital evolution and for accurate predictions of the time and location of a comet's appearance in the sky. *See* CELESTIAL MECHANICS.

Currently, orbital parameters have been determined for roughly 1000 individual comets. Most are classified as long-period comets, that is, comets with

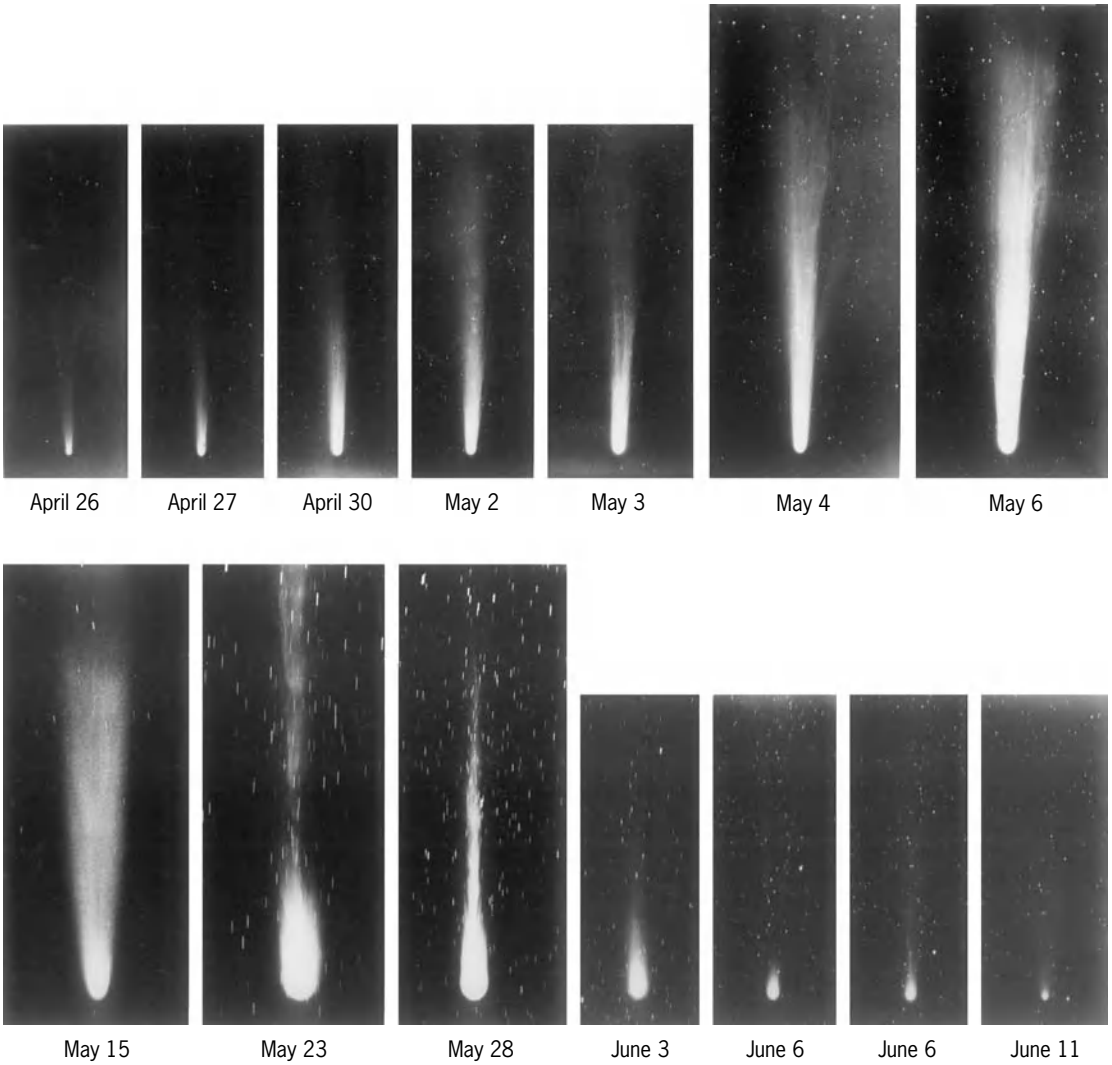

Fig. 3. Sequence of photographs showing the changing appearance of Halley's Comet in 1910. (*Carnegie Observatories*)

orbital periods greater than 200 years. The orbital planes of the long-period comets have approximately random inclinations with respect to the ecliptic. This means that there are as many comets with direct orbits (revolving around the Sun in the same sense as the planets) as with retrograde orbits (revolving in the sense opposed to the planets' motion). Careful examination of the original (that is, inbound) orbits of the long-period comets shows none that are hyperbolas; that is, no interstellar comets have yet been observed. This fact strongly implies that the cloud of comets is gravitationally bound to the Sun and therefore is a part of the solar system.

The short-period comets are mostly in direct orbits with inclinations of less than 30°. The distribution of periods shows a peak between 7 and 8 years, and the majority have an aphelion (point of greatest distance from the Sun) near the orbit of Jupiter.

Composition. The many spectroscopic studies and direct spacecraft measurements on the composition of comets have observed numerous species of atoms, molecules, and ions (see **table**). Fairly complex molecules such as polymerized formaldehyde [$(H_2CO)_n$] and methyl cyanide (CH_3CN) are present.

Observations of the gas composition in Halley's coma have found these values by number of molecules: water, approximately 80%; carbon monoxide, roughly 10%; carbon dioxide, approximately 3.5%; complex organic compounds such as polymerized formaldehyde, a few percent; and trace substances, the remainder. These values were obtained from direct measurements, except for that of carbon monoxide, which was determined from a rocket

Atomic and molecular species observed in comets*

Neutrals	Ions
H, OH, O, S, H_2O, H_2CO	CO^+, CO_2^+,
C, C_2, C_3, CH, CN, CO, CS, S_2	H_2O^+, OH^+,
HCN, CH_3CN, NH	H_3O^+
NH_2, Na, Fe, K	CH^+, N_2^+, Ca^+,
Ca, V, Cr, Mn, Co, Ni, Cu	C^+, CN^+
$(H_2CO)_n$	

*Species observed in coma and tail in spectroscopic studies and direct spacecraft measurements.

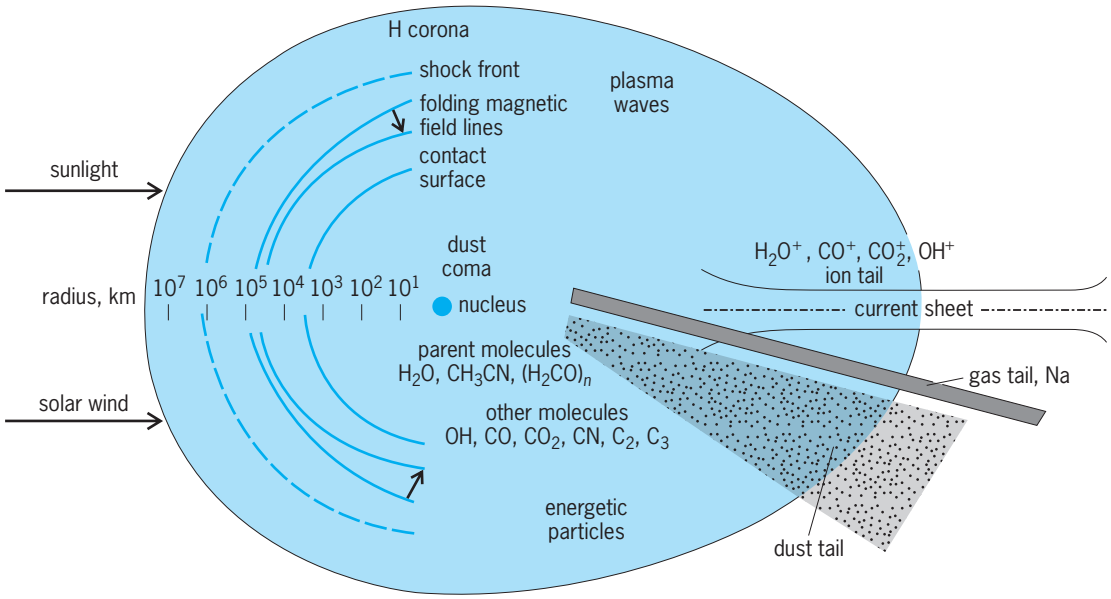

Fig. 4. Cometary features diagrammed on a logarithmic scale. 1 km = 0.6 mi.

experiment. These values are for the outer layers of a comet which has passed through the inner solar system many times. The values of a new comet or the interior of Halley may be different with a higher fraction of carbon dioxide likely.

Measurements of dust composition were made by instruments on the spacecraft sent to Halley's Comet. Some of the particles contain essentially only the atoms hydrogen, carbon, nitrogen, and oxygen; these are the "CHON" particles. Other particles have a silicate composition, that is, they resemble rocks found throughout the solar system. The dominant dust composition is a mixture of these two and resembles carbonaceous chondrites enriched in the light elements hydrogen, carbon, nitrogen, and oxygen.

Structure. The main components of a comet are the nucleus, coma, hydrogen cloud, and tails (**Fig. 4**).

Nucleus. Strong evidence points to the existence of a central nuclear body or nucleus for all comets from which all cometary material, both gas and dust, originates. In the early 1950s, F. L. Whipple proposed an icy conglomerate model of the nucleus. In this model, the nucleus is not a cloud of particles as had previously been thought (the "sandbank" model), but a single mass of ice with embedded dust particles, commonly called a dirty snowball. Such a nucleus could supply an adequate amount of gas to explain cometary phenomena and last through many apparitions because only a relatively thin surface layer would be eroded away (by sublimation) during each passage near the Sun.

The nucleus of Halley's Comet has been directly observed. Its shape was highly irregular, the surface exhibited features and structure, and the albedo (fraction of reflected light) was very low. While the confirmation of the existence of a single, nuclear body is important, Halley is just one comet and generalizations from it may not be valid. *See* ALBEDO.

Fortunately, images of three more nuclei have been obtained. The image of the nucleus of Comet Borelly was taken by the *Deep Space 1* spacecraft on September 22, 2001. The close-up view of the nucleus (**Fig. 5***a*) shows considerable detail and an irregular shape about 5 mi (8 km) long. The image shows smooth terrain around the central part of the nucleus, which is also the part of the nucleus with active jets. The remainder of the terrain is generally inactive and has features identified as pits, troughs, and ridges. The albedo over the surface varies from 0.01 to 0.03; in other words, the surface is very dark.

The image of the nucleus of Comet Wild 2 was taken by the *Stardust* spacecraft on January 2, 2004. The close-up view of the nucleus (Fig. 5*b*) shows detail including an apparent impact crater, and a rounded body with a typical dimension of 2.5 mi (4 km). The albedo is 0.03 ± 0.015; that is, it is also very dark. The surface has high-angle slopes which indicate that the body is cohesive and self-supporting. The images of Comet Wild 2 serve to emphasize the differences with Comets Halley and Borrelly. The diversity of comets is clearly very important.

The generic nuclear body (**Fig. 6**) is irregular, with radius ranging from around 1000 ft (300 m) to 12 mi (20 km). Masses would range in order of magnitude from 10^{11} to 10^{16} kg (10^{11} to 10^{16} lb), with roughly equal parts of snow and dust. The average density would be in the range 0.2 to 2 g/cm^3 (0.2 to 2 times that of water). The interior is probably porous, which means that there are significant voids along with the regions of solid material. Nuclei of comets sometimes split into two or more pieces; Comet West (1976 VI) is an example.

Coma. The coma is observed as an essentially spherical cloud of gas and dust surrounding the nucleus. The principal gaseous constituents are neutral molecules (see table), and the dust composition has

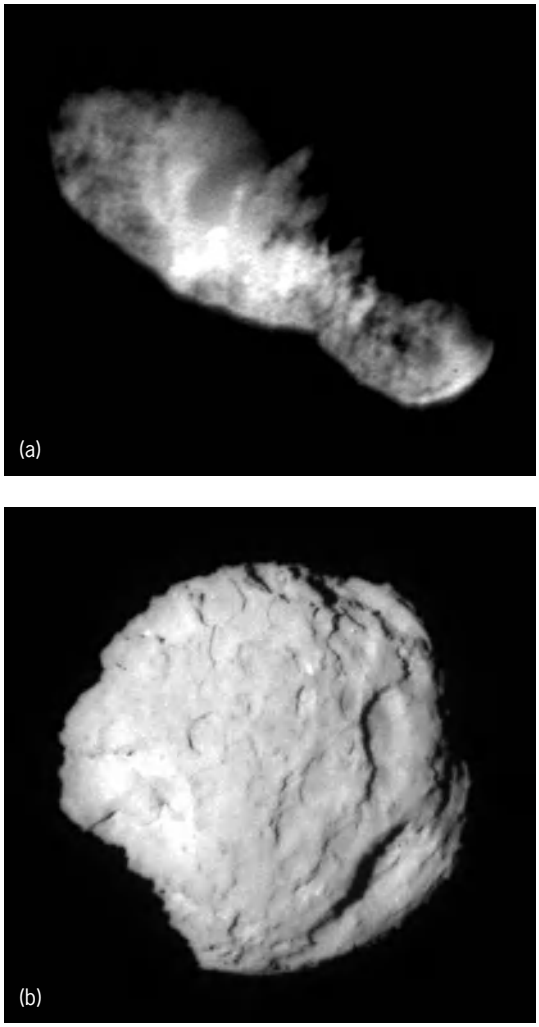

Fig. 5. Images of comet nuclei. (*a*) Close-up image of Comet Borrelly's nucleus taken on September 22, 2001, by the *Deep Space 1* spacecraft (*NASA/JPL*). (*b*) Image of Comet Wild 2's nucleus taken on January 2, 2004, just after closest approach by the *Stardust* spacecraft (*NASA; Stardust mission team*).

been described above. The coma can extend as far as 10^5 to 10^6 km (10^5 to 10^6 mi) from the nucleus, and the material is flowing away from the nucleus at a typical speed of 0.6 mi/s (1.0 km/s). As the gas flows away from the nucleus, the dust particles are dragged along. For comets at heliocentric distances greater than 2.5 to 3 AU, the coma is not normally visible and is presumed not to be present. An exception is Halley's Comet, which displayed an outburst and comalike activity at 14.4 AU.

Hydrogen cloud. In 1970, observations of Comet Tago-Sato-Kosaka and Comet Bennett from orbiting spacecraft showed that these comets were surrounded by a giant hydrogen cloud that extends to distances on the order of 10^7 km (10^7 mi), or a size larger than the Sun. The observations were made in the resonance line of atomic hydrogen at 121.6 nm. Hydrogen clouds have since been observed for many other comets. Fairly bright comets (such as Bennett) have a hydrogen production rate by sublimation from the nucleus at heliocentric distance of 1 AU in the range 3 to 8×10^{29} atoms/s. The size of the hydrogen cloud depends on the velocity of the outflowing hydrogen atoms, and 5 mi/s (8 km/s) has been derived. This velocity would arise (from energy balance considerations) if most of the hydrogen in the cloud were produced by the photodissociation of the hydroxyl radical (OH).

Tails. Photographs of bright comets generally show two distinct types of tails (Fig. 1): the dust tails and the plasma tails. They can exist separately or together in the same comet. In a color photograph, the dust tails appear yellow because the light is reflected sunlight, and the plasma tails appear blue from emission due to ionized carbon monoxide, CO^+.

Studies of features in plasma tails led L. Biermann in 1951 to postulate the existence of a continuous outflow of ionized material from the Sun, which he called the solar corpuscular radiation, now called the solar wind. The interaction of the solar wind and its magnetic field, as suggested by H. Alfvén in 1957, plays an important role in cometary physics and in the formation of plasma tails. The direct measurements on the sunward side by the spacecraft sent to Halley's Comet, on the tailward side by the *International Cometary Explorer* mission to Comet Giacobini-Zinner in 1985, by the *Giotto* spacecraft passing Comet Grigg-Skjellerup in 1992, and by the *Deep Space 1* spacecraft passing Comet Borrelly in 2001 have confirmed these views. *See* SOLAR WIND.

The plasma tails are generally straight and have lengths that range in order of magnitude from 10^7 to 10^8 km (10^7 to 10^8 mi). The Great Comet of 1843 had a plasma tail extending over 2 AU in length. The plasma in these tails is composed of electrons and molecular ions (see table). The dominant visible ion is CO^+. The zone of production for the molecular plasma appears to be in the coma near the sunward side of the nucleus. The material in the plasma tails is concentrated into thin bundles or streamers, and additional structure is found in the tail in the form of knots and kinks. These features appear to move along the tail away from the head at speeds of 6 mi/s (10 km/s) to 120 mi/s (200 km/s). Plasma tails are generally not observed beyond heliocentric distances of 1.5 to 2 AU; an exception is Comet Humason, which showed a spectacular, disturbed plasma tail well beyond normal distances.

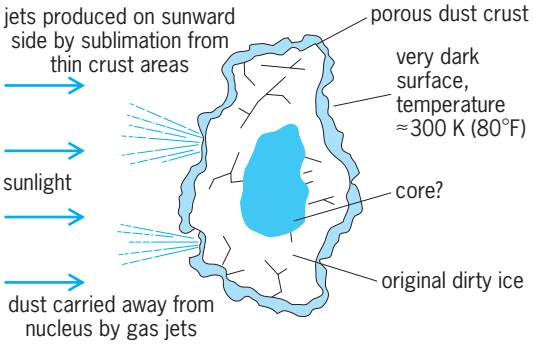

Fig. 6. Schematic model of the cometary nucleus.

The dust tails are usually curved and have lengths that range in order of magnitude from 10^6 to 10^7 km (10^6 to 10^7 mi). Normally, the dust tails are relatively homogeneous; an exception, Comet West, is shown in Fig. 1. Observations indicate that the dust particles are typically 1 micrometer in diameter and are probably silicate in composition. Occasionally dust tails are seen which appear to point in the sunward direction, the so-called antitails. Examples are Comet Arend-Roland in 1957 and Halley's Comet in 1986. These are not truly sunward appendages but are the result of projection effects.

Observations of the bright comet Hale-Bopp have revealed gas tails in the element sodium. One of these is long and narrow (Fig. 4). The other is a diffuse sodium gas tail superimposed on the dust tail. The sodium tails are probably present in all comets but were made more easily visible by Comet Hale-Bopp's large size.

Modern theory. The goal of modern comet theory is to explain the facts about comets outlined above. The broad, theoretical approach appears to be in reasonably good shape and has survived the tests of direct exploration by spacecraft (**Fig. 7**).

Sublimation from nucleus. The cornerstone of the current best ideas is F. L. Whipple's icy conglomerate model of the nucleus. As a comet approaches the Sun on its orbit, sunlight supplies radiant energy to the surface of the nucleus. The energy received heats the nucleus when it is far from the Sun. As the comet

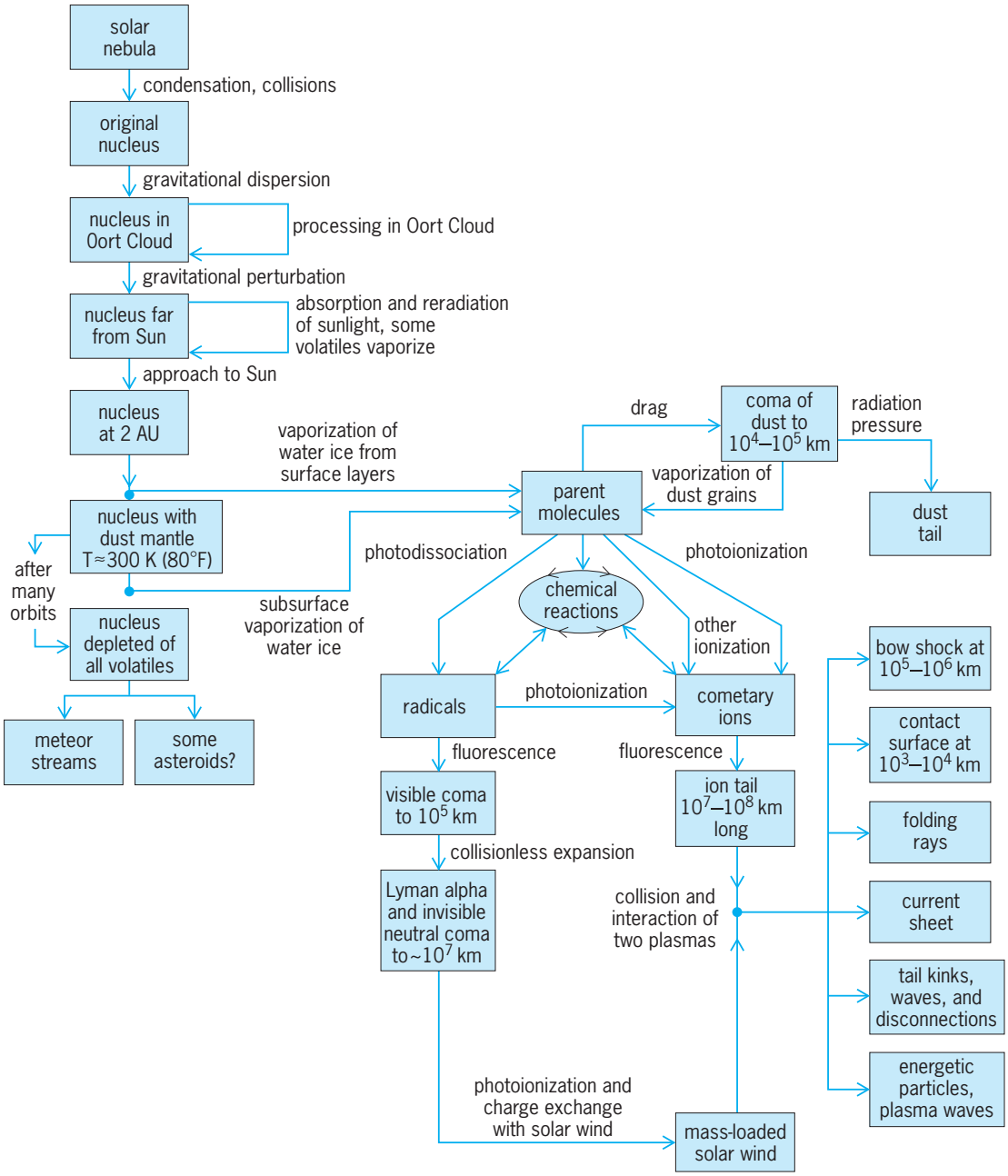

Fig. 7. Summary of processes involved in the formation of comets and the interaction with sunlight and the solar wind. 1 km = 0.6 mi.

continues toward the Sun, the temperature of the surface layers increases to a value, determined by the thermodynamic properties of the ice, where sublimation (passage from the solid state directly to the gaseous state) occurs. Then, most of the incident energy goes to the sublimation of ices. The situation has at least one additional complexity. Sublimation of ice from a dust–ice mixture probably leaves a dust crust which is heated to temperatures higher than the sublimation temperature of water ice. The heat is conducted inward to the ice layers, and the sublimated gases then pass outward through the dust crust. The jets of material seen on Halley's Comet originate in areas of minimal dust crust.

The onset of activity in comets at 2.5 to 3 AU is entirely consistent with water ice as the dominant ice constituent of the nucleus. But a problem still exists. Most other possible constituents are predicted from theory to begin sublimation much farther from the Sun than water ice, but, for example, molecular emissions from CN are visible essentially at the onset of activity.

The solution to this problem lies in the dominant form of water ice found in comets, amorphous ice. Models for the formation of comets treat the process as occurring under low-temperature and low-pressure conditions. When ices form under these conditions, energy is not available to convert the amorphous ice to crystalline ice (a minimum energy state). The transition to crystalline ice takes place at temperatures much higher than the formation temperature. Also, minor constituents are trapped in the amorphous ice. They are partially released when the amorphous ice crystallizes, and the rest is released when the ice sublimates. Some minor constituents may not be trapped in the amorphous ice, and these could be vaporized and released from comets beyond 3 AU, particularly on a comet's first approach to the Sun.

Origin of nongravitational forces. The nongravitational forces have a ready explanation if the cometary nucleus is rotating. The result of radiant energy producing sublimation of ices in an ice-dust mixture is to leave a crust of dusty material. Thus, there is a time lag while the heat traverses the dust layer between maximum solar energy received and maximum loss through sublimation. If the nucleus were not rotating, the maximum mass loss would be directly toward the Sun. For a rotating nucleus, the mass loss occurs away from the sunward direction, toward the afternoon side of the comet. The analogous situation on Earth (that is, the time lag between cause and effect) produces the warmest time of day in the afternoon, not at noon. The mass loss under these circumstances produces a force on the nucleus via the rocket effect (reaction force), and this force can accelerate or retard the motion of the comet in its orbit. Detailed studies of the nongravitational forces in comets show that they are entirely consistent with water ice as the controlling substance.

A possible alternative origin of nongravitational forces is still based on water as the controlling substance. For some comets, simply having a larger mass loss before perihelion compared to after perihelion would produce the nongravitational effects.

Formation of coma. The sublimated gases, mostly neutral molecules, flow away from the nucleus, dragging some of the dust particles with them to form the coma. Close to the nucleus, the densities are high enough that chemical reactions can occur between molecular species. Photodissociation is also important. Thus, the molecules observed spectroscopically far from the nucleus often are not the same as the initial composition.

Formation of dust tails. The dust particles carried away from the nucleus by the flow of coma gases are blown in the antisolar direction by the Sun's radiation pressure to form the dust tails. The detailed theory has been developed through the years, and good agreement can usually be obtained with the observed shapes and sizes of the tails if the emission of dust from the nucleus has a peak before perihelion. The larger particles liberated from the nucleus can orbit the Sun and reflect sunlight to produce the zodiacal light. *See* ZODIACAL LIGHT.

Formation of plasma tails. The gas flowing away from the nucleus has a more involved fate. Under normal circumstances, when a comet's heliocentric distance is about 1.5 to 2 AU or less, significant ionization of the coma molecules occurs (probably by solar radiation), and this triggers a reaction with the solar wind. At the Earth, the solar wind, a fully ionized proton-electron gas, flows away from the Sun at 250–310 mi/s (400–500 km/s) and has an embedded magnetic field. Because of the magnetic field, the ionized cometary molecules cause the solar-wind field lines to slow down in the vicinity of the comet while proceeding at the full solar-wind speed away from the comet. This situation causes the field lines and the trapped plasma to wrap around the nucleus like a folding umbrella, to form the plasma tail. This picture has been completely confirmed by the spacecraft measurements. In addition, the folding can be seen and photographed because of the emission from the trapped ions (such as CO^+) which serve as tracers of the field lines (**Fig. 8**). Thus, while the ionized molecules are indeed swept in the antisolar direction by the solar wind, the plasma tail should be thought of as a part of the comet attached to the near-nuclear region by the magnetic field captured from the solar wind.

Exceptions occur at times apparently when the polarity of the solar-wind magnetic field changes. In terms of solar-wind structure, this polarity reversal takes place when the comet crosses the heliospheric current sheet (HCS). This can disrupt the magnetic connection to the near-nuclear region and literally causes the old plasma tail to disconnect while the new tail is forming. This process is quite common (**Fig. 9**). While the physics of this phenomenon is complex, the observational and theoretical basis is sound.

Fate of comets. When the process of sublimation has been carried out over an extensive period of time, as would be the case for the short-period comets, the ices will be exhausted and the inactive or "dead"

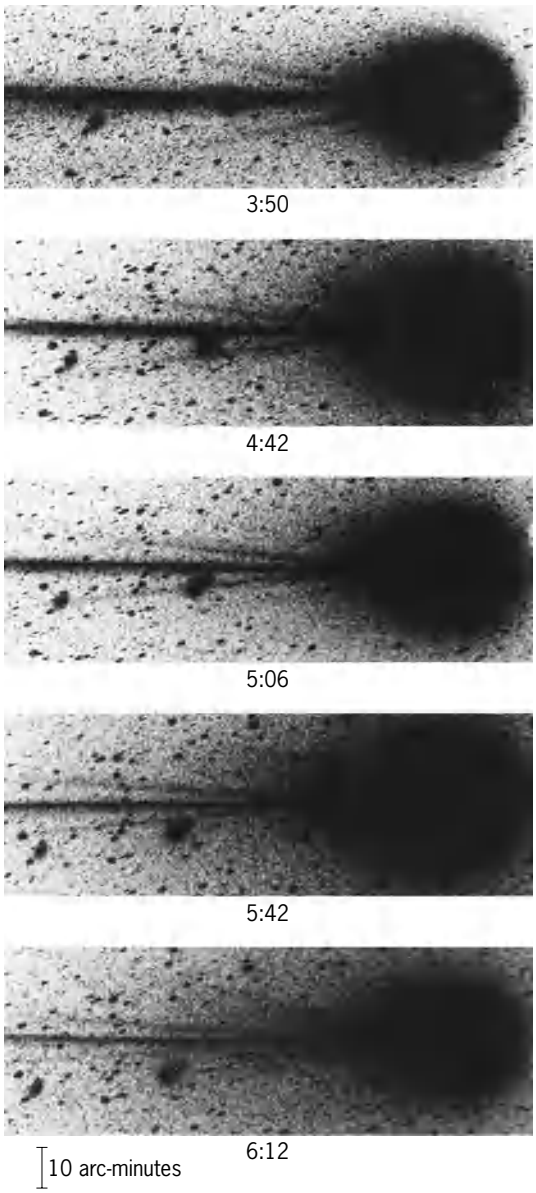

Fig. 8. Sequence of photographs of Comet Kobayashi-Berger-Milon on July 31, 1975, showing the capture of magnetic field from the solar wind. The dominant pair of tail streamers visible on either side of the main tail lengthen and turn toward the tail axis in this sequence. (*Joint Observatory for Cometary Research, operated by NASA–Goddard Space Flight Center and New Mexico Institute of Mining and Technology*)

comet should consist of dust particles and larger-sized rocky material. These remnants are dispersed along the comet's orbit by perturbations, and are the particles responsible for producing meteors when they enter the Earth's atmosphere. Small particles, very probably of cometary origin, have been collected by high-flying aircraft. *See* METEOR.

Comet Shoemaker-Levy 9 and Jupiter. Images of Comet Shoemaker-Levy from the Hubble Space Telescope in 1993 and early 1994 show a string of many bright points (**Fig. 10***a*). The comet then consisted of pieces from a larger body torn apart by a close encounter with Jupiter in 1992. This comet's life ended with a spectacular collision with Jupiter over a period of a few days around July 22, 1994 (Fig. 10*b*). These events dramatically illustrate disruption and collisions with other solar system bodies as fates for cometary bodies. *See* JUPITER.

Comet Hyakutake. In late March 1996, Comet Hyakutake came within about 0.1 AU of Earth, providing spectacular views. On the evening of March 25, the plasma tail stretched from near Polaris well to the south. A length of 55° was easily observed, and experienced observers in dark sites reported lengths of 100°, that is, the tail extended more than halfway across the sky.

The comet was very low in dust content, and the plasma was conspicuous. On March 24–25, a superb disconnection event took place (Fig. 9); the disconnected part of the tail could be traced to a distance of at least 35° from the head. The disconnection event and the disturbed appearance of the plasma tail are characteristic of a comet in the stormy equatorial region of the solar wind. The appearance of Comet Hyakutake contrasted with that of Comet Hale-Bopp (**Fig. 11**).

A major unexpected discovery was that Comet Hyakutake is a bright and variable source of x-rays. Searches of databases of satellites observing in x-rays subsequently showed that x-ray emission is a common property of comets. The x-rays are produced when a highly ionized species in the solar wind (such as oxygen ionized six times) exchanges charge (an electron) with neutral molecules in the coma. The

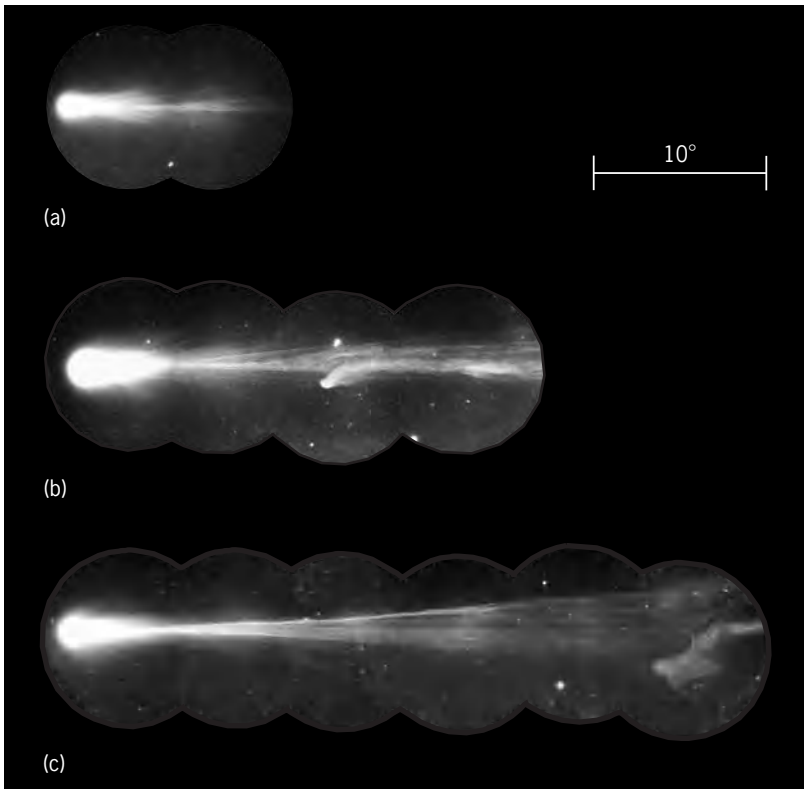

Fig. 9. Photographs of Comet Hyakutake (C/1996 B2) taken on three successive days in March 1996 by Shigemi Numazawa of Niigata, Japan. A spectacular disconnection event is clearly shown, and a new plasma tail immediately forms. (*a*) March 24. (*b*) March 25. (*c*) March 26. (*Courtesy of Sky and Telescope*)

(a) (b)

Fig. 10. Comet Shoemaker-Levy 9. (*a*) Mosaic of three *Hubble Space Telescope* images taken in January 1994. Twenty comet fragments are visible (*H. A. Weaver and T. E. Smith, Space Telescope Science Institute, NASA*). (*b*) Sequence of *Hubble Space Telescope* images of Jupiter showing evolution of impact site of fragment "G," the largest fragment of the comet. From lower left to upper right, the images show the impact plume visible just off the limb about 5 minutes after impact on July 18, 1994; the fresh impact site about 1.5 hours after the impact; the evolved "G" impact site and other impact sites on July 21; and the view on July 23 (*R. Evans, J. Trauger, H. Hammel, HST Comet Science Team, NASA*)

result is a species with reduced ionization (in the previous example, oxygen ionized five times) but in an excited state. These ions spontaneously emit photons in the x-ray wavelength range.

Comet Hale-Bopp. Comet Hale-Bopp provided spectacular views in March and April 1997. Because of the observing circumstances, the comet was often both an evening and a morning object. Its intrinsic brightness can be appreciated by the fact that the comet was never closer to the Earth than 1 AU. Comet Hale-Bopp is clearly the largest comet studied with modern technology, even though no spacecraft was sent to the comet. An extensive body of data was gathered.

Because of the comet's great intrinsic brightness, a large nucleus was suspected. The brightness profile obtained from Hubble Space Telescope images showed a sharp peak that could be identified with light from the nucleus itself. A diameter of about 25 mi (40 km) was inferred. This is a very large nucleus, with roughly 65 times the volume of Comet Halley's nucleus.

Fig. 11. Photograph of Comet Hale-Bopp taken on April 8, 1997, by H. Mikuz and B. Kambik at the Crni Vrh Observatory, Slovenia. The image shows the smooth, white dust tail (right) and the blue plasma tail (left); the latter shows some filamentary structure but is relatively undisturbed. (*Ulysses Comet Watch*)

A wide-field image of Comet Hale-Bopp (Fig. 11) shows large dust and plasma tails. The dust tail was expected because this was a very dusty comet. The appearance of the plasma tail is unusual because of the inclination of the comet's orbit to the plane of the ecliptic of 89.4°. This means that the orbit is almost exactly perpendicular to the Earth's orbit. Comet Hale-Bopp spent months in the calm polar solar wind and was the best-documented comet in the polar region. The relatively undisturbed, almost serene appearance of the plasma tail is a direct consequence of the calm polar-region environment.

Origin. For years, it seemed likely that the population of short-period comets could be produced from the population of the long-period comets—those in Oort's Cloud—by their gravitational interaction with Jupiter. It is now understood that they originate in a ring of comets in the vicinity of the orbits of Neptune and Uranus, and beyond.

The Oort Cloud is the source of the long-period comets. The evidence from the statistics of cometary orbits indicates an essentially spherical cloud of comets with dimensions in the range 10^4 to 10^5 AU. Gravitational perturbations from passing stars and gas clouds have several effects on the cloud. They limit its size and tend to make the orbits random (as observed). Most importantly, the perturbations continually send new comets from the cloud into the inner solar system, where they are observed. Thus, the Oort Cloud can be considered as a steady-state reservoir for new comets. Evidence is mounting for the view that the Oort Cloud is supplied by an inner cloud of comets.

The current consensus on the origin of comets holds that they condensed from the solar nebula at the same time as the formation of the Sun and planets. In other words, although the details are sketchy, comets are a natural by-product of the solar system's origin. They may be remnants of the formation process, and their material may be little altered from the era of condensation to the present time. *See* SOLAR SYSTEM.

The consensus scenario for the origin is as follows. Generally accepted models of the solar nebula have temperature and density conditions suitable for the condensation of cometary materials at solar distances around the orbits of Uranus and Neptune. The condensate could coalesce to produce a ring of comets or cometesimals. Gravitational perturbations by the major planets would disperse the ring of comets, sending some into the inner cloud or Oort's Cloud and sending some inward to become short-period comets.

Some scenarios for formation and evolution assign additional roles to the comets. They may have been a major source of the atmospheres of the terrestrial planets and might even have provided the organic molecules necessary for the evolution of life on Earth. While some of these ideas are speculative, they help show why comets are objects of keen interest and active study, and why several space agencies have launched deep-space probes to comets beginning in the 1980s.

Space missions to comets. Interest in sending a spacecraft to a comet was heightened by the appearance of Halley's Comet in 1985 and 1986. The first mission was the diversion of the third *International Sun-Earth Explorer* (*ISEE 3*) from Earth orbit to pass through the tail of Comet Giacobini-Zinner in September 1985. The spacecraft, renamed *International Cometary Explorer* (*ICE*), measured the properties of the tail, including the magnetic field, and the general plasma environment. For a discussion of the missions to Halley's Comet, also called the Halley armada, *see* HALLEY'S COMET.

One Armada spacecraft, *Giotto*, was diverted to intercept Comet Grigg-Skjellurup in July 1992.

The *Deep Space 1* spacecraft encountered Comet Borrelly in September 2001 and obtained imaging and plasma data, as mentioned above.

The *Stardust* mission was launched in February 1999 and encountered Comet Wild 2 in January 2004. Besides imaging of the nucleus and coma, the mission plan is to bring samples of cometary dust back to Earth. This has been accomplished.

The *Deep Impact* mission to Comet Tempel 1 was launched in January 2005. On July 4, 2005, it sent an 820-lbm (370-kg) impactor spacecraft (mostly copper) into the comet's nucleus. **Figure 12** shows the scene shortly after the impact. Because of the large amount of dust in the plume, the expected crater cannot be seen with certainty. Studies of the plume indicate that the outer layer of the comet (down to tens of meters) consists of fine powdery dust only loosely bound by the gravity of the nucleus.

The *Rosetta* mission to Comet Churyumov-Gerasimenko was launched in March 2004. The journey to the comet takes approximately 10 years, and the plan is highly ambitious. Part of the spacecraft will spend about 2 years near the comet and part will land on the comet's nucleus. The nominal mission end is in 2015.

Comets and the public. There is no scientific evidence for the belief that comets are harbingers, omens, or actual producers of evil, disasters, or natural calamities. Some facts about comets can be distorted to provide an apparent basis for some fears. There are toxic gases in comets, but the amount is very small and, moreover, the cometary gas would not penetrate the Earth's atmosphere. The only real problem would be an impact of a comet on the Earth's surface. The odds of an impact are very small, and studies indicate that the somewhat fragile cometary nucleus would break up high in the Earth's atmosphere. Thus, the Tunguska event, which took place in central Siberia on June 30, 1908, and which produced substantial devastation, is no longer believed to have a cometary origin.

Fig. 12. Image of Comet Tempel 1, 67 s after the impact, taken from the *Deep Impact* flyby spacecraft. The nucleus and the plume of material from the impact are clearly visible. (*NASA/JPL-Caltech/University of Maryland*)

The opportunities for much of the public to view bright comets are becoming increasingly rare due to the combination of atmospheric pollution and bright city lights. In January 1974 there was general public disappointment over the poor visibility of Comet Kohoutek, particularly as viewed from cities in the populous northeastern United States, although it was a conspicuous object to astronomers at mountaintop observatories in the southwestern United States. Halley's Comet was not a spectacular object for public viewing in 1985 and 1986. In part, this was due to the circumstances of the comet's orbit, which led to the comet's times of greatest intrinsic brightness occurring well away from the Earth in contrast to the situation in 1910. The march of light pollution is inexorable, and the interested viewer of comets will require an observing site well away from urban population centers. John C. Brandt

Bibliography. H. Böhnhardt et al. (eds.), *Cometary Science after Hale-Bopp: Invited Review Papers*, 2003; J. C. Brandt, *Comets: A Scientific American Reader*, 1981; J. C. Brandt and R. D. Chapman, *Introduction to Comets*, 2d ed., 2004; J. C. Brandt and R. D. Chapman, *Rendezvous in Space*, 1992; H. Crovisier and T. Encrenaz, *Comet Science*, 2000; W. F. Huebner (ed.), *Physics and Chemistry of Comets*, 1990; The *International Cometary Explorer* Mission to Comet Giacobini-Zinner, *Science*, 232:353–385, 1986.

Comfort heating

The maintenance of the temperature in a closed volume, such as a home, office, or factory, at a comfortable level during periods of low outside temperature. Two principal factors determine the amount of heat required to maintain a comfortable inside temperature: the difference between inside and outside temperatures and the ease with which heat can flow out through the enclosure. *See* CONDUCTION (HEAT).

Heating load. The first step in planning a heating system is to estimate the heating requirements. This involves calculating heat loss from the space, which in turn depends upon the difference between outside and inside space temperatures and upon the heat transfer coefficients of the surrounding structural members.

Outside and inside design temperatures are first selected. Ideally, a heating system should maintain the desired inside temperature under the most severe weather conditions. Economically, however, the lowest outside temperature on record for a locality is seldom used. The design temperature selected depends upon the heat capacity of the structure, amount of insulation, wind exposure, proportion of heat loss due to infiltration or ventilation, nature and time of occupancy or use of the space, difference between daily maximum and minimum temperatures, and other factors. Usually the outside design temperature used is the median of extreme temperatures.

The selected inside design temperature depends upon the use and occupancy of the space. Generally it is between 66 and 75°F (19 and 24°C).

The total heat loss from a space consists of losses through windows and doors, walls or partitions, ceiling or roof, and floor, plus air leakage or ventilation. All items but the last are calculated from $H_l = UA(t_i - t_o)$, where heat loss H_l is in British thermal units per hour (or in watts), U is overall coefficient of heat transmission from inside to outside air in Btu/(h)(ft^2)(°F) (or J/s · m^2 · °C), A is inside surface area in square feet (or square meters), t_i is inside design temperature, and t_o is outside design temperature in °F (or °C).

Values for U can be calculated from heat transfer coefficients of air films and heat conductivities for building materials or obtained directly for various materials and types of construction from heating guides and handbooks.

The heating engineer should work with the architect and building engineer on the economics of the completed structure. Consideration should be given to the use of double glass or storm sash in areas where outside design temperature is 10°F (−12°C) or lower. Heat loss through windows and doors can be more than halved and comfort considerably improved with double glazing. Insulation in exposed walls, ceilings, and around the edges of the ground slab can usually reduce local heat loss by 50–75%. **Table 1** compares two typical dwellings. The 43% reduction in heat loss of the insulated house produces a worthwhile decrease in the cost of the heating plant and its operation. Building the house tight reduces the large heat loss due to infiltration of outside air. High heating-energy costs may now warrant 4 in. (10 cm) of insulation in the walls and 8 in. (20 cm) or more in the ceiling.

Humidification. In localities where outdoor temperatures are often below 36°F (2°C), it is advisable to provide means for adding moisture in heated spaces to improve comfort. The colder the outside air is, the less moisture it can hold. When it is heated to room temperature, the relative humidity in the space becomes low enough to dry out nasal membranes, furniture, and other hygroscopic materials. This results in discomfort as well as deterioration of physical products.

Various types of humidifiers are available. The most satisfactory type provides for the evaporation of the water to take place on a mold-resistant treated material which can be easily washed to get rid of the resultant deposits. When a higher relative humidity is maintained in a room, a lower dry-bulb temperature or thermostat setting will provide an equal sensation

TABLE 1. Effectiveness of double glass and insulation*

Heat-loss members	Area, ft^2 (m^2)	Heat loss, Btu/h†	
		With single-glass weather-stripped windows and doors	With double-glass windows, storm doors, and 2-in. (5.1-cm) wall insulation
Windows and doors	439 (40)	39,600	15,800
Walls	1952 (181)	32,800	14,100
Ceiling	900 (84)	5,800	5,800
Infiltration		20,800	20,800
Total heat loss		99,000	56,500
Duct loss in basement and walls (20% of total loss)		19,800	11,300
Total required furnace output		118,800	67,800

*Data are for two-story house with basement in St. Louis, Missouri. Walls are frame with brick veneer and 25/32-in. (2.0-cm) insulation plus gypsum lath and plaster. Attic floor has 3-in. (7.5-cm) fibrous insulation or its equivalent. Infiltration of outside air is taken as a 1-h air change in the 14,400 ft^3 (408 m^3) of heating space. Outside design temperature is −5°F (−21°C); inside temperature is selected as 75°F (24°C).

†1 Btu/h = 0.293 W.

of warmth. This does not mean, however, that there is a saving in heating fuel, because heat from some source is required to evaporate the moisture.

Some humidifiers operate whenever the furnace fan runs, and usually are fed water through a float-controlled valve. With radiation heating, a unitary humidifier located in the room and controlled by a humidistat can be used.

Insulation and vapor barrier. Good insulating material has air cells or several reflective surfaces. A good vapor barrier should be used with or in addition to insulation, or serious trouble may result. Outdoor air or any air at subfreezing temperatures is comparatively dry, and the colder it is the drier it can be. Air inside a space in which moisture has been added from cooking, washing, drying, or humidifying has a much higher vapor pressure than cold outdoor air. Therefore, moisture in vapor form passes from the high vapor pressure space to the lower pressure space and will readily pass through most building materials. When this moisture reaches a subfreezing temperature in the structure, it may condense and freeze. When the structure is later warmed, this moisture will thaw and soak the building material, which may be harmful. For example, in a house that has 4 in. (10 cm) or more of mineral wool insulation in the attic floor, moisture can penetrate up through the second floor ceiling and freeze in the attic when the temperature there is below freezing. When a warm day comes, the ice will melt and can ruin the second floor ceiling. Ventilating the attic helps because the dry outdoor air readily absorbs the moisture before it condenses on the surfaces. Installing a vapor barrier in insulated outside walls is recommended, preferably on the room side of the insulation. Good vapor barriers include asphalt-impregnated paper, metal foil, and some plastic-coated papers. The joints should be sealed to be most effective. *See* HEAT INSULATION.

Thermography. Remote heat-sensing techniques evolved from space technology developments related to weather satellites can be used to detect comparative heat energy losses from roofs, walls, windows, and so on. A method called thermography is defined as the conversion of a temperature pattern detected on a surface by contrast into an image called a thermogram (see **illus.**). Thermovision is defined as the technique of utilizing the infrared radiation from a surface, which varies with the surface temperatures, to produce a thermal picture or thermogram. A camera can scan the area in question and focus the radiation on a sensitive detector which in turn converts it to an electronic signal. The signal can be amplified and displayed on a cathode-ray tube as a thermogram.

Normally the relative temperature gradients will vary from white through gray to black. Temperatures from −22 to 3540°F (−30 to 2000°C) can be measured. Color cathode-ray tubes may be used to display color-coded thermograms showing as many as 10 different isotherms. Permanent records are possible by using photos or magnetic tape.

Thermograms of building structures: (*a-c*) masonry buildings; (*d*) glass-faced building. Black indicates negligible heat loss; gray, partial loss; and white, excessive loss. (*Courtesy of A. P. Pontello*)

Infrared thermography is used to point out where energy can be saved, and comparative insulation installations and practices can be evaluated. Thermograms of roofs are also used to indicate areas of wet insulation caused by leaks in the roof.

Infiltration. In **Table 2**, the loss due to infiltration is large. It is the most difficult item to estimate accurately and depends upon how well the house is built. If a masonry or brick-veneer house is not well caulked or if the windows are not tightly fitted and weather-stripped, this loss can be quite large. Sometimes, infiltration is estimated more accurately by measuring the length of crack around windows and

TABLE 2. Infiltration loss with 15-mi/h (24 km/h) outside wind

Building item	Infiltration, $ft^3/(ft)(h)$
Double-hung unlocked wood sash windows of average tightness, non-weather-stripped including wood frame leakage	39
Same window, weather-stripped	24
Same window poorly fitted, non-weather-stripped	111
Same window poorly fitted, weather-stripped	34
Double-hung metal windows unlocked, non-weather-stripped	74
Same window, weather-stripped	32
Residential metal casement, $^1/_{64}$ -in. (0.4-mm) crack	33
Residential metal casement, $^1/_{32}$ -in. (0.8-mm) crack	52

doors. Illustrative quantities of air leakage for various types of window construction are shown in Table 2. The figures given are in cubic feet of air per foot of crack per hour.

Design. Before a heating system can be designed, it is necessary to estimate the heating load for each room so that the proper amount of radiation or the proper size of supply air outlets can be selected and the connecting pipe or duct work designed. *See* AIR REGISTER; CENTRAL HEATING AND COOLING; HOT-WATER HEATING SYSTEM; OIL BURNER; RADIATOR; STEAM HEATING; WARM-AIR HEATING SYSTEM.

Heat is released into the space by electric lights and equipment, by machines, and by people. Credit to these in reducing the size of the heating system can be given only to the extent that the equipment is in use continuously or if forced ventilation, which may be a big heat load factor, is not used when these items are not giving off heat, as in a factory. When these internal heat gain items are large, it may be advisable to estimate the heat requirements at different times during a design day under different load conditions to maintain inside temperatures at the desired level.

Cost of operation. Design and selection of a heating system should include operating costs. The quantity of fuel required for an average heating season may be calculated from

$$F = \frac{Q \times 24 \times \mathrm{DD}}{(t_i - t_o) \times \ \mathrm{Eff} \ \times H}$$

where F = annual fuel quantity, same units as H
Q = total heat loss, Btu/h (or J/s)
t_i = inside design temperature, °F (or °C)
t_o = outside design temperature, °F (or °C)
Eff = efficiency of total heating system (not just the furnace) as a decimal
H = heating value of fuel
DD = degree-days for the locality for 65°F (19.3°C) base, which is the sum of 65 (19.3) minus each day's mean temperature in °F (or°C) for all the day's of the year

If a gas furnace is used for the insulated house of Table 1, the annual fuel consumption would be

$$F = \frac{56,500 \times 24 \times 4699}{[75 - (-5)] \times 0.80 \times 1050}$$
$$= 94,800 \ \mathrm{ft}^3 \ \ (2684 \mathrm{m}^3)$$

For a 5°F (3°C), 6–8-h night setback, this consumption would be reduced by about 5%. *See* THERMOSTAT.

Gayle B. Priester

Bibliography. American Society of Heating, Refrigerating, and Air Conditioning Engineers, *Handbook of Fundamentals*, 1993; B. C. Langley, *Comfort Heating*, 4th ed., 1994.

Comfort temperatures

Air temperatures adjusted to represent human comfort or discomfort under prevailing conditions of temperature, humidity, radiation, and wind. Theoretical formulas attempt to compare the rate of heat loss to surroundings with rate of heat production by work and metabolism. Most modern empirical relations, based on relative comfort expressed by human subjects under differing atmospheric combinations, attempt to indicate the temperature at which air at some standard humidity, air motion, and radiation load would be just as uncomfortable (or comfortable). Many former indices, however, had arbitrary scales.

Heat is produced constantly by the human body at a rate depending on muscular activity. For body heat balance to be maintained, this heat must be dissipated by conduction to cooler air, by evaporation of perspiration into unsaturated air, and by radiative exchange with surroundings. Air motion (wind) affects the rate of conductive and evaporative cooling of skin, but not of lungs; radiative losses occur only from bare skin or clothing, and depend on its temperature and that of surroundings, as well as sunshine intensity.

As air temperatures approach body temperature, conductive heat loss decreases and evaporative loss increases in importance. Hence, at warmer temperature, humidity is the second-most important atmospheric property controlling heat loss and hence comfort, and the various sensible temperature formulas incorporate some humidity measure. *See* BIOMETEOROLOGY.

Temperature and humidity. Almost a century ago, wet-bulb temperature was thought to indicate human comfort under warm conditions, especially in deep mines, ship engine rooms, and steel mills and other factories, on the assumption that sweaty skin approximated a wet bulb. It was superseded after 1923 by the effective temperature, which combined dry- and wet-bulb temperatures, at constant low-ventilation rate, as evaluated by 300 trained subjects walking between controlled-environment rooms. In the 1950s this was approximated by the discomfort index, soon renamed temperature-humidity index (THI) and used by the Weather Bureau (now National Weather Service) for several years. In Canada it was called humidity index. It is calculated by either Eq. (1) or Eq. (2), where the air

$$\mathrm{THI} = 15 - 0.4(t_a + t_w) \qquad (1)$$

$$\mathrm{THI} = t_a - 0.55(1 - \mathrm{RH})(t_a - 58) \qquad (2)$$

and wet-bulb temperatures—t_a and t_w respectively—are expressed in °F and relative humidity (RH) is expressed as a decimal fraction. Except in saturated air (RH = 1.0, that is, 100%), the temperature-humidity index is usually lower than air temperature. A different temperature-humidity index was used in 1980

by another National Weather Service office to compute a heat wave index, ranging from 0 (mild) to 15 (extreme).

A simple average of air temperature in °F and relative humidity in percent was called humiture in 1937, and was changed in 1959 to a simple average of dry- and wet-bulb temperatures in °F. Later, this was further changed to an average of air temperature in °F and excess vapor pressure e, above 10 millibars, with the saturation vapor pressure at 45°F. This index is called humidex in Canada. The originator of humidex opposed raising the threshold to 21 mb (saturation at 65°F), which was being used by at least one television weatherperson. No two of these formulations are equivalent, but each is usually less than air temperature. *See* AIR TEMPERATURE; HUMIDITY; PSYCHROMETRICS.

Wind and radiation. Meanwhile, more refined measurements led to more elaborate formulas involving all aspects of heat loss. Operative temperature modifies effective temperature to include radiation at constant wind speed, and standard operative temperature considers also temperature and heat conduction of skin. The 1955 heat stress index, essentially the ratio of evaporation needed to maintain a stable body temperature to the evaporation possible under existing conditions and working rate, has been evaluated by a detailed computer program. The 1979 apparent temperature, a very complete formulation, is calculated by use of a nomograph or computer as the heat index of the National Oceanic and Atmospheric Administration (see **illus.**). It is a nonlinear combination of air temperature and relative humidity, adjusted for wind.

A less detailed graphic computation yields a similar but not equivalent value known as humisery. Another approach uses the shaded comfort-factor thermometer, which is a large bulb half-covered by a wetted cloth so that air temperature, wet-bulb temperature, wind, and radiation are integrated into a single number similar to a Fahrenheit temperature; most people are said to be comfortable at values from 63 to 72.

Cold conditions. Under cold conditions, atmospheric moisture is negligible and wind becomes important in heat removal. Wind chill, given by Eq. (3),

$$H = (10.45 + 10\sqrt{v} - v)(33 - t) \qquad (3)$$

where H represents heat loss in kcal/(m^2)(h^2), does not estimate sensible temperature as such, but rather heat loss in kcal/(m^2)(h) for wind speed v in m/s and air temperature t in °C.

The formula was based on the rate of freezing, at various temperature-wind combinations during the Antarctic winter, of water in a plastic cylinder, not on responses of human subjects. It has never been validated by the latter approach. Nevertheless, the temperature at which the estimated heat loss at some reference wind speed equals that at the existing temperature-wind combination is the wind chill temperature. The reference wind speed is variously 3 mi/h, 3 knots, or 4 mi/h, and in a revised formulation is 0.

Arnold Court

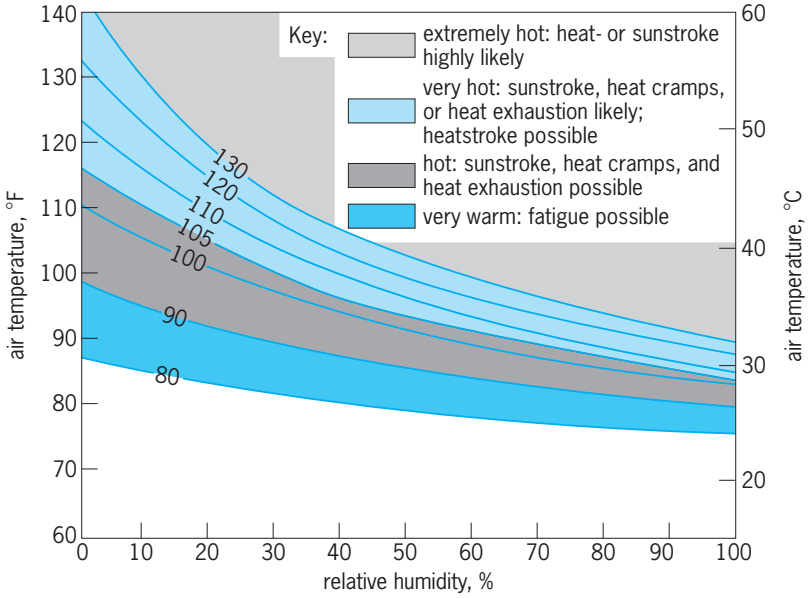

Heat index used by the National Weather Service, National Oceanographic and Atmospheric Administration. The numerical values on each curve represent a heat index, that is, apparent temperature in °F.

Bibliography. P. O. Fanger, *Thermal Comfort*, 1972, reprint 1982; H. Landsberg (ed.), *World Survey of Climatology*, vol. 3: *General Climatology*, 1981.

Commelinales

An order of flowering plants, division Magnoliophyta (Angiospermae), which is included in the subclass Commelinidae in the class Liliopsida (monocotyledons). It consists of 4 families and about 1000 species, the bulk of which is in the families Commelinaceae (about 700 species) and Xyridaceae (about 200 species). The order is marked by its flowers that ordinarily have the perianth well differentiated into sepals and petals but do not have nectaries or nectar. The ovary is consistently superior and the fruit capsular. The wandering Jew (species of *Tradescantia* and *Zebrina* in the Commelinaceae) belongs to this order. *See* COMMELINIDAE; LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Commelinidae

A subclass of the class Liliopsida (monocotyledons) of the division Magnoliophyta (Angiospermae), the flowering plants, consisting of 7 orders, 16 families, and nearly 15,000 species. The orders include Commelinales, Eriocaulales, Restionales, Juncales, Cyperales, Hydatellales, and Typhales. For further information see separate articles on each order.

These monocotyledons are syncarpous (the carpels are united in a compound ovary) or

pseudomonomerous (reduced to a single carpel from a syncarpous ancestry). The endosperm is usually starchy, and the perianth is either well differentiated into sepals and petals or more or less reduced and not petallike. The stomates have two or more subsidiary cells, the pollen is either binucleate or more often trinucleate, and the endosperm may be nuclear. Many of the families have well-developed vessels in all vegetative organs. Several of the orders of Commelinidae have often been treated as a single order Farinosae or Farinales. *See* LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Common cold

An acute infectious disorder characterized by nasal obstruction and discharge that may be accompanied by sneezing, sore throat, headache, malaise, cough, and fever. The disorder involves all human populations, age groups, and geographic regions; it is more common in winter than in summer in temperate climates. Most people in the United States experience at least one disabling cold (causing loss of time from work or school or a physician visit) per year. Frequencies are highest in children and are reduced with increasing age.

Most, or possibly all, infectious colds are caused by viruses. More than 200 different viruses can induce the illness, but rhinoviruses, in the picornavirus family, are predominant. Rhinoviruses are small ribonucleic acid–containing viruses with properties similar to polioviruses. Other viruses commonly causing colds include corona, parainfluenza, influenza, respiratory syncytial, entero, and adeno. *See* ADENOVIRIDAE; ENTEROVIRUS; PARAINFLUENZA VIRUS; RHINOVIRUS.

Cold viruses are spread from one person to another in either of two ways: by inhalation of infectious aerosols produced by the sneezing or coughing of ill individuals, or by inoculation with virus-containing secretions through direct contact with a person or a contaminated surface. Controlled experiments have not shown that chilling produces or increases susceptibility to colds. Infection in the nasopharynx induces symptoms, with the severity of the illness relating directly to the extent of the infection. Recovery after a few days of symptoms is likely, but some individuals may develop a complicating secondary bacterial infection of the sinuses, ear, or lung (pneumonia).

Colds are treated with medications designed to suppress major symptoms until natural defense mechanisms terminate the infection. Immunity to reinfection follows recovery and is most effective in relation to antibody in respiratory secretions. There is no established method for prevention of colds; however, personal hygiene is recommended to reduce contamination of environmental air and surfaces with virus that may be in respiratory secretions. *See* PNEUMONIA.

Robert B. Couch

Bibliography. B. N. Fields et al. (eds.), *Virology*, 3d ed., 1996; G. J. Galasso et al. (eds.), *Antiviral Agents and Human Viral Diseases*, 4th ed., 1997; G. L. Mandell, *Principles and Practice of Infectious Diseases*, 5th ed., 1999.

Communications cable

A cable that transmits information signals between geographically separated points. The heart of a communications cable is the transmission medium, which may be optical fibers, coaxial conductors, or twisted wire pairs. A mechanical structure protects the heart of the cable against handling forces and the external environment. The structure of a cable depends on the application.

Optical communications cables. These cables are used in both terrestrial (**Fig. 1**) and undersea (**Fig. 2**) systems. Optical communications cables for terrestrial use may be installed aerially, by direct burial, or in protective ducts. The terrestrial and undersea cables differ in several respects. First, the terrestrial cable requires only enough longitudinal strength to support its own weight over relatively short pole-to-pole spans, or to allow installers to pull the cable into ducts or lay it in a trench. For the undersea

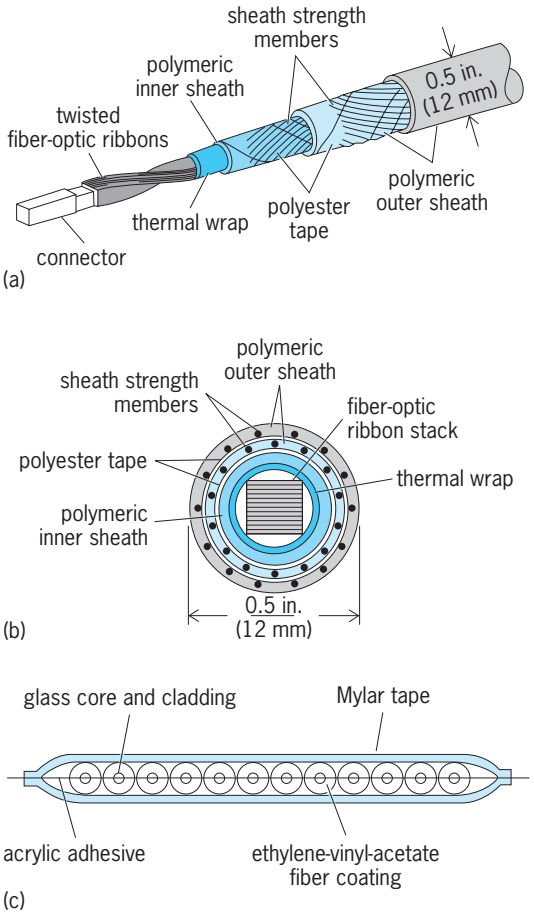

Fig. 1. Multifiber cable for terrestrial use (144 fibers maximum). (*a*) Cutaway view. (*b*) Cross section of cable. (*c*) Cross section of single optical fiber ribbon (12 fibers).

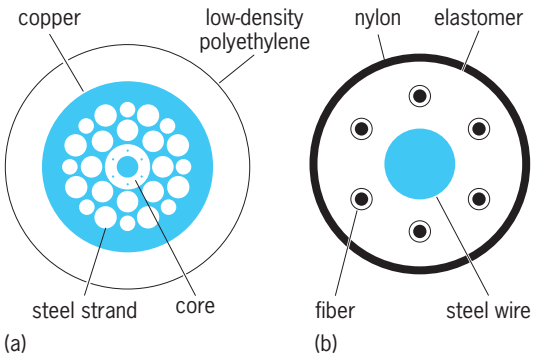

Fig. 2. Optical-fiber cable for undersea systems. (*a*) Cross section. (*b*) Detail of cross section of core.

TABLE 1. Principal characteristics of the first transoceanic fiber-optic United States-to-Europe undersea cable system

Characteristic	Value
Line rate	295.6 megabits/s
Number of service fibers	4
Number of protection fibers	2
Regenerators per repeater	6
Number of voice channels	8000
Number of conversations (with digital circuit multiplication equipment)	40,000
Optical wavelength	1310 nm

cable, the high-strength steel strand allows it to be laid and recovered in ocean depths up to 4.5 mi (7315 m). A second difference is that the undersea cable must deliver as much as 1.6 amperes of direct current to the spaced, series-connected repeaters. Repeaters for the terrestrial system are generally located in buildings, and are powered from locally available sources. A third difference is due to the environment. Terrestrial cables are often installed in ducts. Hence, only limited protection is required against local hazards. The undersea cable, on the other hand, must withstand moderately severe handling in laying, and it sustains the full pressure of ocean depths. At 4.5 mi (7315 m) depth, this pressure is 10,700 lb/in.^{2} (73.8 megapascals). To withstand this pressure, the cable is a filled design (no voids). The thick polyethylene outer layer provides electrical insulation for the power-carrying copper conductor, and is tough enough to resist severe handling. In addition to providing longitudinal strength, the steel strand overlaid with copper acts as a pressure cage to protect the core and fibers. *See* ELECTRICAL INSULATION; POLYOLEFIN RESINS; SUBMARINE CABLE.

Optical communications cables are often used to carry input and output data to computers, or to carry such data from one computer to another. Then they are generally referred to as optical data links or local-area networks. The links are generally short enough that intermediate regeneration of the signals is not needed. *See* FIBER-OPTIC CIRCUIT; LOCAL-AREA NETWORKS.

Optical fibers. Signals in these cables are carried by light pulses which are guided down the optical fiber. In most applications, two fibers make up a complete two-way signal channel. The guiding effect of the fiber confines light to the core of the glass fiber and prevents interference between signals being carried on different fibers. The guiding effect also delivers the strongest possible signal to the far end of the cable. Exceptionally pure silica glass in the fiber minimizes light loss for signals passing longitudinally through the glass fiber. Developers have realized fiber losses of 0.27 dB or less per statute mile (0.17 dB per kilometer), at a wavelength of 1550 nanometers. For such a fiber, the signal can travel 11 mi (18 km) before losing half its power, and can travel 75 mi (120 km) or more before requiring electronic regeneration or amplification. Coaxial cable systems require regeneration at much closer intervals. In the L5 and T4M Coaxial Cable systems, for example, signals are amplified or regenerated every mile (every 1.6 km). *See* OPTICAL FIBERS.

Signal regeneration. Optical cable systems are usually digital. Thus, information is coded into a train of off-or-on light pulses. These are detected by a photodetector at the far end of a cable span and converted into electronic pulses which are amplified, retimed, recognized in a decision circuit, and finally used to drive an optical transmitter. In the transmitter, a laser converts the electric signals back into a train of light pulses which are strong enough to traverse another cable span. By placing many spans in tandem, optical cable systems can carry signals faithfully for thousands of miles.

In TAT-8, the first transoceanic fiber-optic cable, which entered service in 1988 (**Table 1**), six regenerators are mounted in a repeater housing (**Fig. 3**). Thus, the repeater regenerates signals in three pairs of fibers.

Optical-fiber amplifiers. Rather than undersea regenerators, current optical-fiber cable systems use erbium-doped fiber amplifiers (EDFAs) to boost the optical signal on long spans. Conversion from optical to electronic modes and back again is then not needed in the undersea repeaters. Systems using optical amplifiers are simple and reliable, but still deliver a near-perfect, low-noise signal because of signal regeneration at the receiving terminal. With careful design, it is possible to carry a 5-gigabit/s signal across the Pacific Ocean, a 5500-mi (9000-km) distance, with EDFAs before regeneration on shore. Such systems deliver the signal to the distant shore with a bit-error ratio of better than 10^{-10}.

Repeaterless undersea cables. The trend toward lower fiber loss allows the linking of mainland to offshore islands without undersea electronics and optics. Cable systems of over 155 mi (250 km) are feasible (**Table 2**), with future capability trending toward 190 mi (300 km).

Advantages. Optical communications cables have five key advantages over earlier metallic cables: noise- and error-free transmission; greater repeater spacings; enormous information capacity (many individual voice, data, or television channels); immunity

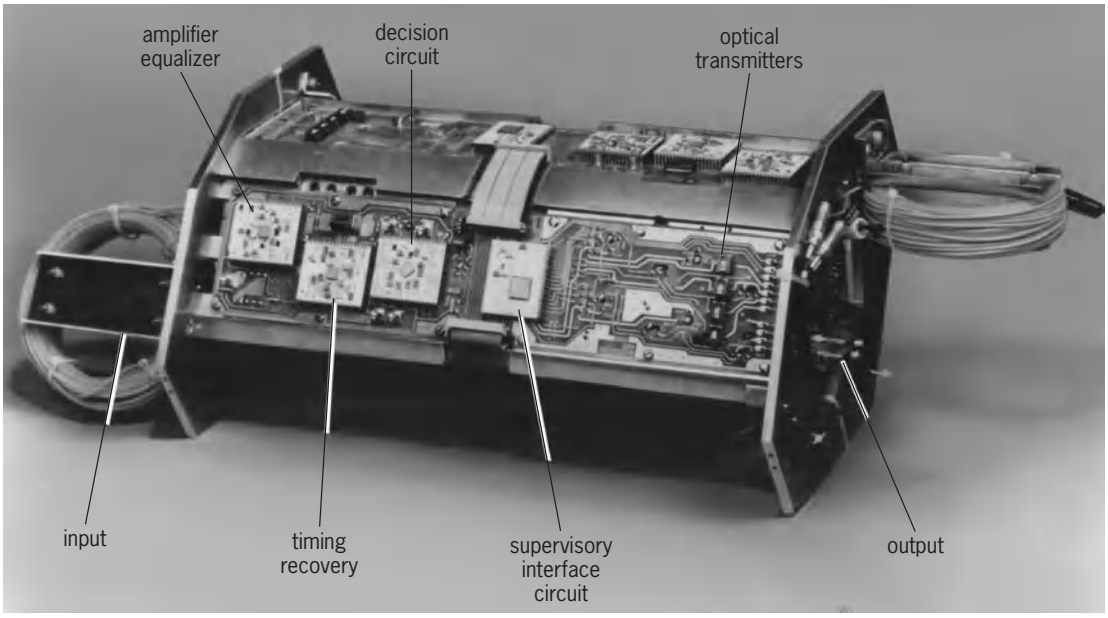

Fig. 3. Regenerator for SL-280 optical undersea cable communications system used in TAT-8.

from outside interference (other signals or noise); and smaller size and lower weight.

Because of these advantages and because optical cable technology is in its infancy, many additional applications can be expected to develop, and the cost of such communications cables can be expected to continue to decrease. *See* OPTICAL COMMUNICATIONS.

Coaxial cable communications systems. Coaxial communication systems evolved before optical systems. Most of these systems are analog in nature. Signals are represented by the amplitude of a wave representing the signal to be transmitted. In a multichannel system, each voice, data, or picture signal occupies its unique portion of a broadband signal which is carried on a shared coaxial conductor or "pipe." In the transmitting terminal, various signals are combined in the frequency-division transmitting multiplex equipment. At the receiving end of a link, signals are separated in the receiving demultiplex equipment. This combining and separation operates much as broadcast radio and television do, and the principles are identical. *See* AMPLITUDE MODULATION; ELECTRICAL COMMUNICATIONS; FREQUENCY MODULATION.

TABLE 2. Repeaterless cable system

Characteristic	Value
System length	188 mi (300 km) maximum
Fiber type	SL, dispersion shifted
Fiber proof stress	200,000 lb/in.^{2} (1.38 GPa)
Cable design	SL, embedded core
Optical wavelength	1.55 μm
Optical source	Buried-heterostructure multiple quantum well laser and EDFA*
Optical detector	EDFA and avalanche photodiode
Line data rate	5 gigabits/s
Line coding	Nonreturn to zero (NRZ)

*EDFA = erbium-doped fiber amplifier.

In order to carry as many channels as possible, the multiplexing and demultiplexing use single-sideband (SSB) transmission. The fixed carrier signals, which do not convey useful information, are suppressed in the transmitting multiplex and reintroduced in the receiving multiplex. This avoids unnecessary load on the system. *See* SINGLE SIDEBAND.

Terrestrial cables. The channel capacity of terrestrial coaxial systems increased dramatically over the years to meet increased long-distance calling (**Table 3**). This resulted in decreased cost per channel, which stimulated still further demand. While demand was the driving force, it is improving technology which allowed such tremendous increases in capacity.

The L1 and L3 systems used electron tubes as the amplifying device, while all later systems used transistors. The L5 system benefited by the use of thin-film integrated circuits. These advances were responsible for the 22-fold increase in the number of channels per coaxial pair (Table 3). Because the number of coaxial pairs per cable was also increased, the total route capacity was increased by a factor of 73 over the 30-year period. *See* ELECTRON TUBE; INTEGRATED CIRCUITS; TRANSISTOR.

Repeater spacing decreases as the capacity rises (Table 3). Multiplying channel capacity by 22 has meant going to higher frequencies, since each additional channel requires another 4 kHz of spectrum. At higher frequencies the cable has higher loss, and therefore such signals must be amplified at closer intervals.

Undersea cables. The history of undersea communications goes back to 1858, when the first transatlantic telegraph cable was completed by C. W. Field. This cable had no repeaters and thus was capable of carrying only a single slow-speed telegraph signal. Nevertheless, it allowed news to be communicated

TABLE 3. L-system characteristics

Characteristic	L1	L3	L4	L5	L5E
Service date	1946	1953	1967	1974	1976
Repeater spacing, mi (km)	8 (12.9)	4 (6.4)	2 (3.2)	1 (1.6)	1 (1.6)
Voice-band channel capacity per 3/8-in. (9.5-mm) coaxial pair	600	1860	3600	10,800	13,200
Coaxial pairs	4	6	10	11	11
Working pairs	3	5	9	10	10
Channel capacity	1800	9300	32,400	108,000	132,000

between North America and Europe without the long delay of surface mail. While telegraph cables were improved over the years, the next major breakthrough occurred in 1956, when the first transatlantic telephone cable was completed from Nova Scotia to Newfoundland, and then across the Atlantic to Scotland.

With the completion of TAT-1 (Transatlantic Telephone-1), callers could for the first time talk across the Atlantic Ocean as easily as they could within the continents. The response was a dramatic increase in use of overseas calling. Since 1956, cables have been installed throughout the world. England, France, Japan, and the United States have been active in designing and installing undersea communications cables.

Low delay is one respect in which undersea cables have a significant edge over satellites in a geosynchronous orbit. Delay is usually evaluated in terms of round-trip message time. For a telephone conversation, it is the extra time added between question and answer because of the transmission time. For the satellite connection, this delay is about 0.5 s. For a 4000-statute-mile (6400-km) undersea cable link, delay is only 0.06 s. Longer delay has two effects: it makes echo more bothersome to voice communications, and it requires more complex and less efficient protocols for data communications. *See* COAXIAL CABLE; TELEPHONE SERVICE; TELEPHONE SYSTEMS CONSTRUCTION.

S. Theodore Brewer

Bibliography. E. Desurvire, *Erbium-Doped Fiber Amplifiers: Principles and Applications*, 2002; R. L. Freeman, *Reference Manual for Telecommunications Engineering*, 3d ed., 2 vols., 2002; J. Gowar, *Optical Communication Systems*, 2d ed., 1993; J. Senior, *Optical Fiber Communications*, 2d ed., 1992, paper 1993; W. Tomasi, *Electronic Communications Systems: Fundamentals Through Advanced*, 5th ed., 2003.

Communications satellite

An artificial device that circumnavigates the Earth, receiving and retransmitting radio-frequency signals to multiple points on Earth. The communications hardware includes antennas, receivers, signal processing, and power amplifiers. The performance is predicted by calculating the power received by the satellite and by the receiving ground station. The satellite has solar arrays and batteries to provide electrical power, thrusters to keep the satellite in the desired position and orientation, and radiators to dissipate heat. The equipment is designed to operate unattended for many years. Many communications satellites are put in a special circular orbit high above the equator.

Communications satellites provide worldwide exchange of voice, text, data, and images to both fixed stations and mobile stations in support of aeronautical, land, and maritime services. With the advent of more powerful satellites and smaller, less powerful, and even portable terminals, satellite broadcasts are accessible to users in both highly populated and remote areas of the world. Applications are proliferating—from distributing network programming to television and radio stations or directly to subscribers, to providing essential communications for disaster relief or medical treatment, supporting Internet access and e-mail, or even to connecting travelers on handheld phones. By providing reliable communications among diverse and geographically distant populations, satellites have helped to make the world a global community. Communications satellites represent at least 40% of the total number of satellites and 30% in cost.

For a communications satellite to provide continuous service 24 h every day, it must remain above the Earth station's horizon, and preferably remain stationary in one spot in the sky so that antennas on the ground do not have to be repointed. As the Earth turns, the satellite moves in a circle, and the center of this circle must be at the center of the Earth. This can happen only if the satellite is above and perpendicular to a spot on the Equator. The height of the satellite is also critical. If the satellite is too low, the attraction of gravity is too strong and the satellite will accelerate toward the Earth. If the satellite is too high, the gravitational attraction is too weak and the satellite will accelerate away from the Earth. If the satellite is 35,786 km (22,237 mi) above the Earth's surface, the attraction of gravity is exactly right. Except for a few small disturbing forces, the satellite will remain at that height, over a point on the Earth's Equator. Such a satellite is called a geostationary satellite and, as the Earth turns, it traverses a path called a geostationary orbit. To an observer on Earth, the satellite remains stationary in the sky.

Basic configuration. Most communications satellites are active repeaters. That is, they receive a signal from an Earth station, amplify it, and retransmit it to one or more earth stations (**Fig. 1**). The signal

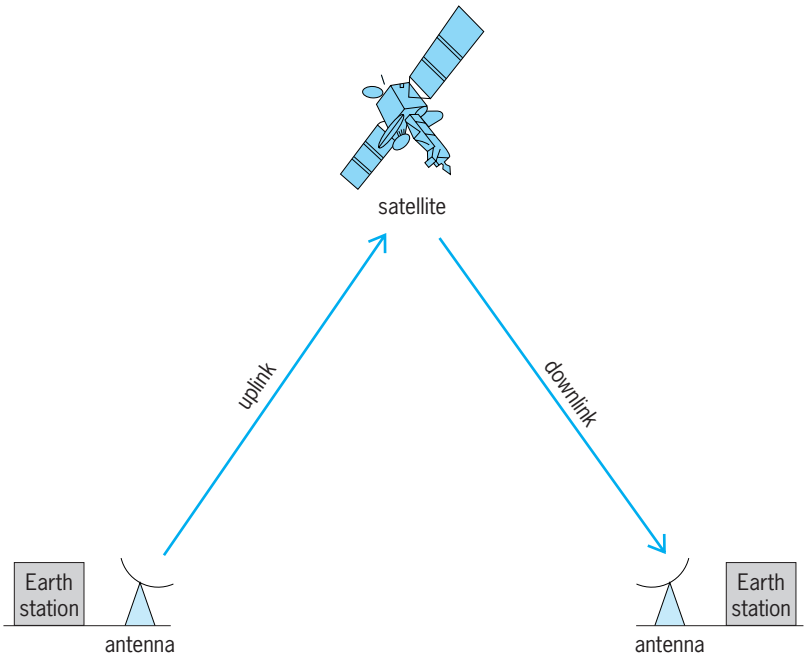

Fig. 1. Communications to and from a satellite.

from the first Earth station to the satellite is called the uplink, and the signal from the satellite back to another Earth station is the downlink. In the satellite, the equipment used to handle the signals is the communications subsystem. The signal is received by an antenna, amplified, processed through many stages, and immediately transmitted through an antenna back to Earth.

A typical communications satellite has a compact, rectangular body, with antennas and solar arrays unfolded in space (**Fig. 2**). The large antennas are used for both transmitting and receiving communications. A smaller antenna is often used to receive commands and to send information about the satellite's status. Before launch, the solar panels are folded next to the body. After launch, they are extended into two long arrays. Other external parts include the thrusters for propulsion, sensors to detect orientation, and radiators to dissipate excess heat.

Inside the satellite, the communications subsystem includes the receiver, the processing equipment, and the power amplifier. The receiver includes a low-noise amplifier to minimize the noise added to the signal, and a downconverter to shift the carrier frequency to another frequency. The processing equipment may include switches, attenuators, and multiplexers. The power amplifiers provide the maximum power to be transmitted.

To support communications, the satellite has several specialized components. These provide the electric power (solar array, batteries, and power conditioning equipment), orientation (sensors and wheels), propulsion (propellant tanks, fuel, and thrusters), structure, and thermal control. Telemetry provides information about the satellite, the command subsystem is used to control the satellite, and tracking is used to help determine its orbit. Since ordinary repairs cannot be done after a satellite is launched, high reliability is essential. Frequently, duplicate equipment is available in the satellite, and used if the primary equipment fails.

Communications Subsystem

The function of a communications satellite is accomplished by the communications subsystem, which receives, amplifies, processes, and transmits

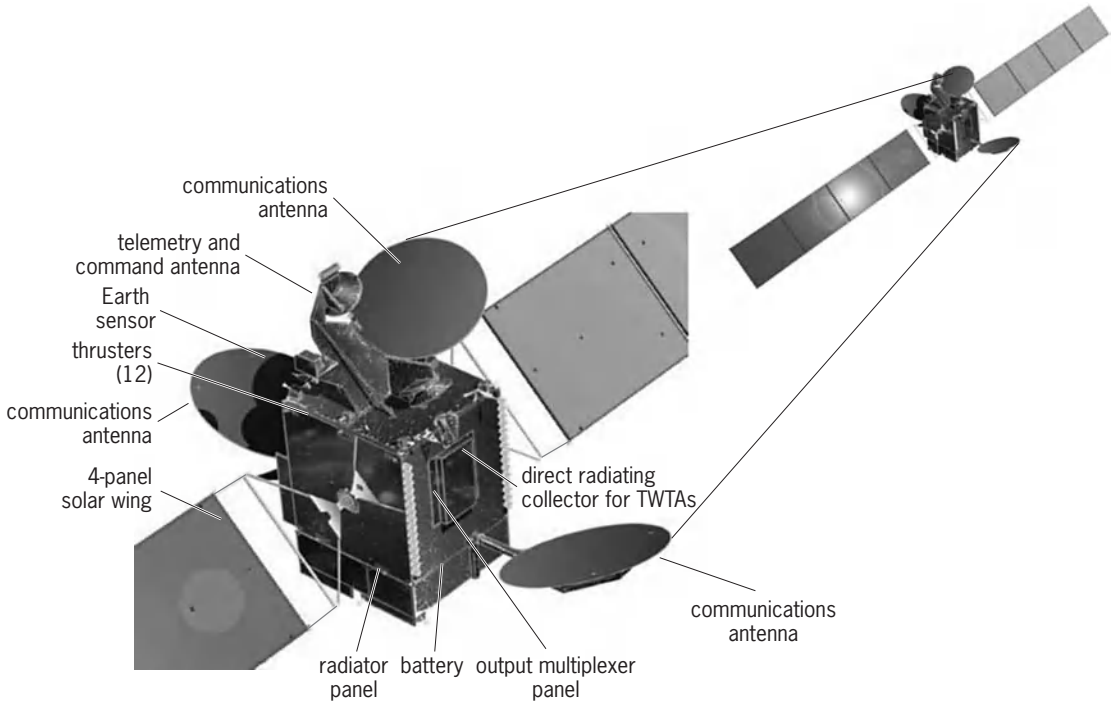

Fig. 2. Typical communications satellite. TWTA = traveling-wave-tube amplifier. (*Space Systems/Loral*)

TABLE 1. Frequencies allocated to communications satellites*

Frequency range, GHz	Letter designation	Typical service usage
1.0–2.0	L-band	Mobile satellite service
1.55–3.9	S-band	Mobile satellite service, digital audio radio service
3.7–6.2	C-band	Fixed satellite service, very small aperture terminal
8.0–12.0	X-band	Military/satellite imagery
11.7–14.5	Ku-band	Fixed satellite service, direct broadcast satellite, very small aperture terminal
17.7–21.2	Ka-band	Fixed satellite service "broadband," intersatellite link
27.5–31.0	Ka-band	Fixed satellite service "broadband," intersatellite link

*From Richard Barnett.

information. Hardware includes antennas, receivers, frequency converters, multiplexers, and power amplifiers. The properties of these components depend on the radio frequencies used, and the way the information is imposed on the carrier frequency.

Radio frequencies. A variety of frequencies are received and transmitted by communications satellites (**Table 1**). The fixed satellite service (FSS) provides the major communications between fixed Earth stations. The mobile satellite service (MSS) communicates with ships, vehicles, and airplanes. Direct broadcast satellite (DBS) transmits television directly to homes. Digital audio radio service (DARS) provides continuous radio programs to vehicles. Very small aperture terminals (VSATs) provide communications to very small antennas, often used by businesses. And intersatellite links carry communications from one satellite to another. *See* SATELLITE RADIO; SATELLITE TELEVISION BROADCASTING.

Communications satellites use frequencies above 1 GHz (1,000,000,000 Hz or 1,000,000,000 cycles per second). The commonly used radio frequencies are lower. The AM radio band is around 1000 kHz (or 0.001 GHz) and the FM radio band is around 100 MHz (or 0.1 GHz). Frequencies from 0.3 to 3 GHz are also called ultrahigh frequencies (UHF), and frequencies from 3 to 30 GHz are called superhigh frequencies (SHF).

The amount of radio spectrum (bandwidth) available for space communications is limited. Bandwidth is a measure of how much information can be transmitted. A telephone call, or a modem connection to the Internet, takes only about 10 kHz. A television channel, or a high-speed connection to the Internet for transmission of pictures, requires much more bandwidth. The demand for satellite communications continues to increase, and the lower frequencies are in high demand. The higher frequencies (above 15 GHz) are more available, but are also more attenuated by rain, which can both decrease the signal power and increase the receiver noise level. *See* BANDWIDTH REQUIREMENTS (COMMUNICATIONS); RADIO-WAVE PROPAGATION.

There are different ways to conserve bandwidth by reusing frequencies in a single satellite. One way is to use narrow antenna beams on a satellite. If one beam covers Europe and the other covers North America, different signals can be sent on the different beams at the same frequency, using the antenna directivity for isolation. Another way is to use polarized beams, with the same frequency used for two different polarizations. Antennas can discriminate against an undesired signal, just as polarizing sunglasses can reduce the glare from horizontal surfaces. In each of these ways, the equipment is designed so that the unwanted signal is no more than 1/1000 of the desired signal.

Multiple satellites can use the same frequencies—if the Earth station antennas can discriminate between adjacent satellites. The Earth station antenna must have a beam that includes only a single satellite. This applies to both the uplink and the downlink. The signal from the ground must be received by only one satellite, and the signal to the ground must be received from only one satellite. In geostationary orbit, some neighboring satellites, using identical frequencies, are operated with only a 2° separation, which is 1470 km (915 mi).

Finally, another method, called compression, packs more information into less bandwidth. This technique takes advantage of redundancy in the signal. Voice activation uses a voice channel only when a person is speaking, typically only 40% of the time. Early television compression made use of the time it took an electron beam to return to the start of another scanning line. More compression is possible using the time between frames. With digital data, more efficient modulation methods allow more information bits per unit of time. Packet switching, used in the Internet, allows a variety of different signals originating and destined for different users to share the same communications channel. With many different technical improvements, a single communications satellite can now broadcast a hundred different television channels to small home antennas. Earlier satellites could broadcast only a few channels to Earth stations with large antennas. *See* DATA COMPRESSION; PACKET SWITCHING.

The International Telecommunications Union (ITU), working with various national agencies such as the Federal Communications Commission (FCC), controls the use of radio frequencies by specifying the frequency ranges and imposing limitations on power. Some frequencies used by satellites were first used by microwave relay systems, which transmit a signal through a series of relay towers, typically on the tops of hills. Since microwave relay systems preceded satellites, satellites must limit the power per

area incident on these towers to avoid causing harmful interference. *See* RADIO SPECTRUM ALLOCATIONS.

Modulation, multiplexing, and multiple access. To send a signal through a communications satellite, the information must be superimposed on a high-frequency (RF) signal (carrier). This is called modulation. If the carrier is not modulated, no information is transmitted and only a constant signal is received. A satellite will normally relay many signals from an Earth station. Multiplexing is used to combine these signals and demultiplexing to separate them. Finally, a satellite often handles signals from more than one Earth station; this is called multiple access.

There are several methods of modulation, multiplexing, and multiple access that may be used in various combinations. Multiplexing and multiple access require sharing the satellite resources to carry various signals. These resources can be shared by using different frequencies, by taking turns (time division), by using different antenna beams, or by using different polarizations. Signals such as telephone and television may be analog or digital. Theoretically, any method can be used for either type of signal. In practice, time division is easier with digital signals, while frequency division is easier with analog signals.

Modulation. There are three basic modulation methods, which change the amplitude, frequency, or phase of a carrier. In amplitude modulation (AM), the amplitude of the carrier is varied in proportion to the input signal. In frequency modulation (FM), the frequency of the carrier is varied in proportion to the input signal. FM is less susceptible to noise and interference than AM. In phase modulation (PM), also called angle modulation, the phase of the carrier is varied in proportion to the input signal. This also varies the frequency of the carrier, but much less than FM, since the variation is never more than one cycle.

Phase-shift keying (PSK) is a digital version of phase modulation in which the phase of the carrier is changed depending on certain bit sequences. This is the most common form of modulation used on communications satellites today. *See* AMPLITUDE MODULATION; FREQUENCY MODULATION; MODULATION; PHASE MODULATION.

Multiplexing. Most communications satellites handle many signals. Multiplexers are used to combine signals before they reach an Earth station, at the Earth station, or in the satellite. The most common forms of multiplexing used in satellite communications are frequency-division multiplexing (FDM) and time-division multiplexing (TDM). As a simple illustration, in the case of a radio or television receiver, different broadcast stations are transmitted at different frequencies (FDM), while different programs may be on one station, but received at different times (TDM).

Frequency-division multiplexing is often used to combine telephone channels. This may be done before the signals reach the Earth station, and then the resulting trunks may be multiplexed together at the station. Each signal is converted to an assigned frequency band, and then the signals can be combined. For example, if two phone channels have frequencies of 0–10 kHz, one channel can stay at the same frequency and another channel can be converted to 10–20 kHz. These two channels can be combined into one 0–20 kHz signal. Many other channels can be converted to other frequency bands and combined. The combination can then be sent up to a satellite and down to another Earth station, where demultiplexers reverse the process and separate the different telephone channels.

Time-division multiplexers were introduced in telephone networks when voice calls were digitized. At 64,000 bits/second, the voice signal can be sampled 8000 times per second, and each sample digitized to 8-bit accuracy. These digitized voice signals can be multiplexed with other voice channels to form digital streams with many more bits/second, extending to standardized streams as high as billions of bits per second, for use in communications satellites, as well as fiber optic and microwave services. Once information is digitized, it can be multiplexed with any other type of digitized information.

Multiple access. Several Earth stations can send signals, including voice, data, teletype, facsimile, television, and digitized data, simultaneously to the same satellite. This is called multiple access. Multiple access is similar to multiplexing signals into a satellite, but the different Earth stations must coordinate their transmissions. There are several ways to share access to a satellite: dividing the frequency band (frequency-division multiple access or FDMA); dividing access in time (time-division multiple access or TDMA); dividing access in space (space-division multiple access or SDMA) via spot beams or polarizations, for example; or by using different codes to divide the access (code-division multiple access or CDMA).

In FDMA, each Earth station has exclusive use of an assigned bandwidth between specified frequencies. The strength of the signals must be balanced, so that a strong signal at one frequency does not interfere with a weak signal at a neighboring frequency. Individual carriers may also have different power, bandwidth, and quality requirements, which all must be balanced so that each station is satisfied.

In TDMA, each Earth station has an assigned time to use the entire bandwidth or a portion of it. A prearranged schedule is used to avoid overlapping transmissions so that at any time the satellite has only one signal present within each TDMA channel. TDMA systems are flexible, and can accommodate traffic variations among stations by increasing or decreasing the duration of each station's bursts to match the traffic flow.

Random multiple access (RMA) is a variation of TDMA in which the available time is not assigned to individual stations. This system, also called ALOHA, was devised at the University of Hawaii. It allows small, inexpensive Earth stations to communicate through a satellite with a minimum of protocols and no network supervision. When a station needs to transmit, it sends a burst of data pulses, and then watches the downlink. If it sees a corrupted signal, it knows that another station is transmitting. It then

waits a short, random time before trying again. If there is not too much traffic, the signal is eventually transmitted with no interference.

In CDMA, the transmitting Earth station uses a unique digital code to spread each signal over a bandwidth much wider than is actually needed. The receiving station uses the same code to retrieve the information. Other networks may operate simultaneously within the same bandwidth, as long as different, noninterfering (orthogonal) codes are used for different signals. *See* MULTIPLEXING AND MULTIPLE ACCESS.

Antennas. Antennas are passive devices used to collect the signal (radio-frequency power) incident on a satellite, and to radiate an amplified signal from the satellite. Inside the satellite, radio-frequency power is often guided by waveguides, or hollow metal pipes with circular or rectangular cross section. If a waveguide is open at one end, it will radiate some power, but not very efficiently. If the open end is flared out like a cone, it will be more efficient as a radiator. *See* WAVEGUIDE.

If an antenna is a good radiator, it is also a good receiver. Its properties are reciprocal. At the same frequency, it has the same pattern for transmitting and receiving. If it transmits a lot of power in a certain direction, it will also be good at receiving power from that same direction.

Types of antennas. A horn is an open waveguide designed to radiate maximum power. Close to the radiating end, the opening increases in size. The angles and surfaces (smooth, steps, or corrugations) are chosen for maximum performance. A horn is simple and reliable. Communications satellites use horns for global beams, which radiate over the entire Earth, facing the satellite. Horns may also be used for telemetry and command signals, which radiate in many directions.

A parabolic antenna is the type most commonly used in communications satellites. It has a feed, similar to a horn, and a parabolic reflector. The feed radiates a wide beam, aimed toward the reflector. The reflector surface has the shape of a paraboloid, made by rotating a parabola around its axis. The feed is at the focus of the parabola so that the reflector focuses the radiation into a narrow beam. Like the reflecting telescopes used by astronomers, parabolic reflectors work over a wide range of frequencies. That is, they have the same focus for different colors (they are free of chromatic aberration).

For Earth stations, large parabolic antennas usually have an on-axis feed located in the reflected beam (**Fig. 3***a*). Since the feed is small compared to the reflector diameter, it obstructs only a small part of the reflected beam. On small antennas, an offset feed may be placed outside of the reflected beam (Fig. 3*b*). The reflector is still part of a paraboloid, and the feed is still at the focus, but the reflector is part of one side of a paraboloid.

Antenna gain. The gain of an antenna is the ratio of how much better it radiates in one direction (beam center) than an ideal isotropic radiator, which radiates equally in all directions. The antenna gain G of a parabolic reflector is given by Eq. (1), where A is

$$G = 4\pi\eta A/\lambda^2 \qquad (1)$$

the physical area of the reflector, η is the antenna efficiency, and λ is the wavelength. The antenna efficiency accounts for imperfections in the antenna, and is always less than unity. The gain can also be calculated as $109\eta D^2 f^2$, where the diameter, D, is in meters (m), and the frequency, f, is in gigahertz (GHz). Thus, the gain of a parabolic antenna is higher for larger reflectors and for higher frequencies. As an example, an INTELSAT Earth station may have a 30-m (98-ft) antenna, transmitting at 6 GHz, with an efficiency of 0.60. Its antenna gain, G, is a ratio greater than 2,000,000.

Antenna beamwidth. A good antenna focuses the power in one direction, but only at the expense of power in other directions. The antenna is a passive device. It does not amplify the total power. With a good parabolic antenna, the power decreases rapidly as the angle from the beam axis increases. A useful definition of beamwidth is the angle over which the power has dropped to one-half of the peak value. This is called a half-power beamwidth. An approximate, but useful, equation for the half-power beamwidth of a parabolic antenna (in degrees) is given by Eq. (2), where, again, frequency, f, is in

$$\theta = 21/(fD) \qquad (2)$$

gigahertz and antenna diameter, D, is in meters. The number 21 is not a true constant, but may vary somewhat for different antennas. The INTELSAT antenna just described would have a beamwidth of 0.1°.

Antenna coverage. An Earth station antenna usually transmits to and receives from a single satellite. The maximum gain is desired in one direction, and a simple parabolic antenna often provides excellent performance. However, a satellite antenna often communicates with several Earth stations. In this case, the aim is to achieve a minimum gain anywhere in a given region. This region may be a country, a group of

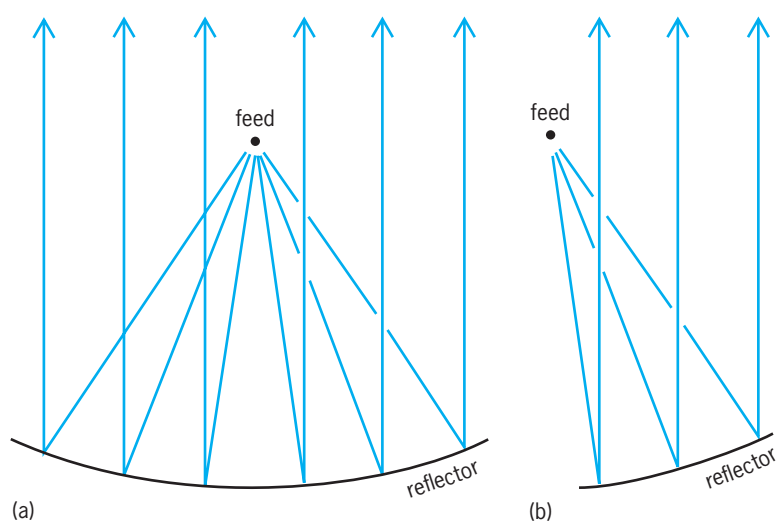

Fig. 3. Parabolic antennas. (*a*) On-axis feed. (*b*) Offset feed.

countries, or a time zone within a country. Because the power outside the desired region is wasted, it should be minimized. Circular and elliptical patterns can be generated with a single feed and a reflector. More complicated patterns can be generated with a single parabolic reflector and several feeds, or a single feed and a specially distorted reflector. In the case of a multifeed antenna, the designer often varies the number of feeds, their location, the power in each feed, and the phase. A computer is used to calculate the pattern for each variation, and the variables are adjusted until a near-optimum pattern is obtained. *See* ANTENNA (ELECTROMAGNETISM).

Receivers. The signal from the Earth station enters the satellite receive antenna and is amplified by the satellite receiver. Because the received signal is at a very low power level, the first few stages should amplify the signal without introducing additional noise. These initial amplifier stages are called a low-noise amplifier. After the signal has been amplified by the initial stages, its higher power level is less likely to be affected by noise.

The receiver also shifts the carrier frequency of the received signal. In a satellite it must be shifted to a different frequency before it is transmitted. Otherwise, the high-power transmitter would overwhelm the sensitive receiver. In an Earth station the frequency is often shifted to a lower frequency for ease of handling. It is most efficient to shift the frequency of the signal at low power levels. After the initial amplification stages, the frequency is changed, followed by additional amplification.

Power amplifiers. Traveling-wave tubes (TWTs) are often used in satellites to generate the high levels of radio-frequency power required. The power output of a traveling-wave tube increases as the power input is increased. As in any amplifier, the power output has a practical maximum; at this maximum, the traveling-wave tube is operating in a nonlinear mode called saturation. In satellite communications, maximum power is often desired, but it sometimes does not result in the best performance. For some operations, it is necessary to "back off" from the maximum, to find a linear portion of the output versus input curve. This decreases the distortion near saturation, and reduces crosstalk, when one user's communications cross over to a different user's channel at a different frequency. *See* CROSSTALK; SATURATION; TRAVELING-WAVE TUBE.

Solid-state power amplifiers (SSPAs) are also used in satellites when they can supply the desired power output. SSPAs have several stages of amplification. The maximum power is often limited by the heat dissipated by the amplifier. This heat can be conducted to larger surfaces and then radiated away. If enough heat cannot be dissipated, the temperature increases and may lead to failure of the amplifier. *See* POWER AMPLIFIER; MICROWAVE SOLID-STATE DEVICES.

Satellite Performance (Link Budget)

A communications satellite is built to be operated with specified Earth stations. Its communications performance, described in terms of a link budget, is evaluated as the power of the received signal relative to the sources of noise in the link between two Earth stations. The total link budget represents both parts of the link: the uplink from Earth station to satellite and the downlink from satellite to a second Earth station.

For each link a "good" signal must be received. The system designer must decide whether to spend more money for a more powerful transmitter, larger antennas, or a better receiver. To find the best system, the designer must work with a budget for each of the two links. One example of a downlink budget starts with a transmitter power of 180 W (**Fig. 4**, right). An antenna increases the power to an apparent 180,000 W, a loss in space drops this to 4×10^{-16} W, and the Earth station antenna increases this to a received power of about 10^{-12} W. A similar calculation is made for the uplink budget (Fig. 4, left). The other part of the calculation for each link, is to see if this is a "good" signal, that is, if it is larger than the noise by a sufficient margin.

Decibels. In both the uplink and the downlink, the power transmitted, P, is a large number of watts, and the carrier power received, C, is usually a very small number of watts. It is easier to express these numbers in decibels rather than in absolute quantities. Ratios were first expressed as decibels in measurements of sound intensity. The bottom of the scale was determined (the minimum intensity that can be heard), and any measured sound intensity could then be expressed in decibels, as an absolute measurement.

To express a quantity in decibels, the common logarithm of the absolute quantity is taken, and the result multiplied by ten, that is, $10 \log_{10} (\ )$. There are two main reasons for using decibels: (1) decibel notation makes it easy to write and understand very small and very large numbers, and (2) it is easier to calculate the cumulative effects of gains and losses along the link path by addition rather than multiplication.

In calculating link budgets, the unit abbreviation is attached to "dB." For example, 10 million watts, or 10^7 W, is expressed as 70 dBW. In calculating satellite performance, not only are many numbers expressed in decibels, but equations are often written in decibel format, with addition and subtraction symbols, instead of multiplication and division notation. *See* DECIBEL; LOGARITHM.

Transmitter performance. The start of any radio-frequency transmission in space is the output of the power amplifier from either the satellite or the Earth station. The transmitter should transmit the maximum power in the direction of the intended receiver. This is done by using an amplifier with maximum output power, P, and a transmit antenna with maximum antenna gain, G_t. The product of the transmitter power (in watts) and the antenna gain (as a ratio) is known as the equivalent isotropically radiated power or EIRP (in watts or dBW). From the receiver's point of view, the transmitter appears to be radiating this amount of power (and, it would be radiating this amount of power if it were radiating equally in all directions).

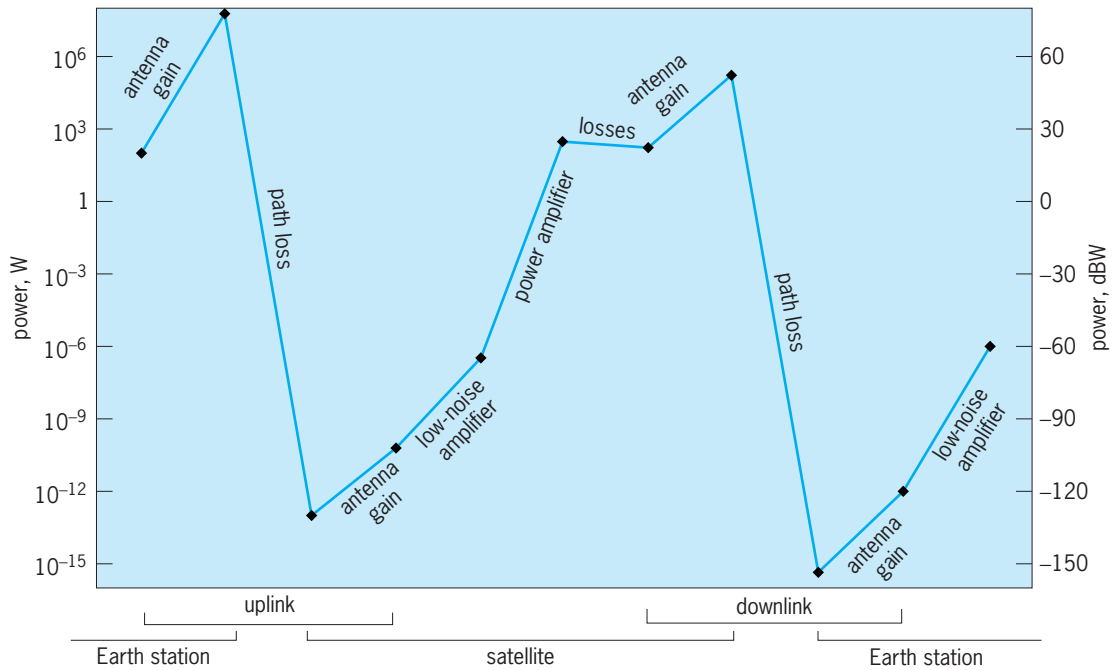

Fig. 4. Power levels for a broadcast satellite (link budget). (*Data from Space Systems/Loral*)

Free-space path loss. In almost any transmission of radio frequencies, most of the power is lost in space. In transmission to and from communications satellites, the distances are larger, so the losses are larger. The distance S from an Earth station to a satellite is enormous—more than 35,000 km (22,000 mi) for a geostationary satellite. The power will spread over a large area of $4\pi S^2$, and the receive antenna only intercepts a small area A, reduced by the receive antenna efficiency η. The received carrier power C is then given by Eq. (3). If the value of ηA from Eq. (1)

$$C = \frac{PG_t\eta A}{4\pi S^2} \tag{3}$$

is substituted in Eq. (3), the received carrier power C can be written as Eq. (4), where P is the transmitter

$$C = PG_t(1/L)G_r \tag{4}$$

power, and G_t and G_r are the gains of the transmit antenna and receive antenna. The transmitter power, P, is increased by the two antenna gains, and reduced by the loss in space. The free-space path loss, L, is found to be $(4\pi S)^2/\pi^2$. Since the wavelength λ of the carrier is very small and the distance S is very large, the free-space path loss is a very large ratio.

The reason for introducing the receive antenna gain G_r into Eq. (3) is to treat both antennas in the same way. Antennas are reciprocal: an antenna used to receive has the same gain as it would to transmit, providing that gain is measured at the same frequency. A numerical example of Eq. (4) is given at the top of **Table 2**. The transmitter power per channel is 180 W. The transmit antenna increases this by a factor of 1000. The reciprocal of the free-space path loss is a factor of 2.3×10^{-21}, which with a receive antenna gain of 2500 results in a carrier power received of 1.04×10^{-12} W.

TABLE 2. Performance of the downlink, from a broadcast satellite*

Symbol	Description	Number	In decibels†
P	Transmitter power per channel	180 W	22.5 dBW
G_t	Transmitter antenna gain	1000	30.0 dB
$1/L$	1/Free space path loss‡	2.3×10^{-21}	−206.3 dB
G_r	Receiver antenna gain	2500	34.0 dB
C	**Carrier power received**	**1.04×10^{-12} W**	**−119.8 dBW**
T_s	Noise temperature	125.0 K	21.0 dBK
k	Boltzmann's constant	1.381×10^{-23} W/Hz-K	−228.6 dBW/Hz-K
B	Bandwidth	24,000,000 Hz	73.8 dBHz
kT_sB	**Noise received**	**4.1×10^{-14} W**	**−133.8 dBW**
C/kT_sB	Carrier-to-noise ratio	25	14.0 dB

*Data from Space Systems/Loral.
†Decibels are explained in the text. Example: $10 \log_{10}(180\text{W}) = 22.5$ dBW.
‡Path is 39,500 km (22,544 mi) from satellite to Earth; radio frequency is at 12.5 GHz.

Noise. The carrier power received is important only when compared to the noise present in the system. Without additional information, it is difficult to determine whether the received carrier power of 1 trillionth of a watt is sufficient. If no noise is present, the signal can be amplified to any desired power; however, if there is too much noise present, the received signal is useless. The noise may come from many sources. The noise may be directly incident on the antenna (such as the Sun in the antenna beam axis), it may leak in from the sides (from antenna sidelobes), it may be added between the antenna and receiver (waveguides may not be perfect), or it may be inherent in the receiver (from the "low-noise amplifier").

In satellite communications, the noise power is often expressed as a system noise temperature T_s, equivalent to that radiated by a blackbody at the specified temperature. This noise is usually present over a broad band of frequencies, and most of it can be filtered out. However, the desired signal occupies a narrow band, and the noise in that band cannot be filtered out. For an audio signal the bandwidth might be only 10 kHz, while for a television channel it might be 24 MHz. When the noise temperature, T_s, the bandwidth, B, and Boltzmann's constant, k (a fundamental physical constant), are multiplied together, the result is the noise received, in watts. In the example in Table 2, the noise temperature is 125 K (−148°C), and the bandwidth is 24 MHz, so the noise power received is 4.1×10^{-14} W. Usually, the noise temperature is lower for a downlink, since the receive antenna on the ground "sees" cold space as background, while in an uplink, the receive antenna "sees" the warmer earth. *See* BOLTZMANN CONSTANT; HEAT RADIATION.

The values in the right column of Table 2 are in decibel notation. To calculate the carrier power received, the first four numbers are added instead of multiplied. To find the noise power, the decibel values of T_s, k, and B are added. The final carrier-to-noise ratio, 14.0 dB, is determined by subtracting the noise (−133.8 dBW) from the carrier power (−119.8 dBW).

In this example, the ratio of carrier to noise is 25, which means that the signal is 25 times stronger than the noise. The minimum ratio required depends on the specific application. During the design phase, if the ratio is too low, some changes must be made in the link to achieve an acceptable value. For example, the transmitter power or the size of one of the antennas might be increased.

System performance. Usually, a communications satellite receives a signal from an Earth station and sends it back to another Earth station. The system performance depends on the uplink (Earth to satellite), and the downlink (satellite to Earth). To combine the two ratios, the noise per carrier for each link must be added. Therefore, the reciprocals of the carrier-to-noise ratios must be added, and then the reciprocal taken of the result. If the uplink carrier-to-noise ratio (C/N_u) is 240 and the downlink carrier-to-noise ratio (C/N_d) is 25, then the total carrier-to-noise ratio (C/N_t) is 22.6. In general, the combined ratio is less than that of either of the links.

Losses. In addition to free-space path loss, there may be many other losses in power. Some are negligible, while others are significant. The transmitter and the transmit antenna are usually connected with a waveguide that may have losses. Similarly, there may be losses between the receive antenna and the receiver; to minimize these, the receiver is mounted as close to the antenna as possible.

Geometric loss results from spreading the signal over a large area. Additional losses may be incurred in going through the Earth's atmosphere. At certain frequencies, especially above 10 GHz, there are absorption bands from oxygen and water vapor in the air. For low-elevation angles from the Earth station to the satellite, these bands produce more absorption because the beam has a longer path in the atmosphere. The presence of raindrops in the atmosphere can increase both the attenuation and the noise level in the signal. Both the decrease in the carrier signal and the increase in the noise result in a decrease of the all-important ratio of carrier to noise power.

Margin. Some losses are fixed, some slowly increase with time, and others fluctuate with time. Some losses are known, and some are unknown. To compensate for all these losses, a system is designed with a margin. This means that the system has more capability (a higher total carrier-to-noise ratio) than needed. When a loss is larger than expected, the margin decreases, but the system continues to operate according to specifications. The power of Earth station transmitters can usually be adjusted, especially when several stations are transmitting to one satellite at the same or adjacent frequencies. The satellite transmitted power can also be adjusted by ground-controlled attenuators.

Space Technology

Much of the space technology used in communications satellites is also used in many other satellites in near-Earth orbit. This section deals primarily with the technology specially for communications satellites. *See* SPACE PROBE; SPACE TECHNOLOGY.

Orientation. Many communications satellites are three-axis body-stabilized, that is, they have one axis pointed north, one axis pointed toward Earth, and the third more or less parallel to the velocity vector along the orbit. The orientation is measured with Sun sensors (the simplest method), Earth sensors (or horizon sensors), or radio-frequency sensors. Any error in attitude is corrected by actuating wheels with rotational inertia. The simplest arrangement to visualize is three reaction wheels, each used to correct for errors around the three axes. For redundancy, four reaction wheels may be used, with each axis at the same angle from the other three (109.5°). Some satellites have either a fixed momentum wheel or a gimballed momentum wheel that always turns in the same direction. Note that gyroscopes are wheels used only as sensors, reaction wheels rotate in either direction, and momentum wheels rotate always

in the same direction. When a wheel reaches its maximum allowed rotation speed, a thruster is fired to exert a torque that causes the wheel to slow down. *See* GYROSCOPE.

Structure. A typical communications satellite has a compact rectangular body, with various antennas, thrusters, and solar arrays mounted on external surfaces. After launch, the two solar arrays are unfolded (deployed) in the north and south directions, and mounted so that they can rotate around this north-south axis to face the Sun continually. Usually, there is no correction for the annual north- south motion of the Sun. The structure needs to be light, strong enough to withstand the acceleration and vibration of launch, and stiff enough to keep antennas pointed accurately. Therefore, it is made of light metals (aluminum, beryllium, magnesium, or titanium) or a composite (boron/epoxy or graphite/epoxy). *See* SPACECRAFT STRUCTURE.

Propulsion. Thrusters are used for both attitude control (satellite orientation) and station keeping (satellite positioning). The most common propellant is hydrazine (N_2H_4). This monopropellant needs no oxidizer; when heated, it decomposes and generates more heat. The simplest system has pressurized nitrogen inside the tank with the hydrazine to expel the fuel from the tank (blow-down system). There is no apparent gravity in orbit, so either a bladder or a surface tension system is needed to ensure that fuel rather than nitrogen comes out of the tank. *See* HYDRAZINE; ROCKET PROPULSION; SPACECRAFT PROPULSION.

Electric power. A communications satellite uses most of the available power for the power amplifiers: either traveling-wave tubes or solid-state power amplifiers. Primary electrical power is provided by solar cells mounted on solar arrays, and secondary power is provided by rechargeable batteries. These components must be connected with appropriate regulation, conversion, and distribution. Deployed solar arrays get cold during eclipse, and generate excess voltage when they come out of eclipse. Batteries require a higher voltage during charge, and provide a lower voltage during discharge. Solid-state power amplifiers need only a few volts. Traveling-wave tubes require a few volts for the filament heater, but over 1000 V for acceleration of the electron beam. Because solar arrays lose power due to particle radiation in space over many years, they are built with excess power. Some batteries have a limited life, but with careful construction, testing, and operation, they provide excellent service. *See* RADIATION DAMAGE TO MATERIALS; SPACE POWER SYSTEMS.

Thermal control. In many communications satellites, the majority of the electric power is used by the power amplifiers. While these amplifiers generate lots of radio-frequency power, they are not 100% efficient, so they also generate heat. Dissipation of this heat is a key element in the thermal control of communications satellites. The north and south faces are often used for heat dissipation, since they receive less sunlight than the other surfaces. These surfaces have low absorptivity to visible light, including sunlight, and high thermal emissivity to radiate the heat into space. The heat is conducted to these faces with solid metal or with heat pipes. The power amplifiers are mounted on or near the inside of these surfaces. With solid-state power amplifiers, the small integrated circuit chip is mounted on a metal heat sink that absorbs and spreads the heat to limit the chip temperature.

Telemetry and command. The objective of the telemetry subsystem is to provide information about the spacecraft for operational use, failure analysis, and prediction of spacecraft performance. The command subsystem receives commands from the spacecraft control center and executes them. Some commands are used to reconfigure the communications subsystem, to fire thrusters, and to switch to redundant equipment when needed.

Since telemetry and command are essential for failure recovery, the antenna coverage must be nearly omnidirectional. This enables recovery even when control of orientation has been lost, and the main antennas no longer point at the Earth. If the available power decreases, other equipment is turned off automatically so that the telemetry and command subsystems will still have power. *See* TELEMETERING.

Tracking. The location of a satellite can be measured by determining its distance from two or three Earth stations. The distance is determined from the transit time of radio signals. This measurement is independent of the processing of telemetry and command signals, but uses the same command receiver and telemetry transmitter. The transit times may be measured by transmitting tones of different frequencies and finding the phase relations of the returned signals. The transit times can also be measured by transmitting a pseudorandom series of bits, and measuring the time delay for which there is most correlation between the transmitted and received signal—when they match. From a series of measurements, the location of the satellite at various times can be determined. Thruster firings are then scheduled to keep the satellite in the desired location.

Reliability. Large velocity changes are needed to launch satellites into space. Once launched, satellites are difficult to repair and expensive to replace. Therefore, satellites must function when they are first launched, and continue to perform for many years. This task is similar to building a car that will go over 200,000 km (over 125,000 mi) without maintenance or repair. Satellites achieve a long life by having inherently reliable components, redundant subsystems, and a rigorous test schedule before launch.

Communications satellites have an excellent record. The launch reliability, including both the launch vehicle and initial operation of the satellite, may be 75% for early launch vehicles and new programs and as high as 95% for mature operations. The probability of their reaching a design lifetime of 12 or 15 years may be calculated as 75%, and is often exceeded. This excellent record, in addition to larger satellites handling lots of traffic, has drastically reduced the cost of communications, including overseas phone calls. For example, a computer

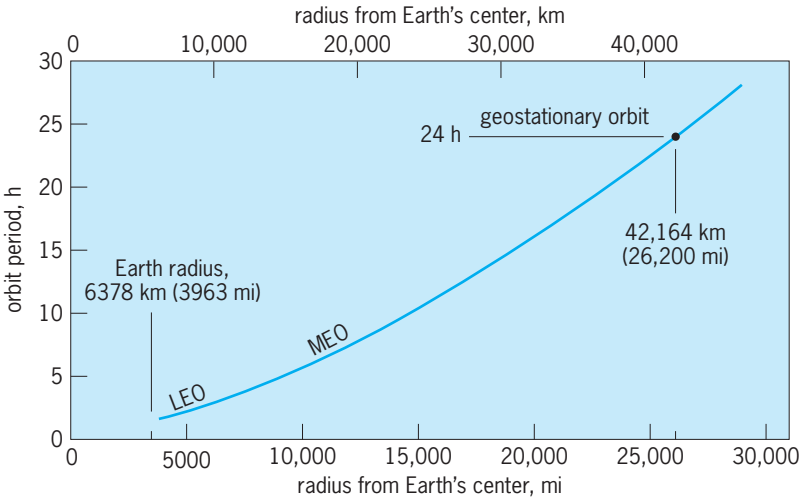

Fig. 5. Orbit periods of circular orbits.

may be purchased at a nearby store, but the technical assistance may be halfway around the world, via satellite. *See* RELIABILITY, AVAILABILITY, AND MAINTAINABILITY.

Orbits

Communications satellites circumnavigate the Earth in a path called an orbit. The majority are in a very special orbit, called a geostationary orbit. Other orbits are also used, due to advantages in launching the satellite, keeping the satellite in place (stationkeeping), or communicating with desired earth stations.

Geostationary orbit. The first characteristic of a geostationary orbit is that the path is circular (eccentricity is near zero), so the satellite is always the same distance (radius) from the center of the Earth. When the radius is larger, the time to go around increases, both because of the larger distance and because the satellite has a smaller velocity (**Fig. 5**). The second characteristic is that the geostationary orbit have a period of rotation exactly equal to the time it takes the Earth to rotate around its axis (23 h, 56 min, 4.1 s). This occurs for a satellite 35,786 km (22,237 mi) above the Earth's surface, or 42,164 km (26,200 mi) from the center of the Earth. For low Earth orbits (LEO), the period is only a few hours; for medium Earth orbits (MEO) it is longer. For a very large radius of 384,300 km (238,860 mi), the orbit period would be 27.3 days, which corresponds to the Moon's orbit around the Earth.

A third characteristic of a geostationary orbit is that the orbit must be in the equatorial plane. Otherwise, the satellite would have a north-south oscillation. Since the satellite orbit period is the same as the time it takes for the Earth to complete a rotation, the satellite is always over the same point (longitude) on the Equator. And, most importantly, since the satellite is always at the same height over the same point on the Equator, it appears to be stationary in the sky to an observer on the Earth. Once an Earth station antenna is pointed at the satellite, it will remain pointed at the satellite.

Different points along the Equator can be identified by their longitude. Longitudes may range from 0 to 360° E, or they can be written from −180° to +180° E. Negative east longitudes are simply west longitudes (west from the Greenwich meridian). The longitude of a geostationary satellite is the longitude of the point on the equator directly underneath the satellite.

Azimuth and elevation. To point an antenna at a satellite, it is useful to know the satellite's azimuth and elevation as seen by an observer on Earth. The azimuth is the direction of the satellite, measured in the horizontal plane from geographic north in a clockwise direction (starting eastward). The elevation is the angle between a satellite and the observer's horizon plane.

For an observer in the Northern Hemisphere, geostationary (GEO) satellites appear placed in an arc in the southern sky, as shown in **Fig. 6**. The top curve in the figure is for an observer at 10° N latitude. Satellites are shown for differences of 2° longitude (satellite longitude minus station longitude). The lower curves are for observers at other latitudes (30° N, 50° N, and 70° N), with satellites having 10° spacing. The azimuths are between 90° and 270°. For an observer in the Southern Hemisphere, the geostationary arc appears in the northern sky, and the azimuths are in the ranges 270° to 360°, and 0° to 90°.

If the longitude of the satellite is equal to the longitude of the observer (Earth station), then an observer in the Northern Hemisphere would find the satellite directly south (azimuth = 180°). When the longitudes of the satellite and observer are different, then the elevations are lower, and the azimuth (direction) is no longer due south.

Stationkeeping. If the Earth were perfectly spherical and there were no other perturbing forces, a satellite inserted into geostationary orbit would remain there indefinitely. But there are perturbing forces, and over time the orbit of a satellite changes. For a communications satellite to perform its mission, small thrusters must be used to counteract these forces and keep the satellite in its desired location. This is called stationkeeping. Without these thrusters, a geostationary satellite would have significant longitudinal drift, north-south oscillations, east-west oscillations, and radial oscillations.

The longitudinal drift is caused by irregularities in the Earth's shape around the Equator, unknown until artificial satellites were launched. Specifically, the equatorial cross section is not circular. This causes a change in the longitudinal drift, which can be a longitude acceleration of as much as 0.002 (degree)/(day)^2. This is small, but over many days the satellite would move to a different longitude. The change in longitude can be pictured as a weight sliding downhill, as in **Fig. 7**. The figure shows two low points, around 75°E and 252°E. A satellite at one of these stable longitudes will stay there, and a satellite near that longitude will oscillate about that longitude. The first commercial communications satellite, *Early Bird*, was stationed at 310°E and kept there by use of its thrusters. When it ran out of fuel, it "slid down the hill" to 252°E, and up the other side to 200°E. Around 200°E, it reversed,

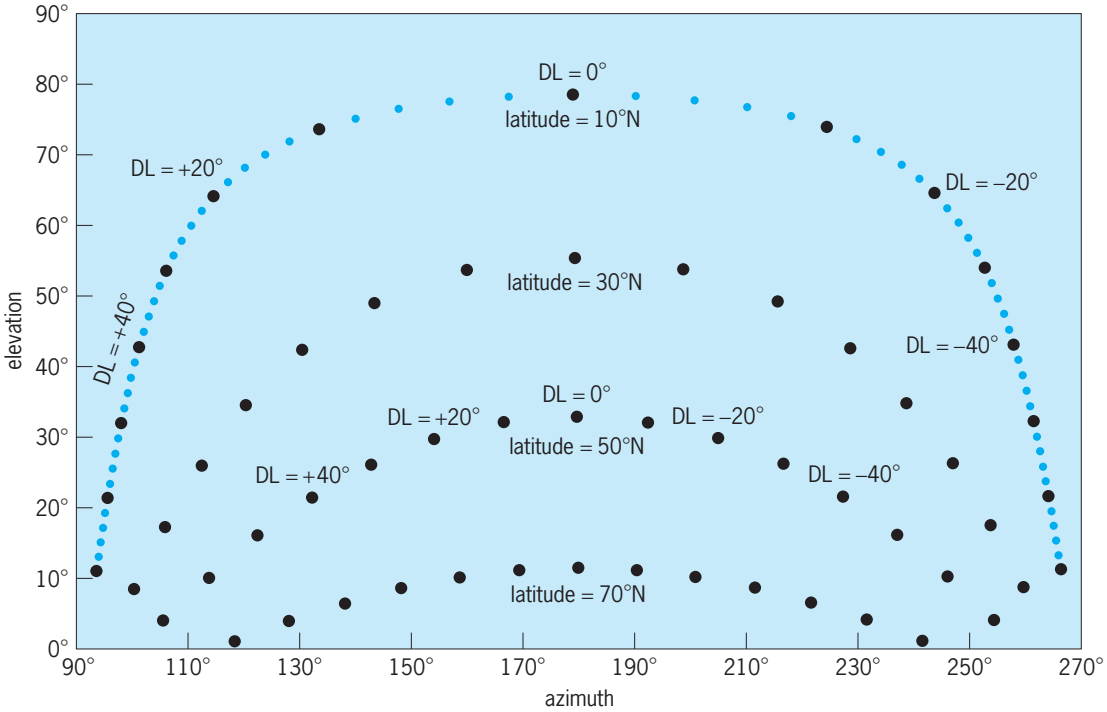

Fig. 6. Azimuth and elevation of geostationary satellites. DL = satellite east longitude − station east longitude.

and came back to 310°E. It continues to go back and forth every 3 years. *See* EARTH, GRAVITY FIELD OF.

The daily north-south oscillations of the satellite are caused by gravitational attractions of the Sun and Moon. These cause the orbit plane to tip with respect to the equatorial plane. Instead of remaining over one spot on the Equator, the satellite shows a daily cycle of going into the Northern Hemisphere and then the southern hemisphere. If unchecked, this north-south oscillation will increase to 15° in about 26 years, and then return to zero. In practice, controlling these north-south oscillations takes the major part of a satellite's fuel. Near the end of the satellite's life, the fuel may be reserved for longitudinal control, and the north-south oscillations allowed to increase.

The daily east-west oscillations of a satellite are caused by solar radiation pressure, which is 4.57×10^{-6} N/m^2 (6.6×10^{-10} psi) at the Sun-to-Earth distance. On one side of the orbit, the force increases the velocity, and on the other side it decreases it. These forces change an initially circular orbit into an elliptical orbit. On one side the satellite is going faster than the Earth, and the satellite moves eastward in longitude, and on the other side, 12 h later, the satellite is moving westward. The eccentricity would continue to increase, except that after 6 months the sunlight comes from the opposite direction, and the effect is reversed. Satellite maneuvers are usually not done just to control eccentricity. When thrusters are fired to control the mean longitude, they also affect the orbit eccentricity. Choosing the best time of day for these maneuvers usually allows the orbit eccentricity to be kept within acceptable limits.

Daily radial oscillations also take place with eccentric orbits, as the distance between the satellite and the center of Earth varies. In terms of actual distance (kilometers or miles), the daily radial oscillation is exactly one-half of the east-west oscillation. For antenna pointing, the radial oscillations are usually not important. In a few applications, involving digital communications, radial oscillations may become significant, as both the frequency received and the time delay have a daily variation.

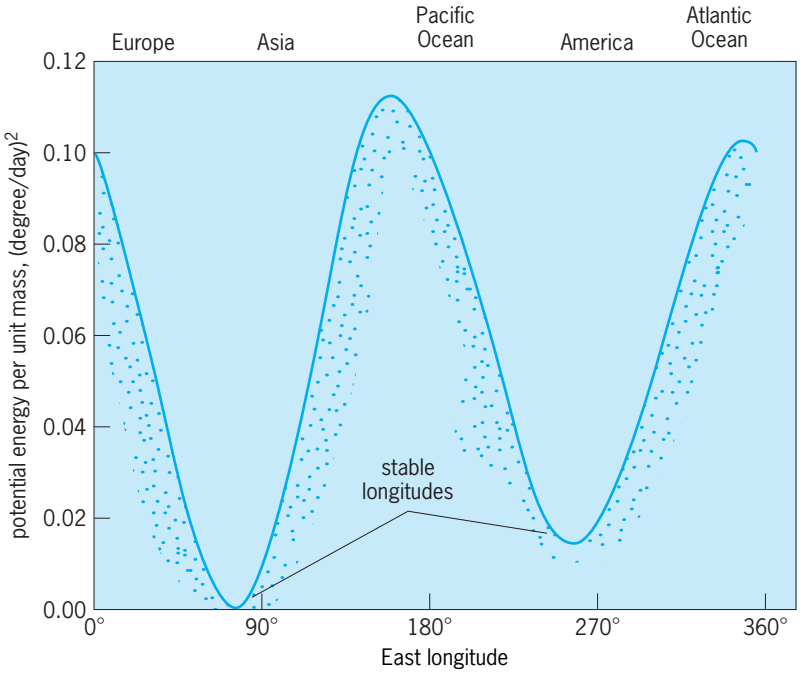

Fig. 7. Longitudinal acceleration pictured as hills and valleys. A satellite tends to "slide downhill." (The vertical scale, which is the potential energy per unit mass, is also the integral of the longitude acceleration with respect to longitude.)

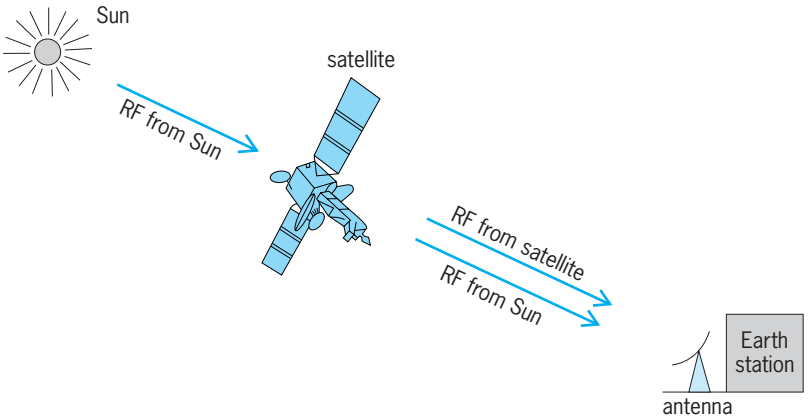

Fig. 8. Sun-station interference. Radio-frequency (RF) signals from the Sun interfere with those from the satellite.

Satellite link delays. A signal goes from Earth to a satellite at the speed of light, which is 299,792 km/s (186,282 mi/s). This means that there is a time delay between sending and receiving a signal. The distance to a communications satellite in geostationary orbit can vary from 35,786 km (22,237 mi) to 41,679 km (25,898 mi), so the time delay is from 0.12 to 0.14 s. The usual phone conversation through a satellite includes an uplink and a downlink from one speaker to another, and then a return through another uplink and downlink. Therefore, there are four time delays in sequence. The total delay is about half a second. This time delay can be a problem when the two speakers start speaking at the same time. It also produces an echo if equipment is not operating properly. Suppressing the echo requires special equipment because local phones have only a pair of wires to transmit and receive. Lower-altitude satellites, not in geostationary orbit, will have less time delay.

Sun, Earth, and Moon effects. The Sun emits radio-frequency radiation at microwave frequencies. This radiation can produce noise in a communications downlink, when the Sun is in the beam of the Earth station antenna receiving the satellite signal. The worst interference occurs when the satellite is directly between the Earth station and the Sun (**Fig. 8**). At different times, this condition may affect the downlink of every geostationary communications satellite to any Earth station. This Sun-generated noise may be a minor annoyance, or may completely overwhelm the transmitted signal. However, interference occurs only a few consecutive days in the spring and a few days in the fall, and it lasts only several minutes each day.

The Earth casts a shadow into space. Except for a few minutes at the start and end of the shadow period, these are total eclipses. When a geostationary satellite is in the shadow, the solar array output drops to zero. To continue operation, power is furnished by rechargeable batteries. These eclipses occur in two 6-week periods in the spring and in the fall. The duration of the eclipse increases to a maximum of 70 min (which occurs on the vernal and autumnal equinox), and then decreases again to zero over that period.

The Moon also casts a shadow in space. When a satellite is in the shadow, the solar array output also drops. These eclipses are mostly partial eclipses (the satellite is in the penumbra). The occurrences are irregular, and can occur during any month of the year. Most last less than an hour. A very small number of eclipses are much longer. On September 23, 2014, for example, an unusual partial eclipse on a geostationary satellite at 180° E will last over 3 h. A few similar eclipses of long duration occur every 18.6 years (when the Moon's orbit plane is closest to the equatorial plane). *See* ECLIPSE; MOON.

Launching into geostationary orbit. A communications satellite is often launched into geostationary orbit in three steps, which requires three firings of rockets or thrusters. The first step is a launch into a parking orbit, a low circular orbit. The second step is from the parking orbit into a transfer orbit, which has a high point (apogee) near geostationary orbit, and a third step circularizes the orbit and puts the satellite into geostationary orbit. Each rocket firing starts with an initial velocity, produces a change in velocity (called Delta-V), and ends with a final velocity (**Table 3**). The change in velocity is related to the mass ratio, that is, the ratio of the combined rocket, fuel, and satellite masses to the rocket and satellite mass after the fuel mass is used up. The actual numbers depend on various factors, including the launch site, launch vehicles, and height of the parking orbit. If a 1000-kg (2200-lb) satellite is needed in geostationary orbit, the mass in parking orbit is about five times as much, or 5000 kg (11,000 lb).

Constellations of satellites at lower altitudes. Compared to satellites in geostationary orbit, satellites at lower altitudes have smaller orbit periods. This means they cannot continuously communicate with an Earth station. Eventually they will disappear below the horizon. With a number of satellites, it is possible to always have one or more above the horizon. The Earth station can maintain continuous communications by switching from one satellite to another, just as a cell phone in a car switches from one tower to another. *See* MOBILE RADIO.

TABLE 3. Launching into geostationary orbit, from Kennedy Space Center (Florida)

	Initial velocity, km/s (mi/s)	Final velocity, km/s (mi/s)	Delta-V, km/s (mi/s)	Mass ratio
Achieve parking orbit	0.4 (0.25)	7.7 (4.8)	7.3 (4.5)	10 to 1
Perigee firing	7.7 (4.8)	10.1 (6.3)	2.4 (1.5)	5 to 2
Apogee firing*	1.6 (1.0)	3.1 (1.9)	1.8 (1.1)	2 to 1

*Cape Kennedy launch requires extra Delta-V of 0.3 km/s (0.2 mi/s) for 30° orbit plane change.

The advantages of these constellations include ease of communication to a satellite that is closer, ease of launch to lower altitudes, and antenna beams that can focus on smaller areas (thus serving higher-density populations). The major disadvantages are the large number of satellites required, and the large initial cost of manufacturing and launching all the satellites. Many different constellations of satellites have been studied. Usually a symmetrical constellation is planned, with each satellite carefully positioned with respect to the others.

One system, called Globalstar, uses 48 satellites. There are eight circular orbits, each with six satellites. The orbits are inclined at 52° to the Equator, and the ascending nodes (where the satellites pass the Equator going from south to north) are separated by 45°. In each orbit, there are six satellites equally spaced around the orbit. Each satellite provides communications between a handheld phone and a base station that connects to landlines, undersea cables, or satellites that handle international traffic. Technically, this system has been successful, but in the initial phases it has not generated the expected traffic and revenue. Another similar system, Iridium, included intersatellite links, so that any communications path could be established through a series of these satellites to any point in the world.

Other orbits. To meet special needs, other orbits are used by some communications satellites. The former Soviet Union established *Molniya* satellites in a Molniya orbit. This elliptical orbit is inclined 62.9° to the Equator, with a high point (apogee) at the most northern latitude. Special parameters were chosen so that the apogee did not precess to the Southern Hemisphere. Also, the orbit period was one-half of a sidereal day (11 h, 58 min), so every other apogee would always be at the same longitude. Earth stations usually had two tracking antennas, so that they could switch between satellites without a break in communications. The *Molniya* orbit has high-elevation-angle coverage of the Northern Hemisphere, and advantages of long dwell times and economical launches. During the Cold War, there were two hotlines between Washington and Moscow, one using a *Molniya* satellite, the other using a geostationary satellite.

More recently, variations on the Molniya orbit have been proposed for certain communications satellites. These all have in common the use of an inclined, elliptical orbit, with the satellites dwelling at the apogee over the geographic area they are intended to serve. In ITU terminology, they are referred to as HEO (highly elliptical orbit) satellites. Some systems have their orbit period as a submultiple of a sidereal day in order to keep the apogee at an constant Earth longitude. For continuous service, more satellites are required in the constellation because the orbit period is reduced (for example, three satellites for an 8-h orbit).

Satellites for mobile communications are sometimes designed without north-south stationkeeping. These are sometimes called geosynchronous satellites because, while the orbit period is equal to a sidereal day, the orbit inclination results in a daily north-south excursion of several degrees. This reduces consumption of satellite fuel. These orbits are applicable to mobile communications because the beamwidth of the mobile Earth stations tends to be wider, and able to encompass the north-south movement of the satellite without losing signal strength. Mobile applications with narrow-beamwidth Earth-station antennas must be equipped to constantly repoint their antennas because the mobile user may move in different directions and sometimes with pitch and roll. In this case, the additional movement of the satellite by a few degrees in a north-south direction can be easily accommodated. The *Marisat* satellite, for example, started with an inclination of several degrees (resulting in north-south oscillations of several degrees), which over the years first decreased to zero and then started increasing.

Stations in polar regions cannot use geostationary satellites. For those very close to a pole, the satellites are below the horizon, and for those a little farther away, it is difficult to communicate with a satellite at a low elevation. After 25 years, the *Marisat* satellite has an inclination of 15°, so each day it is above the horizon for a few hours. It has outlived its design lifetime, but still has some communications capability. Therefore, the research station at the South Pole can use it for a few hours each day, and other similar satellites at other times. *See* CELESTIAL MECHANICS; ORBITAL MOTION.

History and Present-Day Satellites

During World War II, the development of rockets and radar (using microwaves) made satellite communications a feasible concept. Launch vehicles would be required to place the satellites into orbit, and microwave components to make wideband communications possible. In 1945, Arthur C. Clarke wrote about the possibility of communications satellites in geostationary orbit. Transistors had not yet been invented, so he thought "in terms of large manned space stations. . . perhaps half a century" latter. In 1958, a battery-operated tape recorder (project SCORE) transmitted the first satellite message to the whole world—a Christmas message by President Eisenhower from a low-orbit satellite.

Active repeater satellites. In 1962, the first repeating communications satellites were launched: *Telstar* and *Relay*. These satellites received a signal, amplified it, and retransmitted it. For the first time, television signals could be transmitted across the Atlantic Ocean, but these satellites were in low orbit, and after 20 or 30 min disappeared below the horizon. Then *Syncom I* was launched into geosynchronous orbit. Due to launch constraints, it was in a 33° inclined orbit, so that each day the subsatellite point (the point directly under the satellite) oscillated between 33° N and 33° S.

Led by COMSAT, a multinational consortium of nations, the International Telecommunications Satellite (INTELSAT) Organization, was formed to establish global commercial satellite communications. In 1965, *Early Bird* was launched by COMSAT into a geostationary orbit, in which it remained stationed over the mouth of the Amazon River. It could thus

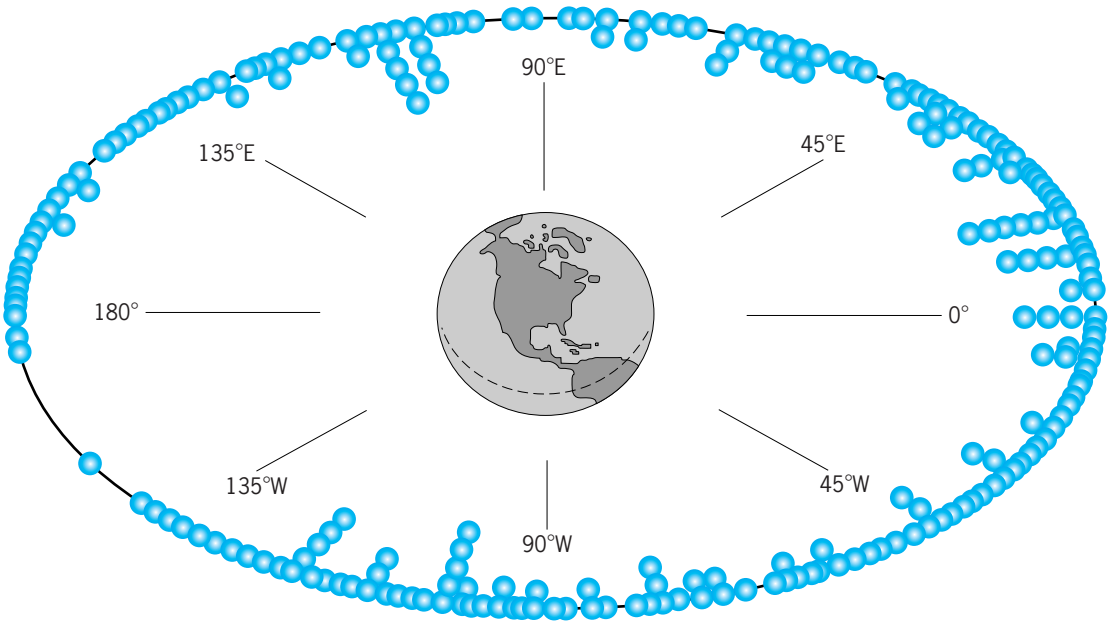

Fig. 9. Commercial communications satellites in geostationary orbit. (*Boeing Company*)

provide continuous communications between the eastern United States and western Europe 24 h a day. An INTELSAT satellite launched in 1967 brought commercial service to the Pacific area. Additional satellites later provided more communications to achieve a commercial global system.

Meanwhile, the U.S. Department of Defense developed the Initial Defense Communications Satellite Program (IDCSP), whose satellites were in near synchronous orbits. The Soviet Union launched communications satellites into Molniya orbits, which provided better coverage at high latitudes (far from the equator). Many Communist countries used the system initiated by the Soviet Union called Intersputnik.

Technical progress. Early earth stations were large and powerful so that they could communicate with small and weak satellites. As the capabilities of satellites increased, the Earth terminals could be smaller and less powerful. From specialized sites they moved to business premises and to homes. From communications to ships they progressed to other mobile communications, with airplanes, cars, and handheld phones. Satellite antennas became larger, with higher gains. Low-noise amplifiers made it easier to detect faint signals. Traveling-wave tubes and solid-state power amplifiers became more powerful. Silicon solar cells tripled in efficiency. Nickel-cadmium batteries were replaced by nickel-hydrogen and other batteries.

The number of active commercial communications satellites in geostationary orbit is now over 200 (**Fig. 9**), not including communications satellites in lower orbit or military communications satellites in geostationary orbit. With a variety of methods of frequency reuse, there is still room for more satellites.

Communications satellites are designed for specific missions, but have many common features (**Table 4**). A large, modern communications satellite may communicate with large fixed Earth stations (INTELSAT), it may broadcast to a large number

TABLE 4. Characteristics of a few communications satellites in use*

Spacecraft type:	Fixed satellite service	Broadcast	Mobile	Ka-band spot	Military
Spacecraft name:	INTELSAT IX	Echostar 9	Thruraya 2	iPSTAR	Wideband Gapfiller
Launch year	2001 to 2003	2003	2003	2005	2007
Procurement agency	INTELSAT	EchoStar	Thruraya/Etisalat	Shin	U.S. Military
Satellite contractor	Space Systems/Loral	Space Systems/Loral	Boeing	Space Systems/Loral	Boeing
Launch vehicle	Ariane-44L, Ariane 5, Proton-M	Zenit-3SL	Zenit-3SL	Ariane-5GS	Delta-4M + (5,4)
Orbital slot	64°E to 34.5°W	121°W	44°E	120°E	
Satellite mass, kg (lb)	4725 (10,395)	4737 (10,421)	5177 (11,389)	6775 (14,905)	5987 (13,171)
Transponders per frequency band	44 C, 12 Ku	2 Ka, 32 Ku, 24 C	128-element L, 2 C	53 Ku, 18 Ka	× Ka
Mission life, years	15	15	12	15	14

*From Space Systems/Loral.

of television users (Echostar and Direct TV), or it may communicate directly with handheld telephones (Thruraya). Satellites may be launched from the Kennedy Space Center in the United State (using the Delta or Atlas launch vehicle), Korou in French Guiana (Ariane), or Russia (Zenit or Proton). Many have a mass of around 5000 kg (10,000 lb), 50 to 100 power amplifiers (transponders), and design lifetimes of 12–15 years.

Since the beginnings of communications satellites in the 1960s, advances in communications have included fiber-optic cables and cellular telephones. However, these technologies are unlikely to completely replace satellites since these services complement each other. Satellites can never provide adequate cell phone service for large cities, but cellular phone towers may never be built in sparsely populated areas, and certainly not in ocean regions. Satellites are well suited to mobile communications, including aeronautical, maritime, and land applications. Fiber-optic cables are ideal for wideband communications from point to point, but when undersea cables break, satellites are often used to restore traffic. And satellites are particularly useful for broadcasting to a large number of points, whether for distributing network programs to television stations, or directly to television subscribers. Satellites also have the flexibility to switch their capability from one area to another. Spot beams can be repositioned to accommodate changing traffic needs for coverage of special events such as the Olympics or compelling news, such as disasters and other national and international developments. Gary D. Gordon; Denis J. Curtin

Bibliography. M. R. Chartrand, *Satellite Communications for the Nonspecialist*, International Society for Optical Engineering (SPIE), Bellingham, WA, 2004; G. D. Gordon and W. L. Morgan, *Principles of Communications Satellites*, Wiley, New York, 1993; G. Maral and M. Bousquet, *Satellite Communications Systems: Systems, Techniques and Technology*, 4th ed., Wiley, New York, 2002; J. N. Pelton, P. Marshall, and R. J. Oslund (eds.), *Communications Satellites: Global Change Agents*, LEA Publishers, Mahwah, NJ, 2004; W. L. Pritchard, R. A. Nelson, and H. G. Suyderhoud, *Satellite Communication Systems Engineering*, Prentice Hall, 1993.

Communications scrambling

The methods for ensuring the privacy of voice, data, and video transmissions. Various techniques are commonly utilized to perform such functions.

Voice. Analog methods typically involve splitting the voice frequency spectrum into a number of sections by means of a filter bank and then shifting or reversing the sections for transmission in a manner determined by switch settings similar to those of a combination lock; the reverse process takes places at the receive end. The strength of these methods depends on the number of switch settings available (often called the key size) and the rate at which the switch settings are changed. *See* ELECTRIC FILTER.

Digital methods first convert the analog voice to digital form by a number of digital coding techniques such as pulse code modulation (PCM), delta modulation, or linear predictive coding (LPC). The digital voice data are then scrambled or encrypted by one of the many methods discussed below and are transmitted digitally. After the data are received, they are descrambled or deciphered and then converted back to analog form by the appropriate voice decader. *See* ANALOG-TO-DIGITAL CONVERTER; DIGITAL-TO-ANALOG CONVERTER; MODULATION; PULSE MODULATION.

Data. A simple data-scrambling method involves the addition of a pseudorandom number sequence to the data at the transmit end. The same sequence must be subtracted at the receive end to recover the data. The strength of the method depends on the period of the sequence and the way the sequence is generated, as well as the number of different pseudorandom number sequences available. Devices using this method are known as stream ciphers. Typically, the pseudorandom number sequences are in binary form; and the addition and subtraction are performed in binary form also.

A second method partitions the data into blocks. Data within a block may be permutated bit by bit or substituted by some other data in a manner determined by the switch setting, which is often called a key. At the receive side, the reverse permutation or substitution is performed to recover the data. Devices using this method are known as block ciphers. An example of block ciphers is the Data Encryption Standard (DES), accepted as a standard by the National Bureau of Standards in 1977. DES, which uses both permutation and substitution iteratively, is approved for nonmilitary, nondiplomatic governmental use.

The National Security Agency established a commercial communication security endorsement program under which encryption devices suitable for different applications are developed and certified by the government. *See* CRYPTOGRAPHY.

Video. Typical video scrambling devices used for cable television applications involve modifying the amplitude or polarity of the synchronization signals, thereby preventing the normal receiver from detecting the synchronization signals. A device capable of undoing the modifications is used to recover the signal. Some devices also invert the polarity of the video signal so that the black and white levels are reversed. *See* CLOSED-CIRCUIT TELEVISION.

A more sophisticated technique, used in satellite transmission, introduces a random delay to the active video signal on a line-by-line basis which is determined by a pseudorandom number sequence. Knowing the sequence, the receiver can compensate for the delay and recover the signal. An even more advanced technique called cut-and-rotate has been proposed. This method selects a random point of time in a scan line and cuts the line into two sections. The second section of the line is transmitted first, followed by the first section. By knowing exactly where the cut takes place, the receive side can perform the reverse process and reconstruct the signal.

Video signals can also be digitized by a number of coding techniques and then scrambled by any of the data-scrambling techniques discussed above to achieve high security. *See* TELEVISION. Lin-Nan Lee

Communications systems protection

The protection of wire and optical communications systems equipment and service from electrical disturbances. This includes the electrical protection of lines, terminal equipment, and switching centers, and inductive coordination, or the protection against interference from nearby electric power lines.

Sources of disturbance. The principal sources of destructive electrical disturbances on wire communications systems are lightning and accidental energization by power lines. Lightning that directly strikes aerial or buried communication cable may cause localized thermal damage and crushing. Simultaneously, it energizes the communication line with a high-level voltage transient that is conducted to terminal equipment. Indirect strikes are more common than direct strikes and, although they normally do not cause mechanical damage, they propagate electrical transients along the line. The rate of occurrence of these transients is geographically dependent, increasing with the level of thunderstorm activity (keraunic level) and earth resistivity. The distance over which a buried cable collects a lightning strike, and the resulting voltage produced within a cable, both increase as resistivity of the earth increases. The duration of the lightning-caused voltage transient on shielded communication cable is typically less than 1 or 2 milliseconds. The magnitude of the voltage between conductors and shield of a cable can reach several thousand volts. In regions of high thunderstorm activity and earth resistivity, a communication line several miles long may experience annually more than 1000 voltage transients of magnitude greater than 250 V. The accompanying current surge usually is less than 100 A, but much higher currents can result from a direct strike to the cable. *See* LIGHTNING.

During fault conditions on commercial power lines, high voltages may occur in nearby communication lines by several mechanisms. The most common is magnetically induced voltage caused by the high unbalanced currents of a phase-to-ground power-line fault. During such an event, an aerial or buried communication line that is parallel to the faulted power line intercepts its time-varying magnetic field, incurring a high longitudinal voltage. In effect, the system is a one-turn to one-turn air-core transformer. This inductively coupled voltage may be several hundred volts, with current limited by the impedance of the communication line and its terminations. Another source of high voltage is a damaged aerial power line that falls and directly contacts an aerial communication cable, or a digging machine that simultaneously contacts power and communication cable jointly buried in a common trench. In these cases, currents conductively enter the communication lines and are limited in magnitude only by fusing capabilities. Finally, communication lines near a power station or transmission-line tower may be subjected to high voltages from the rise in ground potential as power fault currents return through the ground mat of the power station. In each of these mechanisms, power-line fault-clearing devices are intended to limit duration of conduction to a few seconds. However, if the power-line fault impedance is high, the automatic clearing devices may not operate, and the overvoltage can last indefinitely. *See* ELECTRIC POWER SYSTEMS; ELECTROMAGNETIC INDUCTION; GROUNDING; TRANSFORMER; TRANSMISSION LINES.

In city centers, the diversion of lightning strikes by steel-framed buildings, and the shielding effect of the many underground metallic utility systems, considerably reduce the probability of high-voltage transients from lightning or induction from power lines. Since communication and power facilities are routed in separate conduits, the possibility of a power contact is remote. *See* ELECTRICAL INTERFERENCE; ELECTRICAL NOISE.

Electrical protection. The components of a communications system requiring electrical protection include both metallic and optical-fiber lines, terminal equipment, and switching centers.

Metallic lines. Metallic communication lines often are made up of many closely spaced pairs of wires arranged as a cable. A grounded circumferential metallic shield on the cable reduces the magnitude of electrical transients from nearby lightning strikes, and also can intercept a low-current direct strike, minimizing damage to the internal conductors. The effectiveness of the shield is improved if its resistance per unit length is low and if the dielectric strength of the insulation between the shield and the internal conductors is high. Frequent grounding of the shield eases the return of lightning currents to earth, thereby reducing voltages in the cable. Voltage breakdown within the cable can be avoided by installing surge protective devices between the pairs and shield at stress points along the cable. The devices may be air gaps consisting of carbon electrodes separated by a few thousandths of an inch (about 0.1 mm), or gas tube surge arresters, which use metallic electrodes separated by a gap in an enclosed discharge medium. They limit voltages at their terminals to the capabilities of the cable insulation, and are used principally in areas of high lightning activity. *See* COMMUNICATIONS CABLE; ELECTRICAL SHIELDING; GAS TUBE; LIGHTNING AND SURGE PROTECTION; SURGE SUPPRESSOR.

A special case for this type of protection arises on lines that serve antennas of wireless telecommunications systems. Lightning striking an antenna tower propagates voltage surges along the cable. If the cable is not well shielded and insulated, damage can occur, sometimes at considerable distance from the tower. Surge protectors placed between lines and shield at several locations within about a mile (1–2 km) of the tower help avoid this damage.

The localized damage caused by a direct strike to buried shielded cable can be reduced or eliminated by the burial of one or more wires at least 1 ft (0.3 m) above the cable. These shield wires divert all or much of the lightning current from the cable.

Damage to cable plant from a power-line contact is minimized by providing frequent bonds between the cable shield or aerial support strand and the neutral conductor of the power line. These bonds create low-resistance paths for fault currents returning to the neutral, and hasten deenergization by power-line fault-clearing devices. Closely spacing the bonds limits the length of communication cable that is damaged by the contact.

The outer metallic shields of belowground cables may be attacked by electrolytic action and require protective measures against corrosion. Electrolysis occurs where direct current leaves the shield. Severe sources of electrolytic currents include stray direct current from electrified railways and from industries using large amounts of direct current in grounded circuits. An effective remedial measure is to install a gap, known as an insulating joint, in the cable shield to interrupt the current. This gap can hinder the electrical protection and noise mitigation function of the shield unless it is bridged with a capacitor that provides a bypass for lightning and power-frequency currents. *See* CORROSION.

Optical-fiber lines. Optical-fiber communication lines are enclosed in cables that may contain metallic components to provide mechanical strength, waterproofing, local communications, or a rodent barrier. Though the cables otherwise would be immune to the effects of lightning or nearby power lines, such metallic components introduce a measure of susceptibility that is made all the more important by the high information rates carried by the fibers. A direct lightning strike to a metallic component of buried optical-fiber cable can cause localized thermal damage, arcing, and crushing that together may damage the fibers. Electrical protection is provided by cable designs that withstand these effects. As with metallic lines, the probability of this damage can be reduced by burying one or more shield wires at least 1 ft (0.3 m) above the cable.

Alternating currents may be conducted on the metallic sheath components of optical-fiber cables during an accidental contact with power-line conductors. Cable damage is minimized in extent by bonding the sheath to the neutral conductor of the power line, and by providing enough conductivity to carry the currents without damage to the fibers. *See* OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

Some networks provide high-speed data via optical lines between a communication center and a pedestal near the served premises, and then use metallic lines to enter the premises (fiber-to-curb system). Often the power for equipment in the pedestals comes from a centralized powering node over metallic lines. Surge protectors are placed at each end of these powering lines to help avoid equipment damage. Because the power on these lines may exceed those of ordinary telephony, care should be taken to ensure that the protectors can reset after conducting a surge. These protectors, and the protectors on the lines between pedestal and premises, should not have too low a limiting voltage or they may needlessly cause data disturbances during thunderstorms or power-line malfunctions.

Terminal equipment. To protect the users, their premises, and terminal equipment, communication lines that are exposed to lightning or contacts with power lines are usually provided with surge protectors. Article 800 of the National Electrical Code requires that communication lines exposed to contact with power lines of voltages greater than 300 V be equipped with a protector at the entrance to the served premises. Interbuilding lines that are exposed to lightning also have protectors. Although not required, it is common practice to minimize equipment damage by so equipping communication lines in an area of significant lightning exposure even if there is no power-contact hazard. Article 830 extends similar requirements to coaxial circuits that provide network-powered broadband communications. *See* ELECTRICAL CODES.

For the most part, the protector contains a voltage-limiting device, such as a gap-type or solid-state surge protector, connected between each wire of the line and a local ground electrode. Any voltage surge on the wires thereby is limited with respect to local ground. Voltages between the communication lines, power lines, and other grounded systems within the premises are limited by bonds between the protector ground, the power neutral, and other metallic systems.

Sparking between communication lines and nearby objects, and electric shock hazard to the user, are minimized if the insulation of the lines and terminal equipment withstands voltages at least as great as the limiting voltage of the protector (plus an allowance for voltage drop of grounding conductors). Service interruptions are avoided if electronic circuits within the equipment also are able to withstand these surge voltages. Within some equipment, surge immunity is achieved by applying additional voltage limiters, such as varistors, avalanche diodes, or special low-voltage gas tubes. *See* MICROWAVE SOLID-STATE DEVICES; VARISTOR.

Sustained currents from power-line contacts can overheat the network-side wires connected to the protector unless interrupted by a current-limiting device. A short length of fine-gage wire in series with the communication line, or a fuse at the protector, is usually installed for that purpose. If the terminal equipment contains an internal voltage limiter, an additional fuse may be needed to protect the premises wire connecting the equipment to the protector. *See* FUSE (ELECTRICITY).

Switching centers. Switching centers contain electronic equipment that routes communications to their proper destinations. This equipment is protected from the effects of electrical transients appearing at interfaces with external communication lines in a similar way as for terminal equipment. In addition to these effects at external interfaces,

communication equipment within centers may be subjected to transient voltages resulting from lightning strikes to nearby objects or from direct strikes to the building. Such events cause rapidly varying currents in metallic structural members and in the grounding system. These currents can be conductively and magnetically coupled to wires interconnecting equipment in the building. Depending on the relative paths taken by the lightning currents and the wiring, transient differences in voltage may appear between nearby ground and conductors that extend between internal locations.

One method of controlling these voltages is to bond, as often as possible, all metallic equipment frame members and grounded circuits to the floor ground of the building. This technique, known as an integrated ground (or mesh bonding network), reduces potential differences, but permits surge currents to circulate throughout the system. Sensitive electronics can be damaged by the local potential differences caused by these circulating currents, so some designs avoid circulating currents by isolating the entire system from ground except at one location, which is deliberately grounded. This system, known as the isolated ground plane (or isolated bonding network), has found effective use in many switching centers. *See* ELECTRIC PROTECTIVE DEVICES; ELECTROMAGNETIC COMPATIBILITY; ELECTRONIC EQUIPMENT GROUNDING; SWITCHING SYSTEMS (COMMUNICATIONS).

Inductive coordination. Inductive coordination refers to measures that reduce the magnitudes and effects of steady-state potentials and currents induced in metallic communication lines from paralleling power facilities. Where power and communication lines are parallel for long distances, normal-load power currents returning to the power generator via the neutral conductor or earth induce voltages into the communication lines. Malfunction of terminal equipment may result. The induced voltages appear equally in each wire of a pair; electrical asymmetries in the lines or terminal equipment convert them to a voltage between the wires of a pair and may cause background noise or data errors. Remedial measures that reduce the induction effect are (1) balancing the power-line currents to reduce the neutral component returning to the generator; (2) using a grounded metallic shield on the communication cable so that shield currents partly cancel the field from the disturbing power line; (3) using well-balanced communication line and terminal equipment to minimize conversion to voltages between wires; (4) installing reactive components in the communication line to reduce induced currents; and (5) using terminal equipment that is tolerant of induced voltages. *See* INDUCTIVE COORDINATION.

Solar magnetic storms. Solar magnetic storms, which occur with varying intensity over an 11-year cycle, produce strong currents in a predominantly east-west direction of the earth surface. They cause a voltage difference between a conductor and the earth that increases at higher latitudes and with earth resistivity. The voltage typically lasts for several minutes and is low frequency (below 0.1 Hz); it can be treated as a direct-current voltage on telecommunication and power-line circuits.

Long-distance metallic lines are susceptible to these storms to the extent that they have conductive connections to earth, either deliberately or by means of surge protective devices that are operated by the storm. The increasing use of optical-fiber lines for long-distance telephony largely removes this threat. Where protectors are used on long metallic lines, they should be capable of continuously conducting up to 1.0 A without posing a fire hazard. Short metallic lines are unlikely to experience enough voltage to operate surge protectors. If terminals on these lines have conductive connections to earth, ability to operate in the presence of 3-V direct current between ends of the line accommodates all but extreme magnetic storms.

Power utilities have had widespread outages during magnetic storms, and this loss of utility power poses a threat to metallic and optical-fiber communication systems. Reserve power systems using batteries in conjunction with engine alternators help provide uninterrupted telecommunications service for the duration of an extended power outage. *See* GEOELECTRICITY; GEOMAGNETIC VARIATIONS; UNINTERRUPTIBLE POWER SUPPLY.

Nuclear electromagnetic pulse. A nuclear burst at altitudes higher than about 25 mi (40 km) generates an electromagnetic field pulse over a large area of the earth surface; a burst at 300 mi (500 km) would illuminate most of the contiguous United States. The magnitude of the electric field would peak at tens of kilovolts per meter in a few nanoseconds and would last a few hundred nanoseconds. It would affect communication systems by causing high-voltage transients on wires that enter electronic equipment at telecommunications centers. The field would couple to the wires by two mechanisms: direct penetration through walls of the building into the wiring; and coupling of current pulses onto conductors that enter the building (such as power lines, telecommunications cables, and antenna leads), which in turn couple surges to the internal wiring. The surges, though brief, may be much larger than normal signal levels and can cause malfunctions or damage. The protection measures summarized here are among those considered to be reasonable and practical for public communication networks.

Direct field penetration through the structure is reduced by enhancing its shielding. This is done by selecting construction methods with electrically interconnected components in walls, floors, and roof. Poured concrete with interconnected steel reinforcement is preferred. The interconnected steel should be bonded to the common bonding network of the building. Where this construction is unavailable, interior metallic structures, such as cable racks, supports, raceways, and equipment cabinets, should be interconnected and bonded to the common bonding network. Critical equipment and its wiring should be distant from unshielded windows.

The effect of surge currents entering via penetrating conductors is best reduced by diverting to ground much of the surge current before it enters the building. This is largely done by directly bonding to the grounding system all entering telecommunication cable shields and antenna lead-ins, and by installing surge protectors on mains conductors at the service entrance. Because these means do not completely remove penetrating currents, critical equipment and its wiring should be separated from penetrating conductors. *See* ELECTROMAGNETIC PULSE (EMP).

Michael Parente

Bibliography. *Above-Baseline Electrical Protection for Designated Telecommunications Central Offices and Similar-Type Facilities Against High-Altitude Electromagnetic Pulse*, ANSI T1.320–1994, ATIS, 1994; R. H. Golde, *Lightning*, vol. 2, Academic Press, 1977; *Guide for the Application of Gas Tube Arrester Low-Voltage Surge Protective Devices*, ANSI/IEEE C62.42–1992, 1992; *National Electrical Code*, ANSI/NFPA 70, National Fire Protection Association, 1999.

Commutation

The process of transferring current from one connection to another within an electric circuit. Depending on the application, commutation is achieved either by mechanical switching or by electronic switching.

Mechanical switching. Commutation was conceived over a century ago through the invention of the direct-current (dc) motor. When direct current is supplied to a winding on a rotor that is subjected to a stationary magnetic field, it experiences a rotational force and resulting output torque. As the stator north and south poles are reversed relative to the rotating winding, the rotor current is reversed by a commutator in order to maintain the unidirectional torque required for continuous motor action. *See* DIRECT-CURRENT MOTOR; WINDINGS IN ELECTRIC MACHINERY.

The principle is illustrated in **Fig. 1**. In its simplest form, a single rotor winding is connected between two segments of a cylindrical copper commutator which is mounted axially on the rotor. Connection to the external dc supply is through sliding carbon contacts (brushes). The segments have small insulated gaps at *A* and *B*. As *A* and *B* pass the brushes, the current in the rotor winding reverses. In the short interval where the brushes short-circuit the segments, the rotor current decays before building up in the reverse direction. The angular position of the brushes is selected to reverse the current at the appropriate rotor position. *See* COMMUTATOR.

Commutation not only reverses the current but automatically synchronizes the timing of the reversal with the rotor position, irrespective of motor speed. The process can also be considered as a mechanical alternating-current/direct-current converter with the frequency of the rotor current being matched to the speed. The same commutation permits the reverse conversion in a dc generator where alternating-current (ac) voltages induced in a driven rotor are converted to a dc output voltage. *See* DIRECT-CURRENT GENERATOR.

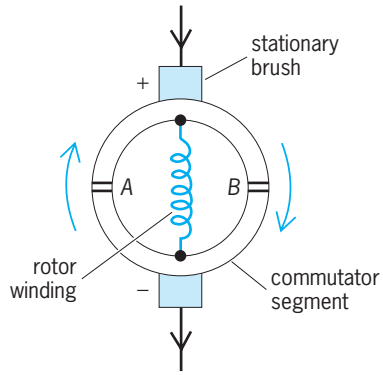

Fig. 1. Basic commutator for a dc motor.

Practical dc motors and generators typically incorporate several winding sections positioned around the rotor and connected to a correspondingly increased number of commutator segments to produce smoother performance.

The same principle of commutation applies to the ac commutator motor and universal ac/dc motor, which are common in variable-speed kitchen appliances and electric hand tools. The commutator still reverses the connection to the rotor to accommodate rotor position. When ac voltage is supplied to both commutator and stator winding, current reversals in the ac waveform are common to both the rotor and the stator thus torque remains unidirectional. *See* ALTERNATING-CURRENT MOTOR; UNIVERSAL MOTOR.

Electronic commutation. The equivalent of mechanical commutation occurs in solid-state converter circuits such as those used for rectifying ac to dc or inverting dc to ac. **Figure 2** shows a three-phase converter widely used in industry. For simplicity, the ac supply network is represented by equivalent phase voltages in series with the effective supply inductance. (Often this inductance is mainly the inductance per phase of a converter transformer that interfaces the converter and the three-phase supply.) Usually, supply resistance is relatively low and plays a negligible role in the converter action. As shown, thyristors 1 and 2 are conducting the dc current from phase *a* to phase *c*. A smooth dc current does not produce a voltage across the inductance *L* in each phase.

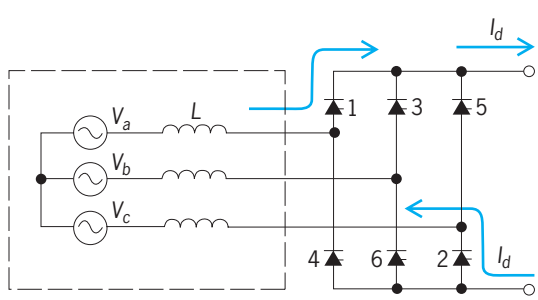

Fig. 2. Three-phase converter.

In the cyclic conduction sequence, the dc current is commutated from phase *a* and thyristor 1 to phase *b* and thyristor 3. To achieve this, thyristor 3 is gated in a region of the ac waveform when its forward voltage is positive. Turning it on applies a reverse voltage to thyristor 1 (phase *b* being more positive than phase *a*), which ceases conduction to complete the commutation of the dc current. This is repeated in sequence for the other thyristors in each ac cycle. *See* CONVERTER; SEMICONDUCTOR RECTIFIER.

In practice, the commutation is not instantaneous because of the inductance, which is referred to as commutation inductance or commutation reactance when considered as an ac impedance. The buildup of current in the incoming phase matches the reduction of current in the outgoing phase over a certain electrical angle until commutation is complete. The angle is referred to as the overlap or commutation angle and may typically be 20–30 electrical degrees. *See* INDUCTANCE.

Meanwhile the dc waveform, which is composed of segments of the ac waveform as selected by the switching action, is affected by the overlap because of the transient commutation voltage drop in the inductances. On average, with consideration of six commutations per ac cycle, the dc voltage may be reduced by several percent, sufficient for commutation to be incorporated into conversion calculations. The overlap angle is a variable and becomes larger for a larger dc current to be commutated, for a larger commutation reactance, and for a smaller ac voltage. It also depends on the actual values of ac voltage at the point-on-wave at which a thyristor is gated to turn on.

Artificial (forced) commutation. The ac supply voltages in Fig. 2 provide a natural commutation voltage. In some other circuits, additional circuit means are necessary to provide a reverse voltage for the outgoing thyristor (or other solid-state switching device). They include charged capacitors, voltage reversals from resonant inductance–capacitance circuits, and additional thyristors.

Electrical noise. Commutation, whether natural or forced, creates abrupt transitions in the converter waveforms, with high-frequency components as an unwanted by-product. This is a factor in assessing radio-frequency interference in the vicinity of power electronic equipment and supply leads. *See* ELECTRICAL INTERFERENCE; ELECTRICAL NOISE.

Commutation with dc converter stations. Converters at the very high power levels of dc transmission are based on natural commutation of the converters shown in Fig. 2. As for many industrial applications of the same circuit, the timing of the commutation process is determined by real-time automatic control. In controlling the dc voltage and current, the commutation must be confined to points in the ac waveform for which the commutating voltage is conducive to successful completion of the commutation. Theoretically, the gating of a thyristor can be delayed by a half cycle (180°) from the point-on-wave where its voltage becomes positive, in order to have a controlled transition from full dc voltage as a rectifier to full reverse dc voltage as an inverter. In practice, commutation imposes constraints at both extremes of control: delay with a minimum value of 5–7° provides a sufficient commutation voltage to avoid hesitant thyristor turn-on; at the other extreme, the delay must still leave sufficient time for the overlap angle together with a period for the outgoing thyristor to recover its ability to hold off conduction. *See* DIRECT-CURRENT TRANSMISSION.

In establishing the maximum permitted delay (firing angle) for inverter operation, the control has to take into account variations in the overlap angle (as noted previously) to maintain the recovery period referred to as the extinction or commutation margin angle, typically 15–18°. Should a delayed commutation extend into this margin, there is the prospect of commutation failure, which implies that the orderly converter switching is disturbed, resulting in a collapse of the converter dc voltage. This power interruption is temporary if the control returns the firing angle to a safe region. A disturbance in the ac system adjacent to a converting station may depress the ac voltage to a level at which commutation cannot be completed. This is also designated as a commutation failure. The recovery of the converter to the previous power level commences when the voltage is restored. John Reeve

Bibliography. J. Arrillaga, *High Voltage Direct Current Transmission*, 2d ed., 1998; B. J. Chalmers, *Electric Motor Handbook* 1988; C. W. Lander, *Power Electronics*, 3d ed., 1994; M. A. Rashid, *Power Electronics*, 3d ed., 2003.

Commutator

That part of a dc motor or generator which serves the dual function, in combination with brushes, of providing an electrical connection between the rotating armature winding and the stationary terminals, and of permitting the reversal of the current in the armature windings. For explanation of the necessity of this function *see* COMMUTATION.

A commutator (**Fig. 1**) is composed of copper bars assembled to form a drumlike cylinder which is concentric with the axis of rotation. Insulation, commonly mica, to provide exceptional mechanical and electrical stability, is placed between commutator bars and between the bars and the shaft.

Conducting brushes (**Fig. 2**), commonly carbon, sufficient in size and number to carry the current, are spaced at intervals of 180 electrical degrees about the surface of the commutator and held in contact with the surface of the commutator by spring tension. In special cases, as when only the top or bottom half of a dc motor is accessible, only one positive brush set and one negative brush set are used when the armature is wave-wound. This is permissible, as wave windings characteristically have two parallel paths between any two commutator bars separated 180 electrical

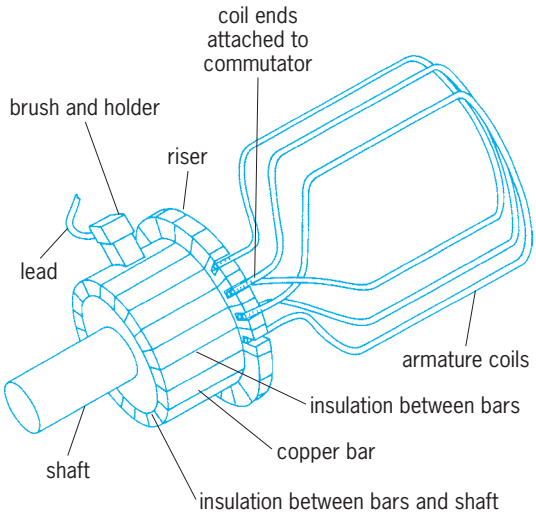

Fig. 1. Commutator and brush assembly with coil connections for lap winding.

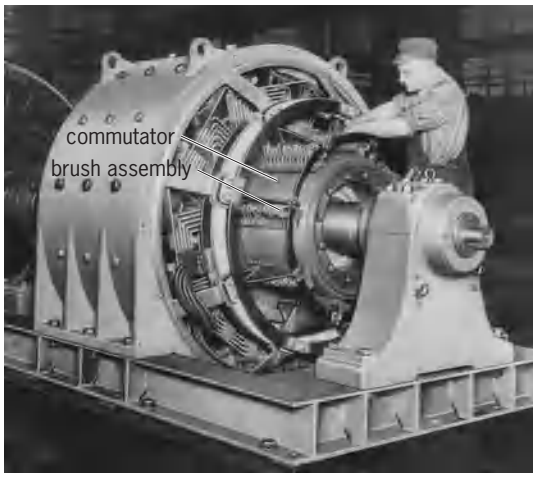

Fig. 2. Direct-current generator showing the typical brush assembly. (*Allis-Chalmers*)

degrees. Machines employing commutators are usually restricted to voltages below the kilovolt range because flashover between the closely spaced commutator segments may result from deposition of carbon dust from the brushes and from uneven voltage distribution between commutator bars under heavy loading. For machines using commutators *see* DIRECT-CURRENT GENERATOR; DIRECT-CURRENT MOTOR; REPULSION MOTOR; UNIVERSAL MOTOR; WINDINGS IN ELECTRIC MACHINERY.

Arthur R. Eckels

Compact disk

A system for data storage in which digitally encoded information in the form of microscopic pits on a rotating disk is accessed by optical readout. The compact disk was originally developed as a music carrier providing high fidelity, random access, convenience, durability, and low cost. Its attributes made it suitable for storing diverse data such as video programs and computer software, and improvements allowed recordability and erasability. Greater storage capacity and more sophisticated integration of features is provided in the DVD optical disk format.

Operating Principles

An acoustic signal waveform is stored on the disk in the form of a binary code, as a series of 0's and 1's. This is done by forming pits along spiral tracks on a transparent plastic disk, overlaying this with a reflective coating, and covering this coating with a protective layer. The light from a semiconductor laser is focused onto the pits from below (**Fig. 1**). The presence or absence of pits within the laser spot changes the intensity of the reflected beam (pits diffract the light, reducing reflected intensity). The reflected light strikes a light-detecting photodiode that converts the varying-intensity light beam into a binary electrical signal. *See* LASER.

With the light focused on the pits, the outer lower surface of the transparent disk is out of focus; hence dust, fingerprints, and minor scratches on this surface are blurred and do not affect the detection process. Even more important, the use of an optical pickup means that there is no mechanical contact and hence no wear on pickup or disk.

Because the focused laser spot size is extremely small, a considerable amount of information can be placed on the data surface. Adjacent tracks of the spiral of pits are nominally 1.6 micrometer apart (track pitch), and hence about 22,200 such tracks are placed across a 35.5-mm radius on a 120-mm-diameter (4.75-in.) disk. The track is optically scanned at a nominal constant linear velocity of about 1.4 m/s. The pits, and the separations between pits along a track, vary in length from about 0.97 to 3.56 μm (**Fig. 2**). Smaller track pitch and lower linear velocities (and shorter pit and land lengths) may be used for greater storage capacity.

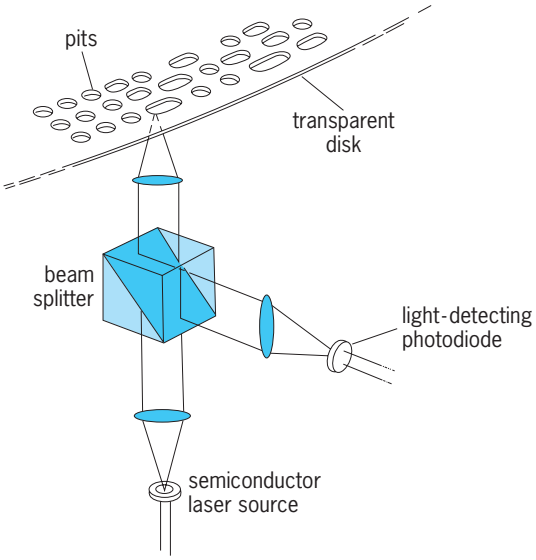

Fig. 1. Optical readout system for an audio compact disk.

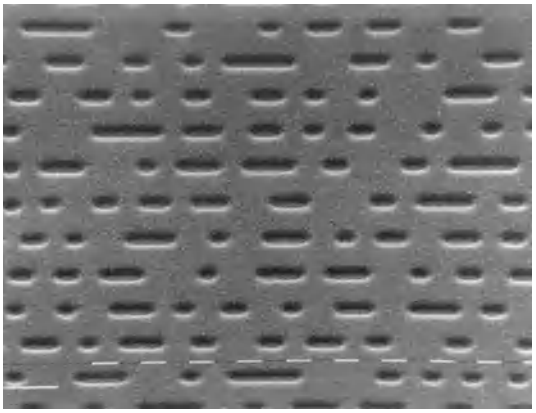

Fig. 2. Pits on an audio compact disk. (*From M. G. Carasso et al., Compact disc: Digital audio, Philips Tech. Rev., 40(6):151–180, 1982*)

Digitally encoded waveform. The human ear is insensitive to sounds with frequencies higher than about 20,000 Hz. The Shannon (or Nyquist) sampling theorem specifies that a waveform can be exactly reproduced from samples taken at a uniform rate of at least twice that of the highest signal frequency. The standard sampling rate for audio compact disks is 44.1 kHz; thus, the highest audio frequency encoded is 22.05 kHz. The amplitude of each sample is represented by a 16-bit binary word. Thus, the amplitude of each original audio sample is approximated to the nearest of 2^{16} (65,536) equally spaced levels in a process known as quantization. With stereo channels, the data rate stored on the disk is $2 \times 16 \times 44{,}100 = 1{,}411{,}200$ bits/s. *See* INFORMATION THEORY; PULSE MODULATION.

An important audio performance figure is the signal-to-noise ratio. If it is assumed that all unwanted noise and interference have been eliminated, the remaining noise will be that due to the quantization process, with an amplitude of the order of one-half of a quantization level. The theoretical signal-to-noise ratio, the amplitude of the largest undistorted signal to the amplitude of the noise, is $S/N = 6.02(16) + 1.76 = \sim 98$ dB. This is comparable, for example, to the figure of 90 dB for the dynamic range of a symphony orchestra. As a comparison, the best signal-to-noise ratio figure for professional magnetic tapes is about 80 dB, and acceptable signal-to-noise ratios for conventional audio systems are about 70–80 dB.

Disk mastering and replication. In the disk mastering and replication process (**Fig. 3**), a glass disk is covered by a uniform coating of photoresist material. A laser exposes portions of the photoresist where pits are to be formed. The photoresist is then developed and washed, leaving the master recording. A nickel mother is derived from this master and is then used as a mold to produce multiple copies of the disk in transparent polycarbonate plastic. These substrates are coated with a thin metallic reflecting layer (usually aluminum), with a protective plastic coating on top of that.

Data is coded on a compact disk using an EFM (eight-to-fourteen modulation) code. In EFM, single binary 1's are represented by both the beginning and end of a pit length; otherwise the signal is a sequence of 0's. This correspondence between the channel bit stream (the signal represented on the disk) and the cross section of a disk is shown in Fig. 3*e* and *f*.

Electromechanical control. Advanced control systems in the compact disk player are designed to retrieve data from the disk surface reliably. The laser spot on the pit track must stay focused to within $+/-0.5\ \mu$m. But since the vertical movement of the disk might be as much as 1000 times this, automatic focusing must be provided. In some designs, this is accomplished by inserting a cylindrical lens in the path of the reflected beam and allowing the beam to impinge on a four-quadrant segmented photodiode. The changing ellipticity of the focused spot is registered as a differential signal between two opposite pairs of photodiode segments, which provides the control signal for the focusing servomechanism. *See* SERVOMECHANISM.

The laser spot must track the pits with a tolerance of $\pm 0.1\ \mu$m. However, track eccentricity might be $100\ \mu$m or more. Thus, an automatic tracking system

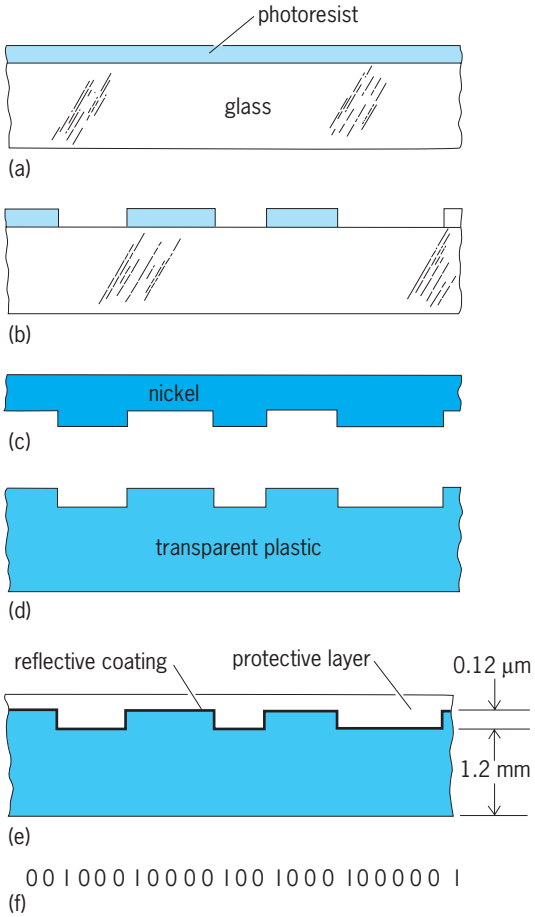

Fig. 3. Mastering and replication process for an audio compact disk. (*a*) Glass disk covered by a uniform layer of photoresist. (*b*) Pits formed by a focused laser in the photoresist, which is developed and washed away, giving the master. (*c*) Nickel mother, formed from the master. (*d*) Transparent plastic disk stamped from the nickel mother. (*e*) Disk with the reflective coating and the protective layer added. (*f*) Associated binary electrical signal (channel bit stream).

is required. In some designs, lateral tracking control is achieved by generating two tracking laser spots, with one straddling each edge of the track when the main spot is properly centered. The differential signal from two photodiodes positioned to pick up the reflection from the tracking spots provides the lateral position control signal.

The constant linear velocity of the beam along the track is ensured by synchronizing the data read off the disk with a quartz-crystal oscillator. *See* OSCILLATOR.

Signal processing. Since the information that is read off the disk is in digital form, as a sequence of 0's and 1's, it can be processed in many ways that are not possible with analog systems. To enable this, information can be passed through a buffer memory, and then output at a rate that is controlled by the player's quartz-crystal (oscillator) clock, hence entirely eliminating the wow and flutter of conventional systems.

Error correction. Errors in the information read off the disk fall into two broad categories: (1) An error in an individual bit may be due to an imperfectly formed pit or random imperfection formed in the plastic disk during the molding process. (2) A burstlike error produced by severe scratches or dirt on the outer layer of the transparent disk affects many adjacent tracks.

In principle, the detection and correction of errors is achieved by adding extra check bits known as parity bits. The bit stream can be divided into blocks and the bits in a block arranged in the form of a rectangular grid or matrix. If a parity bit is assigned to every row and every column, then, on the assumption that only one bit per block is in error, it is possible to identify precisely which bit is in error, and hence to correct it. In practice, extensions of this idea are used in the detection and possible correction of compact disk data errors, but implementation is more complex. In particular, an algorithm known as the cross interleave Reed-Solomon code (CIRC) is used.

Conversion of signal to analog form. To recreate the original music signal, the binary data on the disk must be passed through a digital-to-analog (D/A) converter and a low-pass filter. The digital-to-analog converter accepts each 16-bit sample and outputs a voltage corresponding to its value. The series of samples forms a staircase waveform that is applied to a low-pass filter that removes all frequencies above the half-sampling frequency. In this way, the filter reconstructs the original waveform. A steep (high-order) analog filter can be used, with a flat amplitude to the half-sampling frequency, at which point the amplitude falls to zero. *See* DIGITAL-TO-ANALOG CONVERTER; ELECTRIC FILTER.

In practice, low-pass filtering is accomplished using a technique called oversampling, in which the output sampling rate is increased by a factor of, say, 8, from 44.1 to 352.8 kHz, with a digital filter. This technique has the effect of requiring only a mild (low-order) analog filter while the audio bandwidth remains limited to 20 kHz. Oversampling yields an output signal characteristic with essentially flat frequency response and phase linearity.

Diverse Formats

The compact disk was originally designed as a music carrier. Subsequently, the format was extended to include the CD-ROM (read-only memory) format for computer applications. Newer formats that use the compact disk as their basis include the recordable CD-R and erasable CD-RW formats, and the multimedia DVD format.

CD-R. The CD-R (recordable) format allows users to record audio or other digital data on a compact disk; it is also called the CD-WO (write-once) format. The recording is permanent and may be read indefinitely, but cannot be erased. CD-R disks with up to 80 min (or about 700 megabytes) of playing time are available. A recording is complete when a leadin area (with table of contents), user data, and leadout area have been written.

CD-R disks are built on a polycarbonate substrate and contain a reflective layer and a protective top layer. Sandwiched between the substrate and the reflective layer is a recording layer composed of an organic dye polymer, generally either cyanine or phthalocyanine. To achieve appropriate reflectivity, a gold or silver reflective layer is used. Unlike regular compact disks, CD-R disks are manufactured with a pregrooved track, which is used to guide the recording laser. This practice simplifies recorder hardware design and promotes disk compatibility. The organic dye recording layer is a photoabsorption surface that absorbs energy from the recording laser as heat. The laser passes through the polycarbonate substrate and heats this layer, causing the substrate layer to expand into the absorption layer and mix with the dye materials there. The polycarbonate mixed with dye, which has been decomposed from the heat, acts to form a pit in the substrate. These pits create the change in reflectivity required by compact disk player pickups. During readout, the same laser, reduced in power, is reflected from the pits and its changing intensity is monitored at a photodiode. In this way, CD-R disks can be played in most CD audio and CD-ROM drives.

CD-RW. The CD-RW (rewritable) format allows data to be written, read, erased, and rewritten. A CD-RW drive can read, write, and erase CD-RW media, read and write CD-R media, and read CD audio and CD-ROM media. The data can comprise computer programs, text, pictures, video files, or audio files. Altogether, five layers are built on a transparent polycarbonate substrate: a dielectric layer, a recording layer, another dielectric layer, a reflective metal (aluminum) layer, and a top acrylic protective layer. As in the CD-R, the writing and reading laser follows a pregroove across the disk radius. However, whereas the CD-R uses a dye recording layer, the CD-RW format employs a phase-change mechanism. The recording layer is composed of an alloy of silver, indium, antimony, and tellurium, which exhibits a reversible crystalline-amorphous phase change when recorded at one temperature and erased at another. Data are recorded by directing a laser to heat an area of the crystalline layer to a temperature slightly above

its melting point. When the area solidifies rapidly, it becomes amorphous, and the decreased change in reflectivity can be detected. Because the crystalline form is more stable, the material will tend to change back to this form. Thus, when the area is heated again (at a slightly lower laser power) and cooled gradually, it will return to a crystalline state, providing rewritability. A very low power laser is used to read data.

This phase-change recording technology allows thousands of rewrite cycles. The disk capacity is 650 megabytes. Since the reflectivity of CD-RW disks is only about 15 and 25% in the amorphous and crystalline states respectively, they cannot be played in conventional compact disk players. A CD-RW drive is required, or a multiread drive capable of reading lower-reflectivity disks. The CD-UDF (universal disk format) specification modifies the CD Red Book file format so that nonaudio data can be written in data packets. With CD-UDF-compliant drives, CD-R and CD-RW disks can be used in a personal-computer environment much like other storage media; in addition to audio recording, nonaudio data files can be dragged and dropped from the personal computer to disk storage.

DVD. The DVD (digital video disk or digital versatile disk) standard includes formats for video, audio, and computer applications. DVD-video was the first DVD format to be developed. One disk digitally stores a complete feature film with digital audio sound tracks. As with compact disks, DVD-video disks use a pit-and-land structure to store data. The track pitch is 0.74 μm, compared to 1.6 μm on a compact disk. The pits and lands are as short as 0.4 μm, compared to 0.83 μm on a compact disk. These reduced dimensions are possible because the laser beam uses a visible red wavelength of 635 or 650 nanometers, compared with 780 nm in a compact disk. The standard specifies a lens with a numerical aperture of 0.6, compared with 0.45 in the compact disk standard. These specifications provide a large storage capacity.

Data storage. A DVD-video disk is the same diameter (120 mm or 4.75 in.) and thickness (1.2 mm or 0.05 in.) as a compact disk. Whereas a compact disk uses a single polycarbonate substrate, a DVD-video disk employs two 0.6-mm substrates, bonded together, with the data layers placed near the internal interface.

A single-layer, single-sided DVD-video disk uses one substrate with a data surface, and one blank substrate; the disk holds 4.7 gigabytes of data. Two substrates with data surfaces can be bonded together to form a single-layer, double-sided disk holding 9.4 gigabytes of data; the disk is turned over to access the opposite layer.

The DVD-video standard also allows data to be placed on two layers in a substrate, one embedded beneath the other to create a dual-layer disk that is read from one side. The layers are separated by a semitransparent (semireflective) metal layer. Both layers can be read from one side by focusing the reading laser on either layer. The beam either reflects from the lower semireflective layer or passes through it and reflects from the top reflective layer. Together, the dual layers hold 8.5 gigabytes of data. The DVD-video standard allows for a double-sided, double-layer disk, in which two dual-layer substrates are bonded together, creating a capacity of 17 gigabytes.

Data compression. A motion picture may comprise 200 gigabytes of data; to fit this program on a single disk, data reduction must be employed. The DVD-video standard uses the MPEG-2 (Motion Picture Experts Group-2) algorithm to encode its video program. Image data that are redundant, not perceived, or marginally perceived are removed or coarsely quantized. This analysis is carried out for both individual video frames and series of frames. MPEG-2 allows for a variable bit rate. Simple pictures are given a low bit rate, and complex pictures a high bit rate. The maximum output bit rate is 10.08 megabits per second, and the average bit rate might be 4.7 megabits per second, in contrast to the compact disk's fixed rate of 1.41 megabits per second. With MPEG-2, a single-layer DVD-video disk can store up to 133 min of high-quality digital video with digital audio sound tracks; longer motion pictures can use the dual-layer disk design. The video program is stored at 4:2:0 component video; 4:2:0 refers to the ratios of frequencies used to record the different video components (one luminance and two color-difference), and the luminance and color-difference data are kept separate for higher fidelity. *See* DATA COMPRESSION.

Audio. The audio portion of the DVD-video standard provides both multichannel and stereo sound tracks. DVD-video disks can accommodate three independent 5.1-channel sound tracks, each with five main channels and a low-frequency effects channel. The DVD-video standard recognizes two multichannel formats. Disks can carry a digital sound track with a sampling frequency of 48 kHz; using audio data compression, the normal output bit rate is 384 kilobits per second. Disks also carry linear pulse-code-modulation (PCM) digital stereo sound tracks employing sampling rates of either 48 or 96 kHz, and word lengths of 16, 20, or 24 bits. Up to eight independent pulse-code-modulation channels are permitted.

Features. A DVD-video player is often connected to a home theater system and can also play CD music disks. DVD-video disks cannot be played in compact disk players. The DVD-video standard provides features such as normal (4:3) and wide-screen (16:9) aspect ratios, an automatic pan-scan feature, up to nine camera angles and interactive story lines, up to eight language tracks, and up to 32 sets of subtitles. These features are options, and implementation is left to the content provider. The DVD-video format also uses copy protection to prevent digital or analog dubbing of disks. Disks and players are tagged with regional flags so that players will play only disks intended for certain geographical markets.

Formats. In addition to the DVD-video format (formally known as Book B), the DVD family includes

DVD-read only (Book A), DVD-audio (Book C), DVD-write once (Book D), and DVD-rewritable (Book E) as well as other formats and extensions. The DVD format uses the UDF file structure. The DVD-audio format is used to store high-quality multichannel music programs at a variety of sampling frequencies and quantization word lengths. In addition, the DVD-audio format can employ lossless audio data compression to increase storage capacity effectively without degrading audio file content. The other DVD formats are employed in computer applications; Books A and E are often called DVD-ROM and DVD-RAM (random access memory), respectively. DVD-ROM and DVD-RAM drives are connected to personal computers, and function much like CD-ROM drives.

Other optical disk formats under development provide greater capacity and performance. The compact disk and the DVD will further integrate audio and video technologies with the technology of the computer. *See* COMPUTER STORAGE TECHNOLOGY; OPTICAL RECORDING. Ken C. Pohlmann

Bibliography. A. E. Bell, Next-generation compact disks, *Sci. Amer.*, 275(1):28–32, July 1996; M. G. Carasso et al., The compact disk digital audio system, *Philips Tech. Rev.*, 40(6):149–180, 1982; K. C. Pohlmann, *The Compact Disk Handbook*, 2d ed., 1992; K. C. Pohlmann, *Principles of Digital Audio*, 4th ed., 2000.

Comparator

An electronic circuit that produces an output voltage or current whenever two input levels simultaneously satisfy predetermined amplitude requirements. A comparator circuit may be designed to respond to continuously varying (analog) or discrete (digital) signals, and its output may be in the form of signaling pulses which occur at the comparison point or in the form of discrete direct-current (dc) levels.

Linear or analog comparator. A linear comparator operates on continuous, or nondiscrete, waveforms. Most often one voltage, referred to as the reference voltage, is a variable dc or level-setting voltage and the other is a time-varying waveform. When the signal voltage becomes equal to the reference voltage, a discrete output level is obtained. If the time-varying (signal) voltage approaches the reference voltage from a more negative level the output voltage is of one polarity; if it approaches the reference from a more positive value the output is of the opposite polarity.

Operational-amplifier voltage comparator. A very high-gain operational amplifier whose output is inverting with respect to one input terminal and noninverting with respect to the other and whose output voltage is limited at upper and lower levels (usually at voltages near the supply voltage levels, $+V_{\text{supply}}$) may be used as a voltage comparator (**Fig. 1**). The upper and lower output levels are designated as V_{OH} and V_{OL}, respectively. When the input signal v_i equals V_{REF}, the output is approximately in the center of this range. The change in input voltage around V_{REF} required to drive the output to its respective limits is given by Eq. (1), where A is the voltage gain of the

$$\Delta v_i = \frac{V_{OH} - V_{OL}}{A} \tag{1}$$

amplifier. If A is very large, Δv_i approaches 0; as v_i approaches V_{REF} from the negative side, the output switches abruptly from V_{OL} to V_{OH} as soon as v_i becomes bigger than V_{REF}; and as v_i approaches V_{REF} from the positive side, the output switches abruptly to V_{OL} as soon as v_i becomes less than V_{REF}. If v_i and V_{REF} are interchanged, the output polarity shifts are reversed.

Comparators are often designed with a built-in offset voltage; that is V_{OH} and V_{OL} need not be symmetrical about V_{REF}. For example, high and low output levels may be designed as $V_{OH} = +5$ V and $V_{OL} = 0$, which would be compatible with standard saturated logic levels; or as $V_{OH} = 0$ and $V_{OL} = 15$ V, for outputs compatible with standard emitter-coupled logic levels. *See* LOGIC CIRCUITS.

The high-gain operational-amplifier comparator used in the open-loop mode is designated as a nonregenerative comparator because there is no positive feedback path from the output back to the input.

An operational-amplifier comparator connected in a positive-feedback mode (**Fig. 2**) is referred to as a regenerative comparator. If the voltage gain A is

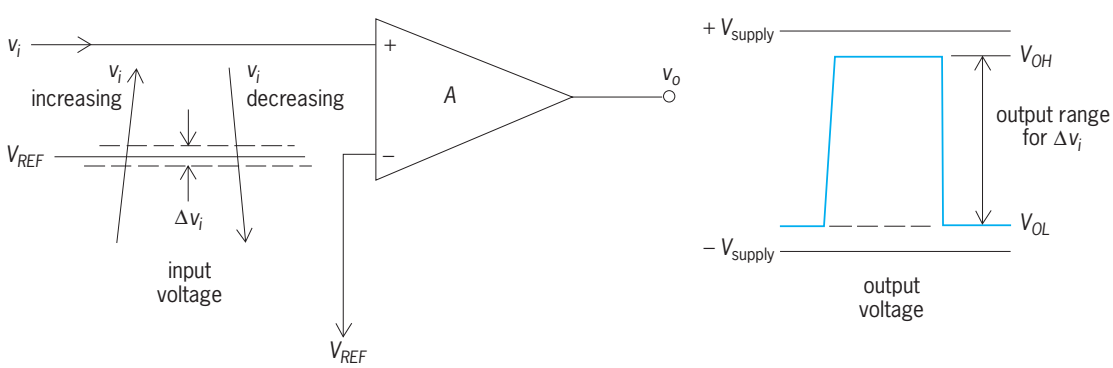

Fig. 1. High-gain nonregenerative operational amplifier as voltage comparator.

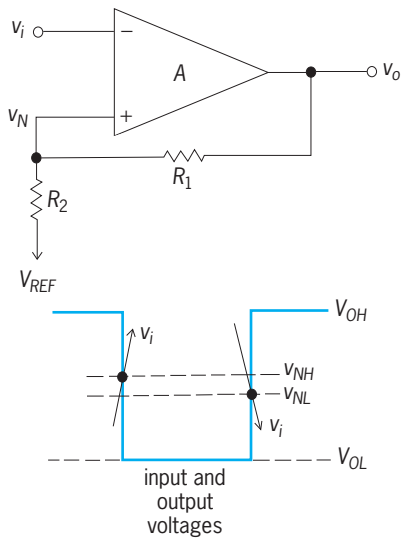

Fig. 2. Regenerative operational amplifier as comparator (Schmitt trigger). The input and output voltages are superimposed, showing hysteresis.

very high, then the feedback voltage v_N is given by Eq. (2) when the output voltage is $v_o = V_{OH}$; and

$$v_N = v_{NH} = \frac{R_2 V_{OH} + R_1 V_{REF}}{R_1 + R_2} \tag{2}$$

when $v_o = V_{OL}$ the feedback voltage is given by Eq. (3). For this circuit there is a difference be-

$$v_N = v_{NL} = \frac{R_2 V_{OL} + R_1 V_{REF}}{R_1 + R_2} \tag{3}$$

tween the input levels which will cause the output to switch, given by Eq. (4). This difference in input lev-

$$v_{NH} - v_{NL} = \frac{R_2}{R_1 + R_2}(V_{OH} - V_{OL}) \tag{4}$$

els is referred to as the hysteresis of the circuit. Such regenerative comparators are generically referred to as Schmitt triggers. Historically, they were first implemented as discrete two-vacuum-tube devices, which were later replaced by transistors and then by the operational amplifier feedback circuit. *See* AMPLIFIER; FEEDBACK CIRCUIT; HYSTERESIS; OPERATIONAL AMPLIFIER.

Digital comparator. The term digital comparator has historically been used when the comparator circuit is specifically designed to respond to a combination of discrete level (digital) signals, for example, when one or more such input signals simultaneously reach the reference level which causes the change of state of the output. Among other applications, such comparators perform the function of the logic gate such as the AND, OR, NOR, and NAND functions. More often, the term digital comparator describes an array of logic gates designed specifically to determine whether one binary number is less than or greater than another binary number. Such digital comparators are sometimes called magnitude or binary comparators.

Applications. Comparators may take many forms and can find many uses, in addition to those which have been discussed. For example, the electronically regulated dc voltage supply uses a circuit which compares the dc output voltage with a fixed reference level. The resulting difference signal controls an amplifier which in turn changes the output to the desired level. In a radio receiver the automatic gain control circuit may be thought of broadly as a comparator; it measures the short-term average of the signal at the output of the detector, compares this output with a desired bias level on the radio-frequency amplifier stages, and changes that bias to maintain a constant average level output from the detector. *See* AUTOMATIC GAIN CONTROL (AGC).

Regenerative comparators (Schmitt triggers) are often used as the key element in the voltage-controlled oscillator (VCO), which in turn is the control element in the phase-locked loop (PLL). Nonregenerative comparators are often used in parallel analog-to-digital converters where a number of comparators (equal to the desired bit rate) are used at a succession of higher V_{REF} levels. *See* ANALOG-TO-DIGITAL CONVERTER; PHASE-LOCKED LOOPS.

Glenn M. Glasford

Bibliography. C. H. Chen (ed.), *Digital Waveform Processing and Recognition*, 2d ed., 1995; G. Glasford, *Analog Electronic Circuits*, 1986; G. Glasford, *Digital Electronic Circuits*, 1988; D. A. Hodges and H. G. Jackson, *Analysis and Design of Digital Integrated Circuits*, 2d ed., 1988; A. S. Sedra and K. C. Smith, *Microelectonic Circuits*, 3d ed., 1992.

Complement

A group of proteins in the blood and body fluids that are sequentially activated by limited proteolysis and play an important role in humoral immunity and the generation of inflammation. When activated by antigen–antibody complexes, or by other agents such as proteolytic enzymes (for example, plasmin), complement kills bacteria and other microorganisms. In addition, complement activation results in the release of peptides that enhance vascular permeability, release histamine, and attract white blood cells (chemotaxis). The binding of complement to target cells also enhances their phagocytosis by white blood cells. The most important step in complement system function is the activation of the third component of complement (C3), which is the most abundant of these proteins in the blood. *See* IMMUNITY.

There are two mechanisms for the activation of C3, the classical and the alternative pathways. The classical pathway requires calcium and is activated when the C1q component binds to immunoglobulin antibodies IgG or IgM. In contrast, activation of the alternative pathway requires magnesium and is not dependent on antibodies. The activation of C3 by either pathway results in the cleavage of C3 to a small fragment, C3a, and a large fragment, C3b. The

latter binds to a bimolecular enzyme, and a new trimolecular complex is formed which cleaves C5 into a small fragment, C5a, and a large fragment, C5b. This is the final enzymatic step in complement activation; the rest of the activation process is nonenzymatic. An amphiphilic macromolecular complex is formed by the binding of one molecule of C5b to one molecule each of C6, C7, and C8, and 12 or more molecules of C9. This so-called membrane attack complex invades the membrane of target cells and kills them. At first these complexes become embedded in target-cell membranes without inflicting any damage. After a molecule of C8 binds to the C5b,6,7 complex, a transmembrane channel (pore) is formed, and the complex can destroy target cells, although very slowly. The process is much accelerated by polymerization of C9 into the complexes, forming a cylindrical structure. The top of the cylinder is hydrophilic and capped by a ring or annulus that remains outside the target cell. The other end is lipophilic; it becomes embedded in the cell membrane and forms large transmembrane channels which coalesce into so-called leaky patches. The target cell loses potassium, imbibes sodium and water, swells, and bursts. This is called cell lysis, or hemolysis in the case of a red blood cell.

As mentioned earlier, several small fragments are enzymatically generated during complement activation, and these act to release histamine from mast cells. The C5a fragment is potent in chemotaxis and in activating phagocytes. The C2b fragment contains a peptide that, upon release by further proteolysis, markedly enhances vascular permeability. Phagocytes and red blood cells have a receptor for fragments C4b and C3b, called complement receptor (CR1), which appears to play an important role in clearing immune complexes from blood. Phagocytes also have a complement receptor 3 (CR3) which, when engaged, stimulates phagocytes to produce superoxide and enhances phagocytic killing of ingested bacteria. A third complement receptor (CR2) is found only on B lymphocytes. *See* ALLERGY; CELLULAR IMMUNOLOGY; HISTAMINE.

Genetic deficiencies of certain complement subcomponents, including properdin, have been found in humans, rabbits, guinea pigs, and mice. All are inherited as autosomal recessives, except for properdin deficiency, which is X-linked recessive, and C1 inhibitor deficiency, which is autosomal dominant. Certain deficiencies lead to immune-complex diseases, such as systemic lupus erythematosus; other deficiencies result in increased susceptibility to bacterial infections, particularly those of the genus *Neisseria* (for example, gonorrhea and meningococcal meningitis), and hereditary angioneurotic edema. *See* COMPLEMENT-FIXATION TEST.

Fred S. Rosen

Bibliography. R. D. Campbell et al., Structure, organization, and regulations of the complement genes, *Annu. Rev. Immunol.*, 6:161–196, 1988; H. J. Muller-Eberhard, The membrane attack complex of complement, *Annu. Rev. Immunol.*, 4:503–528, 1986.

Complement-fixation test

A sensitive reaction used in serology for the detection of either antigen or antibody, as in the diagnosis of many bacterial, viral, and other diseases, including syphilis. It involves two stages: stage I is the binding or fixation of complement if certain antigen-antibody reactions occur, and stage II is detection of residual unbound complement, if any, by its hemolytic action on the sensitized erythrocytes subsequently added. This is shown in the **illustration**. *See* ANTIBODY; ANTIGEN; COMPLEMENT.

In the first stage either the antigen or the antibody must be supplied as a reagent, with the other of the pair as the test unknown. Fresh guinea pig serum is normally used as a complement source. Sheep erythrocytes which are coated with their corresponding antibody (amboceptor or hemolysin) are used in the second stage.

The controls A, B, and C in the illustration demonstrate that sufficient complement is present to effect hemolysis of the sensitized indicator cells and that

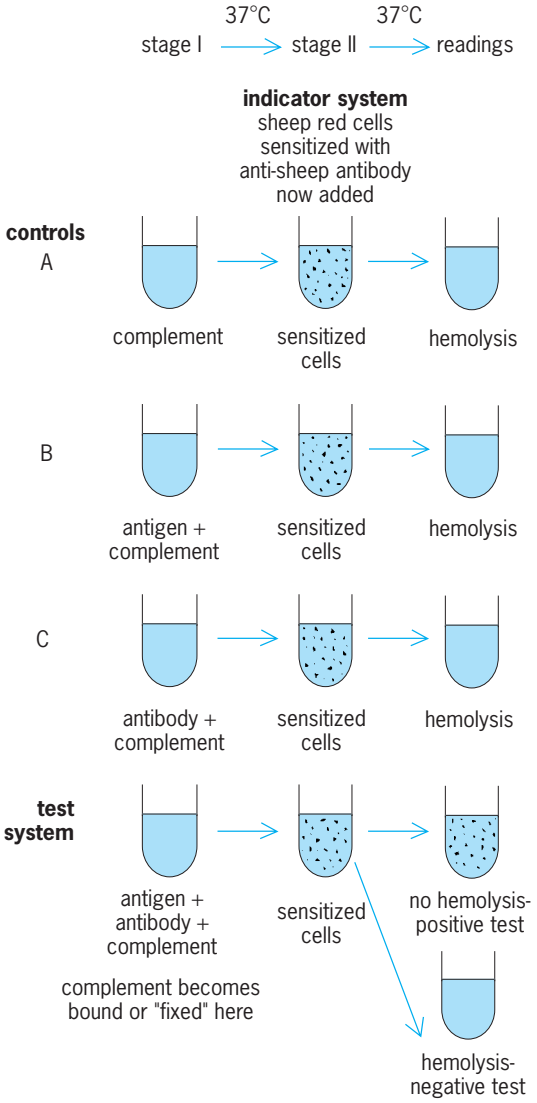

Diagram of the complement-fixation reaction. 37°C = 98.6°F.

neither antigen nor antibody added alone will interfere with this by binding complement. In the test system the combination of a suitable antigen and antibody in the presence of complement will bind the complement to the complex so that the complement becomes unavailable for the hemolysis of the indicator cells added in stage II. If either antigen or antibody is added as a reagent in stage I, then the presence of the other in the test unknown that is added can be detected through its ability to complete the antigen-antibody system. A lack of hemolysis denotes that the complement is bound. This result is termed a positive complement-fixation reaction and, if the system is controlled against false reactions, implies that the antigen (or antibody) assumed to be in the test unknown was actually present. Conversely, complete hemolysis of the sensitized erythrocytes implies, ultimately, absence of antigen or antibody and is termed a negative complement-fixation reaction. In its most sensitive form the microcomplement-fixation test can distinguish the minute antigenic difference that results from the shift in a single amino acid between hemoglobins A and S. *See* IMMUNOLOGY; LYTIC REACTION; SEROLOGY. Henry P. Treffers

Bibliography. W. R. Clark, *The Experimental Foundations of Modern Immunology*, 4th ed., 1991; S. B. Pal (ed.), *Reviews on Immunoassay Technology*, vol. 1, 1988.

Complementation (genetics)

The correction of a cellular defect caused by mutations in a gene by the addition of the normal gene or its product (protein). In cell cultures, complementation can be accomplished by transfer of deoxyribonucleic acid (DNA) or messenger ribonucleic acid (mRNA) that encodes this protein. Another way to "complement" a cellular defect is by joining two chromosome sets in the same cell, each defective due to a different mutation but together able to correct the cellular defect. This principle is used in somatic cell hybridization (or cell fusion), in which cells are fused to generate hybrid cells that possess all chromosomes from both fusion partners. Such hybrids are usually viable and can be examined for the presence of the cellular deficiency via complementation analysis. *See* CHROMOSOME; DEOXYRIBONUCLEIC ACID (DNA); MUTATION.

Complementation analysis. Complementation analysis is used to study the relationships between different genes or alleles which individually cause functional defects in the organism when mutated but in combination restore the wildtype phenotype. It has been very helpful in elucidating the genetic heterogeneity of hereditary diseases, such as Zellweger syndrome, which is caused by malfunctioning peroxisomes (intracellular organelles involved in the metabolism of hydrogen peroxide); xeroderma pigmentosum, which is associated with defective repair of DNA damage caused by ultraviolet light; and Fanconi anemia, a blood disorder associated with chromosomal breakage and a high susceptibility to cancer. These are recessive diseases, which means that both copies of the disease gene are defective; patients inherit a defective copy from each parent, who themselves are carriers of only one defective copy and are asymptomatic.

A condition for carrying out complementation analysis is that a specific disease feature should be recognizable at the cellular level. In the case of Fanconi anemia, for example, this feature is the hypersensitivity of patient cells to DNA cross-linking agents (which promote the joining of two DNA strands or a DNA strand with a protein strand) such as mitomycin C (MMC). When cell lines derived from two unrelated patients with Fanconi anemia are fused, the resulting hybrid cell line may continue to show MMC sensitivity (indicating a lack of complementation), or the cell line may show MMC resistance (indicating that complementation has taken place). *See* PEROXISOME; SOMATIC CELL GENETICS.

Complementation groups. Patients whose cell lines—when combined in a fusion hybrid—fail to complement each other's defect are said to belong to the same "complementation group," whereas when complementation does occur, the cell lines belong to different groups (**Fig. 1**). When multiple complementation groups exist for a disease, each group is thought to represent a distinct disease gene. The explanation is that certain proteins may function as essential components of a specific biochemical pathway and that a defect in any one of these components may cause disruption of that pathway, each defect having essentially the same consequence for the organism. In the case of Fanconi anemia, 12 complementation groups have been distinguished. The disease gene has been identified for 11 of these groups. Eight of these genes encode proteins that aggregate into a multi-subunit complex that protects the cell against certain types of DNA

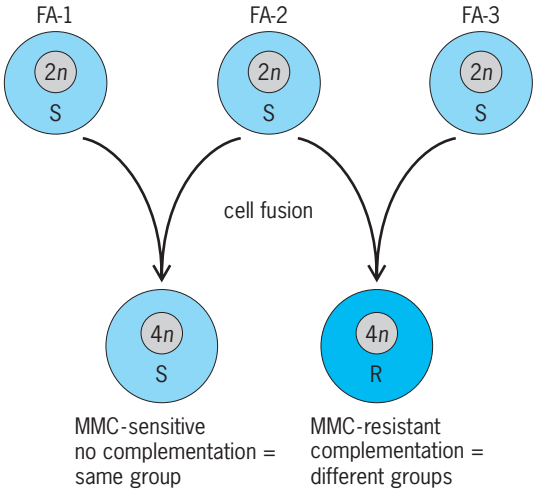

Fig. 1. Complementation analysis by cell fusion. FA, Fanconi anemia; MMC, mitomycin C. [*Reproduced with permission from Nature Reviews Genetics (http://www.nature.com/reviews): H. Joenje and K. J. Patel, The emerging genetic and molecular basis of Fanconi anaemia, Nature Rev. Genet., 2(6):446–457, © 2001 Macmillan Magazines Ltd.*]

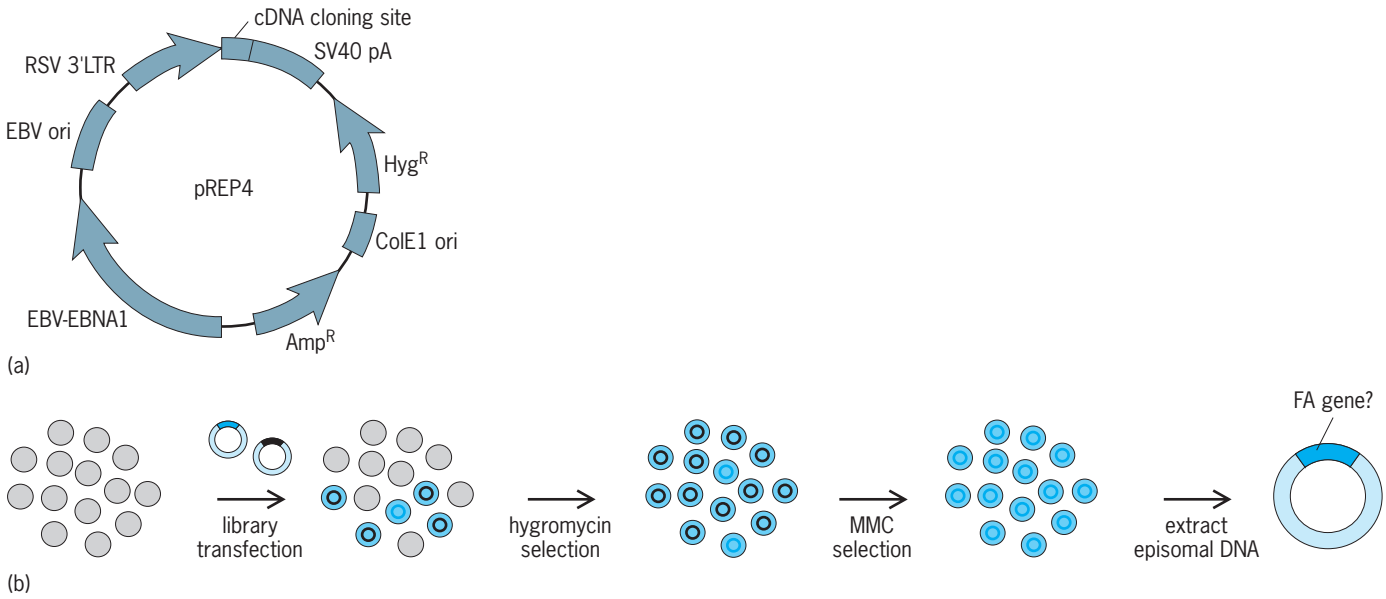

Fig. 2. Example of complementation cloning of an FA gene. (*a*) The vector containing various cDNAs. (*b*) The plasmid with the various inserts is introduced into lymphoblasts derived from a patient with FA (library transfection). Selection with hygromycin causes only cells that contain a plasmid to survive. Selection with MMC causes cells to survive that are complemented for the FA defect. The cDNA present in the plasmid extracted from surviving cells is likely the gene that is defective in the lymphoblasts. [*Reproduced with permission from Nature Reviews Genetics (http://www.nature.com/reviews): H. Joenje and K. J. Patel, The emerging genetic and molecular basis of Fanconi anaemia, Nature Rev. Genet., 2(6):446–457, © 2001 Macmillan Magazines Ltd.*]

damage. Each protein is important for the function of this complex, which explains why defects in any one of them cause Fanconi anemia.

Complementation cloning. A procedure known as complementation cloning has been instrumental in the identification of most of the genes underlying Fanconi anemia. This method makes use of an expression plasmid (a vector that expresses the gene) that is able to replicate in bacteria as well as in human cells (**Fig. 2**). When introduced into human B lymphoblasts (white blood cells), the plasmid is constructed in such a way that it cannot be integrated into the cellular DNA. A wide variety of human genes (cDNAs) are placed into multiple copies of this plasmid, which ensures expression of all these individual genes. After the plasmids with the inserted human genes are introduced into lymphoblasts from a patient with Fanconi anemia, the cells are selected for plasmid uptake by using a selection marker present in the plasmid. Further selection with a concentration of MMC that kills Fanconi anemia cells but not wild-type cells (normal functioning cells) allows survival of cells that have been complemented for the defect. These complemented cells are likely to harbor a plasmid that expresses the gene that is defective in the lymphoblasts. Extraction of the plasmid allows the identification of its insert, which is a candidate for representing the defective gene. Finding pathogenic mutations in this gene in patients belonging to the same complementation group finally should provide proof of identity for this disease gene.

Intragenic complementation. Sometimes combinations of different defects in a single gene may lead to partial or complete complementation (intragenic complementation). This might occur when the protein functions exclusively by forming a homodimer. In such cases, deficiencies in different regions of the individual proteins may be compensated in the dimeric structure, resulting in a preserved function. *See* GENE; GENE ACTION; HUMAN GENETICS.

M. Levitus; H. Joenje

Bibliography. M. Berneburg and A. R. Lehmann, Xeroderma pigmentosum and related disorders: Defects in DNA repair and transcription, *Adv. Genet.*, 43:71–102, 2001; D. Bootsma and J. H. J. Hoeijmakers, The genetic basis of xeroderma pigmentosum, *Ann. Genet.*, 34:143–150, 1991; S. J. Gould and D. Valle, Peroxisome biogenesis disorders: Genetics and cell biology, *Trends Genet.*, 16:340–345, 2000; H. Joenje and K. J. Patel, The emerging genetic and molecular basis of Fanconi anaemia, *Nature Rev. Genet.*, 2:446–457, 2001.

Complex numbers and complex variables

A natural and extremely useful extension of the familiar real numbers. They can be introduced formally as follows. Consider the two-dimensional real vector space consisting of all ordered pairs (a_1, a_2) of real numbers. Geometrically this space can be identified with the ordinary euclidean plane, viewing the real numbers a_1, a_2 as the coordinates of a point in the plane.

The addition of vectors is defined by $(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)$ and is just the usual addition of vectors by the parallelogram law (**Fig. 1**). The multiplication of a vector (a_1, a_2) by a real number c is defined by $c(a_1, a_2) = (ca_1, ca_2)$, and is just the uniform dilation of the plane by the factor c. It

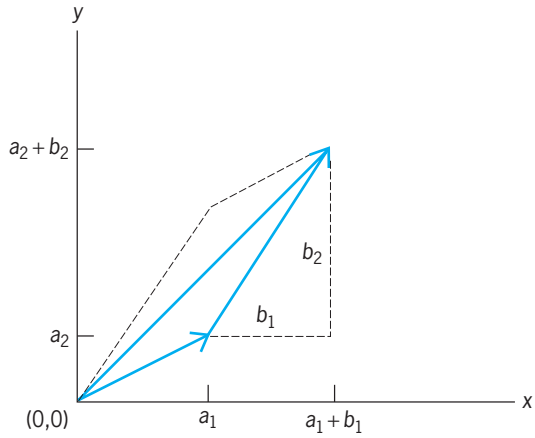

Fig. 1. Vector addition of two complex numbers by the parallelogram law.

may be asked whether it is possible to define a multiplication of one vector by another in such a manner that this multiplication is linear and satisfies the same formal rules as multiplication of real numbers. That is to say, it may be asked whether it is possible to associate to any two vectors (a_1, a_2) and (b_1, b_2) a product vector $(a_1, a_2) \cdot (b_1, b_2)$ in such a manner that:

1 The multiplication is linear in the sense that

$$(a_1, a_2) \cdot [(b_1, b_2) + (c_1, c_2)] = (a_1, a_2) \cdot (b_1, b_2) + (a_1, a_2) \cdot (c_1, c_2)$$

2 It is associative,

$$(a_1, a_2) \cdot [(b_1, b_2) \cdot (c_1, c_2)] = [(a_1, a_2) \cdot (b_1, b_2)] \cdot (c_1, c_2)$$

3 It is commutative,

$$(a_1, a_2) \cdot (b_1, b_2) = (b_1, b_2) \cdot (a_1, a_2)$$

4 There is an identity vector (e_1, e_2) such that

$$(e_1, e_2) \cdot (a_1, a_2) = (a_1, a_2)$$

for all vectors (a_1, a_2).

5 For each vector (a_1, a_2) other than (0, 0) there is a vector (b_1, b_2) for which

$$(b_1, b_2) \cdot (a_1, a_2) = (e_1, e_2)$$

There does exist such a multiplication, given by Eq. (1), and it is essentially unique in the sense that

$$(a_1, a_2) \cdot (b_1, b_2) = (a_1 b_1 - a_2 b_2, a_1 b_2 + a_2 b_1) \quad (1)$$

any multiplication satisfying all the desired properties can be reduced to Eq. (1) by a suitable choice of coordinates in the plane. The plane with the ordinary addition and scalar multiplication of vectors and with the vector multiplication in Eq. (1) is the complex number system. It is not possible to introduce a corresponding multiplication on vector spaces of higher dimensions, although, for example, if the commutativity condition is dropped, there is a multiplication satisfying the remaining conditions on the four-dimensional vector space (the quaternions).

Geometric interpretation. The multiplication in Eq. (1) can be interpreted most easily geometrically by introducing polar coordinates in the plane and writing $(a_1, a_2) = (r \cos \theta, r \sin \theta)$, where r is the length of the vector (a_1, a_2), or the modulus of the complex number (a_1, a_2), defined by $r^2 = a_1^2 + a_2^2$, and θ is the angle between the vector (a_1, a_2) and the first coordinate axis, or the argument of the complex number (a_1, a_2), defined by $\tan \theta = a_2/a_1$. It follows immediately from Eq. (1), by using the familiar addition formulas for the sine and cosine functions, that

$$(r_1 \cos \theta_1, r_1 \sin \theta_1) \cdot (r_2 \cos \theta_2, r_2 \sin \theta_2) = [r_1 r_2 \cos (\theta_1 + \theta_2), r_1 r_2 \sin (\theta_1 + \theta_2)]$$

Thus multiplication of complex numbers amounts to multiplying their moduli and adding their arguments (**Fig. 2**). *See* TRIGONOMETRY.

Real and imaginary parts. Introducing the basis vectors $l = (1, 0)$ and $i = (0, 1)$, any vector (x, y) can be written uniquely as the sum $(x, y) = xl + yi$. Since the multiplication given by Eq. (1) is linear, the product of any two vectors is determined by giving the multiplication table for the basis vectors l and i. It follows readily from Eq. (1) that $l \cdot l = l$, $l \cdot i = i \cdot l = i$, and $i \cdot i = -l$, so that in particular the vector l is the identity element for the multiplication of complex numbers. Furthermore, the addition and multiplication of vectors of the form xl for all real numbers x reduce to the ordinary addition and multiplication of real numbers. There is consequently a natural imbedding of the real numbers into the complex numbers by associating to any real number x the complex number xl. The complex number l can be identified with the ordinary real number 1, and reflecting this identification the notation can be simplified by writing a complex number $z = (x, y) = xl + yi$ merely as $z = x + iy$. The component x is called the real part of the complex number

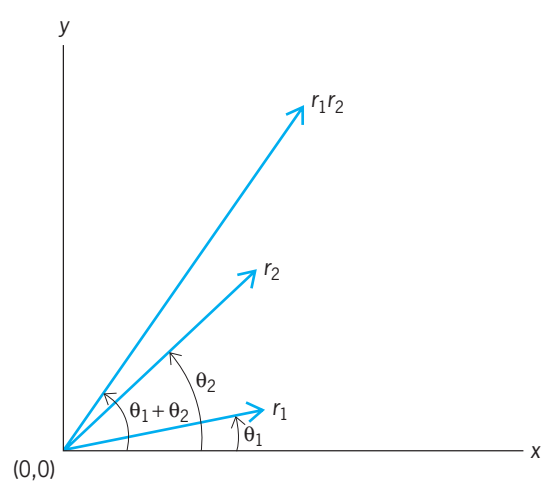

Fig. 2. Multiplication of complex numbers.

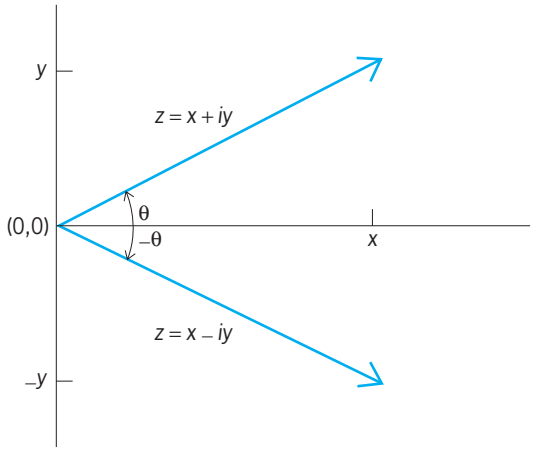

Fig. 3. Complex conjugate numbers.

z, and the component y is called the imaginary part. The modulus of the complex number z is denoted by $|z|$.

Complex conjugates. The vectors l and $-i$ are also a basis and satisfy exactly the same multiplication table as the vectors l and i, so the mapping which associates to any complex number $z = x + iy$ the complex number $\bar{z} = x - iy$ is a one-to-one mapping preserving the algebraic structure of the complex number system; thus $\overline{(z_1 + z_2)} = \bar{z}_1 + \bar{z}_2$ and $\overline{(z_1 \cdot z_2)} = \bar{z}_1 \cdot \bar{z}_2$. The complex number $\bar{z}$ is called the conjugate of z. It has the properties that $z = \bar{z}$ precisely when z is a real number, that $z \cdot \bar{z} = x^2 + y^2 = |z|^2$, and that in polar coordinates $\bar{z}$ is a complex number with the same modulus as z but with the negative argument (**Fig. 3**).

Polynomials. An ordinary real polynomial function $f(x) = x_n + a_1x^{n-1} + \cdots + a_n$ can be extended to a function of a complex variable z in the obvious manner, merely setting $f(z) = z^n + a_1z^{n-1} + \cdots + a_n$ and using the algebraic operations as defined for complex numbers. The advantage of this extension is that the fundamental theorem of algebra then holds: a nontrivial polynomial function over the complex numbers always has a root α, a complex value for which $f(\alpha) = 0$. The same result is true for polynomials with complex coefficients $a_1, \ldots, a_n$. A simple consequence of this result is that any polynomial function of degree n can be written as a product $f(z) = (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_n)$ where the complex numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ are precisely the roots of this polynomial repeated according to multiplicity. Thus passing from the real to the complex number system simplifies the analysis of polynomial functions and clarifies many of their properties. This was in a sense the original motivation for introducing the complex number system.

Power series. In a somewhat similar manner, passing from the real to the complex number system simplifies and clarifies the analysis of more general functions than polynomials. Many of the familiar functions treated in elementary calculus have convergent Taylor series expansions and so can be written for all real numbers x sufficiently near a real number a in the form of a convergent infinite series

$$f(x) = \sum_{n=0}^{\infty} a_n(x - a)^n$$

where $n!a_n = f^{(n)}(a)$. If this series converges when the real number x is replaced by nearby complex numbers z, there results a natural extension of the function f to a complex-valued function of the complex variable z.

The question of convergence can be handled rather easily. A series of the form

$$\sum_{n=0}^{\infty} a_n(z - a)^n$$

where a_n, a, z are complex numbers is called a complex power series centered at the point a. For any such series there is a value r, a nonnegative real number or ∞, such that the power series converges for all complex numbers z satisfying $|z - a| < r$ and diverges for all complex numbers z satisfying $|z - a| > r$; r is called the radius of convergence of the power series. If $r = 0$, the power series diverges for all complex numbers z other than $z = a$, while if $r = \infty$, the power series converges for all complex numbers. Otherwise, the series converges for all complex numbers lying within a circle of radius r centered at the point a and diverges for all complex numbers outside that circle. The series may or may not converge at complex numbers on the circle itself. The radius of convergence can be determined by Hadamard's formula: $1/r = \limsup_{n\to\infty}|a_n|^{1/n}$. Alternatively and generally more simply the radius of convergence satisfies

$$\liminf_{n\to\infty}|a_n|/|a_{n+1}| \le r \le \limsup_{n\to\infty}|a_n|/|a_{n+1}|$$

If r is the radius of convergence of a power series, then that series converges absolutely and uniformly on the set of all complex numbers z satisfying $|z - a| \le \rho$ whenever $\rho < r$.

Using these observations it is easy to see, for example, that the exponential function can be extended to a complex-valued function of the complex variable z defined for any z by

$$e^z = \sum_{n=0}^{\infty} z^n/n!$$

Among the trigonometric functions the sine and cosine functions also can be defined for arbitrary complex numbers z by their Taylor series. Moreover, these functions satisfy the identity $e^{iz} = \cos z + i \sin z$ for all complex numbers z, an interesting relation among the elementary functions only apparent when passing to complex numbers. *See* SERIES.

Analytic functions. A complex-valued function f of a complex variable defined on some region Ω in the complex plane is analytic or holomorphic in Ω if for each point a in Ω the values of the function at all points z sufficiently near a are given by a power series expansion

$$f(z) = \sum_{n=0}^{\infty} a_n (z - a)^n$$

This power series then must be the Taylor series expansion of the function f at the point a and must converge in at least the largest circle centered at a and contained in Ω. An analytic function in a region Ω is necessarily a continuous function in Ω moreover, as a function of the two real variables (x, y), an analytic function has partial derivatives of all orders.

Characterization by differentiability. There are a number of alternative characterizations of analytic functions, all of which are useful. The most common and perhaps the most surprising characterization, which is often taken as the definition of an analytic function, is that a function f is analytic in Ω precisely when the limit given by (2) exists at all points a in Ω. This limit is

$$\lim_{z \to a} \frac{f(z) - f(a)}{z - a} \tag{2}$$

the complex analog of the ordinary derivative and is therefore called the complex derivative and written $f'(a)$. If the power series

$$f(z) = \sum_{n=0}^{\infty} a_n (z - a)^n$$

converges in a disc Δ centered at a, then

$$f'(z) = \sum_{n=0}^{\infty} n a_n (z - a)^{n-1}$$

and this series also converges in the disc Δ. *See* CALCULUS.

Cauchy-Riemann equation. If f is analytic near a point a, then the complex derivative (2) can be calculated by letting z approach a through a sequence of points having the same imaginary value as a, hence $f'(a) = \partial f/\partial x$. Alternatively, the complex derivative can be calculated by letting z approach a through a sequence of points having the same real value as a, hence $f'(a) = -i\, \partial f/\partial y$. Upon comparing these two results it follows that an analytic function f must satisfy the Cauchy-Riemann equation

$$\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} = 0$$

which is sometimes written $\partial f/\partial \bar{z} = 0$. Conversely if f is a complex-valued function defined in a region Ω of the complex plane, and if f has continuous first partial derivatives with respect to the real coordinates x and y, and if they satisfy the Cauchy-Riemann equation at all points of Ω, then f is analytic in Ω. It is not enough merely to require the existence of the first partial derivatives, and so this is not such a clean characterization of analytic functions as the preceding one, but it is nonetheless very useful. Weaker regularity conditions than the continuity of the first partial derivatives are possible. For instance, if the function f has first partial derivatives in the sense of distributions and if they satisfy the Cauchy-Riemann equation, then f is analytic.

Relation to harmonic functions. Setting $f(x, y) = u(x, y) + iv(x, y)$, where u and v are real-valued functions, the Cauchy-Riemann equation can be written equivalently as the pair of equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial u} = -\frac{\partial v}{\partial x}$$

It follows immediately from this that $\partial^2 u/\partial x^2 + \partial^2 v/\partial y^2 = 0$ and similarly for v, so that the real and imaginary parts of an analytic function are harmonic functions. Conversely if u is harmonic in a simply connected region Ω of the complex plane, the pair of real Cauchy-Riemann equations determines a harmonic function v such that $u + iv$ is analytic in Ω, and v is unique up to an additive constant; if Ω is not simply connected, the function v may not be single-valued. *See* DIFFERENTIAL EQUATION; LAPLACE'S DIFFERENTIAL EQUATION.

Morera's theorem. If f is an analytic function in a region Ω of the complex plane, then it follows from the Cauchy-Riemann equation that, using the terminology and notation of differential forms, the differential form $f dz = f dx + i f dy$ is closed, since

$$d(f dx + i f dy) = \left(i \frac{\partial f}{\partial x} - \frac{\partial f}{\partial y} \right) dx \wedge dy = 0$$

Therefore

$$\int_\gamma f dz = 0$$

for any closed path γ homologous to zero in Ω. Conversely Morera's theorem asserts that if f is a continuous complex-valued function in Ω and if

$$\int_\gamma f dz = 0$$

for all closed paths γ homologous to zero in Ω, then f is necessarily analytic in Ω. From this characterization of analyticity it follows immediately that the limit of a uniformly convergent sequence of analytic functions is also analytic.

Conformal mapping. An analytic function f in a region Ω of the complex plane can be viewed as a mapping from the plane region Ω into the complex plane. If $f(z) = u(z) + iv(z)$, this mapping is described in real terms by the pair of functions $u(x, y)$, $v(x, y)$. It follows readily from the Cauchy-Riemann equation that the jacobian determinant of this real mapping is

$$\det \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = u_x^2 + v_x^2 = |f'(z)|^2 \tag{3}$$

Since the jacobian determinant is thus always nonnegative, the mapping always preserves orientation. Moreover, the mapping is locally one-to-one whenever $f'(z) \neq 0$, and its inverse is readily seen also to satisfy the Cauchy-Riemann equation and therefore to be analytic. By analyzing further the condition imposed by the Cauchy-Riemann equation it follows that if $f'(a) \neq 0$, if γ_1 and γ_2 are two differentiable paths through the point a, and if the tangent vectors to γ_1 and γ_2 at the point a are at an angle θ apart, then the tangent vectors to the paths $f(\gamma_1)$ and $f(\gamma_2)$ at the point $f(a)$ are also at an angle θ apart. The mapping

Fig. 4. A conformal mapping that preserves angles at *a*.

f thus preserves angles at a, and is therefore said to be conformal at a (**Fig. 4**). Conversely any continuously differentiable conformal mapping is an analytic function with a nonzero complex derivative, thus providing a very geometric characterization of analytic functions. *See* CONFORMAL MAPPING; PARTIAL DIFFERENTIATION.

For completeness, something should be said about the mapping property of an analytic function at a point at which the complex derivative is zero. The analytic function $f(z) = z^n$, for example, has a zero derivative at $z = 0$ whenever $n > 1$. The mapping is not conformal but increases angles at the origin by a factor of n (**Fig. 5**) and so can be viewed as a

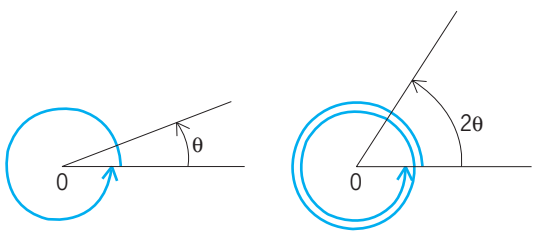

Fig. 5. The mapping z^2 at the origin.

mapping which wraps each circle about the origin in its domain n times about the origin in its range. Now if f is any analytic function at the origin with $f(0) = f'(0) = 0$, then this function can be written in the form $f(z) = g(z)^n$ for some integer $n > 1$, where g is analytic at the origin, $g(0) = 0$ but $g'(0) \neq 0$. Therefore the change of variable $w = g(z)$ reduces the function f to the form $f(w) = w^n$ just considered. It is clear from this that an analytic function takes open sets to open sets even when its derivative is zero and that any analytic function satisfies the maximum modulus theorem: if f is analytic at a and $|f(a)| \geq |f(z)|$ for all points z in some neighborhood of a, then f must actually be a constant. This seemingly simple property has a considerable range of applications.

Cauchy integral formula. An integral representation formula due to A. Cauchy plays a fundamental role in complex analysis and is traditionally the principal tool used in proving the equivalence of some of the alternative characterizations of analytic functions given above. The formula asserts that if γ is the smooth boundary curve of a finite region Ω in the complex plane (**Fig. 6**) and if f is analytic in an open neighborhood of the closed region $\Omega \cup \gamma$ (the union of Ω and γ), then

$$\frac{1}{2\pi i}\int_\gamma \frac{f(\zeta)d\zeta}{\zeta - z} = \begin{cases} f(z) & \text{whenever } z \text{ is a member of } \Omega \\ 0 & \text{whenever } z \text{ is not a member of } \Omega \cup \gamma \end{cases} \tag{4}$$

This shows, for example, that the values of the analytic function f in Ω are completely determined by the values of that function on the boundary of Ω. Equation (4) can be differentiated under the integral sign to yield the companion formula (5)

$$f^{(n)}(z) = \frac{n!}{2\pi i}\int_\gamma \frac{f(\zeta)\,d\zeta}{(\zeta - z)^{n+1}} \tag{5}$$

for the derivatives of the function f at any point z in Ω. In particular, if Ω is a disk of radius r centered at a point a and if $|f(z)| \leq M$ on the boundary γ of Ω, then (5) implies that $|f^n(a)| \leq Mr^{-n}n!$. Thus the derivatives of an analytic function cannot grow faster than $r^{-n}n!$ for some r as n tends to infinity. An immediate consequence of this estimate is Liouville's theorem, the assertion that the only functions analytic and bounded in the entire plane are the constant functions.

Relation to Green's integral theorem. The Cauchy integral formula (4) can be considered as a special case of a more general integral representation formula, following fairly directly from Green's integral theorem. If Ω and γ are as before and if f is any continuously differentiable complex-valued function of the real variables (x, y) in an open neighborhood of the closed region $\Omega \cup \gamma$, then

$$f(z) = \frac{1}{2\pi i}\int_\gamma f(\zeta)\frac{d\zeta}{\zeta - z} + \frac{1}{2\pi i}\int\int_\gamma \frac{\partial f(\zeta)}{\partial \bar{\zeta}}\frac{d\zeta \wedge d\bar{\zeta}}{\zeta - z} \tag{6}$$

is valid for all z in Ω. Since analytic functions are characterized by the Cauchy-Riemann equation $\partial f/\partial \bar{z} = 0$, it is evident that (6) reduces to (4) for analytic functions. The more general formula (6) is useful in investigating solutions of the partial differential equation $\partial f/\partial \bar{z} = g$ for a given function g. This equation has many applications in complex analysis.

Relation to Poisson integral formula. The Cauchy integral formula (4) is also closely related to the Poisson integral formula for harmonic functions, which can be

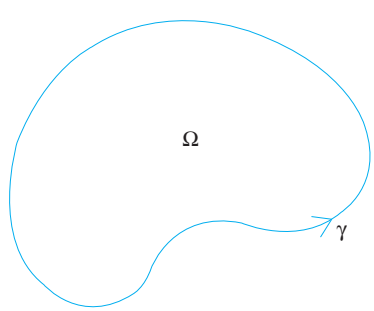

Fig. 6. The geometric situation in which the Cauchy integral formula holds.

derived from Eq. (4). The Poisson formula is useful in solving boundary-value problems for harmonic functions. For any continuous function on the smooth boundary γ of a bounded region Ω, there exists a harmonic function u in Ω with the given boundary values. If Ω is simply connected, the function u is the real part of an analytic function in Ω, so there thus exists a holomorphic function in Ω with an arbitrarily specified real part on γ. The imaginary part of this analytic function is determined uniquely up to an additive constant by the real part, so there cannot exist an analytic function in Ω with arbitrarily specified values on the boundary curve γ. On the other hand, if g is any sufficiently smooth function on γ, the integral

$$\int_\gamma \frac{g(\zeta)}{\zeta - z} d\zeta$$

represents an analytic function $f_i(z)$ at all points z in Ω and another analytic function $f_0(z)$ at all points outside $\Omega \cup \gamma$. These functions have well-defined boundary values on γ, and $f_i(z) - f_0(z) = 2\pi i g(z)$ whenever z is on γ; the function $f_0(z)$ vanishes only when g is the boundary value on γ of an analytic function in Ω.

Isolated singularities. If a function f is analytic at all points of a disk Δ centered at a, except perhaps at the point a itself, a is called an isolated singularity of the function f. There are three mutually exclusive possibilities for the nature of such a singularity.

1. It may happen that $|f(z)|$ remains bounded near the point a(14). In that case the function can be extended to an analytic function even at the point a by suitably choosing the value of $f(a)$; the point a is then called a removable singularity.

2. It may happen that $\lim_{z\to a}|f(z)| = \infty$. In that case a is a removable singularity for the function $g(z) = 1/f(z)$, so that $f(z) = 1/g(z)$ where $g(z)$ is an analytic function vanishing at the point a; and if $g(z) = (z - a)^k g_0(z)$, where $g_0(z)$ is analytic and $g_0(a) \neq 0$, then $f(z) = (z - a)^{-k} f_0(z)$ where $k > 0$ and $f_0(z)$ is analytic and $f_0(a) \neq 0$. The function $f(z)$ is said to have a pole of order k at the point a. The function $f_0(z)$ has a Taylor expansion at the point a as usual, so the function $f(z)$ has an expansion given by Eq. (7).

$$f(z) = a_{-k}(z-a)^{-k} + \cdots + a_{-1}(z-a)^{-1} + a_0 + a_1(z-a) + a_2(z-a)^2 + \cdots \quad (7)$$

The portion

$$\sum_{n=-k}^{-1} a_n (z-a)^n$$

of this expansion is called the principal part or singular part of the function $f(z)$ at the point a, while the remainder of the series represents an analytic function near a.

3. If neither of the two preceding cases arises, the function $f(z)$ is said to have an essential singularity at the point a. A function having an essential singularity admits a Laurent series expansion near the point a of the form

$$f(z) = \sum_{n=-\infty}^{+\infty} a_n (z-a)^n \quad (8)$$

in which infinitely many terms involving negative powers of the variable $(z - a)$ appear. Singularities of this type can occur and do so frequently. For example, whenever $f(z)$ has a pole at a point a, then $e^{f(z)}$ has an essential singularity at the point a. The behavior of a function near an essential singularity is rather complicated. A theorem of K. Weierstrass asserts that in any neighborhood of an essential singularity the values taken by the function come arbitrarily near any complex value. A much deeper theorem of E. Picard asserts that in any neighborhood of an isolated singularity a function actually takes all complex values with at most one possible exception.

Global properties. The zeros of an analytic function are isolated, in the sense that if f is analytic near a point a and $f(a) = 0$, then f has no other zeros in some neighborhood of a. That leads quite easily to the identity theorem: if $f(z)$ and $g(z)$ are analytic in a connected-region Ω and if $f(a_n) = g(a_n)$ for an infinite sequence of points a_n having a limit point inside Ω, then $f(z) = g(z)$ at all points z in Ω. On the other hand, if a_n is any sequence of points in Ω having no limit point inside Ω, there are analytic functions in Ω having zeros precisely at the points a_n. This result is particularly easy in case that Ω is the entire complex plane. Indeed, in that case any analytic function having zeros at the points a_n can be written as a product given

$$f(z) = z^m e^{b(z)} \prod_n \left(1 - \frac{Dz}{a_n}\right) e^{P_n(z)} \quad (9)$$

where $b(z)$ is an analytic function in the entire plane and $P_n(z)$ are suitably chosen polynomials ensuring that the product converges when there are infinitely many points a_n. There is a good deal of quite detailed information relating the growth of the function $f(z)$, the distribution of the zeros a_n, and the canonical product formula (9). This area of investigation is called the theory of entire functions. The infinite product expansions of even the elementary functions are quite interesting. For example, the expansion of $\sin \pi z$ is given by

$$\sin \pi z = \pi z \prod_{n=1} \left(1 - \frac{z^2}{n^2}\right) \quad (10)$$

Meromorphic functions. A function analytic in a region Ω of the complex plane, except possibly for poles at some points of Ω, is called a meromorphic function in Ω. The points at which the function has poles can of course have no limit point inside Ω. Near any pole a function can be written as the quotient of two analytic functions, as already noted. There is an analogous global result: any meromorphic function in Ω can be written as the quotient of two functions each analytic in Ω. If $a_1, a_2, \ldots$ is any sequence of points in Ω having no limit point inside Ω and if

$f_1(z), f_2(z), \ldots$ are principal parts of poles at these points, so that

$$f_v = \sum_{n=-kv}^{-1} a_{v,n}(z - a_v)^n$$

There exists a meromorphic function $f(z)$ in Ω having poles with the specified principal parts at the specified points and no other singularities. This is the analog of the familiar partial fraction expansion for quotients of polynomials. The resulting expansions of even the elementary functions are quite interesting, as for example the expansion

$$\pi \cot \pi z = \frac{1}{z} + \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2} \tag{11}$$

Rational functions. Those functions that have at most finitely many zeros and poles in the entire complex plane, and that either have a well-defined limit or tend to infinity as the variable approaches infinity, are rational functions; that is, they are quotients of polynomial functions. These functions extend naturally to the Riemann sphere, the set that arises by adjoining to the complex plane a single point at infinity. The change of variable $\zeta = 1/z$ transforms the complement of a large circle in the plane of the variable z to the interior of a small circle in the plane of the variable ζ; the point at infinity in the variable z corresponds to the origin in the variable ζ, and that makes it possible to speak of the behavior of a rational function at the point at infinity. The rational functions are precisely those functions that are meromorphic on the entire Riemann sphere, including the point at infinity. Although they are in some senses the simplest of analytic functions, their study leads to extremely complicated problems. Any such function can be viewed as a mapping of the Riemann sphere to itself; and this mapping can be iterated, beginning with the function $f(z)$ and then passing to the iterates $f^2(z) = f(f(z))$, and so on. The behavior of the iterates can be very complicated even for the simplest such functions, second-degree polynomials. The sphere can be decomposed into two disjoint sets: the Fatou set, an open subset of the Riemann sphere on which the iterates form an equicontinuous family of functions and generally converge; and the Julia set, usually a very complicated closed set, even a fractal set, on which the iterates form a family exhibiting the phenomenon of chaos (**Fig. 7**). *See* CHAOS; FRACTALS.

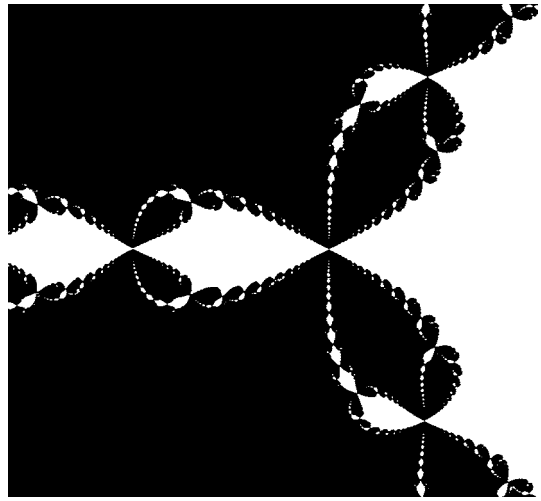

Fig. 7. $z \mapsto z^3 - 1$. **Fractal image.** (***Reprinted with permission from H.-O. Peitgen and P. H. Richter, The Beauty of Fractals, Springer-Verlag, New York, 1986***)

Analytic functions of several variables. A complex-valued function depending on several complex variables $f(z_1, \ldots, z_n)$ is called analytic at a point $(a_1, \ldots, a_n)$ if the values of that function at all points z_j sufficiently near a_j are given by a multiple power series expansion

$$f(z_1, \ldots, z_n) = \sum_{i_1=0}^{\infty} \cdots \sum_{i_n=0}^{\infty} a_{i_1 \ldots i_n}(z - a_1)^{i_1} \cdots (z - a_n)^{i_n} \tag{12}$$

If this series converges at a point $(z^0{}_1, \ldots, z^0{}_n)$ with $|z^0{}_1 - a_i| = r_i > 0$, then it is absolutely and uniformly convergent on the set of all points $(z_1, \ldots, z_n)$ for which $|z_i - a_i| \leq \rho_i$ whenever ρ_i are any positive constants such that $\rho_i < r_i$. This series thus converges to the same value for any ordering of the terms. For example, it can be rewritten as the power series

$$f(z_1, \ldots, z_n) = \sum_{i_n=0}^{\infty} a_{i_n}\left(z_1, \ldots, z_{n-1}\right) \cdot (z_n - a_n)^{i_n} \tag{13}$$

in the variable z_n with coefficients which are analytic functions of the variables $z_1, \ldots, z_{n-1}$. Hence holding $z_1, \ldots, z_{n-1}$ constant, the function $f(z_1, \ldots, z_n)$ is analytic in the variable z_n alone, or more generally the function $f(z_1, \ldots, z_n)$ is analytic in each variable separately. Conversely a surprisingly nontrivial theorem of F. Hartogs asserts that a function $f(z_i, \ldots, z_n)$ that is analytic in each variable separately at all points of a region Ω is an analytic function of all n variables $z_1, \ldots, z_n$ in Ω.

It follows from this result that analytic functions of several variables can be characterized in many of the same ways as are analytic functions of one variable, since it suffices to show that a function is analytic in each variable separately. Thus, if a function of several complex variables has a complex derivative in each variable separately, or satisfies the Cauchy-Riemann equations in each variable separately, then it is analytic. It is convenient to extend the Cauchy-Riemann characterization by associating to any function f of n complex variables the differential form

$$\bar{\partial} f = \sum_{i=1}^{n} \frac{\partial f}{\partial \bar{z}_i} d\bar{z}_i$$

A function of n complex variables that is continuously differentiable in the underlying real variables is analytic if $\bar{\partial} f = 0$. The analysis of this differential operator plays a significant role in a great deal of research on properties of analytic functions of several variables.

Extension of results from one variable. Some results extend quite directly from one to n complex variables: the uniform limit of analytic functions is analytic; the maximum modulus theorem holds; and there is an extension of the Cauchy integral formula, asserting that if $f(z_1, \ldots, z_n)$ is analytic whenever z_j is a member of $\Omega_j \cup \gamma_j$ where Ω_j is a plane domain with smooth boundary γ_j, then

$$f(z_1, \ldots, z_n) = \left(\frac{1}{2\pi i}\right)^n \int_{\zeta_1 \in \gamma_1} \cdots \int_{\zeta_n \in \gamma_n} \frac{f(\zeta_1, \ldots, \zeta_n)}{(\zeta_1 - z_1) \cdots (\zeta_n - z_n)} d\zeta_1 \cdots d\zeta_n \quad (14)$$

holds whenever z_j is a member of Ω_j for all indices j. However, for the most part, the theory of analytic functions of several complex variables is far from being merely an extension of the standard results of the classical theory of functions of a single complex variable, by considering one variable at a time. Indeed, the theory of functions of several complex variables has quite a distinctive character with a considerable variety of results which perhaps are not the expected generalizations of the classical results but which shed new light on the familiar theory of analytic functions of a single variable.

Integral formulas. In addition to the direct extension of the classical Cauchy integral formula from one to several variables in (14), there are a variety of other integral representation formulas that reflect the higher-dimensional situation and that have significant applications. Some of these formulas, such as those due to André Weil and Stefan Bergman, follow the extended Cauchy integral formula (14) in that they involve integration over only part of the boundary of the domain. Others, such as those of Salomon Bochner and E. Martinelli, involve an integration over the full boundary, and are particularly useful for examining the behavior of functions on smoothly bounded domains. Yet others, such as those due to W. Koppelman and I. Lieb, extend the analysis to functions that are not complex analytic, by providing explicit solutions f of the system of partial differential equations $\bar{\partial} f = \phi$, where ϕ is a differential form. These solutions also are important in the analysis of the boundary values of analytic functions of several variables. The integral formulas themselves are rather complicated.

Analytic continuation. An examination of the extended Cauchy integral formula (14) indicates some of the differences between the cases $n = 1$ and $n > 1$. When $n = 1$, either formula (14) or (4) expresses the values of an analytic function inside a domain Ω in terms of the values of the analytic function on the boundary of Ω, whereas when $n > 1$, formula (14) expresses the values of an analytic function inside a domain $\Omega_1 \times \cdots \times \Omega_n$ in terms of the values of the analytic function on a very small piece of the boundary. For example, when $n = 2$ the domain $\Omega_1 \times \Omega_2$ can be viewed as an open set of the four-dimensional euclidean space, its boundary is the three-dimensional subset $(\Omega_1 \times \gamma_2) \cup (\gamma_1 \times \Omega_2)$, but the integration in (14) is extended across the two-dimensional part $\gamma_1 \times \gamma_2$ of the boundary. This difference is not superficial; it reflects the greater possibilities for analytic continuation of functions of $n > 1$ variables than of functions of a single variable.

For any domain Ω in the complex plane there are functions which are analytic in that domain but which cannot be extended as analytic functions across any part of the boundary of Ω. For example, there exist nontrivial analytic functions vanishing at any infinite sequence of points $a_1, a_2, \ldots$ of Ω having no limit point inside Ω, and if every boundary point of Ω is a limit of a subsequence of these points $a_1, a_2, \ldots$ then such an analytic function cannot be continued across any boundary point. However, there are domains Ω in the space of $n > 1$ complex variables such that any function analytic in Ω extends to an analytic function in a properly larger domain. For example, it was shown by Hartogs that if Ω is an open neighborhood of the boundary of a compact subset K in the space $n > 1$ complex variables, then every function analytic in Ω extends to a function analytic in the union of K and Ω (**Fig. 8**).

Thus in the study of functions of $n > 1$ complex variables there arise the problems of characterizing the natural domains of existence of analytic functions (called domains of holomorphy) and of determining for those domains which are not domains of holomorphy the largest domains to which all analytic functions can be extended (called the envelopes of holomorphy). Both problems have been extensively studied, the first one particularly so. Perhaps the most surprising result is that the domains of holomorphy can be characterized purely locally: a domain Ω in the space of n complex variables is a domain of holomorphy if and only if each point on the boundary of Ω has an open neighborhood $\cup$ such that the intersection of $\cup$ and Ω is a domain of holomorphy. The determination of the envelope of holomorphy of a domain is generally rather difficult, the situation being complicated by the fact that the envelope of holomorphy may not be realizable as a domain in the space of n complex variables.

Zeros and singularities. Another and not entirely unrelated difference between functions of $n > 1$ and functions of one complex variable is that in the case $n > 1$ the zeros and singularities of an analytic function are never isolated. The set of zeros of an analytic function of $n > 1$ variables is a set of topological dimension $2n - 2$ and is in most places locally euclidean but can have very complicated singularities. Strange spaces, such as the exotic seven spheres,

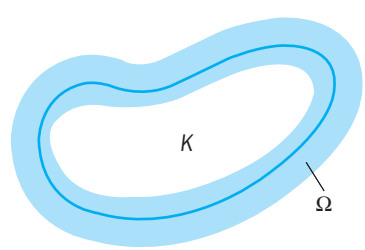

Fig. 8. Analytic continuation in $n > 1$ dimensions according to the theorem of Hartogs.

manifolds which are topologically but not differentiably equivalent to the ordinary seven-dimensional sphere, arise in this context. Even the purely local study of functions of several complex variables thus presents a considerable challenge. *See* TOPOLOGY.

Applications. The general theory of analytic functions of several complex variables has a wide range of applications in other parts of mathematics. It is a standard tool in the examination of function algebras and of special spaces of functions, the subject of harmonic analysis. The techniques also arise in geometric problems in various parts of differential geometry. *See* DIFFERENTIAL GEOMETRY. Robert C. Gunning

Bibliography. L. V. Ahlfors, *Complex Analysis*, 3d ed., McGraw-Hill, 1979; A. F. Beardon, *Iteration of Rational Functions*, Springer, 1991; R. C. Gunning, *Introduction to Holomorphic Functions of Several Variables*, 3 vols., Wadsworth, 1990; L. Hörmander, *An Introduction to Complex Analysis in Several Variables*, 3d ed., Elsevier Science, 1990; S. G. Krantz, *Function Theory of Several Complex Variables*, 2d ed., Brooks Cole Publishing, 1992; J. E. Marsden and M. J. Hoffman, *Basic Complex Analysis*, Freeman, 1999; R. M. Range, *Holomorphic Functions and Integral Representations in Several Complex Variables*, Springer, 1986; W. Rudin, *Real and Complex Analysis*, 3d ed., McGraw-Hill, 1987.

Complexity theory

The theory of complex systems. Complexity theory arises from the need to understand the richness in structure and behavior often seen in large systems. The property that distinguishes complex systems from systems that are large but simple is the emergence of global features from local interactions, as captured in the popular saying "the whole is greater than the sum of its parts." For example, a flock of birds emerges when individual birds coordinate their behavior with each other.

Complexity often appears as unpredictability, or as a tendency for a system to undergo sudden, rapid change. Typically, it arises within systems that have strong interrelationships between components. For instance, the coordinated action of millions of individual cells, all tightly bound together, produces the action of muscles in the body. Epidemics emerge from the spread of infection from individual to individual.

Complexity theory has its roots in many fields of research, including biology, computer science, ecology, engineering control systems, nuclear physics, and operations research. In physics, for instance, there is a need to understand properties that emerge within large ensembles of atoms and other particles. General systems theory, developed by the biologist Ludwig von Bertalanffy in the 1950s, identified the role of internal interactions and processes in self-organization. In the 1960s, W. Ross Ashby and Norbert Wiener developed cybernetics, which concerned communications and feedback in the control of systems. Complexity theory came into its own during the late 1980s, as increasing computing power made large-scale simulation studies practical. The formation of specialist research centers, such as the Santa Fe Institute, did much to promote the new science during the 1990s. *See* CYBERNETICS; ECOLOGY; NUCLEAR STRUCTURE; OPERATIONS RESEARCH.

Several fields of research have become closely linked to complexity theory. The link to graph theory is assured by the network model of complexity (discussed below). The relevance of chaos theory arises because complexity in dynamic systems often leads to nonlinear and chaotic behavior. Another related field is artificial life (Alife), which concerns lifelike features found in artificial systems. By addressing complexity in living systems, this field has influenced the development of theory in evolution, ecology, and sociobiology. Evolutionary computing concerns computational processes that address complex problems by mimicking species evolution. The best-known example is the genetic algorithm, introduced by John Holland in 1975. This solves complex optimization problems by treating key parameters as genes within a population of potential solutions. Complex adaptive systems (CAS) are systems, such as brains or societies, composed of interacting elements that adapt or change in response to experience. *See* CHAOS; EVOLUTIONARY COMPUTATION; GENETIC ALGORITHMS; GRAPH THEORY.

Network model. In most cases, complexity can be understood in terms of networks. This model treats systems as sets of objects ("nodes") with connections ("edges") linking them. For instance, within a flock of birds the nodes would be birds and their interactions would be the edges. In a nuclear chain reaction, the nodes would be atoms and the edges would be moving neutrons. The network model also applies to processes and behavior. In a game of chess, for instance, each arrangement of pieces on the board is a node in a network of board arrangements, and moves by the players provide the edges that link different arrangements. In many games, for instance, a general strategy is to restrict your opponent's moves while maximizing the number of possible moves that you can make.

Strictly speaking, a set of nodes linked by edges forms a graph (for example, people who are neighbors of each other). A directed graph is a graph in which each edge has direction (for example, descendants in a family), and a network is a graph in which the nodes or edges have attributes associated with them (for example, people have names).

In any network, some features arise from the nature of the objects and their relationships, and some emerge from the underlying network. A network is connected if there is a path (a sequence of edges) leading from any one node to any other. For example, the edges A-B, B-C, C-D form a path between nodes A and D. Attributes that are often used to express the structure of different networks include:

Connectivity (the number of edges that must be removed to break the network apart).

Edge density (the number of edges as a fraction of the total number possible).

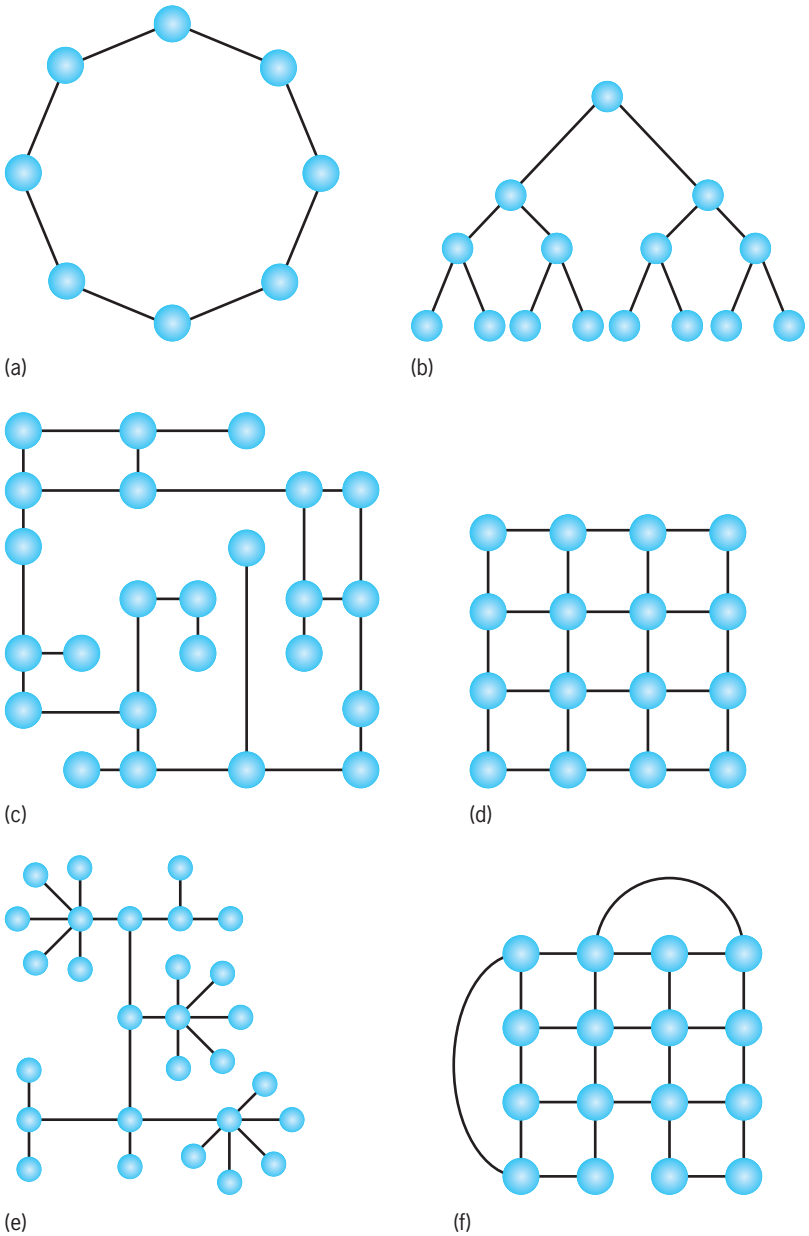

Fig. 1. Structure of some common kinds of networks. (*a*) Cycle. (*b*) Tree. (*c*) Random. (*d*) Regular. (*e*) Scale-free. (*f*) Small world.

Clustering (the density of links between sets of nodes linked to a given node).

Diameter (the greatest separation between pairs of nodes, where separation is taken to mean the shortest path between a given pair of nodes).

Several kinds of network structure are common because they confer important properties:

1. Cycles or circuits (**Fig. 1***a*) are regular networks in which a single path joins every node. The sequence of edges A-B, B-C, C-D, D-A forms a cycle linking the nodes A, B, C, and D.

2. Trees (Fig. 1*b*) are connected networks that contain no cycles. Hierarchies are trees with a root node, a node that all edges lead away from (for example, a common ancestor in a family tree). In computing, objects can form hierarchies in two ways: whole-part hierarchies (for example, in a whole book, the parts are chapters), and gen-spec hierarchies (for example, a novel is a special kind of book, and a book is a special kind of publication).

3. Random networks (Fig. 1*c*) are networks in which edges are distributed randomly between nodes.

4. Regular networks (Fig. 1*d*) are connected networks in which the edges form a pattern (for example, a lattice or grid).

5. Scale-free networks (Fig. 1*e*) form when new nodes in a growing network attach preferentially to existing nodes with highest connectivity. The resulting distribution of links per node obeys an inverse power law. That is, some nodes have many connections, but most have few. Web sites on the Internet, connected by hypertext links, provide a good example.

6. Small worlds (Fig. 1*f*) are connected graphs that contain both local clusters and long-range interactions. They fall between random networks at one extreme and regular networks at the other. Social networks often take this form.

Phase transitions. Complex systems have a tendency to undergo sudden change. Examples include crystallization, water freezing, and epidemics. In all cases, these phase changes occur at a fixed value (the critical point) of some order parameter (for example, water freezes at a temperature of 0°C). Phase transitions can usually be understood as connectivity avalanches in an underlying network (**Fig. 2**). If a network is formed by adding edges at random to a set of N nodes, then a phase change (connectivity avalanche) occurs when the number of edges reaches the critical density $N/2$ (**Fig. 3**): the nodes become absorbed into a unique giant component. This connectivity avalanche translates into many physical processes. In crystallization, for instance, it involves small crystal seeds merging to form large structures. In all epidemic phenomena, the process requires a certain critical level of interaction to sustain itself. Nuclear fission, for instance, starts when the fuel reaches a critical mass. *See* CRITICAL MASS; CRITICAL PHENOMENA; CRYSTALLIZATION; EPIDEMIC; PHASE TRANSITIONS.

Automata whose state spaces lie close to the phase change in connectivity (the so-called edge of chaos) often exhibit interesting behavior, leading to speculation that only these automata are capable of universal computation and the ability to evolve.

Self-organized criticality occurs when internal processes maintain a system in a critical state (for example, a sand pile). Characteristically, the frequency (f) of events in such a system (for example, the sand pile collapses) follows an inverse power law $f(s) = s^{-\tau}$, where s is the size of the event and τ is a self-similarity exponent.

Encapsulation. Encapsulation involves grouping nodes in a network into modules that are highly connected internally, but with limited connections to the rest of the network. In both natural and artificial systems, encapsulation is the most common way of coping with complexity. It appears in many different guises. In human activity, large tasks are usually

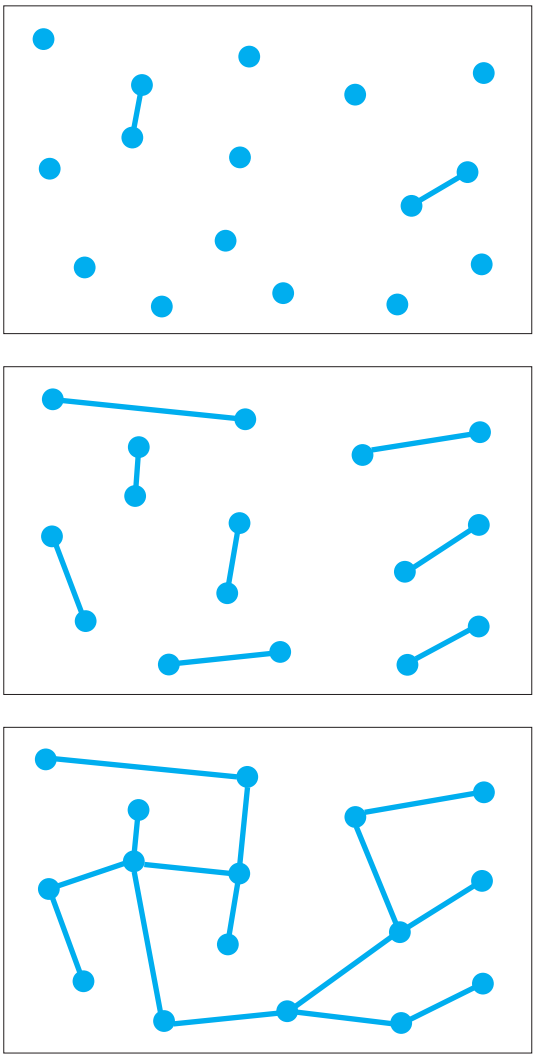

Fig. 2. Formation of a unique giant component as edges are added to a random network.

broken down into managable parts: Large corporations are typically divided into work units such as divisions, departments, and sections. Likewise, filing systems are divided into cabinets, drawers, files, and folders.

Modular design simplifies construction. Computer software typically consists of self-contained, reusable objects and modules. Living trees grow by the repeated addition of modules, such as buds, leaves, and branches. Engineering systems are often designed around modular components. A motor car, for instance, has electrical, fuel, and steering systems, and so on.

Encapsulation reduces complexity by eliminating long-range connections. In a human organization, for instance, specialization reduces the scope of problems that each person needs to deal with. On the other hand, modular designs can be brittle, often failing when long-range connections occur, such as by placing a library book back on the wrong shelf.

Modular structures often form into hierarchies of organization. A book, for instance, is a hierarchy of modules consisting of chapters, sections, paragraphs, sentences, and words. In management, hierarchies arise from the dual desires to simplify complex tasks and to control complex organizations ("divide and rule").

Self-organization. Internal interactions can lead to the emergence of order within complex systems, even in the absence of external influences. Several processes are known to promote self-organization. Ilya Prigogine introduced the term "dissipative systems" to describe systems (including living things) that are far from equilibrium and share energy with their environment. He argued that in dissipative systems local irregularities can grow and spread, allowing large-scale patterns and order to emerge. Crystal growth is an example. *See* CRYSTAL GROWTH.

Feedback plays an important role in self-organization. It occurs where the output of a system is fed back as input to the same system. Negative feedback dampens disturbances and stabilizes local patterns and behavior. Familiar examples include thermostats, speed governors, and other control devices that maintain systems in a steady state (sometimes called homeostasis). Positive feedback amplifies local effects, allowing them to grow and spread. Control systems are designed to avoid positive feedback because of its destabilizing effect. In contrast, positive feedback plays an important role in natural systems by contributing to the formation of large-scale order. In interstellar gas clouds, for instance, gravity leads to the formation of large centers of mass. These centers gather material faster, and at the expense of smaller ones, eventually forming stars and planets. Within an ant colony, piles of food, eggs, and waste grow in a similar way. Eventually this sorting of materials creates the order seen in ant colonies. Firing of a laser occurs when groups of atoms synchronize their outputs by "enslaving" other atoms. *See* CONTROL SYSTEMS; HOMEOSTASIS; LASER; MOLECULAR CLOUD.

Other processes that promote self-organization include encapsulation, adaptation, and evolution within living systems.

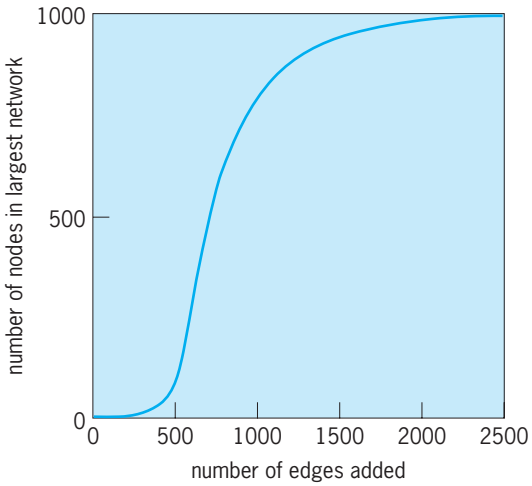

Fig. 3. Connectivity avalanche. Edges, each joining a pair of nodes, are added at random to a set of 1000 nodes. The avalanche begins when about 500 edges are added.

Emergence. In a system of objects, interactions lead to self-organization, which results in the emergence of large-scale features. An ant colony emerges from the joint activity of thousands of individual ants interacting with each other and with their environment. Emergent properties are not predictable from the properties of the individuals, but often involve the imposing of constraints on individuals. In module formation, for instance, individual interactions with the outside could disrupt the module's integrity.

Module formation is a crucial mechanism in the emergence of order in many complex systems. In a growing embryo, for instance, cells differentiate at a very early stage of development, forming modules that grow into separate limbs and organs. *See* CELL DIFFERENTIATION.

A related problem is the paradox of self-replication. This originated during the seventeenth century with attempts to explain human reproduction. Human eggs or sperm were thought to contain homunculi, tiny bodies from which children would grow. However, this idea led to a paradox: a single homunculus would need to have nested within it smaller homunculi from which future generations would spring. The paradox of self-replication arises because some part a self-replicating system apparently needs to contain a copy of the whole system. How is it possible, for instance, for a plant or animal to reproduce without including a complete description of itself? The paradox is resolved by observing that only information about the process of replication needs to be embedded within the system, not the entire system itself. The rest of the information can be derived from interaction of the system with its environment. John von Neumann suggested that there is a threshold of complexity below which automata are too simple to be capable of self-replication. However, later researchers, such as Lionel and Roger Penrose, demonstrated that even certain simple block structures could replicate themselves. Tom Ray's Tierra model provided a practical demonstration of a computational environment within which programs emulated organisms by containing code for reproducing themselves.

Computational complexity. In computing, the complexity of a problem is measured by the time (or space) required to solve it, and the way that time increases as the size of the input increases. For instance, given a list of the locations of n towns, the time needed to find the most northerly town increases in proportion to n [denoted $O(n)$], but the time required to find the shortest distance between pairs of towns increases in proportion to n^2 [denoted $O(n^2)$].

In problem solving, an important advantage of modularity is to reduce combinatorial complexity. In a rail network linking, say, 100 stations, separate fares for every pair of stations would require a table with 4950 entries, but grouping the stations into, say, 3 zones would reduce the table to no more than 6 different fares.

Problems are often classified according to their computational complexity. Conventional terminology defines the class P to consist of all problems that can be solved by a deterministic algorithm for which the time complexity is a polynomial function of the input (for example, the towns examples above). The class NP consists of problems that a nondeterministic machine can solve in polynomial time. An unresolved theoretical question is whether the classes P and NP are identical. *See* CRYPTOGRAPHY.

Hard problems occur when the time complexity increases faster than any polynomial of the input. In the traveling salesman problem, for instance, a salesperson needs to find a route that minimizes the total distance traveled while visiting a number of towns (without doubling back). The solution involves checking all possible orderings of a list of the towns.

Attempts to link complexity to thermodynamics and information theory led to the Kolmogorov-Chaitin definition of algorithmic complexity, which is the length of the shortest program needed to compute a given pattern. For example, a table of sports scores in general contains more or less random entries and cannot be compressed. In contrast, a table of logarithms can be condensed to a formula that calculates all the entries. Describing the path of a flock of birds greatly reduces the detail needed to describe the paths of all the individual birds. These ideas extend into many aspects of computing. Methods such as minimum message length (MML) adapt them to provide practical methods of classification and problem solving. Data compression algorithms typically employ a fixed program and reduce a file to the data needed to regenerate the original pattern. Fractals may be regarded as patterns that are produced by repeating a given algorithm on different scales. *See* DATA COMPRESSION; FRACTALS; INFORMATION THEORY.

As information technology has spread throughout society, the idea of natural computation has been increasingly influential. This idea treats natural processes as forms of computing. For instance, scientific studies sometimes employ computational analogies (for example, animals as programmed robots) to understand complex processes in natural systems. Likewise, computing has borrowed ideas from nature, which has evolved ways of solving complex problems. This practice has resulted in a proliferation of biological terms such as cellular automata, genetic algorithms, neural networks, and swarm intelligence. *See* NEURAL NETWORK.

Role of simulation. Complex systems usually behave in nonlinear fashion. They are often mathematically intractable, and the richness of their behavior can make them inherently unpredictable. Simulation models provide a practical way of representing them. Certain representations are widely used to simulate specific kinds of complex systems, notably cellular automata, spin glasses, and agent-based models.

Cellular automata are grids of cells, each of which is an automaton that interacts with its neighbors. First introduced by John von Neumann, cellular automata became popular during the 1970s as a result of John Conway's cellular automaton Game of Life.

Fig. 4. Cellular automaton simulation of epidemic spread through a landscape. The cells represent areas of the land surface. Filled cells have been infected.

They have been applied to many physical processes, such as flows and spread of epidemics (**Fig. 4**). In a simple cellular automaton model of disease spread, for instance, each cell would denote an area within a landscape, and the cell's state at any given time might be "susceptible," "infected," or "immune." The epidemic would spread from infected cells to neighboring susceptible cells. The virulence of the disease would determine the chances of a susceptible cell becoming infected; otherwise it would become immune. *See* CELLULAR AUTOMATA.

Spin glasses (regular lattices, in which each node has a "spin") have been used to study processes, such as glass formation, involving transitions from incoherent to coherent configurations of spin states in magnetic materials. *See* SPIN GLASS.

Agent-based models consist of many independent agents (for example, people, animals, or companies) that interact with one another and behave according to defined sets of rules. Cellular automata are a subclass of agent-based models. Early examples of this approach included Paulien Hogeweg's work on the emergence of social order within bumblebee colonies and Craig Reynold's "boids" models of flocking behavior in birds.

Simulation models provide several practical advantages for dealing with complexity. One is the potential to perform "virtual experiments" for problems where real-world experiments are impractical. Another is the ability to explore "what-if" scenarios when complexity makes prediction impossible. *See* SIMULATION.

Future challenges. Complexity theory is becoming more important as the problems faced by science and society grow more complex. Is it possible to evolve designs to specification for huge networks that are both robust and efficient? The Internet, for instance, poses enormous challenges as a self-organizing system for gathering, storing, and distributing information. *See* INTERNET.

In biology perhaps the greatest challenge at present is to understand how genetic regulatory networks control growth and development. During the 1970s, Stuart Kauffmann proposed a switching model of genetic control that was based on Boolean networks. Research has increasingly pointed to modular organization within the genome. Development of the eye, for instance, is encoded by a set of genes whose activity is controlled by a single "eyeless gene." *See* DEVELOPMENTAL GENETICS; MOLECULAR BIOLOGY.

One far-reaching impact of communications and globalization is that the world is increasingly rich in interactions. Cascading interactions often mean that events in one part of the world can affect events on the other side of the globe.

David G. Green; Tania Bransden

Bibliography. P. Bak, *How Nature Works: The Science of Self-Organized Criticality*, Oxford University Press, 1996; T. R. J. Bossomaier and D. G. Green (eds.), *Complex Systems*, Cambridge University Press, 2000; P. Coveney and R. Highfield, *Frontiers of Complexity*. Faber and Faber, London, 1995; S. H. Strogatz, *Sync.*, Penguin, London, 2003.

Composite beam

A structural member composed of two or more dissimilar materials joined together to act as a unit. An example in civil structures is the steel-concrete composite beam in which a steel wide-flange shape (I or W shape) is attached to a concrete floor slab (**Fig. 1**). The many other kinds of composite beam include steel-wood, wood-concrete, and plastic-concrete or advanced composite materials–concrete. Composite beams as defined here are different from beams made from fiber-reinforced polymeric materials. *See* COMPOSITE MATERIAL.

There are two main benefits of composite action in structural members. (1) By rigidly joining the two parts together, the resulting system is stronger than the sum of its parts. A simple example is a pair of timber 2 by 4 in. (5 by 10 cm) studs joined by

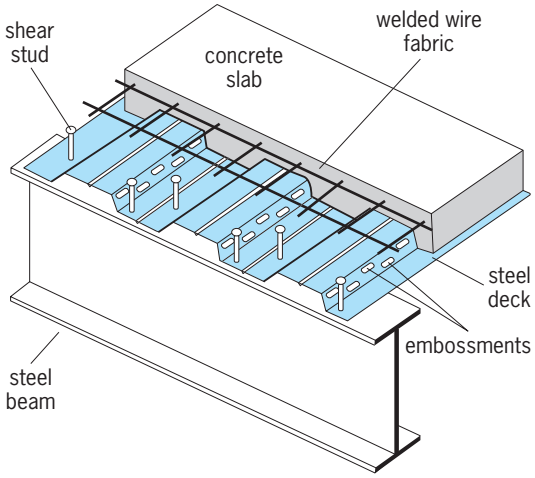

Fig. 1. Typical composite floor system.

nails. The two timber pieces connected together are far stronger and more rigid than the two individual pieces stacked together. A key feature of this system is the connecting elements, nails in this case, which provide the transfer of horizontal (shear) forces in the system. (2) Composite action can better utilize the properties of each constituent material. In steel-concrete composite beams, for example, the concrete is assumed to take most or all of the compression while the steel takes all the tension. Since concrete is very strong in compression but weak in tension (the ratio of tensile to compressive strength being approximately 1:10), and steel is prone to instability (buckling) in compression, the composite beam represents the most efficient use of these two materials. *See* CONCRETE; SHEAR; STEEL.

In general, it is assumed that these composite beams are loaded primarily in flexure and that the steel beam carries all of the tension and the slab part or all of the compression. The bending forces on the beam give rise to large horizontal (shear) forces at the interface of the materials. To transfer this force, several mechanisms can be considered, including adhesion, friction, and bearing. The first two are generally disregarded because of their lack of reliability, and in most composite systems the shear connection is provided by steel elements welded to the steel beam. These elements transfer the shear forces by bearing against the concrete. The most common type of shear connector is a shear stud, which resembles a bolt but has no threads and a smaller head (Fig. 1). Shear studs typically range from $\frac{1}{2}$ to $\frac{3}{4}$ in. (13 to 19 mm) in diameter, and from $3\frac{1}{2}$ to 5 in. (8 to 13 cm) in length. A typical $\frac{3}{4}$ -in. shear stud can transfer about 20 kips (90 kilonewtons) of horizontal force (1 kip = 1000 lb or 454 kg load). The shear stud requires a head to resist any tendency of the slab and beam to separate vertically. The studs typically are distributed evenly along the length of the beam, with spacing ranging from 6 to 16 in. (150 to 400 mm). Because typical steel decks have troughs at 12 in. (150 mm) spacing, this is the most common spacing for shear studs. It is also common to see several studs side by side in the direction perpendicular to the main axis of the beam. This arrangement is necessary in steel bridges, because the size of the steel beam is large and thus the shear forces are large. *See* STRESS AND STRAIN.

There are numerous reasons why steel-concrete composite beams have long been recognized as one the most economical structural systems for both multistory steel buildings and steel bridges. Buildings and bridges require a floor slab to provide a surface for occupants and vehicles, respectively. Concrete is the material of choice for the slab because its mass and stiffness can be used to reduce deflections and vibrations of the floor system and to provide the required fire protection. The supporting system underneath the slab, however, is often steel because it offers superior strength-weight and stiffness-weight ratio, ease of handling, and rapid construction cycles. Since both the steel and concrete are already present in the structures, it is logical to connect them together to better utilize their strength and stiffness. The typical steel-concrete composite beam is 60 to 80% stronger and stiffer than the steel beam alone. The stiffness and strength gains are greater for shallow beams [16 to 24 in. (40 to 60 cm) in depth] and less for deeper sections.

Three important developments have contributed to the widespread use of steel-concrete composite beams. (1) Automated systems were developed to provide the connection between the steel beam and the concrete. This connection was originally provided by welding steel channels, spirals, or serpentines to the top of the steel beam. In the 1950s a system utilizing shear studs that could be welded with a portable system (stud gun) was developed, and is in almost universal use today. (2) In the 1960s and 1970s, steel floor decking was developed. Steel decking consists of thin corrugated metal sheets that are laid between the steel beams and used as the formwork to cast the concrete. This made concrete casting much simpler, as the formwork was left in place to provide much of the tensile capacity of the slab. (3) Alternative types of steel beams were developed. Due to height limitations and the need for complex HVAC (heating, ventilation, and air conditioning), electrical, and communication installations, several variations on the traditional composite beam have been developed. They include composite joists, composite trusses, and stub-girders (**Fig. 2**). These systems are intended to provide high span/depth ratios while retaining flexibility in relocating building services. Systems using joists and trusses are popular since they avoid the need to make expensive penetrations in the beam web. *See* CONCRETE SLAB.

Composite beams in buildings have spans ranging typically from about 20 to 40 ft (6 to 12 m) for traditional beams but can be as large as 80 to 100 ft (24 to 30 m) for systems using trusses rather than I sections for the steel portion. Until recently, most

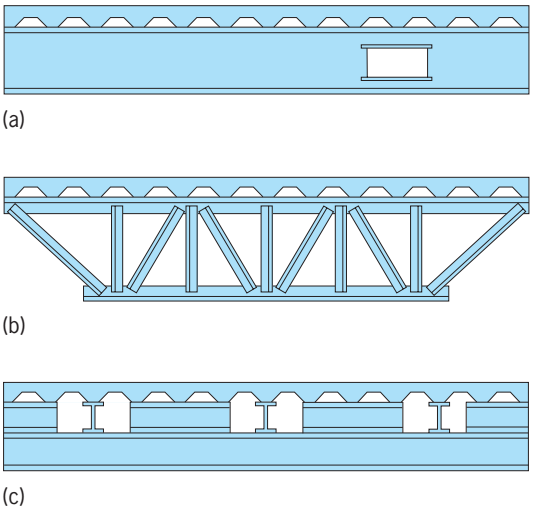

Fig. 2. Types of composite beams. (*a*) Composite beam showing reinforced web opening. (*b*) Composite joist or truss with double angle members. (*c*) Stub-girder system showing perpendicular floor beams.

composite beams were designed for full interaction, meaning that enough studs were provided to transfer the full capacity of the steel beam or concrete slab, whichever was lower. Today, composite beams are often designed for partial interaction; that is, they use fewer studs than needed to transfer the full capacity of the steel or concrete. Unlike composite beams in buildings, where the design of the shear connection is controlled by strength, the design of shear studs for composite beams in bridges is controlled by the fatigue capacity of the studs. Roberto Leon

Bibliography. R. P. Johnson and R. J. Buckby, *Composite Structures of Steel and Concrete*, vol. 2: *Bridges*, 2d ed., Collins, 1986; I. M. Viest et al., *Composite Construction: Design for Buildings*, ASCE/McGraw-Hill, 1997.

Composite laminates

Assemblages of layers of fibrous composite materials (**Fig. 1**) which can be tailored to provide a wide range of engineering properties, including inplane stiffness, bending stiffness, strength, and coefficients of thermal expansion.

The individual layers consist of high-modulus, high-strength fibers in a polymeric, metallic, or ceramic matrix material. Fibers currently in use include graphite, glass, boron, and silicon carbide. Typical matrix materials are epoxies, polyimides, aluminum, titanium, and alumina. Layers of different materials may be used, resulting in a hybrid laminate. The individual layers generally are orthotropic (that is, with principal properties in orthogonal directions) or transversely isotropic (with isotropic properties in the transverse plane) with the laminate then exhibiting anisotropic (with variable direction of principal properties), orthotropic, or quasi-isotropic properties. Quasi-isotropic laminates exhibit isotropic (that is, independent of direction) inplane response but are not restricted to isotropic out-of-plane (bending) response. Depending upon the stacking sequence of the individual layers, the laminate may exhibit coupling between inplane and out-of-plane response. An example of bending-stretching coupling is the presence of curvature developing as a result of inplane loading. *See* COMPOSITE MATERIAL; METAL MATRIX COMPOSITE; POLYMERIC COMPOSITE.

Lamination theory. Classical lamination theory describes the mechanical response of any composite laminate subjected to a combination of inplane and bending loads. The laminate in Fig. 1 uses a global x-y-z coordinate system with z perpendicular to the plane of the laminate and positive downward. The origin of the coordinate system is located on the laminate midplane. The laminate has N layers numbered from top to bottom. Each layer has a distinct fiber orientation denoted θ_k. The z coordinate to the bottom of the kth layer is designated z_k with the top of the layer being z_{k-1}. The thickness, t_k, of any layer is then $t_k = z_k - z_{k-1}$. The top surface of the laminate is denoted z_0, and the total thickness is $2H$.

Assumptions. It is assumed that (1) there is perfect bonding between layers; (2) each layer can be represented as a homogeneous material with known effective properties which may be isotropic, orthotropic, or transversely isotropic; (3) each layer is in a state of plane stress; and (4) the laminate deforms according to the Kirchhoff (1850) assumptions for bending and stretching of thin plates: (a) normals to the midplane remain straight and normal to the deformed midplane after deformation, and (b) normals to the midplane do not change length.

Strain-displacement relationships. The Kirchhoff assumptions require the out-of-plane strains ε_z, γ_{zx}, and γ_{zy} to be identically zero. The nonzero strains, $\{\varepsilon\}$, at any z location can then be expressed in terms of the midplane strains, $\{\varepsilon^0\}$, and curvatures, $\{\kappa\}$, in Eq. (1).

$$\{\varepsilon\} = \{\varepsilon^0\} + z\{\kappa\} \tag{1}$$

Stresses. The stresses, $\{\sigma\}^k$, at any z location in the kth layer are now determined from the strains [Eq. (1)] and the stress-strain constitutive equations expressed in terms of the stiffness matrix $[\overline{Q}]^k$ of the kth layer. The result is shown in Eq. (2).

$$\{\sigma\}^k = [\overline{Q}]^k(\{\varepsilon^0\} + z\{\kappa\}) \tag{2}$$

Inplane forces per unit length. The inplane forces per unit length $\{N\}$ (**Fig. 2**) are defined as the through-thickness integrals of the planar stresses in the laminate. After integration, the results are shown in Eq. (3), where $[A]$ and $[B]$ are coefficient matrices

$$\{N\} = [A]\{\varepsilon^0\} + [B]\{\kappa\} \tag{3}$$

that are functions of the layer material properties and laminate stacking sequence. The $[A]$ matrix represents the inplane stiffness, and the $[B]$ matrix defines the bending-stretching coupling. $[A]$ is a function of the layer thicknesses, but is independent of the stacking sequence of the layers. In contrast, $[B]$ is dependent on both the layer thicknesses and the stacking sequence of the individual layers. If the laminate

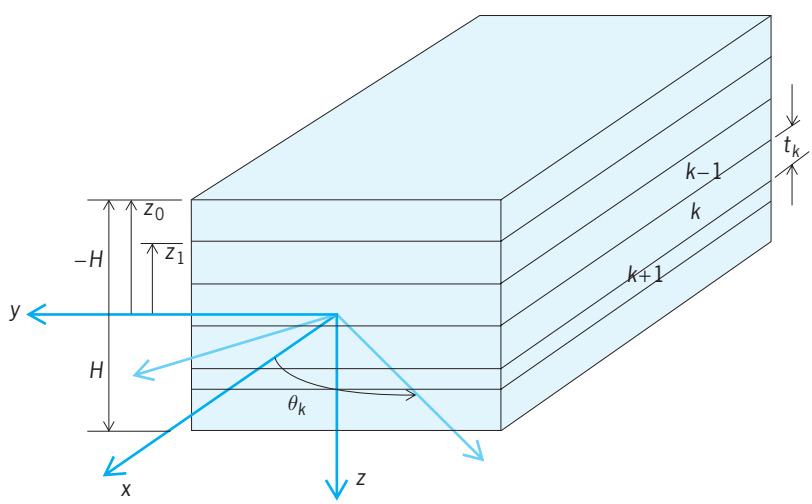

Fig. 1. Composite laminate consisting of layers with varying thickness.

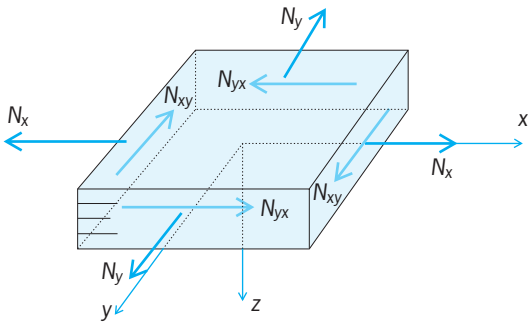

Fig. 2. Inplane forces per unit length, $\{N\}$.

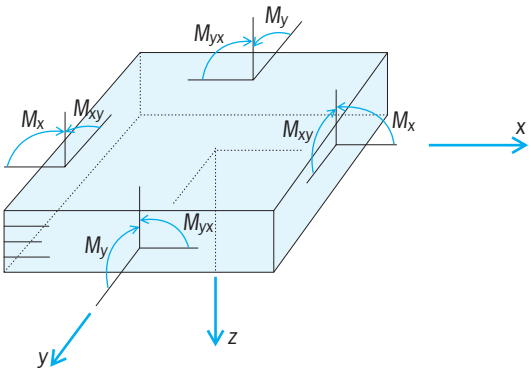

Fig. 3. Moments per unit length, $\{M\}$.

stacking sequence is symmetric about the laminate midplane, then $[B] = 0$.

Moments per unit length. The moments per unit length, $\{M\}$ (**Fig. 3**), are defined as the integrals of the differential forces, $\sigma_i dz$, times the moment arm, z, integrated over the laminate thickness, resulting in Eq. (4). Here the stacking sequence–dependent

$$\{M\} = [B]\{\varepsilon^0\} + [D]\{\kappa\} \tag{4}$$

bending stiffness matrix, $[D]$, is a coefficient matrix similar to $[A]$ and $[B]$.

Laminate constitutive relations. Equations (3) and (4) can be combined to give the fundamental equations of lamination theory, Eqs. (5) and (6). Equations (5)

$$\{N\} = [A]\{\varepsilon^0\} + [B]\{\kappa\} \tag{5}$$

$$\{M\} = [B]\{\varepsilon^0\} + [D]\{\kappa\} \tag{6}$$

and (6) exhibit coupling between the bending and stretching response of a laminate through the $[B]$ matrix. If $[B] = 0$, then the inplane (stretching) response is decoupled from the bending response. Indeed, there is no bending-stretching coupling for laminates which are symmetric about their midplane since $[B] = 0$ for symmetric laminates.

Laminate engineering constants. Expressions for the effective engineering constants of symmetric laminates can be determined from consideration of inplane forces applied individually. The laminate compliance, $[a^*]$, is defined in Eq. (7), where $[A]^{-1}$

$$[a^*] \equiv 2H[A]^{-1} \tag{7}$$

is the inverse of the $[A]$ matrix and $2H$ is the total thickness of the laminate. The engineering constants of symmetric laminates can be expressed as functions of the individual coefficients of the laminate compliance.

Examples demonstrating the variability of the effective engineering constants as a function of fiber orientations and laminate stacking sequence are provided in the following paragraphs for unidirectional off-axis and symmetric angle-ply laminates. A unidirectional laminate (also called a lamina) consists of layers all having the same fiber orientation. A symmetric angle-ply laminate has an equal number of layers at $+\theta$ and $-\theta$ fiber orientations located symmetrically about the laminate midplane. These results demonstrate one of the major advantages of fibrous composite materials, namely, the ability to tailor the material properties through the choice of fiber orientations, layer thicknesses, and materials. Most of the results discussed are for a typical carbon-epoxy composite, T300/5208. This composite consists of T300 carbon fibers in an epoxy matrix (5208).

Axial modulus. Predictions for the effective axial modulus (stiffness in the direction of 0° fiber orientation) E_x of angle-ply laminates are compared to the predictions for unidirectional off-axis lamina. The angle-ply laminate has higher stiffness than the off-axis lamina for fiber angles ranging from 0° to approximately 60°. Experimental results for T300/5208 angle-ply laminates exhibit very good correlation with the theoretical predictions. The higher axial stiffness for the angle-ply laminate is a result of the constraining effect (and resulting multiaxial state of stress) that the adjacent layers of the laminate have on each other.

Axial Poisson's ratio. The coupling between the transverse strain (strain perpendicular to the direction of applied stress) and the strain in the direction of applied stress is called Poisson's ratio. It is an important engineering material property. The comparison of Poisson's ratios (**Fig. 4**) is one of the most interesting

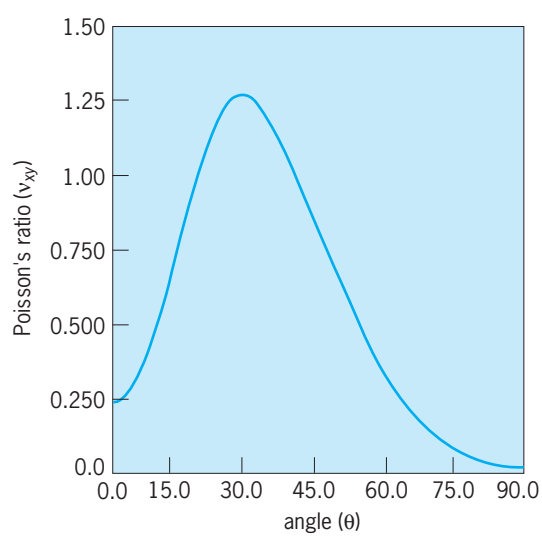

Fig. 4. Poisson's ratio comparisons for the composite laminate T300/5208.

and surprising results for laminated fibrous composites. There is a very large increase in Poisson's ratio when an off-axis ply is laminated with another off-axis ply of opposite sign. This is the case for a wide range of fiber orientations. The angle-ply laminate exhibits a maximum Poisson's ratio in excess of 1.25 (for this particular carbon/epoxy) at an angle of approximately 27°. That the Poisson's ratio can be in excess of 1.0 is very surprising to those accustomed to working with metals, where the maximum Poisson's ratio is 0.5. The high Poisson's ratio in these laminates is due to the internal state of stress in the individual layers. The laminates behave more like structures with variable internal mechanisms as compared with homogeneous materials with uniform stress states. *See* STRESS AND STRAIN.

Through-thickness Poisson's ratio. The ratio of the strain through the thickness of the laminate with that in the direction of the applied stress is called the through-thickness Poisson's ratio. For the unidirectional lamina, the through-thickness Poisson's ratio increases monotonically from the axial value to the transverse value as the fiber orientation angle increases. In sharp contrast, the through-thickness Poisson's ratio for angle-ply laminates initially decreases with increasing angle, passing through zero at approximately 13°, attains a minimum negative value at approximately 25°, then passes through zero again at approximately 40° as it increases to the transverse value at 90° (**Fig. 5**). This is another example of the variety of material properties that are possible with fibrous composites. The minimum value and the angles corresponding to zero through-thickness Poisson's ratios can be altered by modification of the laminate configuration.

Shear modulus. Shear modulus is a measure of the materials resistance (stiffness) to shear loading. The angle-ply laminate is much stiffer than the unidirectional lamina for essentially all fiber orientations. At $\theta = 45°$ where the shear stiffnesses of both the lamina and the laminate are largest, the stiffness of the laminate is more than 3.5 times that of the lamina for the carbon/epoxy under consideration. The results clearly indicate that ±45° fiber orientations are desired in structures requiring high shear stiffness. *See* SHEAR.

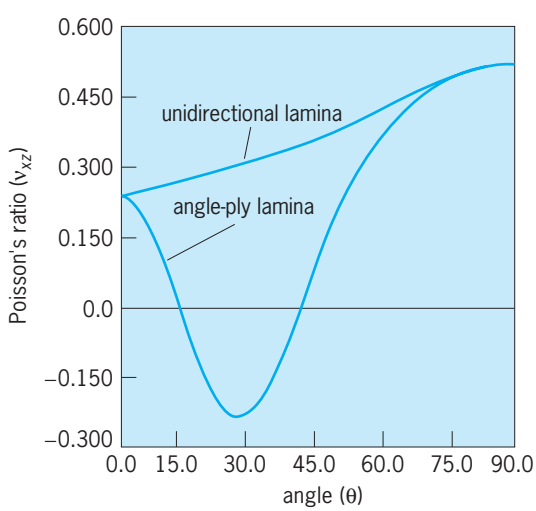

Fig. 5. Through-thickness Poisson's ratios for composite laminate T300/5208.

Coefficients of mutual influence. For anisotropic materials such as fibrous composites, there are additional important engineering properties that describe material behavior. These properties are called coefficients of mutual influence. They are similar to Poisson's ratios in that they provide an indication of the coupling between normal and shear components of strain. This type of coupling is not present in isotropic materials. One of the coefficients of mutual influence for unidirectional off-axis lamina of two different materials (T300/5208 carbon/epoxy and SCS6/Ti-15-3, a metal matrix composite) is the ratio of shear strain to axial strain for applied stress in the axial direction. The most significant features of the results are that the coefficient exhibits vary large gradients and magnitudes for the carbon/epoxy in the vicinity of $\theta = \pm 12°$, and that there is a major difference in maximum values between the two materials. The coefficient of mutual influence can be nearly twice as large as Poisson's ratio. Thus the coupling between axial and shear response can be twice as large as the coupling between axial and transverse response for unidirectional lamina. The effective coefficient of mutual influence is zero for angle-ply laminates because the $+\theta$ and $-\theta$ fiber orientations have the effect of offsetting one another.

Laminate design. A wide variety of effective material properties can be obtained with one type of fibrous composite simply through changes in the stacking arrangement (layer thicknesses and fiber orientations of the individual layers) of the laminate. For example, changing the type of composite provides even more variety in the properties. Thus, material properties of the laminate can be tailored. This is a very important feature of fibrous composites because the material can be designed to have specific material properties.

Coefficient of thermal expansion. The wide variety of coefficients of thermal expansion are possible through changes in the stacking arrangement of a given carbon/epoxy. The coefficient of thermal expansion is the strain associated with a change in temperature of 1°. Most materials have positive coefficients of expansion and thus expand when heated and contract when cooled. The effective axial coefficient of thermal expansion of the carbon/epoxy can be positive, negative, or zero, depending upon the laminate configuration. Laminates with zero coefficient of thermal expansion are particularly important because they do not expand or contract when exposed to a temperature change. Composites with zero (or near zero) coefficient of thermal expansion are therefore good candidates for application in space structures where the temperature change can be 500°F (from −250 to +250°F) [278°C (from −157 to +121°C)] during an orbit in and out of the Sun's proximity. There are many other applications where thermal expansion is a very important consideration.

Carl T. Herakovich

Bibliography. S. B. Dong, K. S. Pister, and R. L. Taylor, On the theory of laminated anisotropic shells and plates, *J. Aerosp. Sci.*, 29:969–975, 1962; C. T. Herakovich, *Mechanics of Fibrous Composites*, John Wiley, New York, 1998; G. Kirchhoff, *J. f. Math.* (*Crelle*), Bd. 40, 1850; K. S. Pister and S. B. Dong, Elastic bending of layered plates, *J. Eng. Mech. Div.*, ASCE, EM 4:1–10, October 1959; E. Reissner and Y. Stavsky, Bending and stretching of certain types of heterogeneous aelotropic elastic plates, *J. Appl. Mech.*, ASME, 28:402–408, 1961.

Composite material

A material system composed of a mixture or combination of two or more macroconstituents that differ in form or material composition and are essentially insoluble in each other. This definition is considered to be too broad by some engineers because it includes many materials that are not usually thought of as composites. For example, in many of the particulate-type composites, such as dispersion-hardened alloys and cermets, the composite structure is microscopic rather than macroscopic. Also, this definition does not draw the line between composite materials and composite structures. However, instead of trying to establish a distinction between materials and structures, it is more useful to make a distinction between mill composites (such as nonmetallic laminates, clad metals, and honeycomb) and specialty composites (such as tires, rocket nose cones, and glass-reinforced plastic boats).

Constituents and Construction

In principle, composites can be constructed of any combination of two or more materials—metallic, organic, or inorganic; but the constituent forms are more restricted. The matrix is the body constituent, serving to enclose the composite and give it bulk form. Major structural constituents are fibers, particles, laminae or layers, flakes, fillers, and matrices. They determine the internal structure of the composite. Usually, they are the additive phase.

Because the different constituents are intermixed or combined, there is always a contiguous region. It may simply be an interface, that is, the surface forming the common boundary of the constituents. An interface is in some ways analogous to the grain boundaries in monolithic materials. In some cases, however, the contiguous region is a distinct added phase, called an interphase. Examples are the coating on the glass fibers in reinforced plastics and the adhesive that bonds the layers of a laminate together. When such an interphase is present, there are two interfaces, one between the matrix and the interphase and one between the fiber and the interface.

Interfaces are among the most important yet least understood components of a composite material. In particular, there is a lack of understanding of processes occurring at the atomic level of interfaces, and how these processes influence the global material behavior. There is a close relationship between processes that occur on the atomic, microscopic, and macroscopic levels. In fact, knowledge of the sequence of events occurring on these different levels is important in understanding the nature of interfacial phenomena. Interfaces in composites, often considered as surfaces, are in fact zones of compositional, structural, and property gradients, typically varying in width from a single atom layer to micrometers. Characterization of the mechanical properties of interfacial zones is necessary for understanding mechanical behavior.

Nature and performance. Several classification systems for composites have been developed, including classification by (1) basic material combinations, for example, metal-organic or metal-inorganic; (2) bulk-form characteristics, such as matrix systems or laminates; (3) distribution of the constituents, that is, continuous or discontinuous; and (4) function, for example, electrical or structural.

There are five classes under the classification by basic material combinations: (1) fiber composites, composed of fibers with or without a matrix; (2) flake composites, composed of flat flakes with or without a matrix; (3) particulate composites, composed of particles with or without a matrix; (4) filled (or skeletal) composites, composed of a continuous skeletal matrix filled by a second material; and (5) laminar composites, composed of layer or laminar constituents.

There is also a classification based on dimensions. The dimensions of some of the components of composite materials vary widely and overlap the dimensions of the microstructural features of common conventional materials (**Fig. 1**). They range from extremely small particles or fine whiskers to the large aggregate particles or rods in reinforced concrete. *See* CRYSTAL WHISKERS; REINFORCED CONCRETE.

The behavior and properties of composites are determined by the composition, form and arrangements, and interaction between the constituents. The intrinsic properties of the materials of which the constituents are composed largely determine the general order or range of properties of the composite. Structural and geometrical characteristics—that is, the shape and size of the individual constituents, their structural arrangement and distribution, and the relative amount of each—contribute to overall performance. Of far-reaching importance are the effects produced by the combination and interaction of the constituents. The basic principle is that by using different constituents it is possible to obtain combinations of properties and property values that are different from those of the individual constituents.

A performance index is a property or group of properties that measures the effectiveness of a material in performing a given function. The values of performance indices for a composite differ from those of the constituents.

Fiber-matrix composites. Fiber-matrix composites have two constituents and usually a bonding phase as well.

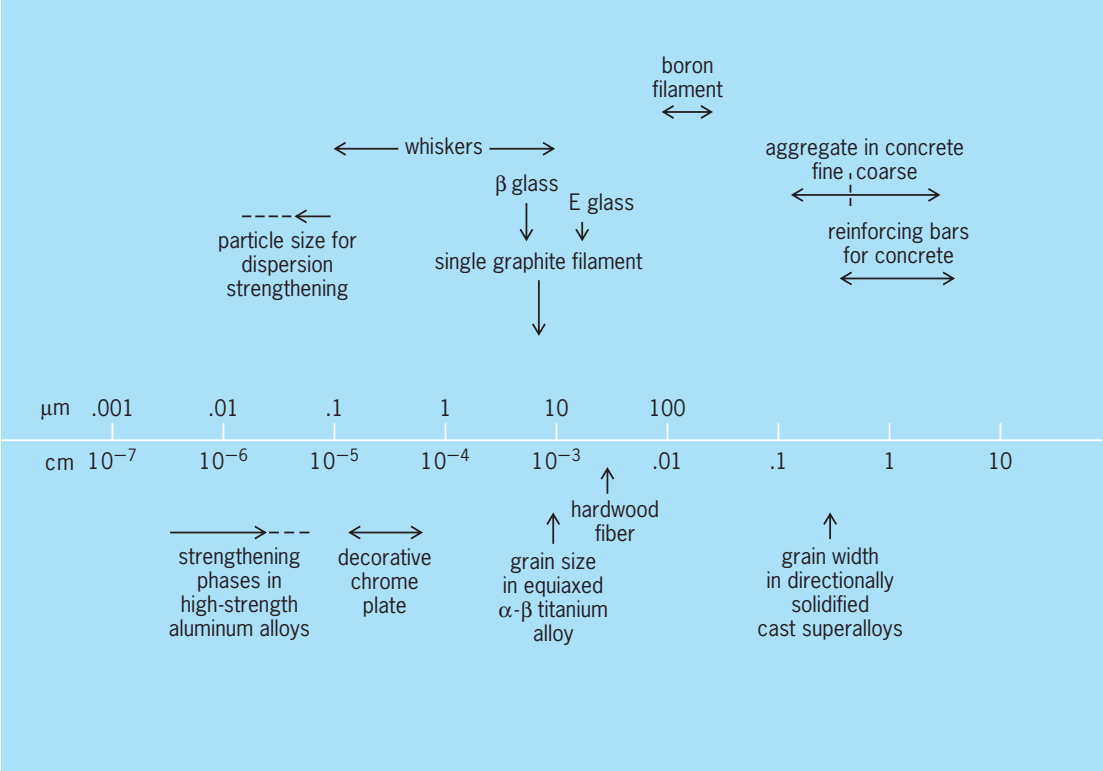

Fig. 1. Dimensional range of microstructural features in composite and conventional materials. Filament and fiber dimensions are diameters. 1 cm = 0.39 in.

Fibers. The performance of a fiber-matrix composite depends on orientation, length, shape, and composition of the fibers; mechanical properties of the matrix; and integrity of the bond between fibers and matrix. Of these, orientation of the fibers is perhaps most important.

Fiber orientation determines the mechanical strength of the composite and the direction of greatest strength. Fiber orientation can be one-dimensional, planar (two-dimensional), or three-dimensional. The one-dimensional type has maximum composite strength and modulus in the direction of the fiber axis. The planar type exhibits different strengths in each direction of fiber orientation; and the three-dimensional type is isotropic but has greatly decreased reinforcing values. The mechanical properties in any one direction are proportional to the amount of fiber by volume oriented in that direction. As fiber orientation becomes more random, the mechanical properties in any one direction become lower.

Fiber length also impacts mechanical properties. Fibers in the matrix can be either continuous or short. Composites made from short fibers, if they could be properly oriented, could have substantially greater strengths than those made from continuous fibers. This is particularly true of whiskers, which have uniform high tensile strengths. Both short and long fibers are also called chopped fibers. Fiber length also has a bearing on the processibility of the composite. In general, continuous fibers are easier to handle but have more design limitations than short fibers.

Bonding. Fiber composites are able to withstand higher stresses than their individual constituents because the fibers and matrix interact, resulting in redistribution of the stresses. The ability of constituents to exchange stresses depends on the effectiveness of the coupling or bonding between them. Bonding can sometimes be achieved by direct contact of the two phases, but usually a specially treated fiber must be used to ensure a receptive adherent surface. This requirement has led to the development of fiber finishes, known as coupling agents. Both chemical and mechanical bonding interactions occur for coupling agents.

Voids (air pockets) in the matrix are one cause of failure. A fiber passing through the void is not supported by resin. Under load, the fiber may buckle and transfer stress to the resin, which readily cracks. Another cause of early failure is weak or incomplete bonding. The fiber-matrix bond is often in a state of shear when the material is under load. When this bond is broken, the fiber separates from the matrix and leaves discontinuities that may cause failure. Coupling agents can be used to strengthen these bonds against shear forces. *See* SHEAR.

Reinforced plastics. Probably the greatest potential for lightweight high-strength composites is represented by the inorganic fiber–organic-matrix composites, and no composite of this type has proved as successful as glass-fiber-reinforced plastic composites. As a group, glass-fiber–plastic composites have the advantages of good physical properties, including strength, elasticity, impact resistance, and dimensional stability; high strength-to-weight ratio; good

electrical properties; resistance to chemical attack and outdoor weathering; and resistance to moderately high temperatures (about 260°C or 500°F).

A critical factor in reinforced plastics is the strength of the bond between the fiber and the polymer matrix; weak bonding causes fiber pullout and delamination of the structure, particularly under adverse environmental conditions. Bonding can be improved by coatings and the use of coupling agents. Glass fibers, for example, are treated with silane (SiH_4) for improved wetting and bonding between the fiber and the matrix.

Generally, the greatest stiffness and strength in reinforced plastics are obtained when the fibers are aligned in the direction of the tension force. Other properties of the composite, such as creep resistance, thermal and electrical conductivity, and thermal expansion, are anisotropic. The transverse properties of such a unidirectionally reinforced structure are much lower than the longitudinal. Seven mechanical and thermal properties are of direct interest in assessing the potential of a new composite: density, modulus, strength, toughness, thermal conductivity, expansion coefficient, and heat capacity; others, such as fracture toughness and thermal diffusivity, are calculated from them.

Advanced Composites

Advanced composites comprise structural materials that have been developed for high-technology applications, such as airframe structures, for which other materials are not sufficiently stiff. In these materials, extremely stiff and strong continuous or discontinuous fibers, whiskers, or small particles are dispersed in the matrix. A number of matrix materials are available, including carbon, ceramics, glasses, metals, and polymers. Advanced composites possess enhanced stiffness and lower density compared to fiber-glass and conventional monolithic materials. While composite strength is primarily a function of the reinforcement, the ability of the matrix to support the fibers or particles and to transfer load to the reinforcement is equally important. Also, the matrix frequently dictates service conditions, for example, the upper temperature limit of the composite.

Reinforcements. Continuous filamentary materials that are used as reinforcing constituents in advanced composites are carbonaceous fibers, organic fibers, inorganic fibers, ceramic fibers, and metal wires. Reinforcing inorganic materials are used in the form of discontinuous fibers and whiskers. *See* STRENGTH OF MATERIALS.

Carbon and graphite fibers offer high modulus and the highest strength of all reinforcing fibers. These fibers are produced in a pyrolysis chamber from three different precursor materials—rayon, polyacrylonitrile (PAN), and pitch. High-modulus carbon fibers are available in an array of yarns and bundles of continuous filaments (tows) with differing moduli, strengths, cross-sectional areas, twists, and plies. *See* CARBON; GRAPHITE.

Almost any polymer fiber can be used in a composite structure, but the first one with high-enough tensile modulus and strength to be used as a reinforcement in advanced composites was an aramid, or aromatic polyamide, fiber. Aramid fibers have been the predominant organic reinforcing fiber; graphite is a close second. *See* MANUFACTURED FIBER; POLYMER.

The most important inorganic continuous fibers for reinforcement of advanced composites are boron and silicon carbide, both of which exhibit high stiffness, high strength, and low density. Continuous fibers are made by chemical vapor deposition processes. Other inorganic compounds that provide stiff, strong discontinuous fibers that predominate as reinforcements for metal matrix composites are silicon carbide, aluminum oxide, graphite, silicon nitride, titanium carbide, and carbon carbide. *See* BORON.

Polycrystalline aluminum oxide (Al_2O_3) is a commercial continuous fiber that exhibits high stiffness, high strength, high melting point, and exceptional resistance to corrosive environments. One method to produce the fibers is dry spinning followed by heat treatment.

Whiskers are single crystals that exhibit fibrous characteristics. Compared to continuous or discontinuous polycrystalline fibers, they exhibit exceptionally high strength and stiffness. Silicon carbide whiskers are prepared by chemical processes or by pyrolysis of rice hulls. Whiskers made of aluminum oxide and silicon nitride are also available. Particulates vary widely in size, characteristics, and function; and since particulate composites are usually isotropic, their distribution is usually random rather than controlled. *See* PYROLYSIS.

Organic-matrix composites. In many advanced composites the matrix is organic, but metal matrices are also used. Organic matrix material are lighter than metals, adhere better to the fibers, and offer more flexibility in shaping and forming. Ceramic matrix composites, carbon-carbon composites, and intermetallic matrix composites have applications where organic or metal matrix systems are unsuitable.

Materials. Epoxy resins have been used extensively as the matrix material. However, bismaleimide resins and polyimide resins have been developed to enhance in-service temperatures. Thermoplastic resins, polyetherketone, and polyphenylene sulfide are in limited use.

The continuous reinforcing fibers for organic matrices are available in the forms of monofilaments, multifilament fiber bundles, unidirectional ribbons, roving (slightly twisted fiber), and single-layer and multilayer fabric mats. Frequently, the continuous reinforcing fibers and matrix resins are combined into a nonfinal form known as a prepreg.

Fabrication. Many processes are available for the fabrication of organic matrix composites. The first process is contact molding in order to orient the unidirectional layers at discrete angles to one another. Contact molding is a wet method, in which the reinforcement is impregnated with the resin at the time of molding. The simplest method is hand

lay-up, whereby the materials are placed and formed in the mold by hand and the squeezing action expels any trapped air and compacts the part.

Molding may also be done by spraying, but these processes are relatively slow and labor costs are high, even though they can be automated. Many types of boats, as well as buckets for power-line servicing equipment, are made by this process.

Another process is vacuum-bag molding, where prepregs are laid in mold to form the desired shape. In this case, the pressure required to form the shape and achieve good bonding is obtained by covering the lay-up with a plastic bag and creating a vacuum. If additional heat and pressure are desired, the entire assembly is put into an autoclave. In order to prevent the resin from sticking to the vacuum bag and to facilitate removal of excess resin, various materials are placed on top of the prepreg sheets. The molds can be made of metal, usually aluminum, but more often are made from the same resin (with reinforcement) as the material to be cured. This eliminates any problem with differential thermal expansion between the mold and the part.

In filament winding, the resin and fibers are combined at the time of curing. Axisymmetric parts, such as pipes and storage tanks, are produced on a rotating mandrel. The reinforcing filament, tape, or roving is wrapped continuously round the form. The reinforcements are impregnated by passing them through a polymer bath. However, the process can be modified by wrapping the mandrel with prepreg material. The products made by filament winding are very strong because of their highly reinforced structure. For example, filament winding can be used directly over solid-rocket-propellant forms.

Pultrusion is a process used to produce long shapes with constant profiles, such as rods or tubing, similar to extruded metal products. Individual fibers are often combined into a tow, yarns, or roving, which consists of a number of tows or yarns collected into a parallel bundle without twisting (or only slightly so). Filaments can also be arranged in a parallel array called a tape and held together by a binder. Yarns or tows are often processed further by weaving, braiding, and knitting or by forming them into a sheetlike mat consisting of randomly oriented chopped fibers or swirled continuous fibers held together by a binder.

Weaving to produce a fabric is a very effective means of introducing fibers into a composite. There are five commonly used patterns (**Fig. 2**). Although weaving is usually thought of as a two-dimensional process, three-dimensional weaving is often employed.

Knitting is a process of interlooping chains of tow or yarn. Advantages of this process are that the tow or yarn is not crimped as happens in weaving, and higher mechanical properties are often observed in the reinforced product. Also, knitted fabrics are easy to handle and can be cut without falling apart.

In braiding, layers of helically wound yarn or tow are interlaced in a cylindrical shape, and interlocks can be produced at every intersection of fibers. During the process, a mandrel is fed through the center of a braiding machine at a uniform rate, and the yarn or tow from carriers is braided around the mandrel at a controlled angle. The machine operates like a maypole, the carriers working in pairs to accomplish the over-and-under sequencing. The braiding process is most effective for cylindrical geometries. It is used for missile heat shields, lightweight ducts, fluid-sealing components such as packings and sleevings, and tubes for insulation.

Carbon-carbon composites. A carbon-carbon composite is a specialized material made by reinforcing a carbon matrix with continuous carbon fiber. This type of composite has outstanding properties over a wide range of temperatures in both vacuum and inert atmospheres. It will even perform well at elevated temperatures in an oxidizing environment for short times. It has high strength, modulus, and toughness up to 2000°C (3600°F); high thermal conductivity; and a low coefficient of thermal expansion. A material with such properties is excellent for rocket motor nozzles and exit cones, which require high-temperature strength as well as resistance to thermal shock. Carbon-carbon composites are also used for aircraft and other high-performance brake applications that take advantage of the fact that carbon-carbon composites have the highest energy-absorption capability of any known material. If a carbon-carbon composite is exposed to an oxygen-containing atmosphere above 600°C (1100°F) for an

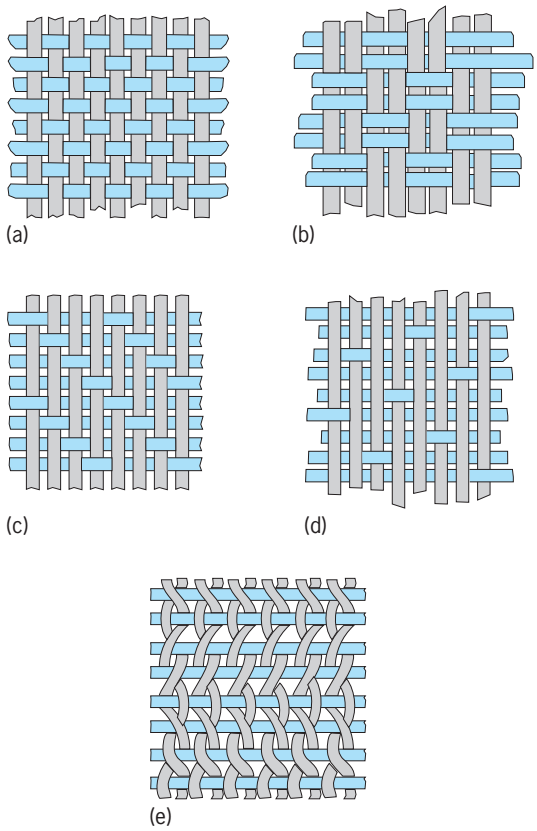

Fig. 2. Common weave patterns: (*a*) box or plain weave, (*b*) basket weave, (*c*) crowfoot, (*d*) long-shaft, and (*e*) leno weave.

appreciable time, it oxidizes, and therefore it must be protected by coatings.

Metal-matrix composites. Metal-matrix composites are usually made with alloys of aluminum, magnesium, or titanium; and the reinforcement is typically a ceramic in the form of particulates, platelets, whiskers, or fibers, although other systems may be used. Metal-matrix composites are often classified as discontinuous or continuous, depending on the geometry of the reinforcement. Particulates, platelets, and whiskers are in the discontinuous category, while the continuous category is reserved for fibers and wires. The type of reinforcement is important in the selection of a metal-matrix composite, because it determines virtually every aspect of the product, including mechanical properties, cost, and processing method. The primary methods for processing of discontinuous metal-matrix composites are powder metallurgy, liquid metal infiltration, squeeze or pressure casting, and conventional casting; however, most of these methods do not result in finished parts. Therefore, most discontinuously reinforced metal-matrix composites require secondary processing, which includes conventional wrought metallurgy operations such as extrusion, forging, and rolling; standard and nonstandard machining operations; and joining techniques such as welding and brazing. *See* BRAZING; MACHINING; METAL CASTING; POWDER METALLURGY; WELDING AND CUTTING OF MATERIALS.

Mel M. Schwartz

Ceramic-matrix composites. In general, ceramics are brittle engineering materials with limited reliability. Brittleness is connected with the structure and chemical bonding of the main constituents, and reliability is connected with the stochastic character of main phases and defect distribution within the polycrystalline ceramic body. In spite of the generally high strength, hardness, and chemical and shape stability of ceramics, these two negative properties disqualify their wider application in industry. Ceramic-matrix composites are designed as materials with higher fracture resistance (less brittleness), higher reliability, and in particular cases higher strength compared to monolithic ceramics. These attributes are required for high technologies, especially in the aircraft, automotive, engineering, and energy industries. *See* BRITTLENESS; CERAMICS.

Ceramic composites are materials with at least two constituents, the ceramic-matrix phase and reinforcing-toughening filaments. The filaments cover a wide range of dimensions, from nanoinclusions, through micro-whiskers, to fibers that are several centimeters to a few meters long.

Ceramic nanocomposites. These composites have at least one of the main constituents at the nanometer scale (from one to several hundred nanometers). Typical examples of such ceramic nanocomposites are silicon carbide–silicon nitride (SiC/Si_3N_4) and silicon carbide–alumina (SiC/Al_2O_3). Benefits from the design of these materials are better mechanical properties at room temperature or high temperature, as well as improved electric and magnetic properties. *See* NANOSTRUCTURE.

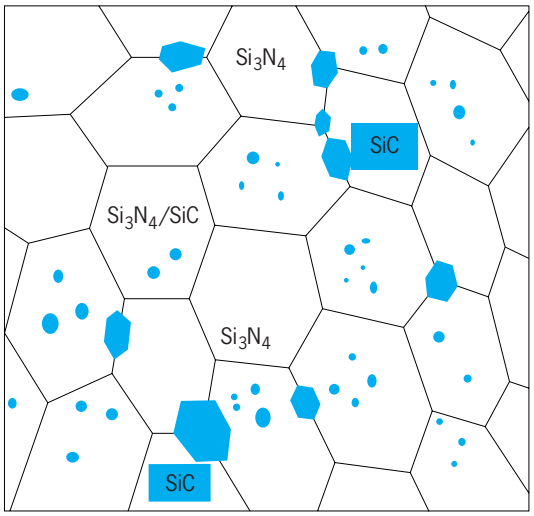

Fig. 3. Schematic of a SiC/Si_3N_4 nanocomposite. This microstructure consists of silicon carbide inclusions within silicon nitride grains, and silicon carbide grains located at the grain boundaries.

Examples of some properties of nanocomposites are as follows: A SiC/Si_3N_4 nanocomposite containing 20 vol % SiC has a bending strength greater than 1 GPa (10000 atm) up to 1400°C (2552°F), and a fracture toughness of 7 $MPa \cdot m^{1/2}$. Silver–ferrite oxide nanocomposite (Ag/Fe_2O_3) exhibits a superparamagnetic state at temperatures greater than 100 K (−173°C; −280°F).

Based on distribution of nanograins within the matrix, ceramic nanocomposites can be formally divided into intra type, inter type, intra/inter type, and nano/nano type. The SiC/Si_3N_4 nanocomposite (**Fig. 3**) can be considered an intra/inter type, because of distribution of SiC grains at grain boundaries as well as within the Si_3N_4 grains.

Whisker/platelet-reinforced composites. This ceramic composite contains whiskers or platelets. The whiskers are randomly distributed within the composite matrix. Silicon carbide or silicon nitride whiskers are usually embedded in a silicon carbide, silicon nitride, or alumina matrix. Improvement of the mechanical properties of these composites is reached by dissipation of the crack tip energy on a whisker or platelet. Whisker length varies between several micrometers and hundreds of micrometers. A typical parameter that determines a whisker shape is the aspect ratio, or the length-to-thickness ratio. Platelets are single crystals of flake shape. Their aspect ratio is the diameter-to-thickness ratio. Despite lower effectivity of platelets in dissipating crack tip energy compared to the whiskers, their application is forced due to the environmental unacceptability of whiskers. The bending strength of Si_3N_4/Si_3N_4 ceramics is 0.8–1 GPa (8000–10,000 atm), and fracture toughness is 6 $MPa \cdot m^{1/2}$ at room temperature.

Ceramic fiber-matrix composites. In principle, these composites are similar to fiber-matrix composites. The difference is in the composition of fibers, the matrices, and the processing routes.

The typical examples of fibers are the silicon carbide or alumina fibers. Fibers are polycrystalline materials with high tensile strength. Silicon carbide fiber tensile strength for Hi-Nicalon-S fiber is greater than 3 GPa (30,000 atm).

Silicon carbide, silicon nitride, and alumina ceramics are typically used as the matrices. Carbon as a constituent of these composites is also used either in the form of fibers or as a matrix.

There are several processes of embedding the continuous fibers into the ceramic matrix: polymer impregnation and pyrolysis, slurry impregnation and hot pressing, and chemical vapor infiltration.

These materials exhibit a high tensile strength when the fibers are oriented in the direction of the tensile force, and a very high work of fracture, which means a high fracture toughness. For example, the tensile strength of a SiC/Si_3N_4 composite is 0.5 GPa (5000 atm), and the fracture toughness is 26.5 $MPa \cdot m^{1/2}$.

Laminate/layered ceramic composites. These ceramic composites consist of two or more different ceramic sheets which are repeated several times through the ceramic body. These materials are usually produced by tape casting in order to built the ceramic "green" body, which is densified by hot pressing or gas pressure sintering. The properties of this type of ceramics are highly anisotropic in different directions, parallel and perpendicular to the layer area. The layers can differ by composition or by microstructure. Mechanical properties of silicon nitride–based layered composite are in particular cases exceptional; for example, a bending strength of 1.2 GPa (12,000 atm) and a fracture toughness of 10 $MPa \cdot m^{1/2}$ are reported. *See* SINTERING.

Layered materials offer a possibility to design materials with multifunctions, exhibiting excellent mechanical properties and improved electrical, thermal, or magnetic properties. Pavol Šajgalík

Applications

The use of fiber-reinforced materials in engineering applications has grown rapidly. Selection of composites rather than monolithic materials is dictated by the choice of properties. The high values of specific stiffness and specific strength may be the determining factor, but in some applications wear resistance or strength retention at elevated temperatures is more important. A composite must be selected by more than one criterion, although one may dominate.

Components fabricated from advanced organic-matrix–fiber-reinforced composites are used extensively on commercial aircraft as well as for military transports, fighters, and bombers. The propulsion system, which includes engines and fuel, makes up a significant fraction of aircraft weight (frequently 50%) and must provide a good thrust-to-weight ratio and efficient fuel consumption. The primary means of improving engine efficiency are to take advantage of the high specific stiffness and strength of composites for weight reduction, especially in rotating components, where material density directly affects both stress levels and critical dynamic characteristics, such as natural frequency and flutter speed.

Composites consisting of resin matrices reinforced with discontinuous glass fibers and continuous-glass-fiber mats are widely used in truck and automobile components bearing light loads, such as interior and exterior panels, pistons for diesel engines, drive shafts, rotors, brakes, leaf springs, wheels, and clutch plates.

The excellent electrical insulation, formability, and low cost of glass-fiber-reinforced plastics have led to their widespread use in electrical and electronic applications ranging from motors and generators to antennas and printed circuit boards.

Composites are also used for leisure and sporting products such as the frames of rackets, fishing rods, skis, golf club shafts, archery bows and arrows, sailboats, racing cars, and bicycles.

Advanced composites are used in a variety of other applications, including cutting tools for machining of superalloys and cast iron and laser mirrors for outer-space applications. They have made it possible to mimic the properties of human bone, leading to development of biocompatible prostheses for bone replacements and joint implants. In engineering, composites are used as replacements for fiber-reinforced cements and cables for suspension bridges. *See* MATERIALS SCIENCE AND ENGINEERING. Mel M. Schwartz

Bibliography. N. P. Cheremisinoff and P. N. Cheremisinoff (eds.), *Handbook of Advanced Materials Testing*, Marcel Dekker, New York, 1995; M. Grayson (ed.), *Encyclopedia of Composite Materials and Components*, Wiley, New York, 1983; G. Lubin (ed.), *Handbook of Composites*, 1982; K. Niihara, New design concept of structural ceramics: Ceramic nanocomposites, *J. Ceramic Soc. Jap.*, 99:974–982, 1991; T. Richardson, *A Design Guide*, 1987; T. Rienhart, *Engineered Materials Handbook*, vol. 1: *Composites, Metals*, 1987; S. J. Schneider, Jr. (ed.), *Engineered Materials Handbook*, vol. 4: *Ceramics and Glasses*, ASM International, 1991; M. M. Schwartz, *Composite Materials Handbook*, 2d ed., 1992; J. W. Weeton (ed.), *Engineer's Guide to Composite Materials*, 1986.

Composition board

A wood product in which the grain structure of the original wood is drastically altered. Composition board may be divided into several types. When wood serves as the raw material for chemical processing, the resultant product may be insulation board, hardboard, or other pulp product. When the wood is broken down only by mechanical means, the resultant product is particle board. Because composition board can use waste products of established wood industries and because there is a need to find marketable uses for young trees, manufacture of composition board is one of the most rapidly developing portions of the wood industry. *See* PAPER.

Fiberboard. One form of fiberboard is produced by loading a batch of wood chips into a chamber which is then heated and pressurized by steam. After about 2 min, the 1000-lb/in.^{2} (6.9-megapascal) pressure is abruptly released to hydrolyze and fluff the chips into a brown fiber. The fiber is refined, washed, and felted into a mat on a wire conveyor so that some of the water can drain out, and then the mat is cut to length for loading on a screen into a press. At controlled temperature in the press, the lignin rebonds the material while water is driven off as steam through the screen. The finished reconstituted wood product is a hard isotropic board as a consequence of the felting of the fibers and the ligneous bonding, possibly augmented by synthetic adhesive.

Alternatively, a similar board is produced by a continuous process. A screw feed delivers wood chips from a hopper to a steam preheater where the chips partially hydrolyze in the vicinity of 150 lb/in.^{2} (0.11 MPa). The hot chips pass between grinding disks to discharge as pulp, which is then formed into sheets essentially as described above. The wood chips may also be processed entirely by grinding. A further variation is to deliver the pulp slurry into a deckle box, in which case most of the water is removed by suction applied below the box before the mat is compressed into the finished sheet.

Instead of conveying the wet fiber or pulp to the mat, it can be dried sufficiently to be conveyed by air; it is then introduced between chrome-plated cauls and pressed. Because of the low-moisture content, no screen is needed; the finished sheet is smooth on both sides. Most new plants use variants of the dry process.

In any of the foregoing methods synthetic resin may be added as a binder before the board is formed. After the board is formed, it may be impregnated with drying oils and heated in a kiln at 300°F (150°C) until the oils are completely polymerized to produced tempered board. *See* DRYING OIL; RESIN.

If insulating board is required instead of hardboard, the material is less compacted, the degree of compaction being described by the specific gravity of the finished board.

Particle board. When formed from wood particles that retain their woody structure, the product is termed particle board. Properties of such boards depend on the size and orientation of the particles, which may be dimensioned flakes, random-sized shavings, or splinters. After the particular type of particles are produced, they are screened to remove fines and to return oversizes for further reduction. Graded particles are dried, mixed with synthetic adhesive and other additives such as preservatives, and delivered to the board-forming machine.

Forming machines are basically of two types. In one type mixed materials are metered by weight or volume onto an open flat tray, prepressed at room temperature, and finally consolidated under heat at high pressure to the finished dimension. Whether carried out as a batch or a continuous process, this flat-press method tends to orient the particles with their long dimensions approximately parallel to the face of the board. In the second type of machine the mixed materials are extruded under pressure between heated platens. Thick boards made this way may have hollow cores produced by rods fixed in the aperture ahead of the platens. Particles in extruded boards are randomly oriented. Boards formed by either method may be layered, with different-sized particles at the surfaces from those in the center, which are also usually coarser.

Characteristics. Hardboards produced by a continuous process can be formed to great length, although a finished length of 16 ft (5 m) is usual stock maximum. Standard thicknesses range from $^1/_8$ to $^5/_{\sqrt{16}}$ in. (3.2 to 7.9 mm); standard densities range from a specific gravity of 0.80 to 1.15, with boards up to 1.45 available if needed.

Tensile strength parallel to the face ranges from 1000 lb/in.^{2} (6.9 MPa) for untreated board to 5500 lb/in.^{2} (37.9 MPa) for tempered hardboard; perpendicular to the face it ranges from 125 to 400 lb/in.^{2} (0.86 to 2.76 MPa). Compression strength parallel to the face ranges from 2000 to 5300 lb/in.^{2} (13.8 to 36.5 MPa).

Lengths and widths of particle boards are similar to those of hardboards; thicknesses range from $^1/_4$ to 1 in. (6 to 25 mm); thicker sheets are most economically produced by laminating sheets of standard thicknesses. Properties of flat-pressed particle board vary more than do those of hardboard. They have low density: 0.4–1.2 specific gravity. Tensile strength parallel to their faces ranges from 500 to 5000 lb/in.^{2} (3.4 to 34.5 MPa) and perpendicular from 40 to 400 lb/in.^{2} (276 kPa to 2.76 MPa). Extruded particle board has low bending strength; it is used chiefly as the core of plywood-type panel. Development of adhesives specifically for composition boards has extended their utilitarian value, and variety of textures has increased their esthetic appeal. *See* WOOD PRODUCTS.

Frank H. Rockett

Compound (chemistry)

A substance composed of two or more elements that do not vary in composition from sample to sample, and that have fixed and definite physical properties, such as density and refractive index. The elements in compounds cannot be separated by simple physical or mechanical means, but only by chemical treatment. When compounds are formed from their elements, heat is generated or absorbed. These properties distinguish them from mixtures.

For example, if iron filings and sulfur powder are mixed together, the two elements can be separated, either by removing the iron filings with a magnet or by dissolving the sulfur in an appropriate solvent, and the individual particles of each element are distinguishable under a magnifying glass. If iron filings and sulfur are heated together, a chemical reaction takes place and a new substance is formed, iron sulfide, with properties totally different from iron or sulfur. *See* DENSITY; ELEMENT (CHEMISTRY).

Most chemical compounds are formed in fixed and definite proportions by weight from their elements, and they obey the laws of chemical combination. However, a number of solid compounds, known as nonstoichiometric compounds, exhibit departures from the law of definite proportions. *See* DEFINITE COMPOSITION, LAW OF; EQUIVALENT WEIGHT; INORGANIC CHEMISTRY; MULTIPLE PROPORTIONS, LAW OF; NONSTOICHIOMETRIC COMPOUNDS; ORGANIC CHEMISTRY.

Thomas C. Waddington

Compressible flow

Flow in which density changes are not negligible. Pressure changes normally occur throughout a fluid flow, and these pressure changes, in general, induce a change in the fluid density. In a compressible flow, the density changes that result from these pressure changes have a significant influence on the flow. The changes in the flow that result from the density changes are often termed compressibility effects. All fluids are compressible. However, compressibility effects are more frequently encountered in gas flows than in liquid flows.

Compressibility. The compressibility of a fluid is defined by Eq. (1), where ρ is the density, and $\Delta\rho$ is the

$$\tau = \frac{\Delta\rho/\rho}{\Delta p} \tag{1}$$

density change that results from a pressure change of Δp. From Eq. (1) it follows that the relative density change is given by Eq. (2). If the product $\tau\,\Delta p$ is very

$$\frac{\Delta\rho}{\rho} = \tau\,\Delta p \tag{2}$$

small, then the change in density, $\Delta\rho$, will be negligible and the effect of the density changes on the flow can be neglected. Such flows that are effectively at constant density are termed incompressible flows. For a flow in which the product $\tau\,\Delta p$ is moderate or large, the changes in density in the flow cannot, in general, be neglected. Such flows in which the density changes are significant are termed compressible flows. As a rough guide, if the density changes by less than 5% in the flow, that is, $\Delta\rho/\rho < 0.05$, the flow can be treated as incompressible. The speed of sound in the fluid, a, is approximately given by Eq. (3). Using this in Eq. (1) gives Eq. (4). The larger the

$$a^2 = \frac{\Delta p}{\Delta\rho} \tag{3}$$

$$\tau = \frac{1}{\rho a^2} \tag{4}$$

value of ρa^2, the less compressible is the fluid. Thus a liquid, which has a large value of ρa^2, will have a much smaller compressibility than a gas, which has a small value of ρa^2. For example, water has a value of ρa^2 that is about 15,000 times greater than that of air and so is much less compressible than air. *See* SOUND.

Mach number. An important dimensionless parameter in compressible flows is the Mach number, M. This is defined by Eq. (5), where a is the speed of

$$M = \frac{V}{a} \tag{5}$$

sound and V is the velocity of the flow. For a gas, the speed of sound is given by Eq. (6), where R is the

$$a = \sqrt{kRT} \tag{6}$$

gas constant, $k = c_p/c_v$, c_p and c_v being the specific heats at constant pressure and constant volume respectively, and T is the temperature. If $M < 0.3$ in a flow, the density changes in the flow will usually be negligible; that is, the flow can be treated as incompressible. Compressible flows are, therefore, as a rough guide, associated with Mach numbers greater than 0.3.

When $M < 1$, the flow is said to be subsonic; when $M = 1$, the flow is said to be sonic; when M varies from slightly below 1 to slightly above 1, the flow is said to be transonic; and if $M > 1$, the flow is said to be supersonic. When the Mach number is very high, this usually being taken to mean $M > 5$, the flow is said to be hypersonic.

Properties. Compressible flows can have features that do not occur in low-speed flows. For example, shock waves and expansion waves can occur in supersonic flows (*see* **illus.**). Another important phenomenon that can occur due to compressibility is choking, where the mass flow rate through a duct system may be limited as a result of the Mach number being equal to 1 at some point in the flow. *See* CHOKED FLOW; SHOCK WAVE; SONIC BOOM.

Another effect of compressibility is associated with the acceleration of a gas flow through a duct. In incompressible flow, an increase in velocity is associated with a decrease in the cross-sectional area of the duct, this in fact being true as long as $M < 1$. However, when $M > 1$, that is, when the flow is supersonic, the opposite is true; that is, an increase in the velocity is associated with an increase in the cross-sectional area. Therefore, in order to accelerate

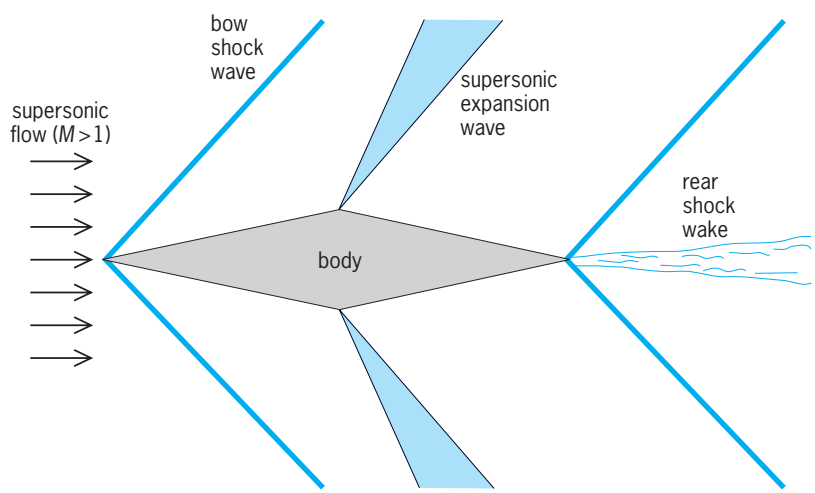

Waves in supersonic flow over a double-wedge body.

a gas flow from subsonic to supersonic velocities in a duct, it is necessary first to decrease the area and then, once the Mach number has reached 1, to increase the area, that is, to use a so-called convergent-divergent nozzle. An example is the nozzle fitted to a rocket engine. *See* FLUID FLOW; MACH NUMBER; NOZZLE; SUPERSONIC FLOW. Patrick H. Oosthuizen

Bibliography. J. D. Anderson, Jr., *Modern Compressible Flow: With Historical Perspective*, 3d ed., McGraw-Hill, 2003; P. H. Oosthuizen and W. E. Carscallen, *Compressible Fluid Flow*, McGraw-Hill, 1997.

Compression ratio

In a cylinder, the piston displacement plus clearance volume, divided by the clearance volume. This is the nominal compression ratio determined by cylinder geometry alone. In practice, the actual compression ratio is appreciably less than the nominal value because the volumetric efficiency of an unsupercharged engine is less than 100%, partly because of late intake valve closing. In spark ignition engines the allowable compression ratio is limited by incipient knock at wide-open throttle. Knock in turn depends on the molecular structure of the fuel and on such engine features as the temperature of the combustible mixture prior to ignition, the geometry and size of the combustion space, and the ignition timing. For example, isooctane, benzene, and alcohol can be burned at much higher compression ratios than *n*-heptane. In compression ignition engines critical compression ratio is that necessary to ignite the fuel and depends on fuel and cylinder geometry. *See* COMBUSTION CHAMBER; DIESEL ENGINE; INTERNAL COMBUSTION ENGINE; SPARK KNOCK; VOLUMETRIC EFFICIENCY. Neil MacCoull

Compressor

A machine that increases the pressure of a gas or vapor (typically air), or mixture of gases and vapors. The pressure of the fluid is increased by reducing the fluid specific volume during passage of the fluid through the compressor. When compared with centrifugal or axial-flow fans on the basis of discharge pressure, compressors are generally classed as high-pressure and fans as low-pressure machines.

Compressors are used to increase the pressure of a wide variety of gases and vapors for a multitude of purposes (**Fig. 1**). A common application is the air compressor used to supply high-pressure air for conveying, paint spraying, tire inflating, cleaning, pneumatic tools, and rock drills. The refrigeration compressor is used to compress the gas formed in the evaporator. Other applications of compressors include chemical processing, gas transmission, gas turbines, and construction. *See* GAS TURBINE; REFRIGERATION.

Characteristics. Compressor displacement is the volume displaced by the compressing element per unit of time and is usually expressed in cubic feet per minute (cfm). Where the fluid being compressed flows in series through more than one separate compressing element (as a cylinder), the displacement of the compressor equals that of the first element. Compressor capacity is the actual quantity of fluid compressed and delivered, expressed in cubic feet per minute at the conditions of total temperature, total pressure, and composition prevailing at the compressor inlet. The capacity is always expressed in terms of air or gas at intake (ambient) conditions rather than in terms of arbitrarily selected standard conditions.

Air compressors often have their displacement and capacity expressed in terms of free air. Free air is air at atmospheric conditions at any specific location. Since the altitude, barometer, and temperature may vary from one location to another, this term does not mean air under uniform or standard conditions. Standard air is at 68°F (20°C), 14.7 lb/in.^{2} (101.3 kilopascals absolute pressure), and a relative humidity of 36%. Gas industries usually consider 60°F (15.6°C) air as standard.

Types. Compressors can be classified as reciprocating, rotary, jet, centrifugal, or axial-flow, depending on the mechanical means used to produce compression of the fluid, or as positive-displacement or dynamic-type, depending on how the mechanical elements act on the fluid to be compressed. Positive-displacement compressors confine successive volumes of fluid within a closed space in which the pressure of the fluid is increased as the volume of the closed space is decreased. Dynamic-type compressors use rotating vanes or impellers to impart velocity and pressure to the fluid.

Reciprocating compressors are positive-displacement types having one or more cylinders, each fitted with a piston driven by a crankshaft through a connecting rod. Each cylinder also has intake and discharge valves, and a means for cooling the mechanical parts (**Fig. 2**). Fluid is drawn into the cylinder during the suction stroke. At the end

Fig. 1. Mobile air compressor unit which is driven by engine in the forward compartment. (*Ingersoll-Rand*)

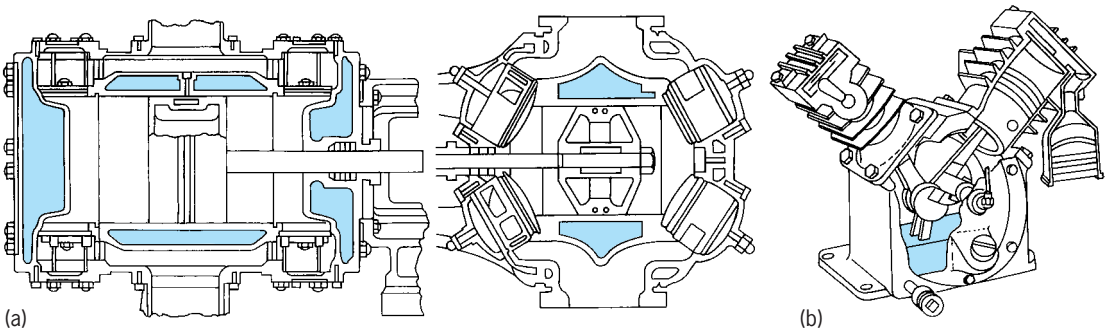

Fig. 2. Compressor cylinders, (*a*) water-cooled and (*b*) air-cooled.

of the suction stroke, motion of the piston reverses and the fluid is compressed and expelled during the discharge stroke. When only one end of the piston acts on the fluid, the compressor is termed a single-acting unit. When both ends of the piston act on the fluid, the compressor is double-acting. The double-acting compressor discharges about twice as much fluid per cylinder per cycle as the single-acting.

Single-stage compressors raise the fluid pressure from inlet to discharge on each working stroke of the piston in each cylinder. Two-stage compressors use one cylinder to compress the fluid to an intermediate pressure and another cylinder to raise it to final discharge pressure. When more than two stages are used, the compressor is called a multistage unit.

Some reciprocating compressors provide variable displacement. Instead of use of a crankshaft, the length of the piston stroke is controlled by a rotating swash plate, or wobble plate, that can change angle while the compressor is running. This design allows the compressor to operate continuously while providing only the displacement needed.

Vertical and horizontal compressors may be single-cylinder or multicylinder units. The angle type is multicylinder with one or more horizontal and vertical compressing elements (**Fig. 3**). Single-frame (straight-line) units are horizontal or vertical, double-acting, with one or more cylinders in line with a single frame having one crank throw, connecting rod, and crosshead. The V or Y type is a two-cylinder, vertical, double-acting machine with cylinders usually at a 45° angle with the vertical. A single crank is used. Semiradial compressors are similar to V or Y type, but have horizontal double-acting cylinders on each side. Duplex compressors have cylinders on two parallel frames connected through a common crankshaft. In duplex-tandem steam-driven units, steam cylinders are in line with air cylinders. Duplex four-cornered steam-driven compressors have one or more compressing cylinders on each end of the frame and one or more steam cylinders on the opposite end. In four-cornered motor-driven units, the motor is on a shaft between compressor frames.

Reciprocating compressors are built to handle fluid capacities ranging to 100,000 ft^3/min (2800 m^3/min); pressures range to over 35,000 lb/in.^{2} (240 megapascals). Special units can be built for larger capacities or higher pressures. Water is the usual coolant for cylinders, intercoolers, and aftercoolers, but other liquids, including refrigerants, may also be used.

Compressor thermodynamics. Compression efficiency in any compressor is compared against two theoretical standards—isothermal and adiabatic. Neither occurs in an actual compressor because of unavoidable losses. Isothermal compression has perfect cooling—air remains at inlet temperature while being compressed. The work input to the compressor is the least possible. Adiabatic compression has no cooling; the temperature rises steadily during compression. Discharge pressure is reached sooner than with isothermal compression. Since air pressure is higher during every part of the piston's stroke, more work input is needed. *See* THERMODYNAMIC PROCESSES.

If compression is divided into two or more steps or stages, air can be cooled between stages. This intercooling brings the actual compression efficiency closer to the isothermal standard.

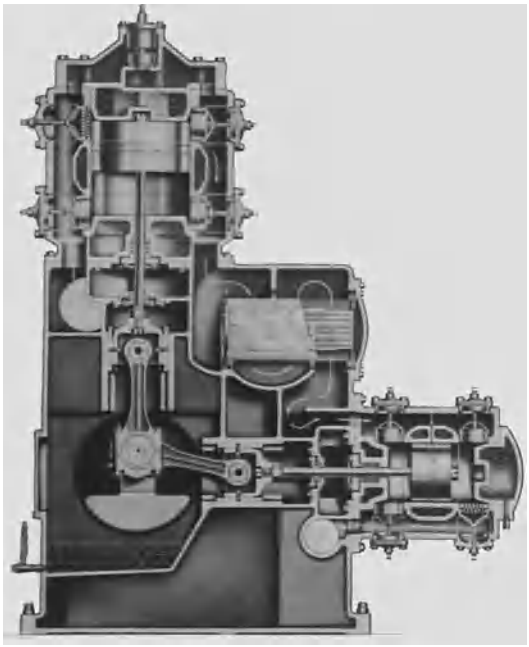

Fig. 3. Cross section of stationary angle-type compressor with first stage vertical and second stage horizontal.

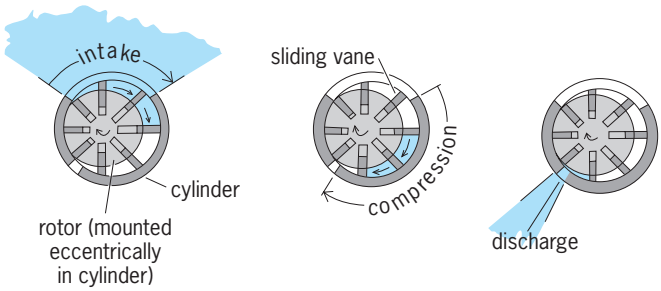

Fig. 4. Operation of sliding-vane rotary compressor.

Water vapor in air entering the compressor leaves as superheated vapor, because its temperature is in excess of that corresponding to its pressure. It can be converted to water only by cooling to a temperature below the saturation temperature corresponding to its pressure. Immediate cooling of air after it leaves the compressor proves best, because this prevents vapor from reaching the distribution system. Either air- or water-cooled heat exchangers, called aftercoolers, are used for this.

The usual single-stage reciprocating compressor is built for pressures to about 150 lb/in.^{2} (1.0 MPa); two-stage types for up to about 500 lb/in.^{2} (3.4 MPa); three-stage for up to about 2500 lb/in.^{2} (17 MPa); four- and five-stage for up to 15,000 lb/in.^{2} (100 MPa), and higher.

Rotary compressors. Other forms of positive-displacement compressor are the rotary types.

Sliding-vane type. In the sliding-vane rotary compressor, fluid is trapped between vanes as the rotor passes the inlet opening (**Fig. 4**). Further rotation of the rotor reduces the volume of the space in which the fluid is trapped. Fluid pressure rises until the discharge port is reached, when discharge occurs.

Depending on design, the compressor may be cooled with atmospheric air, oil, or water. For low pressures one stage of compression is used; for higher pressures, two stages. Capacities range to about 5000 ft^3/min (140 m^3/min). Some compressors of this type have rings around the vanes to keep them from bearing on cylinder walls. Others have vanes bearing against the cylinder.

Lobe type. In the lobed rotary compressor, fluid is trapped between two or more rotors held in fixed relationship to each other. Rotation of the impellers reduces the volume in which the fluid is trapped, producing a pressure rise. Fluid is discharged when the rotors pass the outlet port. Either two or three rotors are used. Capacities range about 5–50,000 ft^3/min (0.1–1400 m^3/min). Pressures above about 15 lb/in.^{2} (100 kPa) are obtained by operating two or more lobe-type compressors in series. Rotors are straight or slightly twisted.

Liquid-piston type. In the liquid piston rotary compressor, a round multiblade rotor revolves in an elliptical casing partly filled with liquid, usually water. When the rotor turns, it carries liquid around with it, the blades forming a series of buckets. Because the liquid follows the casing contours, it alternately leaves and returns to the space between blades (twice each revolution). As the liquid leaves the bucket, the fluid to be compressed is drawn in. When the liquid returns, it compresses the fluid to discharge pressure.

Liquid-piston compressors handle up to about 5000 ft^3/min (40 m^3/min). Single-stage units can develop pressures to about 75 lb/in.^{2} (520 kPa); multistage designs are used for higher pressures. Water or almost any other low-viscosity liquid serves as the compressant. For exacting services, the compressor may be sealed with chilled water to prevent condensation in lines.

Radial-flow units. Dynamic-type centrifugal compressors use rotating elements to accelerate the fluid radially. By diffusing action, velocity is converted to static pressure. Thus, the static pressure is higher in the enlarged section. Centrifugal compressors usually take in fluid at the impeller eye (or central inlet of the circular impeller) and accelerate it radially outward. Some static-pressure rise occurs in the impeller, but most of the pressure rise occurs in the diffuser section of the casing, where velocity is converted to static pressure. Each impeller-diffuser set is a stage of the compressor. Centrifugal compressors are built with from 1 to 12 or more stages, depending on the final pressure desired and the volume of fluid to be handled. The pressure ratio, or compression ratio, of a compressor is the ratio of the absolute discharge pressure to the absolute inlet pressure. *See* BERNOULLI'S THEOREM; DIFFUSER.

In the typical compressor, centrifugal action of the impeller produces a pressure rise and a large increase in air velocity. In the diffuser, velocity energy is converted to static pressure.

Pressure delivered by a centrifugal compressor is practically constant over a relatively wide range of capacities. Pumping limit, also called surge point or pulsation point, is the lower limit of stable operation. Percentage stability is 100 minus the pumping limit in percent of design capacity.

Multistage centrifugal compressors handle from 500 to more than 150,000 ft^3/min (14 to 4300 m^3/min) at pressures as high as 5000 lb/in.^{2} (35 MPa), but are limited to compression ratios in the order of 10.

Axial-flow units. Compressors that accelerate the fluid in a direction generally parallel to the rotating shaft consist of pairs of moving and stationary blade-rows, each pair forming a stage. The pressure rise per stage is small, compared with a radial-flow unit. Therefore, the usual axial-flow compressor has more stages than a centrifugal compressor working through the same pressure range. Single-stage axial compressors have capacities from a few to more than 100,000 ft^3/min (2800 m^3/min) at pressures from less than 1 to several pounds per square inch. Multistage axial-flow compressors compress air to 150 lb/in.^{2} (1 MPa) or more. Some special machines handle over 2,000,000 ft^3/min (60,000 m^3/min). Pressure rise per stage is generally relatively small, so units for higher pressures frequently have a considerable number of stages—20 or more.

While centrifugal machines deliver practically constant pressure over a considerable range of capacities, axials have a substantially constant delivery at

variable pressures. In general, centrifugals have a wider stable operating range than axials. Because of their more or less straight-through flow, axials tend to be smaller in diameter than centrifugals and to be longer. Efficiency of axials usually runs slightly higher.

To prevent surging at extremely low loads, large-capacity axials are sometimes fitted with a blowoff system that discharges excess air to the atmosphere. Then there is always enough air passing through the machine to keep it in its stable range.

Because of difficulty in accurately predicting the performance curves of centrifugal and axial compressors, only one capacity and one discharge-pressure rating, together with corresponding power input, are normally guaranteed. Shape of the curve may be indicated but is never guaranteed. *See* FAN; PUMPING MACHINERY. Tyler G. Hicks; Donald L. Anglin

Bibliography. Business Trend Analysts, *Pumps and Compressors,* 1982; Compressed Gas Association Staff, *Handbook of Compressed Gases*, 4th ed., 1999; *Compressor Handbook for the Hydrocarbon Processing Industries*, 1979; Ingersoll-Rand, *Compressed Air and Gas Data*, 1980.

Compton effect

The increase in wavelength of scattered electromagnetic radiation that is determined by the angle of scattering when gamma rays or energetic x-rays are scattered by electrons. When the scattering electron is assumed to be unbound and at rest, analysis leads to a wavelength increase approximated by Eq. (1).

$$\Delta\lambda = \lambda_2 - \lambda_1 = \frac{h}{m_0 c}(1 - \cos\phi) \qquad (1)$$

Here, λ_1 is the wavelength of the incident radiation, λ_2 is the wavelength of the radiation scattered at the angle ϕ, h is Planck's constant, m_0 is the rest mass of the electron, c is the speed of light, and ϕ is the angle by which the scattered radiation deviates from the direction of the incident radiation.

The quantity $h/m_0c = 2.426310 \times 10^{-12}$ m is called the Compton wavelength of the electron. The Compton wavelengths of other particles are defined in terms of their rest masses in an analogous manner. Thus, for example, the Compton wavelength of the proton is 1.321410×10^{-15} m.

There is a near consensus that the Compton effect provides the most persuasive evidence from any physical phenomenon in support of the photon description of electromagnetic radiation. The evidence provided by the angular dependence of the wavelength shift is not in itself sufficient to establish the reality of photons. However, when supplemented by the experimental evidence that a single scattering electron suffices to conserve both energy and momentum in a binary collision with a photon, it provides compelling support for the photon picture.

A. H. Compton's measurements, carried out in the era of the old quantum theory, challenged the validity of J. J. Thomson's theory of coherent scattering by electrons, which was based on the Maxwell equations of electromagnetism and the wavelike nature of x-rays. The wavelike properties of x-rays were established by the use of crystalline solids as diffraction gratings, leading to the development of wavelength-dispersive x-ray spectrometry. Compton's mastery of this new technology led him to the proof that x-rays also display particlelike properties. *See* MAXWELL'S EQUATIONS; X-RAY DIFFRACTION; X-RAY SPECTROMETRY.

The modern picture of coherent scattering differs from that of Thomson. His calculated differential cross section for unpolarized radiation is given by Eq. (2). Here $d\Omega$ is the solid scattering angle, and r_e is

$$d\sigma = \frac{r_e^2}{2}(1 + \cos^2\phi)d\Omega \qquad (2)$$

the classical electron radius, equal to e^2/m_0c^2, where e is the charge on the electron. This result must be multiplied by the square of an atomic form factor, $f(\phi)$, because the electrons in an atom scatter in a cooperative manner. The resulting Rayleigh cross section depends very strongly on both the atomic number Z and the energy E. Data from low-Z targets such as graphite have the simple form that leads to the Compton analysis. Higher-Z targets such as iron or copper produce data dominated by Rayleigh scattering, showing little shift in wavelength. For example, markedly different energy spectra are obtained when

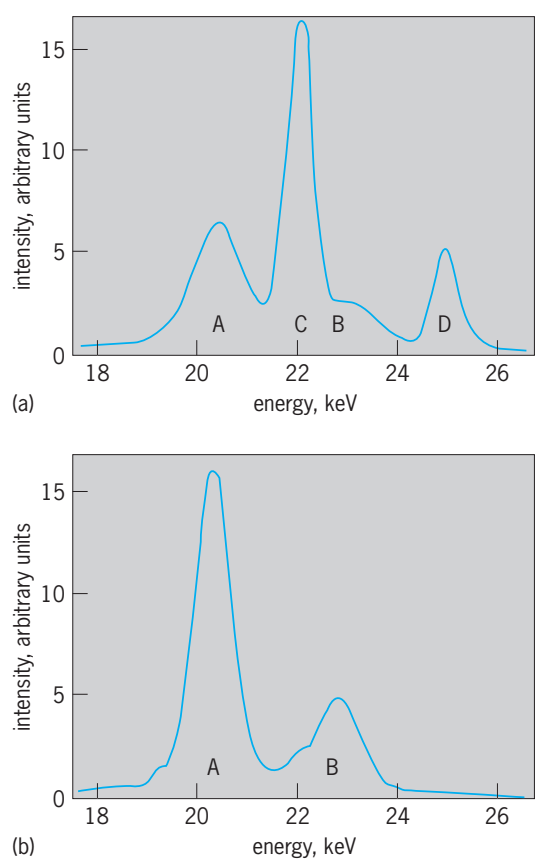

Fig. 1. Energy spectra of K x-rays of silver after scattering at about 170° from targets of (*a*) copper and (*b*) polyethylene (CH_2). The peaks labeled A and B are from Compton scattering; those labeled C and D are from Rayleigh scattering.

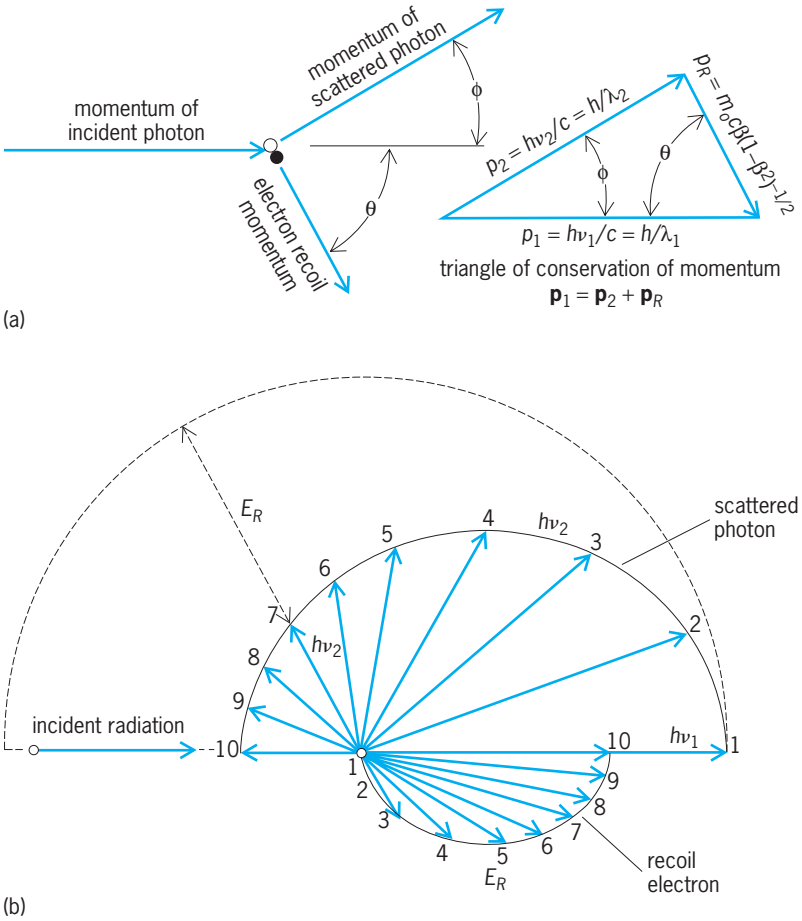

Fig. 2. Geometry of Compton scattering. (*a*) Diagrams for applying energy and momentum conservation to the collision of a photon with an electron, initially at rest. Symbols defined in text. (*b*) Relation of energy and direction of scattered photon, $h\nu_2$, and of recoil electron, E_R, in the special case that initial photon energy, $h\nu_1 = m_0c^2$. Photon and electron energy and direction are plotted for 10 photon scattering angles varying from 0 to 180°. Paired arrows mean, for example, that when the secondary photon goes in direction 5 in the upper half of the diagram, recoil electron follows arrow 5 in the lower half. (*After B. R. Wheaton, The Tiger and the Shark, Cambridge University Press, 1983*)

the K x-rays of silver (with energy of 22.1 keV) are backscattered from polyethylene (CH_2) and copper (**Fig. 1**). These spectral differences, easily revealed by a modern germanium detector, are hidden and unsuspected causes of inconsistency if a detector of low resolving power is used. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

Theory. The essential break with a classical theory of scattering is to adopt a model pioneered by A. Einstein, in which photons of frequency ν travel as discrete quanta with energy $h\nu$ and linear momentum $h\nu/c$. Compton's analysis applies the photon concept to a collision with an unbound electron at rest. The photon transfers some energy and momentum to the electron so the scattered photon must have less energy, $h\nu_2$, than the primary, $h\nu_1$. Since $\nu = c/\lambda$ for a wave of speed c, the wavelength of the scattered photon, λ_2, must exceed the initial wavelength λ_1. *See* PHOTOEMISSION; PHOTON; QUANTUM MECHANICS.

In an elastic collision, both the energy and the linear momentum must be conserved. Application of these two principles to determine the wavelength shift expected for a given scattering angle leads to what is now known as the Compton equation, Eq. (1). *See* COLLISION (PHYSICS); CONSERVATION LAWS (PHYSICS); CONSERVATION OF ENERGY; CONSERVATION OF MOMENTUM.

Scattering process. Compton's analysis applies to the scattering process for a collision between an incident photon of linear momentum p_1 and wavelength λ_1 and a resting, unbound electron (**Fig. 2*a***). After the collision, the scattered photon moves at an angle ϕ and the electron recoils at an angle θ, both angles being with respect to the direction of the incident photon.

Compton's decisive x-ray data were produced at such low energies that the recoil electron traveled at nonrelativistic speeds, but both Compton and P. Debye treated the problem generally. The principle of conservation of energy yields Eq. (3), where β is

$$h\nu_1 = h\nu_2 + m_0c^2\left(\frac{1}{\sqrt{1-\beta^2}} - 1\right) \qquad (3)$$

V/c and V is the speed of the recoil electron. Conservation of linear momentum yields two equations, Eq. (4), for the horizontal component and Eq. (5)

$$\frac{h\nu_1}{c} = \frac{h\nu_2}{c}\cos\phi + \frac{m_0V}{\sqrt{1-\beta^2}}\cos\theta \qquad (4)$$

$$0 = \frac{h\nu_2}{c}\sin\phi - \frac{m_0V}{\sqrt{1-\beta^2}}\sin\theta \qquad (5)$$

for the vertical component. By squaring these equations and substituting $\nu = c/\lambda$, the electron velocity can be eliminated and Eq. (1) is obtained for the shift in wavelength. *See* RELATIVITY.

Many tests of the theory, particularly those using modern energy-dispersive detectors, are checked by reverting to frequency. Equation (1) can be rewritten to give the energy of the scattered photon, and can be recombined with Eq. (3) to obtain an expression for the recoil energy of the initially stationary electron. Following Debye, the nature of these solutions can be illustrated compactly for the important special case that $h\nu_1 = m_0c^2$ by plotting the energy and direction of both the scattered photon and the recoil electron for various scattering angles (Fig. 2*b*). The resulting diagram emphasizes that while the photons may scatter backward as well as forward, the recoil electron is always projected forward in a much narrower cone.

Compton profile. The assumption that the scattering electron is at rest in the laboratory frame of reference is not consistent with the real situation in experiments. The photons are scattered by electrons whose speeds are typically 1% of the speed of light for valence electrons and 10% for core electrons, causing substantial Doppler shifts. Analysis shows that a second term should be added to the right-hand side of Eq. (1) for the wavelength shift. This second term is proportional to the projection of the electron's initial momentum on the scattering vector. Accurate measurement of the complete

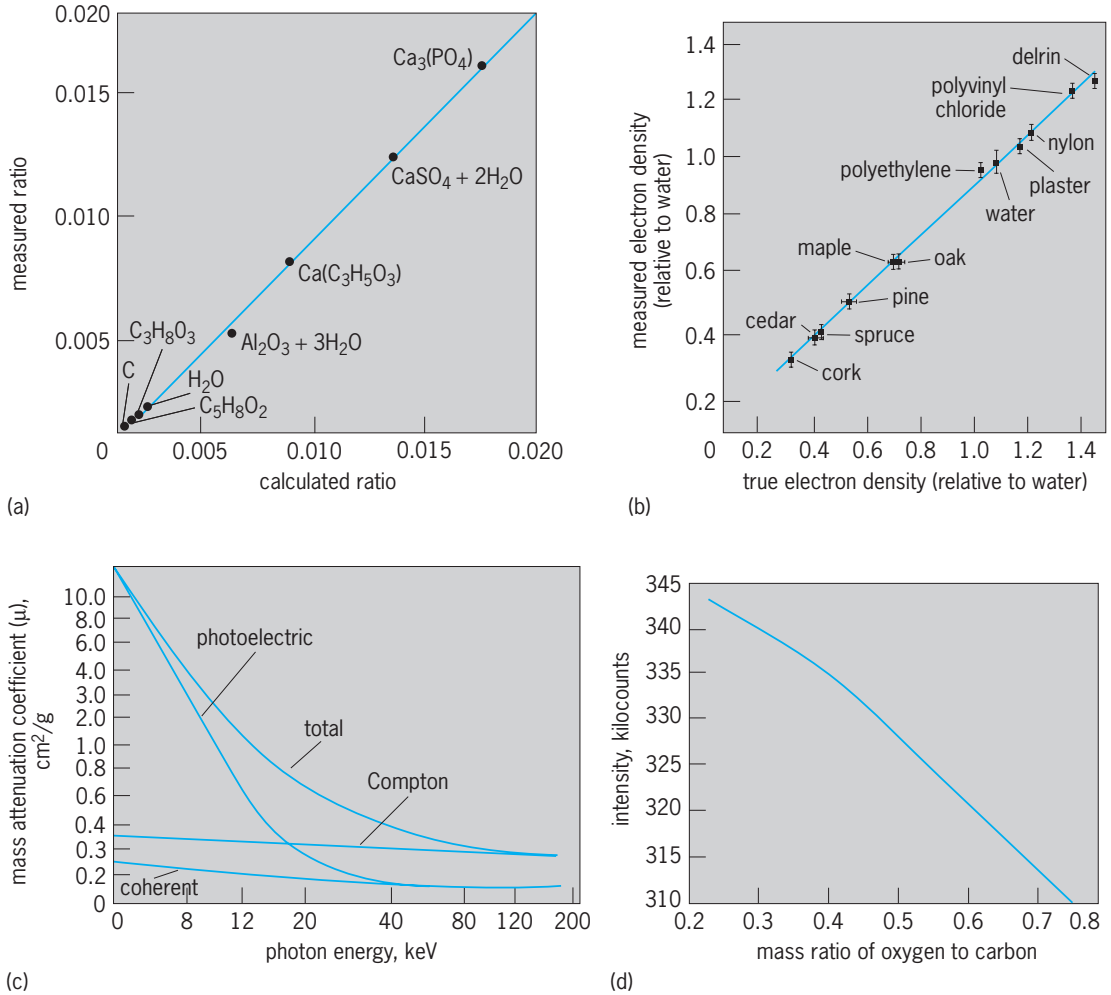

Fig. 3. Examples of nondestructive analysis of low-*Z* compounds using Compton scattering. (*a*) Comparison of measured and calculated Rayleigh-Compton ratios at an energy of 60 keV. Calculated ratio is based on assigning an effective atomic number to the compound. Diagonal line represents equality of measured and calculated ratios. (*b*) Comparison of Compton-measured electron density and true electron density of compounds for energies (662 keV) where Compton effect is totally dominant. Diagonal line represents equality of measured and true densities. (*c*) Energy dependence of the linear attenuation coefficient for the main interactions in soft tissue. Coefficients for total attenuation and components due to photoelectric effect, Compton effect, and coherent scattering are shown. (*d*) Dependence of backscatter intensity on the carbon-oxygen ratio in methanol-water mixtures (which simulate proportions of fat and lean meat) for the L x-rays of neptunium (∼ 17 keV). This dependence results from competition of Compton scattering with photoelectric absorption. (*Parts a, b, c after R. Cesareo, ed., Nuclear Analytical Techniques in Medicine, Elsevier Science, 1988*)

spectrum of scattered wavelengths gives a distribution of one cartesian component of the velocity of the scattering electrons. With a beryllium target, this technique was used to demonstrate that electrons follow Fermi-Dirac rather than Boltzmann statistics. If the incident radiation in this method is provided by ordinary x-ray generators, then the intensity of the scattered radiation is so low as to make most experiments impractical, because of the need to define the scattering angle by collimators. However, the advent of highly luminous synchrotron x-ray sources has made this an active research area. It permits accurate determination of directional variations of velocity spectra in single-crystal specimens and opens new research opportunities in the study of molecular bonds. *See* DOPPLER EFFECT; FERMI-DIRAC STATISTICS; SYNCHROTRON RADIATION.

Intensity distribution. Thomson's theory of x-ray scattering by electrons makes predictions not only of the wavelength shift but also of the intensity as a function of scattering angle. However, his expression depends on the distance of the photon from the scatterer, and Compton's model makes no prediction of this distance. That prediction requires the use of quantum mechanics in its relativistic form, due to P. A. M. Dirac. The result, known as the Klein-Nishina formula, contains the Thomson expression as a first term but goes on to include large frequency- and angle-dependent corrections. *See* RELATIVISTIC QUANTUM THEORY.

Applications. The Compton effect often plays a major nuisance role, producing most of the background in x-ray diffraction and in the field of gamma-ray spectroscopy, where interest focuses on producing energy spectra with narrow lines whose intensity and mean energy can be estimated accurately. Compton scattering of x-rays by materials of low atomic number is so dominant that it forces the

versatile technique of x-ray fluorescence to employ extremely thin targets, thus removing it from the list of nondestructive analytical techniques. However, the development of high-resolution silicon and germanium semiconductor radiation detectors in the 1960s opened new areas for applications of Compton scattering. Those related to the measurement of Compton profiles are discussed above. *See* JUNCTION DETECTOR; X-RAY FLUORESCENCE ANALYSIS.

Noninvasive analysis. Semiconductor detectors make it possible to measure the separate probabilities for Rayleigh and Compton scattering. An effective atomic number has been assigned to compounds that appears to successfully correlate theory with Rayleigh-Compton ratios (**Fig. 3***a*).

Average density can be measured by moving to higher energies where Compton scattering does not have to compete with Rayleigh scattering. At these energies, Compton scattering intensity has been successfully correlated with mass density (Fig. 3*b*). An appropriate application is the measurement of lung density in living organisms.

At low energies and low atomic numbers, the main competition for photon attenuation is between Compton scattering and photoelectric absorption (Fig. 3*c*). This competition can be used to analyze very low Z mixtures such as fat-lean ratios (Fig. 3*d*). *See* NONDESTRUCTIVE EVALUATION.

Gamma-ray astronomy. The ability to put large detectors in orbit above the Earth's atmosphere has created the field of gamma-ray astronomy. This field is now based largely on the data from the Compton Gamma-Ray Observatory, all of whose detectors make use of the Compton effect (although not exclusively). The information produced by this observatory on gamma-ray bursters may change the estimated age of the universe by a factor of 2. *See* GAMMA-RAY ASTRONOMY; GAMMA-RAY DETECTORS; GAMMA RAYS; X-RAYS. Innes K. Mackenzie

Bibliography. R. Cesareo (ed.), *Nuclear Analytical Techniques in Medicine*, 1988; N. A. Dyson, *X-Rays in Atomic and Nuclear Physics*, 2d ed., 1990; N. Gehrels et al., The Compton Gamma-Ray Observatory, *Sci. Amer.*, 269(6):68–77, December 1993; R. S. Holt, Compton imaging, *Endeavour*, 9:97–105, 1985; B. R. Wheaton, *The Tiger and the Shark*, 1983, reprint, 1991.

Computational chemistry

A branch of theoretical chemistry that uses a digital computer to model systems of chemical interest. In this discipline, the essential physics that determines molecular behavior is captured in mathematical models and programmed for digital computers. The models are imperfect but have proven remarkably useful.

The domain of computational chemistry is vast. In principle, almost any quantity that can be measured can be computed. In practice, programs are available that calculate properties of individual molecules, including geometries, charge distributions, and modes of vibration for both ground and excited electronic states. Changes in those properties can also be calculated as the computer transforms compounds from reactants to products during the simulation of a chemical reaction. Properties, such as absorption and emission spectra which involve the interaction of molecules with electromagnetic radiation, can also be computed. Where possible, computational procedures are validated by comparison with experiment, and the confidence gained from such comparisons often encourages application to properties that are at or beyond what can be measured experimentally. Computations can thus provide information about the geometries of short-lived high-energy species or molecules that may be only a glimmer in the eye of a creative chemist.

In addition to estimates of important quantities, computations also provide deep insight into the reasons that chemicals behave as they do. For example, electronic structure calculations provide information about bonding electrons that can explain structural preferences and foretell chemical reactivity. In favorable cases, this information can be used proactively to design molecules to user specifications or aid in the design of catalysts that promote the production of a desired compound.

Because of the increasing power of computers and the simultaneous development of well-tested and reliable theoretical methods, computational chemistry and experimental chemistry are becoming interdependent partners. In addition to the more traditional use of computational chemistry as a tool to explain experimental outcomes, computations are increasingly used to develop plausible hypotheses that are then tested in the laboratory. Hypothetical molecules conceived by a molecular designer are often synthesized and evaluated in the computer long before they are ready to be tested in the laboratory.

The practice of computational chemistry can involve either the application of existing computational tools to new molecules or the development of new computational methods. New software developments are constantly pushing back the frontiers of what is computable. Breakthroughs may involve the incorporation of better physics into the underlying mathematical models, the implementation of algorithms that better utilize computing resources, or the use of time-saving approximations. Popular approximate methods include semiempirical methods that rely partially on experiment to estimate model parameters, and methods that partition a large complex system into two parts, with one part being the focus of the calculation that is treated rigorously and the other being the surroundings that is treated with less rigor. Such approaches coupled with improvements in hardware have made it possible to perform calculations, which required supercomputing capabilities just a few years ago, on a desktop workstation or even a personal computer. *See* ALGORITHM.

Before a computational study is begun, an effective method must first be identified that balances the computational cost against a combination of factors, including accuracy, scalability, interpretability, response speed, and ease of implementation and use.

While keeping this balance in mind, a practitioner's first decision is often between two fundamentally different mechanical treatments of molecules: quantum mechanics and molecular mechanics.

Quantum mechanics. One active area of computational chemistry involves the calculation of molecular properties based on Schrödinger's wave equation

$$H\Psi = E\Psi$$

where H is the Hamiltonian operator for a molecule's electrons as they move in the field of the fixed nuclei, Ψ is a function of the electron coordinates, and E is a constant identified as the electronic energy of the molecule. H typically includes operations that evaluate the kinetic energy of the electrons in terms of the curvature of Ψ, and multiplication operations that account for potential energy. In order for a function Ψ to be a suitable wavefunction, the Hamiltonian operations when applied to Ψ must return Ψ multiplied by the constant E. Solutions of the Schrödinger equation exist for only certain quantized values of E, the lowest of which corresponds to the ground electronic state of the molecule. This energy will change as the positions of the nuclei are moved, and most electronic structure programs feature gradient methods to locate nuclear configurations that correspond to minima in the energy E. The wavefunction Ψ has no physical meaning, but its square is interpreted as the electron density and indicates where the electrons are likely to be found in a molecule. Electron density plays a profound role in determining the physical and chemical properties of a molecule. *See* QUANTUM MECHANICS; SCHRÖDINGER'S WAVE EQUATION.

Regrettably, the Schrödinger equation cannot be solved exactly for any but the simplest molecular systems. Fortunately, useful approximations abound. One popular and successful approximation is to partition Ψ into factors known as molecular orbitals (MOs). Each MO is expressed as a linear combination of atom-centered basis functions known as atomic orbitals (AOs). The coefficients of the AOs in the MOs are determined by an energy minimization process that attributes an orbital energy to each of the MOs and assigns electrons usually in pairs to those MOs that have the lowest energy. *See* MOLECULAR ORBITAL THEORY.

Molecular orbital calculations can take from a few seconds of computer time to many hours, depending on the level of approximation. Computing time is also a steep function of molecular size, increasing roughly as the fourth or fifth power of the number of basis functions (atomic orbitals) used to express the MOs. High-level ab initio calculations can rival the accuracy of experiment in determining ground- and excited-state molecular geometries, electronic-transition frequencies, vibration frequencies, and heats of formation.

Density functional theory is another quantum-mechanical method that has been applied to molecules. Originally developed for problems in solid-state physics, these calculations express the total energy of the system as a function of the ground-state electron density rather than of the molecular wave function. Since the computer time increases only as the cube of the number of basis functions, density functional methods can be used for larger molecules than molecular orbital methods, while often maintaining comparable accuracy. *See* QUANTUM CHEMISTRY.

Electronic structure calculations have been used productively in many chemical disciplines. In the life sciences, electronic structure calculations provide plausible explanations for the biological activity of known molecules, and guidance for the design of new molecules with desired properties. For example, calculated charge distributions for the transition state of an essential biochemical transformation can serve as a template for the design of drugs which act by inhibiting the enzyme responsible for catalyzing this transformation in a pathogenic organism.

Molecular mechanics. The molecular-mechanics model treats the atoms and bonds in a molecule much like a collection of balls joined by springs. Because of its relative simplicity, it can be applied to proteins and other large molecules, where quantum-mechanical methods are impractical. Molecular-mechanics programs can adjust the molecular structure until a conformation of minimum energy is determined. The spring force constants and atomic sphere parameters that define a force field are typically preadjusted by developers so that their programs produce results that nearly agree with a "training set" of experimentally measured or reliably calculated quantities. These parameter values are often assumed to be valid for other molecules and molecular conformations that are outside of the training set. *See* BOND ANGLE AND DISTANCE; CHEMICAL BONDING; MOLECULAR MECHANICS.

Force fields differ mainly by the inclusion of additional terms for interactions such as hydrogen bonding or coupling between the springs. Some force fields have been customized for families of compounds such as proteins and nucleic acids. At one time or another, missing or inappropriate force-field parameters are the bane of every practitioner of molecular mechanics, especially when treating molecules and conformations that are far different from those represented in the training set used to fix the parameters of the force field. In spite of limitations, satisfactory accuracy for many applications and modest computer demands have made molecular mechanics–based methods popular, especially for large molecules.

Gradient techniques used in either molecular or quantum mechanics assure that a low-energy structure at the bottom of an energy well containing the starting point will be found. If the energy surface is complex, other lower-energy structures are likely to exist at remote locations and may not be found unless algorithms that promote broader exploration are exercised. One technique used for broader sampling of potential-energy surfaces is molecular dynamics. *See* CHEMICAL DYNAMICS.

Molecular-dynamics calculations simulate atomic motion by integration of Newton's equations.

Acceleration is directed by the negative gradient of the potential energy, which is calculated with the underlying molecular-mechanics model. The method requires large amounts of computer time just to simulate a few hundred picoseconds to nanoseconds of atomic motion. Various approximations have been developed to decrease the amount of computer time required, while maintaining the accuracy of the calculations. Large systems may be simulated by the imposition of periodic boundary conditions. This means that the box containing the particles is considered to be replicated in all three dimensions, similar in the manner by which strips of wallpaper continue a pattern around a room. The influence of explicit molecules can be replaced by a solvation function whose magnitude depends on the type and fractional exposure of solute atoms to the solvent. In Monte Carlo methods, each particle movement is determined randomly and selectively accepted, based on certain energy criteria. In Langevin dynamics methods, forces between nearby particles are treated exactly, and those due to remote particles, including solvent, are treated stochastically. Software is available for computer simulations of small molecules moving in a variety of solvents, of solutes permeating through lipid membranes, and for various motions of proteins and nucleic acids. *See* MONTE CARLO METHOD; STOCHASTIC PROCESS.

Many important structural and thermodynamic quantities may be computed from simulations of sufficient length by averaging over the ensemble of molecular conformations, which are sampled during the simulation. In order to evaluate the free-energy change associated with a chemical reaction, advantage is often taken of the principle that free energy is a function of state. Consequently in a reaction, changes in free energy F are independent of the path taken to get from the reactants to the products. This principle enables the calculation of the free-energy change of a reaction from the known change for a closely related reference reaction. For example, if the free-energy change ΔF_3 for binding ligand B to a protein receptor R has been measured experimentally (see **illustration**), the free energy change for binding ligand A to the same receptor (ΔF_1) can be estimated as $\Delta F_1 = \Delta F_2 + \Delta F_3 + \Delta F_4$. Molecular simulations can be used to calculate the free-energy changes associated with the kind of structural morphing involved in steps 2 and 4. These transformations, which are an alchemist's dream come true, may be impossible in the laboratory, but are readily simulated in a computer using advanced algorithms based on thermodynamic perturbation or integration theories. In the hands of experienced computational chemists, simulations of this type can be used to find analogs of a reference compound that alter receptor-binding characteristics in a desired fashion. *See* FREE ENERGY; SIMULATION.

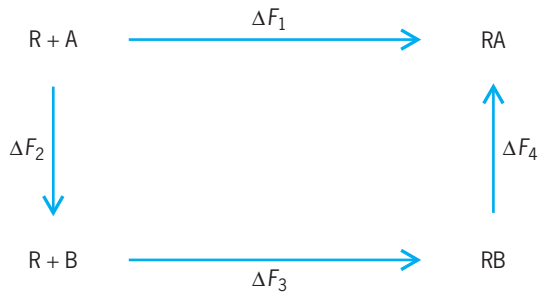

Thermodynamic cycle to calculate free energy of binding of ligand A to a receptor based on a known value for binding ligand B.

Interactive computer graphics. Graphical methods are not necessarily based on a mechanical model, but are often used in conjunction with the mechanical methods described above. Interactive graphics has become so prevalent in the chemical community that it is almost impossible to conceive of calculations without associated molecular manipulation and visualization. Here a picture is worth millions of numbers. When seated at a molecular graphics terminal, the chemist is essentially working with a sophisticated set of molecular models that are stored in the computer memory and displayed on a screen. Molecular orbitals, electron densities, and normal-mode vibrations can be viewed from a variety of perspectives. The sequence of conformations from computer simulations, displayed one frame after another, in effect creates a movie of the thermal motion of the molecules. Moving a mouse or turning a dial moves atoms about the screen as easily as tangible models can be moved physically.

Seeing in three dimensions is often essential for understanding the behavior of a molecule. A sense of depth is produced when a pair of computer-generated left- and right-eye images is seen through one of several types of viewers that aid in merging the images. Stereoscopic viewing is most useful when comparing the structures of two molecules or when studying interactions between them. Structural comparisons in three dimensions enable the user to discern similarities and differences between molecules that might escape a reader of the printed page. Unlike physical models, these graphical renderings of molecules can be easily overlaid in order to further enhance comparisons. Computer graphics can also be used to dock models of small guest molecules into cavities within larger host molecules. This is particularly important for drug designers. For example, when the structure of a target protein is available, computer graphics can be used to study the fit of a drug candidate into potential binding sites in the protein and to anticipate the influence of such binding on the protein's function. Alterations can then be made on the small molecule with the goal of improving the fit to the binding site. The potential of this technology continues to increase as the number of available crystal structures of important enzymes or receptors rises. *See* COMPUTER GRAPHICS; CONFORMATIONAL ANALYSIS; STEREOCHEMISTRY.

Daniel A. Kleier; Zelda R. Wasserman

Bibliography. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, 1991; U. Burkert and N. Allinger, *Molecular Mechanics*, 1982; T. Clark, *A Handbook of Computational Chemistry*, 1985;

C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, 2004; D. M. Hirst, *A Computational Approach to Chemistry*, 1990; F. Jensen, *Introduction to Computational Chemistry*, 1999; A. Leach, *Molecular Modelling: Principles and Applications*, 1996; E. J. Lewars, *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics*, 2003; K. B. Lipkowitz and D. B. Boyd (eds.), *Reviews in Computational Chemistry*, vols. 1–18, 1990–2002.

Computational fluid dynamics

The numerical approximation to the solution of mathematical models of fluid flow and heat transfer. Computational fluid dynamics is one of the tools (in addition to experimental and theoretical methods) available to solve fluid-dynamic problems. With the advent of modern computers, computational fluid dynamics evolved from potential-flow and boundary-layer methods and is now used in many diverse fields, including engineering, physics, chemistry, meteorology, and geology.

Computational methods. The fundamental model of fluid flow, known as the Navier-Stokes equations, is derived from the conservation of mass, momentum, and energy. For example, if the fluid is incompressible, the Navier-Stokes equations can be written as Eqs. (1) and (2), where u, v, and w are the cartesian

$$\text{Continuity}: \qquad \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \qquad (1)$$

Momentum :

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} = -\frac{1}{\rho}\frac{\partial p}{\partial x} + \frac{1}{\mu}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) \qquad (2a)$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z} = -\frac{1}{\rho}\frac{\partial p}{\partial y} + \frac{1}{\mu}\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2}\right) \qquad (2b)$$

$$\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} + w\frac{\partial w}{\partial z} = -\frac{1}{\rho}\frac{\partial p}{\partial z} + \frac{1}{\mu}\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}\right) \qquad (2c)$$

components of velocity, ρ the fluid density, p the pressure, and μ the fluid viscosity. The compressible form of the Navier-Stokes equations, which is of slightly different form because of the variable fluid properties, additionally contains the energy equation, which couples the temperature, velocity, and pressure fields. The complete formulation requires initial and boundary conditions and thermodynamic relations for fluid properties. Numerical solution of this set of nonlinear coupled partial differential equations presents many difficulties, including application of boundary conditions, particularly for free and moving surfaces; grid generation for complex geometries; turbulence modeling; pressure-velocity coupling for incompressible flows; and resolution of shock waves in supersonic flows. A hierarchy of approaches is available for simplifying Eqs. (1) and (2) [see **table**]. *See* DIFFERENTIAL EQUATION; NAVIER-STOKES EQUATION.

Computational fluid dynamics is associated with field-equation solutions of the Euler and Navier-Stokes equations instead of other computational techniques such as boundary-element methods for inviscid flows and momentum-integral methods for boundary-layer flows. The crucial elements of computational fluid dynamics are discretization, grid generation and coordinate transformation, solution of the coupled algebraic equations, turbulence modeling, and visualization.

Discretization. Numerical solution of partial differential equations requires representing the continuous nature of the equations in a discrete form. Discretization of the equations consists of a process where the domain is subdivided into cells or elements (that is, grid generation) and the equations are expressed in discrete form at each point in the grid by using finite difference, finite volume, or finite element methods. The finite difference method requires a structured grid arrangement (that is, an organized set of points formed by the intersections of the lines of a boundary-conforming curvilinear coordinate system), while the finite element and finite volume methods are more flexible and can be formulated to use both structured and unstructured grids

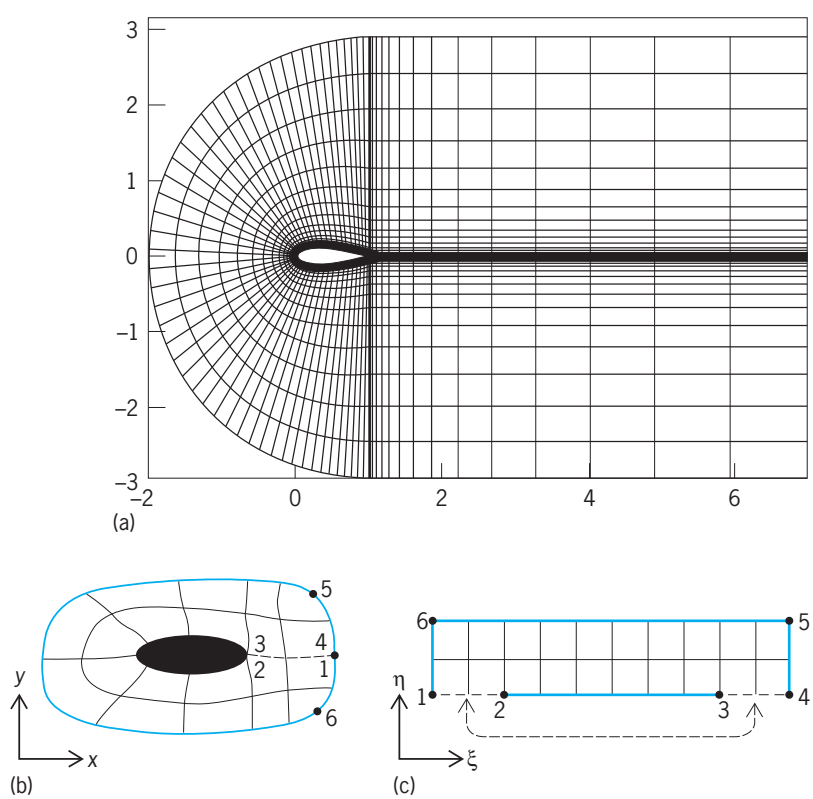

Fig. 1. Computational grid for numerical analysis of flow about a hydrofoil. (*a*) Structured grid. (*b*) Schematic of physical domain. (*c*) Schematic of computational domain.

Hierarchy of approaches to fluid-flow problems*

Mathematical basis	Type of partial differential equation	Physical problems
Laplace equation	Elliptic	Inviscid and irrotational flow
Euler equation	Hyperbolic	Inviscid flow
Boundary-layer equations	Parabolic	Thin viscous layer
Navier-Stokes equations	Mixed	Laminar flow
Reynolds-averaged Navier-Stokes (RaNS) equations + turbulence model	Mixed	Turbulent flow
Large-eddy simulation (Navier-Stokes equations + subgrid model)	Mixed	Turbulent flow
Direct numerical solution (Navier-Stokes equations + grid resolving all length scales)	Mixed	Turbulent flow

*Listed in order of increasing complexity of numerical solution.

(that is, a collection of triangular elements or a random distribution of points). Complex shapes necessitate nonuniform, boundary-conforming grids with grid points concentrated in regions of high gradients, such as inside boundary layers or near shock waves. Structured grids are usually generated by using algebraic interpolation functions or elliptic partial differential equations. For finite difference methods, the equations are rewritten in generalized nonorthogonal coordinates and evaluated on a so-called computational domain, which is defined by the transformation between the boundary-conforming curvilinear-coordinate system in the so-called physical domain and an orthogonal grid with uniform spacing (**Fig. 1**). *See* FINITE ELEMENT METHOD.

Finite difference discretization of Eqs. (1) and (2) is based on the algebraic representation of derivatives by using Taylor-series expansions. For example, the temporal derivative in Eq. (2*a*) expressed as a first-order backward difference is Eq. (3), where (i, j, k) is

$$\frac{\partial u}{\partial t} = \frac{u^n_{i,j,k} - u^{n-1}_{i,j,k}}{\Delta t} + \vartheta[(\Delta t)] \tag{3}$$

the location of the grid point, (n) is the current time level, $(n-1)$ is the previous time level, Δt is the time step between (n) and $(n-1)$, and $\vartheta(t)$ represents the truncated terms in the Taylor-series approximation of the partial derivative $\partial u/\partial t$. (Dropping these terms introduces a source of error, known as a truncation error, which is of an order of magnitude equal to that of the time step Δt.) Due to advantageous stability characteristics, in comparison to central differencing, an upstream-backward difference is typically used for the convective terms. Assuming that the grid is uniform and orthogonal, the first-order upstream difference for the x-component convective term of Eq. (2*a*) is given by Eq. (4), where Δx or the

$$u\frac{\partial u}{\partial x} = u^{n-1}_{i,j,k}\frac{u^n_{i,j,k} - u^n_{i-1,j,k}}{\Delta x} + \vartheta[(\Delta x)] \tag{4}$$

grid spacing, and the grid point locations (i,j,k), are given by a numerical molecule (**Fig. 2**). Usually, the viscous terms in the momentum equations are evaluated by using a central difference. For example, the second-order central difference of the second partial derivative of u in the x direction [from Eq. (2*a*)] is given by Eq. (5).

$$\frac{\partial^2 u}{\partial x^2} = \frac{u^n_{i+1,j,k} - 2u^n_{i,j,k} + u^n_{i-1,j,k}}{(\Delta x)^2} + \vartheta[(\Delta x)^2] \tag{5}$$

Solution of equations. The assembly of the finite difference equations for each grid point results in a large implicit system of algebraic equations for each of the velocity components (u, v, w). For example, the equation for the u component of velocity at the point (i, j, k) is Eq. (6), where A is a matrix of co-

$$\begin{aligned} &A_{i,j,k}u^n_{i,j,k} + A_{i+1,j,k}u^n_{i+1,j,k} + A_{i-1,j,k}u^n_{i-1,j,k} \\ &\quad + A_{i,j+1,k}u^n_{i,j+1,k} + A_{i,j-1,k}u^n_{i,j-1,k} \\ &\quad + A_{i,j,k+1}u^n_{i,j,k+1} + A_{i,j,k-1}u^n_{i,j,k-1} = C^{n-1} \end{aligned} \tag{6}$$

efficients related to the grid and the transformation relations (between the physical and computational domains) and C is a known column vector including pressure, body-force, and velocity terms from the previous time step. Equation (6) is usually solved by iterative schemes, such as the alternating direction implicit (ADI) method, which is based upon splitting Eq. (6) into three tridiagonal systems that correspond to sweeps along each of the index (i,j,k) directions. If Eq. (4) and Eq. (5) are evaluated at time level $(n-1)$ instead of (n), the system of finite difference equations is explicit and the solution at each point can be directly evaluated. However, explicit formulations have restrictions on the ratio of the time step to the grid spacing, which are significant for viscous flows

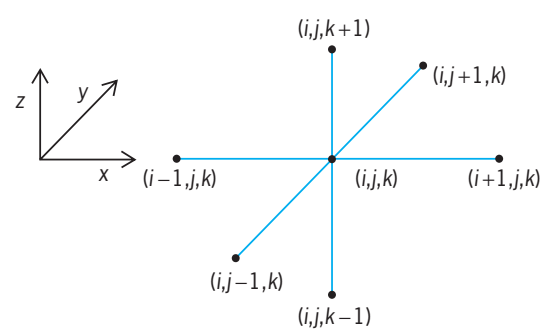

Fig. 2. Numerical molecule, showing nomenclature for neighboring grid points in finite difference discretization.

because of the fine near-wall grid spacing required to resolve the boundary layer. *See* MATRIX THEORY.

For either implicit or explicit approaches, efficient use of computers is required. Current-generation supercomputers are based on vector processors, which are most efficient with large arrays (vectors) of numbers. Optimizing a program, or vectorization, is accomplished by following programming rules so that the compiler can translate the code into vector instructions. *See* SUPERCOMPUTER.

Complete solution of the Navier-Stokes equations requires, for compressible flow, an equation of state such as the ideal gas equation, which relates temperature and density. In contrast, the solution of the incompressible equations presents the problem that the system of equations is lacking an equation for the direct solution of the pressure field. Typically, an equation for pressure is derived which satisfies continuity, and an algorithm must be incorporated which iteratively couples the pressure and velocity fields.

Turbulence modeling. There are a variety of approaches for resolving the phenomena of fluid turbulence. The Reynolds-averaged Navier-Stokes (RANS) equations, which are derived by decomposing the velocity into mean and fluctuating components, contain additional turbulent Reynolds-stress terms which are unknown. Closure of the RANS equations can be accomplished by using the Boussinesq assumption, which relates the Reynolds stresses to the mean rate of strain or the Reynolds-stress equations, which can be solved directly with additional modeling assumptions. An alternative to the RANS equations is large-eddy simulation, which solves the Navier-Stokes equations in conjunction with a subgrid turbulence model. The most direct approach to solving turbulent flows is direct numerical simulation, which solves the Navier-Stokes equations on a mesh that is fine enough to resolve all length scales in the turbulent flow. Unfortunately, direct numerical simulation is limited to simple geometries and low-Reynolds-number flows because of the limited capacity of even the most sophisticated supercomputers. *See* TURBULENT FLOW.

Visualization. The final step is to visualize the results of the simulation. Powerful graphics workstations and visualization software permit generation of velocity vectors, pressure and velocity contours, streamline generation, calculation of secondary quantities (such as vorticity), and animation of unsteady calculations. Despite the sophisticated hardware, visualization of three-dimensional and unsteady flows is still particularly difficult. Moreover, many advanced visualization techniques tend to be qualitative, and the most valuable visualization often consists of simple *x-y* plots comparing the numerical solution to theory or experimental data. *See* COMPUTER GRAPHICS.

Applications. Computational fluid dynamics has wide applicability in such areas as aerodynamics, hydraulics, and geophysical flows, with length and time scales of the physical processes ranging from millimeters (fractions of an inch) and seconds to kilometers (miles) and years.

Vehicle aerodynamics and hydrodyamics, which have provided much of the impetus in the development of computational fluid dynamics, are primarily concerned with the flow around aircraft, automobiles, and ships where the goal is the determination of the lift and drag forces and the moments resulting from these forces. Furthermore, computational fluid dynamics is being applied to explicate basic physics and to examine complete vehicle configurations. *See* AERODYNAMIC FORCE; AERODYNAMICS; HYDRODYNAMICS.

Hydraulics and environmental fluid dynamics also find many applications of computational fluid dynamics. The prediction of pollutant dispersion in rivers, lakes, and estuaries and riverine morphology, including river-bed formation and sediment deposition, have been studied by using computational fluid dynamics. Detailed analysis of hydraulic structures by using computational fluid dynamics can provide data which are useful in new designs and in the modernization of aging facilities. *See* HYDRAULICS; RIVER ENGINEERING; WATER POLLUTION.

The study of atmospheric and oceanic dynamics finds prolific use of computational fluid dynamics. Although long-term prediction of weather is not possible, because of its random nature and the multiplicity of length and time scales, the study of smaller problems (such as thunderstorms and ocean circulation) and low-resolution global-scale phenomena (for example, global climate models) is possible. *See* CLIMATE MODELING; DYNAMIC METEOROLOGY; FLUID FLOW; FLUID-FLOW PRINCIPLES; NUMERICAL ANALYSIS; OCEAN CIRCULATION; SIMULATION; WEATHER FORECASTING AND PREDICTION.

Eric Paterson; Fred Stern

Bibliography. M. Abbott and D. R. Basco, *Computational Fluid Dynamics: An Introduction for Engineers,* 1990; K. A. Hoffmann and S. T. Chiang, *Computational Fluid Dynamics for Engineers*, 2 vols., 2d ed., 1993; Y. Jaluria, *Computer Methods for Engineering*, 1988; J. Von Wendt (ed.), *Computational Fluid Dynamics: An Introduction*, 2d ed., 1996.

Computational intelligence

An approach to designing intelligent systems using mathematical algorithms. It consists of the fields of neural networks, evolutionary computation, and fuzzy sets. The term was first used as an alternative to artificial intelligence in the 1980s. The current popular use of the term as defined above results from an effort by James Bezdek in the early 1990s to differentiate the fields of neural networks, evolutionary computation, and fuzzy sets from classical artificial intelligence (AI). There is significant continuing research in the area of computational intelligence, as well as a number of examples of practical, everyday systems that make use of computational intelligence approaches.

Computational intelligence has been used to enable a car to autonomously navigate from Washington, DC, to San Diego with a human driver

intervening only a few times; to support automatic transmissions and antilock braking systems; and to determine whether one might grant credit card purchase requests. One of the best checkers-playing computer programs developed so far uses a computational intelligence approach.

In the classical approach to intelligent systems, known as artificial intelligence, all reasoning may be done using high-level symbols. For example, one can use rules such as: if it is raining, bring an umbrella. These types of rules (called heuristic rules) may also be used in fuzzy systems with the difference being that approximate matches are allowed. So one might still bring an umbrella on a cloudy overcast day where it was not quite raining at the moment. *See* ARTIFICIAL INTELLIGENCE.

Systems that are described by the term "computational intelligence" make use of lower-level computational features even when they may have a high-level interpretation. For example, in the case of fuzzy systems, which may be expressed as heuristic rules, there will be an underlying mathematical description of sets as well as conjunction, disjunction, and inference operators. Fuzzy rules have been successfully used in the control of antilock brakes.

Fuzzy systems. A fuzzy system is one that makes use of approximate reasoning using the mathematical basis of fuzzy logic. The concept of fuzzy logic is concerned with membership in the classes of true or false. For example, a person with some bald spots might be said to be balding. However, he would completely belong neither to the class of bald people nor to the class of people with hair.

To see how this differs from probability, consider that you have the opportunity to acquire one of two lottery tickets after the winning tickets from the lottery have been chosen. The first is described as having a 90% chance of having won a lot of money, and the second one is described as having a 0.9 membership in the class of tickets which have won a lot of money. In the first case you have a 90% chance of winning a lot of money and a 10% chance of winning nothing. In the second case, you will have approximately won a lot of money. This is likely to translate into something between a lot of money and a fair amount of money. If you had to put a probability distribution on the likely winnings in the second case, you might have a probability of 0 of winning no money and a probability of 0 of winning the entire lottery, but a probability of 80% of being runner-up, 10% of being second runner-up, and 10% of being third runner-up. Most people would choose the ticket with the 0.9 membership in the class of tickets which have won a lot of money because they would be guaranteed a good sum of money. Of course, some gamblers would choose the all-or-nothing route. *See* FUZZY SETS AND SYSTEMS.

Neural networks. Perhaps the most widely used technology within the computational intelligence approaches is that of neural networks. Neural networks are inspired by the underlying processing done in the human brain. There are nodes which correspond (loosely) to neurons in the brain and have interconnections which correspond (loosely) to the synapses that connect neurons in the brain. A simple three-layer neural network is shown in the **illustration**.

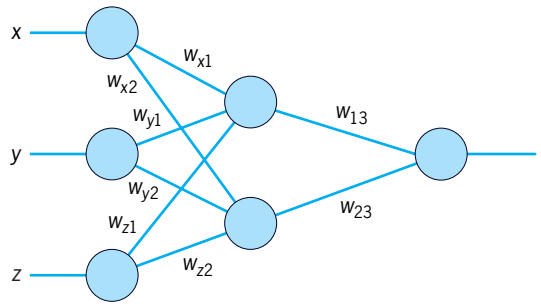

Feedforward neural network with three inputs, one output, and one hidden layer of two nodes. The connection weights are denoted by w_{ij}.

A neural network learns from a set of labeled examples to produce intelligent responses. Neural networks are also used to intelligently group unlabeled data into clusters. Thus, a neural network is one attempt to mimic the underlying learning processes of the human brain. Clearly, a system that can learn to improve its performance will be quite valuable. There are many approaches to training a neural network. Many of them rely on the feedforward–backpropagation approach which was popularized in the mid-1980s. In this approach, examples are presented to the neural network, and the connection weights between nodes are updated to make the outputs of the network closer to those of the examples. The connection weights loosely correspond to the synaptic strengths of connections between neurons in the human brain.

Neural networks have a powerful property of universal approximation. This means that they can arbitrarily closely approximate any function, in theory. This has been proven mathematically. In practice, the theorems do not show how to configure the network, what type of functions should exist within a node to determine its output or, perhaps most importantly, do not guarantee that the appropriate weights can be learned. So it is known that there is a neural network that arbitrarily closely approximates a particular function, but it is not known if this network can ever be discovered.

Consider the function that emulates a person who decides whether a credit card transaction should be allowed. Given what is known about neural networks, it would be possible to develop one that was as good as any human expert if we knew how to do so. It is unknown if this can actually be done, but in practice systems for making such decisions (with capabilities that do not necessarily match those of a human) have been built. The function to be approximated is the binary human decision of whether to grant credit or not grant credit. A highly accurate system will, clearly, save money for the credit card company, and it will have the side effect of preventing people from getting in too much debt.

Fuzzy systems and neural networks have been combined to produce neuro-fuzzy systems. Under certain conditions they have been shown to be universal approximators, also. Typically, a set of fuzzy rules can be extracted from a neuro-fuzzy system which may provide an explanation capability that is lacking in a typical neural network where all information is stored as a set of weights. The extent of the explanation capability will depend on the complexity of the fuzzy rules. There is, typically, a trade-off between complexity and accuracy. So more understandable but less complex rules will be less accurate than the more complex rules in the usual case. *See* NEURAL NETWORK.

Evolutionary computation. Evolutionary computation embodies a set of methods to evolve solutions to problems. Evolutionary computation has typically been applied to optimization problems. For example, excellent solutions to the classical traveling salesman problem have been evolved. This problem is to minimize the distance traveled by a salesperson who has to make calls in a set of cities.

Evolutionary computation considers potential solutions to a problem as members of a population. Mimicking how natural evolution is considered to work, members of the population can be chosen to mate. They then produce offspring, typically with the use of a crossover operator in which parts of each solution become a part of the offspring. After crossover, a mutation operator is often applied to the child which may randomly change some of its characteristics. The offspring compete for survival to the next generation based upon a fitness function. Those that are more fit survive.

There are at least three major paradigms within evolutionary computation. Genetic algorithms, which merge parts of solutions to form new solutions and then apply random mutation to them, can be applied to both binary bit strings and real-valued solutions. Evolutionary strategies make heavy use of the mutation operator. There is also genetic programming in which computer programs are evolved, typically in the Lisp language. The latter approach provides a computer-executable solution to a problem. *See* GENETIC ALGORITHMS.

Evolutionary computation has been used to evolve the weights, and sometimes the structure, of neural networks. It has also been used to evolve fuzzy systems and fuzzy neural networks. Thus, these approaches are complementary. An interesting application of evolutionary computation was in evolving an expert checkers-playing program that did not require any human knowledge. The system evolved weights for a neural network which served as a board evaluation function. After evolution, it was able to beat an expert-level checkers player. *See* EVOLUTIONARY COMPUTATION.

In summary, all the computational intelligence approaches discussed here can be used alone or together to produce systems that exhibit human expertise or intelligence. Lawrence O. Hall

Bibliography. J. C. Bezdek, On the relationship between neural networks, pattern recognition and intelligence, *Int. J. Approx. Reasoning*, 6(2):85–107, February 1992; C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford Press, 1995; D. B. Fogel, *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*, 3d ed., Wiley-IEEE Press, 2006; D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, MA, 1989; H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, 3d ed., CRC Press, New York, 2005.

Computer

A device that receives, processes, and presents information. The two basic types of computers are analog and digital. Although generally not regarded as such, the most prevalent computer is the simple mechanical analog computer, in which gears, levers, ratchets, and pawls perform mathematical operations—for example, the speedometer and the watt-hour meter (used to measure accumulated electrical usage). The general public has become much more aware of the digital computer with the rapid proliferation of the hand-held calculator and a large variety of intelligent devices, ranging from typewriters to washing machines, and especially with exposure to the Internet and the World Wide Web. *See* CALCULATORS; INTERNET; WORLD WIDE WEB.

In the 1950s and 1960s the definition of "computer" included only the provision of input and output, and the ability to process the input to produce the output. Since then, the common understanding of the term (in particular, for digital computers) has come to include storage, and even the ability to store the program of instructions, allowing decision making based on previously computed results. *See* COMPUTER PROGRAMMING; COMPUTER STORAGE TECHNOLOGY.

Analog computer. An analog computer uses inputs that are proportional to the instantaneous value of variable quantities, combines these inputs in a predetermined way, and produces outputs that are a continuously varying function of the inputs and the processing. These outputs are then displayed or connected to another device to cause action, as in the case of a speed governor or other control device. *See* CONTROL SYSTEMS.

Small electronic analog computers are frequently used as components in control systems. Inputs come from sensing devices which output an electrical signal (transducers). These electrical signals are presented to the analog computer, which processes them and provides a series of electronic outputs that are then displayed on a meter for observation by a human operator or connected to an electrical action device to ring a bell, flash a light, or adjust a remotely controlled valve to change the flow in a pipeline system. If the analog computer is built solely for one purpose, it is termed a special-purpose electronic analog computer.

In any analog computer the key concepts involve special versus general-purpose computer designs,

and the technology utilized to construct the computer itself, mechanical or electronic. An analog computer is often combined with a digital computer to form a hybrid computer. *See* ANALOG COMPUTER.

Digital computer. In contrast, a digital computer uses symbolic representations of its variables. The arithmetic unit is constructed to follow the rules of one (or more) number systems. Further, the digital computer uses individual discrete states to represent the digits of the number system chosen.

Electronic versus mechanical computers. The most prevalent special-purpose mechanical digital computers used to be the supermarket cash register, the office adding machine, and the desk calculator. Each of these has been universally replaced by electronic devices allowing much greater logical decision making and greatly increased speed. For example, most products now carry a bar code, the Universal Product Code (UPC). The code is scanned by a light-sensitive device, bringing information about each product into the point-of-sale (POS) terminal that has replaced the mechanical cash register. The POS terminal then computes total charges and provides a receipt for the customer. It may also communicate with a centralized computer system that controls inventory, accounts payable, salaries and commissions, and so on. While a mechanical cash register could carry out only a small number of operations each minute, and some electromechanical devices might handle several hundred operations per second, even a small general-purpose electronic computer can carry out its computations at millions of operations per second.

Stored program operation. A digital computer works with a symbolic representation of variables; consequently, it can easily store and manipulate numbers, letters, images, sounds, or graphical information represented by a symbolic code. Typically, a general-purpose electronic digital computer operates on numbers by using both decimal and binary number systems, and on symbolic data expressed in an alphabet. It contains both an arithmetic unit and a storage unit. As the digital computer processes its input, it proceeds through a series of discrete steps called a program. The storage unit serves to retain both the values of the variables and the program to process those variables. The arithmetic unit may operate on either variables or coded program instructions interchangeably, since both are usually retained in the storage unit in the same form. Thus, the digital computer has the capability to be adaptive, because processing can be determined by the previously prepared program, by the data values supplied as input to the computation, and by the values generated during the course of the computation. Through the use of the stored program, the digital computer achieves a degree of flexibility unequaled by any other computing or data-processing device.

Applications. In the past, most digital computers were confined to standard applications, such as bookkeeping, accounting, engineering design, and test data reduction. However, the advent of the relatively inexpensive and readily available personal computer, and the combination of the computer and communications, such as by the use of networks, have dramatically altered that pattern. The most common application now is probably text and word processing, followed by electronic mail. *See* ELECTRONIC MAIL; LOCAL-AREA NETWORKS; MICROCOMPUTER; WORD PROCESSING.

Ready access to national networks has fostered the growth of communities of computer users who communicate easily and effectively on the World Wide Web, including an explosion of commerce via online communications. The convergence of computers and network communications also has encouraged the development of network computers, smaller and simpler computers that depend on the availability of software and data from around the world arriving via networks in order to accomplish their tasks.

The ability of even modest computer systems to store, organize, and retrieve very large amounts of information has brought about radical changes in the very nature of many business offices. Indeed, many industrial, academic, and governmental processes have been irrevocably changed by the computer. It is often said that the Manhattan Project, although managed by paper and pencil in the 1940s, simply could not have been accomplished at all today without using computers. That is, there has been a change in the very thought processes that people use to accomplish tasks, in order to accommodate computer-based solutions for those tasks.

Personal workstations, used in engineering design and other applications requiring intense computation and sophisticated graphics, have become more powerful than earlier, very large computer systems. This has moved computation, and even the storage of many large data files, directly onto each person's desk, and in many cases, into handheld computers. In addition, applications that were once considered esoteric have led to industrial applications, such as the use of robots on manufacturing assembly lines. Many of the heuristic techniques employed by these robots are based on algorithms developed in such artificial-intelligence applications as virtual reality. *See* ARTIFICIAL INTELLIGENCE; COMPUTER-AIDED DESIGN AND MANUFACTURING; COMPUTER-AIDED ENGINEERING; COMPUTER GRAPHICS; ROBOTICS; VIRTUAL REALITY.

Computers have begun to meet the barrier imposed by the speed of light in achieving higher speeds. This has led to research and development in the areas of parallel computers (in order to accomplish more in parallel rather than by serial computation) and distributed computers (taking advantage of network connections to spread the work around, thus achieving more parallelism). Continuing demand for more processing power has led to significant changes in computer hardware and software architectures, both to increase the speed of basic operations and to reduce the overall processing time. *See* COMPUTER ARCHITECTURE; CONCURRENT PROCESSING; DISTRIBUTED SYSTEMS (COMPUTERS); MULTIPROCESSING; SUPERCOMPUTER.

In the hardware domain, clock rates approaching gigahertz speed have led to ever smaller computers, since electronic pulses cannot travel further than 1 ft (0.3 m) per nanosecond; instruction and video caches have been enlarged for faster processing and display; and central processing units are continually being redesigned to enable more parallelism (and to avoid the "von Neumann bottleneck" inherent in a serial architecture). In the software domain, parallelism is being achieved both locally via multithreading and multiprocessor operating systems, and globally via distributed computation. *See* COMPUTER PERIPHERAL DEVICES; DATABASE MANAGEMENT SYSTEM; DIGITAL COMPUTER; HUMAN-COMPUTER INTERACTION. Bernard A. Galler; John H. Sayler

Bibliography. M. Campbell-Kelly and W. Aspray, *Computer: A History of the Information Machine*, Basic Books, New York, 1996; P. Norton, *Peter Norton's Introduction to Computers*, Macmillan/McGraw-Hill, New York, 1995.

Computer-aided circuit design

The application of specialized computer software to the design, simulation, and physical layout of electronic circuits on semiconductor chips and circuit boards. *See* INTEGRATED CIRCUITS; PRINTED CIRCUIT.

Designing integrated circuits is extremely challenging because of the large amount of detailed technical data associated with circuits containing many transistors and their interconnections. Competition in the integrated circuit marketplace requires that integrated circuit products be optimal in some respect to be profitable. Thus, the integrated circuit design task is necessarily coupled with an optimization task. Readily available transistors make many alternative circuit implementations possible, and optimization means choosing the circuit that maximizes (or minimizes) an objective function of interest, such as performance, power consumption, area, or ease of testing. Because of this complexity, computer-aided design software has evolved into the principal tool used by most designers of high-performance analog and digital electronic devices and systems. Only computer programs can choose optimal designs, can evaluate these designs before prototypes are built, and can simulate circuit performance under realistic component tolerances and environmental conditions.

Classical "breadboarding" and bench testing of design prototypes are infeasible when circuits containing anywhere from hundreds to millions of transistors are involved. Even for circuits containing few components, conventional breadboarding techniques may not evaluate circuit performance adequately. Component positioning on the circuit board, component interconnect routing, and board material properties all can have a significant effect on circuit performance, and if the breadboard does not faithfully replicate all of these final product physical characteristics, its results may be of little use. *See* BREADBOARDING.

Once an analog or digital circuit design has been completed and extensive simulation has shown its performance to be satisfactory, the myriad details required to prepare the device or electronic system for manufacture must be conveyed to the appropriate fabrication facility. Comprehensive computer-aided design systems contain software to generate computer files that describe the physical structure and layout of the circuit components and the routing of their interconnections. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

In semiconductor manufacturing, these files are used to create integrated circuit mask sets. Microelectronic circuits are manufactured using a photolithographic process, by which patterns corresponding to transistors and connection wires are transferred onto a silicon wafer, which is then customized by local physical or chemical reactions. These patterns are contained in a mask set, a detailed draft of the geometries of the transistors and wires on several layers.

In circuit board fabrication, data files are used to generate conductor path-etching masks and to provide positioning data for drilling via and component mounting holes. This software also can generate circuit schematic drawings and prepare component parts lists. Philip V. Lopresti

Design of Digital Circuits

Most digital electronic products employ very large scale integration (VLSI) circuits, often called microelectronic circuits. Such circuits may have over 1 million transistors. This large scale of integration is made possible by advances in semiconductor technologies that allow the manufacture of transistors and wires with dimensions much smaller than 1 micrometer. Raising the scale of circuit integration is advantageous, because increasingly complex functions can be realized on a single integrated circuit, with a corresponding increase in reliability and decrease in overall cost.

There are four stages in the creation of a digital integrated circuit: design, fabrication, testing, and packaging (**Fig. 1**). Design is the refinement of a functional specification into the geometries needed for fabrication. Fabrication means manufacturing semiconductor dies, that is, the unpackaged circuits. Testing means checking for manufacturing defects that may lead to erroneous operation. Working dies are placed into ceramic or plastic packages. Computer-aided-design tools support design and testing, while computer-aided-manufacturing tools support the remaining stages.

Major design tasks. Circuit design can be split into three major tasks: conceptualization and modeling, synthesis and optimization, and validation. Conceptualization consists of casting an idea into a model that captures the function the circuit will perform. Synthesis consists of refining the model from an abstract one to a detailed one with all the features required for fabrication. Validation consists of verifying the consistency of the models used during design, as well as some properties of the original model.

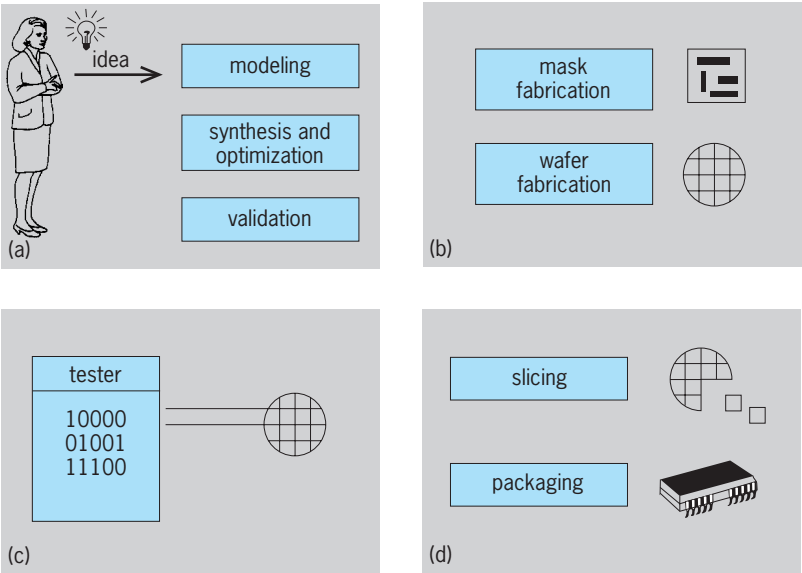

Fig. 1. Four phases in creating a microelectronic chip. (*a*) Design. (*b*) Fabrication. (*c*) Testing. (*d*) Packaging. (*After G. De Micheli, Synthesis and Optimization of Digital Circuits, McGraw-Hill, 1994*)

Modeling. Circuit modeling plays a major role in microelectronic design, because it represents the vehicle used to convey information. Modeling must be rigorous as well as general, transparent to the designer, and machine-readable. Most modeling is now done using hardware description languages (HDLs). Graphic models are sometimes used, such as flow and schematic diagrams. The very large scale nature of the problem forces the modeling style, both textual and graphical, to support hierarchy and abstractions, allowing a designer to concentrate on a particular portion of the model at any given time.

Synthesis. The overall goal of circuit synthesis is to generate a detailed model of a circuit, such as a geometric layout, that can be used to fabricate the chip. This objective is achieved through a stepwise refinement process, during which the original abstract model provided by the designer is iteratively detailed. As synthesis proceeds in refining the model, more information is needed regarding the technology and the desired design implementation style. A functional model of a circuit may be fairly independent from the implementation details, while a geometric layout model must incorporate all technology-dependent specifics, such as wire widths.

Circuit synthesis is usually layered in three stages, related to the modeling abstraction level involved: high-level synthesis, logic-level synthesis, and physical design (**Fig. 2**). High-level synthesis, also called architectural-level synthesis, consists of generating the macroscopic structure of a circuit. This activity corresponds to determining an assignment of the circuit functions to operators, called resources, as well as their interconnection and the timing of their execution. Logic-level synthesis is the task of generating the gate-level structure of a circuit. It entails the manipulation of logic specifications to create logic models as interconnections of logic primitives. Logic design is often eased by the use of libraries of logic gates. The task of transforming a logic model into an interconnection of instances of library cells, that is, the back end of logic synthesis, is often referred to as library binding or technology mapping. Physical design consists of creating the geometric layout of the circuit. It entails the specification of all geometric patterns defining the physical layout of the chip, as well as their position on the die.

Circuit optimization is often combined with synthesis. It entails the selection of some particular choices in a given model, with the goal of raising one or more figures of merit of the design. The role of optimization is to enhance the overall quality of the circuit. Most designers equate quality with circuit performance. Performance involves the time required to process some information, as well as the amount of information that can be processed in a given time period. Circuit performance is essential to competitive products in many application domains. Power consumption is an important quality measure for microelectronic circuits to be used for portable applications, where the autonomous operation time depends on the battery capacity and discharge rate. Another measure of circuit quality is the overall circuit area. An objective of circuit design is to minimize the area, for three main reasons: (1) smaller circuits allow more dies to be placed on each wafer, and therefore enable lower manufacturing costs to be achieved; (2) the manufacturing yield decreases with an increase in chip area; and (3) large chips are more expensive to package. The circuit quality also involves the circuit testability, that is, the ease of testing the chip after manufacturing. Testable chips are clearly desirable because earlier detection of malfunctions in electronic systems enables overall costs to be lowered. *See* OPTIMIZATION.

Validation. Circuit validation consists of acquiring a reasonable certainty level that a circuit will function correctly, under the assumption that no manufacturing fault is present. Circuit validation is motivated by the need to remove all possible design errors before proceeding to expensive chip manufacturing. It can be performed by simulation and by verification

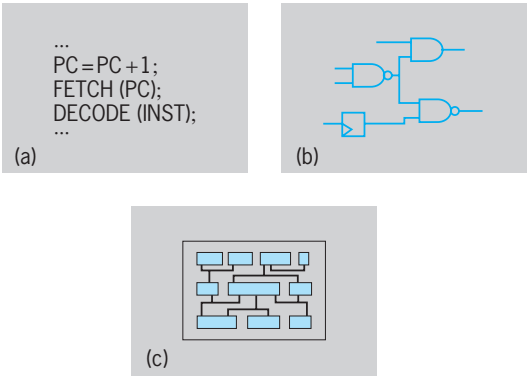

Fig. 2. Three abstraction levels of a circuit representation. (*a*) Architectural level. (*b*) Logic level. (*c*) Geometrical level. (*After G. De Micheli, Synthesis and Optimization of Digital Circuits, McGraw-Hill, 1994*)

methods. Digital circuit simulation consists in analyzing the circuit variables over an interval of time for one specific set (or more sets) of input patterns. Simulation can be applied at different levels, corresponding to the model under consideration. Simulated output patterns must then conform to the expected ones. Even though simulation is the most commonly used way of validating circuits, it is often ineffective for large circuits except for detecting major design flaws. Indeed, the number of relevant input pattern sequences needed to analyze a circuit grows with the circuit size, and designers must be satisfied with monitoring a small subset. Verification methods, often called formal verification methods, consist of comparing two circuit models and detecting their consistency. Another facet of circuit verification is checking some properties of a circuit model, such as whether there are deadlock conditions. *See* SIMULATION.

Testing. Microelectronic circuits are tested after manufacturing to screen fabrication errors. Thus, testing is the verification of correct manufacturing. Since impurities are unavoidable in the air of manufacturing plants, particles from the air can cause defects in random positions on some circuits. There are several testing techniques, and their relation to design is important. In particular, circuits may be designed with a self-testing feature. After manufacturing, the circuit can be set in testing mode, where it will generate, apply, and analyze testing patterns to check proper operation.

Design for testability means addressing the ease of testing during the design stage. Circuit testability affects the circuit quality. A circuit that is not fully testable is less valuable than another one that is, because faulty units may be released on the market. Enhancing testability can be done by appropriate logic and architectural synthesis techniques.

Hardware/software codesign. Digital microelectronic circuits may contain complex building blocks, such as processors, controllers, and memories. Processors and controllers may be designed to execute specific programs stored in memory, which are referred to as embedded software. For integrated circuits containing programmable components executing software, hardware/software codesign techniques are needed to exploit the synergism between hardware and software to achieve the design objectives. Computer-aided codesign tools support a variety of tasks, including performance analysis of hardware/software systems, support for deciding on the hardware/software implementation of system functions, and support for developing software compilers that are retargetable to various processor architectures.

Giovanni De Micheli

Design of Analog and Radio-Frequency Circuits

To meet the challenges of delivering high performance at low power, and to avoid the high cost of multiple design prototyping, designers of analog and radio-frequency circuits utilize very specialized simulation and extraction tools that equate physical design to electrical performance. Mixed analog, radio-frequency, and digital circuit design needs detailed process-dependent active device and passive interconnect characterization data as well as robust circuit simulation tools to accurately predict proper circuit operation under all operating conditions.

Integrated-circuit design process. The analog-integrated circuit design process may be broken down into four principal areas involving multiple design and design support teams. The first area is dedicated to model development and computer-aided design or electronic design automation (EDA) tool support. Team activity focuses on active device model parameter extraction (usually through measurement on actual devices), passive component model extraction (usually through electromagnetic simulation), and component library development and maintenance, including the implementation of parametrized layout structures or "cells." A second design area is dedicated to system-level architectural design and block partitioning. At this high level, designers are responsible for meeting overall system requirements (both digital and analog) and specifying the requirements for individual functional blocks or subcircuits. The third principal design activity is devoted to detailed circuit design and to follow-on system integration. This group uses the electrical models and associated layout information provided by the modeling group. The fourth design team is involved in design testing and verification. This area concerns itself with validating the actual hardware, accounting for physical measurements, test fixturing, and electrical packaging. Sophisticated tool infrastructure and model databases are required to tie together the data flowing among these four design teams.

At high frequencies, component material and geometric attributes play a strong role in electrical performance. Therefore, reliable schematic-to-layout modeling is critical to successful design. Sophisticated design environments that provide a user interface for synchronized schematic and layout capture are fully integrated within several modern analog and radio-frequency simulators. Many of these environments also support parameter tuning, optimization, statistical analysis, and result generation.

Behavioral versus circuit models. In early design phases involving definition of the system architecture and block partitioning, a top-down approach using a high-level behavior model representation is common, since behavior model simulation requires much less computational effort than is needed for transistor-level circuit simulation. Computer models for various function blocks, such as low-noise amplifiers, mixers, voltage-controlled oscillators, or in-phase and quadrature demodulators, typically are used to support early system development.

Behavior model parameters will supply performance simulation accuracy to only a second order of parameter analysis and simulation accuracy within the radio-frequency system. Modern systems are complex and behavioral models may not predict the integrated circuit behavior sufficiently accurately under a wide range of specified operating conditions.

Some of the lost critical circuit detail includes the device terminal impedances, bias and temperature dependencies, transient response times, and signal feedback conditions.

Working with more detailed circuit information can overcome these deficiencies, but doing so comes at the cost of increased characterization effort. Proven existing netlist-based circuit models are an alternative to behavioral models, but circuit-level modeling requires a simulator that can analyze circuits containing a very large number of transistors. These simulators also need to contain libraries of commercial components and manufacturer's device models for standard transistors, diodes, resistors, capacitors, and inductors. The simulator also should be able to support cosimulation with functional blocks described by other circuit or behavioral models, and should include utilities to speed the process of transmission-line design, matching network extraction, and filter synthesis for both lumped and distributed radio-frequency designs.

To address characterization of a wide range of on-chip or packaged physical, analog and radio-frequency/microwave circuit designers rely heavily on a variety of electromagnetic simulation technologies. These tools are used to extract electrical behavior from the structure in the form of transmission-line parameters or equivalent SPICE (Simulation Program with Integrated Circuit Emphasis) models. The resulting extracted electrical models are then inserted into the complete network for cosimulation with transistor-level circuits or high-level behavioral models.

Linear circuit simulation. Linear analysis is used to compute the frequency response of passive circuits or circuits which operate under small-signal conditions. In active circuits operating under small signal conditions, signal level and termination values do not influence the operation of circuit components and the superposition principle holds true. Most high-frequency circuit designers rely on linear analysis to develop circuits to (1) match impedances between cascaded components to ensure optimum power transfer, (2) suppress unwanted signal transmission at specific frequencies, and (3) deliver optimum load terminations for active devices.

Linear simulation results are offered in a wide range of representations including scattering, impedance, and admittance (S,Z,Y) network parameters at specific excitation frequencies. For two-port (terminal) networks, linear analysis also can determine stability, voltage and power gain, noise, and matching element parameters. If there are nonlinear devices in the circuit, a bias-point analysis must first be performed, in which case each device is linearized at its bias point, and linear analysis in the frequency domain proceeds.

Extended linear analysis results may include group delay (delay of the propagation of energy at a given frequency), or even steady-state time-domain responses for impulse, step, and modulated pulses, and nodal voltages at internal circuit nodes. In this case the frequency-domain results from linear-network analysis are transformed into the time domain via the fast Fourier transform (FFT), which presents steady-state (periodic) information about network parameters in the time domain. *See* FOURIER SERIES AND TRANSFORMS.

Harmonic balance analysis (HBA). Harmonic balance analysis supports the large-signal simulation of nonlinear devices and is performed using a spectrum of harmonically related frequencies, similar to measured signals on a spectrum analyzer. This analysis type involves the periodic steady-state response of a fixed circuit to a given predetermined set of fundamental tones. This analysis is limited to periodic responses because the basis set chosen to represent the physical signals in the circuit are sinusoids, which are periodic. *See* SPECTRUM ANALYZER.

In harmonic balance, the circuit is usually divided into two subcircuits connected by wires forming multiports. One subcircuit contains the linear components of the circuit and the other contains the nonlinear device models. The linear subcircuit response is calculated in the frequency domain at each harmonic component and is represented by a multiport Y (admittance) matrix. This is the function performed by linear analysis. The nonlinear subcircuit contains the active devices whose models compute the voltages and currents at the intrinsic ports of the device. (Parasitic elements are linear and absorbed by the linear subnetwork.)

For circuits operating under large signal conditions, device nonlinearity will generate harmonically related signals. Analysis parameters, such as the number of harmonics, specify which combinations of these harmonically related signals will be considered and set the fundamental frequencies used in the analysis of the network. Harmonic balance analysis is often used to examine the compression characteristics and spurious output of radio-frequency circuits such as amplifiers, mixers, and frequency converters.

Since harmonic balance analysis supports frequency domain linear models such as transmission lines as well as nonlinear device models such as transistors, it is well suited to simulate high-frequency networks operating under steady-state conditions.

Transient analysis. Many of today's communication devices, however, as a result of fast switching and on/off states, operate under non-steady-state conditions. Transient circuit analysis is required to understand circuit behavior under nonperiodic conditions. Radio designers, for example, often use transient analysis to investigate the transient performance of oscillators and phase-locked-loop synthesizers.

Transient analysis discretizes time and uses numerical integration methods to solve the circuit equations over each time increment. The equations are derived from the circuit interconnections and circuit device models. The analysis begins with vectors of initial values for the node voltages and branch currents. At the beginning of each iteration, the simulator linearizes the circuit around the solution calculated in the previous iteration, and calculates the

matrix associated with the linearization point. The simulator then uses a matrix solver to calculate a new vector of unknowns. At the end of each iteration, the analysis updates the solution and checks to see if the new solution satisfies the specified tolerances. If the result is within tolerance, the loop terminates. Otherwise, the loop uses the new result to start the next iteration.

Time-domain simulators have waveform sampling requirements that can lead to excessively long run times for certain applications or that can introduce simulation errors. For a device operating under multitone excitation, a time-domain simulator must have a sample rate twice the highest frequency rate, yet must take enough samples to capture the lowest frequency of interest. High sampling rates continuing for a long sampling period result in a lengthy simulation. Capturing the intermodulation products of a radio-frequency circuit under multitone drive, for example, can require prohibitive simulation run times. It also has been shown that the black-box S-parameter and SPICE models used in these simulators has inadequately modeled high-frequency time-domain effects in a number of cases.

High-frequency passive modeling. One of the most fundamental impairments in a radio-frequency circuit or a high-speed channel is the attenuation of the transmitted signal across the printed circuit board caused by transmission line loss. Printed circuit boards with dielectric substrates composed of FR4 (Flame Retardant 4, a fiber glass-epoxyresin combination that is widely used in printed circuit boards) were not originally intended to support signals at multigigabit speeds, and their dielectric loss is very pronounced at high frequency. Typical microstrip differential transmission lines on FR4 exhibit an insertion loss of roughly 0.5 dB/in. at 5 GHz and 0.9 dB/in. at 10 GHz. This loss effectively performs a low-pass filtering on transmitted radio-frequency or digital signals and hence severely limits the distance that they can propagate on printed circuit board transmission lines.

Passive models for integrated circuit components (spiral inductors, for example) and interconnects (transmission lines, vias, and packages) are having increasing impact on the circuit performance of many high-frequency designs. Circuit designers can employ models for these passive components ranging from simple single lumped elements to rigorous measurement or electromagnetic-based broadband response representations.

Electromagnetic simulation is a cost-effective alternative to physical prototyping and is therefore used throughout the high-frequency modeling community. It is employed to both extract electrical behavior and optimize component performance through optimization. Parametric optimization is often used to reduce the electrical losses associated with silicon-based designs.

Various types of electromagnetic simulation such as finite-element (for full-wave three-dimensional extraction), method of moments (for two-dimensional surface currents), and quasistatic solvers (for lower frequencies) are available to characterize the parasitic effects of transmission media operating at various frequencies. The outputs of these solvers are extracted RLGC (resistance, inductance, conductance, capacitance) models (limited bandwidth), w-elements, or S-parameters. (The w-elements are a set of physical parameters needed for the electromagnetic field analysis of circuit conductors. S-parameters essentially provide the expected broadband transmission and reflection properties of the signal propagating through the given structure.) The outputs are represented by data tables with terminal nodes defining the connection points for input/output signals.

Convolution. Properly addressing linear frequency domain models such as transmission lines and S-parameters in transient time-domain simulations requires a time-domain representation of the linear components or network. One method involves a technique known as convolution.

Direct convolution is a relatively slow mathematical process, even at today's computational speeds. Theoretically, in order for the convolution technique to obtain correct results, the sampling range should be set to infinity. However, this is not practical simulation. A finite sampling range will also lead to excessive computation times. For example, if convolution analysis will calculate the spectral components of the impulse response to be at a bandwidth of 0–200 GHz with a resolution of 100 MHz, the number of frequency sampling points is 2000 and the number of time sampling points is 4000 (0–10 ns with a step of 2.5 ps).

Therefore, an alternate method that does not have the computational burden of direct convolution is desired. Such an alternative mathematical method must still recognize the physical behavior of the frequency-dependent losses that the simulator is attempting to model. One option is to select a "state space" representation for the frequency-dependent model in the time domain, in which a set of rational polynomial functions in the time domain provide a least-squares fit to the frequency-domain behavior after a Fourier transform. Thus, performance of the transient simulation does not call upon convolution. The same technique can be applied for obtaining a usable model from frequency-dependent S-parameter inputs, to avoid demanding convolution during the transient simulation to follow.

One challenge for representing frequency-based models in the time domain is preserving causality. Causality requires that "effects" happen after "causes," which means the impulse response for time $t < 0$ must be zero. Beyond proper handling of losses to conform to causality, a state-space conversion of a transmission-line model can become just as numerically intense as direct convolution if not handled properly. For instance, a given transmission-line model might contain a large number of signal conductors to account for the behavior of each line and the crosstalk between them. Computing state-space models for such transmission matrices without extracting out consistent behavior between like terms

would be exceedingly wasteful. This means first capturing the modal delay or propagation delay from the line behavior, before generating the polynomials that represent the loss behavior.

Validation. As the number and complexity of integrated circuit performance metrics continue to proliferate, accurate design verification becomes a greater challenge and is further complicated by the effects of high-density input-output packaging. System metrics such as bit error rate (BER), error vector magnitude (EVM), or intersymbol interference (ISI) must be characterized by exact standards and specifications: Design details of the integrated circuit, the advanced package, or both must now be incorporated in a simulation. For integrated circuit verification, a high-level system test bench environment is often employed to define complex excitations and derive these figures of merit. In such system test benches, complex modulated communication waveforms may be applied directly to the circuit design. Additionally, package effects and test fixture contributions may also be included, either through electromagnetic simulation or direct measurement of existing hardware. David Vye

Bibliography. G. Assobrio and P. Antognetti, *Semiconductor Device Modeling with SPICE*, 2d ed., McGraw-Hill, 1993; G. De Micheli, *Synthesis and Optimization of Digital Circuits*, McGraw-Hill, 1994; G. De Micheli and M. Sami (eds.), *Hardware/Software Co-Design*, Kluwer, 1996; H. Fujiwara, *Logic Testing and Design for Testability*, MIT Press, 1985; K. S. Kundert, J. K. White, and A. L Sangiovanni-Vincentelli, *Steady-State Methods for Simulating Analog And Microwave Circuits*, Kluwer Academic, Boston, 1990; V. Rizzoli and A. Neri, State-of-the-art and present trends in nonlinear microwave CAD techniques, *IEEE Trans. Microwave Theory Tech.*, 36(2):343–365, February 1988; U. L. Rohde and D. P. Newkirk, *RF/Microwave Circuit Design for Wireless Applications*, Wiley-Interscience, 2000; M. Sarrafzadeh and C. K. Wong, *An Introduction to VLSI Physical Design*, McGraw-Hill, 1996.

Computer-aided design and manufacturing

The application of digital computers in engineering design and production. Computer-aided design (CAD) refers to the use of computers in converting the initial idea for a product into a detailed engineering design. The evolution of a design typically involves the creation of geometric models of the product, which can be manipulated, analyzed, and refined. In CAD, computer models and graphics replace the sketches and engineering drawings traditionally used to visualize products and communicate design information.

Engineers also use computer programs to estimate the performance and cost of design prototypes and to calculate the optimal values for design parameters. These programs supplement and extend traditional hand calculations and physical tests. When combined with CAD, these automated analysis and optimization capabilities are called computer-aided engineering (CAE). *See* COMPUTER-AIDED ENGINEERING; OPTIMIZATION.

Computer-aided manufacturing (CAM) refers to the use of computers in converting engineering designs into finished products. Production requires the creation of process plans and production schedules, which describe how the product will be made, what resources will be required, and when and where these resources will be deployed. Production also requires the control and coordination of the necessary physical processes, materials, equipment, and labor. In CAM, computers assist managers, manufacturing engineers, and production workers by automating many production tasks. Computers are used to develop process plans, order and track materials, and monitor production schedules. They also control the machines, industrial robots, test equipment, and systems that move and store materials in the factory.

Advantages of automation. CAD/CAM can improve productivity, product quality, and profitability. Computers can eliminate redundant design and production tasks; improve the efficiency of workers; increase the utilization of equipment; reduce inventories, waste, and scrap; decrease the time required to design and make a product; and improve the ability of a factory to produce different products. Today most manufacturers employ CAD/CAM to varying degrees. *See* AUTOMATION; PRODUCTIVITY.

Several factors account for the rapid development of CAD/CAM technology since the 1960s. First, advances in computer technology produced faster, more capable, smaller, and less expensive machines. Together with improved techniques for software engineering, this makes CAD/CAM increasingly cost-effective even for small companies. Second, world markets for manufactured goods are more competitive. CAD/CAM has become an essential tool for economic success. Third, the design life cycle of many products has been compressed. New products must be developed quickly and brought to market at high levels of quality, reliability, and safety. *See* SOFTWARE ENGINEERING.

Computer-integrated manufacturing. The deployment of CAD/CAM in the manufacturing industry has led to new ideas about organization and production. Traditionally, manufacturing companies, especially large companies, have been organized into functional units such as marketing, design, engineering, quality, and production. These departments were coordinated by a higher level of management, often without a great deal of direct communication among the units themselves.

The fact that CAD, CAE, and CAM work best together has led to the breakdown of many of the traditional barriers between functional and manufacturing units. The goal of computer-integrated manufacturing (CIM) is a database, created and maintained on a factory-wide computer network, that will be used for design, analysis, optimization, process planning, production scheduling, robot

programming, materials handling, inventory control, maintenance, quality assurance and control, and marketing. Although many technical and managerial obstacles must be overcome, computer-integrated manufacturing appears to be the future of CAD/CAM. *See* COMPUTER-INTEGRATED MANUFACTURING; DATABASE MANAGEMENT SYSTEM.

Developing disciplines in CAD/CAM integration include concurrent engineering, design for manufacture and assembly (DFM/A), and factory modeling. In concurrent engineering, interdisciplinary product development teams consider all phases of a product life cycle during the design process. DFM/A software allows a design team to assess the effort and costs involved in the production of a part or assembly and to consider ways to redesign the product for ease of production. Discrete event simulation programs have been developed to model and optimize the processes, equipment, and layout of the factories in which the design will be produced. *See* SIMULATION.

CAD technology. Initially, CAD workstations were terminals connected to a mainframe computer. They allowed graphic displays of the product model and included devices (mouse, joystick, trackball, light pen, and stylus and tablet) for interacting with both the model and its display. Today a CAD workstation can be a personal computer or a graphics workstation (**Fig. 1**), and may be stand-alone or a link on a network. Large machines (mainframes or supercomputers) are often necessary for creating, manipulating, and analyzing complex models and assemblies. Raster graphics (cathode-ray tubes) permit displays of models in full color, immediate updates as the model is changed, sophisticated rendering, and even animation. Interactive devices include the mouse and keyboard, but for manipulating a three-dimensional (3D) model in all its degrees of freedom, a spaceball or other 3D input device is helpful. Virtual reality and gaming environments have led to useful 3D interactive devices (such as data gloves, body suits, control sticks, and steering wheels). Output devices, such as printers, plotters, and various media and storage devices, can also be linked to the computer. *See* COMPUTER PERIPHERAL DEVICES; MICROCOMPUTER; SUPERCOMPUTER; VIRTUAL REALITY.

The CAD system allows the designer to create a geometric model of the product design, display different views of the model, modify it as desired, save different models in a database for later recall, and make printed copies of design notes, engineering drawings, and rendered images as needed.

The appropriate computer model depends on the type of product. In the design of printed wiring assemblies, for example, two-dimensional models are used to lay out circuits and electronic components on the face of a circuit board. In mechanical and civil engineering design, three-dimensional, geometric models represent parts, devices, assemblies, and structures.

Several types of geometric models are possible, including solid models, wire-frame (line) drawings,

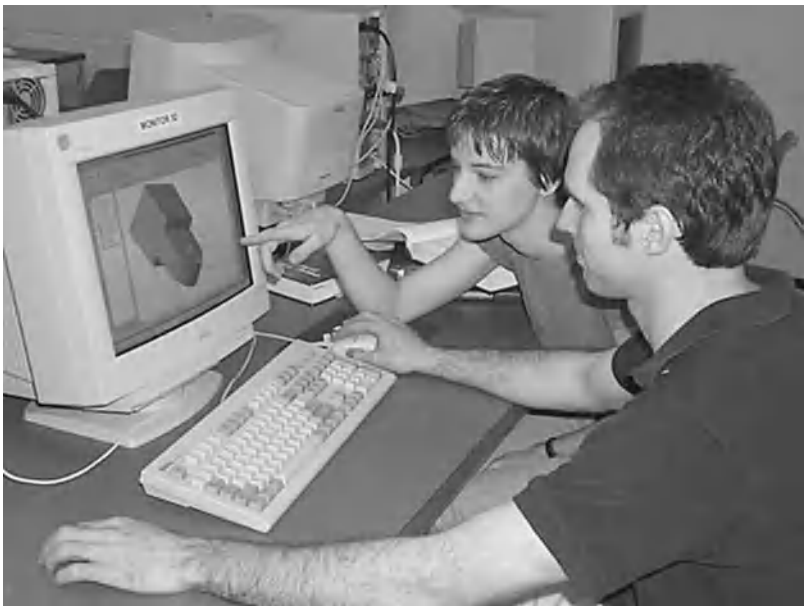

Fig. 1. Computer-aided design workstation where the designer creates and refines a geometric model of a product. (*Center for Computer-Aided Engineering*, *University of Virginia*)

and surface representations, such as planes, patches, or sculptured drawings. Solids modeling is the basis for most current CAD systems, because the information in the model database represents a distinct solid object, completely and unambiguously. There are several modeling schemes for producing complete solid models: boundary representation (b-rep), constructive solid geometry (CSG), cell decomposition, spatial enumeration (octrees), and primitive instancing. Boundary representations define an object in terms of its bounding faces, edges, and vertices, and their relationships. Constructive solid geometry builds complicated solid models by adding or subtracting simpler geometric shapes (cubes, slabs, cylinders, spheres, and cones), called primitives. In cell decomposition, a solid model is composed of a collection of nonoverlapping geometric entities called hyperpatches. Spatial enumeration maps an object into a defined spatial grid and represents whether elements in the grid are filled or empty. Primitive instancing involves parametrized shapes that can be transformed into particular instances. Most commercially available CAD systems are based on boundary representation and constructive solid geometry schemes. Newer solid modelers incorporate design by features, parametric and variational modeling, and constraint-based systems.

Many CAD systems are still used as electronic drafting devices. Although this has many time- and labor-saving advantages, the true benefits of CAD arise from 3D modeling systems. In the traditional drafting environment, several views of a single object must be drawn to capture its geometry, whereas with 3D CAD a single solid model is created from which any particular view can be derived automatically. As the design evolves, each change does not require an entirely new model. In the future, as CAD, CAE, and CAM are integrated in CIM, the CAD model will

become part of a common database for the entire design and manufacturing cycle. *See* COMPUTER GRAPHICS; DRAFTING; ENGINEERING DRAWING.

Analysis programs. A CAD model can be coupled with various analysis programs to predict the behavior of an object or assembly in its operating environment. Finite element analysis (FEA) determines the internal response of an object to applied loads or forces; computational fluid dynamics (CFD) is concerned with flow analysis and visualization; mechanical dynamics programs simulate the behavior of multibody systems (mechanical assemblies) over time; and ergonomic and human factors models assess how people interact with vehicles, machines, and structures. *See* COMPUTATIONAL FLUID DYNAMICS; FINITE ELEMENT METHOD; HUMAN-FACTORS ENGINEERING.

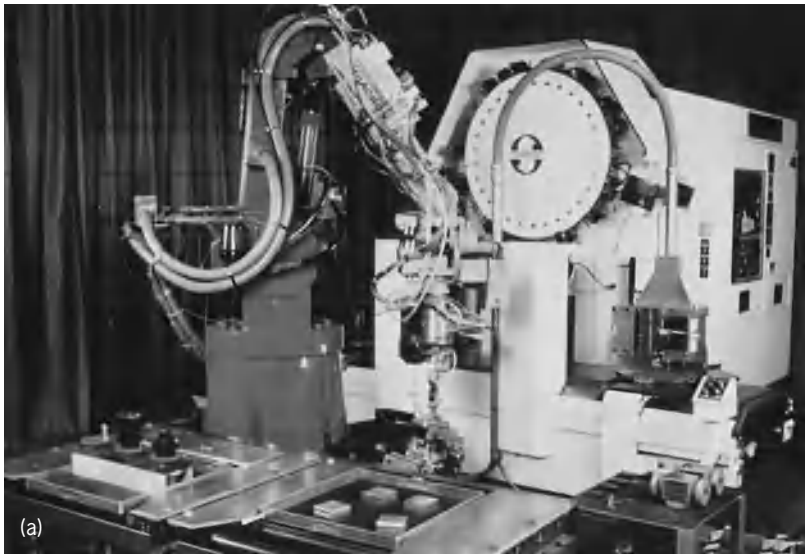

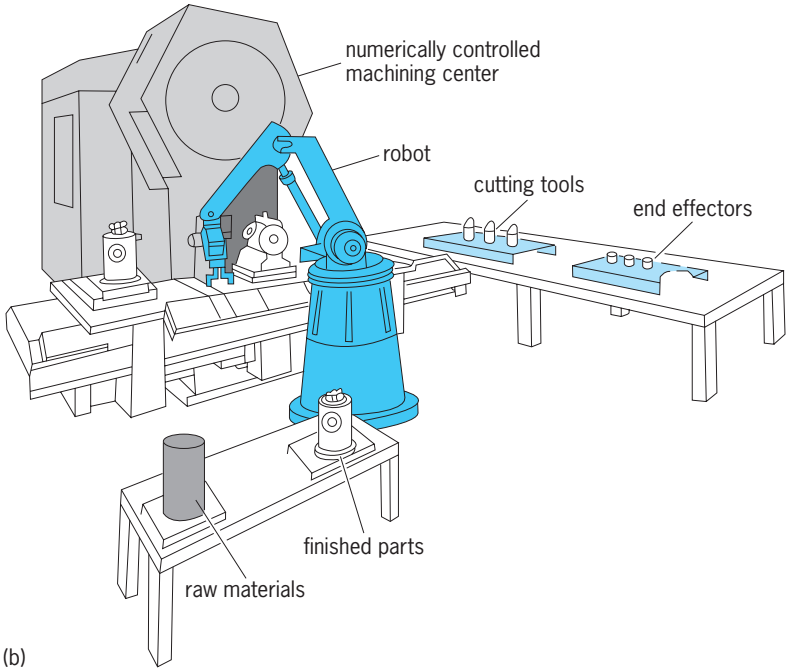

Fig. 2. Horizontal machining workstation. (*a*) Workstation with an industrial robot that loads and unloads parts processed on a numerically controlled machining center. (*b*) Computer simulation display used to verify the robot and part programs which control the workstation. (*Automated Manufacturing Research Facility*, *National Institute of Standards and Technology*)

CAM and numerically controlled machines. The origin of automated manufacturing predates CAM and is usually associated with the introduction of numerically controlled machine tools in the mid-1950s. Machine tools, such as drills, lathes, saws, and milling machines, are used to cut and shape parts from stock. Numerically controlled machine tools use precoded information to operate the machine, rather than a human machinist.

The precoded information, called a program of instructions, contains a detailed set of directions, in symbolic form, that instructs the machine which operations to perform. The program may be punched on paper tape, recorded on magnetic tape, or manually input to the controller unit by a human operator. The more sophisticated numerically controlled systems used in CAM employ direct numerical control, where the program is transmitted electronically to the controller unit over a direct link with the computer.

The numerical controller unit reads the program, interprets the symbols, and sends the appropriate control signals to servomotors and other physical controls on the machine itself. Various sensors send information back to the controller unit. The purpose of this closed-loop or feedback system is to ensure that the actual process is following the program. Most modern numerical control systems are computer–numerically controlled, and the control unit is based on a microprocessor. Microprocessors are also used to monitor and control other types of equipment and processes. The microprocessor control units are called programmable controllers. *See* MICROPROCESSOR; PROGRAMMABLE CONTROLLERS.

Machining centers and cells. Numerically controlled machines vary in complexity from simple, open-loop, tape-controlled drill presses to versatile, multifunction machining centers. A horizontal numerical control machining center is shown in **Fig. 2**, as one component of a still more complex and versatile group of automated machines called a machining workstation or cell. The machining center can perform a variety of functions, such as drilling, boring, reaming, tapping, and milling. The tools required for these operations are stored in a circular tool drum. When the controller calls for a new tool, an automated tool changer removes the current tool from the spindle chuck and returns it to the drum. The drum rotates the new tool to the position in front of the changer, and the changer inserts the new tool in the chuck. The machining center also controls the position and orientation of the part being made, so that it can be machined on all sides.

Automated part programming. Preparing numerically controlled programs is called part programming. Manual part programming is tedious and difficult, because the programmer must calculate and record a long list of the relative positions of tools and parts for each movement along the tool path. Computer-assisted part programming simplifies this task. The

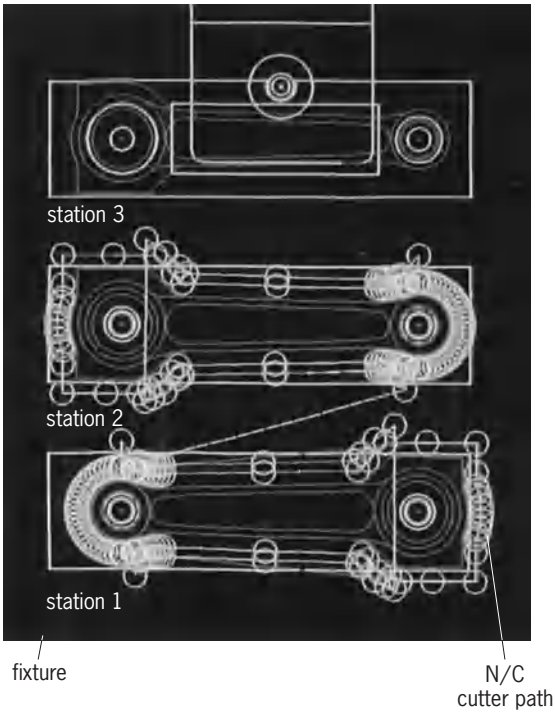

Fig. 3. Computer-aided design and manufacturing display on a system that allows the part programmer to create a cutter location file and preview the cutter path before machining. (*Control Data Corp.*)

programmer specifies the part geometry and machining information (tool sizes, speed, feed rates, and tolerances) in a high-level programming language. The module creates the part program. **Figure 3** shows the display of a numerically controlled machine tool creating a part program for a connecting rod; **Fig. 4** shows this rod being cut out of stock. Automation significantly reduces the time required for part programming and program verification, especially for complex parts with many machining steps. In integrated CAD/CAM systems, the part geometry is specified in the CAD database and does not need to be reentered during programming.

Advances in manufacturing. The development of CAD software for creating true solid models caused designers and manufacturing engineers to think about product models in new ways. They now recognize three approaches to the production of parts: subtractive, additive, and net shape production. Traditional manufacturing relies largely on processes that shape and remove material. Parts are created by subtracting material from stock. Rapid prototyping involves a set of manufacturing techniques based on additive processes; the techniques build up to the desired shape. Several distinct rapid prototyping methods have been developed into commercial systems, including stereolithography (**Fig. 5***a*), selective laser sintering, laminated object manufacturing, ballistic particle manufacturing, fused deposition modeling, (Fig. 5*b*), and 3D printing (Fig. 5*c*). These systems create physical models of parts directly from a CAD database. Such models (prototypes) may be used for presentation, communication, and understanding; as masters for the production of secondary tooling; or for making actual production-quality parts.

The first rapid prototyping system was developed in 1986. The stereolithography apparatus (Fig. 5*a*) uses a CAD driven laser beam to draw individual layers of a part by transforming liquid polymer into solid material using ultraviolet light. It defines the industry standard for rapid prototyping.

Net shape manufacturing techniques produce engineering materials to finish or near-finish dimensions from the start. Working from the CAD model, the goal is to produce a part as close as possible to the final version on the first pass. A variety of fabrication techniques are available: melt processing, shaping from powder, and forming from sheets or billets. Net shape production conserves resources and can reduce costs.

Robotics. Industrial robots are general-purpose, programmable manipulators that perform a variety of production tasks. Robots can move material, load and unload parts on other machines, assemble components, spray paint, and spot-weld. The first industrial robot was installed in a United States automotive diecasting facility in 1961. Figure 2 shows a type of robot that loads and unloads parts and changes tools in a tool drum at a machining center. To spare humans the tasks, robots can do unpleasant, boring, fatiguing, or dangerous jobs, such as working in areas contaminated by fumes, acid, or radiation.

Robots can be reprogrammed for different tasks as needed. In walk-through and lead-through training, the robot arm and end effector (hand or specialized tool) are moved through the desired sequence, and that sequence is recorded in the robot's memory. With a more sophisticated technique, the robot program is created on an off-line computer using a high-level programming language, and then loaded into the robot memory when needed. This eliminates the loss of production time during robot training, and

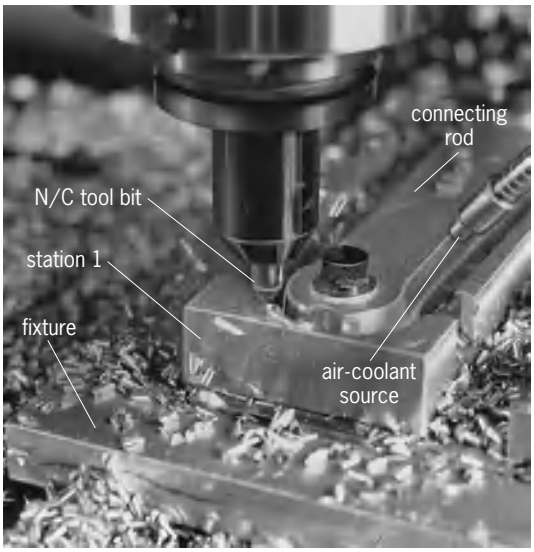

Fig. 4. Numerically controlled machine tool in profiling operation using data from the part program. (*Control Data Corp.*)

Fig. 5. Rapid prototyping systems. (*a*) Stereolithography apparatus, which uses a CAD-driven laser beam to draw individual layers of an object by transforming a photosensitive epoxy plastic into a solid material (*3D Systems*). (*b*) Fused deposition modeling system, which extrudes thermoplastic material from a temperature-controlled head to produce a physical 3D prototype of the CAD model displayed on the graphics monitor (*Stratasys, Inc.*). (*c*) Thermojet 3D printer, whose jets spray a thermoplastic material that builds up in layers to create an object (*3D Systems*).

allows a repertoire of programs to be stored and used interchangeably.

Robots also can be equipped with humanlike senses (vision, touch, and hearing) to perform more complicated tasks. For example, a video camera and a computer equipped with image-processing hardware and software are used for robot vision. This allows the robot to discern objects in the camera field and recognize the position and orientation of familiar objects. In this way the robot can retrieve randomly oriented parts, recognize specific parts in the presence of other objects, perform visual inspections, and assemble products requiring precise alignment. Similarly, touch, force (stress), and proximity sensors let the robot sense contact with another object, gage and control the force of contact (for example, its grip strength), and sense the presence of nearby objects it has not touched. *See* COMPUTER VISION; ROBOTICS.

Automated materials handling. Materials handling refers to the movement, storage, retrieval, and tracking of raw materials, work in process, and completed products. Conveyor belts, hand trucks, fork lifts, and cranes are familiar equipment used to move materials. Shelves, bins, pallets, and silos are used for storage. Inventory sheets, traveler cards, parts lists, and fixed storage locations aid in tracking and retrieving materials.

In CAM, the goal is to automate all materials handling under computer control. Mobile robots, called automated guided vehicles, follow routes on the factory floor defined by wires or tape. Combined with automated pickup and delivery stands, automated guided vehicles move bins or trays of material without human intervention. A typical automated storage and retrieval system uses conveyors and lifts to move bins into and out of specialized inventory shelving. The computer automatically assigns entering bins to empty shelf locations, recording the bin number and shelf location for future retrieval. The bins are labeled with bar codes and can be identified automatically by optical scanners. Large-scale software systems, such as manufacturing resource planning, assist managers in determining what raw materials and components will be required during the current production period. These systems also help the managers to decide when materials and components must be ordered or manufactured to maintain production schedules. *See* MATERIALS HANDLING.

Flexible manufacturing systems. Flexible manufacturing systems are a form of computer-integrated manufacturing used to make small to moderate-sized batches of parts. A typical flexible manufacturing system comprises a cluster of numerical control machine tools, served by an automated materials-handling system and controlled by a central computer. A typical flexible manufacturing system can make from 4 to 100 different kinds of parts in volumes of 40–2000 parts per year, all with minimal human intervention. Such a system is efficient because it is highly automated and requires little downtime for changeovers for making batches of different parts. The flexible manufacturing system is in contrast to the familiar assembly line, which depends for its efficiency on making a single product in very large numbers. *See* FLEXIBLE MANUFACTURING SYSTEM.

Globally distributed manufacturing. With the evolution of the Internet and the World Wide Web,

globally distributed manufacturing is possible. Virtual corporations can be established involving participants from around the world. Design, production, and assembly can occur wherever is most convenient and cost-effective for the enterprise. Developments in CAD/CAM, including advances in product data models and international standards, have made this possible. *See* COMPUTER PROGRAMMING; DIGITAL COMPUTER; ENGINEERING DESIGN; INTERNET; WORLD WIDE WEB. K. Preston White, Jr.; Larry G. Richards

Bibliography. C. R. Asfahl, *Robots and Manufacturing Automation*, 2d ed., 1992; D. D. Bedworth, M. R. Henderson, and P. M. Wolfe, *Computer-Integrated Design and Manufacturing*, 1991; T.-C. Chang, R. A. Wisk, and H.-P. Wang, *Computer Aided Manufacturing*, 2d ed., 1998; C. Machover, *The CAD/CAM Handbook*, 1996; C. McMahon and J. Browne, *CAD/CAM: Principles, Practice and Manufacturing Management*, 2d ed., 1998; P. K. Wright and D. A. Bourne, *Manufacturing Intelligence*, 1988; I. Zeid, *CAD/CAM: Theory and Practice*, 1991.

Computer-aided engineering

Any use of computer software to solve engineering problems. With the improvement of graphics displays, engineering workstations, and graphics standards, computer-aided engineering (CAE) has come to mean the computer solution of engineering problems with the assistance of interactive computer graphics. *See* COMPUTER GRAPHICS.

CAE software is used on various types of computers, such as mainframes and superminis, grid-based computers, engineering workstations, and even personal computers. The choice of a computer system is frequently dictated by the computing power required for the CAE application or the desired level and speed of graphics interaction. The trend is toward more use of engineering workstations. *See* DIGITAL COMPUTER; MICROCOMPUTER.

Design engineers use a variety of CAE tools, including large, general-purpose commercial programs and many specialized programs written in-house or elsewhere in the industry. Solution of a single engineering problem frequently requires the application of several CAE tools. *See* ENGINEERING DESIGN.

A typical CAE program is made up of a number of mathematical models encoded by algorithms written in a programming language. The natural phenomena being analyzed are represented by an engineering model. The physical configuration is described by a geometric model. The results, together with the geometry, are made visible via a user interface on the display device and a rendering model (graphics image). *See* ALGORITHM; COMPUTER PROGRAMMING; MODEL THEORY; PROGRAMMING LANGUAGES.

CAE allows for many more iterations of the analysis-design cycle than was possible by hand computation, especially when the CAE is coupled with optimization systems that drive this cycle automatically. The benefits are translated into improved productivity and quality of design.

Program structure. A CAE program usually consists of a series of mathematical models and a data structure. **Figure 1** illustrates a simplified view of a typical CAE program operating in an engineering workstation environment. First, a mathematical description of the physical phenomena being analyzed is written. This engineering model may consist of equations such as Newton's second law to describe the dynamics of a system or the Navier-Stokes equations to analyze a fluid flow field. Next, a model of the physical configuration is created. This geometric model may consist of two- or three-dimensional (2D or 3D) curves, surfaces, faceted approximations to surfaces, or solid elements. The results of the engineering analysis are frequently displayed on the geometric model by color fringing to show the variation of a scalar parameter. Large amounts of data are created during this modeling phase, and the need for a data structure to store and retrieve them is greater than for the engineering model. *See* DATA STRUCTURE.

Although the engineering and the geometry are fully described, they cannot be viewed on the display until a model for rendering has been formulated and coded. A mathematical description of the lighting conditions, the approximate intensities of

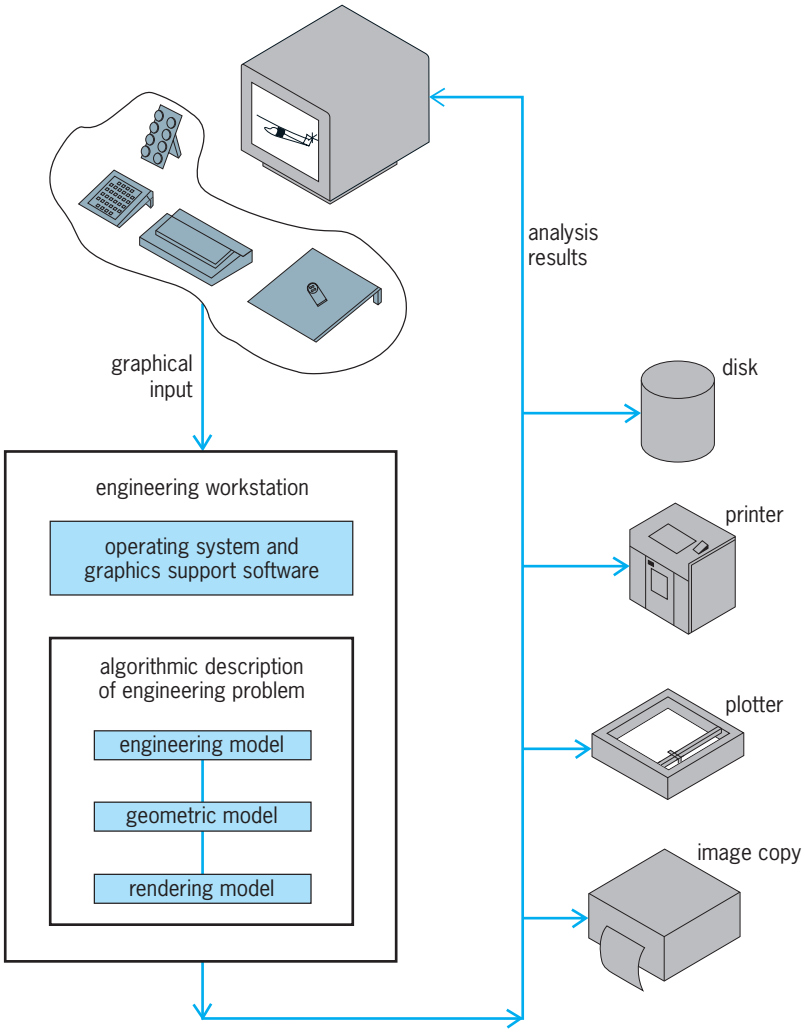

Fig. 1. CAE program in an engineering workstation environment.

light reflected by the nodes of the geometric model, and the corresponding shades of color provide the graphics data necessary for the model to be realistically shaded. Before viewing, the graphics data are transformed from geometric model coordinates to a normalized coordinate system. Parameters are set for an orthographic or perspective projection from 3D model coordinates to 2D coordinates for final transformation to the display screen in device coordinates. The CAE applications programmer accomplishes most of these tasks with the help of precoded algorithms provided by graphics-support software. Interactive communication with the graphics image and the geometric and engineering models occurs through a user interface written with the assistance of the graphics software or windowing software provided with the workstation.

In the past, CAE programs have been predominantly coded in the FORTRAN programming language. The present trend is toward the C, C++, and Java programming languages together with the UNIX and Windows operating systems. *See* OPERATING SYSTEM.

Graphics standards. Portability, the ability to move programs easily from one computer to another, is important for CAE software. Although it has long been possible to make a computational program code portable by using standard programming languages, this was not previously possible for CAE software because of the lack of graphics standards. A proposed 3D graphics standard (CORE), introduced to the American National Standards Institute (ANSI), was superseded by the adoption in 1985 of the 2D Graphical Kernel System (GKS) as an international standard by the International Organization for Standardization (ISO).

A new 3D device-independent graphics standard, the Programmer's Hierarchical Interactive Graphics System (PHIGS), was proposed by ANSI in 1985 and adopted as an international standard by ISO in 1988. Where possible, the concepts and nomenclature of GKS have been used in PHIGS.

The principal limitation to CAE software portability has been the wide variety of graphics hardware and the direct dependence of CAE software on this hardware. Both GKS and PHIGS give the programmer device-independent graphics primitives and coordinate systems as well as a set of logical graphics input devices to replace the wide variety of input hardware. For example, the pick action (selection of a graphics entity on the screen) could be physically accomplished by a light pen, a cursor and tablet, or a mouse. Using graphics standards, the CAE programmer will always specify a logical pick device regardless of the physical device used to achieve the pick.

PHIGS has several important advantages over GKS for CAE software. It is a full 3D system for viewing and modeling transformations, and allows both graphics and nongraphics data to be stored in its data structure. The data structure can invoke other structures and store transformations as attributes. The result is a hierarchical graphics data structure well suited for animation and representation of entities with multiple components. A number of extensions to PHIGS, called PHIGS PLUS (PHIGS+), were adopted as a separate international standard in 1990. PHIGS+ provides support for most of the rendering model and some of the geometric model in a CAE program. PHIGS+ routines address lighting, shading, hidden surface elimination, transparency, nonuniform rational B-spline (NURBS) curves and surfaces, and improved user interaction and control. PHIGS and PHIGS+ were merged into a single international standard, PHIGS, in 1997.

The emergence of PHIGS+ was in response to dramatically improved computer graphics hardware capabilities in the late 1980s and early 1990s, and that were rapidly being incorporated into the new breed of CAE workstations. Vendors did not wait for international committees to address these new capabilities in their standards, but deployed their own proposed extensions in the interim. While this made these new capabilities available to the CAE software developers, it also prevented the CAE software from easily being ported from one hardware platform to another, which is the main purpose of a standard.

Taking advantage of this noncompliance to the adopted international standard and the slow rate at which new technologies and capabilities were addressed by these standards, an industry consortium was established in 1992 to support a new graphics standard known as OpenGL. The aim of OpenGL is to provide a timely integration of new technologies and capabilities as they emerge and to avoid the problem of partial conformance to the standard. GKS, PHIGS, and other standards have been plagued by partially conforming deployments, which makes portability between platforms difficult and defeats a key purpose of standards. The OpenGL Architecture Review Board, which draws its membership from nine leading computer graphics hardware vendors, avoids this problem by permitting only the term OpenGL to be applied to products that conform completely to the OpenGL standard. OpenGL is today the most widely used and supported 2D and 3D graphics application programming interface (API).

Hardware. Computer systems ranging from workstations to mainframe computers are used for CAE software. For example, the aerospace and automotive industries make wide use of very large mainframe computers and large grid-based computer systems to support computer-intensive CAE software such as finite element analysis and computational fluid dynamics. These large computational resources are connected by high-speed data transmission lines to file servers holding the model and resulting data. The precomputation model description and the postcomputation interrogation of the results, and frequently the computations themselves, are usually performed on ultra-high-end workstations capable of managing these vast amounts of data. These workstations are typically 64-bit multiprocessor systems, with 12 to 24 gigabytes (GB) of shared memory

(RAM) and a UNIX operating system, which are connected to the file servers via one or more 1- or 10-gigabit/second network connections. Ultra-high-end workstations are necessary when working with large, complex models and datasets, such as the solid modeling of a complete automobile or aircraft.

For less complex models and datasets, or for subsets of the larger models and datasets, it is typical to use high-end engineering workstations. These workstations are typically 32-bit single- or dual-processor systems with up to 4 GB RAM running a Windows or UNIX operating system, the former being more common. These engineering workstations distinguish themselves from personal computers by having more RAM and faster central processing units (CPU), graphical processing units (GPU), hard disks, network connections, and internal interconnects and data transfer rates. These high-end engineering workstations cost two or three times more than a typical personal computer and are commonly used by engineers. In contrast, ultra-high-end workstations typically cost three to five times more than a high-end engineering workstation, and consequently are less used.

Computer-intensive engineering problems are frequently distributed across multiple CPUs. These CPUs can be connected in a grid structure where each CPU has its own RAM and where communications between these CPUs and their RAM takes place over a computer network. Occasionally these grids are composed of regular personal computers and engineering workstations that are accessed whenever they are idle, such as during nighttime hours. The bottlenecks of such systems are the limited bandwidth between the CPUs and the personnel time required to configure and manage these grids. Grid systems tend to be most successful when the computations can be easily parallelized and when there is minimal need to communicate and move data between the CPUs. Computational problems that do not fall into this category are best performed on large multiprocessor, shared-memory systems with high-speed, high-bandwidth interconnects between the CPUs. Examples of such systems include mainframes and the ultra-high-end workstations described earlier. The advantages of these systems include the flexibility of how data can be accessed simultaneously by multiple CPUs without having to be moved across a network, and the reduced need for manually managing the distribution of the computational resources within the system. *See* MICROPROCESSOR; MULTIACCESS COMPUTER; MULTIPROCESSING; SUPERCOMPUTER.

These large centralized computational resources are often connected to networks of from 10 to 100 client workstations that act as display terminals with local controllers (**Fig. 2**). When a large number of CAE users invoke computer-intensive programs, a significant delay in response usually occurs. For this reason, as well as initial cost considerations, most companies today deploy high-end engineering workstations on a two- or three-year upgrade cycle.

Fig. 2. Solid modeling and simulation on a mainframe via a CAE workstation.

Though many small CAE applications programs are written for personal computers, of which most engineering companies have a wide array, most large, commercial CAE programs do not perform effectively or at all on personal computers. Engineering workstations, on the other hand, have the necessary computer power and local resources, and usually run only one CAE program at a time and serve only one user. As a result of this independence, a demanding CAE code at one workstation node in a computer network will not affect the response time for other users. *See* DISTRIBUTED SYSTEMS (COMPUTERS); LOCAL-AREA NETWORKS.

The personal-computer gaming industry has been instrumental to the recent widespread availability of affordable, high-quality graphical processing units that are capable of supporting the heavy demands of CAE graphics processing. These graphical processing units contain a number of specialized high-speed, parallel graphics engines, display list memory, image memory, buffer memory, and alpha memory to permit the real-time generation of color-shaded images desired for CAE applications. The graphics engines perform graphics modeling and viewing transformations, lighting and shading computations, and scan conversion of models for raster-scan display. Display list memory retains the rendering model structure. The z-buffer provides fast hardware support for hidden surface elimination, whereas alpha memory allows fast hardware-assisted transparency computations. The final rasterized image is stored in one of two image memories for fast transmission to the display screen. Most graphical processing units support the OpenGL graphics standard by performing its operations directly in hardware as opposed to computing them in software. *See* COMPUTER ARCHITECTURE.

The graphical processing unit used in a typical personal computer accounts for approximately 2–4% of the personal computer cost. In contrast, in a high-end engineering workstation, the graphical processing unit typically accounts for 20% of the hardware cost. Where ultra-high-end workstations are also used for graphics processing and visualization, it is increasingly common to use multiple high-end graphical processing units working in parallel, with each unit capable of computing 750 million triangles per second or 8 billion pixels per second.

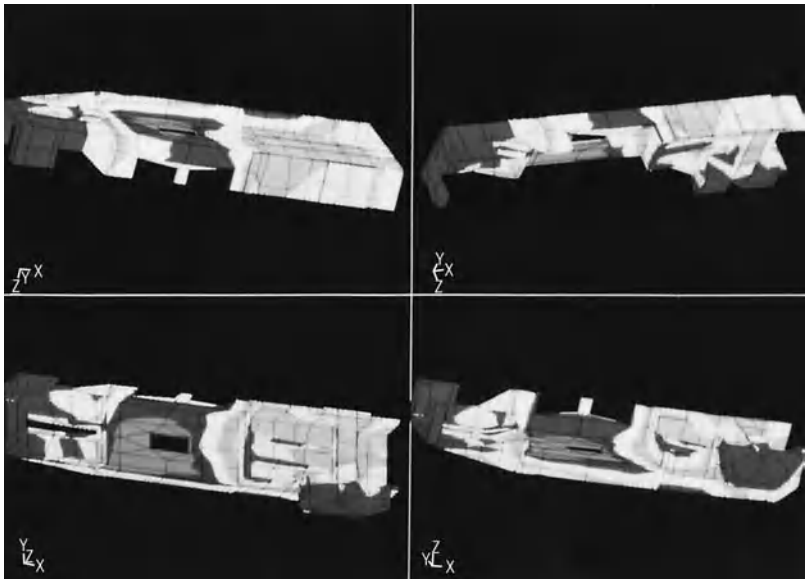

Fig. 3. Finite-element stress analysis using commercial CAE software.

A wide variety of output devices are available for CAE images. The most popular are ink-jet and laser printers for color images and line drawings on paper and Mylar. Others include wall projection and stereovision display systems, often combined with positional and orientation input from the observer. Many types of graphics input devices are available for workstations and mainframe displays. Cursor and tablet, mouse, joystick, ball, glove, function key box, and dial box are the most widely used. Light pens are generally in use on older-technology displays. *See* COMPUTER PERIPHERAL DEVICES.

Commercial CAE software. Most companies use a combination of a few commercial CAE programs and a number of smaller, specialized CAE applications programs typically written in-house or elsewhere in the industry. There are several classes of commercial CAE software. The first large commercial programs used for CAE were dynamic simulation systems, which usually had no graphical output or at best a printer-plotter output of graphs. They required the engineer to give input in the form of ordinary differential equations or building-block elements to describe the dynamic behavior of a real system, in many ways similar to the patching (programming) of analog computers. These systems numerically integrate the set of coupled differential equations, thereby simulating the dynamic behavior of the particular system. Simulation systems are still widely used, and although the graphics output has improved considerably, the description of the input has not changed greatly. *See* ANALOG COMPUTER; NUMERICAL ANALYSIS; SIMULATION.

A need to solve detailed problems, such as stress or deformation analysis, for components of systems led to the development of the finite-element method or finite-element analysis. In this method, a complex object is broken down into simpler elements. With these, a set of equations is formulated which, when solved, predict the behavior of the object as modeled by the set of elements. The modeling of the object is known as finite-element preprocessing, and the validity of the solution is largely dependent on the skill of the modeling engineer. Although very advanced methods of interactive graphics are used in finite-element preprocessing, it is still a very tedious and expensive operation. The solution of the equations is highly computer-intensive. Finite-element methods are applicable for a wide variety of physical phenomena, including mechanical stress and strain, fluid flow, acoustics, heat transfer, and electrical fields. **Figure 3** illustrates the solution of a finite-element stress analysis problem with a commercial CAE code running on a mainframe computer. *See* FINITE ELEMENT METHOD.

Computer-aided design and manufacturing. CAD/CAM systems were created by the aerospace industry in the early 1960s to assist with the massive design and documentation tasks associated with producing airplanes. By the late 1970s, these codes were being distributed to other industries. CAD/CAM systems have been used primarily for detail design and drafting along with the generation of numerical control instructions for manufacturing. Gradually, more CAE functions are being added to CAD/CAM systems. A trend toward open architecture with flexible geometry interfaces is stimulating the addition of more analysis and manufacturing functions. Modeling with CAD/CAM systems has become fairly sophisticated. Most popular commercial systems support 2D and 3D wireframe, surface models, and solid models. Rendered surface models differ from solid models in that the latter have full information about the interior of the object. For solid models, a combination of three types of representation is commonly used: constructive solid geometry, boundary representation, and sweep representation. Complex systems require significant amounts of processing power, primary memory space, and disk space. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Integration of CAE software. Frequently, all of these CAE tools are needed together with specialized CAE applications programs to solve a single engineering problem. The integration of these tools or the communication of data between them is a challenging problem. To successfully integrate these programs, a centralized database is necessary. Data from the various CAE tools are processed into and retrieved from this database through a database management system. If the CAE software is proprietary and the data structure not easily accessible, the proprietary programming interfaces may be used or, sacrificing of functionality and interactivity, a data interchange standard may be used. *See* DATABASE MANAGEMENT SYSTEM.

The Initial Graphics Exchange Specification (IGES) was developed under the leadership of the National Bureau of Standards (today known as the National Institute of Science and Technology) and was accepted as a standard by ANSI in 1981 (Y14.26M-1981). The goal of IGES is to allow the transfer of geometric data between dissimilar CAD systems. IGES is widely available and is generally capable of providing the translation of a snapshot of the model in one CAD

system into a static model that cannot be edited in another CAD system.

The standard for exchange of product model data (STEP) has been adopted as an international standard by the International Organization for Standardization (ISO 10303). The aim of STEP is to provide a neutral, computer-interpretable representation and description of product data throughout the life cycle of a product that is independent from any particular system. It can be used for both data exchange and for archiving data over time. The latter is an important issue when the product life cycle exceeds the life span of the software and hardware that created the product data. Given this extremely broad objective, ISO 10303 is not a single standard but a collection of interrelated documents that form a multipart standard. A number of these documents have been adopted as an international standard, while many others are still in development. The **table** lists of some of the documents that have been adopted as part of ISO 10303.

The software industry's support for adopted ISO 10303 documents has been significantly slower than was the case for PHIGS (graphics standards). In the case of STEP, the software vendors have been hesitant to support the standard, both because of the high upfront cost associated with complying with the complexity of the standard, and because of their desire to provide a comprehensive yet proprietary solution for the customer in which the customer's data is locked into. Thus far, it is primarily the large mechanical CAD/CAM/CAE vendors serving the automotive industry that have given into customer pressure and competition to provide support for STEP (AP 203, AP 214). Recently it appears that AP 210 and AP 212 have also started to receive some initial vendor support.

Adopted ISO 10303 documents

Document number	Title
STEP AP 201	Explicit Draughting
STEP AP 202	Associated Draughting
STEP AP 203	Configuration Controlled Design
STEP AP 204	Mechanical Design using Boundary Representation
STEP AP 207	Sheet Metal Die Planning & Design
STEP AP 209	Composite and Metallic Structural Analysis and Related Design
STEP AP 210	Electronic Assembly, Interconnection and Packaging Design
STEP AP 212	Electrotechnical Design and Installation
STEP AP 214	Core Data For Automotive Mechanical Design Processes
STEP AP 215	Ship Arrangement
STEP AP 216	Ship Moulded Forms
STEP AP 218	Ship Structures
STEP AP 224	Process Planning Using Machining Features
STEP AP 225	Building Elements Using Shape Rep
STEP AP 227	Plant Spatial Configuration
STEP AP 232	Technical Data Packaging: Core Info & Exch.

Applications. The CAE methods for electrical and electronics engineering are well developed. The geometry is generally two-dimensional, and the problems are primarily linear or can be linearized with sufficient accuracy. Chemical engineering makes extensive use of CAE with process simulation and control software. The fields of civil, architectural, and construction engineering have CAE interests similar to mechanical CAE with emphasis on structures. Aerospace, mechanical, industrial, and manufacturing engineering all make use of mechanical CAE software together with specialized software.

An example of CAE is the design of an aircraft landing-gear mechanism. The first step is definition of the problem and creation of a set of performance specifications. Next, the conceptual design phase may be aided by specialized programs to determine a size estimate of the landing gear based on specified loads and deflections. Commercial CAE programs are available for kinematic synthesis of mechanisms based on specified motion requirements, but this landing-gear mechanism will be designed with an in-house program written for the purpose. *See* LANDING GEAR.

The next phase is preliminary design. An applications program is used to analyze the deflection and response of a shock-absorbing, energy-dissipating strut. Dynamic analysis of the guiding mechanism and complete assembly is determined by a commercial code. When the dynamic loads are determined, a finite-element stress analysis of each link of the mechanism is done by using commercial finite-element-method software. Following the stress analysis, some links are changed in size and the dynamic analysis repeated to determine new loads. Using the new loading, another iteration of the finite-element-method software is made to verify that the stresses fall below the strength limits.

The next phase is the final design. All components of the assembly are drawn in 2D on a CAD/CAM system and detailed, giving dimensions, material specifications, and other instructions. An assembly drawing is created from the components, and mating of components are verified. An alternative approach is to create a solid model of each component, assemble the solid components, and run an automatic interference-clearance check; 2D drop-offs are then automatically made of each component and manually detailed (at the workstation) by a drafter. From the final part geometry, instructions for numerically controlled machine tools are generated to produce the part. Some systems may support tooling design and process planning. Finally, a design release is made to the manufacturing department. Jan Helge Bøhn; Arvid Myklebust

Bibliography. J. D. Foley et al., *Introduction to Computer Graphics*, 1993; T. Gaskins, *PHIGS Programming Manual*, 1992; M. K. Gillenson, *Database, Step-by-Step*, 2d ed., 1990; F. R. A. Hopgood et al., *Introduction to the Graphical Kernel System (GKS)*, 2d ed., 1987; M. E. Mortenson, *Geometric Modeling*, 2d ed., 1997; D. F. Rogers, *Pro-*

cedural Elements for Computer Graphics, 2d ed., 1997; D. Shreiner et al., *OpenGL Programming Guide*, 5th ed., 2005; A. Tizzard, *An Introduction to Computer-Aided Engineering*, 1994.

Computer architecture

The art or practice of designing computer systems. Just as the architecture of a building describes its overall structural concept and principal features, the architecture of a computer system consists of a description of the overall layout and major features of a computer system. In both cases, we are concerned with how the object, whose architecture we are describing, appears to the user. In computer architecture, the user is not the final user of the system, but the person who writes computer programs to be run on the system. Compared to the architecture of buildings, which is largely an issue of esthetics, the architecture of a computer system is principally about efficiency and economy of operation. There may be an esthetic dimension to computer architecture, but when present it is distinctly lower in importance.

At present, computer architecture is the highest level of the closely related field of computer design. The field has evolved from being the top level of the design process used by the early developers of computers into a separate area of specialization within the overall hierarchy of computer designers and engineers. Underneath computer architecture are the fields of computer organization and implementation, and computer engineering. *See* COMPUTER; DIGITAL COMPUTER.

Computer organization and implementation. Closely related to the field of computer architecture is a subsidiary discipline known as computer organization and implementation. It is difficult to define precisely where computer architecture stops and computer organization and implementation begins. In general, computer architecture is concerned with the principal design parameters, while computer organization and implementation is concerned with the subsidiary points of design of the system, including the composition of the logic circuits that make up each structural component of the system and the implementation of each of those logic circuits as an arrangement of specific electrical circuits of some particular type or types. Computer organization focuses on issues such as whether an adder circuit will be implemented as a ripple-carry adder or as a carry-lookahead adder and whether the underlying electrical circuits will be implemented in TTL (transistor-transistor logic), bipolar transistors, CMOS (complementary metal-oxide semiconductors), or gallium arsenide (GaAs). *See* INTEGRATED CIRCUITS; LOGIC CIRCUITS.

Major issues. There are several major issues in the design of a computer system covering all or nearly all of the architectural considerations. *See* MICROCOMPUTER.

Instruction-set architecture. Traditionally, the instruction-set architecture—the detailed design of the set of machine instructions implemented in the processor—has been the core issue in defining the architecture of a computer system. For example, how many machine instructions will there be, will the instructions be of uniform length or variable in length, and how many fields will the machine instruction contain. *See* MICROPROCESSOR.

Registers. The typical central processor contains several special-purpose registers (storage units), including an instruction register, a program counter or instruction pointer, a processor status register, perhaps one or more separate index registers, segment (base) registers, one or more general-purpose registers to hold integer or fixed-point data (integer registers), and either the same or a distinct set of registers to hold floating-point data. Although integer registers, a single instruction register, and the program counter or instruction pointer have traditionally been of the same width (number of bits), there is no law requiring such. Floating-point registers may or may not be of different width from integer registers. *See* COMPUTER STORAGE TECHNOLOGY.

Native data types. These are the data types that will be separately and distinctly represented and recognized by the computer hardware. Some computers may natively represent and carry out arithmetic on decimal integers via an underlying binary representation. Floating-point representation may or may not be present. *See* NUMBERING SYSTEMS.

Word size. The most commonly encountered number of bits contained in a single word of memory (word width) are 32 and 64. *See* SEMICONDUCTOR MEMORIES.

Memory addressing scheme. There are two principal issues under the topic of memory addressing scheme. The first is the smallest unit of addressable memory, which is usually the byte, although there are several machines of major significance that are not byte-addressable but word-addressable. The second issue is the order in which the bytes are stored for the several bytes that constitute a single word. In the scheme known as big-endian, the leftmost byte (that is, the byte with the numerically smallest address) contains the most-significant bits of the word content, while the byte containing the bits of least significance possesses the numerically highest address. The opposite arrangement is known as little-endian.

Maximum size of memory. A major issue for the modern computer architect is how large the memory address space should be, since memory has been getting cheaper at a prodigious rate. For every doubling of size of the memory address space, the width of an address increases by one bit. Therefore, the designer is not free to increase the memory address space by an arbitrary amount, as every increase in address width has much wider implications for instruction set architecture, as well as imposing requirements affecting lower levels of computer organization and implementation. On the other hand, insufficiency of memory address space has been one of the principal factors causing an architecture to become outmoded.

Number of buses and their relationship. A modern com-

puter typically has a frontside bus and a backside bus, as well as one or more subsidiary busses. The backside bus connects the registers in the central processing unit (CPU) with the main memory and with the various levels of cache memory. The frontside bus constitutes the main conduit between the CPU and peripheral devices, and one or more subsidiary buses interconnect various groups of peripheral devices. Each subsidiary bus is designed to accommodate the needs of peripheral devices within a particular speed range. A simpler and therefore cheaper computer uses a smaller number of subsidiary buses, and each such bus may operate at a lower efficiency while it is communicating with a peripheral device that is considerably slower than the maximum speed that the bus is capable of. *See* COMPUTER PERIPHERAL DEVICES.

Datapath width. This is the number of bits involved in a single transfer of data within or between units within the central processor, such as its internal registers.

CISC versus RISC. Historically, the mainstream of computer architecture was monopolized by complex instruction-set computers (CISC). During the evolution of CISC machines, a lot of the ingenuity of the designer lay in coming up with a larger repertoire of machine instructions, of which the more recently developed tended to be instructions of ever-greater capability and complexity. However, there is a price to pay for the existence of the extra, more powerful instructions. A more elaborate decoding circuitry is required in the central processing unit to interpret these instructions, or there must be a distinctly longer and wider within-processor special-purpose memory to represent the microprograms for all of the instructions contained in the machine's instruction set architecture (ISA).

During the early 1960s, research carried out at IBM suggested that an alternative approach in which the size of the instruction set was actually diminished to a central core of frequently used instructions might be more efficient. In a reduced instruction-set computer (RISC), the area of the processor chip that is saved by omitting either the extra decoding circuitry or the extra length and width of the microprogramming store necessitated by the presence of the omitted more complex instructions is used instead to increase the number of registers. As a result, there are two forms of time savings involved in the execution of the remaining instructions. First, the smaller size and simplified circuitry of the arithmetic logic unit (ALU) enables this component to significantly faster execute the remaining instructions. Second, the presence of a much larger number of registers enables intermediate data to be retained in a register for significantly longer periods of time, reducing both the number of write operations from registers to memory and read operations from memory to registers, compared to the numbers of such operations that must be performed in a CISC processor because of the need temporarily to vacate a register to enable it to do something else. In fact, a great part of the savings in processing time for the RISC processor is due to the substantial reduction in transfers to and from main memory that is occasioned by the much larger number of registers present in the RISC processor. Among the CISC instructions that are omitted in the RISC processor are those that operate directly upon or that write out their results directly to memory. Instead, there are only two instructions provided to accomplish memory transfer operations: a LOAD instruction and a STORE instruction. For this reason, a RISC processor is sometimes referred to as a Load/Store processor. As with any other design decision, a price must be paid performance-wise in exchange for the gain in execution speed. The remaining, "core" instructions that are still present in the RISC processor are executed with a dramatically improved speed. However, the omitted instructions, when they otherwise would have been called for by the source program, must each be emulated by the compiler by means of a series of shorter instructions. As a result, the execution time for one of these instructions omitted from the processor's instruction set will be dramatically longer than it otherwise would have been. For some source code programs, such as those that used very few of the omitted instructions, and those relatively rarely, there will be a net time savings. But for other source code programs, such as those that used relatively heavily and frequently the instructions that have been omitted, the performance will be slower. *See* COMPUTER PROGRAMMING.

RISC remained largely out of the mainstream of computer architecture until the 1990s, when it began to compete seriously with CISC computers. Some of the pioneers in this area include the IBM RS/6000 and the Digital Equipment Corporation Alpha, followed by a number of other products by these and other vendors. Currently, the CISC and RISC streams of processor development have shown some convergence, with microprocessors of predominantly one or the other design incorporating some features previously associated with the other stream.

It should be noted that the earliest supercomputers designed by the famous computer architect and designer Seymour Cray used some of the same principles of design that we know today as RISC. *See* SUPERCOMPUTER.

Pipelining. Pipelining is a design approach to increase the rate of instruction execution. The traditional operation of a computer processor was described as the fetch-execute cycle or more precisely as the fetch-decode-execute cycle. In the first or fetch portion of the cycle, the instruction pointed to by the instruction pointer is copied into the instruction register. Once it has been copied, electronic circuitry within the processor decodes the instruction, that is, determines what sequence of steps must be performed in order to execute the instruction. An example of such a sequence is to (1) copy the contents of memory location A to register temp01, (2) copy the contents of memory location B to register temp02, (3) add the contents of temp01 and temp02 and write the result to register temp02, and

(4) copy the contents of register temp02 to memory location C. Once the sequence of actions has been identified, it must be carried out. This is the execute portion of fetch-decode-execute. In a pipeline machine, these three phases of executing a machine instruction can be done on three or more instructions in parallel. That is, while the first instruction of a sequence of machine instructions is being executed after having first been fetched and decoded, the second instruction, having previously been fetched, is simultaneously undergoing decoding, while the third instruction in the sequence is being fetched. In addition, the execution of a single instruction usually involves several component steps, and therefore a pipeline can be constructed with more than three stages, thus making it possible for more than three instructions to be present simultaneously in the pipeline at various stages of processing. This last technique is known as super-pipelined architecture.

Other techniques for increasing the processing speed. In addition to instruction pipelining, several techniques can be used to increase the processor speed, principally by increasing the extent to which different machine instructions are executed in parallel. These include branch prediction, speculative execution, and superscalar architecture.

Multiprocessor core. In the competition for higher processor speeds, one of the more recent design techniques is to include within the processor more than one core unit capable of carrying out arithmetic-logic operations in parallel. A substantial speedup is possible, even though the presence of other structures within the overall framework of the processor might not allow the various cores each to operate at full speed, making one or more of the processor cores have to wait for some commonly used resource to become available. *See* CONCURRENT PROCESSING; MULTIPROCESSING.

Compiler as adjunct to processor. All of the techniques that are used for increasing the parallelism of instruction execution possess a common restriction on their application; that is, they can succeed in producing an improvement in performance only with the active cooperation of the compiler to define precisely where in the program each of them can be applied. For example, pipelined processing is successful in increasing processor speed only for those parts of the program where the precise order of instruction execution is certain. When it happens, as on occasion it must, that an unexpected branch of program execution must occur, then many of the instructions already in the pipeline, some of which might have already been partially executed, must all be flushed from the pipeline, and the pipeline must be reloaded from the point to which the program has unexpectedly branched. In fact, in some number of instances it might even be necessary for already executed instructions or parts of instructions to be undone. Thus, there are distinct limits to the extent to which parallel processing can be successfully deployed to increase processor throughput. As the degree of parallelism within the processor has increased with the development of more advanced architectural techniques, the role of the compiler in using the additional opportunities for parallelism has become ever more important. Thus, the modern computer architect must be knowledgeable in the design and construction of compilers as well as of computer hardware.

Economics. What drives computer design more than any other factor is economics. It is always possible to make improvements in the design of a computer. But the main factor that architects and designers must consider is the importance of being able to build a machine for a price that will return a reasonable profit in the market for which it is targeted.

Historically significant computer architects. There have been many computer architects of distinction. John Vincent Atanasoff invented the stored-program computer in something akin to its current form. Even so, the computer architecture he originated is often referred to as the "von Neumann" architecture. J. Prosper Eckert and John W. Mauchly were the designers of the electronic integrator and computer (ENIAC) built at Aberdeen, Maryland, in 1946. They also designed the binary automatic computer (BINAC), the first machine to use magnetic tape instead of punched cards; and the first model of the universal automatic computer (UNIVAC), which was the first machine able to handle both characters and numeric information. Seymour Cray was the inventor of an entire genre, the supercomputer, and was certainly one of the most prolific computer architects of all time. Gene M. Amdahl, in addition to having originated the law of computer architecture that bears his name, was the principal architect for the IBM Series 360 series of mainframes, which to date is the longest-lived computer architecture. The first machines of this series were introduced in 1963, were extended to become the System 370 and the System 390, and are still being produced. Gordon Bell was the principal architect of many members of the PDP series at Digital Equipment Corporation and oversaw the development of its VAX series of machines.

Charles Abzug

Bibliography. R. J. Baron and L. Higbie, *Computer Architecture*, Addison-Wesley, Reading, MA, 1992; R. J. Baron and L. Higbie, *Computer Architecture: Case Studies*, Addison-Wesley, Reading, MA, 1992; J. L. Hennessy and D. A. Patterson, (with contributions by K. Asanovic and D. Goldberg), *Computer Architecture: A Quantitative Approach*, 3d ed., Morgan Kaufmann, San Francisco, 2003; M. D. Hill et al. (eds.), *Readings in Computer Architecture*, Morgan Kaufmann, San Francisco, 2000; M. J. Murdocca and V. P. Heuring, *Principles of Computer Architecture*, Prentice Hall, Upper Saddle River, NJ, 2000; L. Null and J. Lobur, *The Essentials of Computer Organization and Architecture*, Jones & Bartlett, Sudbury, MA, 2003; W. Stallings, *Computer Organization and Architecture: Designing for Performance*, 6th ed., Prentice Hall, Upper Saddle River, NJ, 2003; H. S. Stone, *High-Performance Computer Architecture*, 3d ed., Addison-Wesley, Boston,

Computer-based systems

Complex systems in which computers play a major role. While complex physical systems and sophisticated software systems can help people to lead healthier and more enjoyable lives, reliance on these systems can also result in loss of money, time, and life when these systems fail. Much of the complexity of these systems is due to integration of information technology into physical and human activities. Such integration dramatically increases the interdependencies among components, people, and processes, and generates complex dynamics not taken into account in systems of previous generations. Engineers with detailed understanding of the application domain and computer electronics, software, human factors, and communication are needed to provide a holistic approach to system development so that disasters do not occur.

Complex computer-based systems can contain distributed processing and databases, internal communication systems, heterogeneous processing components, and sophisticated human-machine interfaces. Processing components by themselves can constitute a system, or they may be embedded in a physical system such as an automobile, aircraft, or medical diagnostic system. Both the encompassing system and the processing system are known as computer-based systems. *See* DISTRIBUTED SYSTEMS (COMPUTERS); EMBEDDED SYSTEMS; HUMAN-COMPUTER INTERACTION.

Such systems are normally built by people from diverse disciplines. Each discipline uses terminology, problem structuring, and techniques specific to its field, which prevents information from being shared, increases integration difficulties, and makes interdisciplinary optimization extremely difficult. The lack of communication and interdisciplinary understanding makes it more difficult to keep the parts and interfaces consistent, and to analyze system-level aspects such as security, performance, and safety. In real-time systems, meeting performance requirements is as important as meeting functional requirements. Delayed timing can cause catastrophic results in control systems such as digital flight systems or nuclear power plants.

Engineering activities. The computer-based systems engineer develops a system within a system; the properties of the former have pervasive effects throughout the larger system. The computer-based system consists of all components necessary to capture, process, transfer, store, display, and manage information. Components include software, processors, networks, buses, firmware, application-specific integrated circuits, storage devices, and humans (who also process information). Embedded computer-based systems interact with the physical environment through sensors and actuators, and also interact with external computer-based systems (see **illus.**). The computer-based systems engineer must have a thorough understanding of the system in which the computer-based system is embedded, for example an automobile, medical diagnostic system, or stock exchange.

Computer-based systems engineering is analogous to traditional systems engineering, but the focus and necessary skills are different. To better understand the differences, it is useful to look at how these disciplines work together to develop a sophisticated vehicle. The systems engineer has primary responsibility for specifying and allocating requirements for the vehicle. The computer-based systems engineer works with the systems engineer on system-level requirements that relate to the distributed processing system. One such requirement is "the vehicle operator shall be able to navigate." The systems engineer and the computer-based systems engineer make decisions concerning the system boundary; for example, they decide which navigation functions the system automates and which navigation functions the vehicle operator performs. They also make design decisions; for example, they might decide that the system will include maps and that a global positioning system will provide location coordinates. To architect the

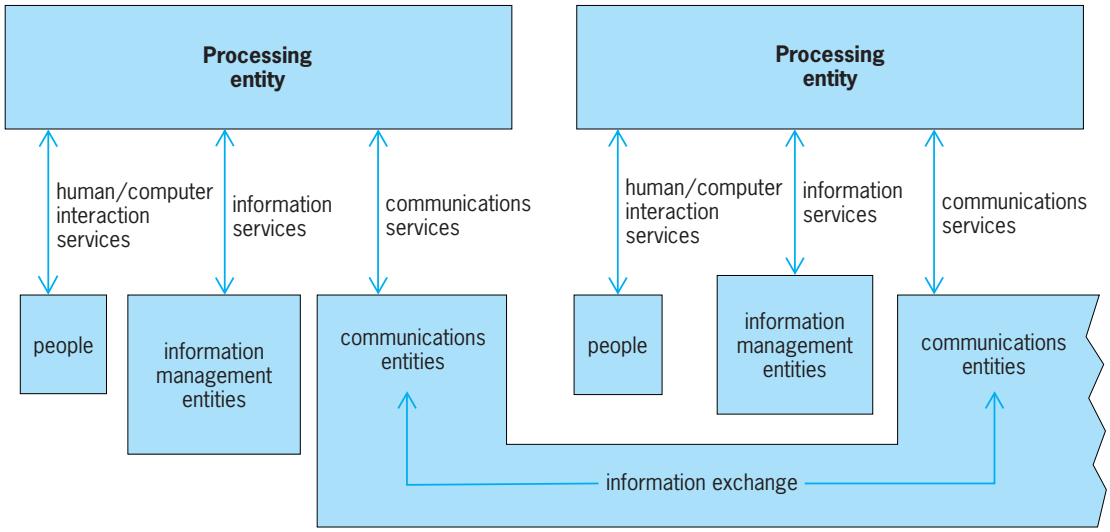

Model of a distributed computer-based system.

computer-based system, the computer-based systems engineer must understand the global positioning system and how the operator will use it, together with the maps, to navigate the vehicle.

The computer-based systems engineer must also determine what is done in software, hardware, and firmware, and make decisions with respect to resources and allocation. These decisions determine the types of processors, communication links, and storage devices that are needed, such as a signal processor for processing data from the proximity sensor and a CD-ROM (compact disk–read-only memory) for storing the map database. Based on dependability needs, the computer-based systems engineer chooses fault tolerance features such as dual buses, standby processors, fault-tolerant operating systems, and software recovery blocks. The computer-based systems engineer also optimizes both the number of standard parts in the computer-based system (commonality) and the system's affordability. Standard software and hardware parts and standard interfaces reduce development and maintenance costs. *See* FAULT-TOLERANT SYSTEMS.

After considering various trade-offs, including those involving processing allocation, the systems engineer and computer-based systems engineer define the subsystems and interfaces. Interfaces should be minimal for ease of change and good maintainability. Software engineers work with computer-based systems engineers to define software requirements and allocate them to processing resources. In the example, some software will be allocated to the signal processor, and other software will be allocated to data processors. Such allocation decisions can be difficult to make and can seriously affect performance. Tools that simulate software functionality running dynamically on different hardware configurations support these trade-offs. *See* SOFTWARE ENGINEERING; SYSTEMS ENGINEERING.

Model-based development. Models are necessary in systems engineering as they support interdisciplinary communication, formalize system definition, improve analysis of trade-offs and decision making, and support optimization and integration. The use of models can reduce the number of errors in the design and thus the system, reduce engineering effort, and preserve knowledge for future efforts. Maintaining models with up-to-date knowledge is a major problem as most systems are not generated from models, although this should be an industry goal. During the later stages of system development and testing, significant schedule pressure makes it difficult to keep the models and manually developed software consistent.

Model types are numerous, and include domain models, process models, data models, system dynamics models, dependability models, and communication models. Integrating diverse models is a major issue in systems engineering science. The Unified Modeling Language (UML) attempts to address this problem. This language is object-oriented and integrates scenario models, state models, and information models. Scenario models describe interactions between the system and its environment, including people who use the system. State models describe system dynamics, changes that occur in the system because of events that occur externally or internally. Information models define the data in the system. *See* OBJECT-ORIENTED PROGRAMMING.

Integrated design environments support tool integration, and are a necessity for integrating the tools that support diverse models. Although integrated design environments are in their infancy, standards such as Computer-Aided Systems and Software Engineering Data Interchange Format (CDIF) are contributing to their development. Stephanie M. White

Bibliography. J. Lavi, *Proceedings of the 1998 IEEE Conference and Workshop on Engineering of Computer-Based Systems*, No. RS00183, IEEE Computer Society Press, Los Alamitos, CA, March 1998; E. Long, A. Misra, and J. Sztipanovits, Increasing productivity at Saturn, *IEEE Computer*, pp. 35–43, August 1998; D. W. Oliver, T. P. Kelliher, and J. G. Keegan, *Engineering Complex Systems with Models and Objects*, McGraw-Hill, New York, 1997; J. Rozenblit, T. Ewing, and S. Schulz, *Proceedings of the 1997 Conference and Workshop on Engineering of Computer-Based Systems*, No. PR07889, IEEE Computer Society Press, Los Alamitos, CA, March 1997; S. Schulz et al., Model-based codesign, *IEEE Computer*, pp. 60–67, August 1998; S. White et al., Systems engineering of computer-based systems, *IEEE Computer*, pp. 54–65, November 1993.

Computer graphics

A branch of computer science that deals with the theory and techniques of computer image synthesis. Computers produce images by analyzing a collection of dots, or pixels (picture elements). Computer graphics is a mature discipline with a growing range of applications. It is used to enhance the transfer and understanding of information in science, engineering, medicine, education, and business by facilitating the generation, production, and display of synthetic images of natural objects with authentic realism almost indistinguishable from photographs. Computer graphics facilitates the production of images that range in complexity from simple line drawings to three-dimensional reconstructions of data obtained from computerized axial tomography (CAT) scans in medical applications. Images of three-dimensional anatomical surfaces can be constructed from the two-dimensional slices generated by computerized tomography or magnetic resonance imaging. These images help physicians to understand the complex anatomy present in the slices, and to evaluate and plan the treatment and even surgery of complex anatomical abnormalities (**Fig. 1**). *See* COMPUTERIZED TOMOGRAPHY; MEDICAL IMAGING.

Scientific data visualization combines computer graphics and data analysis techniques to allow visualization and analysis of large data sets in scientific and engineering computations and thus offer insights that would be difficult or impossible to extract

from numerical results. The creation and interaction that is afforded by computer graphics has become an indispensable tool in automobile, aircraft, and architectural design. New products can be simulated from conception to testing, thereby bypassing several steps and weeks of the traditional design and test cycle. Molecular modeling (**Fig. 2**) and user interaction facilitates and speeds the design of new drugs.

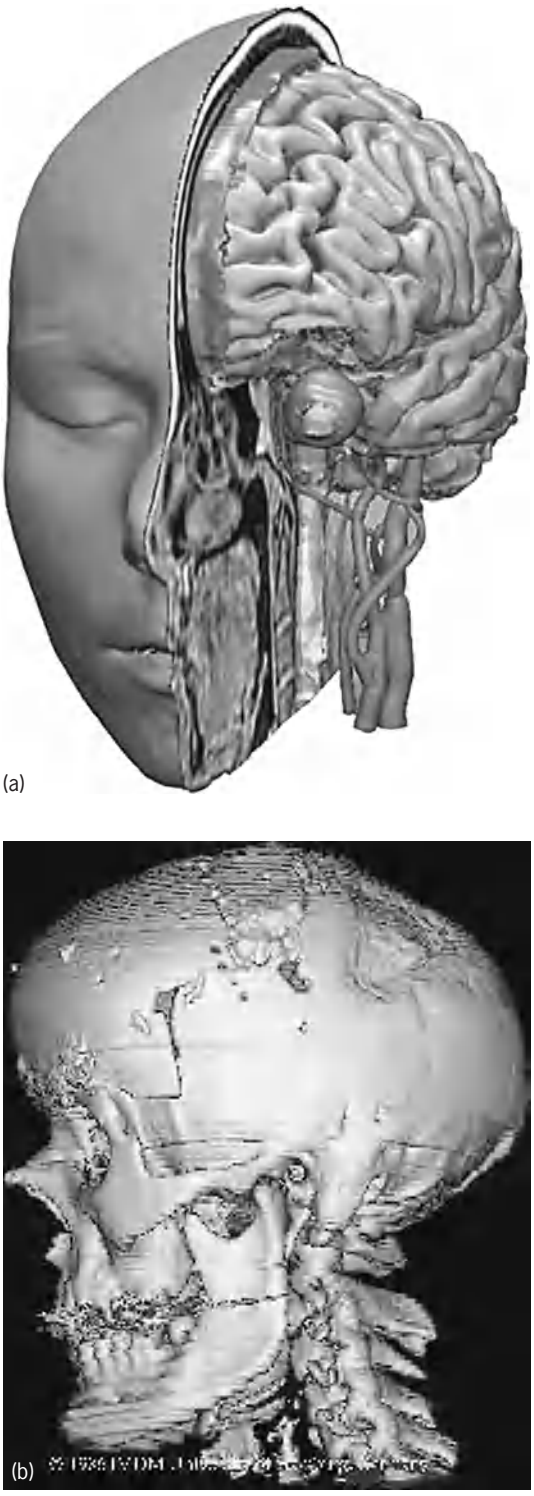

Fig. 1. Volume visualizations of medical image data from computerized tomography and magnetic resonance imaging. Such anatomical images of living subjects are used to create interactive atlases. (*a*) Visualization of brain and skull utilizing the VOXEL-MAN system. (*b*) Post-operative computerized tomography study. (*IMDM, University of Hamburg*)

User interaction can be increased through animation, which conveys large amounts of information by seemingly bringing to life multiple related images. Animation is widely used in entertainment, education, industry, flight simulators, scientific research, and heads-up displays (devices which allow users to interact with a virtual world). Virtual-reality applications permit users to interact with a three-dimensional world, for example, by "grabbing" objects and manipulating objects in the world.

Digital image processing is a companion field to computer graphics, which also deals with images. However, image processing, unlike computer graphics, generally begins with some image in image space, and performs operations on the components (pixels) to produce new images that exhibit some desired feature. *See* IMAGE PROCESSING.

Image presentation and display hardware. Computers are equipped with special hardware to display images. Several types of image presentation or output devices convert digitally represented images into visually perceptible pictures. They include pen-and-ink plotters, dot-matrix plotters, electrostatic or laser-printer plotters, storage tubes, liquid-crystal displays (LCDs), active matrix panels, plasma panels, and cathode-ray-tube (CRT) displays. Images can be displayed by a computer on a cathode-ray tube in two different ways: raster scan and random (vector) scan. *See* CATHODE-RAY TUBE; COMPUTER PERIPHERAL DEVICES.

Liquid-crystal and plasma displays. Liquid-crystal and plasma panel displays are frequently employed in laptop computers. Plasma panels comprise an array of tiny neon bulbs which emit orange light ("in the on" or "intensified" state) under an applied voltage. It remains in this state until it is changed, and thus does not need to be refreshed. Plasma panels typically have about 50–125 cells per inch (20–50 per centimeter) and employ raster technology. The screens are flat, flicker-free, and rugged, making them ideal for military and industrial applications.

A liquid-crystal display is made up of six layers sandwiched together to form a thin panel. The front layer is a vertical polarizer plate. Next is a layer with thin vertical grid wires electrodeposited on the surface adjoining the crystal. This is followed by the thin liquid-crystal layer, a layer of horizontal grid wires, and a horizontal polarizer. The final layer is a reflective layer.

Active matrix panels are liquid-crystal-display panels that have transistors at each grid point. The transistors are used to make the crystals change their state quickly and to control the degree to which the state has been changed. These properties make active-matrix panels attractive display devices for miniature television sets with continuous-tone images. The crystal can also be dyed to provide color. *See* FLAT-PANEL DISPLAY DEVICE; LIQUID CRYSTALS.

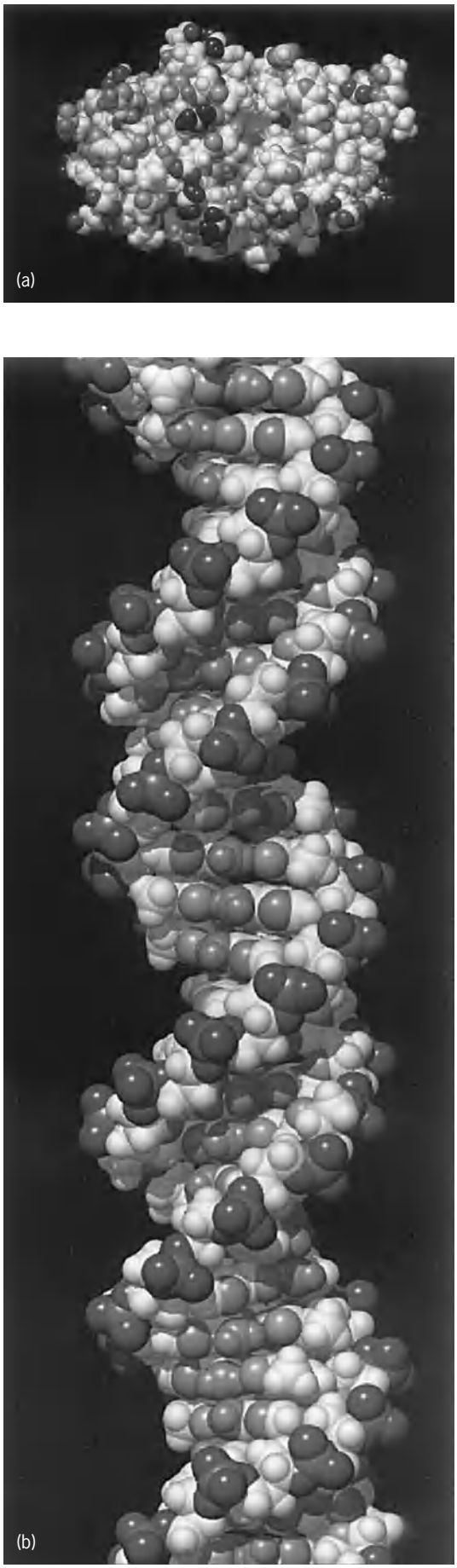

Vector displays. In the 1960s, random scan (vector) displays became a popular technique for generating graphics on a screen without using large amounts of computer memory. Vector images are specified as a series of lines maintained in a list. The lines are specified by the coordinates of their end points. These are stored in a display buffer, a contiguous memory containing all of the information required to draw the picture on a cathode-ray-tube screen. The computer cycles through the list, redrawing each line 30 times a second to maintain a flicker-free picture on the screen. One drawback of most random scan displays is that they do not allow for patterned shading except when it is defined as a series of lines.

Raster-scan displays. The most common display for raster scan images is the cathode-ray tube. The raster scan is very similar to a television set. The beam in this system traces a regular pattern on the screen. It usually starts at the upper left corner of the screen and is directed across the screen in the horizontal direction. After completing one pass across the screen, it moves down and traces the next line. Each line is referred to as a scan line. Images are composed of continuous shaded rows and columns of discrete pixels. Thus horizontal and vertical line segments are represented as a group of adjacent pixels in a row or column. The area of computer graphics that is responsible for converting a continuous figure such as a line segment into its discrete representation is called scan conversion.

Raster-scan systems normally refresh the entire screen with the display pattern 30–60 times a second, recreating a pattern that is stored as a bit map in a special high-speed video memory. Every pixel in the display corresponds to one or more bits of memory. The total number of pixels in an image is a function of the size of the image and the number of pixels per unit length; the latter number is referred to as the resolution of the image. A 2 by 1 in. (5 by 2.5 cm) image at a resolution of 300 pixels per inch (120 pixels per centimeter) would have a total of 180,000 pixels.

The quality of an image is determined by its spatial and contrast resolution. The spatial resolution (sharpness) is determined by the number of rows and columns of pixels. For example, a high-resolution image might consist of about 1000 by 1000 pixels. The contrast resolution is the number of different shades of gray in a monochromatic display or the number of possible colors in a color display. The control over the intensity or color of each pixel is obtained by storing multiple bits for each pixel. For example, three bits will yield eight intensity levels. These bits are then used to control not only

Fig. 2. Arrangements of hydrophobic and hydrophilic atoms which give rise to distinctive properties of living cells. Hydrophobic carbon atoms are white; hydrophilic atoms are various shades of gray. (*a*) Plant enzyme dihydrofolate reductase. (*b*) Double strand of deoxyribonucleic acid (DNA). (*From D. S. Goodsell, Machinery of Life, Springer-Verlag, 1993*)

intensity but color. Higher-resolution images require more data to represent the image and more solid-state memory to store the image. For color images, three times as many bits are needed, one for each of the three additive primary colors from which the image is made; for example, 8 bits for each of the colors red, blue, and green (yield a total of 24 bits in a so-called full-color system. Systems with 24 bits per pixel of color provide for up to 2^{24} or 16,777,216 different colors to produce the smooth shades required of realistic images.

A need for only a small number of colors in a given image is often coupled with a need to change colors from image to image. A look-up table can be employed in such situations. This display device facilitates reductions in storage requirements that never the less satisfy the need to support a sufficient number of simultaneous colors. The look-up table has as many entries as there are pixel values. A pixel's value is not used to control the beam intensity directly but rather as an index into the look-up table. The entry in the table is then used to control the beam's intensity or the color of the screen. For example, a pixel value of, say, 92 is used to access entry number 92 in the table, and the value stored in that entry is then used to control the beam.

Color displays use a shadow-mask cathode-ray tube, which is similar to the television picture tube. Here, the inside of the tube's viewing surface is covered with closely spaced groups of red, blue, and green phosphor dots. The dot groups, arranged in a triad, are so small that light emanating from the individual dots is perceived as a mixture of the three colors. Very high resolution displays require masks to be fine spaced. The triads are placed on about 0.20-mm centers, whereas home television sets are placed on about 0.60-mm centers. This distance is called the pitch of the tube. *See* ELECTRONIC DISPLAY; PICTURE TUBE.

Interactive techniques. Interaction with the object takes place via devices attached to the computer, starting with the keyboard and the mouse. Each type of device can be programmed to deliver various types of functionality. The area of computer graphics that deals with effective use of interactive devices and the protocols that facilitate computer-human interaction is called interactive techniques.

The quality and ease of use of the user interface often determines whether users enjoy a system and whether the system is successful. In critical applications such as air-traffic control and command/control or in "situation display" problems, a poorly designed user interface can contribute to accidents.

Interactive graphics aids the user in the creation and modification of graphical objects and the response to these objects in real-time. The most commonly used input device is the mouse. It is a small hand-held device whose relative motion across a surface can be measured. The roller ball on the underside of the mechanical mouse is converted into digital values that determines its movement in two orthogonal (x, y) directions. The mouse typically specifies locations in image space. Here, the cursor is often displayed as visual feedback to allow the user to see the locations being specified. It could also be used to select an item in a pull-down or pop-up menu. Often in this case the item selected is highlighted or framed in a filled rectangle, since the location of the cursor becomes irrelevant.

Other kinds of interaction devices include the joystick, trackball, light pen, and data tablet. The trackball is often described as an upside-down mouse. The user operates the trackball by dragging the palm across it. The motion of the ball rotating freely within a housing is measured by potentiometers or shaft encoders. Mice are "relative" devices that have no absolute origin—only changes from previous positions are reported—and thus they can be moved, lifted, and moved again. However, the joystick is homed in a center position from which it can be moved left, right, backward, or forward. Here again potentiometers sense the movements. *See* POTENTIOMETER.

A data tablet is a flat surface, from 6 by 6 in. (15 by 15 cm) to 4 by 6 ft. (1.2 by 1.8 m) or larger in size, which is capable of detecting the position of a movable stylus held by the user. Most tablets use magnetic signals generated by electrical pulses applied to a patterned grid which is embedded in the tablet to sense the position of the stylus. The tablet is an absolute device as there is a frame of reference or origin to which the position of the stylus is referred. Through this device, the physical movement of a user's hand can be converted into the movement of a cursor of an object on the display screen.

Voice recognition systems, which free the user's hands for other activities, provide other means of interaction. They analyze the waveforms created when users speak. These systems must, however, be trained to recognize the waveform of a speaker.

Some of these two-dimensional (2D) devices can be modified to extend to three dimensions (3D). The space ball (**Fig. 3**), an example of a three-dimensional device, is a rigid sphere containing strain gages. The user can pull or push the space ball in any direction

Fig. 3. Space ball motion-control input device, which resolves fingertip pressure into translation in three dimensions and rotation, providing six-degrees-of-freedom control. (*Labtec, Inc.*)

to provide three-dimensional translation and orientation. The data glove is a device capable of recording hand movements. It records both the position of the hand and its orientation as well as finger movements. The data glove is capable of a simple gesture recognition and general tracking of hand orientation.

Computer graphics and image processing complement each other. Remarkable visual effects can be achieved by using a combination of these technologies.

Image representation. The representation and description of objects requires the creation of geometric models. Geometric objects have important characteristics such as size, shape, and orientation and certain spatial relationships to each other.

The surfaces of objects in an image can be described as a collection of polygons or patches that are stitched together to make up the scene. This method is called geometric or shape modeling. Polygons work reasonably well for objects that have flat surfaces, like buildings. When surfaces are highly curved, it is generally necessary to cover the surface with many small polygons, which require increases in storage and in execution time. Patches are very efficient at representing surfaces of objects to a given accuracy. Often fewer patches are required, but the equations require more computation than those used for polygons.

The production of a computer-generated image requires additional steps. The designer has to specify the objects in the image and their shapes, positions, orientations, and surface colors or textures. Further, the viewer's position and direction of view (camera orientation) must be specified. The software should calculate the parts of all objects that can be seen by the viewer (camera). Only the visible portions of the objects should be displayed (captured on the film). (This requirement is referred to as the hidden-surface problem.) The rendering software is then applied to compute the amount and color of light reaching the viewer eye (film) at any point in the image, and then to display that point. Some modern graphics work stations have special hardware to implement projections, hidden-surface elimination, and direct illumination. Everything else in image generation is done in software.

Solid modeling is a technique used to represent three-dimensional shapes in a computer. The importance of solid modeling in computer-aided design and manufacturing (CAD/CAM) systems has been increasing. Engineering applications ranging from drafting to the numerical control of machine tools increasingly rely on solid modeling techniques. Solid modeling uses three-dimensional solid primitives (the cube, sphere, cone, cylinder, and ellipsoid) to represent three-dimensional objects. Complex objects can be constructed by combining the primitives. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING; COMPUTER-AIDED ENGINEERING.

Color models. Color specification using the RGB (red, green, and blue) color model is an additive process. The appropriate primary color component is added to black to yield a desired color. This method matches closely the working principles of the display monitor. The complementary color model, called the CMY (cyan, magenta, and yellow) model, defines colors using a subtractive process. This model matches the working principles of the printer. It begins with white, and the appropriate color is subtracted to yield the desired color. *See* COLOR.

Image synthesis. The creation of images by simulating a model of light propagation is often called image synthesis. The goal of image synthesis is often stated as photorealism, that is, the criterion that the image look as good as a photograph. Rendering is a term used for methods or techniques that are used to display realistic-looking three-dimensional images on a two-dimensional medium such as the cathode-ray-tube screen. A three-dimensional image has to be projected before it is displayed on a two-dimensional display medium. The display of a wire-frame image (**Fig. 4***a*) is one way of rendering the object. The most common method of rendering is shading. Generally, rendering includes addition of texture, shadows, and the color of light that reaches the observer's eye from any point in the image. The specification of this color (as determined by each of the red, green, and blue color components) is computed by the rendering software that is used to generate the image.

The concept is somewhat related to operation of a camera that is used to photograph an object that is illuminated by a light source, albeit simplified in the

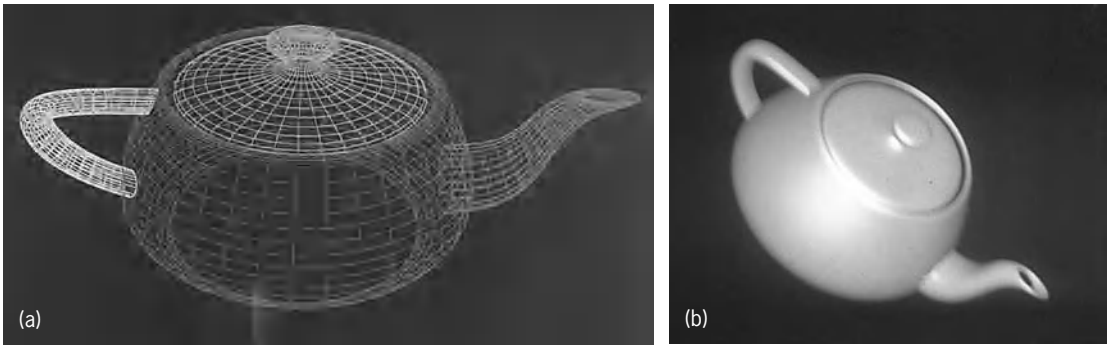

Fig. 4. Renderings of a teapot. (*a*) Wire-frame model with 512 polygons. (*b*) Smooth shading (non-shiny). (*A. Tokuta, Technical Report, Department of Computer Science and Engineering, University of South Florida*)

Fig. 5. Animation frame showing both reflections from and refraction through Geri's eyeglasses in Pixar's Geri's Game. (*Pixar Animation Studios*)

model. Light energy emanating from the source interacts with the physical environment, is reflected off the surface of the object through the camera lens, and is registered on the film, forming an image of the object. The parts of the object closer to the light source should appear brighter, while the parts farther away or directed away from the light source should appear darker. The mathematical model that is used for these types of optical phenomena is called a local illumination model. It tracks the light energy coming directly from the light source to a particular object surface but does not fully account for the energy arriving at the surface. Light energy is also reflected from other objects in the environment, and energy can pass through transparent and translucent objects in the environment (**Fig. 5**). Models of computation that strive to account for the accurate transportation of light are referred to as global illumination models.

Producing realistic images involves both physics and psychology. The psychology of visual perception is extensively documented. The amount of light propagation can be computed by the rendering software, based on simple models. Shading techniques (constant, Gourad, or Phong) can be applied to evaluate the illumination model at various locations. In constant shading, a single color is computed for an entire polygon, based on the position of the light source. This method of shading tends to produce poor results because the human visual system exaggerates changes in intensities that may exist between adjacent polygons that may have different shading.

Consequently, some form of interpolation is often used. For the Gourad shading, the evaluation of the model is done once for each vertex. The intensities are then interpolated across the polygon. The direction of the normal to the surface must be taken into account in the calculation of the intensities; here, the vertex normal that is used is the computed average of the adjacent surface normals. In the case of Phong shading (Fig. 4*b*), the evaluation of intensity is done once for each pixel. The normal vector that is used in the evaluation is computed, however, as the interpolated values of normals at the vertices. This has the added advantage of preserving specular highlights.

The character of the light that is reflected from a surface of an object depends on the composition, direction, and geometry of the source, the orientation of the surface, and the surface properties of the object. It is also characterized as either diffusely or specularly reflected. An ideal mirror reflects light that strikes it, and each ray that hits a point on the surface is reflected in one direction. An observer who happens to be located in that direction, and is looking at the point, should see a reflection of the light source in its original color and not the surface itself. This ideal reflection is called specular. Specular reflection is directional; that is, it depends on the angle of incident light. An ideal dull surface reflects light in all directions because the microfacets constituting the surface are considered to be pointing in different directions. Thus an observer perceives the same light, intensity to be reflected from a given point regardless of the observer's position. This type of reflection is called diffuse. Each point on a surface is thus considered to emit three types of light: diffuse light, specularly reflected light, and transmitted or refracted light. Each ray leaving a surface comprises the sum of these three contributions. Modeling of shadows, requires additional computation. *See* ILLUMINATION; REFLECTION OF ELECTROMAGNETIC RADIATION.

Ray tracing is a technique that is used to produce realistic images but is view-dependent. This means that the entire computation must be repeated when the observer's position is changed. This method, however, interleaves the determination of visibility, illumination, and shading at each pixel. The basic idea is to trace the paths of rays, from the viewpoint of an observer looking through each pixel in the screen into the image. If the ray hits a surface, reflected and possibly refracted rays are spawned, which in turn are traced to see if they intersect any other surfaces. The final color and intensity of each pixel are determined as the summation of the contributions from each spawned ray.

The radiosity method is view-independent. The radiosity of a surface is the rate at which energy leaves the surface, and is the sum of the rates at which the surface emits energy and reflects or transmits it from that surface. Given an image, the calculations are done only once, even before the visible-surface determination is performed from a desired viewpoint of the observer. Once the global illumination has been determined, it is easy to move the observer to different viewpoints and thus create a series of images. This method uses the principle of conservation of energy (by modeling energy equilibrium in a system of surfaces) to compute the light intensity for each surface in the image, and assumes ideal diffuse surfaces composed as either light sources or reflecting surfaces. *See* CONSERVATION OF ENERGY.

Natural and synthetic objects. Computer-generated images are used extensively in the entertainment world and other areas. Realistic images have become essential tools in research and education. Conveying realism in these images may depend on the convincing generation of natural phenomena. A fundamental difficulty is the complexity of the real world. Existing models are based on physical or biological concepts.

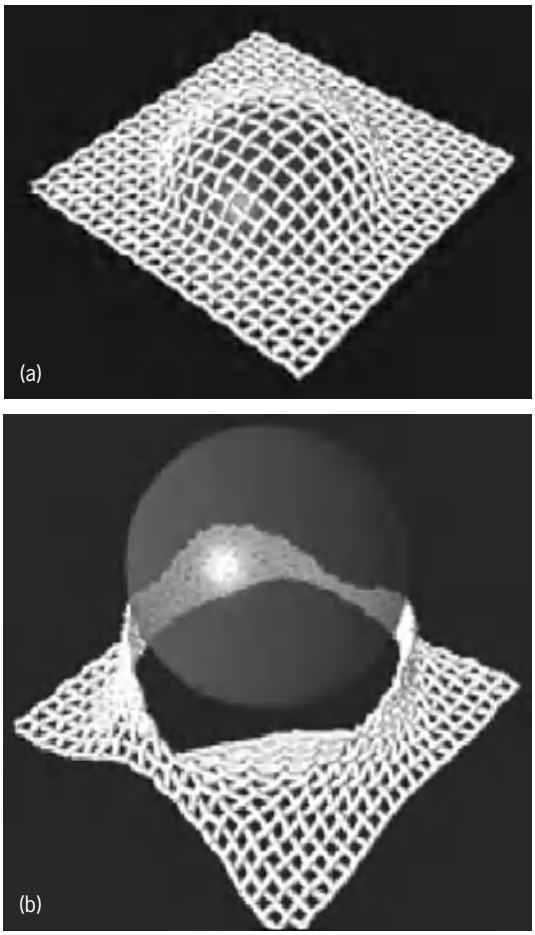

Fig. 6. Example of realistic images based on physical properties. (*a*) Net falling over a physical obstacle. (*b*) Fracture developing in the net. (*D. Terzopoulos and K. Fleischer, Schlumberger Palo Alto Research*)

The behavior of objects can be determined by physical properties or chemical and microphysical properties. (**Fig. 6**).

Many natural phenomena are not efficiently represented by geometric or solid models. Procedural models, fractals, grammar-based models, and particle systems have been used to model a variety of things. Procedural models describe objects that can be modeled from a procedure and a parameter set. The appearances of textured surfaces from wood to marble, and woven material to fur, have been simulated by use of procedural texturing techniques.

Clouds and mountain terrain are examples of objects that have been modeled using fractals. The term "fractal" was introduced by Benoit Mandelbrot to describe a particular attribute of certain natural phenomena. For example, an aerial view of a coastline tends to exhibit a so-called recurrent pattern that is not dependent on the scale of the image. *See* FRACTALS.

Grammar-based models can describe the structure of certain plants and trees. Work in this area is based on a parallel graph grammar called L-system. This system specifies the construction of a plant or tree as a sentence in the language that the grammar generates.

Particle systems have been used to model fuzzy objects such as fire and clouds. An object is represented by a set of particles that evolve over time—each is born, evolves in space, and dies or is extinguished, all at different times.

Volume rendering. Volume rendering combined with animation is a major tool used in visualization. Visualization offers a method to see the unseen. It enables researchers to observe their simulations and computations, thus enriching the process of scientific discovery. A major application of this area is medical imaging (Fig. 1), where volume data are available from positron emission tomography (PET) scanners, for example. *See* SIMULATION. Alade O. Tokuta

Bibliography. M. F. Cohen and J. R. Wallace, *Radiosity and Realistic Image Synthesis*, Academic Press, 1993; J. D. Foley et al., *Computer Graphics: Principles and Practice*, Addison-Wesley, 1996; A. S. Glassner, *Principles of Digital Image Synthesis*, Morgan Kaufmann, 1995; F. X. Sillion and C. Puech, *Radiosity and Global Illumination*, Morgan Kaufmann, 1994; A. Watt, *3D Computer Graphics*, 3d ed., Addison-Wesley, 2000.

Computer-integrated manufacturing

A system in which individual engineering, production, and marketing and support functions of a manufacturing enterprise are organized into a computer-integrated system. Functional areas such as design, analysis, planning, purchasing, cost accounting, inventory control, and distribution are linked through the computer with factory floor functions such as materials handling and management, providing direct control and monitoring of all process operations.

Computer-integrated manufacturing (CIM) may be viewed as the successor technology which links computer-aided design (CAD), computer-aided manufacturing (CAM), robotics, numerically controlled machine tools (NCMT), automatic storage and retrieval systems (AS/RS), flexible manufacturing systems (FMS), and other computer-based manufacturing technology. Computer-integrated manufacturing is also known as integrated computer-aided manufacturing (ICAM). Autofacturing includes computer-integrated manufacturing, but also includes conventional machinery, human operators, and their relationships within a total system. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING; FLEXIBLE MANUFACTURING SYSTEM; ROBOTICS.

Agile manufacturing and lean manufacturing. Flexible manufacturing has been subsumed by the concept of agile manufacturing, which presumes that all operations can be virtually instantly reconfigured to satisfy a changing market demand. For example, only cars with certain options are selling, but the marketing department has directed the manufacturing department to gear up to produce models with other options. Agile manufacturing could efficiently reconfigure operations to satisfy the new requirement, or it could create a virtual manufacturing

facility by electronically combining the operations of many different locations into the optimum configuration required to meet new market demands.

Lean manufacturing reorganizes and reduces operations to the minimum required to produce certain parts or products at the absolutely lowest cost. Lean manufacturing typically focuses on either market-niche-type products or seeks a virtual monopoly on large-volume low-variability products.

Modern manufacturing enterprise. The CIM factory integrates the traditional hard or process-based technology factory (departments) with software of systems-based technology (nondepartments).

Effectiveness, efficiency, and economy. Manufacturing companies seek the optimal combination of effectiveness, efficiency, and economy. Efficiency is most often associated with transfer-line or assembly-line technology. The best-known developer of this technology was Henry Ford. His Model T factory could produce cars at an unprecedented rate for that time, but the consumer had no choice at all about the specifications of the product. Flexibility is most often associated with robotics and other reprogrammable automation. Effectiveness is usually associated with the soft technology and organizational support structure of the enterprise.

Optimization of the CIM factory is based on specific criteria. By considering efficiency, flexibility, and effectiveness, the CIM factory enables economy of scope. In economy of scope, flexibility is at least as important as efficiency, because the factory is seen to be an ever-changing dynamic environment which must always respond quickly to the needs of the marketplace, if not actually lead the need. The traditional factory with both rigid transfer-line technology and a rigid organizational structure cannot satisfy this objective; it strives for maximum volume of each product. In the ideal case, the CIM factory provides for profit with an order quantity of one, that is, when no two products are identical. Within the manufacturing enterprise, the CIM factory combines agility during manufacturing planning and leanness during production.

Organization. The CIM factory concept includes both soft and hard technology. Soft technology can be thought of as the intellect or brains of the factory, and hard technology as the muscles of the factory. The type of hard technology employed depends upon the products or family of products made by the factory. For metalworking, typical processes would include milling, turning, forming, casting, grinding, forging, drilling, routing, inspecting, coating, moving, positioning, assembling, and packaging. For semiconductor device fabrication, typical processes would include layout, etching, lithography, striping, lapping, polishing, and cleaning, as well as moving, positioning, assembling, and packaging. More important than the list of processes is their organization.

Whatever the products, the CIM factory is made up of a part fabrication center, a component assembly center, and a product assembly center. Centers are subdivided into work cells, cells into stations, and stations into processes. Processes comprise the basic transformations of raw materials into parts which will be assembled into products. In order for the factory to achieve maximum efficiency, raw material must come into the factory at the left end and move smoothly and continuously through the factory to emerge as a product at the right end. No part must ever be standing; each part is either being worked on or is on its way to the next workstation.

In the part fabrication center, raw material is transformed into piece parts. Some piece parts move by robot carrier or automatic guided vehicle (AGV) to the component fabrication center. Other piece parts (excess capacity) move out of the factory to sister factories for assembly. There is no storage of work in process and no warehousing in the CIM factory. To accomplish this objective, part movement is handled by robots or conveyors of various types. These materials handlers serve as the focus or controlling element of work cells and workstations (**Fig. 1**). Each work cell contains a number of workstations. The station is where the piece part transformation occurs from a raw material to a part, after being worked on by a particular process.

Components, also known as subassemblies, are created in the component assembly center. Here materials handlers of various types, and other reprogrammable automation, put piece parts together. Components may then be transferred to the product assembly center, or out of the factory (excess capacity) to sister factories for final assembly operations there. Parts from other factories may come into the component assembly center of this factory, and components from other factories may come into the product assembly center of this factory. The final product moves out of the product assembly center to the product distribution center or in some cases directly to the end user. *See* AUTOMATION.

Fig. 1. Typical CIM workstation.

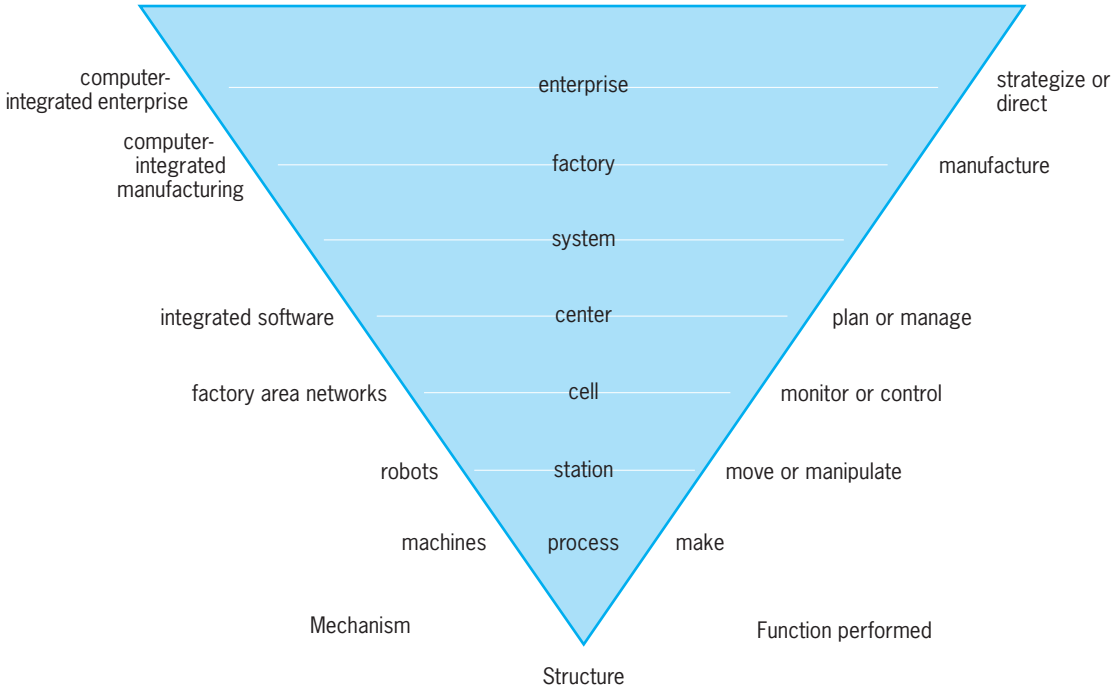

Fig. 2. Diagram of the architecture of manufacturing.

Integration. The premise of CIM is that a network is created in which every part of the enterprise works for the maximum benefit of the whole enterprise. Independent of the degree of automation employed, for example, whether it is robotic or not, the optimal organization of computer hardware and software is essential. The particular processes employed by the factory are specific to the product being made, but the functions performed can be virtually unchanged in the CIM factory no matter what the product. These typical functions include forecasting, designing, predicting, controlling, inventorying, grouping, monitoring, releasing, planning, scheduling, ordering, changing, communicating, and analyzing.

It must be recognized that independent, optimum performance of the individual functions is not as important as their integration with one another, or their integration with the factory floor itself. This integration is brought about in two ways. The first is through an architecture of manufacturing which specifies precisely how and when each function is integrated with the other (**Fig. 2**). The second is through a controller and networks.

For example, a cell controller ensures integration of data between each machine, robot, automatic guided vehicle, and so forth, of the cell. The cell network (**Fig. 3**) actually performs this communication within the cell, and connects cells together into centers. In the cell network, communication is possible between each machine of the network and between different networks. The cell controller itself serves to perform such tasks as downloading part programs or recipes from the CAD system or planning system to each machine in the cell, monitoring actual performance of each machine and comparing this performance to the plan, selecting alternate routing for a part if a machine is not operable, notifying operators of pending out-of-tolerance conditions, archiving historical performance of the cell, and transmitting to the center level on an exception basis the cell performance compared to plan. Typically, the human interface with factory floor operations and intermachine communications is accomplished through networked programmable controllers, personal computers running soft logic, machine and robot controllers, display devices, and programmable logic controllers (PLCs) communicating through Wireless Application Protocol (WAP) devices. These interactions are managed through personal computer–based technology integrated in real time through various communications protocol standards. *See* PROGRAMMABLE CONTROLLERS.

The center level of the CIM factory is all of the policies and procedures that run the factory. These are embodied in computer software, which is in turn based upon an overall integration plan or CIM architecture.

The purpose for the architecture is to provide a blueprint for employing CIM. While there are as many variations of the details of the architecture of the CIM factory as there are factories, the structure of the architecture is generic. The focus is the agile manufacturing system (AMS), which is made up of cells, stations, and processes. The information to be fed back to functions where decisions are made is as important as the product itself. In this example, the functions are defined very broadly and are used only to illustrate the absolute necessity for interaction between functions if the CIM factory is to be made to work efficiently, flexibly, and effectively.

In CIM, as decisions are made, they are tested on a computerized model of the CIM architecture to

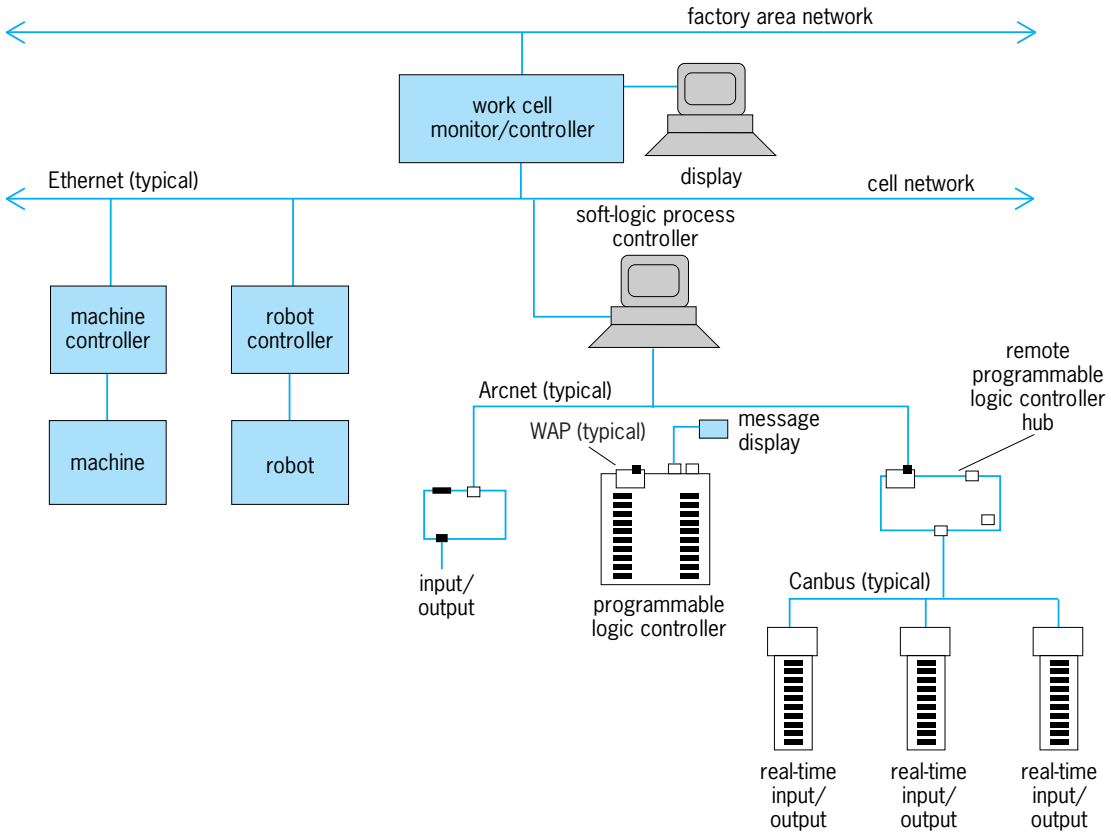

Fig. 3. Organization of a production cell network for computer-integrated manufacturing. Ethernet, Arcnet, Wireless Application Protocol (WAP), and Canbus are international standards for communication.

determine feasibility. Thereby, mistakes are discovered in the computer, not in production in the factory. Simulation and emulation software is used for this purpose. Computer aided software engineering (CASE) software is used to convert the CIM models for decision making into software for operating the CIM factory. Interchangeable and reusable software programs, known as objects, are manipulated by CIM engineers to accomplish this objective. *See* SOFTWARE ENGINEERING. Dennis E. Wisnosky

Bibliography. R. N. Bateson, *Introduction to Control System Technology*, 6th ed., Prentice Hall, 1998; J. L. Fuller, *Robotics: Introduction, Programming and Projects*, 2d ed., Prentice Hall, 1999; R. Hannam, *Computer Integrated Manufacturing: From Concepts to Realisation*, Addison Wesley Longman, 1998; S. K. Vajpayee, *Principles of Computer Integrated Manufacturing*, Prentice Hall, 1998; D. E. Wisnosky and R. C. Feeney, *BPR Wizdom: A Practical Guide to BPR Project Management*, Wizdom Press, 1999.

Computer numerical control

The method of controlling machines by the application of digital electronic computers and circuitry. Machine movements that are controlled by cams, gears, levers, or screws in conventional machines are directed by computers and digital circuitry in computer numerical control (CNC) machines.

Computer numerical control provides very flexible and versatile control over machine tools. Most machining operations require that a cutting tool be fed at some speed against a workpiece. In a conventional machine such as a turret lathe, the turning tool is mounted on a slide with hand-operated infeed and crossfeed slides. The operator manually turns a crank that feeds the cutting tool into the workpiece (infeed) to the desired diameter. Another crank then moves the turning tool along the longitudinal axis of the machine and produces a cylindrical cut along the workpiece. The feed rate of the turning tool is sometimes controlled by selecting feed gears. These gears move the axis slide at the desired feed. A CNC machine replaces the hand cranks and feed gears with servomotor systems. *See* SERVOMECHANISM.

Computer numerical controls allow the desired cut depths and feed rates to be "dialed in" rather than controlled by cranks, cams, and gears. This provides precise, repeatable machine movements that can be programmed for optimal speeds, feeds, and machine cycles. All cutting-tool applications, whether on a lathe, drill press, or machining center, have optimum speeds and feeds, which are determined by carefully weighing the economics of tool life, required production rates, and operator attentiveness. With computer numerical control these parameters are set once, and then they are repeated precisely for each subsequent machine cycle.

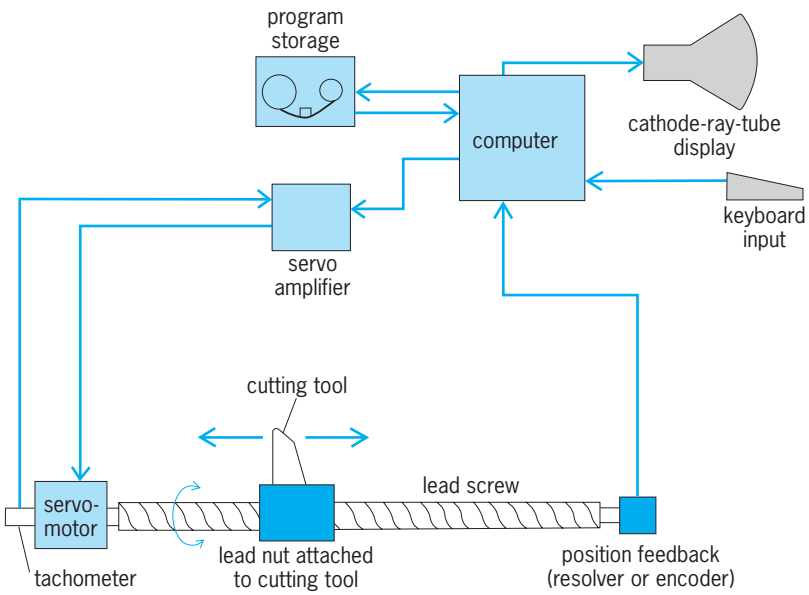

Fig. 1. Schematic diagram of a servo system.

The decision to purchase a CNC machine instead of a conventional machine is a complex one. The CNC machines are usually more complex than conventional machines, which adds to their cost and puts greater burdens on maintenance and support personnel. This added cost is often worthwhile, because of the flexibility and sophistication of the machine. The total cost of purchasing and operating a CNC machine must be weighed against the total cost of purchasing and operating a conventional machine. Complexity of the parts, programming and maintenance costs, and the availability of skilled labor are some of the factors to be considered.

History. Machines with numerical control (NC) became possible in the 1960s, when the development of the transistor made electronic servosystems practical. Earlier NC cutting tools were programmed with punched tape. This technique, while an impressive and productive technology, lacked the sophistication and flexibility that have become available with the development of more advanced systems. Numerical control programs had to be written by hand. Punched tape tended to rip and tear in the hostile environment of the shop floor. Once punched, the programs could not be modified without punching an entirely new tape. Punched tape was also bulky, and if many programs were required, a great deal of space was needed to store the tapes. *See* COMPUTER; TRANSISTOR.

Computer numerical control advanced further with the advent of inexpensive computers in the early 1980s. Instead of being programmed to move to discrete, precalculated locations, the machines can be programmed to perform complex mathematical calculations that determine the locations of the machine's axes. With computer numerical control, milling cutters can be programmed to move in a sinusoidal pattern, for example, by programming a simple sine function into a loop. Previously, each point would have to be explicitly programmed. Programs for computer numerical control can be written parametrically; that is, the same basic motions to be programmed for many different sizes of parts can be controlled with the same program. Programming has become simpler, more powerful, and more user-friendly. It has even become possible to program machine tools graphically at the machine while machining is taking place. *See* PROGRAMMABLE CONTROLLERS.

System components. Machine systems designed for computer numerical control use servosystems to control the position of cutting tools. A typical servosystem (**Fig. 1**) is composed of six parts: a lead screw (ball screw) and lead nut, an electric servomotor, a servoamplifier, a position-feedback device, a velocity-feedback device, and a computer.

Lead screws are precision-ground screws. To prevent any backlash in the screw system, a special lead nut utilizing hundreds of ball bearings tracks the screw. The many ball bearings help equalize and distribute the load forces on the screw. This improves the accuracy, while minimizing friction between the nut and the screw.

Servomotors are specially constructed electric motors. They are designed to operate at very low angular velocities. Although they sometimes drive the lead screw through a gear reduction system, they are usually directly coupled to the lead screw. This direct coupling minimizes any backlash or positioning errors introduced by gears. The motor is required to provide full torque at low, or zero, rpm. Traditionally, servomotors have been direct-current (dc) types. Such motors provide excellent torque at low speeds and have relatively simple electrical requirements, but they have brushes, which are a source of wear and are subject to contamination. The electrical requirements for alternating-current motors are more complex than for dc motors, but the motors themselves are more reliable, especially in hostile environments. *See* ALTERNATING-CURRENT MOTOR; DIRECT-CURRENT MOTOR; GEAR.

The servoamplifier powers the servomotor. Servomotors require precise control of the electrical voltage and current supplied to them. Voltage and current are varied by the servoamplifier. High power is required to accelerate the machine's axes. Moderate power is required to keep them moving. High power (of opposite polarity) is required to decelerate them at the end of a programmed move. Additionally, if the servosystem encounters a physical resistance (such as the load for a heavy cut on a lathe), the servoamplifier must compensate with additional power to keep the lead screw turning at the desired rpm.

The servoamplifier receives a signal voltage from the computer that indicates the desired machine axis direction and speed. A tachometer is mounted on the servomotor and measures the motor's rpm. The output of the tachometer is fed into the servoamplifier, which examines the rpm of the servomotor and varies the voltage and current supplied to the motor to keep it turning at the desired rate. *See* TACHOMETER.

Position feedbacks are provided by resolvers or encoders. These allow the computer to know where the machine's axes are at any time, and to calculate how far the axes must move to reach their programmed positions. Resolvers, and to a lesser extent encoders, are precision devices that must be protected against shock and impact. If not subject to abuse, they generally are very reliable. They are usually directly coupled to the lead screw so that they measure every minute movement of the screw, without the slight backlash that gearing could introduce.

The computer provides overall control of the entire system. It examines the part program, determines the desired positions of the machine axes, then provides special signals to the servoamplifiers to move the servomotors and thereby move the machine's axes. The computer monitors the position feedback provided by the resolvers or encoders. In addition, the computer handles various so-called housekeeping tasks such as displaying the axes' locations on the cathode-ray-tube (CRT). The computer also monitors the program to ensure that dangerous functions, such as opening the chuck in the middle of a cut, have not been programmed.

The speed and computing power of the computer directly affects the quality of the finished workpiece. The computer must examine the axes' locations and velocities, compare their current locations to their programmed locations, and update command signals to the servoamplifiers hundreds or thousands of times each second. Generally, the more often the computer can update in this manner, the smoother and more precise the machine's motions will be. Advances in computing power have led to the construction of CNC machines that can smoothly move six or seven axes simultaneously. This can be accomplished while the computer is still updating the CRT display and checking for errors.

Machine axis layout. Each servo-controlled axis can move either forward or backward in a straight line. The machine tool builder assigns numerical reference (home) locations to each axis. The location of each axis is then determined as a linear measurement from the home position. The home location is usually considered to be 0.0000. A point 2 in. from home would therefore have the coordinate of 2.0000. Programming the machine axis consists of instructing the axis to move to some numerical coordinate.

Most machines require cutting tools to move in at least two directions. On a lathe, this corresponds to the longitudinal and infeed axes (usually labeled Z and X, respectively). Since each axis can move in only one direction, the axes are compounded by being built one on top of the other. In this way, the X axis is mounted on top of the Z axis. A cutting tool mounted to the X axis is therefore also controlled by the Z axis, and has two directions of movement (**Fig. 2**).

Programming. Programming for computer numerical control, at its most basic level, consists of indicating a series of numerical coordinate locations to which the axes will move under program control. Various parameters, such as velocity, are programmed along with the locations of the axes. Ancillary operations, such as turning on both the spindles and the coolant, and specifying the spindle rpm, are also programmed.

Most part programs are broken into lines known as blocks, each of which is a specific program step, executed at one time by the computer. A block may consist of one or more axis locations and a feed rate, a miscellaneous command, or some combination of these. Generally, the execution of each block is completed before the next one is begun.

There are several major classes of part program codes. Although each machine or computer manufacturer is free to define the meaning of each code word, there are several conventions that have been nearly universally adopted.

The block number is coded with an N followed by a number having one to four digits. On older CNC controls, block numbers were necessary, but more modern controls may not require them. Most programmers still use them to assist in reading previously written programs. Block numbers are very useful when part programs must be edited, since they allow specific blocks to be referenced.

The G codes specify how the machine's movements will be made. One of the most common G codes is G01, which activates the linear interpolation mode. In this mode, all the axes move from their current locations to their programmed locations at controlled feed rates, so that they all arrive at the same time. A cutting tool, mounted to both axes, moves in a straight line at the programmed feed rate. If each axis has different distances to traverse in order to reach its programmed location, the one with the longer distance moves faster, so that both axes arrive at the same time. Most part programming consists of moving the axes in coordinated straight lines, and so this function is very common.

A G90 code activates the so-called absolute programming mode. For example, the block N45 G90 X − 2.500 Z − 1.000 will move the X axis to the absolute axis location of −2.500 in., and the Z axis to the absolute axis location of −1.000 in. These

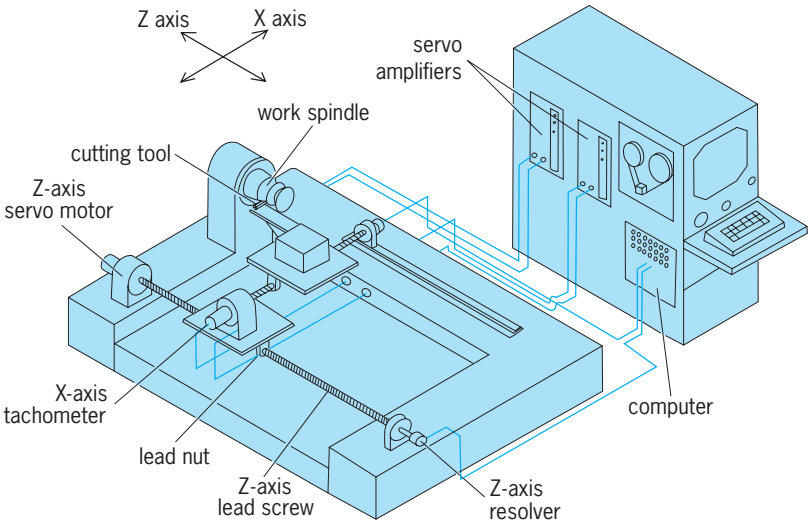

Fig. 2. Diagram of a two-axis computer numerical control lathe.

positions will be reached, regardless of where each axis was before the command was issued.

A G91 code activates the incremental programming mode. In this mode the programmed move is coded as a distance to move from the current location. For example the block N45 G91 X − 2.500 Z − 1.000 will move the X axis 2.500 in. in the negative direction from its current location. The Z axis will move 1.000 in. in the negative direction from its current location. The final locations of the axes will be determined by both their initial locations and the incremental movements programmed.

Just as it is possible to specify the linear interpolation mode, the circular interpolation mode can be specified. This mode causes the axes to move in a controlled circular motion. The programmer merely specifies the radius of the cut and the final location. The computer will automatically calculate the minuscule discrete movements required to move two axes in a coordinated circular fashion.

Parametric programming is the method of programming movements of the axes in terms of computer variables known as parameters. Instead of programming explicit X and Z axes locations in the part program, the programmer indicates that the axes are to move to locations according to specific parameters. Then, by loading the parameter variables with specific values, the computer will move the axes to those locations. This has the advantage of allowing programmers to write programs known as family of parts; these programs have the same basic movements but have differing locations of the axes. The family-of-parts program is logically identical from part to part, but the locations of specific axes may be different.

An example of a parametric program is a program to turn three different truck axle shafts that are identical except for their diameters. A single parametric part program could be written for these three shafts. Instead of specifying the infeed axis locations that correspond to the diameters explicitly, the programmer could have the infeed axis move to the location specified in a parameter variable. By changing the value of the parameter, a shaft with a different diameter can be machined on the lathe with the same program.

In some cases, it is desirable to have all the programs stored in a central computer. The process of direct numerical control can then be used to download the part programs from the central computer to the CNC machine by way of a communication wire. This process eliminates the paper-tape or floppy-disk program storage at the machine. Sophisticated multimachine production lines frequently use this technique to help integrate and automate several machining steps. *See* COMPUTER STORAGE TECHNOLOGY.

In computer-aided manufacturing (CAM), computers are used to assist in programming CNC machines. In sophisticated CNC manufacturing operations, machined parts are first designed on computer-aided-design (CAD) equipment. The same electronic drawing is then used to create the CNC part program automatically. A less advanced version of CAM is the use of high-level part programming languages to write part programs. Special computer programming languages, such as APT and COMPACT II, allow programmers to write complex part programs easily and quickly. For example, indexing a CNC lathe turret may take five or six blocks of a part program. This action can typically be accomplished in just one line of APT or COMPACT II program code. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Typical applications. Computer numerical control machines are used mainly when flexibility is required or variable and complex part geometries must be created. They are used to produce parts in lot sizes of a few pieces to several thousand. Extremely large manufacturing lot sizes frequently call for more product-specific machines, which can be optimized for large production runs.

Complex forms such as pockets, steps, or large radii can be programmed easily on CNC machines. On conventional machines, these features require either special fixturing and tooling or a great deal of operator skill. Often several discrete machining steps can be combined on CNC machines. This saves machine setups and simplifies production processes within a manufacturing plant.

Although CNC machines can help set optimum cutting-tool parameters, the individual cuts made by a CNC machine are usually no faster than those possible on conventional machines of good quality. In fact, sometimes conventional machines can be faster.

Although CNC machines are often simpler to operate than conventional machines, they can sometimes be more difficult. Those CNC machines that allow simultaneous performance of multiple operations require operators who are knowledgable about several production steps. The operator must be able to set up and check more complex parts.

Operators of CNC machines must be able to think numerically. To adjust a machined diameter, the operator usually adds or subtracts a desired amount from the machine's offsets. With conventional machines the operator may only have to turn a crank a few graduations. A CNC machine will require the person to look up an offset on the CRT display, add or subtract the desired diameter change, then type in the new offset. Many people are not adept at such rapid mental arithmetic. In addition, some people find computers intimidating and are not comfortable working with them.

It is often assumed that a machine controlled by a computer is inherently more accurate than an equivalent manual machine, this is not necessarily true. In fact, CNC machines are sometimes less accurate than manual machines. The CNC machines rely on precision-ground slides and screws to position accurately, and normal machine wear can degrade this accuracy. The computer is usually unable to detect or compensate for this wear, while a skilled mechanist can.

Traditionally, CNC machines were mostly lathes and machining centers. Lathes turned axle shafts, drill-bit shanks, or any other multisurfaced cylindrical parts, and machining centers machined engine

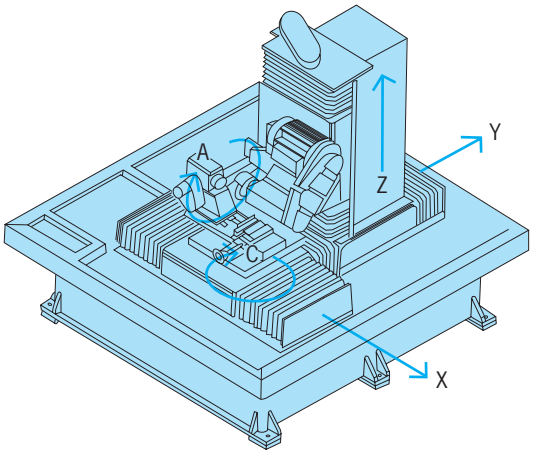

Fig. 3. Five-axis computer numerical control grinder. (*S. E. Huffman Corp.*)

blocks and aircraft rib spars. Now there are complex multiaxis grinders (**Fig. 3**), laser cutters, benders, and many more machines with computer control. Parts such as end mills, airfoils, and universal ball joints can be economically and accurately machined on CNC machines. *See* CONTROL SYSTEMS.

J. R. Casey Bralla

Bibliography. J. J. Childs, *Principles of Numerical Control*, 3d ed., 1982; T. J. Drozda et al. (eds.), *Tool and Manufacturing Engineer's Handbook*, vol. 1, 4th ed., 1983; Electro-Craft Corp., *DC Motors*, *Speed Controls*, *Servo Systems*, 1980; W. W. Luggen, *Fundamentals of Computer Numerical Control*, 3d ed., 1994.

Computer peripheral devices

Any device connected internally or externally to a computer and used in the transfer of data to or from a computer. A personal computer (PC) or workstation processes information. Data (unprocessed information) must get into the computer, and the processed information must get out. Information is entered and displayed on a wide variety of accessory devices called peripherals, which are also known as input/output (I/O) devices. Some peripherals (such as keyboards) are only input devices, others (such as most printers) are only output devices, and some are both. Peripherals may be located inside or outside a computer's case (called an internal or external device, respectively). The I/O bus is the collection of circuits where an internal device is attached. An I/O port (such as parallel or serial port) is a connector where an external device is connected (plugged in). *See* DIGITAL COMPUTER; MICROCOMPUTER.

Monitor. The monitor is the device on which images or text produced by the computer operator or generated by the program are displayed on a cathode-ray tube (CRT) or a liquid-crystal display (LCD). Electron guns in a CRT (one in a monochrome monitor, three in a color monitor) irradiate phosphors on the inside of the vacuum tube, causing them to glow. Aside from the presence or absence of color, monitors are mainly distinguished by the size of the screen and the sharpness of resolution. Screens generally range from about 7 in. (18 cm), measured diagonally, for small portable computers to 21 in. (54 cm) for graphic-intensive work. Sharpness is determined by the number of shades of gray, if any, or colors that the monitor, in conjunction with the computer's circuits and memory, can display; the number of pixels (picture elements) per inch; and the dot pitch, the size of the individual phosphorescent dots.

The flat-panel displays on most portable (notebook, laptop, or tablet) computers are known as liquid-crystal displays. LCD flat-panel monitors are superior to CRTs as computer monitors for desktops for a variety of reasons, among them being that CRTs contain hazardous materials (such as phosphors and lead), consume more power, generate more heat, and are more bulky. In place of the vacuum tube coated with phosphors, LCDs use two polarizing filters with liquid crystals between them to produce the image.

A monitor can be connected to an analog or digital video interface. Some flat-panel LCD monitors can be connected as both an analog monitor to one computer and a digital monitor to another and include a "toggle" switch to connect to either. *See* CATHODE-RAY TUBE; ELECTRONIC DISPLAY; LIQUID CRYSTALS.

Keyboard. The computer keyboard, based on the typewriter keyboard, contains keys for entering letters (usually in the familiar QWERTY arrangement), numbers, and punctuation marks, as well as keys to change the meaning of other keys. For example, pressing the "s" key on any keyboard produces a lowercase s, and pressing the "shift" and "s" keys produces an uppercase S. But pressing the "control" and "s" keys on computers running some programs saves a copy of the file being worked on to the disk or diskette, whereas this function is accomplished by pressing the "command" and "s" keys on other systems' keyboards. Other special keys or combinations of keys move the section of the file being viewed, such as "page up," move the cursor (pointer); or perform other tasks, such as "print screen." The function keys, usually a row across the top or along the left side of the keyboard and numbered F1 through F10 or F12, perform tasks that vary from program to program. In some programs, for example, the F1 key brings a "help" file to the screen. Keyboards generally are physically connected to a computer via one of three interfaces: specialized keyboard interface, a serial port, or a universal serial bus (USB). Keyboards also can connect through a wireless IR (infrared) port or Bluetooth (2.4 GHz) port.

Mouse. A mouse is a device that is rolled on the desktop to move the cursor on the screen. The mouse has one to three buttons. In the syntax of computers, the cursor is moved to select the object on which an action is performed, and then one or more buttons are pressed to perform the action. A ball on the bottom of the mouse translates the device's movements to sensors within the mouse and then through the connecting port to the computer. There are also mice that substitute optical devices for mechanical balls, and mice that use wireless infrared

or Bluetooth technology. Many mouse actions can be accomplished with a keyboard using substitute keystroke sequences.

Trackball. A trackball is essentially an upside-down mouse, with the ball located on the top rather than on the bottom. Since a trackball is attached and therefore does not require an area in which to be moved, it requires less working space than a mouse (although most trackballs occupy more standing space than a stationary mouse).

Touchpad. A touchpad is a device that senses both a user's finger movement and downward pressure exerted on the device (for example, tapping). The movement of the user's finger on the pad results in a corresponding cursor movement on the monitor in a manner corresponding to the movement of a mouse.

Pointing stick. A pointing stick (for example, the IBM TrackPoint™) is a small eraserlike device that is located in the center (between the "g", "h," and "b" keys) of a keyboard. A user applies pressure on this device to the point in the direction desired for the movement of the cursor.

Joystick. A joystick is a pointing device with a vertical handle that is used principally for games. Moving the handle moves the cursor or other simulated object on the screen; pressing the button or buttons on the base or vertical piece performs an action that depends on the software program.

Touch-enabled device. Touch-enabled devices allow an individual to experience a sense of touch. Such a device typically works with a user's computer screen to create a sense that one is actually touching what appears on the screen.

Light pen. A light pen performs the same functions as a mouse or trackball, but it is held up to the screen, where its sensors detect the presence of pixels and send a signal through a cable to the computer. Light pens seem to be a more natural transition from paper and pencil to computers than mice or trackballs, but they have proven to be tiring to use.

Graphic tablet. A graphics, or digitizing, tablet is a pad with electronics beneath the surface on which one can write or draw using a pointed device, called a stylus. The shapes drawn appear on the device's screen. For any but the simplest projects, it is easier to draw with a stylus and tablet than with a mouse or trackball. Tablets are used on PDAs (personal digital assistant) and tablet PCs (a form of notebook computer).

Microphone. A microphone, sometimes called a mic (pronounced "mike"), is a transducer that converts sound into an electrical signal. Microphones are used as input devices in many applications where speech recognition is involved.

Disk drive. The term disk drive has taken on new meanings in recent years. Originally, disk drives were magnetic media devices. Now, there are many devices that connect via a universal serial bus (USB) port and look like a disk drive to the operating system; however, they are devices with no moving parts that use flash memory.

Magnetic media (hard drives and diskettes). The most common input/output, or storage, device in personal computers or workstations is a hard disk drive, a stack of magnetized platters on which information is stored by heads generating an electrical current to represent either 1 or 0 in the binary number system. The device is called hard because the platters are inflexible, and is called a drive because it spins at 3600 revolutions or more a minute, within a sealed case. The heads are only a few millionths of an inch from the platters. Storage capacity of hard disk drives has grown from a few megabytes of data (1 megabyte = 10^6 characters or bytes) to 40 or more gigabytes (10^9 characters). Diskettes, made of flexible plastic (usually polyester) film like that used in recording tape, are stored within a hard shell and are spun by their drives at about 360 revolutions per minute. Their capacity is quite limited, commonly about 1.1 megabytes (3.5-in. diskette), although both smaller and larger ones are available. Since they are removable, they are useful for transporting data.

Bernoulli boxes, zip drives, and jazz drives are similar to a hard drive in storing larger amounts of information and similar to a diskette drive because they can be removed and replaced.

A disk drive as described above is an output device when writing information to the device. When stored programs and data are needed again, it is an input device, reversing the storage procedure to transfer the instructions or information to the computer's working memory for use or further processing. *See* COMPUTER STORAGE TECHNOLOGY.

USB drive. A USB drive, also called a jump drive or flash drive, is a device smaller than a key (it often is placed on a keychain and hence is called a keychain drive) that connects to a computer via the universal serial bus port. USB drives range in size from 16 megabytes to multiple gigabytes of information. They are preferred to floppy disks and zip drives because of their size and capacity. Other devices that connect to a USB through a special adapter and look like a disk drive include compact flash cards and memory sticks. *See* SEMICONDUCTOR MEMORIES.

Optical drives. Most programs and data are stored on magnetic media, including diskettes, hard disks, and tapes. But data can also be stored and retrieved with light, the light of a laser beam reading a pattern of pits on a so-called optical disk.

CD-ROM drive. The most familiar type of optical disk is the CD-ROM (compact disk read-only memory). It is called a compact disk because it is almost identical to audio compact disks; it is read-only because the data can be viewed or heard on the computer but not altered or added to. CD-ROMs have moved from such specialized uses as storing vast amounts of information for businesses and government agencies to consumer products offering software, especially games and reference works, enhanced with video clips, animation, and music and sound effects. CD-R (write) drives are write-only devices, while CD-RW (rewritable) drives can be rewritten numerous times. The capacity of a standard CD-ROM is 650–700 megabytes.

DVD drive. A DVD (digital videodisc) drive is similar to a CD-ROM drive and is backward-compatible in that the can read CD-ROMs. Standard DVDs have a minimum capacity of 4.7 GB of data. There are a number of different formats for DVDs (such as DVD-R, DVD+R, DVD-RW, and DVD+RW). DVD+RW and DVD-RW drives are rewritable. DVDs are excellent backup devices for personal computers. *See* COMPACT DISK; OPTICAL RECORDING; VIDEO DISK.

Tape drives. The early personal computers of the 1970s stored programs and data on tape cassettes in recorders similar or identical to those of audio tape recorders. The drawback with tape storage is that it is sequential; that is, the tape must be gone through from the beginning until the desired data are reached. Disks, whether floppy or hard, are random-access, which means they can go directly to the data. Because of their greater efficiency and speed, disk drives have quickly replaced tape drives as the primary means of data and program storage. Tape drives are still in use for backup storage, copying the contents of a hard disk as insurance against mechanical failure or human error. The relative slowness of tape drives does not matter when they are used for backup because the recording can be scheduled to begin at the end of the workday or work week, and the recovery of the data may never be necessary. Usually, the entire hard disk is backed up in the first session. In later sessions, only files that have changed are recorded (incremental backup). Tape capacity has grown from tens of megabytes of data to more than 100 gigabytes of data, depending on the format and how much data compression is used. *See* MAGNETIC RECORDING.

Camera. A digital camera can be connected to a computer through a USB port. It essentially looks like a USB drive to the computer system. Hence, photos can be moved to and from the camera with simple copy commands or automatically uploaded to the computer via special software supplied with the camera.

A PC or Web camera provides video input to a computer. Such cameras can be used for a variety of applications, including security monitoring. A common use of such cameras is for online "chatting" in which participants can see each other as well hear each other. Such a use requires broadband connections for high-speed transmission and receiving of data.

Scanner. A scanner converts a hard-copy image, such as text, a drawing, or a photograph, into a digital image and sends into the computer for display or further processing. Hand-held scanners are rolled over the image, while flatbed scanners hold the image on a sheet of glass. The image is viewed as a graphics image, not a text image, so it can be altered with a graphics program but not edited with a word-processing program unless the scanner is part of an optical character-recognition (OCR) system. To digitize photographs, a scanner may dither the image (put the dots a varying amount of space apart), or use the tagged image file format (TIFF), storing the image in 16 gray values. Some scanners can use standard video cameras to capture images for the computer. *See* CHARACTER RECOGNITION; IMAGE PROCESSING.

Bar-code reader. Bar-code readers are specialized scanners. A bar-code reader is able to scan and interpret bar codes, the most common of which is a UPC (universal product code) stored as a one-dimensional bar code, where each digit is represented by the ratio of the widths of adjacent stripes and white areas. Such bar-code readers are found on products in grocery stores and other establishments. Two-dimensional bar-code readers are able to read bar-codes that store information along the height as well as the length of the code. Common examples of such bar codes are those printed on income tax forms by tax software.

RFID reader. An RFID (radio-frequency identification) reader is used to poll an RFID tag with radio signals sent via its antenna. The tag in turn responds with identification information sent to the reader. Hence, RFID readers are technically input/output devices.

Printer. This device puts text or other images produced with a computer on paper or other surfaces. In the early days of personal computers, it was widely believed that computers would lead to the paperless office. But because of the convenience of paper copies and the ease of duplicating them, the personal computer and printer have resulted in a great increase in paper documents. Printers are either impact or nonimpact devices.

Impact printers. A daisy-wheel or thimble printer is so called from the shape of the elements bearing raised images of the characters that strike the paper through an inked or carbon ribbon. They produce printouts similar to those from a typewriter. The quality of the printouts is high, but they are limited to the characters contained on the printing element, and their speed, perhaps 30 characters per second, which is considered too slow.

Dot-matrix printers produce their images by striking a series of wire pins, typically 9, 18, or 24, through a ribbon in the pattern necessary to form the letter, number, line, or other character. The quality of the printout varies. A matrix with a small number of pins will produce characters in which the individual dots can be distinguished; one with a large number of pins can produce characters of a crispness that rivals the quality of daisy-wheel or laser printers. Some dot-matrix printers can strike a line a second time, with the dots slightly offset to produce more solid characters but with a penalty of reduced speed.

A thermal printer uses heated pins against heat-sensitive paper. The disadvantage of this type of printer is that it cannot print on standard paper. Many store receipts are printed on small thermal printers.

Line printers generally print a line at a time. There are two types of line printers: drum printers and chain (or "train") printer. Chain printers use a chain with a print "slug" for each character that can be printed. As the chain moves across the printer, a print hammer strikes the slug against a ribbon from the back of the paper resulting in the struck character being printed in that position. A drum printer uses a

similar method of hammers behind the paper striking against a rotating drum. Chain and drum printers can operate at speeds of thousands of lines per minute.

Nonimpact printers. Laser printers are similar to photocopying machines. A laser beam is directed by a rotating mirror to draw the image on a drum, where it is changed to an electrostatic charge that attracts a fine powder, called toner. The paper, also charged, rolls on the drum and attracts the toner's image, which is then fused to the paper by heat. The image can be formed, with a sharpness of 300 dpi (dots per inch) or more, at a speed of four pages or more per minute. The quality of laser printer output is the highest generally available. However, laser printers cannot print through multipart forms or, usually, on wider paper, so dot-matrix and even daisy-wheel printers have not been totally abandoned. *See* PHOTOCOPYING PROCESSES.

Ink-jet printers carry their ink in a well, where it is turned into a mist by heat or vibration and sprayed through tiny holes to form the pattern of the character on paper, much as a dot-matrix printer does by direct impact with the ribbon. The quality of the printout rivals that of laser printers, although ink-jet printers can suffer problems with clogging, smearing, and speed.

Color printers. Printers designed to print images in one color or a full range of colors may use similar technologies to black-and-white printers. Most color printers use ink-jet and laser technology. Some color printers are thermal-transfer printers, in which heat melts colored wax onto the paper. Many color printers also provide for photo printing of digital photos. *See* PRINTING.

Modem. A modem (modulator demodulator) connects one computer to another, ordinarily through the telephone lines, to exchange information. A modem sending data modulates the digital information from the computer into tones, the analog signals that can be transmitted over a telephone line. A modem receiving data demodulates the tones coming from a telephone line back into digital information for its computer. The speed of a modem is measured in bits per second (bps), and the speed for most dial-up modems is 28 to 56 kbps (kilobits per second). [The baud is occasionally and incorrectly used as a synonym for bps.] Higher speeds are preferable not only for convenience but also for reducing telephone charges. The communicating modems operate at the highest speed that they have in common, or the highest speed that the quality of the telephone connection permits. Most modems also send and receive facsimile images. A cable modem operates over a coaxial or fiber-optic cable. The speed of cable modems is over 1 Mbps (megabits per second). *See* MODEM.

Multifunction devices. While printers are typically considered output devices, some are multifunction devices that include a scanner, a laser or color ink-jet printer, and a fax modem. Because the multifunction nature of these devices, they can also operate as a standalone fax machine, a copier, or a color copier (if the printer is a color printer). *See* FACSIMILE.

Malcolm G. Lane; L. R. Shannon

Bibliography. D. Downing, M. A. Covington, and M. Mauldin Covington, *Dictionary of Computer and Internet Terms*, 8th ed., Barrons Educational Series, 2003; *Microsoft Press Computer Dictionary*, 5th ed., 2002.

Computer programming

Designing and writing computer programs, or sequences of instructions to be executed by a computer. A computer is able to perform useful tasks only by executing computer programs. A programming language or a computer language is a specialized language for expressing the instructions in a computer program. A person engaged in computer programming is commonly called a programmer. Two major stages can be identified in a typical process to develop a computer program: problem solving and implementation.

Problem solving. In this stage the programmer gains a full understanding of the problem that the computer program under development is supposed to solve, and devises a step-by step procedure (an algorithm) that, when followed, will solve the problem. Such a procedure is then expressed in a fairly precise yet readily understandable notation such as pseudo code, which outlines the essentials of a computer program using English statements and programming language–like key words and structures. *See* ALGORITHM.

A useful example concerns the problem of finding the greatest common divisor (gcd) of two given positive integers. After an analysis of the problem, a programmer may choose to solve the problem by the procedure described by the following pseudo code. (In the pseudo code, the modulus operator produces the remainder that results from the division of one integer by another; for example, 15 modulus 6 yields 3.)

```
1) Let x and y be the two given integers
2) As long an x and y are greater than 0,
   repeat lines 3–5
3)      if x is greater than y
4)          then replace x by the value of x
            modulus y
5)          else replace y by the value of y
            modulus x
6) If x is 0
7)      then y is the gcd
8)      else x is the gcd
```

See NUMBER THEORY.

Lines 3–5 and 6–8 are examples of the selection control structure, which specifies alternative instructions and enables a computer to choose one alternative for execution and ignore the other. Lines 2–5 form a repetitive control structure (commonly called a loop), which causes the computer to execute certain instructions (lines 3–5) repeatedly.

Programmers indent to show the structural relationship among the statements and to enhance the readability of the program. In the example, indentation makes it clear that lines 4 and 5 are part of the selection structure beginning at line 3, that lines 3–5 are part of the loop that begins at line 2, and that lines 7 and 8 are part of the selection structure that begins at line 6.

An algorithm developed to solve a given problem should be verified to ensure that it will function correctly. The following list shows how the above procedure derives the gcd of 48 and 18 by successively modifying the values of x and y.

x	y	
48	18	(Let x be equal to 48 and y equal to 18)
48	18	(48 modulus 18 is 12, which replaces 48)
12	18	(18 modulus 12 is 6, which replaces 18)
12	6	(12 modulus 6 is 0, which replaces 12)
0	6	(since x is 0, the loop ends and 6 is the gcd)

Implementation. In this stage the pseudo-code procedure developed in the problem-solving stage will be expressed in a programming language. There are numerous programming languages, in two broad classes: low-level languages, which are difficult for human programmers to understand and use but can be readily recognized by a physical computer; and high-level languages, which are easier for human programmers to use but cannot be directly recognized by a physical computer. Nowadays most computer programs are first written in a high-level language and then translated into an equivalent program in a low-level language so that a physical computer can recognize and obey the instructions in the program. The translator itself is usually a sophisticated computer program called a compiler. *See* PROGRAMMING LANGUAGES.

Shown below is the gcd-finding procedure written in a widely used high-level language called C++.

```
/*************************************************************
 * The function findgcd computes the greatest   *
 * common divisor of two int parameters, and    *
 * returns the result as an int value           *
 *************************************************************/
 int findgcd (int x, int y)
 {   while (x > 0 && y < 0)
         if (x > y)
                 x = x % y;
         else    y = y % x;
     if (x == 0)
                 return y;
         else    return x;
 } .
```

The text beginning with /* and ending with */ (the first five lines above) is a program comment. Program comments are not executable instructions and do not affect the functioning of a computer program in any way, but are intended as a way to document a computer program. Appropriate program comments enhance program readability and are considered an important part of a computer program. A more readable program is usually easier to enhance or modify when such needs arise later. The rest of the above text is C++ code (instructions in the language C++) to compute the gcd of two given integers. Indentation in C++ code serves the same purpose as in pseudo code. It is obvious that the above C++ code closely parallels the pseudo code developed previously, as they both express the same abstract procedure.

Syntax errors and logic errors. When a program in a high-level language is being compiled (translated) into machine-executable instructions, the compiler is usually able to detect syntax errors in the program. Syntax errors reflect the incorrect use of a programming language and must be corrected before a program can be successfully compiled. However, error-free compilation of a program does not guarantee that the program will function correctly when it is executed by a computer. Execution of such a program on a computer may not yield the desired result if there is a logic error in the program. Logic errors often require more skill to correct than syntax errors. In order to identify the source of an error in program logic, a programmer may need to revisit both the problem-solving and the implementation stages in the program development process.

Top-down design. Seldom is a large computer program constructed as a single, monolithic entity. The problem that a program is intended to solve can usually be divided into subproblems, each of which can often be further divided into even smaller subproblems. The process can be repeated until all subproblems are sufficiently small and easy to solve. Known as top-down design or stepwise refinement, such repeated decomposition of a large, complex task into smaller, more manageable ones is a widely accepted principle for designing large, complex computer programs.

Object-oriented programming. Object-oriented programming is a way to structure a computer program. The object-oriented paradigm has gained widespread acceptance in recent years and is supported by such widely used languages as C++ and Java. An object includes relevant data and operations on the data as a self-contained entity. Interaction with an object can be made by invoking the object's operations. Each object is an instance of an object class. Object classes may be related by inheritance; one class may be a subclass (specialization) of another class. At the center of an object-oriented program design is a collection of objects that represent entities in the application domain. Identifying the objects and object classes, the relationship among the object classes, and the interactions among the objects is a major issue in designing an object-oriented program. *See* OBJECT-ORIENTED PROGRAMMING.

Language processors. Although many widely used programming languages are processed by compilation into machine-executable instructions, some programming languages (for example, LISP and Prolog) are usually interpreted instead of being compiled. When a computer program is interpreted,

it is directly executed by another program called an interpreter without being translated into low-level, machine-executable instructions. Some other languages are processed by a hybrid approach. For example, a program in the language Java is first compiled into an equivalent program in an intermediate-level language, which is executed by an interpreter. *See* COMPUTER; DIGITAL COMPUTER. Samuel C. Hsieh

Bibliography. F. M. Carrano, P. Helman, and R. Veroff, *Data Abstraction and Problem Solving with C++*, Addison Wesley Longman, 1998; N. Dale, C. Weems, and M. Headington, *Programming and Problem Solving with C++*, Jones and Bartlett, 1997; C. Horstmann, *Computing Concepts with Java Essentials*, John Wiley, 1998.

Computer security

The process of protecting against threats to computing systems. A threat is an event that can cause harm to computers, data or programs, or computations. A failure of computer security occurs because of a vulnerability or weakness in a computing system. A threat agent—person, event, or circumstance—exploits a vulnerability. Computer security involves protecting against failures of availability, integrity or correctness, and confidentiality or privacy.

A failure of availability is also known as denial of service. Partial denial of service is lack of capacity or unacceptable responsiveness. Computer users expect programs and data to be available on demand to meet computing needs. Applications such as power generation, stock trading, and even airplane cockpit navigation and aspects of medical care have become so dependent on computing that loss of availability becomes a serious threat to life or society. Even on a less dramatic level, people have become dependent on computers in aspects of everyday life, and so maintaining expected availability of computers is probably the most important of the three goals of computer security.

Integrity means different things in context: sometimes it means that data or programs should be returned exactly as they were when they were recorded, or that modifications to data or programs should be made only by authorized persons, or by authorized programs, or in certain ways, or that the quality of data should be maintained. Data and programs should meet the demands of the way in which they are to be used.

The third category of failure is loss of confidentiality or privacy or inappropriate disclosure. Privacy of personal information or of data about individuals is a significant concern, as is that of sensitive corporate data (such as trade secrets and proprietary information) and government classified data.

The categories of security failures relate to authorized or acceptable behavior. In order for actions to be authorized or acceptable, a security policy must define who is allowed what access to what resources, what is to be protected, against what threats, to achieve what goals.

Computer users explicitly or implicitly perform a risk analysis by assessing the threats to computing, the likelihoods of these threats, and the harm that could occur from actualization of these threats. Risk must be balanced against the value of computing. Users select controls that prevent exploitation of one or more vulnerabilities as a means of lowering risk. Controls can be physical (such as a lock on a door), procedural (such as a policy that only supervisors can modify critical data items), or technical (such as a program control that blocks access to the password file). Controls can alter the balance among the three categories of failure: a high degree of confidentiality can reduce the availability of a system. *See* RISK ASSESSMENT AND MANAGEMENT.

Technical Controls

Technical controls can be a part of a computing system or application, or they can be added separately. If controls are added to an existing system, they must be integrated properly so that they function as expected.

Security enforcement mechanisms. A computing system must first be able to determine with high assurance the identity of the person who is seeking access (a process called identification and authentication), and the system must then verify that this person's proposed access is consistent with the security policy (a process called access control).

Identification and authentication. Identification is typically performed by logging in or entering a username. But after entering a name, a user may be asked to prove it, so that the system can be certain that one user is not trying to impersonate another. A user can authenticate an identity based on (1) what the user knows, such as a password, (2) who the user is, such as a physical characteristic (for example, a fingerprint), or (3) what the user possesses, such as a token. An authenticator must be something that cannot be easily forged, lost, forgotten, or guessed, while still making it easy for the legitimate owner to use. Techniques can use two or more approaches.

User passwords are commonly employed. Password guessing attacks use computers and actual dictionaries or large word lists to try likely passwords. Brute force attacks generate and try all possible passwords. To block these attacks, users should choose strong passwords.

Physical characteristics can be determined by biometric devices. In addition to fingerprints, voice recognition, retina patterns, and pictures are used. Although biometric authentication devices can be expensive, they are less susceptible to forgery and guessing than other methods, and they tend to be easy for users to adopt. *See* FINGERPRINT.

Tokens are widely used. Identity cards containing magnetic stripes or other elements that can be sensed combine computer access control with physical access control. Some identity cards contain small processors; these cards, called smart cards, can perform computation to guard against forgery. Another form of token resembles a pocket calculator; it is used for challenge-response, in which the user is

given a challenge, a problem to solve, to which the calculator provides part of the answer. Each time the user seeks authentication, a different challenge is set, with a different corresponding answer. Challenge-response is especially desirable with computer network authentication because, even if someone intercepts the challenge and response, the interceptor cannot later reuse the response for impersonation because the challenge will be different.

Access control. The system uses a validated user identity to limit the actions the user can perform. An access control policy is a series of acceptable triples (user, object, action), such as (system administrator, password file, modify), meaning that the user "system administrator" is allowed to perform the action "modify" on the object "password file." An access control list (ACL) is a set of these triples. Access control lists can be represented as a two-dimensional matrix, as a set of rules, or in other ways.

Before permission to access an object is allowed, a reference monitor (also known as a reference validation mechanism or access control mechanism) checks that the access is allowable. A reference monitor must be complete (invoked to validate every reference permission), correct (made to implement the intended access control policy exactly), and tamperproof (unable to be disabled).

Reference monitors can simply process a representation of the access control policy in list or table form. Alternatively, they can process capabilities, which are prevalidated access "tickets." The access control system gives a user a capability to perform a certain access on a particular object, and the user later presents the capability to a reference monitor, which will inspect the capability and allow the access. Capabilities are useful in networked and distributed systems, in which access control may be done at one point and actions on objects may be done elsewhere. The Kerberos system is a popular distributed access control system based on capabilities.

Cryptography. A valuable tool in computer security is cryptography or encryption. Encryption is the process of transforming data so that the data are not readily understandable to unauthorized users; consequently unauthorized people cannot modify encrypted data in a knowledgeable way. Encryption is an important tool in computer security because it provides concealment and immunity to modification, so that it can be used to protect integrity and confidentiality. Cryptography alone is not a security control; it is a technology upon which controls can be based.

Encryption is a transformation function. Data to be encrypted are called plaintext, and the encrypted form is called ciphertext. The encryption function E is denoted $C = E(P)$, where P is a body of plaintext and C is its resulting ciphertext. There may also be a decryption function D that reverses the effect of encryption, so that $P = D(C) = D(E(P))$.

Encryption is a means by which two people can share a secret: one encrypts the secret and the other decrypts it. To facilitate sharing secrets using encryption, people use keyed encryption. A key is an additional parameter that tailors the encryption function, in the same way that a key tailors a basic lock so that it can be opened only by someone possessing the key. The key must be supplied to encrypt or decrypt, and so a single algorithm can represent many different encryptions with different keys.

Key management—arranging the distribution, revocation, and replacement of cryptographic keys—is very difficult, especially in a situation with many users who have no common basis for trust, as is true in many computing systems. The implementation of encryption must be secure in order for the encrypted results to protect the data adequately. *See* CRYPTOGRAPHY.

Security of Programs

Computer programs are both part of the protection and part of the things protected in computer security. Programs implement access controls and other technical security controls. But those same programs must be protected against accesses that would modify or disable their ability to protect. And those programs must be implemented correctly.

Correctness, completeness, and exactness. A computer program is correct if it meets the requirements for which it was designed. A program is complete if it meets all requirements. Finally, a program is exact if it performs only those operations specified by requirements. Computer security requires correct, complete, and exact programs, and nothing more. A program has inevitable side effects. For example, a program inevitably assigns values to internal variables, uses computing time, and causes entries to be generated in audit logs. Although side effects seem benign, they can be used maliciously to leak information. The exactness requirement really concerns only those significant operations specified by requirements, but in security almost any side effect can be significant. Determining which additional actions are security-relevant is difficult, if not impossible.

Correctness and completeness can be determined to some degree by careful testing, although with large or complex systems it may be infeasible to test all possible situations. It is difficult to test security systems appropriately, because they can be large and complex, and because it is hard to simulate all the environments and approaches by which systems can be attacked.

Malicious code. Computing is so fast and complex that users cannot know everything a program is doing. Programs can be modified or replaced by hostile forms, with the replacements seeming outwardly the same as the originals. The general term "malicious code" covers Trojan horses, viruses, worms, and trapdoors. Malicious code has been present in computing systems since the 1960s, and it is increasingly prevalent and serious in impact. Unfortunately, there are no known complete forms of protection against malicious code.

A Trojan horse is a program that has an undocumented function in addition to an apparent function.

For example, a program may ostensibly format and display electronic mail messages while also covertly transmitting sensitive data.

A virus is a program that replicates and transfers itself to another computing system. When executed, each copy can also replicate, so that the infection spreads at a geometric rate. A virus typically inserts its replicated copy into another executable program so that when the other program is executed, so is the copy of the virus. Viruses often perform covert malicious actions.

A worm is a program that, like a virus, seeks to replicate and spread. However, the goal of the worm is only to spread and consume resources. The malicious effect of the worm is denial of service by exhaustion of resources.

A trapdoor is an undocumented entry point into a program. The trapdoor is inserted by a programmer to allow discreet access to a program, possibly with exceptional privileges. A user who had legitimate access at one time might have installed the trapdoor as a means of obtaining access in the future.

All these forms of malicious code are serious security threats for several reasons. First, malicious code can be relatively small, so that it is not readily detected. Second, its actions can be concealed: If a program fails to perform as it did, the change is evident, but an attacker can cause the change to be subtle, delayed, or sporadic, making it very difficult to detect, let alone diagnose and correct. The covert effect of malicious code can be almost anything: It can delete files, transmit messages or files, modify documents or data files, and block a user from accessing a computer system. The attack can be transmitted in pieces that activate only when the entire attack has been delivered. Finally, protecting against malicious code is difficult: The only known totally effective countermeasure is not to accept any executable items from anyone else, a solution that is scarcely acceptable for current networking and information-sharing environments.

Security of code. It is infeasible for a user to determine that a program is secure. The user has little evidence on which to base an opinion, an insecure program may intentionally hide its weaknesses, and many users have little control even over the sources from which programs are derived. Even well-intentioned programmers can fail. Beyond principles of good software engineering, the security field uses rigorous independent evaluation and penetration testing to add confidence that a program is secure. *See* COMPUTER PROGRAMMING; SOFTWARE ENGINEERING.

Independent evaluation. Users need an independent quality assessment, both that the programs meet their stated requirements and that they have no obvious security requirements. An international effort has led to the publication of the *Common Criteria for Information Technology Security Evaluation* (CC). The successor to the U.S. *Trusted Computer System Evaluation Criteria* (TCSEC or "Orange Book") and the European *Information Technology Security Evaluation Criteria* (ITSEC), the CC specifies a structure for defining security and functionality requirements (a protection profile) and a description of a product meeting those requirements (a security target) so that the security target can be evaluated objectively against the protection profile. The evaluation then serves as a seal of approval that, with a stated degree of confidence on the part of the evaluators, the product has met the requirements. The CC scheme has been applied to operating systems, firewalls, intrusion detection systems, cryptographic systems, database management systems, and networking products.

Penetration testing. Because programs intended to enforce security must be robust enough to enforce it rigorously, a special form of security testing is sometimes performed. Ordinary testing is directed toward verifying that the program does what it is expected to do; penetration testing is intended to determine if there is a way to make the program not do what it is expected to do. Penetration testers use known flaw types, common actual (and well-known) attacks, and intuition to study the system, hypothesize flaws that would prevent it from enforcing security, and conduct experiments to confirm or refute the flaws hypothesized.

Security of Operating Systems

Operating systems are the basis of security enforcement in a computing system because the operating system has control over the system resources: the file system, printers, shared buffers, memory, time, and shared programs. The original purpose of an operating system was to promote efficient and equitable sharing of computing resources, which led swiftly and naturally to security enforcement.

Operating system security functions. Operating systems are the first contact of a user with a computing system. Thus, the operating system performs user identification and authentication, as a basis for controlling resource usage and providing each user with a unique computing environment. Operating systems control access to primitive resources, such as files, devices, network connections, programs, and computing time. Most operating systems generate an audit log of what security-relevant events have occurred to the resources they control.

Operating system design for security. Operating systems are the point of fundamental security enforcement. Consequently, for strong security enforcement, security cannot be added successfully to an operating system after the design is partially completed. The reference monitor is the gate that controls access to all security-relevant objects. Several examples of secure operating systems have been implemented based on the reference monitor concept. These examples have involved thoughtful design followed by intense analysis of the design and its implementation. *See* OPERATING SYSTEM.

Database Security

A database is a collection of records containing fields, organized in such a way that a single user can be allowed access to none, some, or all of the data.

Typically the data are shared among several users, although not every user will have access to every item of data. A database is accessed by a database management system that performs the user interface to the database.

Integrity is a much more encompassing issue for databases than for general applications programs, because of the shared nature of the data. Integrity has many interpretations, such as assurance that data are not inadvertently overwritten, lost, or scrambled; that data are changed only by authorized individuals; that when authorized individuals change data, they do so correctly; that if several people access data at a time, their uses will not conflict (for example, two people cannot write a single data item at the same time); and that if data are somehow damaged, they can be recovered.

A database management system is an application program that does not necessarily have a reliable interface to the operating system, and the operating system cannot necessarily be highly trusted. Therefore, a database management system typically maintains its own lists of users and their permitted actions, and the database management system may perform its own identification and authentication of users, quite independent from that performed by the operating system.

Inference and aggregation. Database systems are especially prone to inference and aggregation. Through inference, a user may be able to derive a sensitive or prohibited piece of information by deduction from nonsensitive results without accessing the sensitive information itself. For example, knowledge of any individual's salary may be restricted, but statistical measures (such as mean, median, maximum, and minimum) of the general salary pool may be freely released. If a user can formulate a database query that selects exactly one record (for example, "select all employees who are 6′2″ tall and live in Glenwood, Maryland"), the user can then determine a single individual's salary by asking for the mean salary for the select set (of one employee). Various statistical methods make it very difficult to prevent inference.

A related problem is aggregation, the ability of two or more separate data items to be more (or less) sensitive together than separately. For example, neither the longitude nor the latitude of the location of a secret gold mine is sensitive by itself, but these two values together pinpoint the mine. It is infeasible to release one value to some people and the second value to others (even though no single person is told both) because two people could pool their knowledge. Aggregation is extremely difficult to prevent, since users can access great volumes of data from a database over long periods of time and then correlate the data independently.

Database controls. Encryption is sometimes used in database management systems for both confidentiality and integrity. A device called an integrity lock is used to ensure that the contents of a field are not modified. The field may be stored in plaintext, but a cryptographic checksum of the field, the integrity lock, is also stored. Each time an authorized change to the field is made, a new integrity lock is computed and stored; each time the contents of the field are retrieved (in response to a query), a lock is computed on the value of the field and compared with the stored lock. If the two values do not match, the field has been modified without authorization. To guard against swapping data and their locks, the encrypted data may comprise the actual data of the field together with the record identifier and the name of the field. *See* DATABASE MANAGEMENT SYSTEM.

Network Security

Enforcing security across a network has all the problems of local enforcement, coupled with problems brought on by distance and transmission. Identification and authentication, for example, requires not just verifying the identity of a requesting user but also identifying the system to the user. A user does not want to supply a password (or a credit card number or other sensitive information) to the wrong system. Even if the user is assured initially that the remote system is authentic, the user may need assurance continuously throughout a transaction that the remote system has not been commandeered by an attacker or that the communication has not been diverted to a hostile system.

In a network, the same considerations apply as with nonnetworked systems: identification and authentication, access control, and auditing. The two additional areas of concern are the communications link between the local user and the remote system, and the physical and technical security of the remote system. (In fact, the remote system may be multiple systems, with the user's communication passing from one to another without the user's knowledge.)

Communications security. It is necessary to protect the communications lines by which users communicate with remote systems. Lines are subject to interception, redirection, falsification, and interruption, and the attacks range from very crude (simply cutting a wire) to very sophisticated (intercepting a message, modifying portions of it, and retransmitting the modified message).

Strong encryption is very effective in protecting against interception and modification. It is useful to consider a communication from A, through intermediate points X and Y, and ultimately to B. With end-to-end encryption, the message is encrypted at A and travels in encrypted form all the way to B. End-to-end encryption can be applied from an application to a corresponding application, so that the message is protected even as it is stored temporarily on a file system or as it passes through the operating system. This approach requires that part of the message, the routing and destination information, not be encrypted, so that along the way X and Y, which do not share the encryption between A and B, are able to move the message toward its destination.

An alternative is link encryption, in which a message is encrypted just along one path, say from A to X. At X the message is decrypted and reencrypted for the link from X to Y, and so forth. Link encryption

can be applied at the physical link from A to X, so that all communications along that path are protected. Key management may be simpler than with end-to-end encryption because the encryptor needs to have keys only for its neighbors, the next points to which it sends communications. The individual user does not have to apply the encryption, manage cryptographic keys, or even understand encryption. However, a message is exposed at X and Y and also within the source and destination computing systems up to the application that handles the message.

Virtual private networks. Organizations can benefit from sharing access to computing systems that are not all located together. To do this, organizations have established virtual private networks (VPNs). These networks approach the security of a private network at costs closer to those of shared public resources. The primary security technique used is encryption. A virtual private network involves a set of known participants since an organization knows which systems will form a part of its network. Thus, cryptographic key management is reasonably straightforward: By some secured means, such as postal mail or a telephone conversation, a seed or master cryptographic key is distributed to all the systems that will participate in the private network. One party of the virtual private network then generates a new key, sometimes called a session key, that will be used for communications with that system or for a specific period of time. The session key is encrypted under the master key to be passed on to the other VPN partners.

Internet security. The Internet, or any similar public network, is subject to threats to its availability, integrity, and confidentiality. A complicating feature is that there is effectively no control on transmissions over the Internet. Consequently, a system connected to the Internet is exposed to any malicious attack that any other Internet user wants to launch.

Availability threats. Availability or denial of service threats range from disabling the connection of a system to flooding or saturating the system. The main cause of this type of threat is natural: loss of power, equipment failure, or a failed communications link. Disabling the physical connection requires access to the target system, but disabling the logical connection does not require physical access, and so it is a more likely choice for a malicious outside attacker. It requires only changing the routing information of the Internet so that the target host's address is no longer recognized. Fortunately, the Internet routing information is very well protected, so that this type of attack seldom succeeds.

A flooding attack is fairly common and difficult to protect against. The attacker either transmits (automatically) a vast amount of data to the target host, overwhelming its ability to handle legitimate communication, or the attacker exploits a flaw in the network communications protocols to cause the target host to generate its own massive volume of traffic to itself. A crafty attacker will falsify the source address of malicious communications, so that it is difficult to trace back to the attacker, making it infeasible to block or disregard communications from the malicious source.

Integrity threats. Threats to integrity include unauthorized changes to the data or programs of the target computing system. Integrity attacks seek to modify the data or programs at a target host. Malicious programs can be introduced by commandeering the target host, delivering the malicious program as if it were from a reliable source, or conveying the malicious program through a Trojan horse attack. Additionally, data can be modified by an attacker at any point on the communications link between the source and destination.

Again, encryption becomes a significant tool to control integrity. However, an advantage of the Internet is that it connects millions of sites, but the disadvantage is that there may be no shared encryption between all pairs of those sites. The difficulty is for two previously unknown entities (representing hosts or users) to authenticate each other. A certificate can be used to establish shared encryption with some degree of trust. A certificate contains an identification, an encryption key, and an encrypted field, called a digital signature, from a third party, attesting to the authenticity of the identity. Certificates can be countersigned by trustworthy third parties to add to their credibility. The encryption key in a certificate provides a basis for two unknown entities to establish a shared encryption.

Confidentiality threats. Confidentiality threats in a distributed environment are closely related to integrity threats. With minor modification, integrity attacks can be made to expose the sensitive data at the remote site or in transit. Several network protocols, such as secure http (shttp) or secure socket layer (ssl), use encryption to protect both confidentiality and integrity of communications on the Internet. *See* INTERNET; WORLD WIDE WEB.

Security perimeter. A security perimeter is a logical boundary surrounding all resources that are controlled and protected. The protected resources are called a domain (or enclave or protected subnetwork). There may be overlapping domains of varying protection, so that the most sensitive resources are in the innermost domain, which is the best protected. Protecting the security perimeter may be physical controls, identification and authentication, encryption, and other forms of access control. Two controls that relate especially to the security perimeter are network vulnerability scanning and firewalls.

Scanning. A network vulnerability scan is the process of determining the connectivity of the subnetwork within a security perimeter, and then testing the strength of protection at all the access points to the subnetwork. With a network domain, if a forgotten access point is not secured, its weakness can undermine the protection of the rest of the domain. A network scanner maps the connectivity of a domain, typically by probing from outside the domain, to determine what resources are visible from the outside. Once all outside connections are identified, each is tested with a range of attacks to determine the

vulnerabilities to which it is susceptible and from which it needs to be better protected.

Firewall. It is infeasible to seal off a domain completely against outside access, because a purpose of the domain is to support communication with the outside. For example, a web site may be intended for broad public access. A firewall is a host that functions as a secured gateway between a protected enclave and the outside. The firewall controls all traffic according to a predefined access policy. For example, many firewalls are configured to allow unhindered communication outbound (from the protected domain to a destination outside the domain) but to allow only certain kinds of inbound communication. A firewall can be a separate computer, or firewall functionality can be built into the communications switch connecting the enclave to the external network.

Some firewalls hide internal resources from being visible outside the enclave. An enclave may have registered only the address of the firewall; internal hosts are not registered as part of the Internet or external network, and therefore outsiders have no way to communicate directly with them, only through the firewall. The firewall can be configured to allow only certain types of communications, for example, electronic mail, while blocking communications that could affect the status or integrity or confidentiality of an internal host, such as network management commands. Proxy firewalls simulate the effects of complex interactions across a security perimeter, filtering undesirable information flows.

Unprotected hosts. Hosts that must be accessible to the general public are often situated outside the protected domain, so that they are readily accessible to users without exposing any of the resources inside the domain. An unprotected host is acknowledged to be exposed. Unchanging content of the host can be sealed with integrity checks to detect modification, and if the content is modified, either maliciously or accidentally, the content can be restored readily from a protected copy inside the domain. Data gathered at the open site are moved inside the protected domain for more secure storage.

Intrusion detection. It is most effective to eliminate vulnerabilities, but if that is not possible, it is then desirable to recognize that an attack is occurring or has occurred, and take action to prevent future attacks or limit the damage from the current one. Anti-intrusion technologies are the different ways in which an attack can be detected and countered. Anti-intrusion technologies are prevention, to block an attack; deterrence, to make an attack sufficiently difficult to discourage the attacker; detection, to detect the existence of an attack; deflection, to encourage the attacker to direct the attack elsewhere; and diminution, to limit the negative effect of an attack. Such technologies are often collectively referred to as intrusion detection systems.

Intrusion detection can be either anomaly detection, which seeks to identify an attack by behavior that is out of the norm, or misuse detection, to identify an attack by its attempted effect on sensitive resources. Intrusion detection systems monitor a computing system in order to warn of an attack that is imminent, is under way, or has occurred.

Incident response. When an attack has occurred, the affected organization will generally study the attack in order to learn what has occurred, how to prevent a recurrence, and what is the extent of the damage. The science of computer forensics is the study of evidence from an attack. Computer forensics is much like any other investigative science in that it involves assembling potentially significant pieces of evidence and using insight, experience, and intuition to suggest connections among the pieces.

In order to assemble the knowledge about attacks and their countermeasures, organizations and countries staff computer emergency response teams (CERTs). These teams collect information about significant vulnerabilities and attack patterns and provide warnings and possible countermeasures. *See* LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

Charles P. Pfleeger

Bibliography. B. Cheswick and S. Bellovin, *Firewalls and Internet Security*, Addison Wesley, 1994; D. Denning, *Cryptography and Data Security*, Addison Wesley, 1982; D. Denning, *Information Warfare and Security*, Addison Wesley, 1999; S. Garfinkel and G. Spafford, *Practical Unix and Internet Security*, 2d ed., O'Reilly and Associates, 1996; S. Garfinkel and G. Spafford, *Web Security and Commerce*, O'Reilly and Associates, 1997; C. Pfleeger, *Security in Computing*, 2d ed., Prentice Hall, 1997; A. Rubin et al., *Web Security Sourcebook*, John Wiley, 1997.

Computer storage technology

The techniques, equipment, and organization for providing the memory capability required by computers in order to store instructions and data for processing at high electronic speeds. In early computer systems, memory technology was very limited in speed and high in cost. Since the mid-1970s, the advent of high-density, high-speed random-access memory (RAM) chips has reduced the cost of computer main memory by more than two orders of magnitude. Chips are no larger than $^1/_4$ in. (6 mm) square and contain all the essential electronics to store tens to hundreds of thousands of bits of data or instructions. An analogous increase in magnetic recording density has increased the capacity and reduced the cost per bit of secondary memory. Traditionally, computer storage has consisted of a hierarchy of three or more types of memory storage devices (for example, RAM chips, disks, and magnetic tape units). *See* BIT.

Memory Hierarchy

Memory hierarchy refers to the different types of memory devices and equipment configured into an operational computer system to provide the necessary attributes of storage capacity, speed, access time, and cost to make a cost-effective practical system. The fastest-access memory in any hierarchy is

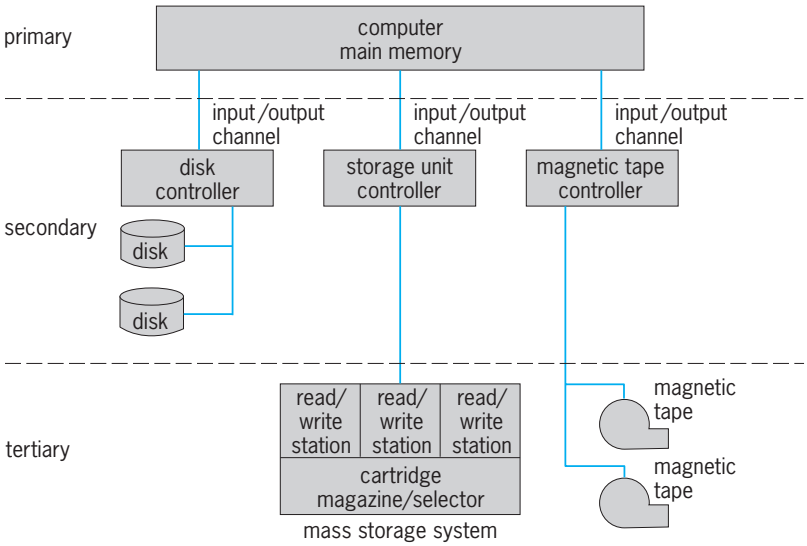

Fig. 1. Memory hierarchy levels and equipment types.

the main memory in the computer. In most computers manufactured after the late 1970s, RAM chips are used because of their high speed and low cost. Magnetic core memories were the predominant main-memory technology in the 1960s and early 1970s prior to the RAM chip. The secondary storage in the hierarchy usually consists of disks. Significant density improvements have been achieved in disk technology, so that disk capacity has doubled every 3 to 4 years. Between main-memory and secondary-memory hierarchy levels, however, there has always been the "memory gap."

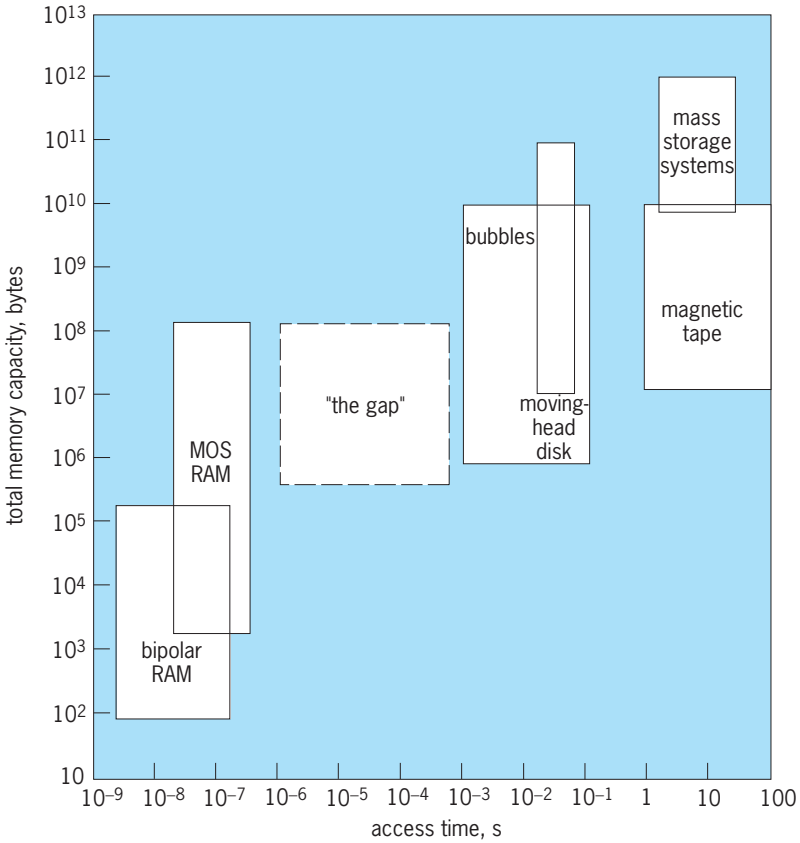

Fig. 2. Memory capacity versus access time for various storage technologies.

The memory gap, noted in the 1940s, still presents a problem for the data-processing system designer. The memory gap for access time is bounded on one side by a 60-nanosecond typical computer main memory cycle time and on the other by the 10-millisecond typical disk-drive access time. Capacity is bounded on the one side by a typical mainframe memory capacity of 256 megabytes and on the other by multiple disks, each with a capacity of from 100 megabytes to over 1 gigabyte.

The last, or bottom, level (sometimes called the tertiary level) of storage hierarchy is made up of magnetic tape transports and mass-storage tape systems (**Fig. 1**). Performance (**Fig. 2**) is usually measured by two parameters: capacity and access time. (Speed or data rate is a third parameter, but it is not so much a function of the device itself as of the overall memory design.) Capacity refers to the maximum on-line user capacity of a single connectable memory unit. Access time is the time required to obtain the first byte of a randomly located set of data.

Memory Organization

The efficient combination of memory devices from the various hierarchy levels must be integrated with the central processor and input/output equipment, making this the real challenge to successful computer design. The resulting system should operate at the speed of the fastest element, provide the bulk of its capacity at the cost of its least expensive element, and provide sufficiently short access time to retain these attributes in its application environment. Another key ingredient of a successful computer system is an operating system (that is, software) that allows the user to execute jobs on the hardware efficiently. Operating systems are available which achieve this objective reasonably well. *See* COMPUTER ARCHITECTURE.

The computer system hardware and the operating system software must work integrally as one resource. In many computer systems, the manufacturer provides a virtual memory system. It gives each programmer automatic access to the total capacity of the memory hierarchy without specifically moving data up and down the hierarchy and to and from the central processing unit (CPU). During the early years of computing, each programmer had to incorporate storage allocation procedures by determining at each moment of time how information would be distributed among the different hierarchic levels of memory, whenever the totality of information was expected to exceed the size of main memory. These procedures involved dividing the program into segments which would overlay one another in main memory. The programmer was intimately familiar with both the details of the machine and the application algorithms of the program. This all changed in the 1970s and improved significantly after 1980, when sophisticated higher-level program languages and database management software became well

established, to provide significantly greater problem-solving capability. Thus manufacturer-supported operating systems evolved, with complete built-in virtual memory support capabilities, which made it possible for the user to ignore the details of memory hierarchy internal software and hardware operations. *See* COMPUTER PROGRAMMING; DATABASE MANAGEMENT SYSTEM; OPERATING SYSTEM; PROGRAMMING LANGUAGES.

In the area of memory organization, two types of memory augmentation have been employed in the more widely used computers to enhance the overall computer performance capabilities. These two memory organization techniques are the cache memory, which speeds up the flow of instructions and data into the central processing unit from main memory, and an intelligent disk controller memory that is used as a staging buffer memory to queue up instructions and data for more rapid access into main memory.

Cache memory. A cache memory is a small, fast buffer located between the processor and the main system memory. Data and instructions in current use are moved into cache, producing two benefits. First, the average access time for the processor's memory requests is reduced, increasing the processor's throughput. Second, the processor's utilization of the available memory bandwidth is thereby reduced, allowing other devices on the system bus to use the memory without interfering with the processor. Cache memory is thus used to speed up the flow of instructions and data into the central processing unit from main memory. This cache function is important because the main memory cycle time is typically slower than the central processing unit clocking rates. To achieve this rapid data transfer, cache memories are usually built from the faster bipolar RAM devices rather than the slower metal oxide semiconductor (MOS) RAM devices.

Performance of mainframe systems is very dependent upon achieving a high percentage of accesses from the cache rather than main memory. In typical systems 80–95% of accesses are to cache memory, with a typical cache size of 256K bytes. (1K bytes is equal to 1024 bytes.) Since the cache contents are a duplicate copy of information in main memory, writes (instructions to enter data) to the cache must eventually be made to the same data in main memory. This is done in two ways: write-through cache, in which write is made to the corresponding data in both cache and main memory; and write-back cache, in which main memory is not updated until the cache page is returned to main memory, at which time main memory is overwritten.

Intelligent disk controller memory. This is used as a cache memory between disk and main memories. Typically it consists of MOS RAM chips which overlap the disk operations, with their longer access time, to mask out the disk access delays so that main memory can execute subsequent tasks more rapidly and efficiently. Intelligent disk controllers provide the latest techniques to fill the memory gap with the best practical memory organization techniques for high performance. Microprocessors are also an integral part of intelligent disk controllers, and carry out many central processing unit operating system functions necessary for disk operations. This off-loads the mainframe computer from doing this kind of overhead processing to a large extent. *See* MICROPROCESSOR.

Main Semiconductor Memory

The rapid growth in high-density very large-scale integrated (VLSI) circuits has advanced to a point where only a few applications require the tens to hundreds of thousands of transistors that can now be placed on a chip. One obvious exception is computer main memory, in which there is a continual demand for higher and higher capacity at lower cost. *See* INTEGRATED CIRCUITS.

In the 1960s and early 1970s, magnetic core memories dominated computer main-memory technology, but these have been completely replaced by semiconductor RAM chip devices of ever-increasing density. This transition started with the introduction of the first MOS 1K-bit RAM memory in 1971. This was followed with the 4K-bit RAM chip in 1973, the 16K-bit chip in 1976, the 64K-bit chip in 1979, the 256K-bit chip in 1982, the 1M-bit chip in 1985, and the 4M-bit chip in the 1990s. (A 1M-bit chip has 1024×1024 or 1,048,576 bits.) The progression of RAM chips (**Fig. 3**) has followed the "rule of four," according to which the cost of development of a new RAM device generation can be justified only by a factor-of-four increase in capacity. *See* SEMICONDUCTOR MEMORIES.

4M-bit RAM chips. The 4M-bit MOS RAM has continued to push photolithographic fabrication techniques with feature sizes of 1 micrometer or less. For RAM chip densities of 1M bits or more per device, the integrated circuit industry has had to make a twofold improvement to maintain volume production and device reliability: better means, such as x-ray step-and-repeat equipment, to achieve features under 1 μm; and plasma-etching or reactive-ion-etching machines to achieve vertical profiles needed as horizontal dimensions decrease. Considerable progress has been made in the use of both x-ray and electron-beam techniques to achieve the submicrometer-size features needed to make even higher-density RAM chips. Most 4M-bit RAMs are dynamic RAMs. The production techniques for dynamic and static RAM chips are identical. Therefore, the cost per unit quickly becomes the cost for mass-producing one chip. Since building costs per chip are about the same whether they store 4K bits or 4M bits, higher densities lead to lower costs per bit.

RAM chip types and technologies. RAM chips come in a wide variety of organizations and types (**Table 1**). Computer main memories are organized into random addressable words in which the word length is fixed to some power-of-2 bits (for example, 4, 8, 16, 32, or 64 bits). But there are exceptions, such as 12-, 18-, 24-, 48-, and 60-bit word-length machines. Usually RAMs contain NK · 1 (for example, 64K · 1) bits, so the main memory design consists of

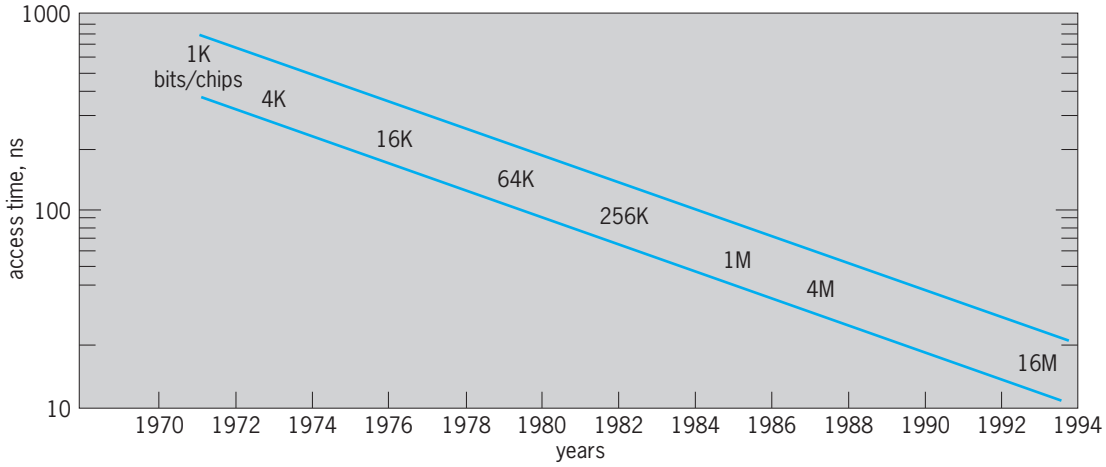

Fig. 3. Trends in performance and density (in bits per chip) of dynamic RAM chips. 1K = 1024 bits; 1M = 1024 × 1024 bits = 1,048,576 bits.

a stack of chips in parallel with the number of chips corresponding to that machine's word length. There are two basic types of RAMs, static and dynamic. The differences are significant. Dynamic RAMs are those which require their contents to be refreshed periodically. They require supplementary circuits on-chip to do the refreshing and to assure that conflicts do not occur between refreshing and normal read-write operations. Even with those extra circuits, dynamic RAMs still require fewer on-chip components per bit than do static RAMs (which do not require refreshing).

Static RAMs are easier to design, and compete well in applications in which less memory is to be provided, since their higher cost then becomes less important. They are often chosen for minicomputer memory, or especially for microcomputers. Because they require more components per chip, making higher bit densities more difficult to achieve, the introduction of static RAMs of any given density occurs behind that of dynamic versions.

There is another trade-off to be made with semiconductor RAMs in addition to the choice between static and dynamic types, namely that between MOS and bipolar chips. Biopolar devices are faster, but have not yet achieved the higher densities (and hence the lower costs) of MOS. Within each basic technology, MOS and bipolar, there are several methods of constructing devices, and these variations achieve a variety of memory speeds and access times, as well as power consumption and price differences. Within the basic MOS technologies there are several types, such as the *n*-channel MOS referred to as NMOS and the complementary MOS solid-state structure referred to as CMOS. For bipolar there are several types such as transistor-to-transistor logic (TTL) and the emitter-coupled logic (ECL). *See* LOGIC CIRCUITS.

Memory cycle and access times. The RAM memory cycle time in a computer is defined as that time interval required to read or write one word of data. Access time is defined as the time interval in which the data are available after the initiation of the read cycle. From the user's point of view, cycle time is also an important characteristic because it has more impact on overall computational speed of the system. A new data read-write cycle does not begin until the previous read cycle is completed. The specified timing, signal overlap, and tolerances allowed vary with each RAM main memory system.

Error checking and correction. Most computer main memories have a memory fault control consisting of a memory controller and a fault control subsystem which automatically corrects all single-bit errors and detects and reports all double-bit or three-bit errors. Each bit in a word is read from a separate RAM

TABLE 1. Summary of characteristics of various memory chip types

Characteristics	Dynamic RAMs	Static RAMs	Static RAMs	ROMs	PROMs	PROMs	EPROMs	EEPROMs (EAROMs)
Chip technology	NMOS CMOS	NMOS CMOS	Bipolar (ECL)	CMOS	Bipolar	CMOS	MOS (UVPROMs)	MOS
Speed (access time)	60–100 ns	50–100 ns	1–5 ns	50–100 ns	5–100 ns	50–200 ns	50–300 ns	50–300 ns
Volatility	Volatile	Volatile	Volatile	Nonvolatile	Nonvolatile	Nonvolatile	Nonvolatile	Nonvolatile
Programmability	Easy	Easy	Easy	During manufacture	Once only	Once only	Many times, long procedure	Many times, easier and faster than EPROMs
Power dissipation	Low	Low	High	Low	High	Low	Medium	Medium
Chip density of currently available devices	4M	1M	4K, 16K, 64K	256K, 1M	16K, 64K	16K, 64K	256K, 1M	64K, 256K, 1M

device. The fault control permits computer operation to continue even if a memory module is malfunctioning, and thus the computer will operate even if a memory chip is removed from the board. Failures are pinpointed to the specific chip by fault-indicating light-emitting diodes built into each array board. *See* LIGHT-EMITTING DIODE.

FIFO and LIFO chips. As memory chips increased in storage density and designers attempted to minimize interconnection circuitry, specialty memory function organizations such as first-in first-out (FIFO) buffers and last-in first-out (LIFO) buffers became available as chips. These kinds of chips are readily used in memory buffering applications to accommodate varying data rates between one element and another. The other complementary feature of FIFO and LIFO chips is the dual data paths (called ports) to allow simultaneous transfer of data into and out of these memory buffer chips. These chip types are commonly used on microcomputer boards and in the computer's peripheral interface and controller electronics.

Gallium arsenide memory chips. All the memory chips discussed above are based on the predominant silicon technology. Reliable gallium arsenide (GaAs) integrated circuits have been produced: 16K-bit and 64K-bit static RAM (SRAM) devices with 1-nanosecond access times. Development of 4M-bit SRAMs has been undertaken for military applications. Considering the need for and emphasis on increasingly rapid computers, there is definite potential for high-speed gallium arsenide logic and memory devices. *See* GALLIUM.

ROMs, PROMs, and EPROMs. Microcomputers have evolved their own special set of semiconductor memory chips (Table 1) to suit their application needs. Whereas large, medium-size, and minicomputers primarily use only RAMs that have read-write capability, microcomputers have found significant use for read-only memory (ROM) chips and programmable read-only memory (PROM) chips. Data and program storage in most microcomputer applications is separately allocated between RAMs and ROMs. ROMs provide protection, since the contents are fixed or hardwired and the chips are completely nonvolatile. During microcomputer program development, PROMs are typically used. A PROM can be written only once and involves the irreversible process of blowing polysilicon fuse links. These PROMs are neither erasable nor reprogrammable. Another kind of device, called an erasable PROM (EPROM), is cleared or rewritten by putting the EPROM chip under an ultraviolet light to zero out its contents and then using a PROM programmer to write in the new bit pattern for the modified program. After the microcomputer application program has completed final tests and acceptance, the final bit pattern in the EPROM can be put into ROMs or PROMs (using a chip mask) to facilitate quantity production. *See* MICROCOMPUTER.

EEPROMs. There are also memory chip devices called electrically erasable programmable read-only memories (EEPROMs). They have an internal switch on the chip to permit a user to electrically rewrite new contents into them. This means EEPROMs do not have to be removed from the circuit in order to clear their contents to put in the new or modified bit pattern representing another program of instructions. Thus EEPROMs have met the need in microcomputer systems for a nonvolatile in-circuit reprogrammable random word access memory device.

EEPROMs use two distinctively different technologies. The more mature technology is metal nitride oxide semiconductor (MNOS), which is a very different gate insulator process technology from that used for MOS. The MNOS technology previously used the generic acronym EAROM (electrically alterable read-only memories) until EEPROMs became common in usage for both technologies. The other technology is floating-gate MOS (also called FAMOS). FAMOS technology was used for ultraviolet erasable EPROMs (UVPROMs) and was subsequently refined to provide an electrically erasable technology. Both technologies rely on Fowler-Nordheim tunneling to move the charge to be stored from the substrate to the gate. FAMOS stores the charge on the gate itself, whereas MNOS devices trap the charge in the nitride layer under the gate.

Secondary Memory

High-capacity, slower-speed memory consists of two major functional types: random-access, which has been provided primarily by disk drives, and sequential-access, which has been provided primarily by tape drives. Since tape drives provide removability of the medium from the computer, tape is used for the majority of off-line, archival storage, although some disks are removable also. The on-line random-access disk devices are classed as secondary, and tape-based systems are classed as tertiary.

Over the history of electronic computers, while the technology for processors and main memory has been evolving from vacuum tubes to transistors to very large-scale integration (VLSI) chips, the predominant technology for secondary memory has continued to be magnetic recording on tape and disk. This has not been due so much to the absence of competing alternatives as to the continuous and rapid progression in magnetic recording capability (**Figs. 4** and **5**).

The current magnetic-disk technology will be discussed, and also the potential of new magnetic recording techniques such as thin-film, vertical recording, and bubble-memory devices; and a nonmagnetic technology, optical recording.

Magnetic disk storage. Conventional magnetic-disk memories consist of units which vary in capacity from the small 360-kilobyte floppy disks (used with microcomputers) to 1-gigabyte disk drives used with large-scale computers. Hard-disk memories are characterized by access times in the 20–80-ms region (versus access times of tens to hundreds of nanoseconds for RAMs). Figure 5 shows the historical progression in several performance factors: capacity, transfer rate, and access time.

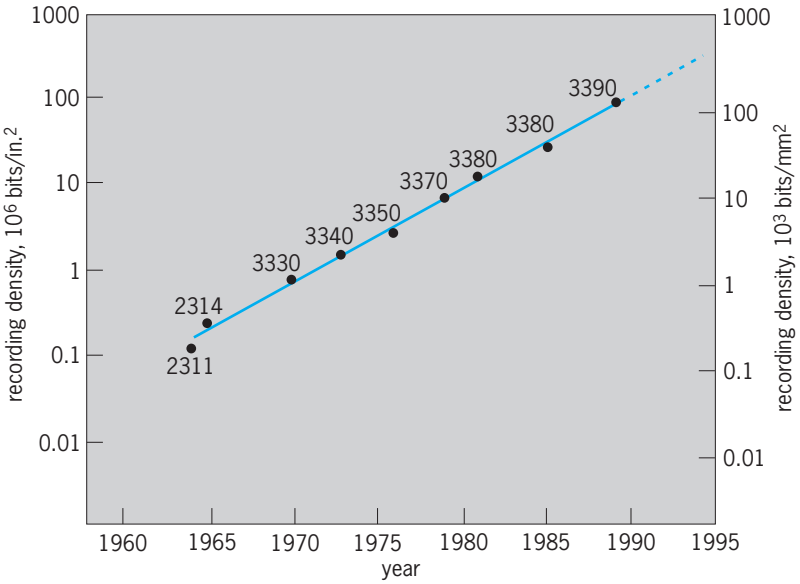

Fig. 4. Increase in recording density on magnetic disk storage. Data points represent IBM models, and model numbers are indicated. Since vertical scale is logarithmic, the straight line represents an exponential increase.

Capacity. The major area of development in disks has been the progressive and even spectacular increases in capacity per drive, particularly in terms of price per byte. There is a substantial economy of scale in storage capacity, in that the cost per byte goes down as the drive capacity goes up. Between small disks and top-end drives, the differences in access time and transfer rate may be only a factor of two to five, while the capacity difference is a factor of as much as 2000, that is, from 20 megabytes to several gigabytes.

Transfer rate. The rotation rate of the disk platter, like the action of the head actuators, is limited by considerations of physical dynamics. Increases in transfer rate will come from greater linear bit density around the track. In an absolute sense, transfer rates are not as often a system performance bottleneck as is access time.

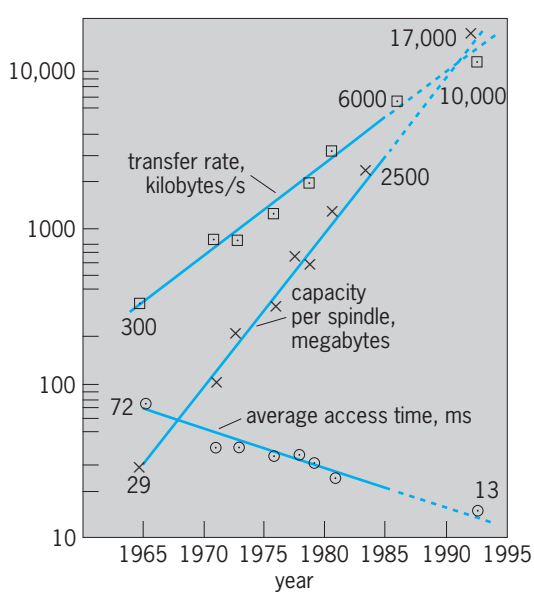

Fig. 5. Development of magnetic disk performance factors.

Average access time. Since the arm supporting the read-write head is one of the few moving mechanical parts in a computer system, it has not shown the multiple-orders-of-magnitude performance improvement over time that the electronic technologies have. Also, average access times for the highest-priced and lowest-prices disks do not vary by an enormous factor, being in the range of 10 to 80 ms for a typical microcomputer hard (as opposed to floppy) disk, and 10 to 30 ms for large disks on mainframe computers. There is nothing which indicates a breakthrough in the technology governing access time, even though greater capacities and transfer rates are still coming. Therefore, for random-access bottlenecks, the solution in the future will be the same as it is now: multiple smaller drives to allow overlapped seeks instead of a single large-capacity drive.

Basic disk technology. Many significant advances in the capabilities of commercially available disk memory technology were made during the late 1970s. One of the major trends was toward larger-capacity disks, progressing from 100 megabytes to several gigabytes per disk drive. The majority of technology improvements to provide higher capacities, lower costs, and better reliabilities resulted from the Winchester technology. Prior to this technology, removable disk packs established themselves as the most advanced technology for large-capacity disks. The Winchester technology is characterized by nonremovable, sealed disk packs, which inherently provide more reliability due to less handling and make more stringent alignment tolerances practical.

The head, or read-write transducer, has undergone substantial refinement. Central to the advance in increasing disk-packing density was the reduction in the flying height of the read-write head. Because flux spreads with distance, reduced separation between the read-write head and the magnetic surface will decrease the area occupied by an information bit. This obviously limits the number of bits that can be defined along an inch of track and increases the minimum spacing between tracks. The ideal would be a direct contact between head and surface, which is the case with magnetic tape. But this is not possible with a rigid aluminum disk whose surface is traveling at rates that can exceed 100 mi/h (45 m/s). By applying a load on the head assembly, proper spacing is maintained. All movable media memories utilize an air-bearing mechanism known as the slider or flying head to space the transducer in proximity to the relatively moving media. The slider, when forced toward the moving surface of the medium, rests on a wedge of air formed between it and the medium by virtue of their relative motion. The thickness of the wedge can be controlled by the applied force, and the wedge in effect is an extremely stiff pneumatic spring. Head-to-medium spacings were well under 10 μin. (0.25 μm) in 1982. *See* TRANSDUCER.

Some standards have been developed for interfaces between the drive and the controller, to permit

compatibility among drives of various capacities and manufacturers. Among these are the storage module device (SMD) interface, enhanced small device interface (ESDI), small computer system interface (SCSI), intelligent peripheral interface (IPI), and high performance parallel interface (HPPI).

SMD drives can run at 3 megabytes per second. ESDI drives can also attain that rate. The IPI Level II drives can operate at 10 megabytes per second. The transfer rates of SCSI vary. As an intelligent interface between peripherals and computer systems, SCSI allows system designers to avoid device-level idiosyncrasies that make quick integration of new peripherals difficult. It is suitable for use on small personal computers with standard asynchronous transfer rates of 1.5 megabytes per second. It can also be used with supermicrocomputers and minicomputers with an optional synchronous transfer rate of 4 megabytes per second. Work has been done toward raising the transfer rate to 8 megabytes per second by tightening the timing specifications of the SCSI interface. Although initial versions of SCSI were designed for small computers, SCSI can meet the high-performance and high-capacity requirements of larger computers. For example, a single SCSI bus can support six SCSI-to-ESDI controllers with 24 disks, providing 7.2 gigabytes of storage, in addition to tape drives and other devices.

The HPPI was intended for interprocessor memory communication and as a data path for high-speed display monitors. It is a half-duplex point-to-point channel and can transmit data at 800 or 1600 megabytes per second. By using HPPI, it is possible to set up a high-performance high-capacity storage system of 256 gigabytes (or more) for supercomputer applications.

In addition to fixed Winchester-type disk units of various capacities, disk storage is used in several removable forms.

Cartridge disks. Removable storage cartridges containing single-disk packs are used extensively with minicomputer systems. The cartridge disk-drive unit has capacities ranging from 1 to 10 megabytes, depending on the recording density. Average access times range from 35 to 75 ms, and data transfer rates from 200 to 300 kilobytes per second. Physically, these plastic-encased cartridges are either top-loaded or front-loaded units. These units, besides having a removable disk, also have a fixed disk. The performance features of cartridge disks compete favorably with the larger-capacity disk pack storage equipment. A disadvantage of removable disk cartridges or disk packs is reduced reliability compared to fixed units.

Floppy disks. Floppy disks have become the widely used random-access secondary memory for microcomputer systems. The floppy disk was originally developed by IBM in 1965 for internal use on its large 370 computers. The disk was designed to be a permanent nonvolatile memory storage device, written at the factory. However, large numbers of new and varied applications evolved for these relatively small, compact disk systems. Floppy disks derive their name from the recording medium, an oxide-coated flexible disk (similar to, but more flexible than, plastic 45-rpm music records) enclosed within a protective plastic envelope. The Mylar flexible disk is $7^3/_4$ in. (195 mm) in diameter, 0.003 in. (0.075 mm) thick, and records 3200 bits per inch (125 bits per millimeter) and 48 tracks per inch (1.9 tracks per millimeter). The disk is permanently enclosed in a protective envelope. The common reference to the size of this disk is 8-in. (200-mm). The protective envelope contains alignment holes for spindle loading, head contact, sector-index detection, and write-protect detection. The capacity of this easily transportable medium is 8 megabits, commonly laid out on 77 tracks, each divided into 26 sectors.

The flexible-disk drive subassembly consists of the metal base frame, the disk-drive motor that rotates the spindle (360 rpm for 8-in. or 200-mm, 300 rpm for $5^1/_4$-in. or 133-mm) through a belt drive, and a stepping motor with an integral lead screw to form the head-positioning actuator. The read-write head is mounted on a carriage that the lead screw moves. Head load is achieved by moving a load pad against the disk and restraining the disk between the head and the load pad to accomplish the head-to-disk compliance.

The removable floppy disk, called a diskette, was originally 8 in. (200 mm), in both single and double densities, and single- and double-sided. A variety of incompatible formats exist, particularly in double density, as the majority of suppliers use a common (IBM) format in single density. The most conspicuous format difference is in sector boundaries, hard-sector referring to boundaries of physical holes in the diskette, and soft-sector to boundaries that are software-formatted. As recording density increased and packaging considerations for microcomputers became important, a smaller, $5^1/_4$-in. (133-mm) diskette was introduced in 1976 with a similar proliferation of formats as in 8-in. (200-mm). Many systems use yet smaller sizes such as 3, $3^1/_4$, and $3^1/_3$ in. (75, 83, and 90 mm). Standard storage capacities range from 720 kilobytes to 1.44 megabytes. The dominant format is the $3^1/_2$-in. (90-mm) high-density (1.44-megabyte) floppy diskette. Floppy diskettes (and drives) with density higher than 1.44 megabytes are available. However, floppy diskettes are often replaced with CD-ROM disks.

Thin-film and vertical recording. Several technologies are in development for disk storage.

The recording technology for conventional magnetic disk recording has been based upon ferrite heads and a recording surface consisting of a coating embedded with ferrous oxide particles. One of the limiting factors is the thickness of the recording layer, which is 20–50 μin. (0.5–1.25 μm). With current products pushing the limits of recording density achievable with that approach, a successor technology has been developed, called thin-film.

The thin-film technique uses a continuous magnetic film of cobalt or nickel, applied by electroplating to a controlled thickness of as little as 1 μin.

(25 nanometers). Either the head or the recording surface or both can use thin-film technology. Most of the major conventional large disks use thin-film heads and an oxide recording surface. Advanced models are designed to use a thin-film recording surface. Thin-film technology is the source for current progress in disk capacity and transfer rate. It was first introduced in small disks, partly due to greater ability to withstand the more portable environment of the small computer. The hardness of the medium which provides this resiliency also reduces both the likelihood and consequences of head crashes.

Another magnetic recording technique is vertical, or perpendicular, recording. In this approach the miniature magnets which are formed in the recording media are oriented downward into the surface rather than longitudinally along the surface. This allows closer packing of the magnetized spots, not just because of spatial compression, but also because of improved magnetic interaction between adjacent spots. The adjacent spots repel each other in the longitudinal case but attract in the vertical case. This technique offers the potential for significantly extending recording density and the accompanying performance parameters.

Recording density is the product of the bit density along the recording track, in bits per inch (or bits per millimeter), and the number of tracks per inch (or tracks per millimeter) across the recording surface (**Table 2**). For a given rotation rate of the disk, the bit density along the track determines the transfer rate.

The point at which vertical recording becomes attractive to develop into commercial products depends in part upon the level of capability which can ultimately be achieved with thin-film technology. The density level achieved by vertical recording technology may represent a limit for magnetic recording. If so, one possible answer to the question of how further levels of performance can be obtained is optical recording.

Optical recording. Optical recording is a nonmagnetic disk technology that uses a laser beam to burn pits in the recording medium to represent the bits of information, and a lower-power laser to sense the presence or absence of pits for reading. This technology has the potential for higher ultimate recording density than magnetic recording. The medium is removable and relatively inexpensive. The removability is significant in light of the nonremovability of most current and projected high-capacity magnetic disks. The recording density increase is apparent when it is noted that 4-gigabyte capacity is achieved on a single disk, whereas current-model large-capacity (2.5-gigabyte) magnetic drives typically consist of a cabinet containing two disk assemblies, each assembly having 16 platters.

One of the first optical disk mass-storage devices has 4 gigabytes on a removable, nonerasable, 14-in. (350-mm) disk. The transfer rate of 1.5 megabytes per second buffered or 3.0 megabytes streaming is similar to that of conventional magnetic disks (3.0 for the IBM 3380). The average access time is slower than that of high-performance magnetic disks at 84.7 ms, versus the 3380's 24 ms, but track-to-track access is a comparatively fast 1.0 millisecond. The corrected bit error rate is 1 in 10^{13}, which is compatible with magnetic disk standards.

In addition to the 14-in. (356-mm) disk for large-capacity systems, 12-in. (300-mm) optical disk systems in the 1-gigabyte capacity range have been produced for use with 16- or 32-bit microcomputer and minicomputer systems.

Disadvantages. Optical technology has several disadvantages. First, the medium is not inherently erasable, although systems which can erase and rewrite are also in development. The primary problem in producing a commercial optical disk system has been getting the bit error rate down to an acceptable level, with sufficient recording density to be competitive with magnetic recording. To be competitive with magnetic tape, a bit error rate, after error detect and correct, of 1 in 10^{10} to 10^{11} is needed, and to compete with magnetic disk, 1 in 10^{13} to 10^{14}. Through error detect and correct techniques, the 1 in 10^{13} level has been attained.

The bit error rate problem is far less critical for storage of documents and images, since bit dropout is seldom noticeable, in contrast with financial information, where it could be disastrous. This is the reason that optical disk recording was first used successfully in consumer and industrial systems for videotape image information, and for document storage and retrieval systems. *See* VIDEO DISK.

Applications. The nonerasability of the technology can be an advantage for some applications. Examples are data which must be kept for historical or audit-trail purposes, and applications where large amounts of data must be shipped among different locations.

When this technology matures, it will be an event of major significance, being a removable, large-capacity, random-access device using a recording technology with large potential for further growth. The potential for development of yet higher

TABLE 2. Storage densities for various recording technologies

	Bit density per track		Track density	
Technology	Bits/in.	Bits/mm	Tracks/in.	Tracks/mm
(Magneto) optical	20,500	800	15,000	600
Magnetic (floppy)	17,500	700	135	5
Magnetic (hard disk)	24,000	960	1,000	40
Floptical (optical floppy)	24,000	960	1,250	50

recording densities is perhaps the most significant factor, but in addition, having large capacity on a removable platter would allow some other kinds of system development. One possible system, referred to as a jukebox, would store on the order of 100 disks in an automatic retrieval and mounting device, so that a given disk could be fetched and accessed in around 5 seconds giving an on-line storage in the range of 500 gigabytes. Previous systems for very large on-line storage which have been delivered in significant quantity are the cartridge tape systems from IBM and CDC and automated tape libraries discussed below. The cartridge tape systems have capacities to 50 megabytes (IBM) and 8 megabytes (CDC) per cartridge; hence the optical system would provide a great improvement in the capacity available on the basic storage unit, and the automated tape libraries of course have the limitation of sequential data access.

CD-ROM and WORM. CD-ROM (compact disk–read-only memory) and WORM (write once, read many) are special types of optical disks. CD-ROM resembles the related audio compact disk technology in that users of CD-ROM can read only prerecorded data on the disk. A 5-in. (125-mm) CD-ROM can hold 500 to 600 megabytes, which is equivalent to 1400 to 1700 (360-kilobyte) diskettes. For example, the text of a 20-volume encyclopedia could be easily stored on a CD-ROM disk. CD-ROM has become an add-on peripheral device for some personal computers and workstations.

The CD-ROM disk and its drive are physically similar to the audio compact disk and its player, but there are several key differences: (1) A CD-ROM drive does not have a digital-to-analog converter. (2) It does have more powerful error correction than the audio system, because computer data requires a lower error rate, typically one error every 10^{15} bits, while the ratio of one error in 10^8 is acceptable in an audio compact-disk system. (3) The CD-ROM drives employ data scrambling to protect the servo control circuits from harmful dc components.

In CD-ROM, data are stored digitally on the surface of the substrate as a series of pits and lands of variable lengths. A land is the surface of the substrate between pits. Instead of using pits and lands to represent ones and zeros, the data in CD-ROM are represented in terms of the transitions from a pit to a land and vice versa (ones) and by the lack of such transitions (zeros). This arrangement, together with a modulation scheme, enhances the resolution with which data are read off the disk from about 1 μm—the diameter of the laser beam—to about 0.6 μm. The result is a 60% increase in the density of the data on the disk.

The pits on the disk surface can be arranged in either concentric circles, like tracks on magnetic disks, or in a spiral track, as on vinyl phonograph records. One common data format uses concentric tracks with the disk rotating at a constant angular speed. This simplifies the servomechanism but wastes space on the outer part of the disk because the bits in the outer tracks are separated from each other by more space than that necessary for adequate resolution. With the spiral track, the disk can be rotated so that the track passes the read head at a constant linear speed. This means that the disk must rotate more slowly when the laser beam is focused on the outer tracks and more quickly when the beam is on the inner tracks. Whether the tracks are concentric or spiral, the intertrack distance is about 1.6 μm, and the width of the pits is about 0.6 μm.

The average access time for CD-ROM drives ranges from 320 ms (for double-speed drives) to 160 ms (for quadruple-speed drives). The average data transfer rates vary from 300 kilobytes per second for double-speed drives to 450 kilobytes per second for triple-speed drives and 600 kilobytes per second for quadruple-speed drives.

While CD-ROM technology is fairly well established, the lack of established standards for WORM technology could delay its acceptance. WORM drives come in different sizes: $3^1/_2$, $5^1/_4$, 8, 12, and 14 (90, 133, 200, 300, and 350 mm). Capacities of $5^1/_4$-in. (133-mm) WORM drives range from 200 megabytes to over 1 gigabyte, while 12-in. (300-mm) WORM drives are in the 2–10-gigabyte range. *See* COMPACT DISK.

Bubble memory devices. Bubble memories are chips rather than disks, but are different from semiconductor memories in that they are magnetic devices, in which the absence or presence of a magnetic domain is the basis for a binary 1 to 0. The performance characteristics of these devices makes them competitive as small-capacity secondary storage. A magnetic bubble is in reality a cylindrical magnetic domain with a polarization opposite to that of the magnetic film in which it is contained. These cylinders appear under a microscope as small circles (which give them their name) with diameters from 80 to 800 μin. (2 to 20 μm). The size of a bubble is determined by the material characteristics and thickness of the magnetic film.

Bubbles are moved or circulated by establishing a magnetic field through a separate conductor mounted adjacent to the bubble chip. A large portion of the bubble chip must be given over to circuitry for generating, detecting, and annihilating the bubbles. Magnetic-domain bubble devices typically operate in an endless serial loop fashion or in an organization with serial minor loops feeding major loop registers. Bubble memories are particularly well suited to applications such as portable recorders because of their physical advantages (low power requirements and light weight) and speed advantage over electromechanical devices such as cassettes and floppy disks. Like their electromechanical counterparts, bubble devices are nonvolatile; that is, they retain their contents when the power goes off.

The controller is central to any bubble memory system, which it serves as the interface between the bubble chip and the system bus. It generates all of the system timing and control functions and supervises the handshaking operations required to access and transfer data between the system bus and the bubble memory module. Usually a controller

can operate more than one bubble chip module, in most cases up to eight modules. The function-driver integrated circuit produces the control currents required to generate and input or ouput tracks on the chip. The sense-amplifier integrated circuit converts the analog out-put signals produced by the bubble detector into a transmittable data stream that passes to the controller. The coil-driver integrated circuits excite the x and y coils of the bubble chip package with out-of-phase signals to produce the rotating field that moves the bubbles.

Several companies are producing 1-megabit and 4-megabit bubble memory chips, and 16-megabit bubble memory chips are in the development stages. Under a cross-license and alternate source agreement, these companies have agreed to jointly adopt a low-height leaded package for complete component level interchangeability. Each company developed its peripheral chip set for the 4-megabit magnetic bubble module that will be compatible with the 1-megabit units.

For portable and other special applications, bubbles have definite advantages such as nonvolatility, low power, and high compactness. Performance capabilities relative to floppy disks are 100 kilobits per second for bubbles versus 200–250 kilobits per second for floppies, and 40 milliseconds average access time for bubbles versus 200–250 milliseconds for floppies.

Magnetic Tape Units

In magnetic tape units, the tape maintains physical contact with the fixed head while in motion, allowing high-density recording. The long access times to find user data on the tape are strictly due to the fact that all intervening data have to be searched until the desired data are found. This is not true of rotating disk memories or RAM word-addressable main memories. The primary use of tape storage is for seldom-used data files and as back-up storage for disk data files.

Half-inch tapes. Half-inch (12.5-mm) tape has been the industry standard since it was first used commercially in 1953. Half-inch magnetic tape drive transports are reel-to-reel recorders with extremely high tape speeds (up to 200 in. or 5 m per second), and fast start, stop (on the order of 1 millisecond), reverse, and rewind times.

Performance and data capacity of magnetic tape have improved by orders of magnitude (**Table 3**). Just prior to the 1970s came the single-capstan tape drive, which improved access times to a few milliseconds. These vacuum-column tape drives have such features as automatic hub engagement and disengagement, cartridge loading, and automatic thread operation. There are two primary recording techniques, namely, nonreturn-to-zero-inverted (NRZI) and phase-encoded (PE), used with packing densities of 800 and 1600 bytes per inch (31 and 63 bytes per millimeter). These typically use a 0.6-in. (15-mm) interrecord gap. In phase encoding, a logical "one" is defined as the transition from one magnetic polarity to another positioned at the center of the bit cell. "Zero" is defined as the transition in the opposite direction also at the center of the bit cell, whereas NRZI would involve only one polarity. The advantage of the phase-encoding scheme over NRZI is that there is always one or more transitions per bit cell, which gives phase encoding a self-clocking capability, alleviating the need for an external clock. The disadvantage of phase encoding over NRZI is that at certain bit patterns the system must be able to record at twice the transition density of NRZI.

Computer-compatible magnetic tape units are available with tape drives of 6250 bytes per inch (246 bytes per millimeter). They have nine tracks, where eight bits are data and one bit is parity. Each track is recorded by a technique called group-coded record (GCR), which uses a 0.3-in. (7.5-mm) record gap. Every eighth byte in conjunction with the parity track is used for detection and correction of all single-bit errors. Thus, GCR offers inherently greater reliability because of its coding scheme and error-correction capability. It does, however, involve much more complex circuitry.

The on-line capacity is strictly a function of the number of tape drives that one controller can handle (eight is typical). A 2400-ft (730-m) nine-track magnetic tape at 6250 bytes per inch (246 bytes per millimeter) provides a capacity of approximately 10^9 bits, depending on record and block sizes.

Quarter-inch tapes. The installation of $^1/_4$-in. (6-mm) streaming tape drives has increased very rapidly. There are several recording-format standards for $^1/_4$-in. (6-mm) tape drives; the most common one is called QIC (quarter-inch compatibility)-24, which connotes a full-size cartridge class with formatted storage capacity of 60 megabytes on nine tracks, using 600 ft (183 m) of tape. The device interface for a QIC-24 drive is either an intelligent interface

TABLE 3. Standard recording densities for magnetic tape

		Bit density along track		Area density	
Year	Number of tracks	bits/in.	bits/mm	bits/in.^{2}	bits/mm^2
1953	7	100	4	1,400	2
1955	7	200	8	2,800	4
1959	7	556	22	7,784	12
1962	7	800	32	11,200	17
1963	9	800	32	14,400	22
1965	9	1600	63	28,800	45
1973	9	6250	246	112,500	174
1990	9	43,200	1,864	35,280,000	50,400

such as QIC-02; the SCSI; or a basic drive interface, QIC-36. Another recording format, QIC-120, provides 125 megabytes on 15 tracks. The recording format QIC-1350 provides 1.35 gigabytes.

Cassettes and cartridges. The most frequently used magnetic tape memory devices for microcomputer systems are cassette and cartridge tape units. Both provide very low-cost storage, although their access times are long and their overall throughput performance is not as great as that of floppy disks. Cassette and cartridge units both use $^1/_4$-in. (6-mm) magnetic tape.

The digital cassette transport was orginally an outgrowth of an audio cassette unit. Unfortunately, the very low-cost audio-designed transport did not meet the endurance needs of true digital computer applications. There are two basic design approaches to digital cassette transports: capstan drive and reel-to-reel drive. Capstan tape drives are better for maintaining long-term and short-term tape-speed accuracy, while reel-to-reel transports have better mechanical reliability.

During the 1960s the first true digital tape cartridge was developed. With the cassette, a capstan must penetrate the plastic case containing the tape in order to make contact. There is no such penetration system with the cartridge because the transport capstan simply drives an elastomer belt by pressing against the rim of the belt's drive wheel. This simplicity eliminates a major source of tape damage and oxide flaking.

Cassettes have undergone evolution as a digital medium, and capacity has increased to 1 gigabyte, at a density of 61,000 bits per inch (2400 bits per millimeter). Transfer rates are 30 kilobytes per second at 30 in. (0.76 m) per second, or 100 kilobytes per second at 90 in. (2.29 m) per second. Some cassettes can transfer data at 400 kilobytes per second. The $^1/_4$-in. (6-mm) streaming tape cartridge has a typical capacity range of 100 to 300 megabytes, with some available with over 1 gigabyte. Transfer rates are 30 to 90 kilobytes per second, at densities of from 6400 to 10,000 bits per inch (252 to 394 bits per millimeter).

A typical mainframe tape cartridge can store data at 38,000 bits per inch (1496 bits per millimeter). At typical blocking factors, a 550-ft (168-m) tape cartridge contains about the same amount of data as a 2400-ft (730-m) reel of tape at 6250 bits per inch (246 bits per millimeter), but it requires less storage space. The data transfer rate between the tape cartridge drive and the process is 4.5 megabytes per second.

Half-inch (12.5-mm) tape is also used for back-up for microcomputer hard disk systems. Configured to standard microcomputer packaging sizes, these give 50-megabyte capacity in $5^1/_4$-in. (133-mm) form, and 300 megabytes in 8-in. (200-mm) form.

Mass Storage Tape Systems

With the gradual acceptance of virtual memory and sophisticated operating systems, a significant operational problem arose with computer systems, particularly the large-scale installations. The expense and attendant delays and errors of humans storing, mounting, and demounting tape reels at the command of the operating system began to become a problem. Cartridge storage facilities are designed to alleviate this problem.

Their common attributes are: capacity large enough to accommodate a very large database online; access times between those of movable-head disks and tapes; and operability, without human intervention, under the strict control of the operating system. The cartridge storage facility is included within the virtual address range. All such configurations mechanically extract from a bin, mount on some sort of tape transport, and replace in a bin, following reading or writing, a reel or cartridge of magnetic tape.

Cartridge storage systems are hardware devices that need operating system and database software in order to produce a truly integrated, practical hardware-software system. Users require fast access to their files, and thus there is a definite need to queue up (stage) files from the cartridge storage device onto the disks. The database software must function efficiently to make this happen. In general, users base their storage device selection on the file sizes involved and the number of accesses per month. Magnetic tape units are used for very large files accessed seldom or infrequently. Mass-storage devices are for intermediate file sizes and access frequencies. Disk units are used for small files or those which are accessed often.

In practice, most users have been satisfied with tapes and disks, and have not chosen to install mass storage systems. During the 1970s, two basic kinds were delivered, although some others were built and installed in the earlier years. In different ways, these units combine the low cost of tape as a storage medium with the operating advantages of on-line access. One is a mechanical selection and mounting unit to load and unload tape reels onto and off standard magnetic tape units; another is a mechanical selection and accessing unit to operate with special honeycomb short-tape units. The first type (sometimes called automated tape libraries) eliminated manual tape-mounting operations. The main objective of the short-tape honeycomb cartridge system is to improve access time. The shorter tape (770 in. or 20 m, versus 2400 ft or 730 m) results in better access time. Operationally, the honeycomb cartridge tape system and the mechanical standard tape-mounting units are capable of handling 100 to 300 cartridge or tape loads per hour.

The first type of fully automated tape library uses standard magnetic tape $^1/_2$ in. or 12.5 mm wide). Under computer control, this equipment automatically brings the tapes from storage, mounts them on tape drives, dismounts the tapes when the job is completed, and returns them to storage. Accessing up to 150 reels per hour, this unit can store up to 7000 standard tapes or 8000 thin-line reels in a lockable self-contained library that can service up to 32 tape drives and that can interface with up to four computers.

The honeycomb cartridge storage system uses a storage component called a data cartridge. Housed in honeycomb storage compartments, these 2 × 4 in. (24 × 50 mm) plastic cartridges can each hold up to 50,000,000 bytes of information on 770 in. (20 m) of magnetic tape approximately 3 in. (75 mm) wide. Whenever information from a cartridge is needed by the computer, a mechanism selects the desired cartridge and transports it to one of up to eight reading stations. There the data are read out and transferred to the staging disk drives. *See* COMPUTER; DIGITAL COMPUTER.

Peter P. Chen; Michael Plesset; Douglas Theis

Bibliography. L. Kempster, *Removable Electronic Storage Media*, 1994; S. Ranade, *Mass Storage Systems,* 1994; C. Sherman, *CD-ROM Handbook*, 2d ed., 1994; H. M. Sierra, *An Introduction to Direct Access Storage Devices*, 1990; E. W. Williams, *The CD-ROM and Optical Disc Recording Systems*, 1996.

Computer-system evaluation

The evaluation of performance, from the perspectives of both developers and users, of complex systems of hardware and software. Modern computer-based information systems have become increasingly complex because of networking, distributed computing, distributed and heterogeneous databases, and the need to store large quantities of data. People are relying increasingly on computer systems to support daily activities. When these systems fail, significant breakdowns may ensue. *See* DISTRIBUTED SYSTEMS (COMPUTERS); LOCAL-AREA NETWORKS.

A computer system can fail in two major ways. First, functional failure occurs when the system fails to generate the correct results for a set of inputs. For example, if an information system fails to retrieve records that match a set of keywords, or if an air-missile tracking system fails to distinguish between a friendly and enemy missile, a functional failure has occurred. Second, performance failure occurs when the system operates correctly but fails to deliver the results in a timely fashion. For example, if an information system takes a longer time than users are willing to wait for the records they requested, the system is said to fail performancewise even though it may eventually retrieve the correct set of records. Also, if the air-missile tracking system fails to detect an enemy missile in sufficient time to launch a counterattack, the system manifests performance failure.

Therefore, in designing a computer system it is necessary to guarantee that the end product will display neither functional nor performance failure. It is then necessary to predict the performance of computer systems when they are under design and development, as well as to predict the impact of changes in configurations of existing systems. This requires the use of predictive performance models.

Performance models. The input parameters of performance models include workload intensity parameters, hardware and system parameters, and resource demand parameters. The outputs generated by performance models include response times, throughputs, utilization of devices, and queue lengths. There are analytic, simulation, and hybrid performance models (see **illus.**). Analytic models are composed of a set of equations, or computational algorithms, used to compute the outputs from the input parameters. Simulation models are based on computer programs that emulate the behavior of a system by generating arrivals of so-called customers through a probabilistic process and by simulating their flow through the system. As these simulated entities visit the various system elements, they accumulate individual and system statistics. Hybrid models combine both analytic and simulation approaches by, for example, replacing an entire subsystem in an analytic model by an equivalent device whose input–output behavior is obtained by simulating the subsystem. Analytic models can be exact or approximate. *See* SIMULATION.

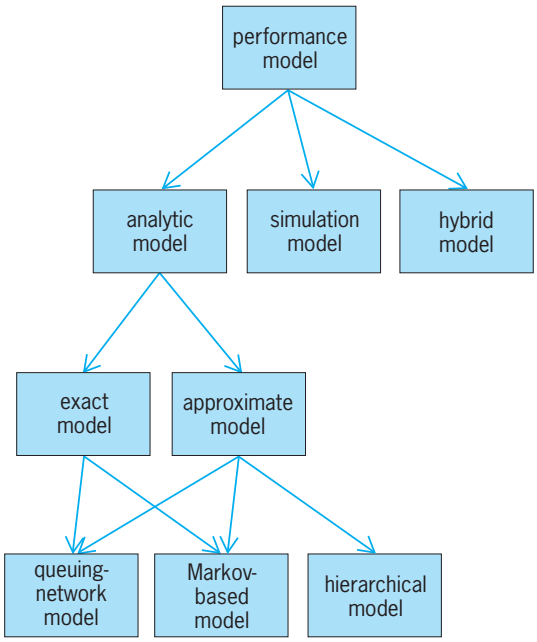

Taxonomy of performance models.

Approximations are needed either when there is no known mathematically tractable exact solution or when the computation of an exact solution is very complex. Modern computer systems are very complex because of ubiquitous networking, distributed processing using client–server architectures, multiprocessing, and sophisticated input–output subsystems using network-attached storage devices. For this reason, most computer system performance models are approximate models. *See* CLIENT-SERVER SYSTEM; MULTIPROCESSING.

Exact performance models. Exact models may be further classified into queuing-network and Markov-chain-based models. Queuing-network models represent a computer system by a network of devices (also called queues) through which customers (for example, transactions, processes, and server requests)

flow. Queues may form at each device because of the finite service rate of the device. Solving a queuing network amounts to finding the probability that the network is found in each of the possible states, or at least finding the performance metrics (for example, response times, throughputs, and queue lengths) associated with the queuing network. The solution to a class of queuing networks, called product-form queuing networks, can be obtained through very efficient computational algorithms such as convolution and mean-value analysis.

In Markov-based models, the system is modeled by a Markov chain, which represents the set of all possible states of the system with the possible transitions between these states. Each state gives rise to a linear flow-equilibrium equation. Solving the Markov chain amounts to solving the resulting system of linear equations, called the global balance equations. The unknowns in the global equations are the probabilities of finding the system at each of the possible states. In some cases, the Markov chain is not explicitly generated from an analysis of the system, but is derived from a higher-level model used to represent the dynamic and temporal behavior of the system. Examples of these higher-level models include generalized stochastic Petri nets (GSPNs), which are timed stochastic extensions to Petri nets. Algorithms exist that generate the Markov chain automatically from the generalized stochastic Petri net. *See* PROBABILITY; STOCHASTIC PROCESS.

Approximate performance models. Product-form queuing networks (PFQNs) cannot model many features of modern computer systems because the conditions for a queuing network to have a product-form (efficient) solution are violated in most cases of interest. For instance, priority-based scheduling disciplines for the devices in the computer system, such as those implemented by many operating systems for scheduling of central processing units, cannot be modeled directly by product-form queuing networks.

Product-form conditions are also violated in a queuing network that exhibits simultaneous resource possession, that is, a situation where a customer is allowed to hold more than one resource at the same time. This occurs when a file is being transferred from a network-attached tape drive to a disk through a HIPPI (High Performance Parallel Interface) switch. In this case, the tape drive is held by the customer simultaneously with the HIPPI switch or the disk during the transfer. Another example of simultaneous resource possession occurs in shared-memory multiprocessors. In this case, a request to access shared memory needs to hold the processor, the bus, and the shared memory bank during the memory access at the same time.

Queuing networks also do not satisfy product-form conditions when fork-and-join synchronization is present. This situation is characterized by a customer being split into more than one customer at the fork point. These customers follow independent paths within the system and are reassembled into a single customer at the join point when all of them arrive. An example of fork-and-join synchronization occurs in the modeling of a RAID-5 disk (a Redundant Array of Inexpensive Disks, with five disks in the array). Here, a read or write request to a RAID-5 disk is split into five requests to the physical disks that compose the RAID-5 disk. *See* COMPUTER STORAGE TECHNOLOGY.

Another case of violation of the product-form conditions occurs when the routing of customers in a queuing network is a function of the number of customers at other parts of the network. This may occur in the modeling of computer network routing algorithms which decide on the next link to which a packet should be sent based on the level of congestion at the set of possible links in the path to the destination.

Finally, another case where product-form conditions do not hold is when there is a constraint on the maximum number of customers in a part of the queuing network. An example occurs in the modeling of multiprogrammed computer systems where there is a maximum degree of multiprogramming.

Many approximate models have been developed to deal with queuing networks with non-product-form solutions. Some of these approximations are based on modifying the equations of the mean value analysis algorithm. Other approximations are based on formulating the problem as a fixed-point equation and solving it iteratively.

Even queuing networks with product-form solutions may be more efficiently solved with the use of approximate algorithms. The reason is that the computational complexity of solving a queuing network is proportional to the product of the number of devices in the computer system times the number of different classes of customers multiplied by the number of customers in each class. This may render models of moderately large computer systems unmanageable. Several approximations have been developed that drastically cut down the number of operations and memory requirements of the algorithms to solve product-form queuing networks. The most widely used approximation technique in this category is the Schweitzer approximation that can be used to derive an efficient iterative algorithm to solve product-form queuing networks.

Approximation techniques are also used when solving large Markov chains with an extremely large number of states. These techniques are aimed at cutting down the number of states without significant loss of accuracy. Finally, an important category of approximate models comprises hierarchical models. In this case, the complete model is decomposed into a hierarchy of models, $M_1, \ldots, M_n$. The output of the model at level i is used as input to the model at level $i + 1$. In some cases, a dependency between models M_n and M_1 is established requiring an iterative fixed-point procedure to solve the series of models. The modeling techniques used at each level may vary. It is possible to use queuing networks at one level and Markov-based models at another level of the hierarchy. Hierarchical models are very useful when dealing with client–server systems.

Performance of software systems. The design and development of complex software systems is a time-consuming and expensive task. Performance modeling techniques must be integrated into the software development methodology. This integrated approach is called software performance engineering. One goal is to estimate the resource consumption of software under development so that performance models can be used to influence the architecture of the software under development. Better estimates on the resource consumption are obtained as the software development process evolves. *See* INFORMATION SYSTEMS ENGINEERING; SOFTWARE ENGINEERING.

Daniel A. Menascé

Bibliography. G. Bolch et al., *Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications*, 2d ed., 2006; P. Fortier and H. Michel, *Computer Systems Performance Evaluation and Prediction*, 2002; D. J. Lilja, *Measuring Computer Performance: A Practitioner's Guide*, 2000; D. A. Menascé et al., *Performance by Design: Computer Capacity Planning by Example*, 2004; C. U. Smith and L. G. Williams, *Performance Solutions: A Practical Guide To Creating Responsive, Scalable Software*, 2001.

Computer vision

The technology concerned with computational understanding and use of the information present in visual images. In part, computer vision is analogous to the transformation of visual sensation into visual perception in biological vision. For this reason the motivation, objectives, formulation, and methodology of computer vision frequently intersect with knowledge about their counterparts in biological vision. However, the goal of computer vision is primarily to enable engineering systems to model and manipulate the environment by using visual sensing. *See* VISION.

Sensing and image formation. Computer vision begins with the acquisition of images. A camera produces a grid of samples of the light received from different directions in the scene. The position within the grid where a scene point is imaged is determined by the perspective transformation. The amount of light recorded by the sensor from a certain scene point depends upon the type of lighting, the reflection characteristics and orientation of the surface being imaged, and the location and spectral sensitivity of the sensor.

Image segmentation. The objective of this early stage is to compress the huge amount of image detail by identifying and representing those aspects of image structure which are salient for later stages of interpretation. Typically, this is accomplished by detecting homogeneous regions in the image or, equivalently, their edges. It must be done for all degrees of regional homogeneity and sizes, yielding a multiscale segmentation. Further, small regions may group together to form a larger structure seen as a homogeneous texture, which may define another basis for characterizing the homogeneity of a segment. The result of segmentation is a partitioning of the image such that each part is homogeneous in some salient property relative to its surroundings.

Three-dimensional interpretation. One central objective of image interpretation is to infer the three-dimensional (3D) structure of the scene from images that are only two-dimensional (2D). The missing third dimension necessitates that assumptions be made about the scene so that the image information can be extrapolated into a three-dimensional description. The presence in the image of a variety of three-dimensional cues is exploited. These cues may occur in a single image taken by a camera at a single time instant, a sequence of images acquired by one camera over a time interval, a set of images taken by multiple cameras from different positions or angles at a single time instant, or a time sequence of images taken by multiple cameras. In each case, the main task is to devise algorithms that can estimate three-dimensional structural parameters from image-based measurements. Some examples of such cues will now be discussed.

Image curves. The boundaries of image regions, or the curves composing a line-drawing image, reveal scene characteristics such as the extent and shape of an object and the occlusion between objects.

Shading. The nature of gradual spatial variation of image values within a region is related to surface shape characteristics such as convex or concave, and planar or curved.

Texture gradient. Variation in the coarseness of image texture is indicative of how the corresponding texture surface is oriented in the scene. For example, the distant flowers in an image of a large field of flowers are packed more densely together than the closer ones.

Stereo. If two cameras are placed at different locations and orientations (like the human eyes), then the coordinates at which a given scene projects within the two images are different, and their disparity is trigonometrically related to the three-dimensional position of the scene point. The disparity variation within an image can be used to estimate the three-dimensional configuration of surfaces in the scene. Three or more cameras can also be used.

Dynamic cues. If there is relative motion between the scene and the camera, the image data consists of a dynamic image sequence. Each of the above-mentioned cues then provides added information since observations about the temporal behavior of the cue are available, in addition to its spatial properties. For example, the temporal variation in image values, called optical flow, can be used to estimate the relative motion, surface shape, and layout of the objects. A moving object gives rise to a decreasing amount of flow as its distance from the camera increases. Similarly, moving image curves, such as those corresponding to orientation discontinuities on an object surface or the silhouette of a rotating object, yield a wealth of information about the shape of an object as well as about its motion. Sequences may be taken by multiple cameras simultaneously, making surface estimation easier because a larger amount of data is

available, that is, the spatial disparity as well its variation over time.

Active image acquisition. If the cameras maintain a single, fixed geometry or optical configuration, the amount of additional information extracted from successive images depends on the changes occurring in the scene, due, for example, to relative motion. It is possible to maximize the increment in scene information obtained from successive images by dynamically reconfiguring the cameras so that the cues in the new images are most informative. Thus, a partial interpretation of the scene is used to dynamically control the sensing parameters so that each stage of image acquisition adds the most to the interpretation. For example, if a scene is too bright, the aperture size may be reduced; if the object is not in sharp focus, the focus setting may be changed; and if the object is not well placed with respect to both cameras in stereo analysis, the cameras may be verged to fixate on the object, all of these actions happening simultaneously (see **illus.**). As a certain part of the scene is satisfactorily interpreted, the results of the interpretation may be used to determine where to point the cameras next, and even to suggest how to analyze these new parts so as to obtain the best final interpretation in the minimum time.

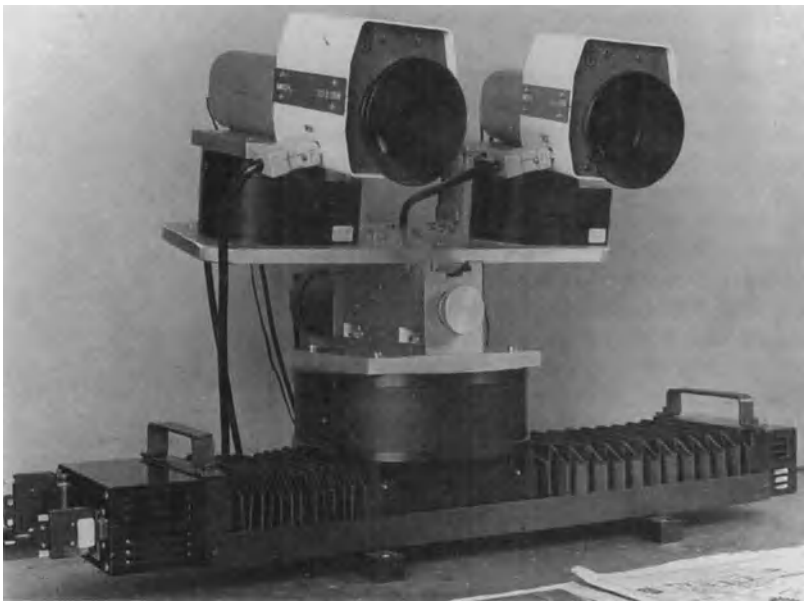

Active camera system that can dynamically vary the part of the scene it has under view, focus on it, adjust the zoom, and open the aperture as necessary. Performing such changes under computer control and on the basis of the contents of the scene allows autonomous visual exploration.

Representation. The two-dimensional structure of an image or the three-dimensional structure of a scene must be represented so that the structural properties required for various tasks are easily accessible. For example, the hierarchical two-dimensional structure of an image may be represented through a pyramid data structure which records the recursive embedding of the image regions at different scales. Each region's shape and homogeneity characteristics may themselves be suitably coded. Alternatively, the image may be recursively split into parts in some fixed way (for example, into quadrants) until each part is homogeneous. This approach leads to a tree data structure. Analogous to two dimensions, the three-dimensional structures estimated from the imaged-based cues may be used to define three-dimensional representations. The shape of a three-dimensional volume or object may be represented by its three-dimensional axis and the manner in which the cross section about the axis changes along the axis. Analogous to the two-dimensional case, the three-dimensional space may also be recursively divided into octants to obtain a tree description of the occupancy of space by objects.

Recognition. A second central objective of image interpretation is to recognize the scene contents. Recognition involves identifying an object based on a variety of criteria. It may involve identifying a certain object in the image as one seen before. A simple example is where the object appearance, such as its color and shape, is compared with that of the known, previously seen objects. A more complex example is where the identity of the object depends on whether it can serve a certain function, for example, drinking (to be recognized as a cup) or sitting (to be recognized as a chair). This requires reasoning from the various image attributes and the derivative three-dimensional characteristics to assess if a given object meets the criteria of being a cup or a chair. Recognition, therefore, may require extensive amounts of knowledge representation, reasoning, and information retrieval.

Visual learning. Visual learning is aimed at identifying relationships between the image characteristics and a result based thereupon, such as recognition or a motor action. For example, the essentials of the visual appearance of a person may be learned so that whenever the person reappears in the scene his or her identity will be recognized despite changes such as in lighting and observer's viewpoint. Alternatively, continuous use of visual feedback to perform a task may be learned. For example, a person trying to pick up a suitcase moves his or her hands toward it while continuously using vision to control and correct the hand configuration and motion to reach the desired location and grip.

Applications. Vision is the predominant sense in humans, and it can play a correspondingly central role in computer-aided systems. In manufacturing, vision-based sensing and interpretation systems help in automatic inspection, such as identification of cracks, holes, and surface roughness; counting of objects; and alignment of parts. Computer vision helps in proper manipulation of an object, for example, in automatic assembly, automatic painting of a car, and automatic welding. Autonomous navigation, used, for example, in delivering material on a cluttered factory floor, has much to gain from vision to improve on the fixed, rigid paths taken by vehicles which follow magnetic tracks prelaid on the floor. Recognition of symptoms, for example, in a chest x-ray, is important for medical diagnosis. Classification of satellite pictures of the Earth's surface to identify vegetation,

water, and crop types, is an important function. Automatic visual detection of storm formations and movements of weather patterns is crucial for analyzing the huge amounts of global weather data that constantly pour in from sensors. *See* ARTIFICIAL INTELLIGENCE; CHARACTER RECOGNITION; COMPUTER GRAPHICS; INTELLIGENT MACHINE; ROBOTICS. Narendra Ahuja

Bibliography. S. Florczyk, *Robot Vision: Video-based Indoor Exploration with Autonomous and Mobile Robots*, 2005; R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2d ed., 2002; A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, vols. 1 and 2, 2d ed., 1999; L. G. Shapiro and G. C. Stockman, *Computer Vision*, 2001; E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*, 1999.

Computerized tomography

An imaging technique which uses an array of detectors to collect information from a beam that has passed through an object (for example, a portion of the human body). The information collected is then used by a computer to reconstruct the internal structures, and the resulting image can be displayed—for example, on a television screen.

Technique. Wave phenomena can penetrate into regions where it is impossible or undesirable to introduce ordinary probes—regions as diverse as the ocean, the human body, and the deep interior of the Earth. Since wave attenuation and velocity are affected by material properties, such as temperature and density, information concerning these properties is accumulated by a wave packet as it travels along its path. Thus, the intensities and travel times of such pulses measured at a receiver represent such accumulations. Although a single such datum does not allow point-to-point variation to be inferred, when many rays crisscross a region the corresponding interrelated data contain constraints that allow an inversion to determine interior structures. Considerable computational power and mathematical sophistication are needed in dealing with indirect, incomplete, and imperfect measurements. The relevant branch of mathematics is called inverse theory. The term computerized tomography for this type of remote sensing arose in medicine, where x-ray attenuation data along many straight paths confined to a plane were used to obtain a map of density structure in a body section. The term tomography refers to methods that display only thin slices or sections through an object or a human body. *See* INVERSE SCATTERING THEORY.

Examining a person or object in the laboratory requires the use of a gantry, composed of an x-ray tube, an array of detectors opposite the tube, and a central aperture in which the person or object is placed. The rigid gantry maintains the proper alignment between the x-ray tube and the detectors.

Earlier computerized tomography scanners had a conventional x-ray tube with a rotating anode and small focal spot. X-rays generated in short bursts, usually lasting 2–3 ms, were picked up by detectors after passing through a section of the body (for example, the abdomen). Invisible images picked up by the detectors were fed to the computer, which reconstructed the image. The thickness of the section to be examined varied from 1 to 10 mm and was determined by collimating the x-ray beam.

Less sophisticated scanners used linear and rotational motions of both x-ray tube and detector. One employs a pencil-shaped x-ray beam and a single detector. The time needed for a single slice or scan was 4.5–5 min. A fan-shaped beam with an array of detectors (as many as 30) and a greater rotation increment (30° instead of 1°) made possible a time interval of 10–90 s.

Sophisticated electronic technology eliminates the need for mechanical movement. An electron beam is used, which is focused on stationary anode rings located in the gantry surrounding the patient, making it possible to generate considerable x-ray output within a very short time and to acquire images while the patient is advancing through the gantry. Up to 34 scans or slices can be obtained within 1 s; a single slice can be obtained in only 50 ms. Some computerized tomography scanners use as many as 1536 measuring channels constituted by xenon detectors with extremely high efficiency. *See* IONIZATION CHAMBER; SCINTILLATION COUNTER.

The computer reconstructs the image from the information collected by the detectors in the following way. The cross section of the object to be reconstructed is divided into tiny blocks, called voxels, giving a square matrix. Matrices can be constructed having voxels of different sizes; for example, 80×80, 160×160, 320×320, 512×512, and 1024×1024 voxels. In the 320×320 matrix, the voxels are simply one-fourth of the volume of those of an 80×80 matrix, and therefore the resolution is greater. The computer assigns each voxel a number proportional to the degree that the voxel has attenuated the beam passing through it. Once the voxel has received a computer number, it is called a pixel (picture element). The linear attenuation coefficient (a measure of the quantity of radiation attenuated by each centimeter of absorber) is used to quantify the attenuation of the beam. Each voxel is "examined" by the beam incident from several positions, generating many "positional" equations. The computer must solve many thousands of equations to determine the linear attenuation coefficient for a single voxel. In order to obtain enough information to calculate one image, at least 90,000 readings (300 pulses and 300 detectors) are needed.

Applications. The ability of computerized tomography to explore the internal structure of objects has made it a valuable tool in medicine and geophysics.

Medical tomography. In medicine, computerized tomography represents a noninvasive way of seeing internal structures. In the brain, for example, computerized tomography can readily locate tumors and hemorrhages, thereby providing immediate information for evaluating neurological emergencies. Equally important is the fact that, in a great majority of cases, computerized tomography offers an alternative approach to the more problematic invasive procedures.

Reductions in scanning time have made it possible to do whole-body scans, and many abdominal tumors, hemorrhages, and infections can now be detected without the need for exploratory surgery. Tumors deep in the body can be biopsied easily with fine needles by using computerized tomography guidance, thereby alleviating the need for surgery. Another advantage of computerized tomography is three-dimensional reconstruction. It is most useful in cases of fracture of the hip or facial bones, helping the surgeon to do reconstructive surgery. Other medical imaging techniques that make use of computerized tomographic methods include magnetic resonance imaging, positron emission tomography, and single-photon emission tomography. *See* MEDICAL IMAGING; RADIOGRAPHY. Nicole-Fr. Bolender

Geophysical tomography. After the success of computerized tomography in medicine, its possibilities in other fields were quickly realized. In the earth, atmospheric, and ocean sciences it has supplemented, but by no means replaced, older methods of remote sensing. Seismic tomography is now an important tool for investigating the deep structure of the Earth, testing theories such as plate tectonics, and exploring for oil. Ocean acoustic tomography is applied to physical oceanography, climatology, and antisubmarine warfare. Atmospheric tomography finds applications to weather, climate, and the environment.

1. *Seismic tomography.* The data for tomographic analysis are the earthquake signal arrival times monitored by the global network of seismic stations. These signals are propagated as elastodynamic waves that can be classified as compressional or shear and as body or surface waves. Compressional waves such as acoustic waves consist of alternating compressions and rarefactions caused by the oscillations of material particles in the direction of wave propagation. Shear waves are of a different (slower) type in which the motion of material particles is at right angles to the wave motion. Surface waves are guided by the Earth's surface; they travel along great circle paths and penetrate the upper mantle, but not deeper. However, body waves are not confined to the surface and can penetrate deeply, even reaching the Earth's core. They can be of the compressional or shear type. There are also two kinds of surface waves: Rayleigh waves, composed of a vertical shear component and a compressional component, and Love waves, which are pure horizontal shear waves. *See* SEISMOLOGY.

The diversity of wave phenomena allows various physical variables to be computed from wave velocities. Temperature, density, and crystalline structure of rock are the most important of these. Body-wave ray paths are not straight as in medical tomography but are curved by refraction in the vertical due to variation in the Earth's mechanical properties. Body and surface ray paths and travel times are calculated from arrival time data at many stations by means of a standard spherically symmetric reference model of the Earth. The travel time associated with thousands of crossing paths are treated as tomographic data and used to draw maps of wave velocity anomalies (deviations from the reference model). These maps reveal much about internal structure and dynamics of the Earth. Since velocity decreases as temperature increases, an anomalously slow velocity may point, for example, to upwelling hot material from the mantle toward the surface. An opposite anomaly may represent subduction of cool surface matter as one tectonic plate is thrust beneath another. Large anomalies near the core have revealed a previously unknown aspherical (bumpy) structure. *See* PLATE TECTONICS.

Smaller-scale seismic tomography serves as a tool for oil exploration. In cross-borehole tomography a seismic source such as a small explosive or impact device creates a mechanical disturbance at a number of different depths in a shaft. A linear array of receivers in a second borehole measures signal arrival times. The propagation paths connecting each source position with each receiver form a tomographic network. Maps of propagation velocity in the region between boreholes are used to detect changes in rock strata. *See* SEISMIC EXPLORATION FOR OIL AND GAS.

2. *Ocean acoustic tomography.* A tomographic array in the ocean consists of a number of acoustic sources and receivers that span an ocean region of interest. Each source transmits a pulse that is detected at each receiver. Because of refraction due to temperature variations and other (smaller) effects, the pulses follow curved paths, which usually reverse direction in the vertical at a number of upper and lower turning points. These paths are determined mainly by the temperature stratification of the ocean and tend to be restricted to a vertical plane because of strong vertical stratification. The elements of each source-receiver pair are connected by many acoustic paths (called multipaths). Since a system typically contains many source-receiver pairs, the ocean volume within the tomographic array is penetrated by hundreds of acoustic paths. Pulse travel times along these paths are the measured data. They are determined by the variation of sound speed and currents along the respective paths. To the first order, sound speed in the ocean is determined by temperature, and the effect of currents is negligible. Thus, the information contained in the data is related mainly to the temperature variation in the ocean, and the monitoring of this distribution has been the primary objective in the development of this technology. However, monitoring of currents is also possible by the use of reciprocal (two-way) transmissions. By subtracting reciprocal travel times, temperature effects are canceled out and the effect of currents becomes dominant.

The operation of a system begins with initialization of the array, referring to the fact that a knowledge of the distribution of ocean variables (mainly temperature) is required at the initial moment of operation in order to identify the acoustic paths that connect sources and receivers. Thereafter, as these paths change position, they can be tracked by the system. In early experiments the necessary initial information was supplied by ship surveys of vertical temperature profiles.

It is not possible in the ocean environment to achieve the extremely fine spatial resolution obtained in medical scanning; a prohibitively large number of sources and receivers would be required.

However, future ocean monitoring systems will include satellite observations, providing global coverage of the ocean surface with excellent horizontal resolution but no significant vertical penetration. Acoustic tomography provides vertical penetration but poorer resolution. These two techniques can play complementary roles in ocean monitoring. *See* OCEANOGRAPHY; UNDERWATER SOUND.

3. *Atmospheric tomography.* Electromagnetic waves interact with the atmosphere in two ways useful for remote sensing. Parts of the atmosphere can reflect incident radiation directly to a detector or can absorb energy and reradiate it. Lidar, a form of radar that uses laser beams, is a suitable source for active tomographic systems because it confines incident energy to narrow beams and because optical frequencies interact strongly with the atmosphere. Alternatively, a system can be based on passive detection of emitted energy from the atmosphere by means of the radiometer, an instrument that measures the intensity of waves arriving from a particular direction. For active systems, measured data are the attenuation of lidar pulses along straight path segments from source to point of reflection to receiving telescope. For passive systems, a measured datum is the total intensity of radiation emitted toward the radiometer by all sources on its line of sight. *See* LIDAR.

A simple example of a configuration that provides the network of crossing paths required for tomography consists of two radiometers in fan beam mode: Each instrument moves its line of sight up and down, forming a fan of receiving beams. The intersection of the two fans forms a network of paths for tomography; this configuration has been used to map the liquid water content of clouds. Other tomographic configurations have been obtained from instruments mounted on satellites, aircraft, surface vehicles, and the ground. *See* AIRGLOW.

As a characteristic tomography, the data represent not the value of physical variables at a point but the integration of variables along a line. Point-to-point mapping is accomplished by inversion algorithms similar to those used in other fields. Although attenuation and emissivity are often of interest in themselves, they can also be related to other physical variables of interest. Thus, atmospheric applications also include mapping of visible airglow, temperature, and properties of aerosols. Thomas J. Eisler; Ronald New

Bibliography. S. C. Bushong, *Computed Tomography*, 2000; H. M. Iyer and K. Hirahara (eds.), *Seismic Tomography: Theory and Practice*, 1993; A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, 2001; W. A. Kalender, *Computed Tomography: Fundamentals, System Technology, Image Quality, Applications*, 2d ed., 2006; V. Kunitsyn and E. Tereshchenko, *Ionospheric Tomography*, 2003; S. K. Lee, S. S. Sagel, and R. Stanley, *Computed Body Tomography: Physical Principles and Image Quality Considerations*, 3d ed., 1998; W. Menke, *Geophysical Data Analysis: Discrete Inverse Theory*, 1984, rev. 1989; W. H. Munk and P. F. Worcester, Ocean acoustic tomography, *Oceanography*, 1:8–10, 1988; F. Natterer, *The Mathematics of Computerized Tomography*, 1986; H. Rigaults et al., A six month clinical evaluation with the Somatom Plus, *Electromedia*, vol. 58, no. 1, 1990; E. Seeram, *Computed Tomography: Physical Principles, Clinical Applications, and Quality Control*, 2d ed., 2001.

Concentration scales

Numerical systems defining the quantitative relations of the components of mixtures. In solutions the concentration is expressed as the mass, volume, or number of moles of solute present in proportion to the amount of solvent or of total solution. Each scale of concentration has significant features of experimental simplicity or of theoretically significant relationship.

Percentage. The simplest scale to measure is percentage; hence it is often used for medicinal or household solutions. Weight percent is the number of parts of weight of solute per hundred parts of weight of solution (total). For example, a 10% saline solution contains 10 g of salt in 90 g of water, that is, 100 g total weight. Gaseous mixtures, being difficult to weigh, are often expressed as volume percent. Thus, air is said to contain 78% nitrogen by volume. Solutions of liquids in liquids (say, alcohol in water) may also be expressed in volume percent.

Molarity. To the chemist, the number of moles of solute is of more significance than the number of grams. In SI units the concentration would be given in moles per cubic meter (mol m^{-3}). The molarity (abbreviated *M*) is the number of moles of solute per liter of total solution. Thus, 12 *M* HCl means that the solution contains 12 formula weights (12 × 36.5), or 438 g, of hydrochloric acid per liter (or, in SI units, mol dm^{-3}). Concentration is an intensive, rather than an extensive, property of solutions. Thus, 0.5 liter of the acid just mentioned is still 12 *M*, although it contains only 6 moles of solute. As used here, the mole is an amount of substance whose weight in grams is numerically the same as the molecular weight. Chemical engineers sometimes use a mole which is a pound-molecular weight.

Molality. Certain solution properties, for example, the lowering of the freezing point of water by addition of salt, require the use of a concentration scale which relates the number of moles of solute to the weight of solvent rather than to the volume of solution. This scale, called molality (abbreviated *m*), indicates the number of moles of solute per 1000 g of solvent. Thus 34.2 g of sucrose ($C_{12}H_{22}O_{11}$, mol wt 342), if dissolved in 200 g of water, has the concentration of 0.5 mole of sucrose per 1000 g of water, and hence is 0.5 *m*. For dilute aqueous solutions, the molality is essentially identical with the molarity. In solutions with densities other than unity, the two scales differ. Molality may be computed from molarity, as shown in the equation below, by first subtracting from the weight of 1 liter (1000 × density *d*)

$$m = M\left(\frac{1000}{1000d - M \times \text{mol wt}}\right)$$

the weight of solute, and then scaling the result by

proportion to the number of moles per 1000 g, where the molecular weight is that of the solute.

Normality. Molarity does not represent reactive capacity for solutes which possess more than one active unit per molecule. Since sulfuric acid (H_2SO_4) molecules yield twice as many hydrogen ions (H^+) as do those of hydrochloric acid, a liter of 0.1 *M* sulfuric acid will neutralize twice as much base as will a liter of 0.1 *M* hydrochloric acid. When it is important to know the reactive capacities of reagents, as in volumetric analysis, the normality scale is used. Normality (abbreviated *N*) is found by multiplying molarity by the number of active units in the formula.

In methathesis (double decomposition) reactions, normality may be an ambiguous concept unless referred to a specific reaction. Phosphoric acid (H_3PO_4) may be reacted with sodium hydroxide (NaOH) to yield NaH_2PO_4, Na_2HPO_4, or Na_3PO_4; thus, phosphoric acid may have three normalities, depending on whether the reaction involves replacing one, two, or three hydrogen atoms. In oxidation-reduction reactions, there is a change in oxidation number, and normality must be calculated on the basis of this change. *See* OXIDATION-REDUCTION.

For reaction in a solution, the product of normality times the volume is the same for both reactants; for a solution of normality N_1, determined by titrating a volume V_1, with volume V_2 of a known solution of normality N_2, this equivalence can be expressed as $V_2N_2 = N_1V_1$. If the volume is expressed in milliliters, the products are in milliequivalents; if in liters, in equivalents.

Formality. The molarity scale is ambiguous when applied to ionic reactions. For example, 1 mole of Na_2HPO_4 in a liter of solution can yield 2 moles of sodium ion (Na^+), 1 mole of H^+, and 1 mole of phosphate ion (PO_4^{3-}), leaving zero mole of undissociated Na_2HPO_4. What the proper molarity is may not be clear. To obviate the confusion, it is possible to disregard moles and to use instead the number of formula weights of solute per liter of solution. This scale is designated the formality scale. In the example given, the solution would be 1 *F* Na_2HPO_4. This identifies the amount of total solute per liter and removes ambiguity in making up a solution; the reactive capability of the solution under various conditions is not considered.

Other scales. For very dilute solutions, for example, hard water, it is useful to express concentration in parts per million (ppm). Parts per billion are given in nanograms (10^{-9} g) per liter. With modern analytical techniques, quantities can be expressed in picogram (10^{-12} g) and femtogram (10^{-15} g) ranges. Particular applications often require special scales; for example, radon concentration in the atmosphere is given in units of picocuries per liter.

Mole percent (mole fraction). Many properties of solutions (for example, vapor pressure of one component) are dependent on the ratio of the number of moles of solute to the number of moles of solvent, rather than on the ratios of respective volumes or masses. The mole fraction (abbreviated N_A or X_A for component *A*) is the ratio of the number of moles of solute to the total number of moles of all components. Thus for 16 g of methanol (0.5 mole) dissolved in 18 g of water (1 mole), the mole fraction of methanol is 0.5/1.5, or 1/3; the mole percent is 33.3. For gases the mole percent is identical with the volume percent. *See* GRAM-MOLECULAR WEIGHT; SOLUTION; STOICHIOMETRY; TITRATION. Allen L. Hanson

Bibliography. D. E. Goldberg, *Fundamentals of Chemistry*, 4th ed., 2003; G. I. Sackheim, *Chemical Calculations*, 17th ed., 2000; K. W. Whitten et al., *General Chemistry*, 7th ed., 2003.

Concrete

A composite material that consists essentially of a binding medium, such as a mixture of portland cement and water, within which are embedded particles or fragments of aggregate, usually a combination of fine and coarse aggregate.

Concrete is by far the most versatile and most widely used construction material worldwide. It can be engineered to satisfy a wide range of performance specifications, unlike other building materials, such as natural stone or steel, which generally have to be used as they are. Because the tensile strength of concrete is much lower than its compressive strength, it is typically reinforced with steel bars, in which case it is known as reinforced concrete. *See* REINFORCED CONCRETE.

Materials. A composite material is made up of various constituents. The properties and characteristics of the composite are functions of the constituent materials' properties as well as the various mix proportions. Before discussing the properties of the composite, it is necessary to discuss those of the individual constituents as well as the effects of the mix proportions and methods of production. *See* COMPOSITE MATERIAL.

Cement. There are many different kinds of cements. In concrete, the most commonly used is portland cement, a hydraulic cement which sets and hardens by chemical reaction with water and is capable of doing so under water. Cement is the "glue" that binds the concrete ingredients together and is instrumental for the strength of the composite. Although cements and concrete have been around for thousands of years, modern portland cement was invented in 1824 by Joseph Aspdin of Leeds, England. The name derives from its resemblance of the natural building stone quarried in Portland, England. *See* CEMENT.

Portland cement is made up primarily of four mineral components (tricalcium silicate, dicalcium silicate, tricalcium aluminate, and tetracalcium aluminoferrite), each of which has its own hydration characteristics. By changing the relative proportions of these components, cement manufacturers can control the properties of the product.

The primary product of cement hydration is a complex and poorly crystalline calcium-silicate-hydroxide gel (or CSH). A secondary product of hydration is calcium hydroxide, a highly crystalline material. A category of siliceous materials known as

pozzolans have little or no cementitious value, but in finely divided form and in the presence of moisture will react chemically with calcium hydroxide to form additional CSH. This secondary hydration process has a generally beneficial effect on the final concrete properties. Examples of pozzolans are fly ash, ground granulated blast-furnace slag, and microsilica or silica fume.

The American Society for Testing and Materials (ASTM) defines five types of cement, specifying for each the mineral composition and chemical and physical characteristics such as fineness. The most common cement is Type I. Type III cement is used if more rapid strength development is required. The other types are characterized by either lower heat of hydration or better sulfate resistance than that of Type I cement.

Aggregate. The aggregate is a granular material, such as sand, gravel, crushed stone, or iron-blast furnace slag. It is graded by passing it through a set of sieves with progressively smaller mesh sizes. All material that passes through sieve #4 [0.187 in. (4.75 mm) openings] is conventionally referred to as fine aggregate or sand, while all material that is retained on the #4 sieve is referred to as coarse aggregate, gravel, or stone. By carefully grading the material and selecting an optimal particle size distribution, a maximum packing density can be achieved, where the smaller particles fill the void spaces between the larger particles. Such dense packing minimizes the amount of cement paste needed and generally leads to improved mechanical and durability properties of the concrete.

The aggregate constitutes typically 75% of the concrete volume, or more, and therefore its properties largely determine the properties of the concrete. For the concrete to be of good quality, the aggregate has to be strong and durable and free of silts, organic matter, oils, and sugars. Otherwise, it should be washed prior to use, because any of these impurities may slow or prevent the cement from hydrating or reduce the bond between the cement paste and the aggregate particles.

Fig. 1. Workers placing and vibrating concrete on a bridge deck including epoxy-coated reinforcing steel. (*Portland Cement Association*)

Admixtures. While aggregate, cement, and water are the main ingredients of concrete, there are a large number of mineral and chemical admixtures that may be added to the concrete. The four most common admixtures will be discussed.

1. Air-entraining agents are chemicals that are added to concrete to improve its freeze–thaw resistance. Concrete typically contains a large number of pores of different sizes, which may be partially filled with water. If the concrete is subjected to freezing temperatures, this water expands when forming ice crystals and can easily fracture the cement matrix, causing damage that increases with each freeze–thaw cycle. If the air voids created by the air-entraining agent are of the right size and average spacing, they give the freezing water enough space to expand, thereby avoiding the damaging internal stresses.

2. Water-reducing admixtures, also known as superplasticizers, are chemicals that lower the viscosity of concrete in its liquid state, typically by creating electrostatic surface charges on the cement and very fine aggregate particles. This causes the particles to repel each other, thereby increasing the mix flowability, which allows the use of less water in the mix design and results in increased strength and durability of the concrete. *See* FLOW OF SOLIDS.

3. Retarding admixtures delay the setting time, which may be necessary in situations where delays in the placement of concrete can be expected. Accelerators shorten the period needed to initiate cement hydration—for example, in emergency repair situations that call for the very rapid development of strength or rigidity.

4. Color pigments in powder or liquid form may be added to the concrete mix to produce colored concrete. These are usually used with white portland cement to attain their full coloring potential. *See* PIGMENT (MATERIAL).

Reinforcing steels. Because of concrete's relatively low tensile strength, it is typically reinforced with steel bars (**Fig. 1**). These bars are produced in standard sizes. In the United States, the identification number of a reinforcing bar refers to the nominal diameter expressed in eighths of an inch. For example, a number 6 bar has a diameter of $^6/_8 = 0.75$ inch. The available bar sizes range in general from 2 to 18. Reinforcing steel usually has a nominal yield strength of 60,000 lb/in.^{2} (414 MPa). To improve the bond strength between the bars and the concrete, the bars are fabricated with surface deformations or ribs. The relatively high cost of steel mandates its sparing use. This means that the concrete is usually assigned the task of resisting compressive forces, while the steel carries primarily the tensile forces. The alkalinity of the cement paste generally provides sufficient protection of the steel against corrosion. However, corrosion protection is often breached, for example, in highway bridge decks with continuous pore

structure or traffic-induced cracks that permit the deicing chemicals used in winter to penetrate the protective concrete cover. Additional protective measures may be necessary, such as using epoxy coatings on the bars, noncorrosive steels, or nonmetallic reinforcement (for example, fiber-reinforced polymers). *See* CORROSION.

Other important concrete terminology can be defined. A mixture of cement and water is called cement paste. Cement paste plus fine aggregate is called mortar or concrete matrix. Mortar plus coarse aggregate constitutes concrete. Concrete reinforced with steel or other high-strength material is known as reinforced concrete. *See* MORTAR.

Production of concrete. The properties of the end product depend not only on the various constituent materials listed above but also on the way they are proportioned and mixed, as well as on the methods of placing and curing the composite.

Mix design. It is not possible to predict the strength and other concrete properties solely based on the properties and proportions of the mix components. Therefore, mixes are designed on an empirical basis, often with the help of trial mixes. The objective of the mix design is to assure that the product has specified properties in both the fresh and hardened state. The most important mix design variable is the weight ratio between water and cement, referred to as the w/c ratio. There is a theoretical minimum amount of water needed for the cement to completely hydrate, which can be determined using the equations of hydration chemistry. Any excess water creates pores which, together with any air-filled pores, do not contribute to the material strength. The result is a drastic decrease in strength as a function of increasing the w/c ratio. On the other hand, too low w/c ratios cause poor workability of the concrete. For practical reasons, the w/c ratio typically varies between 0.4 and 0.6. The other important mix design variables are the cement-to-aggregate ratio and the fine-to-coarse aggregate ratio. Also, the maximum aggregate size is of importance. And since cement is the most expensive bulk ingredient, the mix design will generally aim at the least amount of cement necessary to achieve the design objectives.

Construction practice. The material obtained immediately upon mixing of the various concrete ingredients is called fresh concrete, while hardened concrete results when the cement hydration process has advanced sufficiently to give the material mechanical strength. Concrete that is batched and mixed in a plant and then transported by truck in its fresh, or plastic, state to the construction site for final placement is called ready-mixed concrete. If the resulting structure or highway pavement, for example, remains in place after placement, the concrete is referred to as cast-in-place concrete, whether mixed on-site or off-site. Precast concrete refers to any structure or component that is produced at one site, typically in a precasting plant, and then transported in its hardened state to its final destination. The controlled environment of a precasting plant generally permits higher quality control of the product than is possible with cast-in-place concrete produced at a construction site. *See* CONSTRUCTION METHODS; PAVEMENT.

Code-writing organizations, such as the American Society for Testing and Materials, the American Concrete Institute (ACI), and the American Association of State Highway and Transportation Officials (AASHTO), have published detailed specifications and recommendations for measuring, mixing, transporting, placing, curing, and testing concrete. A proper mix design assures that the concrete mix is well proportioned. The mixing time should be sufficient to assure a uniform mixture. When placing the concrete, care should be taken to avoid segregation. For example, if dropped too far, the heavy or big aggregate particles can settle and lighter mix components, such as water, tend to rise. The concrete is conveyed from the mixing truck to its final destination in dump buckets by cableways or cranes or by pumping through pipelines. In modern high-rise building construction, concrete has been pumped as high as a thousand feet (330 m).

During placement, large amounts of air are entrapped in the mix, which lowers the strength of the hardened concrete. Much of the air is removed by compaction, which is achieved by either immersing high-frequency vibrators into the fresh concrete or attaching them to the outside faces of the formwork (Fig. 1). Care must be taken to avoid excessive vibration; otherwise the heavy aggregate particles settle down and the light mixing water rises to the surface.

For underwater construction, the concrete is placed in a large metal tube, called a tremie, with a hopper at the top and a valve arrangement at the submerged end. For so-called shotcrete applications such as tunnel linings and swimming pools, the concrete mixture is blown under high pressure through a nozzle directly into place to form the desired surface.

Before the concrete sets and hardens, it is relatively easy to give its exposed surfaces the desired finish. Surfaces cast against forms can be given various textures by using form liners or treating the surfaces after forms are removed. Hardened surfaces can be textured by grinding, chipping, bush-hammering, or sandblasting.

Curing. Once the concrete has been placed and compacted, it is critical that none of the mixing water needed for cement hydration is lost. This is the objective of curing. For example, in hot or dry weather large exposed surfaces will lose water by evaporation. This can be avoided by covering such surfaces with sheets of plastic or canvas or by periodically spraying them with water. In precast concrete plants, concrete elements are often steam-cured, because the simultaneous application of hot steam and pressure accelerates the hydration process, which permits high turnover rates for the formwork installations.

Quality control. To assure that the finished material has the specified properties, quality assurance and quality control procedures need to be implemented. From a public safety viewpoint, strength is the most important property. To assure adequate strength,

Fig. 2. Concrete slump test with a slump of 1.5 in., typical for pavement work. (*Portland Cement Association*)

such as determining the time of safe formwork removal, concrete batches are sampled by casting test cylinders at the same time and place as the structure being built. These cylinders are then tested by accredited laboratories to determine their strength. If the in-situ strength of existing structures needs to be evaluated, concrete cores may be drilled from selected parts of the structure and tested in the laboratory. There are also nondestructive test methods available to determine various properties of hardened concrete.

Properties of fresh concrete. The most important property of fresh concrete is its workability or flowability, because this determines the ease with which it can be placed. It is determined using a slump test, in which a standard truncated metal cone form is filled with fresh concrete (**Fig. 2**). The mold is then lifted vertically, and the resulting loss in height of the concrete cone, or the slump value, is indicative of the concrete's workability. For very liquid mixes, the flow test is performed, which is similar to the slump test, except that the mean diameter of the cake formed by the fresh concrete (or mortar) is measured.

A short while after casting, the concrete stiffens and loses its plasticity. The time of setting can be determined by repeatedly dropping a calibrated needle into the fresh concrete and measuring the time when the needle no longer sinks in.

Properties of hardened concrete. By far, the most important property of hardened concrete is its compressive strength. Since this strength continues to increase with continuing cement hydration, it is a function of age which is the time after casting. In the United States, the strength is determined 28 days after casting by loading standardized test cylinders up to failure. In Europe, test cubes are often used. Most commercially produced concrete has compressive strengths between 3000 and 6000 lb/in.^{2} (20 and 40 MPa). If loaded in tension, the material fails at a stress much lower than that, typically of the order of 10% of the compressive strength. Because of this low (and unreliable) tensile strength, concrete is usually reinforced with steel bars. *See* STRESS AND STRAIN.

During hydration and especially if allowed to dry after hardening, the concrete volume decreases by a small amount because of shrinkage. If this shrinkage is restrained somehow, it can lead to cracking. Shrinkage deformations caused by drying can be reversed only partially upon wetting. A concrete member or structure subjected to external load will undergo deformations which, up to a point, are proportional to the amount of applied load. If these loads remain in place for an appreciable time (months or years), these deformations will increase due to a material property called creep. Even for regular concrete mixes, creep deformations can be two or three times as high as the initial elastic deformations, especially if the concrete is loaded at a very young age. When designing concrete structures, such creep and shrinkage deformations must be accounted for. *See* CREEP (MATERIALS); ELASTICITY.

Durability. Durability is the ability of a material (or structure) to maintain its various properties throughout its design or service life. Some concrete structures built by the Romans served for over 2000 years. A material that loses its strength in time, for whatever reason, cannot be considered durable.

There can be numerous causes for loss of durability or deterioration of concrete structures. The most common one is an excessive amount of cracking or pore structure. Most concrete structures contain numerous cracks. But as long as these remain small (of the order of 0.25 mm or less), they are generally invisible to the naked eye, and the concrete remains basically impermeable to salts and other aggressive agents, so that it can continue to protect the reinforcing steel against corrosion. Larger cracks provide easy access for such agents to the steel, thereby promoting corrosion. Since the steel corrosion products occupy a larger volume than sound steel, they produce internal pressure during expansion and can spall off the protective concrete cover, the loss of which may render the structure unsafe to resist loads.

The concrete itself may deteriorate or weather, especially if subjected to many cycles of freezing and thawing, during which the pressure created by the freezing water progressively increases the extent of internal cracking. In addition, carbon in the atmosphere can react chemically with the cement hydration products. This process is known as carbonation. It lowers the pH of the concrete matrix to the point where it can no longer protect the steel against corrosion.

Most types of aggregate used for concrete production are inert; that is, they do not react chemically with the cement or hydration products. However, there are various aggregate types, including those containing amorphous silica such as common glass,

which react chemically with the alkali in the cement. In the presence of moisture, the alkali–aggregate reaction products can swell and cause considerable damage. The deterioration of numerous major structures and highway pavements has been attributed to such reactions, especially alkali–silica reaction, often after years of seemingly satisfactory service. Other common causes of chemical attack are sulfates found in soils, chlorides in seawater, acid rain, and other industrial pollutants. Generally, structures built with well-designed concrete mixes, having low porosity or high density and minimal cracking, are likely to resist most causes of chemical attack, although for service in particularly aggressive environments special countermeasures may have to be taken.

Under repeated load applications, structures can experience fatigue failure, as each successive load cycle increases the degree of cracking and material deterioration to the point where the material itself may gradually lose its strength or the increased extent of cracking is the source of loss of durability.

Thermal and other properties. The heavy weight of concrete [its specific gravity is typically 2.4 g/cm^3 (145 lb/ft^3)] is the source of large thermal mass. For this reason, massive concrete walls and roof and floor slabs are well suited for storing thermal energy. Because of this heat capacity of concrete, together with its reasonably low thermal conductivity, concrete structures can moderate extreme temperature cycles and increase the comfort of occupants. Well-designed concrete mixes are impermeable to liquids and therefore suitable for storage tanks without the need for impermeable membranes or liners. Although steel reinforcing bars conduct electricity and influence magnetic fields, the concrete itself does neither. *See* CONCRETE SLAB; FLOOR CONSTRUCTION; ROOF CONSTRUCTION.

Special concretes and recent developments. Concrete is an engineered material, with a variety of specialty products designed for specific applications. Some important ones are described below.

Lightweight concrete. Although the heavy weight or large mass of typical concrete members is often an advantage, there are situations where this is not the case. For example, because of the large stresses caused by their own heavy weight, floor slabs are often made lighter by using special lightweight aggregate. To further reduce weight, special chemical admixtures are added, which produce large porosity. Such high porosity (in either the matrix or the aggregate particles themselves) improves the thermal resistance of the concrete as well as sound insulation, especially for higher frequencies. However, because weight density correlates strongly with strength, ultralightweight concretes [1.1 g/cm^3 (70 lb/ft^3) and less] are used only for thermal or sound insulation purposes and are unsuitable for structural applications.

Heavyweight concrete. When particularly high weight densities are needed, such as for shielding in nuclear reactor facilities, special heavyweight aggregate is used, including barite, limonite, magnetite, scrap metal, and steel shot for fine aggregate. Weight densities can be achieved that are twice that of normal-weight concrete.

Architectural concrete. Concrete surfaces that remain exposed may call for special finishes or textures according to the architect's desires. Textures are most readily obtained by inserting special form liners before casting the concrete. Sometimes the negative imprint of roughly sawn timber is considered attractive and left without further treatment. Other surface textures are obtained by sandblasting, bush-hammering, and similar treatments. Ordinary portland cement gives concrete the typical gray color. By adding color pigments to the mix, a large variety of colors can be produced, especially in combination with white portland cement. Concrete mixed with specialty aggregate, such as marble, and ground smooth is known as terrazzo concrete, which is very popular for decorative surfaces on floors and walls. Recently, crushed postconsumer glass has been used as aggregate for decorative applications because of the esthetic possibilities, provided suitable countermeasures against alkali–silica reaction are taken (**Fig. 3**).

Fiber-reinforced concrete. The concrete matrix can be reinforced with short, randomly distributed fibers. Fibers may be metallic (primarily steel), synthetic (such as polypropylene, nylon, polyethylene, polyvinyl alcohol, and alkali-resistant glass), or natural (such as sisal, coconut, and rice husk). Such fibers are typically used in addition to conventional steel reinforcement, but in some applications as its replacement. For example, precast glass-fiber-reinforced building façade elements are widely used in the United States. By being uniformly distributed and randomly oriented, the fibers give the concrete matrix tensile strength, ductility, and energy absorption capacities that it otherwise would not have. In particular, when these fibers are engineered to optimize the fracture energy, so-called high-performance fiber-reinforced concrete is obtained, which has remarkable deformational characteristics and extraordinary resistance to blast and impact loads. In the

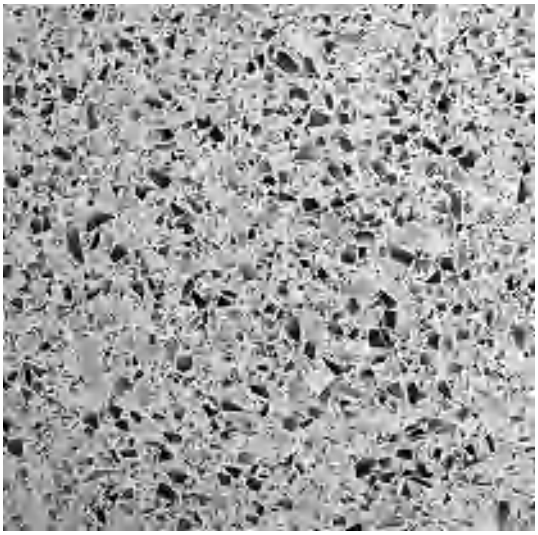

Fig. 3. Glass concrete tile. (*C. Meyer*)

concrete industry, it is very common to add small amounts of polypropylene fibers to reduce the extent of shrinkage cracking.

Textile-reinforced concrete. Whereas in fiber-reinforced concrete the fibers are short [usually no longer than 2 in. (5 cm)] and discontinuous, textile-reinforced concrete contains continuous woven or knitted mesh or textiles. Conceptually, such reinforcement acts similarly to conventional steel reinforcing bars or welded steel wire fabrics. But these fabric materials are noncorrosive and can have mechanical properties that are superior to those of steel. The fabrics can be premanufactured in a wide variety of ways, thereby lending themselves to new applications, especially for repairing or strengthening existing concrete structures. *See* TEXTILE.

Polymer-modified concrete. In polymer-modified concrete, also known as latex-modified concrete, a polymer is added to improve the material's strength, imperviousness, or both. In applications such as highway bridge decks, often a layer of latex-modified concrete is placed on top of a regular reinforced concrete deck for additional protection of the steel reinforcement. In polymer concrete, the hydraulic cement is replaced by an organic polymer as the binder. *See* POLYMER; POLYMERIC COMPOSITE.

Roller-compacted concrete. This type of concrete is formulated with very low contents of portland cement and water and therefore is of relatively low-cost. It is often used for pavements and dams. It can be transported by dump trucks or loaders, spread with bulldozers or graders, and compacted with vibratory rollers. Because the cement content is so low, the heat of hydration does not cause the kind of problems encountered in dams built with conventional concrete.

Ultra-high-strength concrete. Whereas concretes with compressive strengths of 6000 to 12,000 lb/in.^{2} (40 to 85 MPa) can now be categorized as high-strength, a new technology has been developed that results in strengths of 30,000 lb/in.^{2} (200 MPa) and higher. The key ingredient of this ultra-high-strength concrete is a reactive powder; therefore, it is also known as reactive-powder concrete. Other characteristics of this material are low water–cement ratios, carefully selected high-strength aggregates, and small steel fibers.

Self-leveling concrete. The need for good workability has been mentioned. The need for highly skilled workers who can properly compact concrete at the construction site prompted researchers in Japan to optimize the mix design such that the fresh concrete can flow into place without the need for further vibration. The main challenge was to obtain a low-viscosity mix without the threat of desegregation. This innovation is particularly important in applications with dense steel reinforcement, which traditionally have caused severe difficulties of producing high-quality concrete.

"Green" concrete. Concrete is by far the most widely used building material. Well over 10 billion tons are produced worldwide each year, requiring enormous natural resources. Also, it has been estimated that the production of 1 ton of portland cement causes the release of 1 ton of carbon dioxide (CO_2) into the atmosphere, a gas that is known to contribute to global warming. Together with the large amounts of energy required to produce portland cement, the cement and concrete industry has a major impact on the environment worldwide. Efforts are underway to reduce this impact and transform the industry to conform to the principles of sustainable development. The most significant step is the replacement of portland cement by other cementitious or pozzolanic materials, preferably materials that are by-products of industrial processes, such as fly ash (the by-product of coal-burning power plants) and granulated blast furnace slag (a by-product of the steel industry). To reduce the need for virgin aggregate, recycled concrete is the most promising approach, because construction debris, in particular demolished concrete, constitutes a major component of solid waste that fills up sparse landfill capacity. These recent developments are much more advanced in Europe and Japan than in the United States. But the "green" building movement is gaining momentum there as well, and for the concrete industry to maintain its dominant position within the construction industry, it is undertaking major efforts to make concrete a more "green" material.

Christian Meyer

Bibliography. ACI Committee 225, *Guide to the Selection and Use of Hydraulic Cements*, ACI Report 225R-99, American Concrete Institute, Farmingdale Hills, MI, 2001; ACI Committee 304, *Guide to Measuring, Mixing, Transporting, and Placing Concrete*, ACI Report 304R-00, American Concrete Institute, Farmingdale Hills, MI, 2001; B. Mather and C. Ozyildirim, *Concrete Primer*, 5th ed., ACI Spec. Publ. 1, 2002; P. K. Mehta and P. J. M. Monteiro, *Concrete*, 3d ed., McGraw-Hill, 1993; S. Mindess, J. F. Young, and D. Darwin, *Concrete*, 2d ed., Prentice Hall, Englewood Cliffs, NJ, 2003; A. M. Neville, *Properties of Concrete*, 4th ed., Wiley, New York, 1996.

Concrete beam

A structural member of reinforced concrete placed horizontally to carry loads over openings.

Because both bending and shear in such beams induce tensile stresses, steel reinforcing tremendously increases beam strength. Usually, beams are designed under the assumption that tensile stresses have cracked the concrete and the steel reinforcing is carrying all the tension. *See* STRESS AND STRAIN.

Two design theories are used, elastic design and ultimate-load design.

Elastic design. The following assumptions are made for elastic design:

1. Plane sections remain plane after bending and are perpendicular to the longitudinal fibers.
2. The stress-strain curve is a straight line.
3. The ratio n of the modulus of elasticity of steel E_s to that of concrete E_c is a constant $n = E_s/E_c$
4. The concrete does not carry tensile stress.

Transformed section. One approach to elastic design of reinforced concrete beams is to convert the steel to concrete. Because the steel and concrete are assumed to be firmly bonded and thus strained the same amount, the stress in the steel is n times the concrete stress. Hence the steel area may be replaced by an equivalent concrete area which is n times as large (**Fig. 1**).

If the equivalent area is placed at the same level as the steel and the moment of inertia I computed for the transformed section, the bending stresses can be computed from the simple flexural formula $f = Mc/I$ where M is the bending moment and c the distance from the neutral axis to the level at which stresses are to be computed.

Rectangular beams. The following formulas can be derived from the basic assumptions of elastic theory. Equation (1) locates the neutral axis, given steel and

$$\frac{nf_c}{f_s} = \frac{k}{1-k} \tag{1}$$

concrete extreme stresses, where n is the ratio of the modulus of elasticity of steel to that of concrete, f_c is the stress in the extreme fiber of the concrete, f_s is the stress in the steel, and k is the ratio of the distance between the top of the beam and the neutral axis to the distance between the top of the beam and the steel (**Fig. 2**).

The design equation for equal moment resistance of concrete and steel (balanced design) is Eq. (2).

$$k = \frac{1}{1 + f_s/nf_c} \tag{2}$$

Equations (3)–(6) are review equations.

$$k = \sqrt{2np - (np)^2} - np \tag{3}$$

$$j = \frac{1-k}{3} \tag{4}$$

$$M_c = {}^1/_2 f_c k j b d^2 \tag{5}$$

$$M_s = f_s A_s j d \tag{6}$$

Here p is the ratio of effective area of tension reinforcement to effective area of concrete in beams, j is the ratio of lever arm of resisting couple to depth d, M_c is the moment resistance of the concrete, b is the width of the rectangular beam or width of flange of the T beam, d is the depth from the compressive surface of beam or slab to center of longitudinal tension reinforcement, M_s is the moment resistance of the steel, and A_s is the steel area.

When M_s is less than M_c, the capacity of the steel determines the maximum moment that the beam will carry. The beam is called underreinforced. If the beam is loaded to failure, the steel rather than the concrete determines the maximum load that is sustained.

When M_c is less than M_s, the concrete determines the maximum moment the beam will carry. The beam is said to be overreinforced. Usually, overreinforced beams are avoided because they are not considered economical and failure may occur without warning.

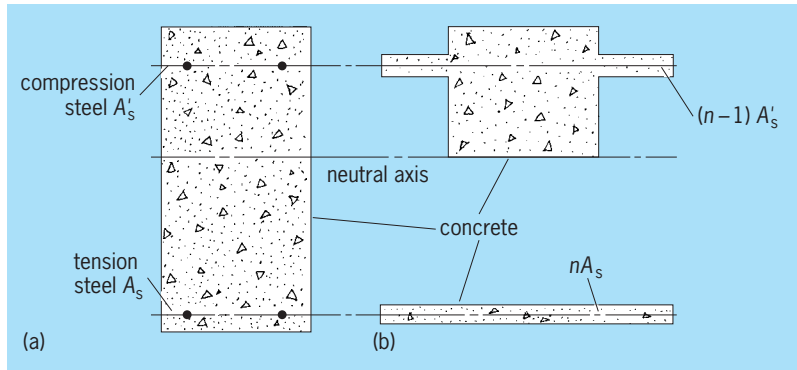

Fig. 1. Sketch of concrete beam. (*a*) Actual beam section. (*b*) Same section transformed. The use of the term $(n-1)A'_s$ prevents compression steel area from being included twice when computing moment of inertia of section; $(n-1)A'_s = nA'_s - A'_s$, because computations require the equivalent steel area to be added to the original concrete area above the neutral axis.

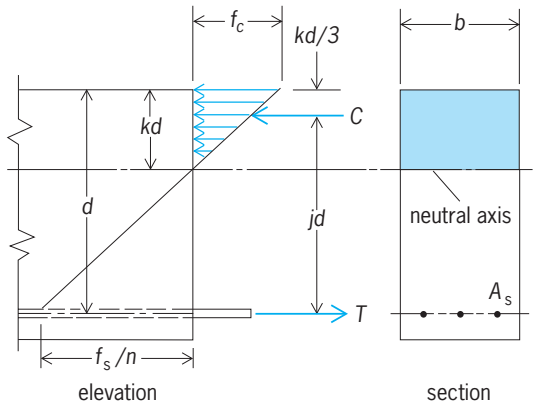

Fig. 2. Stress distribution in a rectangular beam designed according to elastic theory. *C* is total compressive force in concrete and *T* is total tensile stress in longitudinal reinforcement.

To design a rectangular beam by the elastic theory: (1) Select allowable unit stresses and determine k from the formula for balanced design; (2) compute j; (3) assuming the resisting moment of the concrete equal to the bending moment on the section, calculate bd^2 and select values for b and d; and (4) determine the steel area by equating the resisting moment of the steel to the bending moment on the section. This equation indicates that the amount of steel required can be reduced by increasing the depth. If the depth is fixed and an overreinforced beam results, it may be economical to use compression steel.

Compression-reinforced rectangular beams. Reinforcement may be added in the compression zone of concrete beams when the resisting moment of the concrete is less than the bending moment on the section or to avoid an overreinforced design. The compression steel is assumed to act as in plastic design; the Building Code Requirements of the American Concrete Institute allow the steel to take twice the stress given by elastic-theory formulas, provided the allowable tensile stress of the steel is not exceeded. Formulas for design of beams with

both tension and compression reinforcement can be derived from the elastic theory in the same manner as for beams with tension reinforcement only, but they are too complicated for ordinary use. Equations (7)–(12) are either exact or a close approximation.

$$k = \frac{1}{1 + f_s/nf_c} \tag{7}$$

$$f_s' = nf_c \frac{kd - d'}{d - kd} \tag{8}$$

$$A_s = \frac{M}{f_s jd} \tag{9}$$

$$M_c = {}^1/_2 f_c bkd \left(d - \frac{kd}{3}\right) \tag{10}$$

$$M_s' = M - Mc \tag{11}$$

$$A_s' = \frac{M_s'}{2f_s'(d - d')} \tag{12}$$

k = ratio of distance between top of beam and neutral axis to the distance between top of beam andtension steel
f_s = allowable steel stress
f_c = allowable concrete stress
n = ratio of modulus of elasticity of steel to that of concrete
f_s' = stress in compression steel computed from elastic theory formula
d' = distance from top of beam to compression steel
A_s = area of tension steel
M = bending moment on the section
jd = moment arm of the tensile
jd = reinforcement (assumed)
M_c = resisting moment of the concrete
b = width of beam
d = distance from top of beam to tension steel
M_s' = resisting moment of the compression steel
A_s' = area of compression steel

Concrete T beams. A slab cast integrally with a rectangular concrete beam usually is assumed to assist the beam in carrying loads. In regions of positive bending moment, the two act together as a T beam. In regions of negative moment, the beam is designed as a rectangular section because the slab is in tension and is assumed not to be able to resist such stresses, which must be taken by the reinforcing. *See* CONCRETE; CONCRETE SLAB.

If the neutral axis of a T beam is within the slab, it is designed as a rectangular beam, with width b the same as that of the flange, to resist bending moments. For shear, however, only the width of stem b' can be assumed to be effective.

If the neutral axis falls within the stem, the section can be designed as a T beam. However, the compression in the stem is negligible and can be ignored to simplify computations.

The American Concrete Institute's code recommends the following limits for the part of the slab that can be considered effective as the flange: (1) b shall be less than one-fourth the span length, (2) the overhanging width of flange shall not exceed eight times the slab thickness, and (3) the overhanging width shall not exceed one-half the clear distance between beams.

If t is the flange thickness and the remaining symbols are the same as for elastic-theory design of rectangular beams, Eqs. (13) through (20) can be used

$$k = \frac{1}{1 + f_s/nf_c} \tag{13}$$

$$\frac{f_s}{f_c} = \frac{bt(2kd - t)}{2A_s kd} \tag{14}$$

$$kd = \frac{2ndA_s + bt^2}{2nA_s + 2bt} \tag{15}$$

$$z = \frac{t(3kd - 2t)}{3(2kd - t)} \tag{16}$$

$$jd = d - z \tag{17}$$

$$M_s = A_s f_s jd \tag{18}$$

$$M_c = \frac{f_c btjd}{2kd}(2kd - t) \tag{19}$$

$$f_c = \frac{f_s}{n}\left(\frac{kd}{d - kd}\right) \tag{20}$$

for T beams when stem compression is neglected.

The shear unit stress in T beams is computed from Eq. (21), where V is the shear on the section. *See*

$$v = \frac{V}{b'd} \tag{21}$$

ELASTICITY; HOOKE'S LAW; YOUNG'S MODULUS.

Ultimate-load design. Other names for ultimate-load design are ultimate-strength design, limit-load design, and the plastic design. Design assumptions differ from those of the elastic theory principally in that the stress-strain curve is not a straight line. Instead of being designed to carry allowable unit stresses, as in the elastic theory, beams are proportioned to carry at ultimate capacity the design load multiplied by a safety factor.

Among the reasons for using ultimate-load design are the following:

1. The elastic theory is not corroborated with sufficiently great accuracy by beam tests.
2. It is logical to use different safety factors for live and dead loads. Different factors can easily be used with ultimate-load design but not with elastic design.
3. Column design is based on a modified ultimate load theory. To avoid inconsistency, structural members subjected to both bending and compressive stress should also be designed by ultimate-load theory.

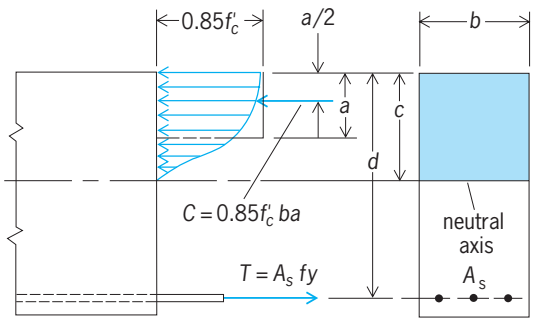

Fig. 3. Stress distribution in a rectangular beam under ultimate load; *T* and *C* are as in Fig. 2.

4. The ultimate strength of beams carrying both bending and axial compression, as determined in tests, conforms closely with ultimate-load theory.

Rectangular beams. The comprehenssive stress distribution may be assumed to be a rectangle, parabola, trapezoid, or any other shape that conforms to test data. Maximum concrete stress is assumed to be $0.85f'_c$, where f'_c is the compressive strength of a standard-test concrete cylinder at 28 days (**Fig. 3**).

When the moment resistance of the steel is less than that of the concrete, the bending moment that a rectangular beam with only tension reinforcement can sustain under ultimate load is Eq. (22).

$$M_u = \phi\left[A_s f_y(d - a/2)\right] = \phi\left[bd^2 f'_c q(1 - 0.59q)\right] \qquad (22)$$

A_s = area of tension reinforcement
f_y = yield point stress of steel
f'_c = comprehensive strength of standard-test concrete cylinder at 28 days
ϕ = reduction factor = 0.90 for this type of beam in flexure
b = width of beam
d = distance from top of beam to centroid of the steel
a = depth of rectangular stress block
$q = A_s f_y / bdf'_c$

Reinforcement ratio $p = A_s/bd$ should not exceed $0.75p_b$, where p_b is given by Eq. (23), and k_1 =

$$p_b = \frac{0.85k_1 f'_c}{f_y}\left(\frac{87,000}{87,000 + f_y}\right) \qquad (23)$$

0.85 for values of f'_c up to 4000 lb/in.^{2}, 0.80 for 5000 lb/in.^{2}, and 0.75 for 6000 lb/in.^{2}.

Compression-reinforced rectangular beams. Based on a nonlinear stress-strain relation, the bending moment that a rectangular beam with both compression and tension reinforcement can sustain under ultimate loads is shown in Eq. (24).

$$M_u = \phi[(A_s - A'_s)f_y(d - a/2) + A'_s f_y(d - d')] \qquad (24)$$

A_s = area of tension steel
A'_s = area of compression steel
f_y = yield point stress of the steel
d = distance from top of beam to centroid of tension steel
d' = distance from top of beam to centroid compression steel
$a = (A_s - A'_s)f_y/0.085f'_c b$
b = width of beam
f'_c = compressive strength of standard-test concrete cylinder at 28 days

Equation (24) holds only if Eq. (25) is true, where

$$(p - p') \geq 0.85k_1\frac{f'_c d'}{f_y d}\left(\frac{87,000}{87,000 - f_y}\right) \leq 0.75p_b \qquad (25)$$

$p = A_s/bd$; $p' = A'_s/bd$; and p_b and k_1 are the same as for beams with tension steel only.

Concrete T beams. Two cases should be considered, one with a relatively thick slab and one with a thin slab. If the flange thickness exceeds $1.18qd/k_1$, where $q = pf_y/f'_c$, the bending moment under ultimate load may be taken to be the same as that for a rectangular beam with tension reinforcement only. The value of p used in computing q should be A_s/bd, with b the width of the flange. For thinner flanges, use Eq. (26) to compute bending moment.

$$M_u = \phi[(A_s - A_{sf})f_y(d - a/2) + A_{sf}f_y(d - 0.5t)] \qquad (26)$$

$A_{sf} = 0.85f'_c(b - b)t/f_y$
f_y = yield point stress of the steel
t = flange thickness
b = width of flange
b' = width of stem
d = distance from top of beam to centroid of tension steel
$a = (A_s - A_{sf})f_y/0.85f'_c b'$

Equation (26) holds only if $p_w - p_f$ does not exceed $0.75p_b$, where $p_w = A_s/bd$, $p_f = A_{sf}/bd$, and p_b is the same as for beams with tension reinforcement only, as in Eq. (23).

Shear and bond. Maximum unit shear stress v acting on a section of a beam subjected to total shear V is given by Eq. (27), where b is the beam width, and

$$v = \frac{V}{bd} \qquad (27)$$

d is depth from top of beam to the tensile steel. The bond stress can be computed from Eq. (28), where

$$u = \frac{V}{\sum_o jd} \qquad (28)$$

Σ_o is the sum of the bar perimeters. This formula applies only to tension steel. Bond stresses will be at a maximum where shear is a maximum and the steel is in the tension side of the concrete.

To develop a given bar stress through bond, a bar

should be embedded a length L at least equal to Eq. (29), where a is the side of a square bar or the

$$L = \frac{f_s a}{4u} \tag{29}$$

diameter of a round bar. This length of embedment is called anchorage. Usually, bars are extended 10 diameters past the section where they are no longer required for bending stress.

Shear in itself is not as important in the design of a concrete beam as the tensile stresses that accompany it on a diagonal plane. To resist these stresses, concrete beams should be reinforced with bent-up bars or with stirrups. The latter are bars placed vertically or on an incline in a beam. They may be the legs of a single U-shaped bar or the sides of a rectangle.

The shear V' taken by stirrups is assumed to be the total shear on the section at which the stirrups are to be placed, less the shear taken by the concrete, vbd. The cross-sectional area of the stirrups needed at a section is given by Eq. (30), where f_v is the allowable

$$A_v = \frac{V's}{f_v d} \tag{30}$$

tensile stress of the steel, and s is the stirrup spacing.

If the stirrups, instead of being vertical, are laid at an angle α with the horizontal, greater steel area is required, as in Eq. (31). Every potential 45° crack

$$A_v = \frac{V's}{d(\sin\alpha + \cos\alpha)f_v} \tag{31}$$

should be crossed by at least one line of reinforcement.

If the area of the stirrups is given, the spacing can be computed from the two formulas above. The first stirrup is usually placed as close to the support as practical, generally 2 in. (5 cm). Stirrups should be placed throughout a beam, even if theoretically they are not needed. They serve also as supports for the longitudinal steel. When not required for diagonal tension, stirrups should be placed at most 18 in. (45 cm) apart.

It is common practice to cut off bars or bend them up where they are no longer needed to resist tension at the bottom of the beam. The bent bars serve as diagonal-tension reinforcement and tensile reinforcement at the top of the beam over the support.

Frederick S. Merritt

Bibliography. American Concrete Institute Staff and American Society of Civil Engineers Staff, *Specifications for Masonry Structures*, 1992; S. E. French, *Reinforced Concrete Technology*, 1994; A. H. Nilson, *Design of Concrete Structures*, 12th ed., 1997; R. R. Schneider and W. L. Dickey, *Reinforced Masonry Design*, 3d ed., 1994.

Concrete column

A structural member subjected principally to compressive stresses. Concrete columns may be unreinforced, or they may be reinforced with longitudinal bars and ties (tied columns) or with longitudinal bars and spiral steel (spiral-reinforced columns). Sometimes the columns may be a composite of structural steel of cast iron and concrete (see **illus.**).

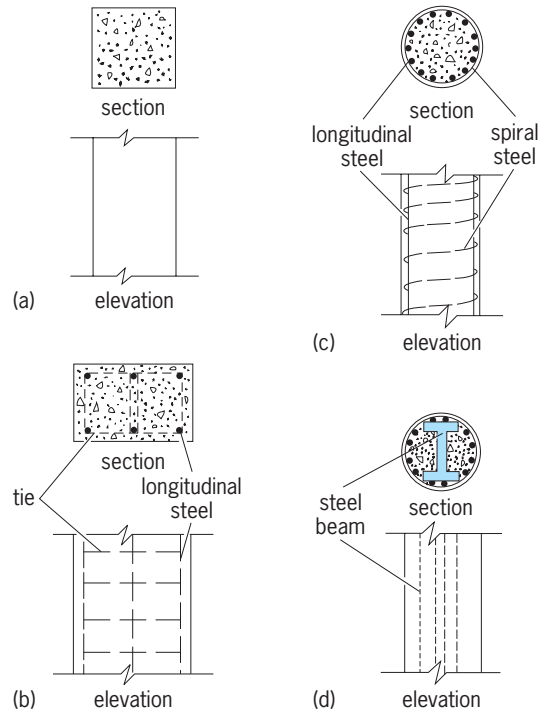

Column types. (*a*) Plain concrete. (*b*) Tied column. (*c*) Spiral-reinforced column. (*d*) Composite column.

Plain concrete columns. Unreinforced concrete columns are seldom used because of transverse tensile stresses and the possibility of longitudinal tensile stresses being induced by buckling or unanticipated bending. Because concrete is weak in tension, such stresses are generally avoided.

When plain concrete columns are used, they usually are limited in height to five or six times the least thickness. Under axial loading, the load divided by the cross-sectional area of the concrete should not exceed the allowable unit compressive stress for the concrete.

Axially loaded reinforced columns. Reinforced concrete columns are designed by ultimate-load theory. Two types of column are considered: short and long. Those whose length is three to ten times their least lateral dimension are called short columns.

For spiral-reinforced short columns, the American Concrete Institute code gives the formula for the allowable load in pounds shown in Eq. (1), where

$$P = 0.25 f'_c A_g + f_s A_s \tag{1}$$

f'_c is the 28-day compressive strength of a standard concrete test cylinder in pounds per square inch (lb/in.^{2}), A_g the gross area of section in square inches, f_s the allowable unit stress for steel in pounds per square inch, and A_s the area of reinforcing steel in square inches.

For a tied column, the allowable load is 85% of that for a spiral-reinforced column. Spiral-reinforced

columns are stronger because columns tend to fail by a lateral bursting of the concrete as the longitudinal bars bend outward, and spiral reinforcing is more effective than ties in restraining the concrete. If spirals are used, the longitudinal bars should be arranged in a circle.

For long columns, the allowable load is reduced from that permitted for short columns because of the possibility of buckling. When h/d, the ratio of unsupported length of column to least lateral dimension, is equal to or greater than 10, one should use the appropriate reduction factor given in Sec. 916 of the Building Code Requirements for Reinforced Concrete (ACI 318–63).

Spiral reinforcement is determined from Eq. (2),

$$p' \geq \frac{0.45(R-1)f_c'}{f_s'} \qquad (2)$$

where p' is the ratio of the volume of spiral reinforcement to the volume of the spiral core (out-to-out of the spiral), R is the ratio of gross area to core area, and f'_s is the yield-point stress of spiral steel, with a maximum of 60,000 lb/in.^{2}

Composite columns. A concrete compression member having a structural steel or cast iron core with a cross-sectional area not exceeding 20% of the gross area of the column is called a composite column. Spiral and longitudinal reinforcing also may be incorporated in the concrete.

The allowable load in pounds on a composite column is given by Eq. (3), where A_c is the net area of

$$P = 0.225A_cf_c' + f_rA_r + f_sA_s \qquad (3)$$

the concrete section (A_g-A_r-A_s), A_s is the area of longitudinal reinforcement other than the metal core, A_r is the cross-sectional area of the structural steel or cast iron core, f_r is the allowable unit stress for the core, and f_s is the allowable unit stress for the longitudinal reinforcement.

If the core area is 20% or more of the gross area and the concrete is at least 2 in. (5 cm) thick over all metal except rivet heads, the members is called a combination column. If the concrete is to be allowed to share the load with the core, the concrete must be reinforced; usually wire fabric is wrapped completely around the core.

Combined bending and axial load. Because columns are designed by ultimate-load theory and this theory has also been developed for beams, it is logical to design columns subjected to both bending and axial loads by ultimate-load theory. The ultimate load is given by Eqs. (4) and (5), where f'_c is the compressive

$$P_u = \phi(0.85f_c'ba + A_s'f_y - A_sf_s) \qquad (4)$$

$$P_ue = \phi[0.85f_c'ba(d - a/2) + A_s'f_y(d - d')] \qquad (5)$$

strength of concrete at 28 days; f_y is the yield strength reinforcement; f_s is the calculated stress in reinforcement when less than yield strength; A'_s is the area of compression reinforcement; A_s is the area of tension reinforcement; e is the eccentricity of axial load at the end of the member measured from centroid of tension reinforcement; a is the depth of equivalent rectangular compression-stress block; b is the width of compression flange; d is the distance from extreme compression fiber to centroid of tension steel; d' is the distance from extreme compression fiber to centroid of compression steel; and ϕ is a capacity reduction factor. It equals 0.75 for spiral-reinforced members and 0.70 for tied members.

For strength reduction factors when length-thickness ratio exceeds 10, one should consult Sec. 916 of the Building Code Requirements for Reinforced Concrete (ACI 318–63). *See* CONCRETE SLAB; REINFORCED CONCRETE.

Frederick S. Merritt

Concrete slab

A shallow, reinforced-concrete structural member that is very wide compared with depth. Spanning between beams, girders, or columns, slabs are used for floors, roofs, and bridge decks. If they are cast integrally with beams or girders, they may be considered the top flange of those members and act with them as a T beam. *See* CONCRETE; CONCRETE BEAM.

One-way slab. A slab supported on four sides but with a much larger span in one direction than in the other may be assumed to be supported only along its long sides. It may be designed as a beam spanning in the short direction. For this purpose a 1-ft width can be chosen and the depth of slab and reinforcing determined for this unit.

Some steel is also placed in the long direction to resist temperature stresses and distribute concentrated loads. The area of the steel generally is at least 0.20% of the concrete area.

Two-way slab. A slab supported on four sides and with reinforcing steel perpendicular to all sides is called a two-way slab. Such slabs generally are designed by empirical methods. A two-way slab is divided into strips for design purposes.

The American Concrete Institute (ACI) code recommends the following design method: Divide the slab in both directions into a column strip and middle strip (**Fig. 1**). If the ratio S/L of the short span to the long span is equal to or greater than 0.5, the width

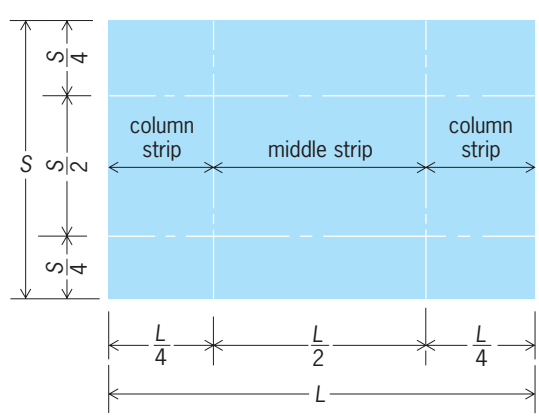

Fig. 1. A two-way slab in strips for design purposes.

of the middle strip extending in the short direction equals $L/2$, as shown. If S/L is less than 0.5, the width of the middle strip in the short direction is $L - S$; the remaining width is divided equally between the two column strips. However, when S/L is less than 0.5, most of the load would be carried in the short direction, and it would be desirable to design the slab as a one-way slab.

A table in the ACI code gives coefficients for calculation of the bending moments in the middle strip for different values of S/L and different types of panels. The moment in the column strip is assumed to be two-thirds that in the middle strip. The reinforcing steel area is determined from $A_s = M/f_s jd$, where M is the bending moment, f_s the tensile unit stress in longitudinal reinforcement, and j the ratio of lever arm of resisting couple to d, the depth from compressive surface of beam or slab to center of longitudinal tension reinforcement.

In the design of the beams that support a two-way slab along its sides, the load is assumed to be uniform over the slab and distributed to the beams as shown in **Fig. 2**. If w is the slab load in lb/ft^2, moments in the beams may be approximated by assuming the beams loaded with the following equivalent uniform loads: for the short span, $wS/3$; for the long span, $wS(3 - S^2/L^2)/6$.

Flat slabs. When a slab is supported directly on columns, without beams and girders, it is called a flat plate or flat slab (**Fig. 3**).

A flat plate generally is of uniform thickness throughout. Usually, the columns also have constant dimensions throughout the story height below the slab (excluding the portion forming capitals, which are discussed later). In designing such construction, it is necessary to investigate shear and diagonal tension in the vicinity of the columns; frequently, reinforcements which are known as shear heads are embedded at the column tops.

In flat-slab construction it is customary to flare out the columns at the top to form capitals so as to give the slab-column junction greater rigidity. The capital is usually sloped at 45°. For exterior columns the capital is sometimes only a bracket that projects inward.

Flat slabs generally are thickened in the region around the columns. The thickened portion, called a drop panel, may be extended until it reaches from

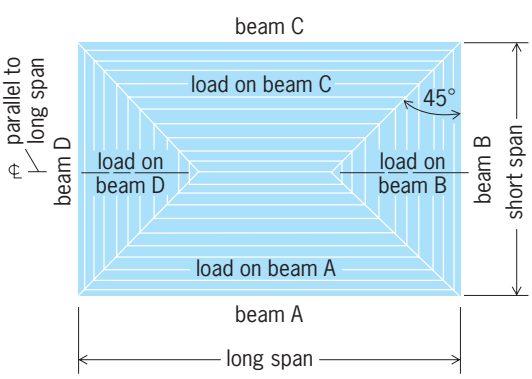

Fig. 2. Load distribution assumed for beams supporting a uniformly loaded two-way slab.

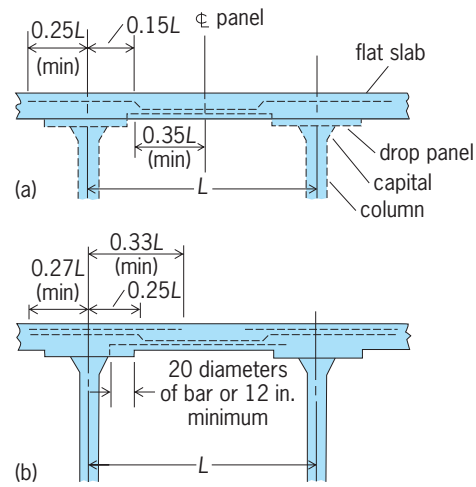

Fig. 3. Details of a flat slab drawn in cross section at two points. (*a*) Middle strip. (*b*) Column strip. *L* is the center-to-center distance between columns.

column to column, forming shallow beams and giving the effect of a paneled ceiling.

Although thicker and more heavily reinforced than slabs in beam-and-girder construction, flat slabs are advantageous because they offer no obstruction to passage of light (as beam construction does); savings in story height and in the simpler formwork involved; less danger of collapse due to overload; and better fire protection with a sprinkler system because the spray is not obstructed by beams.

Flat slabs may be reinforced in several ways:

1. *Two-way system.* When the columns are arranged to form rectangular bays, the reinforcing steel may be placed in two directions, perpendicular to the column lines. Design is based on column strips and middle strips, similar to that for a two-way slab.

2. *Four-way system.* Column strip steel is similar to that for a two-way slab. But middle-strip reinforcing consists of diagonal bands of steel extending over the columns.

3. *Circumferential system.* Bars are placed in the top of the slab in concentric rings around the columns, also radially. Similar radial and circular bars are placed in the bottom of the slab in the central portion.

4. *Three-way system.* When columns are arranged so that lines joining them would form triangles, the reinforcing steel may be laid parallel to the column lines.

Design of flat slabs and flat plates is based on empirical formulas. The ACI code presents design rules that are widely used. *See* CONCRETE COLUMN; REINFORCED CONCRETE.

Frederick S. Merritt

Concretion

A loosely defined term used for a sedimentary mineral segregation that may range in size from inches to many feet. Concretions are usually distinguished from the sedimentary matrix enclosing them by a difference in mineralogy, color, hardness, and

weathering characteristics. Some concretions show definite sharp boundaries with the matrix, while others have gradational boundaries. Most concretions are composed dominantly of calcium carbonate, with or without an admixture of various amounts of silt, clay, or organic material. Less common are the clay-ironstone concretions characteristic of the Carboniferous coal measures in many parts of the world. The latter are mixtures of iron carbonate minerals and iron silicate minerals. Coal balls are calcareous concretions, found in or immediately above coal beds, in which there may be a high percentage of original plant organic matter, showing wonderfully preserved plant fossils in a noncompressed condition. Concretions are normally spherical or ellipsoidal; some are flattened to disklike shapes. Frequently a concretion is dumbbell-shaped, indicating that two separate concretionary centers have grown together. *See* COAL BALLS; SEDIMENTARY ROCKS.

Raymond Siever

Concurrent processing

The simultaneous execution of several interrelated computer programs. A sequential computer program consists of a series of instructions to be executed one after another. A concurrent program consists of several sequential programs to be executed in parallel. Each of the concurrently executing sequential programs is called a process. Process execution, although concurrent, is usually not independent. Processes may affect each other's behavior through shared data, shared resources, communication, and synchronization.

Concurrent programs can be executed in several ways. Multiprogramming systems have one processing unit and one memory bank. Concurrent process execution is simulated by randomly interleaving instructions of the sequential programs. All processes have access to a common pool of data. In contrast, multiprocessing systems have several processing units and one memory bank. Processes are executed in parallel on the separate processing units while sharing common data. In distributed systems, or computer networks, each process is executed on its own processor with its own memory bank. Interaction between processes occurs by transmission of data from one process to another along a communication channel. *See* DISTRIBUTED SYSTEMS (COMPUTERS); MULTIPROCESSING.

Uses and examples. Concurrent programs are inherently more complex to create and to execute than sequential programs. However, concurrent programs are often more appropriate than sequential programs for a particular computing environment. Additionally, concurrent programs often perform the same tasks much more efficiently than sequential programs.

Historically, one of the first uses of concurrent processing was in operating systems. The operating system of a computer is the software that provides an interface between the user, the computer, and input/output devices such as terminals, printers, and disk drives. If the computer is to support a multiuser environment, the operating system must employ concurrent programming techniques to allow several users to access the computer simultaneously. The operating system should also permit several input/output devices to be used simultaneously, again utilizing concurrent processing. *See* MULTIACCESS COMPUTER; OPERATING SYSTEM.

Concurrent programming is also used when several computers are joined in a network. An airline reservation system is one example of concurrent processing on a distributed network of computers. Agents at many different locations on different computers simultaneously book passenger reservations. The concurrent program controlling the system allows this simultaneity while ensuring that the parallel processes do not conflict, for example, that two agents do not book the same seat on the same flight for two different passengers. *See* LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

A simple example of a task that can be performed more efficiently by concurrent processing is a program to calculate the sum of a large list of numbers. A sequential program must iterate through each element of the list, accumulating the sum. In a concurrent program, several processes can simultaneously compute the sum of a subset of the list, after which these sums are added to produce the final total. Depending on the size of the list and the number of processes, the concurrent program may perform the computation in only a small fraction of the time required by the sequential program.

Process interaction. The processes of a concurrent program are usually required to interact and cooperate. An operating system must ensure that processes coordinate the use of shared resources such as printers, memory, and processors. An airline reservation system must ensure that conflicting reservations are not made at different sites. The processes of a concurrent program to sum a list of numbers must cooperate to combine their subtotals into a final total. Many tasks also require concurrent processes to synchronize certain actions. For example, a process controlling a sensor that produces data at a varying rate must be synchronized with a process that reads and analyzes these data.

In multiprogramming and multiprocessing systems, communication and synchronization are achieved through shared variables. Certain memory locations contain data to be read and modified by several processes, such as the number of disk drives available, the next process to send a file to the printer, the number of elements in a buffer, or the total of a sublist of numbers. Reading and writing of shared variable values must be synchronized to avoid undesirable situations. For example, one process may read a shared value in order to update it, and a second process may change the value before the first process is completed. The change in the second process is lost when the first process finally performs its update. This is known as the lost update problem, and it is prevented by ensuring that processes have the ability to gain exclusive access to shared variables for the duration of an operation.

Distributed systems have no shared memory; communication and synchronization are achieved by transmitting messages along shared communication links. Communication can be either synchronous or asynchronous. With synchronous communication, both the sending and receiving processes must be ready to communicate before a message is actually transmitted. With asynchronous communication, the sending process may transmit messages whether or not the receiving process is prepared to receive them. The sent messages are held in a queue until requested by the receiving process. Synchronous communication is used when coordination between processes may be required, as in a series of computers controlling an automated assembly line. Asynchronous communication is used when coordination between processes is not required, as in the transmission of electronic mail messages from one computer to another. *See* DATA COMMUNICATIONS; ELECTRONIC MAIL.

Creating concurrent programs. Concurrent programs can be created explicitly or implicitly. Explicit concurrent programs are written in a programming language designed for specifying processes to be executed concurrently. Implicit concurrent programs are created by a compiler that automatically translates programs written in a sequential programming language into programs with several components to be executed in parallel.

Many programming languages provide facilities for creating concurrent programs. These languages typically include a component for specifying sequential programs and a component for combining and coordinating several sequential programs into one concurrent program. Examples of such languages are Ada, Communicating Sequential Processes (CSP), Concurrent Pascal, Mesa, Modula, and Synchronizing Resources (SR).

Compilers that translate sequential programs into concurrent programs are called parallelizing compilers. They employ a technique called dependency analysis to identify regions of a sequential program that are independent and may therefore be executed in parallel. The regions may be as large as several hundred lines of code or as small as a single arithmetic operation. The object of a parallelizing compiler is to introduce as much parallelism as possible in order to increase the speed of the program. *See* PROGRAMMING LANGUAGES. Jennifer Widom

Bibliography. M. Ben-Ari, *Principles of Concurrent and Distributed Programming*, 1990; A. Burns and G. Davies, *Concurrent Programming*, 1993; T. J. Fountain, *Parallel Computing: Principles and Practice*, 1994; V. Kumar, *Introduction to Parallel Computing*, 1994; A. Silberschatz and P. G. Galvin, *Operating System Concepts*, 5th ed., 1998.

Concussion

Mild traumatic brain injury, due to externally inflicted trauma. While such injury has many symptoms, there is often no physical evidence of damage to the brain. It is clinically manifested by at least one of the following:

Loss of consciousness for 30 minutes or less.

Posttraumatic amnesia for less than 24 hours.

An initial score of 13–15 on the Glasgow Coma Scale, which measures eye, verbal, and motor responses. [Scores can range from 3 (deep coma) to 15 (fully awake).]

Alteration in mental state at time of injury (for example, feeling dazed, disoriented, or confused).

Individuals with concussion do not experience focal neurological deficits (that is, they have no paralysis or sensory loss) and imaging studies (such as computerized tomography and magnetic resonance imaging usually do not show evidence of brain damage. However, after concussion the afflicted persons may experience:

Cognitive deficits (for example, difficulties with concentration and attention, a decrease in information processing speed, and deficits in executive functioning and memory).

Affective disorders (for example, irritability, anxiety, and depression).

Somatic complaints (such as headaches, dizziness, nausea, insomnia, fatigue, and sensory disturbances).

Most of these symptoms disappear within weeks or months, but approximately 15% of individuals continue to experience such symptoms, a condition often referred to as postconcussion syndrome. Actual causes and influencing factors of postconcussion syndrome are much debated but may involve premorbid (preexisting) conditions, psychological reactions, age, mechanism of injury, as well as unidentified organic (physical) disorders.

Since most symptoms of concussion resolve spontaneously and no specific treatment has been validated scientifically, management initially is "wait and see." It is prudent for the individual to promptly undergo a brief neuropsychological assessment and receive appropriate rest. If difficulties persist beyond 6 months, a comprehensive neuropsychological evaluation is indicated. Individuals with persistent complaints months after the injury may benefit from psychotherapy, pain management, cognitive rehabilitation, physical therapy, judicious use of medications, professional counseling, and group support. *See* BRAIN; COGNITION. Kristjan T. Ragnarsson

Bibliography. M. Aubry et al., *Summary and Agreement Statement of the 1st International Conference on Concussion and Sport, Vienna, 2001, The Physician and Sports Medicine*, vol. 30, no. 3, February 2002; T. Kay et al., Definition of mild traumatic brain injury, *J. Head Trauma Rehab.*, 8:86–87, 1993; T. W. McAllister, Mild traumatic brain injury and the post-concussive syndrome, in J. M. Silver, S. C. Yudofsky, and R. E. Hales (eds.), *Neuropsychiatry of Traumatic Brain Injury*, pp. 357–392, American Psychiatric Press, Washington, DC, 1994; J. Whyte et al., Rehabilitation issues in traumatic brain injury, in J. A. DeLisa (ed.), *Physical Medicine*

and Rehabilitation Principles and Practice, pp. 1677–1713, Lippincott Williams & Wilkins, 2005; N. Zasler, Neuromedical diagnosis and management of post-concussive disorders, in L. Horn and N. Zasler (eds.), *Medical Rehabilitation of Traumatic Brain Injury*, pp. 133–170, Henley and Belfuss, Philadelphia, 1996.

Condensation

A phase-change process in which vapor converts into liquid when the temperature of the vapor is reduced below the saturation temperature corresponding to the pressure in the vapor. For a pure vapor this pressure is the total pressure, whereas in a mixture of a vapor and a noncondensable gas it is the partial pressure of the vapor. Sustaining the process of condensation on a cold surface in a steady state requires cooling of the surface by external means. Condensation is an efficient heat transfer process and is utilized in various industrial applications. Condensation of vapor on a cold surface can be classified as film-wise or drop-wise. Direct-contact condensation refers to condensation of vapor (bubbles or a vapor stream) in a liquid or condensation on liquid droplets entrained in the vapor. If vapor temperature falls below its saturation temperature, condensation can occur in the bulk vapor. This phenomenon is called homogeneous condensation (formation of fog) and is facilitated by foreign particles such as dust. *See* GAS; HEAT TRANSFER.

Film condensation. A thin film of liquid forms upon condensation of vapor on a cold surface that is well wetted by the condensate (see **illus.**). The liquid film flows downward as a result of gravity. The temperature drop between the vapor-liquid interface of the film and the surface of the solid on which vapor condenses is representative of the condensation resistance, which is inversely proportional to the condensation heat transfer coefficient. The film is initially laminar in the direction of flow. As the mass flow rate through the film and, in turn, the thickness of the film increase due to addition of condensate, ripples start to form on the free surface of the film. The ripples form as a result of capillary and inertia forces, and generally begin at a Reynolds number of about 30, where the Reynolds number is based on the average velocity of the condensate in the film, four times the thickness of the film, and the kinematic viscosity of the condensate. The film becomes turbulent when the Reynolds number attains a value of about 1800. *See* REYNOLDS NUMBER; SURFACE TENSION; TURBULENT FLOW.

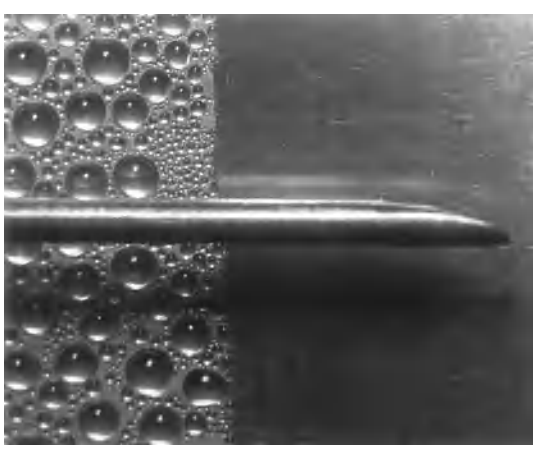

Steam at atmospheric pressure condensing on a vertical copper surface. Film condensation is visible on the right side, and dropwise condensation in the presence of a promoter is visible on the left side. The horizontal tube is a thermocouple. (*J. F. Welch and J. W. Westwater, Department of Chemical Engineering, University of Illinois, Urbana*)

The earliest work on laminar film condensation was performed by W. Nusselt in 1916. He assumed that the vapor was pure, stagnant, and at its saturation temperature; that inertia and convective contributions in the film were small; and that thermophysical properties of the condensate film could be evaluated at the mean temperature of the film. He then obtained the equation below for the dimensionless average heat transfer coefficient, or Nusselt number.

$$\mathrm{Nu} = \frac{\bar{h}L}{k_\ell} = C\left[\frac{\rho_\ell g(\rho_\ell - \rho_v)h_{fg}L^3}{\mu_\ell k_\ell (T_{\mathrm{sat}} - T_W)}\right]^{1/4}$$

Here, $\bar{h}$ is the average heat transfer coefficient for film condensation, L is the plate height over which laminar film condensation occurs, k_ℓ is the thermal conductivity of the liquid, ρ_ℓ is the density of liquid, ρ_v is the density of vapor, h_{fg} is the latent heat of vaporization, μ_ℓ is the molecular viscosity, T_{sat} is the saturation temperature, and T_W is the surface temperature. The latent heat of vaporization and vapor density correspond to the saturation condition in the vapor. From the form of the equation, it follows that average heat transfer coefficients are higher for small bodies (small L). For condensation of steam, typical values of the heat transfer coefficient $\bar{h}$ are of the order of 5000 W/(m^2 · K). *See* CONDUCTION (HEAT); EVAPORATION; VISCOSITY.

The dimensionless constant C in the equation above is of order unity and varies with geometry. For a vertical plate, $C = 0.943$. Other geometries, such as tubes and spheres, can be analyzed. The latent heat term h_{fg} is also modified by convection and subcooling in the film. If ripples form or turbulence ensues in the film, new formulas and correlations are needed.

Bulk motion of vapor in the direction of the flow of the condensate film tends to increase the heat transfer coefficient from that given by the equation above. If the vapor is at a temperature higher than the saturation temperature, the vapor must be cooled to saturation temperature before it is condensed. As such, the vapor superheat tends to increase the condensation heat transfer rate; however, the condensate mass flow rate decreases. The presence of noncondensable gases in the vapor can substantially reduce the condensation heat transfer rate because of the additional resistance imposed by the diffusion of the vapor to the interface. In the

presence of noncondensables, the mass flow rate of the condensate also decreases. In refrigeration and air-conditioning systems, film condensation occurs inside tubes. Since vapor and condensate flow simultaneously through the tube and the fractional cross-sectional area occupied by each phase changes in the direction of the flow, different flow configurations or regimes are encountered which in turn influence the condensation heat transfer coefficients.

Dropwise condensation. On surfaces that are not well wetted, vapor may condense in the form of droplets (see illus.). The droplets form on imperfections such as cavities, dents, and cracks on the surface. The droplets of 10–100 μm diameter contribute most to the heat transfer rate. As a droplet grows to a size that can roll down the surface because of gravity, it wipes the surface of the droplets in its path. In the wake behind the large droplet, numerous smaller droplets form and the process repeats. The heat transfer coefficients with dropwise condensation can be one to two orders of magnitude greater than that for film condensation. Thus, dropwise condensation is preferable to film condensation. Attempts have been made to deposit an ultrathin coating of a nonwetting material such as Teflon on the condensing surface to promote dropwise condensation. However, in heat exchange equipment, because of the buildup of foreign material on the surface or degradation of the promoter layer, dropwise condensation shifts with time to film condensation. Therefore, equipment is designed on the basis of film condensation.

Direct contact condensation. Condensation of vapor bubbles in a host liquid and condensation on droplets entrained in vapor are also very efficient heat transfer processes, especially when the vapor-liquid interface oscillates. The magnitude of the heat transfer coefficient depends on several variables such as the size of the bubble or droplet, liquid subcooling, the relative velocity between the continuous and the discontinuous phase, and the appropriate thermophysical properties of the two phases. V. K. Dhir

Bibliography. Y. Cengel, *Heat Transfer*, McGraw-Hill, 1998; J. G. Collier and J. R. Thome, *Convective Boiling and Condensation*, 3d ed., Oxford, 1996; F. P. Incropera and D. P. Dewitt, *Fundamentals of Heat and Mass Transfer*, 4th ed., Wiley, 1996.

Conditioned reflex

A learned response performed by a trained animal to a signal that was previously associated with an event of consequence for that animal. Conditioned reflex (CR) was first used by the Russian physiologist I. P. Pavlov to denote the criterion measure of a behavioral element of learning, that is, a new association between the signal and the consequential event, referred to as the conditioned stimulus (CS) and unconditioned stimulus (US), respectively. In Pavlov's classic experiment, the conditioned stimulus was a bell and the unconditioned stimulus was sour fluid delivered into the mouth of a dog restrained by harness; the conditioned stimulus was followed by the unconditioned stimulus regardless of the dog's response. After training, the conditioned reflex was manifested when the dog salivated to the sound of the bell.

Classical conditioning. Ideally, certain conditions must be met to demonstrate the establishment of a conditioned reflex according to Pavlov's classical conditioning method. Before conditioning, the bell conditioned stimulus should attract the dog's attention or elicit the orienting reflex (OR), but it should not elicit salivation, the response to be conditioned. That response should be specifically and reflexively elicited by the sour unconditioned stimulus, thus establishing its unlearned or unconditioned status. After conditioned pairings of the conditioned stimulus and the unconditioned stimulus, salivation is manifested prior to the delivery of the sour unconditioned stimulus. Salivation in response to the auditory conditional stimulus is now a "psychic secretion" or the conditioned reflex.

This simple conditioned reflex was important to Pavlov because it presumably demonstrated concretely a new temporary connection in the brain, a neural connection established by training between the auditory apparatus and salivary mechanisms. It provided him with an opportunity to study the psyche objectively. "Here we have exact and constant facts," Pavlov said, "facts that seem to imply intelligence." He had already won the Nobel Prize in Physiology or Medicine for his studies of digestion and its associated secretory reflexes. He knew that the food stimuli contracted oral receptors, and how the impulses were transmitted via centripetal nerves to a specific area of the brainstem. He knew that this reflexive mechanism controlled salivation by centrifugal nerves, so as to match the quality and quantity of saliva precisely with the nature of oral stimulation in accordance with the "wisdom of the body." He understood how the auditory conditioned stimulus might control the head-turning orienting reflex in the dog. But the final mystery was how the auditory conditioned stimulus gained control over the salivary reflex through conditioning or, in other words, how the bell became a stimulus substitute for food. Pavlov's goal was to elucidate these new pathways of control in the brain, and that task was not completed. But to this day, Pavlov's methods provide important guidelines for basic research upon brain mechanisms in learning and memory.

By 1909, other Russian workers had extended the study of the conditioned reflex to other stimulus-reflex combinations in dogs and humans, and since that time scientists all over the world have paired a vast array of stimuli with an enormous repertoire of reflexes to test conditioned reflexes in representative species of almost all phyla, classes, and orders of animals. As a result, classical conditioning is now considered a general biological or psychobiological phenomenon which promotes adaptive functioning in a wide variety of physiological systems in various phylogenetic settings.

Selective learning and homeostatic feedback. The unconditioned stimulus is usually contact with a object or event that is significant for survival, evoking pleasure or pain. This is only a rough guide because some "pleasures" (for example, drugs and alcohol) can be ultimately detrimental. Food, the most popular unconditioned stimulus, is vital for survival. Consumption is followed by digestion which provides feedback (FB) to adjust the palatability of food commensurate with its homeostatic effect. If food is nutritious, it tastes "good" next time. If food is followed by nausea induced by injection, conditioned food aversion is acquired and the food tastes "bad." Therefore the entire conditioning sequence is composed of two different associative processes denoted as CS-US-FB. The first CS-US link and the second US-FB link differ in many ways. For rapid one-trial learning: (1) The CU-US interval must be a matter of seconds, but the US-FB interval can be a matter of hours. (2) The unconditioned stimulus must be presented to an alert subject, but the feedback is completely effective when delivered to a subject asleep under anesthesia. (3) After CS-US training, the subject gains a precise appreciation of both stimuli and their association in time and space and modifies its behavior accordingly. But after conditioned food aversion training the hedonic value of the unconditional stimulus is modified even in the absence of any memory of its association with the nauseous feedback. The subject often approaches the unconditioned stimulus again and rejects it in surprise and disgust.

Furthermore, while many different conditioned stimuli can easily be linked with food unconditioned stimulus, there is a strong selective affinity of the food feedback for taste. In the complete absence of taste, even an aversion to odor paired with poison is absent or weak. With repeated trials, the aversion spreads to other cues associated with feeding. In descending order of effectiveness, these cues are oral stimuli other than taste or odor, visual and tactual aspects of the food, visual and tactual aspects of the food locus, and auditory stimulation paired with food. All these cues are potentiated when accompanied by taste. Natural selection resulted in this specialized US-FB learning system to protect the gut from natural toxins. Charles Darwin was concerned with the mechanisms by which natural toxins protected poisonous insect larvae from predators. Subsequent research revealed that the predators acquired a conditional food aversion for such insects and their mimics; thus the US-FB association has been called Darwinian conditioning.

A brief electric shock impinging on the skin, also a well-known unconditioned stimulus, engages another specialized CS-US-FB system designed by natural selection to protect the skin from predatory attacks. Since the attack is usually swift and preceded by vibrations in the substrate, an immediate noise-shock combination is about as effective for conditioned fear as a delayed taste-toxin combination is for conditioned food aversion. Conversely, taste-shock and noise-toxin are relatively ineffective combinations for rapid conditioning because these stimuli impinge upon two different integrative systems. The shock unconditioned stimulus is followed by an endogenous analgesic feedback that raises the pain threshold. This feedback action is of long duration and is also effective in subjects sleeping under anesthesia. US-FB relationships have not been elucidated for other unconditoned stimulus events, but there is evidence that they exist. Core temperature of the body is apparently feedback for thermal skin receptors. When core temperature is low, a warm unconditioned stimulus is pleasant to the skin. When the core is warmed, the same unconditioned stimulus is unpleasant. In mating behavior of mammals, reproductive success seems to make odors associated with mates more attractive.

Instrumental conditioning. In this procedure the experimenter selects a response (R) emitted by the animal and reinforces it with an unconditioned stimulus. R-US methodology is favored by B. F. Skinner. For example, consider an operant conditioning situation in which a pigeon is placed in the Skinner box with a key or target connected to a food delivery mechanism. Whenever the key illuminates, the pigeon can obtain food by depressing it. Skinner's procedure differs from Pavlov's situation in that the subject is unrestrained, and food is contingent upon depression of the key whenever the light is on. No key press, no food; thus the birds appeared to be working for the pay of food in economic fashion. Skinner, avoiding physiological mechanisms, defined his conditioned reflex as a depression of the key and called it an operant; his pigeon was operating the key to get food. However, high-speed photographs indicated that the pigeon pecked at the target in the same way it pecked at grain. When depressing the lighted key to get water, the pigeon worked the key with the same action it used to drink water. For other birds, the key was disconnected from the food delivery mechanism and seeds were delivered regardless of the pigeon's response, a procedure called autoshaping. These pigeons also learned to peck at the lighted key as if it were a substitute for food. The pigeon's peck, like the dog's salivation, is a Pavlovian conditioned reflex that is characteristic of the consummatory phase of learning.

True instrumental learning is displayed by subjects long before they gain their consummatory objective. For example, E. L. Thorndike confined hungry cats in a puzzle box and placed the food (unconditioned stimulus) outside. Manipulation of a trigger (the response) released the cat, giving access to food. Another procedure is to place a hungry rat in the starting box and to provide food at the end of a maze with cul-de-sacs along the way. With repeated trials in such apparatus, animals will reduce extraneous responses and decrease time from start to finish. With prolonged training under controlled conditions, subjects can combine CS-US and R-US units into long instrumental sequences. However, behavioral units and sequences do not provide a complete answer to the question "What is learned?" Laboratory

studies and detailed observations under natural conditions indicate that animals often forego satisfaction of needs to spontaneously explore their surroundings, acquiring memorial representation or a cognitive map of the spatial and temporal relationships among many stimuli. With this cognitive map well in hand, they will exhibit novel shortcuts, detours, strategies, and expectations, modifying instrumental means to achieve desired ends. After consumption, these incentives to action are automatically reevaluted in accordance with their homeostatic utility. *See* COGNITION; MEMORY. John Garcia

Bibliography. D. L. Alkon and J. Farley (eds.), *Primary Neural Substrates of Learning and Behavioral Change*, 1984; N. Bravemen and P. Bronstein (eds.), *Experimental Assessments and Clinical Applications of Conditioned Taste Aversions*, 1985; M. Cabanac, Sensory pleasure, *Quart. Rev. Biol.*, 54:1–29, 1979; J. Garcia and M. D. Holder, Time, space and value, *Human Neurobiol.*, 4:81–89, 1985; R. A. Hinde and J. Stevenson-Hinde (eds.), *Constraints on Learning: Limitations and Predispositions*, 1973; N. Mackintosh, *The Psychology of Animal Learning*, 1974; A. Trevor and L.-G. Nilsson (eds.), *Aversion, Avoidance, and Anxiety: Perspectives in Aversively Motivated Behavior*, 1989.

Conductance

The real part of the admittance of an alternating-current circuit. The admittance Y of an alternating-current circuit is a complex number given by Eq. (1). The real part G is the conductance. The units

$$Y = G + jB \tag{1}$$

of conductance, like those of admittance, are called siemens or mhos. Conductance is a positive quantity. The conductance of a resistor R is given by Eq. (2).

$$G = \frac{1}{R} \tag{2}$$

In general the conductance of a circuit may depend on the capacitors and inductors in the circuit as well as on the resistors. For example, the circuit in the **illustration** has impedance at frequency ω given by Eq. (3) and admittance given by Eq. (4), so that the conductance, given by Eq. (5), depends on the inductance L as well as the resistance R. *See* ADMITTANCE; ALTERNATING-CURRENT CIRCUIT THEORY; ELECTRICAL IMPEDANCE. J. O. Scanlan

$$Z = R + jL\omega \tag{3}$$

$$Y = \frac{1}{R + jL\omega} \tag{4}$$

$$G = \frac{R}{R^2 + L^2\omega^2} \tag{5}$$

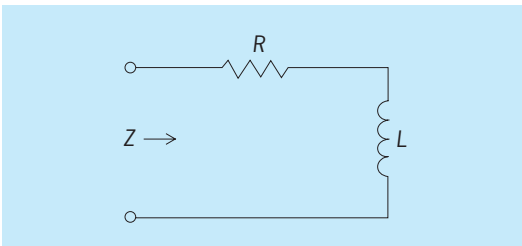

Circuit with a resistor and inductor in series.

Conduction (electricity)

The passage of electric charges due to a force exerted on them by an electric field. Conductivity is the measure of the ability of a conductor to carry electric current; it is defined as the ratio of the amount of charge passing through unit area of the conductor (perpendicular to the current direction) per second divided by the electric field intensity (the force on a unit charge). Conductivity is the reciprocal of resistivity and is therefore commonly expressed in units of siemens per meter, abbreviated S/m. *See* ELECTRICAL RESISTIVITY.

The magnitude of the conductivity of a material or system is determined by two properties: the number n of its charges in unit volume that are free to move in the field, and how effectively the field can move them. Acted on by the field, which accelerates them, and by the resistance of the material to their motion, the free charges achieve a drift velocity that is proportional to the electric field intensity. The mobility μ of the charges is defined as the ratio of this drift velocity to the field. From the above definition of the conductivity σ, it follows directly that $\sigma = ne\mu$, where e is the charge of a free carrier and n is the number of free carriers per unit volume.

Electron and hole conduction. In metals and semiconductors (such as silicon, of which transistors are made) the charges that are responsible for current are free electrons and holes (which, as missing electrons, act like positive charges). These are electrons or holes not bound to any particular atom and therefore able to move freely in the field. Conductivity due to electrons is known as n-type conductivity; that due to holes is known as p-type. *See* HOLE STATES IN SOLIDS; SEMICONDUCTOR.

The conductivity of metals is much higher than that of semiconductors because they have many more free electrons or holes. The free electrons or holes come from the metal atoms. Silver is the highest-conductivity metal, with a room-temperature conductivity of 7×10^7 S/m. Copper has almost as high a room-temperature conductivity, 5.9×10^7 S/m. Semiconductors differ from metals in two important respects. First, the semiconductor atoms do not contribute free electrons or holes unless thermally excited, and second, free electrons or holes can also arise from impurities or defects. The latter process may also require thermal excitation, but the temperature rise needed is smaller than that

for exciting electrons and holes from the semiconductor atoms. These two differences cause the conductivity of semiconductors, in contrast to that of metals, to be quite dependent on temperature and on their purity. Pure silicon, for example, has a room-temperature conductivity of 3×10^{-4} S/m. If impurities such as aluminum or arsenic are added, a process called doping, the conductivity of silicon can be raised to 10^5 S/m and even higher. Gallium arsenide, of which laser diodes are made, when pure has a room-temperature conductivity of 2×10^{-7} S/m, but it can also be doped, for example, with silicon, to 10^5 S/m.

An exception to some of the rules stated above has been found in conjugated polymers. Polyacetylene, for example, although a semiconductor with extremely high resistance when undoped, can be doped so heavily with certain nonmetallic impurities (iodine, for example) that it attains a conductivity comparable to that of copper. *See* ORGANIC CONDUCTOR.

In metals, although the number of free carriers does not vary with temperature, an increase in temperature decreases conductivity. The reason is that increasing temperature causes the lattice atoms to vibrate more strongly, impeding the motion of the free carriers in the field. This effect also occurs in semiconductors, but the increase in number of free carriers with temperature is usually a stronger effect. At low temperatures the thermal vibrations are weak, and the impediment to the motion of free carriers in the field comes from imperfections and impurities, which in metals usually does not vary with temperature. At the lowest temperatures, close to absolute zero, certain metals become superconductors, possessing infinite conductivity. *See* ELECTRICAL CONDUCTIVITY OF METALS; SUPERCONDUCTIVITY.

Although there is no material that does not conduct electricity at all, that is, behaves as a perfect insulator, there are materials that are very poor conductors. Examples are ceramics, with conductivities of 10^{-12} to 10^{-14} S/m; amorphous selenium, with a conductivity of 10^{-14} S/m; and paraffin, with 3×10^{-17} S/m. *See* ELECTRIC INSULATOR.

Ionic conduction. Electrolytes conduct electricity by means of the positive and negative ions in solution. In ionic crystals, conduction may also take place by the motion of ions. This motion is much affected by the presence of lattice defects such as interstitial ions, vacancies, and foreign ions. *See* ELECTROLYTIC CONDUCTANCE; IONIC CRYSTALS.

Ionic conduction can also take place in a gas by using a strong electric field to ionize the gas molecules. If enough ions are formed, there may be a spark. *See* ELECTRIC SPARK; ELECTRICAL CONDUCTION IN GASES.

Vacuum conductivity. Electric current can flow through an evacuated region if electrons or ions are supplied. In a vacuum tube the current carriers are electrons emitted by a heated filament. The conductivity is low because only a small number of electrons can be "boiled off" at the normal temperatures of electron-emitting filaments. *See* ELECTRON EMISSION; ELECTRON MOTION IN VACUUM; VACUUM TUBE.

Esther M. Conwell

Bibliography. J. Beynon, *Conduction of Electricity through Gases*, 1972; E. M. Conwell and H. A. Mizes, Conjugated polymer semiconductors: an introduction, in P. T. Landsberg (ed.), *Handbook on Semiconductors*, vol. 1, 1992; D. Inman and J. Lovering (eds.), *Ionic Liquids*, 1981; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; A. L. Laskar and S. Chandra, *Superionic Solids and Solid Electrolytes: Recent Trends*, 1989.

Conduction (heat)

The flow of thermal energy through a substance from a higher- to a lower-temperature region. Heat conduction occurs by atomic or molecular interactions. Conduction is one of the three basic methods of heat transfer, the other two being convection and radiation. *See* CONVECTION (HEAT); HEAT RADIATION; HEAT TRANSFER.

Steady-state conduction is said to exist when the temperature at all locations in a substance is constant with time, as in the case of heat flow through a uniform wall. Examples of essentially pure transient or periodic heat conduction and simple or complex combinations of the two are encountered in the heat-treating of metals, air conditioning, food processing, and the pouring and curing of large concrete structures. Also, the daily and yearly temperature variations near the surface of the Earth can be predicted reasonably well by assuming a simple sinusoidal temperature variation at the surface and treating the Earth as a semi-infinite solid. The widespread importance of transient heat flow in particular has stimulated the development of a large variety of analytical solutions to many problems. The use of many of these has been facilitated by presentation in graphical form.

For an example of the conduction process, consider a gas such as nitrogen which normally consists of diatomic molecules. The temperature at any location can be interpreted as a quantitative specification of the mean kinetic and potential energy stored in the molecules or atoms at this location. This stored energy will be partly kinetic because of the random translational and rotational velocities of the molecules, partly potential because of internal vibrations, and partly ionic if the temperature (energy) level is high enough to cause dissociation. The flow of energy results from the random travel of high-temperature molecules into low-temperature regions and vice versa. In colliding with molecules in the low-temperature region, the high-temperature molecules give up some of their energy. The reverse occurs in the high-temperature region. These processes take place almost instantaneously in infinitesimal distances, the result being a quasi-equilibrium state with energy transfer. The mechanism for energy flow in liquids and solids is similar to that in gases in principle, but different in detail.

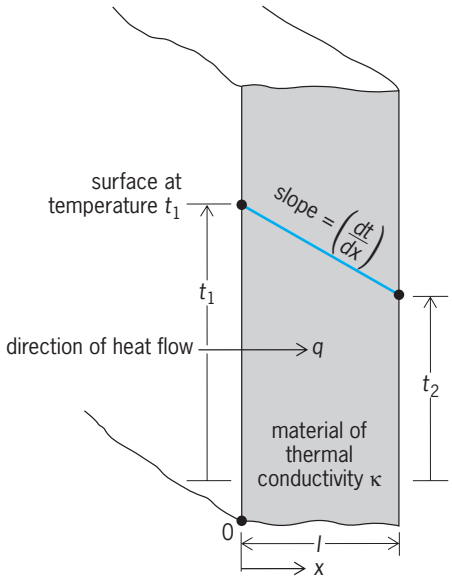

Fig. 1. Heat flow by conduction.

Fourier equation. The mathematical theory as well as the practical calculation of heat conduction is based on a macroscopic interpretation, as contrasted to the basic microscopic mechanism just described. From a physical point of view, it is reasoned that the steady heat flow from a surface (**Fig. 1**) at temperature t_1 to a parallel surface at t_2 is directly proportional to $(t_1 - t_2)$, the area A normal to the direction of flow, and the time of flow τ, and inversely proportional to the distance l between the two planes. These factors are modified by a coefficient κ accounting for the heat-conducting nature of the particular substance between the two planes. Thus, the heat flow Q (in British thermal units, for example) is given by Eq. (1). In terms of the time rate of flow $q = Q/\tau$

$$Q = \kappa A \frac{t_1 - t_2}{l} \tau \tag{1}$$

through an infinitesimally thin layer dx, in which the temperature change is dt, this becomes Eq. (2). The

$$q = -\kappa A \frac{dt}{dx} \tag{2}$$

minus sign is conventionally included to make q positive when heat flows in the increasing direction of x, since dt/dx is then negative. Although this equation was first proposed by J. Biot, it is named after J. Fourier in honor of Fourier's extensive contributions to the theory of heat conduction.

Thermal conductivity. The coefficient κ in Eqs. (1) and (2), called the thermal conductivity, is an important property of matter. It accounts for the heat–conducting ability of a substance, and depends not only on the particular substance involved, but also on the state of that substance. The Fourier equation is essentially a definition of κ which (Fig. 1) can be interpreted as the rate of heat flow per unit of area normal to the direction of flow when a unit temperature difference exists in unit length. Thus from Eq. (2) is derived Eq. (3), where κ is seen to have

$$\kappa = \frac{-q/A}{dt/dx} = \frac{q/A}{(t_1 - t_2)/l} \tag{3}$$

the dimensions of a heat rate per unit area and per unit of temperature gradient. In the cgs system, it can be expressed in cal/(s) · (cm^2)(°C/cm), which is equivalent to cal/(s)(cm)(°C). In engineering, the units most frequently used are Btu/(h)(ft)(°F).

Considerable progress has been made in the interpretation of thermal properties from theories of matter. This is particularly true for gases, where theory involving intermolecular forces has yielded very accurate results. The process of heat conduction in liquids is believed to be similar to that of sound transmission. In dielectric solids, energy is transmitted primarily by means of waves traveling through the atomic lattice; in metals, the electrons behave like an electron gas and provide for energy transfer as well as electrical conduction. This is the basis of the Widemann-Franz law, which states that $\kappa/\sigma T =$ constant (σ is the electrical conductivity and T the absolute temperature). *See* INTERMOLECULAR FORCES; KINETIC THEORY OF MATTER; THERMAL CONDUCTION IN SOLIDS.

For materials occurring as crystalline or amorphous solids, the general trend of κ at atmospheric pressure throughout the three physical states is as shown in **Fig. 2**. The numerical value at the maximum, which occurs near absolute zero in crystalline substances, is comparatively high. For example, the κ of a copper crystal at 20 K has been found to be 7050 Btu/(h)(ft)(°F)—more than 30 times its value at room temperature. Thermal conductivity of solids is discussed in a later section.

Because of the complexity and incomplete understanding of the mechanisms responsible for heat conduction, values of κ are usually determined experimentally. Results for typical gases, liquids, and solids in appropriate temperature ranges are shown in **Fig. 3**. The effect of pressure is significant primarily in gases.

Differential equation of conduction. The evidence of heat flow by conduction through a substance is the variation of the temperature with location and time.

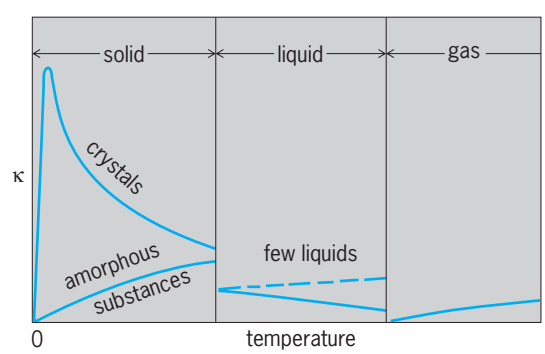

Fig. 2. General variation of thermal conductivity with temperature throughout the three physical states. (*After L. S. Kowalczyk, Trans. ASME, 77:1021–1035, 1955*)

If the temperature as a function of the space coordinates and time is known or can be determined, the heat flow at any location and in any direction can be specified by appropriate differentiation. A given problem is normally attacked by solving the differential equation governing the temperature distribution in a homogeneous substance and making this solution fit the prescribed initial or boundary conditions. This differential equation, essentially an expression of the first law of thermodynamics applied to the heat flow, is derived by making a heat balance on an elemental volume in a medium (**Fig. 4**).

Considering first the x direction, the net heat flow into the element in time $\Delta\tau$ is the difference between the flowing in on the left minus that flowing out on the right; or, applying Eq. (1), Eq. (4) is obtained.

$$\Delta Q_x = -\kappa\ \Delta y\ \Delta z \left(\frac{\partial t}{\partial x}\right)_x \Delta\tau - \left[-\kappa\ \Delta y\ \Delta z \left(\frac{\partial t}{\partial x}\right)_{x+\Delta x} \Delta\tau\right] \quad (4)$$

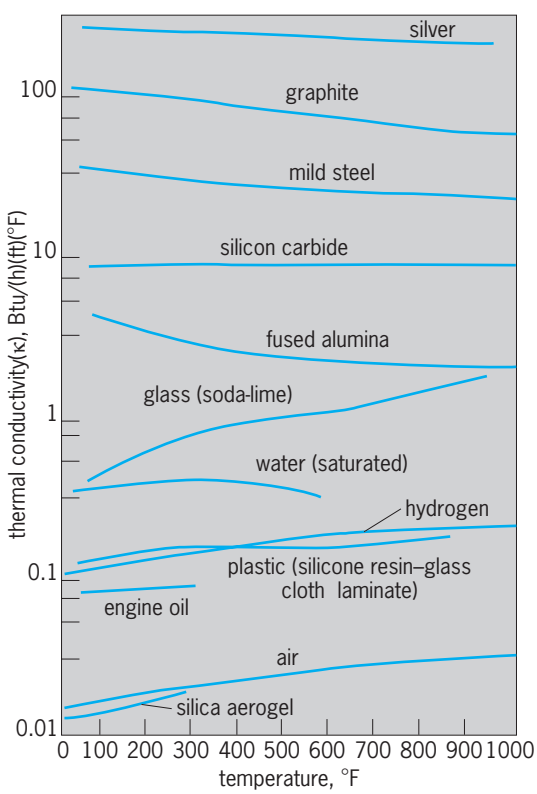

Fig. 3. Thermal conductivities of some typical examples of gases, liquids, and solids. °C = (°F − 32)/1.8.

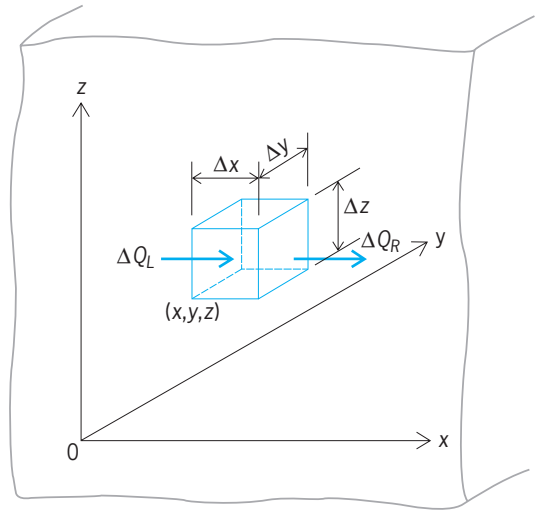

Fig. 4. Diagram of the heat flow through an elemental volume in a homogeneous medium.

tained. Accounting for the variation of κ with temperature is extremely difficult and may make the analytical solution of a problem impossible. Because of this, it is customary to use an appropriate average value which is regarded as constant. Equation (4) can then be rearranged to read as Eq. (5*a*), which, as $\Delta x \to 0$, becomes Eq. (5*b*). Similar expressions apply for the $y$and z directions. Heat generated within the

$$\Delta Q_x = \kappa\ \Delta y\ \Delta z\ \Delta x \frac{\left(\frac{\partial t}{\partial x}\right)_{x+\Delta x} - \left(\frac{\partial t}{\partial x}\right)_x}{\Delta x} \Delta\tau \quad (5a)$$

$$\Delta Q_x = \kappa\ \Delta x\ \Delta y\ \Delta z\ \Delta\tau \frac{\partial^2 t}{\partial x^2} \quad (5b)$$

element at a uniform rate G per unit volume and time would add an amount $G\ \Delta x\ \Delta y\ \Delta z\ \Delta\tau$.

The net heat flow into the element would be manifest as stored energy and would be equal to notation (6), where w is the specific weight (weight per

$$\Delta x\ \Delta y\ \Delta z\ wc\ \Delta t \quad (6)$$

unit volume) of the medium, c its specific heat, and Δt the temperature rise in the time increment $\Delta\tau$. Equating the net flow into the element to that stored and letting Δx, Δy, Δz, and $\Delta\tau \to 0$ leads to Eq. (7).

$$\kappa\left(\frac{\partial^2 t}{\partial x^2} + \frac{\partial^2 t}{\partial y^2} + \frac{\partial^2 t}{\partial z^2}\right) + G = wc\ \frac{\partial t}{\partial \tau} \quad (7)$$

When no heat source is present, Eq. (7) becomes Eq. (8). The ratio $\kappa/wc = \alpha$ is defined as the thermal diffusivity, and is the significant thermal property of a material for transient heat flow (**Fig. 5**).

$$\left(\frac{\partial^2 t}{\partial x^2} + \frac{\partial^2 t}{\partial y^2} + \frac{\partial^2 t}{\partial z^2}\right) = \frac{1}{\kappa/wc}\frac{\partial t}{\partial \tau} = \frac{1}{\alpha}\frac{\partial t}{\partial \tau} \quad (8)$$

Equation (8) will be recognized as the equation governing a potential field written in terms of temperature. Similar equations are satisfied by other potential field phenomena, such as electricity, magnetism, diffusion, and ideal fluid flow. Because of this, solutions to problems in one field are applicable to analogous systems in the others. Also, an experimental solution of a problem in one field may be obtained from an analogous system in another. *See* POTENTIALS.

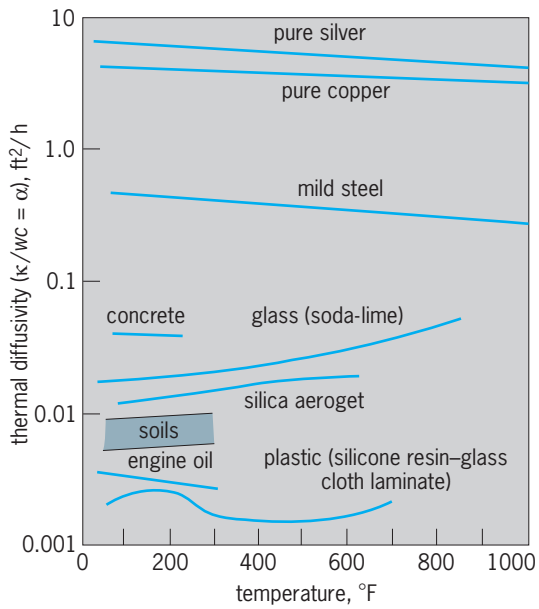

Fig. 5. Thermal diffusivities of some materials. °C = (°F − 32)/1.8.

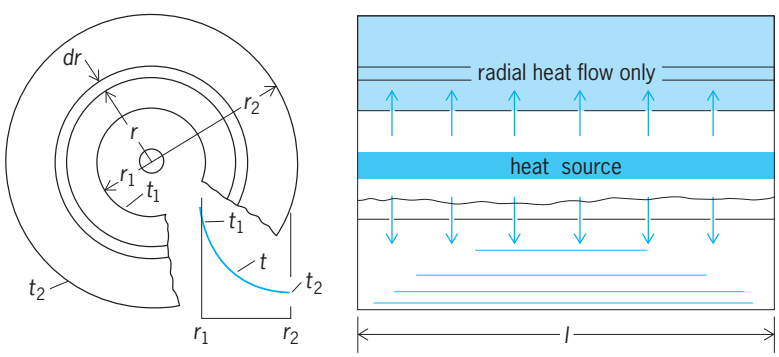

Fig. 6. Steady-state conduction. Diagrams show the temperature distribution for steady radial heat flow through a circular cylinder wall.

Steady-state conduction. When the temperature at all locations is constant with time or, mathematically, when $\partial t/\partial \tau = 0$, steady-state conduction exists. Many important practical problems fall in this category, the most familiar being heat flow through a wall and a hollow cylinder. Referring to Fig. 1 and assuming no heat generation within the wall, Eq. (8) reduces to Eq. (9). The terms $\partial^2 t/\partial y^2$ and $\partial^2 t/\partial z^2$ are

$$\frac{\partial^2 t}{\partial x^2} = 0 \tag{9}$$

eliminated, since t is considered to vary only with x. The desired temperature distribution is obtained by integrating Eq. (9) twice and evaluating the two constants from the known temperatures at $x = 0$ and $x = l$. This leads to Eq. (10). Application of Eq. (2) to obtain q yields Eq. (11).

$$t = t_1 - (t_1 - t_2)\frac{x}{l} \tag{10}$$

$$q = \kappa A \frac{(t_1 - t_2)}{l} \tag{11}$$

In the case of steady radial heat flow through a cylindrical wall (**Fig. 6**), Eq. (2) is applicable to any imaginary thin annular ring in the wall; thus Eq. (12)

$$q = -\kappa A \frac{dt}{dr} = -\kappa 2\pi r l \frac{dt}{dr} \tag{12}$$

is obtained. Integration from $t = t_1$ to t, and r_1 to r leads to Eq. (13), which indicates that the effect

$$t = t_1 - q\frac{\ln (r/r_1)}{2\pi \kappa l} \tag{13}$$

of the increasing area for heat flow is to produce a logarithmic variation in the temperature, as shown by the curve in the left part of Fig. 6.

Heat flow through a cylindrical wall is usually expressed in the same form as Eq. (11), written as Eq. (14). Solving Eq. (13) for q and substituting in

$$q = \kappa A_m \frac{t_1 - t_2}{r_2 - r_1} \tag{14}$$

Eq. (14) shows that the appropriate value for A_m must be as given in Eq. (15), which is called the logarithmic-mean area.

$$A_m = \frac{A_2 - A_1}{\ln (A_2/A_1)} \tag{15}$$

If the rate of heat flow and the inner and outer temperatures of a plane wall or hollow cylinder are measured during steady heat conduction, values of κ can be determined from Eq. (11) or Eq. (14). Because of the simplicity of the equations and the physical systems, most devices for measuring thermal conductivity are based on these types of heat flow.

Interface resistance. Now consider steady-state heat flow through a wall composed of two or more layers of material, each with different uniform thermal properties. If the surfaces of the various layers are very smooth and in very good contact with each other, the temperature distribution will be continuous. At any interface (**Fig. 7***a*), since q is constant, Eq. (16) holds. This shows that there is a discontinuity in the temperature gradient due to the change in κ.

$$\kappa_1 \left(\frac{dt}{dx}\right)_1 = \kappa_2 \left(\frac{dt}{dx}\right)_2 \tag{16}$$

Actual surfaces, even polished ones, are not smooth but have small projections and depressions.

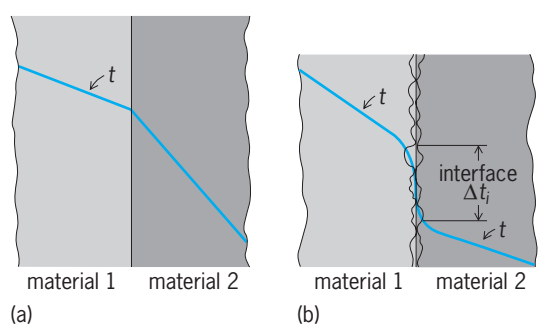

Fig. 7. Temperature distribution through composite wall. (*a*) With perfect interface contact. (*b*) For typical actual surfaces.

Consequently, when two surfaces are brought together, contact occurs primarily at projecting spots, as illustrated in Fig. 7*b*. The resulting contact area is only a small fraction of the nominal contact area. Plastic deformation at the points of contact of one or both materials usually occurs with the application of force to hold them together. Heat flows through both the small contact areas and the substance (usually a gas or liquid) filling the voids between the contacting protuberances.

The impairment to the heat-flow path caused by this imperfect contact is referred to as contact resistance. This is determined by extrapolating the measured temperature distribution in each material to the apparent interface location (Fig. 7*b*). The quotient of the resulting temperature difference Δt_i thus determined and the heat flux defines an interface resistance $R_i = \Delta t_i/q$. R_i depends on the roughnesses of the surfaces, the gas or liquid filling the voids, and the contact pressure. In general, the effect of contact resistance is significant only at low (for example, below 100 lb/in.^{2} or 700 kPa) interface pressures.

Internally generated heat. Conduction of heat generated internally occurs, for example, in the fuel elements used in nuclear reactors. Many of these elements are essentially long, flat plates over which a coolant flows. It is usually necessary to clad the fissionable material to prevent corrosion and keep radioactive particles from entering the coolant (**Fig. 8**). This cladding is undesirable from a heat-transfer standpoint, and is made as thin as possible and of the best adaptable heat-conducting material. Consequently, the temperature drop through it is usually small. *See* NUCLEAR REACTOR.

To illustrate the effect of the heat generation, assume that it is uniform with space and time in the radioactive material of Fig. 8 [that is, G in Eq. (7) is constant]. Heat flow from the ends of the elements and parallel to the flow direction of the coolant is negligible.

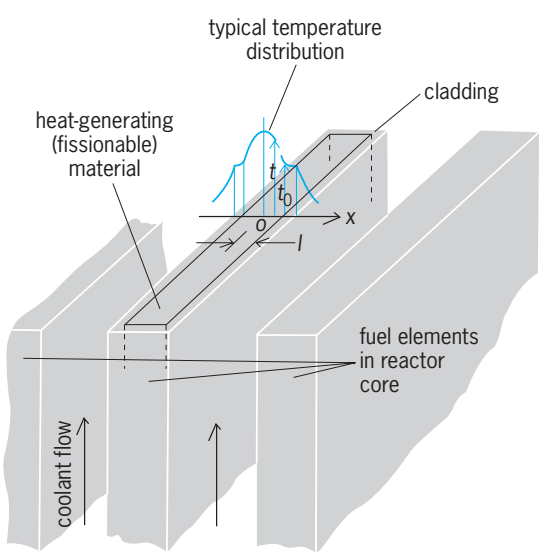

Fig. 8. Temperature distribution in reactor fuel plate.

Therefore, Eq. (7) reduces to Eq. (17). A solution

$$\frac{d^2t}{dx^2} + \frac{G}{\kappa} = 0 \tag{17}$$

for t is obtained by integrating twice and applying the boundary conditions (considering cooling to be the same on each side) $dt/dx = 0$ at $x = 0$ and $t = t_0$ at $x = l$. The result shows the temperature distribution to be parabolic, as in Eq. (18). The rate of heat transfer to the coolant is obtained by differentiating Eq. (18) with respect to x, evaluating at $x = l$, and substituting in Eq. (2).

$$t = t_0 + \frac{G}{2\kappa}(l^2 - x^2) \tag{18}$$

Since the heat generation rate is frequently dependent on the temperature, G will be more complex, possibly making an analytical solution impossible. Numerical methods of solution are then employed.

Periodic and transient conduction. These are the two kinds of non-steady-state heat flow. Periodic means a quasi-steady-state condition in which the temperature and heat flow at any location in a body vary continuosly with time, but pass through the same series of values in a definite period of time, τ_0. A transient state results when the heat flow at any location is momentarily or permanently changed. The duration of the transient period is the time required for the system to return to its original or a new steady-state condition. A transient change may be superimposed on a periodic variation.

Restricting consideration to examples in which no heat generation is present, the fundamental differential equation to be satisfied for either periodic or transient heat flow is Eq. (8). An interesting application is the temperature variation in the Earth due to the diurnal temperature variation, or to a sudden change in surface temperature. Equation (8) reduces in this case to Eq. (19). Assume the existence of a

$$\alpha \frac{\partial^2 t}{\partial x^2} = \frac{\partial t}{\partial \tau} \tag{19}$$

mean temperature of the Earth which is invariable with depth and that its surface temperature has been varying in a steady periodic manner long enough so that the original transient state due to starting the cyclic surface temperature has reached a steady periodic condition. If the surface temperature variation is given by Eq. (20), the appropriate solution to Eq. (19) is given by Eq. (21). This result shows that

$$t = t_0 \cos 2\pi \frac{\tau}{\tau_0} \tag{20}$$

$$t = t_0 e^{-\sqrt{\pi/2\alpha x}} \cos\left(2\pi \frac{\tau}{\tau_0} = \sqrt{\pi/2\alpha}\, x\right) \tag{21}$$

the temperature distribution looks like a wave traveling into the medium with the amplitude decreasing as the factor $e^{\sqrt{\pi/2\gamma}x}$. Computation of the heat flow by determining $(\partial t/\partial t)_{x=0}$ from Eq. (21) indicates that heat flows in during one-half of the period τ_0 and out during the other half.

Calculations of the temperature distribution in the Earth over an 8-h interval, using Eq. (21), are shown

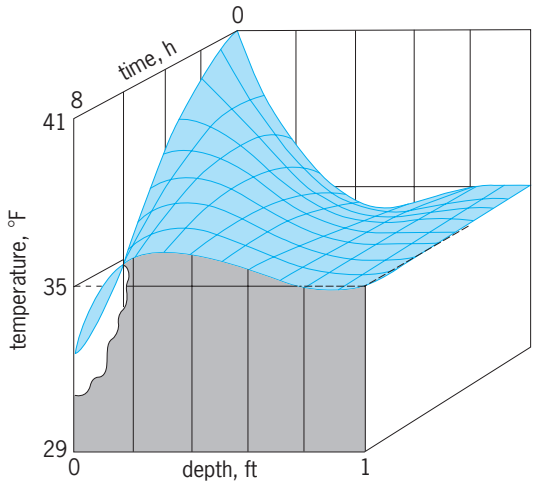

Fig. 9. Temperature distribution in the ground for a daily periodic surface variation. °C = (°F − 32)/1.8.

in **Fig. 9**. In this case the surface temperature varied from 29 to 41°F (−1.6 to 5.0°C) over a 24-h period; the diffusivity of the soil was 0.0065 ft²/h (6.0×10^{-4} m²/h).

An illustration of a transient state is the temperature variation resulting from a sudden change of magnitude t_0 in the temperature at the surface of the Earth. When the initial temperature throughout is uniform and taken as the datum, the applicable solution to Eq. (19) is given by Eq. (22). If a body of soil

$$t = \frac{2t_0}{\sqrt{\pi}} \int_0^{x/2\sqrt{\alpha\tau}} e^{-\beta^2} d\beta \tag{22}$$

in the Earth's surface is at a uniform temperature of 40°F (4.4°C) and the surface temperature suddenly drops to 20°F (−6.7°C), Eq. (22) can be applied to determine the depth at which the temperature will have dropped to freezing (32°F or 0°C) in 12 h. Taking $\alpha = 0.0065$ ft²/h, $t = 32$°F after 12 h is found to occur at $x = 0.27$ ft = 3.2 in. = 8.2 cm.

Variable thermal conductivity. Most materials are sufficiently homogeneous so that their thermal properties are independent of position. Assuming that they vary only with temperature, the heat balance on the element of Fig. 4 leads to Eq. (23). Upon carrying out the indicated differentiation, Eq. (23) becomes Eq. (24). By introducing a new variable θ called the conductivity potential, notation (25), Eq. (24) can be simplified to Eq. (26).

$$\frac{\partial}{\partial x}\left(\kappa\frac{\partial t}{\partial x}\right) + \frac{\partial}{\partial y}\left(\kappa\frac{\partial t}{\partial y}\right) + \frac{\partial}{\partial z}\left(\kappa\frac{\partial t}{\partial z}\right) + G = wc\frac{\partial t}{\partial \tau} \tag{23}$$

$$\kappa\left(\frac{\partial^2 t}{\partial x^2} + \frac{\partial^2 t}{\partial y^2} + \frac{\partial^2 t}{\partial z^2}\right) + \frac{\partial k}{\partial t}\left[\left(\frac{\partial t}{\partial x}\right)^2 + \left(\frac{\partial t}{\partial y}\right)^2 + \left(\frac{\partial t}{\partial z}\right)^2\right] + G = wc\frac{\partial t}{\partial \tau} \tag{24}$$

$$\theta = \int_0^t \kappa\, dt \tag{25}$$

$$\frac{\partial^2\theta}{\partial x^2} + \frac{\partial^2\theta}{\partial y^2} + \frac{\partial^2\theta}{\partial z^2} + G = \frac{1}{\alpha}\frac{\partial\theta}{\partial\tau} \tag{26}$$

Equation (26) has exactly the same form as Eq. (7). In many cases the variation of α with temperature (Fig. 5) is less than that of κ so that a mean constant value may be selected. This is true, for example, of metals at temperatures near absolute zero. In such cases, if G is not a function of t (or no heat source is present), solutions for constant α will apply with θ replacing t, provided that the boundary conditions are specified in terms of t or $\kappa(\partial t/\partial n)$ where n represents the variables x, y, z.

The utility of the conductivity potential is demonstrated in the calculation of the heat leak to a liquid-nitrogen tank through a support rod. Consider steady-state conditions with the exposed end of the rod at 300 K and the other at 77.3 K. Also assume there are no losses from the side of the rod which is 1 cm² in cross section, is 15 cm long, and is made of stainless steel (for which κ decreases from 0.15 at 300 K to 0.08 W/cm K at 77.3 K). At any location q is given by Eq. (27). Since q is constant, Eq. (28) holds. Then Eqs. (29) follows.

$$q = -A\,\kappa\,\frac{dt}{dx} \tag{27}$$

$$\frac{q}{A}\int_0^L dx = -\int_{t_1}^{t_2} \kappa\, dt = \int_0^{t_1} \kappa\, dt - \int_0^{t_2} \kappa\, dt \tag{28}$$

$$\frac{qL}{A} = \theta_1 - \theta_2 \quad q = A\,\frac{\theta_1 - \theta_2}{L} = 1\,\frac{30 - 3.3}{15} = 1.78 \text{ W} \tag{29}$$

Warren H. Giedt

Bibliography. A. Bejan and J. S. Jones, *Modern Heat Transfer*, 1993; B. Gebhart, *Heat Conduction and Mass Diffusion*, 1993; J. P. Holman, *Heat Transfer*, 8th ed., 1997; S. Kakac and Y. Yener, *Heat Conduction*, 3d ed., 1992; F. Krieth and M. S. Bohn, *Principles of Heat Transfer*, 1993; M. N. Ozisik, *Heat Conduction*, 2d ed., 1993; D. Poulikakos, *Conduction Heat Transfer*, 1994.

Conduction band

The electronic energy band of a crystalline solid which is partially occupied by electrons. The electrons in this energy band can increase their energies by going to higher energy levels within the band when an electric field is applied to accelerate them or when the temperature of the crystal is raised. These electrons are called conduction electrons, as distinct from the electrons in filled energy bands which, as a whole, do not contribute to electrical and thermal conduction. In metallic conductors the conduction electrons correspond to the valence electrons (or a portion of the valence electrons) of

the constituent atoms. In semiconductors and insulators at sufficiently low temperatures, the conduction band is empty of electrons. Conduction electrons come from thermal excitation of electrons from a lower energy band or from impurity atoms in the crystal. *See* BAND THEORY OF SOLIDS; ELECTRIC INSULATOR; ELECTRICAL CONDUCTIVITY OF METALS; SEMICONDUCTOR; VALENCE BAND. H. Y. Fan

Conductor (electricity)

Metal wires, cables, rods, tubes, bus-bars, and any other conducting media used for the purpose of carrying electric current. Although any metal assembly or structure can conduct electricity, the term conductor usually refers to the component parts of the current-carrying circuit or system.

Types of conductor. The most common forms of conductors are wires, cables, and bus-bars.

Wires. Wires employed as electrical conductors are slender rods or filaments of metal, usually soft and flexible. They may be bare or covered by some form of flexible insulating material. They are usually circular in cross section; for special purposes they may be drawn in square, rectangular, ribbon, or other shapes. Conductors may be solid or stranded, that is, built up by a helical lay or assembly of smaller solid conductors (**Fig. 1**).

Cables. Insulated stranded conductors in the larger sizes are called cables. Small, flexible, insulated cables are called cords. Assemblies of two or more insulated wires or cables within a common jacket or sheath are called multiconductor cables.

Bus-bars. Bus-bars are rigid, solid conductors and are made in various shapes, including rectangular, rods, tubes, and hollow squares. Bus-bars may be applied as single conductors, one bus-bar per phase, or as multiple conductors, two or more bus-bars per phase. The individual conductors of a multiple-conductor installation are identical. *See* BUS-BAR.

Sizes. Most round conductors less than $^1/_2$ in. (1.3 cm) in diameter are sized according to the American wire gage (AWG)—also known as the Brown & Sharpe gage. AWG sizes are based on a simple mathematical law in which intermediate wire sizes between no. 36 (0.0050-in. or 0.127-mm diameter) and no. 0000 (0.4600-in. or 11.684-mm diameter) are formed in geometrical progression. There are 38 sizes between these two diameters. An increase of three gage sizes (for example, from no. 10 to no. 7) doubles the cross-sectional area, and an increase of six gage sizes doubles the diameter of the wire.

Sizes of conductors greater than no. 0000 are usually measured in terms of cross-sectional area. Circular mil (cmil) is usually used to define cross-sectional area and is a unit of area equal to the area of a circle 1 mil (0.001 in. or 25.4 micrometers) in diameter.

Fig. 1. End views of stranded round conductor.

Wire lengths are usually expressed in units of feet or miles in the United States. Bus-bar sizes are usually defined by their physical dimensions—height and width in inches or fractions of an inch, and length in feet.

Materials. Most conductors are metals, but there is a wide range of conductivities (as much as a 70-to-1 ratio between different metals). Conductors are usually classified as good conductors, such as copper, aluminum, and silver, and poor conductors, such as iron; alloys of nickel, iron, copper, and chromium; and carbon products. In 1914, the U.S. Bureau of Standards made measurements of a large number of representative samples of copper and established standard values of resistivity and temperature coefficients, which have been adopted by the International Electrotechnical Commission. *See* ELECTRICAL CONDUCTIVITY OF METALS; ELECTRICAL RESISTIVITY; METAL.

Most wires, cables, and bus-bars are made from either copper or aluminum. Copper, of all the metals except silver, offers the least resistance to the flow of electric current. Both copper and aluminum may be bent and formed readily and have good flexibility in small sizes and in stranded constructions. Typical conductors are shown in Fig. 1.

Aluminum, because of its higher resistance, has less current-carrying capacity than copper for a given cross-sectional area. However, its low cost and light weight (only 30% that of the same volume of copper) permit wide use of aluminum for bus-bars, transmission lines, and large insulated-cable installations. *See* ALUMINUM; COPPER.

Metallic sodium conductors were used in 1965 on a trial basis for underground distribution insulated for both primary and secondary voltages. Sodium cable offered light weight and low cost for equivalent current-carrying rating compared with other conductor metals. Because of marketing problems and a few safety problems—the metal is reactive with water—the use of this cable was abandoned. *See* SODIUM.

For overhead transmission lines where superior strength is required, special conductor constructions are used. Typical of these are aluminum conductors, steel reinforced (ACSR), a composite construction of electrical-grade aluminum strands surrounding a stranded steel core (**Fig. 2**). Other constructions include stranded, high-strength aluminum alloy and a composite construction of aluminum strands around a stranded high-strength aluminum alloy core (ACAR). *See* TRANSMISSION LINES.

For extrahigh-voltage (EHV) transmission lines, conductor size is often established by corona performance rather than current-carrying capacity. Thus special "expanded" constructions are used to provide a large circumference without excessive weight. Typical constructions use helical lays of widely spaced aluminum strands around a stranded steel core. The space between the expanding strands is

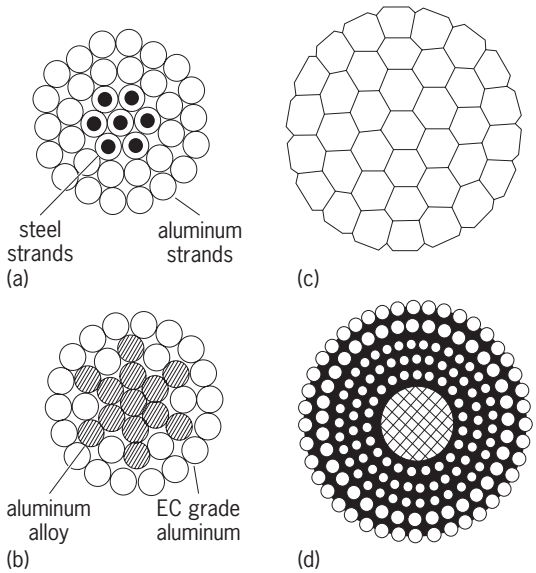

Fig. 2. Aluminum conductors. (*a*) ACSR. (*b*) ACAR. (*c*) Compact concentric stranded conductor. (*d*) Expanded-core concentric stranded conductor.

filled with paper twine, and outer layers of conventional aluminum strands are applied. In another construction the outer conductor stranding is applied directly over lays of widely separated helical expanding strands, without filler, leaving substantial voids between the stranded steel core and the closely spaced outer conductor layers. Diameters of 1.6 to 2.5 in. (41 to 63 mm) are typical. For lower reactance, conductors are "bundled," spaced 6–18 in. (15–45 cm) apart, and paralleled in groups of two, four, or more per phase. Figure 2 shows views of typical aluminum-conductor–steel-reinforced and expanded constructions.

Materials research has focused on amorphous or glassy materials, which include amorphous silicon-based semiconductors and metal alloys that totally lack crystalline structure. *See* AMORPHOUS SOLID.

Bare conductors. Bare wires and cables are used almost exclusively in outdoor power transmission and distribution lines. Conductors are supported on or from insulators, usually porcelain, of various designs and constructions, depending upon the voltage of the line and the mechanical considerations involved. Voltages as high as 765 kV are in use, and research has been undertaken into the use of EHV transmission lines, with voltages as high as 1500 kV.

Bare bus-bars are used extensively in outdoor substation construction, in switchboards, and for feeders and connections to electrolytic and electroplating processes. Where dangerously high voltages are carried, the use of bare bus-bars is usually restricted to areas accessible only to authorized personnel. Bare bus-bars are supported on insulators which have a design suitable for the voltage being carried.

Insulated conductors. Insulated electric conductors are provided with a continuous covering of flexible insulating material. A great variety of insulating materials and constructions has been developed to serve particular needs and applications. The selection of an appropriate insulation depends upon the voltage of the circuit, the operating temperature, the handling and abrasion likely to be encountered in installation and operation, environmental considerations such as exposure to moisture, oils, or chemicals, and applicable codes and standards. *See* ELECTRICAL INSULATION.

Magnet wires, used in the windings of motors, solenoids, transformers, and other electromagnetic devices, have relatively thin insulations, usually of enamel or cotton or both. Magnet wire is manufactured for use at temperatures ranging from 220 to 390°F (105 to 200°C). *See* MAGNET WIRE.

Conductors in buildings. Building wires and cables are used in electrical systems in buildings to transmit electric power from the point of electric service (where the system is connected to the utility lines) to the various outlets, fixtures, and utilization devices. Building wires are designed for 600-V operation but are commonly used at utilization voltages substantially below that value, typically 120, 240, or 480 V. Insulations commonly used include thermoplastic, natural rubber, synthetic rubber, and rubberlike compounds. Rubber insulations are usually covered with an additional jacket, such as fibrous braid or poly(vinyl chloride), to resist abrasion. Building wires are grouped by type in several application classifications in the National Electrical Code. *See* ELECTRICAL CODES.

Classification is by a letter which usually designates the kind of insulation and, often, its application characteristics. For example, type R indicates rubber or rubberlike insulation. TW indicates a thermoplastic, moisture-resistant insulation suitable for use in dry or wet locations; THW indicates a thermoplastic insulation with moisture and heat resistance. Other insulations in commercial use include silicone, fluorinated ethylene, propylene, varnished cambric, asbestos, polyethylene, and combinations of these.

Building wires and cables are also available in duplex and multiple-conductor assemblies; the individual insulated conductors are covered by a common jacket. For installation in wet locations, wires and cables are often provided with a lead sheath (**Fig. 3**).

For residential wiring, the common constructions used are nonmetallic-sheath cables, twin- and multiconductor assemblies in a tough abrasion-resistant jacket; and armored cable with twin- and multiconductor assemblies encased in a helical, flexible steel armor as in **Fig. 4**.

Power cables. Power cables are a class of electrical conductors used by utility systems for the

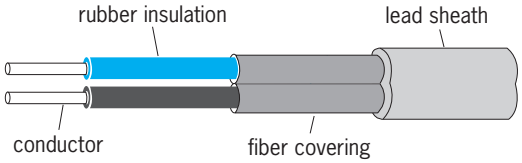

Fig. 3. Rubber-insulated, fiber-covered, lead-sheathed cable, useful for installation in wet places.

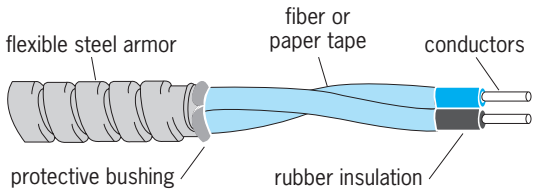

Fig. 4. Two-conductor armored cable.

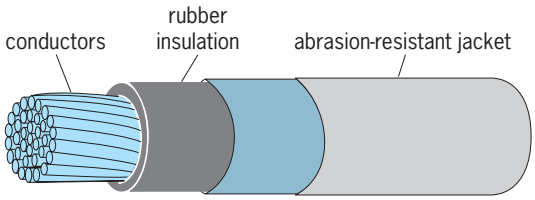

Fig. 5. Rubber-insulated power cable.

distribution of electricity. They are usually installed in underground ducts and conduits. Power cables are also used in the electric power systems of industrial plants and large buildings.

Power-cable insulations in common use include rubber, paper, varnished cambric, asbestos, and thermoplastic. Cables insulated with rubber (**Fig. 5**), polyethylene, and varnished cambric (**Fig. 6**) are used up through 69 kV, and impregnated paper to 138 kV. The type and thickness of the insulation for various voltages and applications are specified by the Insulated Power Cable Engineer Association (IPCEA).

Spaced aerial cable. These cable systems are used for pole-line distribution at, typically, 5–15 kV, three-phase. Insulated conductors are suspended from a messenger, which may also serve as a neutral conductor, with ceramic or plastic insulating spacers, usually of diamond configuration. For 15 kV, a typical system may have conductors with $^{10}/_{64}$-in. (4-mm) polyethylene insulation, 9-in. (23-cm) spacing between conductors, and 20 ft (6 m) between spacers.

High-voltage cable. High-voltage cable constructions and standards for installation and application are described by IPCEA. Insulations include (1) paper, solid type; (2) paper, low-pressure, gas-filled; (3) paper, low-pressure, oil-filled; (4) pipe cable, fully impregnated, oil-pressure; (5) pipe cable, gas-filled, gas-pressure; (6) rubber or plastic with neoprene or plastic jacket; (7) varnished cloth; (8) AVA and AVL (asbestos-varnished cloth).

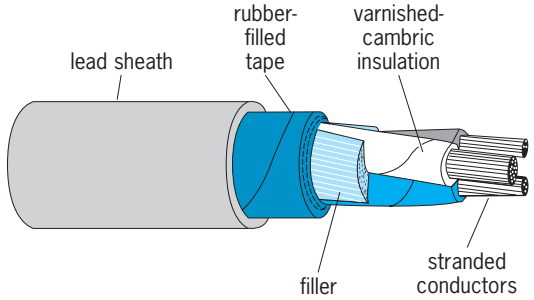

Fig. 6. Varnished cambric-insulated, lead-sheathed power cable, for voltages up to 28 kV.

Underground transmission cables are in service at voltages through 345 kV, and trial installations have been tested at 500 kV. Research has also been undertaken to develop cryoresistive and superconducting cables for transmitting power at high density.

Superconducting cables. It may be possible to apply the phenomenon of superconductivity to help provide more efficient and more compact electrical systems. In the area of electric power transmission, generation, and storage, superconductivity could have a major impact by increasing energy conservation. With the advent, in 1986, of superconducting materials with transition temperatures as high as 125 K (−235°F), there is a possibility of designing superconducting transmission lines that would carry as much as 1 GVA (10^9 W) of power. *See* SUPERCONDUCTING DEVICES; SUPERCONDUCTIVITY.

Busways. Enclosed bus-bar assemblies, or busways, are extensively used for service conductors and feeders in the electrical distribution systems of industrial plants and commercial buildings. They consist of prefabricated assemblies in standard lengths of bus-bars rigidly supported by solid insulation and enclosed in a sheet-metal housing.

Busways are made in two general types, feeder and plug-in. Feeder busways have no provision for taps or connections between the ends of the assembly. Low-reactance feeder busways are so constructed that conductors of different phases are in close proximity to minimize inductive reactance. Plug-in busways have provisions at intervals along the length of the assembly for the insertion of bus plugs.

Voltage drop in conductors. In electric circuits the resistance and (in ac circuits) the reactance of the circuit conductors result in a reduction in the voltage available at the load (except for capacitive loads). Since the line and load resistances are in series, the source voltage is divided proportionally. The difference between the source voltage and the voltage at the load is called voltage drop.

Electrical utilization devices (such as motors, household appliances, and electronic equipment) are designed to operate at a particular voltage or within a narrow range of voltages around a design value. The performance and efficiency of these devices are adversely affected if they are operated at a significantly lower voltage. Incandescent-lamp light output is lowered; fluorescent-lamp light output is lowered and starting becomes slow and erratic; the starting and pull-out torque of motors is seriously reduced.

Voltage drop in electric circuits caused by line resistance also represents a loss in power which appears as heat in the conductors. In excessive cases, the heat may rapidly age or destroy the insulation. Power loss also appears as a component of total energy use and cost. *See* COPPER LOSS.

Thus, conductors of electric power systems must be large enough to keep voltage drop at an acceptable value, or power-factor corrective devices—such as capacitors or synchronous condensers—must be

installed. A typical maximum for a building wiring system for light and power is 3% voltage drop from the utility connection to any outlet under full-load condition. *See* WIRING. H. Wayne Beaty

Bibliography. D. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; B. J. McPartland and J. F. McPartland, *McGraw-Hill's National Electrical Code Handbook*, 25th ed., 2005.

Cone

The solid of revolution obtained by revolving a right triangle about one of its shorter sides is called a cone, or more precisely a right circular cone (see **illus.**).

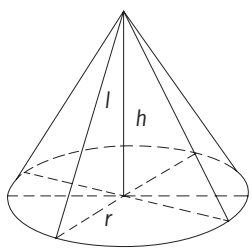

Right circular cone.

More generally, the term cone is used in solid geometry to describe a solid bounded by a plane and a portion of one nappe of a conical surface. In analytic geometry, however, the term cone refers not to a solid but to a conical surface. This is a surface generated by a straight line which moves so that it always intersects a given plane curve, called the directrix, and passes through a point, called the vertex, not in the plane of the directrix. The generating line in each of its positions is called an element of the cone. The vertex divides the surface into two parts, called nappes, that are congruent to each other in the extended sense (under rotation and reflection combined) but may not be superposable by a rigid motion in three-dimensional space.

If the elements of a cone make equal angles with a line through the vertex, the cone is called a cone of revolution, the line is called the axis of revolution, and the angle is called the semivertical angle ϵ. Plane sections of a cone of revolution are called conic sections.

If the plane makes an angle θ with the axis and does not contain the vertex, the resulting conic section is called nondegenerate and is a circle if $\theta = 90^\circ$, an ellipse if $\epsilon < \theta < 90^\circ$, a parabola if $\theta = \epsilon$, or a hyperbola if $\theta < \epsilon$. In the latter case the plane cuts both nappes of the cone, and the conic section consists of two separate congruent arcs called the branches of the hyperbola.

If the cutting plane contains the vertex of the cone, the conic section is degenerate and consists of a point, a line, or a pair of intersecting lines. Cones whose directrices are conic sections are called quadric cones.

The volume of a solid cone is $V = Bh/3$, where B is the base area included within the directrix, and h is the altitude (or height) measured from the vertex to the plane of the base. The volume V and surface area S of a right circular cone are $V = \pi r^2 h/3$ and $S = \pi rl + \pi r^2$, where r denotes the radius of the base and $l = \sqrt{r^2 + h^2}$ denotes the slant height, measured from vertex to base along an element of the cone.

A frustum of a cone is a solid bounded by two parallel planes and a portion of a conical surface. The plane boundaries are called bases. If the bases are circles of radii r and R, the volume of the frustum is $V = (\pi h/3)(r^2 + rR + R^2)$. *See* CONIC SECTION; EUCLIDEAN GEOMETRY. J. Sutherland Frame

Confocal microscopy

A technique that creates high-resolution images of very small objects but differs from conventional optical microscopy in that it uses a condenser lens to focus the illuminating light from a point source into a very small, diffraction-limited spot within the specimen, and an objective lens to focus the light emitted from that spot onto a small pinhole in an opaque screen (see **illus.**). Located behind the screen is a detector capable of quantifying how much light passes through the hole at any instant. Because only light from within the illuminated spot is properly focused to pass through the pinhole and reach the detector, any stray light from structures above, below, or to the side of the spot is filtered out. The image quality is therefore greatly enhanced.

Only the smallest possible spot is illuminated at any one time, and so a coherent image must be built up by scanning point by point over the desired field of view and recording the intensity of the light emitted from each spot. The size of the spot is equal to the ultimate resolution of the instrument and is typically about 0.25 micrometer in diameter and about 0.5 μm deep, although the dimensions vary with the wavelength of the light and the lens system used.

Scanning can be accomplished in several ways, but the most common system is laser scanning. This system uses a pair of mirrors, vibrating at precisely

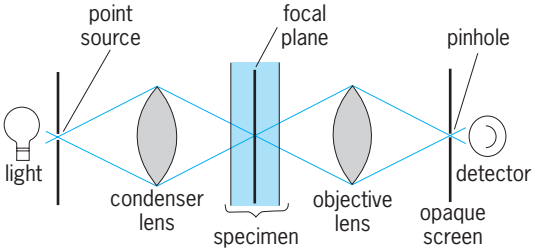

Schematic diagram of a confocal microscope. Light passes through a point source and a condenser lens to a point on the specimen. Only light coming from that spot is focused again by the objective lens onto a pinhole, through which it passes to reach the detector. Light from structures above or below the focal plane is too diffuse to pass through, and light from structures beside the spot is directed to the side of the pinhole. (*After R. J. Taylor, Confocal microscopy sheds new light on the dynamics of living cells, J. NIH Res., 1(1):113–115, 1989*)

controlled frequencies, which are placed in the path of a laser beam and direct it in a sweeping pattern over the specimen. The intensity of light emitted from each spot is digitally recorded, and the complete set of intensities from each point is displayed on a video screen, allowing the total image to be seen at one time. Because the information is already in a digital form, it can easily be passed to a computer for storage or further analysis.

Although confocal microscopy was first described in 1961, instruments were not available commercially until the late 1980s. They have found applications in several fields, including materials science and the semiconductor industry, but they have a number of advantages that make them particularly useful in the biological sciences. Many dyes used in biological studies are fluorescent and emit light at a wavelength different from the illuminating light. The amount of fluorescence from different parts of a cell, which is often directly proportional to the concentration of a specific intracellular constituent, can be measured far more precisely with a confocal microscope, because of the rejection of interference from other parts of a specimen. Haze from out-of-focus structures above the focal plane is eliminated, and so it is possible to see very clear images from fairly deep within a relatively thick slice of tissue without having to disrupt its three-dimensional organization. A stack of images from successive focal planes can be obtained simply by moving the narrow focal plane through the specimen. That technique makes it relatively easy to build up complex, three-dimensional views of cellular interconnections or subcellular structures. *See* FLUORESCENCE MICROSCOPE.

Robert J. Taylor

Bibliography. H. Shuman, J. M. Murray, and C. DiLullo, Confocal microscopy: An overview, *BioTechniques*, 7(2):154–163, 1989; R. J. Taylor, Confocal microscopy sheds new light on the dynamics of living cells, *J. NIH Res.*, 1(1):113–115, 1989; J. G. White and W. B. Amos, Confocal microscopy comes of age, *Nature*, 328:183–184, 1987; J. G. White, W. B. Amos, and M. Fordham, An evaluation of confocal versus conventional imaging of biological structures by fluorescence light microscopy, *J. Cell Biol.*, 105:41–48, 1987.

Conformal mapping

A special operation in mathematics in which a set of points in one coordinate system is mapped or transformed into a corresponding set in another coordinate system, preserving the angle of intersection between pairs of curves.

A mapping or transformation of a set E of points in the xy plane onto a set F in the uv plane is a correspondence, Eqs. (1), that is defined for each

$$u = \varphi(x, y) \qquad v = \psi(x, y) \tag{1}$$

point (x,y) in E and sends it to a point (u,v) in F, so that each point in F is the image of some point in E. A mapping is one to one if distinct points in E are transformed to distinct points in F. A mapping is conformal if it is one to one and it preserves the magnitudes and orientations of the angles between curves (**Fig. 1**). Conformal mappings preserve the shape but not the size of small figures.

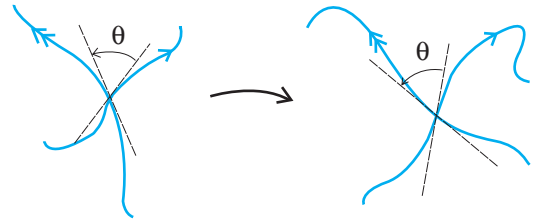

Fig. 1. Angle-preserving property of a conformal mapping.

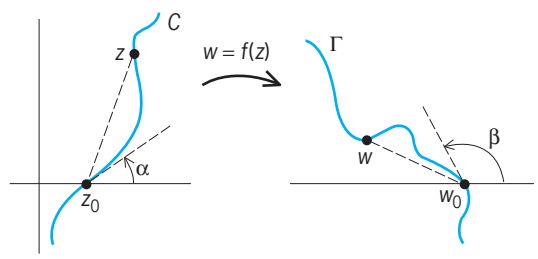

Fig. 2. Illustration for proof that analytic functions preserve angles.

Relation to analytic functions. If the points (x,y) and (u,v) are viewed as the complex numbers $z = x + iy$ and $w = u + iv$, the mapping becomes a function of a complex variable: $w = f(z)$. It is an important fact that a one-to-one mapping is conformal if and only if the function f is analytic and its derivative $f'(z)$ is never equal to zero. This can be seen by considering a function f that is analytic near $z_0 = x_0 + iy_0$, with $f'(z_0) \neq 0$, and mapping a curve C passing through z_0 to a curve Γ passing through $w_0 = f(z_0)$. Then, as a point z tends to z_0 along C, the angle $\arg\{z - z_0\}$ tends to the angle α between C and the line $y = y_0$ (**Fig. 2**). But by assumption, expression (2) is valid. In particular,

$$\frac{f(z) - w_o}{z - z_o} \to f'(z_o) \tag{2}$$

$\arg\{f(z) - w_0\} - \arg\{z - z_0\}$ tends to a limit $\gamma = \arg\{f'(z_0)\}$. In other words, $\arg\{w - w_0\} \to \beta$ as $w \to w_0$ along Γ, where $\beta = \alpha + \gamma$. But any other curve through z_0 is rotated by the same angle γ, so that the mapping preserves the angle between the two curves. In a similar manner, it can be shown that a conformal mapping is necessarily given by an analytic function. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

Examples. Some examples of functions that provide conformal mappings (and one that is not conformal) will now be given.

1. The function $w = (z - 1)/(z + 1)$ maps the right half-plane, defined by $\mathrm{Re}\{z\} > 0$, conformally onto the unit disk, defined by $|w| < 1$.

2. The function $w = \log z$ maps the right half-plane conformally onto the horizontal strip defined by $-\pi/2 < \mathrm{Im}\{w\} < \pi/2$.

3. The function $w = z^2$ maps the upper semidisk, defined by $\mathrm{Im}\{z\} > 0$ and $|z| < 1$, conformally onto the unit disk $|w| < 1$ with the segment $0 \le u < 1$ removed. It doubles angles at the origin, but this is a boundary point which does not lie in the semidisk.

4. The function $w = \bar{z}$ of complex conjugation preserves the magnitudes but not the orientations of angles between curves. It is nowhere conformal.

5. The function $w = z + 1/z$ maps the unit disk $|z| < 1$ conformally onto the extended complex plane (including the point at infinity) with the line segment $-2 \le u \le 2$ removed.

6. The Koebe function $k(z) = z(1 - z)^{-2}$ maps the unit disk conformally onto the complex plane with the half-line $-\infty < u \le -1/4$ removed.

Linear fractional transformations. A linear fractional transformation is a function of the form given by Eq. (3), where *a, b, c*, and *d* are complex constants.

$$w = \frac{az + b}{cz + d} \qquad ad - bc \neq 0 \tag{3}$$

It is also known as a Möbius function. One example is the function $w = (z - 1)/(z + 1)$, discussed above. Simpler examples are magnifications, given by Eq. (4), rotations, given by Eq. (5), and inversion, given by Eq. (6). Every linear fractional transformation is a composition of linear fractional transformations of these three special types. Thus each linear fractional transformation provides a conformal mapping of the extended complex plane onto itself, and in fact the linear fractional transformations are the only such mappings. There is a unique linear fractional transformation which carries three prescribed (distinct) points z_1, z_2, z_3 to prescribed images w_1, w_2, w_3. The most general conformal mapping of the unit disk onto itself is a linear fractional transformation of the form given by Eq. (7).

$$w = az \qquad a > 0 \tag{4}$$

$$w = az \qquad |a| = 1 \tag{5}$$

$$w = 1/z \tag{6}$$

$$w = \frac{z - z_0}{1 - \bar{z}_0 z} \qquad |z_0| < 1 \tag{7}$$

Each linear fractional transformation carries circles to circles and symmetric points to symmetric points. Here a circle means a circle or a line. Two points are said to be symmetric with respect to a circle if they lie on the same ray from the center and the product of their distances from the center is equal to the square of the radius. Two points are symmetric with respect to a line if the line is the perpendicular bisector of the segment joining the two points. As an instance of this general property of linear fractional transformations, the mapping $w = (z - 1)/(z + 1)$ sends the family of circles of Apollonius with symmetric points 1 and -1 (defined by requiring that on each circle the quotient of the distances from 1 and -1 be constant) onto the family of all circles centered at the origin. It carries the orthogonal family of curves, consisting of all circles through the points 1 and -1, onto the family of all lines through the origin (**Fig. 3**).

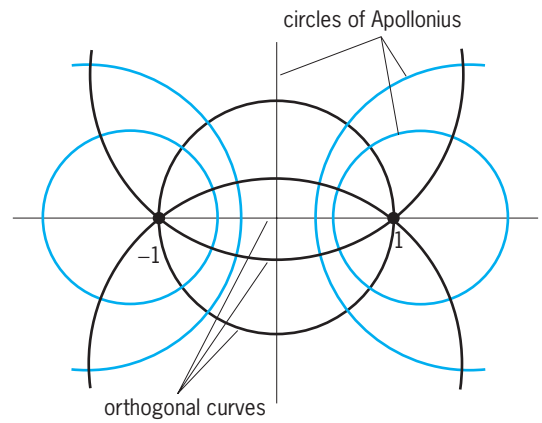

Fig. 3. Circles of Apollonius and members of the orthogonal family of curves.

Applications. Conformal mappings are important in two-dimensional problems of fluid flow, heat conduction, and potential theory. They provide suitable changes of coordinates for the analysis of difficult problems. For example, the problem of finding the steady-state distribution of temperature in a conducting plate requires the calculation of a harmonic function with prescribed boundary values. If the region can be mapped conformally onto the unit disk, the transformed problem is readily solved by the Poisson integral formula, and the required solution is the composition of the resulting harmonic function with the conformal mapping. The method works because a harmonic function of an analytic function is always harmonic. *See* CONDUCTION (HEAT); FLUID-FLOW PRINCIPLES; LAPLACE'S DIFFERENTIAL EQUATION; POTENTIALS.

The term conformal applies in a more general context to the mapping of any surface onto another. A problem of great importance for navigation is to produce conformal mappings of a portion of the Earth's surface onto a portion of the plane. The Mercator and stereographic projections are conformal in this sense. *See* MAP PROJECTIONS.

Riemann mapping theorem. In 1851, G. F. B. Riemann enunciated the theorem that every open simply connected region in the complex plane except for the whole plane can be mapped conformally onto the unit disk. Riemann's proof was incomplete; the first valid proof was given by W. F. Osgood in 1900. Most proofs exhibit the required mapping as the solution of an extremal problem over an appropriate family of analytic univalent functions. (A univalent function is simply a one-to-one mapping.) For instance, the Riemann mapping maximizes $|f'(z_0)|$ among all analytic univalent functions which map the given region into the unit disk: $|f(z)| < 1$. Here z_0 is chosen arbitrarily in the region, and the Riemann mapping has $f(z_0) = 0$.

Multiply connected regions. There is no exact analog of the Riemann mapping theorem for multiply

connected regions. For instance, two annuli, $r_1 < |z| < r_2$ and $R_1 < |w| < R_2$, are conformally equivalent if and only if $r_2/r_1 = R_2/R_1$. Any doubly connected region can be mapped conformally onto a (possibly degenerate) annulus. Any finitely connected region (other than a punctured plane) can be mapped conformally onto the unit disk minus a system of concentric circular arcs, or onto the whole plane minus a system of parallel segments, or radial segments, or concentric circular arcs. Other canonical regions are bounded by arcs of lemniscates or logarithmic spirals, or by full circles.

Distortion theorems. Conformal mappings are often studied by considering the class S of functions $f(z)$ which are analytic and univalent in the unit disk and have the normalizing properties $f(0) = 0$ and $f'(0) = 1$. Alternatively, the class S may be defined as the class of all univalent power series of the form $f(z) = z + a_2z^2 + a_3z^3 + \cdots$ that are convergent for $|z| < 1$. The Koebe distortion theorem gives the sharp bounds, Eqs. (8), for all functions f in S and all points

$$\begin{aligned} &r(1+r)^{-2} \le |f(z)| \le r(1-r)^{-2} \quad r = |z| \\ &(1-r)(1+r)^{-3} \le |f'(z)| \le (1+r)(1-r)^{-3} \end{aligned} \tag{8}$$

z in the disk. The Koebe $^1/_4$-theorem asserts that each function f in S includes the full disk $|w| < {}^1/_4$ in its range. Suitable rotations of the Koebe function (discussed above) show that each of these bounds is the best possible. All of these statements can be deduced from a theorem of L. Bieberbach (1916) that $|a_2| \le 2$ for all functions f in S, with equality occurring only for the Koebe function, given by Eq. (9), or one of

$$k(z) = z(1-z)^{-2} = z + 2z^2 + 3z^3 + \cdots \tag{9}$$

its rotations, given by $e^{-i\theta}k(e^{i\theta}z)$ for some value of θ. Bieberbach conjectured that in general $|a_n| \le n$ for all n. For many years, the Bieberbach conjecture stood as a challenge and inspired the development of powerful methods in geometric function theory. After a long series of advances by many mathematicians, the final step in the proof of the Bieberbach conjecture was taken by L. de Branges in 1984.

Boundary correspondence. The open region inside a simple closed curve is called a Jordan region. Riemann's theorem ensures the existence of a conformal mapping of one Jordan region onto another. C. Carathéodory proved in 1913 that such a mapping can always be extended to the boundary and the extended mapping is a homeomorphism, or a bicontinuous one-to-one mapping, between the closures of the two regions. In fact, Carathéodory proved a much more general theorem which admits inaccessible boundary points and establishes a homeomorphic correspondence between "clusters" of boundary points, known as prime ends. For a conformal mapping of the unit disk onto a Jordan region with a rectifiable boundary, the Carathéodory extension preserves sets of measure zero (or zero "length") on the boundary. It is also angle-preserving at almost every boundary point, that is, except for a set of points of measure zero. If the boundary of the Jordan region has a smoothly turning tangent direction, the derivative of the mapping function can also be extended continuously to the boundary. *See* MEASURE THEORY.

Quasiconformal mappings. A theory of generalized conformal mappings, known as quasiconformal mappings, has evolved. Roughly speaking, a univalent function $w = f(z)$ is said to be K-quasiconformal in a certain region if it maps infinitesimal circles to infinitesimal ellipses in which the ratios of major to minor axes are bounded above by a constant $K \ge 1$. Equivalently, a mapping is quasiconformal if it distorts angles by no more than a fixed ratio. The simplest examples are linear mappings of the form $u = ax + by$, $v = cx + dy$, $ad - bc \ne 0$, where a, b, c, and d are real constants. The 1-quasiconformal mappings are simply the conformal mappings. The notion of a quasiconformal mapping is readily extended to higher dimensions. Peter L. Duren

Bibliography. L. Bieberbach, *Conformal Mapping*, 1952, reprint 2000; R. V. Churchill and J. W. Brown, *Complex Variables and Applications*, 6th ed., 1995; P. L. Duren, *Univalent Functions*, 1983; R. Schinziger and P. A. Laura, *Conformal Mapping: Methods and Applications*, 1991.

Conformal optics

Conformal optical systems have outer surfaces whose shape is chosen to optimize the interaction with the environment in which the optical system is being used. The imaging through such conformal optical windows is likely to suffer from extreme aberration, requiring special techniques for correction. *See* ABERRATION (OPTICS).

For example, a conformal optical window might relate to the nose of an infrared-seeking missile. These conformal surfaces permit a beneficial interface with the environment and allow substantial improvements in the overall performance of the vehicle. Recently there has been considerable research into methods of compensating for the aberrations produced by such windows. Computer-intensive methods of design, fabrication, and testing of optics have reached a level where the development of cost-effective methods for insertion of these conformal optics concepts into operational systems appears to be practical.

Applications. Important applications of conformal optics are found in missile and aircraft systems. Missiles and aircraft carry optical sensors for imaging, detection, and ranging that must look at the world through the outer skin of the vehicle. Traditionally, the windows for viewing through the skin of missiles and aircraft have had simple optical forms, such as flats or spheres, that enable the optical tracking systems to operate by using well-known technology. But these optically advantageous windows degrade the performance of the vehicle through increased drag, aerodynamic heating, or other undesirable effects. One example is the use of an optical tracker or seeker on the front end of a missile. These

optical systems, usually used with infrared sensors, need to have a wide and unobstructed field of view. The classical solution for the window for these optics is a hemispherical dome mounted at the front of the missile. An optical tracker located within the dome is scanned around the field of view to keep track of a target. The spherical-surfaced-dome approach is easy to accommodate in the design, fabrication, and testing of the system, but the blunt spherical form adds considerable aerodynamic drag to the missile. The use of a conformal window, whose shape conforms more closely to the optimal, pointed ogival shape, reduces the drag of the missile and provides significant gains in the missile's performance. Such ogival shapes produce considerable optical aberration, however.

Similar considerations apply to aircraft. The optical sensors of most combat aircraft are mounted in pods that attach to the wings or other outer surfaces. These pods increase drag and reduce maneuverability and stealth properties of the aircraft. Location of the sensors within the body of the aircraft using traditional flat windows limits their field of view. The use of windows that have surfaces conforming to the desirable aerodynamic shapes of the aircraft is necessary to permit internal, drag-free location of these sensors.

One example of the benefits to be obtained through the appropriate use of conformal optics is a typical missile. Replacement of a hemispherical front-end window by an optimal-shaped ogival window can reduce the drag coefficient by as much as 25%. This reduction of the drag will lead to an increase in range for the same fuel consumption. There are additional advantages such as improved maneuverability of the missile.

The **illustration** shows one possible implementation of a conformal window to replace a conventional optical spherical window. It is a diagram of a conformal window, shaped to provide an improved drag profile for the missile. The optical tracker behind the conformal window can be seen looking toward the side of the missile. The optical tracker no longer sees a simple and constant spherical form for the window at any angle of view. The aspheric shape of the window produces a significant amount of aberration that must be compensated at some location in the optical system. The variation of the aberration with direction of view requires that this correction be dynamic.

There are other applications in which general aspheric optical surfaces are needed. Compact illumination systems require many unique forms of conformal surfaces, sometimes for utilitarian reasons and frequently for design or styling reasons. One growing area of interest comprises unique compact headlight systems for automobiles. Unobscured reflecting imaging systems used in spacecraft or microlithographic optics require nonconventional aspheric shapes for packaging or aberration correction purposes.

Challenges. The most desirable form for the front end of a missile is a pointed shape called an ogive. This ogival shape reduces the aerodynamic drag on the missile but produces a greatly distorted view of the scene that is viewed through the ogival dome. An optical tracker looking through the dome will see large amounts of aberration. A conformal optical surface is extremely aspheric and cannot be described by traditional surface prescriptions. Traditional optical design and fabrication methods cannot cope with the large amount of aberration that is generated by such conformal shapes. There are no established guidelines for the incorporation of these extreme aspheric conformal surfaces in the design of optical systems.

Fabrication of these conformal surfaces provides a new set of challenges. Traditional optical fabrication is based on the generation and polishing of spherical or flat surfaces by a lapping process. The surface eventually converges to the final required shape as a natural result of the process. Most aspherics used in lenses are minor deviations from these base spherical surfaces. Conformal surfaces require significant departures from traditional optical surfaces. New techniques for shape generation and polishing need to be developed so that conformal surfaces can be successfully fabricated. *See* OPTICAL SURFACES.

Test and measurement of these surfaces provide a new set of challenges. There are no natural nulling tests, which provide an absolute reference for the surface shape. Testing requires new methods of numerically constructing reference surfaces for defining, aligning, and calibrating the conformal surface. The use of these new surfaces as windows for optical systems also provides some new challenges in operation. The alignment between the window and the tracking optics is very critical. Therefore, alignment during installation and field operation must frequently be carried out by new adaptive optical techniques.

All of these new approaches require novel technical solutions within the framework of economic reality. None of these conformal techniques can be inserted into operational practice unless they are cost effective. Thus the cost of introducing conformal surfaces into missiles and aircraft must not exceed a small fraction of the cost benefit.

Design approach. The approach to design with conformal optics does not call for complete abandonment of current understanding of the image-formation process. Optical design methods are based on the description of the wavefront passing through surfaces by use of numerical ray tracing. The understanding of the aberrations arising at surfaces is obtained from an analytic method for describing the surface and the wavefront to stated levels of accuracy. The aberrations produced by general aspheric surfaces defy simple analytic descriptions but can be obtained by fitting of the numerical ray-tracing results. It is found that for most conformal surfaces the local aberration for a wavefront passing through a conformal window can be well described in terms of combinations of simple aberration shapes such as focus, astigmatism, and coma. The variation of these aberrations with location on the window becomes very difficult to describe analytically, leading to a

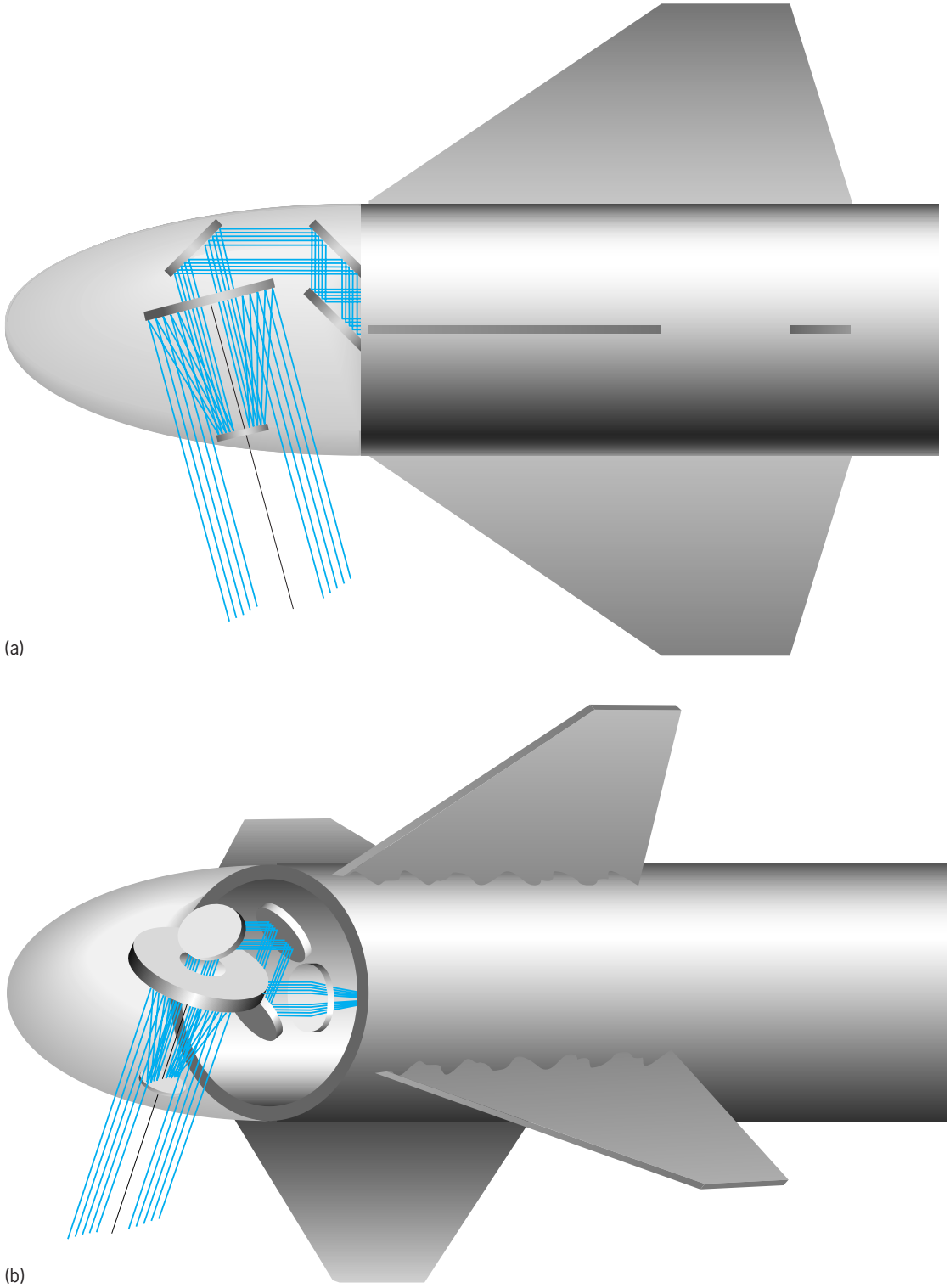

Geometric optics of an optical tracker behind a conformal window that is shaped to provide an improved drag profile for a missile. Paths of rays through the optical system are shown. (*a*) Side view. (*b*) Oblique view. *(Optical Research Associates)*

design process in which the numerical ray tracing is used to establish the aberration level.

An intuitive grasp of the aberration content of conformal windows can be obtained by graphical presentations of the behavior of the aberrations with respect to the angle of view through the window. New design approaches permit expressing the targets of the design programs in terms of combinations of the sets of aberrations that naturally describe the effect of the conformal windows. New types of nonsymmetric aspheric components are required to compensate these sets of aberrations and are incorporated into the tracker optics. New approaches to tolerancing and assembling of these unusual optical systems are being investigated. *See* GEOMETRICAL OPTICS.

Robert R. Shannon

Conformational analysis

The study of the energies and structures of conformations of organic molecules and their chemical and physical properties. Organic molecules are not static entities. The constituent atoms vibrate and groups rotate about the bond axes. *See* STEREOCHEMISTRY.

Linear structures. Rotation about the C-C bond in ethane, for example, results in an infinite number of slightly different structures called conformations, two of these are indicated below.

In conformation (**1**) of ethane, the pairs of C-H

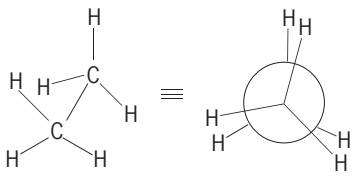

(**1**)

bonds on the two carbon atoms reside in a plane, and are termed eclipsed. (The circle represents the front-most carbon atom, and the long bonds those from that carbon atom to the hydrogen atoms. The bonds from the circle represent those bonds from the rearward carbon to its hydrogen atoms.) Conformation (**1**) is called the eclipsed conformation. In conformation (**2**), the C-H bonds on one carbon re-

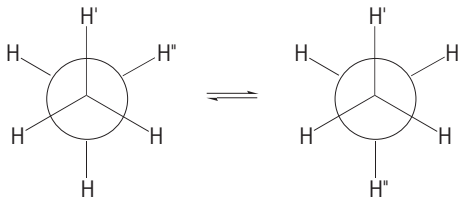

(**2**) (**3**)

side between the C—H bonds on the other carbon. Conformation (**2**) is called the staggered conformation. An infinite number of conformations between (**1**) and (**2**) are possible; however, conformations (**1**) and (**2**) are of greatest interest because they are the maximum and minimum energy structures. In conformation (**1**), the electrons in the C-H bonds and the nuclear charges of the hydrogen atoms repel each other, resulting in a higher energy state. In conformation (**2**), the electrons and hydrogen atoms are at their greatest possible separation and the repulsion is at a minimum. The conversion of (**2**) to (**3**) by rotation of one of the methyl groups by 120° requires the input of energy, approximately 3.0 kcal (13 kilojoules) per mole, in order to pass through the eclipsed conformation. This amount is small compared to the thermal energy at room temperature, and the rotation about the C-C bond in ethane occurs at about 10^9–10^{10} times per second.

Butane provides a more complicated case. There are two different eclipsed (**4** and **6**) and two staggered (**5** and **7**) conformations which are desig-

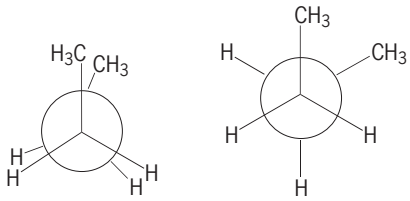

(**4**) synperiplanar (**5**) synclinal

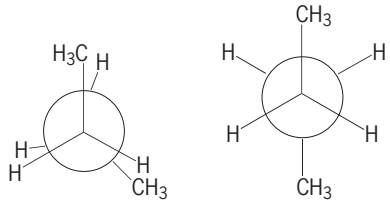

(**6**) anticlinal (**7**) antiperiplanar

nated by the names below the structures. As the methyl group is larger than a hydrogen atom, (**4**) is higher in energy than (**6**), and (**5**) is higher in energy than (**7**) [see **illus.**]. As the molecular weight of an alkane increases, the number of molecular conformations (combinations of all possible individual bond conformations) increases dramatically, although the antiperiplanar conformations are always favored.

1,2-Dichloroethane also exists in conformations similar to those of butane (see interconverting conformations below). In the gas phase, the antiperi-

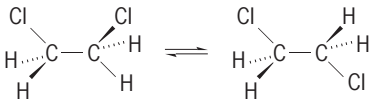

synclinal antiperiplanar

planar conformation predominates. This is not due to a steric repulsion in the synclinal conformation, but to C-Cl bond dipole repulsion. (The bond dipole moment arises from the fact that chlorine is more electronegative than carbon.) In a polar solvent, the energy difference between the antiperiplanar and synclinal conformations is decreased due to

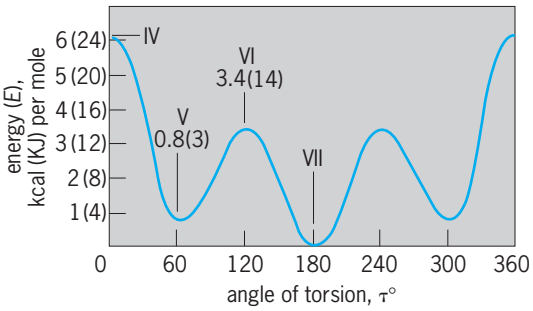

Potential energy diagram for butane. (*After E. L. Eliel et al., Conformational Analysis, Interscience Publishers, 1965*)

stabilization of the synclinal form. In pure liquid 1,2-dichloroethane, the enthalpy difference is zero. Such solvent effects are not observed with nonpolar molecules such as the alkanes.

Cyclic structures. These also exist in various conformations. Cyclopropane, a planar structure, can exist in only one conformation. In cyclobutane, slight twisting about the C-C bonds can occur which relieves some of the C-H eclipsing strain energy. Cyclobutane exists in rapidly interconverting so-called butterfly conformations (**8**). Cyclopentane exists in

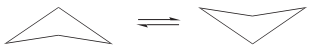

(**8**)

two types of nonplanar conformations: the envelope conformation (**9**) and the half-chair (or twist) conformation (**10**). The flap atom of (**9**), the atom out of the

(**9**) (**10**)

plane of the other four, can migrate around the ring, as can also the twist in (**10**). These motions, called pseudorotation, require very little energy, and cyclopentane presents a very complex conformational system.

Cyclohexane exists predominantly in the chair conformation (**11**), in which all C-C bond conformations are of the staggered type. Interconversion between chair conformations occurs rapidly at room temperature, passing through the high-energy, boat conformation (**12**). In the chair conformation, there

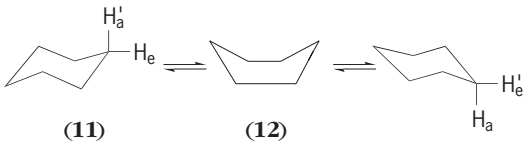

(**11**) (**12**)

are two distinctly different types of hydrogen atoms; one set oriented perpendicular to the general plane of the ring, called axial (a) hydrogens, and one set oriented parallel to the plane of the ring, called equatorial (e) hydrogens. These hydrogens interchange orientation on chair interconversion. Axial and equatorial hydrogens possess different chemical and physical properties, although at room temperature the interconversion occurs very rapidly ($\sim 10^5$ per second), and it is in general not possible to detect the differences. It is easy to design molecules in which the interconversion is not possible and the differences in properties become readily apparent.

A substituted cyclohexane, for example methylcyclohexane, exists in two different chair conformations which are in equilibrium. In the axial conformation (**13**), the methyl group is in proximity with the other two axial hydrogens on the same side of the ring. This steric repulsion disfavors the axial conformation and favors equatorial conformation (**14**). The measurement of conformational equi-

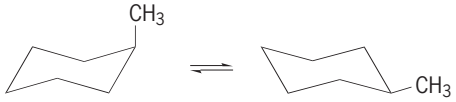

(**13**) (**14**)

librium constants has allowed the construction of a relative size scale for organic functional groups.

Rotation about C-C bond. Rotation about the C-C double bond requires a very large amount of energy, and is rarely observed except under very forcing conditions. Thus, although *cis*- and *trans*-2-butene can be considered to be conformational isomers, practically this does not occur.

An illustration of how the conformation of a molecule affects a chemical reaction is provided by the elimination of hydrogen bromide from 2-bromobutane, which can exist in either of the two staggered conformations (**15**) and (**16**). The elimination of hydrogen bromide requires that the hydrogen and bromine be antiperiplanar. Thus, the elimination from (**15**) produces *trans*-2-butene [reaction (1)], and from **16**, *cis*-2-butene [reaction (2)].

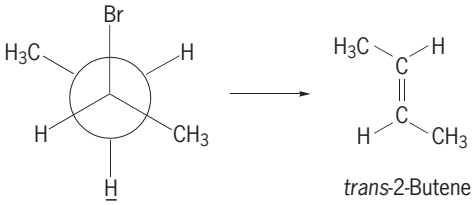

(1)

trans-2-Butene

(**15**)

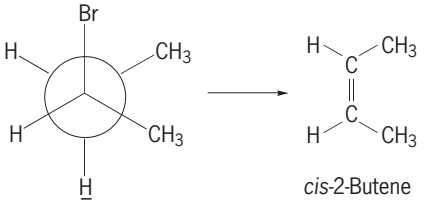

(2)

cis-2-Butene

(**16**)

Conformational analysis represents a very large area of study and has provided important information on the structures and reactivities of complex natural products such as steroids and alkaloids. *See* MOLECULAR ISOMERISM; STEREOCHEMISTRY. Daniel J. Pasto

Bibliography. E. L. Eliel et al., *Conformational Analysis*, 1965; E. L. Eliel and S. H. Wilen, *Stereochemistry of Organic Compounds*, 1994; E. Juaristi, *Introduction to Stereochemistry and Conformational Analysis*, 1991.

Confuciusornithidae

An extinct family of flying birds that represents one of the most primitive avian groups. It comprises three genera (*Confuciusornis*, *Changchengornis*,

and *Jinzhouornis*) and is known exclusively from western Liaoning Province and neighboring areas in northeastern China. Fossils of this group have been collected from the Yixian and Jiufotang formations, which represent freshwater lake deposits that were formed in the Early Cretaceous period 125–120 million years ago.

Confuciusornithids are generally similar in size to *Archaeopteryx* and retain three wing claws and an unreduced phalangeal formula, that is, the phalangeal formula of the hand is 2-3-4 (two phalanges in the first digit, three phalanges in the second digit, and four phalanges in the third digit). The latter may indicate that the hand was used as an aid in climbing or grasping tree trunks and branches, as has also been suggested in the case of *Archaeopteryx* and some of their closely related dinosaurian ancestors. In addition to these primitive features, confuciusornithids share a number of features with modern birds, including a short skeletal tail in which the final few vertebrae are fused together into a pygostyle (or plowshare bone). Tail feathers would have been attached to the pygosyle, suggesting that confuciusornithids were probably better fliers than *Archaeopteryx*. The large and curved toe claws, as well as the toe proportions, suggest that confuciusornithids were arboreal.

The holotype specimen of *Confuciusornis*, the first named confuciusornithid, was discovered in 1993 by an amateur fossil collector and described in 1995 as the earliest known bird to have possessed a horny beak. Unlike living birds, most Mesozoic birds possessed teeth in the jaws; however, all known members of Confuciusornithidae completely lost their teeth. In addition, confuciusornithids have some very specialized skeletal features, such as a hole near the upper end of the humerus (upper arm bone). Although there is as yet no direct evidence concerning the diet of these birds, they probably fed mainly on insects.

Confuciusornis is the only Mesozoic bird that is abundantly represented in the fossil record; at least several hundred specimens are known, due to the exceptional preservation of the fossil assemblages in which they occur. (These deposits, of alternating volcanic ash layers and lake sediments, also yield many other important fossils such as feathered dinosaurs, pterosaur embryos, early mammals, and flowering plants.) In most *Confuciusornis* specimens, the feathers are preserved, including asymmetric flight feathers that indicate powerful flight capability. *Confuciusornis* also displays evidence of sexual dimorphism in the tail feathers; it is generally believed that those with short tail feathers are females, whereas those with a pair of long central tail feathers are males (see **illustration**). In some rare cases, male and female individuals are preserved together on a single slab. The discovery of hundreds of *Confuciusornis* specimens from a relatively small area may also indicate that they were gregarious. Geological evidence, and the abundance of specimens in the same strata, suggests that flocks of these birds may have died in mass mortality events, possibly caused by eruptions of nearby volcanoes. *See* ARCHAEOPTERYX; AVES; FEATHER; FOSSIL. Zhonghe Zhou

Male *Confuciusornis* with a pair of long central tail feathers from the Early Cretaceous of Liaoning, northeastern China.

Bibliography. L. M. Chiappe et al., Anatomy and systematics of the Confuciusornithidae (Aves) from the Mesozoic of Northeastern China, *Amer. Mus. Novitates*, 242:1–89, 1999; A. Feduccia, *The Origin and Evolution of Birds*, 2d ed., Yale University Press, New Haven, 1999; L. Hou et al., A beaked bird from the Jurassic of China, *Nature*, 377:616–618, 1995; Z. Zhou, The origin and early evolution of birds: Discoveries, disputes, and perspectives from fossil evidence, *Naturwissenschaften*, 91:455–471, 2004.

Congenital anomalies

Structural abnormalities of the body that develop during embryogenesis and the fetal period; also called birth defects. Children with significant birth defects need more medical care than other children do, require more frequent hospitalizations, need

community support services, and often require special education programs. Among children, over half of all visits to subspecialty medical clinics and admissions to hospitals are for treatment of disorders resulting from errors in embryonic development, chromosomal abnormalities, and genetic and familial disorders. Two-thirds of the deaths of infants and children in pediatric hospitals in developed countries are caused by underlying congenital anomalies.

Incidence. Although any particular birth defect may be rare, birth defects as a group are common. They are identified in 2–3% of newborns and in about an equal percentage later, in the first year of life, as structural defects and development delays become evident.

The overall incidence of congenital anomalies is about the same throughout the world, but certain geographic and racial groups have higher rates of certain disorders. For example, those developmental defects of the brain and spinal cord known as neural tube disorders are much more common in the British Isles than in North America, occurring in 10 per 1000 live births in Ireland as compared with 1–2 per 1000 in Canada and the United States.

Screening. Screening programs are available to identify fetuses and newborns likely to have disorders such as congenital malformations and genetic diseases. Examples of screening include amniocentesis to detect fetal chromosomal abnormalities in mothers over 35 years of age; measurement of maternal serum alpha-fetoprotein levels at 15–16 weeks' gestation to help identify fetuses with certain malformations; and screening of newborns for phenylketonuria, sickle cell disease, thalassemia, galactosemia, and congenital hypothyroidism. *See* ALPHA FETOPROTEIN; PHENYLKETONURIA; SICKLE CELL DISEASE.

Causes. Birth defects can be caused by genetic factors, exposure to malformation-causing agents (teratogens), or a combination of both. It is estimated that single-gene disorders cause 7.5% of birth defects identified in newborns and structural abnormalities of chromosomes cause 6%. Maternal infections that have been transmitted to the embryo or fetus are responsible for 2–3% of congenital anomalies. Maternal diabetes is associated with 1.5% and other maternal diseases with less than 1.5%. Maternal use of drugs and alcohol is responsible in about 1–2% of cases. In more than 50% of newborns with congenital anomalies, however, the specific etiology is not known. *See* DIABETES; FETAL ALCOHOL SYNDROME; HUMAN GENETICS.

Dysmorphology is the area of medicine and science concerned with the cause of congenital anomalies resulting from errors in embryonic development (dysmorphogenesis). The study of normal and abnormal embryonic development allows identification of the latest time in embryogenesis when a malformation could occur. Examples of such times include 28 days of gestation for neural tube defects, 36 days for cleft lip, and 10 weeks for cleft palate. Because the heart is an embryologically complicated organ, different congenital heart defects develop at different times in embryogenesis. Abnormalities of digestive tract formation can occur as early as 30 days of gestation and as late as 10 weeks.

Classification. To facilitate the study, birth defects are divided into malformations, disruptions, deformations, and dysplasias.

Malformations. Malformations are structural defects that are caused by primary errors in morphogenesis. They are classified as major and minor. Major malformations require medical or surgical intervention or are of substantial cosmetic importance. Various defects that were mentioned above are among such malformations. Minor malformations do not require such treatment or do not greatly affect appearance; examples include protruding ears, wide-set eyes, curvature of the fifth finger, webbing of the toes, and many others. Most individuals have two or three minor malformations, and the same minor malformations are frequently found in various family members. An excess of minor malformations is often associated with major malformations.

Disruptions. These are structural defects resulting from interruption of normal morphogenesis, with consequent destruction of previously existing structures and incomplete development of tissues. For example, if the fetus becomes entangled in the amnion (the inner lining of the placental sac), ring constrictions of the limbs can occur.

Deformations. These are congenital anomalies resulting from external compression of a normally formed part of the fetus. The most common deformations are bowing of the lower legs and in-toeing of the feet. In otherwise normal infants, minor deformations usually improve without treatment. If deformations are major, splints and special shoes may be required for correction.

Dysplasias. These are disorders that result from an abnormal organization of cells into tissues; the morphological result is called dyshistogenesis. Examples of such disorders include hemangiomas and lymphangiomas, which result from overgrowth of blood vessels and lymphatic vessels, respectively. Neurofibromatosis is an autosomal dominant genetic disorder that results in overgrowth of neural tissue.

Syndromes. Some individuals have various major and minor congenital anomalies that together form a recognizable pattern, called a syndrome. The medical diagnosis of a specific congenital syndrome in a child aids in determining etiology, identifying associated anomalies that previously were unrecognized, and establishing prognosis. Often, further diagnostic studies are required to confirm the diagnosis of a syndrome. Syndromes can have both genetic and environmental causes.

Teratology. Tetratology is the study of the effect of environmental agents on the developing embryo and fetus. Teratogens are agents that interfere with normal embryonic development. They can cause miscarriages, retard prenatal growth, and produce congenital anomalies or mental retardation.

There are five general groups of teratogens: (1) infectious diseases and agents, including rubella, toxoplasmosis, syphilis, cytomegalovirus, chickenpox,

Venezuelan equine encephalitis, and human immunodeficiency virus; (2) physical agents, such as radiation, maternal hyperthermia, and abnormalities of the intrauterine environment; (3) drugs and chemical agents; (4) maternal metabolic and genetic factors, such as diabetes, epilepsy, smoking, and nutritional deficiencies; and (5) paternal factors, although rare. (Paternal exposures contribute mainly to decreased fertility rather than birth defects.)

Treatment, detection, and prevention. Treatment of congenital anomalies is specific for each individual. Individuals with severe or numerous abnormalities usually require multidisciplinary treatment, including such measures as medical management, surgical correction, nursing care, special diets, rehabilitation, prosthetic devices, special education, and community support.

If a couple has had a child with a congenital anomaly, tests may be done in subsequent pregnancies to help detect recurrence. Prenatal tests that are sometimes performed include chorionic villus sampling and amniocentesis to detect chromosomal and some metabolic defects, detailed ultrasound imaging of the fetus to identify significant structural malformations, and fetal echocardiograms to detect congenital heart disease.

Measures that help reduce the risk of having a child with congenital anomalies include avoidance of teratogenic exposures, medical treatment of maternal illnesses, good nutrition, and routine obstetrical care. However, much remains to be learned about the cause, detection, and prevention of congenital anomalies. *See* PREGNANCY. Margot I. Van Allen

Bibliography. K. L. Jones, *Smith's Recognizable Patterns of Human Malformation*, 5th ed., 1996; T. H. Shepard, *Catalog of Teratogenic Agents*, 9th ed., 1998; J. S. Thompson and M. W. Thompson, *Genetics in Medicine*, 1992.

Conglomerate

The consolidated equivalent of gravel. Conglomerates are aggregates of more or less rounded particles greater than 0.08 in. (2 mm) in diameter. Frequently they are divided on the basis of size of particles into pebble (fine), cobble (medium), and boulder (coarse) conglomerates. The common admixture of sand-sized and gravel-sized particles in the same deposit leads to further subdivisions, into conglomerates (50% or more pebbles), sandy conglomerates (25–50% pebbles), and pebbly or conglomeratic sandstones (less than 25% pebbles). The pebbles of conglomerates are always somewhat rounded, giving evidence of abrasion during transportation; this distinguishes them from some tillites and from breccias, whose particles are sharp and angular (see **illus.**).

Conglomerates fall into two general classes, based on the range of lithologic types represented by the pebbles, and on the degree of sorting and amount of matrix present. The well-sorted, matrix-poor conglomerates with homogeneous pebble lithology are one type, and the poorly sorted, matrix-rich conglomerates with heterogeneous pebble lithology are the other. The well-sorted class includes quartz-pebbles, chert-pebble, and limestone-pebble conglomerates. The quartz pebbles were derived from long, continued erosion of the source-rock terrain that resulted in the disappearance of all unstable minerals, such as feldspars and ferromagnesian minerals, that make up most igneous and metamorphic rocks. Typically the quartz-pebble conglomerates are thin sheets that overlie an unconformity and are basal to a series of overlapping marine beds. Chert pebble conglomerates are derived from weathered limestone terrains. Limestone pebble conglomerates seemingly are the result of special conditions involving rapid mechanical erosion and short transport distances; otherwise the limestone would quickly become degraded to finer sizes and dissolve. The well-sorted conglomerates tend to be distributed in thin, widespread sheets, normally interbedded with well-sorted, quartzose sandstones. *See* UNCONFORMITY.

The poorly sorted, lithologically heterogeneous conglomerates include many different types, all related in having very large amounts of sandy or clayey matrix and pebbles of many different rock classes. The graywacke conglomerates are the outstanding representatives. They are composed of pebbles, many times only slightly rounded, or many different kinds of igneous and metamorphic rock as well as sedimentary rocks, bound together by a matrix that is a mixture of sand, silt, and clay. They seem to have been formed from the products of rigorous mechanical erosion of highlands and transported in large part by turbidity or density currents.

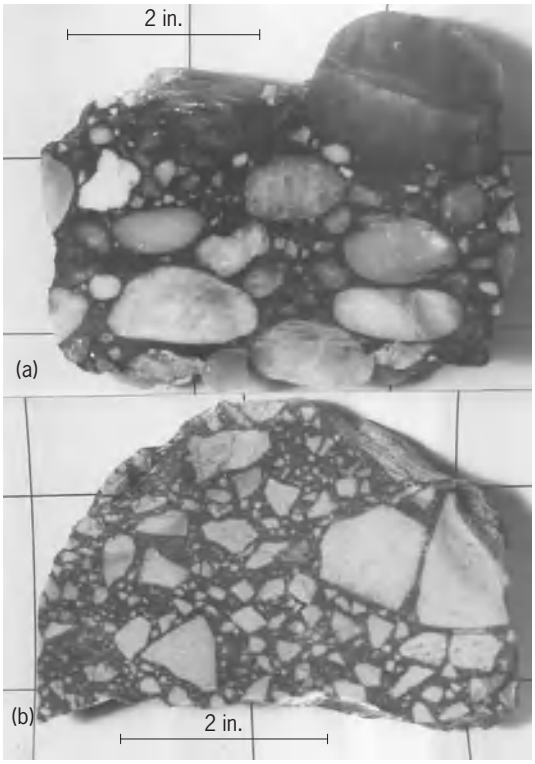

Lithified gravels. (*a*) Conglomerate, composed of rounded pebbles. (*b*) Breccia, containing many angular fragments. 2 in. = 5 cm. (*Specimens from Princeton University Museum of Natural History; photo by Willard Starks*)

Another representative of the poorly sorted class is fanglomerate, a conglomerate formed on an alluvial fan. Fanglomerates in many cases are much better sorted than the graywacke conglomerates but have heterogeneous pebble composition. A tillite is another representative of this class of conglomerate. All of these poorly sorted conglomerates tend to occur in fairly thick sequences, and some of them, typically the fanglomerates, are wedge-shaped accumulations. *See* SEDIMENTARY ROCKS; TURBIDITY CURRENT.

Special types of conglomerates, such as volcanic conglomerates and agglomerates and some intraformational conglomerates composed of shale pebbles or deformed limestone pebbles, do not seem to fall easily into either class. Some of the shale and limestone intraformational conglomerates have formed from the tearing up of previously deposited beds at the sea bottom while they were still relatively soft and only partially consolidated. Because of the angularity of some of these pebbles, they have often been called breccias. *See* BRECCIA; GRAVEL; GRAYWACKE; TILL.

Raymond Siever

Conglutination

A term used in serology to describe the completion or enhancement of an incomplete agglutinating system by the addition of certain substances. Some bacteria or erythrocyte suspensions do not exhibit the visible agglutination ordinarily expected after they have been coated with their specific antibodies and complement. Further addition of a conglutinating agent—normal bovine serum—however, initiates visible agglutination. Complement is an essential component of the system. The conglutination reagent is itself without effect in the absence of antibody and complement. Similar overall actions, such as agglutination of Rh^+ cells, occur when human serum, gelatin, or bovine serum albumin are added to the incomplete Rh antibodies found in some sera, but since this enhancing effect occurs also in the absence of complement, these reactions probably are to be distinguished in mechanism from that of the traditional conglutination. *See* BLOOD GROUPS.

Like the related complement-fixation reaction, the conglutination reaction can be used for a variety of serological diagnostic reactions in bacterial and viral infections. *See* AGGLUTINATION REACTION; COMPLEMENT-FIXATION TEST; SEROLOGY.

Henry P. Treffers

Bibliography. C. W. Bennett, *Clinical Serology*, 1977; N. J. Bryant, *Laboratory Immunology and Serology*, 3d ed., 1992; R. R. A. Coombs, A. M. Coombs, and D. G. Ingram, *The Serology of Conglutination and Its Relation to Disease*, 1961.

Conic section

One of the class of curves in which a plane may cut a cone (surface) of revolution. They were extensively studied by the ancient Greeks. The section is a parabola if the plane is parallel to an element of the cone, an ellipse or circle if the plane cuts all elements of one nappe (but does not go through the apex), and a hyperbola if the plane cuts elements of both nappes (for example, the plane parallel to the cone's axis of revolution) and does not go through the apex (see **illus.**) If a line g intersects the cone's axis perpendicularly at a point distinct from the apex, a plane that revolves about g will cut all three kinds of conics from the cone. If a plane contains the cone's axis, it intersects the cone in a pair of lines (a degenerate hyperbola); if it goes through the apex but does not contain the axis, it has in common with the cone either the apex alone (degenerate ellipse) or an element of the cone (degenerate parabola).

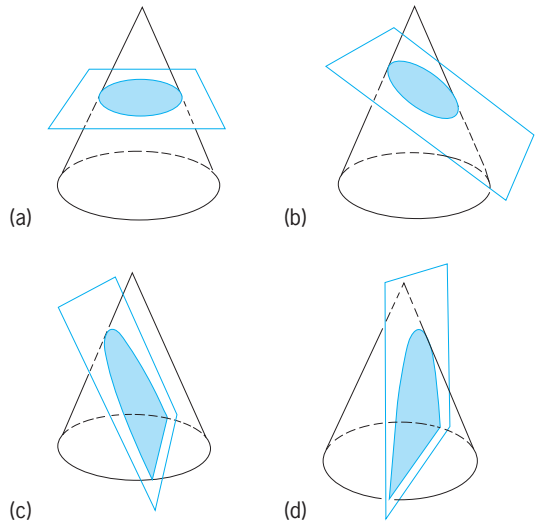

Four conic sections. (*a*) Circle. (*b*) Ellipse. (*c*) Parabola. (*d*) Hyperbola.

Algebraic definitions. After the advent of analytic geometry, the synthetic method used by the Greeks to develop the properties of conics gave way to algebraic procedures which, using other definitions, divorced these curves from their original relationship to cones and regarded them as graphs of second-degree equations in cartesian coordinates x, y. Let plane π cut all the elements of one nappe of a cone C (π not perpendicular to the cone's axis and not passing through the apex), and let E denote the curve of intersection of π and C. To show that E is an ellipse, consider the spheres S_1 and S_2 inside C that are tangent to π and to all the elements of C and that lie on the opposite sides of π. The contact points of S_1 and S_2 with C form circles whose planes π_1 and π_2 are perpendicular to the axis of C. Let F_1 and F_2 denote the contact points of S_1 and S_2, respectively,with π. If P is any point of E, then Eqs. (1) holds, where Q_1, Q_2

$$\text{dist}\,(P, F_1) = \text{dist}\,(P, Q_1) \qquad (1a)$$

$$\text{dist}\,(P, F_2) = \text{dist}\,(P, Q_2) \qquad (1b)$$

are those contact points of S_1, S_2, respectively, that lie on the element of C through P. Since dist (P,Q_1) + dist (P,Q_2) is the (constant) distance between the parallel planes π_1, π_2, then Eq. (2) holds, and so E is

$$\text{dist}\,(P, F_1) + \text{dist}\,(P, F_2) = \text{constant} \qquad (2)$$

an ellipse. The directrices of E are the lines in which π_1, π_2 intersect π. *See* ANALYTIC GEOMETRY.

Conics are basic in astronomy, where it is shown that the newtonian inverse-square law of gravitation implies that the orbits of the planets are ellipses, with the Sun at one focus. Other heavenly bodies, certain comets, for example, have parabolic and hyperbolic paths.

Projective geometry. The third phase in the study of conic sections began with the development of projective geometry. In that subject it is shown that the locus of the points of intersection of two projective (nonperspective), coplanar (nonconcentric) pencils of lines is a nondegenerate (point) conic. All nondegenerate conics are projectively equivalent; that is to say, if C_1, C_2 are any two such conics, there is a projective transformation of their plane into itself that transforms C_1 into C_2. Conics may be classified into types by their behavior with respect to an arbitrarily selected line g, say the ideal line of the plane. In the geometry (affine) obtained by considering only those projective transformations that transform g into itself, a conic is a hyperbola, a parabola, or an ellipse accordingly as the number of points that the conic has in common with g is two, one, or none, respectively. Dual to point conics are line conics, the locus of all lines joining corresponding points of two projective (not perspective), coplanar (not collinear) point ranges. The tangents to a point conic form a line conic, and the contact points of a line conic form a point conic. One of the most famous theorems concerning conics was proved by the French philosopher-mathematician Blaise Pascal in 1640, when only 16 years old. The theorem asserts that if six points A_i, with $i = 1, 2, \ldots, 6$, are on a conic, then the intersections P,Q,R of the lines (A_1, A_2) and (A_4, A_5), (A_2, A_3) and (A_5, A_6), (A_3, A_4) and (A_6, A_1), respectively, lie on a line. This theorem, together with its converse (which is also valid) had numerous interesting consequences. The Brianchon theorem, proved 166 years later, is its dual. It states that if a hexagon circumscribes a conic, the three lines joining the three pairs of opposite vertices are concurrent. *See* CONE; ELLIPSE; HYPERBOLA; PARABOLA; PROJECTIVE GEOMETRY. Leonard M. Blumenthal

Bibliography. G. Fuller and D. J. Tarwater, *Analytic Geometry*, 7th ed., 1993; L. E. Garner, *Outline of Projective Geometry*, 1981; D. F. Riddle, *Analytic Geometry*, 6th ed., 1996; P. Samuel, *Projective Geometry*, 1988.

Conjugation and hyperconjugation

A higher-order bonding interaction between electron orbitals on three or more contiguous atoms in a molecule, which leads to characteristic changes in physical properties and chemical reactivity. One participant in this interaction can be the electron pair in the π-orbital of a multiple (that is, double or triple) bond between two atoms, or a single electron or electron pair or electron vacancy on a single atom. The second component will be the pair of π-electrons in an adjacent multiple bond in the case of conjugation, and in the case of hyperconjugation it will be the pair of electrons in an adjacent polarized σ-bond (that is, a bond where the electrons are held closer to one atom than the other due to electronegativity differences between the two atoms). *See* CHEMICAL BONDING.

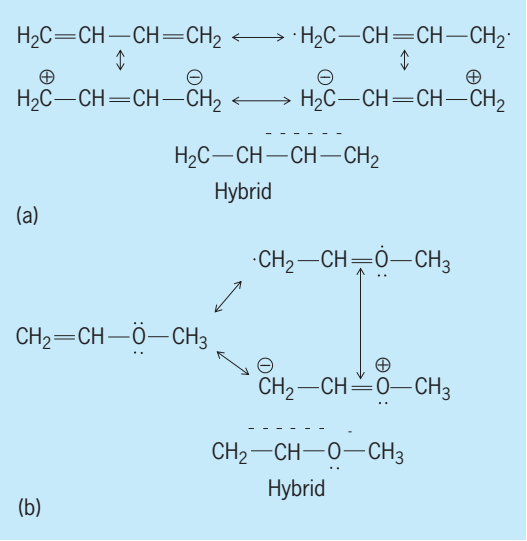

Fig. 1. Conjugated molecules. Broken overbars indicate the effects of conjugation. (*a*) 1,3-Butadiene. (*b*) Methyl vinyl ether.

Conjugation. The conjugated orbitals reside on atoms that are separated by a single bond in the classical valence-bond molecular model, and the conjugation effect is at a maximum when the axes of the component orbitals are aligned in a parallel fashion because this allows maximum orbital overlap. Conjugation thus has a stereoelectronic requirement, or a restriction on how the participating orbitals must be oriented with respect to each other. Two simple examples are shown in **Fig. 1**; in 1,3-butadiene ($H_2C{=}H\underline{C{-}C}H{=}CH_2$) conjugation occurs between the p orbitals (π-bonds) of the two double bonds, and in methyl vinyl ether ($H_2C{=}H\underline{C{-}O}{-}CH_3$) the nonbonding sp^3 orbital on oxygen is conjugated with the p orbitals of the double bond. This interaction is manifest in an effective bond order between single and double for the underlined single bond.

If the component orbitals are orthogonal (at right angles) to each other, or if the atoms bearing these orbitals are separated by more than one single bond or by no bond, the molecule is not conjugated. Two examples of molecules which do not exhibit conjugation are 1,4-pentadiene ($H_2C{=}CH{-}CH_2{-}CH{=}CH_2$) and allene ($H_2C{=}C{=}CH_2$), the former because there is more than one single bond between the component orbitals and the latter because the component orbitals are orthogonal to each other and are not separated by a single bond.

Chemists have used the conjugation rationale since the midnineteenth century, and it has become well accepted. The valence-bond model of conjugation invokes delocalization: the participating

electrons are no longer localized or fixed on a particular atom or between a pair of atoms, but are shared throughout the conjugated orbitals. This increases their entropy, which contributes to a lower overall energy and generally results in greater stabilization for the molecule. From the valence-bond point of view, conjugated molecules are viewed as a weighted average hybrid of two or more valence-bond structures as shown in Fig. 1 for the two examples. Thus butadiene has a major contribution from the standard valence-bond structure, $H_2C{=}CH{-}CH{=}CH_2$, but the diradicals or the structures with separation of charge are valid alternative representations, and the same is true for methyl vinyl ether. The effects of conjugation are suggested by the broken lines on the hybrid structures in Fig. 1. When more orbital units are linked together or conjugated, more resonance structures can be drawn, and the molecule will be more stable and behave less according to the expectation based on the simple structure drawn from the classical valence-bond model. *See* DELOCALIZATION; ENTROPY; RESONANCE (MOLECULAR STRUCTURE); STRUCTURAL CHEMISTRY.

Three of the four resonance contributors for butadiene are drawn with only one multiple bond compared to the two multiple bonds in the standard valence-bond structure (Fig. 1*a*); this is known as sacrificial conjugation, in that some bonds in the standard depiction are sacrificed. In methyl vinyl ether (Fig. 1*b*) and in molecules or chemical intermediates, the different pictures drawn for the resonance contributors have the same number of multiple bonds; this is known as isovalent conjugation and results in a greater resonance energy stabilization than in the case of sacrificial conjugation. Other examples of isovalent conjugation include the allyl cation ($CH_2{=}CH{-}CH_2^+$ and $C^+H_2{-}CH{=}CH_2$) and benzene.

Conjugation is better understood by considering the molecular orbitals which can be approximated by the mathematical combination of the wave functions for the constituent atomic *s* and *p* orbitals (and *d* orbitals for elements beneath the first row of the periodic table). **Figure 2** shows simple pictorial results of complex molecular orbital (MO) calculations for the π-electrons in butadiene; here the σ bonds are neglected for clarity. The combination of four atomic *p* orbitals (two each for each of the double bonds) leads to four molecular orbitals. Each of the molecular orbitals is a quartet of "*p*-like" orbitals of different size (determined by calculation) situated on the four carbon atoms of butadiene (C^1–C^4). The four π-electrons are placed in these orbitals from the bottom up, so that the two lowest-energy orbitals, MO #1 and MO #2, both contain paired electrons. These two molecular orbitals thus account for the π-bonding in butadiene. In MO #1 the continuous bonding across all four lobes can be seen, with the highest orbital density between C^2 and C^3. In MO #2 there is a phase change between C^2 and C^3, with consequent repulsion between the adjacent lobes, but this repulsion is relatively less important because the electron orbital density (depicted as orbital size) is much lower here. In both of these molecular orbitals, the bonding interactions between $C^1{-}C^2$ and $C^3{-}C^4$ are in phase and add up to strong π-bonds. Thus, the π-delocalization interaction across $C^2{-}C^3$ that characterizes conjugation appears very naturally from the molecular orbital description. *See* MOLECULAR ORBITAL THEORY.

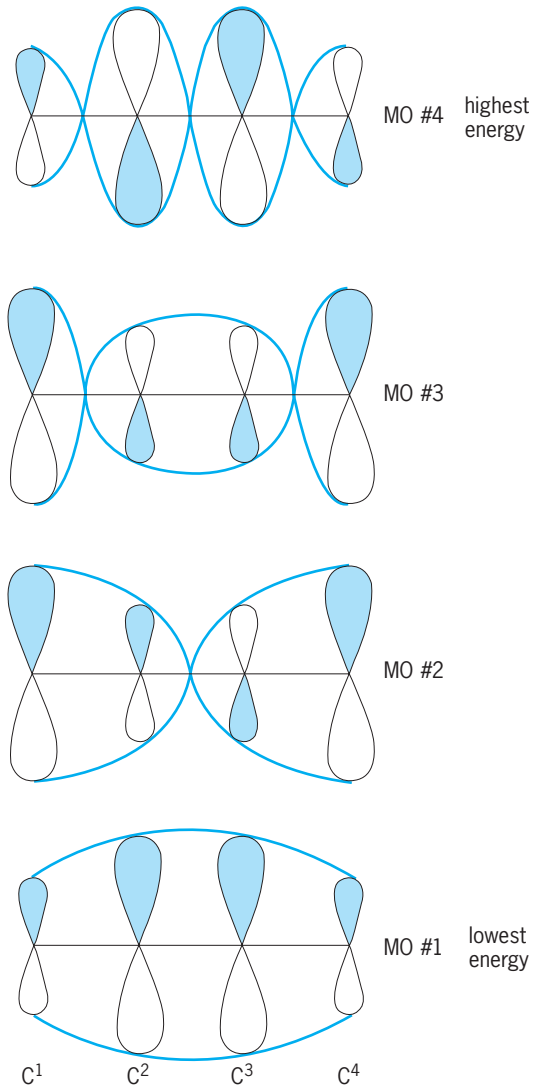

Fig. 2. Molecular orbital (MU) diagram for butadiene as determined by calculation. Algebraic phase is indicated by the presence or absence of tint. Orbital density is depicted as orbital size.

Hyperconjugation. This concept posits the conjugation of polarized σ-bonds with adjacent π-orbitals, and was introduced in the late 1930s by R. S. Mulliken. This rationale was used to explain successfully a wide variety of chemical phenomena; however, the confusing adaptation of the valence-bond model necessary to depict it and its inappropriate extension to some phenomena led to difficulties. The advent of the molecular orbital treatment has eliminated many of these difficulties.

Early on, hyperconjugation was used to explain the stabilization by alkyl groups of carbocations, or positively charged trivalent carbon. **Figure 3** shows how the orbitals of the σ-bonds of a methyl group (H_3C) exert a hyperconjugative stabilizing

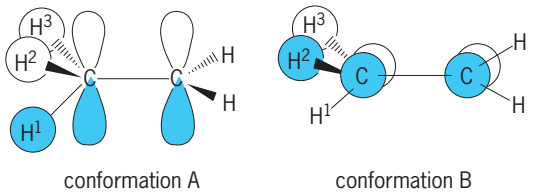

Fig. 3. Diagram showing hyperconjugative effect of orbitals of the σ-bonds of a methyl group (H_3C) on a neighboring *p* orbital of a methylene group (CH_2). Algebraic phrase is indicated by the presence or absence of tint.

effect upon a neighboring *p* orbital on a methylene group (CH_2). The bonding molecular orbitals of the methyl group utilize aspects of its carbon *p* orbitals; in both of the conformations or rotations the methyl group has such an orbital which overlaps in phase with the neighboring *p* orbital on CH_2. As hydrogen is an electropositive atom, the methyl group acts as an electron donor to stabilize an empty *p* orbital and to destabilize a filled *p* orbital on the right-hand carbon. This is in accord with empirical observations. If H^1 is replaced with some other atom X, the hyperconjugative effect of X will be at a maximum in conformation A, where there is orbital density on X, and it will be at a minimum in conformation B, where X is in the nodal or null plane. If X is a more electropositive atom such as silicon, the C-X bond is a donor bond and will stabilize an adjacent empty *p* orbital, such as in a cation. If X is a more electronegative element such as fluorine, the C-X bond is an acceptor bond and will stabilize an adjacent filled *p* orbital such as in an anion; this result accords with experiment as well. The donor and acceptor effects are also important where the single *p* orbital is replaced by a multiple bond. *See* REACTIVE INTERMEDIATES.

Physical properties and chemical reactivity. The most notable differences in physical properties attributed to conjugation are as follows: (1) The single bonds lying between the two orbital components are shortened. For example, the length of the C^2-C^3 bond in butadiene is 0.148 nanometer, as compared to 0.154 nm for the carbon-carbon single bond in ethane. This shortening can also be explained by the changes in orbital hybridization at these carbons. (2) Electronic absorption spectra begin at significantly longer wavelengths for conjugated than for related unconjugated molecules. (3) Ionization potentials are lower than for isomeric unconjugated molecules. (4) Related to points 2 and 3 is the fact that polarizabilities are larger for conjugated than for the corresponding unconjugated dienes. Polarizability is the ease with which the electron orbitals (electron clouds) of the molecule can be distorted through dipole-dipole interactions with other molecules, ions, atoms, or electrons; it is generally a measure of how loosely the electrons are held by the nuclei. (5) As compared with predictions from formulas for unconjugated molecules, conjugated molecules show lower energies, for example, about 6 kcal/mol (25 kilopascals/mol) lower for 1,3-butadiene than for unconjugated dienes; the excess is describable theoretically as resonance energy (valence-bond picture) or delocalization (molecular orbital theory picture). (6) The magnitude and orientation of dipole moments can be derived from hyperconjugation arguments, although the aforementioned differences in hybridization can also explain this. *See* BOND ANGLE AND DISTANCE; DIPOLE MOMENT.

In terms of changes in chemical reactivity, the following are exemplary: (1) The acidity of carboxylic acids stems from the conjugation stabilization of their ionized forms. (2) The addition of many chemical reactants takes place across the conjugated system (conjugate addition) rather than at the isolated double bonds, for example, the addition of bromine to butadiene to give a mixture of 1,4-dibromo-2-butene and 1,2-dibromo-3-butene. (3) The well-recognized orientation rules for the addition of chemical reagents to substituted benzenes and other aromatic systems can be rationalized through conjugated intermediates. (4) The facile formation of an anion from cyclopentadiene and of a cation from cycloheptatriene derives from the conjugation stabilization for those ions. (5) The chemistry of unsaturated organosilicon compounds can be explained only by hyperconjugation effects. (6) Certain changes in chemical reaction rates that occur when hydrogen atoms are substituted by deuterium atoms are clearly hyperconjugative effects.

A number of scientific phenomena depend on the properties of conjugated systems; these include vision (the highly tuned photoreceptors are triggered by molecules with extended conjugation), electrical conduction (organic semiconductors such as polyacetylenes are extended conjugated systems), color (most dyes are conjugated molecules designed to absorb particular wavelengths of light), and medicine (a number of antibiotics and cancer chemotherapy agents contain conjugated systems which trap enzyme sulfhydryl groups by conjugate addition). *See* VALENCE.

Matthew F. Schlecht

Bibliography. I. Fleming, *Frontier Orbitals and Organic Chemical Reactions*, 1976; J. March, *Advanced Organic Chemistry: Reactions, Mechanisms, and Structures*, 4th ed., 1992; R. S. Mulliken, C. A. Rieke, and W. G. Brown, Hyperconjugation, *J. Amer. Chem. Soc.*, 63:41, 1941; V. F. Traven, *Frontier Orbitals and Properties of Organic Molecules*, 1992.

Connecting rod

A link in several kinds of mechanisms. Usually one end of a connecting rod is intended to follow a circular path, while the other end follows a path along a straight line or a curve of large radius. The term is sometimes applied, however, to any straight link that transmits motion or power from one linkage to another within a mechanism. **Figure 1** shows conventional arrangements of connecting rods in typical mechanisms. In some applications (for example, the connecting rod between the crank and an overhead oscillating member or walking beam in a well-drilling

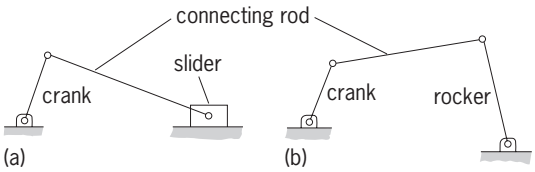

Fig. 1. Connecting rod in (*a*) slider-crank mechanism and (*b*) crank-and-rocker mechanism.

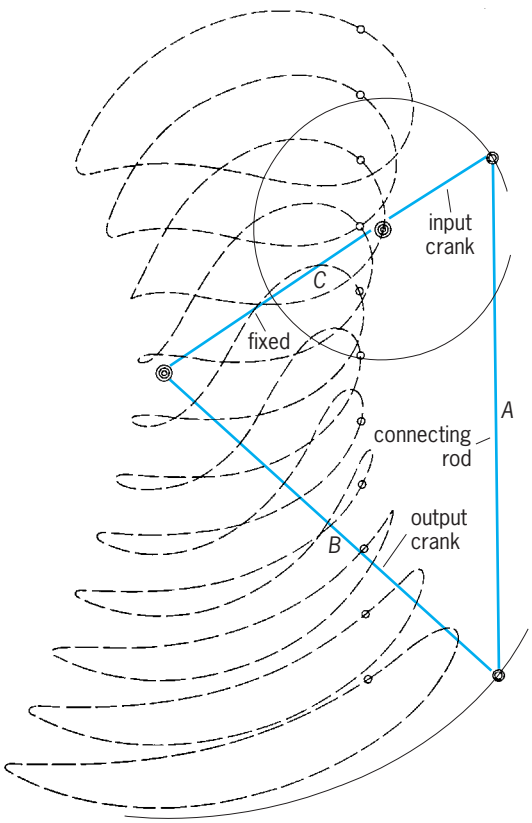

Fig. 2. Paths of points on connecting-rod plane traced by four-bar crank-and-rocker mechanisms for $A = 4.0$, $B = 3.5$, $C = 2.0$. (*After J. A. Hrones and G. L. Nelson, Analysis of Four-Bar Linkage, John Wiley and Sons, 1951*)

rig, or between the steering column and cross-links in an automobile) the connecting rod is called a pitman. *See* AUTOMOTIVE STEERING.

The connecting rod of the four-bar linkage, often called the coupler, has special significance. The motion of its plane can now be synthesized to furnish desired paths for points not necessarily on the straight line *AB* (coupler-point paths), or desired positions of the entire connecting-rod plane. The connecting rod is then not primarily used for transmission of force or motion from input to output crank, but the entire mechanism is employed to impart to the connecting-rod plane certain displacements, and sometimes velocities and accelerations. **Figure 2** illustrates paths traced by various points in the plane of the rod, also revealing some velocities and accelerations there. *See* FOUR-BAR LINKAGE; MECHANISM.

Douglas P. Adams

Bibliography. E. A. Dijksman, *Motion Geometry of Mechanisms*, 1976; G. H. Martin, *Kinematics and Dynamics of Machines*, 2d ed., 1982; A. Ramous, *Applied Kinematics*, 1972.

Connective tissue

One of the four primary tissues of the body. It differs from the other three tissues in that the extracellular components (fibers and intercellular substances) are abundant. It cannot be sharply delimited from the blood, whose cells may give rise to connective tissue cells, and whose plasma components continually interchange with and augment the ground substance of connective tissue. Bone and cartilage are special kinds of connective tissue. *See* BLOOD.

Functions. The functions of connective tissues are varied. They are largely responsible for the cohesion of the body as an organism, of organs as functioning units, and of tissues as structural systems. This cohesive function is achieved through their permeation of other tissues of the body. The connective tissues are essential for the protection of the body both in the elaborate defense mechanisms against infection and in repair from chemical or physical injuries. Nutrition of nearly all cells of the body and the removal of their waste products are both mediated through the connective tissues.

Connective tissues are important in the development and growth of many structures. Constituting the major environment of most cells, they are probably the major contributor to the homeostatic mechanisms of the body so far as salts and water are concerned. They act as the great storehouse for the body of salts and minerals, as well as of fat. The connective tissues determine in most cases the pigmentation of the body. Finally, the skeletal system (cartilage and bones) plus other kinds of connective tissue (tendons, ligaments, fasciae, and others) make motion possible.

Cellular components. The connective tissues consist of cells and extracellular or intercellular substance (see **illus.**). The cells include many

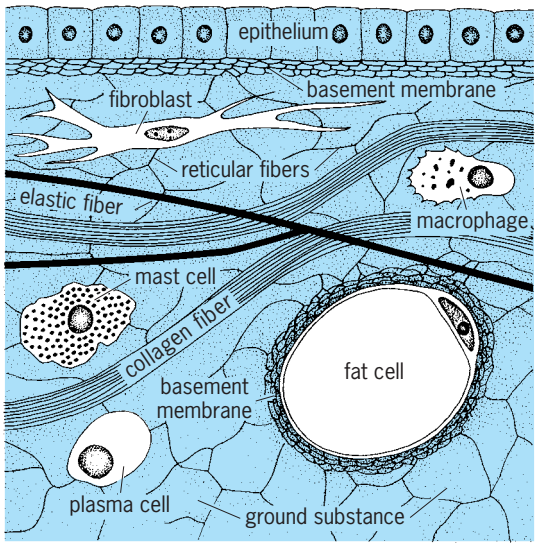

Components of connective tissue.

varieties, of which the following are the most important: (1) fibroblasts—variable in shape, frequently outstretched to form many fine processes, some as small as 20–30 nanometers; (2) macrophages (histiocytes)—variable in shape, strongly motile or outstretched, frequently with numerous minute cell processes, highly phagocytic (the ability to take particulate material into the cytoplasm); (3) mast cells—frequently large, ameboid, nucleus centrally located, cytoplasm very often packed with granules about 0.4 micrometer in diameter; (4) plasma cells—generally ovoid, weakly ameboid, nucleus eccentric, granular cytoplasm staining strongly with basic dyes (toluidine blue, methylene blue, hematoxylin); (5) melanocytes—cytoplasm of highly branched cells filled with minute granules which appear yellow or brown because of the melanin they contain; and (6) fat cells of two kinds: (a) yellow fat cells—very large (80 μm or more), generally spherical, with a thin shell of protoplasm (3.0–0.1 μm or less) enclosing a single enlarged fat droplet which appears yellowish; and (b) brown fat cells—moderately large, generally spherical, with small droplets of variable size (about 2 μm or less), scattered in the cytoplasm. In addition, connective tissues may contain, especially in the vicinity of small blood vessels, undifferentiated mesenchymal cells which may be considered as potentially embryonic, and blood cells which have wandered through the walls of blood vessels, including lymphocytes and polymorphonuclear granular leukocytes. It is especially important to realize that most of the cells of the connective tissue are developmentally related even in the adult; for example, fibroblasts may be developed from histiocytes or from undifferentiated mesenchymal cells.

Extracellular components. The extracellular components of connective tissues may be fibrillar or nonfibrillar. The fibrillar components are (1) reticular fibers which may be stained differentially with certain silver stains; (2) collagenous fibers which have a longitudinally striated appearance with the light microscope and are readily stainable with the acid dyes cosin and aniline blue; and (3) elastic fibers which are highly refractile, appear slightly yellowish, and may be stained differentially by certain dyes like orcein. With the higher resolution achieved with the electron microscope, both reticular and collagenous fibers are seen to be made up of smaller fibrils of about 20–100 nm in diameter, each with an internal cross-banded structure which is repeated on an average of every 64 nm. The unit of structure is tropocollagen. This is an elongated protein molecule about 300 nm long and about 1.3 nm in diameter, and is made up of three subunits. These are wound spirally around each other in very specific manners. The elastic fibers are also resolvable into finer fibrils which may be 8 nm or larger in diameter. All fibers visible with the light microscope are thus composed of bundles of submicroscopic fibrils visible only with the electron microscope. In addition, there are many collagenous and elastic fibers composed of the same fibrillar constituents which are too fine to be visible with the light microscope (fibers that are less than 0.2 μm in diameter). *See* COLLAGEN.

The nonfibrillar component of connective tissues appears amorphous with the light microscope and is the matrix in which cells and fibers are embedded. It consists of two groups of substances: (1) those probably derived from secretory activity of connective tissue cells including mucoproteins, protein-polysaccharide complexes, tropocollagen, and antibodies; and (2) those probably derived from the blood plasma, including albumin, globulins, inorganic and organic anions and cations, and water. In addition, the ground substance contains metabolites derived from, or destined for, the blood.

The ground substance of local origin and the tissue fluid of hematogenous origin are not only coextensive but are dynamically related in a two-phase system. The two phases of this thermodynamic system are probably recognizable in the electron microscope as a series of small, watery vacuoles (about 60 nm in diameter) embedded in a denser matrix. The basement membrane is a specially differentiated sheetlike part of the ground substance (10–50 nm thick) which underlies directly many epithelia, and encloses muscle fibers and fat cells. It is denser than the adjacent ground substance, with which it is continuous, and may contain fibrils. There is some evidence that epithelial cells may contribute some constituents to the basement membrane.

Classification. All the manifold varieties of connective tissue may contain all the cells and fibers discussed above in addition to ground substance. They differ from each other in the relative occurrence of one or another cell type, in the relative proportions of cells and fibers, in the preponderance and arrangement of one or another fiber, and in the relative amount and chemical composition of ground substance. In the classification which follows, some of these differences are pointed out.

1. Irregularly arranged connective tissue—which may be loose (subcutaneous connective tissue) or dense (dermis). The dominant fiber type is collagen, though others may be present. The fibers are irregularly interwoven and the cells irregularly distributed. The looser type has more ground substance than the denser type.

2. Regularly arranged connective tissue—primarily collagenous—with the fibers arranged in certain patterns depending on whether they occur in tendons or as membranes (dura matter, capsules, fasciae, aponeuroses, or ligaments). *See* TENDON.

3. Mucous connective tissue—ground substance especially prominent (umbilical cord).

4. Elastic connective tissue—predominance of elastic fibers or bands (ligamentum nuchae) or lamellae (aorta). *See* LIGAMENT.

5. Reticular connective tissue—fibers mostly reticular, moderately rich in ground substance, frequently numerous undifferentiated mesenchymal cells. In contrast with collagenous fibers which are unbranched, reticular fibers are branched and may be selectively stained by silver methods. These branched fibers, however, consist of submicroscopic

fibrils which have the same submicrostructure as unbranched collagen fibrils.

6. Adipose connective tissue—yellow or brown fat cells constituting chief cell type, reticular fibers most numerous. *See* ADIPOSE TISSUE.

7. Pigment tissue—melanocytes numerous.

8. Cartilage—cells exclusively of one type, derived from mesenchymal cells. The latter persist even in the adult in the perichondrium which encloses cartilage. The extracellular matrix is of three types: (1) hyaline (joint cartilage), containing no bundles of fibers; (2) fibrocartilage (patella), containing numerous coarse bundles of collagen fibers; and (3) elastic cartilage (external ear), containing elastic fibers. Hyaline cartilage does, however, contain collagen fibrils of submicroscopic dimensions. In all three types of cartilage, the fibrils and fibers are embedded in a homogeneous matrix, which is almost entirely composed of protein-polysaccharide complex, whose predominant polysaccharide is chondroitin sulfate. *See* CARTILAGE.

9. Bone—cells predominantly osteocytes, but also include fibroblasts, mesenchymal cells, endothelial cells, and osteoclasts. The fibers are almost entirely collagenous in nature, and may be highly organized into osteones. The fibers are embedded in a ground substance matrix. The outstanding property of bone is its hardness, due to deposition in the extracellular regions of many minute crystals of a calcium mineral, apatite. *See* BONE.

In some forms of connective tissue (fat), blood capillaries are extremely numerous. In others (cartilage and dentine), blood capillaries are lacking. Tendon is nearly avascular, and bone is also poorly supplied with blood capillaries. But all these different kinds of connective tissue must be supplied with nutrients and oxygen and must be drained of metabolic waste products in order to maintain the cells, however sparse, in a functional condition, and to effect turnover of extracellular components. How an adequate nutritive state is maintained is a great puzzle, and is attracting the attention of research workers. *See* HISTOLOGY. Isidore Gersh

Bibliography. G. Bevelander and J. A. Ramaley, *Essentials of Histology,* 8th ed., 1981; L. Weiss and L. Lansing (eds.), *Histology: Cell and Tissue Biology,* 5th ed., 1983.

Connective tissue disease

Any of a group of diseases involving connective tissue; formerly termed collagen vascular disease. These diseases are clinically and pathologically discrete but have overlapping features. The group includes lupus erythematosus, systemic vasculitis, scleroderma (and systemic sclerosis), polymyositis, and Sjögren's syndrome. Each of the diseases can involve multiple organ systems and is often coupled with various immunologic abnormalities. *See* CONNECTIVE TISSUE.

Lupus erythematosus. This set of diseases includes limited, primarily cutaneous disorders (discoid lupus and subacute cutaneous lupus) and a diffuse systemic illness (systemic lupus erythematosus), all of which are of unknown cause. Also, a lupuslike reaction known as drug-induced lupus can be caused by certain therapeutic agents.

Systemic lupus erythematosus. This is an inflammatory, multisystem disorder in which tissue injury is mediated by immune complexes. In this incompletely understood disorder, lymphocytes known as T cells lose some of their control over the antibody production by lymphocytes known as B cells. Thus, in most individuals with systemic lupus, antibodies are found in the blood serum that are not normally present.

The immunologic events of systemic lupus are influenced by multiple factors. In some families, a genetic predisposition seems to be present. That the disease sometimes flares up after excessive exposure to ultraviolet light emphasizes the importance of environmental factors. Finally, hormonal factors also seem to play a contributing role; eight to nine times as many young women compared with men have systemic lupus; and in some animal models of lupus, femaleness is associated with more severe disease. *See* CELLULAR IMMUNOLOGY.

Although systemic lupus can be an acute and fulminating illness, it is far more often a chronic disorder with long periods (years, sometimes even decades) of remission and good health and with infrequent flareups. With the recognition of the immune markers (that is, autoantibodies) in serum, milder forms of the disorder can now be identified.

The manifestations and course of systemic lupus vary greatly from one individual to another and even in one person over time. Fever and other constitutional symptoms (for example, fatiguability, lack of appetite, and weight loss) are prominent. Skin rashes are common, and the so-called butterfly rash over the nose and cheeks is most characteristic. Joint pain and inflammation (arthritis) are among the common features and often are the first manifestations of the disease. Many other parts of the body—such as the lungs, heart, skeletal muscles, kidneys, and nervous system—can be involved. Small blood vessels at any site may be inflamed. Hematologic abnormalities also can occur when antibodies reactive with cell membrane antigens are present. Anemia and low levels of white blood cells and platelets are common. Certain autoantibodies affect blood clotting and can be associated with clots in blood vessels anywhere in the body. *See* ARTHRITIS; HEMATOLOGIC DISORDERS.

The diagnosis of systemic lupus is based on a combination of clinical and laboratory findings; no single clinical feature or laboratory test makes the diagnosis. Even the characteristic immune markers in serum (that is, the autoreactive antibodies) are not specific. For example, the routine screening test for antinuclear antibodies (ANA), which is positive in 95% of patients with systemic lupus erythematosus, can also be positive in those with the other connective tissue disorders and certain other diseases, as well as some healthy family members of patients with systemic lupus. Discoid lupus and subacute cutaneous lupus are variants, primarily with skin involvement.

Photosensitivity is prominent. Systemic features are infrequent and milder than in systemic lupus erythematosus, and kidney and nervous systems are usually spared. ANA positivity is often present.

Drug-induced lupus. This form subsides when the offending agent is discontinued. Many drugs that are chemically unrelated to each other can produce the disorder. Genetic factors play a role in susceptibility to induction of lupus by at least some drugs.

Cause. Except for drug-related lupus, the cause of systemic lupus and its variants is not known; thus, a rational basis for specific curative therapy does not exist. However, the clinical manifestations of systemic lupus can generally be managed effectively for long years or even decades. Prognosis is steadily improving. Drugs that are used to control the disease by suppressing inflammation and altering the immune system include nonsteroidal anti-inflammatory drugs, an antimalarial agent (hydroxychloroquine), corticosteroids, and immunosuppressive agents. Sound general medical management and vigorous treatment of associated illnesses (especially infections) are essential.

Systemic vasculitis. This is a series of different clinical illnesses that are characterized by intense inflammation in the walls of blood vessels, especially arteries (that is, arteritis). These illnesses differ clinically, pathologically, and with respect to therapeutic responsiveness and outcome. Even the age and sex of the typically affected person varies from one to another.

The clinical pattern of illness varies with the sites of blood vessel involvement and the character of vascular injury. Tissue death in the vessel wall can lead to narrowing, or even blockage, of the vessel or weakening of the wall with formation of an aneurysm. Impaired circulation and altered blood flow result. Symptoms, therefore, can arise not only from the tissue inflammation itself but also from the lack of adequate blood supply in the organ supplied by that vessel. Vasculitis can be a systemic condition or can be confined to a single organ or to the skin. *See* ANEURYSM; CIRCULATION DISORDERS.

Polyarteritis. Formerly called polyarteritis nodosa, polyarteritis most commonly occurs in young adult men. In most instances, the cause of the inflammatory reaction, which affects medium-sized and small arteries, is not known. Sometimes, however, the condition appears to be associated with a viral infection (such as hepatitis B) or to be induced by a drug.

The clinical features of polyarteritis include both nonspecific manifestations of diffuse inflammation and reflections of arterial involvement at specific sites. Fever, fatiguability, and loss of appetite and weight are prominent. Arteritis may occur in the kidney (often associated with high blood pressure), the heart, the liver, and the circulation to the bowel. A peripheral nervous system disease is more frequent than involvement of the central nervous system. Joint pain and arthritis, muscle pain, and vasculitic skin lesions are also common. Polyarteritis varies greatly from one person to another; any one of these features, or any combination, can be present and cause clinical problems.

The laboratory manifestations are mostly nonspecific reflections of tissue inflammation, with elevations of the white blood cell count and platelets. The blood sedimentation rate is a standard test for the presence of an inflammatory process of some type but is nonspecific; it is elevated well above the normal in most individuals with polyarteritis. There is no diagnostic blood test specifically for polyarteritis; the diagnosis is established by either a biopsy of an involved site (for example, muscle or peripheral nerve) or demonstration of structural changes in the arteries by x-ray visualization after injection of dye (an angiogram).

Optimal therapy is yet to be defined, but disease control and prognosis have progressively improved, first with the administration of corticosteroids and in combination with a cytotoxic drug.

Granulomatous arteritis. This category encompasses several rare disorders, including Churg-Strauss syndrome, Wegener's granulomatosis, aortic arch arteritis, and giant cell arteritis. These disorders cause very different clinical problems, but all are generally characterized by the presence of granulomas and giant cells in the affected sites. The causes and mechanisms are unknown. These disorders differ from each other in age group and sex predominantly affected, and prognosis. Early recognition and prompt treatment are key to an optimal outcome. Indeed, without treatment these disorders tend to have a poor prognosis.

Churg-Strauss syndrome is rare. It involves small and medium-sized arteries as well as the smallest arterioles and venules. It is usually associated with pulmonary disease, with an allergic asthmatic component often present. A characteristic feature of this syndrome is a marked increase in the number of white blood cells known as eosinophils, both in peripheral blood and in the affected tissue.

Wegener's granulomatosis is characterized by inflammation of small vessels in the respiratory tract, in the kidneys, and elsewhere. Both the upper respiratory tract (nasopharynx and sinuses) and lungs can be involved; biopsy of lung tissue or upper airways best establishes the diagnosis. Addition of a cytotoxic drug, especially cyclophosphamide, to corticosteroid therapy can reverse the pulmonary process and prevent kidney failure.

Aortic arch arteritis, also called Takayasu's arteritis, occurs almost exclusively in young women. Large and medium-sized arteries that branch off the aorta are involved, with consequent narrowing or even occlusion. Impaired blood flow results, with diminished or absent pulses in the upper extremities and neck; thus this condition is sometimes called the pulseless syndrome. Decreased circulation to the head can impair brain function, with recurrent lightheadedness and fainting or even a stroke. Similarly, impaired coronary artery flow may result in heart failure and the pain of angina. These major complications occur late and can be prevented if the process is recognized and treated early. Diagnosis is established by intra-arterial dye studies that demonstrate characteristic structural changes in the large vessels.

Temporal arteritis occurs mainly in women over 50 years of age and especially in those who are elderly. Aching pain and severe stiffness in the proximal limb-girdle regions (that is, shoulder/arm, pelvis/thigh) are common early manifestations. Later, with temporal artery involvement (in the head), symptoms such as headaches, scalp tenderness, and visual abnormalities can occur. A tender, dilated temporal artery may be found. Visual loss is the most feared complication, and is preventable by prompt corticosteroid therapy. The musculoskeletal pain syndrome, which can be progressive and incapacitating, also promptly responds to steroids.

Primary central nervous system arteritis is rare. Headaches, altered mentation and confusion, decreased level of consciousness, and even coma can occur; infrequently, unexplained seizures can occur.

Hypersensitivity angitis. In this disease there is an intense inflammation of the smaller vessels, especially venules, that represents an adverse immunologic reaction to an offending agent. The agent could be a drug, a foreign protein, or an infectious agent. Skin lesions occur in all cases; they vary in type but typically include localized areas of bleeding into the skin from injured vessels over the feet and ankles. Systemic features also can be present. Henoch-Schönlein purpura is a designation applied to one subset of this disorder.

In all types of hypersensitivity vasculitis, the diagnosis is clinical since there are no specific blood tests. A skin biopsy may support the diagnosis if inflammation of small vessels is found. Treatment focuses on withdrawal of the offending antigen, if known, or treatment of any causal infection, if present. Occasionally, use of corticosteroids may be necessary.

Vasculitis associated with specific diseases. Vasculitis can also be a secondary process occurring along with other connective tissue diseases (for example, systemic lupus, polymyositis, and Sjögren's syndrome) and with rheumatoid arthritis, infections, and malignant neoplasms. The existence of this wide array of underlying disorders suggests that various immunologic pathways can lead to vascular injury. Although the appropriate therapy is generally determined by the associated disease process, emergence of vasculitis may indicate the need to administer, or increase the dosage of, corticosteroids. *See* VASCULAR DISORDERS.

Polymyositis. Polymyositis involves intense inflammation in skeletal muscle that, if untreated, can lead to destruction of muscle fibers. When typical skin lesions are present, the term dermatomyositis is applied. The pathology of these two entities differs and suggests different mechanisms of inflammation.

The cause of polymyositis or dermatomyositis is not known in most cases. These conditions can occur alone or in association with other connective tissue disorders (that is, systemic lupus, scleroderma, and Sjögren's syndrome). Dermatomyositis may sometimes accompany malignant tumors.

Polymyositis may be abrupt or insidious in onset. The dominant manifestation is progressive weakness of the proximal limb-girdle muscle groups. Difficulty rising from a chair, in climbing stairs, lifting modest weights, and keeping the arms elevated are common symptoms. Despite intense inflammation in the skeletal musculature, pain and tenderness of the involved muscles are infrequent. Most serious are the potential involvement of pharyngeal musculature which may impair swallowing, or of the chest muscles, which may impair breathing.

Destruction of muscle fibers causes release into the circulation of various enzymes normally contained in muscle; measuring levels of these enzymes in blood thus aids in diagnosis. Also, inflamed muscle is abnormal in its electrical activity; therefore the test known as an electromyogram yields abnormal results. Finally, muscle biopsy firmly establishes the diagnosis; it can be important in distinguishing polymyositis from noninflammatory disorders affecting muscle. Although polymyositis is presumed to be an immune-mediated process, the mechanisms are obscure. Antinuclear antibodies may be found in the serum. Other less frequent but more myositis-specific autoantibodies may also be present.

The inflammatory process is generally responsive to corticosteroid therapy, but in some cases the addition of a cytotoxic drug, especially methotrexate, is required. In the few individuals with tumor-related polymyositis, appropriate treatment of the malignancy is essential; and when polymyositis occurs together with other connective tissue disorders, such as lupus, the additional features of these illnesses must also be controlled.

If diagnosed early, before loss of muscle fibers and replacement by fibrous scarring is extensive, myositis is reversible; recovery with minimal residual loss of strength is to be expected in almost all patients.

Scleroderma. The term scleroderma means "hard skin," designating a disorder in which increased deposition of collagen fibers in the deeper dermis leads to thickened, leathery, bound-down skin. This deposition is variable in extent and degree. When organ involvement is associated with such skin changes, the term systemic sclerosis is used.

Raynaud's phenomenon, a hyperreactivity to cold exposure with blanching and discoloration of the fingers and toes, is the common initial manifestation and may precede sclerodermatous skin changes by decades. Only a small fraction of those with Raynaud's eventually develop scleroderma or systemic illness; but of those who do have scleroderma, at least 80% will have Raynaud's. Skin changes may be confined to the distal digits or become generalized.

Systemically, the gastrointestinal tract is commonly involved. Difficulty in swallowing and mid-chest discomfort result from loss of peristalsis in the esophagus. Diarrhea, inability to absorb nutrients, abdominal pain, and distention can occur with bowel involvement and lead to weight loss and wasting. Scarring of the lungs and decreased pulmonary function occur frequently. The most serious lesions are those of the heart and kidney, and may result from vascular abnormalities. Joint inflammation is mild and uncommon; but advanced skin changes, especially in the hands, may restrict joint mobility.

Polymyositis is occasionally present and is the major indication for corticosteroid therapy in this disorder.

A variant called the CREST syndrome (calcinosis, Raynaud's, esophageal dysmotility, sclerodactyl- or scleroderma-like finger, and telangiectasia) is notable because of its mildness and slow progression. It is also referred to as limited scleroderma.

Multiple autoantibodies (including ANA) may be found in serum. The diagnosis is usually a clinical one, based on characteristic skin tightening. A skin biopsy, if peformed, will demonstrate increased collagen fibers in the deeper dermis without any inflammatory reaction.

Management involves treatment of high blood pressure that can arise because of kidney involvement, use of agents to improve circulation by reducing small-vessel spasm, and reducing gastrointestinal complications with special diets and sometimes antibiotics. The drug D-penicillamine may produce improvement in the skin and possibly systemic features. In a few cases, skin changes may subside spontaneously.

Sjögren's syndrome. This condition is characterized primarily by dryness of the membranes due to excretory gland failure, especially dryness of the eyes (xerophthalmia) and mouth (xerostomia) from loss of tears and saliva, respectively. This sicca (dryness) syndrome reflects the infiltration of lacrimal and salivary glands by immunologically competent cells (lymphocytes).

Sjögren's syndrome may be primary or secondary, the latter occurring in the context of another disorder, such as rheumatoid arthritis or systemic lupus. The typical individual with primary Sjögren's syndrome is a woman later in life. In most cases, the sicca complex dominates, but there can be loss of secretions in the skin, bronchial tree, esophagus and stomach, and vaginal vault. In some people with primary Sjögren's, systemic features may evolve. Diagnosis is primarily clinical, but ocular dryness can be quantitated. Lip biopsy of the minor salivary glands can confirm the infiltration by immune cells. Noninvasive scanning of the parotids can confirm the loss of functional glandular structure. Autoantibodies in the blood are common and can aid in differentiating Sjögren's syndrome from other causes of dry eyes and dry mouth, but in no instance is a blood test diagnostic of either Sjögren's or any companion disorder.

Management involves artificially keeping membranes moist (for example, with artificial tears). Extraglandular inflammatory lesions and associated disorders are treated with anti-inflammatory agents and occasionally cytotoxic agents. *See* IMMUNOLOGY; INFLAMMATION; STEROID.

Mary Betty Stevens; Robert D. Inman

Bibliography. D. L. Gardner, *Pathological Basis of the Connective Tissue Diseases*, 1992; W. J. Koopman and L. W. Moreland (eds.), *Arthritis and Allied Conditions: A Textbook of Rheumatology*, 15th ed., Lippincott Williams & Wilkins, 2005; M. Rajkind (ed.), *Connective Tissue in Health and Disease*, 1989; P. M. Royce and B. Steinman (eds.), *Connective Tissue and Its Disorders, Genetic and Medical Aspects*, 1992.

Conodont

A group of extinct marine animals that are often abundant in strata of Late Cambrian to Late Triassic age, a time span of about 300 million years. Only the mineralized elements, which are usually 0.2 to 2 mm (0.008 to 0.08 in.) in dimension (the largest known reach 14 mm or 0.6 in.), are normally preserved. They are routinely extracted as isolated discrete specimens by chemical degradation of the rock in which they occur. The apatite (calcium phosphate) of which conodont elements are composed is laid down as lamellae. In the earliest euconodonts ("true" conodonts, as opposed to the more primitive, and possibly unrelated, protoconodonts and paraconodonts), the elements comprise an upper crown and a basal body. The basal body occupies a cavity in the base of the crown, but is not present in the majority of post-Devonian species. In advanced conodonts the crown incorporates regular patches of opaque, finely crystalline, white matter.

For many years, conodont taxonomists treated individual element types as separate species. There are three major shape categories, coniform, ramiform, and pectiniform (**Fig. 1**). Coniform elements were dominant in the Cambrian to Early Ordovician and common until the Devonian. They vary in curvature and in cross section, including the development of costae. Ramiform (comblike) elements extend into elongate processes with various arrangements of denticles. Pentiniform elements include straight and arched blades, and may be expanded laterally to form a platform.

Skeletal apparatus. In the 1930s the discovery of "natural assemblages" on bedding surfaces provided direct evidence that each animal possessed a skeletal apparatus consisting of several different kinds of elements. Several hundred such associations have now been found, each representing the skeletal remains of an individual animal. Most are from shales, as their preservation requires that the soft parts of the animal decay undisturbed by currents, so that the skeleton remains intact. Some elements may also become fused into clusters during the diagenesis of the sediments in which they occur, and these, too, give an indication of the configuration of the conodont apparatus. Although most conodont species are still known only from elements scattered in the sediment, it is now conventional practice in conodont taxonomy to reconstruct their skeletal apparatus by using morphological and statistical criteria.

The evidence of bedding-plane assemblages shows that most of the elements in an apparatus were symmetrically paired across the midline. A common type of apparatus in the Upper Paleozoic (Fig. 1*i*) consisted of a set of 7 to 11 ramiform elements (one of which was an unpaired median, bilaterally symmetrical form), followed by a pair of arched blades and, posterior of these, a pair of platform pectiniform

elements. Variations in this arrangement occur; early Paleozoic apparatus often consisted of a series of coniform elements alone.

Soft parts. In the absence of preserved soft parts, the nature of the affinities of conodonts was the subject of considerable speculation and debate. Since the discovery of isolated elements in 1856, conodonts have been variously aligned with algae, higher plants, several wormlike phyla, mollusks, arthropods, lophophorates, chaetognaths, and chordates, or have been assigned to a separate phylum, Conodonta. It was not until 1983 that evidence of the soft parts was described by D. E. G. Briggs, E. N. K. Clarkson, and R. J. Aldridge, on the basis of the first of several specimens discovered in lower Carboniferous rocks near Edinburgh, Scotland.

Features of the soft parts are preserved as a calcium phosphate film (**Fig. 2**). They show that conodonts were elongate animals (40–60 mm or 1.6–2.4 in.). The apparatus lies in the head region, behind a bilobed structure which is preserved as a film darker than the rest of the body (**Fig. 3**). All of the Carboniferous examples discovered to date are compacted in lateral view, and the apparatus lies to one side, in some cases beyond the apparent outline of the body. This suggests that the soft tissue which supported the elements is not preserved. The trunk shows clear evidence of V-shaped structures, the apex of the V's pointing anteriorly. These structures are interpreted as muscle blocks, but the nature of two parallel lines which run the length of the

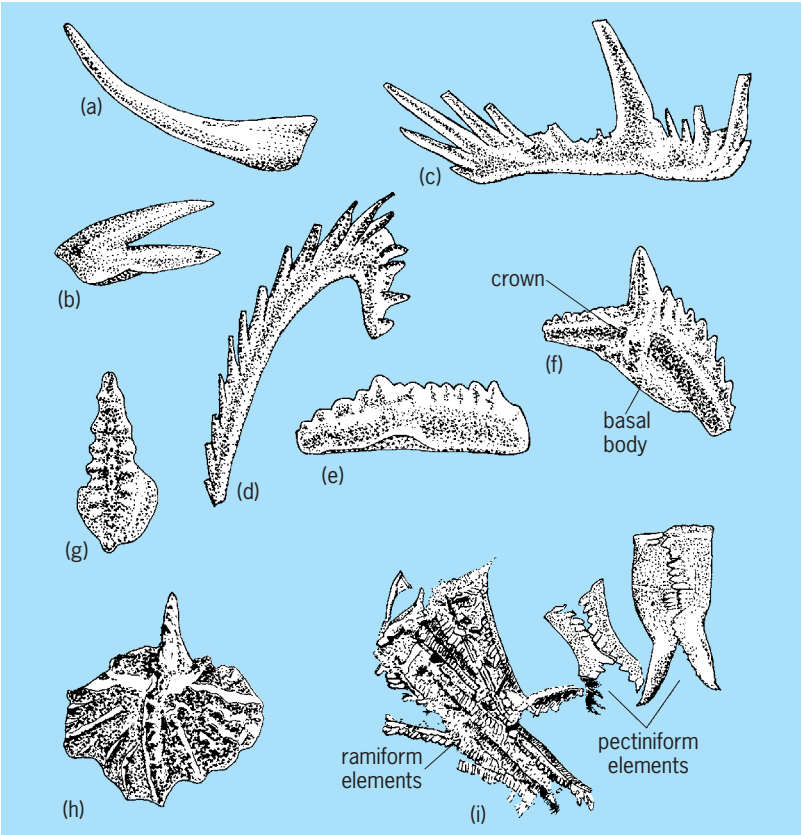

Fig. 1. Conodont elements: (*a*, *b*) coniform elements, (*c*, *d*) ramiform elements, (*e*, *f*) pectiniform blade elements, (*g*, *h*) pectiniform platform elements, (*i*) bedding-plane assemblage.

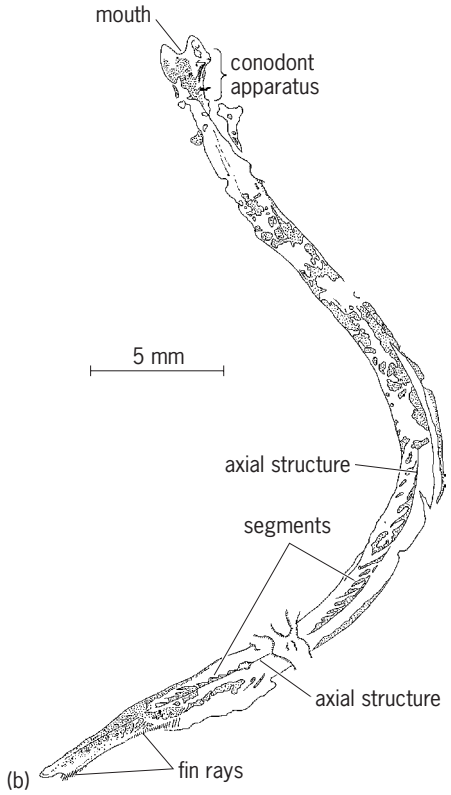

Fig. 2. The first specimen of the conodont animal discovered in Lower Carboniferous rocks in Edinburgh. (*a*) Photograph of specimen. (*b*) Outline of the specimen for comparison. (*From D. E. G. Briggs, E. N. K. Clarkson, and R. J. Aldridge, The conodont animal, Lethaia, 16:1–14, 1983*)

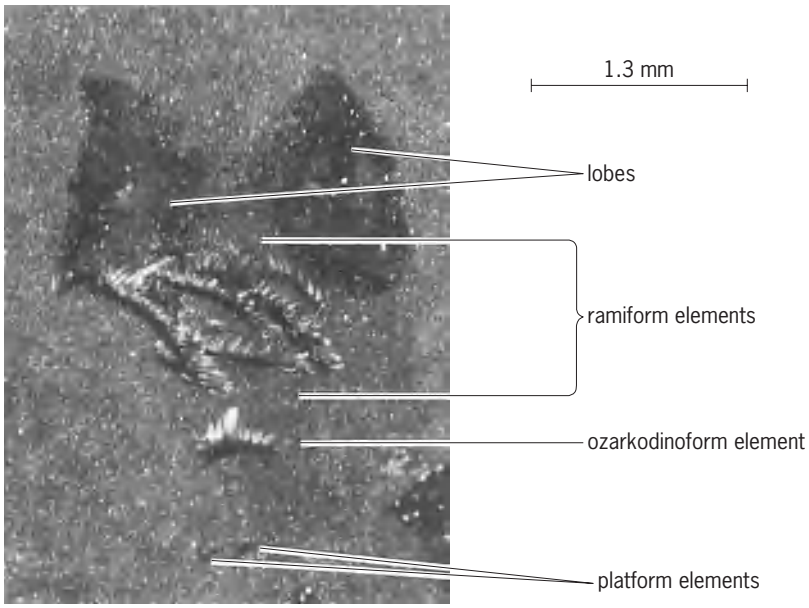

Fig. 3. Head region of the conodont animal, showing the conodont elements on the side of the slab opposing that illustrated in Fig. 2. (*From D. E. G Briggs, E. N. K Clarkson, and R. J. Aldridge, The conodont animal, Lethaia, 16:1–14, 1983*)

trunk is less certain. Fin rays indicate the presence of an asymmetrical fin around the margin of the tail. A single poorly preserved specimen from the Silurian of Wisconsin, the only other example known which shows traces of the soft parts, suggests that the body of some conodonts may have been broader than that in the Scottish specimens.

Biological affinities. The evidence of the soft-part morphology indicates that the conodonts belong within the chordates; it is no longer possible to justify their separation as a phylum, Conodonta. The lack of jaws and of a bony skeleton suggests that they represent a primitive group of jawless craniates, perhaps close to the Myxinoidea (hagfishes). The apparatus is thought to have functioned in food capture. The ramiform elements (or coniform elements in earlier conodonts) appear to be adapted for grasping prey. In later conodonts this would have been processed by the shearing and grinding action of the paired pectiniform elements *See* CHORDATA.

Geological applications. Although the biological affinities of conodonts and the function of the elements were essentially unknown until recently, they have nonetheless been extensively studied because of their important geological applications. Most significant of these is the use of conodont elements in biostratigraphy. Biozonal schemes have been erected for all the systems through which conodonts range, Cambrian to Triassic; the most detailed subdivision is in Upper Devonian strata (conodont subbiozones representing, on average, 0.5 million years). Some conodonts are restricted in their distribution to certain sedimentary environments and can be used to define biofacies. Most conodont biofacies are, at least indirectly, related to depth, as reflected in the contrasts between nearshore and offshore assemblages.

Conodont elements can also be used to determine the thermal history of strata. The color of elements alters from pale yellow through brown to black as their carbon content is fixed during organic metamorphism. This is the basis for a conodont Color Alteration Index (CAI, from 1 to 5) which can be applied in assessing oil and gas potential, and in wider studies of the thermal history of sedimentary basins.

The apatite of which conodont elements are formed is stable and resistant to diagenetic alteration. The trace-element and isotopic-ratio characteristics of the seawater in which the organism lived were retained in the apatite, which can be analyzed to yield significant information about the temperature and chemistry of ancient oceans. Thus an assemblage of conodont elements extracted from a rock sample not only may allow biostratigraphic correlation, but also may give an indication of the paleoenvironment of the locality; the temperature, geochemistry, and oxygen content of the seawater; and the subsequent thermal history of the sedimentary basin. *See* STRATIGRAPHY. Richard J. Aldridge; Derek E. G. Briggs

Bibliography. R. J. Aldridge (ed.), *Palaeobiology of Conodonts*, 1987; R. J. Aldridge et al., The affinities of conodonts: New evidence from the Carboniferous of Edinburgh, Scotland, *Lethaia*, 19:279–291, 1986; D. E. G. Briggs, E. N. K. Clarkson, and R. J. Aldridge, The conodont animal, *Lethaia*, 16:1–14, 1983; R. A. Robison (ed.), *Treatise on Invertebrate Paleontology*, W, suppl. 2, *Conodonta*, 1981.

Consciousness

Accounting for the sensations or "raw feels" which constitute conscious experience has been a persistent problem for scientific approaches to psychology. In addressing this problem, each new scientific approach has put forth a new account of the relationship between those conscious sensations, as they are privately experienced by individuals, and the nervous system, as it is experienced by scientists.

Fechner on consciousness. In 1860 Gustav Fechner published his *Elemente der Psychophysik*, wherein he developed scientific techniques to measure the magnitudes of conscious sensations such as brightness and loudness, and then derived a mathematical equation to describe the relationship between each sensory magnitude and the physical magnitude which stimulated it. Fechner derived this psychophysical equation as a scientific solution to the mind-body problem. He interpreted it as support for panpsychism, the metaphysical position that all matter is conscious of other matter. His interpretation was based on the equation's implication that a very small physical magnitude stimulates a negative sensation: the consciousness of a magnitude beneath the threshold at which neural substances produce awareness of each other's magnitudes, a conscious sensation such as that experienced by flowers and other nonneural substances.

Inasmuch as the mathematical relationship between psychological sensations and physical stimuli

must be derivable from psychoneural and neurophysical equations, Fechner's panpsychist interpretation of his psychophysical equation assumed that the underlying psychoneural equation was a logarithmic equation yielding the negative sensations and that the neurophysical equation was linear. However, critics argued that the logarithmic relationship reflected the neurophysical relationship between the magnitude of the brain state evoked by the stimulus and the magnitude of the physical stimulus itself, both of which are physically and scientifically measurable magnitudes. *See* PSYCHOPHYSICAL METHODS.

Wundt on consciousness. Historically, Wilhelm Wundt convinced the scientific establishment that the methods of science are sufficient to study the sensory data of consciousness. He argued that the inner aspects of conscious experience, the data of psychology, are the only immediately known data of science: that the outer world of physical reality has to be inferred from the data of scientific psychology and, thus, is known only mediately. His scientific studies of perceptual consciousness confirmed that perceptual experience provides no immediate knowledge of external reality.

In logically and experimentally supporting his thesis that the data of scientific psychology constitute the foundation for all sciences, Wundt sought to reduce physical reality to a subset of consciousness itself. Three founders of subsequent approaches to psychology were not satisfied with this effort to rewrite psychophysical and psychoneural equations in purely mental terms. One of them—John Watson, who founded behaviorism around 1913—sought, instead, to reduce mental terms to physical terms. The other two—Sigmund Freud, who founded psychoanalysis around 1900, and Max Wertheimer, who founded Gestalt psychology around 1912—sought to put mental terms on the psychological sides of the psychoneural and psychophysical equations and to put physical terms on the opposite sides.

Freudian reaction to Wundt. By arguing for the existence of unconscious mental events, Freud took issue with both the metaphysical and the methodological underpinnings of Wundt's approach to the psychology of consciousness. If unconscious mental events existed, then much of the mind itself—much of the data of scientific psychology—was no more immediately knowable than the inferred data of physical sciences. Moreover, to the extent that the unconscious mind was actually part of physical reality, Wundt was not justified in treating physical reality as a subset of the mind. In the light of Freud's analysis of the unconscious mind, the psychophysical and psychoneural relationships reemerged as unsolved and problematical relationships.

Behaviorist reaction to Wundt and Freud. In creating and developing behaviorism, Watson introduced both a new methodological approach to psychology and a new metaphysical approach to consciousness. Watson's scientific methods and metaphysical positions were addressed to the psychology of the other. The starting point for behaviorism was the scientist's immediate experience of another organism's behavioral response to a stimulus; metaphysical inferences about conscious events within that other organism were deemed to be unnecessary and unwarranted.

From the standpoint of behaviorism, all scientific conclusions regarding the psychophysical and psychoneural relationships—including Fechner's panpsychist interpretation of his psychophysical equation and Freud's distinction between conscious and unconscious mental events—were meaningless. This did not necessarily imply that mental terms such as consciousness were meaningless. It did imply that the behavioral scientist's physical descriptions of other entities like brains (including the physical description of his or her own brain) could not meaningfully be related to mental descriptions of his or her own consciousness.

From behaviorism to computer models. At the beginning of the twentieth century, Ivan Pavlov and Edward Thorndike initiated psychological research on animals, and Watson's psychology of the other organism provided the rationale for viewing other animals and other people as scientifically equivalent subjects of psychological study. Since the late 1960s, animal-behavior models of psychological processes have been replaced by computer models, and Watson's animal behaviorism has been transformed into a computer behaviorism or a psychology of the other machine. The goal of computer behaviorism is to program computers to produce output responses which observers experience as indistinguishable from human beings' responses. If and when this goal is achieved, computer models of psychological processes will provide logically sufficient explanations of the psychological processes of other human beings. The critical question is whether such computer models of psychological processes would also provide sufficient explanations of the psychological processes of the scientist who consciously experiences the other machine.

Some computer behaviorists argue that neither the scientist's conscious experience nor the computer's lack thereof has any causal effect on behavior. However, a subset of these computer behaviorists believe that any conscious psychological process in the scientist's computerlike brain can actually occur with or without consciousness of that process, implying that the scientist's psychological processes are, in principle, not different from the computer's. Another implication is that the scientist's brain contains a metaphysical subject which can consciously inspect or not inspect the same neural process. Other computer behaviorists adopt the functionalist argument that sensations like blueness and other privately experienced phenomena will "emerge" from any functionally equivalent psychological processes present either in the scientist's computerlike brain or in the other computer studied by the scientist. Metaphysically, computer behaviorism presupposes that the conscious mind is a machine; methodologically, it starts with the scientist's conscious experience of other machines. Another metaphysical and methodological approach, the approach of Gestalt psychology, presupposed also that the mind is a machine,

but started instead with the scientist's phenomenal experience of his or her own machine.

Gestalt reaction to Wundt and behaviorism. While the Gestalt psychologist Max Wertheimer rejected Watson's scientific psychology of the other organism, he also rejected Wundt's contention that nothing in the scientist's immediately known experience corresponds directly with the mediately known physical world. Gestalt psychologists objected that Wundt had restricted his observations to elementary sensations and had not observed the Gestalt, the holistic relationship among the immediately known sensations. According to Gestalt psychology, there was a direct psychophysical correspondence between perceptual experience and physical stimulation.

From Gestalt to mind-brain identity theory. In his effort to rebut Wundt's contention that any psychophysical correspondence is a necessarily indirect, arbitrarily learned, and basically illusory correspondence, Wertheimer adopted the scientist's conscious experience as his starting point and looked for phenomenal evidence to support a direct psychophysical relationship. In so making his case for a direct psychophysical correspondence, he set the stage for the Gestalt psychologist Wolfgang Köhler and other mind-brain identity theorists to argue for a direct psychoneural relationship.

According to mind-brain identity theorists, the scientist's sensations are qualities of particular brain structures in their excited state. Thus, the scientist is "conscious of" neither sensations nor the brain; the scientist is the sensory qualities of his or her brain states. The scientist immediately experiences nothing besides the qualities of the brain states which he or she is; and the scientist's experiential qualities change over time, simply because the patterning of excited structures in his or her brain changes.

According to mind-brain identity theory, a scientist could simultaneously experience his or her own brain both immediately and mediately, by visually experiencing the exposed brain on a television screen. In spite of the fact that such a mediately experienced brain would be a subset of the immediately experienced brain, the scientist could look for mediately experienced brain states which correlate perfectly with immediately experienced sensory qualities. Gestalt psychologists argue that mind-brain identity theory predicts that psychoneural correspondences can be discovered. Scientists adopting a mind-brain identity position have not yet discovered any structural differences in various parts of the brain, much less discovered any correspondences between conscious qualities and brain structures. Robert G. Kunzendorf

Lines of evidence. Psychologists and neuroscientists have developed many productive lines of evidence. Psychological investigations have compared conscious and unconscious processes in selective attention, visual perception, implicit versus explicit learning, conscious availability in problem solving and memory, automatic versus controlled task components, working memory, and so on. Theoretical approaches have emerged, couched in the vocabulary of information processing, cognitive architectures, and adaptive systems. Brain researchers have developed promising lines of evidence, including studies of brain damage (as in amnesia, neglect syndromes, split brain cases, hemispherectomies, and blindsight), as well as normal neurosensory processes, anesthesia, sleep, dreaming, and coma.

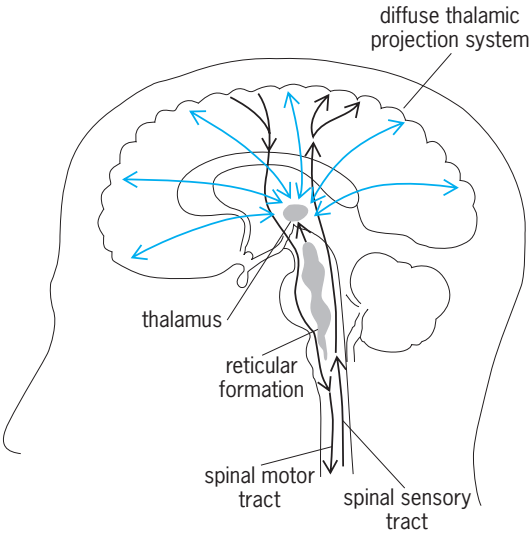

Brain structures most closely associated with the conscious experience. (*After B. J. Baars, A Cognitive Theory of Consciousness, Cambridge University Press, 1988*)

The effects of the presence and absence of consciousness reveal enormous neurobehavioral differences between the brain in deep sleep, stupor, or coma compared to when it is wakeful. The anatomical structures required for waking consciousness are well known. They include the brainstem reticular formation, the reticular nuclei of the thalamus (its outer shell), and the fountain of neuronal fibers projecting from the intralaminar nuclei of the thalamus to all parts of the cortex (see **illus.**). Damage to any of these structures will induce coma, whereas massive lesions elsewhere in the brain generally produce no loss of conscious wakefulness; actually, an entire hemisphere can be removed without producing coma.

In many cases, exact experimental comparisons can be made between mental events that are clearly conscious and very similar events that are not. Everyone has experienced repetitive stimuli fading from consciousness (for example, the repeated sound of a noisy refrigerator pump). Sometimes, when such a highly predictable train of events suddenly stops, the stimulus becomes conscious, so that an individual then perceives that something has changed, even though the person was not aware of it while it was taking place. Scientific observations of such habituation phenomena suggest that the nervous system must maintain a rather accurate, unconscious representation of the repetitive stimulus, since a change in any parameter of the stimulus (for example, an increase or decrease in intensity or a change in the temporal pattern) may bring it to consciousness.

Similarly, if someone is spoken to while reading a book, the person's first impulse may be to ask, "What did you say?", but suddenly the memory of what was said comes to mind. The words must have been stored unconsciously, or they would not be remembered a few seconds later.

A useful distinction can be made between conscious experience, that is, what people can report and act upon, and selective attention, which is the ability to choose between alternative streams of conscious contents. Once an individual attends to one dense stream of speech or visual flow, competing trains of stimulation appear to be processed unconsciously. In a classic dichotic listening task, subjects closely monitored sentences presented to one ear. Information presented to the other ear is not reportable under these conditions; it cannot be acted upon voluntarily, and it can be repeated many times without creating recallable memories. This result is consistent with the subjects' reports that the unattended stream was not conscious. However, the meaning of unattended words is in fact processed, since the listeners' understanding of the attended sentences can be influenced by those words.

In subliminal priming, subjects may be exposed to a printed word via one eye, while a visual mask is presented to the other eye in such a way that the target word cannot be reported. The masked word will nevertheless influence the time needed to recognize a related word presented immediately afterward, again suggesting that the unconscious stimulus must be processed and represented unconsciously.

The question is what is the difference between the mental representation of a stimulus when it is conscious versus unconscious. It has been suggested that novel combinations of stimuli require consciousness; unconscious processing may be rigid and automatic. But the unconscious is not simple. Complex tasks such as linguistic analysis, memory search, motor control, executive planning, and emotional response, which often involve abstract representations and symbolic processes, are performed unconsciously. Unconscious processes can be quite formidable in speed, accuracy, and complexity. The functional capacities of conscious and unconscious processes can be compared as follows.

Conscious processes:

1. They are computationally inefficient. Many errors, relatively low speed, and mutual interference occur between processes.
2. The range of contents is great. There is great ability to relate different conscious contents to each other, and conscious events to their unconscious contexts.
3. In regard to access, largely unconscious executive functions can gain control over unconscious specialists via consciousness. Conscious contents can refer to and sometimes trigger other conscious contents. Conscious contents can trigger unconscious automatisms.
4. There is high internal consistency at any single moment, seriality over time, and limited processing capacity.

Unconscious processes:

1. They are very efficient in routine tasks. Few errors, high speed, and little mutual interference occur between processes.
2. Each routine process has a limited range of contents. It is relatively isolated and autonomous, and is relatively context free.
3. Entirely unconscious processes probably have only routine (highly practiced, hard-wired, or both) access to each other.
4. The set of routine, unconscious processes are diverse, can sometimes operate in parallel, and together have great processing capacity.

Methodology. An individual's understanding of consciousness, while no doubt influenced by personal experience, needs to be based on publicly observable events that can be tested and verified by multiple observers. Conscious processes can be operationally defined as the set of events that are claimed by a subject to be conscious and can be reported and acted upon with verifiable accuracy, under optimal reporting conditions. Optimal reporting conditions imply a minimum delay between the conscious event and the report (for example, freedom from distraction). This definition fits existing practice in the study of perception, short-term memory, explicit problem solving, and imagery. *See* MEMORY; PERCEPTION; PROBLEM SOLVING (PSYCHOLOGY).

By contrast, unconscious processes can be operationally defined as those which are not claimed to be conscious, and cannot be reported or acted on, even under optimal reporting conditions, but whose presence can nevertheless be verified. By this definition, automatic processes are unconscious, as are visually masked stimuli, visual objects in blindsight, implicit memory, unattended information, subliminal effects, the details of language processing, and the effects of priming. Again, there is a good fit between this definition and existing scientific practice.

There is a curious asymmetry in the assessment of conscious and unconscious processes. Obtaining verifiable reports is a viable method for specifying conscious events, but unconscious ones are much more difficult to define. William James suggested that because the evidence for unconscious processes could always be challenged there were no unconscious psychological processes at all. This is known as the zero point problem. For example, if a word shown for a few milliseconds cannot be reported, it does not necessarily mean that the person is truly unconscious of the stimulus. There could have been a fleeting awareness that is soon masked or replaced by succeeding events. It is extraordinarily difficult to be sure of the zero point of consciousness.

Alternatively, a criterion for unconscious events can be applied which does not solve the problem exactly but which gives a reasonable basis for consensus. Suppose that a subject is asked to avoid recently

primed processes. If the subject cannot do so, those processes are plausibly unconscious. James could argue that by this criterion unconscious events could still be fleetingly conscious, however briefly. The criterion only guarantees that the unconscious event cannot be used in working consciousness—the ability to report and act upon the fleeting stimulus.

Limited capacity system. Most cognitive models formerly did not refer to consciousness explicitly but used terms and mechanisms that were very closely associated with it. For example, one of the first models was short-term memory, which is roughly the domain in which one rehearses telephone numbers while walking from the phone book to the telephone. More broadly, it may include the domain in which individuals mentally talk to themselves. It is true, of course, that not all items in short-term memory are conscious at any single instant. Of the seven digits of a new telephone number that can be mentally rehearsed, only one or two are conscious at any time. Thus, short-term memory is not identical to consciousness but seems to be very closely related.

All phenomena associated with conscious experience are characteristically limited in capacity. A person watching a baseball game on television cannot simultaneously experience a football game overlaid on the same screen, nor can a person follow a normal conversation while performing any other unpredictable, skilled task. Short-term memory is characterized by a limit of seven plus or minus two chunks of information. Voluntary control, which is also intimately involved with conscious experience, seems similarly limited. Finally, the flow of conscious information is relatively serial.

There is a curious dichotomy between limited-capacity models and evidence suggesting that the unconscious parts of the nervous system have enormous capacity. The vast capacity of long-term memory, the great size of the linguistic lexicon, and the large set of available automatic skills are known in a psychological sense. The most obvious illustration of enormous unconscious capacity is the sheer size and complexity of the brain. The cerebral cortex alone has some 50 billion neurons, with similar numbers for the large subcortical mass, the spinal cord, and the great network of peripheral neurons. Each neuron sends an electrochemical signal to its neighbors an average of 40 times per second, up to perhaps 1000 Hz. The system is thoroughly interconnected: Neurons have up to 10,000 connections, and a path can be traced from any single neuron in the brain to any other neuron in seven steps or less. This enormous organ, the most complex in the body, operates concurrently rather than serially and performs highly complex and intelligent tasks with minimal conscious involvement. *See* BRAIN.

Society models. The society model of the nervous system suggests that the great bulk of information processing is performed by local networks that are unconscious, though they are complex, specialized, and intelligent, and operate in parallel and normally very efficiently. Such expert societies have great flexibility in problem solving and control. Repetitive events and routine actions can be handled by specialized processors with minimal conscious involvement. It is possible that conscious experience corresponds to a domain of interaction between the multitude of unconscious experts, one that may act as a global "publicity organ'' or meeting place in the society of intelligent processors. Consciousness in this view serves as an indispensable adaptive function. *See* COGNITION; INFORMATION PROCESSING (PSYCHOLOGY); PSYCHOLOGY.

Bernard J. Baars; Katharine McGovern

Bibliography. B. J. Baars, *A Cognitive Theory of Consciousness*, 1988; J. Brzezinski (ed.), *Consciousness: Methodological and Psychological Approaches,* 1985; R. W. Coan, *Human Consciousness and Its Evolution,* 1987; D. Dennett, *Consciousness Explained*, 1992; R. Ellis, *An Ontology of Consciousness*, 1986; A. J. Hobson, *States of Brain and Mind*, 1988; W. James, *The Principles of Psychology*, 1890, 1983; A. J. Marcel and E. Bisiach (eds.), *Consciousness in Contemporary Science*, 1988, reprint, 1992; A. D. Milner and M. D. Ruggs (eds.), *The Neuropsychology of Consciousness*, 1992.

Conservation laws (physics)

Principles which state that the total values of specified quantities remain constant in time for an isolated system. Conservation laws occupy enormously important positions both at the foundations of physics and in its applications.

Realization in classical mechanics. There are three great conservation laws of mechanics: the conservation of linear momentum, often referred to simply as the conservation of momentum; the conservation of angular momentum; and the conservation of energy. It is almost impossible to understand the modern form of the conservation laws without understanding their original and simplest realization in the classical mechanics of particles. Therefore, the conservation laws as they appear within that framework will first be discussed, and then their generalization and modification to cover larger domains.

Conservation of momentum. The linear momentum, or simply momentum, of a particle is equal to the product of its mass and velocity. It is a vector quantity. The total momentum of a system of particles is simply the sum of the momenta of each particle considered separately. The law of conservation of momentum states that this total momentum does not change in time.

An important application of this principle is to collisions. For example, in a collision between two bodies of equal mass, starting with opposite velocities, the total momentum is zero. The conservation of momentum states that it will still be zero after the collision. Thus, if the bodies remain intact they must move off again with opposite velocities—though not necessarily the original ones, for both bodies may have been deflected through a common angle. If the collision causes part of one of the original bodies to adhere to the other, the two bodies will again move off in opposite directions, but the heavier one will

move slower and the lighter one faster, by factors inversely proportional to their new masses. Even if the bodies shatter into several bits, there are three constraints on the motion of the debris that emerges: each component of the total momentum must vanish. *See* COLLISION (PHYSICS).

This application is typical in showing both the power and the limitations of general conservation laws as they are applied to complex processes. They cannot tell what happens in detail, but they constrain the possibilities in ways that are nontrivial and unequivocal. *See* CONSERVATION OF MOMENTUM; MOMENTUM.

Conservation of angular momentum. The angular momentum of a particle is more complicated. It is also a vector quantity. It is defined with respect to some origin of coordinates O. Let the position P of the particle be specified by the position vector $\vec{r} = \overline{OP}$ pointing from the origin to it; and let its mass and velocity be m and $\vec{v}$ respectively. Then the angular momentum of the particle is defined by the vector product of the position and momentum vectors, Eq. (1). The total

$$\vec{L} = \vec{r} \times (m\vec{v}) \tag{1}$$

angular momentum of a system of particles is equal to the sum of the angular momenta of each particle separately. The law of conservation of angular momentum states that the total angular momentum of an isolated system is constant in time.

An extremely important instance of the conservation of angular momentum is Kepler's second law of planetary motion, which states that a line drawn from the Sun to a planet sweeps out equal areas in equal times. This would be a geometrical statement entirely equivalent to the conservation of angular momentum if the planet and Sun formed an isolated system. In reality they do not: A planet is subject to gravitational forces from other planets as well as the Sun, so that Kepler's law is only approximately correct. There are many other important applications of the law of conservation of angular momentum to mechanical problems involving rotation, particularly the motion of tops and gyroscopes. *See* ANGULAR MOMENTUM; KEPLER'S LAWS; RIGID-BODY DYNAMICS.

Conservation of energy. The third great conservation law of classical mechanics, conservation of energy, is perhaps the most important of all (although it was the last to be properly appreciated, for reasons discussed below). Energy is a scalar quantity, and it takes two forms: kinetic and potential. The kinetic energy of a particle is defined to be one-half the product of its mass with the square of its velocity, $T = {}^1/_2 mv^2$, and the total kinetic energy of a system of particles is simply the sum of their separate kinetic energies. The potential energy is loosely defined as the ability to do work; more precisely, it is a function V of the particles' positions, such that the force felt by a particle is minus the derivative of the potential energy with respect to that particle's position. The total energy is the sum of the kinetic and potential energies, $E = T + V$, and according to the conservation law it remains constant in time for an isolated system.

The essential difficulty in applying the conservation of energy can be appreciated by considering once again the problem of two colliding bodies. In general, even if they retain their integrity, the bodies emerge from the collision moving more slowly than when they entered. When this happens, the collision is said to be inelastic. An inelastic collision seems to violate the conservation of energy, until it is recognized that the bodies involved may consist of smaller particles. If these particles are shaken up and rattle around within the larger bodies after the collision, their random small-scale motions will require kinetic energy, which robs kinetic energy from the overall coherent large-scale motion of the bodies that are observed directly.

Such difficulties make it more challenging to extract concrete quantitative consequences from the conservation of energy than from the other conservation laws. One of the greatest achievements of nineteenth-century physics was the recognition that small-scale motion within macroscopic bodies could be identified with the perceived property of heat. By measuring the heat created by a given amount of mechanical work, J. P. Joule was able to define a precise mechanical equivalent of heat—for example, how much work it takes to heat a gram of water by one degree. *See* THERMODYNAMIC PRINCIPLES.

The first law of thermodynamics is simply an expression of the conservation of energy, using this concept that energy can take the form of heat (that is, microscopic, random motion) as well as perceived macroscopic motion or potential energy. The idea of heat as microscopic motion is used to explain some of the properties of gases. Finally, careful observations on the microscopic motions of small bodies have experimentally verified the theoretical predictions for this so-called brownian motion in great detail. *See* BROWNIAN MOVEMENT; CONSERVATION OF ENERGY; ENERGY; KINETIC THEORY OF MATTER; STATISTICAL MECHANICS.

Methodology. In Newton's original scheme of mechanics, the notion of force was primary. Later workers obtained simpler and more powerful mathematical formulations of the laws, without changing their physical content, by regarding energy as the primary notion. In modern physics this change in emphasis has become more pronounced, so that, for example, in quantum mechanics the concept of force is hardly ever encountered, but always energy. *See* HAMILTON'S EQUATIONS OF MOTION; LAGRANGE'S EQUATIONS.

Position in modern physics. As physics has evolved, the great conservation laws have likewise evolved in both form and content, but have never ceased to be important guiding principles, and never more so than now. Only a few highlights will be discussed.

Electromagnetism. In order to account for the phenomena of electromagnetism, it was necessary to go beyond the notion of point particles, to postulate the existence of continuous electric and magnetic fields filling all space. An assembly of charged particles can excite propagating waves in the fields. These are interpreted as electromagnetic radiation,

of which, for example, light and radio waves are special cases. When such radiation occurs, the total energy of the particles is diminished in a definite way, depending on the amount and type of radiation emitted. Thus, to obtain valid conservation laws, energy (and, for similar reasons, momentum and angular momentum) must be ascribed to the electromagnetic fields. *See* ELECTROMAGNETIC RADIATION; MAXWELL'S EQUATIONS; POYNTING'S VECTOR.

Special relativity. In the special theory of relativity, energy and momentum are not independent concepts. A leading idea of this theory is that the laws of physics take the same form when viewed by two observers who move with respect to one another at a constant velocity. In particular, each observer will associate both a conserved energy and a conserved momentum with a given isolated system. However, the energy reported by one observer is equal to a certain combination of the energy and momentum reported by the other.

Einstein discovered perhaps the most important consequence of special relativity, that is, the equivalence of mass and energy, as a consequence of the conservation laws. He considered a thought-experiment wherein an atom emits equal packets of radiation in opposite directions. He showed that the law of conservation of energy could be valid for this process, as viewed by a moving observer, only if, as a result of the radiation, the atom's mass is diminished, according to Eq. (2), where E is the energy

$$\Delta m = \frac{E}{c^2} \tag{2}$$

of the radiation emitted and c is the speed of light. In the nineteenth century, the law of conservation of mass was a major item in any discussion of conservation laws. Now this "law" is understood as an approximate consequence of the conservation of energy (valid in the limit $c \to \infty$). *See* CONSERVATION OF MASS; RELATIVITY.

Connection with symmetries. A remarkable, beautiful, and very fruitful connection has been established between symmetries and conservation laws. Thus the law of conservation of linear momentum is understood as a consequence of the homogeneity of space, the conservation of angular momentum as a consequence of the isotropy of space, and the conservation of energy as a consequence of the homogeneity of time. *See* SYMMETRY LAWS (PHYSICS).

General relativity. The development of general relativity, the modern theory of gravitation, necessitates attention to a fundamental question for the conservation laws: The laws refer to an "isolated system," but it is not clear that any system is truly isolated. This is a particularly acute problem for gravitational forces, which are long range and add up over cosmological distances.

Indeed, according to general relativity, a small region of space-time cannot be regarded as isolated, and the conservation laws are not valid. However, special solutions of the equations of general relativity, including model universes which closely model that of the Earth, support the conservation laws as a good approximation, valid over distances and times that are not too large. The relevant feature of these model universes is that they are approximately homogeneous and evolve slowly.

Fortunately, the scales are such that practical applications of the conservation laws are completely unaffected by this limitation. However, there is an important philosophical consequence: The symmetry of physical laws, which, as mentioned above, is closely tied up with the conservation laws, is actually a more fundamental property than the conservation laws themselves, for the symmetries remain valid while the conservation laws, strictly speaking, fail.

Quantum theory. In quantum theory, the great conservation laws remain valid in a very strong sense. Generally, the formalism of quantum mechanics does not allow prediction of the outcome of individual experiments, but only the relative probability of different possible outcomes. One might therefore entertain the possibility that the conservation laws were valid only on the average. However, the power of the conservation laws is much stronger: Momentum, angular momentum, and energy are conserved in every experiment, not just on the average.

How this can be so is one of the strangest yet simplest and most profound aspects of quantum theory. It is helpful to consider the decay of an excited state of an atom, wherein it radiates a single photon. If an apparatus is set up to detect the photon, then for every possible direction there is a probability that the photon will be detected moving in that direction. Now, if after detecting the photon the observer goes back and examines the atom, it will be found to be moving in the opposite direction in such a way that the total momentum is zero, as it was initially. This raises the question of how the atom could "find out" the result of the measurement of the photon. This is one form of what Einstein called the "spooky action-at-a-distance" in quantum mechanics. To interpret this situation, it must be acknowledged that the atom, like the photon, has a range of potential motions, but that these two potentialities, though they can be measured in either order and quite separately, are rigorously correlated. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; QUANTUM THEORY OF MEASUREMENT.

Conservation laws of particle type. There is another important class of conservation laws, associated not with the motion of particles but with their type. Perhaps the most practically important of these laws is the conservation of chemical elements, that is, the principle that, in a chemical reaction, for each chemical element the total number of atoms of that element in the products is equal to the total number of atoms of that element in the original reactants. From a modern viewpoint, this principle results from the fact that the small amount of energy involved in chemical transformations is inadequate to disrupt the nuclei deep within atoms. It is not an absolute law, because some nuclei decay spontaneously, changing the elemental nature of the atoms

in which they reside. At sufficiently high energies the "law" of conservation of elements is grossly violated; indeed, a well-developed theory traces the origin of chemical elements to processes occurring within minutes after the big bang, and in stars and supernova explosions. *See* BIG BANG THEORY; NUCLEOSYNTHESIS; RADIOACTIVITY.

Several conservation laws in particle physics are of the same character: They are useful even though they are not exact because, while known processes violate them, such processes are either unusually slow or require extremely high energy. Examples of such approximate conservation laws are the conservation of strangeness, of charm, and of isospin. *See* CHARM; ELEMENTARY PARTICLE; I-SPIN; STRANGE PARTICLES.

Two laws of conservation of particle type deserve special mention: the conservation of electric charge and the conservation of baryon number.

Conservation of electric charge. This law is believed to be exact. In principle, it is very simple to state, involving only counting. Every kind of elementary particle has a definite electric charge, which is an integer times the charge of the electron. The conservation law simply specifies that in any reaction the total electric charge of the reactants is equal to the total electric charge of the products.

Conservation of electric charge is deeply connected to the principle of electromagnetic gauge invariance and to the foundations of electromagnetic theory. One aspect is that the total charge inside a volume can be inferred by measurements of the electric field at a distant bounding surface, according to Gauss's law. Thus, if the charge suddenly changed it would require an instantaneous change in the electric field far away, contrary to the principle that the speed of light limits the transfer of information. *See* ELECTRIC CHARGE; ELECTRODYNAMICS; GAUGE THEORY.

Conservation of baryon number. This law is also very easy to state, as a counting rule. Among the common so-called elementary particles, protons and neutrons are assigned baryon number one—and their antiparticles baryon number minus one—while electrons, positrons, photons, and pions are assigned baryon number zero. A more modern and fundamental perspective is that quarks are assigned baryon number one-third, antiquarks baryon number minus one-third, and all other fundamental particles zero. In either formulation, the law of conservation of baryon number is simply the statement that, in any reaction, the total baryon number of the reactants is equal to the total baryon number of the products. *See* BARYON; QUARKS.

The most important evidence for the law of conservation of baryon number is the stability of the proton. There is no principle, other than this (or some similar conservation law, such as the conservation of lepton number), that forbids, for example, the decay [reaction (3)] of a proton into a neutral pion and a

$$p \rightarrow \pi^0 + e^+ \qquad (3)$$

positron. Yet, whereas otherwise similar decays take place within a small fraction of a second, the lifetime of the proton is greater than 10^{30} years. Despite its accuracy, the law of conservation of baryon number, unlike the law of conservation of electric charge, is not tied up with any known profound dynamical principles. Indeed, important ideas in modern particle physics definitely suggest that the conservation of baryon number is not exactly valid, and difficult but important experiments are planned to search for small violations of it. *See* PROTON. Frank Wilczek

Bibliography. T.-P. Cheng, and L.-F. Li, *Gauge Theories of Elementary Particle Physics,* 1984, reprint 1988; C. Lanczos, *The Variational Principles of Mechanics,* 1970, reprint 1986; E. Mach, *The Science of Mechanics,* 6th ed., 1988; R. Tolman, *The Principles of Statistical Mechanics,* 1967, reprint 1980.

Conservation of energy

The principle of conservation of energy states that energy cannot be created or destroyed, although it can be changed from one form to another. Thus in any isolated or closed system, the sum of all forms of energy remains constant. The energy of the system may be interconverted among many different forms—mechanical, electrical, magnetic, thermal, chemical, nuclear, and so on—and as time progresses, it tends to become less and less available; but within the limits of small experimental uncertainty, no change in total amount of energy has been observed in any situation in which it has been possible to ensure that energy has not entered or left the system in the form of work or heat. For a system that is both gaining and losing energy in the form of work and heat, as is true of any machine in operation, the energy principle asserts that the net gain of energy is equal to the total change of the system's internal energy. *See* THERMODYNAMIC PRINCIPLES.

Application to life processes. The energy principle as applied to life processes has also been studied. For instance, the quantity of heat obtained by burning food equivalent to the daily food intake of an animal is found to be equal to the daily amount of energy released by the animal in the forms of heat, work done, and energy in the waste products. (It is assumed that the animal is not gaining or losing weight.) Studies with similar results have also been made of photosynthesis, the process upon which the existence of practically all plant and animal life ultimately depends. *See* METABOLISM; PHOTOSYNTHESIS.

Conservation of mechanical energy. There are many other ways in which the principle of conservation of energy may be stated, depending on the intended application. Examples are the various methods of stating the first law of thermodynamics, the work-kinetic energy theorem, and the assertion that perpetual motion of the first kind is impossible. Of particular interest is the special form of the principle known as the principle of conservation of mechanical energy

(kinetic E_k plus potential E_p) of any system of bodies connected together in any way is conserved, provided that the system is free of all frictional forces, including internal friction that could arise during collisions of the bodies of the system. Although frictional or other nonconservative forces are always present in any actual situation, their effects in many cases are so small that the principle of conservation of mechanical energy is a very useful approximation. Thus for a missile or satellite traveling high in space, the dissipative effects arising from such sources as the residual air and meteoric dust are so exceedingly small that the loss of mechanical energy $E_k + E_p$ of the body as it proceeds along its trajectory may, for many purposes, be disregarded. *See* ENERGY; PERPETUAL MOTION.

Mechanical equivalent of heat. The mechanical energy principle is very old, being directly derivable as a theorem from Newton's law of motion. Also very old are the notions that the disappearance of mechanical energy in actual situations is always accompanied by the production of heat and that heat itself is to be ascribed to the random motions of the particles of which matter is composed. But a really clear conception of heat as a form of energy came only near the middle of the nineteenth century, when J. P. Joule and others demonstrated the equivalence of heat and work by showing experimentally that for every definite amount of work done against friction there always appears a definite quantity of heat. The experiments usually were so arranged that the heat generated was absorbed by a given quantity of water, and it was observed that a given expenditure of mechanical energy always produced the same rise of temperature in the water. The resulting numerical relation between quantities of mechanical energy and heat is called the Joule equivalent, or mechanical equivalent of heat. The present accepted value of the 15-degree calorie is $1\ \text{cal}_{15^\circ} = 4.1855 \pm 0.0004$ J; in thermochemical work it is now more common to use the thermochemical calorie, which is defined as $1\ \text{cal} = 4.184$ J exactly.

Conservation of mass-energy. In view of the principle of equivalence of mass and energy in the restricted theory of relativity, the classical principle of conservation of energy must be regarded as a special case of the principle of conservation of mass-energy. However, this more general principle need be invoked only when dealing with certain nuclear phenomena or when speeds comparable with the speed of light (1.86×10^5 mi/s or 3.00×10^5 km/s) are involved. *See* RELATIVITY.

If the mass-energy relation, $E = mc^2$, where c is the speed of light, is considered as providing an equivalence between energy E and mass m in the same sense as the Joule equivalent provides an equivalence between mechanical energy and heat, there results the relation, $1\ \text{kg} = 9 \times 10^{16}$ joules.

Laws of motion. The law of conservation of energy has been established by thousands of meticulous measurements of gains and losses of all known forms of energy. It is now known that the total energy of a properly isolated system remains constant. Some parts or particles of the system may gain energy but others must lose just as much. The actual behavior of all the particles, and thus of the whole system, obeys certain laws of motion. These laws of motion must therefore be such that the energy of the total system is not changed by collisions or other interactions of its parts. It is a remarkable fact that one can test for this property of the laws of motion by a simple mathematical manipulation that is the same for all known laws: classical, relativistic, and quantum mechanical.

The mathematical test is as follows. Replace the variable t, which stands for time, by $t + a$, where a is a constant. If the equations of motion are not changed by such a substitution, it can be proved that the energy of any system governed by these equations is conserved. For example, if the only expression containing time is $t_2 - t_1$, changing t_2 to $t_2 + a$ and t_1 to $t_1 + a$ leaves the expression unchanged. Such expressions are said to be invariant under time displacement. When daylight-saving time goes into effect, every t is changed to $t + 1$ h. It is unnecessary to make this substitution in any known laws of nature, which are all invariant under time displacement. Without such invariance laws of nature would change with the passage of time, and repeating an experiment would have no clear-cut meaning. In fact, science, as it is known today, would not exist.

Duane E. Roller/Leo Nedelsky

Bibliography. J. D. Cutnell and K. W. Johnson, *Physics*, 6th ed., 2 vols., 2004; D. Halliday, R. Resnick, and K. Krane, *Physics*, 5th ed., 2 vols., 2004; E. R. Jones and R. Childers, *Contemporary College Physics*, 3d ed., 2001; F. J. Keller, W. E. Gettys, and M. J. Skove, *Physics*, 2d ed., 1993; H. D. Young et al., *Sears and Zemansky's University Physics*, 10th ed., 2000.

Conservation of mass

The notion that mass, or matter, can be neither created nor destroyed. According to conservation of mass, reactions and interactions which change the properties of substances leave unchanged their total mass; for instance, when charcoal burns, the mass of all of the products of combustion, such as ashes, soot, and gases, equals the original mass of charcoal and the oxygen with which it reacted.

The special theory of relativity of Albert Einstein, which has been verified by experiment, has shown, however, that the mass of a body changes as the energy possessed by the body changes. Such changes in mass are too small to be detected except in subatomic phenomena. Futhermore, matter may be created, for instance, by the materialization of a photon (quantum of electromagnetic energy) into an electron-positron pair; or it may be destroyed, by the annihilation of this pair of elementary particles to produce a pair of photons. *See* ELECTRON-POSITRON PAIR PRODUCTION; LIGHT; MASS; RELATIVITY.

Leo Nedelsky

Conservation of momentum

The principle that, when a system of masses is subject only to forces that masses of the system exert on one another, the total vector momentum of the system is constant. Since vector momentum is conserved, in problems involving more than one dimension the component of momentum in any direction will remain constant. The principle of conservation of momentum holds generally and is applicable in all fields of physics. In particular, momentum is conserved even if the particles of a system exert forces on one another or if the total mechanical energy is not conserved.

Use of the principle of conservation of momentum is fundamental in the solution of collision problems. *See* COLLISION (PHYSICS).

If a person standing on a well-lubricated cart steps forward, the cart moves backward. One can explain this result by momentum conservation, considering the system to consist of cart and human. If both person and cart are originally at rest, the momentum of the system is zero. If the person then acquires forward momentum by stepping forward, the cart must receive a backward momentum of equal magnitude in order for the system to retain a total momentum of zero.

When the principle of conservation of momentum is applied, care must be taken that the system under consideration is really isolated. For example, when a rough rock rolls down a hill, the isolated system would have to consist of the rock plus the earth, and not the rock alone, since momentum exchanges between the rock and the earth cannot be neglected.

Rocket propulsion. The propulsion of a rocket through space can be explained in terms of momentum conservation. Hot gases produced by the combustion of the fuel are expelled at high speed from the rear of the rocket. Although the total mass of these hot gases may not be large, the gases move with such a high velocity that the total momentum associated with them is appreciable. The momentum of the gases is directed backward. For momentum to be conserved, the rocket must acquire an equal momentum in the forward direction. If the rocket carries all the materials needed for the combustion of its fuel, its propulsion does not require air, and it can move through empty space. *See* PROPULSION; ROCKET.

Exploding bomb. An exploding bomb gives another application of the conservation of momentum. The total resultant vector momentum of all the pieces of the bomb immediately after explosion must equal the momentum of the unexploded bomb just before the explosion. *See* EXPLOSIVE.

Proof of principle. The principle of conservation of momentum follows directly from Newton's second and third laws. While the principle will be proved here only for the straight-line motion of a two-particle system, it can be generalized to systems containing any number of particles. A particle is a mass with dimensions so small that rotational effects are negligible. Momentum will also be conserved for rigid bodies large enough that rotation must be considered, since rigid bodies can be treated as assemblies of many particles.

For the one-dimensional motion of an isolated two-particle system, Newton's third law states that the force F_{12} that particle 1 exerts on particle 2 is equal in magnitude and opposite in direction to the force F_{21} that particle 2 exerts on particle 1. Thus Eq. (1) holds.

$$F_{21} = -F_{12} \tag{1}$$

By use of Newton's second law this equation can be expressed in terms of the momenta m_1v_1 and m_2v_2 of particles 1 and 2, respectively, where m_1, m_2, v_1, and v_2 are the masses and velocities of particles 1 and 2, respectively. Then Eq. (2) holds. Integration

$$m_1\frac{dv_1}{dt} = -m_2\frac{dv_2}{dt} \tag{2}$$

gives Eq. (3), where c is a constant. This equation

$$m_1v_1 + m_2v_2 = c \tag{3}$$

expresses the conservation of momentum for two particles moving in the same straight line.

Finally it should be mentioned that angular and linear momentum are independent quantities. A complete description of a system must include both quantities. The angular momentum of a system is conserved under quite general conditions. *See* ANGULAR MOMENTUM; CONSERVATION OF ENERGY; MOMENTUM.

Paul W. Schmidt

Conservation of resources

Management of the human use of natural resources to provide the maximum benefit to current generations while maintaining capacity to meet the needs of future generations. Conservation includes both the protection and rational use of natural resources.

Natural resources. Earth's natural resources are either nonrenewable, such as minerals, oil, gas, and coal, or renewable, such as water, timber, fisheries, and agricultural crops. Natural resources are the basic goods and services that sustain human societies. Renewable resources play central roles in providing air, water, and food. Nonrenewable resources provide the energy essential for industrial economies and are the source of important products ranging from iron tools to silicon chips.

Although humans are entirely dependent upon Earth's natural resources, the combination of growing populations and increasing levels of resource consumption is now degrading and depleting the natural resource base. The world's population stood at 850 million at the onset of the industrial age, sharing the Earth with life forms nearly as diverse as the planet has ever possessed. The global population has grown to nearly seven times as large (6 billion), and the level of consumption of resources is far greater. This human pressure now exceeds the carrying capacity of many natural resources.

Nonrenewable. Nonrenewable resources, such as fossil fuels, are replaced over geologic time scales of tens of millions of years. Human societies will eventually use up all of the economically available stock of many nonrenewable resources, such as oil. Peak oil production in the United States, for example, occurred in 1984 and has since declined significantly as the most easily accessible oil fields have been drained. Clearly, because the stock of nonrenewable resources is fixed, future generations will inherit a world depleted of those important resources. Conservation thus entails actions to use these resources most efficiently and thereby extend their life as long as possible. By recycling aluminum, for example, the same piece of material is reused in a series of products, reducing the amount of aluminum ore that must be mined. Similarly, energy-efficient products help to conserve fossil fuels since the same energy services, such as lighting or transportation, can be attained with smaller amounts of fuel. *See* HUMAN ECOLOGY.

It may be expected that the biggest challenge of resource conservation would involve nonrenewable resources, since renewable resources can replenish themselves after harvesting. In fact, the opposite is the case. Historically, when nonrenewable resources have been depleted, new technologies have been developed that effectively substitute for the depleted resources. Indeed, new technologies have often reduced pressure on these resources even before they are fully depleted. Fiber optics, for example, has substituted for copper in many electrical applications, and it is anticipated that renewable sources of energy, such as photovoltaic cells, wind power, and hydropower, will ultimately take the place of fossil fuels when stocks are depleted. Renewable resources, in contrast, can be seriously depleted if they are subjected to excessive harvest or otherwise degraded, and no substitutes are available for, say, clean water or food products such as fish or agricultural crops. Moreover, when the misuse of biological resources causes the complete extinction of a species or the loss of a particular habitat, there can be no substitute for the esthetic and moral values that many people associate with that diversity of life.

Renewable. The urgent need to conserve renewable resources, particularly biological resources, can be seen in trends associated with major ecosystems. For example, 60% of the world's most important fish stocks are now overfished. Well over one-half of the world's coral reefs are at risk of degradation from human activities. The world has lost nearly half of its original forest cover, and each year an additional 12 million hectares (29.7 million acres) are deforested. These pressures of overharvesting and habitat loss, combined with other pressures such as pollution and the introduction of exotic species (which often outcompete or otherwise threaten native species), are now depleting the very building blocks of living natural resources—the diversity of genes, species, and ecosystems that constitute living systems. *See* ECOSYSTEM; REEF.

The diversity of life is threatened with a massive extinction that could rival the loss of diversity that took place some 65 million years ago around the time that dinosaurs became extinct. Since the 1600s, nearly 500 animal and 650 plant species (mostly vertebrates and flowering plants) are known to have become extinct, and many more species undoubtedly became extinct before they were even identified. The rate of extinctions in groups such as birds and mammals has increased dramatically during this period. Not all species extinctions are necessarily caused by humans. Species went extinct long before humans appeared on the planet, but the background rate of natural extinction is estimated to be nearly 100 times slower than current rates. If the current rate of loss of closed tropical forest (about 1% globally per year) continues until 2030, the number of species these habitats will support will be reduced by 5–10%. This could amount to an additional tens of thousands of extinctions in groups of organisms such as plants.

Current overuse and misuse of renewable living resources cannot be sustained and poses significant economic costs to societies. A critical element of ensuring a sustainable biosphere—one that continues to provide the goods and services from living systems upon which society depends—is thus the conservation of natural resources. "Conservation" is sometimes used synonymously with "protection." More appropriately, however, it refers to the protection and sustainable use of resources. Critical elements of the effective conservation of natural resources include sustainable resource management, establishment of protected areas, and ex situ (off-site) conservation.

Resource management. Some of the most pressing resource conservation problems stem directly from the mismanagement of important biological resources. Many marine fisheries are being depleted, for example, because of significant overcapacity of fishing vessels and a failure of resource managers to closely regulate the harvest. In theory, a renewable resource stock could be harvested at its maximum sustainable yield and maintain constant average annual productivity in perpetuity. In practice, however, fishery harvest levels are often set too high and, in many regions, enforcement is weak, with the result that fish stocks are driven to low levels. A similar problem occurs in relation to the management of timber resources. Short-term economic incentives encourage cutting as many trees as quickly as possible. Without effective resource management policies and enforcement, sound stewardship of forest resources is difficult. *See* FISHERIES ECOLOGY; FOREST MANAGEMENT.

Early conservation efforts primarily involved regulation of wildlife harvests. Historically, hunters often depleted game or commercial species until they were effectively extinct (or completely extinct in the case of such species as passenger pigeons in the United States or the great auk in the north Atlantic). However, in the last half century the

development and application of principles of wildlife management has led to more effective regulation of harvests. It is now rare that commercial harvest of individual species threatens a species with extinction, although, as in the case of fisheries, it is still common for some species to be harvested at levels that substantially deplete the population. *See* FOOD WEB.

A number of steps are being taken to improve resource conservation in managed ecosystems. (1) Considerable scientific research has been undertaken to better understand the natural variability and productivity of economically important resources. (2) Many national and local governments have enacted regulations for resource management practices on public and private lands. These include laws regulating management practices (for example, laws that require timber companies to leave buffers of trees standing near rivers and creeks), as well as laws protecting specific resources, such as endangered species. (3) Problems of resource mismanagement often have stemmed from the fact that local residents, who would have a greater incentive to manage for long-term production, had little rights to local resources, whereas the companies or individuals with rights to harvest resources found it economically beneficial to deplete the resource and move on to new regions. In some of these regions, programs recently have been established either to involve local communities more directly in resource management decisions or to return to them resource ownership rights. (4) Efforts are under way to manage resources on a regional or ecosystem scale using methods that have come to be known as ecosystem management or bioregional management. Since the actions taken in one location often influence species and processes in other locations, traditional resource conservation strategies were often focused too narrowly to succeed.

Protected areas. Regardless of the success in better managing resources, the fact remains that human activities now have reshaped much of the land surface to meet human needs. Human changes in land use and land cover, for example, have transformed one-third to one-half of Earth's icefree surface. This is an area equal to North America, South America, and Africa combined. Clearly, without significant management interventions, many species that once lived in these transformed regions are likely to become extinct. One of the most effective strategies to protect species from extinction is the establishment of protected areas designed to maintain populations of a significant fraction of the native species in a region. Worldwide, 9832 protected areas, totaling more than 9.25 million square kilometers (24 million square miles), cover about 8% of land on Earth. Although these sites are not all managed exclusively for the conservation of species, they play an essential role in protecting species from extinction.

Many problems remain, however, in ensuring effective protected-area conservation networks. For example, several regions with important biodiversity still lack effective protected-area networks. In addition, where protected areas have been designated, human and financial resources are not always available to effectively manage the areas, creating "paper parks." Particularly in developing countries, the establishment of protected areas has resulted in conflicts with local communities that had been dependent upon the areas for their livelihood. These challenges are now being addressed through international efforts, such as the International Convention on Biological Diversity, which aims to increase the financing available for protected areas and to integrate conservation and development needs.

Ex situ conservation. The most effective and efficient means for conserving biological resources is to prevent the loss of important habitats and to manage resources for their long-term productivity of goods and services. In many cases, effective conservation in the field is no longer possible. For example, some species have been so depleted that only a few individuals remain in their natural habitat. In these cases, there is no alternative to the ex situ conservation of species and genetic resources in zoos, botanical gardens, and seed banks. Ex situ collections play important conservation roles as well as serving in public education and research. Worldwide, zoos contain more than 3000 species of birds, 1000 species of mammals, and 1200 species of reptiles, and botanic gardens are believed to hold nearly 80,000 species of plants. These collections hold many endangered species, some of which have breeding populations and thus could potentially be returned to the wild. Genebanks hold an important collection of the genetic diversity of crops and livestock.

Much can be done to better conserve the world's natural resources, but in the face of growing pressure on these resources the challenge is growing. One of the most important steps to bolster any conservation activity is to learn more about the target resource. Scientists have identified only about 1.7 million of an estimated 13 million species of plants, animals, fungi, and microbes on Earth. Many more species are described each year. Unless basic scientific information on this biodiversity is obtained, it is unlikely that conservation will be effective. *See* LAND-USE PLANNING; MINERAL RESOURCES; SOIL CONSERVATION; WATER CONSERVATION. Walter Reid

Bibliography. A. P. Dobson, *Conservation and Biodiversity*, W. H. Freeman, 1996; R. Prescott-Allen, *Caring for the Earth: A Strategy for Sustainable Living*, World Conservation Union/Earthscan, Gland, Switzerland, 1991; L. R. Roberts (ed.), *World Resources 1998–1999: A Guide to the Global Environment*, Oxford University Press, New York, 1998; E. O. Wilson and F. M. Peters (eds.), *Biodiversity*, National Academy Press, 1988.

Constellation

One of the 88 areas into which the sky is divided. Each constellation has a name that reflects its earliest recognition. Though pictures are associated with

the constellations, they have no official status, and constellations have been depicted differently by different artists.

The identifications of the constellations are lost in antiquity. No doubt ancient peoples associated myths with the heavens and imagined pictures connecting or surrounding the bright stars. The names of some constellations have been handed down by the Chaldeans or Egyptians, but most of those that can be seen from midnorthern latitudes have Greek or Roman origins. Star maps are found in various cultures, and constellation figures and myths usually differ from those found in Greek and Roman sources.

Star catalogs and atlases. The catalog of Ptolemy, in Hellenic Alexandria in the second century of the Christian Era, included over 1000 stars grouped into 48 constellations. Many of today's images of constellations are derived from the beautiful engravings by Alexander Mair in Johann Bayer's *Uranometria* (1603) [**Fig. 1**]. Bayer included the constellations listed by Ptolemy and also named 12 new ones containing stars observed on expeditions to the Southern Hemisphere. Bayer originated the scheme of labeling individual stars in constellations with Greek and other letters, roughly in order of brightness, and the genitive form of the constellation name. For example, the bright star Betelgeuse (the second brightest in the constellation Orion) is alpha Orionis (alpha of Orion), and Sirius, the brightest star in the sky, is alpha Canis Majoris (alpha of the Big Dog). In some cases, Bayer labeled stars in order around figures in the sky, as for the Big Dipper. Bayer also used some uppercase (as in the star P Cygni) and lowercase Latin letters. In 1624, Jakob Bartsch placed three new constellations in gaps in Bayer's atlas and separated the southern constellation Crux from Centaurus. Coma Berenices (Berenice's Hair) was also added at about this time, reportedly by Tycho Brahe.

Johannes Hevelius added nine more southern constellations in his 1690 star atlas, *Firmamentum Sobiescianum sive Uranographia*. His figures represent the view of a celestial globe from the outside and thus are reversed from Bayer's. Nicolas Louis de Lacaille added 14 constellations in 1763 from his expedition to the Cape of Good Hope, using names reflecting the mechanical age, such as air pump, microscope, and telescope.

John Flamsteed (1729) and Johan Elert Bode (1801) produced other elegant star atlases, with engravings of constellations. Since the mid-1800s, Ptolemy's largest constellation, Argo Navis, has been divided into Carina, the keel; Puppis, the stern; and Vela, the sails.

Modern star atlases, from the *Millennium Star Atlas* to the *Guide Star Catalogue* compiled in the late 1980s for the use of observers with the Hubble Space Telescope, do not usually show the figures of tradition. *See* ASTRONOMICAL ATLASES; ASTRONOMICAL CATALOGS.

International agreement. In 1928, the International Astronomical Union formally accepted the division

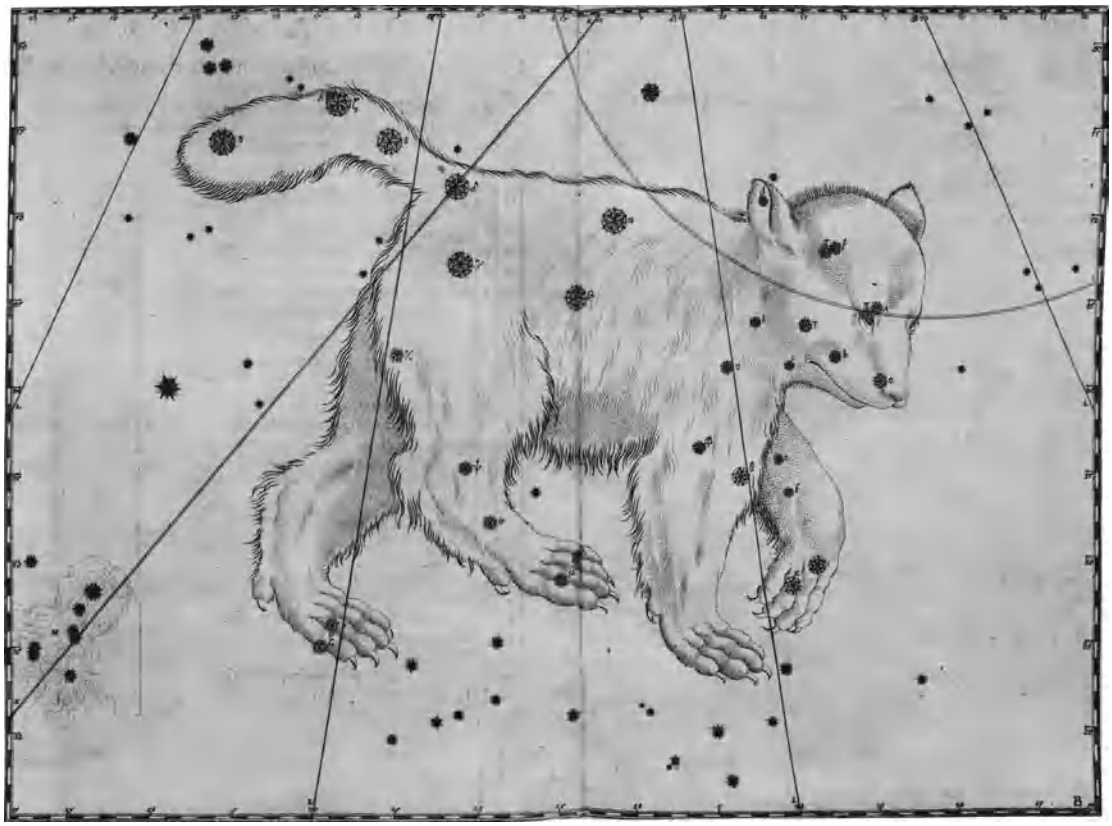

Fig. 1. The constellation Ursa Major from Bayer's star atlas (1603). (*Jay M. Pasachoff collection*)

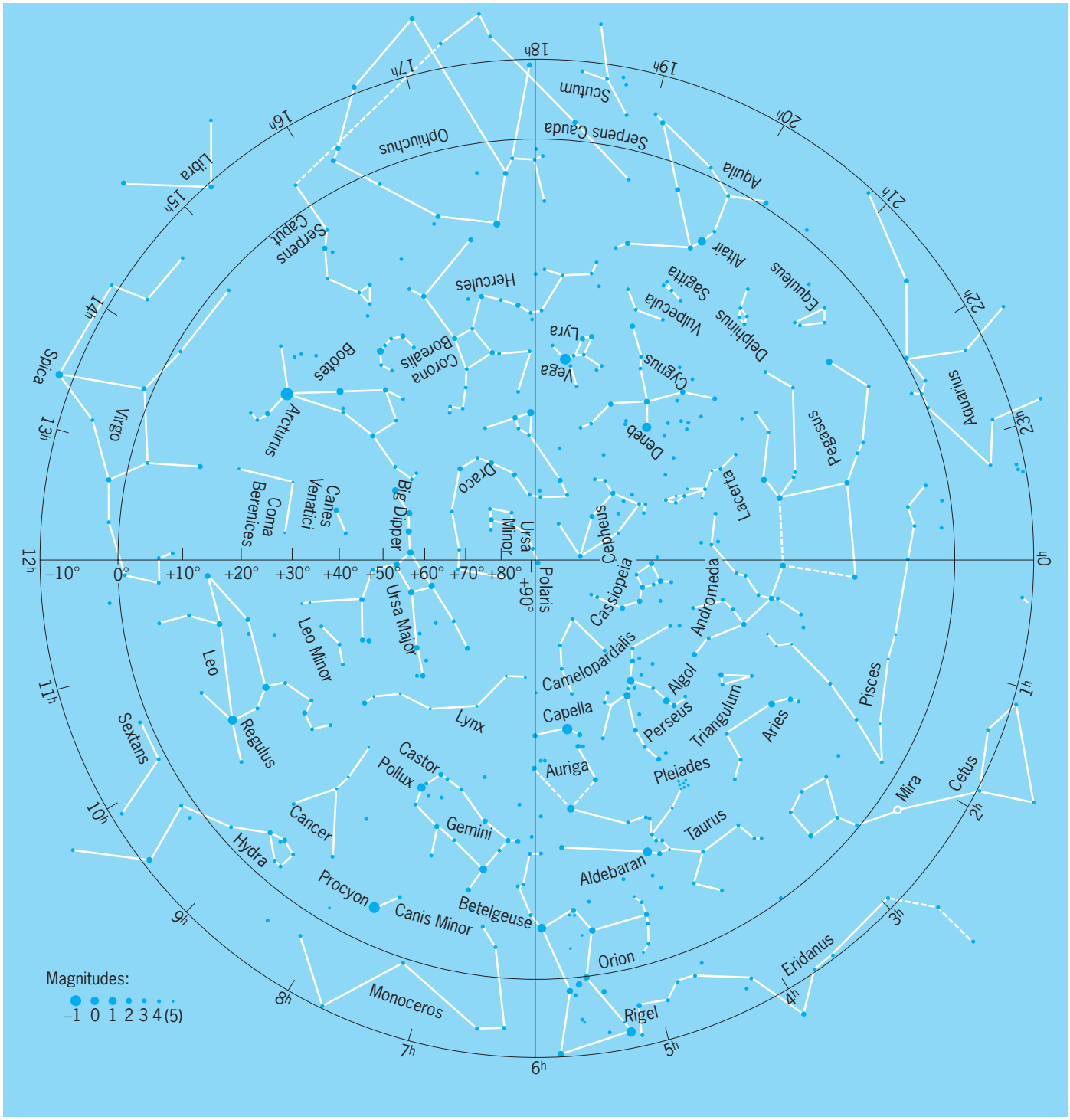

Fig. 2. Constellations of the northern hemisphere; brighter stars are shown by larger dots. (*After J. M. Pasachoff, A Field Guide to the Stars and Planets, Houghton Mifflin, 4th ed., 2000*)

of the sky into 88 constellations (see **table**), with the final list provided 2 years later; each star now falls in only one constellation (**Figs. 2** and **3**). The boundaries follow north-south or east-west celestial coordinates (right ascension and declination lines) from the year 1875; because of precession, the current boundaries do not match rounded values of celestial coordinates. *See* ASTRONOMICAL COORDINATE SYSTEMS; PRECESSION OF EQUINOXES.

Asterisms. Some of the most familiar patterns in the sky are asterisms rather than constellations. For example, the asterism known as the Big Dipper is part of the constellation Ursa Major. The asterism known as the Great Square of Pegasus has three of its corners in Pegasus but the fourth in Andromeda. The Northern Cross is made of stars in Cygnus.

Zodiac. The Sun, the Moon, and the planets move through a band in the sky known as the ecliptic. The

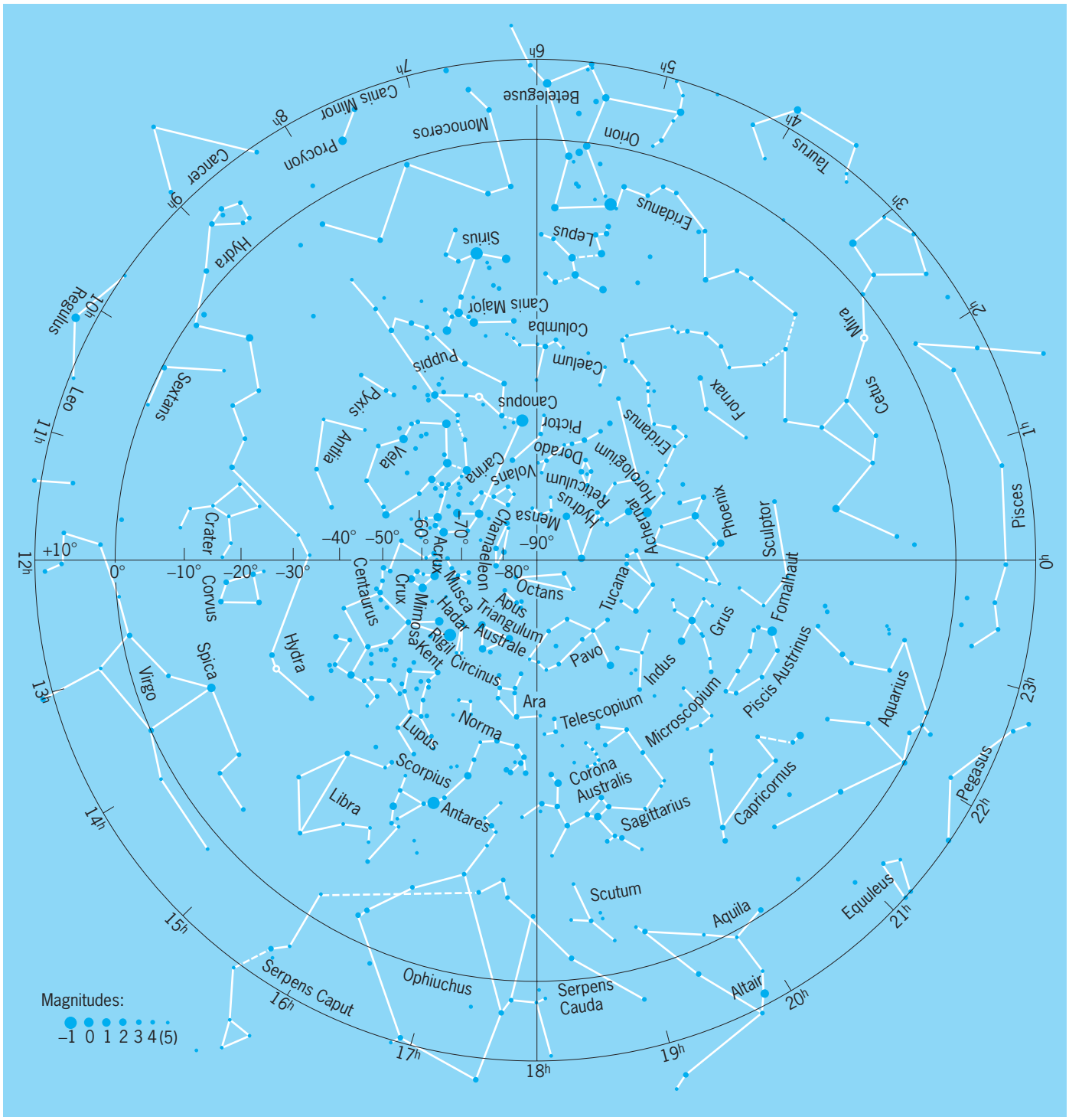

Fig. 3. Constellations of the southern hemisphere; brighter stars are shown by larger dots. (*After J. M. Pasachoff, A Field Guide to the Stars and Planets, Houghton Mifflin, 4th ed., 2000*)

constellations that fall close to the ecliptic through which the Sun traditionally moves are known as the zodiac. Actually, the Sun moves through 13 constellations, and precession has changed the dates at which the Sun passes through the zodiacal constellations. The intersection of the ecliptic and the celestial equator (the extension of the Earth's Equator into the sky) is known as the vernal equinox, or the first point of Aries. It has been used to mark the zero point of the celestial coordinate system. Westward from Aries, the zodiacal constellations are Taurus, the Bull; Gemini, the Twins; Cancer, the Crab; Leo, the Lion; Virgo, the Virgin; Libra, the Scales; Scorpius, the Scorpion; Sagittarius, the Archer; Capricornus, the Seat Goat; Aquarius, the Water Bearer; and Pisces, the Fish. *See* ECLIPTIC; EQUINOX; ZODIAC.

With the advent of equatorial drives on telescopes, astronomers no longer use the constellations to

The constellations*

Latin name	Genitive	Abbreviation	English translation
Andromeda	Andromedae	And	Andromeda†
Antlia	Antliae	Ant	Pump
Apus	Apodis	Aps	Bird of Paradise
Aquarius	Aquarii	Aqr	Water Bearer
Aquila	Aquilae	Aql	Eagle
Ara	Arae	Ara	Altar
Aries	Arietis	Ari	Ram
Auriga	Aurigae	Aur	Charioteer
Boötes	Boötes	Boo	Herdsman
Caelum	Caeli	Cae	Chisel
Camelopardalis	Camelopardalis	Cam	Giraffe
Cancer	Cancri	Cnc	Crab
Canes Venatici	Canum Venaticorum	CVn	Hunting Dogs
Canis Major	Canis Majoris	CMa	Big Dog
Canis Minor	Canis Minoris	CMi	Little Dog
Capricornus	Capricorni	Cap	Goat
Carina	Carinae	Car	Ship's Keel‡
Cassiopeia	Cassiopeiae	Cas	Cassiopeia†
Centaurus	Centauri	Cen	Centaur†
Cepheus	Cephei	Cep	Cepheus†
Cetus	Ceti	Cet	Whale
Chamaeleon	Chamaeleonis	Cha	Chameleon
Circinus	Circini	Cir	Compass
Columba	Columbae	Col	Dove
Coma Berenices	Comae Berenices	Com	Berenice's Hair†
Corona Australis	Coronae Australis	CrA	Southern Crown
Corona Borealis	Coronae Borealis	CrB	Northern Crown
Corvus	Corvi	Crv	Crow
Crater	Crateris	Crt	Cup
Crux	Crucis	Cru	Southern Cross
Cygnus	Cygni	Cyg	Swan
Delphinus	Delphini	Del	Dolphin
Dorado	Doradus	Dor	Swordfish
Draco	Draconis	Dra	Dragon
Equuleus	Equulei	Equ	Little Horse
Eridanus	Eridani	Eri	River Eridanus†
Fornax	Fornacis	For	Furnace
Gemini	Geminorum	Gem	Twins
Grus	Gruis	Gru	Crane
Hercules	Herculis	Her	Hercules†
Horologium	Horologii	Hor	Clock
Hydra	Hydrae	Hya	Hydra† (water monster)
Hydrus	Hydri	Hyi	Sea Serpent
Indus	Indi	Ind	Indian
Lacerta	Lacertae	Lac	Lizard
Leo	Leonis	Leo	Lion
Leo Minor	Leonis Minoris	LMi	Little Lion
Lepus	Leporis	Lep	Hare
Libra	Librae	Lib	Scales
Lupus	Lupi	Lup	Wolf
Lynx	Lyncis	Lyn	Lynx
Lyra	Lyrae	Lyr	Harp
Mensa	Mensae	Men	Table (mountain)
Microscopium	Microscopii	Mic	Microscope
Monoceros	Monocerotis	Mon	Unicorn
Musca	Muscae	Mus	Fly
Norma	Normae	Nor	Level (square)
Octans	Octantis	Oct	Octant
Ophiuchus	Ophiuchi	Oph	Ophiuchus† (serpent bearer)
Orion	Orionis	Ori	Orion†
Pavo	Pavonis	Pav	Peacock
Pegasus	Pegasi	Peg	Pegasus† (winged horse)
Perseus	Persei	Per	Perseus†
Phoenix	Phoenicis	Phe	Phoenix
Pictor	Pictoris	Pic	Easel
Pisces	Piscium	Psc	Fish
Piscis Austrinus	Piscis Austrini	PsA	Southern Fish
Puppis	Puppis	Pup	Ship's Stern‡
Pyxis	Pyxidis	Pyx	Ship's Compass‡
Reticulum	Reticuli	Ret	Net

(*cont.*)

The constellations* (*cont.*)

Latin name	Genitive	Abbreviation	English translation
Sagitta	Sagittae	Sge	Arrow
Sagittarius	Sagittarii	Sgr	Archer
Scorpius	Scorpii	Sco	Scorpion
Sculptor	Sculptoris	Scl	Sculptor
Scutum	Scuti	Sct	Shield
Serpens	Serpentis	Ser	Serpent
Sextans	Sextantis	Sex	Sextant
Taurus	Tauri	Tau	Bull
Telescopium	Telescopii	Tel	Telescope
Triangulum	Trianguli	Tri	Triangle
Triangulum Australe	Trianguli Australis	TrA	Southern Triangle
Tucana	Tucanae	Tuc	Toucan
Ursa Major	Ursae Majoris	UMa	Big Bear
Ursa Minor	Ursae Minoris	UMi	Little Bear
Vela	Velorum	Vel	Ship's Sails‡
Virgo	Virginis	Vir	Virgin
Volans	Volantis	Vol	Flying Fish
Vulpecula	Vulpeculae	Vul	Little Fox

*After J. M. Pasachoff, *Astronomy*: *From the Earth to the Universe*, 6th ed., 2002.
†Proper names.
‡Formerly formed the consellation Argo Navis, the Argonauts' Ship.

locate objects in the sky. Even amateur telescopes are often equipped with computer locators that point the telescope at objects specified by name or catalog number. Jay M. Pasachoff

Bibliography. T. Condos, *Star Myths of the Greeks and Romans*, 1997; G. Lovi and W. Tirion, *Men, Monsters, and the Modern Universe*, 1989; J. M. Pasachoff, *Peterson's First Guide to Astronomy*, 1998; J. M. Pasachoff, *A Field Guide to the Stars and Planets*, 4th ed., 2000; H. A. Rey, *The Stars: A New Way To See Them*, rev. ed., 1988; R. W. Sinnot and M. A. C. Perryman, *Millennium Star Atlas*, 1997; C. Stott, *Celestial Charts: Antique Maps of the Heavens*, 1995; W. Tirion and R. W. Sinnott, *Sky Atlas 2000.0*, 2d ed., 1998; D. J. Warner, *The Sky Explored: Celestial Cartography 1500-1800*, 1979.

Constraint

A restriction on the natural degrees of freedom of a system. If n and m are the numbers of the natural and actual degrees of freedom, the difference $n - m$ is the number of constraints. In principle $n = 3N$, where N is the number of particles, for example, atoms. In practice n is determined by the number of effectively rigid components. *See* DEGREE OF FREEDOM (MECHANICS).

A holonomic system is one in which the n original coordinates can be expressed in terms of m independent coordinates and possibly also the time. It is characterized by frictionless contacts and inextensible linkages. The new coordinates, $q_1, q_2, \ldots, q_m$, are called generalized coordinates. The equations of equilibrium and of motion may be expressed in terms of these coordinates. *See* LAGRANGE'S EQUATIONS.

Nonholonomic systems cannot be reduced to independent coordinates because the constraints are not on the n coordinate values themselves but on their possible changes. For example, an ice skate may point in all directions but at each position it must point along its path. This is a condition between changes (dx,dy) in the two position coordinates (x,y) and the direction angle θ, as in Eq. (1). This cannot

$$dy = \tan\theta \, dx \tag{1}$$

be put in an integrated form, as in Eq. (2), since a

$$f(x, y, \theta) = 0 \tag{2}$$

skater can pass in different directions repeatedly over a point.

The static equilibrium conditions of a constrained system under impressed forces F_1, F_2, and so on are contained in a general statement, the principle of virtual work or virtual displacement. *See* STATICS; VIRTUAL WORK PRINCIPLE.

A moving constraint is one which changes with time, as in the case of a system on a moving platform. Moving constraints differ from the stationary constraints of static equilibrium in being able to do work. For example, a mass constrained to move on the floor of a rising elevator is carried also in the direction of the (vertical) force of constraint which does work on it. Bernard Goodman

Bibliography. H. C. Corben and P. Stehle, *Classical Mechanics*, 2d ed., 1994; H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; D. T. Greenwood, *Principles of Dynamics*, 2d ed., 1988; S. T. Thornton and J. B. Marion, *Classical Dynamics of Particles and Systems*, 5th ed., 2003.

Construction engineering

A specialized branch of civil engineering concerned with the planning, execution, and control of construction operations for such projects as highways, buildings, dams, airports, and utility lines.

Planning consists of scheduling the work to be done and selecting the most suitable construction methods and equipment for the project. Execution requires the timely mobilization of all drawings, layouts, and materials on the job to prevent delays to the work. Control consists of analyzing progress and cost to ensure that the project will be done on schedule and within the estimated cost.

Planning. The planning phase starts with a detailed study of construction plans and specifications. From this study a list of all items of work is prepared, and related items are then grouped together for listing on a master schedule. A sequence of construction and the time to be allotted for each item is then indicated. The method of operation and the equipment to be used for the individual work items are selected to satisfy the schedule and the character of the project at the lowest possible cost. *See* CONSTRUCTION EQUIPMENT; CONSTRUCTION METHODS.

The amount of time allotted for a certain operation and the selection of methods of operation and equipment are generally determined by the equipment that is readily available to the contractor. After the master or general construction schedule has been drawn up, subsidiary detailed schedules or forecasts are prepared from the master schedule. These include individual schedules for procurement of material, equipment, and labor, as well as forecasts of cost and income.

Execution. The speedy execution of the project requires the ready supply of all materials, equipment, and labor when needed. The construction engineer is generally responsible for initiating the purchase of most construction materials and expediting their delivery to the project. Some materials, such as structural steel and mechanical equipment, require partial or complete fabrication by a supplier. For these fabricated materials the engineer must prepare or check all fabrication drawings for accuracy and case of assembly and often inspect the supplier's fabrication.

Other construction engineering duties are the layout of the work by surveying methods, the preparation of detail drawings to clarify the design engineer's drawings for the construction crews, and the inspection of the work to ensure that it complies with plans and specifications.

On most large projects it is necessary to design and prepare construction drawings for temporary construction facilities, such as drainage structures, access roads, office and storage buildings, formwork, and cofferdams. Other problems are the selection of electrical and mechanical equipment and the design of structural features for concrete material processing and mixing plants and for compressed air, water, and electrical distribution systems.

Control. Progress control is obtained by comparing actual performance on the work against the desired performance set up on the master or detailed schedules. Since delay on one feature of the project could easily affect the entire job, it is often necessary to add equipment or crews to speed up the work. *See* PERT.

Cost control is obtained by comparing actual unit costs for individual work items against estimated or budgeted unit costs, which are set up at the beginning of the work. A unit cost is obtained by dividing the total cost of an operation by the number of units in that operation.

Typical units are cubic yards for excavation or concrete work and tons for structural steel. The actual unit cost for any item at any time is obtained by dividing the accumulated costs charged to that item by the accumulated units of work performed.

Individual work item costs are obtained by periodically distributing job costs, such as payroll and invoices to the various work item accounts. Payroll and equipment rental charges are distributed with the aid of time cards prepared by crew forepersons. The cards indicate the time spent by the job crews and equipment on the different elements of the work. The allocation of material costs is based on the quantity of each type of material used for each specific item.

When the comparison of actual and estimated unit costs indicates an overrun, an analysis is made to pinpoint the cause. If the overrun is in equipment costs, it may be that the equipment has insufficient capacity or that it is not working properly. If the overrun is in labor costs, it may be that the crews have too many workers, lack of proper supervision, or are being delayed for lack of materials or layout. In such cases time studies are invaluable in analyzing productivity. William Hershleder

Legal aspects. The construction industry is structured by contracts and other legal obligations. Construction company managers are in daily contact with legal requirements that relate to the contractual provisions of the projects that they are designing or constructing. Other legal considerations relate to subcontractor and materials procurement, the law applied to labor relations, and the legal form of a given company's organization. A firm knowledge of the law is absolutely necessary.

The relationship between owner, designer, and constructor has become so complex and open to misunderstandings that a formal legal framework for performing a given project is essential. Projects are governed by a number of legally binding documents. The construction contract typically consists of an agreement, a description of the general conditions for the project, special or supplementary conditions directly related to the project, and a listing of technical requirements or specifications. These documents, together with the project drawings, provide a legally binding description of the project and establish the relationships between the parties to the contract. On large projects this set of documents can be very large and must be understood in detail by the construction management.

It is not uncommon for disputes to arise regarding the legal interpretation of the contract provisions. In many cases, this has led to an adversarial relationship between the owner, designer, and constructor. The contract itself contains mechanisms for the resolution of disputes. Usually, it is beneficial for all parties to negotiate and to come to a mutual agreement. In extreme cases, the parties become so confrontational that a third party is required to resolve the

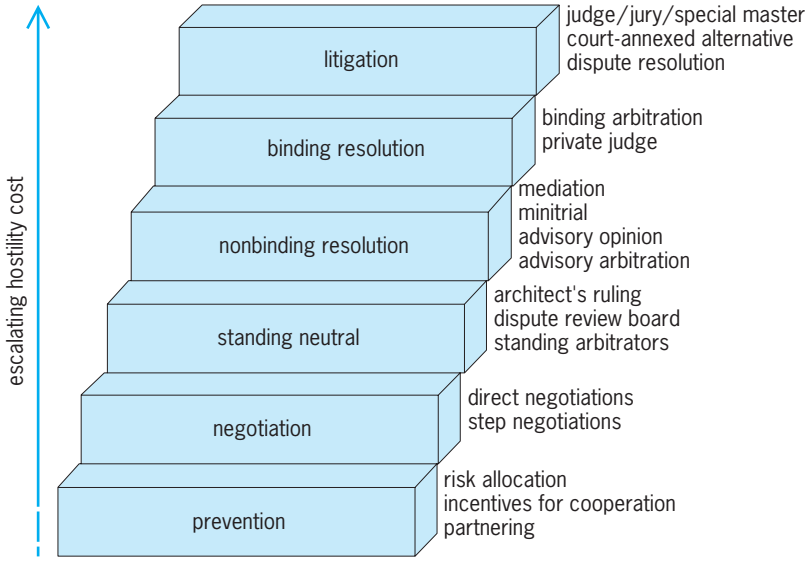

Steps in resolving a construction dispute.

conflict. In the past, this has resulted in the dispute being presented to a civil court in the form of a claim, with a judge or jury deciding the outcome. In some cases, the contract provides that arbitration will be utilized by the parties to settle disputes. In the United States, the American Arbitration Association (AAA) has developed comprehensive procedures for the reconciliation of contract disputes. Other methods of dispute resolution are shown in the **illustration**.

A large body of law regulates the interactions between workers and management as well as working conditions at the construction site. In the United States, labor laws pertain to the operation of labor unions and specific issues such as minimum pay rates on federally funded projects (Davis Bacon Act). The National Labor Relations Act and other laws established the National Labor Relations Board (NLRB), which acts to resolve disputes between labor and management. The construction worksite is regulated by the Occupational Safety and Health Act (OSHA), originally enacted in 1970. This law requires all employers to provide a safe workplace free of hazards that could cause serious harm to employees.

Daniel W. Halpin

Bibliography. D. W. Halpin and R. W. Woodhead, *Construction Management*, 2d ed., John Wiley, New York, 1998; A. L. Iannone and A. M. Civitello, *Construction Scheduling Simplified*, 1985; R. L. Peurifoy, *Construction Planning, Equipment, and Methods*, 4th ed., 1985; J. Sweet, *Legal Aspects of Architecture, Engineering and the Construction Process*, 5th ed., West Publishing, St. Paul, 1994; R. C. Vaughan and S. R. Borgman, *Legal Aspects of Engineering*, 5th ed., Kendall-Hunt Publishing, Dubuque, 1993.

Construction equipment

A wide variety of relatively heavy machines which perform specific construction (or demolition) functions under power. The power plant (which is treated in later paragraphs) is commonly an integral part of an individual machine, although in some cases it is contained in a separate prime mover, for example, a towed wagon or roller. It is customary to classify construction machines in accordance with their functions such as hoisting, excavating, hauling, grading, paving, drilling, or pile driving. There have been few changes for many years in the basic types of machines available for specific jobs, and few in the basic configurations of those that have long been available. Design emphasis for new machines is on modifications that increase speed, efficiency, and accuracy (particularly through more sophisticated controls); that improve operator comfort and safety; and that protect the public through sound attenuation and emission control. The selection of a machine for a specific job is mainly a question of economics and depends primarily on the ability of the machine to complete the job efficiently, and secondarily on its availability.

Hoisting equipment. This class of equipment is used to raise or lower materials from one elevation to another or to move them from one point to another over an obstruction. The main types of hoisting equipment are derricks, cableways, cranes, elevators, and conveyors. *See* BULK-HANDLING MACHINES; HOISTING MACHINES.

Derricks. The two main types of derricks are the guy and the stiff-leg. The former has a mast that is held in a vertical position by guy wires and a boom that can rotate with the mast 360°. In the stiff-leg the mast is tied to two or more rigid structural members, and the rotation of the boom is limited by the position of these members. The derrick is practical only where little mobility is required as in some types of steel erection, excavation of a shaft, or hoisting of materials through a shaft. *See* DERRICK.

Cableway. This is a combination hoist and tram system comprising a trolley that runs on main load cables stretched between two or more towers which may be stationary or tiltable. Since even with towers that can be tilted slightly from side to side the cableway is limited almost entirely to linear movement of materials, its use in construction is restricted to dams, to some dredging and drag-line operations, and in special cases to bridges.

Elevator. The construction elevator, like the passenger elevator, consists of a car or platform that operates within a structural framework and is raised and lowered by cables. Because the elevator can move materials only in a vertical linear direction and because the materials so moved must subsequently be moved horizontally by other means, the elevator is limited to moving relatively light materials on jobs of small areas. *See* ELEVATING MACHINES.

Crane. A crane is basically a fast-moving boom mounted on a frame containing the power supply and mechanisms for moving the boom and for raising and lowering the load-bearing cables that run through sheaves at the top of the boom. Once all cranes in the United States were mobile; that is, they were attached to a chassis mounted on pneumatic tires, crawler treads, or flanged (railroad) wheels. This type of machine is still predominant in all

Fig. 1. Hydraulic truck crane.

operations where mobility is a must. **Figure 1** shows a hydraulic truck crane. However, European and Australian tower cranes and climbing cranes have become popular throughout the United States and the world for high-rise building construction where mobility requirements are negligible. These machines, which carry horizontal or diagonal masts or booms, can "grow" with the construction by climbing with it or by adding sections to their own towers. They have the fast boom and line action of the mobile crane but are not restricted in boom reach.

Conveyor. Occasionally a conveyor is used to raise materials needed for construction of low buildings. This machine, which has a movable endless belt mounted on a frame, can carry only light loads such as bricks and cement sacks and is not practical for buildings of more than one or two floors. A conveyor carrying rock over a railroad yard is shown in **Fig. 2**.

Excavating equipment. This type of equipment is divided into two main classes: standard land excavators and marine dredges; each has many variations.

Standard land excavator. This comprises machines that merely dig earth and rock and place it in separate hauling units, as well as those that pick up and transport the materials. Among the former are power shovels, draglines, backhoes, cranes with a variety of buckets, front-end loaders, excavating belt loaders, trenchers, and the continuous bucket excavator. The second group includes such machines as bulldozers, scrapers of various types, and sometimes the front-end loader.

Power shovel. This has for years been regarded as the most efficient machine for digging into vertical banks and for handling heavy rock. Generally mounted on crawler treads, the power shovel carries a short boom on which rides a movable dipper stick carrying an open-topped bucket. The bucket digs in an upward direction away from the machine and dumps its load by lowering the front of its hinged bottom. **Figure 3** shows power shovels in the process of loading dump trucks.

Front-end loader. This has made heavy inroads into the domain of the power shovel and in time may well replace it in all but the biggest operations. Loaders are built that can handle buckets having 12, 14, and 24 yd^3 (9, 11, and 18 m^3) capacities. The loader, basically an articulated bucket mounted on a series of movable arms at the front of a crawler or rubber-tired tractor, has the advantage of mobility, speed, economy of cost and operation, and light weight. A loader of any given capacity can cost as little as one-half or even one-third as much as a power shovel of the same capacity and requires only one person to run it as opposed to the two or more normally required for the power shovel. **Figure 4** shows a loader operating in a swamp.

Backhoe. This is fundamentally an upside-down power shovel. Its bucket, mounted on a hinged boom, digs upward toward the machine and unloads by being inverted over the dumping point. Backhoes are manufactured as individual machines but more often are attachments mounted on a crawler crane, a rubber-tired truck crane, or a tractor (which usually has a front-end loader bucket at the other end). It is particularly well adapted to the digging of deep trenches.

Dragline. This is a four-sided bucket that is used mainly on soil too wet to support an excavating machine. It is usually carried on a mobile machine mounting a crane boom, but it can also be worked from a cableway where excavating distances are great. The bucket is carried or cast to a point ahead of its support and dropped to the ground. The dragline is designed to excavate and to fill itself as it is drawn across the ground. It empties when its front end is lowered.

Fig. 2. Conveyor carrying rock over a railroad yard.

Fig. 3. Power shovels loading over-the-road rear-dump trucks.

Clamshell. A clamshell is a two-sided bucket that can dig only in a vertical direction. The bucket is dropped while its leaves are open and digs as they close. Formerly, clamshells were suspended from cables on cranes or gantries and worked only under their own weight. Therefore they were imprecise in digging and were practical only in relatively soft soil or loose rock. Presently many clamshell buckets are attached directly to powered booms, are closed hydraulically, and can work with greater precision in dense soils.

Orange peel. This is a multileaved bucket, generally round in configuration. It is normally cable-supported and, like the clamshell, works under its own weight. Like the clamshell, however, it too is being mounted on powered booms and closed hydraulically. One of the commonest uses of the orange peel is the cleaning of small shafts, sewers, and storm drains. Larger sizes are often used for handling broken rock.

Grapple. This group includes a wide variety of special-purpose tined grabs that work on the principle of the clamshell and orange peel. The grapple is used mainly for handling rock, pipe, and logs.

Excavating belt loader. This type of machine is used mainly for loading granular materials from a stockpile, or as a vertical or inclined belt, or as a pair of chains on which are mounted small buckets. Some belt loaders are mobile and can move into a bank under their own power as the material is moved. Others are fixed and are fed by bulldozers which push material onto the belt from above. Their advantage over most excavators lies in their continuous operation.

Trenching machine. This equipment for digging trenches ranges in size from small hand-pushed units

Fig. 4. A loader operating in a swamp.

Fig. 5. A scraper in operation, as it is pulled by a tractor.

used for getting small pipes from streets to private homes to monsters of many tons capable of cutting a 4-ft-wide (1.2-m) swath for 2000 mi (3200 km) of transcontinental pipelines. All, however, work on basically the same principle: A series of buckets, mounted on a chair or a wheel, lift dirt from the ground and deposit it alongside the trench being dug.

Continuous bucket excavator. This excavator works like a trenching machine but is designed to remove earth and loose rock from a wider area and at shallower depths. These machines are self-propelled; they can load trucks in a continuous operation and can switch from a filled to an empty truck in less than 2 s. The continuity of loading and the absence of intermittent activities, such as those necessary in shovel or loader operations, make the continuous bucket excavator a fast, economical machine for large excavating jobs.

Combination excavators and haulers. This classification of construction equipment includes bulldozers, carrying scrapers, self-loading or elevating scrapers, and sometimes front-end loaders.

Bulldozer. This is a curved blade mounted on the front of a crawler or rubber-tired tractor for the purpose of digging and pushing earth and broken rock from one place to another. It is often used in conjunction with a ripper, a heavy tooth or series of teeth mounted on the rear of a tractor to break up rock so that it can be handled by the dozer. Because the amount of dirt the dozer can move at one time is limited by the size of its blade, it is not economical to use the dozer for moving dirt more than a few hundred feet.

Front-end loader. This is sometimes used for moving dirt in the same manner as the dozer, but with the same limitations.

Scraper. This is the most economical of the combination excavating-hauling units for hauls of over a few hundred feet. A wheel-mounted, open-top box or "bowl," the scraper has a hinged bottom that is lowered to scoop up dirt as the machine moves forward. When the bowl is full, the bottom is closed and the unit is moved to the dumping area, where the bottom is lowered and a hydraulic pushing unit in the back of the bowl moves the dirt out through the opening in the bottom, to spread the dirt as needed. Scrapers are pulled by tractors, most of which are an integral part of the unit. A scraper is shown in **Fig. 5**.

Self-loading or elevating scraper. This group of machines does the same work as the conventional scraper, but in front of the bowl is a series of horizontal plates mounted on moving chains. As the chains move, the plates lift the dirt into the bucket in the same manner as a bucket loader. The advantage of the elevating scraper is that it needs less tractor power during the loading operation.

Marine excavator. Usually called a dredge, the marine excavator is an excavating machine mounted on a barge or boat. Two common types are similar to land excavators, the clamshell and the bucket excavator. The suction dredge is different; it comprises a movable suction pipe which can be lowered to the bottom, usually with a fast-moving cutter head at the bottom end. The cutter churns up the bottom so that the pumps on the barge or boat suck up water and the earth suspended in it. When practical, the pumped material is then piped to land, where the solids settle out, allowing the liquids to run back to the body of water being dredged. When distance to land is too great or where there is no empty land on which to discharge the material, the effluent is pumped into barges in which the settled-out solids are towed to a dump area at a distant point.

Hauling equipment. Excavated materials are moved great distances by a wide variety of conveyances. The most common of these are the self-propelled rubber-tired rear-dump trucks, which are classed as

Fig. 6. A 100-ton (90-metric-ton) rear-dump truck.

over-the-road or off-the-road trucks. The main difference between these two is weight and carrying capacity. Many states restrict the total weight of highway-using trucks (truck and load) to as little as 30 tons (27 metric tons). For jobs where trucks remain on site, such as at dams and some highway construction areas, trucks with a carrying capacity of 50 tons (45 metric tons) are common and some capable of carrying loads in excess of 100 tons (90 metric tons) are in use. A 100-ton (90-metric-ton) rear-dump truck is shown in **Fig. 6**. Wagons towed by a rubber-tired prime mover are also used for hauling dirt. These commonly have bottom dumps which permit spreading dirt as the vehicle moves. In special cases side-dump trucks are also used.

Conveyors, while not commonly used on construction jobs for hauling earth and rock great distances, have been used to good advantage on large jobs where obstructions make impractical the passage of trucks. One such conveyor which was utilized on a New Jersey highway job exceeded 2 mi (3.2 km) in length.

Graders. Graders are high-bodied, wheeled vehicles that mount a leveling blade between the front and rear wheels. The principal use is for fine-grading relatively loose and level earth. Commonly considered a maintenance unit, the grader is still an important machine in highway construction. While the configuration of graders has not changed in many years, some models have been built with articulated main frames that permit faster turns of shorter radii and safer operation at the edges of steep slopes.

Pavers. These place, smooth, and compact paving materials. They may be mounted on rubber tires, endless tracks, or flanged wheels. Asphalt pavers embody tamping pads that consolidate the material; concrete pavers use vibrators for the same purpose. Concrete pavers had been required to use forms to contain the material until the advent of the slip-form pavers. These drag their own short forms behind them and consolidate the concrete sufficiently to stand without slumping after removal of the forms. Many high-speed conventional pavers have been produced with highly sophisticated automated control devices that provide extremely accurate and consistent surface elevations.

Drilling equipment. Holes are drilled in rock for wells and for blasting, grouting, and exploring. Drills are classified according to the way in which they penetrate rock, namely, percussion, rotary percussion, and rotary. The first two types are the most common for holes up to 6 in. (15 cm) in diameter and under 40 ft (12 m) in depth. In smaller sizes these drills may be hand-held, but for production work they are mounted on masts which are supported by trucks or, as shown in **Fig. 7**, special tracks mounted on drill rigs. In tunnel work the drills are commonly mounted on platform-supported movable arms or posts that are hydraulically or pneumatically controlled to permit drilling in horizontal or angled positions, with a minimum of effort by the drill operator. The rotary drill is most common in larger sizes, but drills as small as 3 in. (7.5 cm) in diameter are used for coring rock to obtain samples.

Specialized equipment. Among the most common specialized construction equipment are augers, compactors, and pile hammers.

Augers. These are used for drilling wells, dewatering purposes, and cutting holes that can be filled with concrete for foundations. Augers up to 6 ft (1.8 m) in diameter are common.

Compactors. These machines are designed solely to consolidate earth and paving materials to sustain loads greater than those sustained in the uncompacted state. They range in size from small pneumatic hand-held tampers to multiwheeled machines weighing more than 60 tons (50 metric tons). Actual compaction may be achieved by heavily loaded rubber tires or steel rollers. The steel rollers may be solid

Fig. 7. Crawler-mounted drill. (*Gardner-Denver*)

cylinders, or have separate pads or grids, or contain protrusions (sheep-foot roller). Many machines induce a vibratory action into the compacting units so that compaction is achieved by impact force rather than sheer weight.

Pile hammers. These machines are used to install bearing piles for foundations, or sheet piles for cofferdams and retaining walls. The prehistoric pile hammer consisted of a heavy weight lifted by a rope and allowed to drop on top of a pile. This drop hammer is still in use, basically unchanged. More efficient modern hammers are true machines activated by steam, air, oil, hydraulic fluid, or electricty. Conventional hammers contain pistons that are raised by one of these means and then either allowed to fall freely because of gravity (single-acting) or to be driven downward by the same means (double-acting) to impart an impact to the pile. Vibratory hammers have been developed which use electrically activated eccentric cams to vibrate piles into place. Relatively rare are machines that use hydraulic action to install piles.

Road planer. This machine looks like a conventional motor grader, but instead of a blade it carries near its center a horizontal drum that has on its periphery many rows of replaceable hardened steel teeth. As the drum revolves, the teeth pulverize cracked asphalt or decayed concrete pavement to depths of up to 4 in. (10 cm). This permits the placement of a smooth new riding surface on a sound base without raising the road surface elevation, having to raise manhole covers, and reducing the clearances under bridges.

Bore tunneling machine. This is a machine capable of placing water, sewer, or utility pipes accurately for great distances underground. Key to the operation is a series of intermediate hydraulic jacking rings that expand or contract as controlled to permit the movement of any section of pipe independent of the sections before or behind. This independent movement of short sections reduces the friction on the line and therefore the amount of pressure that would be required if the entire pipe were moved from a single push point.

Power plants. Steam is seldom used as a source of power today except in some marine applications and rarely for pile driving. Gasoline and diesel engines of the piston type are the most common source of power for construction machines, with the diesel attaining an increasingly greater prominence for at least two major reasons. First, the size of many construction machines has increased to proportions formerly thought impractical. With this increase in size, the economy of operation and maintenance of the diesel has come to far outweigh its greater initial cost. Second, the increased popularity of the diesel lies in the fact that manufacturers have been able to build lightweight economical diesel engines in very small sizes, thus making practical their application in small compressors, pumps, portable electric power plants, and so on. *See* DIESEL ENGINE.

Electricity is sometimes, though not often, used as a primary source of power for some construction machines, but usually only at the site of large dams and strip-mining operations where mobility is not a primary requirement. Electricity is making its appearance as a secondary source of power on several types of mobile construction machinery, such as trucks and scrapers, which use diesel or gas engines to turn generators that provide power for electric wheels.

A more common type of secondary power, however, is the hydraulic motor which, activated by a gasoline or diesel-powered hydraulic pump, is being used to provide direct power for virtually every movement, including travel, of construction machines.

Some manufacturers have experimented with turbine engines for construction applications. While they are being used with success as primary power sources for large stationary electric power plants on job sites, their practical use in mobile equipment is not evident. *See* CONSTRUCTION ENGINEERING.

Edward M. Young

Bibliography. F. Harris, *Modern Construction and Ground Engineering Equipment*, 2d ed., 1994; J. Russel, *Construction Equipment,* 1985; J. Singh, *Heavy Construction*: *Planning, Equipment, and Methods*, 1994; M. J. Tomlinson, *Pile Design and Construction Practice*, 3d ed., 1987.

Construction methods

The procedures and techniques utilized during construction. Construction operations are generally classified according to specialized fields. These include preparation of the project site, earthmoving, foundation treatment, steel erection, concrete placement, asphalt paving, and electrical and mechanical installations. Procedures for each of these fields are generally the same, even when applied to different projects, such as buildings, dams, or airports. However, the relative importance of each field is not the same in all cases. For a description of tunnel construction, which involves different procedures, *see* TUNNEL

Preparation of site. This consists of the removal and clearing of all surface structures and growth from the site of the proposed structure. A bulldozer is used for small structures and trees. Larger structures must be dismantled.

Earthmoving. This includes excavation and the placement of earth fill. Excavation follows preparation of the site, and is performed when the existing grade must be brought down to a new elevation. Excavation generally starts with the separate stripping of the organic topsoil, which is later reused for landscaping around the new structure. This also prevents contamination of the nonorganic material which is below the topsoil and which may be required for fill. Excavation may be done by any of several excavators, such as shovels, draglines, clamshells, cranes, and scrapers. For a discussion of their application *see* CONSTRUCTION EQUIPMENT.

Efficient excavation on land requires a dry excavation area, because many soils are unstable when wet

Fig. 1. Steel erection with guy derricks on a high structure. (*Bethlehem Steel Co.*)

and cannot support excavating and hauling equipment. Dewatering becomes a major operation when the excavation lies below the natural water table and intercepts the ground-water flow. When this occurs, dewatering and stabilizing of the soil may be accomplished by trenches, which conduct seepage to a sump from which the water is pumped out. Dewatering and stabilizing of the soil may in other cases be accomplished by wellpoints and electroosmosis. *See* WELLPOINT SYSTEMS.

Some materials, such as rock, cemented gravels, and hard clays, require blasting to loosen or fragment the material. Blast holes are drilled in the material; explosives are then placed in the blast holes and detonated. The quantity of explosives and the blast-hole spacing are dependent upon the type and structure of the rock and the diameter and depth of the blast holes.

After placement of the earth fill, it is almost always compacted to prevent subsequent settlement. Compaction is generally done with sheep's-foot, grid, pneumatic-tired, and vibratory-type rollers, which

are towed by tractors over the fill as it is being placed. Hand-held, gasoline-driven rammers are used for compaction close to structures where there is no room for rollers to operate.

Foundation treatment. When subsurface investigation reveals structural defects in the foundation area to be used for a structure, the foundation must be strengthened. Water passages, cavities, fissures, faults, and other defects are filled and strengthened by grouting. Grouting consists of injection of fluid mixtures under pressure. The fluids subsequently solidify in the voids of the strata. Most grouting is done with cement and water mixtures, but other mixture ingredients are asphalt, cement and clay, and precipitating chemicals. *See* FOUNDATIONS.

Steel erection. The construction of a steel structure consists of the assembly at the site of mill-rolled or shop-fabricated steel sections. The steel sections may consist of beams, columns, or small trusses which are joined together by riveting, bolting, or welding. It is more economical to assemble sections of the structure at a fabricating shop rather than in the field, but the size of preassembled units is limited by the capacity of transportation and erection equipment. The crane is the most common type of erection equipment, but when a structure is too high or extensive in area to be erected by a crane, it is necessary to place one or more derricks on the structure to handle the steel (**Fig. 1**). In high structures the derrick must be constantly dismantled and reerected to successively higher levels to raise the structure. For river bridges the steel may be handled by cranes on barges, or, if the bridge is too high, by traveling derricks which ride on the bridge being erected. Cables for long suspension bridges are assembled in place by special equipment that pulls the wire from a reel, set up at one anchorage, across to the opposite anchorage, repeating the operation until the bundle of wires is of the required size.

Concrete construction. Concrete construction consists of several operations: forming, concrete production, placement, and curing. Forming is required to contain and support the fluid concrete within its desired final outline until it solidifies and can support itself. The form is made of timber or steel sections or a combination of both and is held together during the concrete placing by external bracing or internal ties (**Fig. 2**). The forms and ties are designed to withstand the temporary fluid pressure of the concrete.

The usual practice for vertical walls is to leave the forms in position for at least a day after the concrete is placed. They are removed when the concrete has solidified or set. Slip-forming is a method where the form is constantly in motion, just ahead of the level of fresh concrete. The form is lifted upward by means of jacks which are mounted on vertical rods embedded in the concrete and are spaced along the perimeter of the structure. Slip forms are used for high structures such as silos, tanks, or chimneys.

Concrete may be obtained from commercial batch plants which deliver it in mix trucks if the job is close to such a plant, or it may be produced at the job site. Concrete production at the job site requires the erection of a mixing plant, and of cement and aggregate receiving and handling plants. Aggregates are sometimes produced at or near the job site. This requires opening a quarry and erecting processing equipment such as crushers and screens.

Fig. 2. Concrete form. (*Superior Concrete Accesories*)

Concrete is placed by chuting directly from the mix truck, where possible, or from buckets handled by means of cranes or cableways, or it can be pumped into place by special concrete pumps.

Curing of exposed surfaces is required to prevent evaporation of mix water or to replace moisture that does evaporate. The proper balance of water and cement is required to develop full design strength. *See* CONCRETE.

Concrete paving for airports and highways is a fully mechanized operation. Batches of concrete are placed between the road forms from a mix truck or a movable paver, which is a combination mixer and placer. A series of specialized pieces of equipment, which ride on the forms, follow to spread and vibrate the concrete, smooth its surface, cut contraction joints, and apply a curing compound. *See* HIGHWAY ENGINEERING; PAVEMENT.

Asphalt paving. This is an amalgam of crushed aggregate and a bituminous binder. It may be placed on the roadbed in separate operations or mixed in a mix plant and spread at one time on the roadbed. Then the pavement is compacted by rollers.

William Hershleder

Contact aureole

The zone of alteration surrounding a body of igneous rock caused by heat and volatiles given off as the magma crystallized. Changes can be in mineralogy, texture, or elemental and isotopic composition

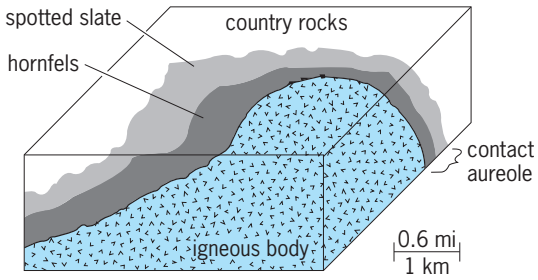

Fig. 1. Zoned contact aureole developed around an igneous body.

of the original enclosing (country or wall) rocks, and progressively increase closer to the igneous contact. The contact aureole is the shell of metamorphosed or metasomatized rock enveloping the igneous body (**Fig. 1**). The ideal contact aureole forms locally around a single magma after it is emplaced.

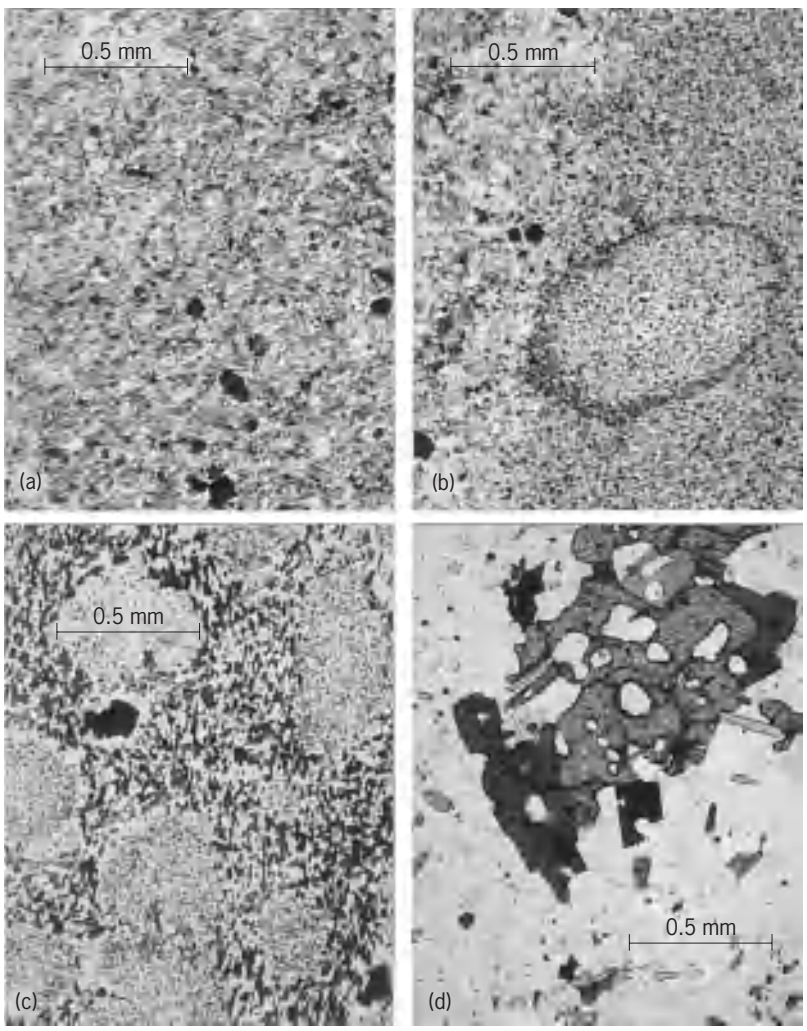

Fig. 2. Microscopic appearance of progressive metamorphism in a contact aureole. (*a*) Original slate country rock comprising chlorite, muscovite, albite, and quartz. (*b*) Spotted slate with the same mineralogy and texture as country rock but with the addition of cordierite plus biotite spots. (*c*) Hornfels with poikiloblastic porphyroblasts of cordierite in a matrix of biotite, muscovite, quartz, and plagioclase. (*d*) Granoblastic hornfels with garnet, biotite, cordierite, alkali feldspar, plagioclase, and quartz.

Metamorphism over a much larger area can result from coalescing of several contact aureoles. This is termed a contact-regional metamorphic aureole and is thought responsible for the regional metamorphism of several mountain areas. Other contact aureoles develop at greater depths and may be physically emplaced to shallower levels along with the igneous body. These are termed dynamothermal aureoles. *See* IGNEOUS ROCKS; MAGMA; METAMORPHIC ROCKS; METAMORPHISM; METASOMATISM; PLUTON.

Extent. The aureole extends from the igneous contact, where the metamorphic effects are the greatest, out into the country rocks to where the temperature or heat energy is insufficient to effect any changes. This temperature lies between 400 and 750°F (200 and 400°C), and actual widths of contact aureoles range from several inches to miles.

Most important in determining the metamorphic intensity and the extent of the aureole are the heat content of the magma and the temperature contrast between the magma and country rock. Initial temperature of the magma depends on its type; metamorphic effects around gabbro magma (~2200°F or 1200°C) are generally more extensive and intensive than those around granite (~1400°F or 750°C). Better developed aureoles are found around larger igneous bodies, magmas which were not extensively crystallized before emplacement, and convecting magmas. Under the right circumstances, the highest temperature in the aureole can approach that of the magma. At deeper crustal levels, country rock temperatures increasingly coincide with those of common magmas, and the contact effects are correspondingly reduced until they are indistinguishable from metamorphic changes brought about by normal burial. Thus contact aureoles are most commonly found around shallow igneous bodies where magma temperatures exceed those of the country rock. They are considered to have low-pressure metamorphism, occurring at pressures less than 5 kilobars (500 megapascals), but usually much less. This corresponds to depths of less than 12 mi (20 km).

Factors of heat transfer also affect the extent and intensity of the metamorphism in the contact aureole. Geometry and orientation of the igneous body are important; intense metamorphism occurs at acute corners of the body and wider aureoles occur at the roof of an igneous body, where most heat is escaping, as opposed to the floor. Wider aureoles and less intense metamorphism occur in country rocks with good thermal conductivity. Heat transfer by conduction produces narrower, more regular, aureoles in contrast to heat transfer by fluid convection. Convective heat transport can produce wide or very narrow aureoles depending on the fluid content of the magma, fluid content of the wall rock, permeability of the system, and direction of fluid flow. Dry magmas and impermeable systems tend to have narrow aureoles. Cool fluid from the wall rocks can flow into the magma, cooling it and preventing development of the aureole. Conversely, hot

fluid from the magma or wall rock fluid heated by the magma flowing outward can produce an extensive aureole.

Aureoles are most pronounced where wall rocks are easily affected by metamorphism because of their composition, texture, or previous metamorphism. Shale and limestone generally show a more conspicuous metamorphism than does sandstone. Aureoles are best developed in fine-grained and previously unmetamorphosed or weakly metamorphosed rocks; they may not be detected in intensely metamorphosed rocks because minerals of the wall rocks are stable at the conditions imposed upon them by the magma. Heat absorption by energy-consuming reactions will lessen the intensity of metamorphism elsewhere in the aureole. If the rate of heating and subsequent cooling exceeds the rate of recrystallization or reaction in the aureole, metamorphic effects will be reduced.

Rocks. Weakly contact-metamorphosed wall rocks may be only discolored or indurated. Contact aureoles with more intense metamorphism are progressively zoned, with appearance of new textures and minerals toward the igneous body (**Fig. 2**). At the outer edge of the aureole, textural changes include development of scattered aggregates of existing or new minerals, which are larger than the matrix minerals. If the host rock is shale, these recrystallized rocks become spotted slates. Further inward there is sequential appearance of new minerals, more extensive recrystallization, and color changes. Average grain size may increase but is still small, usually less than 0.08 in. (2 mm). Fine grain size in contact metamorphism results from short-duration heating and lack of deformation. Well within the aureole, where new mineral growth and recrystallization is complete, older structures (bedding or schistosity) may be eliminated. Mineral grains tend to become equidimensional with little or no preferred orientation and lack their crystal form in an interlocking mosaic (granoblastic texture); the rock is called a hornfels. The term hornfels is applied in a wide sense to any contact-metamorphosed rock and made specific by mineral or rock compositional modifiers. The massive texture of hornfelses results from the lithostatic or hydrostatic nature of the pressure; directed pressure as in regional metamorphism is absent. However, some hornfelses develop a preferred alignment because of intrusion of the adjacent magma. Hornfelses typically have a few larger mineral grains (porphyroblasts) which can contain abundant inclusions of the smaller matrix minerals (poikiloblastic texture). Adjacent to the igneous contact, melting of the wall rocks can occur. *See* HORNFELS; PORPHYROBLAST; ROCK.

Mineral facies. Contact metamorphism can occur over a wide range of temperatures, pressures, or chemical gradients in rocks of any composition. Thus any mineral assemblage or facies of metamorphic rocks can be found. However, the nature of contact aureoles results in minerals characteristic of low to moderate pressures and moderate to high temperatures usually in common rock types: shales, basalt, limestone, and sandstone. Characteristic minerals developed in shales are andalusite, sillimanite, cordierite, biotite, orthopyroxene, and garnet. At the highest temperatures, tridymite, sanidine, mullite, and pigeonite form; whereas in limestone unusual calcium silicates form, including tilleyite, spurrite, rankinite, larnite, merwinite, akermanite, monticellite, and melilite. *See* FACIES (GEOLOGY).

Chemical alterations. Compositional changes in a contact aureole range from none to great, but as a rule, contact metamorphism entails relatively little change in bulk rock composition. Because metamorphic changes are largely brought about by heat, contact aureoles are often termed thermal aureoles. However, there is a tendency for volatiles (water, carbon dioxide, oxygen) and alkalies (sodium, potassium) to be lost from rocks in the aureole. Stable isotope compositions (oxygen, sulfur) change in response to the thermal gradient and flow of fluids through the rocks. In some cases, volatiles (boron, fluorine, and chlorine) and other elements from the crystallizing magma are gained.

Some wall rock compositions, such as limestone, can be greatly changed and form rocks termed skarn. These contact aureoles are economically important because they often contain ore deposits of iron, copper, tungsten, graphite, zinc, lead, molybdenum, and tin. Conversely, the magma can incorporate material from the wall rocks by assimilation or mixing with any partial melts formed. Mixing results in elemental and isotopic contamination of the magma, crystallization of different minerals from the melt, and hybrid rock types at the margin of the igneous body. *See* ORE AND MINERAL DEPOSITS; PNEUMATOLYSIS; SKARN.

J. Alexander Speer

Bibliography. A. Miyashiro, *Metamorphic Petrology*, 1994; V. V. Reverdatto, *The Facies of Contact Metamorphism*, 1970; A. Spry, *Metamorphic Textures*, 1969; H. G. F. Winkler, *Petrogenesis of Metamorphic Rocks*, 5th ed., 1985.

Contact condenser

A device in which a vapor is brought into direct contact with a cooling liquid and condensed by giving up its latent heat to the liquid. In almost all cases the cooling liquid is water, and the condensing vapor is steam. Contact condensers are classified as jet, barometric, and ejector condensers (see **illus.**). In all three types the steam and cooling water are mixed in a condensing chamber and withdrawn together. Noncondensable gases are removed separately from the jet condenser, entrained in the cooling water of the ejector condenser, and removed either separately or entrained in the barometric condenser. The jet condenser requires a pump to remove the mixture of condensate and cooling water and a vacuum breaker to avoid accidental flooding. The barometric condenser is self-draining. The ejector condenser converts the energy of high-velocity injection water

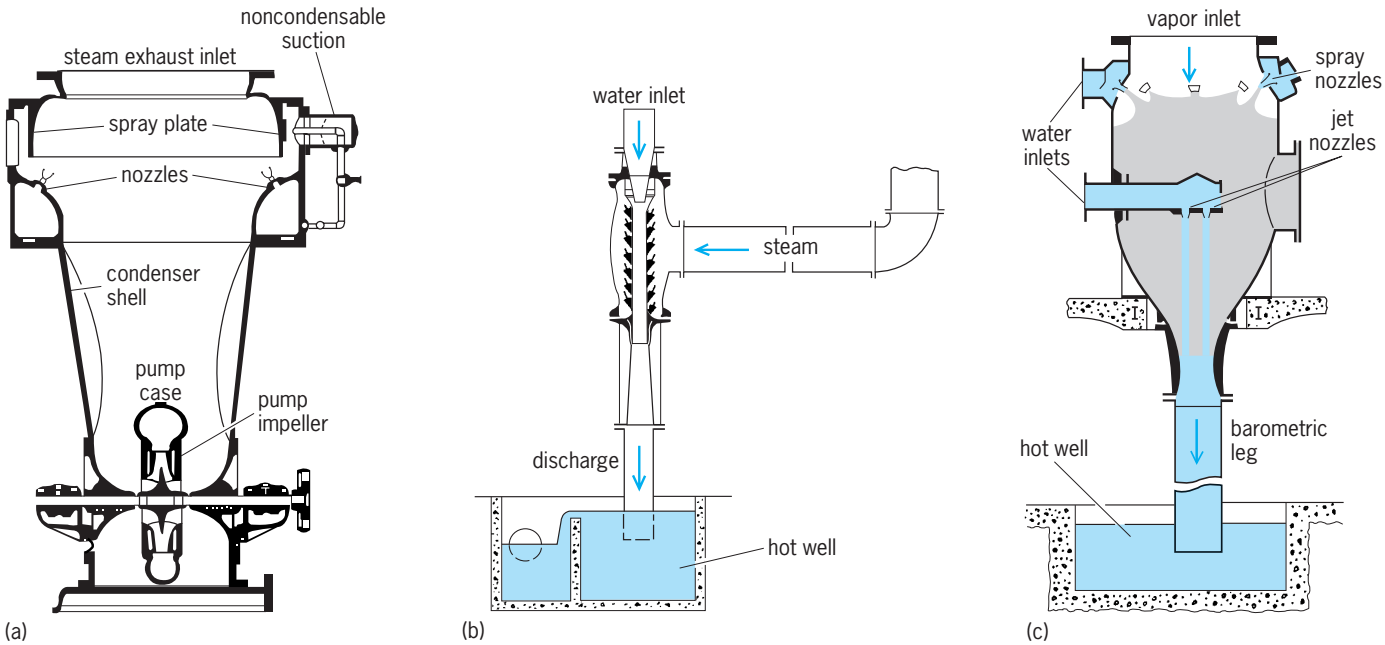

Three basic types of contact condenser: (*a*) low-level jet condenser (*C. H. Wheeler Manufacturing Co.*); (*b*) single-jet ejector condenser; (*c*) multijet barometric condenser (*Schutte and Koerting Co.*).

to pressure in order to discharge the water, condensate, and noncondensables at atmospheric pressure. *See* VAPOR CONDENSER. Joseph F. Sebald

Contact potential difference

An electrostatic potential that exists between samples of two dissimilar electrically conductive materials (metals or semiconductors with different electron work functions) that have been brought into thermal equilibrium with each other, usually through a physical contact. Although normally measured between two surfaces which are not in contact, this potential is called the contact potential difference. Its origin can be described in terms of the process necessary to bring the samples into equilibrium. Initially it is expected that mobile charge carriers (electrons or holes) will migrate from one sample to the other. If there is a net flow of electrons from material A to material B (see **illus.**), material B will become negatively charged and material A will become positively charged, assuming that they were originally neutral. This process is self limiting because a potential difference between the two samples will develop due to the charge separation and will grow to a value sufficient to stop further motion of the electrons from A to B.

In a metal or a semiconductor, the electrons are distributed in energy such that virtually all of them exist at or below a "sea level" called the Fermi level. When any combination of metals and semiconductors are put into equilibrium with one another, the Fermi levels in all will coincide. The contact potential difference between materials is that value necessary to raise or lower the potential energies of the electrons to produce a common Fermi sea level.

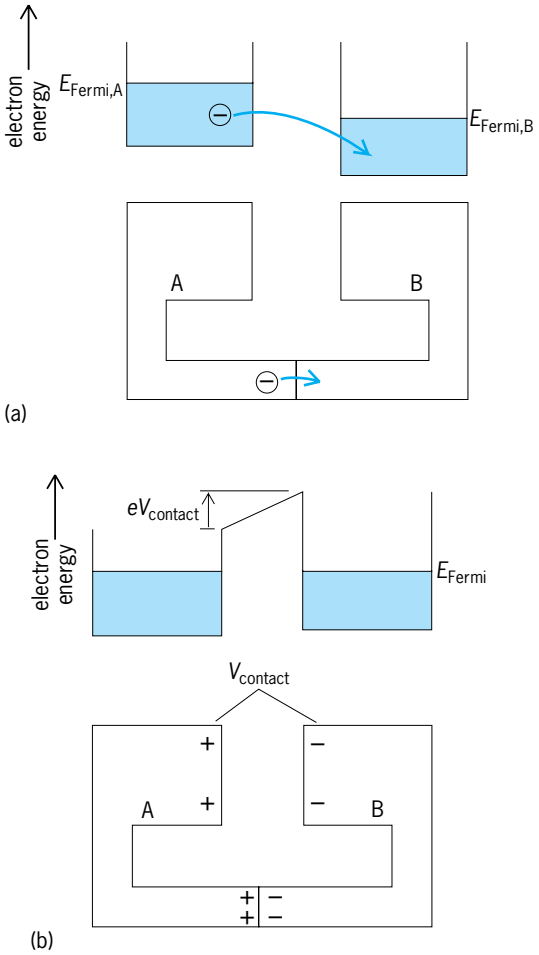

Development of a contact potential as two conductive materials are brought into thermal equilibrium. (*a*) Initial charge transfer. (*b*) Thermal equilibrium. Diagrams show corresponding electron energy distributions. E_{Fermi} = Fermi level; $V_{contact}$ = contact potential difference; e = electron charge.

Since they are then at the same energy, electrons at the Fermi level in either material will have no net force on them, that is, no reason to travel to the other material. Because it causes no net force on the equilibrium distribution of electrons, contact potential difference cannot be directly measured with an ordinary voltmeter. Nevertheless, it profoundly affects the behavior of a number of electronic devices. *See* FREE-ELECTRON THEORY OF METALS; SCHOTTKY BARRIER DIODE; SEMICONDUCTOR; WORK FUNCTION (ELECTRONICS). James E. Nordman

Bibliography. R. H. Bube, *Electrons in Solids*, 1992; L. Solymar and D. Walsh, *Lectures on the Electrical Properties of Materials*, 1993.

Continent

A protuberance of the Earth's crustal shell with an area of several million square miles and with sufficient elevation above neighboring depressions (the ocean basins) so that much of it is above sea level.

The term was originally applied to the most extensive continuous land areas of the globe. The great majority of maps now in use imply that the boundaries of continents are their shorelines. From the geological point of view, however, the line of demarcation between a continent and an adjacent ocean basin lies offshore, at distances ranging from a few to several hundred miles, where the gentle slope of the continental shelf changes somewhat abruptly to a steeper declivity. This change occurs at depths ranging from a few to several hundred fathoms (1 fathom = 6 feet) at different places around the periphery of various continents. *See* CONTINENTAL MARGIN.

On such a basis, numerous offshore islands, including the British Isles, Greenland, Borneo, Sumatra, Java, New Guinea, Tasmania, Taiwan (formerly Formosa), Japan, and Sri Lanka (formerly Ceylon), are parts of the nearby continent. Thus, there are six continents: Eurasia (Europe, China, and India are parts of this largest continent), Africa, North America, South America, Australasia (including Australia, Tasmania, and New Guinea), and Antarctica.

All continents have similar structural features but display great variety in detail. Each includes a basement complex (shield) of metamorphosed sedimentary and volcanic rocks of Precambrian age, with associated igneous rocks, mainly granite. Originally formed at considerable depths below the surface, this shield was later exposed by extensive erosion, then largely covered by sediments of Paleozoic, Mesozoic, and Cenozoic age, chiefly marine limestones, shales, and sandstones. In at least one area on each continent, these basement rocks are now at the surface (an example is the Canadian Shield of North America). In some places they have disappeared beneath a sedimentary platform occupying a large fraction of the area of each continent, such as the area in the broad lowland drained by the Mississippi River in the United States. In each continent there are long belts of mountains in which thick masses of sedimentary rocks have been compressed into folds and broken by faults. *See* CONTINENTS, EVOLUTION OF.

Presumably because of the large volume of acidic aluminum-rich granitic rocks beneath the surface of each continent, the velocity of transmission of earthquake vibrations in continental segments of the Earth's crust is slower than in oceanic segments, in which alkaline, iron-rich basaltic rather than granitic rocks are the dominant type. For the same reason the average density of the rocks composing continents is less than that of the rocks beneath the deep-sea floors. *See* EARTHQUAKE.

Continents are the less dense, subaerially exposed portion of the large and not-so-large plates that make up the Earth's lithosphere, or outer shell of rigid rock material. As such, continents together with part of the ocean's floor are intimately joined portions of the lithospheric plates. As plates rip apart and migrate horizontally over the Earth's surface, so too do continents rip apart and migrate, sometimes colliding with other continental segments scores of millions of years later. Mountain systems, such as the Appalachian-Ouachita, the Arbuckle-Wichita, and the Urals systems, are now believed to represent the sutures of former continents attached to their respective plates which collided long ago. The Red Sea and the linear rift-volcano-lakes district of Africa are also believed by many to manifest continental ripping and early continental drifting. Such continental collision and accretion are believed to have occurred throughout most of the Earth's history. *See* AFRICA; ANTARCTICA; ASIA; AUSTRALIA; EARTH, GRAVITY FIELD OF; EUROPE; ISOSTASY; LITHOSPHERE; NORTH AMERICA; PLATE TECTONICS; SEISMOLOGY; SOUTH AMERICA. Donald Lee Johnson

Bibliography. R. D. Van Der Hilst and W. F. McDonough, eds., *Composition, Deep Structure, and Evolution of Continents (Developments in Geotectonics, 24)*, 2000.

Continental drift

The concept that the world's continents once formed part of a single mass and have since drifted into their present positions. Although it was outlined by Alfred Wegener in 1912, the idea was not particularly new. Paleontological studies had already demonstrated such strong similarities between the flora and fauna of the southern continents between 300,000,000 and 150,000,000 years ago that a huge supercontinent, Gondwana, containing South America, Africa, India, Australia, and Antarctica, had been proposed. However, Gondwana was thought to be the southern continents linked by land bridges, rather than contiguous units.

Wegener's ideas were almost universally rejected at a meeting of the American Association of Petroleum Geologists in 1928. His evidence and timing were undoubtedly wrong in many instances, and only a few scientists accepted his general concepts and remained sympathetic to the ideas. The fundamental objection was the lack of a suitable mechanism. Studies of the passage of earthquake waves through the Earth had shown that, whereas the core

is liquid, the Earth's mantle and crust are solid. It was therefore difficult to find sufficient force to move the continents through the mantle and oceanic crust, and even more difficult to picture them moving without leaving obvious evidence of their passage.

Almost simultaneously with the temporary eclipse of Wegener's theory, Arthur Holmes was considering a mechanism that is still widely accepted. With the knowledge that radioactive minerals occurring within the Earth could result in an internal zone of slippage, Holmes conceived the idea of convective currents within the Earth's mantle which were driven by the radiogenic heat produced by radioactive minerals within the mantle. Since the rates of motion were slow, but vast and inexorable, the Earth appeared solid to the rapid passage of seismic waves, but behaved plastically to such slow motions. This accounted for the apparent discrepancy between the seismic observations and the drift of continents. At that time, Holmes's ideas, like those of Wegener, were largely ignored, as it was no longer thought that continental drift was worthy of further consideration. Nonetheless, several geologists, particularly those living in the Southern Hemisphere, continued to believe the theory and accumulate more data in its support, particularly from the study of the similarities in fossils and rock types on now separated continents.

By the 1950s, convincing evidence had accumulated, with studies of the magnetization of rocks, paleomagnetism, beginning to provide numerical parameters on the past latitude and orientation of the continental blocks. Early work in North America and Europe clearly indicated how these continents had once been contiguous and had since separated. The discovery of the mid-oceanic ridge system also provided more evidence for the geometric matching of continental edges, but the discovery of magnetic anomalies parallel to these ridges and their interpretation in terms of sea-floor spreading finally led to almost universal acceptance of continental drift as a reality. In the 1970s and 1980s, the interest changed from proving the reality of the concept to applying it to the geologic record, leading to a greater understanding of how the Earth has evolved through time. The fundamental change in concept was that not only have the continents drifted, but the continents are merely parts of thicker tectonic plates, comprising both oceanic and continental crust, with 50–300 km (31–186 mi) of the Earth's mantle moving along with them. *See* CONTINENTS, EVOLUTION OF; EARTH, CONVECTION IN; PALEOMAGNETISM; PLATE TECTONICS; GEODYNAMICS.

D. H. Tarling

Bibliography. H. E. LeGrand, *Drifting Continents and Shifting Theories*, 1988; U. B. Marvin, *Continental Drift: The Evolution of a Concept*, 1973; F. Press and R. Siever, *Understanding the Earth*,1993; D. H. Tarling and M. P. Tarling, *Continental Drift*, 2d ed., 1975; T. H. Van Andel, *New Views on an Old Planet*: *A History of Global Change*, 1994.

Continental margin

The submerged portions of the continental masses on crustal plates, including the continental shelf, the continental slope, and the continental rise. All continental masses have some continental margin, but there is great variety in the size, shape, and geology depending upon the tectonic setting.

The most common settings are the trailing-edge margin and the leading-edge margin. The former is associated with tectonic stability, as exemplified by the Atlantic side of the North American landmass (**Fig. 1**). Here the margin is wide and geologically relatively uncomplicated, with thick sequences of coastal plain to shallow marine strata dipping slightly toward the ocean basin. By contrast, the leading-edge margin (for example, the Pacific side of the United States) is narrow, rugged, and geologically complicated. The global distribution of these widespread continental margin types is controlled by the plate tectonic setting in which the landmass resides. Some major landmasses, such as Australia, are surrounded by wide margins, but most, such as North and South America, have some of both types. *See* PLATE TECTONICS.

Any consideration of the continental margin must include a general understanding of global seal-level history over the past few million years. As glaciers expanded greatly just over 2 million years ago, sea level was lowered more than 300 ft (100 m). The cyclic growth and decay of glaciers during this period caused the shoreline to move from near its present position to near the edge of the present continental shelf on multiple occasions. These sea-level changes had a profound effect on the entire continental margin, particularly the shelf and the rise. During times of glacial advance, the coast was near the shelf edge, causing large volumes of river-borne sediment to flow down the continental slope and pile up on the rise; deltas were poorly developed for lack of place for sediment to accumulate. During times of high sea-level stand similar to the present time, little sediment crossed the shelf and large volumes of riverine sediment accumulated in large fluvial deltas. *See* DELTA.

Continental shelf. This part of the margin is simply an extension of the adjacent landmass. It is characterized by a gentle slope and little relief except for shelf valleys (Fig. 1), which are old rivers that were active during times of low sea level. Because of the great variability in the depth, width, and gradient, it is misleading to give average values for these characteristics. The width ranges from a few kilometers to hundreds of kilometers, with the Florida peninsula being a good example of both. The Atlantic shelf is very narrow, and the Gulf side is very wide. The gradient also ranges, generally between 6 and 22 ft/mi (1 and 4 m/km); it is so gentle that it would not be possible for a person walking over it to detect a change. The outer limit of the shelf shows a distinct change in gradient to the much steeper slope. This change may occur at depths of less than 300 ft (100 m) or up to nearly 1000 ft (300 m).

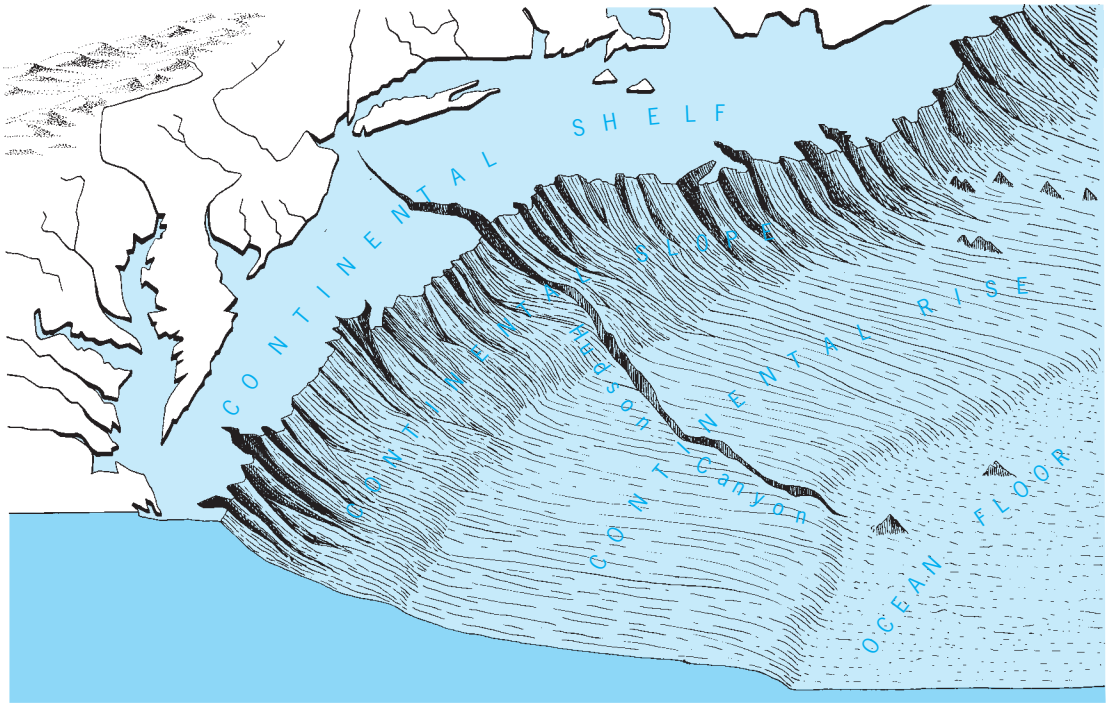

Fig. 1. Continental margin off the northeastern United States.

Unfortunately, legal definitions typically use depth limits to define the shelf for economic purposes without paying attention to the geologic or geomorphic boundaries. The shelf environment may include complete seas, such as the North Sea between Great Britain and Europe.

The modern continental shelf is covered with a combination of Holocene sediments that were deposited as sea level rose over the last 18,000 years, when the last extensive glaciers melted, and older sediments that have yet to be covered. The outer one-half to two-thirds of the shelf has relict sediments that were deposited in coastal environments when sea level was lower, and that were drowned without being buried when sea level rose rapidly in the early Holocene. Intense storms and burrowing organisms can rework these outer shelf sediments into palimpsest sediments. The inner shelf is covered with Holocene sand and mud in thicknesses that vary depending upon sediment flux from the adjacent continent. The shelf also supports an abundant and diverse benthic community that includes both epifauna and infauna. The numerous grazers, burrowers, and filter feeders rework sediment and extract it from suspension. Fecal pellets are a significant part of continental sediments. *See* HOLOCENE; MARINE SEDIMENTS.

Continental slope and rise. The outer continental margin includes the relatively steep slope and the rise that accumulates at the base of the slope (Fig. 1). This continental material has the same general composition as the landmass. Although the slope is steep relative to the shelf and rise, it has an average gradient of only about 4.5° but the range is from less than 2° to more that 40°. The very gentle gradients are areas of large deltas or fans, and the steep ones are associated with faulted margins. The gradient of the continental rise is generally less than 1°. The slope is dissected by numerous deep and steep-sided submarine canyons. These canyons are conduits for sediment transported in various sediment gravity modes to the continental rise. Some have continuations across the shelf where former rivers were located during low stands of sea level. These canyons may also extend across at least part of the rise. During low stands of sea level the rise received great quantities of sediment in rapid rates of accumulation as rivers dumped their loads near the shelf-slope break, whereas sediment supply is now relatively slow. *See* FAULT AND FAULT STRUCTURES; SUBMARINE CANYON.

The sediments carried by sediment gravity processes have a wide range in grain size, from mud to gravel. They accumulated in strata with a wide range of thickness, partly related to their grain size and partly to the volume of sediment introduced. Most of these units display graded bedding, a decreasing-upward grain-size trend.

In addition to sediment gravity processes supplying sediment to the slope and particularly to the continental rise, there are two other sources of sediments for these areas. Contour currents that flow parallel to bathymetric contours carry and deposit sediment along the outer continental margin. These contourites may be hundreds of meters thick. The other sediment supplied to this region is in the form of the tests of planktonic organisms that fall through the water column and accumulate slowly. They include both calcium carbonate (calcite and aragonite) from foraminifera and coccolithophores, and silicon dioxide (opal) from diatoms and radiolaria.

Continental borderland. The leading-edge continental margin that is commonly associated with a crustal

plate boundary displays a very different geology, geomorphology, and bathymetry than the outer continental margin. In this type there is no distinct shelf, slope, and rise. Like the trailing-edge margin, the leading-edge margin exhibits the same characteristics as the adjacent landmass, in this case a structurally complex geology with numerous fault basins and high relief. The borderland is narrow and overall steep. Its geomorphology consists of numerous local basins that receive sediment through numerous submarine canyons. The canyons commonly extend nearly to the beach; there is no shelf as such. Because much of the sediment is trapped in these basins, there is only a modest rise accumulation adjacent to the ocean floor. The best example of this type of continental margin is adjacent to southern California (**Fig. 2**), where dozens of fault-bounded basins exist. These basins range from being essentially filled with sediment to having considerable accommodation space.

Resources. The continental margin contains a vast amount and array of natural resources, most of which are being harvested. The primary fishing grounds around the globe are in shelf waters. The Grand Banks off northeastern North America and the North Sea adjacent to Europe are among the most heavily fished. There are also many mineral resources that are taken from shelf sediments, including heavy minerals that are sources of titanium, phosphate, and even placer gold. Important commodities such as sand, gravel, and shell are also taken in large quantities from the inner shelf. Salt domes that underlie the shelf, especially in the northern Gulf of Mexico, provide salt and sulfur.

Probably the most important resource obtained from the continental margin is petroleum, in the form of both oil and gas. Production is extremely high in some places, ranging from the deltas at the coast across the entire shelf and onto the outer margin, and reserves are high. *See* MARINE GEOLOGY; OIL AND GAS, OFFSHORE. Richard A. Davis, Jr.

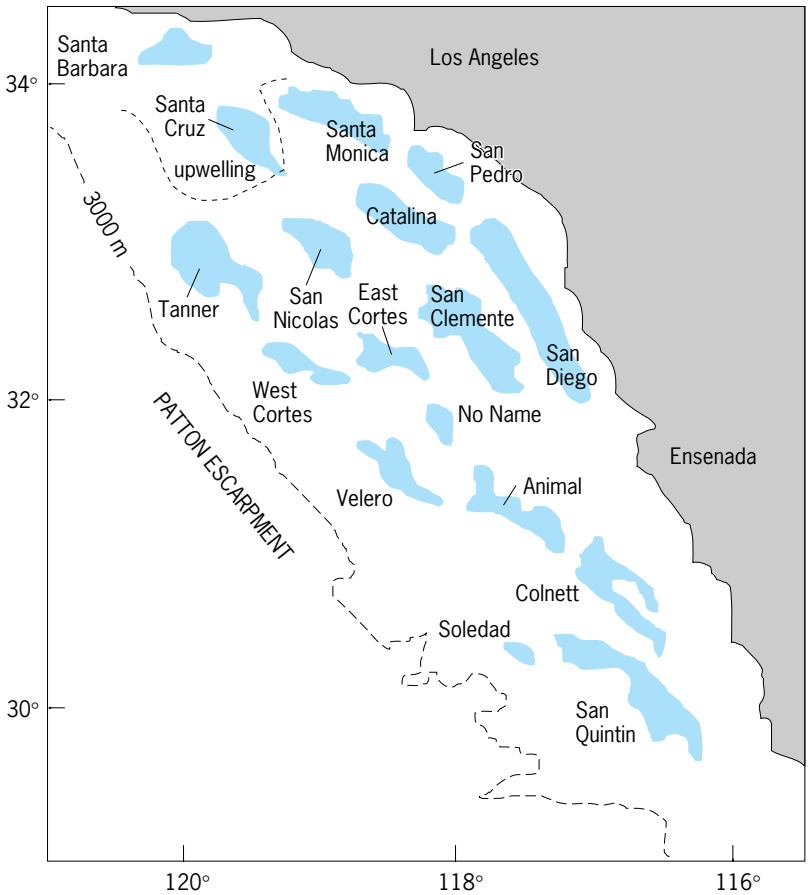

Fig. 2. California continental borderland, showing the outlines of the major basins. Upwelling is centered in the area south of the northern Channel Islands and extends south to 33°N latitude. (*After D. S. Gorsline, J. Sediment. Petrol., 48:1056, 1978*)

Shelf circulation. Shelf circulation is the pattern of flow over continental shelves. An important part of this pattern is any exchange of water with the deep ocean across the shelf-break and with estuaries or marginal seas at the coast. The circulation transports and distributes materials dissolved or suspended in the water, such as nutrients for marine life, freshwater and fine sediments originating in rivers, and domestic and industrial waste.

Water movements over continental shelves include tidal motions, wind-driven currents, and long-term mean circulation. Tidal currents reverse every half tidal cycle (about 6 h for the usually dominant semidiurnal tide), and water particles undergo back-and-forth excursions due to tides amounting to a few miles. Storms generate currents which persist typically for several days and lead to particle excursions of tens of kilometers. A pattern of long-term mean circulation is characteristic of each major subdivision of the continental shelves. Off the North American east coast, for example, the region of the shelf between the Grand Banks of Newfoundland and Cape Hatteras is characterized by a persistent southwestward drift, always present over the outer portions of the shelf (in water at least 150 ft or 50 m deep). The mean shelf circulation in this area is one leg of a larger oceanic gyre lying offshore (**Fig. 3**). Water particles are eventually swept off the shelf and into a deepwater current adjacent to the Gulf Stream. The coupling of shelf circulation to a more massive deep-ocean boundary current is tighter in some other locations, for example in the South Atlantic Bight, where the Gulf Stream controls outer shelf flow. Along the west coast of North America, the seasonally changing upwelling-downwelling regime is coupled to the weaker boundary currents along the eastern margin of the Pacific Ocean. *See* GULF STREAM.

Important for the exchange of water between a shelf region and the deep ocean are differences in motion between surface and bottom layers. At the ocean surface, the wind is a dominant influence, while at depths greater than about 30 ft (10 m) internal forces determine the intensity and direction of motion. In the east coast example discussed above, the surface layers move significantly offshore over the long term, a consequence of westerly winds, in addition to their general southwestward drift. Over shorter periods they move in whatever direction the wind carries them. The bottom layers move shoreward, to replace the water driven offshore by the wind, transporting saltier, nutrient-rich water from the deep ocean onto the shelf.

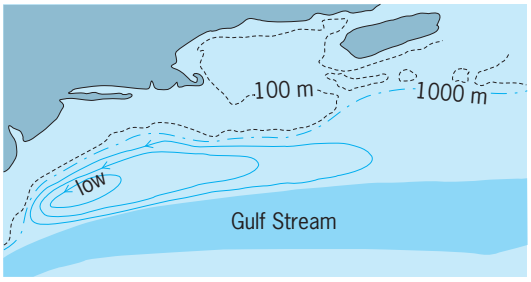

Fig. 3. Deepwater gyre between the Gulf Stream and the North American continent, influencing continental shelf circulation. Georges Bank is at the center of the illustration, separating the Gulf of Maine from deep water, and is subject to intense tidal flow. 1 m = 3.3 ft.

Tidal currents are mainly important for shelf circulation on account of the mixing they cause. While tidal currents are typically 0.6 ft/s (0.2 m/s) in amplitude, they are much stronger in some locations where shallow depth constricts the flow, for example, over Georges Bank where tidal currents are up to 3 ft/s (1 m/s). Currents of this intensity cause strong bottom friction, turbulence, and mixing. Tides on continental shelves are not caused directly by the pull of the Moon and Sun, but propagate in from the deep ocean, where the gravitational attraction of the heavenly bodies acts over a much larger water mass.

The inflow of fresh water from land also contributes to shelf circulation, because such water would tend to spread out on the surface on account of its low density. Rapid nearshore mixing reduces the density contrast, and the Earth's rotation deflects the offshore flow into a shore-parallel direction, leaving the coast to the right. A compensating shoreward flow at depth is deflected in the opposite direction, adding to the complexity of shelf circulation. *See* NEARSHORE PROCESSES; OCEAN CIRCULATION.

G. T. Csanady

Bibliography. J. S. Allen et al., Physical oceanography of continental shelves, *Rev. Geophys. Space Phys.*, 21:1149–1181, 1983; L. P. Atkinson, D. W. Menzel, and K. A. Bush (eds.) *Oceanography of the Southeastern U.S. Continental Shelf*, 1985; C. Burk and C. Drake, *The Geology of Continental Margins*, 1974; J. Kennett, *Marine Geology*, 1982; M. R. Landry and B. M. Hickey, *Coastal Oceanography of Washington and Oregon*, 1989; E. Seibold and W. H. Berger, *The Sea Floor: An Introduction to Marine Geology*, 3d ed., 1996; F. Shepard, *Submarine Geology*, 3d ed., 1973; J. S. Watkins, F. Zhiqiang, and K. J. McMillen (eds.), *Geology and Geophysics of Continental Margins*, 1992.

Continents, evolution of

The process that led to the formation of the continents. The Earth's crust is distinctively bimodal in thickness. Oceanic crust is normally about 4 mi (7 km) thick, varying mainly with the temperature of the mantle beneath the sea-floor spreading ridges when the crust was formed. In contrast, the typical 22–25-mi (35–40-km) thickness of continental crust is controlled ultimately—through the agents of erosion, sedimentation, and isostatic adjustment—by sea level. Oceanic crust is formed at spreading ridges, continental crust at subduction zones. Both are recycled to the mantle, but oceanic crust, being less buoyant, is recycled about 30 times faster than continental crust. Consequently, continents, having a mean age of almost 2 billion years and a maximum age of 4 billion years, provide the only directly accessible record spanning most of Earth history. They are, however, structurally more complex than ocean basins because of their great antiquity and weak rheology. *See* EARTH CRUST; MID-OCEANIC RIDGE.

Constructive processes. Subduction zones are the main factories for making continental crust (**Fig. 1**). The primary constructive processes are trench accretion, arc magmatism, and arc-continent collision. Mantle plumes and lithospheric stretching cause secondary magmatic additions to continental crust. *See* LITHOSPHERE; MAGMA; SUBDUCTION ZONES.

Trench accretion. Where an oceanic plate sinks beneath a continental plate or another oceanic plate, sediment scraped off the top of the descending plate accumulates as an accretionary prism at the leading edge of the overriding plate (Fig. 1). Accretionary prisms are bounded by trenches on one side and fore-arc basins on the other. They are best developed at trenches containing thick accumulations of terrigenous sediment, where the off-scraping process is facilitated by high fluid pore pressures in the sediment. Accretionary prisms may also contain slices of pelagic ooze and underlying oceanic crust, volcanic seamounts and their capping atoll reefs, or far-traveled deep-sea fans. They are typically dismembered: accretion does not proceed smoothly because the top of the descending plate is subject to block faulting related to the bending of the plate. The accreted material is dewatered, deformed, and metamorphosed as the accretionary prism grows. Accretionary prisms are rarely active magmatically. Where they intersect spreading ridges, the descending oceanic crust may be hot enough to partially melt at relatively shallow depths. A distinctive suite of sodic granitoid intrusive rocks record such events. Trench accretion is dominantly a process of crustal reworking, not crustal growth, because the bulk of the sediment is derived from the erosion of preexisting crust. *See* ATOLL; REEF.

Arc magmatism. A volcanic arc is a surface manifestation of partial melting near the tip of the mantle wedge above the subducting slab (Fig. 1). Melting in the wedge is induced by the infiltration of aqueous fluids, which lower the temperature required for melting to begin (hydration melting). The fluids are derived from the breakdown of hydrous minerals associated with the subducting oceanic crust. Depth-dependent release of water controls the location of the volcanic front facing the trench, typically 60–70 mi (100–120 km) vertically above the slab. A fertile source of melting in the mantle wedge is maintained by convective flow driven by the viscous drag of the cold descending slab. Although the main source of arc magmas is the mantle wedge, sediment and ablated crust entrained in the slab make subordinate contributions to arc magmas. The

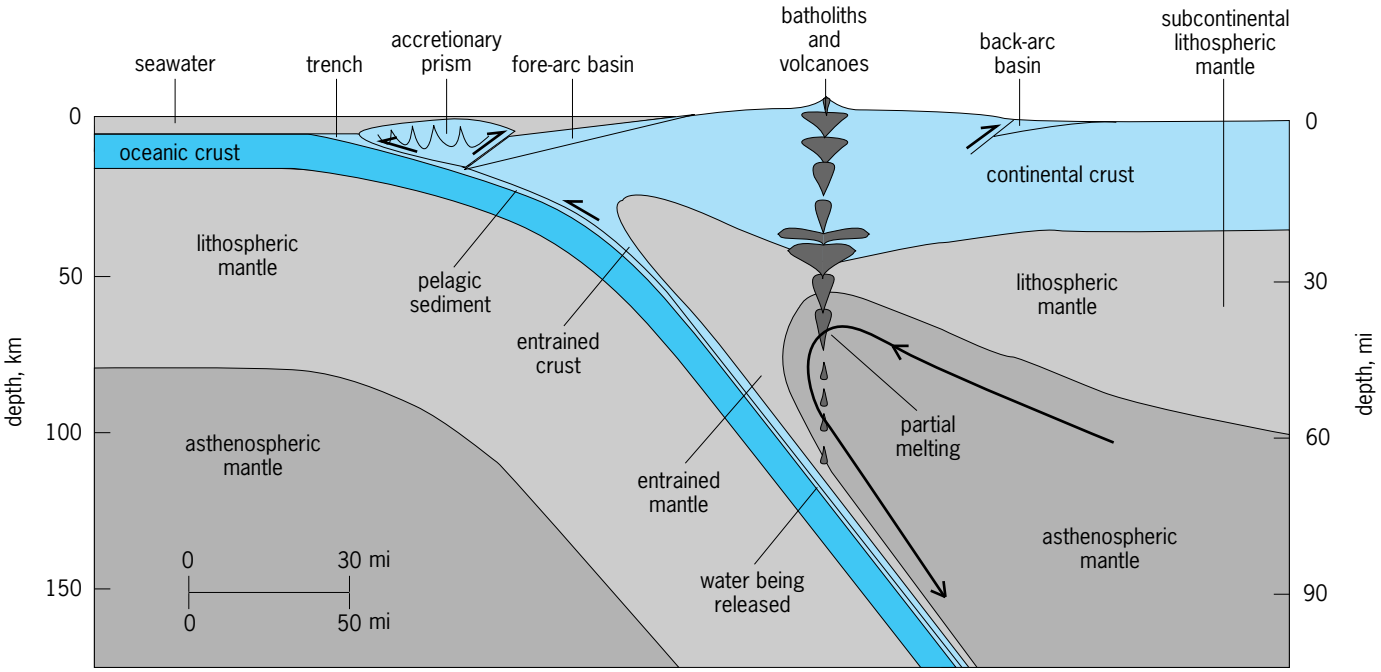

Fig. 1. Recycling of continental crust at a subduction zone. Arrows show direction of movement.

mantle-derived fraction represents new crustal addition. The slab-derived melts serve to speed up (fast-track) the mantle recycling process.

Unlike mid-ocean ridges, magmatic arcs evolve with time because the supply of magma is not balanced by crustal spreading. The crust of volcanic arcs thickens through the addition of new magma. Consequently, the compositions of the magmas evolve as the height of the mantle melting column decreases and the filtering effect of the overlying crust increases. The filtering effect causes crustal differentiation with more buoyantly mobile components concentrating upward. Denser and more refractory components collect in the lower crust. The magmas ingest and assimilate wall rocks and interact with meteoric ground waters, producing hybrid, compositionally diverse, igneous rock suites. Moreover, magmatic arcs are usually the rheologically weakest part of the lithosphere in subduction zones because of their high heat flow. Consequently, they localize tectonic deformation, which generally combines wrenching parallel to the trench (driven by oblique subduction) and extension or compression across the arc (governed by changes in buoyancy of the subducting slab or motion of the upper plate). *See* IGNEOUS ROCKS.

Collisions. Where subduction occurs beneath continental lithosphere, trench accretion and arc magmatism add crust to the continent directly. Continental-type protocrust is also formed at subduction zones (often having complex developmental histories) situated entirely within oceanic lithosphere. Incorporation of such protocrust in a continent is accomplished by arc-continent collision (**Fig. 2**). The reversal of subduction polarity that typically follows arc-continent collision (best exemplified in the modern world by the diachronous collision—a collision progressing from north to south with time—at Taiwan between the Luzon arc of the Philippines and southeast China) is the principal means by which subduction beneath a continent is initiated.

Final closure of an ocean basin results in continent-continent collision, a mainly destructive event. Tectonic crustal thickening produces mountains and greatly accelerated rates of erosion. Much of the resulting sediment is washed off the continent onto deep-sea fans or trenches, and some fraction of that sediment will be subducted. Small crustal gains may be made if the tectonically thickened lithospheric mantle becomes unstable and sinks away. The hotter replacement mantle may partially melt as it rises to take its place (decompression melting), causing small magmatic additions to the already thickened continental crust. *See* BASIN; OROGENY.

Plumes and rifts. Plumes are jets of anomalously hot mantle that partially melts as it reaches the lithosphere, causing the volcanism of Hawaiian-type islands and related seamount chains. Plumes also cause volcanism on continents, for example, Yellowstone in North America. Starting plumes have large heads and relatively small tails through which the plume is maintained. The arrival of starting plume heads at the lithosphere is signaled by continental flood basalt or oceanic plateau volcanism. The largest volumes of melt are generated where lithospheric stretching (upper crustal rifting) occurs in conjunction with a plume. The combination maximizes the potential for decompression melting. In addition to surface volcanism, the melts may pond at or near the base of the crust, causing magmatic underplating. Seamounts and oceanic plateaus are fragmentarily accreted to continents at trenches. *See* BASALT; OCEANIC ISLANDS; SEAMOUNT AND GUYOT; VOLCANOLOGY.

Destructive processes. Subduction zones are sites of continental accretion, but they may also cause

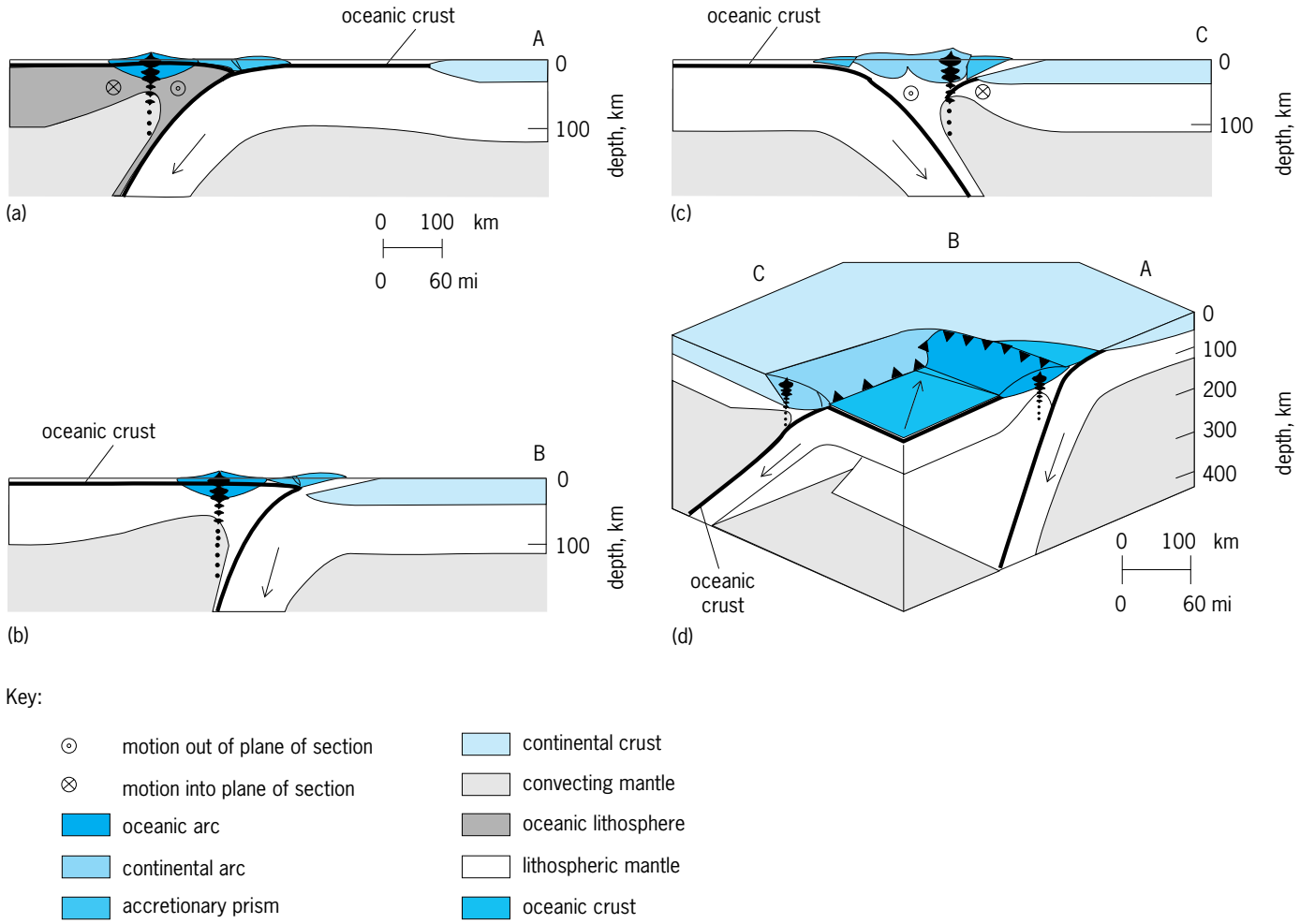

Fig. 2. Stages of an arc-continent collision; arrows give direction of lithospheric movement. (*a*) Passive continental margin converging with an oceanic arc. (*b*) Arc colliding with continent. (*c*) Subduction polarity reversed and continental arc established. (*d*) General diachrony of the collision.

destruction of continental crust (Fig. 1). The destructive processes are sediment subduction and subduction ablation. Constructive and destructive processes may be at work simultaneously, and the net balance may swing one way or the other with time. Selective destruction of lower crust may occur in continent-continent collision zones as a result of convective dripping of tectonically thickened lithosphere.

Sediment subduction. Accretionary prisms are poorly developed or lacking where trenches receive only very sparse amounts of terrigenous sediment. In such cases, pelagic oozes are subducted along with the oceanic crust. Even at subduction zones with well-developed accretionary prisms, some of the pelagic and terrigenous sediments disappears beneath the deformation front of the prism. Some is accreted structurally to the base of the prism, and some is transferred to the magmatic arc by melting. Some escapes melting and sinks deeply into the mantle, where it constitutes an isotopically recognizable component in the source of plume basalts. Sediment subduction, being fed by surface erosion, preferentially destroys the upper crust.

Subduction ablation. It is postulated that crust and lithospheric mantle from the overriding plate may become entrained in the subducting slab, causing tectonic ablation. Like subducted sediment, ablated crust is potentially capable of melting as it passes beneath the convecting mantle wedge. Unlike subducted sediment, ablated material may include lower crust and lithospheric mantle as well as upper crust. The ablation flux is quantitatively unknown.

Continental drip. Where continental lithosphere is tectonically thickened, cold lithospheric mantle is forced downward into hotter convecting mantle. Lithospheric thickening generates lateral thermal gradients, which drive mantle convection. The thickened lithospheric mantle is therefore unstable and may drip away. The missing lithospheric mantle is replaced by hotter asthenospheric mantle, causing melting and surface uplift. Lower crust may be entrained in the drip if it has been transformed to eclogite (a high-density rock chemically equivalent to basalt but composed of garnet and pyroxene) under the load of the tectonically thickened crust. Unlike sediment subduction, upper crust is not lost by dripping; and unlike subduction ablation, continental drips are not susceptible to hydration melting. Drips may sink deeply into the mantle; this view is based on the recognition of distinctive chemical and isotopic signatures of subcontinental lithospheric mantle in plume basalts. *See* ASTHENOSPHERE; ECLOGITE.

Continents and Earth history. The oldest surviving crust on Earth was formed 4 billion years ago. The absence of crust older than 4 billion years is attributed to recycling associated with the early high-impact flux inferred from the dated cratered surface of the Moon. Since the impact flux waned, processes that construct and destroy continental crust have adjusted to slow cooling and convective slowing of the mantle. The estimated magnitude of the secular change in mean upper-mantle temperature since 4 billion years ago is on the order of 360°F (200°C). This is comparable in magnitude to regional temperature variations in the upper mantle at the present time. In effect, mantle plumes in the present simulate more general conditions that prevailed in the early Earth. Furthermore, processes limited to the earlier or later parts of Earth history should be only those that operate in anomalous situations. This reasoning is consistent with empirical comparisons of continental crust of different ages. The basic forms and scales of continental accretion as well as those of continental fragmentation and collision appear to have changed remarkably little over geologic time. *See* MOON.

Oldest crustal relics. Areas of crust between 3.5 and 4 billion years old are small (less than 390 mi^2 or 1000 km^2), and there is nothing extraordinary about the rock suites they contain. Their isotopic characteristics and variability are striking. They show that the contemporaneous mantle was very heterogeneous, containing different sectors strongly depleted in some crustal components and strongly enriched in others. The crust was also heterogeneous and included sectors having long crustal residence times (greater than 200 million years) as well as newly formed sectors. These observations suggest that abundant continental-type crust already existed and that it was subject to vigorous recycling. However, these inferences are widely disputed. *See* ROCK AGE DETERMINATION.

Early juvenile crust. Juvenile crust, in a strict temporal sense, is continental crust recently separated from the mantle. In the broader nontemporal sense used here, it is continental crust not strongly modified by postaccretionary tectonism. Coherent areas up to 1550 mi (2500 km) in diameter of juvenile crust older than 2 billion years are exposed in Precambrian shields. Although rocks and structures typical of accretionary prisms, magmatic arcs, and fore-arc and back-arc basins abound, some systematic differences in comparison with geological younger juvenile crust are evident. The relative abundance of certain sodic granitoid suites in old shields implies a secular decline in the importance of slab (crustal) melting relative to wedge (mantle) melting in subduction zones. Komatiites (extremely hot, magnesium-rich, lava flows) are essentially confined to old shields and are interpreted as representing mantle-plume volcanism in the early Earth. Conversely, blueschists (rocks metamorphosed at high pressure and low temperature) are unknown in old shields. They are products of extremely rapid burial and exhumation in subduction zones having severely depressed geothermal gradients. Blueschists could not have survived even slightly increased geothermal gradients. These and other observations are consistent with the estimated cooling of the Earth and consequent adjustments in the tectonic regime. *See* BLUESCHIST; PRECAMBRIAN.

Cratons and mantle roots. Areas of juvenile crust older than about 2 billion years tend to be rheologically stiffer than younger areas. Such strong blocks are called cratons. Most have survived several rounds of continental fragmentation and collision, their margins battered by repeated openings and closing of ocean basins. Mantle tomography (three-dimensional mantle thermal structure inferred from velocity variations of earthquake-generated seismic waves) shows that cratons are underlain by anomalously cold mantle corresponding to lithosphere more than twice the approximately 60-mi (100-km) thickness of mature oceanic or noncratonic continental lithosphere. The mantle roots must be chemically distinct because, given that cratons tend to have neutral or even positive buoyancy, a low intrinsic density is required to offset the density increase due to thermal contraction. *See* COMPUTERIZED TOMOGRAPHY; CONTINENTAL MARGIN.

Conveniently, rock samples of mantle roots (peridotites) have been delivered to the surface as inclusions in diamond-bearing subvolcanic pipes. The inclusions have appropriate densities and are composed of residual mantle material from which high degrees of partial melt were driven. A class of associated inclusions, called eclogites, consists of rocks corresponding compositionally to oceanic crust but metamorphosed under mantle depth conditions. Isotopic studies reveal that the eclogite inclusions are samples of ancient oceanic crust subducted billions of years ago to depths exceeding 90 mi [150 km (into the diamond stability field)]. There they remained, fixed within drifting lithospheric plates, until fired volcanically to the surface in relatively recent times. If the peridotite and eclogite inclusions are cogenetic, the former should also have originated at seafloor spreading ridges billions of years ago. *See* DIAMOND; PERIDOTITE.

When the Earth was hotter, greater degrees of melting at oceanic spreading ridges (comparable to modern spreading ridges situated above mantle plumes, for example, Iceland) would have generated oceanic lithosphere characterized by thick (greater than 12 mi or 20 km) crust and highly depleted mantle. Such lithosphere might not sink deeply upon subduction because of the buoyancy of the thick crust and because the removal of melt actually lowers the density of the residual mantle. Accordingly, it is postulated that cratonic mantle roots are composed of buoyantly subducted slabs of highly depleted oceanic lithosphere. The presence of mantle roots would have increased the survivability of cratons: they would have been less susceptible to tectonic thickening and destruction by sediment subduction. Therefore, noncratonic continents may have existed on the young Earth but were selectively destroyed.

Continental growth. The mean age of extant continental crust is about 2 billion years. The primary mode in crustal abundance by geologic age is at 2.7 billion years, and perhaps half of all extant crust formed between 1.5 and 3 billion years ago. The conventional interpretation is that there was little continental crust following the period of high impact flux and that continental growth was slow at first, then rose to a peak 2–3 billion years ago, after which it slowly tapered off. An alternative interpretation, however, holds that the volume of continental crust has been in a near steady state since the impact flux waned. The present age distribution is explained by assuming a secular decline in the rate of crustal recycling, presumably modulated by the decreasing vigor of mantle convection as the Earth cooled. The difference in interpretation hinges on the importance assigned to recycling of continental crust. In crustal growth models, there is little crust older than 3 billion years, because little was formed. In steady-state models, much crust was formed early on but little of it survives. The steady-state interpretation is consistent with isotopic data showing that the oldest crustal relics contain highly evolved as well as juvenile components, and that the contemporaneous mantle was also heterogeneous and included strongly depleted regions. Furthermore, the near steady-state alternative is supported by comparative planetology, which indicates that for the hot young mantle not to be thoroughly differentiated by near-surface melting is physically implausible. *See* CONTINENTAL DRIFT; EARTH; EARTH, AGE OF; EARTH, CONVECTION IN; FAULT AND FAULT STRUCTURES; GEOPHYSICS; MARINE GEOLOGY; PLATE TECTONICS; SEISMOLOGY. Paul Hoffman

Bibliography. R. L. Armstrong, The persistent myth of crustal growth, *Austral. J. Earth Sci.*, 38:613–630, 1991; M. Barazangi and L. Brown (eds.), *Reflection Seismology: The Continental Crust*, Geodynamics Series, vol. 14, American Geophysical Union, 1986; E. M. Moores (ed.), *Shaping the Earth: Tectonics of Continents and Oceans* (readings from *Scientific American*), 1990; Special lithosphere issue, *J. Petrol.*, 1988.

Continuous-wave radar

A radar in which the transmitter output is uninterrupted, in contrast to pulse radar, where the output consists of short pulses. Among the advantages of continuous-wave (CW) radar is its ability to measure velocity with extreme accuracy by means of the Doppler shift in the frequency of the echo. The detected, reflected wave is shifted in frequency by an amount which is a function of the relative velocity between the target and the transmitter-receiver. Range data are extracted from the change in Doppler frequency with time. *See* DOPPLER RADAR.

In order to measure the range of targets, some form of frequency modulation (FM) of the continuous-wave output must be used. In one very effective form of modulation, the carrier frequency of the transmitted signal is varied at a uniform rate. Range is determined by comparing the frequency of the echo with that of the transmitter, the difference being proportional to the range of the target that produced the echo. Systems in which this is done are known as FM-CW radars. *See* FREQUENCY MODULATION.

A modified form of FM-CW radar employs long, but not continuous, transmission. This might be regarded as the same as transmitting extremely long pulses on an FM carrier. Systems of this type are referred to as pulse compression radars.

An FM-CW radar used in the radio-proximity fuses of missiles generates functions of the Doppler signal envelope that have identifiable characteristics at certain ranges. The amplitude of the received FM-CW signal is used to determine when the missile has attained the proper altitude. Then the generated functions tell when the missile has arrived at the desired point of burst. Three triangular-wave modulators can be used to enhance the accuracy of an FM-CW radar fusing system.

In advanced short-range air defense radar (ASHORAD) a coded continuous-wave signal reflected from the target is fed to a tapped delay line, where a predetermined delay between adjacent taps is used to establish the range.

Design objectives. Design objectives for continuous-wave radar include protecting the receiver from the transmitter output and close-by return echoes when a single antenna is used; resolving side-lobe ambiguity; resolving range ambiguity; distinguishing between approaching and receding targets; eliminating noise and clutter; simultaneously measuring both target range and velocity; determining the shapes of targets; handling multiple targets; measuring target acceleration; communicating with targets; and increasing the received signal-to-noise ratio.

Receiver protection. One disadvantage of continuous—wave radar is that when a single antenna is used for both reception and transmission it is difficult to protect the receiver against the transmitter because, in contrast to pulse radar, both are on all the time. Use of isolation circuitry gives the receiver protection from a transmitter output up to 200 W. Use of magnetically biased yttrium-iron-garnet (YIG) provides three tuned, tandem-connected power-limiter stages in one X-band FM-CW radar. *See* FERRIMAGNETIC GARNETS.

In continuous-wave surveillance radars where wide bandwidth is required to meet range and range-rate measurement requirements despite clutter and other interfering environments, the transmitter output can be coded by pseudorandom binary signals that vary with range. This is an ideal waveform for permitting range measurements, retaining a capability for range-rate measurements, and rejecting the transmitter output and other nearby echoes.

Side-lobe ambiguity. The problem of distinguishing between echoes from the main lobe and those from the side lobes has been addressed by providing a demodulated range response that has one polarity for the main-lobe echoes and the opposite polarity for side-lobe echoes.

Side-lobe ambiguity is removed in an FM-CW radar fusing system in which the reflected signal is compared with a sample of the continuous-wave output in a pair of FM discriminators displaced from each other in the frequency domain. Thresholding circuits determine in which half of which discriminator the difference frequency signal is concentrated.

Side-lobe ambiguity in the range response of an FM-CW radar can be resolved when a periodic modulation frequency is reflected from a target and combined with a sample of the transmitted signal. Selected harmonics of the combination signal are processed to develop a main-lobe response of one polarity with all side lobes of the opposite polarity.

Range ambiguity. This ambiguity results from reception of echoes from targets beyond the range of interest and of second-time-around echoes. It can be resolved by range gates that make the radar insensitive to targets beyond the range of interest or by filters that put limits on range. Resolution of range ambiguity is very important in radars for personnel surveillance.

Directional ambiguity. To distinguish between approaching and receding targets, a continuous wave beamed by a radar is sampled continuously but delivers an output voltage only when the distance to the target is changing significantly. The voltage is positive if the target is closing, and negative if it is receding.

Noise and clutter. Suppression of noise, clutter, chaff, and jamming has received a great deal of attention. In one approach, a fixed pair of amplitude-modulation (AM) sidebands is added and used to null out the noise.

Echo signal analyzers can be selectively adjusted for a certain moving-target size so that small targets are not detected. A series of digital range cells can be defined such that each cell represents the minimum allowable distance between two reflective points on a target. The radar also can be adaptively adjusted to attend to a desired average signal level.

Use of digitally controlled voltage-tunable radio-frequency generators permits frequency jumping to avoid jamming or spoofing. Pseudorandom-noise-modulated continuous-wave radar can be used to establish perimeter surveillance in a heavily foliaged environment. Pseudorandomly coded continuous-wave radar can also be used to detect targets in the presence of clutter. *See* ELECTRICAL NOISE; ELECTRONIC WARFARE.

Range and velocity sensing. Several techniques are used to obtain simultaneous measurements of range and velocity. One involves transmitting sidebands of linearly increasing and decreasing frequency or using triangular modulation, and comparing the difference between transmitted and received frequencies with the difference between the received frequency and the transmitted frequency delayed by a period of time proportional to the range of the target. Phase-shift keying by a pseudorandom code also permits simultaneous measurement of range and velocity. *See* MODULATION.

A double-sideband linear FM radar obtains simultaneous range and Doppler measurements. Waveforms of both increasing and decreasing frequency are transmitted simultaneously and combined coherently to give an output that is a measure of the target's radial velocity. The received signals are processed separately to obtain range information.

Target shape. Millimeter-wave FM-CW radar can be used to determine the shape of a target. These radars operate at frequencies in excess of 90 GHz. They often use Gunn oscillators as the source. *See* MICROWAVE SOLID-STATE DEVICES.

French and Russian research suggests that in some applications it may be necessary to use space diversity in which two beams illuminate the target in order to compensate for nonlinearities in frequency scanning.

The dependence of radar cross section on transmitted frequency over a wide bandwidth can be used to make the signal amplitude scintillate so that target characteristics can be determined.

Rapid discontinuous fluctuations in the scattering from a moving multielement target is known as the intermittent-contact RADAM (radar detection of agitated metals) effect. This effect is observed in the contact between the drive sprocket and the tread of a tank and can be used to define signatures of various kinds of fighting vehicles.

Multiple targets. Discrete Fourier analysis of the video heterodyne (beat) signals using a digital computer permits multiple-target handling and achieves good resolution between approaching and receding targets. *See* SPECTRUM ANALYZER.

Target acceleration. An indication of the acceleration of a target can be obtained by double subtraction and delay of the returned signal.

Communicating with the target. In radar missile guidance, homing beacons, and identification–friend-or-foe, it is necessary to communicate with the target. This can be done by interrupting the range beat frequency at short intervals near its peak excursion and transmitting AM signals to the illuminated target.

Signal-to-noise ratio. The signal-to-noise ratio can be improved by digitizing the beat frequency signals. Moreover, compressing the received signals to a sine-squared shape can increase the signal level more than 20 decibels. *See* SIGNAL-TO-NOISE RATIO.

Applications. Applications of continuous-wave radar include missile guidance; detection of hostile targets; terrain clearance indication and ground surveillance; laser radar systems; atmospheric studies; automobile safety; surveillance of personnel; ice studies; remote sensing; and reproduction of the shape of a patient's pulse.

Missile guidance. The principal use of continuous-wave radar is in short-range missile guidance. Typically the missile's course is tracked from the ground while the missile is simultaneously illuminated. Continuous-wave radar, in some cases the same radar, can be used for both tracking and illumination, although it is more common for pulse Doppler radar to be used for tracking.

Advances in the design of phased-array antennas have led to pulse Doppler radar becoming more attractive than continuous-wave radar even for illumination. However, at least one system allows the

operator to select either continuous-wave or pulse Doppler radar for illumination.

Illumination radar is for target acquisition. Two signals are of interest. One, the directly received illumination of the missile, is called the rear signal. The other, the signal reflected from the target, is called the front signal. The front signal is shifted in frequency as the missile closes with the target. This shift occurs because of the Doppler effect. The shift, and therefore the range to the target, is obtained by coherently detecting the front signal against the rear signal. In an active guidance system, detection is performed on board the missile. Semiactive missile guidance is more frequently used, and in this system the signals are relayed over a data link for processing at the ground station.

A late version of the MPQ-34 tracking and acquisition radar for the Hawk ground-to-air missile incorporates the choice of either continuous-wave or pulse Doppler radar illumination. The ability of continuous-wave radar to discriminate against clutter made it attractive in this system, which was intended to be used against low-flying aircraft.

Pulsed tracking radar with continuous-wave acquisition radar is used with the Sparrow air-to-air missile system and with the Tartar shipborne missile system.

A miss-distance indicator (MDI) is a special application of continuous-wave radar. It uses radio-frequency signals to determine the relative velocity between two objects, the time of intercept when the distance will be a minimum, and the actual value of that minimum distance. With the availability of real-time intercept information, a prior determination of guidance parameters that will produce a direct hit can be made. *See* GUIDANCE SYSTEMS; GUIDED MISSILE.

Detection of hostile targets. Modulated continuous-wave radar has been used to detect hostile military vehicles and personnel.

The Navy MK-92 fire-control system uses a linear FM-CW radar with multiple resonant filters.

An FM-CW naval radar operating at 95.6 GHz uses pseudorandom-coded phase-shift keying, digital signal processing, and microprocessor-controlled azimuth and elevation scanning. It can display the shape, range, velocity, and direction of vessels. *See* NAVAL ARMAMENT.

In the United States, Canada, and the United Kingdom, there has been interest in continuous-wave radar for detecting nonmetallic mines. One radar operates between 2 and 4 GHz and can detect plastic objects buried in 10 in. (25 cm) of wet sand.

Terrain surveillance. One of the first uses of FM continuous-wave radar was in terrain clearance indicators or radar altimeters. Today intruder aircraft and air-breathing missiles obtain continuous guidance from that kind of equipment. *See* ALTIMETER.

The ability of side-looking radar on reconnaissance aircraft and Earth-observation satellites to resolve targets has been enhanced by the use of synthetic-aperture radar. Although this technique usually requires a huge data rate and great processing complexity, special-purpose processors have been used to extract useful data economically. Such a processor on the *Seasat 1* was able to determine the direction of ocean wave from 800 mi (1300 km) up. *See* REMOTE SENSING.

Earth surveillance has been hampered by some radar reflection phenomena, notably foreshadowing, which distorts distance; layover, which tends to fill in depressions; and shadow, which obscures features. These problems place limiting requirements on the size and contrast of images needed for successful interpretation. These limitations can be partially overcome by training radargrammetrists with simulators.

Laser radar systems. An experimental modulated continuous-wave laser radar was built to track low-flying airborne targets from the ground. The system can determine both range and the rate at which the range is changing. An advantage of this system is its extreme precision. The divergence of the beam is less than 1 milliradian. The transmitter is a carbon dioxide laser operating at a wavelength of 10.6 micrometers. The infrared carrier is frequency-modulated with 10-MHz excursions occurring at a frequency of 1 kHz. *See* LASER.

Atmospheric studies. The ability of FM-CW radar to measure distance and velocity despite smoke, dust, and thermal gradients makes it attractive for many industrial automatic control applications. The Naval Ocean Systems Center uses S-band continuous-wave radar to update refractive index profiles of the troposphere. This work forms the basis for radio interception and over-the-horizon targeting. In France, continuous-wave radar has been used to measure the electron temperatures in the ionosphere. The National Oceanic and Atmospheric Administration (NOAA) has used continuous-wave radar to measure wind in clear air.

Automobile safety. An experimental continuous-wave radar which exploits the Doppler principle has been developed for use in automobiles. The radar can anticipate a crash when an obstacle is 30 ft (10 m) away in order to deploy air-cushion-type passive restraints. It can sense obstacles 500 ft (150 m) away to govern automatic braking and headway control. The source of carrier power is an X-band Gunn-type solid-state oscillator.

Surveillance of personnel. A variant of continuous-wave radar is used to illuminate persons under surveillance by techniques employing semiconductor tracer diodes. These devices are secreted on or implanted in the subject's body without his or her knowledge, or else concealed in objects that are to be protected against theft. Despite the fact that pulsed X-band sources could provide the needed power levels more conveniently, the requirement to reduce clutter makes it desirable to use continuous-wave power.

These techniques are also used to sweep premises to discover clandestine listening devices and to guard protected locations against the introduction of concealed tape recorders.

Implementation of tracer-diode surveillance utilizes a single carrier frequency and looks for reflections of the third harmonic as evidence that the person sought has been acquired by the beam. The

surveillance operations have been carried out from fixed posts, vans, and low-flying aircraft.

An improved version makes use of two carrier frequencies for illumination and depends upon coherent detection of the third and fourth harmonics of the sum and difference frequencies to signify acquisition of the target. *See* ELECTRONIC LISTENING DEVICES.

Other applications. In Finland, a continuous-wave radar operating between 1 and 1.8 GHz has been used to measure ice and frost thickness in lakes and bogs. Continuous-wave radars operating at 1.5 and from 8 to 18 GHz have been used to discriminate among different types of Arctic Sea ice.

Continuous-wave radar has been used in the Netherlands to discriminate among different types of vegetation, crops, and bare soils.

With its antenna placed over a patient's wrist artery, continuous-wave radar has been shown to be able to reproduce the shape of the arterial pulse. *See* RADAR. John M. Carroll

Bibliography. P. L. Bogler, *Radar Principles with Applications to Tracking Systems*, 1990; J. L. Eaves and E. K. Reedy, *Principles of Modern Radar*, 1987; B. Edde, *Radar: Principles, Technology, Applications*, 1992; F. E. Nathanson, *Radar Design Principles*, 2d ed., 1991; M. I. Skolnik, *Radar Handbook*, 2d ed., 1990.

Contour

The locus of points of equal elevation used in topographic mapping. Contour lines represent a uniform series of elevations, the difference in elevation between adjacent lines being the contour interval of the given map. Thus, contours represent the shape of terrain on the flat map surface (see **illus.**). Closely spaced contours indicate steep ground; sparseness or absence of contours indicates gentle slope or flat ground. Contours do not cross each other unless there is an overhang. *See* TOPOGRAPHIC SURVEYING AND MAPPING.

In ground-survey procedures the horizontal positions and elevation of numerous points are observed. Points on the elevations chosen for contouring are then located by interpolation, and each contour is sketched by connecting the points found for its elevation. *See* SURVEYING.

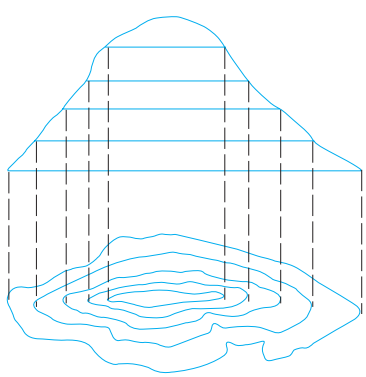

Contour representation; vertical projection indicates the profile of the elevation as it relates to the spacing of the contours.

In photogrammetric mapping with a stereoplotting instrument, the operator sets the plotting table at the chosen contour elevation; the elevation is "hunted" by shifting the table until a floating dot appears to touch the stereoscopic image; the operator then lowers the plotting pencil and traces the contour by keeping the floating dot in apparent touch with the image. *See* CARTOGRAPHY; PHOTOGRAMMETRY. Robert H. Dodds

Bibliography. J. M. Anderson et al., *Surveying: Theory and Practice*, 7th ed., 1997; A. H. Robinson et al., *Elements of Cartography*, 6th ed., 1995.

Control chart

A graphical technique for determining whether a process is or is not in a state of statistical control. Being in statistical control means that the extent of variation of the output of the process does not exceed that which is expected on the basis of the natural statistical variability of the process. Several main types of control charts are used, based on the nature of the process and on the intended use of the data.

Every process has some inherent variability due to random factors over which there is no control and which cannot be eliminated economically. For instance, in a metal fabrication process random factors may include the distribution of impurities and structural faults among the metal molecules, vibrations of the fabrication equipment, fluctuations in the power supply that affect the speed and torque of the equipment, and variations in the operator performance from one cycle to the next. The inherent variability of the process is the aggregate result of many individual causes, each having a small impact.

The control chart technique is applicable to processes that produce a stream of discrete output units. Control charts are designed to detect excessive variability due to specific assignable causes that can be corrected. Assignable causes result in relatively large variations, and they usually can be identified and economically removed. Examples of assignable causes of variations that may occur in the example of metal fabrication include a substandard batch of raw material, a machine malfunction, and an untrained or poorly motivated operator.

A control chart is a two-dimensional plot of the evolution of the process over time. The horizontal dimension represents time, with samples displayed in chronological order, such that the earliest sample taken appears on the left and each newly acquired sample is plotted to the right. The vertical dimension represents the value of the sample statistic, which might be the sample mean, range, or standard deviation in the case of measurement by variables, or in the case of measurement by attributes, the number of nonconforming units, the fraction nonconforming,

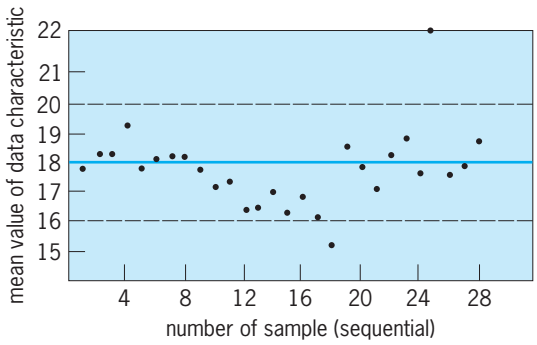

Control chart, showing changes in average of process.

the number of nonconformities, or the average number of nonconformities per unit.

Typically a control chart includes three parallel horizontal lines (see **illustration**): a center line and two control limits. The centerline (CL) intersects the vertical dimension at a value that represents the level of the process under stable conditions (natural variability only). The process level might be based on a given standard or, if no standard is available, on the current level of the process calculated as the average of an initial set of samples. The two lines above and below the centerline are called the upper control limit (UCL) and lower control limit (LCL) respectively, and they both denote the normal range of variation for the sample statistic. The control limits intersect the vertical axis such that if only the natural variability of the process is present, then the probability of a sample point falling outside the control limits and causing a false alarm is very small. Typically, control limits are located at three standard deviations from the centerline on both sides. This results in a probability of a false alarm being equal to 0.0027.

The principle of operation of control charts is rather simple and consists of five general steps:

1. Samples are drawn from the process output at regular intervals.

2. A statistic is calculated from the observed values of the units in the sample; a statistic is a mathematical function computed on the basis of the values of the observations in the sample.

3. The value of the statistic is charted over time; any points falling outside the control limits or any other nonrandom pattern of points indicate that there has been a change in the process, either its setting or its variability.

4. If such change is detected, the process is stopped and an investigation is conducted to determine the causes for the change.

5. Once the causes of the change have been ascertained and any required corrective action has been taken, the process is resumed.

The main benefit of control charts is to provide a visual means to identify conditions where the process level or variation has changed due to an assignable cause and consequently is no longer in a state of statistical control. The visual patterns that indicate either the out-of-control state or some other condition that requires attention are known as outliers, runs of points, low variability, trends, cycles, and mixtures. *See* CONTROL SYSTEMS; QUALITY CONTROL; STATISTICAL PROCESS CONTROL.

Tvi Raz

Bibliography. L. S. Aft, *Fundamentals of Industrial Quality Control*, 3d ed., CRC Press, 1997; J. P. Bentley, *Introduction to Reliability and Quality Engineering*, 2d ed., Addison-Wesley, Harlow, 1999; J. W. Cortada and J. A. Woods, *The McGraw-Hill Encyclopedia of Quality Terms and Concepts*, McGraw-Hill, New York, 1995; J. R. Evans and W. M. Lindsay, *The Management and Control of Quality*, 6th ed., South-Western College Publications, Cincinnati, 2004; V. E. Sower, M. J. Savoie, and S. Renick, *An Introduction to Quality Management and Engineering*, Prentice Hall, 1999; D. Straker, *A Toolbook for Quality Improvement and Problem Solving*, Prentice Hall, London, 1995.

Control system stability

The well-behavedness of the output of a control system in response to either external inputs or internal excitations. Several definitions of stability are in use, each being appropriate to a different application. The objective of stability theory is to obtain information concerning the behavior of a control system without actually solving the equations that describe the system.

Stability theory can be applied in a wide variety of disciplines, such as engineering, mathematics, physics, biology, and economics.

A few examples illustrate such stability theory: (1) The classical problem studied by the Soviet mathematician A. M. Lyapunov was the stability of the solar system. The objective was to determine the ultimate behavior of the trajectories of the various planets, in particular whether they would all spiral into the Sun, fly off into space, or do something else. (2) In digital filters, it is sometimes possible to encounter limit cycles, that is, round off errors that do not decay, perhaps due to the finite word length of the operations. It has been experimentally observed that certain configurations and certain coefficient values never exhibit limit cycles, and a mathematical explanation of this phenomenon would be useful. (3) A voltage source is connected to an electronic circuit containing resistors, capacitors, and operational amplifiers. For a broad but prespecified class of voltage excitations, it would be useful to know whether the resulting current will eventually decay to zero. (4) In a plasma laboratory, a scheme is proposed for confining the plasma to a specified cylindrical region; it would be useful to know whether the proposed scheme will work. (5) A model is developed for the spread of an infectious disease within a population, and it would be useful to be able to determine how the number of survivors is affected by various parameters of the model. *See* ELECTRIC FILTER; NUCLEAR FUSION.

As shown by the discussion below, stability theory encompasses systems described by a wide variety

of models. Thus the equations governing the system may be linear or nonlinear; they may be difference equations or differential equations; they may be ordinary, partial, or functional differential equations; finally, they may be difference equations in more than one parameter. This article deals chiefly with differential equations; the theory for other types of equations is similar in concept, though not in detail.

Routh, Hurwitz, and related tests. Consider a system whose transfer function is of the form given by Eq. (1), where $\hat{n}$ and $\hat{d}$ are polynomials and s denotes

$$\hat{h}(s) = \frac{\hat{n}(s)}{\hat{d}(s)} \tag{1}$$

the variable of Laplace transformation. Such a system exhibits a bounded response to every bounded excitation if and only if the following two conditions hold: (1) $\hat{h}$ is proper, that is, the degree of the polynomial $\hat{n}$ is less than or equal to that of $\hat{d}$; (2) all poles of $\hat{h}$, that is, all zeros of the polynomial $\hat{d}$, must have negative real parts. (A polynomial satisfying these conditions is called a strictly Hurwitz polynomial.) Given a transfer function $\hat{h}$, inspection will verify whether or not the first condition is satisfied, but the second condition is harder to verify. In particular, it would be useful to determine whether or not all zeros of $\hat{d}$ have negative real parts (that is, whether or not $\hat{d}$ is strictly Hurwitz) without actually computing these zeros. The various tests described below are addressed to this problem. *See* LAPLACE TRANSFORM.

Routh test. Suppose the polynomial $\hat{d}$ is of the form given by Eq. (2), where it may be supposed, with-

$$\hat{d}(s) = d_m s^m + d_{m-1}s^{m-1} + \cdots + d_1 s + d_0 \tag{2}$$

out loss of generality, that $d_m > 0$. The Routh test is based on forming the ratio of the even part of $\hat{d}$ over the odd part of $\hat{d}$ (or vice versa), and then carrying out a continued fraction expansion. To illustrate, suppose $m = 2k$ is an even number; then $\hat{d}_{\text{even}}(s)$ and $\hat{d}_{\text{odd}}(s)$ are defined by Eqs. (3) and (4). Then the ratio $\hat{d}_{\text{even}}(s)/\hat{d}_{\text{odd}}(s)$ is expressed in the form given by Eq. (5). [If m is odd, $\hat{d}_{\text{odd}}(s)/\hat{d}_{\text{even}}(s)$ this form.] The

$$\hat{d}_{\text{even}}(s) = \sum_{i=0}^{k} d_{2i}s^{2i} \tag{3}$$

$$\hat{d}_{\text{odd}}(s) = \sum_{i=0}^{k-1} d_{2i} + {}_1 s^{2i+1} \tag{4}$$

$$\frac{\hat{d}_{\text{even}}(s)}{\hat{d}_{\text{odd}}(s)} = b_m s + \cfrac{1}{b_{m-1}s + \cfrac{1}{\cdots + \cfrac{1}{b_1 s}}} \tag{5}$$

basis of the Routh test is in the following fact: $\hat{d}$ is a strictly Hurwitz polynomial if and only if (1) it is possible to completely carry out the continued fraction expansion in Eq. (5), and (2) all the constants $b_m, \ldots, b_1$ are positive. The test also furnishes one additional piece of information. Suppose that the expansion in Eq. (5) is completed, but not all the b_i's are positive. (Then of course $\hat{d}$ is not strictly Hurwitz.) Under these conditions, if none of the b_i's is zero, then the number of zeros of $\hat{d}$ with positive real part is equal to the number of negative b_i's.

Though the above is the theoretical basis of the Routh test, this is not the manner in which the test is actually carried out. To perform the test in an efficient manner, an array is set up, called the Routh array, as shown in Eq. (6), for the case that m is even,

$$\begin{array}{c|cccccc}
s^m & d_m & d_{m-2} & \cdots & d_2 & d_0 \\
s^{m-1} & d_{m-1} & d_{m-3} & \cdots & d_1 & \\
s^{m-2} & a_{m-2,1} & a_{m-2,2} & \cdots & a_{m-2,k} & \\
\cdot & & & & & \\
\cdot & & & & & \\
\cdot & & & & & \\
s^1 & a_{1,1} & & & & \\
s^0 & a_{0,1} & & & &
\end{array} \tag{6}$$

and the entries in the array are recursively defined by Eq. (7). With reference to Eq. (5), it can be shown

$$a_{i,j} = a_{i+2,j+1} - \frac{a_{i+2,1} \cdot a_{i+1,j+1}}{a_{i+1,1}} \tag{7}$$

that the b_i's are given by Eq. (8). Hence the Routh

$$b_i = a_{i,1}/a_{i-1,1} \qquad i = m, \ldots, 1 \tag{8}$$

test can be concisely stated as follows: The polynomial $\hat{d}$ is strictly Hurwitz if and only if all the entries in the first column of the Routh array are positive. Now suppose d is not strictly Hurwitz, but none of the entries in the first column of the Routh array is zero. Then the number of zeros of $\hat{d}$ with positive real part is equal to the number of sign changes in the first column of the Routh array.

Hurwitz test. The Routh test is frequently misnamed the Routh-Hurwitz test. Actually the Hurwitz test is different from the Routh test, though they are equivalent in the sense that they both give the same information. The Hurwitz test is as follows: Given a polynomial $\hat{d}$ of the form in Eq. (2), construct the $m \times m$ matrix $\mathbf{H}$ given by Eq. (9), and let $\Delta_1, \ldots, \Delta_m$ denote

$$\mathbf{H} = \begin{bmatrix}
d_{m-1} & d_{m-3} & \cdots & 0 & 0 & \cdots & 0 \\
d_m & d_{m-2} & \cdots & & & & \\
0 & d_{m-1} & d_{m-3} & \cdots & & & \\
0 & d_m & d_{m-2} & \cdots & & & \\
\cdot & & & & & & \\
\cdot & & & & & & \\
\cdot & & & & & & \\
0 & 0 & \cdots & & & d_3 & d_1 \\
0 & 0 & \cdots & & & d_2 & d_0
\end{bmatrix} \tag{9}$$

the leading principal minors of $\mathbf{H}$. (That is, Δ_i is the determinant of the matrix consisting of the entries in $\mathbf{H}$ that belong to both the first i rows and the first i columns.) Then $\hat{d}$ is a strictly Hurwitz polynomial if and only if $\Delta_i > 0$ for all i. If $\hat{d}$ is not strictly Hurwitz but none of the Δ_i's is zero, then the number of zeros

of $\hat{d}$ with positive real part is equal to the number of negative Δ_i's. *See* DETERMINANT.

The relationship between the Hurwitz determinants and the first column of the Routh array is given by Eq. (10), where $\Delta_0 = 1$.

$$a_{i,1} = \Delta_{m-i}/\Delta_{m-i-1} \qquad (10)$$
$$i = 1, \ldots, m-1$$

Liénard-Chipart test. Given a polynomial $\hat{d}$ of the form in Eq. (5), a necessary but by no means sufficient condition for $\hat{d}$ to be strictly Hurwitz is that all the coefficients d_i be positive. In this case, it is possible to somewhat simplify the Hurwitz test. In particular, if $d_i > 0$ for all i, then $\hat{d}$ is strictly Hurwitz if and only if either of the equivalent conditions (11) and (12) is satisfied. This simplification is known as the Liénard-Chipart test.

$$\Delta_1 > 0, \Delta_3 > 0, \Delta_5 > 0, \ldots \qquad (11)$$

$$\Delta_2 > 0, \Delta_4 > 0, \Delta_6 > 0, \ldots \qquad (12)$$

Tests for zeros all of magnitude less than 1. In analyzing the stability of sampled-data control systems, digital filters, and so forth, it is desirable to determine not whether all zeros of $\hat{d}$ have negative real parts, buy whether all zeros of $\hat{d}$ have magnitude less than 1. One means of doing so is by using the bilinear transformation defined by Eq. (13), which maps the interior of the unit circle in the s plane into the left half of the v plane. Thus, all zeros of $\hat{d}$ in the form (2) have magnitude less than 1 if and only if the associated polynomial given by Eq. (14) is strictly Hurwitz. This

$$s = \frac{v+1}{v-1} \qquad (13)$$

$$p(v) = \sum_{i=0}^{m} d_i (v+1)^i (v-1)^{m-i} \qquad (14)$$

can be determined using the tests described above.

However, there are many tests that can be applied directly to the polynomial $\hat{d}$ without requiring any transformation. One of these is the Schur-Cohn test: Let $\hat{d}(s)$ be a polynomial of the form in Eq. and define the determinants Δ_i by Eq. (15) Δ_i is the determinant of a $2i \times 2i$ matrix. The Schur-Cohn test states that all zeros of $\hat{d}$ have magnitude less than 1 if and only if $(-1)^i \Delta_i > 0$ for all i, that is Δ_i positive for even i and negative for odd i.

The theoretical advantage of the Schur-Cohn test is that it can be used (with slight modifications) even if the coefficients d_i are complex. However, in practical situations, the coefficients d_i are all real, in which case it is possible to simplify the Schur-Cohn test. The Jury-Blanchard test is as follows: Given a polynomial $\hat{d}$ in the form of Eq. (2), define an array of elements $a_{i,j}$, $i = 0$ to $m = 2$, $j = 0$ to $(m - i)$, by the rules of Eqs. (16) and (17). With these definitions, all zeros of

$$a_{0,j} = d_j \qquad j = 0, \ldots, m \qquad (16)$$

$$a_{i,j} = \begin{vmatrix} a_{i-1,0} & a_{i-1,m-i-j} \\ a_{i-1,m-i} & a_{i-1,j} \end{vmatrix} \qquad (17)$$

$\hat{d}$ have magnitude less than 1 if and only if conditions (18) and (19) are satisfied.

$$\hat{d}(1) > 0 \quad (-1)^m \hat{d}(-1) > 0 \quad |d_0| > |d_m| \qquad (18)$$

$$|a_{i,0}| > |a_{i,m-i}| \text{ for } i = 1 \text{ to } m-2 \qquad (19)$$

Nyquist, circle, and Popov criteria. The Nyquist criterion provides a necessary and sufficient condition for the stability of a broad class of linear, time-invariant systems. The criterion can be stated at various levels of generality, and what is given below is one of the simplest versions.

Nyquist plot. Consider a feedback system of the form shown in **Fig. 1**, where the forward element has a transfer function $\hat{g}(s)$ and the feedback contains a constant gain k. It is assumed that the transfer function consists of the sum of three terms, in the form of Eq. (20), where g_0 is a constant; $\hat{g}_a(s)$ is the Laplace

$$\hat{g}(s) = g_0 + \hat{g}_a(s) + \hat{g}_b(s) \qquad (20)$$

transform of an absolutely integrable function, that is, Eq. (21) is valid, and $\hat{g}_b(s)$ is a proper rational

$$\hat{g}_a(s) = \mathcal{L}[g_a(t)] \qquad (21)$$
$$\text{where } \int_0^\infty |g_a(t)|\, dt \text{ is finite}$$

function. Under these conditions, it is possible to determine whether or not the closed-loop system

$$\Delta_i = \left|\begin{array}{ccccc|cccc} d_0 & 0 & 0 & \cdots & 0 & d_m & d_{m-1} & \cdots & d_{m-i+1} \\ d_1 & d_0 & 0 & \cdots & 0 & 0 & d_m & \cdots & d_{m-i+2} \\ d_2 & d_1 & d_0 & \cdots & 0 & 0 & 0 & \cdots & d_{m-i+3} \\ \cdot & & & & & \cdot & & & \\ \cdot & & & & & \cdot & & & \\ \cdot & & & & & \cdot & & & \\ d_{i-1} & d_{i-2} & d_{i-3} & \cdots & d_0 & 0 & 0 & \cdots & d_m \\ \hline d_m & 0 & 0 & \cdots & 0 & d_0 & d_1 & \cdots & d_{i-1} \\ d_{m-1} & d_m & 0 & \cdots & 0 & 0 & d_0 & \cdots & d_{i-2} \\ \cdot & & & & & \cdot & & & \\ \cdot & & & & & \cdot & & & \\ \cdot & & & & & \cdot & & & \\ d_{m-i+1} & d_{m-i+2} & & \cdots & d_m & 0 & 0 & \cdots & d_0 \end{array}\right| \qquad (15)$$

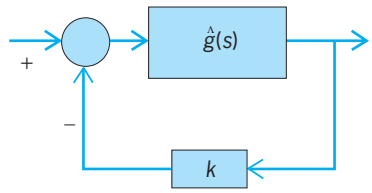

Fig. 1. Feedback system with feedback containing a constant gain *k*.

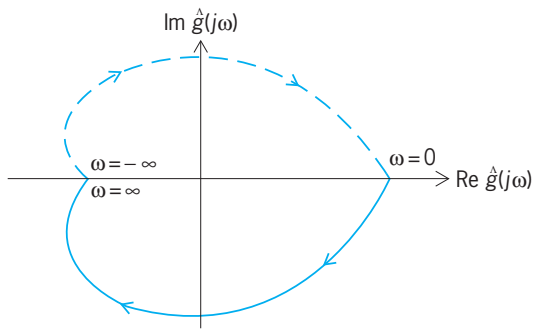

Fig. 2. Nyquist plot of a transfer function $\hat{g}$.

is stable by examining only the behavior of $\hat{g}(j\omega)$, where $j = \sqrt{-1}$ and the frequency parameter ω is allowed to vary along the real axis from $-\infty$ to ∞.

The plot of the real part of $\hat{g}(j\omega)$ versus the imaginary part of $\hat{g}(j\omega)$, as the frequency parameter ω varies from $-\infty$ to ∞, is called the Nyquist plot of $\hat{g}$ (**Fig. 2**). Since $\hat{g}(-j\omega)$ is the complex conjugate of $\hat{g}(j\omega)$, the plot as ω varies from 0 to $-\infty$ is the mirror image, about the real axis, of the plot as ω varies from 0 to ∞.

It sometimes happens that $\hat{g}$ has some poles on the imaginary axis, in which case $\hat{g}(j\omega)$ may be undefined for certain values of ω. In such cases, it is necessary to indent the $j\omega$ axis around the poles of $\hat{g}(j\omega)$, as shown in **Fig. 3**.

Nyquist criterion. With these conventions, the Nyquist criterion can be stated as follows: The system of Fig. 1 is stable if and only if (1) the Nyquist plot of $\hat{g}(j\omega)$ does not intersect the critical point $-(1/k) + j0$, and (2) as ω increases from $-\infty$ to ∞, the Nyquist plot encircles the critical point exactly μ times in the

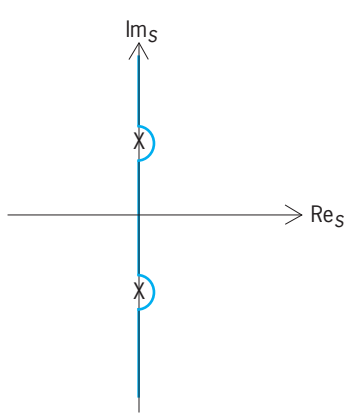

Fig. 3. Modification of Laplace transformation variable to avoid poles of the transfer function $\hat{g}$.

counterclockwise direction, where μ is the number of poles of $\hat{g}$ with positive real parts.

Several advantages of the Nyquist criterion are immediately apparent. (1) It is based on the behavior of $\hat{g}(j\omega)$, which can be readily obtained by frequency response measurements. (2) Since $\hat{g}$ is not required to be a rational function of s (only $\hat{g}_b$ is required to be rational), the criterion is applicable to systems containing transmission lines, time delays, and so forth. (3) For a given transfer function $\hat{g}$, one can readily calculate the range of values of k that result in a stable closed-loop system. (4) Even if the system of Fig. 1 is unstable, it can easily be determined what sort of "compensation" is needed to make the system stable.

Circle criterion. Now consider the system of **Fig. 4**, which is similar to that of Fig. 1, except that the constant gain k has been replaced by a time-varying gain $k(t)$. The assumptions on $\hat{g}(s)$ are the same as before.

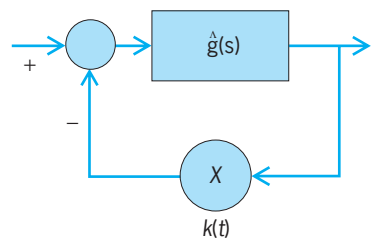

Fig. 4. Feedback system with feedback containing a time-varying gain *k*(*t*).

The circle criterion gives sufficient conditions for the stability of such a system, and can be stated as follows: Suppose the gain $k(t)$ varies in the range $[\alpha,\beta]$, that is, $\alpha \le k(t) \le \beta$ for all t. Let $D(\alpha,\beta)$ denote the critical disk, which passes through $-1/\alpha + j0$ and $-1/\beta + j0$, and is centered on the real axis. Under these conditions, the system of Fig. 4 is stable if one of the following conditions, as appropriate, holds:

Case 1: $0 < \alpha < \beta$. The Nyquist plot of $\hat{g}(j\omega)$ does not intersect the critical disk $D(\alpha,\beta)$ and encircles it in the counterclockwise direction exactly μ times (**Fig. 5**).

Case 2: $0 = \alpha < \beta$. The Nyquist plot of $\hat{g}(j\omega)$ lies in the half-plane $\mathrm{Re}\,\hat{g}(j\omega) < 1/\beta$.

Case 3: $\alpha < 0 < \beta$. The Nyquist plot of $\hat{g}(j\omega)$ is contained within the critical disk $D(\alpha,\beta)$.

Case 4: $\alpha < \beta < 0$. Replace $\hat{g}$ by $-\hat{g}$, α by $-\beta$, β by $-\alpha$, and apply case 1 (since $0 < -\beta < -\alpha$).

The circle criterion shares the same advantages as the Nyquist criterion. In addition, if $\alpha \to \beta$, so that the time-varying gain $k(t)$ approaches a constant gain β, then the critical disk in the circle criterion shrinks to the critical point of the Nyquist criterion.

Popov criterion. Finally, consider the system shown in **Fig. 6**, wherein the feedback contains a memoryless nonlinear element N. Suppose the graph of the nonlinearity lies in the sector $[0,k]$, that is, between the horizontal axis and a line of slope k through the origin (**Fig. 7**). Finally, suppose $\hat{g}(s)$ is strictly proper and that all poles of $\hat{g}$ have negative real parts. The Popov criterion gives a sufficient condition for the stability of such a system.

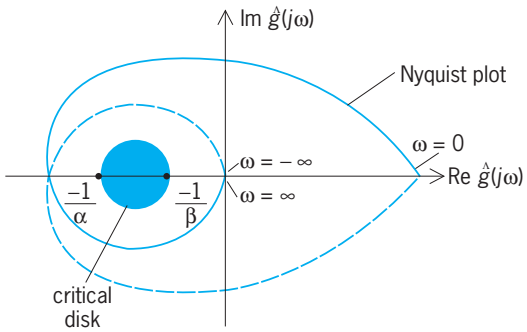

Fig. 5. Circle criterion for $0 < \alpha < \beta$. The Nyquist plot does not intersect the critical disk and encircles it in the counterclockwise direction exactly μ times, where μ is the number of poles of $\hat{g}$ with positive real parts.

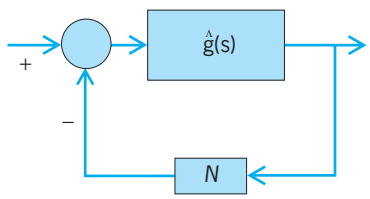

Fig. 6. Feedback system with feedback containing a memoryless nonlinear element *N*.

It states that the system of Fig. 6 is stable, provided one can find a number $q \geq 0$ and a number $\delta > 0$ such that inequality (22) is satisfied. The cri-

$$\text{Re}\,[(1 + j\omega q)\hat{g}(j\omega)] + 1/k \geq \delta > 0 \quad \text{for all } \omega \tag{22}$$

terion (22) can also be given a graphical interpretation by means of a plot of Re $\hat{g}(j\omega)$ versus ω a Im $\hat{g}(j\omega)$, called a Popov plot (**Fig. 8**). It is only necessary to do the plot for $\omega \geq 0$, since both Re $\hat{g}(j\omega)$ and ω Im $\hat{g}(j\omega)$ are even functions of ω. The inequality (22) states that it is possible to draw a line of slope $1/q$ through the point $1/k + j0$ in such a way that the Popov plot of $\hat{g}(j\omega)$ lies strictly to the right of this line (Fig. 8).

Lyapunov theory. This can be used to analyze the behavior of the solutions of the vector differential equation (23), corresponding to various initial con-

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t)) \tag{23}$$

ditions $\mathbf{x}(t_0) = \mathbf{x}_0$ without actually solving the differential equation (23). The form of Eq. (23) is quite general, and is referred to as the set of state equations for a system; it is distinguished by the fact that the system under study is described by a set of first-order differential equations. The form of Eq. (23) is sufficiently general that any set of coupled ordinary differential equations, of any order, can be equivalently expressed in this form. To simplify the exposition, only the case where the time t does not explicitly appear in the state equations will be discussed, that is, where the state equations are of the form of Eq. (24).

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) \tag{24}$$

See DIFFERENTIAL EQUATION; LINEAR SYSTEM ANALYSIS.

Equilibrium. An important concept for the system of Eq. (23) is that of an equilibrium. A vector $\mathbf{x}_0$ is called an equilibrium of the system of Eq. (24) if Eq. (25)

$$\mathbf{f}(\mathbf{x}_0) = \mathbf{0} \tag{25}$$

is satisfied. Equivalently, if $\mathbf{x}_0$ is an equilibrium of Eq. (24), then the solution of Eq. (24) corresponding to the initial condition $\mathbf{x}(0) = \mathbf{x}_0$ is $\mathbf{x}(t) = \mathbf{x}_0$ for all $t \geq 0$.

Asymptotic stability. There are many distinct concepts of "stability" in Lyapunov theory, each of them suitable for a different application, and only one of them will be discussed here. Suppose $\mathbf{x} = \mathbf{0}$ is an equilibrium of Eq. (24) [that is, $\mathbf{f}(\mathbf{0}) = \mathbf{0}$], and let $\mathbf{x}(t,\mathbf{x}_0)$ denote the solution of Eq. (24), corresponding to the initial condition $\mathbf{x}(0) = \mathbf{x}_0$. Finally, let the norm of the vector $\mathbf{x}$ be defined by Eq. (26). Then $\mathbf{0}$ is called

$$|\mathbf{x}| = \left(\sum_i x_i^2\right)^{1/2} \tag{26}$$

an asymptotically stable equilibrium of Eq. (24) if the following two conditions are satisfied: (1) There is a number $\delta_0 > 0$ such that $|\mathbf{x}(t,\mathbf{x}_0)|$ is bounded as a function of t whenever $|x_0| < \delta_0$. Moreover, the function ϕ defined by Eq. (27) is continuous at $r = 0$.

$$\phi(r) = \sup_{|\mathbf{x}_0| \leq r} \sup_{t \geq 0} |\mathbf{x}(t, \mathbf{x}_0)| \tag{27}$$

(2) There is a number $\delta_1 > 0$ $|\mathbf{x}(t,\mathbf{x}_0)| \to 0$ as $t \to \infty$, whenever $|\mathbf{x}_0| < \phi_1$.

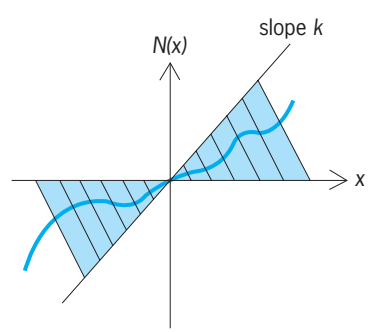

Fig. 7. Graph of nonlinear element *N*.

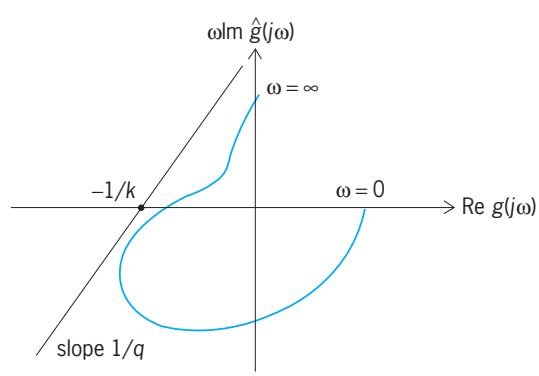

Fig. 8. Popov plot of a transfer function $\hat{g}$.

Basically, condition (1) states that, whenever the magnitude of the initial condition is sufficiently small ($|\mathbf{x}_0| < \delta_0$), small perturbations in the initial condition produce small perturbations in the corresponding solution. Condition (2) states that there is a domain of attraction (the sphere of radius δ_1) such that, whenever the initial condition belongs to this domain, the corresponding solution approaches $\mathbf{0}$ as $t \to \infty$.

Linear systems. In the case of a linear system, Eq. (28), where $\mathbf{x}(t)$ is an $m \times 1$ vector and $\mathbf{a}$ is

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) \tag{28}$$

an $m \times m$ matrix, the equilibrium $\mathbf{x} = \mathbf{0}$ stable if and only if all eigenvalues of $\mathbf{A}$ have negative real parts, that is, the characteristic polynomial $\phi(s) = \det(s\mathbf{I} - \mathbf{A})$ is strictly Hurwitz. Moreover, if $\mathbf{x} = \mathbf{0}$ is asymptotically stable, then the domain of attraction is the entire state space, that is, $|\mathbf{x}(t,\mathbf{x}_0)| \to 0$ as $t \to \infty$ for every $\mathbf{x}_0$.

A test that can be used to determine whether or not all eigenvalues of $\mathbf{A}$ have negative real parts (that is, whether $\mathbf{A}$ is a Hurwitz matrix), without actually computing them, is summarized in the following theorem:

Given an $m \times m$ matrix $\mathbf{A}$, the following three conditions are equivalent: (1) All eigenvalues of $\mathbf{A}$ have negative real parts. (2) For some symmetric positive definite matrix $\mathbf{Q}$, Eq. (29) has a unique solution $\mathbf{P}$;

$$\mathbf{A}^{\mathrm{T}}\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q} \tag{29}$$

moreover, this solution is symmetric and positive definite. (3) For every symmetric positive definite matrix $\mathbf{Q}$, Eq. (29) has a unique solution for $\mathbf{P}$, and this solution is positive definite.

Thus the testing procedure is as follows: Given a matrix $\mathbf{A}$, select a symmetric positive definite matrix $\mathbf{Q}$ arbitrarily, and solve Eq. (29) [known as the Lyapunov matrix equation] for $\mathbf{P}$. If a unique $\mathbf{P}$ cannot be found, or if the unique $\mathbf{P}$ is not positive definite, then $\mathbf{A}$ is not a Hurwitz matrix; if there is a unique $\mathbf{P}$ satisfying Eq. (29) and this $\mathbf{P}$ is positive definite, the $\mathbf{A}$ is a Hurwitz matrix. The test also provides additional information: If Eq. (29) has a unique solution for $\mathbf{P}$, then the number of negative eigenvalues of $\mathbf{P}$ equals the number of eigenvalues of $\mathbf{A}$ having positive real part. *See* EIGENFUNCTION; MATRIX THEORY.

Nonlinear systems. To analyze the nonlinear system of Eq. (24), two general methods are available, namely the direct and indirect methods of Lyapunov. The indirect method is based on expanding $\mathbf{f}(\mathbf{x})$ in a Taylor series around the point $\mathbf{x} = \mathbf{0}$. Specifically, let $\mathbf{A}$ be defined by Eq. (30). Since $\mathbf{f}(\mathbf{0}) = \mathbf{0}$ by assumptions, Eq. (31) follows. Lyapunov's indirect method is

$$\mathbf{A} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\bigg|_{\mathbf{x}=0} \tag{30}$$

$$\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \text{higher} - \text{order terms in } \mathbf{x} \tag{31}$$

based on the following fact: If $\mathbf{A}$ is strictly Hurwitz, then $\mathbf{x} = \mathbf{0}$ is an asymptotically stable equilibrium of the nonlinear system of Eq. (24); if $\mathbf{A}$ has some eigenvalues with positive real parts, then $\mathbf{x} = \mathbf{0}$ is an unstable equilibrium of Eq. (24).

The direct method of Lyapunov is based on constructing an energylike function $v(\mathbf{x})$, known as a Lyapunov function, whose value decreases along the solution trajectories of Eq. (24). The actual details are somewhat complex, but the direct method provides a very powerful tool for analyzing nonlinear systems. *See* NONLINEAR CONTROL THEORY.

Advanced methods. The techniques described here have been extended to systems governed by partial or even functional differential equations, as well as to multidimensional systems (for example, two-dimensional digital filters). In addition, advanced mathematical techniques such as functional analysis and differential geometry have been applied to stability problems. *See* CONTROL SYSTEMS. M. Vidyasagar

Bibliography. E. I. Jury, *Inners and Stability of Dynamic Systems*, 2d ed., 1982; A. M. Lyapunov, *The General Problem of the Stability of Motion*, 1992; P. Parks and V. Hahn, *Stability Theory*, 1992; S. Sastry, *Nonlinear Systems: Analysis, Stability and Control*, 1993; M. Vidyasagar, *Nonlinear Systems Analysis*, 2d ed., 1992.

Control systems

Interconnections of components forming system configurations which will provide a desired system response as time progresses. The steering of an automobile is a familiar example. The driver observes the position of the car relative to the desired location and makes corrections by turning the steering wheel. The car responds by changing direction, and the driver attempts to decrease the error between the desired and actual course of travel. In this case, the controlled output is the automobile's direction of travel, and the control system includes the driver, the automobile, and the road surface. The control engineer attempts to design a steering control mechanism which will provide a desired response for the automobile's direction control. Different steering designs and automobile designs result in rapid responses, as in the case of sports cars, or relatively slow and comfortable responses, as in the case of large autos with power steering.

Open- and closed-loop control. The basis for analysis of a control system is the foundation provided by linear system theory, which assumes a cause-effect relationship for the components of a system. A component or process to be controlled can be represented by a block. Each block possesses an input (cause) and output (effect). The input-output relation represents the cause-and-effect relationship of the process, which in turn represents a processing of the input signal to provide an output signal variable, often with power amplification. An open-loop control system utilizes a controller or control actuator in order to obtain the desired response (**Fig. 1**). *See* BLOCK DIAGRAM.

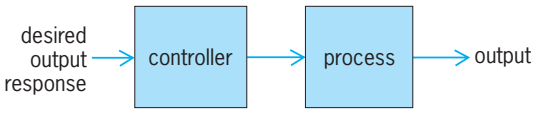

Fig. 1. Open-loop control system.

A closed-loop control system utilizes an additional measure of the actual output in order to compare the actual output with the desired output response (**Fig. 2**). A standard definition of a feedback control system is a control system which tends to maintain a prescribed relationship of one system variable to another by comparing functions of these variables and using the difference as a means of control. In the case of the driver steering an automobile, the driver visually measures and compares the actual location of the car with the desired location. The driver then serves as the controller, turning the steering wheel. The process represents the dynamics of the steering mechanism and the automobile response.

A feedback control system often uses a function of a prescribed relationship between the output and reference input to control the process. Often, the difference between the output of the process under control and the reference input is amplified and used to control the process so that the difference is continually reduced. The feedback concept has been the foundation for control system analysis and design.

Because of the increasing complexity of the system under control and the interest in achieving optimum performance, the importance of control system engineering has grown over the years. Furthermore, as the systems become more complex, the interrelationship of many controlled variables must be considered in the control scheme.

History of automatic control. The first automatic feedback controller used in an industrial process was James Watt's flyball governor (1769) for controlling the speed of a steam engine. Prior to World War II, one main impetus for the use of feedback control in the United States was the development of the telephone system and electronic feedback amplifiers. The frequency domain was used primarily to describe the operation of the feedback amplifiers in terms of bandwidth and other frequency variables. In contrast, the Russian theory, which was dominated by mathematicians, tended to utilize a time-domain formulation using differential equations.

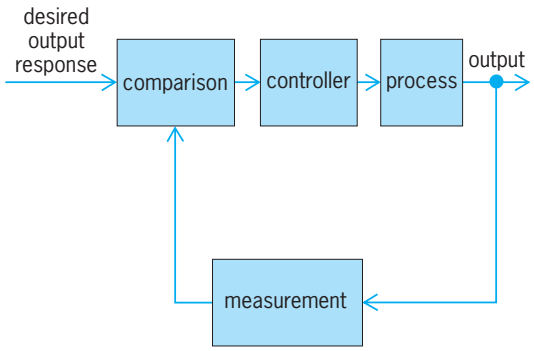

Fig. 2. Closed-loop control system.

A large impetus to the theory and practice of automatic control occurred during World War II, when it became necessary to design and construct automatic airplane pilots, gun-positioning systems, radar antenna control systems, and other military systems based on the feedback control approach. The complexity and expected performance of these military systems necessitated an extension of the available control techniques and fostered interest in control systems and the development of new insights and methods.

Frequency-domain techniques continued to dominate the field of control following World War II, with the increased use of the Laplace transform and the complex frequency plane. During the 1950s, the utilization of both analog and digital computers for control components became possible. These new controlling elements possessed an ability to calculate rapidly and accurately, which was formerly not available to the control engineer. Digital process control computers are employed especially for process control systems in which many variables are measured and controlled simultaneously by the computer. *See* ANALOG COMPUTER; DIGITAL COMPUTER; PROCESS CONTROL.

With the advent of the space age, it became necessary to design complex, highly accurate control systems for missiles and space probes. Furthermore, the necessity to minimize the weight of satellites and to control them very accurately spawned the important field of optimal control. It appears that control engineering must consider both the time-domain and the frequency-domain approaches simultaneously in the analysis and design of control systems. *See* GUIDANCE SYSTEMS; OPTIMAL CONTROL THEORY; SPACE NAVIGATION AND GUIDANCE.

Applications. Familiar control systems have the basic closed-loop configuration as shown earlier. For example, a refrigerator has a temperature setting for desired temperature, a thermostat to measure the actual temperature and the error, and a compressor motor for power amplification. Other examples in the home are the oven, furnace, and water heater. In industry, there are controls for speed, process temperature and pressure, position, thickness, composition, and quality, among many others.

In order to provide a mass transportation system for modern urban areas, a large, complex, high-speed system is necessary. Automatic control is necessary in order to maintain a constant flow of trains and to ensure comfortable deceleration and braking conditions at stations (**Fig. 3**). A measurement of the distance from the station and the speed of the train is used to determine the error signal and therefore the braking signal. *See* RAILROAD CONTROL SYSTEMS.

The electric power industry is primarily interested in energy conversion, control, and distribution. It is critical that computer control be increasingly applied to the power industry in order to improve

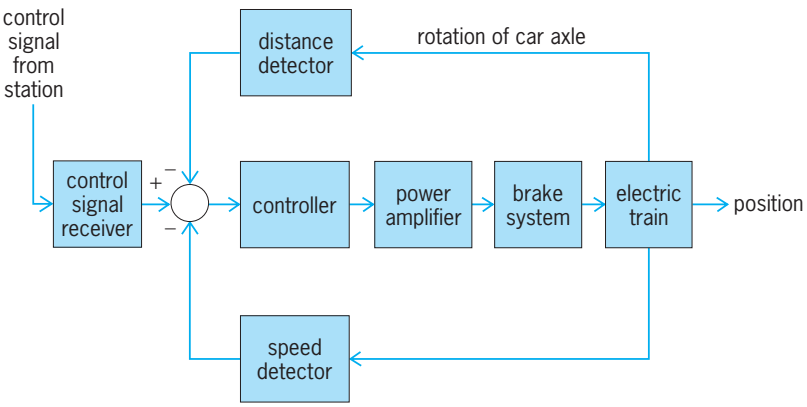

Fig. 3. Train-brake control system.

the efficiency of use of energy resources. Also, the control of power plants for minimum waste emission has become increasingly important. The modern large-capacity plants which exceed several hundred megawatts require automatic control systems which account for the interrelationship of the process variables and the optimum power production. It is common to have as many as 90 or more manipulated variables under coordinated control. Computer controls are also used to control energy use in industry and stabilize and connect loads evenly to gain fuel economy.

Feedback control concepts have been applied to automatic warehousing and inventory control and to the automatic control of agriculture systems (farms). Automatic control of wind turbine generators, solar heating and cooling, and automobile engine performance are other important examples.

Also, there have been many applications of control system theory to biomedical experimentation, diagnosis, prosthetics, and biological control systems. The control systems under consideration range from the cellular level to the central nervous system and include temperature regulation and neurological, respiratory, and cardiovascular control. Most physiological control systems are closed-loop systems. However, they involve not one controller but rather control loop within control loop, forming a hierarchy of systems. The modeling of the structure of biological processes confronts the analyst with a high-order model and a complex structure. Prosthetic devices are designed to provide automatically controlled aids to the disabled. An artificial hand that uses force feedback signals and is controlled by the amputee's bioelectric control signals is an example. *See* MATHEMATICAL BIOLOGY; PROSTHESIS.

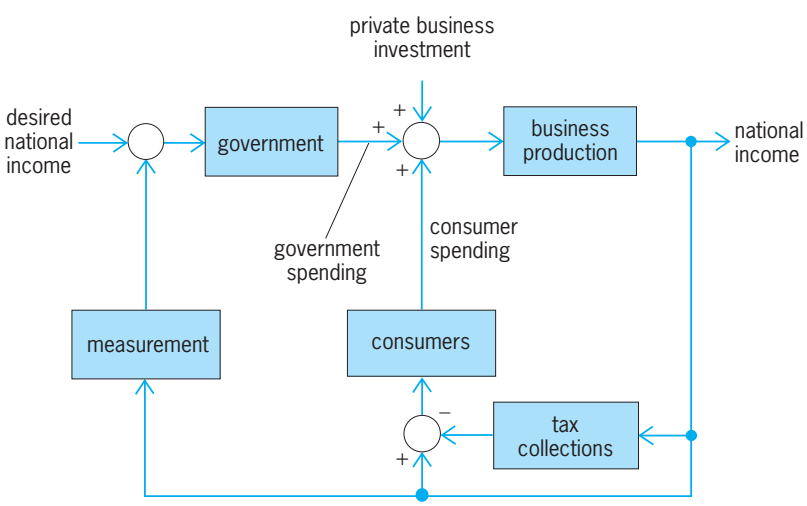

Fig. 4. Feedback control system model of the United States economy.

Finally, it has become of interest to attempt to model the feedback processes prevalent in the social, economic, and political spheres. This approach is undeveloped but appears to have a reasonable future. A simple lumped model of the national-income feedback control system is shown in **Fig. 4**. This type of model helps the analyst to understand the effects of government control—granted its existence—and the dynamic effects of government spending. Of course, many other loops not shown also exist. This type of political or social feedback model, while usually nonrigorous, does impart information and understanding. *See* SYSTEMS ANALYSIS; SYSTEMS ENGINEERING.

Analysis using the Laplace transform. In order to understand a complex control system, it is necessary to analyze the relationships between the system variables and to obtain a quantitative mathematical model of the system. Since the systems under consideration are dynamic in nature, the descriptive equations are usually differential equations. Furthermore, if these equations can be linearized, then the Laplace transform can be utilized in order to simplify the method of solution. In practice, the complexity of systems and incomplete knowledge of the relevant factors necessitate the introduction of assumptions concerning the system operation. Therefore, it is often useful to consider the physical system, delineate some necessary assumptions, and linearize the system. Then, by using the physical laws describing the linear equivalent system, one can obtain a set of linear differential equations. Finally, utilizing mathematical tools, such as the Laplace transform, a solution describing the operation of the system is obtained. In summary, the approach to dynamic system problems can be listed as follows: (1) define the system and its components; (2) formulate the mathematical model and list the necessary assumptions; (3) write the differential equations describing the model; (4) solve the equations for the desired output variables; (5) examine the solutions and the assumptions; and then (6) reanalyze or design. *See* DIFFERENTIAL EQUATION.

The Laplace transform of a function f of time t is given by Eq. (1), where s is a complex variable

$$F(s) = \mathcal{L}\{f(t)\} = \int_0^\infty f(t)e^{-st}\,dt \qquad (1)$$

related to the frequency ω by the equations $s = \sigma + j\omega$ (where $j = \sqrt{-1}$). The analysis thus takes place in the frequency domain. The inverse Laplace

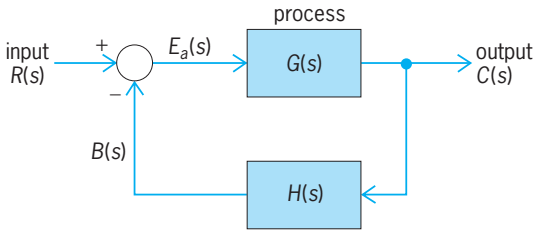

Fig. 5. Negative-feedback control system.

transformation is usually obtained by a partial fraction expansion. *See* LAPLACE TRANSFORM.

The transfer function of a linear system is the ratio of the Laplace transform of the output variable to the Laplace transform of the input variable with all initial conditions assumed to be zero. The block diagram of a negative-feedback control system is shown in **Fig. 5**. The transfer function $G(s)$ of the process is given by Eq. (2). The transfer function of the closed-

$$\frac{C(s)}{E_a(s)} = G(s) \tag{2}$$

loop system can be obtained by the algebraic manipulation of the system equations. The error, $E_a(s)$, is given by Eq. (3), and the output is given by Eq. (4).

$$E_a(s) = R(s) - B(s) = R(s) - H(s)C(s) \tag{3}$$

$$C(s) = G(s)E_a(s) = G(s)[R(s) - H(s)C(s)] \tag{4}$$

Solving for $C(s)$, Eq. (5) is obtained, and therefore $C(s)$ is given by Eq. (6). The closed-loop transfer func-

$$C(s)[1 + G(s)H(s)] = G(s)R(s) \tag{5}$$

$$C(s) = \frac{G(s)}{1 + G(s)H(s)}R(s) \tag{6}$$

tion $T(s) = C(s)/R(s)$ is given by Eq. (7), and the error is then given by Eq. (8). It is clear that in order to re-

$$\frac{C(s)}{R(s)} = T(s) = \frac{G(s)}{1 + G(s)H(s)} \tag{7}$$

$$E_a(s) = \frac{1}{1 + G(s)H(s)}R(s) \tag{8}$$

duce the error, the magnitude of $[1 + G(s)H(s)]$ must be greater than 1 over the range of s under consideration.

Advantages of feedback control. The addition of feedback to a control system results in several important advantages.

Reduction in sensitivity. A process, represented by the transfer function $G(s)$, whatever its nature, is subject to a changing environment, aging, ignorance of the exact values of the process parameters, and other natural factors which affect a control process. In the open-loop system, all these errors and changes result in a changing and inaccurate output. However, a closed-loop system senses the change in the output due to the process changes and attempts to correct the output. The sensitivity of a control system to parameter variations is of prime importance. A primary advantage of a closed-loop feedback control system is its ability to reduce the system's sensitivity.

For the closed-loop case, if $G(s)H(s)$ is much greater than 1 for all complex frequencies of interest, then from Eq. (6) one obtains Eq. (9). That is,

$$C(s) = \frac{1}{H(s)}R(s) \tag{9}$$

the output is affected only by $H(s)$, which may be constant. If $H(s) = 1$, the desired result is obtained; that is, the output is equal to the input. However, this approach cannot be used for all control systems because the requirement that $G(s)H(s)$ be much greater than 1 may cause the system response to be highly oscillatory and even unstable. But the fact that an increase in the magnitude of the loop transfer function $G(s)H(s)$ results in a reduction of the effect of $G(s)$ on the output is an exceedingly useful concept. Therefore, the first advantage of a feedback system is that the effect of the variation of the parameters of the process, $G(s)$, is reduced.

Control of transient response. One of the most important characteristics of control systems is their transient response, which often must be adjusted until it is satisfactory. If an open-loop control system does not provide a satisfactory response, then the process, $G(s)$, must be replaced with a suitable process. By contrast, a closed-loop system can often be adjusted to yield the desired response by adjusting the feedback loop parameters. It is often possible to alter the response of an open-loop system by inserting a suitable cascade filter, $G_1(s)$, preceding the process, $G(s)$. Then it is necessary to design the cascade transfer function $G_1(s)G(s)$ so that the resulting transfer function provides the desired transient response. Therefore, the second advantage of feedback control is control of the transient response of the system.

Reduction of effects of disturbance. The third most important effect of feedback in a control system is the control and partial elimination of the effect of disturbance signals. Many control systems are subject to extraneous disturbance signals which cause the system to provide an inaccurate output. Electronic amplifiers have inherent noise generated within the integrated circuits or transistors, radar antennas are subjected to wind gusts, and many systems generate unwanted distortion signals due to nonlinear elements. Feedback systems have the beneficial aspect that the effect of distortion, noise, and unwanted disturbances can be effectively reduced.

As a specific example of a system with an unwanted disturbance, consider the speed control system for a steel rolling mill. Rolls passing steel through are subject to large load changes or disturbances. As a steel bar approaches the rolls, the rolls turn unloaded. However, when the bar engages in the rolls, the load on the rolls increases immediately to a large value. This loading effect can be approximated by a step change of disturbance torque.

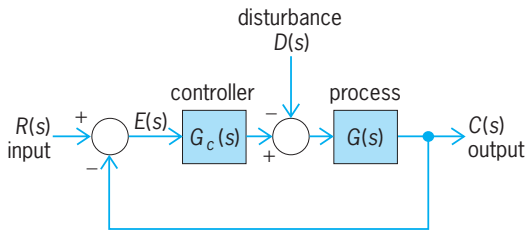

Fig. 6. Feedback control system with a disturbance.

The block diagram of a feedback control system with a disturbance, $D(s)$, is shown in **Fig. 6**, where $H(s) = 1$. Then the output, $C(s)$, may be shown to be given by Eq. (10) when the input, $R(s)$, is set to zero.

$$C(s) = \frac{-G(s)D(s)}{1 + G_c(s)G(s)} \tag{10}$$

Then if $G_c(s)G(s)$ is much greater than one over the range of s, the approximate result given by Eq. (11)

$$C(s) \simeq -\frac{1}{G_c(s)}D(s) \tag{11}$$

is obtained. Therefore, if $G_c(s)$ is made sufficiently large, the effect of the disturbance can be decreased by closed-loop feedback.

Reduction of steady-state error. The reduction of the steady-state error of a control system is another advantage of feedback control.

The error of the closed-loop system when $H(s) = 1$ is given by Eq. (12). The steady-state error can be

$$E(s) = \frac{1}{1 + G(s)}R(s) \tag{12}$$

calculated by using the final-value theorem which is given by Eq. (13), where $e(t)$ is the error variable in

$$\lim_{t\to\infty} e(t) = \lim_{s\to 0} sE(s) \tag{13}$$

the time domain. Therefore, using a unit step input as comparable input, Eq. (14) is obtained for the closed-

$$e(\infty) = \lim_{s\to 0} s\left(\frac{1}{1 + G(s)}\right)\left(\frac{1}{s}\right) = \frac{1}{1 + G(0)} \tag{14}$$

loop system. The value of $G(s)$ when $s = 0$ is often called the dc gain and is normally greater than 1. Therefore, the open-loop system will usually have a steady-state error of significant magnitude. By contrast, the closed-loop system with a reasonably large dc-loop gain $G(0)H(0)$ will have a small steady-state error.

Costs of feedback control. While the addition of feedback to a control system results in the advantages outlined above, it is natural that these advantages have an attendant cost.

Increased complexity. The cost of feedback is first manifested in the increased number of components and the complexity of the system. In order to add the feedback, it is necessary to consider several feedback components, of which the measurement component (sensor) is the key component. The sensor is often the most expensive component in a control system. Furthermore, the sensor introduces noise and inaccuracies into the system.

Loss of gain. The second cost of feedback is the loss of gain. For example, in a single-loop system, the open-loop gain is $G(s)$ and is reduced to $G(s)/[1 + G(s)]$ in a unity negative-feedback system. The reduction in closed-loop gain is $1/[1 + G(s)]$ which is exactly the factor that reduces the sensitivity of the system to parameter variations and disturbances. Usually, there is open-loop gain to spare, and one is more than willing to trade it for increased control of the system response.

However, it is the gain of the input-output transmittance that is reduced. The control system does possess a substantial power gain which is fully utilized in the closed-loop system.

Possibility of instability. Finally, a cost of feedback is the introduction of the possibility of instability. While the open-loop system is stable, the closed-loop system may not be always stable.

Necessity of considering problems. The addition of feedback to dynamic systems results in several additional problems for the designer. However, for most cases, the advantages far outweigh the disadvantages, and a feedback system is utilized. Therefore, it is necessary to consider the additional complexity and the problem of stability when designing a control system.

It is desired that the output of the system $C(s)$ equal the input $R(s)$. However, one might ask, "Why not simply set the transfer function $G(s) = C(s)/R(s)$ equal to 1?" The answer to this question is that the process (or plant) $G(s)$ was necessary in order to provide the desired output; that is, the transfer function $G(s)$ represents a real process and possesses dynamics which cannot be neglected. To set $G(s)$ equal to 1 implies that the output is directly connected to the input. However, it must be recalled that a specific output, such as temperature, shaft rotation, or engine speed, is desired, while the input might be a potentiometer setting or a voltage. The process $G(s)$ is necessary in order to provide the physical process between $R(s)$ and $C(s)$. Therefore, a transfer function $G(s) = 1$ is unrealizable, and a practical transfer function must be adopted.

State variables and the time-domain. A useful approach to the analysis and design of feedback systems was outlined above. The Laplace transform was utilized to transform the differential equations representing the system into an algebraic equation expressed in terms of the complex variable, s. Utilizing this algebraic equation, a transfer function representation of the input-output relationship was obtained. The Laplace transform approach, carried out in terms of the complex variable s and accomplished in the frequency domain, is extremely useful; it is and will remain one of the primary tools of the control engineer. However, the limitations of the frequency-domain techniques and the attractiveness of the time-domain approach require a reconsideration of the time-domain formulation of the equations representing control systems.

The frequency-domain techniques are limited in applicability to linear, time-invariant systems.

Furthermore, they are particularly limited in their usefulness for multivariable control systems because of the emphasis on the input-output relationship of transfer functions. By contrast, the time-domain techniques may be readily utilized for nonlinear, time-varying, and multivariable systems. A time-varying control system is a system for which one or more of the parameters of the system may vary as a function of time. For example, the mass of a missile varies as a function of time as the fuel is expended during flight. A multivariable system is a system with several input and output signals. Furthermore, the solution of a time-domain formulation of a control system problem is facilitated by the availability and ease of use of digital and analog computers.

Definition of state variables. The time-domain analysis and design of control systems utilizes the concept of the state of a system. The state of a system is a set of numbers such that the knowledge of these numbers and the input functions will, with the equations describing the dynamics, provide the future state and output of the system. For a dynamic system, the state of a system is described in terms of a set of state variables $[x_1(t), x_2(t), \ldots, x_n(t)]$. The state variables determine the future behavior of a system when the present state of the system and the excitation signals are known. A set of state variables $(x_1, x_2, \ldots, x_n)$ is a set such that knowledge of the initial values of the state variables $[x_1(t_0), x_2(t_0), \ldots, x_n(t_0)]$ at the initial time t_0, and of the input signals $u_1(t)$ and $u_2(t)$ for times t greater than t_0, suffices to determine the future values of the output and state variables. *See* LINEAR SYSTEM ANALYSIS.

A simple example of a state variable is the state of an on-off light switch. The switch can be in either the on or off position, and thus the state of the switch can assume one of two possible values. Thus if the present state (position) of the switch is known at t_0 and if an input is applied, the future value of the state of the element can be determined.

Example of dynamic system. The concept of a set of state variables which represent a dynamic system can be illustrated in terms of the spring-mass-damper system shown in **Fig. 7**. The number of state variables chosen to represent this system should be as few as possible in order to avoid redundant state variables.

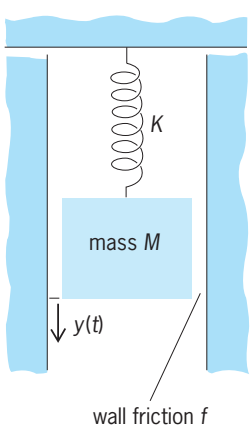

Fig. 7. Spring-and-mass system with wall friction.

A set of state variables sufficient to describe this system is the position $y(t)$ and the velocity of the mass. Therefore, a set of state variables is defined as (x_1, x_2), given by Eqs. (15).

$$x_1(t) = y(t) \qquad x_2(t) = \frac{dy(t)}{dt} \tag{15}$$

The differential equation describing the behavior of the system is usually written as Eq. (16), where M

$$M\frac{d^2y}{dt^2} + f\frac{dy}{dt} + Ky = u(t) \tag{16}$$

is the mass, K is the force constant of the spring, and the coefficient f determines the frictional force exerted by the wall. In order to write Eq. (16) in terms of the state variables, one substitutes the definition of the state variables and obtains Eq. (17). Therefore,

$$M\frac{dx_2}{dt} + fx_2 + Kx_1 = u(t) \tag{17}$$

the differential equations which describe the behavior of the spring-mass-damper system may be written as a set of two first-order differential equations, Eqs. (18) and (19). This set of differential equations

$$\frac{dx_1}{dt} = x_2 \tag{18}$$

$$\frac{dx_2}{dt} = \frac{-f}{M}x_2 - \frac{K}{M}x_1 + \frac{1}{M}u \tag{19}$$

describes the behavior of the state of the system in terms of the rate of change of each state variable.

Matrix differential equation. The state of a system is described by the set of first-order differential equations written in terms of the state variables $(x_1, x_2, \ldots, x_n)$. These first-order differential equations may be written in general form as Eqs. (20), where $\dot{x} = dx/dt$.

$$\begin{aligned}
\dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \cdots \\
&\quad + a_{1n}x_n + b_{11}u_1 + \cdots + b_{1m}u_m \\
\dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \cdots \\
&\quad + a_{2n}x_n + b_{21}u_1 + \cdots + b_{2m}u_m \\
&\;\;\vdots \\
\dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \cdots \\
&\quad + a_{nn}x_n + b_{n1}u_1 + \cdots + b_{nm}u_m
\end{aligned} \tag{20}$$

Thus this set of simultaneous differential equations may be written in matrix form. The column matrix consisting of the state variables is called the state vector and is written as Eq. (21), where the bold-

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \tag{21}$$

face indicates a matrix. The matrix of input signals

is defined as **u**. Then the system may be represented by the compact notation of the system vector differential equation as Eq. (22). The matrix **A** is an

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu} \tag{22}$$

$n \times n$ square matrix, and **B** is an $n \times m$ matrix. The vector matrix differential equation relates the rate of change of the state of the system to the state of the system and the input signals. In general, the outputs of a linear system may be related to the state variables and the input signals by the vector matrix equation (23), where **c** is the set of output signals expressed in

$$\mathbf{c} = \mathbf{Dx} + \mathbf{Hu} \tag{23}$$

column vector form. *See* MATRIX CALCULUS; MATRIX THEORY.

The solution of the matrix differential equation uses the matrix exponential given by Eq. (24). The

$$\phi(t) = e^{\mathbf{A}t} \tag{24}$$

matrix exponential function describes the unforced response of the system and is called the fundamental or transition matrix $\phi(t)$. Therefore, the solution may be written as Eq. (25).

$$\mathbf{x}(t) = \phi(t)\mathbf{x}(0) + \int_0^t \phi(t-\tau)\mathbf{Bu}(\tau)\,d\tau \tag{25}$$

The solution to the unforced system (that is, when $u = 0$) is given simply by Eq. (26). Hence, in order to

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ \cdot \\ \cdot \\ \cdot \\ x_n(t) \end{bmatrix} = \begin{bmatrix} \phi_{11}(t) & \cdots & \phi_{1n}(t) \\ \phi_{21}(t) & \cdots & \phi_{2n}(t) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \phi_{n1}(t) & \cdots & \phi_{nn}(t) \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \\ \cdot \\ \cdot \\ \cdot \\ x_n(0) \end{bmatrix} \tag{26}$$

determine the transition matrix, all initial conditions are set to zero except for one state variable, and the output of each state variable is evaluated. That is, the term $\phi_{ij}(t)$ is the response of the ith state variable due to an initial condition on the jth state variable when there are zero initial conditions on all the other states. The time-domain approach may be used to investigate the performance of a control system as well as to design improved systems.

Time response. The time response of a second-order feedback control system (**Fig. 8**) is illustrative of the normal range of response of a control system. The transfer function of the process of such a system has the general form $G(s) = K/s(s + p)$, where K is the gain of the process, and the transfer function contains a pole at the origin of the s-plane and one pole at $s = -p$. The response to

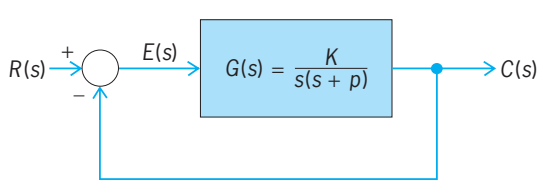

Fig. 8. Closed-loop second-order control system.

a unit step input will be determined. The closed-loop output is given by Eq. (27). Utilizing a generalized notation, Eq. (27) may be rewritten as Eq. (28),

$$C(s) = \frac{G(s)}{1 + G(s)}R(s) = \frac{K}{s^2 + ps + K}R(s) \tag{27}$$

$$C(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n^2 + \omega_n^2}R(s) \tag{28}$$

where ω_n is the natural frequency and ζ is called the damping ratio. With a unit step input, this becomes Eq. (29), for which the transient output is given by Eq. (30), where $\beta = \sqrt[2]{1 - \zeta^2}$ and $\phi = \tan^{-1}\beta/\zeta$.

$$C(s) = \frac{\omega_n^2}{s(s^2 + 2\zeta\omega_n s + \omega_n^2)} \tag{29}$$

$$c(t) = 1 - \frac{1}{\beta}e^{-\zeta\omega}n^t \sin(\omega_n t + \theta) \tag{30}$$

The transient response of this second-order system for various values of the damping ratio ζ is shown in **Fig. 9**. As ζ decreases, the overshoot of the response increases. It is usually desired to have a rapid response to a step input while minimizing the overshoot.

Standard performance measures are usually defined in terms of the step response of a system (**Fig. 10**). The swiftness of the response is measured by the rise time T_r and the peak time T_p. For underdamped systems with an overshoot, the 0–100% rise time is a useful index.

The similarity with which the actual response matches the step input is measured by the percent overshoot and settling time T_s. The percent overshoot, P.O., is defined by Eq. (31) for a unit step input,

$$\text{P.O.} = \frac{M_{pt} - 1}{1} \times 100\% \tag{31}$$

where M_{pt} is the peak value of the time response. The settling time T_s is defined as the time required for the system to settle within a certain percentage δ of the input amplitude. This band of $\pm\delta$ is shown in Fig. 10. For the second-order system with a closed-loop damping constant ω_n, the response remains within 2% after four time constants τ, Eq. (32). Therefore, the settling time is defined as four time constants of the dominant response. Finally, the steady-state error of the system may be measured on the step response of the system as shown in Fig. 10. *See* TIME CONSTANT.

$$T_s = 4\tau = \frac{4}{\zeta\omega_n} \tag{32}$$

Stability of closed-loop systems. The transient response of a feedback control system is of primary interest and must be investigated. A very important characteristic of the transient performance of a system is the stability of the system. A stable system is defined as a system with a bounded system response. That is, if the system is subjected to a bounded input

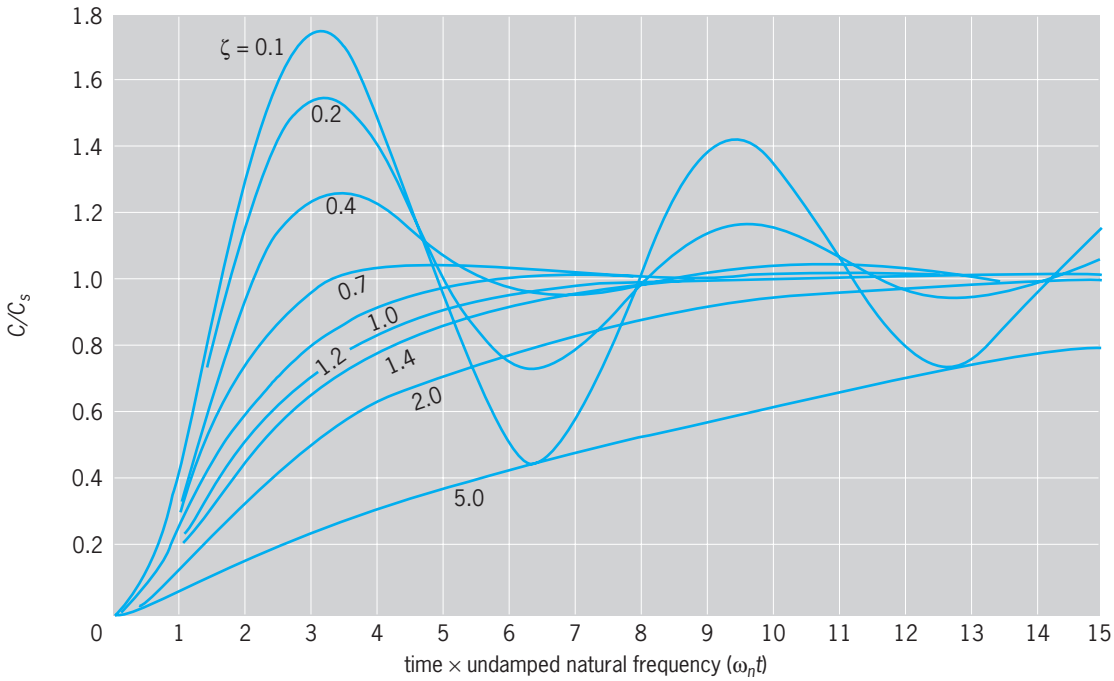

Fig. 9. Response of a second-order system to a step input. C = output at time t; C_s = output when time t is ∞.

or disturbance and the response is bounded in magnitude, the system is said to be stable.

The concept of stability can be illustrated by considering a right circular cone placed on a plane horizontal surface. If the cone is resting on its base and is tipped slightly, it returns to its original equilibrium position. This position and response is said to be stable. If the cone rests on its side and is displaced slightly, it rolls with no tendency to leave the position on its side. This position is designated as neutral stability. On the other hand, if the cone is placed on its tip and released, it falls onto its side. This position is said to be unstable.

The stability of a dynamic system is defined in a similar manner. The response to a displacement, or initial condition, will result in either a decreasing, neutral, or increasing response. Specifically, it follows from the definition of stability that a linear system is stable if and only if the absolute value of its response to an impulse input $g(t)$, integrated over an infinite range, is finite.

There are three general approaches to determining the stability of a system: (1) the s-plane approach, (2) the frequency-plane ($j\omega$) approach, and (3) the time-domain approach. Only the first will be considered here.

The stability requirement of a linear system may be defined in terms of the location of the poles in the closed-loop system transfer function, which may be written as Eq. (33), where $q(s) = \Delta(s)$ is the characteristic polynomial whose roots are the poles of the closed-loop system. The output response for an impulse function input is then given by Eq. (34),

$$T(s) = \frac{p(s)}{q(s)} = \frac{K\prod_{i=1}^{M}(s+z_i)}{S^N\prod_{k=1}^{Q}(s+\sigma_k)\prod_{m=1}^{R}[s^2+2\alpha_m s+(\alpha_m^2+\omega_m^2)]} \tag{33}$$

$$c(t) = \sum_{k=1}^{Q} A_k e^{-\sigma} k^t + \sum_{m=1}^{R} B_m \left(\frac{1}{\omega_m}\right) e^{-\alpha} m^t \sin \omega_m^t \tag{34}$$

when $N = 0$. It follows from Eq. (34) that the poles in the left-hand portion of the s plane result in a decreasing response for disturbance inputs. Similarly, poles on the imaginary axis (the $j\omega$ axis) and in the

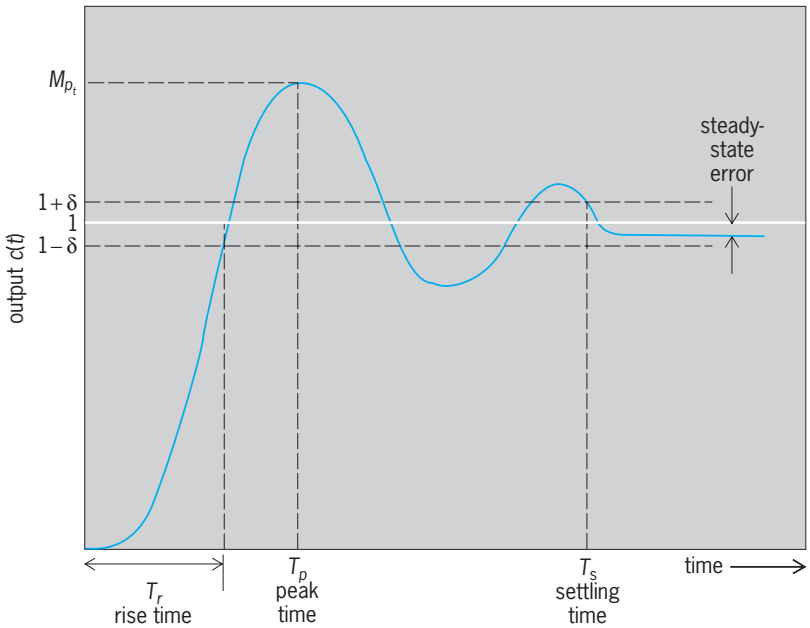

Fig. 10. Step response of a control system.

right-hand plane result in a neutral and an increasing response, respectively, for a disturbance input. Clearly, in order to obtain a bounded response, the poles of the closed-loop system must be in the left-hand portion of the s plane. Thus, a necessary and sufficient condition that a feedback system be stable is that all the poles of the system transfer function have negative real parts.

Therefore, one can ascertain the stability of a feedback control system by determining the roots of the characteristic polynomial $q(s)$. A method for determining whether any of the roots of $q(s)$ lie in the right half of the s plane, and thereby ascertaining the stability of the system, was developed independently by A. Hurwitz and E. J. Routh, in the late 1800s. *See* CONTROL SYSTEM STABILITY.

Design and compensation. The performance of a feedback control system is of primary importance. A suitable control system is stable and results in an acceptable response to input commands, is less sensitive to the system parameter changes, results in a minimum steady-state error for input commands, and finally, is able to eliminate the effect of undesirable disturbances. A feedback control system that provides an optimum performance without any necessary adjustments is rare indeed. Usually one finds it necessary to compromise among the many conflicting and demanding specifications and to adjust the system parameters to provide a suitable and acceptable performance when it is not possible to obtain all the desired optimum specifications.

As discussed above, it is often possible to adjust the system parameters in order to provide the desired system response. However, it is often not possible to simply adjust a system parameter and thus obtain the desired performance. Rather, the scheme or plan of the system must be reexamined, and a new design or plan must be obtained which results in a suitable system. Thus, the design of a control system is concerned with the arrangement, or the plan, of the system structure and the selection of suitable components and parameters. For example, if one desires a set of performance measures to be less than some specified values, one often encounters a conflicting set of requirements. If these two performance requirements cannot be relaxed, the system must be altered in some way. The alteration or adjustment of a control system, in order to make up for deficiencies and inadequacies and provide a suitable performance, is called compensation.

In redesigning a control system in order to alter the system response, an additional component or device is inserted within the structure of the feedback system to equalize or compensate for the performance deficiency. The compensating device may be an electric, mechanical, hydraulic, pneumatic, or other type of device or network, and is often called a compensator. An electric circuit serves as a compensator in many control systems. The transfer function of the compensator is designated as $G_c(s) = E_{out}(s)/E_{in}(s)$, and the compensator may be placed in a suitable location within the structure of the system. The system when the compensator is placed in the forward path is illustrated in **Fig. 11**.

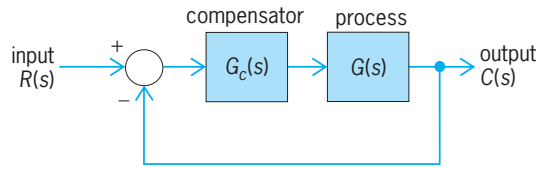

Fig. 11. Cascade compensator in a feedback system.

The performance of a system may be specified by requiring a certain peak time T_p, maximum overshoot, and settling time for a step input. Furthermore, it is usually necessary to specify the maximum allowable steady-state error for several test signal inputs and disturbance inputs. These performance specifications may be defined in terms of the desirable location of the poles and zeros of the closed-loop system transfer function, $T(s)$. Thus, the performance of a control system may be described in terms of frequency-domain performance measures.

The time-domain method, expressed in terms of state variables, may also be utilized to design a suitable compensation scheme for a control system. Typically, one is interested in controlling the system with a control signal $u(t)$, which is a function of several measurable state variables. Then one develops a state-variable controller which operates on the information available in measured form. This type of system compensation is quite useful for system optimization.

Quite often, in practice, the best and simplest way to improve the performance of a control system is to alter, if possible, the process itself. That is, if the system designer is able to specify and alter the design of the process which is represented by the transfer function $G(s)$, then the performance of the system may be readily improved. For example, in order to improve the transient behavior of a servomechanism position controller, one can often choose a better motor for the system. In the case of an airplane control system, one might be able to alter the aerodynamic design of the airplane and thus improve the flight transient characteristics. However, often the process is fixed and unalterable or has been altered as much as is possible and is still found to result in an unsatisfactory performance. Then the addition of compensation networks becomes useful for improving the performance of the system. In the following discussion, it will be assumed that the process has been improved as much as possible and that the $G(s)$ representing the process is unalterable.

Digital computer control systems. The use of a digital computer as a compensator device grew after 1970 as the price and reliability of digital computers improved dramatically. A block diagram of a single-loop digital control system is shown in **Fig. 12**. The digital computer in this system configuration receives the input and output and performs calculations in order to provide an output $u^*(t)$. It may

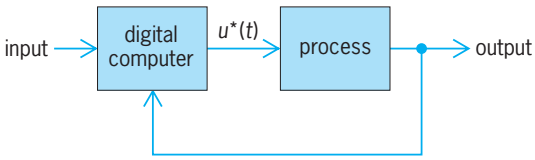

Fig. 12. Digital computer control system.

be programmed to provide an output $u^*(t)$, so that the performance of the process is near or equal to that desired. Many computers are able to receive and manipulate several inputs, so a digital computer control system can often be a multivariable system. *See* DIGITAL COMPUTER; DIGITAL CONTROL.

Within a computer control system, the digital computer receives and operates on signals in digital (numerical) form, as contrasted to continuous signals. The measurement data are converted from analog form to digital form by means of a converter. After the digital computer has processed the inputs, it provides an output in digital form. This output is then converted to analog form through the use of a digital-to-analog converter. *See* ANALOG-TO-DIGITAL CONVERTER; DIGITAL-TO-ANALOG CONVERTER.

Automatic handling equipment for home, school, and industry is particularly useful for hazardous, repetitious, dull, or simple tasks. Machines that automatically load and unload, cut, weld, or cast are used by industry in order to obtain accuracy, safety, economy, and productivity.

Robots are programmable computers integrated with machines. They often substitute for human labor in specific repeated tasks. Some devices even have anthropomorphic mechanisms, including what might be recognized as mechanical arms, wrists, and hands. Robots may be used extensively in space exploration and assembly as well as for various industrial applications. *See* ROBOTICS. Richard C. Dorf

Modeling and simulation. The design of feedback control systems generally assumes that a mathematical model of the system to be controlled is available. The model may have been obtained experimentally by applying test input signals and measuring the response, or analytically by utilizing the laws of physics to derive the equations representing the system. The model would generally be expressed either as a differential equation relating the system's input and output signals or as a transfer function in the complex Laplace variable $s = \sigma + j\omega$ for the special case where the system is linear. The values of the coefficients in the differential equation or transfer function would be functions of the physical quantities in the system. The form of the compensator used in the feedback control system to make it respond satisfactorily to inputs, as well as to possess sufficient stability, would depend on the model for the system being controlled.

At several stages in the control system design process, it is desirable to use a simulation of the control system. Simulation refers in general to the process of numerically representing the equations that describe the system on a digital computer. This allows the designer to analyze the performance of the control system relative to the specifications, and indicates how the design should be modified in order to make its performance more acceptable. Analog computers, although still available for simulation, are now less used than digital computers. Simulation on a computer (analog or digital) permits the system to be analyzed without expending the time and the cost of constructing a physical model, or of testing the system itself. The risk of damage to the equipment by an incorrectly designed compensator is also reduced considerably.

System modeling. When models for physical systems are derived analytically, the equations are developed by applying the appropriate conservation laws. For an electrical system, Kirchhoff's current and voltage laws may be used to model the relationships between variables. For mechanical systems, Newton's second law of motion or Lagrange's equations may be used. An example is the spring-mass damper system of Fig. 7, and the variables and parameters of this simple example. The sum of the applied and reactive forces from the spring, mass, and damper is equal to zero. The reactive forces are given in Eqs. (35)–(37), where

$$F_k = K(y) \quad \text{for spring} \tag{35}$$

$$F_m = M(\ddot{y}) \quad \text{for mass} \tag{36}$$

$$F_f = f(\dot{y}) \quad \text{for damper (wall friction)} \tag{37}$$

the dots over the variables represent differentiation with respect to time. The model for this mechanical system, expressed as a second-order differential equation, becomes, through use of Newton's law, Eq. (16), where $u(t)$ is the total applied force.

Equation (16) is often not convenient in form for control system design. The model of this linear system may be expressed either in state-variable form or as a transfer function. In state-variable form, an nth-order differential equation is expressed as n first-order differential equations. For the model of Equation (16), the state variables $x_1 = y$ and $x_2 = \dot{y}$ and input variable u can be defined. The system model's state equations then become Eq. (38). The output equation for this model is Eq. (39).

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -(K/M) & -(f/M) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ (1/M) \end{bmatrix} u \tag{38}$$

$$y = [\,1 \quad 0\,] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [0]\,u \tag{39}$$

In order to develop the transfer-function model for this system, the Laplace transform is taken of Eq. (16) under the assumption that all initial conditions are zero. This yields Eq. (40). The transfer function is

$$(Ms^2 + fs + K)Y(s) = U(s) \tag{40}$$

defined as the ratio of the Laplace transform of the output to the Laplace transform of the input with

zero initial conditions. Equation (37) thus provides the transfer function, Eq. (41).

$$\frac{Y(s)}{U(s)} = \frac{(1/M)}{s^2 + (f/M)s + (K/M)} \tag{41}$$

Comparison of Eqs. (16), (38), and (41) shows that the coefficients in the various model forms are related to the component parameters in the physical system. For simple systems, such as that in Fig. 7, the state equation or transfer-function models can be developed by inspecting the original differential equation.

In order to model a system experimentally, certain input test signals are chosen and applied to the system, and the output is measured. Perhaps the most common set of input signals for this purpose is composed of sinusoids of various frequencies. Each input signal is of the form given in Eq. (42) for some value

$$u(t) = \sin(\omega t) \tag{42}$$

of angular frequency ω. The input signals are applied singly to the system until a steady-state condition is reached at the output, which will be of the form given in Eq. (43). Therefore, the ratio of the mag-

$$y(t) = A(\omega)\sin[\omega t + \theta(\omega)] \tag{43}$$

nitudes of output to input signal is $A(\omega)$, and the phase shift of the output relative to the input is $\theta(\omega)$. The plots of $A(\omega)$ and $\theta(\omega)$ versus ω are called Bode plots. Generally, $20\log_{10}[A(\omega)]$ is plotted rather than $A(\omega)$, thereby expressing the magnitude ratio in decibels (dB). From known rules of constructing Bode plots, the transfer function for the system can be obtained from the experimental results. To obtain the best plots for developing the transfer function, the frequencies chosen for the input signals should be evenly spaced on a logarithmic scale. **Figure 13** shows the magnitude plot for a particular system. The circles are the values of magnitude ratio that would be determined experimentally if the corresponding sinusoidal frequencies were input to the system. The curve represents the actual magnitude.

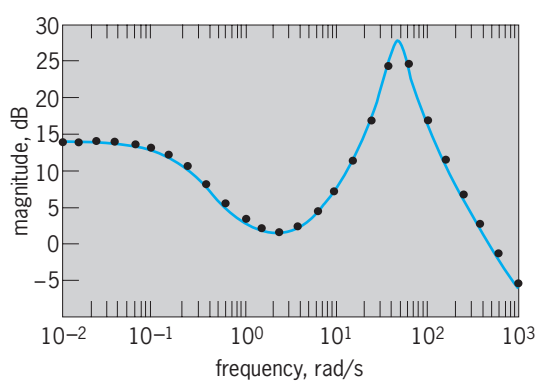

Fig. 13. Frequency domain magnitude for an experimentally determined model.

Equation (44) gives the transfer function for this system.

$$G(s) = \frac{500(s+1)(s+5)}{(s+0.2)(s^2+20s+2500)} = \frac{5(1+s)(1+s/5)}{(1+s/0.2)[1+s/125+(s/50)^2]} \tag{44}$$

The second form of $G(s)$ shows how the transfer function can be obtained from the plot. The dc (zero-frequency) gain in the plot is $14\text{ dB} = 20\log_{10}(5)$. The plot starts to curve down near a pole of the transfer function and curves up near a zero. Underdamped complex poles cause an overshoot in the curve, with a sharper peak for less damping. The quadratic factor in the denominator of Eq. (44) corresponds to an undamped natural frequency $\omega_n = 50$ rad/s and a damping ratio $\zeta = 0.2$. The slopes of the Bode magnitude curve change by integral multiples of ± 20 dB per decade as poles or zeros are encountered.

Care must be taken in determining a system model experimentally. The amplitude of the input signal must be so chosen that the system response remains linear throughout the experiment. Feedback loops must not be broken if the open-loop system has unstable poles.

System simulation. Simulation is used to evaluate the performance of both open- and closed-loop systems. With simulation on a digital computer, either the continuous-time equations forming the mathematical model may be converted into a discrete-time state equation or difference equation, or they may be solved directly by numerical integration. In any case, a sampling period T must be chosen small enough for the important dynamics of the continuous-time system to be reproduced. In converting to a discrete-time system, a rule of thumb is to choose T so that the highest-frequency components of the original system are sampled five to ten times per period. In using numerical integration, T must be so chosen that the products of T with each of the poles of the system all lie inside the region of absolute stability for the particular integration operator used.

If the system model is given in the state-variable form shown in Eqs. (45), then the discrete-time

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \tag{45}$$

model which can be used for simulation is given by Eqs. (46) and (47). Here k takes on values

$$x[(k+1)T] = \Phi x(kT) + \Psi u(kT) \tag{46}$$

$$y(kT) = Cx(kT) + Du(kT) \tag{47}$$

[0, 1, 2, . . .] and the matrices in Eq. (46) are defined by Eqs. (48) assuming that $u(t)$ is constant in

$$\Phi = e^{AT} \qquad \psi = \int_0^T e^{A\tau} B\, d\tau \tag{48}$$

the interval from kT to $(k+1)T$. Equations (46) and (47) can be solved for the output at each discrete point in time kT.

If the simulation is to be performed as an nth-order difference equation, the system model is expressed in the form of a discrete-time transfer function through the use of the Z transform, as shown in Eq. (49). In this form, the sample period T is implicit

$$\frac{Y(z)}{U(z)} = \frac{b_m z^m + b_{m-1} z^{m-1} + \cdots + b_1 z + b_0}{z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0} \tag{49}$$

in the values of the a_i and b_i coefficients. The solution for the output $y(kT)$ is given by Eq. (50). The

$$\begin{aligned} y(kT) = {} & -a_{n-1} y[(k-1)T] - \cdots \\ & - a_1 y[(k-n+1)T] - a_0 y[(k-n)T] \\ & + b_m u[(k+m-n)T] + \cdots \\ & + b_1 u[(k-n+1)T] + b_0 u[(k-n)T] \end{aligned} \tag{50}$$

output is generated recursively as a sum of previous values of the output and present (if $n = m$) and past values of the input.

If the simulation is to be performed by numerical integration, the continuous-time state model of Eq. (45) is used, along with the selected integration operator. Many computer installations have software which can provide highly accurate solutions to linear or nonlinear differential equations. *See* NUMERICAL ANALYSIS; SIMULATION. Guy O. Beale

Bibliography. J. J. D'Azzo and C. H. Houpis, *Linear Control System Analysis and Design*, 4th ed., McGraw-Hill, 1995; C. M. Close and D. K. Frederick, *Modeling and Analysis of Dynamic Systems*, 2d ed., Wiley, 1995; R. C. Dorf and R. H. Bishop, *Modern Control Systems*, 8th ed., Addison-Wesley, 1999; E. Eyman, *Modeling, Simulation, and Control*, 1988; T. T. Hartley, G. O. Beale, and S. P. Chicatelli, *Digital Simulation of Dynamic Systems: A Control Theory Approach*, Prentice Hall, 1994; B. C. Kuo, *Automatic Control Systems*, 7th ed., Prentice Hall, 1995; K. Ogata, *Modern Control Engineering*, 3d ed., Prentice Hall, 1997; C. L. Phillips and R. D. Harbor, *Feedback Control Systems*, 3d ed., Prentice Hall, 1995; J. M. Smith, *Mathematical Modeling and Digital Simulation for Engineers and Scientists*, 2d ed., 1987.

Controlled rectifier

A three-terminal semiconductor junction device with four regions of alternating conductivity type (*pnpn*), also called a thyristor. This switching device has a characteristic such that, once it conducts, the voltage in the circuit in which it conducts must drop below a threshold before the controlled rectifier regains control. Such devices are useful as high-current switches and may be used to drive electromagnets and relays.

The principle of operation can be understood by referring to the **illustration**. The central junction is reverse-biased (positive collector, grounded emitter). The wide n region between collector and base regions prevents holes injected at the collector junction from reaching the collector-to-base barrier by

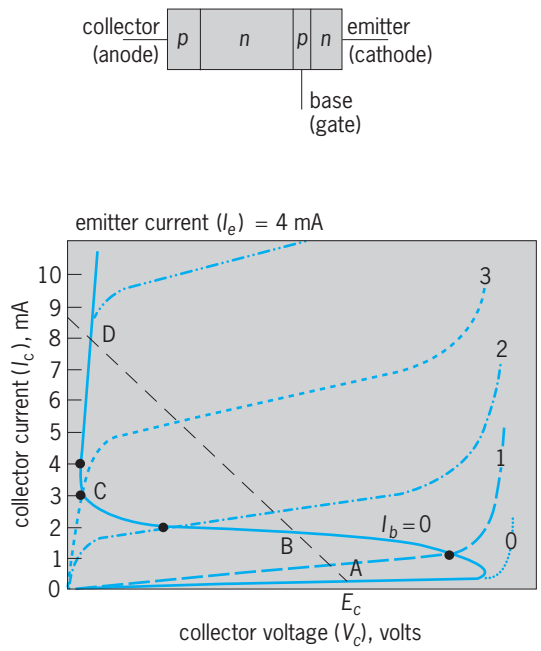

Controlled rectifier diagram and characteristics.

diffusion. The junction between emitter and base is the emitter. When operated as a normal transistor, this device shows a rapid increase of current gain of α (equal to I_c/I_e) with collector current. This effect may be due to a field-induced increase of transport efficiency across the floating n region, or to increased avalanching in the high-field barrier region, or to increased injection efficiency at the two forward-biased junctions, or to a combination of these phenomena.

With a floating-base region ($I_b = 0$), the device is a two-terminal device and collector current I_c must equal emitter current I_e.

By selecting the points on the illustrated characteristics where $I_c = I_e$, the characteristic for the base current $I_b = 0$ is shown as the heavy curve with the negative-resistance characteristic. This characteristic will be found in any transistor which shows an integrated α increasing from below unity at low collector currents to above unity at high collector currents.

If this device is operated as a three-terminal device, the switching between the nonconducting and conducting states can be controlled by the base. If in the grounded-emitter case the collector is biased to $+E_c$ as shown, the device will remain at point A until the base is pulsed positive by at least enough current to carry the emitter current to point B. At this point α exceeds unity and the device will spontaneously switch to point D in the conducting state. To reset the unit, either the emitter must be cut off or the collector voltage must be reduced so that the load line falls below the valley point C. The current of point C is called the holding current. Either of these results can be achieved by appropriate pulses on the base. Modern terminology usually refers to the rectifier terminals as anode and cathode and to the control terminal as the gate. *See* SEMICONDUCTOR RECTIFIER.

Overall current gain α may be maintained below unity for low anode-cathode currents by designing the junction between the anode and the floating *n* region and the junction between the gate and the cathode so each has a low-current injection efficiency below 0.5. For further discussion of four-layer devices *see* TRANSISTOR

Lloyd P. Hunter

Bibliography. J. Singh, *Semiconductor Devices: An Introduction*, 1994; E. S. Yang, *Microelectronic Devices*, 1988.

Controlled/living radical polymerization

A synthetic method which avoids the premature termination of growing radicals and produces polymers with precise compositions, architectures, and functionalities.

Free-radical polymerization is used extensively in industry for the synthesis of a variety of polymeric materials. Its widespread use is due to its compatibility with many functional groups, its tolerance toward water and acidic media, and its simplicity. Free-radical polymerization routinely is done under various conditions and processes, including bulk, solution, or aqueous environments, and is now the leading industrial method for producing vinyl polymers and copolymers. *See* COPOLYMER; FREE RADICAL; POLYMER; POLYMERIZATION; POLYVINYL RESINS.

In spite of its utility, free-radical polymerization is considered unsuitable for the synthesis of polymers with precise compositions, architectures, and functionalities. This is because of the very limited lifetime of the propagating free-radical species (typically 0.1 to 10 s), which suffer from the propensity to undergo diffusion-controlled termination by recombination or disproportionation (**Fig. 1**), unlike growing ionic species which inherently repel each other. The polymers obtained by free-radical means

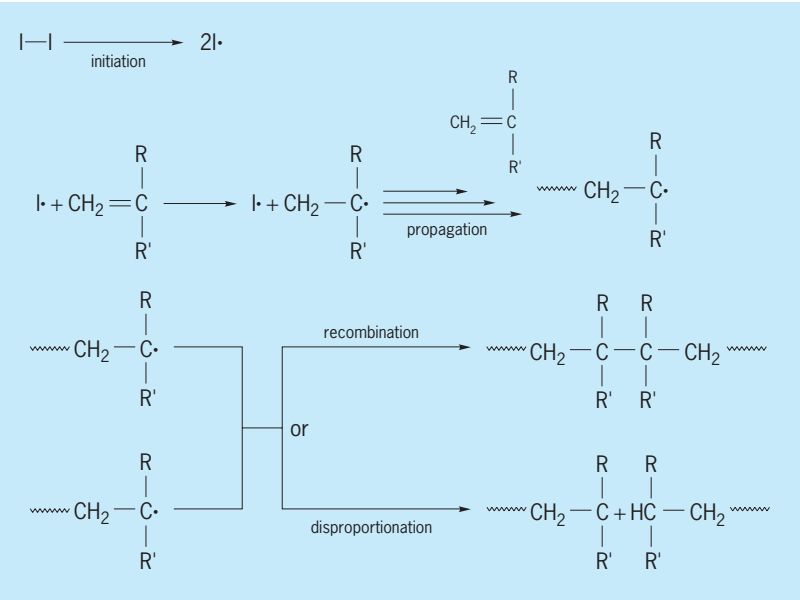

Fig. 1. Free-radical polymerization.

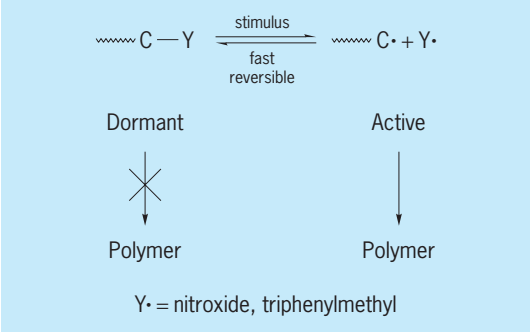

Fig. 2. Active and dormant radical equilibrium.

are therefore generally ill defined and polydisperse (wide molecular-weight distribution), with only limited control over molecular weight or architecture.

Until recently, the only option for efficiently controlling the structure and architecture of vinyl polymers was ionic or coordination polymerizations, which do not suffer from irreversible chain-terminating (deactivating) reactions and are known as "living" polymerizations. Unfortunately, living ionic polymerizations are of limited scope since they require stringent purity conditions as well as the total absence of water and oxygen, and are generally incompatible with monomers with functional groups. The necessity to overcome these limitations fueled the development of a new polymerization procedure that would exhibit all characteristic features of regular free-radical polymerization, as well as "living" character, similar to ionic polymerizations.

When addressing this challenge, the primary difficulty was to find a way to avoid the premature termination of the growing radicals. The idea that eventually predominated was to lower the instantaneous concentration of growing radicals by trapping them in a large reservoir of covalent, dormant species and establish a fast equilibrium between the latter and a minute concentration of active, growing radical species. If such a dynamic equilibrium could be established, the probability for spontaneous bimolecular radical termination would be minimized, and all chains would be given an equal opportunity to grow through constant interconversion between their active and dormant forms (**Fig. 2**). Another interesting outcome would be a nearly uniform molecular weight, as determined by the ratio of the concentrations of converted monomer to that of the dormant species (or the initiator). First conceptualized by T. Otsu in 1982, this seminal idea has given birth to the field of controlled/living radical polymerization, which has grown explosively.

Characterized by different mechanisms and chemicals involved, three main methods (by far the most studied) of control of radical polymerization are (1) nitroxide-mediated radical polymerization (NMP), (2) atom transfer radical polymerization (ATRP), and (3) reversible addition-fragmentation chain transfer (RAFT) polymerization.

Nitroxide-mediated radical polymerization. Nitroxides are stable free radicals that are generally

used in organic chemistry to trap transient radicals. They react at diffusion-controlled rates with carbon-centered radicals and are known not to initiate polymerization. Taking advantage of these features, M. K. Georges and coworkers used the commercially available nitroxide, 2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO), to trap growing polymer chains and demonstrated that the trapping process is reversible when the temperature range is appropriately chosen. The mechanism of control involves the reversible termination of growing radicals by TEMPO, which functions as a counter radical. Samples of polystyrene exhibiting narrow molecular-weight distributions could be obtained by heating styrene at 123°C (253°F) in the presence of TEMPO and a free-radical initiator. However, the reversible radical capping by TEMPO could not be used for most monomers other than styrene, since most other vinyl monomers showed a total incompatibility with this kind of nitroxide. Anticipating that the structure of the stable radical used to trap the growing radicals must play a crucial role in the success or failure of such living radical polymerization, Y. Gnanou and P. Tordo designed a more effective nitroxide to extend the scope of NMP. They developed a α-hydrogen-bearing nitroxide, namely *N-tert*-butyl-*N*-[1-diethylphosphono-(2,2-dimethylpropyl)]nitroxide [DEPN], which is of slightly lower stability than TEMPO and whose main characteristics is to give alkoxyamines with much weaker C—O bonds than TEMPO. Not only could styrene be polymerized at faster rates and under living conditions in the presence of DEPN and a radical initiator, but also the polymerization of acrylates, acrylamides, and more recently methacrylates were found to be living in the presence of DEPN (**Fig. 3**). *See* POLYACRYLATE RESIN; POLYSTYRENE.

In a similar approach, C. Hawker and coworkers developed unimolecular alkoxyamines based on 2,2,5-trimethyl-4-phenyl-3-azahexane-3-nitroxide (TIPNO), another α-hydrogen-bearing nitroxide, and used these as initiators and controlling agents for the living polymerization of the above monomers, as well as for acrylonitrile, butadiene, and so others. The beauty of nitroxide-mediated radical polymerization lies in the fact that the amount of terminated chains eventually formed is exactly equal to the excess nitroxide produced by the system to minimize termination reactions. H. Fischer conceptualized this remarkable phenomenon and called it the persistent radical effect. To minimize the amount of terminated chains to a level lower than that determined by the system, the nitroxide can be used in excess with respect to the initiator.

Atom transfer radical polymerization (ATRP). Discovered by K. Matyjaszewski in 1995, atom transfer radical polymerization entails a reversible one-electron oxidation of a transition metal, abstracting a halogen from an organic halide (RX), thus generating a radical species ($R^{\bullet}$) that can attack an unsaturated compound ($CH_2{=}CHR'$) and further propagate ($R{-}[CH_2{-}CHR']_n{-}CH_2{-}C^{\circ}HR'$), followed by a one-electron reduction of the now oxidized metal, transferring its halogen to the growing chain ($R{-}[CH_2{-}CHR']_n{-}CH_2{-}C^{\circ}HR'$), and thus producing ($R{-}[CH_2{-}CHR']_n{-}CH_2{-}CHR'{-}X$) as a dormant species. Such halide abstraction (k_{act}) and subsequent transfer (k_{deact}) occur repeatedly, and polymer chains grow between each activation/deactivation cycle (**Fig. 4**).

Provided the latter process is rapid, all the chains can grow uniformly. The oxidized metal complexes $X\text{-}M_t^{n+1}$ generated by activation of the dormant alkyl halides behave like persistent radicals and therefore reduce the stationary concentration of growing radicals, minimizing the extent of termination. Like nitroxide-mediated radical polymerization, atom transfer radical polymerization is under the control of the persistent radical effect, which means that a successful atom transfer radical polymerization contains a small amount of terminated chains. Also in 1995, M. Sawamoto independently described metal-catalyzed living radical polymerization (MCRP), which appeared to exhibit the same

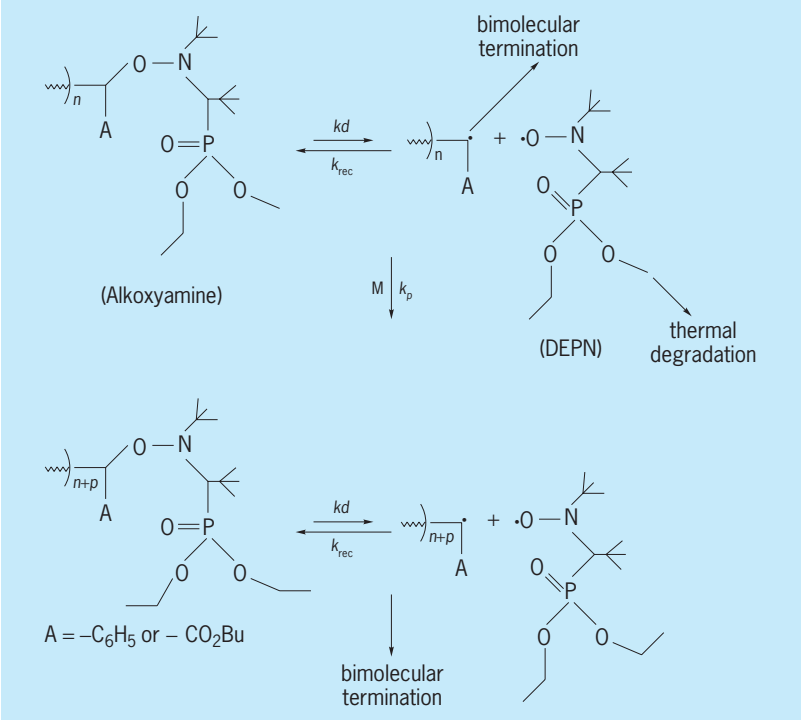

Fig. 3. Nitroxide-mediated radical polymerization of styrene or butylacrylate.

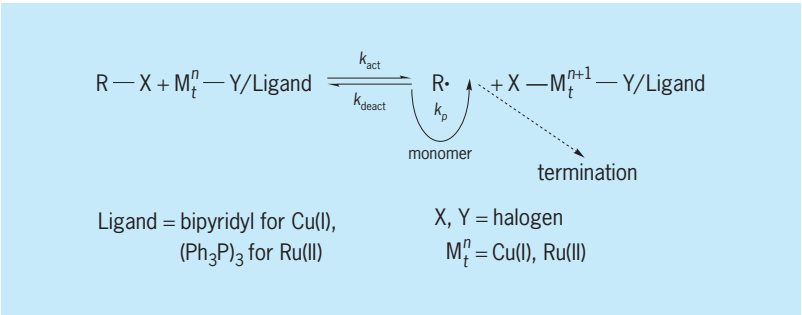

Fig. 4. Atom transfer radical polymerization.

mechanism as that of atom transfer radical polymerization. The initiating systems for both methodologies consist of a transition-metal/ligand complex and an initiator (typically an activated alkyl halide), the role of the former being to generate radical species from the latter by the above mechanism. In choosing the appropriate metal/ligand catalyst for controlling the polymerization, four criteria are critical: (1) the metal center must have at least two accessible oxidation states separated by one electron, (2) it must exhibit affinity toward halogen, (3) the coordination sphere around the metal must be able to accommodate the halogen upon oxidation, and (4) the ligand must form a strong complex with the metal. On these grounds, Matyjaszewski proposed copper(I)bromide associated with a variety of nitrogen-based ligands as the atom transfer radical polymerization catalysts, and Sawamoto associated Ru(II) with phosphine and other ligands as metal-catalyzed living radical polymerization catalysts. As demonstrated by its extensive use in the synthesis of polymers with precise composition and architecture, atom transfer radical polymerization is superior to metal-catalyzed living radical polymerization by its versatility and cost. *See* LIGAND.

Reversible addition-fragmentation chain transfer (RAFT) polymerization. In 1998, E. Rizzardo described a new method for controlling free-radical polymerization that involved the use of thiocarbonylthio compounds reacting by reversible addition-fragmentation transfer (RAFT) with the growing radicals. The effectiveness of these RAFT reagents, mirrored in chain-transfer constants Ctr larger than 1, depends on the nature of their (Z) and (R) groups, Z determining their reactivity toward entering radicals and R its aptitude for fragmentation. The key step in a successful RAFT process is the chain transfer of growing radicals to the RAFT reagent and the subsequent formation of the intermediate radical. The latter then undergoes fragmentation, releasing a radical (R•) which can initiate polymerization. The entering polymer chain now carries a terminal [S═C(Z)S—] moiety and is transformed into a dormant species. After total consumption of the thiocarbonylthio compound, an equilibrium is established between the dormant thiocarbonylthio-carrying chains and active ones, producing a polymer with narrow polydispersity (**Fig. 5**), for large C_{tr} values of the thiocarbonylthio compound.

The molecular weight is given by the ratio of monomer to thiocarbonylthio compound, and the amount of terminated chains correspond to that of the radical initiator used to trigger the polymerization. The main advantage of RAFT over nitroxide-mediated radical polymerization and atom transfer radical polymerization is its applicability to a wider range of monomers, including functional ones such as acrylic acid or those carrying an electron-donating substituent such as vinyl acetate. D. Charmot of Rhodia Recherches, France, showed that xanthates, which are also thiocarbonylthio compounds, are excellent RAFT reagents for vinyl acetate and analogues, and termed its invention MADIX for macromolecular design via interchange of xanthates.

Application. The three methodologies of controlling free-radical polymerization have their merits and limitations. ATRP is easy to carry out but it cannot be used with certain functional monomers and requires a tedious metal-removal step, whereas RAFT or NMP can be applied to functional and nonfunctional monomers but their controlling agents may necessitate complex synthesis.

With the advent of NMP, ATRP, and RAFT, novel polymeric materials—many of which are totally original and exhibit precise composition such as block and gradient copolymers and complex topology such as star-, comb-, and dendrimer-like polymers—can be obtained under the conditions of regular free-radical polymerization. This paves the way for tremendous opportunities in bio- and nanotechnologies and other high-tech areas. Yves Gnanou; Daniel Taton

Bibliography. D. Benoit et al., Kinetics and mechanisms of controlled free-radical polymerization of styrene and *n*-butyl acrylate in the presence of an acyclic β-phosphonylated nitroxide, *J. Amer. Chem. Soc.*, 122:5929–5939, 2000; J. Chefari et al., Living free-radical polymerization by reversible addition-fragmentation chain transfer: The RAFT process, *Macromolecules*, 31(16):5559–5562, 1998; C. Hawker, A. Bosman, and E. Harth, New polymer synthesis by nitroxide mediated living radical polymerization, *Chem. Rev.*, 101:3661–3688, 2001; K. Matyjaszewski and T. P. Davis (eds.), *Handbook of Radical Polymerization*, Wiley-Interscience, 2002; K. Matyjaszewski and J. Xia, Atom transfer radical polymerization, *Chem. Rev.*, 101:2921–2990, 2001.

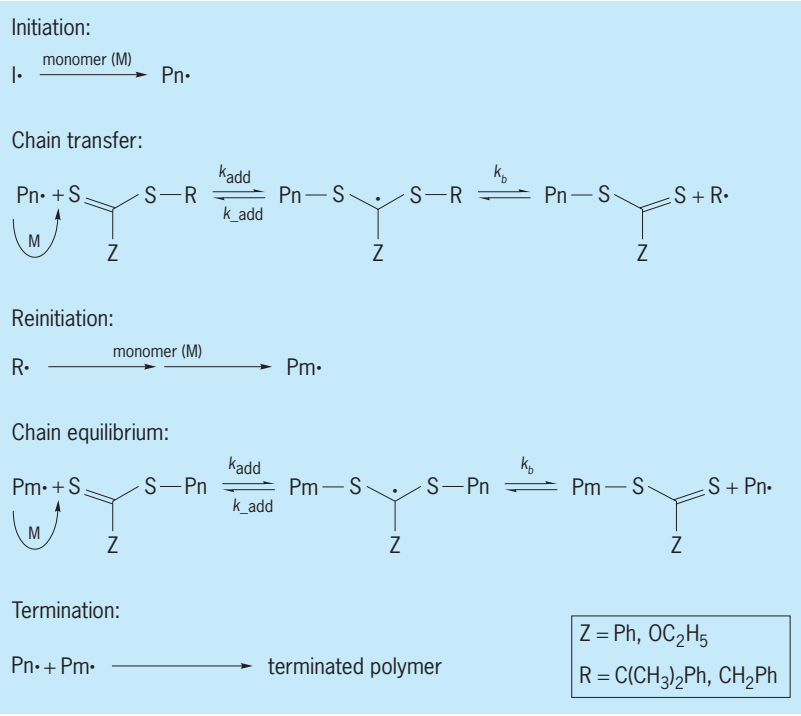

Fig. 5. RAFT polymerization.

Conulariida

A group of extinct invertebrate organisms normally having an elongate, four-sided pyramidal exoskeleton of calcium phosphate composition. Due to poor preservation of key biological structures, the evolutionary affinities and classification of conulariids are uncertain; however, they are often regarded as an unusual group of cnidarians.

Morphology. The conulariid exoskeleton has an aperture at the widest end, and tapers to a closed, slightly rounded apex. The apex is rarely observed in fossils (**Fig. 1***a*). Usually the apical end of the exoskeleton is broken, revealing an internal, slightly bowed partition called the apical wall. Specimens typically range in size from 1 to 30 cm (0.4 to 12 in.). Externally, each of the four sides or faces is crossed by numerous pairs of transverse, curved ridges that meet at a distinct longitudinal midline. On the inside of the exoskeleton, the midline may be represented by a carina, or ridge. Ridges are formed through thickenings of phosphatic skeletal layers (sometimes called rods), and they probably conferred structural strength to an otherwise rather flexible integument. The external surface commonly shows other features such as nodes and small crests. Adjacent sides meet at a longitudinal groove called a corner groove.

Conulariids were evidently attached epibenthic (ocean-bottom-dwelling) organisms. In rare circumstances, the attachment structure, an elongate organic (probably proteinaceous) stalk (Fig. 1*b*), is preserved. The stalk sheaths the apical end of the exoskeleton.

The conulariid exoskeleton was pliable. In shales, they are often found compacted, and in storm bed deposits, they are sometimes oddly bent. Often the faces are bent inward at the aperture (**Fig. 2***b*), and this is interpreted as a taphonomic (postmortem) phenomenon.

Classification. Conulariids are problematic fossils and comprise more than 400 species that are assigned to about 50 genera. Not all these species and genera are undisputed members of the group, however. Because of the rather simple pyramidal or conical morphology of conulariids, it is possible that some fossils have been erroneously categorized. In general, the evolutionary affinities of conulariids and their position within the Linnaean classification system are uncertain because key biological structures were not preserved as fossils. Because of their skeletal shape or composition, at various times they have been classified as worms, mollusks, cnidarians, chordate relatives, and representatives of an extinct phylum. The majority opinion is that these fossils represent an unusual group of cnidarians. The main characters supporting this view include a four-sided symmetry and the presence of a carina along the inside of each side's midline that resembles one of the four interradial septa formed by the ectoderm that divides the enteron (alimentary canal) of a scyphozoan (jellyfish). The pattern of ridge distribution and the phosphatic composition of the conulariid exoskeleton tend to argue against a cnidarian affinity, but not unequivocally. Preserved soft parts of conulariids include an elongate gut tract and various internal ovoid-shaped organs but do not include tentacles. *See* CNIDARIA.

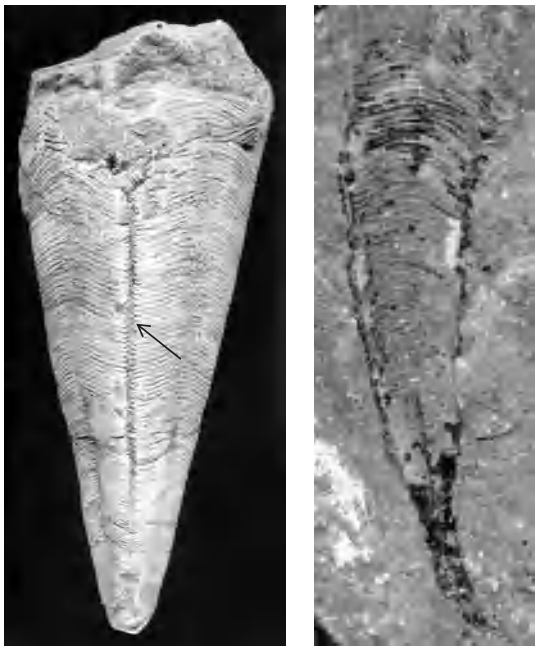

(a) (b)

Fig. 1. Conulariid morphology. (*a*) *Paraconularia byblis*, a well-preserved specimen in siderite (iron carbonate) from the Carboniferous (Mississippian) of Ohio showing two adjacent sides meeting at a deep corner groove (arrow). Ridges along each side meet at a subtle longitudinal midline. Length of specimen: 13 cm (5.1 in.). (*b*) Specimen of *Paraconularia chesterensis* from the Carboniferous (Mississippian) of British Columbia showing attachment stalk (black) extending from the apical end. Length of specimen: 26 mm (1 in.).

Geologic occurrence, paleoecology, and taphonomy. Conulariid fossils are known from marine strata of all continents except Antarctica. In most shale, limestone, and sandstone deposits, articulated conulariids are uncommon to rare. There are notable exceptions, however. Dark, organic-rich shales representing low-oxygen (dysaerobic or exaerobic) environments in the Carboniferous of Kansas and Missouri yield hundreds of specimens per square meter along some bedding planes. Concretions composed of siderite (iron carbonate) in the Carboniferous of Ohio and Kentucky, as well as concretions composed of phosphate in the Carboniferous of Kansas and silica in the Devonian of Bolivia, also yield large numbers of specimens. Normally, deposits rich in conulariid remains have biotas of low overall diversity. Clusters of conulariids (Fig. 2*a*) are a common occurrence. In such circumstances, individuals are usually of approximately the same size. This suggests that the clusters originated from either a spat-fall of larvae or perhaps from asexual budding of a single progenitor.

Following death, conulariid exoskeletons disarticulated into small pieces of integument or even

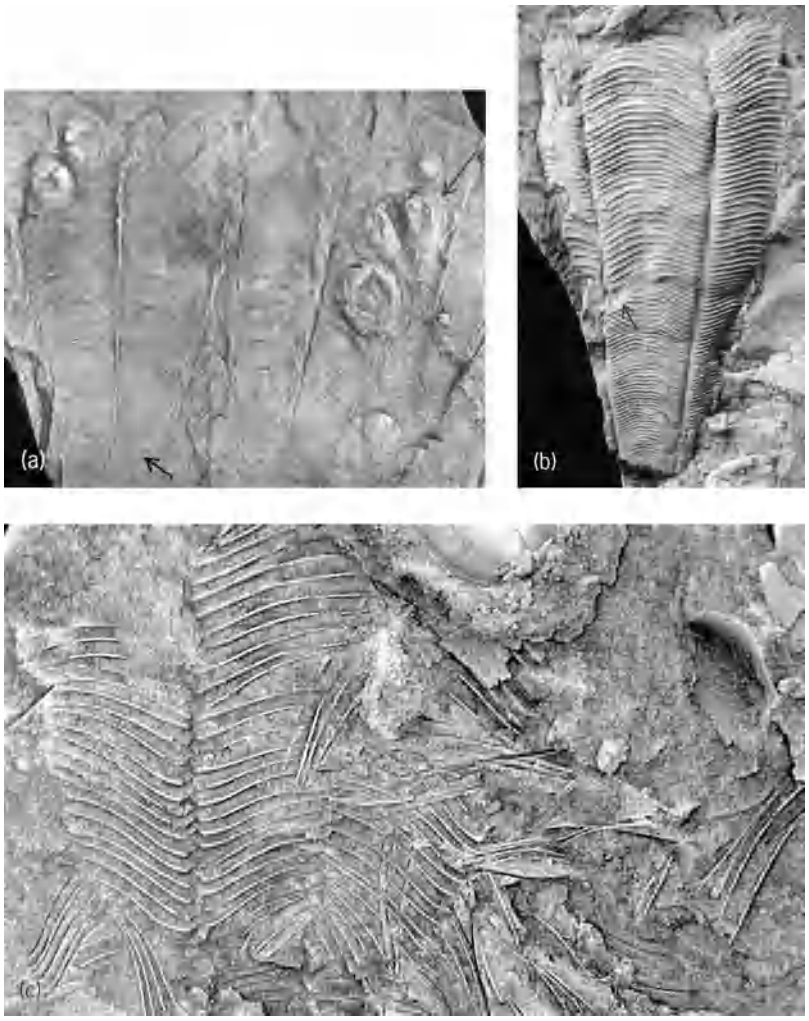

Fig. 2. Conulariid paleoecology and taphonomy. (*a*) Cluster of *Conularia congregata* compacted in shale from the Devonian of New York. The specimens show apical walls (lower arrow), indicating detachment from, or nonpreservation of, their attachment stalks. The conulariids are encrusted by round shells of inarticulate brachiopods (upper arrow). Length of large specimen at left: 6 cm (2.4 in.). (*b*) *Paraconularia chesterensis*, a specimen compacted in shale from the Carboniferous (Mississippian) of Indiana showing one complete, exposed face flanked by two partly exposed sides. Pairs of ridges meeting at the longitudinal midline are exposed on the complete face. At the aperture, the faces have been folded inwardly. A healed bite mark, marked by broken ridges and scar tissue, is present about midway down the left side of the specimen (arrow). Length of specimen: 8 cm (3.1 in.). (*c*) *Paraconularia subulata* from the Carboniferous (Mississippian) of Ohio in mostly disarticulated condition showing fragments of integument (especially at left) and thin, scattered rods. Average length of rods: 3.6 mm (0.14 in.).

into individual rods (Fig. 2*c*). Injuries, both sublethal (Fig. 2*b*) and lethal, inflicted by carnivores such as cephalopods and sharks, have been recorded in conulariids, sometimes in large numbers of specimens from an individual locality. Predation must have led to the breakage of conulariid exoskeletons and may account, in part, for the low numbers of articulated specimens in numerous deposits.

Conulariid skeletons, both in life and afterward, often served as attachment sites for other marine organisms. Inarticulate brachiopods (Fig. 2*a*), edrioasteroid echinoderms, sponges, and other animals have been observed attached to conulariid exoskeletons. *See* FOSSIL; PALEOECOLOGY; TAPHONOMY.

Evolutionary history. Conulariids are a relatively small group that have a stratigraphic range through much of the Paleozoic and the lowermost Mesozoic. Small phosphatic fossils that are usually classified as conulariids occur in the lower part of the Cambrian. These fossils, which include hexaconulariids and hexangulaconulariids, if they are phylogenetically related to Ordovician and younger conulariids, differ from these later examples in having small size (typically less than 1 cm in length) and in having a variable number of faces, instead of just four sides. Some fossils from the Neoproterozoic (Ediacaran Period) have been questionably interpreted as close relatives of conulariids, and even as possible ancestral taxa.

Undisputed conulariids occur in the upper part of the Cambrian but do not become a significant part of fossil assemblages until the Middle Ordovician. Peak diversity was apparently reached in the Devonian through Carboniferous. Conulariids became nearly extinct at the end of the Permian, and barely recovered in the Early Mesozoic. They are represented in the Triassic by only one known genus. *See* ANIMAL EVOLUTION; GEOLOGIC TIME SCALE.

Loren E. Babcock

Bibliography. L. E. Babcock and R. M. Feldmann, *Ann. Carnegie Muse.*, 55:349–410, 411–479, 1986; L. E. Babcock et al., *Nat. Geog. Res.*, 3:210–231, 1987; R. H. Mapes, T. R. Fahrer, and L. E. Babcock, *J. Paleontol.*, 63:34–37, 1989; R. C. Moore and H. J. Harrington, in R. C Moore (ed.), *Treatise on Invertebrate Paleontology*, Pt. F: *Coelenterata*, Geological Society of America and University of Kansas Press, New York and Lawrence, Kansas, 1956; J. G. Richardson and L. E. Babcock, *J. Paleontol.*, 76:391–393, 2002; H. Van Iten, *Palaeontology*, vol. 35, pp. 335–358, 359–372, 1992.

Convection (heat)

The transfer of thermal energy by actual physical movement from one location to another of a substance in which thermal energy is stored. A familiar example is the free or forced movement of warm air throughout a room to provide heating. Technically, convection denotes the nonradiant heat exchange between a surface and a fluid flowing over it. Although heat flow by conduction also occurs in this process, the controlling feature is the energy transfer by flow of the fluid—hence the name convection. Convection is one of the three basic methods of heat transfer, the other two being conduction and radiation. *See* CONDUCTION (HEAT); HEAT RADIATION; HEAT TRANSFER.

Natural convection. This mode of energy transfer is exemplified by the cooling of a vertical surface in a large quiescent body of air at temperature t_∞. The lower-density air next to a hot vertical surface moves upward because of the buoyant force of the higher-density cool air farther away from the surface

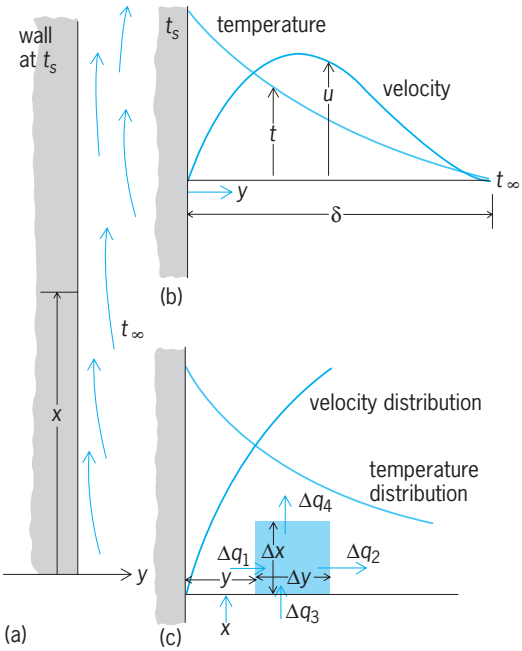

Fig. 1. Temperature and velocity distributions in air near a heated vertical surface. (*a*) Upward movement of hot air. (*b*) Distributions at arbitrary vertical location. The distance δ is that distance at which the velocity and the temperature reach ambient surrounding conditions. (*c*) Distributions in boundary layer.

(**Fig. 1***a*). At any arbitrary vertical location x, the velocity parallel to the surface will increase from zero at the surface to a maximum, and then decrease to zero as ambient surrounding conditions are reached (Fig. 1*b*). In contrast, the temperature of the air decreases from the heated wall value to the surrounding air temperature. These temperature and velocity distributions are clearly interrelated, and the distances from the wall through which they exist are coincident because, when the temperature approaches that of the surrounding air, the density difference causing the upward flow approaches zero.

The region in which these velocity and temperature changes occur is called the boundary layer. Because velocity and temperature gradients both approach zero at the outer edge, there will be no heat flow out of the boundary layer by conduction or convection. *See* BOUNDARY-LAYER FLOW.

An accounting of all of the energy streams entering and leaving a small volume in the boundary layer (Fig. 1*c*) during steady-state conditions yields Eq. (1),

$$\kappa\left(\frac{\partial^2 t}{\partial x^2}+\frac{\partial^2 t}{\partial y^2}\right)=\rho c_p\left(u\frac{\partial t}{\partial x}+v\frac{\partial t}{\partial y}\right) \qquad (1)$$

where κ, ρ, and c_p are the thermal conductivity, the mass density, and specific heat at constant pressure of the air, and u and v are the velocity components in the x and y directions, respectively. Equation (1) states that the net energy conducted into the element equals the increase in energy of the fluid leaving (convected) over what it had entering.

Very close to the surface, the velocity components u and v both approach 0, and the partial derivatives $\partial^2 t/\partial x^2$ and $\partial t/\partial x$ also approach 0. Therefore, the heat transfer from the wall must be by conduction through a thin layer of air which is effectively stationary relative to the surface. Thus, Eq. (2) is obtained,

$$q=-\kappa\left.\frac{\partial t}{\partial y}\right|_{y\to 0} \qquad (2)$$

where q is the time rate of heat flow per unit of surface area through an infinitesimally thin layer of air. At a very small distance from the surface, the velocity becomes finite and some of the energy conducted normal to the surface is convected parallel to it. This process causes the temperature gradient to decrease, eventually to zero. The solution of a problem requires simultaneous determination of the velocity and temperature distributions throughout the boundary layer. From this, the temperature gradient at the wall and the rate of heat flow can be computed.

The effect of energy leaving a surface and remaining in the boundary layer is (1) a gradual increase in temperature of the air in this layer as it moves upward, and (2) diffusion of energy farther from the surface, entraining more air in (thickening) the boundary layer. This effect can be visualized with interference photographs (**Figs. 2** and **3***a*), where the interference fringes indicate lines of constant temperature (isotherms).

Forced convection. If air is blown across a cylinder (Fig. 3*b*), the boundary layer on the forward half of the cylinder will become so thin that it is not possible to resolve the isotherms within it. Although the natural convection forces are still present in this latter case, they are clearly negligible compared with the imposed forces. The process of energy transfer from the heated surface to the air is not, however, different from that described for natural convection. The major distinguishing feature is that the maximum fluid velocity is at the outer edge of the boundary layer. This difference in velocity profile and the higher velocities provide more fluid near the surface to carry along the heat conducted normal to the surface. Consequently, boundary layers are very thin.

The properties of a fluid which influence its heat-convecting ability are the dynamic viscosity μ, heat capacity at constant pressure c_p, and thermal conductivity κ. These combine in a single significant property in the form $\mu c_p/\kappa = \mu/(\kappa/c_p)$, which is the ratio of the fluid viscosity to the quotient of its heat-conducting and heat-storage capacities. With proper units, this ratio is dimensionless and is called the Prandtl number, N_{Pr}. For air N_{Pr} is approximately 1. Fluids such as oils have values greater than 100. Liquid metals, for which $N_{\mathrm{Pr}} \sim 0.01$, are particularly effective for convective heat-transfer applications. *See* FLUIDS.

Heat-transfer coefficient. The convective heat-transfer coefficient h is a unit conductance used for

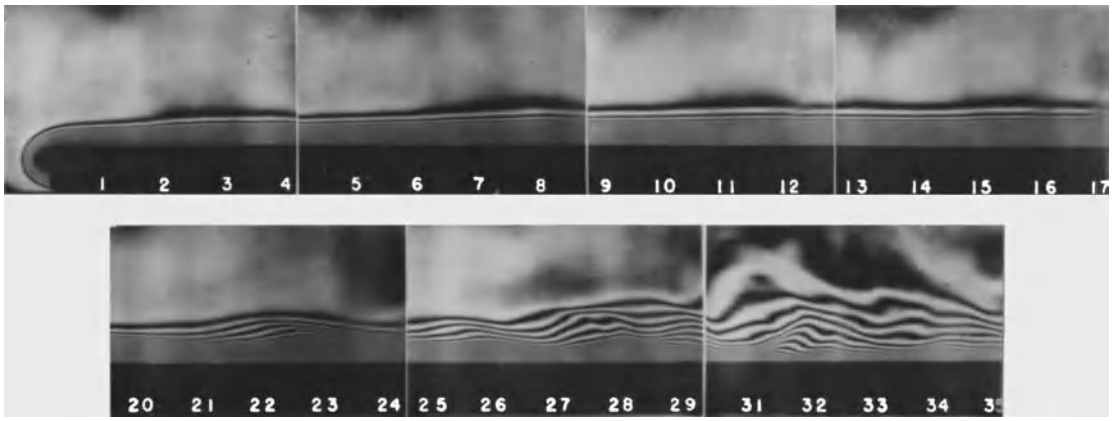

Fig. 2. Laminar and turbulent convection along a vertical plate, as revealed by interference photographs. Interference fringes indicate lines of constant temperature (isotherms). (*Courtesy of E. R. G. Eckert and E. Soehngen*)

calculation of convection heat transfer. It was introduced by Isaac Newton and, until the mechanism of convection was properly interpreted, was thought to be a characteristic of the fluid flowing. To describe quantitatively the cooling of objects in air, Newton suggested Eq. (3). This equation, known

$$q = hA(t_{\text{surface}} - t_{\text{fluid}}) \tag{3}$$

as Newton's law of cooling, is really a definition of h. By comparison with Eq. (2), this coefficient can be seen to be proportional to the slope of the fluid temperature distribution right at the surface, and the thermal conductivity κ of the fluid. Metric units are W/(m^2)(°C); common engineering units are Btu/(h)(ft^2)(°F).

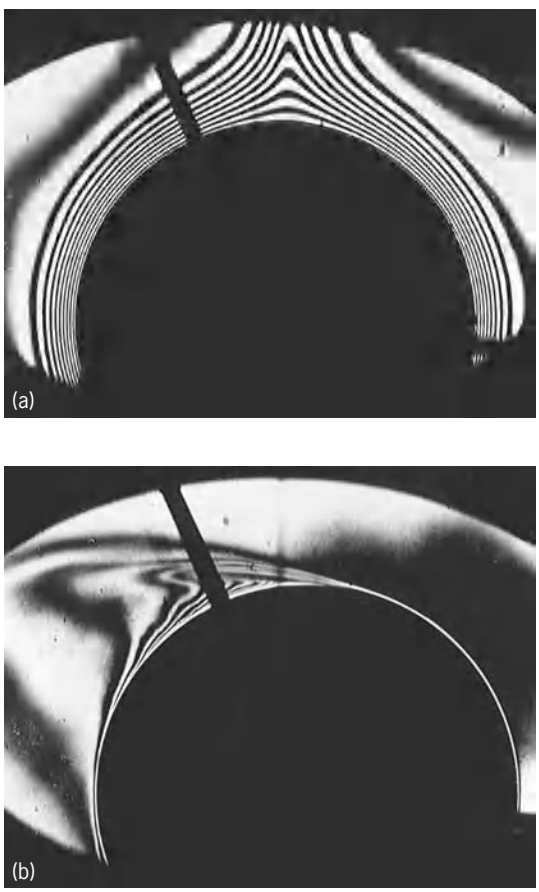

Fig. 3. Interference photographs of isotherms around a heated horizontal cylinder 4 in. (10 cm) in diameter. (*a*) Cylinder in still air. The outer broad fringe indicates the approximate edge of the boundary layer, which is about $^3/_4$ in. (2 cm). (*b*) Cylinder normal to a 33-ft/s (10-m/s) airstream. The boundary layer on the forward half of the cylinder is so thin that the isotherms cannot be resolved. (*Courtesy of D. L. Doughty and W. H. Giedt*)

Local and average coefficients. As the fluid in the heated or cooled boundary layer moves along an isothermal surface, it gradually approaches the temperature of the surface. This causes the temperature gradient in the fluid at the surface (and the rate of heat transfer) to decrease in the direction of flow. Taking, for example, an airstream at 104°F (40°C) moving at 50 ft/s (15 m/s) over a flat plate at 68°F (20°C), the local heat-transfer coefficient h_x decreases in 1 ft (0.3 m) to about one-fifth of its leading-edge value. For practical calculations, an average heat-transfer coefficient h is more useful. It is obtained by integrating h_x over the heat-transfer surface and dividing by the surface area. For this system, h is a function of the plate length l, κ, μ, ρ, and u_∞ of the fluid. For comparing geometrically similar systems involving different fluids, dimensional analysis shows that the specific properties can be combined into dimensionless parameters, conveniently reducing the number of independent variables. For the flat plate, these are the Nusselt number, $N_{\text{Nu}} = hl/\kappa$, the Reynolds number, $N_{\text{Re}} = lu_\infty\rho/\mu$, and the Prandtl number, N_{Pr}. For example, it can be shown that in this case Eq. (4) holds. In the case of free convection, the

$$N_{\text{Nu}} = 0.664 N_{\text{Pr}}^{1/3}\sqrt{N_{\text{Re}}} \tag{4}$$

Grashof number, $N_{\text{Gr}} = \beta g l^3\ \Delta t/\nu^2$ (where β is the coefficient of thermal expansion, and $\nu = \mu/\rho$), replaces the Reynolds number. The product of N_{Gr} and N_{Pr} is known as the Rayleigh number (N_{Ra}), and is used in correlating free convection transfer. Other dimensionless numbers pertinent to convection

include the Stanton number, $N_{St} = h/\rho u c_p = N_{Nu}/N_{Re}N_{Pr}$, which is related to the skin-friction coefficient. *See* DIMENSIONAL ANALYSIS; REYNOLDS NUMBER.

Turbulent flow. Heat convection in turbulent flow is interpreted similarly to that in laminar flow, which has been implied in the above discussion. Rates of heat transfer are higher for comparable velocities, however, because the fluctuating velocity components of the fluid in a turbulent flow stream provide a macroscopic exchange mechanism which greatly increases the transport of energy normal to the main flow direction. For example, if free-stream turbulence caused the boundary-layer flow in the example above to be turbulent, the heat-transfer rate would be about 2.5 times higher. Because of the complexity of this type of flow, most of the information regarding heat transfer has been obtained experimentally. Such results, combined with dimensional analysis, have yielded useful design equations, typical of which is Eq. (5), for predicting the heat transfer in a pipe,

$$\frac{hd}{\kappa} = 0.023\left(\frac{du}{\nu}\right)^{0.8}\left(\frac{\mu c_p}{\kappa}\right)^{0.4} \qquad (5)$$

where d is the pipe diameter.

An illustration of the controlling effect of fluid flow on convective heat transfer is provided by measurements of the local variation around a bluff body such as a cylinder normal to an airstream of velocity u_∞. Major changes occur in the streamline pattern as u_∞ is increased from very low to high values (**Fig. 4**). At very low velocities the flow at the edge of the boundary layer is slowed to zero as it approaches the forward stagnation point. It accelarates as it flows around the cylinder to a maximum at an angle from the stagnation point $\theta = 90°$, and then decelerates to zero at the rear. As the free-stream velocity is increased, the flow separates from the surface before reaching $\theta = 180°$ (Fig. 4*a*). With further increase the separation point moves farther forward (Fig. 4*b*). When the Reynolds number, $N_{Re} = Du_\infty\rho/\mu$ (where D is the cyclinder diameter), reaches about 1×10^5 (Fig. 4*c*), the flow in the boundary layer around $\theta = 90°$ becomes turbulent. Thus, the boundary can remain attached to the surface longer and the separation point moves toward the rear (Fig. 4*d*). This change may be delayed until N_{Re} reaches as high as 5×10^5 in a low-turbulence-level airstream.

Dramatic changes in the local rate of heat transfer are produced by these changes in the flow. For $N_{Re} \geq 140{,}000$, the local heat transfer rises rapidly from a minimum around $\theta = 90°$ to a maximum and then decreases to a second minimum around 140°. This variation is due to the change to turbulent flow before separation on the rear of the cyclinder occurs. *See* LAMINAR FLOW; TURBULENT FLOW.

Condensation and boiling. Condensation and boiling are important phase-change processes involving heat release or absorption. Because vapor and liquid movement are present, the energy transfer is basically by convection. Local and average heat-transfer coefficients are determined and used in the Newton cooling-law equation for calculating heat rates which include the effects of the latent heat of vaporization.

Condensation. Consider a saturated or superheated vapor of a single substance in some region. When it comes in contact with a surface maintained at a temperature lower than the saturation temperature, heat flow results from the vapor, releasing its latent heat and condensing on the surface. The condensation process may proceed in two more or less

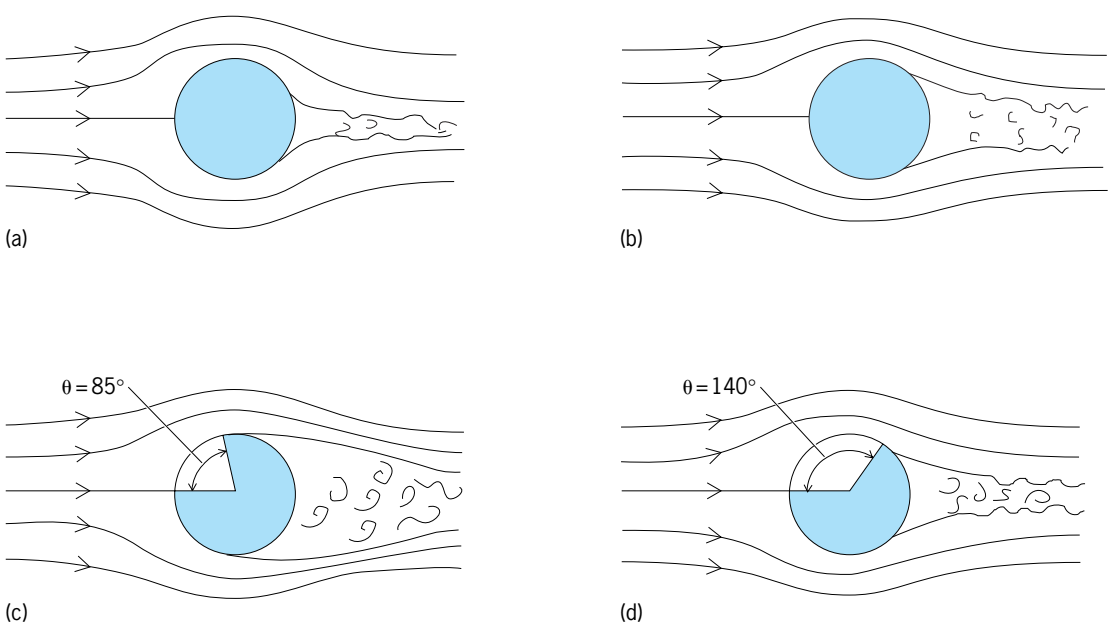

Fig. 4. Streamlines in flow past a cylinder normal to flow at increasing values of the flow velocity (expressed in terms of the Reynolds number, N_{Re}). (*a*) $N_{Re} \approx 10$. (*b*) $N_{Re} \approx 10^4$. (*c*) $N_{Re} \leq 1 \times 10^5$. (*d*) $N_{Re} \geq 1.5 \times 10^5$.

distinct ways. If there are no impurities in the vapor or on the surface (which need not be smooth), the condensate will form a continuous liquid film. If, however, such contaminants as fatty acids or mercaptans are present, the vapor will condense in small droplets. These droplets increase in size until their weight causes them to run down the surface. In doing so, they sweep the surface free for formation of new droplets. For the same temperature difference between the vapor and surface, heat transfer with dropwise condensation may be 15–20 times greater than filmwise condensation. The dropwise type is therefore very desirable, but conditions under which it will occur are not predictable, and designs are limited to systems for which experimental results are available.

Boiling. In boiling, results indicate the existence of several regimes. The important independent variable is the temperature difference Δt between the hot surface and the fluid relatively far from the surface. For values of Δt up to approximately 10°F (6°C), the liquid is being superheated by natural convection, and q/A is proportional to $\Delta t^{5/4}$. With further increase in Δt, bubbles form at active nuclei on the heated surface. These bubbles break away and rise through the pool, their stirring action causing the heat transfer to be much above that due to natural convection. This phenomenon is called nucleate boiling, and q/A varies as Δt^3 to Δt^4. When the rate of bubble formation becomes so rapid that the bubbles cannot get away before they tend to merge, a vapor film begins to form, through which heat must flow by conduction. The rate of heating then decreases with Δt until complete film boiling is reached and heat flows by radiation and conduction through the film.

The high heat rates which occur during boiling make it a very effective means of absorbing the energy capable of being released in furnaces and nuclear reactors. The high heat rates are also one reason why the vapor power-generating cycles have been successful. Warren H. Giedt

Bibliography. L. C. Burmeister, *Convection Heat Transfer*, 2d ed., 1993; A. J. Chapman, *Heat Transfer*, 4th ed., 1984; J. P. Holman, *Heat Transfer*, 8th ed., 1997; F. P. Incropera, *Fundamentals of Heat and Mass Transfer*, 4th ed., 1996; W. M. Kays and M. E. Crawford, *Convective Heat and Mass Transfer*, 3d ed., 1993.

Convective instability

A state of fluid flow in which the distribution of body forces along the direction of the net body force is unstable and will thus break down. Fluid flows are subject to a variety of instabilities, which may be broadly viewed as the means by which relatively simple flows become more complex. Instabilities are an important step in the transition between smooth and turbulent flow, and in the atmosphere they are responsible for phenomena ranging from thunderstorms to low- and high-pressure systems. Meteorologists and oceanographers divide instabilities into two broad classes: convective and dynamic. *See* DYNAMIC INSTABILITY.

In the broadest terms, convective instabilities arise when the displacement of a small parcel of fluid causes a force on that parcel which is in the same direction as the displacement. The parcel of fluid will then continue to accelerate away from its initial position, and the fluid is said to be unstable. In most geophysical flows, the convective motions that result from convective instabilities operate very quickly compared with the processes acting to destabilize the fluid; the result is that such fluids seem to be nearly neutrally stable to convection.

Types. The simplest type of convective instability arises when a fluid is heated from below or cooled from above. This heating and cooling result in a temperature distribution as shown in **Fig. 1**; here a small chunk of fluid at temperature T_3 is shown displaced upward. If it conserves its original temperature, it will be warmer than its environment at the same altitude and will accelerate upward, since warm air is less dense than cold air. The parcel is said to be positively buoyant. Warm air rises and cold air sinks; thus a fluid whose temperature decreases with altitude is convectively unstable, while one in which the temperature increases with height is convectively stable. A fluid at constant temperature is said to be convectively neutral. *See* ARCHIMEDES' PRINCIPLE.

The above description assumes that density depends on temperature alone and that density is conserved in parcels of fluid, so that when the parcels are displaced their density does not change. However, in the Earth's atmosphere, neither of these assumptions is true. In the first place, the density of air depends on pressure and on the amount of water vapor in the air, in addition to temperature. Second, the density will change when the parcel is displaced because both its pressure and its temperature will change. Because of these conditions it is convenient to define a quantity known as virtual potential temperature (θ_v) that is conserved and reflects the actual density of

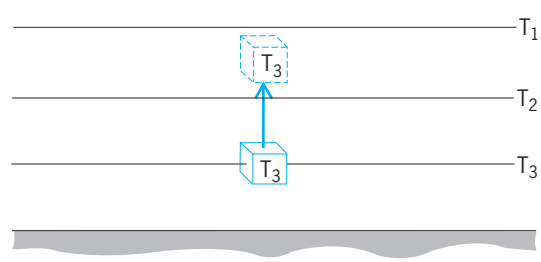

Fig. 1. Diagram showing temperature decreasing upward in a fluid and convective instability. When a small parcel of fluid is displaced upward from its initial position (solid-line box) to a new position (dashed-line box), it will be warmer than its environment and will continue to accelerate upward. *T* denotes temperature, with T_3 the highest and T_1 the lowest.

air. This quantity is given by Eq. (1), where T is the

$$\theta_v \equiv T\left(\frac{1+(r/0.622)}{1+r}\right)\left(\frac{1000}{p}\right)^{0.287} \qquad (1)$$

temperature in kelvins, p is the pressure in millibars, and r is the number of grams of water vapor in each gram of dry air. When θ_v decreases with height in the atmosphere, it is convectively unstable, while θ_v increasing with height denotes stability.

Over the tropical oceans, the θ_v of air adjacent to the sea surface is continually increased by addition of both heat and water vapor from the ocean, while at higher levels θ_v is decreased by cooling due to radiation to space. The result is that a layer roughly 1600 ft (500 m) deep is continually convecting. The convection is so efficient that this layer of the atmosphere is kept very close to a state of convective neutrality. Measurements show that θ_v is constant with height in this layer to within measurement error. A similar situation prevails over land during the day, when the land is heated by the Sun, but at night the land cools and the air becomes convectively stable (θ_v increases with height).

In the oceans, density is a function of pressure, temperature, and salinity; convection there is driven by cooling of the ocean surface by evaporation of water into the atmosphere and by direct loss of heat when the air is colder than the water. It is also driven by salinity changes resulting from precipitation and evaporation. In many regions of the ocean, a convectively driven layer exists near the surface in analogy with the atmospheric convective layer. This oceanic mixed layer is also nearly neutral to convection.

Convective instabilities are also responsible for convection in the Earth's mantle, which among other things drives the motion of the plates, and for many of the motions of gases within other planets and in stars.

Convective motions. Experiments have been performed to determine what kind of motions result from convective instability. Typically these experiments use a fluid contained between two parallel, horizontal plates maintained at constant temperature, with the lower plate warmer than the upper plate. If the temperature difference is small, the viscosity of the fluid prevents convection from occurring. As the temperature difference is increased beyond a critical value, convection begins as steady overturning cells whose width is roughly equal to the distance between the plates. When the temperature difference is increased further, the cells begin to oscillate and eventually the flow becomes chaotic. Variations on this experiment include one in which the flux of heat, rather than the temperature, is fixed at the parallel plates. In this case, the cells tend to be much broader than they are deep.

Virtually all convection in nature is well into the chaotic regime determined by experiments, and yet natural convection is often organized in distinctive patterns. In the atmosphere, these include long rows of clouds known as cloud streets and regular hexagonal arrays of clouds whose width can be more than 50 times their depth. The reason for this is as yet unknown. Little is known about the organization of convecting fluid in the Earth's mantle, which determines the motion of the plates that make up the Earth's crust. Convection on the Sun is also known to be organized in clumps called granules; these are in turn organized in larger clumps, or supergranules.

Moist convection. In the atmosphere, convection is complicated by the phase changes of water substance. When water vapor condenses into the liquid waterdroplets that make up clouds, for example, the heat that was used to evaporate the water in the first place is released, making the air warmer than it would otherwise be. Conversely, when rain evaporates it takes up heat and makes the air cooler. In addition to these effects, suspended condensate (that is, cloud, rain, snow, hail, and so forth) adds directly to the effective density of the air. All of these effects make moist convection, which produces cumulus clouds and thunderstorms and is considerably more exotic than dry convection.

The criterion for convective instability of moist atmospheres has been the subject of much research. A basic tool for examining the stability of moist atmospheres is the thermodynamic diagram, an example of which is presented in **Fig. 2**. This diagram has pressure on the vertical coordinate, and temperature increases downward and to the right. (Since pressure

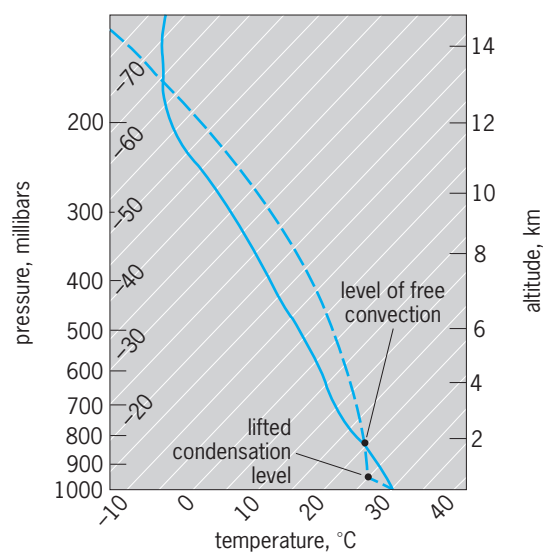

Fig. 2. Thermodynamic diagram with height on the vertical coordinate (with pressure decreasing upward), while temperature increases downward and to the right. The solid line shows an example of temperature measured at various levels in the atmosphere, typically by an ascending instrumented balloon. The broken line shows the temperature that a parcel would have if lifted from near the surface. At the lifted condensation level, water vapor begins to condense, and the parcel cools less rapidly above this point. At the level of free convection, it becomes warmer than its environment and would ascend freely above this point. 1 millibar = 10^2 pascals. 1 km = 0.6 mi. °F = (°C × 1.8) + 32.

decreases with height in the atmosphere, the vertical coordinate also represents height.) The temperature of air measured from an ascending balloon has been plotted as the solid line on this diagram. The thermodynamic diagram may contain various sets of background curves that permit a determination of the temperature of a parcel of air lifted from any point on the sounding. For example, the broken line shows the temperature a parcel of air would have if it were lifted from near the surface. At first, its temperature decreases rapidly with height, since the first law of thermodynamics states that temperature must fall with pressure. Then, when it reaches the lifted condensation level, water vapor begins to condense into cloud, releasing heat, and the temperature does not fall as fast with height.

In the example shown in Fig. 2, the lifted parcel has about the same temperature as the surrounding atmosphere up to its lifted condensation level, reflecting the neutral stability of the layer of air next to the surface. But if the parcel is displaced beyond its level of free convection, it will be warmer than its environment and will accelerate upward. This type of atmosphere is said to be conditionally unstable; that is, it is stable unless near-surface parcels are forcibly displaced beyond their level of free convection. Such an atmosphere will not spontaneously convect, and thus the amount of conditional instability can build up to large values until something triggers convection. The result is severe thunderstorms and sometimes tornadoes. This situation is common in the plains of North America in spring. *See* THUNDERSTORM; TORNADO.

It was formerly believed that the tropical atmosphere is conditionally unstable, but convective stability that is assessed by using thermodynamic diagrams like the one in Fig. 2 omits an important contribution to density: the weight of suspended condensate such as rain, snow, and cloud. When this is included, the tropical atmosphere is nearly neutrally stable to parcels of air displaced upward from near the surface. It is maintained in such a state by the nearly continuous activity of moist convection, which takes the form of cumulus clouds and thunderstorms. In this sense, the whole depth of the tropical troposphere (a layer of air extending up to about 7 mi or 12 km altitude in the tropics) exists in a convectively adjusted state in analogy with the layer near the surface, which has constant θ_v. *See* TROPICAL METEOROLOGY.

Convective instability may be driven by centrifugal as well as by gravitational forces. This is illustrated in **Fig. 3**, which shows the distribution of angular momentum per unit mass (M) in a rotating cylinder of fluid. This angular momentum is defined by expression (2), where r is the radius from the center

$$M \equiv rV \tag{2}$$

and V is the velocity of the fluid rotating about the center. If a ring of fluid centered about the axis of the cylinder is displaced radially inward or outward, it conserves its angular momentum. For example, as

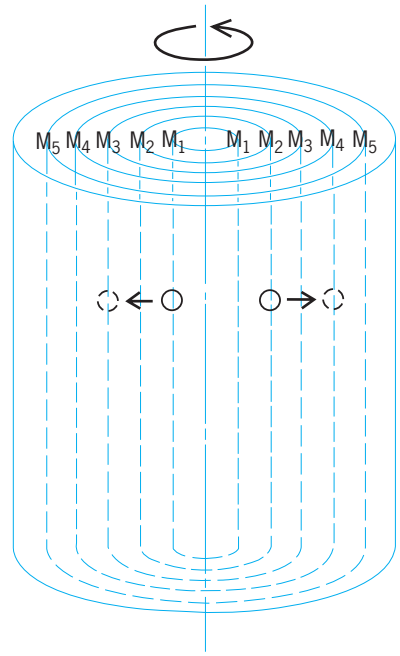

Fig. 3. Diagram of a cylinder of rotating fluid with an angular momentum per unit mass, *M*, which in this case increases outward from the axis (M_1, M_2, M_3, and so forth denote increasing values of *M*). This arrangement is centrifugally stable, since a ring of fluid (solid-line circle) displaced outward to a new position (broken-line circle) will have less centrifugal force acting on it than on its environment; it will therefore accelerate inward toward its initial position.

the ring of fluid in Fig. 3 is displaced outward, its angular momentum is less than that of the surrounding fluid, and the laws of physics can be used to demonstrate that since the centrifugal force on that fluid is less than on the surrounding fluid, it will accelerate inward. The fluid is then said to be centrifugally stable. A corollary to Archimedes' law states that a fluid with low M tends to sink toward the axis of rotation, while fluid with high M rises away from the axis. A fluid whose angular momentum increases outward is centrifugally stable, while a fluid with M decreasing outward is unstable. The nature of this instability is much like that of thermal convection, and this, too, is a type of convective instability. *See* ANGULAR MOMENTUM.

In the atmosphere, it is possible to have a combination of centrifugal and buoyant convection. It turns out that the criterion for this to happen is that parcels displaced upward along angular momentum surfaces must become positively buoyant. These surfaces are usually sloped in the atmosphere because of the increase with height of atmospheric winds. Convection that results from this combined instability (which is also known as symmetric instability) is also generally sloped and is thus called slantwise convection. This form of convection results in bands of rain and snow embedded in winter storms; these bands may be hundreds of miles wide and as much as 600 mi (1000 km) long. The concept of convective neutrality applies here as well, as measurements made by flying instrumented aircraft along angular momentum

surfaces show that parcels displaced along them are almost neutrally buoyant. *See* ATMOSPHERE; ATMOSPHERIC GENERAL CIRCULATION; FLUID FLOW; PRECIPITATION (METEOROLOGY); STORM. Kerry A. Emanuel

Bibliography. A. K. Betts, Saturation point analysis of moist convective overturning, *J. Atmos. Sci.*, 39:1484–1505, 1982; S. Chandrasekhar, *Hydrodynamic and Hydromagnetic Stability*, 1961; K. A. Emanuel, On assessing local conditional symmetric instability from atmospheric soundings, *Mon. Weath. Rev.*, 111:2016–2033, 1983; D. K. Lilly, Severe storms and storm systems: Scientific background, methods and critical questions, *Pure Appl. Geophys.*, 113:713–734, 1975; J. S. Turner, *Buoyancy Effects in Fluids*, 1973.

Converter

A device for processing alternating-current (ac) or direct-current (dc) power to provide a different electrical waveform. The term converter denotes a mechanism for either processing ac power into dc power (rectifier) or deriving power with an ac waveform from dc (inverter). Some converters serve both functions, others only one. *See* RECTIFIER.

Historically, converters were needed to accommodate various needs to match characteristics of the supply voltage. For example, dc motors, which in principle could be connected directly to early dc distribution systems, required a rectifier for operation with ac supplies. *See* DIRECT-CURRENT MOTOR.

While ac has been universally adopted for power distribution and consequently is the standard power supply for industrial, commercial, and domestic purposes, the range of demands for conversion far exceeds the occasional rectifier load. Converters are used for such applications as (1) rectification from ac to supply electrochemical processes with large controlled levels of direct current; (2) rectification of ac to dc followed by inversion to a controlled frequency of ac to supply variable-speed ac motors; (3) interfacing dc power sources (such as fuel cells and photoelectric devices) to ac distribution systems; (4) production of dc from ac power for subway and streetcar systems, and for controlled dc voltage for speed-control of dc motors in numerous industrial applications; and (5) transmission of dc electric power between rectifier stations and inverter stations within ac generation and transmission networks. *See* DIRECT-CURRENT TRANSMISSION.

These are a few examples of the need for converters to be versatile, controllable, and able to handle a wide range of power levels. Furthermore, emphasis on long-term cost of power losses has drawn attention to the efficiency of converters and the ability to contribute to the overall efficiency of their loads. For example, it can be economically advantageous to recapture the kinetic energy of some mechanical loads, which are subject to frequent stops and starts, through a converter power reversal back to the supply (regeneration) rather than losing it as heat in friction braking. This increases operating efficiency in such applications as electric vehicles and steel rolling mills.

Converter development has been furthered by advances in power semiconductors and microprocessors for their control. Beyond the basic converter action, there may be control of voltage, current, power, reactive power, and frequency at the converter output terminals. *See* MICROPROCESSOR.

Types. Until the advent of power semiconductors, converter action was achieved either by rotary converters or mercury-arc valves. In a rotary converter, an electric motor drives a generator. For example, a dc motor drives an ac generator to provide ac voltage at a chosen frequency. The mercury-arc valve is a unidirectional switching device which relies on ionization between a pool of mercury and an anode. Later developments permitted the timing of the start of conduction to be controlled by a grid between the mercury and the anode. Some of the modern semiconductor circuits have evolved from mercury-arc technology.

The introduction of the thyristor (silicon-controlled rectifier) in the 1960s had an immediate effect on converter applications because of its ruggedness, reliability, and compactness. As a controlled, unidirectional, solid-state switch, it replaced mercury-arc devices with ratings in the kilowatt range and, through series–parallel combinations, at high power levels. Power semiconductor devices for converter circuits include (1) thyristors, controlled unidirectional switches that, once conducting, have no capability to suppress current; (2) triacs, thyristor devices with bidirectional control of conduction; (3) gate turn-off devices with the properties of thyristors and the further capability of suppressing current; and (4) power transistors, high-power transistors operating in the switching mode, somewhat similar in properties to gate turn-off devices.

Thyristors are available with ratings from a few watts up to the capability of withstanding several kilovolts and conducting several kiloamperes. All the devices incorporate one or more gates, that is,

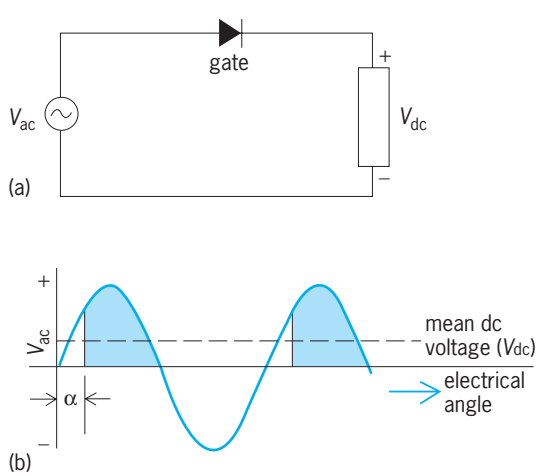

Fig. 1. Thyristor half-wave rectifier. (*a*) Circuit. (*b*) Voltage waveforms. Shaded areas indicate nonsmooth dc output waveform.

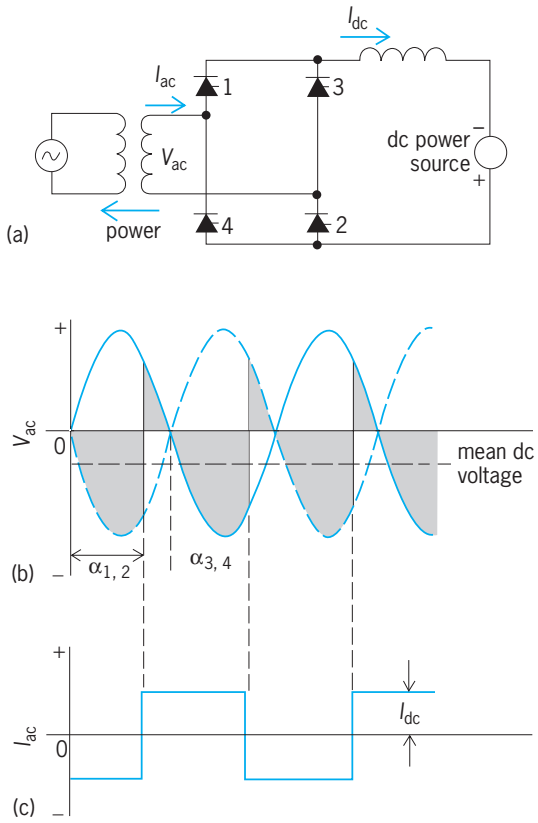

Fig. 2. Inverter operation by thyristor full-wave bridge. (*a*) Circuit. (*b*) Voltage waveforms. Shaded areas indicate nonsmooth dc waveform with negative mean value. (*c*) Current waveform.

low-power input connections to permit point-on-wave control of turn-on relative to the waveform of applied voltage. *See* SEMICONDUCTOR RECTIFIER.

Circuits. Of the many different converter circuits, three types are described.

Half-wave rectifier. The thyristor in the controlled half-wave rectifier (**Fig. 1***a*) conducts on parts of the positive half-cycles of the ac supply to provide a mean dc voltage across the load. Conduction is delayed by the variable control angle α resulting in the nonsmooth dc waveform shown (Fig. 1*b*). In practice, the dc component in the supply current tends to cause saturation problems in supply transformers.

Full-wave circuit. A full-wave circuit (**Fig. 2***a*) has a bidirectional current in the ac supply transformer. Thyristors 1 and 2 are gated together and then, in turn, dc-side current is transferred to thyristors 3 and 4 to provide the negative half-cycle of ac current. The inductor provides some smoothing of the dc waveform. By delaying gating by more than 90 electrical degrees while providing a dc power source (Fig. 2*b*), the sign of the mean dc voltage is reversed so that the circuit now inverts dc power into the ac supply. The ac system must provide an existing ac voltage waveform sufficient to naturally allow the thyristors to turn on and off in the conducting sequence. The ac current waveform is not sinusoidal (Fig. 2*c*); ac filters may be added to reduce the harmonic content. *See* COMMUTATION; ELECTRIC FILTER.

The above examples are called bridge circuits. Inversion into an ac system that does not contain an ac voltage source requires a circuit that either develops internal reverse voltages to turn off thyristors or uses gate turn-off devices. Other types include the cycloconverter which synthesizes a lower frequency from an ac waveform, and the pulse-width-modulated converter which processes a waveform into a series of segments that are recombined into the desired waveform. *See* PULSE MODULATION.

Three-phase bridge. Three-phase ac/dc converters are often used in industrial applications. Very high-power versions are used in power transmission systems.

Converters at the terminal stations of the dc transmission system either rectify three-phase ac voltages to dc or invert the dc voltage to feed power into the adjacent ac system. A three-phase converter bridge is made up of six thyristor valves conducting in sequence as numbered 1, 2, . . . , 6 in **Fig. 3***a*. Bridges are usually connected in series pairs, as shown. The 30° ac phase difference between the converter-side voltages of the wye/wye and the wye/delta converter transformers permits a conduction sequence 1, 1′, 2, 2′, . . . , 6, 6′. This is called 12-pulse operation and gives a smoother dc waveform and less ac harmonic currents.

Each valve is made up of many thyristors in series together with additional components to aid in the voltage sharing between individual thyristors (Fig. 3 *b*). Valves are often combined into single units, such as the quadrivalve in Fig. 3*a*.

Converters range from low-power applications, such as the power supply in a television, up to thousands of megawatts in dc power transmission. *See* ALTERNATING CURRENT; DIRECT CURRENT. John Reeve

Bibliography. C. W. Lander, *Power Electronics*, 3d ed., 1994; M. A. Rashid, *Power Electronics*, 3d ed., 2003.

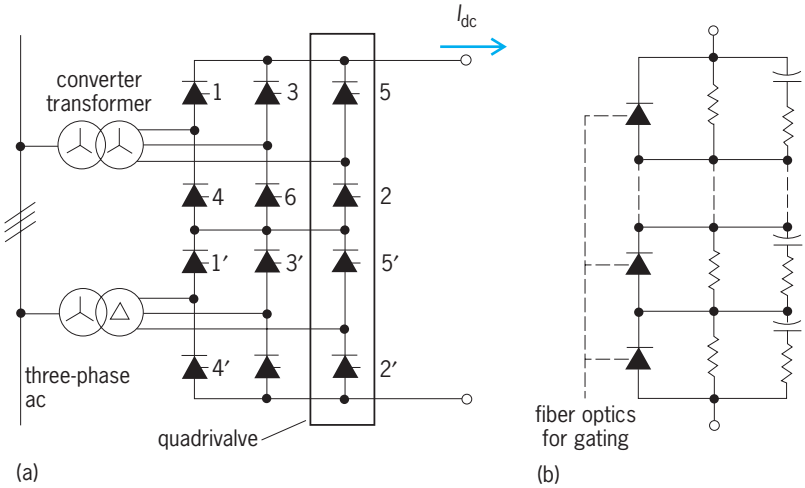

Fig. 3. Converter for dc power transmission. (*a*) Circuit of 12-pulse converter. (*b*) Thyristor valve.

Conveyor

A horizontal, inclined, declined, or vertical machine for moving or transporting bulk materials, packages, or objects in a path predetermined by the design of the device and having points of loading and discharge fixed or selective. Included in this category are skip hoist and vertical reciprocating and inclined reciprocating conveyors; but in the strictest sense this category does not include those devices known as industrial trucks, tractors and trailers, cranes, hoists, monorail cranes, power and hand shovels or scoops, bucket drag lines, platform elevators, or highway or rail vehicles. The more usual basic types of conveyors and their normal, rather than exceptional, operating characteristics

Normal operating characteristics of typical conveyors

	Paths						Typical products											Bulk Materials							
Conveyor Type	horizontal	declined	inclined	vertical	straight	curved	cases, boxes, etc.	barrels and kegs	drums	textile bags	paper bags	bottles and jars	cans	small parts	food products	lumber, pipe, etc.	towing trucks	free flowing	sluggish	dry	wet	cold	hot	nonabrasive	abrasive
GRAVITY CONVEYORS																									
Sliding friction																									
Skid		X			X			X	X																
Chute		X			X	X	X			X	X				X			X		X		X	X	X	
Spiral chute		X				X	X			X	X				X										
Rolling friction																									
Wheel	X	X			X	X	X				X					X									
Roller	X	X			X	X	X	X	X							X									
Spiral wheel		X				X	X																		
Spiral roller		X				X	X	X	X																
POWERED CONVEYORS																									
Continuous belt																									
Fabic	X	X	X		X		X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X
Flexible tube with zipper	X	X	X	X	X	X												X		X		X	X	X	X
Steel band	X	X	X		X		X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X
Woven or flat wire	X	X	X		X																				
Linked rod	X	X	X		X	X																			
Live roller (drive)																									
Flat belt	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X									
V belt	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X									
Sprocket	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X									
Chain																									
Apron (slat)	X	X	X		X		X	X	X	X	X														
Free roller	X	X	X		X		X	X	X																
Pan	X	X	X		X								X	X	X										
Pusher bar	X	X	X		X		X																		
Hinged plate	X	X	X										X	X	X										
Flat top plate	X				X	X						X	X												
Bucket*	X	X	X	X	X													X		X	X	X	X	X	X
Flight	X	X	X	X	X													X		X		X	X	X	
Drag chain																									
Plain links with lugs	X	X	X		X	X	X		X				X			X									
Overhead trolley†	X	X	X		X												X								
Infloor trolley	X				X												X								
Cable																									
Overhead trolley†	X	X	X		X																				
Table-mounted trolley	X	X	X																						
Cableway‡	X	X	X		X																				
Spiral (screw)	X	X	X	X	X													X	X	X	X	X	X	X	X
Vertical																									
Rigid arm				X	X		X	X	X	X	X														
Pendant carriage				X	X		X	X	X	X	X														
Pneumatic																									
Tube (with carriers)	X	X	X	X	X	X								X											
Pressure or suction	X	X	X	X	X	X												X		X		X	X	X	X
Vibrating	X	X	X	X	X													X		X		X	X	X	X

*Some bucket conveyors may use endless rubber belts.
†Most products or bulk materials can be conveyed by selection of attachments to the trolleys.
‡Cableways or serial tramways are used primarily for cross-country conveying of coal, ashes, lumber, and similar products used in construction.

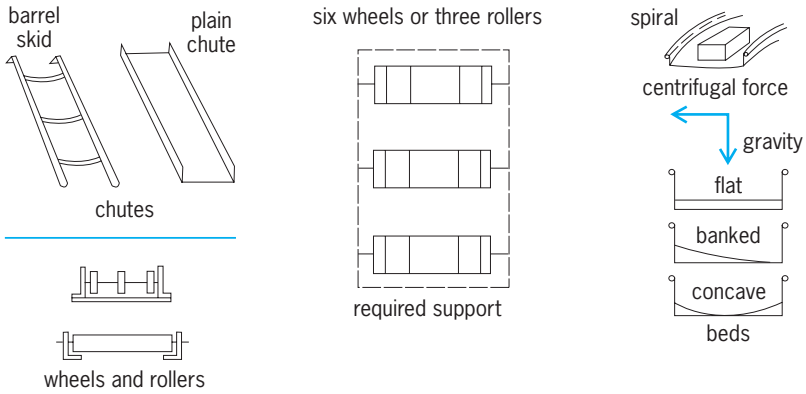

Fig. 1. Elements used in gravity conveyors.

are shown in the **table**. *See* BULK-HANDLING MACHINES.

Gravity conveyors. The economical means for lowering articles and materials is by gravity conveyors. Chutes depend upon sliding friction to control the rate of descent; wheel and roller conveyors use rolling friction for this purpose (**Fig. 1**).

With a body resting on a declined plane, friction F opposes component P parallel to the surface of the plane of weight W of the body. When the angle of elevation equals the angle of repose, the body is just about to start sliding, or to express it another way, when the inclination of the plane equals the angle of repose, F equals P. In this position the tangent of the angle of repose equals the coefficient of friction. For metal on metal the coefficient of friction is 0.15 to 0.25, and for wood on metal it is 0.2 to 0.6, which is why steel tote boxes slide more readily on metal chutes than do wooden cases.

When the body is a smooth-surfaced container and the inclined plane is replaced by rollers, the rollers are spaced so that at least three rollers support the smallest container to be conveyed; then each roller supports one-third the weight of the load. Component P of the weight W which acts downward and parallel to the inclination of the plane produces a turning moment Pr about the roller's axle, where r is the radius of the roller. As the inclination of the conveyor is increased, P increases until, at the point where rolling is about to start, Pr equals Fr. In this case, however, the force of friction is made up of the friction between the contacting surfaces and the friction in the bearings of the rollers. The mass of a tubular roller is concentrated near its circumference; hence rollers have greater starting inertia than wheels, but after they have started rolling, they have a flywheel effect that tends to keep products in motion, especially when packages follow each other in close succession. Inclination (or declination) is expressed as the number of inches of rise or fall per foot of conveyor or per 5-ft (1.5-m) or 10-ft (3-m) length.

Gravity chutes. Gravity chutes may be made straight or curved and are fabricated from sheet metal or wood, the latter being sometimes covered with canvas to prevent slivering. The bed of the chute can be shaped to accommodate the products to be handled. In spiral chutes centrifugal force is the second controlling factor. When bodies on a spiral move too fast, they are thrown out toward a guard rail, and contact with the rail increases the friction thereby causing the bodies to slow down somewhat and settle back into the center of the runway and continue their descent at a controlled rate. Spirals with roller beds or wheels provide smooth descent of an article and tend to maintain the position of the article in its original starting position. Rollers may be of metal, wood, or plastic and can be arranged in an optimum position, depending upon the articles to be carried.

Changes in direction can be effected by turntables, ball transfer tables, or gradual S curves (**Fig. 2**). Generally speaking, wheel-type conveyors are less expensive than the roller varieties; however, the latter can withstand severer service. For successful use of wheeled conveyors, the conveyed article must have a smooth and firm riding surface.

Chain conveyors. There are three basic types of chain conveyors: (1) those that support the product being conveyed, (2) those that carry actuating elements between the two chains, and (3) those that operate overhead or are set into the floor (**Fig. 3**).

In the sliding chain conveyor either plain links or links with such special attachments as lugs are made up into the conveyors to handle cases, cans, pipes, and other similar products.

Slats or apron conveyors are fitted with slats of wood or metal, either flat or in special shapes, between two power chains. The slats handle such freight as barrels, drums, and crates. With the addition of cleats these conveyors carry articles up steep inclines. In this type of conveyor the article rides on the top strand of the slats.

Push-bar conveyors are a variation of the slat conveyor, in which two endless chains are cross-connected at intervals by pusher bars which propel the load along a stationary bed or trough of the conveyor.

Powered conveyors. Gravity conveyors are limited to use in conditions where the material is to be lowered in elevation. The horizontal conveying distance capabilities of the gravity conveyor depend on the difference in elevation and the angle of sliding or rolling friction of the material to be conveyed.

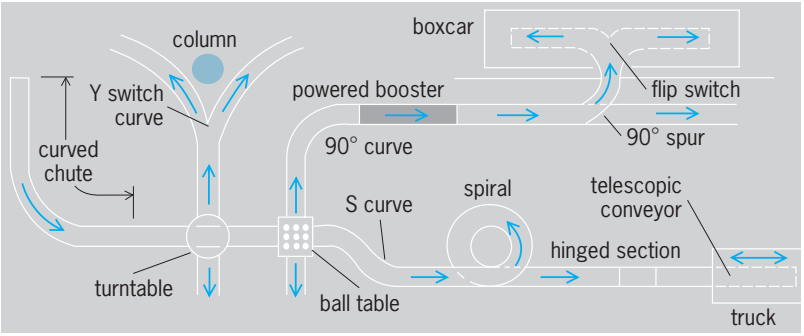

Fig. 2. Gravity components for conveyor lines.

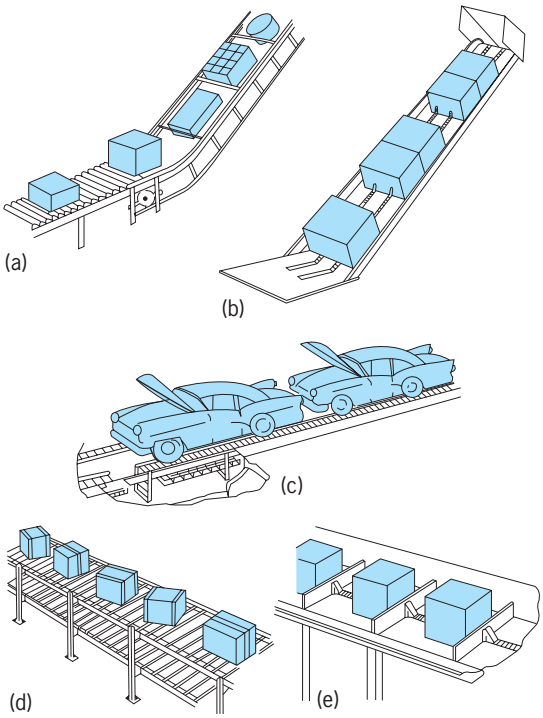

Fig. 3. Five examples of the basic types of chain conveyors. (*a*) Push-bar conveyor. (*b*) Twin chains carrying cartons. (*c*) Heavy-duty plates for assembly conveyor. (*d*) Apron or slat conveyor to carry miscellaneous articles. (*e*) Push-bar or cleat conveyor. (*After Conveyor Equipment Manufacturers Association, Conveyor Terms and Definitions, 3d ed., 1966*)

To move loads on level or inclined paths, or declining paths that exceed the angle of sliding or rolling friction of the particular material to be conveyed, powered conveyors must be employed. The following are various types of powered conveyors.

Belt conveyors. Loads are moved on a level or inclined path by means of power-driven belts. Belt conveyors with rough-top belts make possible inclines up to 28°; cleated belts are limited on degree of incline only by the position of the center of gravity of the conveyed item.

Essential components of belt conveyors in addition to the belt itself are (1) a bed, which may be a combination bed and frame; (2) end rollers; (3) a take-up to adjust belt tension as it varies with age and atmospheric conditions; and (4) a power unit (**Fig. 4**). The bed of a fabric belt conveyor may be wood, or more usually flat or concave sheet metal; or it may be made up of rollers. The former are called slide beds, the latter roller beds. Slide beds provide quiet, smooth conveying, but they require more power than roller beds. The two end rollers are mounted in antifriction bearings with provision for lubrication.

Where the tail roller is mounted in adjustable end arms to provide take-up, the length of the conveyor is not constant. Where this feature is objectionable, automatic gravity, screw, or spring take-ups can be used. Another required adjustment is provided by mounting the rollers so that their bearings can be shifted slightly either forward or backward to train the belt to run true.

Belts can be powered in several ways. A gear-head motor can be connected directly to the head roller, or the drive may be through a chain and sprockets. These arrangements necessitate placing the drive outside the frame. When this is undesirable, one of two drives can be located under the frame. Power units usually provide fixed or variable speeds from a few feet per minute to over 50 ft/min (15 m/min); 50 ft/min is usual with package conveyors. Normally $^1/_4$-hp (190-W) motors powers these units, but there are other, more powerful motors which are available for heavier duty.

Powered package conveyors can be combined in various ways to meet specific conditions. Powered curves are used, but wheel and roller sections are more usually employed to change direction. Inclined belt conveyors usually have a hump at the upper end to reduce the slapping by packages on the belt as they move from inclined to horizontal travel (**Fig. 5**). Cleats can be used, but they complicate the feeding of packages onto the line and require special construction features so that the cleats can pass around pulleys and be properly supported on the return portion of the conveying cycle.

Metal belts made of woven and flat wire, linked rods, and steel bands are engineered to make up conveyors, which are essentially the same as fabric belt conveyors. For example, woven wire belts can be driven by pulleys and these may be lagged, that is, covered with a material to improve traction.

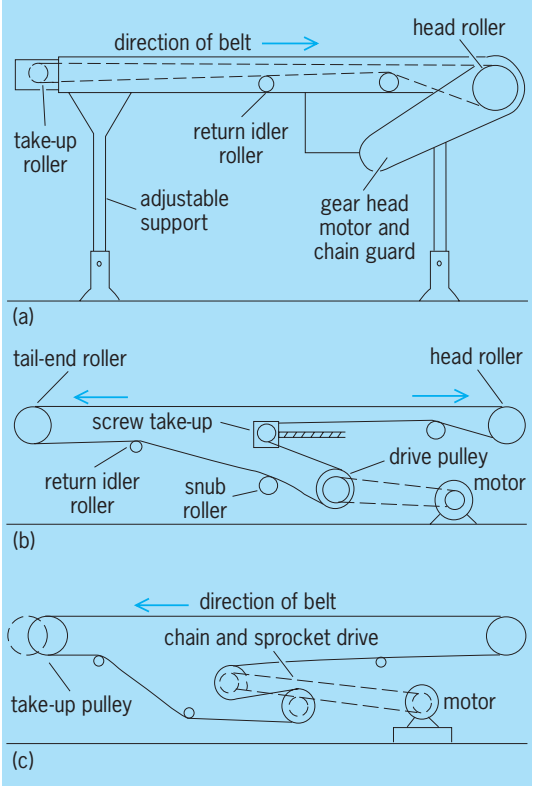

Fig. 4. Diagrams of three driven-belt conveyors. (*a*) Components of a fabric belt conveyor. (*b*) Wrap drive. (*c*) Tandem drive.

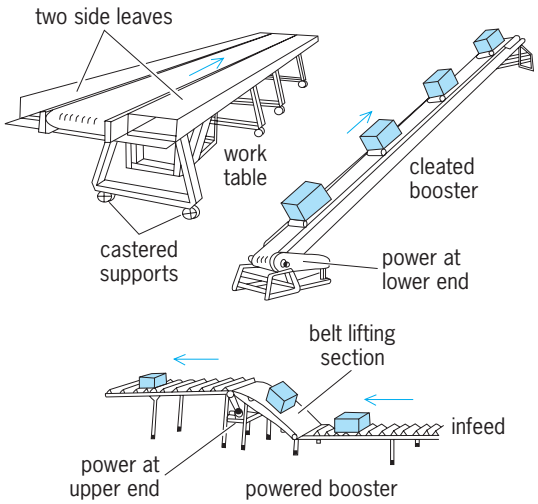

Fig. 5. Fabric belt conveyors.

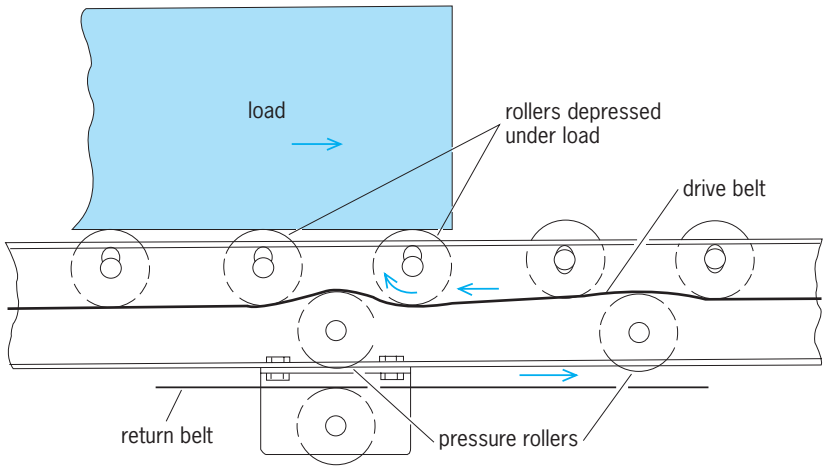

Fig. 6. Live rollers provide power for gravity curves.

Contact with the drive pulley can be increased by snubbing with small-diameter idler pulleys which guide the belt to increase the angle of contact. Flat wire, linked rod, and woven wire belts with chain edges are driven positively by pulleys with sprocket teeth or by true sprockets. Steel band conveyors are driven by pulleys with unusually large diameters. Other components are similar to those used in fabric belt conveyors but are more rugged in construction to handle the heavier carriers and to meet the more exacting conditions under which metal belts operate. Metal belts are used, for example, in high-temperature operations such as baking, annealing, brazing, and heat-treating.

The choice of a belt for a given situation depends on the effect of the belt's surface on the product, amount of drainage required, and similar factors.

Live-roller conveyors. Objects are moved over series of rollers by the application of power to all or some of the rollers. The power-transmitting medium is usually belting or chain. In one arrangement a belt drive running under and in contact with the rollers turns them to propel the load forward (**Fig. 6**). In another variation small sprockets are attached to the ends of some or all of the roller shafts, the sprockets being powered by a chain drive. Arthur M. Perrin

Vibrating conveyors. These mechanical devices are designed to move bulk materials along a horizontal, or almost horizontal, path in a controlled system (**Fig. 7**). They can be used to simply transport material from one point to another or to perform various functions en route, such as cooling, drying, blending, metering, spreading, and, by installing a screen, scalping or dedusting.

They can be fed by a belt conveyor or, more commonly, used to provide a precisely controlled outlet from a bin, hopper, or chute. In all cases they have the ability to take a vertical flow of material and change the direction of movement up to 90° with only a small change in the handling level.

A vibrating conveyor basically consists of a base or reaction mass and a driven mass connected to each other by springs. The base is excited by a power source such as an electromagnetic or electromechanical drive. A trough or pan forms part of the driven mass and serves as the material carrying surface. In operation, the trough is moved alternately forward and up, then down and back. This causes the conveyed material to move along the trough surface in a series of short hops.

Amplitude or frequency of vibration can be adjusted to control conveyor speed and output. Electromagnetic conveyors operating at 3600 vibrations per minute with a low amplitude [a displacement of 0.09 in. (0.2 cm), for example], provide a gentle motion to the conveyed product and are frequently used to handle fragile or friable materials. Electromechanical conveyors normally operate at a lower frequency but a much higher amplitude [a displacement of 0.5 in. (1.2 cm), for example], which results in rougher handling of the product but greatly

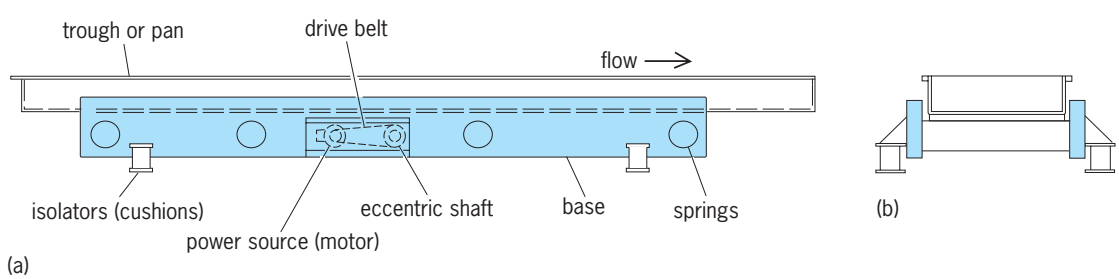

Fig. 7. Vibrating conveyor. (*a*) Side view. (*b*) End view.

increased capacity. The volume of material which can be handled ranges from a few pounds per hour up to thousands of pounds per hour. *See* VIBRATION MACHINE. R. F. Merwin

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; Conveyor Equipment Manufacturers Association, *Belt Conveyors for Bulk Materials*, 5th ed., 1997; J. Fruchtbaum, *Bulk Materials Handling Handbook*, 1986, rev. 1988; F. F. Meyers, *Plant Layout and Material Handling*, 1993.

Cooling tower

A tower- or buildinglike device in which atmospheric air (the heat receiver) circulates in direct or indirect contact with warmer water (the heat source) and the water is thereby cooled. A cooling tower may serve as the heat sink in a conventional thermodynamic process, such as refrigeration or steam power generation, or it may be used in any process in which water is used as the vehicle for heat removal, and when it is convenient or desirable to make final heat rejection to atmospheric air. Water, acting as the heat-transfer fluid, gives up heat to atmospheric air, and thus cooled, is recirculated through the system, affording economical operation of the process.

Basic types. Two basic types of cooling towers are commonly used. One transfers the heat from warmer water to cooler air mainly by an evaporation heat-transfer process and is known as the evaporative or wet cooling tower. The other transfers the heat from warmer water to cooler air by a sensible heat-transfer process and is known as the nonevaporative or dry cooling tower. These two basic types are sometimes combined, with the two cooling processes generally used in parallel or separately, and are then known as wet-dry cooling towers.

Cooling process. With the evaporative process, the warmer water is brought into direct contact with the cooler air. When the air enters the cooling tower, its moisture content is generally less than saturation; it emerges at a higher temperature and with a moisture content at or approaching saturation. Evaporative cooling takes place even when the incoming air is saturated, because as the air temperature is increased in the process of absorbing sensible heat from the water, there is also an increase in its capacity for holding water, and evaporation continues. The evaporative process accounts for about 65–75% of the total heat transferred; the remainder is transferred by the sensible heat-transfer process.

The wet-bulb temperature of the incoming air is the theoretical limit of cooling. Cooling the water to within 5 to 20°F (−15 to −6.7°C) above wet-bulb temperature represents good practice. The amount of water evaporated is relatively small. Approximately 1000 Btu (1055 kilojoules) is required to vaporize 1 lb (0.45 kg) of water at cooling tower operating temperatures. This represents a loss in water of approximately 0.75% of the water circulated for each 10°F (6°C) cooling, taking into account the normal proportions of cooling by the combined evaporative and sensible heat-transfer processes. Drift losses may be as low as 0.01–0.05% of the water flow to the tower (recent performance of 0.001% has been achieved) and must be added to the loss of water by evaporation and losses from blowdown to account for the water lost from the system. Blowdown quantity is a function of makeup water quality, but it may be determined by regulations concerning its disposal. Its quality is usually expressed in terms of the allowable concentration of dissolved solids in the circulating cooling water and may vary from two to six concentrations with respect to the dissolved solids content of the cooling water makeup.

With the nonevaporative process, the warmer water is separated from the cooler air by means of thin metal walls, usually tubes of circular cross section, but sometimes of elliptical cross section. Because of the low heat-transfer rates from a surface to air at atmospheric pressure, the air side of the tube is made with an extended surface in the form of fins of various geometries. The heat-transfer surface is usually arranged with two or more passes on the water side and a single pass, cross flow, on the air side. Sensible heat transfer through the tube walls and from the extended surface is responsible for all of the heat given up by the water and absorbed by the cooling air. The water temperature is reduced, and the air temperature increased. The nonevaporative cooling tower may also be used as an air-cooled vapor condenser and is commonly employed as such for condensing steam. The steam is condensed within the tubes at a substantially constant temperature, giving up its latent heat of vaporization to the cooling air, which in turn is increasing in temperature. The theoretical limit of cooling is the temperature of the incoming air. Good practice is to design nonevaporative cooling towers to cool the warm circulating water to within 25 to 35°F (14 to 20°C) of the entering air temperature or to condense steam at a similar temperature difference with respect to the incoming air. Makeup to the system is to compensate for leakage only, and there is no blowdown requirement or drift loss.

With the combined evaporative-nonevaporative process, the heat-absorbing capacity of the system is divided between the two types of cooling towers, which are selected in some predetermined proportion and usually arranged so that adjustments can be made to suit operating conditions within definite limits. The two systems, evaporative and nonevaporative, are combined in a unit with the water flow arranged in a series relationship passing through the dry tower component first and the wet tower second. The airflow through the towers is in a parallel-flow relationship, with the discharge air from the two sections mixing before being expelled from the system. Since the evaporative process is employed as one portion of the cooling system, drift, makeup, and blowdown are characteristics of the combined evaporative-nonevaporative cooling tower system,

Fig. 1. Counterflow natural-draft cooling tower at Trojan Power Plant in Spokane, Washington. (*Research–Cottrell*)

generally to a lesser degree than in the conventional evaporative cooling towers.

Of the three general types of cooling towers, the evaporative tower as a heat sink has the greatest thermal efficiency but consumes the most water and has the largest visible vapor plume. When mechanical-draft cooling tower modules are arranged in a row, ground fogging can occur. This can be eliminated by using natural-draft towers, and can be significantly reduced with modularized mechanical-draft towers when they are arranged in circular fashion.

Fig. 2. Cross-flow mechanical-draft cooling towers. (*Marley Co.*)

The nonevaporative cooling tower is the least efficient type, but it can operate with practically no consumption of water and can be located almost anywhere. It has no vapor plume.

The combined evaporative-nonevaporative cooling tower has a thermal efficiency somewhere between that of the evaporative and nonevaporative cooling towers. Most are of the mechanical-draft type, and the vapor plume is mitigated by mixing the dry warm air leaving the nonevaporative section of the tower with the warm saturated air leaving the evaporative section of the tower. This retards the cooling of the plume to atmospheric temperature; visible vapor is reduced and may be entirely eliminated. This tower has the advantage of flexibility in operation; it can accommodate variations in available makeup water or be adjusted to atmospheric conditions so that vapor plume formation and ground fogging can be reduced.

Evaporative cooling towers. Evaporative cooling towers are classified according to the means employed for producing air circulation through them: atmospheric, natural draft, and mechanical draft.

Atmospheric cooling. Some towers depend upon natural wind currents blowing through them in a substantially horizontal direction for their air supply. Louvers on all sides prevent water from being blown out of these atmospheric cooling towers, and allow air to enter and leave independently of wind direction. Generally, these towers are located broadside to prevailing winds for maximum sustained airflow.

Thermal performance varies greatly because it is a function of wind direction and velocity as well as wet- and dry-bulb temperatures. The normal loading of atmospheric towers is about 1–2 gal/min (3.7–7.5 liters/min) of cooling water per square foot of cross section. They require considerable unobstructed surrounding ground space in addition to their cross-sectional area to operate properly. Because they need more area per unit of cooling than other types of towers, they are usually limited to small sizes.

Natural draft. Other cooling towers depend for their air supply upon the natural convection of air flowing upward and in contact with the water to be cooled. Essentially, natural-draft cooling towers are chimney-like structures with a heat-transfer section installed in their lower portion, directly above an annular air inlet in a counterflow relationship with the cooling air (**Fig. 1**), or with the heat-transfer section circumscribing the base of the tower in a cross-flow relationship with the cooling airflow (**Fig. 2**). Sensible heat absorbed by the air in passing over the water to be cooled increases the air temperature and its vapor content and thereby reduces its density so that the air is forced upward and out of the tower by the surrounding heavier atmosphere. The flow of air through the tower varies according to the difference in specific weights of the ambient air and the air leaving the heat-transfer surfaces. Since the difference in

specific weights generally increases in cold weather, the airflow through the cooling tower also increases, and the relative performance improves in reference to equivalent constant-airflow towers.

Normal loading of a natural-draft tower is 2–4 gal/(min)(ft^2) [1.4–2.7 liters/(s)(m^2)] of ground-level cross section. The natural-drafting cooling tower does not require as much unobstructed surrounding space as the atmospheric cooling tower does, and is generally suited for both medium and large installations. The natural-draft cooling tower was first commonly used in Europe. Then a number of large installations were built in the United States, with single units 385 ft (117 m) in diameter by 492 ft (150 m) high capable of absorbing the heat rejected from an 1100-MW light-water-reactor steam electric power plant.

Mechanical draft. In cooling towers that depend upon fans for their air supply, the fans may be arranged to produce either a forced or an induced draft. Induced-draft designs are more commonly used than forced-draft designs because of the lower initial cost, improved air-water contact, and less air recirculation (**Fig. 3**). With controlled airflow, the capacity of the mechanical-draft tower can be adjusted for economic operation in relation to heat load and in consideration of ambient conditions.

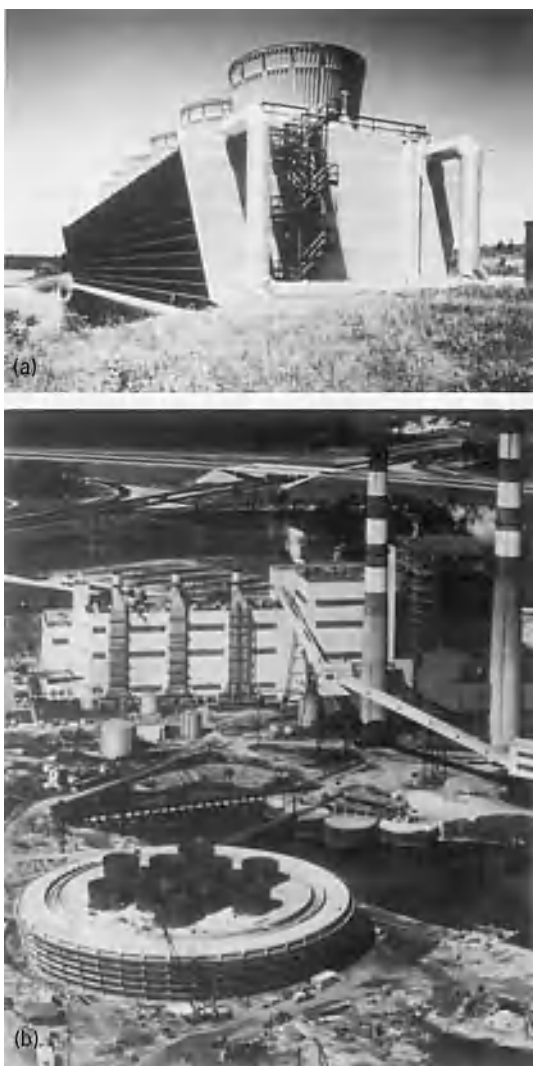

Fig. 3. Mechanical-draft cooling towers. (*a*) Conventional rectangular cross-flow evaporative induced-draft type; (*b*) circular evaporative induced-draft type. (*Marley Co.*)

Normal loading of a mechanical-draft cooling tower is 2–6 gal/(min)(ft^2) [1.4–4.1 liters/(s)(m^2)] of cross section. The mechanical-draft tower requires less unobstructed surrounding space to obtain adequate air supply than the atmospheric cooling tower needs; however, it requires more surrounding space than the natural-draft towers do. This type of tower is suitable for both large and small installations.

Nonevaporative cooling towers. Nonevaporative cooling towers are classified as air-cooled condensers and as air-cooled heat exchangers, and are further classified by the means used for producing air circulation through them. *See* HEAT EXCHANGER; VAPOR CONDENSER.

Air-cooled condensers and heat exchangers. Two basic types of nonevaporative cooling towers are in general use for power plant or process cooling. One type uses an air-cooled steam surface condenser as the means for transferring the heat rejected from the cycle to atmospheric cooling air. The other uses an air-cooled heat exchanger for this purpose. Heat is transferred from the air-cooled condenser, or from the air-cooled heat exchanger to the cooling air, by convection as sensible heat.

Nonevaporative cooling towers have been used for cooling small steam electric power plants since the 1930s. They have been used for process cooling since 1940; a complete refinery was cooled by the process in 1958. The nonevaporative cooling tower was formerly used almost exclusively with large steam electric power plants in Europe. Interest in this type of tower gradually increased in the United States, and a 330-MW plant in Wyoming using nonevaporative cooling towers was completed in the late 1970s.

The primary advantage of nonevaporative cooling towers is that of flexibility of plant siting. There is seldom a direct economic advantage associated with the use of nonevaporative cooling tower systems in the normal context of power plant economics. They are the least efficient of the cooling systems used as heat sinks.

Cooling airflow. Each of the two basic nonevaporative cooling tower systems may be further classified with respect to type of cooling airflow. Both types of towers, the direct-condensing type and the heat-exchanger type, can be built as natural-draft or as mechanical-draft tower systems.

Design. The heat-transfer sections are constructed as tube bundles, with finned tubes arranged in banks two to five rows deep. The tubes are in a parallel relationship with each other and are spaced at a pitch slightly greater than the outside fin diameter, either in an in-line or a staggered pattern. For each section, two headers are used, with the tube ends secured in each. The headers may be of pipe or of a box-shaped cross section and are usually made of steel. The bundles are secured in an open metal frame.

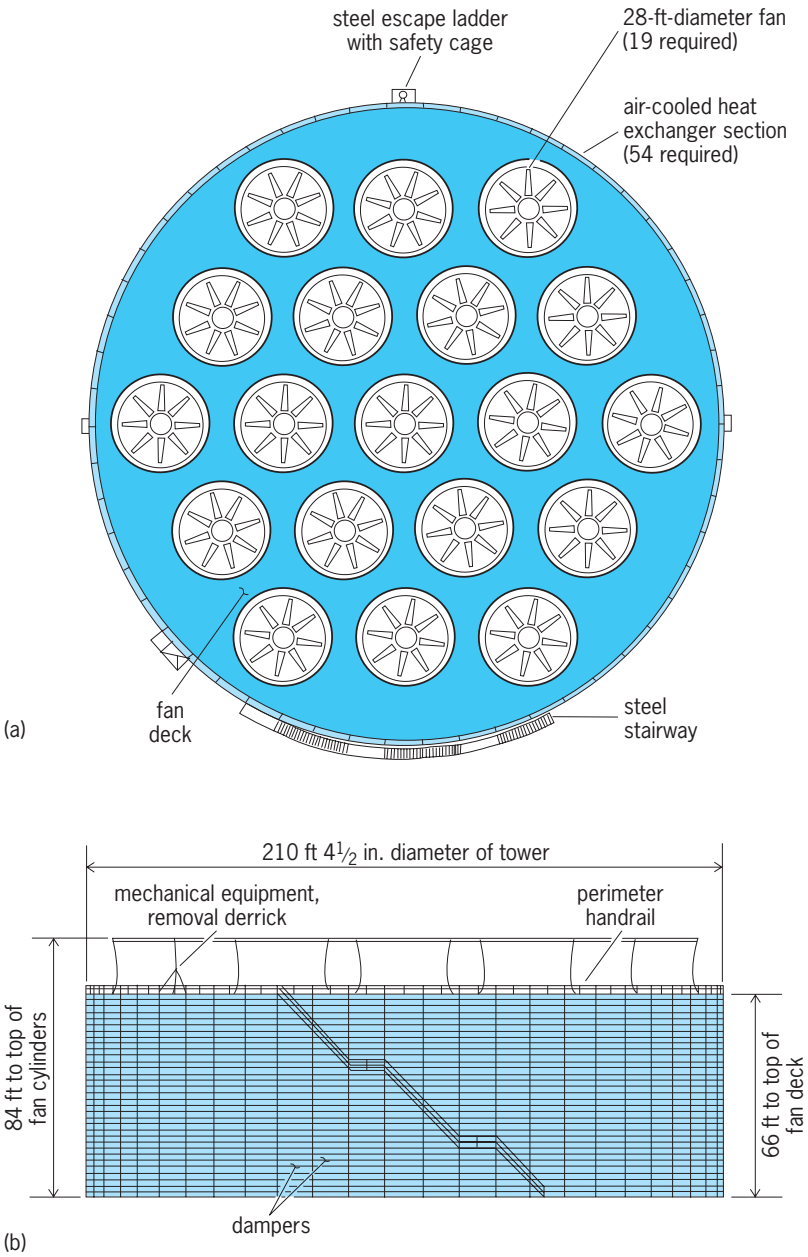

Fig. 4. Multifan circular nonevaporative tower. (*a*) Plan. (*b*) Elevation. 1 ft = 0.3 m; 1 in. = 2.5 cm. (*Marley Co.*)

The assembled tube bundles may be arranged in a V shape, with either horizontal or inclined tubes, or in an A shape, with the same tube arrangement. A similar arrangement may be used with vertical tube bundles. Generally, inclined tubes are shorter than horizontal ones. The inclined-tube arrangement is best suited to condensing vapor, the horizontal-tube arrangement best for heat-exchanger design. The A-shaped bundles are usually used with forced-draft airflow, the V-shaped bundles with induced-draft airflow. With natural-draft nonevaporative cooling towers there is no distinction made as to A- or V-shaped tube bundle arrangement; the bundles are arranged in a deck above the open circumference at the bottom of the tower, in a manner similar in principle to that used in the counterflow natural-draft evaporative cooling tower (Fig. 1). A natural-draft cooling tower with a vertical arrangement of tube bundles has been built, but this type of structure is not generally used.

Modularization of the bundle sections has become general practice with larger units; an installation of a circular module-type unit is shown in **Fig. 4**.

Combined evaporative-nonevaporative towers. The combination evaporative-nonevaporative cooling tower is arranged so that the water to be cooled first passes through a nonevaporative cooling section which is much the same as the tube bundle sections used with nonevaporative cooling towers of the heat-exchanger type. The hot water first passes through these heat exchangers, which are mounted directly above the evaporative cooling tower sections; the water leaving the nonevaporative section flows by force of gravity over the evaporative section.

Fig. 5. Wet-dry cooling tower at Atlantic Richfield Company. (*Marley Co.*)

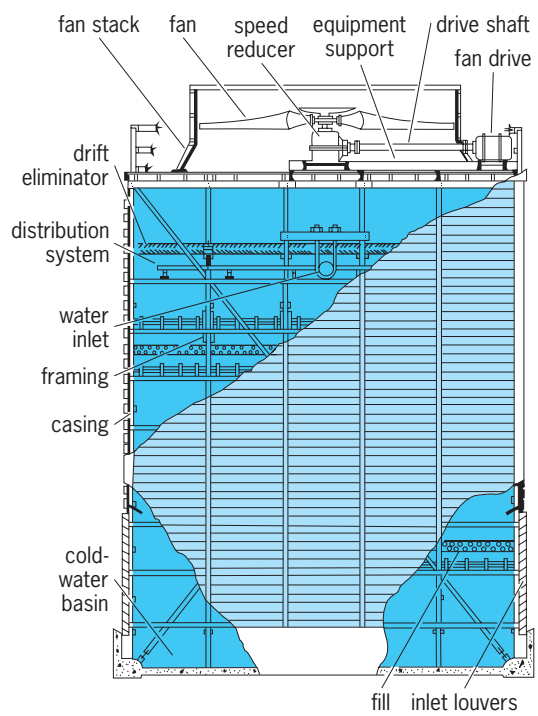

Fig. 6. Induced-draft counterflow cooling tower, showing the component parts.

The cooling air is divided into two parallel flow streams, one passing through the nonevaporative section, the other through the evaporative section to a common plenum chamber upstream of the induced-draft fans. There the two airflow streams are combined and discharged upward to the atmosphere by the fans. A typical cooling tower of this type is shown in **Fig. 5**.

In most applications of wet-dry cooling towers attempts are made to balance vapor plume suppression and the esthetics of a low silhouette in comparison with that of the natural-draft evaporative cooling tower. In some instances, these cooling towers are used in order to take advantage of the lower heat-sink temperature attainable with evaporative cooling when an adequate water supply is available, and to allow the plant to continue to operate, at reduced efficiency, when water for cooling is in short supply.

The more important components of evaporative cooling towers are the supporting structure, casing, cold-water basin, distribution system, drift eliminators, filling and louvers, discharge stack, and fans (mechanical-draft towers only). The counterflow type is shown in **Fig. 6**; the cross-flow type is shown in **Fig. 7**.

Treated wood, especially Douglas-fir, is the common material for atmospheric and small mechanical-draft cooling towers. It is used for structural framing, casing, louvers, and drift eliminators. Wood is commonly used as filler material for small towers. Plastics and asbestos cement are replacing wood to some degree in small towers and almost completely in large towers where fireproof materials are generally required. Framing and casings of reinforced concrete, and louvers of metal, usually aluminum, are generally used for large power plant installations. Fasteners for securing small parts are usually made of bronze, copper-nickel, stainless steel, and galvanized steel. Distribution systems may be in the form of piping, in galvanized steel, or fiber glass–reinforced plastics; and they may be equipped with spray nozzles of noncorrosive material, in the form of troughs and weirs, or made of wood, plastic, or reinforced concrete. Structural framing may also be made of galvanized or plastic-coated structural steel shapes. Natural-draft cooling towers, especially in large sizes, are made of reinforced concrete. The cold-water basins for ground-mounted towers are usually made of concrete; wood or steel is usually used for roof-mounted towers. Fan blades are made of corrosion-resistant

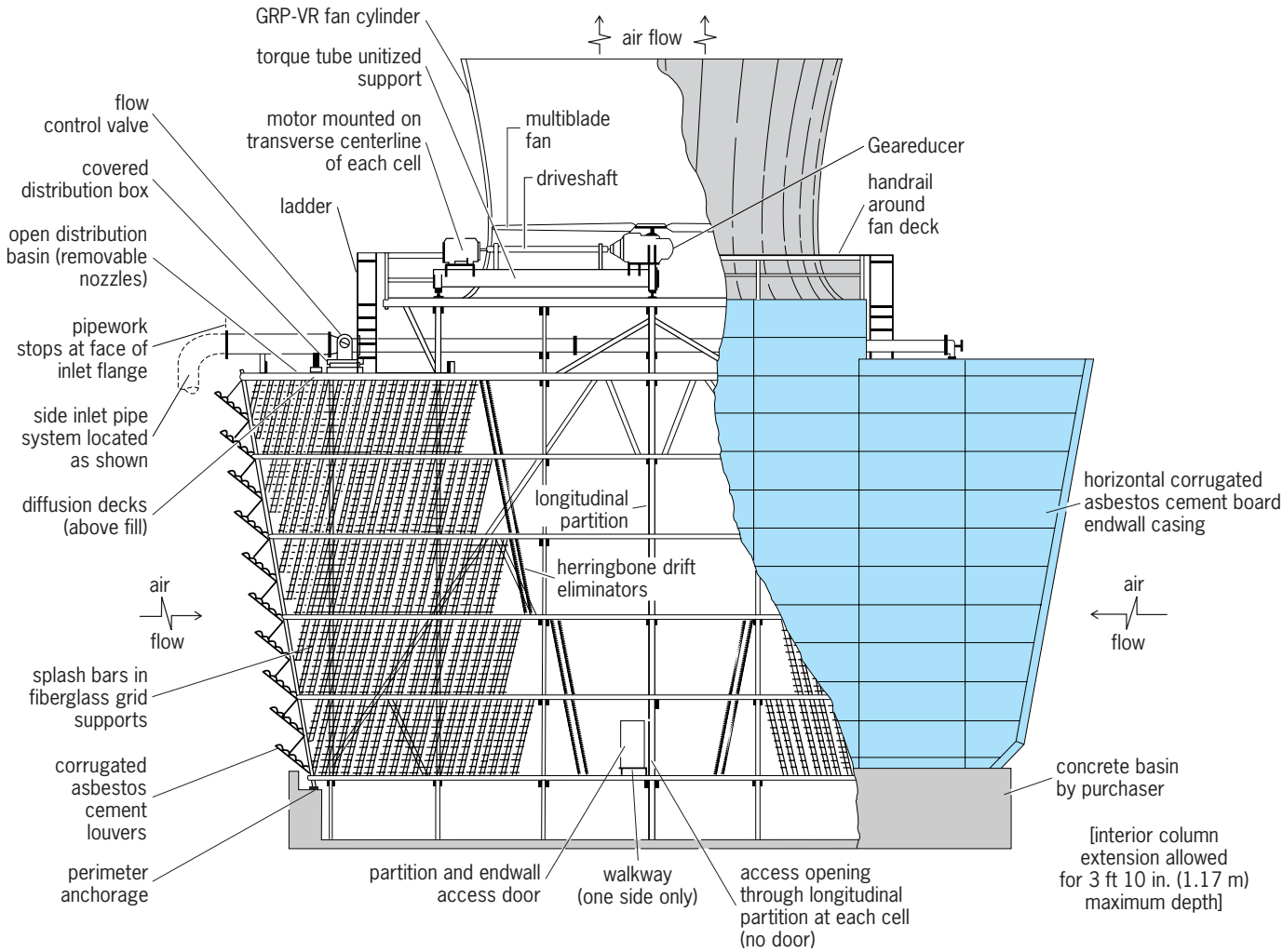

Fig. 7. Transverse cross section of a cross-flow evaporative cooling tower.

material such as monel, stainless steel, or aluminum; but most commonly fiber glass–reinforced plastic is used for fan blades.

The heat-transfer tubes used with nonevaporative cooling towers are of an extended-surface type usually with circumferential fins on the air side (outside). The tubes are usually circular in cross section, although elliptical tubes are sometimes used. Commonly used tube materials are galvanized carbon steel, ferritic stainless steel, and various copper alloys. They are usually made with wrapped aluminum fins, but steel fins are commonly used with carbon steel tubes and galvanized. Most designs employ a ratio of outside to inside surface of 20:25. Outside-diameter sizes range $^3/_4$–$1^1/_2$ in. (19–38 mm).

Tube bundles are made with tube banks two to five rows deep. The tubes are in parallel relationship with each other and secured in headers, either of steel pipe or of weld-fabricated steel box headers. The tube bundle assemblies are mounted in steel frames which are supported by structural steel framework. The bundle assemblies may be arranged in a V pattern, requiring fans of the induced-draft type, or in an A pattern, requiring fans of the forced-draft type (**Fig. 8**). Fans and louvers are similar to those described for evaporative cooling towers.

Nonevaporative cooling towers may also be used with natural airflow. In this case, the tube bundles are usually mounted on a deck within the tower and just above the top of the circumferential supporting structure for the tower. In this application, the tower has no cold-water basin, but otherwise it is identical with the natural-draft tower used for evaporative cooling with respect to materials of construction and design.

Performance. The performance of an evaporative cooling tower may be described by the generally accepted equation of F. Merkel, as shown below,

$$\frac{KaV}{L} = \int_{T2}^{T1} \frac{dT}{h'' - h}$$

where a = water-air contact area, ft^2/ft^3; h = enthalpy of entering air, Btu/lb; h'' = enthalpy of leaving air, Btu/lb; K = diffusion coefficient, lb/(ft^2)(h); L = water flow rate, lb/(h)(ft^2); T = water temperature, °F; T_1 = inlet water temperature, °F; T_2 = outlet water temperature, °F; and V = effective volume of tower, ft^3/ft^2 of ground area. The Merkel equation is usually performed graphically, by Simpson's rule, or with a suitable computer program. *See* PERFORMANCE RATING.

The performance of a nonevaporative cooling tower may be described by the generally accepted equation of Fourier for steady-state unidirectional heat transfer, using the classical summation of resistances formula with correction of the logarithmic temperature difference for cross-counterflow design in order to calculate the overall heat transfer. It is usual practice to reference the overall heat-transfer coefficient to the outside (finned) tube surface.

Evaluation of cooling tower performance is based on cooling of a specified quantity of water through a given range and to a specified temperature approach

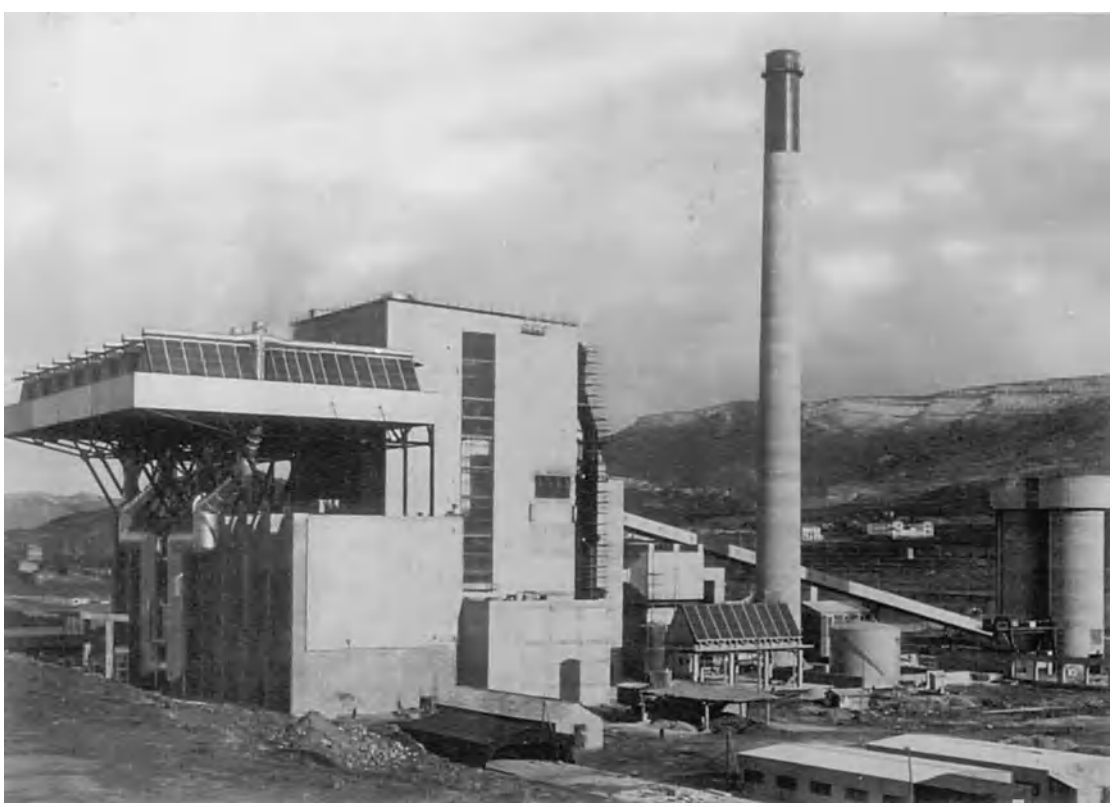

Fig. 8. Nonevaporative cooling tower, Utrillas, Spain. (*GEA*)

to the wet-bulb or dry-bulb temperature for which the tower is designed. Because exact design conditions are rarely experienced in operation, estimated performance curves are frequently prepared for a specific installation. These provide a means for comparing the measured performance with design conditions. Joseph F. Sebald

Bibliography. D. R. Baker, *Cooling Tower Performance*, 1983; N. P. Cheremisinoff and P. N. Cheremisinoff, *Cooling Towers: Selection, Design, and Practice*, 1981, reprint 1989; D. Q. Kern and A. D. Kraus, *Extended Surface Heat Transfer*, 1972; R. K. Miller and M. E. Rupnow, *Cooling Towers*, 1991; National Fire Protection Association Staff, *Water-Cooling Towers*, 1996.

Coordinate systems

Schemes for locating points in a given space by means of numerical quantities specified with respect to some frame of reference. These quantities are the coordinates of a point. To each set of coordinates there corresponds just one point in any coordinate system, but there are useful coordinate systems in which to a given point there may correspond more than one set of coordinates.

A coordinate system is a mathematical language that is used to describe geometrical objects analytically; that is, if the coordinates of a set of points are known, their relationships and the properties of figures determined by them can be obtained by numerical calculations instead of by other descriptions. It is the province of analytic geometry, aided chiefly by calculus, to investigate the means for these calculations.

The most familiar spaces are the plane and the three-dimensional euclidean space. In the latter a point P is determined by three coordinates (x, y, z). The totality of points for which x has a fixed value constitutes a surface. The same is true for y and z so that through P there are three coordinate surfaces. The totality of points for which x and y are fixed is a curve and through each point there are three coordinate lines. If these lines are all straight, the system of coordinates is said to be rectilinear. If some or all of the coordinate lines are not straight, the system is curvilinear. If the angles between the coordinate lines at each point are right angles, the system is rectangular.

Cartesian coordinate system. This is one of the simplest and most useful systems of coordinates. It is constructed by choosing a point O designated as the origin. Through it three intersecting directed lines OX, OY, OZ, the coordinate axes, are constructed. The coordinates of a point P are x, the distance of P from the plane YOZ measured parallel to OX, and y and z, which are determined similarly (**Fig. 1**). In this system the coordinate lines through P are straight lines respectively parallel to the three coordinate axes. Usually the three axes are taken to be mutually perpendicular, in which case the system is a rectangular cartesian one. Obviously a similar construction can be made in the plane, in which case a point has two coordinates (x, y). It is this system that is used in the construction of graphs and charts of various data, whether observed or computed.

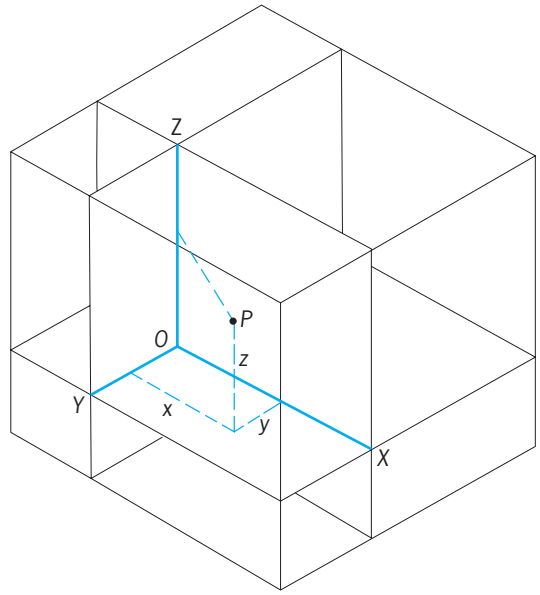

Fig. 1. Cartesian coordinate system.

Polar coordinate system. This system is constructed in the plane by choosing a point O called the pole and through it a directed straight line, the initial line. A point P is located by specifying the directed distance OP and the angle through which the initial line must be turned to coincide with OP in position and direction (**Fig. 2**). The coordinates of P are (r, θ). The radius vector r is the directed line OP, and the vectorial angle θ is the angle through which the initial line was turned, + if turned counterclockwise, − if clockwise. To each pair of values (r,θ) there corresponds just one point, but any point has an endless number of sets of coordinates. The coordinate lines in this case are radial lines through the pole ($\theta =$

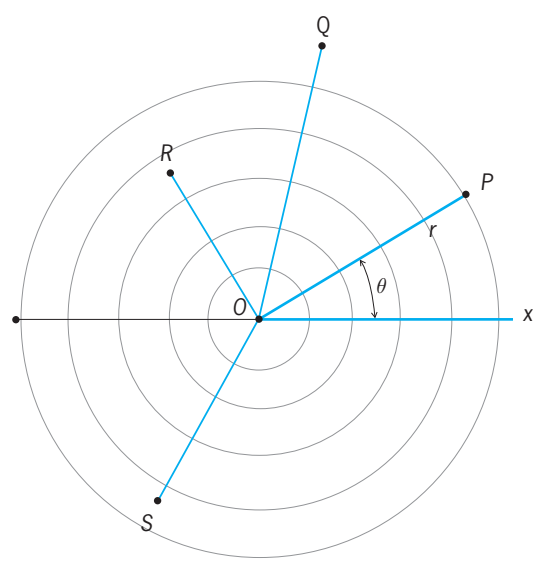

Fig. 2. Polar coordinate system.

constant) and concentric circles with center at the pole (r = constant). In spite of this lack of unique reciprocity between points and their coordinates, the polar system is useful in the study of spirals and rotations and in the investigation of motions under the action of central forces such as those of planets and comets.

Spherical coordinates. In three-dimensional euclidean space this system of coordinates is constructed by choosing a plane and in it constructing a polar coordinate system. At the pole O a polar axis OZ is constructed at right angles to the chosen plane. A point P, not on OZ, and OZ determine a plane. The spherical coordinates of P are then the directed distance OP denoted by ρ, the angle θ through which the initial line is turned to lie in ZOP and the angle $\phi = ZOP$ (**Fig. 3**). The coordinate lines are radial lines through O (θ and ϕ constant), meridian circles (ρ and ϕ constant), and circles of latitude (ρ and θ constant). This is an example of a curvilinear rectangle coordinate system. It is used in locating stars, in the study of spherical waves, and in problems in which there is spherical symmetry.

Cylindrical coordinates. These are constructed by choosing a plane with a pole O, an initial line in it, and a polar axis OZ, as in spherical coordinates. A point P is projected onto the chosen plane. The cylindrical coordinates of P are (r, θ, z) where r and θ are the polar coordinates of Q and $z = QP$ (Fig. 3). This is also a curvilinear rectangular system, the coordinate lines being mutually perpendicular. This system is used in problems of fluid flow and in others in which there is axial symmetry.

What has been done above in the plane and in euclidean space of three dimensions can be extended to curved spaces of any number dimensions. For a space S of n dimensions with some known geometrical or physical properties, a coordinate system in which a point has coordinates $(u_1, u_2, \ldots, u_n)$ is chosen. The coordinate lines are curves along which only one coordinate varies. When the known properties of S are expressed in terms of these coordinates, it is then the province of differential geometry to investigate their consequences. *See* DIFFERENTIAL GEOMETRY.

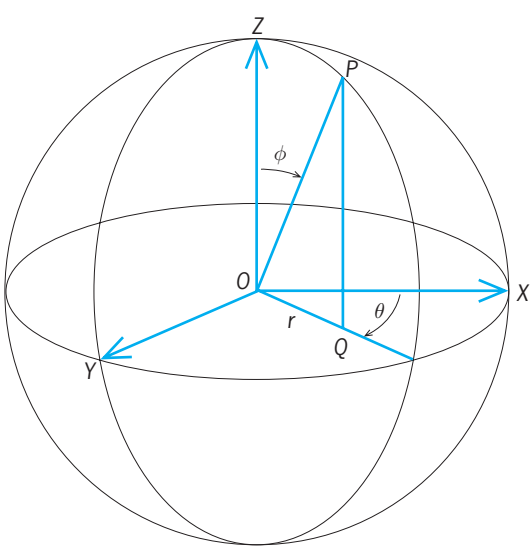

Fig. 3. Spherical coordinate system.

Transformation of coordinates. By means of a system of equations the description of a geometrical object in one coordinate system may be translated into an equivalent description in another coordinate system. Thus in a given space of n dimensions if there is a coordinate system A in which a general point P has coordinates $(u_1, u_2, \ldots, u_n)$ and a coordinate system B in which P has the coordinates $(\upsilon_1, \upsilon_2, \ldots, \upsilon_n)$, the transformation from system A to system B is the set of equations $u_i = f_i(\upsilon_1, \upsilon_2, \ldots, \upsilon_n)$, for $i = 1, 2, \ldots, n$, which expresses each u in terms of the υ's. These functions are obtainable from the relation between the two coordinate systems. They are not completely arbitrary, for they must be single-valued, independent, and such that if in the B system P has another set of coordinates, say $(\upsilon'_1, \upsilon'_2, \ldots, \upsilon'_n)$, then $f_i(\upsilon_1, \upsilon_2, \ldots, \upsilon_n) = f_i(\upsilon'_1, \upsilon'_2, \ldots, \upsilon'_n)$. Then if a geometrical locus in the space S is described by one or more equations of the form $F(u_1, u_2, \ldots, u_n) = 0$ in system A, the equivalent description in system B is one or more of the equations of the form

$$F[f_1(\upsilon), f_2(\upsilon), \ldots, f_n(\upsilon)] = 0$$

Other geometrical objects such as vectors and areas have more complicated laws of transformation, but each such law is expressible in terms of the equations of transformation. As remarked above, a coordinate system is a mathematical language; a transformation of coordinates plays the role of a dictionary that translates from one language to another.

The most important transformations of coordinates are between rectangular cartesian coordinate systems. One such set of coordinate axes is obtainable from another by a translation (in the physical sense) and a rotation. Consider a coordinate system A with coordinate axes OX, OY, OZ, and coordinates of a point $P(x,y,z)$. System B *is* obtained by moving the axes without turning to a point O' with coordinates (a, b, c). The new axes are $O'X'$, $O'Y'$, $O'Z'$ and the coordinates of P in this system are (x', y', z'); then the equations of transformation are $x = x' + a$, $y = y' + b$, $z = z' + c$. If the new system is obtained from A by a rotation of the coordinate axes, let $\alpha_1, \alpha_2, \alpha_3; \beta_1, \beta_2, \beta_3; \gamma_1, \gamma_2, \gamma_3$ be the angles which the new axes OX', OY', OZ' make with the original axes (only three of these angles are independent); then the equations of transformation are

$$x = x' \cos\alpha_1 + y' \cos\beta_1 + z' \cos\gamma_1$$

$$y = x' \cos\alpha_2 + y' \cos\beta_2 + z' \cos\gamma_2$$

$$z = x' \cos\alpha_3 + y' \cos\beta_3 + z' \cos\gamma_3$$

The equations of transformation may be differently interpreted. The equations of transformation make the point P of coordinates (x, y, z), in the same coordinate system, correspond to a point P' of coordinates (x', y', z'). This constitutes a mapping of

the space upon itself which maps a figure into some other figure. In applications one usually seeks a mapping that carries some pertinent loci into simpler loci while preserving some relevant properties. *See* ANALYTIC GEOMETRY; BARYCENTRIC CALCULUS; CALCULUS; CONFORMAL MAPPING; CURVE FITTING; SPHERICAL HARMONICS.

Morris S. Knebelman

Coordination chemistry

A field which, in its broadest usage, is acid-base chemistry as defined by G. N. Lewis. However, the term coordination chemistry is generally used to describe the chemistry of metals and metal ions in their interactions with other molecules or ions. For example, reactions (1)–(3) show Lewis acid-base-type re-

$$Mg^{2+} + 6H_2O \rightarrow Mg(H_2O)_6^{2+} \quad (1)$$

$$Ni + 4CO \rightarrow Ni(CO)_4 \quad (2)$$

$$Fe^{2+} + 6CN^- \rightarrow Fe(CN)_6^{4-} \quad (3)$$

actions, which result in the formation of coordinate covalent bonds. The products formed are coordination complex ions or compounds, and this area of chemistry is known as coordination chemistry.

It follows that coordination complex ions or compounds are ions or compounds that contain a central metal atom or ion and a group of ions or molecules bonded to it. Such an ion or compound tends to retain its identity, even in solution, although partial dissociation may occur. The charge on the coordinated species may be positive, zero, or negative, depending on the charges carried by the central atom and the coordinated groups. These groups are called ligands, and the total number of ligands bound to the central metal atom is called the coordination number. Other names commonly used for these compounds include complex compounds, complex ions, Werner complexes, coordinated complexes, chelate compounds, or simply complexes. *See* ACID AND BASE; CHELATION; COORDINATION COMPLEXES.

Experimental observations as early as the middle of the eighteenth century reported the isolation of coordination compounds. During that time and for the following 150 years, the valence theory could not adequately account for such materials. As a result, they were referred to as complex compounds, a term which is still in common usage, but not for the same reason. The correct interpretation of these compounds was finally given by Alfred Werner in 1893. He introduced the concept of residual or secondary valence, and suggested that elements have this type of valence in addition to their normal or primary valence. Thus, platinum(IV) has a normal valence of 4 but a secondary valence or coordination number of 6. This then led to the formulation of $PtCl_4 \cdot 6NH_3$ as $[Pt(NH_3)_6]^{4+}$, $4Cl^-$ and of $PtCl_4 \cdot 5NH_3$ as $[Pt(NH_3)_5Cl]^{3+}$, $3Cl^-$. The compound with five ammonias has only three ionic chlorides; the fourth is inside the coordination sphere, and therefore is not

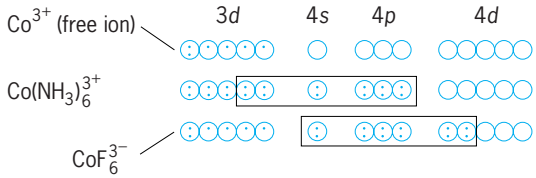

Fig. 1. Cobalt(III) complexes.

readily precipitated upon the addition of silver ion. Although the exact nature of the coordinate bond between metal and ligand remains the subject of considerable discussion, it is agreed that the formulations of Werner are essentially correct.

Coordinate bond. Three theories have been used to explain the nature of the coordinate bond. These are the valence bond theory, the electrostatic theory, including crystal field corrections, and the molecular orbital theory. Currently, the theory used almost exclusively is the molecular orbital theory. The valence bond theory for metal complexes was developed chiefly by Linus Pauling. This theory considers that the pair of electrons on the ligand enter the hybridized atomic orbitals of the metal and that the bond is essentially covalent. Several of the properties of these substances can be explained on the basis of this theory. For example, cobalt(III) complexes are represented by **Fig. 1**. The orbital hybridization of $Co(NH_3)_6^{3+}$ is designated as d^2sp^3, and the complex is referred to as an inner orbital complex. Such a representation is consistent with the diamagnetic properties of this cation, for all electrons are paired. The orbital hybridization of CoF_6^{3-} is designated as sp^3d^2, and it is called an outer orbital complex. This ion is known to be paramagnetic, which is in keeping with the four unpaired electrons in the 3*d* orbitals. The term inner orbital is applied if the *d* orbitals of a lower quantum number than the *s* and *p* orbitals are used in bonding, whereas outer orbital has reference to systems in which the *d* orbitals used have the same quantum number as the *s* and *p* orbitals. *See* ELECTRON CONFIGURATION; MOLECULAR ORBITAL THEORY; QUANTUM NUMBERS.

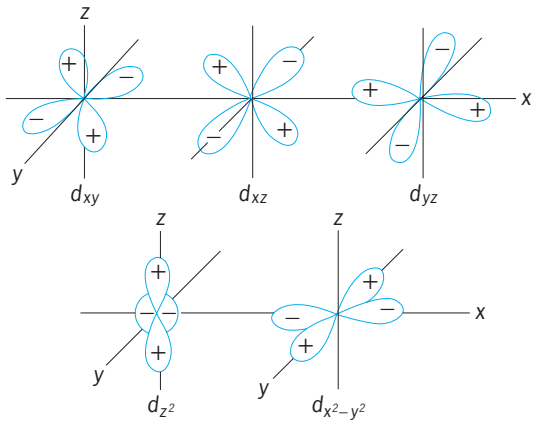

Fig. 2. Orientation of *d* orbitals in space.

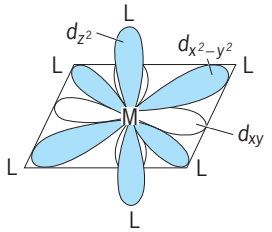

Fig. 3. Spatial orientation of *d* orbitals in coordination complex.

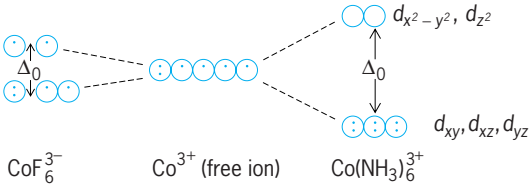

Fig. 4. Cobalt(III) complexes on the basis of crystal field theory.

The electrostatic theory, plus the crystal field theory for the transition metals, assumes that the metal-ligand bond is caused by electrostatic interactions between point charges and dipoles and that there is no sharing of electrons. Physicists have made good use of this theory to explain the properties of ionic crystalline solids, and it has now been extended to the metal complexes. For the nontransition metals, the parameters needed to determine the strength of the metal ligand bond are the charges and sizes of the central ions and the charges, dipole moments, polarizabilities, and sizes of the ligands. In order to give an adequate explanation of the bonding for transition metals, it is also necessary to consider the orientation of the *d* orbitals in space. The five possible spatial configurations are shown in **Fig. 2**. For the gaseous ion M, all five of the *d* orbitals are of equal energy. However, as shown in **Fig. 3**, the $d_{x^2-y^2}$ and d_{z^2} orbitals are pointing directly toward the six ligands at the corners of an octahedron and, because of repulsive interaction with the ligands, are at a higher energy than the d_{xy}, d_{xz}, and d_{yz} orbitals which do not point toward the ligands. *See* LIGAND FIELD THEORY.

Thus, on the basis of the crystal field theory, the cobalt(III) complexes referred to above are designated as in **Fig. 4**. The $Co(NH_3)_6^{3+}$ is called a spin-paired, or low-spin, complex, whereas CoF_6^{3-} is called spin-free, or high-spin. This refers to the fact that the electrons are all paired in the former because of the larger crystal field splitting, Δ_0, whereas in the latter, Δ_0 is small, and the electrons are not paired. In addition to explaining the structure and magnetic properties, this theory affords an adequate interpretation of the visible spectra of metal complexes.

The molecular orbital theory assumes that the electrons move in molecular orbitals which extend over all the nuclei of the metal-ligand system. In this manner, it serves to make use of both the valence bond theory and crystal field theory. The molecular orbital theory is therefore the best approximation of the nature of the coordinate bond because it is sufficiently flexible to permit both covalent and ionic bonding as well as the splitting of *d* orbitals into various energy levels. The complexes CoF_6^{3-} and $Co(NH_3)_6^{3+}$ can be represented by the molecular orbital energy diagrams shown in **Fig. 5**.

It is apparent from these molecular orbital diagrams that this theory combines the desirable features of both the valence bond and the crystal field theories. The covalent bonding of the valence bond theory appears as the sigma (σ) bonded molecular orbitals, designated here as σ_s, σ_p, and σ_d. Likewise the crystal field splitting (Δ_0) of the crystal field

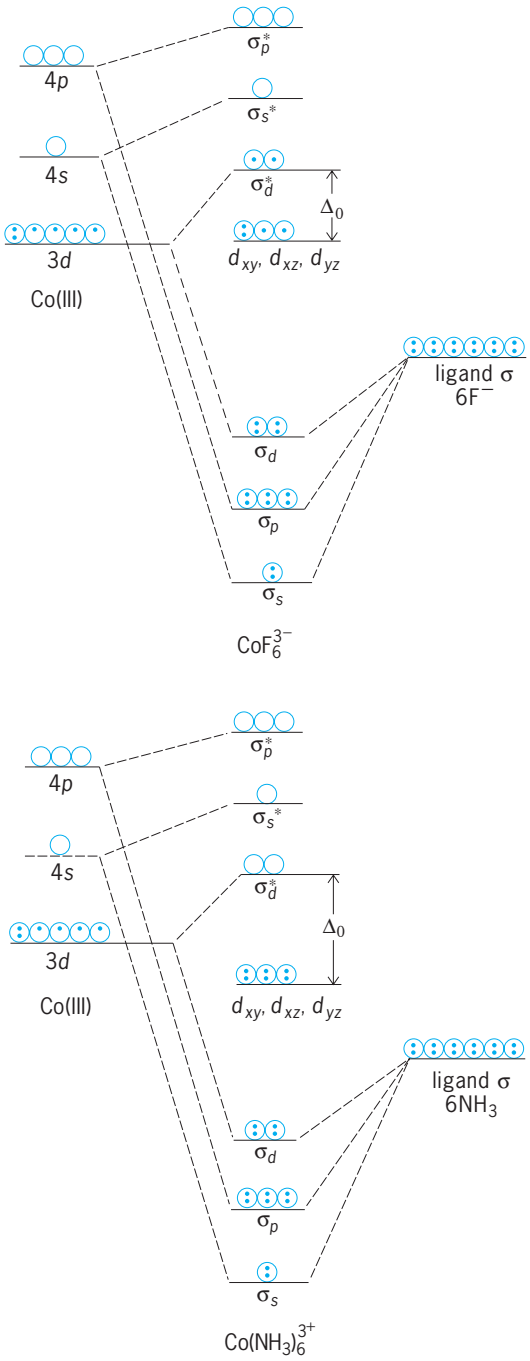

Fig. 5. Cobalt complexes CoF_6^{3-} and $Co(NH_3)_6^{3+}$ according to molecular orbital theory.

theory now is the energy difference between the nonbonding d orbitals d_{xy}, d_{xz}, d_{yz} and the antibonding sigma orbital σ_d^*. More complicated molecular orbital diagrams would include the contribution of π bonding in these systems.

Stability of complexes. The stability of metal complexes depends both on the metal ion and the ligand. There is a great deal of quantitative information on the stability of coordination compounds. These data show that, in general, the stability of metal complexes increases if the central ion increases in charge, decreases in size, and increases in electron affinity. Thus, alkali-metal ions have the least tendency to form complexes, and the highly polarizing transition-metal ions have the greatest tendency. However, even in the least favorable cases of the alkali-metal ions, there is ample evidence that coordination with certain ligands does occur. For the transition-metal ions, it also appears that the electronic configuration of the ion is significant with regard to the stability of its complexes. For example, regardless of the nature of the ligand, the stability of complexes of bivalent transition metals is Mn < Fe < Co < Ni > Cu > Zn. This so-called natural order of stability is explained in terms of the crystal field theory, which indicates that a maximum stabilization effect in spin-free complexes is realized with d^8 systems, such as exists in Ni^{2+}. The platinum metals, because of their large polarizing ability, generally form the most stable metal complexes.

Effect of ligand. Several characteristics of the ligand are known to influence the stability of complexes: (1) basicity of the ligand, (2) the number of metal-chelate rings per ligand, (3) the size of the chelate ring, (4) steric effects, (5) resonance effects, and (6) the ligand atom. Since coordination compounds are formed as a result of acid-base reactions where the metal ion is the acid and the ligand is the base, it follows that generally the more basic ligand will tend to form the more stable complex. Much of the available quantitative data on the formation constants of coordination compounds gives a good linear correlation with the base strength of the ligand. It is also known that a polydentate ligand, one attached to the metal ion at more than one point, forms more stable complexes than does an analogous monodentate ligand. For example, ethylenediamine forms more stable complexes than does ammonia, as shown in the following inequality.

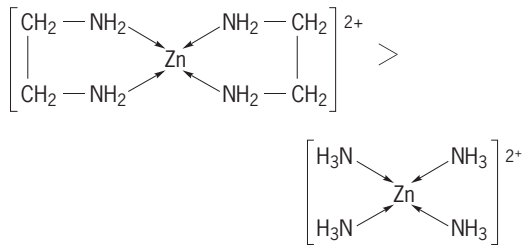

The ethylenediamine complex is referred to as a chelate complex. In general, an increase in the extent of chelation results in an increase in the stability of the complex. Ethylenediaminetetraacetate ion, a sexadentate ligand, is an excellent chelating agent and has found numerous applications as a sequestrant for metal ions. **Figure 6** is the structural formula of the chelate complex of calcium ion and ethylenediaminetetraacetate.

The size of the chelate ring is likewise an important factor. For saturated ligands such as ethylenediamine, five-membered rings are the most stable, whereas six-membered rings are the most stable for chelates containing one or more double bonds, for example (**1**) and (**2**). Examples of smaller and of

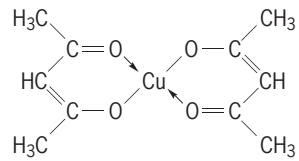

(**1**)

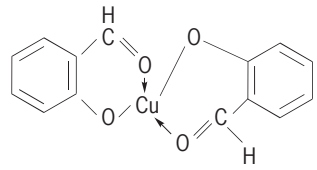

(**2**)

larger chelate ring systems are known, but these are much less common and less stable. It has been observed that for analogous systems an increase in double bonding in the chelate ring increases the stability of the complex. For ligands of comparable base strengths, the acetylacetonate complexes of type (**1**) are more stable than the salicylaldehyde compounds of type (**2**). There are in effect two double bonds per chelate ring in (**1**) compared to one and a half in (**2**). This same phenomenon is believed to be responsible for the marked stability of the porphyrin complexes such as chlorophyll, hemoglobin, and the phthalocyanine dyes.

Steric factors often have a very large effect on the stability of metal complexes. This is most frequently observed with ligands having a large group attached

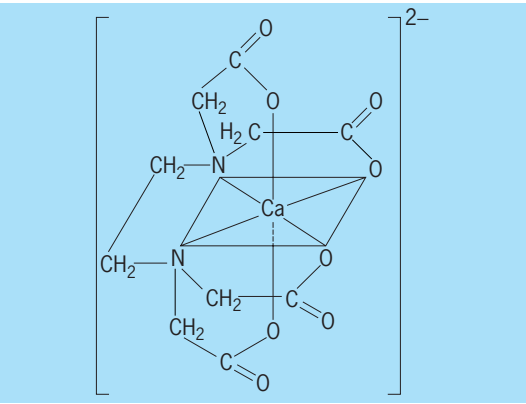

Fig. 6. Structural formula for the chelate complex formed from a calcium ion and ethylenediaminetetraacetate.

to the ligand atom or near it. Thus complexes, of the type shown in **Fig. 7**, with alkyl groups R in the position designated are much less stable than the parent complex where R = H. This results from the steric strain introduced by the size of the alkyl group on or adjacent to the ligand atom. In contrast to this, alkyl substitution at any other position results in the formation of more stable complexes because the ligand becomes more basic, and the bulky group is now removed from a position near the coordination site.

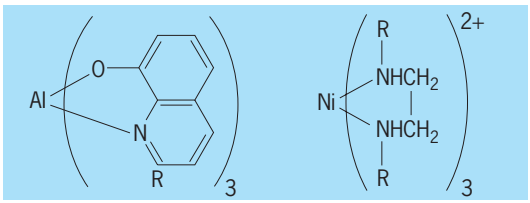

Fig. 7. Structural formulas of metal complexes which are affected by steric factors.

Finally, the ligand atom itself plays a significant role in controlling the stability of metal complexes. For most of the metal ions, the smallest ligand atom with the largest electron density will form the most stable complex. This means that the second period elements form more stable metal-ligand bonds than do other members of the same group, for example, N > P > As > Sb; O > S > Se > Te; F > Cl > Br > I. It is also known that for this same class of metal ions the stability order is N > O > F. These trends are in accord with the coordinate bond strengths expected on the basis of electrostatic interactions. These metal ions and the ligand atoms they prefer are not very polarizable, and are designated as hard Lewis acids and bases, respectively. A different trend in stability is found for a second class of metal ions, for example, Cu(I), Rh(I), Pd(II), Ag(I), Pt(II), and Hg(II). The stability trends observed for these metal ions are N < P > As > Sb, O < S > Se > Te, and F < Cl < Br < I. These metal ions and the ligand atoms they prefer are polarizable, and are designated as soft Lewis acids and bases, respectively. Thus, Pearson has proposed the HSAB theory of acids and bases, which states that hard Lewis acids prefer hard bases and soft Lewis acids prefer soft bases. Since the soft metal ions also have a large number of *d* orbital electrons and the third-period ligand atoms have vacant *d* orbitals, the enhanced stability is attributed to the formation of *dd*-π bonding. This means that the filled *d* orbitals of the metal ion overlap with the empty *d* orbitals of the ligand atom. This is illustrated by **Fig. 8**. This same

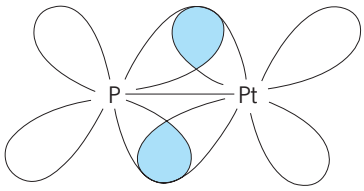

Fig. 8. Overlap of *d* orbitals of platinum with empty *d* orbitals of phosphorus ligand.

class of metal ions is the one that forms olefin complexes such as Zeise's salt, $K[Pt(CH_2{=}CH_2)Cl_3]$. The stability of these compounds is likewise attributed to π bonding.

Stability and reactivity. Often the most stable complex is also the least reactive or most inert. Therefore, the platinum metal complexes as well as cobalt(III) and chromium(III) complexes are usually very slow to react. However, there are exceptions to this, for instance, the rapid exchange of radiocyanide with the coordinated cyanide in the stable tetracyano complexes $Hg(CN)_4^{2-}$ and $Ni(CN)_4^{2-}$. Several factors, such as the electronic configuration of the central metal ion, its coordination number, and the extent of chelation, all have a marked effect on the rate of reaction of a given compound.

There are sufficient data available to permit a fairly reliable classification for the relative reactivities of six-coordinated systems. Metal complexes are said to be labile if they have reacted completely at room temperature within the time of mixing, such as the instantaneous formation of the deep-blue color of $Cu(NH_3)_4^{2+}$ upon the addition of excess ammonia to $Cu(H_2O)_4^{2+}$. Complexes are designated as being inert if their reactions at room temperature proceed at a detectable rate and have half-lives longer than 2 min. On the basis of such a definition, it is possible to classify the labile complexes as being those of either the outer orbital type, such as CoF_6^{3-}, $Al(C_2O_4)_3^{3-}$, $SnCl_6^{2-}$, $Ni(NH_3)_6^{2+}$, or the inner orbital type with one or more vacant *d* orbitals, such as TiF_6^{2-}, $V(CN)_6^{3-}$, $Mo(SCN)_6^{2-}$. The inert complexes are then the inner orbital type with no vacant *d* orbitals; for example, $Co(NH_3)_6^{3+}$, $Cr(H_2O)_6)_6^{3+}$, $RhCl_6^{3-}$. A similar classification based on the crystal field theory is done in terms of the labile systems being those that are not stabilized greatly by the crystal field.

Reactions of complexes. The reactions referred to above are acid-base reactions, where one basic ligand replaces a less basic ligand coordinated to the acidic metal ion, for example, reaction (4). A reac-

$$Co(NH_3)_5Cl^{2+} + H_2O \rightarrow Co(NH_3)_5H_2O^{3+} + Cl^- \qquad (4)$$

tion of this type is called a nucleophilic substitution reaction and is designated by the symbol S_N. There are at least two fundamentally different pathways conceivable for such a reaction. These are the dissociation and displacement mechanisms which are designated as S_N1 and S_N2, respectively. An S_N1 reaction goes by two-step process, where the first step is a slow unimolecular heterolytic dissociation, reaction (5), followed by the rapid coordination of

$$Co(NH_3)_5Cl^{2+} \xrightarrow{\text{slow}} Co(NH_3)_5^{3+} + Cl^- \qquad (5)$$

the entering group, reaction (6). An S_N2 reaction is

$$Co(NH_3)_5^{3+} + H_2O \xrightarrow{\text{fast}} Co(NH_3)_5H_2O^{3+} \qquad (6)$$

one involving a bimolecular rate-determining step in

which one nucleophilic reagent displaces another [reaction (7)]. Examples of both types of substitution are known for the reaction of metal complexes. The specific reaction just cited, as well as analogous reactions of $Co(NH_3)_4Cl_2^+$ and $Co(NH_3)_4OHCl^+$, appears to proceed by an S_N1 mechanism. However, reactions of platinum(II) complexes are believed to involve an S_N2 process, for example, reaction (8).

$$Co(NH_3)_5Cl^{2+} + H_2O \xrightarrow{\text{slow}} H_2O \cdots Co(NH_3)_5 \cdots Cl^{2+} \xrightarrow{\text{fast}} Co(NH_3)_5H_2O^{3+} + Cl^- \quad (7)$$

$$Pt(NH_3)_3Cl^+ + NO_2^- \xrightarrow{\text{slow}} O_2N \cdots Pt(NH_3)_3 \cdots Cl \xrightarrow{\text{fast}} Pt(NH_3)_3NO_2^+ + Cl^- \quad (8)$$

The symbols AID were introduced in order to better designate the mechanisms of these reactions. A = association and relates to S_N2, where the entering group becomes associated with the complex [slow step of Eq. (8)]; I = interchange and signifies that a group just outside the complex interchanges positions with a ligand on the metal; D = dissociation and relates to S_N1, where the leaving group dissociates away from the complex [Eq. (5)]. In general, I_d is similar to S_N1 and designates a dissociative interchange mechanism, whereas I_a is similar to S_N2 and designates an associative interchange mechanism.

Finally, coordination may have the effect of stabilizing an unusual oxidation state of the coordinated metal. For example, ligands such as CO, CN^-, CNR, 2,2′-bipyridine, and PR_3 are most effective in stabilizing low-valent states. Thus, compounds of the type $Na_2[Fe(CO)_4]$, $Na[V(2,2'\text{-bipyridine})_3]$, $K_4[Ni(CN)_4]$, and $Pt(PR_3)_4$ are known where the assigned oxidation states are Fe (2−), V (1−), Ni (0), and Pt (0). It is believed that the low oxidation states in these systems are stabilized by the formation of π-type molecular orbitals between the metal and ligand atoms. If the ligand orbitals are of low energy and vacant, they can accommodate the *d*-orbital electrons of the metal and permit the addition of other electrons leading to the reduction of the complex. In contrast, ligands such as O^{2-}, OH^-, and F^- tend to stabilize high-valence states giving complexes of the type K_2FeO_4, Na_3BiO_4, K_2NiF_6, and K_3CuF_6, where the oxidation states are Fe (6+), Bi (5+), Ni (4+), and Cu (3+). These ligands have no *d* orbitals but only filled *p* orbitals. One explanation for the stabilization of higher-valent states by such ligands is that in these systems the π-type molecular orbitals cannot be formed so the metal electrons are forced into higher-energy antibonding orbitals from which they may be readily removed. *See* CHEMICAL BONDING; MAGNETOCHEMISTRY; SOLID-STATE CHEMISTRY; STEREOCHEMISTRY; SUBSTITUTION REACTION.

Fred Basolo; John L. Burmeister

Bibliography. F. Basolo and J. L. Burmeister (eds.), *On Being Well-Coordinated: A Half-Century of Research on Transition Metal Complexes: Selected Papers of Fred Basolo*, 2003; F. Basolo and R. G. Pearson, *Mechanisms of Inorganic Reactions*, 2d ed., 1967; F. A. Cotton et al., *Advanced Inorganic Chemistry: A Comprehensive Text*, 6th ed., 1999; L. Pauling and Z. S. Herman, Valence-bond concepts in coordination chemistry and the nature of metal-metal bonds, *J. Chem. Educ.*, 61(7):582–587, July 1984; G. E. Rodgers, *Introduction to Coordination, Solid State, and Descriptive Organic Chemistry*, 2d ed., 2002; A. F. Williams, C. Floriani, and A. S. Merbach (eds.), *Perspectives in Coordination Chemistry*, 1992.

Coordination complexes

A group of chemical compounds in which all or a part of the molecular bonding is of the coordinate covalent type.

This article summarizes the different types of compounds that are known and discusses their nomenclature, structure, stereochemistry, and synthesis. For a discussion of the nature of the coordinate bond and the stability and reactivity of complex compounds *see* CHELATION; COORDINATION CHEMISTRY.

Complex compounds contain a central atom or ion and a group of ions or molecules bonded to it. Many simple hydrates, such as $MgCl_2 \cdot 6H_2O$, are best formulated as $[Mg(H_2O)_6]Cl_2$ because it is known that the six molecules of water are bound to the central magnesium ion. Therefore, $[Mg(H_2O)_6]^{2+}$ is a complex ion, and $[Mg(H_2O)_6]Cl_2$ is a complex compound. The charge on this complex ion is 2+, because this is the charge on the magnesium ion and the coordinated water molecules are neutral. However, if the coordinated groups are charged, then the charge on the complex is represented by the sum of the charge on the metal and that of the coordinated ions. See, for example, the progression of charges on the platinum(IV) complexes listed below. Thus the charge is 4+ for $[Pt(NH_3)_6]^{4+}$ because Pt is 4+ and NH_3 is neutral. But the charge is 2− for $[PtCl_6]^{2-}$ because of the 6 Cl^-, that is, $+4 - 6 = -2$ (that is, 2−).

Nomenclature. The metal complex may be a cation, have zero charge, or be an anion, as is exemplified by the following series of complexes:

$[Pt(NH_3)_6]Cl_4$	Hexaammineplatinum(IV) chloride
$[Pt(NH_3)_5Cl]Cl_3$	Chloropentaammineplatinum(IV) chloride
$[Pt(NH_3)_4Cl_2]Cl_2$	Dichlorotetraammineplatinum(IV) chloride
$[Pt(NH_3)_3Cl_3]Cl$	Trichlorotriammineplatinum(IV) chloride
$[Pt(NH_3)_2Cl_4]$	Tetrachlorodiammineplatinum(IV)
$K[Pt(NH_3)Cl_5]$	Potassium pentachloroammineplatinate(IV)
$K_2[PtCl_6]$	Potassium hexachloroplatinate(IV)

These compounds are named according to rules set up by the Nomenclature Committee of the International Union of Pure and Applied Chemistry. Some of the rules are the following: (1) Name the cation first as one word followed by the anion as one word. (2) For a cationic complex, name the negative ligands first with an ending of -o, then the neutral ligands, and finally the metal followed by a Roman numeral in parentheses to designate its oxidation state. Neutral ligands are named as the molecule, except that H_2O is aqua and NH_3 is ammine. (3) The prefixes such as di, tri, and tetra are used before simple names such as chloro, aqua, and oxalato. Prefixes such as bis, tris, and tetrakis are used before complex names such as ethylenediamine, 2,2′-bipyridine, and trialkylphosphine. (4) Neutral complexes are named in the same way except that only one word is required. (5) Anionic complexes are also named according to these same rules except that an -ate ending is used. Additional rules are available to name the more complicated compounds, and some of these will be used in the discussion that follows.

Coordination numbers of 2–10 have been observed for different complex compounds. The most common coordination numbers are 6 and 4. Complexes having coordination number 6 generally have an octahedral structure, but may also be trigonal prismatic. Complexes having coordination number 4 are either square planar or tetrahedral. **Table 1** gives examples of complexes having coordination numbers other than 4 or 6.

Structures. Metal complexes exhibit various types of isomerism. In many ways, inorganic stereochemistry is similar to that observed with organic compounds. Geometrical isomers are common among the substitution inert complexes of coordination numbers 4 and 6. Square planar complexes of the type $Pt(NH_3)_2Cl_2$ exist in two forms, cis and trans. Likewise, cis-trans isomers of the $Pt(NH_2CH_2COO)_2$ type have been isolated (**Fig. 1**). It is apparent from the above examples that the isomer with the same ligands or ligand atoms in adjacent positions is called cis, whereas the trans isomer has its like groups in opposite positions. There are also examples of geometrical isomers of complexes containing four different ligands (**Fig. 2**). Although geometrical isomerism of square complexes is most common with platinum(II), it has also been observed with compounds of nickel(II), palladium(II), and gold(III). *See* STEREOCHEMISTRY.

There are many examples of geometrical isomers for 6-coordinated complexes of cobalt(III), chromium(III), and the platinum metals. Most of the examples are of the type $[Co(NH_3)_4Cl_2]^+$ (**Fig. 3**). If three of the ligands differ from the other three, then only two isomers are possible (**Fig. 4**). Similar isomers are obtained with an unsymmetrical bidentate ligand (**Fig. 5**). The complex with 6 different ligands, such as $[Pt(NH_3)(py)(NH_2OH)(Cl)(Br)(NO_2)]$,

TABLE 1. Metal complexes with coordination numbers other than 4 or 6

Coord. number	Complex	Structure
2	$[Ag(CN)_2]^-$ $[Ag(NH_3)_2]^+$	Linear
3	$[Cu(PR_3)_2I]$ $[Au(AsR_3)_2I]$	Planar
5	$[Fe(CO)_5]$ [Zn(2,2′,2″-terpyridine)Cl_2]	Trigonal bipyramid
	$[Ni(PR_3)_2Br_3]$	Tetragonal pyramid
	{Pd[o-phenylene-bis-(dimethylarsine)]$_2$I}I	
7	ZrF_7^{3-}	Pentagonal bipyramid
	NbF_7^{2-}	Face-centered trigonal prism
	TaF_7^{2-}	
8	$Mo(CN)_8^{4-}$ TaF_8^{3-}	Dodecahedral Archimedean antiprism
9	$Nd(H_2O)_9^{3+}$	Face-centered trigonal prism
10	$Cd_2[Mo(CN)_8(H_2O)_2]$	Unknown structure

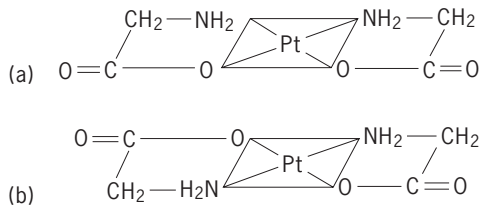

Fig. 1. Structure of (*a*) *cis*-diglycinatoplatinum(II) and (*b*) *trans*-diglycinatoplatinum(II).

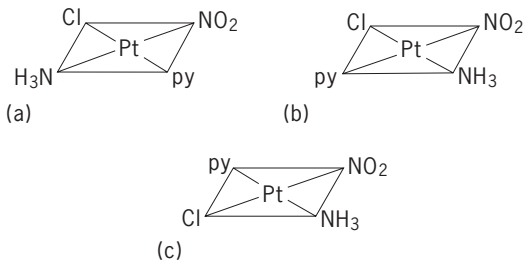

Fig. 2. Isomeric complexes. (*a*) *trans*-Chloropyridine-*trans*-nitroammineplatinum(II). (*b*) *trans*-Chloroammine-*trans*-nitropyridineplatinum(II). (*c*) *trans*-Chloronitro-*trans*-amminepyridineplatinum(II).

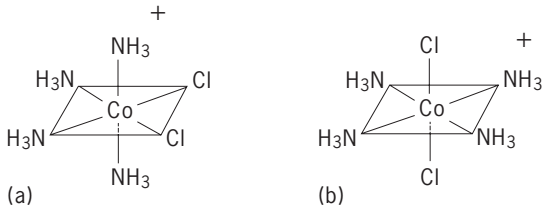

Fig. 3. Structure of (*a*) *cis*-dichlorotetraamminecobalt (III) ion and (*b*) *trans*-dichlorotetraamminecobalt(III) ion.

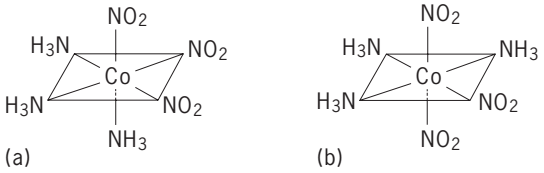

Fig. 4. Structure of (*a*) *cis*- (or facial-) trinitrotriamminecobalt(III) and (*b*) *trans*- (or meridial-) trinitrotriamminecobalt(III).

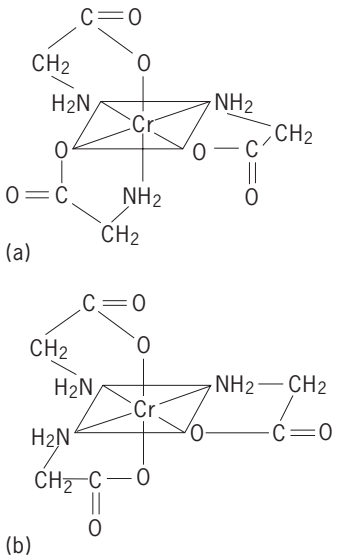

Fig. 5. Structure of (*a*) *fac*-triglycinatochromium(III) and (*b*) *mer*-triglycinatochromium(III).

can exist theoretically in 15 different geometrical forms.

Optical isomerism is also fairly common among these compounds. For example, 4-coordinate tetrahedral complexes containing unsymmetrical bidentate ligands, such as bis(benzoylpyruvato)beryllate (II) ion, have been resolved (**Fig. 6**). Optically active complexes of this type are likewise reported for boron(III) and for zinc(II). Most of the examples of optical activity occur with 6-coordinated systems containing three bidentate ligands, for example, trioxalatorhodiate(III) ion (**Fig. 7**). The resolution of 6-coordinated complexes of this type has been reported for the metal ions Al(III), As(V), Cd(II), Co(III), Cr(III), Ga(III), Ge(IV), Ir(III), Fe(II),(III), Ni(II), Os(II),(III), Pt(IV), Rh(III), Ru(II),(III), Ti(IV), and Zn(II).

The cis isomer of a complex containing two bidentate ligands and two monodentate ligands is asymmetric, and therefore exists in the form of mirror-image isomers, whereas the trans form is symmetrical and cannot be optically active (**Fig. 8**). Since the cis isomer of this type is optically active and the trans isomer is not, the successful resolution of one of the isomers has historically been used as a proof of its cis structure. Other types of isomerism are known for metal complexes (**Table 2**).

Synthesis. The synthesis of metal complexes containing only one kind of ligand generally involves simply the reaction of the metal salt in aqueous solution with an excess of the ligand reagent, reaction (1).

$$[Ni(H_2O)_6](NO_3)_2 + 6NH_3 \longrightarrow [Ni(NH_3)_6](NO_3)_2 + 6H_2O \quad (1)$$

The desired complex salt can then be isolated by removal of water until it crystallizes, or by addition of a water-miscible organic solvent to cause it to separate. Many reactions, such as the one cited above, occur readily at room temperature. For the substitution-inert complexes (those slow to react), prolonged treatment at more drastic conditions is often necessary.

The preparation of geometrical isomers is much more difficult, and in most cases, the approach used

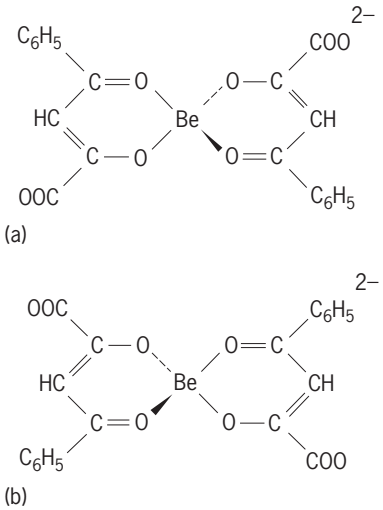

Fig. 6. Optical isomers of bis(benzoylpyruvato)beryllate(II) ion: (*a*) dextro and (*b*) levo.

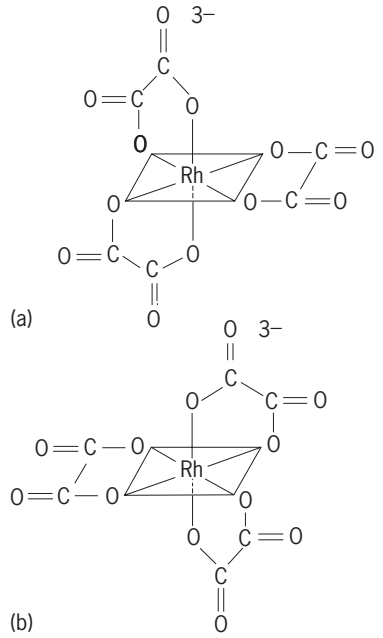

Fig. 7. Optical isomers of trioxalatorhodiate(III) ion: (*a*) lambda and (*b*) delta.

is rather empirical. Generally, reactions yield a mixture of cis-trans products, and these are separated on the basis of their differences in solubility. The trans isomers of platinum(IV) complexes are prepared by the oxidation of the appropriate platinum(II) compound (**Fig. 9**). The cis isomers of cobalt(III) complexes can sometimes be prepared by the reaction of a carbonato complex with the desired acid (**Fig. 10**).

The easiest geometric isomers to prepare are those of platinum(II). For example, the cis and the trans isomers of $[Pt(NH_3)_2Cl_2]$ are prepared as shown in

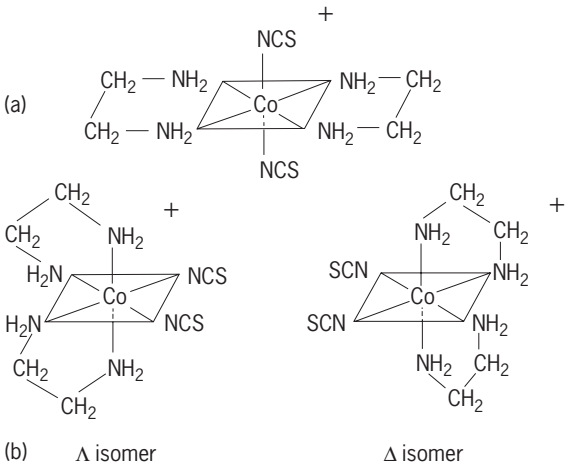

Fig. 8. Symmetric trans form of a complex and asymmetric cis structures. (*a*) *trans*-bis(isothiocyanato) bis(ethylenediamine)cobalt(III) ion. (*b*) Asymmetric forms of lambda, delta; *cis-bis*(isothiocyanato) bis(ethylenediamine) cobalt(III) ion.

TABLE 2. Types of isomerism for complex compounds

Isomerism	Examples
Geometrical	*cis*- and *trans*-$[Pt(NH_3)_2Br_2]$
Optical	Δ-and Ω-$[Ir(NH_2CH_2CH_2NH_2)_3]Cl_3$
Coordination	$[Co(NH_3)_6][Cr(CN)_6]$ and $[Cr(NH_3)_6][Co(CN)_6]$
Coordination position	$[(R_3P)_2Pt(Cl)_2PtCl_2]$ and $[Cl(R_3P)Pt(Cl)_2Pt(PR_3)Cl]$
Hydrate	$[Cr(H_2O)_6]Cl_3$ and $[Cr(H_2O)_5Cl]Cl_3 \cdot H_2O$
Ionization	$[Pt(NH_3)_4Cl_2]SO_4$ and $[Pt(NH_3)_4SO_4]Cl_2$
Linkage	$[(NH_3)_5Co—NO_2]SO_4$ and $[(NH_3)_5Co—ONO]SO_4$
Polymerization	$[Co(NH_3)_3(NO_2)_3]$ and $[Co(NH_3)_6][Co(NO_2)_6]$
Conformation	$[Ni(P(C_6H_5)_2CH_2C_6H_5)_2Br_2]$ square-planar and tetrahedral

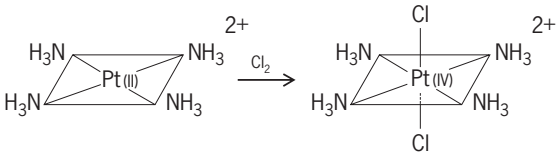

Fig. 9. Reaction in the trans isomer formation of platinum(IV) complexes.

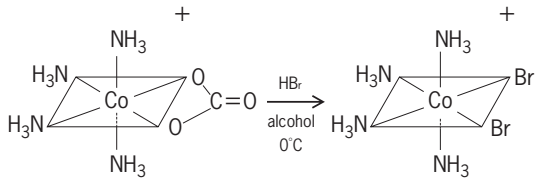

Fig. 10. Cis isomer formation of cobalt(III) complexes.

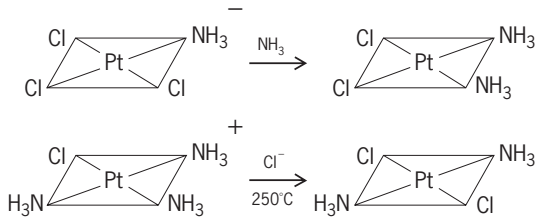

Fig. 11. Cis and trans isomer formation of platinum(II).

reactions (2). The second step in each of these reac-

$$K_2[PtCl_4] + 2NH_3 \longrightarrow cis\text{-}[Pt(NH_3)_2Cl_2] + 2KCl$$
$$[Pt(NH_3)_4]Cl_2 \xrightarrow[(392°F)]{250°C} trans\text{-}[Pt(NH_3)_2Cl_2] + 2NH_3 \qquad (2)$$

tions results in the replacement of the ligand trans to a chloro group (**Fig. 11**). This phenomenon, that a negative ligand often has a greater labilizing effect on a group in the trans position than does a neutral group, for example, $Cl^- > NH_3$, is called the trans effect. Extensive use has been made of this trans effect in the synthesis of desired platinum(II) complexes. The complex *cis*-$[Pt(NH_3)_2Cl_2]$ (Fig. 11) is used as an antitumor drug for certain types of cancer.

Finally, the separation of optical isomers of metal complexes involves techniques similar to those used for organic compounds. The usual procedure is to convert the racemic mixture into diastereoisomers by means of an optically active resolving agent and then to separate the diastereoisomers by fractional crystallization. Nonionic complexes have been resolved by preferential adsorption on optically active quartz or sugars. *See* AMMINE; HYDRATE.

Fred Basolo; John L. Burmeister

Bibliography. F. Basolo and R. G. Pearson, *Mechanisms of Inorganic Reactions*, 2d ed., 1967; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; D. Katakis and G. Gordon, *Mechanisms of Inorganic Reactions*, 1987; S. Kawaguchi, *Variety in Coordination Modes of Ligands in Metal Complexes*, 1988; D. F. Shriver, P. W. Atkins, and C. Langford, *Inorganic Chemistry*, 3d ed., 1999.

Copepoda

A subclass of Crustacea with eight orders. There are over 6000 species, most of which are free-living in aqueous environments, though many are parasitic or symbiotic with other aquatic animals. The free-living forms are the most abundant of all animals in the sea and directly constitute the main food for vast numbers of fishes, many invertebrates, and also at times for baleen whales.

Structure. The structure of copepods, although similar in basic pattern, varies greatly. In free-living forms the body is composed externally of a series of small chitinous cylinders, or segments, some of which may be more or less fused. The head bears two pairs of antennae. The first antennae are used for stabilizing, retarding sinking, olfactory functioning, and usually in the male for grasping the female. The second antennae are used for creating water currents for food straining and slow locomotion in the Calanoida, and for food grasping in the Cyclopoida and Harpacticoida. The feeding appendages are one pair of mandibles, two pairs of maxillae, and one pair of maxillipeds. During the nauplius larval stage, the second antennae are commonly provided with masticatory hooks for feeding. The thorax, consisting of six segments, bears the swimming feet. The genital aperture occurs on the sixth thoracic segment. The abdomen consists of one to five segments and bears no appendages except a pair of terminal processes known as caudal rami. The sexes are separate and sexual dimorphism is frequently evidenced, especially by modification in the male by one or both of the first antennae for clasping, and in many by marked specialization in the fifth pair of feet for the transfer of spermatophores, or packets of spermatozoa, to the female. Free-living species usually range in size from less than 0.02 in. (0.5 mm) to about 0.4 in. (10 mm) in length. A deep oceanic species, *Bathycalanus sverdrupi,* attains a length of 0.68 in. (17 mm). Some parasitic copepods, such as *Penella,* are greatly elongate, reaching up to 10 in. (250 mm) in length. *See* SEXUAL DIMORPHISM.

Classification. Seven orders of copepods are usually recognized. Three of these, Calanoida, Cyclopoida, and Harpacticoida, are wholly or primarily free-living and because of overwhelming numbers are of greatest general interest. The anatomical differences that set these three apart are, among others, the shape of the body and the position of the movable articulation separating the forebody (metasome) from the hindbody (urosome). In the calanoids and the cyclopoids the urosome is distinctly narrower than the metasome (**illus.** *a* and *b*). In the calanoids the movable articulation falls between the fifth and sixth thoracic segments, whereas in the cyclopoids, it is between the fourth and fifth segments. In the harpacticoids the urosome is commonly not narrow (illus. *c*), and the movable articulation falls between the fourth and fifth thoracic segments. In many harpacticoids there is considerable flexibility also between other segments. The length of the first antennae relative to body length is very useful in distinguishing these three orders. The first antennae in the Calanoida have 20–25 segments, usually extending to or beyond the caudal rami. In the Cyclopoida they have 6–17 segments, extending to the posterior border of the metasome or even the first body segment. In the Harpacticoida they have 5–9 segments, usually not extending to the posterior border of the first body segment. *See* CALANOIDA; CYCLOPOIDA; HARPACTICOIDA.

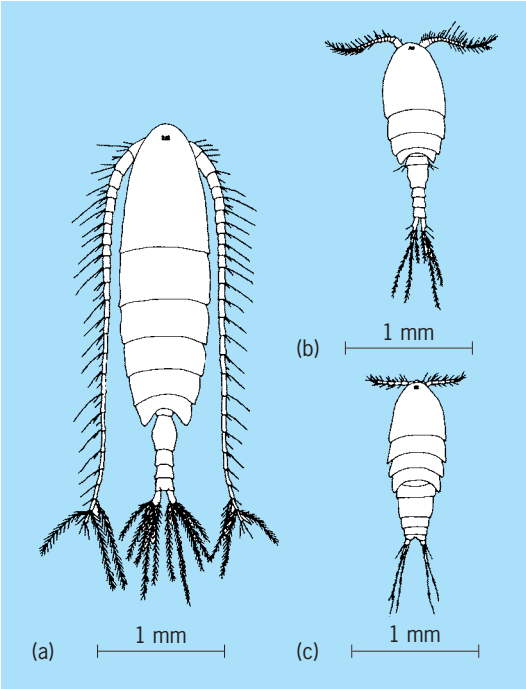

Free-living copepods. (*a*) *Calanus finmarchicus*, a calanoid. (*b*) *Cyclops*, a cyclopoid. (*c*) *Harpacticus*, a harpacticoid.

In the adult state the remaining orders (Poecilostomatoida, Siphonostomatoida, Monstrilloida, Misophrioida, and Mormonilloida) are parasitic at least part of their lives. Many adult stages are so greatly modified that their crustacean relationship is obscure. Their numbers are relatively small, but they may be conspicuous and important as "fish lice," burrowing deep into the flesh or feeding externally upon the skin and gills. *See* MISOPHRIOIDA; MONSTRILLOIDA; MORMONILLOIDA; POECILOSTOMATOIDA; SIPHONOSTOMATOIDA.

Life history. The life histories of copepods differ in details, but in nearly all, development takes place from fertilized eggs which are either attached to the genital segment or are cast free in the water. The spermatozoa, transferred to the female genital segment in spermatophores, are stored in spermathecal sacs in the female. Usually one mating suffices for all subsequent eggs laid by the female. A nauplius larva hatches from the egg. This larva usually undergoes five molts in its growth. Most nauplii are self-sustaining, gathering their own food, but some nauplii hatched from yolky eggs are sustained by

the yolk through one or more stages. The sixth naupliar stage metamorphoses into the first of six copepodid stages, the last of which is the adult. Only a few life histories have been studied completely. In *Calanus finmarchicus,* a calanoid, a summer brood in boreal waters requires about 2–3 months to reach maturity, but broods produced in the autumn live longer because they winter in deep water as subadults. The number of broods produced is related to latitude, fewer being produced in the far north.

In *Tisbe furcata*, a bottom-living harpacticoid, individual females may live for 70 days, producing 5–10 broods averaging 43 eggs per brood. The young may produce the first batch of eggs in 15 days. Because of their high fecundity and early sexual maturity certain harpacticoids have proved useful for genetic studies.

Habitats. Copepod habitats are extremely varied, but the vast majority of species live in the sea. They are mostly planktonic throughout their lives, some occurring only in coastal waters while others live in the open sea. Many frequent only the upper layers to about 650–1000 ft (200–300 m) depth. Others are bathypelagic at great depths. The harpacticoids, however, are primarily a benthic, or bottom-dwelling, group, though many are truly planktonic. A great many harpacticoid species live upon aquatic vegetation to which they cling with tiny claws. Others are found on the bottom feeding upon detritus. Some are especially adapted by slender wormlike bodies to burrow or creep into the bottom among the grains of sand. There are also many fresh-water copepods living in ponds, lakes, and streams in all parts of the world. Their wide dispersal to isolated bodies of water is probably brought about largely through transport on birds' feet. A few copepods live naturally in damp moss.

Economic biology. The role of copepods in the economy of the sea consists mainly of converting plant into animal substance. Many species are adapted to graze directly upon the chief synthesizers of organic material, the microscopic diatoms and dinoflagellates that constitute the great pasturage of the sea. Some species are predatory on other copepods, fish larvae, and other small animals, such as protozoa. Commonly copepods constitute about 70% of the zooplankton, and at times their numbers are so great as to impart a pinkish color to the sea. They are of sufficient size to be fed upon individually by young fishes and to be filtered from the water by the filtering apparatus (gill rakers) of such fishes as the herring, sardines, and others that require animal protein as food. At higher trophic levels these fishes in turn are, directly or indirectly, food for the more predatory fishes and mammals. *See* MARINE ECOLOGY.

Harry C. Yeatman

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; C. B. Wilson, *Copepods of the Woods Hole Region,* U.S. Nat. Mus. Bull. 158, 1932.

Copolymer

A macromolecule in which two or more types of monomer are incorporated into the polymer chain. This contrasts with a homopolymer, which consists of only one type of monomer. Copolymerization offers the ability to alter the properties of a homopolymer by introducing an appropriately chosen second repeating unit, or to combine the desirable properties of two different homopolymers in a single molecule. The properties of a copolymer depend on the nature of the monomers and their distribution in the chain. Thus, monomers A and B can polymerize randomly (ABBAABA), they can alternate (ABAB), they can form blocks (AAABBB), or one polymer can be grafted onto another polymer (**Fig. 1**). Copolymers can be prepared by all the known methods of polymerization, including addition polymerization of vinyl monomers (by conventional or living free-radical, anionic, cationic, or coordination catalysis), ring-opening polymerization, or condensation polymerization. Copolymers can also be formed by the chemical reaction of functional groups (usually end groups) on the constituent polymer chains. *See* FREE RADICAL; MACROMOLECULAR ENGINEERING; POLYMER; POLYMERIZATION; RING-OPENING POLYMERIZATION.

Statistical copolymers. The simplest method of producing a copolymer is to polymerize a mixture of two (or more) different monomers. Assuming that they are both polymerizable via the same mechanism, the different types of repeat unit will be incorporated into the polymer chain according to their concentration in the monomer feed and their relative reactivities. If the comonomers (A and B) have similar reactivities, a random distribution of the alternative repeat units (BABBAABA) is obtained. If A prefers to add to B and B prefers to add to A, an alternating sequence (ABABABAB) is expected. If A prefers to add to A (and/or B prefers to add to B), one might expect a greater concentration of AAAA or BBBB blocks, or in the extreme case no copolymer at all but a mixture of homopolymers. In reality, the copolymer will have a composition and sequence distribution somewhere between these extremes that can be manipulated by varying the ratio of the comonomers in the feed and by varying the reactivities themselves (for example, by varying the temperature, catalyst, or solvent).

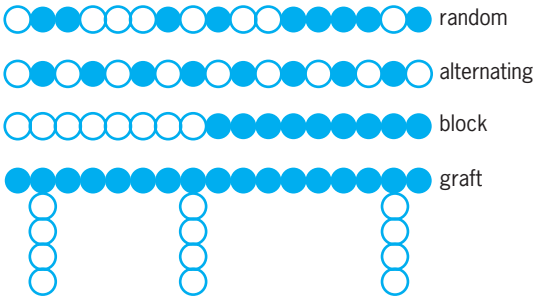

Fig. 1. Different types of copolymers.

Copolymerization models are equations for predicting the composition, sequence distribution, and polymerization kinetics as a function of the monomer feed fractions and a small set of characteristic constants. One of the simplest models is the terminal model. It is assumed that for a given type of polymerization mechanism (radical, anionic, cationic, metallocene-catalyzed, and so on), the reactivity of the propagating species depends only on the nature of its terminal monomer unit. As shown below, there are four types of reaction to consider, with the reacting chains ending in either monomer unit (A* or B*) with each of the comonomers (A or B).

$$\sim\!\sim A^* + \begin{array}{l} A \xrightarrow{k_{AA}} \sim\!\sim AA^* \\ B \xrightarrow{k_{AB}} \sim\!\sim AB^* \end{array} \qquad r_A = \frac{k_{AA}}{k_{AB}}$$

$$\sim\!\sim B^* + \begin{array}{l} A \xrightarrow{k_{BA}} \sim\!\sim BA^* \\ B \xrightarrow{k_{BB}} \sim\!\sim BB^* \end{array} \qquad r_B = \frac{k_{BB}}{k_{BA}}$$

The reactivity ratios r_A and r_B measure the selectivity of each propagating species toward the alternative monomers, and can be determined experimentally by fitting the terminal model equations to composition or sequence distribution data. Hundreds of reactivity ratios have been measured in this way, and numerous studies have attempted to explain their values in terms of the chemical structure. In recent years, it has been shown that the terminal model rarely holds, and much more complicated copolymerization models are generally required. In radical copolymerization, for instance, models that include additional factors, such as the influence of the penultimate unit of the propagating species, the depropagation reaction, or various types of solvent effects, have been introduced to describe various systems.

Tapered copolymers. An interesting feature of statistical copolymerization is the possibility of composition drift. If the comonomers are consumed at different rates, the comonomer feed ratio will vary through the polymerization process. As a result, the composition of the copolymer will also vary as a function of conversion. In a nonliving process, such as conventional free-radical copolymerization, chains are initiated and grow at different stages of the reaction, resulting in compositional variation among the polymer chains. In contrast, in a living process (such as ionic, coordination-catalyzed, or living free-radical copolymerization) the individual chains are initiated together and continue to grow over the entire polymerization process. As a result, changes in the instantaneous copolymer composition are frozen into the growing polymer chain, and a tapered composition results. For instance, if A is a more reactive monomer, the polymer is initially richer in A; but since monomer A is consumed at a faster rate, the polymer becomes progressively poorer in A. As a result, one end of the polymer chain is richer in A than the other end. Depending on the reactivity ratios and the initial monomer feed fractions, a variety of interesting tapered compositions can be produced. Tapered, or gradient, copolymers have properties that are different from either block copolymers or random copolymers, and are particularly useful as blend compatibilizers and vibration-dampening materials.

Random copolymers. Random copolymers were among the first copolymers produced, and remain to this day among the most commercially important as they offer a means of improving the properties of a homopolymer via the incorporation of a small amount of a second type of repeat unit into the polymer chain. The first commercially important copolymers were developed in Germany in the 1930s, where it was discovered that the properties of synthetic rubber could be markedly improved via the copolymerization of butadiene with small amounts of other monomers. Buna N (now known as nitrile rubber or ABR) is a copolymer of butadiene with acrylonitrile, and combines the elastomeric properties of rubber with the solvent resistance of polyacrylonitrile, while Buna S (now known as SBR) is a copolymer of butadiene with styrene that has improved mechanical properties. Both copolymers continue to find widespread commercial use; in particular, SBR (reinforced with carbon black) is used in vehicle tires and still accounts for approximately 40% of the world market for elastomers. *See* ACRYLONITRILE; CARBON BLACK; POLYACRYLONITRILE RESINS; RUBBER; STYRENE.

Countless other applications for random copolymers exist. Low-density polyethylene is a highly branched homopolymer of ethylene obtained via conventional radical polymerization. Various properties of this homopolymer can be improved through copolymerization with small amounts of other monomers, including vinyl acetate which is used to improve clarity and flexibility for food packaging applications, acrylate or methacrylate monomers which can improve thermal stability for applications in cable insulation and medical packaging, and acrylic or methacrylic acids which offer improved abrasion resistance, flexibility, and adhesion for applications such as packing film and laminates. *See* POLYOLEFIN RESINS.

Emulsions of polymers based on acrylic or methacrylic esters are used extensively in adhesives, caulks, paints, and other surface coatings. A significant number of them are copolymers of acrylic and methacrylic esters, and/or combinations of ester and acid monomers. The copolymer composition is varied to optimize the gloss, adhesion, flexibility, impact strength, and other coating properties. *See* EMULSION POLYMERIZATION.

Alternating copolymers. Alternating copolymers are obtained in a statistical copolymerization if the cross-propagation reactions are very strongly favored over the respective homopolymerizations (that is, r_A and

$r_B \approx 0$). In a radical copolymerization, this will be the case if the comonomers are pairs of strongly electron-accepting and electron-donating monomers (such as styrene with acrylonitrile). Unless both reactivity ratios are exactly zero, the ratio of the comonomers also needs to be finely balanced. Alternating copolymers can also be obtained if just one of the comonomers does not readily homopolymerize (such as SO_2 or maleic anhydride), provided the polymerization is carried out in a large excess of this monomer. Pairs of monomers (such as styrene with methyl methacrylate) that might naturally form random copolymers can sometimes be forced to alternate through complexation of one of the comonomers with a Lewis acid. The complexation alters the reactivity of the propagating species and favors the cross-propagation reaction.

Some commercially important examples of alternating copolymers include the propylene-butadiene elastomer, which has improved mechanical properties and resistance to oxidative degradation compared with polybutadiene, and the vinylidene fluoride–hexafluoroisobutylene copolymer, which has improved abrasion resistance and better processability than the tetrafluoroethylene homopolymer (Teflon®). Copolymers of styrene with 10–40% acrylonitrile (SAN) are also of major commercial significance. Strictly speaking, they are statistical copolymers, with a high alternating tendency that peaks at the 50% copolymer composition. In SAN, the acrylonitrile improves the solvent resistance and tensile strength, compared with polystyrene, and raises the temperature of use. Alternating copolymers of chromophore units and saturated segments are showing promise as tunable electroluminescent devices as, unlike their fully conjugated counterparts (like polyacetylene), the conjugation length is well defined, and the optical properties can be more easily adjusted through variation of the chromophore. *See* CONJUGATION AND HYPERCONJUGATION.

Block copolymers. Copolymers containing block sequences can be prepared in a statistical copolymerization if both of the monomers show a tendency to homo-propagate rather than cross-propagate. However, it is difficult to prepare well-defined block copolymers in this way, since typically one monomer either yields statistical copolymers with blocks of varying length or (if there is no cross-propagation at all) a mixture of homopolymers.

There are three main strategies for preparing true block copolymers. Sequential addition of monomers to a living polymerization is perhaps the simplest method of block copolymer formation, but relies upon the two constituent monomers being polymerizable by the same mechanism. Until recently, this was a limitation, owing to the highly selective nature of ionic and coordination catalysts. However, with the development of controlled radical polymerization, a broader range of block copolymers can now be made. A second, more versatile approach is to prepare the first block as an end-functionalized macroinitiator that is added to the second monomer at a later stage. The advantage of this approach is that it is possible to prepare block copolymers from monomers that would not undergo polymerization via the same mechanism. For instance, step polymers (such as polyesters or polysulfones) can be converted to macroinitiators for atom transfer radical polymerization (ATRP) through treatment with 2-bromopropionyl bromide in the presence of pyridine. This introduces bromide end groups that, in conjunction with an appropriate ATRP catalyst, can be used to initiate controlled radical polymerization of styrene or acrylates, resulting in a BAB triblock copolymer. A third approach to preparing block copolymers is to chemically bond the constituent polymer chains as a postpolymerization reaction of functional end groups. *See* CONTROLLED/LIVING RADICAL POLYMERIZATION; POLYESTER RESINS; POLYSULFONE RESINS.

Block copolymers have a wide range of applications. Block copolymers of two otherwise immiscible polymers can function as compatibilizers, while block copolymers of hydrophobic and hydrophilic polymers can self-assemble into stable spherical or cylindrical micelles for use as emulsifiers, dispersants, and even drug delivery devices (**Fig. 2**). Block copolymers of hard (rigid) and soft (flexible) units function as thermoplastic elastomers (**Fig. 3**). *See* MICELLE.

In normal elastomers (such as rubber), the polymer chains are chemically bonded (cross-linked) together to prevent chains from slipping past each other under mechanical stress. This increases the strength of the polymer and allows for rapid recovery after deformation. However, the chemical cross-linking step is irreversible and has to be performed as an additional processing step after fabrication. In the thermoplastic elastomers, the cross-links are not chemical bonds but physical bonds. The hard blocks of these copolymers aggregate together to form rigid domains that act as physical cross-links to keep the rubbery phase in a network structure. At high temperature, the rigid domains melt allowing for reprocessing of the polymer. Thermoplastic elastomers are cheaper to process but offer less resistance to solvents and deformation at high temperatures. As a result, they tend to be used in applications such as wire insulation, footwear, wheels, and adhesives.

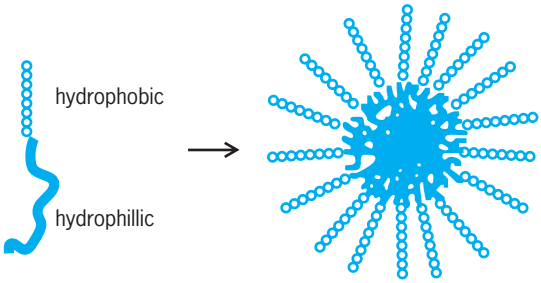

Fig. 2. Micelle formed from block copolymers.

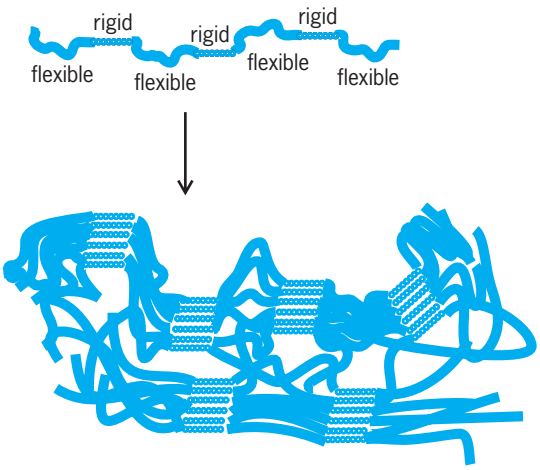

Fig. 3. Thermoplastic elastomer formed from block copolymers.

Graft copolymers. Graft polymers are a type of block copolymer in which one type of polymer chain is chemically attached to the side of the other chain (rather than the chain end). Typically, a single backbone supports several grafts, at either regular or random intervals along the chain. Highly grafted polymers (containing grafts at every repeat unit) are sometimes called bottle-brush polymers, while graft polymers with lower grafting densities are sometimes called comb polymers. *See* BRANCHED POLYMER.

As in the case of block-copolymer formation, grafting offers the possibility of introducing additional favorable properties to a homopolymer (such as improved mechanical properties or adhesion, or special additional functionalities such as antimicrobial action). For example, the impact resistance of polybutadiene (rubber) can be improved by grafting styrene onto it to make high-impact polystyrene, while grafting both styrene and acrylonitrile to make acrylonitrile-butadiene-styrene (ABS) polymers improves solvent resistance. Other commercially important examples include graft copolymers containing acrylic monomers which are used as pressure-sensitive adhesives, self-plasticizing poly(vinyl chloride) [PVC] which is produced by grafting poly(butyl acrylate) onto the PVC backbone, and blend compatibilizers which are obtained by grafting two otherwise immiscible polymers. When copolymers of hydrophilic and hydrophobic blocks are grafted onto a polymer backbone, they can function as unimolecular cylindrical micelles, and are of interest as potential nanoreactors and in drug delivery applications. *See* POLYVINYL RESINS.

Graft copolymers can be prepared using three main strategies: grafting onto, grafting from, and grafting through. In the grafting-onto approach, a preformed polymer with an active end group is attached onto the backbone of another preformed polymer at functional groups along the polymer chain. For example, polystyrene can be chlormethylated via treatment with chloromethyl ether in the presence of a catalyst, and then reacted with living polymeric chains. The grafting-from approach also begins with a polymer backbone but then grows the grafts from initiation sites along this backbone. These sites can be introduced to the backbone during its original preparation, via polymerization or copolymerization of a monomer that contains initiation sites. For example, a graft copolymer can be formed by the copolymerization of propylene with a monomer containing an alkoxyamine functional group. This functional group (which is inert at the temperature of the first polymerization) can then be used to initiate nitroxide-mediated polymerization of styrene in a controlled fashion. Alternatively, one can prepare a normal homopolymer and introduce the initiation sites as a separate reaction. For example, polystyrene can be treated with *N*-bromosuccinimide to brominate the polymer backbone at random intervals; the resulting bromide groups can then act as initiating sites for the polymerization of various vinylic monomers (by ATRP). *See* POLYSTYRENE RESIN.

In the grafting-through approach, one first prepares a macromonomer that is, a polymer with an end group that can be copolymerized with a second monomer. Macromonomers can be prepared using functional initiators or chain terminating agents. For example, polymers containing terminal olefinic bonds [carbon-carbon double bonds (C═C)] can be prepared in conventional radical polymerization in the presence of a catalytic chain-transfer agent. These monomers can then undergo radical copolymerization with a second olefinic monomer to yield graft copolymer.

Step copolymers. In a step polymerization process, a polymer is formed via the stepwise reaction of two different types of functional groups (such as an acid and an alcohol, or an acid and an amine) on multifunctional monomers. In many instances, the different types of functional groups are located on separate types of monomer (for example, a polyester is often made by reacting a diacid with a dialcohol), and a step homopolymer might be thought of as an alternating copolymer. However, in practice the entire ester linkage is defined as the repeating unit, and the term "copolymer" is reserved for polymers containing different types of repeat units.

As in the case of chain polymerization, statistical copolymers are formed from mixtures of monomers, such as two different types of acid with one or more types of alcohol, or two or more types of amine with one or more types of alcohol. For example, the crystallization and processing rates of poly(ethylene terephthalate) [PET] is improved by copolymerizing ethylene glycol and 1,4-butanediol with dimethyl terephthalate, while flame-retardant polycarbonates can be obtained by copolymerizing bisphenol A and its tetrabromo derivative.

In step polymerization, it is possible to make statistical copolymers not only from a mixture of monomers but also from a mixture of homopolymers. For example, if two different polyesters are

heated under vacuum they will transesterify, resulting in the scrambling of their repeat units. Depending on how far the reaction is driven, one can obtain a range of sequence distributions ranging from blocks to fully random copolymers. This scrambling via interchange makes it difficult to synthesize block copolymers from homopolymers containing the same type of functional group. However, as seen above, it is possible to produce block or graft copolymers involving step polymers and various types of chain polymers. One important example is the multiblock copolymers of polyurethanes and butadiene, which function as thermoplastic elastomers.

Michelle L. Coote

Bibliography. C. Barner-Kowollik et al., Copolymerization, in *Encyclopedia of Polymer Science and Technology*, 3d ed., ed. by J. I. Kroschwitz, vol. 9, pp. 394–445, Wiley, New York, 2004; J. M. G. Cowie (ed.), *Alternating Copolymers*, Plenum Press, New York, 1985; N. Hadjichristidis, *Block Copolymers: Synthetic Strategies, Physical Properties, and Application*, Wiley, Chichester, UK, 2002; I. W. Hamley (ed.), *Developments in Block Copolymer Science and Technology*, Wiley, Chichester, UK, 2004; K. Matyjaszewski and T. P. Davis (eds.), *Handbook of Radical Polymerization*, Wiley, New York, 2002; G. Odian, *Principles of Polymerization*, Wiley Interscience, New York, 1991.

Copper

A chemical element, Cu, atomic number 29, atomic weight 63.546. Copper, a nonferrous metal, is the twentieth most abundant element present in the Earth's crust, at an average level of 68 parts per million (0.22 lb/ton or 0.11 kg/metric ton). Copper metal and copper alloys have considerable technological importance due to their combined electrical, mechanical, and physical properties. The discoveries that mixed-valence Cu(II)/Cu(III) oxides exhibit superconductivity (zero electrical resistance) at temperatures as high as 125 K (−234°F; liquid nitrogen, a cheap coolant, boils at 90 K or −297°F) have generated intense international competition to understand these new materials and to develop technological applications. Although some pure copper metal is present in nature, commercial copper is obtained by reduction of the copper compounds in ores followed by electrolytic refining. The rich chemistry of copper is restricted mostly to the valence states Cu(I) and Cu(II); compounds containing Cu(0), Cu(III), and Cu(IV) are uncommon. Soluble copper salts are potent bacteriocides and algicides at low levels and toxic to humans in large doses. Yet copper is an essential trace element that is present in various metalloproteins required for the survival of plants and animals. *See* COPPER ALLOYS.

Copper is located in the periodic table between nickel and zinc in the first row of transition elements and in the same subgroup as the other so-called coinage metals, silver and gold. The electronic

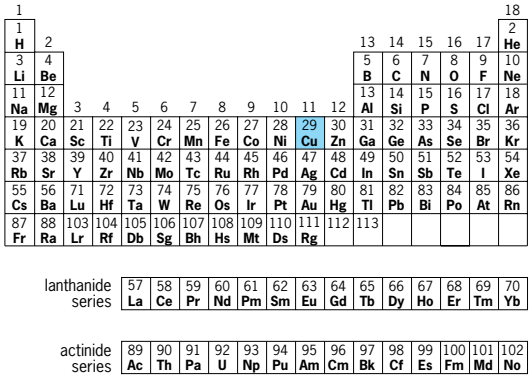

configuration of elemental copper is $[1s^2 2s^2 2p^6 3s^2]$ $3d^{10}4s^1$ or [argon]$3d^{10}4s^1$. At first glance, the sole $4s$ electron might suggest chemical similarity to potassium, which has the [argon]$4s^1$ configuration. However, metallic copper, in sharp contrast to metallic potassium, is relatively unreactive. The higher nuclear charge of copper relative to that of potassium is not fully shielded by the 10 additional d electrons, with the result that the copper $4s$ electron has a higher ionization potential than that of potassium (745.5 versus 418.9 kilojoules/mole, respectively). Moreover, the second and third ionization potentials of copper (1958.1 and 3554 J/mole, respectively) are considerably lower than those of potassium, and account for the higher valence-state accessibility associated with transition-metal chemistry as opposed to alkali-metal chemistry. *See* PERIODIC TABLE; TRANSITION ELEMENTS; VALENCE.

Copper is a comparatively heavy metal. The density of the pure solid is 8.96 g/cm^3 (5.18 oz/in.^{3}) at 20°C (68°F). The density of commercial copper varies with method of manufacture, averaging 8.90–8.92 g/cm^3 (5.14–5.16 oz/in.^{3}) in cast refinery shapes, 8.93 g/cm^3 (5.16 oz/in.^{3}) for annealed tough-pitch copper, and 8.94 g/cm^3 (517 oz/in.^{3}) for oxygen-free copper. The density of liquid copper is 8.22 g/cm^3 (4.75 oz/in.^{3}) near the freezing point.

The melting point of copper is 1083.0 ∓ 0.1°C (1981.4 ∓ 0.2°F). Its normal boiling point is 2595°C (4703°F).

The coefficient of linear expansion of copper is 1.65×10^{-5}/°C at 20°C.

The specific heat of the solid is 0.092 cal/g at 20°C (68°F). The specific heat of liquid copper is 0.112 cal/ g, and of copper in the vapor state about 0.08 cal/g.

The electrical resistivity of copper in the usual volumetric unit, that of a cube measuring 1 cm in each direction, is 1.6730×10^{-6} ohm · cm at 20°C (68°F). Only silver has a greater volumetric conductivity than copper. On a relative basis in which silver is rated 100, copper is 94, aluminum 57, and iron 16.

The mass resistivity of pure copper for a length of 1 m weighing 1 g at 20°C (68°F) is 0.14983 ohm. The conductivity of copper on the mass basis is surpassed

by several light metals, notably aluminum. The relative values are 100 for aluminum, 50 for copper, and 44 for silver.

By far the largest use of copper is in the electrical industry, and therefore high electrical conductivity is its most important single property, although for industrial use this property must be accompanied by suitable characteristics in other respects. *See* CONDUCTOR (ELECTRICITY); ELECTRICAL CONDUCTIVITY OF METALS.

Copper-containing proteins provide diverse biochemical functions, including copper uptake and transport (ceruloplasmin), copper storage (metallothionen), protective roles (superoxide dismutase), catalysis of substrate oxygenation (dopamine β-monooxygenase), biosynthesis of connective tissue (lysyl oxidase), terminal oxidases for oxygen metabolism (cytochrome *c* oxidase), oxygen transport (hemocyanin), and electron transfer in photosynthetic pathways (plastocyanin). *See* BIORGANIC CHEMISTRY; ENZYME. Harvey Schugar

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; J. R. Davis (ed.), *Metals Handbook: Desk Edition*, 2d ed., ASM International, 1998; K. D. Karlin and Z. Tyeklar, *The Bioinorganic Chemistry of Copper*, Springer, 1993; D. R. Likde, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; J. S. Valentine and E. B. Gralla, *Advances in Protein Chemistry: Copper-Containing Molecules*, Academic Press, 2002.

Copper alloys

Solid solutions of one or more metals in copper. Many metals, although not all, alloy with copper to form solid solutions. Some insoluble metals and nonmetals are intentionally added to copper alloy to enhance certain characteristics. Copper alloys form a group of materials of major commercial importance because they are characterized by such useful mechanical properties as high ductility and formability and excellent corrosion resistance. Copper alloys are easily joined by soldering and brazing. Like gold alloys, copper alloys have decorative red, pink, yellow, and white colors. Copper has the second highest electrical and thermal conductivity of any metal. All these factors make copper alloys suitable for a wide variety of products.

Wrought alloys. Copper and zinc melted together in various proportions produce one of the most useful groups of copper alloys, known as the brasses. Six different phases are formed in the complete range of possible compositions. The relationship between composition and phases alpha, beta, gamma, delta, epsilon, and eta are graphically shown in the well-established constitution diagram for the copper-zinc system. Brasses containing 5–40% zinc constitute the largest volume of copper alloys. One important alloy, cartridge brass (70% copper, 30% zinc), has innumerable uses, including cartridge cases, automotive radiator cores and tanks, lighting fixtures, eyelets, rivets, screws, springs, and plumbing products. Tensile strength ranges from 45,000 lb/in.^{2} (310 megapascals) as annealed to 130,000 lb/in.^{2} (900 MPa) for spring temper wire. *See* BRASS; PHASE EQUILIBRIUM.

Lead is added to both copper and the brasses, forming an insoluble phase which improves machinability of the material. Free cutting brass (61% copper, 3% lead, 36% zinc) is the most important alloy in the group. In rod form it has a strength of 50,000–70,000 lb/in.^{2} (340–480 MPa), depending upon temper and size. It is machined into parts on high-speed (10,000 rpm) automatic screw machines for a multiplicity of uses.

Increased strength and corrosion resistance are obtained by adding up to 2% tin or aluminum to various brasses. Admiralty brass (70% copper, 2% aluminum, 0.023% arsenic, 21.97% zinc) are two useful condenser-tube alloys. The presence of phosphorus, antimony, or arsenic effectively inhibits these alloys from dezincification corrosion.

Alloys of copper, nickel, and zinc are called nickel silvers. Typical alloys contain 65% copper, 10–18% nickel, and the remainder zinc. Nickel is added to the copper-zinc alloys primarily because of its influence upon the color of the resulting alloys; color ranges from yellowish-white, to white with a yellowish tinge, to white. Because of their tarnish resistance, these alloys are used for table flatware, zippers, camera parts, costume jewelry, nameplates, and some electrical switch gear.

Copper forms a continuous series of alloys with nickel in all concentrations. The constitution diagram is a simple all alpha-phase system.

Nickel slightly hardens copper, increasing its strength without reducing its ductility. Copper with 10% nickel makes an alloy with a pink cast. More nickel makes the alloy appear white.

Three copper-base alloys containing 10%, 20%, and 30% nickel, with small amounts of manganese and iron added to enhance casting qualities and corrosion resistance, are important commercially. These alloys are known as cupronickels and are well suited for application in industrial and marine installations as condenser and heat-exchanger tubing because of their high corrosion resistance and particular resistance to impingement attack. Heat-exchanger tubes in desalinization plants use the cupronickel 10% alloy.

Copper-tin alloys (3–10% tin), deoxidized with phosphorus, form an important group known as phosphorus bronzes. Tin increases strength, hardness, and corrosion resistance, but at the expense of some workability. These alloys are widely used for springs and screens in papermaking machines.

Silicon (1.5–3.0%), plus smaller amounts of other elements, such as tin, iron, or zinc, increases the strength of copper, making alloys useful for hardware, screws, bolts, and welding rods.

Sulfur (0.35%) and tellurium (0.50%) form insoluble compounds when alloyed with copper, resulting

in increased ease of machining.

Precipitation-hardenable alloys. Alloys of copper that can be precipitation-hardened have the common characteristic of a decreasing solid solubility of some phase or constituent with decreasing temperature. Precipitation is a decomposition of a solid solution leading to a new phase of different composition to be found in the matrix. In such alloy systems, cooling at the appropriate rapid rate (quenching) from an equilibrium temperature well within the all-alpha field will preserve the alloy as a single solid solution possessing relatively low hardness, strength, and electrical conductivity.

A second heat treatment (aging) at a lower temperature will cause precipitation of the unstable phase. The process is usually accompanied by an increase in hardness, strength, and electrical conductivity. Some 19 elements form copper-base binary alloys that can be age- or precipitation-hardened. *See* HEAT TREATMENT (METALLURGY).

Two commercially important precipitation-hardenable alloys are beryllium- and chromium-copper. Beryllium-copper (2.0–2.5% beryllium plus cobalt or nickel) can have a strength of 200,000 lb/in.^{2} (1400 MPa) and an electrical conductivity of less than 50% of IACS (International Annealed Copper Standard). Cobalt adds high-temperature stability, and nickel acts as a grain refiner. These alloys find use as springs, diaphragms, and bearing plates and in other applications requiring high strength and resistance to shock and fatigue.

Copper chromium (1% chromium) can have a strength of 80,000 lb/in.^{2} (550 MPa) and an electrical conductivity of 80%. Copper-chromium alloys are used to make resistance-welding electrodes, structural members for electrical switch gear, current-carrying members, and springs.

Copper-nickel, with silicon or phosphorus added, forms another series of precipitation-hardenable alloys. Typical composition is 2% nickel, 0.6% silicon. Strength of 120,000 lb/in.^{2} (830 MPa) can be obtained with high ductility and electrical conductivity of 32% IACS.

Zirconium-copper is included in this group because it responds to heat treatment, although its strength is primarily developed through application of cold deformation or work. Heat treatment restores high electrical conductivity and ductility and increases surface hardness. Tensile strength of 70,000 lb/in.^{2} (500 MPa) coupled with an electrical conductivity of 88% can be developed. Uses are resistance welding wheels and tips, stud bases for transistors and rectifiers, commutators, and electrical switch gear.

Fabrication. Copper alloys have a long history because of their ease of fabrication into utensils and products.

Melting. Alloys are initially prepared by melting the various elements involved in low-frequency induction furnaces. Usually the melt is protected from oxidation by covering it with charcoal. Copper-rich alloys are also deoxidized by adding elements such as phosphorus, lithium, or boron. The melt is usually poured at 100–200°F (38–93°C) above its melting point into oil- and graphite-dressed, water-cooled copper molds of suitable shape. The dressing protects the mold surfaces, produces a reducing gas that aids in the suppression of oxidation, and brings dirt and dross to the molten surface, preventing their inclusion in the casting.

Continuous casting of copper and its alloys is common practice and is accomplished by pouring the melt into the top of an open-end mold as the solidified casting is withdrawn from the bottom. Various mold designs are used, but all use water to cool the mold proper in addition to cooling sprays and tanks of water below the mold. Casting can be performed in either a vertical or a horizontal plane.

Mechanical properties of castings can be affected by such defects as large grain size, directional grain growth, and segregation. Either hot-working of the casting by rolling, forging, piercing, or extrusion, or cold-working by rolling or drawing with necessary intermediate annealing corrects these defects and markedly improves properties.

Processing. Processing varies with alloy, form, and size. Flat products are cast in heavy cakes for initial hot-rolling or in bars for cold-rolling. Copper and lead-free brasses are readily hot-rolled in a single heating from a 9-in. (23-cm) thick cake to a $^1/_2$ in. (1.25-cm) thick slab. Further rolling is done cold with intermediate anneals. Leaded brasses, tin bronzes, and other hot short alloys (alloys that are brittle when heated above red heat) must be cold-rolled from the casting to final size, using intermediate anneals that heat the particular alloy above its recrystallization temperature.

Round copper rods and wire are started by hot-rolling castings (wire bar) in grooved rolls, but most copper alloys are extruded. Both are subsequently reduced in size by cold-rolling, drawing, or both. In extrusion, a billet placed in a press is forced through a lubricated die by a ram.

Tin and tin-lead bronzes which are hot short are continuously cast in rod form and are cold-rolled or drawn to finish sizes using intermediate anneals as necessary.

Copper and high-copper-content lead-free brass tubes are started by hot piercing wherein the solid billet is driven with a helical motion over a rotating mandrel to form a seamless tube. Tubes of most other alloys are formed by extrusion because the alloys do not stand piercing. In extrusion of tube, a mandrel is pushed through a billet forming the inside of the tube, and a ram pushes the billet over the mandrel and through a die to form a seamless tube.

Cast alloys. Copper alloy castings of irregular and complex external and internal shapes can be produced by various casting methods, making possible the use of shapes for superior corrosion resistance, electrical conductivity, good bearing quality, and other attractive properties. High-copper alloys with varying amounts of tin, lead, and zinc account for a large percentage of all copper alloys used in the

cast form. Tensile strength ranges from 36,000 to 48,000 lb/in.^{2} (250–330 MPa) depending upon composition and size. Leaded tin-bronzes with 6–10% tin, about 1% lead, and 4% zinc are used for high-grade pressure castings, valve bodies, gears, and ornamental work. Bronzes high in lead and tin (7–25% lead and 5–10% tin) are mostly used for bearings. High tin content is preferred for heavy pressures or shock loading, but lower tin and higher lead for lighter loads, higher speeds, or where lubrication is less certain. A leaded red brass containing 85% copper and 5% each of tin, lead, and zinc is a popular alloy for general use.

High-strength brasses containing 57–63% copper, small percentages of aluminum and iron, and the balance zinc have tensile strengths from 70,000 to 120,000 lb/in.^{2} (490–830 MPa), high hardness, good ductility, and resistance to corrosion. They are used for valve stems and machinery parts requiring high strength and toughness.

Copper alloys containing less than 2% alloying elements are used when relatively high electrical conductivity is needed. Strength of these alloys is usually notably less than that of other cast alloys.

Aluminum bronzes containing 5–10.5% aluminum, small amounts of iron, and the balance copper have high strengths even at elevated temperature, high ductility, and excellent corrosion resistance. The higher-aluminum-content castings can be heat-treated, increasing their strength and hardness. These alloys are used for acid-resisting pump parts, pickling baskets, valve seats and guides, and marine propellers.

Additives impart special characteristics. Manganese is added as an alloying element for high-strength brasses where it forms intermetallic compounds with other elements, such as iron and aluminum. Nickel additions refine cast structures and add toughness, strength, and corrosion resistance. Silicon added to copper forms copper-silicon alloys of high strength and high corrosion resistance in acid media. Beryllium or chromium added to copper forms a series of age- or precipitation-hardenable alloys. Copper-beryllium copper alloys are among the strongest of the copper-base cast materials. *See* ALLOY; COPPER; METAL CASTING; METAL FORMING.

Ralph E. Ricksecker

Bibliography. American Society for Testing and Materials, *Annual Book of ASTM Standards*, vol. 02.01: *Copper and Its Alloys*, 1993; J. H. Mendenhall, *Understanding Copper Alloys*, 1977, reprint 1986; E. G. West, *Copper and Its Alloys*, 1982.

Copper loss

The power loss caused by the flow of current through copper conductors. When an electric current flows through a copper conductor (or any conductor), some energy is converted to heat. The heat, in turn, causes the operating temperature of the device to rise. This happens in transformers, generators, motors, relays, and transmission lines, and is a principal limitation on the conditions of operation of these devices. Excessive temperature rises lead to equipment failure. *See* CONDUCTOR (ELECTRICITY).

If the resistance R (in ohms) and the current I (in amperes) are both known, the copper loss in watts may be found from the equation $P = I^2R$. An alternative form is to know the voltage across the resistor V (in volts) and to determine the power loss by the equation $P = V^2/R$. V and I are effective values of the voltage and current, respectively. The resistance R varies with both frequency and temperature. *See* ELECTRICAL RESISTANCE.

In electric machines, there are typically constant losses, as well as those that vary with the square of the load, such as the copper loss. Maximum efficiency is obtained when the load is adjusted so as to make the constant and varying losses equal.

Since the discovery of superconductivity in 1911, much effort has been directed toward the development of devices based on superconducting principles, with superconductors replacing the copper coils in electrical devices and systems in order to reduce or eliminate copper loss. Since superconductors have traditionally operated at temperatures close to absolute zero (0 K or −459.67°F), such devices are practical only if the power required to operate the refrigeration equipment is less than the copper loss that is supplanted. In 1986, a ceramic superconductor, lanthanum barium copper oxide, was discovered whose superconducting temperature was 30 K (−406°F), and by 1988 another ceramic, thallium barium calcium copper oxide, was shown to have a superconducting temperature of 125 K (−235°F). The rapid increase of superconducting temperatures has given new hope to the idea of reducing copper loss. However, the process of going from a scientific discovery to technological realization requires solutions to many technical problems as well as complex problems involving economic, reliability, safety, and related considerations. *See* SUPERCONDUCTING DEVICES; WIRING.

Edwin C. Jones, Jr.

Bibliography. Electric Power Research Institute, *Cost Components of High-Capacity Transmission Options*, Rep. EL-1065, May 1979; K. Fitzgerald, Superconductivity: Fact versus fancy, *IEEE Spectrum*, 25(5):30–41, May 1988; W. Hively, Closer to room temperature, *Amer. Sci.*, 76(6):245–246, May–June 1988; J. L. Kirtley, Jr., and M. Furuyama, A design concept for large superconducting alternators, *IEEE Trans.*, PAS-94(4):1264–1269, July/August 1975.

Copper metallurgy

The economic production of pure copper metal, suitable for fabrication and use, from copper ores containing as little as 0.5% Cu. Over 90% of the consumption of primary copper in the Western world is produced from ores containing sulfide minerals

(chalcopyrite, $CuFeS_2$; chalcocite, Cu_2S; and bornite, Cu_5FeS_4) that can be economically treated only by pyrometallurgical processes. *See* PYROMETALLURGY, NONFERROUS.

Production. The main processes used in the production of copper from sulfide ores are shown in **Fig. 1**. The mined ore (0.5–2.0% Cu) is finely ground, and then concentrated by flotation to form copper concentrates containing 20–30% Cu. The concentrates are then smelted at high temperatures (about 2280°F or 1250°C) to form a molten mixture of copper and iron sulfides called matte. The molten matte

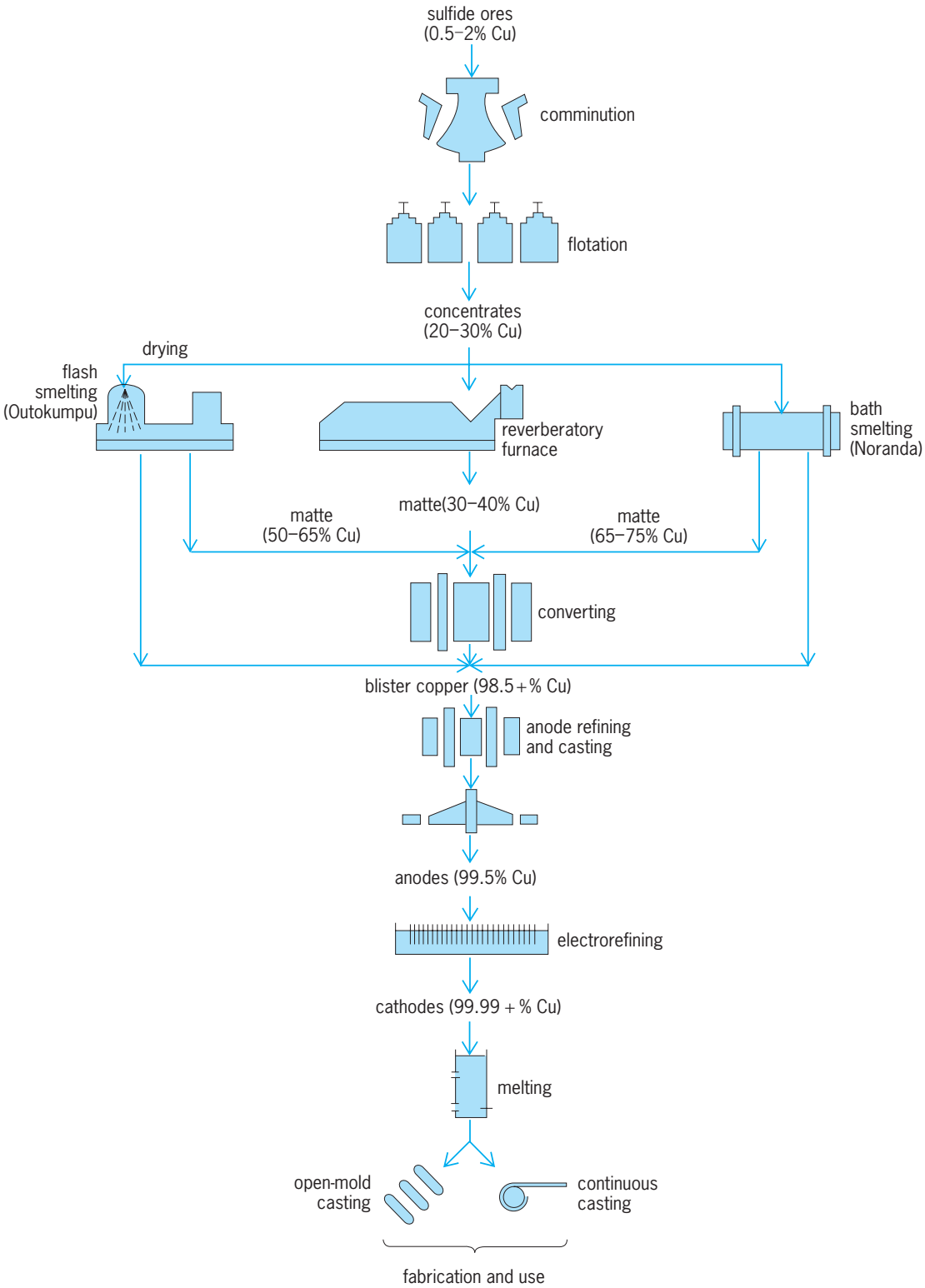

Fig. 1. Flow sheet of copper production from sulfide minerals.

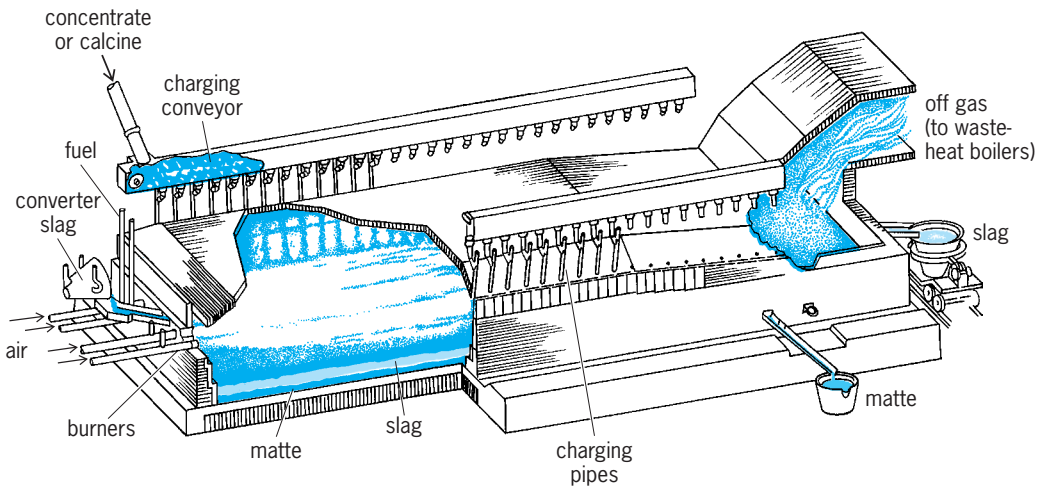

Fig. 2. Cutaway view of a reverberatory furnace for the production of copper matte from sulfide concentrates or roasted calcines. (*After A. K. Biswas and W. D. Davenport, Extractive Metallurgy of Copper, Pergamon Press, 1980*)

is converted to blister copper (about 99% Cu) by oxidizing the remaining iron and sulfur. After removing the residual sulfur and oxygen in an anode furnace, copper anodes are cast and then refined electrolytically to produce high-purity cathode copper (99.99% Cu), which is suitable for most uses. *See* ORE DRESSING.

Smelting processes. The reverberatory furnace (**Fig. 2**) is the oldest and most widely used smelting process. It consists of a refractory-lined chamber, typically 100 ft (30 m) long by 30 ft (10 m) wide, into which copper concentrates and silica flux are charged. Fuel-fired burners melt the charge, driving off the labile sulfur by reaction (1). Little iron sul-

$$2CuFeS_2 \rightarrow Cu_2S \cdot 2FeS(l) + {}^1/_2S_2(g) \qquad (1)$$

fide is oxidized, so that fuel requirements are high, about 5.4×10^6 Btu per ton (6.3×10^6 kilojoules per metric ton) of Cu concentrate. Two molten layers are formed in the furnace: an upper layer slag of iron silicate with little copper ($\leq$0.5% Cu); and a lower layer of matte (30–40% Cu), containing most of the original copper value. The slag and matte are drained separately from the furnace through tapholes into ladles. The slag is discarded, while the matte is transferred to the converting step. The sulfur evolved during smelting leaves the furnace in a 1–2% SO_2 gas stream, which is too dilute for economic treatment.

The reverberatory furnace cannot meet today's requirements for low-energy consumption and stringent environmental control. Therefore more efficient smelting processes have been developed since the 1960s. These processes use much less energy (typically $0.7–1.8 \times 10^6$ Btu per ton or $0.8–2.1 \times 10^6$ kJ per metric ton of Cu concentrate) and produce a strong SO_2 gas stream (10% SO_2) to reduce treatment costs. These processes use oxygen enrichment, and oxidize more iron sulfide to generate more heat and produce mattes with higher copper levels (50–75% Cu). The smelting reaction for these processes can be represented by reaction (2), where

$$2CuFeS_2 + 2\alpha \cdot SiO_2 + (1 + 3\alpha)O_2 \rightarrow$$
$$Cu_2S \cdot 2(1 - \alpha)FeS + 2\alpha \cdot FeO \cdot SiO_2 + (1 + 2\alpha)SO_2 \qquad (2)$$

α is the fraction of FeS reacted (typically, in range $\alpha = 0.5$ to 0.9).

Modern smelting processes fall into two categories: flash smelting (Outokumpu flash smelting furnace) and bath smelting (Noranda process reactor). In flash smelting, dry concentrate is dispersed in an oxidizing gas stream, and the smelting reactions occur very rapidly as the particles fall down a reaction shaft. The molten matte and slag are collected in a hearth, and the SO_2-containing gases exit through an uptake shaft. In bath smelting, moist concentrate is smelted continuously in a molten bath of matte and slag, which is vigorously stirred by the injection of air or oxygen-enriched air. In the Noranda process reactor, the air is injected through tuyeres into a vessel similar to an elongated converter.

Matte converting. The molten matte is converted to blister in Peirce-Smith converters. The converter is a refractory-lined, cylindrical vessel, typically 30 ft (10 m) in length and 13 ft (4 m) in diameter (**Fig. 3**). The converter can be rotated about its axis, and is fitted on one side with a row of about 50 tuyeres through which air is injected. The top of the converter has a large mouth for charging molten matte and removing slag and product copper. Converting is a batch operation; initially several ladles of matte are charged, the air turned on, and the converter rotated until the tuyeres are submerged. The air, bubbling violently through the bath, gradually oxidizes the matte in two stages. In the first or slag-forming stage, iron sulfide is oxidized and fluxed with silica to form a fluid slag by reaction (3). The converter

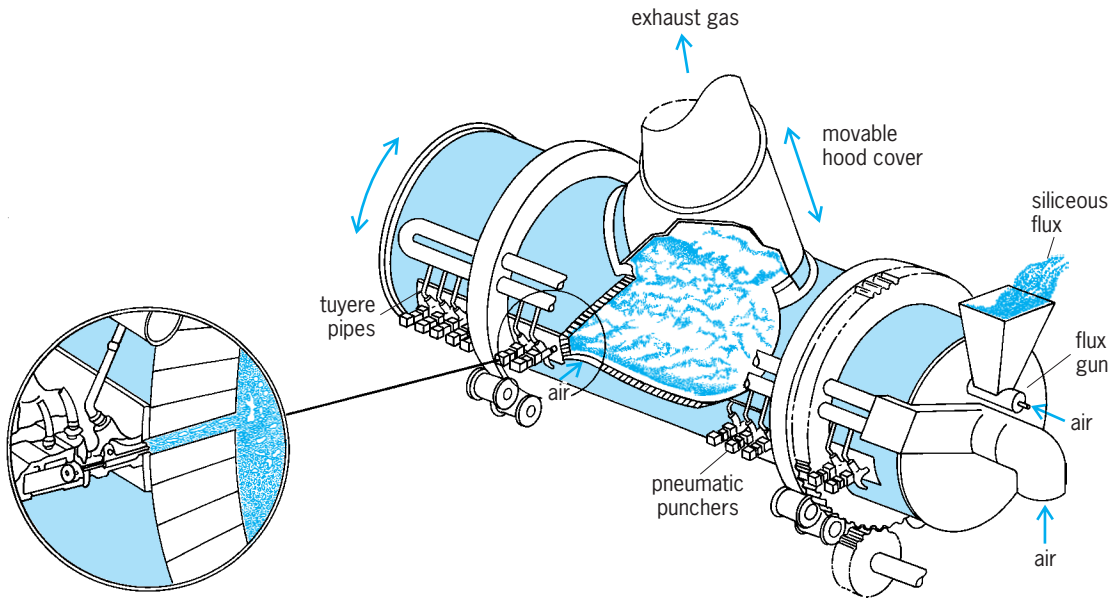

Fig. 3. Cutaway view of a horizontal side-blown Peirce-Smith converter for producing blister copper from matte. (*After A. K. Biswas and W. G. Davenport, Extractive Metallurgy of Copper, Pergamon Press, 1980*)

slag contains some copper (1–5% Cu), and is recy-

$$2FeS(l) + 2SiO_2(s) + 3O_2(g) \rightarrow 2FeO \cdot SiO_2(l) + 2SO_2(g) \quad (3)$$

cled to the smelting furnace or treated in a separate process. When all the iron has been removed, the remaining copper sulfide is further oxidized to blister copper by reaction (4). The converting process

$$Cu_2S(l) + O_2(g) \rightarrow 2Cu(l) + SO_2(g) \quad (4)$$

is sufficiently exothermic that no additional fuel is required. The blister copper from the converter is transferred by ladle to the anode furnace, where the residual sulfur and oxygen levels in the copper are reduced further. The copper is then cast into anodes for electrorefining.

Sulfur fixation. Smelting and converting a typical copper concentrate generates over 0.50 ton SO_2 per ton concentrate (0.50 metric ton SO_2 per metric ton concentrate), and the resulting SO_2 emissions must be controlled to meet local environmental standards. This is generally achieved by converting the SO_2 to sulfuric acid in a contact acid plant, as long as the SO_2 concentration exceeds 4% and a viable market for acid exists. If local conditions are favorable, it is also possible to make liquid SO_2 or elemental sulfur from strong SO_2 gases. *See* SULFUR.

Electrorefining. The objective of electrorefining is to remove the remaining impurities in the anode copper (principally As, Bi, Ni, Pb, Sb, and Se) and produce a pure cathode copper (99.99+% Cu). Also, many copper ores contain appreciable amounts of precious metals (Ag, Au, Pt, and so on), which are concentrated into the anode copper during smelting and are recovered as valuable by-products in electrorefining. The impure anodes are suspended alternately with pure copper cathodes in tanks through which an electrolyte of copper sulfate and free sulfuric acid is continuously circulated. When direct current is applied, the copper in the anodes is electrochemically dissolved and then plated as pure copper on the cathodes. Some of the anode impurities, such as arsenic and nickel, are less noble than copper and dissolve in the electrolyte, but they do not plate out at the cathode as long as their concentrations are controlled. The other impurities, such as silver, lead, and selenium, are virtually insoluble in the electrolyte and fall as slimes to the bottom of the tank. These slimes are recovered and processed for eventual recovery of selenium and the precious-metal values. *See* ELECTROCHEMICAL PROCESS; ELECTROMETALLURGY.

Hydrometallurgical processes. Oxidized copper ores are more effectively treated by hydrometallurgical processes. The ore is crushed, ground if necessary, and leached with dilute sulfuric acid, either by percolation through heaps of ore or by agitation in tanks. Copper is recovered from the resulting solution by either cementation or solvent extraction–electrowinning. In cementation, copper is precipitated by contact with scrap iron to form an impure cement copper, which is smelted, then refined. Solvent extraction–electrowinning has become the preferred process. In solvent extraction special organic reagents are used to selectively extract copper from solution. The resulting copper-containing organic phase is then stripped to give a pure and more concentrated aqueous copper solution for electrowinning. Electrowinning is similar to electrorefining except that an inert anode is used and more energy is required. Although electrowon cathode copper is generally not as pure as electrorefined copper, it is still suitable for many applications. *See* HYDROMETALLURGY; SOLVENT EXTRACTION.

Future trends. Considerable progress has been made in improving the efficiency of copper smelting since the 1960s through new process innovations and the use of oxygen in large tonnages. The trend toward the production of higher-matte grades (up to 75% Cu) is expected to continue in order to minimize energy requirements and SO_2-fixation costs. Pyrometallurgical processes will continue to dominate primary copper production, and it is anticipated that the ultimate goal of direct copper production from chalcopyrite concentrates in a single smelting and converting vessel will be achieved. *See* COPPER.

John G. Peacey

Bibliography. American Institute of Mining and Metallurgical Engineers (AIME), *Copper Smelting: An Update*, 1982; A. K. Biswas and W. G. Davenport, *Extractive Metallurgy of Copper*, 3d ed., 1994; E. G. West, *Copper and Its Alloys*, 1982.

Copulatory organ

An organ utilized by the male animal for insemination, that is, to deposit spermatozoa directly into the female reproductive tract. Various types of copulatory organs are found among the vertebrates, whereas cloacal apposition occurs in most other vertebrates which lack these structures. Amplexus, or false copulation, is peculiar to amphibians.

In fish, internal fertilization is restricted to certain groups. The pelvic fin of male elasmobranchs and holocephalians is modified for the transmission of sperm and is known as the clasper or clasping organ. A pair of anterior claspers (**Fig. 1***a*) and a frontal clasper which protrudes from the head occur among holocephalians in addition to the modified pelvic clasper of elasmobranchs. The anal fin of teleosts, in which copulation occurs, may be elongated to form a gonopodium. Copulatory organs are lacking in amphibians, even in those in which internal fertilization occurs. Among caecelians, the cloaca of the male is muscular and protrusible; however, it is questionable that this structure is a true copulatory organ. Hemipenes which can be everted during copulation are common to both snakes and lizards; however, these structures lack erectile tissue. Turtles and crocodiles possess a single penis with associated erectile tissue, the corpora cavernosa, which becomes distended with blood. A few species of birds have a penis; among these are the ostriches and anseriforms. The penis is present in all mammals.

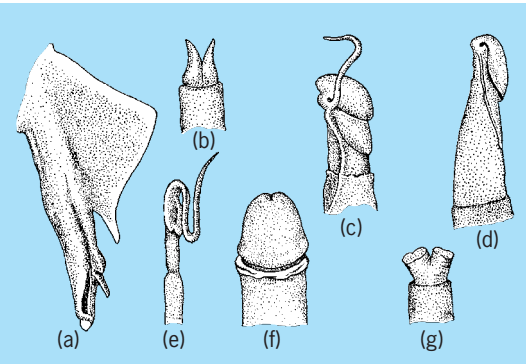

Fig. 1. Vertebrate copulatory organs. (*a*) Clasper of dogfish (*Squalus*). Glans penis of (*b*) opossum, (*c*) ram, (*d*) bull, (*e*) short-tailed shrew, (*f*) man, (*g*) Echidna.

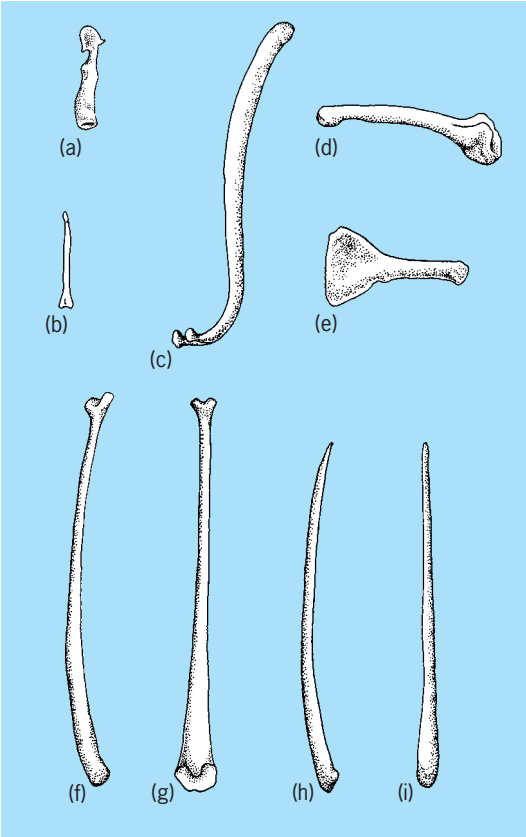

Fig. 2. Bacula of (*a*) fox squirrel, (*b*) cotton mouse, (*c*) otter, (*d*) *Neotoma lepida*, (*e*) *N. micropus*, (*f, g*) *Perognathus hispidus*, showing lateral and ventral views, (*h, i*) *P. baileyi*, showing lateral and ventral views.

This copulatory organ (Fig. 1*b–g*) is variously modified morphologically. Among certain mammals such as rodents, bats, whales, some carnivores, and lower primates, a penis bone, known as the os priapi or baculum (**Fig. 2**), occurs, increasing the rigidity of the penis. *See* FERTILIZATION; PENIS; REPRODUCTIVE SYSTEM.

Charles B. Curtin

Coquina

A calcarenite or clastic limestone whose detrital particles are chiefly fossils, whole or fragmented. The term is most frequently used for an aggregate of large shells more or less cemented by calcite. If the rock consists of fine-sized shell debris, it is called a microcoquina. Encrinite is a microcoquina made up primarily of crinoid fragments (see **illus.**). Some coquinas show little evidence of any transportation by currents; articulated bivalves are preserved in

Coquinoid limestone, encrinite, showing many entire fossils in fine-grained matrix. Note lack of assortment and diversity of types. (*From F. J. Pettijohn, Sedimentary Rocks, 3d ed., Harper and Row, 1949, 1957*)

entirety and the shells are not broken or abraded. *See* LIMESTONE.

Raymond Siever

Coraciiformes

A diverse order of land birds found mainly in the tropics. The interrelationships of several families have been disputed, with no resolution. The closest affinities of the coraciiforms may be with other land birds, such as the Piciformes and Passeriformes. *See* PASSERIFORMES; PICIFORMES.

Classification. The Coraciiformes are divided into four suborders and families: Alcedines, including the families Alcedinidae (kingfishers; 91 species; worldwide, but mainly Old World tropics to Australasia), Todidae (todies; five species; West Indies), and Momotidae (motmots; nine species; New World tropics); Meropes, with the single family Meropidae (bee-eaters; 24 species; Old World to Australia, mainly tropics); Coracii, containing the families Coraciidae (rollers; 12 species; Old World to Australia, mainly tropics), Brachypteraciidae (ground-rollers; five species; Madagascar), Leptosomatidae (cuckoo-rollers; one specie; Madagascar); and Bucerotes, with the families Upupidae (hoopoes; one specie; Old World), Phoeniculidae (wood-hoopoes; eight species; Africa) and Bucerotidae (hornbills; 45 species; Old World tropics to New Guinea).

Several of these families, such as the different groups of rollers, are sometimes merged into a single family, and several groups, such as the kingfishers (**Fig. 1**) and hornbills (**Fig. 2**), are divided into subfamilies or even families.

Fossil record. The fossil record of this group of basically forest birds is rather poor; however, there are a number of specimens known that appear to belong to the Coraciiformes. A better picture of the history and relationships of this order may appear once these fossils are better studied. Most interesting is a fossil tody (a group now restricted to the Greater Antilles) from the Oligocene of Wyoming, and a second fossil tody from the Eo-oligocene of France, and a fossil motmot (a group that is now restricted to the American tropics) from the Oligocene of Switzerland. The identification of these fossils as such is controversial, but, if correct, suggests an earlier, broader distribution of several coraciiform families.

Characteristics. Coraciiform birds are characterized by a syndactyl foot in which the three anterior toes are joined at the base, although the cuckoo-rollers have evolved secondarily a zygodactyl foot

Fig. 1. Malachite kingfisher (*Alcedo cristata*), Africa. (*Photo by H. Vannoy Davis; © 2001 California Academy of Sciences*)

Fig. 2. Southern ground hornbill (*Bucorvus leadbeateri*), Masai Mora Game Reserve, Kenya. (*Photo by Gerald and Buff Corsi; © 2002 California Academy of Sciences*)

(in which the toes are arranged two in front and two behind). Most groups are tropical or warm temperate and are brilliantly colored. Except for bee-eaters and wood-hoopoes, coraciiforms are generally solitary. Most temperate species are migratory; the tropical ones are permanent residents. Most coraciiforms feed on insects and other animal prey, including fish; hornbills are omnivorous and some feed mainly on fruit. All nest, either solitarily or in colonies, in holes burrowed in earthen banks or in tree cavities. Wood-hoopoes and bee-eaters have young adults serving as helpers at the nest. The chicks are born naked and cared for by both parents until they can fly. The pair of central tail feathers of most motmots have racquet tips formed by the falling off of more anterior, weakly attached barbs.

Todies, motmots, and one small subfamily of kingfishers are found in the New World, but all other coraciiforms live in the Old World, suggesting the origin of this order in that part of the world. The kingfishers (Fig. 1) are the largest and most diversified family, with many members catching their prey on dry land. The largest and most famous kingfisher is the kookaburra or laughing jackass (*Dacelo novaeguinaea*) of Australia. Rollers are widespread throughout the Old World tropics, but the ground-rollers and the cuckoo-rollers are restricted to Madagascar. Wood-hoopoes are confined to the tropical forests of Africa, while the related hoopoes are widespread throughout Europe and Asia.

The hornbills (Fig. 2) are the largest of the coraciiforms, feed in forests, and are excellent soarers. A number of species possess a large casque on the base of the upper jaw which may serve in courtship or species recognition. The male helmeted hornbill (*Rhinoplax vigil*) of the Malay Peninsula and Indonesia has a solid casque of a thick rhamphotheca (horny covering of the beak) backed by massive bone, making its skull the heaviest of all known Recent birds. Male hornbills seal up the female with her eggs in the nesting cavity in a large tree, and feed her and the chicks until the young birds are ready to fly. During this period, the female undergoes a complete molt. Some species of hornbills are threatened because of habitat destruction, especially the loss of large trees with suitable nesting cavities. *See* AVES.

Walter J. Bock

Bibliography. C. H. Fry and K. Fry, *Kingfishers, Bee-eaters and Rollers*, Helm, 1992; J. del Hoyo et al. (eds.), *Order Coraciiformes: Handbook of the Birds of the World*, vol. 6, pp. 130–523, Lynx Edicions, 2001; A. Kemp, *The Hornbills*, Oxford University Press, 1995; A. A. Kepler, *A Comparative Study of Todies (Todidae)*, vol. 16, Publications of the Nuttall Ornithological Club, 1977.

Corallimorpharia

An order of cnidarian subclass Hexacorallia commonly known as mushroom anemones. These polyps are morphologically intermediate between members of the hexacorallian orders Scleractinia (corals), with which they share features of internal anatomy, weak musculature, and certain types of nematocysts (intracellular stinging capsules used in food capture and defense), and Actiniaria (anemones), which they resemble in lacking a calcareous skeleton. All three animals have paired, coupled mesenteries that are added around the entire periphery of the polyp. Like Actiniaria and Scleractinia, Corallimorpharia contains some members that are solitary and some that are clonal (reproduce asexually to produce genetically identical individuals that live physically separate from one another); however, by contrast with Scleractinia, Corallimorpharia has no colonial members (clonemates that remain attached to one another). *See* CNIDARIA.

Morphology. There are two distinct morphological variants of corallimorpharians. Members of one group live primarily at high altitudes in shallow water and in deep water (to about 5-km or 3-mi depth) at all latitudes. A typical polyp has a well-defined cylindrical column and an oral disc not much broader than the column. The unbranched tentacles are arranged in radial rows on the disc, and commonly each has a spherical tip (that is, it is capitate). Examples are *Corynactis* (**Fig. 1**), which has a typical oral disc diameter of about 10 mm, and *Corallimorphus*, which may attain a diameter of 100–150 mm (4–6 in.).

In members of the other, primarily tropical, group, the oral disc is much broader than the small column, and the tentacles, which may be arranged in rows or in cycles, can be branched or be mere bumps. Members of this group harbor symbiotic dinoflagellates (zooxanthellae), so they occur only in shallow water. Examples are *Actinotryx* (**Fig. 2**), which may have an oral disc diameter of 20–25 mm (1 in.), and *Amplexidiscus*, the largest corallimorpharian, which can reach an oral disc diameter of 250 mm (10 in.).

Phylogeny. Based on preliminary phylogenetic analyses using both morphological and molecular data, it is likely that the two morphological variants

Fig. 1. *Corynactis californica.* (*Photograph by George Miller*)

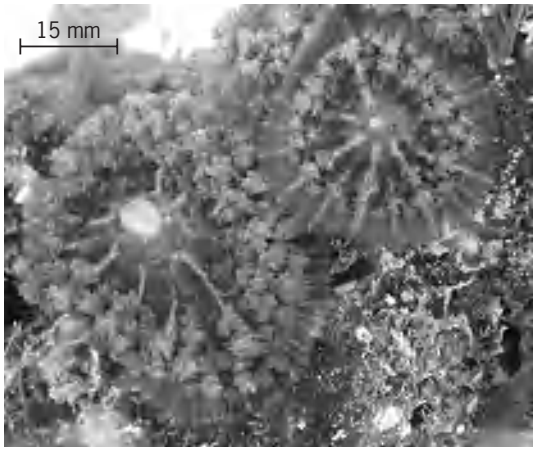

Fig. 2. ***Actinotryx sanctithomae.*** **(*Photograph by Ha-Rim Cha*)**

represent separate evolutionary lineages (that is, are polyphyletic). It seems likely that Corallimorpharia represents at least two taxa, derived independently from different types of scleractinians by loss of the skeleton. Indeed, it may be that corallimorpharians should be considered naked (askeletal) members of the order Scleractinia. Corallimorpharians have no fossil record because the animals lack skeletons. *See* ACTINIARIA; HEXACORALLIA; SCLERACTINIA.

Ecology. Corallimorpharians live in habitats typical of actiniarians and scleractinians. Some species proliferate on stressed coral reefs, attaching to skeletons of dying corals; it is uncertain whether the corallimorpharians merely take advantage of the dead coral skeleton as a substrate on which to grow, but they may kill the corals, for corallimorpharians can aggress toward other cnidarians by extruding their mesenterial filaments (the longitudinal band along the free edge of each mesentery). Like all cnidarians, corallimorpharians eat only animal prey, being opportunistic predators of small animals, such as crustaceans and fishes. The flat, shallow-water species have zooxanthellae, from which they presumably derive photosynthate as well. *See* REEF.

Daphne G. Fautin

Bibliography. D. Doumenc, *Traité de Zoologie: Cnidaires, Anthozoaires*, Masson, Paris, 1987; D. F. Dunn, Cnidaria, pp. 669–706 in S. P. Parker (editor-in-chief), *Synopsis and Classification of Living Organisms*, vol. 1, McGraw-Hill, New York, 1982; H. Erhardt and D. Knop, Order Zoantharia, pp. 269–278 in *Corals: Indo-Pacific Field Guide*, Ikan, Frankfurt, 2005; L. H. Hyman, Class Anthozoa: Subclass Zoantharia, pp. 566–632 in *The Invertebrates: Protozoa through Ctenophora*, vol. 1, McGraw-Hill, New York, 1990.

Corallinales

An order of red algae (Rhodophyceae), commonly called coralline algae. Only one family is recognized, the Corallinaceae, which formerly was assigned to the order Cryptonemiales. These algae are distinguished by the impregnation of cell walls with calcite, a form of calcium carbonate, which causes the thallus to be stony or brittle (see **illus.**). The few other red algae that are calcified are much softer than corallines and are impregnated with aragonite, another form of calcium carbonate. Many early biologists (including C. Linnaeus) thought that these plants were related to corals. *See* RHODOPHYCEAE.

Coralline algae, comprising about 40 genera and 500 species, are widespread, abundant, and ecologically important. They are divisible into two groups on the basis of the presence or absence of uncalcified, moderately flexible joints (genicula) between calcified segments (intergenicula). Those with genicula are called articulated coralline algae, while those totally lacking genicula are called nonarticulated or, with reference to their usual habit, crustose coralline algae. The erect portions (fronds) of articulated corallines usually arise from a crustose base. They vary in height from a few millimeters to 30 cm (12 in.) and are pinnately, dichotomously, or proliferously branched. Segments are cylindrical or flattened.

Structure. The most simple nonarticulated coralline algae are individual crusts of varying extent and thickness (to 20 cm or 8 in. thick). These often become confluent and cover large expanses of substrate. Many crusts bear rounded or pointed, branched or unbranched protuberances that can break off and continue to grow as free nodules known collectively as maerl. In some species, only part of the crust is attached to the substrate, the main portion growing free and resembling lichens or bracket fungi. In three genera of nonarticulated corallines, the thallus is ribbonlike, the linear segments being dichotomously branched, dorsiventrally organized, and flexible.

Coralline algae are cytologically similar to other red algae except for calcification of the cell wall. The wall consists of three layers: a thin, uncalcified inner layer; a thick, calcified middle layer; and a thin,

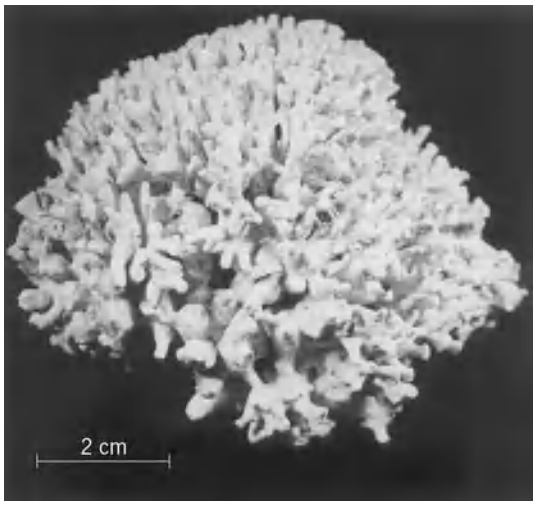

Coralline algae of the Pacific Ocean.

calcified outer layer or middle lamella. The calcite apparently is transported in vesicles from the cell lumen to the outer layers. Growth, which is restricted to uncalcified or decalcified areas or tissues, is usually very slow, especially at sites of low light intensity. Most corallines are perennial.

Reproduction. The life history follows the *Polysiphonia*-type, in which male and female plants alternate with a tetrasporophyte, all three phases being free-living and vegetatively identical or nearly so. Reproductive structures are produced in sunken or elevated chambers (conceptacles) with a roof penetrated by one or many pores. Tetraspores are produced by simultaneous zonate cleavage, a process unique to the order. They are relatively large and in most genera germinate in special patterns of cell division.

Habitat. Coralline algae are exclusively marine, although some species can tolerate a reduction in salinity to 13 parts per thousand. Some species thrive only where light is intense, as at the crest of a coral reef, while others grow only in shaded habitats or in deep water. Most species require constant immersion. Rock is the most common substrate in cold water, with some species preferring particular aspects, such as broad ledges, crevices, cobbles, or pebbles. Other hard substrates include shells, dead coral (the most common substrate in the tropics), and such artifacts as bottles and metal plates. Many corallines are obligate epiphytes, growing on other algae and on seagrasses, and a few overgrow such soft animals as sponges. An inconspicuous layer of corallines often underlies a turf of algae that at first glance appears to be growing directly on rock.

Unlike corals, coralline algae are equally abundant in cold or warm water, occurring from the Arctic to the Antarctic. At high latitudes, crustose corallines pave vast expanses of ocean floor. Along temperate shores, both crustose and articulated corallines are often abundant in tide pools. In areas of rough water, they typically form a broad band in the lowermost intertidal and uppermost subtidal zone. They may dominate deeper sites where herbivores consume fleshy algae. Stands of crustose corallines are made conspicuous by their bright pink or bluish-pink color.

The ecosystem in which coralline algae are most important is the coral reef, where they are primary producers, adding carbon to the ecosystem, adding new material to the reefs, and cementing together other calcareous organisms. They have been engaged in similar activities through the millennia, with modern genera recognizable in limestones at least as old as 150 million years (Jurassic).

Paul C. Silva; Richard L. Moe

Fossils. Coralline algae are important frame builders and sediment producers in tropical reefs throughout Cenozoic time. Although scleractinian corals are volumetrically more important in most living reefs, crustose coralline algae are essential binding and cementing agents. Prominent coralline algae of tropical reefs today are *Neogoniolithon, Porolithon*, and *Lithoporella. Mesophyllum* and *Lithothamnium* are mainly cold-water genera, but they do occur in low-latitude regions in deep-water environments.

Coralline algae have a rather complete fossil record consisting of numerous genera and species. The earliest definite representatives of the Corallinaceae appeared in Jurassic time, although fossil forms with similar morphologies are known from late Paleozoic rocks. Coralline algae are some of the more common and widespread skeletal constituents in Cretaceous and Cenozoic marine rocks. *See* ALGAE; ATOLL; REEF.

John L. Wray

Bibliography. H. W. Johansen, *Coralline Algae: A First Synthesis*, 1981; J. H. Johnson, *Limestone-Building Algae and Algal Limestones*, 1961; J. L. Wray, *Calcareous Algae*, 1977.

Corbino disk

A conducting disk with concentric inner and outer electrical contacts, which can be placed in a magnetic field parallel to its axis. The disk is named after O. M. Corbino, who in 1911 reported magnetoresistance measurements on several metals by using this configuration. In zero applied field the lines of current are simply radial (**illus.** *a*), but in the presence of an axial magnetic field they lengthen by spiraling (illus. *b*). This spiraling occurs because the geometry of the disk is such that the Lorentz force acting on the charge carriers is not counterbalanced by a Hall-effect electric field. The resistance of the disk increases as the field increases, largely as a result of the geometrical magnetoresistance effect associated with the lengthening of the current path.

If B is the magnetic induction, the resistance of the Corbino disk can be expressed by Eq. (1). Here the Hall angle θ is defined by Eq. (2), where ρ_{yx} is

$$R(B) = R(0)(1 + \tan^2\theta)\frac{\rho_{xx}(B)}{\rho_{xx}(0)} \qquad (1)$$

$$\tan\theta = \frac{\rho_{yx}}{\rho_{xx}} = \mu B \qquad (2)$$

the Hall resistivity, ρ_{xx} is the usual resistivity (that is, the resistivity measured on a long wire of the material), and μ is the carrier mobility. For a disk material characterized by high mobility (as is the case for many semiconductors), the factor $1 + \tan^2\theta$, describing the geometrical magnetoresistance effect in Eq. (1), can be very large. The factor $\rho_{xx}(B)/\rho_{xx}(0)$, which measures the intrinsic magnetoresistance, is then relatively much smaller. With conformal mapping methods it is possible to demonstrate that the Corbino disk is equivalent to a very wide or very thin rectangular element of the material in which the Hall voltage is completely shorted and the geometrical magnetoresistance is maximized. Corbino disks have been used only rarely in practical devices

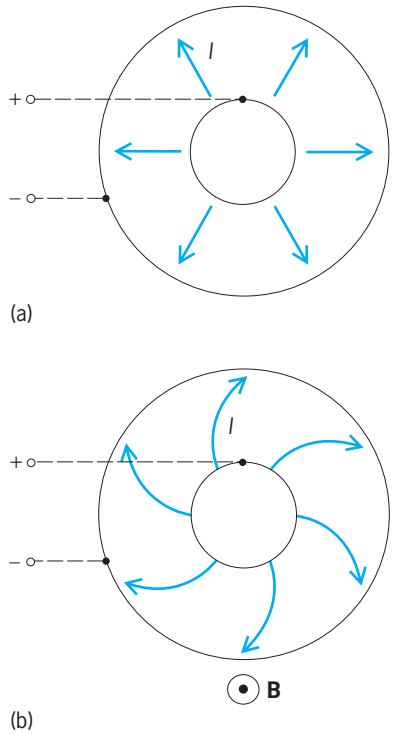

Corbino disk with voltage *V* and current *Γ* between inner and outer electrodes. (*a*) Current paths are radial in the absence of an applied magnetic field. (*b*) With an axial magnetic induction B directed out of the paper, the current paths become spirals, and the increased path length leads to increased resistance.

since they have a low overall resistance and correspondingly large power dissipation. For background information *see* HALL EFFECT; MAGNETORESISTANCE.

J. F. Herbst

Bibliography. J. Heremans, Solid state magnetic field sensors and applications, *J. Phys. D: Appl. Phys.*, 26:1149–1168, 1993; D. A. Kleinman and A. L. Schawlow, Corbino disk, *J. Appl. Phys.*, 31:2176–2187, 1960.

Cordaitales

An extinct order of the class Pinopsida comprising a natural grouping of Paleozoic forest trees or shrubs that first appeared in the Lower Pennsylvanian. They became an important component of the tropical vegetation during the Middle and Upper Pennsylvanian and diminished during the basal Permian. The Cordaitales were divided into three families, Cordaitaceae, Pityaceae, and Poroxylaceae. However, members of the Pityaceae and Poroxylaceae are now known to be either seed ferns or progymnosperms, so only the Cordaitaceae remain. Detailed information about these prominent Paleozoic gymnosperms comes from impression-compression fossils and sandstone casts. However, coal balls, which are mineral nodules in which the plants are preserved with three-dimensional anatomy, have been essential to establishing whole-plant concepts of cordaitean species. *See* COAL BALLS.

Some of the first fossil plant reconstructions were based on the consistent co-occurrence of plant parts at numerous localities. Reconstructions are based on numerous specimens of anatomically preserved, interconnected organs from single localities. This approach has resulted in several species of cordaitean plants constituting two distinct lineages, characterized by a diversity of life cycle and habitat. For example, although cordaites from mesic environments were tall, slender trees, those that grew in or close to peat-forming swamps were scrambling shrubs and small, possibly mangrovelike trees.

Features. *Mesoxylon* and *Cordaixylon* are cordaitean stem genera that have been expanded to whole-plant concepts. There are some features common to both genera. Cordaitean stems typically have a septate pith, although in some species it is solid (**Fig. 1**). Pith casts typically are 1.6–4 in. (4–10 cm) in diameter and provide evidence of the diaphragms separating the pith cavities. The wood is simple and has the basic structure found in modern Araucarian conifers with radial files of tracheids, mostly uniseriate rays, and no resin canals. Cordaites lack growth rings. Main-stem diameters range from 1.75 in. (4.4 cm) [shrubs] to more than 3 ft (1 m) [trees], with trees attaining heights of up to 100 ft (30 m).

Cordaitean leaves are sessile, are helically arranged, and appear to have been coriaceous and heavily cutinized with acuminate-to-rounded tips (**Fig. 2**). Most leaves are linear, up to 40 in. (100 cm) long and 0.75–2 in. (2–5 cm) wide, with numerous parallel veins but no midrib, much like those of the conifer *Agathis* or the angiosperm *Amaryllis*. However, needle leaves with a single vein are produced in the basal regions of branches in some species. There is no evidence of an abscission layer.

Some trees apparently formed highly branched, shallow root systems similar to many living conifers, while the shrublike forms had numerous

Fig. 1. Cross sections of cordaitean stems with simple wood. Left, *Cordaixylon iowensis* with septate pith. Right, *Mesoxylon priapi* with solid pith. (*Photographs by M. L. Trivett*)

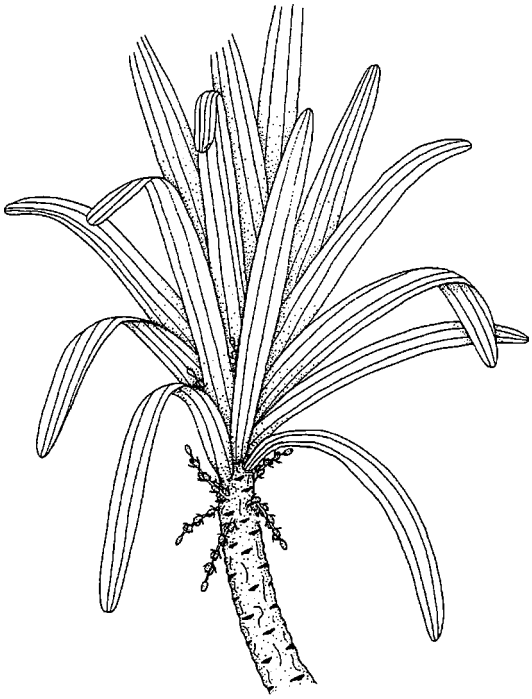

Fig. 2. Shoot of *Cordaixylon dumusum* illustrating strap leaves and nonaxillary *Cordaitanthus*-type cones. (*Courtesy of G. W. Rothwell*)

adventitious roots. Other roots have a combination of aerenchyma, periderm, lenticels, and medullated protosteles reminiscent of the stilt roots of some modern mangrove plants.

Vegetative and fertile branching is almost exclusively axillary. In most species, a single vegetative or fertile branch occurs in the axil of a leaf, although occasionally a vegetative and fertile branch or two fertile branches occur together. In a few species (such as *C. birame*), paired fertile branches are the norm. In others (for example, *C. dumusum*), single fertile branches are borne as nonaxillary, epicormic branches (Fig. 2).

Cordaitean reproductive organs are unisexual compound cones with a bilateral primary axis bearing four ranks of bracts, each subtending a secondary shoot. Secondary axes have numerous helical scale leaves, with some of the apical ones bearing pollen sacs or ovules at their tips.

Features of cauline primary xylem and reproductive structures are important generic characters within the Cordaitales. *Cordaixylon* stems have interconnected cauline strands and leaf traces, while *Mesoxylon* stems lack cauline bundles at the pith margin and have primary xylem present only as leaf traces. *Mesoxylon* cones (*Gothania*) [**Fig. 3**] are larger with more rectangular primary axes than *Cordaixylon* cones (*Cordaitanthus*) [Fig. 2] and produce pollen (*Sullisaccites* or *Felixipollinites*) in a line of pollen sacs. *Cordaixylon* cones produce pollen (*Florinites*) in rings of pollen sacs. *Mesoxylon* ovules (*Mitrospermum*) are winged, and *Cordaixylon* ovules (*Cardiocarpus*) are cordate. *See* FOSSIL SEEDS AND FRUITS.

Growth architecture and ecological interpretations. Integrating depositional information with architectural analyses can provide evidence of plant community structure and ecological partitioning. Delicate organs that show little, if any, transport damage indicate deposition at or near the growth site. Aerial organs of *Mesoxylon priapi* and *C. dumusum* are plentiful. Roots of *C. dumusum* are small but abundant, and *M. priapi* roots are lacking. *Cordaixylon dumusum* is reconstructed as a shrub inhabiting the stressed peat swamp environment, while *M. priapi* trees edged or perhaps grew on hummocks in the swamp. In addition, pollen sacs of *M. priapi* cones matured simultaneously, while those of *C. dumusum* appear to have matured over time, perhaps contributing to its role as a weedy colonizer.

Although most cordaites have leaves uniformly closely spaced, *C. iowensis* leaves are tightly packed in the bud but widely separated on mature stem sections and the pith is disrupted, indicating extensive internodal elongation. Needle leaves develop at the base of some branches. Such features typify rhythmic growth found in many temperate trees, resulting in a somewhat tiered architecture. Although growth rings (evidence of seasonality) are not present, rhythmic growth may have been induced by fluctuating environmental stress. The Cordaitales are an important group in understanding paleocommunity physiognomy, and in determining gymnospermous phylogeny through their relationships to both seed

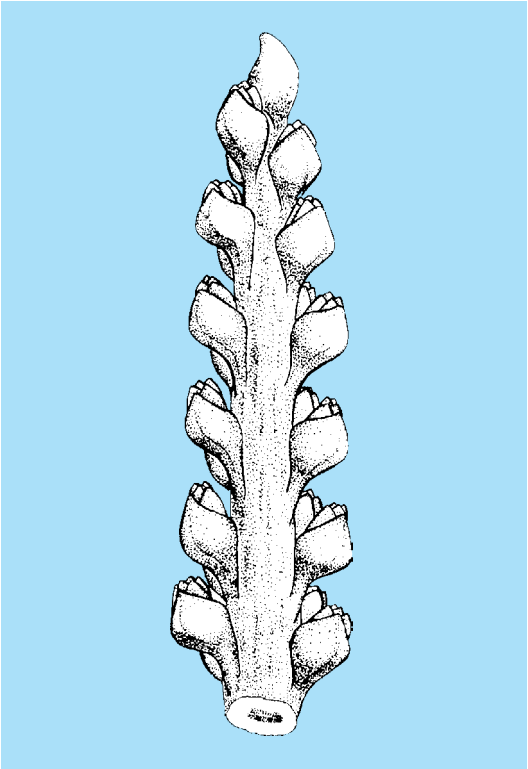

Fig. 3. Compound *Gothania*-type cone of *Mesoxylon priapi* with bracts subtending secondary shoots. (*Courtesy of M. L. Trivett*)

ferns and early conifers. *See* PALEOBOTANY; PINOPHYTA. Mary Louise Trivett

Bibliography. W. N. Stewart and G. W. Rothwell, *Paleobotany and the Evolution of Plants,* 2d ed., 1993; T. N. Taylor and E. L. Taylor, *Biology and Evolution of Fossil Plants,* 1993; M. L. Trivett and G. W. Rothwell, Diversity among Paleozoic Cordaitales: The vascular architecture of *Mesoxylon birame* Baxter, *Bot. Gaz.,* 149:116–125, 1988.

Cordierite

An orthorhombic magnesium aluminosilicate mineral of composition $Mg_2[Al_4Si_5O_{18}]$, $a = 1.710$ nanometers, $b = 0.973$ nm, $c = 0.936$ nm, $Z = 4$, space group *Cccm*. The crystal structure is related to beryl; magnesium is octahedrally coordinated by oxygen atoms, and the tetrahedrally coordinated aluminum and silicon atoms form a pseudohexagonal honeycomb, admitting limited amounts of K^+ and Na^+ ions and water molecules in the open channels. Limited amounts of Fe^{2+} may substitute for Mg^{2+}, and Fe^{3+} for Al^{3+}. The hardness is 7 (Mohs scale), specific gravity 2.6; luster vitreous; cleavage poor; and color greenish-blue, lilac blue, or dark blue, often strongly pleochroic colorless to deep blue (see **illus.**). Transparent pleochroic crystals are used as gem material. The disordered cordierite structure, $Mg_2[(Al,Si)_9O_{18}]$, is called indialite and is hexagonal, space group *P6/mcc*, isotypic with beryl. Osumilite, $KMg_2Al_3 \cdot [(Al,Si)_{12}O_{30}] \cdot H_2O$, is also related but has a structure built of double six-membered rings of tetrahedrons. Cordierite, indialite, and osumilite are difficult to distinguish. Pale colored varieties are often misidentified as quartz, since these minerals have many physical properties in common. *See* BERYL; SILICATE MINERALS.

Cordierite possesses unusually low thermal expansion, and synthetic material has been applied to thermal-shock-resistant materials, such as insulators for spark plugs and low-expansion concrete. Cordierite is an important phase in the system MgO–Al_2O_3–SiO_2. Between 800 and 900°C (1470–1650°F), glasses of cordierite composition crystallize. Below 500°C (930°F) at pressures up to 500 bars (50 megapascals), cordierite in the presence of water breaks down to form chlorite + pyrophyllite. Indeed, it is often found naturally altered to these minerals. Cordierite melts incongruently at 2000 bars (200 MPa) and 1125°C (2057°F) to mullite + spinel + liquid. Natural cordierites break down at temperatures of about 800°C (1470°F) and pressures above 8 kilobars (800 kilopascals) to form enstatite, sillimanite, and quartz.

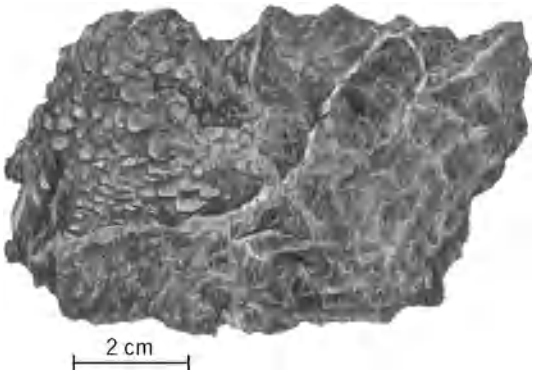

Cordierite, from Tsilaizina, Madagascar. (*Specimen from Department of Geology, Bryn Mawr College*)

Cordierite frequently occurs associated with thermally metamorphosed rocks derived from argillaceous sediments. A common reaction in regional metamorphism resulting in gneisses is garnet + muscovite → cordierite + biotite. It may occur in aluminous schists, gneisses, and granulites; though usually appearing in minor amounts, cordierite occurs at many localities throughout the world. Paul B. Moore

Cordilleran belt

A mountain belt or chain which is an assemblage of individual mountain ranges and associated plateaus and intermontane lowlands. A cordillera is usually of continental extent and linear trend; component elements may trend at angles to its length or be nonlinear.

The term cordillera is most frequently used in reference to the mountainous regions of western South and North America, which lie between the Pacific Ocean and interior lowlands to the east (see **illus.**). The term was first applied to the Andes Mountains of South America in their entirety (Cordillera de los Andes), but individual mountain ranges within the Andean belt are now called cordilleras by some authorities. Farther north, the extensive and geologically diverse mountain terrane of western North America is formally known as the Cordilleran belt or orogen. This belt includes such contrasting elements within the United States as the Sierra Nevada, Central Valley of California, Cascade Range, Basin and Range Province, Colorado Plateau, and Rocky Mountains. *See* MOUNTAIN SYSTEMS.

Cordilleras represent zones of intense deformation of the Earth's crust produced by the convergence and interaction of large, relatively stable areas known as plates. J. F. Dewey and J. M. Bird have analyzed mountain belts in terms of different modes of plate convergence. They contrast cordilleran-type mountain belts, such as the North American Cordillera, with collision-type belts, such as the Himalayas. The former develop during long-term convergence of an oceanic plate toward and beneath a continental plate, whereas the latter are produced by the convergence and collision of one continental plate with another or with an island arc. Characteristics of cordilleran-type mountain belts include their position along a continental margin, their widespread volcanic and plutonic igneous activity, and their tendency to be bordered on both sides by zones of low-angle thrust faulting directed away from the axis of the belt. The deformational, igneous, and metamorphic characteristics of cordilleran belts appear to be largely related to crustal compression, to melting of

The arrangement of the mountains and the continents of the world. (*After P. E. James, An Outline of Geography, Ginn and Co., 1935*)

rocks at depth along the inclined plate boundary, and to attendant thermal effects in the overlying continental plate. *See* OROGENY. Gregory A. Davis

Bibliography. B. C. Burchfiel and G. A. Davis, Nature and controls of Cordilleran orogenesis, western United States: Extensions of an earlier synthesis, *Amer. J. Sci.*, Rodgers Volume, vol. 275A, 1975; J. F. Dewey and J. M. Bird, Mountain belts and the new global tectonics, *J. Geophys. Res.*, vol. 75, no. 14, 1970; C. J. Schmidt and W. J. Perry, Jr. (eds.), *Interaction of the Rocky Mountain Foreland and the Cordilleran Thrust Belt*, 1989.

Core loss

The rate of energy conversion into heat in a magnetic material due to the presence of an alternating or pulsating magnetic field. It may be subdivided into two principal components, hysteresis loss and eddy-current loss. *See* EDDY CURRENT.

Hysteresis loss. The energy consumed in magnetizing and demagnetizing magnetic material is called the hysteresis loss. It is proportional to the frequency and to the area inside the hysteresis loop for the material used. Hysteresis loss can be approximated empirically by using Eq. (1), where K_h is a constant

$$P_h = K_{ith} f B_{\max}^n \tag{1}$$

characteristic of the material, f is the frequency, $B_{\max}$ is the maximum flux density, and n, called the Steinmetz coefficient, is often taken as 1.6 although it may vary from 1.5 to 2.5. *See* MAGNETIZATION.

Most rotating machines are stacked with silicon steel laminations, which have low hysteresis losses. The cores of large units are sometimes built up with cold-reduced, grain-oriented, silicon iron punchings having exceptionally low hysteresis loss, as well as high permeability when magnetized along the direction of rolling.

Eddy-current loss. Induced currents flow within the magnetic material because of variation in the flux. This eddy-current loss may be expressed to close approximation by using Eq. (2), where K_e is

$$P_e = K_e (B_{\max} f \tau)^2 \tag{2}$$

a constant depending on the volume and resistivity of the iron, $B_{\max}$ the maximum flux density, f the frequency, and τ the thickness of the lamination in the core. For 60-cycle rotating machines, core lamination of 0.014–0.018 in. (0.35–0.45 mm) are usually used to reduce eddy-current loss.

Measured core loss. The measured core loss in a rotating machine also includes eddy-current losses in solid structural parts, such as the frame, ventilating duct spacers, pole faces, and damper windings, as well as those due to burrs or other contacts between punchings. A number of precautions are taken to minimize these components, which become appreciable in large machines. Structural parts close to the core, including clamping plates, I-beam spacers, dovetail bars, and shields, may be made of nonmagnetic material. Some of these parts may be shielded from the variable flux by low-resistance plates or by flux traps of laminated, high-permeability, low-loss steel. The core punchings are deburred and are coated with a baked-on insulating varnish. The finished cores are often tested with an ac magnetizing coil to locate and correct hot spots caused by damage during assembly. Pole-face losses are reduced by avoiding excessive slot-width to air-gap ratios, by surface grooving of the poles, and by low-resistance amortisseur windings. *See* ELECTRIC ROTATING MACHINERY.

Leon T. Rosenberg

Coriander

A strong-scented annual herb. Coriander is cultivated in many places throughout the world for both seeds and leaves. The two forms are quite different in taste from one another, and both are used for flavor in a variety of foods. Only one species, *Coriandrum sativum*, is cultivated. Coriander is a member of the carrot family, Apiaceae (Umbelliferae), and is closely related to other spice seed plants such as cumin, caraway, anise, dill, and fennel. A number of distinct cultivars have been developed. Some, with longer maturity times and resulting higher leaf yield, are grown for cilantro, also called Chinese parsley. Other seed-producing types, having a more uniform seed set, and types that resist splitting have been developed in Russia and India. *See* ANISE; APIALES; CARAWAY; CUMIN; DILL; FENNEL.

The growing of cilantro from seed to a harvestable size takes 40–60 days, and full maturity as coriander (dry ripe fruit) is reached in approximately 120 days. Coriander occurs as an erect herbaceous plant with one to many stems that reach about 4 ft (1.2 m) in height. When the plant is young, small deeply segmented leaves are formed, but as the plant ages the leaves become more and more divided until they become feathery in appearance, similar to dill.

The term coriander comes from the Greek *koris,* meaning bedbug, because the leaves and green fruit of this plant have an odor similar to bedbugs. Although some find it objectionable, it imparts a distinctive authentic taste to Mexican and Asian foods. Coriander seeds are used in pickling, curries, and for flavoring alcoholic beverages such as gin. The "seed," which is actually the whole fruit, is a two-seeded schizocarp borne in umbels and is approximately 0.1–0.3 in. (3–7 mm) in diameter. Longitudinal ridges run from top to bottom.

Cultivation is accomplished by planting seed in the spring at a rate of 9–14 lb/acre (10–15 kg/hectare). Pollination and weed control are similar to that of dill. *See* APIALES; SPICE AND FLAVORING.

Seth Kirby

Bibliography. S. Arctander, *Perfume and Flavor Materials of Natural Origin*, 1960; L. H. Bailey, *Manual of Cultivated Plants*, rev. ed. 1975; F. Rosengarten, *The Book of Spices*, 1969.

Coriolis acceleration

An acceleration which arises as a result of motion of a particle relative to a rotating system. Only the components of motion in a plane parallel to the equatorial plane are influenced. Coriolis accelerations are important to the circulation of planetary atmospheres, and also in ballistics. They are so called after G. de Coriolis, the French engineer and mathematician whose analysis of the phenomenon was published in 1835. *See* ACCELERATION; BALLISTICS.

Newton's second law of motion is valid only when the motions and accelerations are those observed in a coordinate system that is not itself accelerating, that

is, an inertial reference frame. In order to utilize familiar concepts in mathematical treatment, the Earth is commonly treated as if it were fixed, as it appears to one observing from a point on the surface, and the Coriolis force is introduced to balance the acceleration observed by virtue of the observer's motion in the rotating frame. *See* FRAME OF REFERENCE; NEWTON'S LAWS OF MOTION.

As with the influenced components of motion, the Coriolis force is directed perpendicularly to the Earth's axis, that is, in a plane parallel to the equatorial plane. Since the direction of its action is also perpendicular to the particle velocity itself, the Coriolis force affects only the direction of motion, not the speed. This is the basis for referring to it as the deflecting force of the Earth's rotation. The magnitude of the Coriolis force is equal to the product of twice the projection of particle velocity into the equatorial plane, with the angular velocity of the Earth's rotation ($7.292 \times 10^{-5}\ s^{-1}$).

The Coriolis force is directed to the right of the direction toward which the horizontal wind blows or the missile flies in the Earth's Northern Hemisphere, and to the left in the Southern. The horizontal component of the deflecting force is zero at the Equator, and a maximum at the poles. While negligible for the dynamics of small-scale phenomena such as cumulus convection and thunderstorms, Coriolis effects are critical to the dynamics of the Earth's large-scale wind currents. Indeed, the general circulations of planetary atmospheres owe their zonal nature and some other characteristics largely to the forces that arise by virtue of planetary rotation.

A simple illustration of a Coriolis effect in the Northern Hemisphere is afforded by a turntable in counterclockwise rotation, and an external observer who moves a marker steadily in a straight line from the axis to the rim of the turntable. The trace on the turntable is a right-turning curve, and obviously this is also the nature of the path apparent to an observer who rotates with the table. Now consider the contrasting case of an air parcel near the North Pole that moves directly south (away from the Earth's axis of rotation) so that its motion to an observer on the Earth is in a straight line. To a nonrotating observer in space, this same motion appears curved toward the east because of the increased linear velocity of the meridian at lower latitudes. The force necessary to produce the eastward acceleration in the inertial frame is equal to the Coriolis force and would be produced by a gradient of air pressure from west to east, and shown by north-south–oriented isobars. In the absence of the pressure gradient force, the Coriolis force would cause the trajectory of the southward-moving air to curve westward on the Earth's surface. Then an air parcel moving uniformly away from the North Pole along a line which appears straight to an observer in space would appear to earthbound observers to curve westward. *See* GEOSTROPHIC WIND; ISOBAR (METEOROLOGY). Edwin Kessler

Bibliography. L. J. Battan, *Fundamentals of Meteorology*, 1979; J. M. Wallace and P. V. Hobbs, *Atmospheric Science, An Introductory Survey*, 1977.

Corn

Zea mays occupies a larger area than any other grain crop in the United States, where 60% of the world production is grown.

Production

Production is widely distributed throughout the United States, but it reaches its greatest concentration in the states of Iowa, Illinois, Indiana, Ohio, Nebraska, Minnesota, Wisconsin, Michigan, and Missouri. This area, called the Corn Belt, is characterized by moderately high temperature, fertile, well-drained soils, and normally adequate rainfall. The Corn Belt accounts for approximately 80% of corn production in the United States. Although corn is grown in the United States primarily for livestock feed, about 10% is used for the manufacture of starch, sugar, corn meal, breakfast cereals, oil, alcohol, and several other specialized products. In many tropical countries, corn is used primarily for human consumption.

Origin and description. From its presumed origin in Mexico or Central America, corn has been introduced into all the countries of the world that have suitable climatic conditions. However, the corn of today is far different from the primitive types found in prehistoric sites excavated in Arizona, Mexico, and Peru. The origin of corn is still unsettled, but the most widely held hypothesis assumes that corn developed from its wild relative teosinte (*Z. mexicana*) through a combination of favorable mutations, recognized and selectively propagated by early humans. Corn migrated from its center of origin and was being cultivated by the Indians as far north as New England upon the arrival of the first European colonists, whose survival was due largely to the use of corn as food.

Botanically, corn is a member of the grass family. Certain seed types within the species *Z. mays* were given subspecific rank by early investigators. Among these types are the dents (*indentata*), flints (*indurata*), soft or flour (*amylaceae*), sweet (*saccharata*), pop (*everta*), and pod (*tunicata*). Each form (botanical variety) is conditioned by fairly few genetic differences, and each may exhibit the full range of differences in color, plant type, maturity, and so on, characteristic of the species. All types have the same number of chromosomes (10 pairs), and all may be intercrossed to produce fertile progeny. Dent corns are the most important in the United States. Kernel texture is less homogeneous in other areas of the world. Soft types tend to predominate at higher elevations, flint types at lower elevations. This pattern is due, in part, to differences in methods of food preparation, but it has also been influenced by the greater resistance of flint types to stored-grain insect pests, which are more prevalent at lower elevations. Sweet corn is grown more extensively in the United States than in any other country. It is eaten as fresh corn or canned or frozen. In other countries, flint, dent, or flour corns may be eaten fresh, but at a much more mature stage

than the sweet corn eaten in the United States. The commercial production of popcorn is almost exclusively American. Pod corn is grown only as a genetic curiosity. *See* CYPERALES; GENETICS; REPRODUCTION (PLANT).

Varieties and types. Corn is a cross-pollinated plant; the staminate (male) and pistillate (female) inflorescences (flower clusters) are borne on separate parts of the same plant (**Fig. 1**). Plants of this type are called monoecious. The staminate inflorescence is the tassel; it produces pollen that is carried by the wind to the silks produced on the ears. Natural cross-pollination permits maintenance of a high degree of genetic variability. As corn was carried by the Indians to different environments in North, Central, and South America, a large array of varieties ultimately was developed. Many of these primitive varieties still exist today. Extensive efforts have been devoted to the collection, classification, and maintenance of this valuable reservoir of genetic variability. Settlers migrating westward took with them their corn seed (botanically fruits). In time, by both planned and natural selection, a large number of varieties and local strains became established and adapted to the new ecological conditions. These strains varied in color of grain, plant, and ear characteristics, and in length of time required to reach maturity. Because of the variability associated with cross-pollination, corn is a highly plastic species, and varieties have been developed that are well adapted to a frost-free growing season of less than 90 days or, in contrast, grow 12–15 ft (3.5–4.5 m) tall and require more than 200 days to reach maturity, as in certain areas in Guatemala. *See* FLOWER; INFLORESCENCE; POLLINATION.

Hybrids. The development of varieties and strains of corn made possible the extension of its culture under diverse soil and climatic conditions. However, modern research methods, used for its further improvement, led to the present widespread use of hybrid corn. Essentially all of the corn grown in the United States is planted to hybrid seed, and this technology is being adopted in other areas of the world as rapidly as suitable hybrids are developed. *See* BREEDING (PLANT); HETEROSIS.

Fig. 1. A corn plant in full tassel and silk. The tassel produces pollen that is blown by wind to the silks. (*Courtesy of J. W. McManigal*)

Hybrid corn is the first generation of cross involving inbred lines. Inbred lines are developed by controlled self-pollination. When continued for several generations, self-pollination leads to reduction in vigor but permits the isolation of types which are genetically pure or homozygous. Intense selection is practiced during the inbreeding phase to identify and maintain genotypes having the desired plant and ear type and maturity characteristics, and relative freedom from insect and disease attacks. After the inbred lines have become fairly pure, they are further evaluated in hybrid combinations. Crosses involving any two unrelated lines will exhibit heterosis, that is, yields above the means of the two parents. Only a very few combinations, however, exhibit sufficient heterosis to equal or surpass hybrids in current commercial use. Hence large-scale experimental testing is required to identify new useful combinations.

The first commercial hybrids were double crosses produced by crossing two single crosses, each, in turn, having been produced by the crossing of two inbred lines. Hybrid seed for commercial planting is produced in special crossing blocks (isolated fields). The field is planted with the parents (inbred lines or single crosses) in alternating groups of rows.

The plants in blocks to be used as female parents are detasseled before silk emerges. Thus all seed produced will be pollinated by the male parent. After drying, sizing, and treatment, the seed harvested from the female rows becomes the hybrid seed of commerce. Since the early 1960s the use of single-cross hybrids has been increasing, and in the Corn Belt today probably accounts for 80% or more of the seed used.

Several techniques have been used to control pollination. Initially, manual removal of the tassels was used exclusively. Then cytoplasmic male sterility was used. This type of male sterility is transmitted only through the cytoplasm of the female parent. Since no fertile pollen is produced, no detasseling is required. With this system, pollen production in the farmers' fields is provided either by blending (artificial mixing of sterile and fertile types) or by the use of genetic fertility-restoring factors introduced through the male parent. In 1970 possibly 90% of the hybrid seed used in the United States utilized a single type of cytoplasmic sterility: the Texas type. This cytoplasmic type proved to be susceptible to a new mutant race (race T) of one of the leaf-blight fungi, *Helminthosporium maydis*, and thus the use of this type of sterility was distincontinued. Investigations of other types of cytoplasmic sterility are under way. Machines which cut or pull tassels in the seed fields have been developed and are being used

extensively. Other types of genetic control of pollen production have been developed, and some attention is being given to the possible use of chemical sterilants.

Companies specializing in hybrid seed production produce and market the entire seed supply. Gains from the use of hybrid seed have been spectacular. The shift from open-pollinated varieties to hybrids, beginning in the 1930s, accounted for a yield increase of 25%. In succeeding years these first hybrids have been replaced by newer, higher-yielding types which have greater disease and insect resistance. Hybrid succession is a continuing process; few hybrids have a commercial life of more than 5 years. The per-acre yield of corn has increased about threefold since the first introduction and use of hybrid seed. *See* AGRICULTURAL SCIENCE (PLANT).

Planting. Planting dates depend upon temperature and soil conditions. Germination is very slow at soil temperatures of 50°F (10°C), and seedling growth is limited at temperatures of 60°F (16°C) or below. Because of these temperature relations, planting begins in Florida and southern Texas in early February and is completed in New York and New England by late May or early June. Planting rates are influenced by water supply, soil type, and fertility and by the maturity characteristics of the hybrid grown. Planting rates for full-season hybrids may vary from 8000 to 12,000 plants per acre (20,000 to 30,000 per hectare) in eastern Texas to 20,000 to 24,000 (50,000 to 59,000) in the central Corn Belt. Check-planting, with hills and rows spaced 38–42 in. (95–105 cm) apart, has been almost completely replaced by drilling. With planting rates above 16,000 plants per acre (40,000 per hectare), drilling in rows 24–36 in. (60–90 cm) apart has become common practice. The use of nitrogen fertilizer has increased greatly; lesser amounts of phosphorus and potash are applied as needed. Soil tests and experience provide useful guides to economic rates of fertilizer use under different environmental conditions.

There has been a marked trend toward reducing the amount of tillage in seedbed preparation. The variations in the planting systems used are influenced by soil type and topography, amount of crop residues on the soil surface, and the weed-control system to be employed. The use of the minimum tillage system, in which all soil preparation as well as planting is performed in one passage across the field, is increasing. This system, and less extreme modifications of the conventional system, reduces total energy costs. *See* AGRICULTURAL SOIL AND CROP PRACTICES.

Weed control. Weed control may be effected by cultivation or herbicides or both. With minimum tillage planting methods, weed control may be more difficult than with conventional seedbed preparation because of crop residues left on or near the surface. Weeds are easiest to control mechanically when they are small, preferably before they emerge. The most efficient and economical tool for killing weeds at this stage is the rotary hoe. Unfortunately this or other mechanical cultivation systems are relatively ineffective during extended periods of rainy weather. Disk or shovel cultivators are used after the corn is 6–8 in. (15–20 cm) in height.

Increasing use has been made of selective herbicides. Most broadleaf and grassy weeds may be controlled without damage to corn. Herbicides may be applied either "preplant" or "preemergence." If applied prior to planting, herbicides may be left on the soil surface or incorporated into the soil. Preemergence applications may be made as the seed is planted, or at any time prior to crop emergence. The choice of the many herbicides available is dependent upon many factors, including the type of weeds to be controlled and the farming system used. Under unfavorable early-season weather conditions, neither mechanical nor chemical control nor their combination may be completely effective, and sizable yield reductions may result. *See* HERBICIDE.

Harvesting. Harvesting of corn has undergone a major revolution since 1930. In the 1930s most corn was husked by hand, and the ears were stored in slatted cribs. Harvesting began when the moisture content of the grain had been reduced, by natural drying, to 20–24%. The mechanical picker supplanted hand harvesting. The mechanical picker, in turn, has been replaced by the picker-sheller or corn combine, which harvests the crop as shelled grain (**Fig. 2**). Energy requirements, however, have increased materially over the years. When harvested as shelled grain, at a relatively high moisture content (20–30%), the grain must be dried artificially for safe storage. High-moisture corn to be used for livestock feed may be stored in airtight silos or may be treated with certain chemical preservatives such as propionic acid. Corn stored under either of these systems is not suitable either for industrial processing or for seed. *See* AGRICULTURAL MACHINERY.

G. F. Sprague

Fig. 2. Two-row mounted picker for harvesting corn.

Productivity. Corn is highly productive largely because it can use solar energy so efficiently. It is classed as a C-4 species, which means that its leaves use the more efficient of the two photosynthetic processes found in crop plants. This is the same photosynthetic system used by sugarcane, sorghum, other tropical grasses, and many persistent weeds. In addition to relatively efficient leaves, corn's growth habit allows planting densities thick enough to almost completely shade the soil. Thus, a well-managed cornfield receives almost all the solar energy the environment provides and uses it as efficiently as any plant known. *See* PHOTOSYNTHESIS.

The higher productivity of modern corn hybrids, as compared with earlier hybrids, when both are grown under disease-free conditions, results to a considerable degree from their ability to tolerate higher plant populations without excessive barrenness. This capability is largely due to the greater plant-to-plant uniformity introduced by single cross hybrids and to the development of varieties with more erect leaves. Such foliage minimizes mutual shading and permits more uniform illumination of all leaf surfaces.

The actual rate of dry matter production by an area of corn with ample water and fertility depends on the average amount of light and on the temperature. The highest value reported, 2 oz (52 g) of dry matter per 11 ft^2 (1.0 m^2) of land area per day, was for high-population corn under cloudless conditions at Davis, California. A more representative range of values for the Corn Belt of the United States would be 0.9–1 oz (26–32 g) of dry matter per 11 ft^2 (1.0 m^2) per day. Day-to-day variation is large because of differences in cloudiness and in temperatures. Such estimates suggest dry matter growth rates of 230–285 lb per acre per day (258 to 319 kg per hectare per day).

The corn plant grows vegetatively until about silking, after which all weight increase is in the form of grain. There is little transfer of earlier growth to grain, so almost the entire grain yield results from photosynthesis during the grain growth period, which runs from silking to maturity. Thus, grain yield results from a combination of the rate of dry matter production after silking and the duration of the grain growth period. Contrary to much popular opinion, grain yields are highest under cool conditions, when the lengthened grain growth period more than compensates for the slower growth rate. Thus, yields for single planting are much greater at higher altitudes in the tropics than at warmer, lower locations. Highest average grain yields are in favored areas of Washington and Oregon in the United States, in the North Island of New Zealand, and in cooler long-season locations in Europe. Consistently high yields in the Corn Belt of the United States result more from favorable soil moisture and excellent management than from climatic superiority.

Relationships among solar radiation, temperature, growing-season length, soil moisture, day length, soil fertility, and corn genotype in producing grain yields are complex and not well understood. Attempts to study the system as a whole, using simulation models on digital computers, may add considerably to knowledge of the subject. William G. Duncan

Diseases

More than 100 diseases of corn are known worldwide; yet damage or yield loss from disease seldom forces a grower out of production. World reduction in grain production due to disease is estimated to average 9% and in the Corn Belt to average 7–17%. The widespread practices of continuous cropping, increased fertilization, high plant populations, and relative genetic uniformity of hybrids have sometimes resulted in epidemics: Stewart's bacterial wilt, northern corn leaf blight epidemics in 1939–1943, the maize dwarf mosaic epidemics in the 1960s, and the southern corn leaf blight epidemic of 1970–1971.

Diseases are caused by bacteria, fungi, nematodes, viruses, several mycoplasma-like organisms, and one parasitic seed plant (*Striga* spp.). Pollutants, pesticides, and unfavorable climatic and soil factors are responsible for still other diseases.

Seedling diseases. Seed decay and seedling blight caused by seed- and soil-borne *Diplodia, Fusarium, Penicillium,* and *Pythium* species are troublesome when grain is planted in cool, moist soil. Planting sound, mature, and chemically treated seed reduces these problems.

Leaf diseases. Of approximately 25 leaf blights, only a few are economically important in the United States. Southern corn leaf blight, caused by *Helminthosporium maydis,* accounted for a 15% loss in yield in the 1970 epidemic in the nation. Yellow leaf blight, caused by *Phyllosticta maydis,* occurred about the same time, and both blights are found most commonly in hybrids with the Texas male-sterile cytoplasm. Northern corn leaf blight, caused by *Helminthosporium* (=*Exserohilum*) *turcicum,* has been destructive and can cause 50% loss in yield if the disease strikes before pollination. With new conservation tillage practices, other leaf diseases have become increasingly important due to fungi that overwinter on crop debris left on the ground. Resistant hybrids and, sometimes, crop rotation are generally used for control. Two of the bacterial leaf blights are of major concern. Stewart's wilt, caused by *Erwinia stewartii,* results in a severe wilt of sweet corn and a leaf blight of dent corn in eastern North America. Control is obtained by use of resistant hybrids or by insecticides that control the flea beetle vector. Goss's bacterial wilt and blight, or leaf freckles, occurs in Colorado, South Dakota, Nebraska, and Kansas, and is caused by *Corynebacterium nebraskense*. All types of corn are susceptible, and various degrees of blighting and stunting occur.

Smuts and rusts. Common or boil smut, caused by *Ustilago maydis,* is found wherever corn is grown, but losses in yield seldom exceed 2% over a wide area. Galls may appear on any part of the plant and are most prevalent in vigorously growing plants in soil high in organic matter and nitrogen, often under dry conditions, and increase with hail and other

Fig. 3. Corn smut on various parts of plant. (*University of Minnesota Agricultural Experimental Station*)

mechanical injuries (**Fig. 3**). Large galls on or above the ear are more destructive than galls below the ear. Control is to plant resistant hybrids. The small cinnamon brown to black pustules of common rust, caused by *Puccinia sorghi,* on leaves are found wherever corn is grown. Most widely used hybrids have good field resistance. Southern rust, *P. polysora,* found mainly in the southeastern United States and other countries, appears to be spreading to northern states of the Corn Belt.

Stalk and root diseases. Stalk rots are universally important and among the most destructive diseases of corn in the world (**Fig. 4**). They are caused by *Fusarium moniliforme, F. roseum* "Graminearum" (*Gibberella zeae*), *Macrophomina phaseolina,* and to a lesser extent by *Diplodia zeae* and *Colletotrichum graminicola,* and several dozen other fungi and bacteria. Losses vary by season and region, but yield losses of 10–20% can occur among susceptible hybrids in the United States and 25–33% in other countries. Conditions favoring disease may cause plants to die before maturity. This results in poorly filled ears, but the greatest loss may come from stalk lodging that interferes with harvesting, or from ears fallen to the ground. Many factors affect stalk rot: unbalanced soil fertility, high plant populations, dry weather, leaf diseases, insect damage, hail, and, sometimes, continuous cropping to corn. The best control is to grow resistant hybrids and to follow good management practices. Root rots are caused usually by the same organisms that cause stalk rot, except that nematodes (eel- or roundworms) are sometimes involved. In fact, more than 40 species of nematodes have been reported to feed on corn roots. They can stunt plants, cause yellowing, and aid root-infecting fungi. *See* NEMATA.

Ear diseases. Considerable damage to ears from ear and kernel infection occurs in humid regions where rainfall is above normal late in the season, and from insect and bird feeding, early frosts, and fallen stalks (**Fig. 5**). Infected grain reduces yield and lowers grain quality, and grain may contain mycotoxins (aflatoxins, estrogens, and vomitoxin) which makes animals ill when used as feed. Infections occur in the field from fungi such as *Diplodia, Fusarium* (*Gibberella*), *Helminthosporium,* and *Nigrospora,* or in storage from such fungi as *Aspergillus, Cladosporium,* and *Penicillium*. Storage fungi can be controlled if grain is stored at 13–15% moisture and at uniform temperatures between 39 and 50°F (4 and 10°C), with provision for adequate aeration.

Viral and mycoplasma diseases. Prior to 1963, viruses were not a factor in corn production except in localized areas outside the United States. That year, maize dwarf mosaic virus (MDMV) caused economic losses in the United States. Almost complete destruction of susceptible corn occurred in some areas. MDMV has been identified in 15–20 states in the Corn Belt, as well as in the East and South. Aphids are vectors of MDMV, but leafhoppers are vectors for maize chlorotic dwarf virus (MCDV). Wheat streak mosaic virus (WSMV), widespread on certain inbred lines and hybrid corns in the Corn Belt, is transmitted by the mite *Eriophyes tulipae*. Other naturally occurring viruses of corn are sugarcane mosaic (a complex of Potyviruses that contain the MDMV subgroup), maize mosaic virus, maize chlorotic mottle virus (MCMV), and maize rayado fino virus. Corn lethal necrosis is a synergistic interaction between MCMV and either MDMV or WSMV, and is potentially destructive in the United States, but is currently confined to Kansas and Nebraska.

Some diseases once thought to be caused by viruses are now known to be caused by a spiroplasma or a mycoplasma-like organism. Corn stunt, transmitted by a leafhopper, occurs in Texas and in the Rio

Fig. 4. Corn stalk rot, one of the most widespread corn diseases. (*a*) Corn lodged from stalk rot. (*b*) Bacterial stalk rot, with healthy corn stem on left. (*Ohio Agricultural Research and Development Center*)

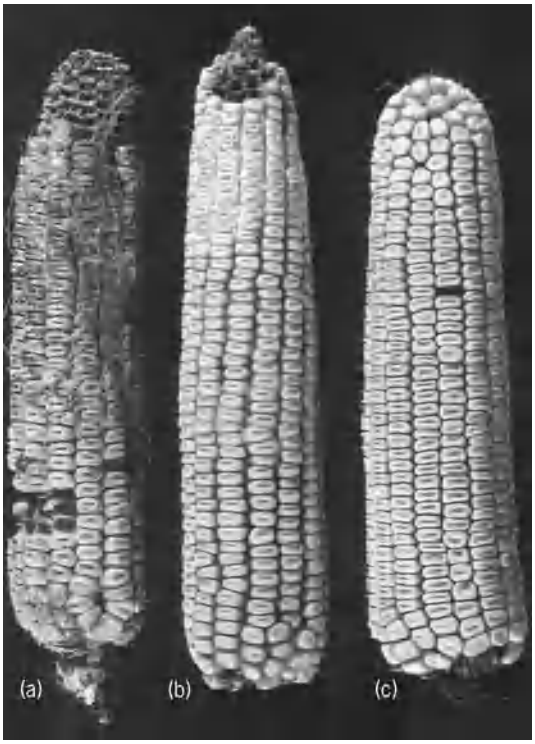

Fig. 5. Ear rots. (*a*) Storage rot. (*b*) Field rot. (*c*) Healthy ear. (*Ohio Agricultural Research and Development Center*)

Grande Valley, and is caused by a spiroplasma. The maize bushy stunt pathogen is a similar organism and has the approximate range of corn stunt in the United States and South America. The economic importance of virus diseases of corn has increasingly been recognized. *See* PLANT PATHOLOGY; PLANT VIRUSES AND VIROIDS.

Thor Kommedahl

Processing

Corn kernels (seeds) are subjected to both wet and dry milling. The goal of both processes is to separate the germ, the endosperm, and the pericarp (hull). The mean moisture content of corn kernels is 17% (range 7–23%). Kernel composition on a percent dry basis is given in **Table 1**. *See* CELLULOSE; HEMICELLULOSE; LIGNIN.

TABLE 1. Composition of corn kernels, on percent dry basis

Components	Mean value, %	Range, %
Starch	71	64–78
Protein	9.9	8–14
Fiber	9.5	8.3–11.9
Hemicelluloses	6.2	5.8–6.6
Cellulose	3.1	3.1–4.1
Lignin	0.2	
Oil	4.5	3.1–5.7
Ash	1.4	1.1–3.9
Sugars	2.6	1.0–3.0

Wet milling. Wet milling separates the chemical constituents of corn into starch, protein, oil, and fiber fractions (**Fig. 6**), the primary objective being to produce refined corn starch. The world production of corn grain in 1996–1997 was approximately 640 million tons (580 million metric tons) with more than 40% being produced in the United States. Products from milling of the United States crop are shown in **Table 2**. The amount wet-milled in 1997 was 16.5% of the crop.

There are variations in the equipment used for wet milling from plant to plant and within processing plants. The process begins with shelled No. 2 yellow dent corn being cleaned on vibrating screens to remove both coarse and fine material. The kernels are then placed in steep tanks that hold 2000–13,000 bushels (60–360 tons or 55–330 metric tons) and covered with steepwater (8–9 gal/bushel or 1.2–1.3 liter/kg) heated to 118–125°F (48–52°C). The steepwater is transferred in a countercurrent system, usually through batteries of about 12 tanks, with as many as 50 being used. The original steepwater contains 0.12–0.20% of added sulfur dioxide (SO_2). Dissolution of sulfur dioxide in water produces sulfurous acid, which assists in breaking down the starch-protein matrix. Kernels release about 6.5% of their dry substance as solubles into the steepwater. From the solubles, bacteria in the steepwater produce lactic acid, an important component for pH control. As the steepwater advances through the tanks in stages, the SO_2 content decreases to 0.01% or less and the lactic acid content increases to

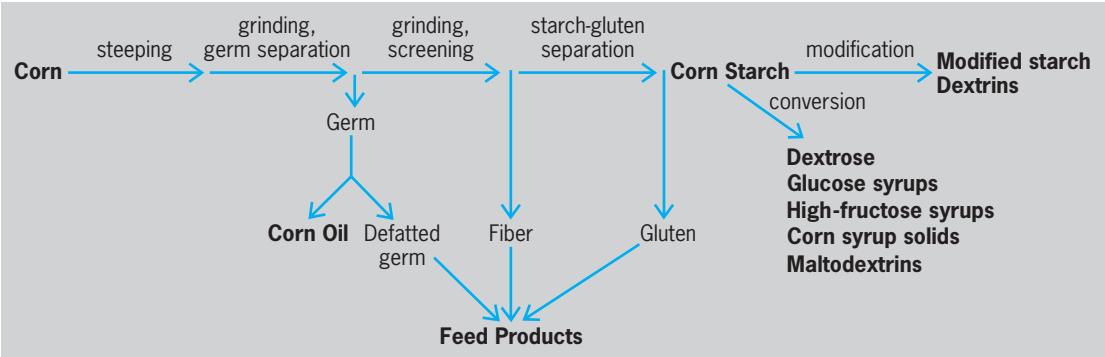

Fig. 6. Corn wet-milling process.

TABLE 2. Products obtained from corn in the United States*

Products	Production, 10^6 lb	Percent of nonfeed use
Wet-milled products	37,885	56.8
Starch products†	5,938	13.2
Conversion products‡	31,947	43.6
Fuel ethanol		28.4
Dry-milled products		7.5
Beverage alcohol		7.3
Coproducts of wet milling§		
Corn oil	1,284	
Corn gluten feed and corn oil meal	12,547	
Corn gluten meal	2,593	
Steepwater	1,118	

*Based on 1997 data; Corn Refiners Association, Washington, DC.
†Corn starches, modified corn starches, and dextrins.
‡Dextrose (d-glucose), glucose syrups, high-fructose syrups, D-fructose, syrup solids, maltodextrins.
§Does not include coproducts from fuel alcohol (ethanol) production.

1.6–1.7%. However, the pH remains a rather constant 4.0–4.1 throughout the process. The steeping process usually takes 20–40 h. At its end, the kernels are hydrated (45% moisture), swollen, and softened. The final steepwater is evaporated to 45–50% solids and mixed with corn fiber. The mixture is sold as corn gluten feed after drying. *See* PH.

The remainder of the process is continuous. The softened corn kernels are passed through coarse grinding mills that have one stationary and one rotating disk. The clearance between the disks is adjusted to minimize both the number of whole kernels and the number of broken germs. The discharge is pumped through banks of hydroclones to remove the germs, which have a lower density. The germs are washed on screens, dewatered in screw presses to about 50% solids, and dried. To recover the oil, the germs are heated to 250°F (120°C), steamed or passed through an extruder, and pressed in an oil expeller which reduces the oil content from 45% to 6%. Alternatively, after reducing the oil content to 13–20%, the germs are flaked and extracted with hexane. Crude corn oil is shipped to refiners. The remaining corn germ meal contains 23–25% protein. *See* FAT AND OIL (FOOD).

The slurry of corn kernels minus their germs is screened to separate "fiber" (mostly the pericarp) from the starch and protein. At this point, the fiber contains some attached starch. The slurry of fiber is again passed through rotating disk mills that free the starch with a minimum of fiber destruction. The milled slurry is again screened to separate the starch from the fiber. The fiber fraction remaining on the screens is dewatered to 42–48% moisture, mixed with evaporated steepwater, dried to 8–11% moisture, and sold as corn gluten feed (21% protein), often after pelletizing.

The starch slurry now contains only starch, protein, and some solubles. Because the protein has a specific gravity of 1.06 and the starch a specific gravity of 1.5, they can be separated in centrifuges. This separation produces a gluten fraction of more than 68% protein, which is dried to about 12% moisture and sold as corn gluten meal. *See* CENTRIFUGATION.

Most of the remaining protein (3–5% of the weight of starch) is removed from the starch by countercurrent washing in stages using banks of hydroclones 0.4 in. (10 mm) in diameter. The slurry of refined starch (40% solids) is centrifuged or vacuum-filtered to produce a cake that is dried in belt or flash driers and milled; or is pumped into tanks for various modifications; or is pumped through jet cookers for various conversions.

A typical composition of commercial common corn starch is starch 88%, moisture 11%, protein 0.35%, ash 0.1%, lipid 0.87%, and crude fiber 0.1%. A typical yield distribution of products obtained from the wet-milling process is starch of 11% moisture (74% yield), corn gluten feed of 21% protein (27%), corn gluten meal of 60% protein (6%), and corn oil (4%). Corn gluten feed and meal are used as animal feed. *See* ANIMAL FEEDS.

While by far the majority of the corn subjected to wet milling is common corn, waxy maize (less than 1% apparent amylose) and high-amylose corn (class V around 55% apparent amylose, and class VII around 73% apparent amylose) are grown under contract for wet millers and used for specialty starch production.

Corn starch. The United States produces almost 60% of the world's starch. Since corn is the source of more than 90% of the starch produced in the United States and starch is also produced from corn in other countries, by far the primary source of starch is corn. The world production of starch was around 40 million tons (36 million metric tons) in 1996, with 80% coming from corn. A large proportion (80–85%) of this starch is converted into other products such as sweeteners. Some is hydrolyzed to D-glucose which is used as the carbon source in fermentation to produce ethanol, a fuel additive; one-third of the starch produced by wet-milling of corn in the United States in 1997 was used to produce fuel ethanol. *See* ALCOHOL FUEL; ETHYL ALCOHOL; STARCH.

Not only does the source of starch vary from country to country, so does the ratio of products going to the various end uses. In the United States the starch of commerce is predominantly corn starch, and the primary market for corn starch, other than for the production of sweeteners and ethanol, is in the paper industry. Manufacture of paper, paper board, and corrugated board products consumes about 60% of the starch produced in the United States. About 20% of the starch is used in prepared foods. Approximately 20% of corn starch products are used as a textile size, in building materials, and in various other consumer and industrial products. The great majority of all corn starch used in any of these applications is first modified by etherification, esterification, acid-catalyzed depolymerization, oxidation, or a physical process (**Fig. 7**). It is not uncommon for a starch product to have been doubly, triply, or quadruply modified to

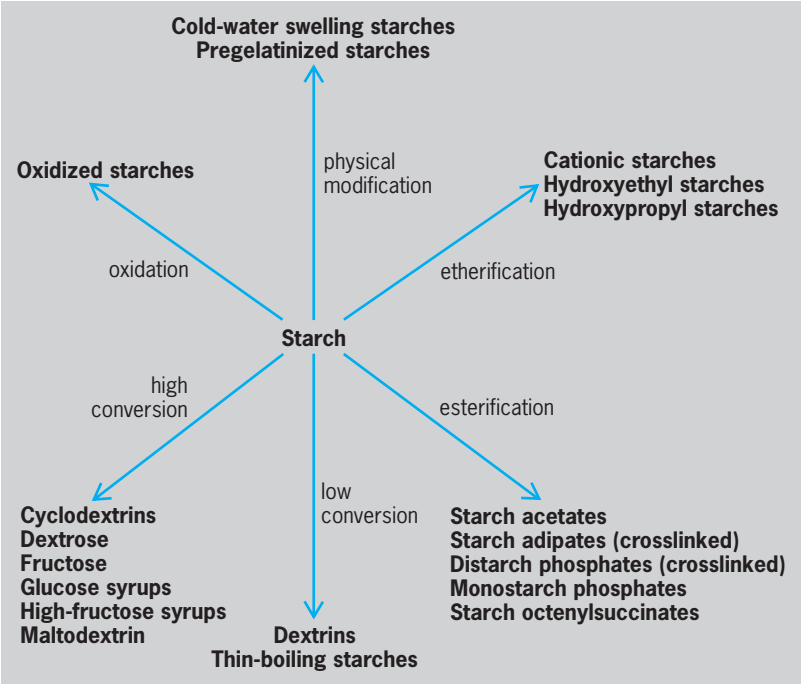

Fig. 7. Starch modification.

give it more desirable attributes for a particular application. Most modified food starch products made in the United States are based on waxy maize starch. For other corn starch products, the base starch is common corn starch.

Sweeteners. Worldwide, the production of nutritive sweeteners is the largest use for the starch obtained from corn. Residents of the United States consume more sweetener derived from corn than they do sucrose (sugar). *See* SUGAR.

Sweeteners can be made from any starch and are essentially identical regardless of which starch is used as the raw material, but in the United States all such products are made from corn starch and are often referred to as refinery products. The manufacture of corn sweeteners begins with the wet milling process. The starch, obtained as a slurry of granules in water, is first cooked, or pasted (**Fig. 8**). Then, the starch polymers are hydrolyzed (depolymerized) to the desired extent using an acid, an enzyme, a combination of enzymes, or an acid-enzyme combination. In all-enzyme conversion processes (the most commonly used), the starch paste is first rapidly thinned with an alpha-amylase (liquefying enzyme), then treated with one or more so-called saccharifying enzymes to convert partially degraded polymers into lower-molecular-weight carbohydrates—for example, to make high-glucose or high-maltose products, that is, solutions containing primarily D-glucose or maltose, respectively. The resulting solutions are refined and concentrated to 70–80% solids. These syrups, of whatever composition, are known worldwide as glucose syrups, but in the United States are often called corn syrups.

The syrups and all other breakdown products of starch are characterized by a value known as dextrose equivalency (DE). The DE value of a hydrolysis product is its reducing power expressed as a percentage of the reducing power of pure dextrose (D-glucose) and is inversely proportional to its

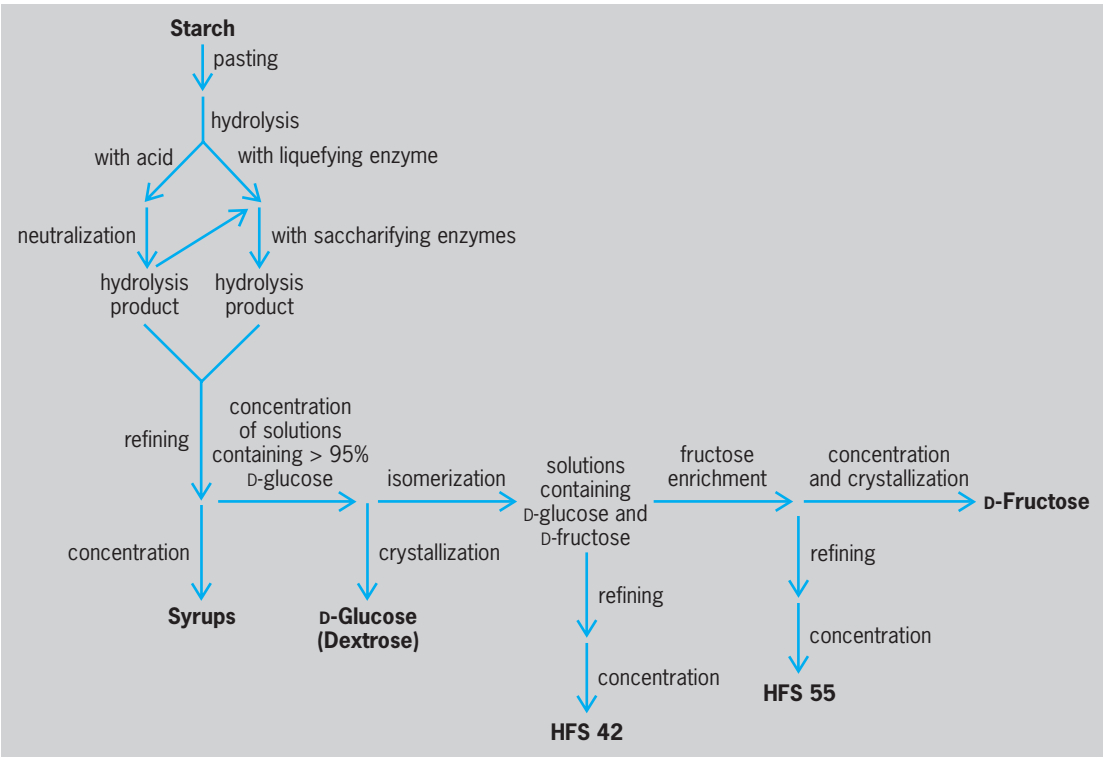

Fig. 8. Sweetener manufacture from starch.

average molecular size. Starch has a DE value of 0, while a solution of D-glucose has a DE value of 100. (Glucose and dextrose are interchangeable terms for the same substance, with D-glucose being the formal chemical name and dextrose being a term of commerce. Commercial products which contain less than 98% of the substance are usually referred to as glucose syrups, and the pure crystalline material is usually called dextrose.) Liquid solutions of nutritive saccharides obtained from starch (glucose syrups) have DE values of 20 or more. Their degree of sweetness is a function of the types and amounts of saccharides they contain. The degree of hydrolysis, that is, the composition of the hydrolyzate, also determines viscosity, fermentability, flavor enhancement, and freezing-point depression.

Glucose syrups are used in a wide variety of products. Besides having different sweetness levels, the functionalities they impart vary with saccharide composition. They may provide body (viscosity) or desirable mouthfeel to certain liquid products such as pancake and waffle syrups and certain beverages. They bind and hold water, providing desirable texture and mouthfeel to other products; for example, a 65 DE syrup is often used in soft cookies to help retain chewiness. Glucose syrups may also be used in bakery products to generate crust color and flavor or to provide texture to reduced-fat products without making them too sweet. They are often used in icings and glazes. A major use can be in candy manufacture, but relative amounts used in various applications vary from country to country. In addition to food uses, glucose syrups are a feedstock for fermentation processes that produce ethanol and other alcohols, vitamins, amino acid feed supplements, citric acids, and antibiotics. High-glucose syrups are hydrogenated to produce the polyol sorbitol; high-maltose syrups are hydrogenated to produce the polyol maltitol; and other syrups are used to produce hydrogenated starch hydrolyzates.

When starch is completely hydrolyzed, that is, converted into its monomer units, the only product is D-glucose (dextrose), which can be crystallized from concentrated solutions. Crystalline dextrose is produced in several different forms. It is about 60% as sweet as sucrose and has limited food applications. Dextrose monohydrate absorbs heat as it dissolves, so it imparts a cooling sensation when used in a sugar coating, and is used in donut glazes and sandwich cookie and sugar wafer fillings. It is used to make polydextrose, a food bulking agent and dietary fiber with reduced caloric value. Dextrose is used in preparing solutions for intravenous injection and for renal dialysis, in tableting, and in oral medications.

Isomerization of some of the D-glucose in a high-glucose hydrolyzate to D-fructose produces high-fructose corn syrups (HFCS), which are known simply as high-fructose syrups (HFS) outside the United States. Fructose is approximately 20% sweeter than sucrose on an equal weight basis. Converting some of the D-glucose to D-fructose makes a product similar to invert sugar, which is the 50:50 mixture of glucose and fructose made by hydrolysis of sucrose and is sweeter than sucrose. Two major products are HFCS 42 and HFCS 55, which contain, respectively, 42% and 55% fructose. HFCS is the sweetener for almost all soft drinks made in the United States that do not use a high-intensity sweetener. Most of the HFCS in colas is HFCS 55, which is made by enriching HFCS 42 with fructose. Many fruit-flavored drinks use HFCS 42. HFCS 42 is also used in most commercially produced yeast-raised breads, rolls, and cakes. The sugars in HFCS are fermentable. When HFCS is used in breadmaking, for example, yeast preferentially ferments glucose, leaving a higher percentage of fructose and the desired slight sweetness. HFCS is used in ice creams, frozen desserts, yogurt, chocolate milk, and other dairy products. It is blended with sucrose and a glucose syrup for use in fruit canning. The world production of high-fructose syrups was 12.5 million tons (11.3 million metric tons) in 1998–1999, with three-fourths of that amount (9.5 million tons or 8.6 million metric tons) being produced in the United States.

Crystalline (pure) fructose provides a rapid but short sweet taste, allowing other flavors to be sensed. For this reason, it is used in some fruit and fruit-flavored products such as jams, jellies, drinks, and sorbets.

Other starch conversion products. Corn starch is less extensively depolymerized to make products other than sweeteners. Very slight hydrolysis makes products known as acid-modified or thin-boiling starches. These products are still granular but cook more easily than the native starch. Their primary use is as film formers and adhesives; they also form gels with increased clarity and strength. A little more modification with an acid produces dextrins. Dextrins have some degree of cold-water solubility and generally have the same applications as thin-boiling starches. One application is as remoistenable adhesives on envelopes.

Hydrolysis catalyzed by acid or enzymes to DE values less than 20 produces starch oligomers, which are known as maltooligosaccharides or maltodextrins. Maltodextrins are easily dispersible and rapidly soluble, form films, have low hygroscopicity, and are generally bland, although some have slight sweetness. They are used extensively in foods for their bulking and binding properties and the protection they give to frozen foods. Special maltodextrins have properties that make them useful as fat replacers or sparers. Hydrolysis to DE 20–60 gives mixtures of breakdown products that, when dried, are known as corn syrup solids. Corn syrup solids dissolve rapidly, are mildly sweet, and are used as bulking materials in food.

Any one and often mixtures of the sweetener and other starch conversion products—that is, mixtures of dextrose, fructose, glucose or corn syrups, corn syrup solids, maltodextrins, hydrogenated sugars or syrups, and polydextrose—may be used to prepare functional, isotonic and oral rehydration beverages, enteral and parenteral fluids, infant formulas, and adult diet supplements.

TABLE 3. Typical compositions of products of corn dry milling in percent

Product	Mesh size	Moisture	Protein	Lipid	Crude fiber	Ash	Starch	Other polysaccharides
Coarse grits	−10 to +15	11.5	7.5	0.7	0.2	0.3	78	1.8
Regular grits	−15 to +30	11.5	7.5	0.7	0.2	0.3	78	1.8
Flaking grits	−3.5 to +6	11.7	7.0	0.6	0.2	0.2	78	2.0
Corn flour	−60 to +325	13.0	5.2	2.0	0.5	0.4	76	2.5

Dry milling. The amount of corn dry-milled in the United States in 1997 was 1.45% of the produced crop. Most of the processing is done by tempering-degerming systems. Corn, usually No. 2 yellow dent, but also white and other corns, is cleaned to remove metal, fines, pieces of cob, and broken corn. The whole kernels are then washed. Cleaned kernels are transferred to a tempering bin, where they are held for various times at various temperatures depending on the miller and the desired product. Tempered kernels at 20% moisture are passed through a degerminator, which removes the bran (pericarp) and germ while leaving the endosperm intact. But some bran and germ remain and must be removed by subsequent milling and aspiration; some endosperm remaining associated with the germ + bran fractions is also subsequently obtained.

Large pieces of endosperm, known as tail hominy, are dried, cooled, and sifted. Part of this fraction is isolated as large flaking grits which are used to make corn flakes breakfast cereal. The remainder is sent through roller mills for reduction to smaller fractions (coarse, medium, and fine grits; meals; flours).

The bran + germ fraction is dried, cooled, and aspirated to remove the bran. Corn oil is obtained from the germ by expelling or hexane extraction. The residue is the germ cake. Fines from the bran + germ fraction, the germ cake, the bran, and broken corn are combined, dried, and milled to produce hominy feed, the main by-product.

The endosperm may be converted into as many as 16 different fractions. The main products are regular grits, coarse grits, flaking grits, and corn flour (**Table 3**). Other products are corn cones and corn meal, −40 to +80 and −30 to +60 Standard U.S. mesh, respectively.

Nixtamalization. Nixtamalization is the process of cooking and soaking corn kernels in water containing calcium hydroxide (lime) to soften the pericarp and hydrate the protein matrix and starch of the endosperm. Subsequent washing removes leached solubles and excess lime. The cooked, steeped product, called nixtamal, is then ground, using stone attrition mills. The product, masa, is sheeted, cut into pieces, and baked, producing tortillas, tortilla chips, taco shells, and corn chips. The temperature, and particularly the time of each step, is varied depending on the desired product. The general process is as follows. Yellow, white, or blue corn (700 lb or 320 kg; alone or in mixtures) is cooked (200–212°F or 95–100°C) for around 4 to more than 30 min under alkaline conditions produced by adding 7 lb (3.2 kg) of powdered lime (calcium hydroxide; 1% based on the weight of corn) to 110 gal (415 liters) of water. Cooking is stopped by adding unheated water to drop the temperature to 150°F (65°C) or below. The corn kernels are then usually steeped for 8–16 h in this solution at about 105–115°F (40–45°C). Rotating drums equipped with pressurized water sprayers are used to wash the swollen kernels and remove the pericarp, lime, and solubles. The washed nixtamal is then ground in attrition mills with one stationary and one rotating stone to form masa. Subsequent processes vary, depending on the product desired. Masa may be dried for later tortilla making. *See* FOOD ENGINEERING; FOOD MANUFACTURING. James N. BeMiller

Bibliography. M. Freeling and V. Walbot (eds.), *The Maize Handbook*, 1993; G. E. Inglett (ed.), *Maize: Recent Progress in Chemistry and Technology*, 1982; R. W. Jugenheimer, *Corn: Improvement, Seed Production, and Uses*, 1985; L. R. Nault and J. C. Rodriguez (eds.), *The Leafhoppers and Planthoppers*, 1985; F. W. Schenck and R. E. Habeda (eds.), *Starch Hydrolysis Products: Worldwide Technology, Production and Application*, 1991; M. C. Shurtleff (ed.), *Compendium of Corn Diseases*, 1980; G. F. Sprague (ed.), *Corn and Corn Improvement*, 3d ed., American Society of Agronomy, 1988; G. M. A. Van Beynum and J. A. Roels (eds.), *Starch Conversion Technology*, 1985; S. A. Watson and P. E. Ramstad (eds.), *Corn: Chemistry and Technology*, 1987; R. L. Whistler, J. N. BeMiller, and E. F. Paschall (eds.), *Starch: Chemistry and Technology*, 1984.

Cornales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the superordinal Asteridae group of Eudicotyledon. The order consists of 5 families (Cornaceae, Grubbiaceae, Hydrangeaceae, Hydrostachyaceae, and Loasaceae), approximately 50 genera, and about 600 species. The members of the order show dramatic differences in habitat, ranging from temperate and tropical trees, to herbaceous perennials (Loasaceae), to submerged aquatics (Hydrostachyaceae). The order is mostly characterized by opposite leaves and flowers with four or five perianth parts that usually grow from the top of the ovary (epigynous) (see **illus.**). The fruit is either a fleshy berry or dehiscent capsule. The various species of dogwoods (*Cornus*) and sour gum (*Nyssa sylvatica*, family Cornaceae) are well known. Various species of Hydrangeaceae (deutzia, hydrangea, mock-orange)

Bunchberry or dwarf colonel (*Cornus canadensis*), a common dwarf, woody plant of the boreal forest of North America. (*a*) Four, large, white bracts surround the cluster of small flowers. (*b*) Each flower produces a red berry at maturity. (*Ken Sytsma, University of Wisconsin*)

are important cultivated shrubs. *See* ASTERIDAE; DOGWOOD; PLANT KINGDOM; TUPELO. K. J. Sytsma

Corner reflector antenna

A directional antenna consisting of the combination of a reflector comprising two conducting planes forming a dihedral angle and a driven radiator or dipole which usually is in the bisecting plane. It is widely used both singly and in arrays, gives good gain in comparison with cost, and covers a relatively wide band of frequency. Corner reflector antennas are used in very high-frequency (VHF) and ultrahigh-frequency (UHF) communications systems, telemetry systems, and electronic intelligence gathering systems.

The **illustration** shows the general configuration for a 90° corner reflector. The distance S from the driven radiator D to edge E need not be critically chosen with respect to wavelength; for a 90° reflector D may lie between 0.25 and 0.7 wavelength. The overall dimensions of the reflector need not exceed 2 wavelengths in order to approximate the performance obtained with very large reflectors. Gain, as compared with an isotropic radiator, is about 12 dB.

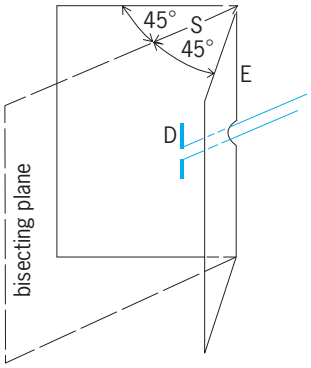

Diagram of corner reflector antenna.

The reflecting planes may be metal sheets or parallel wires separated by a small fraction of a wavelength and extending in the direction of current flow. *See* ANTENNA (ELECTROMAGNETISM). J. C. Schelleng

Bibliography. J. D. Kraus, *Antennas*, 2d ed., McGraw-Hill, 1988.

Corona discharge

A type of electrical conduction that generally occurs at or near atmospheric pressure in gases. A relatively strong electric field is needed. External manifestations are the emission of light and a hissing sound. The particular characteristics of the discharge are determined by the shape of the electrodes, the polarity, the size of the gap, and the gas or gas mixture.

In some cases corona discharge may be desirable and useful, whereas in others it is harmful and attempts are made to minimize it. The effect is used for voltage division and control in direct-current nuclear particle accelerators. On the other hand, the corona discharge that surrounds a high-potential power transmission line represents a power loss and limits the maximum potential which can be used. Because the power loss due to i^2r heating decreases as the potential difference is increased, it is desirable to use maximum possible voltage.

The shape of the electrodes has a very profound effect on the potential current characteristic. If the radius of curvature of the positive electrode is small compared to the gap between electrodes, the transition from the dark current region to the field-sustained discharge will be quite smooth. The effect here is to enable the free electrons to ionize by collision in the high field surrounding this electrode. One electron can produce an avalanche in such a field, because each ionization event releases an additional

Breakdown streamers in positive-point corona crossing from positive point to cathode below. (*From L. B. Loeb, Fundamental Processes of Electrical Discharge in Gases, Wiley, 1939*)

electron, which can then make further ionization. To sustain the discharge, it is necessary to collect the positive ions and to produce the primary electrons far enough from the positive electrode to permit the avalanche to develop. The positive ions are collected at the negative electrode, and it is their low mobility that limits the current in the discharge. The primary electrons are thought to be produced by photoionization (see **illus.**). For a discussion of ionic mobility *see* ELECTRICAL CONDUCTION IN GASES; DARK CURRENT.

The situation at the negative electrode is quite important. The efficiency for ionization by positive ions is much less than for electrons of the same energy. Most of the ionization occurs as the result of secondary electrons released at the negative electrode by positive-ion bombardment. These electrons produce ionization as they move from the strong field at the electrode out into the weak field. This, however, leaves a positive-ion space charge, which slows down the incoming ions. This has the effect of diminishing the secondary electron yield. Because the positive ion mobility is low, there is a time lag before the high field conditions can be restored. For that reason the discharge is somewhat unstable.

From the preceding it may be seen that the electrode shape is important. The dependence on the gas mixture is difficult to evaluate. Electronegative components will tend to reduce the current at a given voltage, because heavy negative ions have a low mobility and are inefficient ionization agents. The excess electron will not be tightly bound, however, and may be released in a collision. The overall effect is to reduce the number of electrons that can start avalanches near the positive electrode. Further, if a gas with low-lying energy states is present, the free electrons can lose energy in inelastic collisions. Thus it is more difficult for the electrons to acquire enough energy to ionize. In electrostatic high-voltage generators in pressurized tanks, it is quite common to use Freon and nitrogen to take advantage of this effect to reduce corona. In a pure monatomic gas such as argon, corona occurs at relatively low values of voltage.

In the potential current characteristic, the corona region is found above the dark current region and is field-sustained. Near the upper end it goes into either a glow discharge or a brush discharge, depending on the pressure. Higher pressure favors the brush discharge. *See* GLOW DISCHARGE.

For still higher potential difference, breakdown takes place and a continuously ionized path is formed. *See* ELECTRIC SPARK. Glenn H. Miller

Bibliography. American Society for Testing and Materials, *Engineering Dielectrics*, vol. 1: *Corona Measurement and Interpretation*, 1979; Institute of Electrical and Electronics Engineers, *IEEE Standard Definitions of Terms Relating to Corona and Field Effects of Overhead Power Lines*, IEEE Std. 539-1990, 1991; Y. P. Raizer, *Gas Discharge Physics*, 1991, reprint 1997.

Coronagraph

A specialized astronomical telescope substantially free from instrumentally scattered light, used to observe the solar corona, the faint atmosphere surrounding the Sun. Coronagraphs can record the emission component of the corona (spectral lines emitted by high-temperature ions surrounding the Sun) and the white-light component (solar photospheric light

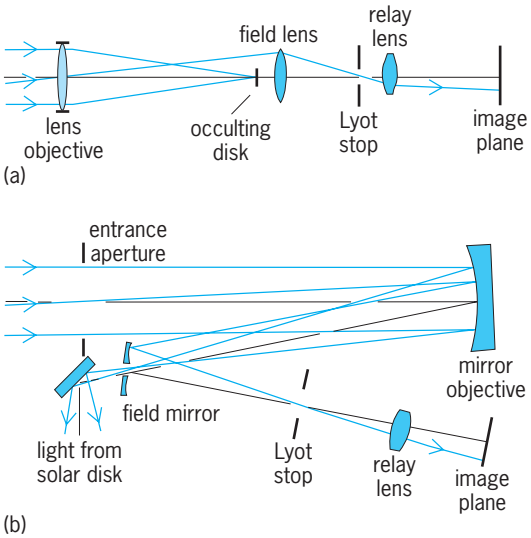

Basic designs of coronagraphs. (*a*) Lyot coronagraph with lens objective. (*b*) Reflecting coronagraph.

scattered by free electrons surrounding the Sun) routinely from high mountain sites under clear sky conditions. The emission and white-light components are typically only a few millionths the brightness of the Sun itself. Hence, the corona is difficult to observe unless the direct solar light is completely rejected, and unless instrumentally diffracted and scattered light that reaches the final image plane of the coronagraph is small relative to the coronal light. *See* SOLAR CORONA; SUN.

The basic design (**illus.** *a*), as invented by B. Lyot, has an occulting disk in the primary image plane of the telescope to block the image of the Sun itself. In addition, the primary objective (a lens or superpolished mirror) is specially fabricated to minimize scattered light. Also, light diffracted by the objective rim must be suppressed. For this, an aperture (Lyot stop) is placed at an image of the objective as produced by a field lens, with the aperture diameter slightly smaller than that of the objective image. A relay lens behind the Lyot stop forms the coronal image at the final image plane.

To observe the emission corona, optical filters or spectrographs are used to isolate the wavelengths corresponding to the various coronal emission lines and thus minimize the contribution due to the sky. To observe the white-light corona with ground-based coronagraphs, polarization subtraction techniques are required to discriminate the coronal electron-scattered light (polarized with the electric vector tangential to the limb) from the variably polarized sky background. Coronagraphs designed for operation on satellites usually have a circular mask out in front of the objective to completely shade it. This external occulting is more efficient in terms of instrumentally scattered light, when no sky background is present, than if only an internal occulting disk is used, but the inner corona is then vignetted, reducing the angular resolution in this region.

The development of superpolished mirror technology allowed the construction of coronagraphs based on mirror objectives, with significant advantages compared with the singlet-lens objectives used in traditional Lyot coronagraphs. For example, achromaticity allows high imaging quality independent of wavelength, simultaneous multiwavelength observations of the inner solar corona can be achieved, and large apertures are possible. In an off-axis design (illus. *b*), an annular field mirror is located at the image plane formed by the primary objective. The image of the solar disk passes through the hole in the annular mirror and is reflected out to the sky, while the field mirror reflects the coronal field to the lens system that forms the final image. A Lyot stop is located at the image of the entrance aperture, as formed by the field mirror. Raymond N. Smartt

Coronatae

An order of the cnidarian class Scyphozoa which includes mainly abyssal species. The exumbrella is divided by a circular or coronal furrow into two parts,

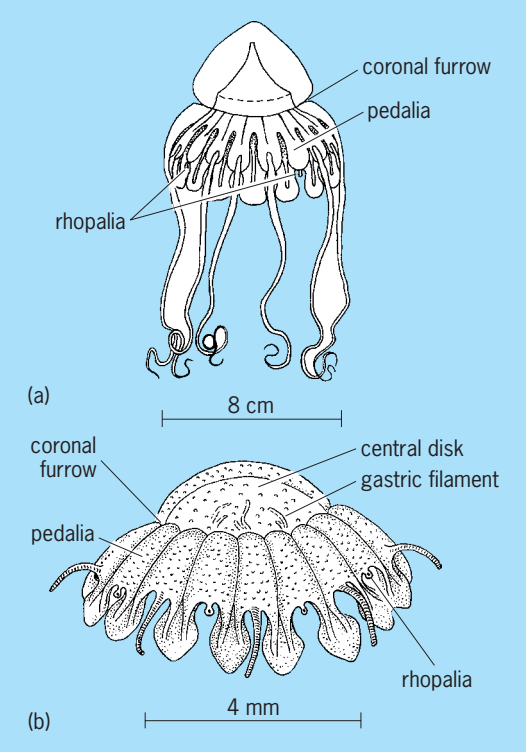

Coronatae. (*a*) *Periphylla*. (*b*) *Nausithoë*. (*After L. H. Hyman, The Invertebrates, vol. 1, McGraw-Hill, 1940*)

an upper central disk and a lower coronal part (see **illus.**). The central disk is usually domelike, but in *Periphylla* it often narrows toward the top. The coronal part has gelatinous thickenings, or pedalia, situated on the radii running from the center of the umbrella to the tentacles and the sensory organs. The pedalia are separated from one another by clefts, running down the radii, that are halfway between each tentacle and rhopalium (a sensory organ). The mouth has four lips. There are four groups of gastral filaments within the central stomach. Abyssal forms of this group are dark brown or reddish purple. The littoral species of *Nausithoë* is reported to show alternation of generations, but the life history of abyssal forms remains unknown. *See* SCYPHOZOA. Tohru Uchida

Coronoidea

A small class of "arm"- and brachiole-bearing, stemmed echinoderms in the subphylum Crinozoa, based on five genera known from the Middle Ordovician to Late Silurian of Europe and North America. Coronoids have a crested theca or body with well-developed pentameral symmetry and plate arrangement very similar to that found in blastoids (see **illus.**). Skeletal plates include three basals (two large and one smaller), five radials extending up to form the crests and bearing central notches for the ambulacra, five small plates supporting the arms, and four regular deltoids plus two anal deltoids in the fifth position around the mouth. The mouth is central on top of the theca and has five ambulacral

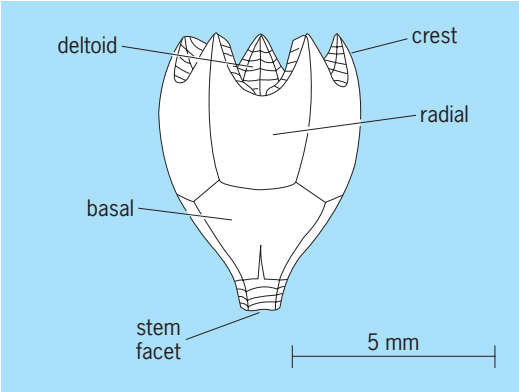

***Stephanocrinus gemmiformis*, Middle Silurian, Tennessee; side view of the theca.**

grooves radiating from it. A coiled arm is attached to a mounting plate at the end of each ambulacral groove; each arm has biserial plating with smaller biserial branches (=brachioles) on both sides. The projecting crests have internal canals that connected with the body cavity and apparently served as respiratory structures. A thin stem supported the theca above the sea floor, allowing coronoids to live as attached, low- to medium-level suspension feeders. Coronoids had previously been assigned to the blastoids, eocrinoids, or crinoids by different researchers, but they have been elevated in rank to a separate class. Coronoids appear to be most closely related to the Blastoidea and may have been the ancestors of this class. *See* BLASTOIDEA; ECHINODERMATA; EOCRINOIDEA. James Sprinkle

Bibliography. C. E. Brett et al., *J. Paleontol.*, 57:627–651, 1983.

Correspondence principle

A fundamental hypothesis according to which classical mechanics can be understood as a limiting case of quantum mechanics; or conversely, many characteristic features in quantum mechanics can be approximated on the basis of classical mechanics, provided classical mechanics is properly reinterpreted. This idea was first proposed by N. Bohr in the early 1920s as a set of rules for understanding the spectra of simple atoms and molecules.

Bohr's formulations. In 1913 Bohr used classical mechanics to calculate the energy levels of the electron in the hydrogen atom, but this result could not be understood in the conventional manner. Although the electron presumably orbits the proton like a planet around the Sun, Bohr had to assume that its electric charge radiates (or absorbs) light only when it makes a transition from one energy level to another. The frequency ν of the light is always the difference between the two energy levels divided by Planck's constant h.

Bohr's original quantization rules are readily applied to the simplest spectra, in particular to the electrons of hydrogenlike atoms and ions, as well as the vibrations and rotations of small molecules. The classical motions in these simple dynamical systems can be understood as composed of independent partial motions, each with its own degree of freedom, like the motion of the electron toward and away from the nucleus, and its motion around the nucleus. Each degree of freedom accumulates its own classical action-integral. This is the integral defined by Eq. (1), calculated along a closed orbit in terms of

$$I = \int p\,dq \tag{1}$$

the position q and momentum p. The classical energy is a function of these action integrals. Every action I is then related to the quantum number n by the Bohr-Sommerfeld rule, given by Eq. (2). A complete

$$I = nh \tag{2}$$

classification of the spectrum in quantum mechanics is obtained in terms of the quantum numbers. *See* ACTION; ATOMIC STRUCTURE AND SPECTRA; DEGREE OF FREEDOM (MECHANICS); MOLECULAR STRUCTURE AND SPECTRA.

The frequency ν of the classical motion for any particular degree of freedom is given by the partial derivative of the energy function with respect to the corresponding action. Bohr noticed that this classical result yields the correct quantum-theoretical result for the light frequency in a transition from one energy level to another, say from $n + 1$ down to n, provided the derivative is replaced by the difference in the energies. Moreover, precise information about the possibility of such transitions and their intensities is obtained by analyzing the related classical motion. This information becomes better as the quantum numbers involved become larger. The apparent inconsistencies in Bohr's quantum theory are thereby overcome by a set of rules that came to be called the correspondence principle.

This approach to the interpretation of spectra turned out to be very useful to the experimenters. Moreover, in 1925 W. Heisenberg's program of replacing all classical derivatives systematically by the corresponding differences led to the new quantum mechanics, as opposed to the old quantum theory. The complete and overwhelming success of the new quantum mechanics, particularly the equivalent approach of E. Schrödinger (wave mechanics), reduced the correspondence principle to a somewhat vague article of faith among physicists.

Use in statistical mechanics. The appeal to classical mechanics is still convenient for some rather crude estimates such as the total number $N(E)$ of levels below a given energy E. As in statistical mechanics, $N(E)$ depends on the volume of phase space, wherein every point gives the complete information on the state of the dynamical system. The phase space of a system with f degrees of freedom has $2f$ dimensions: f for the position in ordinary space, and f for the momentum. If E is sufficiently large, $N(E)$ approaches the volume of phase space with energy less than E divided by h^f. This estimate helps in finding the approximate shape of large atoms and large nuclei in

the Thomas-Fermi model; each particle experiences the average effect of all the others, and only the exclusion principle, not the wavelike nature of the particles, is invoked. *See* EXCLUSION PRINCIPLE; STATISTICAL MECHANICS.

Application to quantum chaos. The correspondence principle, however, has assumed a more profound significance. Experimental techniques in atomic, molecular, mesoscopic, and nuclear physics have improved dramatically. High-precision data for many thousands of energy levels are available where the traditional methods of quantum mechanics are not very useful or informative. However, the basic idea behind the correspondence principle must still be valid: Quantum mechanics must be understandable in terms of classical mechanics for the highly excited states, even in difficult cases like the three-body problem, where the classical motion cannot be resolved into independent partial motions. In these cases, two trajectories that started out at almost the same initial position and with almost the same initial momentum do not stay together in the long run, and their overall behavior seems unpredictable and chaotic.

The challenge of extracting quantum information from classical data is called the problem of quantum chaos. The structure of the trajectories in phase space is fractal; that is, the same patterns occur repeatedly as the scale is reduced. Quantum mechanics does not allow this kind of self-similarity because it describes atoms, objects that cannot be subdivided into ever smaller pieces, and electrons and nuclei are, therefore, represented by waves. Nevertheless, these waves follow the classical rays or trajectories very closely, and their amplitudes act like a bunch of classical particles. These "semiclassical" waves can be superimposed on one another, and thereby manifest all the interference phenomena that are typical of wave optics, and are the essence of quantum mechanics. They yield surprisingly good approximate numerical results as well as intuitive insight, where the standard methods yield only numerical results, often massive calculations. *See* FRACTALS; INTERFERENCE OF WAVES.

The wider application of Bohr's correspondence principle allows many basic but difficult problems to be seen in a new light. The spectra of the hydrogen atom in a strong magnetic field, of the helium atom when both electrons are excited, and of the donor impurity in a semiconductor, and the fluctuations in the electric conductivity for a mesoscopic device are all closely related to the structure of closed orbits in the corresponding classical problem. The statistical properties and the correlations among the energy levels, as well as the fluctuations in the scattering cross sections, for the more complicated nuclei are now understood in terms of the underlying chaos in their classical analogs. *See* CHAOS; MESOSCOPIC PHYSICS; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS. Martin C. Gutzwiller

Bibliography. M. Born, *The Mechanics of the Atom,* 1960; M. C. Gutzwiller, *Chaos in Classical and Quantum Mechanics,* 1991; P. Stehle, *Order, Chaos, Order: The Transition from Classical to Quantum Physics,* 1993; B. L. van der Waerden (ed.), *Sources of Quantum Mechanics,* 1967.

Corrosion

In broad terms, the interaction between a material and its environment that results in a degradation of the physical, mechanical, or even esthetic properties of that material. More specifically, corrosion is usually associated with the oxidation and/or dissolution of a metal, oxide, or semiconductor. For example, immersing a piece of iron in a moderately acidic solution is a typical example of corrosion; the vigorous reaction is the electrochemical oxidation of iron from the zero valence state to divalent ferrous ions (Fe^{+2}) dissolved in the water. Corrosion encompasses principles from diverse fields such as electrochemistry, metallurgy, physics, chemistry, and biology.

Economic impact. The cost of corrosion in the United States and other industrial countries has been estimated to be about 4% of the gross national product. This cost represents both the replacement of structures or components and the protection against corrosion. Such direct losses are relatively easy to assess in comparison to indirect losses, which include the cost of shutdown, loss of product or efficiency, contamination, and even loss of life. Another economic consideration is the source of corrosion-resistant metals. For example, chromium is considered to be of critical importance since it is responsible for the high-temperature oxidation resistance and aqueous corrosion resistance of a wide range of iron- and nickel-based alloys. Much of the chromium being used is considered irreplaceable. Since chromium ores occur primarily in South Africa and Russia, the Western industrialized nations often express concern regarding a perceived vulnerability in supply. *See* CHROMIUM.

Engineering systems for service environments require materials that are chemically stable or that can be made chemically stable. Moreover, modern technology continues to demand materials that can perform in increasingly hostile environments such as high-temperature coal gasification plants, deep sour-gas [natural gas containing hydrogen sulfide (H_2S)] wells, magnetohydrodynamic channels, and fusion reactors. Additionally, as engineering systems continue to evolve, they will use new or advanced materials. Thin-film electronic, magnetic, and optical devices will require special attention to corrosion issues. Likewise, it is necessary to deal with problems associated with traditional systems, such as the aging transportation infrastructure, underground tanks and pipelines, structures exposed to salt water, and conventional coal-burning power stations.

Thermodynamic stability of metals. Metals that are thermodynamically stable in their environment exhibit the best corrosion resistance, but only rarely are such metals of practical use. The stability of a metal in aqueous solutions can be assessed from data on the equilibrium potential that is established by an electrochemical reaction. The relationship between

the equilibrium potential of a reaction and the activities of each species is determined by the Nernst equation. By convention, all electrochemical reactions are written in terms of reduction, reaction (1).

$$Fe^{+2} + 2e^- \rightarrow Fe \qquad (1)$$

The Nernst equation for these reactions is given in general terms by Eq. (2), where E is the equilibrium

$$E = E^\circ + \frac{2.303RT}{nF} \log \frac{\prod[a_O]^q}{\prod[a_R]^p} \qquad (2)$$

potential, E° is the equilibrium potential under standard conditions, $\Pi[a_R]^p$ and $\Pi[a_O]^q$ are the products of the activities of the reduced and oxidized species, respectively, where p and q are the coefficients for each term in the reaction; n is the number of moles of electrons transferred; F is Faraday's constant; R is the gas constant; and T is temperature. The activities are related to concentration, and in Eq. (2) a_R is the activity of the metal ions (for example, Fe^{2+}). The standard condition is unit activity, which means pure metals or 1 mole/liter for dissolved ions.

The standard reduction potentials for various reactions are given in the **table**. Also included are the standard potentials for the hydrogen evolution and oxygen reduction reactions. The equilibrium potential for the hydrogen evolution reaction at pH 0 is arbitrarily designated as zero volts, and all other potentials are referenced to this value.

This table can be used to predict the thermodynamic tendency of a metal to corrode by comparing the equilibrium potential of the metal dissolution reaction to that of a cathodic reaction. For example, comparing gold dissolution and oxygen reduction would lead to the conclusion that gold is thermodynamically stable under standard conditions. Alternatively, the equilibrium potential for iron (Fe^{2+}/Fe) is more negative than the equilibrium potentials for both cathodic reactions (oxygen evolution and hydrogen ion reduction); iron will, therefore, dissolve to form ferrous ions in solution. Copper has an equilibrium potential more positive than that of the hydrogen reaction and will dissolve only with oxygen reduction under standard conditions.

The pH–potential diagram is an alternative way of representing thermodynamic data that also accounts for the important variable pH. For example, **Fig. 1** shows a potential–pH diagram for the iron–water system at 25°C (77°F). Such diagrams are also known as Pourbaix diagrams. The lines on the diagram are derived from the application of the Nernst equation to reactions involving coexisting phases, both solids and dissolved species. Therefore, areas bounded by these lines represent regions of potential and pH values where the various species predominate. For application to corrosion, the activity of dissolved ions is taken (arbitrarily) as 10^{-6} mole/liter, a value small enough to represent very small, or trivial, corrosion. Also superimposed on the diagram are lines, A and B, representing the limit of stability of water. Therefore, from a consideration of the dominant reactions expected for given values of potential and pH, one can infer whether corrosion, thermodynamic stability (immunity), or passivation is likely. For example, in Fig. 1 in the region of stability of ferric [iron (III)] oxide (Fe_2O_3), iron is protected from corrosion by a thin film of this oxide, a phenomenon known as passivation. In regions where the ferrous ions (Fe^{2+}) or other soluble species are stable, iron will corrode.

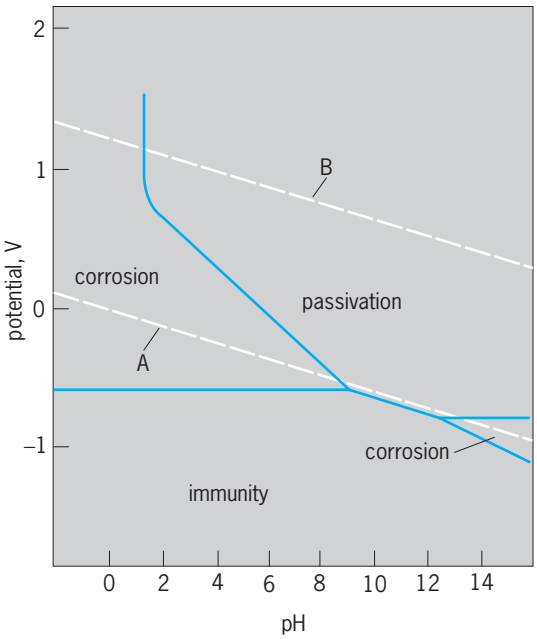

Fig. 1. Potential–pH diagram for the iron–water system: A = hydrogen equilibrium [see reaction (4)]; B = oxygen equilibrium [see reactions (5) and (6)]. ***(After M. Pourbaix, ed., Atlas of Electrochemical Equilibria in Aqueous Solutions, NACE, Houston, 1974)***

Standard reduction potentials* for various reactions

Metal	Electrode reaction	E°, V
Gold	$Au^{3+} + 3e^- \rightarrow Au$	1.50
Oxygen reduction	$O_2 + 2H_2O + 4e^- \rightarrow 4OH^-$	1.23
Copper	$Cu^{2+} + 2e^- \rightarrow Cu$	0.34
Hydrogen evolution	$2H^+ + 2e^- \rightarrow H_2$	0.00
Tin	$Sn^{2+} + 2e^- \rightarrow Sn$	−0.14
Nickel	$Ni^{2+} + 2e^- \rightarrow Ni$	−0.25
Iron	$Fe^{2+} + 2e^- \rightarrow Fe$	−0.44
Chromium	$Cr^{3+} + 3e^- \rightarrow Cr$	−0.74
Zinc	$Zn^{2+} + 2e^- \rightarrow Zn$	−0.76
Aluminum	$Al^{3+} + 3e^- \rightarrow Al$	−1.66

* $T = 25°C$ (77°F), $a = 1$ mol/liter.

Similar diagrams have been compiled for a wide variety of metal–solution systems. However, it is noted that these diagrams allow only thermodynamic predictions and cannot predict the kinetics (rate) of a particular reaction. *See* ACTIVITY (THERMODYNAMICS); CHEMICAL THERMODYNAMICS.

Reactions. Electrolytic corrosion consists of two partial electrochemical reactions: anodic (oxidation) and cathodic (reduction) [**Fig. 2**]. For thermodynamic considerations (as in the table), the cathodic

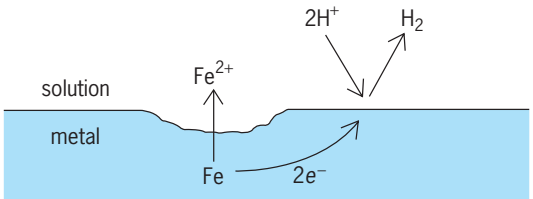

Fig. 2. Diagram of a corrosion cell showing the anodic and cathodic partial processes for a piece of iron in an acidic solution.

reaction is the reverse of the anodic reaction. However, in most aqueous corrosion situations, the anodic reaction is the dissolution of metal and the cathodic reaction is the reduction of a different species—for example, the reduction of dissolved oxygen or hydrogen ions. The loss of metal that is the usual manifestation of the corrosion process is a result of the anodic reaction (3), and the cathodic reaction could be reaction (4). Reaction (4) is often

$$Fe(S) \rightarrow Fe^{+2} + 2e^- \quad (3)$$

$$2H^+ + 2e^- \rightarrow H_2(g) \quad (4)$$

the dominant cathodic reaction in systems at low pH. The hydrogen-evolution reaction can itself cause corrosion-related problems, since atomic hydrogen (H) can enter the metal, causing embrittlement, a phenomenon that results in the degradation of the mechanical properties and can cause catastrophic failure. *See* EMBRITTLEMENT.

At steady state, the rates of each reaction will be such that the electrons (e^-) produced by the oxidation reaction will be consumed by the reduction reaction. The anodic reaction can occur uniformly over a metal surface or can be localized to a specific area. If the dissolved metal ion can react with the solution to form an insoluble compound, a buildup of corrosion products can accumulate at the anodic site. *See* OXIDATION-REDUCTION.

The second important cathodic reaction is the oxygen reduction, given by reactions (5) and (6). These

$$O_{2(dissolved)} + 4H^+ + 4e^- \rightarrow 2H_2O \quad (5)$$

$$O_{2(dissolved)} + 2H_2O + 4e^- \rightarrow 4OH^- \quad (6)$$

represent typical reactions in acidic and alkaline solutions, respectively. Equation (6) is usually dominant in solutions of neutral and alkaline pH. In order for this reaction to proceed, a supply of dissolved oxygen is necessary; hence, the rate of this reaction is sometimes limited by the transport of oxygen to the metal surface.

Equations (4), (5), and (6), in practice, can occur by a sequence of reaction steps. In addition, the reaction sequence can be dependent upon the metal surface, resulting in significantly different rates of the overall reaction on different metals. The cathodic reactions are important to corrosion processes since many methods of corrosion control depend on altering the cathodic process. Although the cathodic reactions can be related to corrosion processes, which are usually unwanted, they are essential for many applications such as energy storage and generation.

Rate of corrosion. Corrosion rates can be expressed in several different ways, reflecting the method of measurement. Loss of thickness per unit time [for example, mils (1/1000 in.) per year (MPY)], is commonly used in published data tables, but equivalent corrosion rates can be expressed as weight change per unit area–time (mg/cm^2-day) or as a current density (amperes per unit area). For many applications, 1–5 MPY is considered excellent corrosion resistance. General (uniform) corrosion rates can vary from centimeters per year [for example, iron in sulfuric acid (H_2SO_4)] to micrometers per year. Relatively large corrosion rates can be tolerated for some large structures, whereas for other structures small amounts of corrosion can result in catastrophic failure. For example, with the advent of technology for making extremely small devices, future generations of integrated circuits will contain components that are nanometers (10^{-9} m) in size, and even small amounts of corrosion could cause device failure. For medical devices implanted in the body, one meaningful measure of corrosion reflects the release rate of certain corrosion products into the body, such as milligrams of Ni^{2+} (from stainless steel) per unit area of device per day. The limitation on corrosion due to toxicity might be much more stringent than limitations based on structural damage.

Passivation. Many metals and alloys that exhibit a thermodynamic tendency to corrode actually have remarkable corrosion resistance in many environments due to the formation of a passive film. Passivation is the formation of a thin protective oxide on the surface of a metal. Passive films are usually less than 10 nm thick, but are highly stable and can prevent severe corrosion of the underlying metal. Stainless steels, such as Fe, 18% Cr, 8% Ni, form a highly stable, thin chromium oxide layer. This is the main reason for the widespread application of stainless steels in industry as corrosion-resistant alloys. Likewise, Al_2O_3 passive films protect aluminum alloys in solutions that are not extremely acidic or basic.

The electrochemical characteristics of a metal or alloy usually are measured by a combination of experimental techniques of which the current density–potential behavior is the most common. **Figure 3** shows a current density–potential diagram for AISI (American Iron and Steel Institute) 304 stainless steel (18–20% Cr, 8.0–10.5% Ni, <0.08% C) in 1 *N* H_2SO_4 solution at 25°C (77°F). The current density is often plotted on a logarithmic scale, as in this figure, due to the exponential relationship between current density and potential for many electrochemical reactions. In Fig. 3, it can be seen that the open-circuit potential, or corrosion potential, is about −0.38 V; and as the potential is increased, the current density increases rapidly to a maximum. At this point, a stable passive film is formed on the surface, and increasing the potential further results in a decrease in current density to a low value. At more positive

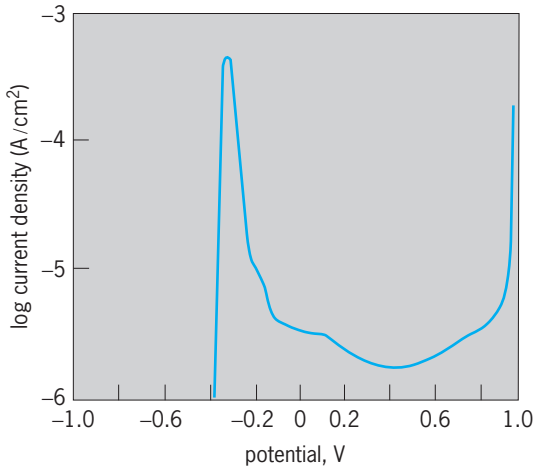

Fig. 3. Current–potential diagram for AISI 304 stainless steel in sulfuric acid (1 *N* H_2SO_4), *T* = 25°C (77°F), scan rate = 0.3 mV/s.

potentials, the current density increases again at about 1 V because of oxygen evolution from the surface. This plot is typical for alloys that exhibit passivation over a range of potential in a given solution.

The continuous stability of the passive film is critical for most corrosion-resistance applications. A passive film can become unstable in an environment in which it can dissolve chemically or in the presence of certain aggressive anions. Halide ions, and in particular chloride ions, are the most common species that can cause passive film breakdown, and consequently many types of corrosion problems are related to exposure to seawater or other chloride-containing environments. This type of corrosion attack usually is characterized by localized breakdown of the film. For many common engineering alloys, the sites of localized attack usually are related to inhomogeneities such as precipitates or secondary phases.

Types. In some situations, corrosion can occur only at localized regions on a metal surface. This type of corrosion is characterized by regions of locally severe corrosion, although the general loss of thickness can be relatively small. Common types of localized corrosion are pitting corrosion and crevice corrosion. Other types of localized corrosion include galvanic corrosion, stress corrosion cracking, and hydrogen embrittlement.

Pitting. This type of corrosion is usually associated with passive alloys, although this is not always the case. Pit initiation is usually related to the local breakdown of a passive film, and pitting is often related to the presence of halide ions in solution (**Fig. 4**). The local environment within a pit can become extremely acidic, resulting in a high rate of corrosion within this small region. For the common situation of a metal in salt water, the dissolution of metal within the pit results in the formation of metal chlorides. When sufficiently concentrated, these salts react with water (hydrolyze) to form hydrochloric acid (HCl) within the pit, as shown in reaction (7).

$$Fe^{+2} + 2Cl^- + 2H_2O \rightarrow Fe(OH)_2 + 2H^+ + 2Cl^- \qquad (7)$$

The decrease in pH within the pit caused by acid formation further increases the rate of metal dissolution, and such self-accelerating reactions are often called autocatalytic.

Crevice corrosion. This type of corrosion occurs in restricted or occluded regions, such as at a bolted joint, and are often associated with solutions that contain halide ions (**Fig. 5**). Crevice corrosion is initiated by a depletion of the dissolved oxygen in the restricted region. For this situation to occur, the dimension of the restricted region is typically less than 1 mm. As the supply of oxygen within the crevice is depleted, cathodic oxygen reduction from the region adjacent to the crevice takes over, allowing the anodic current on the interior metal surface to continue. The ensuing reactions within the crevice are the same as those described for pitting corrosion; that is, halide ions migrate to the crevice, where they are then hydrolyzed to form metal hydroxides and hydrochloric acid. In neutral salt solutions, the pH within a crevice can fall to a value as low as 2.

Galvanic coupling. Corrosion can also be accelerated in situations where two dissimilar alloys are in electrical contact with each other and with the same solution. This form of corrosion is known as galvanic corrosion. Because the two alloys have different corrosion potentials, a voltage difference arises which drives the anodic reaction on one alloy and the cathodic reaction on the other. The possibility of galvanic corrosion and the identification of the alloy that will become the anode (and hence will experience accelerated corrosion rates) can best be determined from a galvanic series, which is a table of corrosion potentials in a particular solution. The alloy with the more negative potential becomes the anode, while the alloy with the more positive potential sustains the cathodic reaction. For example, a copper water pipe connected to a steel water pipe constitutes a galvanic couple, with the water

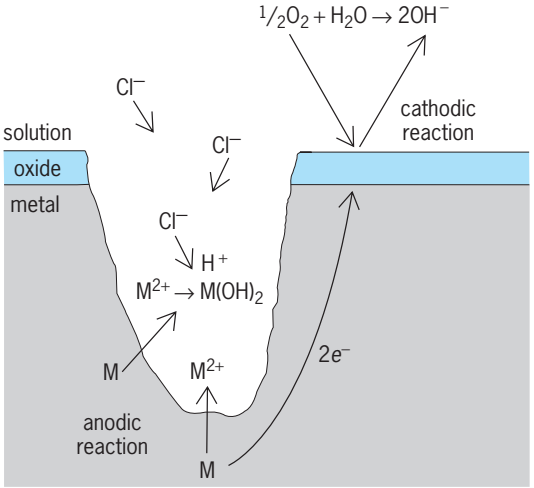

Fig. 4. Diagram of pitting corrosion. The diffusion of chloride ions into the pit to maintain electroneutrality is followed by hydrolysis of metal chlorides, resulting in the formation of hydrochloric acid (HCl).

inside acting as the corroding solution. The steel pipe (anode) will corrode rapidly at locations close to the copper/steel joint. A useful example of galvanic corrosion is the application of a zinc coating to steel (galvanized steel). Zinc has the more negative corrosion potential, so it selectively corrodes, protecting the steel (**Fig. 6**). Another example is the use of magnesium alloy anodes to protect steel ships and steel underground pipelines from corrosion.

Stress corrosion cracking. Stress corrosion cracking (SCC) and hydrogen embrittlement are corrosion-related phenomena associated with the presence of a tensile stress. Stress corrosion cracking results from a combination of stress and specific environmental conditions so that localized corrosion initiates cracks that propagate in the presence of the tensile (stretching) stress. Stress corrosion cracking is particularly dangerous because alloys that normally are quite ductile, and that are used at normally safe design stresses, can fail suddenly without warning by brittle fracture. There are many examples of the specificity of the combination of alloy/environment leading to stress corrosion cracking. For example, mild steels are susceptible to stress corrosion cracking in strongly basic environments (often called caustic cracking). Austenitic stainless steels are susceptible in the presence of chloride ions (Cl^-) and hydroxyl ions (OH^-). For example, in the medical device field stainless-steel implants used to mechanically support broken or diseased bones have been known to fail by stress corrosion cracking, because the environment is a warm, low-oxygen, high-chloride solution. Other alloys that are susceptible under specific conditions include certain brass, aluminum, and titanium alloys.

Hydrogen embrittlement is caused by the entry of hydrogen atoms into a metal or alloy, resulting in a loss of ductility and cracking promotion if the stress is sufficiently high. The source of the hydrogen is usually from corrosion (that is, cathodic hydrogen evolution) or from cathodic protection. Metals and alloys that are susceptible to hydrogen embrittlement include certain carbon steels, high-strength steels, nickel-based alloys, titanium alloys, and some aluminum alloys. *See* ALLOY; STAINLESS STEEL; STEEL.

Other types. Corrosion can be stimulated by biological organisms (such as microbially induced corrosion in soil and pitting induced corrosion beneath barnacles in seawater), where the local chemistry at the metal surface is altered by secretions from the organisms. Contaminants in the air also can accelerate corrosion rates. For example, sulfate particles or acidic droplets (acid rain) accelerate the corrosion of outdoor structures. Dust particles indoors have been shown to corrode unprotected microelectronic circuit boards. *See* ACID RAIN.

Control. In general, the resolution of a corrosion problem starts with the identification of the predominant type of corrosion and a determination of an acceptable design lifetime. Many corrosion problems can be solved simply by modifying the solution composition (pH control) or by selecting a more resistant alloy composition (stainless steel versus mild steel) to achieve a lower rate of corrosion. If large reductions in the rate of corrosion are needed, a useful approach is to separate the aqueous solution from the metal surface with a barrier, for example, using a passivated alloy (stainless steel), applying an overlayer (chrome plating), or adding an inhibitor to the solution. Inhibitors operate by controlling the metal/solution interface. Inorganic salts, such as chromates (CrO_4^{2-}), nitrites (NO_2^-), and phosphates (PO_4^{3-}), either form a barrierlike layer by reacting with the metal surface or stimulate film formation. Organic inhibitors are generally long-chain polar molecules with nitrogen, sulfur, oxygen, or

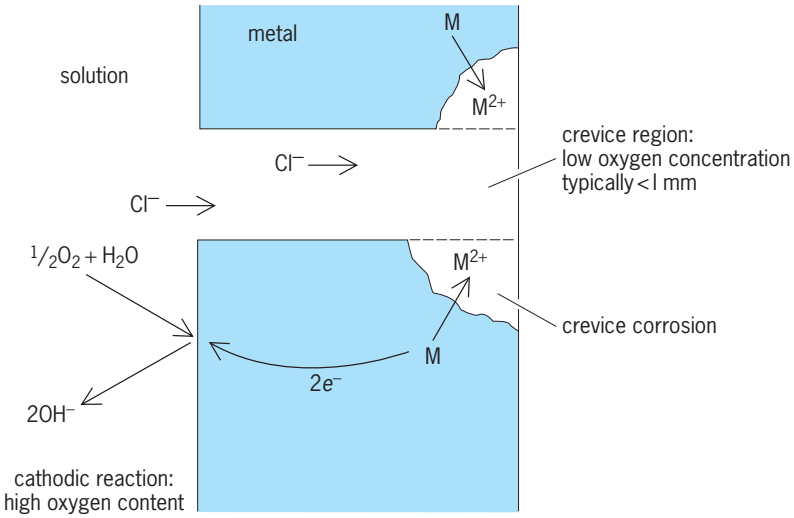

Fig. 5. Diagram of crevice corrosion, showing the mechanisms involved.

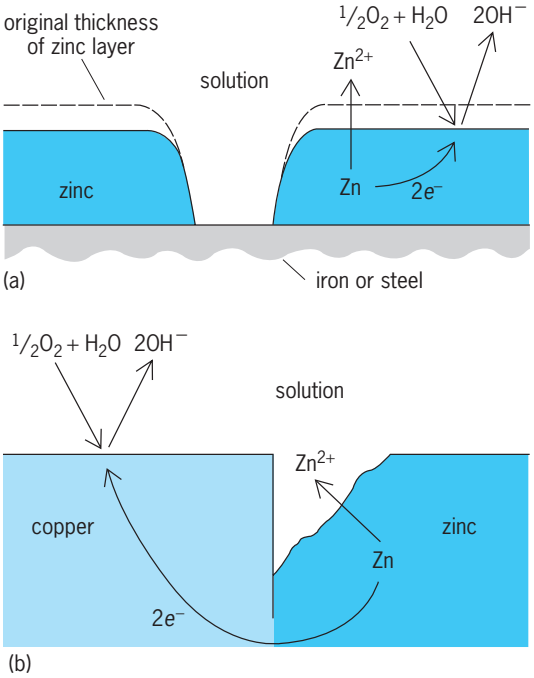

Fig. 6. Reactions of dissimilar metals in contact in the same solution. (*a*) Iron or steel protected from corrosion by a galvanic zinc coating. (*b*) Galvanic corrosion.

hydroxyl groups that strongly chemisorb to the metal surface and thereby separate metal and solution. For some specialized applications, such as protecting parts while they are being stored or shipped, adding a vapor-phase inhibitor to the package can be effective. *See* INHIBITOR (CHEMISTRY); METAL COATINGS.

A different approach is needed for pitting and crevice corrosion problems. Modifying the alloy's surface by cleaning or polishing or by adding a coating may help reduce or avoid the initiation of pits. Crevice corrosion usually can be avoided by redesigning to avoid crevices (for example, joining by welding instead of bolts) or by applying a watertight gasket to the joints. Galvanic corrosion problems can be avoided by choosing alloys with similar corrosion potentials or, if the design allows, by electrically isolating the parts. Designing to avoid stress concentrations (fair out corners) and removing residual stresses that arise from welding can reduce the risk of stress corrosion cracking.

In practice, the application of the above principles often is complicated by additional constraints. Materials selected might need to meet mechanical property specifications and cost restrictions. Clearly, it is impractical to gold-plate automobiles for corrosion resistance, but gold is routinely used for dental crowns, in part because of its corrosion resistance. In some applications, such as the chemical and food-processing industries, the composition of the corroding solution can be quite inflexible, or the solution cannot realistically be altered (for example, seawater). In the biomedical field, the cost of the material chosen for an implant is secondary to achieving the intended function with an appropriate service lifetime; however, the rate of release of potentially toxic corrosion products can be a stringent corrosion limitation. This is one of the newer, challenging areas of corrosion engineering.

Electrochemical techniques also can be used for corrosion control in some applications. For example, cathodic protection is the application of a cathodic current to a structure that prevents it from sustaining an anodic reaction. This can be achieved either by coupling the metal to be protected with a metal having a more negative open-circuit potential (sacrificial anode, such as magnesium) or by applying a cathodic current through an auxiliary electrode (impressed current).

For large structures, such as an underground pipeline, cathodic protection is usually applied in conjunction with a coating, since the current required to protect a large, uncoated structure would be uneconomical. Large marine structures, such as oil drilling platforms, are frequently cathodically protected.

Future. Traditional corrosion engineering problems have included systems such as steam generators, heat exchangers, reaction vessels, bridges, oil platforms, and motor vehicles. These areas will continue to be important in terms of materials selection and corrosion control. Corrosion and its control also will continue to be important in energy systems. The continuing development of alternative energy sources, such as geothermal energy and sour oil and gas wells, require processing in highly corrosive environments. In addition, there is a continuing search for more efficient and less costly materials for fuel cells, batteries, photoelectrolysis, and other emerging technologies.

The current generation of thin-film devices for integrated circuits and magnetic storage, using submicrometer scale ($<10^{-6}$ m) features, pose significant challenges because a small amount of corrosion will cause a device failure. Microprocessors are being used in increasingly aggressive environments, such as process plants and automobiles, where the importance of corrosion control in terms of device packaging will become critical. The current and continuing emphasis on cost reduction is being applied to many different kinds of devices. Corrosion control is increasingly difficult in these circumstances and will require new and ingenious solutions from corrosion scientists and engineers.

All engineering materials are susceptible to some form of environmental degradation. Users of such materials must be aware of the limits of chemical stability of these materials in service environments. *See* ELECTROCHEMISTRY.

David A. Shores; R. M. Latanision; Peter C. Searson

Bibliography. *ASM Handook*, vol. 13A: *Corrosion: Fundamentals, Testing and Protection*, ASM International, Materials Park, OH, 2003; S. L. Chawla and K. K. Gupta, *Materials Selection for Corrosion Control*, 1993; J. R. Davis, *Handbook of Materials for Medical Devices*, ASM International, Materials Park, OH, 2003; D. A. Jones, *Principles and Prevention of Corrosion*, 2d ed., Prentice Hall, Upper Saddle River, NJ, 1996; R. Winston Revie (ed.), *Uhlig's Corrosion Handbook*, 2d ed., Wiley, New York, 2000.

Cortex (plant)

The mass of primary tissue in roots and stems extending inward from the epidermis to the phloem. The cortex may consist of one or a combination of three major tissues: parenchyma, collenchyma, and sclerenchyma. In roots the cortex almost always consists of parenchyma, and is bounded, more or less distinctly, by the hypodermis (exodermis) on the periphery and by the endodermis on the inside.

Cortical parenchyma is composed of loosely arranged thin-walled living cells. Prominent intercellular spaces usually occur in this tissue. In stems the cells of the outer parenchyma may appear green due to the presence of chloroplasts in the cells (**Fig. 1**). This green tissue is sometimes called chlorenchyma, and it is probable that photosynthesis takes place in it.

In some species the cells of the outer cortex are modified in aerial stems by deposition of hemicellulose as an additional wall substance, especially in the corners or angles of the cells. This tissue is called collenchyma, and the thickening of the cell walls gives mechanical support to the shoot.

The cortex makes up a considerable proportion of the volume of the root, particularly in young roots, where it functions in the transport of water and ions from the epidermis to the vascular (xylem and phloem) tissues (**Fig. 2**). In older roots it functions primarily as a storage tissue. The cells of the root cortex generally are similar in appearance and may occur in radial files, but histochemical tests reveal a variety of cell types in root cortex.

In addition to being supportive and protective, the cortex functions in the synthesis and localization of many chemical substances; it is one of the most fundamental storage tissues in the plant. The kinds of cortical cells specialized with regard to storage and synthesis are numerous.

Because the living protoplasts of the cortex are so highly specialized, patterns and gradients of many substances occur within the cortex, including starch, tannins, glucosides, organic acids, crystals of many kinds, and alkaloids. Oil cavities, resin ducts, and laticifers (latex ducts) are also common in the mid-cortex of many plants. *See* SECRETORY STRUCTURES (PLANT).

The young cortex in roots is functional in the selection and passage of solutes and ions into the plant. Experiments with radioactive ions indicate a pileup of solutes in the innermost cells of the cortex next to the endodermis. Radioactive isotopes used in experimental work accumulate just outside the endodermis by a screening action of the endodermis and the innermost cells of the cortex. Cortical cells next to the endodermis apparently act as a selective barrier with the endodermis to the kinds of ions taken into the stele and vascular tissue.

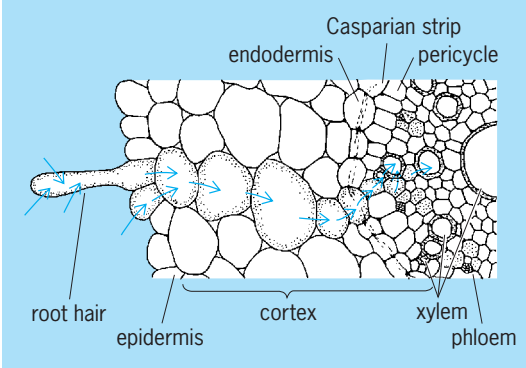

Fig. 2. Transverse section of wheat root, illustrating the kinds of cell that may be crossed by water and salts absorbed from the soil before they reach the tracheary elements of the xylem. The arrows indicate the direction of the movement through a selected series of cells. Among these, the living cells are partly stippled. The Casparian strip in the endodermis is shown as though exposed in surface views of the end walls. (*After K. Esau, Plant Anatomy, 2d ed., John Wiley and Sons, 1967*)

There appears to be complete freedom of movement of water, solutes, and ions through the cortex from the epidermis, but the Casparian strips limit freedom of movement through the endodermis into the stele. The piling up of solutes in the inner cortex limits freedom of movement and may function in the selection of the kinds of ions that pass from the cortex through the endodermis into the stele. Numerous substance gradients, as well as gradients of oxidizing potential across the cortex, may be significant in the uptake of solutes in the active growing portion of young roots. The passage and transport of solutes and ions from the cortex into the vascular system is most pronounced in the cortex just back of the growing tip of the root. In this zone of nutrient absorption, it is believed that passage of solutes into the vascular tissue is effected by an active metabolic transport system in the inner cells of the cortex, with an active system of selection in the endodermis.

The gradient of decreasing oxygen and increasing carbon dioxide across the cortex and into the vascular tissue is associated with a low level of metabolic activity in the cells around the vascular xylem vessels. Cells around the xylem with low metabolic activity would be expected to lose salts, since energy is required to hold salts, into the open vessels of the xylem. Diffusion of solutes back into the cortex is prevented by the impervious Casparian band between all endodermal cells. A unidirectional movement of solutes into the xylem vessels of the plant is a result of the action of the cortex and endodermis. *See* CELL WALLS (PLANT); COLLENCHYMA; ENDODERMIS; HYPODERMIS; PARENCHYMA; ROOT (BOTANY); SCLERENCHYMA; STEM.

D. S. Van Fleet

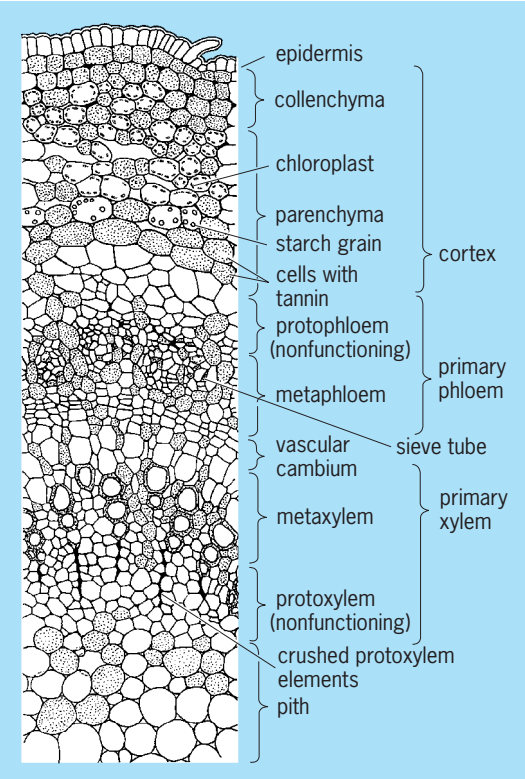

Fig. 1. Transverse section of the *Prunus* stem showing the cortex which is composed of collenchyma and parenchyma. The parenchyma has cells containing chloroplasts, starch grains, and tannin. (*After K. Esau, Plant Anatomy, 2d ed., John Wiley and Sons, 1967*)

Bibliography. J. D. Mauseth, *Plant Anatomy*, 1988.

Cosmic background radiation

The blackbody radiation left over from the origin of the universe, also known as the cosmic microwave background (CMB). This radiation was emitted before any astronomical objects such as stars, galaxies, or quasars existed, and is now observed to be like the thermal emission from a black object with a temperature of 2.725 K (degrees above absolute zero, or −454.8°F). Blackbody radiation is emitted by an isothermal object that absorbs all incident radiation. Its spectrum, or amount of power at different frequencies, depends only on its temperature and is given by the Planck function. The energy density of this radiation, which peaks at wavelengths near 1 mm, is larger than that of any other cosmic radiation field. Its existence was predicted in 1947 and 1948 by George Gamow, Ralph Alpher, and Robert Herman as part of their model for the formation of the elements during a hot dense phase in the early history of the universe. *See* HEAT RADIATION.

The properties of the cosmic microwave background rule out almost all cosmological models except for these hot big bang models. Observations of the spectrum of the cosmic microwave background show that less than 60 parts per million of the energy in the cosmic microwave background was generated later than 2 months after the creation of the universe and thus indicate that the universe started in a very hot and dense state. The anisotropy of the cosmic microwave background—the difference in the temperature of the blackbody radiation from point to point on the sky—reveals the structure of the universe at the time of recombination, 380,000 years after the big bang, when the free electrons and protons of the cosmic plasma combined into transparent hydrogen and helium gas. After recombination, the photons could freely travel very long distances, so that the small inhomogeneities that existed 380,000 years after the big bang became the small anisotropies that are now observed. *See* BIG BANG THEORY.

History of observations. The model for elememt formation developed by Gamow, Alpher, and Herman encountered difficulties accounting for the creation of elements heavier than lithium. These difficulties led to a general neglect of the model's prediction of the cosmic microwave background, even though the development of radar during World War II and the invention of a sensitive radiometer by Robert Dicke provided the tools necessary for the direct detection of the cosmic microwave background, and Dicke had placed a limit of less than 20 K on the temperature of the sky. The first indirect evidence for the cosmic microwave background was the rotational excitation of the interstellar cyanogen (CN) radical seen in 1939 by Andrew McKellar, but this observation was not connected to Gamow's prediction, except in a review by Fred Hoyle of a book by Gamow. Hoyle, who was co-inventor of the steady-state model of the universe, which was then a rival of the big bang model, felt that the 2.3 K measured by McKellar was too low for the big bang model, but Alpher and Herman had published detailed calculations allowing a current temperature of 1 to 5 K. Unfortunately, neither the big bang proponents nor the steady-state proponents followed up this discrepancy. Thus, the effective discovery of the cosmic microwave background was made by Arno Penzias and Robert Wilson in 1965, as they systematically tried to identify all the sources of noise in a very sensitive antenna used by Bell Labs for early communication satellite experiments. In every direction of sky they saw a nearly constant level of excess noise (**Fig 1***a*), and neither instrumental noise nor atmospheric emission could account for this emission. Dicke, James Peebles, Peter Roll, and David Wilkinson were building a radiometer to look for the cosmic microwave background when they heard of the Penzias and Wilson discovery.

Searches were carried out for the anisotropy of the cosmic microwave background, that is, for temperature differences in the cosmic microwave background in different directions on the sky. These searches led to a tentative detection of a dipole pattern by E. K. Conklin in 1969 which was confirmed

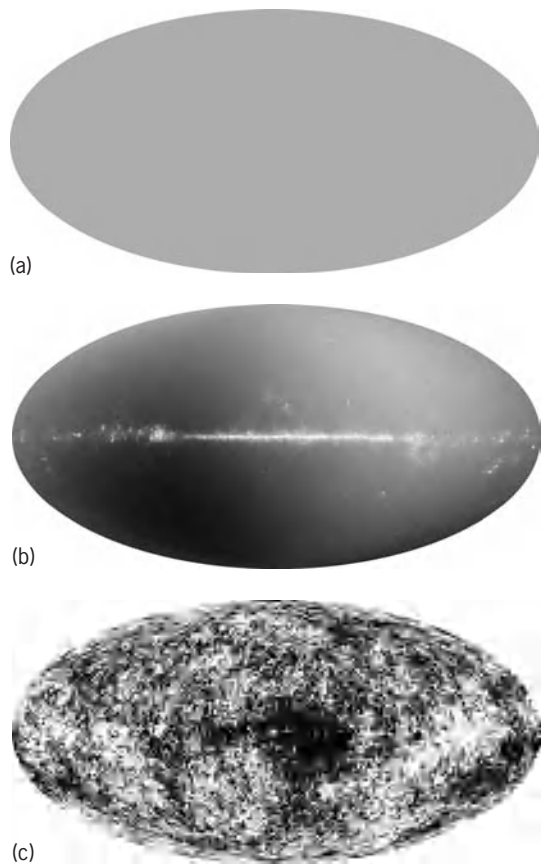

Fig. 1. Plots of the cosmic microwave background (CMB) temperature in a Molleweide projection in galactic coordinates. (*a*) CMB temperature from the *Cosmic Background Explorer* (*COBE*) Differential Microwave Radiometer (DMR), plotted without contrast enhancement. (*b*) Contrast enhanced by a factor of 400 to show the dipole anisotropy and emission from the Milky Way. (*c*) Dipole-subtracted linear combination of different frequency maps chosen to cancel the Milky Way emission, showing the CMB temperature fluctuations with contrast enhanced by a factor of 20,000.

by Paul Henry (a graduate student working with Wilkinson) in 1971, Brian Corey and Wilkinson in 1976, and George Smoot and his colleagues in 1977. A dipole pattern has two poles, with one side of the sky hotter than average, peaking at a hot pole, and the other side cooler than average, giving a cold pole. This dipole pattern is a measure of the motion of the solar system at 368 km/s (823,000 mi/h) relative to the observable universe (Fig 1*b*). After the dipole pattern was subtracted out of the observed temperature map, the anisotropy was less than the sensitivity of these early experiments, which was a few hundred parts per million.

The *Cosmic Background Explorer* (*COBE*) project, proposed in 1974 but not launched by the National Aeronautics and Space Administration (NASA) until 1989, carried an instrument to search for deviations of the cosmic microwave background spectrum from a blackbody, the Far-InfraRed Absolute Spectrophotometer (FIRAS), and an instrument to search for anisotropy, the Differential Microwave Radiometers (DMR). The FIRAS instrument limited distortions of the cosmic microwave background spectrum to less than 50 parts per million root-mean-square. The detection of nondipole anisotropy by the *COBE* DMR at a level of 11 parts per million was described as the "discovery of the century, if not of all time" by Stephen Hawking in 1992.

A second satellite to study the cosmic microwave background was launched by NASA in 2001. This satellite project was led Charles Bennett and Wilkinson, and after Wilkinson's death in 2002, the satellite was renamed the *Wilkinson Microwave Anisotropy Probe* in 2003. *Planck*, a third satellite to study the cosmic microwave background, is scheduled to be launched by the European Space Agency after 2007. *See* WILKINSON MICROWAVE ANISOTROPY PROBE.

Spectrum of CMB. The best measurement of the spectrum of the cosmic microwave background was made by the FIRAS instrument on *COBE*. This spectrometer compared the sky with a very good blackbody on the spacecraft and thus made a very sensitive measurement of deviations from a Planck function. The measured cosmic microwave background spectrum agrees with the spectrum of a blackbody to within 50 parts per million, and the temperature of the blackbody that best matches the sky is measured to be 2.725±0.002 K. **Figure 2** shows the spectrum of the cosmic microwave background plotted such that a blackbody is a horizontal line. The cosmic microwave background is remarkably close to a blackbody.

The expansion of the universe produces a redshift that preserves the blackbody character of the radiation. When the universe was 0.1% of its current size, the photons of the blackbody radiation field had 1000 times more energy than they do now. Since this time the wavelengths of the photons have increased by a factor of 1000, while the size of the universe has also increased by a factor of 1000. This ratio of the final wavelength λ_{obs} to the initial wavelength λ_{em} is used to define z, the redshift, via the formula

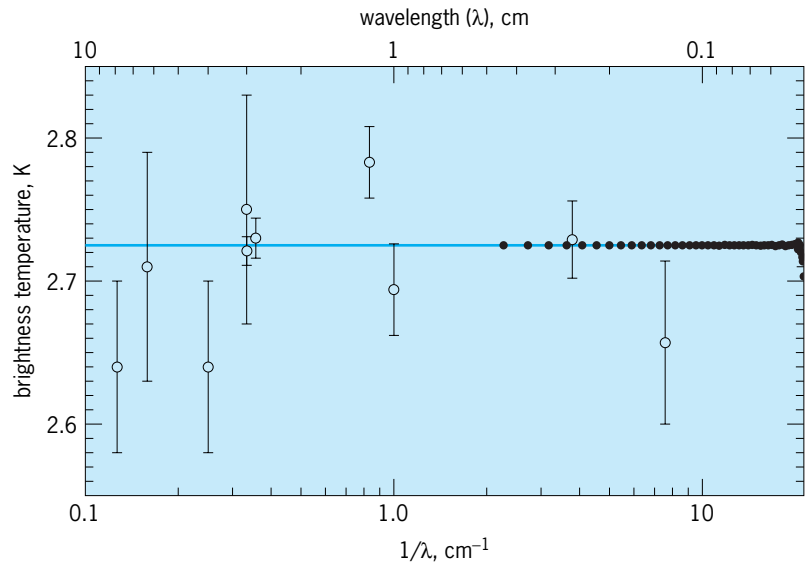

Fig. 2. Brightness temperature of the cosmic microwave background as a function of wavelength λ. For a perfect blackbody, the brightness temperature is independent of wavelength. The horizontal line is a blackbody at a temperature of 2.725 K. The closed points are the very precise Far-InfraRed Absolute Spectrophotometer (FIRAS) data.

$1 + z = \lambda_{obs}/\lambda_{em}$. Thus, when the universe was 0.1% of its current size, the color temperature deduced from the photon energy was 1000 times higher than T_0=2.725 K. However, the number density of the photons was 10^9 times higher than the current number density, and the energy density was 10^{12} times higher than the current energy density. This is exactly the energy density proportional to T^4 law seen for blackbodies. Thus, once a blackbody radiation field is produced, the expansion of the universe acts in a way that preserves the blackbody character of the radiation, with only the temperature changing. *See* REDSHIFT.

Isotropy of CMB. The cosmic microwave background is completely isotropic to about 1 part per thousand and, except for the velocity of the solar system, would be isotropic to better than 40 parts per million. Figure 1*a* shows the temperature of the sky without any contrast enhancement, and no structure can be seen. This is quite a remarkable property that is best explained by the inflationary scenario which postulates an immense expansion of the universe during the first picosecond after the big bang. *See* INFLATIONARY UNIVERSE COSMOLOGY.

Anisotropy of CMB. When the contrast of the temperature map is increased by a factor of 400, as shown in Fig. 1*b*, a systematic variation across the sky can be seen. This figure uses an equal-area Molleweide projection to show the entire sky in an ellipse with a 2:1 aspect ratio, and the equator of the map is the galactic plane, with the center of the Milky Way Galaxy occupying the center of the map. The "dipole" pattern with a hot pole in the upper right and a cold pole in the lower left can be explained by a Doppler shift between the frame of reference of the solar system and a frame of reference at rest relative to the observable universe. *See* DOPPLER EFFECT; MAP PROJECTIONS.

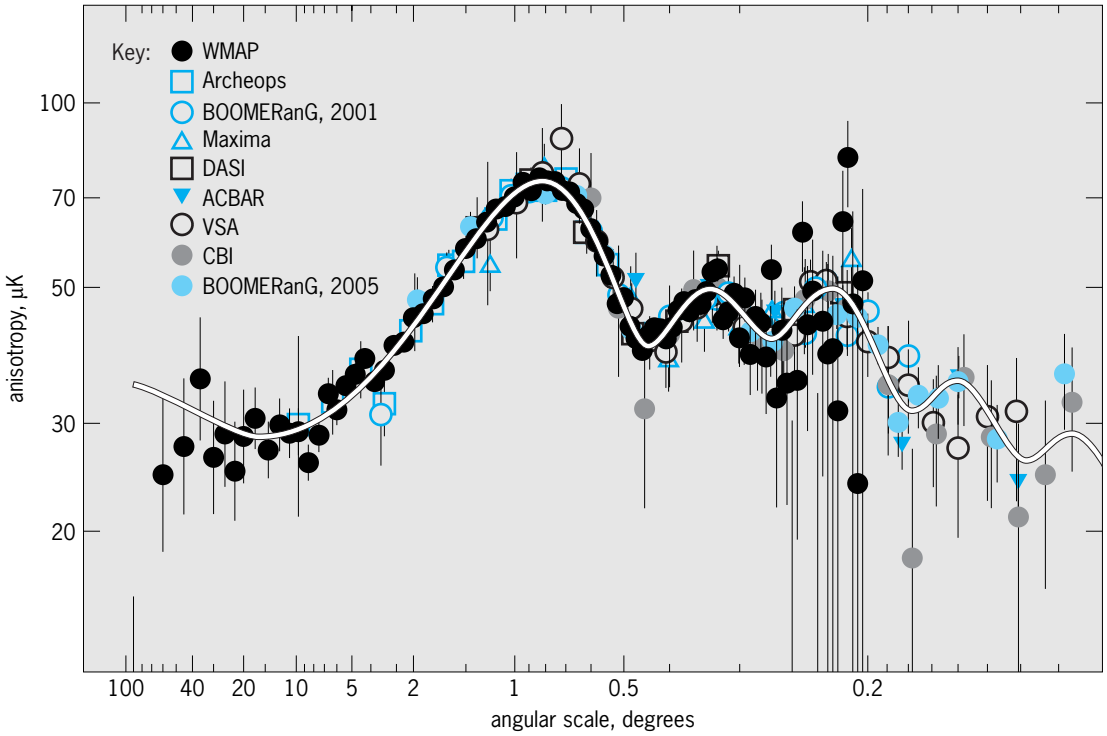

Fig. 3. Angular power spectrum of the cosmic microwave background as measured by the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) and other experiments, including the balloon-borne Archeops, BOOMERanG (data published in 2001 and 2005), and Maxima experiments, the ground-based Very Small Array (VSA) and Cosmic Background Imager (CBI), and DASI and ACBAR at the South Pole. The curve shows the predicted angular power spectrum for a model of the universe with Euclidean (flat) geometry, dominated by vacuum energy (dark energy).

Since the pattern produced on the sky by a Doppler shift can be predicted, the dipole anisotropy can be removed from the map, giving a no-dipole map. After this, the largest signal that remains is due to the radio and millimeter wave emission from the Milky Way. This emission is much stronger at centimeter wavelengths than at millimeter wavelengths, so if a fraction of a long-wavelength map is subtracted from a short-wavelength map, almost all of the emission from the Milky Way can be canceled while still leaving a strong cosmic microwave background anisotropy signal. Figure 1*c* shows such a no-dipole, no-Galaxy map based on data from the *Wilkinson Microwave Anisotropy Probe*.

The dominant cause of the temperature fluctuations for large angular scales is the gravitational redshift. Dense regions have negative gravitational potential wells, and the photons of the cosmic microwave background lose energy as they leave the potential wells. This effect is larger than the normal density effect wherein gases become hotter when compressed. However, at small angular scales the situation changes, in part because the density fluctuations have time to change between the big bang and recombination. The photons and the baryons plus electrons (ordinary matter) are tightly coupled before recombination, so the density fluctuations travel as acoustic waves at the sound speed in the plasma. Because there are billions of photons for every baryon, the radiation pressure from the photons is large compared with the energy density from the baryon rest mass, so the speed of sound c_s is very high, and is comparable to the speed of light c, being given approximately by the formula $c_s \sim c/\sqrt{3}$=170,000 km/s (380,000,000 mi/h). If the wavelength of a fluctuation is set so that it goes through one-half cycle of oscillation before recombination, the effects of the photon density and the gravitational redshift are in phase at recombination instead of partially canceling. This wave has a frequency of 42 femtohertz (4.2×10^{-14} Hz, more than 52 octaves below middle C) and a wavelength of 400,000 light-years. It shows up on the sky as a 1.6° angle with peaks separated by 0.8° from adjacent valleys. Thus, there will be more anisotropy in bumps about 0.8° in size compared to other angular scales. This phenomenon is evident in **Fig. 3**, a plot of the amount of anisotropy versus angular scale, or the angular power spectrum, which has a peak at an angular scale of 0.8°.

Conclusions. Observations of the cosmic microwave background provide a wealth of information about the universe. The cosmic microwave background spectrum gives a picture of the universe from 2 months after the big bang. (Prior to 2 months any spectral distortions were eliminated, while after 2 months the distortions were modified by electron scattering but still survived.) The anisotropy of the cosmic microwave background shows the state of the universe 380,000 years after the big bang, but that state was established during the inflation epoch less than 1 picosecond (10^{-12} s) after the big bang.

From the heights of the peaks and the depths of the troughs in the angular power spectrum (Fig. 3), the density of baryons (protons and neutrons in atomic nuclei) and the density of cold dark matter in the universe have been found: 0.42 ± 0.02 yoctogram (1 yg $= 10^{-24}$ g) per cubic meter for baryons and 2.1 ± 0.2 yg/m^3 for dark matter. In combination with other data, the cosmic microwave background results also show that the universe has Euclidean (Flat) geometry, that 75% of the energy in the universe is vacuum energy (dark energy), and that the universe is 13.7×10^9 years old. *See* COSMOLOGY; DARK ENERGY; DARK MATTER; UNIVERSE. Edward Wright

Bibliography. M. D. Lemonick, *Echo of the Big Bang*, 2003; J. C. Mather and J. Boslough, *The Very First Light*, 1996; B. Ryden, *Introduction to Cosmology*, 2002.

Cosmic rays

Electrons and the nuclei of atoms—largely hydrogen—that impinge upon Earth from all directions of space with nearly the speed of light. Before they enter the atmosphere they are typically referred to as primary cosmic rays, to distinguish them from the particles generated by their interaction with the terrestrial atmosphere. Secondary cosmic rays, comprising a large variety of species of charged and neutral particles, cascade down through the atmosphere all the way to the ground and below. Study of cosmic rays at high energy now is often referred to as particle astrophysics.

Cosmic rays are studied for a variety of reasons, not the least of which is a general curiosity over the process by which nature can produce such energetic nuclei. Apart from this, primary cosmic rays provide the only direct sample of matter from far outside the solar system. Measurement of their composition can aid in understanding which properties of the matter making up the solar system are typical of the Milky Way Galaxy as a whole and which may be so atypical as to yield specific clues to the origin of the solar system. Cosmic rays are electrically charged; hence they are deflected by the magnetic fields which are thought to exist throughout the Milky Way Galaxy, and may be used as probes to determine the nature of these fields far from Earth. Outside the solar system the energy contained in the cosmic rays is comparable to that of the magnetic field, so the cosmic rays probably play a major role in determining the structure of the field. Collisions between cosmic rays and the nuclei of the atoms in the tenuous gas which permeates the Milky Way Galaxy change the cosmic-ray composition in a measurable way and produce gamma rays which can be detected at Earth, giving information on the distribution of this gas. *See* GAMMA-RAY ASTRONOMY.

This modern understanding of cosmic rays evolved through a process of discovery which at many times produced seemingly contradictory results, the ultimate resolution of which led to fundamental discoveries in other fields of physics, most notably high-energy particle physics. At the turn of the century several different types of radiation were being studied, and the different properties of each were being determined with precision. One result of many precise experiments was that an unknown source of radiation existed with properties that were difficult to characterize. In 1912 Viktor Hess made a definitive series of balloon flights which showed that this background radiation increased with altitude in a dramatic fashion. Far more penetrating then any other known at that time, this radiation had many other unusual properties and became known as cosmic radiation, because it clearly did not originate in the Earth or from any known properties of the atmosphere.

Unlike the properties of alpha-, beta-, gamma-, and x-radiation, the properties of cosmic radiation are not of any one type of particle, but are due to the interactions of a whole series of unstable particles, none of which was known at that time. The initial identification of the positron, the muon, the π meson or pion, and certain of the K mesons and hyperons were made from studies of cosmic rays.

Thus the term cosmic ray does not refer to a particular type of energetic particle, but to energetic particles being considered in their astrophysical context. The effects of cosmic rays on living cells are discussed in a number of other articles: for example, *See* ELEMENTARY PARTICLE; RADIATION INJURY (BIOLOGY).

Cosmic-ray detection. Cosmic rays are usually detected by instruments which classify each incident particle as to type, energy, and in some cases time and direction of arrival. A convenient unit for measuring cosmic-ray energy is the electronvolt, which is the energy gained by a unit charge (such as an electron) accelerating freely across a potential of 1 V. One electronvolt equals about 1.6×10^{-19} joule. For nuclei it is usual to express the energy in terms of electronvolts per nucleon, since the relative abundances of the different elements are nearly constant as a function of this variable. Two nuclei with the same energy per nucleon have the same velocity.

Flux. The intensity of cosmic radiation is generally expressed as a flux by dividing the average number seen per second by the effective size or "geometry factor" of the measuring instrument. Calculation of the geometry factor requires knowledge of both the sensitive area (in square centimeters) and the angular acceptance (in steradians) of the detector, as the arrival directions of the cosmic rays are randomly distributed to within 1% in most cases. A flat detector of any shape but with area of 1 cm^2 has a geometry factor of π cm$^2 \cdot$ sr if it is sensitive to cosmic rays entering from one side only. The total flux of cosmic rays in the vicinity of the Earth but outside the atmosphere is about 0.3 nucleus/(cm$^2 \cdot$ s $\cdot$ sr) [2 nuclei/(in.$^2 \cdot$ s $\cdot$ sr)]. Thus a quarter dollar, with a surface area of 4.5 cm^2 (0.7 in.^{2}), lying flat on the surface of the Moon will be struck by $0.3 \times 4.5 \times 3.14 = 4.2$ cosmic rays per second.

Energy spectrum. The flux of cosmic rays varies as a function of energy. This function, called an energy

spectrum, may refer to all cosmic rays or to only a selected element or group of elements. Since cosmic rays are continuously distributed in energy, it is meaningless to attempt to specify the flux at any exact energy. Normally an integral spectrum is used, in which the function gives the total flux of particles with energy greater than the specified energy [in particles/($cm^2 \cdot s \cdot sr$)], or a differential spectrum, in which the function provides the flux of particles in some energy interval (typically 1 MeV/nucleon wide) centered on the specified energy, in particles/[$cm^2 \cdot s \cdot sr \cdot$ (MeV/nucleon)]. The basic approach of cosmic-ray research is to measure the spectra of the different components of cosmic radiation and to deduce from them and other observations the nature of the cosmic-ray sources and the details of where the particles travel on their way to Earth and what they encounter on their journey.

Types of detectors. All cosmic-ray detectors are sensitive to moving electrical charges. Neutral particles (neutrons, gamma rays, and neutrinos) are studied by observing charged particles produced in the collision of the neutral primary with some type of target. At low energies the ionization of the matter through which they pass is the principal means of detection. Such detectors include cloud chambers, ion chambers, spark chambers, Geiger counters, proportional counters, scintillation counters, solid-state detectors, photographic emulsions, and chemical etching of certain mineral crystals or plastics in which ionization damage is revealed. The amount of ionization produced by a particle is given by the square of its charge multiplied by a universal function of its velocity, the Bethe-Bloch relation. A single measurement of the ionization produced by a particle is therefore usually not sufficient both to identify the particle and to determine its energy. However, since the ionization itself represents a significant energy loss to a low-energy particle, it is possible to design systems of detectors which trace the rate at which the particle slows down and thus to obtain unique identification and energy measurement. *See* GAMMA-RAY DETECTORS; GEIGER-MÜLLER COUNTER; IONIZATION CHAMBER; JUNCTION DETECTOR; PARTICLE TRACK ETCHING; PHOTOGRAPHIC MATERIALS; SCINTILLATION COUNTER.

At energies above about 500 MeV/nucleon, almost all cosmic rays will suffer a catastrophic nuclear interaction before they slow appreciably. Some measurements are made using massive calorimeters which are designed to trap all of the energy from the cascade of particles which results from such an interaction. More commonly an ionization measurement is combined with measurement of a physical effect that varies in a different way with mass, charge, and energy. Cerenkov detectors and the deflection of the particles in the field of large superconducting magnets (or the magnetic field of the Earth itself) provide the best means of studying energies up to a few hundred gigaelectronvolts per nucleon. Detectors of x-ray transition radiation are useful for measuring composition at energies up to a few thousand GeV per nucleon. Transition radiation detectors are also used to study electrons having energies of 10–200 GeV which, because of their lower rest mass, are already much more relativistic than protons of the same energies. *See* CERENKOV RADIATION; SUPERCONDUCTING DEVICES; TRANSITION RADIATION DETECTORS.

Above about 10^{14} eV, direct detection of individual particles is no longer practical, simply because they are so rare. Such particles are studied by observing the large showers of secondaries they produce in Earth's atmosphere. These showers are detected either by counting the particles which survive to strike ground-level detectors or by looking at the flashes of light the showers produce in the atmosphere with special telescopes and photomultiplier tubes. It is not possible to directly determine what kind of particle produces any given shower. Because of the extreme energies involved, which can be measured with fair accuracy and have been seen as high as 10^{20} eV (16 J), most of the collision products travel in the same direction as the primary and at essentially the speed of light. This center of intense activity has typical dimensions of only a few tens of meters, allowing it to be tracked (with sensitive instruments) like a miniature meteor across the sky before it hits the Earth at a well-defined location. In addition to allowing determination of the direction from which each particle came, the development of many such showers through the atmosphere may be studied statistically to gain an idea of whether the primaries are protons or heavier nuclei. The main idea behind these studies is that a heavy nucleus, in which the energy is initially shared among several neutrons and protons, will cause a shower that starts higher in the atmosphere and develops more regularly than a shower which has the same total energy but is caused by a single proton. *See* PARTICLE DETECTOR; PHOTOMULTIPLIER.

Atmospheric cosmic rays. The primary cosmic-ray particles coming into the top of the terrestrial atmosphere make inelastic collisions with nuclei in the atmosphere. The collision cross section is essentially the geometrical cross section of the nucleus, of the order of 10^{-26} cm^2 (10^{-27} $in.^2$). The mean free path for primary penetration into the atmosphere is given in **Table 1**. (Division by the atmospheric density in g/cm^3 gives the value of the mean free path in centimeters.)

TABLE 1. Mean free paths for primary cosmic rays in the atmosphere

Charge of primary nucleus	Mean free path in air, g/cm^2*
$Z = 1$	60
$Z = 2$	44
$3 \le Z \le 5$	32
$6 \le Z \le 9$	27
$10 \le Z \le 29$	21

*$1\ g/cm^2 = 1.42 \times 10^{-2}\ lb/in.^2$

When a high-energy nucleus collides with the nucleus of an air atom, a number of things usually occur. Rapid deceleration of the incoming nucleus leads to production of pions with positive, negative, or neutral charge; this meson production is closely analogous to the generation of x-rays, or bremsstrahlung, produced when a fast electron is deflected by impact with the atoms in a metal target. The mesons, like the bremsstrahlung, come off from the impact in a narrow cone in the forward direction. Anywhere from 0 to 30 or more pions may be produced, depending upon the energy of the incident nucleus. The ratio of neutral to charged pions is about 0.75. A few protons and neutrons (in about equal proportions) may be ejected with energies up to a few GeV. They are called knock-on protons and neutrons. *See* BREMSSTRAHLUNG; MESON; NUCLEAR REACTION.

A nucleus struck by a proton or neutron with energy greater than approximately 300 MeV may have its internal forces momentarily disrupted so that some of its nucleons are free to leave with their original nuclear kinetic energies of about 10 MeV. The nucleons freed in this fashion appear as protons, deuterons, tritons, alpha particles, and even somewhat heavier clumps, radiating outward from the struck nucleus. In photographic emulsions the result is a number of short prongs radiating from the point of collision, and for this reason it is called a nuclear star.

All these protons, neutrons, and pions generated by collision of the primary cosmic-ray nuclei with the nuclei of air atoms are the first stage in the development of the secondary cosmic-ray particles observed inside the atmosphere. Since several secondary particles are produced by each collision, the total number of energetic particles of cosmic-ray origin will at first increase with depth, even while the primary density is decreasing. Since electric charge must be conserved and the primaries are positively charged, positive particles outnumber negative particles in the secondary radiation by a factor of about 1.2. This factor is called the positive excess.

Electromagnetic cascade. Uncharged π^0 mesons decay into two gamma rays with a lifetime of about 9×10^{-17} s. The decay is so rapid that π^0 mesons are not directly observed among the secondary particles in the atmosphere. The two gamma rays, which together have the rest energy of the π^0, about 140 MeV, plus the π^0 kinetic energy, each produce a positron-electron pair. Upon passing sufficiently close to the nucleus of an air atom deeper in the atmosphere, the electrons and positrons convert their energy into bremsstrahlung. The bremsstrahlung in turn creates new positron-electron pairs, and so on. This cascade process continues until the energy of the initial π^0 has been dispersed into a shower of positrons, electrons, and photons with insufficient individual energies ($\leq$1 MeV) to continue the pair production. The shower, then being unable to reproduce its numbers, is dissipated by ionization of the air atoms. The electrons and photons of such showers are referred to as the soft component of the atmospheric (secondary) cosmic rays, reaching a maximum intensity at an atmospheric depth of 150–200 g/cm^2 and then declining by a factor of about 10^2 down to sea level. *See* ELECTRON-POSITRON PAIR PRODUCTION.

Muons. The $\pi^\pm$ mesons produced by the primary collisions have a lifetime about 2.6×10^{-8} s before they decay into muons: $\pi^\pm \rightarrow \mu^\pm +$ neutrino. With a lifetime of this order a $\pi^\pm$ possessing enough energy (greater than 10 GeV) to experience significant relativistic time dilatation may exist long enough to interact with the nuclei of the air atoms. The cross section for $\pi^\pm$ nuclear interactions is approximately the geometrical cross section of the nucleus, and the result of such an interaction is essentially the same as for the primary cosmic-ray protons. Most low-energy $\pi^\pm$ decay into muons before they have time to undergo nuclear interactions.

Except at very high energy (above 500 GeV), muons interact relatively weakly with nuclei, and are too massive (207 electron masses) to produce bremsstrahlung. They lose energy mainly by the comparatively feeble process of ionizing air atoms as they progress downward through the atmosphere. Because of this ability to penetrate matter, they are called the hard component. At rest their lifetime is 2×10^{-6} s before they decay into an electron or positron and two neutrinos, but with the relativistic time dilatation of their high energy, 5% of the muons reach the ground. Their interaction with matter is so weak that they penetrate deep into the ground, where they are the only charged particles of cosmic-ray origin to be found. At a depth equivalent of 300 m (990 ft) of water the muon intensity has decreased from that at ground level only by a factor of 20; at 1400 m (4620 ft) it has decreased by a factor of 10^3.

Atmospheric neutrinos. In the late 1990s, detectors became available with sufficient sensitivity to exploit atmospheric neutrinos. Neutrinos of different types are produced in association with muons and electrons, and it is possible to calculate the expected flux of each type with some accuracy. Production of other types of neutrinos is predicted to be quite small. The detected flux of muon neutrinos is significantly lower than the calculation, in analogy with a similar deficit in the neutrino flux from the Sun. Data from the large detector Super Kamiokande in Japan gave the first indication that the atmospheric deficit is due to transformation (known as oscillation) of muon neutrinos into other types of neutrinos. The Sudbury Neutrino Observatory in Canada has confirmed this transformation, demonstrating that the rest mass of the neutrino, while very small, is not zero. Astrophysical consequences of a nonzero rest mass are profound, as a particle with a rest mass interacts gravitationally in a way totally different from that of a particle (such as a photon) with no rest mass. Huge numbers of neutrinos permeate the universe, and details of their gravitational interaction are crucial to the understanding of galaxy formation. *See* NEUTRINO.

Nucleonic component. The high-energy nucleons—the knock-on protons and neutrons—produced by the primary-particle collisions and a few pion collisions proceed down into the atmosphere. They produce

nuclear interactions of the same kind as the primary nuclei, though of course with diminished energies. This cascade process forms the nucleonic component of the secondary cosmic rays.

When nucleon energy falls below about 100 MeV, stars and further knock-ons can no longer be produced. At the same time the protons are rapidly disappearing from the cascade because their ionization losses in the air slow them down before they can make a nuclear interaction. Most of the hadrons in the lower atmosphere are thus neutrons, which are already dominant at 3500 m (11,550 ft), about 300 g/cm^2 (4.3 lb/in.^{2}) above sea level, where they outnumber the protons four to one. Thus the final stages of the cascade involve mainly neutrons in a sequence of low-energy interactions which convert them to thermal neutrons (neutrons of kinetic energy of about 0.025 eV) in a path of about 90 g/cm^2 (1.3 lb/in.^{2}). These thermal neutrons are readily detected in boron trifluoride (BF_3) and helium-3 (^3He) counters. The nucleonic component increases in intensity down to a depth of about 120 g/cm^2 (1.7 g/cm^3), and thereafter declines in intensity, with a mean absorption length of about 200 g/cm^2 (2.8 lb/in.^{2}).

The various cascades of secondary particles in the atmosphere are shown schematically in **Fig. 1**. About 48% of the initial primary cosmic-ray energy goes into charged pions, 25% into neutral pions, 7% into the nucleonic component, and 20% into stars. The nucleonic component is produced principally by the lower-energy (about 5 GeV) primaries. Higher-energy primaries put their energy more into meson production. Hence in the lower atmosphere, a Geiger counter responds mainly to the higher-energy primaries (about 5 GeV) because it counts the muons and electrons, whereas a BF_3 counter detecting thermal neutrons responds more to the low-energy primaries.

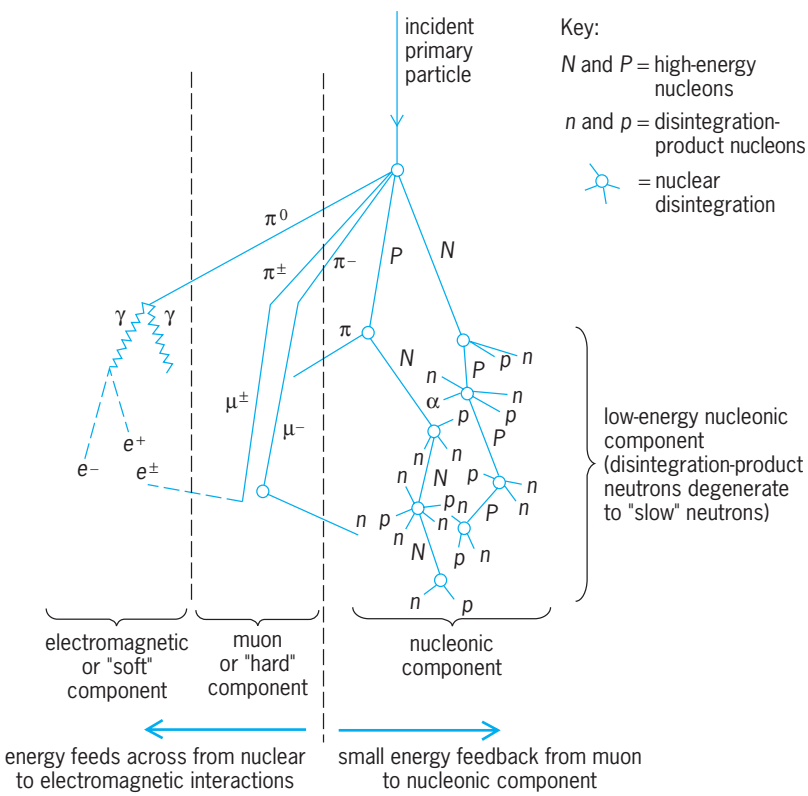

Fig. 1. Cascade of secondary cosmic-ray particles in the terrestrial atmosphere.

Neutrinos. Cosmic neutrinos, detected for the first time from the explosion of the supernova 1987A, provide confirmation of theoretical calculations regarding the collapse of the cores of massive stars. Although neutrinos are produced in huge numbers (over 10^{15} passed through a typical human body from this supernova), they interact with matter only very weakly, necessitating a very large detector. Detectors consisting of huge tanks containing hundreds of tons of pure water located deep underground to reduce the background produced by other cosmic rays recorded less than two dozen neutrino events. Still larger detectors, which are under construction in the Antarctic ice and underwater at several locations, will permit observation of more distant supernovae and allow sensitive searches for point sources of high-energy neutrinos. Additionally, by measuring the fraction of non-neutrino-induced events containing multiple muons, these new detectors can investigate the composition of cosmic rays at energies above 10^{15} eV. Some preliminary measurements indicate that these high-energy cosmic rays may consist primarily of iron nuclei rather than the protons that dominate at lower energies. Much of the interest in the new neutrino observatories derives from the success of the now maturing field of measurement of the flux of solar neutrinos, which is really quite a different problem. With the realization that neutrinos have mass, increasingly precise measurements of the solar neutrino flux, coupled with such techniques as helioseismology, continue to make fundamental contributions to the study of the internal structure of the Sun. *See* NEUTRINO ASTRONOMY; SOLAR NEUTRINOS; SUPERNOVA.

Relation to particle physics. Investigations of cosmic rays continue to make fundamental contributions to particle physics. Neutrino detectors, besides detecting oscillations of atmospheric neutrinos, have set the best limit yet (about 10^{32} years) on the lifetime of the proton. Cosmic rays remain the only source of particles with energies above 1000 GeV. With the continued increase in the size and sensitivity of detectors, study of cosmic rays should continue to provide the first indications of new physics at ultrahigh energies. *See* FUNDAMENTAL INTERACTIONS; PROTON.

Geomagnetic effects. The magnetic field of Earth is described approximately as that of a magnetic dipole of strength 8.1×10^{15} weber-meters (8.1×10^{25} gauss · cm^3) located near the geometric center of Earth. Near the Equator the field intensity is 3×10^{-5} tesla (0.3 gauss), falling off in space as the inverse cube of the distance to the Earth's center. In a magnetic field which does not vary in time, the path of a particle is determined entirely by its rigidity, or momentum per unit charge; the velocity simply determines how fast the particle will move along

TABLE 2. Properties of particles when all have a rigidity of 1 gigavolt

Particle	Charge	Nucleons	Kinetic energy MeV	Kinetic energy MeV/nucleon	Momentum, MeV/*c*
Electron	1	—	1000	—	1000
Proton	1	1	430	430	1000
^{3}He	2	3	640	213	2000
^{4}He	2	4	500	125	2000
^{16}O	8	16	2000	125	8000

this path. Momentum is usually expressed in units of eV/*c*, where *c* is the velocity of light, because at high energies, energy and momentum are then numerically almost equal. By definition, momentum and rigidity are numerically equal for singly charged particles. The unit so defined is dimensionally a volt, but the relationship to electric potential is neither obvious nor particularly useful in practice. **Table 2** gives examples of these units as applied to different particles with rigidity of 1 gigavolt. This corresponds to an orbital radius in a typical interplanetary (10^{-9} tesla or 10^{-5} gauss) magnetic field of approximately 10 times the distance from the Earth to the Moon. *See* RELATIVISTIC ELECTRODYNAMICS.

The minimum rigidity of a particle able to reach the top of the atmosphere at a particular geomagnetic latitude is called the geomagnetic cutoff rigidity at that latitude, and its calculation is a complex numerical problem. Fortunately, for an observer near the ground, obliquely arriving secondary particles, produced by the oblique primaries, are so heavily attenuated by their longer path to the ground that it is usually sufficient to consider only the geomagnetic cutoff for vertically incident primaries, which is given in **Table 3**. Around the Equator, where a particle must come in perpendicular to the geomagnetic lines of force to reach Earth, particles with rigidity less than 10 GV are entirely excluded, though at higher latitudes where entry can be made more nearly along the lines of force, lower energies can reach Earth. Thus, the cosmic-ray intensity is a minimum at the Equator, and increases to its full value at either pole—this is the cosmic-ray latitude effect. Even deep in the atmosphere the variation with latitude is easily detected with BF_3 counters (**Fig. 2**). North of 45° the effect is slight because the additional primaries admitted are so low in energy that they produce few secondaries.

TABLE 3. Geomagnetic cutoff

Geomagnetic latitude	Vertical cutoff, gigavolts
0°	15
±20°	11.5
±40°	5
±60°	1
±70°	0.2
±90°	0

Accurate calculations of the geomagnetic cutoff must consider the deviations of the true field from that of a perfect dipole and the change with time of these deviations. Additionally the distortion of the field by the pressure of the solar wind must often be accounted for, particularly at high latitude. Such corrections vary rapidly with time because of sudden bursts of solar activity and because of the rotation of the Earth. Areas with cutoffs of 400 MV during the day may have no cutoff at all during the night. This day-night effect is confined to particles with energies so low that neither they nor their secondaries reach the ground, and is thus observed only on high-altitude balloons or satellites.

Since the geomagnetic field is directed from south to north above the surface of Earth, the incoming cosmic-ray nuclei are deflected toward the east. Hence an observer finds some 20% more particles incident from the west. This is known as the east-west effect. *See* GEOMAGNETISM.

Solar modulation. Figure 3 presents portions of the proton and alpha-particle spectra observed near the Earth but outside of the magnetosphere in 1973. Below 20 GeV/nucleon the cosmic-ray intensity varies markedly with time. S. Forbush was the first to show that the cosmic-ray intensity was low during the years of high solar activity and sunspot number, which follow an 11-year cycle. This effect is clearly seen in the data of Fig. 2 and has been extensively studied with ground-based and spacecraft instruments. While this so-called solar modulation is now understood in general terms, it has not been

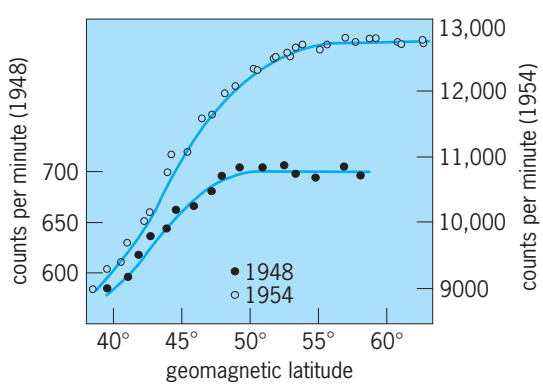

Fig. 2. Latitude variation of the neutron component of cosmic rays in 80°W longitude and at a height corresponding to an atmospheric pressure of 30 kPa (22.5 cm of mercury) in 1948, when the Sun was active, and 1954, when the Sun was deep in a sunspot minimum.

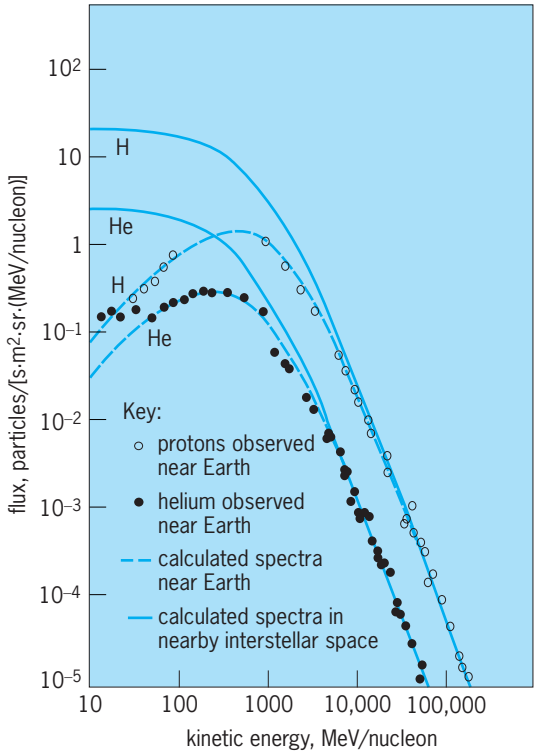

Fig. 3. Spectra of cosmic-ray protons and helium at Earth and in nearby interstellar space, showing the effect of solar modulation. Observations were made in 1973, when the Sun was quiet.

calculated in detail, in large part because there are few direct measurements out of the ecliptic plane and in the outer heliosphere.

The primary cause of solar modulation is the solar wind, a highly ionized gas (plasma) which originates from the solar corona and propagates radially from the Sun at a velocity of about 400 km/s (250 mi/s). The wind is mostly hydrogen, with typical density of 5 protons/cm^3 (80 protons/in.^{3}). This density is too low for collisions with cosmic rays to be important. Rather, the high conductivity of the medium traps part of the solar magnetic field and carries it outward. The rotation of the Sun and the radial motion of the plasma combine to create the observed archimedean spiral pattern of the average interplanetary magnetic field. Turbulence in the solar wind creates fluctuations in the field which often locally obscure the average direction and intensity. This complex system of magnetic irregularities propagating outward from the Sun deflects and sweeps the low-rigidity cosmic rays out of the solar system. *See* SOLAR MAGNETIC FIELD.

In addition to the bulk sweeping action, another effect of great importance occurs in the solar wind, adiabatic deceleration. Because the wind is blowing out, only those particles which chance to move upstream fast enough are able to reach Earth. However, because of the expansion of the wind, particles interacting with it lose energy. Thus, particles observed at Earth at 10 MeV/nucleon actually started out at several hundred megaelectronvolts per nucleon in nearby interstellar space, while those with only 100–200 MeV/nucleon initial energy probably never reach Earth at all. This is particularly unfortunate because at these lower energies the variation with energy of nuclear reaction probabilities would allow much more detailed investigation of cosmic-ray history. Changes in the modulation with solar activity are caused by the changes in the pattern of magnetic irregularities rather than by changes in the wind velocity, which are quite small. *See* MAGNETOHYDRODYNAMICS; PLASMA (PHYSICS).

Heliosphere. Solar modulation is important in a region around the Sun termed the heliosphere, a large bubble formed in the interstellar medium by the solar wind. The density, and therefore the energy and momentum, of the solar wind drop as the material expands with increasing distance from the Sun, eventually becoming too small to push back the interstellar material. The typical distance to the interface is thought to be approximately 100 AU, but the actual distance in any direction is determined by local variations in both the solar wind and the interstellar medium. (1 AU, or astronomical unit, is the average Earth-Sun separation, 1.49×10^8 km or 9.26×10^7 mi.).

The spacecraft *Voyager 1* crossed the termination shock of the solar wind on December 16, 2004, at some 94 astronomical units (AU) or more than 8.7×10^9 miles from the Sun, as evidenced by an abrupt increase in the magnetic field. The termination shock is the innermost, and probably the best-defined, structure in this boundary region. Outside the termination shock several centuries worth of decelerated solar wind are probably piled up, producing a region that is still capable of modulating cosmic-ray intensity. The Sun, carrying the heliosphere with it, is moving through the interstellar medium at approximately 20 km/s (12 mi/s). Eventually all of the solar material blends into the interstellar medium by turbulent interactions. The termination shock had been universally thought to be a prodigious accelerator of particles and *Voyager 1* largely confirmed this. At the shock there is a remarkable increase in particle intensity with a distinctive energy spectrum.

Forbush decreases. Apart from the 11-year modulation cycle, there are many different types of cosmic-ray variation associated with irregularities in the solar wind. The most dramatic is the Forbush decrease, wherein worldwide cosmic-ray intensity may drop as much as 20% in one day, followed by a slow recovery lasting many days or even weeks. Most Forbush decreases are associated with severe magnetic disturbances in the solar wind that result from massive ejections of material from the solar corona into interplanetary space. Often these ejections accompany solar flares. When magnetic disturbances encounter the Earth, they can cause geomagnetic storms and other phenomena that are disruptive to human activity. This complex set of interactions has come to be called space weather. Observing changes in cosmic-ray fluxes from several places on Earth simultaneously is one important tool for investigating the interaction of a magnetic disturbance with Earth. *See* SOLAR WIND; SUN.

Composition of cosmic rays. Nuclei ranging from protons to lead have been identified in the cosmic radiation. The relative abundances of the elements ranging up to nickel are shown in **Fig. 4**, together with the best estimate of the "universal abundances" obtained by combining measurements of solar spectra, lunar and terrestrial rocks, meteorites, and so forth. Most obvious is the similarity between these two distributions. However, a systematic deviation is quickly apparent: the elements lithium-boron and scandium-manganese as well as most of the odd-charged nuclei are vastly overabundant in the cosmic radiation. This effect has a simple explanation: the cosmic rays travel great distances in the Milky Way Galaxy and occasionally collide with atoms of interstellar gas—mostly hydrogen and helium—and fragment. This fragmentation, or spallation as it is called, produces lighter nuclei from heavier ones but does not change the energy/nucleon very much. Thus the energy spectra of the secondary elements are similar to those of the primaries. *See* SPALLATION REACTION.

Calculations involving reaction probabilities determined by nuclear physicists show that the overabundances of the secondary elements can be explained by assuming that cosmic rays pass through an average of about 5 g/cm^2 (0.07 lb/in.^{2}) of material on their way to Earth. Although an average path length can be obtained, it is not possible to fit the data by saying that all particles of a given energy have exactly the same path length; furthermore, results indicate that higher-energy particles traverse less matter in reaching the solar system, although their original composition seems energy independent. *See* ELEMENTS, COSMIC ABUNDANCE OF.

When spallation has been corrected for, differences between cosmic-ray abundances and solar-system or universal abundances still remain. The most important question is whether these differences are due to the cosmic rays having come from a special kind of material (such as would be produced in a supernova explosion), or simply to the fact that some atoms might be more easily accelerated than others. It is possible to rank almost all of the overabundances by considering the first ionization potential of the atom and the rigidity of the resulting ion, although this approach gives no prediction of the magnitude of the enhancement. Relative abundances of particles accelerated in solar flares are also far from constant from one flare to the next. Accounting for these abundance variations is one of the most important constraints on models of solar particle acceleration, the exact mechanism of which remains an unsolved problem. *See* IONIZATION POTENTIAL.

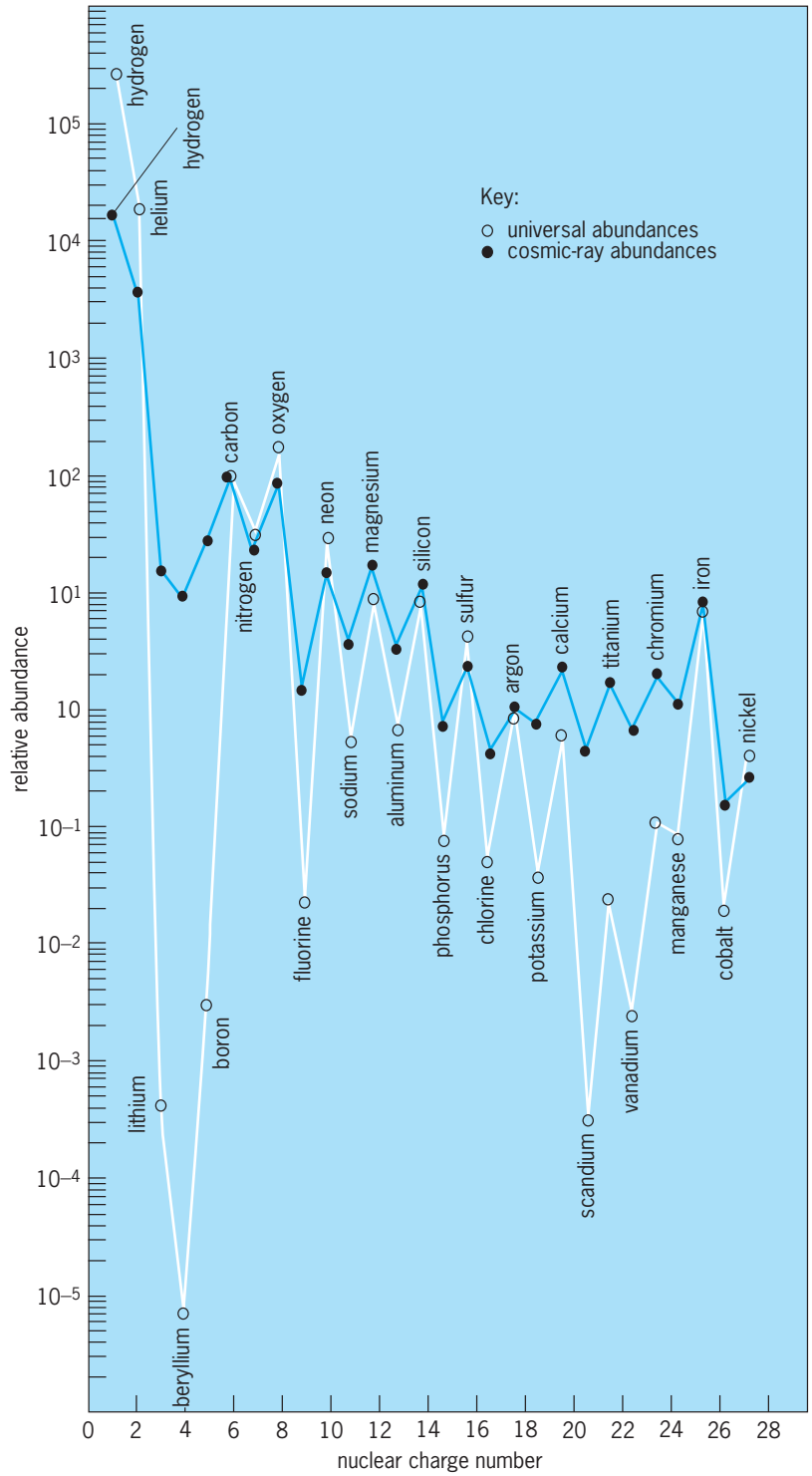

Fig. 4. Cosmic-ray abundances compared to the universal abundances of the elements. Carbon is set arbitrarily to an abundance of 100 in both cases.

Isotopes. Much current cosmic-ray research is concentrated on determining isotopic composition of the elements, partly because this is less likely to be changed by acceleration than the elemental composition and thus is more accurately representative of the composition of the source material. As an example, the low-energy helium data in Fig. 3 are not well represented by the calculation. The excess flux, which is referred to as the anomalous component, is nearly all ^{4}He, whereas higher-energy cosmic rays are nearly 10% ^{3}He. A similar enhancement of low-energy nitrogen is pure ^{14}N, while at higher energies nitrogen is half ^{15}N. Measuring isotopes allows conclusive identification of the anomalous component as a sample of originally neutral interstellar material that has been ionized and energized by processes in the solar wind.

Other variations in the isotopic composition are not currently understood. For example, the ratio of

^{22}Ne to ^{20}Ne in the cosmic-ray sources is estimated to be 0.37, while the accepted solar system value for this number is 0.12, which agrees well with the abundances measured in solar-flare particles. However, another direct sample of solar material—the solar wind—has a ratio of 0.08, indicating clearly that the isotopic composition of energetic particles need not reflect that of their source. Conclusions drawn from the observed difference in the solar and cosmic-ray values must be viewed as somewhat tentative until the cause of the variation in the solar material is well understood. *See* ISOTOPE.

Electron abundance. Cosmic-ray electron measurements pose other problems of interpretation, partly because electrons are nearly 2000 times lighter than protons, the next lightest cosmic-ray component. Protons with kinetic energy above 1 GeV are about 100 times as numerous as electrons above the same energy, with the relative number of electrons decreasing slowly at higher energies. But it takes about 2000 GeV to give a proton the same velocity as a 1-GeV electron. Viewed in this way electrons are several thousand times more abundant than protons. (Electrical neutrality of the Milky Way Galaxy is maintained by lower-energy ions which are more numerous than cosmic rays although they do not carry much energy.) It is thus quite possible that cosmic electrons have a different source entirely from the nuclei. It is generally accepted that there must be direct acceleration of electrons, because calculations show that more positrons than negatrons should be produced in collisions of cosmic-ray nuclei with interstellar gas. Measurements show, however, that only 10% of the electrons are positrons. As the number of positrons seen agrees with the calculated secondary production, added confidence is gained in the result that there is indeed an excess of negatrons. *See* ELECTRON.

Electrons are light enough to emit a significant amount of synchrotron radiation as they are deflected by the 10^{-10}-tesla (10^{-6}-gauss) galactic magnetic field. Measurement of this radiation by radio telescopes provides sufficient data for an approximate calculation of the average energy spectrum of electrons in interstellar space and other galaxies. Comparison of spectra of electrons and positrons measured at Earth with those calculated to exist in interstellar space provides the most direct measurement of the absolute amount of solar modulation. *See* RADIO ASTRONOMY; SYNCHROTRON RADIATION.

Properties of the energy spectrum. At energies above 10^{10} eV, the energy spectra of almost all cosmic rays are approximated over many decades by functions in which the flux decreases as the energy raised to some negative, nonintegral power referred to as the spectral index. Such a power-law relationship is of course a straight line when plotted using logarithmic axes. A steep or "soft" (that is, more rapidly falling with increasing energy) spectrum thus has a higher spectral index than a flat or "hard" spectrum. The straight-line regions of the spectra in Fig. 3 correspond to a variation of flux with a spectral index of -2.7. A spectral index of -2.7 provides a good fit with the data up to 10^{15} eV total energy. Between 10^{15} and 10^{19} eV a steeper spectrum, with an index around -3.0, seems to be well established. Above 10^{19} eV the spectrum hardens once more, returning to an index of about -2.7. The spectral index above 10^{20} eV has not been determined, because particles are so rare that they are almost never seen, even in detectors which cover several square kilometers and operate for many years. At such high energies, the individual particles are not identified, and changes in the measured-energy spectrum could be the result of composition changes. However, the evidence available indicates that the composition is essentially unchanged.

The Pierre Auger Observatory, which began operation in 2005, has a detection area the size of Rhode Island (over 3000 km^2 or 1200 mi^2) located in western Argentina's Mendoza Province. The Auger Observatory is a hybrid detector, employing two independent methods to detect and study high-energy cosmic rays. One technique is ground-based and detects high-energy particles through their interaction with water. Each of the 1600 detectors contains 11,000 L (3000 gals) of ultrapure, deionized water. The other technique tracks the development of air showers by observing ultraviolet light emitted high in the Earth's atmosphere.

Age. Another important result which can be derived from detailed knowledge of cosmic-ray isotopic composition is the "age" of cosmic radiation. Certain isotopes are radioactive, such as beryllium-10 (^{10}Be) with a half-life of 1.6×10^6 years. Since beryllium is produced entirely by spallation, study of the relative abundance of ^{10}Be to the other beryllium isotopes, particularly as a function of energy to utilize the relativistic increase in this lifetime, will yield a number related to the average time since the last nuclear collision. Measurements show that ^{10}Be is nearly absent at low energies, yielding an estimate of the age of the cosmic rays of approximately 10^7 years. An implication of this result is that the cosmic rays propagate in a region in space which has an average density of 0.1–0.2 atom/cm^3 (1.5–3 atoms/in.^{3}). This is consistent with some astronomical observations of the immediate solar neighborhood.

Very high energy particles cannot travel long distances in the 2.7 K blackbody-radiation field which permeates the universe. Electrons of 15 GeV energy lose a good portion of their energy in 10^8 years by colliding with photons via the (inverse) Compton process, yet electrons are observed to energies of 100 GeV and over. A similar loss mechanism becomes effective at approximately 10^{20} eV for protons. These observations are of course not conclusive, but a safe statement is that a cosmic-ray age of 10^7 years is consistent with all currently available data. *See* COMPTON EFFECT; COSMIC BACKGROUND RADIATION.

Several attempts have been made to measure the constancy of the cosmic-ray flux in time. Variations in ^{14}C production, deduced from apparent deviations of the archeological carbon-dating scale from that derived from studies of tree rings, cover a period of about 10^3 years. Radioactive ^{10}Be in deep-sea sediments allows studies over 10^6 years, whereas etching of tracks left by cosmic rays in lunar minerals covers a

period of 10^9 years. None of these methods has ever indicated a variation of more than a factor of 2 in average intensity. There are big differences in these time scales, and the apparent constancy of the flux could be due to averaging over variations which fall in the gaps as far as the time scales are concerned. Nevertheless, the simplest picture seems to be that the cosmic rays are constant in time at an intensity level which is due to a long-term balance between continuous production and escape from the Galaxy, with an average residence time of 10^7 years. *See* COSMOGENIC NUCLIDE; RADIOCARBON DATING.

Although small, variations in cosmic-ray intensity show systematic patterns over historical and geologic time scales. Some of this variation is undoubtedly due to variability in the magnetic field of the Earth that results in changes in the geomagnetic cutoff. However, over the past several thousand years, when the magnetic field has been reasonably constant, cosmic-ray intensity has shown a distinct anticorrelation with measures of overall global temperature. Low cosmic-ray fluxes, indicative of enhanced solar activity, typically go together with warm periods. The relationship is far from understood, but this is one of many observations that must be considered in the complex, nonlinear problem of global warming.

Origin. Although study of cosmic rays has yielded valuable insight into the structure, operation, and history of the universe, their origin has not been determined. The problem is not so much to devise processes which might produce cosmic rays, but to decide which of many possible processes do in fact produce them.

In general, analysis of the problem of cosmic-ray origin is broken into two major parts: origin in the sense of where the sources are located, whatever they are, and origin in the sense of how the particles are accelerated to such high energies. Of course, these questions can never be separated completely.

Location of sources. It is thought that cosmic rays are produced by mechanisms operating within galaxies and are confined almost entirely to the galaxy of their production, trapped by the galactic magnetic field. The intensity in intergalactic space would only be a few percent of the typical galactic intensity, and would be the result of a slow leakage of the galactic particles out of the magnetic trap. It has not been possible to say much about where the cosmic rays come from by observing their arrival directions at Earth. At lower energies (up to 10^{15} eV) the anisotropies which have been observed can all be traced to the effects of the solar wind and interplanetary magnetic field. The magnetic field of the Milky Way Galaxy seems to be completely effective in scrambling the arrival directions of these particles.

For several decades in energy above 10^{15} a smoothly rising anisotropy is measured, ranging from 0.1 to 10%, but the direction of the maximum intensity varies in a nonsystematic way with energy. At these energies, particles have a radius of curvature which is not negligible compared to galactic structures, and thus their arrival direction could be related to where they came from but in a complex way. Above 10^{19} eV the radius of curvature in the galactic magnetic field becomes comparable to or larger than galactic dimensions, making containment of such particles in the disk of the Milky Way Galaxy impossible. Only a few hundred events greater than 10^{19} eV have been detected, with no systematic anisotropy. Clustering of events (of marginal statistical significance) in the data of individual observatories has been reported, but so far never confirmed by other detectors. The Pierre Auger Observatory should, over several years, increase the number of events observed by orders of magnitude and may possibly permit statistically definitive determinations of features in the arrival direction distribution. Actual identification of sources, however, will probably come only from detection of neutral radiation (photons and neutrinos) produced in cosmic-ray interactions.

Cerenkov telescopes, such as HESS in Namibia and MAGIC at the Roque de los Muchachos Observatory in the Canary Islands, have produced remarkable images of objects in the light of 10^{12}-eV gamma rays. Supernova remnants have been clearly resolved to show structures within them, and emission from x-ray binary systems has been definitively observed. Several sources unassociated with known objects are also being studied. It is generally believed that gamma rays of energy greater than 10^{12} eV can be produced only through interactions of high-energy protons or other hadrons. These gamma rays provide important information on the structure and operation of these exotic objects but the key question of whether these particles escape to become galactic cosmic rays remains unanswered. *See* ASTROPHYSICS, HIGH-ENERGY; BINARY STAR; CERENKOV RADIATION; X-RAY ASTRONOMY.

Direct detection of cosmic rays propagating in distant regions of the Milky Way Galaxy is possible by observing the electromagnetic radiation produced as they interact with other constituents of the Milky Way Galaxy. Measurement of the average electron spectrum using radio telescopes has already been mentioned. Proton intensities are mapped by studying the arrival directions of gamma rays produced as they collide with interstellar gas. Unfortunately, the amount of radiation in these processes depends upon both the cosmic-ray flux and the magnetic-field intensity or density of interstellar gas. Areas where cosmic rays are known to exist can be pointed out because the radiation is observed. But where no radiation is seen, it is not known whether its absence results from lack of cosmic rays or lack of anything for them to interact with. In particular, very little radiation is seen from outside the Milky Way Galaxy, but there is also very little gas or magnetic field there. There is therefore no direct evidence either for or against galactic containment.

A major difficulty with the concept of cosmic radiation filling the universe is the large amount of energy needed to maintain the observed intensity in the face of an expanding universe—probably more energy than is observed to be emitted in all other forms put together. *See* COSMOLOGY.

Confinement mechanisms. Three possible models of cosmic-ray confinement are under investigation. All assume that cosmic rays are produced in sources, discrete or extended, scattered randomly through the galactic disk. Most popular is the "leaky box" model, which proposes that the particles diffuse about in the magnetic field for a few million years until they chance to get close to the edge of the Milky Way Galaxy and escape. This is a phenomenological model in that no mechanism is given by which either the confinement time or the escape probability as a function of energy can be calculated from independent observations of the galactic structure. Its virtue is that good fits to the observed abundances of spallation products are obtained by using only a few adjustable parameters. Variations of the model which mainly postulate boxes within boxes—ranging from little boxes surrounding sources to a giant box or static halo surrounding the whole Milky Way Galaxy—can be used to explain variations from the simple predictions. However, all attempts to calculate the details of the process have failed by many orders of magnitude, predicting ages which are either far older or far younger than the observed age.

A second model is that of the dynamical halo. Like the earlier static-halo model, it is assumed that cosmic rays propagate not only in the galactic disk but also throughout a larger region of space, possibly corresponding to the halo or roughly spherical but sparse distribution of material which typically surrounds a galaxy. This model is based on the observation that the energy density of the material which is supposed to be contained by the galactic magnetic field is comparable to that of the field itself. This can result in an unstable situation in which large quantities of galactic material stream out in a galactic wind similar in some respects to the solar wind. In this case the outward flow is a natural part of the theory, and calculations have predicted reasonable flow rates. In distinction to the solar wind, in which the cosmic rays contribute almost nothing to the total energy density, they may provide the dominant energy source in driving the galactic wind.

A third model assumes that there is almost no escape; that is, cosmic rays disappear by breaking up into protons which then lose energy by repeated collision with other protons. To accept this picture, one must consider the apparent 5-g/cm^2 (0.07-lb/in.^{2}) mean path length to be caused by a fortuitous combination of old distant sources and one or two close young ones. Basically, the objections to this model stem from the tendency of scientists to accept a simple theory over a more complex (in the sense of having many free parameters) or specific theory when both explain the data. *See* MILKY WAY GALAXY.

Acceleration mechanisms. Although the energies attained by cosmic-ray particles are extremely high by laboratory standards, their generation can probably be understood in terms of known astronomical objects and laws of physics. Even on Earth, ordinary thunderstorms generate potentials of millions of volts, which would accelerate particles to respectable cosmic-ray energies (a few megaelectronvolts) if the atmosphere were less dense. Gamma rays and neutrons have, in fact, been detected from lightning strikes. Consequently, there are many theories of how the acceleration could take place, and it is quite possible that more than one type of source exists. Two major classes of theories may be identified—extended-acceleration regions and compact-acceleration regions.

Extended-acceleration regions. Acceleration in extended regions (in fact the Milky Way Galaxy as a whole) was first proposed by E. Fermi, who showed that charged particles could gain energy from repeated deflection by magnetic fields carried by the large clouds of gas which are known to be moving randomly about the Milky Way Galaxy. Many other models based on such statistical acceleration have since been proposed, the most recent of which postulates that particles bounce off shock waves traveling in the interstellar medium. Such shocks, supposed to be generated by supernova explosions, undoubtedly exist to some degree but have an unknown distribution in space and strength, leaving several free parameters which may be adjusted to fit the data.

Compact-acceleration regions. The basic theory in the compact-acceleration class is that particles are accelerated directly in the supernova explosions themselves. One reason for the popularity of this theory is that the energy generated by supernovas is of the same order of magnitude as that required to maintain the cosmic-ray intensity in the leaky box model.

However, present observations indicate that the acceleration could not take place in the initial explosion. Cosmic rays have a composition which is similar to that of ordinary matter and is different from the presumed composition of the matter which is involved in a supernova explosion. At least some mixing with the interstellar medium must take place. Another problem with an explosive origin is an effect which occurs when many fast particles try to move through the interstellar gas in the same direction: the particles interact with the gas through a magnetic field which they generate themselves, dragging the gas along and rapidly losing most of their energy. In more plausible theories of supernova acceleration, the particles are accelerated gradually by energy stored up in the remnant by the explosion or provided by the intense magnetic field of the rapidly rotating neutron star or pulsar which is formed in the explosion.

Acceleration of high-energy particles was first observed in the Crab Nebula, the remnant of a supernova observed by Chinese astronomers in 1054. This nebula is populated by high-energy electrons which radiate a measurable amount of their energy as they spiral about in the magnetic field of the nebula. So much energy is released that the electrons would lose most of their energy in a century if it were not being continuously replenished. Pulses of gamma rays also show that bursts of high-energy particles are being produced by the neutron star—the gamma rays coming out when the particles interact

with the atmosphere of the neutron star. Particles of cosmic-ray energy are certainly produced in this object, but it is unlikely that they escape from the trapping magnetic fields in the nebula and join the freely propagating cosmic ray population. As noted above, observation of 10^{12}-eV gamma rays now allows mapping of energetic hadrons in these objects, but the issue of an escape mechanism remains. *See* CRAB NEBULA; NEUTRON STAR; PULSAR.

Acceleration in the solar system. The study of energetic particle acceleration in the solar system is valuable in itself, and can give insight into the processes which produce galactic cosmic rays. Large solar flares, about one a year, produce particles with energies in the gigaelectronvolt range, which can be detected through their secondaries even at the surface of the Earth. It is not known if such high-energy particles are produced at the flare site itself or are accelerated by bouncing off the shock fronts which propagate from the flare site outward through the solar wind. Nuclei and electrons up to 100 MeV are regularly generated in smaller flares. In many events it is possible to measure gamma rays and neutrons produced as these particles interact with the solar atmosphere. X-ray, optical and radio mappings of these flares are also used to study the details of the acceleration process. By relating the arrival times and energies of these particles at detectors throughout the solar system to the observations of their production, the structure of the solar and interplanetary magnetic fields may be studied in detail.

In addition to the Sun, acceleration of charged particles has been observed in the vicinity of the Earth, Mercury, Jupiter, and Saturn—those planets which have significant magnetic fields. Details of the acceleration mechanism are not understood, but certainly involve both the rotation of the magnetic fields and their interactions with the solar wind. Jupiter is such an intense source of electrons below 30 MeV that it dominates other sources at the Earth when the two planets lie along the same interplanetary magnetic field line of force. The anomalous component is known to be local interstellar material, accelerated by a mechanism that is not fully understood but probably is situated near the interface between the solar wind and the interstellar medium. However, there was no evidence in the *Voyager 1* crossing of the termination shock that this production occurs immediately at this particular structure. *See* JUPITER; MERCURY (PLANET); PLANET; SATURN.

Direct observation of conditions throughout most of the solar system will be possible in the next few decades, and with it should come a basic understanding of the production and propagation of energetic particles locally. This understanding will perhaps form the basis of a solution to the problem of galactic cosmic rays, which will remain for a very long time the main direct sample of material from the objects of the universe outside the solar system.

Paul Evenson

Bibliography. V. S. Berezinskii et al., *Astrophysics of Cosmic Rays*, 1991; R. Clay and B. Dawson, *Cosmic Bullets*, 1997; M. W. Friedlander, *Cosmic Rays*, 1989; T. K. Gaisser, *Cosmic Rays and Particle Physics*, 1990; P. K. F. Greider, *Cosmic Rays at Earth: Researcher's Reference Manual and Data Book*, 2001; M. B. Kallenrode, *Space Physics: An Introduction to Plasmas and Particles in the Heliosphere and Magnetospheres*, 2d ed., 2001; M. S. Longair, *High Energy Astrophysics*, vol. 1: *Particles, Photons, and Their Detection*, 2d ed., 1992; J. M. Matthews, *High Energy Astrophysics: Theory and Observations from MeV to TeV*, 1994; L. I. Miroshnichenko, *Solar Cosmic Rays*, 2001.

Cosmic spherules

Solidified droplets of extraterrestrial materials that melted either during high-velocity entry into the atmosphere or during hypervelocity impact of large meteoroids onto the Earth's surface. Cosmic spherules are rounded particles that are millimeter to microscopic in size and that can be identified by unique physical properties. Although great quantities of the spheres exist on the Earth, they are ordinarily found only in special environments where they have concentrated and are least diluted by terrestrial particulates. *See* METEOR.

The most common spherules are ablation spheres produced by aerodynamic melting of meteoroids as they enter the atmosphere. Typical ablation spheres are produced by melting of submillimeter asteroidal and cometary fragments that enter the atmosphere at velocities ranging from 6.5 to 43 mi (11 to 72 km) per second. Approximately 10,000 tons of such particles collide with the Earth each year, and cosmic spheres in the size range of 0.004 to 0.04 in. (0.1 to 1.0 mm) are the most abundant form of this material that survives to reach the Earth's surface. The spheres are formed near 48 mi (80 km) altitude, where deceleration, intense frictional heating, melting, partial vaporization, and solidification all occur in only a few seconds time. During formation, the larger particles can be seen as luminous meteors or shooting stars. Impact spheres constitute a second and rarer class of particles that are produced when giant meteoroids impact the Earth's surface with sufficient velocity to produce explosion craters that eject molten droplets of both meteoroid and target materials.

Impact spheres. These are very abundant on the Moon, but they are rare on the Earth, and they have been found in only a few locations. Impacts large enough to produce explosion craters occur on the Earth every few tens of thousands of years, but the spheres and the craters themselves are rapidly degraded by weathering and geological processes. Meteoritic spherules have been found around a number of craters, including Meteor Crater in Arizona, Wabar in Saudi Arabia, Box Hole and Henbury in Australia, Lonar in India, and Morasko in Poland. They have also been found at the Sikhote-Alin meteorite shower site and at the location of the Tunguska explosion in Russia. Spheres from these craters include ablation spheres as well as true impact spheres produced

either by shock melting of target and meteoroid or by condensation from impact-generated vapor. Impact spheres are also produced by the larger cratering events, some of which may have played roles in biological extinctions. Silica-rich glass spheroids (microtektites) are found in thin layers that are contemporaneous with the conventional tektites. Microtektites are believed to be shock-melted sedimentary materials that were ejected from large impact craters. They were ejected as plumes that covered substantial fractions of the surface of the Earth. The cumulative mass of microtektites in the 35-million-year-old North American tektite field is estimated to be equivalent in mass to 36 mi^3 (100 km^3) of solid rock. Microspherules of a different composition have been found in the thin iridium-rich layer associated with the global mass extinctions at the Cretaceous–Tertiary boundary. *See* TEKTITE.

Ablation spheres. These fall to Earth at a rate of one 0.1-mm-diameter sphere per square meter per year, and every rooftop contains these particles. Unfortunately they are usually mixed in with vast quantities of terrestrial particulates, and they are very difficult to locate. They can, however, be easily found in special environments that do not contain high concentrations of terrestrial particles that could be confused with cosmic spheres larger than 0.004 in. (0.1 mm) in diameter. Such environments include the mid-Pacific ocean floor and certain ice deposits in Greenland and Antarctica. In mid-Pacific sediments the accumulation of terrestrial sediments is only a few meters per million years, and the spherules larger than 0.004 in. (0.1 mm) than are found in concentrations of roughly 10–100 per kilogram of sediment. Cosmic spheres are easily extracted from the sediment because most spherules are ferromagnetic and much larger than typical sediment particles. The highest ablation spherule concentrations on Earth (over one per gram of sediment) are found in the melt zones of the Greenland ice cap, where melting ice leaves dust particles concentrated directly on the ice surface. *See* MARINE SEDIMENTS.

Cosmic ablation spherules can be grouped into two major types: type S (stony) and type I (iron). In polished sections it can be seen that the S spheres are composed of olivine, magnetite, and glass, with textures consistent with rapid crystallization from melt. Both of the type S spherules shown in the **illustration** have elemental compositions similar to those of stony meteorites. In general, types S spheres have compositions that are a close match with chondritic meteorites except for depletion of volatile elements such as sulfur and sodium that are lost during atmospheric entry. Many spheres are also depleted in nickel by a process that may also lead to the formation of the iron spheres. Typical type I spheres consist of iron oxide (magnetite and wustite) surrounding a core of either nickel-iron metal or a small nugget of platinum group elements. Nickel and the platinum group elements are concentrated during brief oxidation when the sphere is molten. Some of the iron spheres may be droplets of iron meteorites, but the most common ones appear to be droplets that separated from the stony spheres during atmospheric melting.

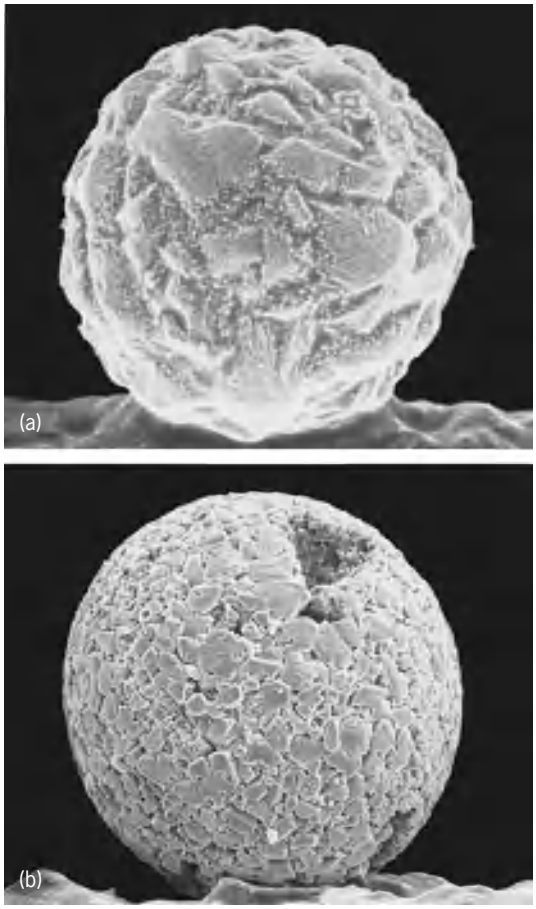

Scanning electron photomicrographs of stony cosmic spherules about 300 μm in diameter collected from the mid-Pacific ocean floor. (*a*) Sphere with so-called turtleback texture, indicative of rapid cooling from a very hot molten droplet. (*b*) Sphere with porphyritic texture, indicative of formation of a particle that was not so strongly heated during hypervelocity entry into the atmosphere as the sphere in *a*.

Identification. Cosmic spherules are of particular scientific interest because they provide information about the composition of comets and asteroids and also because they can be used as tracers to identify debris resulting from the impact of large extraterrestrial objects. The ablation spherules can be positively identified because of several unique properties, including distinctive oxygen isotope compositions and the presence of isotopes that are produced by cosmic rays such as aluminum-26 and manganese-53. Some spheres did not melt entirely, and they also retain high concentrations of noble gases implanted by the solar wind. In general, cosmic spherules can be confidently identified on the basis of their elemental and mineralogical compositions, which are radically different from nearly all spherical particles of terrestrial origin.

Don E. Brownlee, II

Bibliography. D. E. Brownlee, Cosmic dust: Collection and research, *Annu. Rev. Earth Planet. Sci.,* 13:147–173, 1985; C. Emiliani (ed.), *The Sea,* 1981; B. P. Glass et al., North American microtekitites from the Caribbean Sea and their fission track ages, *Earth*

Planet Sci. Lett., 19:184–192, 1973; P. W. Hodge, *Interplanetary Dust,* 1981; J. A. M. McDonnel, *Cosmic Dust,* 1978.

Cosmic strings

Hypothetical thin tubes that carry very large amounts of vacuum energy. They belong to the class of objects known as topological defects. As the universe expands, its temperature decreases and its microscopic properties, such as the strength with which fundamental particles interact, sometime change abruptly in what are called phase transitions. A well-known example of such a transition is when water freezes into ice when the temperature drops below 0°C (32°F). Strings are predicted to form at these transitions. In this context, the lowest energy state of the universe, which is called the "vacuum" although it is not necessarily empty, can have a nontrivial structure in space, and strings appear as defects in this structure. *See* DARK ENERGY; ENERGY LEVEL (QUANTUM MECHANICS); PHASE TRANSITIONS.

Two possible kinds of strings are predicted, namely, local and global strings. The energy of local strings is concentrated within a very short distance of the defect tube, called the core. As a result, local strings have only contact interactions. They are predicted in almost all cosmologically realistic grand unified theories describing all interactions as a single one at high energies. The energy density of global strings, on the other hand, is not well localized along the core, so these strings have long-range interactions. Global strings are a consequence of many models based on superstring theories; they are often called axionic strings because they contain hypothetical particles, called axions, which appear naturally in the low-energy limit of these theories. In practice, both kinds of strings could exist. *See* DARK MATTER; GRAND UNIFICATION THEORIES; SUPERSTRING THEORY.

Formation. Phase transitions are believed to have taken place in the early universe, namely, immediately after the big bang, when the temperature was much higher than it is now. As such a phase transition proceeded in the early universe, regions having different microscopic properties grew independently. When their boundaries met, there could be mismatches in these properties, resulting in defects such as cosmic strings. The process was very similar to a phase transition that is observed in liquid crystals: At high temperature, molecules in such a material move rapidly and their relative orientations are random. When the temperature is lowered below a critical value, the forces between the molecules, which tend to cause them to be aligned, can overcome the thermal motion, and in any given region of space all the molecules end up pointing in the same direction. In uncorrelated locations of the crystal sample, those directions will be different, so that at various interfaces between such regions the molecules cannot have a preferred direction. The original, randomly distributed state persists in a region near such an interface; such a region is called a defect. The geometry of the defects is fixed by the original properties of the system. In the universe, although other defects, such as monopoles, domain walls, and textures, could in principle be formed, they are in practice ruled out as they would induce cosmological consequences that are not observed. *See* CRYSTAL DEFECTS; LIQUID CRYSTALS; SYMMETRY BREAKING.

Strings exist for topological reasons. Therefore, they are stable and cannot have ends, at least in simple models. In more sophisticated theories, strings can be terminated by monopoles or by connecting domain walls, but these theories are usually incompatible with cosmological observations. Therefore, strings can be seen in the form of either closed loops or infinite strings. The latter are often assumed to be, in fact, loops whose dimension is larger than the size of the observable universe, called the Hubble radius. *See* BIG BANG THEORY.

Evolution of string networks. Once a network of local strings has been formed at a phase transition, its cosmological consequences depend on its subsequent evolution. Initially, in what is called the friction-dominated phase, the network interacts very strongly with its environment, namely the very dense primeval plasma. The expansion of the universe makes the plasma density decrease, until the interactions between the strings and the particles become negligible. At this point, another effect comes into play: Whenever two pieces of strings cross each other, they have a large probability (set to unity in actual numerical simulations) to exchange extremities. This contact interaction is called intercommutation or reconnection. Long strings and loops then tend to fragment and produce large numbers of small loops. In turns, these loops lose their energy in gravitational radiation. This energy loss mechanism is compensated by the effect of the expansion, which pulls more and more strings within a Hubble volume as the Hubble radius, namely the size of the observable universe, increases. *See* GRAVITATIONAL RADIATION.

After some relaxation time, the string network eventually reaches a special configuration, known as a scaling solution: Snapshots of the string distribution (see **illustration**) taken at two different times would have similar statistical properties, with characteristic lengths increased because of expansion. The observable universe thus contains at any time an almost fixed number of long strings, somewhere between 10 and 100, separated by a typical distance of the order of the Hubble radius divided by the number of strings. Nowadays, that would be between 10^8 and 10^9 light-years. Thanks to the scaling, the density contained in the string network relative to the overall density of the universe is constant in time; it is evaluated to be of the order of 10^{-5}.

Observational consequences. The weak, strong, and electromagnetic forces appear to be very different at the low energies at which particle physics experiments can be performed. However, it is believed that they are in fact different faces of a single interaction, much in the same way as electric and magnetic effects are merged, at accelerator energies, into the

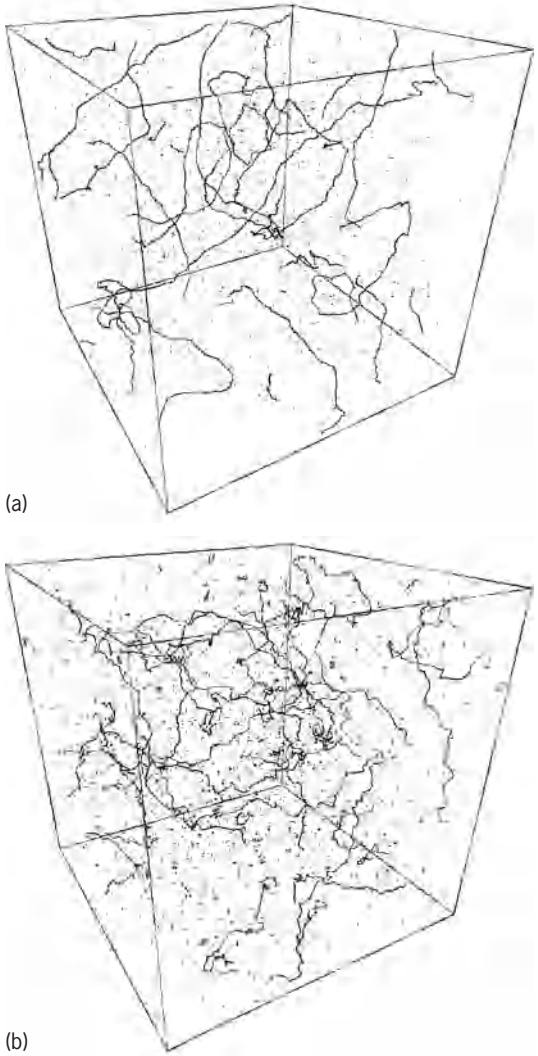

Snapshots of the simulated string distribution, that is, the string network, at two different times. (*a*) Radiation-dominated epoch. (*b*) Matter-dominated era. The size of the simulation box is that of the visible universe. (*C. Ringeval, after D. Bennett and F. Bouchet, High-resolution simulations of cosmic-string evolution: i. Network evolution, Phys. Rev. D, 41:2408–2433, 1990*)

electromagnetic interaction. Above a given energy scale, known as the grand unified scale, the differences among these three interactions disappear. If strings are formed at this grand unified scale, then their energy per unit length is of the order of the equivalent of 10^{15} tons per centimeter, concentrated in a radius of 10^{-29} cm. *See* FUNDAMENTAL INTERACTIONS.

Strings have a tension that is equal to the mass density times the velocity of light squared. As for any kind of string, transverse disturbances travel along strings at a velocity that is given by the square root of the tension divided by the mass per unit length. For a cosmic string, this velocity is therefore the speed of light, and strings are therefore very relativistic objects. Strings also move with respect to the cosmological frame (the frame in which the cosmic microwave background is at rest) with velocities comparable to that of light. As a consequence, the gravitational effects they induce are special to them, leading to specific signatures. In particular, although strings are extremely massive, the gravitational field around a string vanishes. This vanishing means that an observer at rest with respect to a string will experience no gravity. Test particles can, however, feel a gravitational disturbance provided that they move with respect to the string. Because of this disturbance, two particles are instantly projected toward one another at relativistic velocities whenever a string passes between them. This phenomenon was proposed in 1980 as a mechanism capable of seeding large-scale structure formation, namely galaxies, clusters of galaxies, and so on. Nowadays, it is believed that this mechanism is not the chief reason that the primordial featureless soup ended up with all the structures that are now observed, the basic seeds having resulted from another mechanism such as inflation. *See* INFLATIONARY UNIVERSE COSMOLOGY.

Light can be deflected by strings. This deflection is due to the fact that, if one circles around a string, the angle measured falls short of 2π by a very small but nonzero value called the missing angle. The first consequence of this deflection is that moving strings induce a Doppler-like shift in the frequency of a beam of light passing by; this phenomenon is called the Kaiser-Stebbins effect. Such shifts, being independent of the photon wavelength, are seen as temperature fluctuations in the cosmic microwave background radiation, superimposed over the primordial ones. Present studies indicate that these secondary fluctuations cannot contribute at a level higher than 10% of the total fluctuations. The remaining 90% were most probably produced during an inflationary epoch in the early universe and therefore did not result from cosmic strings. *See* COSMIC BACKGROUND RADIATION.

A light ray passing near a string would travel along a trajectory that takes into account the missing angle, and would be bent by half this amount. As a result, if a string is present between an observer and a point source, the source will be lensed and thus seen as two absolutely identical objects by the observer. Unambiguously observing such a double image would be a proof of the existence of cosmic strings. Obtaining such an observation is, however, a difficult task as even the most massive strings produce a missing angle of only a few seconds of arc. Moreover, actual pairs of almost identical objects are very common in the sky, so that, as yet, it has never been possible to distinguish between a coincidence and a lensing effect. Good candidates are currently being examined and the existence of cosmic strings might thus shortly be proven. *See* GRAVITATIONAL LENS.

Superconducting strings. The strings can be coupled to many particles, and in some cases the trapping of those particles in the string cores is energetically favored. In such cases, the trapped particles behave as current carriers, and the strings become superconducting, with currents up to 10^{20} ampères. It has been proposed that such high currents could be the sources for ultrahigh-energy cosmic rays. Current-carrying loops of strings can form

equilibrium states (named vortons), which would not decay into gravitational radiation as do their non-current-carrying counterparts. The energy density they contain would then not be transferred into the radiation fluid, a condition necessary to reach scaling. Vortons would then come to dominate the universe and lead it to a premature collapse, a big crunch. Such scenarios lead to very strong constraints on the theories that predict them. *See* COSMIC RAYS; SUPERCONDUCTIVITY.

Cosmic superstrings. At energies higher than the grand unified scale, particle properties are expected to be described by superstring theory. This theory is based on the idea that the fundamental building blocks of nature are not pointlike particles. Rather, they are extended, one-dimensional objects, also called strings, which are so small that they appear to be pointlike in all conceivable experiments. However, under some conditions, some strings (not to be confused with the axionic strings, also produced by the same theories) could be of cosmological dimensions. In this case, their properties are expected to be similar to those of the local strings described above. The main difference is that the reconnection probability could be much lower, resulting in a network that could have difficulties reaching scaling, and in potentially observable cosmological consequences, currently under investigation. It is hoped that such consequences will perhaps provide the first testable prediction of superstring theory. Patrick Peter

Bibliography. A. Gangui, Cosmology from topological defects, in M. Novello and S. Perez-Bergliaffa (eds.), *Cosmology and Gravitation: Xth Brazilian School of Cosmology and Gravitation*, AIP Press, 2003 (http://www.iafe.uba.ar/relatividad/gangui/xescola/); A. Gangui, Superconducting cosmic strings, *Amer. Scientist*, 88(3):254–263, May-June 2000 (http://www.iafe.uba.ar/relatividad/gangui/SCS.html); P. Peter and A. Gangui, *Des défauts dans l'Univers*, CNRS Éditions, 2003 (in french); P. Peter and J. P. Uzan, *Primordial Cosmology*, Oxford University Press, 2007; A. Vilenkin and E. P. S. Shellard, *Cosmic Strings and other Topological Defects*, Cambridge University Press, 1994.

Cosmochemistry

The science of the chemistry of the universe, particularly that beyond the Earth. As currently practiced, cosmochemistry is concerned primarily with inferences on pre-solar-system events, solar nebular processes, and early planetary processes as deduced from minerals in meteorites and from chemical and isotopic compositions of meteorites and their parts. *See* ISOTOPE; METEORITE.

Minerals in Meteorites

Meteorites provide a great deal of otherwise unobtainable information about the formation and early history of the solar system. The solar system formed from the solar nebula, a large cloud of gas and dust. The chemical composition of the solar system, which can be assumed to be the same as that of the solar nebula, is known in detail from spectroscopic analysis of the Sun, which contains 99.9% of its mass. From a combination of this knowledge, an estimate of temperature and pressure from hydrodynamic models, and thermodynamic data for all possible gas, liquid, and solid species, it can be determined which minerals would have become stable as the gas cloud underwent gravitational collapse, spun down into a disk, and cooled. Pressure in the solar nebula was probably so low that condensation occurred at temperatures below the melting points of the minerals, and they condensed as solids. Different minerals became stable at different temperatures, giving rise to a condensation sequence, which can be calculated (**Table 1**). *See* ASTRONOMICAL SPECTROSCOPY; ELEMENTS, COSMIC ABUNDANCE OF; HYDRODYNAMICS; PROTOSTAR; SOLAR SYSTEM; SUN; THERMODYNAMIC PRINCIPLES.

Detailed chemical information, however, requires samples, and the Earth is too geologically active to provide samples dating to its formation. Meteorites are therefore studied because, in general, they have undergone much less modification than terrestrial rocks.

Carbonaceous chondrite inclusions. A certain type of meteorite, the carbonaceous chondrite, contains millimeter-sized rocklets, or refractory inclusions, that are composed of the very minerals (especially hibonite, perovskite, melilite, spinel, and fassaite; Table 1) believed to have been among the first to condense from the solar nebula. This finding indicates that assumptions used to calculate the condensation sequence are reasonably accurate. Corundum, the first major condensate, is very rare, however. Most of it probably reacted with the nebular gas to form hibonite, or it may have entered clumps of minerals that melted and crystallized other phases upon cooling. The absence of forsterite (FeO-free olivine; Table 1) and other lower-temperature phases from

TABLE 1. Stability ranges of major high-temperature equilibrium solar nebula condensates*

Phase	Formula	Temperature limits, K† Upper‡	Lower§
Corundum	Al_2O_3	1770	1726
Hibonite	$CaAl_{12}O_{19}$	1728	1686
Grossite	$CaAl_4O_7$	1698	1594
Perovskite	$CaTiO_3$	1688	1458
Melilite	$Ca_2Al_2SiO_7$-$Ca_2MgSi_2O_7$	1580	1434
Spinel	$MgAl_2O_4$	1488	1400
Metallic nickel-iron	(Fe, Ni)	1462	
Fassaite	$Ca(Mg, Ti, Al)(Al, Si)_2O_6$	1458	
Forsterite	Mg_2SiO_4	1444	
Anorthite	$CaAl_2Si_2O_8$	1406	
Enstatite	$MgSiO_3$	1366	

*Total pressure = 10^{-3} atm = 100 Pa.
†°F = (K × 1.8) − 459.67.
‡For condensation with decreasing temperature or disappearance with rising temperature.
§For conversion from or to forms stable at lower temperatures.
SOURCE: D. Ebel and L. Grossman, Condensation in dust-enriched systems, *Geochimica et Cosmochimica Acta*, 64:339-366, 2000.

most of these refractory inclusions indicates that they were somehow isolated from the nebular gas at very high temperatures. *See* CORUNDUM; MELILITE; PEROVSKITE; SPINEL.

The most abundant refractory inclusions that are thought to be condensates, found in the C3 group of carbonaceous chondrites, are the so-called fluffy type-A inclusions. These inclusions are mainly composed of melilite, with lesser amounts of spinel, perovskite, and hibonite. Most have very irregular shapes and appear to be aggregates of individually formed nodules. Many contain melilite crystals in which aluminum oxide (Al_2O_3), rather than magnesium oxide (MgO), increases from core to rim. This trend is opposite to that expected for crystallization from a liquid, but it could be produced by condensation from the nebular gas at about 1460 K (2170°F) during a period of falling pressure.

In contrast, type-B inclusions, also found in C3s, are round and contain spinel, melilite, fassaite, and anorthite. Results of experiments in which artificial refractory inclusions were crystallized from liquids at known temperatures and cooling rates suggest that these type B inclusions crystallized from liquids that formed by incomplete melting of preexisting solids, and cooled at rates ranging from a few tenths of a degree to tens of degrees per hour from temperatures of about 1400°C (2550°F). *See* CRYSTALLIZATION.

Unlike terrestrial rocks, most meteorites contain metallic iron-nickel alloys, indicative of formation in a reducing (oxygen-poor) environment. Other minerals within refractory inclusions can be used to estimate the oxygen abundance. On Earth, titanium is most commonly found as Ti^{4+}, but in refractory inclusions there are some minerals (for example, hibonite, fassaite) which also contain titanium in its more reduced form, Ti^{3+}. The Ti^{3+}/Ti^{4+} ratios in these minerals in many cases are quantitatively consistent with their formation in the highly reducing environment of a gas of solar nebular composition.

Formation of chondrites. Condensation probably continued at lower temperatures, after isolation of the refractory inclusions from the nebular gas. The C2 carbonaceous chondrites contain single crystals of forsterite and enstatite (Table 1) that are believed to be condensates, and the C3 chondrites contain ameboid olivine aggregates, which are irregularly shaped clumps of grains that also appear to be condensates. Based on the olivine compositions (FeO-bearing) and the presence of nepheline and sodalite, which contain sodium oxide (Na_2O) or potassium oxide (K_2O), these objects continued to react with the solar nebular vapor to temperatures as low as about 500 K (440°F). There is much olivine in the chondritic meteorites, but most of it is present in objects termed chondrules, which are millimeter-sized spheres of crystals and glass that once were molten droplets. Chondrules are found in carbonaceous and ordinary chondrite meteorites, the latter being aptly named since they are the most abundant type of meteorite. Chondrule formation was thus a major, widespread process in the early solar system. As is the case for many refractory inclusions, chondrules are thought to have crystallized from liquids formed by melting of preexisting solids, but the nature of the heat source that caused the melting is not well understood. *See* ENSTATITE; NEPHELINE; OLIVINE; PYROXENE.

In the C3 carbonaceous chondrites, the refractory inclusions and the chondrules are enclosed in a later-formed, fine-grained matrix dominated by FeO-rich olivine, plus Fe-sulfides and iron-bearing compounds. The matrix must have formed at a much lower temperature than the chondrules and refractory inclusions, because FeO is not sufficiently stable to enter silicates significantly above about 900 K (1160°F) and the carbonaceous materials require low temperatures in order to be stable as well. Thus, diverse components—condensate inclusions, melted inclusions, and chondrules—with different histories were all assembled without much reheating into objects that became carbonaceous chondrites.

In addition, components were efficiently sorted, giving rise to different chondrite groups. The C1 carbonaceous chondrites contain no chondrules or inclusions, just single olivine and pyroxene crystals in a fine-grained matrix of sheet silicates (a water-bearing mineral, similar to the clay minerals). These are the most primitive meteorites, containing solar proportions of all but the most volatile elements. The C2 carbonaceous chondrites contain small refractory inclusions that consist mostly of nonsilicate, oxide minerals (hibonite, perovskite, and spinel), and olivine-rich chondrules in a matrix dominated by sheet silicates. Many of the chondrules and inclusions were altered by some sort of FeO-water-rich fluid. This process may have occurred on the parent body. The C3 carbonaceous chondrites have more abundant chondrules and larger, more silicate-rich refractory inclusions than the C2 chondrites, in an anhydrous, olivine-rich matrix.

Ordinary chondrites consist mainly of chondrules, chondrule fragments, and very rare refractory inclusions in a very small amount of matrix. Many have been heated and have very uniform olivine and pyroxene compositions, especially with respect to their Mg/Fe ratios; these are the equilibrated ordinary chondrites. Others have not been heated and are known as unequilibrated ordinary chondrites. All chondrites contain FeO-bearing olivine, and determining how olivine with such compositions, especially that in ordinary chondrites and carbonaceous chondrite matrices, could have formed in the solar nebula is one of the major unsolved problems in the study of meteorites. Calculations show that, at the (low) temperature at which FeO becomes stable enough to enter olivine in a gas of solar composition, its rate of diffusion into grains would be so slow that negligible amounts, much lower than are observed, would enter olivine before diffusion stops altogether. The stability temperature of FeO would be higher, however, allowing more extensive diffusion, in a more oxygen-rich gas. One way this could have occurred would be if oxygen-bearing dust were added to the nebula, increasing the dust/gas ratio over the solar ratio. But only by using the

maximum dust/gas ratio permitted by nebular models, with oxygen-rich, water ice-bearing dust, and olivine grains with radii of approximately 1 micrometer, do condensation models predict the formation of olivine with FeO contents within the range observed in ordinary chondrites. This set of requirements is not a satisfactory solution to the problem of accounting for the observed FeO contents of olivine in chondrites.

Some meteorites also contain trace amounts of grains of carbon-rich minerals (silicon carbide, diamond, and graphite) whose condensation requires a gas with a carbon/oxygen ratio greater than 1, whereas that of the Sun is 0.5. This finding, along with very large isotopic anomalies for many elements, provides strong evidence that these grains are interstellar in origin, predating the Sun by as much as about 2.4×10^9 years. These grains should have disappeared by reacting with the solar nebular gas, however; why they did not and how they survived long enough to be incorporated into meteorites are not yet understood. *See* DIAMOND; GRAPHITE; INTERSTELLAR MATTER; SILICON.

Meteorites from planets. The presence of a wide variety of objects in chondrites—interstellar grains, refractory inclusions, chondrules, and matrix— which date back to the formation of the solar system and have survived virtually unchanged, shows that these meteorites were generally not subjected to destructive geologic processes on their parent bodies. This fact indicates that those bodies were small and inactive. Some objects, presumably with chondritic bulk compositions, did grow large enough to become geologically active. In such bodies, heat from gravitational compression and radioactive decay caused interior melting. The first liquid produced by melting of chondritic rock consists mostly of metallic nickel-iron and sulfur. Since this liquid is twice as dense as silicates, it will sink, forming an iron-rich metallic core.

Iron meteorites are completely composed of metallic nickel-iron that was once molten. In some cases, as they cooled, the nickel preferentially entered one phase and the iron entered another, forming alternating bands in a process called exsolution. From the bulk composition of a meteorite and the iron-nickel phase diagram, the starting temperature for exsolution can be determined, and from the thicknesses and compositions of the bands, how far it progressed. Then, application of experimentally determined diffusion rates yields the cooling rate. Results for various iron meteorites fall between 1 and 100°C (2 and 180°F) per million years, evidence for origins deep within planets with radii between 100 and 300 km (60 and 180 mi). At least a dozen parent bodies appear to be represented. The pallasite meteorites, with coarse olivine crystals enclosed in a matrix of metallic nickel-iron, probably come from core-mantle boundaries of former planets.

Stony (silicate) meteorites called achondrites, which do not contain chondrules and do not have chondritic (near-solar) compositions, are from the silicate portions of planets that underwent internal melting. Just above the core-mantle boundary of such planets should be metal-poor zones of coarse-grained rocks that formed by settling of large crystals. Ureilites (olivine-rich meteorites) and diogenites (pyroxene-rich meteorites) probably formed in such zones.

The eucrites, another group of achondrites, are essentially basaltic rocks, and microscopic study shows that many of them consist of rock fragments that were broken up by meteorite impacts on their parent body. This finding indicates that they formed at or near the surface, as do terrestrial basalts. Other meteorites, termed howardites, appear to be lithified soils, and consist of rock, mineral, and glass fragments in a fine-grained matrix, as do lunar soils. In howardites the rock fragments are mainly diogenites and eucrites, showing that these meteorite types originated on the same parent body.

The irons and achondrites provide evidence of internal melting in small planets. Radiometric dating indicates that these meteorites are about as old as the solar system itself. Therefore, planet formation took place on a fairly short time scale, with bodies becoming large enough to undergo internal melting soon after the formation of the solar system. *See* MINERAL; PLANETARY PHYSICS.

A group of relatively young meteorites is also very important—the SNC (shergottite-nakhlite-chassignite) group. These are crystalline rocks with ages of less than 1.3 billion years, much younger than other igneous meteorites (4.5 billion years) or lunar samples (the youngest returned samples are about 3.1 billion years old; some unsampled units may be younger). The relatively young ages of the SNCs indicate that they had a fairly large parent body, because larger planets lose their internal heat more slowly and remain geologically active longer than small planets. Because all of the known asteroids are too small to have produced 1.3-billion-year-old internal melts, Mars was proposed as the source of the SNCs. This idea has become widely accepted. Strong evidence that Mars is the source of the SNCs, in addition to their young ages, comes from the rovers *Spirit* and *Opportunity*, which have analyzed rocks with compositions like those of basaltic shergottites. Additional evidence includes similarities between the compositions of shergottites and that of the soil at the *Viking*, *Spirit*, and *Opportunity* landing sites, and a close match between the composition of trapped gas in the SNCs and that of the Martian atmosphere as measured by *Viking* spacecraft. For some years an argument against Mars as the source of the SNCs was the absence of lunar meteorites from collections. But in 1983 the discovery of the first lunar meteorite was confirmed, on the basis of detailed comparison with analyses of lunar samples returned by the *Apollo* astronauts. *See* MARS; SPACE PROBE. Steven B. Simon

Chemical Compositions of Meteorites

The chemical composition of meteorites and independently formed objects within them have taught much about pre-solar-system events, solar nebular processes, and early planetary processes.

Cosmochemical properties of elements. Before discussing examples of the results of meteorite studies, it is necessary to understand a few basic properties of the elements.

Cosmochemical affinity. The elements can be divided into four groups based on their cosmochemical affinity: siderophile or metal-seeking elements; chalcophile or sulfide-seeking elements; lithophile or oxide and silicate-seeking elements; and atmophile elements, which do not condense and are found in atmospheres (**Fig. 1**). This classification of elements applies to the highly reducing conditions of the solar nebula, which is dominated by hydrogen. Conditions on the surface of the Earth and other planets are more oxidizing, so that some elements classified here as siderophile or chalcophile, such as lead, have lithophile characteristics in the Earth's crust and mantle. *See* ELEMENTS, GEOCHEMICAL DISTRIBUTION OF.

Volatility. The elements show a wide range of volatility in a gas of solar composition (**Fig. 2**). The refractory elements are generally considered to be those that are totally condensed in about the first 5% of the total condensable matter of the solar system, corresponding to about 1475 K (2195°F) in a gas of solar composition at a total pressure of 10^{-3} atmosphere (10^2 pascals). A subset of the refractory trace elements are sometimes referred to as the ultrarefractory elements, because they are already completely condensed before the first major mineral, hibonite, condenses at 1743 K (2677°F). The moderately volatile elements show variable degrees of depletion in chondrites and planets compared to cosmic abundances; they have condensation temperatures between 1475 and 700 K (2195 and 800°F). The highly volatile elements are strongly depleted in chondrites and planets and have condensation temperatures below 500 K (440°F). The solar nebula is highly reducing, because of the high abundance of hydrogen compared to other elements. The relative volatilities of some elements change under more oxidizing conditions. For example, under oxidizing conditions cerium becomes a moderately volatile element, whereas all remaining rare-earth elements remain refractory.

Compatibility. Some lithophile elements have large ionic radii and do not fit readily into the crystal structures of the major minerals that make up most rocks: olivine, pyroxene, and plagioclase. These elements are often referred to as incompatible elements. They tend to be enriched in the melt when a rock is partially melted. Examples of incompatible elements are potassium, rubidium, light rare-earth elements, and uranium.

Presolar events. As mentioned above, a number of types of interstellar grains have been separated from primitive meteorites. Presolar grains identified include silicon carbide, graphite, diamond, titanium carbide, zirconium carbide, molybdenum carbide, ruthenium carbide, metallic iron, aluminum oxide, titanium oxide, hibonite, and spinel. While chemistry alone has not unambiguously identified any of these grains as presolar, chemistry does help to understand the formation conditions of objects whose isotopic composition proves a presolar origin. A good example is silicon carbide, most of which is believed, on the basis of isotopic arguments, to come from the circumstellar envelopes of asymptotic

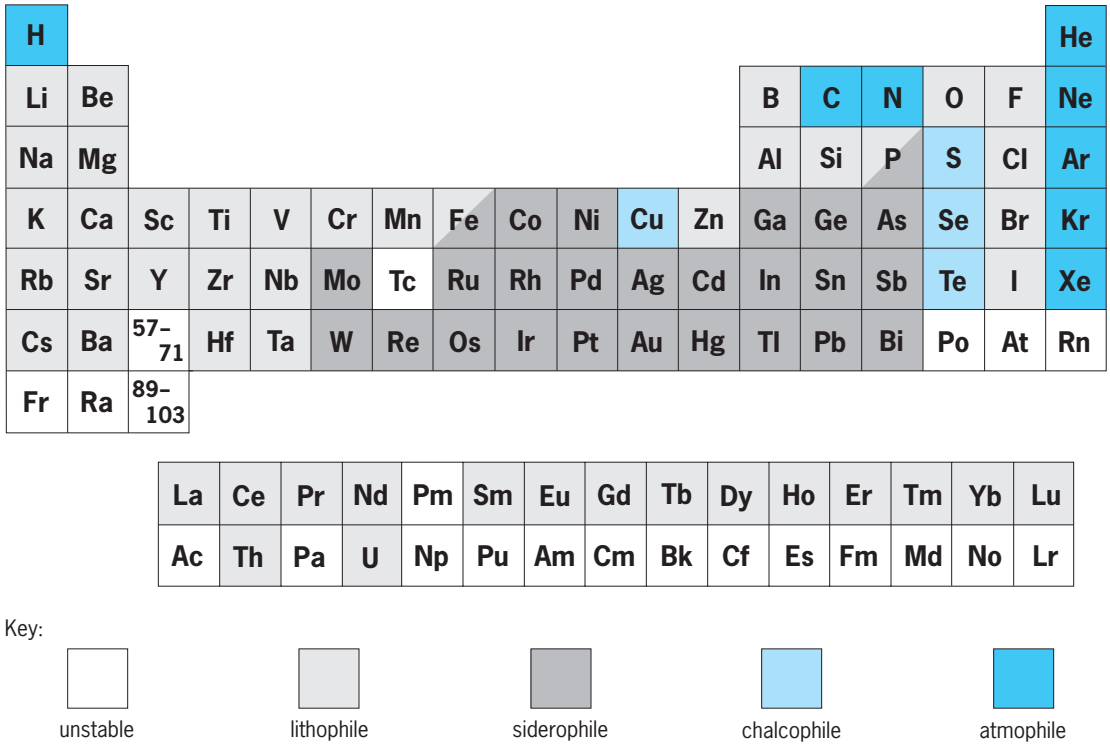

Fig. 1. Periodic table of the elements, showing their cosmochemical affinity under the highly reducing conditions found in the solar nebula.

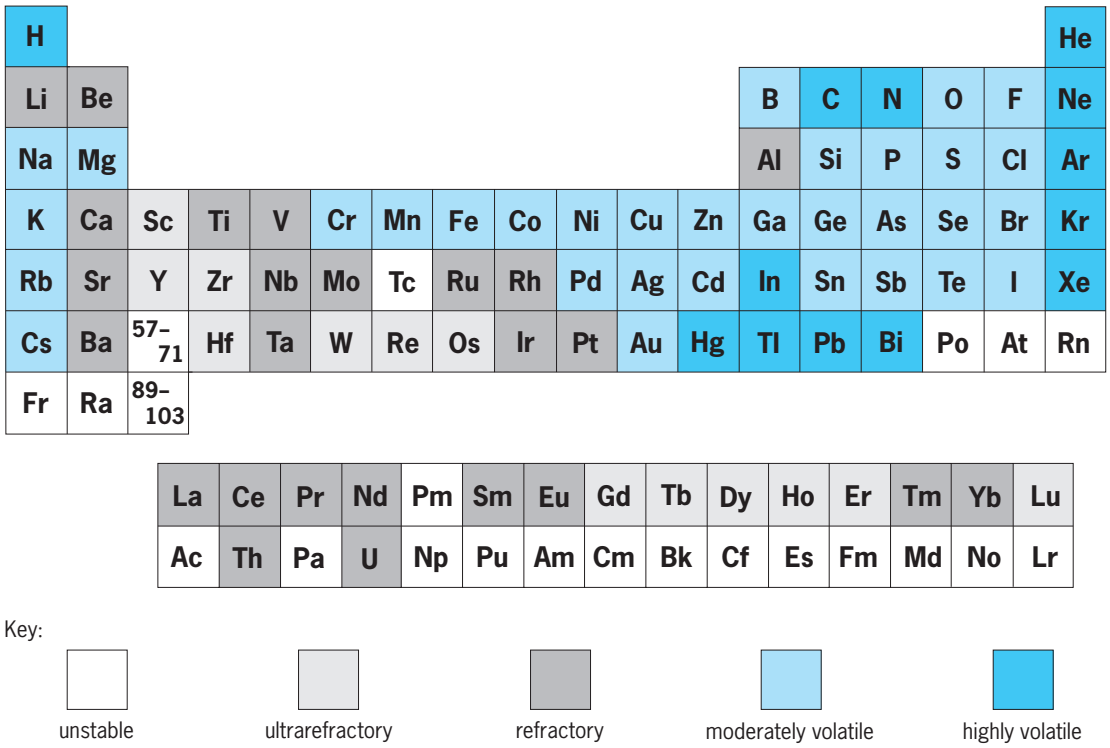

Fig. 2. Periodic table of the elements, with the elements classified according to their volatility in a gas of solar composition.

giant branch stars. The silicon carbide with isotope anomalies that is separated from meteorites is enriched in a number of trace elements that are predicted to condense at high temperatures as carbides in reducing environments, such as zirconium, titanium, molybdenum, strontium, barium, and the rare-earth elements. These enrichments are entirely consistent with the proposed origin of silicon carbide by high-temperature condensation in the circumstellar envelopes of carbon-rich stars. *See* CARBON STAR; GIANT STAR.

Solar nebular processes. Chemistry has been used to infer a number of important processes in the solar nebula. These include high-temperature condensation and vaporization, melting and crystallization of millimeter- to centimeter-sized objects in the solar nebula, separation of metal from silicate grains, and reaction of condensed grains with solar nebular gas. Investigation of these processes helps determine important properties of the solar nebula, such as its temperature, how long it lasted, and what the planets, which formed from the nebula, are made of.

Condensation and evaporation. As mentioned above, most carbonaceous chondrites contain refractory inclusions, which are enriched in refractory elements relative to cosmic abundances. The most common type is enriched by about 20 times relative to cosmic abundances in all refractory elements, regardless of whether they are lithophile or siderophile. Some other refractory inclusions show evidence for volatility fractionation of refractory elements from one another. This fractionation requires higher temperatures and is likely to have occurred within a few degrees of condensation of hibonite, above 1700 K (2600°F). Thermodynamic calculations done to model this fractionation assume equilibrium, so they cannot distinguish between volatilization and condensation.

Kinetically controlled processes can lead to isotopic and chemical fractionation effects that would not have occurred under equilibrium. One example comes from the refractory inclusions. The bulk chemical compositions of refractory inclusions are depleted in magnesium and silicon compared with the calculated bulk composition of material condensing from a gas of solar composition. These inclusions also display isotopic mass fractionation effects, in which they are enriched in heavy isotopes of magnesium and silicon relative to normal solar system material. Both effects have been reproduced in the laboratory by melting and evaporating synthetic refractory inclusions in vacuum or in low-pressure hydrogen. It now seems likely that most refractory inclusions have experienced heating and evaporation in the solar nebula under kinetically controlled conditions. In contrast, chondrules in chondritic meteorites generally do not display isotopic fractionation effects and are likely to have been heated under conditions that were closer to equilibrium.

Depletion of volatile elements. Relative to cosmic abundances, the terrestrial planets and many types of meteorites are depleted in the moderately volatile elements, with the depletions growing larger with increasing volatility. A useful element pair for seeing this fractionation comprises potassium, a moderately volatile element, and uranium, a refractory element. Both are incompatible elements. Within a

single planet or meteorite parent body, the potassium/uranium ratio is relatively constant, because igneous fractionation processes tend to exclude both elements from major minerals. Both have long-lived radioactive isotopes that are gamma-ray emitters, so measurements are available from Mars and Venus. Chondritic meteorites have potassium/uranium ratios near the cosmic value of about 65,000, whereas the Earth, Moon, Venus, Mars, and the parent body of the basaltic achondrites [probably the 500-km-diameter (300-mi) asteroid Vesta] have lower potassium/uranium ratios. The potassium depletion in the terrestrial planets is strong evidence for a nebula-wide event that swept away the volatile elements before they could condense.

Planetary processes. A number of important planetary processes, such as separation and differentiation of a metallic core and differentiation of the silicate portion of a planet into mantle and crust, can be studied through the chemical compositions of meteorites.

Core formation. The siderophile elements are powerful recorders of metal-silicate fractionation processes in parent bodies, that is, core formation. Some elements, such as tungsten, are strongly but not completely partitioned into metal during core formation. In igneous differentiation processes, the tungsten that remains behind in the silicate is highly incompatible, such that its ratio to other incompatible but strongly lithophile elements, such as lanthanum, remains constant. In each major planetary body, such as the Earth, Moon, the parent body of the igneous SNC (shergottite-nakhlite-chassignite) meteorites (probably Mars, as discussed before), and the basaltic achondrite parent body, the ratio of tungsten to lanthanum remains constant, although the absolute concentrations can be quite variable among different samples. There are differences from one parent body to the next in the tungsten/lanthanum ratio, which are due to differences in how metallic cores separated from the silicate mantle of each parent body.

Core differentiation. When the core of a planet forms, it is initially molten (otherwise the mechanical properties of metal and silicate would not have allowed it to separate). The process of crystallization of cores of now-disrupted small parent bodies can be followed through the chemical compositions of some classes of iron meteorites. About 800 iron meteorites are known. About 80% of them can be classified into 12 groups, based primarily on the concentrations of four diagnostic elements: nickel, gallium, germanium, and iridium. During crystallization of an initially molten core, some elements, such as gold, tend to partition into solid metal while others, such as iridium, tend to remain in the liquid. Several of the 12 groups of iron meteorites show correlations between siderophile trace elements that are consistent with slow crystallization of planetary cores.

Asteroid melting. The chondritic meteorites, which have textures indicating that they never melted, show variable depletions of moderately volatile elements, but these depletions are independent of whether a trace element is incompatible or not. In contrast, the basaltic achondrites are composed of basalt, which is a rock that can form by partial melting of a chondritic source composition. The eucrites, which are igneous basaltic achondrites, show a range of trace-element concentrations that indicate that they crystallized from a melt enriched in incompatible elements compared to cosmic abundances. This range of trace-element concentrations is due to fractional crystallization, where, compared to the melt, the first eucrites to crystallize are depleted in incompatible elements and the last eucrites are enriched in incompatible elements. *See* ASTEROID.

Sampling the solar system. Much of what is known about the chemistry of the solar system comes from meteorites, which are believed to have sampled asteroids, Mars, and the Moon. Now, more that 30 years after the Apollo missions to return lunar samples, new missions are sampling the Sun, a comet, and an asteroid, and further sample-return missions are in the planning stages.

Genesis. The first of the new generation of sample-return missions, *Genesis*, exposed high-purity materials such as silicon, aluminum oxide, and diamond film to the solar wind for over 2 years. The spacecraft crashed in Utah when its parachute failed to open, but the solar wind collectors survived. Analyses will measure the chemical and isotopic compositions of many elements in the samples, allowing tests of how well spectroscopic methods measure solar composition and whether the Sun has the same isotopic composition as the Earth and meteorites.

Stardust. The *Stardust* mission flew close to the nucleus of comet Wild 2 in January 2004 and collected dust particles in silica aerogel. The spacecraft returned to Earth in January 2006. Dust grains as small as 1 μm across were extracted from the aerogel and analyzed for their crystal structure and chemical and isotopic composition. The mission should dramatically increase our knowledge of comets. It is currently unknown whether the dust in comets is formed mostly within the solar system or is an aggregation of presolar material that was never significantly modified during formation of the solar system.

Hayabusa. In early February 2006, the *Hayabusa* spacecraft was orbiting the asteroid Itokawa after having landed twice on the surface of the asteroid, with the purpose of collecting a sample to return to Earth. The spacecraft had a number of problems: it may not have collected a sample; it will not return to the Earth until at least 2010, three years later than previously planned; and scientists were attempting to resume communication, which was interrupted in December 2005. It will be of great interest to see how closely asteroids resemble the meteorites in the world's collections, and, regardless of the fate of *Hayabusa*, other robotic sample return missions to asteroids are being proposed and planned. *See* SPACE PROBE.

Andrew M. Davis

Isotopic Compositions of Meteorites

Studies of the relative abundances of isotopes in meteorites provide a unique source of information

about the age of the solar system, the time scales for formation of the first solid bodies in the solar system and the growth and evolution of small planets, the prehistory (nucleosynthesis) of the material out of which the solar system formed, and the interaction of solar and galactic cosmic rays with matter.

Chronology: age of solar system and formation of planets. Meteorites are the oldest objects available for scientists to study in the laboratory and provide the best indication of the age of the solar system. Nineteen radioactive decay schemes with half-lives ranging from a remarkably short 53 days to more than 100 billion years (**Table 2**) have been utilized to determine the ages of samples varying in size from individual mineral grains less than 1 μm across to whole meteorites. The radionuclide systems in Table 2 are generally divided into two classes according to the half-life of the radioactive parent isotope: short-lived radionuclides have half-lives of less than about 100 million years, while long-lived radionuclides have half-lives of greater than 500 million years. This distinction is based on the practical issue of observability. Radioactive species with half-lives shorter than about 100 million years are now extinct in the sense that whatever amount may have existed at the birth of the solar system has completely decayed away. For this reason conventional dating of terrestrial rocks is based solely on the long-lived radionuclides listed at the bottom of Table 2. Most meteorites, in contrast, formed sufficiently early to preserve evidence of the short-lived radionuclide parents, and chronologies based on these short-lived radionuclides provide a powerful tool for examining time scales during the first 50 million years of solar system history.

The oldest meteoritic samples are the refractory, calcium-aluminum-rich inclusions found in carbonaceous chondrites. The 4.567-billion-year age of these objects provides the best estimate for the condensation of the first solid material in the solar system. Chondrules, millimeter-size spherules of glass and crystals, began to form 1 to 3 million years after refractory inclusions as the solar nebula cooled. The oldest basaltic meteorites crystallized from molten lavas on small asteroids within roughly 3 million years following the formation of the carbon-aluminum-rich inclusions (that is, ages of 4.564–4.552 billion years). Most chondritic meteorites have less precisely determined ages encompassing the range covered by chondrules and basaltic achondrites. The tight clustering of ages of the oldest meteorites reflects a period of very rapid growth during the first 5 million years of solar system history, as millimeter-to-centimeter-size objects collided and coalesced to form meter-to-kilometer-size planetesimals. The younger ages of some meteorites indicate that disturbances due to shock and thermal metamorphism began as early as 4.55 billion years ago and continued for several hundred million years. Collisions between planetesimals led to coagulation and differentiation, the growth of larger bodies, and the eventual formation of the terrestrial planets, including the Earth-Moon system, within about 50 million years following the solar system's birth. *See* EARTH, AGE OF; ROCK AGE DETERMINATION.

One unusual group, the SNC meteorites, whose appearance and composition is similar to that of terrestrial basalts, is the exception to the rule with much younger crystallization ages of approximately 1.3 billion years. The young ages of the SNCs require an origin on a planetary body larger than the Moon, on which radioactive heating could sustain igneous activity for at least 3 billion years. It is now generally accepted that this group of meteorites originated on Mars, an argument strengthened by recent evidence that the isotopic compositions of hydrogen, nitrogen, and noble gases in the SNC meteorites and the Martian atmosphere are strikingly similar and distinct from those in the Earth and other meteorites.

The first short-lived radionuclide, ^{129}I, was discovered in the early 1960s by way of the observation that some meteorites have a greater abundance of daughter ^{129}Xe than do others. Twelve short-lived species with half-lives ranging from 53 days to 103 million years (Table 2) are now confirmed to have existed at the birth of the solar system. Suggestive evidence of three additional radionuclides has been reported but awaits confirmation. The evidence of short-lived radionuclides is distributed across all classes of meteorites and provides a compelling argument not only that meteorites are the oldest objects in the solar system but also that they formed rather quickly over a very short interval. Although these nuclides can no longer be observed directly, evidence that they were present during the earliest epoch of solar system history is preserved as variations in the relative abundances of the daughter isotopes of the extinct radionuclides. For example, large excesses of ^{26}Mg, produced by the decay of ^{26}Al (half-life of 0.72 million

TABLE 2. Radionuclides and solar system chronology

Nuclide*	Half-life, 10^9 years	Daughter	Abundance†
^{7}Be	1.5×10^{-10}	^{7}Li	^{7}Be/^{9}Be $\sim 6 \times 10^{-3}$
^{41}Ca	0.00010	^{41}K	^{41}Ca/^{40}Ca $\sim 1.4 \times 10^{-8}$
^{36}Cl	0.00030	^{36}S	^{36}Cl/^{35}Cl $\sim 2 \times 10^{-4}$
^{26}Al	0.00072	^{26}Mg	^{26}Al/^{27}Al $\sim 5 \times 10^{-5}$
^{10}Be	0.0015	^{10}B	^{10}Be/^{11}Be $\sim 1 \times 10^{-3}$
^{60}Fe	0.0015	^{60}Ni	^{60}Fe/^{56}Fe $\sim 1 \times 10^{-6}$
^{53}Mn	0.0037	^{53}Cr	^{53}Mn/^{55}Mn $\sim$ (0.9-4) $\times 10^{-5}$
^{107}Pd	0.0065	^{107}Ag	^{107}Pd/^{108}Pd $\sim 2 \times 10^{-5}$
^{182}Hf	0.009	^{182}W	^{182}Hf/^{180}Hf $\sim 1 \times 10^{-4}$
^{129}I	0.016	^{129}Xe	^{129}I/^{127}I $\sim 1 \times 10^{-4}$
^{244}Pu	0.081	Xe fission	^{244}Pu/^{238}U $\sim$ 0.007
^{146}Sm	0.103	^{142}Nd	^{146}Sm/^{142}Sm $\sim$ 0.008
^{235}U	0.704	^{207}Pb	^{207}Pb/^{204}Pb = 10.29
^{40}K	1.25	^{40}Ar	^{40}Ar/^{36}Ar = 295.5
^{238}U	4.47	^{206}Pb	^{206}Pb/^{204}Pb = 9.307
^{232}Th	13.9	^{208}Pb	^{208}Pb/^{204}Pb = 29.48
^{187}Re	41.6	^{187}Os	^{187}Os/^{188}Os = 0.0956
^{87}Rb	48.8	^{87}Sr	^{87}Sr/^{86}Sr = 0.6988
^{147}Sm	106	^{143}Nd	^{143}Nd/^{146}Nd = 0.5068

*Short-lived radionuclides, listed above in the top of the table, have half-lives shorter than 100 million years; long-lived radionuclides appear in the bottom half.

†Short-lived radionuclides: inferred initial abundance of the parent nuclide 4.566 billion years ago. If the initial abundance is not well established, the range of values is shown. Long-lived radionuclides: initial abundance of daughter isotope.

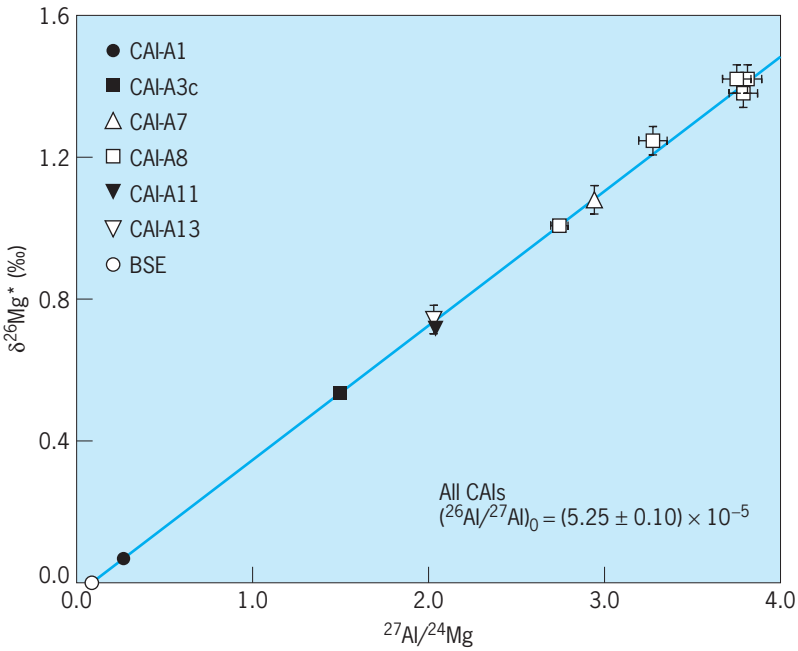

Fig. 3. Correlation of the $^{26}Mg/^{24}Mg$ ratio, expressed as $\delta^{26}Mg$ (the deviation in parts per thousand from normal Mg), with the $^{27}Al/^{24}Mg$ ratio for a group of calcium-aluminum-rich inclusions (CAIs) from the Allende meteorite. The correlation line is generated by the ingrowth of radiogenic ^{26}Mg, produced by the decay of ^{26}Al within the inclusions. The slope of the correlation line, corresponding to $^{26}Al/^{27}Al = 5.25 \times 10^{-5}$, indicates the abundance of the short-lived radionuclide ^{26}Al at the beginning of the solar system, 4.567 billion years ago. The point labeled BSE represents the Mg isotope composition of the silicate portion of the Earth. The fact that all six inclusions lie along a single line indicates that most CAIs formed at the same time, within ~50,000 years. (*After M. Bizzarro, J. A. Baker, and H. Haack, Nature, 431:275–278, 2004*)

years), are found in aluminum-rich, magnesium-poor minerals in refractory inclusions in carbonaceous chondrites (**Fig. 3**). The correlation between the relative abundance of ^{26}Mg and the Al/Mg ratio shown in Fig. 3 indicates that the inclusion crystallized before the ^{26}Al had decayed, while the slope of the line provides the relative ^{26}Al abundance at the time of formation, $^{26}Al/^{27}Al \sim 5 \times 10^{-5}$. The nearly uniform abundance of ^{26}Al among refractory inclusions suggests that the majority of these objects formed over an interval no longer than 100,000 years. No material formed in the solar system is known to have higher $^{26}Al/^{27}Al$, that is, to have formed prior to refractory inclusions.

The short-lived nuclides are not restricted to primitive chondritic meteorites but are also prominent in meteorites produced by planetary-scale melting. Asteroids similar to Vesta must have grown and differentiated within a few million years after the formation of refractory inclusions. A similar time scale is found for parent-body metamorphism of undifferentiated chondritic meteorites. These observations require an efficient and fast-acting heat source to drive planetary melting and differentiation. The long-lived radionuclides ^{40}K, ^{232}Th, ^{235}U, and ^{238}U decay much too slowly to produce melting on small planets within a few million years, and the short-lived nuclides ^{26}Al and ^{60}Fe are the most plausible sources. The abundances of ^{26}Al and ^{60}Fe found, respectively, in refractory inclusions ($^{26}Al/^{27}Al \sim 5 \times 10^{-5}$) and chondrules ($^{60}Fe/^{56}Fe \sim 1 \times 10^{-6}$) are sufficient to melt the interiors of parent bodies larger than a few kilometers. A new wealth of new data from the Hf-W, Fe-Ni, Mn-Cr, and Al-Mg parent-daughter systems in differentiated meteorites provides conclusive evidence for the dominant role of short-lived radionuclides for melting and core formation on small planetesimals.

The discovery of short-lived radionuclides also has important implications for what may be called presolar chronology, or the time-scale of events immediately preceding the birth of the solar system. The widespread presence of short-lived nuclides indicates that the time interval between nucleosynthesis and the formation of the solar system was extremely short. The upper limit to this interval is set by the lifetime of ^{41}Ca (Table 2). The refractory inclusions containing ^{41}Ca must have formed within a few mean lives, that is, a few hundred thousand years of the time of ^{41}Ca production. This time scale is so short that it suggests the formation of the solar system may have been triggered by a shock wave from a nearby exploding star containing freshly synthesized nuclear material. The one radionuclide with an even shorter half-life of 53 days, ^{7}Be, must have been produced by energetic particle interactions between gas and dust within the nascent solar nebula and cannot be used to limit the time since nucleosynthesis.

Isotope anomalies and presolar grains. Except for the variations in isotopic abundances related to decay of radionuclides discussed above, for the most part there is a close similarity between the isotopic compositions of meteorites, the Moon, and the Earth. For many years this apparent homogeneity was taken as evidence that the solar nebula was chemically and physically well mixed before the formation of the first solid bodies. The development of new analytical techniques permitting the determination of isotopic compositions in small (<1 μm) constituents of meteorites has, however, led to the discovery of large differences in isotopic abundances for many elements which cannot be explained in terms of radioactive decay or cosmic-ray interactions. These isotope anomalies provide unequivocal evidence that dust grains that condensed around a variety of different stars were incorporated into primitive meteorites and offer a window through which nucleosynthesis and stellar evolution are revealed with unprecedented clarity.

Two distinct types of meteoritic material with isotope anomalies can be distinguished: (1) samples closely related to chondritic meteorites but containing unusual isotope abundances for selected elements, and (2) authentic interstellar grains. Refractory inclusions are the primary example of the first type of material. These inclusions unquestionably formed within the solar system but exhibit unusual isotope abundances for a number of elements, including oxygen, magnesium, silicon, calcium, titanium, chromium, iron, zinc, strontium, barium, neodymium, and samarium. The variations in O-isotope abundances are particularly significant, because they hold an important clue to understanding

nebular and planetary processes in the early solar system. Oxygen in refractory inclusions is enriched in ^{16}O by up to 4% relative to the average solar composition; no other element exhibits such a large variation in terms of the total number of atoms.

Aside from oxygen, the most prominent nucleosynthetic isotope effects in refractory inclusions are excesses and deficits of the neutron-rich isotopes in elements of the iron abundance peak. Correlated effects for these isotopes provide definitive evidence for the preservation of an exotic component, produced in a supernova during neutron-rich nucleosynthesis. These same inclusions also contain large (percent-level) oxygen, magnesium, and silicon anomalies due to mass-dependent shifts in relative isotope abundance. The connection between isotope anomalies reflecting stellar nucleosynthesis and those due to fractionation processes is poorly understood and remains a subject of current investigation.

Oxygen isotope anomalies are found on all scales studied, from less than a micrometer to many kilometers. The characteristic patterns of oxygen isotopes are similar to those found in refractory inclusions—large variations in the ^{16}O abundance with very small changes in the $^{17}O/^{18}O$ ratio. Different meteorite classes [the H (heavy), L (light), and LL chondrites] have distinct oxygen isotope compositions, requiring at least two different oxygen reservoirs in the solar nebula. The ^{16}O-rich reservoir most plausibly consisted of dust grains, while the ^{16}O-poor reservoir, containing the greater part of the solar system's oxygen, was gaseous. Mixing and exchange between these reservoirs as the solar nebula evolved gave rise to the heterogeneous array of oxygen-isotope compositions seen today. The only group of meteorites to have bulk oxygen isotope abundances similar to those of the Earth and Moon are the enstatite chondrites.

The isotope anomalies found in refractory inclusions represent a mixture of exotic stardust from other stars diluted with normal solar system material; the inclusions themselves clearly are not presolar material. A major new discovery in meteorite research is the identification in primitive chondrites of whole interstellar grains, formed outside the solar system in stellar atmospheres and incorporated into primitive meteorites essentially intact. These grains survived transit through the interstellar medium and injection into the forming solar system and preserve the isotopic and chemical compositions of the stellar nuclear sources for both major and trace elements. Presolar grains show isotope anomalies for all elements, and the size of the anomalies is at least a factor of 100 times larger than that of anomalies found in refractory inclusions.

At least 11 distinct types of presolar grains have been identified in meteorites: diamond, silicon carbide (SiC), titanium carbide (TiC), rutile (TiO_2), graphite, hibonite ($CaAl_{12}O_{19}$), corundum (Al_2O_3), spinel ($MgAl_2O_4$), silicon nitride (Si_3N_4), and, most recently, the common magnesium-rich silicate minerals, olivine and forsterite. These presolar grains have high formation temperatures and, with the exception of diamond, are generally about 1 micrometer in size. The abundance of the different types of grains varies widely among different meteorites, but diamond is always most abundant, while titanium carbide, rutile, and silicon nitride are least abundant. Presolar grains were originally incorporated into all classes of chondritic meteorites. Today, primitive meteorites of low petrologic type (for example, CI or CM chondrites) that never experienced high metamorphic temperatures contain the highest abundance of presolar grains. The pattern of decreasing abundance with increasing petrologic grade of the host meteorite reflects the selective destruction of presolar grain populations during parent body metamorphism. The abundance of presolar grains (diamond and silicon carbide, primarily) now provides a useful measure of the extent of meteorite parent body metamorphism.

With an abundance of approximately 0.05% by mass in carbonaceous chondrites, diamond is by far the most prevalent interstellar material. Interstellar diamonds are exquisitely small (mean size of only about 3 nanometers) and can be analyzed only as aggregates containing hundreds of thousands of grains. The carbon isotope composition of nanodiamonds is very similar to that of average solar system material, and the presolar origin of nanodiamonds is indicated by large enrichments in ^{14}N and Xe-HL (an exotic xenon component characterized by excesses of both the light and heavy Xe isotopes). Despite their small size, diamonds contain 50–90% of the Xe-HL in bulk chondrites. A supernova environment provides the best explanation for the observed properties of presolar diamonds.

Presolar SiC is much less abundant, about 10–15 parts per million (ppm), and much larger (about 0.1–20 μm) than the nanodiamonds. Carbon and nitrogen in SiC exhibit an extraordinary range in isotopic composition; $^{12}C/^{13}C$ ratios extend from 2 to 7000, compared to the range in terrestrial materials of 88 to 90, while $^{14}N/^{15}N$ ratios vary from 5 to more than 20,000, compared to the range in terrestrial materials of 260 to 284. Presolar SiC exhibits isotope anomalies for all elements measured, including helium, neon, magnesium, aluminum, silicon, potassium, calcium, titanium, zirconium, barium, and molybdenum. These isotope abundance patterns cannot be found anywhere in the solar system and indicate that most SiC grains formed as condensates in the atmospheres of carbon-rich, red-giant stars. Approximately 1% of SiC exhibits particularly unusual isotopic properties indicative of supernova origin; the few silicon nitride grains are also related to supernovae.

Interstellar graphite is slightly less abundant, about 10 ppm, than SiC and occurs as micrometer-size spherules which, in turn, often contain even smaller grains of Ti-, Zr-, and Mo-carbide. Presolar graphite exhibits a wide range in carbon and nitrogen isotope compositions, similar to SiC, and also contains essentially pure ^{22}Ne, extreme enrichments in ^{18}O, and large excesses of ^{26}Mg from the decay of extinct ^{26}Al. The inferred initial $^{26}Al/^{27}Al$ ratios range up to

about 0.2, compared to a maximum of 5×10^{-5} in refractory inclusions. The diverse array of isotopic compositions cannot be explained by a single type of stellar source, and requires several different sources, including Wolf-Rayet stars, novae, and carbon-rich giant stars. *See* NOVA; WOLF-RAYET STAR.

Interstellar oxide grains are nearly 1000 times less abundant than presolar SiC, range from 0.5 to 5 μm in size, and are identified as presolar by their extremely anomalous oxygen isotope ratios. The $^{16}O/^{18}O$ ratio extends from 100 to greater than 5000, compared to the solar value of 499. Most presolar oxide grains contain excess ^{26}Mg from the decay of ^{26}Al and have isotope compositions, indicating a red giant branch star (RGB) progenitor.

Of particular interest is the discovery of presolar silicate grains whose chemical composition (rich in magnesium, silicon, and oxygen) is nearly indistinguishable from that of normal meteoritic and terrestrial minerals. Although astronomical observations indicated the presence of silicate grains in dust shells surrounding young stars, previous searches for presolar silicates proved unsuccessful. The presolar silicates were identified using a new type of imaging mass spectrometer to locate small grains displaying extreme oxygen isotope ratios, up to 100-fold enrichments in ^{17}O and ^{18}O relative to solar system matter. These presolar silicates are also found in interplanetary dust particles (IDPs) and have a surprisingly high abundance, ~70 ppm of the whole meteorite, making silicates the second most abundant type of presolar material. Like SiC and graphite, presolar silicates appear to have originated from intermediate-mass asymptotic giant branch stars.

Nucleosynthesis. Studies of isotope anomalies and presolar grains in primitive meteorites have yielded important new clues to nucleosynthesis and the types of stars contributing material to the nascent solar nebula. The very large range of isotope abundances measured in presolar grains cannot be generated by a single stellar source, and is ample proof that the solar system is the product of contributions from disparate stars. Hydrogen and the bulk of the helium in the solar system are relicts of big bang nucleosynthesis, while the elements carbon through iron were produced dominantly in charged particle reactions in massive stars. The radionuclides ^{129}I, ^{232}Th, ^{235}U, ^{238}U, and ^{244}Pu are all produced by neutron-rich nucleosynthesis, the *r*-process, in supernovae. The abundance of short-lived radionuclides requires the late addition of fresh nucleosynthetic material, most plausibly from massive, red-giant stars, to the solar nebula less than a few hundred thousand years prior to the formation of the first meteorites. The short-lived beryllium nuclei, ^{7}Be and ^{10}Be, cannot be produced in stellar nucleosynthesis and must the product of energetic particle interactions in the protosolar nebula. *See* BIG BANG THEORY; NUCLEOSYNTHESIS; STELLAR EVOLUTION; SUPERNOVA.

Ian D. Hutcheon

Bibliography. T. J. Bernatowicz and E. K. Zinner (eds.), *Astrophysical Implications of the Laboratory Study of Presolar Materials*, American Institute of Physics, 1997; D. C. Black and M. S. Matthews (eds.), *Protostars and Planets II*, 1985; R. N. Clayton, R. W. Hinton, and A. M. Davis, Isotopic variations in the rock-forming elements in meteorites, *Phil. Trans. Roy. Soc. London*, 325:483–501, 1988; C. R. Cowley, *An Introduction to Cosmochemistry*, 1995; A. M. Davis (ed.), *Meteorites, Planets, and Comets*, vol. 1 of H. D. Holland and K. K. Turekian (eds.), *Treatise on Geochemistry*, Elsevier-Pergamon, Oxford, 2003; R. Hutchison, *Meteorites: A Petrologic, Chemical and Isotopic Synthesis*, Cambridge University Press, 2004; J. F. Kerridge and M. S. Matthews (eds.), *Meteorites and the Early Solar System*, 1988; A. N. Krot, E. R. D. Scott, and B. Reipurth (eds.), *Chondrites and the Protoplanetary Disk*, Astronomical Society of the Pacific Conf. Series, vol. 341, 2005; D. L. Lambert, The p-nuclei: Abundances and origins, *Astron. Astrophys. Rev.*, 3:201–256, 1992; D. S. Lauretta and H. Y. McSween, Jr. (eds.), *Meteorites and the Early Solar System II*, University of Arizona Press, 2006; E. Levy, J. I. Lunine, and M. S. Matthews (eds.), *Protostars and Planets III*, 1993; H. Y. McSween, Jr., *Meteorites and Their Parent Planets*, 2d ed., 1999; U. Ott, Interstellar grains in meteorites, *Nature*, 364:25–33, 1993; S. R. Taylor, *Solar System Evolution: A New Perspective*, 2d ed., Cambridge University Press, 2001.

Cosmogenic nuclide

A rare nuclide produced by nuclear reactions between high-energy cosmic radiation and terrestrial or extraterrestrial material. Cosmogenic nuclides may be used to examine the history of exposure to cosmic rays, and have numerous applications in earth science and archeology.

Origin of terrestrial nuclides. Primary cosmic radiation is present in space and consists of nuclear particles (mostly protons and alpha particles) with energies that are greater than typical nuclear binding energies. Collisions between cosmic-ray particles and atomic nuclei produce fragments, including cosmogenic nuclides, and also alter the characteristics of the cosmic radiation. As cosmic radiation enters the atmosphere, its flux decreases and its composition becomes dominated by neutrons rather than protons and alpha particles (which react with atmospheric gases). This reduction in flux continues through the atmosphere; at ground level, small but measurable quantities of cosmogenic nuclides are produced in solid material within about 1 m (3.3 ft) of the Earth's surface. The cosmogenic nuclides produced in these two reservoirs (atmosphere and lithosphere) may be employed for examination of different geological processes. Because of their wide range of half-lives and chemical properties, cosmogenic nuclides have a wide range of applications in geological, geomorphological, and biogeochemical studies. *See* ALPHA PARTICLES; COSMIC RAYS; COSMOCHEMISTRY; PROTON.

Production. Production of cosmogenic nuclides varies over both space and time. The flux of cosmic rays, and hence cosmogenic nuclide production rates, decreases exponentially (by a factor of 2

for every ~100 g/cm^2 increase in pressure) with depth into the atmosphere (**Fig. 1**). In addition, the Earth's magnetic field—weakest at the magnetic poles and strongest at the Equator—diverts charged particles before they enter the upper atmosphere. As a result, production rates are greater at high geomagnetic latitude than at low geomagnetic latitude (Fig. 1).

Production of cosmogenic nuclides also varies with time. The three major potential variables are solar activity, the primary flux of galactic cosmic rays, and the Earth's magnetic field. Solar activity has an ~11-year cycle, varying around a mean that is fairly constant over longer periods. The galactic cosmic-ray flux is believed to have been relatively constant for at least the past 10^7 years. Therefore, geomagnetic field variability is probably the most important source of production rate changes over the past 10^3–10^6 years.

The earliest measurements of cosmogenic nuclides utilized radioactive decay counting and established the bases of many present-day studies. The development of exceptionally sensitive techniques (accelerator mass spectrometry and static mass spectrometry) during the 1980s improved detection limits by factors of up to 10^6 and greatly broadened the range of applications of these nuclides. *See* ACCELERATOR MASS SPECTROMETRY.

Atmospheric nuclides. Numerous nuclides are produced in measurable quantities in the atmosphere (**Table 1**). Their half-lives and their chemical reactivities determine their applications in earth science. *See* RADIOISOTOPE (GEOCHEMISTRY).

Short-lived nuclides. Short-lived nuclides, including beryllium-7 (^{7}Be), sodium-22 (^{22}Na), phosphorus-32 (^{32}P), and phosphorus-33 (^{33}P), have been employed in studies of atmospheric circulation, particularly stratospheric-tropospheric exchange. Because production rates are far higher in the stratosphere than in the troposphere, concentrations of these nuclides in atmospheric aerosol particles are related to the extent of exchange processes. *See* STRATOSPHERE; TROPOSPHERE.

These nuclides also have applications in oceanography. Phosphorus is an essential nutrient element in the surface ocean; its isotopes ^{32}P and ^{33}P have half-lives appropriate for studies of surface-layer oceanic biological productivity. The tendency for beryllium in aquatic systems to adsorb onto particle surfaces makes ^{7}Be useful in tracing sediment transport and accumulation, on time scales up to several months, in estuarine and coastal regions.

Longer-lived nuclides. Because of its highly successful applications in dating, carbon-14 (^{14}C; radiocarbon) is the best-known cosmogenic nuclide. Essentially all atmospheric ^{14}C is in the form of gaseous carbon dioxide ($^{14}CO_2$); it thus has an atmospheric residence time long enough for the ratio of ^{14}C to ^{12}C (the stable isotope) to become homogeneous throughout the atmosphere. Living organisms and inorganic carbonates incorporate carbon with an isotope ratio reflecting that of the atmosphere. Upon removal from contact with the atmosphere, the ratio $^{14}C/^{12}C$ decreases through radioactive decay of ^{14}C. With knowledge of the initial ratio and the half-life of ^{14}C, measurement of $^{14}C/^{12}C$ allows calculation of the age of a sample. This technique has been extensively used in archeology and geology for samples as old as 4×10^4 years. High-precision ^{14}C dates of inorganic carbon dissolved in seawater are also used to deduce global oceanic circulation rates and patterns by determining the time since a particular water mass had contact with the atmosphere. *See* OCEAN CIRCULATION; RADIOCARBON DATING; SEAWATER.

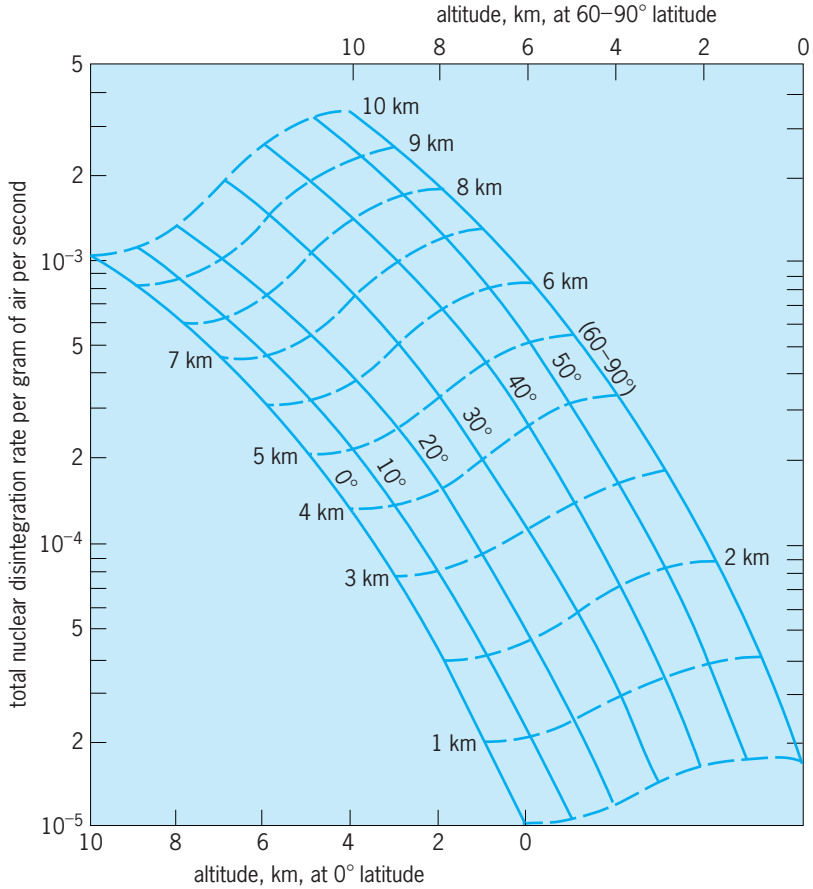

Fig. 1. Total rate of nuclear disintegrations in the atmosphere (with energy release greater than 40 MeV) is plotted as a function of altitude for various geomagnetic latitudes. Curves for each latitude (0–90°, in steps of 10°) are successively displaced by 1 km (0.6 mi) along the abscissa. It can be seen that disintegration rates increase exponentially with altitude and are lower at low geomagnetic latitude. The relationship between the total nuclear disintegration rate and the production rate differs slightly for each nuclide. (*After D. Lal and B. Peters, Cosmic ray produced radioactivity on the earth, in S. Flugge, ed., Handbuch der Physik, vol. 46/2, pp. 551–612, Springer-Verlag, 1967*)

Production rates of cosmogenic nuclides play a key role in setting the initial value of isotope ratios used in calculating ages. The record of the $^{14}C/^{12}C$ ratio in annual tree rings allowed reconstruction of production rate variations over the past 10^4 years. This range has been extended to about 3×10^4 years by measurements of the $^{14}C/^{12}C$ ratio in independently dated corals. These results indicate that the $^{14}C/^{12}C$ ratio in the atmosphere was ~1.4 times the present value during the period between 2–3 × 10^4 years ago. *See* DENDROCHRONOLOGY.

Unlike ^{14}C, most other cosmogenic nuclides have atmospheric residence times on the order of months, and thus they may provide information on cosmogenic nuclide production rate variations on time

TABLE 1. Principal cosmogenic nuclides produced in the atmosphere and their applications

Nuclide	Major target elements	Half-life	Main applications
Beryllium-7 (^{7}Be)	Nitrogen, oxygen	53 days	Tracing of atmospheric mixing Sediment transport and deposition
Beryllium-10 (^{10}Be)	Nitrogen, oxygen	1.5×10^{6} years	Dating and tracing of marine sediments Examination of soil formation Ice core studies
Carbon-14 (^{14}C)	Nitrogen	5730 years	Dating of organic carbon, shells, and so forth Determination of rates of oceanic circulation
Sodium-22 (^{22}Na)	Argon	2.6 years	Tracing of atmospheric mixing
Silicon-32 (^{32}Si)	Argon	130 years	Upper-ocean biological processes and mixing
Phosphorus-32 (^{32}P)	Argon	14.3 days	Tracing of atmospheric mixing Upper-ocean biological processes
Phosphorus-33 (^{33}P)	Argon	25 days	Tracing of atmospheric mixing Upper-ocean biological processes
Chlorine-36 (^{36}Cl)	Argon (chlorine in aerosols)	3.0×10^{5} years	Dating and tracing of ground waters

scales of years to decades. Records of ^{10}Be concentrations in ice cores drilled in Antarctica and Greenland extend as far back as $\sim 1.5 \times 10^{5}$ years. These results are useful for determining precipitation rates and establishing chronologies. However, with careful normalizations they also may be used to examine past variability in production rates. Similarly, a record of beryllium-10 (^{10}Be) fluxes combined from several marine sediment cores provides a 200,000-year history of cosmogenic nuclide production.

Beryllium-10 has also been used for dating marine sediment cores and manganese encrustations up to 1.5×10^{7} years old. If ^{10}Be is homogenized with the stable isotope (^{9}Be) in the oceanic water column, the beryllium incorporated into sediments will have a known isotope ratio. Any changes in the ^{10}Be/^{9}Be ratio down the sediment core can then be assumed to be due to radioactive decay of ^{10}Be. The tendency of beryllium to adsorb onto particles in aqueous systems gives it a relatively short oceanic residence time, making it necessary to consider the possibility of past variation in the oceanic ^{10}Be/^{9}Be ratio. Details of the incorporation of beryllium into marine sediments have been obtained by comparison of ^{10}Be with thorium-230 (^{230}Th) and protactinium-231 (^{231}Pa), other so-called particle-reactive radionuclides. *See* MANGANESE NODULES; MARINE SEDIMENTS.

Atmospheric cosmogenic nuclides can be used to trace mixing between geological reservoirs. For example, the ratio of cosmogenic chlorine-36 (^{36}Cl) to stable chlorine differs greatly in atmospheric, continental, and marine material; chlorine isotopes are thus used to examine the relative contributions of these possible sources to salts in saline lakes and ground waters. Similarly, because ^{10}Be has very high concentrations in marine sediments, its presence within mineral matrices of arc volcanic rocks provides evidence for the role of subducted young sediment in the associated magma. In addition, ^{10}Be concentrations in fluids emitted by ocean-ridge hydrothermal systems can quantify the extent of interaction between those fluids and marine sediments.

Cosmogenic nuclides also have applications in examining soils and ground waters. Chlorine-36 produced in the atmosphere and brought to the Earth's surface in precipitation remains in solution; it is used for tracing and dating ground waters. In contrast, beryllium has very low mobility in neutral-to-alkaline ground waters. The inventory of ^{10}Be in a soil profile is related either to the age of the profile or to the residence time (relative to removal in ground water or by erosion) of beryllium in the profile.

Ground-level nuclides. A small but significant flux of cosmic rays is present at ground level (Fig. 1). This radiation produces cosmogenic nuclides within mineral lattices of exposed rock. In-situ-produced cosmogenic nuclides are used to examine geological problems (**Table 2**). In contrast to the case of cosmogenic nuclides produced in the atmosphere, in which applications are linked to the geochemical behavior of each element as well as the half-life of the nuclide, the uses of in-situ-produced cosmogenic nuclides in exposed rock surfaces depend primarily on the half-lives of the nuclides.

In-situ-produced cosmogenic nuclides are useful for dating periods of surface exposure. Their production rates decrease by a factor of 2 for every $\sim$40 cm (16 in.) depth below an exposed rock surface, so accumulation of cosmogenic nuclides can date geological events that bring material to the Earth's surface. For example, the theoretical evolution of the concentration of ^{10}Be as a function of time for various erosion rates can be calculated (**Fig. 2**). The method has been used to date exposure of rocks brought to the surface by processes including glaciation, volcanic activity, and meteor impact. It has also been used to date the formation of geomorphological features (such as alluvial fans or glacial moraines) deposited over active faults and subsequently offset by fault movement. This provides a quantitative

TABLE 2. Principal cosmogenic nuclides produced in situ within surface rock

Nuclide	Half-life, years	Major target elements	Advantages	Disadvantages
Helium-3 (^{3}He)	Stable	Oxygen, magnesium, silicon, iron	High production rate; low analytical detection limits	Loss by diffusion; corrections for primordial and nucleogenic ^{3}He
Beryllium-10 (^{10}Be)	1.5×10^{6}	Oxygen, magnesium, silicon, iron	Production fairly independent of rock composition	Potential contamination by atmospheric ^{10}Be
Carbon-14 (^{14}C)	5730	Oxygen, magnesium, silicon, iron	Less influence of exposure before current episode	Only for exposure ages under 15,000 years
Neon-21 (^{21}Ne)	Stable	Magnesium, aluminum, silicon, iron	Stable isotope with low diffusive loss rate	Corrections for atmospheric, primordial and nucleogenic ^{21}Ne
Aluminum-26 (^{26}Al)	7.2×10^{5}	Silicon, aluminum, iron	Minimal atmospheric production	Significant noncosmogenic production
Chlorine-36 (^{36}Cl)	3.0×10^{5}	Iron, calcium, potassium, chlorine	Ease of chemical purification	Production rates highly dependent on rock composition

approach to determine slip rates on faults and earthquake recurrence intervals. Such results could not have been obtained by conventional methods. *See* GEOMORPHOLOGY.

Erosional losses of in-situ-produced nuclides limit their use in certain cases. In fact, outside of Antarctica (where erosion rates are exceedingly slow), cosmic-ray-exposure dating is generally applicable only to exposures of less than a few hundred thousand years. However, in cases where independent evidence indicates that exposure has been sufficiently long for nuclides to reach their steady-state concentrations, cosmogenic nuclide concentrations may be used to calculate mean physical erosion rates (Fig. 2). In some cases, this approach may be extended to the scale of a watershed by assuming that the cosmogenic nuclide content of sediment discharged by a river represents the average concentration—weighted for area and denudation rate of the various erosional environments present—for the basin as a whole. In addition, spatial distributions of in-situ-produced nuclides and ratios of nuclides with differing half-lives can provide criteria for understanding whether rocks have undergone previous episodes of exposure and burial. *See* ROCK AGE DETERMINATION.

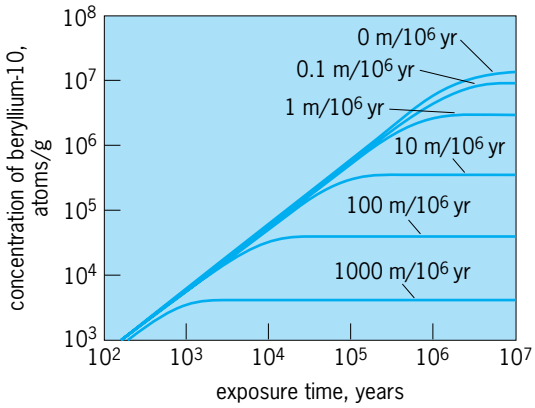

Fig. 2. Theoretical accumulation of in-situ-produced ^{10}Be as a function of time for various erosion rates. These curves are based on calculations using present-day production rates for sea level and high latitude. Concentrations increase with time to reach steady-state values with respect to radioactive decay and erosional loss. Nuclides with other half-lives have similar behavior, but with different time scales. (*After E. T. Brown et al., Examination of surface exposure ages of moraines in Arena Valley, Antarctica using in-situ-produced beryllium-10 and aluminum-26, Geochim. Cosmochim. Acta, 55:2269–2283, 1991*)

Erik Thorson Brown

Meteorites. Collisions among the solid bodies of the solar system produce many small fragments. A fragment from a collision that finds its way to the top of the Earth's atmosphere is known as a meteoroid. A meteoroid becomes a meteorite when it lands on the Earth. While a meteoroid is in space, cosmic rays bombard its surface and penetrate to depths of a few meters. The exposure age of an extraterrestrial object is the time during which cosmic rays bombarded it in space.

Because cosmic rays penetrate just a few meters of solid matter, cosmogenic nuclides cannot form deep in the interiors of large bodies such as asteroids. The accumulation of cosmogenic nuclides begins when a collision lifts deep-seated material to the surface or ejects it into space as part of a small body. When a meteorite hits the Earth, the accumulation of cosmogenic nuclides effectively stops. The reason is that the Earth's atmosphere and magnetic field screen out most cosmic rays.

Identification. The total amounts of the cosmogenic nuclides in a meteorite are related to its exposure age. They also indicate the size of the meteoroid. To get this information, it is necessary to be able to identify cosmogenic nuclides as such and measure their concentrations. Many of the nuclides found in terrestrial materials are radioactive. All of them have lifetimes of a few million years or less. Most bodies in the solar system solidified about 4.5×10^{9} years ago. Thus, even if an asteroid contained some noncosmogenic ^{10}Be when it formed, that ^{10}Be would long since have vanished. Accordingly, when ^{10}Be is found in a meteorite today, it is likely that cosmic rays produced it recently. The same conclusion applies to other radioactive nuclides found in terrestrial materials.

The stable nuclides pose a different kind of problem in identification. For example, the total concentration of purely cosmogenic neon-21 (^{21}Ne) seldom reaches 1 part per billion, an extremely low value. If a meteorite parent body contained much noncosmogenic ^{21}Ne when it formed, then that ^{21}Ne would drown any later contribution due to cosmic rays. As it happens, the parent bodies of meteorites crystallized with very few atoms of neon. Indeed, the low "primordial" abundance of neon is one reason why ^{21}Ne is such an important stable cosmogenic nuclide. Some meteorites do contain measurable concentrations of primordial neon. For them, a general knowledge of isotope abundance patterns can help in sorting out the cosmogenic and noncosmogenic components (**Table 3**). *See* ISOTOPE.

Measurement of exposure ages. The concentrations of cosmogenic nuclides in meteorites are so low that very sensitive techniques are needed to measure them. Formerly, the only practical method for measuring the radioactive nuclides was to observe their rates of decay, a slow process requiring large samples. With the development of accelerator mass spectrometry, it became possible to measure directly the number of atoms of cosmogenic nuclides. Accelerator mass spectrometry has largely replaced decay counting for cosmogenic radionuclides with half-lives greater than a few thousand years. Conventional methods of mass spectrometry give reliable results for the stable cosmogenic nuclides. *See* MASS SPECTROMETRY.

If the concentration of a stable cosmogenic nuclide N grows at a constant rate P in a meteoroid, then after an exposure lasting t years the meteorite will contain a concentration N equal to $P \times t$. The value t for a meteorite can be calculated by measuring the concentration N, provided that the value of the production rate P is known. One way to obtain a value for production rate is to calculate it from basic physical principles. Another is to use the information given by the radioactive cosmogenic nuclides.

The values for meteorite exposure ages range from times too short to measure to 2.5×10^9 years. Stony meteorites generally have exposure ages of 2–20 $\times$ 10^6 years. Iron meteorites have much longer exposure ages, about 5×10^8 years on average, than do stony meteorites, perhaps because iron meteorites resist destruction better. Many groups of meteorites have several members with the same exposure age. The stony meteorites known as H-chondrites provide a striking example of exposure age clustering at about 7×10^6 years. The existence of such clusters suggests that the meteorites involved originated in the same body in a single collision. Interestingly, many of the meteorites known as diogenites, eucrites, and howardites seem to have a common exposure age of about 2×10^7 years. This observation supports the idea that all three groups come from the same object, even though they differ in chemical and mineral composition.

Lunar meteorites have some of the shortest exposure ages measured, no doubt because they had only a short distance to travel. Research related to the problem of meteorite exposure ages is focused on improving the values of production rates, determining the exposure ages of rare but important new meteorites, refining existing values, deciphering exposure histories of objects that underwent more than one episode of irradiation, and understanding the behavior of cosmic rays from the Sun and from the stars over time. *See* METEORITE. G. F. Herzog

TABLE 3. Some important cosmogenic nuclides in extraterrestrial materials

Isotope	Half-life
Hydrogen-3 (^{3}H)	12.3 years
Helium-3 (^{3}He)	Stable
Beryllium-7 (^{7}Be)	53 days
Carbon-14 (^{14}C)	5730 years
Neon-21 (^{21}Ne)	Stable
Sodium-22 (^{22}Na)	2.6 years
Aluminium-26 (^{26}Al)	7.1×10^5 years
Magnesium-26 (^{26}Mg)	Stable
Chlorine-36 (^{36}Cl)	3.0×10^5 years
Argon-36 (^{36}Ar)	Stable
Argon-38 (^{38}Ar)	Stable
Argon-37 (^{37}Ar)	35 days
Argon-39 (^{39}Ar)	269 years
Potassium-40 (^{40}K)	1.3×10^9 years
Potassium-39 (^{39}K)	Stable
Potassium-41 (^{41}K)	Stable
Calcium-41 (^{41}Ca)	1×10^5 years
Manganese-53 (^{53}Mn)	3.5×10^6 years
Manganese-54 (^{54}Mn)	312 days
Nickel-59 (^{59}Ni)	7.6×10^4 years
Iron-60 (^{60}Fe)	1.5×10^6 years
Cobalt-60 (^{60}Co)	5.27 years
Krypton-81 (^{81}Kr)	2.1×10^5 years
Krypton-78 (^{78}Kr)	Stable
Krypton-80 (^{80}Kr)	Stable
Krypton-82 (^{82}Kr)	Stable
Krypton-83 (^{83}Kr)	Stable
Iodine-129 (^{129}I)	1.6×10^7 years
Xenon-124 (^{124}Xe)	Stable
Xenon-132 (^{132}Xe)	Stable

Bibliography. E. Anders, Meteorite ages, *Rev. Mod. Phys.*, 34:287–325, 1962; T. E. Cerling and H. Craig, Geomorphology and in-situ cosmogenic isotopes, *Annu. Rev. Earth Planet. Sci.*, 22:273–317, 1994; D. Elmore and F. Phillips, Accelerator mass spectrometry for measurement of long-lived radioisotopes, *Science*, 236:543–550, 1987; G. Faure, *Principles of Isotope Geology*, 2d ed., 1986; M. Honda and J. R. Arnold, Effects of cosmic rays on meteorites, *Handbuch der Physik*, vol. 46/2, pp. 613–632, 1967; R. C. Reedy, J. R. Arnold, and D. Lal, Cosmic-ray record in solar system matter, *Annu. Rev. Nucl. Part. Sci.*, 33:505–537, 1983; S. Vogt, G. F. Herzog, and R. C. Reedy, Cosmogenic nuclides in extraterrestrial materials, *Rev. Geophys.*, 28:253–275, 1990.

Cosmological constant

The constant gravitational source added to the general theory of relativity by Einstein to balance the gravitational field equations necessary for a static universe. In formulas it is denoted Λ (lambda) and has units of curvature, or $(\text{length})^{-2}$. The existence of the cosmological constant would mean that empty space gravitates as if it were filled by a static, uniform

energy density $\rho = \Lambda c^4/8\pi G$, where G is the newtonian constant of gravitation, and pressure $p = -\rho$. It is speculated to be responsible for dark energy, since a sufficiently large cosmological constant would cause the accelerated expansion of spacetime. However, the physical origin of the cosmological constant is not understood. Quantum theory predicts the existence of the cosmological constant, but with an energy density that is approximately 10^{120} times too large to be compatible with observations. This gross mismatch between theory and observation, known as the cosmological-constant problem, is a deep enigma of astrophysics and physics. It is thought that a resolution requires a quantum theory of gravity. Testing for the existence of the cosmological constant is a major goal of current physics and astronomy research. *See* ACCELERATING UNIVERSE; DARK ENERGY; QUANTUM GRAVITATION; RELATIVITY.

Introduction by Einstein. The cosmological constant was introduced by Albert Einstein in 1917, in the first application of general relativity to cosmology. In order to balance theoretical and practical considerations in constructing a mathematical model of the universe, Einstein found it necessary to invoke a new cosmological term, Λ. He regarded this term as a universal constant whose value would be determined empirically. The effect of adding the Λ was to cut off the gravitational influence of distant matter and balance the field equations for a static universe. This cosmological model was consistent with contemporary observations of the slow kinetic motions of stars and nebulae. Though the 1929 discovery of the cosmic redshift-distance by Edwin Hubble, soon interpreted as expansion, invalidated Einstein's construction, the cosmological constant survived. Later, when discussing the discovery of the expanding universe with George Gamow, Einstein is reputed to have called the introduction of the cosmological constant "the greatest blunder" of his life.

De Sitter model. Shortly after Einstein's introduction of the cosmological constant, Willem de Sitter proposed a theoretical model universe containing the cosmological constant but no other forms of matter. This solution of the gravitational field equations is an idealization of the present-day universe, wherein the stars and galaxies are regarded as test particles that sit atop a perfect sea of static, uniform energy. Since general relativity dictates that pressure gravitates in addition to mass-energy, the negative pressure or tension of the cosmological constant causes spacetime to expand at an accelerating rate. Hence, test particles recede from each other as if under the influence of a repulsive gravitational potential. The theoretical prediction of a cosmological redshift can be traced back to this work, also. In the big bang cosmology, the inflationary universe is well described by de Sitter's solution, as is the distant future of the universe filled with dark energy in the form of the cosmological constant. *See* BIG BANG THEORY; INFLATIONARY UNIVERSE COSMOLOGY.

Quantum vacuum as a source. The quantum vacuum was proposed as the source of the cosmological constant by Yakov Zeldovich in 1967. He showed that the irreducible, quantum zero-point energy of elementary particles—a foaming sea of virtual particles—contributes a static, spatially uniform source of gravitation. Consequently, the cosmological constant is interpreted to represent the cosmological vacuum, the lowest energy state of the universe. The magnitude of this energy density depends sensitively on the maximum particle energy allowed in the quantum theory. Zeldovich found that any reasonable estimate for this energy scale, based on the known properties of particles, results in a cosmological energy density that is too large by many orders of magnitude.

There has been debate whether the cosmological constant is a consequence of the geometry of spacetime or of the gravitating sources, traditionally corresponding to the left- and right-hand sides of the gravitational equations, respectively. Einstein placed Λ on the left-hand side, identifying it as a property of the spacetime. Zeldovich placed Λ on the right-hand side, advocating the view that the cosmological constant, as well as gravitation itself, traced their origins to the quantum properties of particles and fields. The difference is largely semantic, since general relativity dictates that the gravitating sources are balanced by the curvature of spacetime.

Arguments for zero value. Prior to the discovery of the cosmic acceleration, most efforts to solve the cosmological constant problem sought to explain that Λ must be identically zero. One such proposal suggested that under a quantum theory of gravity Λ is a quantum-mechanical variable with a vanishing expectation value. Alternatively, it was argued that boundary conditions at the big bang set Λ to zero. In neither case are there additional, testable consequences of the proposed solution. Supersymmetry, a theory of fundamental interactions beyond the standard model of particle physics, offers a solution which can be tested by laboratory experiments. If this new symmetry of nature is valid, it implies that the cosmological vacuum, and hence Λ, is zero. *See* QUANTUM MECHANICS; SUPERSYMMETRY.

Relation to cosmic acceleration. The discovery of cosmic acceleration puts these arguments in limbo. Current observations are consistent with a nonvanishing cosmological constant, although they are not precise enough to eliminate other possibilities. If observations prove that dark energy can be explained only by the cosmological constant, the challenge will be to explain why the magnitude of Λ is so much smaller than the naive estimate.

Energy scales. The cosmological constant based on Zeldovich's calculation can be expressed as $\Lambda = E^4/2\pi(E_P\hbar c)^2$, where E is the maximum allowed zero-point energy, $\hbar$ is Planck's constant divided by 2π, and c is the speed of light. The natural choice for this energy is the Planck energy, $E_P \equiv \sqrt{\hbar e^5/G} \approx 10^{19}$ GeV, the energy scale at which quantum and gravitational effects are thought to unify. However, the resulting value of Λ is approximately 120 orders of magnitude too large. Consistency with the observed accelerating expansion requires that $\Lambda = 3\Omega_\Lambda H_0^2/c^2$, where $\Omega_\Lambda \approx 0.7$ gives

the fraction of the total energy density of the universe contributed by the cosmological constant, and $H_0 \approx 70$ km/s/Mpc is the Hubble constant. These figures give $\Lambda \approx (3000 \text{ Mpc})^{-2}$ and an energy $E \approx 0.008$ eV $\approx 10^{-21}$ J. The significance of this energy is unclear, since virtual particles exist at energies well above this scale. The proximity of this energy to an energy scale associated with neutrino masses has prompted speculation of new physics. Other proposals have linked this energy scale with novel theories of gravity beyond Einstein's general relativity. *See* HUBBLE CONSTANT; NEUTRINO.

Consequences of nonzero value. The existence of the cosmological constant would imply the existence of a cosmological event horizon, a spherical boundary centered on any observer marking a distance beyond which no causal contact is possible. A light signal sent from the center could never reach the boundary because the expansion of spacetime is so rapid. This horizon is similar to the event horizon surrounding a black hole, viewed from within. If the dark energy is a cosmological constant, the horizon is located at a radius $r = \sqrt{3/\Lambda} \sim 5000$ Mpc. Correspondingly, no rocket or light signal that we may send can reach objects currently beyond a redshift $z \approx 1.7$ within the lifetime of the universe. *See* BLACK HOLE; REDSHIFT.

Prospects. Precision measurements of cosmic distances and the evolution of the universe will help refine the evidence for the cosmological constant. Current evidence is ambiguous as to whether Λ is dark energy. If not, it will likely be accepted in practice that Λ is zero. Confirmation of the existence of a nonzero cosmological constant would have a tremendous impact on physics and our view of the universe. *See* COSMOLOGY; UNIVERSE.

Robert Caldwell

Bibliography. L. M. Krauss and M. S. Turner, A cosmic conundrum, *Sci. Amer.*, 291:52–59, 2004; S. Weinberg, The cosmological constant problem, *Rev. Mod. Phys.*, 61:1–23, 1989.

Cosmology

The study of the structure and the origin of the universe, including the origin of galaxies, the elements, and matter itself.

Big bang model. When Albert Einstein proposed his theory of general relativity, the universe was believed to be static. Einstein had to modify his equations so that general relativity would allow a static universe by adding a cosmological constant term. Today, this term is incorporated into the current cosmological model. *See* COSMOLOGICAL CONSTANT; RELATIVITY.

At the time of Einstein's work, astronomers were uncertain about the nature of the faint spiral nebulae and the size of the visible universe. Some contended that they were regions of star formation that were part of the Milky Way Galaxy, while others argued that they were distant galaxies much like the Milky Way. This debate was settled in 1923 by E. Hubble's discovery of 12 Cepheid variable stars in M31, the Andromeda Nebula. There is a simple empirical relationship between the period of the Cepheids and their luminosity. Thus, Hubble's observations allowed him to determine that the Andromeda Nebula was at a very large distance and was a galaxy much like the Milky Way. *See* ANDROMEDA GALAXY; CEPHEIDS; GALAXY, EXTERNAL; MILKY WAY GALAXY.

Hubble continued his study of galaxies and found that most were receding from the Earth. Hubble proposed a simple linear relationship between the distance to a galaxy and its recessional velocity Eq. (1),

$$v = Hr \tag{1}$$

where v is the recessional velocity of the galaxy, usually measured in kilometers per second, and r is the distance to the galaxy, usually measured in megaparsecs (1 Mpc $= 3.26 \times 10^6$ light-years $= 3.1 \times 10^{19}$ km $= 1.9 \times 10^{19}$ mi). Various techniques now yield consistent measurements of the Hubble constant, $H \sim 70$ (km/s)/Mpc. *See* HUBBLE CONSTANT.

Hubble's observations of the expanding universe then implied that the addition of a cosmological constant to the equations of general relativity was unnecessary. In the 1920s, G. Lemaître and A. Friedmann independently proposed a general relativistic model of the expanding universe. One of the simplest solutions to Einstein's relativity equation, this model assumes that the universe is homogeneous and expanding. When Lemaître and Friedmann made their proposal, there was no real evidence for their simplifying assumptions. Only since the late 1980s have observations become sensitive enough to confirm them.

The Friedmann-Lemaître model, often called the big bang model, implies that the universe began in an extremely dense state and expanded and cooled. In this model, the Hubble law is predicted as an approximate description of the expansion valid for galaxies within a few hundred megaparsecs of the Milky Way Galaxy. The model implies that radiation is redshifted as the universe expands. Thus, radiation from distant objects should appear at lower frequencies than those at which it was emitted. Observations of atomic lines from distant quasars confirm that radiation is redshifted just as predicted. *See* QUASAR; REDSHIFT.

The Friedmann-Lemaître model, while fully relativistic, can be described in the language of newtonian physics. The Hubble law implies that a shell of galaxies of radius R and mass m expands with velocity HR. Thus, the kinetic energy of the shell is $m(HR)^2/2$. If M is the mass interior to the shell, then the gravitational binding energy of the shell is GMm/R, where G is the newtonian constant of gravitation. The total energy E of the shell is therefore given by Eq. (2). Since it has been assumed that the

$$E = \frac{m(HR)^2}{2} - \frac{GMm}{R} \tag{2}$$

universe is uniform, the mass M within the shell can

be replaced with the quantity $4\pi R^3/3$, where ρ is the density of the universe and $4\pi R^3/3$ is the volume of a shell of radius R. Then Eq. (2) can be rewritten as Eq. (3), where Ω is the ratio of the density of

$$\frac{E}{mR^2} = \frac{H^2}{2}\left(1 - \frac{8\pi G\rho}{3H^2}\right) = \frac{H^2}{2}(1 - \Omega) \quad (3)$$

the universe to the critical density of the universe, $3H^2/(8\pi G)$. If $\Omega < 1$, the total energy of the shell is positive, and the universe will continue to expand forever. If $\Omega > 1$, the total energy of the shell is negative, gravity will eventually stop the expansion, and the universe will eventually collapse. Some physicists speculate that this so-called big crunch will be followed by a future big bang. If $\Omega = 1$, the total energy is zero, and the universe stands on the balance between open and closed and corresponds to a special solution called the Einstein-de Sitter model. *See* GRAVITATION.

This simple newtonian model, while accurately describing the dynamics of the expanding universe, can lead to a conceptual error. The newtonian shell has a center, since newtonian theory cannot deal with a uniform mass density. General relativity, however, allows the universe to be isotropic and to be expanding uniformly without having a special center point. In the Friedmann-Lemaître model, the Milky Way Galaxy is not a special place in the universe.

Geometry of the universe. The density of the universe determines not only the final fate of the universe but also its geometry. If $\Omega > 1$, the universe is closed and its geometry is that of a three-dimensional sphere. (A circle is a one-sphere and the surface of the Earth is a two-sphere.) If two pilots leave New York in opposite directions, one heading east and the other heading west, they will meet somewhere over the Pacific. If the Earth were flat, the two pilots would continue to head directly away from each other. Similarly, if the universe is closed, two light rays sent off in opposite directions will eventually bend toward each other. If $\Omega = 1$, the universe is flat and the two light rays will continue to move away from each other. If $\Omega < 1$, the geometry of the universe is hyperbolic, much like that of a saddle. A common misconception is that if the universe is open it must be spatially infinite. This relationship need not be true, as general relativity does not constrain the topology of the universe. If $\Omega < 1$, the universe could be either infinite or finite and periodic.

Microwave background radiation. One of the most dramatic discoveries of modern physics was the detection of the microwave background radiation by A. Penzias and R. Wilson. Most cosmologists believe that this radiation is the leftover heat from the big bang.

In the hot big bang model, the universe started in an extremely hot dense state. In this state, the universe was composed of electrons, positrons, quarks, neutrinos, and photons. As the universe expanded, most of the matter annihilated with antimatter into photons. These photons then cooled as the universe expanded. Thermal physics predicts that the spectrum of radiation from the big bang would be similar to that emitted by a blackbody, a so-called Planck spectrum. The Friedmann-Lemaître model predicts that this radiation should be uniform since the big bang started in a uniform state. *See* HEAT RADIATION.

The observations of the *Cosmic Background Explorer* (*COBE*) satellite, launched in 1989, provided strong confirmation of the hot big bang model. The observed spectrum of the microwave background radiation agrees closely with the predicted Planck spectrum. The *COBE* experiment also confirmed that the microwave background radiation is uniform to nearly 1 part in 50,000, consistent with the homogeneity assumption of the hot big bang model.

Observations of small fluctuations in the temperature of the microwave background have become a powerful tool for studying the basic properties of the universe. Since most of the microwave background photons last interacted with matter when the universe was only 500,000 years old, these tiny fluctuations in temperature are directly probing variations in the density of the universe soon after the big bang. *See* COSMIC BACKGROUND RADIATION.

Nucleosynthesis. Within the context of the hot big bang model, the conditions in the universe now can be extrapolated back to the first moments after the big bang. The temperature of the cosmic background radiation now is 2.73 kelvins above absolute zero. When the universe was half its present size, the background temperature was twice as high. One second after the big bang, the universe was only 3×10^{-11} of its present size and the temperature of the microwave background was roughly 10^{10} K (1.8×10^{10} °F). At this high temperature, the universe consisted of a thermal sea of photons, electrons, positrons, and neutrinos. In addition, there were a handful of protons and neutrons. There was approximately 1 baryon (proton or neutron) for every 10^{10} photons. *See* BARYON; ELECTRON; NEUTRINO; NEUTRON; POSITRON; PROTON.

As the universe cooled, the protons and neutrons combined to make deuterium. Most of this deuterium then interacted to make helium, while trace amounts combined to make lithium. One of the successes of the big bang model is its ability to account for the observed abundances of light elements. The universe today consists of roughly 28% helium, 70% hydrogen, and 2% other elements, calculated by their masses. Nucleosynthesis in stars produces roughly equal amounts of helium and heavier elements from hydrogen. Thus, it is difficult to understand the helium abundance without big bang nucleosynthesis. *See* DEUTERIUM; HELIUM; HYDROGEN; LITHIUM.

The amounts of deuterium, helium, and lithium produced in the big bang depend sensitively on the number of baryons per photon in the early universe and on the number of neutrino flavors. Based on the observed abundances of this isotope and these elements in old stars, it was predicted that there should be only three flavors of neutrino. Experiments at particle accelerators have confirmed this prediction based on the first moments of the early universe. *See* FLAVOR.

The observed light-element abundances also lead to a prediction for the density of protons and neutrons. The current best estimates suggest that the density in atoms is 5×10^{-31} gram/cm^3. Measurements of the amplitude of microwave background fluctuations also measure the density in atoms as the pattern of fluctuations depends on the atomic density. The best fit value for the density of atoms based on microwave background fluctuations agrees with the estimate based on light-element abundances to better than 10% (and well within the current statistical errors). *See* BIG BANG THEORY; NUCLEOSYNTHESIS.

Lambda cold dark-matter model. Over the past few decades, a simple cosmological model called the lambda cold dark-matter (lambda CDM) model has emerged as the best fit to the current observational data. The lambda CDM model is a version of the standard hot big bang model. The big bang model assumes that general relativity is valid on cosmological scales and that the distribution of matter on large scales is homogeneous. This set of assumptions implies that the universe is expanding: the distances between galaxies are constantly growing. Since the universe is expanding, it is constantly getting cooler and less dense. In the standard big bang model, the shape and evolution of the universe depend on its composition.

The lambda CDM model assumes that the universe is composed of atoms, dark matter, and vacuum energy. The total energy density of the universe is very close to the critical density so that the geometry of the universe is flat. *See* DARK ENERGY; DARK MATTER.

This model has five basic parameters: the expansion rate of the universe (the Hubble constant), the density of matter in the universe, the density of atoms in the universe, the amplitude of the primordial density fluctuations, and how the amplitude of fluctuations varies with scale. With these five parameters, the model can fit a host of astronomical data, including measurements of positions of millions of galaxies, observations of microwave background fluctuations, measurements of distances to both nearby galaxies and distant supernovae, as well as a host of other astronomical observations. While simple, the lambda CDM model posits the existence of two novel forms of matter: cold dark matter and vacuum or dark energy.

Evolution of structure. The inflationary universe model is an attractive scenario for simultaneously explaining the homogeneity of the universe on large scales and the fluctuations in density observed on the smaller scales. In this scenario, the universe in its first few moments underwent a period of exponential expansion. This rapid expansion, which took place before nucleosynthesis, stretched out any initial density variations and eliminated any monopoles or other unwanted objects that formed before the inflationary epoch. During this inflationary epoch, quantum fluctuations generated new density fluctuations that are predicted to be the initial source of structure. The inflationary scenarios predict that the density parameter Ω should be extremely close to 1. *See* INFLATIONARY UNIVERSE COSMOLOGY.

Since different regions of space underwent different amounts of inflationary expansion, this process generated large-scale variations in the density of the universe. The inflationary universe model predicts that the variations in density should be gaussian random-field fluctuations. This implies that the statistical properties of both fluctuations in the microwave background and the fluctuations in the galaxy density distribution are characterized by its overall amplitude and its dependence on scale.

Observations of the microwave background radiation probe these conditions in the early universe. In most cosmology models, the microwave photons that are detected on Earth last interacted with electrons when the universe was one-thousandth its present size. Thus, fluctuations in the microwave background temperature reflect the density fluctuations in the early universe.

During its first 500,000 years, the early universe was a hot plasma of electrons, protons, electromagnetic and radiation (photons). The density fluctuations generated by inflation set up sound waves that propagated as long as the universe was ionized. When the universe cooled and atoms formed, radiation was able to travel freely to our telescopes and satellites. These density variations and the sound waves that they generated produced a characteristic pattern of fluctuations. Since the early universe was nearly uniform, simple linear physics describes the generation and evolution of these fluctuations. The characteristic size of these fluctuations equals the distance that sound waves can travel in 500,000 years. The amplitude of the fluctuations depends on the amount of matter and the density of atoms in the early universe.

The cosmic microwave background temperature fluctuations are very small: variations of only a few millionths of a degree. Over 25 years of searches for these fluctuations were rewarded in 1992, when the *COBE* satellite detected fluctuations in the microwave background temperature of about 2 parts in 100,000. NASA's *Wilkinson Microwave Anisotropy Probe* (*WMAP*) satellite and several ground and balloon-based measurements have been making ever more accurate measurements of these fluctuations (**Fig. 1**). With the launch of the European Space Agency's *Planck* satellite and a new generation of more sensitive ground-based experiments in development, the accuracy of these measurements will continue to rapidly improve. *See* WILKINSON MICROWAVE ANISOTROPY PROBE.

These inflationary density fluctuations grew to form the observed pattern of galaxies. Regions that were slightly overdense expanded more slowly; thus, these overdense regions collapsed and formed galaxies and clusters. On the other hand, underdense regions expanded more rapidly than the surrounding universe. These underdense regions eventually developed into voids. Astronomers using wide-field surveys such as the Sloan Digital Sky Survey and the 2dF

Galaxy Redshift Survey have now measured the positions of several hundred thousand galaxies. They have found that the clustering pattern matches the predictions of the lambda CDM model. *See* SLOAN DIGITAL SKY SURVEY.

Cosmologists now use massive computer simulations to model the clustering of matter and the formation of galaxies. These simulations are compared with observations of the large-scale distribution of galaxies and with observations of galaxy and cluster properties. Simulations based on the lambda CDM model match a wide range of astronomical observations. Remarkably, the inflationary model successfully describes the statistical properties of both galaxies and the microwave background and fits astronomical observations over a range of 10^4 in scale and 10^{12} in mass (**Fig. 2**).

Measuring the size of the universe. Measuring the distance to astronomical objects remains a great scientific challenge. Astronomers are able to measure the distances to the nearest stars by using parallax. They then must rely on empirical properties of stars to extrapolate to more distant objects. Observations of globular clusters and open clusters in the Milky Way Galaxy are needed to calibrate the luminosity–period relation for Cepheid variable stars. *See* PARALLAX (ASTRONOMY); STAR.

Primary extragalactic distance indicators are used to measure the distance to nearby galaxies such as the Magellanic Clouds and the Andromeda Galaxy. Variable stars, such as Cepheids and RR Lyrae stars, are still important tools in determining the distances to these objects. Observations of the expansion of the dust shell around Supernova 1987A in the Large Magellanic Cloud have provided an alternative method for measuring distances. All these techniques imply that the Large Magellanic Cloud is at a distance of roughly 50,000 parsecs. *See* MAGELLANIC CLOUDS; SUPERNOVA; VARIABLE STAR.

Secondary distance indicators are then used to determine the relative distance to the Virgo Cluster. For example, the brightest red and blue stars in a distant galaxy in the Virgo Cluster are assumed to have the same luminosity as the brightest stars in Andromeda. Other secondary distance indicators include the brightness of typical globular clusters or planetary nebulae. *See* VIRGO CLUSTER.

Tertiary distance indicators are then used to extrapolate from the Virgo Cluster to more distant clusters. Galaxies in the Virgo Cluster are observed to have a simple relationship, noted by B. Tully and R. Fisher, between their gas velocities and their luminosity. This relation is used to measure the relative distance to distant galaxies. Together with Doppler-shift measurements of recessional velocity, this technique yields estimates of the Hubble constant. *See* DOPPLER EFFECT.

Observations using the *Hubble Space Telescope* have significantly improved our measurements of the Hubble constant. These observations suggest that the Hubble constant is 65–80 (km/s)/Mpc. If the lambda CDM model is assumed, the *Wilkinson Microwave Anisotropy Probe's* measurements of primordial fluctuations imply a Hubble constant of 68–76 (km/s)/Mpc. Estimates of the Hubble constant using other techniques (for example, time delays from multiple imaged quasars) yield consistent results.

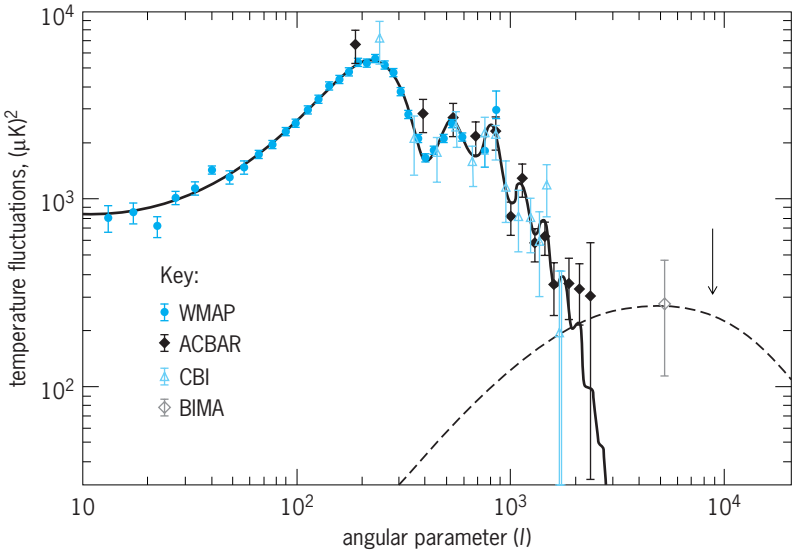

Fig. 1. Measurements of the amplitude of temperature fluctuations as a function of angular scale. The solid line is the predictions of the lambda CDM model. The data points are from the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) and three ground-based experiments. The broken line shows an estimate of the expected level of contamination from hot gas in clusters. The horizontal axis is an angular parameter, *l*, and the angular scale measured is 180°/*l*. The vertical axis shows the amplitude of the temperature fluctuations in millionths of a degree (μK). (*A. Readhead; CBI team*)

Observations of supernovae have become a powerful tool for measuring distances to distant galaxies. Type Ia supernovae, supernovae produced by the explosive burning of a white dwarf star, have nearly the same luminosity. By measuring the time evolution of these supernovae, astronomers can correct for variations in their absolute luminosities and use these standard candles to determine the distances to galaxies. By measuring the distances to supernovae as a function of their redshifts, astronomers have been able to determine the geometry and expansion history of the universe. Current data suggest that the universe is accelerating and is well fit by the basic parameters of the lambda CDM model. *See* ACCELERATING UNIVERSE.

Ground-based supernova observations have found that supernovae at redshifts of 0.3–0.8 are systematically dimmer than predicted in cosmological models without dark energy (**Fig. 3**). The observations imply that the universe is now accelerating. Hubble telescope observations of supernovae at redshifts greater than 1 (not shown in Fig. 3) find that these more distant supernovae are actually brighter than expected in a universe without dark energy. The observations of these more distant supernovae imply that, earlier in its history, the expansion of the universe was slowing and that dark energy has been dominating the expansion only during the past 5×10^9 years. The current supernova data fit the lambda CDM model well and are one of the key pieces of evidence for the existence of dark energy.

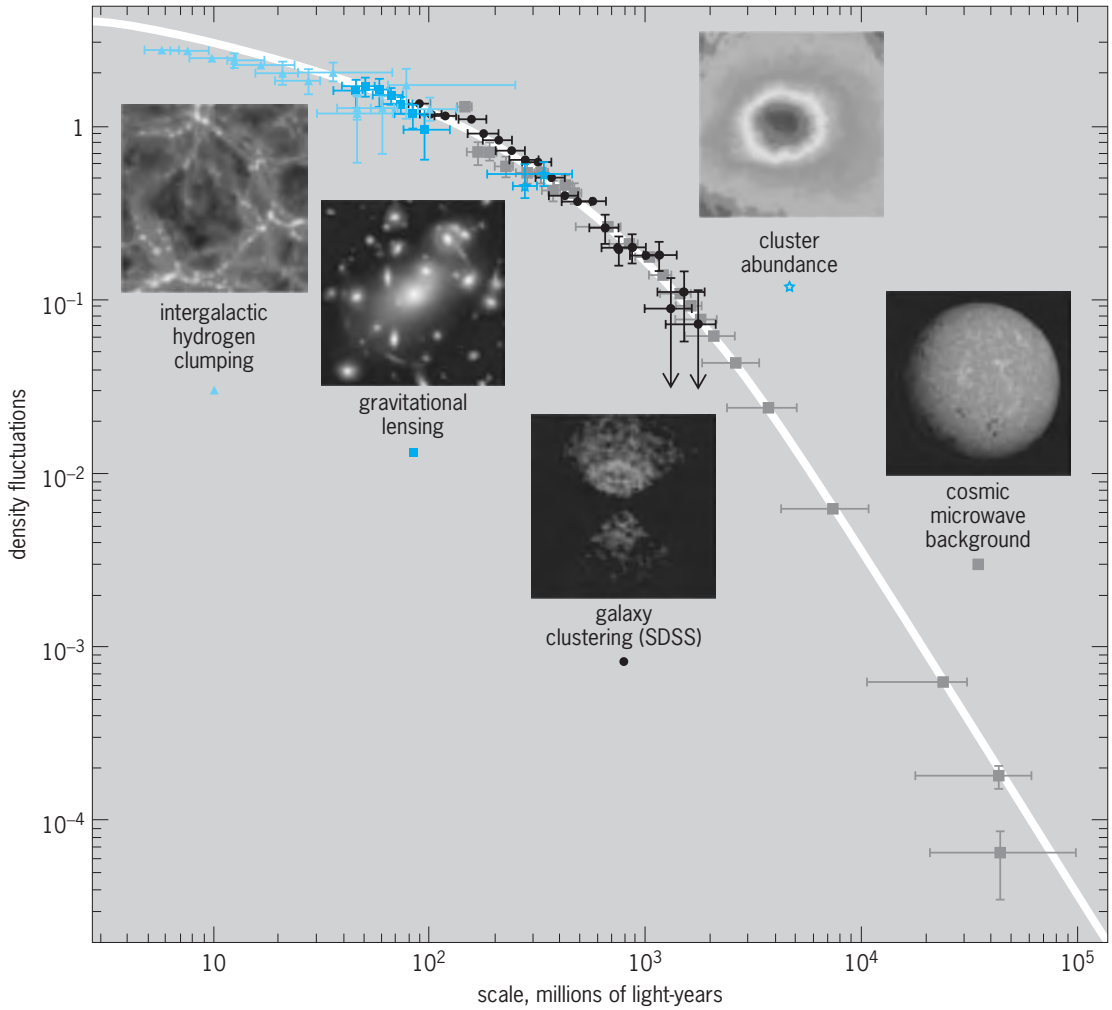

Fig. 2. Amplitude of density fluctuations as a function of scale. The insets indicate the various types of observations that measure the fluctuations and the data points that represent them in this figure. The microwave background observations measure the amplitude of fluctuations 500,000 years after the big bang; the clumping of gas measures the amplitude of fluctuations 2×10^9 years after big bang; and the galaxy and clustering measurements and gravitational lensing observations measure the amplitude of clustering in the nearby universe. The galaxy clustering measurements are based on the Sloan Digital Sky Survey (SDSS). The intergalactic hydrogen abundance is based on measurement of the absorption by its Lyman alpha line of light from distant quasars and galaxies. Cluster abundance is illustrated by an image of x-ray emission from a cluster. The amplitudes have been extrapolated to today using the lambda CDM model. The curve shows the prediction of the model. (*Max Tegmark*)

Age of the universe. The universe ought to be older than any visible star; thus, the inferred ages of the stars in the Milky Way Galaxy place a lower limit on the age of the universe. The oldest known stars are in globular clusters, dense concentrations of roughly 100,000 stars, believed to be 12–20 $\times 10^9$ years old. These age estimates are based on models of stellar evolution that predict when a star of a given mass becomes a red giant. *See* STAR CLUSTERS; STELLAR EVOLUTION.

Distant elliptical galaxies are composed almost entirely of old stars. Measurements of their spectra constrain the age of the universe as a function of redshift. Measurements of the age of the universe provided the first strong hints of the existence of dark energy. The current measurements of the Hubble constant imply that a flat universe without dark energy would be only 9×10^9 years old, much younger than some of the stars in our galaxy. In a flat universe without dark energy, light from objects with a redshift of 2 was emitted when the universe was only 1.7×10^9 years old, which is difficult to reconcile with observations showing the existence of 2.5×10^9-year-old stars in galaxies with this redshift. On the other hand, models with dark energy provide a good fit to measurements of stellar and galactic ages (**Fig. 4**).

Just as radioactivity lifetimes can be used to date archeological artifacts, nuclear dating provides an alternative independent measure of the age of the Milky Way Galaxy. By using these techniques, the minimum age of the Galaxy has been estimated to be 10×10^9 years. The oldest white dwarf stars detected are 9×10^9 years old, and the Sun is believed to be 4.5×10^9 years old. These age measurements place a firm lower limit on the age of the universe. *See* SOLAR SYSTEM; WHITE DWARF STAR.

Observations of the microwave background

fluctuations yield a direct geometric measurement of the age of the universe. The fluctuations have a characteristic size set by the distance that sound waves can travel in 500,000 years. By measuring the angular size of the typical fluctuation, cosmologists have determined an age of the universe as 13–14 × 10^9 years.

Dark-matter problem. Astronomers have been in an embarrassing situation for several decades: the amount of mass measured dynamically in galaxies by observing stellar and gas motions is roughly 10 times the mass observed in dust, gas, and known stars. This discrepancy suggests that either there is a basic flaw in newtonian physics or 90% of the mass in galaxies is in some yet unknown form, usually referred to as dark matter.

Evidence for dark matter. Some of the strongest evidence for existence of dark matter comes from radio observations of hydrogen gas and optical observations of star motions in spiral galaxies like the Milky Way Galaxy. The neutral gas and stars in the disks of galaxies are moving in nearly circular orbits. The centrifugal acceleration of the gas (and stars), v^2/r outward (where v is the velocity of motion and r is the orbit radius), must balance the gravitational acceleration of the galaxy, yielding Eq. (4), where $M(r)$

$$v^2 = \frac{GM(r)}{r} \tag{4}$$

is the mass of the galaxy within radius of the gas or stellar orbit. Thus, observations of the gas or stellar velocity are a direct measure of the mass interior to its orbit. Most of the light in a galaxy like the Milky Way Galaxy is located in the inner 15,000 parsecs. (The Sun is roughly 8000 parsecs from the galactic center.) If mass were distributed as light, the gas rotation velocity should begin to fall off outside the optical edge of the galaxy. In all optical and radio observations of spiral galaxies, the gas rotation velocity is not falling but is either flat or slowly rising. This situation implies that there is mass where there is no light.

Observations of hot x-ray-emitting gas in elliptical galaxies provide a method for measuring their mass distribution. Since the gas pressure gradients must balance the force of gravity, gas density and temperature profiles can be directly related to the elliptical galaxy mass profiles. Data from x-ray satellites suggest that the galactic dark-matter problem is ubiquitous: all galaxies, whether spiral or elliptical, dwarf or normal, seem to have halos of dark matter. *See* X-RAY ASTRONOMY.

Observations of groups and clusters of galaxies provide an alternative method of weighing galaxies. The velocities of galaxies in clusters can be used to determine the mass of the entire cluster. This technique for measuring galactic mass, first applied in the 1930s, also implies a dynamical galactic mass that vastly exceeds the mass in luminous material.

General relativity implies that mass curves space. Thus, the path of light rays moving through a cluster of galaxies is bent by the mass within the cluster (**Fig. 5**). Hence, dense clusters of galaxies can act as gravitational lenses that distort the images of galaxies behind the cluster into arcs. Measurements of these distorted arcs confirm the evidence for enormous amounts of dark matter in clusters. *See* GRAVITATIONAL LENS.

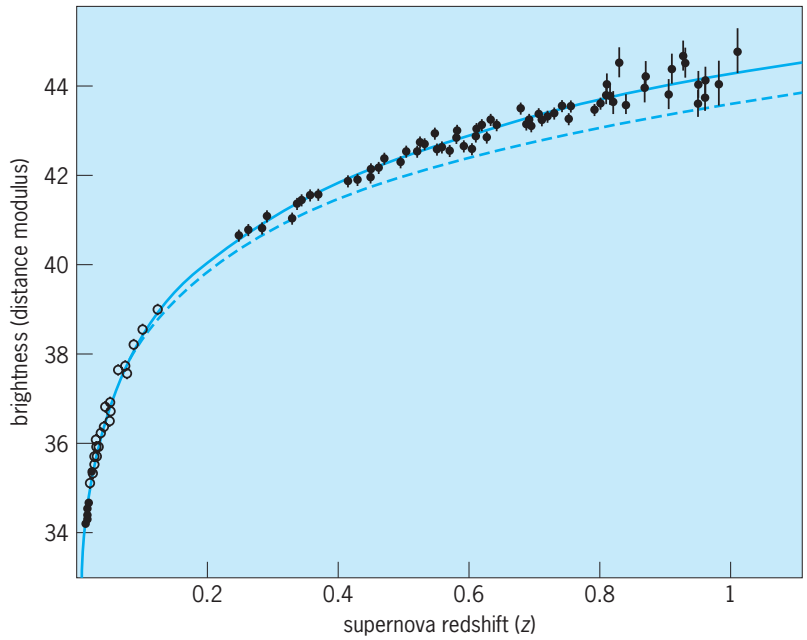

Fig. 3. Brightness of supernovae as a function of their redshift. The supernova brightnesses are measured in terms of the distance modulus, which is related to the magnitude (it is the difference between the apparent and absolute magnitudes), and are corrected for variations in supernova properties. The broken line shows the predictions for a flat universe with no dark energy. The solid line shows the predictions of the lambda CDM model. (*P. Astier; Supernova Legacy Survey*)

Candidates for dark matter. Candidates for the dark matter range in mass from microelectronvolt axions to 10^6-solar-mass black holes. Many groups of physicists and astronomers are actively searching for various candidates for the dark matter.

Baryons (such as ordinary protons and neutrons) make up most of the mass of the Earth and the Sun. Baryons are the most obvious dark-matter candidate. These baryons cannot be bound into luminous stars, nor can they be in either hot or cold gas. Cold gas can be detected through hyperfine and molecular lines. Hot gas is detectable by ultraviolet and x-ray satellites. Current observations constrain the gas mass to be much less than the mass needed to account for the dark matter. Within our galaxy, baryons could have escaped direct detection if they were bound together into clumps, either as comets, bound by atomic forces, or as planets or very low mass stars, bound by gravitational forces. However, microwave background observations provide a more direct total accounting of the composition of the universe. These observations imply that the total density in atoms is only one-sixth of the density in matter, so ordinary matter is no longer considered a likely candidate for the dark matter.

Black holes have been proposed as another possible candidate for the dark matter. If the first generation of star formation consisted of very massive stars, these objects would rapidly burn hydrogen to

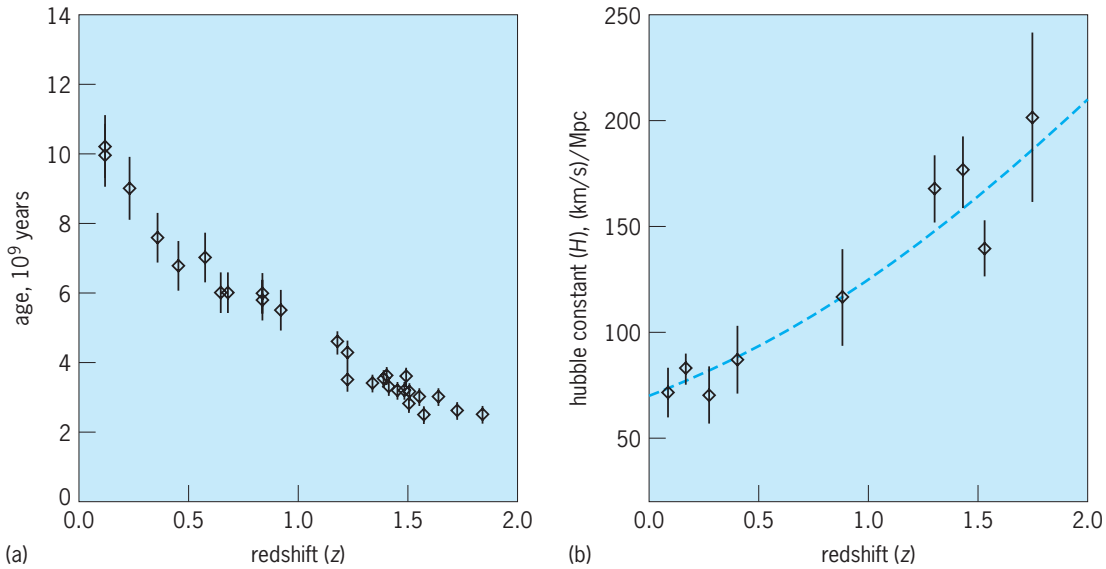

Fig. 4. Comparison of galactic ages with predictions of the lambda CDM model. (*a*) Ages of the oldest galaxies as a function of redshift. (*b*) Data points and bars show the Hubble constant inferred from the galactic ages. The curve shows the predictions of the lambda CDM model. (*J. Simon, L. Verde, R. Jimenez*)

heavier elements and then collapse to black holes. Observational constraints on element abundance require that these massive stars not lose much of the carbon, oxygen, or iron produced during their stellar evolution to the interstellar medium. Models of very massive star evolution suggest that these stars could swallow most of their mass and form a halo of black holes. This scenario is rather tentative, since very little is understood about pregalactic star formation. These black holes could also be detected in the current gravitational lensing searches. *See* BLACK HOLE.

The currently favored hypothesis is that the dark matter consists of some as yet undetected particle. Particle physicists have suggested several possible candidates for the dark matter. The most popular of these particles is predicted in an extension of standard physics called supersymmetry. This particle, named a WIMP (weakly interacting massive particle), could potentially be detected in deep underground experiments or created in particle accelerators. The upcoming generation of dark-matter experiments at CERN may be able to produce these particles in the laboratory. *See* ELEMENTARY PARTICLE; SUPERSYMMETRY; WEAKLY INTERACTING MASSIVE PARTICLE (WIMP).

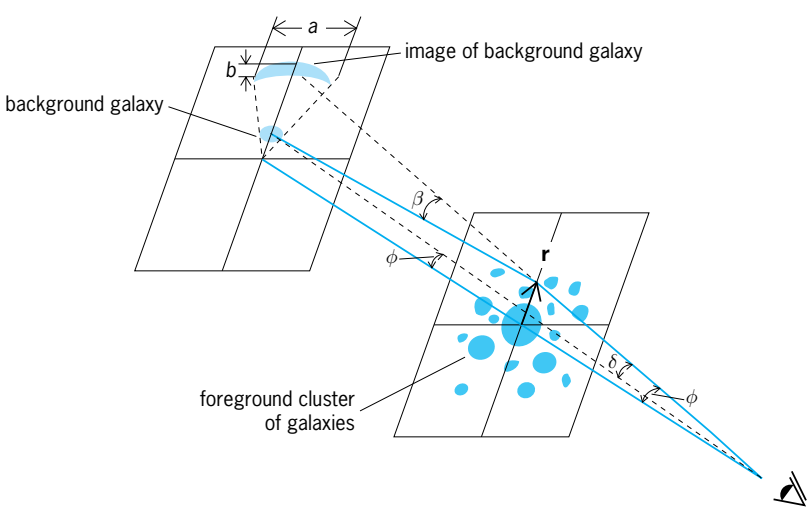

Fig. 5. Gravitational displacement and distortion (lensing) of a distant background galaxy by a compact foreground cluster of galaxies. A light ray passing the cluster plane at an impact parameter vector r is gravitationally bent through an angle β. Thus, it is seen displaced (through an angle δ) from ϕ, its true angular distance from the cluster centroid, to the larger angle ϕ. Because of its finite width, the image also is distorted into a circular arc (of length *a* and width *b*) concentric with the cluster. (*After A. Tyson, Mapping dark matter with gravitational lenses, Phys. Today, 45(6):24–32, June 1992*)

Dark energy. Measurements of distances to supernovae, galaxies, and the microwave background imply that most of the energy in the universe is not in the form of matter but dark energy. Measurements of the evolution of the clustering of galaxies confirm the existence of this dark energy and imply that it does not cluster gravitationally.

The most popular current idea for the dark energy is that it is energy associated with the vacuum. Einstein first posited the existence of this vacuum energy (or cosmological constant); however, the current model of fundamental physics cannot explain its measured value.

Other possible explanations for the dark energy include extremely weakly interacting ultralight fields or modifications of general relativity. Understanding the dark energy and measuring its properties is one of the most active areas of research not only in cosmology but also in all of physics.

Alternatives to the hot big bang. There have been a variety of models proposed as alternatives to the hot big bang. Historically, the most important alternative model was the steady-state model, which assumes a homogeneous expanding universe that is not evolving because matter is continuously being created out of the vacuum. This model has great difficulty accounting for the observed

shape of the microwave background radiation and cannot account for the observed abundance of helium.

Another alternative is the cold big bang model in which the universe began in a big bang that started at absolute zero. It has been suggested that in this model the microwave background radiation could be due to reemitted light from iron needles. This model requires special dust grains and also cannot explain the abundances of light elements.

In H. Alfvén's plasma universe, the Milky Way Galaxy is part of a finite cloud of material expanding into empty flat space. It is very difficult for this model to explain the observed uniformity of the microwave background spectrum or the light-element abundances.

E. Segal has proposed an alternative to cosmology based on general relativity. In this chronometric theory, there is a quadratic rather than a linear relationship between redshift and distance. Modern measurements of the distances to rich clusters find that these objects obey Hubble's linear relationship rather than Segal's quadratic law. Observations of the correlation between jet expansion velocity and redshift are a direct test of the chronometric hypothesis, which it fails.

David N. Spergel

Bibliography. S. Dodelson, *Modern Cosmology*, Elsevier Science, 2003; A. Liddle, *An Introduction to Modern Cosmology*, 2d ed., Wiley, 2003; J. A. Peacock, *Cosmological Physics*, Cambridge University Press, 1999; P. J. E. Peebles, *Principles of Physical Cosmology*, Princeton University Press, 1993; D. N. Spergel et al., First-Year *Wilkinson Microwave Anisotropy Probe* (*WMAP*) observations: Determination of cosmological parameters, *Astrophys. J. Suppl.*, 148:175–194, 2003.

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

5 COT-EAT

www.MHEST.com

Cotton

A fiber obtained from the cotton plant *Gossypium*, of the order Malvales. It has been used for more than 3000 years. It is the most widely used natural fiber, because of its versatility and comparatively low cost to produce and to manufacture into finished products. Cotton traditionally has been used alone in textile products, but blends with artificial fibers have become increasingly important. *See* MALVALES; NATURAL FIBER; TEXTILE.

Unlike most other fibers obtained from plants, the cotton fiber is a single elongated cell. Under the microscope, it resembles a flattened, spirally twisted tube with a rough surface. The fiber cell has many convolutions, with a collapsed inner canal (lumen). Chemically, cotton is essentially pure cellulose. In its original state, cotton contains 3–5% natural waxes and gums, mostly in the outer wall of the cells. The natural waxes on the fiber surface act as a finish which facilitates spinning. Cotton is hygroscopic and contains 6–8% moisture under normal atmospheric conditions. *See* CELLULOSE.

Cultivation and harvesting. To mature, cotton requires about 180 days of continuously warm weather with adequate moisture and sunlight; frost is harmful and may kill the plant. The ground must be thoroughly plowed and the soil pulverized. In the United States, usually in March or April, carefully selected seeds are planted in rows. The plants require systematic fertilization. When they are about 3 in. (7.5 cm) high, they are weeded and thinned. The plants begin to bloom in June or July, about 4 months after planting. Creamy white flowers appear first, change to a reddish-purple in about 2 days, and then fall off, leaving seed pods that grow to full size by August or September. The cotton capsules or bolls must be protected against the boll weevil or other insects. *See* ENTOMOLOGY, ECONOMIC; FERTILIZER; FERTILIZING; FLOWER; FRUIT; SEED.

When fully grown, the cotton plant may be 3–6 ft (1–2 m) in height. Its wide green leaves partially conceal some of the bolls until they burst and expose the fleecy white fiber, which indicates that the cotton is ready for harvesting (**Fig. 1**). Not all cotton bolls open at the same time, but the ripening period has been shortened and pickings have been reduced to one or two. When the raw cotton is harvested, it contains seeds, leaf fragments, and dirt that must be removed before baling. The cotton seeds alone constitute approximately two-thirds of the weight of the picked cotton.

Products and processing. When the bolls open, the fiber and seed, or "seed cotton," is harvested mostly by machines. The fiber, or lint as it is then called, is separated from the seed by gins. The lint is compressed into bales, covered with jute bagging, and bound with metal bands for ease of handling. Bales weigh about 500 lb (225 kg) each. The seed (except

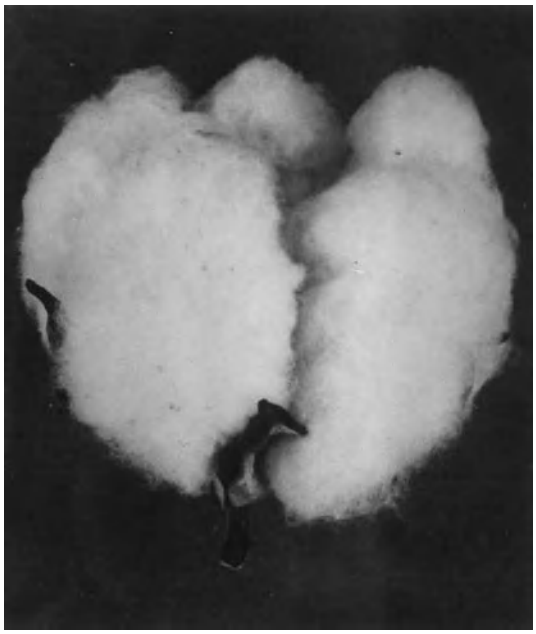

Fig. 1. Open cotton boll ready to harvest.

Fig. 2. Productive field of open cotton. (*K. Bilbrey, County Agent, Mississippi County, Arkansas*)

that portion needed for planting the next crop and for other farm uses) is transferred from the gins, usually by trucks, to oil mills for processing. After baling, the lint is sampled, graded, and sold. The bales then pass through commercial channels and eventually reach cotton mills, where they are broken open and the lint is blended, cleaned, carded, and spun into yarns for many uses. In the oil mill processing industry, the cottonseed is separated into fuzz or linters, hulls, oil, and protein cake. Each of these products is converted to several subproducts. The oil and cake are the most valuable products. The oil subproducts are largely used as human food; the protein cake, either as cracked cake or ground meal, is used as livestock feed. Hulls are also fed to livestock, and the linters are converted into chemical cellulose. *See* ANIMAL FEEDS; FAT AND OIL (FOOD).

Distribution and production. The cotton plant is one of the world's major agricultural crops. It is grown in all countries lying within a wide band around the Earth. The limits of this band in the New World are at about 37°N latitude and at about 32°S latitude. In the Old World the band spreads northward to 47°N in the Ukraine, but only to about 30°S in Africa and Australia. In addition to the effects of latitude, suitability of climate for the growth of cotton is also regulated by elevation, wind belts, ocean currents, and mountain ranges. As a result of the effects of these climatic factors, the topography of the land, the nature of the soil, and the availability of irrigation water when needed, cotton culture actually is carried out in an irregular and greatly extended world pattern. The regions of more intensive culture comprise the Cotton Belt of the United States, the northern valleys of Mexico, India, West Pakistan, eastern China, Central Asia, Australia, the Nile River Valley, East Africa, South Africa, northeastern and southern Brazil, northern Argentina, and Peru.

Stimulants to production. The United States has made three major contributions to world cotton production. These were the development of Upland cotton, the invention of the saw gin, and the development of knowledge of cotton culture.

Upland cotton arose from an annual form of *G. hirsutum* native to the plateaus of southern Mexico. This species was introduced into what is now southern United States at different times from the colonial period to the first decade of the twentieth century. The name Upland is derived from the upland country in southeastern United States, where this stock was first grown commercially. This cotton was shipped to England under the name of American Upland. From the time of its early development as a crop in the American Cotton Belt, it has been very hardy and productive, and versatile in its many uses (**Fig. 2**).

Upland production was handicapped at first by the difficulty of ginning. The seed coat has a tight covering of short fuzz hairs (**Fig. 3**). Only sawlike teeth can penetrate and remove the longer coat of lint from the seed and leave the fuzz behind with the seed coat undamaged. The saw gin, incorporating

Fig. 3. Ginned seed of Upland cotton, showing dense and tight covering of short fuzz hairs on seed coat. (*J. O. Ware, Journal of Heredity*)

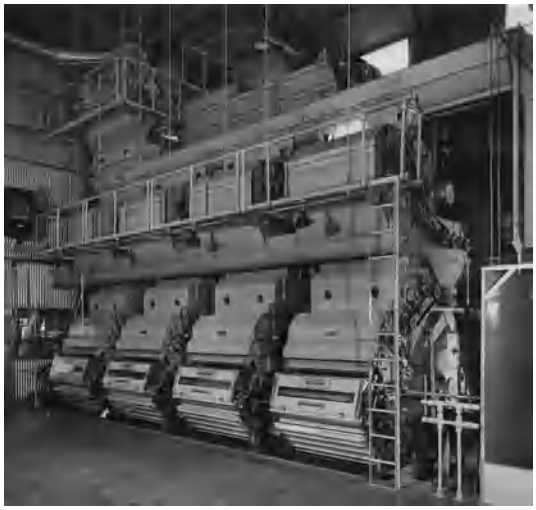

Fig. 4. Modern cotton gin, showing the four stands and the accessory equipment for cleaning and conveying the cotton. (*Murray Company of Texas*)

Fig. 5. Planting cotton with a four-row tractor-propelled planter. (*John Deere Co.*)

Fig. 6. Two-row, self-propelled cotton picker in operation. (*International Harvester Co.*)

this principle, was gradually developed into the modern gin plant (**Fig. 4**).

In the history of cotton culture, simple but efficient farm tools were developed and the mule was generally the source of farm power. The cotton fields were well tilled, the weeds controlled, the cotton carefully picked by hand, and the best seed selected for the next crop. However, when mechanical power began to replace animal power, the United States quickly mechanized both the cultural and harvesting phases of cotton production. (**Figs. 5** and **6**). Great progress in research, plant breeding, technology, and mechanics paved the way for advancements when economic conditions were ready for their adoption. Upland cotton and some of the improvements made in handling this crop spread to most cotton-growing countries as they began commercial production. Upland cotton provides about 85% of the world supply of raw cotton for factory use. *See* AGRICULTURAL SCIENCE (PLANT); BREEDING (PLANT).

The other 15% of world production is supplied from three other cultivated species: *G. barbadense, G. arboreum,* and *G. herbaceum*. The extra-long staple type includes Egyptian, Tanguis, American-Pima, and Sea Island. Egyptian is grown in Egypt and the Sudan, Tanguis in Peru, American-Pima in southwestern United States, and Sea Island in the West Indies. Sea Island cotton formerly was grown on islands and coastal areas of the southeastern United States. *Gossypium arboreum* and *G. herbaceum* are commonly referred to as Asiatic cottons and include the shortest-staple types of cultivated cottons. The growth of the former is largely confined to India and China and of the latter to Central Asia, western India, and the Near East. The saw gins usually used for Upland cotton in other countries are simpler and carry less cleaning equipment than those used in the United States. Most other countries still employ much handwork in cotton culture; most of the crop is harvested by hand.

Most commercial cottons are now annuals. However, in the tropics some races of the regular cultivated species are perennial. All cottons at one time were tropical and perennial.

Wild cotton. Besides the four cultivated species, *Gossypium* includes about 30 wild and lintless species that are widely scattered and occur mostly in sparsely covered desert and tropical areas of the world. All of the wild species may ultimately be of plant breeding interest. Some of them already have been used in crosses to improve quality values in cultivated cottons. Elton G. Nelson

Diseases. Pathogens may attack or invade the cotton plant through the roots or aboveground parts. Diseases reduce the yield potential 15 to 20%, and also have undesirable effects on quality of fiber and seed.

Root diseases are caused by soil-borne organisms. Severity of disease is proportional to the number of infective pathogen propagules in soil and to the degree of susceptibility in varieties. Minimizing losses involves using resistant varieties and practices that reduce pathogen numbers in the soil.

The seedling disease complex occurs wherever cotton is grown. Symptoms are seed rot, preemergence damping-off, postemergence damping-off, and damage to roots of plants that survive. The primary organisms are the fungi *Rhizoctonia solani, Pythium* spp., *Thielaviopsis basicola,* and *Glomerella gossypii* (**Fig. 7**), and the root-knot nematode *Meloidogyne incognita*.

Using high-quality seed treated with protectant fungicides is essential to the control of seedling

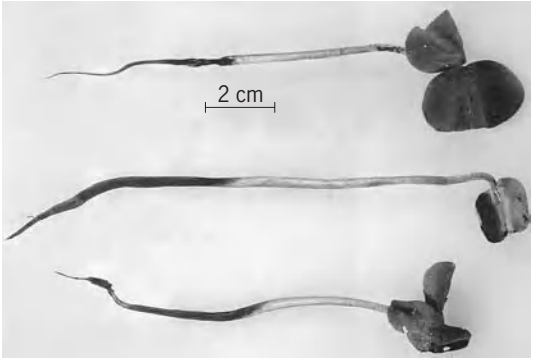

Fig. 7. Diseased cotton seedlings which have been affected by *Glomerella gossypii*.

Fig. 8. Verticillium wilt of young cotton plant.

diseases. Application of soil fungicides at planting is an additional practice that gives more effective control. Cotton varieties with resistance to seedling pathogens, preservation of seed quality, and with improved performance under early-season cool-wet conditions are now available.

Although the wilt-inducing fungus *Fusarium oxysporum* f. *vasinfectum* or the root-knot nematode *M. incognita* (Fig. 9*c*) can damage cotton in single infections, losses due to double infections are greater. Management practices effective against both organisms must therefore be used. The most economical practice is the use of resistant varieties, but it is best to use resistant varieties and rotations with root-knot–resistant crops such as small grains, corn, sorghum, or some grasses.

Verticillium wilt, caused by *Verticillium alboatrum*, occurs in the cooler regions of the world (**Fig. 8**). Effective control can be achieved with integrated management. High populations of 25 to 30 plants per meter of row help reduce losses. Too much nitrogen causes wilt to be more severe. Serious losses are usually limited to irrigated cotton, cotton in high-rainfall regions, and in poorly drained areas of fields. Reducing the amount of irrigation water by one-third will often give higher yields. The wilt fungus survives as microsclerotia in the soil and in undecomposed plant debris. Resistant crops such as barley and sorghum in rotation with cotton have been effective in reducing the inoculum potential. Many resistant varieties are available, and their use, along with other practices that reduce disease severity, gives effective control.

Root rot, caused by *Phymatotrichum omnivorum*, occurs in alkaline soils in the southwestern and western regions of the United States and in Mexico. The soil in infested areas has more sclerotia, lower bacterial populations, less organic matter, and less sodium than other soils. Control depends on reducing sclerotia production and preventing survival of the fungus between cropping seasons.

Burial of barnyard manure or residues of such legumes as Hubam clover and winter peas helps control the disease. Turning the upper 8–12 in. (20 to 30 cm) of soil with a moldboard plow a few days after harvesting a crop of cotton reduces sclerotia production. Rotation of cotton with grain crops also helps to keep numbers of sclerotia low.

Planting early in the season with fast-maturing varieties permits early boll set and maturation of the crop before soil temperatures reach levels for the pathogen to become highly active. Restoration of sodium to a level equal to that in noninfested soil prevents production of new sclerotia, and is now being used for control. Best control comes from using short-season management with practices that reduce sclerotia production and survival.

Diseases of aboveground plant parts may be caused by bacteria or fungi. These organisms may survive in undecomposed plant debris, on alternative hosts, or in and on seed. Sanitation of planting-seed and shredding and burial of residues from a cotton crop reduce the ability of foliage pathogens to survive the winter.

Bacterial blight, caused by *Xanthomonas malvacearum*, affects all plant parts and occurs wherever cotton is grown (**Fig. 9***a* and *b*). The bacterium survives in and on seed and as dried exudates on undecomposed plant debris. Chemical delinting

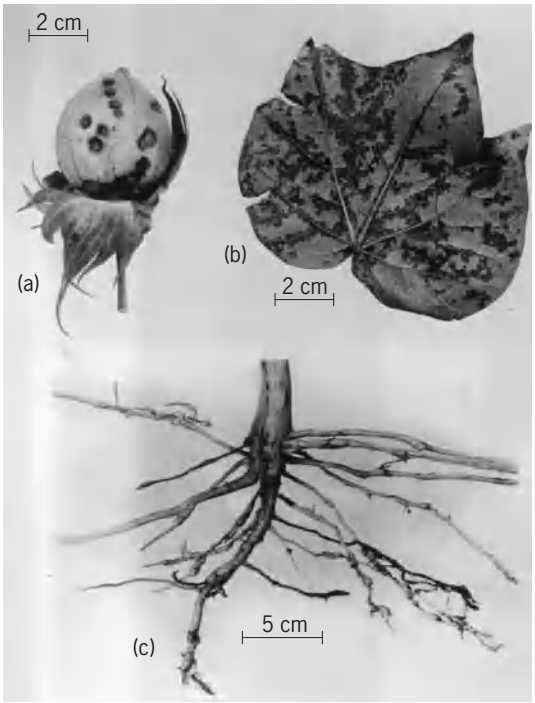

Fig. 9. Bacterial blight on cotton plant. (*a*) Bacterial blight lesions on bolls. (*b*) Bacterial blight lesions on leaves. (*c*) Root-knot galls on cotton roots.

removes the bacterium from the seed surface, but no treatment is available for eliminating internal transmission. Using seed from disease-free fields ensures no transmission. Management to ensure debris decomposition eliminates crop residue as a means of transmission. Varieties with high horizontal resistance to the 18 races of the pathogen are available for many cotton-growing regions of the world.

Boll rot may be caused by a number of fungi, bacteria, or yeasts. Some, such as *G. gossypii* and *X. malvacearum*, are primary pathogens. Many gain entrance through natural openings such as boll sutures and nectaries; others enter through wounds caused by insects or primary pathogens.

Control hinges on using practices that reduce inoculum densities and promote dryness within the leaf canopy. Varieties with okra-shaped leaves and frego bracts (which are strap-shaped leaves) are significant in reducing boll rot. Sanitation should be practiced to reduce transmission on seed and survival of organisms on debris. *See* PLANT PATHOLOGY.

Luther S. Bird

Bibliography. M. E. Selsom, *Cotton*, 1982; G. M. Watkins, *A Compendium of Cotton Diseases*, American Phytopathology Society, 1981.

Cotton effect

The characteristic wavelength dependence of the optical rotatory dispersion curve or the circular dichroism curve or both in the vicinity of an absorption band.

When an initially plane-polarized light wave traverses an optically active medium, two principal effects are manifested: a change from planar to elliptic polarization, and a rotation of the major axis of the ellipse through an angle relative to the initial direction of polarization. Both effects are wavelength dependent. The first effect is known as circular dichroism, and a plot of its wavelength (or frequency) dependence is referred to as a circular dichroism (CD) curve. The second effect is called optical rotation and, when plotted as a function of wavelength, is known as an optical rotatory dispersion (ORD) curve. In the vicinity of absorption bands, both curves take on characteristic shapes, and this behavior is known as the Cotton effect, which may be either positive or negative (**Fig. 1**). There is a Cotton effect associated with each absorption process, and hence a partial CD curve or partial ORD curve is associated with each particular absorption band or process. *See* OPTICAL ROTATORY DISPERSION; POLARIZED LIGHT.

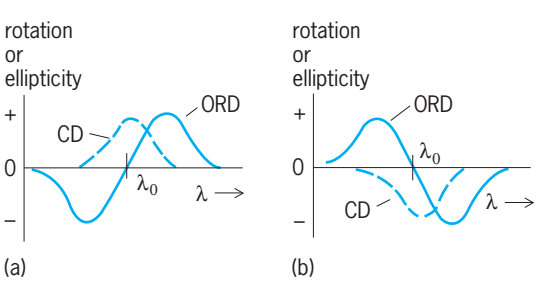

Fig. 1. Behavior of the ORD and CD curves in the vicinity of an absorption band at wavelength λ_0 (idealized). (*a*) Positive Cotton effect. (*b*) Negative Cotton effect.

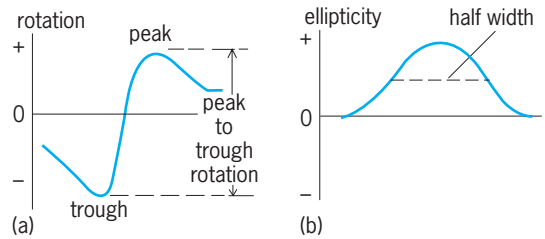

Fig. 2. Curves used to determine relative rotatory intensities. (*a*) Partial ORD curve. (*b*) Partial CD curve.

Measurements. Experimental results are commonly reported in either of two sets of units, termed specific and molar (or molecular). The specific rotation $[\alpha]$ is the rotation in degrees produced by a 1-decimeter path length of material containing 1 g/ml of optically active substance, and the specific ellipticity θ is the ellipticity in degrees for the same path length and same concentration. Molar rotation $[\varphi]$ (sometimes $[M]$) and molar ellipticity $[\theta]$ are defined by Eqs. (1) and (2). For comparisons among

$$[\varphi] = [\alpha]M/100 \tag{1}$$

$$[\theta] = \theta M/100 \tag{2}$$

different compounds, the molar quantities are more useful, since they allow direct comparison on a mole-for-mole basis.

The ratio of the area under the associated partial CD curve to the wavelength of the CD maximum is a measure of the rotatory intensity of the absorption process. Moreover, for bands appearing in roughly the same spectral region and having roughly the same half-width (**Fig. 2**), the peak-to-trough rotation of the partial ORD curve is roughly proportional to the wavelength-weighted area under the corresponding partial CD curve. In other words, relative rotatory intensities can be gaged from either the pertinent partial ORD curves or pertinent partial CD curves. A convenient quantitative measure of the rotatory intensity of an absorption process is the rotational strength. The rotational strength R_i of the ith transition, whose partial molar CD curve is $[\theta_i(\lambda)]$, is given by relation (3).

$$R_i \approx 6.96 \times 10^{-43} \int_0^\infty \frac{[\theta_i(\lambda)]}{\lambda} d\lambda \tag{3}$$

Molecular structure. The rotational strengths actually observed in practice vary over quite a few orders of magnitude, from $\sim 10^{-38}$ down to 10^{-42} cgs and less; this variation in magnitude is amenable to stereochemical interpretation. In this connection it is useful to classify optically active chromophores, which are necessarily dissymmetric, in terms of two limiting types: the inherently dissymmetric chromophore, and the inherently symmetric but dissymmetrically perturbed chromophore. *See* OPTICAL ACTIVITY.

A symmetric chromophore is one whose inherent geometry has sufficiently high symmetry so that the isolated chromophoric group is superimposable on its mirror image, for example, the carbonyl group $\rangle C{=}O$. The transitions of such a chromophore can become optically active, that is, exhibit a Cotton effect, only when placed in a dissymmetric molecular environment. Thus, in symmetrical formaldehyde, $H_2C{=}O$, the carbonyl transitions are optically inactive; in ketosteroids, where the extrachromophoric portion of the molecule is dissymmetrically disposed relative to the symmetry planes of the $\rangle C{=}O$ group, the transitions of the carbonyl group exhibit Cotton effects. In such instances the signed magnitude of the rotational strength will depend both upon the chemical nature of the extrachromophoric perturbing atoms and their geometry relative to that of the inherently symmetric chromophore. In a sense, the chromophore functions as a molecular probe for searching out the chemical dissymmetries in the extrachromophoric portion of the molecule.

The type of optical activity just described is associated with the presence of an asymmetric carbon (or other) atom in a molecule. The asymmetric atom serves notice to the effect that, if an inherently symmetric chromophore is present in the molecule, it is almost assuredly in a dissymmetric environment, and hence it may be anticipated that its erstwhile optically inactive transitions will exhibit Cotton effects. Moreover, the signed magnitude of the associated rotational strengths may be interpreted in terms of the stereochemistry of the extrachromophoric environment, as compared with that of the chromophore. But an asymmetric atom is not essential for the appearance of optical activity. The inherent geometry of the chromophore may be of sufficiently low symmetry so that the isolated chromophore itself is chiral, that is, not superimposable on its mirror image, for example, in hexahelicene.

In such instances the transitions of the chromophore can manifest optical activity even in the absence of a dissymmetric environment. In addition, it is very often true that the magnitudes of the rotational strengths associated with inherently dissymmetric chromophores will be one or more orders of magnitude greater ($\sim 10^{-38}$ cgs, as opposed to $<10^{-39}$ cgs) than those associated with inherently symmetric chromophores. In the spectral regions of the transitions of the inherently dissymmetric chromophore, it will be the sense of handedness of the chromophore itself that will determine the sign of the rotational strength, rather than perturbations due to any dissymmetric environment in which the inherently dissymmetric chromophore may be situated.

The sense of handedness of an inherently dissymmetric chromophore may be of considerable significance in determining the absolute configuration or conformations of the entire molecule containing that chromophore. Accordingly, the absolute configuration or conformation can often be found by focusing attention solely on the handedness of the chromophore itself. For example, in the chiral molecule

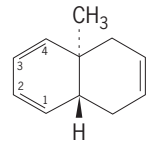

Fig. 3. Structural formula of (+)-*trans*-9-methyl-1,4,9,10-tetrahydronaphthalene.

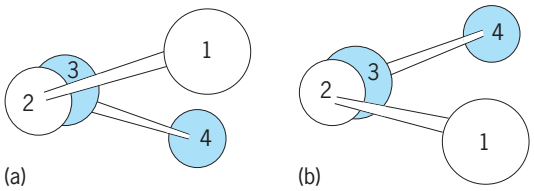

Fig. 4. Schematic representation of the twisted diene chromophore showing the two possible handednesses; the numbering is as indicated in Fig. 3. (*a*) Right-handed. (*b*) Left-handed.

shown in **Fig. 3** there is a one-to-one correspondence between the sense of helicity of the nonplanar diene chromophore present and the absolute configuration at the asymmetric carbon atoms. Hence there exists a one-to-one correspondence between the handedness of the diene and the absolute configuration of the molecule. Since it is known that a right-handed diene helix (**Fig. 4**) associates a positive rotational strength with the lowest diene singlet transition in the vicinity of 260 nanometers, by examination of the pertinent experimental Cotton effect (positive), the absolute configuration of the molecule is concluded to be as shown.

Other examples of inherently dissymmetric chromophores are provided by the helical secondary structures of proteins and polypeptides. Here the inherent dissymmetry of the chromophoric system arises through a coupling of the inherently symmetric monomers, which are held in a comparatively fixed dissymmetric disposition relative to each other through internal hydrogen bonding. The sense of helicity is then related to the signs of the rotational strengths of the coupled chromophoric system. The destruction of the hydrogen bonding destroys the ordered dissymmetric secondary structure, and there is a concomitant decrease in the magnitude of the observed rotational strengths. Albert Moscowitz

Bibliography. E. Charney, *The Molecular Basis of Optical Activity*, 1975, reprint 1985; S. F. Mason, *Molecular Optical Activity and the Chiral Discriminations*, 1982.

Coulomb excitation

Nuclear excitation caused by the time-dependent long-ranged electric field acting between colliding nuclei. Theoretically, the Coulomb force between the positively charged colliding nuclei is well understood, and the interaction is exactly calculable. Coulomb excitation usually is the dominant reaction in nuclear scattering, and even occurs at low bombarding energies where the separation of

the nuclei is sufficiently large that the short-ranged nuclear force does not act. *See* COULOMB'S LAW.

Coulomb excitation plays a vital role in probing the response of both shape and volume collective modes of motion as well as the interplay of single-particle degrees of freedom of the nuclear many-body system. The goal of this work is to develop better models of nuclear structure and to elucidate the underlying nuclear force.

Collective nuclear modes of motion. The residual interaction between nucleons bound in the nucleus leads to coherent collective modes of motion of the nuclear surface and volume. Such coherent motion of many nucleons in the nucleus is of considerable interest in understanding the physics of many-body quantal systems. Collective rotation and vibration of deformed shapes of the nuclear surface is a dominant and ubiquitous feature of the low-lying structure in nuclei. Quite separate from these low-lying surface collective modes are volume collective modes that lead to high-frequency giant resonances at 10–30 MeV in excitation energy. The nuclear charge of the nucleons involved in this coherent motion produces considerably enhanced electromagnetic properties for collective nuclear states. The electric multipole moments of the nuclear states are a direct and sensitive measure of nuclear deformation. For example, the electric quadrupole moments are a direct measure of quadrupole deformation, such as football-shaped deformation; the electric octupole moments are sensitive to octupole deformation, such as pear shapes; while electric hexadecapole moments are sensitive to more complicated hexadecapole-shaped deformation. *See* GIANT NUCLEAR RESONANCES; NUCLEAR MOMENTS.

The considerable importance of Coulomb excitation lies in the fact that it is the preeminent probe of collective-shape degrees of freedom in nuclei. That is, Coulomb excitation selectively populates modes of motion of the collective shape with cross sections that are a direct and sensitive measure of the electric moments, and these electric moments can be measured with considerable precision since the electromagnetic interaction is exactly calculable.

One-step and multistep excitation. Coulomb excitation was first observed in the 1950s and played a pivotal role in showing that many nuclei have prolate deformation like a football. The initial experiments used beams of protons or alpha particles for which the electromagnetic interaction is weak and only simple one-step excitation occurs. The introduction of high-atomic-number (high-Z) projectiles dramatically advanced exploitation of Coulomb excitation. With such projectiles, the electromagnetic excitation probability for surface modes can approach unity and multistep excitation dominates, leading to the population of excited states with up to 34 units of angular momentum. Such multistep Coulomb excitation can determine the electromagnetic properties of many low-lying collective states in a nucleus, making it a powerful probe of collective motion in nuclear structure. *See* ALPHA PARTICLES; PROTON.

Technical advances. Three technical advances have greatly enhanced the power of Coulomb excitation as a probe of nuclear structure. The first is the development of heavy-ion accelerators that can provide copious beams of stable nuclear isotopes throughout the periodic table, including high-Z projectiles such as uranium ($Z = 92$). Also, radioactive beam facilities are being built that expand the arsenal of beams to include unstable nuclear species. The second advance is the fabrication of arrays of large intrinsic-germanium high-resolution gamma-ray detectors that surround the target. These have high detection efficiency and extremely high sensitivity for resolving the gamma rays emitted during the subsequent decay of excited states populated by multiple Coulomb excitation. In addition, large-solid-angle arrays of heavy-ion detectors are used to detect the coincident scattered ions and to determine unambiguously the nuclei excited and the scattering trajectories. The third advance is the development of a Coulomb excitation least-squares search computer code that makes it possible to extract the hundreds of electromagnetic matrix elements that couple the many states excited in multiple Coulomb excitation.

These advances allow population and study of complete sets of states for low-lying collective bands in a nucleus. Coulomb excitation has allowed detailed mapping of the collective-shape degrees of freedom in nuclei. The moments of inertia of collective rotation bands are derived from the excitation energies, while the nuclear shapes are derived from the electric moments. *See* GAMMA-RAY DETECTORS; PARTICLE ACCELERATOR.

Studies of low-lying modes. Coulomb excitation has produced a wealth of information on low-lying collective-shape degrees of freedom in nuclei. It has allowed study of rotational bands up to high angular momentum in nuclei that are not easily populated by other reaction mechanisms, such as neutron-rich stable nuclei, transuranic nuclei, and radioactive neutron-rich nuclei produced at radioactive beam facilities. Such studies have shown that collective motion is richer than early theoretical models had suggested. In many nuclei, Coulomb excitation has identified complete sets of states in rotational bands that result from rotation of football shapes with axis ratios of about 1.5 to 1. These strongly deformed prolate quadrupole shapes have electric quadrupole transition strengths that are over 200 times greater than those produced by a single proton. Other collective nuclear states have been found corresponding to the rotation and vibration of nearly oblate quadrupole deformed shapes, but where all three spatial axes of the nuclear shape differ in length. Collective states attributed to both one and two units of quadrupole or octupole vibration have been discovered. More complicated motion, such as bands of states corresponding to pear-shaped octupole vibration about rotating prolate deformed shapes, have been studied. Coexistence of rotational-vibrational bands having very different deformation also has been discovered in certain nuclei. These

observations are being used to refine models of nuclear structure.

Studies of giant resonances. The above studies involve Coulomb excitation of low-lying rotational and vibrational collective surface modes. Scattering of much faster heavy ions at very small scattering angles can lead to distances of closest approach that still are large enough to ensure that the interaction is dominated by the electromagnetic interaction. The shorter electromagnetic impulse in such fast Coulomb excitation makes it possible to excite the high-frequency collective volume modes, that is, giant resonances. Fast Coulomb excitation has been used to map giant resonances corresponding to both in-phase and out-of-phase vibrations of the protons and neutrons as well as double-phonon giant resonances. These studies elucidate the interplay of collective and single-particle degrees of freedom in nuclear structure. *See* NUCLEAR STRUCTURE; SCATTERING EXPERIMENTS (NUCLEI). Douglas Cline

Bibliography. K. Alder and A. Winther, *Electromagnetic Excitation: Theory of Coulomb Excitation with Heavy Ions*, 1975; T. Aumann, P. F. Bortignon, and H. Emling, Multiphonon giant resonances in nuclei, *Annu. Rev. Nucl. Part. Sci.*, 48:351–399, 1998; D. Cline, Collective modes studied by Coulomb excitation, *Acta Physica Polonica B*, 30:1291–1308, 1999; D. Cline, Nuclear shapes studied by Coulomb excitation, *Annu. Rev. Nucl. Part. Sci.*, 36:683–716, 1986.

Coulomb explosion

A process in which a molecule moving with high velocity strikes a solid and the electrons that bond the molecule are torn off rapidly in violent collisions with the electrons of the solid; as a result, the molecule is suddenly transformed into a cluster of charged atomic constituents that then separate under the influence of their mutual Coulomb repulsion. The initial velocity of the molecule is typically greater than 3×10^6 ft/s (10^6 m/s), and it takes on the order of 10^{-17} s for electrons to be torn off the molecule. Typically, it takes about 10^{-15} s for the initial Coulomb potential energy of the cluster to be converted into kinetic energy as the charged fragments recede from one another. *See* COULOMB'S LAW.

Coulomb explosions are most commonly studied using a particle accelerator, normally employed in nuclear physics research (Van de Graaff generator, cyclotron, and so forth), to produce a beam of fast molecular ions that are directed onto a solid-foil target. The Coulomb explosion of the molecular projectiles begins within the first few tenths of a nanometer of penetration into the foil, continues during passage of the projectiles through the foil, and runs to completion after emergence of the projectiles into the vacuum downstream from the foil. Detectors located downstream make precise measurements of the energies and charges of the molecular fragments together with their angles of emission relative to the beam direction. The Coulomb explosion causes the fragment velocities to be shifted in both magnitude and direction from the beam velocity. The corresponding shifts in energy and angle are small, but if the foil target is thin (approximately 10 nm) and of light material (for example, carbon), the blurring effects of energy-loss straggling and multiple scattering in the foil can be kept small relative to the Coulomb explosion effects. *See* CHARGED PARTICLE BEAMS; PARTICLE ACCELERATOR.

Consider a beam of 3-MeV HeH^+ ions incident on a 10-nm-thick carbon foil. Upon striking the foil, each projectile produces a 2.4-MeV alpha particle and a 600-keV proton separated by about 80 picometers (the bond length for HeH^+). The Coulomb explosion causes the separation to grow to about 100 pm during traversal of the foil. Downstream, the fragments achieve asymptotic shifts in energy and angle that are determined by the initial orientation of the molecule. The maximum energy shift (± 8.4 keV) is obtained when the internuclear vector in the projectile is parallel to the beam direction. There is then no angular shift. The sign of the energy shift for a given fragment depends on whether it is leading or trailing its partner. If the internuclear vector is perpendicular to the beam, the maximum angular shift (0.4° for the proton, 0.1° for the alpha particle) is achieved, and there is no shift in energy. A joint distribution in energy and angle of protons from similar dissociations (with a 19.5-nm-thick foil) is shown in the **illustration**. The width of the "rim" reflects variations in the internuclear distance of the incident molecular ions due to their vibrational motion.

Coulomb explosion experiments of this type serve two main purposes. First, they yield valuable information on the interactions of fast ions with solids. For example, it is known that a fast ion generates a polarization wake that trails behind it as it traverses a solid. This wake can be studied in detail by using diatomic molecular-ion beams, since the motion of a trailing fragment is influenced not only by the Coulomb explosion but also by the wake of its partner (rather like the way a boat is affected by the wake of another in front of it). The nonuniform distribution of intensity around the ring in the illustration can be accounted for quantitatively in terms of forces due to the polarization wakes. Second, Coulomb-explosion

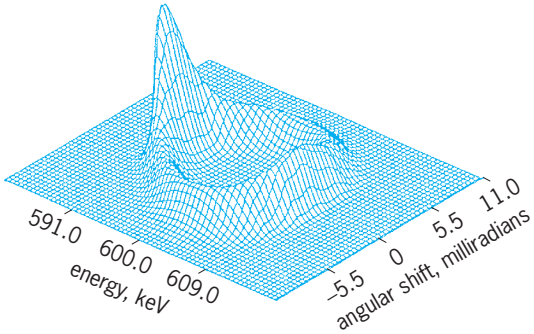

Joint distribution in energy and angle for protons from the dissociation of 3-MeV HeH^+ ions incident on a 19.5-nm-thick carbon foil.

techniques can be used to determine the stereochemical structures of molecular-ion projectiles. For example, with this method it was demonstrated experimentally for the first time that the H_3^+ molecule is equilateral-triangular. *See* ELECTRON WAKE; MOLECULAR STRUCTURE AND SPECTRA; STEREOCHEMISTRY.

Donald S. Gemmell

Coulomb's law

For electrostatics, Coulomb's law states that the direct force F of point charge q_1 on point charge q_2, when the charges are separated by a distance r in free space, is given by $F = k_0 q_1 q_2 / r^2$, where k_0 is a constant of proportionality whose value depends on the units used for measuring F, q, and r. It is the basic quantitative law of electrostatics. In the International System (SI) of units, $k_0 = 1/(4\pi\epsilon_0)$, where ϵ_0 is called the permittivity of empty space and has the value 8.85×10^{-12} farad/m. Thus, Coulomb's law is as in the equation below where q_1 and q_2 are expressed

$$F = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r^2}$$

in coulombs, r is expressed in meters, and F is given in newtons. *See* ELECTRICAL UNITS AND STANDARDS.

The direction of F is along the line of centers of the point charges q_1 and q_2. The force is one of attraction if the charges are opposite in sign and one of repulsion if the charges have the same sign. For a statement of Coulomb's law as applied to point magnet poles. *See* MAGNETOSTATICS.

Experiments have shown that the exponent of r in the equation is very accurately the number 2. Ernest (Lord) Rutherford's experiments, in which he scattered alpha particles by atomic nuclei, showed that the equation is valid for charged particles of nuclear dimensions down to separations of about 10^{-12} cm. Nuclear experiments have shown that the forces between charged particles do not obey the equation for separations smaller than this.

The direct force that one charged particle exerts on another is unaffected by the presence of additional charge and, in any electrostatic system, the equation gives this direct force between q_1 and q_2 under any conditions of charge configuration, including that in which intervening and surrounding matter is present and the molecules of the matter are polarized so that their charges contribute to this configuration. The total force on any one charge, say q_1, is the vector sum of the separate direct forces on q_1 due to q_2, q_3, q_4, and so on, each force computed separately by use of the equation as if all other charges were absent.

The permittivity ϵ of a medium is defined by $\epsilon = \epsilon_r \epsilon_0$, where ϵ_r is the relative permittivity of the medium. It was formerly known also as the relative dielectric constant or specific inductive capacity. *See* PERMITTIVITY.

If two free point charges q_1 and q_2 are immersed in an infinite homogeneous isotropic dielectric, the total force on one of them, say q_1, is given by $F = q_1 q_2/(4\pi\epsilon r^2)$ and the use of ϵ (in place of ϵ_0) takes proper account of the forces on q_1 due to the polarization charges of the dielectric molecules. *See* ELECTRIC CHARGE; ELECTROSTATICS.

Ralph P. Winch

Bibliography. B. I. Bleaney and B. Bleaney, *Electricity and Magnetism*, 3d ed., 1989; Coulomb's Law Committee, *Amer. J. Phys.*, 18:6–11, 1950; E. M. Pugh and E. W. Pugh, *Principles of Electricity and Magnetism*, 2d ed., 1976; E. M. Purcell, *Electricity and Magnetism*, Berkeley Physics Course, vol. 2, 2d ed., 1985.

Coulometer

Electrolysis cell in which a product is obtained with 100% efficiency as a result of an electrochemical reaction. The quantity of electricity, that is, the number of coulombs of electricity (Q), can be determined very accurately by weighing the product that is deposited on an electrode in the course of the electrochemical reaction. The relationship between the weight of the product formed in the coulometer and the quantity of electricity used is given by Faraday's laws of electrolysis. When a constant current of i amperes flows through the electrolyte in the coulometer for t seconds, the number of coulombs passed is given by Eq. (1). If the current varies in the course of the

$$Q = it \tag{1}$$

electrolysis, the simple current-time product in Eq. (1) is replaced by the current-time integral, Eq. (2). When Q coulombs of electricity are passed

$$Q = \int_0^t i\,dt \tag{2}$$

through the electrolyte, the weight in grams of the material that is deposited on the electrode (w) is given by Eq. (3), where n is the number of electrons

$$w = \frac{QM}{Fn} \tag{3}$$

transferred per mole of material deposited, M is its molecular weight, and F is the Faraday constant, $96{,}487 \pm 1.6$ coulombs.

Equation (3) is fundamental in coulometry and is a mathematical statement of Faraday's laws. This equation is used for the accurate determination of Q, the current-time integral, by weighing or measuring a product that is formed at an electrode by an electrochemical reaction that occurs with 100% current efficiency. The electrolysis cell that is used for this purpose is a coulometer.

Only a few electrode reactions proceed with the 100% current efficiency that is required for the use of Eq. (3). The deposition of silver or copper (in a silver or copper coulometer), the evolution of oxygen and hydrogen (in a gas coulometer), and the oxidation of iodide to iodine (in an iodine coulometer) are examples of electrode reactions that have been successfully employed. One coulomb of electricity will deposit 1.1180 mg of silver at the cathode in

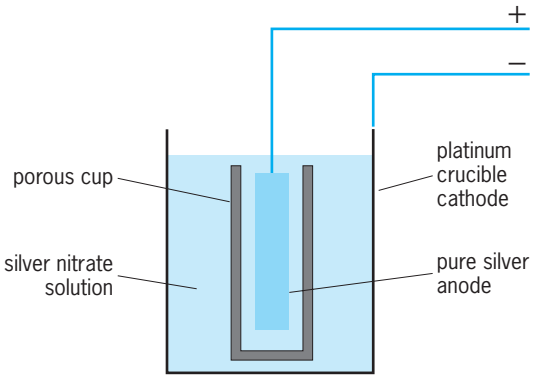

Fig. 1. Silver coulometer. The porous cup catches particles of silver that fall off the anode.

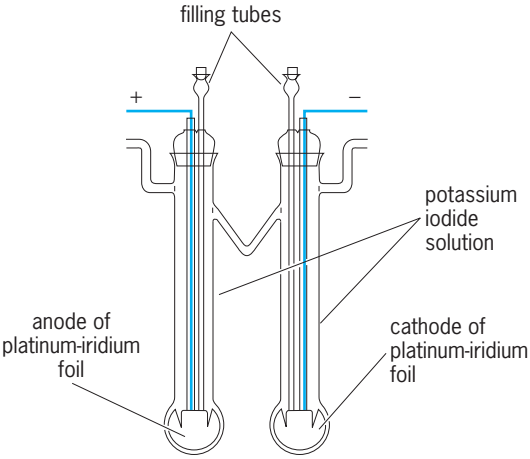

Fig. 2. Iodine coulometer.

a silver coulometer (**Fig. 1**) or liberate 1.315 mg of iodine at the anode in an iodine coulometer (**Fig. 2**). Although these classical chemical coulometers are capable of measuring the quantity of electricity with high precision and accuracy, their use is time-consuming and inconvenient; and they have been largely replaced by operational amplifier integrator circuits or digital circuits that display in a direct readout the number of coulombs passed during electrolysis. *See* ELECTROCHEMICAL EQUIVALENT; ELECTROLYSIS. Quintus Fernando

Bibliography. S. A. Borman (ed.), *Instrumentation in Analytical Chemistry*, 1982; B. G. Liptak, *Analytical Instrumentation*, 1994; G. W. Milner and G. Phillips, *Coulometry in Analytical Chemistry*, 1968.

Countercurrent exchange (biology)

Engineers have known for decades that efficient, almost complete heat or other exchange could be achieved between two fluids flowing in opposite directions in separate tubes. Such countercurrent systems have been "invented" numerous times by living organisms for all types of exchange function. They are most commonly found in the circulatory, respiratory, and excretory (kidney) systems, serving in heat, oxygen, and ion exchange. Biological countercurrent systems can be classified into two main types: downhill exchanges and hairpin multipliers. In both cases, the basic mechanism is the same—exchange of substance between fluids flowing in opposite directions—but the consequences are very different.

Downhill exchanges. These countercurrent systems are commonest in the circulatory system where their morphological structure is a rete mirabile (a wonderful net) of closely apposed sets of small arteries and veins. They are also found in gills of fish and in the minute air tubules of the avian lung. The principle of downhill exchanges is simple, as shown in **Fig. 1**. Fluids flow in opposite directions in separate tubes with the possibility of exchange, for example, heat flow or diffusion of oxygen, between them. The fluid entering one tube is warmest at that end, while that entering the second tube is coolest at the other end. Heat flows from higher to lower temperature. As heat flows from the warmer to the cooler tube, the fluid in the warmer tube cools down slightly and moves down along the tube. But as the slightly cooler fluid has moved further, it comes into contact with still cooler fluid in the second tube, and additional heat flow can occur. Thus, as the warmer fluid flows down the tube, it constantly loses heat, but always comes into contact with even cooler fluid in the second tube; the reverse is true for the initially cool fluid as it receives heat and warms up. Although the temperature differential between the two fluids is small at any point along the length of the countercurrent system, almost all the heat contained in the warmer tube is transferred to the cooler tube. Exchange of heat or oxygen occurs by passive diffusion. Most of the heat that entered the countercurrent system at one end leaves the system at the same end.

Retia of blood vessels thus serve as thermal isolating mechanisms within the body. They are found in appendages of mammals and birds (for example,

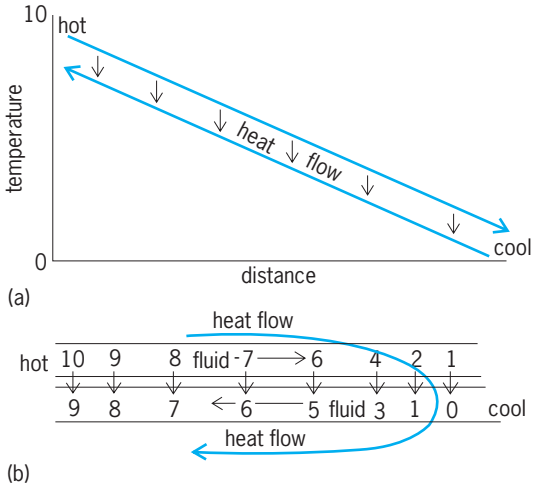

Fig. 1. Downhill exchange countercurrent system. (*a*) Graph showing the relationship between distance and temperature in a downhill system. (*b*) Schematic showing the two tubes carrying hot fluid and cool fluid and the mechanism by which heat exchange takes place.

whale flippers, the tail of beavers, and legs of gulls) to prevent excessive heat loss from these uninsulated parts. In reverse, masses of warm muscles in rapidly swimming fish, such as mackerel, tuna, and the mako shark, are isolated from the rest of the body and the gills, where heat loss would occur, by sets of retia. In desert mammals a rete located between the veins draining evaporating (that is, cooling) surfaces and the carotid artery cools the arterial blood before it reaches the brain.

Downhill exchange systems in the gills of fish and in the air tubules of birds permit maximum exchange of oxygen from the environment into the blood. Blood in respiratory capillaries flows against the water or air current and thus can pick up most of the oxygen contained in the external fluid. The advantage of downhill exchangers is that they achieve greater efficiency without extra energy cost simply by arranging flow in a countercurrent rather than in a concurrent fashion.

Hairpin multipliers. These exchange systems take their name from the structure of the tubes, which have a hairpin turn between the afferent (descending) and the efferent (ascending) limbs. Hairpin countercurrent systems are found in the nephron (the loop of Henle) of the kidney and in the capillary system of the gas gland in the swim bladder of many fish. In contrast to downhill systems, which operate by passive transport, hairpin multipliers must employ active transport of materials. These are always materials pumped out of the efferent limb of the system.

Operation of a hairpin multiplier, such as the loop of Henle, is shown in **Fig. 2**. As the kidney filtrate flows down the afferent limb and up the efferent limb, sodium ions (Na^+) are actively transported out of the efferent limb. The sodium diffuses back into the afferent limb and is carried once again around the hairpin turn, together with additional sodium that constantly enters the system via the afferent limb. Continual active transport of sodium along the entire length of the efferent limb and its diffusion back into the afferent limb will result in the accumulation of sodium at the bottom of the loop and in the surrounding interstitial fluid, and in the establishment of a steep osmotic gradient from the top to the bottom of the hairpin system. Yet at no point within this multiplier mechanism are individual cells exposed to excessive osmotic pressure. The steep gradient of osmotic pressure is used for final concentration of the urine as it passes down the collecting ducts. As water is drawn from the urine by the osmotic pressure of the interstitial fluid, the more concentrated urine passes into an area of even greater osmotic pressure—an example of downhill exchange. Thus, both types of countercurrent systems operate in the kidney to recover water from the urine. Desert mammals improve their water-conserving abilities by simply increasing the length of the loop of Henle; hence the strength of the osmotic pressure at the lower end of the hairpin is also increased.

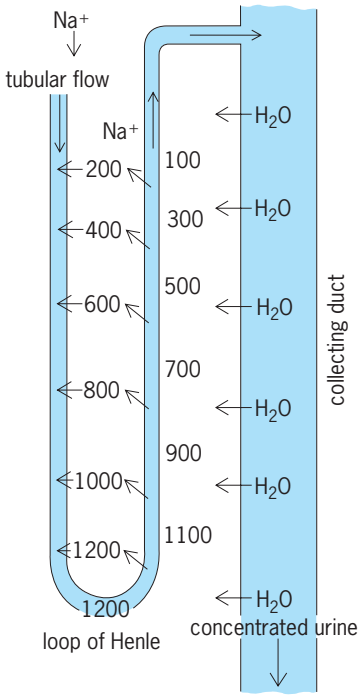

Fig. 2. Hairpin multiplier countercurrent system (loop of Henle of nephron).

The hairpin multiplier in the gas gland of the swim bladder of deep-sea marine fish serves to concentrate oxygen. Lactic acid is produced in the gas gland located at the bend of the hairpin capillaries and secreted into the blood. Lactic acid drives oxygen from hemoglobin faster than it can recombine with the hemoglobin molecule. The oxygen diffuses out of the efferent limb and into the afferent limb. With the constant addition of oxygen by arterial blood in the efferent limb, a very steep concentration gradient increase of over 1000-fold is achieved, thereby filling the swim bladder with gaseous oxygen against the great water pressures at depth of 330 ft (1000 m) or more. *See* KIDNEY; RESPIRATORY SYSTEM; SWIM BLADDER.

Walter J. Bock

Bibliography. C. L. Prosser (ed.), *Comparative Animal Physiology*, 4th ed., 1991; K. Schmidt-Nielsen, *Animal Physiology: Adaptation and Environment*, 5th ed., 1997; K. Schmidt-Nielsen et al. (eds.), *Primitive Mammals*, 1980.

Countercurrent transfer operations

Industrial processes in chemical engineering or laboratory operations in which heat or mass or both are transferred from one fluid to another, with the fluids moving continuously in very nearly steady state or constant manner and in opposite directions through the unit. Other geometrical arrangements for transfer operations are the parallel or concurrent flow, where the two fluids enter at the same end of the apparatus and flow in the same direction to the other end, and the cross-flow apparatus, where the two fluids flow at right angles to each other through the apparatus. These two arrangements are ordinarily not as efficient as countercurrent flow, but do

find certain applications in industry and the laboratory.

Heat transfer. In heat transfer there can be almost complete transfer in countercurrent operations. The limit is reached when the temperature of the colder fluid becomes equal to that of the hotter fluid at some point in the apparatus. At this condition the heat transfer is zero between the two fluids. The amount of actual heat transfer is determined by economical design, that is, by comparing the value of the transferred heat with the cost of the heat exchange equipment. The economically optimum heat transfer has been studied for many years in engineering and is changing constantly as the costs of basic forms of energy increase.

Most heat transfer equipment has a solid wall between the hot fluid and the cold fluid, so the fluids do not mix. Heat is transferred from the hot fluid through the wall into the cold fluid. Another type of equipment, fewer in number but significant in size and use, does use direct contact between the two fluids—for example, the cooling towers used to remove heat from a circulating water stream. Cooling towers are of the countercurrent type and the cross-flow type. For more discussion of heat transfer *see* COOLING TOWER; HEAT BALANCE; HEAT EXCHANGER; HEAT TRANSFER.

Mass transfer. This process involves the changing compositions of mixtures, and is done usually by physical means instead of chemical reactions because of the lower costs. Even as heat is transferred from a region of high temperature to one of lower temperature, a material is transferred within a single phase from a region of high concentration to one of lower concentration by processes of molecular diffusion and eddy diffusion. In typical mass transfer processes, at least two phases are in direct contact in some state of dispersion, and mass (of one or more substances) is transferred from one phase across the interface into the second phase. Similar to heat transfer, mass transfer takes place between two immiscible phases until equilibrium between the two phases is attained. In heat transfer, equilibrium denotes an equality of temperature in the two phases, but in mass transfer there is seldom an equality of concentration in the two equilibrium phases. This means that a component may be transferred from a phase at low concentration (but at a concentration higher than that at equilibrium) to a second phase of greater concentration. The approach to equilibrium is controlled by diffusion transport across phase boundaries. Because this is a relatively slow process, the transport rate is increased by increasing the total interfacial area, by decreasing the thickness of the near-stagnant films adjacent to the interface, and by more frequent renewals of the interfacial films.

Although the two phases may be in concurrent flow or cross-flow, usual arrangements have the phases moving in countercurrent directions. The more dense phase enters near the top of a vertical cylinder and moves downward under the influence of gravity. The less dense phase enters near the bottom of the cylinder and moves upward under the influence of a small pressure gradient. In some cases, centrifugal force is used instead of gravity to provide phase separations.

For discussions of specific mass transfer operations *see* ADSORPTION; CHEMICAL SEPARATION TECHNIQUES; CRYSTALLIZATION; DISTILLATION; DRYING; ELECTROPHORESIS; EXTRACTION; LEACHING.

Frank J. Lockhart

Bibliography. R. H. Perry and D. W. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; R. E. Treybal, *Mass Transfer Operations*, 3d ed., 1980; J. Wesslingh and R. Krishna, *Mass Transfer*, 1991.

Couple

A system of two parallel forces of equal magnitude and opposite sense (**Fig. 1**). Forces P at normal distance P constitute a counterclockwise couple C_z (viewed from $+Z$) in the OXY plane; forces Q at arm q constitute a clockwise couple C_x (from $+X$) OYZ.

Under a couple's action a rigid body tends only to rotate about a line normal to the couple's plane. This tendency reflects the vector properties of a couple. *See* RIGID-BODY DYNAMICS.

Vector properties. The total force of a couple is zero. The total moment **C** of a couple is identical about any point. Accordingly, **C** is the moment of either force about a point on the other and is perpendicular to the couple's plane. *See* RESULTANT OF FORCES; STATICS.

In Fig. 1, $\mathbf{C}_z$, at the origin for convenience, is the moment of couple C_z. Its magnitude $|C_z| = +P_p$. Its sense, by the convention of moment, is $+Z$. Also $\mathbf{C}_x$, of sense $-X$, represents couple C_x; $|C_x| = +Q_q$.

Scalar moment. The moment of a couple about a directed line is the component of its total moment in the line's direction. For example, the moment of couple C_z about line L is $M_L = |C_z| \cos \theta_L = +P_p \cos \theta_L$. Also $M_x = 0$, $M_y = 0$, and $M_z = +P_p \cos 0^\circ = +P_p$.

Equivalent couples. Couples are equivalent whose total moments are equal. In **Fig. 2**, the paired linear forces and the counterclockwise curl represent

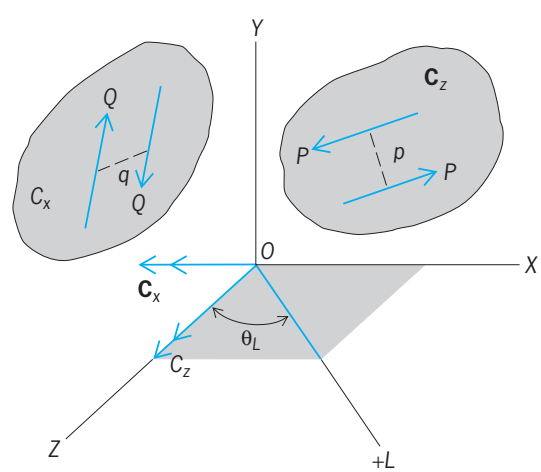

Fig. 1. Vector properties of a couple.

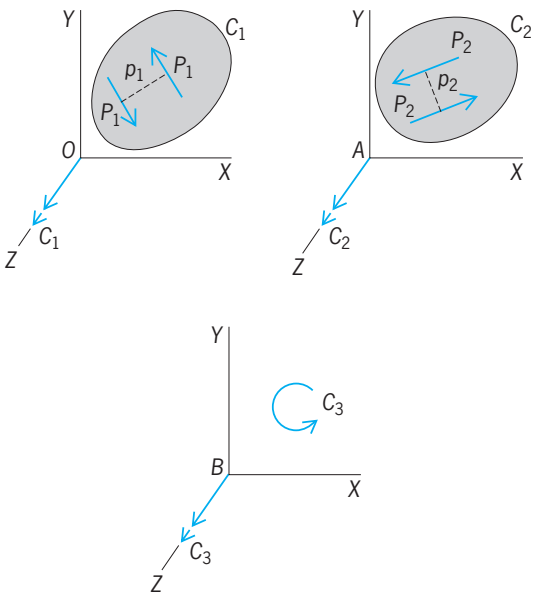

Fig. 2. Equivalent couples.

counterclockwise couple C_1, C_2, and C_3 in parallel planes OXY, AXY, and BXY. When their total moments are directed the same ($+Z$), and their magnitudes $|C_1| = P_1p_1$, $|C_2| = P_2p_2$, and $|C_3|$ are equal, then $\mathbf{C}_1 = \mathbf{C}_2 = \mathbf{C}_3$ and these couples are equivalent. Thus $\mathbf{C}_1$ can represent these or any number of other equivalent couples. Nelson S. Fisk

Coupled circuits

Two or more electric circuits are said to be coupled if energy can transfer electrically or magnetically from one to another. If electric charge, or current, or rate of change of current in one circuit produces electromotive force or affects the voltage between nodes in another circuit, the two circuits are coupled.

Between coupled circuits there is mutual inductance, resistance, or capacitance, or some combination of these. The concept of a mutual parameter is based on the loop method of analysis. A mutual parameter can be one that carries two or more loop currents; such a network has conductive coupling because electricity can flow from one circuit to the other. *See* NETWORK THEORY.

Also, there can be purely inductive coupling, which appears if the magnetic field produced by current in one circuit links the other circuit. A two-winding transformer is an application of inductive coupling, with energy transferred through the magnetic field only.

It is also possible to have mutual capacitance, with energy transferred through the electric field only. An example is the capacitive interference between two transmission lines, as a power line and a telephone line, that run for a considerable distance side by side.

Polarity. With inductive coupling, polarity may need to be known, particularly if two circuits are coupled in more than one way. Do two kinds of coupling produce voltages that add or subtract?

There are several ways to show the relative polarities of inductive coupling. **Figure 1** shows, somewhat pictorially, two coils wound on the same core. Current flowing into the upper end of coil 1 would produce magnetic flux upward in the core, and so also would current flowing into the upper end of coil 2. For this reason the upper ends of the two coils are said to be corresponding ends.

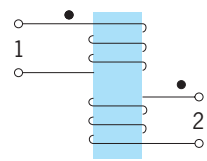

Fig. 1. Two coils wound on a single core, illustrating polarity.

If a wiring diagram is drawn with some such semipictorial sketch of the coils, it is not difficult to determine which are corresponding ends. However it is easier, both in representation and in interpretation, to indicate the corresponding ends symbolically. For this purpose dots are placed on a diagram at the corresponding ends of coupled coils. Such dots are shown in Fig. 1, though they are not needed there. Dots are also shown in **Fig. 2**, where they give the only means of identifying corresponding ends of the coils.

Note that the bottom ends of the coils shown in Fig. 1 are also corresponding ends, and that the two lower ends might have been dotted instead of the two upper ends; it makes no difference. But if the sense of winding of either coil 1 or coil 2 in Fig. 1 were reversed, then one dot (either one) would have to be shifted correspondingly.

Voltage equations. In addition to showing corresponding ends of two coils, Fig. 2 indicates that the two coils couple to each other. This coupling is identified as L_{12}. There is voltage from a to b in Fig. 2 if current i_1 is changing through the self-inductance L_{11} of the coil in circuit 1. There is additional voltage from a to b if current i_2 is changing in circuit 2 because of the mutual inductance L_{12} between the circuits. Thus, with circuit 1 coupled to circuit 2 and to any number of other circuits, Eqs. (1) and (2) hold. Here

$$v_{ab} = \left(L_{11}\frac{di_1}{dt} \pm L_{12}\frac{di_2}{dt} \pm \cdots\right) \tag{1}$$

$$v_{cd} = \left(\pm L_{21}\frac{di_1}{dt} + L_{22}\frac{di_2}{dt} \pm \cdots\right) \tag{2}$$

L_{11} and L_{22}, the self-inductances of the circuits, are positive numbers (inductances measured in henrys).

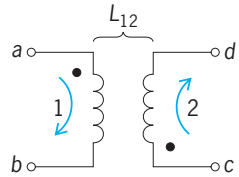

Fig. 2. Mutual inductance. Dots identify corresponding ends of coils.

With regard to the mutual terms, two questions of polarity must be asked. Are the mutual inductances such as L_{12} and L_{21} always positive numbers? And are the signs before the mutual terms, shown as $\pm$ in Eq. (1), actually $+$ or are they $-$? Answers to these questions are not always the same, and may be different as given by different authors, but the most usual simple procedure to answer them is as follows:

1. Draw a circuit diagram, such as Fig. 2, with nominal positive directions of currents shown by arrows and with dots at corresponding ends of coupled coils.
2. Let each mutual inductance such as L_{12} be a positive number.
3. Write equations such as Eqs. (1) and (2) with the following signs: First, if both arrows enter dotted ends of a pair of coupled coils, or if both arrows enter undotted ends, use $+$ before the corresponding mutual-impedance term. Second, if one arrow enters a dotted end and the other an undotted end, use $-$ before the corresponding mutual-impedance term.

However, this simple procedure sometimes fails, for it may be impossible to dot corresponding ends of all pairs of coupled coils in a network if there are three or more coils. A more general method follows:

1. Draw a circuit diagram with nominal positive directions of currents shown by arrows; place dots at coil ends arbitrarily, as may be convenient.
2. Determine corresponding ends of pairs of coupled coils, considering the coils two at a time. For the mutual inductance between a pair of coils with corresponding ends dotted, use a positive number, such as $L_{12} = 5$. For mutual inductance between a pair of coils with noncorresponding ends dotted, use a negative number, such as $L_{12} = -5$.
3. Write equations according to rule 3, above.

Steady-state equations. The differential Eqs. (1) and (2) are quite general. For steady-state operation at a single frequency, it is often simpler to have phasor or transform equations. Such equations can be in terms of reactances instead of inductances ($X = \omega L = 2\pi fL$ where f is frequency in cycles per second, or hertz). With the usual interpretation of phasor or transform equations, the steady-state relations corresponding to Eqs. (1) and (2) are shown in Eqs. (3) and (4). The rules in the foregoing paragraphs determine the choice of $+$ or $-$ in these equations; also, a mutual reactance such as X_{12} is a positive number (of ohms) if the corresponding mutual inductance such as L_{12} is a positive number (of henrys), as determined by the above rules. *See* ALTERNATING-CURRENT CIRCUIT THEORY.

$$\begin{aligned} V_{ab} &= j\omega(L_{11}I_1 \pm L_{12}I_2 \pm \cdots) \\ &= j(X_{11}I_1 \pm X_{12}I_2 \pm \cdots) \end{aligned} \tag{3}$$

$$\begin{aligned} V_{cd} &= j\omega(\pm L_{21}I_1 + L_{22}I_2 \pm \cdots) \\ &= j(\pm X_{21}I_1 + X_{22}I_2 \pm \cdots) \end{aligned} \tag{4}$$

Equality of mutual inductance. Because of this equality it is not uncommon, if there are only two coupled coils in a network, to use the letter M in place of both L_{12} and L_{21}. This use of M will be adopted in the following paragraphs.

Coefficient of coupling. The coefficient of coupling between two circuits is, by definition, Eq. (5). If the

$$k = M/\sqrt{L_{11}L_{22}} \tag{5}$$

circuits are far apart or, because of orientation, have little mutual magnetic flux, they are loosely coupled and k is a small number. Values of k for circuits with loose coupling may be in the range between 0.01 and 0.10. For closely (or tightly) coupled circuits with air-core coils, k may be around 0.5. In a transformer with a ferromagnetic core, k is very nearly 1.00.

Ideal transformers. An ideal transformer is defined as one in which primary and secondary currents are related inversely as the number of turns in the windings; this is shown in Eq. (6). Voltages across the

$$I_1/I_2 = N_2/N_1 \tag{6}$$

primary and secondary windings are in direct proportion to the numbers of turns, as shown in Eq. (7).

$$V_1/V_2 = N_1/N_2 \tag{7}$$

An actual transformer may be almost ideal or not at all ideal, depending on how it is made, and construction in turn depends on the purpose for which it is to be used. In an ideal transformer $k = 1$; in any actual transformer k is less than 1. In an actual transformer there is magnetizing current, so Eq. (6) is not exact. Also, an actual transformer has resistance and leakage reactance, so Eq. (7) is not exact. Nevertheless, the relations of ideal transformers are closely approximated by 60-Hz power transformers, and these relations are more or less close for other transformers that have ferromagnetic cores. *See* TRANSFORMER.

Whereas coupling k is desirably as close as possible to unity in a transformer that couples power from one circuit to another, k may purposefully be considerably less than unity in a transformer used for another purpose. In an oscillator, coupling need only be sufficient to sustain oscillation. In a band-pass amplifier, coupling is determined by bandwidth requirements. *See* OSCILLATOR; RESONANCE (ALTERNATING-CURRENT CIRCUITS).

Equivalent circuits. It is often convenient to substitute into a network, in place of a pair of inductively coupled circuits, an equivalent pair of conductively coupled circuits. Circuits so substituted are equivalent if the network exterior to the coupled circuits is unaffected by the change; in many cases this requirement implies that input current and voltage and output current and voltage are unaffected by the change.

Voltages and currents at the terminals of the coupled circuits of **Fig. 3** are related by Eqs. (8) and (9).

$$(R_1 + j\omega L_{11})I_1 - j\omega MI_2 = V_{ab} \tag{8}$$

$$(R_2 + j\omega L_{22})I_2 - j\omega MI_1 = V_{cd} \tag{9}$$

Voltages and currents at the terminals of the network of **Fig. 4** are also related by Eqs. (8) and (9), so the

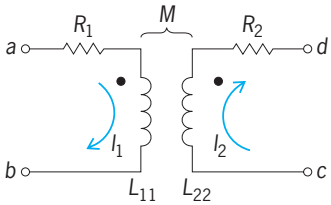

Fig. 3. A pair of coupled circuits.

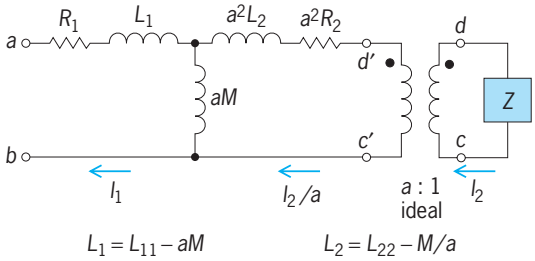

Fig. 4. A network equivalent to coupled circuits, feeding a load of impedance *Z*.

network of Fig. 4 is equivalent to the coupled circuits of Fig. 3. This is not immediately obvious, but it appears when loop equations for the network are simplified to a form that shows the requirement that $L_{11} = L_1 + aM$ and $L_{22} = L_2 + M/a$ and that $V_{c'd'} = aV_{cd}$, where a is the turn ratio of the ideal transformer shown, equal to N_1/N_2 in Eqs. (6) and (7).

It seems at first that this substitution, resulting in the network of Fig. 4, has produced something more complicated than the circuit of Fig. 3, but the value of substituting will now be shown.

Mathematically, a could be any number, but practically there are two particularly advantageous values for a. When the coupling between coils is loose so that k, the coefficient of coupling, is small, and if the coils have somewhere near the same number of turns, it is advantageous to let $a = 1$. With $a = 1$ the network of Fig. 4 is simplified to **Fig. 5**, leaving only the conductively coupled circuits with parameters R_1 and L_1, M, and L_2 and R_2, where $L_1 = L_{11} - M$, and $L_2 = L_{22} - M$. Many loosely coupled circuits particularly in radio circuits, are conveniently represented by this equivalent network with $a = 1$.

On the other hand, if coupling is close and especially when the two coupled coils have widely different numbers of turns (a situation that is typical of transformers), it is more convenient to let a equal the actual turn ratio of the coils. If, for example, there are 10 times as many turns in the primary winding of a transformer as there are in the secondary winding, it is well to let $a = 10$. (Letting $a = 1$ in such a transformer would result in a negative value for L_2 which, though correct for analysis, is difficult to visualize.)

With a equal to the actual transformer turn ratio, the following interpretation of Fig. 4 is usual and convenient. All causes of power loss and voltage drop have been put into the equivalent T network. Only the turn ratio, the actual transforming function, remains in the ideal transformer. It now becomes possible to consider, to study, and even to design the separate functions independently.

Transformers. The equivalent circuit of Fig. 4 is so commonly used in transformer work that a number of the quantities have been given special names. With the concept that power is supplied to one winding of the transformer, this is called the primary winding; the other, the secondary winding, provides power to a load. There may be a third, or tertiary, winding, perhaps providing power to another load at a different voltage or with a different connection, and even other windings, on the same transformer core.

However, it is only when a transformer has become part of a system that there is any meaning in designating the windings as primary and secondary, for a two-winding transformer can be used to carry power in either direction. The terms "high-voltage side" and "low-voltage side" are preferable for a transformer that is not part of a system. In the following discussion the words primary and secondary really mean nothing more than the windings of the transformer that are designated by the subscripts 1 and 2.

Speaking of a transformer with N_1 primary turns and N_2 secondary turns, let a be the turn ratio N_1/N_2. Then, referring to Fig. 4, R_1 is primary resistance; L_1 is primary leakage inductance; aM is primary exciting inductance; a^2L_2 is secondary leakage inductance referred to the primary side; and a^2R_2 is secondary resistance referred to the primary side. Usually, operation of power transformers is described in terms of these leakage inductances (or leakage reactances) and in terms of resistances, inductances, or reactances all referred to one side of the transformer or the other.

Transformation of impedance. **Figure 6** is the same as Fig. 4 except that the ideal transformer with turn ratio of a:1 has been eliminated and the impedance of the load has been changed from Z to a^2Z. This conversion makes the network of Fig. 6 equivalent to that of Fig. 4 at the input terminals, and it suggests a concept that is quite useful in communications.

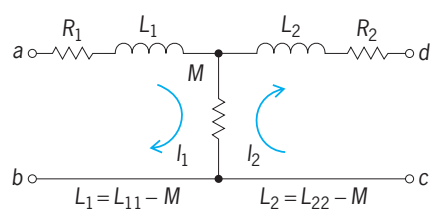

Fig. 5. A conductive network equivalent to the coupled circuits of Fig. 3 ($a = 1$).

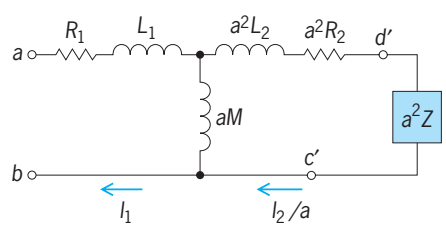

Fig. 6. Transformation of impedance from Fig. 4.

If a load with impedance Z is preceded by an ideal transformer with turn ratio a, the input impedance to the transformer is a^2Z. Thus a load of impedance Z can be made to act like a load with impedance a^2Z by using an ideal transformer with turn ratio a. A useful application is to connect a load of one impedance to an incoming line that has a different terminal impedance through an impedance-matching transformer. By this means a transformer with turn ratio a will match a load with resistance R to a source with resistance a^2R, thereby providing for maximum power transfer to the load.

The preceding paragraph assumes an ideal transformer to provide a perfect impedance match. Figure 6 shows the T network of resistances and inductances that an actual transformer introduces into the network. In practice, the transformer used must be good enough so that these resistances and inductances are acceptable.

Core loss. Neither the equations nor the circuits of this article have taken into account the power loss in a ferromagnetic core. Both eddy-current loss and hysteresis loss may be appreciable; these iron losses are commonly great enough to affect the economics of power transformers, and even to prohibit the use of metal cores with ratio-frequency current. *See* CORE LOSS.

The relations of core loss are not linear and cannot accurately be included in linear equations, but a satisfactory approximation is used with power transformers for which frequency and applied voltage change little if at all. With these restrictions, core loss can be represented by the loss in a resistance shunted around mutual inductance aM of Fig. 4. This approximation of loss is much better than neglecting core loss entirely, and is usual in work with power transformers. *See* CIRCUIT (ELECTRICITY).

Hugh H. Skilling

Bibliography. N. Balabanian, *Electric Circuits*, 1994; R. L. Boylestad, *Introductory Circuit Analysis*, 8th ed., 1996; L. P. Huelsman, *Basic Circuit Theory*, 3d ed., 1991; J. W. Nilsson and S. A. Riedel, *Electric Circuits*, 7th ed., Prentice Hall, 2004; S. Ramo, J. R. Whinnery, and T. Van Duzer, *Fields and Waves in Communications Electronics*, 3d ed., 1994.

Coupling

A mechanical fastening device for connecting the ends of two shafts together. There are three major coupling types: rigid, flexible, and fluid.

Rigid coupling. This connection is used only on shafts that are perfectly aligned. The flanged-face coupling (**Fig. 1***a*) is the simplest of these. For this type of coupling the flanges must be keyed to the shafts. The clamp, or keyless compression, coupling (Fig. 1*b*)has split cylindrical elements which clamp the shaft ends together by direct compression, through bolts, and by the wedge action of conical sections. This coupling is generally used on line shafting to transmit medium or light loads. *See* MACHINE KEY.

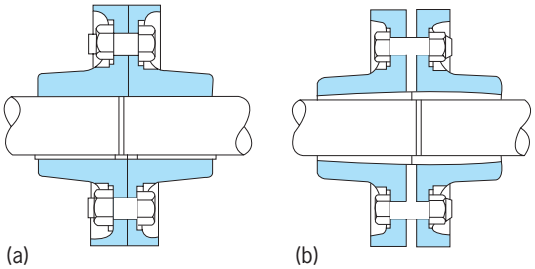

Fig. 1. Rigid coupling. (*a*) Flanged-face coupling. (*b*) Clamp coupling. (*After E. A. Avallone and T. Baumeister III, eds., Marks' Standard Handbook for Mechanical Engineers, 9th ed., McGraw-Hill, 1987*)

Flexible coupling. This connection is used for shafts which are misaligned either laterally or angularly. It also absorbs the impact due to fluctuations in shaft torque or angular speed. The Oldham, or double-slider, coupling (**Fig. 2***a*) may be used to connect shafts that have only lateral misalignment. Because

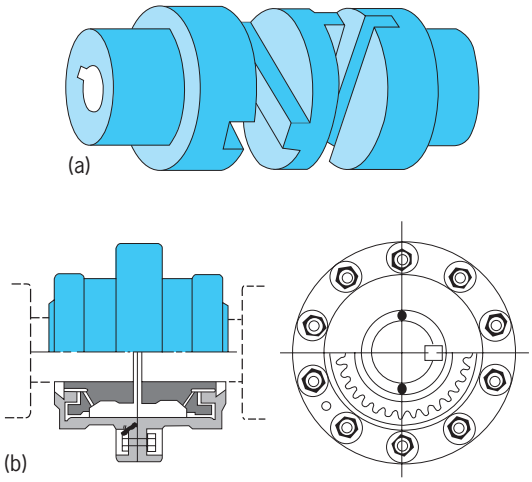

Fig. 2. Flexible coupling. (*a*) Oldham (double-slider) coupling. (*b*) Geared "fast" flexible coupling. (*After E. A. Avallone and T. Baumeister III, eds., Marks' Standard Handbook for Mechanical Engineers, 9th ed., McGraw-Hill, 1987*)

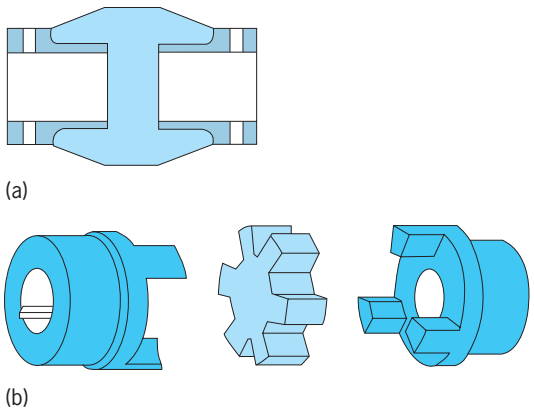

Fig. 3. Rubber flexible coupling. (*a*) Shear type. (*b*) Compression type. (*After E. A. Avallone and T. Baumeister III, eds., Marks' Standard Handbook for Mechanical Engineers, 9th ed., McGraw-Hill, 1987*)

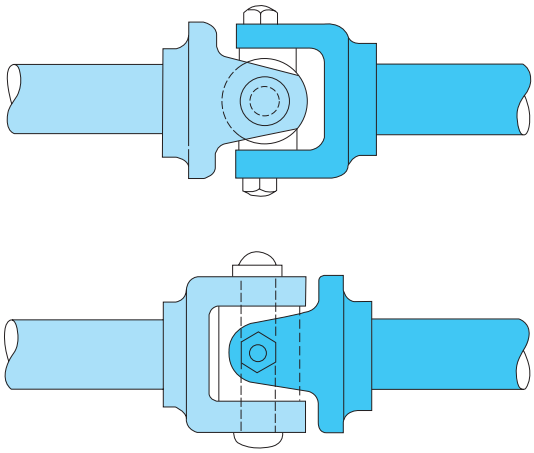

Fig. 4. Hooke's universal joint. (*After E. A. Avallone and T. Baumeister III, eds., Marks' Standard Handbook for Mechanical Engineers, 9th ed., McGraw-Hill, 1987*)

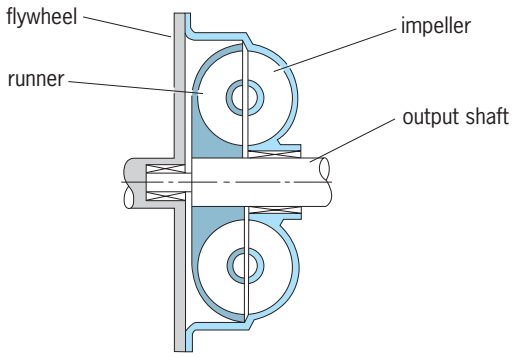

Fig. 5. Fluid coupling. (*After E. A. Avallone and T. Baumeister III, eds., Marks' Standard Handbook for Mechanical Engineers, 9th ed., McGraw-Hill, 1987*)

the tongues move about in the slots, the coupling must be well lubricated and can be used only at low speeds. The geared "fast" flexible coupling (Fig. 2*b*) uses two interior hubs on the shafts with circumferential gear teeth surrounded by a casing having internal gear teeth to mesh and connect the two hubs. Considerable misalignment can be tolerated because the teeth inherently have little interference. This completely enclosed coupling provides means for better lubrication, and is thus applicable for higher speeds. The rubber flexible coupling (**Fig. 3***a*) is used to transmit the torque through a comparatively soft rubber section acting in shear. The type shown in Fig. 3*b* loads the intermediate rubber member in compression. Both types are recommended for light loads only.

Universal joint. This is a flexible coupling for connecting shafts with much larger values of misalignment than can be tolerated by the other types of flexible coupling. Shaft angles up to 30° may be used. The initial universal joint, credited to Robert Hooke (**Fig. 4**), was a swivel arrangement by which two pins at right angles allowed complete angular freedom between two connected shafts. However, it suffers a loss in efficiency with increasing angles. *See* UNIVERSAL JOINT.

Fluid coupling. This type has two basic parts: the input member, or impeller, and the output member, or runner (**Fig. 5**). There is no mechanical connection between the two shafts, power being transmitted by kinetic energy in the operating fluid. The impeller is fastened to the flywheel and turns at engine speed. As this speed increases, fluid within the impeller moves toward the periphery because of centrifugal force. The circular shape of the impeller directs the fluid toward the runner, where its kinetic energy is absorbed as torque delivered by shaft. The positive pressure behind the fluid causes flow to continue toward the hub and back through the impeller. The toroidal space in both the impeller and runner is divided into compartments by a series of flat radial vanes. *See* FLUID COUPLING; SHAFTING. Y. S. Shin

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996.

Cover crops

Unharvested crops grown to improve soil quality or enhance pest management. Increasing numbers of farmers plant cover crops as a means of conserving soil, enhancing production, and reducing off-farm inputs. The sustainable agriculture movement has been a driving force for the increased use of cover crops. Cover cropping can accomplish a wide range of desired benefits, although there can be some drawbacks.

Benefits. Cover crops are often used as a means of preventing soil erosion (**Fig. 1**). Soil erosion is detrimental due to the loss of nutrients and crop rooting depth. Also, the washed-away sediment pollutes streams, rivers, and estuaries and increases the flooding potential downstream.

Cover crops play a vital role in controlling erosion by (1) shielding the soil surface from the impact of falling raindrops; (2) holding soil particles in place; (3) preventing crust formation; (4) improving the soil's capacity to absorb water; (5) slowing the velocity of runoff; and (6) removing subsurface water between storms through transpiration.

Fig. 1. Early spring view of a rolling corn field protected by stalk residues and domestic ryegrass covers.

Cover crops also improve soil structure and water infiltration, both of which are important for crop growth. Production of polysaccharide gums by plant roots and associated microorganisms increases aggregation of soil particles because the gums act as glues to bind particles together. Aggregation of soil particles improves soil tilth, aeration, and drainage of water. Cover crops with a higher carbon-to-nitrogen (C/N) ratio, such as grasses, usually promote greater polysaccharide production and thus greater aggregate stability than do cover crops with lower carbon-to-nitrogen ratios, such as legumes. Cover crops also improve soil structure by adding organic matter. In northern areas it is not unusual for a cover crop like annual ryegrass (*Lolium multiflorum*) to return 4 or more tons (dry weight) of organic matter to the soil. In warmer areas, 5 or more tons can be produced.

Legume cover crops add nitrogen to the soil, which can then be used by crops. They are able to utilize nitrogen from the atmosphere through a symbiotic relationship with bacteria in the genus *Rhizobium*. Through this process, called nitrogen fixation, these bacteria form nodules on legume roots (**Fig. 2**), where gaseous nitrogen is converted into nitrate or ammonium forms usable by plants. Because rhizobia occur in low levels naturally in most soils, the bacteria must be placed on legume seed before planting, either by the seed mill or the farmer; this process is called inoculation. After incorporation of the plants into the soil, the nitrogen becomes available to succeeding crops or to nearby trees or vines. Up to 150–200 lb (68–90 kg) of nitrogen can be made available by sowing legume cover crops. In many cases, legumes can fully satisfy the cash crop's demand for nitrogen.

Fig. 2. Nodules created by nitrogen-fixing rhizobia on berseem clover. Most of the nitrogen fixed in the roots is transported to the tops. (*Photo by Jack Kelly Clark*)

Insect, weed, and nematode management can be affected by cover cropping. Cover crops provide a food source to many beneficial insect species—either nectar, pollen, or other insects. In turn, these beneficial insects may feed on adjacent crop pests. For example, adult parasitic wasps feed on nectar and pollen, while their larvae feed on and kill insect pests, such as aphids. Predator insects, such as lady beetles and green lacewings, are frequently attracted to the cover crop to feed on soft-bodied insects. Many cover crops can effectively suppress weed emergence by shading the weeds or by releasing naturally occurring plant toxins in a process called allelopathy. Species commonly used to suppress weed growth include annual ryegrass, cereal rye (*Secale cereale*), and vetches (*Vicia* spp.). Nematode populations can also be reduced by the use of cover crops. For example, some varieties of cowpea (*Vigna unguiculata*) have been shown to suppress root-knot nematode (*Meloidogyne incognita*).

Potential disadvantages. Cover crops require knowledge and management to attain the desired benefit. If they are not properly selected or managed, there are drawbacks to their use, including depletion of soil moisture, competition with the adjacent crop (when present) for soil moisture and nutrients, increased frost hazard in orchards and vineyards, increased insect, nematode, and weed pests, and added costs to purchase and plant seeds.

Uses and management. Cover crops can be readily grown in humid climates and in arid and semiarid climates where irrigation provides sufficient water that cover crops do not rob the soil of needed moisture. Cover crops that reseed themselves are often sown on rangeland. Long-lived grasses such as blue grama (*Bouteloua gracilis*), buffalo grass (*Buchloë dactyloides*), and the blue-stems (*Andropogon* sp.) have been widely seeded on many different soil types in the Great Plains. In California, annual clovers (*Trifolium* spp.), annual medics (*Medicago* spp.), soft chess (*Bromus hordeaceus*), annual fescue (*Vulpia myuros*), and several perennial grass species are used on rangelands.

In cold climates, many common spring-sown crops can be advantageously used as cover crops. Spring oats (*Avena sativa*) sown in the fall are killed by cold weather but provide excellent ground cover during the fall and winter. Buckwheat (*Fagopyrum esculentum*) can be used as a cover crop to protect the soil and as a smother crop to control weeds; it also improves the soil upon incorporation. Cereal rye is widely seeded as a cover crop following the harvest of the main crop. In mild climates, summer crops such as sorghum-sudangrass (*Sorghum vulgare* × *S. sudanense*) are sometimes planted in late summer, then are killed by frost in late autumn. This practice

Fig. 3. California native perennial grass mixture planted in a drip-irrigated San Joaquin Valley vineyard. This cover crop goes dormant in the summer, reducing competition with the vines for water and nutrients. (*Photo by Jack Kelly Clark*)

protects the soil, suppresses weeds, and scavenges excess soil nitrogen to prevent it from leaching below the root zone.

Cover cropping in orchard and vineyard middles is often easier than in annual crops because they are grown between the crop spatially rather than temporally. A wide range of species, mixes, and management systems are used in orchards and vineyards. For maximum soil protection and wheel traction, perennial grasses such as tall fescue (*Festuca arundinacea*) or ryegrass (*Lolium perenne*) can be used, although these species use substantial soil moisture. In the arid west, some tree and vine crop growers sow native perennial grass species such as California brome (*Bromus carinatus*) and blue wildrye (*Elymus glaucus*) [**Fig. 3**]. These species offer excellent soil protection and go dormant in the summer, using little water. Reseeding annual clovers or grasses are also sown in no-till orchards and vineyards. Many organic growers sow high-biomass legume species such as vetches, clovers, or field peas (*Pisum sativa*), sometimes in combination with cereal cover crops. Upon incorporation, these "green manure" cover crops add nitrogen to the soil for use by the trees or vines. *See* AGRICULTURAL SOIL AND CROP PRACTICES; NITROGEN FIXATION; SOIL CONSERVATION.

Chuck Ingels; Paul J. Zwerman

Bibliography. C. A. Ingels et al. (eds.), *Cover Cropping in Vineyards: A Grower's Handbook*, University of California, Division of Agriculture and Natural Resources, Oakland, 1998; *Managing Cover Crops Profitably*, 2d ed., Sustainable Agriculture Network, Beltsville, MD, 1998; M. Sarrantonio, *Northeast Cover Crop Handbook*, Rodale Institute, Emmaus, PA, 1994.

Cowpea

The legume *Vigna unguiculata* ssp. *unguiculata*, also called southern pea, blackeye pea, or blackeye bean (United States), or niébé (French-speaking Africa). It is an important source of dietary protein for human consumption and of animal feed in the tropics, especially in Africa, Brazil, and India where cowpeas are grown mostly as a subsistence crop for home consumption and are not sold in markets. The cowpea is adapted to hotter, more arid climates and more infertile soils than other food legume crops. Its symbiotic nitrogen-fixing abilities help maintain soil fertility in peasant cropping systems. Over 65% of the cowpea crop [2.5 million metric tons (5.5 billion pounds) from 8 million hectares (19.8 million acres)] is produced in Africa, Nigeria and Niger producing 50% of the world supply. Brazil produces over 25%. The United States is the only developed country producing large amounts of cowpea. Yields in developing countries are only about 551–882 lb/ha (250–400 kg/ha), principally because of insect infestations but also because of diseases and low soil fertility. In the United States, average yields reach 4410 lb/ha (2000 kg/ha) because of heavier use of inputs, and are comparable to yields of other food legumes.

Origin. The cowpea was domesticated in Africa from one of several wild taxa that belong to the same species as the cultivars and are classified as *V. unguiculata* ssp. *dekindtiana*. The wild relatives are widely distributed in the Sahel, eastern and southern Africa. The precise location of domestication is not known, although northern Nigeria and Benin, Ethiopia, and Botswana have been suggested as possible areas. At least by 2000 years ago, the cowpea was introduced into India, probably with African cereals such as pearl millet and sorghum. In Asia, new cultivar types were selected in addition to the original unguiculata type, grown mostly for dry seeds. The sesquipedalis type, also known as the yard-long or asparagus bean, is grown in India, China, and southeastern Asia for its succulent young pods and leaves. Its pods are up to 18–30 in. (45–75 cm) long. The biflora or catjang type is grown for fodder and dry seeds mostly in India and Sri Lanka. An additional, minor cultivar type is the textilis type, which was grown in northern Nigeria but is now rarely found in farm fields. It is a primitive type, grown for its fibers in the inflorescence peduncles. Cowpea was introduced into the Americas (principally Brazil and the United States) through the colonial and slave trades. In the United States, more than 65% of the production originates in California and Texas (dry seeds) and Georgia (fresh peas).

Characteristics. Cultivars generally have dark-green, glabrous, shiny leaves. Growth habits include climbing, prostrate, and bush types. Flowers and pods are distributed mostly above the leaf canopy, in the upper part of the canopy, or throughout the canopy. Flowers are white or violet with various color patterns. Most cultivars are self-pollinating; however, some wild cowpeas in Africa are cross-pollinated, producing larger, odoriferous flowers. Pods can be erect or pendant (hanging from the peduncle). Their shape ranges from straight to curved to coiled (similar to alfalfa pods). Seeds can be kidney or egg shaped, spherical or rhomboid. The seed coat ranges from smooth to wrinkled. Pigmentation covers the entire seed (self-colored) or surrounds the hilum (eye). Colors are mottled, speckled, or solid

and include white, cream, brown, red, pink, green, and black.

Cultivation. In Africa, the savannas of South America, and southern and southeastern Asia, cowpeas are intercropped with cereals (sorghum, pearl millet, maize) or cotton, or in some cases they are grown as a monoculture. In very dry areas, such as northeastern Brazil, they are grown with spineless forage cactus (*Opuntia* spp.). In the moist Amazon areas of Peru and Brazil, climbing cowpeas are grown. In rubber and tea plantations in Asia, cowpeas are grown as a cover crop or a green manure. In the United States, they are grown in monocultures. Breeders are developing cultivars with improved growth habits; greater resistance to pests, diseases, and parasitic plants (*Striga* and *Alectra*); better tolerance to heat and drought stress; and enhanced symbiotic nitrogen fixation. Novel, low-cost methods are being developed to control cowpea seed weevils in developing countries that do not have adequate pesticide use and crop management programs. *See* LEGUME FORAGES; MULTIPLE CROPPING.

Utilization and nutritional qualities. The most important utilization of cowpeas is that of seeds, whether mature (dry seeds or blackeye bean) or immature (southern pea). Seeds are marketed as a dry pack or as a canned or frozen product. Other uses, such as for the yard-long beans, are becoming more available, especially in specialty food markets. Cowpea seeds contain 23–30% protein. As with other food legumes, these proteins are deficient in the essential amino acids methionine and cystine, but they complement cereal proteins, which are deficient in lysine. *See* LEGUME. P. Gepts

Diseases. Diseases caused by fungi and viruses are primary constraints in cowpea production. Roots can be destroyed by the fungi *Rhizoctonia solani*, *Sclerotium rolfsii*, and *Pythium* sp. *Cercospora cruenta* and *C. canescens* cause leaf spots, which sometimes result in defoliation of mature cowpea plants. Stem anthracnose and pod spot are caused by *Colletotrichum* sp. *Fusarium oxysporum* infects vascular tissue and causes wilting.

Virus diseases of cowpea are characterized by leaf discolorations (mottle, mosaic, chlorosis), stunted growth, and reduced yields. Malformed leaves and distorted growth are caused by cowpea mosaic and southern bean mosaic viruses. Although cucumber mosaic virus causes a very mild disease in cowpea, it frequently occurs in a mixed infection with blackeye cowpea mosaic virus and results in a synergistic interaction with severely stunted plants (see **illus.**). *See* PLANT VIRUSES AND VIROIDS.

The bacterium *Xanthomonas vignicola* can cause both leaf blight and stem canker. In sandy soils a root-knot disease caused by nematodes (*Meloidogyne* sp.) can be particularly injurious to general plant growth.

The most common method to control cowpea diseases is biological. Resistant cultivars are known for many of the diseases, and chemical control usually is not applied. Breeding programs have attempted to incorporate resistance to the major diseases into single cowpea lines. *See* BREEDING (PLANT); PLANT PATHOLOGY. Cedric W. Kuhn

Severe mosaic and stunting of cowpea caused by a mixed infection of cucumber mosaic virus and blackeye cowpea mosaic virus. (*From G. Pio-Ribeiro, S. D. Wyatt, and C. W. Kuhn, Cowpea stunt: A disease caused by synergistic interaction of two viruses, Phytopathology, 68:1260–1265, 1978*)

Bibliography. J. C. Forbes and D. Watson, *Plants in Agriculture*, 1992; S. R. Singh and K. O. Rachie, *Cowpea Research, Production and Utilization*, 1985; J. J. Vorst, *Crop Production*, 1992.

Coxsackievirus

A large subgroup of the genus *Enterovirus* in the family Picornaviridae. The coxsackieviruses produce various human illnesses, including aseptic meningitis, herpangina, pleurodynia, and encephalomyocarditis of newborn infants. *See* ENTEROVIRUS; PICORNAVIRIDAE.

Coxsackieviruses measure about 28 nanometers in diameter; they resemble other enteroviruses in many biological properties, but differ in their high pathogenicity for newborn mice. Inapparent infections can be induced in adult mice, chimpanzees, and cynomolgus monkeys. Some types, especially A7, produce severe central nervous system lesions in monkeys. Coxsackieviruses of group A and group B differ in their effects on suckling mice; those of group A produce widespread myositis in skeletal muscle, whereas group B viruses produce necrosis of embryonic fat pads (brown fat). At least 23 antigenically distinct types in group A are now recognized, and 6 in group B. Several types (B1–6, A9) grow well in tissue culture. *See* TISSUE CULTURE.

After incubation for 2–9 days, during which the virus multiplies in the enteric tract, clinical manifestations appear which vary widely.

Diagnosis is by isolation of virus in tissue culture or infant mice. Stools are the richest source of virus. Neutralizing and complement-fixing antibodies form during convalescence and are also useful in diagnosis. *See* ANTIBODY; COMPLEMENT-FIXATION TEST; NEUTRALIZING ANTIBODY.

The coxsackieviruses have worldwide distribution. Infections occur chiefly during summer and early fall, often in epidemic proportions. Spread of virus, like that of other enteroviruses, is associated with family contact and contacts among young children. *See* ANIMAL VIRUS; VIRUS CLASSIFICATION.

Joseph L. Melnick; M. E. Reichmann

Bibliography. D. O. White and F. Fenner, *Medical Virology*, 4th ed., 1994.

CPT theorem

A fundamental ingredient in quantum field theories, which dictates that all interactions in nature, all the force laws, are unchanged (invariant) on being subjected to the combined operations of particle-antiparticle interchange (so-called charge conjugation, *C*), reflection of the coordinate system through the origin (parity, *P*), and reversal of time, *T*. The operations may be performed in any order; *TCP*, *TPC*, and so forth, are entirely equivalent. If an interaction is not invariant under any one of the operations, its effect must be compensated by the other two, either singly or combined, in order to satisfy the requirements of the theorem. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM FIELD THEORY.

The *CPT* theorem appears implicitly in work by J. Schwinger in 1951 to prove the connection between spin and statistics. Subsequently, G. Lüders and W. Pauli derived more explicit proofs, and it is sometimes known as the Lüders-Pauli theorem. The proof is based on little more than the validity of special relativity and local interactions of the fields. The theorem is intrinsic in the structure of all the successful field theories. *See* QUANTUM STATISTICS; RELATIVITY; SPIN (QUANTUM MECHANICS).

Significance. *CPT* assumed paramount importance in 1957, with the discovery that the weak interactions were not invariant under the parity operation. Almost immediately afterward, it was found that the failure of *P* was attended by a compensating failure of *C* invariance. Initially, it appeared that *CP* invariance was preserved and, with the application of the *CPT* theorem, invariance under time reversal. Then, in 1964 an unmistakable violation of *CP* was discovered in the system of neutral *K* mesons. *See* PARITY (QUANTUM MECHANICS).

One question immediately posed by the failure of parity and charge conjugation invariance is why, as one example, the π^+ and π^- mesons, which decay through the weak interactions, have the same lifetime and the same mass. It turns out that the equality of particle-antiparticle masses and lifetimes is a consequence of *CPT* invariance and not *C* invariance alone. A casual proof can be obtained by applying the *CPT* operation to a free particle. This results in an antiparticle with precisely the same momentum and total energy and therefore, from special relativity, the same rest mass as the original particle. This obvious result can be generalized heuristically to include equality of lifetimes by including an imaginary component in the mass.

Experimental tests. The validity of symmetry principles such as the *CPT* theorem should not be assumed beyond the limits of the experimental tests. In that the *CPT* theorem predicts the equality of particle-antiparticle properties such as mass, lifetime, and the magnitude of the magnetic moment, measurements which compare these quantities constitute tests of the theorem. For example, it is known from studies of the neutral *K*-meson system that the mass of the K^0 is the same as the $\overline{K^0}$ to within 1 part in 10^{18}. The magnitude of the magnetic moment of the electron is known to be the same as the positron to 1 part in 10^{10}. The lifetime of the π^+ and π^- mesons is the same to better than 0.1%, and so with the μ^+ and μ^- mesons.

The discovery of *CP* violations in 1964 in the system of neutral *K* mesons reopened the question of *CPT* invariance. That system has been analyzed in great detail, and it is known now that at least 90% of the observed *CP* violation is compensated by a violation of time-reversal invariance. That is, not more than 10% of the observed *CP* violation can be credited to a violation of *CPT*. These are all experimental limits and do not in any way suggest a violation of *CPT*.

The remarkable sensitivity of the *K* mesons to possible departures from *CPT* invariance, as demonstrated by the precision with which the masses of the K^0 and $\overline{K^0}$ are known to be the same, suggests that this system is perhaps the best one for further experimental tests of the validity of the *CPT* theorem. *See* ELEMENTARY PARTICLE; MESON; SYMMETRY LAWS (PHYSICS).

Val L. Fitch

Crab

The name applied to arthropods in sections Anomura and Brachyura of the reptantian suborder of the order Decapoda, class Crustacea. The term crab is also sometimes used for two species of sucking lice (order Anoplura) which prey upon humans. *Phthirus pubis* is the crab louse which inhabits the pubic region.

Section Anomura includes over 1400 species in 12 different families. Commonly called hermit, king, sand, or mole crabs, the anomurans have lobsterlike abdomens which bend beneath the cephalothorax in crablike manner, setting them apart in a somewhat ill-defined group (see **illus.**).

Section Brachyura encompasses the so-called true crabs, with reduced abdomens folded snugly beneath the cephalothorax. More than 4500 species making up 28 families have been grouped into subsections. In addition to subsections Dromiacea, Gymnopleura, and Oxystomata, the Hapalocarcinidea has one family of the same name containing three

Representative species of an anomuran, the hermit crab.

genera of small crabs with elongated abdomens. Young coral gall crab females settle on coral and become enveloped by growing coral. Small openings in the coral, maintained by the crab's water currents, are large enough to permit passage of miniature males (less than 0.08 in. or 2 mm long) and larvae; the female is stuck there for life.

Subsection Brachygnatha is the largest group of brachyurans and includes the most typical crabs. It has 20 families encompassing over 3700 species which occupy a variety of habitats, undergo different types of development, and exhibit contrasting patterns of behavior. Family Cancridae includes popular edible species of the genus *Cancer*. The Portunidae are the swimming crabs, most of which have paddle-shaped modifications of the terminal segments of their fifth pereopods. This structure, advantageous in swimming, is disadvantageous to the blue crab, *Callinectes sapidus*, in that the outline of the new exoskeleton in the thin paddle allows identification of premolt crabs. Such crabs, called "peelers," are kept in cypress wood floats until they molt, then sold as soft-shelled delicacies; the meat of hard-shelled crabs is also delectable. Blue crab fishing and marketing is one of the largest of the shellfish industries on the Atlantic coast of the United States.

Families Potomidae, Pseudothelphusidae, and Trichodactylidae include freshwater species. These crabs are unique in that, like crayfish, the hatching young resemble adults; there is no series of larval stages culminating in metamorphosis. The only other freshwater crabs are four species in the family Grapsidae. These crabs hatch as zoeae which complete their nonfeeding larval development rapidly, using yolk retained from the eggs for nutrition. One Jamaican species, *Metapaulias depressus*, lives in water contained in bromeliad plants. Other grapsid crabs are marine or semiterrestrial forms which invade fresh waters but must undergo postembryonic development in the sea.

The Gecarcinidae, or land crabs, live in humid, tropical regions. *Cardisoma* and *Gecarcinus* may migrate several kilometers from the sea to dig burrows in moist areas. Although the land crabs retain gills, portions of the moist inner surface of the branchial chamber have become highly vascularized to allow respiration in air. A female returns to the sea only to release larvae as they hatch from the fertilized egg mass she carries.

Most Pinnotheridae are tiny, therefore called pea crabs, and live commensally with other organisms. Different species live in the mantle cavities of oysters or other mollusks, in polychaete tubes or burrows, in branchial cavities of ascidians, in cloacas of sea cucumbers, or in the burrows of other crustaceans. After establishing a commensal habitat, many undergo a molt, after which the exoskeleton remains soft.

Other crabs commonly seen at the coast belong to the family Xanthidae (the mud crabs) or the family Ocypodidae. These are amphibious burrowers. *Uca* spp. (Ocypodidae) are called fiddler crabs because males have one enlarged cheliped, the weight of which may be equal that of the rest of the animal. In different species, this claw is used to beat the ground, creating vibrations to threaten, or to wave, an activity which may be a part of courtship. *Ocypode*, a ghost crab, does not have an enlarged cheliped but makes a variety of sounds by stridulation, rubbing the ridged inner surface of the claw against a tubercle on the base of the same cheliped.

Many of the other brachygnathans are marine and inhabit deeper waters. The best known of these are the Majidae, decorator or spider crabs. The Japanese spider crab, *Macrocheira kaempferi*, can have a span up to 11.9 ft (3.6 m) [body width of 18 in. or 45 cm] between the tips of its chelipeds; it is the largest living crab. *See* ARTHROPODA; CRUSTACEA; DECAPODA (CRUSTACEA); MEROSTOMATA.

Marilyn S. Kerr

Crab Nebula

The Crab Nebula is the remnant of a tremendous stellar explosion witnessed by Chinese astronomers in 1054. The explosion, called a supernova, occurred at a distance of about 2000 parsecs from the Earth (1 parsec = 1.9×10^{13} mi = 3.1×10^{13} km = 3.26 light-years). Even so, it was bright enough to be visible in the daytime sky for 3 weeks and in the nighttime sky for 2 years. *See* SUPERNOVA.

Neutron star. The Crab now consists of three components. At the heart of the nebula is what is left of the core of the Crab's giant stellar progenitor. This neutron star has twice the mass of the Sun concentrated into an object only about 20 km (12 mi) across, giving it a density of 10^9 tons per cubic centimeter. The neutron star is spinning at 30 times a second, whipping its powerful (10^8 tesla) magnetic field around with it. Radiation formed in this extreme environment is concentrated into two intense beams directed away from the neutron star's two magnetic poles. As these beams sweep past the direction of the Earth like the beam from a lighthouse, the star appears to wink on and off, earning it the name "pulsar." The Crab pulsar has been seen in all parts of the electromagnetic spectrum from gamma rays through radio waves. *See* CERENKOV

RADIATION; GAMMA-RAY ASTRONOMY; NEUTRON STAR; PULSAR; RADIO ASTRONOMY; X-RAY ASTRONOMY.

Synchrotron nebula. The Crab pulsar's rotational period is slowing by 34 nanoseconds a day, and as it slows it loses 5×10^{31} joules per second, or 100,000 times more power than is radiated away by the Sun. Most of this energy is carried away from the pulsar by a wind of electrons and positrons moving at close to the speed of light. The wind feeds a vast cloud of highly relativistic particles about 3 parsecs × 2 parsecs in size. This cloud is called the Crab synchrotron nebula because, as the particles spiral through the nebula's $3–5 \times 10^{-8}$ tesla magnetic field, they give off the sort of radiation emitted by a synchrotron particle accelerator. The synchrotron nebula is bright in x-rays and ultraviolet radiation as well as the lower-energy visible and radio parts of the spectrum. *See* SYNCHROTRON RADIATION.

Gaseous filaments. The third component of the Crab is a complex of filaments made up of gas ejected by the explosion itself. The filaments, which together have a mass of about twice that of the Sun, are moving away from the site of the explosion at up to around 1500 km/s (900 mi/s). This gas traps the synchrotron "bubble." The pressure of the synchrotron bubble in turn pushes on and compresses the gas into filaments, much like the droplets in an inverted bottle of oil and vinegar salad dressing. Together the mass of the neutron star and the mass of the filaments add up to about four times that of the Sun, yet the understanding of supernova explosions is that the star which exploded must have had at least twice this much mass. The remaining mass from the progenitor may be distributed in a very diffuse cloud of expanding ejecta surrounding the bright nebula (see **illus.**).

The filaments are ionized by ultraviolet radiation from the synchrotron nebula, causing them to glow like the gas in a fluorescent light bulb. Virtually all of the atoms in the universe that are more massive than helium (including the atoms of which humans are made) were formed by nuclear reactions in the interiors of stars and then ejected back into space. Analysis of the light from the Crab filaments shows that they are rich in these more massive elements, the chemical legacy of the fusion furnace that powered the Crab's progenitor until the moment of its cataclysmic death. *See* ASTROPHYSICS, HIGH-ENERGY; FLUORESCENCE; NUCLEOSYNTHESIS. Jeff Hester

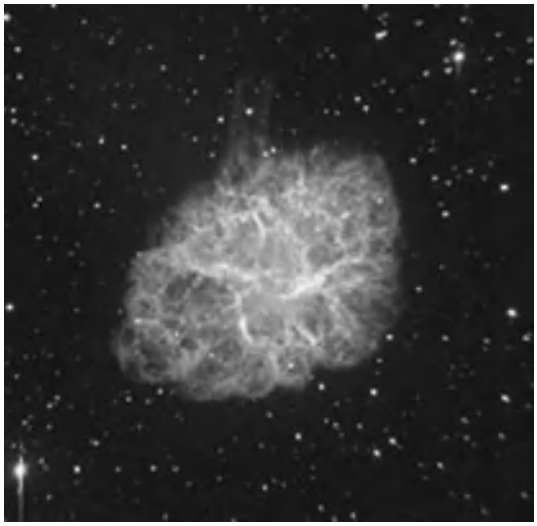

Crab Nebula, the 900-year-old remnant of a cataclysmic stellar explosion.

Crabapple

A fruit, represented commercially by such varieties as Martha, Hyslop, and Transcendent, comprising hybrids between *Malus baccata* (Siberian crabapple) and *M. domestica* (cultivated apple).

Trees may grow to 40 ft (12 m) and are very hardy. At one time the fruit of the crabapple was esteemed because of its high pectin content and its usefulness in jam and jelly manufacture. With the introduction of commercial pectin preparations, demand for the crabapple declined sharply. Except for its use as a pickled product, there is little commercial interest in this fruit. *See* PECTIN.

Twenty or more species of Oriental flowering crabapples, some of which produce edible fruit of indifferent quality, are listed. *See* APPLE; FRUIT; FRUIT, TREE; ROSALES. Harold B. Tukey

Cracking

A process used in the petroleum industry to reduce the molecular weight of hydrocarbons by breaking molecular bonds. Cracking is carried out by thermal, catalytic, or hydrocracking methods. Increasing demand for gasoline and other middle distillates relative to demand for heavier fractions makes cracking processes important in balancing the supply of petroleum products.

Thermal cracking depends on a free-radical mechanism to cause scission of hydrocarbon carbon-carbon bonds and a reduction in molecular size, with the formation of olefins, paraffins, and some aromatics. Side reactions such as radical saturation and polymerization are controlled by regulating reaction conditions. In catalytic cracking, carbonium ions are formed on a catalyst surface, where bond scissions, isomerizations, hydrogen exchange, and so on, yield lower olefins, isoparaffins, isoolefins, and aromatics. Hydrocracking, a relative newcomer to the industry, is based on catalytic formation of hydrogen radicals to break carbon-carbon bonds and saturate olefinic bonds. Hydrocracking converts intermediate- and high-boiling distillates to middle distillates, high in paraffins and low in cyclics and olefins. Hydrocracking also causes hydrodealkylation of alkyl-aryl components in heavy reformate to produce benzene and naphthalene. *See* HYDROCRACKING.

Thermal cracking. This is a process in which carbon-to-carbon bonds are severed by the action of heat alone. It consists essentially in the heating of any

fraction of petroleum to a temperature at which substantial thermal decomposition takes place through a thermal free-radical mechanism followed by cooling, condensation, and physical separation of the reaction products. *See* FREE RADICAL.

There are a number of refinery processes based primarily upon the thermal cracking reaction. They differ primarily in the intensity of the thermal conditions and the feedstock handled.

Visbreaking is a mild thermal cracking operation (850–950°F or 454–510°C) where only 20–25% of the residuum feed is converted to mid-distillate and lighter material. It is practiced to reduce the volume of heavy residuum which must be blended with low-grade fuel oils.

Thermal gas-oil or naphtha cracking is a more severe thermal operation (950–1100°F or 510–593°C) where 45% or more of the feed is converted to lower molecular weight. Attempts to crack residua under these conditions would coke the furnace tubes.

Steam cracking is an extremely severe thermal cracking operation (1100–1400°F or 593–760°C) in which steam is used as a diluent to achieve a very low hydrocarbon partial pressure. Primary products desired are olefins such as ethylene, propylene, and butadiene.

Fluid coking is a thermal operation where the residuum is converted fully to gas-oil products boiling lower than 950°F (510°C) and coke. The thermal conversion is carried out on the surface of a fluidized bed of coke particles.

The fluid coking process has been extended to include gasification of the product coke in a vessel that is heat-integrated into the plant. Low-Btu gas which can be desulfurized is produced rather than a product coke stream.

Delayed coking is a thermal cracking operation wherein a residuum is heated and sent to a coke drum, where the liquid has an infinite residence time to convert to lower-molecular-weight hydrocarbons which distill out of the coke drum, and to coke which remains in the drum and must be periodically removed.

In fluid coking and delayed coking, there is total conversion of the very heavy high-boiling end of the residuum feed.

Although there are many variations of visbreaking and thermal cracking, most commonly a feedstock that boils at higher temperatures than gasoline is pumped at inlet pressures of 75–1000 pounds per square inch gage (0.52–6.9 megapascals) through steel tubes so placed in a furnace as to allow gradual heating of the coil to temperatures in the range 850–1100°F (454–593°C). The flow rate is controlled to provide sufficient time for the required cracking to lighter products; the time may be extended by subsequently passing the hot products through a reaction chamber that is maintained at a high temperature. To achieve optimum process efficiency, part of the overhead product ordinarily is returned to the cracking unit for further cracking (**Fig. 1**).

Crude oils differ in their compositions, both in molecular weight and molecular type of hydrocarbon. Since refiners must make products in harmony with market demand, they often need to alter the molecular structure of the hydrocarbons. The cracking of heavy distillates and residual oils increases the yields of gasoline and the light intermediate distillates used primarily as domestic heating oils, as well as providing low-molecular-weight olefins needed for the manufacture of chemicals and polymers.

Beginning in 1912, thermal cracking proved for many years to be eminently suitable for this purpose. During the period 1920–1940, more efficient automobile engines of higher compression ratios were developed. These engines required higher-octane-number gasolines, and thermal cracking operations in the United States were expanded to meet this need. Advantageously, thermal cracking reactions produce olefins and aromatics, leading to gasolines generally of higher octane number than those obtained by simple distillation of the same crude oils. The general nature of the hydrocarbon products and the basic mechanism of thermal cracking is well described by the free-radical theory of the pyrolysis of hydrocarbons. *See* PYROLYSIS.

In the early 1930s, the petrochemical industry began its growth. Olefinic gases from thermal cracking operations, especially propylene and butylenes, were used as the chief raw materials for the production of aliphatic chemicals. Stimultaneously, practical catalytic processes were invented for polymerizing propylene and butylenes to gasoline components, and for dimerizing and hydrogenating isobutylene to isooctane (2,2,4-trimethylpentane), the prototype 100-octane fuel. Just prior to World War II, the alkylation of light olefins with isoparaffins to produce unusually high-octane gasoline components was discovered and extensively applied for military aviation use. These developments resulted in intense engineering efforts to bring thermal cracking to maximum efficiency, as exemplified by a number of commercial processes made available to the industry.

Since World War II, thermal cracking has been largely supplanted by catalytic cracking, both for the manufacture of high-octane gasoline and as a source of light olefins. It is, however, still used widely for the mild cracking of heavy residues to reduce their viscosities and for the final cracking of low-grade residuum.

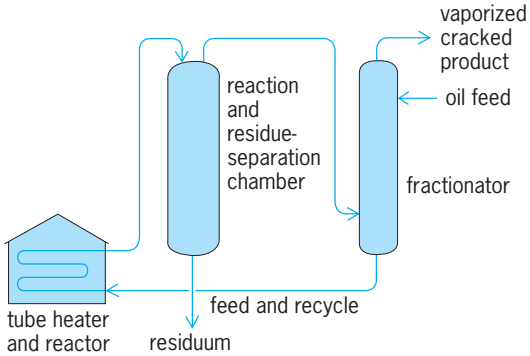

Fig. 1. Thermal cracking unit.

Since carbon-to-hydrogen bonds are also severed in the course of thermal cracking, two hydrogen atoms can be removed from adjacent carbon atoms in a saturated hydrocarbon, producing molecular hydrogen and an olefin. This reaction prevails in ethane and propane cracking to yield ethylene and propylene, respectively. Methane cracking is a unique case wherein molecular hydrogen is obtained as a primary product and carbon as a coproduct. These processes generally operate at low pressures and high temperatures and in some cases utilize regenerative heating chambers lined with firebrick, or equipment through which preheated refractory pebbles continuously flow. Such conditions also favor the production of aromatics and diolefins from normally liquid feedstocks and are applied commercially to a limited extent despite relatively low yields of the desired products.

Catalytic cracking. This is the major process used throughout most of the world oil industry for the production of high-octane quality gasoline by the conversion of intermediate- and high-boiling petroleum distillates to lower-molecular-weight products. Many aspects involve proprietary technology. In general, oil heated to within the lower range of thermal cracking temperatures (850–1025°F or 454–551°C) reacts in the presence of an acidic inorganic catalyst, typically a silica-alumina zeolite, under low pressures (10–35 psig or 70–240 kilopascals). Gasoline of much higher octane number is obtained than from thermal cracking, a principal reason for the widespread adoption of catalytic cracking. All nonvolatile carbonaceous materials are deposited on the catalyst as coke and are burned off during catalyst regeneration. *See* ZEOLITE.

In contrast to thermal cracking, low-grade residual oils are not generally processed, because excessive amounts of coke are deposited on the catalyst, and inorganic components of these oils contaminate the catalyst. Preferential feedstocks are heavier-boiling distillate gas oils. However, in recent years with catalyst improvement, residual stocks can also be processed. These stocks are generally good-quality residuals in low concentration as they yield 2–3 barrels of gas oil feed when processed in a full unit.

Many processes have become available to improve catalytic cracking feedstocks. These include hydrogenation, deasphalting (including the Rose process), deep-vacuum distillation, and separations or low-conversion processes such as the Art process. These processes generally reject the poorer components of the feed, such as Conradson carbon (a residual carbon), metallic contaminants, and possibly heavy aromatics. Hydrogenation can upgrade feedstocks by saturation or aromatics, and remove metals to considerably improve catalytic cracker product selectivity.

Catalytic cracking, as conceived by E. J. Houdry in France, reached commercial status in 1936 after extensive engineering development by American oil companies. In its first form, the process used a series of fixed beds of catalyst in large steel cases. Each of these alternated between oil cracking and catalyst-coke burning at intervals of about 10 min and provided for heat and temperature control.

Successful operation led to major engineering improvements, and the goals of much improved efficiency, enlarged capacity, and ease of operation were achieved by two different systems. One employs a moving bed of small pellets or beads of catalyst traveling continuously through the oil-cracking vessel and subsequently through a regeneration kiln. The beads are lifted mechanically or by air to the top of the structure to flow down through the vessels again. This process has two commercially engineered embodiments, the Thermofor (or TCC) and the Houdriflow processes, which are similar in general process arrangements.

These moving-bed processes are limited in size and became technically obsolete. They generally have been replaced by another type of unit, the fluid-solids, as dictated by economic considerations, and no new moving-bed units are being constructed. In the fluid-solids unit, a finely divided powdered catalyst is transported between oil-cracking and air-regeneration vessels in a fluidized state by gaseous streams in a continuous cycle. This system employs the principle of balanced hydrostatic heads of fluidized catalyst between the two vessels. Catalyst is moved by injecting heated oil vapors into the transport line from the regenerator to the reactor, and by injecting air into the transport line from the reactor stripper to the regenerator. Large amounts of catalyst can be moved rapidly; cracking units of total oil intake as great as 180,000 bbl/day (28,800 m^3/day) are in operation (**Fig. 2**).

In both the moving-bed and fluidized systems, the circulating catalyst provides the cracking heat. Coke deposited on the catalyst during cracking is burned at controlled air rates during regeneration; heat of combustion is converted largely to sensible heat of the catalyst, which supplies the endothermic heat of cracking in the reaction vessel.

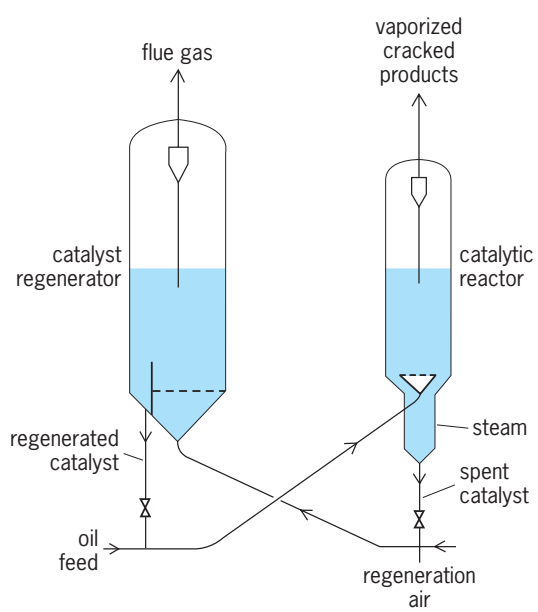

Fig. 2. Fluid catalytic cracking unit.

Gasoline of 90–95 research octane number without tetraethyllead is rather uniformly produced by catalytic cracking of fractions from a wide variety of crude oils, compared with 65–80 research octane number via thermal cracking, the latter figures varying with crude oil source. *See* OCTANE NUMBER.

Although the primary objective of catalytic crackling is the production of maximum yields of gasoline concordant with efficient operation of the process, large amounts of normally gaseous hydrocarbons are produced at the same time. The gaseous hydrocarbons include propylene and butylenes, which are in great demand for chemical manufacture. Isobutane and isopentane are also produced in large quantities and are valuable for the alkylation of olefins, as well as for directly blending into gasoline as high-octane components.

The other chief product is the material boiling above gasoline, designated as catalytically cracked gas oil. It contains hydrocarbons relatively resistant to further cracking, particularly polycyclic aromatics. The ligher portion may be used directly or blended with straight-run and thermally cracked distillates of the same boiling range for use as a diesel oil component and heating oils. Part of the heavier portion can be recycled with fresh feedstock to obtain additional conversion to lighter products. The remainder is withdrawn for blending with residual oils to reduce the viscosity of heavy fuel, or else subjected to a final step of thermal cracking.

Thus, the catalytic cracking process is used in refineries to shift the production of products to match swings in market demand. It can process a wide variety of feeds to different product compositions. For example, light gases, gasoline, or diesel oil can be emphasized by varying process conditions, feedstocks, and boiling range of products (see **table**).

To account for the difference between the product compositions obtained by catalytic and thermal cracking, the mechanism of cracking over acidic catalysts has been investigated intensively. In thermal cracking, free radicals are reaction intermediates, and the products are determined by their specific decomposition patterns. In contrast, catalytic cracking takes place through ionic intermediates, designated as carbonium ions (positively charged free radicals) generated at the catalyst surface. Although there is a certain parallelism between the modes of cracking of free radicals and carbonium ions, the latter undergo rapid intramolecular rearrangement reactions prior to cracking. This leads to more highly branched hydrocarbon structures than those from thermal cracking, and to important differences in the molecular weight distribution of the cracked products. Furthermore, the cracked products undergo much more extensive secondary reactions with the catalyst. *See* REACTIVE INTERMEDIATES.

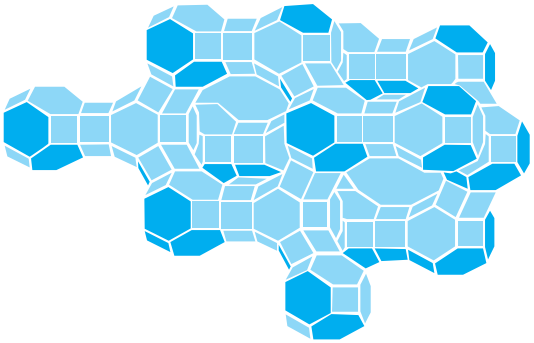

Fig. 3. Model of zeolite type Y.

The catalytic cracking mechanism also favors the production of aromatics in the gasoline boiling range; these reach quite high concentrations in the higher-boiling portion. This characteristic, together with the copious production of branched aliphatic hydrocarbons especially in the lower-boiling portion, is largely responsible for the high octane rating of catalytically cracked gasoline.

Cracking catalysts must have two essential properties: (1) a chemical composition capable of maintaining a high degree of acidity, preferably as readily available hydrogen ions (protons); and (2) a physical structure of high porosity so that these active sites are available for cracking. Mechanical durability is also necessary for industrial use.

Cracking catalysts are essentially silica-alumina compositions. A dramatic improvement in catalytic unit performance occurred with the switch from acid-treated clays (montmorillonite or kaolinite) to synthetic silica-aluminas. After 1960, a new group of aluminosilicates, molecular sieve zeolites, were introduced into the catalyst formulation. These crystalline materials (**Fig. 3**) have cracking activity 50 to 100 times the previous amorphous catalyst. They permit cracking to greater conversion levels, producing more gasoline, less coke, and less gas. *See* ZEOLITE.

As the catalyst particles pass through the reactor regenerator system every 3 to 15 min they are gradually deactivated through loss of active surface area

Representative yield structures for three different processing objectives in catalytic cracking

Process variables	Light gases	Gasoline	Light cycle oil
Feed	Light gas oil	Gas oil	Gas oil
Reactor temperature, °F (°C)	990 (532)	950–990 (510–532)	890–900 (477–482)
Light gases, wt %	4.5	2.8	1.6
Propane/propylene, vol %	15	10.0	7.5
Butane/butylene, vol %	22	16.4	11.2
Gasoline, vol %	46	69.5	32.6
Catalytic diesel oil, vol %	18	10	43.6
Bottom, vol %	5	5	5

by the effect of heat and steam and through contamination by the trace metallic components on the feedstocks, mainly sodium and vanadium. The catalyst particles also undergo mechanical attrition, and fines are lost in the reactor and regenerator gases. To compensate, fresh catalyst is added.

Cracking catalysts have evolved considerably since the mid-1960s, when zeolite crystals were incorporated into the matrix. The sodium and hydrogen ions in the crystals have been replaced with rare-earth elements to provide thermal stability. Surface area of the catalysts has been reduced to provide increased resistance to contaminant metals (Ni and V) and to reduce coke-make of the reaction.

With increased levels of metals on the catalysts, processes have been developed to passivate the effect of the metals. One process involves addition of antimony to the catalytic feedstock. Antimony thus incorporated into the catalyst can reduce hydrogen and carbon yield attributable to the metals.

The new zeolite catalysts have resulted in considerable change in the process itself. The catalysts are made more resistant to thermal degradation primarily by the exchange of sodium and hydrogen ions in the zeolite with rare-earth elements. Regenerator temperatures can be safely raised to the 1350°F (732°C) or greater level. The carbon on regenerated catalyst is reduced to the 0.05 wt % level resulting in improved gasoline yields. In addition, all the carbon monoxide produced at lower generation temperatures (10% concentration in the regenerator flue gas) can be combusted in the regenerator if desired, making for a more efficient recovery of combustion heat and reduced atmospheric pollution. The effluent CO concentration can be reduced to less than 0.05 vol %, if desired.

Additives have been developed to improve the efficiency of regeneration. Combustion promoters can ensure complete combustion of the flue gas to CO_2 even at relatively low temperatures of 1250°F (695°C). These promoters generally contain an element from the platinum group in extremely low concentrations, the ppm range. They may be added as separate powders or in admixtures with the catalyst. Other additives are being produced which are claimed to reduce sulfur oxides in the flue gas.

The high-activity zeolite catalyst has permitted units to be designed with all riser cracking (**Fig. 4**), wherein all the cracking reaction takes place in a relatively dilute (less than 2 lb/ft^3 or 3.2 kg/m^3) catalyst suspension in a 2- or 5-s residence time. No dense-bed (10–15 lb/ft^3 or 16–24 kg/m^3) cracking exists in these units.

Many old units are being converted to riser crackers, and virtually all new units feature riser cracking. Some riser crackers are also provided with small dense beds to achieve an optimum yield pattern. Some have riser-type regenerators directed to improve regenerator efficiency.

Over the years, there have been continued improvements to the process in many areas. Feeds available for processing have become much heavier, and include stocks such as vacuum residua containing contaminant metal up to 30–40 ppm. Many catalyst formulations have been developed to increase catalytic activity and to resist the high level of contaminant metals on the catalyst. Levels as high as 10,000 ppm of nickel and vanadium on catalyst have been run commercially. Nickel, vanadium, and copper catalyze detrimental competing dehydrogenation reactions that produce coke and light gases (hydrogen and methane). These heavier stocks have produced larger amounts of coke which must be burned in the regenerator, causing regenerator temperatures to rise. Staged regeneration is reported to permit temperatures as high as 1500°F (815°C), since water of combustion is liberated in the lower-temperature first stage, and thus hydrothermal deactivation is not as much a factor in the second stage. Additionally, facilities have been provided internally and externally to the regenerator to remove excess heat of combustion.

With more emphasis on environmental protection, complex facilities are provided to remove pollutants from the effluent regeneration gases. Third-stage cyclone collectors, electrostatic precipitators, and scrubbers are used commercially to meet regulations.

To conserve energy in the process, flue gas expanders are used in the flue gas system to provide more than enough energy to compress the regeneration air.

With the continuing emphasis on conservation of petroleum resources, catalytic cracking assumed a more important role. The heavier products from a refinery could be replaced by coal-, shale-, or tar-sand-derived products. This creates a need for greater conversion of petroleum to gasoline, with

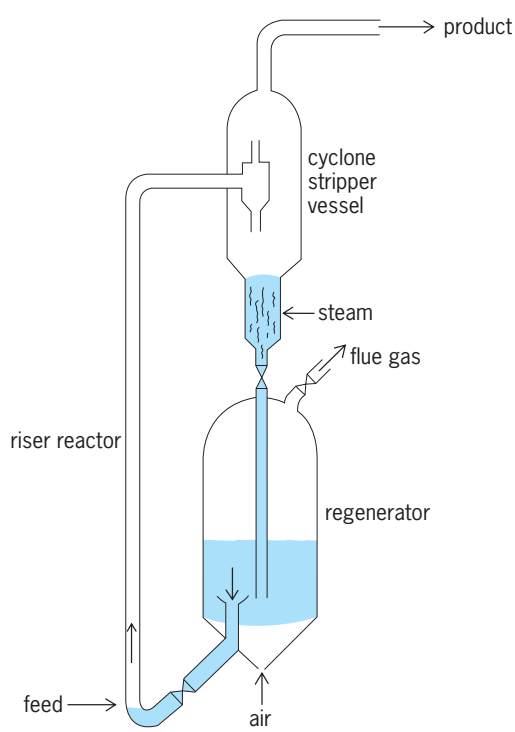

Fig. 4. Riser catalytic cracker in "pure" transfer-line configuration.

emphasis on catalytic cracking since it is the cheapest major conversion process. *See* ALKYLATION (PETROLEUM); AROMATIZATION; COAL LIQUEFACTION; FIRED HEATER; HETEROGENEOUS CATALYSIS; ISOMERIZATION; PETROLEUM PROCESSING AND REFINING.

Edward Luckenbach

Bibliography. G. D. Hobson, *Modern Petroleum Technology*, 5th ed., 1983; J. J. McKetta, *Petroleum Processing Handbook*, 1992; J. C. Speight, *The Chemistry and Technology of Petroleum*, 3d ed., revised, 1999.

Cranberry

The large-fruited American cranberry, *Vaccinium macrocarpon*, a member of the heath family, Ericaceae, is a native plant of open, acid peat bogs in northeastern North America. Selections from the wild have been cultivated since the early nineteenth century. It is an evergreen perennial vine producing runners and upright branches with conspicuous terminal flower buds (**Fig. 1**). *See* ERICALES.

Cultivation. A cranberry bog is made by removing the vegetation on a maple, cedar, or brown-brush swamp, draining it by cutting "shore," "lateral," and "main" ditches, spreading 2 or 3 in. (5 or 8 cm) of sand over the leveled peat, and inserting cuttings of the selected variety of vines through the sand into the peat. Rooting occurs readily within a month, but 3 or 4 years of growth and care are required before the first commercial crop is produced.

Care includes frost protection in spring and fall, now largely provided by solid-set, low-gallonage sprinkler systems which effect protection much quicker and with only a tenth of the water formerly required with flood frost protection. Flooding is used in winter to protect the cranberry vines, not from cold but from desiccation when the root zone is frozen. Early-season floods are used for drowning pest insects.

Because well-tended cranberry bogs continue to produce annual crops for a century or more without replanting, a half-inch (1.3 cm) layer of sand is spread every 3 or 4 years over the vines (by spreading on winter ice when thick enough) to cover the accumulating cranberry leaves on the soil surface where the cranberry girdler insect would otherwise breed and multiply.

Fig. 1. Cranberry uprights in full bloom.

Harvesting. Commercial cranberry growing is confined to Massachusetts, New Jersey, Wisconsin, Washington, and Oregon and to several provinces in Canada. Only about 20% of the cranberries are sold as fresh fruit, and most of these are dry-harvested by machines. Most of the processing fruit is harvested in flood waters, either by machines which pick and deliver the fruit into towed plastic boats or by water reels which detach the berries to float and be driven by wind to shore where they are elevated into bulk trucks. Berries for juice manufacture are usually vine-ripened and deep-frozen for a month or more prior to thawing and extraction. The older form of hand harvest with wood-toothed scoops is now gone, the scoops sometimes being used for ditch-edge picking.

Good-quality fresh cranberries can be stored for several months, refrigerated, with very little loss to decay. Good-quality cranberries can be kept in deep-freeze storage for several years with only minor moisture loss. Frozen berries on thawing are soft and juicy, unlike the firm fresh berry, and must be utilized promptly.

Cranberry bogs in full bloom in early July are a sight to remember because they have 40,000,000 to 50,000,000 white or pink flowers per acre (100,000,000 to 125,000,000 per hectare), each of which must be visited by pollinating honeybees or bumblebees to set the berries. The control of frost and winter injury, the control of weeds, insects, and diseases, and the development of more efficient harvest techniques have seen productivity rise. Supplies of the fruit are ample to satisfy both domestic and export demand.

The special requirements of the cranberry plant for low fertility, acid soil, and winter protection make it a poor choice for home garden cultivation.

Chester E. Cross

Diseases. Fruit rots are the most economically important diseases of cranberry. Eight species of fungi cause almost all of the loss attributed to these diseases.

Symptoms of all fruit rots are almost identical, and include the development of a discolored spot on the surface of the fruit; the spot enlarges rapidly until the entire berry is affected (**Fig. 2**). Later the infected berry shrivels into a raisinlike mummy.

Several factors influence the amount of fruit rot that occurs. Higher temperatures during the growing season increase fungal activity and the level of disease. High moisture levels also increase disease, because spore release, spore germination, and penetration of the host plant are all dependent on the presence of adequate moisture. Bruising during harvest, sorting, or packing significantly increases the amount of postharvest fruit rot.

Infection before harvest is common in the production areas with warmer climates and higher rainfall. In such areas fungicide sprays, at midbloom and

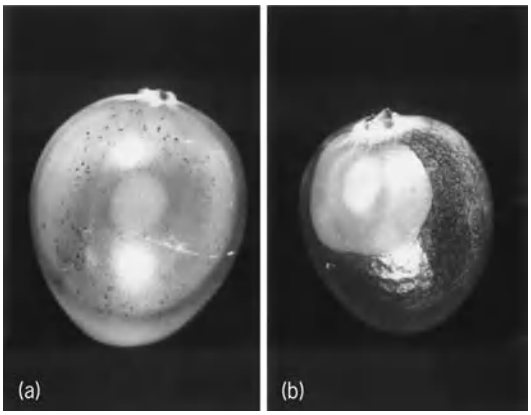

Fig. 2. Symptoms of cranberry disease. (*a*) Early rot symptoms on fruit; black specks in the affected area are the fruiting bodies of the fungus. (*b*) Early symptoms of blotch rot on fruit; all fungal fruit rots appear identical at this stage.

every 14 days until all blossoms have fallen, have provided excellent control.

Practices that reduce fruit rots include rapid drying of the fruit, reducing fruit temperatures to 36–40°F (2–4°C), and reducing the time the fruit remains in storage. Dry harvest results in less disease than wet harvest. *See* PLANT PATHOLOGY. John K. Springer

Bibliography. H. F. Bergmann, Disorders of cranberries, *Yearbook of Agriculture*, pp. 789–796, 1953; E. G. Hall and K. J. Scott, *Storage and Market Diseases of Fruit*, 1982.

Cranial nerve

Any peripheral nerve which has its central nervous system connection with the brain, as opposed to the spinal cord, and reaches the brain through a hole (foramen) in the skull. Nerve fibers are sensory if they carry information from the periphery to the brain, and motor if they carry information from the brain to the periphery. Sensory fibers are classified as somatic sensory if they come from the skin or muscle sense organs, visceral sensory if they come from the viscera, and special sensory if they come from special sense organs such as the eye and ear. Motor fibers are classified as somatic motor if they carry information to somatic striated muscles; general visceral motor if they carry information to glands, smooth muscle, or cardiac muscle; and special visceral motor if they carry information to visceral striated muscle. A cranial nerve may have only one fiber type or several; cranial nerves with several fiber types are called mixed nerves.

In mammals, 12 pairs of cranial nerves, numbered I through XII, are usually described (see **table**). However, mammals also have an anterior unnumbered nerve, the terminal nerve. Many vertebrates also have two pairs of lateral line nerves (unnumbered), and lack discrete nerves XI and XII. *See* BRAIN; NERVOUS SYSTEM (VERTEBRATE). Douglas B. Webster

Crank

In a mechanical linkage or mechanism, a link that can turn about a center of rotation. The crank's center of rotation is in the pivot, usually the axis of a crankshaft, that connects the crank to an adjacent link. A crank is arranged for complete rotation (360°) about its center; however, it may only oscillate or have intermittent motion. A bell crank is frequently used to change direction of motion in a linkage (see **illus.**). *See* LINKAGE (MECHANISM).

In mechanisms where energy input fluctuates, it is

Cranial nerves of vertebrates

Number	Name	Fiber types	Peripheral origin or destination	Vertebrates possessing this nerve
—	Terminal	Somatic sensory	Anterior nasal epithelium	Almost all
I	Olfactory	Special sensory	Olfactory mucosa	All
—	Vomeronasal	Special sensory	Vomeronasal mucosa	Almost all
II	Optic	Special sensory	Retina of eye	All
III	Oculomotor	Somatic motor	Four extrinsic eye muscles	All
IV	Trochlear	Somatic motor	One extrinsic eye muscle	All
V	Trigeminal	Special visceral motor	Muscles of mandibular arch derivative	All
		Somatic sensory	Most of head	All
VI	Abducens	Somatic motor	One extrinsic eye muscle	All
—	Anterior lateral line	Special sensory	Lateral line organs of head	Fish and larval amphibians
VII	Facial	Special visceral motor	Muscles of hyoid arch derivative	All
		General visceral motor	Salivary glands	All
		Somatic sensory	Small part of head	All
		Visceral sensory	Anterior pharynx	All
		Special sensory	Taste, anterior tongue	All
VIII	Vestibulocochlear	Special sensory	Inner ear	All
	Posterior lateral line	Special sensory	Lateral line organs of trunk	Fish and larval amphibians
IX	Glossopharyngeal	Special visceral motor	Muscles of third branchial arch	All
		General visceral motor	Salivary gland	All
		Somatic sensory	Skin near ear	All
		Visceral sensory	Part of pharynx	All
		Special sensory	Taste, posterior tongue	All
X	Vagus	Special visceral motor	Muscles of arches 4±6	All
		General visceral motor	Most viscera of entire trunk	All
		Visceral sensory	Larynx and part of pharynx	All
		Special sensory	Taste, pharynx	All
XI	Spinal accessory	Special visceral motor	Some muscles of arches 4–6	Reptiles, birds, mammals
XII	Hypoglossal	Somatic motor	Muscles of tongue and anterior throat	Reptiles, birds, mammals

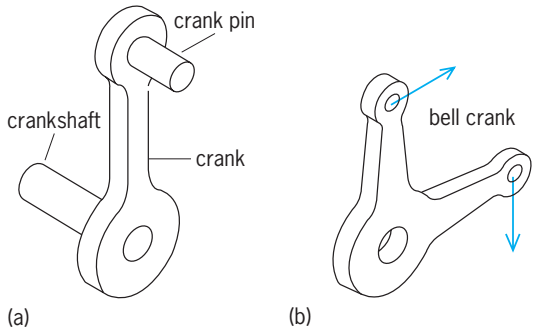

Cranks (*a*) for changing radius of rotation, and (*b*) for changing direction of translation.

sometimes useful to design a crank stoutly, as a form of flywheel, so that considerable rotational energy is stored in it, tending thereby to make the output smoother inasmuch as the inertia of such a crank absorbs peak input and releases energy more smoothly to output.

Input and output cranks of a four-bar linkage may express in their relative angular placement a relation, $\beta = \beta(\alpha)$, which is excellent in control operations where electric current computers would be vulnerable. *See* FOUR-BAR LINKAGE. Douglas P. Adams

Craton

A large, relatively stable portion of the Earth's crust. Although ocean basins originally were considered low cratons, today the term applies only to continents. Continental (high) cratons are the broad heartlands of continents with subdued topography, encompassing the largest areas of most continents. Cratons experience only broad (epeirogenic) warping and occasional faulting in contrast to the much more structurally mobile or unstable zones of continents, which include mountain ranges, such as the Himalaya, and rift zones, such as those of East Africa. The terminology used today to express these contrasts was first proposed by the German geologist L. Kober in 1921. *Kratogen* referred to stable continental platforms, and *orogen* to mountain or orogenic belts. Later authors shortened the former term to kraton or craton. A complementary term, taphrogen, coined in the 1940s, encompasses the rift structures.

The present North American craton comprises the low continental interior extending from the young Rocky Mountains east to the older Paleozoic Appalachian Mountains, and north to the Paleozoic Franklin mountain belt along the Arctic margin of Canada and Greenland. The Canadian shield is merely that part of the craton where pre-Paleozoic rocks are widely exposed today. In contrast, the remainder of the craton in the plains of western Canada and most of the central United States has little-deformed younger strata at the surface. Such Phanerozoic strata must have covered most of the shield in the past but have been eroded. Most continents have similar large Phanerozoic cratons—for example, all of South America east of the Andes Mountains; Australia west of the Great Dividing Range; Siberia in northern Asia; Europe from the Baltic Sea to the Ural Mountains; all of Africa except the northern and southern edges (orogens) and East Africa (rift valleys); India south of the Himalaya; and all but the Pacific side of Antarctica (see **illus.**). China and Southeast Asia are exceptional today because of the ongoing collision of India with Asia, which has caused intense fragmentation of former cratons there by great, still-active faults. *See* CONTINENTS, EVOLUTION OF; EARTH CRUST.

The recognition of fundamental differences in structural behavior between cratons and orogens long preceded their formal classification. In the 1850s, J. Hall recognized a profound difference between Paleozoic strata in the lowland interior of North America and the Appalachian Mountains. In the mountains, such strata were 10 times thicker, were much deformed, and were metamorphosed in the eastern ranges. Not only were the equivalent strata in the interior still nearly flat-lying and thinner, but there were also subtle contrasts in lithology. Hall postulated a simplistic cause-and-effect relationship, with greater thickness of strata along the eastern continental margin causing subsidence of the crust, which in turn resulted in crumpling of those same strata. In 1873 J. Dana proposed instead that cooling of the earth had caused buckling of the margin between the continental and oceanic crust. The Appalachian region was a downbuckled zone termed a geosyncline, which could accommodate the thick sediments. Eventually this zone could bend downward no more and failed due to compression, which caused the crumpling and elevation of the strata to form the Appalachian Mountains. European geologists also were highlighting the structural contrast between cratons (or forelands) and orogens. Soon the theory of continental drift, proposed in 1910 by F. Taylor and in 1912 by A. Wegener, presented a new mechanism for explaining orogens as the result of the collision between two continental cratons or between a craton and some smaller tectonic element such as an island archipelago. This new explanation eventually evolved into the modern theory of plate tectonics. *See* CONTINENTAL DRIFT; CONTINENTAL MARGIN; OROGENY; PLATE TECTONICS.

Today, cratons are believed to comprise bits of old continental crust, which have had long and complex histories involving the overprinting of many tectonic events. Continents can be viewed as great collages of tectonic elements amalgamated together at different times and in different ways (see illus.). Therefore, the delineation of both cratons and orogens has an important temporal element: What is now a craton may be made up of interlaced orogens of differing ages, each of which may include bits of mangled older cratons within them. For example, 1-billion-year-old orogens may have been fused to or may have cut across orogens of 2-billion-year age, which in turn may surround a cratonic kernel more than 3 billion years old. Returning to the present, most of the North American craton has been a relatively

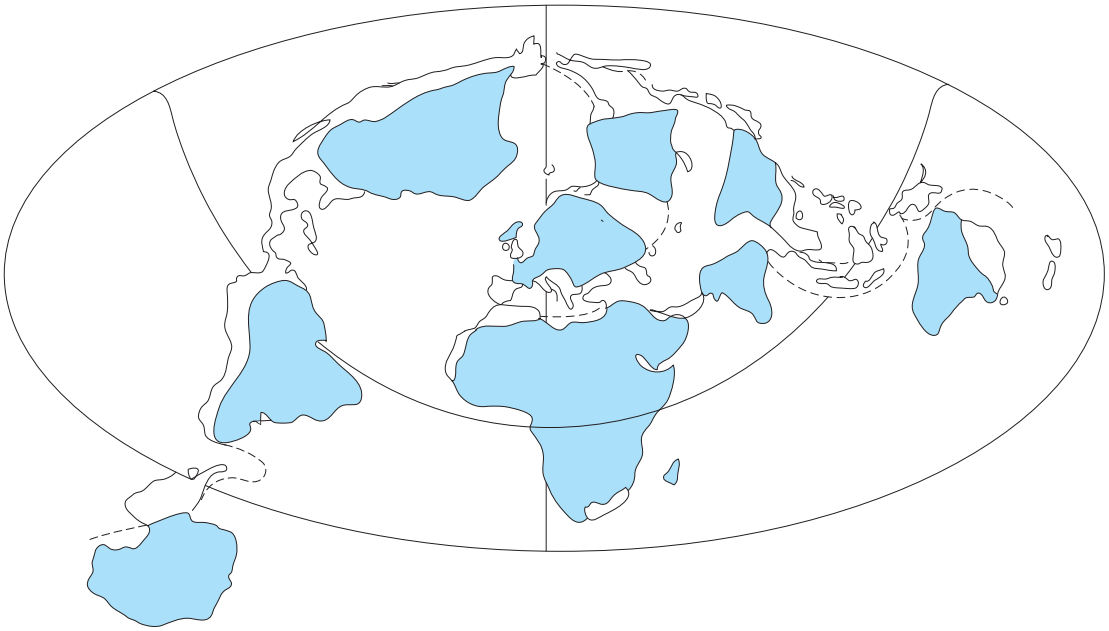

World map showing Phanerozoic continental cratons (colored areas) of the last 500–600 million years. Each craton is an amalgam of older tectonic elements, which include fragments of complexly interlaced pre-Phanerozoic cratons and orogens. (*After R. Dott and R. L. Batten, Evolution of the Earth, 8th ed., McGraw-Hill, 1988*)

stable element for at least the past 600 million years. Prior to that, the configuration among cratons and orogens was very different, and only the ghosts of some of the fragmentary components of the present continent are discernible. Refined studies of seismic waves that have penetrated the Earth's deep interior indicate that beneath cratons there is anomalous mantle with distinctive compositional or thermal characteristics. Such roots imply a long-term linkage between the evolution of early continental crust and underlying mantle, which somehow contributed to the survival of fragments of cratons older than 2.5–3 billion years. Robert H. Dott, Jr.

Bibliography. R. Dott, Jr., James Hall's discovery of the craton, *Geol. Soc. Amer. Cent. Spec. Pub.*, 1:157–167, 1985; M. Kay, North American geosynclines, *Geol. Soc. America Mem.*, no. 48, 1951; E. Suess, *Das Antlitz der Erde*, vol. 3, transl. by H. Sollas, *The Face of the Earth*, Oxford, Clarendon, 1908.

Creep (materials)

The time-dependent strain occurring when solids are subjected to an applied stress. *See* STRESS AND STRAIN.

Kinds of phenomena. Some of the different kinds of creep phenomena that can be exhibited by materials are shown in the **illustration**. The strain $\epsilon = \Delta L/L_0$, in which L_0 is the initial length of a body and ΔL is its increase in length, is plotted against the time t for which it is subjected to an applied stress. The most common kind of creep response is represented by the curve A. Following the loading strain ϵ_0, the creep rate, as indicated by the slope of the curve, is high but decreases as the material deforms during the primary creep stage. At sufficiently large strains, the material creeps at a constant rate. This is called the secondary or steady-state creep stage. Ordinarily this is the most important stage of creep since the time to failure t_f is determined primarily by the secondary creep rate $\dot{\epsilon}_s$. In the case of tension creep, the secondary creep stage is eventually interrupted by the onset of tertiary creep, which is characterized by internal fracturing of the material, creep acceleration, and finally failure. The creep rate is usually very temperature-dependent. At low temperatures or applied stresses the time scale can be thousands of years or longer. At high temperatures the entire creep process can occur in a matter of seconds. Another kind of creep response is shown by curve B. This is the sort of strain-time behavior observed when the applied stress is partially or

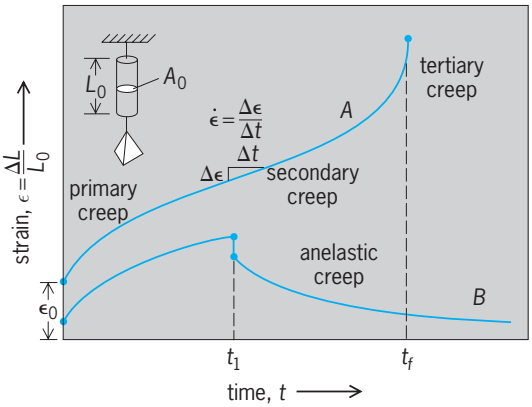

Typical creep curves for materials.

completely removed in the course of creep. This results in time-dependent or anelastic strain recovery.

Examples. Creep of materials often limits their use in engineering structures. The centrifugal forces acting on turbine blades cause them to extend by creep. In nuclear reactors the metal tubes that contain the fuel undergo creep in response to the pressures and forces exerted on them. In these examples the occurrence of creep is brought about by the need to operate these systems at the highest possible temperatures. Creep also occurs in ordinary structures. For example, when a bolt is tightened to fasten two parts together, the tension stress in the bolt can be quite high. Over a long period of time the bolt creeps, with the result that the force holding the two parts together diminishes. Another example of stress relaxation caused by creep is found in prestressed concrete beams, which are held in compression by steel rods that extend through them. Creep and stress relaxation in the steel rods eventually leads to a reduction of the compression force acting in the beam, and this can result in failure. *See* PRESTRESSED CONCRETE.

There are many examples of creep that are not related to high technology. The window panes in very old homes or cathedrals are often thicker at the bottom than they are at the top due to creep under the force of gravity. Similarly, when a glacier flows down the side of a mountain the ice undergoes extensive creep deformation.

Mechanism. The mechanism of creep invariably involves the sliding motion of atoms or molecules past each other. In amorphous materials such as glasses, almost any atom or molecule within the material is free to slide past its neighbor in response to a shear stress. In plastics, the long molecular chains can slide past each other only to a limited extent. Such materials typically show large anelastic creep effects (curve B in the illustration).

For crystalline materials, creep deformation also involves the sliding of atoms past each other, but here the sliding can occur only within the cores of crystal dislocations. Thus, creep of metals and ceramics is usually governed by the motion of dislocations. In addition to gliding along their slip planes, the dislocations can climb out of their slip planes by the absorption or emission of lattice vacancies. This process is controlled by the rate at which lattice vacancies diffuse to or from climbing dislocations and gives rise to the strong temperature dependence of creep.

Materials with creep resistance. Understanding of the mechanisms that control creep deformation makes it possible to design materials with superior creep resistance. When solute atoms are added to metals, they are attracted to the strain fields of the dislocations. There they inhibit dislocation motion and in this way improve the creep resistance. Many of the aluminum alloys used for aircraft structures are strengthened in this way. The addition of second-phase particles to alloys is another way to improve the creep resistance. The most effective strengthening phases are oxides, carbides, or intermetallic phases, because they are usually much stronger than the host metal and therefore create strong obstacles to dislocation motion. Materials containing finely dispersed, strong particles of a stable phase are usually very creep-resistant. Nickel-based superalloys, used in gas-turbine engines, derive their creep resistance from these effects. *See* ALUMINUM ALLOYS; CRYSTAL DEFECTS; DIFFUSION; HIGH-TEMPERATURE MATERIALS; METAL, MECHANICAL PROPERTIES OF; PLASTICITY; RHEOLOGY.

William D. Nix; Jeffery C. Gibeling; Kevin J. Hemker

Bibliography. M. F. Ashby and L. M. Brown (eds.), *Perspectives in Creep Fracture*, 1983; Z. P. Bazant (ed.), *Mathematical Modelling of Creep and Shrinkage of Concrete*, 1989; A. C. Cocks (ed.), *Mechanics of Creep Brittle Materials*, 1991; H. E. Evans, *Mechanisms of Creep Fracture*, 1984; F. R. Nabarro and H. Filmer, *Physics of Creep and Creep-Resistant Alloys*, 1993; J. P. Poirier, *Creep of Crystals*, 1985; J. J. Skrzypek, *Plasticity and Creep: Theory, Examples, and Problems*, 1993; B. Wilshire, *Introduction to Creep*, 1993.

Creeping flow

Flow in which the Reynolds number is very small. Fluid dynamics encompasses an intriguing array of phenomenologically rich processes, from weather prediction in the atmosphere to the swimming of the tiniest microorganism in its natural environment. A qualitative measure of the expected behavior in a given fluid dynamics system is found from the Reynolds number, Re = density · length · velocity/viscosity. In the case where Re is very small (much less than 1, usually), the inertia effects (momentum, acceleration, and body forces) become negligible in comparison to the viscous resistance. This situation is described as creeping flow. *See* REYNOLDS NUMBER; VISCOSITY.

Creeping flows exist typically when the fluids involved are very viscous (oil, syrup) or the length scales involved are very small (swimming microorganisms, microelectronic machines, flows in cracks or pores, the transport of sediment or pollutants). In many cases of creeping flow, it is possible to significantly simplify the governing Navier-Stokes equations, which are highly nonlinear, by setting the Reynolds number identically to zero. Under this simplification, a set of equations is obtained called the Stokes equations, which have the decided advantage of being linear. As a result, it is possible to draw upon the principle of linear superposition as well as complex variable theory in search of mathematical solutions. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES; LINEARITY; NAVIER-STOKES EQUATIONS; SUPERPOSITION PRINCIPLE.

An example of one creeping flow solution is the flow of a uniform velocity field $\mathbf{u} = (U, 0, 0)$ past a stationary sphere of radius a. The Stokes stream function ψ is given by Eq. (1), where (r, θ) are the

$$\psi(r, \theta) = \frac{1}{2} U \left(\frac{a^3}{2r} - \frac{3ar}{2} + r^2 \right) \sin^2 \theta \qquad (1)$$

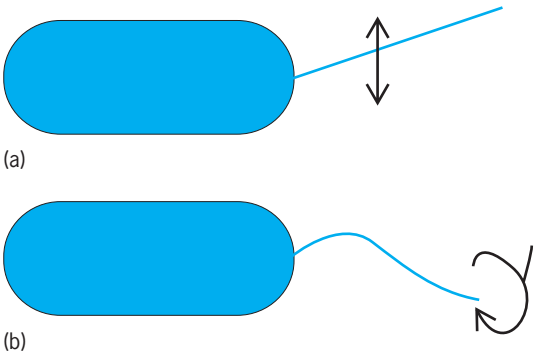

Aquatic propulsion by (*a*) a mechanical fish and (*b*) a swimming microorganism.

polar coordinates about the sphere and lines of constant values of ψ are stream lines for the flow. From this solution, the drag D acting on this sphere can be shown to be given by Eq. (2), where μ is the fluid

$$D = 6\pi\mu Ua \tag{2}$$

viscosity. For small values of Re, inertial effects can be incorporated through a small correction to this drag formula, known as the Oseen correction. *See* STREAM FUNCTION.

Another consequence of linearity is reversibility: streamlines remain unchanged when the direction of motion is reversed. This has profound implications for aquatic animal propulsion in creeping flow. For example, a mechanical "fish" that has a tail which flaps back and forth (**illus.** *a*) may swim quite readily in the bath tub, but will make no progress in a very viscous fluid such as corn syrup. Since most swimming microorganisms exist in a low-Reynolds-number environment, they must adopt a non-time-reversible motion, such as having a corkscrew-type tail (illus. *b*). *See* FLUID FLOW; FLUID-FLOW PRINCIPLES; FLUID MECHANICS. Mary-Catherine Kropinski

Bibliography. S. W. Churchill, *Viscous Flows: The Practical Use of Theory*, Butterworth, Stoneham, MA, 1988; V. N. Constantinescu, *Laminar Viscous Flow*, Springer-Verlag, New York, 1995; J. Happel and H. Brenner, *Low Reynolds Number Hydrodynamics*, Kluwer Academic, Dordrecht, The Netherlands, 1991; W. E. Langlois, *Slow Viscous Flow*, Macmillan, New York, 1964.

Creodonta

An extinct order of mammals containing the dominant carnivores in North America, Europe, Asia, and Africa during much of the Cenozoic. Its members roamed the Earth for more than 50 million years, then became extinct less than 9 million years ago. Creodonts were not part of the order Carnivora, but they independently evolved carnivorous specializations in their teeth and limbs. They ranged from tiny (*Isohyaenodon* from Africa was the size of a small weasel) to gigantic (*Hyainailouros* from Europe and Asia was one of the largest land-dwelling mammalian carnivores ever). The wolflike genus *Hyaenodon* is probably the most familiar creodont among the more than 180 known species. Most lived in North America and Europe, but new species from Africa and Asia are being described at an incredible rate. *See* DENTITION.

There are two major families within Creodonta: Hyaenodontidae and Oxyaenidae. Hyaenodontids are the more diverse, with more than 140 known species. They are first recognized from the mid-Paleocene of eastern Asia. They had emigrated to North America and Europe by the earliest Eocene. Their greatest diversity was during the latest Eocene and Oligocene, when more than 30 species of the genus *Hyaenodon* (see **illus.**) were doglike predators, feeding on the variety of hoofed, plains-dwelling mammals in North America, Asia, and Europe. The dentition of *Hyaenodon* was catlike, and it has been estimated that the diet of these animals was 70–100% meat. Their bodies, however, were wolflike, and they may have hunted in a manner more common to canids than felids. Early hyaenodontids were more diverse: *Prolimnocyon* was an arboreal, genet-like animal, *Prototomus* was a generalized, raccoon-like creature, and *Pterodon* was large and bearlike. Some hyaenodontids were carnivores (*Prototomus*), others were probably omnivores (*Masrasector*), some were insectivorous (*Thereutherium*), and some may have been mollusk eaters (*Teratodon* and *Apterodon*). The range of locomotory and dietary adaptations in Hyaenodontidae was probably almost as great as in Carnivora, although there were no fully aquatic creodonts (*Apterodon* from Africa may have been partially aquatic). The last hyaenodontid, *Dissopsalis*, lived in southern Asia and eastern Africa and became extinct only about 8 million years ago. Oxyaenids were less diverse (only about 40 species are known) and reached their peak in the latest Paleocene and earliest Eocene. By the late Eocene, the family had become completely extinct. While oxyaenids were specialized, the range of niches they occupied was not as great as hyaenodontids. Many oxyaenids, such as *Oxyaena* and *Dipsalidictis*, were like large wolverines, although the largest, *Sarkastodon* from Mongolia, was almost bearlike. There were a few saber-toothed oxyaenids, including *Apataelurus* from the mid-Eocene of North America; unfortunately, little is known about these.

The formal classification of Creodonta has changed considerably, largely because the relationships among the various creodont groups and other placental mammals are still poorly understood. In the 1950s, the order included many mammalian groups that are now considered part of Carnivora (true carnivores), Lipotyphla (insectivores), or Ungulata (hoofed mammals). Creodonts later formed the core of the order Deltatheridia; however, that grouping is no longer used, largely because the name-bearing genus *Deltatheridium* is now known to be related to marsupials. Improved knowledge about the relationships of placental mammals has resulted in many former creodont families being moved to other orders

so that only oxyaenids and hyaenodontids remain. It is unclear how Hyaenodontidae and Oxyaenidae are related to one another, and their grouping as Creodonta may turn out to be unjustified.

It was long thought that the extinction of the creodonts was caused by competition from better-adapted members of the order Carnivora. However, the timing of creodont evolution does not support this theory. It appears that carnivores evolved before creodonts, the former being first known from the earliest Paleocene, while the latter first evolved in the mid to late Paleocene. Creodonts and carnivores diversified side by side through the Eocene, apparently playing different ecological roles. The persistence of creodonts until the end of the Miocene (more than 50 million years of coexistence with Carnivora) also belies the competitive scenario. Furthermore, studies of creodont adaptations do not support the long-touted idea that they were inferior to contemporary true carnivores. It is more likely that chance and changing conditions, including the evolution and extinction of creodont prey species, explain their slow demise. *See* CARNIVORA; MAMMALIA. P. David Polly

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988; J. Mellett, The Paleobiology of North American Hyaenodon, *Contrib. Vert. Evol.*, 1:1–134, 1977; R. J. G. Savage and M. R. Long, *Mammal Evolution: An Illustrated Guide*, 1986.

Cress

A prostrate hardy perennial crucifer of European origin belonging to the plant order Capparales. Watercress (*Nasturtium officinale*) is generally grown in flooded soil beds and used for salads and garnishing. Propagation is by seed or stem cuttings. High soil moisture is necessary. Leafy stems are cut usually 180 days after planting. Virginia is an important producing state.

Garden cress (*Lepidium sativum*) is a cool-season annual crucifer of western Asian origin grown for its flavorful leaves. Propagation is by seed, and leaves are harvested approximately 2 months after planting.

Upland or spring cress (*Barbarea verna*) is a biennial crucifer of European origin, grown and harvested similarly to garden cress but of lesser commercial importance. *See* CAPPARALES. H. John Carew

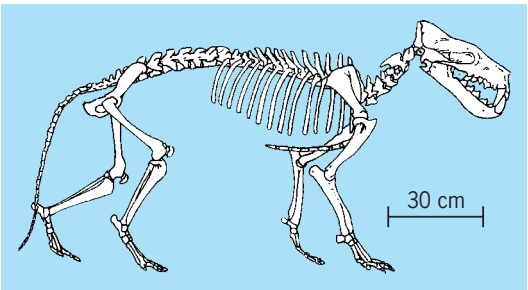

The creodont *Hyaenodon*. (*After A. S. Romer, Vertebrate Paleontology, 3d ed., University of Chicago Press, 1966*)

Cretaceous

In geological time, the last period of the Mesozoic Era, preceded by the Jurassic Period and followed by the Tertiary Period. The rocks formed during Cretaceous time constitute the Cretaceous System. Omalius d'Halloys first recognized the widespread chalks of Europe as a stratigraphic unit. W. O. Conybeare and W. Phillips (1822) formally established the period, noting that whereas chalks were remarkably widespread deposits at this time, the Cretaceous System includes rocks of all sorts and its ultimate basis for recognition must lie in its fossil remains. *See* CHALK; FOSSIL; JURASSIC; ROCK AGE DETERMINATION; STRATIGRAPHY; TERTIARY.

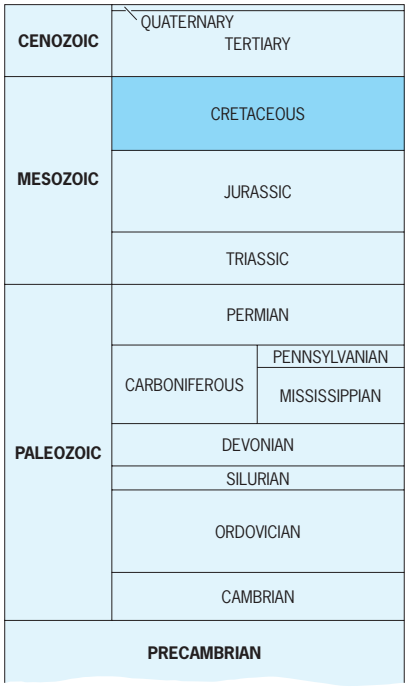

The parts of the Earth's crust that date from Cretaceous time include three components: a large part of the ocean floor, formed by lateral accretion; sediments and extrusive volcanic rocks that accumulated in vertical succession on the ocean floor and on the continents; and intrusive igneous rocks such as the granitic batholiths that invaded the crust of the continents from below or melted it in situ. The sedimentary accumulations contain, in fossils, the record of Cretaceous life. The plutonic and volcanic rocks are the chief source of radiometric data from which actual ages can be estimated, and suggest that the Cretaceous Period extended from 144 million years to 65 ± 0.5 million years before present (**Fig. 1**).

Subdivisions and time markers. Changes in life, by the development of new species and by the loss of old, are not uniform but somewhat step-like. Larger steps led to the recognition of periods such as the Cretaceous, while lesser steps within the Cretaceous delineate 12 globally recognized subdivisions, or stages (Fig. 1). In marine

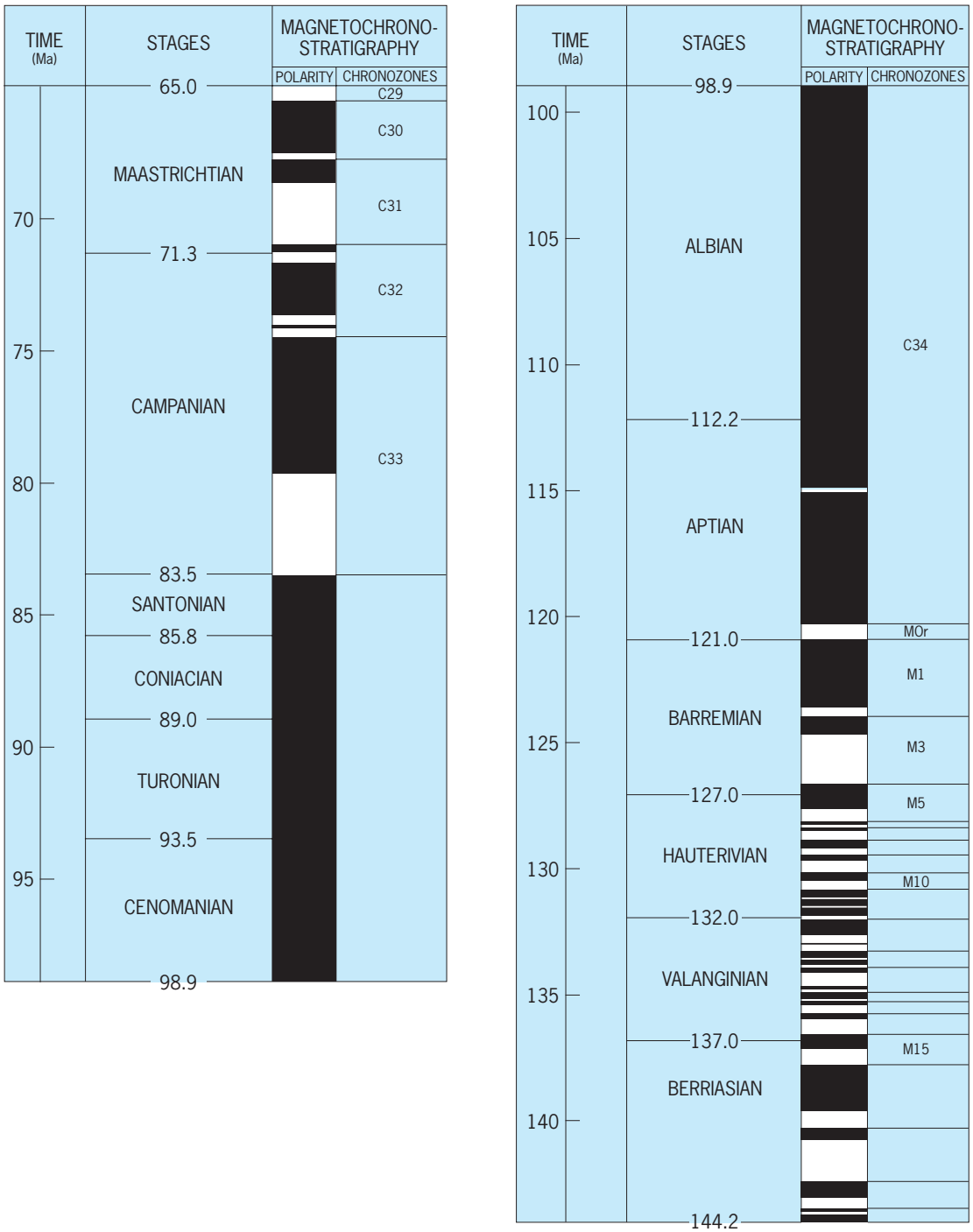

Fig. 1. Stages of the Cretaceous Period, their estimated ages in years before present, and polarity chrons representing the alternation between episodes of normal (black) and reversed (white) orientations of the Earth's magnetic field. (*After F. M. Gradstein et al., in W. A. Berggren et al., eds., Geochronology, Time Scales and Global Stratigraphic Correlations, SEPM Spec. Publ., no. 54, 1995*)

sediments the appearance and disappearance of individual, widely distributed species allows further time resolution by so-called zones. Initially these zones were largely based on ammonites, a species-rich and rapidly evolving, now extinct group of cephalopods, closest to the squids and octopuses but resembling the pearly nautilus in possession of a coiled calcareous hydrostatic shell. Over 100 successive ammonite zones have been recognized, but due to the provinciality of some species and the rarity of many, the dating of Cretaceous marine sediments now mainly devolves on microscopic fossils of the calcareous plankton. The most important of these are protozoans—the infusorial (tintinnid) calpionellids and the sarcodine planktonic foraminiferans. Their shells, in the range of 0.1–1 mm, occurring by hundreds if not thousands in a handful of chalk, have furnished about 38 pantropical zones. Next in importance are the even tinier (0.01-mm) armor plates of "nanoplanktonic" coccolithophores—

prymnesiophyte algae, best studied by electron microscopy. Thousands may be present in a pinch of chalk, and 24 zones have been recognized. *See* CEPHALOPODA; COCCOLITHOPHORIDA; FORAMINIFERIDA; MARINE SEDIMENTS; MICROPALEONTOLOGY; PALEONTOLOGY; PHYTOPLANKTON; PROTOZOA; ZOOPLANKTON.

The Earth's magnetic field reversed about 60 times during Cretaceous time, and the resulting polarity chrons have been recorded in the remanent magnetism of many rock types (Fig. 1). The actual process of reversal occurs in a few thousand years and affects the entire Earth simultaneously, providing geologically instantaneous time signals by which the continental and volcanic records can be linked to marine sequences and their fossil zonation. Noteworthy here is the occurrence of a very long (32-million-year) interval during which the field remained in normal polarity. *See* PALEOMAGNETISM.

Geography. During Cretaceous time the breakup of Gondwana, the great late Paleozoic-Triassic supercontinent, became complete (**Fig. 2**). For pre-Cretaceous time the positions of continents can be roughly derived from the remanent magnetic vectors in their rocks, which record the directions toward and the distance to the poles. But for Cretaceous and later times the positions may be more precisely tracked by the "growth lines" of crustal plates, visible in magnetic maps of the oceanic areas. As the Earth's magnetic field reversed, the continuous growth of ocean floors along the mid-oceanic ridge system occurred alternately in normal and in reversed polarity. Imprinted in the rocks, the remanent magnetism reinforces the present magnetic field above the crust grown in normal polarity and diminishes it over that grown in reversed fields.

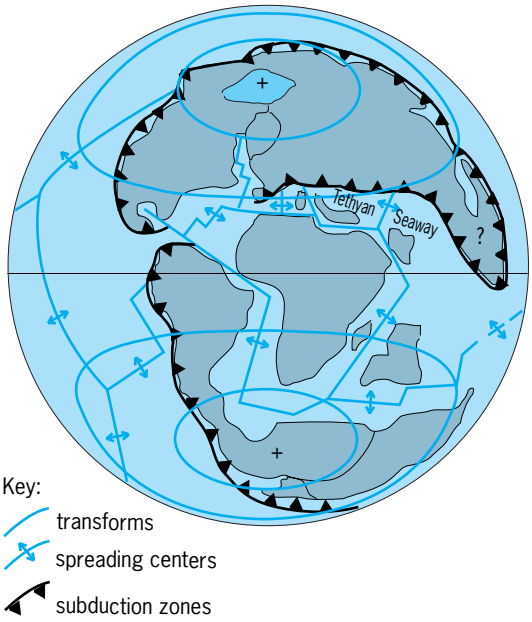

Fig. 2. Late Cretaceous global tectonics, showing inferred lithosphere plate configurations with spreading, subduction, and transform boundaries; generalized continental outlines are shown for reference. (*After R. H. Dott, Jr., and R. L. Batten, Evolution of the Earth, 3d ed., McGraw-Hill, 1988*)

Laurasia had already separated from Africa by the development of Tethys and became split into North America and Eurasia by the opening of the North Atlantic (though tenuous land connections continued to exist in the far north). These new, deep oceanic areas continued to grow in Cretaceous time. India broke away from Australia and Australia from Antarctica. South America tore away from Africa by the development of the South Atlantic Ocean, while India brushed past Madagascar on its way north to collide with southeast Asia. As these new oceanic areas grew, comparable areas of old ocean floor plunged into the mantle in subduction zones such as those that still ring the Pacific Ocean, marked by deep oceanic trenches and by the development of mountain belts and volcanism on adjacent continental margins (**Fig. 3**). *See* CONTINENTS, EVOLUTION OF; PLATE TECTONICS; SUBDUCTION ZONES.

The face of the globe was also affected by changes in sea level. Sea level at times in the early Cretaceous stood at levels comparable to the present, but subsequently the continents were flooded with relatively shallow seas to an extent probably not attained since Ordovician-Silurian times. Maximal flooding, in the Turonian Stage, inundated at least 40% of present land area. Cretaceous seas covered most of western Europe, though old mountain belts such as the Caledonides of Scandinavia and Scotland remained dry and archipelagos began to emerge in the Alpine belt. In America (Fig. 3), seas flooded the southeastern flank of the Appalachian Mountains, extended deep into what is now the Mississippi Valley, and advanced along the foredeep east of the rising Western Cordillera to link at times the Gulf of Mexico with the Arctic Ocean. In the far west, accretionary prisms of subducted Cretaceous deep-water fans and oceanic sediments, partly mixed with ophiolites in the Franciscan melange, became juxtaposed with forearc sediments.

Large seas extended over parts of Asia, Africa, South America, and Australia. The wide spread of the chalk facies is essentially due to this deep inundation of continents, combined with the trapping of detrital sediments near their mountain-belt sources, in deltas or in turbidite-fed deep-water fans such as the classical Alpine Flysch. At the same time, carbonate platforms were still widespread, and the paratropical dry belts were commonly associated with evaporite deposits. *See* PALEOGEOGRAPHY; SALINE EVAPORITES.

Tectonic and igneous activity. The North American plate, in the process of separating from Europe, continued to have a western convergent margin of growing mountains, and a southeastern passive margin. Convergent or active margins such as those surrounding the Pacific Ocean (then as now a "ring of fire") became intruded by great batholiths of granitic composition, such as those of the British Columbia Coast Range, the Sierra Nevada, the Peninsular Ranges of southern and Baja California, and the great Andean batholiths. Above these there accumulated superstructures of andesitic volcanics. Exotic

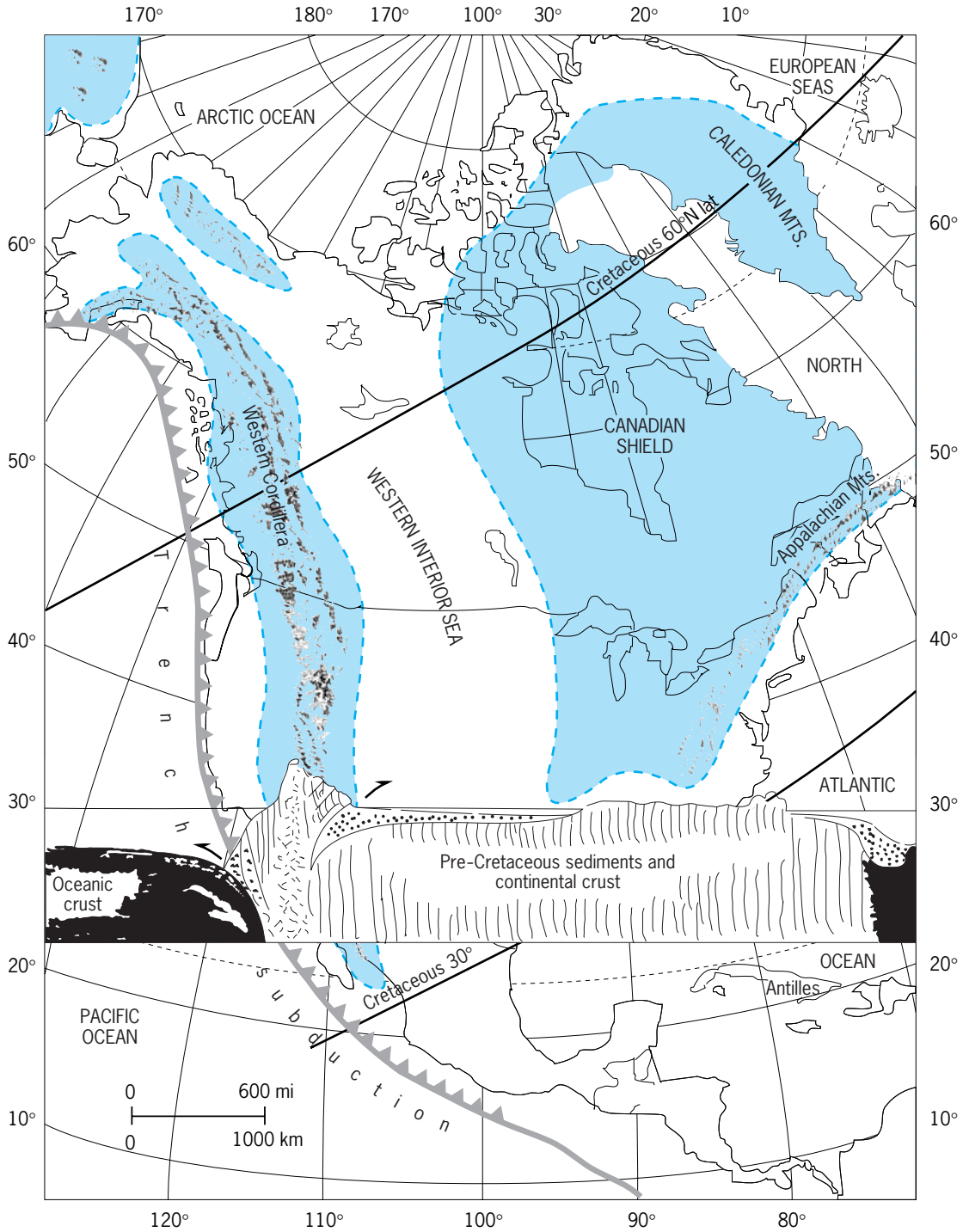

Fig. 3. Schematic view of North America in Cretaceous time. Cretaceous seas are shaded, and Cretaceous sediments in cross section (in the black-and-white strip) are dotted. Note the counterclockwise rotation of the continent shown by the lines of Cretaceous latitude.

island arc terranes riding on the Pacific plate but resisting subduction came to be welded to this margin of North America. The eastern margin of this Western Cordillera overthrust and incorporated the margin of the subsiding Western Interior Foredeep, presaging development of the Rocky Mountains (Fig. 3). Passive continental margins, such as those bordering the Atlantic, came to be sites of massive sediment accumulations, the classical geosynclines of older literature, from which mountain systems have developed locally as in the Caribbean or have yet to arise. *See* GEOSYNCLINE.

Flood basalts of Cretaceous time include the early Aptian outpourings of basalts on the mid-Pacific floor, perhaps the largest on record. Early Cretaceous emplacement of the Rajmahal basalts of India and Bangladesh was followed by the eruption of the Deccan basalts of India, emplaced during the

short magnetic interval (chron 29R) that includes the Cretaceous/Tertiary (K-T) boundary. The beginnings of the "Brito-Arctic" basalts also go back to Maastrichtian time.

Climate and oceans. In parts of early Cretaceous time, ice extended to sea level in the polar regions, as shown by glendonites, concretionary calcite pseudomorphs of the cold-water mineral ikaite; and in marine mudstones containing exotic pebbles, dropstones melted out of icebergs. But during most of Cretaceous time, climates were in the hothouse or greenhouse mode, showing lower latitudinal temperature gradients. Tropical climates may have been much like present ones, and paratropical deserts existed as they do now, but terrestrial floras and faunas suggest that nearly frost-free climates extended to the polar circles as did abundant rainfall, and no ice sheets appear to have reached sea level.

The cause for these climatic changes may be sought in variations in the heat-retaining character of the atmosphere (abundance of greenhouse gases such as carbon dioxide or methane) or in oscillations of solar luminosity. The hydrologic cycle appears to have been intensified, with more heating and therefore more water evaporation in the tropics, leading to massive transport of water vapor to the polar regions warmed by its condensation.

With a mean temperature of about 38°F (3°C), the present ocean lies 27°F (12°C) below the mean surface temperature of the Earth. This refrigerator is maintained by the sinking of cold waters in the circumpolar regions. Temperatures calculated from the oxygen isotope ratios in Cretaceous deep-water foraminiferans yield values as high as 62°F (16°C), and suggest an ocean much closer to mean surface temperature. In this ocean circulation, visualized long ago by T. C. Chamberlin, warm saline waters of the dry paratropics were the main source of dense waters and sank to the bottom when only moderately cooled. The implication is not only altered current directions and deep-water temperature regimes but also a diminished oxygen supply to the depths. This diminished oxygen is documented by the widespread development of black shales, deposited on bottoms from which most or all scavengers were excluded by lack of oxygen. The resulting excess burial of organic matter yielded an abundance of petroleum source beds, reflected in the large petroleum reserves in Jurassic and Cretaceous rocks. Occasional episodes of black shale deposition, such as that of the Bonarelli event near the end of the Cenomanian Stage, became circumglobal and buried enough organic matter to alter the carbon isotope ratios in the entire ocean-atmosphere system, but the circumstances of their origin remain unknown. *See* OCEANOGRAPHY; OIL SHALE; PALEOCLIMATOLOGY; PETROLEUM RESERVES.

Life. A walk along a Cretaceous beach or a visit to a surf-beaten cliff would have yielded many snail and bivalve shells and sea urchins differing from living ones only at the level of species. Yet there are some notable differences. Two families of bivalves rose to particular prominence in Cretaceous time. The inoceramids were a widespread and diversified group that developed species 5 ft (1.5 m) wide, the largest of all known clams, while the rudistids came to resemble corals in form and largely replaced corals as reef builders in Cretaceous time. Among cephalopods, ammonites were prominent swimmers in Cretaceous seas, as were belemnites—squids with a heavy calcareous skeleton. The calcareous microplankton had only in latest Jurassic time reached an abundance sufficient to produce widespread chalks, and in the Cretaceous two major elements of this, the prymnesiophyte coccolithophorids and the planktonic foraminiferans, reach maximal species diversity. Bony fishes evolved extensively in Cretaceous time, but prominent among marine predators were reptiles, some of them gigantic. Marine turtles have persisted, the dolphinlike ichthyosaurs died out in Turonian-Coniacian time, while the massive, sea-lion-like plesiosaurs and the lizard-derived mosasaurs lasted to the end of the period. Flightless birds populated some of the late Cretaceous seas. *See* BIVALVIA; GASTROPODA; REPTILIA.

On land, flowering plants (angiosperms) first appeared in early Cretaceous time, as opportunistic plants in marginal settings, and then spread to the understory of woodlands, replacing cycads and ferns. In late Cretaceous time, evergreen angiosperms, including palms, thus came to dominate the tropical rainforests. Evergreen conifers maintained dominance in the drier midlatitude settings, while in the moist higher latitudes forests of broad-leaved deciduous trees dominated. Grasslands, however, were not developed until Tertiary time. Insects became highly diverse, and many modern families have their roots in the Cretaceous. Amphibians and small reptiles were present. Larger land animals included crocodiles and crocodilelike reptiles, turtles, and dinosaurs. Dinosaurs, derived from reptilian stock in Triassic time, rose to become the largest land animals of all time. The very size of these animals implies complex circulatory, digestive, and respiratory systems unlike those of reptiles. Evidence for temperature control is emerging, and it is speculated that some dinosaurs were feathered. Thus, there is growing support for classification of dinosaurs not as a branch of the class Reptilia but as a separate class. The mammals, another reptilian offshoot in the Triassic, remained comparatively minor elements in the Cretaceous faunas. In early Cretaceous time, egg-laying and marsupial mammals were joined by placentals, but Cretaceous mammals were in general small, and lack of color vision in most modern mammalians suggests a nocturnal ancestry and a furtive existence in a dinosaurian world. Birds had arisen, from dinosaurs in Jurassic time, but their fossil record from the Cretaceous is poor and largely one of water birds. More common are the remains of flying reptiles, the pterosaurs, which in Cretaceous time reached a wingspan of 35 ft (11 m), the largest known flying animals by far. *See* DINOSAURIA; MAMMALIA; MARSUPIALIA; PTEROSAURIA.

K/T boundary crisis. Most species of the inoceramid and rudistid bivalves died out in a crisis within the Maastrichtian Stage. Some millions of years later,

during the reversed magnetic interval known as chron 29R, a collision with an asteroid or comet showered the entire Earth with impact debris, preserved in many places as a thin "boundary clay" enriched in the trace element iridium. This event coincided with the great wave of extinctions—the K/T crisis—which serve to bound the Cretaceous (Kreide) Period against the Tertiary.

The Chicxulub crater is located at the northwestern edge of what was then the Yucatán Carbonate Platform and now the Yucatán Peninsula. Buried under some thousands of feet of Tertiary limestone, it forms a geophysical anomaly explored for oil by a number of boreholes. With a diameter of 110 mi (180 km) excluding possible peripheral rings, this is the largest crater known from Phanerozoic time. Like comparable craters on the Moon and Mars, it has a central peak and peripheral ejecta blanket. It is filled with partly melted impact debris (suevite) derived from sediments and the underlying igneous-metamorphic complex. Quartz grains with shock lamellae confirm the impact origin, and abundant glass yields the K/T boundary age of 65 $\pm$ 0.5 million years. The ejecta blanket, where exposed at the surface in the Mexico-Belize border country 300 mi (400 km) distant, consists mainly of middle to late Cretaceous carbonates in blocks up to 10 ft (3 m), mixed with carbonate lapilli and altered glass lumps. Rare cobbles of limestone and chert were melted and quenched. In surrounding regions, deeper-water marls of latest Cretaceous and earliest Tertiary age are separated by a few feet of unusual sediment: a thin bed contains chips of limestone, blebs of glass generally altered to clay, and shocked quartz. In the Gulf of Mexico region this is followed by cross-beds interpreted as deposits of a "tidal wave" (tsunami) that swept repeatedly across the Gulf. A centimeters-scale layer of clay, extending around the world, contains microscopic silicates, while electron microscopy reveals crystals of spinel with excessive degrees of oxidation. Shocked quartz, abundant and relatively coarse in America, becomes scarce and fine with distance.

The iridium content associated with the fallout implies an impactor with a composition similar to that of planetary interiors. The size, estimated at 6 mi (10 km), implies an asteroid or a correspondingly larger comet. Traveling at 16 mi/s (25 km/s), the body would have struck with the explosive force of 10^{14} tons of TNT. Quite aside from local and regional devastation, global effects must have included earthquake shock many orders of magnitude greater than any found in human history; associated land slips and tidal waves; a dust blackout of sunlight that must have taken many months to clear; a sharp drop in temperatures that would have brought frost to the tropics; changes in atmospheric and water chemistry; and disturbance of existing patterns of atmospheric and oceanic circulation.

It is possible that earthquakes influenced volcanic eruptions. The first flows of the Deccan basalts predate polarity chron 29R and the impact, but the great body of basalt poured out within that brief chron, possibly emitting enough sulfur to add to the crisis.

Different biotic communities were affected to different degrees. The pelagic community, sensitive to photosynthetic productivity, was severely struck, with coccolithophores and planktonic foraminiferans reduced to a few species, while ammonites, belemnites, plesiosaurs, and mosasaurs were eliminated. Yet dinoflagellates, endowed with the capacity to encyst under stress, suffered no great loss. Benthic life was only moderately damaged beyond the loss of species, excepting destruction of the reef community. While North American trees underwent far more extinction at the specific level than formerly believed, land floras escaped with little damage, presumably because they were generally equipped to handle stress by dormancy and seed survival. The plant-fodder-dependent dinosaurs perished, as did their predators and scavengers. The fresh-water community, buffered by ground water against temperature change and food-dependent mainly on terrestrial detritus, was little affected. While a great many individual organisms must have been killed by the immediate effects of the impact, the loss of species and higher taxa must have occurred on land and in shallowest waters mainly in the aftermath of darkness, chill, and starvation, and in the deeper waters in response to changed regimes in currents, temperatures, and nutrition.

The Cretaceous crash led above all to an evolutionary outburst of the mammals, which in the succeeding tens of millions of years not only filled and multiplied the niches left by dinosaurs but also invaded the seas, to become successors to plesiosaurs and mosasaurs. Humans, among others, owe their existence to that devastating event of 65 million years ago.

Alfred G. Fischer

Bibliography. W. Alvarez et al. (eds.), *Geochronology, Time Scales and Global Stratigraphic Correlation*, SEPM Spec. Publ., no. 54, 1995; R. H. Dott, Jr., and D. R. Prothero, *Evolution of the Earth*, 1992; L. A. Frakes, J. E. Francis, and J. I. Syktus, *Climatic Modes of the Phanerozoic*, 1992; J. L. Powell, *Night Comes to the Cretaceous*, 1998.

Criminalistics

Criminalistics is the science and profession dealing with the recognition, collection, identification, individualization, and interpretation of physical evidence, and the application of the natural sciences to law-science matters. The term originated from the book *Handbuch für Untersuchungsrichter als System der Kriminalistik* (3d ed., 1898) by Hans Gross, an investigating magistrate and professor of criminology at the University of Prague. He described the need for a scientifically trained investigator who could undertake certain technical aspects of an investigation and could also serve as liaison between scientific specialists who might assist in the investigation of criminal activity. This concept was popular in Europe, where a number of forensic science institutes were developed to apply the tools and techniques of the natural sciences to the investigation

of crime and, generally, in official governmental inquiries.

In the United States and the United Kingdom, with legal systems fundamentally different from those of Europe, the criminalistics profession has developed in a different fashion. Criminalistics laboratories were established in a few police departments in the early decades of the twentieth century (in Chicago, Los Angeles, and New York) and in the Federal Bureau of Investigation (FBI). In the 1970s an infusion of money from the federal government resulted in the formation of numerous laboratories associated with state and local law enforcement agencies.

The majority of criminalists work in laboratories associated with governmental agencies charged with enforcing local, state, or federal laws or regulations. However, a number of criminalists engage in private practice or are employed by academic institutions.

Education. Criminalists require a broad education in the natural sciences, including inorganic, organic, and analytical chemistry; physics; mathematics and statistics; and biology and biochemistry. In addition, curricula in criminalistics or forensic science leading to bachelor's or master's degrees include courses dealing with specific types of commonly encountered evidence materials (such as dried biological fluid stains, impression evidence, or trace evidence), with specific analytical procedures (such as microscopy or instrumental analysis), and with applicable legal issues.

As a consequence of the small number of graduates with degrees in criminalistics or forensic science,

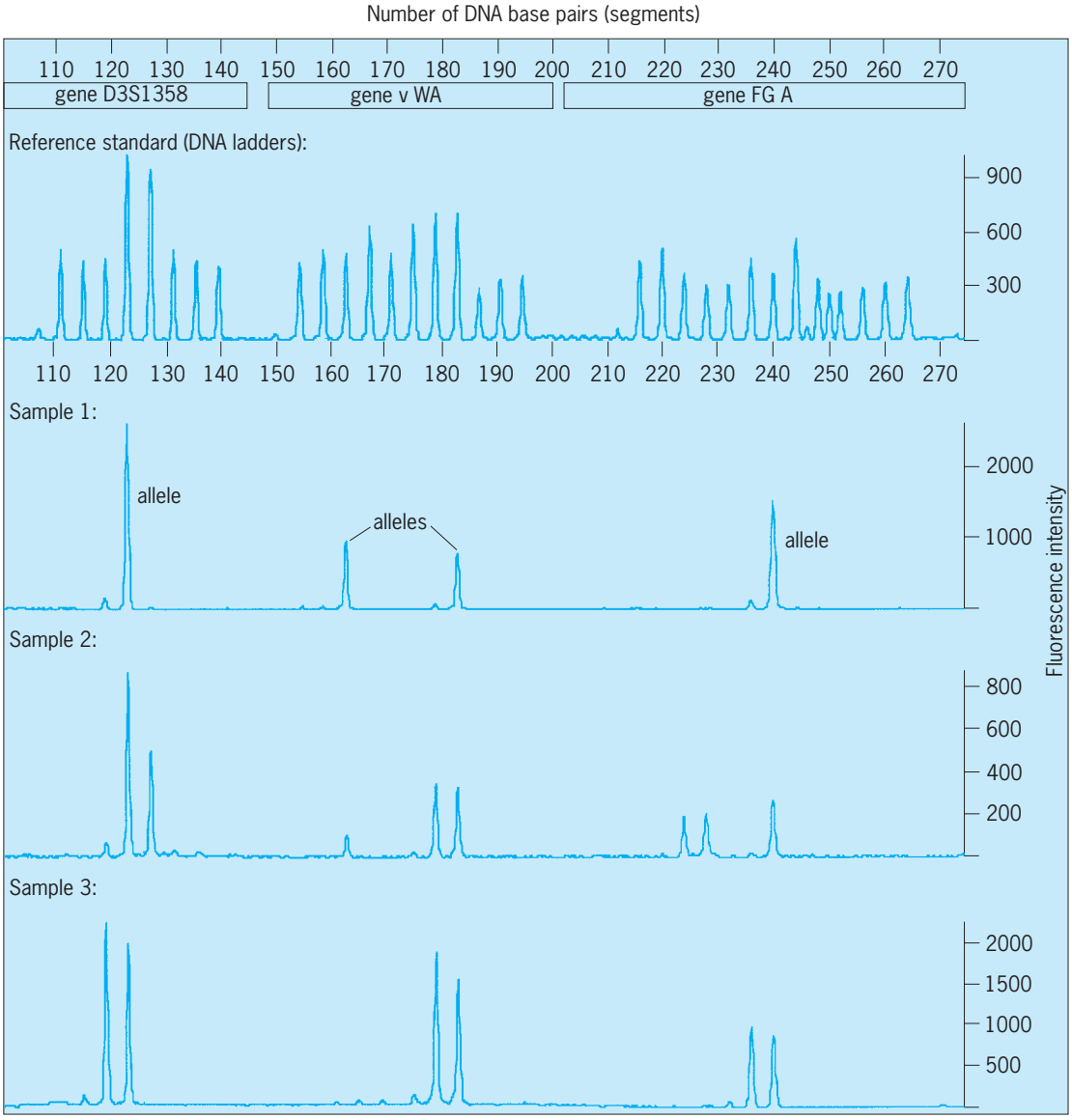

Fig. 1. Capillary electropherogram of DNA isolated from biological evidence. DNA, isolated from biological material such as blood, semen, hair, or tissue, can be compared with DNA obtained from suspects or victims of crimes. The polymerase chain reaction (PCR) is used to increase minute amounts of DNA isolated from evidenciary specimens. Analysis by capillary electrophoresis, or other techniques, allows the criminalist to determine, with a very high degree of certainty, that the evidenciary sample is or is not from a particular individual. (*Forensic Science Associates*)

many criminalists have undergraduate or graduate degrees in scientific disciplines such as chemistry, biochemistry, genetics, or biology, and have chosen a career in criminalistics after completing their education. A wide variety of in-service training is available, including training through the FBI, the California Department of Justice (through the California Criminalistics Institute), and various professional associations.

Specialists. It is not possible for a single person to become proficient in the examination and analysis of all types of physical evidence. Increasingly, criminalists and other workers in forensic science laboratories are specializing in the examination of one or a few types of physical evidence.

Forensic biologists. Forensic biology is the analysis of the biological or genetic properties of evidence. Criminalists specializing in forensic biology are involved in the identification of biological evidence and attempts to determine its source. Traditional methods for the analysis of blood group antigens and genetically variant proteins present in blood, semen, and other biological materials have given way to the analysis of DNA (deoxyribonucleic acid) which can be recovered from such materials. (**Fig. 1**). The ability to identify the individual from whom a blood or semen sample has come has revolutionized the science of criminalistics. *See* DEOXYRIBONUCLEIC ACID (DNA); FORENSIC MEDICINE.

Trace evidence analysts. These specialists analyze material that is transferred between two objects which come into contact (**Fig. 2**). The Locard exchange principle, credited to the French criminalist Edmond Locard, states that whenever two objects come into contact, portions (or traces) of one object will be transferred to the other. Finding these traces, identifying what they are, and determining their origin through the process of individualization (determination of the parameters of a sample which will allow it to be distinguished from other or all similar items) is the job of the trace evidence analyst. This analysis is based on the chemical or physical properties of the material. Use of the optical microscope and the electron microscope and highly sensitive methods of chemical analysis (such as Fourier transform infrared spectroscopy, energy dispersive x-ray analysis, mass spectroscopy, and neutron activation analysis) are typically used by trace evidence specialists. *See* MASS SPECTROSCOPE; SCANNING ELECTRON MICROSCOPE; SPECTROSCOPY; X-RAY FLUORESCENCE ANALYSIS.

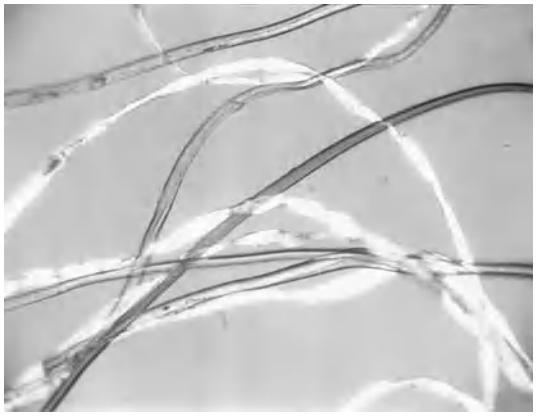

Fig. 2. Fiber photomicrographs. Trace evidence such as fibers, hairs, soil, glass, paint, or botanical material may be transferred between two objects which come into contact. A variety of analytical techniques, including polarized light microscopy, fluorescence microscopy, infrared spectroscopy, scanning electron microscopy, energy-dispersive x-ray spectroscopy, and microspectrophotometry are used for the analysis of such materials. The finding of material from one object on another object provides evidence that the two objects were in contact. (*Forensic Science Associates*)

Firearms and toolmark examiners. These specialists examine firearms, ammunition components, and tools and marks left by them. The underlying principle is that when a tool acts on some object it will leave a mark which is unique due to the configuration of the cutting edge. The uniqueness of each tool is a result of manufacturing processes and postmanufacturing wear and damage. The tool may be a screwdriver which was used to pry open a door or window, or a gun barrel which produced distinctive markings on a fired bullet. The firearms and toolmark specialist will compare the marks on a recovered bullet or a toolmark on a window with guns or tools recovered during the investigation. It is often possible to conclude that a recovered bullet was fired from a particular firearm or that a mark on a window was made by a particular screwdriver.

Other evidence specialists. At a crime scene, physical evidence is routinely encountered (for example, shoe or tire impressions, gunpowder residues on the body or clothing of victims or perpetrators in shootings, fragments of bombs or destructive devices recovered from scenes of sabotage, or pieces of botanical material). There are people who specialize in some of these areas, but often a criminalist is called upon to develop a method for examination of unique evidence, or to consult with a scientist from an industrial or academic laboratory with expertise in an unusual field (such as forensic anthropology).

Criminalists are often involved in the analysis of suspected illegal drugs and narcotics, the examination of questioned documents, or fingerprint identification. Although they are familiar with the techniques used for the examination of these types of evidence, the examinations are usually performed by specialists. *See* FINGERPRINT; FORENSIC CHEMISTRY; FORENSIC TOXICOLOGY.

Analytical techniques. A wide variety of techniques are used by criminalists for the location and collection of evidence at crime scenes as well as for the examination and analysis of that evidence in the laboratory. Crime scene techniques may involve the use of lasers or other light sources to locate biological stains or minute fibers or paint particles, chemical tests for lead around suspected bullet holes, electrostatic devices to recover a dusty shoe sole impression from a floor, or special reagents for the development of latent fingerprints.

Many techniques used in the forensic laboratory are the same ones that are used by analytical

chemists, molecular biologists, materials scientists, and so on. Often these techniques are adapted to the special requirements of the forensic science laboratory. Infrared spectroscopy, mass spectrometry, gas chromatography, optical and electron microscopy, and a host of other standard analytical chemistry techniques find common use by criminalists. *See* GAS CHROMATOGRAPHY.

Routine techniques and procedures have been developed by forensic scientists which have little or no application outside the forensic laboratory. Examples are the determination of genetic markers in minute fragments of dried biological material, the determination of the refractive index of microscopic glass fragments, the microscopic comparison of individual human hairs, and the microscopic comparison of markings on the surface of bullets. Peter D. Barnett

Bibliography. P. R. DeForest et al., *Forensic Science: An Introduction to Criminalistics*, McGraw-Hill, 1983; C. Meloan, *Criminalistics*, Prentice Hall, 1997; A. Moenssens et al., *Scientific Evidence in Civil and Criminal Cases*, 4th ed., Foundation Press, New York, 1995; R. Saferstein, *Criminalistics: An Introduction to Forensic Science*, 6th ed., Prentice Hall, 1997; R. Saferstein, *Forensic Science Handbook*, vols. 1–3, Prentice Hall, 1982; H. Tuthill, *Individualization: Principles and Procedures in Criminalistics*, Mass Market Paperback, 1994; J. Zonderman, *Beyond the Crime Lab: The New Science of Investigation*, Wiley, New York, 1998.

Crinoidea

A class of exclusively suspension-feeding echinoderms with long, slender arms arranged radially around the calyx, a rigid cuplike structure composed of calcareous plates. The radial arm arrangement gives crinoids a flowerlike appearance (**Fig. 1**). Two basic adult body types are recognized: the sea lilies, with a long, anchored stem vertically supporting the calyx and arms above the sea bottom; and the stemless featherstars, or comatulids, with a whorl of flexible appendages on the calyx. Crinoids have a worldwide distribution and can be found in all seas except the Black and Baltic. They occupy depths ranging from just below sea level to a depth of over 9000 m (29,500 ft). Sea lilies are found only at depths greater than 100 m (330 ft), whereas comatulids are most abundant and diverse in shallow, tropical coral reef environments. *See* BLASTOIDEA; PELMATOZOA.

Over 500 species of living comatulids have been described, and many of these can be extremely abundant locally, yet fewer than 100 species of living sea lilies are known. These patterns contrast starkly with the fossil record of crinoids, which is dominated by stemmed crinoids: of the more than 5000 described fossil species, more than 90% have stems.

Living crinoids have no economic importance and are not used for food by humans. However, their fossil remains are often the dominant constituent of building limestone (for example, Indiana limestone) which is highly valued.

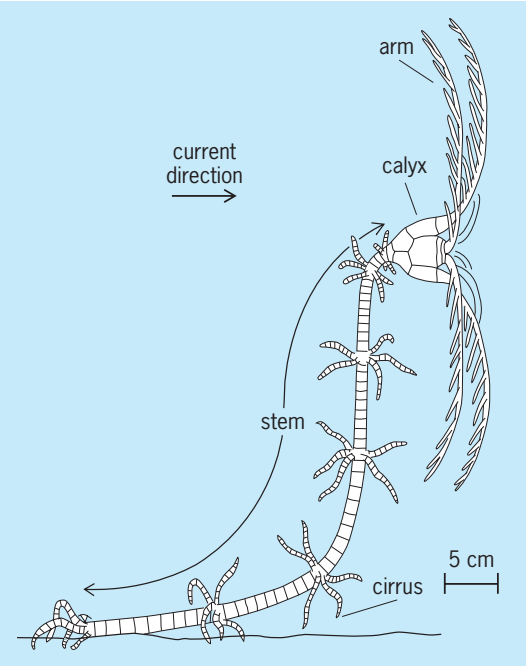

Fig. 1. Stemmed crinoid in feeding posture.

Morphology. Adult crinoids range in size from a few centimeters for some of the stemless forms to several meters from base of the stem to tip of the arms for the largest stemmed crinoids. They also vary in color from the largely bland whites and grays of the deep-water forms to the brilliant reds, yellows, and purples of the shallow-water tropical featherstars. In spite of the size range, color differences, and stemmed or stemless condition as adults, all crinoids share a suite of morphological traits that can be traced back to the Ordovician age, nearly 500 million years ago. *See* ORDOVICIAN.

As in other echinoderms, the crinoid skeleton is composed of many calcareous plates, often numbering in the thousands. Each plate is an intricate meshwork of interconnected microscopic rods and bars of calcium carbonate. The skeletal plates are joined to each other by fibers of connective tissues that penetrate and attach to the array of rods and bars. Plates can be bound to adjacent plates by ligaments, muscles, or both, and the strength and flexibility of each articulation is determined largely by the type and arrangement of the soft tissues.

The calyx consists of tightly sutured plates arranged in circlets of five with arm plates articulated to the uppermost circlet (**Fig. 2**). A flat to dome-like structure, called the tegmen, forms a lid over the cuplike calyx. The tegmen may be armored with calcareous plates or, as in most living crinoids, may consist only of a leathery skin.

The arms are constructed of one or two series of plates; in arms consisting of two series of plates, adjacent plates interdigitate with each other. Each of the five arms may remain undivided, or may branch one or numerous times producing as many as 200 branches. Arm plates may bear short, lateral

branches, called pinnules, composed of serially arranged plates (**Fig. 3**).

As juveniles, all crinoids possess a stem. However, comatulids lose the stem early in ontogeny, such that in adults the lowest plate of the calyx bears only cirri, a set of long, slender articulated appendages with a clawlike terminal element used for grasping the substrate. In adult sea lilies, the stem is retained. One end of the stem articulates to the base of the calyx, and the opposite end anchors the animal to the bottom. The stem is composed of flat, disklike plates, called columnals, that are circular to pentagonal in outline, perforated at the center, and stacked plate on plate. Ligament fibers connect adjacent columnals. The anchoring end of the stem may consist of a rootlike structure that penetrates soft bottom sediments or that cements the stem to a hard substrate, or, most commonly, a set of grasping cirri. In extant sea lilies with cirri, called isocrinids, these appendages are arranged in whorls of five that are regularly spaced along the entire length of the stem. *See* ONTOGENY.

Crinoid viscera are contained in the coelom that forms the body cavity in the calyx and also extends into the arms, pinnules, stem, and cirri. The digestive system consists of a U-shaped gut with its terminal ends, the mouth and anus, penetrating the tegmen. A water vascular system, consisting of an array of fluid-filled canals in the calyx and all appendages, performs many functions associated with feeding and respiration. Its slender movable podia-like structures, called tube feet, extend along the arms and pinnules; they are sites of respiratory gas exchange and the primary food capturing organs.

Ecology and function. Crinoids are exclusively passive suspension feeders, extracting food particles from the ambient water. Food capture occurs when a microscopic particle carried by the current strikes and adheres to a mucus-coated tube foot extended from the pinnule (**Fig. 4**). Subsequently, the tube foot deposits the particle into the food groove that lines the oral side of pinnules and arms. In the food groove, other tube feet and ciliary currents transport the particle toward the mouth.

Crinoids are indiscriminate feeders, and the tube feet capture organic and inorganic particles with a median size of about 50 micrometers and rarely larger than 500 μm. The organic food component consists primarily of phytoplankton, protozoa, and crustacea.

Crinoid morphology and behavior strongly reflect their total reliance on water movement for nutrient supply. Crinoids avoid slack-water environments, living in areas dominated by currents, wave action, or multidirectional flows. Comatulids typically occupy exposed sites on reef crests, and in deep water they are often perched on corals, sponges, and sea lily stems, away from the lower velocities characterizing the boundary layer. However, some live semicryptically within the reef infrastructure. Sea lilies are found in areas dominated by currents, even in ocean depths.

Among comatulids, tube-foot morphology and spacing along pinnules varies in relation to flow velocities: species with longer, more widely spaced tube feet live in environments with lower velocities, and those with shorter, more closely spaced tube feet occupy exposed sites with higher velocities. Crinoid feeding postures vary depending on flow velocities and flow patterns. When exposed to unidirectional currents, crinoid arms, pinnules, and tube feet are aligned into a nearly planar array oriented perpendicular to flow. To avoid having the ambient current directly impinge on food grooves in which particles are transported toward the mouth, crinoids arrange all their arms with the food groove, oral side facing downstream and the opposite, aboral side facing upstream.

Under oscillating water movements produced by wave action, arms are arranged in a bidirectional fan with half the arms facing one direction and the others facing in the opposite direction. The downstream orientation of the pinnular food groove is often maintained by the weathervanelike swiveling of pinnules with their extended tube feet.

Semicryptic comatulids living within the reef infrastructure experience multidirectional flows.

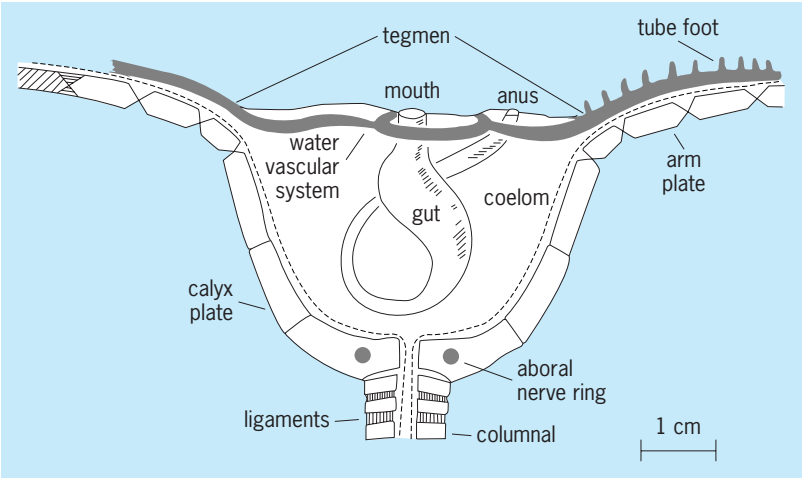

Fig. 2. Calyx and proximal stem and arm morphology, vertical section.

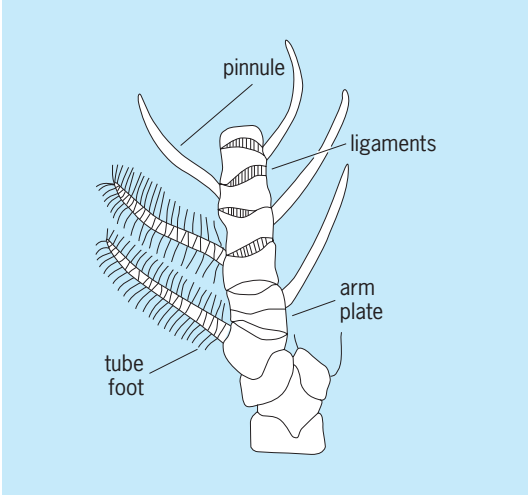

Fig. 3. Arm morphology, back (aboral) view.

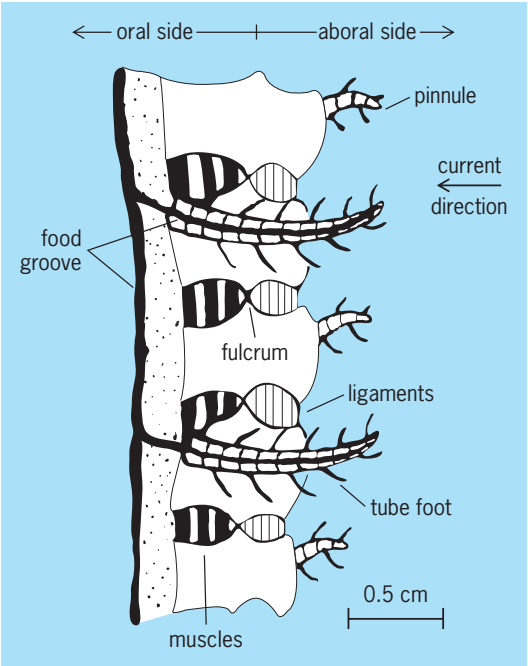

Fig. 4. Arm morphology, side view.

These crinoids do not form planar fans with their arms, but extend their arms into three-dimensional bushlike structures. Pinnules are held in two or three planes along each arm, and swiveling of pinnules occurs in response to changes in current direction.

The coordination and flexibility required for achieving the appropriate feeding posture is attained with the complex nervous system and the ligamentary and muscular arm articulations. Arm movements are generated by muscles and viscoelastic ligaments acting antagonistically across a fulcrum that lies between two adjoining arm plates. In comatulids, arm movement can be highly coordinated and rapid, allowing crawling and, in some, short bursts of swimming. Crawling also has been observed among the dominant group of living stemmed crinoids, the isocrinids, which can move slowly along the bottom by pulling with their arms and dragging their stem behind.

Some comatulids are nocturnal and use their crawling abilities to move between reef crevices and exposed sites on the reef that they occupy at night. In general, crawling allows crinoids to move to better feeding microhabitats and to avoid disturbance.

Crinoid ligaments can undergo rapid and reversible changes in stiffness, a property that is important to crinoid function and behavior. In their compliant state, ligaments allow movement of body parts with a minimum amount of muscular force. In their stiff state, ligaments can withstand high external forces, such as those due to current drag, with little deformation, allowing crinoids to maintain stable feeding postures. Through the rapid loss of stiffness and strength, ligaments can irreversibly disintegrate, allowing the animal to shed body parts. The shedding of body parts, called autotomy, can occur in response to external physicochemical stimuli, such as rapid temperature changes or predatory attacks.

Several interactions between crinoids and other animals are known to occur in the modern seas. The most common are parasitic and commensal associations between polychaete worms (Myzostomida), small snails (Eulimidae, Styliferidae), and crustaceans (Ascothoracida) that infest crinoids. These parasites can feed from the ambulacral grooves or the tegmen and can induce malformations such as cysts. Clingfishes (Gobiesocidae) are also known to feed on comatulid pinnules and on worms and copepods associated with crinoids.

Geologic history. Crinoids have a long and rich fossil record, and at times in Earth history they were numerically one of the dominant members of the benthic marine ecosystem. The first undisputed crinoids were found in rocks of the Ordovician Period, although an enigmatic fossil, *Echmatocrinus*, of Middle Cambrian age has been described as a crinoid. Crinoid diversity waxed and waned throughout the roughly 250 million years of the Paleozoic. The Mississippian Period, known as the Age of Crinoids, represents the peak in their abundance and diversity. Thick, extensive deposits consisting almost exclusively of crinoid remains, called encrinites or criquinites, are especially common in the Mississippian rock record, when crinoid diversity reached its peak of over 100 genera. Crinoids remained an important component of marine communities until the Permo-Triassic extinction event that signaled the end of the Paleozoic Era. This event led to a great reduction in crinoid diversity. A single genus, *Holocrinus*, found in Lower Triassic rocks, is the most primitive member of the Articulata, a subclass that includes all post-Paleozoic crinoids. *See* CAMBRIAN; MISSISSIPPIAN; PALEOZOIC; TRIASSIC.

Crinoids rediversified rapidly, and by the Late Triassic their range of morphologies and ecologies was as high as at any time in the Paleozoic, although the number of Mesozoic or Cenozoic genera never approached that of the Mississippian. The Mesozoic also represents an interval during which several lineages lost the stem. Some of these lineages became planktonic or pseudoplanktonic, while the comatulids, which first appeared in the latest Triassic, achieved a great level of vagility but remained in the benthos. Overall, levels of vagility not known in the Paleozoic characterize post-Paleozoic crinoids, a consequence of the highly muscularized arms and stem-shedding abilities of the latter group.

After the Mesozoic, crinoids became a much less common and less diverse part of the fossil record. Cenozoic rocks contain only bottom-dwelling crinoids, with the stemless comatulids and the cirri-bearing isocrinids most common. Whereas stemmed crinoids had been an important element of shallow-water environments throughout the Paleozoic and Mesozoic, they became restricted to bathyal and abyssal depths during the Cenozoic, which is still true today. The timing of the displacement of stemmed crinoids into deeper waters during the Cenozoic was coincident with the diversification of shell-crushing fishes. Today significant fish predation

on crinoids does not occur. The disappearance of stemmed crinoids from shallow water environments and the success there of the stemless comatulids that are highly mobile, often semicryptic, and possess certain other antipredatory features, argue for fish predation as an important ecological and evolutionary agent. *See* ARTICULATA (ECHINODERMATA); CAMERATA; CENOZOIC; CRETACEOUS; ECHINODERMATA; FLEXIBILIA; INADUNATA; JURASSIC; MESOZOIC; REGENERATION (BIOLOGY). Tomasz K. Baumiller

Bibliography. M. Jangoux and J. M. Lawrence (eds.), *Echinoderm Nutrition*, A. A. Balkema, Rotterdam, 1982; R. C. Moore and C. Teichert (eds.), *Echinodermata 2: Crinoidea*, pt. T of *Treatise on Invertebrate Paleontology*, Geological Society of America and University of Kansas Press, Lawrence; C. R. C. Paul and A. B. Smith (eds.), *Echinoderm Phylogeny and Evolutionary Biology*, Clarendon Press, London, 1988; J. A. Waters and C. G. Maples, Geobiology of Echinoderms, *Paleontol. Soc. Pap.*, vol. 3, Paleontological Society, Pittsburgh, 1997.

Critical care medicine

The treatment of acute, life-threatening disorders, usually in intensive care units. Critical care medicine has been practiced informally for many decades in trauma centers, postanesthesia recovery rooms, coronary care units, delivery rooms, emergency rooms, and postoperative areas. The facilities and trained personnel available in the intensive care unit (ICU) permit extensive monitoring of physiological variables, organization of complex, multidisciplinary diagnostic and therapeutic plans, administration of therapy to predetermined goals, and expert nursing care.

Critical care thus runs counter to the traditional division of specialities by organ or organ system. Specialists in critical care undergo training beyond a primary qualification (internal medicine, surgery, anesthesia, or pediatrics), and must be able to manage acute respiratory, cardiovascular, metabolic, cerebral, and renal problems, as well as infections. The patients may be newborns, children, or adults suffering from trauma or acute life-threatening disease. Patients having failure of multiple organs, complicated medical problems, disorders falling into several medical specialities, or a need for 24-h care often become the responsibility of the critical care specialist.

The intensive care unit is the most labor-intensive, technically complex, and expensive part of hospital care. Of the average daily cost of a bed, about 70% is for personnel. The intensive care unit, however, may be crucial to the patient's survival.

Historical background. Although there were probably precedents, it was during the Crimean War in 1856 that Florence Nightingale placed the most serious casualties nearest the nursing station for the sake of better and more efficient care. Operational efficiency became important in the deployment of scarce medical resources and personnel, particularly in combat. In the civilian area, a special room was used for the postoperative care of neurosurgical patients at Johns Hopkins Hospital in 1923. In 1930 a postanesthesia recovery room was installed at Tübingen, Germany, as well as a unit that applied new knowledge and techniques to postoperative needs. Casualty facilities and trauma hospitals were developed in Europe around World War II. Shock wards were established in field hospitals to resuscitate battle casualties before and after their operations; the importance of blood transfusions, early operation, and postoperative care were recognized. Subsequently, trauma units, coronary care units, and postoperative care were recognized. Trauma units, coronary care units, and neonatal intensive care units were established in the United States to promote expediency and efficiency. In community hospitals the need for intensive care units was mainly the result of the shortage of nurses during and after World War II. It became essential to group patients according to the severity of their illness.

A historical example illustrates these developments. In the poliomyelitis epidemic of 1952 in Copenhagen, an estimated 5000 persons were afflicted; about 10% of these had respiratory paralysis from bulbar polio. Tubes placed through surgically prepared openings in the trachea and manual ventilation were instituted for patients with respiratory paralysis. To assist patients who could not breathe, medical teaching and research were suspended and the students worked day and night in shifts to ventilate patients by hand. This extraordinary effort reduced the mortality rate from 85 to 40%. This episode stimulated clinical research toward the development of positive-pressure mechanical ventilation. An early model of the Engstrom respirator proved to be reliable; through subsequent developments the iron lung, with its negative-pressure whole-body ventilation, was superseded. Mechanical ventilation was used for patients having crushed chest injuries and was employed during thoracic surgery and abdominal surgery. Respiratory and multidisciplinary intensive care units were rapidly developed for patients whose respiration was inadequate for any of various reasons.

Internists became involved in critical care when it was discovered that apparent death from heart attack could be treated by electric shock: dysrhythmias (abnormal heart rhythms) were found to be the most common cause of death from myocardial infarction in the first 12 h after the onset of symptoms. Resuscitation was found to be possible only if a well-equipped, trained team capable of inserting the tracheal tube and performing defibrillation arrived within minutes of cardiac arrest. This finding gave impetus to the development of paramedic transport of patients (prehospital emergency care) and in-hospital coronary care units.

In the late 1950s, modern emergency resuscitation developed. It was based on the observation that the expired breath of the rescuer could be used to breathe for the patient with ventilatory arrest. Mouth-to-mouth breathing and external cardiac massage led to the technique of cardiopulmonary

resuscitation (CPR) and subsequently to protocols for basic life support and advanced life support.

The introduction of cardiac catheterization, its application to the diagnosis and management of congenital and acquired heart disease, and its later application to cardiovascular monitoring have been basic to managing circulatory problems in patients in the intensive care unit. The advent of plastic catheters facilitated such procedures as intravascular blood sampling, infusion of fluids and titration of drugs, and circulatory measurements. In the early 1970s, the flow-directed, balloon-tip pulmonary artery catheter was introduced, which in combination with a thermodilution system permits estimation of cardiac output.

Although renal failure had been managed with peritoneal dialysis beginning in 1923, hemodialysis (dialysis of blood through an external membranous coil) came into wide clinical use only after World War II. Today, in addition to being used in outpatient and other inpatient settings, hemodialysis is employed in intensive care units, for example, to treat patients who have multiple organ failure.

Many advances in medicine have increased the need for critical care. Among them are heart operations, transplantation of various organs, and cancer chemotherapy.

Life-support systems. Perhaps the single most useful function of the intensive care unit is to provide life-support systems for desperately ill patients who would not survive without them. Such systems are briefly discussed below.

Mechanical ventilation. A mechanical ventilator automatically delivers oxygen-enriched air at positive pressure through an endotracheal tube. The machine may be driven by pressure (pressure-regulated) or volume (volume-regulated). Adjustments may be made to control rate of ventilation, maximum airway pressure, peak flow, ratio of time of inspiration to that of expiration, and oxygen concentration. Ventilator modes determine how much effort the patient makes and how much assistance the ventilator provides; the ventilator can entirely control ventilation. Different clinical conditions require somewhat different ventilator modes.

Cardiopulmonary resuscitation. Resuscitation for cardiopulmonary arrest must be begun at the site and time of arrest whether in the intensive care unit, elsewhere in the hospital, or in the community. Arrest may occur more frequently in the intensive care unit than elsewhere because the sickest patients are usually there. Out in the field, without equipment, cardiopulmonary resuscitation consists of (a) airway control (positioning of head, inflation of lungs, manual clearing of mouth and throat), (b) breathing support [mouth-to-mouth (or nose) ventilation], and (c) circulation support (control of external hemorrhage, mechanical chest compressions if the patient is pulseless, and circulatory support with intravenous fluids if available). When equipment is available, cardiopulmonary resuscitation consists of the above plus (a) pharyngeal suctioning, tracheal intubation, or, if this is not immediately successful, creation of an opening into the trachea in the neck; (b) ventilation manually (using a bag and mask) or mechanically, with oxygen added; and (c) mechanical chest compressions, or if blood pressure and pulse do not respond immediately, open-chest massage with direct cardiac compressions.

Advanced life support consists of the use of cardiac drugs, fluids, electrocardiographic monitoring for cardiac dysrhythmias, defibrillation by electrical shock with external direct current of 200–400 joules, and management of subsequent complications and multiple organ support in an intensive care unit.

Peritoneal dialysis and hemodialysis. In the intensive care unit, peritoneal dialysis and hemodialysis are used to help treat such challenging conditions as renal failure, multiple organ failure, drug overdose, sepsis (widespread infection), posttrauma conditions, and postoperative renal failure. Blood is withdrawn through intravascular catheters and pumped through a dialysis chamber and back to the patient's vascular system. Among items removed by dialysis are the waste products, urea and creatinine; excess water, potassium, and sodium; drugs; and organic acids and other noxious metabolic end-products.

Circulatory support with intraaortic balloon pumping. The intraaortic balloon pump is the temporary device that is most widely used to assist the failing heart. It is employed when the ventricle of the heart cannot pump enough blood to meet the minimum needs of the body. The balloon, which is placed in the descending aorta, is rapidly inflated in diastole (the time of cardiac relaxation) and rapidly deflated at the onset of systole (the time of cardiac contraction). When the pump is properly synchronized with the heart, the balloon deflation produces a vacuumlike effect that helps the left ventricle to contract. The device also augments diastolic pressure, which improves blood flow and oxygen supply to the heart. Combining the mechanical assistance with the ventricle's own limited ability can permit adequate circulation.

Extracorporeal membrane oxygenation. Extracorporeal membrane oxygenation equipment is a modification of the apparatus that is routinely used to oxygenate the blood outside the body during open heart surgery. It is used mainly in near-term infants who do not respond to maximum ventilatory and medical support.

Life-sustaining therapy. Another aspect of critical care is the provision of life-sustaining therapy. Components include administration of intravenous fluids, provision of nutritional support, and control of infections.

Fluid and electrolyte administration. Fluids are given mainly to replace fluid losses and to provide circulatory support in patients with conditions such as shock or circulatory deficiencies. Intravenous and intraarterial catheters are placed to allow the monitoring of arterial and central venous pressures and the rapid administration of fluids, such as blood, blood components, and various solutions.

Drugs often are used in conjunction with fluid therapy. Such drugs include agents that stimulate cardiac contraction, agents that increase blood

pressure, and agents that reduce vascular resistance to cardiac output.

Nutritional support. In life-threatening crises, nutritional requirements must be maintained. A patient who cannot take nourishment by mouth receives it either via a tube introduced through the nose into the stomach or duodenum, or parenterally (through a vein). Glucose, fat, amino acids, and other substances are given. Total parenteral nutrition can sustain life indefinitely in patients who are unable to eat and have higher than normal nutritional requirements because of sepsis, injury, or prior malnutrition.

Control of infection. A large percentage of patients in intensive care units have infections that are part of their primary disease or were acquired in the hospital. The latter infections are particularly difficult to eradicate because they are usually resistant to most antibiotics. Infectious agents are identified by culturing bronchial secretions and the fluids from body cavities, wounds, and elsewhere. The sensitivity of each bacterial agent to an array of antibiotics must be tested.

The control of the environment in the intensive care unit is a major problem. Cross contamination between various patients and between patients and personnel must be constantly monitored to prevent outbreaks of infection. *See* HOSPITAL INFECTIONS.

Other acute crises. Various types of crises are managed by critical care specialists. The following are some examples. Upper gastrointestinal hemorrhage may occur because of bleeding duodenal ulcers, esophagitis, gastritis, perforations of the stomach or esophagus, or other conditions. Lower gastrointestinal bleeding may arise from conditions such as colon tumors, colon diverticula, and hemorrhoids. Therapy usually consists of rapid blood replacement and then, once the patient's condition is stable, operative control of the bleeding source. *See* HEMORRHAGE.

Acute abdominal crises. Acute catastrophic abdominal conditions can include acute pancreatitis, intestinal obstruction, perforation of a duodenal ulcer, stones of the kidney or ureter, and many other conditions. Many patients with these conditions are admitted to the intensive care unit during their preoperative diagnostic work-up and for stabilization of their cardiorespiratory and fluid and electrolyte status; they may be readmitted postoperatively for management.

Cardiac emergencies. Among the cardiac emergencies sometimes treated in intensive care units are serious abnormalities of heart rate and rhythm, toxicity from drugs used to treat heart disease, and cardiac tamponade (compression of the heart by blood or fluid in the sac surrounding it). For some conditions treatment can include emergency placement of a pacemaker.

Coma management. The sudden onset of altered mental status or coma may result from head injury, cerebrovascular accidents (stroke), drug overdose, postoperative conditions after neurosurgery, meningitis, brain tumors, and various other causes. Patients with altered mental status or coma require special nursing to prevent complications. A major consideration, especially after neurosurgical operations and head injuries, is to monitor the level of consciousness; a change in consciousness is usually the first warning of increasing intracranial pressure, which if not promptly corrected may lead to irreversible brain damage.

Acute endocrine disorders. Critical care is used in treating a number of acute endocrine catastrophies that are uncommon but potentially fatal. Examples include severe conditions caused by excessive or deficient thyroid hormone secretion, hypoparathyroidism after thyroid surgery, and adrenocortical insufficiency (Addison's disease). *See* ENDOCRINE MECHANISMS.

Injuries. Trauma from automobile accidents, penetrating injuries from violent crimes, and blunt trauma from falls or assaults pose serious technical and logistic problems that must be rapidly addressed. The injured patient must promptly receive appropriate diagnostic tests, monitoring, and therapy. Sometimes surgery must be undertaken while resuscitation efforts are still under way.

Hepatic failure and metabolic problems. Critical care also is used in managing hepatic failure and metabolic problems. Hepatic insufficiency or failure can result from conditions such as alcoholic cirrhosis, viral hepatitis, drug reactions, and poisoning. Often infection is present, as immunocompetence is impaired in liver failure. *See* CIRRHOSIS; LIVER.

Pulmonary problems. Patients whose breathing is seriously impaired by conditions such as pneumonia, asthma, pulmonary embolism (blood clot in the lung), smoke inhalation, and near-drowning may receive critical care. Measures available to assist such patients include use of mechanical ventilators (discussed above), administration of humidified supplemental oxygen through a mask or nasal prongs, and use of bronchodilator drugs. William C. Shoemaker

Bibliography. J. B. Hall et al., *Principles of Critical Care*, 2d ed., 1997; W. C. Shoemaker et al. (eds.), *Textbook of Critical Care*, 2d ed., 1989.

Critical mass

The amount of fissile material (uranium-233, uranium-235, or plutonium-239) that supports a self-sustaining nuclear chain reaction. It ranges from as little as 1 lb (0.45 kg) for plutonium-239 dissolved in light water to 2 tons (1800 kg) for uranium-235 fueling a 1-GWe (1 gigawatt of electric power) nuclear power reactor. *See* CHAIN REACTION (PHYSICS).

Critical mass is a definitive feature of nuclear reactors and nuclear explosives. It is increased by the presence of such neutron-absorptive materials as admixed uranium-238, aluminum pipes for flow of coolant, zirconium fuel-element cladding, and boron, cadmium, or gadolinium control rods. It is reduced by a moderator, such as graphite, heavy water, or light water, which slows down the neutrons, inhibits their escape, and thus improves the probability of producing fission in the fuel. Without a

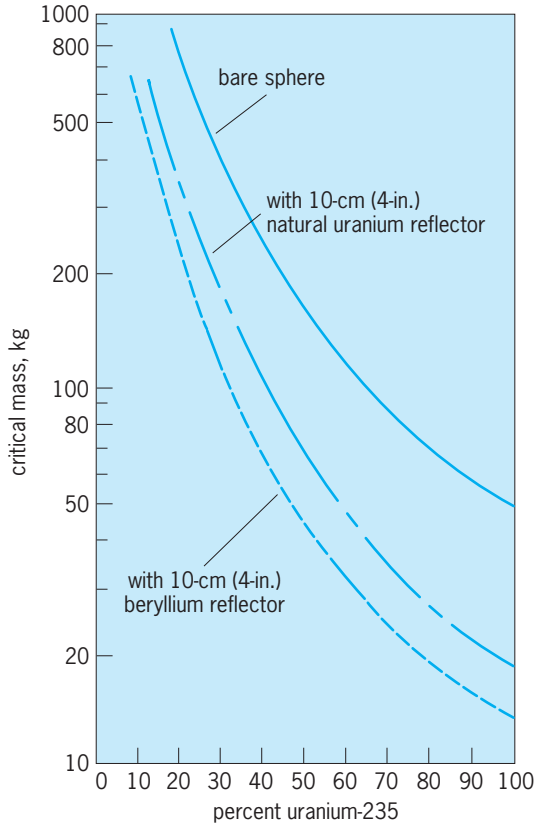

Critical mass of a uranium sphere as a function of the uranium-235 fraction for three reflector configurations. 1 kg = 2.2 lb.

moderator, natural uranium cannot reach criticality; about 1 ton (900 kg) is used to fuel a heavy-water-moderated 1-GWe reactor. Most power reactors, especially fast reactors, are loaded with core-distributed fissile inventories that add up to many latent critical masses. Additional fissile material is also bred into the core fissile material through neutron absorption in admixed fertile material, such as in the conversion of uranium-238 to fissile plutonium-239. *See* NUCLEAR REACTOR.

Critical mass also depends on the density of the fissile component, its geometry, and the immediate surroundings. The **illustration** shows its relationship to the fissile fraction for a metallic sphere enriched in uranium-235. It also shows how a surrounding layer of suitable material (called a reflector), such as beryllium or natural uranium, can significantly reduce the critical mass. The critical mass varies for different nuclear isotopes (see **table**).

Reactor criticality. In a reactor the critical mass is not a static quantity because specific nuclear phenomena activate and influence the chain reaction. When the reactor is just critical, the reactor fission rate is stable, a condition called delayed critical. When the overall fission rate is declining or increasing, subcritical or supercritical conditions respectively exist. Subcriticality ensues if, under prevailing conditions, the fuel loading is insufficient to sustain criticality. Supercriticality (or prompt criticality) occurs at the instant that the reactive fissionable content exceeds the requirements for nuclear equilibrium.

Delayed criticality is sustained in all reactors by adjusting operating parameters that regulate the effective critical mass. Heat production for power generation, which consumes fissionable material, can be maintained by compensating with the removal of neutron absorbers through the movement of control rods, changes in the external geometry provided by reflectors, or continuous replacement of fissionable material when liquid fuel is employed. Burnable poisons and the deliberate conversion of fertile isotopes are also used to overcome reactivity losses from power-producing fuel consumption. Reactivity loss occurs also from the production of neutron-absorbing fission products, through the power-production cycle.

The criticality state of a reactor is described mathematically by the multiplication constant k, which is a ratio comparing the number of neutrons born to those consumed or lost. When the reactor is just critical, the effective value of k is unity. In calculating the value of k, all neutron sources and sinks are taken into account. The sources are (1) intrinsic yields of fission neutrons in the fissile isotope and (2) augmented fission in the fertile component. The sinks are (1) nonfission captures in the fuel, cladding, coolant, fission products, and structure and (2) leakage from the reactor.

By exploiting inherent nuclear and thermodynamic properties of core and reflector materials, reactors are designed to be almost self-regulating. The effective critical mass and state of criticality at any given moment are determined by delayed neutrons, expansion coefficients, void fractions, and resonance widths for neutron capture reactions in the reactor. Reactors are normally operated at delayed critical, with redundant features to avoid prompt criticality. Safety and power generation are dependent upon proper reactor management and conformance to operating standards that are based on these inherent design attributes. *See* DELAYED NEUTRON; REACTOR PHYSICS.

Nuclear explosions. An explosive, supercritical condition can be deliberately created by rapid introduction of reactivity in excess of the requirements for a critical mass. This is the basis for nuclear explosions, which rely on sudden explosively driven densification of a subcritical fissile component.

Approximate critical masses and sizes of fissionable spheres surrounded by 10-cm (4-in.) uranium reflectors

Isotopes	Density, g/cm^3	Critical mass, kg (lb)	Radius, cm (in.)
^{233}U	18.9	5.7 (12.6)	4.2 (1.7)
^{235}U	18.9	15.7 (34.6)	5.8 (2.3)
^{239}Pu	19.5	4.5 (9.9)	3.8 (1.5)
^{240}Pu	19.5	19.6 (43.2)	6.2 (2.4)
^{242}Pu	19.5	73 (161)	8.4 (3.3)
20% ^{235}U, 80% ^{238}U	18.9	350 (770)	9.7 (3.8)
30% ^{239}Pu + ^{241}Pu, 70% ^{240}Pu + ^{242}Pu	19.5	6 (13)	4.2 (1.7)

Tritium boosting with thermonuclear reactions is used to amplify the explosive yield of a fission chain reaction, thus reducing the fissile mass inventory of a weapon that must be compressed in order to become supercritical. *See* ATOMIC BOMB; HYDROGEN BOMB; NUCLEAR EXPLOSION.

Safety issues. Because critical mass is dependent on extrinsic factors such as geometry, moderator, and coolant flow, reactor safety issues are of extreme importance. Reactors can become prompt critical but, lacking physical mechanisms for intense compression, they are unable to explode like nuclear weapons. Early in the development of the first atomic bomb (the Manhattan Project), a few workers lost their lives in test-reactor-core prompt-criticality accidents. In 1986, at the Chernobyl reactor in the former Soviet Union, fatal consequences occurred after operators lost control of reactor criticality. The Three Mile Island reactor in Pennsylvania suffered a loss-of-coolant accident in 1979, but criticality control was maintained, no lives were lost, and only a small amount of radioactivity escaped the primary containment. *See* NUCLEAR POWER.

When subcritical quantities of nuclear materials are handled, packaged, or transported, safe geometrical configurations and environments are worked out in great detail and retained through careful management. For example, it should not be possible for a nuclear mass to become critical even if its container is flooded with water. No criticality accidents related to fuel transport have been recorded. A. DeVolpi

Bibliography. A. DeVolpi, Denaturing fissile materials, *Prog. Nucl. Energy*, 10:161–219, 1982; F. J. Rahn et al., *A Guide to Nuclear Power Technology*, 1984.

Critical phenomena

The unusual physical properties displayed by substances near their critical points. The study of critical phenomena of different substances is directed toward a common theory.

Critical points. Ideally, if a certain amount of water (H_2O) is sealed inside a transparent cell and heated to a high temperature T, for instance, $T > 647$ K (374°C or 705°F), the enclosed water exists as a transparent homogeneous substance. When the cell is allowed to cool down gradually and reaches a particular temperature, namely the boiling point, the enclosed water will go through a phase transition and separate into liquid and vapor phases. The liquid phase, being more dense, will settle into the bottom half of the cell. This sequence of events takes place for water at most moderate densities. However, if the enclosed water is at a density close to 322.2 kg · m^{-3}, rather extraordinary phenomena will be observed. As the cell is cooled toward 647 K (374°C or 705°F), the originally transparent water will become increasingly turbid and milky, indicating that visible light is being strongly scattered. Upon slight additional cooling, the turbidity disappears and two clear phases, water and vapor, are found. This phenomenon is called the critical opalescence, and the water sample is said to have gone through the critical phase transition. The density, temperature, and pressure at which this transition happens determine the critical point and are called respectively the critical density ρ_c, the critical temperature T_c, and the critical pressure P_c. For water $\rho_c = 322.2$ kg · m^{-3}, $T_c = 647$ K (374°C or 705°F), and $P_c = 2.21 \times 10^7$ pascals. *See* OPALESCENCE.

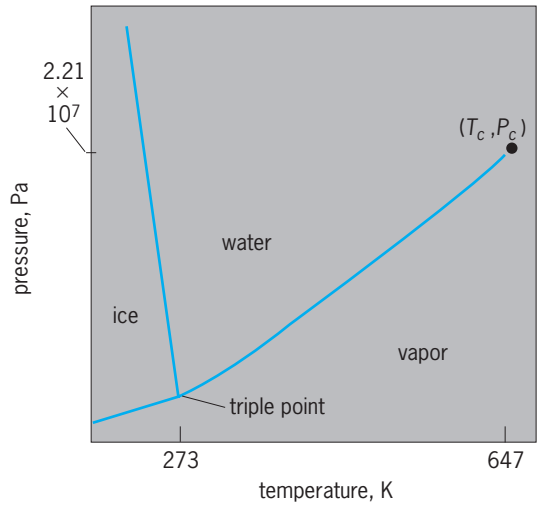

Fig. 1. Phase diagram of water (H_2O) on the pressure-temperature (*P-T*) plane.

Different fluids, as expected, have different critical points. Although the critical point is the end point of the vapor pressure curve on the pressure-temperature (*P-T*) plane (**Fig. 1**), the critical phase transition is qualitatively different from that of the ordinary boiling phenomenon that happens along the vapor pressure curve. In addition to the critical opalescence, there are other highly unusual phenomena that are manifested near the critical point; for example, both the isothermal compressibility and heat capacity diverge to infinity as the fluid approaches T_c. *See* THERMODYNAMIC PROCESSES.

Many other systems, for example, ferromagnetic materials such as iron and nickel, also have critical points. The ferromagnetic critical point is also known as the Curie point. As in the case of fluids, a number of unusual phenomena take place near the critical point of ferromagnets, including singular heat capacity and divergent magnetic susceptibility. The study of critical phenomena is directed toward describing the various anomalous and interesting types of behavior near the critical points of these diverse and different systems with a single common theory. *See* CURIE TEMPERATURE; FERROMAGNETISM.

Order parameters. One common feature of all critical phase transitions is the existence of a quantity called the order parameter. The net magnetization M is the order parameter for the ferromagnetic system. At temperatures T above T_c and under no external field, there is no net magnetization. However, as the temperature of the system is cooled slowly through T_c, a net magnetization M will appear precipitously. The increase in M is more gradual as temperature is reduced further. The nonzero magnetization is due

TABLE 1. Order parameters, theoretical models, and classification according to universality class of various physical systems*

Universality class	Theoretical model	Physical system	Order parameter
$d = 2$ $n = 1$	Ising model in two dimensions	Adsorbed films	Surface density
$n = 2$	XY model in two dimensions	Helium-4 films	Amplitude of superfluid phase
$n = 3$	Heisenberg model in two dimensions		Magnetization
$d > 2$ $n = \infty$	Spherical model	None	
$d = 3$ $n = 0$	Self-avoiding random walk	Conformation of long-chain polymers	Density of chain ends
$n = 1$	Ising model in three dimensions	Uniaxial ferromagnet	Magnetization
		Fluid near a critical point	Density difference between phases
		Mixture of liquids near consolute point	Concentration difference
		Alloy near order-disorder transition	Concentration difference
$n = 2$	XY model in three dimensions	Planar ferromagnet	Magnetization
		Helium-4 near superfluid transition	Amplitude of superfluid phase
$n = 3$	Heisenberg model in three dimensions	Isotropic ferromagnet	Magnetization
$d \leq 4$ $n = -2$		None	
$n = 32$	Quantum chromodynamics	Quarks bound in protons, neutrons, etc.	

*From K. Wilson, Problems in physics with many scales of length, *Sci. Amer.,* 241(2):158–179, August 1979.

to the partial alignment of the individual spins in the ferromagnetic substance. M is called the order parameter since the state with partial alignment of spins is more ordered than that with no alignment. The density difference in the liquid and vapor phases, $(\rho_l - \rho_v)/\rho_c$, is the proper order parameter in the fluid system; ρ_c in the denominator is the critical density. This order parameter has temperature dependence similar to that for the net magnetization M of a ferromagnetic system. For $T > T_c$ the order parameter is equal to zero, since there is only one homogeneous phase in the fluid. As the system is cooled through T_c, the fluid system phase separates with a precipitous increase in the difference of the liquid and vapor densities. A number of critical systems and their respective order parameters are listed in **Table 1**.

The order parameters assume power law behaviors at temperatures just below T_c. In the fluid and ferromagnetic systems, for $t < 0$, they follow Eqs. (1) and (2), and for $t > 0$ they obey Eqs. (3).

$$M = B(-t)^{\beta} \tag{1}$$

$$\frac{\rho_l - \rho_v}{\rho_c} = B(-t)^{\beta} \tag{2}$$

$$M = 0 \qquad \frac{\rho_l - \rho_v}{\rho_c} = 0 \tag{3}$$

In these equations, $t \equiv (T - T_c)/T_c$ reduced temperature, and β and B are respectively the critical exponent and amplitude for order parameter.

Measurement of the order parameter of fluid neon near the critical temperature [$T_c = 44.48$ K (-228.57°C or -379.43°F), $\rho_c = 484$ kg · m^{-3}, $P_c = 2.72 \times 10^6$ Pa] is shown in **Fig. 2**. The densities of the liquid (upper branch) and vapor (lower branch) phases as a function of temperature are deduced by measuring the dielectric constant of neon in the bottom and top halves of a sample cell. Data at reduced temperatures t between -4×10^{-4} and -2×10^{-3} are shown as broken lines. Careful data analysis shows that the simple power function given in Eq. (2) is not adequate to describe the data over the entire temperature range. This is the case because even in the range of reduced temperatures t between -4×10^{-4} and -4×10^{-2} the fluid is not yet inside the asymptotic or true critical region, and correction terms are needed.

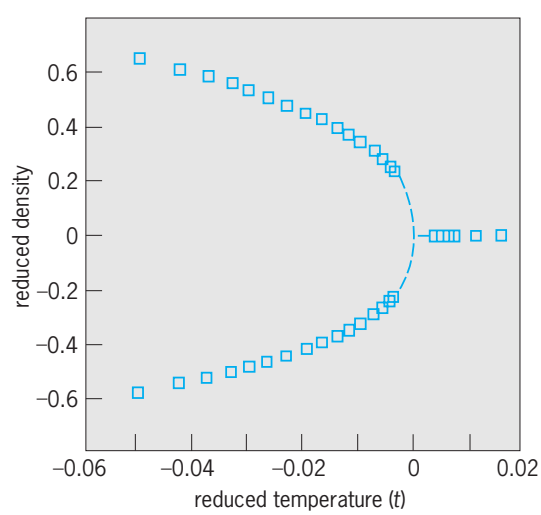

Fig. 2. Densities of liquid (upper branch) and vapor (lower branch) neon near its critical temperature. Reduced density is defined as $(\rho - \rho_c)/\rho_c$ and reduced temperature is defined as $t = (T - T_c)/T_c$. (*After M. Pestak and M. H. W. Chan, Equation of state of N_2 and Ne near their critical points: Scaling, correction to scaling, and amplitude ratios, Phys. Rev., B30:274-288, 1984*)

TABLE 2. Critical exponents and power laws in pure fluids and ferromagnets*

Thermodynamic quantity	Fluid	Ferromagnet	Power law
Specific heat	C_v	C_H	$\sim(t)^{-\alpha}$ for $t > 0$ $\sim(-t)^{-\alpha'}$ for $t < 0$
Order parameter = S	$(\rho_l \pm \rho_v)\rho_c^{-1}$	M	= 0 for $t > 0$
Response function	$K_t = \frac{1}{\rho}\left(\frac{\partial\rho}{\partial P}\right)_T$	$X = \left(\frac{\partial M}{\partial H}\right)_T$	$\sim(-t)^{\beta}$ for $t < 0$ $\sim(t)^{-\lambda}$ for $t > 0$ $\sim(-t)^{-\lambda'}$ for $t < 0$
Critical isotherm	$P \pm P_c$	H	$\sim S\|^{\delta} \cdot$ sign of S
Correlation length	ξ	ξ	$\sim(t)^{-\upsilon}$ for $t > 0$ $\sim(-t)^{-\upsilon'}$ for $t < 0$
Critical correlation function of fluctuation	$G(r)$	$G(r)$	$\sim(r)^{-(d-2+\eta)}$

*d = spatial dimensionality of the critical system; H = magnetic field strength; C_v = specific heat at constant volume; C_H = specific heat at constant magnetic field strength; K_t = isothermal compressibility; X = susceptibility.

When the density data are analyzed according to the predicted functional form given by Eq. (4)

$$\frac{\rho_l - \rho_v}{\rho_c} = B(-t)^{\beta}\,[1 + B_1\,|t|^{0.5} + B_2\,|t| + \cdots] \quad (4)$$

the exponent is found to be 0.327 ± 0.002, in excellent agreement with the theoretically predicted value of 0.325 as discussed below. The value of β found for other simple fluids also converges toward the theoretical value.

Power law behavior of other quantities. The anomalous behavior of other thermodynamic quantities near the critical point can also be expressed by power laws. These can be characterized by a set of critical exponents, labeled (besides β) α, γ, δ, ν, and η: α characterizes the divergent heat capacity; γ the susceptibility (of magnets) and isothermal compressibility (of fluids); δ the critical isotherm; ν the correlation length; and η the critical correlation function. The functional forms of these power laws for the fluid and magnet systems are shown in **Table 2**. The critical opalescence phenomenon discussed above is closely related to the strong density fluctuations induced by the divergent isothermal compressibility near T_c. When the density fluctuations are correlated at lengths comparable to the wavelength of visible light, intense scattering occurs.

Mean field theories. The earliest attempts to understand the critical behavior were the van der Waals model for fluids (1873) and the Weiss model for ferromagnets (1907). These are mean field theories in the sense that the state of any particular particle in the system is assumed to be determined by the average properties of the system as a whole. In these models, all particles can be considered to contribute equally to the potential at each site. Therefore, the mean field theory essentially assumes the intermolecular interaction to be of infinite range at all temperatures. The mean field theories are qualitatively quite successful in that they predict the existence of critical points and power law dependence of the various thermodynamic quantities near the critical point. They are not quantitatively correct because the predicted values for the various exponents are not in agreement with exact model calculations or with experimental results (**Table 3**).

Scaling hypothesis. Theoretical efforts in the study of critical phenomena have been centered on predicting correctly the value of these critical exponents. One of the most important developments is the hypothesis of scaling. This hypothesis is model-independent and applicable to all critical systems. The underlying assumption is that the long-range correlation of the order parameter, such as the spin fluctuation in the ferromagnetic system and the density fluctuation in the fluid system near T_c, is responsible for all singular behavior. This assumption leads to a particular functional form for the equation of state near the critical point. With this simple assumption, it has been demonstrated that a number of inequalities among critical exponents that can be proved rigorously by thermodynamic arguments are changed into equalities. These equalities, or scaling laws, show that there are only two independent critical exponents; once two exponents in a system are given, all other exponents can be determined. What is truly impressive about this simple hypothesis is that the scaling laws, Eqs. (5)–(9), have been shown

$$\alpha = \alpha', \quad \gamma = \gamma', \quad \nu = \nu' \quad (5)$$

$$2 = \alpha + \gamma + 2\beta \quad (6)$$

$$2 = \alpha + 2\beta\delta - \gamma \quad (7)$$

$$\nu d = 2 - \alpha \quad (8)$$

$$\gamma = \nu(2 - \eta) \quad (9)$$

to be correct in almost all real and theoretical model critical systems.

The meaning of these exponents is explained above and also in Table 2; d is not an exponent but the spatial dimensionality; the primed and unprimed exponents represent the value below and

TABLE 3. Values of critical exponents

Systems	α	β	γ	δ	ν	η
Mean field model	0 (discontinuity)	$^1/_2$	1	3	$^1/_2$	0
Ising model (exact) ($d = 2$, $n = 1$)	0 (logarithmic discontinuity)	$^1/_8$	$^7/_4$	15	1	0.25
Ising model (approx.) ($d = 3$, $n = 1$)	0.110	0.325	1.24	4.82	0.63	0.03
Heisenberg model (approx.) ($d = 3$, $n = 3$)	−0.10	0.36	1.38	4.80	0.705	0.03
Fluids*						
Xe	0.08 ± .02	0.325 − 0.337	1.23 ± 0.03	4.40		
SF_6		0.321 − 0.339	1.25 ± 0.03			
Ne		0.327 ± 0.002	1.25 ± 0.01			
N_2		0.327 ± 0.002	1.233 ± 0.010			
Ferromagnets* (isotropic)						
Iron, Fe	−0.09 ± .01	0.34 ± .02	1.33 ± 0.02			0.07 ± 0.07
Nickel, Ni	−0.09 ± .03	0.37 ± .03	1.34 ± 0.02	4.2 ± 0.1		

*Experimental data are the averaged measured values of a number of experiments.

above T_c, respectively. A large number of theoretical and experimental activities are concerned with putting the scaling hypothesis on a firm fundamental basis and understanding its universality and limitations.

Model systems. A great deal of insight has been gained by the construction and solution of model systems that can be solved exactly. The most famous one is the two-dimensional ($d = 2$) Ising model solved by L. Onsager in 1944. In this model, spins (little magnets) on a lattice are allowed to point in either the up or down directions, and it is assumed that only pairs of nearest-neighboring spins can interact. Onsager found a critical point for this system and calculated the values of the various critical exponents. The solution of this model is important because this is one of the very few model systems with exact solutions, and it is often used to check the validity of approximation techniques. There are many other model systems similar to the Ising model. They are distinguished from each other by the spatial (d) and spin (n) dimensionality. The spin can be oriented along an axis ($n = 1$, Ising model), or in any direction on a plane ($n = 2$, XY model), or in any direction in space ($n = 3$, Heisenberg model). *See* ISING MODEL.

These models are essentially simplified versions of real physical systems. The three-dimensional Heisenberg model ($d = 3$, $n = 3$), for example, clearly resembles the isotropic ferromagnets; the three-dimensional Ising model ($d = 3$, $n = 1$) can be found to correspond to the pure fluid system. The correspondence can be shown if the space accessible to the fluid is divided into lattice sites, and at each site the spin parameter is considered to be up if the site is occupied and down if it is not. There are other physical systems beside pure fluids that resemble the three-dimensional Ising model, for example, the binary fluid near its consolute mixing point, the uniaxial ferromagnet, and an alloy system near the order-disorder transition. A great deal of effort and ingenious mathematical techniques have been employed to obtain approximate solutions to these model systems.

Universality hypothesis. It has been observed that the measured values for the critical exponents are rather close to the calculated ones of the corresponding model system. This observation leads to the hypothesis of critical universality. According to this hypothesis, the details of the particle-particle interaction potential in the vicinity of the critical point are not important, and the critical behavior is determined entirely by the spatial dimensionality d and the spin dimensionality n. All systems, both model and real, that have the same value of d and n are said to be in the same universality class and have the same critical exponents. The hypotheses of scaling and universality are closely related: since the length of correlation between density or spin fluctuation diverges as one approaches the critical point, and any interaction potential between particles is finite in range, the details of interparticle potential are expected to be increasingly less important as one approaches the critical point. It has been shown that scaling laws can be derived from the universality hypothesis. Classification of model and physical systems according to universality classes is shown in Table 1.

Renormalization group. The renormalization group (RG) method, originally used in quantum field theory, has been introduced in the study of critical phenomena. By employing a set of symmetry transformations, the ideas contained in the principle of universality and in the scaling hypothesis can be reformulated and incorporated much more economically. As a result, a fully operational formalism is obtained from which critical exponents can be calculated explicitly. Beside the success in critical phenomena, the renormalization group method is also found to be a very useful technique in many diverse areas of theoretical physics. *See* QUANTUM FIELD THEORY.

The validity and indeed the elegance of the concept of universality and the theory of critical phenomena are borne out by experiments. As stated above, the order parameter exponents for simple fluids were found in a number of experiments to be equal to 0.327 ± 0.002. This is in excellent

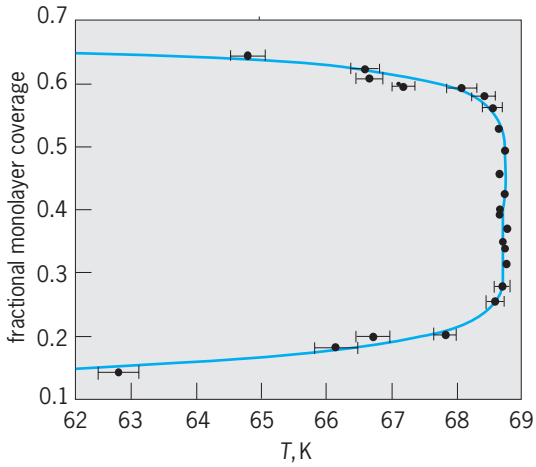

Fig. 3. Surface densities (values of fractional monolayer coverage) of monolayer methane adsorbed on graphite near its two-dimensional critical temperature. (*After H. K. Kim and M. H. W. Chan, An experimental determination of a two-dimensional liquid-vapor critical point exponent, Phys. Rev. Lett., 53:170–173, 1984*)

agreement with the value of 0.325 calculated for the three-dimensional Ising model ($n = 1$, $d = 3$) by renormalization-group and high-temperature series expansion methods. **Figure 3** shows the measured densities of the liquid and vapor phases of a two-dimensional fluid. The two-dimensional fluid is achieved by adsorbing methane molecules on an atomically flat graphite surface at monolayer and submonolayer coverages. The curve in Fig. 3, which traces the boundary separating the liquid-vapor coexistence region from the hypercritical fluid region, also reflects the liquid (upper half) and vapor (lower half) densities. It is qualitatively different from that for a bulk or three-dimensional fluid as shown in Fig. 2. When the curve is subjected to an analysis according to the power law given in Eq. (2), the resulting exponent β is found to be 0.127 ± 0.020, in excellent agreement with the value of $^1/_8$ as predicted by Onsager for the two-dimensional Ising model ($d = 2$, $n = 1$).

Dynamical effects. Besides the static critical phenomena, there are many interesting dynamical effects near the critical point, including critical slowing down, the dynamics of density and spin fluctuations, thermal and mass transport, and propagation and attenuation of sound. Understanding of these effects is far from complete.

Most critical phenomena experiments have been performed on three-dimensional systems, but a number have been done in two-dimensional and quasi-two-dimensional systems. The experiments include the order-disorder transition and the continuous melting transition of gases bound to a graphite surface and the superfluid transition of ^{4}He films on substrates. There has also been considerable interest in the influence of disorder on fluid and magnetic systems. These experiments provide interesting physical realizations of the various model systems in two dimensions. *See* PHASE TRANSITIONS; STATISTICAL MECHANICS.

Moses H. W. Chan

Bibliography. J. J. Binney et al., *The Theory of Critical Phenomena*, 1992; C. S. Domb et al. (eds.), *Phase Transitions and Critical Phenomena*, vols. 1–16, 1972–1994; H. E. Stanley, *Introduction to Phase Transition and Critical Phenomena*, paper 1987; D. I. Uzunov, *Introduction to the Theory of Critical Phenomena*, 1993; K. Wilson, Problems in physics with many scales of length, *Sci. Amer.*, 241:(2)158–179, August 1979.

Crocodile

The common name used for 15 species of large reptiles included in the family Crocodylidae, one of the three families of the order Crocodylia, which also includes the alligators, caimans, and gharials (also known as gavials). Like all crocodilians, the crocodiles are distributed throughout the tropical and subtropical regions of the world. Species occur in both saltwater and freshwater habitats. Crocodiles are generally carnivorous, feeding on invertebrates, fish, other reptiles, amphibians, birds, and mammals. A few, very narrow snouted species are believed to subsist primarily on fish. Crocodiles are primarily aquatic and nocturnal, leaving the water only to bask by day or to build their nests. Some species construct burrows into the banks of rivers or lakes where they spend part of their time. *See* ALLIGATOR; CROCODYLIA; GAVIAL; REPTILIA.

Anatomy. These animals are powerful predators with large teeth and strong jaws. Large adults of some species may exceed 20 ft (6 m) in length and are capable of overpowering and eating large grazing mammals such as deer and cattle and even, occasionally, humans. The webbed feet, laterally compressed tail, and placement of the nostrils, eyes, and ears on raised areas of the head are adaptations for an aquatic existence (see **illustration**). The raised nostrils, eyes, and ears allow the animals to float almost completely submerged while still monitoring their environment. The body, particularly the back, is covered with a series of bony plates (osteoderms) which are unfused and have no direct connection with the

American crocodile (*Crocodylus acutus*) from Ding Darling National Wildlife Refuge, Florida. (*Photo by Gerald and Buff Corsi; copyright © 2001 California Academy of Sciences*)

endoskeleton. The crocodile does not molt the epidermis in a single unit or in large pieces, as do most reptiles.

Size. There is considerable controversy over just how large crocodiles may grow. The saltwater crocodile (*Crocodylus porosus*) of the Indo-Pacific and Australia; the Nile crocodile (*C. niloticus*) of Africa; the American crocodile (*C. acutus*; see illustration) of Central and South America, many Caribbean islands, and Florida; and the Orinoco crocodile (*C. intermedius*) of the Orinoco drainage in South America are reliably reported to approach or slightly exceed 20 ft (6 m) in total length. Johnson's crocodile (*C. johnsoni*) of Australia; the Mugger crocodile (*C. palustris*) of India and Sri Lanka; the Siamese crocodile (*C. siamensis*) of Malaysia and Indonesia; the New Guinea and Philippine crocodiles (*C. novaeguineae* and *C. novaeguinea mindorensis*) of Papua New Guinea and the Philippines; Morelet's crocodile (*C. moreletti*) of Mexico, Guatemala, and Belize (British Honduras); the Cuban crocodile (*C. rhombifer*) of Cuba; and the narrow-snouted crocodile (*C. cataphractus*) of Africa have a maximum length of between 9 and 14 ft (3 and 4 m). The false gavial (*Tomistoma schlegeli*) of Malaysia, Borneo, and Sumatra may occasionally reach 16 ft (5 m) in length. The broad-snouted crocodiles (*Osteolaemus tetraspis*) of west-central Africa seldom exceed 6 ft (2 m) in total length.

Reproduction. Reproduction in crocodiles, and in all crocodilians, is elaborate and has been well studied in a few species. Courtship and mating occur in the water. The female digs a hole in the soil for the 10 to 50 eggs, or, in the case of several species, constructs an elevated nest of available soil and vegetation near the water. The decaying vegetation gives off heat that aids in the incubation of the eggs. Nests may be as large as 12 ft (3.5 m) across and 4 ft (1.2 m) high and are often guarded by the female, which may also construct wallows adjacent to the nest where she remains for much or all of the incubation period. Incubation takes approximately 2 months. In some species, at least, the hatchlings croak or grunt from within the nest, and the female opens the nest to assist their emergence. The young may remain together as a pod with the female for a year or more. Hearing and vocal communication are well developed in the crocodiles, and a variety of bellows, snarls, and grunts are utilized in their elaborate social behavior. *See* REPRODUCTIVE SYSTEM.

Conservation and ecology. Many factors have contributed to the overall decline of crocodiles worldwide. Crocodile hides are considered valuable, and the decline of many species may be attributed to hunting and poaching. The loss of habitat through land development for other uses also contributes significantly to the decline of crocodile species. Most countries with native crocodile populations are implementing conservation measures, and international efforts are being made to regulate trade in crocodile products. Papua New Guinea has developed a conservation-farming program for its species which, if successful, could serve as a model for the conservation of crocodiles and all crocodilians throughout the world.

The Siamese crocodile (*C. siamensis*), found in Thailand, is considered to be the most endangered crocodile in the world and is listed as "critically endangered" by the International Union for the Conservation of Nature and Natural Resources (IUCN). In the United States, the American alligator (*Alligator mississippiensis*) and the American crocodile (*C. acutus*; see illustration) are both considered to be in decline. The American alligator is listed as "federally threatened" in certain parts of its range, whereas the American crocodile, found in the coastal contact zone of southern Florida, is listed as "critically endangered." Conservation efforts are hindered by the lack of knowledge about the population biology of crocodiles. Historically, they existed in many areas in large numbers and were evidently of considerable importance in the energetics and nutrient cycling in the habitats where they occurred. The overall decline of some species makes studies of their ecological importance difficult. Howard W. Campbell; W. Ben Cash

Bibliography. P. Brazaitis, *The Identification of Living Crocodilians*, 1974; C. A. W. Guggisberg, *Crocodiles: Their Natural History, Folklore, and Conservation*, 1972; E. Jarvik, *Basic Structure and Evolution of Vertebrates*, 2 vols., 1981; W. T. Neill, *The Last of the Ruling Reptiles: Alligators, Crocodiles, and Their Kin*, 1971; F. Pough et al., *Herpetology*, 3d ed., Prentice-Hall, Upper Saddle River, 2004; G. R. Zug, L. J. Vitt, and J. P. Caldwell, *Herpetology*, Academic Press, San Diego, 2001.

Crocodylia

An order of the class Reptilia (infraclass Archosauria) which is aquatic species including the alligators, caimans, crocodiles, and gavials. The group has a long fossil history from the Late Triassic, and its members are the closest living relatives of the extinct dinosaurs and the birds. The 23 living species are found in tropic areas of Africa, Asia, Austria, and the Americas. One form, the saltwater crocodile (*Crocodylus porosus*), has traversed oceanic barriers from the East Indies as far east as the Fiji Islands. *See* ARCHOSAURIA; REPTILIA.

Morphology. The order is distinguished from other living reptiles in that it has two temporal foramina, an immovable quadrate, a bony secondary palate, no shell, a single median penis in males, socketed teeth, a four-chambered heart, and an oblique septum that completely separates the lung cavities from the peritoneal region.

Certain of these unique features and other salient characteristics of the Crocodylia are intimately associated with their aquatic life. The secondary palate, composed of medial expansions of premaxillary, maxillary, palatine, and pterygoid bones, divides the mouth cavity into two separate regions. The area above the bony palate forms an air passage extending from the external nostrils at the tip of the snout

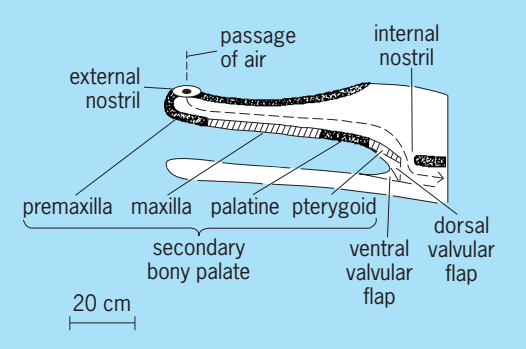

Fig. 1. Sagittal section of crocodylian head to show secondary palate and air passage.

posteriorly to near the orbital region. The lower space retains the usual functions of the mouth and is bordered above by the secondary palate, and below by the lower jaw. A special pair of fleshy flaps are found at the posterior end of the mouth cavity and form a valvular mechanism which separates the mouth from the region where the air passage opens into the throat (**Fig. 1**). This complex arrangement allows crocodylians to breathe even though most of the head is under water, or the mouth is open holding prey or full of water. Inasmuch as these great reptiles are extremely active when submerged and may remain under water for considerable periods, the increased efficiency of respiration provided by the oblique septum is an obvious adaptation. The significance of the four-chambered heart also lies in its contribution to rapid cellular respiration through an increased efficiency of circulation. Interestingly, all three of these structural characteristics are also found in mammals, although differing in detail. The Crocodylia, however, are not closely related to any of the reptiles that ultimately gave rise to the mammal stock.

Other aquatic adaptations include valvular external nostrils that are closed during immersion, a recessed eardrum that can be covered by a skin flap underwater, webbed toes, and a long, compressed, muscular tail that propels the crocodylian through the water with strong lateral thrusts. In addition, the eyes, adapted for nocturnal activity by a vertically elliptical pupil, and the nostrils are mounted in raised areas of the head so that the animal may see and breathe without exposing much of its body above water. The method of locomotion on land is by progression on all fours, with the belly and head held off the ground and the tail dragging behind. In the water, movement is produced by lateral undulations of the tail and the forelimbs are held flat against the sides of the body.

Feeding. Crocodylians are carnivorous and prey upon insects, mollusks, fishes, and mammals. Usually only the larger adults attack terrestrial prey, which is usually caught at the water's edge and after being grasped in the powerful jaws is dragged down to be drowned. The jaws, numerous sharp teeth, and the habit of twisting and rolling underwater after the prey is seized all contribute to the particular reputation crocodilians share. Even on land they are imposing creatures, especially when utilizing the muscular tail for defense. As with any large predator, crocodilians should be treated with respect and caution should be used in approaching individuals in the wild.

Reproduction. During the breeding season male crocodilians set up territories on land which they defend against intruders of the same species. During this period their loud roars are frequently heard at night. Fertilization is internal and the shelled eggs are deposited in excavations in the sand or in large nests of decaying vegetation, depending upon the species. In some forms the female guards the nest and several females may take turns in protecting a communal nest. It has been reported that when the young hatch the female may liberate them from the nest and lead them to the nearest water.

Classification. The living species of crocodilians are placed in two families and eight genera. The family Crocodylidae contains two subgroups: the true crocodiles, Crocodylinae, including the genera *Crocodylus* found in all tropic areas, *Osteolaemus* in central Africa, and the false gavial (*Tomistoma*) in Malaya and the East Indies; the alligators and caimans, Alligatorinae, including the genera

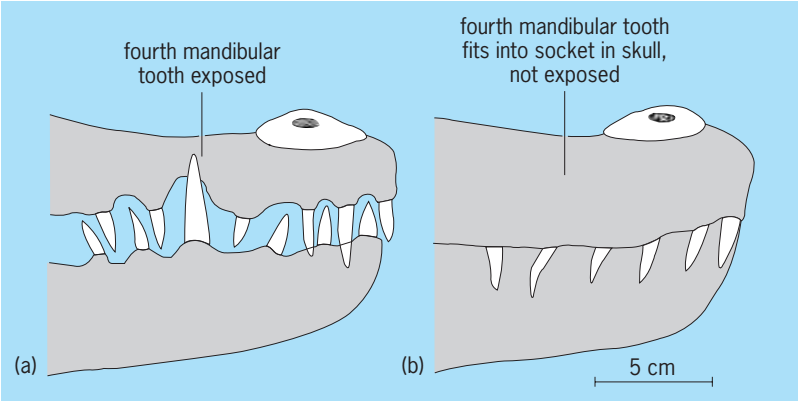

Fig. 2. Comparative lateral view of crocodylian snouts, showing tooth arrangement of (*a*) Gavialidae and Crocodylinae, and (*b*) Alligatorinae.

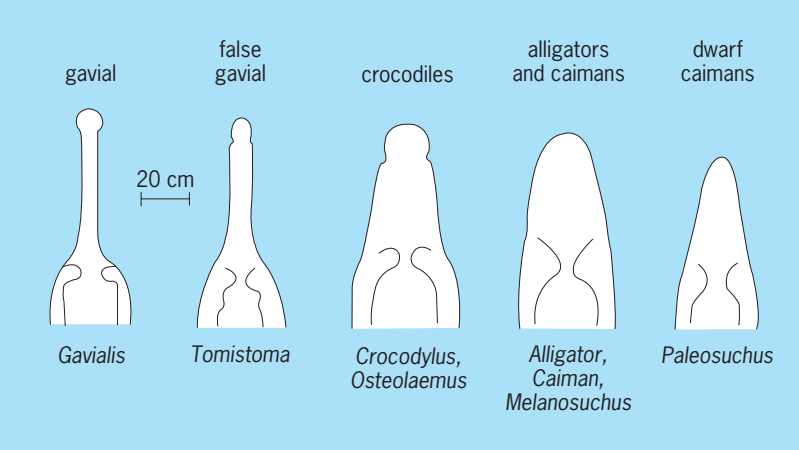

Fig. 3. Dorsal view of crocodylian heads.

Alligator of the southeastern United States and near Shanghai, China, the *Caiman* from Central and South America, and *Melanosuchus* and *Paleosuchus* of South America. The gavial (*Gavialis gangeticus*) of India and north Burma is the only living member of the family Gavialidae. Crocodiles differ most obviously from alligators and caimans in head shape and in the position of the teeth, although other technical details also separate them (**Figs. 2** and **3**). The gavial differs from other living forms in its extremely long and narrow snout, with 27 or more teeth on the upper jaw and 24 or more on the lower. Even the false gavial (*Tomistoma*), which resembles the gavial in head shape, has no more than 21 teeth on the upper jaw and 20 on the lower. *See* ALLIGATOR; CROCODILE; GAVIAL. Jay M. Savage; W. Ben Cash

Bibliography. E. H. Colbert, *The Age of Reptiles*, 1965; E. Jarvik, *Basic Structure and Evolution of Vertebrates*, 2 vols., 1981; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; F. Pough, *Herpetology*, 3d ed., Prentice-Hall, Upper Saddle River, 2004; A. S. Romer, *Vertebrate Paleontology*, 3d ed., 1966; K. P. Schmidt and R. F. Inger, *Living Reptiles of the World*, 1957; S. W. Williston, *Water Reptiles of the Past and Present*, 1914; G. R. Zug, L. J. Vitt, and J. P. Caldwell, *Herpetology*, Academic Press, San Diego, 2001.

Crocoite

A mineral with the chemical composition $PbCrO_4$. Crocoite occurs in yellow to orange or hyacinth red, monoclinic, prismatic crystals with adamantine to vitreous luster (see **illus.**); it is also massive granular. Fracture is conchoidal to uneven. Hardness is 2.5–3 on Mohs scale and specific gravity is 6.0. Streak, or color of the mineral powder, is orangish-yellow. It fuses easily.

Needlelike crocoite crystals in association with cerussite and dundasite, Dundas, Tasmania. (*Specimen courtesy of Department of Geology, Bryn Mawr College*)

Crocoite is a secondary mineral associated with other secondary minerals of lead such as pyromorphite and of zinc such as cerussite. It has been found in mines in California and Colorado. *See* LEAD. Edward C. T. Chao

Crossing-over (genetics)

The process whereby one or more gene alleles present in one chromosome may be exchanged with their alternative alleles on a homologous chromosome to produce a recombinant (crossover) chromosome which contains a combination of the alleles originally present on the two parental chromosomes. Genes which occur on the same chromosome are said to be linked, and together they are said to compose a linkage group. In eukaryotes, crossing-over may occur during both meiosis and mitosis, but the frequency of meiotic crossing-over is much higher. This article is concerned primarily with meiotic crossing-over. *See* ALLELE; CHROMOSOME; GENE; LINKAGE (GENETICS).

Crossing-over is a reciprocal recombination event which involves breakage and exchange between two nonsister chromatids of the four homologous chromatids present at prophase I of meiosis; that is, crossing-over occurs after the replication of chromosomes which has occurred in premeiotic interphase. The result is that half of the meiotic products will be recombinants, and half will have the parental gene combinations (**Fig. 1**). Using maize chromosomes which carried both cytological and genetical markers, H. Creighton and B. McClintock showed in 1931 that genetic crossing-over between linked genes was accompanied by exchange of microscopically visible chromosome markers. *See* RECOMBINATION (GENETICS).

During meiosis, crossing-over occurs at the pachytene stage, when homologous chromosomes are completely paired. At diplotene, when homologs separate, the sites of crossing-over become visible as chiasmata, which hold the two homologs of a bivalent together until segregation at anaphase I. Each metaphase I bivalent will necessarily have at least one chiasma. In favorable material, such as grasshopper spermatocytes, it is possible to observe that each diplotene chiasma involves a crossover of two of the four chromatids at one site.

Where two or more crossovers occur in one bivalent, they usually do not cluster together but are widely separated; this is known as chiasma interference. The occurrence of one crossover event appears to preclude the occurrence of a second crossover in the immediate vicinity. In addition, the distribution of occurrence of chiasmata along a chromosome may be localized; the probability that a crossover will occur is higher in some chromosome segments and lower in other segments.

In general, the closer two genes are on a chromosome, that is, the more closely linked they are, the less likely it is that crossing-over will occur between them. Thus, the frequency of crossing-over between different genes on a chromosome can be used to produce an estimate of their order and distances apart; this is known as a linkage map. *See* GENETIC MAPPING.

Molecular mechanisms. Since each chromatid is composed of a single deoxyribonucleic acid (DNA) duplex, the process of crossing-over involves the

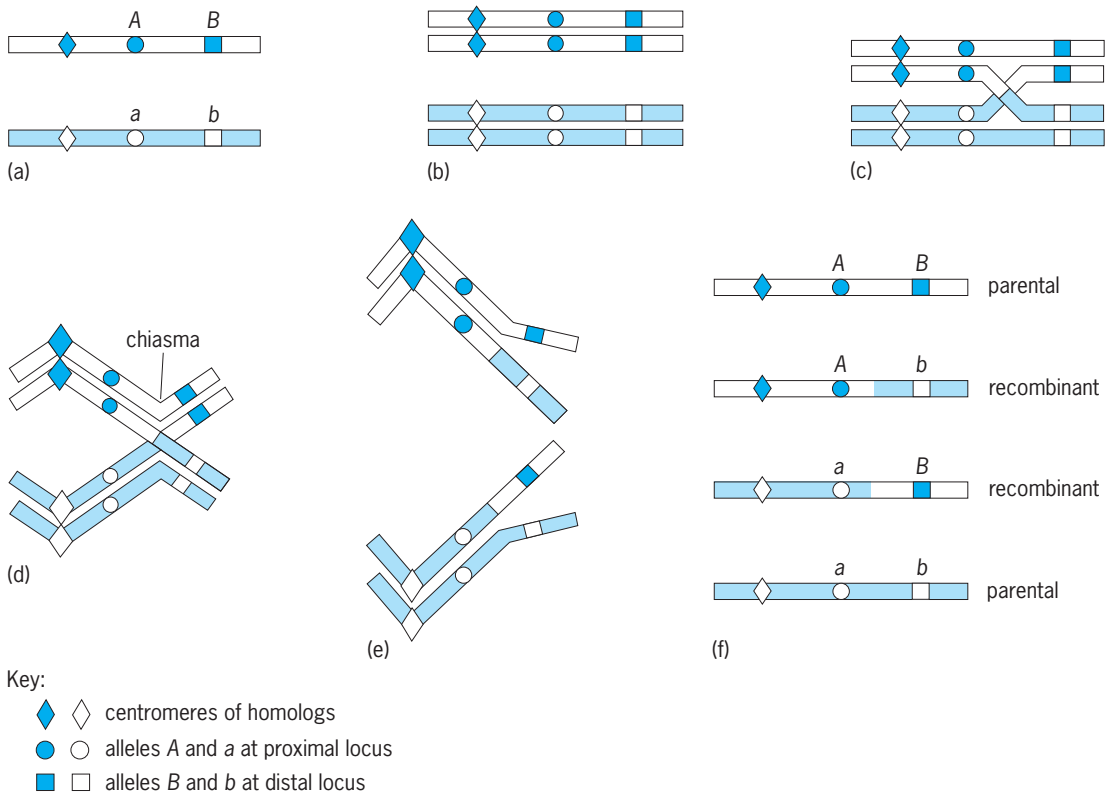

Fig. 1. Meiotic crossing-over between two gene loci *A* and *B* on homologous chromosomes with different alleles at each locus. (*a*) Prior to replication. (*b*) After replication in premeiotic interphase. (*c*) Crossing-over at pachytene between nonsister chromatids of paired homologs (bivalent). (*d*) Diplotene bivalent held together by a chiasma at crossover point. (*e*) Anaphase I segregation of chromosomes in bivalent. (*f*) Four meiotic products after anaphase II segregation. In general, centromeres and loci proximal to the chiasma (crossover) segregate at first division, while loci distal to the chiasma segregate at second division.

breakage and rejoining of DNA molecules. Although the precise molecular mechanisms have not been determined, it is generally agreed that the following events are necessary: (1) breaking (nicking) of one of the two strands of one or both nonsister DNA molecules; (2) heteroduplex (hybrid DNA) formation between single strands from the nonsister DNA molecules; (3) formation of a half chiasma, which is resolved by more single-strand breakages to result in either a reciprocal crossover, a noncrossover, or a nonreciprocal crossover (conversion event).

Two molecular models of recombination which have gained credence are those of R. Holliday and of M. Meselson and C. Radding. Holliday's model postulates nicks in both chromatids at the initiation of crossing-over (**Fig. 2**). Meselson and Radding postulate single-strand cut in only one DNA strand. Repair synthesis displaces this strand, which pairs with its complement on the other chromatid, thereby displacing and breaking the other strand of that DNA molecule. Following pairing and ligation of the two remaining broken ends, a half chiasma is formed. Other models have been postulated in which recombination is initiated by a double-stranded break in one chromatid. In all the above models, gene conversion can occur in the middle region of the molecules (with or without outside marker crossing-over) by mismatch repair of heteroduplex DNA.

Ultrastructural cytology. Pachytene, the meiotic stage at which crossing-over is considered to occur, corresponds with the period of close pairing or synapsis of homologous chromosomes. Electron microscopy has revealed that proteinaceous structures, the synaptonemal complexes (**Fig. 3**), are involved in the synapsis of chromosomes. A synaptonemal complex forms during zygotene by pairing of axial elements from two homologous chromosomes. It is present along the whole length of each pachytene bivalent and disappears at diplotene. Evidence from inhibitor studies and mutant stocks shows that the synaptonemal complex is necessary for meiotic crossing-over to occur. However, in cases such as desynaptic mutants, some hybrids, and the female silkworm, complete pachytene synaptonemal complexes have been observed, but no crossing-over occurs, showing that the synaptonemal complex alone is not sufficient to cause crossing-over.

In *Drosophila melanogaster* oocytes, the occurrence at pachytene of dense spherical bodies bridging the central region of the synaptonemal complex has been described. These bodies coincided in number and position with expected crossover events, and therefore were named recombination nodules. A variety of oval and bar-shaped recombination nodules (Fig. 3) have also been found in organisms as diverse as fungi, humans, rat, silkworm, and maize. In many cases their number correlates with crossover

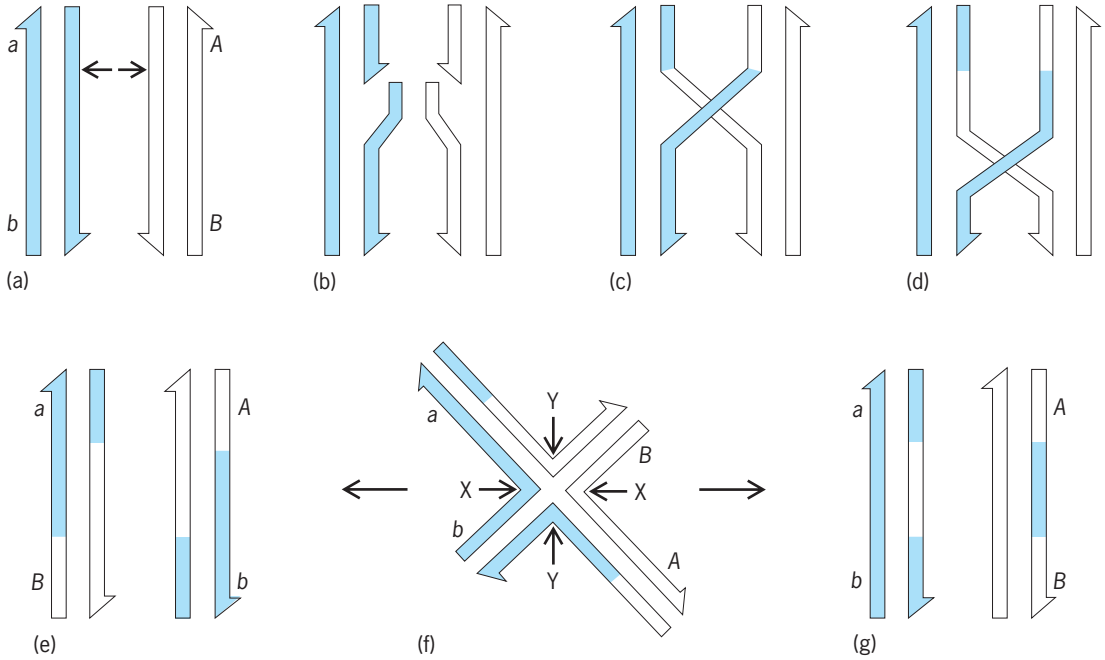

Fig. 2. Molecular model of recombination, based on that of R. Holliday. Only the two recombinant chromatids are shown. (*a*) Two homologous nonsister chromatid DNA molecules, with gene loci *ab* and *AB*, respectively; arrows are paired nicking sites. (*b*) After single-strand nicks at equivalent sites in both chromatids, nicked strands separate. (*c*) Nicked strands are displaced and reanneal to opposite duplex forming a half chiasma. (*d*) Migration of half chiasma increases length of heteroduplex DNA. (*e*) Isomerization of structure in *d* by rotation gives open half-chiasma form; paired nicks occur in two strands. (*f*) Nicks at X,X resolve the half chiasma into reciprocal crossover chromatids *aB* and *Ab*, with heteroduplexes in middle regions; or (*g*) nicks at Y,Y resolve the half chiasma into noncrossover chromatids *ab* and *AB*, with heteroduplexes in middle regions.

frequency. It has been suggested that recombination nodules are prerequisites for crossing-over. If this is so, the recombination nodule may represent a complex of enzymes involved in the early events of recombination (nicking, strand separation, repair synthesis).

DNA repair synthesis has been observed during pachytene in lily microsporocytes, and has been shown to be reduced in an achiasmatic mutant.

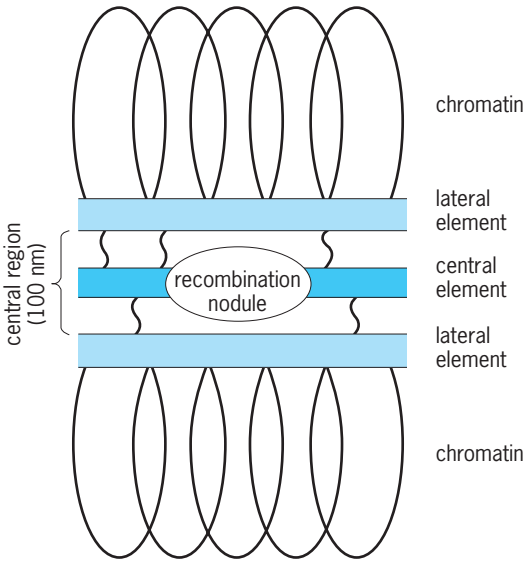

Fig. 3. Longitudinal section of a synaptonemal complex with a recombination nodule.

Prophase I of lilies is characterized by the presence of several proteins which could have a role in crossing-over, for example, DNA binding protein, endonucleases, ligases, and kinase. Inhibition of protein synthesis at zygotene-pachytene results in failure of crossing-over. Thus both DNA synthesis and protein synthesis appear necessary for meiotic crossing-over in lily. *See* MEIOSIS; MITOSIS.

Sex chromosomes. The differentiated X and Y sex chromosomes in human males and many animals (Z and W chromosomes in female birds) have small regions near one tip which undergo pairing and crossing-over at meiotic prophase I. Electron microscopy of the pachytene XY reveals the formation of a short synaptonemal complex segment with a recombination nodule in the majority of cases; the presence of a chiasma between the X and Y at metaphase I indicates the occurrence of crossing-over. An obligatory crossover in the XY bivalent is necessary to ensure regular segregation of X and Y to opposite poles at anaphase I. The pairing region contains a few gene loci on both X and Y chromosomes which exhibit an autosomallike inheritance pattern. Recombination between genes and DNA sequences in this pseudoautosomal region confirms the occurrence of obligatory crossing-over. The rare occurrence of XX males in some cases is accounted for by abnormal recombination events outside the pseudoautosomal region which have transferred the male sex-determining gene from the Y to the X chromosome. *See* SEX DETERMINATION; SEX-LINKED INHERITANCE.

C. B. Gillies

Bibliography. D. G. Catcheside, *The Genetics of Recombination*, 1977; C. B. Gillies (ed.), *Fertility and Chromosome Pairing: Recent Studies in Plants and Animals*, 1989; P. B. Holm and S. W. Rasmussen, Human meiosis, VI. Crossing-over in human spermatocytes, *Carlsberg Res. Commun.*, 48:385–413, 1983; B. Lewin, *Genes*, 3d ed., 1987; P. B. Moens (ed.), *Meiosis*, 1987; F. Rouyer et al., A gradient of sex linkage in the pseudoautosomal region of the human sex chromosomes, *Nature*, 319:291–295, 1986.

Crosstalk

Interference in a communications channel (disturbed channel) caused by activity in other communications channels (disturbing channels). The term was originally used to denote the presence in a telephone receiver of unwanted speech signals from other conversations, but its scope has been extended by common usage to include like effects in other types of communications.

Causes. The cause of crosstalk is some form of coupling mechanism between the disturbed channel and the disturbing channels. Communications channels are normally segregated by space, frequency, time, code, or polarization division, or by some combination of the five, to avoid such coupling, but economic and other constraints often preclude complete segregation.

Space division. In space-division segregation, each communication channel is assigned its own transmission medium, for example, a pair of wires is a multipair cable (see **illus.**) or a separate radio propagation path. Coupling between channels is caused by the physical proximity and relative orientation of the transmission media. The coupling is usually electromagnetic, a linear phenomenon that is independent of signal level in the channels. *See* ELECTROMAGNETIC COMPATIBILITY.

Frequency division. In frequency-division segregation, a single, common transmission medium (for example, a coaxial cable system) is used to provide many communications channels, and each channel is assigned a separate frequency band within the medium. A signal is placed in its channel by amplitude modulation (AM) or angle modulation frequency modulation (FM) or phase modulation (PM). With AM systems, nonlinearities in the transfer characteristics of system elements create harmonic, or intermodulation, distortion. Typically, nonlinearities are introduced by system amplifiers that are used to compensate for loss in the medium. Intermodulation products (distortions) have frequencies that are combinations (sums, differences, multiples) of the frequencies of the signals being carried by the system. Thus, nonlinearities spawn many coupling paths among the channels of the system. The coupling is nonlinear, that is, the losses of the paths are functions of signal amplitudes. In angle-modulation systems, the principal cause of intermodulation distortion is frequency dependency in the amplitude and delay characteristics of the transmission medium (including amplifiers), but variation of the delay characteristic of the medium as a function of signal level (AM to PM conversion) can also cause intermodulation distortion. The resulting coupling paths among the channels of the system are similar to those in an AM system. *See* AMPLITUDE MODULATION; DISTORTION (ELECTRONIC CIRCUITS); FREQUENCY MODULATION; MODULATION; PHASE MODULATION.

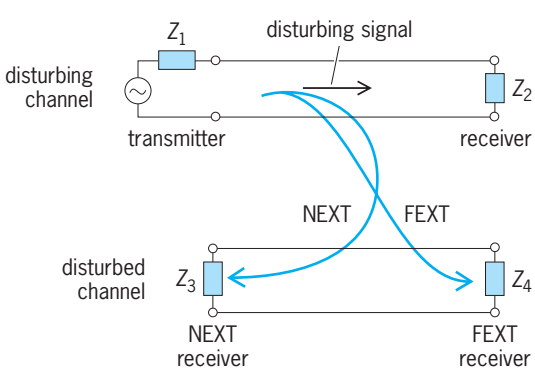

Crosstalk in a space-division system. Z_1, Z_2, Z_3, and Z_4 designate impedances.

Time division. In time-division segregation, a single common transmission medium is time-shared to provide many communications channels, with each channel allotted a separate time interval. These may be assigned on the basis of relative activity among the channels, or each channel may be periodically assigned a like interval regardless of activity, which is usually the case with digital systems. Interchannel coupling is the result of the presence of vestigial energy in the transmission medium from previous channel assignments. The coupling tends to be linear (the loss tends to be independent of signal amplitude), and often involves an impedance that is common to both the disturbing and disturbed channels. *See* PULSE MODULATION.

Code division. In code-division segregation, which is generally used by spread-spectrum communications systems, each communications channel is modulated by a coding sequence that is orthogonal to the coding sequences of the other channels. Perfect orthogonality is generally not achieved due to imperfections, like finite-length coding sequences, which cause some crosstalk. *See* SPREAD SPECTRUM COMMUNICATION.

Polarization division. In polarization-division segregation, transmissions are isolated from each other by use of opposite polarizations (for example, right-hand versus left-hand circular polarizations or vertical versus horizontal linear polarizations). Perfect cross-polarization isolation is not attainable due to antenna effects, antenna alignments, and variations of the transmission medium, leading to some degree of crosstalk. *See* POLARIZATION OF WAVES.

Classification. Crosstalk is classified in a variety of ways. The type of coupling (electromagnetic, intermodulation, or common impedance) indicates the

mechanism. The terms near-end crosstalk (NEXT) and far-end crosstalk (FEXT) indicate the relative directions of signal propagation in the disturbed and disturbing channels (see illus.). The terms direct crosstalk, where the disturbing channel couples directly to the disturbed channel, and indirect crosstalk, where the coupling path between disturbing and disturbed channels involves a third, or tertiary, channel, are often used to further describe crosstalk caused by electromagnetic coupling. Interaction crosstalk (IXT) is a term used to further describe indirect crosstalk that couples from the disturbing channel to the tertiary channel at one place, propagates along the tertiary channel, and subsequently couples into the disturbed channel at another place. Transverse crosstalk is a term that includes all direct and indirect crosstalk that is not interaction crosstalk.

Intelligible crosstalk is understood by the receiver of the disturbed channel, whereas nonintelligible crosstalk is not. Intelligible crosstalk most often occurs between channels carrying similar analog signals, such as analog speech signals. Intelligible crosstalk between channels, one of which is carrying a digital signal, or between channels carrying different types of analog signals (a voice signal and a video signal) is less likely to occur because the receiver does not comprehend the information in the disturbing signal. Nonintelligible crosstalk often affects the receiver in much the same way as noise, but in some cases it may be nearly intelligible, for example, cross-talk from one analog speech channel to another where intermodulation coupling has inverted the frequency spectrum without masking the syllabic nature of the disturbing speech signal. *See* ELECTRICAL NOISE.

Quantitative measures. The magnitude of crosstalk that becomes disruptive to a communications channel depends on many factors: the type of signals involved, the intelligibility of the crosstalk, the acuity of the receiver, the activity factors for the channels, and so on. In some cases crosstalk affects the disturbed channel in much the same way as noise, and signal-to-noise ratios are applicable measures. Digital channels tend to be of this nature. Typically, binary (two-level) digital signals require peak signal–to–root-mean-square noise ratios equal to, or better than, about 20 dB (voltage or current ratios of 10:1 or better). Crosstalk from one analog video channel to another produces an effect much different from noise, but a signal-to-interference ratio is an appropriate measure because the primary signal is usually present. Peak signal-to-peak interference ratios of 60 dB or better (voltage or current ratios of 1000:1 or better) are generally required for good reception of such signals. *See* SIGNAL-TO-NOISE RATIO.

In other cases, such as intelligible crosstalk in telephony, signal-to-crosstalk ratios are not applicable measures because the primary signal is not always present to mask the interference. In telephony, detection of intelligible crosstalk by the receiver is particularly disruptive—it does not really distort the primary signal, but it represents a breach of privacy—the detection of nonintelligible crosstalk that seems to be intelligible (of a syllabic nature) is almost as disruptive. Ever-present noise in the disturbed channel tends to mask crosstalk detection. To prevent detection, the noise power must exceed the crosstalk power by about 10 dB (noise-to-crosstalk power ratio of 10:1). When a disturbed telephony channel has very little noise, the typical receiver can detect crosstalk levels of about −85 dBm (−85 decibels below 1 milliwatt, or 3×10^{-9} mW). For comparison, a typical level for the primary signal at the receiver is about −25 dBm (3×10^{-3} mW). Objectives for the crosstalk performance of a telephony network are often expressed by crosstalk index, a measure of the probability of the detection of intelligible crosstalk during the course of a telephone call. Values for crosstalk index are expressed in percentages, and typical values are less than 1%. *See* DECIBEL.

Remedies. Most remedies for reducing crosstalk entail some technique for decreasing coupling among the communications channels involved. The use of twisted pairs in multipair cables, of shielding for each pair, or of coaxial conductors and optical fibers in place of pairs are common techniques for reducing electromagnetic coupling where space division alone is not adequate. Improved control of signal levels and improved linearity in amplifiers are effective for frequency-division systems. Often such improvements are made possible by advances in technology. Selection of the appropriate type of modulation (AM, FM, PM) is also important. Crosstalk within multichannel digital systems that transport digital versions of analog signals can be reduced with the help of a separate coder-decoder for each analog channel in place of a common, time-shared coder-decoder for all channels. *See* ELECTRICAL SHIELDING; OPTICAL FIBERS.

Signal processing in the disturbed channel can sometimes be effective in reducing crosstalk. Syllabic compandors, which are sometimes used with telephony channels on analog carrier systems to reduce noise in silent intervals during conversations, are effective in reducing crosstalk as well. Pulse shaping (modification of the signal spectrum) can be effective in reducing crosstalk between digital transmission systems and in reducing intersymbol interference within a digital transmission system. One of the most important tools for control of crosstalk is spectrum management, a systematic approach to the assignment of frequency bands, signal levels, and the like in such way as to offer the most efficient use of the transmission media. *See* ELECTRICAL COMMUNICATIONS; ELECTRICAL INTERFERENCE; RADIO SPECTRUM ALLOCATIONS; TELEPHONE SERVICE.

Jonathan W. Smith

Bibliography. Bell Telephone Laboratories, *Transmission Systems for Communications*, 5th ed., 1982; Bellcore and Bell Operating Companies, *Telecommunications Transmission Engineering*, vol. 1: *Principles*, 3d ed., 1990; C. Walker, *Capacitance, Inductance, and Crosstalk Analysis*, 1990.

Crown gall

A neoplastic disease of primarily woody plants, although the disease can be reproduced in species representing more than 90 plant families. The disease results from infection of wounds by the free-living soil bacterium *Agrobacterium tumefaciens* which is commonly associated with the roots of plants.

The first step in the infection process is the site-specific attachment of the bacteria to the plant host. Up to half of the bacteria become attached to host cells after 2 h. At 1 or 2 weeks after infection, swellings and overgrowths take place in tissue surrounding the site of infection, and with time these tissues proliferate into large tumors (see **illus.**). If infection takes place around the main stem or trunk of woody hosts, continued tumor proliferation will cause girdling and may eventually kill the host. Crown gall is therefore economically important, particularly in nurseries where plant material for commercial use is propagated and disseminated.

Crown gall tumors can be cultured free of bacteria on synthetic, defined media and, under proper conditions, can reinitiate tumors when grafted back onto healthy plants. Unlike healthy normal cells, crown gall tumor cells do not require an exogenous source of phytohormones (auxins and cytokinin) for growth in culture because they readily synthesize more than sufficient quantities for their own growth. The tumor cells also synthesize basic amino acids, each conjugated with an organic acid, called opines. Octopine and nopaline are opines that are condensation products of arginine and pyruvic acid, and arginine and 2-ketoglutaric acid, respectively. The tumor cells also grow about four times faster and are more permeable to metabolities than normal cells.

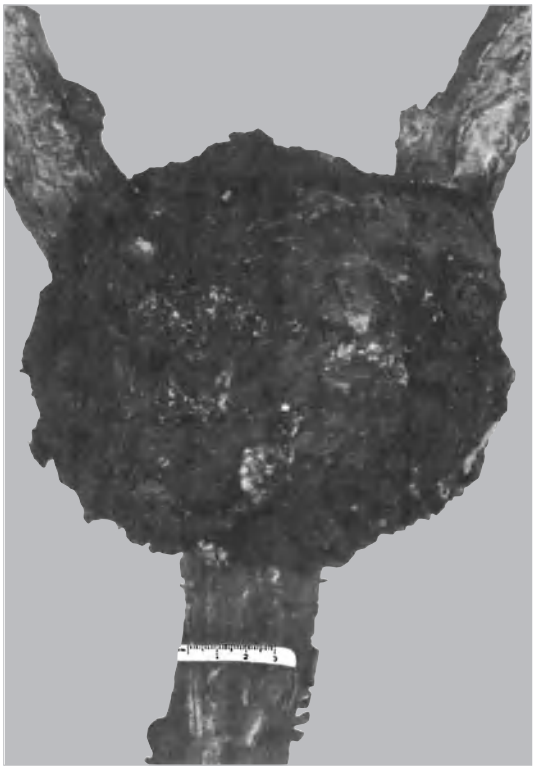

Crown gall on peach.

These cellular alterations, such as the synthesis of opines and phytohormone regulation, result from bacterial genes introduced into host plant cells by *A. tumefaciens* during infection. Although it is not understood how these genes are introduced into the plant cell, the genes for the utilization of these opines and for regulating phytohormone production have been found to be situated on an extrachromosomal element called the pTi plasmid. This plasmid, harbored in all tumor-causing *Agrobacterium* species, also carries the necessary genetic information for conferring the tumor-inducing and host-recognition properties of the bacterium. It is thus possible to convert avirulent strains of *A. tumefaciens* and related species to virulent forms by introducing the pTi plasmid into them through standard microbial genetic procedures. Also, the host range of a given *Agrobacterium* can be modified by introducing a pTi plasmid from a bacterial strain having a different host range.

Analyses made of crown gall nuclear genetic material (DNA) have revealed the presence of a specific segment of pTi plasmid DNA known as the T-DNA. Several distinct genes such as chloramphenicol acetyltransferase, and β-galactosidase have been inserted into the T-DNA by recombinant DNA methods. These genes can be introduced into plants through infection by *A. tumefaciens* carrying a pTi plasmid with one of these genes inserted in an appropriate location within the T-DNA. These molecular and genetic manipulations have made it feasible to genetically modify plants for desirable traits, a procedure popularly called genetic engineering.

Transfer of the T-DNA requires a set of genes that are clustered in another region of the pTi plasmid, known as the *vir* region. *Vir* region genes are essential for virulence of *A. tumefaciens*, and seem to respond to signals from wounded plant tissue. Although these signals remain to be identified, the interactions with the *vir* genes suggest a long-time association and dependency of *A. tumefaciens* with its host plants. Consequently, crown gall is a result of this unique bacteria-plant interaction, whereby *A. tumefaciens* genetically engineers its host to produce undifferentiated growth in the form of a large tumor, in which there is the synthesis of a unique food source in the form of an opine for specific use by the bacterial pathogen. *See* BACTERIAL GENETICS; GENETIC ENGINEERING; PLANT HORMONES; PLANT PATHOLOGY.

Clarence J. Kado

Crushing and pulverizing

The reduction of materials such as stone, coal, or slag to a suitable size for their intended uses such as road building, concrete aggregate, or furnace firing. Crushing and pulverizing are processes in ore dressing needed to reduce valuable ores to the fine size at which the valueless gangue can be separated

Crushing specifications

Category	Feed	Product	Reduction ratio	Equipment used
Primary crushing	27–12 in. (69–30 cm)	9–4 in. (23–10 cm)	3:1	Jaw, gyratory, cone
Secondary crushing (one or two stages)	9–4 in. (23–10 cm)	1–$^1/_2$ in. (25–13 mm)	9:1	Hammer mill, jaw, gyratory, cone, smooth roll, and toothed rolls
Pulverizing	1–$^1/_2$ in. (25–13 mm)	60–325 mesh	60:1	Ball and tube, rod, hammer, attrition, ball race, and roller mills

from the ore. These processes are also used to reduce cement rock to the fine powder required for burning, to reduce cement clinker to the very fine size of portland cement, to reduce coal to the size suitable for burning in pulverized form, and to prepare bulk materials for handling in many processes. *See* MATERIALS-HANDLING EQUIPMENT; ORE DRESSING.

Equipment suitable for crushing large lumps as they come from the quarry or mine cannot be used to pulverize to fine powder, so the operation is carried on in three or more stages called primary crushing, secondary crushing, and pulverizing (see **table**). The three stages are characterized by the size of the feed material, the size of the output product, and the resulting reduction ratio of the material. The crushing-stage output may be screened for greater uniformity of product size.

Reduction in size is accomplished by five principal methods: (1) crushing, a slow application of a large force; (2) impact, a rapid hard blow as by a hammer; (3) attrition, a rubbing or abrasion; (4) sudden release of internal pressure; and (5) ultrasonic forces. The last two methods are not in common use.

Crushers. All the crushers used in primary and secondary crushing operate by crushing as defined above except the hammer mill, which is largely impact with some attrition.

Primary crushers. The Blake jaw crusher using a double toggle to move the swinging jaw is built in a variety of sizes from laboratory units to large sizes having a feed inlet 84 by 120 in. (213 by 305 cm; **Fig. 1**). Large units have a capacity of over 1000 tons/h (900 metric tons/h) and produce a 9-in. (23-cm) product. The Dodge jaw crusher uses a single toggle

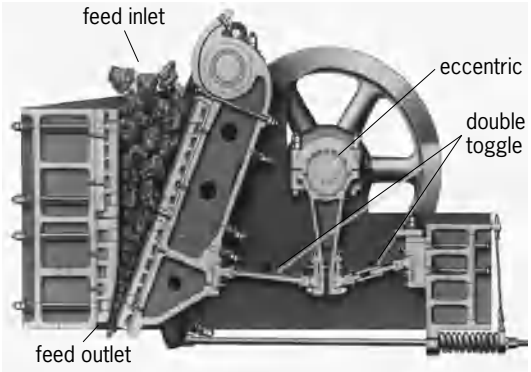

Fig. 1. Blake-type jaw crusher. (*Allis Chalmers Co.*)

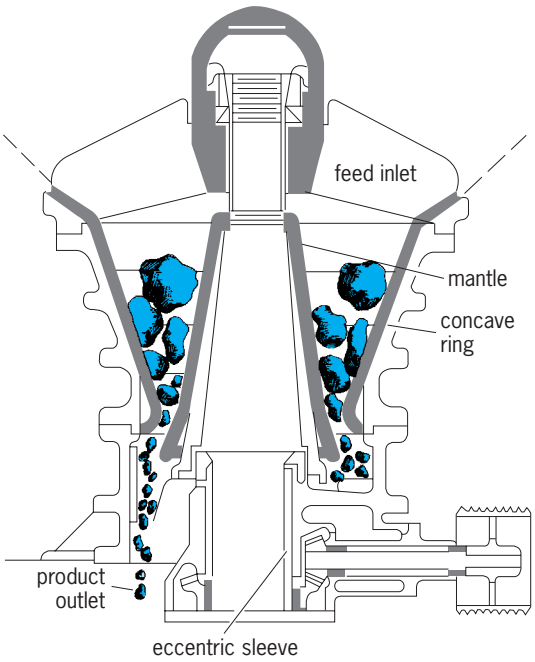

Fig. 2. Gates-type gyratory crusher. (*Allis Chalmers Co.*)

or eccentric and is generally built in smaller sizes.

The Gates gyratory crusher has a cone or mantle that does not rotate but is moved eccentrically by the lower bearing sleeve (**Fig. 2**). A 42- by 134-in. (107-by 340-cm) crusher has a capacity of 850 tons/h (770 metric tons/h), crushing rock from 27 to 8 in. (69 to 20 cm).

The Symons cone crusher also has a gyratory motion, but has a much flatter mantle or cone than does the gyratory crusher (**Fig. 3**). The top bowl is spring-mounted. It is used as a primary or secondary crusher.

Angle of nip is the largest angle that will just grip a lump between the gyratory mantle and ring, between the smooth jaws of a jaw crusher, or between the pair of smooth rolls. Depending on the material, it is 8–48°, but 32° is commonly used.

Secondary crushers. The single-roll crusher and the double-roll crusher have teeth on the roll surface and are used mainly for coal (**Fig. 4**). Single-roll crushers 36 in. (91 cm) in diameter by 54 in. (137 cm) long have a capacity of 275 tons/h (250 metric tons/h) crushing run-of-mine coal to $1^1/_4$-in. (32-mm) ring size. Smooth rolls without teeth are sometimes used

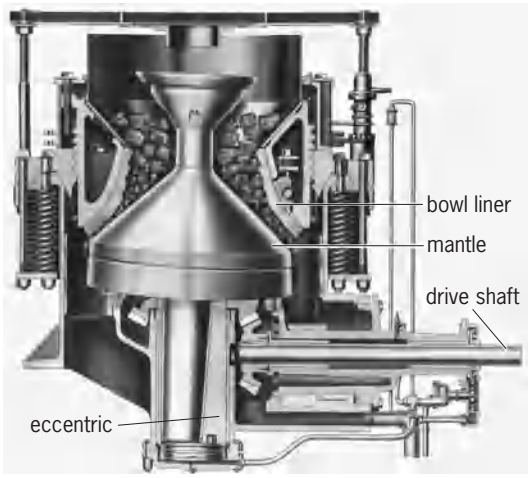

Fig. 3. Symons cone crusher. (*Nordberg Co.*)

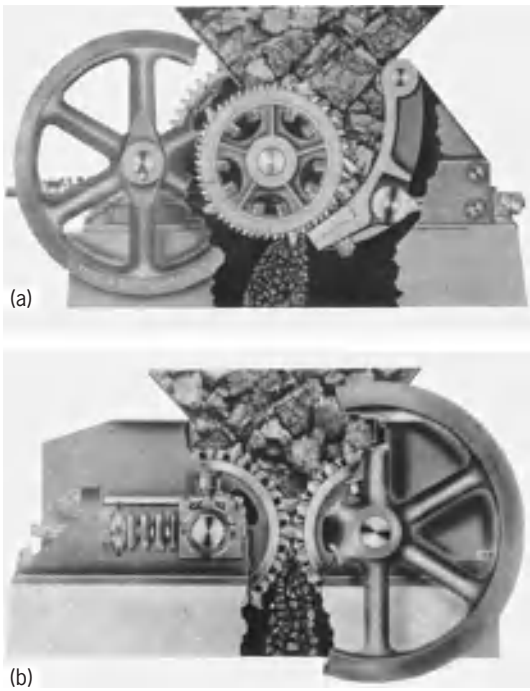

Fig. 4. Types of roll crusher. (*a*) Single-roll. (*b*) Double-roll. (*Link Belt Co.*)

for crushing ores and rocks.

The hammer crusher is the type of secondary crusher most generally used for ore, rock, and coal (**Fig. 5**). The reversible hammer mill can run alternately in either direction, thus wearing both sides of the hammers (**Fig. 6***a*). A hammer mill 42 in. (107 cm) in diameter by 82 in. (208 cm) long crushes 300 tons/h (270 metric tons/h) of coal to $^1/_4$ in. (6 mm) for coke oven feed. The ring coal crusher is a modification of the hammer mill using rings instead of hammers (Fig. 6*b*).

Pulverizers. In open-circuit pulverizing, the material passes through the pulverizer once with no removal of fines or recirculation.

In closed-circuit pulverizing, the material discharged from the pulverizer is passed through an external classifier where the finished product is removed and the oversize is returned to the pulverizer for further grinding.

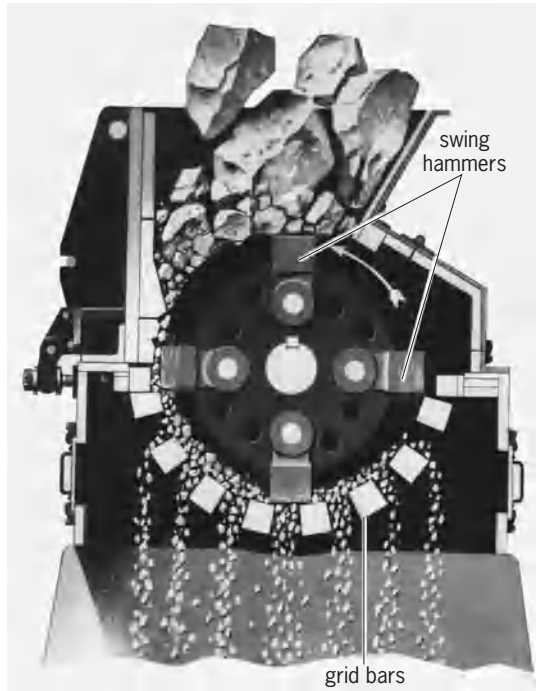

Fig. 5. Hammer crusher, a secondary crusher for ore, rock, and coal. (*Jeffrey Co.*)

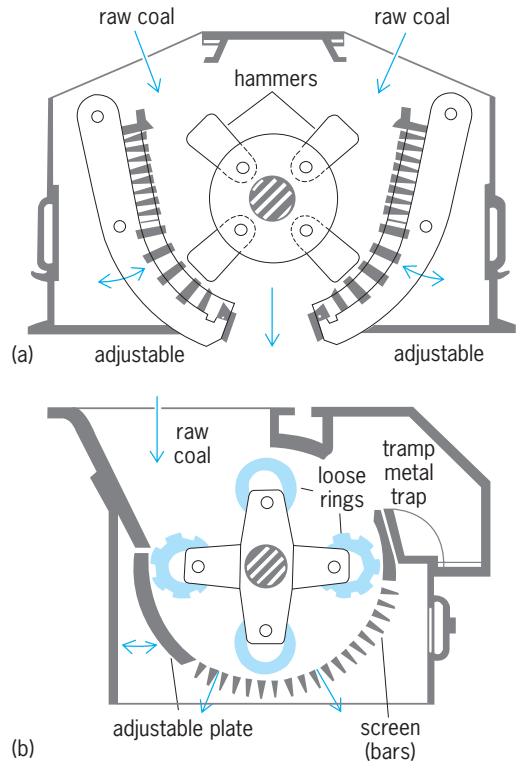

Fig. 6. Two examples of coal crusher. (*a*) Hammer mill type. (*b*) Ring type. (*Penna Crusher*)

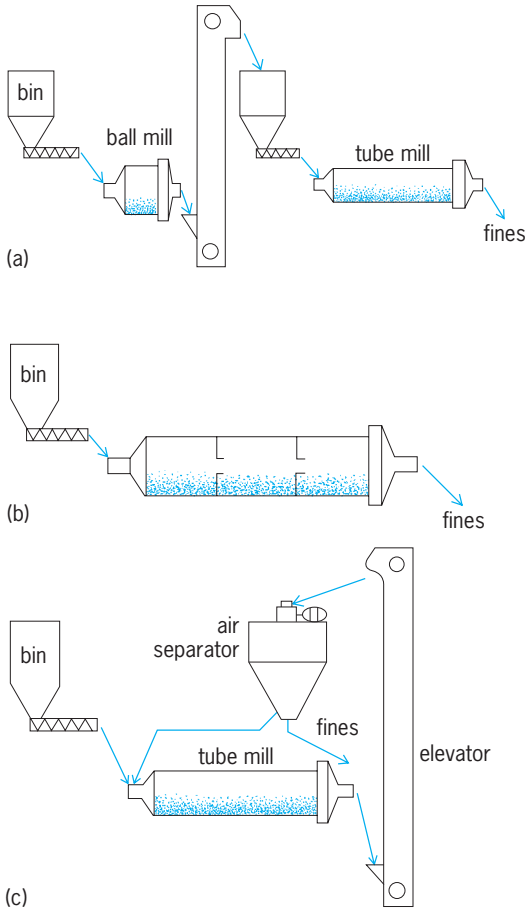

Fig. 7. Pulverizers. (*a*) Open-circuit type with preliminary ball mill stage followed by tube mill. (*b*) Open-circuit three-compartment tube mill. (*c*) Closed-circuit tube mill with an air separator.

Ball and tube mills, rod mills, hammer mills, and attrition mills are pulverizers operating by impact and attrition. In ball race and roller pulverizers, crushing and attrition are used.

Buhrstone pulverizer. The Buhrstone pulverizer is one of the oldest forms of pulverizers; a few still operate in flour and feed mills. The pulverizing is done between a rotating lower stone and a floating upper stone, which is restrained from rotating. Modern pulverizers use tumbling or rolling action instead of the grinding action of such earlier mills. For example, smooth double rolls are used for pulverizing flour and other materials. *See* BUHRSTONE MILL.

Tumbling pulverizers. There are two principal classes of tumbling pulverizer. Both consist of a horizontal rotating cylinder containing a charge of tumbling or cascading steel balls, pebbles, or rods. If the length of the cylinder is less than 2–3 diameters, it is a ball mill. If the length is greater than 3 diameters, it is a tube mill. Both types are extensively used in ore and rock pulverizing where the material is abrasive, as it is easy to add balls during operation. *See* TUMBLING MILL.

Ball mills charged with large steel balls 2–3 in. (5–8 cm) in diameter are often used as preliminary pulverizers in series with a tube mill in open circuit (**Fig. 7***a*). The Krupp mill is a preliminary ball mill which discharges the material through a screen which forms the cylindrical surface.

Pebble mills use balls and liners of quartz for application where contamination of the product with iron would be detrimental. For products coarser than 6 mesh, steel rods are sometimes used instead of balls. *See* PEBBLE MILL.

In some cement plants and many mining operations, ball, tube, rod, and pebble mills operate wet, water being added with the feed and the product being discharged as a slurry. Rake or cone type classifiers are used in closed circuits that employ wet grinding.

Tube mills for fine pulverization are used extensively in open circuit, or as compartment mills, or as closed-circuit mills (Fig. 7). A three-compartment, grate-discharge tube mill has larger balls in the first, medium-size balls in the second, and smaller balls in the third compartment (**Fig. 8**).

Cascade pulverizers are another form of tumbling pulverizer; they use the large lumps to do the pulverizing. The Aerfall pulverizer is built in diameters up to 16 ft (5 m). Feed is 12- to 18-in. (30- to 46-cm) lumps and the unit is airswept with an external classifier. A few steel balls are sometimes used but not over $2^1/_2$% of the volume.

The airswept conical ball mill is used in closed circuit with an external air classifier (**Fig. 9**). The

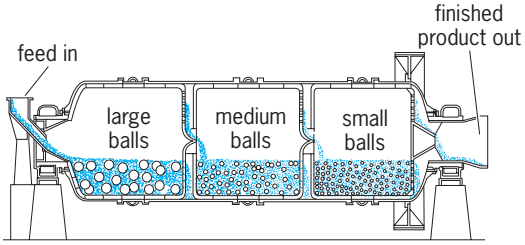

Fig. 8. Three-compartment tube mill pulverizer, containing three sizes of balls. (*Hardinge Co., Inc.*)

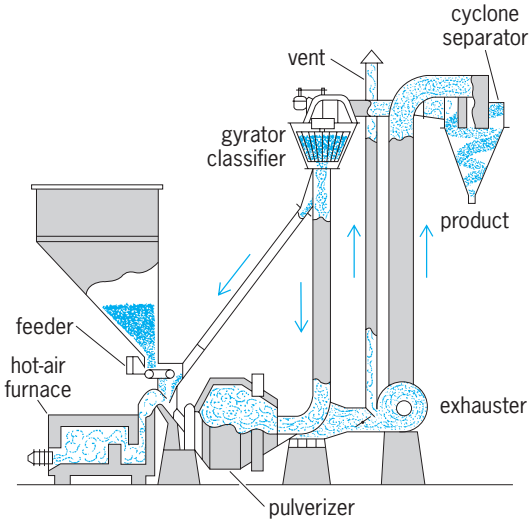

Fig. 9. Conical ball mill pulverizer in closed circuit with classifier. (*Hardinge Co., Inc.*)

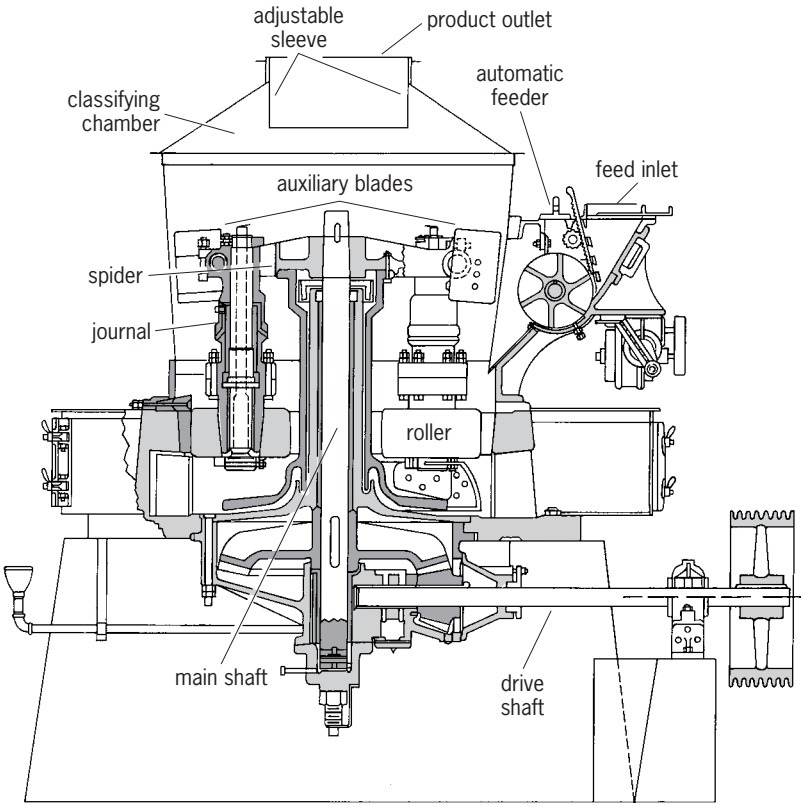

Fig. 10. A pulverizer of the airswept pendulum type. (*Raymond Pulverizer Division*, *Combustion Engineering Co.*)

conical shape of the mill classifies the balls, keeping the large balls at the feed end, where they are needed to crush the larger lumps, and the smaller balls at the small discharge end where the material is finer. This conical ball mill is also used extensively in open circuit in rocks and ores.

Roller pulverizers. The airswept pendulum roller pulverizer has rollers pivoted from a rotating spider on a vertical shaft; the centrifugal force of the rollers supplies the pressure between rollers and ring. Hot air enters below the rollers and picks up the fines. The rotating rollers and spider provide a centrifugal classifying action which returns the oversize to the pulverizing zone. For higher fineness a rotating classifier is added in the top housing (**Fig. 10**).

The ball race pulverizer uses large steel balls between races (**Fig. 11**). The bottom race is fixed, the intermediate race is rotated through driving lugs, and the top races are restrained from rotating by stops and are loaded by the springs. The material flows by centrifugal force from the center outward through the $12^1/_4$-in. (31-cm) and $9^1/_2$-in. (23.5-cm) balls. A ball race unit in closed circuit with an air separator is used for pulverizing cement rock.

Screens. Material is separated into desired size ranges by screening. For products 2 in. (5 cm) or larger, an open-end perforated cylindrical rotary screen is used. It is on a slight inclination, and the undersize goes through the holes and the oversize passes out through the open lower end.

For finer products from 2 in. (5 cm) down to fine

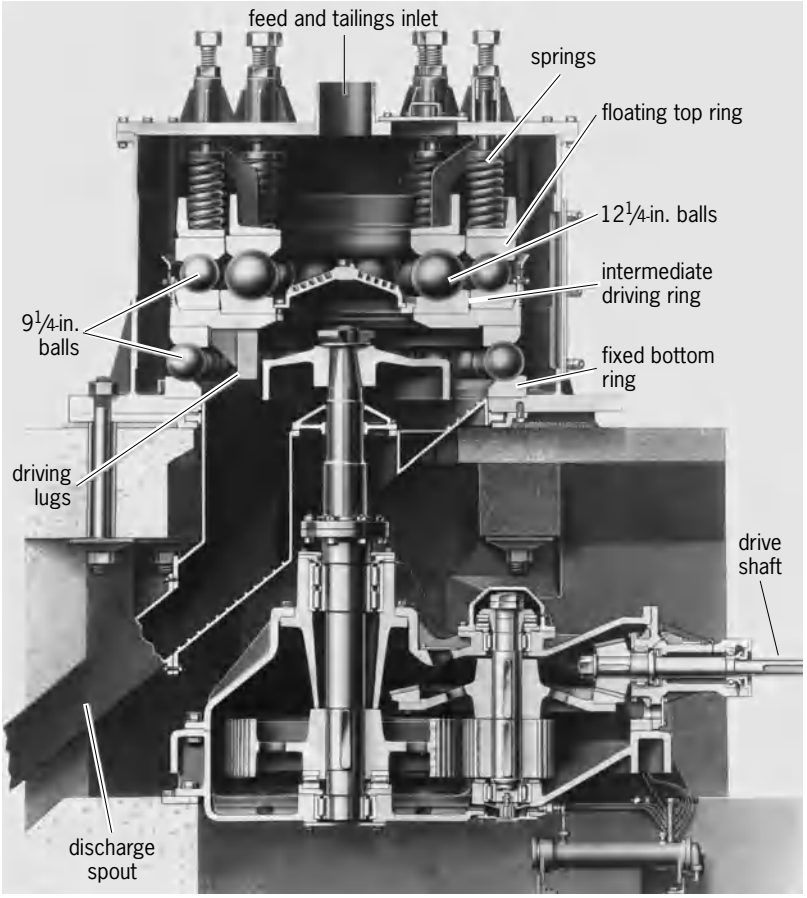

Fig. 11. Ball race rock pulverizer for use in closed circuit. (*Babcock and Wilcox Co.*)

powder passing through a 200-mesh sieve, shaking or vibrating screens of woven wire are used.

Shaking screens have either an eccentric drive or an unbalanced rotating weight to produce the shaking. Vibrating screens have electric vibrators (**Fig. 12**). They are often hung by rods and springs from overhead steel, giving complete freedom to vibrate. In some applications the screen is double decked, thus producing three sizes of product.

Direct-fired pulverizers for coal. Most modern large coal-fired furnaces are fed directly by the pulverizer, the coal rate being regulated by the rate of pulverizing. Air at 300–600°F (150–315°C) is supplied to the pulverizer to dry and preheat the coal. In a bowl-type medium-speed pulverizer, the springs press the pivoted stationary rolls against the rotating bowl grinding ring, crushing the coal between them (**Fig. 13**).

Fig. 12. Vibrating screen classifier. (*Jeffrey Co.*)

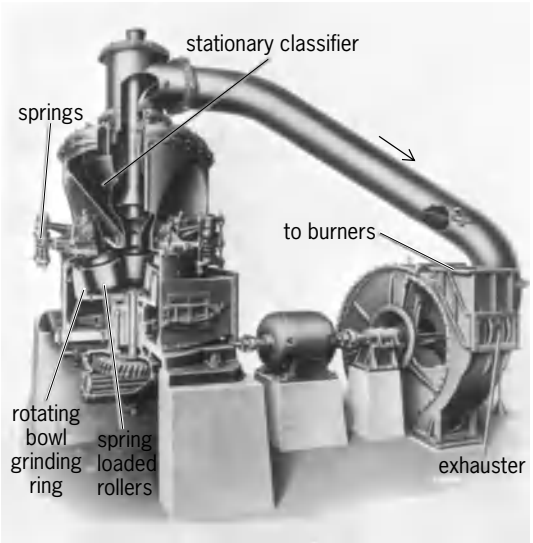

Fig. 13. Medium-speed bowl-mill coal pulverizer, which feeds burners. (*Combustion Engineering Co.*)

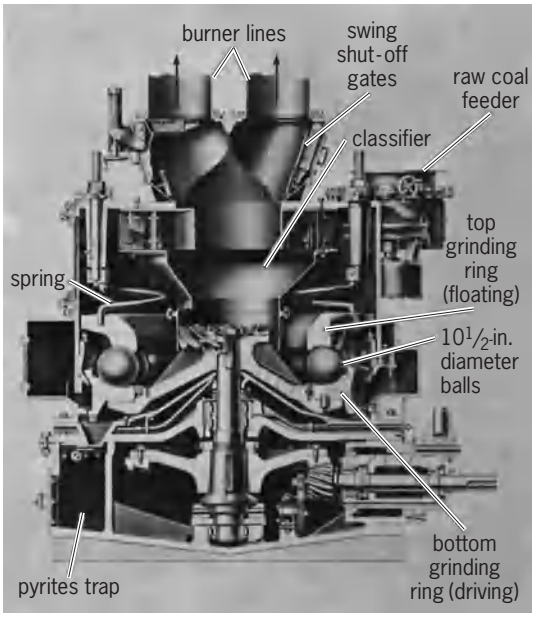

Fig. 14. Medium-speed ball race coal pulverizer, operated under pressure. 10½ in. = 27 cm. (*Babcock and Wilcox Co.*)

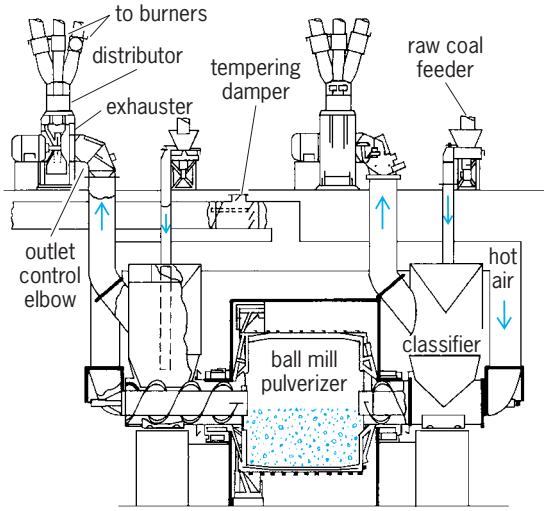

Fig. 15. Slow-speed coal pulverizer of the ball mill type; classifier sorts out oversize. (*Foster Wheeler Corp.*)

The self-contained stationary cone classifier returns the oversize to the pulverizing zone.

The medium-speed pulverizer of the ball race type uses closely spaced 18-in.-diameter (46-cm) balls between a lower rotating race and a floating top race (**Fig. 14**). The single coil springs restrain the top race and also apply the pressure needed. A stationary cone classifier is used. This pulverizer operates under pressure, using a blower on the clean inlet air instead of an exhauster on the coal-laden outlet air.

Slow-speed pulverizers of the ball mill type have classifiers and return the oversize for regrinding on both ends (**Fig. 15**).

High-speed pulverizers usually incorporate multiple stages such as a preliminary hammer mill and a secondary pulverizing stage. A built-in exhaust fan

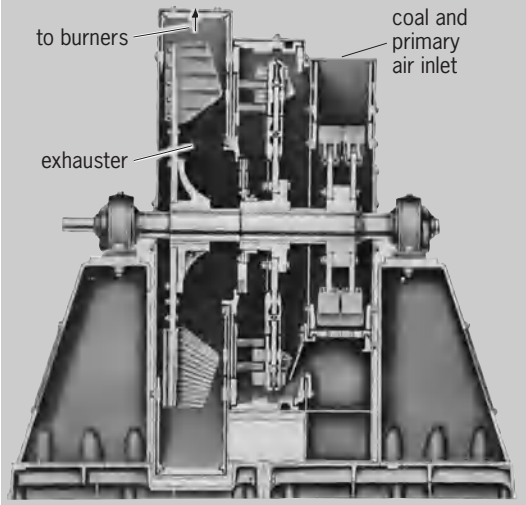

Fig. 16. High-speed coal pulverizer. (*Riley Stoker Corp.*)

exhausts the fuel and air from the unit (**Fig. 16**). *See* BALL-AND-RACE-TYPE PULVERIZER. Ralph M. Hardgrove

Bibliography. R. H. Perry and D. W. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; C. L. Prasher, *Crushing and Grinding Process Handbook*, 1987.

Crustacea

An extremely diverse, species-rich group of arthropods that have inhabited marine environments since the beginning of the Cambrian Period. In some ways, they may be thought of as the marine equivalents of the insects. Within the marine realm the crustaceans occupy as varied a spectrum of habitats as the insects inhabit on land. The Crustacea currently are regarded as a subphylum in the monophyletic phylum Arthropoda, although debate about this rank continues among carcinologists.

The Crustacea are one of the most complex animal groups to define because of their great diversity of structure, habit, habitat, and development. No one character or generalization will apply equally well to all. The adults of some aberrant or highly specialized parasitic species are not recognizable as crustaceans, and evidence of their kinship rests on knowledge of their life histories. Virtually every form of marine animal life, as well as a number of freshwater and even terrestrial animals, has its crustacean parasites.

Characterization

Unlike the other arthropod subphyla, the Crustacea exhibit not one but several body plans (*Baupläne*) indicative of higher taxa within the subphylum. In addition to the chitinous exoskeleton—which may vary from thin, almost membranous and transparent to thick and impregnated with calcium carbonate—all crustaceans, unless unduly modified for a parasitic lifestyle, have a segmented body differentiated into dorsal tergites and ventral sternites, with each somite bearing a pair of appendages. Cephalic segmentation varies little among the crustacean classes, consisting of five, often fused somites with paired antennules, antennae, mandibles, maxillules, and maxillae on somites 1–5, respectively. The eyes are borne on an acronal region considered by some to be presegmental and by others to represent a sixth cephalic somite. Traditionally, the postcephalic body "trunk" was divided into a thorax (with a variable number of somites in lower, or less morphologically advanced, crustaceans and eight in higher taxa) and an abdomen (also with a variable number of somites in lower crustaceans but only six or seven in higher taxa). However, recent advances in developmental genetics have broadened the interpretation of trunk segmentation. Based at least in part on *Hox* gene expression patterns, the crustacean trunk now is delineated by the presence or the anteriormost expression of the *Hox* gene *Abd-B*. The trunk is divided into a thorax, defined as the region just posterior to the cephalon, which is differentiated from posterior regions by a structurally distinct set of appendages, and which is anterior to or coincident with the anteriormost expression of *Abd-B*; and a pleon, defined as the region posterior to the thorax, which is differentiated by a structurally distinct set of limbs (which may be reduced or absent), and which is typically posterior to the gonopores (and the postulated anteriormost expression of *Abd-B*.) Posterior to the trunk is the abdomen, defined as the region posterior to the trunk that lacks appendages; it is posterior to the expression of *Abd-B* and exhibits no expression of *Hox* genes. *See* DEVELOPMENTAL GENETICS.

Distribution and Ecology

The Crustacea are ubiquitous. They live at almost all depths and levels of the sea, in freshwaters at elevations up to 12,000 ft (3658 m), in melted snow water, in the deepest sea abysses more than 6 mi (9 km) down, and in waters of 0°C (32°F) temperature. They also live on land, in trees, and in mountains, although most of these species must descend to salt water areas again to spawn their young. Some live in strongly alkaline waters and others in salt water that is at the saturation point; still others exist in hot springs and hydrothermal vents with temperatures in excess of 55°C (131°F). Species such as shrimps, prawns, crabs, or lobsters (**Fig. 1**) are familiar. However, there are many more with vernacular names such as water fleas, beach fleas, sand hoppers, fish lice, wood lice, sow bugs, pill bugs, barnacles, scuds, slaters, and krill or whale food, and many with no vernacular names at all.

All crustaceans of sufficient size, if not directly eaten by people somewhere in the world, are used as bait for food. Only a very few, if properly prepared, have proved to be poisonous. Myriads of small fry are the sustenance of larger aquatic animals, chiefly fish and whales. Crustaceans come in all sizes, from minute and microscopic (less than 1 millimeter in total length) to very large with spans in excess of 12 ft (3.7 m) from tip to tip of the laterally extended legs. Individuals of the giant xanthid crab of Australia have been recorded weighing 20 to 30 lb (9 to

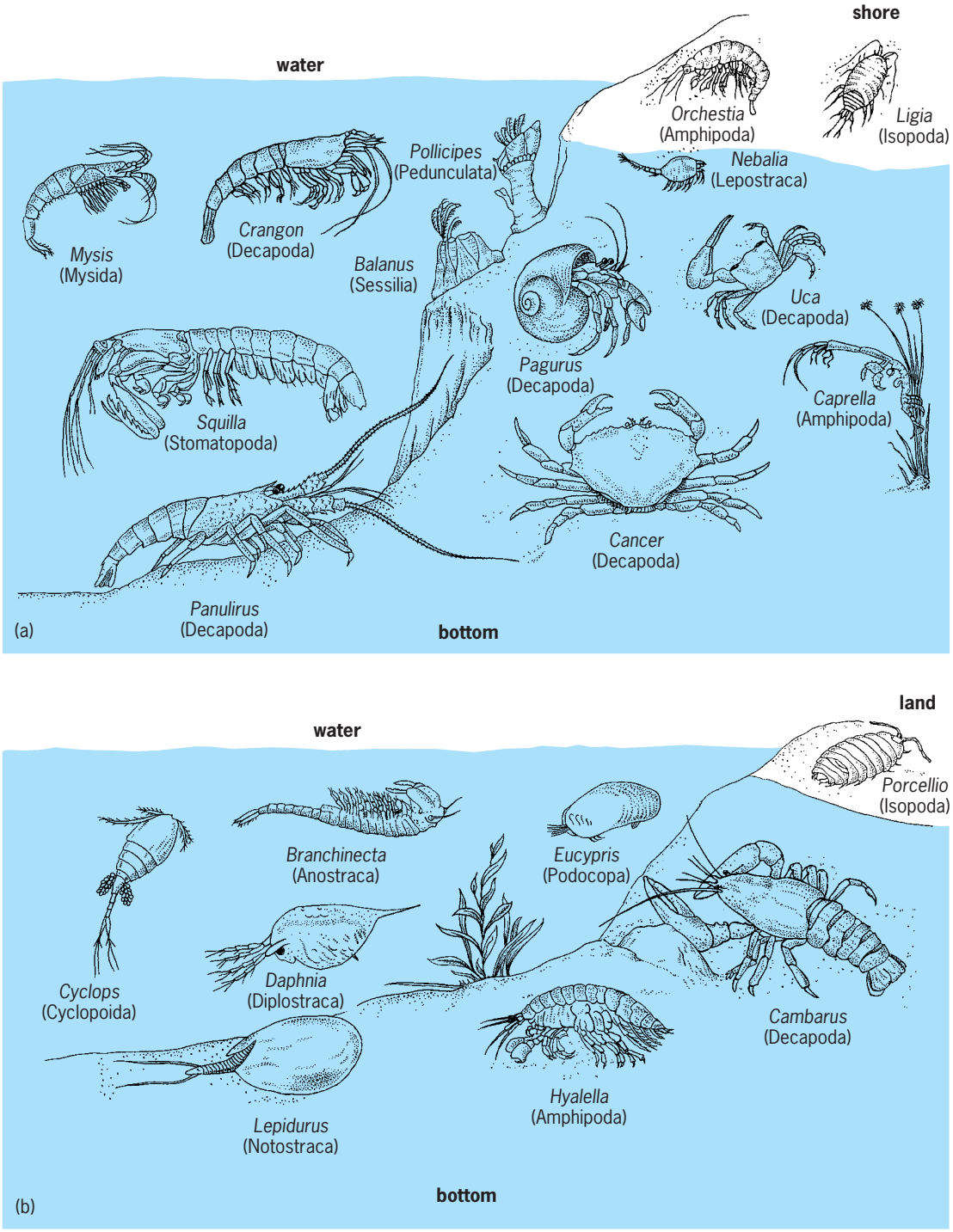

Fig. 1. Some crustaceans in their respective habitats, mostly reduced but not to same scale. Orders are in parentheses. (*a*) Marine forms and common names: *Pollicipes*, goose barnacle; *Balanus*, acorn barnacle; *Mysis*, opossum shrimp; *Orchestia*, sand hopper; *Caprella*, skeleton shrimp; *Crangon*, shrimp; *Panulirus*, spiny rock lobster; *Pagurus*, hermit crab; *Cancer*, edible crab; *Uca*, fiddler crab; *Squilla*, mantis shrimp. (*b*) Freshwater forms and common names if recognized. *Branchinecta*, fairy shrimp; *Daphnia*, water flea; *Eucypris*, seed or mussel shrimp; *Lepidurus*, tadpole shrimp; *Cyclops*, copepod; *Porcellio*, sow bug (on land); *Cambarus*, crayfish. (*After T. I. Storer et al., General Zoology, 6th ed., McGraw-Hill, 1979*)

14 kg), and the American lobster, the heaviest crustacean known, can weigh 44 lb (20 kg).

Most crustaceans are omnivorous and essentially scavengers. Many are filter feeders and screen particulate life, plankton, and organic detritus from the waters in which they live; others are largely carnivorous; and some are vegetarian. Among the vegetarians are the grazers of the ocean meadows, which convert the microscopic plant life (diatoms) into flesh and food for larger animals, which in turn are harvested as food for humans. *See* FEEDING MECHANISMS (INVERTEBRATE).

Taxonomy

Although disagreement persists regarding elements of the hierarchical classification of the Crustacea, in recent years a consensus, for the most part, has been reached. Still contentious is the class Maxillopoda, for which no clearly defined, single "good" character is shared by the various assigned taxa. Nonetheless, until such time as molecular, fossil, and morphological studies on members of the class provide some level of compromise, it seems in the best interest of nomenclatorial stability to continue to recognize the class Maxillopoda. *See* MAXILLOPODA.

The current crustacean classification scheme (based on Martin and Davis, 2001) is as follows (see separate articles on many of these texa):

Subphylum Crustacea
- Class Branchiopoda
  - Subclass Sarsostraca
    - Order Anostraca
  - Subclass Phyllopoda
    - Order: Notostraca
      - Diplostraca
- Class Remipedia
  - Order Nectiopoda
- Class Cephalocarida
  - Order Brachypoda
- Class Maxillopoda
  - Subclass Thecostraca
    - Infraclass Fascetotecta
    - Infraclass Ascothoracida
      - Order: Laurida
        - Dendrogastrida
    - Infraclass Cirripeda
      - Superorder Acrothoracica
        - Order: Pygophora
          - Apygophora
      - Superorder Rhizocephala
        - Order: Kentrogonida
          - Akentrogonida
      - Superorder Thoracica
        - Order: Pedunculate
          - Sessilia
  - Subclass Tantulocarida
  - Subclass Branchiura
    - Order Arguloida
  - Subclass Pentastomida
    - Order: Cephalobaenida
      - Porocephalida
  - Subclass Mystacocarida
    - Order Mystacocaridida
  - Subclass Copepoda
    - Infraclass Progymnoplea
      - Order Platycopioida
    - Infraclass Neocopepoda
      - Superorder Gymnoplea
        - Order Calanoida
      - Superorder Podoplea
        - Order: Misophrioida
          - Cyclopoida
          - Gelyelloida
          - Mormonilloida
          - Order: Harpacticoida
            - Poecilostomatoida
            - Siphonostomatoida
            - Monstrilloida
- Class Ostracoda
  - Subclass Myodocopa
    - Order: Myodocopida
      - Halocyprida
  - Subclass Podocopa
    - Order: Platycopida
      - Podocopida
- Class Malacostraca
  - Subclass Phyllocarida
    - Order Leptostraca
  - Subclass Hoplocarida
    - Order Stomatopoda
  - Subclass Eumalacostraca
- Superorder Syncarida
  - Order: Bathynellacea
    - Anaspidacea
- Superorder Peracarida
  - Order: Spelaeogriphacea
    - Thermobaenacea
    - Lophogastrida
    - Mysida
    - Mictacea
    - Amphipoda
    - Isopoda
    - Tanaidacea
    - Cumacea
- Superorder Eucarida
  - Order: Euphausiacea
    - Amphionidacea
    - Decapoda

One group of parasites now acknowledged as probably crustacean is the Pentastomida. Until quite recently, the Pentastomida had been classified as a distinct phylum, all adult members of which are parasites of the respiratory tracts of vertebrates. However, molecular and spermatological evidence of a phylogenetic relationship between pentastomids and branchiurans accumulated over the past quarter-century—with additional supporting evidence from similarities in embryogenesis and details of cuticular and nervous system structures—has resulted in the acceptance of this very distinctive group of parasites as crustacean in origin. The primary larvae are autoinfective to the definitive host, but they can also migrate to the host's gut and pass out with the feces. Many pentastomids require an intermediate host, which may be any kind of invertebrate. Larvae bore through the gut wall of the intermediate host, where they undergo further development to the infective stage. Once the intermediate host is consumed by a definitive host, the parasite makes its way from the host's stomach to its final destination in the lungs, nasal passages, or bronchi.

If this current phylogenetic assessment is accurate, crustacean morphological diversity and extremes in lifestyle are taken to an even greater level, with representatives adapted to life in the respiratory passages of crocodilians, reindeer, and lions. However,

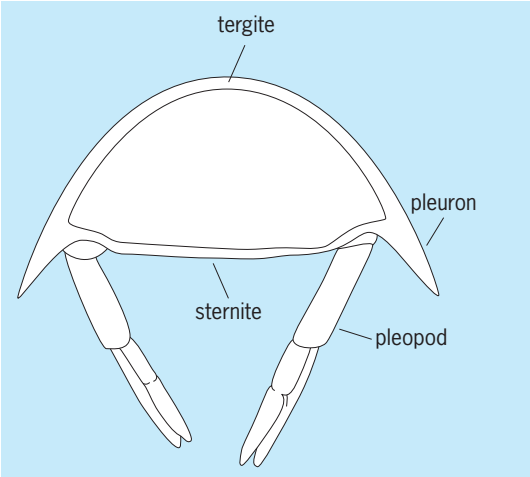

Fig. 2. Diagram of the fifth pleonal somite of the crayfish, with its paired fifth pleopods.

the placement of the Pentastomoida within the Crustacea is not universally accepted. The discovery of fossils from the Middle Cambrian limestone that are very similar to extant pentastomids would appear to dispute this classification. *See* PARASITOLOGY; PENTASTOMIDA.

General Morphology

The true body segments, the somites or metameres, are usually somewhat compressed or depressed (**Fig. 2**). Rarely are they circular rings. The somite's dorsum, or upper side, is the tergum or tergite; the lower or ventral surface, the sternum or sternite; and either lateral portion, the pleuron or epimeron. When prolonged or extended downward to overhang the attachment of the biramous appendages, of which one pair is a part of every typical crustacean somite, the structure is termed the epimeron. Between the embryonic anterior or prostomial element and the telson at the posterior end, the linear series of somites making up the body of a crustacean are more or less distinctly organized into three or four regions, or tagmata: the head (cephalon), thorax, and abdomen in lower crustaceans (for example, Branchiopoda, Cephalocarida, Maxillopoda) and head, thorax, pleon, and abdomen in malacostracans. The somites are variously fused with one another in diagnostic combinations in different groups of the Crustacea.

Body. The head region of the simplest type is found in the characteristic crustacean larva, the nauplius. When the nauplius hatches, it has only three pairs of appendages—the antennules, antennae, and mandibles—which correspond to the three fused somites of the primary head of the lower Crustacea. In the nauplius, there is a single median eye (known as the nauplius eye on the head that persists to become the sole visual organ in many of the lower Crustacea. In some it disappears; occasionally it remains as a vestige in species with paired compound eyes.

In more advanced forms, and at later developmental stages, two additional somites bearing the maxillules and maxillae become coalesced with the protocephalon to form the most frequently encountered crustacean head, or cephalon. Thoracic somites frequently are fused with the cephalon to form a cephalothorax. A dorsal shield, or carapace, of variable length arises from the dorsum of the third cephalic somite and covers the cephalon and cephalothorax to a varying extent (**Fig. 3**). In some crustaceans the carapace develops in the form of a bivalved shell; in sessile barnacles it is modified as a fleshy mantle reinforced by calcareous plates. The carapace reaches its greatest development in the malacostracan Decapoda where, including the head, it fuses dorsally with the eight thoracic somites and extends down laterally over the attachment of the appendages to enclose the branchial chamber on either side of the cephalothorax.

The chitinous cuticle covering the crustacean body is its external skeleton (exoskeleton). Ingrowths for muscle attachment develop from this cuticular exoskeleton as apodemes. Further development of the apodemes results in the formation of the endophragmal skeleton in the higher Crustacea. The chitin is flexible at the joints, in foliaceous appendages, and throughout the exoskeletons of many small and soft-bodied species, but it is often thickened and stiffened in others. It becomes calcified in many species as a result of the deposition of lime salts, and can be rock-hard in some crabs.

Appendages. The paired appendages of the cephalon and thorax (**Fig. 4**) are typically biramous and consist of two branches: the endopod and exopod. These arise from the distal end of a peduncle, the protopod, which is composed of two segments: the coxa, articulated with the body, and the basis; occasionally a precoxa may also be present. Either the endopod or exopod or both may be reduced or absent, or both may be coalesced, as in the foliaceous limbs of lower Crustacea. The protopod frequently has inner lobes (endites) or outer lobes (exites) or branchial structures (epipods). The

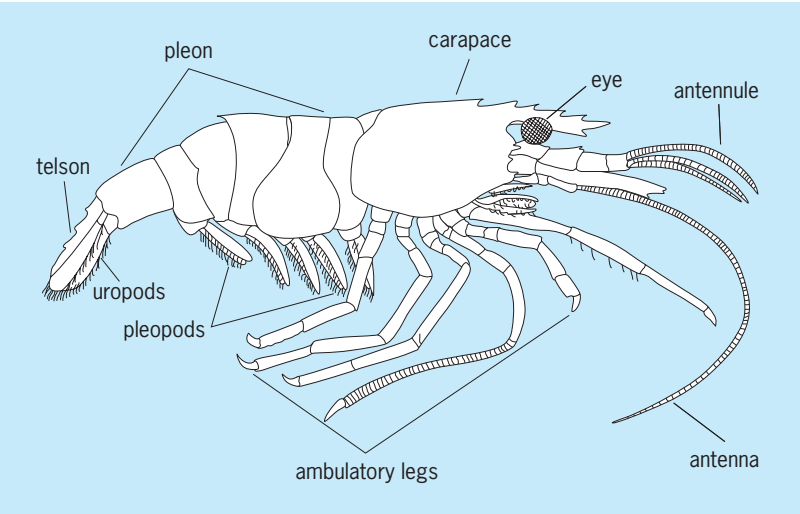

Fig. 3. External morphology of a shrimp. (*After P. A. McLaughlin, Comparative Morphology of Recent Crustacea, W. H. Freeman, 1980*)

endopod is definitely segmented in the higher Crustacea. The segments, which are usually present in full complement from the peduncular basis outward, are the ischium, merus, carpus, propodus, and dactyl. The endopods are variously modified to serve a number of functions and needs, such as sensory perception, respiration, locomotion, prehension (holding, seizing, or grasping) and manipulation of food, cleansing, defense, offense, reproduction, and sex recognition and attraction. If retained in the adult, exopods may remain leaf- or paddlelike, or become flagellated structures, facilitating swimming or aiding respiration. Pleonal appendages develop only in the Malacostraca; they are commonly used in locomotion and/or reproduction. The nonmalacostracan crustacean abdomen lacks appendages.

Digestive system. The digestive system is composed of a gut divided into three distinct regions and its accompanying glands and diverticula. Appendages adapted for feeding frequently include not only the mandibles, maxillules, and maxillae but also a number of thoracic appendages and occasionally the antennae.

In its simplest form, the gut consists of a foregut with an anteroventrally directed mouth, midgut, and hindgut with a posterior or posteroventrally directed anal orifice. The anterior part of the foregut often is esophageal in nature, whereas the posterior and larger portion usually consists of two parts. (In the anterior foregut of the Eucarida, a complex and elaborate grinding mechanism, the gastric mill, is developed). The posterior chamber is divided into dorsal and ventral filtering compartments for the straining of food. The midgut in most cases contains several ceca or diverticula, which produce digestive secretions or serve as organs for the absorption of food. When these ceca are present in considerable number, they become organized into a sizable gland, the midgut gland of higher Crustacea. The hindgut is typically short and terminates, with few exceptions, in a muscular anus on the underside of the telson. *See* DIGESTION (INVERTEBRATE).

Circulatory system. The crustacean circulatory system is, with few exceptions, an open system through which the blood or hemolymph flows freely throughout the hemocoelic cavity. Not all crustaceans have a heart, but in those that do, it is a simple single-chambered organ lying dorsally above the gut and enclosed in a pericardial sinus. The heart is perforated by openings, or ostia, that admit venous blood from the pericardial sinus. The heart may be elongated and tubular and extend through the greater part of the body, but generally it is a more compact organ. It pumps blood through an arterial system or through connecting sinuses or lacunae within the body tissues. In Crustacea lacking a heart, blood is moved by the muscular movements of the animal or its alimentary tract, circulating it through the body cavities or sinuses. The blood of malacostracan crustaceans is bluish because it contains the respiratory pigment hemocyanin. A few crustaceans have red blood due to the presence of hemoglobin. *See* HEART (INVERTEBRATE); RESPIRATORY PIGMENTS (INVERTEBRATE).

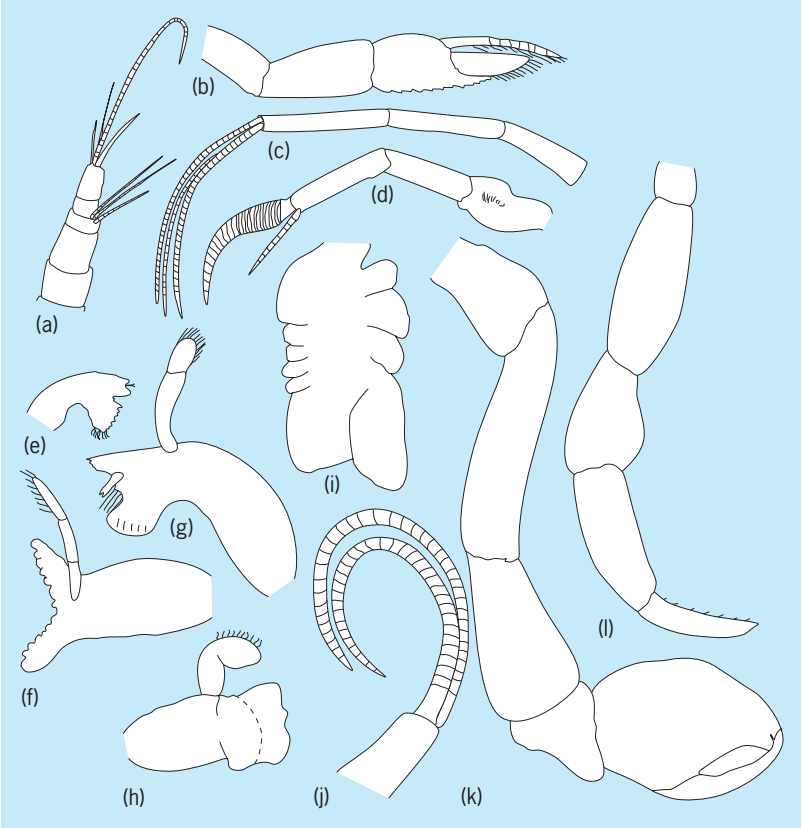

Fig. 4. Cephalic and thoracic appendages of representative crustaceans: (*a–d*) antennules; (*e–h*) mandibles; (*i–k*) thoracopods. (*a*) Brachypoda; (*b*) Leptostraca; (*c*, *d*) Decapoda; (*e*) Brachypoda; (*f*) Stomatopoda; (*g*) Mysida; (*h*) Decapoda; (*i*) Anostraca; (*j*) Sessilia; (*k*) Stomatopoda; (*l*) Decapoda. (*After P. A. McLaughlin, Comparative Morphology of Crustacean Appendages, in D. E. Bliss, ed., The Biology of Crustacea, vol. 2, 1982*)

Respiratory system. The majority of nonmalacostracan crustaceans have no specialized respiratory structures; respiration takes place across the integumental surfaces. However, in malacostracans oxygen uptake is accomplished principally by means of gills, although the general body surface, or special areas of it, also may participate. Some of the few species that have become more or less terrestrial in their habits have developed modifications of their branchial mechanism (such as villi, ridges, or water-retaining recesses), which when sufficiently moist enable them to breathe air. Some sow or pill bugs have special tracheal developments in their abdominal appendages for the same purpose.

Nervous system. As do other arthropods, crustaceans possess a dorsally located brain connected to a double ventral nerve cord by paired circumesophageal connectives. Both the brain and the ventral nerve cord are segmental structures, and each unit (neuromere) is joined by paired longitudinal connectives. This segmental aspect of the central nervous system is most evident in embryonic stages, and is generally obscured by condensation of the neuromeres into ganglia during later embryonic or larval development. In decapod embryos, the ventral nervous system consists of eight thoracic and eight pleonal neuromeres, whereas in many lower

crustaceans a much higher number of trunk neuromeres is present. *See* NERVOUS SYSTEM (INVERTEBRATE).

Organs of special sense. The organs of special sense are the eyes, antennules, and antennae. Crustacean eyes (photoreceptors) are of two types: sessile and stalked. Sessile eyes, found in numerous lower crustaceans and in larval stages of numerous stalk-eyed species, often are simple structures consisting of light-sensory elements, although some are weakly to moderately well-developed compound eyes. Stalked eyes are unique among the Crustacea; the stalks provide support and movement for the eyes, and most also contain optic ganglia and important endocrine organs. The crustacean compound eye contains from few to many structural units, or ommatidia.

The antennules and antennae are provided with a variety of sensory structures for the reception of chemical and mechanical stimuli. Specific structures on the antennules, called aesthetascs, act as organs of smell. Taste chemoreceptors are usually found on the mouthparts and dactyls of the pereopods (or pereiopods), or walking legs. Many of the hairs (setae) and bristles found on the crustacean body and appendages act as mechanoreceptors. Organs of balance (statocysts) are also present on basal segments of the antennules in many crustaceans, but on the uropods in most mysids. Although crustaceans may not hear in the accepted sense, they are sensitive to sound waves and vibrations; many have stridulating or sound-producing mechanisms on appendages or on somites overlapping one another on the carapace, which produce snapping, clicking, rasping, or rubbing noises. Crustacea have setae variously located on their bodies which are sensitive to various external stimuli. *See* CHEMORECEPTION; EYE (INVERTEBRATE); PHOTORECEPTION.

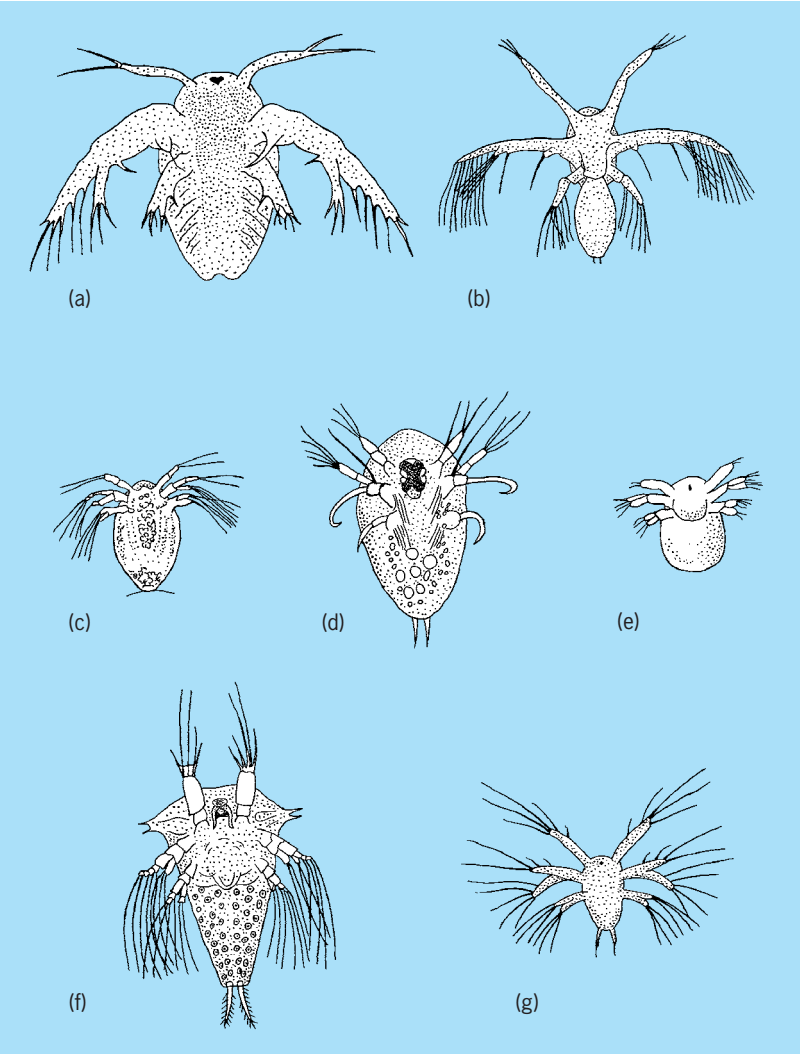

Fig. 5. Nauplii of (*a*) *Triops cancriformis* (Notostraca); (*b*) *Branchinecta occidentalis* (Anostraca); (*c*) *Lernaeocera branchialis* (Siphonostomatoida); (*d*) *Haemocera danae* (Monstrilloida); (*e*) *Meganyctiphanes norvegica* (Euphausiacea); (*f*) *Sacculina carcini* (Kentrogonida); (*g*) *Litopenaeus setiferus* (Decapoda). Enlarged but not to same scale. (*After R. E. Snodgrass, Crustacean Metamorphoses, vol. 131, no. 10, Smithson. Inst. Publ. 4260, 1956*)

Glands and glandular systems. Those glands recognized as definite excretory organs are the maxillary and antennal glands in the adult crustaceans. They are located in the cephalon, and rarely are both present at the same time. The antennal gland is best developed in the Malacostraca. It is often called the green gland and opens to the exterior on the underside of the coxal segment of the antennal peduncle. Many crustaceans, especially deep-sea forms, have phosphorescent or luminous organs. Barnacles and some other crustaceans have cement glands. This may also be true of tube-building species that construct their homes of extraneous materials. Species that produce encysted or drought-resistant eggs have other glands of special secretion. The X organ-sinus gland complex of the eyestalks is part of a neurosecretory system that produces hormones that control color change and pattern, molting cycle, oogenesis, and egg development within the ovary. *See* ENDOCRINE SYSTEM (INVERTEBRATE); NEUROSECRETION.

Reproductive system. The sexes are separate in most Crustacea and usually can be differentiated from each other by secondary sexual characters. Chief among these characters are the size and shape of the body, appendages, or both, and placement of the genital apertures. Hermaphroditism is the rule in the Cephalocarida, Remipedia, some ostracodes, and sessile Cirripedia (barnacles), in isolated cases in other crustaceans, and in certain parasitic forms. Parthenogenesis (eggs developing and hatching without prior fertilization) occurs frequently in some of the lower crustaceans such as the Branchiopoda and Ostracoda, which have what might be called an alternation of generations. The parthenogenetic generations (and there are usually very many of them in succession) alternate with a generation produced by fertilized eggs.

Protandric hermaphroditism, in which individuals reach maturity as males but subsequently become functional females, has been observed in some decapods. Protogyny, a condition in which the development is reversed, has been reported in tanaids and some carideans. Sexual differentiation is under control of the androgenic gland. This is a reversal of sex

in the individual or in the species that reach maturity as males but become female in their third year and function as females.

The eggs of most crustaceans are carried attached to the female until hatched. Some females develop brood pouches in which the young are retained for a time. A nutrient secretion, which sustains the young until they are released, is produced in some species having a brood chamber. Penaeid shrimp and a few of the lower Crustacea deposit their eggs in the medium in which they live, in some cases attaching them to aquatic vegetation.

Development. The nauplius larva has long been considered characteristic of Crustacea (**Fig. 5**). Its oval, unsegmented body has three pairs of appendages indicative of three embryonic somites. These are the antennules (which are uniramous in the nauplius), the biramous antennae, and mandibles. In addition to serving other functions such as sensation and feeding, all three pairs of appendages are the organs of locomotion of this free-swimming larva. There is a single dorsal, median, nauplius eye composed of three closely united, similar parts that form a relatively simple compound eye and a frontal organ. The integument of the upper surface of the nauplius body has the appearance of a dorsal shield, as in most crustaceans.

Typical as this first larval stage may be, it does not appear in all crustaceans. It is common in the lower forms, but among malacostracan crustaceans, free-living nauplius larvae occur in only the Euphausiacea and the decapod suborder Dendrobranchiata (penaeid shrimp). This larval form is replaced in many malacostracans by the so-called egg-nauplius, an embryonic stage characterized by an early germ band marking the ventral side of the embryo, which is attached to a relatively large amount of yolk and covered by a few extraembryonic cells. The anterior region is formed by the paired head lobes, which give rise to the lateral compound eyes and the protocerebrum. Posteriorly, these lobes are followed by the characteristic, well-developed antennules, antennae, and mandibles. The occurrence of an egg-nauplius is always correlated with a relatively yolky egg but is not necessarily correlated with direct development. Stomatopods and most decapods hatch as advanced planktonic larvae (zoeae) with at least one well-developed thoracic appendage. These larvae resemble the advanced stages of euphausiid and dendrobranchiate decapod larvae that hatch at the free-living naupliar stage.

Life histories vary from the simple to the complex within the different groups of Crustacea. Following the naupliar stage, whether it is passed in the egg or not, very remarkable larval stages follow in some groups the free-swimming cypris larvae of barnacles; the phyllosomae of the spiny lobsters; the pseudozoeal larvae of the mantis shrimps (Stomatopoda); and the zoeae and megalopae of crabs (**Fig. 6**).

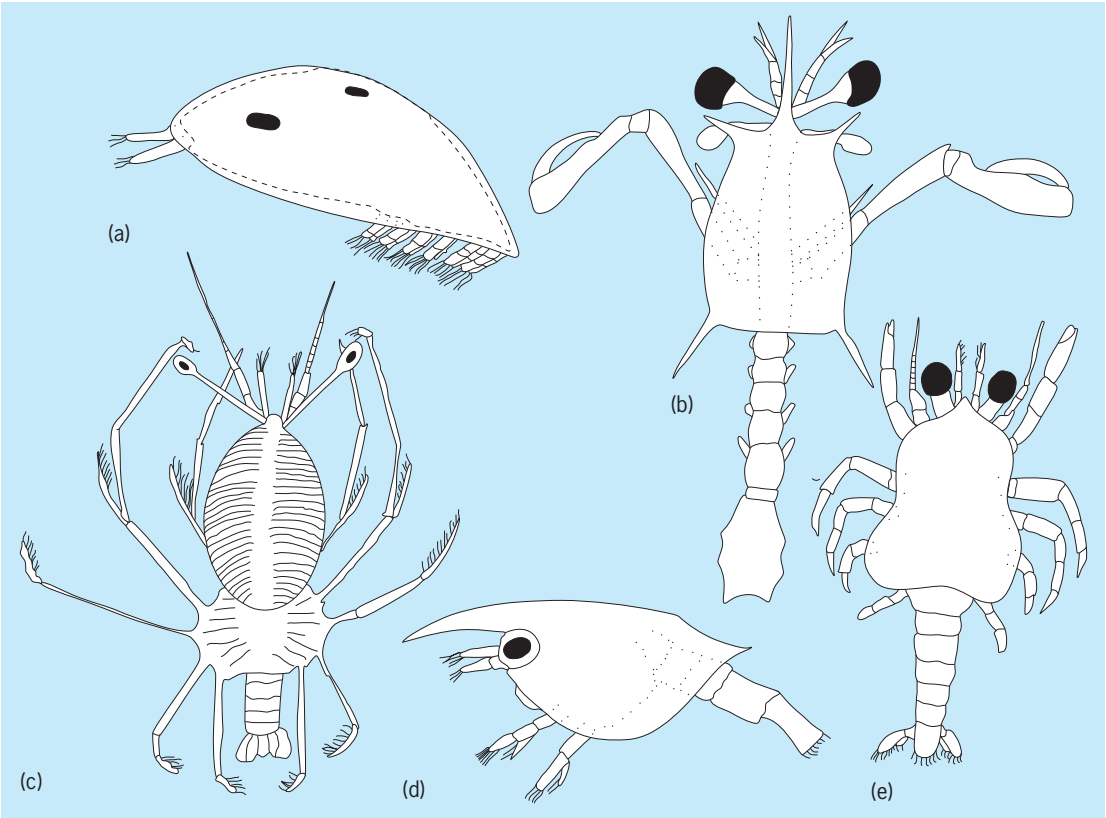

Fig. 6. Larval stages of representative crustaceans: (*a*) cyprid larva of Sessilia; (*b*) pseudolarva of Stomatopoda; (*c*) phyllosoma larva of Decapoda; (*d*) zoeal larva of Decapoda; (*e*) megalopa of Decapoda. (*After P. A. McLaughlin, Comparative Morphology of Recent Crustacea, W. H. Freeman, 1980*).

Physiological Mechanisms

Because of its tough, chitinous, or calcified exoskeleton, a crustacean does not usually grow without molting. Crustaceans have also developed the remarkable feature of autotomy and regeneration to minimize injury and hemorrhage from loss of limb through accident or to an enemy.

Molting (ecdysis). This process involves several steps: (1) preparation, which includes some degree of resorption of the old cuticle; (2) the formation of a new, temporarily soft and thin one within it; (3) the accumulation and storing of calcium in the midgut gland or as lenticular deposits (gastroliths), one of which occurs on either side of the foregut in certain forms such as crayfishes and lobsters. The preparatory period is less complicated in the thinly chitinous forms. The actual molt follows. The old shell or cuticle splits at strategic and predetermined places, permitting the crustacean within, already enclosed in new but still soft exoskeleton, to withdraw. Most likely, a temporary absorption of water enables the animal to split or crack its housing. Upon withdrawal of the entire animal, absorption of water again rapidly takes place with a pronounced increase in body size. The chitinous lining of the foregut and hindgut as well as the endophragmal skeleton are shed along with the exoskeleton. In the immediate postmolt period the crustacean is quite helpless and, if unsheltered, is at the mercy of its enemies. It is in the soft-shelled stage immediately after molting that the East Coast blue crab (*Callinectes*) becomes a much-sought seafood delicacy. The tender new cuticle is reinforced rapidly by the resorbed chitin, and hardened by whatever reserves of calcium the animal may have stored, supplemented and extended by the far more plentiful supplies in solution in the sea, which may be absorbed or ingested by the growing crustacean. Molting takes place quite frequently in the larval stages when growth is rapid but becomes less frequent as the animal ages. In many species there is a terminal molt at maturity. Hormones from the eyestalk play an important role in both initiating and inhibiting molting. *See* GASTROLITH.

Autotomy and regeneration. The mechanisms of autotomy and regeneration are developed in some crustaceans to minimize injury or loss to an enemy. When an appendage is broken, it is cast off or broken at the fracture or breaking plane (preferred breakage point or PBP). In lower crustaceans and even malacostracans such as amphipods and isopods, no preferred breakage point has been found, but its occurrence in brachyuran decapods is well documented. In most species the breaking point is localized at the joint between the basis and ischium, which is usually a nonfunctional articulation. Microcanals spanning the whole depth of the cuticle are more abundant here than in other regions of the appendage. No muscle develops through the breaking point; instead a specialized autotomizer muscle is inserted just proximal to it. The only nerve developing through the breaking point is locally tapered and, thus, weakly resistant. In addition, the blood vessels crossing the breaking point have a valve, preventing bleeding. Autotomy is well documented in some brachyuran decapods. Contraction of the autotomizer muscle pulls the anterior-basal part of the basis into the coxa. The mechanical resistance of the distal margin of the coxa against the anterior surface of the basi-ischium results in breaking of the integument between the basis and ischium; the nerve and blood vessels are then cut.

Crustacea also have the ability to regenerate lost parts. Although the regenerated parts are not always the same size as the original in the first molt after injury, increase in size in successive molts soon restores a lost limb to virtually its former appearance. During the regenerative stages among alpheid shrimps and lobsters, which generally exhibit marked dimorphism in the chelipeds when the major chela is lost, the minor one often may take on more or less the character of the former major chela. Then, the chela replacing the lost one assumes the habitus of the former minor chela during successive molts. *See* REGENERATION (BIOLOGY).

Geologic History

Exceptionally well-preserved arthropods have been discovered in Late Cambrian rocks. These arthropods are referred to as stem crustaceans, a group which is judged to have appeared at least by the Early Cambrian. The stem crustaceans are related to the true crustaceans, to which the former gave rise. Their morphology indicates that the Crustacea is most likely monophyletic, that is, derived from a single common ancestor. The oldest true crustaceans are ostracodes, dating from 570 million years ago (Ma), documented by species in several families. Thus, although the ostracodes are not the most primitive crustaceans, their remains constitute the oldest authentic record of the Crustacea. The more primitive classes (**Fig. 7**) have few hard parts likely to be preserved, and therefore they are far less well known. The Cephalocarida are known only from the Recent, and the Branchiopoda do not appear as fossils until the Early Devonian, about 395 Ma. The Remipedia first appear in the fossil record in the Carboniferous, about 330 Ma. The appearance of more primitive crustaceans later in geologic history, the reverse of what would be expected, is clearly more related to their preservation potential than to their evolutionary history. Furthermore, it is likely that the agnostids, a group of trilobitelike organisms, are more closely allied with the Crustacea than with the Trilobita. The more advanced classes, the Maxillopoda and the Malacostraca, appeared in the fossil record, respectively, in the Middle Cambrian, about 530 Ma, and the Late Devonian, about 370 Ma. *See* CEPHALOCARIDA; MALACOSTRACA.

The isopods and the decapods are among the most commonly known crustaceans. Isopods inhabit marine, freshwater, and terrestrial habitats and are known from the fossil record as far back as the Triassic, 230 Ma. The common pill bug represents one of

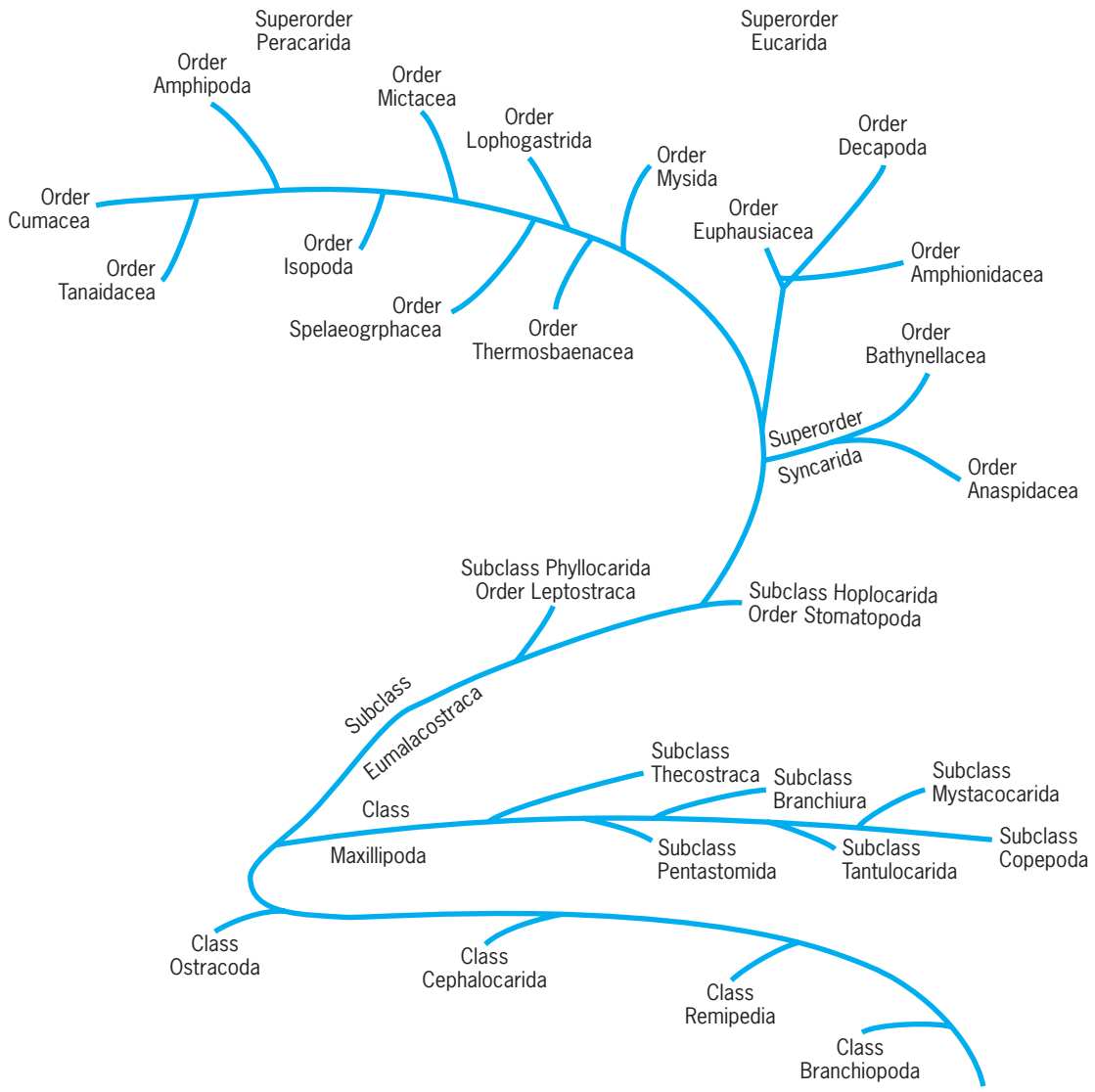

Fig. 7. Chart of the classification of Crustacea; no indication of phylogenetic relationships is implied. (*Modified from D. E. Bliss, ed., The Biology of Crustacea, vols. 1–10, Academic Press, 1982–1985*)

the few totally terrestrial isopods (Fig. 1). The decapod crustaceans, including the modern shrimp, crabs, and lobsters (Fig. 1), trace their beginnings to a lobsterlike form in the Late Devonian, 350 Ma (**Fig. 8**). Thus, the geologic history of the crustaceans is robust and diverse, and the marine environment, in particular, has been home to this group since the early part of the Paleozoic Era. *See* DECAPODA (CRUSTACEA); ISOPODA.

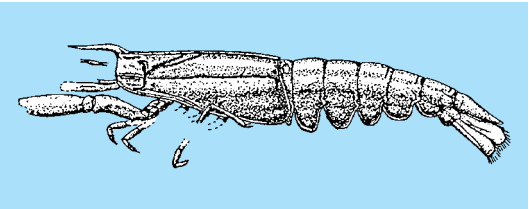

Fig. 8. Late Devonian malacostracan *Palaeopalaemon newberryi* Whitfield, considered to be the earliest decapod crustacean. (*After F. R. Schram, R. M. Feldmann, and M. J. Copeland, J. Paleontol., 52:1381, 1978*)

Waldo L. Schmitt; Patsy A. McLaughlin; Rodney M. Feldmann

Bibliography. H. Balss, *Decapoda and Stomatopoda*, in H. G. Bronn (ed.), *Klassen und Ordnungen des Tierreichs*, vol. 5, 1940–1957; D. E. Bliss (ed.), *The Biology of Crustacea*, vols. 1–10, 1982–1985; J. W. Martin and G. E. Davis, *An Updated Classification of the Recent Crustacea*, Natural History Museum of Los Angeles County, Science Ser. 39, 2001; D. Maruzzo et al., Appendage loss and regeneration in arthropods: A comparative review, in S. Koenemann and R. A. Jenner, Crustacea and Arthropod Relationships, *Crustacean Issues*, vol. 16, pp. 215–245, Francis & Taylor, London, 2005; P. A. McLaughlin et al., *Common and Scientific Names of Aquatic Invertebrates from the United States and Canada: Crustaceans*, American Fisheries Society Spec. Publ. 31, Bethesda, MD, 2005; S. P. Parker

(ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; W. L. Schmitt, *Crustaceans*, 1964; F. R. Schram, *Crustacea*, 1986; F. R. Schram and S. Koenemann, Developmental genetics and arthropod evolution: On body regions of Crustacea, in G. Scholtz, Evolutionary Developmental Biology of Crustacea, *Crustacean Issues*, vol. 15, pp. 75–92, A. A. Balkema, Lisse, 2004.

Crux

The Southern Cross, a small southern constellation. Because of the absence of a star at the center of the cross, this constellation really looks like a kite (see **illustration**). To its side, the Coalsack is a giant dark silhouette of dust against the Milky Way Galaxy. In it the colorful open star cluster known as the Jewel Box is found. It contains dozens of hot, bright, blue stars plus a prominent, central red supergiant. *See* STAR CLUSTERS.

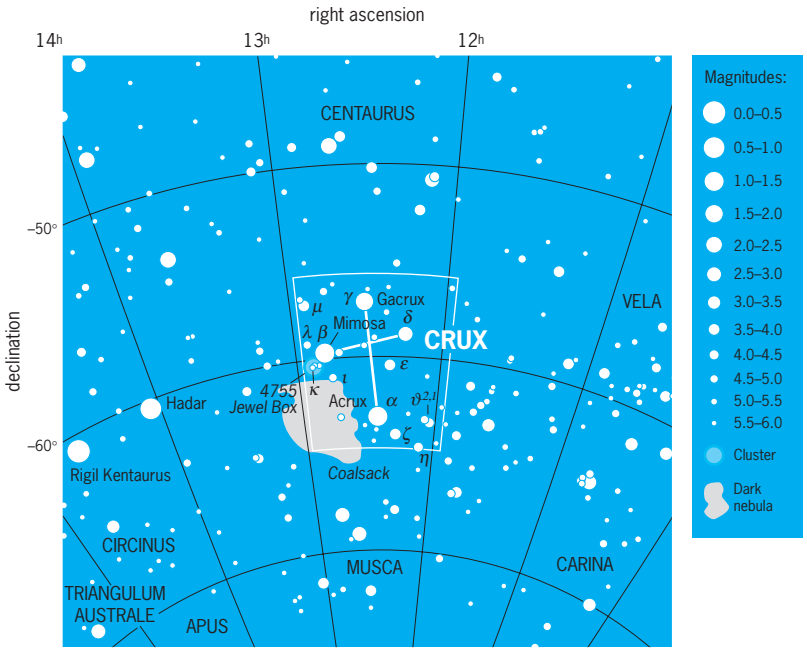

Modern boundaries of the constellation Crux, the Southern Cross. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

The first three stars of the cross, clockwise from the brightest, are Acrux, Mimosa, and Gacrux.

Crux, whose stars were originally part of Centaurus but which were considered separately when they were used for navigation by sixteenth-century explorers, is pictured on the flags of Australia and New Zealand. The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CENTAURUS; CONSTELLATION.

Jay M. Pasachoff

Cryobiology

The use of low-temperature environments in the study of living plants and animals. The principal effects of cold on living tissue are destruction of life and preservation of life at a reduced level of activity. Both of these effects are demonstrated in nature. Death by freezing is a relatively common occurrence in severe winter storms. Among cold-blooded animals winter weather usually results in a comalike sleep that may last for a considerable time. *See* HIBERNATION AND ESTIVATION.

While these natural occurrences suggest the desired ends of cryobiology, activities in this field differ in that much lower temperatures are employed than are present in natural environments. The extreme cold of liquid nitrogen (boiling at −320°F or −196°C) can cause living tissue to be destroyed in a matter of seconds or to be preserved for an indefinite time, certainly for years and possibly for centuries, with essentially no detectable biochemical activity.

The result achieved when heat is withdrawn from living tissue depends on processes occurring in the individual cells. Basic knowledge of the causes of cell death, especially during the process of freezing, and the discovery of methods which circumvent these causes have led to practical applications both for long-term storage of living cells or tissue (cryopreservation) and for calculated and selective destruction of tissue (cryosurgery). *See* CELL SENESCENCE AND DEATH.

Fundamental processes. The biochemical constituents of a cell are either dissolved or suspended in water. During the physical process of freezing, water tends to crystallize in pure form, while the dissolved or suspended materials concentrate in the remaining liquid. This principle has been suggested as one means for producing fresh water from seawater. In the living cell, however, this process is quite destructive. *See* WATER DESALINATION.

In a relatively slow freezing process ice first begins to form in the fluid surrounding the cells, and the concentration of dissolved materials in the remaining liquid increases. The equilibrium between the fluid inside the cell and the fluid surrounding the cell is upset. A concentration gradient is established across the cell wall, and water moves out of the cell in response to the osmotic force. As freezing continues, the cell becomes quite dehydrated. Salts may concentrate to extremely high levels. For example, sodium chloride may reach 200 g/liter (7 oz/qt) concentration in the final stages of freezing. In a similar manner the acid-base ratio of the solution may be altered during the concentration process. Dehydration can affect the gross organization of the cell and also the molecular relationships, some of which depend on the presence of water at particular sites. Cellular collapse resulting from loss of water may bring in contact intracellular components normally separated to prevent destructive interaction. Finally, as the ice crystals grow in size, the cell walls may be ruptured by the ice crystals themselves or by

the high concentration gradients which are imposed upon them.

By speeding the freezing process to the point that temperature drop is measured in degrees per second, some of these destructive events can be modified. For example, ice crystals will form within the cells so that less water will be lost, and the size of each crystal will be minimized, tending to reduce incidence of destruction of the cell walls. However, most of the previously mentioned destructive processes will prevail.

To prevent dehydration, steps must be taken to stop the separation of water in the form of pure ice so that all of the cell fluids can solidify together. The chief tools used to accomplish this are agents that lower the freezing point of the water. Typical materials used for industrial antifreeze agents are alcohols or glycols (polyalcohols). Glycerol, a polyalcohol which is compatible with other biochemical materials in living cells, is frequently used in cell preservation.

Besides the antifreeze additive, refrigeration procedures are designed to control the rate of decline in temperature to the freezing point, through the liquid-solid transition, and below, to very low temperatures. As a result of these procedures, the materials in the cell and in the surrounding fluid solidify together in a state more nearly resembling a glass than a crystalline material. In this form cells can be preserved apparently for as long as one may desire and will regain all their normal functions when properly thawed.

In the case of preservation of sizable aggregates of cells in the form of whole organs, the problems are significantly increased. With a large mass, control of the cooling rate is much more difficult. Additives may not reach cells inside a large organ in sufficient quantity to depress the freezing point adequately. The organ may suffer from oxygen deprivation or from other interruption of its normal environment quite apart from the freezing process itself. While there are signs of progress, whole-organ preservation is not yet a normally successful procedure.

Uses of cryopreservation. The earliest commercial application of cryopreservation was in the storage of animal sperm cells for use in artificial insemination. In the early post–World War II period artificial insemination of dairy herds was practiced widely. However, sperm were viable for only a few days when cooled but not frozen. The development of techniques for sperm freezing, storage, and transport using liquid nitrogen refrigeration, which occurred in the mid-1950s, greatly increased the use of artificial insemination, both for breeding stock and for human reproduction. In the frozen state sperm may be transported great distances. However, of even greater value is the fact that sperm cells can be stored for many years. *See* BREEDING (ANIMAL).

In an entirely similar manner the microorganisms used in cheese production can be frozen, stored, and transported without loss of lactic acid–producing activity. Pollen from various plants can be frozen for storage and transport. Normally this is of little value since fertile seeds can also be conveniently stored and transported without special preservation techniques. However, in plant-breeding experiments it is sometimes desired to crossbreed plants that do not grow in similar environments. In such instances the transport of frozen pollen can be utilized.

In biological research the growth of a pure strain or culture of a particular species is a necessity. For long-term studies these cultures must be continually renewed by growth on a suitable medium. Given enough time, these cultures undergo mutations; therefore, the species characteristics may change during the course of the investigation. By preparing a sufficiently large culture at the start of the investigation and freezing most of it, the investigator is assured of a continuous supply of unaltered microorganisms for his study no matter how long it may take. Banks of carefully established cell lines are stored in frozen condition for use by future generations of biologists.

Another application of cryopreservation is the storage of whole blood or separated blood cells. Frozen storage permits the buildup of the large stocks of blood required in a major catastrophe. It is also useful in storing the less common types of blood. Through frozen blood storage a person can actually develop a stock of their own blood for autotransfusion, a practice which is almost essential for those individuals with rare blood types. *See* BLOOD.

Freezing and preservation equipment. While other refrigerant materials or mechanical refrigeration systems can be used, most cryobiology apparatus has been designed to use liquid nitrogen as the refrigerant because it is convenient, biologically inert, relatively cheap, and plentiful. *See* NITROGEN.

When a controlled-rate freezer is used, the material to be frozen is placed in a well-insulated chamber. Liquid nitrogen is admitted to another portion of the

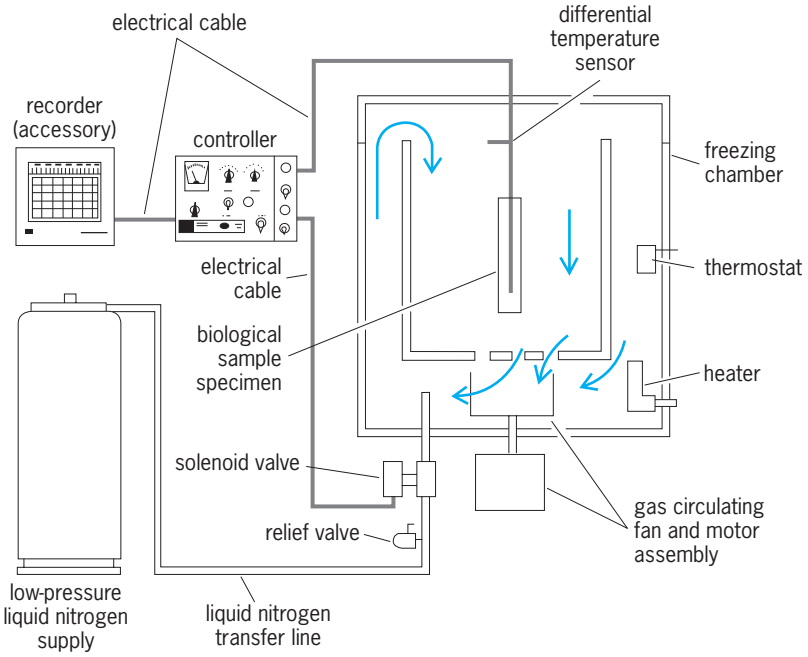

Fig. 1. Schematic representation of freezing system operation.

chamber at a rate controlled by an electrically operated valve. The liquid nitrogen is immediately vaporized by the air that is circulated through the chamber by a motor-driven fan. A thermocouple inserted in a sample of the material to be frozen is connected to the controller. Using the temperature of the thermocouple as a guide, the controller adjusts the flow of liquid nitrogen and the output of an electrical heater to obtain just the cooling rate desired. This procedure is outlined in **Fig. 1**. Working storage of sperm for artificial insemination or for small laboratories employs specially designed liquid nitrogen containers as illustrated in **Fig. 2**. A large opening (2.5-in. or 6.3-cm neck-tube diameter in diagram) is provided so that materials can be inserted and removed. Since the container is designed to allow very little leakage of heat into the inner container, one charge of liquid nitrogen can last for 60 days in active use. The frozen material is normally stored in small glass ampules. Six or eight ampules are clipped onto a vertical rod, and ten rods are mounted in a canister 2 in. (5 cm) in diameter. A typical storage container holds six canisters, or 360–480 ampules. A porous plug fits into the neck tube to hold the canister handles in place and to reduce the leakage of heat into the container.

While containers of the type just mentioned can be used to ship frozen biologicals by all normal means of transport, there are materials that would not conveniently fit into a container with such a small opening. Furthermore, some types of shipment require that the container be able to assume any position during transport. For these applications the equipment illustrated in **Fig. 3** has been developed. The entire top of this equipment is removable, leaving an opening as large as 14 in. (35 cm) in diameter. The liquid nitrogen refrigerant is absorbed in disks of porous material. The porous material can easily be shaped to conform to the contours of the object to be stored. Because the refrigerant liquid is absorbed in the porous packing disks, the container can be shipped in any attitude, although somewhat better results are obtained in the upright position.

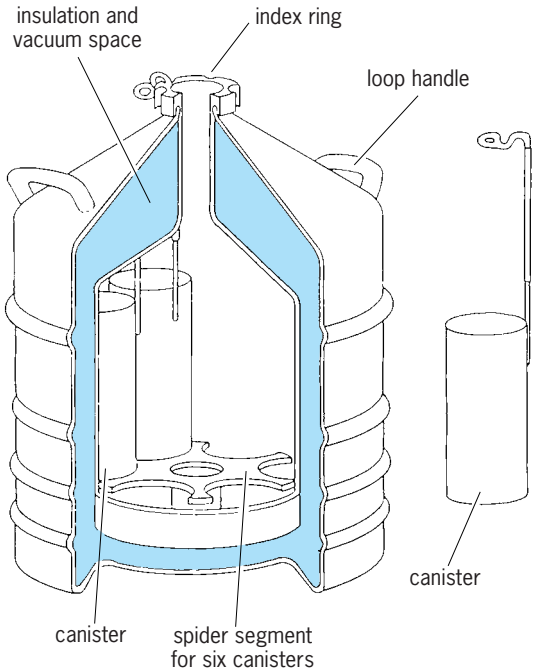

Specifications:

height = 23 3/8 in.
diameter = 17 3/4 in.
neck tube diameter = 2 1/2 in.
ampule capacity = 360 (1.2-cc ampules, 6/rack)
= 484 (1.0-cc ampules, 8/rack)
maximum liquid nitrogen capacity = 30 liters
working holding time = 60 days

Fig. 2. Liquid nitrogen container. 1 in. = 2.5 cm. 1 liter = 1.06 qt.

There is a container made for the stationary storage of large amounts of frozen material equipment. While this unit has much the same general appearance as a chest-type frozen food storage refrigerator, it is far superior in insulation performance. The unit can be left untended for a week without loss of effectiveness. The capacity of a large refrigerator is about 90,000 ampules (250 times the capacity of the portable container in Fig. 2), and several hundred units of frozen blood can be stored in it. Each item to be stored is mounted in a vertical rack. Stored material is retrieved by lifting the rack high enough to reach the desired package and then returning it to position. The brief exposure to ambient temperatures does not warm the other packages more than 1 or 2 degrees above their storage temperature of −320°F (−196°C).

Cryosurgery. As was previously mentioned, the causes of cellular destruction from freezing can be effected deliberately to destroy tissue as a surgical procedure. A technique for controlled destruction of tissue in the basal ganglia of the brain was developed for relief from tremor and rigidity in Parkinson's disease.

One significant advantage of cryosurgery is that the apparatus can be employed to cool the tissue to the extent that the normal or the aberrant function is suppressed; yet at this stage the procedure can be reversed without permanent effect. When the surgeon is completely satisfied that the exact spot to be destroyed has been located, the temperature can be lowered enough to produce irreversible destruction. This procedure is of particular assistance in neurosurgery.

A second major advantage of cryosurgery is that the advancing front of reduced temperatures tends to cause the removal of blood and the constriction of blood vessels in the affected area. This means that little or no bleeding results from cryosurgical procedures.

A third major advantage of cryosurgery is that the equipment employs a freezing apparatus (which is about the size of a large knitting needle) that can be placed in contact with the area to be destroyed with a minimum incision to expose the affected area.

The equipment used in cryosurgery utilizes the same operating principles as the controlled-rate freezer. The flow of liquid nitrogen is controlled by an electrically operated valve whose function is directed by a controller that is based on the temperature sensed by a thermocouple located at the

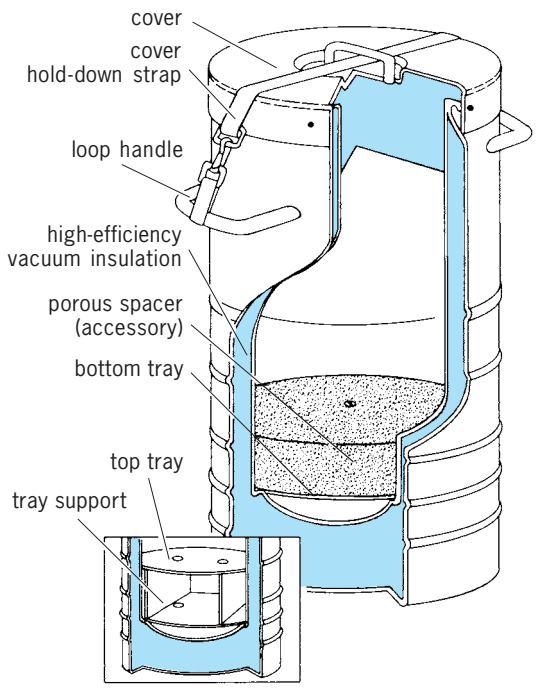

Specifications:
height = 32 in.
diameter = 17 3/4 in.
neck tube diameter = 14 in.
ampule capacity = 5400 (1.2-cc ampules, 6/rack)
= 7200 (1.0-cc ampules, 8/rack)
working liquid
nitrogen capacity = 40.5 liters
working holding time = 19 days
bulk storage capacity = 1.6 ft
LD-40 used with accessory porous spacers
spacer thickness = 3 in.
spacer diameter = 13 1/2 in.
liquid nitrogen capacity
per spacer = 6.3 liters

Fig. 3. Liquid nitrogen container which is used for shipment of large samples. 1 in. = 2.5 cm. 1 ft = 0.3 m. 1 liter = 1.06 qt.

freezing element. Since it is desirable to freeze only the area immediately surrounding the tip of the instrument, the liquid nitrogen lead, the gaseous nitrogen return, the thermocouple leads, and the electrical supply for a small heater are all encased in a vacuum-insulated jacket.

Other uses. Besides the preservation or destruction of life, low temperatures can be employed for other purposes. For example, it may be desirable in the formulation of drugs to release antibodies or other active agents from their cells. One method by which this can be achieved is by grinding the cells while they are frozen. In frozen form the cells are hard and brittle and can be fractured easily without resulting in any chemical changes in the enzyme or antibody that is being released. A. W. Francis

Bibliography. *CIBA Foundation Symposium 52: The Freezing of Mammalian Embryos*, 1978; P. Dowzou, *Cryobiochemistry: An Introduction*, 1977; L. K. Lozina-Lozinskii, *Studies in Cryobiology*, 1974; J. E. Rash and C. S. Hudson, *Freeze-Fracture: Methods, Artifacts, and Interpretations*, 1979.

Cryogenics

The science and technology of phenomena and processes at low temperatures, defined arbitrarily as below 150 K (−190°F). Phenomena that occur at cryogenic temperatures include liquefaction and solidification of ambient gases (**Table 1**); loss of ductility and embrittlement of some structural materials such as carbon steel; increase in the thermal conductivity to a maximum value, followed by a decrease as the temperature is lowered further, of relatively pure metals, ionic compounds, and crystalline dielectrics (diamond, sapphire, solidified gases, and so forth); decrease in the thermal conductivity of metal alloys and plastics; decrease in the electrical resistance of relatively pure metals; decrease in the heat capacity of solids; decrease in thermal noise and disorder of matter; and appearance of quantum effects such as superconductivity and superfluidity. *See* ELECTRICAL RESISTIVITY; SPECIFIC HEAT OF SOLIDS; SUPERCONDUCTIVITY; SUPERFLUIDITY; THERMAL CONDUCTION IN SOLIDS.

Production of low temperatures. Cryogenic environments are maintained with cryogens (liquefied gases) or with cryogenic refrigerators. The temperature afforded by a cryogen (Table 1) ranges from its triple point to slightly below its critical point. Commonly used cryogens are liquid helium-4 (down to 1 K), liquid hydrogen, and liquid nitrogen. Less commonly used because of their expense are liquid helium-3 (down to 0.3 K) and neon. The pressure maintained over a particular cryogen controls its temperature. Heat input—both the thermal load and the heat leak due to imperfect insulation—boils away the cryogen, which must be replenished. *See* LIQUID HELIUM; THERMODYNAMIC PROCESSES; TRIPLE POINT.

A variety of techniques are available for prolonged refrigeration if such is favored by economic considerations or if very low temperatures are required. Down to about 1.5 K, refrigeration cycles involve

TABLE 1. Physical constants of low-temperature liquids

Liquid	Triple point Temp., K	Triple point Pressure atm†	Normal boiling point, K	Critical point Temp., K	Critical point Pressure, atm†
Helium-3			3.2	3.3	1.2
Helium-4			4.2	5.2	2.3
n-Hydrogen	14.0	0.07	20.4	33.2	13.0
n-Deuterium	18.7	0.17	23.7	38.4	16.4
n-Tritium	20.6	0.21	25.0	42.5*	19.4*
Neon	24.6	0.43	27.1	44.7	26.9
Fluorine	53.5	< 0.01	85.0	144.0	55.0
Oxygen	54.4	< 0.01	90.2	154.8	50.1
Nitrogen	63.2	0.12	77.4	126.1	33.5
Carbon monoxide	68.1	0.15	81.6	132.9	34.5
Argon	83.8	0.68	87.3	150.7	48.0
Propane	85.5	< 0.01	231.1	370.0	42.0
Propene	87.9	< 0.01	226.1	365.0	45.6
Ethane	89.9	< 0.01	184.5	305.4	48.2
Methane	90.7	0.12	111.7	191.1	45.8

*Calculated.
†1 atm = 1.013×10^5 pascals.

compression and expansion of appropriately chosen gases. At lower temperatures, liquid and solids serve as refrigerants. Adiabatic demagnetization of paramagnetic ions in solid salts is used in magnetic refrigerators to provide temperatures from around 4 K down to 0.003 K. Nuclear spin demagnetization of copper can achieve 5×10^{-8} K. Helium-3/helium-4 dilution refrigerators are frequently used for cooling at temperatures between 0.3 and 0.002 K, and adiabatic compression of helium-3 (Pomeranchuk cooling) can create temperatures down to 0.001 K. *See* ADIABATIC DEMAGNETIZATION.

Cryogenic refrigeration principles. In principle, any substance that undergoes an isentropic temperature change (meaning at constant entropy, without heat transfer or frictional-type losses) with a change in field intensity (pressure, stress, magnetic field intensity, or electric field intensity) can be used as a refrigerant. In practice, gases (for example, helium, hydrogen, and nitrogen) and magnetic substances (paramagnetic materials such as gadolinium sulfate or gadolinium gallium garnet) are usually employed.

An ideal refrigeration cycle, which has the maximum possible thermodynamic efficiency, is the Carnot cycle. This cycle can be illustrated with a temperature-entropy (*T-S*) diagram (**Fig. 1**), on which the state of the substance traces a rectangular path in a counterclockwise direction. Starting with the state of greatest temperature and entropy, the corners of the rectangle may be labeled *A, B, C*, and *D*. The *T-S* diagram is useful for understanding the thermodynamics of refrigeration cycles because the refrigeration (heat removed) is the area under the base of the rectangle (area *C-D-D′-C′*) and the net work input required is the area enclosed by the rectangle (area *A-B-C-D*). The magnetic substance (or gas) is magnetized (or compressed) isothermally from *A* to *B*; demagnetized (or expanded) isentropically from *B* to *C*, with cooling from temperature T_B to T_C; demagnetized (or expanded) further (*C* to *D*) at constant temperature with heat *Q* absorbed from the load given by Eq. (1); and finally magnetized (or compressed) back to the original state (*D* to *A*). The net work required to run the Carnot cycle is given by Eq. (2); therefore, the ratio of work input to refrigeration produced is given by Eq. (3). If T_A is the ambient temperature, Eq. (2) gives the ideal, or minimum, work required to supply refrigeration at temperature T_C. A useful, common measure of an actual cryogenic refrigerator's efficiency compares the ideal (Carnot) work to the actual work of refrigeration to give a so-called percent-of-Carnot, Eq. (4).

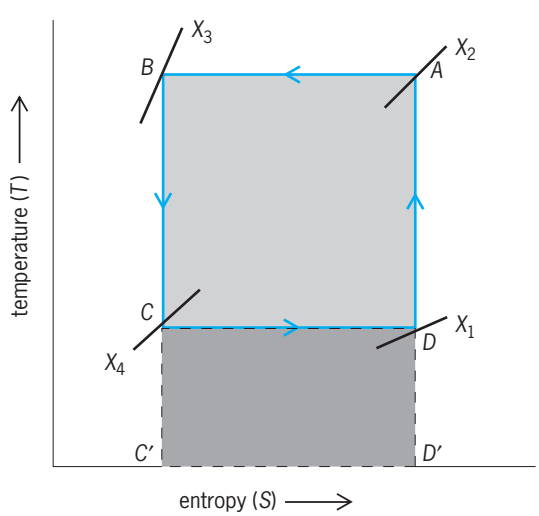

Fig. 1. Ideal Carnot refrigeration cycle on a temperature-entropy (*T-S*) diagram. The lines X_1, X_2, X_3, and X_4 represent lines of constant magnetization for a magnetic substance or constant pressure for a gas.

$$Q = T_C(S_D - S_C) \tag{1}$$

$$W_C = (T_B - T_C) \cdot (S_D - S_C) \tag{2}$$

$$\frac{W_C}{Q} = \frac{(T_B - T_C)}{T_C} \tag{3}$$

$$\eta_c = \frac{W_c}{W_{\text{actual}}} \tag{4}$$

See CARNOT CYCLE; ENTROPY; THERMODYNAMIC CYCLE.

Approximations to the Carnot cycle are actually used in magnetic refrigerators, with the upper temperature (T_B) ranging from 15 to 0.3 K, depending on the heat sink used (cryocooler, helium-4 bath, or helium-3 bath). Percent-of-Carnot efficiencies range up to about 50%. For gas refrigeration machines, however, the Carnot cycle is not practical because of heat transfer between the gas and the cylinder walls and the excessive pressures required to reach low temperature.

Two ideal thermodynamic cycles with practical application to gas refrigeration machines, however, are the reverse Stirling cycle and the reverse Brayton cycle with isothermal expander. These cycles may again be represented as counterclockwise paths on a temperature-entropy diagram (**Fig. 2**), labeled in the same way as before. Gas is isothermally compressed (*A-B*); cooled (*B-C*) in counterflow heat exchange by the return flow (*D-A*) in a thermal regenerator at constant volume (Stirling cycle) or counterflow heat exchanger at constant pressure (Brayton cycle); expanded isothermally (*C-D*), with the absorption of heat again given by Eq. (1); and then warmed at constant volume (Stirling cycle) or constant pressure (Brayton cycle) by counterflow heat exchange (*D-A*) with the incoming high-pressure (density) stream (*B-C*). The counterflow heat exchanger or regenerator is the fundamental element for attaining cryogenic temperatures with gas refrigerating machines, since it allows bootstrapping of the incoming flow to low temperature by heat exchange with the outgoing flow. Modest pressure ratios are possible because cooling by gas expansion is required for the thermal load (and irreversibilities in an actual machine) only, but not for initial attainment of the low temperature. *See* BRAYTON CYCLE; STIRLING ENGINE.

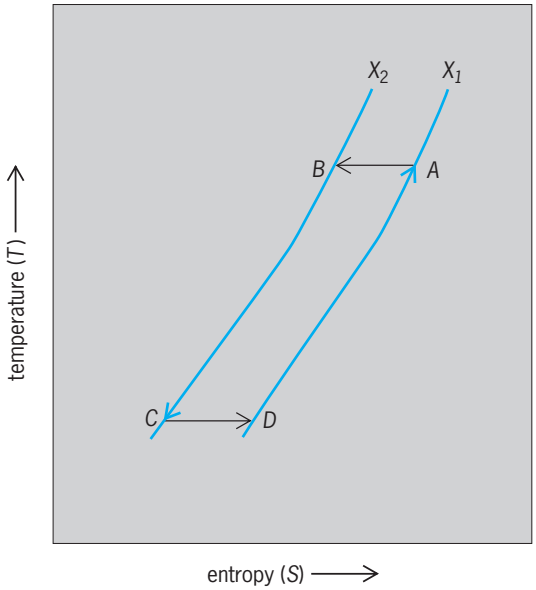

Fig. 2. Ideal practical gas refrigeration cycle on a temperature-entropy (*T-S*) diagram. The curves X_1 and X_2 are lines of constant volume in the reverse Stirling cycle and lines of constant pressure in the reverse Brayton cycle with isothermal expander.

Most modern cryogenic refrigerations are based on the reverse Stirling and reverse Brayton cycles. Reverse Stirling, Gifford-McMahon (G-M), and pulse-tube cryocoolers use thermal regenerators and expand the gas in the low-pressure part of the cycle with pistons or displacers (a gas piston for the pulse tube). Ordinarily, these machines are used for lower refrigeration loads (a fraction of a watt to a few hundred watts). Their lowest useful temperature is about 10 K because at lower temperatures the regenerator effectiveness becomes poor as a result of the rapidly diminishing heat capacity of the regenerator matrix. Experimental coolers reach as low as 4 K by using special high-heat-capacity regenerator materials. The reverse Brayton and the related Claude cycle use reciprocating piston expanders or, especially at higher capacities, turbo expanders. Temperatures as low as 1.8 K are typical for these machines. The percent-of-Carnot efficiency varies from as high as 40% for machines with refrigeration capacities in excess of 10 kW, to less than 1% for some fractional-watt machines. Higher refrigeration temperatures generally give slightly better efficiencies than lower refrigeration temperatures. Two refrigerators connected in series (for example, a magnetic refrigerator rejecting heat to a Stirling cooler) have an overall efficiency approximately equal to the product of the individual refrigerator efficiencies.

Gas liquefaction principles. Both the latent heat of vaporization and the sensible heat of the gas (heat content of the gas) must be removed to liquefy a gas. Of the total heat that must be removed to liquefy the gas, the latent heat is only 1.3% for helium and 46% for nitrogen. Consequently, an efficient liquefier must supply refrigeration over the entire temperature range between ambient and the liquefaction point, not just at the liquefaction temperature (**Table 2**).

The Collins-Claude refrigeration cycle (**Fig. 3**) forms the basis (with a multitude of variations) of most modern cryogenic liquefiers. Gas is compressed isothermally (1-2) and cooled (2-5) in a counterflow heat exchanger by the colder return stream of low-pressure gas (4′-1). During this cooling, a fraction of the high-pressure stream (equal to the rate of liquefaction) is split off (at points 2, 3, and 4) and cooled by the removal of work (energy) in expansion engines or turbines (which expand it to 2′, 3′, and 4′). This arrangement provides the cooling for the removal of the sensible heat. At the end of the counterflow cooling (point 5), the remaining high-pressure stream is expanded in either a Joule-Thomson valve or a wet expander to give the liquid product (point 6) and the return stream of saturated vapor (point 4′). *See* LIQUEFACTION OF GASES.

The work input required to produce refrigeration is commonly given in terms of watts of input power per watt of cooling, that is, W/W (Table 2). Cooling with a refrigerator is more efficient (that is, requires a lower W/W) than cooling with evaporating liquid supplied from a Dewar because the refrigerator does not discard the cooling available in the boil-off gas. *See* REFRIGERATION; REFRIGERATION CYCLE; THERMODYNAMIC CYCLE.

Storage and handling of cryogens. When gases are cooled from ambient to cryogenic temperatures, and liquefied, they undergo large increases in density, ranging from a 648-fold increase for methane to an 867-fold increase for hydrogen. Such an increase

TABLE 2. Ideal (reversible) power requirements for cryogenic liquefaction and refrigeration

Fluid	Temp., K	Liquefaction, W·h/liter*	Evaporating liquid refrigeration, W/W†	Ideal refrigerator, W/W‡
Helium	4.2	236	326	70.4
Hydrogen	20.4	278	31.7	13.7
Neon	27.1	447	15.5	10.1
Nitrogen	77.4	173	3.87	2.88
Fluorine	85	238	3.26	2.53
Argon	87.3	185	2.95	2.44
Oxygen	90.2	195	2.89	2.33
Methane	111.5	129	2.15	1.69

*Ideal (minimum) work required to change gas at ambient temperature and pressure to liquid.
†Ideal work to provide refrigeration at the cryogenic normal boiling point (column 2) by evaporating liquid without recovering the cooling available in the boil-off gas. This value is obtained by dividing the ideal work of liquefaction by the fluid's heat of vaporization.
‡Ideal work to provide refrigeration at the cryogenic normal boiling point with an ideal refrigerator such as Carnot or Stirling cycle machines.

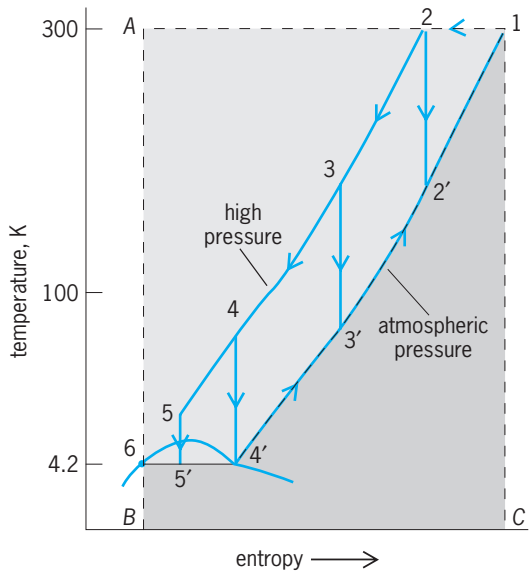

Fig. 3. Nearly ideal Collins-Claude gas liquefication cycle on a temperature-entropy (*T*-*S*) diagram. Area 1-4′-6-*B*-*C*−1 represents heat removed to liquefy the gas; area 1-*A*-6-4′−1 represents ideal (minimum) work required.

often makes it attractive to transport helium, hydrogen, oxygen, nitrogen, and natural gas as liquids, even when the end use requires a gas. The storage capacity of a trailer or rail car is about 10 times greater for liquid storage than gas storage.

Storage. Storage of cryogens requires special insulated containers. For helium and hydrogen, whose low heats of vaporization make them the most difficult to store, Dewars with evacuated multilayer insulation and boil-off gas-cooled radiation shields are used. Vacuums of 10^{-5} torr (10^{-3} pascal) or better virtually eliminate conduction heat ingress, while multilayer insulation, consisting of 100 or more layers of 0.25-mil (6-micrometer) aluminum-coated polyester film, and gas-cooled radiation shields greatly reduce radiation heat leak. Typical evaporative losses range from about 2% per day for 25-gallon (100-liter) liquid-helium Dewars to a fraction of a percent for large liquid-helium Dewars. For higher-boiling-point cryogens, with correspondingly higher heats of vaporization, such as liquid oxygen, liquid nitrogen, and liquefied natural gas (LNG), smaller containers use evacuated multilayer insulation, whereas larger vessels use evacuated perlite (silica powder) insulation or, for very large LNG tanks, unevacuated silica powder or foamed plastic. *See* DEWAR FLASK; HEAT INSULATION.

Transfer and metering. Transfer of cryogens requires pipes with insulations varying from evacuated multilayer insulation for liquid-helium service to foam insulation for large LNG lines. Transfer is by pressure, reciprocating pumps, or centrifugal pumps. For superfluid helium, the thermomechanical pump, which uses the unique ability of superfluid to flow without friction through porous plugs with micrometer size-channels, is also available. *See* PUMP.

The large economic value of cryogens, as well as the requirements of process control, make accurate metering important. Many standard meters such as orifice, venturi, turbine, positive-displacement, and coriolis are used, although two-phase flow and cavitation can reduce both accuracy and reliability. Liquid-level measurement in calibrated tanks is also used, especially for smaller vessels. *See* FLOW MEASUREMENT; LEVEL MEASUREMENT.

Applications. Cryogenic technology has manifold applications, including liquefaction of oxygen, nitrogen, and natural gas; transportation; aerospace; medicine; physics research; and electric power systems.

Oxygen and nitrogen. Every year the chemical, steel, and medical industries use millions of tons of oxygen and nitrogen. Although the end use is usually in the form of gas for breathing oxygen and applications such as basic oxygen steel making and production of ammonia and methanol, delivery is frequently in liquid form from cryogenic air-separation plants. *See* AIR SEPARATION; NITROGEN; OXYGEN.

Liquefied natural gas. Natural gas (methane with depending on the source, varying percentages of ethane, propane, butane, and nitrogen) is liquefied and shipped in insulated tankers with capacities up to 860,000 bbl (135,000 m^3). With the refrigeration power of some LNG plants exceeding 450 MW, these are the world's largest refrigeration systems. Storage vessels up to 1.6×10^6 bbl (250,000 m^3) receive LNG at the unloading sites.

Utility companies use LNG for either base load (continuous supply) or peak shaving (peak requirements). At some smaller peak-shaving plants, LNG is produced and stored during low-usage periods and returned as natural gas to the pipeline system during peak-usage periods. Some base-load installations contain air-separation facilities, so that refrigeration available from gasification of LNG can be applied to air liquefaction. *See* LIQUEFIED NATURAL GAS (LNG).

Transportation. LNG has considerable potential as a vehicle fuel because of its cost, efficient combustion, and especially low emissions. Compressed natural gas–powered vehicles are becoming commonplace. Liquid hydrogen has been demonstrated as both an aircraft and vehicle fuel, having the advantages of cleanliness and absence of carbon dioxide emission associated with the greenhouse effect. Prototype magnetically levitated trains, supported by superconducting magnets, are under evaluation. *See* AIRCRAFT FUEL; MAGNETIC LEVITATION.

Aerospace. As exemplified by the space shuttle, liquid hydrogen and liquid oxygen are the principal propellants for space flight because they have the highest specific impulse of any common fuel-oxidizer combination. Slush hydrogen, a mixture of liquid and solid hydrogen which coexist at 13.8 K (−434.8°F), has been studied extensively as a fuel for space probes and hypersonic aircraft-space vehicles. It offers a 16% greater density than normal-boiling-point hydrogen, as well as a greater capacity for cooling engine and aerodynamic surfaces. Supercritical hydrogen and oxygen are used on space vehicles to power fuel cells and to provide breathing oxygen.

Liquid oxygen, stowed in lightweight, well-insulated containers, is used in military aircraft for breathing oxygen and offers significant weight and space advantages over pressurized gas storage. *See* FUEL CELL; ROCKET PROPULSION.

Cryogenic wind tunnels, injecting liquid nitrogen to cool the flow, make possible matching of both Reynolds number and Mach number on aerodynamic models in the transonic region. Independent variation of the Reynolds and Mach numbers and lower power consumption are further advantages over conventional wind tunnels. *See* WIND TUNNEL.

Cooling infrared detectors in space satellites has been the impetus for extensive development of cryocoolers (small cryogenic refrigerators) which operate unattended for up to 5 years and supply refrigeration at 80 K (−316°F) and below. Considerable progress has been made, mostly with Stirling-cycle coolers, in attaining that required operating life. Superfluid helium at about 1.8 K is used as a coolant for infrared telescopes in astronomical satellites, and as a refrigerant in satellite experiments.

Medicine. Using liquid nitrogen, cryopreservation of human blood and bone and of both cattle and human sperm for artificial insemination is commonplace. Cryosurgery, involving destruction of tissue by freezing, is used to treat Parkinson's disease, cancers, and tumors (for example, pulmonary metastases, oral squamous cell carcinoma, malignant melanoma, and cancers of the brain, stomach and prostate). Routine cryotherapy removes warts without scarring. *See* CRYOBIOLOGY.

Magnetic resonance imaging (MRI) uses a large-bore superconducting magnet, cooled by liquid helium, to produce medical images of the body similar to x-ray tomography, but with better resolution of soft, primarily hydrogenic body tissue. Clinical use of these devices is widespread. *See* MEDICAL IMAGING.

Physics research. Liquid helium–cooled superconducting magnets, operating at 4 K (−452°F), are frequently used in high-energy particle accelerators for beam bending and focusing. Radio-frequency, cavities, cooled to about 2 K with superfluid helium, are used in electron-beam accelerators. Tokamak fusion reactors employ large, liquid helium–cooled magnets for confinement of the plasma. They are fueled by injecting solid deuterium pellets at speeds up to 1.2 mi/s (2 km/s) or more. These projects are quite large and employ some of the world's largest helium refrigerators and liquefiers. *See* NUCLEAR FUSION; PARTICLE ACCELERATOR.

Low-temperature physics focuses on superfluidity, superconductivity and, at temperatures below 1 millikelvin, the investigation of matter under conditions where thermal disorder is small. *See* LOW-TEMPERATURE PHYSICS.

Electric power systems. There have been a number of feasibility demonstrations of superconducting power systems cooled by liquid and supercritical helium. These include superconducting power transmission lines, superconducting generators and motors, low-loss superconducting transformers, and superconducting magnetic energy storage. *See* ENERGY STORAGE; SUPERCONDUCTING DEVICES.

Other applications. Additional applications of cryogenic technology include cryopumps for trapping gases on refrigerated surfaces or molecular sieves in vacuum systems, high-field-gradient magnetic separation and purification of materials, food preservation, hydrogen isotope separation using low-temperature distillation, metal treatment, and scrap processing. *See* FOOD PRESERVATION; ISOTOPE SEPARATION; VACUUM PUMP.

Superconductivity use. A number of the cryogenic applications discussed above use superconducting devices, especially magnets. Cooling these devices is a major area of cryogenic technology. The discovery of high-temperature superconductors, some with transition temperatures over 150 K (−190°F), shows promise of making economically feasible a number of applications that are currently uneconomical with low-temperature, liquid helium–cooled superconductors. *See* MAGNET. D. E. Daney

Bibliography. *Advances in Cryogenic Engineering*, semiannually; A. Bejan, *Advanced Engineering Thermodynamics*, 2d ed., 1997; T. M. Flynn and K. D. Timmerhaus, *Cryogenic Process Engineering*, 1989; S. W. Van Sciver, *Helium Cryogenics*, 1986; G. Walker, *Cryocoolers*, vols. 1 and 2, 1983; J. Wilks and D. S. Betts, *An Introduction to Liquid Helium*, 2d ed., 1987, paper 1989.

Cryolite

A mineral with chemical composition Na_3AlF_6. Although it crystallizes in the monoclinic system, cryolite has three directions of parting at nearly 90°, giving it a pseudocubic aspect. Hardness is $2^1/_2$ on Mohs scale and the specific gravity is 2.95. Crystals are usually snow-white but may be colorless and more rarely brownish, reddish, or even black (see **illus.**). The mean refraction index is 1.338, approximately that of water, and thus fragments become invisible

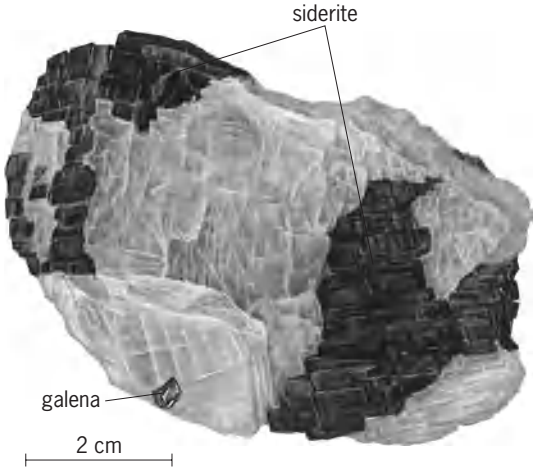

White translucent cryolite in association with siderite and galena, Ivigtut, Greenland. (*Specimen courtesy of Department of Geology, Bryn Mawr College*)

when immersed in water. Cryolite, associated with siderite, galena, and chalcopyrite, was discovered at Ivigtut, Greenland, in 1794. This locality remains the only important occurrence.

Cryolite was once used as a source of metallic sodium and aluminum, but now is used chiefly as a flux in the electrolytic process in the production of aluminum from bauxite. *See* ALUMINUM; BAUXITE.

Cornelius S. Hurlbut, Jr.

Cryotron

A current-controlled switching device based on superconductivity for use primarily in computer circuits. The early version has been superseded by the tunneling cryotron, which consists basically of a Josephson junction. In its simplest form (see **illus.**)

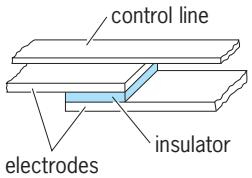

Diagram of the tunneling cryotron.

the device has two electrodes of a superconducting material (for example, lead) which are separated by an insulating film only about 10 atomic layers thick. For the electrodes to become superconducting, the device has to be cooled to a few degrees above absolute zero. The tunneling cryotron has two states, characterized by the presence or absence of an electrical resistance. They can be considered as the "on" and "off" states of the switch, respectively. Switching from on to off is accomplished by a magnetic field generated by sending a current through the control line on top of the junction. The device can switch in a few picoseconds and has a power consumption of only some microwatts. These properties make it an attractive switching device for computers, promising performance levels probably unattainable with other devices. *See* JOSEPHSON EFFECT; SUPERCONDUCTING DEVICES; SUPERCONDUCTIVITY.

P. Wolf

Bibliography. W. Anacker, Computing at 4 degrees Kelvin, *IEEE Spectrum*, 16:26–37, May 1979; W. Krech (ed.), *Superconductivity and Cryoelectronics*, 1991.

Cryptobiosis

A state in which the metabolic rate of an organism is reduced to an imperceptible level. The several kinds of cryptobiosis (hidden life) include anhydrobiosis (life without water), cryobiosis (life at low temperatures), and anoxybiosis (life without oxygen). Of these, most is known about anhydrobiosis; therefore, the discussion will be restricted to that type.

States of anhydrobiosis occur only in the early developmental stages of various organisms, including plant seeds, bacterial and fungal spores, cysts of certain crustaceans, and certain insect larvae. They occur in both developmental and adult stages of certain soil-dwelling invertebrates (rotifers, tardigrades, and nematodes), mosses, lichens, and certain ferns.

Induction of anhydrobiosis. Many organisms that are capable of entering an anhydrobiotic state require a period of induction of the morphological and biochemical alterations. For example, rotifers, tardigrades, and nematodes include species that live in the water associated with soil. When this water dries up, the animals dry up also, but are not killed by the dehydration. When they come in contact with water again, they rapidly swell and assume active life, often within minutes. In order to enter this remarkable state the animals require slow dehydration, during which they undergo pronounced morphological changes. Tardigrades and rotifers contract longitudinally, assuming a so-called tun configuration (**Fig. 1**). Nematodes typically coil into tight spirals (**Fig. 2**). In all these animals, the morphological changes characteristic of cryptobiosis are thought to reduce the rate of evaporative water loss by reducing the animal's surface area accessible to the surrounding

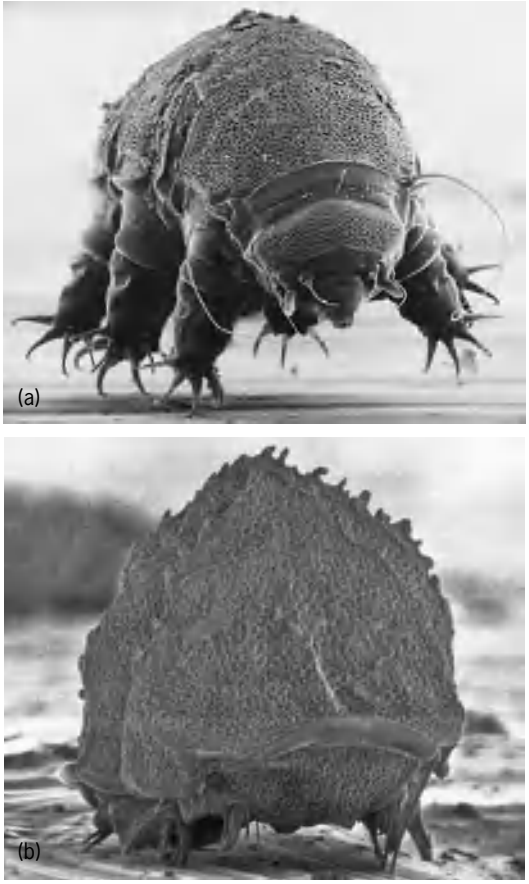

Fig. 1. Scanning electron micrographs of tardigrades: (*a*) active and (*b*) anhydrobiotic. (*From J. H. Crowe and A. F. Cooper, Jr., Cryptobiosis, Sci. Amer., 225:30–36, 1971*)

Fig. 2. Scanning electron micrograph of anhydrobiotic nematode. (*From J. H. Crowe and K. A. C Maden, Anhydrobiosis in nematodes: Evaporative water loss and survival, J. Exp. Zool., 193:323–334, 1975*)

air. In addition, contraction of the body is thought to entail ordered packing of organ systems and intracellular contents, thereby obviating mechanical damage.

Electron microscopy studies on dry organisms lend support to these suppositions. During the slow dehydration the animals show changes in their chemical composition. The nematode *Aphelenchus avenae*, for which there is the most information, synthesizes two carbohydrates, free glycerol and trehalose, that are thought to be important to its survival of dehydration. It appears that these molecules are made from storage lipids by way of the glyoxylate cycle, a biochemical pathway that is unusual among animals. Once the animals have produced these compounds, they can be dehydrated to less than 2% water content and can be kept in this dehydrated state for years. Workers in this field originally believed that both these compounds might be involved in stabilizing dry cells, but other studies have shown that while trehalose appears to be present at high concentrations in all animals that survive extreme dehydration, glycerol is not always present. Trehalose is also found at high concentrations in lower plants, such as yeast cells and the desert resurrection plant, that survive dehydration, but the analog of trehalose in higher plants appears to be sucrose. Thus, the data available from comparative biochemistry strongly suggest that the disaccharides trehalose and sucrose are involved in protecting dry organisms from damage due to dehydration.

In contrast with the above examples, cysts of crustaceans like the brine shrimp *Artemia* and seeds of plants seem to require no inductive period prior to anhydrobiosis. An *Artemia* cyst, for example, is an embryo covered with a capsule that is produced by the mother. The encapsulated embryo is released from the mother in a partially dehydrated state, and may afterward be rapidly reduced to low water content. Female *Artemia* also produce embryos without capsules that proceed with normal development without an intervening period of anhydrobiosis. In fact, these nonencapsulated embryos appear incapable of surviving dehydration. Encapsulated embryos contain large amounts of free glycerol and trehalose, similar to the amounts found in anhydrobiotic nematodes. Embryos produced without capsules contain insignificant amounts of glycerol and far less trehalose. Thus, induction of anhydrobiosis certainly occurs in *Artemia*, but the inductive mechanisms, which are poorly understood, probably act on the parent rather than directly on the embryo. Similar arguments can be made concerning the induction of anhydrobiosis in the seeds of higher plants and bacterial and fungal spores.

Biochemical adaptation and anhydrobiosis. Aside from gross mechanical stresses, dehydration results in a variety of other stresses that the chemical composition of anhydrobiotes may inhibit or obviate altogether. For example, it is well established that most macromolecules and cell membranes are maintained in a functional state by hydrophobic interactions. When water is removed, molecules and membranes lose their functionality; in proteins hydrophobic residues are exposed, and membranes undergo phase transitions that are likely to result in displacement of integral proteins from the membrane. A considerable body of evidence is accumulating which suggests that disaccharide compounds like trehalose may interact with macromolecules and membranes, stabilizing them at low water activities. For example, it has been possible to dry membranes in the presence of trehalose and to obtain membranes that show normal biological activity and structure upon rehydration. By contrast, when the same membranes are dried without trehalose, extensive structural damage occurs and all biological activity is lost. Similarly, proteins that are usually labile to drying can be dried in the presence of trehalose without loss of activity. *See* CELL MEMBRANES.

Another possible consequence of dehydration is that molecules, normally separated by bulk water in hydrated cells, would come into proximity, potentially resulting in chemical reactions between these molecules. Trehalose could serve as a substitute for bulk water, preventing such reactions. Yet another possible consequence of dehydration arises from the observation that oxygen is deleterious to anhydrobiotes. The organisms survive much longer in the dry state if oxygen is excluded.

It has been shown that damage to the dry organisms is probably due to direct oxidation of their chemical constituents, resulting in the formation of highly reactive free radicals; and that free glycerol inhibits such oxidation, which may explain the presence of glycerol in many dry but viable organisms. Finally, the reaction between reducing sugars and protein in the dry state, which is known as the browning reaction, could have serious consequences for dry tissues. This reaction involves the formation of a covalent bond between the protein and sugar, resulting in an insoluble product and denatured protein. It is possible that glycerol inhibits this reaction. It is not known whether sugars like trehalose show a browning reaction with dry proteins, since it is a nonreducing disaccharide of glucose.

Thus, it would be an advantage to the anhydrobiote to store low-molecular-weight carbohydrates in this form.

Longevity of anhydrobiotes. It is not known just how long anhydrobiotes can survive. Some nematodes have been revived after being kept dry for 39 years. A museum specimen of moss that had been kept dry for over 120 years yielded a number of viable rotifers and tardigrades. Cysts of *Artemia* survive under appropriate conditions (under vacuum and at temperatures below 32°F or 0°C) for at least 10 years and probably considerably longer. Possibly the most celebrated (and most debated) case of extreme longevity of anhydrobiotes is that of the viable lotus seeds recovered from a dry peat bed in Manchuria by Professor I. Ohga in 1927. These seeds were initially reported to be about 400 years old, but radiocarbon dating suggested the seeds were no more than 150 years old. Other, similar reports have appeared, but the greatest authenticated longevity for a seed was obtained from the accidental hydration and subsequent germination of a seed in the British Museum in the course of an air raid in 1942. The seed had been mounted on a herbarium sheet in 1702. Thus, though there is no definitive data on the limit of longevity of anhydrobiotes, it seems likely to be impressive—on the order of centuries.

A factor that may be related to the longevity of anhydrobiotes is their phenomenal resistance to environmental extremes. In the 1920s researchers found that anhydrobiotic rotifers and tardigrades could be exposed to temperatures as high as 304°F (151°C) and as low as −328°F (−200°C), provided their water content was low. It has also been shown that anhydrobiotic rotifers and tardigrades survived exposure to a temperature at the brink of absolute zero (0.008 K). The reported LD_{50} (median lethal dose) for x-ray exposure of anhydrobiotic tardigrades is 570,000 roentgens; by contrast, the LD_{50} for humans is about 500 roentgens. Tardigrades and nematodes will survive exposure to high vacuum (less than 10^{-6} mmHg or 10^{-4}Pa). In fact, the tardigrade shown in Fig. 1*b* and the nematode in Fig. 2 were each kept at such a vacuum and were bombarded with electrons for 15 min or longer in a scanning electron microscope. When the animals were moistened after removal from the microscope they both recovered, although the tardigrade died shortly afterward. Similar observations have been made on cysts of *Artemia*, seeds, and spores.

The resistance of anhydrobiotes to environmental extremes has important economic consequences. As a result, considerable research has been devoted to killing anhydrobiotes if they are detrimental to humans and their various agricultural endeavors or to preserving the anhydrobiotes if they are beneficial. For example, many soil-dwelling nematodes that exhibit anhydrobiosis are important agricultural pests. In the anhydrobiotic state they are remarkably resistant to chemical control measures. *Artemia* cysts and seeds, on the other hand, are beneficial to humans, and the fact that they can be stored and transported dry is particularly convenient. *Artemia* cysts have become vital links in aquaculture systems. The cysts are hydrated and hatched, and the young brine shrimp are fed to fish, crustaceans, and other aquatic animals maintained in culture.

Other applications of anhydrobiosis. The discovery that biological membranes and proteins can be dried in the presence of trehalose without loss of activity has led to applications that may be important in human welfare. For instance, liposomes are artificial membranes enclosing an aqueous compartment. These structures are being used for delivery of water-soluble drugs to cells. The liposomes are injected intravenously and are transported through the circulatory system. They fuse with cell membranes, depositing the drug in the cytoplasm of the cell. Ultimately, it may be possible to target specific cells with liposomes. One difficulty with this application is that liposomes are not stable structures. However, it is now possible to dry the liposomes with trehalose and to store them in this state. When the liposomes are needed, water can be added and they are ready for immediate use. There are similar applications in other areas of the pharmaceutical industry, medicine, and agriculture.

Metabolism of anhydrobiotes. A central question in the study of anhydrobiosis has been whether metabolism actually ceases. Most investigations in this area have involved attempts at measuring the metabolic rate of the dry organism directly. Increasingly sophisticated measurements have been made, using, for example, oil-filled cartesian divers to record oxygen uptake, radiotracers to study metabolism of radiolabeled metabolic intermediates, and gas chromatography to determine emission of carbon dioxide from the anhydrobiote. In all these studies the rate of metabolism recorded is low or undetectable. However, such observations cannot be construed to imply that metabolism is lacking altogether. The argument can always be made that metabolism does proceed in the anhydrobiote but the techniques used are not sufficiently sensitive to measure it. The most convincing argument is the one based on the hydration properties of proteins in laboratory cultures: If all the water contained by an anhydrobiotic organism were associated with protein, the protein would still exist in a semicrystalline state, and enzymes would not possess catalytic activity. Thus, enzyme-mediated metabolism must almost certainly cease at the low water contents that these organisms can possess. However, some chemical reactions do proceed in anhydrobiotes; mostly they are deleterious ones, such as the browning reaction. Such reactions could be construed as metabolism. But would it then follow that chemical reactions in rocks, air, and dead animals are metabolism, or even that a cosmic metabolism comprises the total chemical reactions of the universe.

Thus, the available evidence strongly suggests that dry anhydrobiotes are ametabolic. If that is the case, a philosophical question immediately arises concerning the nature of life. If metabolism is absent, the organism is generally referred to as dead. By this line of reasoning, anhydrobiotes would therefore be dead,

returning to life when they are rehydrated. But it is also known that some anhydrobiotes die while in the dry state, in the sense that an increasing proportion of the population fails to resume activity upon rehydration. It follows, then, that they must have "died" while they were "dead."

This philosophical quandary can be avoided by applying the definition of life adopted by most students of anhydrobiosis: An organism is alive, provided its structural integrity is maintained. When that integrity is violated, it is dead. *See* METABOLISM; PLANT METABOLISM. John H. Crowe; Lois M. Crowe

Bibliography. J. H. Crowe et al., *Biochim. Biophys. Acta Membranes Rev.*, 947:367–384, 1988; J. H. Crowe and J. S. Clegg, *Anhydrobiosis*, 1973; J. H. Crowe and J. S. Clegg, *Dry Biological Systems*, 1978; C. Leopold, *Membranes, Metabolism and Dry Organisms*, 1986; S. Weisburd, Death-defying dehydration, *Sci. News*, 133:107–110, February 13, 1988.

Cryptography

The various mechanisms for data confidentiality, data authentication, and nonrepudiation (that is, the property to prevent the denial of previous commitments or actions). Cryptography, with a long history of being an art of secret writing, has evolved into a science. As society is increasingly dependent upon computers, the vast amounts of data communicated, processed, and stored within computer systems and networks often have to be protected, and cryptography is a means of achieving this. It provides confidentiality services for transmitted data as well as stored data. Cryptographic techniques can also be used for message authentication, entity authentication, and digital signature verification for electronic funds transfer and credit card transactions. *See* COMPUTER SECURITY; DATA COMMUNICATION; DATABASE MANAGEMENT SYSTEM; DIGITAL COMPUTER; ELECTRICAL COMMUNICATIONS.

Basics. Data confidentiality is the property that only those authorized can access the data. Cryptography achieves data confidentiality by transforming meaningful messages, called plaintext, into unintelligible forms, called ciphertext or cryptogram. The process of transforming plaintext into ciphertext is called encryption or encipherment, and the opposite process is called decryption or decipherment. (An encryption scheme is also referred to as a cipher.) As shown in **Fig. 1**, the encryption algorithm is controlled by a parameter k', and the decryption algorithm is controlled by a secret k, known as a key. For confidentiality purposes, k should be kept

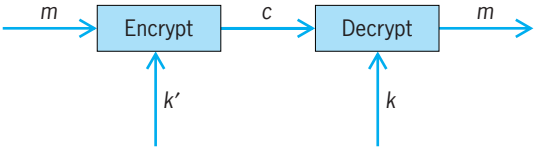

Fig. 1. The confidentiality model.

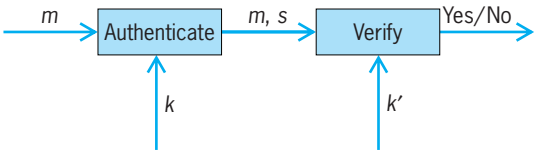

Fig. 2. The authentication model.

secret. If k' of a cipher can be used to derive k, then k' should also be kept secret, and such a cipher is called a symmetric key encryption. If it is computationally hard (see below) to derive k from k', then k' can be made public and the corresponding cipher is called an asymmetric encryption or a public key encryption. In most symmetric key encryption schemes, k' is the same as k and they are called symmetric keys. In public key cryptography, k' is called a public key and k is called a private key.

In a symmetric key encryption, if each encryption operation processes just one character or one bit, it is called a stream cipher. A block cipher, on the other hand, encrypts a block of fixed size, usually multiple bytes, each time.

In addition to data confidentiality, cryptography provides both data-origin authentication and entity authentication services. Data-origin authentication allows a message receiver to verify the message origin and its integrity. Entity authentication allows one to check the identity of another entity that possesses a certain secret. Given a message m, cryptography achieves data-origin authentication by appending to the message some additional bits (such as "s" in **Fig. 2**), called authentication tags, which are computed from m and a secret k. The receiver of the message and the authentication tags can verify the message origin and integrity by applying a verification parameter k' to the received data, including the authentication tags and the message. In most symmetric key cryptosystems, k and k' are the same and they should be kept secret. In public key cryptography, k' can be made public and k is kept secret; it is computationally hard (see below) to derive k from k'. Authentication tags generated by symmetric-key cryptography are called message authentication codes (MAC), which can be generated by both the sender and the receiver as both of them know k. In contrast, authentication tags by public key cryptography, called digital signatures, are stronger: since only the sender knows k and can produce a valid authentication tag under k, he cannot deny having sent the message in a later situation. Thus authentication tags by public key cryptography are the digital equivalent of a handwritten signature and have the potential for the nonrepudiation service.

Data confidentiality and data authentication are orthogonal in that one can exist without the other. On one hand, the authentication model given in Fig. 2 does not provide data confidentiality since message m is sent in the clear. On the other hand, encryption alone is insufficient to ensure that information is not altered during transmission. This fact is most evident when encryption with a public key algorithm is used. With a public key system, anyone can encipher

data by using the public enciphering key. Thus, any system user or node can masquerade as any other system user or node.

The opposite of cryptography is cryptanalysis, which studies techniques to defeat cryptographic methods. Cryptanalysts are those people doing cryptanalysis. The security of an encryption/authentication algorithm should depend solely on the secrecy of keys, not on the secrecy of algorithms. It is often assumed that the enemy cryptanalysts know the details of the algorithm employed.

In summary, symmetric block ciphers, symmetric stream ciphers, and public key encryption ciphers are used for data confidentiality, while MAC and digital signature algorithms are used for data authentication. We will first review the fundamental theory of these cryptographic algorithms and then give the details of some practical cryptographic algorithms.

Unbreakable algorithms. Unbreakable ciphers are possible, but the key must be randomly selected and used only once, and its length must be equal to or greater than that of the plaintext to be enciphered. Therefore such long keys, called one-time tapes, are not practical in data-processing applications.

Similarly, unconditionally secure authentication schemes (that is, schemes whose security does not depend on the computational power of the adversary) are possible. But the key must be randomly selected and used only once, which makes such keys impractical.

To work well, a key must be of fixed length, relatively short, and capable of being repeatedly used without compromising security. In theory, any algorithm that uses such a finite key can be analyzed; in practice, the effort and resources necessary to break the algorithm would be unjustified.

Strong algorithms. Fortunately, to achieve effective data security, construction of an unbreakable algorithm is not necessary. However, the work factor (a measure, under a given set of assumptions, of the requirements necessary for a specific analysis or attack against a cryptographic algorithm) required to break the algorithm must be sufficiently great. Included in the set of assumptions is the type of information expected to be available for cryptanalysis. For encryption schemes, this could be ciphertext only, plaintext (not chosen/known) and corresponding ciphertext, chosen plaintext and corresponding ciphertext, or chosen ciphertext and corresponding recovered plaintext.

A strong cryptographic algorithm must satisfy the following conditions: (1) The algorithm's mathematical complexity prevents, for all practical purposes, solution through analytical methods. (2) The cost or time necessary to unravel the message or key is too great when mathematically less complicated methods are used, either because too many computational steps are involved or because too much storage space is required.

To be strong, an encryption algorithm must satisfy the above conditions even when the analyst has the following advantages: (1) Relatively large amounts of plaintext (specified by the analyst, if so desired) and corresponding ciphertext are available. (2) Relatively large amounts of ciphertext (specified by the analyst, if so desired) and corresponding recovered plaintext are available. (3) Large high-speed computers are available for cryptanalysis.

Similarly, to be strong, an authentication algorithm must satisfy the above conditions even when the analyst has been given large amounts of messages (specified by the analyst, if so desired) and their corresponding authentication codes.

Computational complexity. The strength of a cryptographic scheme can be measured by the computational complexity of the task of cryptanalysis. The complexity of a task is the least possible number of elementary operations used by any program to accomplish this task. However, nontrivial lower bounds on the complexity of a problem are very hard to obtain. This is a fundamental problem in the design of cryptographic systems: it is very difficult to ensure that a system is sufficiently hard to crack. Without a good lower bound, the possibility that someone will find a fast algorithm for cryptanalyzing a given scheme must always be anticipated.

Problem classes P and NP. An important direction of theoretical work concerns the consideration (from computer science) of P versus NP. The class P consists of those problems which can be solved in polynomial time. That is, there are constants c and k such that, if the input to the problem can be specified in N bits, the problem can be solved on a sequential machine in time $c \times N^k$. Roughly speaking, these are the tractable problems. They include multiplication of two large numbers, exponentiation modulo a large prime, running the Data Encryption Standard (discussed below), and roughly any straightforward problem that does not involve searching.

The class NP (nondeterministic polynomial time) consists of problems which can be solved by searching. Roughly speaking, a possible solution to a problem in NP is to guess in turn each of 2^N possible values of some N-bit quantity, do some polynomial-time work related to each guess, and if some guess turns out to be correct, report the answer.

An example of a problem in NP is the knapsack problem: Given a set of integers $\{A_1, A_2, \ldots, A_n\}$ and a target integer B, can a subset of the A_i be selected without repetition (say $\{A_1, A_3, A_8\}$) such that their sum $(A_1 + A_3 + A_8)$ is the target B? One algorithm for solution is to try all possible subsets and just see whether any has the desired property. This algorithm requires exponential time, so called because the size of the input (n) occurs in the exponent of the formula expressing the running time (in this case, roughly 2^n). In fact, all known algorithms for this problem require exponential time. But there may be an unknown algorithm that runs in polynomial time; no proof prohibiting this is currently known.

Certainly, any problem in P is also in NP. A major outstanding question in computer science is whether P equals NP or whether NP is strictly larger.

There is a particular collection of problems in NP, including the knapsack problem, which are termed NP-complete. These can be thought of as the hardest

problems in NP. More precisely, an important mathematical result states: If there are any problems in NP but not in P, then each NP-complete problem is also in NP but not in P. This class has particular significance for cryptography.

In a good cryptographic system, certain operations (like decryption) are easy for those in possession of the key and difficult for those without the key. The cryptanalyst who could first guess and verify the correct key would be able to decrypt easily (in polynomial time). This could be done by searching over the space of possible keys and attempting to verify each one in polynomial time. Since the problem can be solved by searching, decryption is in NP.

If P = NP, then decryption would also be in P, and a good cryptographic system would most likely be difficult to design. If P $\neq$ NP, then the NP-complete problems might form a good starting point for cryptographic system design. Unfortunately, mathematicians and cryptographers have not yet learned how to transform an NP-complete problem into a secure cryptographic system.

Even if an NP-complete problem is eventually transformed successfully into a cryptographic system, a proof of the difficulty of cryptanalysis of that system (or any other) can be expected to be taxing. Such a proof would probably also prove that P $\neq$ NP, and this conjecture has eluded computer scientists so far. Thus, for now, the designers of cryptographic systems must rely on experience rather than rigorous proof of the difficulty of cryptanalysis.

Examples. Many practical and strong cryptographic algorithms have been developed. The Data Encryption Standard (DES), a block cipher, would require massive attacks to break it. The only known attacks are key exhaustion, which involves trying all $2^{56} \simeq 10^{17}$ possible keys, or attacks involving prohibitively large amounts (10^{14} bytes) of known plaintext and corresponding ciphertext. There is no guarantee against a more efficient, analytic attack, although none is known at present. A 56-bit key is not considered secure these days. A more advanced block cipher, Advanced Encryption Standard (AES), uses longer keys and larger data blocks and is assumed to be more secure and more efficient.

On the public key encryption vein, the RSA encryption algorithm, also discussed below, is based on the difficulty of factoring large numbers. However, a family of algorithms has been developed for factoring such numbers. If the modulus involved has, say, 2048 bits, then RSA should be secure in the foreseeable future. Other popular public key encryption algorithms include ElGamal public key encryption and Elliptic-curve (EC) public key encryption, both of which are based on the difficulty of the discrete logarithm problem.

Efficient MAC algorithms include those based on the cipher block chaining (CBC) mode of block ciphers, called CBC-MAC, and the MAC algorithms based on cryptographic hash algorithms, called HMAC.

Among the efficient digital signature algorithms are the RSA digital signature algorithm, the Digital Signature Standard (DSS, discussed below), and the Elliptic-curve Digital Signature Algorithm (ECDSA).

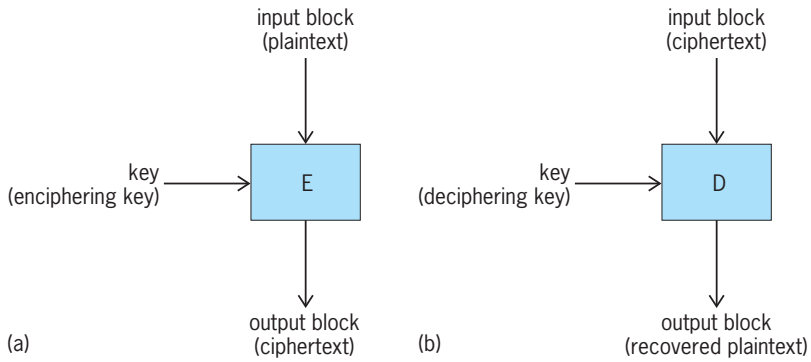

Fig. 3. Block cipher. (*a*) Enciphering. (*b*) Deciphering.

Block ciphers. A block cipher (**Fig. 3**) transforms a string of input bits of fixed length (called an input block) into a string of output bits of fixed length (called an output block). In a strong block cipher, the enciphering and deciphering functions are such that every bit in the output block jointly depends on every bit in the input block and on every bit in the key. This property is called intersymbol dependence.

Data encryption standard. Regardless of the application, a cryptographic system must be based on a cryptographic algorithm of validated strength if it is to be acceptable. The data encryption standard (DES) was such a validated conventional algorithm available in the public domain. It was accepted as a standard by the NIST in 1997 and was officially retired in 2004.

The DES enciphers a 64-bit block of plaintext into a 64-bit block of ciphertext under the control of a 56-bit key. Conceptually, it can be thought of as a huge key-controlled substitution box (S-box) with a 64-bit input and output. With such an S-box, 2^{64}! different transformations or functions from plaintext to ciphertext are possible. The 56-bit key used with DES thus limits the number of usable functions to 2^{56}.

A single, huge S-box is impossible to construct. Therefore DES is implemented by using several smaller S-boxes (with a 6-bit input and a 4-bit output) and permuting their concatenated outputs. By repeating the substitution and permutation process several times, cryptographic strength "builds up." The DES encryption process (**Fig. 4**) consists of 16 iterations, called rounds. At each round a cipher function (f) is used with a 48-bit key. The function is composed of the substitution and permutation. The 48-bit key, which is different for each round, is a subset of the bits in the externally supplied key.

The externally supplied key consists of 64 bits (56 bits are used by the algorithm, and up to 8 bits may be used for parity checking). By shifting the original 56-bit key, a different subset of 48 key bits is selected for use in each round. These key bits are labeled $K_1, K_2, \ldots, K_{16}$. To decipher, the keys are used in reverse order (K_{16} is used in round one, K_{15} in round two, and so on).

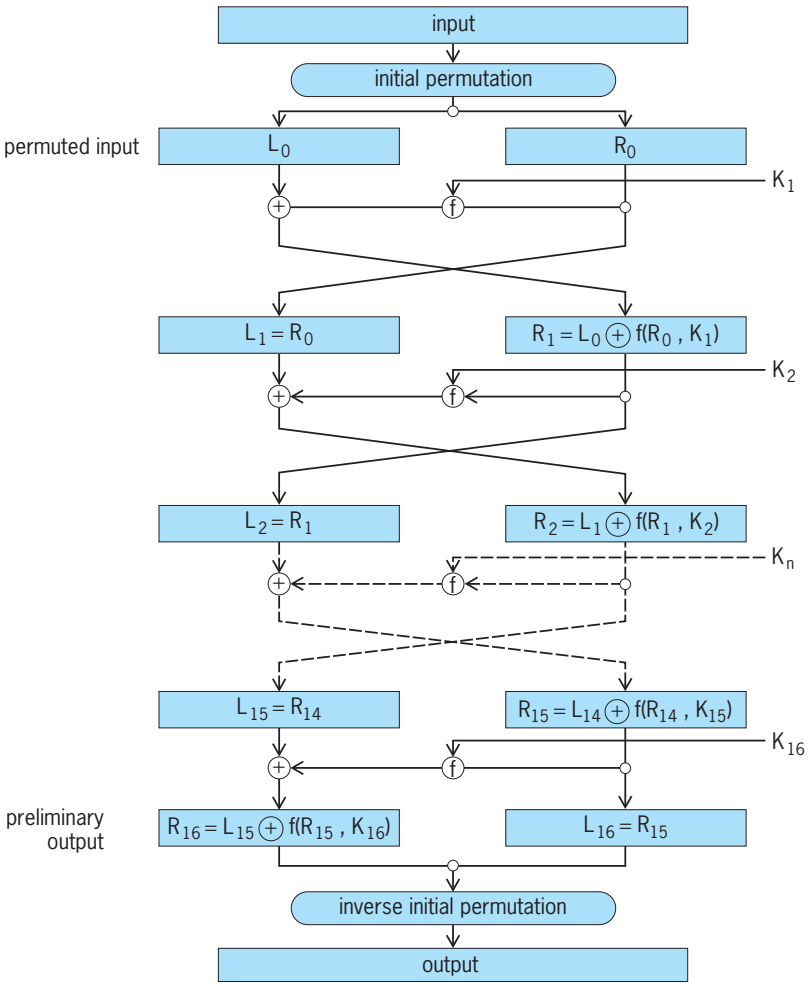

Fig. 4. Enciphering computation in the Data Encryption Standard.

At each round (either encipherment or decipherment), the input is split into a left half (designated L) and a right half (designated R) [Fig. 4]. R is transformed with f, and the result is combined, using modulo 2 addition (also called the EXCLUSIVE OR operation; $0 \oplus 1 = 1$, $1 \oplus 0 = 1$, $0 \oplus 0 = 0$, $1 \oplus 1 = 0$), with L. To consider one round of encipherment (**Fig. 5**), the input block (X) may be denoted $X = (L_0, R_0)$, where L_0 and R_0 are the left and right halves of X, respectively. Function f transforms R_0 into $f_{K1}(R_0)$ under control of cipher key K_1. L_0 is then added (modulo 2) to $f_{K1}(R_0)$ to obtain R_1, as in Eq. (1). The round is then completed by setting L_1 equal to R_0.

$$L_0 \oplus f_{K_1}(R_0) = R_1 \qquad (1)$$

The above steps are reversible, without introducing any new parameters or requiring that f be a one-to-one function. The ciphertext contains L_1, which equals R_0, and therefore half of the original plaintext is immediately recovered (**Fig. 6**). The remaining half, L_0, is recovered by recreating $f_{K1}(R_0)$ from $R_0 = L_1$ and adding it (modulo 2) to R_1, as in Eq. (2).

$$R_1 \oplus f_{K_1}(R_0) = [L_0 \oplus f_{K_1}(R_0)] \oplus f_{K_1}(R_0) = L_0 \qquad (2)$$

However, to use the above procedure for encipherment as well as decipherment, the left and right halves of the output must be interchanged; that is, the ciphertext (Y) is defined by Eq. (3). This modified procedure easily extends to n rounds, where the keys used for deciphering are $K_n, K_{n-1}, \ldots, K_1$.

$$Y = [L_0 \oplus f_{K_1}(R_0)], R_0 \qquad (3)$$

Advanced Encryption Standard (AES). The key length of DES is 56 bits, which is considered weak. Also, the 64-bit DES data block size is too small for modern applications. To overcome these difficulties, the National Institute of Standards and Technology (NIST) adopted an adapted version of the Rijndael algorithm, a symmetric block cipher, as the Advanced Encryption Standard (AES). AES uses 128-bit data block size and its keys can be 128, 192, and 256 bits, called *AES-128*, *AES-192*, and *AES-256*, respectively. Like DES, AES has multiple rounds: AES-128 has 10 rounds, AES-192 has 12 rounds, and AES-256 has 14 rounds. Different from DES, AES is not based on the Feistel cipher structure (Fig. 5).

In AES, the 16-byte plaintext is organized into a 4-by-4-byte array, called the State. The plaintext bytes are filled into the byte array in a column-wise way: the first column first, then the second column, the third column, and the last column. During the AES encryption, this state is continuously transformed and the end state is the ciphertext.

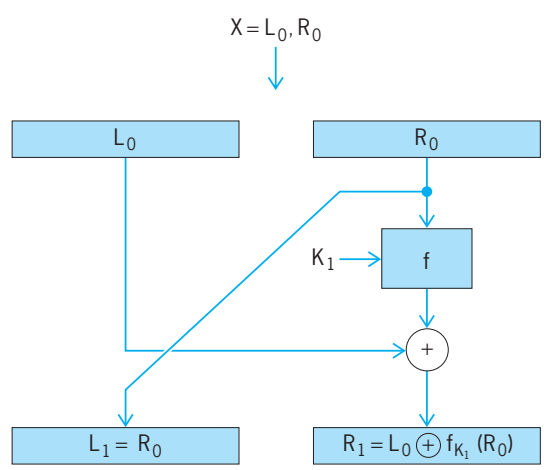

Fig. 5. Transformation of input block (L_0, R_0).

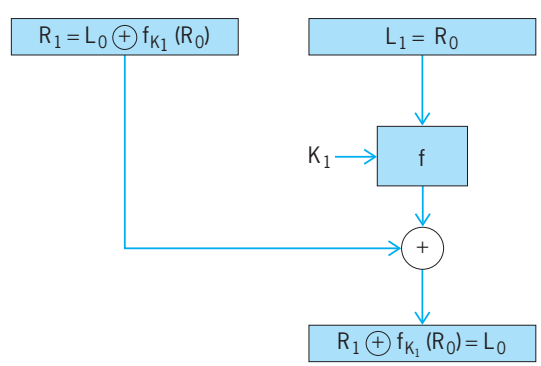

Fig. 6. Recovery of L_0.

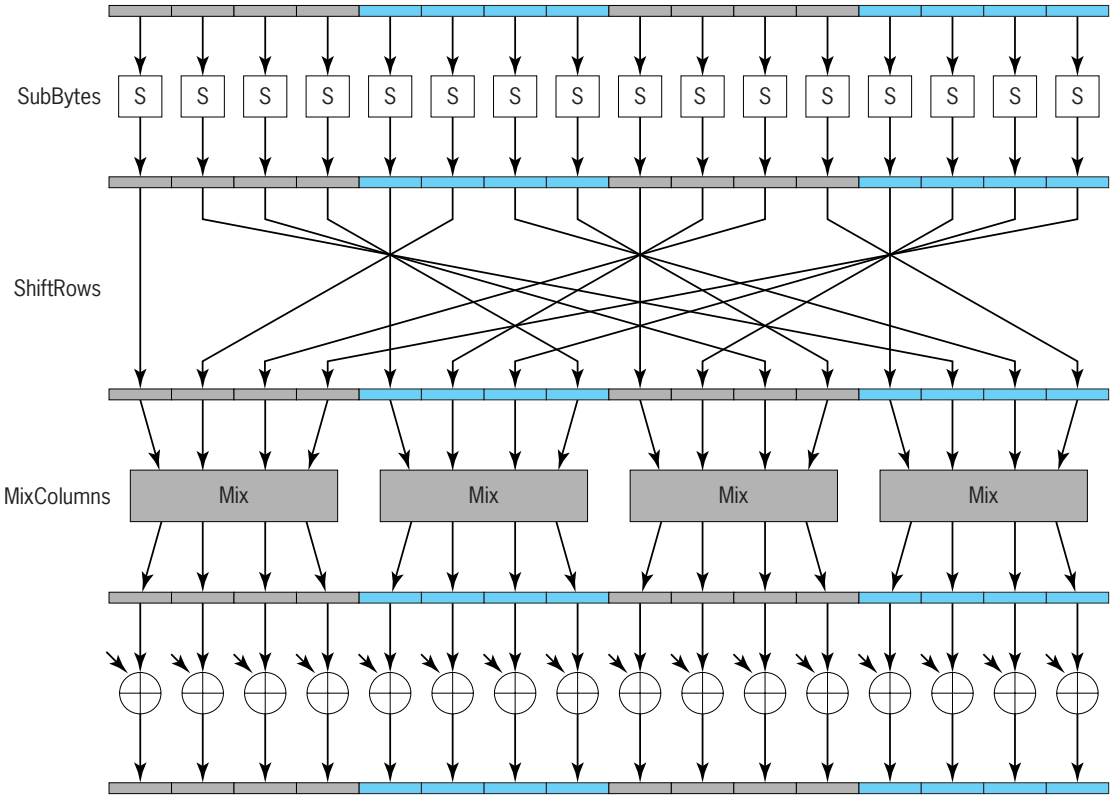

Fig. 7. AES round function.

For AES-128, AES-192, and AES-256, from the given cipher key, the AES key expansion operation derives, respectively, 11, 13, and 15 128-bit round keys.

The AES encryption starts with an initial round key addition, in which the first round key is added to the State by a simple bitwise XOR operation. Then, the State array is transformed by a round function multiple times (10 times for AES-128, 12 times for AES-192, and 14 times for AES-256), with the final round differing slightly with the other rounds. Each time the round function consumes one round key.

Figure 7 shows the details of the AES round function, which comprises four transformations, SubBytes, ShiftRows, MixColumns, and AddRoundKey. The SubBytes transformation is a nonlinear byte substitution (called S-box) and it works independently on a single byte. The ShiftRows transformation shifts around the bytes within the same rows of the State. The MixColumns transformation mixes the bytes within the same column of the State. The AddRoundKey transformation applies the corresponding round key to the State by a simple bitwise XOR operation. In the last AES round, the MixColumns transformation does not appear.

All the four AES transformation operations are reversible, which makes AES decryption possible. More specifically, AddRoundKey is its own inverse; the inverse of SubBytes is InvSubBytes, in which an inverse S-box is applied to the each byte of the State; the inverse of the ShiftRows transformation is InvShiftRows; and the inverse of the MixColumns transformation is InvMixColumns. The AES decryption is performed through a sequence of the inverse transformations, in reverse order, with the corresponding round keys generated from the same key expansion operation. This makes the AES decryption structure completely different from the AES encryption structure. AES decryption can also be structured in a similar way as the AES encryption, but a different key expansion operation should be employed.

Modes of operation. As a rule, block encryption is used to protect keys. A different method, called chained block encryption, is used to protect data. In chaining, the process of enciphering and deciphering is made dependent on other (prior) data, plaintext, ciphertext, and the like, also available at the time enciphering and deciphering take place. Thus every bit in a given output block depends not only on every bit in its respective input block and every bit in the key, but also on any or all prior data bits, either inputted to or produced during the enciphering or deciphering process.

Sometimes data to be enciphered contain patterns that extend beyond the cipher's block size. These patterns in the plaintext can then result in similar patterns in the ciphertext, which would indicate to a cryptanalyst something about the nature of the plaintext. Thus, chaining is useful because it significantly reduces the presence of repetitive patterns in the ciphertext. This is because two equal input blocks encipher into unequal output blocks.

A recommended technique for block chaining, referred to as cipher block chaining (CBC), uses a

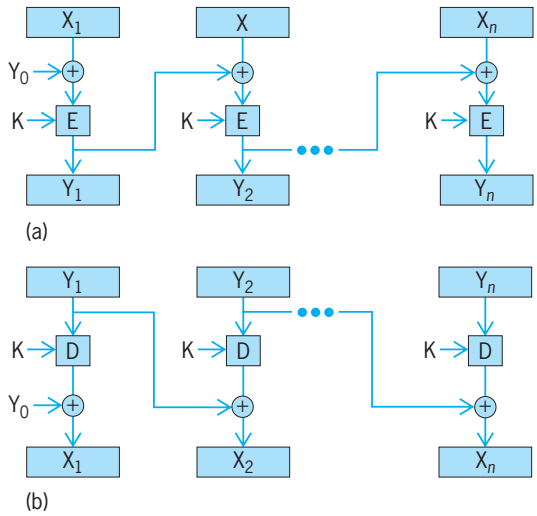

Fig. 8. Block chaining with ciphertext feedback. (*a*) Enciphering. (*b*) Deciphering.

ciphertext feedback (**Fig. 8**). Let $X_1, X_2, \ldots, X_n$ denote blocks of plaintext to be chained using key K; let Y_0 be a nonsecret quantity defined as the initializing vector; and let $Y_1, Y_2, \ldots, Y_n$ denote the blocks of ciphertext produced. The *i*th block of ciphertext (Y_i) is produced by EXCLUSIVE ORing Y_{i-1} with X_i and enciphering the result with K, as in Eq. (4) , where $\oplus$ denotes the EXCLUSIVE OR operation, or modulo 2

$$E_K(X_i \oplus Y_{i-1}) = Y_i \quad i \geq 1 \tag{4}$$

addition. Since every bit in Y_i depends on every bit in X_1 through X_i, patterns in the plaintext are not reflected in the ciphertext.

The *i*th block of plaintext (X_i) is recovered by deciphering Y_i with K and EXCLUSIVE OR-ing the result with Y_{i-1}, as in Eq. (5). Since the recovered

$$D_K(Y_i) \oplus Y_{i-1} = X_1 \quad i \geq 1 \tag{5}$$

plaintext X_i depends only on Y_i and Y_{i-1}, an error occurring in ciphertext Y_j affects only two blocks of recovered plaintext (X_j and X_{j+1}).

Other chained block encryption methods include cipher feedback (CFB), output feedback (OFB), and counter mode (CTR).

Stream ciphers. A stream cipher (**Fig. 9**) employs a bit-stream generator to produce a stream of binary digits R (traditionally termed the key stream), which is then combined either with plaintext X to produce ciphertext or with ciphertext Y to recover plaintext. In most stream ciphers, an EXCLUSIVE OR operation, or modulo 2 addition, is used to combine the respective bit streams. Thus encipherment and decipherment are defined by $X \oplus R = Y$ and $Y \oplus R = X$, respectively.

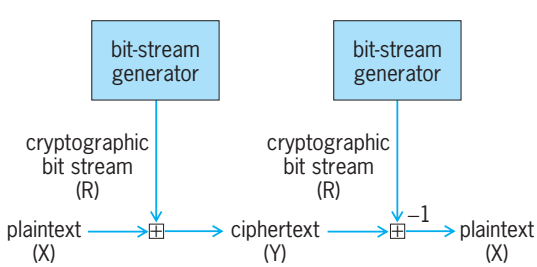

Fig. 9. Stream cipher concept.

If the bit-stream generator were truly random, an unbreakable cipher could be obtained by EXCLUSIVE ORing the plaintext and the cryptographic bit stream R. The cryptographic bit stream would be used directly as the key, and would be equal in length to the message. But in that case the cryptographic bit stream must be provided in advance to the communicants via some independent and secure channel. This introduces insurmountable logistic problems for heavy data traffic. Hence, for practical reasons, the bit-stream generator must be implemented as an algorithmic procedure. Then both communicants can generate the same cryptographic bit stream—provided that their algorithms are identically initialized (**Fig. 10**).

When modulo 2 addition is used as the combining operation, each bit in the output ciphertext (recovered plaintext) is dependent only upon its corresponding bit in the input plaintext (ciphertext). This is in marked contrast to the block cipher. Both approaches, however, have comparable strength.

For security purposes, a stream cipher must never predictably start from the same initial condition, thereby producing the same cryptographic bit stream. This can be avoided by making the cryptographic bit stream dependent on a nonsecret quantity Z (known as seed, initializing vector, or fill), which is used as an input parameter to the ciphering algorithm (Fig. 9). *See* INFORMATION THEORY.

The most famous symmetric stream cipher might be RC4.

RSA public key algorithm. The RSA algorithm (named for the algorithm's inventors, R. L. Rivest, A. Shamir, and L. Adleman) is a public key encryption algorithm based on the fact that factoring large composite numbers into their prime factors involves an overwhelming amount of computation.

To describe the RSA algorithm, the following quantities are defined:

p and q are primes	(secret)
$N = p \cdot q$	(nonsecret)
$\varphi(N) = (p-1)(q-1)$	(secret)
d is the private key	(secret)
e is the public key	(nonsecret)
m is the plaintext	(secret)
c is the ciphertext	(nonsecret)

Based on an extension of Euler's theorem (6) the

$$m^{x\varphi(N)+1} \equiv m \bmod N \tag{6}$$

algorithm's public and private keys (e and d) are chosen so that Eq. (7) or, equivalently, Eq. (8) [which is

$$d \times e = x\varphi(N) + 1 \tag{7}$$

$$d \times e = 1 \bmod \varphi(N) \tag{8}$$

used here in place of a more complicated equation for ease of understanding] is satisfied. By selecting two secret prime numbers p and q, the user can calculate $N = p \times q$, which is made public, and $\varphi(N) = (p - 1) \times (q - 1)$, which remains secret and is used to solve Eq. (8). (Tests are available to determine with a high level of confidence if a number is prime or not.) To obtain a unique solution for the public key (e), a random number, or secret key (d), is selected that is relatively prime to $\varphi(N)$. (Integers a and b are relatively prime if their greatest common divisor is 1.) e is the multiplicative inverse of d, modulo $\varphi(N)$, and e can be calculated from d and $\varphi(N)$ by using Euclid's algorithm. Equation (6) can therefore be rewritten as Eq. (9), which holds true for any plain-

$$m^{d \times e} \equiv m \bmod N \tag{9}$$

text (m). Encipherment and decipherment can now be interpreted as in Eqs. (10) and (11). Moreover, be-

$$E_e(m) = c \equiv m^e \bmod N \tag{10}$$

$$D_d(c) \equiv c^d \bmod N \equiv m^{e \times d} \bmod N \equiv m \bmod N \tag{11}$$

cause multiplication is a commutative operation ($e \times d = d \times e$), encipherment followed by decipherment is the same as decipherment followed by encipherment. Thus the RSA algorithm can be used for both confidentiality and digital signatures (see below).

Finally, since $m^e(\bmod N) \equiv (m + xN)^e(\bmod N)$ for any integer x, $E_e(m) = E_e(m + xN)$. Thus the transformation from plaintext to ciphertext, which is many-to-one, is made one-to-one by restricting m to the set $\{0, 1, \ldots, N - 1\}$.

Message authentication codes. Message authentication codes are based on symmetric key cryptography, and they assume that the message sender and receiver share a common symmetric key.

The strength of a MAC algorithm depends on its output length, its key length, and the algorithm itself. The strongest attack against a MAC algorithm is the key recovery attack, in which an attacker can recover the symmetric key from some messages and their corresponding MACs under that key. A weaker attack is the forgery attack: an attacker might not be able to find the symmetric key used, but can still forge a valid MAC on a given new message, using help from some known/chosen messages and their corresponding MACs under a certain key. A secure MAC algorithm should be secure against both types of attacks.

Two widely used MAC algorithms are the CBC-MAC, a MAC algorithm based on CBC mode of block ciphers, and HMAC, a MAC algorithm based on cryptographic hash algorithms such as SHA-1. In a CBC-MAC, the symmetric key is used as a block cipher key to encrypt the given message in CBC mode; the last block ciphertext is the MAC code (other ciphertext blocks are thrown away).

In a HMAC, a cryptographic hash algorithm, h, is used. Given a symmetric key k and a message m, its HMAC code is computed as $h(k \oplus \text{opad} \parallel h(k \oplus \text{ipad} \parallel m))$, where opad and ipad are specified constants.

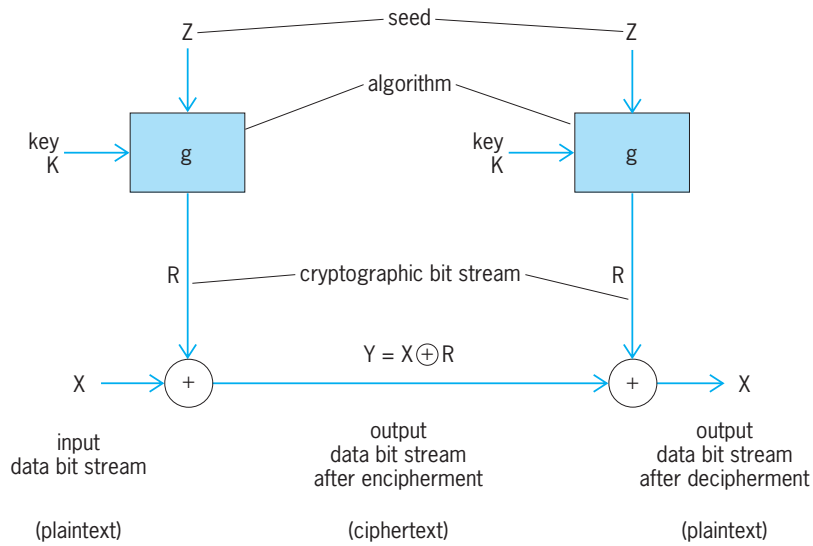

Fig. 10. Stream cipher using an algorithmic bit stream generator, modulo 2 addition, and seed.

Digital signatures. Digital signatures authenticate messages by ensuring that the sender cannot later disavow messages, the receiver cannot forge messages or signatures, and the receiver can prove to others that the contents of a message are genuine and that the message originated with that particular sender. The digital signature is a function of the message, a secret key or keys possessed by the sender and sometimes data that are nonsecret or that may become nonsecret as part of the procedure (such as a secret key that is later made public).

Digital signatures are more easily obtained with public key than with conventional algorithms. When a message is enciphered with a private key (known only to the originator), anyone deciphering the message with the public key can identify the originator. The originator cannot later deny having sent the message. Receivers cannot forge messages and signatures, since they do not possess the private key.

Since enciphering and deciphering keys are identical in a conventional algorithm, digital signatures must be obtained in some other manner. One method is to use a set of keys to produce the signature. Some of the keys are made known to the receiver to permit signature verification, and the rest of the keys are retained by the originator in order to prevent forgery.

The RSA algorithm previously discussed can be used for digital signatures. The National Institute of Standards and Technology (NIST) proposed another method for generating digital signatures, based on a 1985 proposal by T. El Gamal. Its security relies on the difficulty of computing logarithms over finite fields, or discrete logarithms, that is, of solving the following problem: Given a prime p and integers g, y, find an integer x satisfying $g^x \equiv y \,(\bmod\, p)$. In this scheme, all users share a large prime number p and a generator g (that is, an integer whose powers give all nonzero residues modulo p). Each user A has a

secret key x_A and a public key $y_A \equiv g^{(x_A)} \pmod{p}$. To each message m is associated an integer $h(m)$, $0 < h(m) < p$. User A "signs" the message by producing a pair of integers (r, s) between 0 and p satisfying Eq. (12). [The signature is computed by choosing a

$$g^{h(m)} \equiv y_A^{\,r} r^s \tag{12}$$

random integer k between 0 and $p - 1$, relatively prime to $p - 1$, computing $r \equiv g^k \pmod{p}$, and solving for s in Eq. (13): different random choices k will

$$h(m) \equiv x_A r + ks \bmod (p - 1) \tag{13}$$

give different signatures.] This "signature" is verified by using modular exponentiation of large integers. But a forger, without knowledge of the secret quantity x_A, apparently cannot produce such a signature without first solving the difficult discrete logarithm problem. *See* NUMBER THEORY.

A signature in the ElGamal scheme consists of two integers, (r, s), each about as large as the prime p. Since p must have at least 512 bits to make the discrete logarithm problem infeasible, the signature would then require 1024 bits.

A modification incorporated into the NIST proposal allows for shorter signatures. Here p is chosen such that $p - 1$ has a moderate prime factor $q \simeq 2^{160}$, and g is chosen so that its powers mod p form a subgroup of size q, so that s can be specified with fewer bits (160). The discrete logarithm problem is now restricted to that subgroup, which may or may not make cryptanalysis easier.

A caution in the design of digital signature schemes is the so-called birthday paradox: if signatures are n bits long, then among a collection of $(2^n)^{1/2} = 2^{n/2}$ messages there are likely to be two with the same signature. An opponent who finds it feasible to generate $2^{n/2}$ messages and their signatures can use the two that match to generate a signature. To thwart an opponent whose computing power might be 2^{64}, a digital signature of at least 128 bits is needed. The ElGamal, NIST, and RSA schemes need larger signatures than that because of algebraic shortcuts.

Xunhua Wang; Don Coppersmith;
Stephen M. Matyas; Carl H. Meyer

Bibliography. N. Ferguson and B. Schneier, *Practical Cryptography*, Wiley, 2003; W. Mao, *Modern Cryptography: Theory and Practice*, Prentice Hall PTR, 2003; A. Menezes, P. van Oorschot, and S. Vanstone, *Handbook of Applied Cryptography*, CRC, 1996; C. H. Meyer and S. M. Matyas, *Cryptography: A New Dimension in Computer Data Security*, Wiley, 1982; G. J. Simmons (ed.), *Contemporary Cryptology: The Science of Information Integrity*, IEEE Press, 1992; D. R. Stinson, *Cryptography: Theory and Practice*, CRC, Boca Raton, 1995.

Cryptomonadida

An order of the class Phytamastigophorea, also known as the Cryptomonadina. Cryptomonads are considered to be protozoa by zoologists and algae by botanists. Most species occur in fresh water. They are olive-green, blue, red, brown, or colorless. All flagellated members have two subequal flagella inserted in a gullet opening through an obliquely truncate anterior end (see **illus.**). In *Nephroselmis* and *Protochrysis* the gullet is laterally displaced and a contractile vacuole opens into it. The nucleus is posterior. Oil and starch or a related compound are formed and pyrenoidlike bodies are present. *Cyathomonas* and *Chilomonas* are colorless; the latter has been widely used in nutritional studies, and extensive studies have been made of its cytology, demonstrating trichocysts around the gullet and muciferous bodies under the pellicle.

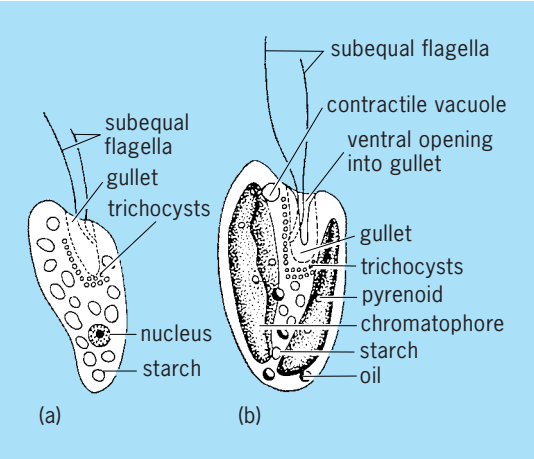

Examples of cryptomonads. (*a*) *Chilomonas paramecium*. (*b*) *Cryptomonas erosa*.

Phaeoplax is a palmelloid type with nonmotile cells embedded in a gelatinous matrix. The reproductive cells are flagellated. Radiolarians have symbiotic yellow cells, zooxanthellae, which are cryptomonads. Although few in genera and species, Cryptomonadida form extensive marine and freshwater blooms. *See* CILIA AND FLAGELLA; PHYTAMASTIGOPHOREA; PROTOZOA.

James B. Lackey

Cryptophyceae

A small class of biflagellate unicellular algae (cryptomonads) in the chlorophyll *a–c* phyletic line (Chromophycota). In protozoological classification, these organisms constitute an order, Cryptomonadida, of the class Phytomastigophora. Cryptomonads are 4–80 micrometers long, ovoid or bean-shaped, and dorsoventrally flattened. The cell is bounded by a moderately flexible periplast comprising the plasmalemma and underlying rectangular or polygonal proteinaceous plates. A tubular invagination (gullet, groove, or furrow) traverses the ventral cytoplasm and opens just below the apex of the cell. A pair of subequal flagella, which are covered with hairs and small scales, arise from the center or apical end of the gullet.

Cryptomonads may be photosynthetic, osmotrophic, or phagotrophic. In photosynthetic species,

chloroplasts occur singly or in pairs. The cryptomonad that is symbiotic in the ciliate *Mesodinium* is exceptional in having numerous chloroplasts. The photosynthetic lamellae, which are composed of loosely appressed pairs of thylakoids, are surrounded by four membranes. As in other members of the Chromophycota, the inner pair constitutes the chloroplast envelope while the outer pair represents endoplasmic reticulum confluent with the outer membrane of the nuclear envelope. The periplastidial compartment (space between the pairs of membranes on one side of the chloroplast) is more extensive than in other chromophytes. In other algae the compartment contains a dense network of tubules and vesicles, but in cryptomonads it contains starch grains, ribosomes, and a unique structure called a nucleomorph, in addition to scattered tubules and vesicles. The nucleomorph, which is smaller than the chloroplast, is bounded by a double membrane with numerous pores. It contains fibrillogranular inclusions. One nucleomorph, which always divides before the cell enters mitotic prophase, is closely associated with each chloroplast. If it can be shown that the nucleomorph is a vestigial nucleus, as it seems to be, the hypothesis that cryptomonad chloroplasts are evolutionarily derived from an endosymbiont will be supported.

Photosynthetic pigments include chlorophyll *a* and *c*, α-carotene and β-carotene (the former being predominant, an unusual ratio for algae), and red and blue phycobiliproteins closely related to those in red algae but occurring in the intrathylakoidal space rather than forming phycobilisomes. The color of the chloroplasts, and thus the color of the cryptomonad, depends upon pigment composition and may be green, olive, brown, yellow, red, or blue. Pigment composition is affected by environmental conditions. Thus, red cryptomonads at a low nitrate level gradually and reversibly bleach to green. Pyrenoids, usually capped with starch, are present in the chloroplasts of most species while eyespots are present in a few.

A single wormlike or reticulate mitochondrion, differing from those of other chromophytes in having flattened rather than tubular or inflated cristae, is present in each cell. Most cryptomonads have two sizes of ejectile organelles (trichocysts or ejectosomes)—those large enough to be seen with the compound microscope form rows along the gullet, and those much smaller are situated at corners of the periplast plates. Discharge of trichocysts imparts a characteristic darting movement to the cell. Contractile vacuoles may be present in fresh-water species.

Reproduction is by longitudinal binary fission. Sexual reproduction has not been confirmed.

Cryptomonads are found in fresh, brackish, marine, and hypersaline bodies of water, sometimes in great abundance. Several species are endosymbionts of marine ciliates and invertebrates, and of fresh-water dinoflagellates. They do not appear to be closely related to any other group of algae and thus are sometimes assigned to their own division. In addition to typical cryptomonads (order Cryptomonadales), of which about 200 species have been described, there are two monospecific genera with a life history that includes only a brief motile phase. In these genera, which constitute the order Tetragonidiales, the cells usually are nonflagellate and sedentary and are said to have a cellulose wall. They form zoospores, however, which are typically cryptomonad in structure. *See* ALGAE; CHROMOPHYCOTA; CRYPTOMONADIDA. Paul C. Silva; Richard L. Moe

Bibliography. R. W. Butcher, *An Introductory Account of the Smaller Algae of British Coastal Waters*, pt. 4: *Cryptophyceae*, 1967; R. E. Norris, Cryptophyceae, in S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Cryptostomata

An extinct order of Bryozoa in the class Stenolaemata. Cryptostomes had basally attached, erect, calcified colonies made of sheets or delicate branches composed of short tubular or box-shaped feeding zooids (individual units that constitute these colonial organisms) and commonly extrazooidal skeletal deposits. *See* BRYOZOA; STENOLAEMATA.

Morphology. Most cryptostomes have a sharp bend at the transition from the thin-walled inner portion to the thick-walled outer portion of zooids, and this transition point commonly has one or more shelflike structures called hemisepta that generate pronounced constrictions in the width of the zooids. Complete transverse partitions (diaphragms) crossing the interior of zooids may be present. The thick outer walls were built of superposed laminae typically interrupted by small, transverse rod-shaped structures (styles). Overall colony morphology is the basis for dividing cryptostomes into four subgroups, discussed below, that commonly are themselves treated as orders. Small "space-filling" zooids were common, but other types of polymorphic zooids were uncommon.

Classification and history. Rhabdomesina is characterized by arborescent colonies with cylindrical or nearly cylindrical branches, typically having feeding zooids located around the entire perimeter of the branches; some colonies tend to have feeding zooids concentrated toward one side of branches and an extrazooidal skeletal band on the opposite side. Colonies either may be rigidly calcified or may have calcified branch segments joined at flexible, uncalcified joints. Rhabdomesina ranged from the Early Ordovician into the Late Triassic.

Ptilodictyina is characterized by bifoliate colonies in which zooids grew back to back and were oriented toward opposite surfaces of relatively flattened branches or erect sheets. Branches typically divided dichotomously. Closely spaced branching and branch fusion produced reticulated colonies in some ptilodictyines. Others with more widely spaced branching generated arborescent colonies. A few grew as a rigid erect structure above a flexible joint at the base, some others had flexible joints at

branch junctions, but most were completely rigid. Ptilodictyine colonies that grew as sheets had regularly spaced interruptions on the surfaces such as small groups of depressed or protuberant zooids, or spots where feeding zooids were absent. Ptilodictyina existed from the Early Ordovician through the late Carboniferous.

Fenestellina is characterized by colonies with very narrow branches that had feeding zooids oriented toward one surface and massive extrazooidal skeletal deposits on the opposite surface. Branches were closely spaced and were typically linked, either by bars made of extrazooidal skeleton or by anastomosis, into netlike (fenestrate) colonies. Polymorphic zooids were uncommon. These organisms ranged from the Early Ordovician through the late Permian.

Timanodictyina is characterized by cylindrical or bifoliate branches that had very thick walls, containing abundant tiny styles, in outer portions of zooids. They existed from the lower Carboniferous through the late Permian.

Cryptostomes were exclusively marine and lived in a wide variety of relatively low-energy environments on continental shelves and seas. They were important contributors to the growth of organic mounds during the Late Ordovician and early Carboniferous. Some Carboniferous limestones are made predominantly of fenestellid skeletons, and may be petroleum reservoirs where space between the skeletal fragments was not filled by mud or precipitates.

Phylogenetic relationships among the four subgroups are not clear. Probably the Rhabdomesina is paraphyletic and gave rise to the other three groups, each of which apparently is monophyletic.

Frank K. McKinney

Crystal

A solid in which the atoms or molecules are arranged periodically. Within a crystal, many identical parallelepiped unit cells, each containing a group of atoms, are packed together to fill all space (see **illus.**). In scientific nomenclature, the term crystal is usually short for single crystal, a single periodic arrangement of atoms. Most gems are single crystals. However, many materials are polycrystalline, consisting of many small grains, each of which is a single crystal. For example, most metals are polycrystalline. *See* SINGLE CRYSTAL.

In electronics, the term crystal is restricted to mean piezoelectric crystal. Piezoelectric crystals contract or expand under application of electric voltages, and conversely they generate voltages when compressed. They are used for oscillators, pressure sensors, and position actuators. *See* PIEZOELECTRICITY.

The anisotropic microscopic structure of a crystal is often reflected in its external form, consisting of flat faces and sharp edges. Crystal structure is generally determined via diffraction of x-rays, neutrons, or electrons. Unlike disordered materials such as glasses or liquids, the diffraction pattern of a periodic array of atoms consists of individual sharp spots. The symmetry and structure of the crystal can be inferred from the symmetry of the diffraction pattern and the intensities of the diffracted beams. *See* ELECTRON DIFFRACTION; NEUTRON DIFFRACTION; X-RAY DIFFRACTION.

A crystal can be characterized by the symmetry operations that leave its structure invariant. These can include rotation about an axis through a specific angle, reflection through a plane, inversion through a point, translations by a unit cell dimension, and combinations of these. It can be shown that, for a periodic structure, the only allowable rotational symmetries are twofold, threefold, fourfold, and sixfold (see illus.).

A quasicrystal is a solid which yields a sharp diffraction pattern but has rotational symmetries (such as 5-fold or 10-fold) which are inconsistent with a periodic arrangement of atoms. This apparent paradox can be explained if the structure is assumed to consist of two or more unit cells, arranged according to precise rules. *See* QUASICRYSTAL.

A plastic crystal is generally composed of organic molecules which are rotationally disordered. The centers of the molecules lie at well-defined, periodically spaced positions, but the orientations of the molecules are random. Plastic crystals are often very soft and may flow under their own weight.

A liquid crystal is a material which is intermediate in structure between a liquid and a solid. Liquid crystals usually flow like liquids but have some degree of internal order. They are generally composed of rodlike organic molecules, although in some cases they are composed of disklike molecules. In a nematic liquid crystal, the rods all have the same general orientation, but the positions of the rods are disordered. In a smectic liquid crystal, rodlike molecules are ordered into sheets, within which there is only liquidlike order. A smectic can thus be thought of as being crystalline in one dimension and liquid in

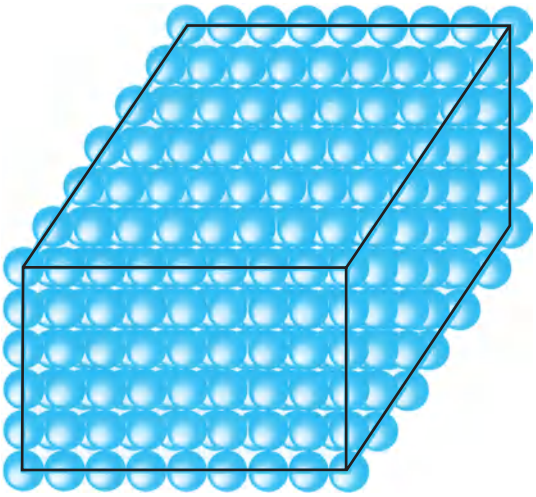

Structure of a simple crystal. Spheres, representing atoms, are packed together into a cubic lattice. This crystal has 4-fold symmetry axes passing through the front face; after a 90° rotation the structure appears unchanged.

the other two. In a discotic liquid crystal, disklike molecules are ordered into columnar arrays; there is short-range liquidlike order within the columns, but the columns form a two-dimensional crystal. *See* CRYSTAL DEFECTS; CRYSTAL GROWTH; CRYSTAL STRUCTURE; CRYSTALLOGRAPHY; LIQUID CRYSTALS.

Paul A. Heiney

Bibliography. B. D. Cullity, *Elements of X-ray Diffraction*, 2d ed., 1978; D. P. DiVincenzo and P. J. Steinhardt (eds.), *Quasicrystals: The State of the Art*, 1991; J. R. Hook and H. E. Hall, *Solid State Physics*, 2d ed., 1995; H. V. Keer, *Principles of Solid State*, 1993; J. N. Sherwood (ed.), *The Plastically Crystalline State (Orientationally-Disordered Crystals)*, 1979.

Crystal absorption spectra

The wavelength or energy dependence of the attenuation of electromagnetic radiation as it passes through a crystal, due to its conversion to other forms of energy in the crystal. When atoms are grouped into an ordered array to form a crystal, their interaction with electromagnetic radiation is greatly modified. Free atoms absorb electromagnetic radiation by transitions between a few electronic states of well-defined energies, leading to absorption spectra consisting of sharp lines. In a crystal, these states are broadened into bands, and the cores of the atoms are constrained to vibrate about equilibrium positions. The ability of electromagnetic radiation to transfer energy to bands and ionic vibrations leads to broad absorption spectra that bear little resemblance to those of the free parent atoms. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; ATOMIC STRUCTURE AND SPECTRA; BAND THEORY OF SOLIDS; LATTICE VIBRATIONS.

The absorption spectrum of a crystal includes a number of distinct features. These are shown in the **illustration** for a semiconducting crystal, gallium arsenide. Absorption spectra of insulating crystals are similar except that electronic absorption is shifted to higher energy. Intrinsic absorption and extrinsic absorption differ according to whether energy would be extracted from electromagnetic radiation even if the crystal were perfect, or if defects or impurities are required. Intrinsic absorption includes lattice absorption in the infrared, electronic absorption by transitions from bonding to antibonding states (valence and conduction bands) in the visible and ultraviolet, and electronic absorption by transitions from core levels in the x-ray region. Extrinsic absorption is much weaker and is generally observable only in the region of transparency of semiconductors and insulators. Free-carrier absorption may be either intrinsic or extrinsic: it is intrinsic to metal and semimetal crystals with their partially filled bands, and extrinsic to semiconductors and insulators with their normally filled or empty bands. Free-carrier processes dominate the absorption spectra of metals from the lowest frequencies throughout the ultraviolet, with lattice and band-to-band absorption processes being much less pronounced. *See* ELECTRIC INSULATOR; METAL; SEMICONDUCTOR.

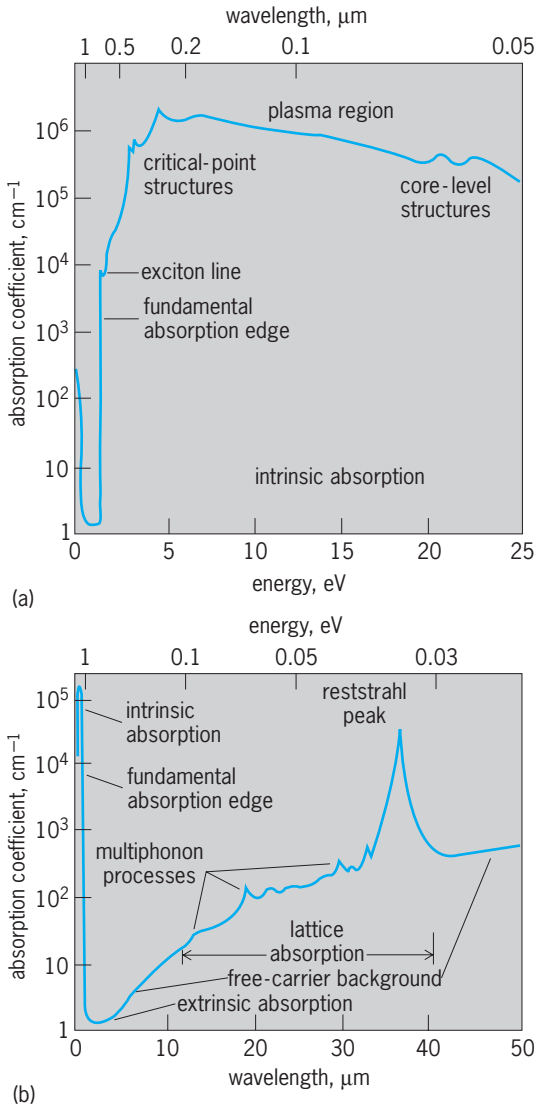

Absorption spectrum of GaAs. The region in which each type of absorption process is important is indicated. (*a*) Spectrum at higher energies. (*b*) Spectrum at lower energies.

Properties. Absorption is measured in terms of the absorption coefficient, which is defined as the relative rate by which electromagnetic radiation is attenuated per unit length. The absorption coefficient is typically expressed in units of cm^{-1}. Absorption coefficients of solids range from less than 2×10^{-7} cm^{-1} in the exceedingly highly refined glasses used in fiber-optic communications systems to higher than 10^6 cm^{-1} for band-to-band electronic transitions in semiconductors and insulators. Absorption is related to reflection through a more fundamental quantity called the dielectric function. Absorption is also related to the index of refraction by a relationship called the Kramers-Kronig transformation, which can be evaluated if the spectral dependence of the absorption coefficient is known over a sufficiently wide frequency range. *See* DISPERSION RELATIONS; REFLECTION OF ELECTROMAGNETIC RADIATION.

Because energy is extracted from electromagnetic radiation by work (force × distance) performed on the crystal, considerable insight into absorption processes can be obtained from simple mechanical considerations. Although force is just equal to the electric field times the charge, the distance that the charge moves depends inversely on mass and frequency and on the bonding configuration. The mass dependence suggests that absorption associated with the heavy ions of the lattice is much weaker than that associated with the light electrons that hold the lattice together, as can immediately be verified from the illustration. The bonding dependence suggests that absorption associated with the free electrons of a metal such as chromium is different from that associated with the tightly bound electrons of an insulator such as glass, as is immediately verified from observation.

Lattice absorption. Lattice absorption arises from the excitation of phonons, which are collective vibrations of the atomic cores about their equilibrium lattice positions. Lattice absorption occurs in the infrared and is responsible for the intrinsic structure in the absorption spectra of crystals below about 0.1 eV (wavelengths above about 10 micrometers). In a diatomic, partially ionic crystal such as gallium arsenide, an electric field that forces the gallium cores one way simultaneously forces the arsenic cores the other way, leading to the generation of transverse optic phonons. The process becomes very efficient at the natural mechanical resonance of the lattice at $8.8 \times 10^{12}\ s^{-1}$ (energy of 36 meV, wavelength of 34 μm), and causes the strong reststrahl absorption peak in the spectrum of illus. *b*. Since the lattice-restoring forces are anharmonic, it is also possible for electromagnetic radiation to create several phonons simultaneously at higher multiples of the reststrahl frequency. These multiphonon processes are responsible for the structure shown from 10 to 30 μm in illus. *b*. *See* IONIC CRYSTALS; MOLECULAR STRUCTURE AND SPECTRA; PHONON.

Bonding-to-antibonding transitions. Electronic transitions between bonding (filled valence) bands and antibonding (empty conduction) bands yield the crystal analog of the line absorption spectra of free atoms. These intrinsic absorption processes dominate the optical behavior of semiconductors and insulators in the visible and ultraviolet regions of the electromagnetic spectrum, as binding energies of most materials are of the order of magnitude of electronvolts per atom. Although analogous transitions also occur in metals, these are generally overwhelmed by free-carrier effects with the spectacular exceptions of copper, silver, and gold, where free-carrier and band-to-band absorption processes are of comparable importance.

The energy of an electron within a band is not constant but is a function of the velocity or momentum of the electron relative to the crystal lattice. This dependence on momentum causes optical structure within a bonding-to-antibonding absorption band. The most readily apparent such feature is the fundamental absorption edge of semiconductors and insulators, which marks the boundary between the range of transparency at lower energies and the strong electronic absorption that occurs at higher energies. The energy of the fundamental absorption edge is determined by the forbidden gap, which is the energy difference between the highest filled valence and lowest empty conduction level of the crystal. The forbidden gap may be nonexistent, as in metals such as iron, lead, and white tin, and in semimetals such as arsenic, antimony, and bismuth; zero, as in the semiconductors gray tin and one of the mercury-cadmium-telluride alloys; in the infrared, as in the semiconductors onductors germanium, silicon, gallium arsenide, and lead sulfide; in the visible or near-ultraviolet, as in the semiconductors gallium phosphide, cadmium sulfide, and gallium nitride; or in the ultraviolet, as in insulators such as diamond, strontium titanate, sodium chloride, silicon dioxide, and lithium fluoride.

Bonding-to-antibonding absorption processes are of two types, direct or indirect, according to whether the initial and final electronic states have the same momentum or whether the electron must absorb or emit a phonon to conserve momentum in the process. The need to involve phonons makes indirect transitions about 100–1000 times less probable than direct transitions. As absorption coefficients for indirect transitions are correspondingly reduced relative to those for direct transitions, it is not surprising that indirect absorption is typically seen only in semiconducting crystals such as germanium, silicon, gallium phosphide, and aluminum arsenide where the highest valence and lowest conduction band states occur at different momenta. In other semiconducting crystals, for example, indium antimonide, gallium arsenide, and zinc oxide, the highest valence and lowest conduction-band states occur at the same momentum and the fundamental absorption edge is direct.

The detailed dependence of absorption on frequency in the vicinity of the fundamental absorption edge depends not only on the variation of electron energy with momentum but also on many-body effects, that is, interactions between the electron excited into the conduction band and the vacancy, or hole, left behind. Because electrons and holes are charges of opposite sign, there is a strong tendency for an electron and hole to bind together in a state like a hydrogen atom. This state is called an exciton. As a result of the positive binding energy of the exciton, the energy needed to create it is less (by the binding energy) than that of the forbidden gap itself. Consequently, electromagnetic radiation is absorbed at energies less than what would be expected in the absence of electron-hole interactions. The lowest possible exciton absorption process produces the single absorption line seen at the fundamental edges of most semiconductors and insulators at low enough temperatures, and is analogous to the lowest absorption line of a hydrogen atom. Excitonic effects also substantially enhance optical absorption

at energies immediately above the fundamental edge. *See* EXCITON.

Direct thresholds analogous to the fundamental absorption edge also occur at higher energies, and cause additional structure in bonding-to-antibonding absorption spectra. These features are generally described in terms of critical points. In three-dimensional energy bands, four types of critical points occur. Local or absolute minima, including the fundamental absorption edge itself and other thresholds where transitions between new band pairs become possible, are of type M_0. Saddle points where the interband energy reaches local minima in two dimensions and a local maximum in the third are of type M_1. Saddle points where the interband energy reaches a local minimum in one dimension and maxima in the other two are of type M_2. Points where the interband energy reaches a local maximum in all three dimensions are of type M_3.

Critical point structure begins to die out at energies greater than 10 eV, as the crystal potential in which the excited electrons propagate becomes small compared to their kinetic energy. This is the plasma region shown in illus. *a*. Although small groups of absorption features typically appear at higher energies (at 20 eV in illus. *a*), these high-energy structures no longer involve bonding states but are associated with the electronic levels of the atomic cores. Being highly localized, core electrons are virtually unaffected by the crystal and (apart from minor shifts in electrostatic potential) can be viewed as being the same as in the free parent atoms. The core-level features of gallium arsenide seen near 20 eV in illus. *a* originate from the gallium 3*d* core electrons; the 3*d* core electrons of arsenic give rise to similar features (not shown) near 45 eV. Deeper-lying core levels give corresponding absorption features at higher energies. Core-level features in both crystalline and free-atom absorption spectra tend to be broad owing to the short lifetimes of core holes created by the absorption process.

Extrinsic absorption. Extrinsic absorption processes involve states associated with deviations from crystal perfection, such as vacancies, interstitials, and deliberately or inadvertently introduced impurities. Since the concentration of imperfections is typically low relative to the concentration of host atoms (10^{12}–10^{19} cm^{-3} compared to 10^{23} cm^{-3}), extrinsic absorption is weak relative to intrinsic absorption. Consequently, as with indirect absorption, extrinsic absorption is typically seen only in spectral regions where the crystal is otherwise transparent. However, as electromagnetic radiation can interact with many absorbing centers in passing through macroscopic lengths of ordinarily transparent material, extrinsic absorption is an important phenomenon. Extrinsic processes are responsible for the colors of most gemstones, for example, the red of ruby and blue of sapphire, and also for the poor optical quality of industrial diamonds. Extrinsic processes are also vitally important in luminescence, which depends almost entirely on the types and concentrations of impurities in a crystal. *See* LUMINESCENCE.

Free-carrier absorption. Unbound charges such as conduction electrons and valence holes can interact with electromagnetic radiation at all frequencies. However, the absorption of electromagnetic radiation requires that the motion induced by the electromagnetic field be coupled to dissipative mechanisms involving charge-charge or charge-lattice collisions. Thus, although all common metals have large densities of free carriers (of the order of magnitude of 10^{23} cm^{-3}), it is the existence of highly efficient loss mechanisms in transition metals such as iron, nickel, chromium, platinum, and palladium that gives them their characteristic neutral gray appearance. In contrast, the loss mechanisms in aluminum and in the noble metals copper, silver, and gold are much less efficient, leading among other effects to substantially higher reflectances for these materials.

Free-carrier effects are typically described phenomenologically by the simple Drude model with electron density and scattering lifetime as the only two parameters. At frequencies below the reciprocal of the scattering lifetime, absorption is roughly independent of frequency and proportional to the direct-current conductance. Above this value, absorption drops roughly as the square of the frequency. The intrinsic absorption properties of electrons in transition metals and the extrinsic absorption properties of free carriers in semiconductors can both be described approximately by the Drude model. However, the noble metals copper and gold cannot be so described, as a result of the strong interband thresholds in the visible that also give these metals their characteristic colors. Free-electron absorption in high-temperature superconductors may be affected similarly by a strongly frequency-dependent loss mechanism, although the detailed nature of this mechanism is still uncertain. *See* FREE-ELECTRON THEORY OF METALS; SUPERCONDUCTIVITY.

David E. Aspnes

Bibliography. M. L. Cohen and J. R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors*, 1988; H. Haug and S. W. Koch, *Quantum Theory of the Optical and Electronic Properties of Semiconductors*, 2d ed., 1993; B. Henderson and G. F. Imbusch, *Optical Spectroscopy of Inorganic Solids*, 1989; K. C. Lee, P. M. Hui, and T. Kushida (eds.), *Optical Properties of Solids*, 1991; G. Martinez (ed.), *Optical Properties of Semiconductors*, 1993.

Crystal counter

A device, more correctly described as a crystal detector, that detects ionizing radiations of all types and is adaptable to measuring neutrons. The sensitive element is a single crystal with a dc resistance normally higher than 10^{12} ohms. The crystals are small and are cut or grown to volumes ranging from less than

1 mm^3 (6×10^{-5} in.^{3}) to approximately 200 mm^3 (1.2×10^{-2} in.^{3}).

Crystal detectors fall into two categories: Certain crystals act as thermoluminescent detectors, of which lithium fluoride (LiF), lithium borate ($Li_2B_4O_7$), and calcium sulfate ($CaSO_4$) are among the best known. Other crystals, for example, cadmium telluride (CdTe) and mercury iodide (HgI_2), act as ionization chambers that deliver either pulses or a dc signal, depending upon the associated electronic circuitry. *See* IONIZATION CHAMBER; THERMOLUMINESCENCE.

Diamond deserves special mention. It is a unique crystal that functions as a thermoluminescent detector or, if suitable contacts are made, as an ionization chamber. The efficiency of the diamond detector in the thermoluminescent or ionization chamber mode is strongly dependent on the impurity atoms included within the crystal lattice, with nitrogen and boron playing dominant roles. Not all diamonds are good detectors; only the rare and expensive natural types IB or IIA are appropriate. Besides being stable and nontoxic, diamond has an additional attractive feature as a detector. As an allotrope of carbon, it has the atomic number $Z = 6$. Human soft tissue has an effective $Z = 7.4$, so that diamond is a close tissue-equivalent material, an essential characteristic for biological dosimetry, for example, in measurements in living organisms. Because of diamond's desirable properties as a detector, economical techniques for the synthesis of diamonds with the correct types and concentrations of impurity atoms are being actively sought for both thermoluminescent and ionization chamber modes. *See* DIAMOND.

Good crystal detectors are insulators and therefore have significant band-gap energies. A large band gap impedes the spontaneous excitation of charge carriers between the valence and conduction bands, thus lowering leakage currents and movement of charge carriers to trapping centers. Room temperature devices are consequently possible. Band gaps for some typical crystal detectors are listed in the **table**. *See* BAND THEORY OF SOLIDS; ELECTRIC INSULATOR; TRAPS IN SOLIDS.

Thermoluminescence. Thermoluminescence can be explained in terms of the band theory of solids. It is to be understood within the framework of the distinctions between fluorescence, which is luminosity continuing only as long as the radiation is continued, and is essentially temperature-independent, and phosphorescence, which is observable luminosity even after the removal of the radiation source.

Band gaps of crystal detectors

Crystal	Band gap, eV	Mode
Diamond	5.4	Thermoluminescence and conduction
LiF	11.0	Thermoluminescence
CdTe	1.45	Conduction
HgI_2	2.15	Conduction

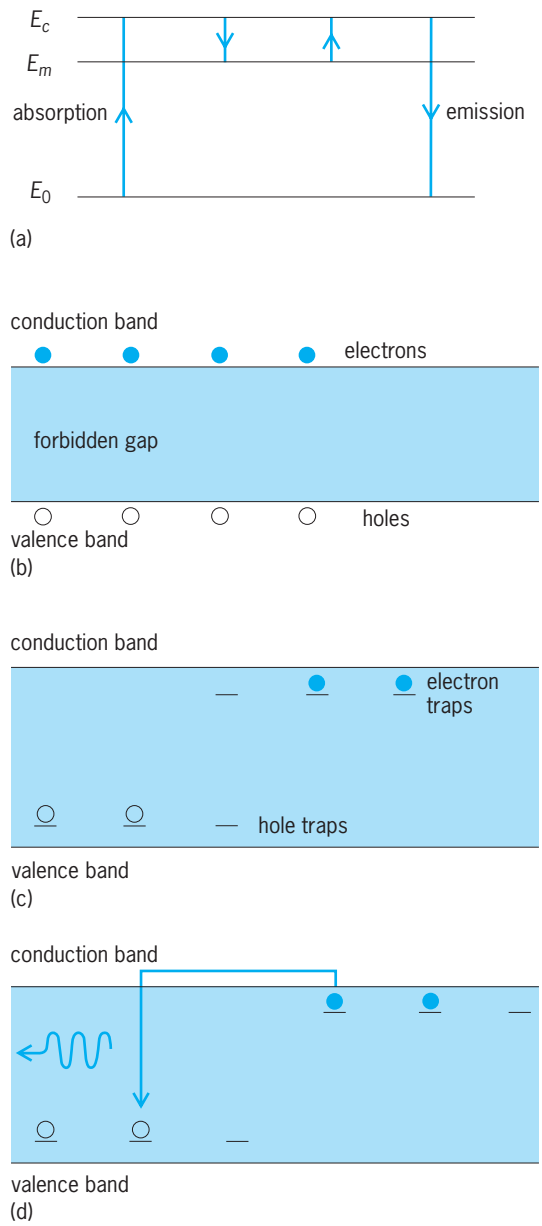

Fig. 1. Crystal band theory of luminescence: (*a*) energy levels; (*b*) conduction and valence bands; (*c*) hole and electron traps; (*d*) photon emission.

Phosphorescence has an intensity decay time which is highly sensitive to the crystal temperature. This situation can arise when the atom or molecule is excited by ionizing radiation—for example, from its ground state E_0 to a metastable state E_m, from which the immediate return to the ground state is prohibited by selection rules (**Fig. 1***a*). If, however, energy $E_c - E_m$ can be supplied to the system, then it can be further excited to a state E_c from which it can return to the ground state directly with the emission of a photon. *See* FLUORESCENCE; LUMINESCENCE; PHOSPHORESCENCE.

Crystals always contain defects in the lattice. Some, such as impurity atoms, can be introduced intentionally or unintentionally during the growing phase. Others are distortions of the lattice. These

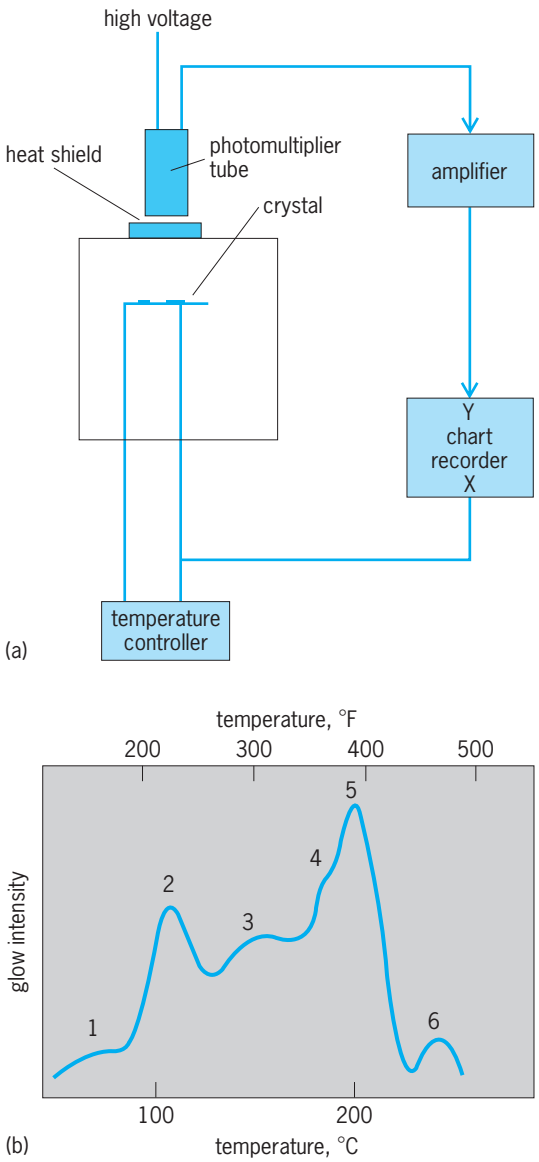

Fig. 2. Thermoluminescence: (*a*) measurement; (*b*) glow curve.

defects perturb the periodicity of the lattice and constitute localized energy levels within the band gap between the valence and conduction bands. After irradiation and excitation of the crystal, the carriers moving through their respective bands (electrons through the conduction band and holes through the valence band; Fig. 1*b*) can become trapped at these defect lattice sites (Fig. 1*c*). Upon heating of the crystal, the trapped electrons (holes) may gain sufficient energy to escape from the trapping centers to the conduction (valence) band, from which they may make a direct transition back to the valence (conduction) band, become retrapped, or combine with a trapped hole (electron). The last alternative, called recombination, can lead to energy release, one form of which is the emission of a photon (Fig. 1*d*). The trapped carriers are known as recombination centers if the emission of photons dominates the energy release of the recombination. The degree of heating required to empty the traps is a function of their depth, and the rate at which they are emptied is a function of the temperature. *See* CRYSTAL DEFECTS; ELECTRON-HOLE RECOMBINATION; HOLE STATES IN SOLIDS.

The crystal is heated at a controlled rate on a metal tray by means of an electric current. The photon emission from the crystal is monitored by a photomultiplier the output of which is directed toward an appropriate recording device (**Fig. 2***a*). The result is a "glow curve" (Fig. 2*b*), the area of which correlates with the number of traps depopulated. These traps were, of course, the traps which were populated during irradiation, their number being directly related to the radiation-field intensity. The integrated light output therefore becomes a direct measure of the total radiation dose. The peaks 1–6 (Fig. 2*b*) are indicative of the various trapping levels within the crystal. *See* DOSIMETER; PHOTOMULTIPLIER.

Conduction mode. An essential feature of the crystal operating as a detector in the conduction mode is the forming of the contacts. The contacts must be ohmic and should not form a Schottky barrier, which would severely impede the charge signal from the crystal. A polarizing voltage is applied, usually to opposite faces of the crystal, and a field of 10^4–10^5 V/cm is used to collect the charges (**Fig. 3**).

A charged particle entering the crystal transfers its kinetic energy to the bulk of the crystal by creating charge carriers (electron-hole pairs). A photon of sufficient energy interacts with the crystal atoms, losing all or part of its energy through the photoelectric effect, the Compton effect, or pair production. In each of these processes, electrons are either liberated or created, and they in turn have their energy dissipated in the bulk of the crystal by creating charge carriers. *See* COMPTON EFFECT; ELECTRON-POSITRON PAIR PRODUCTION; GAMMA-RAY DETECTORS; PHOTOEMISSION.

When the carrier pair is created, the individual carriers move under the influence of the electric field toward the oppositely charged contacts. On arriving at the contacts, the charges can be measured at the output point either as a dc or as a pulse signal, depending upon the circuitry. It is, however, necessary for full

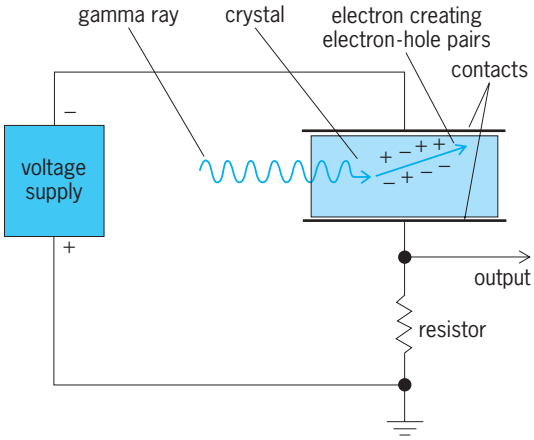

Fig. 3. Crystal detector in conduction mode.

efficiency of the counting system that both types of carriers are collected equally. This makes the property of ion mobility within the crystal very important, and both types should have equal or nearly equal mobility. A condition called hole trapping arises if this criterion is not satisfied. It leads to poor resolution of the overall detecting system, which is also a consequence of high leakage currents or electronic noise within the circuit elements. Because of the high resistance of the crystal, the length of the leads and the screening of all connections must be carefully controlled to reduce the inherent input capacitance of the front end of the system. Resolution can also be improved by cooling the crystal.

Sensitivity and resolution. In the thermoluminescent mode the crystals measure the total dose of the applied radiation, whereas in the conduction mode they measure the instantaneous dose rate; in both cases it is ultimately the crystal itself that limits the sensitivity and resolution of the system. Present methods of synthesis for crystals permit the detection of radiation fields down to nearly background values (0.1 microgray/h or 10^{-5} rad/h) even with crystals as small as 1 mm^3 (6×10^{-5} $in.^3$). Small crystals make detectors possible that are capable of very high spatial resolution. This feature is important in electron radiation therapy, for example, where the isodose curves lie very close together. *See* PARTICLE DETECTOR; RADIATION BIOLOGY; RADIOLOGY.

R. J. Keddy

Bibliography. C. F. Delaney and E. C. Finch, *Radiation Detectors: Physical Principles and Applications*, 1992; K. Kleinknecht, *Detectors for Particle Radiation*, 2d ed., 1999; G. F. Knoll, *Radiation Detection and Measurements*, 3d ed., 1999.

Crystal defects

Departures of a crystalline solid from a regular array of atoms or ions. A "perfect" crystal of sodium chloride (NaCl), for example, would consist of alternating Na^+ and Cl^- ions on an infinite three-dimensional simple cubic lattice, and a simple defect (a vacancy) would be a missing Na^+ or Cl^- ion. There are many other kinds of possible defects, ranging from simple and microscopic, such as the vacancy and other structures (**Fig. 1**), to complex and macroscopic, such as the inclusion of another material, or a surface.

Natural crystals always contain defects, often in abundance, due to the uncontrolled conditions under which they were formed. The presence of defects which affect the color can make these crystals valuable as gems, as in ruby [chromium replacing a small fraction of the aluminum in aluminum oxide (Al_2O_3)]. Crystals prepared in the laboratory will also always contain defects, although considerable control may be exercised over their type, concentration, and distribution.

The importance of defects depends upon the material, type of defect, and properties which are being considered. Some properties, such as density and elastic constants, are proportional to the concentration of defects, and so a small defect concentration will have a very small effect on these. Other properties, such as the color of an insulating crystal or the conductivity of a semiconductor crystal, may be much more sensitive to the presence of small numbers of defects. Indeed, while the term defect carries with it the connotation of undesirable qualities, defects are responsible for many of the important properties of materials, and much of solid-state physics and chemistry and materials science involves the study and engineering of defects so that solids will have desired properties. A defect-free silicon crystal would be of little use in modern electronics; the use of silicon in devices is dependent upon small concentrations of chemical impurities such as phosphorus and arsenic which give it desired electronic properties.

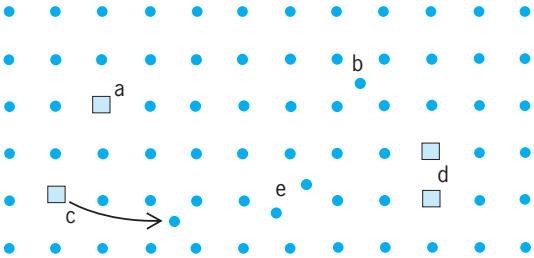

Key:
a = vacancy (Schottky defect)
b = interstitial
c = vacancy-interstitial pair (Frenkel defect)
d = divacancy
e = split interstitial
□ = vacant site

Fig. 1. Some simple defects in a lattice.

Surfaces and interfaces. The importance of surfaces depends very much upon the properties which are being considered and upon the geometry of the specimen. In a specimen no more than a few hundred atoms thick, a significant fraction of the atoms is close to the surface, while in a sample 0.4 in. (10^{-2} m) thick or greater, a much smaller fraction of the atoms is close to the surface. In the latter case, it makes sense to consider bulk properties which are surface-independent and shape-independent. Even in this case, surface effects may be important, as in metals which reflect visible light from their surfaces, but they may generally be separated from bulk effects.

Surfaces are of great importance in determining the chemical properties of solids, since they are in contact with the environment. Such properties include chemical reactions involving the surface, such as corrosion; they also include the role of solids as catalysts in chemical reaction. It is often found that surface atoms reconstruct; that is, they displace and form bonds different from those existing in the bulk. Experimental studies of surfaces are difficult, because of the high vacuum which must be maintained and the relatively small number of atoms

which are involved. *See* CATALYSIS; CORROSION; SURFACE PHYSICS.

An interface is the boundary between two dissimilar solids; it is in a sense an interior surface. For example, a simple metal-oxide-semiconductor (MOS) electronic device has two interfaces, one between metal and semiconductor, and the other between semiconductor and oxide. (In this case the oxide is not crystalline.) Interfaces tend to be chemically and electrically active and to have large internal strains associated with structural mismatch.

Interfaces are of great importance in microelectronics. Modern electronic devices involve very small features and layers of thin semiconducting and insulating films. The success of silicon technology is based in large part on the quality and integrity of the interface between silicon and silicon dioxide. Much of the challenge in developing new generations of semiconductor devices using materials other than silicon lies in finding ways to form high-quality interfaces between different semiconductors and between semiconductors and insulators. *See* INTEGRATED CIRCUITS.

Chemical impurities. An important class of crystal defect is the chemical impurity. The simplest case is the substitutional impurity, for example, a zinc atom in place of a copper atom in metallic copper. Impurities may also be interstitial; that is, they may be located where atoms or ions normally do not exist.

In metals, impurities usually lead to an increase in the electrical resistivity. Electrons which would travel as unscattered waves in the perfect crystal are scattered by impurities. Thus zinc or phosphorus impurities in copper decrease the conductivity. Impurities can play an important role in determining the mechanical properties of solids. As discussed below, an impurity atom can interact with an extended structural defect called a dislocation to increase the strength of a solid; hydrogen, on the other hand, can lead to brittle fracture of some metals. *See* ELECTRICAL CONDUCTIVITY OF METALS; ELECTRICAL RESISTIVITY.

Impurities in semiconductors are responsible for the important electrical properties which lead to their widespread use. The electronic states of all solids fall into energy bands, quasicontinuous in energy. In nonmetals there is a band gap, a region of energy in which no states exist, between filled (valence) and empty (conduction) bands. Because electrons in filled bands do not conduct electricity, perfect nonmetals act as insulators. Impurities (and other defects) often introduce energy levels within the forbidden gap (**Fig. 2**). If these are close in energy to the valence or conduction band, electrons may be thermally excited from a nearby filled level (donor) into the conduction band, or from the valence band into a nearby empty level (acceptor), in the latter case leaving an unoccupied state or hole in the valence band. In both cases, conduction will occur in an applied electric field. Semiconductor devices depend on the deliberate addition (doping) of appropriate impurities. Other impurities may have energy levels which are not close in energy to either

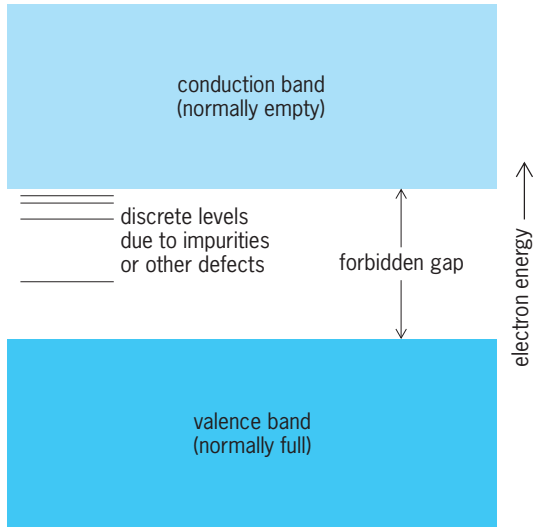

Fig. 2. Band structure of an imperfect nonmetallic crystal, including defect states which occur in the normally forbidden energy gap.

valence or conduction band (so-called deep traps). These may be undesirable if they trap the conducting particles. Thus careful purification is an important part of semiconductor technology. *See* ACCEPTOR ATOM; BAND THEORY OF SOLIDS; DONOR ATOM; HOLE STATES IN SOLIDS; SEMICONDUCTOR; TRAPS IN SOLIDS; ZONE REFINING.

The energy levels associated with impurities and other defects in nonmetals may also lead to optical absorption in interesting regions of the spectrum. Ruby and other gems are examples of this; in addition, the phenomenon of light emission or luminescence is often impurity-related. Solid-state lasers generally involve impurity ions. *See* CRYSTAL ABSORPTION SPECTRA; LASER; LUMINESCENCE.

Structural defects. Even in a chemically pure crystal, structural defects will occur. Thermal equilibrium exists when the Gibbs free energy G is minimized. G is given by Eq. (1), plus a pressure-times-

$$G = E - TS \tag{1}$$

volume term which can often be neglected. Here E is the energy, T is the absolute temperature, and S is the entropy. A perfect crystal has a lower energy E than one containing structural defects, but the presence of defects increases the entropy S. Thus at $T = 0$ K, a perfect crystal is stable, but at higher values of T, the quantity G has its lowest value when structural defects exist. *See* FREE ENERGY.

Crystals may be grown by solidifying the molten substance. If a crystal is cooled slowly (annealed), the number of defects will decrease. However, as the temperature decreases it becomes more difficult for atoms to move in such a way as to "heal" the defects, and in practice there are often more defects than would be expected in thermal equilibrium.

In thermal equilibrium the number of defects at temperatures of interest may be nonnegligible. Application of the methods of statistical mechanics yields

Eq. (2). Here n is the number of defects, N is the

$$\frac{n}{N} = Ae^{-BE/kT} \qquad (2)$$

total number of possible atomic sites, E is the energy required to create a defect, and k is Boltzmann's constant. The quantities A and B are dimensionless constants, generally of order 1, whose actual value depends on the type of defect being considered. For example, for vacancies in monatomic crystals such as silicon or copper, $A = B = 1$. E is typically of order 1 eV; since $k = 8.62 \times 10^{-5}$ eV/K, at $T = 1000$ K (1340°F), $n/N = \exp\ [-1/(8.62 \times 10^{-5} \times 1000)] \simeq 10^{-5}$, or 10 parts per million. For many purposes this fraction would be intolerably large, although as mentioned this number may be reduced by slowly cooling the sample. *See* STATISTICAL MECHANICS.

Simple defects. Besides the vacancy, other types of structural defects exist (Fig. 1). The atom which left a normal site to create a vacancy may end up in an interstitial position, a location not normally occupied. Or it may form a bond with a normal atom in such a way that neither atom is on the normal site but the two are symmetrically displaced from it. This is called a split interstitial. The name Frenkel defect is given to a vacancy-interstitial pair, whereas an isolated vacancy is a Schottky defect. (In the latter case the missing atom may be thought of as sitting at a normal site at the surface.)

Simple defects are often combined. For example, when a small amount of alkaline-earth halide is melted with an alkali halide and the resulting mixture is recrystallized, the mixed solid is found to have a large number of alkali-ion vacancies, approximately one for each alkaline-earth ion. Since the alkaline-earth ions are divalent, the alkali-ion vacancies act as "charge compensators" to make the crystal neutral. Such a crystal has a relatively high electrical conductivity associated with the movement of positive ions into the vacancies (or, alternatively, movement of the vacancies through the crystal). This general phenomenon of ionic conductivity is of practical interest, for example, in attempts to develop solid-state batteries. *See* IONIC CRYSTALS; SOLID-STATE BATTERY.

Simple defects tend to aggregate if thermally or optically stimulated. The F-center in an alkali halide is an electron trapped at a halogen-ion vacancy. Under suitable conditions of optical or thermal stimulation, it is found that F-centers come together to form F_2- or M-centers (two nearest F-centers), F_3- or R-centers (three nearest F-centers), and so on. *See* COLOR CENTERS.

Extended defects and mechanical properties. The simplest extended structural defect is the dislocation. An edge dislocation is a line defect which may be thought of as the result of adding or subtracting a half-plane of atoms (**Fig. 3***a*). A screw dislocation is a line defect which can be thought of as the result of cutting partway through the crystal and displacing it parallel to the edge of the cut (Fig. 3*b*). The lattice displacement associated with a dislocation is called the Burgers vector $\vec{b}$.

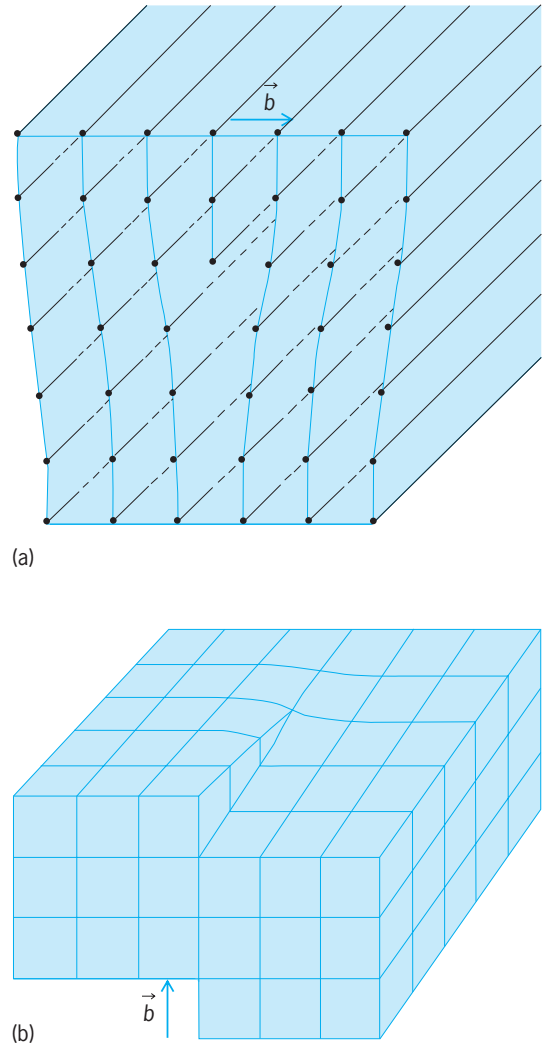

Fig. 3. Extended structural defects. (*a*) Edge dislocation. (*b*) Screw dislocation. In each case the Burgers vector is denoted by $\vec{b}$.

Dislocations generally have some edge and some screw character. Since dislocations cannot terminate inside a crystal, they either intersect the surface or an internal interface (called a grain boundary) or form a closed loop within the crystal.

Dislocations are of great importance in determining the mechanical properties of crystals. A dislocation-free crystal is resistant to shear, because atoms must be displaced over high-potential-energy barriers from one equilibrium position to another. It takes relatively little energy to move a dislocation (and thereby shear the crystal), because the atoms at the dislocation are barely in stable equilibrium. Such plastic deformation is known as slip. *See* PLASTIC DEFORMATION OF METAL.

To extend the earlier argument about the difficulty in preparing perfect crystals, it should be anticipated that most crystals will contain dislocations in ample numbers and that special care would have to be taken to prepare a dislocation-free crystal. The latter is in most cases impractical (although in fact some electronic materials are dislocation-free). Although practical metallurgy developed centuries before

dislocations were known to exist, most strengthening processes are methods for dealing with dislocations.

Dislocations will not be mechanically deleterious if they cannot move. Three major ways of hindering dislocation motion are pinning by dissolved impurities, blocking by inclusions, and tangling. Each of these is described briefly below.

Impurities tend to collect near dislocations if the solid is sufficiently hot that the impurities can move. The presence of these impurities along a dislocation removes the energetic equivalence of the dislocation in different positions; now it requires energy to move the dislocations unless the cloud of dissolved impurities can move as well; and to move them will involve overcoming a large energy barrier. This approach's success is lessened at higher temperatures, when the impurities tend to migrate away.

Small particles of a second substance can "block" dislocation motion. Steel is iron with small inclusions of iron carbide, whose major function is dislocation blocking.

A large, randomly oriented collection of dislocations can tangle and interfere with their respective motions. This is the principle behind work-hardening.

Dislocations have other important properties. The growth of crystals from the vapor phase tends to occur where screw dislocations intersect the surface, for in these regions there are edges rather than planes, and the attractive forces for atoms will be greater. Thus, spiral growth patterns are derived from screw dislocations (**Fig. 4**). *See* CRYSTAL GROWTH.

Acommon planar structural defect is the stacking fault. This can be thought of as the result of slicing a crystal in half, twisting the two pieces with respect to each other, and then reattaching them. Another planar defect, the grain boundary, is the boundary between two crystallites. In some cases it can be described in terms of a parallel set of edge dislocations. *See* TWINNING (CRYSTALLOGRAPHY).

Alloys. A solid containing more than one element is called an alloy, especially when the solid is metallic. There are a vast number of possible alloys, and the present discussion will treat two-component or binary systems as examples.

The term stoichiometry is used to describe the chemical composition of compounds and the resulting crystal structure. The alkali halides are particularly simple, forming compounds with well-defined compositions (for example, NaCl). The situation becomes more complicated for compounds in which one or more components can take on more than one valence state. In particular, in the transition-metal oxides different compositions can occur, with different crystal structures and defects. For example, titanium dioxide (TiO_2) represents a well-defined perfect crystal, whereas TiO_{2-x} would have an oxygen deficiency whose value is described by the subscript x. In this case the deficit oxygens are associated with neither oxygen vacancies nor titanium interstitials, but rather with defect aggregates called shear planes.

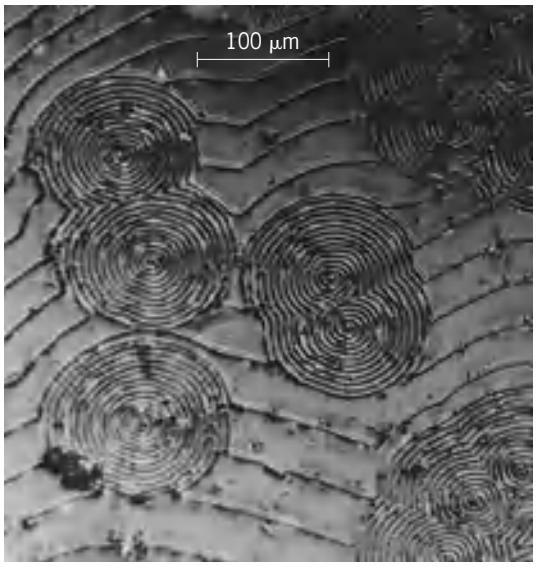

Fig. 4. Crystal-growth spirals on the surface of a silicon carbide crystal. Each spiral step originates in a screw dislocation defect. From the side the surface looks like an ascending ramp wound around a flat cone. (*General Electric Co.*)

If two elements are mixed uniformly throughout the solid, the atomic arrangement may be described as ordered or disordered. For example, an alloy composed of equal numbers of two metal atoms may form an ordered lattice with a regular periodic arrangement of the two atoms with respect to each other, or a disordered lattice in which some of the atoms are on the "wrong" site (**Fig. 5**). The ordered arrangement is most stable at very low temperatures. The temperature at which the structure becomes disordered is called the transition temperature of the order-disorder transformation. *See* ALLOY STRUCTURES.

Defect chemistry. Most of the preceding discussion has treated the simpler types of defects and their static properties, that is, their properties under conditions such that the nature and location of the defects do not change with time. Situations have also been alluded to in which the number, locations, or types of defects change with time, either because of thermal effects (for example, slowly cooling a hot sample and annealing vacancies) or because of

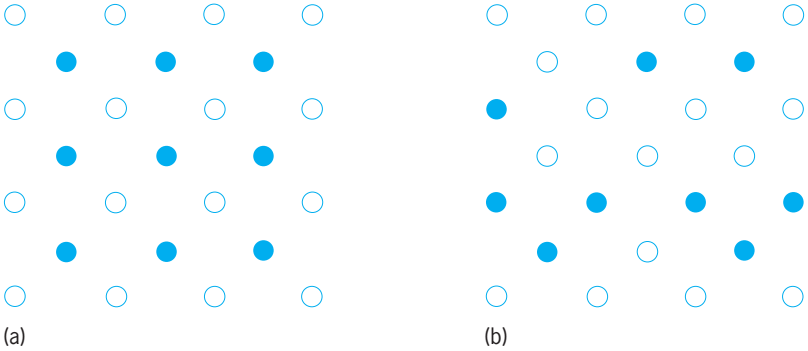

Fig. 5. Binary alloy. (*a*) Ordered lattice. (*b*) Disordered lattice.

external irradiation or defect interactions (causing, for example, the aggregation of alkali-halide *F*-centers).

For both scientific and practical reasons, much of the research on crystal defects is directed toward the dynamic properties of defects under particular conditions, or defect chemistry. Much of the motivation for this arises from the often undesirable effects of external influences on material properties, and a desire to minimize these effects. Examples of defect chemistry abound, including one as familiar as the photographic process, in which incident photons cause defect modifications in silver halides or other materials. Properties of materials in nuclear reactors is another important case. A few examples will be considered here.

Thermal effects. Thermal effects are the easiest to understand from a microscopic point of view: raising the temperature increases the amplitude of atomic vibrations, thereby making it easier for atoms to "hop" from place to place over potential-energy barriers. Almost all dynamic processes are temperature-dependent, occurring more readily at high temperatures. *See* DIFFUSION; LATTICE VIBRATIONS.

Defect creation by irradiation. Radiation can have profound effects on materials in both direct and indirect ways. For example, a beam of high-energy electrons, neutrons, or atomic nuclei (such as alpha particles) can create defects by simple collision processes: the incident particle imparts a portion of its momentum and energy to a lattice atom, thereby releasing the atom from its normal postion. If the released atom has enough energy, it may collide with and release a second lattice atom, thereby creating a second vacancy, and so on; thus one incident particle may lead to the production of a number of defects. Other processes also exist. *See* CHARGED PARTICLE BEAMS; RADIATION DAMAGE TO MATERIALS.

A beam of photons may also create defects through a variety of processes. If the photon energy equals or exceeds the band gap, ionization can occur; that is, electrons can be excited from core or valence bands into the conduction bands as the photons are absorbed. Very energetic photons (gamma rays) may be absorbed or scattered, generating fast electrons, which in turn eject atoms through collision processes such as those occurring with incident particles. *See* GAMMA RAYS.

Lower-energy photons (x-rays, ultraviolet, visible, and so on) have insufficient energy to create defects by such processes. In some solids simple ionization results in defect creation; for example, in alkali halides excitation of an electron on a halogen ion leads after atomic rearrangements to the formation of an excited halogen molecule-ion which is unstable with respect to a translation along its axis. This leads to the creation of a halogen ion vacancy plus an electron (*F*-center) next to a molecule-ion which is crowded into a normal halogen ion position (*H*-center). A replacement sequence can occur (**Fig. 6**) in which each halogen displaces its neighbor in the indicated direction. Eventually this replacement sequence ends, and the result is a pair of isolated *F*- and

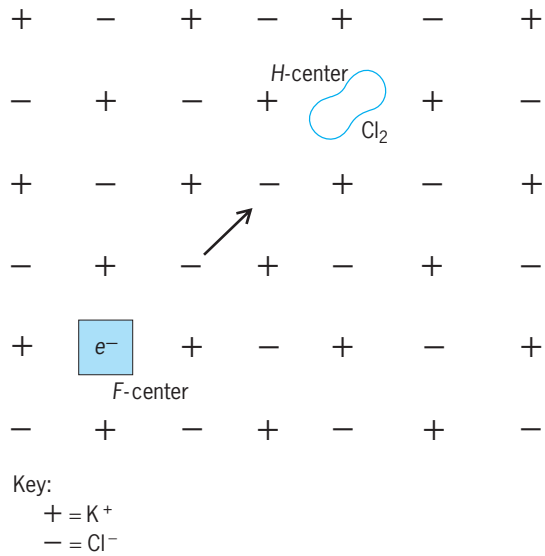

Fig. 6. *F*- and *H*-center formation in KCl after a replacement sequence of halogen-halogen collisions along the direction indicated by the arrow.

H-centers. The first stages of this process can occur in very short times (on the order of 10^{-12} s).

Laser irradiation is often observed to result in crystal destruction, even when the crystal is an insulator and the photon energy is well below the band gap. One way this can occur is first by ionization of defects by the laser, producing a few conduction electrons. These conduction electrons can in turn be accelerated, by the large electric field present in the laser beam, to sufficiently high energy to ionize normal atoms; thus more electrons are available to be accelerated, and a type of cascade process occurs which leads to destruction of the crystal. *See* LASER-SOLID INTERACTIONS.

Effects of radiation on preexisting defects. More subtle effects must often be considered when studying the effects of radiation on device performance. This involves the interaction of radiation with existing (and desirable) defects. Aggregation of *F*-centers occurs when alkali halides are irradiated in the *F*-absorption band. The process involves several steps: (1) The excited *F*-center loses its electron to the conduction band, leaving behind a positively charged vacancy. (2) The conduction electron is subsequently trapped at another *F*-center, forming a (negative) *F′*-center. (3) Under the attractive Coulomb force, the vacancy moves toward the *F′*-center. (4) The extra electron of the *F′*-center returns to the vacancy either by ionization or by tunneling, leading to two *F*-centers closer than normal. (5) The process repeats, until the two *F*-centers are adjacent.

Externally activated defect processes are also found in semiconductors, typically (but not solely) associated with irradiation. Charge-state processes involve the motion of preexisting defects following a change in the charge of the defect. For example, band-gap light generates electrons and holes in silicon, which can be trapped, for example, at a vacancy *V* in silicon. The doubly negative vacancy (V^{2-})

has a considerably smaller activation energy for motion than less negatively charged vacancies, and will consequently migrate more readily. Presumably differences in locations of the neighboring atoms are responsible for the different activation energies of motion.

Recombination processes may also occur. Here an electron and hole recombine at a defect and transfer their (electronic) energy to vibrational energy of the defect, leading to its motion. This effect has been observed, for example, at a recombination junction in gallium arsenide (GaAs), where defects introduced by proton bombardment were destroyed under charge injection. *See* ELECTRON-HOLE RECOMBINATION.

A third possibility is excited-state effects. In this case a defect in an excited electronic state and its surroundings relax into a configuration in which defect motion is easier; in some instances, there may be a zero barrier energy to overcome. One of the best examples of this is the type II F_A-center in alkali halide crystals, an F-center with an alkali impurity (usually Li) as one of its six nearest alkali neighbors (**Fig. 7**). After a photon is absorbed, the system relaxes into a "double-well" configuration with a Cl^- which was initially adjacent to both the Li^+ and the vacancy moving into a "saddle-point" position half the way to the vacancy. The excited electron is then shared by the two small equivalent "wells." When the system returns to the ground state, there is equal probability of the Cl^- returning to its original site or to the original vacancy; in the latter case the vacancy will have moved. Similar processes are thought to occur in semiconductors.

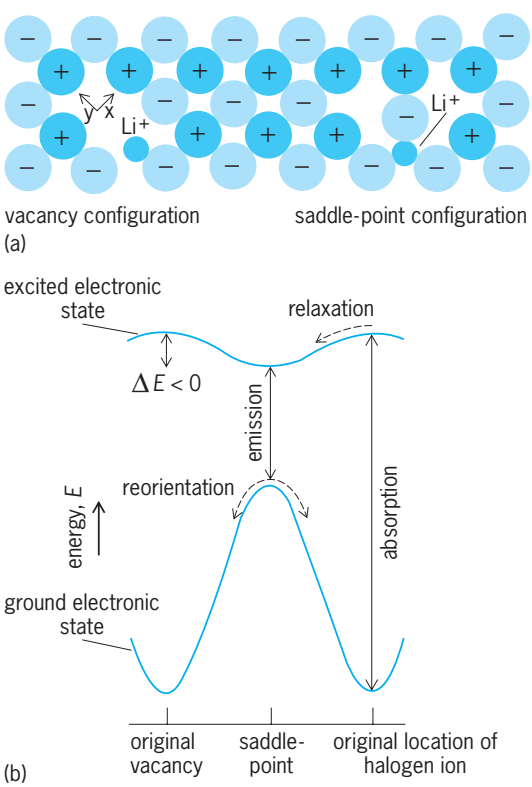

Fig. 7. Reorientation of the type II F_A-center in alkali halides. (*a*) Vacancy and saddle-point configurations. (*b*) Schematic potential energy functions for motion of the neighboring (halogen)$^-$ ion. (*After W. Beall Fowler, ed., Physics of Color Centers, Academic Press, 1968*)

Bistability and metastability. Many semiconductor defects are found to exhibit bistability or metastability. An example is the defect formed in silicon from an interstitial carbon–substitutional carbon pair. This defect has three stable charge states: it can be neutral, it can trap an electron (−), or it can give up an electron (+). Furthermore, it is found to have two distinctly different atomic configurations (whose structures have been proposed but are not known with certainty). Configuration A is stable when the defect is in either the (+) or (−) charge state, while configuration B is stable when the defect is neutral. This phenomenon is called bistability. In addition, in each case the other configuration also has a potential-energy minimum but at a slightly higher energy; this is an example of metastability. Thus when A is stable, B is metastable, and vice versa. At low temperature a metastable state can have a long lifetime if the energy barrier to atomic rearrangement is sufficiently high.

Negative U. Some semiconductor defects possess a characteristic called negative U. This term comes about because the defect behaves as if the electron–electron interaction (U) were attractive. An example is interstitial boron in silicon: this defect can exist in either (−) or (+) charge states, but the neutral state is not stable. This means that more energy is gained when a neutral boron traps an electron than was lost when that electron was removed from another neutral boron: the total energy of the system is lower if half the defects are (+) and half (−) than if all are neutral. In chemistry this phenomenon is known as disproportionation. The occurrence of negative U in semiconductors involves several factors: First, the net Coulomb repulsion between defect electrons is reduced by the interactions with other electrons in the solid (polarization and hybridization). Second, the energies gained from atomic rearrangements can sometimes exceed the remaining Coulomb repulsions, thereby favoring the more ionic charge states. W. Beall Fowler

Topological defects. Defects such as vacancies and interstitials have been known for many years, as have topological defects such as dislocations. The study of two-dimensional systems and their associated defects has led to the discovery of a new phase of matter which has no crystalline positional order and behaves like a fluid under shear, but has well-defined axes like a crystal. An example is the layered liquid-crystal compound 650 BC, which has properties intermediate between a liquid and a crystal.

Order in crystals. To understand topological defects, it is first necessary to understand the nature of order in crystals that have both translational and bond-orientational order. Translational order means that the atoms of the crystal form a well-defined periodic lattice, while orientational order is weaker and means that the directions of the bonds between neighboring atoms are the same throughout the

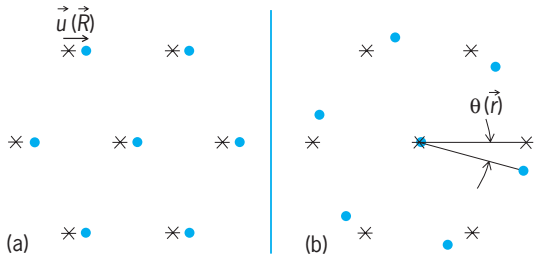

Fig. 8. Translations and rotations of a crystal. The ideal positions are denoted by crosses. (*a*) A patch of crystal translated by $\vec{u}$ from its ideal position with no change of the crystal axes. (*b*) A patch rotated by an angle θ relative to the perfect crystal.

system. A translation of a patch of crystal from its ideal position will spoil translational but not orientational order (**Fig. 8***a*), whereas a rotation will disrupt both (Fig. 8*b*). In principle, orientational order may exist in the absence of translational order, but not vice versa, and a system of point atoms or spherical molecules can exist in three phases: a fully ordered crystal, an intermediate anisotropic fluid with orientational order only, and an isotropic fluid with neither. Systems of nonspherical molecules such as liquid crystals may have many more complex phases. In the great majority of real three-dimensional systems, only the first and last possibilities are realized, with an abrupt melting or freezing transition between states of complete order and complete disorder. However, the intermediate phase with long-range orientational but short-range translational order corresponds to the hexatic phase observed in various liquid-crystal systems and also in xenon adsorbed on graphite. The name hexatic comes from the underlying hexagonal structure of the parent crystal. The three phases can be distinguished by diffraction measurements.

Dislocations and disclinations. A topological defect is one that disrupts a particular order. A crystal, which has two types of order, has two types of topological defect. A translational defect is a dislocation (Fig. 3), and an orientational defect is a disclination. A dislocation may be regarded as a disclination dipole (**Fig. 9**). Circling around the $+\pi/2$ ($-\pi/2$) disclination in a square lattice requires making five (three) right-angle turns and the orientation of the axes is quickly lost. Circling around the dipole involves four right angles, and so this does not disrupt orientational order. However, an attempt to go around the dipole while keeping count of the number of lattice spacings on a path which would close if the lattice were perfect results in failure to return to the starting point by the Burgers vector $\vec{b}$. In other words, the dislocation disrupts translational order only, while a disclination disrupts both translational and orientational order.

Two-dimensional systems. Topological defects do not occur to any significant extent in most bulk crystals in thermal equilibrium. They are present in general as frozen-in metastable states and have a dramatic effect on mechanical properties. In films a few atoms thick and in three-dimensional systems of very weakly coupled layers, they are, however, present in thermal equilibrium and lead to some unusual effects. A graphic way of visualizing a two-dimensional crystal and its excitations and defects is to imagine the crystal lattice drawn on an elastic sheet. This can be distorted in a number of ways, all of which are present to some extent depending on the energy involved, as in Eq. (2). There are the obvious smooth distortions created by stretching and compressing different regions without cutting. Interstitials and vacancies result from the addition or removal of small pieces containing a single atom and regluing the edges. Dislocations are introduced by cutting the sheet, displacing the edges by one lattice spacing, and rejoining the edges, after adding or removing some material as necessary. In a real system this requires diffusing in interstitials or vacancies. Disclinations involve the addition or removal of an angular segment contained between two crystal axes and clearly require a huge amount of energy in the crystalline state.

The equilibrium crystalline state is one with some smooth distortions, some vacancies and interstitials, and a small density of dislocation pairs of equal and opposite Burgers vectors (**Fig. 10***a*). The essential ideas of the role played by topological defects can be understood by focusing on the members of a single

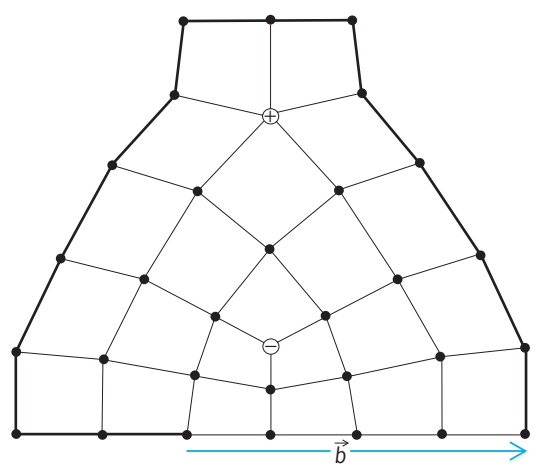

Fig. 9. Disclination dipole in a square lattice. The atom denoted by $\oplus$ ($\ominus$) has five (three) neighbors and corresponds to a $+\pi/2$ ($-\pi/2$) disclination. The heavy line is a perfect lattice closed path.

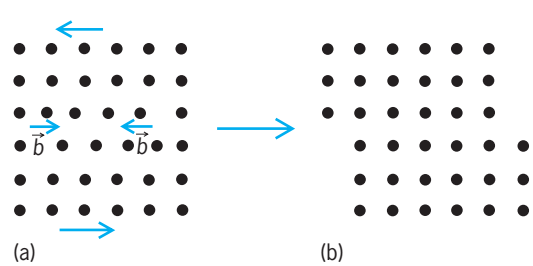

Fig. 10. Defect unbinding transition in a two-dimensional crystal. (*a*) A dislocation pair of equal and opposite Burgers vectors $\vec{b}$ in a square lattice. Above the melting temperature, an infinitesimal shear stress, indicated by the arrows above and below the lattice, will cause the dislocations to move and the crystal to distort. (*b*) Result of this motion.

bound pair. As the temperature is raised, the average spacing between them increases until it equals the spacing between pairs. The density of dislocations remains small. At this stage, the dislocations become independent of each other, positional order is lost, and no force is required to move them about the crystal. The system will behave exactly like a very viscous fluid under a shear stress (Fig. 10*b*).

Provided the density of dislocations and point defects is sufficiently low, the resulting fluid has orientational order. As the temperature is further increased, pairs of equal and opposite disclinations appear in the fluid and their separation will eventually diverge. Above this temperature, all orientational order is lost and the fluid is isotropic.

This discussion assumes that the density of defects remains quite low; in that case the melting of the crystal is via two stages of defect unbinding transitions of the Kosterlitz-Thouless type. Should the density of local defects, such as vacancies, interstitials, and closely bound dislocation pairs, be large, then the crystal melts by an explosive proliferation of these directly to the isotropic fluid. This is the route taken by the great majority of melting transitions both in the bulk and in films, but the defect unbinding mode has been observed experimentally. The theoretical discovery of the hexatic phase led to the experimental identification of the bulk hexatic B and smectic I and F liquid-crystal phases. *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY; LIQUID CRYSTALS.

John M. Kosterlitz

Bibliography. F. Aguillo-Lopez and C. R. Catlow, *Point Defects in Materials*, 1988; G. Benedek (ed.), *Point and Extended Defects in Semiconductors*, 1989; W. Hayes and A. M. Stoneham, *Defects and Defect Processes in Nonmetallic Solids*, 1985; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; S. Mrowec, *Defects and Diffusion in Solids*: *An Introduction*, 1980; T. Mura, *Micromechanics of Defects in Solids*, 2d ed., 1987; F. R. N. Nabarro, *Theory of Dislocations*, 1967, reprint 1987; P. S. Pershan, *Structure of Liquid Crystal Phases*, 1988; K. J. Standburg, Two dimensional melting, *Rev. Mod. Phys.*, 60:161–207, 1988.

Crystal growth

The growth of crystals, of which all crystalline solids are composed. This article discusses the processes by which crystal growth occurs, and methods of artificially growing crystals.

Growth Processes

Crystal growth generally comes about by means of the following processes occurring in series: (1) diffusion of the atoms (or molecules, in the case of molecular crystals such as hydrocarbons or polymers) of the crystallizing substance through the surrounding environment (or solution) to the surface of the crystal, (2) diffusion of these atoms over the surface of the crystal to special sites on the surface, (3) incorporation of atoms into the crystal at these sites, and (4) diffusion of the heat of crystallization away from the crystal surface. The rate of crystal growth may be limited by any of these four steps. The initial formation of the centers from which crystal growth proceeds is known as nucleation. *See* CRYSTALLIZATION; NUCLEATION.

Increasing the supersaturation of the crystallizing component or increasing the temperature independently increases the rate of crystal growth. However, in many physical situations the supersaturation is increased by decreasing the temperature. In these circumstances the rate of crystal growth increases with decreasing temperature at first, goes through a maximum, and then decreases. Often the growth is greatly retarded by traces of certain impurities.

If nucleation occurs in several locations in the medium, the crystals grow isolated from one another for a time. However, if several differently oriented crystals are growing, they may finally impinge on one another, and intercrystalline (grain) boundaries will be formed. At relatively high temperatures, the average grain size in these polycrystalline aggregates increases with time by a process called grain growth, whereby the larger grains grow at the expense of the smaller.

Regular growth. During its growth into a fluid phase, a crystal often develops and maintains a definite polyhedral form which may reflect the characteristic symmetry of the microscopic pattern of atomic arrangement in the crystal. The bounding faces of this form are those which are perpendicular to the directions of slowest growth. How this comes about is illustrated in **Fig. 1**, in which it is seen that the faces *b*, normal to the faster-growing direction, disappear, and the faces *a*, normal to the slower-growing directions, become predominant. Growth forms, like that shown in the figure, are not necessarily equilibrium forms, but they are likely to be most regular when the departures from equilibrium are not large. *See* CRYSTAL STRUCTURE.

The atomic binding sites on the surface of a crystal can be of several kinds. Thus an atom must be more weakly bound on a perfectly developed plane of atoms at the crystal surface (site A in **Fig. 2**) than at a ledge formed by an incomplete plane one atom thick (site B). Atom A binds with only three neighboring atoms, whereas atom B binds to five neighbors.

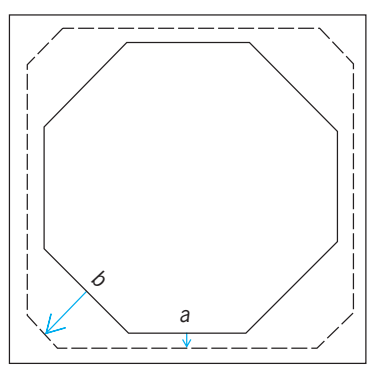

Fig. 1. Schematic representation of cross section of crystal at three stages of regular growth.

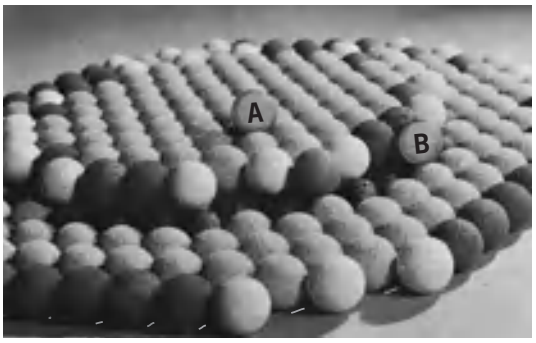

Fig. 2. A model of a crystal surface showing the inequivalent binding sites at which the atoms labeled A and B are located. (*General Electric Co.*)

(An atom in the middle of the island monolayer has bonds with nine neighbors.) Therefore, the binding of atoms in an island monolayer on the crystal surface will be less per atom than it would be within a completed surface layer.

The potential energy of a crystal is most likely to be minimum in forms containing the fewest possible ledge sites. This means that, in a regime of regular crystal growth, dilute fluid, and moderate departure from equilibrium, the crystal faces of the growth form are likely to be densely packed and atomically smooth. There will be a critical size of monolayer, which will be a decreasing function of supersaturation, such that all monolayers smaller than the critical size tend to shrink out of existence, and those which are larger will grow to a complete layer. The critical monolayers form by a fluctuation process. Kinetic analyses indicate that the probability of critical fluctuations is so small that in finite systems perfect crystals will not grow, except at substantial departures from equilibrium. That, in ordinary experience, finite crystals do grow in a regular regime only at infinitesimal departures from equilibrium is explained by the screw dislocation theory. According to this theory, growth is sustained by indestructible surface ledges which result from the emergence of screw dislocations in the crystal face. *See* CRYSTAL DEFECTS.

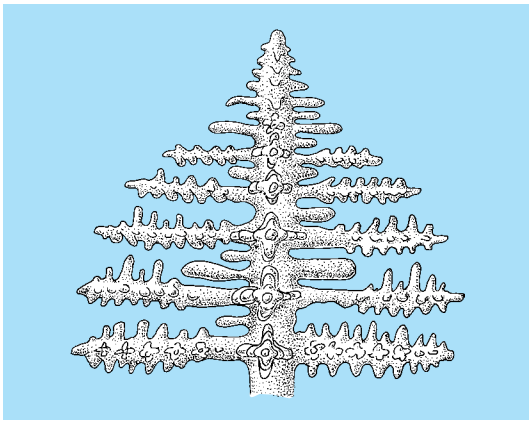

Fig. 3. A dendrite. Growth direction is upward for primary protuberances; side branches grow perpendicularly; side branches on the side branches are also seen growing perpendicularly. (*After W. Kurz and D. J. Fisher, Fundamentals of Solidification, Trans Tech Publications, 1986*)

Irregular growth. When the departures from equilibrium (supersaturation or undercooling) are sufficiently large, the more regular growth shapes become unstable and cellular (grasslike) or dendritic (treelike) morphologies develop. Essentially, the development of protuberances on an initially regular crystal permits more efficient removal of latent heat or of impurities, but at the cost of higher interfacial area and the associated excess surface energy. In the cellular structure, the protuberances align, forming an advancing front and rejecting impurities laterally into the spaces between protuberances. In the dendritic morphology, these protuberances become narrower and develop side branches, which are other protuberances growing laterally (**Fig. 3**). When the supersaturation becomes so great that the energy associated with the increase in interfacial area is unimportant, protuberances proliferate and the crystal grows in a multibranched form that is even more complicated; its shape is characterized by fractal geometry. This limit, in which atoms stick permanently wherever they land, is called diffusion-limited aggregation because steps 2–4 in the series of processes outlined above occur very rapidly, and the rate of crystal growth is limited by diffusion (step 1). *See* FRACTALS.

Michael J. Aziz; David Turnbull

Growth Methods

The advent of semiconductor-based technology generated a demand for large, high-quality single crystals, not only of semiconductors but also of associated electronic materials. With increasing sophistication of semiconductor devices, an added degree of freedom in materials properties was obtained by varying the composition of major components of the semiconductor crystal over very short distances. Thin, multilayered single-crystal structures, and even structures that vary in composition both normal and lateral to the growth direction, are often required.

Bulk single-crystal growth. Bulk single crystals are usually grown from a liquid phase. The liquid may have approximately the same composition as the solid; it may be a solution consisting primarily of one component of the crystal; or it may be a solution whose solvent constitutes at most a minor fraction of the crystal's composition. **Figure 4** shows the phase diagram and corresponding component partial pressures for a hypothetical binary compound. The stoichiometry range of Fig. 4*a* results from the deviation from the 1-to-1 composition caused by point defects in the crystal structure. Nonstoichiometry, high partial pressures, and relatively high point-defect concentrations compared to growth at lower temperatures are all disadvantages of melt growth. Nevertheless, melt growth is used for virtually all bulk semiconductor crystals and for many other industrial crystals. *See* NONSTOICHIOMETRIC COMPOUNDS; PHASE EQUILIBRIUM.

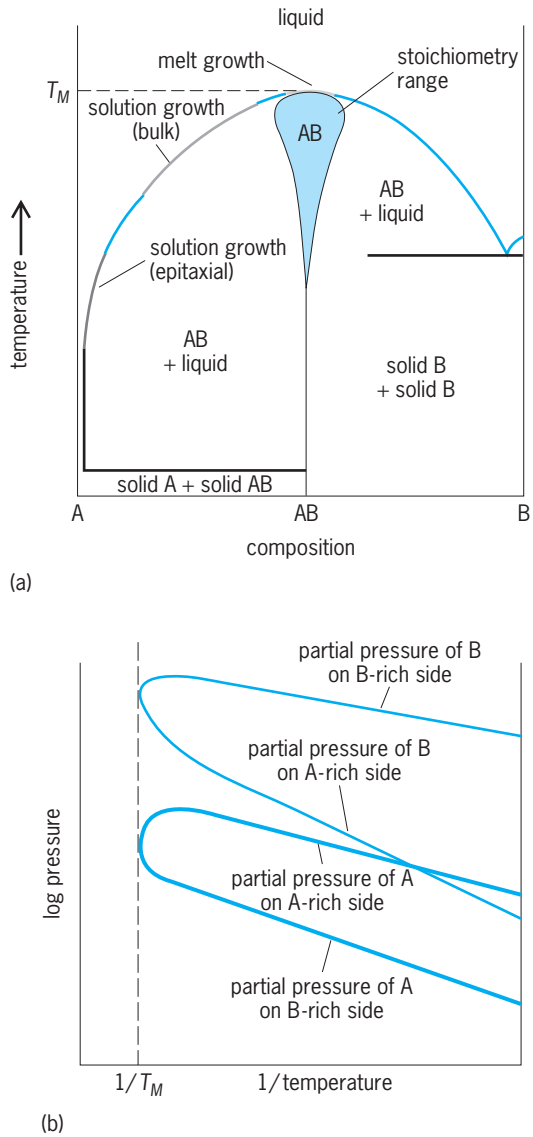

Fig. 4. Hypothetical binary system AB having simple compound formation. (*a*) Phase diagram. (*b*) Corresponding partial pressure curves. The extent of the stoichiometry range is greatly exaggerated for illustration.

Czochralski method. The most important bulk crystal growth technique is the crystal-pulling or Czochralski method, in which a rotating seed crystal is dipped into the melt (**Fig. 5**). Rotation reduces radial temperature gradients, and slow withdrawal of the rotating seed results in growth of a cylinder of single-crystal material. The conditions for optimum growth vary widely, and pulling rates range up to a few inches per hour. Crystal diameter and length depend upon the details of the temperature and pulling rate, and the dimensions of the melt container. Crystal quality depends very critically upon minimization of temperature gradients that enhance the formation of dislocations. Pulled silicon crystals 6 in. (15 cm) in diameter are important for the semiconductor industry. Ruby, sapphire, and group 13–15 compound semiconductor crystals are among the many crystals that are routinely grown by the Czochralski technique.

Melt-growth variations. There are several other variations of the melt-growth technique. One of them is the zone-refining method, in which a narrow molten region is passed through a solid rod of single-crystal or polycrystalline material. Another is the Bridgman technique, in which single-crystal growth is achieved for some materials by moving the melt through a temperature gradient (**Fig. 6**). Generally the container has a restricted region where the first solidification takes place as the liquidus temperature moves past it. An initially formed crystallite in that region propagates with the moving liquidus front, and a single

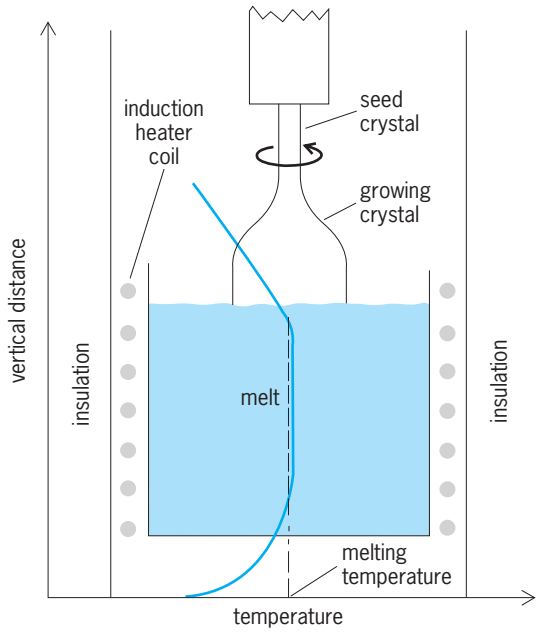

Fig. 5. Czochralski crystal growth and temperature distribution.

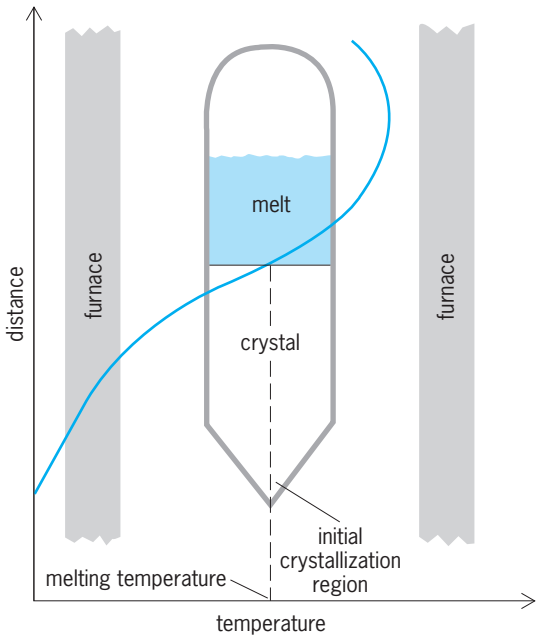

Fig. 6. Vertical Bridgman crystal and temperature distribution.

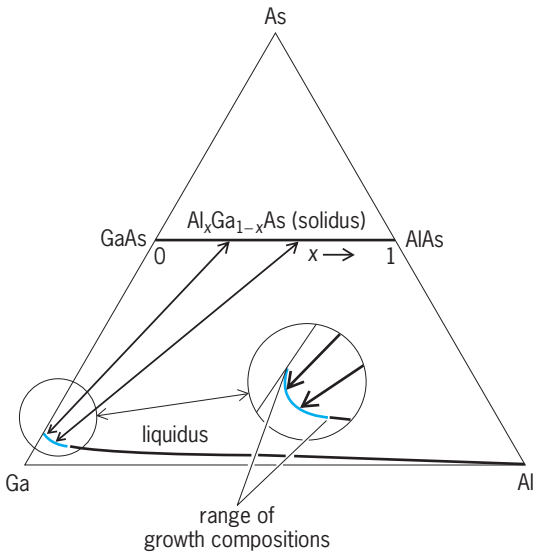

Fig. 7. An isotherm in the aluminum-gallium-arsenic (Al-Ga-As) ternary phase diagram. Each corner represents a pure element and each point within the diagram a unique composition, all at a given temperature. Each composition on the liquidus is in equilibrium with one composition on the solidus. The tie lines (lines with arrowheads) show two such sets of points.

crystal consisting of the entire volume of the melt can result. For some variants of this technique, the liquid is rich in the less volatile component of the solid compound. Growth is driven by maintaining the partial pressure of the more volatile component as required to maintain the liquidus composition. *See* ZONE REFINING.

Solvent methods. Other solution growth methods utilize a solvent in which the crystal components may have extensive solubility, but which itself is not usually appreciably soluble in the crystal. Often, small seed crystals are used to initiate growth which is maintained by subsequent cooling of a saturated solution or diffusion of components through a temperature gradient from a source nutrient. Garnets and other magnetic oxides are often grown from molten salts; quartz crystals are routinely grown from aqueous solution at high temperature and pressure. *See* CRYSTAL WHISKERS; FERRIMAGNETIC GARNETS; SINGLE CRYSTAL.

Epitaxial growth. The evolution of methods for the growth of very thin but very high-quality epitaxial layers has resulted largely from the need for such layers of semiconductors and magnetic garnets.

Liquid-phase epitaxy. The technique most closely related to the methods used for bulk crystal growth is liquid-phase epitaxy. For a typical binary semiconductor, growth is done onto a substrate single-crystal wafer from a solution rich in the component with the lowest partial pressure (Fig. 4). For a binary compound, the grown layers may differ only in impurity concentrations to modify their electrical characteristics. More often, multilayered structures with layers differing in major component composition but having the same crystal structure and lattice parameter are required. The simplest example of liquid-phase epitaxy with major composition changes in layers is the growth of layers of aluminum gallium arsenide ($Al_xGa_{1-x}As$; $1>x>0$) on a gallium arsenide (GaAs) substrate. In this example, for all x, the solid solution $Al_xGa_{1-x}As$ has essentially the same lattice parameter, so that this is a naturally lattice-matched system. An isotherm of the ternary phase diagram for this system is shown in **Fig. 7**. A typical range of growth compositions can yield layers with very different solid compositions.

Growth by liquid-phase epitaxy is done in an apparatus in which the substrate wafer is sequentially brought into contact with solutions that are at the desired compositions and may be supersaturated or cooled to achieve growth. For crystalline solid solutions other than $Al_xGa_{1-x}As$, very precise control over solution compositions is required to achieve a lattice match. Typically, structures grown by liquid-phase epitaxy have four to six layers ranging widely in composition and having thickness from 10^{-7} to 10^{-6} m.

Molecular-beam epitaxy. The desirability of highly reproducible growth and even thinner epitaxial layers of 13–15 compounds on large wafer areas has led to the development of molecular-beam epitaxy and several forms of chemical-vapor deposition. Molecular-beam epitaxy is an ultrahigh-vacuum technique in which beams of atoms or molecules of the constituent elements of the crystal provided by heated effusion ovens, impinge upon a heated substrate crystal (**Fig. 8**). It has been used for epitaxial layers as thin as 0.5 nanometer (**Fig. 9**). Molecular-beam epitaxy has also been used for group 12–16 compounds and for silicon. *See* ARTIFICIALLY LAYERED STRUCTURES; MOLECULAR BEAMS.

The ability to achieve epitaxy of compounds under vacuum conditions requires that it be possible to obtain a nominally stoichiometric crystal from

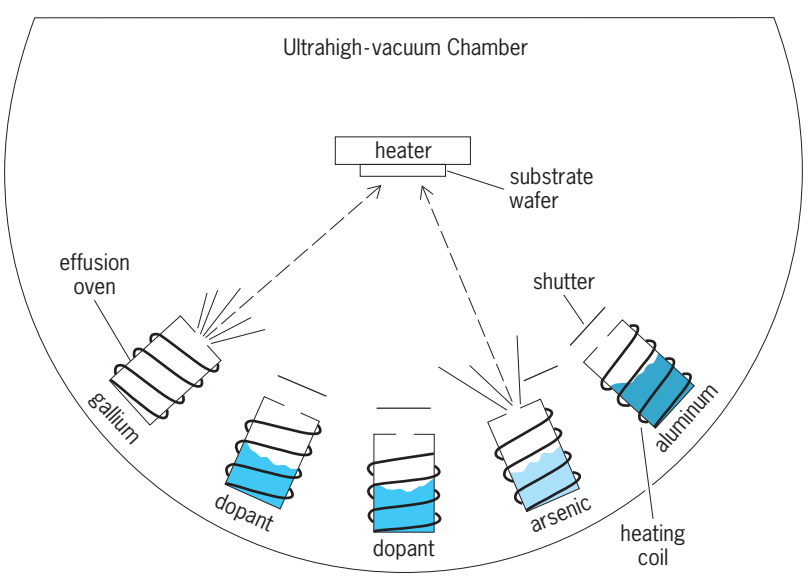

Fig. 8. Molecular-beam epitaxy apparatus.

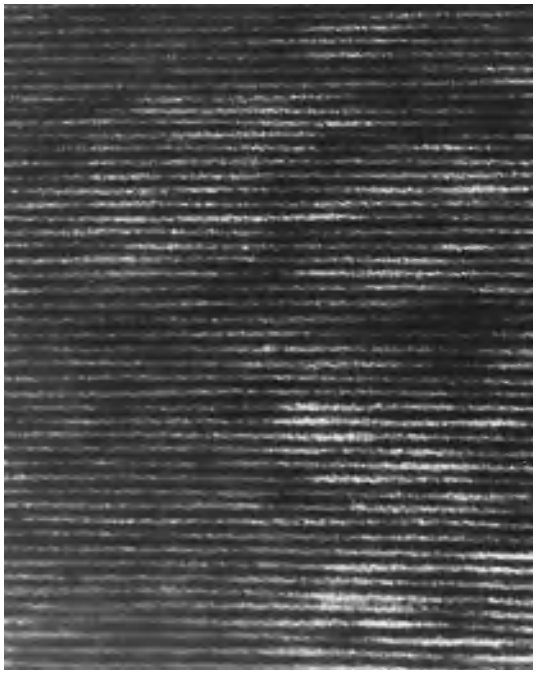

Fig. 9. Cross section of a gallium arsenide–aluminum arsenide (GaAs–AlAs) single crystal with individual gallium arsenide and aluminum arsenide layers each about 5 × 10^{-10} m thick.

impingement of a flux of component atoms and molecules whose net composition is far from that at stoichiometry. It is also required that surface adsorption and mobility of the impinging components be sufficient to ensure that atoms finally reside in their most stable lattice sites. How this can happen is best illustrated with gallium arsenide. Growth of gallium arsenide is achieved by impinging Ga atoms and As_2 or As_4 molecules onto a heated GaAs wafer. The Ga atoms essentially all stick to the surface because of the relatively low Ga vapor pressure at the growth temperature of 900–1300°F (500–700°C). Only those As atoms that interact with a Ga atom stay on the surface, so that the growing layer is stoichiometric provided the As-to-Ga ratio striking the surface is somewhat greater than unity. Crystalline solid solutions such as $Al_xGa_{1-x}As$ are obtained by adding aluminum (Al) to the beams with an Al-containing effusion oven (Fig. 8). Shutters are used to block the beams and permit changes in the beam composition for multilayer growth. *See* GALLIUM.

Chemical-vapor deposition. Chemical-vapor deposition of epitaxial layers is also often used for semiconductors. Again, the group 13–15 compounds are of major importance. The growth of GaAs will be used for illustration. One of the most common chemical-vapor deposition methods is a hot-wall procedure in which the growing layer is generated by reaction (1). The

$$As_4 + 4GaCl + 2H_2 \rightleftharpoons 4GaAs + 4HCl \qquad (1)$$

As_4 is generated either by the thermal decomposition of arsine, AsH_3, as in reaction (2), or by reaction (3). (GaCl) is generated by reaction (4). The

$$4AsH_3 \longrightarrow 6H_2 + As_4 \qquad (2)$$

$$4AsCl_3 + 6H_2 \longrightarrow As_4 + 12HCl \qquad (3)$$

$$2HCl + 2Ga \longrightarrow 2GaCl + H_2 \qquad (4)$$

reactants are introduced into a heated fused-silica reaction tube with a carrier gas, usually molecular hydrogen (H_2) or nitrogen (N_2). A temperature gradient is maintained so that the substrate seed is downstream of the decomposition region for reactions (1), (2), and (3), and is at a temperature that favors growth of GaAs by reaction (1). Crystalline solid solutions of a variety of 13–15 compounds can be grown by this process. Layers of $Al_xGa_{1-x}As$ are not generally grown this way because of thermodynamic limitations on composition control, and the very high reactivity of Al-containing vapor species with most hot-wall reactor materials.

Metal organic chemical-vapor deposition. An alternate process is the metal organic (MOCVD) process, in which metal organic compounds are decomposed near the surface of the heated substrate wafer. This method depends upon the fact that a very large variety of MR_x compounds (where M is a metal and R is an organic radical) such as $Ga(CH_3)_3$, $As(CH_3)_4$, and $Zn(CH_3)_2$ are available. The organic radical may be a methyl, ethyl, or more complex group. Generally, a compound is available as a liquid with a convenient vapor pressure so that it can be transported by saturating a stream of H_2. Most often, group 15 elements are transported as the hydrides. Thus, typically, indium phosphide (InP) is grown by reaction (5).

$$In(C_2H_5)_3 + PH_3 \longrightarrow InP + 3C_2H_6 \qquad (5)$$

Growth by MOCVD is most often done in a cold-wall chamber containing an induction-heated platform for the substrate wafer. Decomposition of the reactants apparently occurs in a boundary-layer region near the crystal surface. The MOCVD technique has proved highly successful for the growth of GaAs, $Al_xGa_{1-x}As$, InP, and $Ga_xIn_{1-x}P_yAs_{1-y}$, among others, and its versatility for multilayer growth of very thin, very high-quality layers is comparable to that of molecular-beam epitaxy. *See* SEMICONDUCTOR; SEMICONDUCTOR HETEROSTRUCTURES. Morton B. Panish

Bibliography. H. Arend and J. Hulliger (eds.), *Crystal Growth in Science and Technology*, 1990; J. C. Brice, *Crystal Growth Processes*, 1986; D. F. J. Hurle (ed.), *Handbook of Crystal Growth*, 1993; I. V. Markov, *Crystal Growth for Beginners*, 1995; M. Panish and H. Temkin, *Gas Source Molecular Beam Epitaxy*, 1993; H. Pfeiffer, T. Kwpsch, and W. Haubereisser, *Microscopic Theory of Crystal Growth*, 1990; L. M. Sander, Fractal growth processes, *Nature*, 322:789–793, 1986; J. P. van der Eerden, *Fundamentals of Crystal Growth*, 1993; A. W. Vere, *Crystal Growth: Principles and Progress*, 1988.

Crystal optics

The study of the propagation of light, and associated phenomena, in crystalline solids. The propagation of light in crystals can actually be so complicated that not all the different phenomena are yet completely understood, and not all theoretically predicted phenomena have been demonstrated experimentally.

For a simple cubic crystal the atomic arrangement is such that in each direction through the crystal the crystal presents the same optical appearance. The atoms in anisotropic crystals are closer together in some planes through the material than in others. In anisotropic crystals the optical characteristics are different in different directions. In classical physics the progress of an electromagnetic wave through a material involves the periodic displacement of electrons. In anisotropic substances the forces resisting these displacements depend on the displacement direction. Thus the velocity of a light wave is different in different directions and for different states of polarization. The absorption of the wave may also be different in different directions. *See* DICHROISM; TRICHROISM.

In an isotropic medium the light from a point source spreads out in a spherical shell. The light from a point source embedded in an anisotropic crystal spreads out in two wave surfaces, one of which travels at a faster rate than the other. The polarization of the light varies from point to point over each wave surface, and in any particular direction from the source the polarization of the two surfaces is opposite. The characteristics of these surfaces is opposite. The characteristics of these surfaces can be determined experimentally by making measurements on a given crystal.

For a transparent crystal the theoretical optical behavior is well enough understood so that only a few measurements need to be made in order to predict the behavior of a light beam passing through the crystal in any direction. It is important to remember that the velocity through a crystal is not a function of position in the crystal but only of the direction through the lattice. For information closely related to the ensuing discussion. *See* CRYSTAL STRUCTURE; POLARIZED LIGHT; REFRACTION OF WAVES.

Index ellipsoid. In the most general case of a transparent anisotropic medium, the dielectric constant is different along each of three orthogonal axes. This means that when the light vector is oriented along each direction, the velocity of light is different. One method for calculating the behavior of a transparent anisotropic material is through the use of the index ellipsoid, also called the reciprocal ellipsoid, optical indicatrix, or ellipsoid of wave normals. This is the surface obtained by plotting the value of the refractive index in each principal direction for a linearly polarized light vector lying in that direction (**Fig. 1**). The different indices of refraction, or wave velocities associated with a given propagation direction, are then given by sections through the origin of the coordinates in which the index ellipsoid is drawn. These sections are ellipses, and the major and minor axes of the ellipse represent the fast and slow axes for light proceeding along the normal to the plane of the ellipse. The length of the axes represents the refractive indices for the fast and slow wave, respectively. The most asymmetric type of ellipsoid has three unequal axes. It is a general rule in crystallography that no property of a crystal will have less symmetry than the class in which the crystal belongs. In other words, if a property of the crystal had lower symmetry, the crystal would belong in a different class. Accordingly, there are many crystals which, for example, have four- or sixfold rotation symmetry about an axis, and for these the index ellipsoid cannot have three unequal axes but is an ellipsoid of revolution. In such a crystal, light will be propagated along this axis as though the crystal were isotropic, and the velocity of propagation will be independent of the state of polarization. The section of the index ellipsoid at right angles to this direction is a circle. Such crystals are called uniaxial and the mathematics of their optical behavior is relatively straightforward. *See* CRYSTALLOGRAPHY.

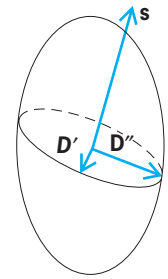

Fig. 1. Index ellipsoid, showing construction of directions of vibrations of D vectors belonging to a wave normal s. (*After M. Born and E. Wolf, Principles of Optics, 6th ed., Pergamon Press, 1980*)

Ray ellipsoid. The normal to a plane wavefront moves with the phase velocity. The Huygens wavelet, which is the light moving out from a point disturbance, will propagate with a ray velocity. Just as the index ellipsoid can be used to compute the phase or wave velocity, so can a ray ellipsoid be used to calculate the ray velocity. The length of the axes of this ellipsoid is given by the velocity of the linearly polarized ray whose electric vector lies in the axis direction. *See* PHASE VELOCITY.

The ray ellipsoid in the general case for anisotropic crystal is given by Eq. (1), where α, β, and γ are the

$$\alpha^2x^2 + \beta^2y^2 + \gamma^2z^2 = 1 \qquad (1)$$

three principal indices of refraction and where the velocity of light in a vacuum is taken to be unity. From this ellipsoid the ray velocity surfaces or Huygens wavelets can be calculated as just described. These surfaces are of the fourth degree and are given by

Eq. (2). In the uniaxial case $\alpha = \beta$ and Eq. (2) factors

$$(x^2+y^2+z^2)\left(\frac{x^2}{\alpha^2}+\frac{y^2}{\beta^2}+\frac{z^2}{\gamma^2}\right) -\frac{1}{\alpha^2}\left(\frac{1}{\beta^2}+\frac{1}{\gamma^2}\right)x^2-\frac{1}{\beta^2}\left(\frac{1}{\gamma^2}+\frac{1}{\alpha^2}\right)y^2 -\frac{1}{\gamma^2}\left(\frac{1}{\alpha^2}+\frac{1}{\beta^2}\right)z^2+\frac{1}{\alpha^2\beta^2\gamma^2}=0 \quad (2)$$

into Eqs. (3). These are the equations of a sphere and

$$x^2+y^2+z^2=\frac{1}{\alpha^2} \qquad \gamma^2(x^2+y^2)+\alpha^2z^2=1 \quad (3)$$

an ellipsoid. The z axis of the ellipsoid is the optic axis of the crystal.

The refraction of a light ray on passing through the surface of an anisotropic uniaxial crystal can be calculated with Huygens wavelets in the same manner as in an isotropic material. For the ellipsoidal wavelet this results in an optical behavior which is completely different from that normally associated with refraction. The ray associated with this behavior is termed the extraordinary ray. At a crystal surface where the optic axis is inclined at an angle, a ray of unpolarized light incident normally on the surface is split into two beams: the ordinary ray, which proceeds through the surface without deviation; and the extraordinary ray, which is deviated by an angle determined by a line drawn from the center of one of the Huygens ellipsoidal wavelets to the point at which the ellipsoid is tangent to a line parallel to the surface. The construction is shown in **Fig. 2**. The two beams are oppositely linearly polarized. When the incident beam is inclined at an angle ϕ to the normal, the ordinary ray is deviated by an amount determined by Snell's law of refraction, the extraordinary ray by an amount which can be determined in a manner similar to that of the normal incidence case already described. The plane wavefront in the crystal is first found by constructing Huygens wavelets as shown in **Fig. 3**. The line from the center of the wavelet to the point of tangency gives the ray direction and velocity.

The relationship between the normal to the wavefront and the ray direction can be calculated algebraically in a relatively straightforward fashion. The extraordinary wave surface, or Huygens wavelet, is given by Eq. (3), which can be rewritten as Eq. (4),

$$\epsilon^2(x^2+y^2)+\omega^2z^2=1 \quad (4)$$

where ϵ is the extraordinary index of refraction and ω the ordinary index. A line from the center of this ellipsoid to a point (x_1, y_1, z_1) on the surface gives the velocity of a ray in this direction. The wave normal corresponding to this ray is found by dropping a perpendicular from the center of the ellipsoid to the plane tangent at the point (x_1, y_1, z_1). For simplicity consider the point in the plane $y = 0$. The tangent

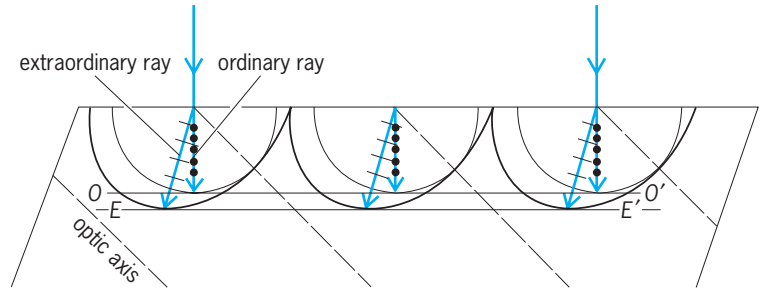

Fig. 2. Huygens construction for a plane wave incident normally on transparent calcite. If one proceeds to find the common tangents to the secondary wavelets shown, the two plane waves labeled *OO′* and *EE′* are obtained. (*After F. A. Jenkins and H. E. White, Fundamentals of Optics, 4th ed., McGraw-Hill, 1976*)

at the point x_1z_1 is given by Eq. (5). The slope of the

$$\epsilon^2xx_1+\omega^2zz_1=1 \quad (5)$$

normal to this line is given by Eq. (6). The tangent

$$\frac{z}{x}=\frac{\omega^2z_1}{\epsilon^2x_1} \quad (6)$$

of the angle between the optic axis and the wave normal is the reciprocal of this number, as shown in Eq. (7), where ψ is the angle between the ray and the

$$\tan\varphi=\frac{\epsilon^2x_1}{\omega^2z_1}=\frac{\epsilon^2}{\omega^2}\tan\psi \quad (7)$$

optic axis. The difference between these two angles, π, can be calculated from Eq. (8). This quantity is a maximum when Eq. (9) holds.

$$\tan\tau=\frac{\tan\varphi-\tan\psi}{1+\tan\varphi\tan\psi} \quad (8)$$

$$\tan\varphi=\pm\frac{\epsilon}{\omega} \quad (9)$$

One of the first doubly refracting crystals to be discovered was a transparent variety of calcite called Iceland spar. This uniaxial crystal cleaves into slabs in which the optic axis makes an angle of 45° with one pair of surfaces. An object in contact with or a few inches from such a slab will thus appear to be doubled. If the slab is rotated about a normal to

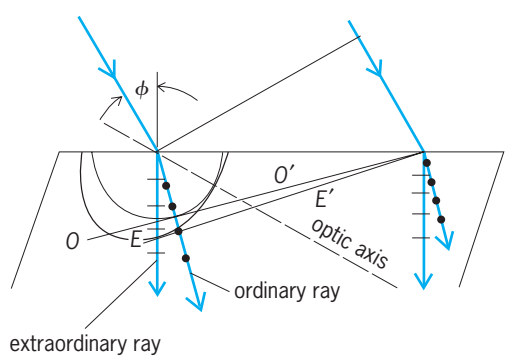

Fig. 3. Huygens construction when the optic axis lies in the plane of incidence. (*After F. A. Jenkins and H. E. White, Fundamentals of Optics, 4th ed., McGraw-Hill, 1976*)

the surface, one image rotates about the other. For the sodium D lines at 589 nanometers, the indices for calcite are given by Eqs. (10). From these, the

$$\epsilon = 1.486 \qquad \omega = 1.659 \tag{10}$$

maximum angle τ_{max} between the wave normal and the ray direction is computed to be 6°16′. The wave normal at this value makes an angle of 41°52′ with the optic axis. This is about equal to the angle which the axis makes with the surface in a cleaved slab. Accordingly, the natural rhomb gives nearly the extreme departure of the ray direction from the wave normal.

Interference in polarized light. One of the most interesting properties of plates of crystals is their appearance in convergent light between pairs of linear, circular, or elliptical polarizers. An examination of crystals in this fashion offers a means of rapid identification. It can be done with extremely small crystals by the use of a microscope in which the illuminating and viewing optical systems are equipped with polarizers. Such a polarizing microscope is a common tool for the mineralogist and the organic chemist.

In convergent light the retardation through a birefringent plate is different for each direction. The slow and fast axes are also inclined at a different angle for each direction. Between crossed linear polarizers the plate will appear opaque at those angles for which the retardation is an integral number of waves. The locus of such points will be a series of curves which represent the characteristic interference pattern of the material and the angle at which the plate is cut. In order to calculate the interference pattern, it is necessary to know the two indices associated with different angles of plane wave propagation through the plate. This can be computed from the index ellipsoid. *See* INTERFERENCE OF WAVES.

Ordinary and extraordinary indices. For the uniaxial crystal there is one linear polarization direction for which the wave velocity is always the same. The wave propagated in this direction is called the ordinary wave. This can be seen directly by inspection of the index ellipsoid. Since it is an ellipsoid of revolution, each plane passing through the center will produce an ellipse which has one axis equal to the axis of the ellipsoid. The direction of polarization will always be at right angles to the plane of incidence and the refractive index will be constant. This constant index is called the ordinary index. The extraordinary index will be given by the other axis of the ellipse and will depend on the propagation direction. When the direction is along the axis of the ellipsoid, the extraordinary index will equal the ordinary index.

From the equation of the ellipsoid, Eq. (11), one

$$\frac{x^2}{\omega^2} + \frac{y^2}{\omega^2} + \frac{z^2}{\epsilon^2} = 1 \tag{11}$$

can derive the expression for the ellipse and in turn Eq. (12) for the extraordinary index n_e as a function

$$n_e = \frac{\omega\epsilon}{(\epsilon^2 \cos^2 r + \omega^2 \sin^2 r)^{1/2}} \tag{12}$$

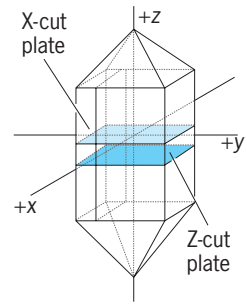

Fig. 4. X-cut and Z-cut plates in $NH_4H_2PO_4$ crystal.

of propagation direction. Since the ellipsoid is symmetrical it is necessary to define this direction only with respect to the ellipsoid axis. Here r is the angle in the material between the normal to the wavefront and the axis of the ellipsoid, n_e is the extraordinary index associated with this direction, ω is the ordinary index, and ϵ is the maximum value of the extraordinary index (usually referred to simply as the extraordinary index).

A slab cut normal to the optic axis is termed a Z-cut or C-cut plate (**Fig. 4**). When such a plate is placed between crossed linear polarizers, such as Nicol prisms, it gives a pattern in monochromatic light shown in **Fig. 5**. The explanation of this pattern is obtained from the equations given for the indices of refraction. To a first approximation the retardation for light of wavelength λ passing through the plate at an angle r can be written as Eq. (13), where d is

$$\Gamma = \frac{(n_e - \omega)d}{\lambda \cos r} \tag{13}$$

the thickness of the plate. The axis of the equivalent retardation plate at an angle r will be in a plane containing the optic axis of the plate and the direction of light propagation. Wherever Γ is a whole number

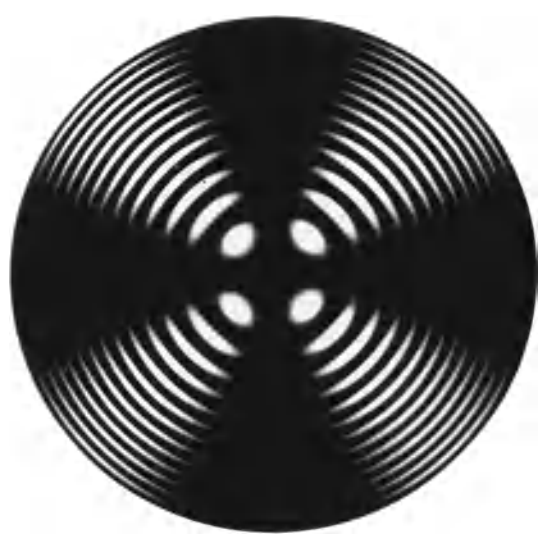

Fig. 5. Interference figure from fluorspar that has been cut perpendicular to the optic axis and placed between crossed Nicol prisms. (*After M. Born and E. Wolf, Principles of Optics, 6th ed., Pergamon Press, 1980*)

of waves, the light leaving the plate will have the same polarization as the incident light. The ordinary index is constant. The extraordinary index is a function of r alone, as seen from Eq. (12). Accordingly, the locus of direction of constant whole wave retardation will be a series of cones. If a uniform white light background is observed through such a plate, a series of rings of constant whole wave retardation will be seen. Since the retardation is a function of wavelength, the rings will appear colored. The innermost ring will have the least amount of color. The outermost ring will begin to overlap and disappear as the blue end of the spectrum for one ring covers the red end of its neighbor. The crystal plate has no effect along the axes of the polarizers since here the light is polarized along the axis of the equivalent retardation plate. The family of circles is thus bisected by a dark cross. When the crystal is mounted between like circular polarizers, the dark cross is not present and the center of the system of rings is clear. During World War II crystal plates were used in this fashion as gunsights. The rings appear at infinity even when the plate is close to the eye. Lateral movement of the plate causes no angular motion of the ring system. The crystal plate could be rigidly fastened to a gun mount and adjusted so as to show at all times the direction in space in which the gun was pointing.

Uniaxial crystal plates in which the axis lies in the plane of the plate are termed X-cut or Y-cut. When such plates are observed between crossed polarizers, a pattern of hyperbolas is observed.

When crystal plates are combined in series, the patterns become much more complex. In addition to the fringes resulting from the individual plates, a system of so-called moiré fringes appears. *See* MOIRÉ PATTERN.

Negative and positive crystals. In calcite the extraordinary wave travels faster than the ordinary wave. Calcite and other materials in which this occurs are termed negative crystals. In positive crystals the extraordinary wave travels slower than the ordinary wave. In the identification of uniaxial minerals one of the steps is the determination of sign. This is most easily demonstrated in a section cut perpendicular to the optic axis. Between crossed linear polarizers the pattern in convergent light, as already mentioned, is a series of concentric circles which are bisected by a dark cross. When a quarter-wave plate is inserted between one polarizer and the crystal, with its axis at 45° to the polarizing axis, the dark rings are displaced outward or inward in alternate quadrants. If the rings are displaced outward along the slow axis of the quarter-wave plate, the crystal is positive. If the rings are displaced inward along the axis, then the crystal is negative.

If a Z-cut plate of a positive uniaxial crystal is put in series with a similar plate of a negative crystal, it is possible to cancel the birefringence so that the combination appears isotropic.

Biaxial crystals. In crystals of low symmetry the index ellipsoid has three unequal axes. These crystals are termed biaxial and have two directions along which the wave velocity is independent of the polarization direction. These correspond to the two sections of the ellipsoid which are circular. These sections are inclined at equal angles to the major axes of the ellipsoid, and their normals lie in a plane containing the major and intermediate axes. In convergent light between polarizers, these crystals show a pattern which is quite different from that which appears with uniaxial crystals. A plate cut normal to the major axis of the index ellipsoid has a pattern of a series of lemniscates and ovals. The directions corresponding to the optic axes appear as black spots between crossed circular polarizers. The interference pattern between crossed linear polarizers is shown in **Fig. 6**.

Fig. 6. Interference figure from Brazil topaz. (*After M. Born and E. Wolf, Principles of Optics, 6th ed., Pergamon Press, 1980*)

Angle between optic axes. One of the quantities used to describe a biaxial crystal or to identify a biaxial mineral is the angle between the optic axes. This can be calculated directly from the equation for the index ellipsoid, Eq. (14), where α, β, and γ

$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} + \frac{z^2}{\gamma^2} = 1 \tag{14}$$

are the three indices of refraction of the material. The circular sections of the ellipsoid must have the intermediate index as a radius. If the relative sizes of the indices are so related that $\alpha > \beta > \gamma$, the circular sections will have β as a radius and the normal to these sections will lie in the xz plane. The section of the ellipsoid cut by this plane will be the ellipse of Eq. (15). The radius of

$$\frac{x^2}{\alpha^2} + \frac{z^2}{\gamma^2} = 1 \tag{15}$$

length β will intersect this ellipse at a point x_1z_1 where Eq. (16) holds. The solution of these two

$$x_1^{\,2} + z_1^{\,2} = \beta^2 \tag{16}$$

equations gives for the points x_1z_1 Eqs. (17). These

$$x_1 = \pm\sqrt{\frac{\alpha^2(\gamma^2-\beta^2)}{\gamma^2-\alpha^2}}$$
$$z_1 = \pm\sqrt{\frac{\gamma^2(\alpha^2-\beta^2)}{\alpha^2-\gamma^2}} \qquad (17)$$

points and the origin define the lines in the xz plane which are also in the planes of the circular sections of the ellipsoid. Perpendiculars to these lines will define the optic axes. The angle Ω between the axes and the z direction will be given by Eq. (18).

$$\tan^2\Omega = \frac{{z_1}^2}{{x_1}^2} = \frac{(1/\alpha^2)-(1/\beta^2)}{(1/\beta^2)-(1/\gamma^2)} \qquad (18)$$

The polarization direction for light passing through a biaxial crystal is computed in the same way as for a uniaxial crystal. The section of the index normal to the propagation direction is ordinarily an ellipse. The directions of the axes of the ellipse represent the vibration direction and the lengths of the half axes represent the indices of refraction. One polarization direction will lie in a plane which bisects the angle made by the plane containing the propagation direction and one optic axis and the plane containing the propagation direction and the other optic axis. The other polarization direction will be at right angles to the first. When the two optic axes coincide, this situation reduces to that which was demonstrated earlier for the case of uniaxial crystals.

Conical refraction. In biaxial crystals there occurs a set of phenomena which have long been a classical example of the theoretical prediction of a physical characteristic before its experimental discovery. These are the phenomena of internal and external conical refraction. They were predicted theoretically in 1832 by William Hamilton and experimentally demonstrated in 1833 by H. Lloyd.

As shown earlier, the Huygens wavelets in a biaxial crystal consist of two surfaces. One of these has its major axis at right angles to the major axis of the other. The two thus intersect. In making the geometrical construction to determine the ray direction, two directions are found where the two wavefronts coincide and the points of tangency on the two ellipsoids are multiple. In fact, the plane which represents the wavefront is found to be tangent to a circle which lies partly on one surface and partly on the other. The directions in which the wavefronts coincide are the optic axes. A ray incident on the surface of a biaxial crystal in such a direction that the wavefront propagates along an axis will split into a family of rays which lie on a cone. If the crystal is a plane parallel slab, these rays will be refracted at the second surface and transmitted as a hollow cylinder parallel to the original ray direction. Similarly, if a ray is incident on the surface of a biaxial crystal plate at such an angle that it passes along the axes of equal ray velocity, it will leave the plate as a family of rays lying on the surface of a cone. The first of these phenomena is internal conical refraction; the second is external conical refraction. Equation (19) gives the

$$\tan\psi = \frac{\beta}{\sqrt{\alpha\gamma}}\sqrt{(\beta-\alpha)(\gamma-\beta)} \qquad (19)$$

half angle ψ of the external cone. Bruce H. Billings

Bibliography. F. D. Bloss, *An Introduction to the Methods of Optical Crystallography*, 1961, reprint 1989; M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; P. P. Ewald, *On the Foundations of Crystal Optics*, 1991; F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976; E. A. Wood, *Crystals and Light*, 2d ed., 1977.

Crystal structure

The arrangement of atoms, ions, or molecules in a crystal. Crystals are defined as solids having, in all three dimensions of space, a regular repeating internal unit of structure. This definition, which dates from early studies of crystals, expresses very concisely what crystals are.

The interior of crystals has been studied by the use of x-rays, which excite signals from the atoms that inhabit the crystal. The signals are of different strengths and depend on the electron density distribution about atomic cores. Thus, there is no probe analogous to a constant light that illuminates a landscape uniformly. Light atoms give weaker signals and hydrogen is invisible to x-rays. However, the mutual atomic arrangements that are called crystal structures can be derived once the chemical formulas and physical densities of solids are known, based on the knowledge that atomic positions are not arbitrary but are dictated by crystal symmetry, and that the diffraction signals received are the result of systematic constructive interference between the scatterers within the regularly repeating internal unit of pattern.

This article discusses fundamental concepts, the dimensionality dependence of crystal structure determination, the structures found for elements and chemically simple compounds, and the conclusions drawn. The relationship between low coordination and structural variability is illustrated, as is a noncrystallographic packing. For related information *See* CRYSTALLOGRAPHY; POLYMORPHISM (CRYSTALLOGRAPHY); X-RAY CRYSTALLOGRAPHY; X-RAY DIFFRACTION.

Fundamental Concepts

Crystals should be imagined as three-dimensional solids. They have fundamental units of structure—atoms, ions, or molecules—mutually arranged to form a geometrically regular unit of pattern, or motif, that continuously repeats itself by translations in three dimensions. Crystals are thus defined in terms of space, population, and mutual arrangement. Crystal space is represented as an indefinitely extended lattice of periodically repeating points. The periodicity of the lattice is defined by the lengths and mutual orientations of three lattice vectors that enclose the pattern. Population is defined as the total number

and kind of fundamental units of structure that form the pattern.

Ordering and states of matter. It is the chemical ordering to make a pattern, and the geometric regularity with which that pattern is repeated in space, that distinguishes among the three classical states of matter—gas, liquid, and solid—and also distinguishes between the crystals and noncrystalline solids. Gases are freely moving atoms or molecules that have no interaction that maintains a contact between them. They are totally disordered. Liquids of various viscosities are essentially incompressible. This means that there are fundamental units of structure that maintain a contact distance between them as they move freely in space. This distance is called the short-range spatial order of the liquid. Short-range order is similar to the order in a bag of marbles. *See* GAS; LIQUID.

Glasses are similar to liquids but have rigid structures because contact between fundamental units is maintained. Their geometric regularity of repetition extends to a few diameters of the fundamental unit. This is an intermediate range of spatial order, larger than that of a liquid. *See* AMORPHOUS SOLID; GLASS.

The order and periodicity of crystals must extend to about 100 nanometers in all three dimensions of space to give the sharply defined diffraction signals required for mapping structural details by x-rays. Intermediate states of order are seen in liquid crystals, which have long molecules as fundamental units of structure. These are arranged with their lengths parallel to each other, but without periodicity, in the nematic state. In the smectic state there is orientation in equally spaced planes but no sideways periodicity, like traffic moving freely on a multilane highway. *See* LIQUID CRYSTALS.

Requirement for three-dimensional order. The necessity of meeting the order and periodicity requirements in three dimensions for exact mapping of crystal structure is well illustrated by elemental carbon. In the graphite structure (**Fig. 1**) the atoms are arranged in six-member, planar, hexagonal rings, which form two-dimensional extensive and periodic layers by sharing their edges. Every carbon atom is trigonally bonded to three others in the same sheet, with separations of 0.142 nm. Normal to this sheet, the layer spacing is 0.342 nm, but this does not define a period. There are no binding forces between carbon atoms in different sheets, except for the weak van der Waals attractive forces; the layers therefore slide over each other like cards in a deck. There is therefore no lattice periodicity in the third dimension. The structure is fundamentally two-dimensional, and was more difficult to define than the three-dimensional, tetrahedrally bonded diamond structure, where the carbon-carbon separation is 0.154 nm. There are no sharp diffraction signals from which the interlayer spacing could be determined precisely because, lattice periodicity being absent in this direction, there is no fixed carbon-carbon separation. *See* GRAPHITE.

Arrangement of atoms and molecules. **Figure 2** illustrates the lattice symmetry and the fundamental unit of pattern and symmetry operations of a crystal structure that generate clones of the pattern in different orientations. The link between a crystal structure and either physical properties or biological function requires examination of the mutual arrangement of all the atoms and molecules, the distances that separate atom centers, and the distribution of empty space.

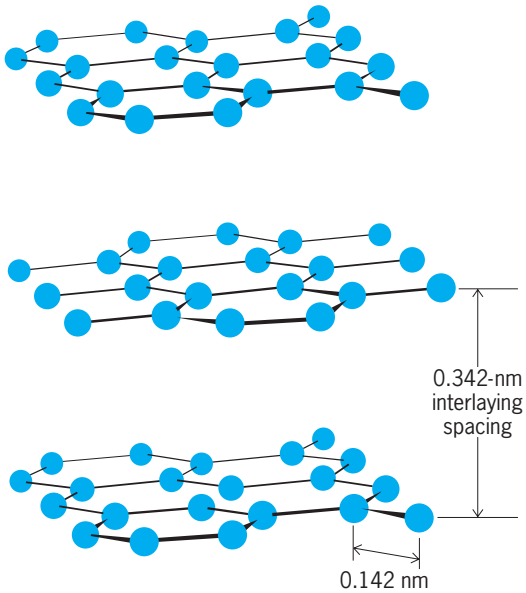

Fig. 1. Crystal structure of graphite.

Figure 2 shows the crystal structure of anthracene ($C_{14}H_{10}$), an organic dye with blue fluorescence. The crystal symmetry is monoclinic with three unequal lattice vectors ($a = 0.856$ nm, $b = 0.604$ nm, $c = 1.116$ nm), and one angle, $\beta \neq 90^\circ$ ($\beta = 124.7^\circ$). There are 14 carbon atoms, arranged as three fused hexagonal rings in each unit of pattern and two units of pattern in each unit cell: eight at the cell vertices, each shared by eight cells ($8/8 = 1$), and two halves within each cell. The long axes of the molecules are parallel, inclined at about 9° from the c axis of the crystal. The 10 hydrogen atoms, invisible to x-rays

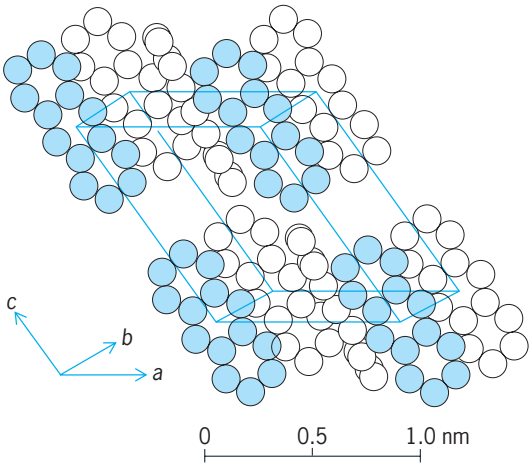

Fig. 2. Structure of anthracene ($C_{14}H_{10}$).

but seen by neutron diffraction since they have a magnetic moment, symmetrically surround each unit of pattern. *See* NEUTRON DIFFRACTION.

The cohesive forces that cause these nonpolar molecules to come together and pack themselves into a crystal are explained by quantum mechanics as instantaneously induced dipole moments of one molecule on another, due to the noncoincidence of the centers of charge of an atom's nucleus and the electron cloud. Their average over time is zero, so that they are not observed as a net dipole moment. They provide the small cohesive energy of the crystal and the rapidly varying, minute locations of lines of communication within the crystal. *See* CHEMICAL BONDING; COHESION (PHYSICS); INTERMOLECULAR FORCES.

Crystal imperfection. The realities of crystal space are, in general, not as they are found in perfect diamonds, nor as they are modeled to be, even in crystals that appear to be freely formed, as in the case of crystallization from supersaturated solutions or from a melt. If scaled up, these crystals would win no prizes in crafting; not in quiltmaking for accurate replication of details within a unit of pattern; nor in wall building for precise translations of identical unit-cell bricks; nor in architecture for the exact symmetries of their extensive designs. The growth process is characterized by constraints and turbulences, and by the dynamic interaction between the substance that is crystallizing and the medium that provides the components of growth and the growth environment. The process is reflected within the structures formed as an assemblage of atoms is collected and made relatively immobile by releasing the energy known as the heat of crystallization. The resulting crystal lattices resemble a mosaic of slightly misaligned adjacent regions. This is actually fortunate for research in x-ray crystallography. Perfect alignment would result in subtraction of energy by interference with the primary beam, due to a 180° phase reversal of the reflected beam (primary extinction). Internally diffracted beams would also be attenuated by internal reflection from regions above them (secondary extinction). *See* CRYSTAL GROWTH.

Detailed structure and advanced probes. Each of the spatially misaligned mosaic blocks of a single crystal is assumed to maintain lattice periodicity within it. This assumption is confirmed by the sharp diffraction patterns observed. There are some "wrong" atoms, vacant lattice sites, trapped gas atoms, and so forth, and the atomic occupants jiggle about while also vibrating cooperatively and synchronously in complex internal modes of motion. Intricate patterns of electron exchange are enacted, and systematic changes in spin orientations can occur for an atom with a magnetic moment. *See* CRYSTAL DEFECTS; EXCHANGE INTERACTION; LATTICE VIBRATIONS.

Details like these are important for understanding the relationships between structure determination on the atomic and molecular levels and the cooperative behavior that determines bulk properties and functions. They are explored by methods other than conventional x-ray crystallography. For example, the power and tunability of synchrotron x-radiation makes it possible to discriminate against a specific atom by not exciting it to generate a signal. In effect, it is possible to sequentially "light up" different atoms and resolve specific details by parts. Intricate structures with very large atomic populations thereby yield maps with details resolved. *See* SYNCHROTRON RADIATION.

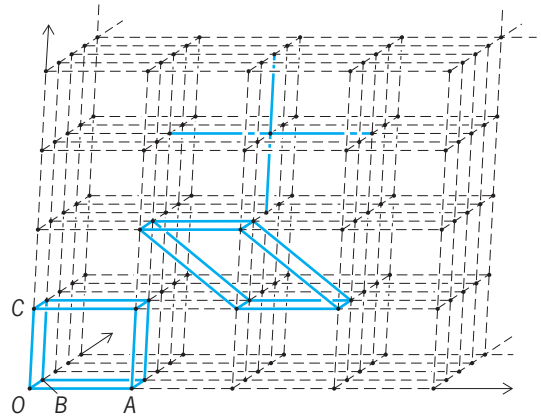

Fig. 3. A space lattice, two possible unit cells, and the environment of a point.

Space Lattices and Symmetry Operations

Figure 3 shows a rectangular space lattice with two possible cells outlined. These have the same cell volumes but different symmetries. Since crystallographic unit cells are completely defined by three lattice vectors, the crystal symmetry referenced to this lattice can be no higher than orthorhombic: $a \neq b \neq c$ ($OA \neq OB \neq OC$), and all angles equal to 90°. This and a possible monoclinic cell, with the same vectors a and b (OA and OB) and one angle not equal to 90°, are outlined. If the OAB plane is rotated and the vector a (OA) is extended to terminate at the next lattice point, then all angles differ from 90° and the crystal symmetry represented becomes triclinic. The mutual arrangement and atom coordinates of the cell population must be such that the environment, seen from every point of the space lattice, remains the same.

Screw axes. These combine the rotation of an ordinary symmetry axis with a translation parallel to it and equal to a fraction of the unit distance in this direction. **Figure 4** illustrates such an operation, which is symbolically denoted 3_1 and 3_2. The translation is respectively 1/3 and 2/3. The helices are added to help the visualization, and it is seen that they are respectively right- and left-handed. The projection on a plane perpendicular to the axis shows that the relationship about the axis remains in spite of the displacement. A similar type of arrangement can be considered around the other symmetry axes, and the following possibilities arise: 2_1, 3_1, 3_2, 4_1, 4_2, 4_3, 6_1, 6_2, 6_3, 6_4, and 6_5.

If screw axes are present in crystals, it is clear that the displacements involved are of the order of a few tenths of nanometer and that they cannot be

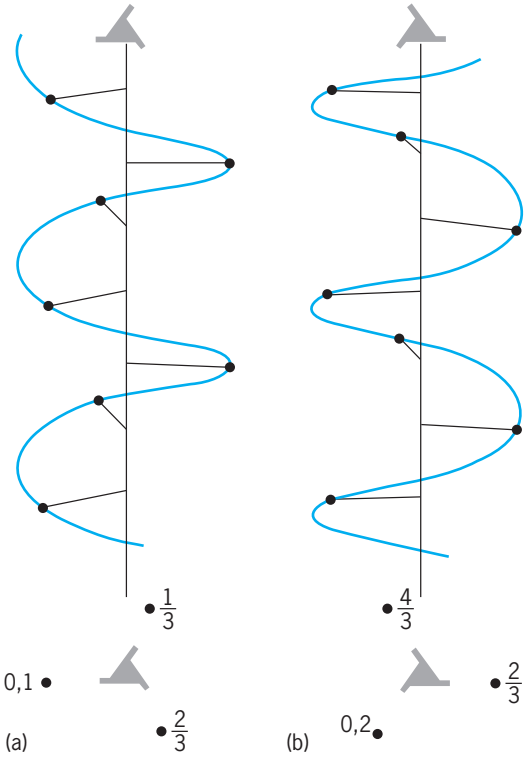

Fig. 4. Screw axes (*a*) 3_1 and (*b*) 3_2. Lower parts of figure are projections on a plane perpendicular to the axis. Numbers indicate heights of points above that plane.

distinguished macroscopically from ordinary symmetry axes. The same is true for glide mirror planes, which combine the mirror image with a translation parallel to the mirror plane over a distance that is half the unit distance in the glide direction.

The handedness of screw axes is a very important feature of many biological and mineral structures. They come in enantiomorphic pairs, with one member of the pair, the left-handed one, preferred, in the biological molecules found on Earth. In concentrated sugar solutions this screw symmetry operation is apparently maintained in some fashion since different directions for rotation of the plane of polarized light are observed. *See* OPTICAL ACTIVITY; STEREOCHEMISTRY.

The helices of deoxyribonucleic acid (DNA) and of quartz are alike in that they contain what is essentially a structural backbone that can both flex and change partners with molecules that are not in the structure while maintaining structural continuity. In quartz, this is reflected by the direction of negative expansion when heated. The helical chain of $(SiO_4)^{4-}$ vertex-linked tetrahedra changes shape as two linked tetrahedra rotate about their common oxygen vertex. The overall configurational change is like that of a squashed mattress coil when it is sat upon. *See* DEOXYRIBONUCLEIC ACID (DNA); SILICA MINERALS.

Space groups. These are indefinitely extended arrays of symmetry elements disposed on a space lattice. A space group acts as a three-dimensional kaleidoscope: An object submitted to its symmetry operations is multiplied and periodically repeated in such a way that it generates a number of interpenetrating identical space lattices. The fact that 230 space groups are possible means, of course, 230 kinds of periodic arrangement of objects in space. When only two dimensions are considered, 17 space groups are possible.

Space groups are denoted by the Hermann-Mauguin notation preceded by a letter indicating the Bravais lattice on which it is based. For example, P $2_12_12_1$ is an orthorhombic space group; the cell is primitive and three mutually perpendicular screw axes are the symmetry elements. J. D. H. Donnay and D. Harker have shown that it is possible to deduce the space group from a detailed study of the external morphology of crystals.

Common Structures

A discussion of some common structural systems found in metals, crystalline compounds, and minerals is given below.

Metals. In general, metallic structures are relatively simple, characterized by a close packing and a high degree of symmetry. Manganese, gallium, mercury, and one form of tungsten are exceptions. Metallic elements situated in the subgroups of the periodic table gradually lose their metallic character and simple structure as the number of the subgroup increases. A characteristic of metallic structures is the frequent occurrence of allotropic forms; that is, the same metal can have two or more different structures which are most frequently stable in a different temperature range.

The forces which link the atoms together in metallic crystals are nondirectional. This means that each atom tends to surround itself by as many others as possible. This results in a close packing, similar to that of spheres of equal radius, and yields three distinct systems: close-packed (face-centered) cubic, hexagonal close-packed, and body-centered cubic.

Close packing. For spheres of equal radius, close packing is interesting to consider in detail with respect to metal structures. Close packing is a way of arranging spheres of equal radius in such a manner that the volume of the interstices between the spheres is minimal. The problem has an infinity of solutions. The manner in which the spheres can be most closely packed in a plane *A* is shown in **Fig. 5**. Each sphere is in contact with six others; the centers form a regular pattern of equilateral triangles. The cavities between the spheres are numbered. A second, similar plane can be positioned in such a way that its spheres rest in the cavities 1, 3, 5 between those of the layer *A*. The new layer, *B*, has an arrangment similar to that of *A* but is shifted with respect to *A*. Two possibilities exist for adding a third layer. Its spheres can be put exactly above those of layer *A* (an assembly *ABA* is then formed), or they may come above the interstices 2, 4, 6. In the latter case the new layer is shifted with respect to *A* and *B* and is called *C*. For each further layer two possibilities exist, and any sequence such as *ABCBABCACBA* . . . in which two successive layers have not the same denomination is a solution of the problem. In the vast majority of

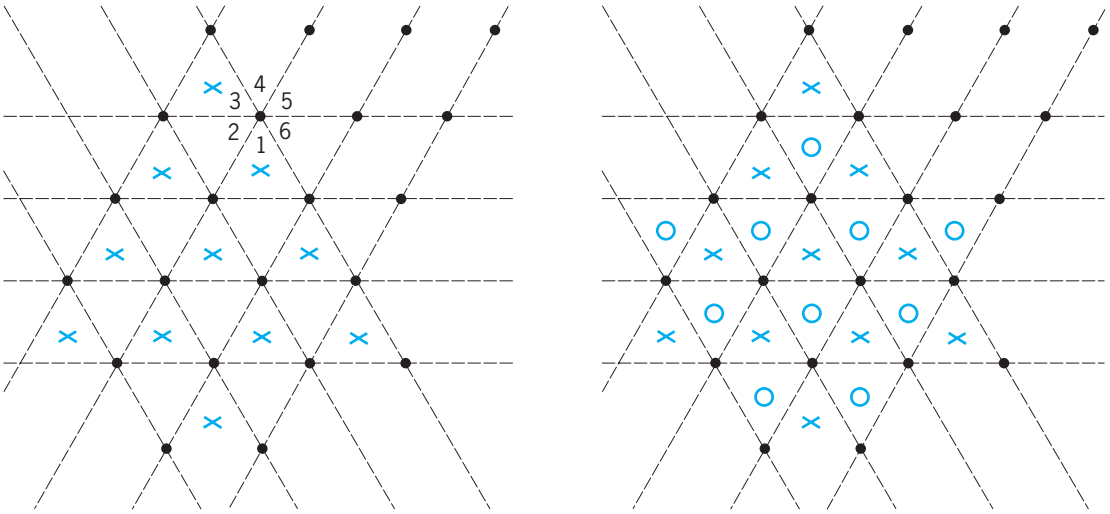

Fig. 5. Close packing of spheres of equal radius in a plane. The centers for the *A*, *B*, and *C* layers are indicated by a dot, a cross, and an open circle, respectively. Cavities between spheres are numbered.

cases, periodic assemblies with a very short repeat distance occur.

Face-centered cubic structure. This utilizes close packing characterized by the regular repetition of the sequence *ABC*. The centers of the spheres form a cubic lattice, as shown in **Fig. 6***a*. This form contains four sets of planes. The densely packed planes of the type *A, B, C* are perpendicular to the threefold axis and can therefore be written as $\{111\}$. Since these are the closest packed planes of the structure, d_{111} is greater than any other d_{hkl} of the lattice. The densest rows in these planes are $\langle 110 \rangle$.

It is relatively easy to calculate the percentage of the volume occupied by the spheres. The unit cell has a volume a^3 and contains four spheres of radius R, their volume being $16\pi R^3/3$. The spheres touch each other along the face diagonal (Fig. 6*a*); $a\sqrt{2}$ is therefore equal to $4R$, or $R = a\sqrt{2}/4$. Substitution gives for the volume of the spheres $\pi\sqrt{2a^3/6}$, which is 73% of the volume of the cube. It is clear that the same percentage of the unit volume is filled in all other close-packed assemblies, that is, assemblies corresponding to other alternations of the planes *A, B, C*.

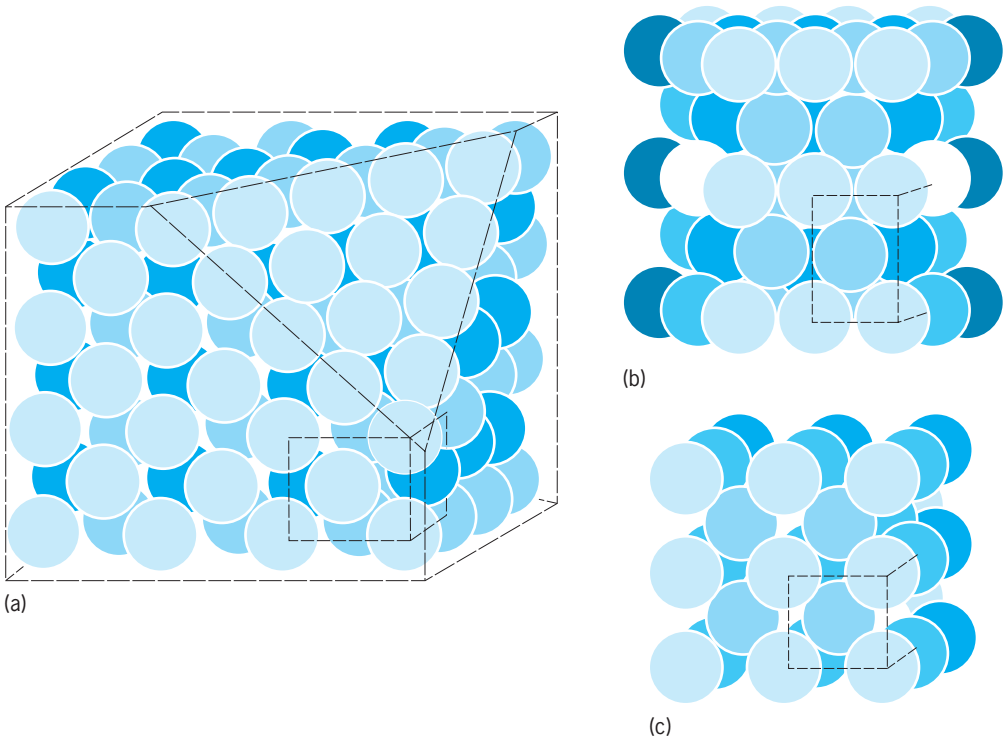

Fig. 6. Close packing of spheres in space. (*a*) Cubic close packing (face-centered cubic). One set of close-packed 111″ planes (*A*, *B*, or *C*) is shown. (*b*) Hexagonal close packing. (*c*) Body-centered cubic arrangement.

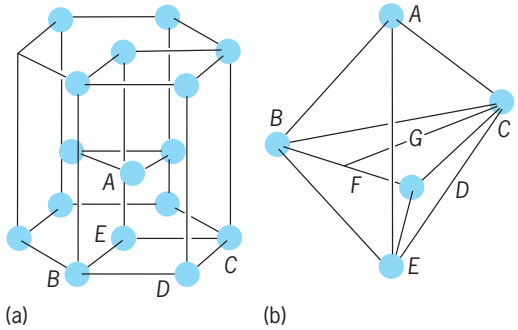

Fig. 7. Hexagonal close-packed structure. (*a*) Three unit cells, showing how the hexagonal axis results. One of the cells is fully outlined. (*b*) Calculation of the ratio *c*/*a*. Distance *AE* is equal to height of cell.

Hexagonal close-packed structure. This is a close packing characterized by the regular alternation of two layers, or *ABAB* . . . . The assembly has hexagonal symmetry (**Fig. 7**). Six spheres are at the corners of an orthogonal parallelepiped having a parallelogram as its base; another atom has as coordinates (1/3, 1/3, 1/2). The ratio c/a is easily calculated. The length $BD = a$, the edge of the unit cell. The height of the cell is $AE = 2AG = c$, defined in Eqs. (1) and (2).

$$AG = \sqrt{a^2 - \left(\frac{2}{3}\frac{a\sqrt{3}}{2}\right)^2} = a\sqrt{2/3} \qquad (1)$$

$$c = 2AG = a\sqrt{8/3} \qquad c/a = \sqrt{8/3} = 1.633 \qquad (2)$$

This latter value is important, for it permits determination of how closely an actual hexagonal structure approaches ideal close packing.

Body-centered cubic structure. This is an assembly of spheres in which each one is in contact with eight others (Fig. 6*c*).

The spheres of radius R touch each other along the diagonal of the cube, so that measuring from the centers of the two corner cubes, the length of the cube diagonal is $4R$. The length of the diagonal is also equal to $a\sqrt{3}$, if a is the length of the cube edge, and thus $R = a\sqrt{3/4}$. The unit cell contains two spheres (one in the center and 1/8 in each corner) so that the total volume of the spheres in each cube is $8\pi R^3/3$. Substituting $R = a\sqrt{3/4}$ and dividing by a^3, the total volume of the cube, gives the percentage of filled space as 67%. Thus the structure is less dense than the two preceding cases.

The closest-packed planes are {110}; this form contains six planes. They are, however, not as dense as the *A, B, C* planes considered in the preceding structures. The densest rows have the four ⟨111⟩ directions.

Tabulation of structures. The structures of various metals are listed in **Table 1**, in which the abbreviations fcc, hcp, and bcc, respectively, stand for face-centered cubic, hexagonal close-packed, and body-centered cubic. The structures listed hcp are only roughly so; only magnesium has a c/a ratio (equal to 1.62) that is very nearly equal to the ratio (1.63) calculated above for the ideal hcp structure. For cadmium and zinc the ratios are respectively 1.89 and 1.86, which are significantly larger than 1.63. Strictly speaking, each zinc or cadmium atom has therefore not twelve nearest neighbors but only six. This departure from the ideal case for subgroup metals follows a general empirical rule, formulated by W. Hume-Rothery and known as the 8-*N* rule. It states that a subgroup metal has a structure in which each atom has 8-*N* nearest neighbors, *N* being the number of the subgroup.

Table 2 compares metallic (elemental) and ionic radii of the same elements. The first and second elements of their respective periods are listed. Both radii are determined from families of compounds. Metallic radii are always larger—for the light metals, very much larger—than the ionic radii. Thus, although the metallic structures are close-packed arrangements, the atoms obviously do not touch each other as they do in the sphere-packing models.

Crystalline compounds. Simple crystal structures are usually named after the compounds in which they were first discovered (diamond or zinc sulfide, cesium chloride, sodium chloride, and calcium fluoride). Many compounds of the types A^+X^- and $A^{2+}X_2^-$ have such structures. They are highly

TABLE 1. Metal structures

Metal	Modification	Stability range	Structure
Cadmium	α	To melting point	hcp
Chromium	α	To melting point	bcc
	β	Electrolytic form	Hexagonal
Cobalt	α	To 788°F (420°C)	hcp mixed with fcc
	β	788°F (420°C) to melting point	fcc
Gold	α	To melting point	fcc
Iron	α	To 1668°F (909°C)	bcc
	γ	1668–2557°F (909–1403°C)	fcc
	δ	2557°F (1403°C) to melting point	bcc
Magnesium	α	To melting point	Nearly hcp
Nickel	α	To melting point	fcc
	β	Electrolytic form	Hexagonal
Silver	α	To melting point	fcc
Zinc	α	To melting point	hcp
Zirconium	α	To 1584°F (862°C)	hcp
	β	1584°F (862°C) to melting point	bcc

TABLE 2. Metallic and ionic radii in crystals

Period	Element	Atomic number Z	Metallic structure: Structure type	Metallic structure: Radius, nm	Ionic structure: Radius, nm	Ionic structure: Charge
First short	Lithium (Li)	3	bcc	0.152	0.07	1^+
	Beryllium (Be)	4	hcp	0.111	0.035	2^+
Second short	Sodium (Na)	11	bcc	0.186	0.10	1^+
	Magnesium (Mg)	12	hcp	0.160	0.07	2^+
First long	Potassium (K)	19	bcc	0.231	0.133	1^+
	Calcium (Ca)	20	fcc	0.197	0.100	2^+
Second long	Rubidium (Rb)	37	bcc	0.240	0.150	1^+
	Strontium (Sr)	38	fcc	0.210	0.110	2^+
Third long	Cesium (Cs)	55	bcc	0.266	0.170	1^+
	Barium (Ba)	56	bcc	0.217	0.135	2^+

symmetrical, the unit cell is cubic, and the atoms or ions are disposed at the corners of the unit cell and at points having coordinates that are combinations of 0, 1, 1/2, or 1/4.

Sodium chloride structure. This is an arrangement in which each positive ion is surrounded by six negative ions, and vice versa. The arrangement is expressed by stating that the coordination is 6/6. The centers of the positive and the negative ions each form a face-centered cubic lattice. They are shifted one with respect to the other over a distance $a/2$, where a is the repeat distance (**Fig. 8***a*). Systematic study of the dimensions of the unit cells of compounds having this structure has revealed that:

1. Each ion can be assigned a definite radius. A positive ion is smaller than the corresponding atom and a negative ion is larger. Figure 8*b* shows the effective sizes of ions in sodium chloride.

2. Each ion tends to surround itself by as many others as possible of the opposite sign because the binding forces are nondirectional.

On this basis the structure is determined by two factors, a geometrical factor involving the size of the two ions which behave in first approximation as hard spheres, and an energetical one involving electrical neutrality in the smallest possible volume. In the ideal case all ions will touch each other; therefore, if r_A and r_X are the radii of the ions, $4r_X = a\sqrt{2}$ and $2(r_A + r_X) = a$. Expressing a as a function of r_X gives $r_A/r_X = \sqrt{2} - 1 = 0.41$. When r_A/r_X becomes smaller than 0.41, the positive and negative ions are no longer in contact and the structure becomes unstable. When r_A/r_X is greater than 0.41, the positive ions are no longer in contact, but ions of different sign still touch each other. The structure is stable up to $r_A/r_X = 0.73$, which occurs in the cesium chloride structure. *See* IONIC CRYSTALS.

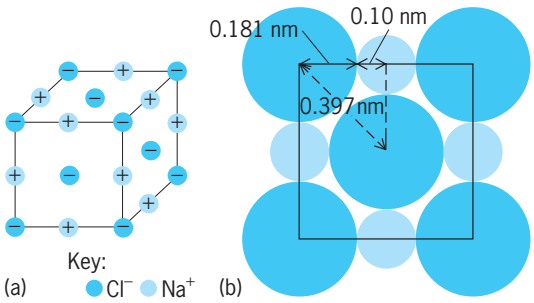

Fig. 8. Structure of sodium chloride. (*a*) Cubic unit cell. (*b*) Basal plane of the unit cell. Effective sizes of ions are shown.

Cesium chloride structure. This is characterized by a coordination 8/8 (**Fig. 9***a*). Each of the centers of the positive and negative ions forms a primitive cubic lattice; the centers are mutually shifted over a distance $a\sqrt{3/2}$. The stability condition for this structure can be calculated as in the preceding case. Contact of the ions of opposite sign here is along the cube diagonal (Fig. 9*b*).

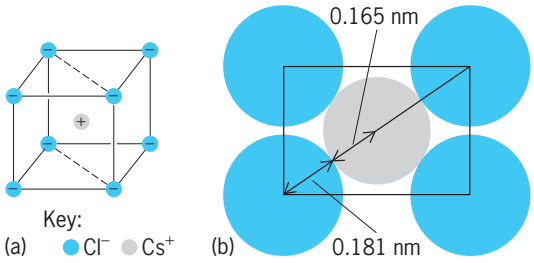

Fig. 9. Structure of cesium chloride. (*a*) Cubic unit cell. (*b*) Section of the cell through a diagonal plane, showing the close packing of Cs^+ and Cl^- ions, and effective sizes of ions.

Diamond structure. In this arrangement each atom is in the center of a tetrahedron formed by its nearest neighbors. The 4-coordination follows from the well-known bonds of the carbon atoms. This structure is illustrated in **Fig. 10***a*. The atoms are at the corners of the unit cell, the centers of the faces, and at four points having as coordinates (1/4, 1/4, 1/4), (3/4, 3/4, 1/4), (3/4, 1/4, 3/4), (1/4, 3/4, 3/4). The atoms can be divided into two groups, each forming a face-centered cubic lattice; the mutual shift is $a\sqrt{3/4}$.

Zinc blende structure. This structure, shown in Fig. 10*b*, has coordination 4/4 and is basically similar to the diamond structure. Each zinc atom (small circles in Fig. 10*b*) is in the center of a tetrahedron formed by sulfur atoms (large circles), and vice versa. The zinc atoms form a face-centered cubic lattice, as do the sulfur atoms.

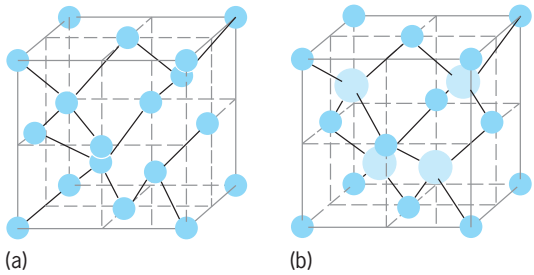

Fig. 10. Tetrahedral crystal compound structures. (*a*) Diamond. (*b*) Zinc blende.

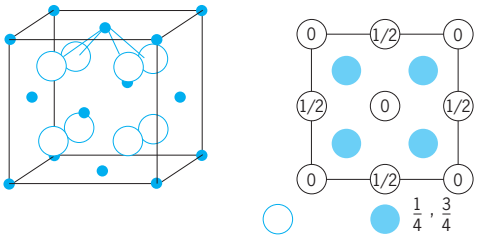

Fig. 11. Calcium fluoride structure.

Calcium fluoride structure. **Figure 11** shows the calcium fluoride structure. If the unit cell is divided into eight equal cubelets, calcium ions are situated at corners and centers of the faces of the cell. The fluorine ions are at the centers of the eight cubelets. There exist three interpenetrating face-centered cubic lattices, one formed by the calcium ion and two by the fluorine ions, the mutual shifts being (0, 0, 0), (1/4, 1/4, 1/4), (3/4, 3/4, 3/4).

Minerals. Silica (SiO_2) is the most abundant material in the Earth's crust. It is surpassed only by carbon in variety of architectural structures. As in carbon, these structures also have widely ranging, dynamic properties. **Figure 12** indicates schematically how this abundance of structures can evolve from the 4:2 coordination of silica where each silicon atom has four tetrahedrally disposed oxygen atoms about it. Figure 12*a* represents an isolated tetrahedron with unsatisfied charge about each oxygen atom. Joining tetrahedra to make a single chain (Fig. 12*b*) in effect subtracts oxygen atoms and lowers the charge. Joining two such chains (Fig. 12*c*) extends the structure and also reduces the charge since another oxygen is subtracted. It is readily seen how the single chain of Fig. 12*b* could be wrapped as a helix, thereby extending the charged regions into the third dimension. *See* SILICATE MINERALS.

Figure 13 shows a polyhedral structural model of the zeolite faujasite. The composition is $NaCa_{0.5}(Al_2Si_5O_{14}) \cdot 10H_2O$. Every vertex of the truncated octahedra with hexagonal faces is the location of a tetrahedrally coordinated silicon or aluminum atom. The structure continues to fill space indefinitely. However, while model space is filled according to lattice requirements, physical space is quite empty. This is a charged framework structure in which mobile species, such as water (H_2O) molecules and cations, can freely flow through the large pipes that are an integral part of the structure. Cation exchange occurs with this flow. *See* MOLECULAR SIEVE; ZEOLITE.

Structures and Periodicity

The requirement that structures have three-dimensional periodicity that can be referenced to a Bravais lattice of mathematical points with three lattice vectors, that every point of the Bravais lattice have the same environment about it, and therefore that identical unit cells repeat periodically by translations in all three directions, is a requirement for accurate crystal structure determination by x-rays. As was demonstrated by the example of graphite, it is not a requirement imposed upon physical reality. Compositionally modulated structures exist; indeed, they are to be expected. As growth occurs, local composition changes, the heat of crystallization is released, and growing surfaces become available as templates for fresh deposition. Nor does the fact that crystals, as defined, give sharp x-ray diffraction patterns imply that all diffraction patterns come from crystals, as defined.

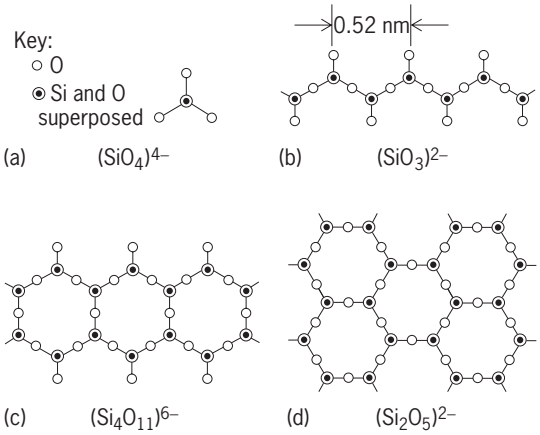

Fig. 12. Structures of silicates. (*a*) Fundamental SiO_4 tetrahedron. (*b*) Linear chain of tetrahedra. (*c*) Band formed by two linked linear chains. (*d*) Plane of linked tetrahedra.

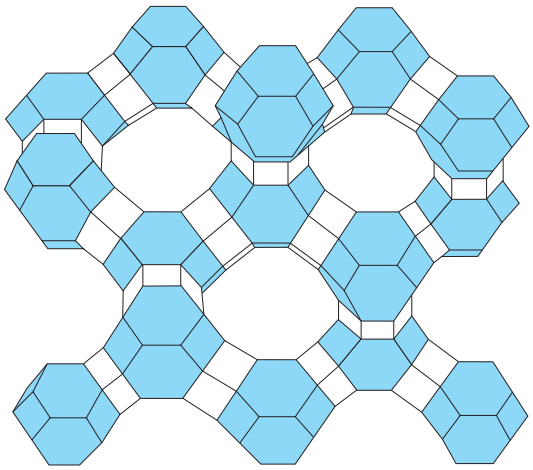

Fig. 13. Crystal structure of faujasite. The silicon and aluminum atoms are situated at the apices of the truncated octahedra, which are joined to form a three-dimensional framework.

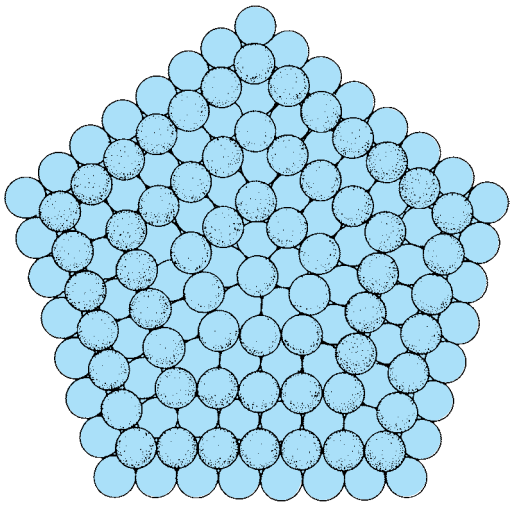

Fig. 14. Sphere packing with fivefold symmetry. (*After A. F. Wells, Models in Structural Inorganic Chemistry, Oxford University Press, 1970*)

Finally, the random packing of spheres in contact, resembling a bag of marbles, is very dense, while close-packed arrangements need not be physically dense. **Figure 14** shows how a dense-sphere packing, with one-dimensional periodicity only, along a fivefold symmetry axis, will fill space. A two-layer construction is shown in which the first layer forms concentric pentagons with odd numbers of spheres on each side, and a second layer is added in which the corresponding numbers of spheres are even. The compact two-layer constructions extend indefinitely in two dimensions and can be stacked to form an infinite packing along the fivefold axis. The density is 0.724, which is denser than the body-centered cubic close packing of spheres. Doris Evans

Bibliography. C. S. Barrett and T. B. Massalski, *Structure of Metals*, 3d ed., 1980; J. D. Brock, Liquids, crystals and liquid crystals, *Phys. Today*, 42(7):52–59, July 1989; A. Guinier, *The Structure of Matter*, 1984; B. Hyde and S. Andersson, *Inorganic Crystal Structures*, 1989; C. Janot, *Quasicrystals: A Primer*, 2d ed., 1995; A. F. Wells, *Structural Inorganic Chemistry*, 5th ed., 1984.

Crystal whiskers

Single crystals that have grown in filamentary form. Such filamentary growths have been known for centuries and can form by a variety of mechanisms. Some grow from their base: either these are extruded to relieve pressure in the base material or they grow as a result of a chemical reaction at the base. In both cases, the growth occurs at a singularity in the base material. Other crystal whiskers grow at their tip, into a supersaturated medium, and the filamentary growth results from a strong anisotropy in the growth rate. *See* SINGLE CRYSTAL; SUPERSATURATION.

The growth of extruded metal whiskers was studied in the 1950s, when it was discovered that tin whiskers growing from solder were causing electrical short circuits. These whiskers were typically a few micrometers in diameter and a few millimeters in length. This shorting problem was solved by adding more lead to the solder. Great interest in the whiskers developed after it was discovered that the strength exhibited by some whiskers approached that expected theoretically for perfect crystals. This great strength results from the internal and surface perfection of the whiskers, whereas the strength of most materials is limited by defects. The interest in the high strength of the whiskers centered on the possibility of using them in composites to increase the strength of more ductile matrix materials. Fibers of silica, boron, and carbon, which are much easier to fabricate in large quantity than whiskers, also exhibit similarly high strengths, and are now used in composites. *See* COMPOSITE MATERIAL; CRYSTAL DEFECTS.

The controlled growth of whiskers of a wide variety of materials is possible by means of the vapor-liquid-solid mechanism (see **illus.**). Here, a liquid droplet containing impurities rides on the tip of the whisker, providing a preferential site for condensation. Many electronic materials are prone to whisker formation by the vapor-liquid-solid process during chemical vapor deposition, so that control of the deposition to eliminate whisker formation is essential. Whiskers of many substances have also been grown in supersaturated solutions, but the mechanism of growth is not understood in detail. *See* CRYSTAL GROWTH. Kenneth A. Jackson; Richard S. Wagner

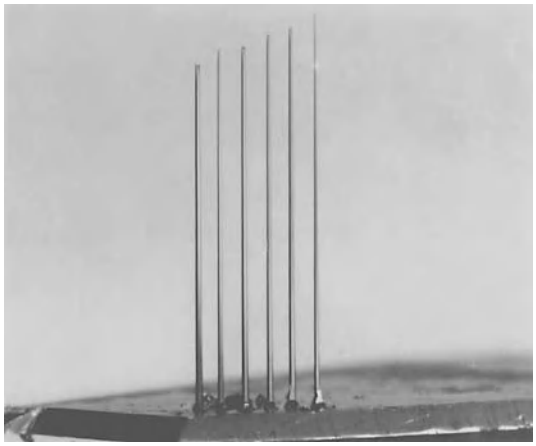

Crystal whiskers grown by the vapor-liquid-solid process. (*AT&T Bell Laboratories*)

Bibliography. S. S. Brenner, Growth and properties of "whiskers," *Science*, 128:569–575, 1958; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; A. P. Levitt, *Whisker Technology*, 1970.

Crystallization

The formation of a solid from a solution, melt, vapor, or a different solid phase. Crystallization from solution is an important industrial operation because

of the large number of materials marketed as crystalline particles. Fractional crystallization is one of the most widely used methods of separating and purifying chemicals. This article discusses crystallization of substances from solutions and melts. *See* GLASS; PHASE EQUILIBRIUM; POLYMER; SINGLE CRYSTAL; SOLUTION; SUBLIMATION.

Solutions. In order for crystals to nucleate and grow, the solution must be supersaturated; that is, the solute must be present in solution at a concentration above its solubility. Different methods may be used for creating a supersaturated solution from one which is initially undersaturated. The possible methods depend on how the solubility varies with temperature. Two examples of solubility behavior are shown in the **illustration**. Either evaporation of water or cooling may be used to crystallize potassium nitrate (KNO_3), while only evaporation would be effective for NaCl. An alternative is to add a solvent such as ethanol which greatly lowers the solubility of the salt, or to add a reactant which produces an insoluble product. This causes a rapid crystallization perhaps more properly known as precipitation. *See* PRECIPITATION (CHEMISTRY).

Nucleation. The formation of new crystals is called nucleation. At the extremely high supersaturations produced by addition of a reactant or a lower-solubility solvent, this nucleation may take place in the bulk of the solution in the absence of any solid surface. This is known as homogeneous nucleation. At more moderate supersaturations new crystals form on solid particles or surfaces already present in the solution (dust, motes, nucleation catalysts, and so on). This is called heterogeneous nucleation.

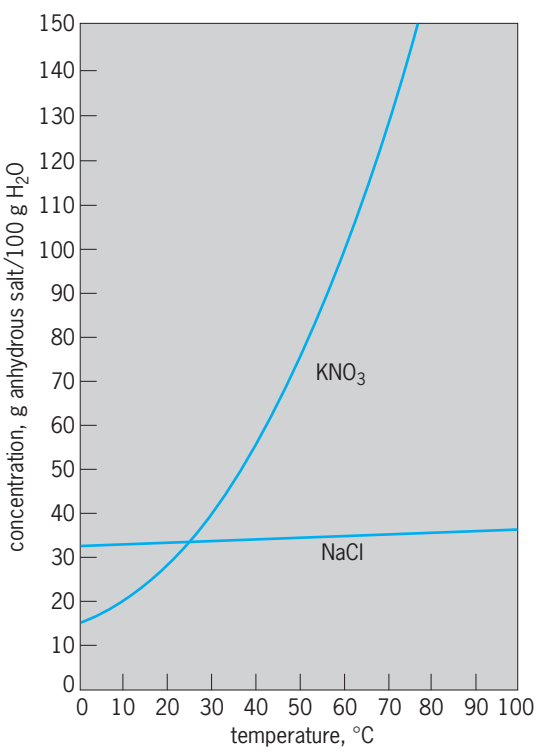

Temperature-solubility curves for two salts. °F = (°C × 1.8) + 32.

When solutions are well agitated, nucleation is primarily secondary, that is, from crystals already present. Probably this is usually due to minute pieces breaking off the crystals by impact with other crystals, with the impeller, or with the walls of the vessel. *See* NUCLEATION.

Crystal size. Generally, large crystals are considered more desirable than small crystals, probably in the belief that they are more pure. However, crystals sometimes trap (occlude) more solvent as inclusions when they grow larger. Thus there may be an optimal size. The size distribution of crystals is influenced primarily by the supersaturation, the amount of agitation, and the growth time. Generally, the nucleation rate increases faster with increasing supersaturation than does the growth rate. Thus, lower supersaturations, gentle stirring, and long times usually favor large crystals. Low supersaturations require slow evaporation and cooling rates.

Crystal habit. Often the habit (shape) of crystals is also an important commercial characteristic. In growth from solutions, crystals usually display facets along well-defined crystallographic planes, determined by growth kinetics rather than by equilibrium considerations. The slowest-growing facets are the ones that survive and are seen. While habit depends somewhat on the supersaturation during growth, very dramatic changes in habit are usually brought about by additives to the solution. Strong habit modifiers are usually incorporated into the crystal, sometimes preferentially. Thus the additive may finally be an impurity in the crystals.

Fractional crystallization. In fractional crystallization it is desired to separate several solutes present in the same solution. This is generally done by picking crystallization temperatures and solvents such that only one solute is supersaturated and crystallizes out. By changing conditions, other solutes may be crystallized subsequently. Occasionally the solution may be supersaturated with respect to more than one solute, and yet only one may crystallize out because the others do not nucleate. Preferential nucleation inhibitors for the other solutes or seeding with crystals of the desired solute may be helpful. Since solid solubilities are frequently very small, it is often possible to achieve almost complete separation in one step. But for optimal separation it is necessary to remove the impure mother liquor from the crystals. Rinsing of solution trapped between crystals is more effective for large-faceted crystals. Internally occluded solution cannot be removed by rinsing. Even high temperatures may not burst these inclusions from the crystal. Repeated crystallizations are necessary to achieve desired purities when many inclusions are present or when the solid solubility of other solutes is significant.

Melts. If a solid is melted without adding a solvent, it is called a melt, even though it may be a mixture of many substances. That is the only real distinction between a melt and a solution. Crystallization of a melt is often called solidification, particularly if the process is controlled by heat transfer so as to produce a relatively sharp boundary between the solid

and the melt. It is then possible to slowly solidify the melt and bring about a separation, as indicated by the phase diagram. The melt may be stirred to enhance the separation. The resulting solid is usually cut into sections, and the purest portion is subjected to additional fractional solidifications. Alternatively the melt may be poured off after part of it has solidified. Zone melting was invented to permit multiple fractional solidifications without the necessity for handling between each step. Fractionation by the above techniques appears to be limited to purification of small batches of materials already fairly pure, say above 95%. *See* ZONE REFINING.

Fractional crystallization from the melt is also being used for large-scale commercial separation and purification of organic chemicals. It has also been tested for desalination of water. Rather than imposing a sharp temperature gradient, as in the above processes, the melt is relatively isothermal. Small, discrete crystals are formed and forced to move countercurrent to the melt. At the end from which the crystals exit, all or part of the crystals are melted. This melt then flows countercurrent to the crystals, thereby rinsing them of the adhering mother liquor. *See* CRYSTAL GROWTH; CRYSTAL STRUCTURE; CRYSTALLOGRAPHY. William R. Wilcox

Bibliography. S. J. Jancic and P. A. Grootscholten, *Industrial Crystallization*, 1984; J. W. Mullin (ed.), *Industrial Crystallization*, 3d ed., 1993; J. Nývlt, *Industrial Crystallization: The Present State of the Art*, 2d ed., 1983; J. Nývlt and O. Sohnel, *The Kinetics of Industrial Crystallization*, 1985; A. D. Randolph and M. A. Larsen, *Theory of Particulate Processes: Analysis and Techniques of Continuous Crystallization*, 2d ed., 1988.

Crystallography

The branch of science that deals with the geometric forms of crystals. The field of crystallography originated in the prehistoric observation of geometric forms of minerals, such as a collection of quartz (SiO_2) crystals (**Fig. 1**). These crystals grow as hexagonal pyramids, with growth being terminated by the coming together of external faces to form a pinnacle. Equally beautiful forms can be made by very simple procedures. How to describe, classify, and measure such forms are the first questions of crystallography. Revealing the forces that made them and the activities within them are the modern directions of the field. This article deals with the rigor, scope, and impact of crystallography, its interdisciplinary nature, and the key results and successes that have made it historically continuous and vital. For other aspects of crystallography *see* CRYSTAL STRUCTURE; POLYMORPHISM (CRYSTALLOGRAPHY).

Crystallography is essential to progress in the applied sciences and technology and developments in all materials areas, including metals and alloys, ceramics, glasses, and polymers, as well as drug design. It is equally vital to progress in fundamental physics and chemistry, mineralogy and geology, and computer science, and to understanding of the dynamics and processes of living systems. This vitality and power are reflected in the rapid exploitation and commercial development of its methods and the support given to data files that record and make available its results.

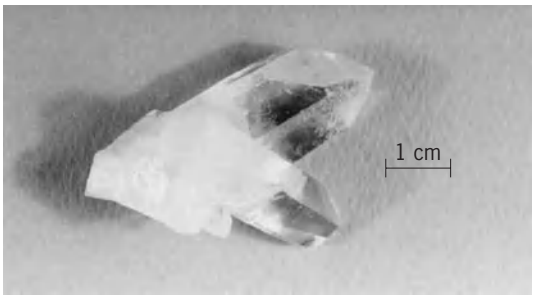

Fig. 1. Natural, intergrown quartz crystals.

Classical Fundamentals

Understanding of crystals with well-developed faces can be obtained with visible light, a simple mechanical device, mathematics, and the imagination as the only tools.

Faces and zones. Slow changes in the composition of the in-feeding substances and in the environment of growth (temperature, pressure, and atmospheric composition) yield many minerals that have large crystals that are optically transparent and chemically well defined, with well-defined shapes. The external morphology reflects growth rates in different directions. These directions remain constant during the course of the growth process, and are represented mathematically as the normals to sets of parallel planes that are imagined as being added on as growth proceeds. The faces that meet and define an edge belong to a zone, a zone being a set of planes that share one common direction, the zone axis. The invariance of interfacial angles, measured by rotation about an axis that is defined by the zone direction, was discovered in the seventeenth century. *See* CRYSTAL GROWTH.

Goniometer. Figures 2 and **3** show the optical goniometer and the method of tagging and

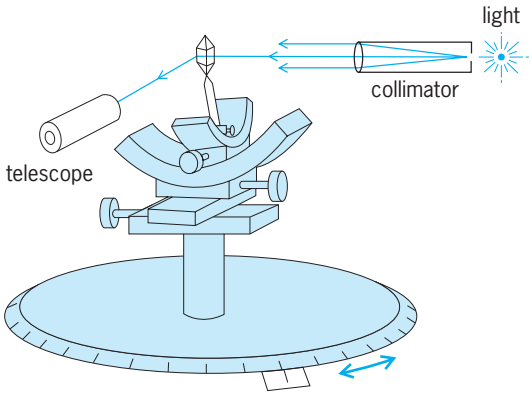

Fig. 2. Reflecting goniometer.

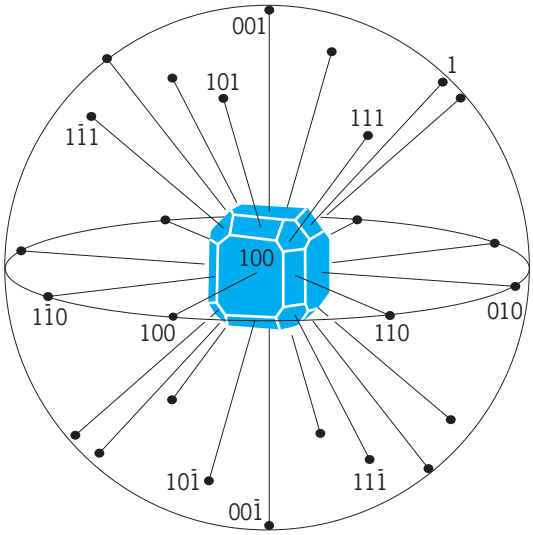

Fig. 3. Spherical projection of normals to crystal faces.

measuring morphological geometries. Translation in two orthogonal directions aligns a face of the crystal so that it reflects light into a telescope when a zone axis is vertical, and the position is recorded. The crystal is then reoriented by rotation about the goniometer axis and two calibrated circular arcs, horizontal or equatorial and vertical or azimuthal, to reflect light from another face into the telescope.

Spherical projection. Interfacial angles are calculated from spherical geometry. Figure 3 illustrates the procedure for a crystal having well-developed faces of which three are mutually perpendicular. The normals to these faces are the natural directions for constructing an orthogonal frame of reference for measurement. The crystal is imagined to be shrunk and placed at the center of a sphere with coordinates (0,0,0). The face normals, labeled [100], [010], and [001], define the directions of an orthogonal reference system. Normals to the same set of planes, but oppositely directed, are labeled [$\bar{1}$00], [0$\bar{1}$0], [00$\bar{1}$]. The reversal of sign indicates that the crystal must be rotated 180° to obtain the same view. Rotation

Bravais lattice cells	Axes and interaxial angles	Examples
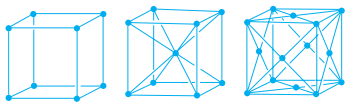 Cubic P Cubic I Cubic F	Three axes at right angles; all equal: $a = b = c$; $\alpha = \beta = \gamma = 90°$	Copper (Cu), silver (Ag), sodium chloride (NaCl)
Tetragonal P Tetragonal I	Three axes at right angles; two equal: $a = b \neq c$; $\alpha = \beta = \gamma = 90°$	White tin (Sn), rutile (TiO_2), β-spodumene ($LiAlSi_2O_6$)
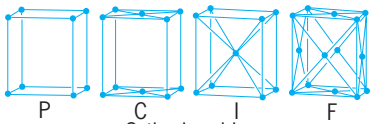 P C I F Orthorhombic	Three axes at right angles; all unequal: $a \neq b \neq c$; $\alpha = \beta = \gamma = 90°$	Gallium (Ga), perovskite ($CaTiO_3$)
Monoclinic P Monoclinic C	Three axes, one pair not at right angles, of any lengths: $a \neq b \neq c$; $\alpha = \gamma = 90° \neq \beta$	Gypsum ($CaSO_4 \cdot 2H_2O$)
Triclinic P	Three axes not at right angles, of any lengths: $a \neq b \neq c$; $\alpha \neq \beta \neq \gamma \neq 90°$	Potassium chromate (K_2CrO_7)
Trigonal R (rhombohedral)	Rhombohedral: three axes equally inclined, not at right angles; all equal: $a = b = c$; $\alpha = \beta = \gamma \neq 90°$	Calcite ($CaCO_3$), arsenic (As), bismuth (Bi)
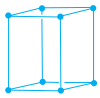 Trigonal and hexagonal C (or P)	Hexagonal: three equal axes coplanar at 120°, fourth axis at right angles to these: $a_1 = a_2 = a_3 \neq c$; $\alpha = \beta = 90°$, $\gamma = 120°$	Zinc (Zn), cadmium (Cd), quartz (SiO_2) [P]

Fig. 4. The 14 Bravais lattices, derived by centering of the seven crystal classes (P and R) defined by symmetry operators.

about the [001] direction interchanges the positions of [100] and [$\bar{1}$00] faces and their bounding edges. Rotation about [010] turns these faces upside down. Correct designations for group movements and symmetry operations are clearly essential for establishing and maintaining orientation in crystal space. The directions of face normals determine points at which the imagined sphere is pierced. The solid angles between an array of such points, all lying on the same great circle of the sphere, belong to a zone.

Bravais lattices. Optical measurements and stereographic projections established the constancy of interfacial angles, independent of how well developed the faces are. Such properties as the cleavage of large rhombohedral crystals of calcite ($CaCO_3$) into little rhombs suggested that the large crystal could be represented by geometrically identical smaller units stacked together, by translation, to fill space. The 14 lattices of Bravais (**Fig. 4**) enlarged the seven crystal systems of optical mineralogy by adding centering points to them: body (I), face (F), and base (C) centers. The 14 lattices define three-dimensional distributions of mathematical points such that the environments of all points of the lattice are identical. They also define the symmetries of frameworks for constructing mathematical models to represent the observed and measured realities—models made from cells of the smallest volume, but also highest symmetry, that stack together by translation to fill space.

Stacking of model cells does not imply that a crystal grows by stacking identical bricks; a lattice of identically surrounded mathematical points does not imply that any real objects, atoms or molecules, are located at the points; and filling space by translation of identical cells does not imply that the space defined by the cells is filled. Rather, the Bravais lattices are a formalism for representing observed geometries and symmetries of real crystals by three-dimensional lattices of identically surrounded points. **Figure 5** shows a two-dimensional square array of points with imaginary cells outlined. The four different shapes with one lattice point per cell have the same area and one elementary motif per cell (P in Fig. 4). The square with two units of motif, one within it and the four that are common to each of four cells ($4/4 = 1$), has twice the area of the other shapes; it is centered. The lengths of three edges, and the angles between them, represent crystal symmetry with reference to three-dimensional lattices of such points.

Miller indices. The lattices also provide the means to identify imaginary planes within the cell; these are called Miller indices (h,k,l). They consist of small whole numbers. For example, each of the six faces of a simple cube, with the origin of a coordinate frame of reference at the cube body center, is normal to one of the reference axes and parallel to the plane defined by the two others. The six faces are indexed as their normals in Fig. 3—(100), ($\bar{1}$00), (010), (0$\bar{1}$0), (001), and (00$\bar{1}$)—to represent a face that intercept the $x,\bar{x}$ axis but not the y and z; the y, $\bar{y}$ axis but not the x and z; and so forth. Hypothetical parallel planes with $^1/_2$ the interplanar spacing are represented as (200), ($\bar{2}$00), (020), and so forth.

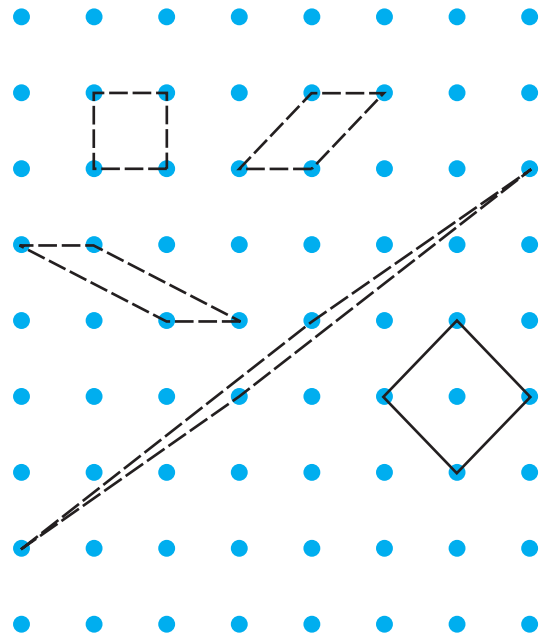

Fig. 5. Two-dimensional lattice of points. All cells fill space by translation; that is, the surroundings of all lattice points are the same. The four cells indicated by broken lines are P, with one unit of pattern and the same areas. The cell indicated by solid lines is C, with two units of pattern and twice the area of the others.

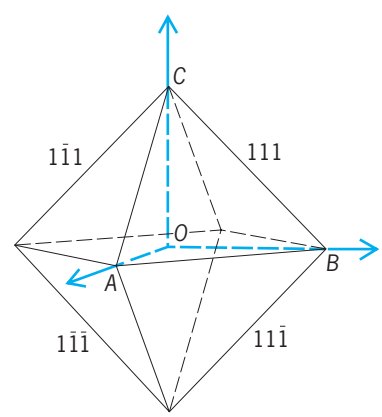

Fig. 6. Indexing of the four front faces of an octahedron.

Figure 6 shows the indexing of faces for an octahedon—the solid obtained by truncating the eight vertices of a cube until all the original faces disappear and a new regular solid with six vertices and eight faces, the dual of the cube with eight vertices and six faces, emerges. The truncation simulates the relative change in growth rates that exposes interior planes as bounding faces, thereby creating the architecture of crystals. Nothing has changed except the faces exposed. The octahedron has full cubic symmetry. The eight bounding faces have intercepts (111), ($\bar{1}$11), (1$\bar{1}$1), (11$\bar{1}$), ($\bar{1}\bar{1}\bar{1}$), (1$\bar{1}\bar{1}$), ($\bar{1}$1$\bar{1}$), and ($\bar{1}$1$\bar{1}$).

Symmetry operations and space groups. A complete mathematical formalism for modeling an external morphological form, and the symmetry relations between imagined units of structure within it, was in place in the nineteenth century. The symmetry operators include rotation axes, glide and mirror

planes, and left- and right-handed screw axes which will simultaneously rotate and translate a three-dimensional object to create its clone in a different spatial position and orientation. The operators minimize the detail required to specify the spatial arrangements of patterns and objects that fill two-dimensional and three-dimensional space.

Symmetry is easily detected in common objects like a car wheel, where one nut and either a five- or sixfold rotation axis will generate other nuts holding the wheel in place. Wallpaper and quilts are useful for determining the minimal unit of pattern that must be specified within a cell, the combination of symmetry operations required to generate the complete pattern within a cell, and different shapes to outline two-dimensional cells that can be moved to fill the plane by translation. Quilts have well-defined symmetry operations and cells and identical shapes of nonidentical clones, if made from different fabrics. These are analogous to the so-called color space groups of crystallography, space groups that greatly increase the number of distinguishably different symmetries beyond the classical 230 by adding a fourth coordinate to the three space coordinates. This is done to encode a real difference that will be manifested in some property. The different directions of the magnetic moments of chemically identical atoms of an element such as iron provide an example of the need for representing a difference on the atomic level between cells that are otherwise identical.

Properties and models. Many physical and chemical properties were known, and models that accounted for them were proposed, before there was a means of getting inside crystals and mapping the components and mutual arrangements of components. Physical properties include the opacity and ductility of metals; the flakiness of mica; the hardness of diamond; the lubricity of graphite; double refraction from calcite, which gives two images of an object; pyroelectricity, where a thermal gradient builds up opposite charges on different faces; the piezoelectric effect in quartz; and the inherent anisotropy of crystals and the handedness of quartz, a handedness seen as either a counterclockwise or clockwise rotation of the plane of polarization of light. Polymorphs were known, as were chemical isomers, chemical compositions with the same molecular formula but different properties, such as ammonium cyanate (NH_4CNO) and urea [$CO(NH_2)_2$]. *See* BIREFRINGENCE; CRYSTAL OPTICS; DIAMOND; GRAPHITE; METAL, MECHANICAL PROPERTIES OF; PIEZOELECTRICITY; POLARIZED LIGHT; PYROELECTRICITY; QUARTZ.

Chemically identical but conformationally different (chair and boat) geometries had been proposed for cyclohexane. Steric hindrance was a concept, and structure models had been proposed and built to explain observed properties. Several were later confirmed by x-ray crystallography. Models based on the packing of ion spheres for sodium chloride (NaCl) and potassium chloride (KCl) were particularly insightful. Compositions, densities, and refractive indices in different directions had been measured, molecular formulas conjectured, and magnetic effects explored. A mathematically rigorous framework for mapping the internal mutual arrangements existed before there was a probe to enter into materials with light (radiation) that would send comprehensible signals revealing the species that built the structures, how these species were arranged, and what the arrangements implied regarding the forces between atoms. *See* CONFORMATIONAL ANALYSIS; STEREOCHEMISTRY.

Modern Crystallography

A suitable light or radiation to serve as a probe came with W. C. Roentgen's discovery of x-rays in 1895. The nature of x-radiation, whether waves or particles, was not immediately known, but the wavelength expected if x-rays were waves could be calculated to be of the order of 0.1 nanometer, comparable to the dimensions of atoms and 10^3–10^4 shorter than visible light. *See* ELECTROMAGNETIC RADIATION; X-RAYS.

Laue theory. The great insight of M. von Laue followed, leading to quantum physics and the reconciliation of the dual wave–particle nature of electromagnetic radiation, and also providing the means to get within the structures that chemists and mineralogists were modeling. Laue asked what would happen if x-rays really were waves; a crystal was really a regular, periodic, three-dimensional array of atoms; and the two entities were to meet. He concluded that the atoms would be excited to vibrate cooperatively, generating interference phenomena analogous to that of an optical diffraction grating. The same phenomenon can be observed by looking at a distant light through the mesh of a fine fabric or sieve. Spots are seen, which get larger as the mesh spacing gets smaller, regularly arranged about the direction of the incident beam.

Equations (1)–(3) must be satisfied to give

$$a(\alpha - \alpha_0) = h \quad (1)$$

$$b(\beta - \beta_0) = k \quad (2)$$

$$c(\gamma - \gamma_0) = l \quad (3)$$

constructive interference from a regular three-dimensional array of scattering lattice points; here a, b, c are the repeat distances of the space lattice; α and α_0, β and β_0, and γ and γ_0 are the direction cosines of the diffracted and incident beams; and h, k, and l are integers defining the order of a particular diffracted beam. Experiment showed all speculations to be correct: scattered x-rays combine with each other as waves do, crystals are three-dimensional arrays of regularly spaced scatterers, and the atoms and molecules of chemical models are real scatterers with 0.1-nm separation between their centers. Broadened diffraction spots reveal considerable distortion from lattice ideality, while very sharp spots come from crystals that are perfect within the resolution of the beam. *See* X-RAY CRYSTALLOGRAPHY; X-RAY DIFFRACTION.

Bragg spectrometer and modeling. The Bragg x-ray spectrometer uses the goniometer of Fig. 2 to study the x-ray spectrum, with the light replaced by x-rays, the collimator by a slit system, and the telescope by an ionization chamber. With mica and other crystals, the beam from a platinum target was shown to consist of continuous radiation, analogous to white light, and superimposed monochromatic radiation of various wavelengths that results from specific transitions between electron shells. The simple three-dimensional cubic chessboard model for structures of alkali halides (sodium chloride and potassium chloride), proposed many years before by optical crystallographers, was confirmed by W. L. Bragg by modeling all the spots found in Laue patterns, recorded on film by a stationary crystal, as well as the intensity differences seen when potassium chloride is substituted for sodium chloride. The following assumptions were used: (1) the mathematical planes of the lattice can act as x-ray mirrors; (2) the x-ray beam is heterogeneous; and (3) the scattering power of an atom is dependent upon its atomic weight.

With a two-dimensional chessboard lattice arrangement, Avogadro's number, and the experimental densities, the distance between the sodium and chlorine atoms in a base-centered square cell containing two atoms (sodium and chlorine) per cell can be calculated. This distance is the sum of the radii of hypothetical atoms of sodium and chlorine. The required probe thus exists for entering structures and obtaining interpretable signals from atoms. The intensities of these signals can be understood in terms of interference from atoms whose distribution in space is modeled by idealized lattice and symmetry theory. The apparent sizes of scatterers—atoms, ions, or molecules—can be measured by assuming spheres in contact. The physical density and molecular formula determine the cell contents. Systematic work recording patterns for the same element in different chemical combinations and decoding the patterns by matching them with hypothetical models of scatterers, whose mutual arrangements are assigned from existing lattice and symmetry theory, gives values for the radii of spherical atoms.

Figure 7 shows the Laue pattern of calcium fluoride (CaF_2), recorded on flat photographic film. The spots are indexed by the projection method of Fig. 3. The distribution of spots reveals the geometric regularity and lattice symmetry of arrangement. The pattern has a threefold rotation axis (spots repeat every $360/3 = 120^\circ$) and a vertical mirror plane, normal to the film, which in effect reflects the spots. Spots on the left are seen on the right as if reflected by a mirror.

Bragg equation. The Bragg equation simplifies the Laue conditions to be met for constructive interference by using what had already been surmised for minerals: that crystals grow as if parallel planes were being added on.Equation (4) makes use of this to

$$\lambda = 2d_1 \sin\theta \tag{4}$$

decode observed patterns as if they originated from

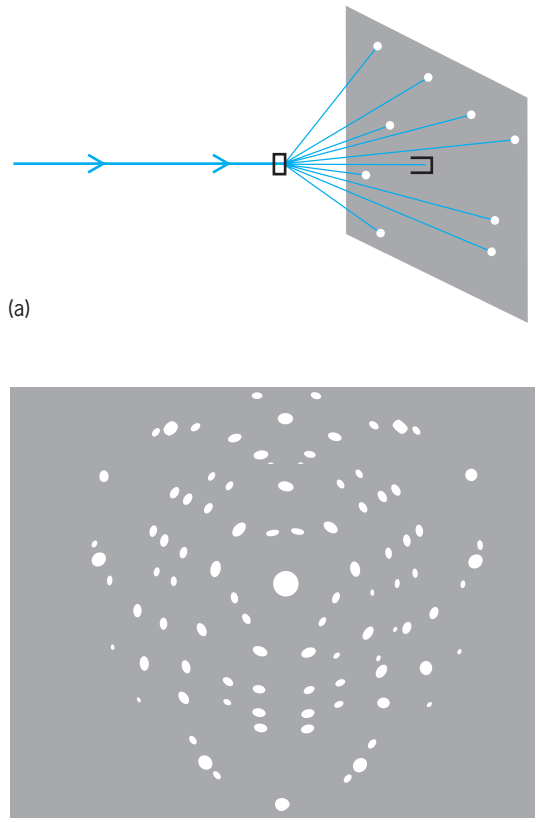

Fig. 7. Laue pattern of calcium fluoride (CaF_2), which has cubic symmetry. The x-ray beam is along a body diagonal (111). (*a*) Experimental setup for generating the pattern. (*b*) Pattern on photographic film. (*Photographic Library, Palais de la Découverte*)

scattering by imaginary planes reflecting the x-ray light as if the planes were mirrors within a unit cell. Here, λ is the x-ray wavelength, $d' = d/n$, where d is the true spacing of lattice planes and n is the number of imaginary planes between them, and θ is the reflection angle.

For orthogonal crystals, Eq. (5) relates the Miller

$$\frac{1}{d^2} = \frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2} \tag{5}$$

indices of a plane, h, k, and l, the spacing d, and the lattice unit cell dimensions, a, b, and c. For the cubic system this simplifies to Eq. (6). More complex equa-

$$\frac{1}{d^2} = \frac{h^2 + k^2 + l^2}{a^2} \tag{6}$$

tions, involving angles, apply to lower symmetry.

X-ray spectra and synchrotron radiation. X-ray spectra have two origins. The beam ordinarily used in x-ray crystallography consists of sharp line spectra emitted by a target material when target atoms are ionized by removing electrons from innermost orbits. A limited heterogeneous spectrum of continuous (white) radiation is emitted during the sudden reduction of energy of the bombarding electrons. This is the radiation used in x-ray spectroscopy. The same phenomenon, reduction of electron beam

energy, but with enormous energies and highly parallel beams, is the source of synchrotron radiation. Synchrotron beams add an extremely powerful probe to x-ray crystallography. They are tunable x-ray sources; therefore, some atoms in a structure can often, in effect, be switched off by using wavelengths that will not excite them. In addition, the high energies give diffraction patterns in very short times. Therefore, the problem of degradation of biologically active materials during long x-ray exposures is eliminated. *See* BREMSSTRAHLUNG; SYNCHROTRON RADIATION; X-RAY SPECTROMETRY.

Technological and Scientific Applications

The sharp x-ray line spectra, characteristic of the bombarded element, are the primary probes for determining interior structural detail.

Techniques. Following the introduction of beams of this radiation, experimental and analytical techniques developed rapidly. Cameras with cylindrical film and enclosed powdered samples record all diffraction lines as arcs of concentric circles. This fundamental powder method has endured since 1917 and is now employed with improved beam purity and optics, improved diffractometers which couple sample and detector rotation, electronic detection, rapid sequential recording, and computer indexing programs that provide patterns of compounds, mixtures of phases, and dynamic changes that occur when crystals are subjected and respond to external stress. The method is applied to single crystals, polycrystalline aggregates, and multiple-phase mixtures, randomly disposed either in space or in geometrically designed composite materials. *See* X-RAY POWDER METHODS.

Data collection. The x-ray patterns of virtually every material that can be handled have been recorded and indexed. Comprehensive collections of experimental and model calculation results for metals, minerals, and inorganic compounds are kept current and accessible through the Joint Committee for Powder Diffraction Study (JCPDS). More sharply focused data files for structures of organic and biological interest are also extensive and report molecular configurations.

Basic structures and properties. Counting the number of neighbors about an atom and measuring the distances between their centers from the diffraction patterns of elements and compounds yield the basic elements of structure: atoms, molecules, ions, and the modern concept of the chemical bond. These procedures also yield explanations for some of the properties of materials.

Many metals fill space with a close-packed periodic arrangement of spheres in contact, but the space is really quite empty. The positively charged atom cores repel each other electrostatically and maintain large equilibrium separations through a sea of mobile electrons, which is the source of metallic conductivity. The mobility of the electron sea, the emptiness of the structures, and the mutual repulsions of the atom cores explain metallic ductility. The effective radii of atoms of the elementary metals are seen to be much greater than those found when the metals are in combination with other elements. Effective radii are also different in compounds having different numbers of neighbors around the same element, smaller for fewer near neighbors, larger for more. L. Pauling's theory of the chemical bond is based upon such crystallographic data. *See* CHEMICAL BONDING; ELECTRICAL CONDUCTIVITY OF METALS; FREE-ELECTRON THEORY OF METALS.

Relation of structure and composition. Decoded patterns of materials of very different compositions and even from widely different origins show surprising similarities when viewed from within the framework of their structural architectures. For example, crystallographic patterns reveal much similarity of structural architecture between fibrous protein materials, such as wool, hair, and the myosin of muscle, which have very different compositions. *See* FIBROUS PROTEIN.

An example of materials of different origins is provided by silica (SiO_2) and ice (H_2O), which are extraordinarily alike in their polymorphic forms, the geometry of their building units, and the connectivity schemes that link the units to form the solids. There are four tetrahedrally arranged oxygen atoms about a silicon atom in silica and four hydrogen atoms about an oxygen atom in all the forms of ice. The variety of possible forms is seen as the many spatial arrangements that can be made by connecting each tetrahedron at its corners to other tetrahedra. *See* SILICA MINERALS; WATER.

In contrast to these architectural similarities between chemically different substances, other patterns show different structural architectures for identical compositions. Furthermore, fundamental differences in apparent size (as mentioned above) and mutual arrangements are observed among atoms of the same element. These differences depend upon the composition that the atoms are part of and also on the conditions of formation. The division between atomic, ionic, and covalently bonded arrangements is established by systematic measurements on the periodic lattice that models the observed diffraction patterns.

Studies of response to change. Modeling crystallographic results with space-filling unit cells of the proper density, independent of the mutual arrangement of atoms or molecules within the cell, provides the means for following, in the laboratory, the degrees of freedom of the overall structure, which are the types of adjustment that the structure makes to accommodate change without disruption. These are determined in several ways. Static determinations, which involve varying the composition systematically and comparing the patterns and densities of a family of compositions, reveal solid solutions and phase relations important to materials development. Dynamic determinations involve changing the temperature or pressure between a sequence of pattern recordings. Thermal expansion and compressibility are important both in developing new materials and properties and in understanding geological processes.

Dynamic studies. The foundation of space-filling lattice theory allows internal changes to be studied under laboratory conditions. Even when the details of mutual rearrangements are not known exactly, the size and shape of unit cells with the proper physical density can often be obtained from indexed patterns, density measurements, and a correct chemical formula. Subjecting a polycrystalline powder to external sources of stress is a dynamic way of focusing on structural details, such as the directions of strongest bonding, that are strongly implied but not directly measured. Heating, cooling, and the application of pressure cause changes in the size and shape of the unit cell as responses to the stress applied.

Changes in relative diffracted intensities indicate that the phase relationships that build up the total intensity are changing. This is often a signal that new modes of interatomic motion have been generated. Dynamic applied crystallography reveals many fundamental phenomena. A solution of elements that were initially uniformly distributed can coalesce as one-dimensional strings or two-dimensional platelets in specific crystallographic directions within the host lattice. A lively pattern of cooperative movements at the atomic level of structure can be observed crystallographically.

Crystals always sense and respond to environmentally imposed stress. They strive to equilibrate their structures to the new environment. Permanent and irreversible processes, as well as reversible ones, occur frequently. Detecting these is critical to the fundamental understanding of the solid state as well as the development of compositions and processes that produce materials that survive under conditions and environments of use.

Structure of glass. The surprising observation that there are true glasses that fill space uniformly with nonperiodic arrangements stimulated a reexamination of scattering theory and improvements in the detection of weak signals. The atomic structure of glass is explained by radial density distributions of scatterers in physical space such that all identical atoms in the structure have the same number and relative arrangement of their nearest neighbors. A geometrically irregular network that distributes spherical concentric shells of appropriate scatterers at larger distances accounts for the background pattern of diffuse scattering. A physical model of identical tetrahedral units, completely connected at their vertices and randomly arranged in space, fits the experimental scattering data. *See* AMORPHOUS SOLID; GLASS.

Living systems and other challenges. The dynamics of living systems, the difficulties in distinguishing light elements, and the inherent ambiguities of measuring, decoding, and mapping are continuing challenges. Major achievements of crystallography include the determination of the structures of deoxyribonucleic acid (DNA), proteins, other biological compounds, and boranes; the development of direct methods of phase determination; and the determination of the structure and mechanism of a photosynthetic center. *See* BORANE; DEOXYRIBONUCLEIC ACID (DNA); PHOTOSYNTHESIS; PROTEIN.

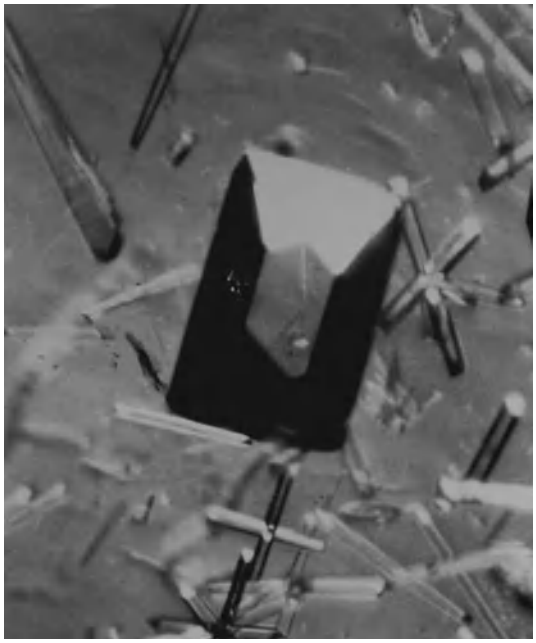

Fig. 8. Large porcine elastomer crystal grown in Spacelab (1988). (*Charles Bugg, University of Alabama*)

The problems encountered in the study of biological systems include the large atomic populations of these systems and the small scale of resolution needed to observe how components arrange themselves and how their configurations readjust, in response to subtle signals, with great speed and over extensive distances. **Figure 8** illustrates an ongoing challenge in the study of living systems: producing good crystals from material that does not ordinarily crystallize, and could not perform its biological function if it did crystallize. Although the crystal of porcine elastomer is considered large, it is 0.5 mm in the largest dimension, about 1/100 the size of the crystals in Fig. 1. It grew aboard Spacelab in 1988. Since convection forces on Earth disrupt the growth process, a gravity-free environment is required to grow a crystal of this protein, see inside it, and map its internal architecture. Only then can the rapid communication and readjustments that take place in living systems be modeled and understood. *See* SPACE PROCESSING.

Advanced probes and detailed modeling. Many kinds of radiation are used to enter within and excite signals from a crystal structure. All are employed to determine detail and fix positions for molecules and atoms within a space that is modeled as the real interior. Success in modeling depends upon extracting regularities and symmetries from signals that originate in the inner space of a material. These regularities are used to reconstruct both the summation of many amplitudes that yields the observed intensities, and the kind and mutual arrangement of the scatterers that locate these intensities in their observed positions. Relating these arrangements to internal activities is the goal of all crystallography. *See* ELECTRON DIFFRACTION; NEUTRON DIFFRACTION.

Icosahedral crystals. Conditions of formation in some three-component metal melts produce very small and morphologically well-defined icosahedral crystals. The idealized interior space of these crystals may not be correctly mapped in terms of diffraction patterns referenced to the three vectors of a Bravais lattice, but require so-called higher-dimensional spaces. *See* QUASICRYSTAL. Doris Evans

Bibliography. C. Giacovazzo (ed.), *Foundations of Crystallography*, 1992; J. P. Glusker, *Crystallography*, 1993; A. Guinier and R. Jullien, *The Solid State: From Superconductors to Superalloys*, 1989; T. Hahn (ed.), *International Tables for Crystallography*, vol. A: *Space-Group Symmetry*, 4th revised ed., 1995; C. Hammond, *Introduction to Crystallography*, 1990; A. Holden and P. Morrison, *Crystals and Crystal Growing*, 1982; J. Karle, Macromolecular structure from anomalous dispersion, *Phys. Today*, 42(6):22–29, June 1989.

Ctenophora

A phylum of exclusively marine organisms, formerly included in the jellyfish and polyps as coelenterates. These animals, the so-called comb jellies, possess a biradial symmetry of organization. They are characterized by having eight rows of comblike plates (*ctenos* = comb) as the main locomotory structures. Most of these animals are pelagic, but a few genera are creeping. Many are transparent and colorless; others are faintly to brightly colored. Many of these organisms are hermaphroditic. Development is biradially symmetrical, with a cydippid larval stage. Five orders constitute this phylum as follows:

Phylum Ctenophora
Order: Cydippida
Lobata
Cestida
Platyctenida
Beroida

Morphology. The body is gelatinous and extremely fragile; its form may be globular, pyriform, or bell- or helmet-shaped. Some species resemble a ribbon. Their size ranges from $^1/_8$ to 20 in. (3 mm to 50 cm). All are biradially symmetrical. The axis of symmetry is determined by the mouth and the organ of equilibrium, or statocyst (**Fig. 1**). The mouth leads into the flattened, elongated pharynx. The sagittal plane is thus referred to as the stomodeal plane. The other plane of symmetry is perpendicular to the sagittal plane and is marked (Beroida excepted) by tentacles. This is known as the tentacular, transverse, or equatorial plane. Eight meridional comb-plate rows or ribs stretch from the aboral pole on the surface of the body.

Comb plates. These are the most characteristic structure of the ctenophores, possessed by all members of the phylum. Each plate consists of a great number of very long related cilia. Successive plates are arranged as meridional ribs or costae which extend from the aboral pole toward the mouth. Each rib is connected to the statocyst by a thin ciliary furrow. Of the eight ribs, four are stomodeal and are located on both sides of the stomodeal plane. The remaining four are subtentacular ribs and are situated on both sides of the tentacular plane. Comb rows are secondarily modified in some specialized ctenophores, and ribs may be entirely absent in the adult stages of a few genera of Platyctenida.

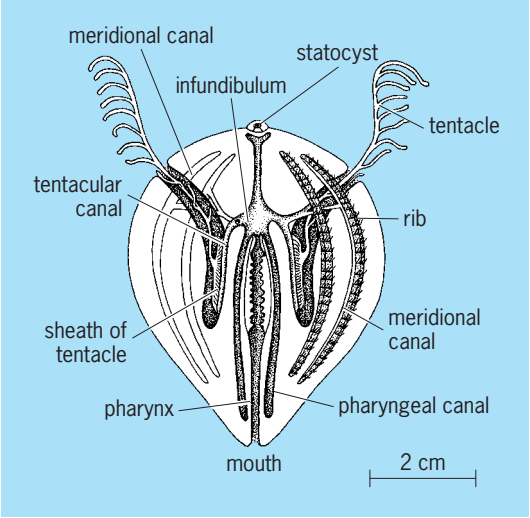

Fig. 1. Structure of a cydippid ctenophore.

Tentacles. With the exception of Beroida, all ctenophores possess two highly extensile tentacles bearing many lateral branches (tentilla). The base of each tentacle is anchored on the bottom of a sheath which is formed by a deep depression in the body surface, and into which the tentacular apparatus may be completely retracted. Tentacles and tentilla are thickly covered with adhesive cells (colloblasts) which are exclusive to the phylum. Each colloblast is connected by the nervous system. A colloblast consists of a cell which looks like a mushroom (**Fig. 2**), the foot of which is firmly anchored in the fibrillar jelly surrounding the neuromuscular axis of the tentillum. This foot contains a helicoidal twisted thread which reaches the expanded head of the colloblast. Here, the helical thread divides into a great number of radii. Each radius terminates by a granule of glue. When a prey collides with a tentillum, some colloblasts burst and their granules stick the prey to the helical threads. These threads apparently act as shock absorbers of the captured prey's movements rather than as springs projecting the head. Tentacles are organs in continuous growth. Their formative tissue is located at their base, in the bottom of the tentacular sheath. In the adult stages of the Lobata and Cestida, the main tentacles, as described above, are reduced, and a series of tentilla with colloblasts develop on the margin of the mouth.

Gastrovascular system. The pharynx is connected aborally with a small chamber, the stomach, or infundibulum, from which a canal system common to all ctenophores ramifies throughout the mesoglea. The most important of the canals are the meridional

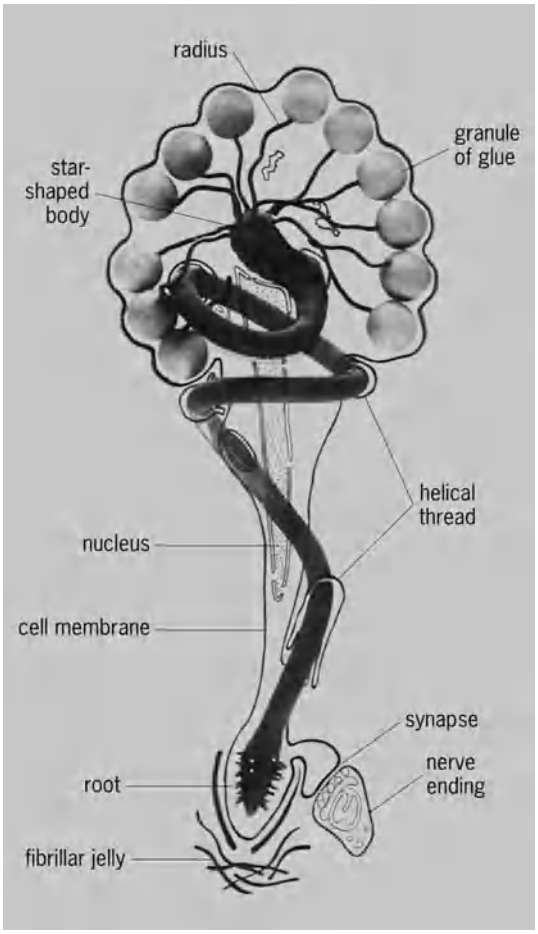

Fig. 2. Colloblast from the tentacle of a ctenophore.

canals which underlie the ribs, the pharyngeal (stomodeal) canals which run along the pharynx, and the tentacular canals (when tentacles exist) which go to the base of the tentacles. While the pharynx has an ectodermal lining, the stomach and the canal system are lined with endodermal epithelium which includes digestive cells. The peripheral canals bear special structures, the cell rosettes, consisting of the superposition of two circles of eight ciliated cells. The peculiarity of these structures lies in a contractile diaphragm which allows complete opening or closing between the mesoglea and the lumen of the gastrovascular canals. The cell rosettes may be considered peculiar excretory or hydrostatic organs, since reversible liquid flows were observed through them with changes in seawater salinity.

Mesoglea. The maintenance of shape in the ctenophores depends on the mesoglea. This inner jelly, very rich in water (95%) is mainly made of protein-carbohydrate complexes. Collagen was biochemically detected, but has not been related to definite mesogleal structures. The mesoglea is crossed by numerous smooth muscle cells which are among the longest (2 in. or 5 cm) of the animal kingdom. Nerve processes connecting the muscle cells have been observed in the mesoglea, where scattered mesenchymal cells are also found.

Nervous system and sensory organs. Ctenophores possess a nervous system organized in a subectodermic synaptic network, which is more condensed along the ribs, under the statocyst, and in the tentacles. There is some evidence for the existence of a cholinergic system together with a monoaminergic one. Although intramesogleal neurons exist, innervation of the endoderm has not been observed. Sensory cells sensitive to water movement are located over the entire surface of the body and are connected to the nervous network. They are most abundant along the tentacles, the lips, and at the aboral pole. The nerve net propagates a localized excitation over the entire body of the animal and complex behavior patterns may take place (feeding behavior, for example). The statocyst is situated in a hollow of the aboral pole. The main feature of this equilibrium organ (**Fig. 3**) is the statolith, a mass of cells containing calcareous granules, held by and linked to four S-shaped tufts of cilia, the balancers. Each balancer, interradially located, is connected with the subpharyngeal and subtentacular costae of the same quadrant by the two corresponding ciliary furrows. The floor of the hollow contains secretory cells involved in the formation of the statolith and sensory cells, presumably photoreceptors and baroreceptors. A fence of incurved cilia forms a transparent dome over this aboral sensory area. Paired structures called the polar fields lie close to the statocyst, in the pharyngeal plane, and consist of linear or crescentic ciliated areas, the function of which is unknown.

Gonads. Many ctenophores are hermaphroditic. Gonads of endodermal origin develop in the outer side of the wall of the meridional canals or their branches. Both the ovaries and spermaries generally develop as continuous tracts of germ cells. The ovaries occupy the perradial and the spermaries the interradial side of the wall. Contrary to the spermatozoa of Porifera and Cnidaria, the ctenophore spermatozoon possesses a well-differentiated acrosome. Gonducts seem to be of general occurrence. Functional gonads in the larval stage occur throughout the phylum.

Embryology. Fertilization usually occurs in seawater. Pelagic ctenophores are self-fertile, but cross-fertilization might also take place inside a swarm of ctenophores. Ctenophore eggs provide the unique example of a biradial pattern of cleavage (**Fig. 4**). The blastomeres are arranged in respect to the future

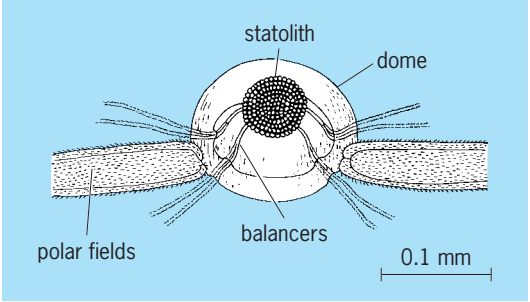

Fig. 3. Features of a statocyst.

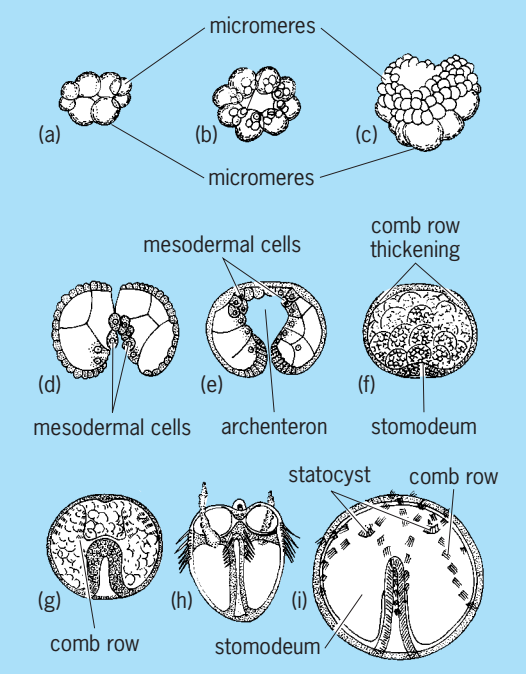

Fig. 4. Embryology. (*a*) Eight-cell stage. (*b*) Micromeres, viewed from above. (*c*) Later stage in division of micromeres. (*d*) Gastrula, section. (*e*) Invagination. (*f*) Stomodeum and thickenings for comb rows. (*g*) Stomodeum invaginated; comb rows differentiated. (*h*) Cydippid larva. (*i*) Embryo with statocysts and comb rows.

two main planes of symmetry of the animal. Gastrulation is first epibolic (the macromeres are gradually covered by the multiplication of the micromeres), but ends with invagination of the last micromeres produced by macromeres. Development of the embryo is determinate. The claimed endodermal formation of a true mesoderm, giving rise to muscular and mesenchymal cells, is still debatable. The earliest larva formed in the development of all types of ctenophores is the cydippid, with a globular body, a pair of primary tentacles (or rudiments), and eight costae of comb plates. In the Lobata, Cestida, and Platyctenida, the larva undergoes secondary changes in later development.

Physiology. The ctenophores move through the sea with the oral pole forward, using comb plates as oars. The beating of successive plates occurs along each costa by metachronal waves and is initiated for the two costae of the same quadrant by the corresponding balancer of the statocyst. Balancer and comb-row activities and some muscular contractions are integrated by the nervous system.

Almost all ctenophores are luminescent. This capacity, linked to a calcium-dependent photoprotein, is assumed by specialized cells underlying the comb plates.

The ctenophores feed on zooplankton. Prey are trapped with the tentacles (Cydippida, Cestida) or oral extensions of the body (Lobata), or are engulfed directly through the widely opened mouth (Beroida). Extracellular digestion begins in the pharynx, where gland cells secrete enzymes. The partially digested material is then carried throughout the gastrovascular system, where it is phagocytized by digestive cells.

Ctenophores have high powers of regeneration. Asexual reproduction (in a few platyctenid genera) is by regeneration of an entire organism from a small piece of the adult body.

Ecology. Ctenophores are themselves important plankton organisms. They are quite common and have a worldwide distribution in all seas, where they can appear in enormous swarms. Some genera stand great changes in the seawater salinity. Because of their voracity as predators of zooplankton, they play an important role in plankton equilibrium and in fisheries.

Phylogeny. The Ctenophora constitute a well-defined phylum. Among its orders, the Cydippida are undoubtedly the most primitive, with radiating evolution of the other orders. The ancestor of the phylum is possibly to be found among the Trachylina (Cnidaria, Hydrozoa). However, because of such forms as *Coeloplana* (Platyctenida), ctenophores have also been suspected of affinity with the polyclad Turbellaria. *See* POLYCLADIDA; TRACHYLINA.

J.-M. Franc

Bibliography. G. R. Harbison, L. P. Madin, and N. R. Swanberg, On the natural history and distribution of oceanic ctenophores, *Deep-Sea Research*, 25:233–256, 1978; L. H. Hyman, *The Invertebrates: Protozoa through Ctenophora*, 1940; A. G. Mayer, *Ctenophores of the Atlantic Coast of North America*, Carnegie Institution of Washington, Publ. 162, 1912; S. C. Morris et al. (eds.), *The Origins and Relationships of Lower Invertebrates*, 1985; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Ctenostomata

A paraphyletic order of Bryozoa in the class Gymnolaemata. Ctenostomes have entirely cuticular body walls, and colonies consist of feeding zooids (individual bryozoans) only or of feeding zooids plus spine- or stolon-shaped zooids. *See* BRYOZOA; GYMNOLAEMATA.

Morphology. Feeding zooids possess tentacles that collectively form a circular, feeding, bell-shaped lophophore centered on the mouth. The mouth is located atop a cuticular tube enclosing the initial portion of the U-shaped digestive tract, and the terminal portion of the digestive tract ends as an anus part of the way up the cuticular tube. When the lophophore is retracted into its protected position within the main body of the zooid, the cuticular tube inverts and surrounds the lophophore as a tentacle sheath, and the opening through which it protrudes to feed is closed by a sphincter muscle. At the muscular diaphragm of the inverted cuticular tube, ctenostomes have a cuticular collar referred to as either a setigerous or pleated collar (see **illustration**). The collar extends as a cone around the lower parts of the lophophore when the zooid is feeding but

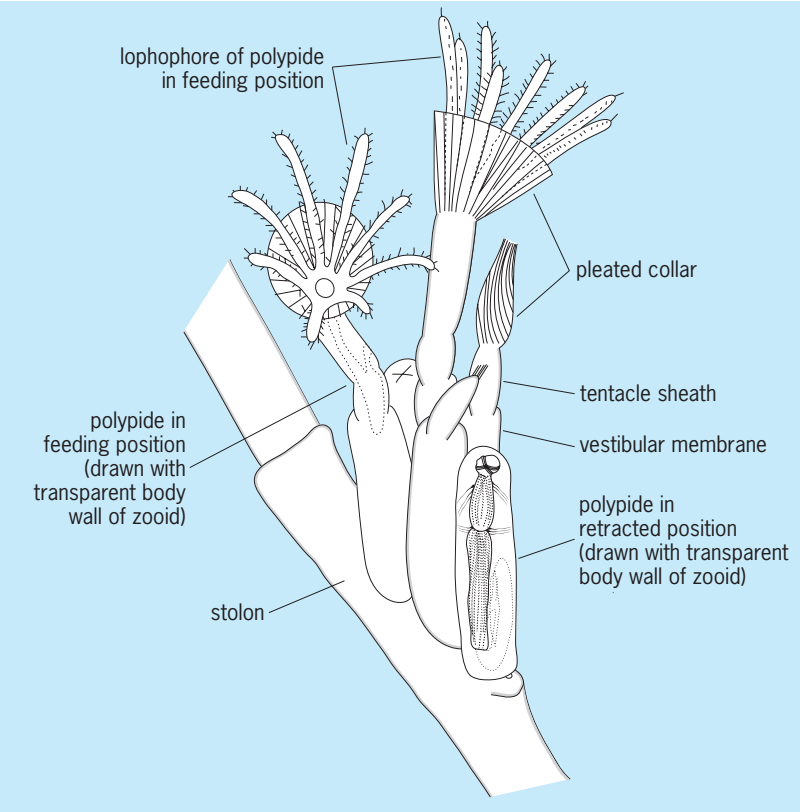

Cluster of zooids of the ctenostome genus *Amathia*.

at other times is collapsed into a narrow, longitudinally pleated tube that extends from the region of the sphincter muscle at the outer tip of the zooid (see illustration). Many, perhaps all, cuticular collars have several longitudinal rods of thickened cuticle that project from the outer end of the cone to give a comblike appearance; hence the name of the group (cteno- = comb; stome = mouth).

Most ctenostomes grow as encrusting or erect, ramifying, thin branches. Some ctenostomes consist entirely of feeding zooids (for example, *Victorella*), others of individual feeding zooids (for example, *Bowerbankia*) or feeding zooid groups (for example, *Amathia*; see illustration) connected by zooids specialized as stolons. Other ctenostome colonies, such as *Alcyonidium*, consist entirely of densely packed feeding zooids that form flexible encrustations or thick, erect, branched growths. A few branching ctenostomes are endolithic borers, creating their living space by dissolving carbonate substrates such as shells just below the surface.

Life cycle. Colonies are hermaphroditic. Fertilized eggs are usually brooded as embryos in the space normally occupied by the retracted tentacles of feeding zooids. In other species, brooding may be within specialized invaginations or within the coelom. A few species release fertilized eggs directly into the sea. The free-swimming larvae attach to a substrate for metamorphosis (including absorption and histolysis of all larval organs), resulting in the formation of the first feeding zooid of the colony, from which all subsequent zooids within the developing colony are budded.

Most ctenostomes are marine, although several species are tolerant of lower salinity than any other marine bryozoans and a few live in freshwater. Ctenostomes range from the low intertidal zone to the abyss.

History and classification. The fossil record of endolithic ctenostomes extends back to the Late Ordovician. The lack of a mineralized skeleton diminishes chances of preservation of nonboring ctenostomes. However, where rapidly growing skeletons of organisms such as oysters extend over ctenostomes, the ctenostome shapes may be molded in surprising detail on the undersurfaces of the overgrowing organism by a process called bioimmuration. Bioimmured ctenostomes are known back to the Jurassic. *See* JURASSIC; ORDOVICIAN.

Ctenostome bryozoans are relatively primitive in their lack of calcification, small number of zooid types, lack of brooding in some species, and other features. The oldest known Cheilostomata, of the Late Jurassic, have zooidal shapes and encrusting colony growth habits that are close to morphologies known in extant ctenostomes and apparently are derived from ctenostomes. The class Stenolaemata may have originated from ctenostomes by splitting off the muscle-bearing mesodermal layer from the original outer body wall as calcification rigidified the lateral body walls of cylindrical zooids, leaving the flexible terminal wall and the mesodermal layer of the lateral walls involved in protrusion of tentacles to the feeding position. *See* STENOLAEMATA. Frank K. McKinney

Bibliography. P. J. Hayward, *Ctenostome Bryozoans*, E. J. Brill/Dr. W. Backhuys, London, 1985; J. S. Ryland and J. S. Porter, The identification, distribution and biology of encrusting species of *Alcyonidium* (Bryozoa: Ctenostomatida) around the coasts of Ireland, in *Biology and Environment*; *Proceedings of the Royal Irish Academy*, in press; J. A. Todd, The central role of ctenostomes in bryozoan phylogeny, pp. 104–135, in A. Herrera Cubilla and J. B. C. Jackson (eds.), *Proceedings of the 11th International Bryozoology Association Conference, Balboa, Panama*, Smithsonian Tropical Research Institute, 2000.

Ctenothrissiformes

A small order of teleostean fishes that are of particular interest because they seem to be on or near the main evolutionary line leading from the generalized, soft-rayed salmoniforms to the spiny-rayed beryciforms and perciforms, dominant groups among higher bony fishes.

Ctenothrissiformes lack fin spines but they share the following characters with primitive beryciforms: pelvic fins that are thoracic, with seven or eight rays, and pectoral fins that are placed well up on the

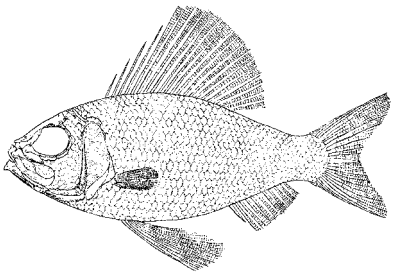

Cretaceous ctenothrissid, *Ctenothrissus radians*; length 10–12 in. (25–30 cm). (*After C. Patterson, Phil. Trans. Roy Soc. London, Ser. B, Biol. Sci., no. 739, vol. 247, 1964*)

flank (see **illus.**); a subocular shelf; a scaly opercle; a mesocoracoid lacking; a similar premaxilla; maxillary teeth; an orbitosphenoid; and 17 branched caudal rays. Except for the last three these features persist in generalized Perciformes and indicate a major evolutionary trend away from typical soft-rayed fishes. The scales are ctenoid or cycloid, and there are nine branchiostegals. *See* SCALE (ZOOLOGY).

The few fishes confidently assigned to the Ctenothrissiformes, *Ctenothrissus* (Ctenothrissidae) and *Aulolepis* and *Pateroperca* (Aulolepidae), are known as Upper Cretaceous marine fossils from Europe and southwestern Asia. *See* BERYCIFORMES; PERCIFORMES; SALMONIFORMES. Reeve M. Bailey

Bibliography. C. Patterson, A review of Mesozoic Acanthopterygian fishes, with special reference to those of the English Chalk, *Phil. Trans. Roy. Soc. London*, Ser. B, Biol. Sci., no. 739, vol. 247, 1964.

Cube

A parallelepiped whose six faces are all squares. The cube (see **illus.**) is one of the five regular solids known to the ancient Greeks, who proposed the

A cube, a type of parallelepiped.

famous problem (now proved to be unsolvable) of constructing, with the use of the compass and unmarked straightedge alone, a cube whose volume is twice that of a given cube. The edge of the desired cube may be found by use of conic sections. A cube has eight vertices and twelve edges. Each vertex is on three edges and three faces, each edge is on two vertices and two faces, and each plane is on four vertices and four edges. *See* PARABOLA; POLYHEDRON; REGULAR POLYTOPES. Leonard M. Blumenthal

Cubeb

The dried, nearly ripe fruit (berries) of a climbing vine, *Piper cubeba*, of the pepper family (Piperaceae). This species is native to eastern India and Indomalaysia. The crushed berries are smoked, and the inhaled smoke produces a soothing effect in certain respiratory ailments. Cubebs are used in medicine as a stimulant, expectorant, and diuretic. Medicinal properties of cubeb are due to the presence of a volatile oil which formerly was thought to stimulate healing of mucous membranes. *See* PIPERALES. P. D. Strausbaugh; Earl L. Core

Cubozoa

An order of the class Scyphozoa. They are good swimmers and their distribution is mostly tropical. The umbrella is cubic, and a pedalium, a gelatinous leaf-shaped body, is present at the base of each ridge of the exumbrella (see **illus.**). A well-developed

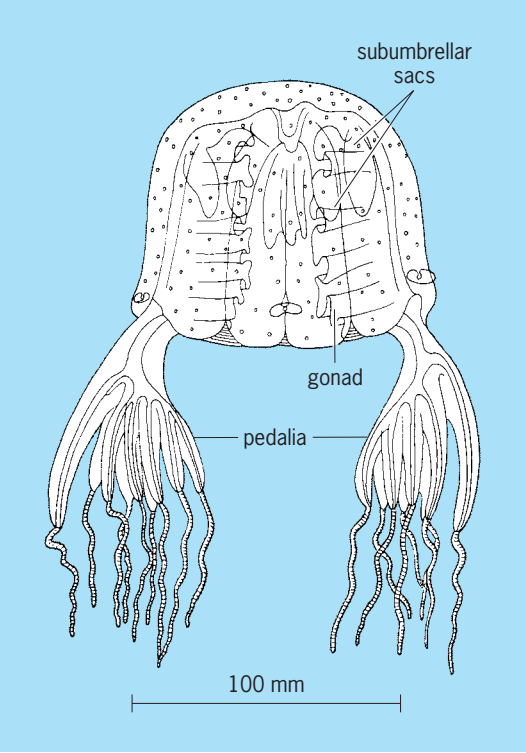

Cubomedusan, *Chiropsalmus quadrumanni*. (*After L. Hyman, The Invertebrates, vol. 1, McGraw-Hill, 1940*)

tentacle, or cluster of tentacles, arises from each pedalium. The sensory organs are on the perradii. They are the most highly developed in the Coelenterata, each one having six ocelli and a statocyst. The manubrium, a peduncle connecting the mouth with the stomach, has four lips. The cubic stomach connects with four radial pockets, and between these

pockets are the interradial septa. The leaf-shaped gonads develop from the septa and extend into the pockets. No connecting stage between polyp and medusa is known. *See* SCYPHOZOA. Tohru Uchida

Cuculiformes

An order of birds having zygodactyl feet (two toes are in front and two behind) in which the fourth toe is reversed, but not having any close relationship to other birds with this foot structure, such as the parrots or piciforms. The relationship of the touracos to the cuckoos has been strongly debated and is still unresolved. However, this conclusion is very controversial because the hoatzin possesses an anisodactyl pattern of the toes (three toes are in front and one is in back); thus, many authors still place the hoatzin in its own order next to the Cuculiformes. *See* OPISTHOCOMIFORMES; PICIFORMES; PSITTACIFORMES.

Classification. The order Cuculiformes is divided into two suborders: Musophagi, containing the single family Musophagidae (touracos; 18 species), and Cuculi, with the large and diverse family Cuculidae (cuckoos; 129 species). Cuculidae is divided into six distinct subfamilies, some of which have a limited distribution, such as the couas (Couinae; 10 species) which are confined to Madagascar.

Fossil record. The fossil record of the Cuculiformes is poor and comprises only scattered finds. Although extant members of the Musophagidae are found only in Africa, several fossils of this family have been found in the Oligocene and Miocene of Europe, indicating that this group once had a wider distribution.

Touracos. Touracos are African woodland birds with a short- to medium-length crest on their heads, and a loose green, gray, or brown plumage with patches of bright yellow, red, or violet. Most species feed largely on fruit, although several species also consume insects. One species, the Great Blue touraco (*Corythaeola cristata*), also eats leaves, especially late in the day; it feeds regurgitated leaves to its young after their first week of life. The green pigment (touracoverdin) in the feathers is unique and is the only true green pigment known in birds; the red pigment (touracin) is water-soluble and copper-based and is unique to birds. The pigment in the feathers is protected from washing away by overlying feathers. The nest is a bulky platform placed in a tree. Incubation and care of the chicks is shared by both parents. The young leave the nest at an early age (about 10 days) and crawl about the trees before being able to fly.

Roadrunner (*Geococcyx* sp.). (*Photo by Marguerite Gregory, © 2004 California Academy of Sciences*)

Cuckoos. Cuckoos are a morphologically and behaviorally diverse worldwide group, as indicated by the six recognized subfamilies, of which three are found only in the Old World and three only in the New World. Their size varies from small (30 g or 0.07 lb) to large (channel-billed cuckoo, *Scythrops novaehollandiae*, at 600 g or 1.32 lb). They are mostly arboreal, but members of three subfamilies are largely or entirely ground-dwelling; some, such as the American roadrunners (*Geococcyx*; see **illustration**), are fast runners, with top speeds of about 15 mi/h (25 km/h), and rarely fly. Most cuckoos are secretive birds, sulking in heavy vegetation. They are strong fliers but do not fly often. Temperate species are migratory. Cuckoos feed on insects, small vertebrates (including poisonous snakes taken by roadrunners, *Geococcyx*), and other animals, and less commonly on fruit; no species feeds entirely on fruit. Most are solitary, but the New World anis and guiras (Crotophaginae) live in flocks and build a communal nest housing several females.

All species of Old World cuckoos (Cuculinae) are nest parasites, as are three species of American ground-cuckoos (Neomorphinae). The common cuckoo of Europe is the most studied nest parasite, and has many adaptations in both the adult female and the young to ensure success in having the host species rear the young. Females of these species belong to different "egg varieties," in which the eggs resemble those of different hosts. The eggs laid by cuckoo females in host nests are small compared to the size of the adult bird. Incubation is short, and the young cuckoo pushes the other eggs or young out of the nest. The host parents feed the parasitic chick, which grows to a gigantic size compared with the adult host. Most American species build their own nests and care for their own young.

Cuckoos are valuable as insect eaters, especially of toxic and hairy caterpillars rejected by other birds. The European cuckoo has held an important place in European folklore, literature, and music, as has the roadrunner in the southwestern United States and in Mexico. *See* AVES. Walter J. Bock

Bibliography. N. B. Davies, *Cuckoos, Cowbirds and Other Cheats*, T & A D Poyser, 2000; R. B. Payne, *The Cuckoos*, Oxford University Press, 2005; D. A. Turner, Family Musophagidae (Turacos), in J. del Hoyo et al. (eds.), *Handbook of the Birds of the World*, vol. 4, pp. 480–506, 1997.

Cucumber

A warm-season annual cucurbit (*Cucumis sativus*) native to India and belonging to the plant order Campanulales. *See* CAMPANULALES.

Production. The cucumber is grown for its immature fleshy fruit which is used primarily for pickling and for slicing as a salad (**Fig. 1**). The development of gynoecious (bearing only female flowers) and hybrid varieties (cultivars) with increased disease resistance is resulting in a continuing change in the varieties of both pickling and slicing cucumbers. In addition, the advent of mechanical harvesting and chemical weed control for pickling cucumbers is altering traditional methods of culture, for example, higher plant populations per acre and single harvests instead of multiple pickings. Important cucumber-producing states are Florida and South Carolina for fresh market and Michigan and Wisconsin for pickling. *See* SQUASH.

H. John Carew

Fig. 1. Cucumber fruits. (*Asgrow Seed Co.*; *Upjohn Co.*)

Diseases. Diseases of cucumbers reduce plant growth, fruit yield, and product quality in both field and greenhouse plantings. Because cucumbers are grown all months of the year, it is necessary to guard against seeding in unfavorable environments and to control seed decay and seedling blight.

Bacterial diseases. Bacterial wilt is a serious disease in the midwestern, north-central, and northeastern regions of the United States, but it is not important in the South or West. Plants infected with the causal bacterium *Erwinia tracheiphila* first show a wilting of one or a few leaves which remain green, but later all the leaves wilt and the plant dies. A white, stringy, viscid bacterial ooze shows in freshly cut wilting stems. Bacteria-infested 12-spotted and striped cucumber beetles infect the plant while feeding upon it. The bacteria multiply in the plant and become distributed throughout the vascular system. The bacteria overwinter in the body of the insect. Because the only way in which the plants can become infected is through the feeding of the cucumber beetles, control depends upon early and prompt destruction of these pests by insecticides. Antibiotics have sometimes proved useful against the bacterium.

Angular leaf spot, incited by the bacterium *Pseudomonas lachrymans*, causes water-soaked spots on leaves and bordered, irregular lesions on fruits. The lesions later develop into a brown firm rot extending into the flesh (**Fig. 2**). In moist conditions, bacteria ooze from the lesion in "tears" which dry into a whitish residue. The causal bacterium overwinters on infected vine refuse, is seed-borne, and is spread by rain and surface water. Control consists of using California seed, which is low in infection, soaking it for 5 min in a mercuric chloride solution (1 part in 1000 parts of water), rinsing the seed in water, and planting promptly. Fixed copper sprays can achieve control, but must be used carefully as they can be phytotoxic to young plants. There are some resistant varieties available.

Fungal diseases. Downy mildew is a destructive disease in the eastern and southern states where the weather is warm and moist, is less damaging in the north-central states, and rarely occurs in the Southwest. It is incited by the leaf-inhabiting fungus *Pseudoperonospora cubensis*, which causes angular, yellowish spots; the spots later turn brown when the

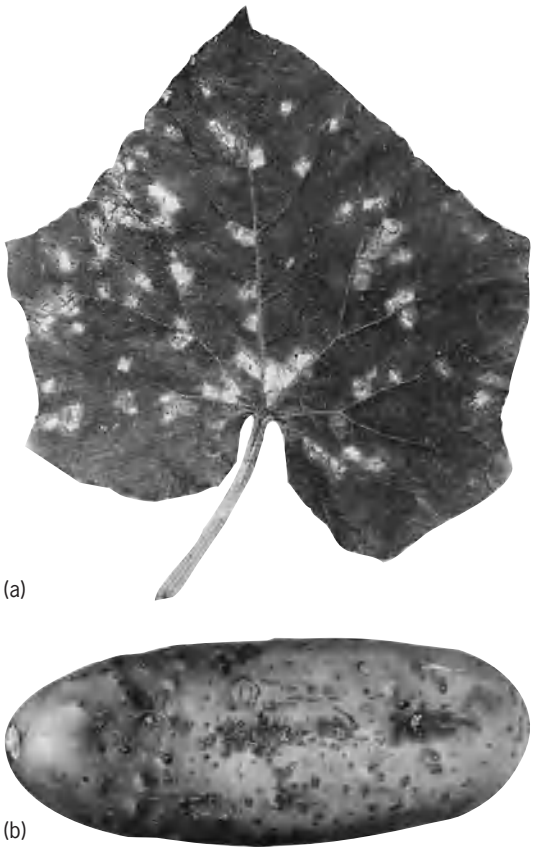

Fig. 2. Angular leaf spot disease of cucumber caused by bacterium *Pseudomonas lachrymans*. (*a*) Leaf spots. (*b*) Fruit lesions. (*Courtesy of J. C. Walker*)

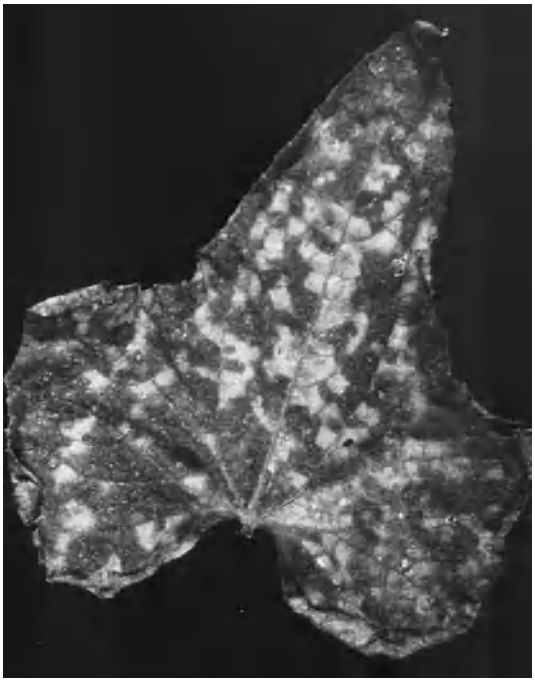

Fig. 3. Downy mildew of cucumber leaf showing yellow angular blotches induced by *Pseudoperonospora cubensis.* (Courtesy of D. E. Ellis)

leaves shrivel (**Fig. 3**). In moist conditions, the lower leaf surface is covered with the fruiting layer of the fungus. Entire vines may be killed. The loss of foliage interferes with normal flower set and fruit development. Maturing fruits fail to color properly, are tasteless, and are usually sunburned. Planting of resistant varieties, such as Ashley, Palmetto, Palomar, Santee, and Stono, and treating plants with copper fungicides will provide maximum protection from downy mildew.

Scab is a serious disease in the northern and northeastern states, especially during cool moist weather. The causal fungus, *Cladosporium cucumerinum*, produces small, circular, halo-bordered, water-soaked lesions on leaves and stems. The greatest damage is to the fruit and appears as sunken, dark-brown spots on immature fruit and as rough corky lesions on mature fruit. The fungus survives in plant refuse and on the seed. Two-year crop rotation, seed treatment, and well-drained soil are important in control of scab. Control is achieved by planting resistant varieties, such as the slicing variety Highmoor or the pickling varieties Wisconsin SR 6 and SMR 12.

Viral diseases. Cucumber mosaic is the most destructive disease caused by a virus. The pathogen causes mottling of terminal leaves and dwarfing of vines, as well as mottling and varying malformations of the fruits. Plants can be infected at any growth stage. The virus is harbored in a number of perennial hosts and is spread by aphids. The best means of control is provided by planting resistant varieties, such as the slicing varieties Burpee Hybrid, Niagara, Ohio MR 200, Sensation Hybrid, Shamrock, and Surecrop Hybrid, and pickling varieties Ohio MR 17, Ohio MR 25, Wisconsin SMR 12, and Yorkstate.

See HEMIPTERA; PLANT PATHOLOGY; PLANT VIRUSES AND VIROIDS.

Frank L. Caruso

Culture

The cultivation of cells in the laboratory. Bacteria and yeasts may be grown suspended in a liquid medium or as colonies on a solid medium (usually solidified by 1.5–2.5% agar); molds grow on moist surfaces; and animal and plant cells (tissue cultures) usually adhere to the glass or plastic beneath a liquid medium. Cultures must provide sources of energy and raw material for biosynthesis, as well as a suitable physical environment.

The materials supplied determine which organisms can grow out from a mixed inoculum. Some bacteria (prototrophic) can produce all their constituents from a single organic carbon source; hence they can grow on a simple medium (including K^+, NH_3, Mg^{2+}, Fe^{3+}, PO_4^{3-}, and SO_4^{2-}). Other cells (auxotrophic) lack various biosynthetic pathways and hence require various amino acids, nuclei acid bases, and vitamins. Obligatory or facultative anaerobes grow in the absence of O_2; many cells require elevated CO_2. Cultures isolated from nature are usually mixed; pure cultures are best obtained by subculturing single colonies. Viruses are often grown in cultures of a host cell, and may be isolated as plaques in a continuous lawn of those cells.

In diagnostic bacteriology, species are ordinarily identified by their ability to grow on various selective media and by the characteristic appearance of their colonies on test media. These tests include hemolytic reactions with sheep or other red cells embedded in the solid medium. Indicator dyes in the medium are used to detect colonies that ferment various sugars.

In ordinary cultures the cells are at all possible stages in their division cycle, and the composition of the medium changes continually as a result of their metabolism (until growth ceases, in the stationary phase of the culture). On transfer of an inoculum to fresh medium, there may be a lag phase, without multiplication, followed by a phase of exponential growth. Synchronous cultures are achieved by blocking growth or harvesting cells at a specific stage; the cells then divide in synchrony for several generations. In continuous cultures, fresh medium flows into the vessel and full-grown culture flows at the same rate (such as in a chemostat); the cells are therefore harvested from a medium of constant composition. *See* BACTERIAL GROWTH; CHEMOSTAT.

Laboratory cultures are often made in small flasks, test tubes, or covered flat dishes (petri dishes). Industrial cultures for antibiotics or other microbial products are usually in fermentors of 10,000 gallons (37,850 liters) or more. The cells may be separated from the culture fluid by centrifugation or filtration.

Bernard D. Davis

Technique. Specific procedures are employed for isolation, cultivation, and manipulation of microorganisms, including viruses and rickettsia, and for

propagation of plant and animal cells and tissues. A relatively minute number of cells, the inoculum, is introduced into a sterilized nutrient environment, the medium. The culture medium in a suitable vessel is protected by cotton plugs or loose-fitting covers with overlapping edges so as to allow diffusion of air, yet prevent access of contaminating organisms from the air or from unsterilized surfaces. The transfer, or inoculation, usually is done with the end of a flamed, then cooled, platinum wire. Sterile swabs may also be used and, in the case of liquid inoculum, sterile pipets.

The aqueous solution of nutrients may be left as a liquid medium or may be solidified by incorporation of a nutritionally inert substance, most commonly agar or silica gel. Special gas requirements may be provided in culture vessels closed to the atmosphere, as for anaerobic organisms. Inoculated vessels are held at a desired constant temperature in an incubator or water bath. Liquid culture media may be mechanically agitated during incubation. Maximal growth, which is visible as a turbidity or as masses of cells, is usually attained within a few days, although some organisms may require weeks to reach this stage. *See* CHEMOSTAT; EMBRYONATED EGG CULTURE; TISSUE CULTURE.

Media. Combinations of many nutrients are used for the growth of microorganisms as well as plant and animal tissues. Water is the chief constituent of most media. A medium may consist of natural materials, such as enzymatic digests or extracts of yeast or beef. Milk, potato slices, and chick embryos are other common examples. Artificial media are prepared by mixing various ingredients according to particular formulas. A complex medium contains at least one crude ingredient derived from a natural material, hence of unknown chemical composition. A chemically defined, or synthetic, medium is one in which the chemical structure and amount of each component are known.

All media furnish minerals, sources of nitrogen, carbon, and energy, and sometimes vitamins or other growth factors. An ingredient furnishing carbon—usually the chief source—is called the substrate for the utilizing organism.

Selective medium. Usually this medium contains an individual organic compound as the sole source of carbon, nitrogen, or sulfur for growth of an organism. The organism oxidizes the carbon compound, thereby obtaining energy for vital processes. Strictly inorganic media are utilized by some bacteria, the chemoautotrophs. They obtain carbon exclusively from carbon dioxide and energy from the oxidation of ammonia, nitrate, hydrogen gas, or reduced, inorganic, sulfur compounds. Algae and a few bacterial types are photoautotrophs, securing their energy from light.

An agar medium of a given composition permits selection, through development of colonies, or organisms capable of utilizing the specific sources of carbon, energy, nitrogen, or sulfur available. A colony on such a medium is prima-facie evidence that such organisms have been selected, but this must be confirmed by restreaking (subculturing on agar medium) on a medium of the same composition.

Enrichment media. These are liquid media of a given composition which permit preferential emergence of certain organisms that initially may have made up a relatively minute proportion of a mixed inoculum. This enrichment of the population in the desired type expedites isolation by direct plating on the corresponding selective solid medium. Because of the numerous subtle, competitive forces, enrichment cultures yield only a relatively small proportion, that is, the fastest growers, of all the organisms potentially capable of developing in a given enrichment medium.

Indicator media. These media, usually solid, contain substances capable of changing color in the vicinity of the colony which has effected a particular chemical change, such as fermenting a certain sugar. Thus, organisms with distinctive types of metabolism may be readily detected by the color change in the medium, even when the surface of the medium contains many other colonies. Indicators may be used in liquid media to characterize a pure culture of organisms.

Elective culture. No single culture medium or environment will support growth of more than a small proportion of the known types of microorganisms. Each organism has characteristic nutritional and environmental requirements. Thus, in a culture composed of more than one type of organism, the medium and conditions are selective for only one type of organism present and the other types are suppressed. These selective, or elective, culture principles and techniques are invaluable for the detection and isolation of specified organisms from natural sources. Elective culture techniques are more effective for obtaining bacteria from natural sources than for obtaining fungi, actinomycetes, and other groups; the faster growth rates, larger numbers, and greater variety of bacteria usually confer a competitive advantage upon them. In general, elective culture procedures are applicable to organisms capable of utilizing particular sources of energy, carbon, nitrogen, and other nutrients in a given physicochemical environment. There is no principle whereby elective cultures may be employed to obtain selectively organisms capable of producing particular compounds. Basically, two procedures are available, direct isolation and liquid enrichment.

Direct isolation. This requires a solid medium, usually agar, in petri plates. The surface is lightly sprinkled with finely divided particles of soil or other inoculum material. Only organisms favored by that medium and those conditions will produce colonies of appreciable size; these may then be purified. Literally thousands of other kinds of bacteria fail to develop, although they might develop in other selective media. Thus, the direct isolation technique is a means of selecting certain minority members from natural, mixed microbial populations. Since most of the colonies are on a solid medium, develop independently, and are physically separate, competition rarely is a factor determining the emergence of

the various colonies. Although certain fast-growing and characteristically spreading colonies may tend to overgrow neighboring colonies, the chief advantage of the procedure is that slow-growing bacteria have a chance to reveal themselves. This technique is, therefore, the best for obtaining the largest variety of bacteria capable of developing on a given medium. Colonies from a inoculum appearing under any one environmental condition are a small representation of a wider spectrum of bacteria capable of developing. For example, two sets of plates incubated, respectively, in the presence and absence of oxygen will yield different organisms, as will plates adjusted to different hydrogen ion concentrations (pH) and plates incubated at various temperatures.

Elective media containing common organic compounds, such as carbohydrates or amino acids, will yield large numbers of different organisms. Similarly, most such organisms are capable of utilizing a number of different but related compounds. However, certain substrates and conditions are highly specific for a few organisms or even for a single kind of organism. Occasionally, growth of a colony occurs at the expense of degradation of biosynthetic products or accessory growth factors which are liberated from one or more adjacent colonies and diffuse into the medium.

Although the direct-isolation method does not permit preferential development of colonies capable of forming specific products, such colonies as may be desired may be detected among those on an agar plate. Thus, acid producers may be distinguished by a zone of color change of an indicator in the medium or by dissolution of powdered calcium carbonate suspended in the medium. Clear zones result from the destruction of other insoluble substrates, such as starch, proteins, and lipids, by colonies producing the corresponding enzymes. Other metabolic products may be detected by flooding or spraying colony-containing plates with chemical or biological reagents, giving a distinctive reaction with the desired product.

The specificity can be enhanced by incorporation of substances that selectively inhibit growth of certain kinds of organisms, especially of those in the majority. For example, in a medium adjusted to pH 4.0–4.5 most bacteria and actinomycetes are inhibited, and thus filamentous fungi and yeasts may be selectively obtained. Other chemicals are used to inhibit undesirable organisms in specimens of body fluids, thereby facilitating isolation of the pathogenic bacteria in the minority.

Liquid enrichment. A small amount of soil or other natural material is added to a prescribed fluid medium which then is incubated. After appreciable growth of bacteria occurs, a drop of the culture is used to inoculate a second flask of the same medium. Organisms originally representing small minorities are thus preferentially selected by the medium and the conditions, and by the time turbidity develops in the secondary culture, they represent practically the majority populations. The chief disadvantage of this procedure is that only a small proportion of all the bacterial types potentially capable of developing in a given medium will be found. Potentially competent organisms are in direct competition, and the specific conditions usually select relatively few of the fastest growers. Liquid-enrichment cultures also contain organisms unable to attack the original substrate; these scavenge products of degradation and biosynthesis, as well as living and dead cells, of the primary organisms.

Pure culture. A pure culture contains cells of one kind, all progeny of a single cell. The concept of a pure culture is an important one in microbiology because most considerations of microorganisms require dealing with only one kind of a cell at a time. Many techniques have been evolved for the isolation and maintenance of pure cultures.

Isolation of aerobic bacteria. Isolation of pure cultures may be achieved by a variety of methods, the main differences being in isolation of aerobic versus anaerobic bacteria.

1. *Pour-plate procedures.* These involve a thorough distribution of a microbial mixture in cooled, melted agar medium. After solidification of the medium, individual cells are embedded and proliferate to form a colony. Purity of the colony cannot be taken for granted since it may have arisen from two adjacent cells (mixed colony); so the colony is dispersed in sterile water and the pour-plate procedure repeated. Well-isolated colonies in the second plating ordinarily can be shown to be pure. This method is based on the use of a homogeneous suspension of individual cells. The great majority of the colonies on secondary plates should be similar and, seen microscopically, the cells of such colonies should be one kind. However, certain organisms mutate characteristically, producing different colony types (dissociation).

2. *Streak plates.* The surface of the solid medium in a petri dish is streaked with an inoculating wire or glass rod in a continuous movement so that most of the surface is covered. The later parts of the streak should not intersect with the earlier parts. In later portions of the streak, the separate colonies are the outgrowth of individual cells. The chief advantage of this method is isolation of majority members of a mixed population.

3. *Roll-tube technique.* This technique, employed chiefly in tissue culture, achieves a homogeneous gaseous and nutrient environment by the elimination of diffusion gradients. During incubation the tubes are held in a wheellike instrument at an angle of about 15° from the horizontal. The wheel rotates vertically about once every 2 min. Constant liquid film formation over the inner surface of the tube facilitates oxygen transfer to the medium.

4. *Submerged culture.* In this method the microorganisms are grown in a liquid medium subjected to continuous, vigorous agitation. A homogeneous cellular suspension results and all cells in the medium are simultaneously provided with nearly optimal rates of diffusion of oxygen and nutrients. The mechanical agitation may be achieved by continuous bubbling of sterile air through the medium, by a

propeller system, by combination of both, or by continuous rotary or reciprocating shaking machines.

Isolation of anaerobic bacteria. Special cultural conditions may be provided for anaerobes, organisms that do not require oxygen or that grow only in the complete absence of oxygen. Fluid thioglycollate medium, the Brewer anaerobic petri dish, and the Brewer anaerobic jar are some of the methods used (see **illus.**). Fluid thioglycollate medium contains sodium thioglycollate, a reducing compound which lowers the oxygen content of the medium. In the anaerobic petri dish, the thioglycollate agar medium and the petri dish cover, with its restricted air space and close fit to the medium surface, give anaerobic conditions. The anaerobic jar is operated by replacing the air in the jar with hydrogen or illuminating gas. A platinized catalyst heated by electricity causes the oxygen remaining in the jar to combine with hydrogen to form water or, in the case of illuminating gas, water and carbon dioxide. In this way free oxygen is removed from the jar.

1. *Shake culture.* This is a method for isolating anaerobic bacteria. Various dilutions of an inoculum are throughly mixed with cooled, melted agar medium occupying about two-thirds the volume of test tubes. After solidification a 1-in. (2.5-cm) layer of melted vaspar, a 50:50 mixture of paraffin wax and petrolatum, is added. When the mixture has hardened, the culture is sealed against diffusion of oxygen during incubation. In tubes containing higher dilutions of inoculum, separated discrete colonies develop in different positions in the agar. With soft glass tubes, the glass is cut and the agar column broken in the region of a desired colony, which is then isolated. Sometimes the seal is first removed by warming that area of the tube; the column of agar is then slowly ejected into a sterile container by compressed air carried to the bottom of the tube through a fine capillary tube introduced between the agar and the glass. Colonies are reached by slicing the agar with a sterile scalpel.

2. *Anaerobic plates.* Pour or streak plates are incubated in a closed vessel in which the air is replaced by an inert oxygen-free gas such as helium or nitrogen. Alternatively, the oxygen in a closed vessel is absorbed chemically by mixing pyrogallol and alkali. Specially designed petri dishes permit oxygen absorption in individual dishes.

Single-cell isolation. Single-cell isolation employs microscopic techniques to verify the presence of a single cell, which is then used to initiate a culture.

1. *Little-drop technique.* Minute drops of dilutions of a cellular suspension are examined microscopically on a glass slide. A drop containing a single cell is transferred with a capillary pipet to an appropriate culture medium. This procedure is best suited for isolation of large cells, such as yeasts and mold spores, which can be located quickly with the high-dry lens of the compound microscope. Evaporation problems and the difficulty of obtaining drops small enough to examine with the oil immersion lens limit this procedure for bacteria. A variation of this procedure was originated by

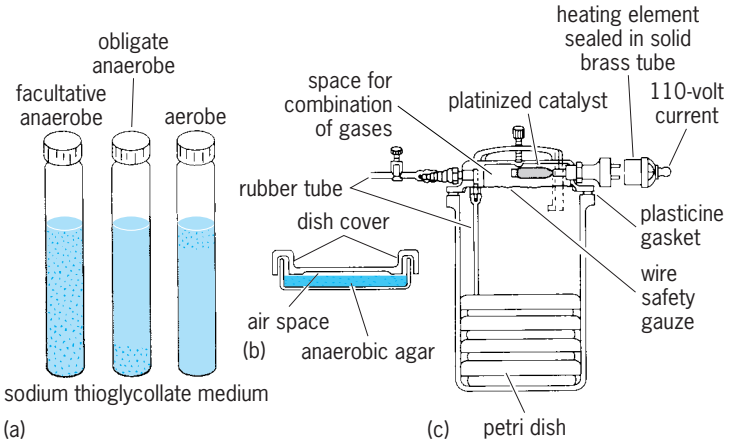

Some techniques for the cultivation of anaerobic bacteria. (*a*) Fluid thioglycollate medium showing aerobic, anaerobic, and facultative anaerobic growth. (*b*) Brewer anaerobic petri dish. (*c*) Brewer anaerobic jar. (*Baltimore Biological Laboratory, Inc.*)

J. Lederberg. Little drops of a diluted cellular suspension are deposited on a slide covered with a thin film of oil. The Lederberg technique, by permitting rapid isolation of single bacterial cells, obviates to some extent the problems enumerated before.

2. *Micromanipulation.* The disadvantages of the little-drop technique for the isolation of single bacterial cells are eliminated with this technique. A micromanipulator permits study of drops in a moist chamber on the microscope stage, reducing evaporation difficulties and enabling the use of the oil immersion lens. The manipulations are performed by remotely controlled capillary micropipets activated by air pressure. *See* MICROMANIPULATION.

Maintenance. This is concerned with the preservation of culture viability and is accomplished by practices that reduce the basal metabolism of the cells, thereby enhancing longevity and eliminating the need for frequent transfers.

1. *Lyophilization.* This procedure, also called freeze-drying, is the best method for preserving culture viability. Freshly grown cells are suspended, in a glass tube, in a small volume of a sterile fluid, such as blood serum or skim milk. A small amount of a reducing agent such as ascorbic acid may help prolong the viability of the culture. The suspension is quickly frozen at approximately −94°F (−70°C) by immersion in a dry ice–acetone mixture. The frozen material is then dehydrated, without melting, under high vacuum, or the gas space may be filled with nitrogen or helium prior to sealing. Storage in the absence of oxygen may preserve the viability of a culture for years, even without refrigeration. The mortality, during lyophilization and storage, varies with the culture and the suspending fluid. Lyophilized cells are rejuvenated by inoculation into fresh nutrient medium.

2. *Slants.* Slants are the most convenient method of maintenance for periods of 1–2 months. Cotton-plugged tubes containing an agar medium are slanted after sterilization until the agar solidifies. The inoculum is streaked on the agar surface. After maximal growth at the appropriate temperature, the slant culture is stored in a refrigerator and used repeatedly as

inoculum. With long storage periods, the tube should be sealed to retard desiccation. Viability of slant cultures not used frequently may be prolonged by filling the tube with sterile mineral oil to about 1 in. (2.5 cm) above the agar. This retards water loss from and oxygen access to the culture.

3. *Stabs.* Stabs are used to maintain anaerobic bacteria. A tube about two-thirds filled with agar medium is pierced vertically for about 2 in. (5 cm) with an inoculating wire. Growth develops in the anaerobic depths of the agar along the stab.

4. *Strict anaerobe maintenance.* Strict anaerobes, in liquid or solid media, may be maintained in an atmosphere from which the oxygen has been absorbed by alkaline pyrogallol. Also, the air may be replaced by an inert gas, such as helium or nitrogen. The most convenient procedure employs liquid media containing approximately 0.1% of a reducing agent, such as sodium thioglycollate, to keep the oxidation-reduction potential low enough for growth of anaerobes and to protect it from oxygen diffusing from the air. Special vessels or atmospheres are not required, but fully developed cultures are sealed against oxygen and refrigerated.

5. *Use of sterile soil.* Sterile soil may be used to carry stock cultures of sporeforming organisms. Autoclaved soil supports limited growth of the organism, which then sporulates. The spores remain viable indefinitely when kept refrigerated.

Jackson W. Foster; R. E. Kallio

Cumacea

A crustacean within the Malacostraca: Peracarida. All have a well-developed carapace which is fused dorsally with at least the first three thoracic somites and which overhangs the sides to enclose a branchial cavity. The telson may be free or may appear lacking because it is fused with the last abdominal somite. Eyes are sessile and, except for a few genera, are coalesced in a single median organ or wholly wanting. There are eight pairs of thoracic appendages, the first three modified as maxillipeds and the last five as peraeopods. The first maxillipeds bear an epipodite forming a respiratory organ with a forward-directed siphon and a posterior part carrying branchial lobes. Natatory exopodites may be present on the third maxillipeds and the first four pairs of peraeopods or some of them. The uropods are styliform. There is no free larval stage. In the few species whose life history has been investigated there are five to seven molts after release from the brood pouch.

There is always some sexual dimorphism in adults, but immature males resemble females until the later stages. The second antennae are always vestigial in females and nearly always well developed with a long flagellum in males (see **illus.**). Females often have a reduced number of thoracic exopodites. Pleopods are never present in the female but there may be one to five pairs in the male. The form and sculpturing of the body and the number of spines and setae on the appendages often differ between the sexes. Oostegites are present on the third to sixth thoracic limbs of the female.

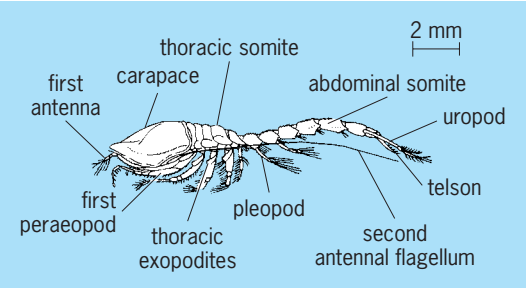

Typical adult male cumacean.

All known species have a characteristic shape owing to the inflated carapace and slender abdomen. Most are small, 0.04–0.48 in. (1–12 mm) in length; the largest reach 1.4 in. (35 mm). The order is rather uniform and classification is based on trivial characters. Seven families are recognized. Affinities are probably closer to the Mysidacea than to the Tanaidacea. A few fossils are now known.

Cumacea are found in all seas and in some estuaries and areas of brackish water, including the Caspian Sea. Most inhabit water shallower than 650 ft (200 m), but many occur in great depths. They normally burrow in sand or mud with the front protruding. Coastal species, and especially the more active adult males, may often be captured above the bottom. They feed on organic detritus or very small organisms and may themselves form part of the diet of various fishes. About 740 species have been described and placed in 82 genera. Many more remain to be found, as they have been poorly investigated in most areas. *See* MALACOSTRACA; SEXUAL DIMORPHISM.

Norman S. Jones

Bibliography. H. G. Bronn, *Klassen and Ordnungen des Tierreichs*, vol. 5, bk. 4, 1941; N. S. Jones, *Galathea Reports*, vol. 10, 1969; R. Lankester, *A Treatise on Zoology*, pt. 7, fasc. 3, 1909; J. W. Martin and G. E. Davis, *An Updated Classification of the Recent Crustacea*, Nat. Hist. Mus. Los Angeles County Sci. Ser., no. 39, 2001; D. Maruzzo et al., Appendage loss and regeneration in arthropods: A comparative review, in S. Koenemann and R. A. Jenner, Crustacea and arthropod relationships, *Crustacean Issues*, 16:215–245, Francis & Taylor, 2005; P. A. McLaughlin et al., *Common and Scientific Names of Aquatic Invertebrates from the United States and Canada: Crustaceans*, Amer. Fisheries Soc. Spec. Publ., no. 31, 2005; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; T. H. Waterman (ed.), *The Physiology of Crustacea*, 2 vols., 1960, 1961.

Cumin

Cuminum cyminum, a plant for which the whole or ground dried ripe fruit, commonly called seed, is a popular spice. It is a major ingredient in both chili

powder and curry, and is added to meat sauces, rice, bread, pickles, soups, and other foods.

Roman caraway is another common name for this member of the parsley family (Apiaceae). The only species in its genus, cumin exhibits a variety of plant types depending on the seed source. A small annual herb about 1–2 ft (0.3–0.6 m) tall, cumin grows upright as a single slender stem with many branches. The deep green leaves are 0.5–2 in. (1.3–5 cm) long, thin, and finely divided. *See* APIALES.

The strong, pungent green-spicy odor and flavor of cumin is attributable largely to cuminaldehyde, the main constituent of the essential oil, and other aldehydes.

This herb is native to the Mediterranean region. Presently cumin is commercially grown in Iran, southern Russia, China, India, Morocco, and Turkey. The three major types of cumin seed, Iranian, Indian, and Middle Eastern, vary in seed color, essential oil quantity, and flavor. Cumin does best in a mild growing climate, free from hot dry periods, on well-drained sandy loam soil. In 90–120 days the plant matures; the small fragile plant is generally hand-harvested. Commonly cumin is planted at 30–35 lb/acre (33.7–39.3 kg/ha), and yields 400–500 lb seed/acre (450–560 kg/ha). Upon distillation cumin seed yields 2.5–5% essential oil, used in both perfumery and flavoring liqueurs. Cumin is also used medicinally. *See* SPICE AND FLAVORING.

Seth Kirby

Cummingtonite

An amphibole (a double-chain silicate mineral) with the idealized chemical formula $Mg_7Si_8O_{22}(OH)_2$ that crystallizes with monoclinic symmetry. Naturally occurring samples generally are solid solutions between $Mg_7Si_8O_{22}(OH)_2$ and the corresponding iron (Fe)-end member with the general formula $(Mg,Fe)_7Si_8O_{22}(OH)_2$. The name cummingtonite (derived from the location Cummington, Massachusetts) is applied to all solid solutions with $Mg/(Mg + Fe^{2+}) \geq 0.5$, whereas those with $Mg/(Mg + Fe^{2+}) < 0.5$ are termed grunerite. Up to a total of 1.0 (Ca + Na) atom per formula unit may also be present in cummingtonite.

The amphibole anthophyllite also has the idealized formula $Mg_7Si_8O_{22}(OH)_2$, but crystallizes with orthorhombic symmetry. The conditions that favor formation of one amphibole over the other are not fully understood, but both are known to coexist in the same rock. It has been proposed that cummingtonite is stable at higher temperature, and inverts, on cooling, to anthophyllite by a process that involves unmixing of a more calcium-rich amphibole (tremolite) and the formation of an intermediate metastable monoclinic form of cummingtonite.

Cummingtonite commonly occurs as aggregates of fibrous crystals, often in radiating clusters. It is transparent to translucent, varies in color from white to green to brown, and may be pale to dark depending primarily on the iron content. Hardness is 5–6 on Mohs scale; density is 3.1–3.6 g/cm^3 (1.8–2.1 oz/in.^{3}), increasing with increasing iron content. Cummingtonite has distinct, prismatic (110) cleavage that produces cross sections with approximately 124° and 56° angles, which are definitive for amphibole group minerals. Chemical, optical, or x-ray techniques usually are necessary to unambiguously distinguish cummingtonite from other amphibole minerals.

Cummingtonite is generally considered to be a metamorphic mineral, but it has been found in silicic volcanic rocks and, rarely, plutonic igneous rocks. It occurs in a variety of metamorphic rock types (amphibolite, schist, gneiss, granulite) that have undergone medium- to high-grade metamorphism. It commonly occurs in the calcium- and aluminum-poor environment of metamorphosed iron formation. It can also be a constituent of metamorphosed mafic and ultramafic igneous rocks, where it may coexist with other amphibole minerals such as hornblende, tremolite, and anthophyllite. With increasing intensity of metamorphism, cummingtonite is commonly replaced by pyroxene-bearing mineral assemblages. *See* AMPHIBOLE; HORNBLENDE; IGNEOUS ROCKS; METAMORPHIC ROCKS; TREMOLITE.

Robert K. Popp

Bibliography. C. Klein and C. S. Hurlbut, Jr., *Manual of Mineralogy*, Wiley, New York, 1993; B. E. Leake et al., Nomenclature of Amphiboles: Report of the Subcommittee on Amphiboles of the International Mineralogical Association, Commission on New Minerals and Mineral Names, *Amer. Mineralog.*, 82:1019–1037, 1997; D. R. Veblen, Amphiboles and other Hydrous Pyriboles—Mineralogy, *Rev. Mineralogy*, vol. 9A, 1981; D. R. Veblen and P. H. Ribbe, Amphiboles: Petrology and Experimental Phase Relations, *Rev. Mineralogy*, vol. 9B, 1982.

Curie temperature

The critical or ordering temperature for a ferromagnetic or a ferrimagnetic material. The Curie temperature (T_C) is the temperature below which there is a spontaneous magnetization M in the absence of an externally applied magnetic field, and above which the material is paramagnetic. In the disordered state above the Curie temperature, thermal energy overrides any interactions between the local magnetic moments of ions. Below the Curie temperature, these interactions are predominant and cause the local moments to order or align so that there is a net spontaneous magnetization.

In the ferromagnetic case, as temperature T increase from absolute zero, the spontaneous magnetization decreases from M_0, its value at $T = 0$. At first this occurs gradually, then with increasing rapidity until the magnetization disappears at the Curie temperature (see **illus.**). Many physical properties show an anomaly or a change in behavior at the Curie temperature. There is a peak in specific heat and in magnetic permeability, and there are changes in the behavior of such properties as electrical

Ordering temperatures of magnetic materials*

Ferromagnets	Curie temperature (T_C), °F (K)	Ferrimagnets	Curie temperature (T_C), °F (K)	Antiferromagnets	Néel temperature (T_N), °F (K)
Co	2039 (1388)	Fe_3O_4	1085 (858)	NiO	620 (600)
Fe	1418 (1043)	$MnFe_2O_4$	572 (573)	Cr	100 (311)
$CrBr_3$	−393 (37)	$Y_3Fe_5O_{12}$	548 (560)	$FeCl_2$	−417.0 (23.7)
$GdCl_3$	−455.7 (2.2)				

*From F. Keffer, *Handbuch der Physik*, vol. 18, pt. 2, Springer, 1966.

resistivity, elastic properties, and thermal expansivity. In ferrimagnetic materials the course of magnetization with temperature may be more complicated, but the spontaneous magnetization disappears at the Curie temperature. The Curie temperature can be determined from magnetization versus temperature measurements or from the related anomalies in other physical properties. The Curie temperature can be altered by changes in composition, pressure, and other thermodynamic parameters.

Well above the Curie temperature the magnetic susceptibility follows the Curie-Weiss law, and in fact the characteristic temperature θ in that the law is called the paramagnetic Curie temperature. Generally θ is slightly above T_C. The susceptibility deviates from Curie-Weiss behavior as the temperature approaches T_C and θ. *See* CURIE-WEISS LAW; FERRIMAGNETISM; FERROMAGNETISM; MAGNETIC SUSCEPTIBILITY; MAGNETIZATION; PARAMAGNETISM.

In antiferromagnetic materials the corresponding ordering temperature is termed the Néel temperature (T_N). Below the Néel temperature the magnetic sublattices have a spontaneous magnetization, though the net magnetization of the material is zero. Above the Néel temperature the material is paramagnetic. *See* ANTIFERROMAGNETISM.

The ordering temperatures for magnetic materials vary widely (see **table**). The ordering temperature for ferroelectrics is also termed the Curie temperature, below which the material shows a spontaneous electric moment. *See* FERROELECTRICS; PYROELECTRICITY.

J. F. Dillon, Jr.

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; S. Chikazumi and S. H. Charap, *Physics of Magnetism*, 1964, reprint 1978; A. H. Morrish, *Physical Principles of Magnetism*, 1966, reprint 1980.

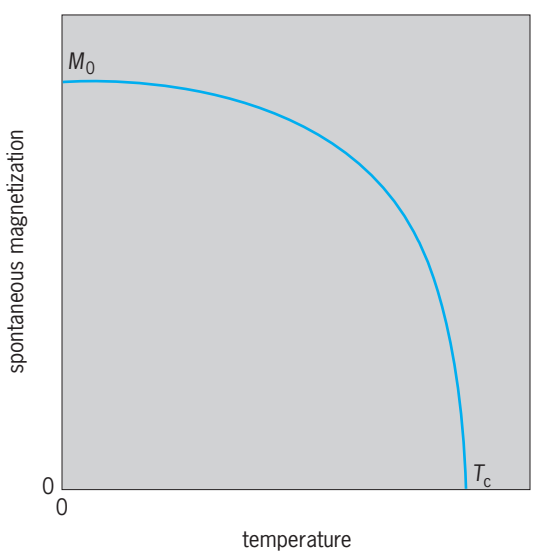

Variation of spontaneous magnetization of a ferromagnetic material with temperature.

Curie-Weiss law

A relation between magnetic or electric susceptibilities and the absolute temperature which is followed by ferromagnets, antiferromagnets, nonpolar ferroelectrics and antiferroelectrics, and some paramagnets. The Curie-Weiss law is usually written as the equation below, where χ is the susceptibility, C is a

$$\chi = C/(T - \theta)$$

constant for each material, T is the absolute temperature, and θ is called the Curie temperature. Antiferromagnets and antiferroelectrics have a negative Curie temperature. The Curie-Weiss law refers to magnetic and electric behavior above the transition temperature of the material in question. It is not always precisely followed, and it breaks down in the region very close to the transition temperature. Often the susceptibility will behave according to a Curie-Weiss law in different temperature ranges with different values of C and θ. *See* CURIE TEMPERATURE; ELECTRIC SUSCEPTIBILITY; MAGNETIC SUSCEPTIBILITY.

The Curie-Weiss behavior is usually a result of an interaction between neighboring atoms which tends to make the permanent atomic magnetic (or induced electric) dipoles point in the same direction. The strength of this interaction determines the Curie temperature θ. In the magnetic case the interaction is caused by Heisenberg exchange coupling.

In many paramagnetic salts, especially those of the iron group, the Curie-Weiss behavior is due to crystalline distortions and their effect upon the atomic orbital angular momenta which contribute to the magnetization. For further discussion and derivation of the Curie-Weiss law. *See* ANTIFERROMAGNETISM; FERROELECTRICS; FERROMAGNETISM.

Elihu Abrahams; Frederic Keffer

Curium

A chemical element, Cm, in the actinide series, with an atomic number of 96. Curium does not exist in the terrestrial environment, but may be produced artificially. The chemical properties of curium are so similar to those of the typical rare earths that, if it were not for its radioactivity, it might easily be mistaken for one of these elements. The known isotopes of curium include mass numbers 238–250. The isotope ^{244}Cm is of particular interest because of its potential use as a compact, thermoelectric power source, through conversion to electrical power of the heat generated by nuclear decay. *See* PERIODIC TABLE.

1												13	14	15	16	17	18
1 **H**	2																2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Metallic curium may be produced by the reduction of curium trifluoride with barium vapor. The metal has a silvery luster, tarnishes in air, and has a specific gravity of 13.5. The melting point has been determined as $2444 \pm 72°F$ ($1340 \pm 40°C$). The metal dissolves readily in common mineral acids with the formation of the tripositive ion.

A number of solid compounds of curium have been prepared and their structures determined by x-ray diffraction. These include CmF_4, CmF_3, CmC_3, $CmBr_3$, CmI_3, Cm_2O_3, and CmO_2. Isostructural analogs of the compounds of curium are observed in the lanthanide series of elements. *See* ACTINIDE ELEMENTS; TRANSURANIUM ELEMENTS. Glen T. Seaborg

Bibliography. S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; G. T. Seaborg (ed.), *Transuranium Elements: Products of Modern Alchemy*, 1978; G. T. Seaborg and W. D. Loveland, *The Elements Beyond Uranium*, 1990.

Currant

A fruit (berry) in the genus *Ribes* in the family Saxifragaceae. Cultivated black and red currants and gooseberries all belong to this genus. *Ribes* species having prickles or spines are called gooseberries, and those that do not are called currants. The berries are produced in clusters on bushes, and cultivars ripen in midsummer (**Fig. 1**).

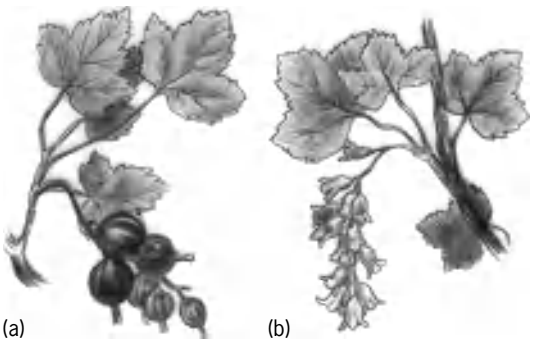

Fig. 1. Black-currant (*Ribes nigrum*). (*a*) Branch with cluster of fruits. (*b*) Branch with flowers.

Varieties. *Ribes* are not widely grown commercially for fruit in the United States, but red currants and gooseberries are popular in home gardens for use in jellies and pies. Several *Ribes* species, particularly golden currant (*R. aureum*) and fuchsia-flowered gooseberry (*R. speciosum*), are used as shrubs in landscaping. In the United States important red currant cultivars include Cherry, Minnesota 71, Perfection, Red Lake, White Imperial, and Wilder; popular gooseberry cultivars include Fredonia, Pixwell, Poorman, Red Jacket, and Welcome. In Europe cultivars of the black currant (*R. nigrum*) are extensively grown commercially for juice concentrate. Wild *Ribes* species occur widely in the United States, and desirable, edible types are sometimes gathered during the season when the fruit is locally abundant. *See* GOOSEBERRY.

Currant and gooseberry diseases. Many of the same diseases affect both currants and gooseberries. Several virus diseases occur in Europe in these crops, reducing yield and fruit quality, often without showing symptoms on the leaves. In Great Britain and several other European countries where *Ribes* species are important crops, local growers can purchase virus-tested, certified plants from nurseries, thus reducing the likelihood of planting virus-infected stock. Tomato ringspot virus, causing leaf-spotting patterns and reduced vigor in red currants in the United States, is spread in the soil by the American dagger nematode (*Xiphinema* sp.), and can be controlled by appropriate soil fumigation prior to planting.

Anthracnose is a leaf- and fruit-spotting fungus disease caused by *Pseudopeziza ribis* (**Fig. 2**). It may be serious in wet seasons, causing leaves to yellow and drop by midsummer. Control involves elimination of overwintered leaves that harbor the fungus; pruning to allow rapid drying of interior foliage; and use of approved fungicides, usually fixed coppers, at bloom and harvest.

Powdery mildew (caused by *Sphaerotheca morsuvae*) is a fungus disease that affects cultivated *Ribes*, particularly gooseberries, producing white, powdery coatings that later darken on young shoots, leaves, and fruits. Powdery mildew often stunts infected plants and renders the fruit visually unappealing. The disease is hard to control, but sanitation and spraying with approved fungicides just as the buds open in spring and during the growing season often help.

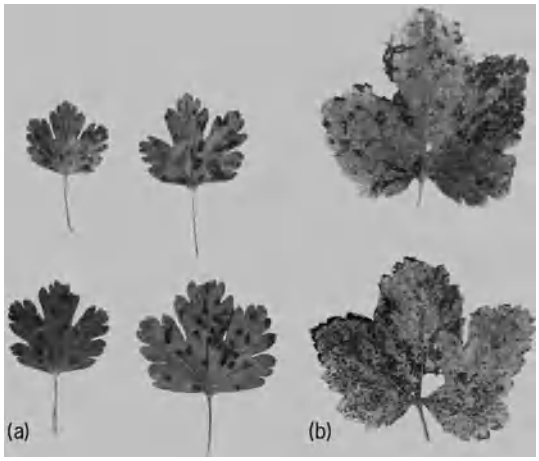

Fig. 2. Anthracnose disease on (*a*) gooseberry (small leaves) and (*b*) currant (large leaves).

Other important fungus diseases of *Ribes* include white pine blister rust (caused by *Cronartium ribicola*) on black currant, leaf spot (caused by *Mycosphaerella ribis*), and cane blight (caused by *Botryosphaeria ribis*). *See* PLANT PATHOLOGY; ROSALES.

Richard H. Converse

Bibliography. H. W. Anderson, *Diseases of Fruit Crops*, 1956; N. W. Frazier (ed.), *Virus Diseases of Small Fruits and Grapevines*, 1970; J. S. Shoemaker, *Small Fruit Culture*, 1975.

Current balance

An apparatus with which force is measured between current-carrying conductors. The purpose of the measurement was to establish the value of the ampere in terms of its International System (SI) definition; therefore the relative locations of the conductors had to be measured so that the force between them could be computed from their geometry. The difficulty of this measurement limited the accuracy of even the best instruments to a few parts in a million.

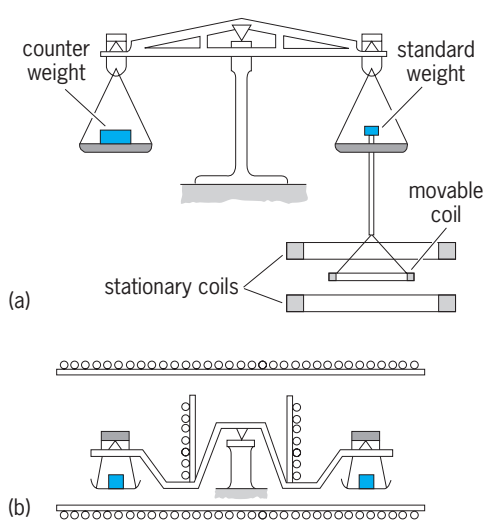

Current balances. (*a*) Rayleigh. (*b*) Pellat.

Therefore, current balances have now been superseded by other methods. *See* ELECTRICAL UNITS AND STANDARDS; PHYSICAL MEASUREMENT.

Operation. In a current balance, force is measured between fixed and movable coils (see **illus.**). If the coils carry the same current, Eq. (1) holds for a linear displacement, or Eq. (2) for an angular displacement, where F is the mechanical force or Γ is the mechanical torque, M is the coefficient of mutual inductance, and I is the current being measured. The coil construction must be such that $\partial M/\partial x$ or $\partial M/\partial\theta$ can be computed from dimensional measurements. The movable coil is attached to a balance, and the force between the coils is balanced by the gravitational force on a known mass.

$$F = I^2 \frac{\partial M}{\partial x} \tag{1}$$

$$\Gamma = I^2 \frac{\partial M}{\partial \theta} \tag{2}$$

Types. Three types of current balance have been used in ampere determinations. In two of these, the Rayleigh and the Ayrton-Jones balances, the fixed and movable coils are coaxial, and the motion is linear along the coil axis. In the third type, the Pellat balance, the axes of the fixed and movable coils are at right angles, and the movable coil turns about its center, the pivot point of the weighing balance.

In the Rayleigh balance the coils are circular and have a small square winding section. The fixed coils are located above and below, but coaxial with, the smaller, movable coil. The coil spacing is such that the force on the movable coil is maximum. Because of the small winding section, the ratio of radii of fixed and movable coils ratio can be calculated from the ratio of ampere-turns which produce cancellation of the magnetic flux densities at the center of the coil system. The radius of the movable coil is about half that of the fixed coils because for this ratio the exact dimensions of the winding cross section are of reduced importance.

In the Ayrton-Jones and Pellat balances, the fixed and movable coils are single-layer solenoids and the coil construction is such that the diameter and winding pitch can be measured accurately.

The current being measured was passed through a standard resistor, and the resulting voltage drop was compared to the emf of a standard cell. The resistor and cell were used to preserve the values of the electrical units in the intervals between SI ampere and ohm determinations. *See* CURRENT MEASUREMENT.

Bryan P. Kibble

Current comparator

A device that measures very accurately an unknown current in terms of a known one. Current comparators are based on Ampère's law, Eq. (1), where the

$$\sum I_p \cdot n_p = \oint \vec{H} \cdot d\vec{l} \tag{1}$$

left side of the equation is the sum of currents, each calculated as the current I_p in a conductor multiplied by the number of times n_p it threads a closed path in space (this product is the so-called ampere-turns or current linkage), and the right side is the integral of the magnetic field strength $\vec{H}$ that the currents produce around the entire path. If an instrument could detect when the result of the path integral was exactly zero, then (supposing that only two currents I_1 and I_2 thread the path n_1 and n_2 times, respectively) we would know that there was accurate ampere-turn balance, that is, Eq. (2) is valid. The shape and position of the current circuits is in principle immaterial. *See* AMPÈRE'S LAW; MAGNETISM.

$$I_1 \cdot n_1 = I_2 \cdot n_2 \tag{2}$$

Room-temperature current comparators. Ideal flux line–integral detectors do not exist. If alternating currents are being compared, detecting the voltage induced at the contiguous ends of a solenoid uniformly close-wound on the surface of a tube that follows the path, which is usually circular (**Fig. 1**), is an approximation which will be limited in accuracy by imperfections of the winding to a fraction of a percent. Sensitivity can be greatly enhanced if the tube contains a highly magnetically permeable material. If the currents are steady direct currents, the magnetic material and the coil will be operated as a fluxgate magnetometer. To attain a million times better accuracy than a percent requires in practice that the current circuits are so positioned and shaped that the resulting magnetic field follows the circular path as closely as possible. In room-temperature current comparators, this is accomplished by winding the current-carrying conductors on nested toroids that enclose the detector's path. Additionally, this path is contained within a toroidal shield of magnetically highly permeable material (Fig. 1). This suppresses all field components except those following the detector's toroidal path. *See* MAGNETOMETER.

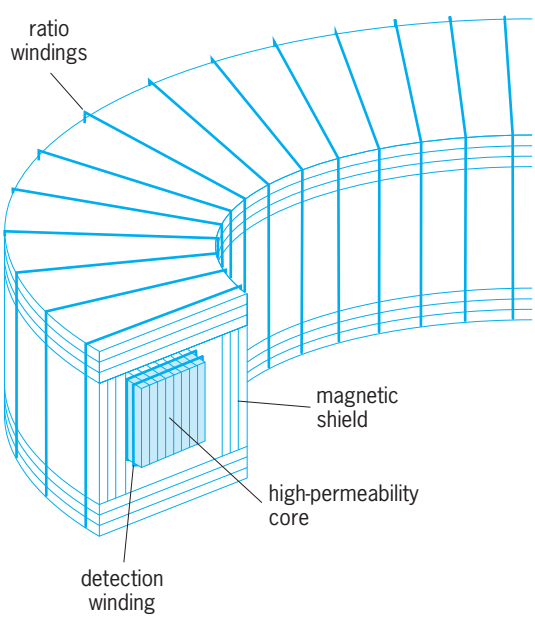

Fig. 1. Room-temperature alternating-current (AC) comparator.

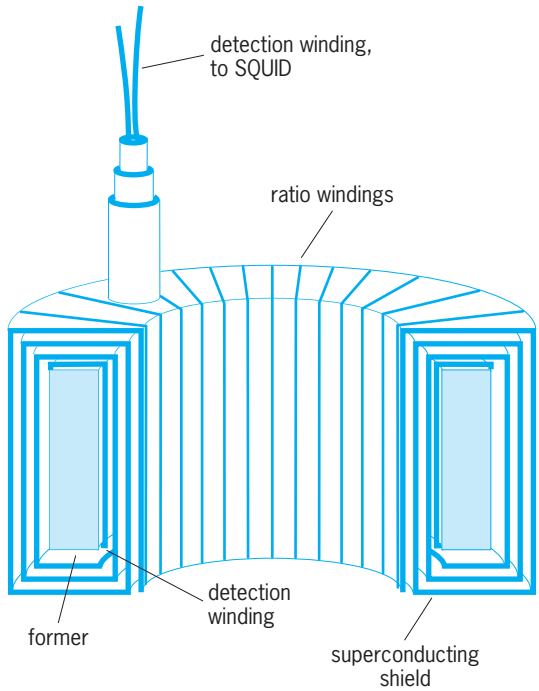

Fig. 2. Cryogenic current comparator.

Cryogenic current comparators. In superconducting cryogenic comparators (**Fig. 2**), which are principally used in national measurement laboratories in order to attain the highest possible level of demonstrable accuracy, flux shaping is achieved by making use of the Meissner effect. Several types of cryogenic comparators exist, but in all the current-carrying windings thread a toroidal superconducting shield, consisting of a tube folded back on itself, creating an overlap of several turns (Fig. 2). The Meissner effect ensures that the magnetic field inside the superconductor decreases exponentially with distance from the surface, rapidly becoming negligibly small, and consequently the flux line integral along a path well within the superconductor is also negligibly small. It follows from Ampère's law that the total current linkage enclosed by the path must also be negligible. As a consequence, screening currents flow on the surface of the superconductor, which oppose the current linkage imposed by any imbalance in the currents in the windings. On the surface that faces the windings, the distribution of the surface currents depends on the position of the current-carrying windings; but on the other side, well away from the ends, the current (and the field) distribution becomes very nearly independent of the position of windings. A detection winding in this region is therefore sensitive only to the total current linkage of the ratio windings. The perfection of the flux shaping is such that the pick-up coil need not be wound with as much precision as is needed for the room-temperature instrument. *See* MEISSNER EFFECT.

Flux detection in cryogenic current comparators is usually achieved by using a superconducting detection winding connected directly to the input of a SQUID. Compared to a room-temperature direct-current (DC) comparator, this has the advantage that no excitation winding is required, as a direct-current in the windings results in a DC signal at the output of the SQUID system. On the other hand, operation of a current comparator with a SQUID, which is very sensitive to electrical interference, requires the use of purpose-built low-noise electronics. A cryogenic comparator can achieve an accuracy exceeding 1 part in 10^9. *See* SQUID; SUPERCONDUCTING DEVICES.

In addition to imperfect flux shaping, another kind of error results from the electric fields associated with the potential drops down the resistive conductors carrying the currents. If the currents change in time, for example, if alternating currents are being compared, the time dependence of the electric fields results in displacement currents whose magnetic effects are indistinguishable from currents flowing in conductors. Fortunately, displacement currents are zero if steady direct currents are being compared, and their effect is very small if the frequency of the alternating current is only a few hundred hertz and the resistance of the windings is not unduly great. *See* DISPLACEMENT CURRENT.

Applications. Applications of current comparators range from the measurement of very large currents (of the order 1000 A for room-temperature comparators), to medium currents, such as used in the most accurate resistance measurements, and small currents (of the order 1 nA), such as generated by single-electron transport. *See* CURRENT MEASUREMENT; ELECTRICAL MEASUREMENTS; RESISTANCE MEASUREMENT. Peter Kleinschmidt; Bryan P. Kibble

Bibliography. K. Grohmann et al., Ironless cryogenic current comparator for AC and DC applications, *IEEE Trans. Instrum. Meas.*, IM-23(4):261–263, 1974; B. P. Kibble et al., *A Guide to Measuring Direct and Alternating Current and Voltage below 1 MHz*, Institute of Measurement and Control/National Physical Laboratory, 2003; W. J. M. Moore and P. N. Miljanic, *The Current Comparator*, 1988; J. M. Williams and A. Hartland, An automated cryogenic current comparator resistance ratio bridge, *IEEE Trans. Instrum. Meas.*, 40(2):267–270, 1991.

Current density

A vector quantity equal in magnitude to the instantaneous rate of flow of electric charge, perpendicular to a chosen surface, per unit area per unit time. If a uniform wire of cross-sectional area A carries a current I, the current density J is I/A. The units of J in the International System (SI) are amperes per square meter.

The concept of current density is useful in treating the flow of electricity through a conductor of nonuniform cross section, and in electromagnetic theory. In terms of current density Ohm's law can be written $J = E/\rho$, where E is electric field strength and ρ is the local resistivity. Current density is related to the number per unit volume n, the charge e, and the average velocity υ of the effective charge carriers in a conductor by the formula $J = ne\upsilon$. John W. Stewart

Current measurement

The measurement of the rate of passage of electric charges in a circuit. The unit of measurement, the ampere (A), is one of the base units of the International System of Units (SI). It is defined as that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross section, and placed 1 m apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newton per meter of length.

Determination of the ampere. In order to establish an electrical unit in accordance with the SI definition, it is necessary to carry out an experimental determination. Obviously the ampere cannot be realized exactly as defined. Electromagnetic theory has to be used to relate a practical experiment to the definition. For many years the current balance was used, but its accuracy is limited to about 10 parts per million by the difficulties of the mechanical measurements to define coils and convection currents. *See* CURRENT BALANCE.

An alternative determination of the SI unit of current, devised by B. P. Kibble, depends on the availability of a satisfactory SI standard of resistance. The unit of resistance can be set up through the calculable capacitor, which is employed in several national standards laboratories. Through the concept of frequency (the reciprocal of time, another SI base quantity), the unit of impedance obtained from a calculable capacitor can be used to calibrate a resistor in SI units. In Kibble's virtual power balance, a coil is suspended partly in a strong permanent magnetic field. The experiment has two parts. First, a current I is passed through the coil and the resulting force F is measured by weighing. The current and force are proportional and related by a factor depending on the geometry of the apparatus. Second, the coil is moved at a controlled speed and the voltage induced in the coil is measured. The voltage V and the rate of displacement (dx/dt) are also proportional and depend on the same geometrical factor. The results of the two operations can be combined to eliminate the geometrical factor, giving the result shown in the equation below, where electric power = mechanical power.

$$V \times I = F \times \frac{dx}{dt}$$

The result of the experiment, when combined with the SI unit of resistance (V/I), allows the ampere to be determined and all the electrical units to be set up in accordance with the SI definitions. *See* RESISTANCE MEASUREMENT.

Current or virtual power balances are not convenient for use as working standards. Since

January 1, 1990, working standards of voltage and resistance have provided the foundations of practical electrical measurements. The standard of voltage is based on the alternating-current (ac) Josephson effect, in which voltage is related to frequency. By international agreement the value 483 597.9 GHz/V for the Josephson constant is now used throughout the world. The working unit of resistance is maintained through the quantum Hall effect, with an agreed value of 25 812.807 ohms for the voltage-to-current ratio obtained under certain defined experimental conditions. These values have been chosen to provide the best known approximations to the SI units and have the advantage of reproducibility at the level of 1 part in 10^8. The working standard of current is derived from measurements of voltage across a known resistor. *See* ELECTRICAL UNITS AND STANDARDS; HALL EFFECT; JOSEPHSON EFFECT.

Direct and low-frequency currents. The moving-coil (d'Arsonval) meter measures direct currents (dc) from 10 microamperes to several amperes. The accuracy is likely to be a few percent of the full-scale indication, although precision instruments can achieve 0.1% or even better. Above 1 milliampere a shunt usually carries the major part of the current; only a small fraction is used to deflect the meter. Since the direction of deflection depends on the direction of the current, the d'Arsonval movement is suitable for use only with unidirectional currents. Rectifiers are used to obtain dc and drive the meter from an ac signal. The resulting combination is sensitive to the rectified mean value of the ac waveform.

In the moving-iron meter, two pieces of soft magnetic material, one fixed and one movable, are situated inside a single coil. When current flows, both pieces become magnetized in the same direction and accordingly repel each other. The moving piece is deflected against a spring or gravity restoring force, the displacement being indicated by a pointer. As the repulsive force is independent of current direction, the instrument responds to low-frequency ac as well as dc. The natural response of such a movement is to the root-mean-square (rms) value of the current. The accuracy of moving-iron meters is less than that of moving-coil types. *See* AMMETER.

Radio-frequency currents. For radio-frequency applications it is essential that the sensing element be small and simple to minimize inductive and capacitive effects. In a thermocouple meter the temperature rise of a short, straight heater wire is measured by a thermocouple and the corresponding current is indicated by a d'Arsonval movement. In a hot-wire ammeter the thermal expansion of a wire heated by the current is mechanically enhanced and used to deflect a pointer. Both instruments, based on heating effects, respond to the rms value of the current. Above 100 MHz, current measurements are not made directly, as the value of current is likely to change with position owing to reflections and standing waves. *See* MICROWAVE MEASUREMENTS; THERMOCOUPLE.

Large currents. Above 50 A the design of shunts becomes difficult. For ac, current transformers can be used to reduce the current to a level convenient for measurement. At the highest accuracy, current comparators may be used in which flux balance is detected when the magnetizing ampere-turns from two signals are equal and opposite.

Direct-current comparators are available in which dc flux balance is maintained and any unbalance is used to servo a second, or slave, current signal. For the highest accuracy, second-harmonic modulators are used, and for lower precision, Hall effect sensors. Electronically balanced ac and dc current comparators make clip-around ammeters possible, in which an openable magnetic core can be closed around a current-carrying conductor. This allows the meter to be connected into and removed from the circuit without breaking it or interrupting the current. *See* CURRENT COMPARATOR; INSTRUMENT TRANSFORMER.

Very small currents. The obvious method for measuring a very small current is to determine the voltage drop across a large resistor. A sensitive voltage detector having very low offset current is required, for example, an electrometer. Electrometers based on MOSFET (metal-oxide-semiconductor field-effect transistor) devices have overtaken other designs where the very highest resolution is required, as they can have offset current drifts less than 10^{-16} A. In order to provide a low impedance to the measured current, it is preferable to use the electrometer device in an operational-amplifier configuration (see **illus.**). The input becomes a virtual ground, and so stray capacitance across the input connection does not degrade the rate of response of the circuit as seriously as in the simple connection. *See* TRANSISTOR.

The operational-amplifier configuration can be extended to provide a variety of ranges from each high-value resistor. Driving the resistor from the amplifier output through a voltage divider of ratio N gives a similar result to the use of a resistor N times higher in value. This also has the advantage that extremely high-value resistors can be avoided.

Operational amplifiers suitable for this application are readily available, based not only on field-effect transistors but also on varactors. The latter are semiconductor equivalents to the mechanical vibrating capacitor electrometers. *See* AMPLIFIER;

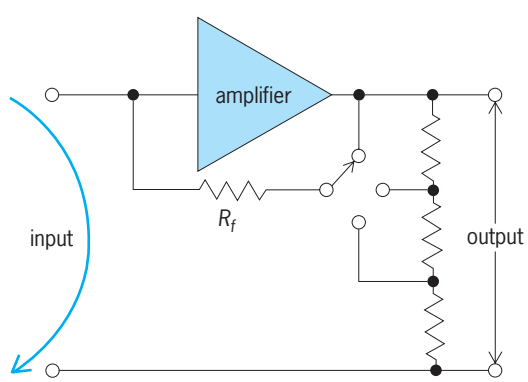

Measurement of very small currents by using an operational amplifier.

ELECTRICAL MEASUREMENTS; ELECTROMETER; OPERATIONAL AMPLIFIER; VARACTOR; VOLTAGE MEASUREMENT. R. B. D. Knight

Bibliography. D. A. Bell, *Electronic Instrumentation and Measurements*, 2d ed., 1997; S. Geczy, *Basic Electrical Measurements*, 1984; B. P. Kibble et al., *A Guide to Measuring Direct and Alternating Current and Voltage below 1 MHz*, Institute of Measurement and Control/National Physical Laboratory, 2003; M. A. Laughton and D. F. Warne (eds.), *Electrical Engineer's Reference Book*, 16th ed., 2003; L. Schnell (ed.), *Technology of Electrical Measurements*, 1993.

Current sources and mirrors

A current source is an electronic circuit that generates a constant direct current which flows into or out of a high-impedance output node. A current mirror is an electronic circuit that generates a current which flows into or out of a high-impedance output node, which is a scaled replica of an input current, flowing into or out of a different node.

Basic circuit configurations. Most specifications of analog integrated circuits depend almost uniquely on the technological parameters of the devices, and on the direct or alternating current that flows through them. The voltage drop over the devices has much less impact on performance, as long as it keeps the devices in the appropriate mode of operation (linear or saturation). High-performance analog integrated-circuit signal processing requires that currents be generated and replicated (mirrored) in an accurate way, independent of supply voltage and of those device parameters that are least predictable (for example, current gain β in a bipolar transistor). Hence, current sources and mirrors occupy a large portion of the total die area of any analog integrated circuit. They are also used, but less often, in discrete analog circuits.

Two basic circuit configurations are shown in **Fig. 1**. The circuit in Fig. 1*a* is suitable for implementation in bipolar integrated-circuit technology, or in complementary metal-oxide-semiconductor (CMOS) integrated-circuit technology if Q_1 and Q_2 are replaced by enhancement metal-oxide-semiconductor field-effect transistors (MOSFETs). Formulas will be derived here only for the bipolar configuration, but can be extended easily to MOS circuit configurations. The collector-base short circuit guarantees that transistor Q_1 will be always in the linear range. If it is assumed, for the moment, that the current gain β and the output impedance r_o of the bipolar transistors are infinite, then the currents are determined by Eqs. (1)–(3). Here, I_{C1} is the collector current of

$$I_{C1} = I_{S1} \cdot \left(\exp \frac{V_{BE1}}{U_T} - 1 \right) = \frac{V_{DD} - V_{SS} - V_{BE1}}{R} \tag{1}$$

$$I_{\text{out}} = I_{S2} \cdot \left(\exp \frac{V_{BE2}}{U_T} - 1 \right) \tag{2}$$

$$V_{BE1} = V_{BE2} = 0.7 \text{ volt} \tag{3}$$

Q_1; V_{BE1} and V_{BE2} are the base-emitter voltages of Q_1 and Q_2; V_{CC} and V_{EE} are the collector and emitter supply voltages; I_{S1} and I_{S2} are the saturation currents of Q_1 and Q_2, which are proportional to the emitter areas A_{E1} and A_{E2}; and U_T is the thermal voltage kT/q, where k is Boltzmann's constant, T is the absolute temperature, and q is the electron charge. (At room temperature, $U_T \simeq 16$ mV.) For two transistors, side by side on the same integrated circuit, the ratio of the two saturation currents I_{S1}/I_{S2} can be realized with an error of less than 1%; Eqs. (1)–(3) can be rewritten as Eq. (4). This equation implies that

$$I_{\text{out}} = \frac{A_{E2}}{A_{E1}} \cdot \frac{V_{DD} - V_{SS} - 0.7 \text{ V}}{R} = \frac{A_{E2}/I_{\text{ref}}}{A_{E1}} \tag{4}$$

the output current will be independent of the technological parameters of the bipolar transistors but dependent on the value of resistor R and on the supply voltages. Even if resistor R or the supply voltages are varied, I_{C1} and I_{C2} always remain proportional: I_{C1} is mirrored in I_{C2}. Equations (1)–(3) demonstrate an important principle in analog integrated-circuit design: replica biasing, in which I_{C1} is forced to produce a base-emitter voltage V_{BE} on transistor Q_1. This voltage creates exactly the right current in Q_2, under the assumption that both halves of the current mirror are exact replicas. The circuit of Fig. 1*a* thus acts as a current source characterized by Eq. (4) if the resistor R is present, and as a current mirror also characterized by Eq. (4) if the resistor is replaced by an input current source of value I_{ref}.

A simple and elegant one-transistor current source can be realized with a junction field-effect transistor (JFET) or depletion MOSFET (Fig. 1*b*). Indeed, the source-to-gate connection guarantees a constant and predictable transistor biasing, since the threshold voltage V_T is negative, as in Eq. (5). The output

$$V_{GS} - V_T = -V_T > 0 \tag{5}$$

current realized in this way is often denoted by I_{DSS}.

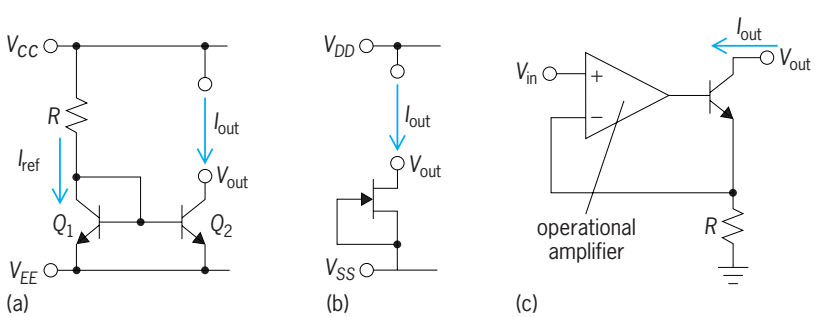

Fig. 1. Basic current-source circuit configurations. (*a*) Circuit suitable for implementation in bipolar or enhancement metal-oxide-semiconductor (MOS) integrated-circuit technology, for example, complementary MOS (CMOS) technology. (*b*) Circuit suitable for implementation in depletion field-effect transistor (FET) integrated-circuit technology, for example, negative MOS (NMOS), bipolar-junction FET (BiFET), or gallium arsenide metal-semiconductor FET (MESFET) technologies. (*c*) More general active circuit (both integrated-circuit and discrete applications).

Finally, a more general circuit configuration (Fig. 1*c*) employs an operational amplifier to create a constant current given by Eq. (6). This circuit

$$I_{\text{out}} = V_{\text{in}}/R \tag{6}$$

has been successfully applied in both integrated and discrete circuits to create constant or voltage-dependent current sources.

Variations. Although the circuit in Fig. 1*a* is the most common circuit configuration, with applications in bipolar and CMOS technology [with bipolars replaced by negative MOS (NMOS) transistors], it has some serious drawbacks that are addressed in the improved circuits in **Fig. 2**.

The output current I_{out} in the circuit of Fig. 1 is affected by the finite current gain β in the bipolar transistors as given by Eq. (7). The circuit in Fig. 2*a*

$$\frac{\Delta I_{\text{out}}}{I_{\text{out}}} = \frac{1}{1 + (2/\beta)} \tag{7}$$

has a considerably lower error due to finite β.

Current mirror ratios I_{S2}/I_{S1} different from unity can be realized by transistor pairs with unequal emitter areas; however, manufacturing tolerances can now lead to significant systematic errors. Integer mirror ratios can be realized without accuracy degradation, if arrays of identical unit transistors are used (3 to 1 in Fig. 2*b*).

The basic current source or current mirror of Fig. 1 has an output impedance equal to the dynamic output resistance r_o of the transistors. Cascodes (Fig. 2*c*) enhance the output resistance of the current source

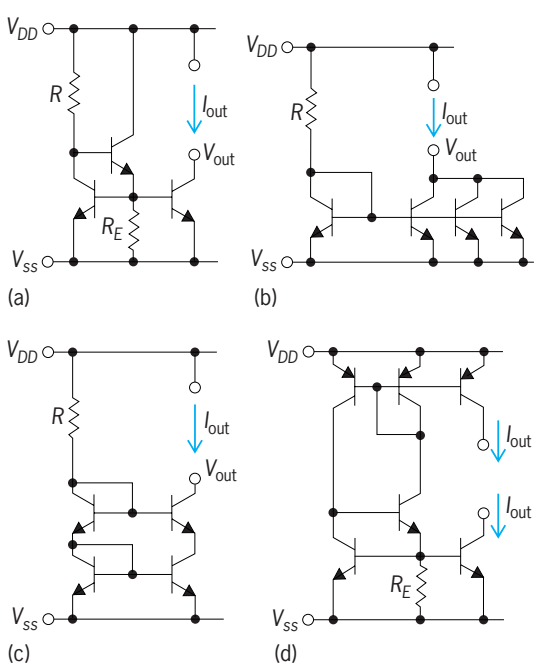

Fig. 2. Improved current-source configurations. (*a*) Circuit for reduction of effect of finite β. (*b*) Replica biasing for integer current mirror ratio. (*c*) Cascoded current mirror for enhanced output impedance. (*d*) Power-supply-independent current source.

according to Eq. (8), where G_m is the transconduc-

$$R_{\text{out,source}} = r_o \cdot (G_m \cdot r_o) \tag{8}$$

tance of the cascode transistor. The increase in R_{out} is usually by a factor of about 100.

It is often important that analog circuit specifications be independent of the applied supply voltages. Figure 2*d* shows a simple circuit in which the current is independent of the supply voltages. The realized current I_{out} can be found by solving transcendental equation (9).

$$I_{\text{out}} = V_T \cdot \frac{\log_e(I_{\text{out}}/I_S)}{R_E} \tag{9}$$

This circuit has two operating points: I_{out} given by Eq. (9) or I_{out} equal to zero. To avoid having the circuit lock into the currentless mode during power-up, a startup circuit must be added. *See* INTEGRATED CIRCUITS; TRANSISTOR.

Peter M. VanPeteghem

Bibliography. P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3d ed., 1992; R. Gregorian and G. C. Temes, *Analog MOS Integrated Circuits for Signal Processing*, 1986.

Curve fitting

A procedure in which the basic problem is to pass a curve through a set of points, representing experimental data, in such a way that the curve shows as well as possible the relationship between the two quantities plotted. It is always possible to pass some smooth curve through all the points plotted, but since there is assumed to be some experimental error present, such a procedure would ordinarily not be desirable. *See* INTERPOLATION.

The first task in curve fitting is to decide how many degrees of freedom (number of unspecified parameters, or independent variables) should be allowed in fitting the points and what the general nature of the curve should be. Since there is no known way of answering these questions in a completely objective way, curve fitting remains something of an art. It is clear, however, that one must make good use of any background knowledge of the quantities plotted if the above two questions are to be answered. Therefore, if it was known that a discontinuity might occur at some value of the abscissa, one would try to fit the points above and below that value by separate curves.

Against this background knowledge of what the curve should be expected to look like, one may observe the way the points fall on the paper. It may even be advantageous to make a few rough attempts to draw a reasonable curve "through" the points.

A knowledge of the accuracy of the data is needed to help answer the question of the number of degrees of freedom to be permitted in fitting the data. If the data are very accurate, one may use as many degrees of freedom as there are points. The curve can then be made to pass through all the points, and

it serves only the function of interpolation. At the opposite extreme when the data are very rough, one may attempt to fit the data by a straight line representing a linear relation, Eq. (1), between y and x.

$$y = ax + b \tag{1}$$

Using the above information and one's knowledge of the functions that have been found useful in fitting various types of experimental curves, one selects a suitable function and tries to determine the parameters left unspecified. At this point there are certain techniques that have been worked out to choose the optimum value of the parameters.

One of the most general methods used for this purpose is the method of least squares. In this method one chooses the parameters in such a way as to minimize the sum, Eq. (2), where y_i is the ordinate of ith

$$S = \sum_{i=1}^{n} [y_i - f(x_i)]^2 \tag{2}$$

point and $f(x_i)$ the ordinate of the point on the curve having the same abscissa x_i as this point.

Method of least squares. If one attempts to fit the data with a straight line, one lets y equal the values shown in Eq. (3), and Eq. (2) becomes (4).

$$y = f(x) = ax + b \tag{3}$$

$$S = \sum_{i=1}^{n} [y_i - ax_i - b]^2 \tag{4}$$

Since S is to be a minimum,

$$\frac{\partial S}{\partial a} = \frac{\partial S}{\partial b} = 0$$

and therefore Eqs. (5) hold.

$$\begin{aligned} a \sum_i x_i^2 + b \sum_i x_i &= \sum_i x_i y_i \\ a \sum_i x_i + bn &= \sum_i y_i \end{aligned} \tag{5}$$

The solution of these equations is shown as Eqs. (6), where the term Δ is defined in Eq. (7).

$$\begin{aligned} a &= \frac{1}{\Delta}\left(\sum_i x_i \sum_i y_i - n \sum_i x_i y_i\right) \\ b &= \frac{1}{\Delta}\left(\sum_i x_i \sum_i x_i y_i - \sum_i x_i^2 \sum_i y_i\right) \end{aligned} \tag{6}$$

$$\Delta = \left(\sum_i x_i\right)^2 - n \sum_i x_i^2 \tag{7}$$

These values of a and b are those that minimize S and therefore represent the proper choice of the parameters of a straight line according to the criterion of least squares. *See* LEAST-SQUARES METHOD.

Method of averages. This method has less theoretical justification, but it is easier to apply than the method of least squares.

If all the points (x_i, y_i for $i = 1, 2, \ldots, n$) representing the data were to lie on a straight line, one could require that a and b satisfy all n of Eqs. (8). Any two

$$y_i = ax_i + b \qquad i = 1, 2, \ldots, n \tag{8}$$

of these equations could then be used to determine a and b. One could also add these equations together in two groups to obtain Eqs. (9).

$$\begin{aligned} \sum_{i=1}^{k} y_i &= a \sum_{i=1}^{k} x_i + bk \\ \sum_{i=k+1}^{n} y_i &= a \sum_{i=k+1}^{n} x_i + b(n-k) \end{aligned} \tag{9}$$

Suppose now that the points do not lie on a straight line; then Eqs. (8) are no longer valid. One may, however, assume that a best line through the points is given by Eq. (3) with a and b determined by requiring that Eqs. (9) be satisfied. There is some arbitrariness in the choice of k, but it is usually taken to be a whole number near $n/2$.

Consider the data below:

x	y
0.5	0.4
1.0	1.3
1.5	1.6
2.0	2.3
2.5	3.2
3.0	3.6
3.5	4.2
4.0	5.2

Letting $k = 4$, one obtains the following values for Eqs. (9):

$$5a + 4b = 5.6$$
$$13a + 4b = 16.2$$

Subtracting, one has $8a = 10.6$ or $a = 1.325$, and therefore $b = -0.256$; thus, the best straight line is given by the equation

$$y = 1.325x - 0.256$$

Graphical method. In this method one first plots the points (x_i,y_i) on coordinate paper using suitable scales along the X and Y axes (see **illus.**). One then takes a transparent straightedge and adjusts it so the points lie as near the edge as possible with about the same numbers above as there are below. The straight line corresponding to the position of the edge is drawn in, and the coordinates of two points (x',y') and (x'',y'') near each end of the line are read off. The straight line drawn is then given by Eq. (3), with a and b determined by the simultaneous equations (10).

$$\begin{aligned} y' &= ax' + b \\ y'' &= ax'' + b \end{aligned} \tag{10}$$

The coordinates of points P' and P'' in the

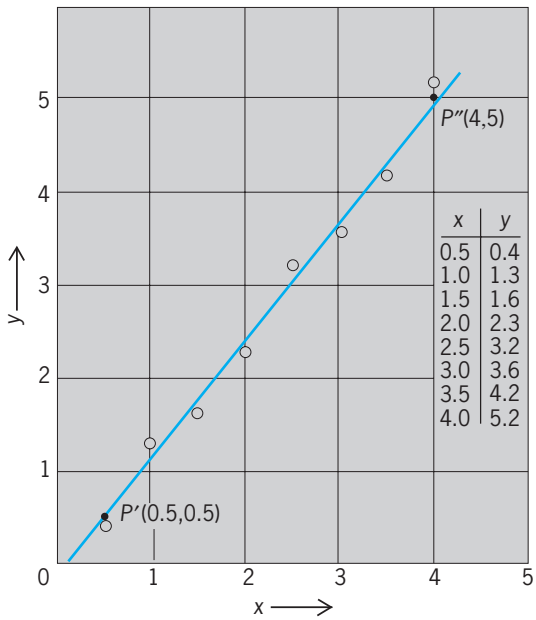

Graphical method of curve fitting.

illustration are, respectively, (0.5,0.5) and (4,5); therefore, $\alpha = 1.286$ and $b = -0.144$.

Use of nonlinear scales. Suppose one tries to fit the data points (x_i,y_i) with any general functional relation $y = f(x,\alpha,\beta)$ having two undetermined parameters α and β. It often happens that it is possible to introduce two new variables, Eqs. (11), that will satisfy linear Eq. (12). From the given data points (x_i,y_i)

$$X = F(x) \qquad Y = G(y) \tag{11}$$

$$Y = aX + b \tag{12}$$

one can obtain a new set of data points (X_i,Y_i) by the use of Eqs. (11). One may then fit these points by a straight line in the XY plane, given by Eq. (12), using any of the techniques described above. The required equation in x and y is then $G(y) = aF(x) + b$, which will determine the original constants α and β.

For example, suppose one assumes that Eq. (13)

$$y = \beta e^{\alpha x} \tag{13}$$

holds. Then, by taking the natural logarithm of both sides, one has Eq. (14). Thus a linear equation as

$$\log y = ax + \log \beta \tag{14}$$

given in Eq. (12) exists, as shown in Eqs. (15) and (16). A graphical solution involves use of semilog

$$X = x \qquad Y = \log y \tag{15}$$

$$a = \alpha \qquad b = \log \beta \tag{16}$$

paper in which the vertical displacement of points is proportional to $\log y$. On such paper, therefore, Eq. (14), or its equivalent, Eq. (13), plots as a straight line.

Consider next fitting data by Eq. (17), which may

$$y = \beta x^{\alpha} \tag{17}$$

also be written $\log y = \alpha \log x + \log \beta$. Thus again the linear equation labeled (12) connects, as shown in Eqs. (18), and Eqs. (16) again hold for the constants.

$$X = \log x \qquad Y = \log y \tag{18}$$

Logarithmic paper makes it simple to give the data points the rectangular coordinate values X and Y. The best line through these points then determines a and b and, therefore, also α and β.

If the plotting on logarithmic paper seems to lie about a curved line, one may try an equation of the form of Eq. (19), which may be written as Eq. (20).

$$y = \beta x^{\alpha} + \gamma \tag{19}$$

$$\log(y - \gamma) = \alpha \log x + \log \beta \tag{20}$$

One therefore works with a linear relation between $\log x$ and $\log (y - \gamma)$.

To obtain γ one draws a smooth curve through the points (x_i,y_i) and picks out two points (x',y') and (x'',y'') near the ends of this curve. Next one chooses $x''' = {}^1\!/_2(x'x'')$ and finds the point on the curve (x''',y''') with this value of x. It can be shown quite simply that Eq. (21) holds. The values of a and

$$\gamma = \frac{y'y'' - y'''}{y' + y'' - 2y'''} \tag{21}$$

b and therefore of α and β are found as before.

If one decides to fit the data by a parabolic curve represented by Eq. (22), one may first draw a smooth

$$y = \alpha + \beta x + \gamma x^2 \tag{22}$$

curve through the points and select some point $(x'y')$ on this curve. The parabola will pass through this point if Eq. (23) holds. From Eqs. (22) and (23)

$$y' = \alpha + \beta x' + \gamma x'^2 \tag{23}$$

one has Eq. (24), which shows that $(y - y')/(x - x')$

$$\frac{y - y'}{x - x'} = \beta + \gamma(x + x') \tag{24}$$

is a linear function of x. Therefore, for each data point (x_i,y_i) one may plot a new point (X_i,Y_i), as shown in Eqs. (25). The best straight line through these points

$$X_i = x_i \qquad Y_i = \frac{y_i - y'}{x_i - x'} \tag{25}$$

given by Eq. (12) can be determined as before and this determines β and γ in Eq. (24). The value of α is then determined by use of Eq. (23). *See* COORDINATE SYSTEMS; EXTRAPOLATION; GRAPHIC METHODS.

Kaiser S. Kunz

Bibliography. S. Arlinghaus et al., *Handbook of Practical Curve Fitting*, 1994; P. Lancaster and K. Salkauskas, *Curve and Surface Fitting*, 1986; J.-G. Reich, *Curve Fitting and Modeling for Scientists and Engineers*, 1992.

Cutaneous sensation

The sensory quality of skin. The skin consists of two main layers: the epidermis, which is the outermost protective layer, and the dermis, which consists of a superficial layer called the papillary dermis and a deeper layer called the reticular dermis (**Fig. 1**). Beneath the dermis is a layer of loose connective tissue, the subcutaneous tissue, which attaches the skin to underlying structures. *See* SKIN.

Sensory Receptors

Sensory receptors in or beneath the skin are peripheral nerve-fiber endings that are differentially sensitive to one or more forms of energy. The sensory endings can be loosely categorized into three morphological groups (**Fig. 2**): endings with expanded tips, such as Merkel's disks found at the base of the epidermis; encapsulated endings, such as Meissner's corpuscles (particularly plentiful in the dermal papillae), and other organs located in the dermis or subcutaneous tissue, such as Ruffini endings, Pacinian corpuscles, Golgi-Mazzoni corpuscles, and Krause's end bulb; and bare nerve endings that are found in all layers of the skin (some of these nerve endings are found near or around the base of hair follicles).

Available evidence indicates that the encapsulated and expanded tip endings and the hair follicle endings are mechanoreceptors and that the specialized end organs serve to influence the dynamic range of mechanical stimuli to which the receptor responds, but not the type of energy to which it is most sensitive (for example, mechanical as opposed to thermal). The response specificity of a receptor is determined by some as yet unidentified property of the membrane of the nerve ending. Thus, receptors having bare nerve endings still differ in the specificities of their responses to certain stimuli.

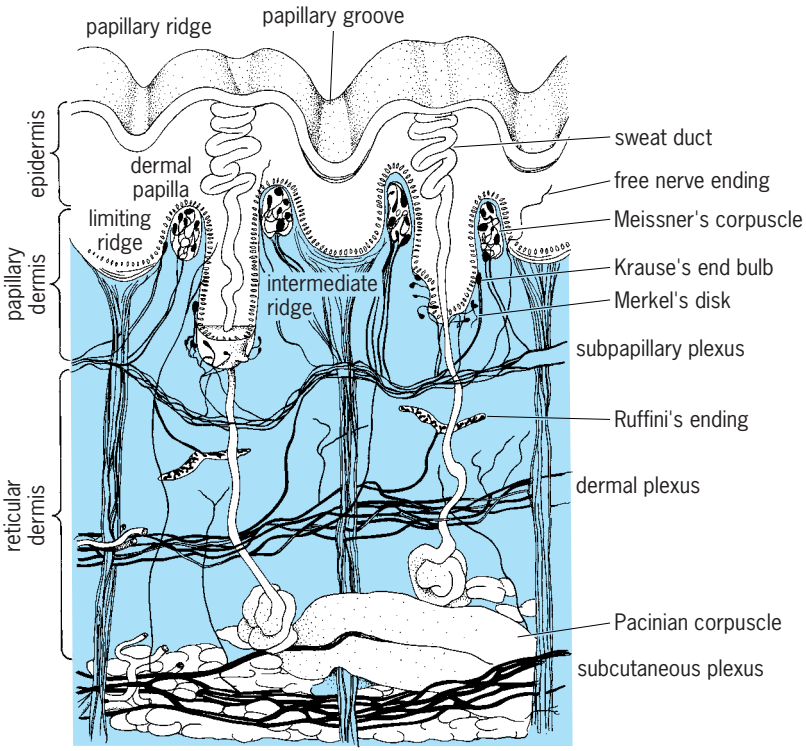

Fig. 1. A section of glabrous (nonhairy) human skin cut vertically through the papillary ridges to show locations of sensory receptors. (*After V. B. Mountcastle, ed., Medical Physiology, vol. 1, 14th ed., C. V. Mosby, 1980*)

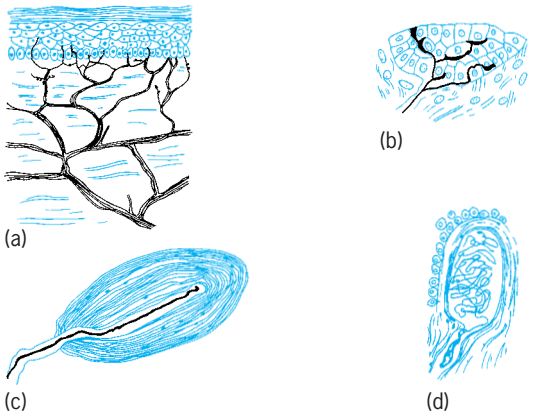

Fig. 2. Types of receptor endings in the skin: (*a*) free nerve endings; (*b*) Merkel's disk; (*c*) Pacinian corpuscle; (*d*) Meissner's encapsulated ending. (*After A. Brodal, Neurological Anatomy in Relation to Chemical Medicine, Oxford University Press, 1969*)

Specificities receptors. There is a remarkable relationship between the response specificities of cutaneous receptors and five primary qualities of cutaneous sensation, the latter commonly described as touch-pressure, cold, warmth, pain, and itch. Each quality is served by a specific set of cutaneous peripheral nerve fibers. More complex sensations must result from an integration within the central nervous system of information from these sets of nerve fibers. For example, tickle sensations may result for the temporal and spatial sequence of stimulating a succession of touch receptors with a slowly moving stimulus lightly applied to the skin. Exploration of the skin surface with a rounded metal point reveals that there exist local sensory spots on the skin, stimulation of which (by any type of energy and by any temporal pattern of stimulation) evokes only one of the five qualities of sensation. Thus may be plotted maps of pressure, warm, cold, pain, or itch spots. *See* PAIN.

There is also a correlation between different qualities of sensation and the sizes (diameters) of the sensory peripheral nerve fibers serving them. The larger-diameter nerve fibers of 6–12 micrometers serve the sense of touch-pressure (and such temporal variations as vibratory sensations or tickle sensations). The sense of coolness and the sense of pricking pain are served by different sets of small myelinated fibers (1–5 μm diameter); sensations of warmth, burning or aching pain, or itch are each mediated by different sets of unmyelinated nerve fibers (0.2–1.5 μm diameter). Since the conduction velocity of nerve impulses along a peripheral nerve is inversely related to the fiber's diameter, there may be a temporal sequence of sensations evoked by an abruptly delivered stimulus. For example, stubbing the toe or pricking

the skin of the foot or hand with a pin may elicit first a sense of touch (mediated by the faster-conducting myelinated fibers), followed almost immediately by a sharp pain of short duration (smaller myelinated fibers) and, after as long as a second or so, a more agonizing pain of longer duration (mediated by the slowly conducting unmyelinated fibers).

Local anesthetic infiltrated around a peripheral nerve first blocks conduction in unmyelinated fibers (with a corresponding loss of the capacity to sense warmth, burning pain, and itch), followed by a block of myelinated fibers in ascending order of their diameters with concomitant losses of the senses of coolness, pricking pain, tickle, and touch-pressure. The exact opposite sequence of sensation loss occurs when a nerve is blocked by asphyxia produced by a pressure cuff around the arm, with the larger-diameter fibers blocked first and the unmyelinated fibers last.

The best evidence for specificity of sensory peripheral nerve fibers and their receptor endings is provided by experiments in humans and animals in which a recording electrode is brought in contact with a portion of the nerve (such as following microdissection of the nerve or by inserting the electrode through the skin and into the nerve), and the electrical signs of impulse discharges of single nerve fibers are recorded in response to cutaneous stimulation. Evidence is that each of the elementary qualities of sensation is served by a different set of peripheral nerve fibers.

Mechanoreceptors. The sense of touch-pressure is mediated by large, myelinated fibers. One type, believed to terminate in Merkel's disks, responds with increasing frequency of discharge to increases in pressure applied to the skin. The frequency of discharge is linearly related to the amount of skin indentation produced by a punctate stimulus and linearly related to numerical estimates, made by human observers, of the perceived magnitude of pressure sensations. Furthermore, the capacity of this mechanoreceptor to respond differentially to gaps of different sizes in a smooth surface applied to the passive skin can account for the human capacity for spatial resolution. The perceptual threshold (minimal separation) for spatially resolving gaps applied to the passive fingertip is a gap of 0.8 mm. Spatial resolution is very good on the fingertip, where the density of mechanoreceptors is high, and relatively poor on the back, where receptor density is sparse.

Other types of mechanoreceptors with large diameter myelinated fibers respond to mechanical movement of the skin and can be considered as velocity detectors. One type ends in Meissner corpuscles and responds best (that is, requires the least amplitude of skin displacement) to mechanical oscillations of 30 Hz. Another type ends in the Pacinian corpuscle and is even more sensitive, but only at higher frequencies (best frequency is about 225 Hz). The combined response properties of receptors with Meissner's or Pacinian corpuscles provide the necessary information required to account, at the level of the receptor, for the capacities of humans to detect and discriminate between mechanical oscillatory stimuli of different frequencies and amplitudes applied to the skin. Such receptors probably have an important role in conveying information about surface texture when the finger is moved back and forth over the surface of an object. *See* MECHANORECEPTORS.

Thermoreceptors. There are two types of thermoreceptors responsive to temperature changes in human skin. Warmth is served by a set of unmyelinated nerve fibers with bare nerve endings. These fibers have an ongoing discharge (in warm skin) that is eliminated by gentle cooling. Their frequencies of discharge increase nearly linearly with increases in stimulus intensity over the range of temperatures described as increasingly warm by human observers. The sense of coolness is served by a set of thinly myelinated nerve fibers with bare nerve endings. These have an ongoing discharge (in cool skin) that is eliminated by gentle warming. They increase their rates of discharge nearly linearly with decreases in temperatures described as increasingly cool by human observers. Neither type of thermoreceptor responds to mechanical stimulation. The response properties of these types of receptors have been shown to provide the information required to account for human capacities to discriminate small differences in the temperatures of stimuli applied to the skin.

Nociceptors. These receptors have bare nerve endings in or beneath the skin and are selectively responsive to noxious stimuli described as painful when applied to human skin. There are different sets of thinly myelinated nerve fibers and unmyelinated fibers whose endings are differentially responsive to only noxious thermal stimuli (hot or cold, or both) or only to noxious mechanical stimuli (such as pricking or pinching the skin). Other types of nociceptors are less specific, and the most common of these is one which responds to any noxious stimulus, whether it be chemical, thermal, or mechanical, and thus is called the polymodal nociceptor. It typically has an unmyelinated nerve fiber, but those with thinly myelinated fibers have been found as well. The minimal intensities of heat required to evoke a response in the typical polymodal nociceptor with an unmyelinated fiber has been shown to elicit minimal sensations of pain in humans. Furthermore, the frequencies of discharge in these receptors in response to heat stimuli increase as a slightly positively accelerating function of increasing temperature over the range of approximately 109–124°F (43–51°C) [higher intensities are not commonly tested], as do numerical estimates by humans of the magnitude of pain evoked by such stimuli. Following injury to the skin, such as that produced by a superficial burn, the threshold responses to mechanical or thermal stimuli of many nociceptors may be greatly lowered. This is believed to account, at least in part, for the increased painfulness of stimuli applied to injured skin. In such a state of heightened sensitivity to pain, termed hyperalgesia, normally innocuous stimulation such as gentle rubbing or warming can evoke a sensation of pain.

Itch. Itch sensations are believed to be mediated by a set of unmyelinated nerve fibers different from those serving pain. Itch can be evoked by mechanical, electrical, and chemical stimuli. Certain chemical stimuli are the most effective in producing itch, and these may be externally applied to the skin or released within the skin under certain conditions. For example, enzymes called proteinases may be found in certain plants and, when delivered to the skin via spicules, are powerful itch-producing agents. A naturally occurring substance is histamine, which is released from mast cells in the skin following local injury or in response to foreign substances that trigger an allergic reaction. Some chemicals produce itch by causing histamine to be released in the skin, while others act independently to evoke itch and can do so in regions of skin previously depleted of histamine.

Robert Lamotte

Touch Sensations

The hands and the oral cavity, including the lips, are the body structures primarily used to tactilely explore small objects in the environment. Touch sensations can easily be aroused from other areas of the body, for example, the stimulation of body hairs, but the resolution of fine detail by these areas is not as great as that of the hands or the oral cavity.

Stimulus characteristics. A mechanical event applied to the skin may be divided into static and dynamic components. During indentation, mechanical energy is expended to move the skin from one position to another. The quantity of energy can be expressed in terms of the applied force per unit dimension, for example, tension (g/mm) or pressure (g/unit area), or in terms of work (force × indentation depth). Work, compared to pressure or tension, has proved to be a better stimulus measurement to account for variations in touch sensations that are the result of variations in the mechanical properties of the skin. The depth of skin indentation produced by a given force is related to the elastic properties of the skin, which are, in turn, related to age, site of stimulation, and the underlying tissue. If the force is made large and indentation depth controlled, much of the confounding that results from the varying mechanical properties of the skin can be avoided. In this way indentation depth becomes a convenient measurement by which to express the intensities of cutaneous mechanical stimuli.

The velocity, or rate, of indentation has also been shown to be related to threshold tactile sensations. The rate of indentation has little effect on the absolute tactile threshold when the rate of indentation is greater than about 0.3 mm per second. At slower rates of indentation, however, the depth of indentation required to produce a just-detectable tactile sensation increases dramatically. Rate of indentation also has a small but consistent effect on the judged intensity of tactile sensations. Acceleration and deceleration in indentation rate have so far not been implicated as stimulus variables for tactile sensations.

Absolute sensitivity. Sensitivity to tactile stimuli varies from one part of the body to another. The nose, lips, cheek, and forehead are the most sensitive parts, followed closely by the fingertips and palms, but even here there are large differences. Single tactile points on the palmar aspects of the fingers are highly sensitive with a median indentation depth threshold of 0.011 mm, whereas the tactile points in the center of the palm and the lateral aspects of the fingers have a median indentation depth threshold of 0.036 mm.

Intensity of sensations. As the intensity of a touch stimulus (for example, depth of indentation) increases, the intensity of the resulting tactile sensation increases. However, the intensity of the tactile sensation appears to increase in two stages, at least when rates of indentation of 0.4 mm per second or greater are used. Tactile sensations have been judged to be just about the detection level for indentation depths of 0.005 mm and rates of 1 or 10 mm per second. As indentation depth increases, the resulting sensation is judged to increase slowly, but when indentation depths greater than about 0.5 mm are used, the sensation increases considerably more with each stimulus intensity increment. This two-stage response probably results from the cooperation of two different receptor mechanisms to produce the touch sensations.

Spatial discriminations. There are two traditional means of assessing the passive, tactile, spatial, discriminatory capabilites of the skin. These are point localization, the ability to locate the point on the skin that has been touched, and the two-point threshold, the distance separating two pointed stimulators to experience two rather than one point of stimulation. Measurements of the accuracy of point localization show the index finger to have the smallest error (1.4 mm) while the back shows the largest error (12.5 mm). The two-point threshold has been found to be smallest on the middle-finger pad (2.5 mm, the approximate separation of braille dots). The next smallest two-point threshold was on the index-finger pad (3.0 mm), and largest on the calf of the leg (47.0 mm). This means that the finger pads can detect much smaller (finer) features of an object than can other skin areas.

Temporal discriminations. When rapidly vibrating mechanical stimuli are applied to the skin, another stimulus dimension is added—the frequency of repetition. When applied to the index-finger pad, the resulting sensation, for repetition frequencies up to about 40 Hz, may best be described as a flutter. At higher frequencies the flutter changes to a buzzing sensation.

The lower limit to which the skin is sensitive is defined as that frequency at which the individual pulses fuse to form a single continuous sensation. This occurs, as it does in audition, at about 15 Hz. The upper limit is defined as that frequency at which the buzzing sensation disappears and the sensation becomes that of steady skin indentation, about 640 Hz.

Thresholds of vibratory sensations vary with the frequency of the vibration. There is little change in sensitivity for frequencies up to about 40 Hz. As the

frequency of vibration increases beyond 40 Hz, the threshold decreases; that is, the receptors are more sensitive to the higher frequencies, until they reach their peak sensitivity at about 250 Hz, beyond which their sensitivity decreases. Sensitivity to increased frequency once again shows two stages. Two separate receptor processes, each with different rules of operation, combine to form the range of response.

For many years scientists and engineers have sought ways in which the skin might be used as a substitute means of communication for the blind or deaf. There are difficulties, such as cost, ease of learning, and the cosmetic appearance, in the development of a suitable device. In addition, the skin presents its own difficulties in terms of the amount of information that it can convey to the brain. Three information channels are possible: body site of stimulation, frequency, and intensity of the vibratory stimuli. The reliability of point localization is poor except on the fingertips and face, and varies widely with body site. Frequency requires discrimination by the skin—vibrations of 40 and 44 or 200 and 220 Hz applied to the finger pad can be discriminated about 75% of the time, but like point localization, this varies with the site of stimulation. The range of useful frequencies is also severely limited—up to about 600 Hz. Discrimination between intensities of vibration not only varies with the site of stimulation but also varies over a wide range with the frequency of the vibration.

Temperospatial sensations. When a light tactile stimulus, such as the tip of a feather, is drawn across the skin so as to stimulate successive touch points in temporal succession, a tickle sensation is often experienced. This is a different sensation than that of itch, although the response elicited may be the same—rubbing or scratching. Little is known of the necessary and sufficient stimulus conditions to arouse a tickle.

Active touch. The limits of fine discrimination (two-point discrimination) discussed above involved the finger pads resting passively on the feature to be detected. Much finer features (texture) can be detected, however, if the finger pads are actively moved over the surface. Scanning an object with the finger increases the information flow in a tactile sensation. It is not the movement of the finger per se over the object, for just as fine a discrimination can be made if the object is moved under the stationary finger. Perhaps small vibrations, shearing forces on the skin, the temporal pattern of tactile receptor stimulation, or some similar dimensions have been added by the movement.

Another, more complex aspect of active touch is object recognition, or stereognosis. Here, not only are the fine features of the object distinguished, but so is the object's overall form. Kinesthesis, the muscle and joint sense, is used.

Thermal Sensations

The thermal sense is unique among the senses in that an absence of sensation occurs at about 86–97°F (30–36°C). This represents a considerable amount of thermal energy, but this is the temperature domain in which homeothermic systems operate. Changes in temperature imposed from outside change the thermal environment of the (warm and cold) receptors, presumably producing changes in their metabolism and hence their response rates, without an exchange of energy between the stimulus temperature (change) and the receptor. In other sensory systems the absence of the appropriate energy form leads to the absence of sensation. In these systems stimulation is the result of an exchange of energy between the stimulus and the receptor.

The thermal sense is unique in another way. Its primary function—an important input to the body temperature-regulating system—does not reach awareness. Only relatively innocuous sensations of warm and cool are perceived, and only when the skin temperature drops below about 64°F or 18°C (cold-pain) or rises above 113–115°F or 45–46°C (heat-pain) does the sensation become noxious. This represents engagement of a different system—the nocifensor system—and is not properly a part of the thermal sensing system.

Stimulus characteristics. There are six principal stimulus variables that alter the responsiveness of the human temperature-sensing system: (1) the temperature to which the skin has been adapted; (2) the direction of the temperature change; (3) the intensity of the temperature change; (4) the rate of the temperature change; (5) the area of the skin stimulated; and (6) the site of application of the thermal stimuli.

The band of skin temperatures between about 86 and 97°F (30 and 36°C) is called the zone of physiological zero (**Fig. 3**). When skin temperatures do not exceed the limits of this band, given time, thermal sensations adapt completely; that is, they disappear. At lower skin temperatures a cool sensation persists no matter how long that temperature is maintained—the stimulus adapts but the sensation never disappears completely. With further reductions in skin temperature the sensation would be described as cold, and finally at about 64°F (18°C) is replaced by a dull, aching sensation—cold-pain. When the skin temperature is raised above the zone of physiological zero, much the same series of sensations occur except that their qualities are warm to hot, and finally at about 113°F (45°C) the hot is replaced by a stinging sensation—heat-pain.

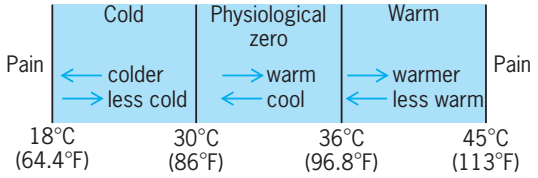

Fig. 3. The relationship between skin temperature, physiological zero, and the direction of temperature change as they affect thermal sensations. (*From D. R. Kenshalo, Sr., Biophysics and psychophysics of feeling, in E. C. Carterette and M. P. Friedman, eds., Handbook of Perception, vol. 6B, chap. 2, Academic Press, 1978*)

The direction of a temperature change, when skin temperature is within the zone of physiological zero, produces thermal sensations consistent with the direction of the change—an increase produces a warm sensation and a decrease produces a cool sensation. A small increase in temperature from a skin temperature below physiological zero produces a decrease in the persisting cool sensation, not a warm sensation. Similarly, a small decrease in temperature from a skin temperature above physiological zero produces a decrease in the persisting warm sensation, not a cool sensation.

The intensity of a temperature change, when skin temperature is within the zone of physiological zero, produces sensations of proportional intensity. When skin temperature is below the zone of physiological zero, sensitivity to decreases in temperature remains unchanged from that in the zone of physiological zero. Sensitivity to temperature increases is reduced, and the farther the skin temperature is below the lower limit of physiological zero, the less sensitive it becomes to temperature increases. A similar circumstance occurs for temperature decreases when the skin temperature is above the upper limit of physiological zero.

When the rate of a temperature change is less than about 0.05°F (0.03°C) per second, sensitivity to that change decreases. At the extreme, skin temperature can be changed from one to the other limit of physiological zero without producing a thermal sensation if done sufficiently slowly, about 0.013°F (0.007°C) per second. Sensitivity to temperature changes does not increase at rates of change faster than 0.05°F (0.03°C) per second, at least up to rates of 2.7°F (1.5°C) per second.

The size of the area of the skin stimulated by temperature changes markedly influences sensitivity. For temperature increases, areal summation is almost complete for areas up to about 30 in.^{2} (200 cm^2). Thus, there is almost complete reciprocity between the area stimulated and the intensity of the stimulus, so that if area is doubled the stimulus intensity required to reach detectability will be halved. This means that the whole palm of the hand is a much more sensitive heat detector than just the fingertip. Areal summation for temperature decreases is somewhat less than that for temperature increases but is still considerable. Such generous summation is not without a price. The thermal sensing system is almost devoid of any ability for spatial localization. Subjects are able to judge only with 80% accuracy whether radiant heat stimuli were delivered to the back or abdomen. This is easy to understand when one considers that the primary function of the cutaneous temperature-sensing system is to inform the central temperature-regulating system of the heat load experienced at the body surface rather than to specify the spatial location of warm and cold objects.

Various parts of the body differ in their sensitivity to temperature increases and decreases. For small temperature increases the forehead is the most sensitive, followed by the chest, forearm, calf, thigh, and abdomen. For small temperature decreases the body trunk is the most sensitive, followed by the limbs and finally the face. With relatively large changes in skin temperature (both increases and decreases) such regional differences in sensitivity are markedly reduced.

Paradoxical thermal sensations. Some cold points, when touched by a hot thermal probe (113°F or 45°C), will yield a cold sensation. The phenomenon of a hot probe giving rise to a cold sensation is called paradoxical cold. Paradoxical cold sensations do not occur, however, when a hot stimulus is applied over a large area. The necessary and sufficient conditions for paradoxical warmth (arousal of a warm sensation by application of a cold probe to a warm point) have not been established.

Heat sensations. The sensation of heat has been held to be uniquely different from either warm or heat-pain sensations. Its threshold ranges from 104 to 115°F (40 to 46°C) with a mean at about 108–109°F (42–43°C).

Two theories dealing with the mechanisms of heat sensations have been advanced. The first maintains that heat sensations are synthesized from the simultaneous stimulation of warm and paradoxically stimulated cold receptors. The second theory maintains that heat sensations arise from their own specifically sensitive receptors. Neither theory has been satisfactorily verified or refuted. *See* NERVOUS SYSTEM (VERTEBRATE); SENSATION; SOMESTHESIS.

Dan R. Kenshalo, Sr.

Bibliography. E. R. Kandel et al. (eds.), *Principles of Neural Science*, 2000; D. Katz and L. E. Kruger (eds.), *The World of Touch*, 1989.

Cyanamide

A term used to refer to the free acid, H_2NCN, or commonly to the calcium salt of this acid, $CaCN_2$, which is properly called calcium cyanamide. Calcium cyanamide is manufactured by the cyanamide process, in which nitrogen gas is passed through finely divided calcium carbide at a temperature of 1000°C (1830°F).

Most plants are equipped to produce the calcium cyanamide from the basic raw materials of air, limestone, and carbon according to reactions (1)–(3).

$$CaCO_3 \xrightarrow{\Delta} CaO + CO_2 \quad (1)$$

$$CaO + 3C \longrightarrow CaC_2 + CO \quad (2)$$

$$CaC_2 + N_2 \rightleftharpoons CaCN_2 + C \quad (3)$$

Most calcium cyanamide is used in agriculture as a fertilizer; some is used as a weed killer, as a pesticide, and as a cotton defoliant.

Cyanamide, H_2NCN, is prepared from calcium cyanamide by treating the salt with acid. It is used in the manufacture of dicyandiamide and thiourea. *See* CARBON; NITROGEN.

E. Eugene Weaver

Cyanate

A compound containing the —OCN group, typically a salt or ester of cyanic acid (HOCN). The cyanate ion is ambidentate, that is, it has two reactive sites, because it can bind through the oxygen (O) or the nitrogen (N). Cyanate is commonly N-bonded with most nonmetallic elements, presumably because of the small charge density on the oxygen. Cyanic acid has the structure H—O═C≡N, but it may exist in an isomeric form known as isocyanic acid, H—N═C═O. The cyanates are isomeric with fulminates, where the carbon and the nitrogen are transposed (—ONC).

Cyanic acid is a volatile liquid, and it polymerizes upon standing to form cyamelide acid and cyanuric acid. In water cyanic acid undergoes hydrolysis to ammonia and carbon dioxide. However, dilute solutions in ice water may be kept for several hours. Both cyanic acid and isocyanic acid are prepared from cyanuric acid. The linear —OCN ion may be prepared by mild oxidation of aqueous cyanide ion (CN^-) using lead oxide (PbO) and potassium cyanide (KCN).

The primary use for cyanates is in the synthesis of a number of organic compounds such as unsubstituted carbamates (R—NH═C═OOR′). These compounds are used as derivatives of alchols. Other important reactions include addition to amines to give substituted ureas. *See* CARBON; CYANIDE; NITROGEN; THIOCYANATE.

Thomas J. Meade

Cyanide

A compound containing the —CN group, for example, potassium cyanide, KCN; calcium cyanide, $Ca(CN)_2$; and hydrocyanic (or prussic) acid, HCN. Chemically, the simple inorganic cyanides resemble chlorides in many ways. Organic compounds containing this group are called nitriles. Acrylonitrile, CH_2CHCN, is an important starting material in the manufacture of fabrics, plastics, and synthetic rubber.

HCN is a weak acid, having an ionization constant of 1.3×10^{-9} at 18°C (64°F). In the pure state, it is a highly volatile liquid, boiling at 26°C (79°F).

HCN and the cyanides are highly toxic to animals, a lethal dose to humans usually considered to be 100–200 mg (0.0035–0.007 oz). Death has been attributed, however, to as low a dose as 0.57 mg of HCN per kilogram (9×10^{-6} oz per pound) of body weight. As poisons, they are rapid in their action. When these compounds are inhaled or ingested, their physiological effect involves an inactivation of the cytochrome respiratory enzymes, preventing tissue utilization of the oxygen carried by the blood.

The cyanide ion forms a variety of coordination complexes with transition-metal ions. Representative complexes are those of gold $[Au(CN)_2]^-$, silver $[Ag(CN)_2]^-$, and iron $[Fe(CN)_6]^{4-}$. This property is responsible for several of the commercial uses of cyanides. In the cyanide process, the most widely used method for extracting gold and silver from the ores, the finely divided ore is contacted with a dilute solution of sodium or potassium cyanide. The metal is solubilized as the complex ion, and, following extraction, the pure metal is recovered by reduction with zinc dust. In silver-plating, a smooth adherent deposit is obtained on a metal cathode when electrolysis is carried out in the presence of an excess of cyanide ion. *See* ELECTROPLATING OF METALS; GOLD METALLURGY; SILVER METALLURGY.

$Ca(CN)_2$ is extensively used in pest control and as a fumigant in the storage of grain. In finely divided form, it reacts slowly with the moisture in the air to liberate HCN.

In case hardening of metals, an iron or steel article is immersed in a bath of molten sodium or potassium cyanide containing sodium chloride or carbonate. At temperatures of 750°C (1380°F) and above, the cyanide decomposes at the surface, forming a deposit of carbon which combines with and penetrates the metal. The nitrogen also contributes to the increased hardness at the surface by forming nitrides with iron and alloying metals. *See* SURFACE HARDENING OF STEEL.

NaCN and KCN are produced commercially by neutralization of hydrogen cyanide with sodium or potassium hydroxide. The neutralized solution must then be evaporated to dryness in a strictly controlled manner to avoid undue losses. $Ca(CN)_2$ is produced primarily from calcium cyanamide by reaction with carbon in the presence of sodium chloride at 1000°C (1830°F). *See* COORDINATION CHEMISTRY; FERRICYANIDE; FERROCYANIDE; TOXICOLOGY.

Francis J. Johnston

Bibliography. J. Berg and S. Lippard, *Principles of Bioorganic Chemistry*, 1994; R. H. Dreisbach and W. O. Robertson, *Handbook of Poisoning*, 12th ed., 1987.

Cyanobacteria

A large and heterogeneous group of photosynthetic microorganisms, formerly referred to as blue-green algae. They had been classified with the algae because their mechanism of photosynthesis is similar to that of algal and plant chloroplasts; however, the cells are prokaryotic, whereas the cells of algae and plants are eukaryotic. The name cyanobacteria is now used to emphasize the similarity in cell structure to other prokaryotic organisms. *See* ALGAE; CELL PLASTIDS.

General characteristics. All cyanobacteria can grow with light as an energy source, generating energy and reducing power through oxygen-evolving photosynthesis; carbon dioxide (CO_2) is fixed into organic compounds via the Calvin cycle, the same mechanism used in green plants. Thus, all species will grow in the absence of organic nutrients. However, some species will assimilate organic compounds into cell material if light is available, and a few isolates are capable of growth in the dark by using organic compounds as carbon and energy sources. Some cyanobacteria can shift to a different mode of photosynthesis, in which hydrogen sulfide rather than water serves as the electron donor. Molecular

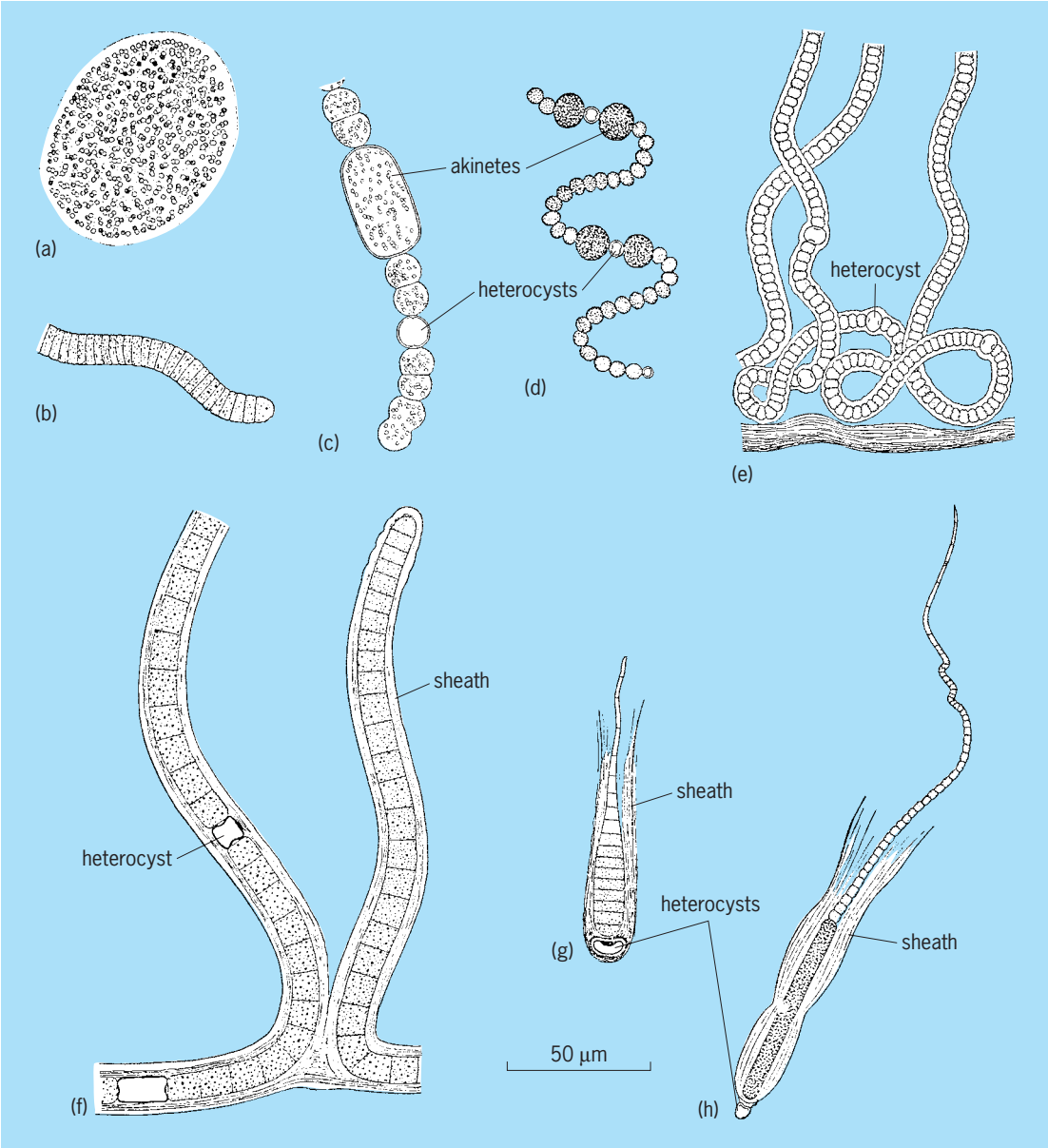

Morphological diversity of cyanobacteria. (*a*) *Microcystis* sp.; (*b*) *Oscillatoria* sp.; (*c, d*) *Anabaena* sp.; (*e*) *Nostoc* sp. with gelatinous capsule; (*f*) *Scytonema* sp.; (*g*) *Calothrix* sp.; (*h*) *Gleotrichia* sp. (*After L. H. Tiffany and M. E. Britton, The Algae of Illinois, University of Chicago Press, 1952*)

oxygen is not evolved during this process, which is similar to that in purple and green photosynthetic sulfur bacteria. *See* PHOTOSYNTHESIS.

The photosynthetic pigments of cyanobacteria include chlorophyll *a* (also found in algae and plants) and phycobiliproteins. One of these latter pigments, *c*-phycocyanin, is blue, and led to the name cyanobacteria (Greek cyano-, blue), and in combination with the green chlorophyll pigment gives cells a blue-green coloration. However, other species also contain a red phycobiliprotein, phycoerythrin, and these are red or brown in color. The coloration of the organisms can also be affected by the presence of orange or yellow carotenoid pigments. *See* CAROTENOID; CHLOROPHYLL.

Form and structure. Cyanobacteria have limited metabolic versatility in comparison with other bacteria, but they are extremely diverse morphologically. Species may be unicellular or filamentous (see **illus.**). Both types may aggregate to form macroscopically visible colonies. The cells range in size from those typical of bacteria (0.5–1 micrometer in diameter) to 60 μm in *Oscillatoria princeps*. The latter is among the largest-sized cells known in prokaryotes.

When examined by electron microscopy, the cells of cyanobacteria appear similar to those of gram-negative bacteria. The cell wall contains an outer lipid membrane and a layer of peptidoglycan. Many species produce extracellular mucilage or sheaths

that promote the aggregation of cells or filaments into colonies.

The photosynthetic machinery is located on internal membrane foldings called thylakoids. Chlorophyll *a* and the electron transport proteins necessary for photosynthesis are located in these lipid membranes, whereas the water-soluble phycobiliprotein pigments are arranged in particles called phycobilisomes which are attached to the lipid membrane.

Several other types of intracellular structures are found in some cyanobacteria. Gas vesicles are rigid shells of protein constructed in a way to exclude water from the interior. The hollow structures, which have a diameter of about 70 nanometers and lengths up to 1000 nm, have densities much less than that of water and thus may confer buoyancy on the organisms. They are often found in cyanobacteria that grow in the open waters of lakes. Polyhedral bodies, also known as carboxysomes, contain large amounts of ribulose bisphosphate carboxylase, the key enzyme of CO_2 fixation via the Calvin cycle. Several types of storage granules may be found; glycogen and polyphosphate granules are also found in other organisms, but cyanophycin granules, which consist of polymers of the amino acids aspartic acid and arginine, are unique to cyanobacteria. These storage granules serve as reserves of carbon and energy, phosphorus, and nitrogen, respectively. They are synthesized when the particular nutrient is present in excess in the environment, and the reserves are utilized when external sources of the nutrient are unavailable.

Specialized cells. Some filamentous cyanobacteria can undergo a differentiation process in which some of the normal vegetative cells become heterocysts. The process occurs when combined nitrogen (ammonia and nitrate) is depleted in the environment. Heterocysts contain the enzyme nitrogenase, which catalyzes the conversion of nitrogen gas to ammonia. Nitrogen fixed in the heterocysts is translocated to the vegetative cells in the filament, to allow their continued growth in the absence of external ammonia or nitrate. Other changes occur during the differentiation into heterocysts. The cell wall becomes thicker, and the heterocyst develops intracellular connections with adjacent vegetative cells. In addition, the heterocysts lose the portion of the photosynthetic apparatus responsible for oxygen evolution. These changes are necessary for nitrogen fixation to occur, because nitrogenase is inactivated by oxygen. Thus, the heterocyst becomes an anaerobic site for nitrogen fixation, and depends upon translocation of organic compounds from vegetative cells to provide the reducing power to drive nitrogen fixation. *See* NITROGEN FIXATION.

In some filamentous species, resting spores called akinetes are formed. These cell forms are more resistant to periods of drying, freezing, or darkness than are the normal vegetative cells.

Occurrence. Cyanobacteria can be found in a wide variety of freshwater, marine, and soil environments. They are more tolerant of environmental extremes than are eukaryotic algae. For example, they are the dominant oxygenic phototrophs in hot springs (at temperatures up to 72°C or 176°F) and in hypersaline habitats such as may occur in marine intertidal zones. In these areas, they are often components of microbial mats, layers of microbial protoplasm up to several centimeters in thickness. In addition to cyanobacteria, other phototrophic bacteria and heterotrophic bacteria proliferate in these mats.

Cyanobacteria are often the dominant members of the phytoplankton in fresh-water lakes that have been enriched with inorganic nutrients such as phosphate. Up until the 1970s, cyanobacteria were not thought to be important members of the oceanic plankton, but it is now known that high population densities of small, single-celled cyanobacteria occur in the oceans, and that these are responsible for 30–50% of the CO_2 fixed into organic matter in these environments. *See* PHYTOPLANKTON.

Some cyanobacteria form symbiotic associations with eukaryotic photosynthetic organisms. These associations are based on the capacity of heterocystous cyanobacteria to fix nitrogen. *Nostoc* species reside in specialized cavities on the undersurface of liverworts, such as *Anthoceros*. In this environment, 30–40% of the *Nostoc* cells become heterocysts (a frequency tenfold higher than in free-living species), and 90% of the nitrogen fixed in the heterocysts is assimilated by the liverwort. Associations also occur with the water fern *Azolla*, some cycads, and the angiosperm *Gunnera*.

Lichens are symbiotic associations between a fungus and a photosynthetic microorganism. About 8% of the lichens involve a cyanobacterium, which can provide both fixed nitrogen and fixed carbon to the fungal partner. *See* LICHENS.

Dispersal. Cyanobacteria in open waters or in microbial mats can adjust their vertical position in the environment. Planktonic cyanobacteria do so by regulating the buoyancy conferred by gas vesicles. These cells can lose buoyancy by accumulating dense polymers such as glycogen. Filamentous cyanobacteria in microbial mats are often capable of gliding motility, a process in which cells lacking flagella can move in contact with surfaces. In both instances, the organisms adjust their vertical positions to avoid high, damaging light irradiances. However, their behavior is also modulated by the presence of inorganic nutrients or toxic substances such as hydrogen sulfide.

Evolution. Cyanobacteria are thought to be the first oxygen-evolving photosynthetic organisms to develop on the Earth, and hence responsible for the conversion of the Earth's atmosphere from anaerobic to aerobic about 2 billion years ago. This development permitted the evolution of aerobic bacteria, plants, and animals. *See* BACTERIA; CYANOPHYCEAE; PREBIOTIC ORGANIC SYNTHESIS. Allan Konopka

Bibliography. Y. Cohen, R. W. Castenholtz, and H. O. Halvorson (eds.), *Microbial Mats: Stromatolites*, 1984; W. M. Darley, *Algal Biology: A Physiological Approach*, 1982; T. Platt and W. K. W. Li (eds.),

Photosynthetic Picoplankton, 1987; B. A. Whitton and N. G. Carr (eds.), *The Biology of Cyanobacteria*, 1982.

Cyanocarbon

A derivative of hydrocarbon in which all of the hydrogen atoms are replaced by the —C≡N group. Only two cyanocarbons, dicyanoacetylene and dicyanobutadiyne, were known before 1958. Since then, tetracyanomethane, hexacyanoethane, tetracyanoethylene, hexacyanobutadiene, and hexacyanobenzene were synthesized.

The term cyanocarbon has been applied to compounds which do not strictly follow the above definition: tetracyanoethylene oxide, tetracyanothiophene, tetracyanofuran, tetracyanopyrrole, tetracyanobenzoquinone, tetracyanoquinodimethane, tetracyanodithiin, pentacyanopyridine, diazomalononitrile, and diazotetracyanocyclopentadiene

Tetracyanoethylene, the simplest olefinic cyanocarbon, is a colorless, thermally stable solid, having a melting range of 198–200°C (388–392°F). It is a strong π-acid, readily forming stable complexes with most aromatic systems. These complexes absorb radiation in or near the visible spectrum.

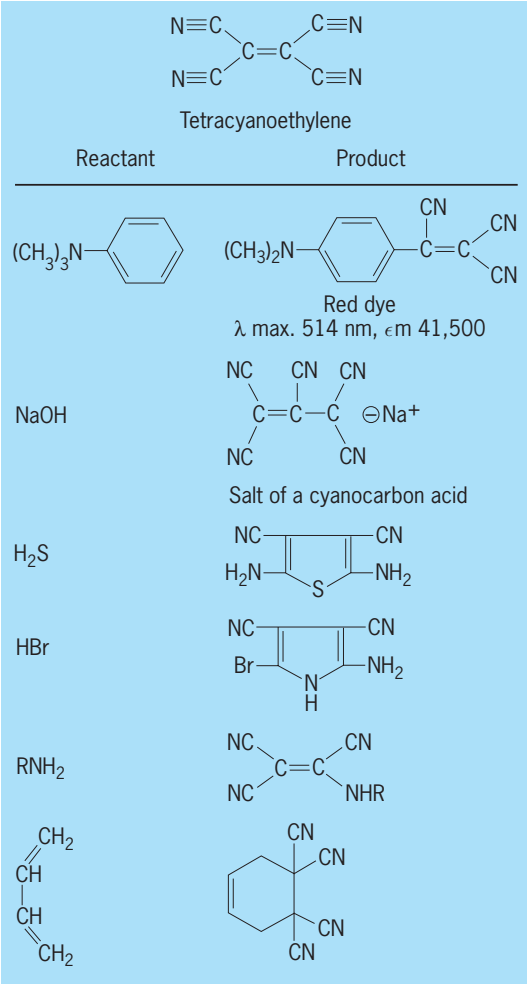

Reactions of tetracyanoethylene.

Charge-transfer salts derived from tetracyanoquinodimethane and tetrathiafulvalene are highly conducting and have attracted much interest as possible replacements for metal conductors.

Cyanocarbon acids are among the strongest protonic organic acids known and are usually isolated only as the cyanocarbon anion salt. These salts are usually colored and stabilized by resonance of the type indicated in the reaction below.

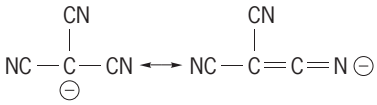

Twenty-five similar forms can be written for the 2-dicyanomethylene-1,1,3,3-tetracyanopropanediide ion, $(NC)_2C{=}C[C(CN)_2]$, which accounts for the similarity between the acid strength of the free acid and that of sulfuric acid.

The reactivity of cyanocarbons is illustrated by the facile replacement of one of the cyano groups in tetracyanoethylene by nucleophilic attack under very mild conditions and by its addition to dienes. Examples of these reactions are shown in the **illustration**. *See* ORGANIC CHEMISTRY.

Owen W. Webster

Bibliography. R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; R. T. Morrison and R. N. Boyd, *Organic Chemistry*, 7th ed., 2001; Z. Rappaport (ed.), *The Chemistry of the Cyano Group*, 1970, and Suppl. C ed. by S. Patoi and Z. Rappaport, 1983.

Cyanogen

A colorless, highly toxic gas having the molecular formula C_2N_2. Cyanogen belongs to a class of compounds known as pseudohalogens, because of the similarity of their chemical behavior to that of the halogens. Liquid cyanogen boils at −21.17°C (−6.11°F) and freezes at −27.9°C (18.2°F) at 1 atm (10^5 Pa). The density of the liquid is 0.954 g/ml at the boiling point.

Cyanogen may be prepared by prolonged heating of mercuric cyanide at 400°C (750°F), reaction (1), or by allowing a solution of copper sulfate to

$$Hg(CN)_2 \rightarrow Hg + C_2N_2 \qquad (1)$$

flow slowly into a solution of potassium cyanide, reaction (2). The copper(II) cyanide is unstable, decomposing to give cyanogen and copper(I) cyanide, reaction (3).

$$2KCN + CuSO_4 \rightarrow Cu(CN)_2 + K_2SO_4 \qquad (2)$$

$$2Cu(CN)_2 \rightarrow 2CuCN + C_2N_2 \qquad (3)$$

Cyanogen is also formed by the reaction in the gas phase of hydrogen cyanide and chlorine at

400–700°C (750–1300°F) in the presence of a catalytic agent such as activated carbon, reaction (4).

$$2HCN + Cl_2 \rightarrow C_2N_2 + 2HCl \qquad (4)$$

Cyanogen chloride, ClCN, is an intermediate in this reaction. When heated to 400°C (750°F), cyanogen gas polymerizes to a white solid, paracyanogen $(CN)_x$.

Cyanogen reacts with hydrogen at elevated temperatures in a manner analogous to the halogens, forming hydrogen cyanide, reaction (5). With hydrogen sulfide, H_2S, cyanogen forms thiocyanoformamide or dithiooxamide. These reactions are shown by (6) and (7).

$$C_2N_2 + H_2S \longrightarrow NC-\overset{\overset{\displaystyle S}{\|}}{C}-NH_2 \qquad (6)$$

$$C_2N_2 + 2H_2S \longrightarrow H_2NC-\overset{\overset{\displaystyle S}{\|}}{C}\overset{\overset{\displaystyle S}{\|}}{C}NH_2 \qquad (7)$$

Cyanogen burns in oxygen, producing one of the hottest flames known from a chemical reaction. It is considered to be a promising component of high-energy fuels.

Structurally, cyanogen is written N≡C—C≡N. *See* CYANIDE. Francis J. Johnston

Bibliography. R. J. Fressenden and J. S. Fressenden, *Organic Chemistry*, 5th ed., 1994; Z. Rappaport (ed.), *The Chemistry of the Cyano Group*, 1971.

Cyanophyceae

A class of prokaryotic organisms coextensive with the division Cyanophycota of the kingdom Monera. Because these organisms have chlorophyll *a* and carry out oxygen-evolving photosynthesis, they have traditionally been aligned with algae and, with regard for their characteristic color, called blue-green algae. Microbiologists have emphasized the prokaryotic structure of these organisms and aligned them with bacteria, as the Cyanobacteria. Other names applied to these organisms include Cyanophyta at the level of division and Myxophyceae or Schizophyceae at the level of class. Blue-greens range in form from unicells 1–2 micrometers in diameter to filaments 10 cm (4 in.) long. *See* CYANOBACTERIA.

Classification. The taxonomy of the blue-greens is in an especially unsteady state, polarized by two viewpoints. At one pole (Geitlerian system) nearly every morphological variant encountered in nature is accorded taxonomic recognition. At the other pole (Drouetian system) great morphological plasticity is attributed to each species. Geitler recognized 4 orders, 22 families, 140 genera, and more than 1200 species, while Drouet recognized 2 orders, 6 families, 24 genera, and 61 species. Most specialists of Cyanophyceae steer a middle course. A further major consideration has recently been introduced by microbiologists, who have proposed their own system of classification based solely on characters revealed in pure cultures.

Unicellular forms, which may aggregate in colonies or loosely constructed filaments and which reproduce by binary fission or spores, constitute the order Chroococcales, from which two additional orders, Chamaesiphonales and Pleurocapsales, are sometimes segregated. Filamentous forms, which reproduce by hormogonia, constitute the order Nostocales (= Hormogonales or Oscillatoriales). The Nostocales may be restricted to unbranched or falsely branched forms (in which the ends of a trichome adjacent to a rupture grow out as a pair of branches), while those forms in which cells divide in more than one plane (true branching) constitute the order Stigonematales.

Habitat. Geographically and ecologically, blue-green algae are nearly as ubiquitous as bacteria. They are especially abundant in the plankton of neutral or alkaline eutrophic fresh waters and tropical seas, often forming blooms. Habitats for benthic forms include hot springs, snow and ice, soil, rocks, tree trunks, and buildings. Cyanophyceae live symbiotically with a large variety of animals and plants, including sponges, diatoms, liverworts, cycads, the aquatic fern *Azolla*, and the flowering plant *Gunnera*. They constitute the phycobiont of many lichens.

In addition to contributing to food chains, blue-green algae play specific beneficial roles. Nitrogen-fixing forms greatly enrich rice paddies. *Spirulina*, a traditional food in parts of Mexico and central Africa, is grown commercially and marketed widely as a high-protein dietary supplement. On the other hand, blue-green algae are often a nuisance. They clog filters, impart undesirable tastes and odors to domestic water supplies, and make unusable or at least unattractive many swimming pools, aquariums, and fountains. Cyanophycean blooms are often toxic to fish, birds, and livestock.

Blue-green algae were pioneers on Earth and are known from rocks at least as old as 2.3 billion years. They are believed to have been responsible for the accumulation of oxygen in the primeval atmosphere and to have been involved in the formation of laminated reeflike structures called stromatolites. To judge from similar modern structures (as at Shark Bay, Western Australia), stromatolites were produced by algal mats that entrapped detrital sediments and sometimes deposited calcium carbonate. *See* ALGAE; STROMATOLITE. Paul C. Silva; Richard L. Moe

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae: Structure and Reproduction*, 1985; F. Drouet, Revision of the Stigonemataceae with a summary of the classification of the blue-green algae, *Beih. Nova Hedwigia*, Heft 66, 1981; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Cybernetics

The study of communication and control within and between humans, machines, organizations, and society. This is a modern definition of the term cybernetics, which was first utilized by N. Wiener in 1948 to designate a broad subject area he defined as "control and communication in the animal and the machine." A distinguishing feature of this broad field is the use of feedback information to adapt or steer the entity toward a goal. When this feedback signal is such as to cause changes in the structure or parameters of the system itself, it appears to be self-organizing. *See* ADAPTIVE CONTROL.

Wiener's thinking was influenced by his work during World War II on the aiming of an antiaircraft gun ahead of its target in order to achieve a successful intercept. He was also familiar with the work of physiologists and believed that similar mechanisms are exhibited in biological systems to provide for goal-directed behavior. He developed the statistical methods of autocorrelation, prediction, and filtering of time-series data to provide a mathematical description of both biological and physical phenomena. The use of filtering to remove unwanted information or noise from the feedback signal mimics the selectivity shown in biological systems in which imprecise information from a diversity of sensors can be accommodated so that the goal can still be reached. *See* ESTIMATION THEORY; HOMEOSTASIS; STOCHASTIC CONTROL THEORY.

This similarity between physical and biological systems was also recognized by a number of scientists and engineers in the 1940s, 1950s, and 1960s, such as W. McCulloch, who described the brain as a digital computer. Computing systems that exhibit behavior somewhat akin to that of humans are said to possess artificial intelligence. The field of neural networks is a subset of artificial intelligence in that intelligent behavior can be realized with a collection of computing elements (like biological neurons) connected together to merge various inputs together into a single signal, generally through use of a weighted sum of the individual signals. Usually, a threshold is set and, if this weighted sum is above the set threshold, a neuron fires, indicating that something of importance has happened. Neural networks that adjust the weights by using internal feedback appear to be self-organizing, and are thus consistent with Wiener's view of cybernetic systems. *See* ARTIFICIAL INTELLIGENCE; CONTROL SYSTEMS; ELECTRICAL COMMUNICATIONS; HUMAN-FACTORS ENGINEERING; HUMAN-MACHINE SYSTEMS; NEURAL NETWORK; SYSTEMS ENGINEERING.

Donald W. Bouldin

Bibliography. T. Kohonen, *Self-Organization and Associative Memory*, 2d ed., 1988; W. McCulloch, *Embodiments of Mind*, 1965, paper 1988; M. Singh (ed.), *Systems and Control Encyclopedia*, 1987; N. Wiener, *Cybernetics*, 2d ed., 1961; N. Wiener, *Extrapolation: Interpolation and Smoothing of Stationary Time Series with Engineering Applications*, 1949.

Cycadales

An order of the class Cycadopsida of the division Pinophyta (gymnosperms) consisting of four families, Cycadaceae, Stangeriaceae, Zamiaceae, and Boweniaceae, with perhaps 100 species. The order dates from the upper Carboniferous and has few living representatives. The cycads were distributed worldwide in the Mesozoic, but today are restricted to subtropical and tropical regions, with the plants occurring in small colonies; few have broad distributions.

The cycads, often incorrectly referred to as palms, range from a few inches (*Zamia*) to 65 ft (20 m; *Macrozamia*) tall. The stems are cylindrical and often branched; in some the stem is subterranean, in others it is mainly above the ground. The pinnate (or bipinnate in *Bowenia*) leaves are borne at the apex of the stems. Microsporophylls and megasporophylls are borne in cones of highly varied appearance. Male and female cones appear on separate plants (dioecious). *See* CYCADOPSIDA; PINOPHYTA; PLANT KINGDOM.

Thomas A. Zanoni

Cycadeoidales

An order of extinct plants that formed a conspicuous part of the landscape during the Triassic, Jurassic, and Cretaceous periods. The Cycadeoidales (or Bennettitales) had unbranched or sparsely branched trunks with a terminal crown of leaves (family Cycadeoidaceae), or they were branched, at times profusely (family Williamsoniaceae).

The squat, barrel-shaped trunks of members of the Cycadeoidaceae were covered with a dense armor of persistent leaf bases, and were terminated by a crown of pinnately divided leaves (see **illus.**).

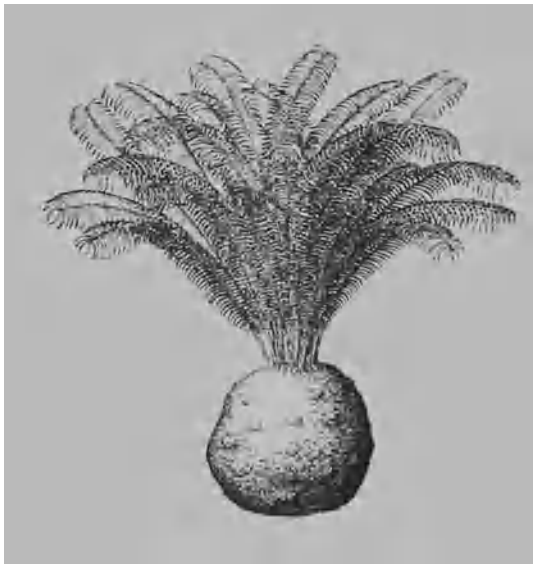

Reconstruction of a plant of the genus *Cycadeoidea*, showing terminal crown of leaves, persistent leaf bases, and positions of the fruiting structures. (*Courtesy of T. Delevoryas*)

On the trunk surface, among the leaf bases, were borne complex fruiting structures, each with a central seed-bearing receptacle, surrounded by a whorl of pollen-bearing organs that were derived from modified leaves. Dicotyledonous embryos have been observed in the seeds.

In members of the Williamsoniaceae, fruiting structures were often stalked, and were flowerlike in appearance. They had both pollen-bearing and seed-bearing structures in the same "flower," or these organs may have been borne in separate "flowers." Seeds were borne on the surface of a domelike receptacle, while the pollen-bearing organs were fused at the base to form a whorl of simple or pinnate structures. Foliage was either pinnately divided or entire. No known relatives of the Cycadeoidales exist at the present time, although members of the order Cycadales show some resemblances to the extinct group. *See* CYCADALES; EMBRYOBIONTA; PALEOBOTANY; PLANT KINGDOM. Theodore Delevoryas

Cycadopsida

A class of the division Cycadophyta consisting of a single order, Cycadales, dating back to the Triassic and most abundant in the Jurassic. There are 100–150 living species of cycads in two families, Cycadaceae with one genus and Zamiaceae with ten genera. Most cycads are tropical, but some extend into warm temperate regions. Wild cycads are threatened by collection for horticulture so are protected by international treaties. They were used as traditional foods in some places but contain potent carcinogens and neurotoxins. *See* CYCADALES; GNETALES; WELWITSCHIALES.

Pinnately compound leaves on massive, usually unbranched trunks emerge in flushes, sometimes with the rachis or leaflets coiled like fiddleheads. Large pollen and seed cones crown different individuals. Pollen cones have numerous wedge-shaped scales, each covered below with clustered microsporangia. Seed cones of Zamiaceae are similar, with two seeds on each scale, but megasporophylls of Cycadaceae are leaflike, with two to eight seeds. As in *Ginkgo*, fertilization is by swimming sperm. *See* PINOPHYTA; PLANT KINGDOM. James E. Eckenwalder

Bibliography. K. J. Norstog and T. J. Nichols, *The Biology of the Cycads*, Cornell University Press, Ithaca, 1998.

Cyclanthales

An order of flowering plants, division Magnoliophyta (Angiospermae), sometimes called Synanthae or Synanthales, in the subclass Arecidae of the class Liliopsida (monocotyledons). The order consists of the single tropical American family Cyclanthaceae, with about 180 species. They are herbs or, seldom, more or less woody plants, with characteristic leaves that have a basal sheath, a petiole, and an expanded, usually bifid (cleft into two equal parts) blade which is often folded lengthwise (plicate) like that of a palm leaf. The numerous, small, unisexual flowers are crowded into a spadix which serves as a sort of pseudanthium. One of the species, *Carludovicia palmata*, is a principal source of fiber for panama hats. *See* ARECIDAE; LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM. Arthur Cronquist

Cyclic nucleotides

Cyclic nucleotides, derivatives of nucleic acids, control the activity of several proteins within cells to regulate and coordinate metabolism. They are members of a group of molecules known as intracellular second messengers; their levels are regulated by hormones and neurotransmitters, which are the extracellular first messengers in a regulatory pathway. Cyclic nucleotides are found naturally in living cells, from bacteria, to yeast, to human cells.

Formation and degradation. Two major forms of cyclic nucleotides are characterized: 3′,5′-cyclic adenosine monophosphate (cyclic AMP or cAMP) and 3′,5′-cyclic guanosine monophosphate (cyclic GMP or cGMP). Like all nucleotides, cAMP and cGMP contain three functional groups: a nitrogenous aromatic base (adenine or guanine), a sugar (ribose), and a phosphate. Cyclic nucleotides differ from other nucleotides in that the phosphate group is linked to two different hydroxyl (3′ and 5′) groups of the ribose sugar and hence forms a cyclic ring. This cyclic conformation allows cAMP and cGMP to bind to proteins to which other nucleotides cannot.

Specific enzymes are responsible for the formation of cyclic nucleotides. Adenylate cyclase forms cAMP from the nucleotide adenosine triphosphate (ATP), and guanylate cyclase forms cGMP from the nucleotide guanosine triphosphate (GTP) [see **illus.**]. These enzymes form cyclic nucleotides at a very slow rate when cells are in a basal, unstimulated state. However, when certain hormones and neurotransmitters bind to specific receptors on the surface of cells, the cyclases are stimulated more than a hundredfold in their rate of cyclic nucleotide synthesis. The result is a large increase in the intracellular concentration of either cAMP or cGMP.

Cyclic nucleotides are degraded by cyclic nucleotide phosphodiesterases, which cleave the phosphoester bond between the phosphate and the 3′ hydroxyl group of the ribose to generate the 5′-monophosphate nucleotide (see illus.). Many different types of phosphodiesterases are found in various tissues of the body, and several clinically useful drugs inhibit the activity of phosphodiesterases so that cyclic nucleotide concentrations are increased in specific tissues. An example is theophylline, used to treat bronchial asthma.

The increase in cAMP or cGMP triggered by hormones and neurotransmitters can have many different effects on any individual cell. The type of effect is dependent to some extent on the cellular proteins to which the cyclic nucleotides may bind. Three

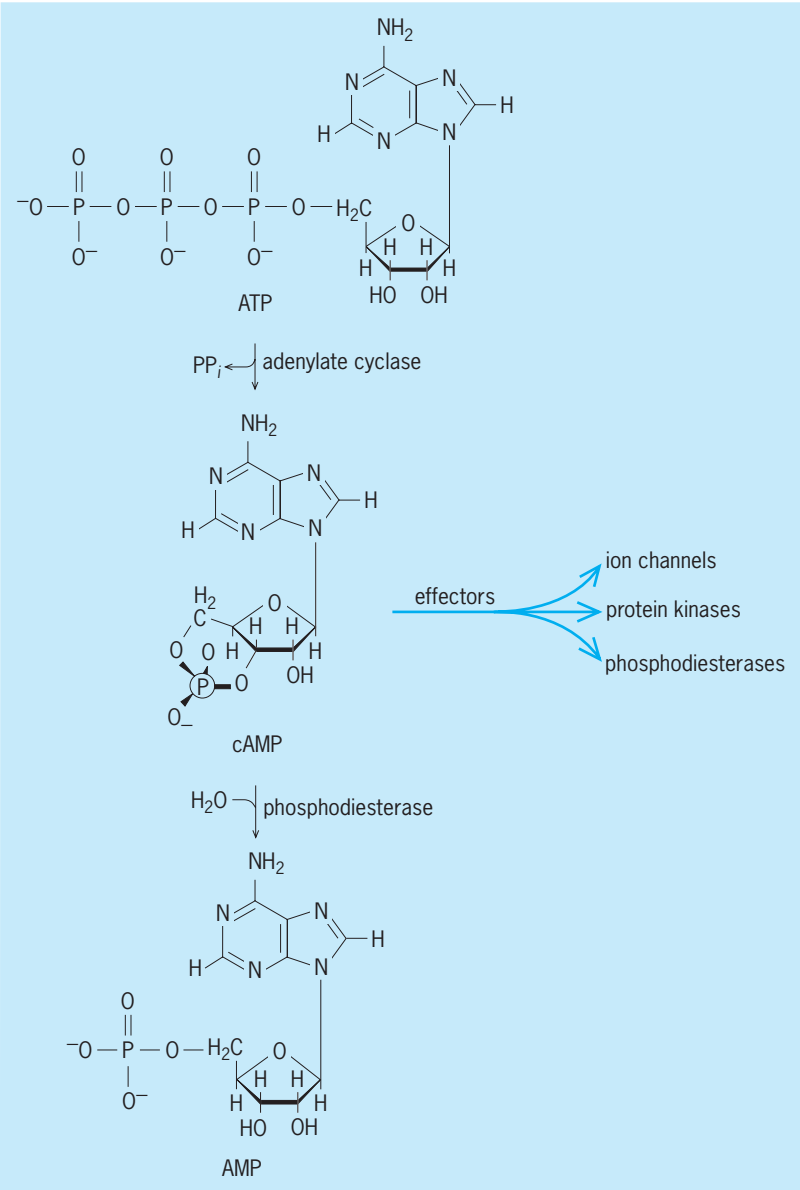

Biochemical structures of ATP, cAMP, and AMP. In cAMP, a cyclic ring involves the 3′ and 5′ hydroxyl groups of the ribose with the phosphate. There are three major effector mechanisms for the second messenger action of cAMP in human cells.

types of effector proteins are able to bind cyclic nucleotides.

Effector proteins. Protein kinases constitute the first type of effector protein regulated by cyclic nucleotides. They are enzymes which are able to transfer a phosphate group to (phosphorylate) individual amino acids of other proteins. This action often changes the function of the phosphorylated protein. For example, the enzyme phosphorylase is activated by phosphorylation, resulting in the release of glucose into the bloodstream from glycogen stored in the liver. Cyclic nucleotides activate the cAMP-dependent protein kinases (PKA) and the cGMP-dependent protein kinases (PKG). PKA is found in all cell types, where it regulates the flow of ions across the cell membrane, the activity of metabolic enzymes such as phosphorylase, the growth and division of cells, and the rate at which individual proteins are synthesized. PKA phosphorylates distinct cellular proteins involved in each of these aspects of cell function. PKG is found only in certain cells of the body, including smooth muscle cells lining the blood vessels where it regulates blood flow, and neurons of the brain where it regulates their response to neurotransmitters.

The second type of protein regulated by cyclic nucleotides comprises ion channels. These proteins are found in the outer plasma membrane of some cells, and the binding of cyclic nucleotides can alter the flow of sodium ions across the cell membranes. These ion channels are found in greatest abundance in the sensory neurons of the retina and nose as well in some regions of the brain. The binding of cyclic nucleotides to these ion channels alters the sensitivity of the neurons to neurotransmitters. In the retina, cGMP plays a critical role in vision by regulating an ion channel.

The third type of protein to which cyclic nucleotides bind are the cyclic nucleotide phosphodiesterases. These enzymes are responsible for the degradation of cyclic nucleotides. The degradation rate can be affected by the amount and nature of the cyclic nucleotides present in a cell. For example, high levels of cGMP can inhibit the phosphodiesterase that is responsible for the degradation of cAMP.

In bacteria, cAMP can bind to a fourth type of protein, which can also bind to deoxyribonucleic acid (DNA). This catabolite gene activator protein (CAP) binds to specific bacterial DNA sequences, stimulating the rate at which DNA is copied into ribonucleic acid (RNA) and increasing the amount of key metabolic enzymes in the bacteria. This type of regulation does not occur in eukaryotic or animal cells.

In humans, cyclic nucleotides acting as second messengers play a key role in many vital processes and some diseases. In the brain, cAMP and possibly cGMP are critical in the formation of both long-term and short-term memory. In the liver, cAMP coordinates the function of many metabolic enzymes to control the level of glucose and other nutrients in the bloodstream. Cyclic GMP mediates the effects of nitroglycerin on smooth muscle cells in the arteries of the heart to alleviate the lack of blood flow associated with angina. Cholera results in abnormally high levels of cAMP in the intestinal epithelial cells and the loss of large amounts of fluid, resulting in diarrhea. In some cases of cystic fibrosis, the inability of cAMP to regulate an ion channel causes the disease. *See* ADENOSINE TRIPHOSPHATE (ATP); ENZYME; NUCLEIC ACID; NUCLEOPROTEIN; NUCLEOTIDE; PROTEIN.

Michael D. Uhler

Bibliography. F. Burns, A. Z. Zhao, and J. A. Beavo, Cyclic nucleotide phosphodiesterases: Gene complexity, regulation by phosphorylation, and physiological implications, *Adv. Pharmacol.*, 36:29–48, 1996; M. D. Houslay and G. Milligan, Tailoring cAMP-signalling responses through isoform multiplicity, *Trends Biochem. Sci.*, 22:217–224, 1997; G. Siegel (ed.), *Basic Neurochemistry*, 6th ed., Raven Press,

1998; D. Voet and J. G. Voet, *Biochemistry*, 2d ed., John Wiley,1997; W. N. Zagotta and S. A. Siegelbaum, Structure and function of cyclic nucleotide-gated channels, *Annu. Rev. Neurosci.*, 19:235–263, 1996.

Cycliophora

A recently described phylum in the animal kingdom. Only one microscopic species, *Symbion pandora*, has been described. Its sessile stage is approximately 0.3 mm in length and 0.1mm wide. Surprisingly, Cycliophora was not discovered in an exotic biotope such as a deep-sea vent or marine caves, but on the Norwegian lobster. All known occurrences are from the crustacean mouth limbs, which can be totally encrusted with thousands of the sessile stage of the microscopic animal (**Fig. 1**). Two other undescribed species, very similar to the type species, are found on the American lobster (*Homarus americanus*) and the common lobster (*H. gammarus*).

The sessile animal has a circular mouth surrounded by a ring of cilia (**Fig. 2**). The mouth ring is used for filtering small food particles, such as bacteria or algae. The anatomy is relatively simple, with a U-shaped gut, similar to what is found in bryozoans or some sessile rotiferans. It is the extremely complex life cycle of *S. pandora* that makes the phylum unique. The life cycle consists of an asexual and a sexual generation with no less than two types of free-swimming larva, dwarf male and female; two stages of sessile feeding animals, which brood the male and female; and one type of larva inside a brooding chamber (called a marsupium). Furthermore, the sessile animal has internal budding, whereby it grows by loss of the head itself and replacement of the old gut and feeding system with a new bud coming from embryonic cells in the basal part of the body. The larva developed inside the feeding animal is called the Pandora larva. The feeding animal thus has developed both a new bud with a head and the Pandora larva. However, inside the posterior segment of the Pandora larva, a new internal bud with the head of a new feeding animal is developed before the larva leaves the sessile mother animal.

When the Pandora larva emerges, it settles on the setae of the lobster very close to the maternal individual, and the internal bud pops up from the rear to form a new feeding-stage animal. In this way the *S. pandora* can multiply very fast and completely cover the mouth limbs in a whitish cast. After several cycles, or before the lobster molts or dies, the feeding stage produces a single male or female in the marsupium in the same way as the Pandora larva. It is thought that the molting hormones of the lobster may trigger the sexual life cycle.

The dwarf male is released from the marsupium with an undeveloped testis and without a penis. It settles on a feeding animal. If the feeding animal is brooding a female inside, the newly settled male begins to make internal buddings and it forms up to three testes and penial structures. The female looks very similar to a Pandora larva except that she lacks the internal budding; rather she has a single large egg posteriorly. This egg is fertilized inside the female before release from the marsupium. The fertilized egg (zygote) immediately begins to cleave inside the female when it escapes the brooding chamber. She settles very close to her mother on the same host and now broods a totally new type of larva, the chordoid larva inside herself. In fact, the larval tissue engulfs all her tissue, so that only her cuticle remains. The chordoid larva hatches from the female

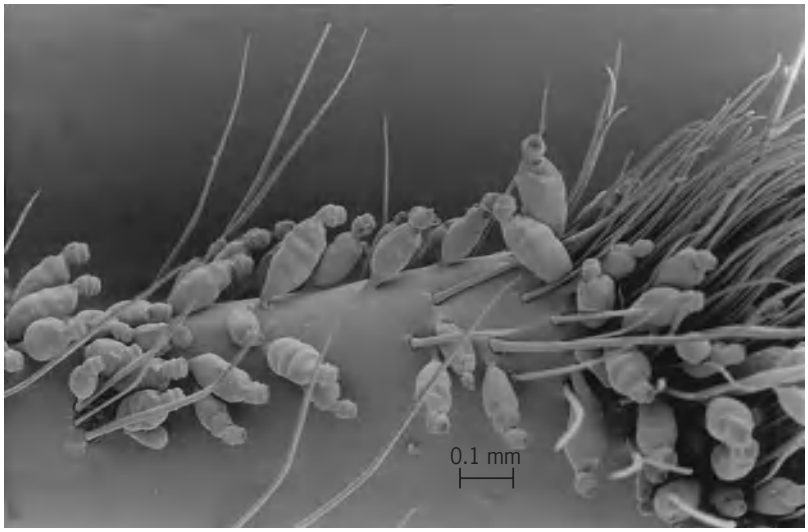

Fig. 1. Scanning electron micrograph of the Norwegian lobster's mouthpart with numerous feeding stages of *Symbion pandora*. (*Courtesy of P. Funch and R. M. Kristensen*)

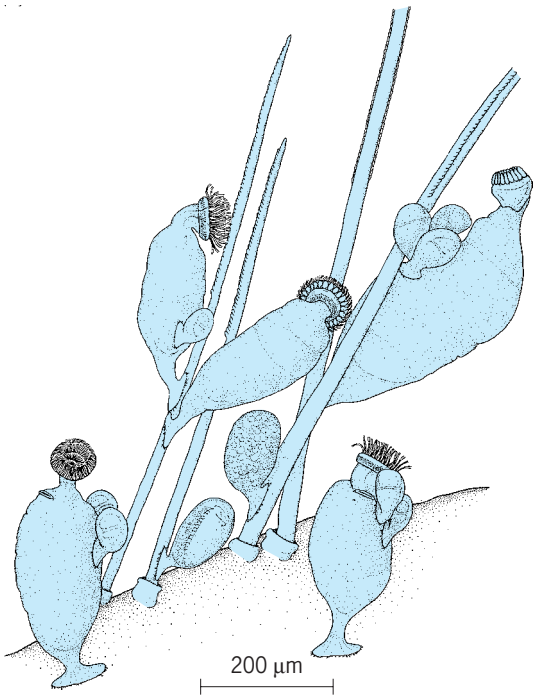

Fig. 2. Drawing of the first described cycliophoran, *Symbion pandora*, from the Norwegian lobster, *Nephrops norvegicus*. (*Drawn by Stine Elle*)

exuvium head first and swims away. The chordoid larva is the dispersal stage. It is a good swimmer and has a strong chorda-like structure similar to that found in the larva of tunicates, although it is not believed that these structures are homologous in the two groups. The chordoid larva settles on a new host and forms a new feeding animal; a new life cycle can begin.

Most of the life cycle of *S. pandora* was elucidated in the Faeroe Islands in the North Atlantic, but there is no reason why the Cycliophora should not be found on other invertebrates worldwide.

Reinhardt Møbjerg Kristensen

Bibliography. S. Conway Morris, A new phylum from the lobster's lip, *Nature*, 378:661–662, 1995; P. Funch, The chordoid larva of *Symbion pandora* (Cycliophora) is a modified trochophore, *J. Morphol.*, 230:231–263, 1996; P. Funch and R. M. Kristensen, Cycliophora is a new phylum with affinities to Entoprocta and Ectoprocta, *Nature*, 378:711–714, 1995; P. Funch and R. M. Kristensen, Cycliophora, in F. W. Harrison and R. M. Woollacott (eds.), *Microscopic Anatomy of Invertebrates*, vol. 13: *Lophophorates, Entoprocta, and Cycliophora*, Wiley/Liss, New York, 1997; P. Funch and R. M. Kristensen, Cycliophora, in E. Knobil and J. D. Neil (eds.), *Encyclopedia of Reproduction*, vol. 1, Academic Press, 1999; B. M. H. Winnepenninckx, T. Backeljau, and R. M. Kristensen, Relations of the new phylum Cycliophora, *Nature*, 393:636–638, 1998.

Cycloamylose

Any of a group of cyclic oligomers of glucose in the normal C-1 conformation in which the individual glucose units are connected by 1,4 bonds. They are called Schardinger dextrins (after the discoverer), cyclodextrins (to emphasize their cyclic character), or cycloamyloses (to emphasize both their cyclic character and their amylose origin). The most common ones are cyclohexaamylose (with six glucose units in a cyclic array) and cycloheptaamylose (with seven glucose units in a cyclic array). The toroidal shape of these molecules has been determined by x-ray analysis (**Fig. 1**). The molecular weights are around 1000.

In water or in a mixture of water and dimethyl sulfoxide, cycloamyloses form 1:1 complexes with many organic molecules and most benzenoid derivatives. X-ray and spectroscopic evidence indicates that the guest molecule (usually a small organic molecule) is bound tightly in the cavity of the host molecule (the cycloamylose) just like a key fitting into a lock. The result is an inclusion complex, which is what is formed by enzymes when they first bind the molecules whose reactions they subsequently catalyze. *See* CATALYSIS; CLATHRATE COMPOUNDS.

Inclusion complex formation is probably the most important similarity between cycloamyloses and enzymes. In the cycloheptaamylose-catalyzed hydrolysis of *p*-nitrophenyl acetate, a nonlinear increase in the rate of reaction is obtained when one increases the concentration of cycloamylose. However, a linear increase is obtained when the same data are plotted in a double reciprocal manner. These two concentration dependencies are exactly those seen in enzyme-catalyzed reactions, and show that enzymes and cycloamyloses increase the rate of reaction in the same way.

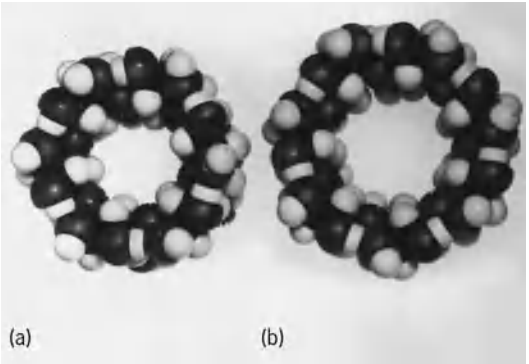

Fig. 1. Molecular models of cycloamyloses viewed from the secondary hydroxyl side of the torus. (*a*) Cyclohexaamylose. (*b*) Cycloheptaamylose. (*From M. L. Bender and M. Komiyama, Cyclodextrin Chemistry, Springer, 1978*)

Specificity is one of the most important aspects of enzyme action. A particular enzyme can catalyze a given reaction but cannot catalyze a closely related reaction. The cycloamyloses provide the answer to how an enzyme recognizes a given compound but ignores a closely related compound.

The importance of the stereochemistry of binding may be seen easily in the models of complexes of corresponding meta- and para-substituted esters with cyclohexaamylose: cyclohexaamylose *m-t*-butylphenyl acetate and cyclohexaamylose *p-t*-butylphenyl acetate. In the former complex, a glucosidic hydroxyl group is practically adjacent to the ester carbonyl group with which it is to react, leading to a facile reaction (**Fig. 2***a*). In the latter complex, the complex, the glucosidic hydroxyl group is several tenths of a nanometer away from the ester carbonyl group

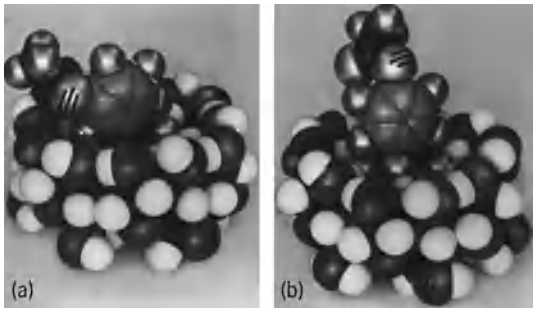

Fig. 2. Molecular models of cycloamylose complexes. (*a*) Cyclohexaamylose-*m-t*-butylphenyl acetate complex. (*b*) Cyclohexaamylose-*p-t*-butylphenyl acetate complex. (*From R. L. Van Etten et al., Acceleration of phenyl ester cleavage by cycloamyloses: A model for enzymatic specificity, J. Amer. Chem. Soc., 89:3249, 1967*)

with which it must react, leading to a slow reaction (Fig. 2*b*).

Cycloamyloses have been used to mimic many enzymes, such as ribonuclease, transaminase, and carbonic anhydrase. The catalysis of the hydrolysis of phenyl acetates discussed above is a mimicry of the enzyme chymotrypsin by a cycloamylose. Both cycloamyloses and chymotrypsin catalyze via three steps: binding; acylation of the catalyst; and deacylation of the acyl-catalyst. The mechanisms by which cycloamyloses and chymotrypsin catalyze this reaction are identical: cycloamyloses form a noncovalent complex followed by the formation of an acyl-cycloamylose intermediate, while chymotrypsin forms a noncovalent enzyme-substrate complex followed by an acyl-chymotrypsin intermediate. Cycloamyloses, like chymotrypsin, catalyze the hydrolysis of both esters and amides.

Cycloamyloses have also been used to effect selectivity in several chemical reactions. Anisole can be chlorinated by HOCl in the presence of the cyclohexaamylose exclusively in the position para to the methoxy group, although the absence of cyclohexaamylose gives the usual mixture of ortho and para chlorinated products. The ortho position of anisole is buried inside the cavity when the anisole is bound in the cyclohexaamylose, and so reaction takes place in the para position only, although the methoxy group is an ortho, para director in aromatic electrophilic substitution. In a similar sense, it was found that the introduction of aldehyde groups into many phenols by chloroform and sodium hydroxide leads solely to the para-substituted benzaldehyde because the dichlorocarbene attacking agent is prevented from approaching the ortho position by the cycloheptaamylose.

Cycloamyloses are the premier example of chemical compounds that act like enzymes in inclusion complex formation followed by reaction. *See* ENZYME; STEREOCHEMISTRY; STEREOSPECIFIC CATALYST; STERIC EFFECT (CHEMISTRY).

Myron L. Bender

Bibliography. M. L. Bender et al., *The Bioinorganic Chemistry of Enzymatic Catalysis*, 1984; M. L. Bender and M. Komiyama, *Cyclodextrin Chemistry*, 1978; R. Breslow, Artificial enzymes, *Science*, 218:532–537, 1982; D. W. Griffiths and M. L. Bender, Cycloamyloses as catalysts, *Adv. Catalysis,* 23:209–261, 1973.

Cyclocystoidea

An extinct class of small, disk-shaped echinozoans in which the lower surface of the body probably consisted of a suction cap for adhering to substrate, and the upper surface was covered by separate plates arranged in concentric rings. The mouth lay at the center of the upper surface, and the anus, also on the upper surface, lay some distance between the margin and the mouth. Little is known of the habits of cyclocystoids. They occur in rocks of middle Ordovician to middle Devonian age in Europe and North America. *See* ECHINODERMATA; ECHINOZOA.

Howard B. Fell

Bibliography. R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, pt. U, 3(1):188–210, 1966.

Cycloid

A curve traced in the plane by a point on a circle that rolls, without slipping, on a line. If the line is the x axis of a rectangular coordinate system, at whose origin O the moving point P touches the axis, parametric equations of the cycloid are $x = a(\theta - \sin\theta)$, $y = a(1 - \cos\theta)$, when a is the radius of the rolling circle, and the parameter θ is the variable angle through which the circle rolls (see **illus.**). One arch is obtained where θ assumes all values from 0 to 2π. The length of an arch is $8a$, and the area bounded by an arch and the x axis is $3\pi a^2$ (three times the area of the rolling circle).

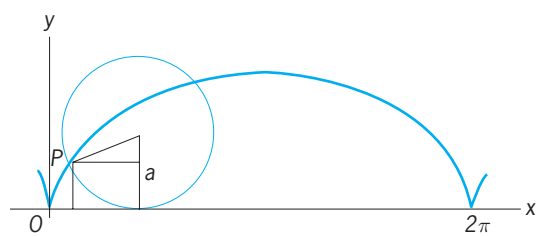

Diagram of cycloid; symbols explained in text.

Two important properties of the cycloid are the following: (1) It is isochronous, that is, if the line is horizontal, then the time of descent of a particle starting from rest at a point of a cycloidal runway and falling to its lowest point is independent of the starting point. (2) It is a brachistochrone; that is, if P_1, P_2 are two points in a vertical plane, but not in a vertical line, the curve down which a particle will fall from P_1 to P_2 in the shortest time is an arc of a cycloid. *See* ANALYTIC GEOMETRY.

Leonard M. Blumenthal

Cyclone

A vortex in which the sense of rotation of the wind about the local vertical is the same as that of the Earth's rotation. Thus, a cyclone rotates clockwise in the Southern Hemisphere and counterclockwise in the Northern Hemisphere. In meteorology the term cyclone is reserved for circulation systems with horizontal dimensions of hundreds of kilometers (tropical cyclones) or thousands of kilometers (extratropical cyclones). For such systems the Coriolis force due to the Earth's rotation and the centrifugal force due to flow curvature are both directed to the right of the direction of the flow. Away from the surface of the Earth these two forces are approximately balanced by the pressure gradient force, which is directed toward low pressure (**Fig. 1**), so that there

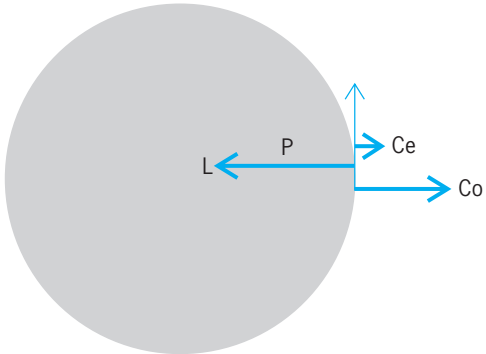

Fig. 1. Force balance for gradient flow about a cyclone in the Northern Hemisphere. Thin arrow indicates wind velocity; L = low-pressure center; Co = Coriolis force; Ce = centrifugal force; and P = pressure gradient force.

must be a pressure minimum at the center of the cyclone, and cyclones are sometimes simply called lows.

This three-way balance is referred to as the gradient wind. In large-scale extratropical cyclones the centrifugal force is of secondary importance. The balance is then primarily between the Coriolis force and the pressure gradient force. In this case the flow, which is parallel to the isobars with high pressure to the right of the direction of motion (in the Northern Hemisphere), is said to be in geostrophic balance. *See* AIR PRESSURE; CORIOLIS ACCELERATION; GRADIENT WIND; VORTEX.

Extratropical cyclones. These are the common weather disturbances which travel around the world from west to east in mid-latitudes. They are generally associated with fronts, which are zones of rapid transition in temperature. In a typical mid-latitude cyclone (**Fig. 2**) the air poleward of the frontal boundary is cool and dry, while the air equatorward is warm and moist. The poleward and upward motion of warm moist air south of the warm front, which extends eastward from the low-pressure center, is responsible for much of the precipitation associated with cyclonic storms. The equatorward and downward motion of cool dry air behind the cold front, which extends southwest of the low-pressure center, is responsible for the fine weather which generally follows the passage of a midlatitude cyclone.

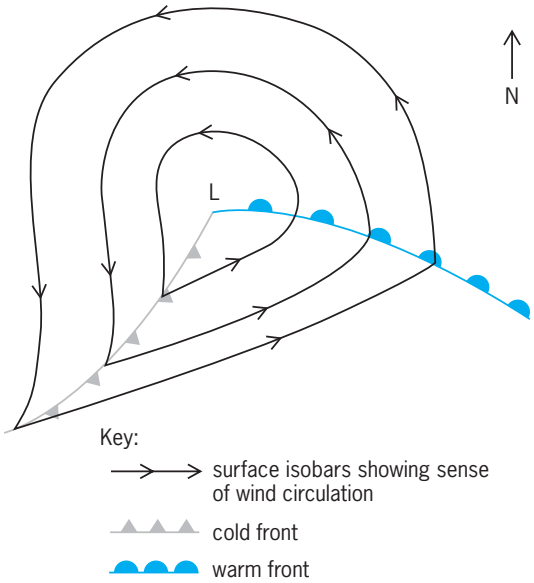

Fig. 2. Plane view of an extratropical cyclone in the Northern Hemisphere as depicted on a surface weather map. The L designates a pressure low.

Formation of extratropical cyclones is referred to as cyclogenesis. In the classic Norwegian cyclone model, which is favored by many weather forecasters, cyclogenesis begins with the development of a wavelike disturbance on a preexisting frontal discontinuity separating a cold polar air mass from a warm subtropical air mass (**Fig. 3***a*). The circulation associated with the wave then distorts the frontal surface to produce a closed vortical circulation with a warm front ahead of the vortex and a trailing cold front (Fig. 3*b*). In the later stages the cold front may overtake the warm front to produce an occlusion, a region where at the surface the warm air is limited to a narrow band (Fig. 3*c*).

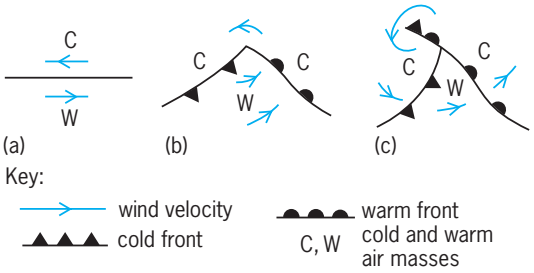

Fig. 3. Classic Norwegian cyclone model showing three stages in the life of a polar front cyclone. (*a*) Frontal discontinuity, (*b*) distorted frontal surface, (*c*) occlusion. (*After B. J. Hoskins, Dynamical processes in the atmosphere and the use of models, Quart. J. Roy. Meteorol. Soc., 109:1–21, 1983*)

A more modern theoretical model of cyclogenesis, on the other hand, suggests that most extratropical cyclones develop as dynamical instabilities in broad regions of enhanced horizontal temperature gradients and vertical shear of the geostrophic wind associated with the subtropical jet stream. In this instability process, which is referred to as baroclinic instability, the cyclone grows by converting potential energy associated with the horizontal temperature gradient in the mean flow into kinetic energy. As baroclinically unstable systems develop, the disturbance motion tends to intensify temperature contrasts near the surface so that warm and cold fronts develop. Thus, in this theory, fronts are considered to be secondary features produced by cyclogenesis, rather than acting as causes.

Cyclogenesis occurs over very broad areas of the Earth. However, it is most frequent over the western oceans (such as in the regions east of Japan and off

the east coast of North America) and in the lee of major mountain barriers (such as the Rockies and the Alps).

Occasionally, over maritime regions during the cold season, extratropical cyclones can develop rapidly into dangerously intense storms with winds comparable to those found in hurricanes. If decreases in surface pressure occur at a rate exceeding 1 millibar (100 pascals) per hour for a period of greater than 24 h, the development is called explosive cyclogenesis. Explosive cyclogenesis tends to occur in strong baroclinic zones when a weak surface disturbance receives reinforcement from a cyclonic disturbance aloft, and the static stability of the atmosphere (that is, the resistance to vertical motions) is weak so that strong cumulus convection occurs. *See* BAROCLINIC FIELD; FRONT; GEOSTROPHIC WIND; JET STREAM.

Tropical cyclones. By contrast, tropical cyclones derive their energy from the release of latent heat of condensation in precipitating cumulus clouds. The moisture necessary to sustain this precipitation is supplied by the frictionally driven flow of moist air toward low pressure near the Earth's surface and its subsequent ascent into the cumulus clouds. Over the tropical oceans, where moisture is plentiful, tropical cyclones can develop into intense vortical storms (hurricanes and typhoons), which can have wind speeds in excess of 200 mi/h ($100\ m \cdot s^{-1}$). *See* HURRICANE; PRECIPITATION (METEOROLOGY); STORM; TROPICAL METEOROLOGY; WIND.

J. R. Holton

Bibliography. J. R. Holton, *An Introduction to Dynamic Meteorology*, 3d ed., 1992; B. J. Hoskins, Dynamical processes in the atmosphere and the use of models, *Quart. J. Roy. Meteorol. Soc.*, 109:1–21, 1983; R. J. Reed and M. D. Albright, A case study of explosive cyclogenesis in the Eastern Pacific, *Month. Weath. Rev.*, 114:2297–2319, 1986; J. M. Wallace and P. V. Hobbs, *Atmospheric Science: An Introductory Survey*, 1977.

Cyclone furnace

A water-cooled horizontal cylinder in which fuel is fired and heat is released at extremely high rates. When firing coal, the crushed coal, approximately 95% of which is sized at 14 in. (0.6 cm) or less, is introduced tangentially into the burner at the front end of the cyclone (see **illus.**). About 15% of the combustion air is used as primary and tertiary air to impart a whirling motion to the particles of coal. The whirling, or centrifugal, action on the fuel is further increased by the tangential admission of high-velocity secondary air into the cyclone. *See* FURNACE CONSTRUCTION.

The combustible is burned from the fuel in the cyclone furnace, and the resulting high-temperature gases melt the ash into liquid slag, a thin layer of which adheres to the walls of the cyclone. The incoming fuel particles, except those fines burned in suspension, are thrown to the walls by centrifugal force and caught in the running slag. The secondary air sweeps past the coal particles embedded in the slag surface at high speed. Thus, the air required to burn the coal is quickly supplied and the products of combustion are rapidly removed.

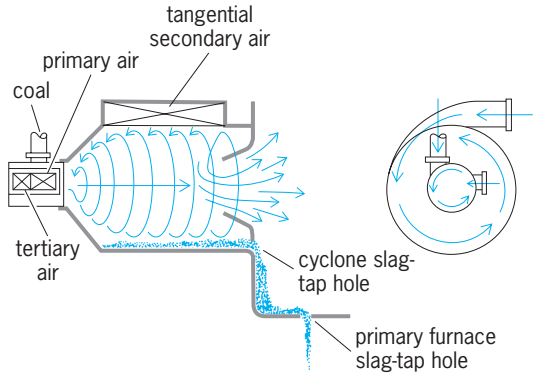

Schematic diagram of cyclone furnace. (*After E. A. Avallone and T. Baumeister III, eds., Marks' Standard Handbook for Mechanical Engineers, 9th ed., McGraw-Hill, 1987*)

The products of combustion are discharged through a water-cooled reentrant throat at the rear of the cyclone into the boiler furnace. The part of the molten slag that does not adhere to the cyclone walls also flows toward the rear and is discharged through a taphole into the boiler furnace.

Essentially, the fundamental difference between cyclone furnaces and pulverized coal–fired furnaces is the manner in which combustion takes place. In pulverized coal–fired furnaces, particles of coal move along with the gas stream; consequently, relatively large furnaces are required to complete the combustion of the suspended fuel. With cyclonic firing, the coal is held in the cyclone and the air is passed over the fuel. Thus, large quantities of fuel can be fired and combustion completed in a relatively small volume, and the boiler furnace is used to cool the products of combustion.

Gas and oil can also be burned in cyclone furnaces at ratings and with performances equaling those of coal firing. When oil is the fuel, it is injected adjacent to the secondary air ports and directed downward into the secondary airstream. The oil is picked up and sufficiently atomized by the high-velocity air. Gas is fired similarly through flat, open-ended ports located in the secondary air entrance. *See* BOILER; STEAM-GENERATING FURNACE; STEAM-GENERATING UNIT.

G. W. Kessler

Cyclophane

A molecule composed of two building blocks: an aromatic ring (most frequently a benzene ring) and an aliphatic unit forming a bridge between two (or more) positions of the aromatic ring. Structure (**1**),

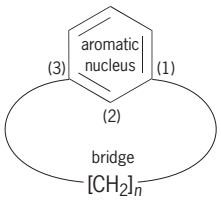

(**1**)

an [*n*](1,3)cyclophane, provides a general example for a simple cyclophane and its nomenclature. This cyclophane is also known as [*n*]metacyclophane. In structure (**1**), *n* is the number of [CH_2]-groups in the bridge, and (1) and (3) are the starting positions of the bridge.

Nomenclature. Those cyclophanes that contain heteroatoms in the aromatic part are generally called phanes, as "cyclo" in cyclophanes more strictly means that only benzene rings are present in addition to aliphatic bridges.

Phane molecules containing hetero atoms in the aromatic ring (for example, nitrogen or sulfur) are called heterophanes, while in heteraphanes they are part of the aliphatic bridge [structures (**2**) and (**3**)]. By coupling meta-xylylene dibromide with sodium (Wurtz coupling) the [2.2]metacyclophane (**4**) is ob-

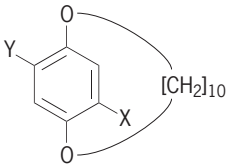

(**2**) X = Y = H
(**3**) X = CO_2H, Y = Br

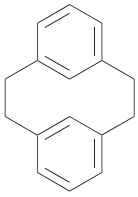

(**4**)

tained. The name is formed by putting the lengths (2 atoms) of the 2 bridges in square brackets, and the (1) and (3) positions at the benzene ring where the bridge starts and ends, in brackets. As the 1,3 positions are also called meta, this hydrocarbon (**4**) is usually named [2.2] metacyclophane. *See* ARTICLE: AROMATIC HYDROCARBON.

The first heteracyclic [*n*]paracyclophane (**2**) that was synthesized, the [12]paracyclophane, is called ansa compound [ansa (Latin), handle] because of its geometry. The tight embracing of structure (**2**) has stereochemical significance. Compound (**3**), with its hindered free rotation of the aliphatic bridge, exists as atropisomers (axial chirality) and can be separated into its stable enantiomers.

Synthesis. Modern syntheses of cyclophanes began with the first synthesis of [2.2]paracyclophane (**5**). The hydrocarbon was obtained by extraction

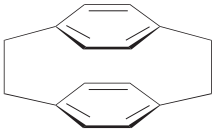

(**5**)

from pyrolysis products of *p*-xylene.

The second major advance in cyclophane chemistry resulted from the work of D. J. Cram, the first directed synthesis of [2.2], [3.2], and [4.2]paracyclophanes (**7**)–(**9**) that used a high-dilution intramolecular Wurtz reaction with molten sodium in xylene starting from bis(brommethyl)bibenzyl (**6**), as shown in reaction (1). The Wurtz coupling is a radical re-

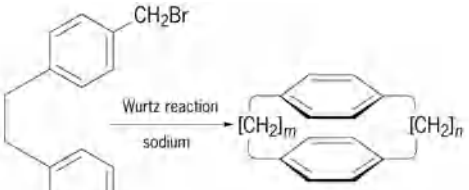

(1)

(**6**)

	m	n
(**7**)	2	2
(**8**)	2	3
(**9**)	2	4

action induced by molten sodium forming carbon-carbon single bonds starting from alkyl halides (for example, **6**). It is necessary to use a low concentration of starting materials in a high volume of solvent (high dilution)to suppress intermolecular side reactions and to obtain mainly the intramolecular coupling shown in reaction (1).

In the 1970s, new synthetic methods of forming highly strained cyclophanes became available. One type, sulphone pyrolysis (also known as sulfur dioxide extrusion), can be applied to the synthesis of many cyclophanes, and it has become one of the most widely applicable routes [reaction (2)]. By

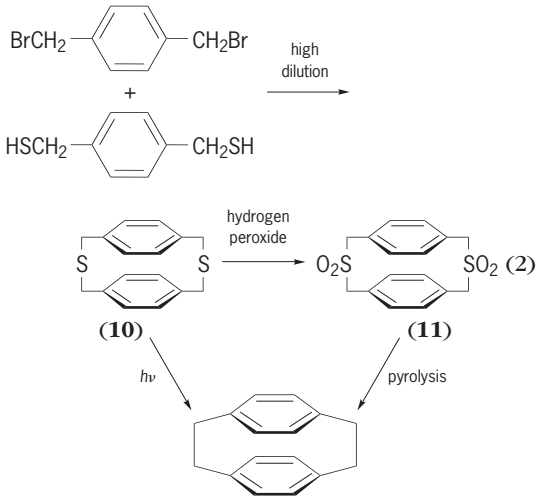

(**5**)

eliminating sulfur photochemically or by pyrolysis, strain is introduced into the ring system starting from molecules like (**10**) and (**11**). Compound (**5**) can be obtained in high yield by this synthetic approach. *See* ARTICLE: ORGANIC SYNTHESIS.

Properties. Cyclophanes, with their remarkable structure, often exhibit peculiarities in electronic and spectroscopic properties, the investigation

of which is of fundamental scientific value. For example: Cyclophanes usually are not at all planar molecules; they exhibit an interesting stereochemistry (arrangement of atoms in three-dimensional space); molecular parts are placed and sometimes fixed in unusual orientations toward each other, and often the molecules are not rigid but (conformationally) flexible. Ring strain and, as a consequence, deformation of the benzene rings out of planarity are often encountered. Electronic interactions between aromatic rings fixed face to face [like the two benzene rings in (**5**)] can take place. In addition, influence on substitution reactions in the aromatic rings results; that is, substitutents in one ring [for example, of (**5**)] induce transannular electronic effects in the other ring, often leading to unexpected products. *See* CONFORMATIONAL ANALYSIS; STEREOCHEMISTRY.

Aromatic rings in many cyclophanes are not planar but distorted through bending (referred to as bent and battered benzene rings). It has been shown by x-ray crystal structure analysis that benzene rings can be distorted into boat, chair, and twist forms by clamping them in cyclophane structures by short bridges. These stereochemical aberrations result in spectroscopic and chemical consequences that are of fundamental interest to all chemists, particularly in connection with aromatic units such as benzene, naphthalene, azulene, or pyridine. *See* BENZENE.

Tightly clamped cyclophanes are usually highly strained. The ring strain reaches an extreme in highly deformed molecules such as [6]paracyclophane (**12**). This molecule is the smallest stable

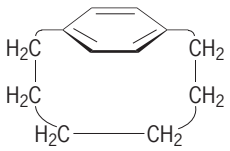

(**12**)

member of [*n*]paracyclophanes that can be isolated. In ring-substituted [6]paracyclophanes, it was found by x-ray crystal structure analysis that the distortion angle is up to 22.4° and the bond angle in the aliphatic bridge was expanded to 126.5° (the ideal value in an unstrained molecule is 109.5°). Some of the bridging protons in the [6]paracyclophane are located above the plane of the aromatic ring and therefore shifted strongly upfield to δ[ppm]: −0.6. In ^{1}H-NMR spectroscopy, the aromatic protons generally appear in a region of δ[ppm]: 7–9. Some bridging protons of the [6]paracyclophane are oriented above the aromatic ring, and as a consequence these protons are strongly shifted upfield to high magnetic field strength to δ[ppm]: −0.6 (beyond the internal standard substance tetramethylsilane, which by definition equals 0) influenced by the aromatic π system. Cyclophane chemistry provides the opportunity to gradually increase the distortions and to measure the resulting effects. *See* X-RAY CRYSTALLOGRAPHY.

Benzene rings can be arranged into multilayer cyclophanes, and transannular electronic effects through more than two aromatic rings can be detected spectroscopically. Furthermore, cyclophane molecules can orient themselves as planar-chiral and helical arrangements; and the resulting enantiomers, once separated, can be examined with respect to their properties toward circularly polarized light.

Cyclophane chemistry, with the three-dimensional geometry of its molecules, has revealed many unusual molecules. It is the wide variety of arrangements of partial structures, fixed or conformationally flexible, that affords such a rich field for research into molecular design, synthesis, structural analysis, physical investigations, chemical reactions, and all the way to supramolecular chemistry. *See* DENDRITIC MACROMOLECULE.

Configuration. Cyclophanes have been synthesized in both fairly simple structures and more elaborate structures. An interesting representative of the cyclophane family is the rather simple hydrocarbon [2.2]metaparacyclophane (**13**), which shows a dy-

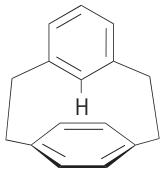

(**13**)

namic flexibility of the meta-substituted benzene ring that can be studied by nuclear magnetic resonance spectroscopy. The compound [2.2]anthracenophane (**14**) can easily undergo a smooth, photo-

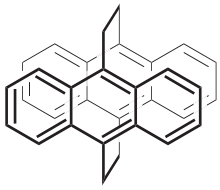

(**14**)

chemically induced cyclization to yield a cage compound. Completely bridged aromatic units lead to [2_6](1,2,3,4,5,6)cyclophane (**15**), also called super-

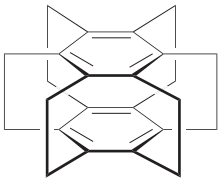

(**15**)

phane, which was synthesized by stepwise formation of the bridges. In sharp contrast to the strained

cyclophanes, which are conspicuous for their face-to-face and bent-and-battered benzene rings, the two benzene rings in superphane (**15**) are planar and undistorted.

Interest in calixarenes like structure (**16**) stems

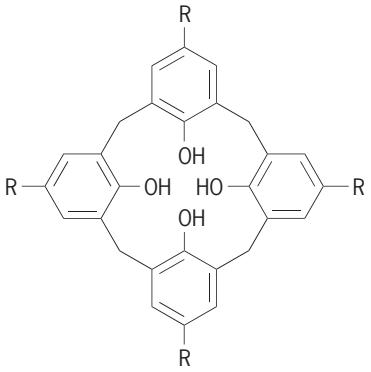

(**16**)

from their basketlike shape being adapted to host-guest or receptor-substrate chemistry and the fact that they are easily synthesized. Some members are commercially available, and have been shown to function as hydrolysis catalysts and oxygen carriers. They were named calixarenes because of the chalice-like shape of the molecules and the aromatic (arene) components of the macrocycle. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

As the number of benzene ring layers is increased is multilayered cyclophanes, the ultraviolet absorption of the donor-acceptor complexes is shifted to longer wavelengths, indicating an increasing donor character of the cyclophane. These spectral characteristics are explained as arising from electronic interaction between the benzene rings. With increasing donor character, the stability of the charge-transfer complex increases. This enhanced π-donor effect is still noticeable in the quintuple-layered cyclophane complex (**17**), with tetracyanoethene as

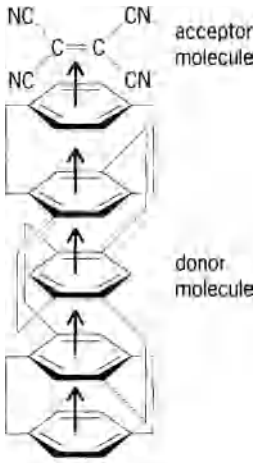

(**17**)

the acceptor molecule. There is no more additional difference seen in the transannular electron-donor capacity in the corresponding sextuple-layered cyclophane. *See* CHEMICAL BONDING.

The quinone-capped porphyrinophane (**18**) is of

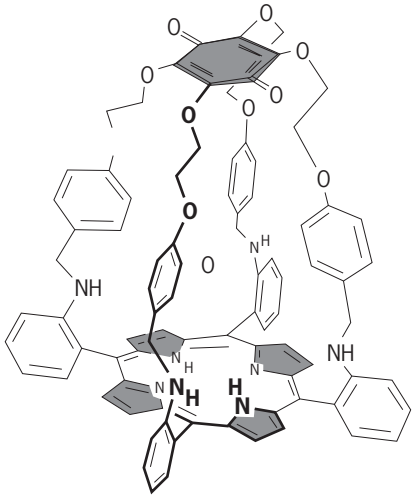

(**18**)

interest as a model compound in connection with biological charge-transfer interactions and provides an example for surprisingly high yields in macrocyclization reactions involving rigid subunits and reversible reaction types. The spheriphane (**19**; $C_{36}H_{36}$), with

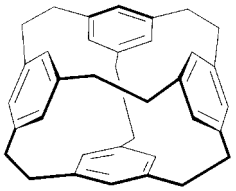

(**19**)

its four tetrahedrally arranged benzene rings, constitutes a concave hydrocarbon and a cage molecule with its possibility for encapsulating and transporting small molecules or ions. *See* CAGE HYDROCARBON.

Molecular recognition. Cyclophanes have become important in host-guest chemistry and supramolecular chemistry as they constitute host compounds for guest particles because of their cavities. Recognition processes at the molecular level are understood to mean the ability to design host molecules to encompass or attach selectively to smaller, sterically complementary guests in solution—by analogy to biological receptors and enzymes. For this reason, water-soluble cyclophanes like host (**20**) were synthesized; they have the ability to include nonpolar guest molecules (for example, tetramethylbenzene) in aqueous solution. The stoichiometric inclusion was proven by x-ray crystal structure analysis. *See* MOLECULAR RECOGNITION; POLAR MOLECULE.

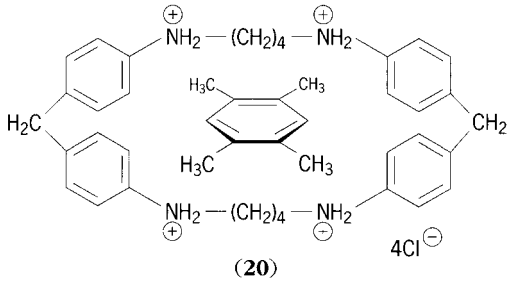

(**20**)

Applications. Cyclophanes also exist in nature as alkaloids, cytotoxic agents, antibiotics, and many cyclopeptides. With the possibility of locating groups precisely in space, cyclophane chemistry has provided building units for molecular niches, nests, hollow cavities, multifloor structures, helices, macropolycycles, macro-hollow tubes, novel ligand systems, and so on. Cyclophane chemistry has become a major component of supramolecular chemistry, molecular recognition, models for intercalation, building blocks for organic catalysts, receptor models, and variation of crown ethers and cryptands. Apart from the various research possibilities and their role as cyclo-oligomers for high-molecular-weight compounds, some cyclophanes find application in the high-tech field. Examples are the [2.2.2]paracyclophane, a cyclophane that exhibits a remarkable high and selective extraction capability for silver ions, which can be extracted from a water phase into a lipophilic solvent; and [2.2]paracyclophane, which has been used for the mechanistic investigation of catalytic hydrogenation, a very important industrial process. *See* MACROCYCLIC COMPOUND; SUPRAMOLECULAR CHEMISTRY.

F. Vögtle; M. Böhme

Bibliography. F. Diederich, *Cyclophanes*, 1991; P. M. Keehn and S. M. Rosenfeld (eds.), *Cyclophanes*, vols. 1 and 2, 1983; F. Vögtle, *Cyclophane Chemistry*, 1993; F. Vögtle, *Fascinating Molecules in Organic Chemistry*, 1992.

Cyclopoida

An order of small copepod Crustacea. The abundant fresh-water and salt-water species form an important link in the food chains of aquatic life, consuming tiny plants, animals, and detritus and, in turn, furnishing food for small fish, some large fish, and other organisms. Some species are important intermediate hosts for human parasites. Females of the family Ergasilidae are parasitic on aquatic animals, while the males are free-swimming. *See* FOOD WEB.

The front part of the body is oval and sharply separated from the tubular hind end, which bears two caudal rami with distinctly unequal setae (**Fig. 1**). Usually 10 body segments are present in the male, and 9 in the female because of fusion of two to form a genital segment. Appendages (**Fig. 2**) are the first and second antennae, mandibles, first and second maxillae, maxillipeds, and five pairs of swimming legs, the last of which are alike and rudimentary. Males usually have rudimentary sixth legs. Appendages are usually biramous, with a 1- or 2-segmented basal portion. One branch is missing from the first antennae, and usually from the second antennae, the mandibles, and fifth legs. The first antennae are never much longer than the anterior portion of the body and consist of 6–17 segments. Usually there are either 12 or 17. Both first antennae of the males are modified for grasping females. The digestive tract consists of a narrow, dilatable esophagus, an expanded stomach, a long intestine, and a short rectum opening on the dorsal side of the last body segment.

A heart is absent. Colorless blood containing ameboid corpuscles is circulated by general body movements and rhythmic movements of the digestive tract. External respiration is through the body surface. Striated muscles are highly developed. The genital systems are shown in Fig. 1. In mating, the male grasps the female with the first antennae and by means of the swimming legs places two kidney-shaped spermatophores at the opening of the female's seminal receptacle. Egg sacs are paired and lateral. Rarely, they are subdorsally located. There are six nauplius and six copepodid stages, the last being the adult. Some species encyst in the copepodid stages to survive adverse conditions. A small brain in the head is connected, around each side of the esophagus, to a ventral chain of ganglia. There is a median dorsal eye. Cyclopoids are intermediate hosts of the parasitic guinea worm (*Dracunculus medinensis*), and sometimes the fish tapeworm

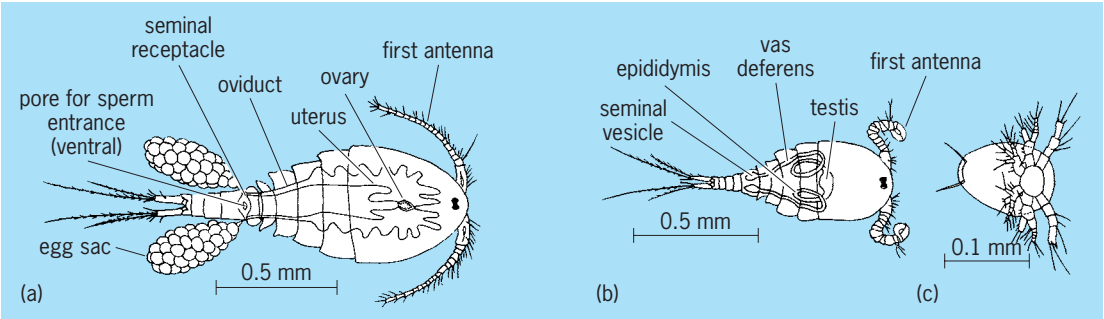

Fig. 1. ***Cyclops vernalis*. (*a*) Dorsal view of female. (*b*) Dorsal view of male. (*c*) Ventral view of the first of the six nauplius stages.**

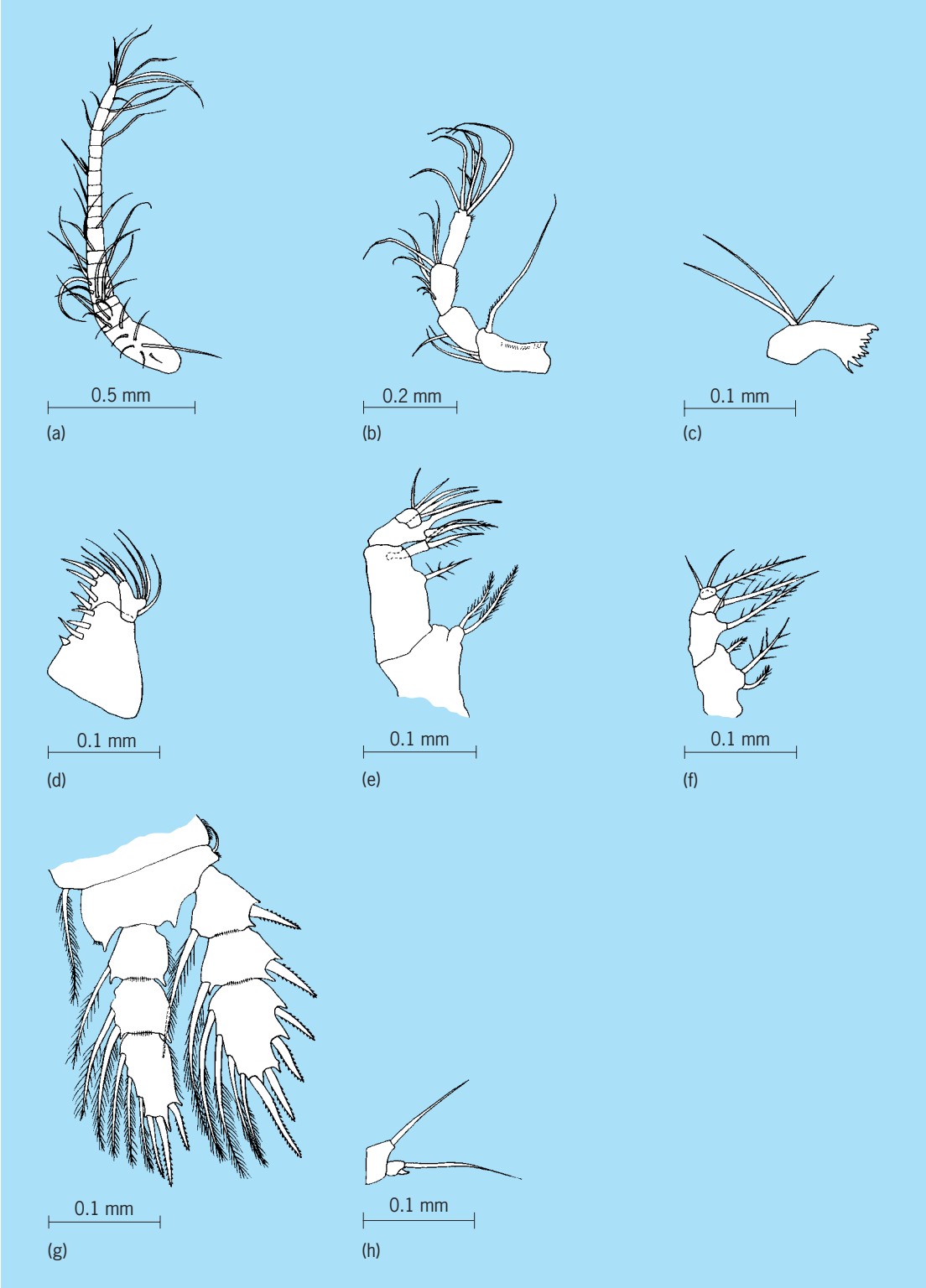

Fig. 2. Representative appendages of *Cyclops*. (*a*) First antenna of female. (*b*) Second antenna. (*c*) Mandible. (*d*) First maxilla. (*e*) Second maxilla. (*f*) Maxilliped. (*g*) Typical swimming leg (right third leg). (*h*) Fifth leg.

(*Dibothriocephalus latus*). Most fresh-water species live in shallow water, swimming from plant to plant, but salt-water species are generally water-treaders. Food is not secured by filtration, but is seized and eaten with the biting mouthparts. Many species have a worldwide distribution. *See* COPEPODA.

Harry C. Yeatman

Bibliography. B. Dussart, *Les Copepodes des Eaux Continentales*, vol. 2, 1969; R. Gurney, *British Fresh-water Copepoda*, vol. 3, 1933; F. Kiefer, *Ruderfuss-Krebse* (*Copepoden*), 1960; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; G. O. Sars, *Crustacea of Norway*, vol. 6, 1918.

Cyclostomata (Bryozoa)

An order of bryozoans in the class Stenolaemata. Cyclostomes tend to have delicate colonies composed of relatively isolated, loosely bundled or tightly packed, comparatively simple, long, slender, and tubular zooecia (individual living chambers), with thin, highly porous, calcareous walls. When regionated, their colonies display a very gradual transition between rather indistinct endozones and exozones. *See* BRYOZOA; STENOLAEMATA.

Morphology. Cyclostome colonies commonly are small and delicate, but some are moderately large and massive. They may be inconspicuous encrusting threadlike networks (*Corynotrypa*, *Hederella*), delicate prostrate branches (*Tubulipora*, *Diaperoecia*), thin encrusting disks (*Lichenopora*) or sheets (*Berenicea*), nodular masses (*Ceriopora*), or erect tuftlike (*Crisia*), twiglike (*Tretocycloecia*), or frondlike (*Hornera*) growths. Cyclostome colonies may sometimes bear rhizoids (rootlike structures that reinforce the colony's attachment to a substrate), nanozooecia (dwarf zooecia), mesopores (tiny pores), and various small coelomic chambers, but lack many of the specialized morphologic structures characteristic of other bryozoan orders. Many colonies also display well-developed ovicells (brood chambers) that have evolved by complete modification of initially normal-appearing zooids (gonozooids).

Individual cyclostome zooecia are straight to slightly curved long tubes, each containing a polyp with tentacles (see **illustration**). The walls of adjacent zooecia may be either distinctly separate or fused together. The zooecial walls, generally thin throughout the colony and having finely granular or finely laminated microstructure, consist of calcareous material (calcite) perforated by many tiny pores which are either open (interzooidal or communication pores) or filled (in living forms) with soft tissues (pseudopores). In addition, some species may possess a chitinous cuticle covering the external surface of the outermost zooecial walls in the colony. The space within each zooecium usually lacks transverse calcareous partitions, but sometimes is crossed by a few diaphragms. Round or circular in outline, the zooecial aperture lacks an operculum, but in some forms is surrounded by a raised rim (peristome). Cyclostome zooids lack epistomes, have few to moderate numbers of tentacles, and (like the colony overall) can be either hermaphroditic or dioecious.

Life cycle. Immediately after fertilization, the cyclostome zygote enters an ovicell and divides into as many as 100 embryos (polyembryony), each of which develops into a larva. After escaping from the ovicell, the larva swims freely for less than a day before settling down to the bottom and attaching to a substrate. There, the larva metamorphoses (by means of complete degeneration of its internal structures) into an ancestrular (primary) zooid from which repeated asexual budding produces the rest of the colony (but no resistant resting bodies).

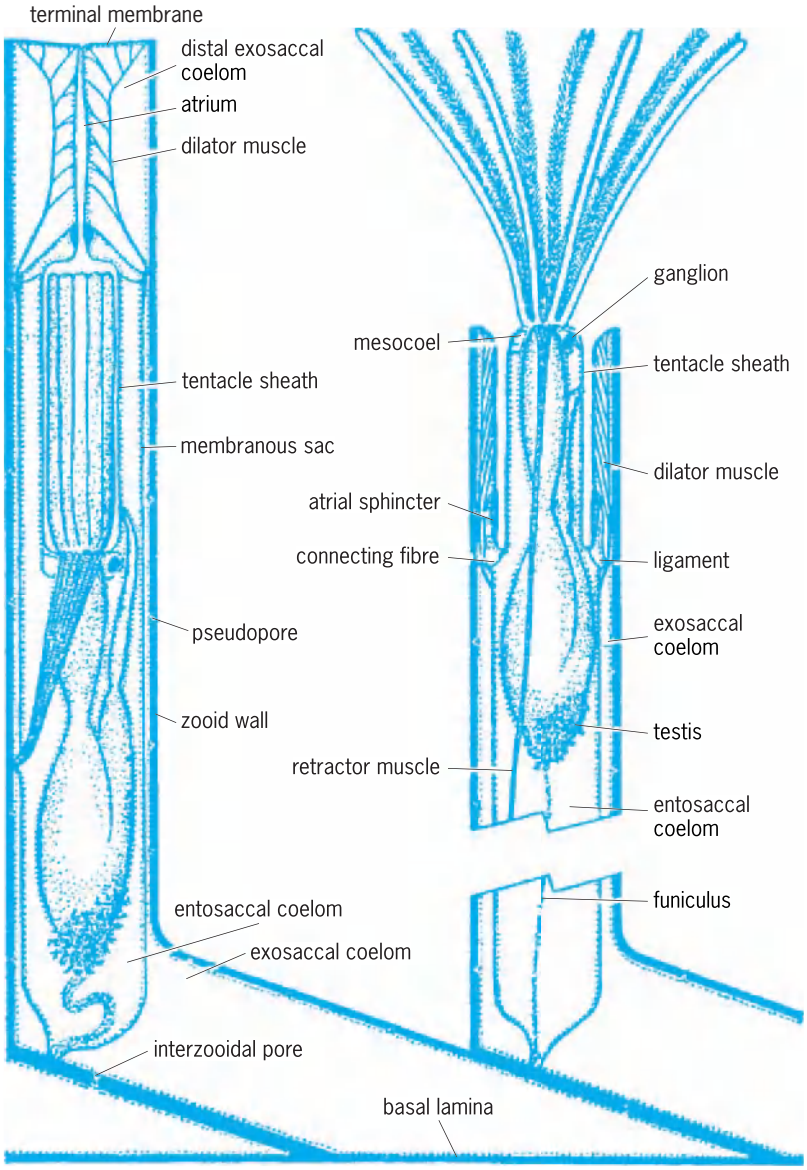

Cyclostome zooids with tentacles retracted versus extended. (*Reprinted with permission from R. J. Cuffey and J. E. Utgaard, Bryozoans, in R. Singer, ed., Encyclopedia of Paleontology, Fitzroy Dearborn Publishers, 1999*)

History and classification. Exclusively marine, cyclostomes (recently termed Tubuliporata by some) are known first from the Early Ordovician, when they may have evolved from ctenostomes, and possibly gave rise to the other Paleozoic stenolaemate orders. Rare and inconspicuous throughout the rest of the Paleozoic and Triassic, cyclostomes became moderately common in Jurassic and Cretaceous seas, but have since (including today) been less numerous. However, cyclostomes are of great scientific interest because, as the only living stenolaemates, they can shed much light on related extinct forms. *See* CTENOSTOMATA.

Roger J. Cuffey

Bibliography. R. J. Cuffey and J. E. Utgaard, Bryozoans, in R. Singer (ed.), *Encyclopedia of Paleontology*, vol. 1, pp. 204–216, Fitzroy Dearborn, Chicago, 1999; P. J. Hayward and J. S. Ryland, *Cyclostome*

Bryozoans, Linnean Society & Brill/Backhuys, London, 1985; F. K. McKinney and J. B. C. Jackson, *Bryozoan Evolution*, Unwin Hyman, Boston, 1989; P. D. Taylor, Bryozoa, in M. J. Benton (ed.), *The Fossil Record 2*, pp. 465–488, Chapman & Hall, London, 1993.

Cyclothem

A vertical sequence of several different kinds of distinctive sedimentary rock units that is repeated upward through the stratigraphic succession. Originally defined in the rock succession of Pennsylvanian age in the Illinois Basin in the 1930s, the rock types include coal, limestone, sandstone, and several types of shale and mudstone. Cyclothems were soon recognized elsewhere in rocks of this age in the central and eastern United States. Those in the Midcontinent (Kansas and states to the northeast) are dominated by several types of limestone and shale, with less coal and sandstone. Those in the Appalachian region are dominated by coal, mudstone, shale, and sandstone, with less limestone. Some workers generalized the term, and simple cyclothems became recognized in coal-bearing successions of other ages, but the term never achieved wide usage outside of rocks of Pennsylvanian age. *See* PENNSYLVANIAN; SEDIMENTARY ROCKS; STRATIGRAPHY.

The original definition stated that this sequence of rocks was deposited during a single sedimentary cycle of the type that prevailed during the Pennsylvanian Period, thus implying a strong genetic relation among such sequences in strata of this particular age. Three major types of origin were proposed: (1) periodic uplift (tectonism) of the active Appalachian Mountains followed by subsidence and encroachment of the sea; (2) periodic rise and fall of sea level (eustacy) driven by repeated episodes of continental glaciation in the southern continents of that time; and (3) shifting of deltas when flooding broke the levee and established a new delta in a different but overlapping area. The most recent work on Pennsylvanian cyclothems (when complex cyclothems were best developed) has resulted in the recognition of well-developed widespread paleosols (ancient soils) beneath the coals nearly everywhere and the widespread lateral continuity of the marine limestones and shales in the Midcontinent. This, in conjunction with the absence of deltas in most of that region, has ruled out delta-shifting as the basic cause. The estimates of rather great water depth attained during marine inundation in the Midcontinent (which was lower on the ancient shelf) along with estimates of greater frequency than is

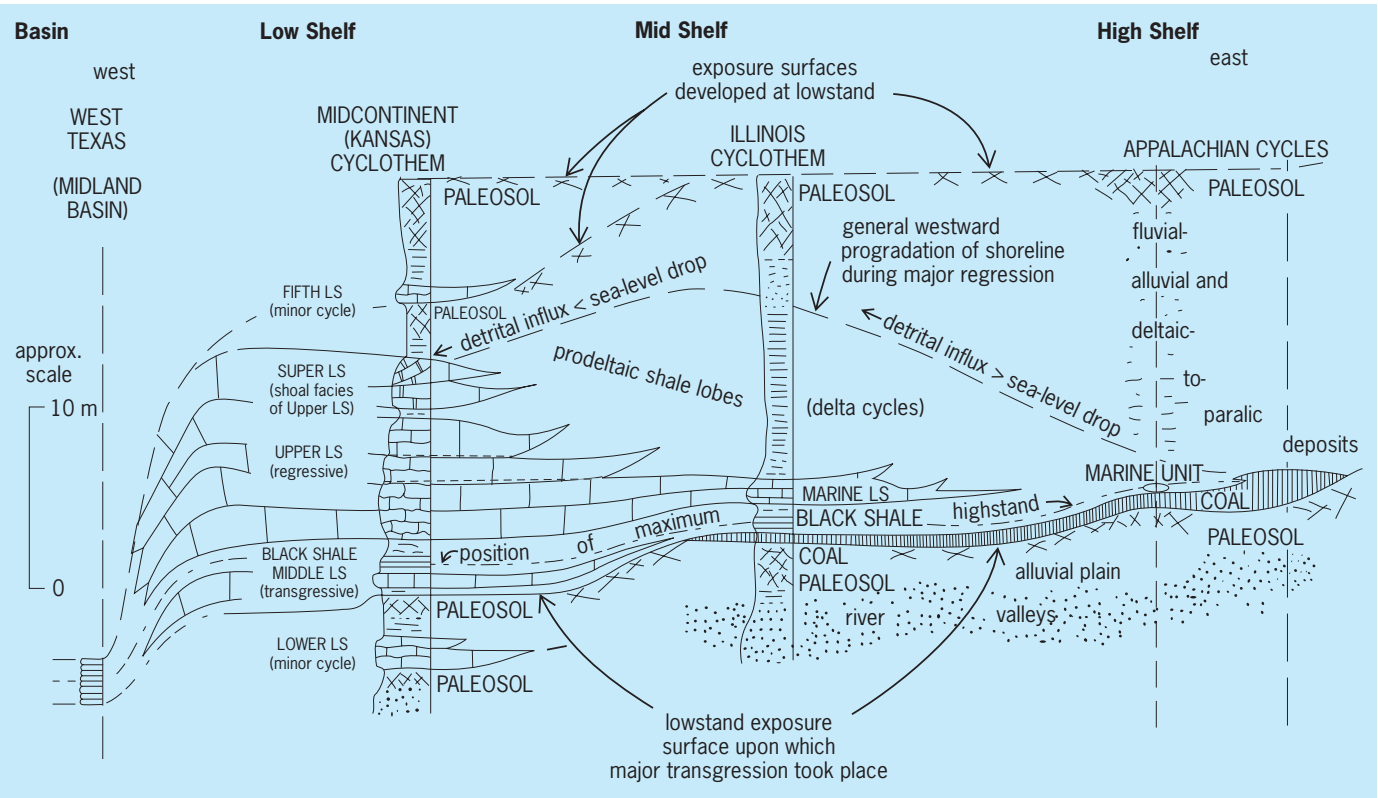

Restored cross section showing relations of cyclothems in Midcontinent, Illinois, and Appalachian regions. LS = Low Shelf. (*After P. H. Heckel, Evaluation of evidence for glacio-eustatic control over marine Pennsylvanian cyclothems in North America and consideration of possible tectonic effects, in Tectonic and Eustatic Controls on Sedimentary Cycles, ed. by J. M. Dennison and F. R. Ettensohn, SEPM Concepts in Sedimentology and Paleontology, 4:65–87, 1994; and P. H. Heckel, Glacial-eustatic base-level–climatic model for late Middle to Late Pennsylvanian coal-bed formation in the Appalachian Basin, J. Sed. Res., B65:348–356, 1995*)

likely with tectonic fluctuations of sea level have ruled out tectonism as a general cause. Estimates of cyclothem frequency and magnitude of water-depth changes are both within the ranges detected for the well-known Pleistocene continental glaciations, and thus further support the glacial eustatic origin for cyclothems. For this reason (combined with the specific original definition), the term is probably most usefully applied only to glacially driven repeating sequences of rock units. This would include strata from latest Mississippian through Pennsylvanian to mid-Permian age (most of late Paleozoic time), and the term could apply to similar genetic sequences of rock in glacially influenced late Cenozoic time, which includes the Pleistocene ice ages and modern times. *See* CENOZOIC; MISSISSIPPIAN; PALEOSOL; PERMIAN; PLEISTOCENE.

The sequence of events that explains Pennsylvanian cyclothems of the central and eastern United States is shown in the **illustration**. At sea-level lowstand, sandstones were deposited in ancient valleys while soils formed over the exposed land surface on the areas between the rivers. As the glaciers melted, sea level rose causing coal swamps to form in poorly drained areas landward of the encroaching sea. These were followed by shallow-water marine limestone, which in the Midcontinent was followed by deep-water black phosphatic shale deposited when the sediment surface passed below the photic zone in this deeper area, and sea-circulation changes diminished bottom oxygen and concentrated phosphate. In the shallower Appalachian region, the marine limestone was the deepest water deposit. Sea-level fall then caused outbuilding of deltas in the Appalachian region, followed by subaerial exposure and development of paleosols, but caused formation of shoaling-upward limestone above the black shale in the Midcontinent, an area that was not affected by delta outbuilding or subaerial exposure until later.

Cyclothems were first noticed in the United States because of the many distinctive rock types involved. These resulted from the interplay between the then tropical to equatorial humid climate with ready access to detrital sand and mud and the marine circulation changes caused by the changing water depth of the sea. Coal-rich cyclothems were soon recognized in strata of the same age in western Europe. Currently, cyclothems without coal are being identified in late Paleozoic strata elsewhere in the world by the recognition of thin paleosols and exposure surfaces in more monotonous limestone successions in the more subtropical regions of that time, where the climate was drier and lithic differences not as conspicuous. Complex cyclothems are also being recognized in repeating sequences of entirely subtidal marine rock units in late Paleozoic pelagic realms. *See* PALEOZOIC.

Philip H. Heckel

Bibliography. P. H. Heckel, Evaluation of evidence for glacio-eustatic control over marine Pennsylvanian cyclothems in North America and consideration of possible tectonic effects, in Tectonic and Eustatic Controls on Sedimentary Cycles, ed. by J. M. Dennison and F. R. Ettensohn, *SEPM Concepts in Sedimentology and Paleontology*, 4:65–87, 1994; P. H. Heckel, Glacial-eustatic base-level–climatic model for late Middle to Late Pennsylvanian coal-bed formation in the Appalachian Basin, *J. Sed. Res.*, B65:348–356, 1995; P. H. Heckel, Sea-level curve for Pennsylvanian eustatic marine transgressive-regressive depositional cycles along midcontinent outcrop belt, North America, *Geology*, 14:330–334, 1986.

Cyclotron resonance experiments

The measurement of charge-to-mass ratios of electrically charged particles from the frequency of their helical motion in a magnetic field. Such experiments are particularly useful in the case of conducting crystals, such as semiconductors and metals, in which the motions of electrons and holes are strongly influenced by the periodic potential of the lattice through which they move. Under such circumstances the electrical carriers often have "effective masses" which differ greatly from the mass in free space; the effective mass is often different for motion in different directions in the crystal. Cyclotron resonance is also observed in gaseous plasma discharges and is the basis for a class of particle accelerators. *See* BAND THEORY OF SOLIDS; PARTICLE ACCELERATOR.

The experiment is typically performed by placing the conducting medium in a uniform magnetic field H and irradiating it with electromagnetic waves of frequency ν. Selective absorption of the radiation is observed when the resonance condition $\nu = qH/2\pi m^* c$ is fulfilled, that is, when the radiation frequency equals the orbital frequency of motion of the particles of charge q and effective mass m^* (c is the velocity of light). The absorption results from the acceleration of the orbital motion by means of the electric field of the radiation. If circularly polarized radiation is used, the sign of the electric charge may be determined, a point of interest in crystals in which conduction may occur by means of either negatively charged electrons or positively charged holes. *See* HOLE STATES IN SOLIDS.

For the resonance to be well defined, it is necessary that the mobile carriers complete at least $^1/_2\pi$ cycle of their cyclotron motion before being scattered from impurities or thermal vibrations of the crystal lattice. In practice, the experiment is usually performed in magnetic fields of 1000 to 100,000 oersteds (1 oersted = 79.6 amperes per meter) in order to make the cyclotron motion quite rapid ($\nu \approx$ 10–100 gigahertz, that is, microwave and millimeter-wave ranges). Nevertheless, crystals with impurity concentrations of a few parts per million or less are required and must be observed at temperatures as low as 1 K in order to detect sharp and intense cyclotron resonances.

The resonance process manifests itself rather differently in semiconductors than in metals. Pure, very cold semiconductors have very few charge carriers; thus the microwave radiation penetrates the sample uniformly. The mobile charges are thus exposed to

radiation over their entire orbits, and the resonance is a simple symmetrical absorption peak.

In metals, however, the very high density of conduction electrons present at all temperatures prevents penetration of the electromagnetic energy except for a thin surface region, the skin depth, where intense shielding currents flow. Cyclotron resonance is then observed most readily when the magnetic field is accurately parallel to the flat surface of the metal. Those conduction electrons (or holes) whose orbits pass through the skin depth without colliding with the surface receive a succession of pulsed excitations, like those produced in a particle accelerator. Under these circumstances cyclotron resonance consists of a series of resonances $n\nu = qH/2\pi m^*c$ ($n = 1, 2, 3, \ldots$) whose actual shapes may be quite complicated. The resonance can, however, also be observed with the magnetic field normal to the metal surface; it is in this geometry that circularly polarized exciting radiation can be applied to charge carriers even in a metal.

Cyclotron resonance is most easily understood as the response of an individual charged particle; but, in practice, the phenomenon involves excitation of large numbers of such particles. Their net response to the electromagnetic radiation may significantly affect the overall dielectric behavior of the material in which they move. Thus, a variety of new wave propagation mechanisms may be observed which are associated with the cyclotron motion, in which electromagnetic energy is carried through the solid by the spiraling carriers. These collective excitations are generally referred to as plasma waves. In general, for a fixed input frequency, the plasma waves are observed to travel through the conducting solid at magnetic fields higher than those required for cyclotron resonance. The most easily observed of these excitations is a circularly polarized wave, known as a helicon, which travels along the magnetic field lines. It has an analog in the ionospheric plasma, known as the whistler mode and frequently detected as radio interference. There is, in fact, a fairly complete correspondence between the resonances and waves observed in conducting solids and in gaseous plasmas. Cyclotron resonance is more easily observed in such low-density systems since collisions are much less frequent there than in solids. In such systems the resonance process offers a means of transferring large amounts of energy to the mobile ions, a necessary condition if nuclear fusion reactions are to occur. *See* NUCLEAR FUSION; PLASMA (PHYSICS). Walter M. Walsh

Bibliography. C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996.

Cydippida

The largest order of the phylum Ctenophora, comprising 5 families (Bathyctenidae, Haeckelidae, Lampeidae, Mertensidae, and Pleurobrachiidae) and 11 genera. The body of cydippids is usually globular or cylindrical in shape (see **illustration**); sizes range from a few millimeters to about 15 cm (6 in.). Most cydippids are colorless and transparent, but some deep-sea species are pigmented dark red. The comb rows are of equal length, and the eight meridional canals lying beneath them end blindly at the oral end. Adult cydippids retain the larval morphology common to all ctenophores (except Beroida) and are usually thought to be the most primitive order. All forms have two main tentacles, which have side branches (tentilla) in some species. Cydippids catch prey on the outstretched tentacles, contracting them to bring it to the mouth. In three families (Bathyctenidae, Haeckelidae, and Lampeidae) the tentacles emerge from the body near the oral end; these cydippids have extensile mouths and catch large gelatinous animals. In the Pleurobrachiidae and Mertensidae, the tentacles exit near the aboral end, have numerous side branches, and catch mostly small crustacean prey. The well-known genus *Pleurobrachia* is widely distributed in temperate coastal waters and estuaries, where it can be a major predator of copepods and larval fish. Many other species of cydippids occur in oceanic waters, from the surface to mesopelagic depths. *See* BEROIDA; CTENOPHORA.

Laurence P. Madin

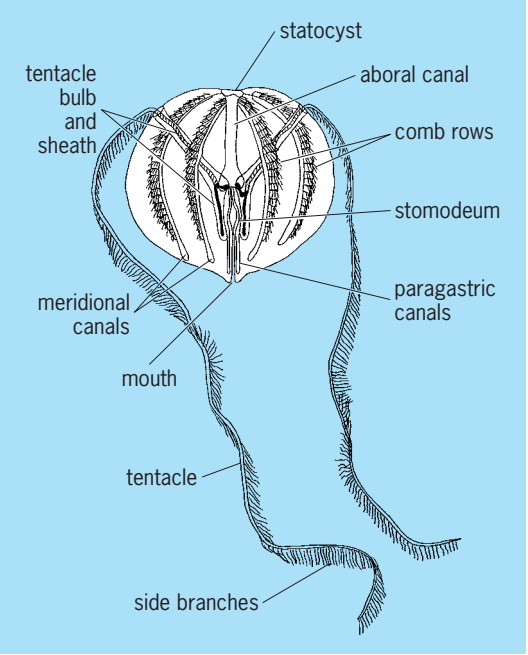

Typical *Pleurobrachia* sp.

Cygnus

The Swan, a large constellation lying prominently along the northern Milky Way (see **illustration**). It contains the stars that define the asterism (star pattern) known as the Northern Cross. Near Deneb, its brightest star and the nineteenth brightest star in the sky, the North America Nebula shows its glowing gas and dark outlining dust. Deneb marks the swan's tail while Albireo, a double star whose components beautifully contrast in color, marks the swan's head.

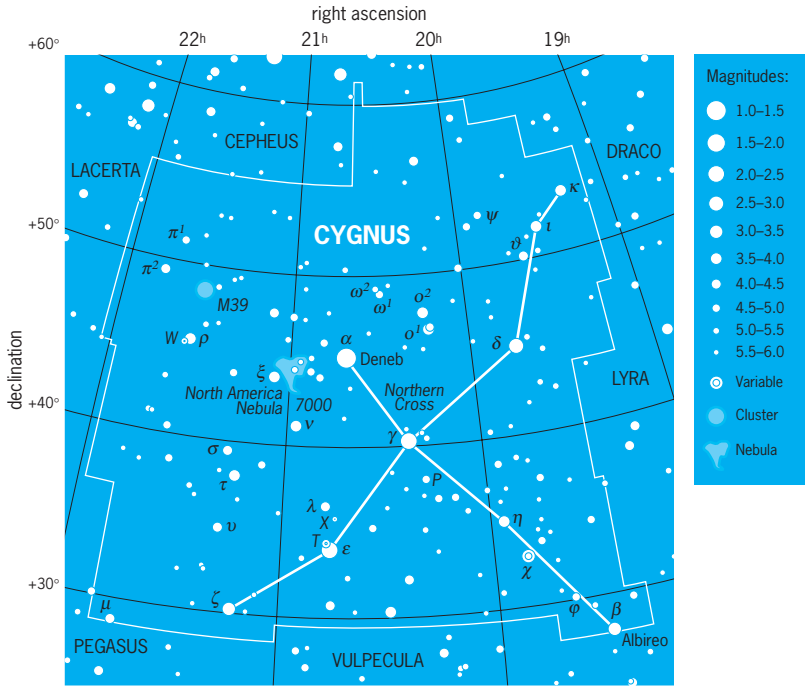

Modern boundaries of the constellation Cygnus, the Swan. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

Deneb in Cygnus, Vega in Lyra, and Altair in Aquila make the "summer triangle." *See* LYRA.

In Greek mythology, Phaethon, the son of the sun god, was allowed to drive the chariot of the Sun across the sky, but lost control. Zeus saved Earth from disaster by hurling a thunderbolt at Phaethon, who fell into the Eridanus River. Eridanus is now another constellation. A devoted friend of Phaethon dived into the water to search for his friend. Zeus changed him into the swan, changed his name to Cygnus, and elevated him to the heavens.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Cylinder

The solid of revolution obtained by revolving a rectangle about a side is called a cylinder, or more precisely a right circular cylinder. More generally the word cylinder is used in solid geometry to describe any solid bounded by two parallel planes and a portion of a closed cylindrical surface. In analytic geometry, however, the word cylinder refers not to a solid but to a cylindrical surface (see **illus.**). This is a surface generated by a straight line which moves so that it always intersects a given plane curve called the directrix, and remains parallel to a fixed line that intersects the plane of the directrix. (Part or all of the directrix may consist of straight-line segments.) The generating line in each of its positions is called an element of the cylinder, and a section of the cylinder made by a plane perpendicular to the elements is called a right section. All right sections of a cylinder are congruent. A circular cylinder (cylindrical surface) is one whose right sections are circles. Its other sections made by planes cutting all the elements are ellipses. Cylinders whose right sections are ellipses, parabolas, or hyperbolas are called elliptic cylinders, parabolic cylinders, or hyperbolic cylinders, respectively. All these are quadric cylinders. The lateral surface of a prism is a special type of cylindrical surface whose directrix and right sections are polygons.

Diagram of cylindrical surface.

The volume of a solid cylinder is $V = Bh$, where B denotes the area of the base, and h its altitude (height) measured perpendicular to the base. For a right circular cylinder whose base is a circle of radius r, the volume is $V = \pi r^2 h$, and the total surface area is $S = 2\pi r^2 + 2\pi rh$. *See* EUCLIDEAN GEOMETRY.

J. Sutherland Frame

Cyperales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Commelinidae of the class Liliopsida (monocotyledons). The names Glumiflorae, Graminales, and Poales have also been used for this order. There are only two families, the Poaceae (Gramineae), with about 8000 species, and the Cyperaceae, with nearly 4000. The Cyperales are Commelinidae with reduced, mostly wind-pollinated or self-pollinated flowers that have a unilocular, two- or three-carpellate ovary bearing a single ovule. The flowers are arranged in characteristic spikes or spikelets representing reduced inflorescences. The perianth is represented only by a set of bristles or tiny scales, or is completely missing. The leaves generally have a well-defined sheath and a narrow blade, often with a small adaxial appendage (the ligule) at the junction of the two. The stomates have two supporting cells and are arranged in straight files or rows of one or two, all oriented in the same direction. The pollen is uniformly trinucleate, and vessels are present in all vegetative organs.

The Cyperaceae, or sedge family, includes the bulrushes (*Scirpus*) and the papyrus (*Cyperus papyrus*) of Egypt, as well as the sedges (*Carex*). The Poaceae, or grass family, embraces all true grasses, including bamboos and the cereal grains such as wheat, maize, oats, and rye. The two families differ in a number of more or less consistent technical characters of the inflorescence, fruits, stems, and leaves. The well-developed intercalary meristem at the base of the blade in the grass leaf permits the plants to withstand grazing, and may be largely responsible for the dominance of grasses in large areas of the world with a suitable climate (seasonal rainfall, with the dry season too severe to permit forest growth). *See* BAMBOO; BARLEY; CEREAL; COMMELINIDAE; CORN; FLOWER; GRASS CROPS; LILIOPSIDA; MAGNOLIOPHYTA; MILLET; OATS; PLANT KINGDOM; RICE; RYE; SORGHUM; SUGARCANE; TIMOTHY; WHEAT.

Arthur Cronquist; T. M. Barkley

Cypress

The true cypress (*Cupressus*), which is very close botanically to the cedars (*Chamaecyparis*). The principal differences are that in *Cupressus* the cones are generally larger and the cone scales bear many seeds, whereas in *Chamaecyparis* each cone scale bears only two seeds; and in *Cupressus* the branchlets are not flattened as in *Chamaecyparis*. *See* CEDAR; PINALES.

All *Cupressus* species in the United States are western and are found from Oregon to Mexico. The Arizona cypress (*Cupressus arizonica*) found in the southwestern United States and the Monterey cypress (*Cupressus macrocarpa*) of California are medium-sized trees and are chiefly of ornamental value. The Italian cypress (*Cupressus sempervirens*) and its varieties are handsome ornamentals, but usually do well only in the southern parts of the United States. Other trees are also called cypress, such as the Port Orford cedar (*Chamaecyparis lawsoniana*) known also as the Lawson cypress, and the Alaska cedar (*Chamaecyparis nootkatensis*), known also as the Nootka cypress or cedar.

The bald cypress (*Taxodium distichum*) is an entirely different tree that grows 120 ft (36 m) tall and 3–5 ft (1–1.5 m) in diameter. It is found in the swamps of the South Atlantic and Gulf coastal plains and in the lower Mississippi Valley. The soft needle-like leaves and short branches are deciduous; hence, they drop off in winter and give the tree its common name. Also known as the southern or red cypress, this tree yields a valuable decay-resistant wood used principally for building construction, especially for exposed parts or where a high degree of resistance to decay is required as in ships, boats, greenhouses, railway cars, and railroad ties. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Cypriniformes

An order of actinopterygian fishes in the superorder Ostariophysi with about 2700 species, which ranks second in number only to the Perciformes. Cypriniformes is the largest order of freshwater fishes in the world. Six families of cypriniforms occupy the freshwaters of most of the major landmasses of the world, but are notably absent from South America, Madagascar, and the Australian realm. North America is represented by two native families, Cyprinidae and Catostomidae.

Cypriniforms differ from other fishes in having the following combination of characters: The fifth ceratobranchial (a pharyngeal bone) is enlarged and teeth (pharyngeal teeth) are ankylosed, or fused, to the bone; the pharyngeal teeth are opposed to a pair of enlarged posterior processes of the basioccipital bone; the mouth (jaws and palate) is toothless; the premaxillae have an ascending process and are usually protractile; four vertebrae are in the Weberian apparatus (a series of modified vertebrae that form a chain connecting the swim bladder with the inner ear); there are three branchiostegal rays; an adipose fin is lacking (except in some loaches); and, with a few exceptions, the head is scaleless. *See* OSTARIOPHYSI.

Cyprinidae. Shiner, minnow, chub, carp, and dace are the most common names applied to North American cyprinids. Rudd, ide, tench, and especially carp are cyprinid common names heard in Eurasia (see **illustration**).

With about 210 genera and more than 2000 species (and counting), this is the largest freshwater family of fishes, indeed, with the possible exception of the Gobiidae, the largest of all the vertebrate families. About 300 native species in 52 genera, in addition to seven genera and eight established nonnative species, occur in North America. Seven subfamilies are recognized but only one, the

subfamily Leuciscine, represents the native North America cyprinids.

Cyprinids are known from the Eocene in Asia and from the Oligocene in Europe and North America. The majority of the species, and morphologically the most diverse, are in Eurasia, especially in China and southeastern Asia.

Morphology. Cyprinids are distinguished from other cypriniforms in the following ways: pharyngeal teeth are in one to three rows, with never more than eight in any row; the upper jaw is bordered by premaxillae and is usually protrusible; the lips are usually thin, and neither plicate (pleated) or papillose; barbels are either present or absent; the first dorsal fin is spikelike in some species; and intermuscular bones are present. Most cyprinids are small fishes; the smallest, *Danionella transluida*, a species from Burma, is not known to exceed 12 mm (less than $^1/_2$ in.). The largest species are *Catlocarpio siamensis* of Thailand, whose length is no less than 2.5 m (probably to 3 m), and *Tor putitora* of eastern India, which is about 2.7 m. The largest native North American minnows are *Ptychocheilus lucius* (Colorado pikeminnow), which is reported to have reached 1.8 m (5.9 ft) and *P. oregonensis* (northern pikeminnow), 1.22 m (4 ft). However, most cyprinids of the world are probably less than 25 cm in length.

Reproduction. Reproductive strategies vary greatly. The large carps typically spawn in large free-flowing rivers where they broadcast millions of pelagic eggs which develop while carried by the river current. The eggs hatch within a few days and the fry reach flooded backwaters where they mature. The smaller species produce far fewer eggs, but compensate by preparing a nest in the form of a depression excavated in the substrate, or by selecting protective cover, such as crevices in rocks or bark of fallen trees, for secure sites to deposit their fertilized eggs. This strategy calls for parental care, usually provided solely by the male parent. Some species deposit their fertilized eggs in the interstices of gravel and pebbles, and no parental care is provided.

Many cyprinids are silvery and without prominent color markings, but during the breeding season some species acquire conspicuous red, yellow, and blue colors, which may be accompanied by iridescences. Also, during the breeding season males, and some females, develop tubercles on the head, body, and fins. Depending on size, number, and placement, the tubercles may function in defense, nest building, or to hold breeding pairs close together during spawning.

It is not unusual for a female minnow to produce in a single reproductive season a mass of mature eggs that exceeds several times her body weight. Such fecundity is accomplished by producing eggs in multiple clutches, the clutches being spaced to take advantage of environmental conditions likely to promise a successful spawn.

Yet another reproductive strategy is seen in the bitterling (*Rhodeus*) of Central Europe, which ensures the safety of its offspring by depositing the eggs within the valves of mussels. When the female is ready to spawn the oviduct is drawn out as a long tube that acts as an ovipositor, and the eggs are fertilized after they are given over to the mussel. The respiratory currents of the mussel provide essential aeration necessary for development of the eggs. The mussel's service is rewarded—its own embryos attach to the gill of the bitterling, where, safe from predators, they undergo early stages of development. *See* CARP.

Koi carp (*Cyprinus carpio carpio*). (*Photo © by J. J Photo*)

Psilorhynchidae. Members of this small family (two genera and five species), known from the mountain streams of Nepal and adjacent India to western Burma, have a small subterminal mouth whose jaws have sharp horny edges and fleshy lips; a pharyngeal bone with one row of four teeth; no barbels; narrow gill openings; a flattened ventral head surface; and reduced swim bladder.

Gyrinocheilidae. This family, known as algae eaters, is from Southeast Asia. The single genus (*Gyrinocheilus*), with four species, has several distinguishing characters which adapt it to an existence in high gradient streams. Rare among cypriniforms is the absence of pharyngeal teeth. The mouth is modified as a sucking device for attaching to objects; the gill opening is reduced to two small apertures, an upper incurrent and lower excurrent; and gill rakers number about 140. These fishes feed exclusively on algae, hence are popular as glass cleaners for home aquaria.

Catostomidae. The suckers are North American and eastern Asian fishes comprising 13 genera and about 72 species, two of which are thought to be extinct. North America has the greatest diversity and would appear to be the center of origin of catostomids. However, the evidence seems to indicate that catostomids originated in Asia from cyprinid ancestry.

Catostomids are characterized by having a single row of pharyngeal teeth numbering 16 or more. In some species the pharyngeal teeth are numerous, slender, very closely set (baleenlike), and function as a sieve; in other species the teeth are more or less comblike and function for mastication, or molarlike and used for grinding. The jaws lack teeth, the lips are thick and plicate or papillose, and both the premaxilla and maxilla border the upper jaw. Barbels are absent and scales are absent from the head.

The family has three rather distinctive subfamilies.

The Ictiobinae, with two genera, *Carpiodes* (quillbacks and carpsuckers) and *Ictiobus* (buffalo suckers), have a deep, relatively strongly compressed body; large scales; a complete lateral line; 115 to 190 pharyngeal teeth on each pharyngeal arch; a long dorsal fin base; and a swim bladder with two chambers. They occur from Canada to Guatemala. The Cycleptinae, having an elongate body in the adults, very small head, dorsal fin falcate (deeply concave) and with a long base, and an air bladder with two lobes, is represented in China by one species, *Myxocyprinus asiaticus*, and in North America (Mississippi River and adjacent gulf coastal drainages) by two described species of *Cycleptus*. The Catostominae have a slender terete (nearly cylindrical) body, small head, short dorsal fin, and swim bladder in two or three chambers. The lateral line is complete, incomplete, or absent, depending on the genus. Two genera, *Catostomus* (tribe Catostomini) and *Moxostoma* (tribe Moxostomatini), with about equal number of species, account for 42 of the 62 species of the family. Members of the tribe Catostomini live mostly in western North America, however one, *Catostomus catostomus*, ranges from New York State to eastern Siberia, and another, *C. commersoni* (white sucker), ranges from Georgia to British Columbia. All but one of the 30 species of Moxostomatini live in North America, east of the Continental Divide.

Most catostomids feed on the bottom on a variety of foods. Depending on the species, the food may be algae, detritus, aquatic insect larvae, or mollusks. Typically catostomids make spring runs up small streams, where they select clean gravel in flowing water as a spawning site. No parental care is given the young, except among the chubsuckers (genus *Erimyzon*), in which the males defend a nest, often an unoccupied nest of the stoneroller or creek chub.

Cobitidae. The loaches consist of about 110 species in 18 genera in the freshwaters of Eurasia plus one genus in Morocco. The greatest diversity is in southern Asia. This family and the river loaches are the most atypical of the cypriniforms. Cobitids are identified by their eel-like to fusiform shape, a subterminal mouth, pharyngeal teeth in one row, and an erectile spine below the eye. Two subfamilies are recognized: the Corbitinae, with one pair of rostral barbels and caudal fin usually rounded or slightly emarginate, and the Botiinae, with two pairs of rostral barbels and caudal fin deeply forked. Loaches are bottom-dwelling fishes, with a maximum length about 40 cm. A number of species are popular aquarium fishes.

Balitoridae. This family (river loaches) are similar to the cobitids but are recognized as a different family primarily by the Weberian apparatus, which has a Y-shaped tripus (the posteriormost Weberian ossicle). Also, the balitorids have three or more pairs of barbels near the mouth. Two rather distinctive subfamilies are recognized. The Nemacheilinae contains about 350 species of Eurasian fishes, with the greatest concentrations of species in India, Indochina, and China. They lack a spine under the eye; they have two pairs of rostral barbels and one pair of maxillary barbels, a subterminal mouth, and an adiposelike fin in some species; and scales are present or absent. Distinguishing characters of the subfamily Balitorinae are paired fins that are enlarged with adhesive pads on the ventral surface and oriented horizontally; pelvic fins, in some species, that are united under the belly; and a flattened head and body. These adaptations fit them well for occupying the fast mountain streams from India throughout Southeast Asia.

Herbert Boschung

Bibliography. P. M. Bănărescu and B. W. Goad, Cyprinids of Eurasia, pp. 127–155, in I. J. Winfield and J. S. Nelson (eds.), *Cyprinid Fishes: Systematics, Biology and Exploitations*, Chapman and Hall, London, 1981; T. M. Berra, *An Atlas of Distribution of the Freshwater Fish Families of the World*, University of Nebraska Press, 1986; S. V. Fink and W. L. Fink, Interrelationships of the ostariophysan fishes (Teleostei), *J. Linn. Soc. (Zool.)*, 72(4):297–353, 1981; R. L. Mayden, Cyprinid of the New World, pp. 240–263, in I. L. Winfield and J. S. Nelson (eds.), *Cyprinid Fishes: Systematics, Biology and Exploitations*, Chapman and Hall, London, 1891; R. L. Mayden (ed.), *Systematics, Historical Ecology, and North American Freshwater Fishes*, Stanford University Press, 1992; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994; I. L. Winfield and J. S. Nelson (eds.), *Cyprinid Fishes: Systematics, Biology and Exploitations*, Chapman and Hall, London, 1891.

Cyprinodontiformes

An order of teleost fishes comprising eight families, 88 genera, and about 810 species that inhabit fresh and brackish waters of tropical and temperate North and Middle America, Eurasia, and Indo-Malaysia. Several features unify the order: a symmetrical caudal skeleton of one epural supporting an unlobed (round) caudal fin; low-set pectoral fins with a large, scalelike postcleithrum; sensory pores chiefly on the head; and the upper jaw bordered by protractile premaxillae. Sexual dimorphism is common, with the males brightly colored, and, in the livebearers, an anal fin that is modified as a gonopodium (intromittent organ) for internal fertilization. The top of the head is usually flat and the mouth is usually terminal to superior, features that adapt the fishes to feed on or near the surface of the water.

Aplocheilidae. These fishes, known as rivulines, differ from all other cyprinodontiforms by having the pelvic fin bases inserted close together (versus pelvic fins not close together), a metapterygoid bone (versus no metapterygoid), and three basibranchials (versus two basibranchials). Two subfamilies are recognized: the Aplocheilinae (Old World rivulines) of Africa (south of the Sahara Desert to South Africa) Madagascar, Seychelles, India, Sri Lanka, and Indo-Malaysian Archipelago to Java; and the Rivulinae (New World rivulines) of southern Florida and

Middle America to Uruguay. In the Aplocheilinae, comprising eight genera and about 185 species, the supracleithrum is fused to the posttemporal and the first postcleithrum is present, whereas in the Rivulinae, comprising 12 genera and no less than 125 species, the supracleithrum is not fused to the posttemporal and the first postcleithrum is absent. *Rivulus marmoratus* (mangrove rivulus) of southern Florida and the West Indies is a self-fertilizing hermaphrodite. Another oddity in this group is *R. nudiventris* of Brazil, which lacks pelvic fins and a pelvic girdle. Reproduction is oviparous, and in only a few species is fertilization internal.

Profundulidae. This family of one genus, *Profundus*, and five species occurs in the Atlantic and Pacific slopes of Mexico, Guatemala, and Honduras.

Goodeidae. Two subfamilies are recognized. The Goodeinae, comprising 43 species in 16 genera, are centered in the Rio Lerma River basin of Mexico. They are viviparous fishes, and internal fertilization is accomplished by a pseudophallus fashioned from crowded short anterior anal fin rays that are slightly separated by a notch from the rest of the fin. The eggs are small, with little yolk, but placenta-like structures provide the embryos with nutrients and oxygen. The four species in the subfamily Empetrichthyinae lack a pelvic girdle and pelvic fins, as well as a pseudophallus, and fertilization is external. They are further distinguished from the other goodeids by lacking a lateral line and by the placement of the dorsal and anal fins far back on the body. These fishes are limited to a few pools and springs in southern Nevada. Two Mexican species and one Nevada species, the Ash Meadows poolfish, are thought to be extinct.

Fundulidae. Topminnows and killifishes comprise five genera and 48 species. Forty species occur in North America, 35 of which are in the genus *Fundulus* (see **illustration**). Most of these fishes are elongate, and somewhat compressed and flattened about the head and nape. The mouth is small and turned upward; the jaws are protractile and the lower extends beyond the upper; the dorsal and anal fins are small, similar in size and shape and posteriorly placed nearly opposite each other; the anal fin is unmodified; and fertilization is external. The flat head and superior mouth allow them to feed at the surface, as well as to utilize the oxygen-rich layer at the atmosphere-water interface without greatly altering their usual swimming posture. Consequently, these fishes can survive in habitats where the subsurface waters are essentially depleted of oxygen. They inhabit the lowlands of North and Middle America from southern Canada to Yucatan, as well as Bermuda and Cuba. Some of the fundulids are very euryhaline and temperature tolerant. The diamond killifish (*Adinia xenica*) can tolerate temperatures as high as 38°C (100°F) in freshwater with a salinity of 54 parts per thousand. It is not unusual to find Gulf killifish (*Fundulus grandis*) in tidal pools where the salinity reaches 76 ppt, as well as in the freshwater Brazos River 400 km (249 mi) from the Gulf of Mexico.

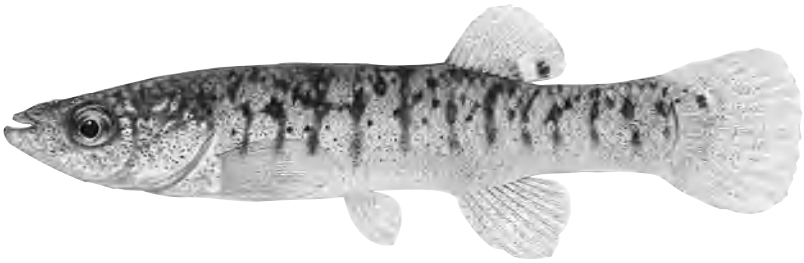

Female marsh killifish (*Fundulus confluentus*). (*Illustration © by Joseph R. Tomelleri*)

Valenciidae. This is a family of one genus and two species in the freshwaters of Spain, Italy, and western Greece.

Cyprinodontidae. Known as the pupfishes, this is a family comprised of nine genera and about 107 species of small freshwater, brackish water, and coastal marine inhabitants of North and Middle America, the West Indies, parts of South America, northern Africa, and the Mediterranean Anatolian area. Cyprinodontids are deep bodied and usually have a deep, strongly compressed caudal peduncle, an upturned mouth, and cycloid scales. The anal fin is unmodified, fertilization of eggs is external, and reproduction is oviparous. The dorsal process of each maxillary is expanded laterally, and the two processes nearly meet in the midline forming a distinct groove.

Five genera and 44 species of pupfishes occur in North America, 36 of which are in the genus *Cyprinodon*. The majority inhabit springs, sinkholes, salt marshes, and saline creeks in the deserts of Mexico and the southwestern United States. Various species of pupfishes are known to survive in pools that reach temperatures of 45°C (113°F), have salinity of 142 ppt (open oceanic water is about 35 ppt), and oxygen concentrations as low as 0.13 mg/liter. No other fishes could survive these harsh conditions. Some species have very restricted ranges (even a single small spring), which makes them especially vulnerable to extinction because of environmental changes, such as lowering ground water. Indeed, four species are extinct, three in Mexico and one, the Santa Cruz pupfish, which was described as new in 2002.

Anablepidae. This is a family of three genera and eight species inhabiting freshwater and brackish water of southern Mexico to southern South America. Unlike other cyprinodontiforms, the insertion of the pelvic fins is well behind the tip of the pectoral fins. Two of the genera have a tubular gonopodium formed from the anal fin rays, and in some males the gonopodium can be moved only to the right, in others only to the left. In females the genital aperture is open only to the right or only to the left. This probably means that a sinistral male only mates with a dextral female and vice versa. Fertilization is internal and the young are born alive. In the genus *Anableps* the eyes bulge above the top of the head and are divided longitudinally into upper

and lower portions, each portion with a pupil; thus, with the water line in the center of the eye, the fish can focus on images simultaneously from above and below. There are three species of the four-eyed fishes in the lowlands of the Pacific slope of southern Mexico to Honduras and the Atlantic slope of northern South America.

Poeciliidae/Poeciliinae. The livebearing subfamily Poeciliinae is considered by some authors to be equivalent to the family Poeciliidae. It comprises 22 genera and 192 species ranging from the eastern United States south to the Rio de la Plata drainage in northern Argentina, with the greatest diversity in Central America, Mexico, and the West Indies. Twelve genera and 93 species occur in North America. The males have a gonopodium formed from three anal rays. To accommodate the elongate gonopodium, the anal fin is displaced forward, often to a point under the pelvic fins. Depending on the species, poeciliid embryos are nourished by yolk deposited before the eggs are fertilized (ovoviviparity), or the embryos are nourished by nutrients transferred from the mother via a placenta (viviparity). A phenomenon commonly seen in poeciliids is superfetation, a process by which two or more broods at different stages of embryonic development occur simultaneously in the same female. The Amazon molly, *Poecilia formosa*, from Texas and northern Mexico, is unisexual, all female. They are of hybrid origin and depend on males of bisexual species for sperm to stimulate embryonic development. Many of the poeciliids are popular aquarium fishes, such as sailfin mollies, guppies, swordtails, and platyfishes. The mosquitofishes, *Gambusia affinis* and *G. holbrooki*, have been introduced into numerous countries for the purpose of mosquito control; however, their presence is of questionable value. As a result of predation on eggs, larvae, and juveniles of other fishes, mosquitofishes often significantly reduce or eliminate native fishes. Herbert Boschung

Bibliography. J. H. Huber, *Killi-Data 1996* [updated checklist of taxonomic names, collecting localities and bibliographic references of oviparous Cyprinodont fishes (Atherinomorpha, Pisces)], Société Française d'Ichtyologie, Muséum National d'Histoire Naturelle, Paris, 1996; G. K. Meffe and F. F. Snelson, Jr., *Ecology and Evolution of Livebearing Fishes* (*Poeciliidae*), Prentice Hall, Englewood Cliffs, 1989; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994; L. R. Parenti, A phylogenetic and biogeographical analysis of Cyprinodontiform fishes (Teleostei, Atherinomorpha), *Bull. Amer. Mus. Nat. Hist.*, 168(4):335–557, 1981; J. J. Scheel, *Atlas of Killifishes of the Old World*, T.F.H. Publications, Neptune City, NJ, 1990.

Cyrtosoma

A subphylum of the Mollusca comprising members of the classes Monoplacophora, Gastropoda, and Cephalopoda. The use by paleontologists of subphyla, including Cyrtosoma and Diasoma, in the classification of the phylum Mollusca has not been generally adopted by students of living forms. There is no disagreement over the further division of the phylum into eight clearly distinguishable classes.

Primitive members of the Cyrtosoma have, or had, a conical univalved shell, often twisted into a spiral. The relatively small single-shell aperture forced the anus to lie close to the mouth, either behind it (Monoplacophora and Cephalopoda) or ahead of it in those forms which have undergone torsion (Gastropoda; **illus.** *f*). Studies of the comparative anatomy and history of living and fossil mollusks indicate that monoplacophorans are the most primitive cyrtosomes and that the other two classes were derived independently from the Monoplacophora (see illus.). *See* CEPHALOPODA; GASTROPODA; MONOPLACOPHORA.

Although both gastropods and monoplacophorans are found in earliest Cambrian rocks in Siberia and China, intermediate forms which link the two classes (*Pelagiella*; illus. *e*) survived through the Cambrian. Apparently asymmetric coiling of the shell of some early monoplacophorans caused the body to begin to rotate with respect to the shell, thus initiating torsion. When this process was completed, the shell was carried so that it coiled posteriorly, and the first gastropods had evolved. Other early monoplacophorans, which remained bilaterally symmetrical, gave rise to animals with a variety of symmetrical shell forms, including those of a diverse extinct group, the Bellerophontida. The bellerophonts have usually been placed in the Gastropoda, but there is mounting evidence that they were untorted and hence monoplacophorans.

The third cyrtosome class, the Cephalopoda, did not appear until the end of the Cambrian. Its first members (illus. *b*) had small conical shells with numerous calcareous partitions (septa) at the apical end. Similar septate monoplacophorans are known from slightly older rocks, but these lacked a siphuncle, used for buoyancy control. It was the appearance of the siphuncle which enabled the first cephalopods to swim, and it has been responsible for the great post-Cambrian success of that class.

Each cyrtosome class has, in turn, been the most diverse molluscan group. In the Ordovician the cephalopods succeeded the monoplacophorans as the most diverse molluscan class, and they held this position until the close of the Mesozoic. They were followed by the gastropods, which by then had diversified into a great variety of aquatic and terrestrial habitats. It has been estimated that about 80% of the species of living mollusks are gastropods.

Although primitive cyrtosomes, including the only living shelled cephalopod, *Nautilus*, are shell-bearing animals, many advanced cyrtosomes have either lost or greatly reduced their shells (examples include slugs, nudibranchs, squids, and octopuses). One rare family of gastropods has even evolved a truly bivalved shell. Consequently, it is not any

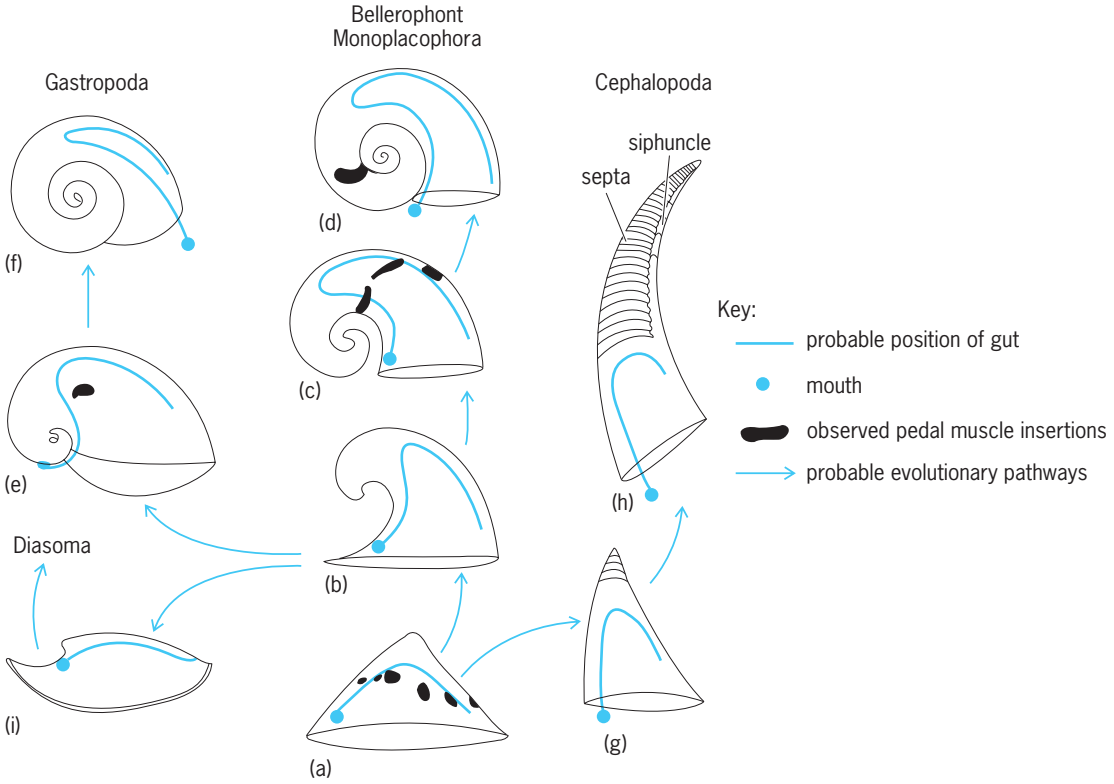

Evolution of the Cyrtosoma. (*a*) *Lenaella*. (*b*) *Latouchella*. (*c*) *Cyrtolites*. (*d*) *Bucania*. (*e*) *Pelagiella*. (*f*) *Aldanella*. (*g*) *Knightoconus*. (*h*) *Plectronoceras*. (*i*) *Mellopegma*. Organisms *a–e*, *g*, and *i* are monoplacophorans, *f* is an early gastropod, and *h* is the oldest cephalopod.

particular shell form which unites all members of the Cyrtosoma; rather, it is a knowledge of their evolutionary history. *See* MOLLUSCA.

Bruce Runnegar

Bibliography. J. Pojeta and B. Runnegar, *The Paleontology of Rostroconch Mollusks and the Early History of the Phylum Mollusca*, U.S. Geol. Surv. Prof. Pap. 968, 1976.

Cystoporata

An extinct order of bryozoans in the class Stenolaemata. Cystoporates tend to have robust colonies composed of relatively simple, long, slender, and tubular zooecia (with thin, somewhat porous calcareous walls), separated by blisterlike vesicles (cystopores) stacked upon one another. Their colonies display a very gradual transition between comparatively indistinct endozone and exozone regions. *See* BRYOZOA; STENOLAEMATA.

Morphology. Cystoporate colonies range from small and delicate to large and massive; they can be thin encrusting sheets, tabular or nodular masses, or erect twiglike or frondlike growths. Some cystoporate colonies are not regionated; others display indistinct endozone and exozone regions. Lacking known ovicells, stolons, acanthorods, vibracula, and avicularia, cystoporate colonies may sometimes include mesopores in addition to the ubiquitous cystopores. Many colonies exhibit monticules or maculae or both.

Individual cystoporate zooecia are straight to markedly curved long tubes. Because the zooecia are generally separated from one another by cystopores or vesicles, the walls of nearby zooecia remain distinctly separate and thin throughout the colony. The zooecial walls are calcareous, finely granular in microstructure, and perforated in some species by a few minute pores. The space within each zooecium may be crossed by diaphragms or may lack transverse calcareous partitions. Round or oval in outline, the zooecial aperture lacks an operculum, but often is surrounded by a raised rim (peristome) or partially covered over by a hoodlike lunarium.

History and classification. The cystoporates may possibly have evolved from the earliest ctenostomes. Cystoporates first appeared late in the Early Ordovician, somewhat earlier than representatives of other stenolaemate orders; thus, the cystoporates may possibly have in turn given rise to those orders, although all these forms may simply share a common ancestor further back in time. Apparently marine, the cystoporates (or ceramoporoids or fistuliporoids; formerly included in the Cyclostomata) became moderately common by Mid-Ordovician time—contributing to building small reefs, as well as to level-bottom communities—and remained so until they died out in the latest Permian. *See* CTENOSTOMATA; CYCLOSTOMATA (BRYOZOA).

Roger J. Cuffey

Cytochemistry

The science concerned with the chemistry of cells. Specifically, the macromolecules of subcellular structures can be treated chemically to form microsopically visible end products. For example, proteins, enzymes, lipids, carbohydrates, and nucleic acids can be directly visualized in cell nuclei, membranes, and organelles by cytochemical methods which generate images that can be viewed by either bright field or light, fluorescence, confocal, or electron microscopes. *See* ELECTRON MICROSCOPE; FLUORESCENCE MICROSCOPE; INTERFERENCE MICROSCOPE; PHASE-CONTRAST MICROSCOPE; SCANNING ELECTRON MICROSCOPE.

Enzyme cytochemistry. Enzyme cytochemical methods detect enzymes associated with subcellular structures (see **illus.**). The method employs a substrate specifically cleaved by an enzyme in the cell to liberate a product which is transformed into a visible precipitate. For example, to localize the enzyme acid phosphatase in lysosomes, a subcellular organelle, a lead sulfide precipitate is produced by splitting a phosphate (product) from cytidine monophosphate (substrate) and transforming the phosphate product initially into lead phosphate and subsequently into lead sulfide; the latter is visible as a brown-black precipitate in the light microscope. Other methods employ azo dye or formazan precipitates rather than lead to produce a visible product. The number of enzymes that can be detected are limited by the availability and specificity of substrates. *See* ENZYME.

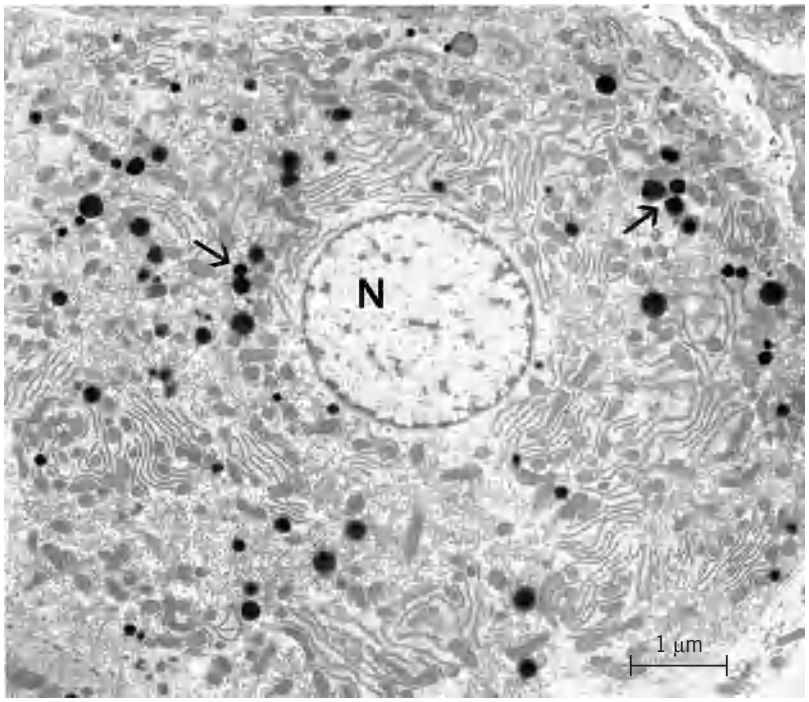

Electron micrograph of a hepatocyte from rat liver showing the localization of catalase in peroxisomes by using an enzyme cytochemical method. The catalase activity in peroxisomes was revealed by incubation in a diaminobenzidine medium after fixation with aldehydes. Peroxisomes (arrows) appear as black spherical organelles distributed throughout the cytoplasm. The nucleus is labeled N. (*P. M. Novikoff and A. Yam*)

Immunocytochemistry. Immunocytochemical methods are used to detect specific proteins associated with subcellular structures. The methods are based on the immunologic principle of the built-in specificity of antibody-antigen interactions; that is, an antibody (a protein called immunoglobulin) has a unique binding site that recognizes an antigen. An antigen in a tissue or a cell is detected by employing either a direct or indirect method. The direct method requires a specific antibody linked to a molecule that produces a visible signal; the antibody-molecule complex is applied to a tissue or cells for antigen binding. The indirect method employs an unlabeled primary antibody which is applied to a tissue or cells. The antibody binds to the antigen. The bound primary antibody is then detected by application of a secondary antibody that yields a visible signal. For example, to localize acid mannosidase in lysosomes by the indirect method, a specific antibody (monoclonal or polyclonal) against acid mannosidase (antigen) is applied to a tissue section or to cells and interacts with the intracellular antigen (acid mannosidase) to form a specific antibody-antigen complex, which is then visualized by employing a labeled second antibody that binds to the first antibody. Second antibodies can be labeled with fluorescent probes (such as fluorescein, rhodamine, Texas Red, Cy 3, and Cy 5), enzymes (including peroxidase and alkaline phosphatase), metals (such as gold and ferritin), and high-affinity complexes (for example, avidin-biotin) that are detectable with the appropriate microscope (such as bright-field light, fluorescence, confocal, and electron), in some cases, after application of other reagents to generate a visible product.

Double or triple label immunocytochemistry methods allow multiple antigens to be localized within a tissue or cells. Multiple antigens can be detected at the same time by using different antibodies labeled with probes that fluoresce at different wavelengths; for example, in triple label fluorescence immunocytochemistry one antibody is labeled with fluorescein (520 nanometers wavelength), a second with Cy 3 (570 nm), and another with Cy 5 (670 nm). Multiple antibodies can also be labeled with different enzymes (for example, peroxidase and alkaline phosphatase) or with different size gold particles (such as 5 or 10 nm). When the indirect method is employed, the different primary antibodies used must be from different species so that they can be distinguished using secondary antibodies. Fluorescent probes are detected with conventional fluorescence microscopy or with confocal microscopy using either appropriate filters or lasers. Enzyme probes and gold probes are detected with light and electron microscopy. *See* IMMUNOCHEMISTRY; IMMUNOFLUORESCENCE; PROTEIN.

In-situ hybridization. In-situ hybridization cytochemical methods detect specific deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) sequences in specific cells or intact nuclei, chromosomes, or cytoplasmic subcellular organelles. The method involves the hybridization of known DNA or RNA sequences

(probe) with complementary sequences of DNA or RNA (target) in intact cells. When the target is messenger RNA, the direct product of gene expression, which encodes a specific protein, the purpose of the method is usually to identify the tissues and cells in which a particular gene is expressed. When the target is DNA, the purpose is frequently to map, or localize, a specific gene on a particular chromosome or to detect the presence of a viral genome in infected cells. The binding or hybridization of probe and target is based on the matching of nucleic acid base pairs adenine (A) to thymine (T), and guanine (G) to cytosine (C), within contiguous stretches of nucleic acid sequences. The known probe can be labeled with fluorescent molecules, enzymes, avidin-biotin, or radioactive elements with detection of the probe-target complex accomplished using fluorescence microscopy, cytochemistry, immunocytochemistry, or autoradiography, respectively. Polymerase chain reaction (PCR) also can be performed on the tissue and cells to amplify specific sequences prior to in-situ hybridization to detect single copies of DNA or low levels of RNA. *See* NUCLEIC ACID.

Lectin cytochemistry. Lectin cytochemical methods are employed to detect specific sugar molecules on the plasma membrane and intracellular membranes of cells. Lectins are proteins with binding sites that recognize and attach tightly to a specific sequence of sugar moieties. The lectins exhibit a high degree of specificity for a wide spectrum of sugars. For example, the lectin concanavalin A will detect mannose-rich oligosaccharides such as those in glycoproteins that reside in the endoplasmic reticulum of the cell; wheat germ agglutinin will detect the sugar glucosamine, characteristic of many proteins located in the distal regions of the Golgi apparatus; and peanut agglutinin will detect the sugar galactosamine on cell surfaces. Lectins which are generally isolated from plants do not have enzymatic activity but can be labeled with a variety of molecules (such as enzymes, fluorescent molecules, or metals) allowing their detection by microscopy. *See* CARBOHYDRATE.

Autoradiography. Autoradiographic methods employ radioactive labeled molecules to identify sites within cells where synthesis of macromolecules occurs. This method is based on the property that an isotope emits ionizing particles during radioactive decay. Radioactive labeled precursor molecules (building blocks of molecules) become incorporated into specific cellular sites after injection into living organisms or cells, producing a radioactive product. To establish the location of the radioactive product in tissues or cells, a photographic emulsion is placed over tissues or cells in the dark for a period of time, followed by a photographic developer solution to reveal sites containing the radioactive product. The radioactive product is produced by the interaction of silver bromide crystals suspended in the emulsion, with the ionizing radiation given off by the decaying isotope. The product is made visible by a photographic development solution and appears as black dots or grains. The black dots or grains indicate sites of the radioactive product; these sites are visible in the light and electron microscopes. Autoradiography has determined that DNA synthesis occurs in the nucleus by using a radioactive labeled precursor molecule of DNA tritiated (^{3}H)-thymidine; that collagen is synthesized by fibroblasts by using ^{3}H-proline, an amino acid that is very abundant in collagen; and that thyroglobulin, an iodinated protein from which thyroid hormone (which retains the iodine atoms) is derived by proteolytic degradation, is synthesized by the thyroid and stored in the lumen of thyroid by injecting animals with radioactive iodine. *See* AUTORADIOGRAPHY.

Fluorescence analog. Fluorescence analog cytochemical methods are employed to determine the function of proteins in living cells (for example, during growth, cell division, or migration). A purified protein is coupled to a fluorescent probe and microinjected into living cells, and its fate can be followed in a conventional fluorescent microscope or confocal microscope. For example, fluorescent labeled tubulin, a subunit of microtubules, can be injected into a living cell, and the incorporation of the labeled tubulin into microtubules during cell migration can be followed second by second. Multiple proteins can also be followed in cells by labeling the proteins with a fluorescent probes that emit light of different wavelengths (see **colorplate**).

Measuring pH. Measuring the pH of different organelles involves determining the concentration of hydrogen ions in intracellular organelles in living cells. One method uses fluorochrome probes that fluoresce at different wavelengths depending on the concentration of hydrogen ions. For example, the pH of lysosomes was determined to be 5.0 by using the fluorescent probe fluorescein. *See* PH.

Marker enzymes. Marker enzymes are employed to identify subcellular organelles and to distinguish subcellular organelles from each other in intact tissues and cells and in cell fractions of organelles using enzyme cytochemical methods. For example, acid phosphatase is used as a marker enzyme for lysosomes, catalase for peroxisomes (see illus.), thiamine pyrophosphatase for the Golgi apparatus, nucleoside diphosphatase for the endoplasmic reticulum, and 5′-nucleotidase for the plasma membrane. The marker enzyme concept (that each subcellular organelle possesses one or a set of unique enzymes detectable by enzyme cytochemical methods) has been greatly enhanced by the introduction of immunocytochemistry, which permits the visualization of a larger number of proteins specific for subcellular organelles.

Epitope-tagged proteins. The detection of epitope-tagged proteins introduced by transfection is a molecular genetic-immunocytochemical method to localize proteins for which antibodies are not available. The method involves the construction of an expression vector in which the gene for the protein of interest (for which there is no available antibody) and the DNA sequence encoding an epitope tag (such as a small peptide c-myc protein for which there is an antibody) are linked so that the tandem gene sequences are translated to produce a single protein.

Transfection of the vector into living cells by one of a variety of methods results in production of a fused protein (protein of interest containing the linked epitope tag) in the cells. The fused protein is detected in cells by immunocytochemistry using an antibody to the epitope tag protein (for example, c-myc protein). Commonly used epitope tags include a short peptide-segment from c-myc, a cellular protein normally located in the nucleus, or a peptide from the envelope proteins of the influenza virus. Transfection of living cells is commonly accomplished by an electroporation method which makes the cell membrane transiently permeable to the vector by the application of a brief but intense electric shock.

Histologic dyes. Histologic dye methods use organic dyes to stain cell components in tissue sections of normal and pathologic tissues. The dye methods are used extensively in clinical pathology laboratories for diagnosis because they can be performed easily and quickly. The dyes hematoxylin and eosin are used in a routine staining procedure in pathology laboratories to establish overall architecture or histology of tissues, and they are extremely valuable for determining whether a tissue is normal or pathologic during surgery. Dye methods are available for staining neutral lipids (oil red O or sudan black), for staining nuclear DNA (methyl-green), for staining the nucleolus and ribosomes (pryonin), and for staining acid mucopolysaccharides (alcian blue or ruthenium red). *See* LIPID.

Laser scanning confocal microscopy. Laser scanning confocal microscopy uses lasers to obtain thin optical sections of the entire depth of fluorescent-labeled cells or tissues with an optical system that integrates the focus of the optical section and the focus of illumination to achieve confocality (that is, same focus). This results in the production of fluorescent images of superior clarity because the out-of-focus fluorescence above or below the focused area of a cell or tissue is eliminated. Confocal microscopy is extremely useful for performing double or triple immunofluorescence in which as many as four antigens can be simultaneously viewed: three antigens by fluorescent probes and one by a reflectance probe (such as gold). Images are subsequently processed by computer programs to produce three-dimensional reconstruction, volume rendition, and stereo projections. Study of fluorescent-labeled living cells can be performed optimally by using real-time scanning confocal microscopy, which combines optical sectioning capability of a confocal microscope with rapid scanning and collection of images in real time.

Specimen preparation. Cytochemical methods can be applied to various types of specimens, including cells in culture, sections of tissues, and cell fractions of isolated organelles. For most cytochemical procedures, except those noted above for living cells, the specimen requires fixation with formaldehyde or glutaraldehyde, which functions to keep macromolecules of interest in place. When a tissue is to be used for cytochemical procedures, the tissue is fixed by immersion or perfusion, and then sections are prepared with a cutting instrument (microtome). Depending on the enzyme or antigen to be localized and the level of microscopy (light or electron) used for analysis, tissues can be sectioned without embedding (frozen or nonfrozen) or with embedding (in paraffin for light microscopic pathologic studies or in epon for electron microscopic study). The objective is to preserve the enzyme or antigen of interest as well as the morphology of the tissue or cells. In cases where antigen or enzymes are inhibited by fixation, unfixed frozen sections have to be prepared. The purpose of the experiment will dictate the kind of fixation and sectioning instrument. *See* CELL (BIOLOGY); CELL, SPECTRAL ANALYSIS OF; MICROTECHNIQUE.

Phyllis M. Novikoff

Bibliography. J. E. Beesley (ed.), *Immunocytochemistry: A Practical Approach*, 1993; R. F. Bonner et al., *Science*, 78:1481, 1997; E. Holtzman, Lysosomes, in P. Siekevitz (ed.), *Cellular Organelles*, 1989; R. D. Lillie, *Histopathologic Technic and Practical Histochemistry*, 1965; A. B. Novikoff and P. M. Novikoff, Cytochemical staining reactions for enzymes in cytoplasmic organelles, in L. A. Manson (ed.), *Biomembranes*, vol. 2, 1971; G. J. Nuovo, *PCR in situ Hybridizaton: Protocols and Applications*, 1997; J. B. Pawley (ed.), *Handbook of Biological Confocal Microscopy*, 1995; J. M. Polak and I. M. Varndell (eds.), *Immunolabelling for Electron Microscopy*, 1984; Y. L. Wang and D. L. Taylor (eds.), *Fluorescence Microscopy of Living Cells in Culture*, 1989.

Cytochrome

A member of a class of proteins containing hemes as prosthetic groups that catalyze oxidation/reduction reactions. Cytochromes are distinct from other heme-containing proteins, such as myoglobin and hemoglobin, which function in the binding and/or transport of molecular oxygen; catalase, which eliminates hydrogen peroxide from cells; and peroxidases, which use hydrogen peroxide to oxidize various substrates.

Types. Cytochromes are identified by the type of heme prosthetic group they contain. Heme a, heme b, and heme c differ from one another by the chemical groups attached to the general iron–protoporphyrin structure (**Fig. 1**). Individual cytochromes can also be further identified by the wavelength of light that the heme group maximally absorbs, which is affected by the protein environment of the heme group. For example, cytochrome b_{562} contains heme b that maximally absorbs light with a wavelength of 562 nm.

Oxidation and reduction. Cytochromes are able to catalyze oxidation/reduction reactions because their heme prosthetic groups can exist in two forms: the oxidized form, in which the iron is in the Fe(III) oxidation state, and the reduced form, in which the iron is in the Fe(II) oxidation state. By switching back and forth between the oxidized and reduced forms, cytochromes can transfer electrons from one substrate to another using the following steps. First, the oxidized form of the cytochrome binds to the reduced

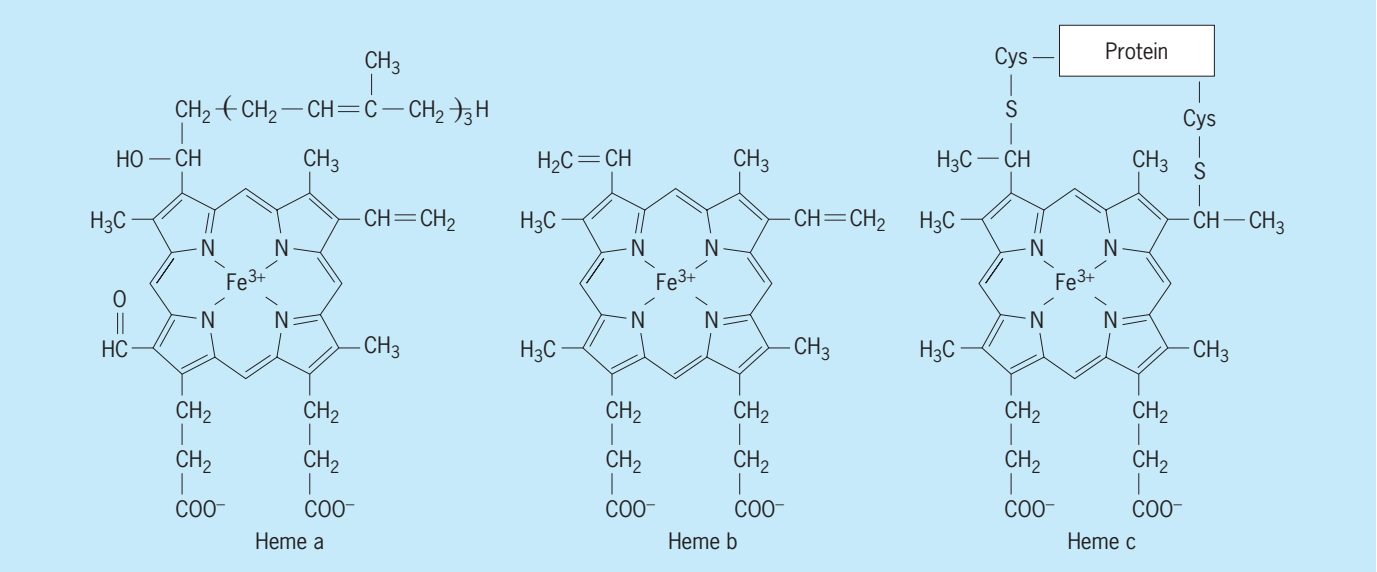

Fig. 1. Structures of the heme a, heme b, and heme c iron-protoporphyrin groups found in cytochromes.

form of one substrate. One electron is then transferred from this substrate to the heme group. This generates the oxidized form of this substrate and the reduced form of the cytochrome. Next, the reduced form of the cytochrome binds to the oxidized form of a second substrate. One electron is then transferred form the heme group to the second substrate. This generates the reduced form of the second substrate and the oxidized form of the cyctochrome.

Electron transport system. Cytochromes play essential roles in oxidative phosphorylation in all organisms except for a few anaerobic bacteria. Oxidative phosphorylation is a very complicated process by which energy released by oxidation/reduction reactions is used to fuel the synthesis of the energy-rich compound adenosine triphosphate (ATP). In the first stage of oxidative phosphorylation, the electron transport system takes the energy released by oxidation/reduction reactions and stores it in the form of a proton gradient across the inner membrane of the mitochondria. In the second stage of oxidative phosphorylation, an enzyme called the ATP synthase takes the energy stored in the proton gradient and uses it to create ATP. Cytochromes are involved only in the first stage of oxidative phosphorylation, the electron transport system that establishes the proton gradient. *See* ADENOSINE TRIPHOSPHATE (ATP).

The electron transport system comprises four protein complexes found within the inner mitochondrial membrane (**Fig. 2**).

Complex I. Complex I is the NADH–coenzyme Q reductase complex. This complex takes electrons from hydrogenated (reduced) nicotinamide adenine dinucleotide (NADH) and transfers them to a membrane-soluble molecule called coenzyme Q. In this process, four protons are transferred across the inner mitochondrial membrane for every two electrons transferred to coenzyme Q. This complex contains no cytochromes. *See* NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD).

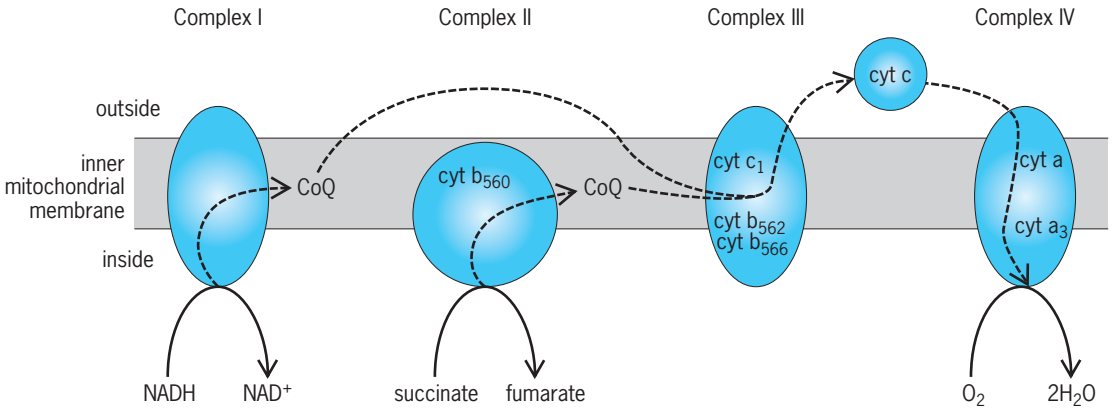

Fig. 2. Diagram of the mitochondrial electron transport system. The dashed arrows represent the flow of electrons as they enter the system either at complex I (from NADH) or at complex II (from succinate) and exit the system at complex IV (to molecular oxygen). The purpose of this process is to pump protons across the inner mitochondrial membrane, establishing a gradient that is used to generate ATP.

Complex II. Complex II is the succinate–coenzyme Q reductase complex. This complex comprises at least four different protein subunits. Part of this complex is the enzyme succinate dehydrogenase, which faces the inner side of the membrane. This enzyme catalyzes the conversion of succinate (a metabolic intermediate found inside the mitochondria) to fumarate, and thereby it transfers electrons from the succinate to an FAD covalently bound to the complex. This complex also contains one 4Fe-4S iron–sulfur cluster, two 2Fe-2S iron–sulfur clusters, and a cytochrome b_{560}. While the path taken by the electrons through this complex is not clear, the electrons ultimately end up transferred to coenzyme Q.

Complex III. Complex III is the coenzyme Q–cytochrome c reductase complex. This complex comprises at least 11 different protein subunits. One protein subunit contains two hemes, b_{562} and b_{566}, another protein subunit contains a 2Fe-2S iron–sulfur cluster, and yet another protein subunit is cytochrome c_1. This complex takes electrons from coenzyme Q. These electrons follow a complex, defined pathway through the iron–sulfur cluster and three heme groups, and ultimately end up transferred to another protein called cytochrome c. In this process, four protons are transported across the inner mitochondrial membrane from inside the mitochondria for every pair of electrons transferred from coenzyme Q to cytochrome c.

Cytochrome c. Cytochrome c is a small cytochrome protein that is not found within the inner mitochondrial membrane. Instead, cytochrome c floats around between the inner and outer mitochondrial membranes. It picks up electrons from complex III and transfers them to complex IV.

Complex IV. Complex IV is the cytochrome c oxidase complex. This complex comprises 13 different protein subunits. One subunit contains a 2Cu-2S copper–sulfur cluster, and another subunit contains two hemes, a and a_3. This complex takes electrons from cytochrome c and ultimately transfers them to molecular oxygen to form water. The electrons travel from cytochrome c to the 2Cu-2S cluster, then to the heme a group, then to the heme a_3 group, and then finally to the bound oxygen molecule. In this process, two protons are transported across the inner mitochondrial membrane from inside the mitochondria for every pair of electrons transferred from cytochrome c to molecular oxygen. *See* BIOLOGICAL OXIDATION; COENZYME; MITOCHONDRIA.

Photosynthesis. Cytochromes also play an essential role in photosynthesis. In bacteria, photosynthesis occurs on the cell membrane at protein complexes called the reaction center, which contains chlorophyll. Light causes electrons from the reaction center to be transferred from the chlorophyll to a quinone molecule. Quinone moves through the membrane and transfers electrons to a cytochrome bc_1 complex. The cytochrome bc_1 complex then transfers these electrons to either a cytochrome c_2 protein (in the case of purple bacteria) or a cytochrome c_{553} protein (in the case of green sulfur bacteria). These soluble cytochrome proteins then return the electrons to the chlorophyll of the reaction center for another cycle. During this process, a proton gradient is established across the cell membrane that is used to fuel ATP generation.

In plants, photosynthesis occurs in the thylakoid membrane of the chloroplasts. In a proton complex called photosystem II, light causes electrons (derived from the oxidation of water to molecular oxygen) to be transferred from chlorophyll to plastoquinones, which move through the membrane and transfer the electrons to the cytochrome b_6f complex. The cytochrome b_6f complex transfers these electrons to plastocyanin, which in turn transfers them to chlorophyll in another protein complex called photosystem I. In photosystem I, light causes these electrons to be passed through another series of carriers until they ultimately reduce $NADP^+$ to NADPH. The NADPH is needed by plant cells to reduce compounds during many important biosynthetic pathways. *See* PHOTOSYNTHESIS.

P-450s. Much recent research has focused on a class of cytochromes called P-450s, which contain the heme b prosthetic group. Their name is derived from the fact that the reduced forms of these proteins (in which the iron in the heme is bound to carbon monoxide) maximally absorb light with a wavelength of 450 nm. In higher animals, all cytochrome P-450s are found within membranes—usually within the membrane of the endoplasmic reticulum but also within the mitochondrial membrane. Cytochrome P-450s are components of monoxygenase systems that catalyze the addition of a hydroxyl group to an organic compound using molecular oxygen. This reaction requires electrons passed to cytochrome P-450 from NADH by a protein called cytochrome P-450 reductase.

The human genome encodes 57 cytochrome P-450 proteins, each thought to be specific for hydroxylating a different organic compound. Many of these cytochrome P-450s are involved in hydroxylating xenobiotics (compound such as drugs or toxins that are not ordinarily found within the cell). For example, the carcinogen benzo[*a*]pyrene found in cigarette smoke is hydroxylated and detoxified by a cytochrome P-450. Other cytochrome P-450s are involved in the synthesis of bile acids and steroid hormones. Some are involved in fatty acid metabolism, particularly the omega oxidation of fatty acids, in which the methyl carbon at the end of the fatty acid hydrocarbon chain is hydroxylated. The biological functions of many cytochrome P-450s encoded in the human genome have still not been determined.

M. Todd Washington

Bibliography. D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 4th ed., 2004; D. Voet and J. G. Viet, *Biochemistry*, 3d ed., 2004.

Cytokine

Any of a class of soluble proteins (molecular weight less than 100,000 daltons, or 100 kD) that are released by a cell to send messages that are

delivered to the same cell (autocrine), an adjacent cell (paracrine), or a distant cell (endocrine). The cytokine binds to a specific receptor and causes a change in function or in development (differentiation) of the target cell. Cytokines are involved in reproduction, growth and development, normal homeostatic regulation, response to injury and repair, blood clotting, and host resistance (immunity and tolerance). Unlike cells of the endocrine system, many different types of cells can produce the same cytokine, and a single cytokine may act on a wide variety of target cells. Further, several cytokines may produce the same effect on a target, so the loss of one type of cytokine may have few if any consequences for the organism; this situation is called redundancy. Finally, the response of a target cell may be altered by the context in which it receives a cytokine signal. The context includes other cytokines in the milieu and extracellular matrix. Thus has developed the concept of cytokines as letters of an alphabet that combine to spell words which make up a molecular language.

Cytokines may be divided into six groups: interleukins, colony-stimulating factors, interferons, tumor necrosis factor, growth factors, and chemokines.

Interleukins. Interleukins are proteins that are produced primarily by one type of lymphocyte (a type of white blood cell, or leukocyte) or macrophage (a large phagocytic cell of the immune system) and that act on other leukocytes. However, it is now known that production of interleukins is not confined to lymphocytes and macrophages. Interleukins may be grouped into families with similar structure and function. The plethora of interleukins suggests considerable redundancy, although it is suspected that each interleukin may have some special purpose. Thirty-four types are described below.

Interleukin (IL)-1α is composed of two protein subunits produced by macrophages, a specialized antigen-presenting cell called a dendritic cell, B (bone marrow–derived) and T (thymus-derived) lymphocytes (also known as B and T cells). IL-1β is also a dimer of two proteins and is produced by B and T cells, macrophages, dendritic cells, skin epithelium, fibroblasts, blood vessel endothelium, joint-lining cells, cartilage cells, bone marrow and circulating leukocyte cells, liver cells, adrenal gland cells, pancreatic islet β cells (which secrete insulin), fertilized eggs (early embryos), and cells of the nervous system. IL-1α and IL-1β bind to the same receptor, which may occur on a variety of target cells, for example, T lymphocytes (stimulating immunity), those of the nervous system (causing fever), macrophages and dendritic cells, and endothelial cells (activating inflammation and blood clotting). Soluble IL-1 receptors have been identified and synthesized. They antagonize IL-1 action, and can prevent shock in response to bacterial endotoxin. *See* CELLULAR IMMUNOLOGY.

Interleukin-2 is a glycoprotein produced by a subset of T lymphocytes called T-helper type 1 cells (Th1), which "help" in immune responses. IL-2 is a prototypic example of a cytokine produced by Th1 lymphocytes. Th1-type cytokines promote cell-mediated immunity and inflammation (so IL-1 is included in the Th1 group); they inhibit generation and production of Th2-type cells and Th2-type cytokines. Conversely, Th2 cytokines inhibit Th1 responses. Hence, there is reciprocal inhibition and Th1/Th2 cytokine balance. IL-2 may also be produced by B cells, natural killer cells, and oligodendrocytes (which myelinate nerve cells) in the brain. IL-2 stimulates the development of cytotoxic (killer) T lymphocytes and B lymphocytes (which develop into antibody-secreting cells). IL-2 also activates nonspecific (unlike T and B cells) cytotoxicity in natural killer cells. Natural killer cells participate in the early defense response to infection and may reject certain types of cancers early in their development. IL-2 may also be made by fetal trophoblast cells that lie at the feto-maternal interface in the placenta. Receptors for IL-2 are found on lymphoid cells and on astroglia of the central nervous system.

Interleukin-3 is a growth factor for multiple lineages of hemopoietic cells (which develop into blood cells). It was originally defined by its ability to stimulate the growth of mast cells (a large leukocyte that stores various inflammatory mediators). IL-3 is produced by Th2 cells, as well as by nonlymphoid cells (including natural killer cells, mast cells, endothelial cells, and skin keratinocytes) and can act as an immunoregulatory factor. For example, IL-3 (a Th2 cytokine) prevents IL-2 (a Th1 cytokine) from activating killer activity in natural killer cells.

Interleukin-4 is also a product of Th2 T cells, but may be produced by nonlymphoid cells such as fetal trophoblast. IL-4 promotes antibody responses by B lymphocytes, inhibits cell-mediated immunity and natural killer cell activation by IL-2, and contributes to granuloma formation (possibly by acting on macrophages and endothelial cells). Granulomas are largely composed of macrophages and represent a particular type of chronic inflammation that can lead to tissue scarring and, hence, disease.

Interleukin-5, another Th2-type cytokine, promotes the development of B cells which ultimately produce immunoglobulin E, a class of antibody involved in immune responses to certain parasites and in allergies such as asthma and hay fever. IL-5 may arise from eosinophils and basophils (a type of granulocyte) and also stimulates the bone marrow to produce eosinophils, which are involved in allergies; eosinophils contain a 12-kD major basic protein which is highly toxic to certain parasites.

Interleukin-6, also known as β-2 interferon and liver cell growth-stimulating factor, is a Th2-type cytokine produced by a wide variety of cell types including T cells, B cells, and macrophages. IL-6 contributes to B lymphocyte (antibody) responses and stimulates the liver to produce protease inhibitors (such as α1-antirypsin and acute phase proteins) which dampen inflammation. IL-6 may also stimulate fetal trophoblast cell function.

Interleukin-7 is a polypeptide that acts as a growth and maturation factor for precursors of B and T

lymphocytes. IL-7 is produced by stromal cells in the thymus as well as by bone marrow cells, and is particularly important for the development of a type of T cell at mucosal surfaces which express $\gamma\delta$ T cell receptor (TcR) instead of the usual $\alpha\beta$-type receptor that is found on most circulating T cells and T cells in the lymph nodes and spleen. The $\gamma\delta$ TcR allows the T cell to recognize certain types of antigens at mucosal surfaces which pose a threat. Recognition can occur without the need for antigen presentation by a macrophage or dendritic cell, which present the antigen as a peptide fragment in the groove of conventional class I or class II major histocompatibility complex (MHC)-self molecules. Some antigenic peptides are recognized via $\gamma\delta$ TcR in association with atypical class I MHC-type antigens. IL-7 also enhances the generation of IL-2-activated nonspecific killer cells and IL-2 activated antigen-specific cytotoxic T cells, and thus is viewed as a Th1-type cytokine.

Interleukin-8 is actually a group of peptides produced by a variety of cell types, including vascular endothelial cells, macrophages, and fibroblasts. IL-8 activates and recruits polymorphonuclear leukocytes in the inflammatory process, and is involved in the initiation of labor and delivery in pregnant women. IL-8 is not found in the mouse.

Interleukin-9 is produced by a subset of Th2 cells and stimulates hemopoietic cells that develop into red blood cells. Mast cells and B cells may also be a target.

Interleukin-10 is produced by the Th2 subset of T cells as well as by B lymphocytes, macrophages, skin keratinocytes, nonepithelial nonleukocytic cells (stroma) in the endometrium, and fetal trophoblast cells that form the placenta. IL-10 inhibits secretion (and function) of proinflammatory macrophages and activation of vascular endothelium and polymorphonuclear leukocytes (eosinophils, basophils, and neutrophils). Hence, inflammation and clotting are suppressed. IL-10 protects against lethal shock triggered by bacterial endotoxin in sepsis by inhibiting production of proinflammatory (Th1) cytokines and other factors that contribute to low blood pressure.

Interleukin-11 is a factor produced by bone marrow–derived fibroblasts and supports the growth and development of certain types of cells, such as B lymphocytes, neutrophils, and platelet-producing megakaryocytes.

Interleukin-12 is a Th1-type produced by macrophages and some dendritic cells. The p20 form of IL-12 stimulates Th1 lymphocyte development and activation on natural killer–type cells that produce interferon γ. IL-12 may substitute for IL-2 in Th1-type responses. The p40 form of IL-12 acts on macrophages.

Interleukin-13 is produced by activated Th2 T cells as well as by natural killer cells and mast cells. IL-13 inhibits proinflammatory cytokine production induced by bacterial endotoxin. It binds to one of the two types of receptors for IL-4, suppresses production of tumor necrosis factor (TNF)-α, is a chemoattractant for vascular endothelial cells and eosinophils (which it stimulates), and promotes B lymphocyte production of IgE, which is important in allergy.

Interleukin-14 is a Th2-type cytokine produced by microvascular endothelial cells (along with IL-3,7,8,11,15, tumor necrosis factor-α and transforming growth factor-β). It may also be produced by T and B cells, and like other Th2-type cytokines, stimulates B cells.

Interleukin-15 is a Th1-type IL-2-like cytokine that binds to IL-2 receptor but is made primarily by monocytes/macrophages, fibroblasts in joint lining (synovia), and keratinocytes in inflammatory responses. Brain astrocytes and microglia stimulated by IL-1, interferon (INF)-γ and TNF-α may produce IL-15. It is also stimulated by IL-10, and activates INF-γ and IL-4 production in lymphoid cells within mucosal epithelia. It can activate natural killer cells, T lymphocytes, B lymphocytes, and neutrophils (that then produce IL-8).

Interleukin-16 is a proinflammatory (Th1) cytokine produced by T cells, fibroblasts, eosinophils, and mast cells. It stimulates migration of $CD4^+$ T cells (that recognize antigen in association with Class II MHC), upregulates (increases surface expression of) receptors for IL-2, stimulates monocytes, eosinophils, and precursors of B cells.

Interleukin-17 is a Th1-type cytokine produced almost exclusively by memory T cells (defined by their $45RO^+$ surface marker). IL-17 induces production of proinflammatory Th1-type cytokines by a variety of cells, and is involved in angiogenesis (origin and development of blood vessels) and recruitment of neutrophils to tissues.

Interleukin-18 is a Th1-type cytokine similar to IL-12. It is produced by macrophage-type cells and skin cells, and acts to stimulate production of interferon γ by T lymphocytes and natural killer cells.

Interleukin-19 is a product of B cells and macrophages and can act on T cells and macrophages to increase production of IL-4,5,10,13 and TNF-α.

Interleukin-20 is produced by macrophages and skin keratinocytes and acts on keratinocytes.

Interleukin-21 is derived from T cells and stimulates T and natural killer cells.

Interleukin-22 is a Th1-type cytokine produced by Th1 cells and macrophages that stimulates inflammation.

Interleukin-23 is a product of dendritic cells that acts on T cells.

Interleukin-24 is a Th2 type cytokine that can also derive from macrophages, natural killer cells, and fibroblasts. IL-24 stimulates megakaryocytes.

Interleukin-25 is a Th2 cytokine produced from nonlymphoid stroma. It can stimulate certain bone marrow cells not committed to a particular differentiation pathway.

Interleukin-26 is derived from T and natural killer cells and stimulates epithelial cells to make IL-8 and IL-10.

Interleukin-27 is a product of dendritic cells that acts on natural killer and T cells.

Interleukin-28A and B are interferon-related

molecules produced by dendritic cells and exert an antiviral effect.

Interleukin-29 is similar to IL-18A and IL-18B.

Interleukin-30 is a product of antigen-presenting cells (for example, macrophages and dendritic cells) that acts on natural killer and T cells.

Interleukin-31 is a T cell–derived factor that acts on macrophages and epithelial cells.

Interleukin-32 is a product of T and natural killer cells that stimulates macrophages to make TNF-α, IL-8, and other proinflammatory proteins.

Fibroleukin is a soluble fibrinogen-related protein produced by a subpopulation of T cells and by a wide variety of other cell types. Fibroleukin appears to alter maturation of dendritic cells; therefore, its role may be to regulate (that is, suppress) immune responses. It has not yet been given an interleukin number. A membrane-associated form of fibroleukin can act directly as a prothrombinase, leading to clotting. *See* CELLULAR IMMUNOLOGY; FIBRINOGEN; HISTOCOMPATIBILITY; INFLAMMATION.

Colony-stimulating factors. Colony stimulating factors (CSFs) are produced by lymphoid and nonlymphoid cells. These factors provide a mechanism whereby cells that are distant from bone marrow can call for different types of hemopoietic progeny. There are also growth-promoting actions of locally produced CSFs within the bone marrow to stimulate progenitors to differentiate into macrophages (CSF-M = CSF-1), granulocytes (CSF-G), or colonies containing both cell types (CSF-GM). Generation of erythrocyte colonies is stimulated by erythropoietin (EPO), which is produced by the kidney. Certain interleukins may affect particular steps in the growth and differentiation pathways activated by CSFs. Macrophage-lineage progenitors of specialized antigen-presenting dendritic cells that activate or suppress responses of T and B lymphocytes; proliferation of dendritic cell precursors is driven by the cytokine flt3 ligand. CSF-1 may promote Th2 over Th1 type responses. CSF-G is administered to patients receiving cytotoxic anticancer drugs to enhance recovery of blood neutrophils that fight infection. In the uterus, CSFs are produced by epithelium and stromal cells, with epithelium dominating. CSF-1 enhances trophoblast growth, probably via stimulation of macrophages. CSF-GM appears to have a direct effect on implantation, and stimulates preimplantation division of embryonic cells within its shell of zona pellucida. *See* EMBRYONIC DIFFERENTIATION; FERTILIZATION.

Interferons. Interferons classically interfere with the virus-replication mechanism within cells. Interferon-α (produced by leukocytes) and interferon-β (produced by fibroblasts) activate cytotoxicity in natural killer cells. Interferon γ also activates natural killer cells and is a potent activator of macrophages as well. In certain domestic animal species, fetal trophoblast cells of the placenta produce novel types of interferon (interferon tau), which signals the mother to continue to make steroid hormones necessary for the pregnancy to continue. In primates, such as the human, this signaling is mediated by chorionic gonadotropin. IL-28A, and IL-28B and IL-29 have been designated as IFN-λ1, -λ2, and -λ3.

Tumor necrosis factors. Tumor necrosis factor (TNF)-α is produced by a variety of cell types, but activated macrophages represent the dominant source. TNF-α activates natural killer cell cytotoxicity, enhances generation of cytotoxic T lymphocytes (if there is inhibition by transforming growth factor-β, as described below), and activates natural killer cells to produce interferon γ. TNF-α also acts on vascular endothelium to promote inflammation and thrombosis (formation of a blood clot in a blood vessel); these effects may be enhanced by the copresence of IL-1 and interferon γ. TNF-α may also induce apoptosis (programmed cell death), and this also may be enhanced by copresence of interferon γ. Indeed, a synergistic effect of TNF-α and IFN-γ has been described in a variety of disorders; for the cytokines to act (in some cases), a Toll-like receptor (TLR) signal is also required. Toll-like receptors are related to the fruit fly protein *Toll*; they enable cells of the innate immune system to recognize "danger," for example, TLR4 binds bacterial endotoxin.

At low levels, TNF-α may serve as a growth and differentiation signal to trophoblasts and other cell types. Soluble receptors to TNF-α may bind and neutralize biological function. TNF-β is a product of Th1 T cells; in addition to providing help in proinflammatory cell–mediated immune responses, these cells produce delayed-type hypersensitivity reactions, in which macrophages are locally recruited and activated to kill intracellular pathogens, for example, certain bacteria. TNF-β has interferon-type activity and a narrower spectrum of action than TNF-α. Originally, TNF-β was called lymphotoxin as it was produced by lymphocytes and could attack certain types of cancer cells.

Other growth factors. Transforming growth factors (TGFs) have the ability to promote unrestrained proliferation of cells that are otherwise benign. TGFs have therefore been implicated in development of cancer. There are two groups of TGFs, TGF-α and TGF-β. TGF-α is a 5-kD peptide produced by a variety of cells. It collaborates with TGF-β, a 25-kD peptide, in promoting unrestrained tumorlike growth. Fibroblasts with a benign phenotype must attach to a surface to grow in culture (usually to a plastic disk), but in the presence of TGF-α and TGF-β, some of the cells (if suspended in an agar gel) will grow into colonies. This "anchorage-independent" growth is an identifying property of transformed/malignant cancer cells.

Epidermal growth factor, produced by the salivary gland, can substitute for TGF-α. TGF-β has potent pleiotropic effects on a wide variety of tissues and is a potent fibrogenic and immunosuppressive agent. Classical TGF-βs in mammals are divided into β1, β2, and β3, and have some functional differences. For example, TGF-β2 is less inhibitory to developing bone marrow cells compared with TGF-β1, and TGF-β3 may inhibit scar formation (fibrogenesis) induced by TGF-β1. There are other TGF-βs produced by tumor

cells and found in the uterus in pregnancy that differ in molecular size and certain functional properties.

More distantly related molecules such as activin and inhibin are involved in regulation of the endocrine and reproductive systems. Insulin, produced by β cells in the islets of the pancreas, is a growth factor for many cells, and belongs to the endocrine system. Related are insulinlike growth factors (somatomedins) which may be produced by a variety of cell types. A distinct fibroblast growth factor and platelet-derived growth factor have also been defined. *See* GROWTH FACTOR; TUMOR.

Chemokines. Chemokines are chemoattractant cytokines of small size (7–14 kD); they are subdivided into four families: CXC, CC, C, and CX_3C. Chemokines are produced by macrophages stimulated by bacterial endotoxins, and they control the nature and magnitude of cell infiltration in inflammation. They may affect T cell subsets, monocytes, natural killer cells, basophils, neutrophils, eosinophils, dendritic cells, and mast cells. The mouse chemokine known as MIP-1α is the functional equivalent of IL-8 in the human.

Wound healing. Wound healing is probably the most common phenomenon in which the importance of cytokines is seen. Cytokines ensure that the restorative sequences are carried out in the appropriate order by signaling blood cells and vascular endothelium to coagulate and fill in a wound opening, recruiting and signaling macrophages and neutrophils to engulf microbes, and guiding protective skin epidermal cells to grow over the wounded area (a TGF-β effect). If the damage is more extensive, cytokines stimulate production of new skin cells, blood vessels, connective tissue, and bone. It has been suggested that TGF-β, which inhibits certain enzymes (proteases) that break down tissues, may promote more rapid healing of surgical wounds. Such treatment could reduce the postsurgery recovery period of many individuals and animals. Inhibiting excessive scar formation will be beneficial in plastic surgery. Manipulating the levels of specific cytokines in the body is likely to enhance our ability to treat many different types of illnesses in the future. The immunosuppressive cytokines may be particularly helpful in more successful transplantation of organs (kidney, heart, liver, skin). *See* IMMUNOSUPPRESSION.

David A. Clark

Bibliography. A. Angeli et. al., The overtraining syndrome in athletes: A stress-related disorder, *J. Endocrin. Invest.*, 27:603–612, 2004; S. P. Cobbold, T cell tolerance in transplantation: Possibilities for therapeutic intervention, *Exp. Opin. Therap. Targets*, 6:583–599, 2002; C. Deans and S. J. Wigmore, Systemic inflammation, cachexia and prognosis in patients with cancer, *Curr. Opin. Clin. Nutrit. Metab. Care*, 8:265–269, 2005; S. T. Holgate, Cytokine and anti-cytokine therapy for treatment of asthma and allergic disease, *Cytokine*, 28:152–157, 2004; J. Imitola, T. Chitnis, and S. J. Khoury, Cytokines in multiple sclerosis: From bench to bedside, *Phar. Therap.*, 106:163–177, 2005; Y. Matsuzawa, Adipocytokines: Emerging therapeutic targets, *Curr. Atherosc. Rep.*, 7:58–62, 2005; O. C. Stokkers and D. W. Hommes, New cytokine therapeutics for inflammatory bowel disease, *Cytokine*, 28:167–173, 2004; L. Ulloa and K. J. Tracey, The "cytokine profile": A code for sepsis, *Trends Mol. Med.*, 11:56–63, 2005.

Cytokinins

A group of plant hormones (phytohormones) that, together with other plant hormones, induces plant growth and development. The presence of an unknown compound in the phloem that stimulated cell proliferation of potato tubers was demonstrated in the early 1900s by G. Haberlandt. In the 1940s J. van Overbeek found that the milky endosperm from immature coconuts is rich in compounds that promote cell proliferation. Then, in 1954, during investigations of the nutritional requirements of tobacco callus tissue cultures, F. Skoog and C. O. Miller discovered a very active substance that stimulated cell division and was formed by the partial breakdown of aged or autoclaved herring sperm deoxyribonucleic acid (DNA). The compound was named kinetin because it induced cytokinesis of cultured tobacco cells. Kinetin has not been found in plants, and it can be formed from a mixture of autoclaved furfuryl alcohol and adenine. In 1961 Miller obtained an active substance that stimulates cell division from the milky endosperm of corn (*Zea mays*) seeds; in 1964 D. S. Letham, identified this active compound and named it zeatin. Cytokinin is now the accepted term for those naturally occurring and synthetic compounds that promote the cell division (cytokinesis) and growth of various plant tissues.

Analogs. The discovery of kinetin stimulated the synthesis of over 150 different species of kinetin analogs. Tobacco pith tissue culture bioassays indicate that most of the highly active cytokinins are N6-substituted adenines containing 5-C alkyl side chains (**Fig. 1**). Two of the most potent synthetic cytokinins are kinetin and N6-benzyl adenine. *N,N′*-diphenylurea and its derivatives are also considered cytokinins. These compounds induce cell division and growth but are yet to be confirmed as naturally occurring cytokinins.

Since the isolation of the first naturally occurring cytokinin, zeatin, more than 25 different cytokinins have been isolated from plants. A few cytokinin bases such as isopentenyladenine and *cis*-zeatin have been found in transfer ribonucleic acid (tRNA) of various organisms. The most active form of cytokinin is *trans*-zeatin isomer; the *cis* isomer of zeatin, on the other hand, shows little cytokinin activity.

Each cytokinin can exist as a free base, that is, a nucleoside in which a ribose group is attached to the nitrogen atom of position 9 or a nucleotide in which phosphate is esterified to the 5′ carbon of ribose. In most cases, bases have been shown to be more active than their corresponding nucleosides or nucleotides in bioassays. However, it is still not clear

what the number of active forms of cytokinins may be, because cytokinin bases can be converted to the corresponding nucleosides and nucleotides in plant cells within a short period of time. *See* BIOASSAY.

A group of cytokinin analogs has been reported to be highly potent inhibitors of the cytokinin-requiring callus tissues. On the basis that growth inhibition was abolished or reduced in the presence of higher concentrations of cytokinin, these analogs are named anticytokinins or cytokinin antagonists. One of the most potent anticytokinins is 3-methyl-7-(pentylamino)pyrazolo[4,3-*d*]-pyrimidine (**Fig. 2**).

Biosynthesis and metabolism. Since certain tRNA species contain cytokinins, it has been suggested that turnover of tRNA serves as a source of free cytokinins, but it is now generally accepted that tRNA turnover is not a major source of free cytokinin under normal conditions. Indeed, it has been demonstrated that plant tissues contain enzymes that are capable of the direct synthesis of free cytokinins. The enzyme Δ^2-isopentenylpyrophosphate: AMP-Δ^2-isopentenyl- transferase, or isopentenyltransferase, which has been isolated from various organisms, catalyzes the formation of Δ^2-isopentenyladenosine-5′-monophosphate, a cytokinin nucleotide, from Δ^2-isopentenylpyrophosphate and adenosine-5′-monophosphate. Once the cytokinin nucleotide is formed, other enzymes in plant cells convert it to the corresponding nucleoside and base. Oversynthesis of cytokinins in plant tissue causes abnormal growth: crown gall tumor disease caused by the bacterium *Agrobacterium tumefaciens* is an example of excessive production of cytokinins in local tumor tissue. Tissue from crown gall tumors can grow on a simple medium lacking plant hormones because the tumor tissue overproduces both cytokinin and auxin. This is due to the insertion of a piece of bacterial plasmid DNA into the plant nuclear genomes causing activation of the gene responsible for the regulation of cytokinin production. Roots have been shown to be the major site of cytokinin biosynthesis, but stems and leaves are also capable of synthesizing cytokinins. It is possible that all actively dividing cells are capable of cytokinin biosynthesis. *See* CROWN GALL.

Cytokinins applied to plant tissues are rapidly converted to a variety of metabolites by enzymes which catalyze side-chain modifications, side-chain cleavage, or conjugation of cytokinins with sugars or amino acids. The formation of glycosides and amino acid conjugates is a very common metabolic response of plant tissues to externally applied cytokinins.

Detection. The qualitative and quantitative aspects of cytokinin extracted from plant tissues vary with the types of extraction solvent and the procedure employed, but, in general, the purification and identification of naturally occurring cytokinins has been facilitated by the development of high-performance liquid chromatography, gas chromatography–mass spectrometry, and monoclonal antibodies. The levels of the major cytokinins in plant tissues range from 0.01 to 150 nanomoles per 100 g of fresh tissues.

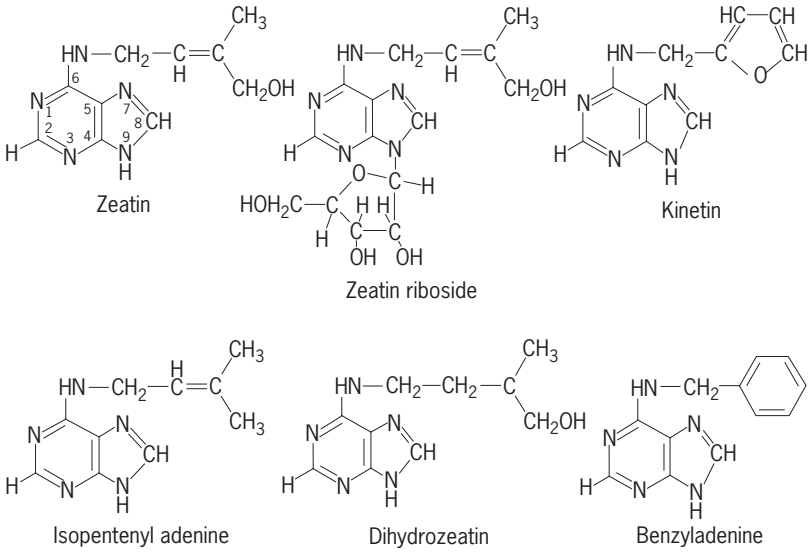

Fig. 1. Structures of naturally occurring (zeatin, zeatin riboside, isopentenyl adenine, and dihydrozeatin) and synthetic (kinetin and benzyladenine) cytokinins. All of these compounds are adenine derivations in which the purine ring is numbered as shown for zeatin. Zeatin can exist in the trans (as shown) or cis (with CH_3 and CH_2OH groups interchanged) configuration.

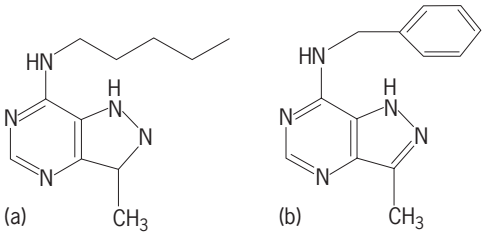

Fig. 2. Structures of two synthetic anticytokinins: (*a*) 3-methyl-7-(pentylamino)pyrazolo(4,3-*d*)pyrimidine and (*b*) 3-methyl-7-(benzylamino)pyrazolo(4,3-*d*)pyrimidine.

However, most of the reported cytokinin levels in plant tissues are rough estimates, because the various forms of cytokinins are not accurately measured. Certain cytokinins are also present in tRNAs of plants, animals, and microorganisms.

Roles of the cytokinins. Cytokinins exhibit a wide range of physiological effects when applied externally to plant tissues, organs, and whole plants. Exogenous applications of this hormone induce cell division in tissue culture in the presence of auxin. The formation of roots or shoots depends on the relative concentrations of auxin and cytokinin added to the culture medium. High auxin and low cytokinin concentrations lead to root formation, while low auxin and high cytokinin concentrations give shoots. Tissue culture techniques have been employed by plant biotechnologists to grow genetically engineered plant cells into whole plants. Cytokinins appear to be necessary for the correlated phenomena of mitosis and nucleic acid synthesis. Cytokinins delay the aging process in detached leaves by slowing the loss of chlorophyll. Retention of chlorophyll is accompanied by chlorophyll synthesis. Exogenously applied cytokinins are effective on fresh flowers, vegetables, and fruits as post-harvest preservatives; however, the use of cytokinins for the preservation of plant materials is not practiced widely.

Cytokinin effects also include breaking of dormancy, promotion of seed germination, stimulation and nutrient mobilization, enhanced anthocyanin and flavanoid synthesis, increased resistance to disease, and stimulation of the opening of stomates. Under some conditions, cytokinin enhances shoot branching and the expansion of cells in leaf disks and cotyledons. In a mammalian system, cytokinin nucleotides have been shown to suppress leukemic cell growth, but the mechanism by which this occurs is unknown. Certain cytokinins located adjacent to the anticodon of tRNA have been shown to influence the attachment of amino acyl tRNA to the ribosome–messenger ribonucleic acid (mRNA) complex and may be involved in protein synthesis. *See* AUXIN; DORMANCY; PLANT PHYSIOLOGY; TISSUE CULTURE.

Cytokinin-binding molecules. It is a widely held view that there are macromolecules such as proteins which interact between cytokinin molecules and the cytokinin-responsive genes. According to this hypothesis, the initial event of hormone action involves the binding of a hormone to a receptor, with the hormone–receptor complex then binding to a specific region of gene. There have been numerous papers reporting the isolation of cytokinin-specific and relatively high-affinity cytokinin-binding proteins. However, the physiological role of any of the reported cytokinin-binding proteins is unknown, and these proteins cannot yet be considered to be cytokinin receptors.

Cytokinin-modulated gene expression. The mechanism of action of cytokinin on plant growth and development is poorly understood. It has been demonstrated that specific proteins are induced, enhanced, reduced, or suppressed by the hormone. Some of the enzymes or proteins regulated by cytokinins have been identified. The activities of acid phosphatase, ribulose 1,5-bisphosphate carboxylase, invertase, light-harvesting chlorophyll *a*/*b*-binding protein, and hydroxypyruvate reductase are all enhanced by the hormone, but the details of how these enzymes are stimulated by cytokinins remain to be studied. Applications of cytokinin to cytokinin-responsive tissues also regulate the synthesis of specific mRNAs; some of the mRNAs are induced and some are suppressed. By using modern techniques, plant molecular biologists are able to show that mRNA changes occur within 60 min of cytokinin application to plant cells. Initial evidence suggests that cytokinins regulate, at least in part, the transcriptional process of gene expression by turning on or off specific genes and stimulating or suppressing the synthesis of specific mRNAs. Another possible action of cytokinins is the regulation of a posttranscriptional process such as stabilization of mRNA. Scientists also demonstrated that cytokinins specifically increase the rate of protein synthesis and the effect seems to be on the translation of mRNA into proteins. *See* PLANT GROWTH; PLANT HORMONES.

Chong-maw Chen

Bibliography. S. D. Kung and C. J. Arntzen (eds.), *Plant Biotechnology*, 1988; D. S. Lethan and L. M. S. Palni, The biosynthesis and metabolism of cytokinins, *Annu. Rev. Plant Physiol.*, 34:163–197, 1983; H. F. Linskens and J. F. Jackson (eds.), *Modern Methods of Plant Analysis, New Series*, vol. 5, 1987; C. O. Morris, Genes specifying auxin and cytokinin biosynthesis in phytopathogens, *Annu. Rev. Plant Physiol.*, 37:509–538, 1986; F. Skoog and D. J. Armstrong, Cytokinins, *Annu. Rev. Plant Physiol.*, 21:359–384, 1970; M. B. Wilkins (eds.), *Advanced Plant Physiology*, 1984.

Cytolysis

An important immune function involving the dissolution of certain cells. There are a number of different cytolytic cells within the immune system that are capable of lysing a broad range of cells. The most thoroughly studied of these cells are the cytotoxic lymphocytes, which appear to be derived from different cell lineages and may employ a variety of lytic mechanisms. Cytotoxic cells are believed to be essential for the elimination of oncogenically or virally altered cells, but they can also play a detrimental role by mediating graft rejection or autoimmune disease. There are two issues regarding cytotoxic lymphocytes that are of concern: one is the target structure that is being recognized on the target cell, that is, the cell that is killed, which triggers the response; and the other is the lytic mechanism. *See* CELLULAR IMMUNOLOGY.

Cytotoxic lymphocytes. When freshly isolated, large granular lymphocytes from peripheral blood are tested in cytotoxicity assays, they spontaneously lyse certain tumor cells. These cytotoxic cells are called natural killer cells, and they are important mediators of innate immunity as a first line of defense against invading pathogens. They are unique in that no previous sensitization is required for them to kill. It now appears that a number of different receptors on natural killer cells are capable of activating the lytic machinery. Recently, it has been found that these cells also express inhibitory receptors that actually inhibit cell lysis, thus adding another level of complexity to the regulation of cytolysis by these cells.

Another killer cell, called the lymphokine-activated killer cell, is a product of culturing large granular lymphocytes in relatively high concentrations of certain growth factors. They lyse any target cell, including cells from freshly isolated tumors, and are being employed in cancer therapy. Lymphokine-activated killer cells may also be important in mounting a vigorous response under conditions of extreme immunological stress. Very little is known about the mechanisms by which these cells recognize and lyse the target cell.

The last group of cytotoxic cells is the cytotoxic T lymphocyte. These are T cells that can lyse any target cell in an antigen-specific fashion. That is, as a population they are capable of lysing a wide range of target cells, but an individual cytotoxic T lymphocyte is capable of lysing only those target cells which bear the appropriate antigen. Cytotoxic T lymphocytes recognize peptide antigens in the context of major

histocompatibility complex class I antigens. These are truly immune cells in that they require prior sensitization in order to function. These cells are thought to mediate graft rejection, mount responses against viral infections and intracellular bacterial infections, and play a major role in tumor destruction.

Cytolytic cycle. Cytotoxic cell-mediated lysis is divided into three distinct steps (see **illus.**). The first step is conjugation, when the killer cell determines if the target cell expresses the appropriate antigen and binds to it via a complex array of adhesion molecules. The second step involves the programming for lysis in which the lytic event is triggered. The third step is the destruction of the target cell. This step occurs after the cytotoxic T lymphocyte has broken free from the target cell. The cytotoxic cell is then free to repeat multiple cycles of killing. The lytic cycle is extremely rapid, and the process can be completed within 20–30 min after mixing the target with the killer cell.

Any proposed explanations of the mechanism of cytolysis must take the above steps and the following observations into consideration. It is clear that direct cell contact between the target cell and the killer cell is absolutely required for initiating the lytic mechanism; if cell contact is prevented, no killing occurs. Another interesting aspect of this process is that the killer cells remain unscathed during the lytic event, suggesting that the killer cell must either employ a unidirectional lytic mechanism or be resistant to the lytic mechanism. Also, when many, but not all, target cells die after cytotoxic T-lymphocyte interaction, nuclear damage with rapid deoxyribonucleic acid (DNA) fragmentation precedes detectable plasma membrane damage. *See* ANTIGEN; HISTOCOMPATIBILITY.

It is becoming clear that cytotoxic lymphocytes employ multiple mechanisms designed to initiate target cell destruction. If cytolysis exists to protect the organism from invading pathogens, there should be redundancies in the system so that, if the pathogen has a mechanism for escaping one cytolytic pathway, alternative mechanisms would still be functional. For example, some viruses have evolved strategies to evade the immune system, including the specific inhibition of certain cytolytic pathways. Therefore, there may be a number of ways in which killer cells can trigger cell death in the target cells that have yet to be recognized. Two mechanisms, which are the best studied to date, are the degranulation of cytolytic granules and the triggering of death receptors found on target cells.

Cytolytic granules. In the degranulation mechanism of killing, cytotoxic cells release the contents of cytotoxic granules after specific interaction with the target cell (see illus.). After target cell binding, the Golgi apparatus, which is an organelle that prepares proteins for secretion, and the granules containing the cytolytic molecules orient toward the target cell in preparation for secretion. Vectorally oriented granule exocytosis then occurs, with the release of cytolytic molecules into the intercellular space between the killer and target cells. This results

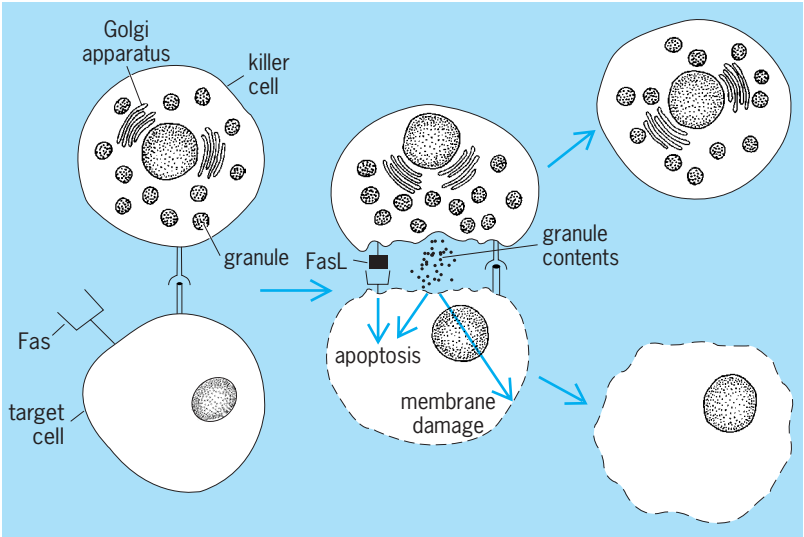

Lytic cycle. The killer cell binds to the appropriate target cell, triggering the expression of FasL and the reorientation of the Golgi apparatus and the cytolytic granules toward the target cell. The cytolytic granules then fuse with the external membrane, releasing the granule contents, including perforin and granzyme B, into the intercellular space. The granule contents trigger apoptosis and membrane damage. Death receptors, such as Fas, expressed on the target cell are also engaged, leading to apoptosis. After the lytic pathways are initiated in the target cell, the killer cell releases from the target cell and goes on to repeat the entire process.

in the leakage of salts, nucleotides, and proteins from the target cell, leading to cell death. *See* GOLGI APPARATUS.

The cytolytic granules contain a variety of proteins, including a well-characterized protein called perforin and a group of enzymes called serine proteases. It is thought that the perforin can polymerize into a pore-forming complex, perhaps to act as a channel that contributes to the redistribution of small molecules leading to cell death; however, direct evidence that the formation of a pore plays a role in cell-mediated cytotoxicity is lacking. Mice deficient in perforin have been generated, and the killing capacity of killer cells isolated from these mice is severely compromised, as are the immune responses against some pathogens. Although the exact function of this protein remains to be defined, these results demonstrate that perforin is a critical component in the cytolytic mechanism.

Serine proteases, the best characterized being granzyme B, are also released from the cytolytic granules and play an important part in triggering a cellular suicide process referred to as apoptosis or programmed cell death. In apoptosis, a cytotoxic cell triggers an endogenous suicide pathway within the target cell, resulting in nuclear disintegration, DNA fragmentation, and mitochondrial dysfunction. The target cell plays an active role in its own demise, which is consistent with much of the data available on cytotoxic T-lymphocyte-mediated killing.

Upon entering the target cell, granzyme B cleaves other proteases called caspases, triggering a cascade of enzymatic reactions leading to changes in the mitochondria and plasma membrane and DNA fragmentation, ultimately culminating in cell death. It is not exactly understood how granzyme B enters the cell

to initiate the caspase cascade. It was initially thought that perforin formed pores through which the proteases entered to initiate the death cascades; however, a number of observations are inconsistent with this hypothesis, and recent evidence suggests that additional mechanisms for serine protease entry into the cells may exist. Mice made deficient of granzyme B have compromised killer-cell-dependent immune responses, indicating that granzyme B is important for cell-mediated cytotoxicity.

Although it is not clear how the granule contents cause cell death, substantial evidence suggests that these granules play an important role in mediating target cell lysis by killer lymphocytes by both apoptotic and nonapoptotic pathways. Future research efforts will be focused on determining exactly how the granule contents, and specifically granzyme B and perforin, cause cell death in the target cell.

Death receptors. On virtually all cells of the body a number of receptors have been identified that are able to trigger apoptosis when engaged. These receptors are called death receptors. The most studied, and relevant to cytolytic cells, is a receptor called Fas. Fas, when engaged by its ligand (FasL), triggers the caspase cascade leading to the hallmark signs of apoptosis, which include membrane blebbing, chromosomal condensation, nuclear disintegration, DNA fragmentation, and cell death. Thus, the cellular suicide pathway triggered by Fas engagement is similar in some respects to the pathways triggered by granzyme B. If FasL were expressed constitutively on killer cells, this would mean that the killer cell could indiscriminately lyse cells; however, FasL is expressed only after the target cell has been recognized. Since FasL is very rapidly expressed (for example, after the T-cell receptor for antigen is engaged), Fas is triggered only on the target cell that has been identified as a cell destined for death. Fas is expressed on essentially all cells of the body, indicating that all cells can be easily triggered to undergo cell death by the appropriate killer cell. It makes sense for all of the cells in the body to have a built-in self-destruct mechanism so that if a cell becomes pathogenically altered it can be easily triggered to die.

There appears to be no single mechanism that killer cells use to eliminate tumor cells or infected cells. The belief that killer cells employ multiple mechanisms is consistent with the plasticity of the immune system and is important given the wide range of pathogenic organisms that the body must eliminate. Unfortunately, these killer cells may also play a role in certain autoimmune diseases and graft rejection after transplantation. Understanding how these killer cells function and are regulated will be essential for developing therapies to either enhance advantageous or inhibit deleterious cytotoxic lymphocyte responses. *See* IMMUNOLOGY.

H. L. Ostergaard; M. M. Barry

Bibliography. M. Barry and G. McFadden, Apoptosis regulators from DNA viruses, *Curr. Opin. Immunol.*, 10:422–430, 1998; A. J. Darmon and R. C. B. Bleackley, Proteases and cell-mediated cytotoxicity, *Crit. Rev. Immunol.*, 18:255–273, 1998; E. Meinl et al., Anti-apoptotic strategies of lymphotropic viruses, *Immunol. Today*, 19:474–479, 1998; M. E. Peter and P. H. Krammer, Mechanisms of CD95 (APO-1/Fas)-mediated apoptosis, *Curr. Opin. Immunol.*, 10:545–551, 1998; S. Shresta et al., How do cytotoxic lymphocytes kill their targets?, *Curr. Opin. Immunol.*, 10:581–587, 1998; J. A. Trapani, V. R. Sutton, and M. J. Smyth, CTL granules: Evolution of vesicles essential for combating virus infections, *Immunol. Today*, 20:351–356, 1999.

Cytomegalovirus infection

A common asymptomatic infection caused by cytomegalovirus, which can produce life-threatening illnesses in the immature fetus and in immunologically deficient subjects.

The virus. Cytomegalovirus is a member of the herpesvirus group, which asymptomatically infects 50–100% of the normal adult population. Such infections usually take place during the newborn period when the virus can be transmitted from the mother to the baby if the virus is present in the birth canal or in breast milk. Toddlers may also acquire the infection in nurseries. Later in life, the virus may be transmitted by heterosexual or male homosexual activity. After infection, cytomegalovirus remains latent in the body because it cannot be completely eradicated even by a competent immune system. It may be activated and cause illnesses when there is a breakdown of the immune system.

Pathogenesis. Congenital or transplacental cytomegalovirus infection, like asymptomatic infection after birth, is also a fairly common event. With rare exceptions, it too is usually asymptomatic. Congenital cytomegalovirus disease was the first disease recognized to be caused by this virus. It results from transplacental transmission of the virus usually from a mother undergoing initial or primary cytomegalovirus infection during pregnancy. Its manifestations range from subtle sensory neural hearing loss detectable only later in life, to a fulminating multisystem infection and eventual death of the newborn. This important congenital disease occurs in about 1 in 1000 pregnancies.

The only cytomegalovirus illness clearly described in mature, immunologically normal subjects is cytomegalovirus mononucleosis. This is a self-limited illness like infectious mononucleosis, the main manifestation of which is fever. *See* INFECTIOUS MONONUCLEOSIS.

Otherwise, cytomegalovirus illnesses are usually seen only when cellular immunity is deficient. They constitute the most important infection problem after bone marrow and organ transplantations. Manifestations vary from the self-limited cytomegalovirus mononucleosis to more serious organ involvement such as pneumonia, hepatitis, gastrointestinal ulcerations, and widespread dissemination. The virus causing these illnesses may come from activation of the patient's own latent infection, or it may be transmitted from an outside source, usually from latent

cytomegalovirus infecting the graft from a donor. *See* IMMUNOLOGICAL DEFICIENCY; TRANSPLANTATION BIOLOGY.

Cytomegalovirus illnesses are also serious, fairly frequent complications of the acquired immunodeficiency syndrome (AIDS). One reason is that most individuals with human immunodeficiency virus (HIV) infection are already infected with cytomegalovirus. Disease manifestations are similar to what is seen in transplant cases, except they may be more severe. Cytomegalovirus retinitis is a typical problem associated with advanced AIDS. Without treatment, the retina is progressively destroyed such that blindness of one or both eyes is inevitable. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS).

Treatment. Cytomegalovirus diseases can be treated with two antivirals, ganciclovir or foscarnet, with varying degrees of success. Cytomegalovirus pneumonia in the bone marrow transplant recipient cannot be cured by antivirals alone because it probably has an immunopathologic component. Cytomegalovirus diseases in persons with AIDS can be contained but not cured by specific treatment. For example, ganciclovir treatment of cytomegalovirus retinitis is effective only as long as maintenance therapy is continued. *See* ANIMAL VIRUS. Monto Ho

Bibliography. Y. Becker, G. Darai, and E. S. Huang (eds.), *Molecular Aspects of Human Cytomegalovirus Diseases*, 1992; M. Ho, *Cytomegalovirus: Biology and Infection*, 2d ed., 1991.

Cytoplasm

That portion of living cells bordered externally by the plasma membrane (cell membrane) and internally by the nuclear envelope. In the terminology of classical cytology, the substance in living cells and in living organisms not compartmentalized into cells was called protoplasm. It was assumed at the time that the protoplasm of various cells was similar in structure and chemistry. Results of research on cell chemistry and ultrastructure after about 1960 showed that each cell type had a recognizably different "protoplasm." Primarily for that reason, the term protoplasm gradually fell into disuse in contemporary biology. The terms cytoplasm and nucleoplasm have been retained and are used descriptively; they are used almost synonymously with the terms cytosome (body of cytoplasm) and nucleus, respectively.

Observation by optical microscopy. Optical microscopes are limited in resolving power at the level of about 0.2 micrometer. Therefore only the larger structures in the cell could be observed and named by early cytologists. Structures visible in the light microscope include the nucleus, with its chromosomes and nucleolus, and the cytoplasm, with its larger organelles such as chloroplasts (in plants only), mitochondria, Golgi bodies, and centrosome (containing two centrioles), as well as inclusions such as lipid droplets, pigment granules, yolk (in eggs), starch grains, and vacuoles. The main impression that early microscopists gained from observations of cells with an ordinary bright-field microscope was that cells were transparent and replete with particles of different types, and that many exhibited motion of some kind, usually Brownian motion (thermally induced collisions between molecules and granules in the cell).

Modern optical microscopes operate in modes that produce vastly improved contrast. Dark-field and phase-contrast microscopy greatly increases the visibility of cytoplasmic particles of all kinds. Interference microscopes show the shapes of cells by representing differences in optical path (refractive index difference times the thickness) in brightness contrast or color contrast. Differential interference microscopes render all cytoplasmic surfaces shadowcast, and provide a more detailed view of the various interfaces in cells caused by the existence of organelles and inclusions, vacuolar membrane systems, and various kinds of filamentous structures. Polarizing microscopes reveal ordered or crystalline structures (such as the mitotic spindle, stress fibers, and myofibrils) when they are doubly refracting, or birefringent. Fluorescence microscopes reveal naturally fluorescent substances (fluorochromes) such as chlorophyll in plants, or fluorchrome-labeled antibodies to specific protein molecules. These optical probes have made it possible to discover the location of enzymes, hormones, and structural proteins in the cytoplasm, sometimes even while cells are alive. *See* MICROSCOPE.

Regional differentiation. Many cells, especially the single-celled organisms or protistans, have regional cytoplasmic differentiation. The outer region is the cortex or ectoplasm, and the inner region is the endoplasm. In some echinoderm eggs, for example, endoplasmic organelles and inclusions can be sedimented at moderate centrifugal accelerations, while cortical granules that are denser than most other particles remain fixed in the cortex. In many cases the cortical layer is a gel made up of a meshwork of cytoskeletal fibers, especially F-actin. In many ciliates the cortical layer contains organelles peculiar to that region. For example, the cortex of *Paramecium* contains the basal bodies of cilia and a system of fibrils associated with them, as well as rows of trichocysts which are fired like tiny javelins when the *Paramecium* is injured. In amebae, movement requires that the cell maintain an outer ectoplasmic gel surrounding the more fluid flowing endoplasm, which nevertheless also has some gellike properties.

Apparently the cytoplasm of nearly all cells is contractile when stimulated. In some cells the machinery of contraction is localized in the cortex (for example, the ciliates *Stentor, Spirastomum*, and *Vorticella;* the flagellates *Euglena* and *Peranema;* the testacean *Difflugma;* and many sporozoans).

Centrifugal sedimentation of cytoplasm. The organelles and inclusions of some cells can be caused to sediment into layers based on their density by the application of centrifugal acceleration. Particles with high density are sedimented to the centrifugal pole, while less dense (buoyant) particles are moved to the centripetal pole. In classical experiments by

E. N. Harvey and E. B. Harvey in the 1930s, it was shown that eggs of the sea urchin *Arbacia* could be stratified while under observation on a cushion of sucrose solution in a special microscope-centrifuge, actually a centrifuge with the microscope built into it. If the eggs were centrifuged while in a sucrose density gradient, they were pulled into two halves by the force caused by the centrifugal acceleration acting upon the density difference between the upper and lower halves of the egg and the local density in the gradient.

Centrifugation experiments clearly showed that organelles and inclusions fall into classes differentiated by their density, and focused attention on the properties of the ground cytoplasm (or hyaloplasm) in which the organelles and inclusions were bathed. The early centrifugation experiments led to later efforts directed at measuring the consistency of the ground cytoplasm, and to isolation and purification of fractions of cytoplasmic particles. *See* CENTRIFUGATION.

Observation by electron microscopes. The fine structure of cells could not be studied until the development of electron microscopes and a technique for cutting very thin sections (slices) of cells embedded in a suitable plastic. The first ultramicrotome for slicing cells embedded in plastic was developed by K. R. Porter. The scanning electron microscope (SEM) permits a detailed view of the cell surface or into a cell that has been opened or rendered partially transparent by treatment with electrically conductive fixative compounds. The resolution of the SEM is limited at between 2 and 5 nanometers, depending on instrument design. *See* SCANNING ELECTRON MICROSCOPE.

The conventional transmission electron microscope (CTEM or TEM) is used to produce images and electron micrographs of thin sections (about 40–60 nm thick). The resolving power of the TEM is about 0.2 nm, about a thousand times greater than any light microscope.

The high-voltage electron microscope (HVEM) is similar in design to the TEM but operates with an electron beam of $1–3 \times 10^6$ eV, about 10 times greater voltage than the CTEM. The higher penetrating power and reduced specimen damage at higher voltages make it possible to obtain stereo images of thick sections or whole cells up to 2–5 micrometers in thickness. The HVEM has been very important in obtaining a comprehensive view of the relationships of parts of cells not appreciated from thin sections. *See* ELECTRON MICROSCOPE.

Cytoplasmic ultrastructure. Electron microscopy revolutionized concepts of cell structure and function. Organelles too small to be seen clearly with light microscopes were discovered, named, and characterized after 1955. Organelles previously known were characterized as to their detailed inner morphology.

Centrioles. The two centrioles within the centrosome or central body of the cytoplasm appear as barely visible spherical dots in the optical microscope. In the electron microscope the two centrioles can be seen to be short bundles of 27 tubules arranged in a characteristic pattern, each bundle perpendicular to the other. Centrioles are present in most animal cells and absent in most plant cells, and are organizing centers for the microtubular structures in the mitotic spindles of dividing animal cells and in cilia and flagella. The basal bodies of the latter are related to centrioles. *See* CENTRIOLE.

Endoplasmic reticulum. In cells known to synthesize proteins for secretion (for example, cells of the pancreas), the cytoplasm just outside the nucleus was known from optical microscopy to be basophilic. This region of the cell, called the ergastoplasm, has an affinity for basic stains long known to mark the presence of nucleic acids. In the electron microscope, this region of the cell was found to contain massive amounts of rough endoplasmic reticulum consisting of stacks of flat membranous cisternae covered with ribosomes, particles of ribonucleoprotein which play an important role in protein synthesis. Face-on views of the endoplasmic reticulum show the ribosomes in lines connected by delicate strands of messenger ribonucleic acid, macromolecules that serve as an intermediate carrier of genetic information for the synthesis of specific protein molecules. The rough endoplasmic reticulum has been shown to be a temporary repository for proteins to be exported from the cell or incorporated into membranes. Smooth endoplasmic reticulum is more likely to be tubular than flat, and lacks attached ribosomes; it is involved in the synthesis and storage of other macromolecules. *See* ENDOPLASMIC RETICULUM; RIBOSOMES.

Golgi complex. The Golgi complex is another class of organelles consisting of stacks of membranous sacs or cisternae that are slightly curving in shape. The Golgi complex serves to package proteins that the cell will secrete. The edges of the Golgi cisternae break off and form Golgi vesicles containing the secretions. *See* GOLGI APPARATUS.

Mitochondria. Mitochondria, usually rod-shaped particles 0.5–3 μm long, consist of two membranes, the inner one thrown into folds called christae, on which are located assemblies of respiratory pigments and enzymes. Mitochondria are often spoken of as the powerhouses of the cell. Their matrix surrounding the inner membrane folds contains the enzymes that catalyze the oxidative phase of glycolysis (carbohydrate breakdown). Mitochondria each contain a circular strand of deoxyribonucleic acid (DNA), which contains the genetic information coding for some, but not all, of the mitochondrial enzymes and structural proteins. *See* MITOCHONDRIA.

Lysosomes. Lysosomes are cytoplasmic particles in the same size range as mitochondria, and can be recognized by an electron-dense matrix with no folded inner membrane. Lysosomes differ in density from mitochondria, and thus can be separated from them by density gradient centrifugation. Lysosomes contain several hydrolytic enzymes that can catalyze the breakdown of various classes of macromolecules. When cells ingest foreign material (endocytosis), lysosomes fuse with the vacuole containing the ingested material and supply enzymes required for its

intracellular digestion. *See* ENDOCYTOSIS; LYSOSOME.

Microbodies. Another class of cytoplasmic particles are the microbodies, consisting of peroxisomes and glyoxisomes, particles that contain different groups of oxidative enzymes.

Chemical deposits. Cells contain inclusions that are not functional entities but, rather, deposits of certain chemicals. Many cells contain small droplets of fat or oil that are buoyant and can be displaced by centrifugation to the centripetal pole of the cell. Certain large amebas contain bipyrimidal or flat crystals of triuret, a nitrogen excretion product of purine and pyrimidine metabolism. In the same cells are many so-called refractile bodies containing calcium and magnesium polyphosphate compounds. Plant cells contain starch particles which can be broken down to glucose to obtain chemical energy; many animal cells contain glycogen deposits that play a similar role. The inclusions found in various cell types are more numerous and varied than the organelles. They can be removed by centrifugation and microsurgery without harming the cell, and are therefore nonessential to life.

Role in cell division. When the cell divides, a new organelle, the mitotic spindle, is formed, which divides the replicated chromosomes into equal groupings for the daughter nuclei to be formed. In some cells the mitotic spindle forms, at least in part, while the chromosomes are still within the nuclear envelope. In other cells the spindle itself is entirely intranuclear.

Cytoplasmic division (cytokinesis) in animal cells and many protists is accomplished by a contractile ring of microfilaments composed of the protein F-actin. Contraction involves the interaction of F-actin with another protein, myosin. In plant cells cytokinesis is accomplished by the transport of Golgi vesicles containing precursors of the new cell plate along microtubules of the phragmoplast, a reorganized remnant of the mitotic spindle. *See* CELL DIVISION; CYTOKINE.

Cytochemistry. Cytoplasm contains mostly water, from 80 to 97% in different cells, except for spores and other inactive forms of living material, in which water may be present in lesser amounts. The dry mass of cells consists mainly of macromolecules: proteins, carbohydrates, nucleic acids, and lipids associated with membranes. The small molecules present in cells are mainly metabolites or metabolic intermediates. The principal ions other than the hydrogen and hydroxyl ions of water are the cations of potassium, sodium, magnesium, and calcium, and the anions chloride and bicarbonate. The concentrations of these ions are maintained within narrow physiological limits by ion-pumping mechanisms and the presence of natural chelators or ion-binding substances. Therefore analyses of whole cells for calcium ion concentration often give values two to three orders of magnitude higher than the concentration of free calcium ions in cytoplasm.

Many other elements are present in cytoplasm in smaller amounts. Iron is found in cytochrome pigments in mitochondria; magnesium is present in chlorophyll in chloroplasts; copper, zinc, iodine, bromine, and several other elements are present in trace quantities.

The presence of certain chemical compounds in the cell can be measured quantitatively by spectroscopic techniques applied either to living cells or to fixed cells subjected to chemical reactions in place. Cytochrome pigments, for example, can be demonstrated in living mitochondria by their absorption bands in the blue region of the spectrum. The same holds for different forms of chlorophyll and accessory photosynthetic pigments in plant cells. Nucleic acid absorption can be detected in the ultraviolet region of the spectrum.

By carrying out chemical reactions in fixed cells, it is possible to localize proteins, carbohydrates, and the two basic types of nucleic acids, DNA and ribonucleic acid (RNA). Enzymes can be localized by chemical reactions that demonstrate the product of the reactions they catalyze.

One of the most powerful techniques in cytochemistry is cell fractionation. Cells are first homogenized or otherwise disrupted in a medium that keeps organelles and inclusions intact and, if possible, functional. The resulting suspension of cell organelles and inclusions is centrifuged first at low accelerations to remove nuclei and other heavy particles, and then the homogenate is centrifuged at a higher acceleration in a manner designed to separate particle classes on the basis of density. In this way it is possible to obtain a nuclear fraction, a mitochondrial fraction, and so on. The purity of each fraction is monitored by optical or electron microscopy, and when the desired criterion of purity has been met, fractions are submitted to chemical analysis. In this way it has been possible to construct lists of substances, especially enzymes, known to be present in different organelles or cell compartments.

The many proteins that lack enzymatic function may be regulatory or structural in function. The best way to identify and localize these is by fluorescent-antibody techniques. A given protein is first isolated and purified. Repeated injections of the purified protein (an antigen) into rabbits causes the production of a rabbit antibody to that protein. Antibody isolated from the blood serum of the rabbit and subsequently purified is then used as a probe to look for the antigenic protein by using fluorescent antibodies to rabbit serum proteins produced by some other animal, such as a goat or sheep. This technique is known as the indirect immunofluorescent technique or the "sandwich" technique. *See* CYTOCHEMISTRY; IMMUNOFLUORESCENCE.

Cytoplasmic rheology. In the first two decades of this century a controversy arose as to whether cytoplasm was fluid or gel. Efforts to resolve the controversy by objective physical experiments were not at first successful, because there was insufficient knowledge of the complexity of cytoplasmic organization, and because certain "simplifying" assumptions turned out to be misleading.

Fluids in which the rate of shear (flow) is proportional to the applied shear stress (force) are

called newtonian fluids. Therefore fluids with a nonlinear force-flow relationship are designated nonnewtonian. The slope of the force-flow relationship for a fluid is defined as its fluidity, while the reciprocal of the slope is viscosity, a measure of fluid resistance. Fluids with nonlinear force-flow relations obviously exhibit different slopes and therefore different viscosities at different rates of shear; that is, they exhibit a different fluid resistance at different rates of flow. Thixotropy is a reduction in viscosity at higher flow rates, and indicates a breakage of affinities or bonds contributing to fluid resistance. Dilatancy is the opposite effect: an increase in resistance with increased rate of flow. *See* BIORHEOLOGY.

The first objective methods used in an attempt to gain quantitative information about the consistency of cytoplasm relied on the assumption, which seemed reasonable at the time, that cytoplasm might be a newtonian fluid. Since this assumption disregarded the possibility that cytoplasm might be structured, the fact that the methods used applied forces that destroyed cytoplasmic structure was overlooked.

It is now clear that cytoplasm is highly nonlinear in rheological behavior due to the presence of cytoskeletal proteins in the form of microfilaments, microtubules, intermediate filaments, and various other cytoskeletal proteins in dynamic association with one another. These skeletal elements can confer sufficient structure on cytoplasm in some cells so that a measurable yield force is required to cause any flow at all. The presence of a yield point classifies any fluid as a gel. Some gels can be strong, as in the case of gelatin or agar, while other cytoplasmic gels are weak in comparison. Even weak gel structure is important, because the forces acting upon the cytoplasm are also weak. Cytoplasm in some cells also shows the property of viscoelasticity. Viscoelastic fluids have a "memory" for their original shape, and when deformed tend to recoil toward their original shape.

It is very difficult to characterize the rheological properties of cytoplasm quantitatively, because they depend on the force applied in order to make the measurement. Reliable viscosity values have been found for cytoplasm in the range of 5 centipoise to 10 poise or 0.005 to 1.0 pascal · second (water has a viscosity of about 1 centipoise or 0.001 Pa · s).

In some cells it has been possible to differentiate between the macroviscosity, the viscosity of the cytosol and cytoskeleton, and the microviscosity, the viscosity of the cytosol alone. The former is always an order of magnitude (or more) more viscous than the latter.

Important experiments carried out by Y. Hiramoto showed that the yield point and viscoelastic constants of cytoplasm could be measured by the displacement of an iron sphere inside a cell by an electromagnet. These studies showed the complex nature of the rheological properties of cytoplasm, and the fact that these properties can change with the life cycle of the cell.

Ground cytoplasm: cytosol and cytoskeleton. Sedimentation of cells by centrifugation shows that organelles and inclusions can be separated from the ground cytoplasm, the fluid phase of the cytoplasm in which they are suspended. The ground cytoplasm in turn has been shown to consist of a cytoskeletal network and the cytosol, the fluid in which the cytoskeleton is bathed.

The cytoskeleton consists of several biopolymers of wide distribution in cells. Microtubules have been observed in electron micrographs of a vast number of different cell types. They are about 25 nm in diameter and very long. They consist of the protein tubulin, and are frequently covered by a fuzzy layer of microtubule-associated proteins. Microfilaments of more than one type have been observed. The most common and ubiquitous are those composed of F-actin, one of the most common proteins in muscle of all kinds. Microfilaments are usually 6–7 nm in diameter and several micrometers long. If composed of F-actin, they can be decorated in a characteristic arrowhead pattern by the S-1 subfragment of skeletal muscle myosin to reveal their polarity. Intermediate filaments of several types are known. Most have a diameter of about 10 nm, and are sometimes called 10-nm filaments. At least some of these intermediate filaments are related to the protein keratin, which is present in skin, nails, hair, and horns.

Many proteins have been found that appear to have either a regulatory or cross-linking (gel-forming) function. They bind to one or more of the cytoskeletal proteins and act to control contractility, motility, or cytoplasmic consistency. The patterns of distribution of cytoskeletal proteins have been demonstrated by fluorescent antibody probes which show the presence of a given specific protein by labeling it with a fluorescent marker using the specificity of an antigen-antibody reaction. *See* CYTOSKELETON.

Cytoplasmic motility. In most cells the smaller particles exhibit brownian motion due to thermal agitation. In some cells lacking extensive cytoskeletal structure, particles can be moved freely around the cell by brownian motion. In others they are restricted by the cytoskeletal elements. *See* BROWNIAN MOVEMENT.

Particles of various types may also undergo saltatory motions which carry them farther than brownian motion possibly could. Such excursions are usually either linear or curvilinear, and result from the interaction of a particle with an element of the cytoskeleton such as one or more microtubules or microfilaments. Single particles or trains of particles may exhibit saltation, which is a jumping motion.

When masses of particles move together, they are usually accompanied by a flow of the cytoplasm around them called cytoplasmic streaming. Streaming of cytoplasm is best seen in cells of *Nitessa* or *Chara* (green algae), in amebas and paramecia (protists), in cellular slime molds (fungi), and in many cells (for example, eggs and leukocytes) of higher animals. There is evidence to indicate that cytoplasmic streaming involves the interaction of the proteins F-actin (microfilaments) and myosin. However, the manner of interaction and the cellular regulatory

systems appear to differ widely. The movement of cells of both protists and higher organisms involves contractile, cytoskeletal, and regulatory proteins which exhibit differences from species to species and from cell to cell. *See* CELL MOTILITY.

Perhaps the most intricate and important form of motility is that which occurs in neurons (nerve cells) of all types: axoplasmic transport. Neurons synthesize transmitter substances near their nuclei in the cell body. Vesicles containing these transmitter substances must be transported long distances (up to 6 ft or 2 m in humans) in order to be released at the synapses where one nerve cell excites the next. Anterograde axoplasmic transport brings the vesicles to the nerve endings, whereas retrograde transport carries chemical messages back to the nerve cell body. Failure of this transport mechanism in certain diseases (for example, Alzheimer's disease and senility of old age) is due to the failure of this important process. *See* ALZHEIMER'S DISEASE.

Muscular movement is made possible by virtue of paracrystalline interdigitating arrays of F-actin thin filaments and myosin thick filaments in the myofibrils in the cytoplasm of muscle fibers. Muscle contraction is a highly specialized form of contractility, which is one of the basic properties of cytoplasm. *See* CELL BIOLOGY; MUSCLE; MUSCLE PROTEINS. Robert Day Allen

Bibliography. C. De Duve, *A Guided Tour of the Living Cell*, vols. 1 and 2, 1984; E. D. P. De Robertis and E. M. De Robertis, Jr., *Cell and Molecular Biology*, 8th ed., 1987; H. Herrmann, *Cell Biology*, 1990; S. Seno and Y. Okada, *International Cell Biology 1984*, 1991.

Cytoskeleton

A system of filaments found in the cytoplasm of cells which is responsible for cell shape, cell locomotion and elasticity, interconnection of the major cytoplasmic organelles, cell division, chromosome organization and movement, and the adhesion of a cell to a surface or to other cells.

Since this filamentous system is too small to be resolved by light microscopy, its several subcomponents, structural features, and cytoplasmic associations were not discovered until the electron microscope was developed. With this instrument, which is capable of magnifying a section of a cell many thousands of times, three major classes of filaments could be resolved on the basis of their diameter and cytoplasmic distribution: actin filaments (or microfilaments) each with an average diameter of 6 nanometers, microtubules with an average diameter of 25 nm, and intermediate filaments whose diameter of 10 nm is intermediate to that of the other two classes. In muscle cells, an additional class of thick filaments, known as myosin filaments, is found whose function is to interact with actin filaments and generate the force necessary for muscle contraction. The presence of this system of filaments in all cells, as well as their diversity in structure and cytoplasmic distribution, has been recognized only in the modern period of biology.

A technique that has greatly facilitated the visualization of these filaments as well as the analysis of their chemical composition is immunofluorescence applied to cells grown in tissue culture. In this technique, the protein that is to be localized is injected into an animal to elicit an antibody response. The resulting antibody is highly specific for the protein. Following several purification steps, the antibody is chemically attached to a fluorescent dye. The cells, which are usually grown on a glass cover slip, are fixed with a cross-linking agent such as formaldehyde in order to stabilize their cytoskeletons. They are subsequently rendered permeable to the fluorescent antibody by exposure to a detergent or to a polar solvent such as ethanol or acetone. The cells are then treated with the fluorescent antibody solution for a few minutes; unbound antibodies are washed away, and the cover slip is examined under a regular light microscope equipped with fluorescence optics. The fluorescent antibody binds to the protein it is specific for, and upon excitation of the fluorescent dye, the location of that protein is revealed. Extensions of this technique involve the use of antibodies and electron microscopy. In this technique, known as immunoelectron microscopy, the resolution of subcellular localization is increased appreciably. Furthermore, it can be used to detect the binding site of one protein to another, thereby expanding the detail at which cytoskeletal molecular interactions among each other and in the cytoplasm can be understood. *See* IMMUNOFLUORESCENCE.

Actin filaments. Actin, a protein highly conserved in evolution, is the main structural component of actin filaments in all cell types, both muscle and nonmuscle. As revealed by immunofluorescence and electron microscopy, actin filaments assume a variety of configurations depending on the type of cell and the state it is in. They extend a considerable distance through the cytoplasm in the form of bundles (**Fig. 1***a*). These bundles are also known as stress fibers since they are important in determining the elongated shape of the cell and in enabling the cell to adhere to the substrate and spread out on it. In a single migratory or resting cell, most of the actin filament bundles are located just below the cell membrane and are frequently in close association with it. The sites where the bundles make contact with the cell membrane, known as adhesion plaques, correspond to the sites of cell-substratum contact. It has been estimated from time-lapse motion pictures that most of the actin filament bundles are oriented parallel to the long axis of the cell and shift their orientation as the cell changes direction; the bundles reorient themselves over a period of 15 to 60 min.

Actin filaments can exist in forms other than straight bundles. In rounded cells that do not adhere strongly to the substrate (such as dividing cells and cancer cells), the filaments form an amorphous meshwork that is quite distinct from the highly organized bundles. This filamentous meshwork is usually located in the cell cortex, which is the area just beneath the cell membrane. It is also found in areas of membrane ruffling in motile cells.

Membrane ruffles are sheetlike extensions of the cytoplasm which function as the main locomotory organ of the cells (**Fig. 2**). The membrane ruffle forms transient contacts with the substrate on which the cell rests, stretching the rest of the cell among several temporary adhesions. When the cell is spread out to its fullest extent, it maintains one or two primary ruffles, each of which tends to lead the cell in a particular direction. When the transient contacts of a membrane ruffle are converted to permanent contacts, they become adhesion plaques and function to organize long actin filament bundles. Thus, the two filamentous states, actin filament bundles and actin filament meshworks, are interconvertible polymeric states of the same molecule. They appear to be the extremes of a gradient, and the ratio between them depends on a variety of factors, including the extent of cell motility, the age of the cell, the degree of adhesion to the substrate, and the stage of the cell cycle. Bundles give the cell its tensile strength, adhesive capability, and structural support, while meshworks provide elastic support and force for cell locomotion.

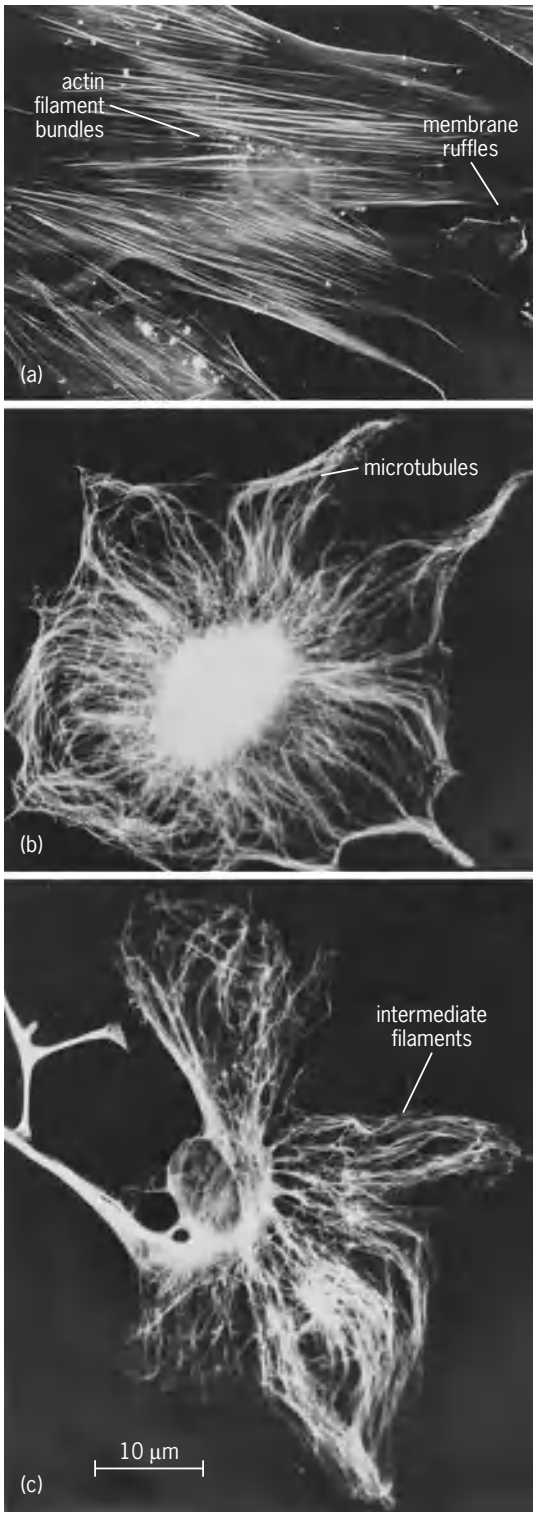

Fig. 1. Immunofluorescence with human cells grown in tissue culture. (*a*) Distribution of actin in filament bundles and membrane ruffles using fluorescent antibodies to muscle actin. (*b*) Distribution of microtubules using fluorescent antibodies to their major subunit tubulin. (*c*) Distribution of intermediate filaments using fluorescent antibodies to their major subunit from fibroblasts.

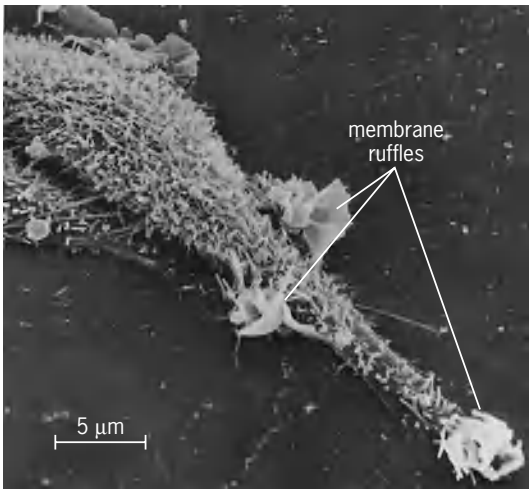

Fig. 2. Scanning electron micrograph of the surface of a human fibroblast grown in tissue culture. The small protrusions on the cell surface are known as microvilli or microspikes and are composed of a core of 20–30 actin filaments.

Actin filaments are involved in two main cellular functions: cell motility, and cell shape and elasticity (Fig. 1*a*). These broad types of function are mediated by specific accessory proteins that interact with actin to change the properties, functional associations, or the polymeric state of actin filaments. Thus, the cell has molecules which regulate the polymerization of actin to single filaments, molecules which regulate the assembly of single actin filaments to filament bundles, molecules which interconvert actin filament bundles and actin filament meshworks, molecules which mediate the interaction of actin filaments with the plasma membrane, and finally molecules, such as myosin, which enable actin filaments to move the cell as a whole or mediate specific motile activities to the cytoplasm, such as cell cleavage. In turn, the spatial and temporal interaction of these proteins with actin is regulated, in part, by the cytoplasmic hydrogen ion concentration (pH) and the calcium ion concentration. In the cytoplasm these regulatory mechanisms coexist in a dynamic equilibrium,

many of them operating simultaneously, to govern cell motility and maintenance of cytoplasmic structure.

The importance of accessory actin-binding proteins in the maintenance of cell shape and elasticity has been demonstrated through the study of certain severe human hemolytic anemias, such as hereditary spherocytosis, pyropoikilolytosis, and elliptocytosis. In these diseases one or more of the actin-binding proteins are defective. Normally, this battery of proteins binds to and crosslinks actin filaments, forming a pliable and deformable network under the plasma membrane of the red blood cell. Mutations in any one of these accessory proteins cause structural alterations in this skeletal network, resulting in the fragmentation of the plasma membrane as the cells deform their shapes during circulation. The consequence is cell lysis and hemolysis. A number of these accessory proteins are now known to be common to red blood cells, lens cells, muscle cells, and nerve cells, among others, in association with actin.

A main function of actin filaments is to interact with myosin filaments to generate the force that allows muscles to contract, and many nonmuscle cells to move. The interaction of these two filament systems is extremely well coordinated by a different set of proteins that interact with either actin filaments or myosin filaments. The interaction of these proteins with their respective filaments is regulated by Ca^{2+} ions, by the H^{+} ion concentration in the cytoplasm, and by direct covalent modifications such as phosphorylation. *See* MUSCLE PROTEINS.

Microtubules. These slender cylindrical structures exhibit a cytoplasmic distribution distinct from actin filaments. Microtubules originate in structures that are closely associated with the outside surface of the nucleus known as centrioles. Fluorescent antibodies specific to the major structural protein of these filaments, known as tubulin, reveal a wavy system of filaments originating in the centriole area and extending to the cell's outer membrane (Fig. 1*b*). Unlike the other two classes of filaments, microtubules are highly unstable structures and appear to be in a constant state of polymerization-depolymerization. During mitosis, they rapidly disaggregate and reorganize around the mitotic spindle, where they function in moving the chromosomes to the two newly forming daughter cells. In the resting cell, microtubules are involved in the determination of a cell's shape and in cytoplasmic transport of vesicles and other small organelles. In conjunction with the centrioles, they also function as the "information processing center" of the cell since they translate external stimuli into decisions about the direction of locomotion of a cell. In more specialized cells, like the sperm cell, they function to propagate the cell through body fluids. *See* CENTRIOLE; MITOSIS.

Intermediate filaments. These filaments function as the true cytoskeleton. Unlike microtubules and actin filaments, intermediate filaments are very stable structures. If a cell is exposed to a mild nonionic detergent, most of the microtubules, actin filaments, and various cytoplasmic organelles are rapidly removed, but intermediate filaments and nuclei are left insoluble. In the intact cell, they anchor the nucleus, positioning it within the cytoplasmic space. During mitosis, they form a filamentous cage around the mitotic spindle in a fixed place during chromosome movement. As seen using immunofluorescence (Fig. 1*c*), intermediate filaments have a cytoplasmic distribution independent of actin filaments and microtubules. However, their distribution in the cytoplasm depends very much on the type of cell and its stage in differentiation. For example, in undiffferentiated muscle cells these filaments are found scattered in the cytoplasm, but in mature, fully differentiated cells they wrap around the individual contractile units, the myofibrils, and function to interconnect them to the plasma membrane. In chicken red blood cells, which contain a nucleus, these filaments function to anchor the centrally located nucleus to the plasma membrane. In mammalian red blood cells, however, which lose their nuclei, these filaments are removed prior to enucleation. In all epithelial cells, and epidermal cells in particular, a special class of intermediate filaments, known as keratin filaments, interconnect the nuclei to special areas of cell-cell contact at the plasma membrane known as desmosomes. In this fashion, an epithelial layer is integreted mechanically both intercellularly and intracellularly to withstand tension and possess elastic properties. Therefore, intermediate filaments are a fundamental component of the morphogenesis of certain differentiating cells. *See* MUSCLE.

Elias Lazarides

Bibliography. A. D. Bershadsky and J. M. Vasiliev, *Cytoskeleton*, 1988; G. G. Borisy et al. (eds.), *Molecular Biology of the Cytoskeleton*, 1984; A. B. Fulton, *The Cytoskeleton: Cellular Architecture and Choreography*, 1985; L. Goldstein and D. M. Prescott (eds.), *Cell Biology: A Comprehensive Treatise*, vol. 4, 1980; D. Marme (ed.), *Calcium and Cell Physiology*, 1985; J. W. Shay (ed.), *Cell and Molecular Biology of the Cytoskeleton*, 1986; P. Traub, *Intermediate Filaments*, 1985.

D

Dacite — Dysprosium

Dacite

An igneous volcanic rock that is the extrusive equivalent of granodiorite, according to the IUGS (International Union of Geological Sciences) classification scheme. The overall color of dacite lava is generally light to medium gray but can appear pinkish or tan when oxidized or weathered. The color index of dacite averages 12–14%, denoting that mafic or dark minerals (hornblende, biotite, hypersthene, augite, titanomagnetite, and ilmenite) make up 12–14% of the rock. As an extrusive igneous rock, dacite is usually porphyritic in texture, meaning it contains 0–50% phenocrysts (visible crystals), in a fine-grained or glassy matrix called the groundmass (**Fig. 1**). The phenocryst minerals typically include some combination of plagioclase, hypersthene, augite, hornblende, biotite, quartz, potassium feldspar, titanomagnetite, and ilmenite, depending on chemical composition and equilibration conditions (pressure, temperature, and oxygen fugacity) of the magma prior to eruption. Dacite may also occur as aphyric (no visible crystals) lava, composed entirely of microcrystalline matrix minerals or glass (vitrophyric). *See* GRANODIORITE; IGNEOUS ROCKS; LAVA; MAGMA; PHENOCRYST.

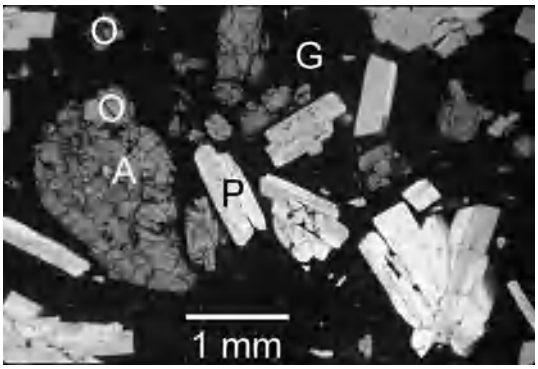

Fig. 1. Photomicrograph of a dacite from the Zitacuaro–Valle de Bravo region of central Mexico. P = plagioclase, O = orthopyroxene (hypersthene), A = augite, G = glassy groundmass.

Fig. 2. Dacite domes from the Zitacuaro volcanic complex of central Mexico. (*Photograph by D. Blatter*)

Dacite is intermediate between andesite and rhyolite in the basalt-andesite-dacite-rhyolite continuum and is classified as having 63–70 wt % silica (SiO_2). Rocks within the dacite SiO_2 range that contain greater than 2.4 wt % potassium oxide (K_2O) are known as latites, and rocks with total alkalis ($Na_2O + K_2O$) greater than ~7 wt % are known as trachytes. Dacite alumina (Al_2O_3) concentrations average about 16 wt %. *See* ANDESITE; LATITE; RHYOLITE; TRACHYTE.

Dacite magma is generated from basaltic or andesitic magmas through magma mixing, assimilation of crustal material, or fractional crystallization. These processes are most common at convergent margins where dacite eruptions have recently occurred, such as at the volcanoes Mount Pinatubo and Mount St. Helens. Because of the high SiO_2 contents and high crystallinities (up to 50% visible crystals) of many dacites, this lava type erupts to form steep-sided domes (**Fig. 2**) and is often associated with explosive eruptions, which form pyroclastic and air-fall deposits. *See* BASALT; PYROCLASTIC ROCKS; VOLCANO; VOLCANOLOGY.

Dawnika Blatter

Bibliography. M. G. Best and E. H. Christiansen, *Igneous Petrology*, Blackwell Science, 2001; D. W. Hyndman, *Petrology of Igneous and Metamorphic*

Rocks, McGraw-Hill, 1985; A. R. Philpotts, *Principles of Igneous and Metamorphic Petrology*, Prentice Hall, 1990.

Dairy cattle production

A section of animal agriculture concerned with the production of milk and milk products. The mammary gland in mammals was developed for the nourishment of their young. Over the centuries humans have selected some mammals, especially the cow, goat, and water buffalo, for their ability to produce quantities of milk in excess of the requirements of their young. At the same time humans have improved the management and nutrition of the animals so that they can produce milk more efficiently. The production of milk includes selection of animals, breeding, raising young stock, feed production, nutrition, housing, milking, milk handling, sanitation, disease control, and disposal of milk, surplus animals, and manure.

The cattle industry dates back many centuries. In Libya, Africa, friezes dating back to 9000 B.C. show domesticated cows. Written records as far back as 6000 B.C. of Old Mesopotamia (now Iraq) indicate that dairying was very highly developed at that time. These people were the first to make butter. Egyptian records show that milk, butter, and cheese were used as early as 3000 B.C.

Dairy development in America. Dairying in the United States had its principal origin from northern Europe. The first cattle were brought to the West Indies by Columbus in 1495. The first importations to the United States were made to Jamestown, Virginia, in 1611, and the Plymouth Colony in 1624.

In the United States up to 1850 most families kept a cow and the milk and its products were consumed by the family. Excess milk was used for cheese and butter, but little was sold since transportation was by horse and wagon and no refrigeration was available.

After 1850 dairying became an industry as a result of several major factors. In 1841 the first regular shipment of milk by rail was made from Orange County, New York, to New York City. The first cheese factory was built in 1851 in Oneida, New York. In 1856 Gail Borden invented the condensed milk process. The first butter factory was started in New York State in 1856, and the first creamery was established in Iowa in 1871. The separator was invented by Gustav DeLaval in 1878, and in 1890 the Babcock test (to determine fat content) was introduced, which allowed for the impartial buying of milk. The first pasteurizing machine was introduced in 1895, and the first compulsory pasteurization law was passed in Chicago in 1908. Refrigeration came into use in 1880. Contributing to the development of the dairy industry was the establishment of experiment stations, which conducted research on feeding, breeding, and management of cattle. This was accompanied by the organization of the agricultural extension service and other educational facilities to carry information to the dairy farmers. Artificial insemination was introduced in 1938 and is now an important segment of the dairy cattle industry.

World dairy production. The leading milk-producing countries in the world, in decreasing order of annual milk production, are Russia, United States, France, Germany, United Kingdom, Netherlands, Italy, and Canada. Figures are not available for most of the Asian countries, which include the mainland of China and India. Most of the leading milk-producing countries are located in the temperate regions. Dairy cattle numbers and milk production are somewhat more limited in the tropical and semitropical regions. Some countries, namely, New Zealand and Denmark, are not listed among the 10 top countries for milk production but are major export countries because of their large per capita production in comparison to consumption.

Dairy production in United States. The five leading dairy states are Wisconsin, California, New York, Minnesota, and Pennsylvania. The major milk-producing areas are the Midwest and Northeast.

Even though there has been a marked decrease in the number of milk cows since 1940, the annual milk production has increased because of the marked increase in production per cow. A decrease in the number of dairy farms producing milk has been accompanied by an increase in the number of cows per farm and in the number of cows handled by each dairy farmer. Smaller inefficient producers have generally been forced out of the dairy industry, and only the larger, more specialized farms remain.

Milk is used for many products. Approximately 38% goes into fluid milk and cream. Of the remaining milk, about 19% is used for butter, 28% for cheese, and 15% for other products, with frozen dairy products being the most important. *See* BUTTER; CHEESE; ICE CREAM; MILK.

Dairy breeds. Most of the milk produced in the United States is by five major dairy breeds: Ayrshire, Brown Swiss, Guernsey, Holstein-Friesian, and Jersey. Approximately 12% of the dairy cattle are registered purebreds. The first breed registry association was organized by the Jersey dairy producers in 1868, and all breeds had registry associations by 1880. The purposes of these organizations are to preserve the purity of the breed through registration; to keep official records of the breed; to improve the breed through production testing programs, type classification, and dairy cattle shows; and to promote the breed.

The introduction of the dairy herd improvement associations (DHIA) in 1905 to measure the production of cows greatly aided the improvement of breeds, especially nonpurebred animals. These testing programs produced the needed data for the selection of replacement animals and for culling low-producing animals. The principal features of each of the five dairy breeds are given below.

Holstein-Friesian. This is the largest breed in number and in body size. A mature dairy cow weighs about 1500 lb (675 kg) and the bull 2200 (990 kg). The breed was developed in the North Holland and West Friesland provinces of the Netherlands. Holsteins

A champion milk producer, Harborcrest Rose Milly, the all-time, all-American Holstein-Friesian aged cow. (*Owned by Paclamar Farms, Louisville, Colorado*)

were first imported to the United States in 1621 and 1625 by the Dutch settlers; however, there are no descendants of these animals in the country. In the mid-1850s cattle were imported into Massachusetts from which descendants are now present. Importations in large numbers were made up to 1905.

The Holstein is a distinct black and white, is noted for its rugged conformation, and has a decided advantage in the production of beef and veal, as well as milk. The breed produces the largest volume of milk with the lowest milk fat percentage, averaging about 3.7%. Since the late 1940s the percentage of Holstein cattle in the United States has increased, because the milk-pricing systems have favored breeds that produce large volumes of milk with relatively low butterfat. One of the outstanding cows in the breed (see **illus**.) was classified Excellent with 97 points, had four records of over 23,000 lb (10,350 kg) of milk on twice-a-day milking, and had a lifetime production of over 138,000 lb (62,100 kg) of milk.

Jersey. The Jersey is the smallest of the dairy breeds. The mature cow weighs about 1000 lb (450 kg) and mature bull 1500 lb (675 kg). The Jerseys were developed on the Island of Jersey, one of the Channel Islands between France and England. The breed was first imported into the United States in 1850, with large importations occurring between 1870 and 1890. The percentage of Jerseys has been decreasing primarily due to the milk-pricing systems and a decreased consumption of butter. The Jersey ranges from a light fawn to black; some animals have white markings. The breed has the highest milk-fat content, averaging 4.9%, and a high solids-not-fat content. The milk is notably yellow, which is due to a high carotene content.

Guernsey. The Guernsey breed is slightly larger than the Jersey and has many of the latter's common characteristics. The mature female weighs 1100 lb (495 kg) and mature bull 1700 lb (765 kg). The Guernsey varies from an almost red to a very light fawn, with various sizes of light markings. The breed was developed on the Island of Guernsey, another Channel Island. They were first imported into the United States in 1830; while some importations still occur, they have decreased in numbers. The Guernseys average 4.7% milk fat and have a high solids-not-fat content. The milk also is a deep yellow due to the high carotene content.

Ayrshire. The Ayrshire breed was developed in the County of Ayr in Scotland around 1750. The mature cow weighs approximately 1200 lb (540 kg) and the bull 1850 lb (830 kg). The Ayrshire is red and white and noted for its general beauty, straightness of top line, symmetry of udder, ruggedness, and excellent grazing ability. The Ayrshires were first imported from Scotland to Massachusetts in 1837, with frequent importations made until 1860.

The average milk-fat content of the Ayrshire is 3.9%. The breed is known for its longevity and lifetime production rather than for its high lactation records.

Brown Swiss. The Brown Swiss is one of the larger breeds with the mature cow weighing about 1400 lb (630 kg) and the mature bull 2000 lb (900 kg). The Brown Swiss varies from a light fawn to almost black, with a mouse color predominating. Brown Swiss have a quiet temperament but a tendency to be stubborn. It is the slowest-maturing breed and is noted for its longevity. Because of size and ruggedness, they are well adapted for beef and veal purposes.

Brown Swiss is one of the oldest breeds. It was developed in the mountain territory of northeastern Switzerland, where in the winter the cows are housed in valleys and barns and in the spring are pastured on mountain slopes. This developed the ruggedness and excellent grazing ability. The cattle (less than 200) were first imported into the United States in 1869. The milk averages 4.1% milk fat and is well adapted to market milk, condensed milk, and cheese.

Dairy cattle feeding. Research on dairy cattle feeding has provided information on feeding animals according to their requirements to the dairy industry, interested in meeting nutritional needs at a minimum cost. The nutrient requirements of the dairy cow depend upon a daily maintenance need, milk production and milk-fat test, pregnancy, and growth. The two most important requirements to be met are energy and protein. Various energy standards are used, but the commonest ones in the United States are the total digestible nutrient (TDN) and net energy (NE) standards. The mineral requirements of the cow, particularly calcium, phosphorus, sodium, chloride, and iodine, are another concern of the dairy farmer. The other minerals and vitamins are found in sufficient quantities in the feeds.

The main components of the dairy ration are hay, silage, pasture, and grain. Grain is the most expensive portion of the ration and is fed according to the milk production of the cow. Hay and silage are usually fed to the cows free-choice, and a cow consumes daily about 2 lb of hay or 6 lb of silage per 100 lb of body weight. This provides enough energy and protein to meet the maintenance requirements of the cow plus 10–20 lb (5–10 kg) of milk each day. The rest of the energy and protein must come from grain. Grass and corn are used as silage. *See* ANIMAL FEEDS.

Dairy cattle breeding. The breeding program involves two major aspects. One is the selection and breeding of cows to be used in the dairy herd. The second is reproduction and maintenance of a regular calving interval for maximum milk production. Many commercial dairy producers place most emphasis on production in the breeding program and select for type only if the conformation of the animal presents management problems, such as a poorly attached udder. Some purebred dairy cattle breeders who sell replacement stock emphasize the overall type of animal.

To make progress in a breeding program, the dairy producer must use records. DHIA records are most commonly used for the selection of milk production. Selection for type is based on official classification scores conducted by the breed associations or the dairy farmer's own appraisal. Most selection in the breeding program is done at two times. One occurs in the selection of the bull or bulls to be used to breed cows for development of future herd replacements; the second is selecting animals to be culled and replaced by new animals. Approximately 32% of the cows leave the herd each year, because of disease, reproductive difficulties, and low production. With the advent of artificial insemination (AI) and the use of sire proving programs by AI associations, desirably proven bulls are available almost everywhere. Approximately 55% of the dairy cows in the United States are bred by artificial insemination. *See* BREEDING (ANIMAL).

Maintaining a good breeding program requires that cows calve at regular intervals. Most dairy workers try to have a 13–14-month calving interval. The ideal is 1 year; however, this is usually not possible. Cows calving at 13–14-month intervals produce more milk during their lifetimes than cows calving at wider intervals.

Dairy cattle management. Almost all cows in the United States are milked by machine. The recommended milking procedure involves a series of steps based upon the physiology of the cow, as well as the production of clean milk and disease control. The usual recommended steps in a milking operation are: Wash the cow's udder with a disinfectant solution and a paper towel, use a strip cup to detect abnormal milk, apply the milking machine 1–2 min after the cow's udder has been washed, remove the machine as soon as milk flow stops, and dip the teat ends in a disinfectant solution. Good milking procedures are required to maintain a low incidence of mastitis, which is the most costly dairy cattle disease. *See* AGRICULTURAL SCIENCE (ANIMAL); DAIRY MACHINERY; PASTEURIZATION.

G. H. Schmidt

Bibliography. D. L. Bath et al., *Dairy Cattle*: *Principles*, *Practices*, *Problems*, *Profits*, 3d ed., 1985; J. R. Campbell and R. T. Marshall, *The Science of Providing Milk for Man*, 1975; W. M. Etgen and P. M. Reaves, *Dairy Cattle Feeding and Management*, 7th ed., 1987; H. O. Gravert (ed.), *Dairy-Cattle Production*, 1988; G. H. Schmidt and L. D. Van Vleck, *Principles of Dairy Science*, 1974.

Dairy machinery

Equipment used in the production and processing of milk and milk products, including milking machines, cream separators, coolers, pasteurizers, homogenizers, butter-making equipment, evaporators and dryers, and related items of equipment. The equipment must be easy to clean and designed to prevent contamination of the milk or milk products from dirt, oil, soluble metals, insects, and other foreign materials.

Construction and design. Stainless steel, an alloy of chromium and steel, is widely used and is highly satisfactory for direct contact with milk and other food products. If properly used, stainless steel does not affect the flavor and is corrosion-resistant to food products. However, corrosion of stainless steel may be caused by prolonged contact with food or by removal of the protective oxide layer, which must be maintained to provide corrosion resistance. The layer is removed by prolonged contact with chlorine. Stainless steel surfaces must be cleaned regularly after use with detergent solutions, and should be sanitized before, rather than after, use so that excessive chlorine contact will be avoided. Surfaces are sanitized by using a hypochlorite solution of approximately 200 ppm of available chlorine. *See* STAINLESS STEEL.

Most modern dairy equipment is designed to be cleaned in place (CIP), without disassembly, by pumping detergents and cleaning solutions through the entire system. Equipment is specially designed to avoid pockets, to provide smooth surfaces, and to avoid burning-on or buildup of the product on food-contact surfaces.

Certain standards of design, construction, operation, and sanitation of milk and food equipment, known as the 3-A standards, have been established; they were developed by three associations and are widely used throughout the industry. The 3-A standards specify such items as composition of the material, thickness of material (gage), surface finish, design of corners and joints, size, power requirements, and method of utilization.

The science of hydraulics is basic to the fluid-flow processes of pumping, piping, agitating, mixing, centrifuging, and homogenizing. *See* HYDRAULICS.

Electric motors are used to supply power for these processes. Electricity is used for heating for small equipment and for instrumentation and control of processing equipment. Boilers furnish hot water and steam for cleaning (equipment and plant) and pasteurizing operations. Steam is the principal utility used for heating. Heating is a major utility necessary for pasteurization, sterilization, sanitation, and production of evaporated and dried milk. Refrigeration is provided through direct-expansion systems, by sweet water (cooled tap water) systems, or with brine systems for cooling of incoming product, storage, ice cream production, and for control of plant environment.

Refrigeration and heating require the major portion of the energy in dairy and food plants. In a medium-size plant the following basic utilities are

required per 1000 gal (3.78 kiloliters) of processed product: 130 kWh of electricity, 72 ton-h of refrigeration, 350 lb (158 kg) of steam, and 8000 lb (3630 kg) of water. *See* REFRIGERATION; THERMODYNAMIC PRINCIPLES.

Milking machines. These units extract milk from the cow's udder and deliver the milk to an adjacent container or, for a pipeline milking system, directly to a central cooling tank. Teat cups extract milk by intermittent vacuum action on a flexible inner wall called an inflation. A strip cup should be used before milking to check the condition of the milk. Machine milking requires 3.5–5 min, which is less time than for hand milking. A milking machine consists basically of a unit for providing a vacuum with a power unit, pipes or hoses to the animal, a device for providing pulsations, and a container for collecting the product.

A piston or vane-type pump, usually driven by an electric motor, is used for developing a vacuum of 10–15 in. Hg (34–51 kilopascals); a vacuum storage tank next to the outlet of the vacuum pump reduces the fluctuations of the piston. Intermittent action of 45–55 cycles/min is secured by the pulsator unit. Milk may be pumped or transferred to a central point by pipeline using a vacuum. *See* VACUUM PUMP.

Types. Milking machines can be classified as portable and pipeline units. The motor and vacuum pump of the portable unit can be either a part of the milker head or on a cart which is moved along behind the cows. The pipeline type can include either a pipeline for the air supply from the vacuum pump to the milker head or a pipeline from the cow to the milk house for transport of milk. With the latter type, there is usually a separate air line. About 1.5 kWh per cow per month is required for electricity for milking using a portable unit.

Operation. Two hoses go to the teat-cup units. One hose, called the air hose, extends from the pulsator, which alternately provides air at atmospheric pressure and vacuum (low pressure) to the teat cups. The other hose is the milk hose, which is under vacuum and moves the milk, caused to flow from the udder at atmospheric pressure, from the teat cups to the milker pail, which is under vacuum, or to the pipeline for a pipeline milking machine (**Fig. 1**). The rubber inflation surrounds the teat and provides the action for milking by the air-vacuum relationship set up by the pulsator. The rubber inflation is set in the metal cup, and the space between the inflation and the metal cup is alternately connected to air and vacuum, thus resulting in a massaging action. The space inside the inflation is always under vacuum when the teat cups are attached to the cow. The uniform vacuum on the inside of the inflation holds the teat cups in place.

Sanitization. Thorough cleansing of all equipment parts in contact with milk is essential to the production of milk of acceptably low bacteria count. Washing should be done immediately after using by first rinsing with cold water, then hot water using a wetting agent, rinsing with boiling water, and dismantling of equipment. A sanitizing solution may be run through the reassembled equipment just before milking. Local and federal public health ordinance requirements must be met for equipment and for the milk house.

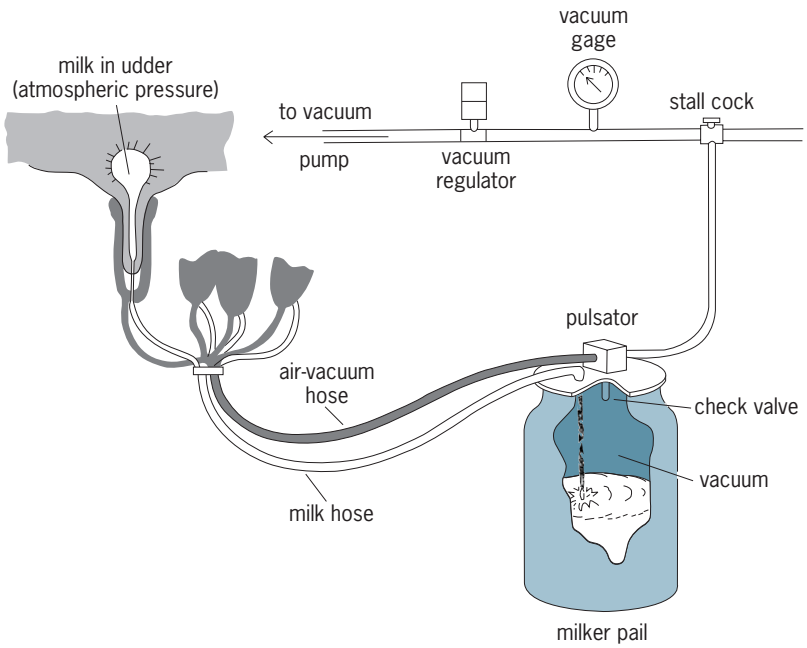

Fig. 1. Arrangement of the component parts of a milking machine.

Farm milk coolers. Milk must be rapidly cooled after production to avoid rapid growth of microorganisms in the product. Milk is normally about 93–95°F (34–35°C) as it leaves the cow. Most ordinances require that milk be cooled to 45–55°F (7–13°C) within 2 or 3 h after milking.

Surface coolers may be used for cooling the milk coming from the cow before it is put in the tanker. The surface cooler may be supplied with cool well water with the final cooling done by mechanical refrigeration. For milk which has been cooled over a surface cooler, the cans may be placed in refrigerated storage to maintain the temperature. Can coolers may be used for cooling the milk coming from the cow to below 50°F (10°C). Water is sprayed over the surface of the can or the can is immersed in cold water to provide the cooling. Bulk milk coolers are now widely used for refrigerating and holding milk on farms. Tanks meeting 3-A standards cool milk to 50°F (10°C) within 1 h and to 40°F (4.5°C) within 2 h after milking. The temperature of the product should not rise above 50°F when warm milk is added. The milk is normally held at 40°F or below. Milk is normally held 1 or 2 days in the farm bulk tanks (**Fig. 2**).

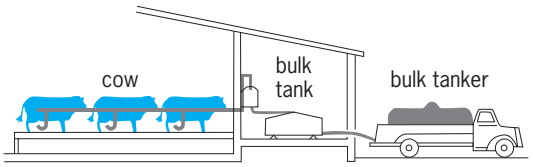

Fig. 2. Handling of milk from cow to tanker.

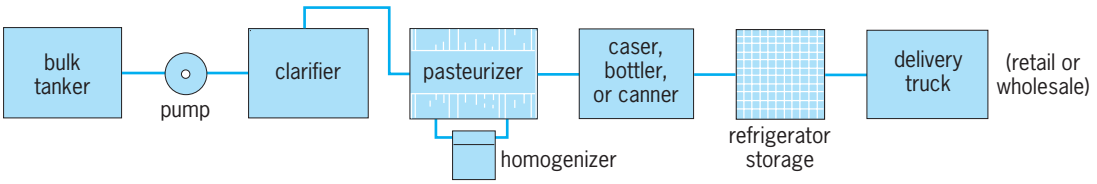

Fig. 3. Processing plant operations.

Bulk milk tankers. Tankers for bulk milk have insulated cylindrical tanks mounted in a horizontal position, as a trailer or on a truck, and are used for pickup of bulk milk. The tanker driver is responsible for sampling, weighing, checking quality and temperature, and handling the milk. Tolerances for measuring the quantity of product have been established by the National Bureau of Standards. The milk is pumped from the farm bulk tank to the tanker to be delivered to the processing plant. At the plant, a pump is usually used for emptying the tanker so that the milk can be transferred from the bulk tanker to a storage tank in the plant. The principal operations in the processing plant are shown in **Fig. 3**.

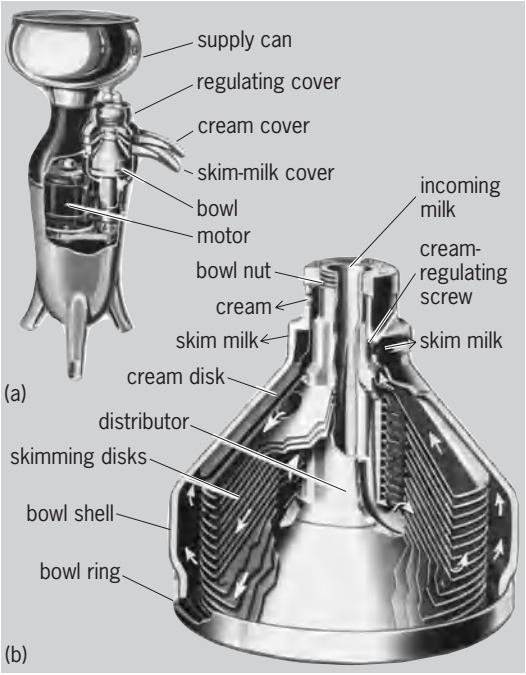

Fig. 4. Cream separator driven by electric motor. (*a*) Entire unit. (*b*) Cross-sectional view of the separating bowl. (*DeLaval Separator Co.*)

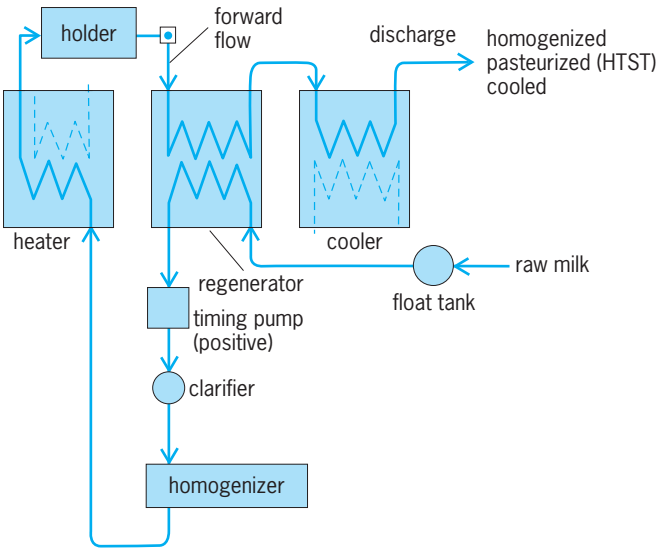

Fig. 5. Schematic representation of pasteurization processes. Details of bypass, relief lines, equalizer, and check valves are not included.

Clarifier. A clarifier removes foreign particles from milk by centrifugal force. A filter is sometimes used for this operation but is less efficient than a clarifier. In appearance and operation the clarifier is similar to a separator, but instead of a cream spout, which is on a separator for removal of a portion of the product, the clarifier accumulates the sediment in the bowl. The sediment must be removed periodically in order to maintain product flow.

The milk is normally clarified before homogenization in the bottled product so that the pasteurized product will not become contaminated. Clarification may be at temperatures of 50–95°F (10–35°C). Generally, a lower temperature of clarification is desired to avoid possible foaming from incorporation of air, to eliminate extraneous material before it is dissolved in the milk, thereby preventing off-flavors, and to reduce removal of components of the product and buildup of sediment in the centrifugal bowl.

Separator. Mechanical centrifugal devices used to extract fat from milk are called cream separators. Milk normally contains 3–6% fat in the form of globules which are lighter in density than the fluid skim milk. These particles rise to the surface when fresh milk is allowed to stand quiescent for several hours and form a layer of cream which may be skimmed from the surface.

In the centrifugal cream separator (**Fig. 4**), milk is moved into the bottom of an airtight bowl turning at 6000–10,000 rpm, which subjects the milk to radial centrifugal forces up to 500 *g*. The bowl contains stacked disks in the form of inverted cones which divide the milk into thin layers and provide the fluid friction necessary to bring the milk to the needed rotational speed. The angle of the inverted cone is such that the lighter cream particles tend to be forced up the center, where they are removed, and the heavier skim milk flows down the outside of the bowl and into the top, where it is collected. An efficient separator will leave no more than 0.1% fat in the milk. *See* CENTRIFUGATION.

The usual method of centrifugally separating the cream from the milk is to heat the milk to at least 90°F (32°C) and as high as 160°F (71°C). As the temperature of the milk increases, the viscosity decreases and

more efficient separation is obtained. Separators specially designed for efficient separation of cold milk have been developed.

Homogenizer. A homogenizer is a high-pressure pump, which is adjustable in pressure, with three or five pistons, and which forces the milk through small openings to break up and reduce the size of the fat globules. The size of the particles is regulated in general by adjusting the pressure of the valve. One- or two-stage units can be used, with the product passing through one or two homogenizing valves. Pressures from 2000 to 2500 lb/in.^{2} (14 to 17 megapascals) are normally used for single-valve milk homogenizers.

Milk must be heated before homogenization to inactivate lipase activity and to provide efficient breakup of the fat particles. The homogenizer may also be used as a pump.

Pasteurizer. This type of equipment is used to heat milk to a predetermined temperature and hold it at that temperature long enough to kill the pathogenic organisms that may be present, and, equally important, to inactivate enzymes. Satisfactory pasteurization can be obtained at various time-temperature combinations, ranging from at least 30 min at no less than 145°F (63°C), for batch pasteurization, to at least 15 s at no less than 161°F (72°C), for high-temperature short-time (HTST) systems. Flash pasteurization at 230°F (110°C), followed by discharge into a vacuum chamber for rapid cooling, is also used. This proce-dure is also termed the vacuum process. Batch-type pasteurizers are heated externally with hot water and require thorough agitation of the milk to heat the milk uniformly. Continuous-flow pasteurizers of the HTST type use regeneration as a means of increasing thermal efficiency by transferring heat from the hot outgoing milk to the cool incoming milk (**Fig. 5**). Cream and milk products to which sugars or other materials have been added must be given a higher temperature or longer time of pasteurization.

The trend is to higher temperatures and shorter times. Sterilized milk not needing refrigeration is now being produced. Sterilizing milk followed by aseptic packaging provides a means of minimum heat treatment and maximum product quality as contrasted with sterilization in the can.

Pasteurized products are placed in glass, metal, paper, or plastic cartons. Some products are placed in bulk containers which are usually placed in cases that are stacked automatically and moved by conveyor into storage. The storage is maintained at 40°F (4.5°C) or less. Milk products are often stored in the cooler, in truck lots, so that the retail or wholesale delivery trucks can rapidly make up their loads from the cooler storage. Automatic equipment for stacking and unstacking of cases, for emptying and filling cases with bottles or cartons or both, and for inspecting the cleanliness of containers is used.

Milk is submitted to a vacuum treatment to maintain a uniform flavor and odor by removal of low-boiling-point volatile components. Feed flavors and variation of product throughout the year, often caused by volatiles, are removed by submitting the product to a vacuum of 10–20 in. Hg (34–68 kPa) in one or two units. If used, this equipment is normally a part of the overall system of milk pasteurization.

Evaporation and drying. Water is removed from milk to reduce the bulk and decrease the cost of transportation and storage.

Evaporator. Evaporated milk is produced by removing water to condense the product to 3:1 or 2:1 volume ratios. Condensed milk is often treated by adding sugar. A vacuum pan is operated at a temperature of 130–140°F (55–60°C) or at 25 in. Hg (85 kPa) vacuum. For large operations, multiple-effect evaporators with one, two, or three stages of moisture removal are used. Decreased steam cost can be obtained with multiple-effect evaporators, a reduction that must be balanced against increased equipment cost. Evaporation usually precedes drying operations. The cost of removing water from the product is less in the evaporator than in the dryer.

Dryer. Dryers are generally of two types: the drum, or roller, dryer and the spray dryer (**Fig. 6**). The drum

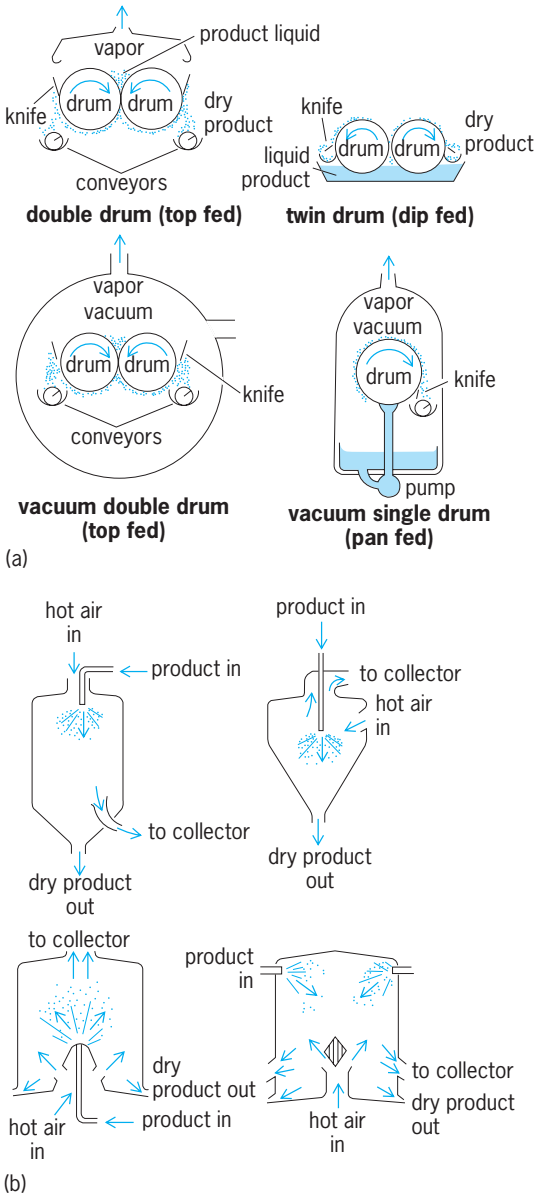

Fig. 6. Types of milk dryers and method of operation. (*a*) Drum dryers. (*b*) Vertical chamber spray dryers.

dryer is less expensive for moisture removal but has a greater heat effect on the product. With the drum dryer a thin film of product is moved over the steam-heated drum. As the drum turns, moisture is removed and a knife scrapes the product from the drum.

The spray dryer has an atomizing nozzle or a centrifugal device that breaks the product into small droplets. Heated air is forced around the atomizing head and evaporates moisture from the droplets. The dry product is then removed from the dryer. The spray-dried product has better solubility and provides, in general, a more satisfactory product for food. Instantizing equipment provides a procedure whereby dry particles are wetted slightly and agglomerates of particles formed. These agglomerates are then dried. The agglomerates have much more rapid solubility than the agglomerated products and provide many products that are known in the trade as instant products or instant powders. Improvements in methods of processing and drying milk have led to rapid increase in utilization of spray-dried and instantized products for the beverage milk industry. This practice will lead to wide distribution of milk products throughout the United States and the world. *See* BUTTER; CHEESE; DAIRY CATTLE PRODUCTION; ICE CREAM; MILK.

Carl W. Hall

Bibliography. A. W. Farrall, *Engineering for Dairy Food Products*, 2d ed., 1980; C. W. Hall and D. C. Davis, *Processing Equipment for Agricultural Products*, 2d ed., 1979; W. J. Harper and C. W. Hall, *Dairy Technology and Engineering*, 1976; R. K. Robinson (ed.), *Modern Dairy Technology*, 2d ed., 1994.

D'Alembert's paradox

A theorem in fluid mechanics which states that no forces act on a body moving at constant velocity in a straight line through a large mass of incompressible, inviscid fluid which was initially at rest (or in uniform motion). This seemingly paradoxical theorem can be understood by first realizing that inviscid fluids do not exist. If such fluids did exist, there would be no internal physical mechanism for dissipating energy into heat; hence there could be no force acting on the body, because work would then be done on the fluid with no net increase of energy in the fluid.

The viscosity of many fluids is very small, but it is essential in explaining the forces that act on bodies moving in them. The action of viscosity creates a rotation of the fluid particles that come near the surface of a moving body. This vorticity, as it is called, is convected downstream from the body so that the assumption or irrotationality of the fluid motion, made in the proof of D'Alembert's theorem, does not correspond to reality for any known fluid. For a winglike body this viscous action sets up a circulation around the body which creates a lifting force. Viscosity also gives rise to tangential stresses at the body surface, called skin friction, which result in a drag force on the body. The work done on the fluid by moving a body through it shows up first as kinetic energy in the wake behind the body, which is gradually dissipated into heat by further action of viscosity.

D'Alembert's theorem does not preclude the possibility of a couple acting on the body, and in fact the irrotational, inviscid fluid theory does predict such a couple. This couple is almost always such as to cause the body to present its greatest projected area in the direction of motion. *See* FLUID-FLOW PRINCIPLES.

Arthur E. Bryson, Jr.

D'Alembert's principle

The principle that the resultant of the external forces $\mathbf{F}$ and the kinetic reaction acting on a body equals zero. The kinetic reaction is defined as the negative of the product of the mass m and the acceleration $\mathbf{a}$. The principle is therefore stated as $\mathbf{F} - m\mathbf{a} = 0$. While D'Alembert's principle is merely another way of writing Newton's second law, it has the advantage of changing a problem in kinetics into a problem in statics. The techniques used in solving statics problems may then provide relatively simple solutions to some problems in dynamics; D'Alembert's principle is especially useful in problems involving constraints. *See* CONSTRAINT.

If D'Alembert's principle is applied to the plane motion of a rigid body, the techniques of plane statics can be used. The principal advantage is that in a dynamics problem the torques must be calculated about a fixed point or about the center of mass, while in statics the torques can be calculated about any point.

Paul W. Schmidt

Bibliography. E. A. Desloge, *Classical Mechanics,* 1982, reprint 1989; H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; R. Matzner and L. Shepley, *Classical Mechanics,* 1991.

Dalitz plot

Pictorial representation in high-energy nuclear physics for data on the distribution of certain three-particle configurations. Many elementary-particle decay processes and high-energy nuclear reactions lead to final states consisting of three particles (which may be denoted by a, b, c, with mass values m_a, m_b, m_c). Well-known examples are provided by the K-meson decay processes, Eqs. (1) and (2), and by the K- and $\overline{K}$-meson reactions with hydrogen, given in Eqs. (3) and (4). For definite total energy

$$K^+ \rightarrow \pi^+ + \pi^+ + \pi^- \tag{1}$$

$$K^+ \rightarrow \pi^0 + \mu^+ + \nu \tag{2}$$

$$K^+ + p \rightarrow K^0 + \pi^+ + p \tag{3}$$

$$K^- + p \rightarrow \Lambda + \pi^+ + \pi^- \tag{4}$$

E (measured in the barycentric frame), these final states have a continuous distribution of configurations, each specified by the way this energy E is shared among the three particles. (The barycentric frame is the reference frame in which the observer

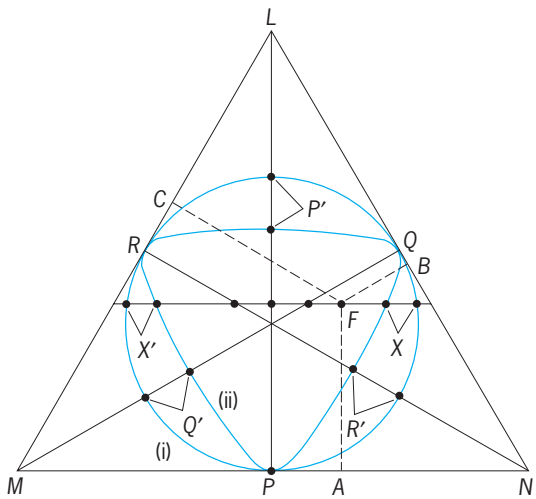

Fig. 1. A three-particle system (*abc*) in its barycentric frame is specified by a point *F* so that perpendiculars *FA*, *FB*, *and FC* to the sides of an equilateral triangle *LMN* (of height *Q*) are equal in magnitude to the kinetic energies T_a, T_b, T_c, where *Q* denotes their sum. See Eq. (5). Curve (i) encloses all points *F* which correspond to physically allowed configurations, for equal masses and nonrelativistic kinematics; curve (ii) corresponds to curve (i) when relativistic kinematics are used, appropriate to the decay process ω(785 MeV) $\rightarrow \pi^+\pi^-\pi^0$.

finds zero for the vector sum of the momenta of all the particles of the system considered.) *See* ELEMENTARY PARTICLE.

Equal mass representation. If the three particles have kinetic energies T_a, T_b, and T_c (in the barycentric frame), Eq. (5) is obtained. As shown in **Fig. 1**,

$$T_a + T_b + T_c = E - m_a c^2 - m_b c^2 - m_c c^2 = Q \quad (5)$$

this energy sharing may be represented uniquely by a point F within an equilateral triangle LMN of side $2Q/\sqrt{3}$, such that the perpendiculars FA, FB, and FC to its sides are equal in magnitude to the kinetic energies T_a, T_b, and T_c. This exploits the property of the equilateral triangle that $(FA + FB + FC)$ has the same value (Q, the height of the triangle) for all points F within it. The most important property of this representation is that the area occupied within this triangle by any set of configurations is directly proportional to their volume in phase space. In other words, a plot of empirical data on this diagram gives at once a picture of the dependence of the square of the matrix element for this process on the a, b, c energies.

Not all points F within the triangle LMN correspond to configurations realizable physically, since the a, b, c energies must be consistent with zero total momentum for the three-particle system. With nonrelativistic kinematics (that is, $T_a = p_a^2/2m_a$, etc.) and with equal masses m for a, b, c, the only allowed configurations are those corresponding to points F lying within the the circle inscribed within the triangle, shown as (i) in Fig. 1. With unequal masses the allowed domain becomes an inscribed ellipse, touching the side MN such that MP:PN equals m_b:m_c (and cyclical for NL and LM). More generally, with relativistic kinematics ($T_a = \sqrt{(m_a^2C^4 + p_a^2c^2} - m_ac^2$, etc.), the limiting boundary is distorted from a simple ellipse or circle. This is illustrated in Fig. 1 by the boundary curve (ii), drawn for the $\omega \rightarrow 3\pi$ decay process, where the final masses are equal. This curve was also calculated by E. Fabri for Eq. (1), and this plot is sometimes referred to as the Dalitz-Fabri plot. In the high-energy limit $E \rightarrow \infty$, where the final particle masses may be neglected, the boundary curve approaches a triangle inscribed in LMN.

The following points (and the regions near them) are of particular interest:

1. All points on the boundary curve. These correspond to collinear configurations, where a, b, c have parallel momenta.
2. The three points of contact with the triangle LMN. For example, point P corresponds to the situation where particle c is at rest (and therefore carries zero orbital angular momentum).
3. The three points which are each farthest from the corresponding side of the triangle LMN. For example, point P' on Fig. 1 corresponds to the situation where b and c have the same velocity (hence zero relative momentum, and zero orbital angular momentum in the bc rest frame).

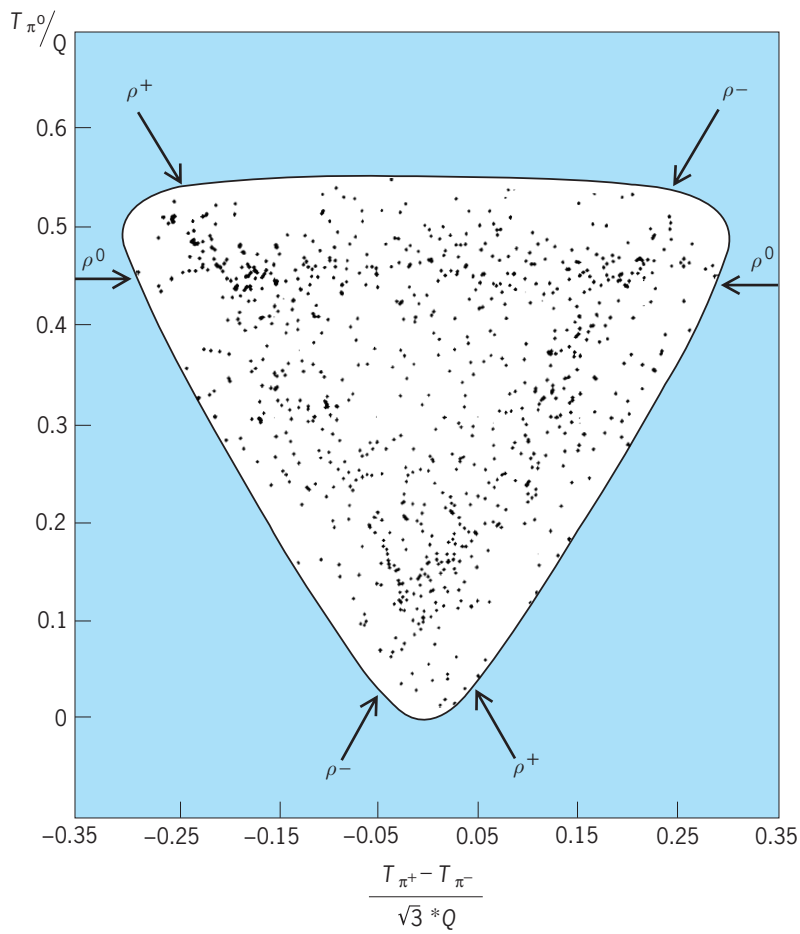

Fig. 2. Dalitz plot for 823 examples of the antiproton annihilation process, $p + p \rightarrow \pi^+ + \pi^+ + \pi^0$. Arrows show positions expected for ρ(765)-meson resonance bands, which appear clearly for the ($\pi^+\pi^-$), ($\pi^-\pi^0$), and ($\pi^0\pi^+$) systems. Distribution is symmetrical between the six sectors obtained by drawing the three axes of symmetry. The authors interpret this distribution as being due to (1) the reaction $\bar{p}p \rightarrow \rho\pi$ occurring in the $I = 0$ $S_1{}^3$ initial state, (2) the reaction $pp \rightarrow 3\pi$ (*s*-wave pions) occurring in the $I = 1$ $S_0{}^1$ initial state, with roughly equal intensities. (*After C. Baltay et al., Annihilation of antiprotons in hydrogen at rest, Phys. Rev., 140:B1039, 1965*)

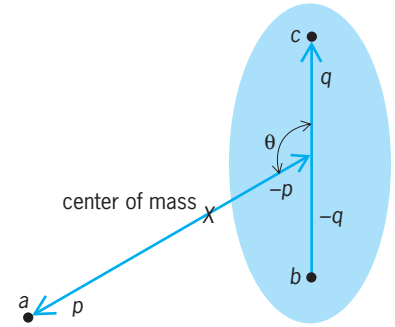

Fig. 3. Coordinate system for relativistic three-particle system. For given total energy *E*, the two momenta, *q* and *p*, are related in magnitude by the following equations: $E = \sqrt{(m_a^2+p^2)} + \sqrt{(M_{bc}^2+p^2)}$ and $M_{bc} = \sqrt{(m_b^2+q^2)} + \sqrt{(m_c^2+q^2)}$.

If the process occurs strongly through an intermediate resonance state, say $a + (bc)^*$ where $(bc)^* \rightarrow b + c$, there will be observed a "resonance band" of events for which T_a has the value appropriate to this intermediate two-body system. Such a resonance band runs parallel to the appropriate side of the triangle [the side *MN* for the case $(bc)^*$, and cyclically] and has a breadth related with the lifetime width for the resonance state.

Antiproton annihilation. The Dalitz plot shown for equal masses in Fig. 1 has been especially useful for three-pion systems, since it treats the three particles on precisely the same footing. Points placed symmetrically with respect to the symmetry axis *PL* represent configurations related by the interchange of π_b and π_c. The symmetry axes *PL*, *QM*, and *RN* divide the allowed regions into six sectors; the configurations in each sector can be obtained from those corresponding to one chosen sector (for example, the sector such that $T_a \geq T_b \geq T_c$) by the six operations of the permutation group on three objects. These operations are of particular interest for three-pion systems, since pions obey Bose statistics; the intensities in the six sectors are related with the permutation symmetry of the orbital motion in the three-pion final state. The Dalitz plot shown in **Fig. 2**, for the antiproton capture reaction $\bar{p}p \rightarrow \pi^+\pi^-\pi^0$, illustrates these points. Three ρ-meson bands [ρ(765 MeV) $\rightarrow \pi\pi$] are seen, corresponding to intermediate states $\pi^+\rho^-$, $\pi^0\rho^0$, and $\pi^-\rho^0$; the six sectors have equal intensity. *See* BOSE-EINSTEIN STATISTICS.

Relativistic three-particle system. **Figure 3** depicts a less symmetric specification for a three-particle system. The momentum of *c* in the (*bc*) barycentric frame is denoted by **q**, the momentum of *a* in the (*abc*) barycentric frame by **p**, and the angle between **p** and **q** by θ. For fixed energy T_a, the points *F* on Fig. 1 lie on a line parallel to *MN*; as cos θ varies from +1 to −1, the point *F* representing the configuration moves uniformly from the left boundary X' to the right boundary *X*. If cartesian coordinates are used for *F*, with origin *P* and *y* axis along *NM* (as has frequently been found useful in the literature), then Eqs. (6) hold. Note that, for fixed *x*, *y* is linearly related with cos θ.

$$x = T_a/Q \qquad y = (T_b - T_c)/Q\sqrt{3} \tag{6}$$

Unsymmetrical plot. The Dalitz plot most commonly used is a distorted plot in which each configuration is specified by a point with coordinates (M^2_{ab}, M^2_{bc}) with respect to right-angled axes. This depends on the relationship given in Eq. (7) and its

$$M_{bc}^2 = (E - m_a c^2)^2 - 2ET_a \tag{7}$$

cyclic permutations, for the total barycentric energy M_{bc} of the two-particle system *bc*. This plot may be obtained from Fig. 1 by shearing it to the left until *LM* is perpendicular to *MN*, and then contracting it by the factor $\sqrt{3/2}$ parallel to *MN* [which leads to a cartesian coordinate system (T_c, T_a), finally reversing the direction of the axes [required by the minus sign in Eq. (7)] and moving the origin to the point $M_{ab}^2 = M_{bc}^2 = 0$. The plots shown in **Fig. 4** correspond in this way to the relativistic curve (ii) in Fig. 1 for two values of the total energy *E*. This distorted plot retains the property that phase-space volume is directly proportional to the area on the plot. As shown in Fig. 4, the $(ab)^*$ and $(bc)^*$ resonance bands have a fixed location on this plot; data from experiments at different energies *E* can then be combined on the same plot to give a stronger test concerning the existence of some intermediate resonance state. On the other hand, the $(ca)^*$ resonance bands run across the plot at 135° and move as *E* varies, so that a different

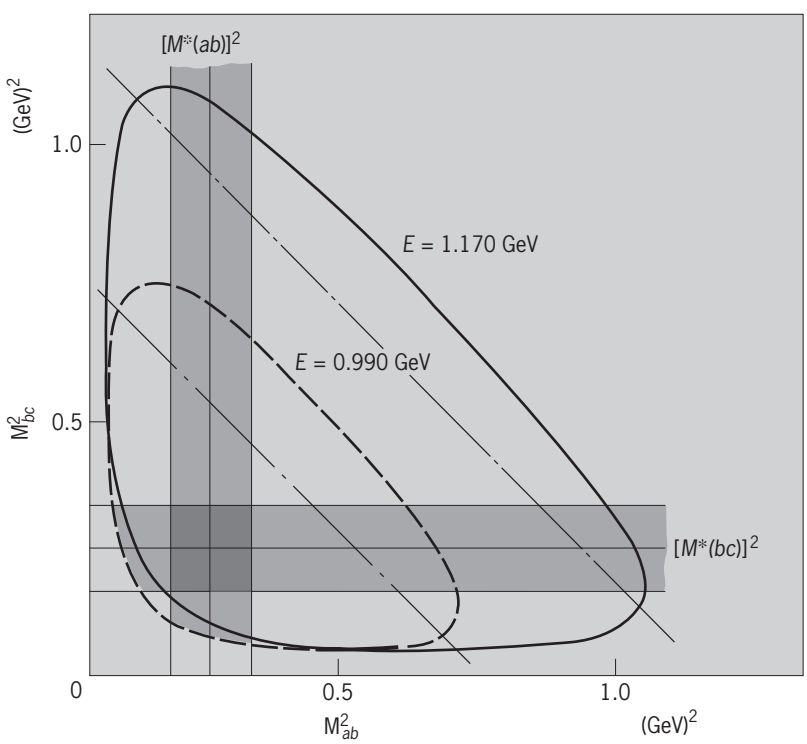

Fig. 4. An unsymmetrical Dalitz plot. Configuration of system (*abc*) is specified by a point (M^2_{ab}, M^2_{bc}) in a rectangular coordinate system, where M_{ij} denotes the barycentric energy of the two-particle system (*ij*). Kinematic boundaries have been drawn for equal masses $m_a = m_b = m_c = 0.14$ GeV and for two values of total energy *E*, appropriate to a three-pion system ($\pi^+\pi^-\pi^+$). Resonance bands are drawn for states (*ab*) and (*bc*) corresponding to a (fictitious) π-π resonance mass 0.5 GeV and full width 0.2 GeV. The dot-dash lines show the locations a (*ca*) resonance band would have for this mass of 0.5 GeV, for the two values of the total energy *E*.

choice of axes [say (M^2_{ca}, m^2_{ab})] is more suitable for their presentation.

Conclusion. It must be emphasized that the Dalitz plot is concerned only with the internal variables for the system (*abc*). In general, especially for reaction processes such as Eqs. (3) and (4), there are other variables such as the Euler angles which describe the orientation of the plane of (*abc*) relative to the initial spin direction or the incident momentum, which may carry additional physical information about the mechanism for formation of the system (*abc*). The Dalitz plots usually presented average over all these other variables (because of the limited statistics available); this sometimes leads to a clearer picture in that there are then no interference terms between states with different values for the total spin-parity. However, it is quite possible to consider Dalitz plots for fixed values of these external variables, or for definite domains for them. *See* EULER ANGLES; GOLDHABER TRIANGLE. Richard H. Dalitz

Bibliography. G. Kane, *Modern Elementary Particle Physics*, rev. ed., 1993; D. P. Roy, *Phenomenology of Elementary Particle Physics*, 1993.

Dallis grass

A general term for a genus of grasses of which the most important species is the deeply rooted perennial *Paspalum dilatatum*. Dallis grass is widely used in the southern United States, mostly for pasture, and remains productive indefinitely if well managed. Dallis grass does best on fertile soils and responds to lime and fertilizer. On heavier soils it remains green throughout the winter unless checked by heavy frosts. Seed production is hampered by infection with ergot, a fungus that invades developing seeds and produces purplish-black horny bodies. Ergot-bearing seed heads are very toxic to livestock, whether in pasture or in hay, and Dallis grass must be so managed as to prevent consumption of infected heads by livestock. *See* CYPERALES; ERGOT AND ERGOTISM; FERTILIZER; GRASS CROPS; PLANT PATHOLOGY. Howard B. Sprague

Dalton's law

The total pressure of a mixture of gases is the sum of the partial pressures of each gas in the mixture. The law was established by John Dalton (1766–1844). In his original formulation, the partial pressure of a gas is the pressure of the gas if it alone occupied the container at the same temperature. Dalton's law may be expressed as $P = P_A + P_B + \cdots$, where P_J is the partial pressure of the gas J, and P is the total pressure of the mixture; this formulation is strictly valid only for mixtures of ideal gases. For real gases, the total pressure is not the sum of the partial pressures (except in the limit of zero pressure) because of interactions between the molecules.

In modern physical chemistry the partial pressure is defined as $P_J = x_J P$, where x_J is the mole fraction of the gas J, the ratio of its amount in moles to the total number of moles of gas molecules present in the mixture. With this definition, the total pressure of a mixture of any kind of gases is the sum of their partial pressures. However, only for an ideal gas is the partial pressure (as defined here) the pressure that the gas would exert if it alone occupied the container. *See* AVOGADRO'S NUMBER; GAS; KINETIC THEORY OF MATTER; THERMODYNAMIC PRINCIPLES. P. W. Atkins

Bibliography. P. W. Atkins, *Physical Chemistry*, 6th ed., 1998; L. Jones and P. W. Atkins, *Chemistry: Molecules, Matter, and Change*, 4th ed., 2000.

Dam

A barrier or structure across a stream, river, or waterway for the purpose of confining and controlling the flow of water. Dams vary in size from small earth embankments for farm use to high, massive concrete structures for water supply, hydropower, irrigation, navigation, recreation, sedimentation control, and flood control. As such, dams are cornerstones in the water resources development of river basins. Dams are now built to serve several purposes and are therefore known as multipurpose (**Fig. 1**). The construction of a large dam requires the relocation of existing highways, railroads, and utilities from the river valley to elevations above the reservoir. The two principal types of dams are embankment and concrete. Appurtenant structures of dams include spillways, outlet works, and control facilities; they may also include structures related to hydropower and other project purposes. *See* ELECTRIC POWER GENERATION; IRRIGATION (AGRICULTURE); WATER SUPPLY ENGINEERING.

Dams have been built since ancient times, and improvements were made at varying intervals as engineering technology developed. However, very rapid advances occurred in the twentieth century as a result of developments in the use of concrete, soil mechanics, and construction equipment. In the early 1900s, concrete dams became thinner, and a new era of thin arch dams began. Earth and rock–fill embankment dams became economical during and after World War II. In 1980, an innovative method of using earth-moving and compacting equipment to place dry concrete (roller-compacted concrete) greatly improved the economics of concrete dams. As of December 1997, there was a total of 100,781 dams in the United States. Most of these are small recreational projects and farm ponds. Of this total, 6389 dams have a height of 50 ft or more and 1586 dams have a height of 100 feet or more. Numerous dams have been constructed in various countries worldwide (**Table 1**). Many dams possess considerable height, volume, and reservoir capacity (**Table 2**).

Purposes. Dams are built for specific purposes. In ancient times, they were built only for water supply or irrigation. Early in the development of the United States, rivers were a primary means of transportation, and therefore navigation dams with locks (**Fig. 2**) were constructed on the major rivers. Dams

Fig. 1. John Day Lock and Dam, looking upstream across the Columbia River at Washington shore. In the foreground the navigation lock is seen, then the spillway beyond it, and then the powerhouse. The John Day multiple-purpose project has the highest single-lift navigation lock in the United States. (*U.S. Army Corps of Engineers*)

have become more complex to meet large power demands and other needs of modern countries (**Figs. 3** and **4**). Although recreation is a popular purpose of small private dams, it is planned as a benefit with an assigned monetary value in federal projects in the United States. A typical summary of purposes of nonfederal dams is shown in **Table 3**.

Features. In addition to the standard impounded reservoir and the appurtenant structures of a dam (spillway, outlet works, and control facility), a dam with hydropower requires a powerhouse, penstocks, generators, and switchyard. The inflow of water into the reservoir must be monitored continuously, and the outflow must be controlled to obtain maximum benefits. Under normal operating conditions, the reservoir is controlled by the outlet works, consisting of a large tunnel or conduit at stream level with control gates. Under flood conditions, the reservoir is maintained by both the spillway and outlet works.

The reservoir level of a flood control dam is maintained as low as possible to create the maximum amount of storage space for use in the flood season. For an irrigation project, the reservoir is filled as high as possible in the winter and early spring, and it is maintained at that level for maximum release of

Fig. 2. Lock and Dam 1 on the Mississippi River between St. Paul and Minneapolis, Minnesota. Built in 1917 with a single lock to provide barge traffic to Minneapolis, the dam failed in 1929 and was rebuilt with twin locks in 1932. (*U.S. Army Corps of Engineers*)

TABLE 1. Countries with the largest number of dams*

Country	Number of dams†
China	18,595
United States	6,389
Japan	2,142
India	1,085
Spain	690
Republic of Korea	628
Canada	580
United Kingdom	529
Brazil	489
Mexico	487
France	432
Italy	408
Australia	374

*From "ICOLD Register of Dams," updated 1996.
†Dams at least 50 ft (15 m) high.

TABLE 2. Landmark dams of the world

Order	Name	River	Country	Type*	Height ft	Height m	Year completed†
			HIGHEST DAMS				
1	Rogun	Vakhsh	Tajikistan	E-R	1099	335	1985
2	Nurek	Vakhsh	Tajikistan	E	984	300	1980
3	Xiaowan	Lancang	China	A	958	292	UC
4	Grand Dixence	Dixence	Switzerland	G	935	285	1961
5	Inguri	Inguri	Georgia	A	892	272	1980
6	Vajont	Vajont	Italy	A	860	262	1961
7	Manuel M. Torres	Grijalva	Mexico	R	856	261	1980
8	Tehri	Bhagirathi	India	E-R	856	261	2006
9	Alvaro Obregon	Tennasco	Mexico	G	854	260	1946
10	Mauvoisin	Drance de Bangnes	Switzerland	A	820	250	1957
11	Sayano-Shushensk	Yenisei	Russia	A	804	245	1989
12	Alberto Lieraso	Guavio	Colombia	R	797	243	1990
13	Mica	Columbia	Canada	E	794	242	1972
14	Ertan	Yangtze/Yalong	China	A	787	240	1998
15	La Esmeralda	Bata	Colombia	R	777	237	1975
16	Kishau	Tons	India	G	774	236	UC
17	Oroville	Feather	United States	E	770	335	1968
18	El Cajon	Comayagua	Honduras	A	768	234	1984
19	Chirkei	Sulak	Russia	A	764	233	1978
20	Shuibuya	Qingjiang	China	R	764	233	UC

			LARGEST EMBANKMENT DAMS BY VOLUME		Volume 10^3 yd^3	Volume 10^3 m^3	Year completed
1	Chapeton	Paraná	Argentina	E-G	387,400	296,200	UC
2	Pati	Paraná	Argentina	E-G	311,411	238,100	UC
3	Tarbella	Indus	Pakistan	E-R	194,223	148,500	1967
4	Fort Peck	Missouri	United States	E	125,624	96,050	1937
5	Lower Usuma	Usuma	Nigeria	E	121,635	93,000	1990
6	Cipasang	Cimanuk	Indonesia	E-R	117,842	90,100	UC
7	Tucurui	Tocantis	Brazil	E-R	84,098	85,200	1984
8	Ataturk	Firat	Turkey	R	111,172	85,000	1990
9	Yacyreta	Parana	Paraguay/Argentina	E	105,944	81,000	1997
10	Guri	Caroni	Venezuela	R-G	102,014	78,000	1986
11	Rogun	Vakhsh	Tajikistan	E-R	98,746	75,500	1985
12	Oahe	Missouri	United States	E	91,996	70,339	1958
13	Parambikulam	Parambikulam	India	E-G	90,461	69,165	1967
14	Mangla	Jhelum	Pakistan	E	85,497	65,370	1967
15	Gardiner	S. Saskatchewan	Canada	E	85,014	65,000	1968
16	Afsluitdijk	Zuiderzee	Netherlands	E	82,921	63,400	1932
17	Mangla	Jhelum	Pakistan	E	82,893	63,379	1967
18	Oroville	Feather	United States	E	77,997	59,635	1968
19	San Luis	San Luis	United States	E	77,897	59,559	1967
20	Nurek	Vakhsh	Tajikistan	E	75,858	58,000	1980

			LARGEST ARTIFICIAL RESERVOIRS IN TERMS OF CAPACITY	Capacity 10^3 acre ft‡	Capacity 10^5 m^3	Year completed
1	Owen Falls	Victoria Nile	Uganda	166,033	296,200	1954
2	Kariba	Zambezi	Zimbabwe/Zambia	146,090	180,200	1959
3	Bratsk	Angara	Russia	137,010	169,000	1964
4	High Aswan	Nile	Egypt	131,335	162,000	1970
5	Akosombo	Volta	Ghana	119,952	147,960	1965
6	Daniel Johnson	Manicouagan	Canada	115,000	141,851	1968
7	Guri	Caroni	Venezuela	109,446	135,000	1986
8	Krasnoyarsk	Yenisei	Russia	59,425	73,300	1967
9	WAC Bennett	Peace	Canada	57,000	70,309	1967
10	Zeya	Zeya	Russia	55,452	68,400	1978
11	Cabora Bassa	Zambezi	Mozambique	51,075	63,000	1974
12	La Grande 2	La Grande	Canada	50,033	61,715	1978
13	Chapeton	Paraná	Argentina	49,129	60,600	UC
14	La Grande 3	La Grande	Canada	48,659	60,020	1981
15	Ust-Ilim	Angara	Russia	48,075	59,300	1977
16	Boguchany	Angara	Russia	47,183	58,200	UC
17	Kuibyshev	Volga	Russia	47,021	58,000	1955
18	Serra da Mesa	Tocantins	Brazil	44,103	54,400	UC
19	Caniapiscau Barrage KA 3	Caniapiscau	Canada	43,608	53,790	1980
20	Upper Wainganga	Wainganga	India	41,103	50,700	UC

*E = embankment, earth-fill. R = embankment, rock-fill. E-R = embankment, earth and rock-fill. G = gravity. A = arch.
†UC = under construction.
‡1 acre foot = 1 acre of water 1 foot deep.

Fig. 3. Guri Dam, Venezuela, showing construction under way on the second stage of the concrete gravity dam. Embankment dams are shown on the left and right abutments. The concrete batch plant is shown at left with a trestle for transporting concrete shown on the downstream face of the dam. The second powerhouse with an excavated outlet channel is shown at the center of the photo. (*C.V.G. Electrification del Caroni, C. A. EDELCA, Caracas, Venezuela*)

water during the dry season. The reservoir level of a hydropower dam is maintained as constant as feasible to create a uniform head for power generation. Water quality is an important ingredient in sustaining a balance in nature and is taken into account in modern dam design, construction, and operation. The chemical quality and temperature of the water are monitored in the reservoir. Intake ports at various depths allow selective withdrawal and mixing to produce the desired temperature and oxygen content, in order to enhance downstream environmental conditions. Fish ladders, that is, stepped series of elevated pools, are provided at many dams to allow free passage of fish upstream and downstream. Screens are used to keep fish out of the turbines. *See* RESERVOIR.

The discharge from modern dams must be managed carefully and continuously. During floods, reservoir inflows may exceed maximum discharges and cause reservoir levels to rise. To prevent dams from overtopping and possibly failing, spillways are provided to pass floodwater safely. They are commonly built at elevations just below dam crests and without gates. These uncontrolled, ungated spillways are designed to allow all of the excess water to pass.

Fig. 4. Itaipu Dam, on the Paraná River between Brazil and Paraguay. An operating spillway with a flip bucket on the Paraguay side of the river is shown at left. The hollow concrete gravity dam with powerhouse construction at the downstream toe is shown at center. (*G. S. Sarkaria, International Engineering Company, Inc.*)

TABLE 3. Primary purpose of nonfederal dams in the United States*

Purpose	Percent of total
Recreation	35
Flood control	15
Water supply	10
Irrigation	10
Farm and livestock water	15
Power	3

*From "National Dam Inventory," 1996 update.

Fig. 5. Aerial view of North Fork Dam, a combination earth-and-rock embankment of the North Fork of Pound River, Virginia. The channel-type spillway (left center) has a simple overflow weir. (*U.S. Army Corps of Engineers*)

In other cases, spillways are constructed at even lower levels and contain gates that are operated from the control facilities. The tops of these gates are lower than the dam crests, thus allowing some control of floodwater.

Penstocks (usually steel pipes or concrete-lined tunnels) are used to convey water from the reservoir through or around the dam to the powerhouse. The penstocks are connected to the turbines, and the water flow is controlled by valves. The number and size of the penstocks vary, depending on the number of generators and amount of water needed.

All the features of a dam are monitored and operated from a control room. The room contains the necessary monitors, controls, computers, emergency equipment, and communications systems to allow project personnel to operate the dam safely under all conditions. Standby generators and backup communications equipment are necessary to operate the gates and other reservoir controls in case of power failure. Weather conditions, inflow, reservoir level, discharge, and downstream river levels are also monitored. In addition, the control room monitors instrumentation located in the dam and appurtenant features that measures their structural behavior and physical condition.

Requirements. All dams are designed and constructed to meet specific requirements. First, a dam should be built from locally available materials when possible. Second, the dam must remain stable under all conditions, during construction, and ultimately in operation, both at the normal reservoir operating level and under all flood and drought conditions. Third, the dam and foundation must be sufficiently watertight to control seepage and maintain the desired reservoir level. Finally, it must have sufficient spillway and outlet works capacity as well as freeboard to prevent floodwater from overtopping it.

Types. Dams are classified by the type of material from which they are constructed. In early times, the materials were earth, large stones, and timber, but as technology developed, other materials and construction procedures were used. Most modern dams fall into two categories: embankment and concrete. Embankment dams are earth or rock-fill; other gravity dams and arch and buttress dams are concrete.

Earth-fill dam. Earth is the predominant material in this type of embankment dam. Earth dams are further classified by the construction method: hydraulic-fill or rolled-fill. A hydraulic-fill dam is one in which the soil is excavated, transported, and placed by flowing water. A large dredge operating in the river or other borrow area pumps a slurry of earth and water to the damsite. Here the coarse-grained materials settle on the outside portion of the embankment, and the remaining slurry is allowed to pond at the center, where the very fine-grained clay-size particles settle to form the impervious portion of the dam.

Advances in earth-moving construction equipment during World War II led to widespread construction of rolled-earth-fill dams. Economic advantages of this type of embankment often include the use of material available from the site excavation as embankment material, and the ready availability of fill material at or near the damsite. Other advantages of earth-fill dams include their adaptability to a wide variety of site configurations and their tolerance of weak foundations.

At various stages during excavation and placement of the fill, the moisture content of the soil may be adjusted by wetting or drying in order to optimize its performance in the finished embankment. The soil is spread on the embankment in uniform layers 8–12 in. (20–30 cm) thick and compacted with sheepsfoot or rubber-tired rollers. The rollers make from four to eight passes, depending on the desired

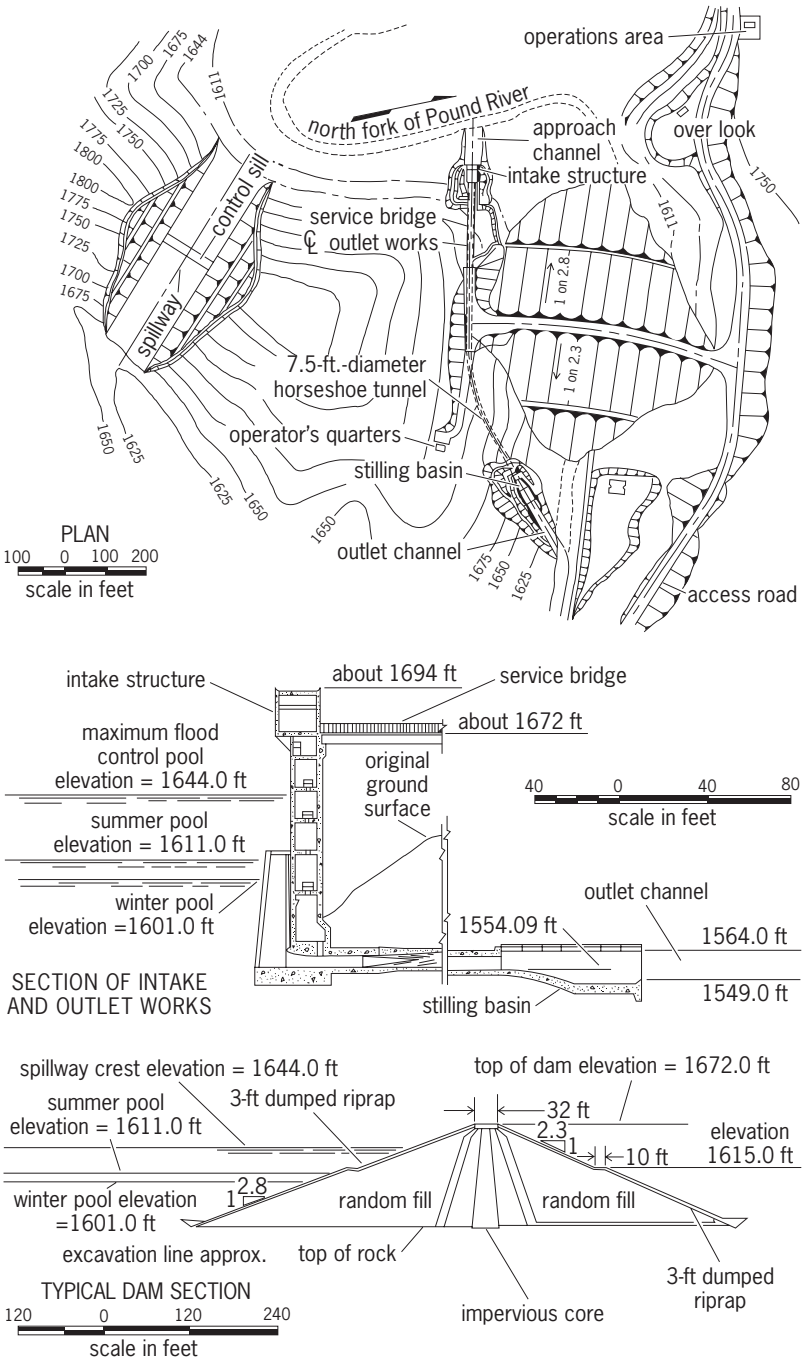

Fig. 6. Plan and sections of North Fork of Pound Dam, Virginia. 1 ft = 0.3 m. (*U.S. Army Corps of Engineers*)

density. Typical dry densities for rolled earth fill range 100–130 lb/ft^3 (1602–2083 kg/m^3).

Seepage control is an important aspect of earth-dam design. In early times, earth embankments were homogeneous, and seepage would emerge on the downstream slope just above ground level. If uncontrolled, such seepage can move soil particles and cause failure. In 1940, filter criteria were developed after careful scientific tests on all types of soil and on the sands and gravel to be used as filter material. These criteria allow engineers to design internal drains to collect and remove seepage. When an earth dam is built on a site where bedrock is at considerable depth, the foundation must be treated to control seepage. Typical treatment includes one or a combination of several things: upstream impervious blanket, cutoff wall, drainage blanket, gravel drains excavated into the foundation at the downstream toe, and relief wells. *See* FOUNDATIONS.

Earth-fill dams are by far the most popular type in the world. They make up 78%, or 27,260, of all those dams at least 50 ft (15 m) high. The earth dam's spillway is usually located in adjacent terrain rather than in the dam itself. The outlet works are either a conduit in the valley or a tunnel in one of the abutments (**Figs. 5** and **6**). Excavation for the spillway and outlet works usually produces large quantities of rock. As a result, the use of both earth and rock in an embankment is a common practice.

Rock-fill dam. A rock-fill dam is a rolled fill embankment composed of fragmented rock with an impervious zone or membrane located on the upstream face or incorporated in the center of the embankment. The impervious membrane is typically a concrete slab or asphalt layer on the upstream face. The impervious zone is typically a thin internal core of earth fill. The earth core is separated from the rock shell (the structural mass of the dam) by zones of small rock fragments or gravel, to prevent the earth from washing into the rock fill, and a drain to control seepage. In Europe, it is common to use asphalt for the impervious zone.

Rock-fill dams require solid rock foundations and sites where large quantities of rock are available. Seepage through rock foundations is prevented or minimized by grout curtains. Rock-fill dams are usually more economical than concrete gravity dams at sites having wide valleys and adequate foundations. Spillways and outlet works are at locations similar to those of earth-fill dams. The primary advantage of rock-fill dams is that they require less material. Rock fill has a higher shear strength than earth fill and therefore permits steeper exterior slopes.

Rock for the embankment is normally excavated by drilling and blasting. Hole spacing and powder charges are set to produce a particular gradation of rock fragments for the dam. The rock is placed and spread in the same manner as earth but in thicker layers, 18–36 in. (45–90 cm). The material is normally compacted by weighted or vibratory steel drum rollers. Dry densities of rock fill are normally in the range 110–145 lb/ft^3 (1762–2323 kg/m^3).

Rock-fill dams became popular in the United States during the California gold rush in the 1860s and 1870s, when many dams were built in remote locations to store water for use in hydraulic sluicing. Of the 34,780 dams in the world that are at least 50 ft (15 m) high, 1590 are rock-fill embankments.

Concrete gravity dam. Concrete gravity dams are massive structures, characterized by vertical or near-vertical upstream faces and steep downstream faces (**Figs. 7** and **8**). They are designed with enough weight to resist being overturned or moved by the force of the water in the reservoirs. They are economical only at sites with shallow, high-strength rock foundations. Because of the large volumes of

concrete involved, adequate sources of high-quality aggregates must be available near the sites. Concrete is composed of water, cement, pozzolan, aggregates, and entrained air. These ingredients are proportioned to produce concrete of the desired workability, durability, and strength as economically as possible. The density of concrete in dams typically ranges 140–160 lb/ft^3 (2243–2563 kg/m^3). An important feature of gravity dams is the simplicity with which safe spillways and outlet works can be provided. *See* CONCRETE.

The design of a concrete gravity dam is controlled by stability considerations and internal stresses. The structure must be able to resist water, sediment, and ice pressures from the reservoir, as well as earthquake forces. Computers permit rapid solutions of complex equations for determining the magnitude and distribution of internal stresses. These dams are built in monolithic units by using the block method of construction (**Fig. 9**). This promotes dissipation of heat produced by hydration of the cement (chemical combination with the water) and thus helps minimize the volume changes associated with overheating that cause tensile stresses and cracking. The blocks are separated by construction joints. In building a block, the concrete is placed in horizontal layers and vibrated to eliminate voids. The monoliths are cast on top of firm rock foundations that have been cleaned with water and treated by placement of cement and water slurry or grout in the cracks and joints.

A concrete gravity dam usually contains an internal gallery large enough to allow for physical inspection and for collection of drainage from downstream drain holes drilled into the foundation. Grout holes to reduce seepage in the foundation are also drilled from the gallery in the vertical or upstream direction. The grout is injected under pressure to force it into all joints and openings encountered at depth.

Arch dam. The arch dam is a thin concrete dam that curves upstream from each abutment (**Figs. 10** and **11**). Such dams are classified as thin, medium, or thick arch, depending on the ratio of structural height to base thickness. The ratio is 0.2 or less for a thin arch, 0.25 for a medium arch, and 0.3 or greater for a thick arch. The arch transmits the water pressure and other loads directly to the abutments and foundation. It contains significantly less concrete than a concrete gravity dam of the same height and length. Relatively narrow canyons favor the use of arch dams.

The shape of early arch dams was controlled by construction materials available at the time, and by less sophisticated understanding of structural behavior and the way that loads were transmitted through the curved structures to the foundations. As a result, arches were simple masonry structures with curved alignments and near-vertical upstream faces. This type was popular among water companies supplying domestic and irrigation water.

Beginning in the 1900s, improved structural analysis and actual performance records led to the use of variable-thickness arch dams. Varying the thickness can reduce the volume of concrete required. Measurement of the physical properties of concrete began in the late 1920s. This led to improved design procedures and measurement of actual performance with such instruments as strain gages. The concept of working stresses emerged in the late 1920s. The double-curvature shape (curved top to bottom as well as transversely) emerged in the mid-1950s. Vertical curving and shaping of the arch improves stress distribution. Making the compressive stresses levels throughout the dam as close as possible to the maximum allowable stress results in the minimum volume of concrete. A symmetrical profile is desirable. This may require excavation on one abutment if the canyon is not symmetrical. The economic upper limit of the length-to-height ratio of an arch dam lies between 4:1 and 6:1.

Fig. 7. Green Peter Dam, a concrete gravity type on the Middle Santian River, Willamette River Basin, Oregon. A gate-controlled overflow-type spillway is constructed through the crest of the dam; the powerhouse is at the downstream toe of the dam. (*U.S. Army Corps of Engineers*)

Buttress dam. The buttress dam consists of two principal structural elements: a sloping upstream deck that retains the water, and buttress walls that support the deck and transmit the loads to the foundation. Traditionally, buttress dams have been classified into three categories: flat slab, multiple arch, and massive head. The flat-slab type consists of a reinforced concrete flat slab inclined at about 45° and connected to buttresses. The multiple-arch type is a series of concrete arches spanning the buttresses. The massive-head type has a large mass of concrete in the section upstream from the buttresses.

In 1918, the flat-slab design was patented in the United States. About 200 buttress dams of all three categories have been built in the United States. Many are less than 150 ft (45 m) high. Some landmark

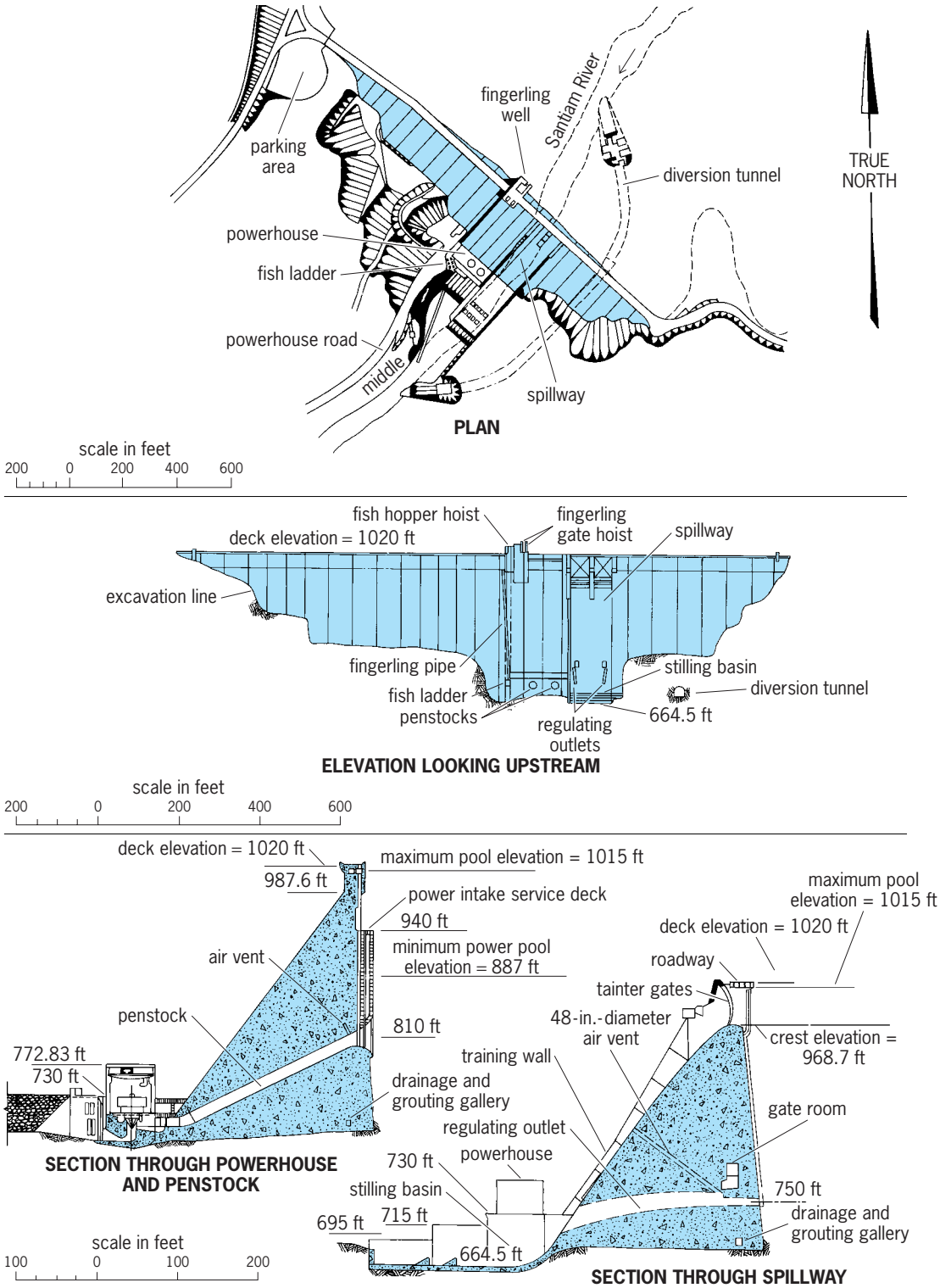

Fig. 8. Plan and sections of Green Peter Dam. 1 ft = 0.3 m, 1 in. = 2.5 cm. (*U.S. Army Corps of Engineers*)

buttress dams are the Daniel Johnson in Canada (1986), 702 ft (214 m) high, the world's highest multiple-arch buttress dam; and the José M. Oriol in Spain (1969), 426 ft (130 m), the world's highest flat-slab buttress dam.

Site and type selection. The type of dam for a particular site is selected on the basis of technical and economic data and environmental considerations. In the early stages of design, several sites and types are considered. Drill holes and test pits at each site provide soil and rock samples for testing physical properties. In some cases, field pumping tests are performed to evaluate seepage potential. Preliminary designs and cost estimates are prepared and reviewed by hydrologic, hydraulic, geotechnical, and structural engineers, as well as geologists. Environmental quality of

Fig. 9. Block method of construction on a typical concrete gravity dam. (*U.S. Army Corps of Engineers*)

the water, ecological systems, and cultural data are also considered in the site-selection process.

Factors that affect the type are topography, geology, foundation conditions, hydrology, earthquakes, and availability of construction materials. The foundation of the dam should be as sound and free of faults as possible. Narrow valleys with shallow sound rock favor concrete dams. Wide valleys with varying rock depths and conditions favor embankment dams. Earth dams are the most common type. *See* ENGINEERING GEOLOGY; FAULT AND FAULT STRUCTURES.

Construction process. Hydraulic-fill operations over a 4-year period at Fort Peck Dam, on the Missouri River, the largest embankment by volume in the United States (**Fig. 12**), dredged 156,000,000 yd^3 (119,340,000 m^3) of material. Of this volume, 122,000,000 yd^3 (93,333,000 m^3) was retained in the embankment. Large conventional excavation operations can produce hourly volumes of 2000–3000 yd^3 (1530–2293 m^3). Processing, hauling, placement, and compaction operations for earth or rock result in daily placement rates that vary from as low as 2500 yd^3 (1911 m^3) on small dams to 6500 yd^3 (4969 m^3) on larger dams.

The materials and construction procedures for concrete dams evolved gradually from the early dams in Asia and Europe to the modern massive concrete dams. Prior to 1900, portland cement used in the United States was imported from England. Thus, the early concrete dams built in the United States were masonry. Generally, the concrete was mixed and transported in wheelbarrows. In the case of cyclopean masonry, large irregular blocks of rock with mortar, small derricks were erected, and the maximum rate of placement approached a few hundred cubic yards a day. There was no attempt to cure the concrete. Between 1900 and 1930, concrete was placed by towers and chutes. Portland cement had become available in the United States, and placement rates improved. However, little attention was given to the mix design, and wet mixes that could easily flow in chutes were widely used.

Hoover Dam, on the Colorado River, was a major turning point in both the design and construction of concrete dams. Its unprecedented size,

Fig. 10. East Canyon Dam, a thin-arch concrete structure on the East Canyon River, Utah. There is an uncontrolled overflow-type spillway through the crest of the dam at the right. (*U.S. Bureau of Reclamation*)

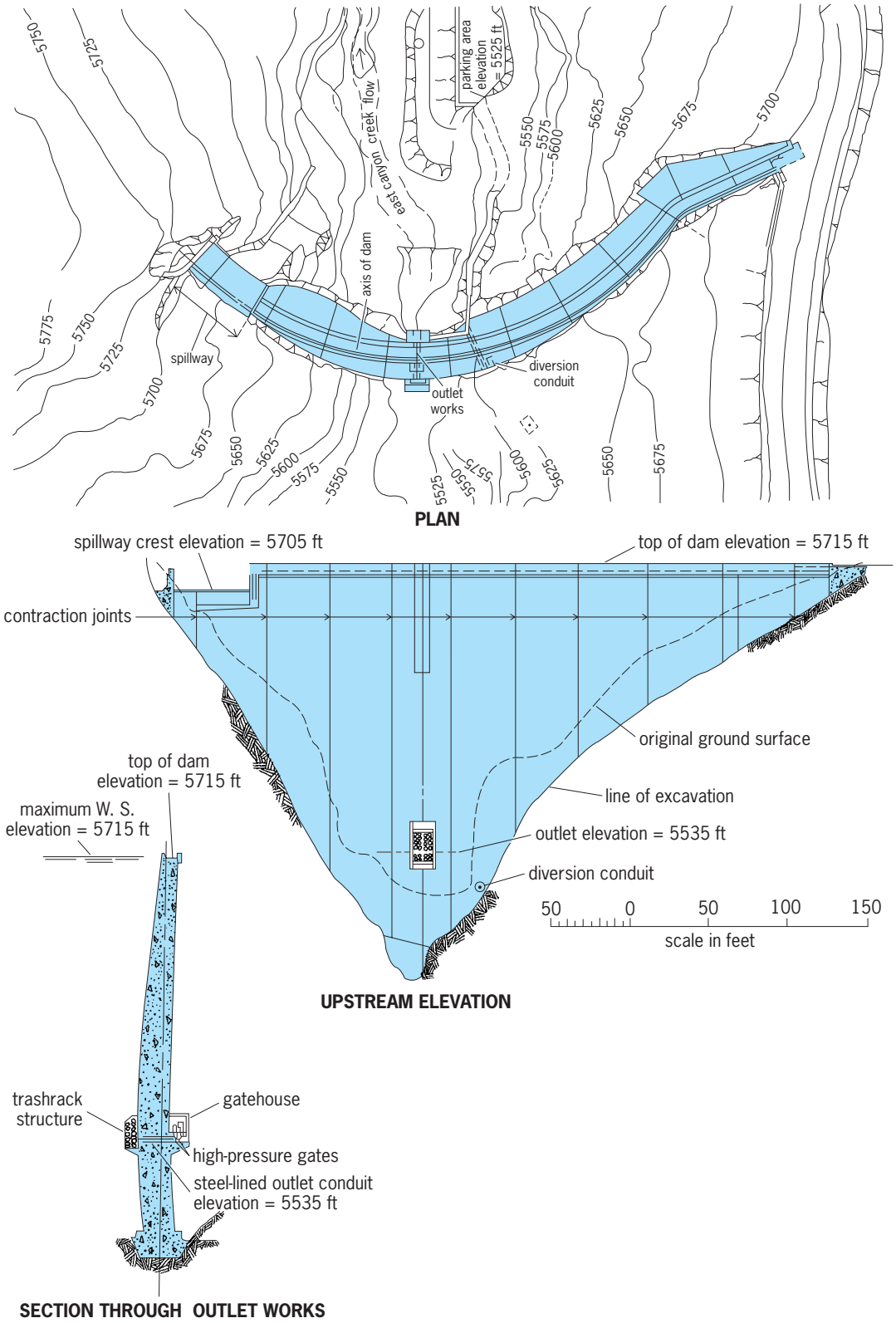

Fig. 11. Plan and sections of East Canyon Dam, Utah. 1 ft = 0.3 m. (*U.S. Army Corps of Engineers*)

4,400,000 yd^3 (3,363,800 m^3), led to the introduction of mass concrete placement. Average placement rates of 10,000 yd^3 (7645 m^3) per day were achieved. Advances in design resulting from the Hoover project led to the construction of Grand Coulee, on the Columbia River, 10,585,000 yd^3 (8,099,000 m^3). Two large concrete plants were used that supported a maximum placement of 20,680 yd^3 (15,810 m^3) per day and an average rate over the construction period of 6000 yd^3 (4587 m^3) per day.

Since 1980, the technology of placing dry concrete with paving equipment and compacting it with rollers has gained wide acceptance (**Fig. 13**). This construction method is known as roller-compacted concrete. By 1997, 30 dams at least 50 ft (15 m) high had been constructed in the United States using the roller-compacted concrete method. As of January 1998, there were 16 RCC dams in the world having a height of 100 m or greater. This method produces the high placement rates usually associated with earth-fill construction and results in economical structures. It was initially used in 1975 in the tunnel repairs at Tarbela Dam, Pakistan, and placement reached a maximum rate of 24,000 yd^3 (18,343 m^3).

An outstanding example of the rapid construction achieved by using roller-compacted concrete is Copperfield Dam in Australia. This 131-ft-high (40-m) dam contains 183,000 yd^3 (140,000 m^3) and required only 10 months from initial design to completion. The project was originally designed as an earth- and rock-fill dam, but it was switched to roller-compacted concrete for greater economy.

River diversion during construction. The designers of a dam must consider the stream flow around or through the damsite during construction. Stream flow records provide the information for use in determining the largest flood to divert during the selected construction period. One common practice for diversion involves constructing the permanent outlet works, which may be a conduit or a tunnel in the abutment, along with portions of the dam adjacent to the abutments, in the first construction period. In some cases, a temporary channel is built at a preferred diversion location, and levees are built to control the flow of water through the damsite. After the outlet works and main dam are completed to an appropriate level, the stream is diverted into the outlet works by a cofferdam high enough to prevent overtopping during construction. A downstream cofferdam is also required to keep the damsite dry. In the final construction period, the entire dam is brought to full height. *See* COFFERDAM.

Operation and maintenance. Personnel responsible for operation and maintenance of the dam become involved during the final design and construction to become familiar with design details that relate to operation. The operating instructions and maintenance schedule are published in a formal document for each dam. A schedule is established for collection and reporting of data for climatic conditions, rainfall, snow cover, stream flows, and water quality of the reservoir, as well as the downstream reaches. All these data are evaluated for use in reservoir regulation. Another schedule is established for the collection of instrumentation data used to determine the structural behavior and physical condition of the dam. These data are evaluated frequently.

Routine maintenance and inspection of the dam and appurtenant structures are ongoing processes. The scheduled maintenance is important to preserve the integrity of the mechanical equipment.

Periodic inspection and evaluation. Upon completion of construction, the project is inspected in detail by a team made up of the designers, construction managers, operations personnel, and other experts. The purpose is to ensure that the dam has been built as designed and can safely impound water and that all systems are ready for the initial reservoir filling and operation. In addition, the same team conducts an in-depth inspection once a year for about 5 years after completion and at 5-year intervals thereafter. Design criteria and performance of the dam as measured by instruments are reviewed during the life of the dam, and structural reanalyses are made when necessary. Photographs are taken to record rates of deterioration.

Fig. 12. Fort Peck Dam, Missouri River, Montana, the largest embankment dam by volume in the United States, 125,628,000 yd^3 (95,625,000 m^3). A hydraulic-fill dam, it was built between 1935 and 1939 for flood control, hydropower, irrigation, and navigation. (*U.S. Army Corps of Engineers*)

Fig. 13. Galesville Dam, Oregon, showing roller-compacted concrete construction; completed in 1985, it has a height of 157 ft (48 m) and a volume of 161,000 yd^3 (123,100 m^3).

The intake structure, trash racks, emergency gates, outlet conduit or tunnel, and stilling basin are normally under water and therefore require special procedures such as dewatering prior to the inspection. At normal velocity, the flowing water can severely erode soil and rock in the approach and discharge channels. High-velocity flow over small irregularities can cause a phenomenon known as cavitation, which can lead to rapid erosion of metal and concrete and can threaten the safety of the outlet works. *See* CAVITATION.

Instrumentation. As the technology of dam design and construction progressed, the need to measure performance and structural behavior became important in order to verify the design. Advances in instruments starting in the 1950s gave the designer a valuable tool. Instrumentation gave the engineer knowledge of how the temperatures from hydration in concrete varied and the effect on strength. Pressure cells were developed that gave information about the interaction between soil backfill and a concrete wall or structure as well as the actual load distribution. Piezometers (devices to measure water level), settlement plates, and slope indicators are used in measuring the performance of embankment dams. Plumb lines, strain gages, and uplift cells are used for the same purpose in concrete dams. In addition, instruments are used to measure vertical and horizontal movement, alignment and plumb, stresses, strains, water pressure, seismic effects, and the quantity and clarity of seepage.

Instrumentation for a dam is installed at first in the design phase to establish baseline data, then during construction and throughout the life of the dam as conditions warrant. The frequency with which instrumentation data are obtained is an extremely important issue and depends on operating conditions. Timely collection and evaluation of data are critical for periods when the loading changes, such as during floods and after earthquakes. Advances in applications of remote sensing to instrumentation have made real-time data collection possible. This is a significant improvement for making dam safety evaluations.

Safety. Throughout history there have been instances of dam failure and discharge of stored water, sometimes causing considerable loss of life and great damage to property. Failures have generally involved dams that were designed and constructed to engineering standards acceptable at the time. Most failures have occurred with new dams, within the first five years of operation.

As dam technology advanced with increasing knowledge of design principles and better understanding of foundation and material properties, dams became safer. There is no question that they can be built and operated safely. The major issue is to monitor deterioration as the structures and equipment get older. In earlier times, the sizes of spillways and outlet pipes had to be determined by judgment. As a result, overtopping was the main cause of dam failure. Little was known about soil mechanics and slope stability, and so slides and slope failures were common. Beginning in the 1930s, statistical methods were used to predict floods. Advances in soil mechanics in the later 1930s and early 1940s produced new methods of stability analysis that revolutionized the slope design for excavations and earth embankment dams. Historical data indicate that the causes of failure (in the order of their significance) are piping, overtopping, slope instability, conduit leakage (outlet works), and foundation failures. *See* PIPELINE; ROCK MECHANICS; SOIL MECHANICS.

TABLE 4. Some major dam failures in history

Dam	Country	Year
San Ildefonso	Bolivia	1626
Puentes	Spain	1802
Bradfield	England	1864
Johnstown	United States	1889
Walnut Grove	United States	1890
Bouzey	France	1895
Austin	United States	1911
Bila Densa	Czechoslovakia	1916
Tigra	India	1917
Glenco	Italy	1923
St. Francis	United States	1928
Alla S. Zerbimo	Italy	1935
Fred Burr	United States	1948
Malpasset	France	1959
Kuala Lumpur	Malaya	1961
Babi Yar	Soviet Union	1961
Baldwin Hills	United States	1963
Vaiont	Italy	1963
Buffalo Creek	United States	1972
Teton	United States	1976
Kelly Barnes	United States	1977
Machhu II	India	1979

It is estimated that about 150,000 dams around the world present a potential hazard to life or property; there have been 200 failures since 1900. Many of these have involved small dams. **Table 4** lists some major failures that resulted in considerable loss of life.

Dam failures cause loss of life and property damage in downstream reaches that are beyond the control of the dam owner or local government. For this reason, and because dam safety practices should apply to all dams, national governments have become involved in order to provide supervision and standardize regulations. The United States government published "Federal Guidelines for Dam Safety" in June 1979. This initiated a coordinated effort in management practices among federal agencies and set an example for private organizations that own dams. The International Commission on Large Dams (ICOLD) was formed in 1928 by 6 countries with the purpose of developing and exchanging dam design experience, and it has grown to 76 member countries. In 1982 ICOLD established a committee on dam safety to define common safety principles, integrate efforts, and develop guidelines, and in 1987 ICOLD published "Dam Safety Guidelines." Arthur H. Walz, Jr.

Bibliography. Federal Emergency Management Agency, *Federal Guidelines for Dam Safety*, FEMA 93, June 1979; International Commission on Large Dams, *Dam Safety Guidelines*, Bull. 59, ICOLD, 1987; International Commission on Large Dams,

Dam Failures, Statistical Analysis, 1995; International Commission on Large Dams, *World Register of Dams*, 1997; R. B. Jansen, *Advanced Dam Engineering for Design, Construction, and Rehabilitation*, 1988; R. B. Jansen, *Dams and Public Safety*, U.S. Department of the Interior, 1980; E. B. Kollgaard and W. L. Chadwick (eds.), *Development of Dam Engineering in the in the United States*, 1988; J. Sherard et al., *Earth and Earth-Rock Dams*, 1967; U.S. Bureau of Reclamation, *Design of Small Dams*, 3d ed., 1987.

Damping

A term broadly used to denote either the dissipation of energy in, and the consequent decay of, oscillations of all types or the extent of the dissipation and decay. The energy losses arise from frictional (or analogous) forces which are unavoidable in any system or from the radiation of energy to space or to other systems. For sufficiently small oscillations, the analogous forces are proportional to the velocity of the vibrating member and oppositely directed thereto; the ratio of force to velocity is $-R$, the mechanical resistance. For the role of damping in the case of forced oscillations, where it is decisive for the frequency response, *see* FORCED OSCILLATION; HARMONIC MOTION; MECHANICAL VIBRATION; OSCILLATION; RESONANCE (ACOUSTICS AND MECHANICS).

Damped oscillations. An undamped system of mass m and stiffness s oscillates at an angular frequency $\omega_0 = (s/m)^{1/2}$. The effect of a mechanical resistance R is twofold: It produces a change in the frequency of oscillations and it causes the oscillations to decay with time. If u is one of the oscillating quantities (displacement, velocity, acceleration) of amplitude A, then Eq. (1) holds in the damped case, whereas in the undamped case Eq. (2) holds. The reciprocal

$$u = Ae^{-\alpha t} \cos \omega_d t \tag{1}$$

$$u = A \cos \omega_0 t \tag{2}$$

time α in Eq. (1) may be called the damping constant.

In Eqs. (1) and (2), the origin for the time t is chosen so that $t = 0$ when $u = A$. The damped angular frequency ω_d in Eq. (1) is always less than ω_0; its value will be given later. According to Eq. (1), the amplitude of the oscillation decays exponentially; the time, given in Eq. (3), is that required for the amplitude to

$$1/\alpha = 2m/R \tag{3}$$

decrease to the fraction $1/e$ its initial value.

A common measure of the damping is the logarithmic decrement δ, defined as the natural logarithm of the ratio of two successive maxima of the decaying sinusoid. If T is the period of the oscillation, then Eq. (4) holds, so that Eq. (1) becomes Eq. (5). Thus

$$\delta = \alpha T \tag{4}$$

$$u = Ae^{-\delta t/T} \cos \omega_d t \tag{5}$$

$1/\delta$ is the number of cycles required for the amplitude to decrease by the factor $1/e$ in the same way that $1/\alpha$ is the time required.

The Q of a system is a measure of damping usually defined from energy considerations. In the present case, the stored energy is partly kinetic and partly potential; when the displacement is a maximum, the velocity is zero and the stored energy is wholly potential, while at zero displacement, the energy is wholly kinetic. The Q is π times the ratio of peak energy stored to energy dissipated per cycle. In the present example, this reduces to Eq. (6). The damped fre-

$$Q = \omega_0 m/R = \pi/\delta \tag{6}$$

quency ω_d (1) is related to the undamped frequency ω_0 of Eq. (2) by Eq. (7), so that for high-Q (lightly

$$(\omega_d/\omega_0)^2 = 1 - (1/2Q^2) \tag{7}$$

damped) systems, it is only slightly less than ω_0. *See* ENERGY.

Overdamping; critical damping. If α in Eq. (1) exceeds ω_0, then the system is not oscillatory and is said to be overdamped. If the mass is displaced, it returns to its equilibrium position without overshoot, and the return is slower as the ratio α/ω_0 increases. If $\alpha = \omega_0$ (that is, $Q = {}^1/_2$), the oscillator is critically damped. In this case, the motion is again nonoscillatory, but the return to equilibrium is faster than for any overdamped case.

Distributed systems. An undamped, one-dimensional wave of frequency $\omega/2\pi$ propagated in the positive direction of x is represented by Eq. (8), c

$$u = A \cos \omega(t - x/c) \tag{8}$$

being the velocity of the wave. If the vibration is maintained at $x = 0$ at the value $u = A \cos \omega t$, then the damping manifests itself as an exponential decrease of amplitude with distance x. Equation (8) is replaced by Eq. (9). The attenuation α' may depend

$$u = Ae^{-\alpha' x} \cos [\omega(t - x/c)] \tag{9}$$

on frequency. If the medium is terminated, the wave will be reflected from the ends, and a system of standing waves will be set up. Examples are a rod carrying sound waves, a piece of electrical transmission line or waveguide, and a vibrating violin string. Such a system has a number of natural frequencies $\omega_n/2\pi$, at each of which it behaves like the lumped system of the previous sections. The decay of a vibration is characterized by Q in Eq. (10), where $\lambda_n = 2\pi c/\omega_n$ is the wavelength.

$$Q = \omega_n/2\alpha' c = \pi/\alpha' \lambda_n \tag{10}$$

Hysteresis damping. At a given instant, the elongation (strain) of a metal bar which is under periodic, alternating stress is not determined exactly by the instantaneous value of the stress existing at that time. For example, the elongation is less at a given stress value when the stress is increasing than when it is decreasing. This phenomenon, which is known as

mechanical hysteresis, causes an undesirable energy loss. A vibration problem of serious nature exists in the blades of jet engines and other steam and gas turbines. The blade material itself exhibits a mechanical hysteresis damping which holds the vibrations in check. When the stress is small, the hysteresis damping is very small in all metals, but it rises suddenly when the stress reaches a certain value. Unfortunately, in most metals the stress at which hysteresis damping becomes large and that at which the metal fails because of fatigue are very close together. However, a much higher hysteresis damping at safe stresses than that of ordinary steel is exhibited by certain alloys. *See* HYSTERESIS; MECHANICAL VIBRATION; STRESS AND STRAIN.

Oscillating electrical circuits. A simple series electrical circuit consisting of an inductance L, resistance R, and capacitance C is exactly analogous to the mechanical system described by Eqs. (1) through (6). The inductance, resistance, and elastance ($1/C$) correspond to the mass, mechanical resistance, and stiffness, respectively. A distributed electrical circuit, such as a section of transmission line or wave guide, is analogous to a vibrating rod or disk. *See* CIRCUIT (ELECTRICITY).

In the ordinary electrical oscillator, the frequency is controlled by an electrical resonator (tank) lumped at the lower, and distributed at the higher, frequencies. Good frequency stability is associated with a high-Q tank. For frequencies of tens of megahertz (MHz) and below, mechanical resonators can be constructed which have a much higher Q than the equivalent electrical tanks. Thus, very stable electrical oscillators have mechanical resonators as their frequency-determining elements. Such an electromechanical system can operate only if there is some coupling between the electrical and mechanical aspects of the system. *See* OSCILLATOR.

The coupling can be arranged in various ways. In some materials, such as quartz, the constitutive relations involve the mechanical and electrical variables jointly; thus, for example, an electric field may produce a strain in the absence of any stress. Thus, the coupling is inherent in the quartz itself. Quartz crystals are used for frequency control of oscillators in the range from kilohertz to perhaps 100 MHz; the Q of a high-frequency crystal may be several million. Some low-frequency oscillators are controlled by tuning forks having a Q of several hundred thousand, the action being similar to that of the electric bell or buzzer. *See* PIEZOELECTRICITY.

The electrostatic motor-generator effect provides the coupling in such mixed systems as the condenser microphone and electrostatic loudspeaker, and the electromagnetic motor-generator effect plays the same role in the dynamic microphone, dynamic loudspeaker, and in various electrical instruments. *See* ALTERNATING-CURRENT CIRCUIT THEORY; ELECTRIC TRANSIENT; ELECTROSTATICS; LOUDSPEAKER; MICROPHONE; RESONANCE (ALTERNATING-CURRENT CIRCUITS); VIBRATION DAMPING. Martin Greenspan

Bibliography. A. D. Helfrick and W. D. Cooper, *Modern Electronic Instrumentation and Measurement Techniques*, 1990; S. G. Kelly, *Fundamentals of Mechanical Vibrations*, 2d ed., 2000; L. E. Kinsler et al., *Fundamentals of Acoustics*, 4th ed., 2000; C. Kittel, W. D. Knight, and M. A. Ruderman, *Mechanics*, Berkeley Physics Course, vol. 1, 2d ed., 1973; H. J. Pain, *The Physics of Vibrations and Waves*, 5th ed., 1999; W. T. Thompson and M. D. Dahleh, *Theory of Vibration with Applications*, 5th ed., 1997.

Daphniphyllales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Hamamelidae of the class Magnoliopsida (dicotyledons). The order consists of a single family with but one genus, *Daphniphyllum*, containing about 35 species. All are dioecious trees or shrubs native to eastern Asia and the Malay region. The plants produce a unique type of alkaloid (daphniphylline group); they often accumulate aluminum and sometimes produce iridoid compounds. The wood has vessels with scalariform perforations. The leaves are simple and entire, alternate or sometimes closely clustered at the ends of the branches. The flowers are small and inconspicuous, unisexual, regular, and hypogynous. Usually there are 2–6 sepals, or sometimes the sepals are absent; petals are lacking. The pollen is triaperturate, and the pistil has 2(–4) united carpels. *Daphniphyllum* has sometimes been included in the Euphorbiales of the subclass Rosidae, but structural details such as the very tiny embryo make it highly aberrant there. *See* HAMAMELIDAE; MAGNOLIOPSIDA; PLANT KINGDOM. T. M. Barkley

Dark current

An ambiguous term used in connection with both gaseous-discharge devices and photoelectric cells or tubes. In gaseous-conduction tubes it refers to the region of operation known as the Townsend discharge. The name is derived from the fact that photons produced in the gas do not play an important part in the production of ionization. The initial ionization arises from independent effects such as cosmic rays, radioactivity, or thermionic emission. When applied to photoelectric devices, the term applies to background current. This is current which may be present as the result of thermionic emission or other effects when there is no light incident on the photosensitive cathode. *See* ELECTRICAL CONDUCTION IN GASES; TOWNSEND DISCHARGE. Glenn H. Miller

Dark energy

An entity that comprises the majority of the energy of the universe and is responsible for the accelerating cosmic expansion. Its name derives from the inference that it is nonluminous. The nature of dark energy is speculative. Leading theories propose that dark energy is a static, cosmological constant

consisting of quantum zero-point energy, and a dynamical condensate of a new, low-mass particle. Alternatively, it has been proposed that the dark energy phenomena are due to a change in the form of gravitation on cosmological length scales. Confirmation of any one of these ideas would have a profound impact on physics. Dark energy appears to be distinct from dark matter, nonluminous particles which make up the majority of the mass of galaxies and galaxy clusters. Determining the nature of dark energy is widely regarded as one of the most important problems in physics and astronomy. *See* ACCELERATING UNIVERSE; COSMOLOGICAL CONSTANT; DARK MATTER.

Observational evidence. The empirical evidence for dark energy is indirect. Yet the evidence is manifest in what is, perhaps, the single most remarkable feature of the universe—the cosmic expansion. According to Albert Einstein's general theory of relativity, a cornerstone of the standard big bang cosmology, the rate of cosmic expansion is determined by the average energy density of all forms of matter and radiation in the universe. Measurements of the expansion, discovered by Edwin P. Hubble in 1929 and refined by the National Aeronautics and Space Administration's Hubble Space Telescope Key Project, therefore determine a critical energy density, $\rho_{\text{crit}} = 0.88\ (\pm 0.19) \times 10^{-9}\ \text{J/m}^3$ (68% confidence level), to which all forms of matter and radiation must sum. For comparison, the energy density of water is nearly 30 orders of magnitude larger. *See* BIG BANG THEORY; HUBBLE CONSTANT; RELATIVITY.

The first piece of evidence for the existence of dark energy arises from measurements of the gravitational mass density of galaxies and galaxy clusters. Diverse measurements accumulated over many decades, from Fritz Zwicky's 1933 observations to recent galaxy-redshift surveys, indicate that the fractional energy density in the form of atoms and even dark matter falls far short of the critical density, with $\Omega_m \equiv \rho_m/\rho_{\text{crit}} \approx 0.3$. The fraction in baryons, referring to all atoms or nuclear material, is a mere $\Omega_b \approx 0.04$, implying that the nature of ~96% of the energy of the universe (26% dark matter and 70% dark energy) is unknown.

The second piece of evidence derives from measurements of the mean curvature of space. Just as mass and energy distort space-time, a mean curvature of space itself contributes an equivalent energy density. A space that is sufficiently curved so as to explain the dark energy must thereby possess a geometry in which initially parallel light rays traversing cosmological distances are distorted and eventually diverge. But detailed measurements of the cosmic microwave background radiation, most recently by NASA's *Wilkinson Microwave Anisotropy Probe*, have eliminated this possibility in confirming that the mean curvature of space is negligible. The ~70% gap in energy must be due to something other than the curvature or dark and baryonic matter. *See* COSMIC BACKGROUND RADIATION; WILKINSON MICROWAVE ANISOTROPY PROBE.

The third piece of evidence, which provides the strongest clue to the nature of the dark energy, derives from the 1998 discovery of the accelerating cosmic expansion. Cosmological distance-redshift measurements of the scale of the universe as a function of look-back time reveal that the rate of expansion has been increasing during the last several billion years. In general relativity, the acceleration of a test particle due to the gravitational field of a fluid sphere is proportional to $\rho + 3p$ of the fluid, where ρ is the fluid density and p is the pressure. In the cosmic context, acceleration implies that the universe is under tension, like a fluid with a negative effective pressure. (An example is the stress parallel to a uniform electric or magnetic field.) Because no other cosmologically abundant forms of energy support a sufficiently negative pressure across cosmic scales, the dark energy must carry the cosmic pressure responsible for the acceleration. The strength of the pressure is characterized by an equation-of-state parameter, w, defined as the ratio of the homogeneous pressure to the energy density of the dark energy, $w \equiv p/\rho$. Formally, acceleration requires $w < -1/3(1 - \Omega_m)$, with current measurements giving $-1.2 \leq w \leq -0.8$ (95% confidence level).

These are the main lines of evidence for dark energy. Numerous other observations and measurements—including the age of the universe, the variation of distances with redshift, and distortions of the cosmic microwave background radiation by the accelerated expansion—are consistent with the interpretation that some form of dark energy with negative pressure is responsible for the majority of the energy of the cosmos.

Theoretical solutions. The leading hypothesis is that the dark energy is a cosmological constant (Λ), a uniform sea of positive energy and negative pressure with $w = -1$. A detailed comparison between theoretical predictions and observations finds that a cosmological constant supplying ~70% of the critical energy density, or $\Omega_m \approx 0.3$, fits current data. In the absence of further clues, a cosmological constant provides the most economical description of dark energy.

The cosmological constant was introduced in 1917 by Einstein as a mathematical device in his ground-breaking application of general relativity to the problem of cosmology. A physical explanation of its origin was offered in 1967 by Yakov Zeldovich, who demonstrated that the vacuum of quantum-mechanical particles is equivalent to a cosmological constant. However, the predicted amplitude is at least 60 orders of magnitude too large—a mismatch regarded as one of the most enigmatic problems in theoretical physics. It is speculated that a complete understanding of the cosmological constant requires a theory of quantum gravity. *See* QUANTUM GRAVITATION.

The need to adjust or finely tune the cosmological constant to a value so many orders of magnitude below the predicted value has prompted speculation that perhaps $\Lambda = 0$ or that it does not contribute to the gravitational field equations, and some other mechanism is responsible for the accelerated expansion.

A widely studied possibility is that the dark energy is dynamical, consisting of a condensate of a new species of particles of very small mass. In the simplest version of this theory, these particles have mass $m \approx 10^{-68}$ kg and are identified as excitations of a cosmic field which oscillates with wavelength $\lambda \approx 10^{23}$ km, just within the size of the observable universe. This form of dark energy is also referred to as quintessence, as distinct from other fields and forms of matter or radiation. In contrast to a cosmological constant, the equation of state of quintessence is time varying, with $w > -1$, and its energy density and pressure fluctuate in response to gravitational field variations spanning length scales larger than λ. Through measurements of such time variation or fluctuations it is hoped to distinguish among competing theories of dark energy.

The theory of quintessence dark energy has two significant, although unproven, precedents. First, the dark energy phenomenon is very similar to primordial inflation, an epoch of accelerated expansion conjectured to have taken place in the early universe. According to inflation theory, the mechanism that drives the expansion is a cosmic field identical to quintessence except scaled up in mass and energy to match the conditions in the early universe. Second, the quintessence field is very similar to the axion particle, a dark matter candidate particle, originally proposed to resolve a fine-tuning problem in quantum chromodynamics. Despite these motivations for quintessence dark energy, any such theory also faces the challenge that it must explain why the new particle does not readily interact with known particles and mediate a new force, and why the new particles are so light compared to known particles. *See* INFLATIONARY UNIVERSE COSMOLOGY.

More speculative theories of dynamical dark energy with equation of state $w < -1$ have been proposed. One such example, called phantom dark energy for a similarity to so-called ghost fields in quantum field theory, is also a cosmic field but with very different properties from quintessence. The energy density of the phantom field grows with time, rather than decays as for quintessence. The viability of this theory has implications for quantum gravitational processes.

Numerous other theories of dynamical dark energy have been proposed. Most can be characterized simply by an equation of state of the quintessence ($w > -1$) or phantom ($w < -1$) type. In all cases, the magnitude of the cosmic acceleration is determined by the combination $-[1 + 3w(1 - \Omega_m)]/2$ so that, for equal matter density, quintessence gives the weakest acceleration and a phantom gives the strongest. Important clues to the nature of dark energy can be gained from measurements of the equation of state and any possible time variation.

Gravity. It has been proposed that the dark energy phenomena are due to a change in the form of gravitation on cosmological scales, as alternative to a cosmological constant or quintessence. Einstein's general relativity is well-tested within the solar system, and underpins the successes of the big bang cosmology. However, a modification of the gravitational field equations required to balance the expansion against the energy density and pressure of the cosmological fluid could obviate the need for the above-described theories of dark energy.

New theories of gravitation which introduce extra gravitational degrees of freedom under the rubric of scalar-tensor gravity, or which incorporate the effects of additional dimensions of space-time based on string theory or other fundamental theories of physics, are a subject of intense study. Most such theories proposed to explain the dark energy phenomena include a new cosmic field which is similar to quintessence but also modulates the strength of gravitation, thereby resulting in variability of Newton's constant, the amount of space-time curvature produced per unit mass or energy, and the degree to which the curvature produced by a massive body exceeds the sum of the curvatures due to its individual constituents. Most simple extensions of general relativity are heavily constrained by solar system observations in such a way that viable cosmological models are essentially indistinguishable from general relativity with a cosmological constant or quintessence dark energy. *See* SUPERSTRING THEORY.

Fate of the universe. The discovery of the cosmic acceleration suggests that the future evolution and fate of the universe will be determined by the nature of the dark energy. Previously, the main speculation was whether the big bang would terminate in a big crunch, as the universe collapsed under its own weight. Now a big chill appears as a likely fate: In a universe containing a cosmological constant or quintessence dark energy which maintains the inequality $-1 \leq w < -1/3(1 - \Omega_m)$, the expansion will continue to accelerate forever. Only gravitationally bound structures today will survive into the distant future. Furthermore, the acceleration will prevent any communication across distances greater than $\sim 10^{23}$ km due to the appearance of a cosmological event horizon, and the universe will grow cold and rarefied. But in a universe containing phantom dark energy which maintains the inequality $w < -1$, the future would appear to be a big rip: the catastrophic end to the universe. In this case, the expansion would undergo a runaway process resulting in the gravitational disruption of clusters, galaxies, and then both gravitationally and nongravitationally bound objects on smaller and smaller scales. The Milky Way would be torn apart $\sim$50 million years before the big rip. The universe would be shredded to its fundamental constituents, unless quantum gravity effects intervene, and space-time itself would end $\sim$50 billion years in the future.

Prospective observations and experiments. The current challenge to cosmology is to tighten the evidence for dark energy, refine measurements of the cosmic expansion so as to narrow the bounds on the equation of state w, and continue to test the framework of general relativity and big bang cosmology. The main impact of the dark energy is on the rate of cosmic expansion, so that new clues to

the physics of dark energy may be obtained from measurements of distance, volume, and age versus redshift; the evolution and abundance of galaxies and clusters; and the distortion of the cosmic microwave background radiation. Numerous experiments and observatories dedicated to investigating dark energy are planned. Robert Caldwell

Bibliography. R. R. Caldwell, Dark energy, *Phys. World*, 17(5):37–42, 2004; W. L. Freedman and M. S. Turner, Measuring and understanding the universe, *Rev. Mod. Phys.*, 75:1433–1447, 2003; J. P. Ostriker and P. J. Steinhardt, The quintessential universe, *Sci. Amer.*, 284(1):46–53, 2001; P. J. E. Peebles and B. Ratra, The cosmological constant and dark energy, *Rev. Mod. Phys.*, 75:559–606, 2003.

Dark matter

Particles or objects that exert a gravitational force but do not emit any detectable light. Dark matter is the dominant form of matter in our Galaxy. Astronomers have detected the presence of dark matter through its gravitational effects and have shown that dark matter is not composed of ordinary atoms. Particle physicists have suggested several plausible candidates for dark matter; planned experiments are capable of detecting these new particles.

Astronomical evidence for dark matter. A variety of astronomical observations imply that dark matter is ubiquitous; it is detected in dwarf galaxies, in spiral galaxies, and in elliptical galaxies. It is the dominant form of matter in galaxy clusters and leaves clear signatures in the large-scale distribution of galaxies and in the microwave background. Astronomers infer the presence of matter through its gravitational effects. Since they have not been able to detect any light directly associated with this matter, they have labeled it "dark matter."

Dark matter is also sometimes called the "missing matter." This is a misnomer since astronomers detect the mass but they are unable to detect the light associated with the matter.

Galaxy motions in clusters. Many galaxies are bound together in galaxy clusters. These galaxy clusters contain hundreds and sometimes thousands of galaxies. These galaxies are moving with very high relative velocities. Typical velocities are roughly 1000 km/s (600 mi/s), and some galaxies that appear to be bound to the cluster move at velocities as high as 2000 km/s (1200 mi/s). Within an individual galaxy, stars and gas move much slower with typical velocities of only 200 km/s (120 mi/s).

As early as the 1930s, F. Zwicky recognized that clusters must be very massive to have a strong enough gravitational field to keep these high-velocity galaxies bound to the cluster. Zwicky noted that the inferred mass was much larger than the mass in stars. The result was so surprising that most of his colleagues ignored his conclusion. Current estimates suggest that stars comprise only 1% of the mass of the cluster (**Fig. 1**). *See* GALAXY, EXTERNAL.

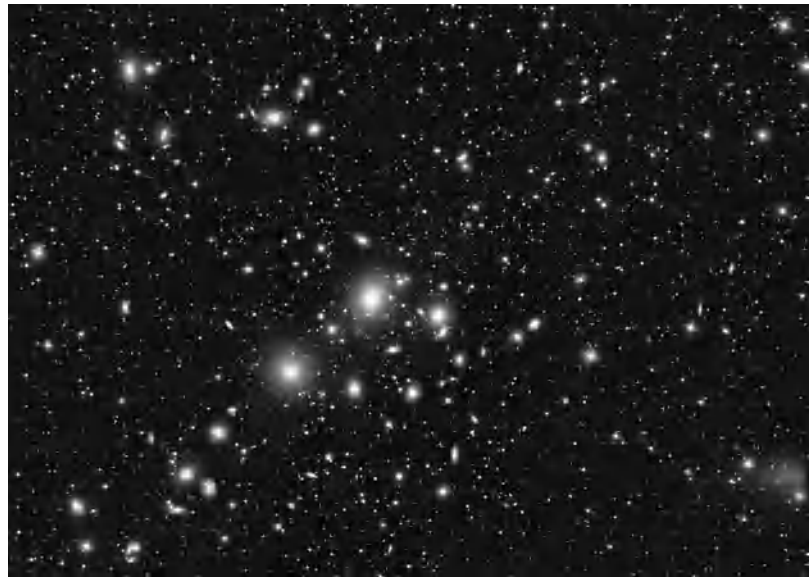

Fig. 1. Sloan Digital Sky Survey image of the Perseus galaxy cluster. Galaxies in the cluster are moving relative to each other at velocities of thousands of kilometers per second. Since the cluster appears to be gravitationally bound, this implies that the cluster has a mass of 10^{11} solar masses, roughly 100 times the mass contained in stars. (*Robert Lupton; SDSS Collaboration*)

Galaxy rotation curves. Most galaxies, including our own Milky Way, are spiral or disk galaxies, where cold atomic and molecular gas (primarily hydrogen) settles into a rotating disk. Much of the molecular gas is found in spiral arms. These spiral arms are active sites of star formation. *See* MILKY WAY GALAXY.

The gas and stars in disk galaxies move on nearly circular orbits. The Sun is moving on a nearly circular orbit around the center of the Milky Way Galaxy. To keep material moving on a circular orbit, centrifugal acceleration must balance gravitational acceleration. This balance can be expressed by the equation below, where V is the velocity of the gas or stars,

$$\frac{V^2}{R} = \frac{GM}{R^2}$$

R is the distance from the center of the galaxy, G is Newton's constant, and $M(R)$ is the mass contained within radius R. *See* CENTRIPETAL FORCE; CENTRIFUGAL FORCE.

Astronomers can measure the velocity of gas and stars in other galaxies through the Doppler effect. In the 1960s and 1970s, Vera Rubin pioneered optical observations of the motions of stars near the visual edges of galaxies. She found that the stars were moving rapidly. This large velocity implied the presence of hundreds of millions of solar masses of matter where little light was detected. Radio observations of the motions of neutral gas beyond the visual edge of spiral galaxies confirmed this remarkable result: Most of the mass in a galaxy is not in its disk of young stars, but rather in a halo composed of dark matter (**Fig. 2**). *See* DOPPLER EFFECT.

Hot gas in elliptical galaxies and clusters. X-ray satellites have detected copious amounts of hot gas in elliptical galaxies and in galaxy clusters. Since the thermal

Fig. 2. Sloan Digital Sky Survey image of galaxy M101. Measurements of the velocities in the disk of the galaxy imply the presence of enormous amounts of mass out toward the edge of the disk. (*Robert Lupton; SDSS Collaboration*)

pressure in this 1–10 million degree gas balances the gravitational field of the galaxy, measurements of the density and temperature profiles of the gas can be used to directly measure the mass distribution in these galaxies. These measurements reveal that the dark matter problem is ubiquitous: Stars can account for only a small fraction of the mass in elliptical galaxies. Observations of x-ray gas in clusters confirm Zwicky's inference that stars account for only a small fraction of the mass in a cluster (**Fig. 3**). *See* X-RAY ASTRONOMY.

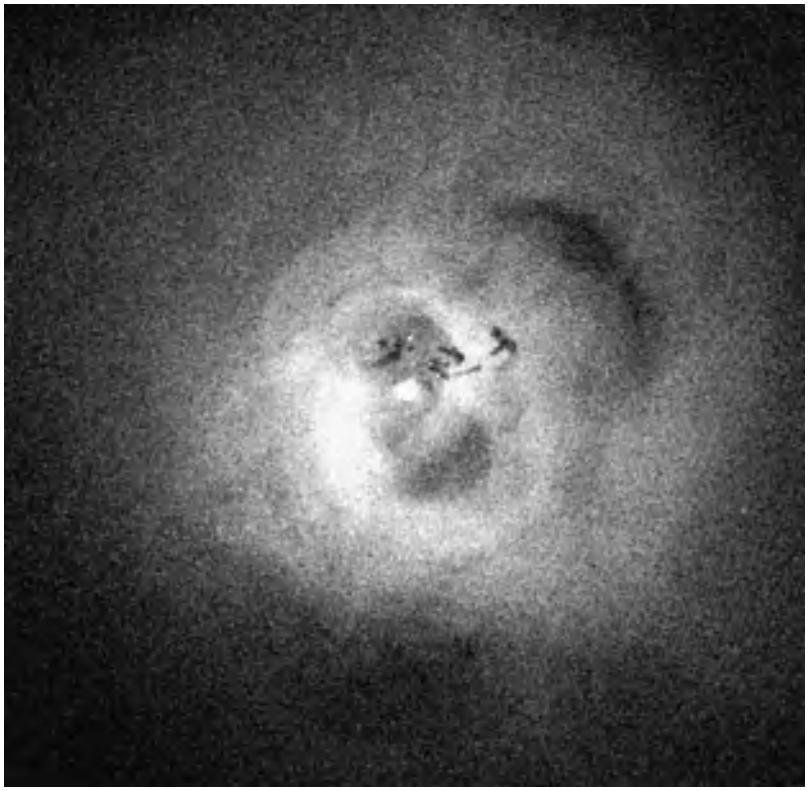

Fig. 3. *Chandra X-ray Observatory* image of hot gas in the Perseus galaxy cluster, indicating that stars account for only a small fraction of the mass of the cluster. (*NASA/Chandra X-ray Center/Institute of Astronomy/A. Fabian et al.*)

Gravitational lensing. John Wheeler has remarked that "General relativity is simple: Matter tells space how to curve; the curvature of space tells matter how to move." Since mass curves space and distorts the path of light, images of galaxies that are behind rich clusters can be used to infer the distribution of mass in clusters. The masses in clusters distort background galaxies into arcs whose orientations are very sensitive to the distribution of matter in the cluster. Observations of gravitational lenses detect the presence of significant amounts of dark matter, and are consistent with the x-ray and optical observations of the same clusters (**Fig. 4**). *See* GRAVITATIONAL LENS; RELATIVITY.

Fig. 4. *Hubble Space Telescope* image of a lensed background source in the form of an extended arc about an elliptical lensing galaxy. By measuring the distortions in the background, the mass of the foreground lensing galaxy can be inferred. (*Kavan Ratnatunga, Carnegie Mellon University; NASA*)

Novel nature of dark matter. The astronomical evidence discussed above shows that most of the mass in galaxies is not in the form of luminous stars and is consistent with a wide range of plausible candidates. Most astronomers assumed that the dark matter was similar to the "ordinary stuff" that makes up stars: protons, neutrons, and electrons. Cosmologists often refer to this "ordinary stuff" as baryonic matter. (Unfortunately, cosmologists use the term "baryonic matter" to refer to protons, neutrons, and electrons, while particle physicists, more correctly, refer to only protons and neutrons as baryonic matter.) Astronomers have searched for various possible forms of ordinary matter, including gas, dust, and low mass stars, but they have been unable to detect it. *See* BARYON.

Cosmological observations can measure the density of atoms in the early universe. These

observations imply that baryonic matter can account for only one-sixth of the mass in galaxies. These observations imply that most of the mass in galaxies is in the form of some novel and not yet identified form of matter.

Gas. Gas can be detected through either emission or absorption. Astronomers have measured the mass in atomic gas through observations of the 21-cm line and molecular gas through millimeter observations of carbon monoxide (CO) clouds. These observations suggest that only 1% of the mass of the Milky Way Galaxy is in cold gas in the galactic disk. *See* INTERSTELLAR MATTER.

X-ray satellites can detect 1–100 million degree gas through its thermal bremsstrahlung emission. These observations reveal that most of the baryonic matter in rich clusters is not in stars but in hot gas. In some rich clusters, there is ten times as much mass in hot gas as in stars. While significant for the evolution of the cluster, only 10–20% of the mass in a cluster is in hot gas. Dark matter accounts for most of the remaining 80–90%. *See* BREMSSTRAHLUNG.

Low-mass stars. While massive stars can be easily detected through their light, very low mass stars and planet-size objects are so dim that they can evade detection by optical telescopes. Astronomers, however, can detect these stars through their gravitational effects.

When a low-mass star passes in front of a distant star, it serves as a gravitational lens. This lensing brightens the image of the background star. Using large-area cameras capable of monitoring millions of stars, astronomers have searched for and detected these effects as low-mass stars pass in front of stars in the Large Magellanic Clouds and in front of the Milky Way bulge. While these observations have detected hundreds of low-mass stars through their gravitational effects, the number of events are consistent with models of the Milky Way Galaxy and imply that there are not enough low-mass stars to account for the dark matter.

Deuterium and helium. During the first minutes of the big bang, all of the neutrons in the universe combined with protons to make deuterium (one proton and one neutron). Two deuterium nuclei then collided to make helium (two protons and two neutrons). Most of the helium and deuterium in the universe was produced during the first few minutes of the big bang, and thus deuterium and helium can be said to be fossils from that time. The cosmological abundance of deuterium depends on the density of baryons (protons and neutrons) in the early universe: the more baryons, the lower relative abundance of deuterium. Astronomers use measurements of the ratio of the relative number of deuterium nuclei to provide a direct estimate of the density of atoms in the universe. *See* BIG BANG THEORY.

These observations imply that ordinary matter makes up only 4% of the energy density of the universe. These observations suggest that dark matter is not composed of protons and neutrons (or of nuclei made of protons and neutrons).

Microwave background fluctuations. Measurements of microwave background fluctuations are an important probe of the composition of the universe. The tiny variations in temperature seen by the *Wilkinson Microwave Anisotropy Probe* (and by ground and balloon-based microwave experiments) were produced during the first moments of the big bang. The statistical properties of these fluctuations depend on the composition of the universe: the more atoms, the smoother the shape of the fluctuations; the more matter, the higher the relative amplitude of small-scale fluctuations. The current observations confirm that ordinary matter makes up only 4% of the energy density of the universe and also imply that dark matter comprises 25% of the energy density of the universe *See* COSMIC BACKGROUND RADIATION; WILKINSON MICROWAVE ANISOTROPY PROBE.

Dark matter candidates. The big bang theory implies that the early universe was very hot and very dense, the perfect environment for creating particles through collisions. The densities and temperatures in the first microsecond of the big bang exceed the energies achieved in even the most powerful particle accelerators. Many cosmologists suspect that the dark matter is composed of some yet undiscovered fundamental particle that was produced in copious numbers during the first moments of the big bang.

Supersymmetry. The most popular dark matter candidate is the neutralino, a new particle posited by the theory of supersymmetry. Supersymmetry is an extension of the standard model of physics and is a vital element in almost all attempts to unify the forces of nature and in attempts to connect gravity and quantum mechanics.

Modern particle physics is based on the symmetries observed in nature. For example, all particles have antiparticles. The positron is the antiparticle of the electron. The antiproton is the antiparticle of proton. While seemingly bizarre, antimatter has not only been produced in the laboratory and detected in space, but now plays a major role in medical imaging (PET scans). *See* ANTIMATTER; MEDICAL IMAGING; SYMMETRY LAWS (PHYSICS).

Particle physicists have speculated on the existence of a new, not yet confirmed symmetry: supersymmetry. Supersymmetry implies that particles such as electrons and positrons have partners called selectrons and spositrons. These superparticles have the opposite statistical properties to ordinary matter. Selectrons behave like photons, while the supersymmetric partner of the photon, the photino, behaves quantum mechanically like an electron. This new symmetry has many aesthetic attractions and helps to explain the relative strength of the fundamental forces of nature. These new particles interact weakly with ordinary matter and have so far escaped detection. (They also may not exist.) *See* QUANTUM STATISTICS; SUPERSYMMETRY.

Supersymmetry implies the existence of a new stable particle: the neutralino. The neutralino would have been produced in abundance during the first moments of the big bang. During the first nanosecond of the big bang, there were as many neutralinos

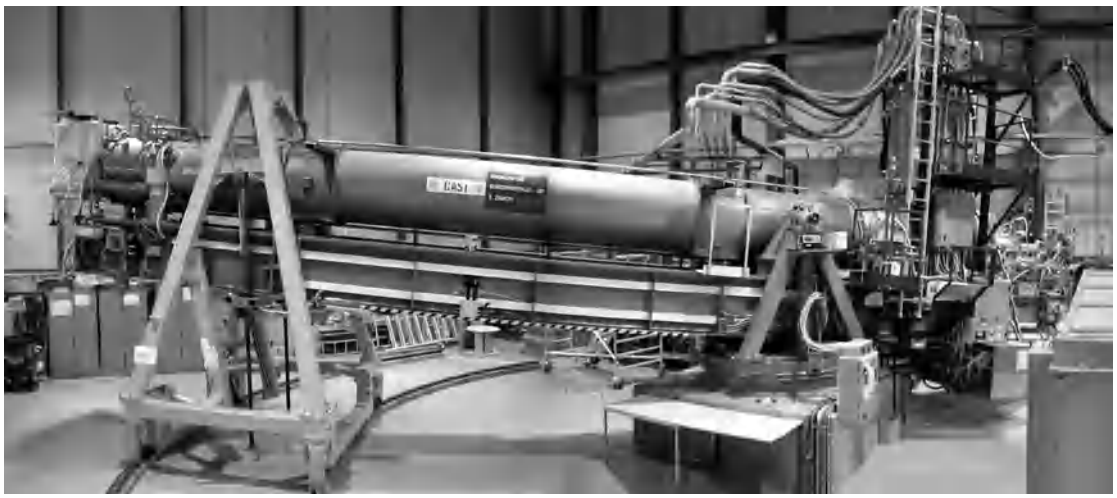

Fig. 5. CERN Axion Solar Telescope (CAST), which uses a 9-tesla, 10-m (33-ft) prototype dipole magnet for CERN's Large Hadron Collider, enhanced by the use of a focusing x-ray telescope, to search for axions from the Sun's core. (*CERN*)

as photons in the universe. While the number of neutralinos was reduced as they annihilated into ordinary matter, the current best estimates of their predicted residual abundance imply that they could be the dark matter. *See* WEAKLY INTERACTING MASSIVE PARTICLE (WIMP).

The primary scientific goal of the new Large Hadron Collider (LHC) at CERN in Switzerland, scheduled for start-up in 2007, is to detect the experimental signatures of supersymmetry. By colliding nuclei at velocities very close to the speed of light, the LHC aims to recreate some of the conditions in the big bang. The LHC may be able to provide direct evidence for the existence of supersymmetry and may even be able to produce dark matter particles in the laboratory. *See* PARTICLE ACCELERATOR.

Deep-underground experiments may also detect dark matter. If the neutralino is the dark matter, then hundreds of millions of these particles are streaming through our bodies every second. Because the neutralino interacts so weakly with ordinary matter, our bodies (and our detectors) are nearly transparent. These particles, however, do have rare weak interactions. In a kilogram detector, the dark matter particles may have a few collisions per day. These collisions are difficult to detect since each collision deposits only 10^{-3} (10^{-15} erg) of energy. Physicists have built a number of very sensitive low-background experiments capable of directly detecting neutralinos. Because the experiments need to be shielded from cosmic rays and other terrestrial sources of background signal, they must operate in deep-underground mines and tunnels.

Gamma-ray satellites may also be able to indirectly detect the neutralino. While the neutralinos interact weakly, occasionally one neutralino collides with another neutralino. This collision produces a shower of particles and radiation. Several on-going and planned experiments are looking for this dark matter annihilation signal. *See* GAMMA-RAY ASTRONOMY.

Axions. The axion is another hypothetical particle, which was invented to explain one of the symmetries seen in the strong nuclear interaction. If this explanation is correct, then axions would have been produced copiously during the first microsecond of the big bang. While axions interact very weakly with ordinary matter, they can be converted into ordinary photons in the presence of a very strong magnetic field. Experimentalists are using very strong fields to search for the axion (**Fig. 5**).

Black holes. If black holes were produced in large numbers during the first moments of the big bang, then they would be plausible candidates for the dark matter. Currently, there are no viable scenarios for producing large numbers of black holes in the early universe. However, since there are significant uncertainties in our understanding of physics during the first microsecond of the big bang, some astronomers consider black holes to be a viable dark matter candidate.

If the dark matter were composed of black holes and the black holes were significantly more massive than the Sun, then the black holes could be detected through their gravitational effects on galaxies and globular clusters, or through gravitational microlensing (discussed above in connection with the detection of low-mass stars). However, low-mass black holes could easily evade any current (and planned) observational searches. *See* BLACK HOLE.

Failure of general relativity. Many astronomers have speculated that the discrepancy between the mass seen in stars and gas and the mass inferred by applying general relativity (and Newtonian gravity, its low-velocity simplification) to astronomical observations is not the signature of some new exotic particle, but instead is the observational signature of the breakdown of general relativity. Classical Newtonian physics works well in daily life: Quantum mechanics is usually important only at very small scales and general relativity is only important on very large scales and near massive objects such as black holes and neutron stars. Perhaps the dark matter problem is the observational signature of the failure of general relativity. The discovery of dark energy

strengthens the motivation for considering these alternative models.

Modified Newtonian dynamics (MOND), developed by Mordehai Milgrom and Jacob Bekenstein, is the most carefully examined alternative gravitational theory. It has the advantage of explaining galaxy rotation curves without resorting to dark matter. However, MOND has difficulties fitting the data implying the existence of dark matter in clusters and is inconsistent with WMAP's observations of microwave background fluctuations. *See* COSMOLOGY; UNIVERSE.

David N. Spergel

Darmstadtium

The eighteenth of the synthetic transuranium elements, symbol Ds, atomic number 110. Darmstadtium should be a heavy homolog of the elements

1												13	14	15	16	17	18
1 **H**	2																2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

platinum, palladium, and nickel. It is the eighth element in the 6*d* shell. The various isotopes of darmstadtium have been reliably documented by the observation of at least four isotopes with half-lives in the range of milliseconds, a time too short for the application of chemical methods, verifying its position in the periodic system. *See* HALF-LIFE; NICKEL; PALLADIUM; PLATINUM; RADIOISOTOPE.

Searches for this element began in 1985. With the finding of the isotope ^{269}Ds in late 1994, its discovery became conclusive. Early experiments performed at FLNR (Flerov Laboratory for Nuclear Reactions), Dubna, Russia, GSI (Gesellschaft für Schwerionenforschung), Darmstadt, Germany, and LBNL (Lawrence Berkeley National Laboratory), Berkeley, California, failed to provide reliable evidence for a successful synthesis. In irradiations of uranium-235 and uranium-236 with beams of argon-40 at FLNR, a 10-ms spontaneous fission activity was observed in 1986 and assigned to the isotope ^{272}Ds. The method used for separation was a rotating wheel collecting the recoils combined with spontaneous fission detection. A production cross section of about 10^{-35} cm^2 was deduced. An irradiation of ^{235}U with ^{40}Ar at GSI in 1986 using the recoil separator technique reached the cross-section limit given by the FLNR experiment, but did not reveal any indication of darmstadtium formation. In 1991 at LBL bismuth-209 was irradiated by cobalt-59. A single sequence of signals observed after separation in a gas-filled separator was assigned to ^{267}Ds. But the correlation to known daughter isotopes of elements seaborgium to mendelevium could not be reliably established. Unsuccessful experiments to synthesize darmstadtium during 1984–1991 showed that production cross sections below 10^{-35} cm^2 should be made readily accessible by the equipment to be developed for future synthesis reactions. The technique was successfully applied at GSI for the synthesis of elements bohrium, hassium, and meitnerium. Local adaptations and improvements made this the standard technique in the 1990s. *See* ARGON; BOHRIUM; COBALT; HASSIUM; LEAD; MEITNERIUM; MENDELEVIUM; NUCLEAR REACTION; SEABORGIUM; URANIUM.

Separating heavy-ion products. Improved beam intensity and quality, improvement of the detection efficiency, and a new detector system allowing nearly complete chain reconstruction made the discovery of darmstadtium possible (**Fig. 1**). The total sensitivity for finding a new species was increased by a factor of 20. The efficiency to transport, select, and detect an atom produced as a fusion evaporation residue is 40%. The fusion evaporation residues observed in reactions of heavy-ion beams in thin targets (about 1 mg/cm^2) are generated with a momentum in direction of the beam of projectiles inducing the fusion reaction. Their recoil is well defined by momentum conservation and is large enough to allow the fusion evaporation residue to leave the target as a multiply charged heavy ion flying in the direction of the projectile beam with a velocity smaller than the velocities of any other reaction product. *See* ION.

At GSI the recoils are separated from the projectiles by a combination of electric and magnetic fields acting as a filter for a selected velocity. Two identical separation stages, each reducing the projectile beam by a factor of 10^5–10^6, isolate the fusion evaporation residue. About 3×10^{18} projectiles enter the velocity filter in a 10-day experiment, the time to detect one atom of an element produced with the cross section of 10^{-36} cm^2. The remaining 3×10^7 ions together with the hidden fusion evaporation residue leaving the filter are analyzed by a detector system. Once more the velocity of all ions is measured by a time-of-flight detector before they are implanted into a position-sensitive silicon detector of high energy resolution. The measurement of the deposited energy and the time of flight indicates the mass value for each of the ions. The fusion evaporation residue has the highest mass of all possible reaction products, and its implantation energy is defined by momentum conversation. Mass and energy of candidates for a fusion evaporation residue observed in their expected ranges are registered with the arrival time and the position. *See* ELECTRIC FIELD; ENERGY; ELECTROMAGNETIC FIELD; MASS.

The nuclei of the heaviest elements are radioactive. The main decay modes are alpha decay and spontaneous fission. Silicon detectors are best suited for detecting these decay modes. An alpha decay

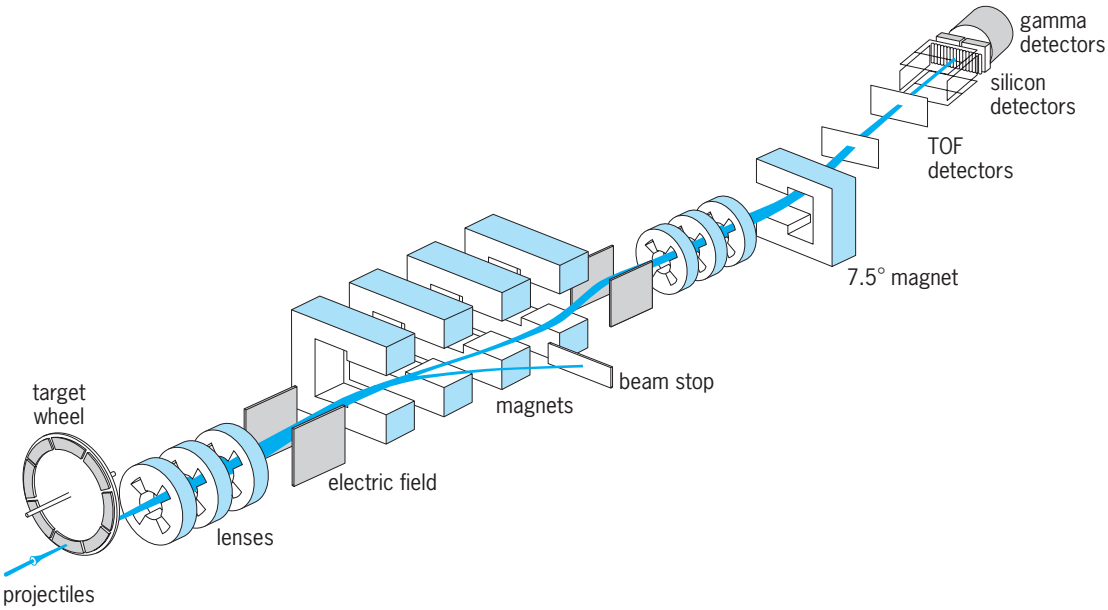

Fig. 1. A velocity filter is used as a separator for heavy-ion reaction products (SHIP). Fusion products travel at known velocities and are forced to follow the beam direction. Products are separated and identified by silicon detectors. TOF = time of flight.

does not change the position of the emitter in the bulk of the detector material within the accuracy of the position measurement. Correlated decays from one parent nucleus must have the same position coordinates. Spontaneous fission breaks the nucleus and ends its decay chain. As all correlation methods are restricted by the occurrence of random sequences, which are dependent on the rates in the windows defined for the different generations in the decay chain, the identification of an atom by its correlated decay chain becomes increasingly uncertain for long correlation times. Dividing the total measuring cycle into a time of irradiation with higher background rates and a measuring time between irradiations helps to reduce random correlations. Isotopes in the half-life range of 10^{-5} s to 1 min could be analyzed in GSI experiments in *x*-chains of up to six generations. For shorter decay chains, the maximum possible correlation times are in the few-seconds range.

The combination of recoil separation and single-decay-chain analysis allows detection of a new element from a single event observed. The method is fast (microseconds), efficient (40%), and sensitive (10^{-36} cm^2).

Decay chain. After an extensive search for the optimum bombarding energy on November 9, 1994, a decay chain was observed which proved the existence of the isotope ^{269}Ds. The isotope was produced in a fusion reaction of a nickel-62 projectile with a lead-208 target nucleus. The fused system, with an excitation energy of 13 MeV, cooled down by emitting one neutron and forming ^{269}Ds, which by sequential alpha decays transformed to hassium-265, seaborgium-261, rutherfordium-257, and nobelium-253. All of these daughter isotopes were already known, and four decay chains observed in the following 12 days corroborated without any doubt the discovery of the element. **Figure 2** shows the first decay chain observed, which ended in ^{257}Rf. The isotope ^{269}Ds has a half-life of 0.2 ms, and is produced with a cross section of about 3×10^{-36} cm^2. *See* ALPHA PARTICLES; NEUTRON; NUCLEAR STRUCTURE; RUTHERFORDIUM.

A second isotope, ^{271}Ds, was produced in a subsequent 12-day experiment by fusion of nickel-64 and lead-208. Nine atoms, with an excitation energy of 12 MeV, were produced, and by sequential alpha decay they transformed to the known isotopes

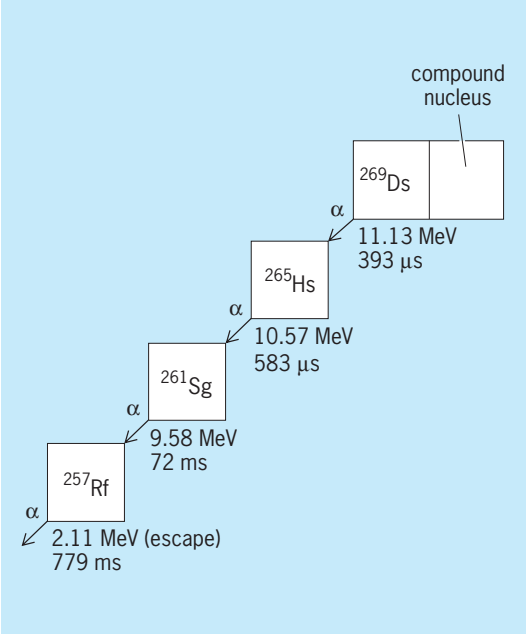

Fig. 2. Sequence of decay chains that document the discovery of darmstadtium. Numbers below boxes are alpha energies and correlation times. Darmstadtium is produced in the reaction ^{62}Ni + ^{208}Pb → ^{269}Ds + 1*n*.

hassium-267, seaborgium-263, rutherfordium-259, and nobelium-255. The half-life of ^{271}Ds is 1.1 ms, and its production cross section amounts to 1.5×10^{-35} cm^2. *See* ATOM.

In the decay chain of ^{277}112, the isotope ^{273}Ds was discovered in 1996. Its half-life is 0.1 ms, and the production cross section is 5×10^{-37} cm^2. The isotope ^{270}Ds with a half-life of 0.1 ms was synthesized at GSI-Darmstadt in the fusion of ^{207}Pb and ^{64}Ni. *See* ELEMENT 112. Peter Armbruster

Bibliography. S. Hofmann et al., Production and decay of ^{269}110, *Z. Phys. A*, 350:277–280, 1995.

Dasheen

Common name for the plant *Colocasia esculenta*, including the variety *antiquorum* (taro). These plants are among the few edible members of the aroid family (Araceae). Native to southeastern Asia and Malaysia, the plants supply the people with their most important food. The edible corms (underground stems) support a cluster of large leaves 4–6 ft (1.2–1.8 m) long. A main dish in the Polynesian menu is poi, a thin, pasty gruel of taro starch, often fermented, which is frequently formed into cakes for baking or toasting. The raw corms are baked or boiled to eliminate an irritating substance present in the cells. *See* ARALES. Perry D. Strausbaugh; Earl L. Core

Dasycladales

An order of green algae (Chlorophyceae) in which the plant body (thallus) is composed of a nonseptate axis, attached by rhizoids and bearing whorls of branches. The walls utilize a partially crystalline mannan (polymer of mannose) as the skeletal component and usually are impregnated with aragonite, a form of calcium carbonate, causing these algae to be easily fossilized. Among green algae, only the Dasycladales have a fossil record sufficient to permit meaningful phylogenetic speculation. About 140 extinct genera have been described (**Fig. 1**) from limestones as old as the Ordovician, with three peaks of abundance and diversity—Carboniferous, Jurassic-Cretaceous, and Eocene. Only about 11 genera comprising 50 species are extant, with 3 of the genera extending from the Cretaceous. All are confined to warm marine waters except *Batophora*, which occurs as a relict in brackish sinkholes in New Mexico in addition to having a normal Caribbean distribution.

The life history is known for only a few genera, but the one that has been most thoroughly investigated—in *Acetabularia*—is unique. The stalk (**Fig. 2**), which may be as long as a few centimeters, bears a terminal whorl of sterile branches, which is superseded in the second growing season by another sterile whorl and in the third season by a fertile whorl. Prior to reproductive maturity, the thallus contains a single nucleus, which remains in a rhizoid and increases in diameter from about 2 to 150 micrometers. Increase in size is accompanied by development of a large nucleolar mass, which may be sausage-shaped or distributed among as many as 24 units. During the formation of the fertile whorl (whose component branches often cohere laterally to form a disk or cap), the nucleus undergoes profound but poorly understood changes, including reduction in size of the nucleus itself and in the size and number of nucleolar units. The residual nucleus

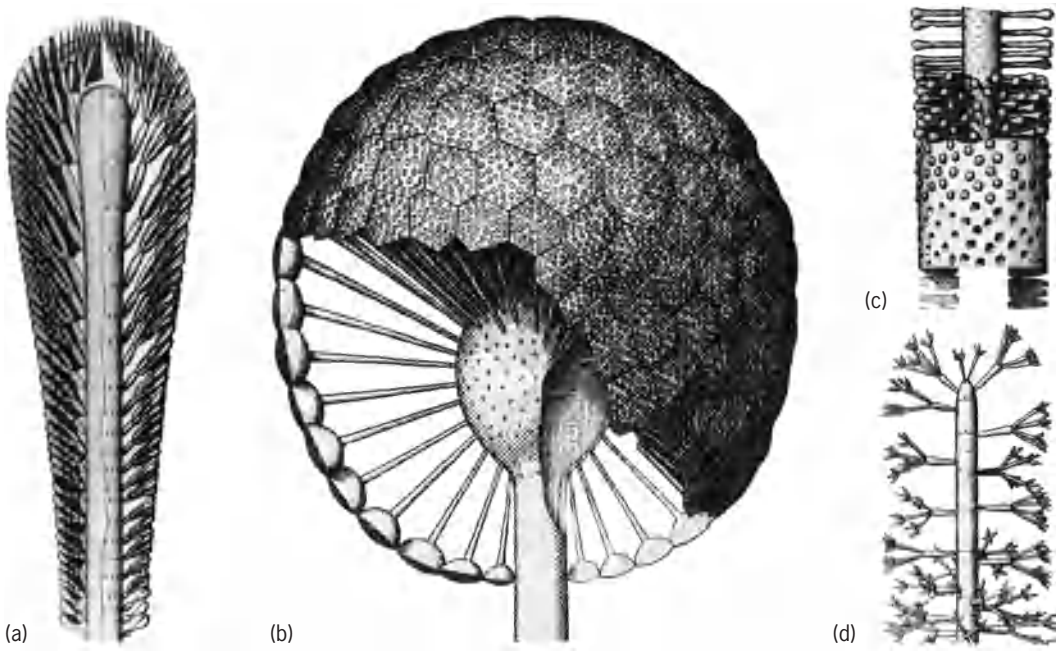

Fig. 1. Reconstructions of fossil Dasycladales. (*a*) *Palaeodasycladus* (Lower Jurassic, Italy); longitudinal section showing whorled branching to third order. (*b*) *Cyclocrinus* (Ordovician, Baltic); stalk terminating in head of unbranched laterals bearing reproductive cysts which cohere laterally to form a cortex. (*c*) *Rhabdoporella* (Upper Silurian, Sweden); single order of spirally arranged branches. (*d*) *Primicorallina* (Ordovician, New York); spiral branching to third order.

is still relatively large (about 40 μm in diameter), but has only one nucleolus. About 3 days following the completion of nuclear reduction, a gigantic intranuclear spindle is formed (to 120 μm long at late telophase). Although only two daughter nuclei are formed, chromosome counts support the belief that the division is meiotic.

A sequence of mitotic divisions results in the formation of numerous small secondary nuclei, which are carried by cytoplasmic streaming into the rays or branches (gametangia) that constitute the fertile cap. Each secondary nucleus together with surrounding protoplasm initiates the formation of an operculate cyst. Mitotic divisions prior to the onset of dormancy and immediately preceding germination result in numerous nuclei, which are incorporated in biflagellate isogametes. The cysts may become calcified and be liberated only when the parent plant disintegrates. Gametes from a single cyst are sexually incompatible. The quadriflagellate zygote swims briefly, then settles and develops into an irregular, prostrate, siphonous structure from which the primary axis arises.

Acetabularia (mermaid's wineglass) is the best-known genus of Dasycladales, partly because it occurs in the Mediterranean (as well as in all tropical seas), where its striking form gained the attention of early naturalists. Its extraordinary cytology and life history have made it a favorite subject of investigation by experimental morphologists, biochemists, enzymologists, biophysicists, and geneticists. The nucleus-containing rhizoid can be excised and the stalk cut into viable sections. Experiments that involve intraspecific and interspecific grafting of segments with or without a nucleus have provided important information concerning the nuclear control of morphogenesis.

In *Bornetella* (**Fig. 3**) and *Neomeris* the stalk is covered by lateral cohesion of the inflated apices of second-order branches. The thallus of *Bornetella* is spherical and only slightly calcified, with the gametangia enclosed within the continuous cortex. The thallus of *Neomeris* is cylindrical or clavate and heavily calcified, with the gametangia exposed at the base of the thallus by disintegration of the cortex. *Batophora* is exceptional in being completely uncalcified, its repeatedly divided branches remaining free. The thallus of *Cymopolia* is an articulate series of cylindrical, beadlike, calcified segments. This genus is exceptional in having dichotomous or irregular branching and coenocytic segments. Unlike *Acetabularia*, it produces gametangia continuously in uncalcified apical segments throughout growth.

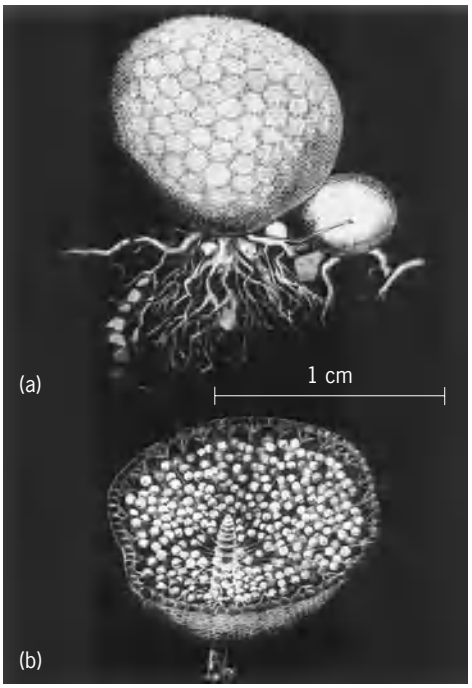

Fig. 3. ***Bornetella sphaerica*** **(extant). (*a*) Habit of thallus, showing sand grains trapped in rhizoids. (*b*) Oblique view into thallus, showing delicate primary branches bearing globose gametangia and whorls of secondary branches joined laterally to form a cortex.**

The extant Dasycladales are usually placed in a single family, Dasycladaceae, but those forms in which fertile and sterile whorls alternate or only one fertile whorl is produced may be segregated as the Acetabulariaceae (or Polyphysaceae). *See* ALGAE; CHLOROPHYCEAE.

Paul C. Silva; Richard L. Moe

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae: Structure and Reproduction*, 1985; L. E. Egerod, An analysis of the siphonous Chlorophycophyta with special reference to the Siphonocladales, Siphonales, and Dasycladales of Hawaii, *Univ. Calif. Publ. Bot.*, 25:325–453, 1952.

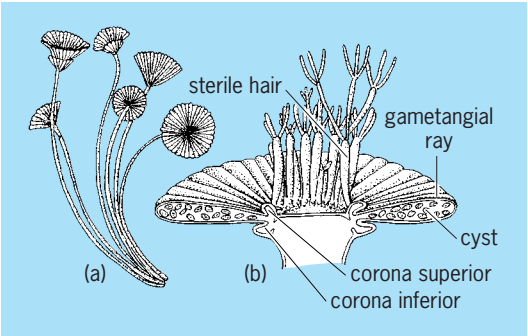

Fig. 2. ***Acetabularia*****. (*a*) Thallus of *A. crenulata*, showing stalk bearing fertile cap or disk. (*b*) Section of fertile disk of *A. acetabulum*. (*After G. M. Smith*, *Cryptogamic Botany*, *vol. 1*, *2d ed.*, *McGraw-Hill*, *1955*)**

Data communications

The conveyance of information from a source to a destination. Data means the symbolic representation of information, generally in a digital (that is, discrete) form. (Analog information refers to information encoded according to a continuous physical parameter, such as the height or amplitude of a waveform, while digital information is encoded into a discrete set of some parameter, for example, a voltage level.) Usually, this digital information is composed of a sequence of binary digits (ones and zeros), called bits. The binary system is used because its simplicity is universally recognizable and because digital data have greater immunity to electronic noise than analog information and

allow flexible processing of the information. Groups of eight bits create a data byte or character. These characters make up the so-called alphabets (including alphabetic, numeric, and special symbols) which are used in data communications. The most common data sources and destinations are computers, computer peripherals, and other data devices [such as digital phones, cellular phones, personal digital assistants (PDAs), and digital imaging applications], and the data represent groups of characters to form text, hypertext (text organized according to topic rather than linear sequence), or multimedia information, including audio, graphics, animation, and video information. *See* BIT; COMPUTER; COMPUTER GRAPHICS; COMPUTER PERIPHERAL DEVICES; DIGITAL COMPUTER; ELECTRICAL NOISE; MULTIMEDIA TECHNOLOGY; NUMBERING SYSTEMS.

However, the term data communications more specifically deals with the methods by which this digital information is conveyed. The importance of data communications continues to grow. As society becomes increasingly information intensive and connected more applications require the best methods for conveying this information. Data communications may be accomplished through two principal functions, data transmission and data switching or routing.

Transmission and the data channel. Data transmission always involves at least three elements: a source of the information, a channel for the transmission of the information, and a destination for the information. In addition, sometimes the data are encoded. The codes can be used for error detection and correction, compression of the digital data, and so forth. This encoding and the corresponding decoding are specific to data in digital form. Depending on the application, the data can be transmitted point-to-point (for example, from a computer to its printer) or can be broadcast to many receiving points (similar to a television broadcast). *See* DATA COMPRESSION.

Transmission media. The communications channel is carried over a transmission medium. Such media can be wired, as in the cases of twisted-pair telephone wires, coaxial cables, or fiber-optic cables, or they can be wireless, where the transmission is not confined to any physical medium, such as in radio, satellite, or infrared optical transmission. Sometimes, even when the source of the information is digital, the transmission medium requires analog signaling, and modems (modulators-demodulators) are required to convert the digital signals to analog, and vice versa. For example, data communication between personal computers transmitted over telephone lines normally uses modems. Other transmission media (such as fiber systems) can support digital transmission directly. The choice of media frequently depends on the desired performance and cost criteria. *See* COAXIAL CABLE; COMMUNICATIONS CABLE; COMMUNICATIONS SATELLITE; MODEM; OPTICAL COMMUNICATIONS; OPTICAL FIBERS; RADIO; TELEPHONE SERVICE.

Parameters. The transmission is characterized by a number of parameters, including transmission directionality, whether or not the link is wired, whether bits are transmitted serially or in parallel, available capacity (bandwidth), data transmission rate (bits per second), channel error rate (the probability of errors per bit transmitted), channel reliability (how often the channel fails), and channel security (how easy it is to eavesdrop on the channel).

The directionality of the information can be either one-way (simplex communications) or two-way. Two-way communications can be either half-duplex (information goes both ways over the communications link, but not at the same time) or full-duplex (information goes both ways at the same time). The speed of transmission can vary greatly, depending on the application, from a few bits per second (for applications such as alarm monitoring or point-of-sale terminals) to advanced fiber-optic transmission systems, where many billions of bits per second can be transmitted along a single optical fiber. (For comparison, a typical 400-page book, entirely text, contains approximately 15–25 million bits.)

The data channel can be a serial channel, in which the bits are transmitted one after another across a single physical connection; or a parallel channel, in which many bits are transmitted simultaneously (for instance, over parallel wires). Fiber-optic communications, for example, is usually serial, whereas the backplane of a computer, where data are exchanged between memory and processing elements, is usually parallel. Generally, parallel channels are used for short-distance links (less than 300 ft or 100 m), whereas serial links are used for larger distances and high data rates.

At low data rates (less than a few hundred kilobits per second) communications channels are typically dedicated, whereas at higher data rates, because of the cost of high-speed transmitters and receivers, shared channels are used by multiplexing the data streams. For example, two independent data streams with constant data rates of 10 megabits per second (Mb/s) could use a shared channel having a data-rate capability of 20 Mb/s. The multiplexing system would select one bit from each of the two channels to time-multiplex the data together. Similarly, at the destination the channels could be demultiplexed to restore the independent 10-Mb/s channels (**Fig. 1**). Here, the available information-carrying capacity of the channel (20 Mb/s) is shared by the two users through multiplexing. There are many different types of multiplexing used to share the channels; for example, there can be multiplexing in frequency, time, coding, or amplitude. *See* MULTIPLEXING AND MULTIPLE ACCESS.

In choosing the technology for transmission, the appropriate parameters depend on the application requirements. For example, the link between a personal computer and a printer requires lower performance (and lower cost) than that connecting two high-performance computers.

Switching. In some applications, the data communications link can be a permanent, dedicated connection, for example, from a computer to its printer. However, in many cases, the source, the destination, and the path taken by the data may vary; thus switching is required.

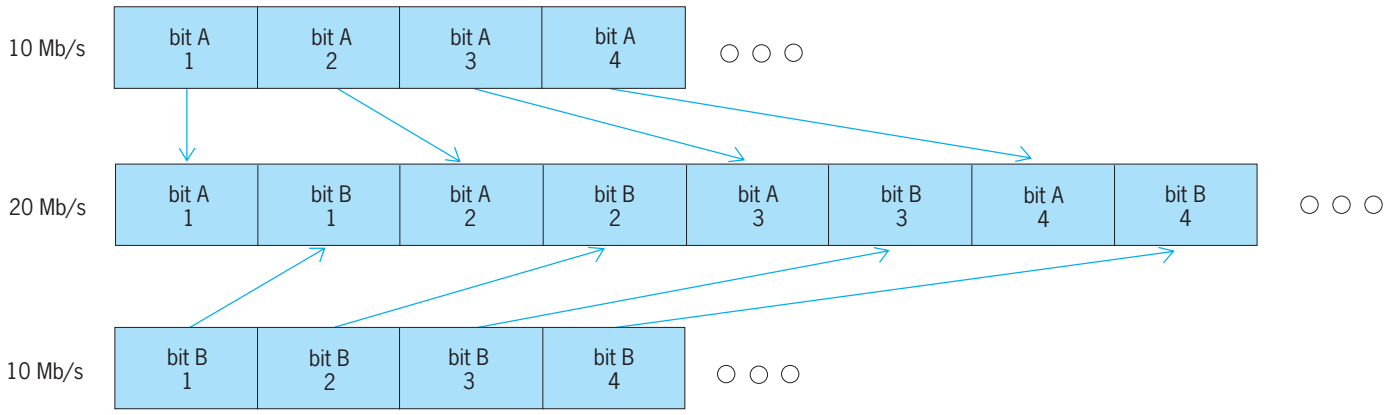

Fig. 1. Bit-stream multiplexing.

Types. The two primary types of switching employed in data communications are circuit switching and packet switching. In circuit-switched data communications, an end-to-end connection is established prior to the actual transmission of the data and the communications channel is open (whether or not it is in use) until the connection is removed. Since the full information-carrying capacity of the channel (the bandwidth) is available to the user, no additional connections along that channel are allowed (that is, new requests for the use of the channel are blocked) as long as that connection is in place.

A packet is a group of data bytes which represents a specific information unit with a known beginning and end. The packet can be formed from either a fixed or variable number of bytes. Some of these bytes represent the information payload, while the rest represent the header, which contains address information to be used in routing the packet. The packet is somewhat analogous to an envelope. The information payload is the content of the envelope, whereas the header is like the address on the envelope.

In packet switching, unlike circuit switching, the packets are sent only when information transmission is required. The channel is not used when there is no information to be sent. Sharing the channel capacity through multiplexing is natural for packet-switched systems. Furthermore, packet switching enables users to access the full capacity of the communications links, even when the links are already in use; that is, the packet switches allow for temporary loading of the network beyond the transmission capacity of the channel. This information overload is stored (buffered) in the packet switches and sent on when the channel becomes available. (Thus packet switching is sometimes referred to as store-and-forward switching.) Of course, such network overloads must be temporary.

A characteristic of data communications traffic (the amount of data being transmitted through a link) is that it is often bursty. That is, the traffic is not constant in time, but has intense periods and very light periods. Thus, packet switching is well suited to this type of traffic. An important characteristic of the information being sent is the allowable delay and loss characteristics. Voice-telephone traffic and video traffic, for example, require constant bit rates without variable delay. If a several-second pause in the transmission were to occur, it would be disturbing to a phone call or video transmission. In contrast, for most data traffic a few-second delay may not even be noticed (for example, from a point-of-sale charge-card terminal to a bank). However, data signals require very low losses (a single bit error can cause retransmission of a whole packet), whereas a few bits lost for audio or video signals may be easily tolerated.

Although circuit switching is widely used in the voice-telephone network, where calls are established for holding times typically on the order of minutes, modern data communication networks are largely structured around packet-switching technologies. Packet switching provides flexibility between various types of equipment with inherently different data rates, as well as efficient sharing of the transmission capacity between bursty data sources.

Protocols. In order to transfer information from a sender to a receiver, a common physical transmission protocol must be used. Protocols can range from very simple to quite complex. For example, a protocol widely used for low-data-rate serial data transmission from a computer to and from its low-speed peripherals is defined by a common set of electrical specifications referred to as RS-232-C. (Actually, the RS-232-C protocol uses analog signaling between digital devices.) Today the most widely used data link protocol is the Ethernet protocol. In the local-area network (LAN) it is used between 10 and 100 Mb/s, and between Ethernet switches it is used at either 1 gigabit per second (Gb/s) or 10 Gb/s. More than 95% of all data packets originate on Ethernets. The most widely used protocol for data communications over a vast range of technologies is the Internetwork Protocol (also known as the Internet Protocol or IP).

Existence of connections. Data communications can be connection-oriented or connectionless. In connection-oriented systems, each time that information is required, a connection is established and the information is passed over the connection path. In connectionless systems, the header information in the data packet is used by each of the packet switches along the path to route the packet to its destination.

Packet-switched networks. A network is composed of a number of communication links working together. The complexity of the data communications can depend on the interconnection and network size. At the shortest distances, communications occur between integrated circuits (chip-to-chip interconnection). This is followed by interconnections between circuit boards containing the chips (board-level interconnection). Groups of boards are interconnected (backplane interconnection) to form computers and other digital processing systems, and computers are interconnected to form local-area networks (LANs). Local-area networks are interconnected to form metropolitan-area networks, and these are interconnected to form wide-area networks (WANs). Finally, the wide-area networks are interconnected to form national-scale networks, which are interconnected to form global networks spanning the world. *See* LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

Protocol layering. As network complexity grows, a hierarchy of layers is used to better manage and control the network. The layer definitions and the protocol definitions are portions of the network architecture. The International Standards Organization developed a widely used model for these layers, called the Open Systems Interconnect (OSI) model (**Fig. 2**). The essential feature of the OSI model is that it reduces protocol complexity by breaking the protocol into smaller independent functional units which operate in conjunction with matched functional units at a peer-to-peer level. Each layer performs functions for the next higher layer by building on the functions provided by the layer below. The advantage of performing communications based on this model is that the protocol at each layer can be defined independent from the other layers. Standard protocols at each layer can be applied, enabling internetworking.

At the lowest layer, there is physical transmission, which is concerned with transmitting raw data bits. Important issues at this layer include the transmission data rate, the bit error rate, and the delay through the links. Above this physical layer are virtual communications layers, designed to simplify the management of complex functions. They are virtual in the sense that there is no additional physical connection at any layer above the physical transmission. The data-link layer takes the physical transmission of bits and makes it appear to the higher layers that the communications are error free. This may be accomplished by defining some bytes in the information as check bytes, to detect transmission errors. Building on this, the network layer has the job of converting the information into packets and determining the routing of the packets. This layer has to deal with the flow of the traffic in the network and possible network congestion.

The transport layer accepts information from the user layers, apportions it into smaller units for transmission through the network, and is responsible for regrouping the units to ensure that they arrive at the destination in the correct order and without errors. The transport layer is responsible for end-to-end flow control. For example, in the communications between a computer and a printer, the transport layer controls the flow of data out of the computer so that the printer is not overwhelmed by it. This layer also determines what type of service to provide to the higher layers. The data-communications-networking functionality is largely confined to these first four layers in the network hierarchy. The highest layers are related to the end-user applications.

Ethernet protocol. Some protocols take explicit advantage of the bursty nature of data communications. For example, the Ethernet protocol (carrier sense multiple access and collision detect, or CSMA-CD) is based on a common bus (ether) in which every transmission is broadcast to all users. The Ethernet protocol works as follows: When a host computer wants to transmit a packet, it first listens to the physical medium (the ether) to see if it is currently being used. If not, it transmits its data packet. If the ether is busy, it waits a specific time interval, then tries to transmit again. This same protocol can be used for all the hosts on the ether. The host computers

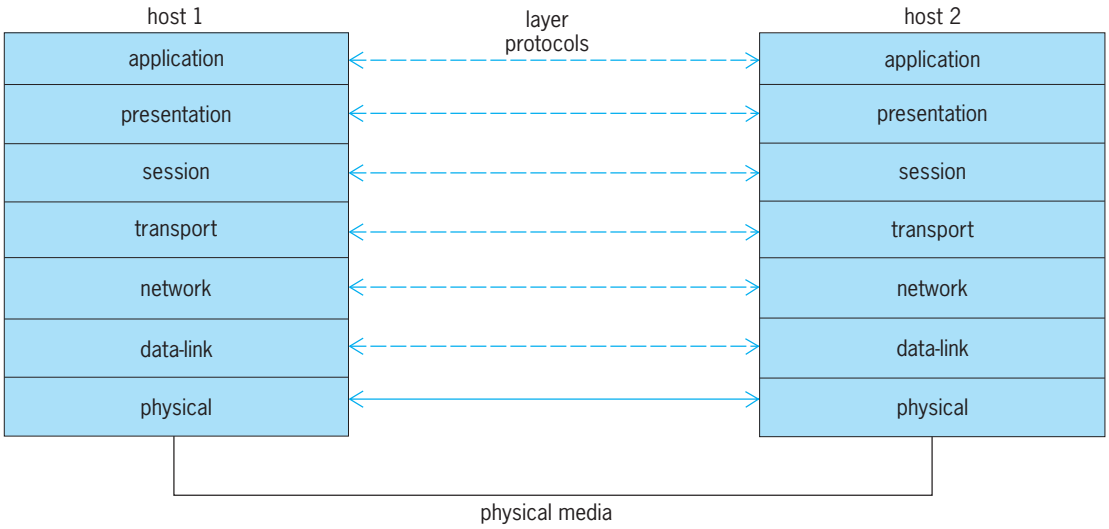

Fig. 2. Open Systems Interconnect (OSI) model for network protocols.

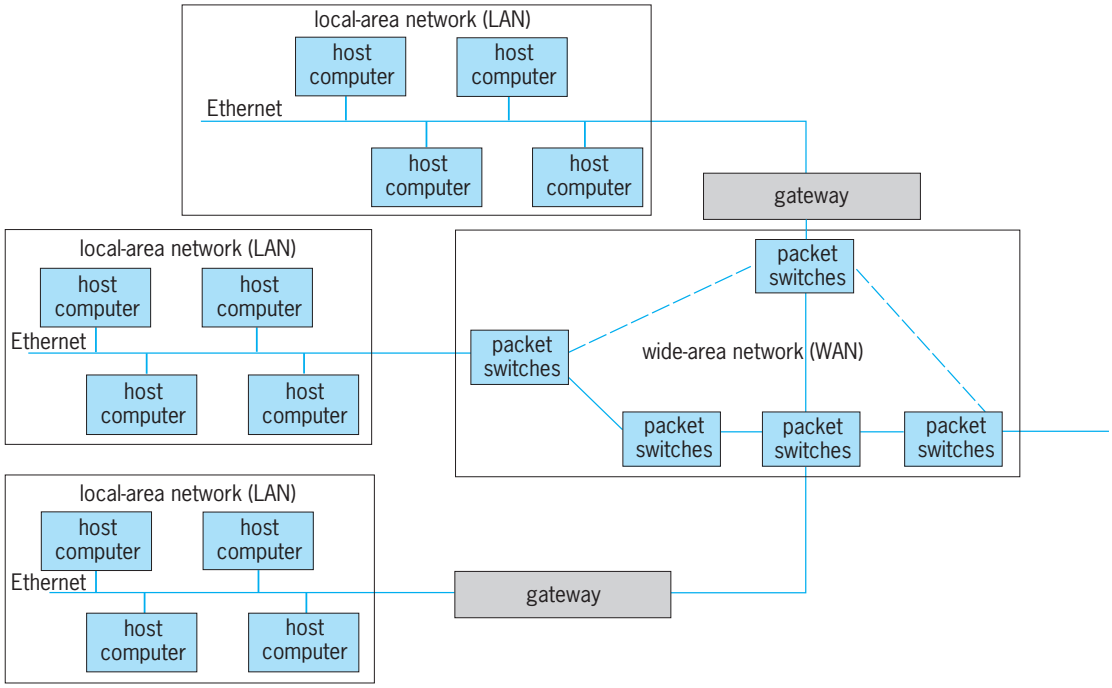

Fig. 3. Data communications network.

must be able to detect packet collisions caused by two hosts simultaneously transmitting on the ether. Collision recovery is accomplished by aborting the transmission, waiting a random length of time, and then trying again.

While the Ethernet (CSMA-CD) would not make sense for telephone traffic, it is very efficient for bursty data communications within a LAN. This type of system is feasible only if the packet propagation time (the time to travel from the source to the destination) is short relative to the packet transmission time (the time required to send the packet). Thus, this type of protocol is well suited to LANs, but would not be useful in large-scale networks (for example, over WANs or satellites), where the propagation time can be large relative to the transmission time because of the great distances involved. However, the original Ethernet protocol has been extended to a switched Ethernet protocol in which there is bandwidth reservation rather than packet collision detection as described above. This protocol can be used at data rates from 10 and 100 Mb/s to more than 10 Gb/s and over large distances.

Network interconnection. To form wide-area networks, multiple LANs are usually interconnected by using gateways or bridges, which provide communications paths so that data may be exchanged between networks (**Fig. 3**). To achieve integration of these subnetworks into a larger framework of networks, these gateways must support internetworking protocols. The most common internetworking protocol is based on a system known as TCP/IP (transmission control protocol/internetwork protocol). The IP protocol provides the networking protocol portion of the OSI model and TCP provides the transport layer.

Internet. The Internet is an immense data network of networks. It has grown from an experimental network called the ARPAnet, originally operated by the U.S. Defense Advanced Research Projects Agency and using the TCP/IP protocol, into the world's largest network of networks. The Internet now supports a huge variety of protocols in addition to the TCP/IP protocol (many of them using the TCP/IP protocol). The Internet has had pervasive impact on all aspects of communications. It has revolutionized communications for commerce, science, and art, and has had a large-scale impact on communications in daily life. It has given rise to many other protocols, including hypertext transmission protocol (HTTP) and the World Wide Web (WWW), and the many familiar browsers and higher-level applications.

Addressing in the Internet is accomplished through a specific protocol version. While the version known as IPv4 is widely used, an evolution is taking place to greatly extend the original addressing scheme. IPv6 is a new version of IP which is designed to be an evolutionary step from IPv4. IPv6 is designed to run well on high-performance networks (such as Gigabit Ethernet, OC-12, and ATM) and at the same time still be efficient for low-bandwidth networks (such as wireless networks). In addition, it provides a platform for new Internet functionality that will be required in the near future. IPv6 includes a transition mechanism which is designed to allow users to adopt and deploy IPv6 in a highly diffuse fashion and to provide direct interoperability between IPv4 an IPv6. The IPv6 transition allows users to upgrade their hosts to IPv6, and the network operators to deploy IPv6 in routers, with very little coordination between the two.

The Internet enables the integration and convergence of voice, data, and video transmission, and it is commonly used for Voice over IP (VoIP) and streaming video transmission as well as for its originally

intended computer data networking applications. The Internet also supports many higher-layer applications such as the World Wide Web. There is no single coordinating or operating group for the Internet, but many of the developments and enhancements to the Internet originate from the Internet Engineering Task Force (IETF), a large international group of network architects, designers, equipment vendors, and carriers, who recommend evolutionary changes to the Internet. Use of the Internet is the largest, most pervasive electronic connectivity in human history—it is the networking parallel of the development of microelectronics. *See* INTERNET; VOICE OVER IP; WORLD WIDE WEB.

Other networking protocols. One advanced protocol, frame relay, supports packets (frames) of variable length, from a few bytes to around 10,000 bytes. Frame-relay functions are connection-oriented; that is, frames of individual virtual connections are transferred by means of address information in the control part of the frame. Error correction is implemented by retransmission of faulty frames end to end. The network switches only frames for which no errors were found, and discards all faulty frames. Frame relay is typically used for applications at bit rates up to about a few megabits per second. Frame relay, while suitable for a range of data communications applications, has limited applicability to multimedia data involving time-sensitive information (such as audio or video), which must be delivered without large, variable delays.

Asynchronous transfer mode. Communication between powerful desktop computing platforms generates large volumes of data communications traffic, which requires high transmission speeds. Furthermore, demands are placed on LANs and other networks for multimedia traffic, and include variable-bit-rate data as well as continuous-bit-rate voice and video. At these higher data rates, packet-switching protocols based on the asynchronous transfer mode (ATM) are used. Variable-length packets can be difficult and slow to forward through networks, and may cause performance limitations to continuous-bit-rate services. ATM addresses this problem through a packet-switching protocol where fixed-length, 53-byte packets (cells) are used. The information payload is 48 bytes, with 5 bytes for header information. ATM may be used with systems at data rates ranging from low (around 1.5 Mb/s) to very high (up to 2.5 and 10 Gb/s). ATM may be suitable for both bursty and continuous-bit-rate traffic. Bursty traffic is assigned a lower priority so that it takes advantage of the statistical multiplexing possible for such traffic, while reasonable performance is guaranteed for continuous-bit-rate services which typically have end-to-end delay requirements. While ATM is widely used in the public switched telephone network and some data networks, due to its complexity, it is being replaced by Switched Ethernet.

Switched Ethernet. As described above, Ethernet began as a LAN technology, but it has continued to evolve and is also used as a metro- and wide-area-networking technology. These capabilities began with the development of 1-Gb/s Ethernet capability, and 10-Gb/s Ethernet has become prevalent. The advantages of Ethernet are its low cost due to its widespread use and large market in the LAN and enterprise areas, and its relative simplicity.

Multiprotocol label switching. Another set of protocols is called Multiprotocol Label Switching (MPLS), which is an IETF initiative to integrate layer 2 information (bandwidth, utilization, and so forth) about network links to layer 3 and to simplify and improve the performance at layer 3 (the IP layer). Simply put, when a packet enters an MPLS network, label edge routers give it a label. These labels contain not only information on how to route the packets but also information about the types of service that are being requested and the protocols that are being used.

MPLS enables routing of the information based on this label information, so that different qualities of service may be supported. MPLS gives network operators a great deal of flexibility to divert and route traffic around link failures, congestion, and bottlenecks. It essentially adds "superhighways" and can help address network congestion at specific routers. There are a number of generalizations and variants of MPLS as well.

Wireless protocols. The need for mobility in data communications has given rise to the development of a large number of wireless networking protocols. Some of the most common are the IEEE 802.11 suite of protocols. These layer 3 and 4 protocols address the unique requirements of the wireless environment. *See* ELECTRICAL COMMUNICATIONS; INTEGRATED SERVICES DIGITAL NETWORK (ISDN); PACKET SWITCHING; SWITCHING SYSTEMS (COMMUNICATIONS); TELEPROCESSING; WIRELESS FIDELITY (WI-FI).

Matthew S. Goodman

Bibliography. D. P. Bertsekas and R. Gallager, *Data Networks*, 2d ed., 1991; M. de Prycker, *Asynchronous Transfer Mode: Solution for Broadband ISDN*, 3d ed., 1995; D. Greenblatt, *The Call Heard 'Round the World: Voice over Internet Protocol and the Quest for Convergence*, 2003; I. Pepelnjak and J. Guichard, *MPLS and VPN Architectures*, 2000; W. Stallings, *Computer Networking with Internet Protocols and Technology*, 2003; W. Stallings, *Data and Computer Communications*, 7th ed., 2003; A. S. Tanenbaum, *Computer Networks*, 4th ed., 2003.

Data compression

The process of transforming information from one representation to another smaller representation from which the original, or a close approximation to it, can be recovered. The compression and decompression processes are often referred to as encoding and decoding. Data compression has important applications in the areas of data storage and data transmission. Besides compression savings, other parameters of concern include encoding and decoding speeds and workspace requirements, the ability to access and decode partial files, and error generation and propagation.

Basic Principles

The data compression process is said to be lossless if the recovered data are assured to be identical to the source; otherwise the compression process is said to be lossy. Lossless compression techniques are requisite for applications involving textual data. Other applications, such as those involving voice and image data, may be sufficiently flexible to allow controlled degradation in the data.

Data compression techniques are characterized by the use of an appropriate data model, which selects the elements of the source on which to focus; data coding, which maps source elements to output elements; and data structures, which enable efficient implementation.

Source data. The source data typically are presented in a format in which each datum is represented by a fixed number of bits. For textual data, each character is typically represented in eight bits using, for example, the extended American Standard Code for Information Interchange (ASCII) mapping. For speech, each datum consists of a real-valued sample of the waveform expressed to a set precision. For images, each picture element or pixel typically is represented by 1 bit (for black-white), 8 bits (for gray scale), or 24 bits (for color images, 8 bits intensity for each of three primary colors).

Models. Information theory dictates that, for efficiency, fewer bits be used for common events than for rare events. Compression techniques are based on using an appropriate model for the source data in which defined elements (source values in general or within a particular context) are not all equally likely. The encoder and the decoder must agree on an identical model. The encoder determines the relevant elements under the agreed model and outputs sufficient information to enable the decoder to determine these elements. *See* INFORMATION THEORY.

A static model is one in which the choice of elements and their assumed distribution is invariant. For example, the letter "e" might always be assumed to be the most likely character to occur. A static model can be predetermined with resulting unpredictable compression effect, or it can be built by the encoder by previewing the entire source data and determining element frequencies. This latter process suffers from the necessity of the encoder making two passes over the data, one pass for determining frequencies and another for encoding, and also from the delay imposed on producing any output until after the entire source has been scanned. In addition, the relevant aspects of the determined frequency distribution must be sent to the decoder, which would otherwise have no means of determining it. The benefits of using a static model include the ability to decode without necessarily starting at the beginning of the compressed data.

An alternative dynamic or adaptive model assumes an initial choice of elements and distribution and, based on the beginning part of the source stream that has been processed prior to the datum presently under consideration, progressively modifies the model so that the encoding is optimal for data distributed similarly to recent observations. Some techniques may weight recently encountered data more heavily. Dynamic algorithms have the benefit of being able to adapt to changes in the ensemble characteristics. Most important, however, is the fact that the source is considered serially and output is produced directly without the necessity of previewing the entire source.

In a simple statistical model, frequencies of values (characters, strings, or pixels) determine the mapping. In the more general context model, the mapping is determined by the occurrence of elements, each consisting of a value which has other particular adjacent values. For example, in English text, although generally "u" is only moderately likely to appear as the "next" character, if the immediately preceding character is a "q" then "u" would be overwhelmingly likely to appear next.

In a sequential model, the source is considered in order of its (one-dimensional) presentation. In a hierarchical model, the source is recursively partitioned into smaller subparts. The quadtree representation, which recursively splits a square into four subsquares, exemplifies the use of this model.

Coding techniques. The use of a model determines the intended sequence of values. (For example, if a dictionary model is used for text compression, the output may be indices into a hash table. If a gray image is compressed, the output may correspond to gray-scale values, typically in the range 0–255.) An additional mapping via one coding technique or a combination of coding techniques is used to determine the actual output. Several commonly used data coding techniques will be described. The decoder is made aware of which coding technique has been employed either by prearrangement or by an inserted codeword.

Dictionary substitution. In using this technique, a repeated sequence of characters is replaced with a reference to an earlier occurrence. A table to possible values for reference is maintained; a value from this table is presented as an index into the table. In ordered table indexing, the table of possible values is ordered by some predetermined criteria. Actual frequency can be used, necessitating either a prescan or the maintenance of a corresponding dynamic frequency table. Alternatively, a strategy for self-organizing a list, such as transpose or move-to-front, may be used. The method of static dictionary substitution is useful for representing alphanumeric representations of information, such as gender or month. In a Lempel-Ziv (or dynamic dictionary) approach, as additional source characters are scanned, some table entries may be deleted and appropriate candidates may be considered for addition to the table. The variations on this method differ in terms of how the references to earlier occurrences are presented and how the deletions and additions to the table are selected.

A typical Lempel-Ziv procedure is as follows. The table is initialized to contain the set of singleton characters. Iteratively the longest string s in the table that matches the source is determined, and the corresponding table index is output. If the following

nonmatched source character is c, then the concatenation of s followed by c is added to the table if space remains. When the table becomes full, this method reverts to static dictionary substitution. Alternatively, after the table becomes full and the saving from compression is observed to degrade below a trigger level, the table is reinitialized. Another approach maintains a sliding window in which string references may occur. For example, a longest match of the to-be-compressed source is sought within the immediately proceding 4095 characters of the already-compressed source.

Run-length encoding. With this approach, a sequence of values is presented as a sequence of pairs (c,v), where c is the count of the number of values in a row that have the same value v or are within a predefined margin of allowable error.

Predictive coding. For this technique, a sample value is presented as the error term formed by the difference between the sample and its prediction. In difference mapping, it is expected that the next value is close to the current value, and so the predicted value is defined to be equal to the previous sample. In a linear predictive model, the expectation of the next value is the extrapolation of the present and previous values. Thus, if the previous value was 200 and the current value is 220, the next value will be 240. There are many other predictive models, both one- and two-dimensional, in use.

Huffman coding. The essence of this coding technique is that a frequency or probability distribution is associated with the set of candidate values. A code tree is constructed in such a manner as to minimize the expected weighted path length of the tree. A value is presented as a path from the root of the code tree to a leaf, expressed as a sequence of 0's and 1's denoting left and right tree branches.

Arithmetic coding. This technique is based on the concept that associated with the set of candidate values is a frequency or probability distribution from which can be determined a cumulative frequency distribution and, ultimately, associated subintervals of the unit interval. For example, if the only characters are X with probability 20%, Y with probability 70%, and end-of-string (EOS) with probability 10%, then X occupies the interval 0 to 0.2, Y occupies 0.2 to 0.9, and EOS occupies 0.9 to 1. Accordingly, associated with a sequence of values will be an increasingly smaller subinterval of the unit interval. A sequence of values is presented as a real number within that ultimate small subinterval. Continuing with the earlier example, the string YX would be associated with the interval 0.326 to 0.34 because the first character (Y) focuses on the interval 0.2 to 0.9, the second character (X) focuses on the first 20% of that interval (0.2 to 0.34), and the string end focuses on the last 10% of that. Similarly, the string XY would be associated with the interval 0.166 to 0.18. The code is the final real number to whatever precision is required to distinguish it from other contenders.

Variable-length representation of integers. A number of different coding techniques have been developed that, depending on the relative expected frequency of integers, can present the values of the integers in compact form. Such codes include Elias codes, Fibonacci codes, and start-step-stop codes.

Scalar quantization. In this technique, a value is presented (in approximation) by the closest, in some mathematical sense, of a predefined set of allowable values. In some cases, the allowable values may be nonuniformly spaced. For example, a real number may be presented rounded to the nearest three decimal places or truncated to five significant digits.

Vector quantization. In this method, a sequence of k (some fixed number) values is presented as resembling the template (from among the choices available in a given codebook) that minimizes a distortion measure. Under this approach, the encoder must output information sufficient for the decoder to be aware of the contents of the codebook as well as the selection.

Transform coding. For this technique, a block transform (such as discrete cosine transform) is applied to a vector (or a subimage) of sampled values. The transformed coefficients are then scalar quantized.

Daniel S. Hirschberg

Digitized Audio and Video Signals

The information content of speech, music, and television signals can be preserved by periodically sampling at a rate equal to twice the highest frequency to be preserved. This is referred to as Nyquist sampling. However, speech, music, and television signals are highly redundant, and use of simple Nyquist sampling to code them is inefficient. Reduction of redundancy and application of more efficient sampling results in compression of the information rate needed to represent the signal without serious impairment to the quality of the remade source signal at a receiver. For speech signals, redundancy evident in pitch periodicity and in the format (energy-peaks) structure of the signal's spectrum along with aural masking of quantizing noise are used to compress the information rate. In music, which has much wider bandwidth than speech and far less redundancy, time-domain and frequency-domain masking are principally used to achieve compression. For television, redundancy evident in the horizontal and vertical correlation of the pixels of individual frames and in the frame-to-frame correlation of a moving picture, combined with visual masking that obscures quantizing noise resulting from the coding at low numbers of bits per sample, are used to achieve compression. *See* SPEECH; TELEVISION.

Compression techniques may be classified into two types: waveform coders and parametric coders. Waveform coders replicate a facsimile of a source-signal waveform at the receiver with a level of distortion that is judged acceptable. Parametric coders use a synthesizer at the receiver that is controlled by signal parameters extracted at the transmitter to remake the signal. The latter may achieve greater compression because of the information content added by the synthesizer model at the receiver.

Waveform compression. Waveform compression methods include adaptive differential pulse-code modulation (ADPCM) for speech and music signals, audio masking for music, and differential encoding

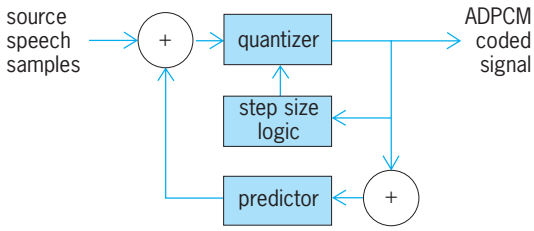

Fig. 1. Adaptive differential pulse-code modulation (ADPCM) encoder.

and sub-Nyquist sampling of television signals.

ADPCM. For coding of speech signals, this method (**Fig. 1**) incorporates predictors at the transmitter and receiver having response functions representative of the short- and long-term correlation inherent in the signal. The difference between the predicted signal and the input signal is coded by a signal quantizer that adapts to difference signal amplitude. The coded difference values are transmitted to the receiver and used to reconstruct the speech signal. The quality of the ADPCM-coded speech signal recovered at the receiver at 32 kilobits per second (kbps) is equivalent to that of pulse-code modulation at 64 kbps, and is suitable for public switched telephone service. ADPCM at 16 kbps has a lesser yet acceptable quality which is suitable for mobile telephone service. *See* PULSE MODULATION.

Coding for a bandwidth up to 7 kHz for speech and music is provided by an ADPCM coder that splits a 16-kHz-sampled, 14-bit, linearly quantized source signal into two 8-kHz critically sampled components, one for the low subband and the other for the high subband. Coding of the low subband uses 6 bits per sample yielding 48 kbps, and the high subband 2 bits per sample yielding 16 kbps, for a total of 64 kbps.

CD-quality coding. Good to excellent compact-disk (CD) quality for the reproduction of 14-, 20-, and 22-kHz bandwidth music signals has been achieved at coding rate rates of 2 bits per sample. The coding method is based on audio (music) signal masking of the quantizing noise. The input signal is coded by using a 16-bit-per-sample linear quantizer at 44.1 kHz for a 20-kHz-bandwidth input signal. The signal is windowed and converted to 28 equally spaced subbands, each having 750-Hz width, by a polyphase processor. The subband signals, sampled at 1.5 kHz, are quantized with a number of bits sufficient to keep the quantizing noise masked in each subband. A masking level in each subband is calculated by using a fast-Fourier-transform spectrum analyzer operating in parallel. The coded signal, comprising the sample bits and identifying the bit quantizer used for each subband, is sent to the receiver where an inverse polyphase processor remakes the digital audio signal. *See* COMPACT DISK; FOURIER SERIES AND TRANSFORMS; SPECTRUM ANALYZER.

Differential encoding. Television pixels are strongly correlated from line to line and from frame to frame. By coding these differences, the bit rate needed to transmit television can be reduced in excess of 4:1 depending on source material. At the receiver the differences progressively update a stored version of the frame that is displayed after each update. Motion prediction used in the difference coding can double the compression ratio. Differential coding can also be used with orthogonal transform coding as discussed below.

Sub-Nyquist sampling. This refers to sampling at a rate lower than Nyquist to preserve signal content without aliasing distortion. It can be applied to a television signal by sampling it at a rate less than Nyquist and at an odd multiple of the frame rate. This places the aliasing components into periodically spaced voids in the television spectrum where they can be removed by a comb filter at the receiver.

Parametric encoders. These include vocoders for speech signals and encoders using orthogonal transform techniques for television.

Vocoders. A vocoder analyzes the source speech signal at the transmitter to extract parameters suitable for control of a synthesizer that remakes speech at the receiver. A typical speech analyzer-synthesizer is based on a linear predictive coder (LPC) model of the human vocal cavity (**Fig. 2**). The synthesizer consists of an excitation source that presents a combination of periodic pitch pulses and noise to a dynamic vocal cavity model. The pitch pulses represent the excitation presented to the human vocal cavity by the vibration of the larynx, and the noise represents the excitation created as air flows through constrictions in the vocal cavity. The synthesizer has a transmission response which, under control of the parameters extracted at the transmitter, tracks the size and shape of the vocal cavity in the process of speech production. Compression occurs because these

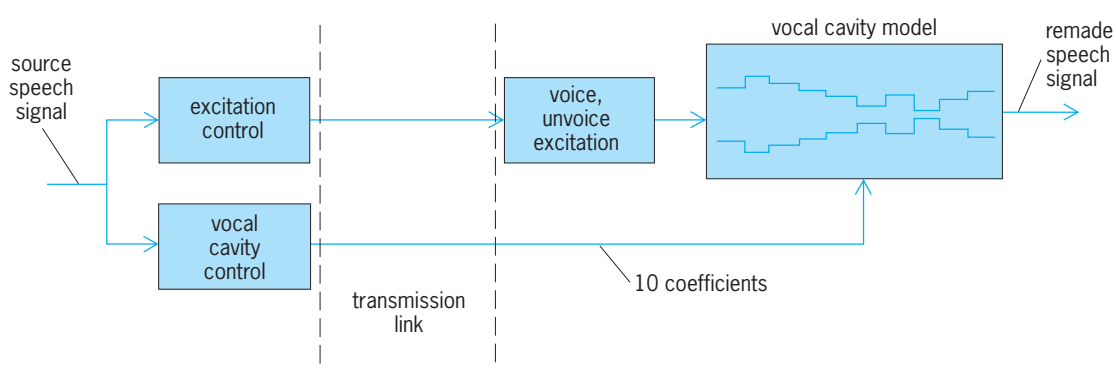

Fig. 2. Linear predictive coded (LPC) vocoder.

parameters vary at low rates compared to the high frequencies that appear in the source speech. Typical sampling rates for the parameters are 40–50 Hz. The number of bits per sample required to specify the parameters and the excitation is between 96 and 192, yielding transmission bit rates of 4800–9600 bits per second (bps) at 50-Hz sampling rate, compared with 64,000 bps typical of Nyquist sampling of telephone-quality voice signals.

The LPC synthesizer is equivalent to a cascade of tube sections, each section of fixed length and varying cross-sectional area, that approximates the size and shape of the vocal cavity. The remade speech of the LPC using 10 or 12 tube sections and excited by a multiband excitation (MBE) method is very natural at bit rates of 4800 and 9600 bps, but it still bears a trace of quality impairment. Multiband excitation dynamically assigns pitch excitation to some bands and noise excitation to other bands to excite the synthesizer. *See* VOICE RESPONSE.

Orthogonal transform techniques. Orthogonal transforms, such as discrete Fourier, Hadamard, and Haar, are used to convert blocks of $N \times N$ pixels into the coefficients of a set of N^2 orthogonal functions which are summed at a receiver to reconstruct a television picture. A two-dimensional discrete cosine transform (DCT), which is a Fourier transform, is widely used to code frame-to-frame differences for teleconferencing and broadcast-quality television. Typically, an array of 16×16 pixels, each pixel coded by 8 bits, is transformed into 256 DCT coefficients. In the reconstruction of the signal, some DCT coefficients must be coded with 10–12 bits, others with a lesser number, and some can be dropped altogether. On the average, 1 or 2 bits per coefficient are required. This represents a compression of 4:1 to 8:1, depending on the image content and the reproduction quality objective. When combined with frame-to-frame difference and motion prediction coding, overall compression ratios of 40:1 and greater result. Such techniques can also be used to achieve significant reductions in the transmission bit rate needed for high-definition television (HDTV). *See* INTEGRAL TRANSFORM.

S. Joseph Campanella

Bibliography. T. Bell, J. G. Cleary, and I. H. Witten, *Text Compression*, 1990; K. B. Benson and J. Whitaker, *Television Engineering Handbook: Featuring HDTV Systems*, rev. ed., 1992; G. Held, *Data Compression: Techniques and Applications, Hardware and Software Considerations*, 3d ed., 1991; N. S. Jayant and P. Noll, *Digital Coding of Waveforms*, 1984; S. Prentiss, *HDTV: High Definition Television*, 2d ed., 1993; J. A. Storer, *Data Compression: Methods and Theory*, 1988.

Data mining

The development of computational algorithms for the identification or extraction of structure from data. This is done in order to help reduce, model, understand, or analyze the data. Tasks supported by data mining include prediction, segmentation, dependency modeling, summarization, and change and deviation detection. Database systems have brought digital data capture and storage to the mainstream of data processing, leading to the creation of large data warehouses. These are databases whose primary purpose is to gain access to data for analysis and decision support. Traditional manual data analysis and exploration requires highly trained data analysts and is ineffective for high dimensionality (large numbers of variables) and massive data sets. *See* DATABASE MANAGEMENT SYSTEM.

A data set can be viewed abstractly as a set of records, each consisting of values for a set of dimensions (variables). While data records may exist physically in a database system in a schema that spans many tables, the logical view is of concern here. Databases with many dimensions pose fundamental problems that transcend query execution and optimization. A fundamental problem is query formulation: How is it possible to provide data access when a user cannot specify the target set exactly, as is required by a conventional database query language such as SQL (Structured Query Language)? Decision support queries are difficult to state. For example, which records are likely to represent fraud in credit card, banking, or telecommunications transactions? Which records are most similar to records in table A but dissimilar to those in table B? How many clusters (segments) are in a database and how are they characterized? Data mining techniques allow for computer-driven exploration of the data, hence admitting a more abstract model of interaction than SQL permits.

Data mining techniques are fundamentally data reduction and visualization techniques. As the number of dimensions grows, the number of possible combinations of choices for dimensionality reduction explodes. For an analyst exploring models, it is infeasible to go through the various ways of projecting the dimensions or selecting the right subsamples (reduction along columns and rows). Data mining is based on machine-based exploration of many of the possibilities before a selected reduced set is presented to the analyst for feedback.

KDD process. Data mining is a single step in the larger knowledge discovery in databases (KDD) process of identifying valid, novel, potentially useful, and ultimately understandable structure in data. Here, structure refers to either patterns or models. A pattern is an expression representing a parsimonious description of a subset of the data. A model is a representation of the source generating the data. The term "process" implies that knowledge discovery in databases comprises many steps. The step of data mining consists of the enumeration or identification of structures over the data. To be deemed knowledge, the derived structure must pass certain criteria such as utility (for example, gain, the amount of money saved by improved predictions) and validity (for example, estimated prediction accuracy). More difficult criteria to quantify include novelty and understandability. Sometimes various measures are combined as a single interestingness measure.

Density estimation and model selection. The fundamental problem of data mining is density estimation. If a density model of the joint probability density of the source that generated the data is available, then all questions about the data can be answered. Estimating the density, while a useful unifying view, is a difficult problem, and hence more specific tasks are often solved instead.

The other basic question is: Given a fixed set of data D, and many possible models for it, which model is to be selected? Formally, it is desirable to choose the model M for which the probability $\Pr(M|D)$ is maximized. Rewriting this probability using Bayes rule yields the equivalent expression given in

$$\Pr(M|D) = \frac{\Pr(D|M)\,\Pr(M)}{\Pr(D)}$$

which is known as the maximum a posteriori probability (MAP). The denominator of this expression is constant for the specified set of data, and $\Pr(M)$ is known as the prior probability of a model before the data are examined. The third term, $\Pr(D|M)$, is a measure of degree of the fit of the data to the model. When the prior probability is ignored, then maximum likelihood estimation is being performed, whereby the model is chosen that best fits the data. This procedure often leads to a problem known as overfitting the data with unjustifiably complex models. Maximum a posteriori probability hence gives a trade-off between model likelihood and fit to data. In statistics this approach appears under various names, including bias-variance trade-off. Taking logarithms of the expression in the above equation yields an equivalent form known as the minimum message (description) length, which calls for choosing the model that gives the minimal encoding of the data. The cost of minimally encoding an object M is at least $-\log[\Pr(M)]$ bits. Hence, minimizing $\log[\Pr(D|M)] + \log[\Pr(M)]$ prescribes choosing the model that minimizes the number of bits required to specify the model and the data, given the model. *See* BAYESIAN STATISTICS; PROBABILITY; STATISTICS.

Client-server decomposition. Algorithms for data mining build their models based on sufficient statistics or sufficient statistics in the data. Deployment of scalable data mining systems that can operate on large databases requires decomposing the operation into server-based operations that collect sufficient statistics (counts and data-specified parameters on models) and client-based operations that consume the sufficient statistics to produce models. This breakdown (see **illus.**) primarily avoids the unnecessary and expensive movement of large data sets. Unfortunately, SQL does not provide required sufficient statistics efficiently; hence, the server layer performs specialized data reduction operations. Traditional statistical or pattern recognition tools for data analysis fail on large databases since they assume that data fit in memory and hence do not scale. *See* CLIENT-SERVER SYSTEM.

Methods. Data-mining methods include predictive modeling, segmentation, data summarization, dependency modeling, and change and deviation detection.

Predictive modeling. This method predicts some field or fields in a database based on other fields. Regression refers to prediction of numeric (continuous) variables. Classification refers to prediction of categorical fields. The problem is to determine the most likely value of the variable being predicted, given the other fields (inputs), the training data (in which the target variable to be predicted is given for each record), and assumptions representing prior knowledge.

Segmentation. This method, also known as clustering, separates the data items into subsets that are similar to each other. A two-stage search is employed: an outer loop over possible numbers of clusters, and an inner loop to fit the best possible clustering for a given number of clusters.

Data summarization. Extracting compact descriptions of subsets of the data can be done by taking horizontal (cases) or vertical (fields) slices of the data. Taking horizontal slices provides summaries of subsets (for example, sufficient statistics, or logical conditions that hold for subsets); taking vertical slices finds relations between fields as opposed to predicting a specified field (classification) or grouping cases (clustering). One common method is called association rules. Associations are rules that state that certain combinations of values occur with other combinations of values with a certain frequency and associated confidence. A common application of this is market basket analysis, where the aim is to summarize which products are bought with what other products. The number of possible rules increases exponentially with the number of products, but due to data sparseness, only a few such rules satisfy given support and confidence thresholds. Scalable algorithms find all such rules in linear time (that is, in a time proportional to the size of the data set) for reasonable threshold settings. *See* ALGORITHM.

Dependency modeling. Insight into data is often gained by deriving causal structure: either probabilistic (statements about the probability distribution

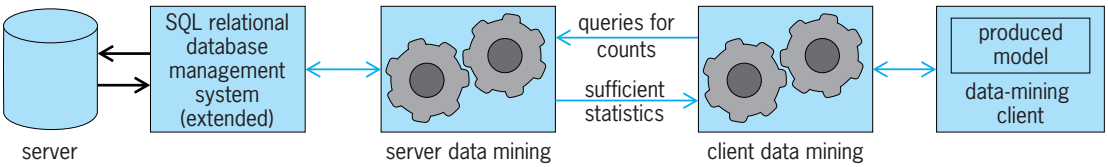

Client-server decomposition for data mining.

governing the data) or deterministic (functional dependencies between fields). In addition to density estimation, there are methods for explicit causal modeling (for example, belief network representations of density functions).

Change and deviation detection. These methods account for sequence information: time-series or some other ordering (for example, protein sequencing in genome mapping). Ordering of observations is important and must be accounted for. Scalable methods for finding frequent sequences in databases are exponential in complexity in worst-case scenarios. However, they appear to execute efficiently because of the sparseness of data in real-world transactional databases.

Usama Fayyad

Bibliography. R. Brachman et al., Mining business databases, *Commun. ACM*, 39(11):42–48, November 1996; U. Fayyad et al. (eds.), *Advances in Knowledge Discovery and Data Mining*, MIT Press, 1996; U. Fayyad, D. Haussler, and P. Stolorz, Mining scientific data, *Commun. ACM*, 39(11):51–57, November 1996; H. Mannila, H. Toivonen, and A. I. Verkamo, Discovery of frequent episodes in event sequence, *Data Mining and Knowledge Discovery*, vol. 1, no. 3, 1997.

Data reduction

The techniques used to transform massive datasets into a more suitable, or smaller, form for data analysis, while preserving the intrinsic characteristics of the data and minimizing the loss of accuracy. Data reduction is used for scientific computation, statistical analysis, control of industrial processes, and business applications, as well as in data mining. Many data reduction techniques focus on obtaining an approximate representation of the data, others on reducing the size of the original data. *See* DATA MINING; DATABASE MANAGEMENT SYSTEM.

Data reduction includes both parametric and nonparametric techniques. Parametric techniques assume a model for the data and attempt to estimate the model parameters that produce a best fit of the data, while nonparametric techniques represent, or categorize, the data without making any assumptions on the data model. *See* STATISTICS.

Dimensionality reduction techniques. Dimensionality reduction techniques concentrate on reducing the number of dimensions (attributes) of the data, while preserving the original information as much as possible. One class of methods for dimension reduction seeks to project the data into a smaller subspace than the original data space. The most widely used method of this class is principal component analysis (PCA). Given a set of data points or a vector in a d-dimensional data space, PCA finds a set of orthonormal vectors that point in the direction of the variability of the data known as principal components. A subset of the principal components that represents the largest variability of the data constitutes a base in a subspace of lower dimension. The projection of a point into that lower-dimensional subspace is an approximation of the point with a minimum loss of accuracy. There are several methods to compute the principal components; for instance, single-value decomposition is one of the most popular methods.

Feature selection is another class of methods that achieves dimensionality reduction by attempting to determine the relevant attributes of the data without applying any transformation. These methods use a heuristic approach to determine a minimum set of relevant attributes in the data according to certain evaluation criterion. Feature selection techniques are classified as supervised whenever class information is available for the data or unsupervised otherwise. Some evaluation criteria for unsupervised feature selection include maximum information compression index and correlation between attributes. Dependency measures and mutual information are examples of evaluation criteria used in supervised feature selection. One simple approach used in supervised feature selection is variable ranking. In this approach, a score is assigned to each attribute to rank the relevance of the attribute with respect to the class distribution of the data. Those attributes that show the highest relevance are selected. Two other main strategies used by both supervised and unsupervised feature selection methods are stepwise forward selection and stepwise backward selection. Stepwise forward selection methods start with an empty set of attributes and build the set of relevant attributes incrementally. At each step, the methods add the best attribute according to the evaluation criterion. Conversely, stepwise backward selection methods begin with the set of all attributes, and at each step remove the worst attribute according to the evaluation criterion. A hybrid method combines both strategies by adding the best of the remaining attributes and removing the worst of the current selected relevant attributes.

Numerosity reduction. There are two main approaches for numerosity reduction; one focuses on estimating a representation of the data by fitting the data to some model, while the other concentrates on finding a smaller representation of the data. Some methods for numerosity reduction, such as regression, are parametric, while others, such as sampling, clustering, and histograms, are nonparametric.

Regression models. Parametric models can be used to generate an appropriate approximation of the data. In this context, data reduction is achieved by storing the parameters of the model that best fit the data. Among several parametric techniques, regression models represent the data as a function of a set of independent variables. The simplest regression model is known as multiple linear regression. In this model, a dependent target variable Y is modeled as a linear function of several independent variables $X_1, X_2, \ldots, X_n$ as $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$. The parameters $\beta_0, \beta_1, \ldots, \beta_n$ are estimated using a least-squares fitting procedure. Also, regression models can use nonlinear functions of the dependent variables. After a proper redefinition of the variables, several nonlinear regression models can be transformed into a linear regression model. However,

there are nonlinear regression models that cannot be transformed into linear models. Although, the least-squares estimate of the parameters of some of these nonlinear regression models can be computed, many of them are intractable.

Histograms. Histograms are an effective nonparametric technique to represent an approximation of the data distribution by partitioning the data into "buckets." A bucket is a multidimensional cube, with each dimension consisting of a continuous range of values that are grouped (binned) at equal intervals. Typically, the frequencies of the data values contained in a bucket are considered to be the same and are computed by averaging the frequency of the binned data values in the bucket. The buckets are determined by a partitioning rule. Several types of histograms have been proposed according to different partitioning rules such as equiwidth, equidepth, and v-optimal. Equidepth histograms partition the data into a predefined number of buckets, such that all the buckets have the same width. An equidepth histogram splits the data into a number of buckets in such a way that the frequencies of the data values are the same in each bucket, while a v-optimal histogram partitions the data into a fixed number of buckets to minimize the variance of frequencies.

Sampling. Sampling is another nonparametric technique that reduces the size of the original data by selecting a representative sample of the data. There are several sampling techniques: simple random sample without replacement (SRWOR), simple random sample with replacement (SRWR), stratified sampling, and cluster sampling. SRWOR draws a sample from the data, where each data point has an equal chance of being selected but can be selected only once. SRWR is similar to SRWOR except that a selected data point is placed back before the next data point in the sample is drawn; therefore, a data point can be selected more than one time. Stratified sampling partitions the data into mutually disjoint groups known as "strata." A sample of each strata is extracted using simple random sampling. Typically, the size of the sample for each strata should be proportional to the size of the strata. In cluster sampling, the data are grouped into disjoint groups or "clusters." Some of the clusters are selected using simple random sampling; the resulting sample includes all the data points belonging to the selected clusters.

Clustering. Clustering is a nonparametric technique that attempts to identify similar groups in the data according to certain criterion. In the context of data reduction, a representative of each cluster is selected to represent the data points in the cluster. Two types of clustering methods are hierarchical clustering, which can be agglomerative or divisive, and partitioning clustering. Agglomerative clustering methods build a hierarchy in a bottom-up fashion. They begin by considering each data point as a cluster. At each step, the methods select two clusters to merge according to the evaluation criterion. In contrast, divisive clustering methods start with one cluster that contains all the data points. At each step, it splits a selected cluster in two clusters. The process continues until no further splitting is possible. Partitioning clustering methods try to find the best partition of the data in a predefined number of clusters.

Julia Couto

Bibliography. D. Barbará et al., The New Jersey Data Reduction Report, *IEEE Data Eng. Bull.*, 20(4), 1997; J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kauffman, 2001; T. Hastie et al., *The Elements of Statistical Learning*, Springer, 2001; H. Liu, H. Motoda, and L. Yu, *Feature Extraction, Selection, and Construction, The Handbook of Data Mining*, pp. 409–423, Lawrence Erlbaum Associates, Inc., 2003.

Data structure

A means of storing a collection of data. Computer science is in part the study of methods for effectively using a computer to solve problems, or in other words, the entire process of determining exactly the problem to be solved. This process entails (1) gaining an understanding of the problem; (2) translating vague descriptions, goals, and contradictory requests, and often unstated desires, into a precisely formulated conceptual solution; and (3) implementing the solution with a computer program. This solution typically consists of two parts: algorithms and data structures.

Relation to algorithms. An algorithm is a concise specification of a method for solving a problem. A data structure can be viewed as consisting of a set of algorithms for performing operations on the data it stores. Thus algorithms are part of what constitutes a data structure. In constructing a solution to a problem, a data structure must be chosen that allows the data to be operated upon easily in the manner required by the algorithm. Because there are different ways to organize data and devise algorithms, it is important to develop criteria for choosing among them. Thus attention is also given to the analysis of the behavior of algorithms under various conditions.

Data may be arranged and managed at many levels, and the variability in algorithm design generally arises in the manner in which the data for the program are stored, that is (1) how data are arranged in relation to each other, (2) which data are calculated as needed, (3) which data are kept in memory, and (4) which data are kept in files, and the arrangement of the files. An algorithm may need to put new data into an existing collection of data, remove data from a collection, or query a collection of data for a specific purpose. *See* ALGORITHM.

Abstract data types. Each data structure can be developed around the concept of an abstract data type that defines both data organization and data handling operations. Data abstraction is a tool that allows each data structure to be developed in relative isolation from the rest of the solution. The study of data structure is organized around a collection of abstract data types that includes lists, trees, sets, graphs, and dictionaries. *See* ABSTRACT DATA TYPE.

Primitive and nonprimitive structures. Computers must have data before they can perform operations on them. Ultimately, data must be processed and organized in a manner that facilitates processing.

Data can be structured at the most primitive level, where they are directly operated upon by machine-level instructions. At this level, data may be character or numeric, and numeric data may consist of integers or real numbers.

Nonprimitive data structures can be classified as arrays, lists, and files. An array is an ordered set which contains a fixed number of objects. No deletions or insertions are performed on arrays. At best, elements may be changed. A list, by contrast, is an ordered set consisting of a variable number of elements to which insertions and deletions can be made, and on which other operations can be performed. When a list displays the relationship of adjacency between elements, it is said to be linear; otherwise it is said to be nonlinear. A file is typically a large list that is stored in the external memory of a computer. Additionally, a file may be used as a repository for list items (records) that are accessed infrequently.

Internal and host-language structures. In general, applications use representations of data based on internal data types defined by the computer architecture. For example, "char" is a single byte that is capable of holding or storing one character in the local character set; "int" is an integer, typically reflecting the natural size of integers on the host machine, 16 or 32 bits. Other applications work with host-language data types which are predefined by the language in which they are written. *See* COMPUTER PROGRAMMING; PROGRAMMING LANGUAGES.

Linear data structures. An array is a means of allocating space in a computer memory for data. Array allocation is static because once it is done it never changes. Static memory allocation works well when the number of data items to be operated upon, and thus the amount of physical storage that will be required for the task, is known. Dynamic allocation allows the number of data items (elements) to grow. Memory is allocated as needed, by request from the application, provided it is still available. When memory or storage is requested, the appropriate size (block) and type required by the data items will be set aside.

An array may be envisioned geometrically as a number of adjacent cells of equal size. **Figure 1** shows a linear array of integers with the name thisarray. Each element of the array has 32 bits or 2 bytes of data and the first element of the array is referred to as thisarray[1], the second element as thisarray[2], and so on. The *k*th element is thisarray[k]. In this example, the location of the first element is loc, the second element is at loc + 2, and so on. Thus the second element is 2 bytes removed from the first.

In the most general form of a linear list, deletion and insertion of an element can be made at any position in the list. A subclass of lists permits insertion or deletion to occur only at one end (**Fig. 2**). The last information in such a list is processed first, that is, on a last-in first-out (LIFO) or a first-come last-served

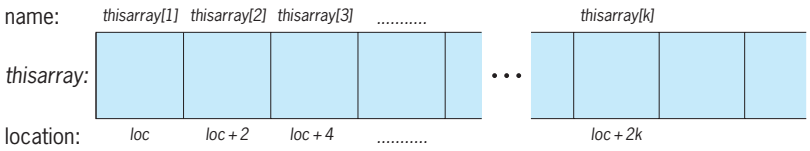

Fig. 1. Linear array of integers.

(FCLS) basis. A linear list belonging to this class is called a stack. The insertion operation is known as push, and the deletion operation is known as pop. A stack can be envisioned as a pile of trays. A tray is removed only from the top of the stack (pop) and returned only to the top of the stack (push). The most and least accessible elements are the top and bottom of the stack respectively, and the last tray in is the first taken out.

Another subclass of lists permits insertion at one end and deletion at the other (**Fig. 3**). The information in such a list is processed in the same order it was received, that is, on a first-in first-out (FIFO) or first-come first-served (FCFS) basis. This type of list is referred to as a queue. Each deletion removes the oldest remaining item in the list or structure. A line of patrons at a ticket window serves to illustrate a queue. Entry is made at the rear, exit at the front, and the person at the head of the list is served first.

Variations of this structure exist. A double-ended queue deque is a linear list in which insertions and deletions are permitted from either end of the structure. There are also two variations of the deque: the input-restricted deque, which permits insertions at only one end, and the output-restricted deque, which permits deletions from only one end.

These structures can be described by using the adjacent-allocation method of storage, which works well in situations when the storage requirement of the application is known. There are applications, however, for which the storage requirements are

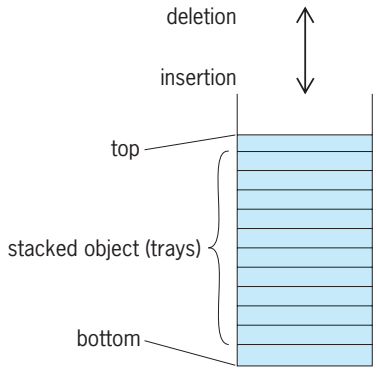

Fig. 2. A stack, a linear list in which elements are inserted or deleted only at one end.

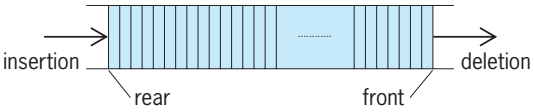

Fig. 3. A queue, a linear list that permits insertion at one end and deletion at the other.

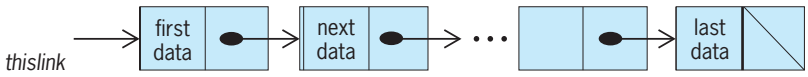

Fig. 4. Singly linked list or one-way chain.

unpredictable and cannot be determined at the time that the program is written; and those for which extensive data manipulations or operations such as insertions and deletions of data are required. In these cases the linked-allocation method of storage can result in the efficient use of both computer storage and time.

Linked allocation. The use of links to refer to elements of a linearly ordered data structure means that elements that are logically adjacent because of the linear ordering may not necessarily be adjacent in computer memory. These types of allocation are referred to as linked allocation. Each element in the list is represented by a node. In a singly linked linear list or a one-way chain (**Fig. 4**), each node of the list references the subsequent node, and thus, starting with the first node in the list, other nodes can be processed by repeatedly traversing the link or reference to the next node. Below thislink in Fig. 4 is a pointer which indicates the first node of the list. This first node contains data for the first element. The purpose of its link is to specify the next node in the linear ordering. The last node indicates the end of the list, and therefore its link address is NULL, which is indicated by a slash.

It is possible for a list to have no nodes at all. Such a list is referred to as an empty list, and the value of thislink in the example would be NULL.

Nonlinear data structures. Nonlinear data structures are capable of representing more complex relationships than that of physical adjacency. One important nonlinear data structure is a tree. A tree is a type of graph.

Directed graphs. A directed graph **G** can be defined as a pair **G** = (**V**,**E**), where **V** is a finite set of points called the nodes or the vertices of the graph, and **E** is a set of ordered pairs of **V**, called the edges of the graph. Thus a directed graph can be represented by a diagram (**Fig. 5**) in which the vertices are shown as points and the edges are shown as oriented line segments corresponding to the pairs of vertices. For example, if (u, v) is an edge, then in a directed graph the line segment uv is drawn with an arrowhead at the point v. Figure 5 shows the directed graph with the vertices given by Eq. (1) and the edges given by Eq. (2). *See* GRAPH THEORY.

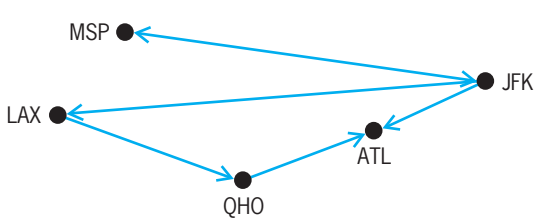

Fig. 5. Diagram representing a directed graph.

$$\mathbf{V} = \{\text{MSP, JFK, LAX, ATL, QHO}\} \tag{1}$$

$$\mathbf{E} = \{(\text{MSP,JFK}), (\text{JFK,MSP}), (\text{JFK,LAX}), (\text{JFK,ATL}), (\text{LAX,QHO}), (\text{QHO,ATL})\} \tag{2}$$

Paths. A path **P** is said to exist from a node u to another node v if there exist nodes $n_1, n_2, \ldots, n_k$ such that the pairs (u, n_1), (n_1, n_2), $(n_2, n_3), \ldots,$ (n_{k-1}, n_k), (n_k, v) are all edges of the directed graph. The collection of edges is called a path of the graph from node u to node v. Thus, in Fig. 5, there is a path **P**(LAX,ATL) from Los Angeles (LAX) to Atlanta (ATL) through Houston (QHO) given by the edges {(LAX,QHO), (QHO,ATL)}.

In the case of directed edges, the two possible edges between a pair of nodes which are opposite in direction are considered distinct. Thus like a two-way street, there is a connection between New York (JFK) and LAX and from LAX to JFK.

Trees. A directed graph is said to be a tree if (1) there exists a distinguished node R called the root of the tree that has no edge going into it, and (2) for every other node u in the graph, there exists exactly one path from the root R to the node u. In **Fig. 6**, for example, the path from R to u_6 is given by Eq. (3). Trees are useful for describing any structure

$$\begin{aligned}\mathbf{P}(R, u_6) &= R \rightarrow u_1 \rightarrow u_2 \rightarrow u_6 \\ &= \{(R, u_1), (u_1, u_2), (u_2, u_6)\}\end{aligned} \tag{3}$$

which involves hierarchy. Thus u_2, u_3, and u_4 are children or immediate successors of node u_1. Every node which is reachable from a node, n say, is called a successor of n. Also, u_5 is a successor of u_1, and u_1 is the immediate predecessor or parent of u_2, u_3, and u_4. Any node with no successor is called a leaf or terminal node. Thus nodes u_3, u_4, u_5, u_6, 3, and 4 are all leaves.

If the root R and the edges connecting it to the next level are deleted (removed) from the graph, a set of disjoint trees is obtained, rooted in the nodes which were the children of R. These trees are subtrees of the original configuration. A set of disjoint trees is a forest.

Although the nodes in a general tree may contain any number of pointers to other tree nodes, a large number of data structure applications involve binary trees, in which each node can have at most two pointers (successors) to other tree nodes. Various storage representations for trees are based on the linked allocation technique. In one linked representation of binary trees, there are two pointers from each parent node to its left and right subtrees, respectively (**Fig. 7**). An empty subtree is represented by a pointer value of NULL.

File structures. Not all information that is processed by a computer necessarily resides in immediately accessible memory because some programs and their data cannot fit into the main memory of the computer. Large volumes of data or records and archival data are commonly stored in external memory as entities called files. Any storage other than

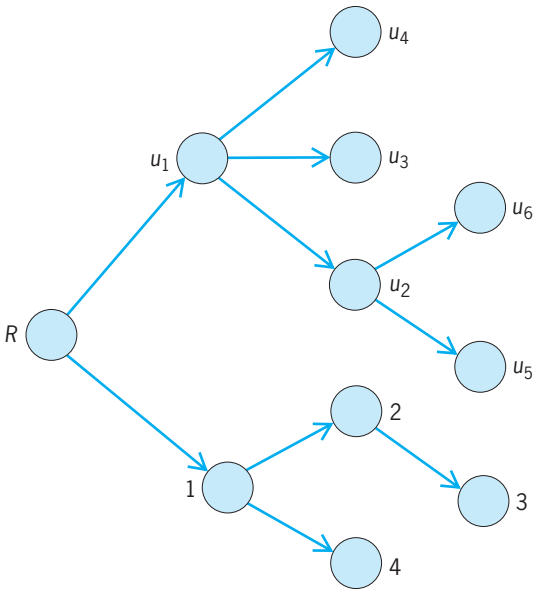

Fig. 6. Example of a tree.

main memory may be loosely defined as external storage. This includes tapes, disks, and so forth. A record item that uniquely identifies a record in a file is called a key. The most common types of file organizations are sequential, indexed sequential, and direct. Access of components of a file can be sequential or direct (random) access. *See* COMPUTER STORAGE TECHNOLOGY.

In a sequential file, records are stored one after the other on the storage device. In sequential access the values can be accessed only in the order in which they are stored in the file, while in random access the values can be accessed as the program desires. The operations performed on a sequential file may differ depending on the storage device used. A file on a magnetic tape, for example, can be opened for either input or output but not both at the same time.

Sequential processing is the accessing of records, one after another, in ascending order by a key or index item of the record. Sequential processing and serial processing are most effective when a high percentage of the records in a file must be processed; since every record in the file must be scanned, a relatively large number of transactions could be batched together for processing. Indexed sequential access combines a sequential key organization with indexed access. A disk can be considered to be divided up into a number of pages, which are contiguous blocks of information that can be accessed by the disk hardware, and each page generally holds many records. To vastly improve the speed of access, an index is maintained that specifies which keys belong to which pages on the disk.

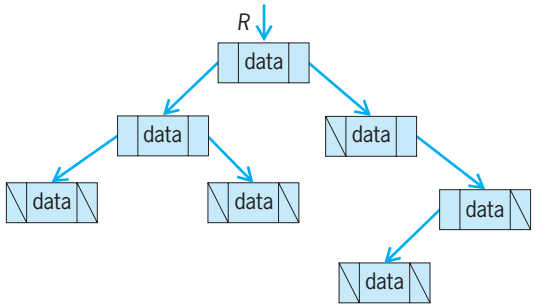

Fig. 7. Linked storage representation for a binary tree.

In a direct (random) access file, a transformation is required from the key of a record to the address of the storage location in which the record is resident. The transformation or mapping is referred to as a hashing technique and is made up of (1) a hash function, which defines a map from the key space to the address space, and (2) a collision-resolution technique to resolve conflicts when two or more record keys are mapped to the same location in a table.

Virtual memory. This is a system that provides an extension to main memory in a logical sense. Most modern computer systems provide a large virtual memory capability. In a virtual system, all currently active programs and data are allocated space or virtual addresses in virtual memory. The program and data may not in fact reside in main memory but in an external storage. References to virtual addresses are translated dynamically by the operating system into real addresses in main memory. This strategy relies on the fact that many programs have relatively small locality of reference; that is, each reference of a program to memory is likely to be in an area of memory that is relatively close to the other recently referenced areas, and thus transfer from external storage to main memory is infrequent. *See* DIGITAL COMPUTER.

Alade O. Tokuta

Bibliography. P. Helman and R. Veroff, *Intermediate Problem Solving and Data Structures: Walls and Mirrors*, 2d ed., Benjamin/Cummings, 1991; B. W. Kernighan and D. M. Ritchie, *The C Programming Language*, 2d ed., 1988; R. Sedgewick, *Algorithms*, Addison-Wesley, 1983; R. Sedgewick, *Algorithms in C, Parts 1–4: Fundamentals, Data Structures, Sorting, Searching*, 3d ed., Addison-Wesley, 1998; R. Sedgewick, *Algorithms in C++, Parts 1–4: Fundamentals, Data Structures, Sorting, Searching*, 3d ed., Addison-Wesley, 1999; J.-P. Tremblay and P. Sorenson, *An Introduction to Data Structures with Applications*, McGraw-Hill, 1984.

Database management system

The software that manages and controls access to data in a database. In a database management system (DBMS), the mechanisms for defining its structures for storing and retrieving date efficiently are known as data management.

A database is a collection of logically related data with some inherent meaning. In any modern enterprise where data and information are essential assets, databases play an important role. The power of databases comes from their conveniently and efficiently managing large amounts of data. For example, a telephone book can be considered a type of database, where the data concerning the names of

individuals and businesses are related to their corresponding telephone numbers and addresses.

Nowadays, a database generally refers to a computerized database with a collection of related data and software package or DBMS that manages and controls access to the data. A computerized database can be considered as an electronic "filing cabinet" for storing data/information and allowing users to manipulate that data/information on demand. Examples of computerized databases include airline reservation systems, banking systems, company payroll, and employee information systems.

The first commercial databases appeared in the late 1960s. These databases evolved from traditional file systems. Databases are also used in many different scientific or human endeavors such as the Human Genome Project, pharmaceutical research, anthropologic investigations, and nuclear medicine, to name some.

Database components. A computerized database has hardware, software, data, users, and procedural components.

Hardware. The hardware of a database is the actual computer system used for storing, accessing, and processing the data, as well as the different storage media on which the data are stored. In large organizations, the hardware includes a network with one or more servers and many client programs. Clients are programs that interact with the DBMS and may run on personal computers at the user's end. The server is generally the computer where the DBMS resides (**Fig. 1**). In general, this server has a more powerful processor than the client machines because it handles the data-retrieval operations and most of the actual data manipulation. *See* CLIENT-SERVER SYSTEM; COMPUTER SYSTEMS ARCHITECTURE; MICROPROCESSOR.

The secondary storage media where the data reside can accommodate large amounts of data and are nonvolatile; that is, any storage medium that, unlike main memory, can retain information in the absence of power. The most commonly used media are magnetic disks (hard disks) that can store large amounts of data at a lower cost than main memory. To improve their reliability, hard disks may be organized as redundant arrays of independent disks (RAID), where data are stored with some amount of redundancy across several disks. In the event of a disk failure, data can be recovered from the data on the other disks that form the system. Other media commonly used to store data include magnetic tapes, flash memory, optical storage, and DVDs. *See* COMPUTER STORAGE TECHNOLOGY; MAGNETIC RECORDING; OPTICAL RECORDING.

Software. The main software component of a database is its DBMS. This software package is the mediator or intermediate layer that allows users to retrieve and manipulate (create, delete, update, and modify) the data. In general, all operations on the data are performed through the capabilities of the DBMS. The DBMS also controls simultaneous access to the data and maintains its consistency or integrity.

The DBMS runs on a computer under a particular operating system. A DBMS and its clients can reside in a computer where it can be accessed by a single user (single-user systems) or by several users simultaneously (multiuser systems). In a network, the DBMS allows for data-handling programs to reside on the server and clients programs on each desktop. A DBMS also makes use of utilities, such as report writers, application development tools, and other design aids. A database is generally referred to by the name of its DBMS, for example, an Oracle® database. *See* OPERATING SYSTEM; SOFTWARE.

Data. The data component is the heart of any DBMS. There are two kinds of data. First, and most obvious, is the set of related information needed by the organization. The second kind is information about the database itself, known as metadata. This information is usually kept in the part of the DBMS called the data dictionary, or catalog. For every piece

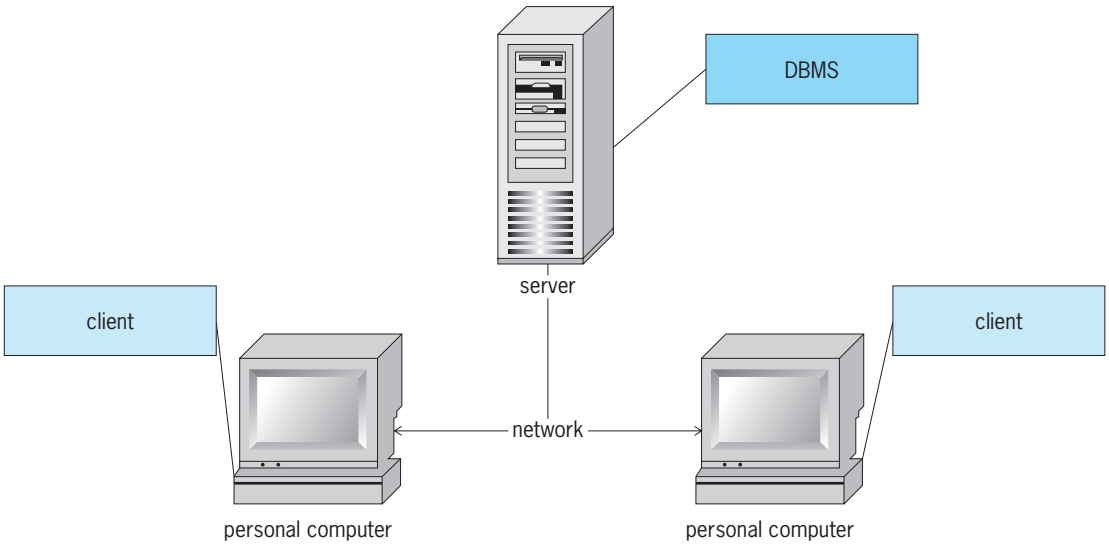

Fig. 1. Components of a database.

of data, the data dictionary holds information such as name, type, size, format, and any possible constraint (the conditions that the data must meet or satisfy). The data dictionary also contains information about all the users and the access type and actions that they are allowed to perform (privileges) in the database. *See* METADATA.

One of the major features offered by a DBMS is that the actual data are separated from the programs that use that data. The set of facts represented in a database is called the universe of discourse (UOD). The UOD is the representation of the "world" that is relevant to the database users. For this reason, databases are designed and built for a particular audience and for a specific purpose.

In a database, the things that are of interest to its users are called entities. The characteristics that describe or qualify an entity are called its attributes. For example, in a database that holds information about students at a particular university, the basic entity is the student. Among the attributes that describe a student entity are the student's first and last names, date of birth, student ID, major, department, and grade point average. There may be some other pieces of information that can be gathered about a student such as eye color, race, and height, but this information may not be relevant for this particular database.

The entities and attributes that describe them are part of the UOD and are stored in the data dictionary. The set of possible values that an attribute can have is called the domain of the attribute. For example, the attribute describing a student's field of study at a particular university may have values such as medicine, architecture, electrical engineering, or computer science. Each attribute of an entity must have a domain.

The entire description of all objects in the database is called the schema. The description of this schema depends on the data model (organizational guidelines) used and is specified during the design phase of the database. The schema should reflect the requirement specification of the database.

The data present in the database at any particular moment in time are called an instance or snapshot. The instance changes whenever new data are inserted or when the existing data are modified or deleted, keeping in mind that the data stored in the database may be required to satisfy certain conditions or constraints. For example, a constraint may be that every student must have a last name, a first name, and middle initial; that is, these attributes cannot be "left blank." For students who do not have a middle name, this attribute may be "filled in" with a predetermined value such as the letter "X." Another example of data constraint may be that the grade point average cannot be negative or exceed a predetermined value, say, 4.00.

Users. Even though databases are used in many fields and for many purposes, there is no such a thing as a general database for everything. Databases are built for specific purposes and for an intended user. For example, it is impossible to use a database built to operate an airline reservation system to manage a payroll system.

In a database, users can access or retrieve data on demand using the applications or the interfaces provided by the DBMS. In general, the users of a database can be categorized as application programmers, end users, or database administrators. Application programmers interact with the database through programs (applications) written in programming languages such as C++, Visual Basic, and Java. Applications generally extend or customize the original functionalities of the DBMS. *See* PROGRAMMING LANGUAGES.

End users are the people at workstations that have access to the data through the facilities provided by the DBMS. They may respond to requests from within or outside an organization. End users can take advantage of some of the interfaces provided by the DBMS to retrieve or update the data stored in the database. Some of these interfaces are menu-driven or forms-driven. Although somewhat different in appearance, these types of interfaces allow the user to select one type of action from a set of choices or to specify certain criteria that need to be met by the data being retrieved. In general, an end user is required to provide some type of input information to complete the action that has been selected. In a database of telephone subscribers, for example, a user can retrieve the address of a particular customer by inputting the telephone number.

Users retrieve data from a database by posing questions through the DBMS. These questions are called queries. Interfaces that allow the end users to query the database in plain English or any other spoken language are called natural language interfaces. This type of interface is still in development due to the inherent difficulties associated with having a computer communicate directly with a person. A somewhat primitive application of this type of database can be found in the services provided by some cellular telephone companies. Instead of dialing a number, users of these services talk on the telephone and mention the name of the person they want to call. Voice recognition software translates this verbal request into a search for a name in a database. *See* SPEECH RECOGNITION.

Another database interface frequently used by cashiers or automatic teller machines (ATM) is the parametric interface. Parametric interfaces, often used for simple and repetitive operations, generally require that the user press buttons on a screen or on special keypads. Each button or key triggers the execution of a procedure designed to perform a particular function.

Procedural components. Database administrators are the users responsible for defining the internal structures to store the data, the policies for manipulating the information, and the security mechanisms to protect the data. Database administrators also serve as liaison between end users and application programmers. The database administrator can be a single individual for small databases or a group of two or more people for large corporations.

An integral part of any database management system is the set of procedures that control the behavior of the system, that is, the actual practices the users follow to obtain, enter, maintain, and retrieve the data. For example, in a payroll system, a procedure may determine how salaries are calculated for the different types of employees that may exist in the organization. Part of the job of the database administrator is to verify that the procedures are clearly defined and comply with the guidelines of the organization.

Size of a database. The size of a database is generally measured by the size of its data. Databases can range from a few megabytes (10^6 bytes) for personal databases to gigabytes (10^9 bytes) or even terabytes (10^{12} bytes) for large corporate databases. Petabyte (10^{15} bytes) and exabyte (10^{18} bytes) databases will be common in the future.

Access. A DBMS provides a programming interface and simultaneous access to the data by more than one user. Applications programmers or data administrators can use the programming interface to manipulate or maintain the database. The DBMS allows concurrent access to data while making sure that the users do not interfere with one another.

Characteristics of the data in a database. In any database, the data are both integrated and shared. Integrated means that the database can be thought of as the unification of several distinct data files with redundancy between these files kept to a minimum. For example, a company database might contain data about its employees in an EMPLOYEE data file and information about the employees' dependents in a DEPENDENT data file. The employees' data might include employee ID, last name, first name, department, and salary. Data about the employees' dependents might include employee ID, dependent ID, dependent last name, dependent first name, relationship to employee, dependent date of birth, and so on. Information about the dependents of a particular employee is obtained by combining information from these two files using the common piece of data that they share (employee ID). The employee ID (or any other unique piece of identification) is necessary in this case to associate an employee with dependents since they are kept in separate files.

Shared data means that different users can use the same data for different purposes. For example, the human resources department might use the date of birth of the dependents to send them birthday cards, whereas the accounting department might use the same piece of data for tax purposes.

Data models. In a database, data have certain physical organization on the storage media on which they are kept. The data also have a logical organization as seen at the user interface of the DBMS. A major purpose of a database is to provide users, through its DBMS, with an abstract representation of the data. That is, users can conceptually picture the data in a particular manner while the system implements and hides all details on how the data are physically stored. For this purpose, databases are based on a data model.

A data model is a set of guidelines for representing the logical organization of data in a database. In other words, a data model aids database users by helping them think of, visualize, and communicate information about the data. A data model consists of named logical units and relationships among the data. One of several data models can be used for a particular database. The main difference between these data models is the manner in which they represent certain relationships among the data. Most data models handle attribute relationships in a similar manner but handle relationships (associations) between the entities in different ways.

Data models can be categorized according to the concepts that they use to describe the database. High-level or conceptual data models provide concepts that are close to the way users perceive the data, whereas low-level or physical data models provide concepts that describe the details of how the data are stored in the computer. In relational databases, for instance, all the data are viewed by the users as tables, while the data can be stored in a variety of complex data structures that permit efficient retrieval.

In addition, there is a class of implementation data models that use concepts understood by the end users with the data organized similarly within the computer. Implementation models are the most widely used in traditional commercial databases. These data models include the widely used relational data model and some of the legacy (older) models such as network and hierarchical models. However, the complexity of database applications has increased. A new family of higher-implementation data models has been proposed to address issues such as new data types for storing images, more complex structures for user-defined objects, and the need to define nonstandard application-specific operations. Another reason for new data models is to ensure their integration with object-oriented programming languages. The need for additional data modeling features has been recognized by some of the relational DBMS vendors who now offer extended or object-relational DBMSs. *See* DATA STRUCTURE; OBJECT-ORIENTED PROGRAMMING.

Three-level architecture. To facilitate the differentiation of the conceptual views of the data from the physical implementation, the American National Standard Institute (ANSI) has defined a general framework known as the three-level architecture (**Fig. 2**). Each level of this architecture, described by its own schema, serves a particular purpose. The internal level or physical level is concerned with the way the data are physically stored and accessed. The external level or user logical level is concerned with the way the data are seen by individual users. The conceptual level or community view describes what data are stored in the database and the relationships among the data. This architecture is a framework for describing the data and it is not the actual data. The data are physically stored in the computer system.

The three-level architecture can also be used to explain the concept of data independence—the ability

to change the schema at any level of the database without having to change the schema at the next higher level. There are two types of data independence: logical and physical. Logical data independence allows changing the conceptual schema without having to change the external schemas or applications programs. The conceptual schema may change by adding or deleting constraints to the existing data. Physical data independence allows changing the internal schema without having to change the conceptual schema. Changes to the internal schema may be the result of the reorganization of some of the physical files for storing the data.

Most DBMSs do not make explicit separation or support the three-level architecture in its entirety. In general, those DBMSs that support user views may support external schemas through the same data model that describes the conceptual level. The process of transforming requests and results between levels is called mapping. All information about the schemas and mapping between them is stored in the data dictionary. Whenever the schema at a lower level is changed but not the one at the next level, the notion of data independence is achieved by only changing the mapping between these two levels.

Data languages. Database designers and administrators use data languages to specify conceptual and internal schemas, as well as any mapping. A data language is a set of operations that permit access to the data as organized by a data model. A complete programming language that exists only to support the data language is called a self-contained data language. For example, DBMSs organized according to the relational data model use the self-contained data language known as the Structured Query Language (SQL). English-like, self-contained data languages are sometimes called query languages. This type of language allows the user to query the database using commands that follow a simple but somewhat rigid syntax. SQL is the de facto industry standard query language.

The subset of commands of the data language that allows the definition of the conceptual schema is called a data definition language (DDL). Another subset of the language that allows the definition of internal schema is called the storage definition language (SDL). Within the data language, the set of commands that allow the retrieval, insertion, deletion, and modification of the data is known as the data manipulation language (DML). If DML commands are used within an application program written in a high-level programming language, such as Java or C++, this programming language is referred to as the host language and the DML is called a data sublanguage (DSL).

Database classification. DBMSs can be classified according to criteria such as single-user or multiuser systems, as well as on the basis of the data model that they support. In addition, DBMSs can be classified according to where the DBMS and data reside. When both reside on the same computer system, the database is said to be centralized. If the DBMS and the data are distributed over many sites connected by a computer network, the database is said to be distributed. A distributed database is called homogeneous if all the DBMSs at the sites are the same; otherwise, the database is called heterogeneous. A federated database is a group of autonomous and independent databases that are connected through a network. A database can also be classified as special purpose if, for performance reasons, it is built for a specific application such as a medical registration database or telephone directory system. *See* DISTRIBUTED SYSTEMS (COMPUTERS); LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

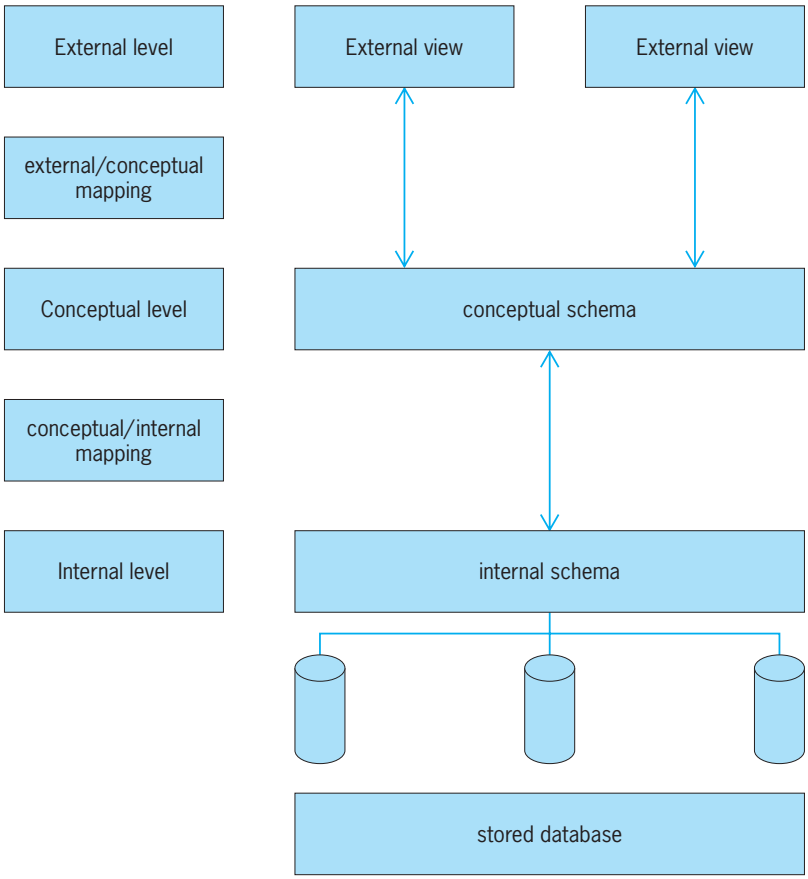

Fig. 2. Three-level architecture. (*Elmasri and Navathe*)

Client-server architecture. In client-server or two-tier architecture, client programs communicate with the DBMS across a network. The client program resides at the user's computers and the DBMS resides at the server computer. The client program can execute independently on its host computers. When a client program needs to communicate with the DBMS, it does so by using a special type of software that lets the client and the DBMS "talk" to each other. This special software has at least two components: one that runs on the client machine and another that runs on the server. These software components follow a predefined standard that defines an application program interface (API). An API consists of a set of library functions that implement many of the common operations that programmers require. Some of these operations include connecting to the database, making requests to the server, and obtaining results.

The industry standard API, based on the SQL language and developed by Microsoft, is a technology known as Open Database Connectivity (ODBC). The ODBC standard provides a common interface for accessing heterogeneous SQL databases. Through this interface, a single application can access data in different databases with a common set of codes. A specialized piece of software often used to facilitate this process is known as a driver. Drivers are provided by the different DBMS vendors and work exclusively with the DBMS for which they were created. A similar standard, called Java Database Connectivity (JDBC), has been created for accessing relational DBMS using Java applications. JDBC defines an API that supports basic SQL functionality and provides access to a large variety of relational DBMS products.

Many Web applications use a three-tier architecture where an additional layer of software is added between the clients and the server. This new layer, depending on the type of application, is called the application server or Web server. The Web server (which runs on its own computer) is the depository of procedures and constraints that have been defined to access the data stored at the database server. The Web server accepts the requests from the client, processes the request, and sends appropriate commands to the DBMS. The Web server receives the results from the DBMS and passes them to the client for further processing before being presented to the user. The Web server can also be used to secure the database. It can do this by checking the validity of the user's requests before sending them to the server or by using encryption/decryption mechanisms to protect the data. *See* COMPUTER SECURITY; WORLD WIDE WEB.

Relational model. The relational database management system has become the most popular of all the DBMSs. It was initially proposed by E. F. Codd in the early 1970s. In this model, all data are logically structured as relations or two-dimensional tables. Each entity of the universe of discourse corresponds to a table; the attributes of this entity form the columns of the table (**Fig. 3**). Each row or "tuple" of the table represents an instance of the entity. The values under each column must belong to the domain of that attribute but, in some cases, these values may be absent or "null." The table name and the column names help in interpreting the meaning of the values of each row. The total number of columns of the table is called its degree. The total number of rows in the table is called its cardinality.

The relational model is based on the mathematical concept of relation and uses some of the elements of predicate logic and set theory. However, some of these definitions, when compared to their mathematical counterparts, have been relaxed to accommodate the logical information conveyed by the table. *See* LOGIC; SET THEORY.

Any relation must have a set of one or more attributes, called the primary key, that uniquely identify each row. A foreign key is a column in a table that matches the primary key column in some other table. Key values in any table must be unique and cannot be null. When representing a relation as a two-dimensional table, the attributes that comprise the key are generally underlined.

Keys, with foreign keys, can also be used to maintain the consistency of the data among the tuples of two relations. In Fig. 3, the EMP_DEPT (employee department) attribute of the EMPLOYEE table has been defined as a foreign key for this table. This EMP_DEPT attribute references the primary key (ID) of the DEPARTMENT table. What this implies is that for any employee inserted into the EMPLOYEE table, the values under the department column must be one of the values listed in the key attribute of the DEPARTMENT table. The foreign key concept allows the EMP_DEPT to be NULL. However, if this NULL value is ever changed it has to be to one of the values listed in the ID column of the DEPARTMENT table.

A basic set of table operations, known as relational algebra, allow the user to specify basic retrieval requests. A SELECT operation extracts a subset of table rows that satisfy a given condition. A PROJECTION operation extracts some table columns. A JOIN operation combines related rows from two or more tables (not necessarily different) into single rows. Other operations on tables are the INTERSECTION, UNION, and DIFFERENCE. These table operations behave just as their counterpart set operations.

If for security reasons there is a need to hide the identity of the tables that participate in a query, users can create virtual tables or views. A view is a relation that can be produced on demand but does not necessarily exist in the database. Technically, views are queries stored in the data dictionary that users can "run." Views are a basic security mechanism.

Emerging technologies. Relational DBMSs are the most popular of all DBMSs. Estimated yearly sales of RDBMs and related products are around U.S. $20 billion. In the foreseeable future, relational DBMS will continue being the dominant DBMS. However, there are some new advances in database technology. Among the most promising technologies are mobile, multimedia, and geographical information systems. Mobile databases comprise portable computing devices, which, coupled with wireless communication, allow users to access data from virtually anywhere and at any time. Multimedia databases store and process data, including text, graphics, animations, video, audio, or any combination of these.

Department

ID	Name	Location
10	Accounting	New York
40	Sales	Miami

Employee

EMP ID	EMP_NAME	EMP_MGR	TITLE	EMP_DEPT
1234	Green		President	40
4567	Gilmore	1234	Senior VP	40
1045	Rose	4567	Director	10
9876	Smith	1045	Accountant	10

Fig. 3. Example table for a relational database.

See MULTIMEDIA TECHNOLOGY.

Geographical information systems (GIS) collect, model, store, and analyze information describing the physical properties of the world. GIS manipulate spatial data (such as maps, images from satellites, roads, weather, and soil) and nonspatial (such as census results, sales or marketing results). *See* GEOGRAPHIC INFORMATION SYSTEMS. Ramon A. Mata-Toledo

Bibliography. C. J. Date, *An Introduction to Database Systems*, 8th ed., Addison-Wesley, 2003; R. Elmasri and S. Navathe, *Fundamentals of Database Systems*, 4th ed., Addison-Wesley, 2003.

Dataflow systems

An alternative to the sequential, or von Neumann–based, model of computing, offering an inherently parallel metaphor. The dataflow concept is characterized by an orientation toward producing and consuming data values rather than updating storage cells. Such systems are made up of dataflow machines which execute programs expressed as a partial ordering of operations. This partial ordering, called a dataflow graph, is created by high-level language compilers. Dataflow systems are capable of exploiting all of the parallelism inherent in an algorithm. The economies of scale of very large scale integration can be exploited quite naturally by using the dataflow approach. *See* ALGORITHM; INTEGRATED CIRCUITS.

Basic concepts. Dataflow systems employ an underlying execution model which differs substantially from the conventional sequential model; specifically, the dataflow model is concerned with naming and manipulating data values as opposed to naming and manipulating the containers for data values, that is, storage cells. This simple yet fundamental change in perspective has a profound impact on issues both at the level of the programming language and at the level of the computing machinery. Moreover, the shift of focus from where and when values should be stored to merely what the values should be has a character which is mathematically more direct and appealing. This value-oriented view is called definitional: each step in the computation defines some entity, such as a number or a data structure, based on other such definitions or on input values. The traditional view is called imperative: each step in the computation not only defines something to be computed but also constrains where the resulting value should be stored and when such storing should take place. Because value containers can be reused within an imperative framework, additional constraints may be imposed implicitly on the time ordering of these steps.

Within the definitional framework, the statement

$$A = B + C$$

simply means that, in the scope of the identifier A, its value is defined as the sum of the values of B and C. It indicates nothing about where this value should be kept, nor even that the value must be computed at all, in the most rigorous sense. References to A within its scope all denote the same value, regardless of the time-ordering of the actual operations required to produce and consume the value. Given this view, statements such as

$$A = A + 1$$

are obviously meaningless as definitions in that they have no time-invariant denotation. This property in a programming language is sometimes called single assignment and is a key element of the dataflow model.

Another aspect of the definitional framework is that all operators behave functionally; the output value they produce is entirely defined by the input values with no reliance on the time at which the operation is performed (state independence). Moreover, operators and their compositions are constrained to be free of side effects.

These properties, when present in a programming language, allow compilers to deduce the essential dependences between operations and to direct dataflow machines to execute instructions accordingly. When the compiler can prove that two operations have no dependence of one upon the other, a dataflow machine can execute them in parallel.

Dataflow graphs. In the traditional model of computing, compiled instructions are represented as a sequential list. In the dataflow model, instructions are not totally ordered in this way but must be expressed as a partial ordering. Conceptually, each instruction can be viewed as a node in a graph. Directed arcs connect any instruction to those subsequent instructions which depend upon its value. Executing a program as a dataflow graph can be viewed as a process of propagating data-carrying tokens along the arcs of the graph. Tokens consist of a value and a tag. The tag serves to route the token from the instruction which produced the value to the instruction which is to consume the value. Tags denote an instruction within a graph, and in some classes of dataflow machines further specify a unique instantiation name which allows multiple invocations of the instruction to coexist. It is the dataflow machine's responsibility to execute, at any given time, all and only those instructions in a graph for which a complete set of identically tagged input tokens exists (so-called enabled instructions). An enabled instruction represents the complete specification of some processing activity: input values and function. Executing an instruction means computing the function denoted by the instruction (for example, *ADD*) by using the values on the input tokens, and fabricating a result token containing the new value and a new tag. This new tag denotes the instruction, called the destination, which is to receive the result. To the extent that there are many such enabled instructions at any given time, the machine is free to execute them in any temporal order whatever or, most interestingly, to execute them all in parallel. This execution rule is in contrast with the traditional method based on program counters.

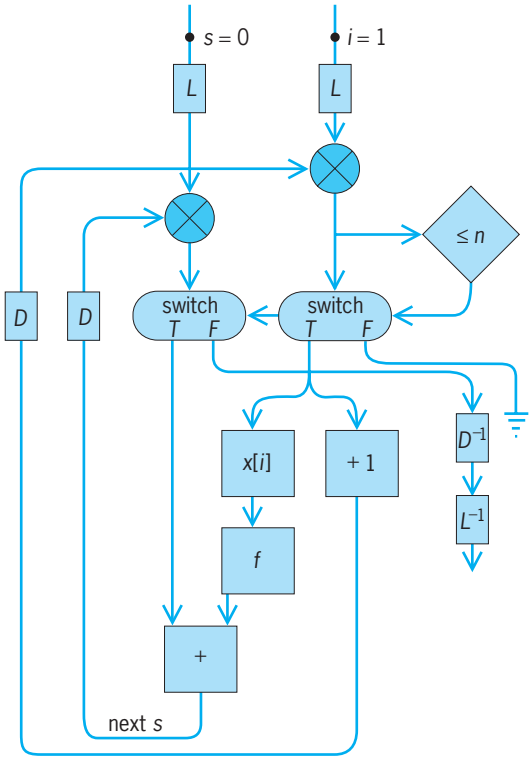

Fig. 1. Dataflow graph for compilation of the loop expression for $\sum_1^n f(x_i)$. (*After Arvind and R. A. Iannucci, Two Fundamental Issues in Multiprocessing: The Data Flow Solution, Computation Structures Group Memo 226-3, MIT Laboratory for Computer Science, 1985*)

Figure 1 is an example dataflow graph, compiled from a program fragment(shown below) in Id, a high-

```
{ s = 0
  in
  {for i from 1 to n do
   next s = s + f x[i]
   finally s }}
```

level language designed specifically for data-flow systems. The fragment applies the function f to successive elements of a one-dimensional array x and sums the results. The graph is somewhat stylized: the box marked f represents the subgraph necessary for computing function f. Instructions D, D^{-1}, L, and L^{-1} are tag-manipulation instructions which manage dynamic instantiation information on the tokens; they act as identity instructions for values. The *SWITCH* instruction routes a data token (top input) to either the T (true) or F (false) output, depending upon the value of a boolean token (right input). The remainder of the operators are arithmetic and relational instructions whose function should be self-evident. (The circled X represents the merging of two arcs and does not represent an instruction.)

Architectures. Clearly, a program could be designed to interpret dataflow graphs on a conventional architecture with the attendant inefficiencies that come with interpretation. Dataflow machines, on the other hand, compute by using encoded dataflow graphs as their machine language. Such a machine is shown in **Fig. 2**. As a token arrives at the waiting-matching unit, its tag is compared to those of the other tokens which are already waiting. Conceptually, this can be done with associative matching, although hash-based and direct-mapped techniques are used in practice. If no match is found, the token joins the waiting pool. If a match is found, both tokens are extracted and fed, along with their common tag, to the instruction fetch unit. Here, the tag is used to look up the correct instruction from the program memory. The values on the two tokens, along with the operation code from the instruction, are passed on to the arithmetic-logic unit (ALU), while the tag and the destination information are forwarded to the compute tag section. For most instructions, new tags are derived from old ones simply by substituting destination instruction information. Instructions such as D and L in Fig. 1 cause more sophisticated transformations on the instantiation-specific part of the tag, effecting the allocation and deallocation of instantiation names. The new value from the arithmetic-logic unit is combined with the new tag from the compute tag section, the result being a new token which is passed on to the waiting-matching unit. Parallelism can be exploited by distributing the dataflow graph across multiple processors. The routing network coordinates the task of shuttling tokens to the correct processors, based on their tags.

In this type of architecture, called dynamic dataflow, storage for the intermediate results of the computation, tokens in particular, is allocated at run time. Because of this, dynamic procedure invocation, including recursion, can be supported. For certain classes of computations wherein the intermediate storage issues can be worked out at compile time, a more restricted class of architecture, called static dataflow, is suitable. By preallocating storage and other resources, static machines avoid the need to carry instantiation information in the tags. *See* COMPUTER SYSTEMS ARCHITECTURE.

Suitability for parallel processing. The dataflow model of computation differs significantly from that of conventional systems; definitional semantics displace imperative semantics, and the data-driven notion of instruction scheduling displaces the notion of sequential control flow. It has been argued that the dataflow model is inherently parallel, whereas the conventional model is inherently sequential and that as a result dataflow is a more suitable basis for general-purpose parallel computing. A general-purpose parallel computer is, by its very nature, a collection of processors and memories distributed in space. If a single computation is to be mapped onto this set of processors, the fraction of the total computation (instructions and data) local to any one processor must necessarily decrease as the number of processors is increased. There is a direct relationship between this fact and the increase in latency as the degree of parallel processing increases. Latency is the time that elapses between making a request for a remote value, for example, and receiving the associated response. Because latency is

unavoidable, an important measure of any parallel processing architecture is its ability to tolerate long latencies.

Just as physically partitioning the machine (processors and memories) into n component parts gives rise to this fundamental issue of latency, so also decomposing the program into asynchronously communicating fragments gives rise to the fundamental issue of synchronization. This is not an issue in a conventional single-processor system, because there is a single locus of control. All instruction executions are implicitly synchronized through sequential instruction dispatching. Using such conventional machines in a parallel processor, however, invalidates this assumption. Synchronization may be provided, at some loss of generality, through a globally synchronous parallel processor architecture. A more universal approach necessitates the emulation, in software, of the synchronization means. Emulated synchronization is costly and must be used sparingly. The upshot for parallel processing is that, except for the most trivially parallelizable applications, seeking to minimize synchronization is tantamount to seeking to minimize the exploitation of parallelism. The most general approach, adopted under the dataflow paradigm, is to integrate synchronization features into the lowest levels of the individual processor in such a way that explicit synchronization takes place prior to the invocation of each instruction. This architectural step having been taken, the cost of synchronization is not a first-order programming concern. Providing efficient hardware means for synchronization is, however, a very significant hardware concern. *See* CONCURRENT PROCESSING.

Advantages and limitations. Dataflow systems offer an approach to scalable, general-purpose parallel processing which has been shown to address satisfactorily the fundamental issues of latency and synchronization. The approach relies on the data-driven dispatching of instructions and the manipulation of values, rather than value containers, to decouple the issue of temporal computing behavior (the time-interleaving of parallel activities) from the correctness of the computed result. Compared to conventional systems in which synchronization must be provided in software, dataflow machines provide synchronization which, for better or worse, is inherent in the scheduling of every instruction. This has a profound impact on the economics of exploiting high degrees of parallelism. Dataflow languages obviate the need for explicit parallelism notations in programs and the attendant concerns over correctness.

As in any parallel processing system, conventional or dataflow, the issues of dynamic resource management, such as the allocation and deallocation of storage and processors, must be considered. The exploitation of parallelism in a given application usually requires more resources than the sequential execution of the same application. Moreover, in the dataflow model, the problem of bounding or limiting the instantaneous resource requirements during execution is of great importance. Because instruction executions are partially, and not totally, ordered when parallelism is available, it is possible that resource requirements for executing a given program on a given set of data may vary from run to run. Running a program at one time may result in peak resource demands which exceed the machine's capacity (the computation cannot be completed), whereas running the same program another time may not.

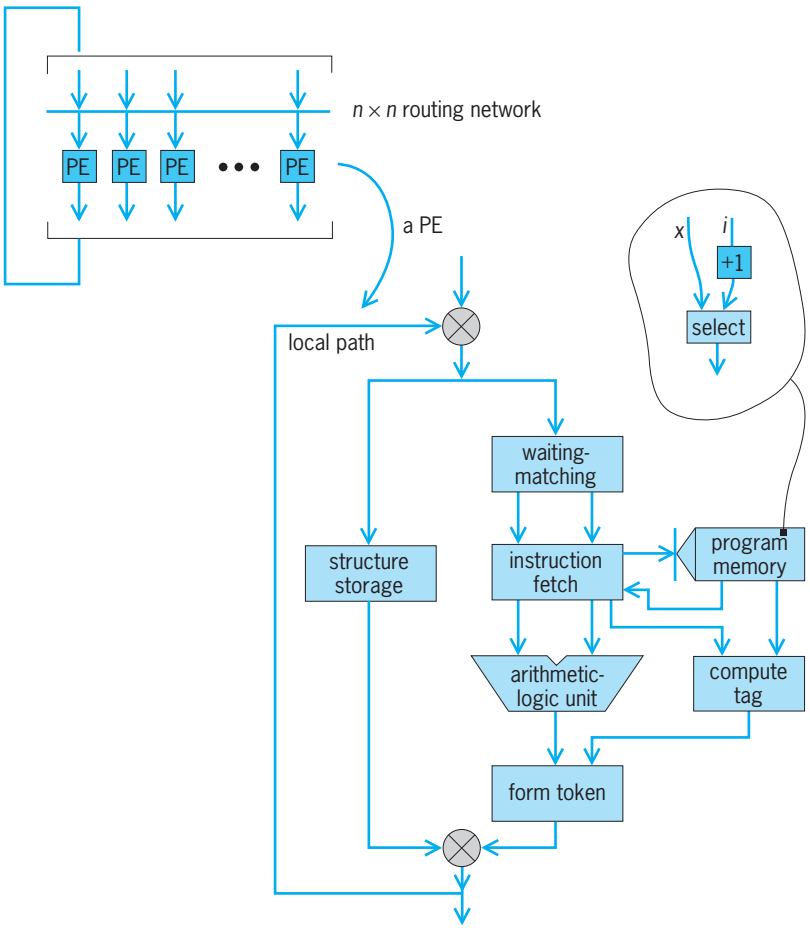

Fig. 2. Dynamic dataflow machine. PE = processing element. ***(After Arvind and R. A. Iannucci, in R. Dierstein, D. Müller-Wichards, and H.-M. Wacker, eds., Parallel Processing in Science and Engineering, Springer-Verlag, 1988)***

Known realizations of dataflow machines suffer from an inability to process efficiently sequential streams of instructions during periods of low-degree parallelism. This is attributable to the data-driven instruction execution rule in the presence of deep pipelines. Solutions to this problem involve combining conventional and dataflow-based instruction dispatching hardware. Such systems have come to be known as hybrid dataflow architectures. *See* COMPUTER PROGRAMMING; PROGRAMMING LANGUAGES.

Robert A. Iannucci

Bibliography. Arvind and D. E. Culler, Data flow architectures, *Annu. Rev. Comput. Sci.*, 1:225–253, 1986; R. Dierstein, D. Müller-Wichards, and H.-M. Wacker (eds.), *Parallel Processing in Science and Engineering*, 1988; J. A. Sharp (ed.), *Data Flow Computing: Theory and Practice*, 1992.

Date

The date palm (*Phoenix dactylifera*) is one of the oldest cultivated tree crops. It provides a staple food for many populations in the Middle East and North Africa, and also is highly valued for feed, fiber, and shelter. Beginning in the 1890s, the U.S. Department of Agriculture (USDA) plant collectors began importing date offshoots from Algeria, Tunisia, Egypt, and Iraq into the Salt River Valley in Arizona and then into the Coachella Valley in California, and the United States date industry was born. Date production and culture continue to challenge farmers around the world.

Botany. The date palm is in the family Palmaceae. *Phoenix* is distinguished from other palm species by the production of offshoots, columnar-circular trunks, pinnate leaves, and distinctively furrowed seeds. Date palms are monocotyledons and have a single bud or growing point. The date trunk also lacks the cambium growth layer typically found in dicotyledon fruit trees, which is important for true secondary growth.

The date palm species is dioecious: male palms produce staminate flowers and female palms produce pistillate flowers. Both male and female flower clusters are enclosed in protective sheaths. Bees and insects go to the male flowers but not to the female blooms. The date palm is dependent on humans for pollination. Pollen from the staminate flowers must be collected and manually applied to the pistallate inflorescence every season for commercial date production. Windblown pollination results in poor fruit set and in fruit that do not mature properly. One male tree can provide enough pollen for 30–40 female trees. Because the inner layer of the fruit wall is fleshy, the date fruit is classified as a berry, like a tomato.

Culture. The date palm grows under a wide range of temperatures, but fruit production requires 6000 to 7000 degree-days (Fahrenheit). Degree-days are the sum of the number of degrees of daily mean temperature above 50°F (10°C) during flowering and fruit ripening. In the Coachella Valley the medium temperature frequently exceeds the 110°F (43.3°C) and has been as high as 123°F (50.6°C). Temperatures below 20°F (−6.6°C) can cause injury to mature date palms; offshoots are injured at temperatures under 25°F (−3.9°C). Other requirements for date production are low rainfall and low humidity because prolonged periods of humidity affect developing fruit and reduce quality.

The date garden (a grower term) is planted with offshoots collected from established palms. Vegetative propagation is necessary because date seedlings do not duplicate the female parent, resulting in nonuniform fruit of inferior quality. Male palms should also be selected from offshoots for superior pollen source. Offshoots are connected to the parent tree by a short, woody base a few inches in diameter. Within the woody base is the terminal bud, it is protected by the lower leaf bases, and care must be taken to protect it from damage during offshoot removal. Offshoot size varies depending on variety, ranging 40–100 lb (18.1–45.4 kg) in weight and 8–14 in. (20.3–35.6 cm) in diameter. Commercial date plantings are spaced 30 by 30 ft (9.1 by 9.1 m) apart to provide sufficient sunlight and ventilation. Dates begin to fruit at around 5 years of age and are in full production at 10–15 years; they can remain productive for about 50 years. The root system is composed of long lateral roots that can extend as far as 20–30 ft (6.1–9.1 m) from the trunk, and secondary smaller roots. Root hairs, which are the main uptake of water and nutrients in most plants, are absent in date palms. Irrigation can be by borders or drip irrigation. Careful attention to irrigation is necessary to maintain palm growth and high fruit yields. Date palms can require up to 10 acre-feet (4 hectares) of water a year. Irrigation can be reduced during the harvest season in order to facilitate the harvest.

Fertilization. A total of 4–6 lb (1.8–2.7 kg) of actual nitrogen per palm is adequate in most soils. However, in light-textured soils where leaching is a concern, more nitrogen may be required. The nitrogen is used in split applications during the growing season. The source of the nitrogen can be from animal manures or winter or summer cover crops. High nitrogen applications may increase yields, but result in lowered quality.

Fruit ripening stages. Fruit thinning of bunches on the trees is necessary to increase fruit size, and it can be accomplished by reducing the number of fruits per bunch (bunch thinning) or by reducing the number of bunches per palm (bunch removal). The amount of quality fruit that a palm can safely carry depends on the age, size, vigor, and variety of the palm and on the number and size of green leaves that it has. Even after bunch thinning and removal, there may still be unequal ripening, within the bunch and with the other bunches on the tree.

Arabic terms are used in referring to the developmental stages of the fruit. "Kimri" refers to very young green-colored dates. The dates increase in size and weight and begin to rapidly accumulate reducing sugars. At this stage, dates begin to reduce in acidity and slightly gain moisture. In the "Khalal" stage the fruit just about reaches its maximum size and turns yellow, pink, red, or scarlet, or yellow spotted with red, depending on the cultivar. The "Rutab" or wet stage follow, and little or no sugar accumulates; the fruit continues to lose water. The fruit begins to soften and merges into the "Tamar" or cured stage; the fruit are fairly firm in consistency and have fully ripened. The sugar-to-water relationship in the fruit now prevents fermentation.

Grading. Each palm bunch must be hand-harvested. Field bins are transported to the packing houses for grading. The dates are cleaned and sorted into uniform lots and classes. The several lots comprise dates that are uniform in color, texture, size, and moisture content. Most packing houses have USDA inspectors during operating hours. Inspection covers adherence to grade standards, which include cultivar characteristics; uniformity of color; and freedom from insects, diseases, and various other defects; and size uniformity for the grade.

Varieties. In 1997 there were about 5640 acres (2282 ha) planted in California's Coachella Valley with an annual production value of $62 million. There are also 600 acres (243 ha) planted in Bard near Yuma, Arizona. The most popular of the many date varieties will be mentioned. The Deglect Noor cultivar is a semidry date and is the most popular variety in the United States; it is originally from Algeria. The Medjool is a soft date, originally from Morocco, that is prized for its extralarge size. The Zahidi is a semidry date from Iraq and has proved to be well adapted to California and Arizona. The Halawy is a soft date from Iraq, and its fruit is light amber to translucent golden brown when ripe. The Barhee, whose fruit is amber to deep golden brown on ripening, is also a soft date from Iraq. The Khadrawy is a soft date that is greenish amber on ripening and reddish brown when cured, and is also originally from Iraq.

Disorders. A common disorder is blacknose, a darkening and shriveling of the tip resulting in severe skin cracking on the fruit. During the growing stages the Banks grass mite causes serious drying and scarring of developing fruit. Fruit beetles can also attack ripe fruit, preferring soft and sour dates. These beetles lay eggs on the fruit, which the larvae feed on, making it unfit for consumption. If dates are not properly stored, fermenting yeasts and mold fungi decrease the quality. Physiological deterioration can also cause darkening and loss of flavor. Temperature, moisture content, and date variety affect this deterioration.

Pitted dates are now packaged for greater consumer acceptance and convenience. United States supermarkets stock pitted, whole, and chopped domestic dates. Because of the wide range of date textures available, packinghouses also produce a wide variety of products for use in the baking and institutional food-service industry. *See* ARECALES; FOOD ENGINEERING; FRUIT; FRUIT, TREE; INFLORESCENCE; REPRODUCTION (PLANT). Jose L. Aguiar

Bibliography. S. Aslan et al., *Soil, Water, and Climactic Considerations in Selecting Date Palm Planting Sites in the Coachella Valley*, Coachella Valley Resource Conservation District, 1991; J. B. Carpenter and H. S. Elmer, *Pests and Diseases of the Date Palm*, U.S. Government Printing Office, Handb. 527, 1978; R. E. Hardenburg, A. E. Watada, and C. Y. Wang, *The Commercial Storage of Fruits, Vegetables, and Florist and Nursery Stocks*, U.S. GPO; Handb. 66, 1986; W. Nixon and J. B. Carpenter, *Growing Dates in the United States*, U.S. GPO, Bull. 207, 1978; G. L. Rygg, *Date Development, Handling and Packing in the United States*, U.S. GPO, Handb. 482, 1975.

Dating methods

Relative and quantitative techniques used to arrange events in time and to determine the numerical age of events in history, geology, paleontology, archeology, paleoanthropology, and astronomy. Relative techniques allow the order of events to be determined, whereas quantitative techniques allow numerical estimates of the ages of the events. Most numerical techniques are based on decay of naturally occurring radioactive nuclides, but a few are based on chemical changes through time, and others are based on variations in the Earth's orbit. Once calibrated, some relative techniques also allow numerical estimates of age. *See* ARCHEOLOGY; ASTRONOMY; GEOLOGY; RADIOISOTOPE.

Relative Dating

Relative dating methods rely on understanding the way in which physical processes in nature leave a record that can be ordered. Once the record of events is ordered, each event is known to be older or younger than each other event. In most cases the record is contained within a geological context, such as a stratigraphic sequence; in other cases the record may be contained within a single fossil or in the arrangement of astronomical bodies in space and time. The most important relative dating methods are stratigraphic dating and paleontologic dating. *See* SEQUENCE STRATIGRAPHY.

Stratigraphic dating. Stratigraphers rely on only a few basic principles to determine the order of events: (1) strata were horizontal (or nearly so) when initially deposited; (2) strata were laterally continuous over their original area of deposition; (3) younger strata lie on top of older strata; (4) a geological body must be younger than the particles of which it is composed; and (5) crosscutting features (such as faults and dikes) are younger than the rocks they cut.

Application of these principles allows determination of the order of events in local or regional geological sections and at archeological sites through determination of the physical relations between individual units. Exceptions and apparent exceptions to these principles are known but do not detract from their general applicability. Physical stratigraphic methods alone, however, do not allow the determination of relative ages of events recorded in sections that are widely separated geographically. *See* STRATIGRAPHY.

Paleontologic dating. Through physical stratigraphy, and close attention to the fossils contained within rocks, the principle of fossil succession was enunciated. This principle states that fossils succeed one another in a definite and determinable order, and that the period of geological time to which a rock body belongs can be recognized by the fossils included within it. Biological evolution is the fundamental process that makes this method work. Throughout geological history, different assemblages of organisms existed on the Earth at different times, and each organism has a definite time of appearance and (if extinct) disappearance. Paleontologic studies are the basis for the geologic time scale that is applicable to stratified rocks worldwide, for fossils allow correlation of rock units from one continent to another. The taxa present, the stage of evolution of related taxa, and the morphologic features within particular fossil taxa all change through time, and all may have temporal significance. Where quantitative techniques have been used to determine the temporal

range of a taxon, paleontological dating can be used to determine the age of rocks if direct quantitative methods cannot be applied. *See* FOSSIL; GEOLOGIC TIME SCALE; PALEONTOLOGY; SEDIMENTARY ROCKS.

Paleomagnetic dating. Throughout Earth history the magnetic field has changed its polarity, direction, and intensity. Most rocks contain sufficient iron oxide minerals to make them slightly or strongly magnetic, and they preserve a record of the orientation of the rock with respect to the Earth's magnetic field at that time. A principal use of these variations is chronological, and it is based on the fact that the polarity of the global magnetic field has reversed many times at highly irregular intervals. The timing of geomagnetic polarity reversals is reasonably well known from the Jurassic to the present. The present normal period does not differ from any that occurred in the past, so an initial estimate of the age of a rock sequence based on isotopic dates or paleontology is required before magnetic polarity zonation can yield additional chronological information. In a thick stratigraphic section where the rate of sedimentation was more or less constant, enough magnetozones may be encountered that their relative positions will allow a fit to the polarity-reversal scale with only rudimentary initial knowledge of the age. Note that certain rock types are susceptible to remagnetization, and spurious magnetozones may be encountered. Such magnetozones have no chronological significance because they arise from effects other than field reversal.

Scales of intensity and direction of the Earth's magnetic field have also been developed for many areas over the last few thousand years. These are principally of interest to archeologists (hence the term "archeomagnetism"), because pottery and bricks record the direction and intensity of the magnetic field at the time they were made. Thus, determining the direction and intensity recorded in such artifacts allows placement of the artifact in time once the scales of intensity and direction have been calibrated locally. *See* PALEOMAGNETISM.

Annual counting methods. In subtropical climates, trees produce a layer of spring wood and summer wood in response to the annual temperature cycle, thus making tree rings. In glacial lakes, sedimentation is controlled annually by the freeze-thaw cycle, so that a distinctive layer called a varve is produced. Counting back from the present, very accurate ages of tree rings can be determined, which have found use in calibration of the radiocarbon time scale. In geological sections that contain annual varves, precise estimates of duration of deposition are possible, even if the numerical age of the section is not known. *See* DENDROCHRONOLOGY; VARVE.

Tephrostratigraphy. Volcanic ashes from large silicic eruptions may be deposited in a very brief time over areas of subcontinental dimension. Each of these ashes is recognizable by the chemical composition of its glass phase and, once dated by quantitative techniques in one locality, allows a date to be assigned to a stratigraphic section wherever else it is identified. *See* VOLCANO.

Chemical methods. Several chemical processes occur slowly, producing changes over times of geological interest; among these are the hydration of obsidian, and the conversion of L- to D-amino acids (racemization or epimerization). Determination of age requires measurement of a rate constant for the process, knowledge of the temperature history of the material under study, and (particularly for amino acid racemization) knowledge of the chemical environment of the materials. *See* AMINO ACID DATING; OBSIDIAN; RACEMIZATION.

Natural volcanic glasses (obsidians) contain little water, but when a fresh surface is exposed through natural (breakage during transport; loading) or human (flaking) process, the surface absorbs water to form a hydrated layer, which thickens through time. With reasonable estimates of the temperature at which hydration occurs, and the composition of the obsidian, ages of obsidian artifacts and of geologically fractured obsidian can be obtained. The method has been applied principally in archeology, but also to obsidian fractured by glacial loading.

In biological systems, only left-handed (L-) amino acids are utilized, but right-handed (D-) amino acids are equally stable. In geological environments over time, there is a slow conversion of L- to D-amino acids (and back) until half of the molecules are right-handed. Thus, the ratio D/L changes from 0 to 1 (where only one asymmetric carbon is present in the amino acid). In theory, it should be possible to measure the age of ancient organic material by measuring the D/L ratio. However, the conversion process is very sensitive to temperature history and chemical environment, so the method has found use primarily as a stratigraphic tool. Few environments are constant in temperature over periods of several hundred thousand years, but the deep sea is one, and studies of deep-sea cores demonstrate that the method is useful in those environments.

While the bone-fluorine relationship is not strictly a dating method, it has some utility. Bones that have been buried and fossilized take up fluorine from ground water. Thus, older bones contain, in general, more fluorine than younger ones. By measuring fluorine, it may be determined whether an assemblage of bones is likely to be of the same age. Application of this method was instrumental in exposing the Piltdown hoax.

Calibrated isotopic scales. As with paleomagnetic polarity, other changes have occurred through Earth history that leave a record in the rocks. Among these are oxygen, carbon, and strontium isotopic ratios. Once the record of these changes has been calibrated by quantitative methods, measurement of the isotopic composition of materials in a geological section may date the material. The principal use of these methods has been in rocks of Tertiary or later age. *See* TERTIARY.

Rocks of different ages and rubidium content have different strontium isotope ($^{87}Sr/^{86}Sr$) ratios, and as they weather, the isotopic composition of strontium in the ocean changes. In the modern ocean, the $^{87}Sr/^{86}Sr$ ratio is uniform, and was presumably uniform in the past. Over the last 40 million years, there

has been a rise in the ratio of ^{87}Sr to ^{86}Sr in the oceans. As marine organisms build their shells of calcium carbonate, they incorporate small amounts of strontium, thus providing a record of the $^{87}Sr/^{86}Sr$ ratio in ocean water at the time they form. The $^{87}Sr/^{86}Sr$ ratio of shells or tests less than 40 million years old determines their age within about 1 million years. *See* STRONTIUM.

As water is removed from the oceans and locked up in glacial ice, the oceans become isotopically heavier because water ($H_2{}^{16}O$) evaporates more readily than heavy water ($H_2{}^{18}O$). Organisms that extract carbonate from the oceans record these changes, so that scientists were able to determine a curve of the isotopic composition of the oceans. As glaciers wax and wane, so does the isotopic composition of the ocean. Therefore, there are several times at which the isotopic composition of the ocean is identical. Thus, a good estimate of the age of specimens is needed before the oxygen isotope ($^{18}O/^{16}O$) ratio yields additional temporal information. However, knowing whether the $^{18}O/^{16}O$ ratio reflects glacial or interglacial times restricts the possible ages of a specimen to well-defined intervals.

Quantitative Methods

Unlike chemical methods, in which changes depend both on time and on environmental conditions, methods which are based on radioactive decay depend only on time. A parent nuclide may decay to one stable daughter in a single step by simple decay [for example, rubidium decays to strontium plus a beta particle ($^{87}Rb \rightarrow {}^{87}Sr + \beta$)]; to two daughters by branched decay through different processes [for example, potassium captures an electron to form argon, or loses a beta particle to form calcium ($^{40}K + e^- \rightarrow {}^{40}Ar$; $^{40}K \rightarrow {}^{40}Ca + \beta$)]; to one stable daughter through a series of steps (chain decay); or into two unequal-sized fragments by fission. In all cases, the number of parent atoms decreases as the number of daughter atoms increases, so that for each method there is an age-sensitive isotopic ratio of daughter to parent that increases with time. Many different isotopes have been exploited for measuring the age of geological and archeological materials. For example, the **table** shows parent isotopes, their half-lives, and the resulting daughter products. *See* HALF-LIFE.

Isotopic methods. All dating methods based on radioactive decay use simple equations (1)–(3) that

$$N = N_0\, e^{-\lambda t} \tag{1}$$

$$t = \frac{1}{\lambda}\ \ln \frac{N_0}{N} \tag{2}$$

$$t = \frac{1}{\lambda}\ \ln\left(\frac{D - D_0}{N} + 1\right) \tag{3}$$

relate time (t) to the number of parent atoms originally present (N_0), the number of parent atoms that remain (N) after t, and the number of daughter atoms produced (D), accounting for any daughter atoms initially present (D_0); λ is the decay constant ($\log_e$ 2/half-life), and e is a fundamental-constant. To determine an age, N and D are measured, and D_0 is measured or estimated. In order for an age to reflect the time of the event in question, changes in the number of parent and daughter atoms must result only from radioactive decay. *See* RADIOACTIVITY.

Potassium-argon dating. Potassium is the dominant cation in micas, many alkali feldspars, and leucite. It occurs in high concentration in nepheline and many volcanic glasses, and in lower concentration in amphiboles, plagioclase, and basic volcanic rocks. The natural isotopic composition of potassium is constant; potassium-40 (^{40}K) decays either to calcium-40 (^{40}Ca) by beta emission or to argon-40 (^{40}Ar) by electron capture or positron emission. As a result, the $^{40}Ar/^{40}K$ ratio increases with time in potassium-bearing phases, and measurement of that ratio provides a measure of time. The method has been applied principally to igneous rocks and minerals, and has a useful range from about 100,000 years to the age of the Earth (4.5 billion years).

Argon is quantitatively retained by many solids at low temperatures, but diffuses rapidly through solid materials at high temperatures, so initial argon is not generally a problem. The event that is dated is the time of cooling to a temperature at which argon is effectively trapped, or the time at which low-temperature minerals crystallize.

In conventional potassium-argon dating, no initial ^{40}Ar is assumed to have been present in the sample. In most cases, this assumption is justified, but some samples have been shown to contain initial argon, and yield potassium-argon ages that are too old. Other samples have been shown to have lost argon, and yield ages that are too young.

In $^{40}Ar/^{39}Ar$ dating, a sample is first irradiated with neutrons to produce ^{39}Ar from ^{39}K. The ^{39}Ar content then becomes a measure of the potassium content of the sample. An advantage of this method is that the entire analysis is done on a single sample and not with separate aliquots for potassium and argon analysis. The age may be measured by extracting all of the gas at once, which produces an age comparable to that of a conventional potassium-argon determination. Alternatively, the gas may be extracted in a series of steps at higher and higher temperatures, and an age computed for each fraction of the gas; this method is referred to as step heating or incremental release. Ages are commonly plotted against temperature of release or the fraction of ^{39}Ar released, to

Principal parent and daughter isotopes used in radiometric dating

Radioactive parent isotope	Stable daughter isotope	Half-life, years
Carbon-14	Nitrogen-14	5730
Potassium-40	Argon-40	1.25×10^9
Rubidium-87	Strontium-87	4.88×10^{10}
Samarium-147	Neodymium-143	1.06×10^{11}
Lutetium-176	Hafnium-176	3.5×10^{10}
Rhenium-187	Osmium-187	4.3×10^{10}
Thorium-232	Lead-208	1.4×10^{10}
Uranium-235	Lead-207	7.04×10^8
Uranium-238	Lead-206	4.47×10^9

show a spectrum of ages. In this method, the gas released at low temperature comes from poorly retentive sites within the material, but gas released at higher temperatures comes from more retentive sites and therefore yields ages nearer the time of initial cooling. In an elegant variant of the method, a focused laser is used to heat a single crystal to extract the argon, from which an age is computed; this is called the single-crystal laser fusion $^{40}Ar/^{39}Ar$ method.

Rubidium-strontium dating. Rubidium-strontium dating has been widely applied to igneous and metamorphic rocks. The method is based on the beta decay of rubidium-87 (^{87}Rb) to strontium-87 (^{87}Sr), but because ^{87}Sr is present in significant amounts in most earth materials, it is essential that it be measured. This is accomplished by isotopic analysis of three or more cogenetic minerals or rocks, and plotting $^{87}Sr/^{86}Sr$ versus $^{87}Rb/^{86}Sr$. As ^{86}Sr is a nonradiogenic stable isotope of strontium, these ratios differ only because of differences in the original rubidium content of the material, and because of the formation of ^{87}Sr by radioactive decay. A plot of these ratios for an undisturbed system results in a straight line whose slope (m) is proportional to age [$t = m(e^{\lambda t} - 1)$], and whose y-intercept is the initial $^{87}Sr/^{86}Sr$ for the system. Ages measured by this method range from about 2 million to about 4 billion years.

Uranium-lead dating. Although they are chain decays, for ancient systems the decay of uranium and thorium isotopes to lead isotopes (^{238}U to ^{206}Pb, ^{235}U to ^{207}Pb, and ^{232}Th to ^{208}Pb) can be considered as occurring in a single step. Lead isotopic ratios are measured with respect to ^{204}Pb which is stable and is not a product of radioactive decay. In uranium- and thorium-bearing systems, the $^{206}Pb/^{204}Pb$, $^{207}Pb/^{204}Pb$, and $^{208}Pb/^{204}Pb$ ratios increase with time, and their measurement, together with measurement of the uranium and thorium content, allows computation of the time that has elapsed since the minerals crystallized. The methods are applied chiefly to zircon, a common accessory mineral in igneous and metamorphic rocks. If both the ^{238}U-^{206}Pb and ^{235}U-^{207}Pb systems have remained closed, both systems will yield the same age and are said to be concordant. The locus of all values of $^{206}Pb/^{238}U$ and $^{207}Pb/^{235}U$ that yield concordant ages is a line called concordia. Many ages are not concordant, and a series of determinations from the same rock may lie on a straight line (discordia) that leads from concordia to the origin, suggesting that the discordant ages are caused by recent loss of lead. In other cases, discordia intersects concordia twice, and may indicate both the time of original crystallization and a time of loss of lead in the past. For ages less than 200 million years, concordia is nearly a straight line, and Tera-Wasserburg diagrams are used instead, in which $^{207}Pb*/^{206}Pb*$ versus $^{238}U/^{206}Pb$ is plotted, where the asterisk implies radiogenic lead. These diagrams enhance the curvature of concordia for young ages. Uranium-lead methods are principally used on rocks older than 100 million years, but in recent years high-precision ages down to about 10 million years have been reported. *See* LEAD ISOTOPES (GEOCHEMISTRY); MASS SPECTROMETRY; ZIRCON.

Common lead dating. The time of isolation of lead from uranium-bearing systems into phases that contain no uranium can be determined by measurement of lead isotope ratios of sulfides, knowledge of the age of the Earth, and a reasonable estimate of the U/Pb ratio of the system from which the lead derived. This is called the common lead method, and has been applied to many ore bodies that are older than 100 million years in order to estimate the time of mineralization. *See* EARTH, AGE OF.

Uranium series disequilibrium methods. Many elements form as uranium decays by alpha and beta emission to lead. If a mineral containing uranium is closed to all elements in the decay series, each element will eventually be present in an amount related to its half-life. This condition is known as radioactive equilibrium, in which the activities (λN) of all elements are equal. Intermediate daughters may be separated from each other by geological or geochemical processes, leading to geochemical systems that are out of equilibrium. The quantitative departure of the activities from equality is a measure of the time that the series was disrupted. Several different isotopic pairs in the uranium decay series have been used for dating in the range from a few years to about 10^6 years. Most of these are in the ^{238}U decay series, but protactinium-231 (^{231}Pa) in the ^{235}U decay series has also proven useful for some systems. *See* PROTACTINIUM.

The principal use of uranium-disequilibrium dating systems has been in the study of deep-sea sediments; but some systems, notably $^{234}U/^{238}U$, have been applied to nonmarine carbonates with reasonable success. In water and minerals on continents the $^{234}U/^{238}U$ ratio is greater than unity, because ^{234}U is preferentially leached from minerals damaged by previous decays in the chain leading to it. Uranium is incorporated into calcium carbonate and other secondary minerals, and thereafter the excess ^{234}U decays more rapidly than it is formed, until the activity ratio returns to unity. The half-life for this process is 248,000 years. Another method for dating terrestrial carbonates is based on the $^{230}Th/^{234}U$ ratio. This is possible because calcium carbonate excludes thorium but includes significant quantites of uranium. As time passes, the amount of ^{230}Th increases because of the decay of both ^{234}U and ^{238}U in the sample. *See* DEEP-MARINE SEDIMENTS.

Other rock-dating methods. Many other methods based on decay of rare elements with long half-lives have been developed and applied since about 1980 to rocks older than about 10^7 years. Examples include the lutetium-hafnium (^{176}Lu-^{176}Hf) method applied to several common minerals, the rhenium-osmium (^{187}Re-^{187}Os) method applied principally on molybdenite and other sulfides, and the samarium-neodymium (^{147}Sm-^{143}Nd) method. Isochron techniques similar to those used in rubidium-strontium dating are normally applied.

Carbon-14 method. Radiocarbon dating is useful over the past 70,000 years and is based on the decay of

carbon-14 (^{14}C) to nitrogen-14 (^{14}N) by beta emission. ^{14}C is produced in the upper atmosphere by cosmic-ray bombardment of nitrogen, is oxidized to carbon dioxide, and enters biological systems through the photosynthetic reaction and metabolic reactions of herbivores; thus, living organisms are radioactive. Carbonates precipitated from water are also radioactive, because atmospheric radiocarbon exchanges with carbonate in solution in water bodies. When an organism dies or when carbonate minerals form, carbonates are removed from the active carbon cycle, and the ^{14}C within them begins to decay. To calculate an age, an initial activity (A_0) of ^{14}C is assumed, the activity (A) of the sample is measured, and ages are computed from Eq. (4), where

$$t = \frac{1}{\lambda} \ln\left(\frac{A_0}{A}\right) \tag{4}$$

t is time, and λ is the decay constant. Uncertainties in the age arise from analytical error, and from variations in the assumed initial activity. The latter are corrected by using curves calibrated against tree rings and by measuring ^{13}C content to correct for isotopic fractionation. Radiocarbon dates are most commonly measured on charcoal, wood, and carbonate minerals. *See* RADIOCARBON DATING.

Tritium-helium dating. This method is useful for dating ground waters, and depends on the decay of tritium (^3H) to helium (^3He); ^3H has a half-life of only 12.43 years, so the method is limited to water less than 40 years old. In this method, ^3He formed from decay of ^3H must be distinguished from ^3He derived from the atmosphere and other sources. A correction is made for atmospheric ^3He by measuring the concentration of dissolved neon; a correction for ^3He produced by subsurface nuclear reactions [primarily fission of lithium-6 (^6Li)] is made by measuring radiogenic ^4He in a sample. In this method, ^3H is normally measured by ingrowth of ^3He following complete degassing and a 6-month holding period. Tritium-helium ages have been used to evaluate ground-water recharge rates, timing of contaminant spills, ground-water flow rates, and as ground-water flow models. *See* TRITIUM.

Other methods based on radiation. Several methods of dating exploit the effects produced in crystals by particles emitted during radioactive decay. Thus, radioactivity is required, but trails of damage (fission track method) or trapped electrons (thermoluminescence, optically stimulated luminescence, and electron spin resonance methods) are measured rather than measuring the abundance of daughter nuclides, as is the case in the methods discussed above.

Fission track method. ^{238}U decays by alpha emission and also by spontaneous fission. The fission products disrupt the physical structure of the material in which the uranium is contained, and leave trails of damage that can be enlarged by chemical etching using strong reagents such as hydrofluoric acid or sodium hydroxide. The density of these fission tracks (number per unit area) is dependent on the age of the specimen and the uranium concentration; hence an age can be calculated by measurement of the track density and the uranium content. If materials containing fission tracks are heated, the tracks fade as the damaged parts are annealed. Thus, the event being dated is the time of cooling of the material below a temperature at which track fading is important.

To date a sample, a polished surface is prepared, the sample is etched, and the fission tracks are counted over a determined area. The uranium content is measured while irradiating the sample with a known dose of thermal neutrons that induce ^{235}U atoms to fission. As the $^{235}U/^{238}U$ ratio is known, the amount of ^{238}U may be calculated from the abundance of new tracks.

The most important mineral for fission track dating is zircon ($ZrSiO_4$), which contains relatively high concentrations of uranium and is very resistant to track fading. These qualities make it nearly ideal for production and retention of tracks. Sphene ($CaTiSiO_5$) has also been used, as have micas, but the latter have relatively low uranium contents and thus have low track densities in young rocks. Apatite [$Ca_5(PO_4)_3OH$] contains enough uranium to be useful, but its tracks anneal at low temperatures, and track fading is often a problem. Volcanic glass differs widely in its ability to retain tracks, but has been used for dating as well.

The method has been applied to artifacts as young as a few hundred years and to geological samples ranging from about 1 million to 1 billion years in age. *See* APATITE; FISSION TRACK DATING; MICA.

Thermoluminescence method. Some electrons freed from the valence bands of atoms by ionizing radiation from radioactive decay become trapped in crystal defects in many nonconducting materials. The number of such trapped electrons is a measure of the total radiation dose that a material has received; and if the annual dose rate is known, an age can be calculated by the equation t = dose/annual dose rate. When such materials are heated, electrons are freed from their traps, combine across luminescence centers, and emit light. The amount of light is a measure of the number of trapped electrons. The greatest difficulty with the method is determining the annual dose rate, but the method has found wide application in dating pottery, burned flints, and sedimentary deposits up to about 200,000 years old.

If light, rather than heat, is used to free electrons from their traps, the method is termed optically stimulated luminescence. *See* CRYSTAL DEFECTS; GEOCHRONOMETRY; THERMOLUMINESCENCE.

Electron spin resonance (ESR). As with thermoluminescence, this method measures the number of trapped electrons by measuring their absorption of microwave radiation. Paramagnetic centers with a static magnetic field are formed by traps occupied by a single electron. The energy of a microwave-frequency electromagnetic field applied to a sample is absorbed as the electrons in the traps oscillate between being oriented with and against the field, so measuring the amount of energy absorbed is a measure of the

number of electrons in traps. As with thermoluminescence dating, the greatest difficulty is in determining the annual dose rate. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY.

Cosmogenic nuclide dating. As cosmic rays interact with rock and soil at the Earth's surface, a variety of isotopes are produced by spallation, muon capture, neutron activation, and alpha-particle interaction. These isotopes include ^{3}He, ^{10}Be, ^{14}C, ^{21}Ne, ^{26}Al, and ^{36}Cl. The production rates are quite small, ranging from less than 10 to about 200 atoms per gram per year. Further, production rates vary with latitude, altitude, target composition, and time, and decrease rapidly with depth below the surface. The abundance of these cosmogenic nuclides in geological materials increases with time, and has been used to estimate surface exposure ages, erosion rates, and sedimentation rates. *See* COSMOGENIC NUCLIDE.

Astronomical Methods

These methods involve use of Milankovitch cycles to date various geological periods which have occurred on the Earth, and cosmochronology to determine the age of the universe.

Milankovitch cycles. Periodic changes in deposition in the deep sea are caused by small perturbations of the Earth's orbit around the Sun and the orientation of its rotational axis that result in small differences in energy flux to the Earth's surface, thus influencing climate. These changes occur at intervals of approximately 20,000, 41,000, and 100,000 years, at times reinforcing one another and at times opposing one another. The effects are expressed differently at different latitudes for cycles of different length. This technique has been used to refine the ages of the geomagnetic polarity time scale.

Cosmochronology. Astronomers have estimated the age of the universe, and of the Milky Way Galaxy, by various methods. It is well known that the universe is expanding equally from all points, and that the velocity of recession of galaxies observed from Earth increases with distance. The rate of increase of recession velocity with distance is called the Hubble constant; and knowing the recession rate and distance of galaxies at some distance, it is simple to find how long it took them to get there. Initial estimates for the age of the universe were approximately 20 billion years; but as the rate of expansion decreases with time, revised estimates are nearer 13 billion years. By contrast, the Earth and other bodies in the solar system are only about 4.5 billion years old. Comparison of present-day osmium isotope ratios with theoretically estimated initial ratios yields estimates of 8.6–15.7 billion years for the age of the Galaxy. *See* COSMOCHEMISTRY; COSMOLOGY; MILKY WAY GALAXY; UNIVERSE.

Francis H. Brown

Bibliography. M. Aitken, *An Introduction to Optical Dating: The Dating of Quaternary Sediments by the Use of Photon-Stimulated Luminescence*, 1998; R. Bowen, *Isotopes in Earth Sciences*, 1988; L. A. Currie (ed.), *Nuclear and Chemical Dating Techniques*, 1982; G. B. Dalrymple, *The Age of the Earth*, 1991; A. P. Dicken, *Radiogenic Isotope Geology*, 1997; G. Faure, *Principles of Isotope Geology*, 2d ed., 1986; J. Hess, M. L. Bender, and J.-G. Schilling, Evolution of the ratio of strontium-87 to strontium-86 in seawater from Cretaceous to Present, *Science*, 231:979–984, 1986; M. Ivanovich and R. S. Harmon, *Uranium Series Disequilibrium*, 1982; I. McDougall and T. M. Harrison, *Geochronology and Thermochronology by the ^{40}Ar/^{39}Ar Method*, 1999.

Datolite

A mineral neosilicate, composition $CaBSiO_4(OH)$, crystallizing in the monoclinic system. It usually occurs in crystals showing many faces and having an equidimensional habit. It may also be fine granular or compact and massive. Hardness is 5–5$^1/_2$ on Mohs scale; specific gravity is 2.8–3.0. The luster is vitreous, the crystals colorless or white with a greenish tinge. Datolite is a secondary mineral found in cracks and cavities in basaltic lavas or similar rocks associated with zeolites, apophyllite, prehnite, and calcite. It is found in the Harz Mountains, Germany; Bologna, Italy; and Arendal, Norway. In the United States fine crystals have come from Westfield, Massachusetts; Bergen Hill, New Jersey; and various places in Connecticut. In Michigan, in the Lake Superior copper district, datolite occurs in fine-grained porcelainlike masses which may be coppery red because of inclusions of native copper. *See* SILICATE MINERALS.

Cornelius S. Hurlbut, Jr.

Dawsoniidae

A subclass of the true mosses (Bryopsida) largely limited to the South Pacific. The subclass consists of a single genus, *Dawsonia*, of nine species. The gametophytes are remarkably similar to those of the Polytrichidae, but the origin and structure of the peristome are very different.

The plants are coarse and rigid in loose dark green or brownish tufts. They are simple or occasionally forked, grow erect from a procumbent base, and sometimes achieve a height of 26 in. (65 cm), much more than any other mosses. The stems may have a central strand of homogeneous fiberlike cells or both xylemlike and phloemlike cells. The leaves are long-linear or long-lanceolate from a relatively short, oblong-sheathing base. The limb, which may be as long as 1.6 in. (40 mm), is concave-tubulose and spreading or flat and spirally twisted when dry. It is coarsely toothed at the margins and at back. The costa occupies most of the limb, which is covered on its upper surface by 40–100 long green ridges, often with differentiated marginal cells. The cells of the sheath are long-linear, and those at the junction of sheath and limb are short and thick-walled. The inflorescences are dioecious and terminal. The perigonia are discoid with scalelike bracts. Male plants repeatedly form new growth through the perigonial bud. The setae are elongate and smooth or papillose. The

capsules are inclined, broadly ovoid-asymmetric, laterally compressed, and two-angled. The peristome teeth are long-filiform and somewhat twisted together in a bushy tuft. They consist of whole cells derived from eight concentric rows of amphithecial cells. They are united at the base into bundles of two to four U-shaped cells. The calyptrae are cucullate and matted with hairs. The haploid chromosome number is 7. *See* BRYOPHYTA; BRYOPSIDA; POLYTRICHIDAE. Howard Crum

Bibliography. B. O. van Zanten, A taxonomic revision of the genus *Dawsonia* R. Brown, *Lindbergia*, 2:1–48, 1973.

Day

A unit of time equal to the period of rotation of Earth. Different sorts of day are distinguished, according to how the period of rotation is reckoned with respect to one or another direction in space. A day is normally defined as 86,400 SI (Système International) seconds [86,400 s/d = (60 s/min) × (60 min/h) × (24 h/d)], where SI seconds are measured by atomic processes. *See* PHYSICAL MEASUREMENT.

Solar day. The apparent solar day is the interval between any two successive meridian transits of the Sun. It varies through the year, reaching about 24 h 30 s of ordinary clock time in December and about 23 h 59 min 39 s in September.

The mean solar day is the interval between any two successive meridian transits of an imagined point in the sky that moves along the celestial equator with a uniform motion equal to the average rate of motion of the Sun along the ecliptic. The difference between mean solar time and time measured by solar meridian transits is known as the equation of time. It leads to the analemma, the figure-eight seen on some globes. *See* EQUATION OF TIME.

Ordinary clocks are regulated to advance 24 h during a mean solar day. Because of irregularities in the Earth's rotation, leap seconds are occasionally added to keep the day in terms of seconds of atomic time (Coordinated Universal Time, or UTC) in step with the day based on rotation (UT1), though as of 2000 this system is under evaluation.

Sidereal day. The sidereal day is the interval between any two successive upper meridian transits of the vernal equinox. Similarly, as for the solar day, a distinction is made between the apparent sidereal day and the mean sidereal day which, however, differ at most by a small fraction of a second. The difference between sidereal time and solar time arises from the Earth's revolution around the Sun, which results in a year containing one more sidereal day than solar day. A mean sidereal day comprises 23 h 56 min 4.09053 s of a mean solar day.

The period of rotation of Earth with respect to a fixed direction in space is 0.0084 s longer than a sidereal day; the difference is caused by precession. No special name has been given to this kind of day, and although of theoretical interest, it is not used in practice.

Variations in duration. The mean solar day, the sidereal day, and the day mentioned in the preceding paragraph all vary together in consequence of variations in the speed of rotation of Earth, which are of three sorts: seasonal, irregular, and secular. The seasonal variations are probably caused, at least in part, by the action of winds and tides; the effect is to make the day about 0.001 s longer in March than in July, and is nearly repetitive from year to year. The irregular variations are probably the result of interactions between motions in the core of Earth and the outer layers; the effect is to cause more or less abrupt changes of several thousandths of a second in the length of the day, which persist for some years. The secular variation is the result of tidal friction, mainly in shallow seas, which causes the duration of the day to increase about 0.001 s in a century. *See* EARTH ROTATION AND ORBITAL MOTION; TIME. Gerald M. Clemence; Jay M. Pasachoff

Bibliography. H. Jeffreys, *The Earth: Its Origin, History and Physical Constitution*, 6th ed., 1976; D. D. McCarthy and W. W. Carter (eds.), *Variation in Earth's Rotation*, 1990; J. Mitton, *Cambridge Dictionary of Astronomy*, 2001; P. K. Seidelmann et al., *Explanatory Supplement to the Astronomical Almanac*, 1992; F. R. Stephenson, *Historical Eclipses and the Earth's Rotation*, 1997.

Dead reckoning

A form of navigation that determines position of a craft by advancing a previous position to a new one on the basis of assumed distance and direction moved. The name probably stems from the early practice of determining speed by throwing overboard a buoyant object, called a Dutchman's log, and noting the time needed for a known length of the vessel to pass the floating object, or attaching a line to the object (when the whole device became known as a chip log) and noting the amount of line paid out in a given time. In either case, the floating object was assumed to remain dead in the water, thus providing an indication of speed through the water. The reckoning of future positions of the vessel by means of this speed was known as dead reckoning.

Parameters. The parameters of dead reckoning are direction of motion and distance traveled. The intended direction of travel, the course, may differ from the direction steered because of the anticipated offsetting effect of wind (called leeway) or current, or both. When it is desired to distinguish between the two directions, mariners call the second the course steered, or heading, while aviators refer to it as the heading.

A compass is used to indicate direction. Distance is usually determined indirectly by measurement of speed and time, but it may be measured directly.

Measurement of direction. Direction is expressed as angular distance from a reference direction, usually north for courses and headings. This angular distance is customarily stated in integral degrees, from 000°

Fig. 1. Marine magnetic compass.

at north, clockwise to 360°. The reference direction used for advancing a position by dead reckoning is true north, the direction of the North Pole. The reference directions that may be used for courses steered and heading are magnetic north, the direction north along a magnetic meridian; compass north, the direction north as indicated by a magnetic compass; or gyro north, the direction north as indicated by a gyrocompass. For another reference direction, grid north, used in polar regions, *see* POLAR NAVIGATION.

Magnetic compass. For many centuries the primary direction instrument has been the magnetic compass. The first such compass probably consisted of a magnetic needle freely floated in water by means of a straw. This was replaced by the dry compass, and this in turn by the modern liquid compass in which the north-seeking magnetic element is placed in a compass bowl completely filled with a liquid which will not freeze at the temperatures in which the compass might reasonably be expected to operate.

For a typical modern marine magnetic compass (**Fig. 1**), the directional element, consisting essentially of several bundles of slender magnetized rods, is mounted by means of an almost frictionless bearing so that the compass will be responsive to weak magnetic fields.

Attached to the magnetic element is an annulus of lightweight material on which compass graduations are placed. This is the compass card, which remains aligned with the Earth's magnetic field as the vessel turns. Heading is indicated by the compass graduation aligned with a lubber's line in the longitudinal axis of the craft, generally on the forward side of the compass bowl aboard ship, and the after side in aircraft and land vehicles. Bearings can be measured by some compasses by means of a suitable attachment which is provided with sighting vanes.

Magnetic compasses are subject to compass error, the difference between true north and compass north. This error is the algebraic sum of two components. One of these, called variation by the navigator and magnetic declination by the magnetician, is the angle between magnetic and geographic meridians. It is related to the magnetic field of the Earth, and is subject to a small daily or diurnal change and to a slow progressive secular change. It is affected, also, by magnetic storms.

The second component is deviation, the angle between the magnetic meridian and the axis of the compass card. It is the result of local magnetic influences, particularly those within the craft. Permanent magnetism in metal of the craft, and transient magnetism induced there by the Earth's magnetic field are important contributors. Direct currents in electrical wiring also influence the compass. *See* MAGNETIC COMPASS.

The various disturbing influences can be largely neutralized by establishing additional magnetic fields of equal strength but opposite polarity. This process is called compass adjustment, or sometimes compass compensation, especially with respect to aircraft compasses. In the U.S. Navy the latter expression is used to indicate the process of neutralizing the effects that degaussing currents exert on a magnetic compass. These currents may be used to change a ship's magnetic characteristics to protect against magnetic mines.

In some compass installations, particularly in aircraft, the magnetic compass is installed in a position relatively free from magnetic disturbances, and its indications are transmitted electrically to locations throughout the craft. Such an instrument is called a remote-indicating compass.

Compass error is determined by comparing the compass with one having a known error, with the azimuth of a celestial body, or with the bearing of a landmark.

When deviation is reduced to a minimum, the residual deviation on a number of headings is recorded on a deviation table or deviation card (**Fig. 2**) kept near the compass.

The directive force of the compass is the horizontal component of the Earth's magnetic field at the compass. This becomes progressively weaker as

#2 MASTER COMPASS
SWUNG: 7-24-58 BY: KB

TO FLY	STEER	TO FLY	STEER
N	359	180	181
15	15	195	196
30	30	210	211
45	44	225	226
60	60	240	241
75	75	255	256
90	90	270	271
105	105	285	286
120	121	300	301
135	135	315	316
150	151	330	330
165	166	345	345

Fig. 2. Typical aeronautical deviation card.

the magnetic poles are approached. Within a few hundred miles of these poles the magnetic compass is unreliable. *See* GEOMAGNETISM.

Gyrocompass. In addition to several magnetic compasses, nearly all naval vessels and ocean liners are equipped with one or more north-seeking gyrocompasses. This instrument has the disadvantage of requiring a source of electrical power, but it is not subject to variation or deviation. It may have a small gyro error, but in most modern installations this is not large enough to be significant, except in high latitudes. The compass tends to align its rotational axis with that of the Earth but, since it is constrained to remain in the horizontal, its directive force decreases as the horizontal becomes more nearly perpendicular to the Earth's axis. Beyond some latitude, generally about 70°, the gyro error becomes large and erratic and the instrument requires frequent checking. Near the geographical poles the gyrocompass becomes unreliable.

Great progress has been made in the development of better gyroscopes. As a result, gyrocompasses have become smaller, lighter, more accurate, and practical for smaller vessels. Gyrocompasses have replaced magnetic compasses as the primary source of directional information on many modern vessels. *See* GYROCOMPASS.

Directional gyro. A directional gyro, used in some aircraft and submarine compass systems, tends to maintain a fixed position in a great-circle plane. Better instruments are precessed to allow for rotation of the Earth, so that a craft following their indications tends to follow a great circle on the Earth. Since the instrument does not seek anything, it is subject to a slow, cumulative error called wander. This is allowed for or corrected at intervals by comparison with a directional reference, generally a magnetic compass or the azimuth of a celestial body. *See* AIRCRAFT COMPASS SYSTEM.

Several types of devices provide directional guidance in polar regions.

Distance or speed measure. Distance is generally stated by navigators in units of nautical miles, usually to integral miles by air navigators and to tenths of a mile by marine navigators. For this and other linear units of navigation *See* NAVIGATION.

Aboard ship, distance or speed is measured by means of a log or by an engine revolution counter. The primitive Dutchman's log and chip log were mentioned earlier. The latter is of particular interest because a series of knots was tied in the log line, and the number of knots which passed through the seaman's hands in a given time interval represented the speed of the vessel in nautical miles per hour. This is the origin of the name knot applied to the unit of speed still commonly used by marines.

About the middle of the seventeenth century, mechanical logs appeared. These had rotators towed through the water, first with the indicator also in the water, but later at the taffrail, giving the name to the taffrail log widely used for many years.

The pitometer log uses a pitot-static tube. The Forbes log uses a small rotor in a tube projecting below the bottom of the vessel. An electromagnetic log has a sensing element which produces a voltage directly proportional to speed through the water. Accurate so-called measured miles are marked out at various places on the beach to provide means for calibrating or checking logs.

In aircraft, speed through the air is measured by means of an airspeed indicator or Mach meter. The latter provides an indication of speed in units of the speed of sound, which varies with density of the atmosphere. For measurement of air speed a pitot-static tube is generally used with a suitable registering device. Distance is determined indirectly from speed and time. Distance, and indirectly speed, may also be determined by means of two fixes, the length of the line between them being the distance.

Practice. To determine position by dead reckoning, air and land navigators, and some marine navigators, use the best estimate of direction and distance traveled over the surface. Many marine navigators, however, prefer to use course steered and estimated speed through the water, without allowance for leeway, for their dead reckoning; they consider positions determined by allowance of estimated effects of wind and current as estimated positions.

The uncertainty of a dead-reckoning position, however determined, increases with time and, if there is an error in direction measurement, it also increases with distance traveled. From time to time an independent determination of position is made by means of external references. When a reliable position, called a fix, is so obtained, a new dead reckoning is started from this point.

The total combined effect of all errors since the last fix can be determined by comparing the new fix with the position that the craft would have occupied if there had been no errors. It is for this reason, in part, that many marine navigators prefer to use course steered and speed at which a vessel is propelled by its primary source of power for their dead reckoning, thus perpetuating the original concept of reckoning relative to an object dead in the water (motionless with respect to the water). For convenience, they consider the entire offset effect as the result of current. The true direction from the dead-reckoning position to a fix at the same time is called the set of the current. The distance between these two positions in nautical miles, divided by the time in hours since the last fix, is called the drift (speed) of the current in knots. Air navigators consider the entire offset the result of wind, determined by comparison of a no-wind position, also called an air position, with a fix at the same time. The no-wind position is the point at which the aircraft would have arrived if there had been no offsets.

Determination of position. Several different methods have been used for determination of position by dead reckoning.

Computation. Before reliable charts became generally available, dead reckoning was commonly performed by computation using various forms collectively known as sailings. This method is seldom used

by modern navigators except in small craft where chart work is difficult, or when computation of direction and distance between widely separated points is desired. A computer or hand-held calculator may be used to make a computation. *See* CALCULATORS; DIGITAL COMPUTER.

Plotting. As reliable charts and plotting sheets came into general use, plotting became the common method of performing dead reckoning, a method still widely used. Marine navigators usually measure the plotted direction by means of parallel rules (a device designed to move over the chart or plotting sheet parallel to itself) and a compass rose, or a drafting machine (which combines parallel motion with direction indication). Dividers are usually used with the latitude graduations on the chart or plotting sheet for measuring distance. Air navigators usually use some form of plotter, one type of which is shown in **Fig. 3.**

Dead-reckoning computers. In many large naval vessels and commercial ships the dead reckoning is performed automatically by a device that receives inputs of direction from a gyrocompass and speed from a log and continuously computes dead-reckoning position, which is displayed on dials or traced on a chart or plotting sheet.

Doppler navigation. The Doppler effect, a frequency shift that is proportional to the speed of relative motion between transmitter and receiver or reflector of radiant energy, either acoustic or electromagnetic, is used in a system to accomplish dead reckoning, automatically. In the ship version, called a Doppler sonar navigator, ultrasonic energy is transmitted obliquely downward (typically 30° from the vertical) and the frequency of the return echo is noted. By using four beams separated 90° laterally, the system provides an indication of vessel speed in both the fore-and-aft and athwartship directions, so that total speed and direction of motion can be determined if the device is properly oriented. When reflections are from the sea bottom (bottom return mode), true ground speed (speed relative to the solid Earth) is measured. When reflections are from suspended particulate matter in the water (volume reverberation mode), the speed is relative to the water, similar to speed relative to an object dead in the water. In either mode the speed is integrated to determine distance from a starting point. Doppler sonar navigation has proved particularly useful in survey and geophysical exploration vessels. *See* SONAR.

Similar systems called Doppler navigators, which use electromagnetic energy, have been used in aircraft, but these systems have largely been replaced by inertial navigators. *See* DOPPLER EFFECT; DOPPLER RADAR.

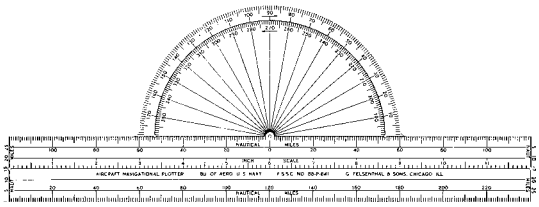

Fig. 3. A plotter in use among air navigators.

Inertial navigation. Most aircraft that are used for long overwater flights, and some others, as well as some ships, notably submarines, are equipped with one or more inertial navigators. This device, when properly aligned, provides a continuous indication of speed, position, and heading by means of appropriate inertial sensors. Gyroscopes are used to sense angular motions of the craft and maintain accelerometers in the correct orientation to sense linear accelerations, or changes in speed. Single integration of the accelerations provides a measure of speed, and double integration produces a measure of distance. The term inertial stems from the fact that measurements are made with respect to inertial space. Provision must be made to eliminate effects related to the rotation of the Earth because data relative to the Earth are desired.

An inertial navigator is self-contained, because it is independent of external reference points, and passive, because no energy is transmitted to obtain information from an external source. It is free from effects of wind and current, but like all dead-reckoning systems its output degrades with elapsed time and distance traveled, and must be reset periodically. Its use is particularly attractive in aircraft because of their high speed and hence relatively short time in transit. *See* INERTIAL GUIDANCE SYSTEM.

Continuous position indication. Knowledge of the real-time present position of a craft is generally considered essential to safe navigation. Dead reckoning, always available in some form, provides this capability, giving information useful in the evaluation of all other navigational data received. Satellite navigation systems such as the Global Positioning System can provide an essentially continuous position fix when enough satellites are in view. An integrated navigation system that combines satellite observations with measurements from inertial instruments, Doppler sonar, or other navigation sensors can use this redundancy to maintain accurate positioning during brief intervals of satellite outages. *See* CELESTIAL NAVIGATION; ELECTRONIC NAVIGATION SYSTEMS; PILOTING; SATELLITE NAVIGATION SYSTEMS. Alton B. Moody

Bibliography. N. Bowditch, *American Practical Navigator*, *An Epitome of Navigation*, 2002; T. J. Cutler, *Dutton's Nautical Navigation*, 15th ed., 2003; H. H. Shufeldt, G. D. Dunlap, and B. A. Bauer, *Piloting and Dead Reckoning*, 4th ed., 1999.

Death

Cessation of life functions. This can involve the whole organism (somatic death), individual organs (organ death), individual cells (cellular death), and individual parts of cells (organelle death). Although the next smaller level or organization, the macromolecules that make up the cell organelles, may also cease to function, their disintegration is ordinarily not spoken of as death.

Pathogenesis

Figure 1 represents a useful conceptualization of the dynamics of injury, death, and necrosis. A normal animal, cell, or organelle can be assumed to exist in a normal state of homeostasis or a normal "steady state." Following a disturbance of the normal steady state, referred to as an "injury," the living system can be assumed to undergo a change in the steady-state level, that is, to exhibit an altered steady state. Such altered steady states can be regarded as having increased, decreased, or unaltered levels of ability to maintain homeostasis. An altered steady state with increased homeostatic ability might include the hypertrophy of a muscle cell resulting from continued exercise; an altered steady state with decreased homeostatic ability is the atrophy or shrinking of cells that occurs following disuse or lack of stimulation by hormones or suitable nutrients. In the case of injuries that do not result in the death of the living system, the system may persist in one of the new or altered steady states for long periods up to many years; that is, although the system differs significantly from the normal steady state, it is able to continue its existence and to maintain homeostasis. On the other hand, certain severe injurious stimuli result in an altered state that ultimately passes a point beyond which recovery of homeostatic ability is impossible even if the injurious stimulus or situation is removed. This may be considered as a point of no return or the point of death of the living system. Following this point, the system undergoes exponential decay, leading to complete equilibrium with its environment. This period of decay, which is characterized by reactions resulting in the breakdown of structural components of the system, is referred to as necrosis. *See* HOMEOSTASIS.

Somatic death. Somatic death has different meanings for the biologist, the theologian, the physician, and the lawyer. In the biological sense, somatic death refers to the cessation of characteristic life functions. In higher organisms, death can be assumed to result from an interplay between the genetic background of the individual and the effects of the environment. Death is, in a sense, the price of differentiation; in another sense, protoplasm is immortal and all life springs from preexisting life. The undifferentiated amebas need never die, only divide. It would appear that death is inevitable for higher organisms such as mammals and, in a sense, life is a series of little deaths.

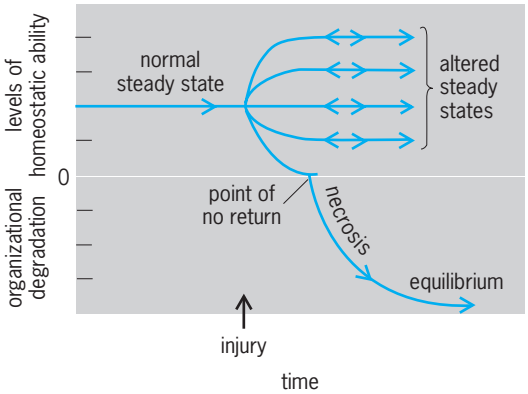

Fig. 1. Changes that can occur in living systems following injury. In the case of lethal or fatal injuries, the system can eventually be assumed to pass a point, shown as zero level of homeostatic ability, beyond which it is unable to restore a normal state of homeostasis even if the injurious situation is removed. (*After E. Bajusz and G. Jasmin, eds., Methods and Achievements in Experimental Pathology, vol. 4, Karger, Basel, Switzerland, 1969*)

Forensic medicine. Developments in medicine and associated technology involving supportive therapy, resuscitation, hypothermia, extracorporeal circulation, and organ transplantation reopened serious questions regarding the definition of when an individual is "dead." The vagueness and inconsistency in previous legal and medical definitions have led to serious dilemmas. These are of particular importance in the selection of donors for transplantation of vital organs such as the heart. It has been proposed that death be defined as irreversible cessation of all of the following: (1) total cerebral function, including determination by electroencephalography; (2) spontaneous function of the respiratory system; and (3) spontaneous function of the circulatory system. *See* FORENSIC MEDICINE.

Technology and prolonged survival. Medical techniques have shown that people can be sustained for prolonged periods without spontaneous respiratory or cardiac movements as well as without electrical evidence of central nervous system activity. These techniques have also achieved the resuscitation of individuals who would formerly have been considered dead and have clearly shown the survival of transplanted organs from cadavers.

Because of the present ability to maintain cardiac and respiratory function, even in the presence of complete absence of cerebral activity, special emphasis has been placed on laboratory techniques such as electroencephalography, coupled with clinical evaluation, to determine whether irreversible cessation of cerebral function has, in fact, occurred. If intravital brain death can be fully established clinically as well as by the electroencephalogram, further attempts to maintain the patient, even if successful, result only in incubation of a "dead" brain at the body temperature of the patient. Just as organs may survive after the death of the organism, the organism, properly supported by mechanical or other means, can survive the death of organs, including the brain.

Postmortem changes in body. The following are changes that occur in the body after somatic death. The rate of progression and time of occurrence of these changes are of some use in determining the time and circumstances of death and, accordingly, are important in forensic medicine.

1. *Cooling (algor mortis).* This refers to the cooling of the body that occurs postmortem. The rate of algor mortis is affected by environmental temperature, conductivity of the environment, amount and type of clothing, and amount of body fat.

2. *Rigidity (rigor mortis).* This refers to the stiffening of the musculature. Although its time of onset is somewhat dependent on the degree of muscular

activity at the time of death, it usually begins within 5–10 h. The muscles of the jaw are first affected. Rigor mortis begins to disappear after 3 or 4 days.

3. *Staining (livor mortis)*. This refers to the reddish-blue discoloration that occurs in the dependent portions of the body and results from the gradual gravitational flow of unclotted blood.

4. *Clotting of blood*. Blood begins to clot shortly after death and sometimes prior to cessation of cardiac function. Such clots are not attached to the vessel wall and can be distinguished by this, as well as by their appearance, from antemortem thrombi.

5. *Autolysis*. This refers to the necrosis that occurs in the cells of the body following somatic death. The changes are indistinguishable from those occurring following ischemic (deprivation of blood supply) injury in the body, although an inflammatory reaction, a response to necrosis in tissues of the body, is absent.

6. *Putrefaction*. This refers to the bacterial and enzymatic decomposition that begins shortly after death, caused by microorganisms, including bacteria and fungi that enter the body, usually from the gastrointestinal tract. It results in production of gases and a greenish discoloration of the tissues.

Organ death. Following somatic death, all organs do not die at the same rate. Thus, certain vital neurons in the brain survive for approximately 5 min, the muscle cells in the heart for approximately 15 min, the kidney cells for about 30 min, and the liver cells for 45 min. If a blood supply and a suitable environment are restored to these organs prior to the point of no return, they will recover normal function. Length of survival can be prolonged by reducing the temperature of the organs to near 32°F (0°C), which appreciably retards metabolic function and extends survival time. This survival of organs after the death of the host forms the basis for organ transplantation from cadavers.

Cellular death. Death and subsequent necrosis of cells form an important part of many reactions of cells to injury in disease states. Thus, in the heart after a coronary occlusion, the cells of the myocardium die and undergo disintegration. In this case, the physiological consequences, cardiac failure and death of the organism, are the direct consequences of the death and disintegration of individual muscle fibers. The death and disintegration of cells within a living animal incite a vascular reaction called an inflammatory reaction on the part of the host. It is characterized by a series of vascular phenomena which include the migration of leukocytes (white blood cells) from the blood capillaries into the area of tissue damage with subsequent digestion and phagocytosis of the cellular debris resulting from cellular necrosis.

Cellular changes. The essential features of necrosis are similar following a variety of types of lethal injury to cells. Thus, the changes in the necrotic heart muscle cells mentioned above are essentially similar to those in the neurons of the brain after a stroke, to those in the cells of the lung in a tuberculous infection, and to those of the liver cells during viral hepatitis. The cell and its organelles show a characteristic series of changes involving the plasma membrane, the mitochondria, the endoplasmic reticulum, the lysosomes, and the nucleus (**Figs. 2** and **3**).

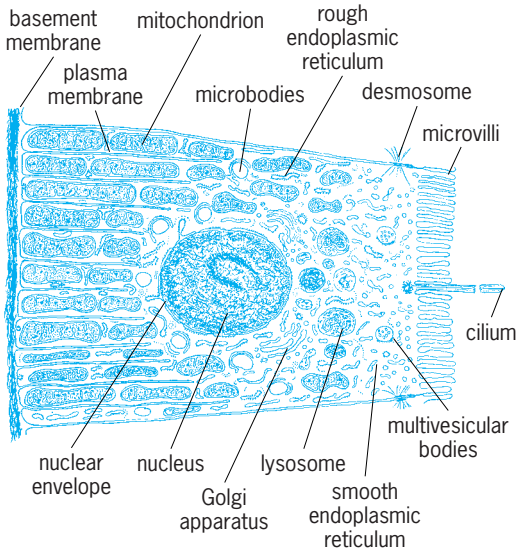

Fig. 2. Diagrammatic representation of a normal kidney tubule cell as seen with the electron microscope. (*After F. L. Ginn, J. Shelburne, and B. F. Trump, Disorders of cell volume regulation, I: Effects of inhibition of plasma membrane adenosine triphosphatase with ouabain, Amer. J. Pathol., 53:1041–1047, 1968*)

In a normal kidney tubule cell, at the apex of the cell adjacent to the tubular lumen, the plasma membrane forms numerous microvilli; at the base of the cell, the plasma membrane is highly invaginated and forms complex compartments that contain the elongate mitochondria. The mitochondria are limited by two membranes; the inner membrane forms a series

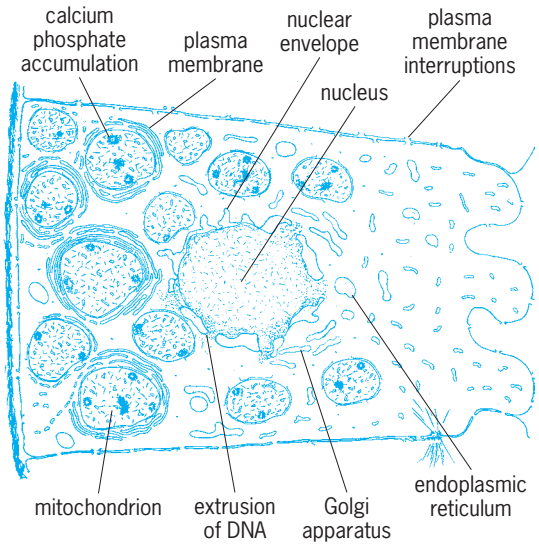

Fig. 3. Diagram showing a necrotic kidney tubule cell. (*After F. L. Ginn, J. Shelburne, and B. F. Trump, Disorders of cell volume regulation, I: Effects of inhibition of plasma membrane adenosine triphosphatase with ouabain, Amer. J. Pathol., 53:1041–1047, 1968*)

of invaginations called cristae, which are responsible for production of adenosine triphosphate (ATP). The nucleus, in the center of the cell, is surrounded by the double-walled nuclear envelope, the lumen of which is continuous with that of the endoplasmic reticulum. The endoplasmic reticulum has two portions, smooth- and rough-surfaced. The rough-surfaced endoplasmic reticulum is covered with particles called ribosomes, on which synthesis of protein for secretion from the cell occurs. The Golgi apparatus is in the supranuclear region. Small organelles called microbodies appear to be responsible for a primitive form of cellular respiration.

A characteristic feature of necrotic cells includes massive swelling of the cytoplasm, often referred to in the past as cloudy swelling. This swelling results from movements of ions, such as sodium and water, into the cell as a result of the loss of active transport mechanisms at the cell membrane. The swelling involves specific organelles such as the endoplasmic reticulum and the mitochondria. The endoplasmic reticulum becomes greatly dilated, and the ribosomes along the surface of the reticulum are disorganized, with loss of ability to synthesize protein. The plasma membrane is whorled near the base of the cell and becomes interrupted elsewhere. The mitochondria exhibit massive swelling of their inner compartment, and in this stage have lost the ability for production of ATP. Mitochondria also contain accumulations of calcium phosphate. The nucleus undergoes lysis, often called karyolysis, and the nuclear deoxyribonucleic acid (DNA) extrudes into the cytoplasm through the nuclear pores. The lysosomes in this stage rupture, with release of digestive enzymes into the cytoplasm. Such enzymes are evidently responsible for the disintegration of cell components that occurs during the necrotic phase.

Subcellular changes. The elucidation of the myriad of subcellular changes that occur in injured, dying, and dead cells is possible through the use of the electron microscope and associated techniques. With this instrument, which is capable of resolutions much greater than those obtainable by light microscopy, a close correlation between the structure of organelles, or parts of organelles, and the physiological and biochemical functions of these same structures is possible. Use of the electron microscope permits the demonstration of cellular damage at very early intervals following application of an injury. Such studies were impossible by light microscopy alone, and in the past it was necessary to wait for many hours for the characteristic changes to develop following death of the cell. The normal, the injured, and the dead cell are difficult to delineate clearly and exist along a continuum of cell structure and function.

Biochemical parameters. It is believed that the limiting factors in the survival of most cells following a lethal injury are the integrity of the plasma membrane and interference with ATP synthesis by mitochondria or other organelles. Although the lysosomes, which are membrane-bound vacuoles containing hydrolytic enzymes, are sometimes thought of as "suicide bags" which may kill the cell by premature release of enzymes, most experiments indicate that release of hydrolytic enzymes from these bodies does not occur until after cellular death, at which time such release catalyzes hydrolytic reactions resulting in disintegration of cell components. *See* LYSOSOME.

Surviving elements. After death of the cell as a unit, some parts survive longer than others. Just as the heart may continue to beat for a short time after respiration has ceased, protein synthesis within the endoplasmic reticulum may occur after cessation of respiration in mitochondria. Because of the great complexity of changes that can occur within the cell, and because of the dissociation between their time of occurrence, an exact definition of cell death is extremely difficult.

Organelle death. Following the death of the cell, organelles continue to function for variable periods of time, depending upon the manner in which death occurred. Thus, in cells dying from ischemia, the mitochondria continue to respire for an hour or more, although the synthesis of ATP is impaired. Also, portions of the endoplasmic reticulum exhibiting polyribosomes (rough-surfaced endoplasmic reticulum) may continue to synthesize protein. The lysosomes maintain their enzyme activity for as long as 2 days following death of the cell at 98.6°F (37°C). Their enzymes are thought to be important in the removal of the products of dead tissue by hydrolytic digestion.

The survival of organelles after the death and disintegration of the cell is well illustrated by experiments in which cells are destroyed by mechanical homogenization. In such experiments, cell organelles, including fragments of the endoplasmic reticulum or mitochondria, can then be maintained for relatively long periods of time in the test tube. It has been shown that fragments of the endoplasmic reticulum can continue to synthesize protein following disintegration of the cell and that mitochondria can respire, synthesize ATP, and even continue to replicate after the cell has been utterly destroyed. Such experiments illustrate the relative independence of organelles such as mitochondria and even suggest the possibility of organelle transplantation.

Autophagocytosis. During the life of most, if not all, cells, portions of the cytoplasm are enclosed from time to time in vacuoles that come to contain hydrolytic enzymes. These vacuoles are known as autophagic vacuoles or autolysosomes. Death and disintegration of the enclosed organelles occur in these vacuoles. Such autophagic vacuoles may be an important site of normal organelle turnover and may provide the cell with energy sources in times of nutritional deficiency or starvation.

Necrobiosis. This is a term used to indicate physiological death of cells and their replacement within the living organism. Examples of death and turnover of cells in normal living animals are seen in aging, in shedding of the superficial cells of the epidermis of the skin or the epithelium of the gastrointestinal

tract, and in the involution of the breast and uterus following pregnancy. Benjamin F. Trump

Issues

The advancements of modern scientific medicine have resulted in significant changes in the care and treatment of individuals with life-threatening illness, and challenge concepts and definitions of death. Considerable efforts are being made toward establishing clarity by developing a more precise definition of death and the corresponding criteria for determining death. These changes are accompanied by a cluster of persistent ethical issues such as the right to refuse life-saving treatment, the circumstances under which it is morally right not to prolong life, and whether extreme situations justify a deliberate act of euthanasia or physician-assisted suicide. Legislation, court decisions, and public referendums are changing the shape of public policy surrounding death and dying in the United States, even though a clear moral consensus may be difficult to reach.

As the determination of death becomes increasingly complex in a technological society, the definition of death becomes correspondingly more complicated. The meaning of death is a philosophical concept which lies outside the scientific point of view. Traditionally, death has been understood by the general public, and by the physician as well, as the cessation of life or ceasing to exist. Although dying is a process that may continue over a considerable period of time, for all practical purposes death occurs at that point where there is an irreversible loss of one's essential human characteristics. Such loss indicates the death of the organism as a whole, the death of a person.

Historical perspective. Up until the seventeenth century in western civilization, death was considered a discrete moment in which the soul left the body. This point was determined by the absence of visible manifestations of life such as were apparent to the ordinary person, absence of breathing, and cessation of circulation.

In the eighteenth and nineteenth centuries a concern arose over premature burial as people began to doubt the accuracy of judgment of those who made the assessment that death had occurred. During this period life signals were installed on caskets with the belief that a person mistakenly believed to be dead should have an opportunity to sound an alarm until burial had been accomplished. Laws arose requiring physicians to examine the body prior to burial, and interment was sometimes delayed until the signs of putrefaction proved death undeniably. In some cases, boiling liquids or surgical incisions were administered as final proof.

In earlier times, when heart or lung function ceased, the brain also died within a few moments. Conversely, when the brain died, heart and lung function ceased as well. Today, however, the capabilities of modern medical technology to support or replace failing organ functions has completely changed the inevitability of this cause-effect relationship.

In the past, there was no evidence of concern for a condition in which persons appearing to be alive were in actuality dead. On the contrary, the concern was uniformly for false positives. Yet it is just this concern, that of the false negatives, which is being examined by physicians, philosophers, theologians, and lawyers. The general public has also demonstrated growing concern for this issue.

There are four major factors which seem to demand the more precise definition of death: (1) modern medicine's technological ability to sustain life in the absence of spontaneous heartbeat or respiration; (2) the advent of successful organ transplantation capabilities, which creates a demand for viable organs from recently deceased donors; (3) the enormous expenditure of resources which would be wasted if persons who in actuality are dead are being treated medically as though they are alive; (4) respect for persons requiring that the distinction be made when death has occurred, and that the dead person is treated appropriately in the light of this change in status.

New concepts. Throughout history, the concept of death has been tied to one or possibly two factors: the loss of the soul from the body, or the loss of vital-fluid circulation in the body. Efforts to establish a more precise definition of death have moved beyond the spontaneous function of the respiratory and circulatory system. The body's capacity for integrating its function is identified as the essentially significant indication of life. In the suggested new concept of death, it is not the death of one or more organs which is definitive, but the death of the whole person. When the point is reached in which these essential integrative characteristics are irreversibly lost, death occurs.

A still further refinement is offered by those working toward a more precise definition. The human capacity for integration can be divided into two categories. The first is the integration of the internal bodily mechanisms which are governed unconsciously, such as respiration and the regulation of blood pressure. The second is the integration of the person with the social environment through consciousness. This capacity for consciousness permits self-reflection and self-awareness as well as social interaction with others. Thus, this further refinement would claim that when the irreversible loss of the capacity for consciousness occurs, that which is essential to life is lost and the person is then dead. An even more radical concept is that death is the irreversible loss of the capacity for social interaction. *See* CONSCIOUSNESS.

Locus of death. Depending upon the philosophical concept of death that has been accepted, it becomes more specifically apparent as to where to look to determine the loss of the essential characteristics of human life. Thus, in terms of a vitalist concept, death is the loss of the circulation of vital fluids, and it would be appropriate to examine and evaluate the function of the heart and lungs. If the concept of integrative function is accepted as being the essential characteristic, it would be appropriate to evaluate the total brain. If, on the other hand, death is the loss of the embodied capacity for social interaction,

then all evidence today points to the neocortex as the locus of consciousness and social interaction.

New criteria. Whereas the concept of death is a philosophical issue, the determination of the irreversible loss in the relevant locus of the body is primarily the task of the biomedical sciences.

An interdisciplinary committee at the Harvard Medical School accepted the task of examining possible new criteria for determining death. In 1968 they cited four criteria as determinative of a permanently nonfunctioning brain: unreceptivity and unresponsivity to externally applied stimuli and inner need; absence of spontaneous movements of breathing; absence of reflexes other than spinal column reflexes, and fixed dilated pupils; and a flat electroencephalogram (EEG) to be used as a confirmatory but not mandatory test.

These tests were to be repeated after 24 h, and in the absence of hypothermia or depressant drugs, were to be considered a confirmation of brain death. It has been conclusively demonstrated that when the flow of blood to the brain has been interrupted for approximately 15 min there is a uniform necrosis and liquefaction of the brain. There has been no documented example in which the Harvard criteria have been shown to be invalid.

Ethical issues. These new criteria for death determination have important consequences of an ethical and social nature; there is concern for the relationship of the criteria to organ transplantation and the role of physicians in establishing the new criteria, and fears concerning future updating of the criteria.

With regard to the issue of organ transplantation, it is contrary to moral intuitions that a dying person might be prematurely defined as dead in order that someone else might receive the organs. Yet, it has been claimed that because anencephalic infants born alive cannot survive because they lack a brain, they should be declared dead and placed on a respirator to maintain circulation, so that any transplantable organs that are needed to sustain the lives of other infants can be removed. However, this is not practical because it is held that a moral obligation exists to treat the living as such until death occurs. But this does not rule out the possibility that a more precise definition of death is needed for the reasons mentioned earlier.

To avoid a conflict in the role of the physician, it has become common practice that the physician for the recipient of an organ donation is not party to the judgment that the donor has died. The needs of a patient for a viable organ must not contaminate the judgment that another person has died. Many states however, have drafted "required request" laws that obligate the hospital to develop a policy to ensure that the opportunity for the family to donate the organs of the newly deceased is not overlooked.

In light of the growing demand for organs suitable for transplant, it has been suggested that the criteria for death be changed. Some physicians and philosophers believe that the Harvard criteria are too stringent. They would, in most cases, favor a definition of cerebral death, a condition in which the neocortex shows a flat (isoelectric) reading in an EEG. Falling into this category would be those who view consciousness and the capacity for social interaction as the essential characteristics of being a living human. It was reported that this criterion for death was used in the former Soviet Union. There is an apparent reluctance to accept this criterion in the West, and the division is largely attributable to the philosophical debate about whether the human being is to be characterized in terms of the biological organism or as a rational and personal being. Thus, there is by no means agreement on this issue, and many physicians and others find it morally abhorrent to consider a person clinically dead according to the neocortex definition when there is still spontaneous breathing and circulation.

Social and legal responses. In addition to the ethical issues stemming from the articulation of a new definition of the criteria for determining death, there have been legal and social implications. Many states, either by legislation or by court opinion, have determined that the irreversible loss of brain function can be used as legal evidence of death.

In 1977 the American Medical Association's House of Delegates adopted the position that "permanent and irreversible cessation of the brain constitutes one of the various criteria which can be used in the medical diagnosis of death."

In August 1978 the National Conference of Commissioners on Uniform State Laws approved a Uniform Brain Death Act and recommended it to the states for adoption. It reads as follows: "For legal and medical purposes, an individual who has sustained irreversible cessation of all functioning of the brain, including the brain stem, is dead. This judgment is to be made in accordance with reasonable medical standards."

Prolonging life. The ability to define death with more precision is important in cases involving the irreversibly comatose individual. It is generally accepted as both a moral and legal right that competent persons have the right to refuse medical treatment. The law of informed consent is based upon this right, and the law has upheld the right of competent adults to decline treatment, even at the cost of their lives. Increasingly, dying individuals claim a more active role in self-determination regarding medical treatment in the final stage of life.

Every state and the District of Columbia have passed laws allowing individuals to either develop a living will or name an agent to hold durable power of attorney for health care decisions at the end of life, should they lose capacity for decision making. The intent is generally to provide the opportunity for individuals to extend their rational choice to refuse certain types of life-prolonging treatment. A signed directive is filed for this purpose and becomes authoritative if the individual is in an unconscious state and death is deemed imminent.

A federal law known as the Patient Self-Determination Act requires all institutions receiving federal funds to notify individuals at the time of

admission to a hospital, hospice, nursing home, or retirement facility of the right to a Natural Death Act Directive (living will), and to designate a person to hold durable power of attorney for health care decisions. While the establishment of new public policy may be helpful in some cases, it does not resolve all of the issues at stake. Many physicians believe a Natural Death Act is helpful, but primarily as an encouragement to opening communication within families. Such a law leaves unanswered the question as to whether persons who fail to sign such a directive beforehand might be at greater risk to have their lives prolonged against their will than before the law. Also, what of special cases such as that of a renal dialysis patient who finds life on the kidney machine intolerable, but for whom death is not imminent if treatment continues?

Even before these legislative innovations, the issue received considerable attention from ethicists and theologians. Pope Pius XII ruled that there was not a moral duty to prolong life in the face of imminent death when the provision of such treatment would be difficult to obtain, dangerous, painful, unusual, or costly and burdensome to the individual or the family. Herein is a moral persuasion, an appeal not to law but to the principles of reasonableness in the light of the prognosis and the relevant human factors in the case. Thus, when the burdens of treatment outweigh the benefits for the individual, it is considered morally correct to stop treatment. Court opinion has often upheld the lawfulness of discontinuing the use of respirators, tube feedings, and hydration.

Euthanasia. Finally, one of the more radical issues is that of actively ending the life of a suffering person or helping another to die through a positive action. Some ethicists see no moral distinction between removing life-support systems or lifesaving medications which allow a patient to die and taking a direct action which causes a patient to die, since the intention and the consequence are the same in both cases. Others disagree, claiming that there is indeed a morally significant difference if a person dies of the underlying disease and is not made to endure still greater suffering by prolonging the dying process than if the person dies because another has caused the death through an intentional and fatal action.

Two events have occurred regarding the controversial issue of physician-assisted suicide. A federal district judge held that the state law making assisted suicide a class C felony was unconstitutional. The ruling is limited to cases involving mentally competent, terminally ill patients who are not under undue influence and who wish to have physician assistance in committing suicide. The second major shift in public policy occurred in Oregon and provides legal protection for physicians who assist dying patients to commit suicide.

Opponents of assisted suicide and active euthanasia are not convinced that removing legal prohibitions are good for society. They prefer to work toward providing suffering individuals with adequate pain control and comfort measures during the last stages of life. They argue that even if in some cases a moral justification could be found for actively ending the life of a suffering dying individual the social consequences that would arise from the general tolerance of such a practice would be more problematic than good. Thomas R. McCormick

Bibliography. I. Bowen and R. Lockshin, *Cell Death in Biology and Pathology*, 1981; D. Carnevali and M. Patrick (eds.), *Nursing Management for the Elderly*, 3d ed., 1993; D. W. Myers, *Medico-Legal Implications of Death and Dying, Supplement*, 1981; M. Patrick et al. (eds.), *Medical-Surgical Nursing*, 2d ed., 1991; W. J. Smith, *Dying in the Human Life Cycle*, 1985.

De Broglie wavelength

The wavelength $\lambda = h/p$ associated with a beam of particles (or with a single particle) of momentum p; $h = 6.63 \times 10^{-34}$ joule-second is Planck's constant. The same formula gives the momentum of an individual photon associated with a light wave of wavelength λ. This formula, along with the profound proposition that all matter has wavelike properties, was first put forth by Louis de Broglie in 1924, and is fundamental to the modern theory of matter and its interaction with electromagnetic radiation. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS. Edward Gerjuoy

Decapoda (Crustacea)

One of the more highly specialized orders of the subphylum (or class) Crustacea. This order includes the shrimps, lobsters, hermit crabs, and true crabs. The order is so diverse that satisfactory definition is difficult, but a few characters are common to nearly all decapods. The most obvious is a head shield, or carapace, which covers and coalesces with all of the thoracic somites and which overhangs the gills on each side. The first three of the eight pairs of thoracic appendages are specialized as maxillipeds and closely associated with the true mouthparts. The last five pairs are walking legs or chelae (claws); this feature is the basis of the name Decapoda (= ten feet). The gills are usually well developed and arranged in several series.

Adult decapods vary in size from less than 1/2 in. (13 mm) in length to that of the giant Japanese crab (*Macrocheira*), the largest living arthropod, a spider crab which may span more than 12 ft (3.6 m) between the tips of the outstretched claws. Although most decapods are found in the sea, they are by no means restricted to that habitat. Crayfishes are well-known inhabitants of freshwater streams and ponds, as are several kinds of shrimps and some true crabs. A number of crabs and hermit crabs have become well adapted to a terrestrial existence far from water; they return to the sea only seasonally to hatch their eggs. Many crayfishes burrow in the ground, and one

species of crab spends its entire life in the tops of lofty trees.

In a general way, the decapods may be divided into two groups: the long-tailed and the short-tailed forms. The long-tailed species, such as the shrimps and lobsters, have a more or less cylindrical, or laterally compressed, carapace that often bears a head spine or rostrum. The large, muscular abdomen permits the shrimp or lobster to dart quickly backward out of danger, but the succulence of the tail makes the animals prey to numerous enemies, including humans. In the short-tailed species, the crabs, the carapace is often broadened and flattened dorsally and usually does not form a rostrum. The much-reduced and feeble abdomen is tucked under the thorax, where it serves the female as a brood pouch for the eggs.

Appendages

Like other malacostracans, nearly all decapods have pairs of appendages on each of 19 segments: 5 pairs on the head, 8 on the thorax, 6 on the abdomen (described in sequence from anterior to posterior ends). Each of the first pair of antennae (also called antennules) has a three-segmented peduncle and, typically, two flagella, the outer one being bifurcated near the base in some shrimps. The second antennae have up to five segments in the peduncle, a single flagellum, and often a scalelike outer branch or exopod (**Fig. 1**).

Mouth. There are six pairs of mouthparts in the decapods, the posterior three pairs being modified thoracic legs. The mandibles sometimes have distinct incisor and molar processes. These are sometimes indistinguishably fused, or the incisor process may be lacking. There is usually a three-segmented palp (**Fig. 2**). The first maxillae (maxillulae), second maxillae (maxillae), and the first maxillipeds are usually foliaceous. The second maxillipeds are also composed of broad, flattened segments, but they are less modified from the typical thoracic appendages than the first maxillipeds. The third maxillipeds are elongate and leglike in the shrimps but are greatly modified, forming an operculum over the preceding mouthparts in most of the crabs.

Thorax. The five remaining pairs of thoracic appendages, the pereiopods (or pereopods), are basically walking legs although one or more pairs may be variously modified. The most common modification is seen in the formation of pinching chelae. In the long-tailed decapods, one to three (rarely four or all five) pairs of pereiopods may be chelate. In the short-tailed forms, only the first pair is usually thus modified. Some of the pereiopods may be broadened and flattened as swimming paddles, especially in the swimming crabs. In the shrimps, the pereiopods are basically composed of seven segments, of which some are occasionally multiarticulate. In most of the other decapods, however, the second and third segments are fused. Occasionally, one or two pairs of pereiopods are vestigial or entirely lacking.

Abdomen. The first five pairs of abdominal appendages, the pleopods, are swimming organs in

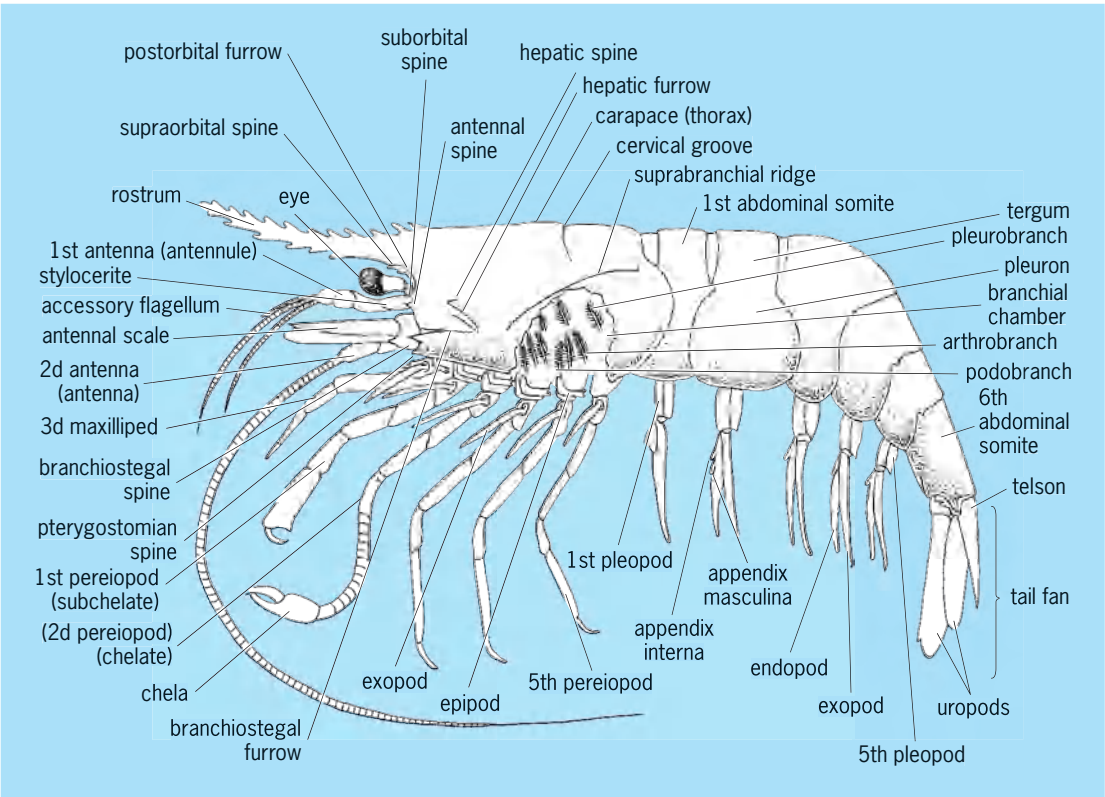

Fig. 1. Lateral view of caridean shrimp showing external morphology.

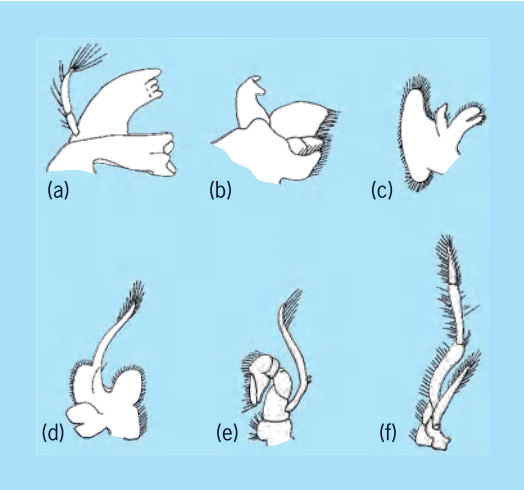

Fig. 2. Mouthparts of caridean shrimp, *Palaemon*. (*a*) Mandible showing incisor and molar processes and palp. (*b*) First maxilla (maxillule). (*c*) Second maxilla (maxilla). (*d*) First maxilliped. (*e*) Second maxilliped. (*f*) Third maxilliped.

most of the long-tailed decapods, but the first one or two pairs are usually modified as sexual appendages (gonopods) in the male and are often reduced and sometimes absent in the female. In the short-tailed forms, the pleopods (except those in the male which are modified as male sexual organs) are usually reduced or modified in the female for the attachment of eggs. In both sexes of the long-tailed species, the last pair of abdominal appendages, the uropods, forms a tail fan with a terminal median structure, the telson. In the hermit crabs, the uropods are thickened and twisted to hold the animal in the gastropod shell or other structure chosen as a shelter. In practically all of the crabs, the uropods are absent.

Nervous System and Sense Organs

The nervous system is somewhat variable. Although usually fewer, there may be as many as 11 ganglia, 5 thoracic and 6 abdominal, in the ventral nerve chain posterior to the subesophageal ganglion (**Fig. 3**). In most of the crabs, all of the ventral ganglia are concentrated in a rounded mass in the thorax, from which the nerves radiate outward.

The sense organs include a pair of compound eyes and statocysts, olfactory filaments, and tactile setae. The eyes are well developed in most decapods, but they may be reduced or entirely absent in species living in the deep sea, in caves, or in burrows. They are borne on limblike eyestalks located near the antennae and composed of two, rarely three, segments, either of which may be abnormally elongate. The cornea is usually distinctly faceted, the facets being square or hexagonal. The statocysts, or balancing organs, occur on the first segment of the antennules. They are cavities, either open or closed to the exterior, lined with sensory hairs. If open, they usually contain foreign particles, such as sand grains. If closed, they may contain secreted bodies, but usually there is no inclusion. Olfactory filaments are especially numerous on the outer flagellum of the antennules. Tactile setae, or hairs, occur on various parts of the body and appendages.

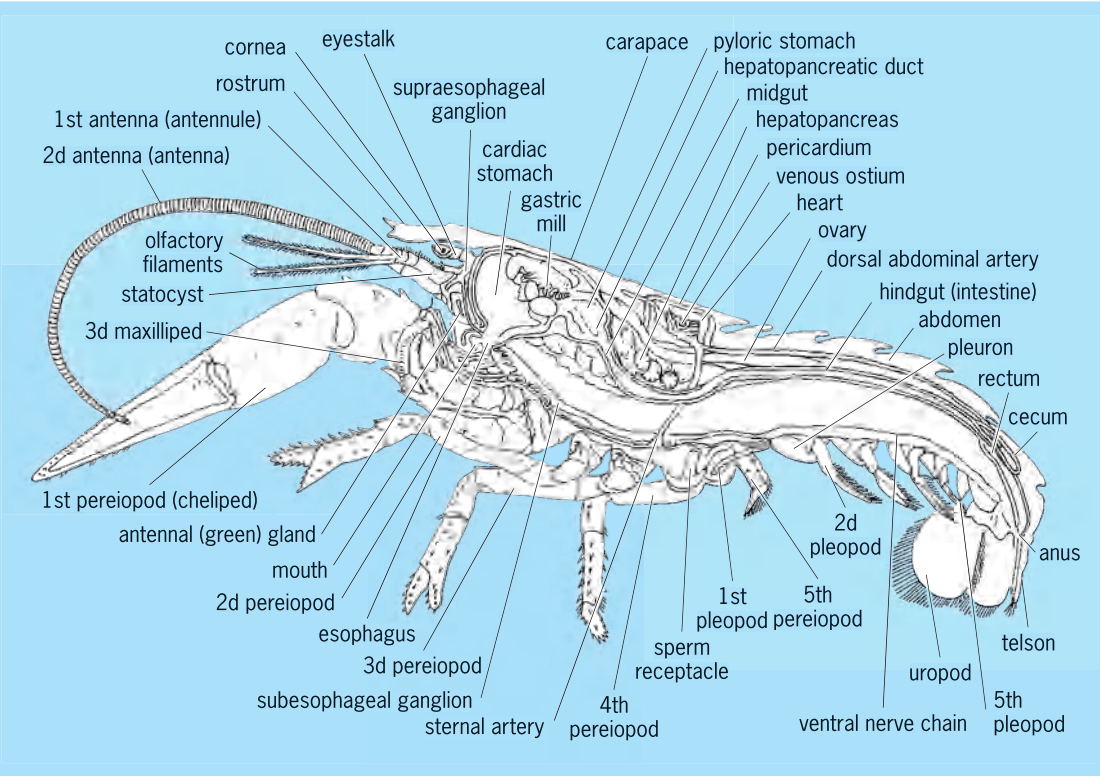

Fig. 3. Median section of lobster, *Homarus*, showing general anatomy.

Alimentary and Circulatory System

As in most crustaceans, the alimentary tract is a more or less straight tube divided into three parts: the foregut or stomodeum, the midgut or mesenteron, and the hindgut or proctodeum. The first and last have a chitinous, and sometimes calcified, lining. The foregut, in most decapods, is dilated to form a stomach. Its anterior, or cardiac, portion is at least partially lined with movably articulated, calcified ossicles, controlled by a complex system of muscles called the gastric mill. The posterior, or pyloric, part is provided with hairy ridges which combine to form a straining apparatus. The midgut is variously furnished with blind tubes, and a mass of minutely ramified tubules, known as the hepatopancreas or liver, empties through one or more ducts into the anterior end of the midgut.

The heart is short and polygonal, and situated under the posterior part of the carapace. It has three or more pairs of ostia (holes with valves) and two pairs of arteries lead from it, in addition to the median anterior and posterior arteries. The blood collects in a venous sinus in the midventral line, passes to the gills, and then back to the pericardial sinus surrounding the heart.

Respiration

In practically all decapods, respiration is accomplished by a series of gills, attached to the lateral walls of the thoracic somites and to the basal segments of the thoracic appendages. A complete set of gills consists of the pleurobranchs, attached to the lateral walls of the thoracic somites; the arthrobranchs, fastened to the articular membranes between the body wall and the first segment of the appendage; and podobranchs, attached to the first segment of the appendage. The gills (**Fig. 4**) may be one of three types: trichobranchiate, with filamentous branches arranged in several series around the axis; phyllobranchiate, with flattened branches usually arranged in two opposite series; and dendrobranchiate, with extensive, sometimes complex, branching of the two primary series.

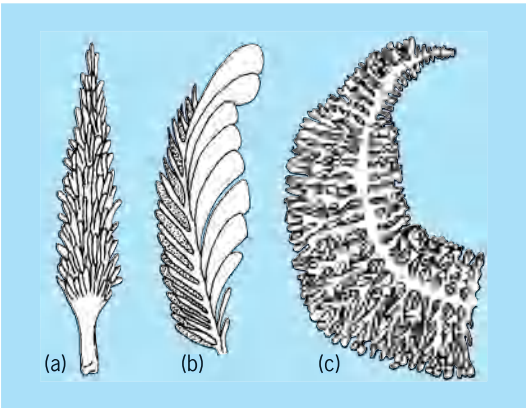

Fig. 4. Decapod gills. (*a*) Trichobranchiate gill of stenopodid shrimp, *Spongicola*. (*b*) Phyllobranchiate gill of mud shrimp, *Thalassina*. (*c*) Dendrobranchiate gill of penaeid shrimp, *Benthesicymus*.

A continuous current of water is drawn through the branchial chamber and over the gills by the fluttering motion of the outer lobe (scaphognathite) of the second maxilla. The current usually moves forward, but this direction may be periodically reversed in burrowing species. In the latter, also, the exhalant channels may be extended anteriorly by the apposition of the left and right elements of certain appendages, such as the first and second antennal flagella, the antennal scales, the mandibular palps, and the third maxillipeds.

In those crabs best adapted for life on land, the action of the gills is supplemented by the lining membrane of the branchial chamber. This chamber is covered with minute villi and is unusually well supplied with blood vessels, and thus functions as a lung. A similar highly vascularized area occurs in the delicate abdominal skin of the terrestrial hermit crabs.

Excretion

The antennal or "green" gland is the chief excretory organ in all decapods. Maxillary glands are often present in the larval stages, but they never persist in the adult. The antennal gland is compact and more complex than in any other crustacean group. It is often divided into regions: the saccule (end sac), which is usually partitioned or ramose; the labyrinth, a spongy mass with a complex system of folds; and a coiled canal and bladder with a duct leading to the exterior (**Fig. 5**). The bladder may be a simple vesicle

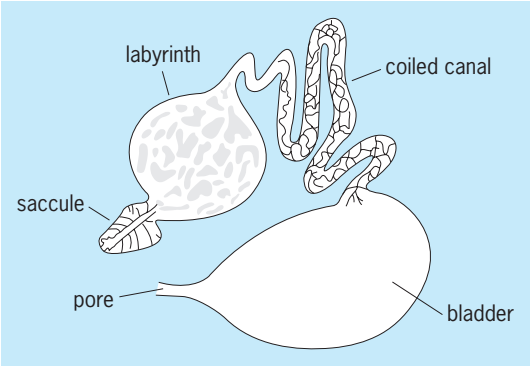

Fig. 5. Antennal gland of crayfish (section).

with a short duct, as in the crayfish; or it may be variously lobate, as in most of the shrimps, crabs, and hermit crabs. In the latter case, two diverticula from the bladder extend through the entire abdomen. The external opening is usually in a small elevation on the first segment of the second antennae. In the crabs, the aperture is covered by an operculum which may be opened and closed.

Reproductive System

The sexes are separate in most decapods, although protandric hermaphroditism (in which the animal develops first as a male and subsequently changes to a female) occurs in some shrimps. The testes most often lie partly in the thorax and partly in the abdomen, and they are usually connected across the

midline. In some hermit crabs, however, they lie wholly within the abdomen, and they may be either unconnected or fused into a single organ. The vas deferens usually opens on the first (most proximal) segment of the last pair of thoracic legs, but in some crabs it perforates the last thoracic sternum. The ovaries resemble the testes in shape and position. The oviducts usually open on the first segment of the third pereiopods; in most crabs, however, they open on the sternum.

Development

In only one group of shrimps (suborder Dendrobranchiata) are all of the presumably primitive larval stages represented (**Fig. 6**), some with multiple instars separated by molts. The first stage is a typical crustacean nauplius with an unprotected oval body bearing a median simple eye and three pairs of appendages that are used in swimming. This stage is followed by a metanauplius, in which four more pairs of appendage rudiments appear. The third pair of the original appendages becomes less of a swimming organ and more like a pair of primitive mandibles. In the third stage, the protozoea, the seven pairs of appendages become more highly developed, a carapace covers the anterior part of the body, the abdomen is clearly formed though unsegmented at first, the rudiments of paired eyes appear, and the heart is formed. In the next stage, the zoea, the eyes become movable and the carapace develops a rostrum. All thoracic appendages are present, at least in rudimentary form, and those of the abdomen appear, especially the uropods. In later zoeal instars (mysis stage), the well-developed thoracic

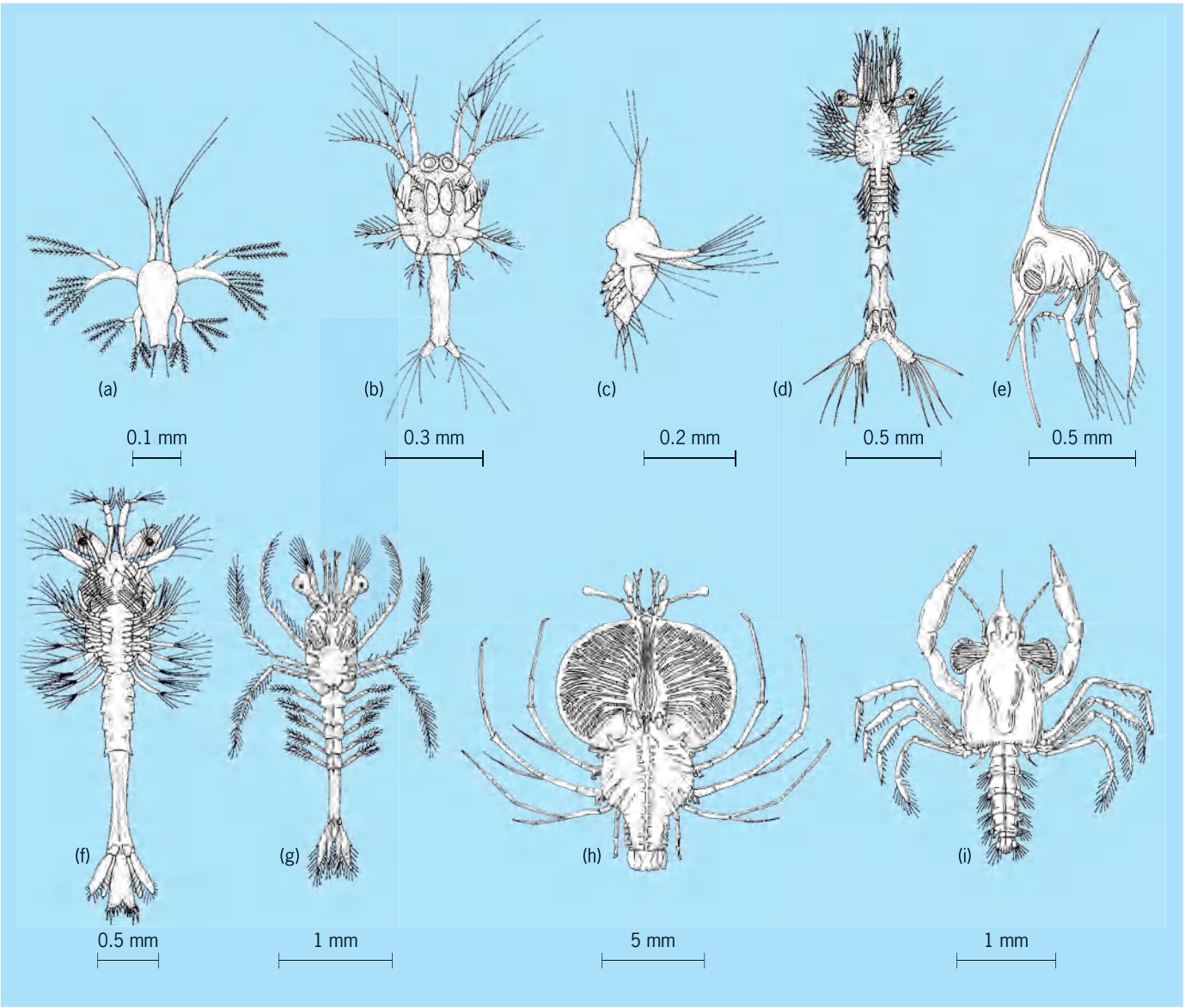

Fig. 6. Decapod larval stages. (*a*) Nauplius larva of penaeid shrimp (dorsal view). (*b*) Protozoea larva of penaeid shrimp. (*c*) Metanauplius larva of penaeid shrimp (lateral view). (*d*) Metazoea larva of penaeid shrimp. (*e*) Zoea larva of crab. (*f*) Mysis larva of penaeid shrimp. (*g*) Mysis (acanthosoma) larva of sergestid shrimp. (*h*) Phyllosoma larva of scyllarid lobster. (*i*) Megalopa larva of swimming crab.

appendages replace the antennae as the chief swimming organs. The abdomen increases markedly in size and takes on a form similar to that of the adult.

In most decapods (constituting the suborder Pleocyemata), the nauplius stage is passed while in the egg, and the animals hatch as zoeae, often molting several times as they differentiate and grow. The last zoeal instar metamorphoses into a postlarva or megalopa. This is similar to the adult, except that all appendages are present, the abdomen is prominent, and the pleopods are used for swimming.

Decapods that live in deep-water or polar habitats tend to exhibit a reduction in the number of larval instars; in freshwater species there may be none. Most freshwater shrimps hatch in a late zoeal (mysis) stage, and the crayfishes and freshwater crabs hatch as miniature adults. Terrestrial forms, on the other hand, typically migrate to the sea to hatch their eggs, which then pass through the usual larval stages of marine species.

Phylogeny

No true decapod fossils are known with certainty from Paleozoic deposits, but decapod shrimps are not uncommon in Triassic and, especially, Jurassic rocks. Lobsterlike forms related to Recent deep-sea species also appeared in the Triassic; true lobsters are known from the Jurassic, and crayfishes from the Cretaceous. Crabs are well known as fossils, with primitive groups appearing in the Jurassic; several more specialized groups are found in the Cretaceous, and all of the major existing groups were represented in the Tertiary. *See* CRETACEOUS; JURASSIC; TERTIARY.

The evolution of the decapods has not been satisfactorily worked out and may never be known completely, but several hypotheses based on cladistic analysis of morphology and deoxyribonucleic acid (DNA) sequences have been put forth recently. The presence of a nauplius larva in the development of the dendrobranchiate shrimps suggests a link with the more primitive malacostracan crustaceans, such as the Euphausiacea (krill), but it can hardly be doubted that the Decapoda, as a group, are highly specialized.

Classification

The classification of the approximately 18,000 living species of decapods presents a difficult problem that has not yet been entirely solved. The division of the order into long-tailed (suborder Natantia) and short-tailed (suborder Reptantia) groups, alluded to previously, was in favor in the eighteenth century. This scheme was generally abandoned, however, when it was discovered that some of the long-tailed forms (known in older classifications as section Macrura, including the lobsters and crayfish) were obviously more closely related to the short-tailed group than they were to other long-tailed species. The chief disparity seems to lie between the shrimps that possess dendrobranchiate gills and all of the other decapods (Pleocyemata), in which the gills are either trichobranchiate or phyllobranchiate. The current, generally accepted classification scheme, follows.

Order Decapoda
- Suborder Dendrobranchiata
- Suborder Pleocyemata
  - Infraorder: Caridea
    - Stenopodidea
    - Palinura
    - Astacidea
    - Thalassinidea
    - Anomura
    - Brachyura

Dendrobranchiata. This suborder (equivalent to the section Penaeidea of older classifications) contains the superfamilies Penaeoidea and Sergestoidea, which make up most of the shrimps caught or "farmed" worldwide (**Fig. 7**). Collectively they are distinguished by the following characteristics: the pleura (skeletal plates along the sides) of the first abdominal somite overlap those of the second; the third legs are nearly always chelate but are no stouter than the first pair; the first pleopods of the male bear a complicated flaplike appendix, the petasma, while the females have a spermatophore receptacle, the thelycum, on the ventral surface of the posterior thoracic somites. Unlike those of all other decapods, the

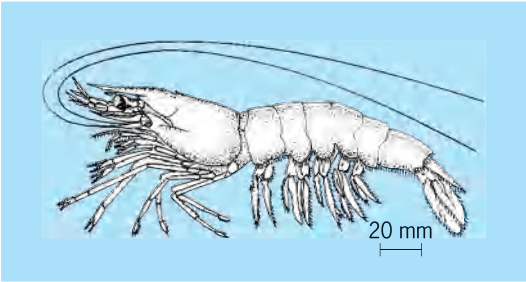

Fig. 7. Penaeidean shrimp, *Litopenaeus setiferus*, a primitive decapod.

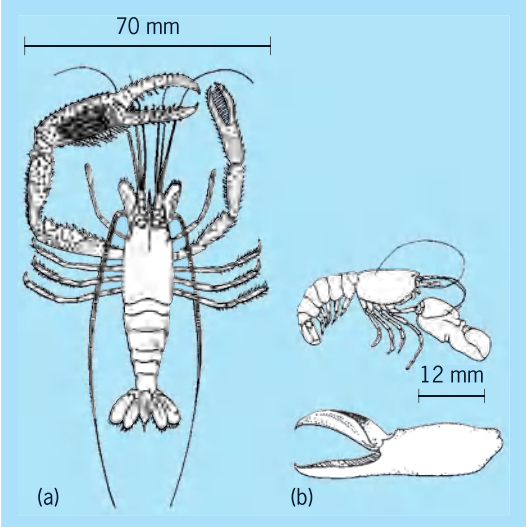

Fig. 8. Caridean shrimp. (*a*) River shrimp, *Macrobrachium faustinum*. (*b*) Snapping shrimp, *Alpheus heterochaelis*, with minor claw enlarged.

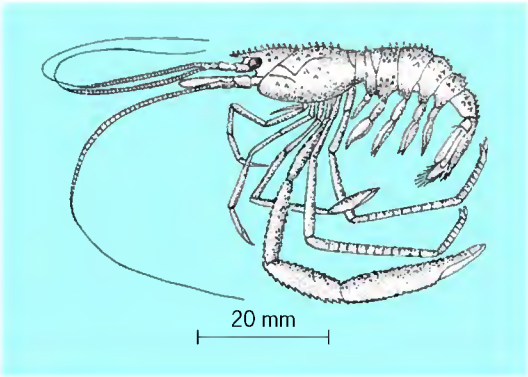

Fig. 9. Stenopodid shrimp, *Stenopus hispidus*, a marine decapod.

eggs are not attached to the abdominal appendages but are almost always shed free into the water. The Penaeoidea share with the Euphausiacea the presence of a petasma; in both groups, the eggs are rarely carried, and they hatch at an early larval stage, usually as nauplii. The best-known penaeoideans, including most of the commercially important shrimps or prawns of the warmer seas (family Penaeidae), live on muddy bottoms in shallow or moderate depths. However, they also occur both in the deep sea and as pelagic organisms in the middepths (around 500 m or 1640 ft). In this latter region are found the Sergestidae, in which the last two pairs of thoracic legs are reduced or absent; some have luminescent organs (photophores). There are

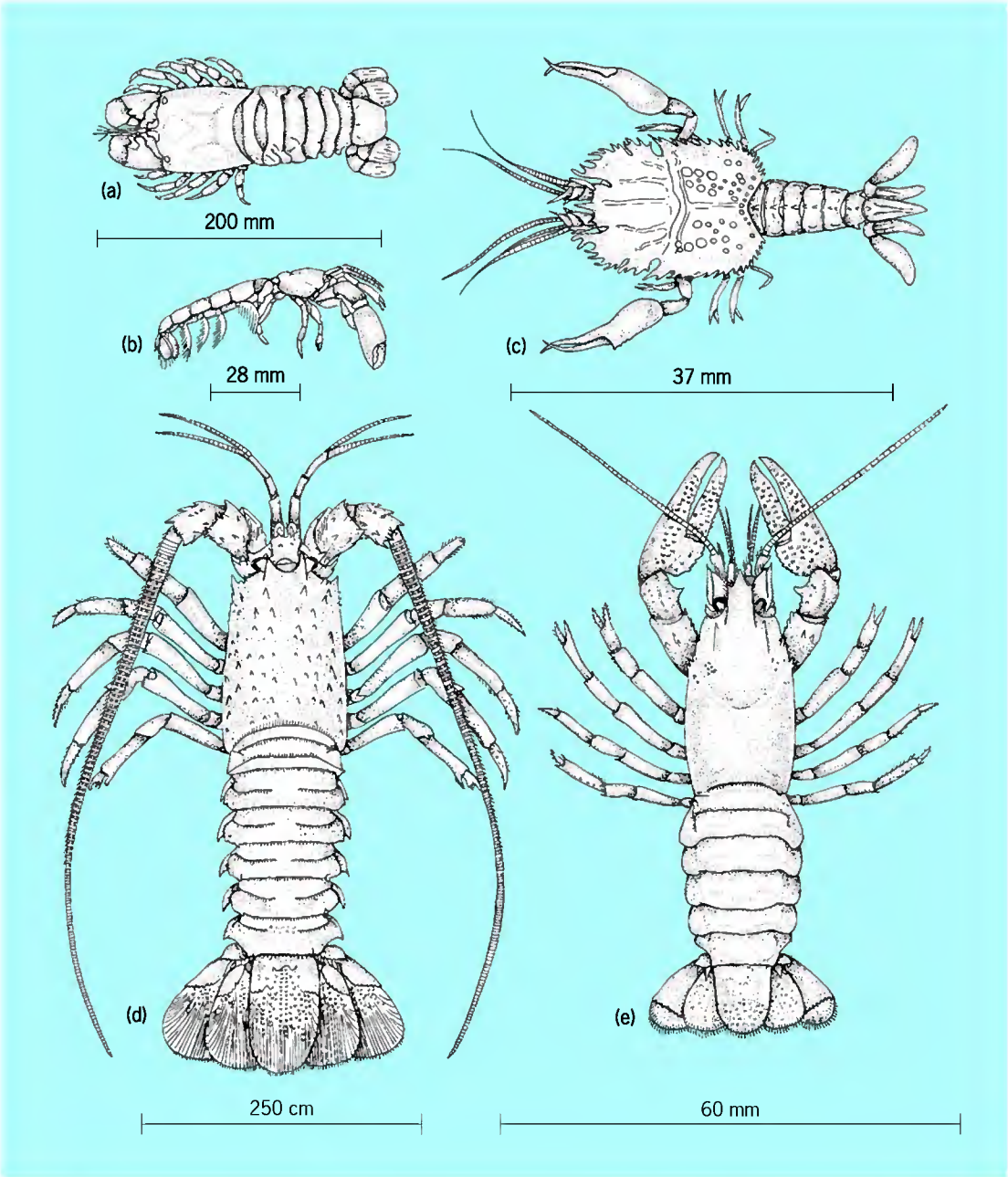

Fig. 10. Macruran decapods. (*a*) Spanish lobster, *Scyllarides aequinoctialis*. (*b*) Mud shrimp, *Callianidea laevicauda*. (*c*) Eryonid, *Polycheles crucifer*. (*d*) Spiny lobster, *Panulirus interruptus*. (*e*) Crayfish, *Orconectes limosus*.

about 500 species of living Dendrobranchiata.

Pleocyemata. This suborder contains the infraorders Caridea, Stenopodidea, Palinura, Astacidea, Thalassinidea, Anomura, and Brachyura.

Caridea. This is the largest and most diverse group of shrimps and prawns. In this group, the pleura of the second abdominal somite overlap those of the first; the third legs are never chelate; there is no petasma or thelycum; and the gills are phyllobranchiate. The sexes can usually be distinguished by the presence in the male of two small stalks, an appendix masculina and an appendix interna, located on the inner edge of the inner branch (endopod) of the second pleopods; only the appendix interna occurs in the female. Carideans are found in all parts of the sea, often in association with other marine animals. Members of at least two families, the Atyidae and Palaemonidae, are also widespread in freshwater. The edible shrimps and prawns of northern Europe and North America, of the genera *Crangon* and *Pandalus*, belong to the Caridea, as do the large freshwater prawns of the tropical genus *Macrobrachium*. Some of the latter are the largest natantians known, attaining a length of more than a foot and having abnormally long second pereiopods, which when fully grown may be as long as the body (**Fig. 8***a*). The snapping shrimps, or Alpheidae, make a loud popping noise by cavitation, using a socket-and-plunger mechanism on an enlarged first pereiopod (claw) [Fig. 8*b*]. The resulting crackling noise from vast populations in tropic seas sometimes interferes with underwater sound equipment. Most shrimps and prawns are favorite foods of fishes, but some carideans maintain "service stations" for fishes where they remove parasites from the outer skin, mouths, and gills of the fish. More than 2800 species of Caridea are known.

Stenopodidea. This is a small group of shrimps which superficially resemble the Penaeoidea (**Fig. 9**). The third pereiopods are chelate but are much longer and stouter than the first pair; the pleura of the second abdominal somite do not overlap those of the first; there is no petasma or thelycum; and the gills are trichobranchiate. All of the 50 or more living species are marine; some are closely associated with sponges and corals.

Palinura, Astacidea, and Thalassinidea. The infraorders Palinura, Astacidea, and Thalassinidea, constituting the Macrura of older classifications, are characterized by an extended abdomen, well-developed tail fan, and trichobranchiate gills (**Fig. 10**). They are capable of swimming, as shrimps do, by flexion of abdominal muscles. However, like crabs, they are better adapted for crawling over the substrate, with strong legs, or burrowing into it.

Palinura is a traditional infraorder (which is probably not monophyletic) containing two major groups (superfamilies), Palinuroidea and Eryonoidea, and about 200 species. The Palinuroidea (Achelata in some contemporary classifications) include the spiny lobsters or langoustes (family Palinuridae) [Fig. 10*d*] and the Spanish, or shovel-nosed, lobsters (Scyllaridae) [Fig. 10*a*]. They are heavily armored like the true lobsters but are distinguished by the absence of a rostrum and of chelae, except occasionally on the last pereiopod of the female. They are abundant in shallow and moderate depths of warm and temperate seas, where they are often of considerable commercial importance. Frozen tails of spiny lobsters are imported to the United States from South Africa, Australia, and New Zealand, and the fresh lobster meat is the basis of fisheries in the Caribbean, southern Europe, and Japan. The Eryonoidea (Polychelida in some recent classifications) are rather thin-shelled, blind inhabitants of the depths of the sea (Fig. 10*c*). The carapace is flattened dorsally and considerably expanded laterally. The first four, or all five, pairs of pereiopods are chelate, with the first pair much elongated. This group was probably more numerous in ancient seas.

Astacidea are the true lobsters and crayfishes, characterized by a rostrum and by chelae on the first three pairs of pereiopods, with the first pair being noticeably larger than the others (Fig. 10*e*). Most lobsters (Nephropidae) are found in cool seas or in the cool, offshore waters of the tropics. The most familiar lobsters (*Homarus*) are those found along the Atlantic coasts of Europe and North America. The Norway lobster of Europe (*Nephrops*) is also of some commercial importance but is smaller and less meaty. The crayfishes (Cambaridae, Astacidae, Parastacidae) are widespread in freshwaters of the temperate regions of all continents except Africa. They are of commercial importance in southern Europe and Australia. More than 700 living species of lobsters

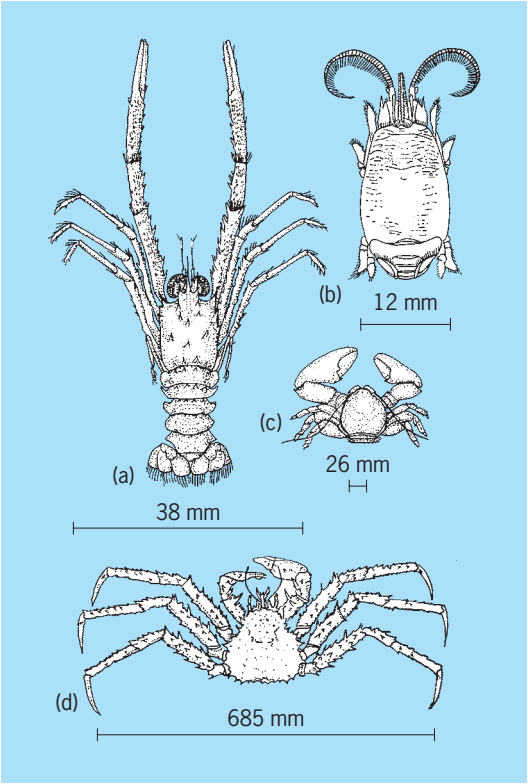

Fig. 11. Anomuran decapods. (*a*) Galatheid, *Munida evermanni*. (*b*) Mole crab, *Emerita talpoida*. (*c*) Porcellanid crab, *Petrolisthes tridentatus*. (*d*) King crab, *Lithodes maia*.

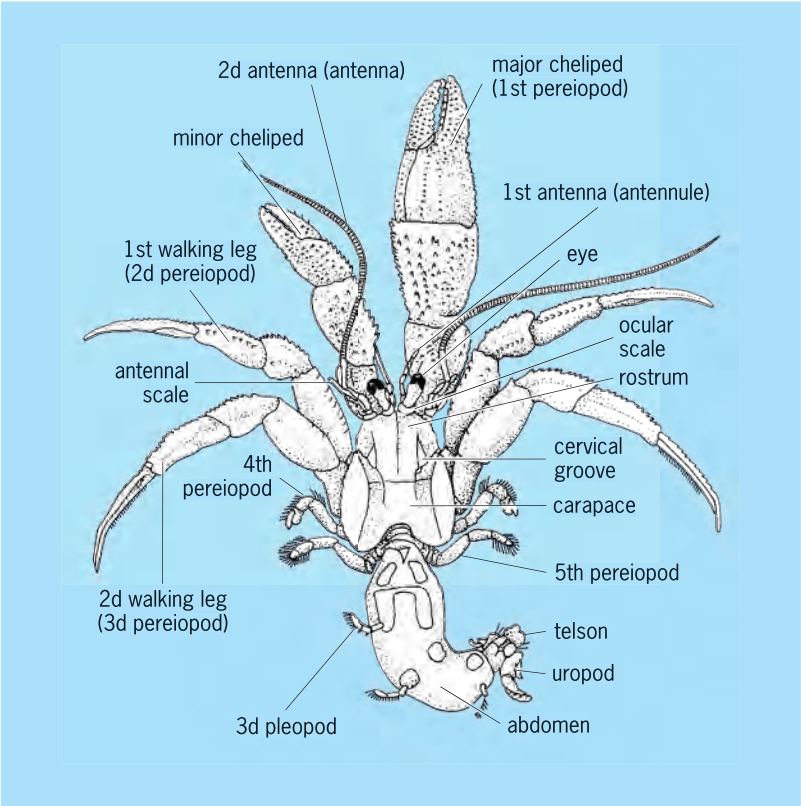

Fig. 12. Dorsal view of hermit crab, *Pagurus*, showing external morphology.

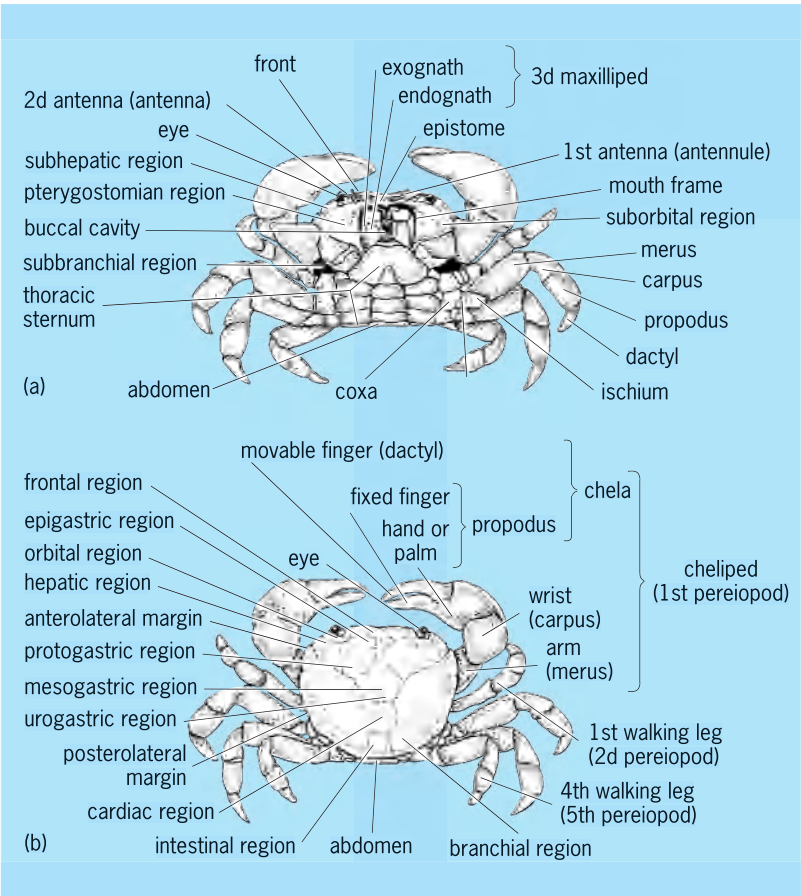

Fig. 13. External morphology of grapsoid crab. (*a*) Ventral view. (*b*) Dorsal view.

and crayfishes are known, more than half of them being from the freshwaters of the United States.

Thalassinidea are usually thin-shelled, burrowing crustaceans with large, chelate or subchelate first pereiopods, and no chelae on the third pereiopods (Fig. 10*b*). Known as mud or ghost shrimps (Upogebiidae and Callianassidae, respectively), the more than 500 living species are found in shallow and deep seas throughout the world, especially in tropical and warm temperate regions.

Anomura. This possibly unnatural group (renamed Anomala in some newer classifications) contains about 1900 species that are morphologically intermediate between the Macrura and the Brachyura (**Fig. 11**). It includes the galatheids and porcelain crabs (superfamily Galatheoidea), the hermit and king crabs (Paguroidea), and the mole crabs or hippas (Hippoidea). In nearly all of these diverse crustaceans, the first, and sometimes the last, pereiopods are chelate or subchelate. The abdomen is usually bent forward ventrally or is asymmetrical, soft, and twisted. Most anomurans possess phyllobranchiate gills.

The Galatheoidea include those anomurans with a symmetrical abdomen bent upon itself and provided with a well-developed tail fan. Most species are found in the sea at shallow or considerable depths, but one aberrant genus (*Aegla*) inhabits freshwater streams of temperate South America. The more primitive species (Galatheidae) are somewhat lobsterlike (Fig. 11*a*); the porcelain crabs or rock sliders (Porcellanidae) resemble the true crabs but may be distinguished by the much reduced and chelate fifth pereiopods and by the well-developed tail fan (Fig. 11*c*). They are most abundant in the intertidal zone and in moderate depths of the warmer seas; some species live in symbiotic relationships with other animals such as tube-dwelling polychaete worms and sponges.

In the hermit and king crabs (Paguroidea), the abdomen is nearly always asymmetrical, being either soft and twisted or bent under the thorax. The uropods, when present, as in the hermit crabs, are adapted for holding the body in a portable shelter. The first, and often the last, pereiopods are chelate. The hermit crabs (Paguridae and Diogenidae; **Fig. 12**) are found everywhere in the sea and are especially abundant in the intertidal zone. A few species (Coenobitidae) are also found on land in the tropics. Most hermit crabs live in the shells of dead gastropod mollusks which they must replace periodically as they grow. They sometimes place a sea anemone on the shell. The king crabs (Lithodidae) [Fig. 11*d*] are usually large crablike anomurans which live in offshore marine waters. They can be distinguished from the spider crabs, which they superficially resemble, by the reduced last pereiopods and by the asymmetrical arrangement of the abdominal plates in the female. The king crab of the North Pacific (*Paralithodes*) is of considerable commercial importance as food.

The third group of the Anomura, the Hippoidea (Fig. 11*b*), includes cylindrical or squarish

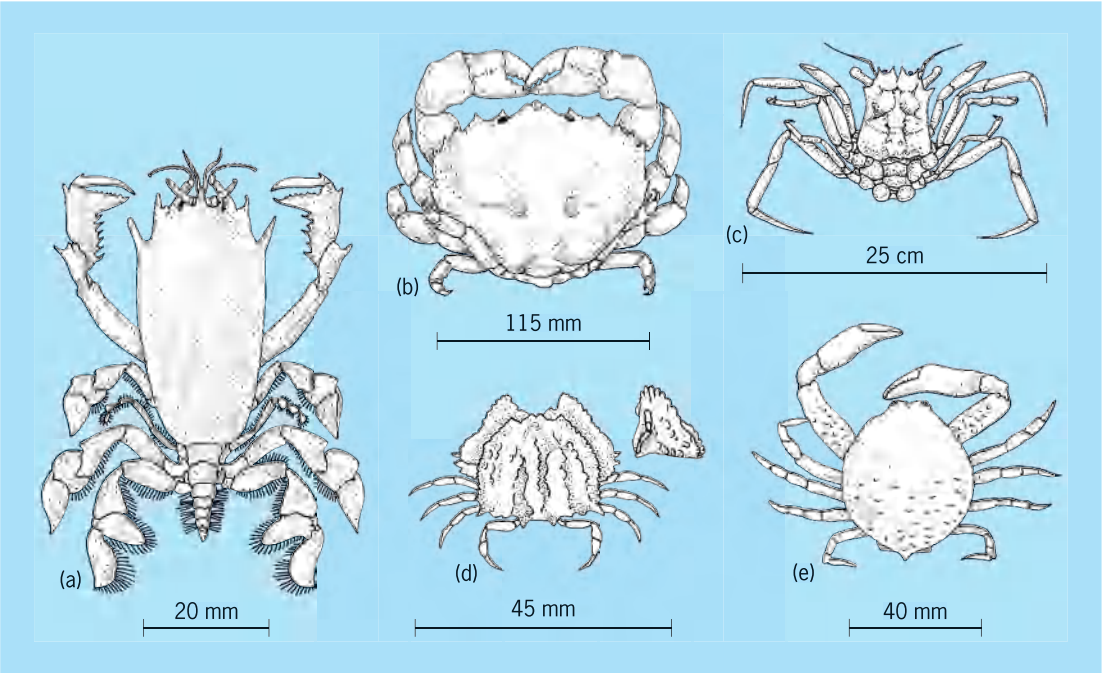

Fig. 14. Common brachyuran crabs. (*a*) Raninoid crab, *Raninoides louisianensis*. (*b*) Dromiid crab, *Dromia erythropus*. (*c*) Mask crab, *Ethusa mascarone americana*. (*d*) Box crab, *Calappa sulcata*. (*e*) Purse crab, *Persephona punctata*.

burrowing crustaceans in which the abdomen is symmetrical and bent under the thorax; the tail fan, if present, is not adapted for swimming; and the first pereiopods are either styliform or subchelate. Common hippas live in sand in the surf zone of tropical and temperate shores, and they travel up and down the beach with the tide. Using their flattened legs, they quickly burrow their entire body into the sand except for their feathery antennae, which strain food particles as the water of receding waves washes over them.

Brachyura. These are the true crabs (**Figs. 13**, **14**, and **15**), with dorsoventrally flattened and widened bodies. The abdomen is symmetrical, without a tail fan, and bent under the thorax. The first pereiopods are always chelate or subchelate, and the gills are phyllobranchiate. The true crabs are as numerous as all other decapods combined, numbering more than 10,000 species. They are divided into about 70 families, grouped principally on the basis of the location of the male and female gonopores (openings of the reproductive system).

Two primitive groups (collectively known as the Podotremata, with gonopores at pereiopod bases) are the Raninoidea (equivalent to the Gymnopleura of older classifications) and the Dromiacea. The carapaces of raninoids are more or less trapezoidal or elongate, the first pereiopods are subchelate, and some or all of the remaining pereiopods are flattened and expanded for burrowing (Fig. 14*a*). In dromiid crabs (Fig. 14*b*), the first pereiopods are chelate; the last pair is dorsal in position and modified for holding objects (such as sponges, tunicates, and bivalve mollusk shells) over the crab. The mouth frame is quadrate.

Most families of crabs are contained within the subsection Heterotremata, defined by the presence of female gonopores on the thoracic sternum rather than on the legs (as in males). In three relatively primitive families, collectively known as the Oxystomata in older classifications, the first pair of pereiopods are chelate and the last pair are either normal or modified as in the Dromiacea. The mouth frame is triangular and forward, over the epistome. The oxystomes include the mask crabs (Dorippidae), box crabs (Calappidae), and purse crabs (Leucosiidae) [Fig. 14*c*, *d*, *e*].

In the remaining heterotremate families (including, in part, the Oxyrhyncha and Brachygnatha of older classifications), the mouth frame is quadrate and the last pereiopods are rarely reduced or dorsal in position. To this group belong the spider, or decorator, crabs (Majidae) [Fig. 15*a*], slow-moving animals that often conceal themselves by attaching seaweeds and various sessile animals to their carapaces. Some cancroid crabs (for example, *Cancer*) are large and abundant enough to be of commercial importance in northern Europe and North America. The rock, or Jonah, crabs of New England and the Dungeness crab of the Pacific Coast are familiar edible crabs of this genus. In the swimming crabs (Portunidae) [Fig. 15*f*], the last pereiopods are modified as swimming paddles; a well-known example is the edible blue crab (*Callinectes*) of the shallow waters of the tropical and temperate Americas. The ubiquitous little mud crabs (Xanthidae) [Fig. 15*e*], well known to every visitor to rocky shores, are characterized by strong, often black-tipped claws. The stone crab (*Menippe*), another xanthoid (Fig. 15*g*), is highly esteemed by connoisseurs of seafood. Also worthy of mention are freshwater crabs (Potamidae, Pseudothelphusidae, and

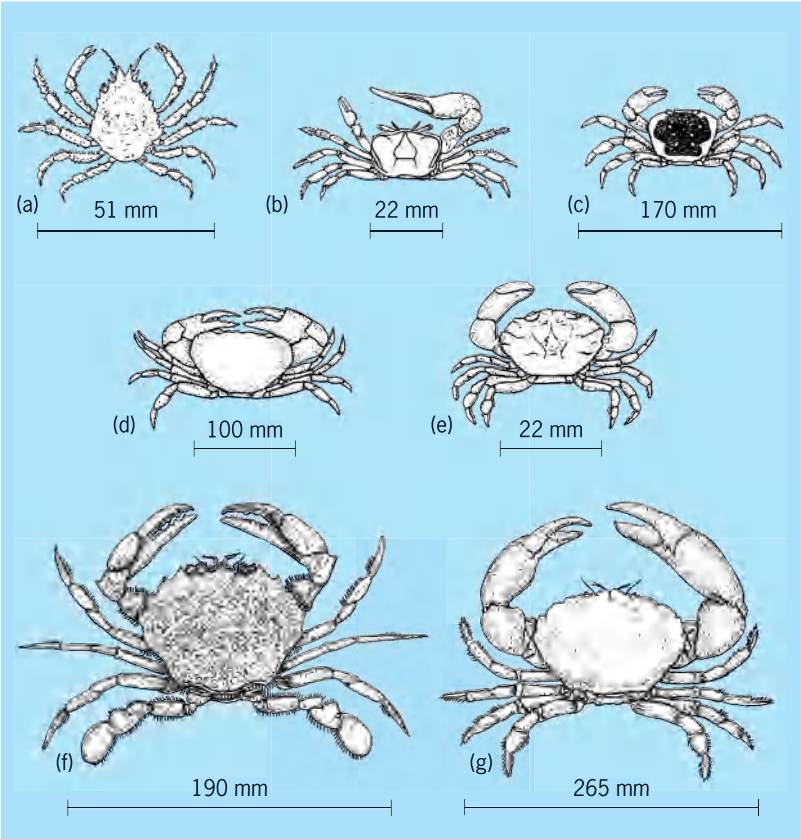

Fig. 15. Other common brachyuran crabs. (*a*) Spider crab, *Mithrax acuticornis*. (*b*) Fiddler crab, *Uca pugilator*. (c) Land crab, *Gecarcinus lateralis*. (*d*) Freshwater crab, *Epilobocera sinuatifrons*. (e) Mud crab, *Eurypanopeus abbreviatus*. (*f*) Swimming crab, *Ovalipes ocellatus*. (*g*) Stone crab, *Menippe mercenaria*.

Trichodactylidae) [Fig. 15*d*], found in tropical and some temperate regions, usually in areas not inhabited by crayfishes.

The last few families are assigned to the subsection Thoracotremata. In this group, the gonopore is located on the sternum in both sexes. The Pinnotheridae (pea crabs) are almost always small (less than 1 cm), pale parasites or commensals found on or in bivalve mollusks (especially oysters), tube-dwelling polychaetes, and other invertebrates. The square-backed crabs (Grapsidae) [Fig. 13] are characteristic of warm, marshy areas and rocky shores. They mark the trend toward the true land crabs (Gecarcinidae) [Fig. 15*c*], the depredations and intrusions of which are familiar to all who live in the tropics. Similar to gecarcinids, members of the family Ocypodidae (Fig. 15*b*) occupy burrows during the day or high tide. Ghost crabs (*Ocypode*) scuttle almost unseen over sandy beaches and dunes at night. Fiddler crabs (*Uca*) emerge at low tide, often in large numbers, to feed on mud and sand deposited in embayments; males also engage in courtship and combat displays using a greatly inflated major chela. *See* ARTHROPODA; CRUSTACEA.

Fenner A. Chace, Jr; Robert E. Knowlton

Bibliography. K. Anger, The Biology of Decapod Crustacean Larvae, in R. Vonk (ed.), *Crustacean Issues*, vol. 14, 2001; R. T. Bauer, *Remarkable Shrimps: Adaptations and Natural History of the Carideans*, 2004; D. E. Bliss, *Shrimps, Lobsters and Crabs: Their Fascinating Life Story*, 1990; F. W. Harrison and A. G. Humes (eds.), *Microscopic Anatomy of Invertebrates: Decapod Crustacea*, vol. 10, 1992; T. H. Huxley, *The Crayfish: An Introduction to the Study of Zoology*, 1977; J. W. Martin and G. E. Davis, An updated classification of the Recent Crustacea, *Nat. Hist. Mus. Los Angeles Co., Sci. Ser.*, 39:1–124, 2001; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; G. F. Warner, *A Biology of Crabs*, 1977; A. B. Williams, *Shrimps, Lobsters, and Crabs of the Atlantic Coast of the Eastern United States, Maine to Florida*, 1984.

Decca

A hyperbolic radio navigation system. The Decca Navigator hyperbolic radio-navigation system was invented in the United States in 1937 and developed in 1944 by the British Decca Company. More than 50,000 ships of all classes, including a large and growing number of recreational craft, use the system, and about 750 helicopters and fixed-wing aircraft; coverage is shown in **Fig. 1**.

The system employs radio transmitting stations grouped in chains of four (in a few cases, three) established at precisely known positions. A chain comprises a central master station with outlying slave stations, all radiating unmodulated continuous-wave signals on spot frequencies between 70 and 130 kHz. In a given chain, each of the slave transmissions is locked in a known phase relationship with the master. Assuming a constant speed of signal propagation, the phase difference between the signals received from the master and a given slave is a measure of the difference in the distances from the receiver to the two stations. The phase difference can thus locate the receiver on a line of position in the form of a hyperbola with the stations as foci. The observed phase difference between the master and the signal from a second slave similarly identifies an intersecting line of position, fixing the position.

In its simplest form, the on-board receiver displays the phase differences as numerical line-of-position readings, which the user plots on a marine chart overprinted with a lattice of correspondingly numbered hyperbolic curves. A four-station chain produces three overlapping line-of-position patterns, printed on charts in red, green, and purple. For manual plotting (**Fig. 2**), the charts normally carry the two patterns cutting at the best angle and omit the third. Computer-based Decca receivers can make use of all three patterns to form the fix. Such receivers also convert the hyperbolic fix into latitude and longitude, position with respect to a waypoint, and other desired information.

Lanes and lane identification. The line-of-position readings are expressed in units and fractions of a lane, where a lane is the space corresponding to a change of 360° in the master–slave phase difference. Typically the receiver measures the phase difference to 3.6° (0.01 lane), corresponding to a line-of-position resolution of about 16 ft (5 m) on the

master-to-slave baselines where the lane width is minimum. The measurement is, however, ambiguous, since there is no way of determining directly the number of whole cycles in the phase difference. Accordingly the receiver automatically keeps count of the successive lanes that it passes through, once the correct lane numbers at the starting position have been inserted. These numbers are supplied periodically by additional transmissions which identify, for each pattern, the correct lane within a relatively large group of lanes known as a zone. The zones are about 6 mi (10 km) wide on the baselines, so that in practice there is a spacing of 6 mi (10 km) or more between ambiguous lines of position. The lane identification values are either displayed as numerical readings or, in computer-based receivers, processed automatically.

Frequencies. Each chain is assigned a set of four frequencies (see **table**), which are multiples of a fundamental value f between 14.0 and 14.4 kHz. The harmonic relationship allows the receiver to derive, by frequency multiplication, a common frequency from the master and a given slave so that their phase can be compared. For example, master $6f \times 3$ and green slave $9f \times 2$ are each equal to $18f$, and the green line-of-position pattern is the result of phase comparison at that frequency.

For lane identification, each station transmits in turn, in a cycle repeated every 20 s, a half-second burst comprising the four phase-coherent harmonics $5f$, $6f$, $8f$, and $9f$. The other stations cease transmitting on these frequencies for the half-second. From the fourfold signals, the receiver derives, by one of several methods depending on the software, a high-integrity signal having the fundamental frequency f. Comparing master and slave phases at that frequency generates what amounts to a coarse

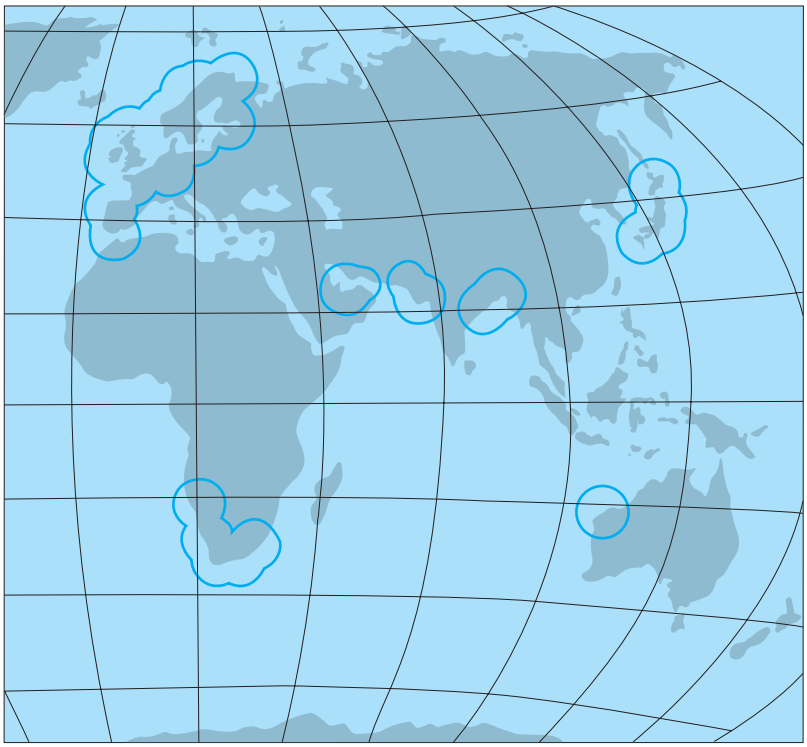

Fig. 1. Decca coverage areas. (*After Racal-Decca Marine Navigation Ltd, Decca Navigator Coverage Diagrams, 1988*)

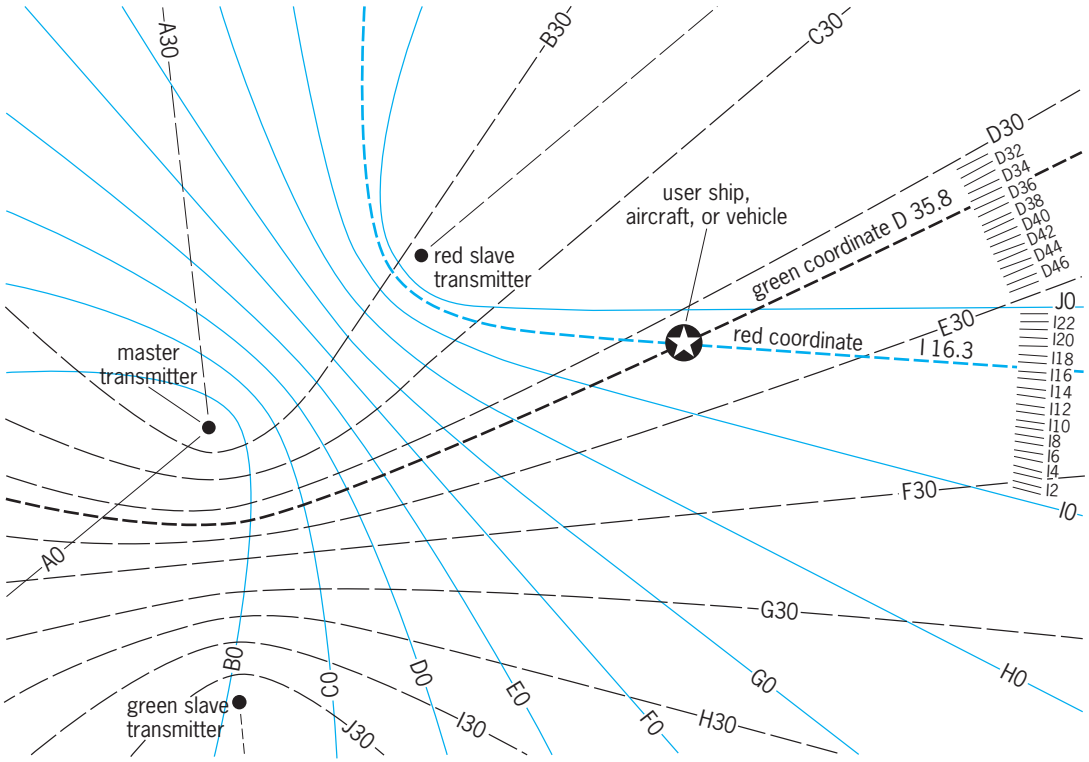

Fig. 2. Fix plotted from line-of-position readings Red I 16.30, Green D 35.80, where I and D denote the zones; lanes are shown at top right. (*After C. Powell, Radio Position Fixing for Yachtsmen, Adlard Coles Ltd, London,1986*)

Decca transmitted and comparison frequencies*

	Frequency, kHz	Harmonic, *f*	Lane or zone width on baseline, ft (m)
Transmitted frequencies			
Master station	85.0000	6	
Purple slave	70.8333	5	
Red slave	113.3333	8	
Green slave	127.5000	9	
Phase-comparison frequencies			
Purple lanes pattern	425.00	30	1,155 (352.1)
Red lanes	340.00	24	1,443 (440.1)
Green lanes	255.00	18	1,925 (586.8)
Zones (all patterns)	14.1666	1	34,652 (10,562.0)

*For Decca chain 05B.

hyperbolic pattern superimposed upon the lanes and serving to identify the correct lane for each pattern, assuming that the zone is known. Some computer-based receivers treat the continuous and lane identification signals in a relationship that differs from the foregoing basic description.

Range and accuracy. The range of a Decca chain is typically some 240 nautical miles (440 km) by night and about twice that distance by day. Repeatability of position fixing is as high as a few tens of meters in favorable coverage during summer daytime, falling to a few miles at night near the limit of range. The variation with time occurs through contamination of the wanted ground-wave component of the received signals by the sky wave reflected by the ionosphere, which is strongest at night and varies randomly in phase. Published accuracy data include predicted levels of repeatability (random position-fixing error), taking seasonal and diurnal sky-wave propagation variations into account. The composite lane identification signal is relatively insensitive to the effects of sky-wave contamination, as a result of the fourfold frequency diversity of the transmission. Data are also published on fixed (systematic) position-fixing errors observed in coastal areas. *See* ELECTRONIC NAVIGATION SYSTEMS; HYPERBOLIC NAVIGATION SYSTEM; MARINE NAVIGATION. Claud Powell

Bibliography. *Admiralty List of Radio Signals,* vol. 5, updated periodically; C. Powell, *Radio Position Fixing for Yachtsmen*, 1986; Racal-Decca Marine Navigation Ltd, *Decca Navigator Coverage Diagrams*, 1988; Racal-Decca Marine Navigation Ltd, *Marine Data Sheets*, updated periodically.

Decibel

A logarithmic unit used to express the magnitude of a change in level of power, voltage, current, or sound intensity. A decibel (dB) is $^1/_{10}$ bel.

In acoustics a step of 1 bel is too large for most uses. It is therefore the practice to express sound intensity in decibels. The level of a sound of intensity I in decibels relative to a reference intensity I_R is given by notation (1). Because sound intensity is proportional to the square of sound pressure P, the level in decibels is given by Eq. (2). The reference pressure is usually taken as 0.0002 dyne/cm^2 or 0.0002 microbar. (The pressure of the Earth's atmosphere at sea level is approximately 1 bar.) A sinusoidal alternation in pressure at a frequency of 1000 Hz is barely audible to the average person when it has a root-mean-square sound pressure of 0.0002 microbar. By this definition such a tone has a sound pressure level of 0 dB. *See* SOUND PRESSURE.

$$10 \log_{10} \frac{I}{I_R} \qquad (1)$$

$$10 \log_{10} \frac{P^2}{P_R{}^2} = 20 \log_{10} \frac{P}{P_R} \qquad (2)$$

The neper is similar to the decibel but is based upon natural (napierian) logarithms. One neper is equal to 8.686 dB. *See* NEPER; VOLUME UNIT (VU).

Karl D. Kryter

Bibliography. American National Standards Institute, *Acoustical Terminology*, ANSI S1.1-1994 (R2004), 2004.

Deciduous plants

Woody higher plants that regularly lose their leaves at the end of each growing season. Dropping of the leaves occurs at the inception of an unfavorable season characterized by either cold or drought or both. Many woody plants of temperate climates have the deciduous habit, and it may also occur in those of tropical regions having alternating wet and dry seasons. Many deciduous trees and shrubs of regions with cold winters become evergreen when grown in a warm climate. Conversely, such trees as magnolias, evergreen in warm areas, become deciduous when grown in colder climates. *See* LEAF; PLANT PHYSIOLOGY; PLANT TAXONOMY. Nelle Ammons

Decision analysis

An applied branch of decision theory. Decision analysis offers individuals and organizations a methodology for making decisions; it also offers techniques for modeling decision problems mathematically and finding optimal decisions numerically. Decision models have the capacity for accepting and

quantifying human subjective inputs: judgments of experts and preferences of decision-makers. Implementation of models can take the form of simple paper-and-pencil procedures or sophisticated computer programs known as decision aids or decision systems.

Principles and methodology. The methodology is rooted in postulates of rationality—a set of properties which preferences of rational individuals must satisfy. One such property is transitivity: if an individual prefers action a to action b and action b to action c, he or she should prefer a to c. From the rationality postulates, principles of decision-making are derived mathematically. The principles prescribe how decisions ought to be made, if one wishes to be rational. In that sense, decision analysis is normative.

The methodology is broad and must always be adapted to the problem at hand. An illustrative adaptation to a class of problems known as decision-making under uncertainty (or risk) and consists of seven steps (**Fig. 1**):

1. The problem is structured by identifying feasible actions $a_1, \ldots, a_k, \ldots, a_n$, one of which must be decided upon; possible events $\omega_1, \ldots, \omega_j, \ldots, \omega_m$, one of which occurs thereafter; and outcomes $z_{11}, \ldots, z_{kj}, \ldots, z_{nm}$, each of which results from a combination of decision and event. Problem structuring can be facilitated by displays such as decision trees, decision matrices, and influence diagrams.

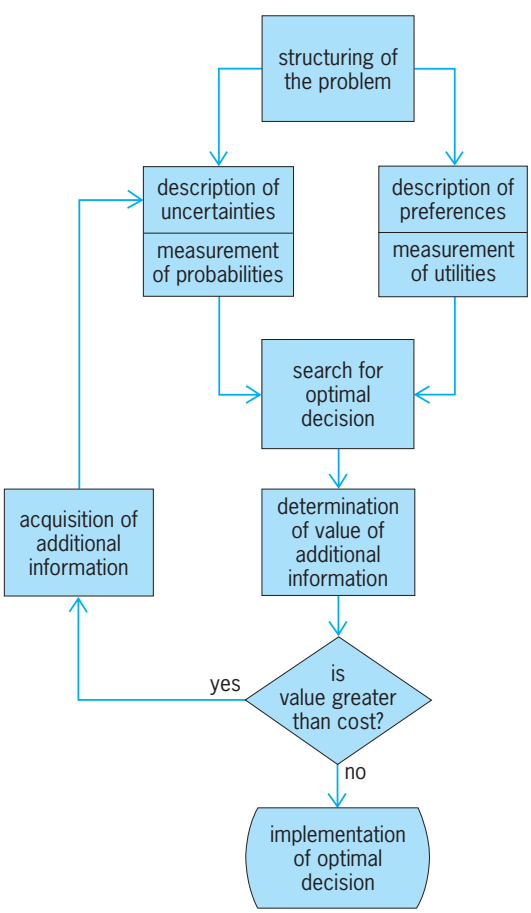

Fig. 1. A methodology of decision analysis.

2. At the time of decision-making, the event that will actually occur cannot be predicted perfectly. The degree of certainty about the occurrence of an event, ω_j, given all information at hand, is quantified in terms of the probability of the event, $P(\omega_j)$. *See* PROBABILITY.

3. Preferences are personal: the same outcome may elicit different degrees of desirability from different individuals. The subjective value that a decision-maker attaches to an outcome is quantified and termed the utility of outcome, $u(z_{kj})$.

4. The preceding steps conform to the principle of decomposition: probabilities of events and utilities of outcomes must be measured independently of one another. They are next combined in a criterion for evaluating decisions. The utility of a decision $U(a_k)$ is defined in Eq. (1) as the expected utility of

$$U(a_k) = \sum_{j=1}^{m} u(z_{kj})P(\omega_j) \tag{1}$$

the outcome. The optimal, or the most preferred, decision a^* is one with the maximum utility, as given in Eq. (2), where the symbol max indicates the oper-

$$U(a^*) = \max_{k} \; U(a_k) \tag{2}$$

ation of finding the maximum by varying the index k from 1 to n. Relations (1) and (2) constitute the principle of utility maximization.

5. The probability $P(\omega_j)$ encodes the current state of information about a possible event. Often, additional information can be acquired in the hope of reducing the uncertainty. The monetary value of such information is computed before purchase and compared with the cost of information. Thus, one can determine whether or not acquiring information is economically rational.

6. The source of information may be a real-world experiment, a laboratory test, a mathematical model, or the knowledge of an expert. The informativeness of the source is described in terms of a probabilistic relation $P(I|\omega_j)$ between information I and event ω_j. This relation, known as the likelihood function, makes it possible to revise the prior probability $P(\omega_j)$ of the event and to obtain a posterior probability $P(\omega_j|I)$ of the event, conditional on additional information. The revision is carried out via Bayes' rule, as shown in Eq. (3).This rule (**Fig. 2**) is also known

$$P\left(\omega_j \,|I\right) = \frac{P\left(I\,|\omega_j\right) P\left(\omega_j\right)}{\sum_{j=1}^{m} P\left(I\,|\omega_j\right) P\left(\omega_j\right)} \tag{3}$$

as the Bayesian learning model. *See* BAYESIAN STATISTICS; ESTIMATION THEORY.

7. Given the additional information, prior probabilities can be replaced by posterior probabilities, and the analysis can be repeated from step 4 onward. Steps 4–6 may be cycled many times, until the cost of additional information exceeds its value, at which moment the optimal decision is implemented.

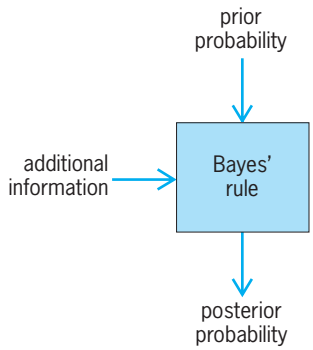

Fig. 2. Model for revising the state of information or knowledge.

Situations and techniques. A spectrum of technique exists for modeling and solving various tasks within a decision analysis; these tasks vary with situations. Six elements of decision situations are highlighted below along with the associated modeling tasks.

1. The decision-maker may be either an individual or a group, in which case utility functions of group members must be aggregated.

2. The stakeholder (one who is directly affected by the outcome of a decision) may be the decision-maker, another individual, or a group; a decision-maker may wish to include in the analysis a utility function of the stakeholder in addition to his or her own.

3. Events leading to outcomes either may be known with certainty or be uncertain at the time of decision-making. Accordingly, decision-making under conditions of certainty and uncertainty (or risk) requires different techniques.

4. Information about possible events may be acquired from single or multiple sources; in the latter case, information must be aggregated.

5. Outcomes may be one-dimensional (for example, monetary damage caused by a flood) or multidimensional (for example monetary damage and lives lost); they may be quantitative or qualitative (for example, verbal descriptors of product quality or the state of one's health). Outcome dimensions are often called attributes, and a decision-maker's preferences are described in terms of a single-attribute or multiattribute utility function. The latter is usually a composition of single-attribute utility functions and parameters that encode preferred trade-offs between outcome dimensions.

6. Time scale: decision problems are static when one decision must be made at a specified time; they are dynamic when the optimal timing of action must be decided (a stopping problem) or when many decisions must be made at predetermined times (a sequential decision problem). Dynamic problems may have a finite or infinite horizon, and they require two additional models. First, it is necessary to describe a sequence of events as a stochastic process. Second, it is necessary to describe preferences for sequences of outcomes; this may involve discounting the value of future outcomes relative to the present ones. *See* STOCHASTIC PROCESS.

Multiobjective decision analysis. A static decision problem under conditions of certainty and with multiple objectives is structured as follows. Objectives of import to a decision-maker are identified (for example, maximize economic efficiency, maximize customer satisfaction, minimize environmental impact). To measure the attainment of the objectives, attributes $z_1, \ldots, z_i, \ldots, z_r$ are defined, one or more for each objective. For each feasible action a_k, outcomes $z_{k1}, \ldots, z_{ki}, \ldots, z_{kr}$, which represent levels of the attributes, are determined.

Suppose that for each attribute a higher outcome is preferred to a lower outcome. In a comparison of two actions, action a_k is said to dominate action a_l if $z_{ki} \geq z_{li}$ for $i = 1, \ldots r$, with at least one strict inequality. Any action dominated by some other action is eliminated from further analysis. To rank the remaining, nondominated actions, preferences of the decision maker must be modeled.

The subjective values that the decision maker attaches to multidimensional outcomes are quantified in terms of a multiattribute utility function. The simplest form is additive, as shown in Eq. (4). Here,

$$U(a_k) = \sum_{i=1}^{r} \lambda_i u_i(z_{ki}) \tag{4}$$

the utility of an action $U(a_k)$ equals the sum of scaled single-attribute utilities $u_i(z_{ki})$, where the scaling constants λ_i (also termed the trade-off coefficients) quantify the trade-offs preferred by the decision maker in the following sense: λ_i is the utility of an action that yields the most desirable outcome in attribute z_i and the least desirable outcomes in the remaining attributes. By ordering the utilities of actions $U(a_k)$, $k = 1, \ldots, n$, from the highest to the lowest, the most preferred ranking of actions is decided.

Measurement of probability and utility. Measurement of probability and utility functions is guided by principles of decision theory, statistical estimation procedures, and empirical laws provided by behavioral decision theory—a branch of cognitive psychology.

Statistical techniques for estimating probabilities are applicable when the available information consists of past observations of events, outputs from models simulating behavior of a system that generates the events, or other numerical data. When information is in the form of qualitative statements and images or general knowledge of experts, probabilities are assessed via human judgment. For instance, meteorologists assess probabilities of precipitation occurrence, and physicians assess probabilities of correctness of their diagnoses. Several questions arise: How do people perform such judgmental tasks? How can they be aided in these tasks? How can their performance be evaluated? How can they be trained to become good probability assessors?

These questions are studied within behavioral decision theory. Although far from complete, the

theory has gathered vast empirical knowledge about human judgmental capabilities. When processing information intuitively, people rely on heuristics—simple subjective rules for inference. Heuristics are at best approximations to normative inference rules and frequently lead to predictable biases. For instance, a heuristic known as anchoring and adjustment operates as follows: Information first received is used to form a prediction, which is next adjusted in light of new information. Because adjustments are usually insufficient, the performance of an expert does not increase commensurately with information. Yet, the expert's confidence in his prediction increases steadily. At the end of information processing, most experts are overconfident.

Understanding heuristics and biases is important for designing probability measurement protocols that lessen erroneous predictions. Methods exist for evaluating the performance of probabilistic predictions and providing experts with feedback. Empirical evidence suggests that possessing professional expertise does not assure good probability assessment. Training in probabilistic thinking usually improves an expert's performance.

Techniques for measuring utilities are numerous. Most often, a decision-maker is presented with two simple alternatives, and his task is to choose. For example, the choice may be between an alternative that yields outcome y with certainty and an alternative that yields outcome x with probability p or outcome z with probability $1 - p$, a gamble. Preferences expressed in such simple situations supply information for constructing a utility function. A decision model (1)–(2) with this utility function applied to simple situations should replicate the decision-maker's choices. In that sense, the decision model should be descriptive.

The main premise of the analysis, however, is that the model will be applied to complex situations, and will prescribe decisions that are optimal according to the preferences encoded in the utility function. Thus, the model guarantees that all decisions, in simple and complex situations, satisfy the postulates of rationality, that is, all decisions are coherent.

Experimental evidence points to the limited human capacities for making coherent decisions intuitively, particularly in situations with uncertainties and multidimensional outcomes. Human preferences can easily be affected by the format of information display, the ways questions are framed and sequenced, and other factors which, from the normative point of view, should not matter. Research is in progress toward a better understanding of cognitive processes underlying preference formation and intuitive decisions. The ultimate objective is to improve techniques for interfacing a decision analysis with a human decision-maker whom the analysis seeks to serve. *See* COGNITION; DECISION THEORY; OPERATIONS RESEARCH; OPTIMIZATION; SYSTEMS ANALYSIS; SYSTEMS ENGINEERING. Roman Krzysztofowicz

Bibliography. R. T. Clemen, *Making Hard Decisions: An Introduction to Decision Analysis*, 2d ed., 1996; S. French, *Decision Theory: An Introduction to the Mathematics of Rationality*, 1988; R. M. Hogarth, *Judgement and Choice*, 2d ed., 1987; R. L. Keeney and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, 1976, reprint 1993; D. V. Lindley, *Making Decisions*, 2d ed., 1991; D. von Winterfeldt and W. Edwards, *Decision Analysis and Behavioral Research*, 1986.

Decision support system

A system that supports technological and managerial decision making by assisting in the organization of knowledge about ill-structured, semistructured, or unstructured issues. A structured issue has a framework comprising elements and relations between them that are known and understood. Structured issues are generally ones about which an individual has considerable experiential familiarity. A decision support system (DSS) is not intended to provide support to humans about structured issues since little cognitively based decision support is generally needed for these experientially familiar situations. Artificial-intelligence based expert systems are very appropriate for these tasks.

Scope of application. The primary components of a decision support system are a database management system (DBMS), a model-base management system (MBMS), and a dialog generation and management system (DGMS) [see **illus.**]. Emphasis in the use of a decision support system is upon provision of support to decision makers in terms of increasing the effectiveness of the decision-making effort. This support involves the systems engineering steps of formulation of alternatives, the analysis of their impacts, and interpretation and selection of appropriate options for implementations. Efficiency in terms of time required to evolve the decision, while important, is usually secondary to decision effectiveness.

Types of decisions and systems. Decisions may be described as structured or unstructured, depending upon whether or not the decision-making process can be explicitly described prior to its execution. Generally, operational performance decisions are more likely than strategic planning decisions to be prestructured. Thus, expert systems are usually more appropriate for operational performance and operational control decisions, while decision support systems are more appropriate for strategic planning and management control. Most expert systems are based upon extensive use of what are generally called production rules. This rule-based reasoning may not cope well with situations that require either formal knowledge-based reasoning or skill-based expert reasoning. *See* EXPERT SYSTEMS.

Systems engineering steps. Fundamental to the notion of a decision support system is the formal assistance provided to humans in assessing the situation of potential interest, identifying alternative courses of action, formulating the decision situation, structuring and analyzing the decision situation in terms of obtaining the impacts of alternative potential

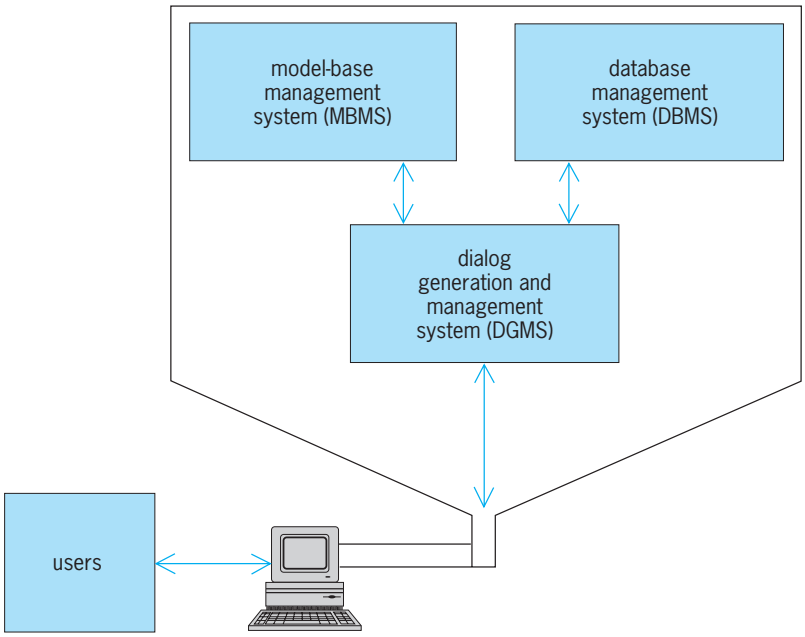

Components of a decision support system (DSS).

courses of action, and interpreting the results of analysis of the alternatives in terms of the value system of the decision maker. These are the basic formulation, analysis, and interpretation steps of systems engineering in expanded form. *See* SYSTEMS ENGINEERING.

Database management system. An appropriate database management system must be able to work with both data that are internal to the organization and data that are external to it. If there are multiple decision makers, then personal databases, local databases, and systemwide databases are generally needed. A database management system should be able to cope with a variety of data structures that allow for probabilistic, incomplete, and imprecise data, and data that are unofficial and personal as contrasted with official and organizational. It should also be capable of informing the support-system user of the types of data that are available and how to gain access to them. *See* DATABASE MANAGEMENT SYSTEM.

Model-base management system. The need to provide recommendation capability in a decision support system leads to the construction of model-base management systems. Such systems provide sophisticated analysis and interpretation capability in a decision support system. A model-base management system enables the decision maker to explore the decision situation through use of the database by a model base of algorithmic procedures and associated model management protocols. This process may involve use of modeling statements in some procedural or nonprocedural language; model subroutines, such as mathematical programming packages, that are called by a management function; and data abstraction models. This last approach is close to the expert-system approach in that there exist element, equation, and solution procedures that together make up an inference engine. Advantages include ease of updating and use of the model for explanatory and explication purposes.

The model-base management system should provide flexibility upon system-user request through a variety of prewritten models, such as linear programming and multiattribute decision analysis models, and procedures to use these models. The system should also allow for the development of user-built models and heuristics that are developed from established models. It should also be possible to perform sensitivity tests of model outputs, and to run models with a range of data in order to obtain the response to a variety of what-if questions. *See* DECISION ANALYSIS; LINEAR PROGRAMMING; MODEL THEORY.

Dialog generation and management. The dialog generation and management system is designed to satisfy knowledge representation, and control and interface requirements of the decision support system. It is responsible for presentation of the information outputs of the database management system and model-base management system to the decision makers, and for acquiring and transmitting their inputs to these two systems. Specifically, it is responsible for producing the output representations of the decision support system, for obtaining the decision-maker inputs that result in the operations on the representations, for interfacing to the memory aids, and for explicit provision of the control mechanisms that enable the dialog between user input and output and the database management system and model-base management system. *See* DECISION THEORY. Andrew P. Sage

Bibliography. J. A. Lawrence, Jr., and B. A. Pasternack, *Applied Management Science: A Computer-Integrated Approach for Decision Making*, 1998; H. R. Parsaei, S. Kolli, and T. R. Hanley (eds.), *Manufacturing Decision Support Systems*, 1998; A. P. Sage, *Decision Support Systems Engineering*, 1991; R. H. Sprague, Jr., and H. J. Watson, *Decision Support for Management*, 1996; E. Turban and J. E. Aronson, *Decision Support Systems and Intelligent Systems*, 1998.

Decision theory

A broad range of concepts which have been developed to both describe and prescribe the process of decision making, where a choice is made from a finite set of possible alternatives. Normative decision theory describes how decisions should be made in order to accommodate a set of axioms believed to be desirable; descriptive decision theory deals with how people actually make decisions; and prescriptive decision theory formulates how decisions should be made in realistic settings. Thus, this field of study involves people from various disciplines: behavioral and social scientists and psychologists who generally attempt to discover elaborate descriptive models of the decision process of real humans in real settings; mathematicians and economists who are concerned with the axiomatic or normative theory of decisions; and engineers and managers who may be concerned with sophisticated prescriptive decision-making procedures.

Concepts and terminology. Decisions have various degrees of importance in the life of an individual, a group, or an organization. The importance of a decision depends upon its consequences or outcomes as determined by: (1) controlled factors, represented by the alternative courses of action or alternate strategies available to the decision maker; and (2) uncontrolled factors, or states of nature, which interact with the strategies.

A decision is an allocation of resources. It is generally irrevocable, although a later decision may sometimes reverse the effects of an initial decision. The decision maker is the person who has authority over the resources that are being allocated. This person makes a decision in order to further the achievement of some objective which is felt to be desirable.

A decision maker should make decisions that are consistent with values, or those things that are important to the decision maker, or those parties represented by the decision maker. The most common value is economic, the effort to increase wealth; but such personal or social values as happiness, security, and equity are also important. Sometimes a distinction is made between a goal and an objective. A goal refers to a specific degree of satisfaction of a given objective. For example, the objective of a decision to exercise or diet might be to lose weight, and the goal might be to weigh 10 lb (4.5 kg) less in a month.

The important features of a decision situation are (1) a decision to be made and course of action to be taken, (2) the unknown events and outcomes which can affect the result, and (3) the obtained result. Decision theory is based on construction of models which represent logical mathematical representations of the relationships among these three features. Use of these models leads to estimation of the possible implications of each alternative course of action that might be taken, and thereby to understanding of the relationships between actions and objectives, and to selection of the most appropriate decision.

At the time a decision is made or taken, the decision maker has at least two alternatives or courses of action (otherwise, no decision is possible). When an alternative is chosen and committed to, the decision has been made or taken, and uncertainties (uncontrollable chance elements) generally arise. Different alternatives may be associated with different uncertainties. It is the combination of alternatives with uncertainties that ultimately leads to an outcome. Thus, an outcome is a generally uncertain result that follows from the decision situation, whose impacts should be measured on the decision maker's value scale.

A simple decision is a situation in which there is only one decision to be made, even though there might be many alternatives. If there are two or more decisions to be made, a problem of strategy arises. Sometimes, several decisions need to be made at the same time, and this special case is called the portfolio problem. Often, the various decisions associated with the strategy are of a similar nature, and there are not sufficient resources to enable selection of all alternatives. The prototypical example is an investment portfolio. Here, the decision maker is aware of a number of investments that might potentially be made, but there is a finite budget. In such a situation, the decision maker must accomplish a prioritization of the various opportunities. If one opportunity has a higher priority than another, the decision maker would prefer to invest in the alternative that leads to this opportunity. In other cases, sequential decisions are involved. Here, the decision maker is able to observe the outcome that results from one decision before having to make a subsequent decision.

The consequences of the chosen strategies are dependent quantities commonly termed outcomes. These are often measured by a variety of quantitative or qualitative parameters, such as the number of dollars returned from an investment, or the quality of life that results from a successful medical procedure, and by the probability of the various possible consequences following a selected course of action. The aggregate outcome that follows from a given decision is known as the expected relative value or expected utility.

Categories of problems. Efforts in decision theory may be divided into five categories:

1. Decision under certainty issues are those in which each alternative action results in one and only one outcome and where that outcome is sure to occur. The decision situation structural model is established and is correct. There are no parametric uncertainties.

2. Decision under probabilistic uncertainty issues are those in which one of several outcomes can result from a given action depending on the state of nature, and these states occur with known probabilities. The decision situation structural model is established and is correct. There are outcome uncertainties, and the probabilities associated with these are known precisely. The utility of the various event outcomes can be precisely established. *See* PROBABILITY.

3. Decision under probabilistic imprecision issues are those in which one of several outcomes can result from a given action depending on the state of nature, and these states occur with unknown or imprecisely specified probabilities. The decision situation structural model is established and is correct. There are outcome uncertainties, and the probabilities associated with the uncertainty parameters are not all known precisely. The utility of the event outcomes can be precisely established.

4. Decision under information imperfection issues are those in which one of several outcomes can result from a given action depending on the state of nature, and these states occur with imperfectly specified probabilities. The decision situation model is established but may not be fully specified. There are outcome uncertainties, and the probabilities associated with these are not all known precisely. Imperfections in knowledge of the utility of the various event outcomes may exist as well.

5. Decision under conflict and cooperation issues are those in which there is more than a single decision maker, and where the objectives and activities of

one decision maker are not necessarily known to all decision makers. Also, the objectives of the decision makers may differ.

Problems in any of these groupings may be approached from a normative, descriptive, or prescriptive perspective. Problems in category 1 are those for which deterministic principles may be applied. This condition of known states of nature ignores the overwhelming majority of issues faced in typical realistic decision issues.

Problems in category 5 are game theory–based problems, and will not be considered here. Most decision theory developments are concerned with issues in category 2. *See* GAME THEORY.

The basic normative decision problem in category 2 is conceptually straightforward: From a set of alternative strategies (S_n), choose the one which when combined with the relevant states of nature (N_j) will produce the outcomes (O_{nm}) which best satisfy the goals or objectives of the decision maker. Some type of penalty or benefit criterion, or multiple criteria, are used to evaluate the utility of the outcome O_{nm}. Normative decision theory implies that all four factors should be known to make a completely rational and meaningful decision. Behavioral or descriptive decision theory tries to explain why real people often makes seemingly irrational decisions.

Sample problem. By considering the relatively simple decisions which typically have to be made before getting to work, the possible difficulties can be demonstrated. In **Fig. 1**, it is assumed that the decision maker has three basic decisions to make on the way to work: (1) to get up at 7:00 A.M. (choice 1) or 7:15 A.M. (choice 2), with the associated consequence that after dressing he or she would depart at 7:15 or 7:30 A.M., respectively; (2) to choose routes A or B (choices 3 and 4) with the expected outcomes that route A always takes 30 min while route B, resulting from choice 4, can take 15, 30, or 45 min (*a*, *b*, or *c*), depending on traffic conditions; and (3) to stop for breakfast, which takes 15 min, or skip breakfast (choices 5 and 6).

The decision sequence is illustrated (Fig. 1) in the form of a decision tree. It shows all the possible intermediate and final outcomes which would result from all possible choices of alternatives posed by this example. The decision nodes at which decisions have to be made are marked *D*.

Several interesting observations can be made about the decision tree. (1) Although there are 16 possible outcomes, there are only 8 possible initial strategies since the decision outcomes *a*, *b*, and *c* are mutually exclusive and determined by the state of nature, traffic congestion, at the time. (2) The decision problem has been bounded by enumerating all admissible strategies, states of nature, and outcomes. This has already involved some judgment in excluding other strategies, states of nature, and outcomes on the basis that they are unlikely to occur, or because the decision maker is specifically interested only in the alternatives presented. (3) Although the alternative strategies and outcomes are readily apparent from this particular decision tree, the fourth ingredient of a decision problem, the objectives of the decision maker, have not yet been defined. The objectives which generate the criteria for evaluating the decision problem represent the critical link in reaching the best decision for that particular decision maker at this particular time. In Fig. 1 it is clear that the decision relative to choice of best strategy will depend on the decision maker's objectives. These objectives are translated into values, to a related concept called utility, and to the expected value of the utilities in most decision theory–based approaches.

If the objective is to definitely arrive at work at 7:45 A.M. sharp, it is possible to immediately eliminate all strategies which do not have 7:45 A.M. as their final outcome as inadmissible strategies. In this case the only admissible strategy would be (1-3-6) since any other strategy often would result in coming to work late.

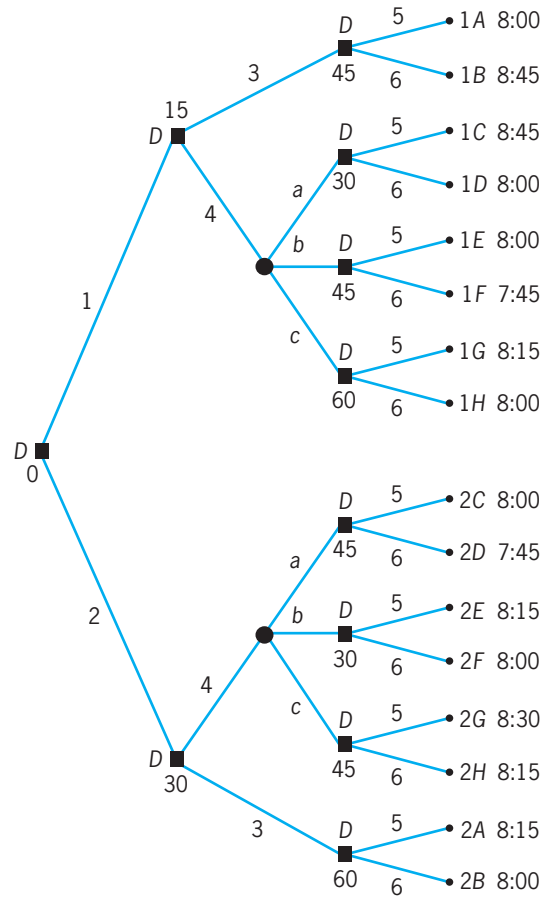

Key to decisions:

1 = Get up and dress when alarm rings
2 = Sleep 15 min more, then get up and dress
3 = Choose route A to work
4 = Choose route B to work
5 = Eat breakfast
6 = Skip breakfast

Fig. 1. Decision tree for the problem given in the text. Numbers at nodes represent time past alarm (7:00 A.M.), and numbers at branches refer to the decision mode or state of nature.

If, however, the decision maker decides that breakfast is a vital function and will always insist on breakfast, all strategies which have a branch 6 in the decision tree will be eliminated. If the objective is still to get to work as early as possible, the decision must now be made between strategy (1-3-5) and strategy (1-4-5), which may under certain circumstances result in arrival at work earlier than (1-3-5), that is, outcome 1*C*. These two strategies dominate all other admissible strategies because under all circumstances they will result in arrival at work earlier than any of the other strategies. Thus, when one strategy shows dominance over some other strategies, all but the dominant strategies can be eliminated from consideration. Strategies (1-3-5) and (1-4-5) do not dominate each other since under certain circumstances one or the other will have the more desirable outcome.

The choice of strategy (1-3-5) leads to a decision-under-certainty outcome since the outcome for this strategy is uniquely predictable and all states of nature are known. The outcome of strategy (1-4-5) depends on the traffic conditions (*a*, *b*, or *c*) or states of nature which are not under the decision maker's control. If no information is available about the likelihood or probability of occurrence of states *a*, *b*, or *c*, the decision maker will be faced with decision making under probabilistic imprecision. If the probability of the states of nature *a*, *b*, or *c* can be predicted from past experience or from observations, a probability can be associated with each of the possible outcomes. This represents decision making under probabilistic uncertainty if it is assumed that the values of the various outcomes are also known.

Value or utility theory. Except in very well defined decision problems in which the outcomes are clearly represented by such an objective parameter as dollars or time, there are conflicting and incommensurable objectives which require subjective value judgments in order to determine the value or utility of each objective. This inescapable aspect of decision theory has led to the development of value or utility theory. Utility theory strives to evaluate the subjective relative value of outcomes to individuals and groups. For example, the utility or value of $100,000 to a multimillionaire in a gambling decision is obviously quite different from its value to a person whose total possessions amount to $100,000. Similarly, the question of the relative values of 15 min extra sleep versus no breakfast versus coming to work late requires some resolution before the objectives and evaluation criteria for the prior decision problem can be established.

Bases of normative decision theory. The general concepts of axiomatic or normative decision theory formalize and rationalize the decision-making process. The major contributions of decision theory are to structure the decision problem and to define the scope and importance of components so that each part of the decision can be evaluated before synthesizing the available information into a rational decision.

Normative decision theory depends on the following assumptions:

1. Past preferences are valid indicators of present and future preferences.
2. People correctly perceive the values of the uncertainties that are associated with the outcomes of decision alternatives.
3. People are able to assess decision situations correctly, and the resulting decision situation structural model is well formed and complete.
4. People make decisions that accurately reflect their true preferences over the alternative courses of action, each of which may have uncertain outcomes.
5. People are able to process decision information correctly.
6. Real decision situations provide people with decision alternatives that allow them to express their true preferences.
7. People accept the axioms that are assumed to develop the various normative theories.
8. People make decisions without being so overwhelmed by the complexity of actual decision situations that they would necessarily use suboptimal decision strategies.

Given these necessary assumptions, there will exist departures between normative and descriptive decision theories. A principal task of those aiding others in decision making is to retain those features from the descriptive approach which enable an acceptable transition from normative approaches to prescriptive approaches. The prescriptive features should eliminate potentially undesirable features of descriptive approaches, such as flawed judgment heuristics and information processing biases, while retaining acceptable features of the normative approaches. *See* DECISION ANALYSIS; DECISION SUPPORT SYSTEM.

Decision matrices. The information necessary for a normative decision can be summarized in a decision matrix (**Fig. 2**). The alternative strategies, pure (independent from each other) or mixed (combinations of pure strategies), are listed in the first column. These strategies (S_1 to S_n) can be generated by the decision maker, or by an advisor to the decision maker often called a decision analyst, and should contain all possible courses of action which the decision maker can identify as feasible.

Across the top row, all possible states of nature (N_1 to N_m) which might affect the outcome of the various strategies are enumerated. The decision matrix can become quite large if, for example, it is to represent a complete decision tree, such as the tree of Fig. 1. In that case, the trade-off between time and money expended on a decision and the value of a more or less reliable decision can become a decision problem in itself. There are computational approaches for resolution of decision trees that involve procedures from dynamic programming and optimization theory, and often involve such computational strategies as "average out and fold back." *See* OPTIMIZATION.

The probabilities of occurrence of various states of nature (p_1 to p_m) must be estimated and listed with the states of nature. These critical values should be as reliable as possible. In value theory a distinction is made between the subjective probability of

States of nature		N_1	N_2	— —	N_j	— — — — — —	N_m	Expected value
Probabilities of occurrence		p_1	p_2	— —	p_j	— — — — — —	p_m	
Alternate strategies	S_1	O_{11}	O_{12}	— —	O_{1j}	— — — — — —	O_{1m}	EV_1
	S_2	O_{21}	O_{22}	— —	O_{2j}	— — — — — —	O_{2m}	EV_2
	|	|	|		|		|	|
	|	|	|		|		|	|
	S_i	O_{i1}	O_{i2}	— —	O_{ij}	— — — — — —	O_{im}	EV_i
	|	|	|		|		|	|
	|	|	|		|		|	|
	|	|	|		|		|	|
	S_n	O_{n1}	O_{n2}	— —	O_{nj}	— — — — — —	O_{nm}	EV_n

Fig. 2. Form of a decision matrix. O_{ij} = outcome corresponding to the interaction of a strategy (S_i) and a state of nature (N_j).

occurrence which reflects the decision maker's opinion on what the probability is, and the objective probability which is derivable mathematically or as a result of observations. Thus, an optimist may consider the probability of being delayed in traffic very small and make a decision based on this assumption, while observation of traffic conditions over a period of time may indicate that such delays will occur 90% of the time. Probability and statistics play important roles in this effort. *See* STATISTICS.

Corresponding to each interaction of a strategy (S_i) and a state of nature (N_j), a consequence or outcome (O_{ij}) is determined in terms of some value or utility parameters. The outcome may sometimes be easily computed, for example, the interest cost on a business loan, but quite often must be based on insufficient data, past experience, or questionable assumptions, such as would be the case in trying to assess the outcome of a limited war. The generation of reliable outcome data can be expensive and time-consuming, and a probability of accuracy is often associated with each outcome. How to assess the relative value or utility of mixed parameters to present a single parameter for the outcome is still a paramount question in decision and utility or value theory, that is, establishing the utility of two apples versus one apple and one orange versus two oranges and so forth.

The information in the decision matrix is used to generate the last column which describes the overall expected outcome, expected value (EV), expected utility (EU), or expected profit for each strategy.

The decision tree of Fig. 1 and the decision matrix of Fig. 2 have been discussed assuming that the probabilities are known, such that these appear to represent the case 2, decisions under probabilistic uncertainty. Alternatively, if various probabilities can be assumed to be 0 or 1, a decision-under-certainty situation is obtained. If the probabilities are inexactly specified, then case 3, probabilistic imprecision, results. Finally, if the outcome values or utilities are specified inexactly, case 4 results.

Determination of utility. When choosing among alternatives, the decision maker must be able to indicate preferences among decisions that may result in diverse outcomes. In simple situations when only money is involved, an expected-value approach might be suggested, in which a larger expected amount of money is preferred to a smaller amount. However, in many situations the utility associated with money is not a linear function of the amount of money involved.

Risk aversion issues also need to be considered. One decision maker may have an aversion toward risk, while another may be risk-prone and make decisions that are associated with high-impact unfavorable consequences, perhaps because of very large rewards if the chance outcome does turn out to be favorable. Not all consequences are easily quantified in terms of monetary value, especially when there are multiple and potentially incommensurate objectives. A scalar utility function makes it possible, in principle, to attempt to resolve questions which involve uncertain outcomes with multiple attributes. The notion of a lottery provides a basis for the utility concept.

A utility function is a mathematical transformation which maps the set of outcomes of a decision problem into some interval of the real line, thereby assigning a numerical value to each outcome of a decision problem. It conforms with the basic set of four assumptions, conditions, or axioms needed to establish a utility theory. Axiom 1 is often called the orderability condition. It ensures that the preferences of the decision maker impose a complete ordering and a transitive preference ordering across outcomes. Axiom 2 is a monotonicity condition. It ensures that the preference function is smooth and that increasing preferences are associated with increasing rewards.

Axiom 3 is a continuity condition that ensures comparability across outcomes with equivalent preferences. Axiom 4 is a decomposition condition that ensures transformation of lotteries and substitutability of equivalent lotteries.

While utility theory is not complicated, successfully practicing the theory is very complicated. A number of important theoretical issues can easily be overlooked.

According to expected utility theory, the decision maker should seek to choose the alternative a_i which makes the resulting expected utility the largest possible. The utility u_{ij}, of choosing decision a_i and obtaining outcome event e_j, will also depend upon the particular value of the probabilistically uncertain random variable e_j as conditioned on the decision path that is selected. So, the best that the decision maker can do here is to maximize some function, such as the expected value of utility, as in Eq. (1), where

$$\max_i \mathrm{EU}\{a_i\} = \max_i \sum_{j=1}^{n} u_{ij} P(e_j \mid a_i) \qquad (1)$$

the maximization is carried out over all alternative decisions, and $P(e_j \mid a_i)$ is the probability that the state of nature is e_j given that alternative a_i is implemented. The abbreviated notation $\mathrm{EU}\{a_i\}$ is often used to mean the expected utility of taking action a_i. Generally, this is also called the subjective expected utility (SEU). "Subjective" denotes the fact that the probabilities may be based on subjective beliefs and the utilities may reflect personal consequences.

Generation of a utility curve. A lottery is a chance mechanism which results in an outcome with a prize or consequence $e(1)$ with probability $P(1)$, an outcome with prize or consequence $e(2)$ with probability $P(2), \ldots,$ and an outcome with prize or consequence $e(r)$ with probability $P(r)$. The probabilities $P(i)$ must be nonnegative and such that they sum to one, so that Eq. (2) holds. The lottery is denoted as $L = [e_1, P_1; e_2, P_2; \ldots; e_N, P_N]$.

$$\sum_{j=1}^{M} P(j) = P(1) + P(2) + \cdots + P(M) = 1 \qquad (2)$$

A utility curve for money can be generated in the following manner. All the outcomes of a decision are arranged in order of preference. The most preferred outcome is denoted W and the least preferred outcome L. The utilities $U(W) = 1$ and $U(L) = 0$ are arbitrarily assigned. Next, the decision maker is encouraged to answer the question: "Suppose I owned the rights to a lottery which pays W with probability 1/2 and L with probability 1/2. For what amount, say $X_{0.5}$, would I be willing to sell the rights to this lottery?" Since the decision maker is indifferent between definitely receiving $X_{0.5}$ and participating in the lottery $[W, 0.5; L, 0.5]$, Eq. (3) must be satisfied (**Fig. 3**).

$$U[X_{0.5}] = 0.5U(W) + 0.5U(L) = 0.5 \qquad (3)$$

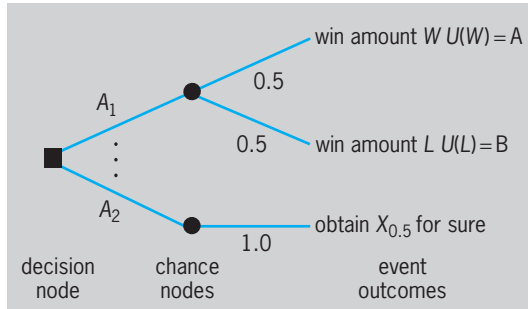

Fig. 3. Decision tree structure to determine the subjective utility of lottery outcomes.

Any other lottery with outcomes W and L could have been picked, such as $[W, 0.01; L, 0.99]$. There are several analytical reasons, however, for choosing a lottery which gives W or L with equal probability, or $[W, 0.5; L, 0.5]$. It may be conceptually quite simple for the decision maker to imagine a lottery which gives W or L with equal probability. In this case, the decision maker need only be concerned with the relative preferences of W and L. Also, the value of the lottery $[W, 0.5; L, 0.5]$ occurs midway on the U axis between W and L and therefore facilitates plotting the utility curve. This choice of 0.5 probability for the assessment may, however, be quite unrealistic. For example, W might mean some enormous increase in profit to the firm and L might mean going bankrupt. A business person might so abhor a situation in which there is a 50% chance of going bankrupt that it would not be realistically possible to think in terms of such extreme utilities.

Next, the decision maker is encouraged to answer the question: "Suppose I owned the rights to a lottery which paid W with probability 0.5 and $X_{0.5}$, which has a utility of 0.5, with probability 0.5. For what amount, call it $X_{0.75}$, would I just be willing to sell the rights to this lottery?" Here, the decision maker is indifferent between receiving $X_{0.75}$ for certain and participating in the lottery $[W, 0.5; X_{0.5}, 0.5]$, so Eq. (4) holds. Next, the lottery $[X_{0.5}, 0.5; L, 0.5]$

$$U(X_{0.75}) = 0.5U(X_{0.5}) + 0.5U(L) = 0.75 \qquad (4)$$

is considered, and the amount to be determined is $X_{0.25}$. In this case, Eq. (5) is obtained. This process

$$U(X_{0.25}) = 0.5U(X_{0.5}) + 0.5U(L) = 0.25 \qquad (5)$$

could be continued indefinitely. However, the three points generated here and the two end points are often sufficient to give an idea of the shape of the utility curve, which is usually concave downward, concave upword, or flat (**Fig. 4**). The utilities that have been obtained are indicators of preference, and are not absolute measurements. As a result, they are unique only up to a linear transformation of the form $U'(X) = aU(X) + b$, where $a > 0$.

A decision maker is rarely risk-prone, and most people are risk-averse in almost all situations. One of the few exceptions to this behavior is situations where there is a very small probability of receiving a very large reward, such as a probability of 10^{-6} of

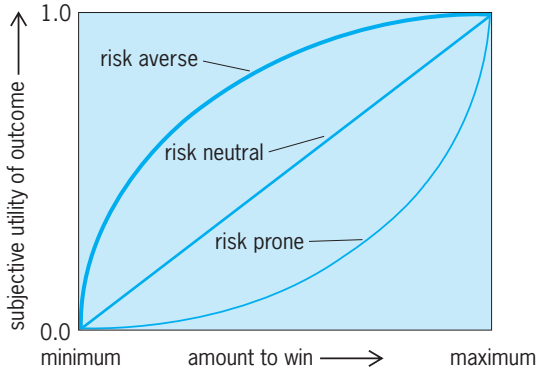

Fig. 4. Hypothetical subjective utility curves.

receiving 10^6, and a person is willing to pay \$5 for such a lottery. Thus, the typical utility curve is concave downward. There are other notable exceptions to this generally risk-averse nature of the decision maker, especially when the utility function is not required to be symmetric about losses and gains.

Utility assessment. The first step in a utility assessment process ideally results in a definition of the scope of the assessment procedure. Elements of the value system are identified and often structured, at least partially, in the issue formulation step. Attributes, according to which alternatives are to be assessed, are identified and structured. The attributes are next related to measurable characteristics of the alternatives. Much of this effort is accomplished in the formulation and analysis steps, and the results should be organized to match the requirements of the utility assessment portion of the interpretation step. The final part of the utility assessment process concerns either elicitation of the weights of each attribute and aggregation of the information to facilitate selecting the best alternative, or the use of a number of holistic judgments and regression analysis to identify these weights. *See* SYSTEMS ENGINEERING.

Utility functions are fully useful only when they are intelligently obtained and thoughtfully applied. Structuring of a decision assessment issue is a crucial part of an overall decision assessment process. The final result will, in practice, depend upon the truthfulness of both the decision situation structural model and the parameters within this structure.

Once the decision maker's utility curve has been defined, the standard tenets of axiomatic decision theory suggest that it may be employed in any decision problem as long as all the outcomes fall between the extremes used to establish the utility function. In principle, the decision maker can now delegate authority to an agent, who can employ this utility curve to arrive at the same decisions that the decision maker would make. However, in reality a person's utility curves continually change, as the person learns new things, has new experiences, or develops new perspectives. Similarly, a utility curve that applies to one kind of problem may offer little insight into another. Thus, in many situations it may not be appropriate to suggest that someone else can learn the decision maker's utility curves, and then operate as if these could be rigidly applied in changing circumstances.

Andrew P. Sage

Bibliography. R. T. Clemen, *Making Hard Decisions: An Introduction to Decision Analysis*, 2d ed., Duxbury Press, Belmont, CA, 1996; J. S. Hammond, R. L. Keeney, and H. Raiffa, *Smart Choices: A Practical Guide to Making Better Decisions*, Harvard Business School Press, Boston, 1999; C. W. Kirkwood, *Strategic Decision Making: Multiobjective Decision Analysis with Spreadsheets*, Duxbury Press, Belmont, CA, 1997; G. Klein, *Sources of Power: How People Make Decisions*, MIT Press, Cambridge, MA, 1998; A. P. Sage, *Systems Engineering*, John Wiley, New York, 1992; A. P. Sage and J. E. Armstrong, *Introduction to Systems Engineering*, John Wiley, New York, 2000.

Decomposition potential

The electrode potential at which the electrolysis current begins to increase appreciably. Decomposition potentials are used as an approximate characteristic of industrial electrode processes. *See* ELECTROCHEMICAL PROCESS; ELECTROLYSIS.

Decomposition potentials are obtained by extrapolation of current-potential curves (see **illus.**).

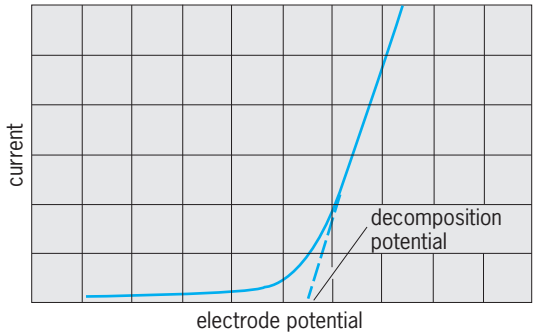

Determination of decomposition potential.

Extrapolation is not precise because there is a progressive increase of current as the electrode potential is varied. The decomposition potential, for a given element, depends on the range of currents being considered. The cell voltage at which electrolysis becomes appreciable is approximately equal to the algebraic sum of the decomposition potentials of the reactions at the two electrodes and the ohmic drop, or voltage drop, in the electrolytic cell. The ohmic drop term is quite negligible for electrolytes with high conductance. For a discussion of current-potential curves. *See* ELECTROLYTIC CONDUCTANCE; OVERVOLTAGE.

Paul Delahay

Decompression illness

Symptoms in humans which result from a sudden reduction in atmospheric pressure. It is also called dysbarism, caisson disease, the bends, and compressed-air illness. It is most commonly seen in two groups

of subjects: those who rapidly ascend in nonpressurized airplanes to altitudes in excess of 18,000 ft (5500 m); and divers, scuba divers, sandhogs, and professional workers in hyperbaric chambers who work under increased ambient pressures and are decompressed. *See* DIVING.

The onset of symptoms may occur at any time from a few minutes to several hours after decompression. The most common manifestation is pain in the joints and muscles. However, skin, respiratory, and neurologic symptoms are not uncommon. The skin manifestations are itching, discoloration, and edema (swelling). Respiratory symptoms are coughing and dyspnea (difficulty in breathing). The neurologic symptoms are of a more grave nature and vary from mild paresthesia (sensations of tingling, crawling, or burning of the skin) and weakness to total paralysis; loss of bladder and rectal sphincter control is common. The severe forms of this illness are followed by circulatory failure, paralysis, coma, and death.

This condition is caused by the formation of nitrogen bubbles in the tissues and blood vessels. In the body the nitrogen normally dissolved in body fluids forms bubbles when the atmospheric pressure is reduced. These bubbles plug vessels and expand in tissue spaces, such as muscles and joints, producing the characteristic symptoms and signs of this illness. Helium, when present in high concentrations utilized by divers or found in some spacecraft, can also produce a similar condition.

Treatment is recompression, followed by gradual decompression to normal atmospheric pressure. Prognosis is generally good except in those subjects who show central nervous system damage. *See* AEROSPACE MEDICINE; BIOPHYSICS; SPACE BIOLOGY.

McChesney Goodall

Decontamination of radioactive materials

The removal of radioactive contamination which is deposited on surfaces or spread throughout a work area. Personnel decontamination is included. Radioactive contamination is a potential health hazard and, in addition, may interfere with the normal operation of plants, particularly when radiation detection instruments are used for control purposes. Thus, the purpose of decontamination is the detection and removal of radioactive contaminants from unwanted locations to locations where they do not create a health hazard or interfere with plant operation. The objective of a good decontamination operation is to remove the radioactive contamination permanently from the human environment, with minimum radiation exposure.

There are four ways in which radioactive contaminants adhere to surfaces, and these limit the decontamination procedures which are applicable. The contaminant may be (1) held more or less loosely by electrostatic forces or surface tension, (2) absorbed in porous materials, (3) mechanically bonded to surfaces through oil, grease, tars, paint, and so on, or (4) adsorbed or chemically bonded to the surface. Typically, the last of these provides the most adherent deposits, often requiring chemical decontamination methods.

Methods. Decontamination methods are mechanical or chemical. Commonly used mechanical methods are vacuum cleaning, sand blasting, blasting with solid carbon dioxide, flame cleaning, scraping, ultrasonic cleaning, vibratory finishing, and using lasers to vaporize contaminants. Scraping includes surface removal—for example, removal of a contaminated layer from concrete floors with an air hammer or scrabbling device, which scrapes off the top layer of concrete mechanically. Strippable coatings are sprayed onto a painted, metal, or concrete surface; contamination is removed when the dried coating is peeled off. Hydrolyzing with high-pressure water is another effective technique. The interior surfaces of piping systems can be cleaned out using rotating brushes (wire or plastic with abrasive tips) or water- or air-propelled abrasive devices.

Chemical methods are used primarily to decontaminate components and tools that are immersed in a tank, either by means of a chemical solvent to dissolve the contaminant, or by using electropolishing techniques to remove the surface layer, including contaminants, from metals. Chemical decontamination methods are also used to remove radioactive deposits from the interior surfaces of piping, pumps, heat exchangers, and boilers. For these applications, the solvent is pumped or flushed through the system, dissolving the radioactive deposits. The solution itself is then radioactive, and the contaminants are typically removed using filters or ion-exchange resins. The use of this approach to clean the coolant systems of nuclear reactors has become common. Dilute chemical reagents, including organic acids, are used to decontaminate the primary coolant systems of operating nuclear power plants to minimize radiation exposure of the workers. An important consideration for plants that are returning to service is to avoid any corrosion of the system components; in these cases, typically 90–95% of the radioactivity is removed, thereby reducing radiation levels by a factor of 10–20. For permanently shutdown plants, more aggressive solvents can be used; complete removal of contaminants is possible, permitting recycling of the metal components.

Collection and disposal of the radioactive residue from mechanical decontaminations can be more difficult, and this is often a major consideration in determining which process to use. In some cases, dust is produced, which has to be removed from the atmosphere using high-efficiency filters. The wet processes, such as hydrolyzing, result in large quantities of water, which then has to be treated. In all cases, it is important that the volume of the final radioactive waste be as small as possible, because large volumes cost more to store and transport to a repository, and disposal charges are greater.

Other considerations, besides waste treatment and disposal, that influence the choice of decontamination method include (1) cost, (2) hazards involved

in the procedure, (3) impact on other plant systems and operations (for example, the need to isolate areas for dust and debris removal), and (4) permanency of the removal (for example, some processes may leave the surface in a rougher condition, resulting in rapid recontamination). For objects that are contaminated with short-lived radioactive material, such as fission products, it may be more cost-effective to postpone the decontamination to allow for radioactive decay. In cases with alpha-emitting contamination, painting of the surface to fix the contaminant in place may provide a permanent solution. *See* ALPHA PARTICLES.

Personnel. Personnel decontamination methods differ from those used for materials primarily because of the possibilities of injury to the subject. Procedures used for normal personal cleanliness usually will remove radioactive contaminants from the skin, and the method used will depend upon the form of the contaminant and associated dirt (grease, oil, soil, and so on). Soap and water (sequestrants and detergents) normally remove more than 99% of the contaminants. If it is necessary to remove the remainder, chemical methods which remove the outer layers of skin can be used. These chemicals, such as citric acid, potassium permanganate, and sodium bisulfite, should be used with caution and preferably under medical supervision because of the risk of injury to the skin surface. The use of coarse cleansing powders should be avoided for skin decontamination, because they may lead to scratches and abraded skin which can permit the radioactive material to enter the body. The use of organic solvents should be avoided because of the probability of penetration through the pores of the skin. It is very difficult to remove radioactive material once it is fixed inside the body, and the safest and most reliable way to prevent radiation exposure from contamination is the application of health physics procedures to prevent entry of radioactive material into the body.

Environment. Air contaminants frequently are eliminated by dispersion into the atmosphere. Certain meteorological conditions, such as prevailing wind velocity and direction, and inversion layers, seriously limit the total amount of radioactive material that may be released safely to the environment in any time period. Consequently, decontamination of the airstream by high-efficiency filters, cyclone separators, scrubbing with caustic solutions, cryogenic removal, and entrapment on charcoal beds is often implemented. The choice of method is guided by airflow requirements, the cost of heating and air conditioning, the cost of waste storage and disposal, the hazards associated with the airborne radioactive material, and the isolation of the operation from other populated areas. *See* AIR FILTER; DUST AND MIST COLLECTION; FILTRATION.

Water decontamination processes can use either maximum dilution or maximum concentration (and subsequent removal) of the contaminant, using water purification processes such as ion exchange, chemical precipitation, flocculation, filtration, and evaporation. Contaminated soil is typically cleaned by washing with dilute chemical solutions of chelating agents, such as ethylenediaminetetraacetic acid (EDTA), or is removed to a regulated landfill. *See* ION EXCHANGE; WATER TREATMENT.

Certain phases of radioactive decontamination procedures are potentially hazardous to personnel. Health physics practices include the use of protective clothing and respirators, localized shielding, isolation or restriction of an area, and the use of remotely controlled equipment. *See* HAZARDOUS WASTE; HEALTH PHYSICS; RADIATION INJURY (BIOLOGY); RADIOACTIVE WASTE MANAGEMENT; RADIOACTIVITY.

Christopher J. Wood

Bibliography. *Dose Control at Nuclear Power Plants*, National Council on Radiation Protection and Measurements, NCRP Rep. 120, 1994; M. Eisenbud and T. F. Gesell, *Environmental Radioactivity*, 4th ed., Academic Press, 1997; J. Shapiro, *Radiation Protection*, 3d ed., Harvard University Press, 1990; J. E. Turner, *Atoms, Radiation, and Radiation Protection*, 2d ed., Wiley, 1995; *Workshop To Define Decontamination and Decommissioning Needs for the U.S. Department of Energy*, Federal Energy Technology Center, 1996.

Deep inelastic collisions

Either highly energetic collisions of elementary particles, namely, leptons and nucleons, which probe the nucleons' internal structure; or collisions between two heavy ions in which the two nuclei interact strongly while their nuclear surfaces overlap.

Elementary-particle collisions. Many collisions between elementary particles either are elastic, in which the particles emerge unscathed, or proceed via the creation of a resonance, in which the particles briefly coalesce and the resonance subsequently decays. Deep inelastic collisions of elementary particles are very energetic collisions between leptons such as electrons or neutrinos and nucleons (that is, protons or neutrons, typically in a nucleus) in which the target nucleon breaks up into many particles and the lepton is scattered through a large angle in the center-of-mass frame. These collisions are akin to the Rutherford scattering experiments in which most alpha particles went through a thin gold foil undeflected but some were deflected through large angles. In both cases, the explanation for large deflections is that the incident particle encounters not a uniform sphere of material but a few hard or pointlike objects inside the target. The alpha particles encounter gold nuclei, while leptons strike quarks inside the nucleons. *See* ALPHA PARTICLES; LEPTON; NUCLEON; RESONANCE (QUANTUM MECHANICS); SCATTERING EXPERIMENTS (NUCLEI).

Discovery of quarks. Quarks were originally postulated by M. Gell-Mann and others in the early 1960s mainly as a mathematical device to explain the observed symmetries in elementary particles known at the time. Deep inelastic scattering provided the first hard evidence of their actual existence inside protons and neutrons, giving quarks a new measure of reality.

In a deep inelastic collision of a lepton of energy E with a nucleon of mass M, the scattering is usually described as a function of two variables: Q^2, which may be thought of as the square of the energy of the exchanged virtual particle, and y, which is the fraction of energy lost by the lepton. In the late 1960s J. Bjorken predicted the onset of scaling at very high energy transfers, that is, the scattering would become a function not of Q^2 and y independently but of the ratio $x \equiv Q^2/(2MEy)$ alone. This was dramatically confirmed in a series of historic electron scattering experiments conducted around 1968. R. Feynman soon showed that an interpretation of the observed collisions in terms of elastic scattering of leptons from pointlike parts of nucleons, which he called partons, leads directly to the same Bjorken scaling. The variable x is simply the fraction of the nucleon's momentum carried by the parton. With these experiments and Feynman's interpretation came the realization that protons and neutrons are not fundamental but are in turn composed of smaller objects that are now called quarks. A whole new substructure of matter had been discovered. *See* QUARKS.

Structure functions. Deep inelastic scattering experiments are now conducted to study the structure of protons and neutrons. In each collision the fraction x of the nucleon's momentum carried by the struck quark is measured, and thus the x distributions of quarks inside a proton are directly measured. These are known as structure functions. Studies of these have shown that the momentum of a proton is not carried entirely by quarks. In fact, only about half the momentum can be ascribed to quarks. The other half is believed to be carried by gluons, which are carriers of the strong force which binds the quarks within nucleons and other hadrons. By studying the evolution of quark structure functions as Q^2 changes, the structure functions of gluons may be inferred by using the theory of strong interactions called quantum chromodynamics (QCD). *See* GLUONS; QUANTUM CHROMODYNAMICS.

The quark structure functions may also be used to deduce the number of valence quarks inside a nucleon. A valence quark, unlike a sea quark, does not arise from a gluon creating a quark-antiquark pair and hence does not have a partner antiquark. Thus, when the number of antiquarks is subtracted from the number of quarks, only the number of valence quarks remains. This relation is called the Gross-Llewellyn-Smith sum rule, and experiments confirm that there are three valence quarks in a nucleon as expected in the quark model of hadrons. Similarly, other sum rules are tested by using structure functions measured by deep inelastic scattering experiments. Attention has focused on measuring the structure functions for quarks with spin aligned or antialigned to the direction of the proton's spin. The surprising conclusion is that only a small fraction of the spin is carried by valence quarks. *See* SUM RULES.

Deep inelastic scattering may also be viewed as a collision between the virtual particle exchanged (see **illus.**) and the nucleon. When viewed in this fashion,

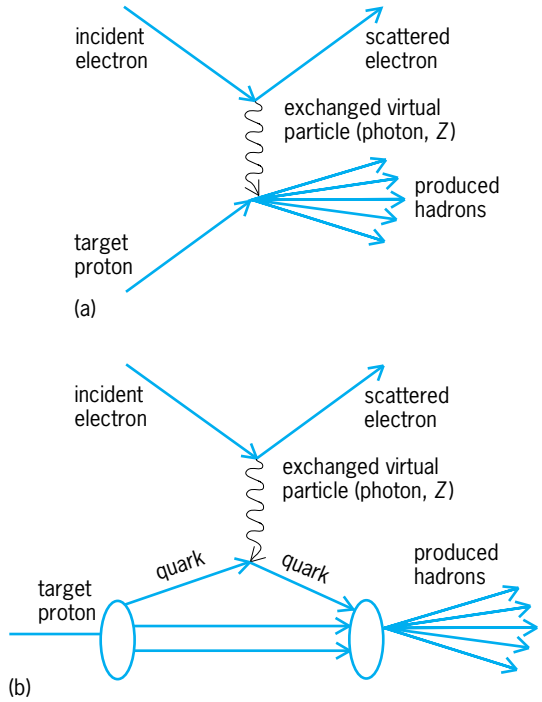

Deep inelastic collision between an electron and a proton where the nucleon is viewed either (*a*) as a blob that breaks up or (*b*) as composed of quarks, one of which is ejected by elastic scattering with the electron.

it is possible to measure the ratio of the scattering cross section from (hypothetical) spin-zero partons and spin-half partons by using the measured cross section for longitudinally and transversely polarized virtual probes. This ratio is expected to be zero and is indeed measured to be close to zero. The difference is understood in QCD as being due to the effect of gluon radiation.

Experiments. Modern-day experiments in deep inelastic scattering often use neutrinos or muons as probes. Neutrinos have the advantage that they are not affected by the electric charge of the target nucleus and hence scatter directly off the quarks, making the study of structure functions with neutrinos particularly attractive. Muons are easy to detect and identify. However, the highest-energy deep inelastic collisions are carried out by using an electron-proton collider. These experiments aim to study structure functions at very low values of x where some models predict new behavior.

It is natural to ask whether there exists yet another layer of tiny particles, smaller than even the quarks. Since quarks cannot be isolated for such a study, the highest-energy deep inelastic collisions use electrons as probes to search for quark substructure.

Milind V. Purohit

Heavy-ion collisions. Deep inelastic collisions of heavy ions are characterized by features that are intermediate between those of comparatively simple quasielastic, few-nucleon transfer reactions and those of the highly complex compound-nucleus reactions. These deep inelastic or damped collisions often occur for heavy-ion reactions at center-of-mass energies less than 5 MeV per nucleon above the

Coulomb barrier. During the brief encounter of the two nuclei, large amounts of kinetic energy of radial and orbital motion can be dissipated. The lifetime of the dinuclear complex (analogous to a chemical molecule) corresponds to the time required for the intermediate system to make a partial rotation (10^{-22} s to 5×10^{-21} s). On separation, the final total kinetic energies of the two reaction fragments can be well below those corresponding to the Coulomb repulsion of spheres, indicating that the fragments are highly deformed in the exit channel, as is known to be the case for fission fragments. *See* NUCLEAR FISSION.

Characteristic features. The principal properties of damped or deep inelastic reactions at bombarding energies of a few megaelectronvolts per nucleon above the interaction barrier are well established. Among these characteristic features are the following: (1) The reactions are binary with two massive reaction fragments in the exit channel. Subsequently, the excited primary fragments, which for large-energy losses have nearly the same temperature, decay via fission or the emission of light particles and gamma rays. (2) Large amounts of kinetic energy of radial and orbital motion can be dissipated in a reaction. (3) Fragment mass and charge distributions are bimodal and are centered close to the masses and charges of projectile and target nuclei. The second moments of these distributions are strongly correlated with kinetic energy dissipation. However, the first moment in $A(Z)$ is nearly independent of energy loss. (4) The angular distributions of the reaction fragments change in a characteristic way with the charge product $Z_P \cdot Z_T$ of the projectile-target combination and the bombarding energy.

Reaction fragment distributions. Experimental studies of the first and second moments of the mass, charge, and isobaric and isotopic distributions of damped reaction fragments and the variation of the moments with total kinetic energy loss and scattering angle have yielded new insight into the microscopic reaction mechanism for heavy ions. Evidence for the magnitude of the charge and mass flow is manifested by the second moments of these distributions which increase with increasing energy loss (where energy loss is a direct measure of the reaction time). Studies of the second moment (variance) in the charge distribution for fixed mass asymmetry (isobaric distribution) have shown that the charge equilibration degree of freedom, like the excitation energy, is relaxed in less than 2×10^{-22} s. In contrast to the short relaxation times for the above two intrinsic degrees of freedom, the mass asymmetry degree of freedom relaxes far more slowly.

Light-particle emission. Studies of light-particle emission from deep inelastic reactions are important because these particles can, in principle, be emitted at any stage of the reaction and hence carry signatures of the dominating mechanisms at work at the different reaction stages. In damped reactions with bombarding energies up to 5 MeV per nucleon above the Coulomb barrier, where hundreds of megaelectronvolts of energy can be dissipated, neutron spectra are consistent with evaporation from the two fully accelerated heavy fragments. For kinetic energy losses greater than 50 MeV, analysis of the neutron spectra indicates that the two fragments have essentially the same temperature as a function of energy loss. In addition, analysis of the differential cross sections of emitted neutrons leads to the conclusion that the excitation energy is divided between the fragments in proportion to their masses. These interesting and important results are, at first, somewhat puzzling when interpreted in terms of simple nucleon exchange models, where the lighter fragment is predicted to have the higher temperature. However, such an initial temperature gradient over the dinuclear complex can be expected to decay through subsequent interactions. The nucleon exchange mechanism itself serves to produce a strong temperature-equalizing effect over the two components of the dinuclear complex with a temperature gradient. The dynamical competition between generation of a temperature difference and its decay can only be estimated from realistic model calculations.

Interpretation. The ultimate goal of experimental and theoretical investigations of damped or deep inelastic reactions between complex nuclei is to achieve an understanding of the reaction properties in terms of the underlying microscopic mechanisms. However, the development of microscopic quantal reaction theories has not progressed far enough to yield a consistent explanation of the observed reaction features. In this situation, a phenomenological interpretation appears to be useful, where certain reaction patterns are analyzed in terms of macroscopic or microscopic concepts, making it possible to study the dependence of model parameters on important reaction conditions. Comparisons of a variety of experimental results with dynamical transport calculations based on one-body nucleon exchange give strong evidence that stochastic exchange of independent nucleons accounts for the dominant part of the dissipative and fluctuative phenomena observed in damped reactions. However, the exact fraction of the energy loss that is due to nucleon exchange and the fraction due to other, more minor processes, including collective modes of energy transformation, remain to be determined. *See* NUCLEAR REACTION; NUCLEAR STRUCTURE.

John R. Huizenga

Bibliography. R. A. Broglia and A. Winther, *Heavy Ion Reactions*, 1991; F. E. Close, *An Introduction to Quarks and Partons*, 1979; R. P. Feynman, *Photon-Hadron Interactions*, 1972, reprint 1998; M. Jacob and P. Landshoff, Inner structure of the proton, *Sci. Amer.*, 242(3):66–75, March 1980.

Deep-marine sediments

The term "deep marine" refers to bathyal sedimentary environments occurring in water deeper than 200 m (650 ft), seaward of the continental shelf break, on the continental slope and the basin (**Fig. 1**). The continental rise, which represents that part of the continental margin between continental

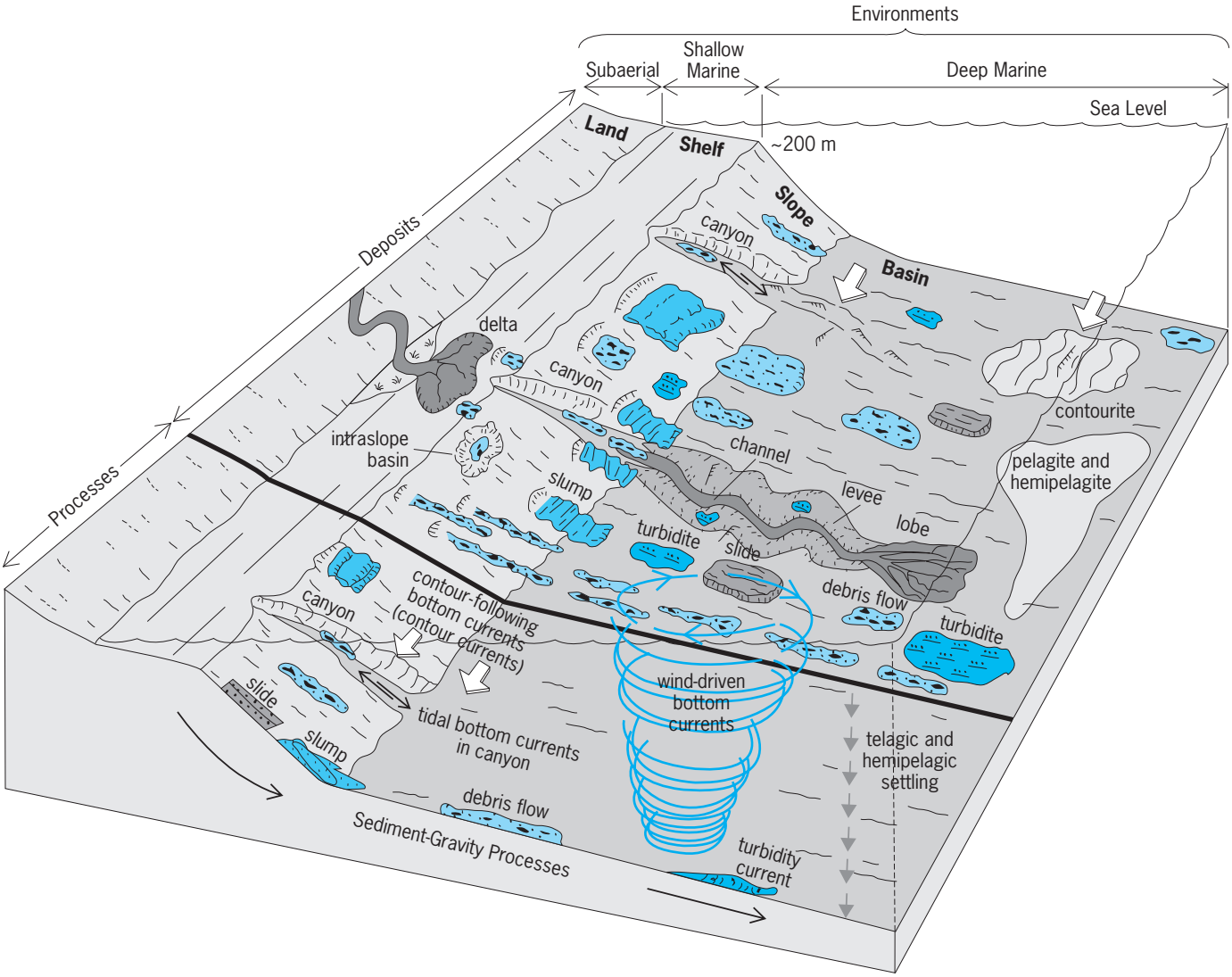

Fig. 1. Slope and basinal deep-marine sedimentary environments occurring at water depths greater than 200 m (650 ft). Slides, slumps, debris flows, turbidity currents, and various bottom currents are important processes in transporting and depositing sediment in the deep sea. Note the complex distribution of deep-marine deposits.

slope and abyssal plain, is included under the broad term "basin." An example of well-developed shelf, slope, and basin settings can be seen in the modern Gulf of Mexico (**Fig. 2**). On the slope and basin environments, sediment-gravity processes (slides, slumps, debris flows, and turbidity currents) and bottom currents are the dominant depositional mechanisms, although pelagic and hemipelagic deposition is also important. *See* BASIN; CONTINENTAL MARGIN; GULF OF MEXICO; MARINE SEDIMENTS.

In 1960s and 1970s, most deep-marine sedimentary systems were thought to be dominated by submarine-fan deposits formed by turbidity currents regardless of whether they were fan-shaped in morphology or whether they were turbidites in origin. A critical evaluation of deep-marine sedimentary systems revealed that, in fact, they are quite complex and that simple fan models are not realistic. Furthermore, the basic tenets of understanding deep-marine processes and their deposits have recently been challenged. In light of these developments, this article focuses on types of deep-marine processes and their deposits; and discusses aspects of deep-marine sedimentary environments in the categories of (1) submarine slopes, (2) submarine canyons and gullies, (3) submarine channels, (4) submarine fans, and (5) submarine basin plains. *See* SUBMARINE CANYON; TURBIDITE; TURBIDITY CURRENT.

Types of processes. The mechanics of deep-marine processes is critical in understanding the nature of transport and deposition of sand and mud in the deep sea. In deep-marine environments, gravity plays the most important role in transporting and depositing sediments. Sediment failure under gravity near the shelf edge commonly initiates gravity-driven deep-marine processes, such as slides, slumps, debris flows, and turbidity currents (Fig. 1). Sedimentary deposits reflect only depositional mechanisms, not transportational mechanisms. Therefore, general characteristics of each process listed below should be used only to interpret depositional mechanisms. *See* GRAVITY; SEDIMENTOLOGY.

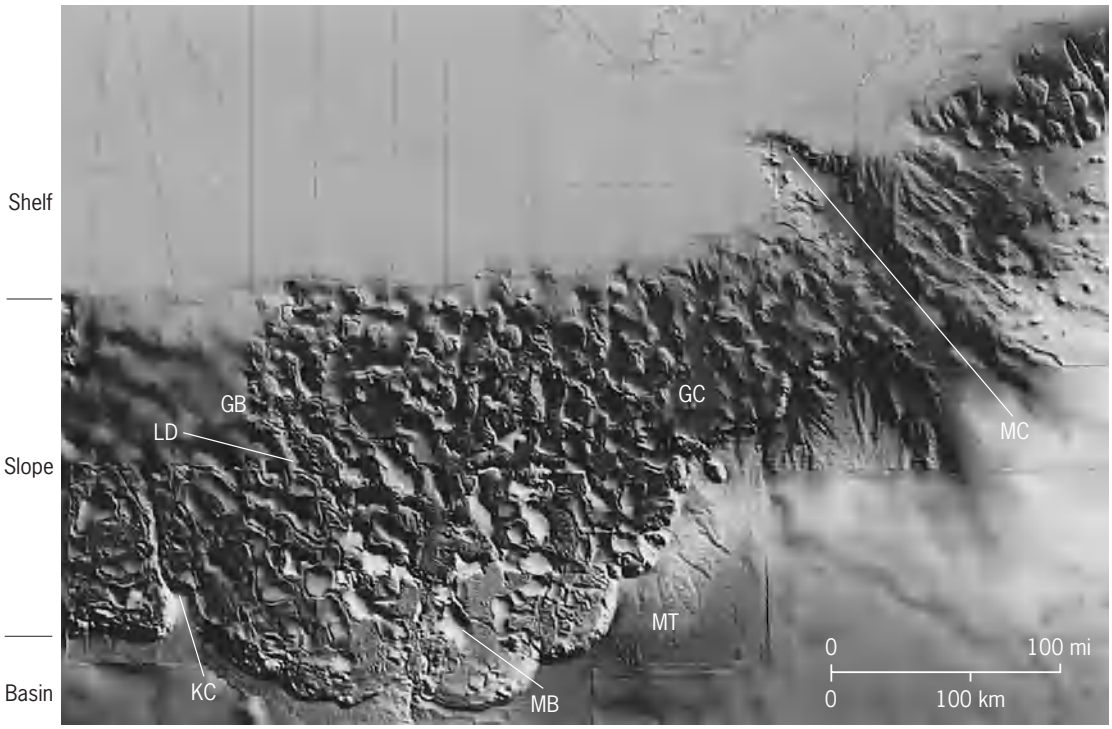

Key:

GB = Garden Banks Area
MT = Mass Transport (slide/slump/debris flow)
MB = Mini Basin
GC = Green Canyon Area
MC = Mississippi Canyon
KC = Keathley Canyon
LD = Linear Depression

Fig. 2. Bathymetric display of the northern Gulf of Mexico based on high-resolution bathymetric swath data. The data are displayed to give a 3-D perspective of present-day shelf, slope, and basin settings. Note the highly irregular sea-floor topography and intraslope basins controlled by salt tectonism. The intraslope basins commonly range in width from 10 to 20 km (6.2 to 12.4 mi).

Slides. A slide is a coherent mass of sediment that moves along a planar glide plane and shows no internal deformation (Fig. 1). Slides represent translational movement. Submarine slides can travel hundreds of kilometers. For example, the runout distance of Nuuanu Slide in offshore Hawaii is 230 km (143 mi). Long runout distances of 50–100 km (31–62 mi) of slides are common.

General characteristics of submarine slide deposits are:

- Gravel-to-mud lithofacies
- Large dimensions, up to hundreds of kilometers long
- Glide plane scar on the main shear surface, and basal zone of shearing
- Upslope areas contain tensional faults
- Downslope edges may show compressional folding or thrusting (that is, toe thrusts)
- Common in areas of tectonics, earthquakes, salt movements, and rapid sedimentation
- Sheetlike geometry with irregular thickness and shape
- Occur on slopes as gentle as 1–4°
- Travel hundreds of kilometers down the gentle slopes (**Fig. 3**)

Examples of submarine slides reveal that modern ones are an order of magnitude larger than their known ancient counterparts. The modern Storegga Slide on the Norwegian continental margin has a width of 290 km (180 mi), a thickness of 430 m (1411 ft), and a width-to-thickness ratio of 675:1 (see **table**). A seismic-scale slide sheet was described, 6 km (3.7 mi) wide and 440 m (1444 ft) thick with a width-to-thickness ratio of 14:1, from the Ablation Point Formation (Kimmeridgian) exposed in Alexander Island of Antarctica. Individual sandy slide sheets reach a thickness of 50 m (164 ft) and a length of 1000 m (3281 ft).

Slumps. A slump is a coherent mass of sediment that moves on a concave-up glide plane and undergoes rotational movements causing internal deformation (Fig. 1).

General characteristics of submarine slump deposits are:

- Gravel-to-mud lithofacies
- Basal zone of shearing
- Slump folds
- Contorted bedding
- Steeply dipping and truncated layers
- Undeformed units above deformed units
- Lenticular-to-sheetlike geometry with irregular thickness

For example, one study reported a thickness of slumps varying from 10 to 374 m (33 to 1227 ft).

Lateral dimensions (width/length) reached up to 64 km (40 mi). Width-to-thickness ratio varied from 10:1 to 3600:1 (see table). In the Cretaceous of the Norwegian North Sea, abrupt lateral changes in slump facies thickness occur across a distance of 29 m (95 ft) between two adjacent wells. Such abrupt terminations of slump deposits have been observed in outcrops in Antarctica. Slump and slide deposits show large vertical dimensions. *See* ANTARCTICA; CRETACEOUS.

Debris flows. A downslope increase in mass disaggregation results in the transformation of slumps into debris flows (Fig. 1). Sediment is now transported as an incoherent viscous mass, as opposed to a coherent mass in slides and slumps. A debris flow is a sediment-gravity flow with plastic rheology (that is, fluids with yield strength) and laminar state. Deposition from debris flows occurs through freezing. The term "debris flow" is used here for both the process and the deposit of that process. The terms "debris flow" and "mass flow" are used interchangeably because each exhibits plastic flow behavior with shear stress distributed throughout the mass. Although only muddy debris flows (debris flows with mud matrix) received attention in the past, recent experimental and field studies show that sandy debris (debris flows with sand matrix) flows are equally important. Rheology is more important than grain-size distribution in controlling sandy debris flows, and the flows can develop in slurries of any grain size (very fine sand to gravel), any sorting (poor to well), any clay content (low to high), and any modality (unimodal and bimodal). *See* RHEOLOGY.

General characteristics of debris-flow deposits (both muddy and sandy types) are:

- Gravel-to-mud lithofacies
- Floating or rafted mudstone clasts near the tops of beds
- Inverse grading of floating mudstone clasts
- Inverse grading, normal grading, inverse-to-normal grading, and no grading of matrix
- Floating quartz pebbles and inverse grading of granules
- Pockets of gravels
- Planar clast fabric
- Preservation of fragile shale clasts
- Irregular, sharp upper contacts and lateral pinch-out geometries
- Side-by-side occurrence of garnet granules (density 3.5–4.3) and quartz granules (density 2.65)
- Lenticular-to-sheetlike geometry

Using outcrop measurements, it can be established that debris-flow deposits are likely to have a width-to-thickness ratio of 30–50:1. In the examples used in this study, thicknesses of debris-flow deposits commonly range from 1 to 60 m (3 to 197 ft), but unusually thick deposits may occur in association with large slides. An example is the Storegga Slide in the Norwegian continental margin, 430 m (1411 ft) thick and 290 km (180 mi) wide, in which debris flow is a dominant facies. Debris-flow facies have width-to-thickness ratios of up to 500:1 or more. This is generally the result of amalgamation of depositional units.

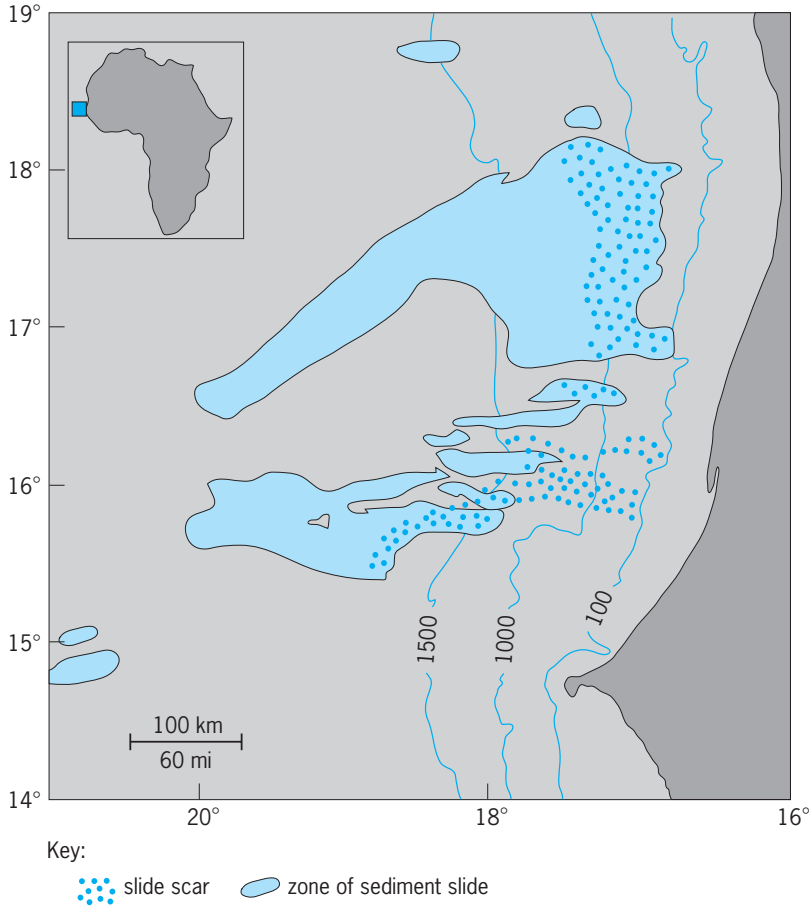

Fig. 3. Map showing long-distance transport of modern slides of about 300 km (186 mi) downslope from the shelf edge (100-fathom, 183-m, or 600-ft contour line) off northwestern Africa. Thickness of slide = 25 m (82 ft). (*From R. D. Jacobi, Sediment slides on the northwestern continental margin of Africa, Mar. Geol., 22:157–173, 1976*)

Turbidity currents. With increasing fluid content, plastic debris flows tend to become newtonian turbidity currents (Fig. 1). Although turbidity currents can constitute a distal end member occurring in basinal areas, they can occur in any part of the system (proximal and distal). Turbidity currents can also occur above debris flows due to flow transformation in density-stratified flows. A turbidity current is a sediment-gravity flow with newtonian rheology (that is, fluids without yield strength) and turbulent state. Deposition from turbidity currents occurs through suspension settling. Deposits of turbidity currents are called turbidites. Although turbidity currents have received a lot of emphasis in the past, other processes are equally important in the deep sea (Fig. 1). In terms of transporting coarse-grained sediment into the deep sea, sandy debris flows and other mass flows appear to play a greater role than turbidity currents.

General characteristics of turbidites are:

- Fine-grained sand-to-mud lithofacies
- Sharp or erosional basal contact
- Gradational upper contact

Dimensions of deep-marine facies

Example	Width:thickness ratio
Slide	
Lower Carboniferous, England	7:1 (100 m wide/long, 15 m thick)
Cambrian-Ordovician, Nevada	30:1 (30 m wide/long, 1 m thick)
Jurassic, Antarctica	45:1 (20 km wide/long, 440 m thick)
Modern, U.S. Atlantic margin	40–80:1 (2–4 km wide/long, 50 m thick)
Modern, Gulf of Alaska	130:1 (15 km wide/long, 115 m thick)
Middle Pliocene, Gulf of Mexico	250:1 (150 km wide/long, 600 m thick)
Slide/slump/debris flow/turbidite	
5000–8000 years B.P., Norwegian continental margin	675:1 (290 km wide/long, 430 m thick)
Slump	
Cambrian-Ordovician, Nevada	10:1 (100 m wide/long, 10 m thick)
Aptian-Albian, Antarctica	10:1 (3.5 km wide/long, 350 m thick)
Slump/slide/debris flow	
Lower Eocene, Gryphon Field, U.K.	20:1 (2.6 km wide/long, 120 m thick)
Paleocene, Faeroe Basin, north of Shetland Islands	28:1 (7 km wide/long, 245 m thick)
Slump	
Modern, southeast Africa	170:1 (64 km wide/long, 374 m thick)
Carboniferous, England	500:1 (5 km wide/long, 10 m thick)
Lower Eocene, Spain	900–3600:1 (18 km wide/long, 5–20 m thick)
Debris flow	
Modern, British Columbia	12:1 (50 m wide/long, 4 m thick)
Cambrian-Ordovician, Nevada	30:1 (300 m wide/long, 10 m thick)
Modern, U.S. Atlantic margin	500–5000:1 (10–100 km wide/long, 20 m thick)
Quaternary, Baffin Bay	1250:1 (75 km wide/long, 60 m thick)
Turbidites (depositional lobes)	
Cretaceous, California	165:1 (10 km wide/long, 60 m thick)
Lower Pliocene, Italy	1200:1 (30 km wide/long, 25 m thick)
Turbidites (basin plain)	
Miocene, Italy	11,400:1 (57 km wide/long, 5 m thick)
16,000 years B.P., Hatteras Abyssal Plain	125,000:1 (500 km wide/long, 4 m thick)

Normal grading without complications (that is, no floating clasts or granules)

Sheetlike geometry in basinal settings. Lenticular geometry may develop in channel settings

In the geologic record, units showing normal grading commonly occur directly above units with inverse grading. Such inverse–to–normally graded beds have traditionally been interpreted as high-density turbidites. However, the concept of high-density turbidity currents and their deposits have come under severe criticism because there are at least five different ways that the concept of high-density turbidity currents has been defined: (1) flow density, (2) grain size, (3) driving force, (4) velocity, and (5) rate of deposition. These five concepts are not in agreement in terms of fluid dynamics. As a result, there are no standard and meaningful ways of interpreting deep-water rocks as high-density turbidites. In terms of fluid rheology and flow state, high-density turbidity currents are nothing but sandy debris flows. *See* FLUID MECHANICS; GEOLOGIC TIME SCALE.

Submarine fan lobes (lobate deposits that accumulate at the mouths of canyons and channels) show width-to-thickness ratios of 165:1 to 1200:1. Lateral dimensions of turbidites deposited on modern abyssal plains can be unusually large. For example, the "Black Shell turbidite" on the Hatteras Abyssal Plain in the western North Atlantic Ocean is 4 m (13 ft) thick and extends for 500 km (311 mi), with a width-to-thickness ratio of 125,000:1. It covers an area of 44,000 km^2 (16,984 mi^2). Classic examples of basin-plain turbidites, such as those exposed along the foreshore at Zumaya, Spain, are usually in the range of 10 cm to 1 m (4 in. to 3 ft) in thickness, but they can be traced for several kilometers. Extensive sheetlike turbidites deposited in the open oceans of Atlantic-type margins are seldom preserved in the geologic record.

Bottom currents. In large modern ocean basins, such as the Atlantic, thermohaline-induced geostrophic bottom currents within the deep and bottom water masses commonly flow approximately parallel to bathymetric contours (that is, along the slope; Fig. 1). They are generally referred to as contour currents. However, because not all bottom currents follow regional bathymetric contours, it is preferred that the term "contour current" be applied only to currents flowing parallel to bathymetric contours, and other currents be termed bottom currents. For example, wind-driven surface currents may flow in a circular motion (Fig. 1) and form eddies that reach the deep-sea floor, such as the Loop Current in the Gulf of Mexico, and the Gulf Stream in the North Atlantic. Local bottom currents that move up- and downslope can be generated by tides and internal waves, especially in submarine canyons. These currents are quite capable of erosion, transportation, and redeposition of fine-to-coarse sand in the deep sea. *See* GULF STREAM; OCEAN CIRCULATION.

Bottom currents (1) generally persist for long time intervals and can develop equilibrium conditions;

(2) transport sand primarily by traction (bedload movement-sliding), rolling, and saltation; (3) are sometimes free of sediment, and are termed clear-water currents; (4) entrain and transport passive sediment particles; (5) are driven by thermohaline, wind, wave, or tidal forces; and (6) commonly flow parallel to the strike of the regional slope (**Fig. 4**) but can also flow in circular motions (gyres) unrelated to the slope. These characteristics clearly discriminate deep-sea bottom currents from turbidity currents. Bottom currents operate parallel to the strike of the slope in most deep-marine settings independently of downslope turbidity currents, debris flows, and slumps. As a result, bottom currents can constantly rework sands introduced into the basin episodically by downslope gravity processes. *See* TIDAL POWER.

Deposits of contour currents (bottom currents) have been termed contourites. However, the general term "bottom-current-reworked sands" is preferred for all types of reworking in the deep sea.

General characteristics of bottom-current-reworked sands are:

- Fine-grained sand and silt lithofacies
- Thin-bedded to laminated sand (usually less than 5 cm or 2 in.) in deep-marine mud
- Rhythmic occurrence of sand and mud layers
- Sharp (nonerosional) upper contacts and sharp-to-gradational bottom contacts
- Internal erosional surfaces
- Well-sorted sand and low depositional matrix (clean sand)
- Inverse size grading (coarsening upward) at various scales
- Horizontal lamination and low-angle cross lamination
- Cross bedding
- Lenticular bedding/starved ripples
- Current ripples and wave ripples
- Mud offshoots in ripples
- Mud-draped ripples
- Alternating traction and suspension structures
- Absence of associated thin units with normal grading
- Double mud layers (tidal)
- Sigmoidal cross bedding (tidal)
- Flaser bedding
- Occurrence of sand layers with traction structures in discrete units, but not as part of a vertical sequence of structures, such as the Bouma Sequence with basal graded division
- Lenticular-to-sheetlike geometry

Pelagic and hemipelagic settling. Pelagic and hemipelagic processes generally refer to settling of mud fractions derived from the continents and shells of microfauna down through the water column throughout the entire deep-ocean floor (Fig. 1). Hemipelagites are deposits of hemipelagic settling of deep-sea mud in which more than 25% of the fraction coarser than 5 micrometers is of terrigenous, volcanogenic, or neritic origin. Although pelagic mud and hemipelagic mud accumulate throughout the entire deep-ocean floor, they are better preserved in parts of abyssal plains (Fig. 1). Rates of sedmentation vary from millimeters to greater than 50 cm (20 in.) per 1000 years, with the highest rates on the upper continental margin.

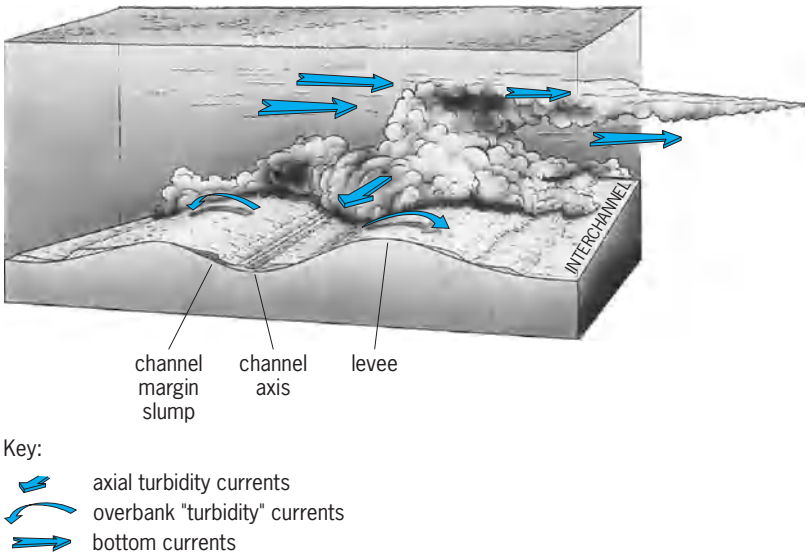

Fig. 4. Conceptual model showing interaction between channelized (axial) turbidity currents, overbank turbidity currents, and bottom currents (contour currents) in deep-sea environments. (*From G. Shanmugam, T. D. Spalding, and D. H. Rofheart, Process sedimentology and reservoir quality of deep-marine bottom-current reworked sands (sandy contourites): An example from the Gulf of Mexico, Amer. Ass. Petrol. Geol. Bull., 77:1241–1259, 1993*)

General characteristics of pelagites and hemipelagites are:

- Mud lithofacies
- Parallel lamination
- Faint normal grading
- Bioturbation
- Deep-marine body fossils and trace fossils
- Sheetlike geometry conformable to underlying sea-floor topography (drape)

Submarine slope environments. Submarine slopes are considered to be of the sea floor between the shelf-slope break and the basin floor (Fig. 1). Modern continental slopes around the world average 4°, but slopes range from less than 1° to greater than 40°. Slopes of active margins (for example, California and Oregon, about 2°) are relatively steeper than those of passive margins (for example, Louisiana, about 0.5°; **Fig. 5**). On constructive continental margins with high sediment input, gravity tectonics involving salt and shale mobility and diapirism forms intraslope basins of various sizes and shapes (for example, Gulf of Mexico; Fig. 2). Erosional features, such as canyons and gullies, characterize intraslope basins. Deposition of sand and mud occurs in intraslope basins. Slope morphology plays a major role in controlling deep-marine deposition through (1) steep versus gentle gradients, (2) presence or absence of canyons and gullies, (3) presence or absence

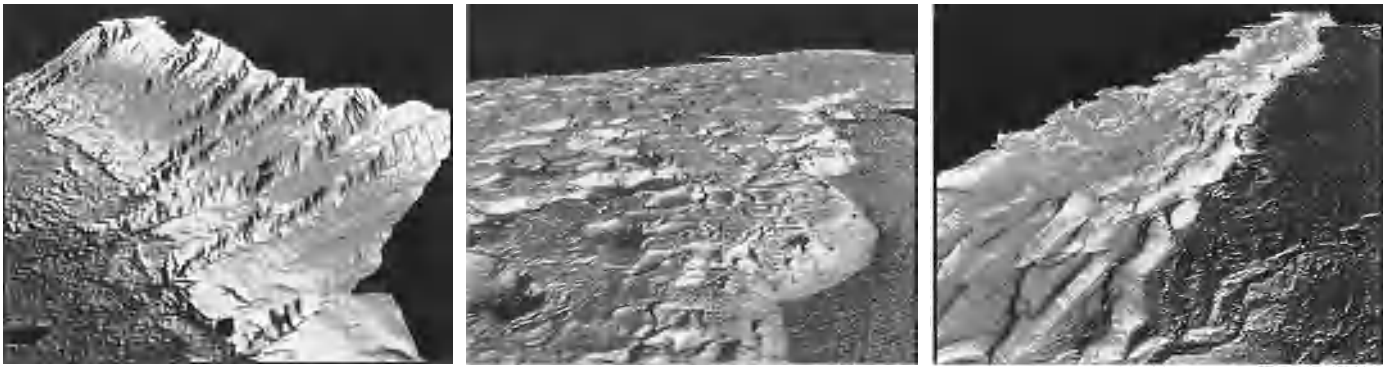

(a) (b) (c)

Fig. 5. Modern United States continental slope showing variability in slope values among different areas: (*a*) New Jersey-Maryland (2.5°); (*b*) Louisiana (0.5°); and (*c*) Oregon (2.0°). (*From L. F. Pratson and W. F. Haxby, What is the slope of the U.S. continental slope?, Geology, 24:3–6, 1996*)

of intraslope basins, and (4) influence of salt tectonics. *See* DIAPIR; EROSION.

Submarine canyon and gully environments. Submarine canyons and gullies are erosional features that tend to occur on the slope. Although canyons are larger than gullies, there are no standardized size criteria to distinguish between them. Submarine canyons are steep-sided valleys that incise the continental slope and shelf. They serve as major conduits for transporting sediment from land and the continental shelf to the basin floor. Modern canyons are relatively narrow, deeply incised, steeply walled, often sinuous valleys with predominantly V-shaped cross sections. Most canyons originate near the continental shelf break and generally extend to the base of the continental slope. Canyons commonly occur off the mouths of large rivers such as the Hudson and Mississippi, although many others, such as the Bering Canyon in the southern Bering Sea, have developed along structural trends. *See* BERING SEA.

Modern submarine canyons vary considerably in their dimensions. Their average length of canyons has been estimated to be about 55 km (34 mi), although the Bering Canyon, the world's longest, is nearly 1100 km (684 mi). The shortest canyons are those off the Hawaiian Islands, with average lengths of about 10 km (6.2 mi).

The water depths in which heads of canyons begin vary from a few meters (near the shoreline) to several hundred meters (shelf break and upper slope). Canyons on the Pacific margin develop at shallower depths than canyons on the Atlantic margin. For example, heads of California canyons begin at an average depth of about 35 m (115 ft), whereas canyons of the east coast of the United States begin at depths greater than 100 m (328 ft). The average depth of canyon termination has been estimated to be 2000 m (6562 ft). The average relief of canyon walls is over 900 m (559 ft). The submarine canyon with the greatest relief is the Great Bahama Canyon, where wall relief is up to 4285 m (14,058 ft).

Physical and biological processes common in submarine canyons are mass wasting, turbidity currents, bottom currents, and bioerosion. Mass wasting is a general term used for the failure, dislodgment, and downslope movement of sediment under the influence of gravity. Common examples of mass wasting are slides, slumps, and debris flows. Major slumping events can lead to formation of submarine canyons. The Mississippi Canyon in the Gulf of Mexico, for example, is believed to have been formed by retrogressive slumping during the Late Pleistocene fall in sea level. During the Holocene rise in sea level, it was partially filled. *See* HOLOCENE; MASS WASTING; PLEISTOCENE.

Francis Shepard made significant contributions to understanding the processes that operate and modify modern submarine canyons. His current-meter records from various submarine canyons around the world revealed that bottom currents (rarely exceeding 50 cm/s or 25 in./s) flow almost continuously up or down canyon axes. According to Shepard, tidal and internal wave forces are the major causes for the semidiurnal up- and downcanyon flow reversals observed in submarine canyons. Such currents can move sediment up and down canyons.

Sea-level changes and tectonic settings are important controlling factors in the development of submarine canyons. Geologic evidence suggests that downcutting of submarine canyons by sediment-gravity flows took place primarily during periods of glacially lowered sea levels. *See* GLACIOLOGY.

Submarine channel environments. According to E. Mutti and W. R. Normark, "A channel is the expression of negative relief produced by confined turbidity current flow, and represents a major, long-term pathway for sediment transport." In other words, all submarine channels are of turbidity current origin, and they all must be long-lived in order to qualify as channels. However, it is not always possible to establish whether a channel in the rock record was cut by a turbidity current or by some other process because processes that cut channels are not necessarily the same processes that fill channels. For example, many channels are filled with deposits of debris flows; however, debris flows are generally not capable of forming erosional channels. Another problem is determining whether an ancient channel acted as a pathway for a long or a short period of time. There are no standardized definitions of long and short periods

of time in deep-water sedimentation. Dimensions of submarine channels vary highly in their widths and depths.

Submarine channels can be erosional, aggradational, or both. Erosional channels commonly develop on the slope, whereas aggradational channels with levees tend to develop on basin floors having gentle gradients. Again, recognition of the type of submarine channel is not easy. Also, there are no objective criteria to distinguish erosional canyons and gullies from erosional channels. Many channels that began as an erosional feature became an aggradational feature. Overbank sediments are characteristic of the aggradational type, and are fine-grained, thin-bedded, and current-laminated. Graded mudstones are dominant in overbank deposits. Persistent overbanking can result in the development of positive relief in the form of levees.

Submarine channels on mature passive-margin fans tend to be relatively long, bifurcating, low-gradient, and largely sinuous. By contrast, channels on active-margin fans are short, steep, and low-sinuosity. The relatively fine-grained (mud-rich) character of the transported sediment associated with channels on mature passive-margin fans, such as the Amazon and Mississippi, gives rise to excellent bank stability and favors the development of a single, largely sinuous commonly meandering channel. Channel shifting in such a system is probably mainly by periodic changes in course (avulsion). However, the origin of sinuous and meandering channels in submarine setting is still poorly understood because, unlike prolonged fluvial currents that develop sinuous channel morphology on land, turbidity currents are episodic events. Therefore, it is difficult to envision how these events could develop the highly meandering channel plan forms observed in the deep sea. Some sinuous channels have been attributed to structural origin (that is, they are controlled by fault motions). *See* GEOMORPHOLOGY.

The sand-rich character and steep gradient of active-margin settings should, by contrast, favor development of a braided channel system. A long-range side-scan sonar (GLORIA) survey of the Orinoco deep-sea fan (mixed setting), which abuts against the deformation front of the Barbados Outer Ridge, also reveals a braided distributary system. *See* SONAR.

High-amplitude reflection packets. In the Amazon Fan in the equatorial Atlantic, channel bifurcation through avulsion is thought to initially lead to deposition of unchannelized sandy flows in the interchannel area (**Fig. 6***a*). Subsequent channel and levee development and progradation over these sandy deposits (Fig. 6*b*) produces a sheetlike geometry at the base of the new channel-levee system that returns high-amplitude reflection packets (HARPs) on seismic data. These sheetlike HARPs overlain by a channel-levee system (gull-wing geometry) are in many ways identical in appearance to a basin-floor fan overlain by a slope fan in a sequence-stratigraphic framework. However, there is a major difference between a basin-floor fan and a HARP. For example, a basin-floor fan is formed by progradation primarily during lowstands of sea level (allocyclic process), whereas HARPs are the results of channel bifurcation (autocyclic process). More importantly, the basin-floor fan and the slope fan are not contemporaneous, whereas a HARP and its overlying channel-levee system are essentially contemporaneous features. This again illustrates that caution must be exercised in interpreting seismic geometries in terms of processes. *See* SEISMIC STRATIGRAPHY; SEQUENCE STRATIGRAPHY.

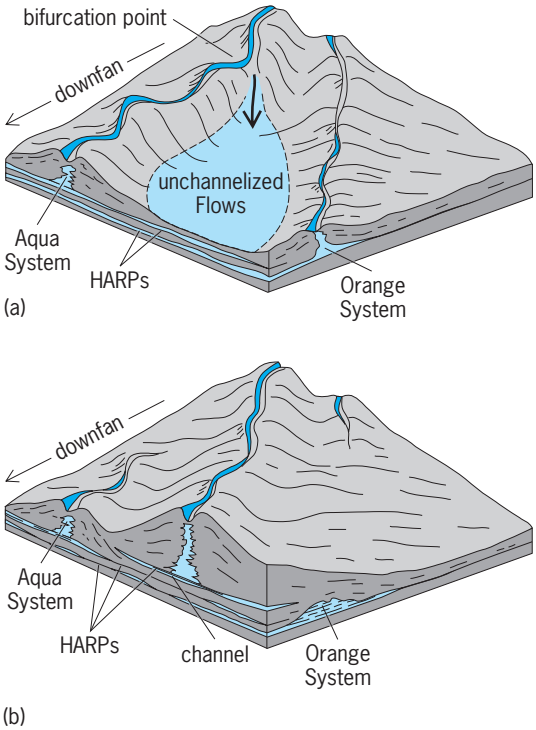

Fig. 6. Amazon Fan. (*a*) Channel bifurcation through avulsion on the deep-sea Amazon Fan results in unchannelized sandy flows by breaching their confining levee through a crevasse and spreading out initially as unchannelized flows into a lower interchannel areas. (*b*) Reestablishment of a new channel above these sandy deposits can result in a sheetlike geometry that returns high-amplitude reflections (HARPs) on seismic data. (*From R. D. Flood et al., Seismic facies and Late Quaternary growth of Amazon submarine fan, in P. Weimer and M. H. Link, eds., Seismic Facies and Sedimentary Processes of Submarine Fans and Turbidite Systems, pp. 415–433, Springer-Verlag, New York, 1991*)

Submarine-fan environments. Submarine fans represent fan-shaped or lobate deposits located in front of submarine canyons or channels. Submarine fans are considered to be formed primarily by turbidity currents coming from a point source. However, not all submarine canyons or channels develop submarine fans. Submarine-fan models, based on turbidite concepts, have been the most influential sedimentologic tools in the petroleum industry for interpreting deep-marine environments.

Sedimentologic fan models. Normark presented the first widely used sedimentologic model for modern submarine fans based on studies of small, sand-rich fans such as the San Lucas and Navy fans off California. He introduced the term "suprafan" to describe the

lobe-shaped bulge on modern fans found immediately downfan of the termination of the major feeder channel. This morphologic feature was presumably formed by rapid deposition of coarse sediment by turbidity currents at the termination of the upper-fan valley. The suprafan lobe was thought to exhibit an overall mounded, hummocky morphology in bathymetric data.

Mutti and Ricci Lucchi proposed submarine-fan models based on outcrop studies in Italy and Spain that popularized the concept of submarine fans with channels in the middle-fan setting and depositional lobes (lobate deposits) in the lower-fan setting. The general characteristics of the depositional lobes of ancient submarine fans include the following: (1) they are considered to develop at or near the mouths of submarine-fan channels analogous to distributary mouth bars in deltaic systems; (2) they show an absence of basal channeling; (3) they usually display thickening-upward depositional cycles composed of classic turbidites; (4) their common thickness range is 3–15 m (9.8–49 ft); and (5) they exhibit sheetlike geometry.

Roger Walker combined the major elements of Normark's model for modern fans with facies concepts of ancient submarine fans of Mutti and Lucchi, and advocated a general fan model with a single feeder channel in the upper-fan area and suprafan lobes in the middle/lower fan areas. Subsequently, this general fan model became influential in hydrocarbon exploration and production because of its predictive capabilities.

Sequence-stratigraphic fan models. Based partly on the suprafan-lobe model of Normark, seismic facies and geometries are used to classify deep-marine systems into basin-floor fans and slope fans in a sequence-stratigraphic framework, and in turn these models are used to predict deposits of specific depositional processes (for example, turbidity currents). However, "turbidity current" has precise meanings in terms of rheology (newtonian) and flow state (turbulence). Evidence for newtonian rheology and flow turbulence cannot be established directly from seismic-reflection profiles or wireline-log motifs; rather, these properties can be ascertained only from actual sediment facies in cores or outcrops. This is because depositional features in cores and outcrops are only centimeters in scale, and these features are too small to be resolved in seismic data. Furthermore, the interpretation of specific seismic facies and geometries (for example, sheet, mounded, continuous, hummocky) as to sediment processes may vary from one worker to the next depending on one's experience. More importantly, a single depositional facies (sandy debris flows) can generate a variety of external seismic geometries (for example, mounded, sheet, and lateral pinchout) and internal seismic reflection patterns (for example, bidirectional downlap, hummocky/chaotic, and parallel/continuous).

Abandonment of submarine-fan models. By the early 1980s, so many deep-marine sedimentary systems were studied that the simple submarine models discussed earlier were proved to be obsolete. Because the morphologic characteristics of modern suprafan lobes are either not preserved in the rock record or cannot be planimetrically mapped in outcrops, Normark abandoned his "suprafan lobe" concept altogether. Walker also recently abandoned his general fan model as "a submarine fan model of the channel-depositional lobe type, influential in its time, but now obsolete because it ignored external controls, especially sea level fluctuations."

The conceptual basin-floor fan model, characterized by mounded seismic facies, predicts sheetlike turbidite sands. However, one study provides a detailed description and interpretation of about 3600 m (12,000 ft) of core through a number of mounded seismic forms of "basin-floor fans" in the North Sea and Norwegian Sea. In these mounded features, turbidites are extremely rare (<1%). Mass-transport deposits, especially slumps, slides, and debris flows, are predominant in the core (50–100%) taken from mounded seismic facies. Recent data also suggest that some of these sands are laterally discontinuous. While features identified as basin-floor fans may occur at specific and predictable stratigraphic positions within a depositional sequence and produce characteristic seismic facies and reflection patterns on seismic data, this core study indicates that basin-floor fans do not represent specific depositional facies (for example, turbidites) and geometries (that is, sheetlike) as the model predicts. Accordingly, seismic mound models (for example, basin-floor fan) should be abandoned because they are based on a model (suprafan-lobe) that is no longer used.

Submarine basin plain environments. The term "abyssal plain" refers to a flat region of the modern deep ocean floor (**Fig. 7**), where the gradient is less than 1:1000. The geologic community, in referring to ancient examples, commonly uses the more general term "basin plain." Basin plains form in response to filling and leveling sea-floor topography by ponding of turbidity currents and by other processes. Because of their flat nature, basin plains favor sheetlike geometries for deposits of turbidity currents and other suspension settling processes (pelagic and hemipelagic). Basin plains are also potential areas for sheetlike slides and debris flows. The Enderby abyssal plain occupies a vast area covering nearly 3.7×10^6 km^2 (1.4×10^6 mi^2; Fig. 7).

Paradigm shift. Although the turbidite paradigm began in the 1950s, understanding of depositional processes in deep-marine environments is still in its infancy. No single facies model can adequately represent all deep-marine sedimentary systems. Submarine fan studies are presently in a state of flux. The days of interpreting complex deep-marine sequences as channels and lobes using fan models are over. Sedimentologic and sequence-stratigraphic lineages of fan concepts dominated by lobes show that their popularity escalated in the 1970s and 1980s but declined in the 1990s to the point of abandonment. The turbidite paradigm reached full circle in the 1990s, completing a remarkable scientific journey. Things are much improved in terms of available marine geological data, core, and outcrop studies,

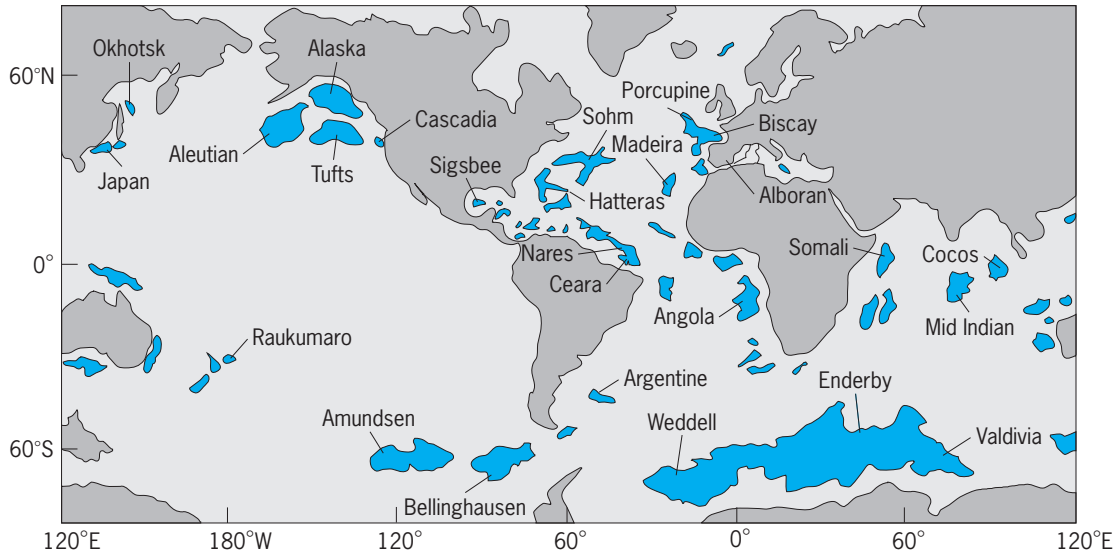

Key:

Alboran: 2600 km^2 (1004 mi^2)
Hatteras: 44,000 km^2 (16,984 mi^2)
Madeira: 54,000 km^2 (20,844 mi^2)
Angola: 1,001,000 km^2 (386,386 mi^2)
Enderby: 3,703,000 km^2 (1,429,358 mi^2)

Fig. 7. Sizes of modern abyssal plains. (*From P. P. E. Weaver, J. Thomson, and P. M. Hunter, Introduction, in P. P. E. Weaver and J. Thomson, eds., Geology and Geochemistry of Abyssal Plains, Geol. Soc. London Spec. Publ., no. 31, pp. vii–xii, 1987*)

theoretical considerations, and flume experiments. Deep-marine systems are quite complex in terms of sea-floor topography, depositional processes, geometries, and stacking patterns. As a result, no single facies model can possibly explain all variations in the complex deep-sea environments.

Mass transport processes (slides, slumps, sand flows, and debris flows) have been observed in modern oceans; however, convincing direct observations of turbidity currents in modern oceans are lacking. It is ironic that there are numerous deep-marine facies models for deposits of turbidity currents that have never been observed, but there are no facies models for deposits of mass flows that are observed. This is perhaps due to the simplicity of turbidity current concepts and submarine-fan models, and the historical association between turbidites and sheet geometries. It is true that basinal turbidites are sheetlike in geometry; however, these turbidite sands are commonly thin-bedded and fine-grained and contain large amounts of mud. In comparison to turbidites, slope sands of debris flow origin are thicker-bedded and coarser-grained and contain lower amounts of mud. The current trend in the petroleum industry is to routinely apply submarine-fan models, developed for base-of-slope settings with smooth sea floors, to intraslope settings with highly irregular sea floors such as in the Gulf of Mexico. However, there is a need to develop separate models for slopes emphasizing slope processes and products. The conventional wisdom that slopes are areas of "bypassing" of sand is not always valid. Slopes are important future target areas where major petroleum reservoirs are waiting to be found. Some slopes undergo extensive gravity tectonic deformation, which leads

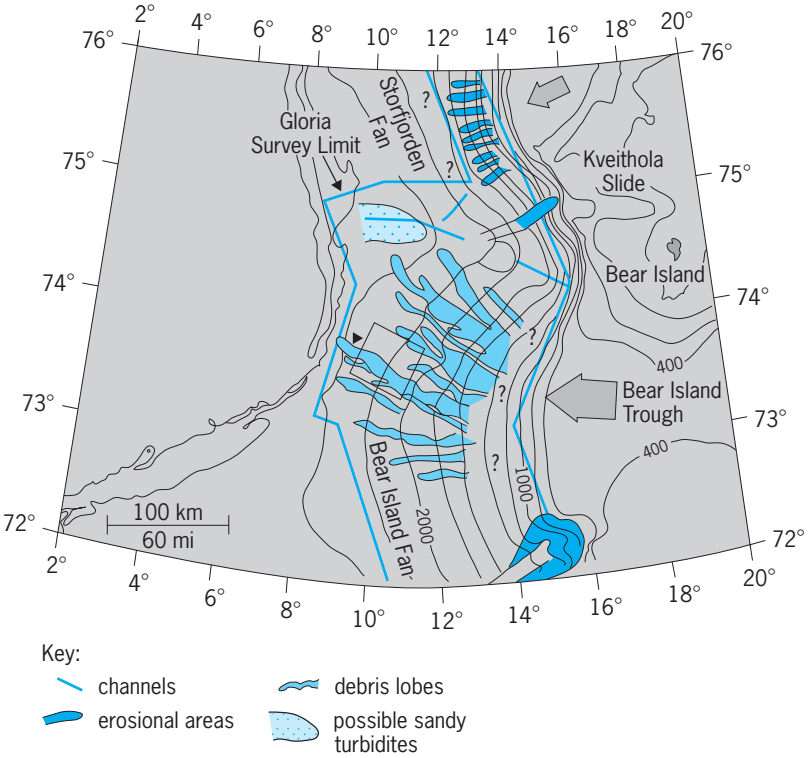

Fig. 8. GLORIA long-range side-scan sonar imagery showing fingerlike patterns of debris flows in the Bear Island Fan on the continental margin of the modern Norwegian-Barents Sea. These fans are composed of cohesive debris flows. Properties of the Bear Island Fan are: runout distance 100–200 km (62–124 mi); width 5–25 km (3–16 mi); thickness 10–50 m (33–164 ft); volume 10–30 km^3 (2.4–7.2 mi^3); slope angle—midslope 0.5°, lower slope 0.2°. (*From A. Elverhoi et al., On the origin and flow behavior of submarine slides on deep-sea fans along the Norwegian-Barents Sea continental margin, Geo-Mar. Lett., 17:119–125, 1997*)

to development of diapirs and intraslope basins with erosional features, such as canyons and gullies (for example, northern Gulf of Mexico, Niger Delta). Deposition of thick sand bodies can occur in such intraslope basins. Future models should take into account the great wealth of information available on modern and ancient slope processes and products such as ancient sandy slides, muddy slides, slumps, and debris flows. Debris flows can travel hundreds of kilometers on gentle gradients. On continental margins, debris flows tend to develop fingerlike patterns (**Fig. 8**). These details are seldom included in popular deep-marine models. *See* PETROLEUM ENGINEERING.

A new paradigm for deep-marine systems will emerge that will be more inclusive in terms of debris flows and bottom currents than just turbidity currents and fan models. This paradigm shift will emphasize that deep-marine systems are extremely complex and each case is unique. G. Shanmugam

Bibliography. J. D. Clark and K. T. Pickering, *Submarine Channels: Processes and Architecture*, Vallis Press, London, 1996; A. Elverhoi et al., On the origin and flow behavior of submarine slides on deep-sea fans along the Norwegian-Barents Sea continental margin, *Geo-Mar. Lett.*, 17:119–125, 1997; R. D. Flood et al., Seismic facies and Late Quaternary growth of Amazon submarine fan, in P. Weimer and M. H. Link (eds.), *Seismic Facies and Sedimentary Processes of Submarine Fans and Turbidite Systems*, pp. 415–433, Springer-Verlag, New York, 1991; D. I. M. Macdonald, A. C. M. Moncrieff, and P. J. Butterworth, Giant slide deposits from a Mesozoic fore-arc basin, Alexander Island, Antarctica, *Geology*, 21:1047–1050, 1993; L. F. Pratson and W. F. Haxby, What is the slope of the U.S. continental slope?, *Geology*, 24:3–6, 1996; G. Shanmugam, The Bouma Sequence and the turbidite mind set, *Earth-Sci. Rev.*, 42:201–229, 1997; G. Shanmugam, Fifty years of turbidite paradigm (1950s–1990s): Deep-water processes and facies models—A critical perspective, *Mar. Petrol. Geol.*, 17:285–342, 2000; G. Shanmugam, High-density turbidity currents: Are they sandy debris flows?, *J. Sed. Res.*, 66:2–10, 1996.

Deep-sea fauna

The deep sea may be regarded as that part of the ocean below the upper limit of the continental slopes (**Fig. 1**). Its waters fill the deep ocean basins, cover about two-thirds of the Earth's surface, have an average depth of about 12,000 ft (4000 m), and provide living space for communities of animals that are quite different from those inhabiting the land-fringing waters which overlie the continental shelves (neritic zone). *See* ECOLOGICAL COMMUNITIES.

The systematic exploration of the deep sea began with the voyage of the HMS *Challenger* (1872–1876). Since that time there have been numerous large-scale, deep-sea expeditions. In 1948 the Swedish Deep Sea Expedition in the Atlantic developed new techniques for trawling and the winch used by later expeditions. *See* MARINE BIOLOGICAL SAMPLING.

The deep-sea fauna consists of pelagic animals (swimming and floating forms between the surface and deep-sea floor) and below these the benthos, or bottom dwellers, which live on or near the ocean bottom. Pelagic animals can be divided into the usually smaller forms that tend to drift with the currents (zooplankton) and the larger and more active nekton, such as squids, fishes, and cetaceans. Pelagic, deep-sea animals are frequently termed bathypelagic in contrast to the epipelagic organisms of the surface waters (Fig. 1).

Bathypelagic fauna. All animal life in the sea, pelagic and benthic, depends on the growth of microscopic plants (phytoplankton). From the surface down to a maximum depth of about 300 ft (100 m) there is sufficient light for photosynthesis and vigorous phytoplanktonic growth. This layer is known as the photic zone. In the deep sea, plants can exist only as saprophytes. The productivity of the plants, however, is reflected down to the deepest parts of the sea through complex food chains. These consist of zooplankton that graze on phytoplankton, carnivorous species that feed on zooplankton, and large predators that eat the other animals. The typical bathypelagic animals (**Fig. 2**) begin to appear below depths of about 600 ft (200 m). *See* FOOD WEB; PHYTOPLANKTON; ZOOPLANKTON.

Zooplankton. The planktonic or drifting forms of animal life in the ocean include the Protozoa, larval stages of deep-sea fishes, and even larger organisms with limited powers of movement.

1. *Protozoa*. Included in this group are various families of Foraminiferida and Radiolaria, the skeletons of which form an important part of the deep-sea sediments. *See* PROTOZOA.

2. *Coelenterata*. Scyphomedusae such as *Atolla* and *Periphylla* are not uncommon. Other jellyfishes include various Trachymedusae (*Crossota* and *Colobonema*) and Narcomedusae. Siphonophorida, particularly the diphyids, are found down to depths of at least 9000 ft (3000 m) but are more common in the upper several hundred feet. *See* CNIDARIA.

3. *Nemertea*. This group of worms has bathypelagic species belonging to some 10 families.

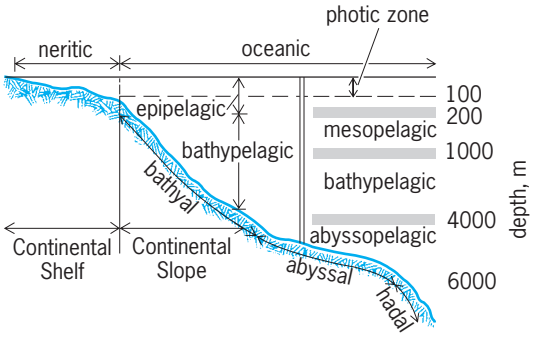

Fig. 1. Classification of marine environments. Right side of diagram illustrates the proposal to divide the bathypelagic zone into mesopelagic, bathypelagic, and abyssopelagic zones. Division of benthic region into bathyal, abyssal, and hadal zones also is shown. 1 m = 3.3 ft.

4. *Crustacea*. In numbers of species and individuals, the small Copepoda are the dominant group of crustaceans in the ocean. There are numerous bathypelagic species. Certain of the Ostracoda (*Gigantocypris*) are purely bathypelagic, as are some of the Amphipoda (the gammarid genera *Cyphocaris* and *Hyperiopsis* and most species of the hyperiid families Scinidae and Lanceolidae). *See* AMPHIPODA; COPEPODA; OSTRACODA.

The larger and more active pelagic crustaceans (Euphausiacea, various Mysidacea, and prawns) are usually classed as plankton but might well be called "micronekton," a group intermediate between thrusting nekton and feebler-swimming plankton. Of the euphausiid shrimps, *Benthenphausia* and various species of *Thysanopoda*, *Nematoscelis*, and *Stylocheiron* have centers of abundance in the bathypelagic zone. Deep-water genera of mysids include *Gnathophausia*, *Lophogaster*, and *Eucopia*, while the prawn families Hoplophoridae and Sergestidae have numerous bathypelagic representatives. *See* CRUSTACEA.

5. *Chaetognatha*. Certain species of arrow-worms, such as *Eukrohnia fowleri* and *Sagitta macrocephals*, are predominantly bathypelagic. *See* CHAETOGNATHA.

6. *Echinodermata: Holothuroidea*. The genera *Pelagothauria*, *Enypiastes*, and *Galatheathuria* are bathypelagic, the first two being medusalike. *See* ECHINODERMATA.

7. *Protochordata: Thaliacea*. While they are more abundant in the surface layers, the salps, doliolids, and pyrosomes have been fished down to 9000 ft (3000 m).

Deep-sea nekton. This group consists largely of squids, octopods, and fishes. The sperm whale also enters the deep sea, where it finds some of the squid upon which it feeds.

1. *Mollusca: Cephalopoda*. Several families of squids form an important part of the deep-sea nekton together with a few octopods, such as *Cirrothauma*, *Amphitretus*, *Vitreledonella*, and *Vampyroteuthis*. *See* CEPHALOPODA.

2. *Fishes*. Apart from a few squaloid sharks, the bathypelagic fish fauna consists of teleosts. The most diverse groups are the stomiatoids (Elupeiformes), with about 300 species; Myctophidae (lantern fishes; Salmoniformes), about 250 species; and the ceratioid angler fishes (Lophiiformes), about 90 species. The few species forming the orders Anguilliformes (gulper eels) and Cetomimiformes (whale fishes) are entirely bathypelagic, as are certain of the eels (Cyemidae, Nemichthyidae) and Beryciformes (for example, Melamphaidae). *See* TELEOSTEI.

Distribution. The bathypelagic fauna is most diverse in the tropical and temperate parts of the ocean. Numerous species are found in all three temperature zones, but many appear to have a more limited distribution.

Each species also has a definite vertical occurrence. Findings suggest that there are three main vertical zones, each with a characteristic community. Here the term bathypelagic is used for the fauna between about 3000 and 6000 ft (1000 and 2000 m), that above (between 600 and 3000 ft or 200 and 1000 m) being called mesopelagic and that below 6000 ft (2000 m) abyssopelagic (Fig. 1). The typical forms of the mesopelagic fauna (stomiatoids and lantern fishes) live in the twilight zone of the deep sea (between the 68 and 50°F or 20 and 10°C isotherms), while the bathypelagic species (ceratioid angler fishes and *Vampyroteuthis*) occur in the dark, cooler parts below the 50°F (10°C) isotherm.

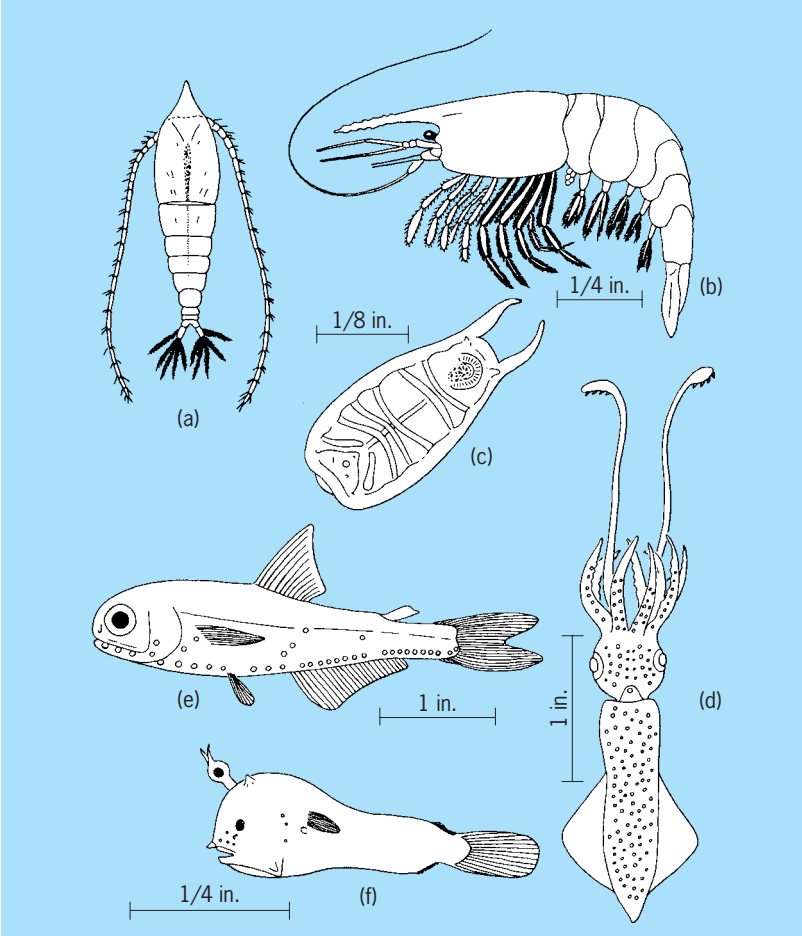

Fig. 2. Pelagic animals of the deep sea. (*a*) Copepod (*Haloptilus acutifrons*). (*b*) Prawn (*Acanthephyra multispina*). (*c*) Salp (*Salpa*, or *Thalia*, *democratica*). (*d*) Squid (*Abraliopsis morisii*). (*e*) Lantern fish (*Myctophym punctatum*). (*f*) Anglerfish (*Lophodolus acanthognathus*).

Lastly, numerous species of mesopelagic animals, such as euphausiids, prawns, squids, and fishes (particularly lantern fishes), undertake extensive diurnal, vertical migrations, moving upward into the productive surface layers to feed at night. Toward sunrise they begin to descend to their daytime levels. *See* SCATTERING LAYER.

Bioluminescence. Perhaps the most conspicuous feature of pelagic deep-sea life is the widespread occurrence of luminescent species bearing definite light organs (photophores). Many of the squids and fishes have definite patterns of such lights, as do some of the larger crustaceans (hoplophorid and sergestid prawns and euphausiids). Investigations have shown that flashes from luminescent organisms could be

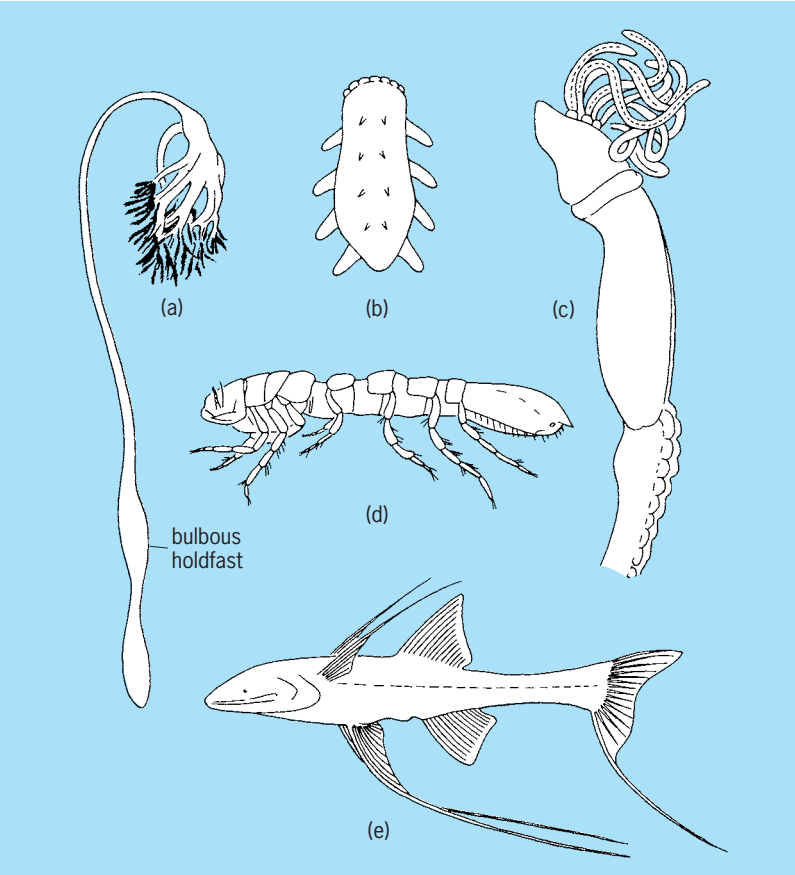

Fig. 3. Benthic animals of the deep sea. (*a*) Sea-pen (*Umbellula*) showing bulbous holdfast. (*b*) Sea cucumber (*Elpidia glacialis*). (*c*) Head and forepart of the trunk of a pogonophoran (*Birsteinia witjasi*). (*d*) Isopod crustacean (*Macrostylus hadalis*). (*e*) Bathypteroid fish (*Benthosaurus*).

detected down to depths of 12,300 ft (3750 m). *See* BIOLUMINESCENCE; PHOTOPHORE GLAND.

Benthic fauna. There are two main ecological groups of bottom-living animals (**Fig. 3**) in the ocean: organisms that attach to the bottom and those that freely move over the bottom.

Attached benthic organisms. This group consists of species that attach themselves to the sediments, rocks, or to other organisms. The more typical forms include the following: Hexactinellida (Porifera) (glass sponges), with about 375 species; certain hydroids, gorgonians, pennatulids (sea fans), antipatharians (black corals), actiniarians (sea anemones), and madrepore corals (*Lophohelia* and *Amphihelia*); Cirripedia (barnacles), such as *Scalpellum* and *Verruca* spp.; numerous species of stalked crinoids living in the deep sea together with a number of unstalked forms; and Pogonophora (beard-bearers) and certain ascidians (*Culeolus* sp.).

Benthic crawlers and swimmers. This group comprises the freely moving animals, those that swim or crawl over the bottom or burrow into the sediments, the upper layer of which has a rich bacterial flora. (1) Annelida: a few species of Polychaeta (bristle worms). (2) Gephyrea: certain species of echiuroid and sipunculoid worms. (3) Crustacea. In numbers of species and individuals the most important group of benthic deep-sea crustaceans is the Peracarida, represented by various species of cumaceans (*Bathycuma*, *Macrocylindrus*), isopods (*Ischnomesus* and *Eurycope*), amphipods, and tanaids (*Apseudes*, *Neotanais*). Of the Eucarida, the most prominent groups are the penaeid prawns and the Eryonidae. There are also a number of crabs (*Platymaia*, *Geryon*, *Ethusa*, and *Scyramathia*) and hermit crabs (numerous species of Axiidae). (4) Pycnogonida (sea spiders): numerous species of the families Colossendeidae and Nymphonidae. (5) Mollusca. Certain of the Octopodidae (octopuses) and the cirromorph octopods live on the deep-sea floor, as do various gastropods, scaphopods, and lamellibranchs. (6) Echinodermata. These form an important part of the benthic fauna, particularly the sea cucumbers (Holothuroidea) of the orders Elasipodida and Molpadida. Among the sea urchins (Echinoidea), the order Cidaroida and the suborder Meridosternata mainly consist of deep-sea species. There are also various brittle stars (Ophiuroidea) and starfishes (Asteroidea). (7) Fishes. The species of one group of cartilaginous fishes (Holocephali) live over the continental slope. The main groups of benthic, deep-sea teleosts are the Bathypteroidae (Salmoniformes); Halosauridae and Notacanthidae (Notacanthiformes); Macrouridae (rattails) and Morinae (deep-sea cods) (Gadiformes); and Brotulidae, Liparidae, and Zoarcidae (Perciformes). *See* ECHIURIDA; POLYCHAETA; PYCNOGONIDA; SIPUNCULA.

Distribution. The benthic fauna is most diverse in the temperate and tropical ocean, although the arctic and antarctic areas have their characteristic species. As in the pelagic fauna, certain species occur in all three oceanic zones, while others appear to have a more restricted occurrence.

While a number of species—particularly among the polychaete worms, gastropod mollusks, and the brittle stars (Ophiuroidea)—range from littoral to abyssal regions, most forms tend to live within smaller ranges of depth. Data suggest that there are typical communities of animals over the continental slopes (Fig. 1) extending down to about 9000 ft (3000 m; bathyal zone); others occur below this in the abyssal zone. Danish and Russian exploration also suggested that the deep-sea trenches (with depths over 21,000 ft or 7000 m) form another ecological zone (hadal zone) having certain characteristic species—those capable of living under pressures of 700–1000 atm (70–100 megapascals; barophilic species). This work also showed that life could exist at the very bottom of the ocean (down to depths of more than 30,000 ft or 10,000 m) and that species of certain groups, such as sea anemones, echiuroid and polychaete worms, bivalves, isopod and amphipod crustaceans, sea cucumbers, and Pogonophora, occurred at depths beyond 27,000 ft (9000 m).

Lastly, there is a decrease in the numbers of species and individuals with depth. Russian biologists found that at depths of 24,000–30,000 ft (8000–10,000 m)

the weight of animals per square meter of sea floor was about one-fifth to one-fifteenth the weight at depths of 3000–12,000 ft (1000–4000 m). As the deep-sea benthic fauna is dependent on organic matter originating in the upper, plant-bearing waters and as the amount reaching the bottom must decrease with depth, the above findings are comprehensible. It is also interesting that there are very few carnivorous animals, such as crabs, brittle stars, and starfishes, below a depth of 21,000 ft (7000 m). It is the particle-catchers, such as the Pogonophora, and ooze-eaters, such as sea cucumbers and echiuroid worms, that make up most of the hadal fauna. *See* MARINE ECOLOGY; SEAWATER FERTILITY.

Norman B. Marshall

Bibliography. C. P. Idyll, *Abyss: The Deep Sea and the Creatures That Live in It*, 3d rev. ed., 1976; N. B. Marshall, *Aspects of Deep Sea Biology*, 1977; N. B. Marshall, *Developments in Deep-Sea Biology*, 1980.

Deep-sea trench

A long, narrow, characteristically very deep and asymmetrical depression of the sea floor, with relatively steep sides. Oceanic trenches characterize active margins at the ocean-basin–continent or ocean-basin–island-arc boundaries. They contain the greatest oceanic depths and are associated with the most active volcanism, largest negative gravity anomalies, most frequent shallow seismicity, and almost all of the intermediate and deep-focus earthquake activity. As the surface expression of the widely accepted process of subduction by which oceanic crustal material is returned to the upper mantle, they are key elements in current models of plate tectonic evolution on Earth and possibly on Venus.

Occurrence. Deep-sea trenches are the signature relief form of the Pacific; in a counterclockwise direction, they occur from southern Chile to just northeast of North Island, New Zealand (**Fig. 1**). A secondary or outer branch trends southward from near Tokyo Bay in a festoon of arcs to south of Palau. The principal gaps in the circum-Pacific chain are from Baja California to southeastern Alaska, and off the northern coast of New Guinea. From eastern New Guinea to southern Vanuatu the trenches lie southwest, or "inside," the island chains; otherwise their characteristics are like those facing the Pacific. The Indian Ocean contains only the very long contorted Sunda Trench that appears near the northwestern end of Sumatra and extends southeast and east past Timor, to curve north and west near Aru and end adjacent to Buru. In the Atlantic, the Puerto Rico–Antillean trench system extends outside the island arc from eastern Hispaniola around to Trinidad; but south of 14°N, off Barbados, the trench is filled with sediment. In the far South Atlantic a typical island-arc–trench complex extends from near South Georgia through the South Sandwich Archipelago.

Mid-to-late-nineteenth century exploration indicated that the deepest areas of the oceans lay near continents or island arcs, in elongated depressions. Understanding of the trenches' fundamental significance to global tectonics and basin evolution has come only during the twentieth century, with the development and application of precise or indicative techniques of measurement and observation. Examples include pinpointing from land observatories the distribution and focal mechanism of deep earthquakes; laborious shipboard collections to relate geophysical data on crustal structure, gravity, magnetic patterns, and precise topography; and geological-geochemical data from igneous rocks and sediment samples obtained by dredging, coring, or drilling.

Geophysical exploration. A series of pioneering gravity observations with pendulum instruments on Dutch submarines during the 1920s and 1930s established that the East Indian trenches, and several others, were characterized by a belt of negative gravity anomalies of 150–200+ milligals, that is, values 150–200 parts per million less than normal, interpretable as deficiency of mass near and at their axes. Investigators attributed the deficiency in gravity to a symmetrical downbuckle of lighter crust into the mantle, a result of compression at unstable boundaries that might have produced the trenches. This postulated root, called a tectogene, was produced in physical model studies and was widely accepted until the development of shipboard seismic refraction techniques and more precise station seismology foci measurements after World War II.

Seismologists in California and Japan independently noted the association of great oceanic deeps such as Tonga Trench, the Peru-Chile Trench, and the Western Pacific trenches with sloping zones of shallow–intermediate–deep-focus earthquakes, and attributed a fault origin to oceanic deeps, postulating reverse faulting with the continental or island-arc side as the overthrusting member. In some areas, for example, within the "interior" Melanesian complex, the directions were reversed, but relative displacement was by thrust faulting. This gross process, as a specific mechanism for this phenomenon, had considerable appeal.

Other seismologists, interpreting the reduced gravity differently, postulated a tensional origin for trenches, modeling a thinning of the crust there. With the application of two-ship seismic refraction methods, it was established that oceanic crust is thin, that crust under island arcs is thicker, and its layers display different sound transmission velocities, indicating different composition. Shipboard studies in the Middle America, Tonga, Cedros, Aleutian, Peru-Chile, and Sunda trenches established that the characteristic oceanic crustal layer [that is, 6.8–7.0 km/s (4.1–4.2 mi/s) compressional wave velocity] does not end or thin under the trench; rather, it may thicken slightly but does deepen steeply as it passes beneath the island arc or continental slope by the process that has come to be known as subduction. The shoreward flank of the trench consists of complexes of materials of intermediate velocity, such as

Fig. 1. Deep-sea trenches in the Pacific-Melanesian-Indonesian hemisphere. Broken line indicates filled trench.

volcanics or contorted sediments. Measurements of heat flow through the sea floor on the outer flanks and on the trench floors were lower than normal for deep-sea areas, suggesting progressive cooling toward the trench, or possibly blanketing by sediments there. *See* EARTH, GRAVITY FIELD OF; EARTH, HEAT FLOW IN; EARTH CRUST; FAULT AND FAULT STRUCTURES; OCEANIC ISLANDS; SEISMOLOGY; SUBDUCTION ZONES.

Sea-floor topography and surficial processes. Topographic studies established that, in cross section, trenches are asymmetrical, with the nearshore flank the steeper, often 4–8° on the shallower foreslope becoming 10–16° on the lower slopes, and locally with precipices of 45° or more. Benches or small flat sediment ponds behind pinnacles are common, providing evidence of talus movement or mass wasting of sediment downward, probably triggered by earthquakes. Bottom photography, specifically of the Tonga Trench, shows evidence of slumps, ripple marks on sediment surfaces at great depths, and angular boulders or slickensided outcrops on the deeper reaches of the shoreward flank. The offshore flanks of trenches almost always are gentler in slope, still steepening downward, often with evidence of fracturing under tension, specifically grabens by normal faulting. Trenchward tilting of the beveled summits of at least two flat-topped seamounts on the outer slopes of the Tonga and Aleutian trenches provides direct evidence of depression and flexing on the outer slope. A low crustal swell outside the trench, required by some models, frequently is recognized. Bathymetric exploration of the trenches has been sufficient to establish a definitive ranking of their maximum depths (see **table**). *See* GRABEN; MARINE GEOLOGY; MARINE SEDIMENTS; MASS WASTING; SEAMOUNT AND GUYOT.

Crustal composition: evidence from dredging. Only with the development of ultrastrong dredging cable and stroboscopic-timed equipment-positioning devices has it been possible to collect well-located specimens of the consolidated sediment and

Trenches ranked by bathymetric depth

Trench	Maximum depth, m*
Challenger Deep, Mariana Trench, western Pacific	10,920±10
Tonga, southwestern Pacific	10,800±5
Philippine, western Pacific	10,057±5
Kermadec, southwestern Pacific	10,047±50
Izu-Ogasawara, western Pacific	9,780±10
Kuril, northwestern Pacific	9,550±5
North New Hebrides, Melanesia	9,175±10
New Britain, Melanesia	8,940±10
Yap, western Pacific	8,650±10
Puerto Rico, western Atlantic	8,605±5
South Sandwich, southern Atlantic	8,325±10
South Solomon, Melanesia	8,322±10
Peru-Chile, southeastern Pacific	8,170±10
Palau, western Pacific	8,055±10
Aleutian, northern Pacific	7,679±?
Ryuku (Nanseishoto), western Pacific	7,460±10
Sunda, northeastern Indian Ocean	7,125±10
Middle America, eastern Pacific	6,662±10

*1 m = 3.3 ft.

volcanic-plutonic igneous rocks exposed on the deep flanks of the great Pacific trenches. Meanwhile high-energy seismic reflection systems with complicated arrays and computerized data-processing methods have revealed the structural complexities—fault traces, contorted layer interfaces, rock slivers, shingling, sediment packets—that characterize other trenches' onshore flanks. There, accretion of upper crustal material from offshore by scraping, or by slumping or mass wasting of upper slope sediments, are dominant processes; evidence comprises the filled or filling trenches in the southern Caribbean Antilles, off southern Chile, off Central America, or even exposed in the islands off Sumatra by uplift.

The three deepest trenches—Mariana, Tonga, and Philippine—all display exposures or in-place outcrops of igneous rock on the deeper 5000 m (16,500 ft) of their nearshore flanks. In the Philippine Trench at its deepest off Mindanao there are volcanic rocks—basalts, basaltic andesite, and andesite transitional from arc-tholeiitic to arc-calcakaline series—as flows or dikes. These volcanics appear characteristic of the Samar–East Mindanao arc massif; there is no evidence of accretion of material from the subducted Philippine Sea oceanic crust, or even a deformed volcanoclastic apron. The nearshore flank of the Mariana Trench's Challenger Deep, the oceans' greatest abyss, also shows only rocks of island-arc affinity throughout: both andesitic and basaltic andesite volcanics, cherts, serpentinites, diabase dikes, and at lower middepths sparse plutonic granodiorites and gabbros. Again there is no evidence of accretion; apparently subduction erosion and nonaccretion prevail.

Very clear geological field evidence for a crudely layered trench-flank–island-arc foundation, for the oceanic crust being subducted, and for possible subduction erosion and nonaccretion has come from dredging in the 9500–10,000+ m deep (31,000–33,000+ ft) segment of Tonga Trench. Findings on the deepest 5000 m (16,500 ft) of both flanks are bolstered by open-ocean drilling on the upper foreslope. Dredging of 150–600-m (500–2000-ft) traverses on the nearshore flank at depths of 5200–9700 m (17,000–32,000 ft) suggest a crude layering: (1) at 5000–7000 m (16,500–23,000 ft), silicic volcanoclastics, pumice, and a siltstone carapace intruded by diabase; (2) at 7000–8500 m (23,000–28,000 ft), altered basalts overlying gabbros; (3) at 8500–9200 m (28,000–30,000 ft), coarse-grained exceedingly fresh (that is, unaltered) harzburgite and dunite; (4) locally, vesicular intermediate volcanic rocks (seamount fragment?); and (5) at 9200–10,000 m (30,000–33,000 ft), fresh to altered oceanic tholeiitic basalt (pillow fragments) from the Pacific plate crust. Two hauls on faulted offshore slopes (graben) at 8800–9000 m (29,000–30,000 ft) yielded oceanic basalt and minor diabase. An interpretive section based on these dredged rocks, seismic reflection and multibeam surveys, and seismic refraction data offshore shows oceanic crust passing beneath the peridotitic bases of the Tonga arc massif (**Fig. 2**). The precipitous lowermost flanks on both sides may represent a graben in the downflexed oceanic crust; its narrow bottom is flattened by turbidites and siliceous ooze. *See* PETROLOGY; PLUTON.

Evolution: evidence from drilling. Drilling into the sedimented forearcs (upper slopes seaward of the island chains) of several Pacific trenches elucidated characteristics of volcanism in the initiation–early subduction phases of island-arc–trench systems. A key example is the Izu-Bonin-Mariana forearc; a drilled hole in the central Tonga arc near Eua yields similar results but differs in detail. In both, the earliest volcanism, in mid to late Eocene time, constructed a wide platform of depleted (tholeiitic) arc extrusives that erupted in an extensional environment. This volumetrically great extrusion was built upon, or replaced, preexisting oceanic crust, some evidence of which remains in the Tonga forearc. Locally the forearc was subjected to rapid subsidence and subduction erosion, as evidenced by normal faulting and extension in an overall convergent system. Old lavas are found nearer the trench. With establishment of a mature subduction system (after 10,000,000+ years), normal arc volcanics (andesites and dacites) were erupted from a series of points; these evolving magmas incorporated sedimentary debris and volatile components obtained from the subducted oceanic slab. Such is the volcanism that builds the island chains. Serpentinite diapirs, produced by hydration of igneous rocks, occur at moderate depths in the Izu-Bonin forearc.Volcanic dikes of mid to late Tertiary age from island-arc magma sources bake sediments in both drilled forearcs; other sediments are hydrothermally mineralized by percolating fluids. This model, in which subduction-assimilation-magma evolution and subduction erosion are major elements, may

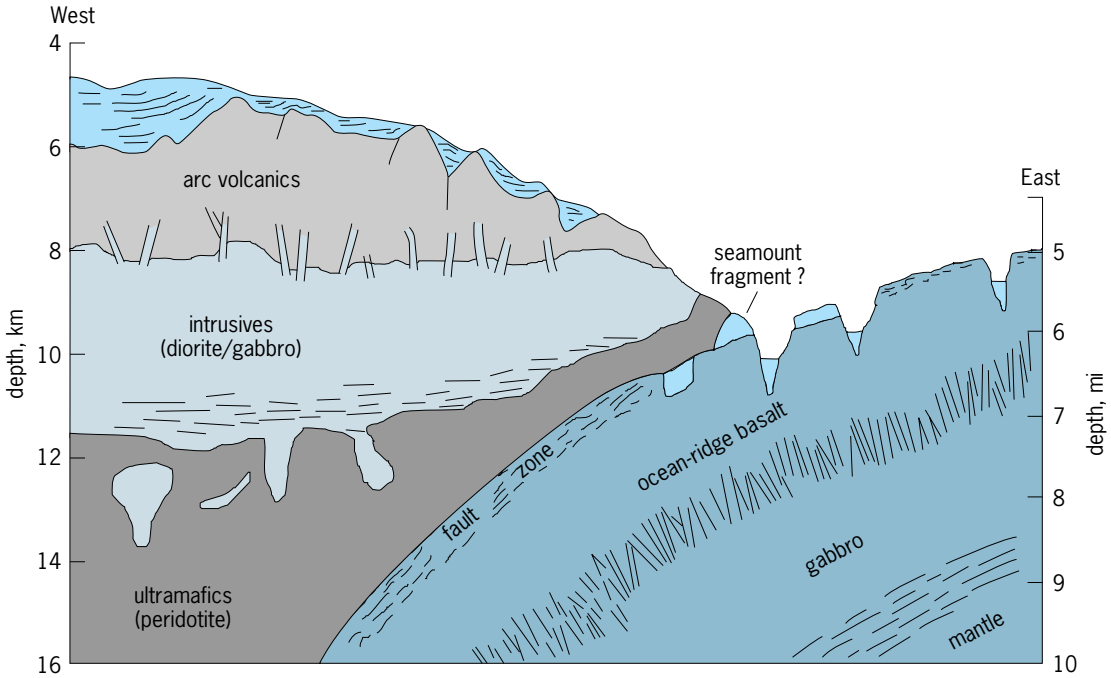

Fig. 2. Interpretive section across Tonga Trench near 20°S, based on dredged rock distribution, seismic refraction and reflection surveys, and seabeam data; vertical exaggeration 4×. (*After S. H. Bloomer and R. L. Fisher, Petrology and geochemistry of igneous rocks from the Tonga Trench: A non-accreting plate boundary, J. Geol., 95:469–495, 1987*)

characterize the sediment-starved, steady-state megatrenches of the western Pacific. Drilling and multichannel reflection profiles in regions of abundant sediment supply, for example, off Barbados or Washington State, have revealed the complicated interactions of sediment-volcanic mélanges and fluid flow in accretionary prisms, "transient" trench locales where borders advance and islands have risen at the expense of the ocean. *See* HYDROTHERMAL VENT; PLATE TECTONICS; VOLCANOLOGY.

Robert L. Fisher

Bibliography. H. Benioff, The fault origin of oceanic deeps, *Bull. Geol. Soc. Amer.*, 60:1837–1866, 1949; S. H. Bloomer and R. L. Fisher, Petrology and geochemistry of igneous rocks from the Tonga Trench: A non-accreting plate boundary, *J. Geol.*, 95:469–495, 1987; M. H. Hill (ed.) *The Sea: Ideas and Obervations,* vol. 3, 1963; R. Von Huene and D. S. Scholl, Observations at convergent margins concerning sediment subduction, subduction erosion, and the growth of continental crust, *Rev. Geophys.*, 29:279–316, 1991.

Deer

The name for a significant family (Cervidae) of 41 species of antler-bearing, ruminant, even-toed ungulates (Order Artiodactyla). These mammals vary in size from small and delicate to large and massive, and range from the dimensions of a rabbit to those of a large horse. Members of this family include the musk deer, Chinese water deer, barking deer, fallow deer, axis deer, sambar deer, sika, red deer, roe deer, American wapiti (elk), white-tailed deer (**Fig. 1**), mule deer, black-tailed deer, marsh deer, moose, and caribou (reindeer). The smallest living deer species is the Northern Pudu (*Pudu mephistopheles*) of Peru, Ecuador, and Colombia with a shoulder height of 32–36 cm (13–14 in.). The largest living species is the Alaskan moose (*Alces alces*) which stands 140–235 cm (4.6–7.8 ft) at the shoulder, weighs nearly a ton, and has palmate antlers that can span 2 m (6.5 ft). In former times, deer were even larger: the now extinct great Irish elk (*Megaceros giganteus*) had antlers that stretched 3.3 m (11 ft) or more from tip to tip.

Distribution and habitat. Members of the family Cervidae are spread throughout the Northern Hemisphere, including Asia, Europe, and North America. In the Southern Hemisphere, however, their range is limited almost entirely to South America. With the exception of certain kinds of red deer that are found in the Barbary Coast area, this family is absent from Africa, and no native species exist in Madagascar or Australia (see **table**). Deer may inhabit lowland swamps, grassy plains, and sparsely covered brush country as well as dense woodland. They are found in mountainous terrain and even in the tundra. They may feed on grasses, herbs, leaves, buds, shoots, fruit, twigs, bark, fungi, mosses, and lichens. Most species are more or less solitary, but can occur in small groups or herds. Most species are not territorial. Some undertake regular, annual migrations.

Morphology. Deer have short tails, large ears, and a muzzle that ranges from being naked to almost entirely covered with hair. The pelage, which varies

Fig. 1. Male white-tailed deer (*Odocoileus virginianus*).

from being short and smooth to dense and long, ranges in color from light gray to brown to black. Skin glands, such as preorbital, interdigital, tarsal, and frontal, are common. The feet have four toes: the third and fourth digits are well-developed, whereas the second and fifth are small and do not touch the ground. The dental formula is I 0/3, C 1/1, PM 3/3, M 3/3 × 2 for a total of 34 teeth. Deer have no central incisors in their upper jaw, their place being taken by a pad of hard gum tissue. Enlarged canines are present in musk deer (*Moshus moschiferus*), Chinese water deer (*Hydropotes inermis*), muntjac deer (*Muntiacus* sp.), and tufted deer (*Elaphodus cephalophus*). The body length in deer ranges 80–310 cm (2.6–10.3 ft) with the tail length ranging 1–50 cm (0.4–20 in.) and a weight ranging from 7 to about 800 kg (15 to about 1760 lb).

Antlers. Sexual dimorphism is evident in most species of deer with the most characteristic feature being the presence of antlers on the males of most species. Antlers are one of the most remarkable features in the entire animal kingdom. Generally, they are found only on the male. Exceptions include the musk deer (**Fig. 2**) and Chinese water deer, which lack antlers in both sexes, and the caribou in which both males and females possess antlers. Antlers are solid bony growths that develop from permanent bases on the frontal bones of the skull. The deer sheds them annually, usually in midwinter or early spring; it is controlled by male sex hormones. Several weeks after the old antlers are dropped, a round fur-like ball (burr) begins to rise from each base. These small growths rapidly swell and expand into the curving and branching structures that will soon grace the head. They can grow up to 20 mm (0.8 in.) a day. During this stage, the antlers are soft and spongy and are covered with a highly vascular skin with a coat of fine hair (velvet). Growth continues until the supply of blood and nourishment slows down. Finally, it stops altogether as the blood vessels shrink and cease to function. The antlers now harden to bonelike consistency, while the overlying "velvet" dries up and peels off. The animals burnish and polish their antlers by rubbing them against tree trunks and branches—a process probably prompted by the itching sensation caused by the drying up of the living tisssues. This periodic loss and replacement of antlers every year is amazing, with even the enormous antlers of the moose, which may weigh up to 27 kg (60 lb), falling off and reappearing every year. While antlers vary quite a bit in size and shape, those of each species of deer have their typical pattern. They are used as formidable weapons against other kinds of animals and as weapons against rival males in battles for possession of females.

Rumination. Deer are ruminants. When a deer browses, fresh food is partially masticated and only coarsely broken down before being swallowed. It passes into the first (rumen) and second (reticulum) portions of the four-compartment stomach. Here the food is partially predigested by bacteria. Later, the

Representative species of Eurasian and South American deer and their geographical distributions

Species	Distribution
Musk deer *Moschus moschiferus*	Asia
Chinese water deer *Hydropotes inermis*	Eastern Asia
Montjac, or Barking deer *Muntiacus* sp. *Megamuntiacus vuquangensis*	Southern Asia
Tufted deer *Elaphodus cephalophus*	Southern Asia
Axis deer *Axis* sp.	Southern Asia, Phillipines
Fallow deer *Dama dama*	Southern Europe, Southeast Asia
Pere David's deer *Elaphurus davidianus*	China
Sambar *Cervus unicolor*	Southern Asia, Borneo, Sumatra
Barasingha *Cervus duvauceli*	Southern Asia
Red deer *Cervus elaphus*	Europe
Roe deer *Capreolus capreolus*	Europe, Asia
Pampas deer *Ozotoceros bezoarticus*	South America
Marsh deer *Blastocerus dichotomus*	South America
Guemal *Hippocamelus* sp.	South America
Pudu *Pudu* sp.	South America
Brocket deer *Mazama* sp.	South and Central America

Fig. 2. Adult male Siberian musk deer (*Moschus moschiferus moschiferus*). (© *Copyright © by Brent Huffman, 2001*)

animal can contract the rumen and reticulum and return some of the food—now called the cud—to its mouth where it is thoroughly chewed and swallowed again. This time the cud bypasses the rumen and reticulum and passes through the other two compartments (omasum and abomasum) before entering the small intestine.

Reproduction. Gestation ranges from 180 days in the muntjac to 262 days in the wapiti and usually results in the birth of one to two young annually. Muntjacs can mate twice in the same year. Weaning occurs between 2 and 12 months with sexual maturity being reached between 6 months (muntjacs) and 3 years (red deer, wapiti) of age. Life spans range from about 10 to 30 years.

North American Deer. Members of the family Cervidae inhabiting North America include the white-tailed deer (*Odocoileus virginianus*; Fig. 1), mule or black-tailed deer (*O. hemionus*), wapiti or elk (*Cervus elaphus*), moose (*Alces alces*), and caribou or reindeer (*Rangifer tarandus*). The white-tailed deer is one of the most adaptable animals in the world. Their range extends from the southern tip of Hudson Bay in Canada well into South America. They live in areas where climatic factors vary from humid, tropical jungle to dry, hot desert to northern subarctic conditions. This species has probably extended its range northward as a result of forestry and agricultural practices. White-tails are very tolerant of people and their practices. A small subspecies of white-tailed deer, the Key deer (*O. v. clavium*), occurs in the Florida Keys and is classified as endangered. The mule deer inhabits most of temperate North America between the Pacific coast and the 100th meridian. Mule deer may be distinguished from white-tailed deer by a number of characteristics, including the rather long black-tipped tail, the long ears, the shape and position of metatarsal glands, the forked antlers having two tines of equal size and branching equally above the base to form four major tines, and their behavior. Antlers of adult white-tailed deer consist of a series of tines arising from a continuous, forecurving main beam. The elk or wapiti is a group of subspecies of the red deer complex that is distributed across North America. The entire species is distributed across Eurasia as well as North America. Elk once occurred in virtually all of temperate North America, excluding the Great Basin in Nevada and the southeastern United States. Merriam's elk and the eastern elk are now extinct, and the prairie and decidous hardwood portions of the original range are no longer occupied, except for small, isolated introduced populations. Moose, the largest living members of the deer family, are found from the Rocky Mountain region to Maine and north through Canada and Alaska. The long legs and short neck make it impossible for a moose to reach short grass unless it kneels; in order to drink, it must wade into the water. The names caribou and reindeer refer to the same species: in the wild state these animals are known as caribou, whereas the domesticated form is known as reindeer. With the exception of the musk ox, caribou live farther north than any other hoofed animal. They range beyond the timber line into the Arctic Circle in both the New World and the Old World. *See* ARTIODACTYLA; CARIBOU; MOOSE.

Donald W. Linzey

Bibliography. G. A. Feldhamer, B. C. Thompson, and J. A. Chapman (eds.), *Wild Mammals of North America: Biology, Management, and Conservation*, 2d ed., Johns Hopkins University Press, 2003; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford Ltd., 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Definite composition, law of

The law that a given chemical compound always contains the same elements in the same fixed proportions by weight. Thus, whatever its source, silver chloride always contains 108 g (3.81 oz) of silver to every 35.45 g (1.251 oz) of chlorine. If a compound is formed by the union of m atoms of one element, each weighing a, with n atoms of another element, each weighing b, the composition by weight of one molecule of the compound is in the ratio $ma:nb$. This must be the composition of any mass of the

compound, provided that all atoms of the same kind have the same weight. It is now known that this is not usually the case but that the atoms of an element may consist of a number of isotopes, having different masses. However, as long as any sample of the element always contains the same relative proportions of the isotopes, the law still holds. *See* ATOMIC MASS; ISOTOPE.

This is not the case for lead. Lead is the final product of the decay of three radioactive series, the atomic weights from the three series being 206 from radium, 208 from thorium, and 207 from actinium. Hence both the atomic weight of lead and the proportion of lead in its compounds will vary with the source of the lead. *See* RADIOACTIVITY.

Much more widespread and serious departures from the law of definite composition occur in a large variety of solid compounds (the nonstoichiometric compounds). Nonstoichiometry arises for a number of differing reasons. In the silicate minerals, such as olivine, it occurs as a result of isomorphous replacement. Thus olivine is $(Mg,Fe)_2SiO_4$ and the proportion of magnesium to iron may vary widely from sample to sample. Other solids are simply deficient in metal atoms. Thus ferrous sulfide, FeS, rarely has an atomic ratio Fe/S of precisely unity. Density measurements have shown that the lattice of the sulfur atoms remains intact but that some of the iron atoms are missing. *See* CRYSTAL STRUCTURE; NONSTOICHIOMETRIC COMPOUNDS; SILICATE MINERALS; STOICHIOMETRY.

Thomas C. Waddington

Defoliant and desiccant

Defoliants are chemicals that cause leaves to drop from plants; defoliation facilitates harvesting. Desiccants are chemicals that kill leaves of plants; the leaves may either drop off or remain attached; in the harvesting process the leaves are usually shattered and blown away from the harvested material. Defoliants are desirable for use on cotton plants because dry leaves are difficult to remove from the cotton fibers. Desiccants are used on many seed crops to hasten harvest; the leaves are cleaned from the seed in harvesting.

True defoliation results from the formation of an abscission layer at the base of the petiole of the leaf. Most of the chemicals bring about this type of defoliation, and the leaves abscise and drop from the plant. Certain chemicals kill plant leaves at low application rates, resulting in desiccation but not abscission. *See* ABSCISSION.

The most common agency of defoliation in nature is frost; frosted leaves of cotton dry up and fall off after a few days. With the rapid increase in mechanical harvesting of crops, new chemicals and new processes have been introduced; these are termed harvest aids. Most harvest-aid chemicals are growth regulators that function at low concentrations or at low application rates; in some way they stimulate or inhibit growth. Some serve to hold fruits such as apples and pears upon the trees, thus preventing preharvest drop. Others cause loosening of fruits, facilitating harvest by shaking; examples are sour cherries and grapes.

Because of the advantages of harvest aids, much effort has gone into the search for new chemicals.

A list of active ingredients in some harvest aids includes (trade names are given in parentheses) ametryn (Evik), amino triazole, ammonia gas, ammonium nitrate, ammonium thiocyanate, arsenic acid, cacodylic acid, calcium cyanamide, diethyl dithio-bisthionoformate, diquat, endothal (Accelerate, Des-I-Cate), ethephon (Ethrel), 4,6-dinitro-*o*-*sec*-butylphenos, hexachloroacetone, magnesium chlorate, nitrophen, oxyfluorfen, paraquat, pentachlorophenol, petroleum solvents, sodium borate, sodium cacodylate (Bollseye), sodium chlorate (Tumbleleaf), sodium *cis*-3-chloroacrylate, sodium naphthalene acetate (Fruitone N), tributylphosphorothioate (Folex), tributylphosphorotrithioate (Def-6), 2,4,5-TP (Silvex), triethanolamine salt of Silvex (Fruitone T).

Crops upon which harvest aids are used include alfalfa, apples, blackberries, blueberries, clovers (Alsike, Ladino), cantaloupes, castor beans, cherries, cranberries, flax, figs, filbert nuts, grapes, guar, hops, lemons, onions, oranges, peas, potatoes, rice, safflower, sudan grass, sorghum, soybeans, sunflowers, tangerines, tomatoes, walnuts, and wheat. Tomato plants are sometimes cut below the soil surface and allowed to dry out; this brings the fruits to a uniform ripeness for mechanical harvesting.

Modern agricultural technology, which has greatly increased production in the United States, is dependent upon harvest-aid chemicals. These aids bring fruits to uniform ripeness for harvest by machinery or by well-managed crews. They reduce or eliminate much labor, save time, and make for reliable management of harvest. They, along with fertilizers and improved crop varieties, are responsible for the great abundance of foods in developed countries.

It was in an effort to find such chemical agents that 2,4-D was discovered in 1942 by E. J. Kraus. For decades 2,4-D and 2,4,5-T have been important compounds used for weed control.

Defoliants and desiccants have also been used during war to destroy crops. *See* HERBICIDE.

Alden S. Crafts

Degeneracy (quantum mechanics)

A term referring to the fact that two or more stationary states of the same quantum-mechanical system may have the same energy even though their wave functions are not the same. In this case the common energy level of the stationary states is degenerate. The statistical weight of the level is proportional to the order of degeneracy, that is, to the number of states with the same energy; this number

is predicted from Schrödinger's equation. The energy levels of isolated systems (that is, systems with no external fields present) comprising an odd number of so-called fermions (for example, electrons, protons, and neutrons) always are at least twofold degenerate. *See* EXCLUSION PRINCIPLE.

Relation to symmetries. Except for so-called accidental degeneracy, degeneracy is associated with special symmetries of the physical system, and can be removed by destroying this symmetry. For example, an energy level of total angular momentum J (in units of $(\hbar/2\pi$, where $\hbar$ is Planck's constant) has a $(2J + 1)$-fold degeneracy that results from the circumstance that the $2J + 1$ allowed values of J_z, the z component of total angular momentum, all have the same energy. In a magnetic field H along z, this level splits into $2J + 1$ energy levels, since the energy now depends on J_z, that is, on the angle between H and the total angular momentum. *See* ANGULAR MOMENTUM.

Relation to perturbation theory. In applications of quantum mechanics to complicated systems, such as many-electron atoms (in atomic physics applications) or heavy atomic nuclei (in nuclear physics applications), the ability to associate degenerate or nearly degenerate levels with underlying symmetries is crucial to successful calculations of the needed stationary-state energies and wave functions via perturbation theory, which often is the most convenient calculational method for such systems. In stationary-state perturbation theory the full hamiltonian H is written in the form $H = H_0 + V$, where H_0 is regarded as the unperturbed hamiltonian and V is the perturbing interaction, which, presumably, is comparatively small. Typically, however, the actual "smallness'' of the perturbation—and therefore the validity of the perturbation theory for computing the energy and wave function of a stationary state i of H starting from the corresponding unperturbed energy (also indexed by i) and wave function—is measured by ratios of the form $V_{ij}/(E_i - E_j)$, where j is another stationary state of the unperturbed hamiltonian, V_{ij} is the matrix element of V between the unperturbed states i and j, and E_i and E_j denote respectively the unperturbed energies of the states i and j. Consequently, stationary-state perturbation theory breaks down when there are states for which $E_i - E_j = 0$, that is, when the energy shift from the unperturbed energy is sought for a state i belonging to a group of degenerate unperturbed states, unless the original set of unperturbed states are chosen so that every matrix element $V_{ij} = 0$ whenever $E_i - E_j = 0$. In the magnetic field illustration discussed above, for instance, in order to calculate the level splitting in heavy atoms (where the total orbital angular momentum and the total spin generally no longer are good quantum numbers) by perturbation theory, it is necessary to choose the $2J + 1$ unperturbed originally degenerate states so that J_z is a good (and different) quantum number in each of those unperturbed states; in other words, each of those $2J + 1$ originally degenerate states must be an eigenstate belonging to a different value of J_z in the allowed range $-J \leq J_z \leq J$.

For a state i of the full hamiltonian corresponding to a nondegenerate state of the unperturbed hamiltonian, choosing a set of unperturbed states for which every $V_{ij} = 0$ is not a prerequisite for use of stationary perturbation theory. The above-stated measure of the validity of stationary-state perturbation theory suggests, however, and careful analysis confirms, that the energy shift of a perturbed state i from its unperturbed value (whether degenerate or nondegenerate) generally is much more strongly influenced by its "interaction" V_{ij} with any state j whose unperturbed energy E_j lies close to E_i than by interactions with states whose unperturbed energies are quite different from E_i unless V_{ij} is zero or exceptionally small. Similarly, unless V_{ij} happens to be zero or exceptionally small, the wave function of the perturbed state i (whether degenerate or nondegenerate) will be much more strongly "mixed'' with the unperturbed wave function of any state j having a nearby unperturbed energy E_j than with the wave functions of states whose unperturbed energies are quite different from E_i.

In quantum mechanics and in other branches of mathematical physics, the term degeneracy is employed also to characterize the eigenvalues of operators other than the energy operator. *See* EIGENVALUE (QUANTUM MECHANICS); NONRELATIVISTIC QUANTUM THEORY; PERTURBATION (QUANTUM MECHANICS); SYMMETRY LAWS (PHYSICS). Edward Gerjuoy

Bibliography. S. Gasiorowicz, *Quantum Physics*, 2d ed., 1995; L. Pauling and E. B. Wilson, *Introduction to Quantum Mechanics*, 1935, reprint 1985.

Degree-day

A unit used in estimating energy requirements for building heating and, to a lesser extent, for building cooling. It is applied to all fuels, district heating, and electric heating. Origin of the degree-day was based on studies of residential gas heating systems. These studies indicated that there existed a straight-line relation between gas used and the extent to which the daily mean outside temperature fell below 65°F (18°C).

The number of degree-days to be recorded on any given day is obtained by averaging the daily maximum and minimum outside temperatures to obtain the daily mean temperature. This procedure was found to be adequate and less time consuming than the averaging of 24 hourly temperature readings. The daily mean so obtained is subtracted from 65°F (18°C) and tabulated. Monthly and seasonal totals of degree-days are available from local weather bureaus. They use the 30-year period 1931–1960 as their basis.

The base temperature of 65°F (18°C) was established by observation of the outside temperatures at which the majority of homeowners started their gas heating systems in the early fall. Interpreted in another way, the internal heat gains in residences,

which appear to be about 5000 Btu/h (1500 W), are sufficient to provide heat down to a balance temperature of 65°F (18°C).

Provided that the efficiency of the heating equipment remains constant, the fuel usage of a given house will be proportional to the number of degree-days. For purposes of comparing one heating season with another, constant efficiency is a reasonable assumption. However, efficiency of the heating equipment decreases markedly in the warmer portions of the heating season. Thus the fuel use per degree-day will be higher in spring and fall than in winter. The same considerations apply to use of degree-day data in warmer climates compared to cooler regions.

A frequent use of degree-days for a specific building is to determine before fuel storage tanks run dry when fuel oil deliveries should be made.

Number of Btu which the heating plant must furnish to a building in a given period of time is Btu required = heat rate of building × 24 × degree-days where "Btu required" is the heat supplied by the heating system to maintain the desired inside temperature. "Heat rate of building" is the hourly building heat loss divided by the difference between inside and outside design temperatures.

To determine fuel requirement from this relationship, two decisions are needed. First, the correct seasonal efficiency of the heating system must be obtained or estimated; second, appropriate base temperature for the degree-day data must be calculated. Seasonal efficiency is generally less than the test efficiency of the heating unit although the load efficiency curves are surprisingly level down to quite low loads. Base temperature for degree-day values may be chosen from consideration of internal heat gains, as in a well-lighted office building, then degree-day data on other than a 65°F (18°C) basis are required. Balance temperatures for both day operation and night operation may also need to be taken into consideration. Then it would be proper to replace the term degree-days in the equation with the term adjusted degree-days. *See* AIR CONDITIONING; CLIMATOLOGY; COMFORT HEATING; COMFORT TEMPERATURES; PSYCHROMETRICS.

C. George Segeler

Degree of freedom (mechanics)

Any one of the number of independent ways in which the space configuration of a mechanical system may change. A material particle confined to a line in space can be displaced only along the line, and therefore has one degree of freedom. A particle confined to a surface can be displaced in two perpendicular directions and accordingly has two degrees of freedom. A particle free in physical space has three degrees of freedom corresponding to three possible perpendicular displacements. A system composed of two free particles has six degrees of freedom, and one composed of N free particles has $3N$ degrees. If a system of two particles is subject to a requirement that the particles remain a constant distance apart, the number of degrees of freedom becomes five. Any requirement which diminishes by one the degrees of freedom of a system is called a holonomic constraint. Each such constraint is expressible by an equation of condition which relates the system's coordinates to a constant, and may also involve the time. When applied to systems of particles, a holonomic constraint frequently has the geometrical significance of confining a particle to a specified surface, which may be time-dependent. *See* CONSTRAINT.

A mechanical system may be conceived as a set of N particles in space subject to a certain number K of holonomic constraints which reduce the number of degrees of freedom to $3N - K$. In the case of a rigid structure consisting of more than two noncollinear particles, the number of degrees of freedom is independent of the number of particles and is equal to six.

Russell A. Fisher

Bibliography. A. P. Arya, *Introduction to Classical Mechanics*, 2d ed., 1997; H. Goldstein, C. P. Poole and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; R. Matzner and L. Shepley, *Classical Mechanics*, 1991.

De Haas-van Alphen effect

An oscillatory behavior of the magnetic moment of a pure metal crystal with changes in the applied magnetic field B, at very low temperatures. It is named after its discoverers W. J. de Haas and P. M. van Alphen. This effect has its origin in the Bohr-Sommerfeld quantization of the orbits of conduction electrons under the influence of the magnetic field.

Origin. A free electron in a magnetic field B moves in a circular orbit with an angular frequency given by the cyclotron resonance frequency, Eq. (1), where e

$$\omega_c = \frac{eB}{m} \tag{1}$$

and m are the charge and mass of the electron. But the energy E of this simple harmonic motion must be quantized according to the Planck prescription, Eq. (2), where $\hbar$ is Planck's constant divided by 2π,

$$E_n = (n + {}^1/_2)\hbar\omega_c \tag{2}$$

and n numbers the quantized energy levels.

In a metal at low temperatures at zero magnetic field, electrons are not at rest, but have a continuous distribution of energies up to some limiting energy E_F called the Fermi energy; this arises because the Pauli exclusion principle keeps the electrons from all falling into the lowest energy state. Those electrons with energy close to E_F dominate the behavior of the metal. In a magnetic field the allowed electron energies are not continuous, but must correspond to one of those given by Eq. (2). Suppose for some value of n, B_n an electron energy level matches

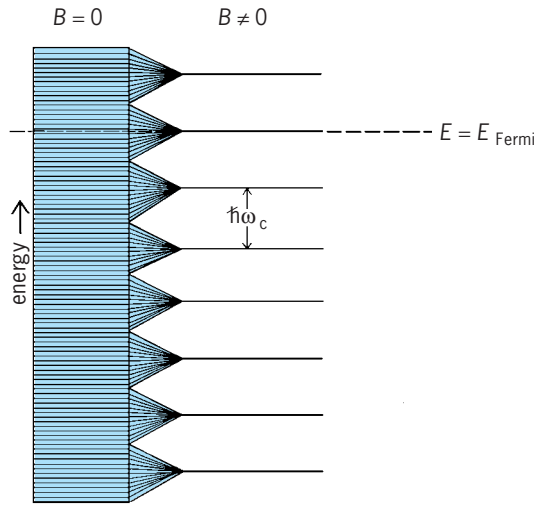

Distribution of electron energy in a metal at zero and finite magnetic field *B*, showing the condition that an electron energy level matches the Fermi energy.

the Fermi energy E_F, that is, $E_n = E_F$ (see **illus.**); from Eqs. (1) and (2), this corresponds to Eq. (3).

$$E_F = \frac{(n + {}^1/_2)\hbar e}{m} B_n \tag{3}$$

See CYCLOTRON RESONANCE EXPERIMENTS; EXCLUSION PRINCIPLE; FREE-ELECTRON THEORY OF METALS; NONRELATIVISTIC QUANTUM THEORY.

In order for n to increase by 1, the magnetic field must change so that Eq. (4) is satisfied. Eliminating n between Eqs. (3) and (4) yields Eq. (5) for the amount

$$E_F = \frac{(n + {}^3/_2)\hbar e}{m} B_{n+1} \tag{4}$$

$$\Delta\left(\frac{1}{B}\right) \equiv \frac{1}{B_{n+1}} - \frac{1}{B_n} = \frac{e\hbar}{mE_F} \tag{5}$$

by which the magnetic field must change in order for n to increase by 1. The quantity $\Delta(1/B)$ determines the "period" of the de Haas–van Alphen oscillations (it is a period in magnetic field, not time). Since $E = p^2/2m$, where p is the electron momentum, Eq. (6)

$$\Delta\left(\frac{1}{B}\right) = \frac{2e\hbar}{p_F^2} \tag{6}$$

is valid, where p_F is the limiting momentum corresponding to the energy E_F. Thus, a measurement of the period of the oscillations permits a determination of the limiting or Fermi momentum of the electrons. A detailed quantum statistical-mechanical treatment shows that many of the properties of a metal, including the magnetization, oscillate with field according to Eq. (6). *See* MAGNETIZATION.

Measurement of effective mass. The assumption that the electrons occupy only states with E_n equal to or less than E_F involves the implicit assumption that the metal is cooled to absolute zero. At any finite temperature T, states with E_n greater than E_F will also be occupied. The ratio of the number of electrons in two neighboring energy levels E_n and E_{n+1} is governed by the Boltzmann occupation factor exp $[(E_{n+1} - E_n)/k_BT]$, where k_B is Boltzmann's constant. If T is large enough such that k_BT is much greater than $\hbar\omega_c$, this factor becomes unity, and in this case there is nothing particularly significant about the condition $E_F = E_n$; that is, the amplitude of the oscillations is expected to be greatly attenuated. On the other hand, at very low temperatures where $\hbar\omega_c$ is greater than k_BT, strong oscillations are expected. Thus, a measurement of the temperature dependence of the oscillation amplitude permits a determination of the cyclotron frequency, or equivalently the electron mass. In a metal, as in a semiconductor, electrons behave with an effective mass m^* rather than the free electron mass, and the de Haas–van Alphen effect thus allows a measurement of m^* in metals.

Determination of Fermi surface. When electrons move in a metal, the energy-momentum relation in general differs radically from that of a free particle. The momenta $\vec{p}_F$ which correspond to the Fermi energy depend on the direction $\hat{p}_F$ within the crystal; that is, the shape of the Fermi surface differs from a sphere.

Application of the Bohr-Sommerfeld quantization rules shows that Eq. (6) for the de Haas–van Alphen period must be generalized to Eq. (7), where A_p is

$$\Delta\left(\frac{1}{B}\right) = \frac{2\pi\hbar e}{A_p} \tag{7}$$

the area of the orbit in momentum space, that is, the cross section of the Fermi surface.

For a given Fermi surface, the cross-sectional area is a function of the momentum parallel to the field and ranges between various extremal values. In principle, all of these sections make an oscillatory contribution. However, for all sections other than extremal sections, the oscillatory contributions at neighboring sections tend to cancel. Thus in Eq. (7) A_p is the area of an extremal section (or sections for metals whose Fermi surface supports multiple extremal cross sections). By studying the dependence of A_p on the direction B of the magnetic field, sufficient information can be obtained to construct the detailed shape of the Fermi surfaces of a metal. A further condition for observation, in addition to low temperatures, is sample purity: the electron-scattering time τ must be longer than, or at least comparable to, a cyclotron period $2\pi/\omega_c$. *See* RELAXATION TIME OF ELECTRONS.

Virtually all metals available in sufficient purity have been studied by the de Haas–van Alphen effect. It is also the most powerful probe of Fermi surface properties in alloys and intermetallic compounds.

The intermetallic compounds of the greatest interest are the so-called heavy-fermion materials involving actinide and rare-earth atoms. Effective masses greater than 1000 have been seen in electronic

heat-capacity measurements. In one of these materials, UPt_3, a mass of 200 electron masses was observed, the largest value ever measured by using the de Hass–van Alphen effect.

Prescriptions for deconvolution of extremal areas to Fermi momenta are well developed. Furthermore, measurements of the angular-dependent effective mass m^* can be used to determine the velocity of electrons at the Fermi surface. *See* BAND THEORY OF SOLIDS; FERMI SURFACE. J. B. Ketterson

Bibliography. A. Isihara, *Condensed Matter Physics*, 1991; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996.

Dehumidifier

Equipment designed to reduce the amount of water vapor in the atmosphere. *See* HUMIDITY.

The atmosphere is a mechanical mixture of dry air and water vapor, the amount of water vapor being limited by air temperature. Water vapor is measured in either grains per pound of dry air or pounds per pound of dry air (7000 gr = 1 lb).

There are three methods by which water vapor may be removed: sorbent materials; cooling to the required dew point; and compression with aftercooling.

Sorbent type. Sorbents are materials which are hygroscopic to water vapor; they are available in both solid and liquid forms. Solid sorbents include silica gels, activated alumina, and aluminum bauxite. Liquid sorbents include halogen salts such as lithium chloride, lithium bromide, and calcium chloride, and organic liquids such as ethylene, diethylene, and triethylene glycols and glycol derivatives.

Solid sorbents may be used in static or dynamic dehumidifiers. Bags of solid sorbent materials within packages of machine tools, electronic equipment, and other valuable materials subject to moisture damage constitute static dehumidifiers. An indicator chemical may be included to show by a change in color when the sorbent is saturated. The sorbent then requires reactivation by heating at 300–350°F (149–177°C) for 1–2 h before reuse.

A dynamic dehumidifier for solid sorbent consists of a main circulating fan, one or more beds of sorbent material, reactivation air fan, heater, mechanism to change from dehumidifying to reactivation, and aftercooler.

A single-bed dehumidifier (**Fig. 1**) operates on an intermittent cycle of dehumidifying for 2–3 h and then switches to the reactivation cycle for 15–45 min. No dehumidification is obtained during the reactivation cycle. A single-bed unit is used for small areas where moisture is a problem. The moist reactivation air is discharged to the outside.

The dual-bed machine is larger in capacity than the single-bed unit and has the advantage of providing a continuous supply of dehumidified air (**Fig. 2**). While one bed is dehumidifying, the other bed is reactivating. After a preset time interval, the air cycle is switched to pass the air through the reactivated bed for dehumidification and to reactivate the saturated bed.

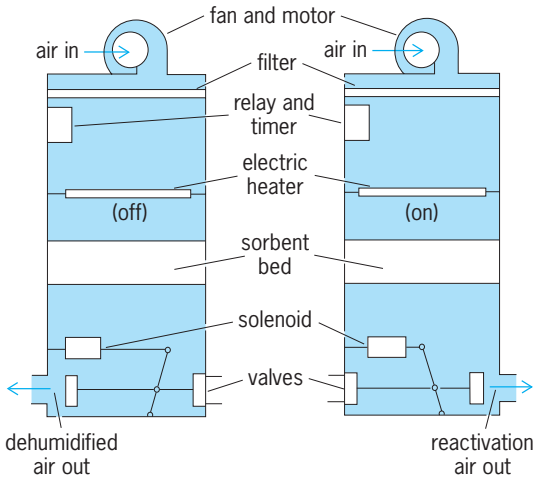

Fig. 1. Single-bed solid-sorbent dehumidifier. Dehumidifying cycle on left and reactivation cycle on right.

The dew point of the effluent air of a fixed-bed machine is lowest at the start of a cycle immediately after the reactivated bed has been placed in service. The dew point gradually rises as the bed absorbs the water vapor and eventually would be the same as the entering dew point when the vapor pressure of the sorbent reached the vapor pressure of the air and could no longer absorb moisture from the air. The cutoff point at which the absorbing bed is changed over to reactivation is fixed by the maximum allowable effluent dew point. *See* DEW POINT.

A multibed unit with short operating cycles will reduce the range of effluent dew point to within a few degrees. A multibed unit with rotating cylindrical bed maintains a reasonably constant effluent dew point.

The liquid-sorbent dehumidifier consists of a main circulating fan, sorbent-air contactor, sorbent pump,

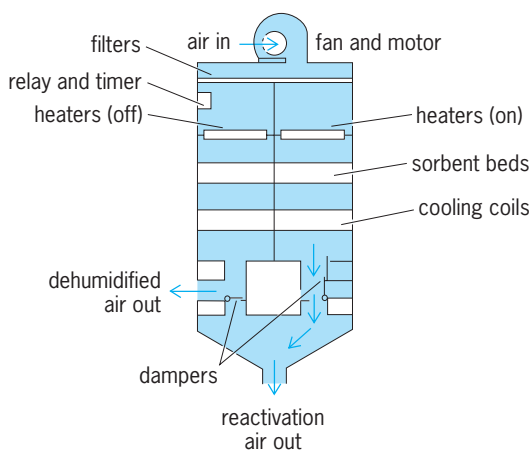

Fig. 2. Dual-bed solid-sorbent dehumidifier. Air is being dehumidified through the left bed at the same time that the right bed is being reactivated.

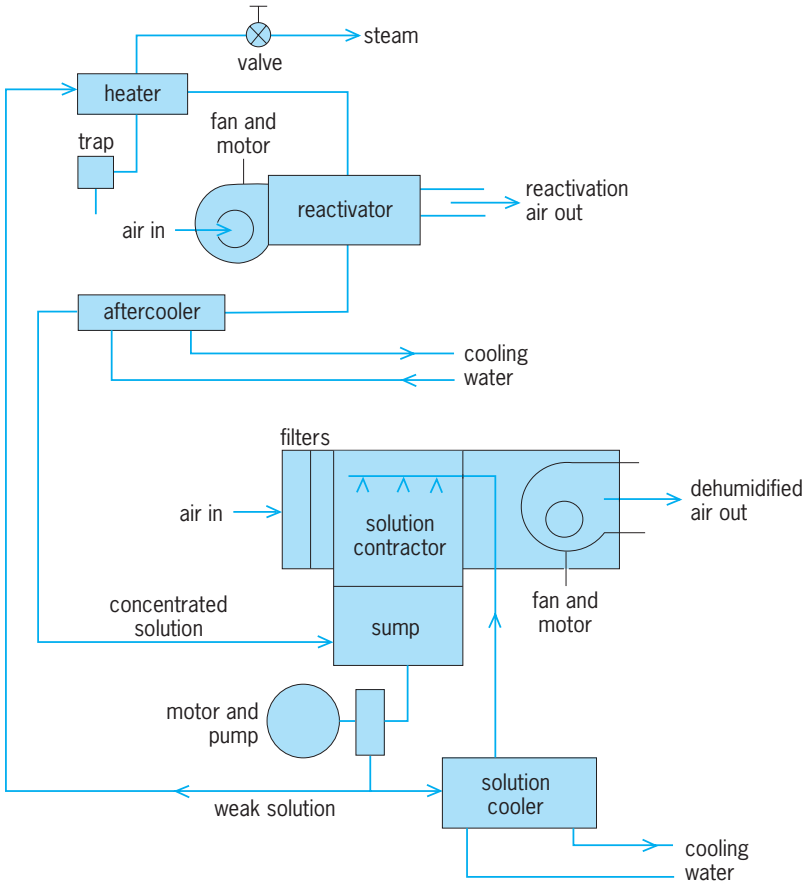

Fig. 3. Liquid-sorbent dehumidifier with continuous dehumidifying and reactivation.

and reactivator including contactor, fan, heater, and cooler (**Fig. 3**). This unit will control the effluent dew point at a constant level because dehumidification and reactivation are continuous operations with a small part of the sorbent constantly bled off from the main circulating system and reactivated to the concentration required for the desired effluent dew point.

Cooling type. A system employing the use of cooling for dehumidifying consists of a circulating fan and cooling coil. The cooling coil may use cold water obtained from wells or a refrigeration plant, or may be a direct-expansion refrigeration coil. In place of a coil, a spray washer may be used in which the air passes through two or more banks of sprays of cold water or brine, depending upon the dew-point temperature required.

When coils are used, the leaving dew point is seldom below 35°F (1.7°C) because of possible buildup of ice on the coil. When it is necessary to use coils for temperatures below 35°F (1.7°C), as in cold-storage rooms, either two coils are used so one can be defrosted while the other is in operation, or only one coil is used and dehumidifying is stopped during the defrost period.

A brine-spray dehumidifier or brine-sprayed coil can produce dew-point temperatures below 35°F (1.7°C) without frosting if properly operated and maintained. *See* AIR COOLING.

Compression type. Dehumidifying by compression and aftercooling is used when the reduction of water vapor in a compressed-air system is required. This is particularly important, for example, if the air is used for automatic control instruments or cleaning of delicate machined parts.

If air is compressed and the heat of compression removed to bring the temperature of the air back to the temperature entering the compressor, condensation will take place and the remaining water vapor content will be directly proportional to the absolute pressure ratio of the compressed air (**Fig. 4**).

For example, if saturated air at 70°F or 39°C (111 gr/lb of dry air) is compressed from atmospheric pressure (14.7 lb/in.^{2} absolute) to 88 lb/in.^{2} absolute (6:1 compression ratio) and cooled to 70°F (39°C), the remaining water vapor in the compressed air will be 111/6 = 18.5 gr/lb of dry air. If the air is expanded back to atmospheric pressure and 70°F (39°C), the dew point will be 24°F (13°C). *See* PSYCHROMETRICS.

The power required for compression systems is so high compared to power requirements for dehumidifying by either the sorbent or refrigeration method that the compression system is not an economical one if dehumidifying is the only end result required. *See* COMPRESSOR; DRYING. John Everetts, Jr.

Bibliography. American Society of Heating, Refrigerating, and Air-Conditioning Engineers, *Equipment Handbook,* 1992; ASHRAE, *Heating, Ventilating, Air-Conditioning Guide,* annually.

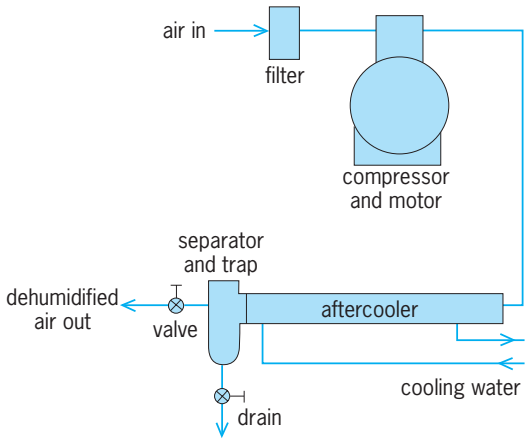

Fig. 4. Dehumidifying by compression and aftercooling.

Dehydrogenation

A reaction in which hydrogen is detached from a molecule. The reaction is strongly endothermic; therefore, heat must be supplied to maintain the reaction temperature. Since carbon-carbon bonds are easier to break than carbon-hydrogen bonds, highly selective reaction conditions are required to cause the molecule to dehydrogenate rather than break apart into small molecules (for example, methane) and carbon (coke). Superheated steam (725°C or 1337°F) is

often mixed with the dehydrogenation reactor feed to provide the needed heat of reaction, lower the partial pressure and enhance the equilibrium conversion, and decoke the catalyst to extend the reactor onstream time. In addition, excess hydrogen is sometimes added to the feed to depress coke formation and diminish the breakup of large molecules (for example, styrene) into smaller molecules of much less economic value.

Industrial applications. Dehydrogenation reactions are of major industrial importance. They account for the production of tremendous quantities of organic compounds that enter into the manufacture of textiles, gasoline, lubricants, explosives, plastics, plasticizers, agricultural chemicals, perfumes, detergents, elastomers, and pharmaceutical products.

Most commercial dehydrogenation reactions are catalytic and take place in the vapor phase. However, the noncatalytic, thermal dehydrogenation of isobutane is practiced to produce propylene and isobutylene. Also, alcohols with high boiling points can be dehydrogenated in the liquid phase in the presence of powdered catalysts (for example, active Raney nickel). The primary catalysts used for dehydrogenation reactions are the transition metals: Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, Pd, W, and Pt. *See* CATALYSIS; HETEROGENEOUS CATALYSIS.

Reactions. The primary types of dehydrogenation reactions are:

1. Vapor-phase conversion of primary alcohols to aldehydes, as in the preparation of acetaldehyde from ethyl alcohol in the presence of a silver catalyst at 550°C (1020°F), reaction (1).

$$CH_3CH_2OH + {}^1/_2O_2 \longrightarrow CH_3CHO + H_2O \qquad (1)$$

2. Vapor-phase conversion of secondary alcohols to ketones, as in the preparation of acetone from isopropyl alcohol, reaction (2), and methyl ethyl ketone from secondary alcohol in the presence of copper catalyst at 300°C (570°F), reaction (3).

$$CH_3CHOHCH_3 \longrightarrow CH_3COCH_3 + H_2 \qquad (2)$$

$$CH_3CHOHCH_2CH_3 \longrightarrow CH_3COCH_2CH_3 + H_2 \qquad (3)$$

3. Dehydrogenation of paraffins to monoolefins, as in the Houdry Catofin process to manufacture propylene from propane, in the presence of chromia-alumina catalyst at 550–600°C (1020–1110°F), reaction (4).

$$CH_3CH_2CH_3 \longrightarrow CH_2{=}CHCH_3 + H_2 \qquad (4)$$

4. Conversion of monoolefins to diolefins, practiced in the conversion of 1-butene to 1,3-butadiene, over chromia-alumina or noble-metal catalysts at 550–600°C (1020–1110°F), reaction (5).

$$CH_2{=}CHCH_2CH_3 \longrightarrow CH_2{=}CHCH{=}CH_2 + H_2 \qquad (5)$$

5. Dehydrogenation of a side chain, commercialized in the Monsanto/Lummus and Cosden/Badger processes for the manufacture of styrene from ethylbenzene in the presence of iron oxide catalyst at 600°C (1110°F), reaction (6).

$$C_6H_5C_2H_5 \longrightarrow C_6H_5CH{=}CH_2 + H_2 \qquad (6)$$

6. Manufacture of aromatics by the dehydrogenation of cycloparaffins, as in the catalytic reforming of naphthas and naphthenes in the presence of a platinum catalyst at 500°C (930°F) and 100–1000 lb/in.^{2} (700–7000 kilopascals) for the production of toluene and high-octane gasoline.

Oxidative dehydrogenation. When the detached hydrogen from the dehydrogenation reaction is immediately oxidized, two benefits accrue: (1) the conversion of reactants to products is increased because the equilibrium concentration is shifted toward the products (law of mass action), and (2) the added exothermic oxidation reaction supplies the needed heat of reaction. An example is the oxidative dehydrogenation of 1-butene to 1,3-butadiene practiced in the PetroTex Oxo-D process, reaction (7).

$$CH_2{=}CHCH_2CH_3 + {}^1/_2O_2 \longrightarrow CH_2{=}CHCH{=}CH_2 + H_2O \qquad (7)$$

An innovative concept for the oxidative dehydrogenation of paraffins, which has only been demonstrated experimentally, is the use of molten halide catalysts on porous supports. This type of catalyst uses a binary eutectic mixture of alkali metal halides (for example, KCl-LiCl) with added transition-metal promoter (for example, Mn) on a low-surface-area support (for example, 15 m^2/g alumina or 4500 ft^2/oz). Under reaction conditions, the catalytic material exists as a thin, molten film which covers the inner surface area of the support. (This is similar to the sulfuric acid catalyst, which exists as a vanadium pyrosulfate melt lining the porous pellets during reaction.)

Oxidative dehydrogenation reaction systems suffer one major drawback: the added oxygen will react with the hydrocarbon as well as the hydrogen, generating carbon dioxide and other undesired oxides. Unless the oxygen-hydrocarbon reactions can be depressed, an oxidative dehydrogenation system will not be economical.

Dual-function catalysts. Dual-function catalysts have been developed which promote both dehydrogenation and a subsequent reaction, to reduce the total reactor volume and processing steps required to manufacture the desired end product from the raw material. Generally, the catalytic agent (for example, Pt) for dehydrogenation is bound to a support which has acid sites (for example, halogen-activated alumina). One application is the dehydrogenation and isomerization of naphthenes in catalytic reforming. Another example is dehydrocyclodimerization, the conversion of small (C_3–C_5) paraffins or olefins to aromatics. Xylenes, ethylbenzene, toluene, and benzene are produced from butanes by this process.

Biochemical dehydrogenation. Dehydrogenation reactions are essential steps in the metabolism and

biosynthesis of living organisms. The catalysts for these reactions are dehydrogenases, which catalyze the removal of hydrogen atoms from fuel molecules and their transfer to coenzymes. These coenzymes, in turn, carry the energy-rich electrons from catabolic reactions to electron-requiring biosynthesis reactions. *See* CHEMICAL CONVERSION; ENZYME; HYDROGENATION; METABOLISM; OXIDATION PROCESS. J. W. Fulton

Bibliography. R. T. Morrison and R. N. Boyd, *Organic Chemistry*, 7th ed., 2000; D. L. Nelson and M. M. Cox, *Lehninger's Principles of Biochemistry*, 3d ed., 2000.

Delay line

A transmission line (as nearly dissipationless as possible) or an electric network approximation of it which, if terminated in its characteristic impedance, will reproduce at its output a waveform applied to its input terminals with little distortion but at a time delayed by an amount dependent upon the electrical length of the line.

If a transmission line is dissipationless, which will be the case if its series resistance is zero and its shunt conductance is also zero, it will have a characteristic impedance Z_0, as shown in Eq. (1), where L is the

$$Z_0 = \sqrt{L/C} \tag{1}$$

series inductance and C the shunt capacitance per unit length of the line. *See* TRANSMISSION LINES.

The velocity v of propagation of a signal along the line is shown by Eq. (2).

$$v = \frac{1}{\sqrt{LC}} \tag{2}$$

Therefore the time required for the pulse to propagate a distance x along the line is shown by Eq. (3). Such a line, terminated in its characteristic

$$T_d = x\sqrt{LC} \tag{3}$$

impedance, is shown in **Fig. 1**; and output pulses reproduce input at a delayed time T_d.

The lumped-circuit approximation of the transmission line is shown in **Fig. 2**. If the inductance and capacitance per section are L_1 and C_1, then the total time delay is shown by Eq. (4), where n is the number of sections.

$$T_d = n\sqrt{L_1C_1} \tag{4}$$

If the delay line is not terminated in its characteristic impedance, there is multiple reflection back and forth along the line. For example, if the receiving end of the line is unterminated, as in **Fig. 3**, but the sending end is terminated in its characteristic impedance, the receiving-end reflection coefficient is positive, and a delayed pulse appears at the output, as given by Eqs. (3) and (4). In addition, a pulse

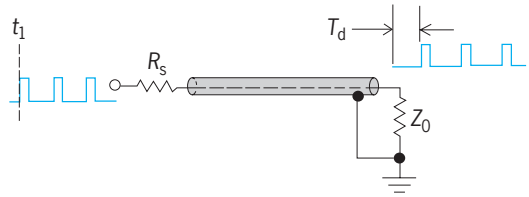

Fig. 1. Transmission line as delay line.

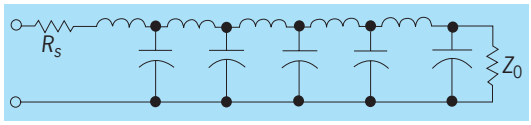

Fig. 2. Lumped-circuit delay line.

of the same polarity appears at the input terminals, delayed by twice this amount (the time required for a pulse to travel to the end and back).

If the receiving end is terminated in a short circuit as shown in **Fig. 4**, the receiving-end reflection coefficient is negative, and no pulse appears at the output, although an inverted pulse will appear at the input terminals with the same time delay as before.

Various applications make use of the short circuit and open circuit of delay lines, including line-controlled pulse generators. *See* PULSE GENERATOR.

Delay lines are also used for establishing a time sequence for the occurrence of events. A delay line with a total length equal to the greatest time delay required in a system may be used as a basic element. Pulses occurring at intermediate times may be obtained from taps at various points along the line. A specific application is found in the synchronizing signal generator of the television system. Also, the lumped-circuit delay line is an essential element of the wide-band distributed amplifier.

When a signal is digital in nature, or consists of a series of pulses, the series of pulses may be delayed by using a shift register, which might, for

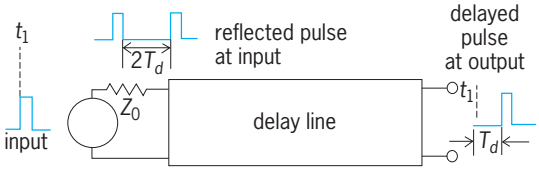

Fig. 3. Reflection due to unterminated receiving end.

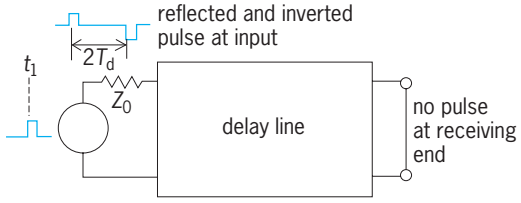

Fig. 4. Reflection due to short-circuited receiving end.

example, consist of a chain of cascaded type D flip-flops. If the register has n stages, the pulse series will appear at the output delayed by a time $(n - 1)T$, where T is the interval between consecutive pulses of the system timing clock. The same function can be realized by using switched-capacitor circuits or an array of charge-coupled devices. *See* CHARGE-COUPLED DEVICES; SWITCHED CAPACITOR; TRANSMISSION LINES.

Glenn M. Glasford

Bibliography. T. F. Bogart, Jr., *Linear Electronics*, 1993; C. H. Chen (ed.), *Digital Waveform Processing and Recognition*, 2d ed., 1995; C. G. Fonstad, *Microelectronic Devices and Circuits*, 1994; G. M. Glasford, *Digital Electronic Circuits*, 1988.

Delayed neutron

A neutron emitted spontaneously from a nucleus as a consequence of excitation remaining from a preceding radioactive decay event. Analogously, delayed emission of protons and alpha particles is observed, but the known delayed-neutron emitters are more numerous, and some of them have practical implications. In particular, they are important in the control of nuclear chain reactors. *See* CHAIN REACTION (PHYSICS); NUCLEAR REACTOR.

The delayed neutrons were discovered in 1939, just after the discovery of fission. The first measure of their abundance and the first resolution of their decay periods were made in 1942 and were instrumental in the decision to attempt the first nuclear chain reaction in the midst of a large city. The first chemical identifications of the delayed-neutron emitters were made in 1942 and published in 1945.

Nuclear origin. Radioactive transformation by beta decay often leaves a product nucleus with internal energy in excess of that associated with its stable state, and usually the energy is radiated as gamma-ray quanta. In exceptional cases, the energy may exceed that required to remove one or more of the constituent neutrons from the product nucleus; neutron emission then rapidly takes place and at least partially deexcites the nucleus. The delayed neutrons are accordingly a radioactive emission, occurring essentially at the time of decay of the beta-emitting precursor. If the precursor is formed, for example, by fission, the delayed neutrons may be emitted seconds, or even minutes, after the moment of fission (because of the intervening beta decay), and are to be distinguished from the prompt neutrons that are emitted at the time of fission. On the other hand, the delayed neutrons are emitted so quickly after the completion of the beta decay that the half-lives and the chemical behavior of the delayed-neutron radioactivities are, in practice, those of the precursors.

Occurrence. Delayed-neutron emission occurs in neutron-rich nuclei, that is, in the heavier isotopes of various elements. These isotopes have several neutrons in excess of the proton-neutron mixture associated with the element's stable isotopes; in consequence, their beta-decay half-lives are short. In relation to the shell model of the nucleus, they are likely to be associated also with nuclides having one or a few neutrons in excess of a closed neutron shell, because such neutrons are expected to be loosely bound. *See* NUCLEAR STRUCTURE.

When a group of nuclei of a given precursor undergoes its radioactive decay, only a fraction of the product nuclei will emit neutrons because many of the decays may leave the product nuclei in an insufficiently high state of excitation. Hence, there is a neutron yield probability associated with each decay. After neutron emission, the product nucleus may be stable [for example, precursor ^{137}I beta-decays to ^{137}Xe (excited), which emits a neutron going to ^{136}Xe (stable)], but more commonly the product nucleus will undergo further beta decay, and may itself be a precursor for a subsequent delayed-neutron emission.

Table 1 lists known precursors of delayed-neutron emission, together with the respective half-lives $T_{1/2}$ in seconds and the neutron yield probability in percent. Those in the lighter elements have been made by using particle accelerators, and of these, double, as well as single, delayed-neutron emission has been observed following the beta decay of ^{11}Li, ^{30}Na, ^{31}Na, and ^{32}Na. The main groups are in the fission products, but the longest-lived precursor (^{210}Tl) happens to be in natural radioactivity.

The mean energies of the delayed-fission neutrons are lower than those of the prompt neutrons, being usually in the range of a few hundred keV. The actual delayed-neutron spectra usually have several energy groups, depending upon the nuclear energy levels involved in the beta decay of the precursor.

Role in reactor control. In a ^{235}U nuclear reactor, about 0.7% of the fission neutrons are delayed, the others being prompt. In conventional moderated reactors, the prompt neutrons are born, slowed down, and reabsorbed to produce the next fission in a cycling time of 10^{-3} to 10^{-4} s. (In a fast-neutron reactor the time is much shorter.) Consequently, if the reactor were to become overcritical (more neutrons generated per millisecond than are absorbed or leak out), the chain reaction would exponentiate or "run away," and the reactor might quickly overheat and cause a dangerous accident unless the control rods could respond within a few milliseconds to correct the situation. The fortunate presence of the delayed neutrons eases the situation, because so long as the reactor operates within the margin of 0.7% (delayed critical), the control rods can take as long as several seconds to respond; thus the chain reaction comes within the range of easy and leisurely manual or automatic control.

Within the delayed-neutron margin, advantage can be taken of the reactor's behavior to derive a fundamental characteristic of its operation, namely, its reactivity, which expresses the fractional excess over unity of the chain reaction's effective multiplication

TABLE 1. Delayed-neutron precursors

Nuclide	Half-life ($T_{1/2}$), s	Fraction of beta decays that lead to neutron emission, %
Accelerator-produced light elements		
^{8}He	0.12	12
^{9}Li	0.18	35
^{11}Li	0.008	61
^{12}Be	0.011	—
^{13}B	0.017	0.28
^{16}C	0.75	99
^{17}N	4.2	95
^{19}N	0.42	—
^{27}Na	0.30	0.08
^{28}Na	0.031	0.58
^{29}Na	0.043	15
^{30}Na	0.054	33
^{31}Na	0.017	31
^{32}Na	0.014	15
Fission products, light group		
^{79}Ga	2.63	0.098
^{80}Ga	1.66	0.84
^{81}Ga	1.23	12
^{82}Ga	0.60	21.4
^{83}Ga	0.31	43
^{83}Ge	1.9	—
^{84}Ge	1.2	—
^{84}As	5.6	0.61
^{85}As	2.03	53
^{86}As	0.9	12
^{87}As	0.73	44
^{87}Se	5.6	0.19
^{88}Se	1.52	0.6
^{89}Se	0.41	5.0
^{91}Se	0.27	21
^{87}Br	55.6	2.48
^{88}Br	16.0	6.55
^{89}Br	4.4	13
^{90}Br	1.9	22.5
^{91}Br	0.54	16
^{92}Br	0.36	22
^{92}Kr	1.85	0.033
^{93}Kr	1.29	2.0
^{94}Kr	0.21	5.7
^{92}Rb	4.5	0.011
^{93}Rb	5.85	1.3
^{94}Rb	2.76	9.9
^{95}Rb	0.38	8.4
^{96}Rb	0.22	13.3
^{97}Rb	0.17	25.1
^{98}Rb	0.116	15
^{99}Rb	0.066	14
Fission products, light group (cont.)		
^{100}Rb	0.05	—
^{97}Sr	0.43	0.3
^{98}Sr	0.80	0.36
^{97m}Y	1.13	0.06
^{98}Y	2.1	3.4
^{99}Y	1.4	1.2
Fission products, heavy group		
^{123}Ag	0.39	—
^{127m}In	1.12	<0.04
^{127}In	3.76	0.68
^{128}In	0.84	0.06
^{129m}In	0.59	0.25
^{129}In	1.26	2.5
^{130}In	0.58	1.40
^{131}In	0.28	1.72
^{132}In	0.16	4
^{133}Sn	1.47	—
^{134}Sn	1.04	17
^{134}Sb	10.4	0.11
^{135}Sb	1.71	15.5
^{136}Sb	0.82	23
^{136}Tl	17.5	0.9
^{137}Tl	2.8	2.5
^{138}Tl	1.4	6.3
^{137}I	24.5	6.5
^{138}I	6.5	5.3
^{139}I	2.31	9.3
^{140}I	0.60	8.8
^{141}I	0.44	26
^{141}Xe	1.73	0.04
^{142}Xe	1.24	0.42
^{143}Xe	0.30	—
^{141}Cs	24.9	0.037
^{142}Cs	1.71	0.093
^{143}Cs	1.78	1.61
^{144}Cs	1.00	3.1
^{145}Cs	0.58	13.6
^{146}Cs	0.335	13.3
^{147}Cs	0.21	25.4
^{147}Ba	0.71	5.2
^{148}Ba	0.47	23.9
^{147}La	4.1	0.5
Natural radioactivity		
^{210}Tl (Ra C″)	78	0.007

constant, k_{eff}; that is, reactivity $= (k_{\mathrm{eff}} - 1)/k_{\mathrm{eff}}$. The reactivity is used by the operators to follow the response of the reactor to physical changes, as in fuel loading, the insertion or extraction of neutron-absorbing samples, or a change in temperature.

If after a period of steady operation (reactivity = 0) a control rod is inserted farther, the reactor's response will consist of two parts: a quick reduction in power associated with the reduction of the population of neutrons undergoing diffusion, with their millisecond times of residence; and a slow response that eventually becomes the half-life of the longest-lived precursor, ^{87}Br. If, on the other hand, the control rod is withdrawn to a new position, the reactor's response is stronger because a positive exponential is involved. There is a brief, transient jump in power due to an increase in the prompt neutron population, but in the delayed-neutron component there is a balance between the decay of existing precursors made at the original neutron flux level and the growth of new precursors made at the new (higher) neutron flux level. This has a controlling effect because the population of delayed-neutron precursors greatly exceeds the population of diffusing neutrons with their millisecond lifetimes. After the transient

TABLE 2. Classification of delayed-neutron emitters into six groups

Group	1	2	3	4	5	6
$(T_{1/2})_i$, s	55	22	6	2	0.6	0.2
λ_i, s^{-1}	0.0126	0.0316	0.115	0.346	1.15	3.46
β_i, yield per 10^4 fissions	5.2	34.6	31.0	62.4	18.2	6.6

passes, the reactor exhibits an exponential rate of increase in power, and the time required for the power to increase by a factor of e (that is, 2.718) is called the stable period, T. This period may be as long as an hour or more, because the balance between precursors in formation and precursors decaying is delicate. The period in turn is related to the reactivity through the inhour equation (inhour = inverse hour, or hour^{-1}), which in one form is expressed by the equation below. Here l is the life time

$$\text{Reactivity} = \frac{l/T + \sum_i \left[\beta_i/(1+\lambda_i T)\right]}{1 - \sum_i \left[\beta_i/(1+\lambda_i T)\right]}$$

of diffusing neutrons, and l/T represents the transient effect. The summation represents the delayed-neutron effect, λ_i being the decay constant of the ith precursor [$\lambda_i = 0.693/(T_{1/2})_i$] and β its fractional intensity. (Physically, β_i would include both the fractional decay yields listed in Table 1 and the fission yield of the associated mass chain.) If l and T are expressed in hours, and the λ_i in hours^{-1}, the reactivity will be stated in inhours.

Evaluation of the summation is complex if all of the precursors are listed individually, and, in practice, a grouping has been used. The original resolution of the overall delayed-neutron decay curve (1942) yielded four groups, decaying exponentially with half-lives ranging between 55 and 2.5 s, with associated relative intensities. Subsequent work, particularly upon the short-lived emitters, has expanded this to six groups. As an example, six groups that account for the delayed-neutron decay curve in the case of slow-neutron fission of ^{235}U are given in **Table 2**. In this case the total β is 158 delayed neutrons per 10^4 fissions.

To match this grouping, the summation in the inhour formula is taken over six terms, each with its associated λ_i and β_i.

As related to Table 1, group 1 represents almost entirely ^{87}Br, group 2 is mostly a combination of ^{88}Br and ^{137}I, and the other groups each contain several or many precursors having half-lives close to that associated with their group.

As a further simplification (and a rougher approximation), the six groups can be compressed into one. The total β is then used, and the mean λ is an average of the λ_i, each weighted proportionally to $(T_{1/2})_i$. On the other hand, it is also possible to make the summation over the 80 or more individual precursors, and when this is done, reasonable agreement is found with the six groups of the experimentally measured total delayed-neutron decay curve. *See* REACTOR PHYSICS.

Other uses. Delayed neutrons have been used for the detection and assay of fissile materials in several ways. In prospecting for uranium, for example, if a neutron source and a neutron detector are lowered into a borehole, the source can be activated for an interval and then turned off or removed; the detection of delayed neutrons returning to the counter would then show the presence of fissile ores. Alternatively, if an ore sample is exposed to the neutron flux of a reactor and then moved to a neutron detector, its assay of fissile material can be quickly determined; a sensitivity of 1000 delayed-neutron counts per microgram of uranium content has been reported. The content of nuclear fuel in fuel rods can be nondestructively measured by using the delayed neutrons. They are used for localization of cladding failures in reactor fuel elements; when the cladding fails, fission products appear in the cooling stream, and a neutron detector placed downstream, outside the reactor shield, can signal the presence of a cladding failure with much greater certainty than can a gamma-ray detector. (In fact, triangulation methods have been devised whereby the channel that contains the failure can be localized.)

Finally, the delayed neutrons are of some interest in cosmogony, in connection with the r-process. This is a rapid buildup of the elements and isotopes by successive neutron capture and beta decay in an event such as a supernova explosion, in which an immense density of neutrons is temporarily generated. As the nuclides decay toward stability after the event, the presence of delayed-neutron activities can modify the final distribution of relative abundances, and in fact, the inclusion of this refinement into the r-process calculations seems to improve the agreement with the isotope abundances found in nature. *See* NUCLEAR FISSION; NUCLEOSYNTHESIS; RADIOACTIVITY.

Arthur H. Snell

Bibliography. D. R. Alexander and Y. K. Peng, Calculated time-dependent delayed neutron group yields for thermal-neuron fission of U-235, *Nucl. Sci. Eng.*, 70:184–191, 1979; S. Glasstone and A. Sesonke, *Nuclear Reactor Engineering*, 4th ed., 1993; P. L. Reeder, Status of and outstanding problems in delayed neutron data, P_n values and energy spectra, U.S. DOE Rep. PNL-SA-8831 (Conf. 800979-13), Battelle Pacific Northwest Laboratories, NTIS, September 1980; A. H. Snell et al., The cyclotron group at the Metallurgical Laboratory, Chicago, 1940–44, *Amer. J. Phys.*, 48:971–978, 1980.

Deliquescence

The absorption of atmospheric water vapor by a crystalline solid until the crystal eventually dissolves into a saturated solution. This behavior is well known for certain salts such as hydrated calcium chloride, $CaCl_2 \cdot 6H_2O$, and zinc chloride, $ZnCl_2$, but it is a property of all soluble salts in air of sufficiently high humidity.

Themodynamically, the condition for deliquescence is that the partial pressure of the water vapor in the air exceed the vapor pressure (aqueous tension) of the water in the saturated solution of the salt. Then the reaction solid + water vapor → aqueous solution will occur spontaneously. The speed of the process depends upon the rate of diffusion of water vapor into the crystal lattice, crystal size, and other factors. The process will stop when the water vapor in the atmosphere is depleted to the point at which its partial pressure equals that of the saturated solution.

In general, substances which are highly soluble in water have a greater tendency to deliquesce, since concentrated solutions will have a lower vapor pressure. At 77°F (25°C) the vapor pressure of pure water is 23.8 mmHg (3173 pascals), whereas that of a saturated solution of $CaCl_2 \cdot 6H_2O$ is only 7.0 mmHg (933 pascals); this salt then will deliquesce in an atmosphere where the relative humidity exceeds 7.0/23.8 or 30%. For $ZnCl_2$ the situation is much more extreme; it will deliquesce at a relative humidity of 10%. Ordinary sugar (sucrose) at 77°F (25°C) will deliquesce at humidities above 85%.

Crystalline solids also may absorb water by increasing their water of hydration if the dissociation pressure of the hydrated species to be formed is less than the partial pressure of the water vapor. It is this process, not deliquescence, which is the opposite of efflorescence.

Deliquescent substances can be used to remove water vapor from air, although they have no special advantage over substances which merely add water of hydration and remain crystalline. *See* DESICCANT; DRYING; EFFLORESCENCE; SOLUTION; VAPOR PRESSURE.

Robert L. Scott

Delocalization

A phenomenon in which the most loosely held bonding electrons of some molecules serve to bind not two but several atoms. This contrasts with localization, a characteristic of ordinary single bonds, as in the normal paraffins, for example, methane (CH_4); in ethane, (C_2H_6); or in the molecules water (H_2O) and ammonia (NH_3). *See* CHEMICAL BONDING.

The prototype completely delocalized system is the ideal crystalline metal, in which the valence electrons are spaced uniformly over a periodic lattice of positive-ion cores. This confers a special stability to the metal, and it also accounts for its high electrical conductivity and other metallic properties. *See* METAL.

The aromatic and conjugated molecules of organic chemistry contain delocalized electronic systems. The benzene molecule (C_6H_6) is considered the archetypical aromatic molecule. Benzene possesses an underlying single-bonded planar framework $(C_6H_6)^{6+}$ plus six additional electrons. Available for these electrons are six $2p$ orbitals, one on every carbon atom, each perpendicular to the molecular plane. The six electrons, called pi electrons, are ascribable to the six carbon atoms in such a way that neither structure (**1**) nor structure (**2**) is an accurate

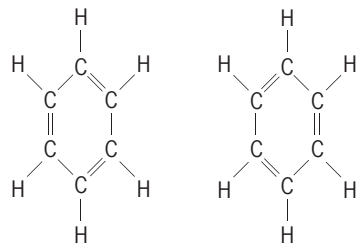

(**1**) (**2**)

representation, but instead a structure that can be described in the valence bond language as a resonance hybrid of the two, schematically written as (**3**), where the circle stands for the delocalized electrons. All molecules containing such rings possess an extra stability associated with the delocalization phenomenon. As a consequence, they also have a low propensity for chemical reactivity.

(**3**)

Delocalization is also present, though to a lesser extent, in conjugated molecules such as butadiene (C_4H_6) which contain alternating single and double bonds. Butadiene has the principal localized structure (**4**), but this structure is in resonance with a somewhat less stable structure (**5**). The resonance

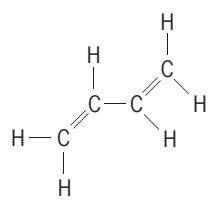

(**4**)

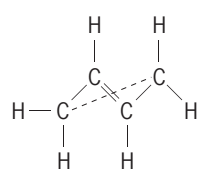

(**5**)

gives a certain amount of delocalization. *See* CONJUGATION AND HYPERCONJUGATION; RESONANCE (MOLECULAR STRUCTURE).

The greater the length of a path of electron delocalization, the more the wavelength of light that a molecule can absorb will shift to the red. Thus

butadiene is colorless, but β-carotene is colored; benzene is colorless, but triphenylmethyl dye molecules are colored. Another manifestation of the mobile nature of the pi electrons in aromatic systems is the ring-current effects they produce on proton chemical shifts in nuclear magnetic resonance spectra. *See* MAGNETIC RESONANCE.

When all the bonding in a molecule is localized, it can be stated to good accuracy that many molecular properties (for example, the energy) are sums of bond contributions, and that the properties of individual bonds (for example, polarity and length) can be transferred from one molecule to another. When the bonding is delocalized, on the other hand, simple additivity and transferability rules do not hold.

Quantititative understanding of localization and delocalization has been achieved by using calculational techniques from valence bond and molecular orbital theories of chemical bonding. In the valence bond approach, substantial resonance signifies the presence of delocalization. In the molecular orbital approach, wherever delocalization is large, the highest occupied molecular orbitals will turn out necessarily to extend over several atoms. *See* MOLECULAR ORBITAL THEORY. Robert G. Parr

Bibliography. C. A. Coulson, *Coulson's Valence*, 3d ed., 1980; M. J. S. Dewar, *The Molecular Orbital Theory of Organic Chemistry*, 1969; R. McWeeny, *Methods of Molecular Quantum Mechanics*, 2d ed., 1989, reprint 1992.

Delta

A deposit of sediment at the mouth of a river or tidal inlet. The name, from the Greek letter Δ, was first used by Herodotus (fifth century B.C.) for the triangular delta of the Nile. It is also used for storm washovers of barrier islands and for sediment accumulations at the mouths of submarine canyons. *See* FLOODPLAIN; NEARSHORE PROCESSES.

Structure and growth. The shape and internal structure of a delta depend on the nature and interaction of two forces: the sediment-carrying stream from a river, tidal inlet, or submarine canyon, and the current and wave action of the water body in which the delta is building. This interaction ranges from complete dominance of the sediment-carrying stream (still-water deltas) to complete dominance of currents and waves, resulting in redistribution of the sediment over a wide area (no deltas). This interaction has a large effect on the shape and structure of the delta body.

Most of the sediment carried into the basin is deposited when the inflowing stream decelerates. If there is little density contrast, this deceleration is sudden and most sediment is deposited near the mouth of the river. If the inflowing water is much lighter than the basin water, for example, fresh water flowing into a colder sea, the outflow spreads at the surface over a large distance away from the outlet. If the inflow is very dense, for instance, cold muddy water in a warm lake, it may form a density flow on or near the bottom, and the principal deposition may occur at great distance from the outlet. A good example is Lake Mead, where 134 ft (41 m) of sediment have been deposited against Hoover Dam at 75 mi (120 km) from the inlet of the Colorado River. However, not all long-distance transport of river sediment in the sea can be attributed to deep density flow; almost half of the total load of the Amazon River is deposited more than 1000 mi (1600 km) away near the Orinoco delta and in the southeast Caribbean due to longshore transport by currents and waves.

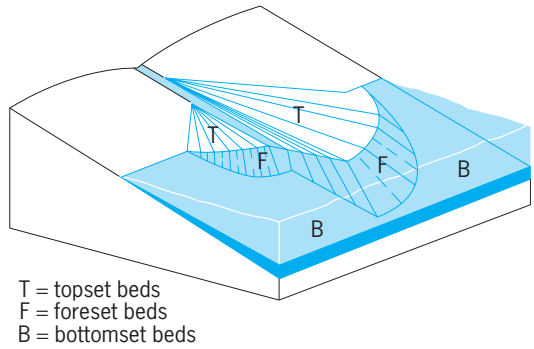

Fig. 1. Schematic diagram showing two stages of growth of a typical Gilbert-type delta. (*After P. H. Kuenen, Marine Geology, John Wiley and Sons, 1950*)

Three principal components make up the bodies of most deltas in varying proportions: topset, foreset, and bottomset beds. They were defined originally by G. K. Gilbert for lake and shallow marine deltas; there, sizable differences in the depositional slopes of the three units, which form part of the definition, are readily seen (**Fig. 1**). The concept has been redefined to apply to most deltas, and as now understood, the topset beds comprise the sediments formed on the subaerial delta: channel deposits, natural levees, floodplains, marshes, and swamp and bay sediments. The foreset beds are those formed in shallow water, mostly as a broad platform fronting the delta shore, and the bottomset beds are the deep-water deposits beyond the deltaic bulge. In marine deltas the fluviatile influence decreases and the marine influence increases from the topset to the bottomset beds. Concurrently, the mean grain size and lithological variability also decrease. The topset beds are variable in grain size and composition over short distances, ranging from coarse channel deposits through layered levee silts and clays to the fine clays of the floodplains and marshes. The foreset beds consist of a uniform blanket of laminated silts and silty clays reflecting the seasonal influence of currents and waves on the shallow platform, while the bottomset beds are homogeneous silty clays, often characterized by abundant plant fragments. The boundary between bottomset beds and neritic shelf clays is completely arbitrary. Similarly, the faunas incorporated in the sediments range from great local variability in the topset beds, with dominance of barren sediments and local pockets of rich brackish faunas, to the almost completely marine fossils uniformly distributed in the bottomset beds.

Through progressive outbuilding, the delta can become overextended with long river courses and very low gradients. Eventually, shorter and steeper paths to the sea will be developed and the existing subdelta will be abandoned in favor of a shorter course. The Mississippi delta shows good examples of such subdelta migration, and the present active delta would be abandoned for a new one off the Atchafalaya River, if artificial control did not keep the flow in check. The abandoned delta body gradually submerges and is eroded by waves and wind. The submergence results from two factors: temporary continuation of subsidence caused by loading of the deltaic sediments on the substrate, and compaction of the fine sediments. These are overcompensated during delta growth, but cause it to sink below sea level after abandonment. Winnowing of the sediments by waves produces a lag deposit of sand at the advancing edge of the sea, which ultimately results in a thin and discontinuous blanket of sand, often bordered by a string of sand bars and small islands near the previous delta shore. This sequence is called the destructive phase of delta formation; a good example is the Breton Sound-Chandeleur Island area of the Mississippi delta, which marks the site of the old St. Bernard subdelta.

In a different way, deltas can be viewed as being composed of three structural elements: (1) a framework of elongate coarse bodies (channels, river-mouth bars, levee deposits), which radiate from the apex to the distributary mouths (sand fingers); (2) a matrix of fine-grained floodplain, marsh, and bay sediments; and (3) a littoral zone, usually of beach and dune sands which result from sorting and longshore transport of river-mouth deposits by waves, currents, tides, and wind. The relative proportions of these components vary widely. The Mississippi delta consists almost entirely of framework and matrix (**Figs. 2** and **3**); its rapid seaward growth is the result of deposition of river-mouth bars and extension of levees, and the areas in between are filled later with matrix. This gives the delta its characteristic bird-foot outline. A different makeup is presented by the Rhone delta (Fig. 3), where the supply of coarse material at the distributary mouths is slow, and dispersal by wave action and longshore drift fairly efficient, so that nearly all material is evenly redistributed as a series of coastal bars and dunes across a large part of the delta front. This delta advances as a broad lobate front, while the present Mississippi delta grows at several localized and sharply defined points.

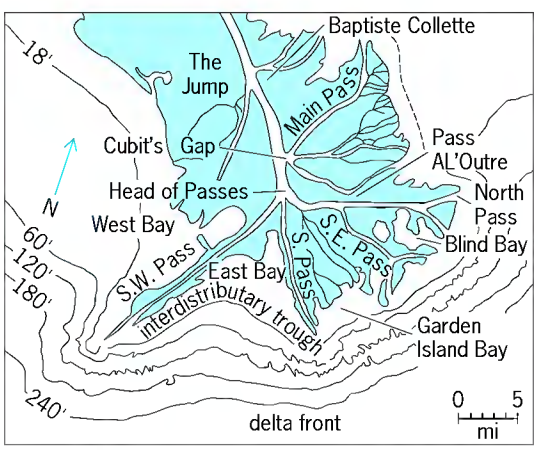

Fig. 2. Modern Mississippi bird-foot delta. 1 mi = 1.6 km. (*After H. N. Fisk et al., Sedimentary framework of the modern Mississippi delta, J. Sediment. Petrol., 24:76–99, 1954*)

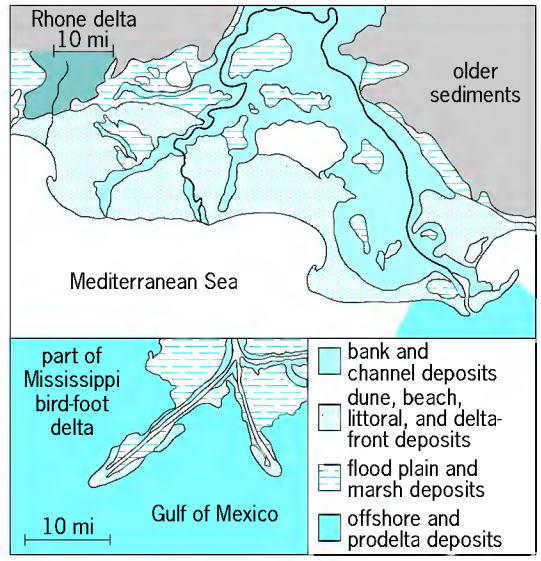

Fig. 3. Examples of deltaic sediment distribution. 1 mi = 1.6 km.

Pattern of deposition. Many different types of deltas exist, few of which have been examined in detail. It appears that most of them can be explained by the interplay, in various proportions, of two major controlling forces: the coarseness and the rate of sediment supply by the river, and the rate of reworking and dispersal along the delta front by marine currents and waves. Other factors, such as the shape of the preexisting coastline and of the sea floor, seasonal fluctuations of marine and river forces, and the tidal range, cannot as yet be evaluated, but appear to be secondary. The Mississippi delta (**Fig. 4*a***) is dominantly fine-grained and consists in large part of thick sand fingers with intervening matrix clays. It builds rapidly outward, is almost entirely controlled by river action, and has a typical bird-foot shape.

In the Rhone delta (Fig. 4*b*) the rate of sediment supply is much lower, but the sediment is coarser, and the rate of reworking and dispersal by waves and currents is higher. Consequently, this delta is more sandy, and its principal growth takes place through the accretion over a broad front of bars and beaches, resulting from winnowing of sand at the river mouth and dispersal by wind and waves. The advance of the delta is slow, but occurs over a broadly lobate front, and the bulk of the delta body consists of the deposits of the sandy strand plain. Matrix and framework deposits are restricted to the topset beds in the upper delta. Littoral marine forces control to a large extent the shape of this delta.

A broad coastal plain is also found in the Orinoco

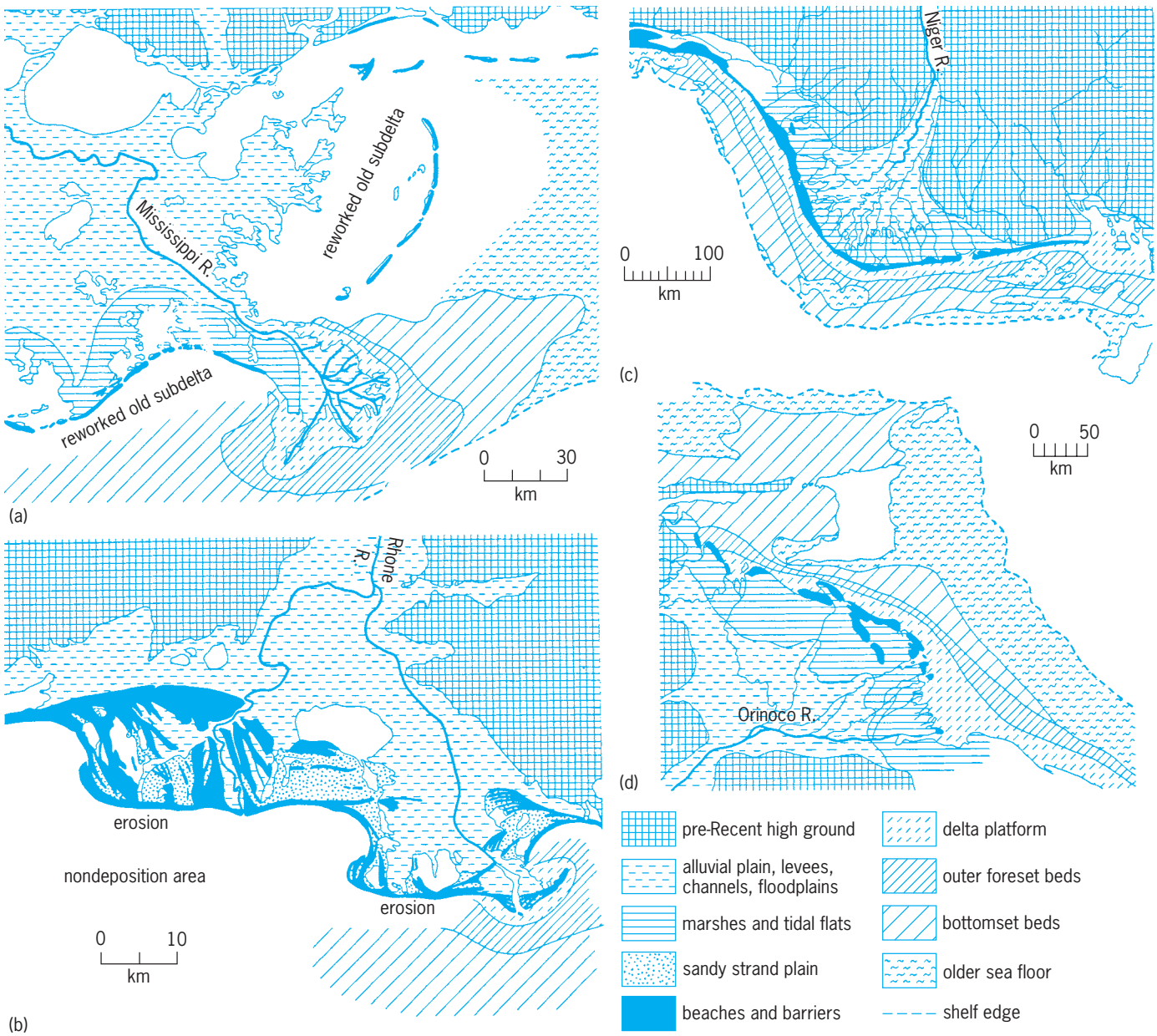

Fig. 4. Sketch maps of the sediments of the (*a*) Mississippi, (*b*) Rhone, (*c*) Niger, and (*d*) Orinoco deltas. 1 km = 0.6 mi. (*After T. H. van Andel, The Orinoco delta, J. Sediment. Petrol., 37(2)297–310, 1967*)

delta (Fig. 4*d*). The large supply of very fine sediment of this river is augmented by Amazon mud brought along the shore from the east. It builds rapidly in a sheltered sea with little wave action, but longshore drift to the northwest is marked. Consequently, the coastal plain, although also littoral marine, is not wave-built and sandy, but formed by tidal deposition on mud flats. The delta, as a result of the longshore currents, advances over its entire perimeter, and the action of the river plays an insignificant part in its development.

Delta advance over the entire perimeter also takes place in the Niger delta (Fig. 4*c*), which resembles the Orinoco delta in shape and distribution of sediments. However, enough sediment is supplied by this river to form, under the influence of moderate wave action on an exposed shelf, a marginal zone of sand bars, beaches, and barrier islands. Behind this barrier, tidal sedimentation of a muddy coastal plain takes place, and framework-matrix units are restricted to the upper delta. The delta advances over a broad front as a result of redistribution of sediments by currents. In a sense it is intermediate between the Orinoco and Rhone deltas.

Characteristics of modern deltas. While over 150 major deltas are formed today, not all rivers, or even all major ones, have deltas. This is the result of a rise in sea level following the last glacial period, which produced deep estuaries in many parts of the world which have not yet been filled (for example, the Amazon estuary). Delta thicknesses vary widely: The Nile is depositing a layer 50 ft (15 m) thick in a

Statistics pertaining to modern deltas*

River	Dimension of subaerial delta, statute mi		Amount of sediment discharged		Annual extension of subaerial delta	
	Length	Breadth	River water by weight (avg), ppm	Annual volume of sediment, mi^3	Measurement period, years	Approximate distance, ft
Mississippi's present bird-foot delta	12	30	550	0.068	1838–1947	250
Hwang-Ho	300	470†	50,000‡		1870–1937	950
Ganges-Brahmaputra	220	200	870	0.043 (Ganges only)		
Rhone into Mediterranean Sea	30	47	400–590	0.005	1737–1870	190
Danube	46	46	310	0.008		40
Nile (prior to barrages)	96	145	1600	0.001	1100–1870	45
Colorado above Hoover Dam	43	0.05–0.6	8300	0.032	1936–1948	3.6 mi (gorge)
Euphrates-Tigris	350	90			1793–1853	180

*1 mi = 1.6 km. 1 mi^3 = 4.17 km^3. 1 ft = 0.3 m.
†Includes 100 mi of nondeltaic Shantung Peninsula.
‡Maximum is 400,000 ppm.

shallow embayment, whereas the Mississippi, building out in deeper water, has constructed a delta body more than 850 ft (260 m) thick. This thick wedge exerts pressure on the underlying beds, causing downbending of the shelf and producing mudlumps in areas where the underlying sediment is soft (see **table**).

Engineering problems. Despite difficult engineering problems, many cities, such as Calcutta, Shanghai, Venice, Alexandria (Egypt), and New Orleans, were constructed on deltas. These problems include shifting and extending shipping channels; lack of firm footing for construction except on levees; steady subsidence, which may reach a rate of 5 ft (1.5 m) per century; poor drainage; and extensive flood danger. Submergence to a depth of 15 ft (4.6 m) during hurricanes or typhoons is not uncommon. Moreover, in certain deltas the tendency of the main flow to shift away to entirely different areas, with resulting disappearance of the main channels utilized for water traffic, is a constant problem that is difficult and costly to counter. *See* ESTUARINE OCEANOGRAPHY; MARINE SEDIMENTS. Tjeerd H. Van Andel

Bibliography. G. Friedman, J. E. Sanders, and D. C. Kopaska, *Principles of Sedimentary Deposits: Stratigraphy and Sedimentation*, 1992; M. R. Leeder, *Sedimentology: Process and Product*, 1982; K. Pye (ed.), *Sediment Transport and Deposition*, 1994.

Delta electrons

Energetic electrons ejected from atoms in matter by the passage of ionizing particles. In every primary ionizing collision between a charged particle and an atom, one or more electrons are ejected. Delta electrons are, by definition, that small fraction of these emitted electrons having energies which are large compared to the ionization potential. The name is a traditional one—comparable to alpha particles, for energetic helium nuclei, and beta particles, for energetic electrons emitted in radioactive decays. *See* ALPHA PARTICLES; BETA PARTICLES; CHARGED PARTICLE BEAMS; IONIZATION.

The energy W of the delta electron emitted at angle θ in the laboratory is given by Eq. (1), where v is

$$W = 2mv^2 \cos^2 \theta \tag{1}$$

the velocity of the incident charged particle, and m is the electron mass. The electrons in the forward direction are the most energetic and have an energy near $4(m/M)E$, where M and E are the mass and energy of the incident charged particle. The cross section for the ejection of a delta electron in the energy range between W and $W + dW$ given by Eq. (2), where e

$$d\sigma = \frac{2\pi e^4 z^2}{mv^2} \cdot \frac{dW}{W^2} \tag{2}$$

is the electron charge and ez is the charge of the incident particle, and the cross section for finding an electron between θ and $\theta + d\theta$ is given by Eq. (3).

$$d\sigma = \frac{2\pi e^4 z^2}{m^2 v^4} \cdot \frac{\sin\theta \, d\theta}{\cos^3 \theta} \tag{3}$$

The number of electrons emitted per centimeter of path length is obtained by multiplying these cross sections by the number of electrons per cubic centimeter in the stopping material.

Delta electrons are responsible for the "hairy" appearance of charged particle tracks when they are observed in cloud chambers or in photographic emulsions (see **illus.**). In studies of super-high-energy particles in cosmic radiation and from the highest-energy accelerators, observation of the number of delta electrons per centimeter of path length has been shown to lead to a reliable determination of the charge of the energetic particle. Biological

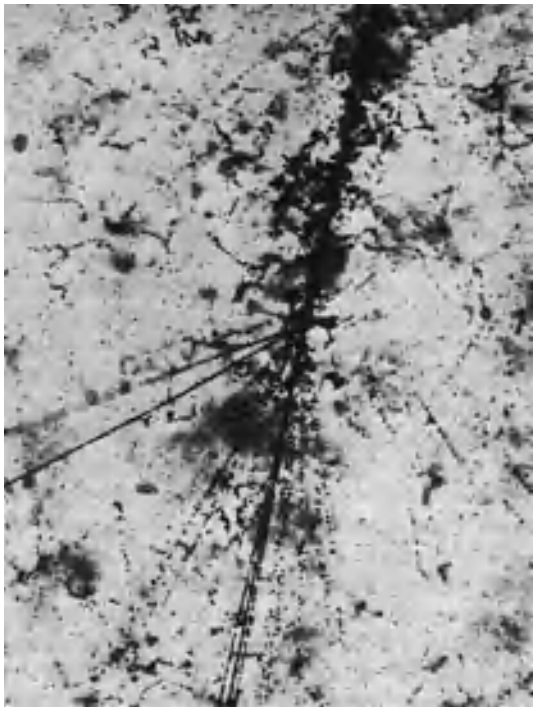

Nuclear emulsion showing iron nucleus ($Z = 26$) in cosmic radiation entering from upper right and producing large numbers of delta electrons, giving its track a heavy, "hairy" appearance.

response to ionizing radiation is related to the production of delta electrons by the primary ionizing particle. D. Allan Bromley

Delta resonance

A member of a class of subatomic particles called baryons, which exists in four electric charge states and has a total spin of $J = {}^3/_2$. In the underlying quark model, the delta resonance (Δ) consists of three quarks whose intrinsic spins of ${}^1/_2$ are lined up in the same direction. The Δ is closely related to the more familiar nucleon constituents of atomic nuclei, the neutrons (n) and protons (p). *See* NUCLEON; QUARKS.

The Δ was first observed as a resonant interaction of a beam of pi mesons (π) with a proton target. The probability of a scattering interaction between the π and the proton is strongly dependent on energy, attaining a maximum at the Δ mass of 1236 MeV/c^2 (where c is the speed of light). The formation of the very short-lived Δ (with a lifetime on the order of 10^{-23} s) is followed immediately by its decay back into pion and nucleon. *See* MESON; SCATTERING EXPERIMENTS (NUCLEI).

The understanding of some nuclear phenomena requires explicit treatment of π or Δ degrees of freedom. For instance, the Δ modifies the polarizability of the nucleus, and thus contributes to the quenching of the transition strength observed in charge exchange reactions [that is, (p,n) reactions] on nuclei. The Δ also plays a role in the quenching of the beta-decay process ($n \rightarrow p + e^- + \bar{\nu}_e$, where e^- denotes an electron and $\bar{\nu}_e$ an antineutrino) in the nucleus. High-momentum components are induced in the nucleus through the presence of Δ's. These are probed by a variety of transfer processes. The effect of the Δ is also seen in electromagnetically induced reactions [(γ,p), (γ,π^0), $(e,e'p)$, and so forth] involving the interaction of a real or virtual photon (γ) with the nucleus. *See* NUCLEAR REACTION; NUCLEAR STRUCTURE; RADIOACTIVITY.

The Δ thus plays an important role in a wide variety of nuclear phenomena, even under conditions of low energy and momentum transfer. The study of these phenomena reveals much about the presence of pions in nuclei, in addition to neutrons and protons. *See* BARYON; ELEMENTARY PARTICLE.

Carl Dover

Bibliography. B. H. Bransden and R. G. Moorhouse, *The Pion-Nucleon System*, 1973; P. Catillon, P. Radvanyi, and M. Porneuf (eds.), *Nuclear Structure*, 1982; J. M. Eisenberg and D. S. Koltun, *Theory of Meson Interactions with Nuclei*, 1980, paper, 1989; M. Rho and D. Wilkinson (eds.), *Mesons in Nuclei*, vol. 1, 1979.

Demagnetization

The process of reducing the retained level of magnetization (remanence) in an object. Demagnetization is necessary because the ferromagnetic elements iron, nickel, and cobalt and their alloys can become permanently magnetized if they are exposed to a magnetic field. There is a vast range of sizes of objects whose magnetization would degrade their performance. Examples range from mine-sweeping ships, though magnetic shields for cathode-ray tubes, down to even smaller objects. The level retained and the field required to demagnetize depend on their ferromagnetic behavior. Permanent magnets are called "hard" and are required to retain magnetization since it is this property that makes them useful. The magnetic materials used in transformers are called "soft" and here the retention of magnetization could detrimentally alter the performance of the transformer.

Mechanical factors such as residual stress levels also contribute to magnetic behavior. The dimensions of an object are important as are its material properties. If certain permanent magnets are made too thin, they will lose their magnetization. In this case, this unwanted demagnetization is caused by an internal magnetic field whose magnitude is related to the ratio of the cross section of the magnet poles to the thickness. Sometimes when the magnetic field associated with an object must be reduced to an acceptable level and demagnetization in impracticable, coils are added to the object and the current in them is adjusted so as to oppose the field. *See* FERROMAGNETISM; MAGNET; MAGNETIC MATERIALS; MAGNETISM; MAGNETIZATION; TRANSFORMER.

Need for demagnetization. When permanent magnetization remains in a material the object behaves like a magnet and produces a magnetic field in the

surrounding space. In situations where controlled magnetic fields are to be produced, such as magnetic resonance imaging (MRI) scanners, this behavior will be a problem. Also, if the environment in which a magnetometer that is sensitive to very small magnetic fields is housed produces its own magnetic field, then the magnetometer will not work as required. These are examples of situations where, as far as possible, the material selected should not retain magnetization. *See* MAGNETOMETER; MEDICAL IMAGING.

Techniques. Magnetic materials can exhibit an enormous range of behavior, and it is this variety that presents a challenge if they need to be demagnetized. A process that is suitable for one object may not be as effective for another, even if it is only the geometry of the object that has changed.

Use of magnetic fields. When ferromagnetic materials are incorporated in an object, a way of reducing inadvertent magnetization to an acceptable level is required. To demagnetize them, it is necessary to generate a magnetic field that is of sufficient magnitude and can be reduced to a very low value in small steps. The magnitude of the magnetic field should be such that the magnetization of the material is taken as close as possible to saturation. If the applied magnetic field is not strong enough, the success of demagnetizing will be limited. It is necessary to reduce the magnetic field through cycles that take the magnetization through reversals such that a north pole becomes a south pole and vice versa and then back to a north pole. An alternating magnetic field is usually employed and the frequency should be low enough that complete penetration of the magnetic field occurs. If the material is thick (thicker than 20 mm or 0.8 in.) and the electrical conductivity is high, then the frequency used should be low (not greater than 10 Hz). *See* SKIN EFFECT (ELECTRICITY).

A wire carrying a current can produce the magnetic field needed to demagnetize an object. Coils consisting of a number of turns are typical, and the object is placed inside the coil, where the magnetic field is higher. If the object is large, then the coil diameter will need to be large and this reduces the magnetic field at the center of the coil for a given current. If large currents cannot be produced, then the number of turns can be increased. The magnetic field produced will be nonuniform and it may be necessary to use other coaxial coils to improve the uniformity to ensure that the magnetic flux reaches the required value throughout the material volume. If the properties of the material vary with direction, then coils along more than one axis may be necessary.

When reducing the magnetic field it is very important that the change in the amplitude between one cycle and the next is not too large; otherwise, the demagnetization will not be complete. Usually, it is necessary to repeat the demagnetization process a number of times, typically three, and each additional attempt after the first will improve the demagnetization. To determine if the demagnetization is sufficient, the object should be used in the required manner and its performance should be assessed. If unacceptable magnetization remains, then the demagnetization process should be altered with changes made to the frequency and to the rate at which the amplitude of the magnetic field is reduced to zero.

Use of heating. Another way of demagnetizing a material is to heat it above the Curie temperature. When the material is returned to room temperature, the magnetization will have been reduced. When this approach is used, it is essential to assess whether heating to the required temperature irreversibly alters other properties of the material. Some alloys can undergo an internal structure change at a temperature lower than the Curie temperature that would make them unsuitable for the desired application. *See* CURIE TEMPERATURE.

Michael Hall

Bibliography. B. I. Bleaney and B. Bleaney, *Electricity and Magnetism*, 3d ed., 1990; D.-X. Chen, J. A. Brug, and R. B. Goldfarb, Demagnetizing factors for cylinders, *IEEE Trans. Magnet.*, 27(4):3601–3619, 1991; S. Chic Kazumi and S. H. Charap, *Physics of Magnetism*, 1964, reprint 1978; M. J. Hall et al., Low permeability reference standards with improved high magnetic field strength performance, *IEE Proc. Sci. Meas.Technol.*, 145(4):181–183, 1998; *Methods for the Determination of the Relative Magnetic Permeability of Feebly Magnetic Materials*, British Standards Institute, BS 5884:1999, 1999; *Methods of Measurement of d.c. Magnetic Properties of Magnetically Soft Materials*, International Electrotechnical Committee, IEC 60404 part 4:1997.

Demodulator

A device that recovers from a transmitted signal information that has been embedded in the signal by a process of modulation. The modification of a usually sinusoidal electrical carrier signal (or simply carrier) for the purpose of transmitting information that is carried on a second electrical modulating signal or information signal is called modulation. The device or circuit that performs the modulation is called a modulator. The information may be from any of a variety of sources such as sensor data, speech signals, or image data. The modulator varies the carrier in amplitude, frequency, phase, or some combination to embed the information for efficient transmission. After reception, this information is recovered from the received signal by a demodulator. During modulation, information is shifted in frequency from baseband to the carrier frequency for efficient transmission. Demodulation is the reverse process by which the information is restored in frequency to its baseband. *See* CARRIER; MODULATION; MODULATOR.

One primary class of modulation (and demodulation) is amplitude modulation, which is subdivided into the classes double-sideband transmitted carrier (variously called DSTC, AMTC, or "standard" AM); the more complex but power-efficient double-sideband suppressed carrier (DSSC);

the even more power- and bandwidth-efficient (and complex) single-sideband suppressed carrier (SSBSC, or simply SSB); vestigial sideband (VSB), in which one sideband is only partially removed, leaving only a vestige of itself; and quadrature amplitude modulation (QAM), in which two amplitude-modulated (AM) signals are transmitted simultaneously on orthogonal carriers (that is, 90° out of phase with one another), that consequentially stay out of one another's way while sharing a common spectrum. The other major modulation methods are frequency modulation and phase modulation; both are highly resistant to amplitude noise because the amplitudes of the received signal are forced to a constant value. *See* AMPLITUDE MODULATION; FREQUENCY MODULATION; PHASE MODULATION; SINGLE SIDEBAND.

The information signals usually are of analog origin but may be converted to digital format for more exact processing. Transmission is always analog or continuous, but within the receiver the signal may be converted to digital format by an analog-to-digital converter (ADC) after the IF (intermediate-frequency) amplifier stage because digital signals can be more precisely processed for superior clarity. The post-IF-stage demodulator may be either analog or, if an ADC is used, digital. Most digital demodulation methods mimic their analog counterparts because of the simplicity, reliability, and thoroughly understood behavior. *See* ANALOG-TO-DIGITAL CONVERTER; INTERMEDIATE-FREQUENCY AMPLIFIER; RADIO RECEIVER.

Demodulation of AM-signals. DSTC signals can be recovered by either envelope detection (also known as asynchronous or incoherent detection) or product detection (also known as synchronous or coherent detection). DSTC signals are of the form $\cos(2\pi f_c t) \cdot [A + B \cdot s(t)]$, where f_c is the carrier frequency, t is time, $s(t)$ is the modulating signal with a maximum amplitude of unity, and A and B are positive, real constants, where $A > B$, so that the bracketed quantity is always positive. For that reason, rectification in the recovery process will not introduce signal distortion. Envelope detection is a two-stage process that first rectifies the signal (that is, all negative values are either eliminated or converted to positive values) to produce the information signal, plus the rectified carrier. The rectified carrier appears as noise, which is then removed by lowpass filters. In product detection, the received signal is multiplied by a replica of the carrier (discussed below), thereby generating both the information signal and a double-frequency carrier signal, which must be removed by a lowpass filter. All AM signals, including single-sideband signals, can be demodulated by product detection followed by a lowpass filter. *See* AMPLITUDE-MODULATION DETECTOR; ELECTRIC FILTER; RECTIFIER.

Demodulation of QAM signals. A QAM transmitter modulates the analog or digital signal in both amplitude and phase. For digital signals, a so-called M-ary (that is, each data symbol can be represented by one of M possible states in the complex plane) digital QAM signal occupies the same bandwidth as a phase-shift keying (PSK) signal (discussed below), but the power efficiency is greater. The two carrier-replica demodulation signals are orthogonal or in quadrature (that is, 90° out of phase with one another). An M-ary digital QAM signal is generated by segmenting the input-signal data into M-bit lengths. These are divided into two sets, A and B, of $M/2$ bits each. A common segmentation approach is to steer the odd-numbered bits of the input-signal data into the A set and the even-numbered bits into the B set. This pair of $M/2$-bit words is modified by complementing the most significant bit of each word and appending a "1" at the end to provide a pair of $M/2 + 1$ bit words which will be treated as 2's complement signed numbers, passed through a pair of $M/2 + 1$ bit digital-to-analog converters (DACs), and then multiplied (modulated) by $\cos(2\pi f_c t)$ for the A set and by $\sin(2\pi f_c t)$ for the B set. These two sets of analog information are then summed for transmission. *See* DIGITAL-TO-ANALOG CONVERTER.

This seemingly arcane exercise provides a rectangular M-by-M array of evenly spaced points in the complex plane, centered at the origin, that represent the M^2 possible data states. Of course, $M = 2$ is binary. Digital QAM with $M = 4$ is robust and commonly used. As M increases, both the channel capacity and the sensitivity to noise similarly increase.

A QAM digital demodulator, or detector, synchronously demodulates the received signal in two parallel signal paths with the sine and cosine carrier replicas, integrates them for $M/2$ clock periods, samples the signals at the end of the integration intervals, passes the signals through a pair of ADCs which then strip off the least-significant bits and complement the most-significant bits, and interleaves the two resulting A and B bit streams to reconstruct the original digital data.

Phase-locked loops. All synchronous demodulation or detection methods depend upon the availability of a replica of the carrier. Exact frequency is demanded. Small phase errors which cause signal loss can be tolerated except in the case of QAM, where phase inaccuracies cause signals from each channel to bleed into the other. The most powerful and successful device to provide the necessary reference signals is the ubiquitous phase-locked loop (PLL).

Briefly, a PLL is a closed-loop system consisting of a voltage-controlled or numerically controlled oscillator (VCO or NCO) whose output is the sought-after frequency reference for synchronous demodulation. The oscillator output is combined with the signal to be demodulated in a phase detector. The phase-detector output, which is proportional to the phase difference between the two signals, is applied to a loop filter, which provides servo stability for the loop and integrates the signal to provide the command signal for the controlled oscillator. Depending on the sophistication of the design, the PLL can be as perfect in performance as the designer wishes. The gain of the loop filter determines the speed of response of the PLL. For amplitude demodulation, a

slow loop that clings tightly to the carrier frequency and ignores all modulation information is desirable. *See* PHASE-LOCKED LOOPS.

Demodulation of FM signals. Frequency-modulated (FM) signals are in the form below, where the con-

$$\cos\left\{2\pi \cdot \int_0^t [f_c + k \cdot s(\tau)]\, d\tau\right\}$$

stant k is the modulation index. A useful tool is based on the observation that the time derivative of cos (xt) is $x \cdot \sin(xt)$, which is just the 90° phase-shifted signal scaled in amplitude by the frequency. Consequently, one form of FM demodulation, or discrimination (because this process involves discriminating between frequencies), is time differentiation followed by amplitude detection. A suitable approximation to time differentiation over a narrow frequency band of interest is a linear filter whose amplitude response increases linearly with frequency. A second form of FM demodulation uses a wide-bandwidth PLL which is able to follow the FM signal. The PLL oscillator control input signal is the recovered frequency-demodulated signal, $s(t)$.

Transmission of digital data by frequency modulation is called frequency-shift keying (FSK). The frequency of the transmitted signal is determined by the digitized information signal. M-ary FSK employs M distinct frequencies. Binary FSK demodulation consists of determining if frequency A or frequency B is present. The simplest noncoherent approach (no carrier reconstruction is required) applies the signal to a pair of band-pass filters, one tuned to A and the other to B; each filter is followed by an energy detector, such as the cascade of a rectifier, integrator, and threshold. The greater output, A or B, is declared present. M-ary FSK demodulation is an obvious extension. *See* FREQUENCY-MODULATION DETECTOR.

Demodulation of PM signals. Phase-modulated (PM)—sometimes called phase-angle modulated—signals are in the form $\cos[2\pi f_c t + ks(t)]$. Since phase is the time integral of frequency, the demodulated PM signal can be recovered simply by integrating the output of an FM demodulator. Alternatively, the output of the PLL phase detector is proportional to the sine of $s(t)$. The PLL phase detector can be made into a linear phase demodulator by converting its inputs from sinusoids to square waves, but at the price of introducing a small amount of phase jitter in the PLL performance.

Transmission of digital data by phase-angle modulation is called phase-shift keying (PSK), and employs a single frequency. The phase of the transmitted signal is determined by the digitized information signal. Each data symbol in an M-ary PSK system may be represented by a point on a unit circle because the amplitude is held constant. The points are separated in angle (phase) by $360°/M$. In general, PSK demodulation is more difficult than FSK demodulation because of synchronization problems, but changes in state can be detected simply by delaying the signal by one clock period and comparing the delayed and undelayed signals in a correlator. A high output means that no transition has occurred. *See* PHASE-MODULATION DETECTOR.

Modems. Integrated digital communications sets are referred to as modems, for modulator/demodulator. Data modems are highly specialized devices. Almost all data modems today are adaptive, which means that they automatically adapt to variations in the transmission medium. *See* MODEM.

Stanley A. White

Bibliography. G. A. Breed (ed.), *Wireless Communications Handbook*, Cardiff, Englewood, 1992; S. Gibilisco, *Handbook of Radio and Wireless Technology*, McGraw-Hill, New York, 1999; J. G. Proakis, *Digital Communications*, 4th ed., McGraw-Hill, New York, 2001; T. S. Rappaport, *Wireless Communications*, 2d ed., Prentice Hall, Upper Saddle River, 2002; M. S. Roden, *Analog and Digital Communication Systems*, 4th ed., Prentice Hall, Englewood Cliffs, 1995; M. E. Van Valkenburg and W. M. Middleton (eds.), *Reference Data for Engineers: Radio, Electronics, Computer and Communications*, 9th ed., Butterworth-Heinemann, 2002.

Demospongiae

A class of the phylum Porifera, including sponges with a skeleton of one- to four-rayed siliceous spicules or of spongin fibers or both. Several genera lack a skeleton, and it is through a study of these seemingly primitive forms that the complicated structure of most adult Demospongiae may be understood. The Demospongiae constitute the most abundant and widely distributed group of sponges, occurring in the sea from the tidal zone down to abyssal depths (at least to 18,000 ft or 5500 m). Three families have invaded fresh water. The species vary in size from thin encrustations several centimeters in diameter to huge cake-shaped forms which may measure up to as much as 6.6 ft (2 m) in diameter.

Comparative studies of the embryology and early attached stages of sponges of the class Demospongiae suggest at least two evolutionary lines within this group. One line of development, the subclass Ceractinomorpha, is characterized by the presence of incubated parenchymella larvae; the other, the subclass Tetractinomorpha, includes oviparous species with parenchymella larvae. However, the more primitive families have incubated amphiblastulalike larvae.

Ceractinomorpha. Among the Ceractinomorpha, the genus *Halisarca*, lacking skeletal elements, is a primitive form. The larva of *Halisarca* is a diploblastula or parenchymella with an outer layer of flagellated cells and an inner mass of presumptive ectomesenchymal cells (**Fig. 1***a*). The out flagellated cells lose their flagella, migrate into the interior (Fig. 1*b* and *c*), and later differentiate into choanocytes. Other cell types characteristic of the adult sponge differentiate, and inhalant canals begin to form. The young sponge soon develops a single internal cavity lined with choanocytes (Fig. 1*d*), and an oscular opening breaks through at the apex (Fig. 1*e*). At

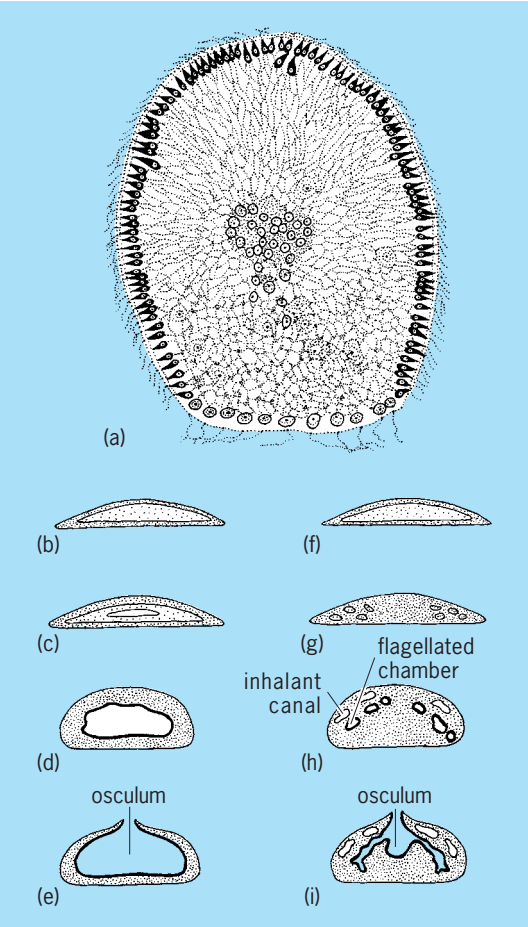

Fig. 1. Development of ceractinomorph sponges. (*a*) Free-swimming parenchymella larva of *Halisarca*. (*b*–*e*) Metamorphosis of *Halisarca*: (*b*) newly settled parenchymella; external flagellated cells migrate internally; (*c*) internal cavity appears; (*d*) choanocytes line the central cavity; and (*e*) osculum breaks through and young asconoid stage is formed. (*f*–*i*) Metamorphosis of *Aplysilla*: (*f*) newly settled parenchymella; external flagellated cells migrate internally; (*g*) islands of choanocytes form in the internal mass of cells; (*h*) flagellated chambers and inhalant canals appear; and (*i*) flagellated chambers join exhalant canal system which opens through osculum; young syconoid sponge is formed.

this stage of development, called the rhagon stage, the young *Halisarca* is essentially identical with the asconoid grade of construction found in some Calcarea. Later, folds in the choanocytal layer lead to the formation of flagellated chambers, and a transitory syconoid grade of construction exists. Eventually the flagellated chambers are isolated from one another through the appearance of exhalant canals which converge on the oscula. The adult sponge has a simple leuconoid structure with elongate flagellated chambers having wide openings into exhalant canals and communicating with the surface pores by means of an inhalant canal system. *Aplysilla*, a closely related genus of sponges with a branching, nonanastomosing fibrous skeleton, has a similar developmental history (Fig. 1*f*–*i*), except that the earliest rhagon stage is syconoid in structure with a folded choanocytal layer. *See* CALCAREA.

Metamorphosis from the larval condition in the Demospongiae characteristically involves a transitory rhagon stage with a simple leuconoid canal system. During further development, the flagellated chambers are isolated between the inhalant and exhalant canals to produce leuconoid canal systems of varying complexity (**Fig. 2**).

Ultrastructural studies of the choanocytes and choanocyte chambers of representative genera of the Demospongiae have revealed differences among the orders of the class. In some genera of the orders Haplosclerida and Petrosiida, belonging to the subclass Ceractinomorpha, the choanocyte chambers lie across the inner ends of the inhalant canals (**Fig. 3**) and are separated therefrom by a pinacocyte layer that is perforated beneath the prosopyles of the

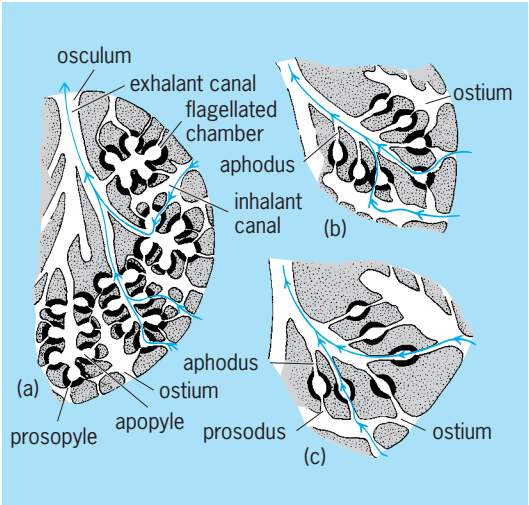

Fig. 2. Types of leuconoid canal systems. (*a*) Eurypylous chambers opening directly into exhalant canal. (*b*) Aphodal chambers; a narrow canal or aphodus leads from the chamber to exhalant canal. (*c*) Diplodal chambers; narrow canal or prosodus intervenes between inhalant canal and chamber, as well as between chamber and exhalant canal. Arrows show direction of water flow.

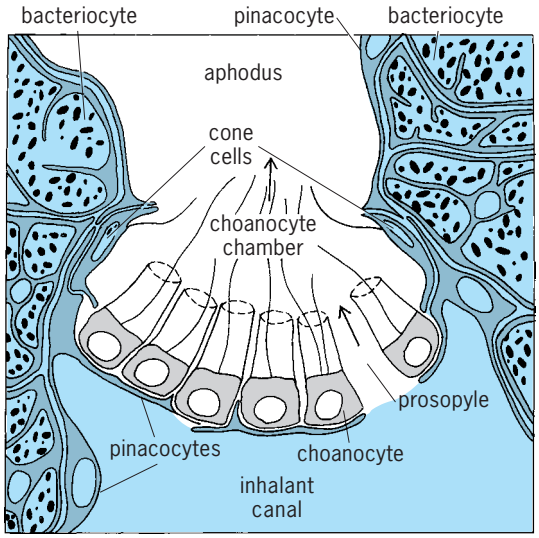

Fig. 3. Section of choanocyte chamber of *Petrosia ficiformis*, showing pathway of water flow from end of inhalant canal through chamber to aphodus.

chamber to allow water to pass from the inhalant canal into the chamber. Between the chamber and the exhalant canal lies a pore that allows water to escape from the choanocyte chamber into an aphodus (Fig. 2*b*) of the exhalant canal system. This pore is surrounded by a ring of cells (called cone cells from their shape) that are part of the choanocyte chamber. This quite complicated structure is, so far as is known at present, restricted to demosponges of the orders mentioned above.

In form, ceractinomorph sponges vary from encrustations, either thin or massive, to lobate and upright branching colonies. The shallow-water species tend to be more plastic in form than the deep-water species, which usually exhibit little intraspecific variation in shape.

Tetractinomorpha. In the other subclass of Demospongiae, the Tetractinomorpha, the skeletonless genus *Oscarella* has a primitive structure. Cleavage results in a solid mass of cells (morula) which later becomes hollow by cytolysis of the interior cells, rich in food reserves. Upon being freed from the parent, the hollow larva is made up of a single layer of flagellated cells and is known as an amphiblastula (**Fig. 4***a*). After a short free-swimming period, the larva attaches to the substrate by its anterior pole and flattens out (Fig. 4*b*). Gastrulation occurs at this point as the anterior half of the larva invaginates (Fig. 4*c*). The blastopore closes as the edges of the now double-layered larva push together against the substratum (Fig. 4*d*). The internal layer of cells is now thrown into folds which pinch off to form cavities which will become the flagellated chambers. Simultaneously a depression forms in the apical ectoderm, and the rudiment of the exhalant canal system appears (Fig. 4*e*). The depression deepens to form a cavity giving off branches which push their way among the flagellated chambers (Fig. 4*f*). The latter finally open into these cavities which become the exhalant canals; inhalant canals push in from the surface of the sponge and join the flagellated chambers. The metamorphosed larva of *Oscarella* assumes the leuconoid grade of construction from the start, with isolated flagellated chambers communicating with inhalant and exhalant canals. The adult *Oscarella* and a related genus, *Plakina*, with two-, three-, and four-rayed spicules, retain the simple leuconoid structure. The sponge consists of a folded wall, each fold made up of a dermal layer and a group of flagellated chambers opening into an exhalant canal. In a more complicated stage, as seen in *Plakortis*, the dermal membrane spreads over the outer ends of the folds, and subdermal cavities are developed. In many species of Tetractinomorpha an extensive cortex is developed, consisting of a network of fiber cells, called desmacytes, which in some cases is overlaid by a thick gelatinous layer containing amebocytes. In form, the species of these orders may be thinly encrusting or massive, but often they have spherical or ovoid shapes. Branching species rarely occur.

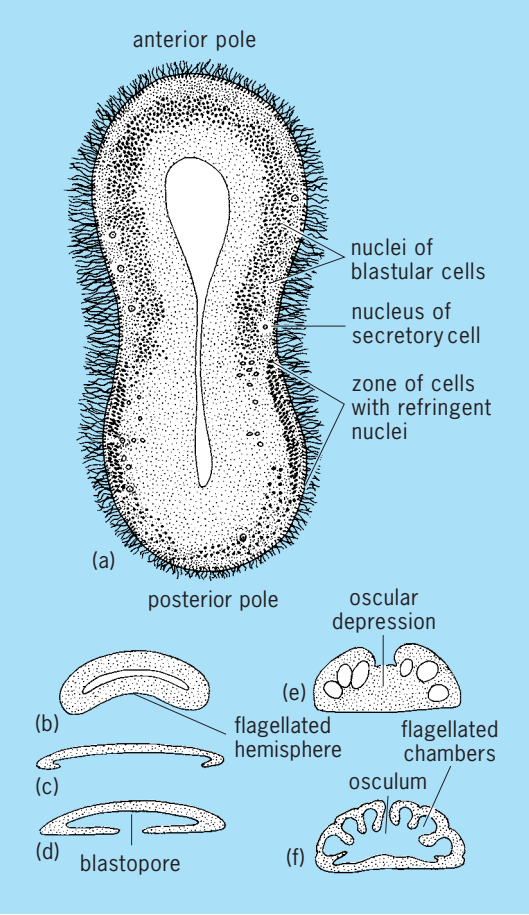

Fig. 4. ***Oscarella*. (*a*) Free-swimming amphiblastula larva. (*b*) Newly settled larva. (*c*) Gastrulation begins by invagination of flagellated hemisphere. (*d*) Later stage of gastrulation; blastopore closing. (*e*) Formation of osculum. (*f*) Young leuconoid stage is formed.**

Skeleton. The skeletal system consists of spicules (sclerites), spongin fibers, or both.

Spicules. The spicules of Demospangiae are intracellular secretions of scleroblasts, cells derived from archeocytes. Each monaxonid spicule is laid down within a single cell by continuous deposition of hydrated silicon dioxide around an organic axial fiber. Above a minimal concentration of silica in the surrounding medium, the length of spicules is independent of the amount of silica present. The thickness of the spicules varies up to a maximal value in correlation with silica concentration, however. Tetraxonid spicules are apparently also formed in individual scleroblasts. Microscleres are formed in special scleroblasts; they require a higher minimal content of silica for formation than do megascleres (**Fig. 5***a–d*).

Fibrous skeletal elements. The reticulum of prominent spongin fibers that forms the skeleton of members of the order Dictyoceratida (bath sponges and related forms) is composed of a felt-work of collagen fibrils, without periodic striation, laid down by specialized cells, the spongioblasts (Fig. 4*e*). Similar spongin fibers enclose or join together siliceous spicules in sponges of the orders Haplosclerida and Poecilosclerida. Intercellular collagen fibrils occur in Demospongiae as well. These fibrils, laid down by collencytes, lophocytes, or pinacocytes, have a

periodic striation and are oriented in a parallel fashion in loose fascicles.

Ceractinomorph skeletal elements. Among the ceractinomorph sponges, spongin tends to be of common occurrence. In the orders Dendroceratida and Dictyoceratida, the latter of which includes the commercially valuable bath sponges, spicules are absent, and the skeleton is formed of spongin fibers only. In the orders Haplosclerida and Poecilosclerida, varying quantities of spongin occur along with siliceous monaxonid spicules. In some species a network of spongin fibers occurs in which the spicules are embedded; in others spongin serves as an interspicular cement. In the order Halichondrida, spongin is rarer in occurrence but is always present in the form of short tracts or as a cement. The cement helps to unite the irregularly arranged siliceous monaxonid spicules.

Ceractinomorph microscleres are commonly C- or S-shaped (sigmas), bow-shaped (toxas), or anchor-shaped (chelas), or are fine and hairlike (raphides).

Tetractinomorph skeletal elements. Tetractinomorph sponges of the orders Homosclerophorida, Choristida, and Spirophorida have little or no spongin and almost always have some tetraxonid siliceous megascleres to which may be added monaxonid types. When triaenes are present, they usually occur in tracts radiating from the central part of the sponge to the surface.

The orders Hadromerida and Axinellida are characterized by a skeleton of monaxonid siliceous megascleres accompanied by varying quantities of spongin. The spicules characteristically occur in tracts radiating from the central regions of the sponge, or in the case of those with an abundance of spongin a plumose arrangement is found. Microsclere types include asters, with numerous rays diverging from a central point; streptasters, spiny rods, often twisted spirally; sigmas; and raphides. The same range of shapes occurs as in other tetractinomorph orders, and in addition many species have an upright branching form. *See* HADROMERIDA.

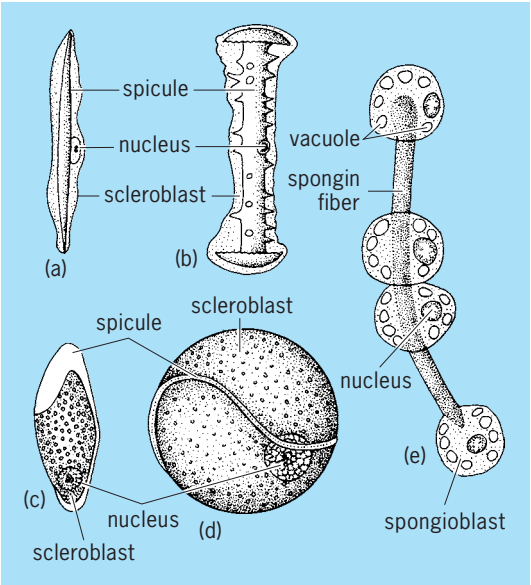

Fig. 5. Skeleton formation in Demospongiae. (*a*) Diactinal megasclere. (*b*) Amphidisc forming within a scleroblast. (*c*) Chela formation within scleroblasts. (*d*) Sigma formation within scleroblasts. (*e*) Spongin fiber in process of formation by spongioblasts in *Haliclona*.

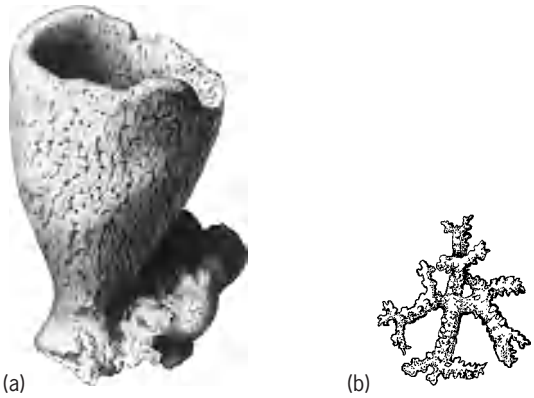

Fig. 6. *Discodermia ornata*, a lithistid sponge. (*a*) Whole sponge. (*b*) Desma.

Lithistid sponges. Some species of Demospongiae are characterized by the presence of spicules called desmas (**Fig. 6**). These are formed by the secondary deposition of silica around ordinary monaxonid or tetraxonid spicules. Supplementary knobby branches often develop, and articulating processes may occur by which neighboring desmas become interlocked to form a stony or lithistid skeleton. Because of the rigidity of the skeletons of lithistid sponges, they are commonly preserved as fossils and are the best-known Demospongiae in the fossil record. Paleontologists have tended to classify such sponges in an order Lithistida. It is apparent from studies of Recent species with a lithistid skeleton that this modification has arisen many times in the evolution of the Demospongiae and that the order Lithistida has no validity. It is difficult to place fossil lithistids among the several orders of the class Demospongiae, however, unless developmental stages of the peculiarly modified spicules are present.

Sclerosponges. While some authors regard the sclerosponges as a subclass of the Demospongiae, they are here regarded as a separate class.

Willard D. Hartman

Dendritic macromolecule

A large molecule having a well-defined three-dimensional structure. Dendritic macromolecules play a crucial role in the chemistry of living systems. In contrast to the high level of structural precision that characterizes many biologically active macromolecules, the sizes and shapes of macromolecules made by polymer chemists are usually far less controlled. Most synthetic polymers are best described as statistical mixtures. However, chemists have sought to develop ways to prepare large molecules with more control over their architecture. If properly designed, such molecules might be capable of performing chemical or physical functions

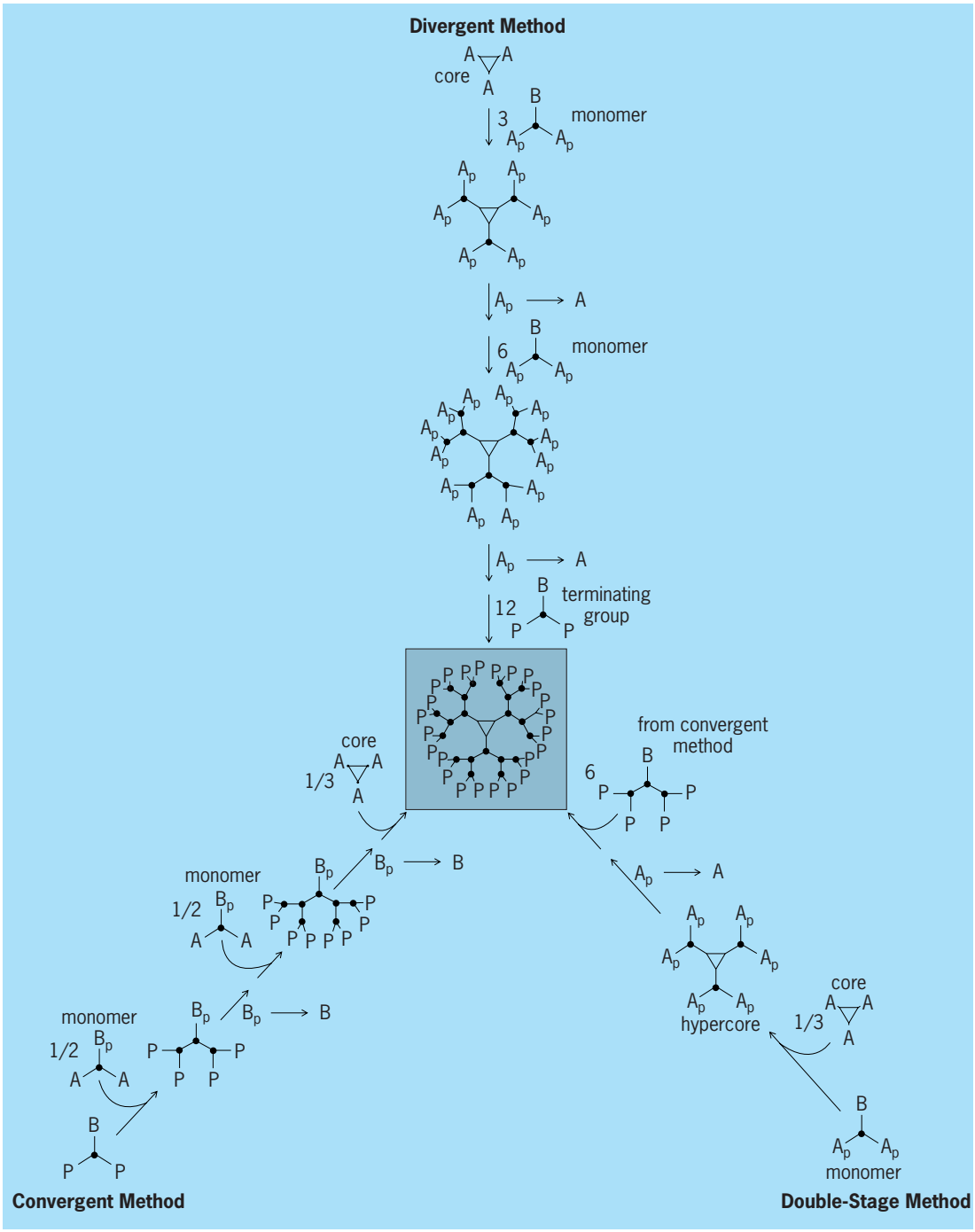

Fig. 1. Diagram of the three synthetic methods used to prepare structure-controlled dendrimers with triconnected branch junctures. Symbols are explained in the text.

reminiscent of the macromolecules found in living systems.

Terminology. Dendritic macromolecules are characterized by a highly branched molecular connectivity, whereby each repeat unit (•) forms a branch juncture (**Fig. 1**). The monomers used to prepare dendrimers possess three or more functional groups, and are of the type AB_2, AB_3, and so forth, where A and B represent a functionality (a site of chemical activity) that can combine to form a new covalent bond. Monomer chemistry is thus similar to that used to make condensation polymers except that the functionality is higher (there are more sites of chemical activity). Repeat units in dendritic macromolecules can be viewed as being arranged in concentric shells centered on a core (∇). By analogy to a genealogical tree, each shell represents a generation of repeat units. The generation farthest from the core makes up the peripheral functionality (P). Dendrimers that emanate from a core of single, double, triple, etc., connectivity are referred to as mono-, di-, tri-, etc., dendrons, respectively. Monodendrons

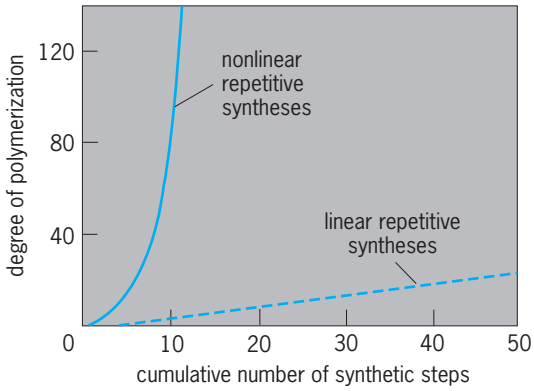

Fig. 2. Plot of the degree of polymerization versus the number of synthetic steps for a typical dendrimer synthesis: nonlinear repetitive synthesis. The efficiency of the dendrimer synthesis is compared to a typical linear repetitive synthesis in which only a constant number of monomers is added during each generation.

have also been called dendritic wedges, and the monodendron core is known as the wedge focal point. *See* POLYMER.

Synthetic strategies. There are two distinct categories of synthetic processes lead to dendritic macromolecules. In the first category, uncontrolled self-condensation of AB_n ($n \geq 2$) monomers generates dendriticlike macromolecules (the one-step method). In the second category, protecting-group chemistry is used to harness monomer condensation so that macromolecules with potentially perfectly defined composition and constitution can be produced (the stepwise, repetitive method). In the stepwise process, monomers are added one generation at a time. The number of monomers added per generation increases in a nonlinear fashion, making the synthesis of these large molecules not only feasible but highly efficient (**Fig. 2**). This type of growth process can be contrasted with a linear repetitive synthesis, in which a constant number of monomer units is added in each round of the synthesis.

Many different dendritic systems have been made by the one-step approach, including several systems that have counterparts as engineering resins. The list includes polyesters, polyamides, poly(siloxysilanes), polyethers, and poly(ether ketones). Like traditional synthetic macromolecules, the size and shape of dendrimers made by the one-step method are best described as statistical mixtures. These dendrimers can be viewed as structurally irregular dendritic wedges with multiple reactive B groups on their periphery and a single reactive A group at their focal point. In these systems an important chemical defect that is thought to limit molecular weight involves intramolecular cyclization of the focal-point functional-group backbiting (an intramolecular chemical reaction) onto one of the many peripheral B groups.

Divergent method. Stepwise synthesis of structurally well-defined dendritic macromolecules typifies the divergent method of dendrimer preparation (Fig. 1). Growth begins at the core and proceeds outward. Because all of the A functional groups on the monomer are protected and therefore not reactive (that is, A_p), growth stops once all of the unprotected factional groups have been consumed. The next round of synthesis can begin only after the A_p groups on the dendrimer's periphery are transformed to their reactive counterparts (that is, $A_p \rightarrow A$). Many different types of dendritic macromolecules have been prepared by the divergent method, including polyamides, poly(amido alcohols), poly(amido amines), polyethers, poly(siloxanes), poly(carbosilanes), and poly(alkanes).

Convergent method. A second route for preparing structure-controlled dendrimers by the stepwise process is the convergent method (Fig. 1). In this method, growth begins from the periphery and proceeds inward, first yielding a series of monodendrons. At the end of the synthesis, the focal point of these monodendrons can be coupled around a core of desired functionality to provide the final dendrimer. The advantage of the convergent method is that the number of reactive groups is constant throughout the synthesis while the molecular weight doubles at each stage. A disadvantage is that the focal point tends to become sterically encumbered, making the synthesis of higher generations more difficult. Dendritic systems prepared by the convergent approach include polyethers, polyesters, polyamides, poly(phenylenes), and poly(phenylacetylenes).

To overcome the problem of steric crowdedness associated with the convergent method, a hybrid approach known as the double-stage convergent method has been described (Fig. 1). In this method, smaller dendritic wedges prepared by the normal convergent approach are coupled to a multifunctional core (hypercore). The hypercore itself can be prepared by either a divergent or convergent process.

An alternative convergent method that alleviates the problem of steric crowdedness involves the use of monomers of progressively larger size at each generation of dendrimer growth. One advantage of this method is that dendrimers having very large dimensions can be assembled rapidly, as has been demonstrated by the ten-step synthesis of a phenyl-acetylene dendrimer that spans more than 12.5 nanometers.

Architecture. The architectures of dendritic macromolecules are dramatically different from those of conventional macromolecules. Variations in molecular size, shape, and flexibility are all possible, depending on the monomer's chemical structure and geometry, the branch-point multiplicity, and the number of repetitive cycles carried out. The architecture most typical of dendritic macromolecules is a globular shape where segment density increases in going from the core to the periphery. The unusual structure of dendritic macromolecules, yet to be fully investigated, results in rheological and solubility characteristics that are dramatically different from linear macromolecules. *See* MACROCYCLIC COMPOUND; ORGANIC SYNTHESIS.

Jeffrey S. Moore

Bibliography. F. Vögtle, *Fascinating Molecules in Organic Chemistry*, 1992.

Dendrochronology

The science that uses annual tree rings dated to their exact year of formation for dating historical and environmental events and processes. Trees are intimately bound to the environment, making them one of the most consistent and dependable recorders of processes and events that occur in nature. Trees, like most plants, are sensitive to both natural (precipitation and temperature patterns) and human-related (air and water pollution) events that trigger certain responses in the vigor of the tree as seen in its growth rate. In most geographic regions, climate patterns in any year cause a response by trees in the volume of wood the tree produces, and often leave indelible "fingerprints" in certain physical and chemical properties of the wood. These fingerprints can be seen in the varying widths of tree rings. In some years, environmental conditions may be favorable for tree growth, allowing trees to produce greater volumes of wood. In other years, climate conditions may be generally unfavorable for tree growth, causing a reduction in the volume of wood produced. *See* DENDROLOGY; TREE; TREE GROWTH.

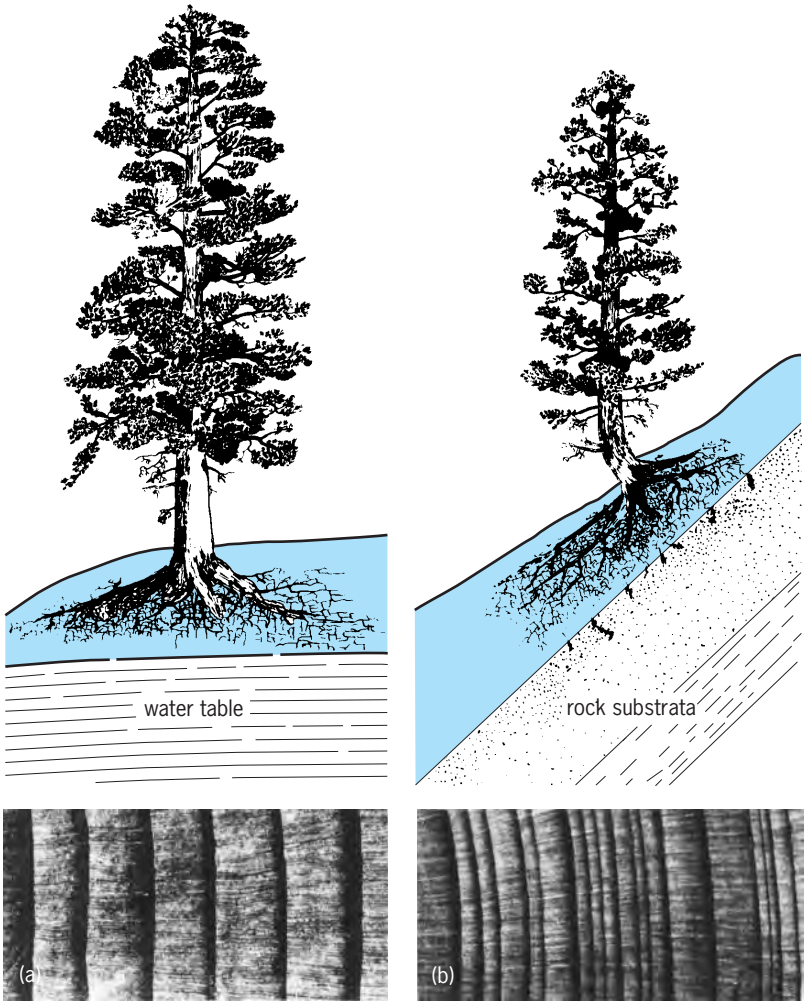

Fig. 1. Ring patterns showing complacent and sensitive ring series. (*a*) Climate has minimal effect on ring widths when trees have ample moisture. (*b*) Maximum variations occur on sites where factors favorable to growth are limited. (*Laboratory of Tree-Ring Research, University of Arizona*)

History. In the early 1900s, Andrew E. Douglass, an astronomer at the University of Arizona, was analyzing the relationship between sunspots and climate. He speculated that any change in solar energy caused by sunspot activity would cause changes in the amount of energy received by the Earth, and therefore would affect Earth's climate. Because climate affects the growth of plants, Douglass reasoned that sunspot cycles should be reflected in the growth of trees. His subsequent studies established a link between tree-ring widths and climate, and he soon confirmed the cross-matching of tree-ring widths among different trees, a technique known as crossdating. In 1914, Clark Wissler of the American Museum of Natural History noted Douglass's successful attempts at dating tree rings, and asked him if he could use tree-ring techniques to date the Puebloan ruins of the American Southwest. Fifteen years later, dendrochronology matured into an accepted science when Douglass overlapped his chronology developed from living trees with chronologies developed from trees used in many southwestern archeological sites. This accomplishment, published in *National Geographic Magazine* in 1929, provided archeologists with dates to the exact year when many southwestern pueblos were built. The importance of dendrochronology eventually led to the establishment of numerous tree-ring laboratories throughout the United States and Canada, Europe, South America, Asia, and Australia. *See* CLIMATOLOGY; DATING METHODS.

Annual ring formation. A tree grows by forming a new sheath or layer of woody tissue each year, much like a stack of cones one on top of the other. This growth occurs in a thin layer of cells that completely shrouds the tree just inside the bark, and is called the vascular cambium. Here, cambial cells divide, with outer cells contributing to the formation of phloem, and inner cells contributing to the formation of woody tissue. In temperate and subpolar regions, most trees break from winter dormancy and begin growth by forming new cells of wood using nutrient reserves stored from the previous growing season. In conifer trees, these cells are large and less dense, and have thin walls, producing light-colored wood called earlywood. Toward the end of the growing season, before the tree begins its dormancy period, smaller, denser, thick-walled cells are formed that are darker in color and called latewood. Taken together, these two bands of wood form an annual tree ring. In certain hardwoods, such as oaks, large cells called vessels are formed in the earlywood, while the latewood often lacks such vessels. *See* WOOD ANATOMY.

The formation of annual rings holds true for most trees in temperate and subpolar latitudes, but recent studies have shown some tropical and subtropical tree species also may produce annual rings. Occasionally, a tree may form latewood toward the end of the growing season, but a new flush of growth

due to favorable environmental conditions may again form earlywood cells. This intraannual band of latewood between earlywood bands is a false ring, but often is anatomically different from a true tree ring when magnified. In contrast, environmental conditions during some years may be so unfavorable that some trees may not form new wood, resulting in a missing ring, or a locally absent ring in the trunk of the tree. In most cases, the locations of false and locally absent rings are identified by dendrochronologists using the process of crossdating.

Principles. To search for and collect tree-ring data, dendrochronologists are guided by the following principles.

Limiting factors. Rates of plant processes are constrained by the primary environmental variable that is most limiting. For example, precipitation is often the most limiting factor to plant growth in arid and semiarid areas. In these regions, tree growth cannot proceed faster than that allowed by the amount of precipitation, causing the width of the rings (that is, the volume of wood produced) to be a function of precipitation.

Ecological amplitude. Plant processes function at their optimum within a defined geographic range constrained by one or more interacting limiting factors. Species are more sensitive to changes in environmental factors along the periphery of this geographic range. Tree species useful for dendrochronology are often found near the margins of their natural range.

Site selection. Sites useful to dendrochronology can be identified based on criteria that will produce tree-ring series sensitive to the environmental variable being examined. For example, trees especially responsive to drought conditions are usually found where the water supply is limiting, such as rocky outcrops and cliffs, mountain slopes, and arid and semiarid regions (**Fig. 1**). *See* PLANT-WATER RELATIONS.

Replication. The strength of the environmental signal being investigated can be maximized by collecting more than one sample per tree and more than one tree per site. Obtaining numerous trees from one site, and perhaps several sites in a region, ensures that the amount of noise (environmental or physiological factors not being studied) in a tree-ring series is minimized.

Collecting tree rings. After appropriate sites and trees have been located, dendrochronologists collect tree-ring data using an increment borer, a threaded hollow metal tube that is screwed into the trunk of a tree (**Fig. 2**). An extractor is inserted under the tree core that now resides inside the borer shaft, and is gently pulled from the tree. The hole left by the increment borer rarely does harm in conifer trees, but may cause internal wood discoloration and the formation of fungi that decay wood in hardwood tree species. Usually, at least two cores are extracted from opposite sides of the tree, and are placed in paper straws or plastic tubes to prevent breakage. In the laboratory, the pencil-thin cores are glued to grooved wooden mounts and sanded with progressively finer sandpaper (usually with a belt or orbital sander), until a high polish is obtained and the cell structure of the wood is easily discernible under a microscope. Occasionally, certain projects require complete cross sections from the trunks of dead trees, and a chainsaw is most often used.

Crossdating. Crossdating is the primary guiding principle in dendrochronology, and concerns the matching of patterns of ring widths from one tree with corresponding patterns for the same years from another tree (**Fig. 3**). Crossdating allows scientists to accurately assign calendar dates to tree rings by matching the sequence of tree-ring widths against a known reference chronology. Crossdating is possible because climate is largely a regional phenomenon, affecting trees in a like manner, so that similar patterns of ring widths are produced among many trees. A unique pattern of wide and narrow rings formed during a 50-year period is unlikely to be formed during any other 50-year period, because climate varies from year to year. Furthermore, crossdating helps identify false and locally absent rings that may otherwise be recorded as true rings.

In many laboratories, crossdating is accomplished with graphical comparisons called skeleton plots (**Fig. 4**) that accentuate the importance of narrow rings. The pattern of narrow rings is matched, or crossdated, against a reference chronology until a firm match is found. The tree-ring widths are then measured, and the crossdating accuracy is statistically confirmed using computer software. To ensure a high level of confidence in the dates assigned to tree rings, dendrochronologists look for high levels of statistical significance, higher than those normally used in statistics to identify a probable match. Only after a tree-ring series has been crossdated precisely,

Fig. 2. Small core being taken from a tree with an increment borer. (*Laboratory of Tree-Ring Research, University of Arizona*)

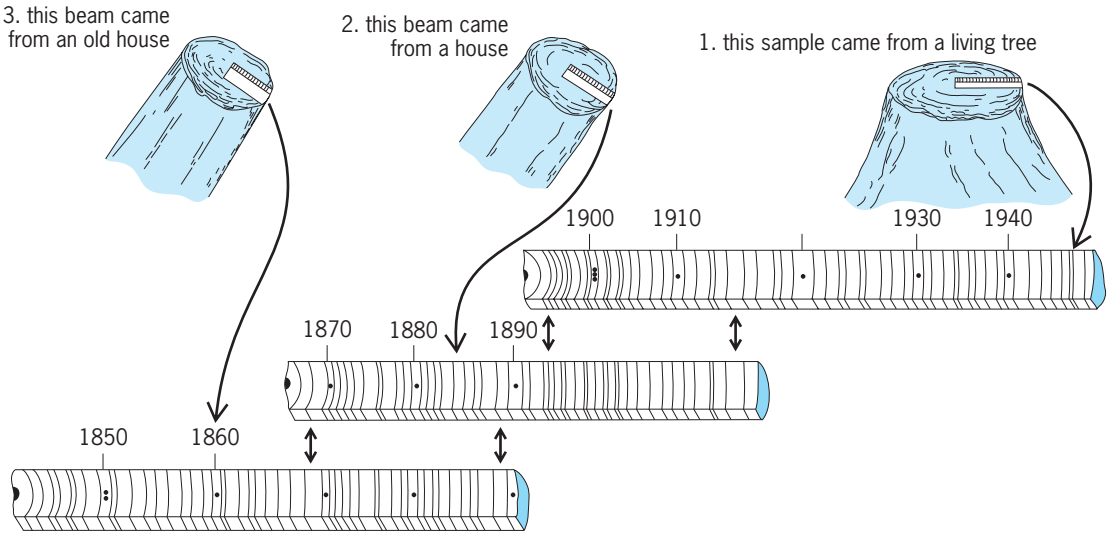

Fig. 3. Chronology building. Chronologies are extended backward in time by matching patterns of wood samples whose ages overlap.

both graphically and statistically, can dates for a wood sample be used.

Measuring tree rings. The widths of individual tree rings are measured using a measuring stage accurate to a hundredth of a millimeter. This stage is interfaced to a microscope, screen display, and a computer running software that records the measurement of each ring. Other physical properties of tree rings can be measured and used, including the average density of the tree ring, the minimum density of the earlywood, or the maximum density of the latewood, as well as the widths of the earlywood or latewood themselves. Density measurements are obtained using x-ray images made of tree cores that have been thin-sectioned with a microtome. In recent years, scanning techniques, image analysis, and sophisticated software have provided new techniques to easily capture tree-ring data. The information from scanned images of tree rings produces data similar to the changing density of the individual cells that constitute the tree rings.

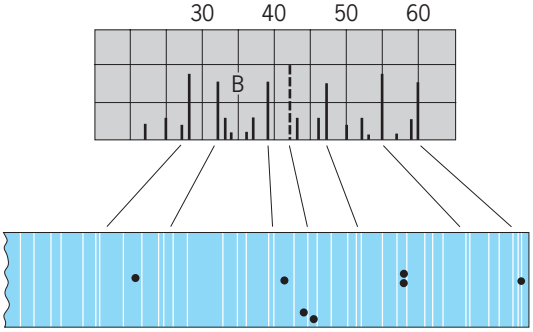

Fig. 4. Skeleton plot from an idealized tree-ring sequence. Ring widths are represented by vertical lines in an inverse proportion; that is, tall lines represent narrow rings. An exceptionally wide ring is designated by a B (big). The location of a missing ring, marked on the specimen by offset pinholes, is identified by a broken line. The number scale indicates years. (*After R. Berger, ed., Scientific Methods in Medieval Archeology, UCLA Center for Medieval and Renaissance Studies Contribution, IV, University of California Press, 1970*)

Chronology building. The measurements for individual years create a continuous time series for the full length of the individual cores. However, the raw measurements are rarely suitable for developing a site chronology because (1) trends are apparent simply due to physiological aging, (2) individual trees experience differences in growth rates, and (3) possible localized ecological disturbances (for example, the death of an individual tree) may impact growth of a subset of the sampled trees. Dendrochronologists are interested in average growth rates of trees at a site and must control for variation not due to climate. To accomplish this, all measurement series are separately standardized using complex mathematical expressions (for example, negative exponential curves) that provide predicted values of tree growth in any given year. The actual measurement is then divided by the predicted value (or sometimes the predicted value is subtracted from the actual measurement) to create an index of tree growth. The indices from all series for any given year are then averaged together to develop an average index. The final tree-ring index chronology represents the collective information from numerous trees within a site, and captures the majority of variability in tree-ring widths for a particular year.

Applications. Dendrochronology has become a useful tool in many areas of research.

Dendroarcheology. This science uses tree rings to date wood material from archeological sites or artifacts, and is most often applied in the southwestern United States and Europe. Dendroarcheology has established the exact year of construction for individual rooms in many ancient Native American ruins throughout the Southwest, providing important

Fig. 5. Bristlecone pines, which attain ages of more than 4000 years, form the basis for the longest chronology yet developed. (*Laboratory of Tree-Ring Research, University of Arizona*)

clues on changes in human populations hundreds of years before Euro-American settlement. In Europe, dendrochronologists often date the period of construction for barns, manors, cathedrals and churches, Roman bridges, wells and fountains, pile dwellings in lakes, and Neolithic settlements. Recently, dendrochronologists announced the development of a multimillennial tree-ring chronology that helped clarify Mediterranean archeological history. *See* ARCHEOLOGY.

Dendroclimatology. This science was first developed by Edmund Schulman, who was a student of Douglass and reconstructed climate for the American Southwest. Because temperature, precipitation, and other climatic variables affect tree growth, the climatic information can be mathematically extracted from the tree-ring record and reconstructed back in time for the length of the tree-ring record. For example, bristlecone pines growing in the White Mountains of eastern California have provided information on both temperature and precipitation fluctuations for the past 8000 years (**Fig. 5**). Reconstructions of annual rainfall for the American Southwest over the past 2000 years have provided new clues concerning century-scale climate change and its probable impact on ancient Native American cultures. Dendroclimatologists also use tree-ring records to quantify the rising levels of atmospheric carbon dioxide to better understand global warming.

Dendroecology. This science analyzes changes in ecological processes over time using tree-ring information. Forest trees are subject to a variety of ecological processes and disturbances that may alter forest structure and composition. For example, periodic insect outbreaks (such as the eastern spruce budworm) can defoliate or kill large stands of trees. Defoliation can be recognized in the tree-ring record as a group of continuous and progressively narrower tree rings. Other applications that fall under dendroecology include analyzing (1) the effects of air, water, and soil pollution on tree growth and forest health, (2) the age, maturity, and successional status of forest stands, and (3) the effects of human disturbances and management on forest vitality. *See* ECOLOGY.

Dendropyrochronology. This science reconstructs the history of wildfires from tree rings. Humans have greatly affected the natural course of wildfires, largely by fire suppression, often not realizing the beneficial nature of wildfires to forest ecosystems. In recent years, many land management agencies have sought to restore fire. This field takes advantage of the fact that a low-intensity surface fire will kill a part of the living cambium on the lower portion of the tree trunk which subsequent growth will preserve, thus leaving a fire scar in the tree-ring record. Dendrochronologists date tree rings that contain these fire scars to learn the frequency of past fires before widespread and pervasive human disturbances (such as fire suppression) occurred.

Dendrogeomorphology. This science studies earth surface processes using tree-ring data. Mass movements, such as landslides and snow avalanches, can kill or bend the trees in the movement path or along the edges of the movement. Tree-ring scientists can date when these trees were killed (by dating the outer ring of the tree) or when they were bent (by analyzing when dramatic changes in tree growth occurred). Therefore, a chronology of past landslide or avalanche events can be developed back in time to help assess the hazard potential for the low-lying areas at the base of these slopes. Tree rings are also used to learn about glacier movements and past climate by dating the moraines created along the sides and at the terminus of these glaciers. *See* GEOMORPHOLOGY.

Dendrohydrology. This science uses tree-ring data to investigate and reconstruct hydrologic properties, such as streamflow and riverflow, runoff, and past lake levels. This field is closely aligned with dendroclimatology, and requires statistical calibrations of tree-ring data, such as ring width and density, with hydrologic properties recorded during the historical period. For example, estimated flow for the Colorado River using information from gaging stations has been reconstructed from tree-ring data, and has shown that the flow of the river has fluctuated greatly prior to the twentieth century. Riverflow amounts based on twentieth-century observations are therefore not an accurate indicator of potential riverflow in the future. *See* HYDROLOGY.

Dendrochemistry. This important, emerging field of dendrochronology analyzes the chemical composition of tree rings, especially the mineral elements. Because trees record events in the environment, changes in the amounts of naturally occurring or

human-manufactured chemicals are recorded in the tree-ring record. Trees can provide a temporal record of toxic pollutants, such as cadmium, cesium, sulfur, and aluminum, to better understand the impact of pollution in the air, surface-water, or groundwater supply in ecosystems. The record of such trace elements is obtained by a variety of analytical methods that include neutron activation analysis, proton-induced x-ray emission, and atomic absorption spectrometry. *See* ANALYTICAL CHEMISTRY.

Henri D. Grissino-Mayer

Bibliography. J. Alestalo, Dendrochronological interpretation of geomorphic processes, *Fennia*, 105:1–140, 1971; E. R. Cook and G. C. Jacoby, Potomac River streamflow since 1730 as reconstructed by tree rings, *J. Climate Appl. Meteorol.*, 22(10):1659–1672, 1983; A. E. Douglass, The secret of the Southwest solved by talkative tree rings, *Nat. Geog. Mag.*, 56(6):736–770, 1929; H. C. Fritts, Dendroclimatology and dendroecology, *Quaternary Res.*, 1:419–449, 1971; H. C. Fritts, *Tree Rings and Climate*, 1976; H. C. Fritts and T. W. Swetnam, Dendroecology: A tool for evaluating variations in past and present forest environments, *Adv. Ecol. Res.*, 19:111–188, 1989; G. C. Jacoby and R. D. D'Arrigo, Tree rings, carbon dioxide, and climatic change, *Proc. Nat. Acad. Sciences USA*, 94(16):8350–8353, 1997; F. H. Schweingruber, *Tree Rings: Basics and Applications of Dendrochronology*, 1987; M. A. Stokes and T. L. Smiley, *An Introduction to Tree Ring Dating*, 1968, 1996; T. W. Swetnam, Fire history and climate change in giant sequoia groves, *Science*, 262:885–889, 1993.

Dendrology

The division of forestry concerned with taxonomy of trees and other woody plants. Dendrology, called forest botany in some countries, usually is limited to taxonomy of trees but may also include shrubs and woody vines. This basic subject in the training of foresters teaches how trees are named (nomenclature), described (morphology), and grouped (classification); how to find the name of an unknown tree and recognize important forest species (identification); and where trees occur both by geographic ranges of species and by forest types (distribution).

In forestry a tree is defined as a woody plant with a single erect perennial stem or trunk at least 3 in. (7.5 cm) in diameter at breast height (412 ft or 1.4 m above the ground), a more or less definitely formed crown or foliage, and a height of at least 13 ft (4 m). A shrub is a woody plant, generally lower growing than a tree, and frequently having several slender perennial stems arising at or near the ground.

Common names of trees and lumber in the United States have been standardized by various forest agencies. The scientific name of a tree is in Latin and consists of two parts, a genus and a species; for example, *Pinus ponderosa*, ponderosa pine.

Trees are described and distinguished in botanical terminology largely in terms of their characteristics of form and structure (morphology). The principal parts useful in identification are leaves, flowers, fruits, seeds, buds, twigs, branches, trunk, bark, and wood. See separate articles on trees listed under common names.

Classification. A common but artificial classification groups plants into trees, shrubs, and herbs. However, the 50,000 or more tree species in the world, as well as the numerous other species making up the plant kingdom, are arranged scientifically according to natural relationships as indicated by evolutionary evidence. For example, 10 genera, having a number of characteristics in common, compose the pine family (Pinacea). These and trees of related families make up the gymnosperms (Pinopsida), generally referred to as softwoods by lumber workers. Most tree species, however, are flowering plants, angiosperms (Magnoliophyta), and nearly all of these are dicotyledons (Magnoliopsida), often called hardwoods. Palms and bamboos are monocotyledons (Liliopsida). *See* PLANT ANATOMY; PLANT KINGDOM.

Native tree species in the continental United States total approximately 679 species in 216 genera and 73 families; 69 species are naturalized. However, only about 175 species in 50 genera are commercially important for lumber. About 100 are subtropical and tropical species of southern Florida. Many others are of limited occurrence, small size, or low-quality wood. Hawaii, with its unique tropical flora, has about 300 native tree species, nearly all not occurring elsewhere.

Identification. The correct scientific name of a recorded species may be determined by means of printed keys or manuals or by comparison with a known tree or with mounted specimens in a herbarium. Even in winter, leafless trees usually can be identified from keys by studying the bud, twig, and bark characteristics. Nearly every state publishes a popular, inexpensive, illustrated pocket guide or bulletin for the identification of the trees of that state. Regional floras are covered in other publications. *See* PLANT KEYS.

Distribution. Each tree species has its own natural distribution or range. A few tree species of the United States extend across the continent, while others are local and rare in distribution. About two-thirds of the important forest trees are eastern, or southeastern, whereas one-third are western, some in the Rocky Mountains, and others in the Pacific Coast region. Trees also have an altitudinal distribution and zonation in high mountains. *See* FOREST ECOSYSTEM.

Forest stands of similar composition, appearance, and structure are grouped together into areas characterized by major forest types or formation, and are named from the predominant or characteristic species (see **illus.**). The Society of American Foresters defined 90 forest cover types of eastern

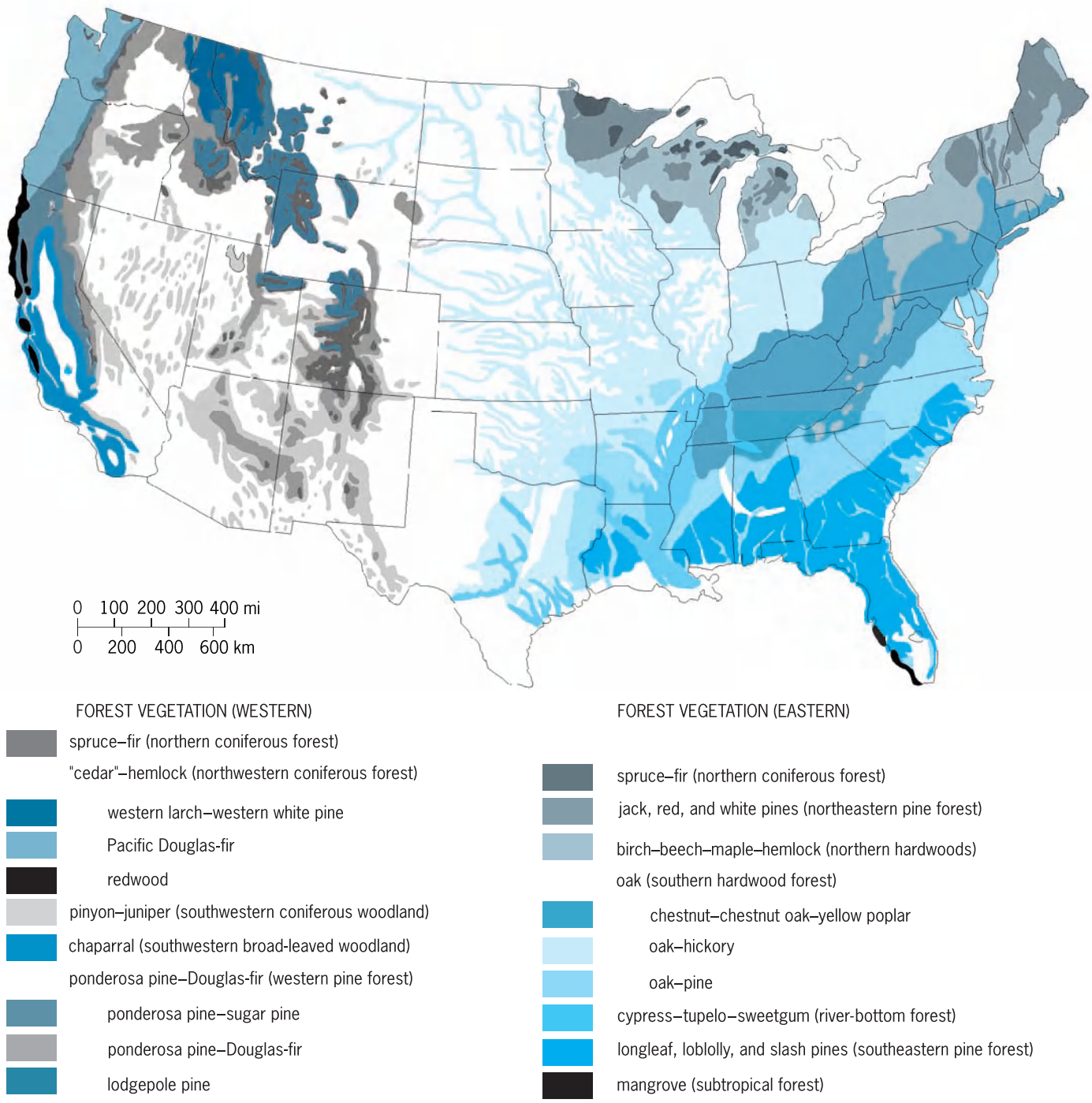

Areas characterized by major forest types in the United States. (*U.S. Forest Service*)

North America (exclusive of Mexico) and 55 of western North America. *See* CHAPARRAL; FOREST AND FORESTRY; FOREST MANAGEMENT; FOREST MANAGEMENT; PLANT TAXONOMY. Elbert L. Little, Jr.

Bibliography. W. M. Harlow et al., *Textbook of Dendrology*, 8th ed., 1995; Society of American Foresters, *Forest Cover Types of the United States and Canada*, 1980; U.S. Forest Service, *Atlas of United States Trees*, 6 vols., 1971–1981; U.S. Forest Service, *Checklist of United States Trees*, 1979.

Density

The mass of an object divided by its volume. Density is conventionally represented by the Greek letter rho: ρ. The most common units for density are g/cm^3 or specific gravity, which is dimensionless. Specific gravity for liquids or solids is usually defined as the density of a material divided by the density of water at 4°C (40°F), which is 1.0000 g/cm^3. Liquid water has a density close to 1 g/cm^3 at any reasonable

TABLE 1. Selected gas densities.

Material*	Temperature, °C (°F)	Density, kg/m^3
Hydrogen	−252 (−422)	1.440
Hydrogen	20 (68)	0.090
Helium	20 (68)	0.179
Water (100°C)	100 (212)	0.588
Methane	20 (68)	0.716
Dry air	20 (68)	1.204
Oxygen	20 (68)	1.428
Sulfur dioxide	20 (68)	2.858
Chloroform	62 (144)	4.341
Xenon	20 (68)	5.857
Sulfur hexafluoride	20 (68)	6.072
Uranium hexafluoride	57 (135)	13.00

*In gaseous state, at a pressure of 1 atm.

temperature [it drops to 0.958 g/cm^3 at 100°C (212°F), or boiling]. In U.S. customary units, at 40°F, a gallon of water weighs 8.345 lb (3.785 kg), while a cubic foot of water weighs 62.43 lb (28.32 kg). In modern SI units, 1 g/cm^3 is frequently expressed as 1000 kg/m^3. There are also a variety of archaic, industry-specific systems in which density is measured in degrees. An example of this is °API, still used in the petroleum industry. Degree measures of density are generally used as quantitative expressions of quality or composition, not as explicit measures of density. *See* HYDROMETER; MASS.

Gases. The density of a typical gas is about one thousandth that of a typical liquid, or about 0.001 g/cm^3. However, the density of gases (**Table 1**) varies over a much wider range than that of liquids and solids. Gas density is dependent on three key factors: the material, the pressure, and the temperature. The dominant factor is pressure.

As pressure approaches zero (a vacuum), the density of any gas approaches zero. As pressure is increased, the density increases in one of two ways. At temperatures below the critical point, the gas will begin to condense into liquid or solid form at sufficiently high pressure. Liquids and solids are always considerably more dense than gases, so condensation is accompanied by a rapid increase in density. At temperatures above the critical point, the gas will simply increase in density indefinitely, without liquefying or solidifying. Each material has its own critical point: for water it is 374°C (705°F), while for nitrogen (N_2) it is 126 K (−147°C or −233°F).

At low pressures, the density ρ of most gases is very well approximated by the ideal gas law, Eq. (1),

$$\rho = \frac{MP}{RT} \tag{1}$$

where M is the molecular weight of the gas, P is the pressure, R is the ideal gas constant, and T is the absolute temperature. This equation is simplest in SI units, where ρ is in kilograms per cubic meter (kg/m^3), M is in grams per mole (g/mol), P is in kilopascals (kPa), T is in kelvins (K), and R has the fixed value 8.3145 joules per mole per kelvin [J/(mol · K)]. For example, the molecular weight of air is about 29 g/mol, so that air at 293.15 K (20°C or 68°F) and atmospheric pressure (101.3 kPa) has a density of 1.2 kg/m^3, or 0.0012 g/cm^3. *See* METRIC SYSTEM; PHYSICAL MEASUREMENT; UNITS OF MEASUREMENT.

At higher pressures, it is necessary to account for interactions between molecules, and this is especially complicated near the critical point, where the densities of gases and liquids are comparable. There are several empirical correlations available for determining the density of nonideal gases; a particularly popular one is the van der Waals equation. The study of gas behavior under high pressure or near-critical conditions is an important part of the science of thermodynamics. *See* CRITICAL PHENOMENA; GAS; INTERMOLECULAR FORCES; VAN DER WAALS EQUATION; VIRIAL EQUATION.

Liquids. Liquid densities are more difficult to predict than those of gases. The density of a typical liquid is 1 g/cm^3, although organic solvents often have lower density (octane: 0.703 g/cm^3 at 20°C or 68°F) while liquid metals generally have much higher densities (mercury: 13.534 g/cm^3 at 25°C or 77°F). The densities of liquids are far more dependent on the particular material than temperature or pressure (**Table 2**). Almost all liquids increase in density with decreasing temperature, but the effect is much more moderate than for gases. Over small ranges of temperature, the density of liquids is well represented by Eq. (2), where β is the thermal expansion coefficient, also known as the coefficient of cubic expansion, and ρ_0 is the density at the temperature T_0. An alternative expansion coefficient, the coefficient of linear expansion, α, is related by $\beta = 3\alpha$. The thermal expansion coefficient for most organic solvents is on the order of $\beta = 0.001\ \mathrm{K}^{-1}$. Over larger ranges of temperature, densities of liquids are more difficult to predict. Almost all liquids solidify at some point as temperature is decreased. The density generally increases significantly at the point of solidification. *See* THERMAL EXPANSION.

$$\rho = \rho_0[1 + \beta(T - T_0)] \tag{2}$$

Water is the most familiar liquid, but it is anomalous in terms of density. Water has a maximum density at 4.0°C (39.2°F). Below that temperature, water

TABLE 2. Selected liquid densities.

Material*	Temperature, °C (°F)	Density, g/cm^3
Hydrogen	−253 (−423)	0.071
Nitrogen	−195 (−319)	0.807
Octane (gasoline)	20 (68)	0.703
Ethanol (ethyl alcohol)	20 (68)	0.789
Water	20 (68)	0.997
Glycerol (glycerin)	20 (68)	1.264
Chloroform	20 (68)	1.484
Perchloroethylene (PCE)	20 (68)	1.623
Lithium	181 (358)	0.534
Gallium	30 (86)	6.095
Lead	328 (622)	10.66
Mercury	20 (68)	13.55

*In liquid state, at a pressure of 1 atm.

TABLE 3. Selected solid densities.

Material*	Temperature, °C (°F)	Density, g/cm^3
Paraffin wax	20 (68)	0.90
Ice (water)	0 (32)	0.92
Sodium chloride (salt)	20 (68)	2.17
Silicon dioxide (quartz)	20 (68)	2.65
Lithium	20 (68)	0.53
Aluminum	20 (68)	2.70
Titanium	20 (68)	4.54
Stainless steel	20 (68)	7.9
Copper	20 (68)	8.96
Lead	20 (68)	11.35
Gold	20 (68)	19.3
Osmium	20 (68)	22.57

*In solid state, at a pressure of 1 atm.

actually decreases in density with decreasing temperature. Furthermore, when water solidifies into ice (at 0°C or 32°F) it decreases in density significantly. At atmospheric pressure and 0°C, the density of liquid water at freezing is 0.9998 g/cm^3 while the density of solid ice is 0.9167 g/cm^3. Most liquids freeze from the bottom; the buoyancy of ice makes frozen lakes and oceans possible without complete freezing. The density of water is often modified by soluble or particulate content. Typical ocean water (15°C or 59°F) has a density of 1.026 g/cm^3, while normal human blood (37°C or 98.6°F) has a density of about 1.060 g/cm^3 (it contains salts, biological materials, and red blood cells). At low concentrations, the change in density (from pure water) is directly proportional to the concentration of soluble or particulate content. *See* LIQUID; SEAWATER; WATER.

Solids. Solid densities (**Table 3**) are similar to those of liquids, except that solids are generally slightly denser, and solid densities are somewhat less dependent on pressure and temperature. The density of nonporous solids is primarily dependent on the particular material, and secondarily on the structure, which may be amorphous or crystalline. Transitions in solid structure are generally accompanied by significant changes in density. For example, diamond and graphite are both forms of elemental carbon; the density of diamond is 3.2–3.5 g/cm^3, that of graphite, 1.9–2.3 g/cm^3. For solids of fixed structure, Eq. (2) may be used over fair ranges of temperature. Typical thermal expansion coefficients are $\beta = 5 \times 10^{-5}\ \mathrm{K}^{-1}$ for copper (over a broad range of temperatures), and $\beta = 2 \times 10^{-4}\ \mathrm{K}^{-1}$ for ice (at 0°C). When thermal expansion coefficients are mismatched between adjacent solids, extreme forces (thermal stresses) can be generated by changes in temperature. This is the reason why small gaps are typically left between different construction materials, and why vinyl siding is usually only loosely fixed to wooden houses. *See* THERMAL STRESS.

Many solids are commonly encountered in granular or porous form (soils, for example). The density of such materials is difficult to define when the granules or pores are too small to be easily distinguished from the bulk material. Many modern materials incorporate features at the molecular scale (zeolites and nanotubes, for example). The density of the nonporous solid is generally significantly different from porous materials infused with air (such as powders and foams) or water (such as slurries and clays). *See* CARBON; NANOSTRUCTURE; ZEOLITE.

The most dense ordinary solids (and liquids) are heavy metal elements and their alloys. The densest elements are the metals osmium and iridium, with densities of 22.6 and 22.4 g/cm^3 (at 20°C), respectively. The least dense metal is lithium, at 0.53 g/cm^3 (at 20°C). The least dense solids are a kind of sol-gels known as aerogels, generated by evaporating off a solvent from a gelatinous liquid solution. One common type of aerogel, based on silica, can have a density as low as 0.003 g/cm^3. The low density of aerogels makes them excellent insulators. *See* GEL.

Extremes. While the most extreme densities found on Earth are air and heavy ores, interstellar space and stars exhibit a far broader range of density. The density of open space is practically zero, but there are measurable amounts of matter even in intergalactic space (the exact nature of such matter is currently the topic of much debate). At extremely high densities, quantum-mechanical effects become very significant. Densities in ordinary stars such as the Sun are not extreme by terrestrial standards, despite the extreme pressures and temperatures. The mean density of the Sun is only 1.4 g/cm^3. However, stars eventually collapse into white dwarfs, neutron stars, or even black holes. The densities of a typical white dwarf and neutron star are 10^6 and 10^{15} g/cm^3, respectively. *See* BLACK HOLE; NEUTRON STAR; SUN; WHITE DWARF.

Measurement. Density measurement devices are commonly referred to as densitometers, and density measurement is sometimes referred to as densitometry. The American Society for Testing and Materials (ASTM) defines a broad range of standard tests for measuring density, which should be used wherever standardization is critical. There are a wide variety of handbooks with extensive data on the densities of common substances and materials.

Most modern density measurement techniques are indirect, as weighing materials tends to be imprecise. However, certain liquids and solids are easily weighed and have well-defined volumes. Liquid density can generally be determined by placing a sample of fluid in a volumetric flask of known volume and weight; the difference between the weight of the flask with and without the liquid divided by the volume gives the density directly. Nonporous solid volumes are easily determined using Archimedes' method: Place the solid in a full container of water or another liquid, and measure the volume displaced (overflow), which is the volume of the solid. The mean solid density follows from the weight. *See* ARCHIMEDES' PRINCIPLE.

Where gas densities cannot be predicted, they are typically measured via momentum. Bernoulli's law states that flowing gas—through a tube, venturi (**Fig. 1**), orifice, or turbine (**Fig. 2**)—generates a pressure difference that is directly proportional to

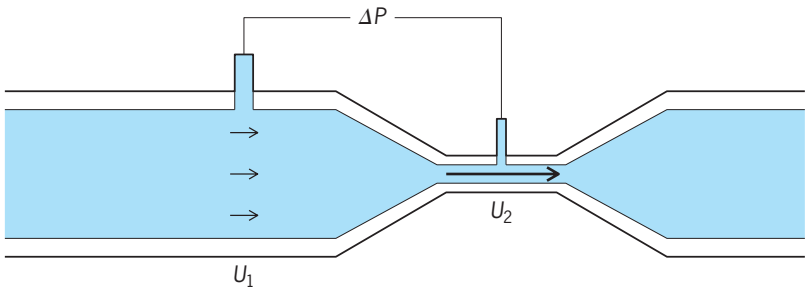

Fig. 1. Venturi flow constriction, used to measure gas or liquid density, showing ports for measurement of the pressure difference ΔP, which is determined by the density and the flow velocities, U_1 and U_2, at the ports.

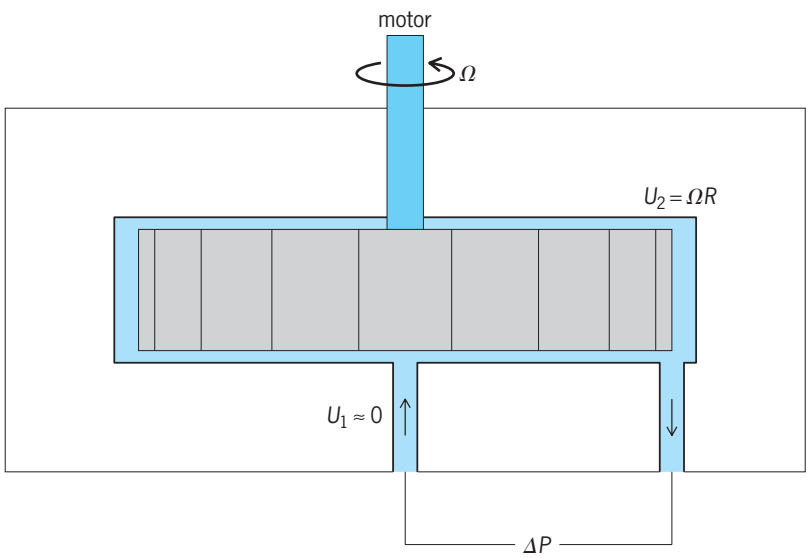

Fig. 2. Turbine driven by motor, used to measure gas density, showing ports for measurement of the pressure difference ΔP. This quantity is determined by the density and the flow velocities, $U_1 \approx 0$ and $U_2 = \Omega R$, at the ports, where R and Ω are the radius and angular velocity of the turbine.

the density of the fluid. It is given by Eq. (3), where

$$\Delta P = \rho \left(U_1^2 - U_2^2\right) = K\rho \tag{3}$$

ΔP is the measured pressure difference, U_1 and U_2 are flow velocities (normally held constant for density measurements), and K is a calibration constant. Devices with turbines are especially convenient as they can act as a measurement device and a pump simultaneously. Gas densities are sometimes expressed as specific gravity with respect to dry air at a specific temperature, as air is a convenient calibration material with a well-defined density. Liquid density can also be measured via pressure differences, but many viscous liquids only follow Bernoulli's law [Eq. (3)] only at prohibitively high flow rates. *See* BERNOULLI'S PRINCIPLE; FLOW MEASUREMENT; TURBINE.

Liquid density can be measured very precisely using buoyancy effects. Objects immersed in liquid behave as if their density has been reduced by that of the liquid: $\rho_{\text{effective}} = \rho - \rho_{\text{liquid}}$. Astronauts use this principle in special pools where they are neutrally buoyant (same density as the pool liquid), which mimics zero gravity well. Hydrometers (**Fig. 3**) are flotation devices in which the level of the float above the liquid surface gives density directly. Other buoyancy-based devices measure forces on a float, where the force is zero when the float and liquid densities are equal. *See* BUOYANCY.

The density of solids and liquids is often measured via absorption of radiation. For materials of identical atomic composition, the absorption rate of x-rays or gamma rays is proportional to the density of the material. One advantage of radiation densitometers is that they can be mounted external to pipes or containers. Disadvantages include the clearances and safety precautions normally required to use radioactive materials in practice. *See* RADIOACTIVITY AND RADIATION APPLICATIONS.

Vibrating densitometers make use of the fact that springs and other mechanisms vibrate with periods proportional to the square-root of mass. For a fixed volume of material in a container of fixed mass, the period of vibration is given by Eq. (4), where P is the

$$\rho = AP^2 - B \tag{4}$$

period, and A and B are calibration constants. There are many types of vibrating densitometers, and most have a relatively broad range of measurement with a high degree of precision. Unfortunately the calibration constants tend to be dependent on temperature and can also drift with time. *See* VIBRATION.

Many materials are simple in form but complicated in composition. Many common materials bond

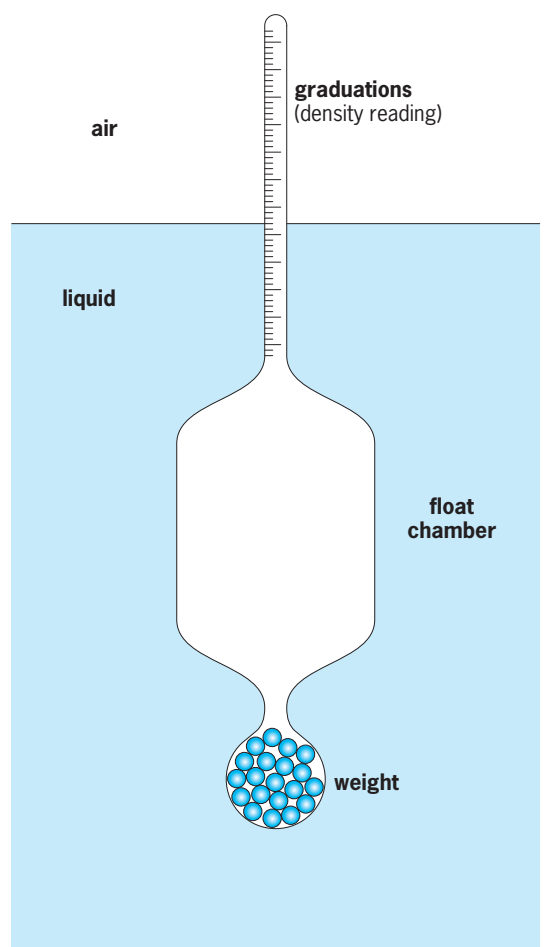

Fig. 3. Hydrometer.

chemically with water, as well as absorbing additional amounts into their structure. Wood, for example, can contain significant amounts of water without appearing wet; some of this is bound chemically while the rest is held physically. The density of wood and other wet materials is generally highly dependent on water content, and "dry" materials are typically defined as materials where all of the physically bound water has been removed (but not the chemically bound water). Porous or contaminated materials pose additional challenges. Often pores must be cleared by washing with water or other solvents and then evaporating off the cleaning material. Cleaning generally changes the density of a material in numerous ways, by dissolving soluble salts or collapsing pores, for example. The cleaning material may also associate chemically or physically with the solid, and must be chosen carefully to avoid contamination. Sol-gels are especially prone to modification; silica-based aerogels collapse when exposed to liquid water or even high humidity. *See* HYDRATE; HYDRATION.

Brian J. Lowry

Bibliography. F. Cardarelli, *Scientific Unit Conversion*, Springer, 1996; D. R. Lide (ed.), *CRC Handbook of Chemistry and Physics*, 87th ed., CRC Press, 2006; B. G. Liptak (ed.), *Instrument Engineers' Handbook*, 4th ed., vol. 1, CRC Press, 2003; M. J. O'Neill et al. (eds.), *The Merck Index*, 13th ed., Merck, 2001.

Density matrix

A matrix which is constructed as the most general statistical description of the states of a many-particle quantum-mechanical system. The state of a quantum system is described by a normalized wave function $\psi(x,t)$ [where x stands for all coordinates of the system, and t for the time], which satisfies the Schrödinger equation (1), where H is the hamilto-

$$H\psi(x,t) = \frac{\hbar}{i}\frac{\partial\psi(x,t)}{\partial t} \tag{1}$$

nian of the system, and $\hbar$ is Planck's constant divided by 2π. Furthermore, $\psi(x,t)$ may be expanded in terms of a complete orthonormal set $\{\varphi(x)\}$, as in Eq. (2). Then, the density matrix is defined by Eq. (3), and this density matrix describes a pure state.

$$\psi(x,t) = \sum_n a_n(t)\varphi_n(x) \tag{2}$$

$$\rho_{mn}(t) = a_n^*(t)a_m(t) \tag{3}$$

Examples of pure states are a beam of polarized electrons and the photons in a coherent beam emitted from a laser. *See* LASER; NONRELATIVISTIC QUANTUM THEORY.

In quantum statistics, one deals with an ensemble of N systems which have the same hamiltonian. If the αth member of the ensemble is in the state ψ^α in Eq. (4), the density matrix is defined as the ensemble average, Eq. (5). In general, this density ma-

$$\psi^\alpha(x,t) = \sum_n a_n^\alpha(t)\varphi_n(x) \tag{4}$$

$$\rho_{mn}(t) = \frac{1}{N}\sum_\alpha [a_n^\alpha(t)]^* a_m^\alpha(t) \tag{5}$$

trix describes a mixed state, for example, a beam of unpolarized electrons or the photons emitted from an incoherent source such as an incandescent lamp. The pure state is a special case of the mixed state when all members of the ensemble are in the same state. *See* STATISTICAL MECHANICS.

Diagonal elements. The diagonal element $\rho_{mn}(t) = |a_n(t)|^2$ of the density matrix for a pure state gives the probability of finding the quantum system in the state $\varphi_n(x)$ at time t. The set $\{\varphi(x)\}$ is complete; property of the probabilities implies that $\Sigma_n\rho_{nn}(t) = \mathrm{Tr}$ unity also for a mixed state. Since all diagonal elements are nonnegative, it also follows that $0 \le \rho_{nn} \le 1$.

Observables. The ensemble average of an observable A (an operator that represents an observable quantity) can be written as Eq. (6). Since $\overline{\overline{A}}$ is the

$$\overline{\overline{A}} = \frac{1}{N}\sum_\alpha\sum_{mn}\left(a_n^\alpha\right)^*\left(a_m^\alpha\right)\int \varphi_n^* A\varphi_m dx$$

$$= \sum_{mn}\rho_{mn}A_{nm} = \mathrm{Tr}\ \left(\rho A\right) \tag{6}$$

trace of a matrix, its value is independent of the choice of the orthonormal set $\{\varphi(x)\}$.

Pure and mixed states. For a pure state, Eq. (7)

$$(\rho^2)_{mn} = \sum_p a_p^* a_m a_n^* a_p = a_n^* a_m = \rho_{mn} \tag{7}$$

holds, or, equivalently, $\rho^2 = \rho$ or $\mathrm{Tr}\ \rho^2 = 1$. But, for a mixed state, $\mathrm{Tr}\ \rho^2 < 1$ always holds. The necessary and sufficient condition for a state to be a pure state is $\rho^2 = \rho$.

Eigenvalues and eigenvectors. The density matrix is hermitian, $\rho^*_{mn} = \rho_{nm}$, and is diagonalizable. The eigenvalues of ρ for a pure state are all zero except one, which has the value unity. The eigenvector corresponding to the nonzero eigenvalue is the pure state the system is in. For a mixed state, the eigenvectors are the possible pure states the system may have, and the eigenvalues are the probabilities of these states. Thus all eigenvalues lie between 0 and 1. The extreme case of a mixed state is the random state for which all N members of the ensemble are in different states. All eigenvalues corresponding to the occupied states have the value $1/N$. *See* MATRIX THEORY.

Liouville equation. The density matrix satisfies the Liouville equation (8), where H is the hamil-

$$i\hbar\frac{\partial\rho}{\partial t} = [H,\rho] \tag{8}$$

tonian of the system. In a state of equilibrium $\partial\rho/\partial t = 0$. Then, Eq. (8) implies that ρ is a function of the hamiltonian H alone. For an ensemble in thermal equilibrium, ρ is related to temperature

T according to Eq. (9), where k is the Boltzmann

$$\rho = \frac{e^{-H/kT}}{\mathrm{Tr}\, e^{-H/kT}} \tag{9}$$

constant. This result, together with Eq. (6), allows a complete determination of the equilibrium thermodynamic properties of the system. For example, the entropy S is proportional to the expectation value of $\ln \rho$ as given in Eq. (10).

$$S = -k\, \overline{\overline{\ln \rho}} = -k\, \mathrm{Tr}\, (\rho\ \ln\ \rho) \tag{10}$$

The relationship in Eq. (10) between the entropy and the density matrix holds for nonequilibrium systems as well. On the basis of knowledge of the eigenvalues of the density matrix, it can be deduced that the entropy of a pure state is zero, that of a mixed state is nonnegative, and that of a random state has the highest possible value of $k \ln N$ for an ensemble of N systems. Hence, entropy is a measure of the randomness of the state of an ensemble. *See* ENTROPY.

Effect of measurement. Another consequence of the Liouville equation, Eq. (8), is that, for an isolated system whose hamiltonian is time-independent, a pure state cannot evolve into a mixed state, or vice versa, as time goes on. The act of a physical measurement, however, can interrupt the time development of a system so that a pure state can be turned into a mixed state. Consider a Stern-Gerlach experiment which employs a nonuniform magnetic field to separate a beam of electrons into two beams which have spins parallel and antiparallel to the field. The direction of the field is chosen as the z axis. The incoming electrons are assumed to be polarized along an axis different from the z axis. The wave function of the incoming beam is then $\psi = a\psi_p + b\psi_-$, where ψ_p, ψ_- are the wave functions of the spin-up and spin-down states and $|a|^2 + |b|^2 = 1$. The density matrix is shown in Eq. (11). This matrix satisfies the requirement

$$\rho = \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix} \tag{11}$$

$\rho^2 = \rho$ for a pure state. After passing through the Stern-Gerlach apparatus, the density matrix becomes Eq. (12), which represents a mixed state. The measurement has resulted in an increase in the entropy or randomness of the electron beam.

$$\rho' = \begin{pmatrix} |a|^2 & 0 \\ 0 & |b|^2 \end{pmatrix} \tag{12}$$

Perturbations. The Liouville equation is also very useful in determining the effect of a perturbation on the density matrix and, hence, on the properties of the system. Let the unperturbed hamiltonian be H_0 and the perturbing hamiltonian be H_1. The density matrix is similarly written as the sum of the unperturbed part ρ_0 and a correction ρ_1. Furthermore, it is assumed that the system is in equilibrium state before the perturbation is applied. Then Eq. (8) takes the form of Eq. (13). If the perturbation is weak, the last term of Eq.(13) can be ignored because it is of the second order of the perturbation. When this is done, the correction term can be solved to obtain Eq. (14). The Heisenberg operator $H_1(t - t')$ is defined in Eq. (15). The lower limit of the integration

$$i\hbar \frac{\partial \rho_1}{\partial t} = [H_0, \rho_1] + [H_1, \rho_0] + [H_1, \rho_1] \tag{13}$$

$$\rho_1 = -\frac{i}{\hbar} \int_{t_0}^{t_1} [H_1(t - t'),\ \rho_0]\, dt' \tag{14}$$

$$H_1(t - t') = \exp\left[\frac{i}{\hbar} H_0(t - t')\right] H_1 \cdot \exp\left[-\frac{1}{\hbar} H_1(t - t')\right] \tag{15}$$

in Eq. (14) is the time at which the perturbation is turned on. *See* PERTURBATION (QUANTUM MECHANICS).

From this result the first-order correction to the expectation value of an observable A can be calculated as in Eq. (16). Notice that ρ_1 involves only the perturbation and the equilibrium density matrix. Thus, the first-order correction, called the linear response of the system to the perturbation, can be completely calculated from the knowledge of the equilibrium state. This is the starting point of the modern theory of transport phenomena. *See* KINETIC THEORY OF MATTER.

$$\overline{\overline{A}}_1 = \mathrm{Tr}\, (\rho_1 A) \tag{16}$$

S. H. Liu

Bibliography. K. Blum, *Density Matrix Theory and Applications*, 1981; N. N. Bogoliubov and N. N. Bogoliubov, Jr., *An Introduction to Quantum Statistical Mechanics*, 1994; K. Gottfried, *Quantum Mechanics*, 1966, reprint 2000; L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics*, 1990; L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory*, 3d ed., 1977, paper 1997; L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3d ed., 1980; G. H. Wannier, *Statistical Physics*, 1966, reprint 1987.

Dental anthropology

The scientific study of people, with their living and extinct primate relatives, using the evidence of teeth. Dental anthropologists include not only those trained in anthropology but also practicing dentists, anatomists, radiologists, forensic scientists, biochemists and geneticists, archeologists, paleontologists, and zoologists.

Dentition. The deciduous (milk) teeth and permanent (adult) teeth make up the dentition. They have their own distinctive patterns of development, physiology, size and shape, disease, and changes with age, all very different from the bones of the skeleton. In particular, the three tissues of the teeth—enamel, dentine, and cement—do not turn over continuously as do bone and most other tissues of the body. Once formed in childhood, the main structure of a tooth changes very little throughout adult life, although it is modified by wear or disease, and there is slow continued deposition of dentine within the pulp chamber, and of cement on the root surface. This means

that, even in adults, tooth form is more clearly related to the gene-mediated control of development than a structure such as a bone, which is modified throughout life in response to changing physiological factors and mechanical loadings. Many features of the tooth crown have a high heritability. Before modern molecular genetics, dental morphology was proposed as a simple method for establishing the relationship of twins. It has become possible to demonstrate the activity of particular genes in the differentiation of tooth form. Another effect of this conservation of original tooth structure is that the layered pattern of growth in dental tissues is preserved throughout life, so that childhood development can be studied even in adults. For archeologists and paleoanthropologists, there is a further advantage in the study of teeth—they are exposed in the mouth. Their adaptation to this harsh, abrasive environment in life ensures that they are the toughest part of the body, and survive relatively unchanged in a wide variety of burial environments. The majority of fossils are teeth and jaws. In addition, teeth were in contact with all the food eaten, and were used not only for chewing but for initial food preparation, grooming, acting as a third hand, and so on. Therefore, teeth bear many traces of the diet and life-style of their owners. Furthermore, teeth are easily studied in living people and animals by direct observation, by taking dental impressions (casts), or by examining teeth extracted in dental surgeries. Direct comparisons can be made far more easily than with bony fossils.

Morphology. For many dental anthropologists, the core of the subject is tooth morphology. Identification of different primate species is most readily done from tooth form in archeological and fossil material. One feature of human evolution has been a reduction in tooth crown size. A rapid dental reduction has taken place over the last 50,000 years in many parts of the world. This may be part of a general change in body proportions or may relate to particular dietary or technological changes. Among the higher primates as a whole, males tend to have larger teeth than females, particularly the canine teeth. Even modern humans show this to a small extent, and it helps to distinguish between the sexes in archeological and forensic remains. It does, however, complicate the definition of species from the fossils of extinct primates. The cusps, ridges, and furrows that decorate the crown surface also vary within different species of primates, together with the number and form of tooth roots. The different forms have characteristic frequencies for different populations of living people, and are used to investigate their relationships and history in comparison with archeological groups. Teeth additionally vary in their placement and in the way in which they fit together—the dental occlusion. Occlusal anomalies have become more common with the adoption of modern westernized life-styles and diet.

Development. The timing of formation has been established for the different types of teeth in the dentition from x-ray studies of living children, and of the young of several species of other primates. The dentition develops in a well-defined series of stages, so the state of dental development at death is an important guide to a child's age in archeological and forensic cases. It is more difficult to apply the concept to extinct primates, so attempts have been made to calibrate their development sequence by counting regular microscopic structures developing during the formation of enamel. This technique makes it possible to produce a development chronology for the dentition of one individual, providing such detail that episodes in which growth is disturbed (caused by factors such as childhood fevers or dietary deficiencies) are recorded as defects in the sequence. Estimation of age in the remains of adults is much more difficult, but teeth still provide important methods for forensic and archeological studies. Gradual microscopic changes take place in the dentine and cement, and the tooth crowns are progressively worn. Wear rate was much greater among people in the past than in most modern societies. Not only did the food involve more energetic chewing and contain more abrasive particles, but the teeth were used for a variety of other tasks. These factors leave their mark in a variety of anomalous wear patterns. At a microscopic level, the pattern of wear scratches on the enamel (microwear) has been used to compare the diet of different living primate species and to reconstruct the diet of extinct species. The nature of the diet also leaves its mark in the epidemiology of dental diseases. There are differences, for example, among the great apes in the frequency of dental caries (decay) in relation to the proportion of sugar-containing fruit in their diet. Among humans, hunter-gatherer groups typically had very low caries rates and, for example, there was a clear increase associated with the adoption of maize agriculture by native Americans. The much higher caries rates of the last two centuries are associated with high sugar consumption. Associated dental treatment leaves its mark, and one of the most important methods for personal identification in forensic cases is through the matching of this evidence with dental records. *See* DENTAL CARIES; DENTISTRY; TOOTH.

Simon Hillson

Bibliography. S. W. Hillson, *Dental Anthropology*, Cambridge University Press, 1996; M. A. Kelley and C. S. Larsen (eds.), *Advances in Dental Anthropology*, Wiley-Liss, 1991; G. R. Scott and C. G. Turner II, *The Anthropology of Modern Human Teeth: Dental Morphology and Its Variation in Recent Human Populations*, Cambridge Studies in Biological Anthropology, Cambridge University Press, 1997; D. R. Swindler, *Dentition of Living Primates*, Academic Press, 1976.

Dental caries

A disease in which the mineralized tissues of the tooth undergo progressive destruction from the outside surface of the tooth. It is caused by bacteria that colonize the tooth surface and, under certain conditions, produce sufficient acids to demineralize the enamel covering of the tooth crown or the cementum covering the root, and then the underlying dentin. As the destruction of the dentin progresses,

along with breakdown of the organic components, the bacteria invade the dead tissue and enter the pulp chamber. The pulpal tissue becomes infected and the typical toothache may ensue. The infection can ultimately destroy the pulpal tissue and extend out through the apical openings of the roots and into the surrounding periodontal tissues.

Anatomy. The carious lesion begins on the surface of the enamel of the tooth crown. If the gum tissue has receded, exposing the surface of the root, the lesion may begin on the cementum surface of the exposed root. Since the bacteria that cause caries reside in dental plaque that forms on the tooth surface, sites that retain plaque are at the greatest risk for the development of carious lesions. Thus, dental caries most often develops in the pits and grooves or fissures of the occlusal (biting) surfaces of the premolars and molars, on the smooth surfaces of the crown just below the contact areas of the adjacent surfaces, and on the smooth surfaces of the facial surfaces (facing lips or cheeks) or lingual surfaces (facing the tongue) adjacent to the margin of the gum tissue. The progression of the carious lesion through dental enamel follows the direction of the enamel rods that extend from the surface of the crown to the dentinoenamel junction, that is, the junction of dentin and enamel. As a result, carious lesions originating in the occlusal pits and fissures of posterior tooth crowns have very small surface defects that expand in size as they progress toward the dentinoenamel junction. Those originating in the enamel of the smooth surfaces of the tooth are broadest at the surface and decrease in breadth as they approach the dentinoenamel junction. All carious lesions expand laterally at the dentinoenamel junction and progress within the dentin as a conical lesion with its apex directed inward.

One of the most important anatomical features of caries is the very early development of an enamel lesion, also referred to as the incipient lesion. As acid is produced in dental plaque next to the surface of the enamel, it penetrates the enamel between the crystals of mineral and begins to dissolve these crystals, liberating mainly calcium and phosphate ions. As these ions diffuse out of the enamel, the mineral reprecipitates in the surface layer because of the high calcium and phosphate content of the dental plaque overlying the area. This complex demineralization-remineralization process leads to the development of a lesion that is demineralized, but covered superficially by a relatively intact surface zone. Since the crystals of mineral within the partially demineralized subsurface zone are not totally dissolved, this lesion can remineralize in the absence of further acid attack. Hence, the early enamel lesion is regarded as a reversible condition. Trace amounts of the fluoride ion (tenths of a part per million) are known to catalyze the remineralization of incipient lesions, so it is now believed that this is one of a number of ways in which fluoride prevents dental caries. If the lesion progresses beyond the incipient stage and there has been cavitation with total loss of mineral crystals, the lesion is no longer reversible and the tooth must be restored with filling material such as plastic, amalgam, or gold.

Etiology and pathogenesis. The sites of caries development have been correlated with the presence of dental plaque containing one of a number of cariogenic streptococci referred to as mutans streptococci—notably in humans, *Streptococcus mutans* and *S. sobrinus*. These microorganisms are cariogenic because of their ability to generate a considerable amount of acid as a result of their metabolism of carbohydrates, and to survive in an acid environment. The acid environment is only slowly neutralized since its dilution and neutralization is hindered by the fact that the acid is within the dental plaque and thereby shielded from the saliva. Each time a carbohydrate-rich substance is ingested, acid is formed. Therefore, the frequency of ingestion and physical consistency of fermentable carbohydrates are important factors in caries formation. *See* STREPTOCOCCUS.

The progression of development of a carious lesion depends on the amount of time that a site on a tooth is subjected to acidic conditions (low pH) relative to the amount of time that the site is at neutral pH. If demineralization exceeds remineralization, the carious site progresses from the incipient lesion stage to the stage of cavitation where the surface layer of enamel is lost because of excessive mineral destruction. From this point, the lesion can increase in size and involve the underlying dentin, and finally the pulpal tissue. As the carious lesion approaches the pulpal tissue, bacterial toxins and other products gain access to this tissue and elicit an inflammatory reaction which intensifies as the bacteria themselves enter and infect the pulp. Inflammation of the pulp (pulpitis) is usually followed by death of this tissue (pulpal necrosis) and in most cases the infection extends through the apical foramina of the roots and into the adjacent periodontal ligament space.

Prevention. Dental caries can be prevented by making the tooth less susceptible to acid attack, removing cariogenic bacteria from the teeth, and limiting ingestion of cariogenic substrate.

The most effective and least expensive means of rendering the tooth less susceptible to the caries attack is through fluoridation, either by systemic means such as water fluoridation, or by topical application of fluoride to the tooth surfaces. Water fluoridation (approximately 1 ppm) by itself reduces the incidence of caries by 50–60%. Topical fluoridation can be instituted by many means, ranging from professional applications in the dental office to self-applied fluoride in the form of toothpastes, mouth rinses, and gels. The mode of action of fluoridation is not fully understood, but the benefit to the tooth structure is probably a result of a number of actions. It is believed that fluoride, when incorporated into the mineral phase of enamel and the other hard tissues, results in a better crystalline structure that is less susceptible to acid dissolution. Alternatively, the anticaries effect of fluoride may be due to fluoride's ability to inhibit demineralization and promote remineralization. In addition to fluoridation,

means of improving the tooth's resistance to dental caries include the application of plastic pit and fissure sealants which eliminate retentive areas on the biting surfaces of the posterior teeth that are more susceptible to carious lesions.

Cariogenic bacteria in the dental plaque can be reduced or eliminated by physical means such as daily toothbrushing and flossing. Although antibiotics are known to be effective against dental caries, the use of agents such as penicillin cannot be justified because of the danger of causing hypersensitivity to the antibiotic, or promoting the development of antibiotic-resistant strains that could be much more dangerous than the normal ones found in the mouth. The antiseptic chlorhexidine has been shown to be effective in reducing dental plaque and caries in children when used in the form of a mouth rinse.

Since cariogenic bacteria cannot form acid in the absence of fermentable carbohydrates, the risk of developing caries can be diminished significantly by controlling the dietary intake of carbohydrate, mainly sucrose. Not only is sucrose an excellent substrate with which plaque microorganisms form acid by-products, but it also promotes the accumulation of mutans streptococci on the teeth. Acid formation continues as long as there is fermentable carbohydrate available to the bacteria. *See* MICROBIOLOGY; TOOTH. Murray R. Robinovitch

Bibliography. W. H. Bowen and L. A. Tabak, *Cariology for the Nineties,* 1993; J. R. Mellberg and L. W. Ripa, *Fluoride in Preventive Dentistry: Theory and Clinical Applications,* 1983; E. Newbrun, *Cariology,* 3d ed., 1989; J. Slots and M. A. Taubman, *Contemporary Oral Microbiology and Immunology,* 1992.

Dentistry

The branch of biomedical science concerned with the embryology, development, and structure of the teeth, jaws, oral cavity, and adjacent structures, and with the prevention, diagnosis, and treatment of their abnormalities.

The practice of general dentistry requires a thorough knowledge of not only the structure, growth, function, and pathologic conditions of the oral cavity and related structures, but also the relationship of the oral cavity to other parts of the body in health and disease, and a recognition of oral manifestations of various systemic diseases. It is highly demanding because of the wide spectrum of diseases and abnormalities encountered. Nevertheless, most of the dentist's time and effort are devoted to the treatment of dental caries (decay) and the periodontal diseases. In the United States the prevalence of dental caries, especially in children and young adults, has decreased rapidly; a similar trend is occurring with gingivitis (inflammation of the mucous membrane surrounding the tooth sockets) and periodontitis (inflammation of the tissues surrounding the teeth). As a consequence, other conditions are becoming relatively more important in general dentistry. These include orofacial and temporomandibular joint pain; oral mucosal lesions associated with immunosuppresion, irradiation, and chemotherapy in patients with cancer or organ transplants; and unusual mucosal and periodontal lesions in patients with acquired immune deficiency syndrome (AIDS).

Restorative dentistry attempts to preserve and restore decayed, defective, missing, and traumatized teeth. In addition to biological and medical science, restorative dentistry calls upon engineering, metallurgy, and polymer chemistry, coupled with a high degree of manual dexterity. Restorative treatment can result in improved chewing, better appearance, and correction of some speech abnormalities; it can also alleviate joint and muscle pain. Advances include the development of composite resins strong enough to withstand biting and chewing; some of these can be bonded directly to the teeth.

The treatment of gingivitis and periodontitis requires the removal of bacterial deposits (calculus and plaque) from the crowns and roots of affected teeth (known as scaling and root planing) and adoption of hygienic practices by the patient. In some cases, especially in moderate to severe periodontitis, surgery is needed to reach diseased root surfaces or to eliminate periodontal pockets. Special care is required by patients with joint or cardiac valve replacements or vascular grafts or with a history of such systemic diseases as rheumatic fever, diabetes mellitus, hepatitis, or AIDS. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); CANCER (MEDICINE); DIABETES; HEPATITIS; RHEUMATIC FEVER.

Specialization

The complexity of modern methods of diagnosis and treatment has led to the recognition of eight branches of specialization by the American Dental Association: oral and maxillofacial surgery, endodontics, orthodontics, pedodontics, periodontics, prosthodontics, oral pathology, and public health dentistry. Subspecialties include oral medicine, dental radiology, oral diagnosis, and gerodontology.

Oral and maxillofacial surgery. This branch of dentistry treats diseases and abnormalities of the maxillofacial region by surgical means. Oral surgeons treat a wide variety of problems by removing teeth, cysts, tumors, and growths; reducing bone fractures; and correcting congenital anomalies and malformations of the maxillofacial region. Arthroscopy using an endoscope to examine the interior of a joint to diagnose for diseases of the temporomandibular joint, and replacement of missing teeth with implants are areas of importance.

Endodontics. Teeth contain a pulp chamber occupied by nerves, blood vessels, and connective tissue. The tooth pulp can die as a result of infected carious lesions or of trauma. Endodontists treat teeth that are nonvital or made so during restoration. With files and reamers, the pulp tissue is removed and root canals are enlarged, sterilized, and filled, usually with guttapercha (a heat-sensitive plastic substance). Following root canal therapy, the tooth dentin is without sensitivity, although the cementum (calcified tissue)

covering the tooth surface and attaching the tooth to the alveolar bone remains healthy.

Orthodontics. Orthodontists deal with with abnormalities in tooth and jaw relationships that result in facial distortion and malfunction. Orthodontic treatment attempts to establish normal closing of the mouth and facial harmony. The teeth are positioned and the jaws modified by fixed or removable appliances. Some habits such as tongue thrust must be corrected. The teeth are held in the new position during modification of the bone and muscles. Successful treatment results in normal shape and expression of mouth and lips, better enunciation, and good mastication. A major advance has been the use of agents that directly attach appliances to the teeth, in contrast to metal bands. The subspecialty surgical orthodontics is concerned with correcting the position and relationship of teeth and jaws by surgery.

Pedodontics. Pedodontics, the practice of dentistry in children, requires specialized knowledge and skills. The deciduous or primary teeth are small and differ in shape from adult teeth; special procedures are required for their treatment, which is made difficult by the immaturity of the patient. Caries are extremely prevalent in some groups of children, and premature tooth loss has serious effects on the permanent dentition. Harmful habits in children include thumbsucking, which displaces the teeth, and the consumption of large quantities of sugary foods, which causes decay. Caries and periodontal disease can be prevented by tooth cleaning, in which it is important to train children. In addition to the usual dental education, pedodontists are trained in behavioral science, hospital dentistry, anesthesiology, and human growth and development.

Periodontics. This branch of dentistry is devoted to diseases of the gingiva (gum tissue), alveolar bone, periodontal ligament, and cementum. Periodontal diseases include gingivitis, periodontitis (sometimes called pyorrhea), primary and secondary occlusal traumatism, gingival hyperplasia, and periodontal atrophy. The first two are the most important. Both are caused by bacteria at the junction between the tooth and the gingival tissue. Several types of anaerobic microorganisms are associated with periodontal destruction, including serotypes and species of *Actinobacillus*, *Bacteroides*, *Eikenella*, *Wolinella*, and *Fusobacterium*. Gingivitis is diagnosed when the inflammation is confined to the gingiva; in periodontitis the inflammation extends into the bone and periodontal ligament. In chronic periodontitis, pockets form between the tooth surface and the periodontal tissues, the periodontal ligament and alveolar bone are destroyed, an abscess may form, and the tooth can be lost.

The objectives of periodontal therapy are to eliminate the infection and to prevent its recurrence. Therapy consists initially of removal of microbial deposits (calculus and plaque) from the surfaces of the affected teeth by scaling and root planing with or without opening flaps, and adoption of meticulous oral hygiene. In cases of early-onset aggressive periodontitis, antibiotics may be used in addition to scaling and root planing. Nonsteroidal anti-inflammatory drugs may be useful adjuncts. In some cases, surgery may be needed to gain access to diseased tooth roots and eliminate periodontal pockets to prevent recurrent infection. Periodontics includes implantology. *See* PERIODONTAL DISEASE.

Prosthodontics. Prosthodontics is the branch of dentistry devoted to artificial substitutes for oral structures. The replacement of teeth and other oral structures may be necessitated by abnormalities, loss of teeth from disease or trauma at birth, and destruction of teeth or jaws or other parts of the mouth by surgical treatment of neoplasms or trauma. Prosthodontics has two subdivisions: removable prosthodontics, in which missing teeth are replaced by dentures; and fixed prosthodontics, in which teeth are replaced by bridges. Bridges serve when healthy teeth or implants are present at the ends of an edentulous (toothless) space. The support of bridges by implants is an important technique. In addition, prosthodontists use splints to connect teeth that have been loosened by periodontal disease. They also construct devices that modify the relationships of the biting surfaces of the teeth to alleviate facial and joint pain.

Oral pathology. Oral pathology is concerned with the diseases of the teeth, oral cavity, and jaws, and with the oral manifestations of systemic diseases. Oral pathologists correlate clinical and radiographic features with the history of the tissue changes associated with disease to reach a diagnosis and in some cases to suggest treatment. *See* PATHOLOGY.

Pubic health dentistry. This is the science of promoting dental health through community effort. It comprises research, education, prevention, diagnosis, prescription, and the evaluation of community dental care. Public health dentists must have skill in public and human relations, a knowledge of administrative practice, and the ability to determine the needs of the community for dental care facilities and to assist in their design, development, administration, and support. They must assist in the development of dental health education programs, and possess a broad understanding of the social and behavioral sciences as these affect dental disease in the community. In many communities they have sponsored the fluoridation of water systems for controlling caries. *See* MOUTH DISORDERS; TOOTH DISORDERS.

Roy C. Page

Dental Materials

Dental materials is an interdisciplinary area that applies biology, chemistry, and physics to the development, understanding, and evaluation of materials used in the practice of dentistry. It is principally involved in restorative dentistry, prosthodontics, pedodontics, and orthodontics, and to a smaller extent in the other areas of practice. The field of dental materials has advanced rapidly with the application of technologies such as genetic engineering of filling materials, computer designing and machining of restorations, adhesive bonding to teeth, dental implants, and new types of restorative materials.

Restorative materials. The restoration of missing tooth structure is one of the most challenging areas

for the use of dental materials. The enamel that covers the crown of the tooth is the hardest substance in the human body and is very resistant to physical and chemical attack. Direct restorations are put in place and harden by a setting reaction, whereas indirect restorations require a replica of the tooth which has to be fabricated.

Gold. Pure gold or gold foil has long been used as a restorative material. Pellets of pure gold are formed from very thin gold sheets or foils. The gold is compacted with compressed air or manually, and contoured and polished to the desired shape. These restorations are long-lasting and biologically inert. They are placed in low-stress areas since gold is soft.

Amalgam. Dental amalgam is the most frequently used direct restorative material for posterior (molar and premolar) teeth. It is formed by mixing mercury with an alloy powder of silver, tin, and copper. After mixing, there is time to insert and shape the material before it hardens. Silver–copper compounds and elements such as indium improve the alloy. Its advantages are ease of application, durability, and cost; the disadvantages are chiefly esthetic. *See* AMALGAM.

Resins. Composite resins are tooth-colored filling materials that were first developed for restoring anterior teeth (incisors and canines) but have found increased use as an attractive alternative to dental amalgam in posterior teeth. A composite resin consists of an organic polymer matrix reinforced with up to 80% of an inorganic filler such as silica or glass. Composite resin from a single paste is placed on the tooth and then polymerized with the help of a fiber-optic visible light source. This allows more time for contouring and gives a restoration with improved color stability. Composite resin restorations are strengthened by etching the enamel with 37% orthophosphoric acid prior to placement of the resin in order to create a strong attachment to the enamel.

Glass ionomer. The only dental restorative material that forms a durable chemical bond to dentin is formed by the reaction of aluminosilicate glass with polyacrylic acid. Glass ionomer is particularly useful for restoring eroded or carious areas on exposed root surfaces with little or no enamel for bonding with composite resin. This material slowly releases fluoride to provide some additional protection from new decay. It is brittle and not used in areas of heavy stress.

Metal castings. Crowns and inlays are made of metal castings and used when there is insufficient tooth structure to support filling. A negative of the tooth is prepared with synthetic rubber or agar, into which high-strength gypsum (plaster) is poured. This forms the replica of the tooth, to which wax is applied. An indirect lost-wax process is used in fabricating dental castings. Alloy ingots used for the castings cover a range of compositions from 80% of noble metals (gold and the platinum group) combined with copper and other minor elements to base metal alloys of chrome, cobalt, and nickel. Traditionally, only noble alloys were used, but as the cost of gold increased, other alloys were developed. For making the impression and attaching the casting and adjusting the bite, at least two sessions are required. *See* METAL CASTING.

Porcelain. Porcelain jacket crowns and metal crowns to which porcelain is bonded are used in visible areas. In one type an inner core of high-strength aluminous (40–50% alumina) porcelain is covered by feldspathic porcelain. The latter is applied to a refractory gypsum model covered with platinum foil and sintered under a vacuum at a high temperature. The porcelain powder contains metal oxides to color the crown. The other type is made from a glass ceramic of silica that is cast by the lost-wax process. After casting and contouring, the crown is heat-treated to form an opaque crystalline material, mica. Finally, color is applied by firing oxide stains on the surface. Porcelain and cast ceramics are used for a cosmetic veneer of the front surface of anterior teeth. The thin veneers are bonded to etched enamel with a composite resin cement. The third type of crown has a cast metal substructure for strength, onto which feldspathic porcelain is bonded by the same process described above. These crowns resist chewing most successfully, whereas the jacket crowns are considered more attractive because they resemble enamel in their translucency.

Provisional restoration. For provisional restorations different materials are frequently used. While a permanent crown is being fabricated, a temporary or provisional crown protects the tooth. For best appearance, a custom-formed poly(methyl methacrylate) or polycarbonate crown is lined with acrylic resin. Aluminum crowns lined with acrylic resin can be used with posterior teeth along with the temporary crowns. These crowns are easily removed, since a cement is used that sets when pastes containing the active ingredients zinc oxide and eugenol are combined. A zinc oxide–eugenol cement reinforced with poly(methyl methacrylate) beads is also used as a temporary direct filling material for patients with deep decay. *See* POLYACRYLONITRILE RESINS.

Prosthodontic dental materials. The materials used with fixed prosthodontics are similar to those described above under "Metal castings" and "Porcelain." For small areas where teeth have been lost, a fixed partial denture (bridge) restores the missing teeth. This typically consists of crowns on the teeth next to the toothless space, rigidly connected with a casting suspended over the gap. Given multiple toothless areas on the same arch or long spaces, a removable partial denture can be placed that consists of a thin chrome–cobalt framework that rests on existing teeth. Highly cross-linked poly (methyl methacrylate) resin, pigmented to match the gingival tissue, is attached to the framework in the toothless areas for additional support of the artificial teeth and for appearance. Plastic, composite resin, or porcelain denture teeth are bonded to the resin base to recreate the missing occlusion. In the case of a completely toothless area, a complete removable denture is inserted, consisting of the acrylic resin base and teeth mentioned above. There has been increasing use of carefully placed titanium implants into bone where teeth have been lost. An individual crown or a partial or complete denture is placed on the implants

projecting through the gingival tissue. The crown or denture is secured with screws so that the artificial teeth can be removed for cleaning, inspection, and maintenance.

Orthodontic materials. The movement and stabilization of teeth can be controlled by several unique dental materials. Stainless steel alloys are commonly used for brackets and bands on teeth, but plastic brackets are also being used to improve appearance. Brackets are bonded to etched enamel with the help of composite resin without inorganic filler. Bands encircling posterior teeth are made of stainless steel and are generally secured with a cement formed by combining zinc oxide and phosphoric acid. Wires in different configurations are adapted to the dental arch to create the desired static forces for controlled tooth movement. The wires are commonly formed from wrought stainless steel, but alloys of nickel–titanium, titanium–molybdenum, and cobalt–chromium–nickel are also used. At one time, thin stainless steel wires were used to secure the archwire to the brackets, but today rubber elastic bands are preferred. Removable appliances made of poly(methyl methacrylate) are frequently used to move or stabilize individual teeth, or expand the dental arch. Stainless steel wires protruding from the acrylic generate active or passive forces on the teeth.

G. H. Johnson

Bibliography. R. G. Craig, *Dental Materials*, 7th ed., 1999; R. G. Craig, *Restorative Dental Materials*, 9th ed., 1993; *Dentists' Desk Reference*: *Materials, Instruments, and Equipment*, 2d ed., 1983; P. H. Jacobsen (ed.), *Conservative Dentistry*: *An Integrated Approach*, 1990; R. W. Phillips, *Skinner's Science of Dental Materials*, 9th ed., 1992; S. W. Redding and M. T. Montgomery (eds.), *Dentistry in Systemic Disease*: *Diagnostic and Therapeutic Approach to Patient Management*, 1990.

Dentition

The arrangement, type, and number of teeth which are variously located in the oral or the pharyngeal cavities, or in both. Teeth are found in areas where there is an underlying supporting structure of cartilage or bone and where stomodeal ectoderm is present. The bones with which these structures are usually associated are the mandible (both the cartilaginous bones and the dentary), the premaxillaries, and the maxillaries. However, in certain vertebrates, the vomerine, palatine, parasphenoid, and pterygoid bones may be involved; and in many fishes, teeth are also found on the branchial arches.

Attachment and replacement. Attachment of teeth is variable among vertebrates. They may be inserted in sockets in the jawbones (thecodont condition), fused to the edge of the bone proper (acrodont condition), or attached to the inner surface of the jawbone (pleurodont condition). In polyphyodont animals, teeth may be constantly replaced. The polyphyodont condition is characteristic of fishes and the lower tetrapods in which the teeth are replaced as they are worn out. Replacement may be random, as in many fishes (sharks are an excellent example), or may follow a definite pattern, such as the waves of replacement seen in many reptiles. Most mammals have two sets of teeth during their lifetime, which is the diphyodont condition. These sets are the deciduous, or milk, teeth and the permanent dentition. Monophyodont dentition is the development of only one set of teeth.

Dental formulas of some mammals

Animal	Teeth I	C	Pm	M	Total
Human	2/2	1/1	2/2	3/3	32
Cony	3/3	1/1	4/4	4/4	48
Beaver	1/1	0/0	1/1	3/3	20
Cat	3/3	1/1	3/2	1/1	30
Dog	3/3	1/1	4/4	2/3	42
Sheep	0/3	0/1	3/3	3/3	32
Lynx	3/3	1/1	2/2	1/1	28
Rat	1/1	0/0	0/0	3/3	16
Horse	3/3	1/1	4/4	3/3	44
Mole	3/3	1/1	4/4	3/3	44
Squirrel	1/1	0/0	2/1	3/3	22
Reindeer	0/3	0/1	3/3	3/3	32
Pig	3/3	1/1	4/4	3/3	44
Common seal	3/2	1/1	4/4	1/1	34
Skunk	3/3	1/1	3/3	1/2	34
Raccoon	3/3	1/1	4/4	2/2	40
Bear	3/3	1/1	4/4	2/3	42

Patterns. Teeth may have a similar form in all regions where they occur. Such teeth are homodont, as distinct from those that are variable in shape, called heterodont dentition. The homodont condition is characteristic of nonmammalian vertebrates, although the anterior teeth may differ from those lying in the cheek region of the jaws. A heterodont dentition began to evolve in the reptilian ancestors of mammals, but became fully developed in the mammals. The mammalian heterodont condition is frequently described by a dental formula expressing both the number and kind of teeth in each half jaw, both upper and lower (see **table**). For humans the formula is I 2/2, C 1/1, Pm 2/2, M 3/3. This can be read for either the right or left side of the jaw; there are 2 upper and 2 lower incisors, 1 upper and 1 lower canine, and 2 premolars and 3 molars in both the upper and lower jaws. According to this formula, humans have a total of 32 teeth, 8 located in each half of the upper jaw and 8 in each half of the lower jaw.

The initial letter may be omitted in the dental formula, and if a tooth does not appear in the group, a zero so indicates. Unfortunately, the formula does not indicate which of the several molars or premolars may be absent in a species, nor does it provide information as to the morphology of the several tooth types. Such data are equally important. Information about the dentition is an essential part of mammalian systematics at all levels for both fossil and living species. The dental formula and other dentition information are included in the description of all species of living mammals. Because teeth are very hard structures and fossilize well, and because teeth are usually

characteristic of the mammalian taxa (even species in many cases), the dentition is one of the most valuable features in mammalian paleontology.

Evolution. Study of dentitions has revealed that the general evolutionary trend among vertebrates has been a reduction of the numerous successive dentitions in favor of fewer sets of teeth, and a change from the usual homodont dentition to the heterodont dentition of mammals with its several characteristic types of teeth. The latter trend has reversed in some mammalian groups, such as the whales, which have a very uniform homodont dentition. *See* TOOTH.

Frederick S. Szalay

Denudation

The term "denudation" was introduced into the earth sciences in the early nineteenth century to refer to the removal of surface material that leads to the exposure of underlying strata and rock units. Now, denudation refers to all the weathering and erosional processes that contribute to the lowering of the land surface, potentially to the production of a peneplain as the product of a long period of erosion. Denudation is, thus, the complement of deposition, the accumulation of the products of denudation in sedimentation basins. *See* BASIN; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; EROSION; WEATHERING PROCESSES.

Contemporary records of sediment and solute flux through river systems, representing the mass removed from the continental surfaces, have been used to estimate rates of denudation since the mid-nineteenth century. These are usually expressed as an average rate of lowering of the land surface, in units of millimeters per thousand years (mm/ka), and are often differentiated as mechanical denudation and chemical denudation. The former should include both suspended and bed load transport, but is often estimated from the suspended sediment yield alone, or with an added fixed proportion (10–15%) to account for bed load. The latter includes material removed from the landmasses as solutes within river discharges and usually needs to be corrected for those constituents which are derived from the atmosphere. *See* SEDIMENTOLOGY.

Even when estimated for large drainage basins, rates of denudation are highly variable. Mechanical denudation in the world's largest river systems varies between 4 mm/ka and more than 500 mm/ka. Chemical denudation rates on a global scale are both lower and less variable, falling within the range 2–30 mm/ka. This variability is greatly increased when denudation rates are calculated for small drainage basins, especially those in regions of variable relief. Environmental controls on contemporary rates of denudation have often been sought in climate, as reflected in effective precipitation, runoff, or vegetation. Other studies emphasize the importance of topography and local relief as controls of denudation rates, especially in tectonically active mountain systems.

On a geologic time scale, denudation should lead to the development of planation surfaces of low relief that were classically interpreted as providing a denudation chronology based on relative age. Until recently, such studies suffered from a dearth of procedures capable of defining a true chronology of erosional landforms. That has greatly changed in recent decades with the development of methods which allow the definition of denudation rates over geologic time scales. *See* GEOLOGIC TIME SCALE.

Attempts to invert sedimentary and stratigraphic records, and so define rates of erosion on the contributing continental surface, have greatly benefitted from improved seismic surveys, core extraction, and modeling of sedimentation basin processes. On the land surface itself, estimates of denudation rates based on the dissection of surfaces of known age have improved with the use of digital terrain models and absolute dating techniques. Equivalent estimates show that very rapid rates of bedrock denudation (up to 15 mm/yr) have been maintained in the western Himalaya during the past few million years. *See* DATING METHODS.

However, developments in the dating of bedrock surfaces and surficial material in the 1990s represent the greatest improvement in defining rates of denudation during the late Cenozoic era. Potassium-argon (K-Ar) dating of lavas and the magnitude of dissection on volcanic cones had been used earlier to define erosion rates. More recently, exposure dating of bedrock through the accumulation of cosmogenic nuclides produced in situ within them is a development that allows the estimation of average surface lowering rates over long time intervals. Cosmogenic nuclides have also been used to define rates of regolith production and erosion, contributing to denudation, over more recent time periods. Fission track analysis of apatite is another recent development that has had success in defining the long-term tectonic and denudational history of continental margins and mountain systems. *See* APATITE; CENOZOIC; COSMOGENIC NUCLIDE; FISSION TRACK DATING; REGOLITH.

Nel Caine

Bibliography. J. M. Harbor (ed.), Cosmogenic isotopes in geomorphology, *Geomorphology*, 27:1–172, 1999; J. D. Milliman and J. P. M. Syvitski, Geomorphic/tectonic control of sediment discharge to the ocean: The importance of small mountainous rivers, *J. Geol.*, 100:525–544, 1992; M. A. Summerfield and N. J. Hulton, Natural controls of fluvial denudation rates in major world drainage basins, *J. Geophys. Res.*, 99(B7):13871–13883, 1994.

Deoxyribonucleic acid (DNA)

A linear polymer made up of a specific sequence of deoxyribonucleotide repeating units linked by 3′,5′-phosphodiester bonds; it is the carrier of genetic information. The set of DNA molecules that contains all genetic information for an organism is called its genome. DNA is found primarily in the nuclei of eukaryotic cells and in the nucleoid of bacteria. Small amounts of DNA are also found in organelles (such

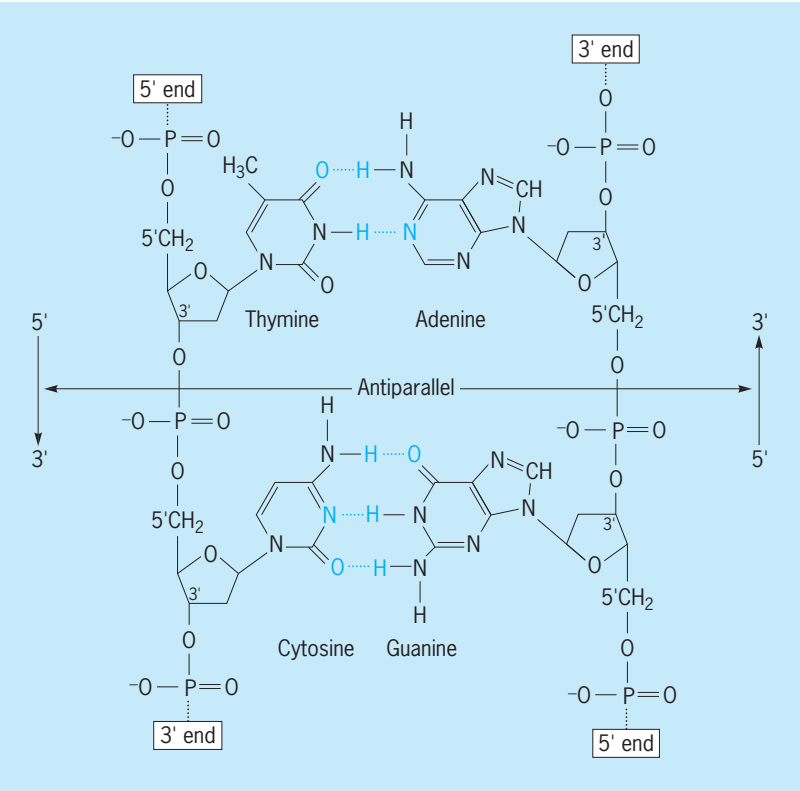

Fig. 1. Diagram of the DNA duplex. Each of the two strands comprises repeating units of a sugar–phosphate polymer. Each nucleotide on one strand is connected to the next by a phosphodiester bond. The linkage is always made from the the 3′ position of one sugar to the 5′ position of the next. The two strands are aligned antiparallel to each other, with deoxyribose on the outside and purines and pyrimidines on the inside of the duplex. Purines and pyrimidines share hydrogen bonds, with adenine binding to thymidine and cytosine binding to guanine. (*Reprinted with permission of John Wiley & Sons, Inc., from T. M. Devlin, Textbook of Biochemistry with Clinical Correlations, 4th ed., 2002. Copyright © 1997 Wiley-Liss, Inc.*)

as mitochondria and chloroplasts) that contain their own genomes, in autonomously maintained DNAs called plasmids, and in viruses. *See* NUCLEIC ACID.

DNA is composed of two long polymer strands of the sugar 2-deoxyribose, phosphate, and purine and pyrimidine bases. The backbone of each strand is composed of alternating 2-deoxyribose and phosphate linked together through phosphodiester bonds (**Fig. 1**). A DNA strand has directionality: each phosphate is linked to the 3′ position of the preceding deoxyribose and to the 5′ position of the following deoxyribose. The four bases found in DNA are adenine, thymine, guanine, and cytosine. Each 2-deoxyribose is linked to one of four bases in the 1′ position via a covalent glycosidic bond. A base linked to a sugar is called a nucleoside, and a base linked to a sugar phosphate is called a nucleotide. The numbering of atom positions of the sugar is accompanied by a prime sign to distinquish it from the positional numbering of the bases. The sequence of these four bases allows DNA to carry genetic information. Bases can form hydrogen bonds with each other. Adenine forms two bonds with thiamine, and cytosine forms three bonds with guanine. These two sets of base pairs have the same geometry, allowing DNA to maintain the same structure regardless of the specific sequence of base pairs (Fig. 1). *See* DEOXYRIBOSE; PURINE; PYRIMIDINE.

Structure. DNA can be single-stranded or aligned to form two-, three-, or even four-stranded structures. It is commonly found in its double-stranded form in chromosomes. In double-stranded DNA, the two strands wrap around each other to form a double helix (**Fig. 2**). The two strands are held together by base pairing and are antiparallel. Thus if one strand

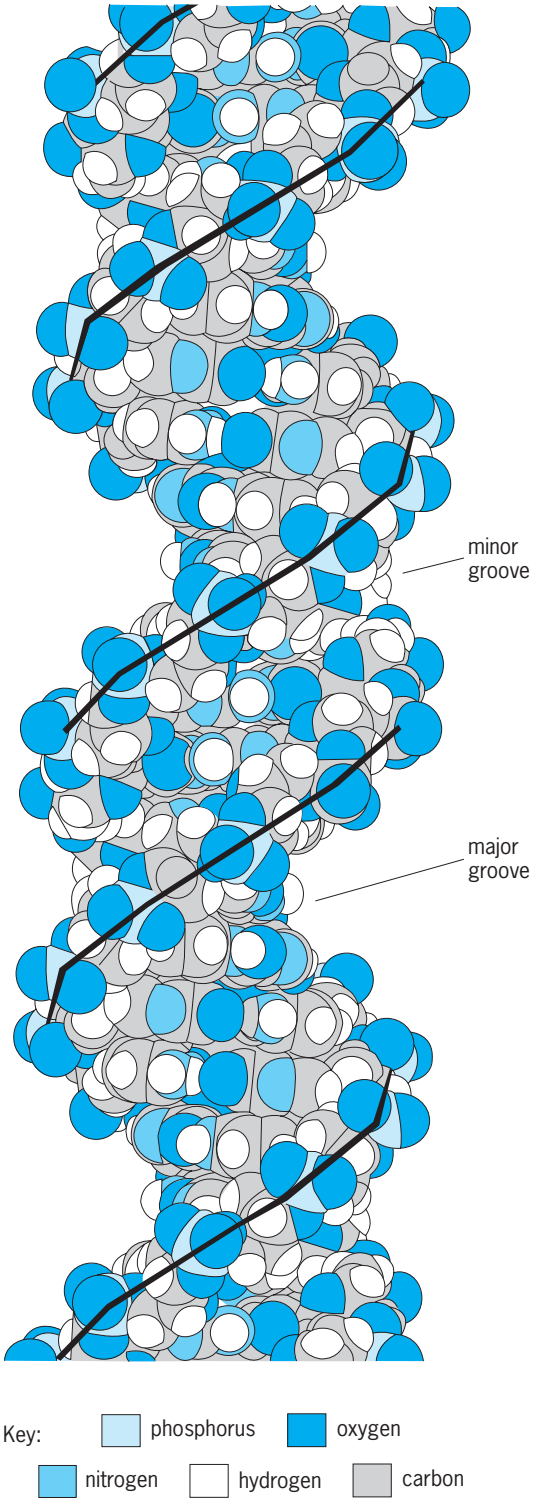

Fig. 2. B-DNA structure. The helical axis is vertical, and a solid line is drawn from phosphate to phosphate along the backbone.

is oriented in the 5′ to 3′ direction, the other strand will be 3′ to 5′. This double-helical structure of DNA was first proposed in 1954 by J. D. Watson and F. H. C. Crick.

The most common form of DNA is the B-form, which is a right-handed double helix with 10.4 base pairs per turn (Fig. 2). In a B-form double helix, the ribose–phosphate backbones run along the outside of the helix, and the base pairs are roughly perpendicular to the axis of the helix and lie just offset from the center axis of the helix. This offset causes one of the two grooves that run along the helix to be larger than the other—thus the designations, major groove and minor groove. Proteins that bind to DNA usually bind specifically to one of the two grooves. Less common forms of DNA include the A-form, which is a right-handed double helix that has 11 base pairs per turn and has a wider diameter than the B-form, and the Z-form, which is a narrow, irregular left-handed double helix (**Fig. 3**).

Right-handed conformation. Following Watson and Crick's proposal in 1954, experiments supported a right-handed form of the double helix in which there were approximately 10 base pairs per turn. In this form, each base pair is similar to the neighboring base pair, except that there is a rotation of approximately 36° between the base pairs. The conformation of DNA was determined by x-ray diffraction studies of fibers. A dried DNA fiber produces a type A pattern, in which the base pairs are tilted somewhat away from the helix axis. However, a wet or hydrated DNA fiber produces a B-type pattern, in which the bases are perpendicular to the helix axis (Fig. 2).

Left-handed conformation (Z-DNA). In 1979 it was discovered that DNA can also form a left-handed double helix with the same hydrogen bonding between the two bases (Fig. 1). This was discovered by solving the structure of crystals of short DNA fragments; crystals can yield a more detailed view of the molecule. A solid line drawn between the phosphates in the left-handed conformation follows a zigzag path (Fig. 3); hence it is called Z-DNA. In contrast to B-DNA, Z-DNA has only one deep helical groove extending almost to the axis of the helix. The base pairs in left-handed Z-DNA are located near the outside of the helix. B-DNA has a one-base-pair repeating unit, whereas Z-DNA has a dinucleotide repeating pattern. Both forms of DNA have the two sugar–phosphate backbones oriented in opposite directions (Fig. 1). Z-DNA has been identified in DNA fibers as well as inside living cells.

Conformational changes. In biological systems, DNA is normally subjected to a torsional strain that tends to unwind the double helix. This torsional strain tends to untwist right-handed B-DNA and sometimes is great enough to stabilize certain regions of DNA in the left-handed Z-DNA conformation. The conversion from right-handed B-DNA to left-handed Z-DNA is associated with changes in the relationship of the bases to the sugar backbone. Every other base in Z-DNA is rotated about the bond connecting it to the sugar. This altered form is common in adenine or guanine residues (purines) but less common

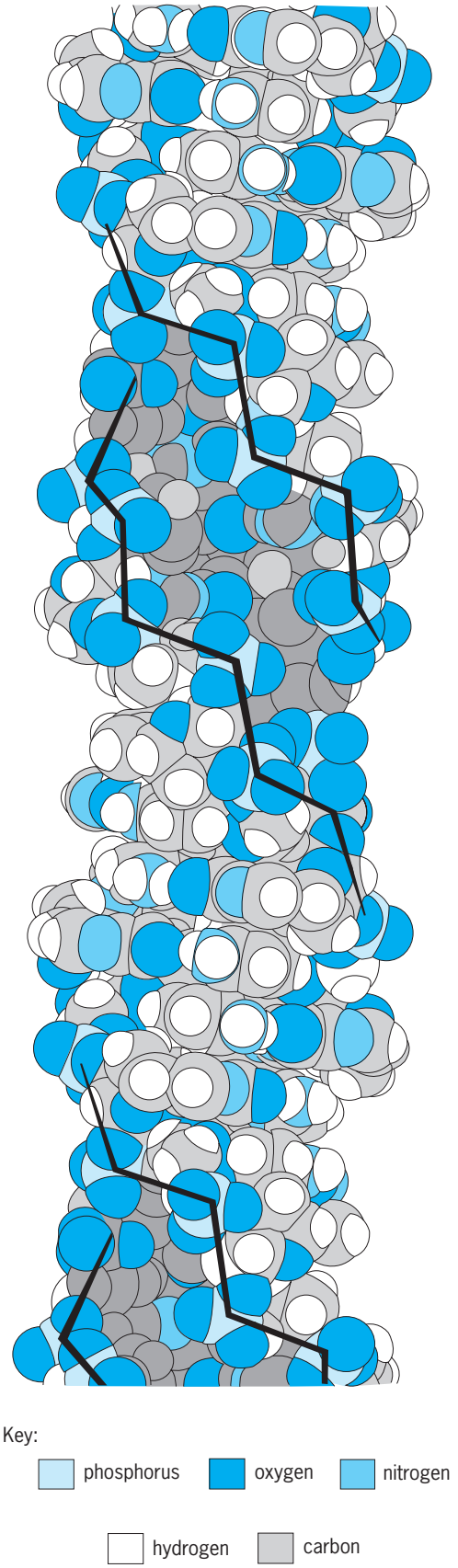

Fig. 3. Z-DNA structure. Darker gray area is the deep, narrow minor groove. A solid line is drawn from phosphate to phosphate along the backbone.

in thymine or cytosine residues (pyrimidines); thus, sections that form Z-DNA often have alternations of purines and pyrimidines.

Although right-handed B-DNA is the most common form of DNA, left-handed Z-DNA occurs in a number of places and may play an important role in the regulation of gene expression. In addition, Z-DNA appears to be an important component of the process of crossing-over or recombination of DNA molecules. Z-DNA is important in carrying out special functions that are in general associated with special proteins that bind to Z-DNA and not to B-DNA. *See* CROSSING-OVER (GENETICS); GENE ACTION.

Marc S. Wold; Alexander Rich; Daniel L. Weeks

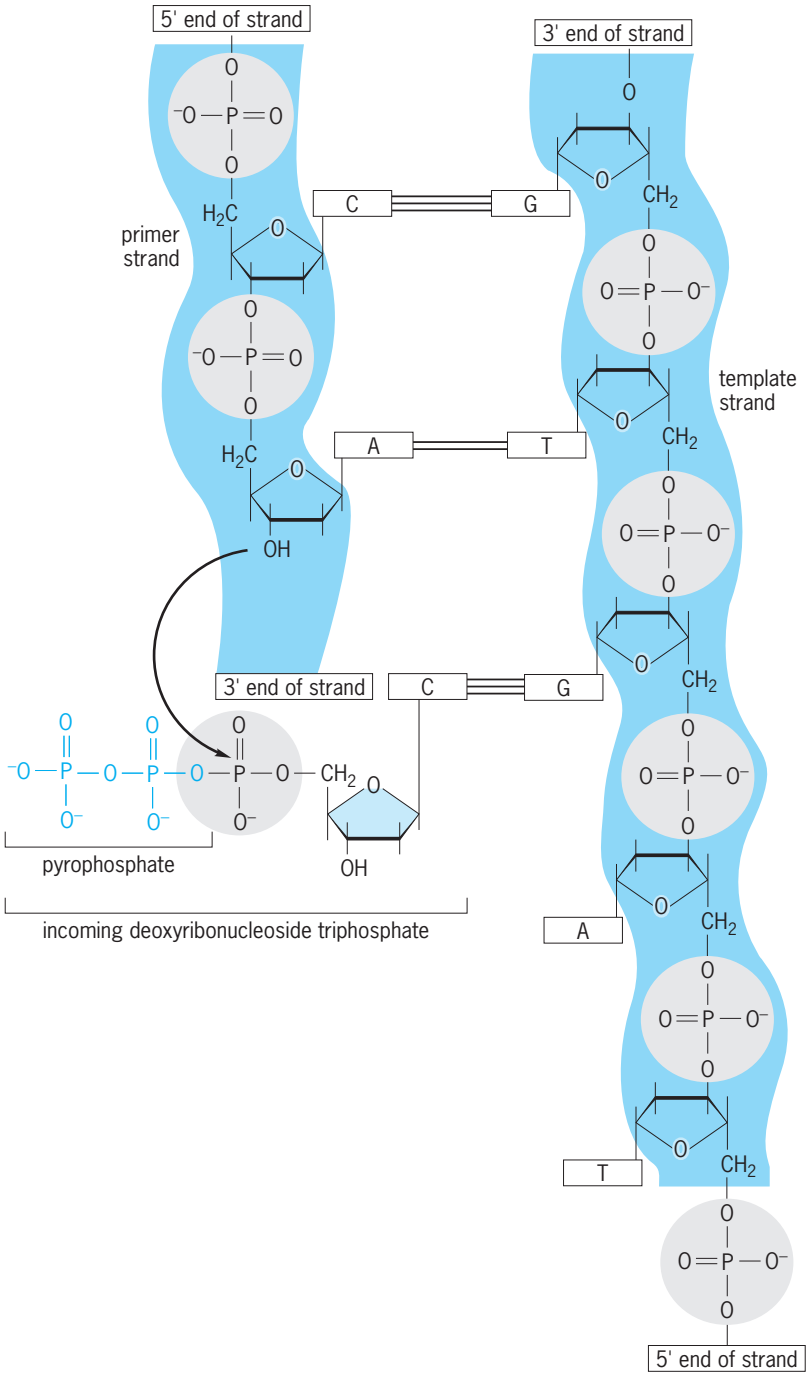

Fig. 4. Replication of DNA. The replication shown here highlights the selection of new base insertion by using Watson-Crick (A-to-T and G-to-C) base pairing, the addition of new bases to the 3′ end of an existing chain, and the attack of the 3′OH on the alpha phosphate of a nucleoside triphosphate to add the newest base. (*Copyright © 2002. From Molecular Biology of the Cell, 4th ed., by B. Alberts et al. Reproduced with permission of Garland Science/Taylor & Francis LLC*)

Function. For cells to live and grow, the genetic information in DNA must be (1) propagated and maintained from generation to generation and (2) expressed to synthesize the components of a cell. These two functions are carried out by the processes of DNA replication and transcription, respectively. *See* GENETIC CODE.

Replication. Each of the two strands of a DNA double helix contains all of the information necessary to make a new double-stranded molecule (**Fig. 4**). During replication the two parental strands are separated, and each is used as a template for the synthesis of a new strand of DNA. Base incorporation is directed by the existing DNA strand; nucleotides that base-pair with the template are added to the nascent DNA strand. The product of replication is two complete double-stranded helices, each of which contains all of the genetic information (has the identical base sequence) of the parental DNA. Each progeny double helix is composed of one parental and one nascent strand. DNA replication is very accurate. In bacteria the mutation rate is about 1 error per 1000 bacteria per generation, or about 1 error in 10^9 base pairs replicated. This low error rate is due to a combination of the high accuracy of the replication process and cellular pathways which repair misincorporated bases.

Synthesis of the nascent DNA strands is carried out by a family of enzymes called DNA polymerases. These enzymes synthesize a new phosphodiester bond between the 3′ hydroxyl position of the terminal deoxyribose of an existing DNA (or ribonucleic acid, RNA) chain and a deoxynucleotide triphosphate. All DNA polymerases (1) synthesize new DNA in a 5′ to 3′ direction, (2) require a primer (they must add on to the 3′ hydroxyl position of DNA or RNA), and (3) require a DNA molecule as a template. DNA polymerases, in conjunction with multiple accessory proteins, can rapidly synthesize new DNA strands. (The mandatory participation of the 3′ hydroxyl forms the rationale for the use of nucleoside analogues that lack a 3′ hydroxyl as antiviral and anticancer drugs, as they lead to premature chain termination.) A bacterial replication fork moves at a rate of approximately 1000 nucleotides per second, while in eukaryotic cells forks move at a rate of about 100 nucleotides per second. *See* MUTATION.

Transcription. In transcription, DNA acts as a template directing the synthesis of RNA. RNA is a single-stranded polymer similar to DNA except that it contains the sugar ribose instead of 2-deoxyribose, and the base uracil instead of thymidine (**Fig. 5**). The two strands of DNA separate transiently, and one of the two single-stranded regions is used as a template to direct the synthesis of an RNA strand. As in DNA replication, base pairing between the incoming ribonucleotide and the template strand determines the sequence of bases incorporated into the nascent

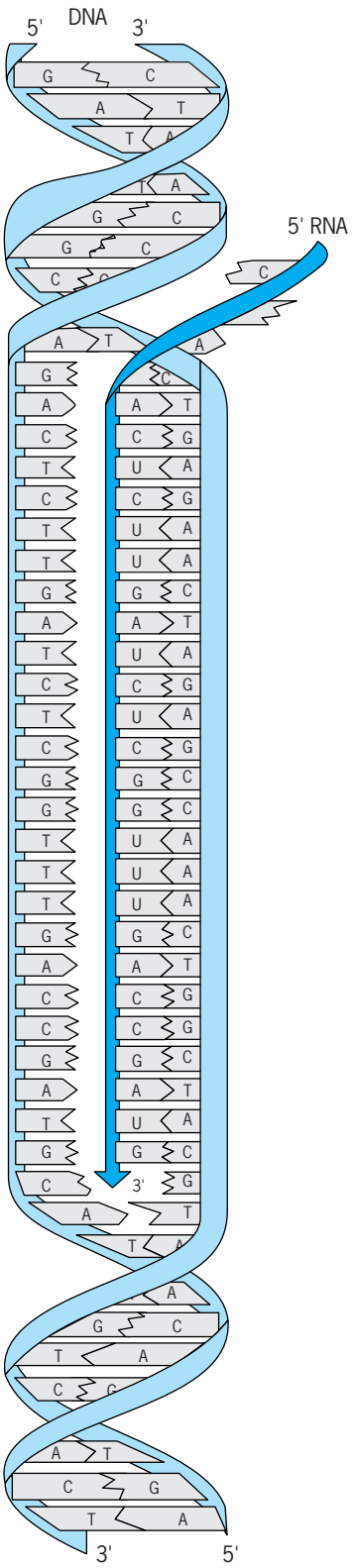

Fig. 5. Transcription of DNA. The double-stranded DNA is transiently separated to allow the synthesis of RNA. Using a mechanism similar to that discussed for DNA replication (Fig. 4), ribonucleotides are added to the 3′ end of a growing chain, and the selection of bases to insert is guided by Watson-Crick base pairing. (*Reprinted with permission from G. M. Cooper, The Cell: A Molecular Approach, 2d ed., 2000. © Sinauer Associates, Inc.*)

RNA. Thus, genetic information in the form of a specific sequence of bases is directly transferred from DNA to RNA in transcription. After the RNA is synthesized, the DNA reverts to a double-stranded form. Transcription is carried out by a family of enzymes called RNA polymerases. In prokaryotes, RNA polymerase interacts directly with the sites at which RNA transcription start (promoters), while in eukaryotes large complexes of proteins bind to promoters and then direct RNA polymerase to start RNA synthesis. Newly synthesized RNA is ultimately used to direct protein synthesis by ribosomes in a process called translation. *See* PROTEIN; RIBONUCLEIC ACID (RNA); RIBOSOMES.

Genetic variation. There is a great deal of variation in the DNA content and sequences in different organisms (**Fig. 6**). Because of base pairing, the ratios of adenine to thiamine and cytosine to guanine are always the same. However, the ratio of adenine and thymine to guanine and cytosine in different organisms ranges from 1:4 to 3:4. There is also large variation in the amount of DNA in the genome of various organisms. The simplest viruses have genomes of only a few thousand base pairs, while complex eukaryotic organisms have genomes of billions of base pairs. (The lungfish has one of the largest genomes known, about 100 billion base pairs.) This variation partially reflects the increasing number of genes necessary to encode more complex organisms, but mainly reflects an increase in the amount of DNA that does not encode proteins. A large percentage of the DNA in multicellular eukaryotes is noncoding or is repetitive DNA (sequences that are repeated many times). This DNA can serve structural or regulatory functions, it can be related to transposable elements (DNA sequences that can insert into other unrelated sequences), or in many cases it has no known function. Rarely, genes are present in multiple copies to allow high levels of expression of certain cellular components. For example, most cells have hundreds or thousands of copies of the genes encoding the components needed for protein synthesis to ensure sufficient levels of expression of these components.

In most eukaryotes the DNA sequences that encode proteins are not continuous but have other sequences interspersed within them. The sequences that are expressed in protein are called exons, and the interspersed sequences are called introns. The initial transcript synthesized by RNA polymerase contains both exons and introns and can be many times the length of the actual coding sequence. The RNA is then processed, and the introns are removed through a mechanism called RNA splicing to yield messenger RNA (mRNA), which is translated to make protein.

Recombinant technology. Techniques have been developed to allow DNA to be manipulated in the laboratory. These techniques, collectively called molecular biology or molecular genetics, have led to a revolution in biotechnology. This revolution began when methods were developed to cleave DNA at specific sequences and to join pieces of DNA together. Another major component of this technology is the

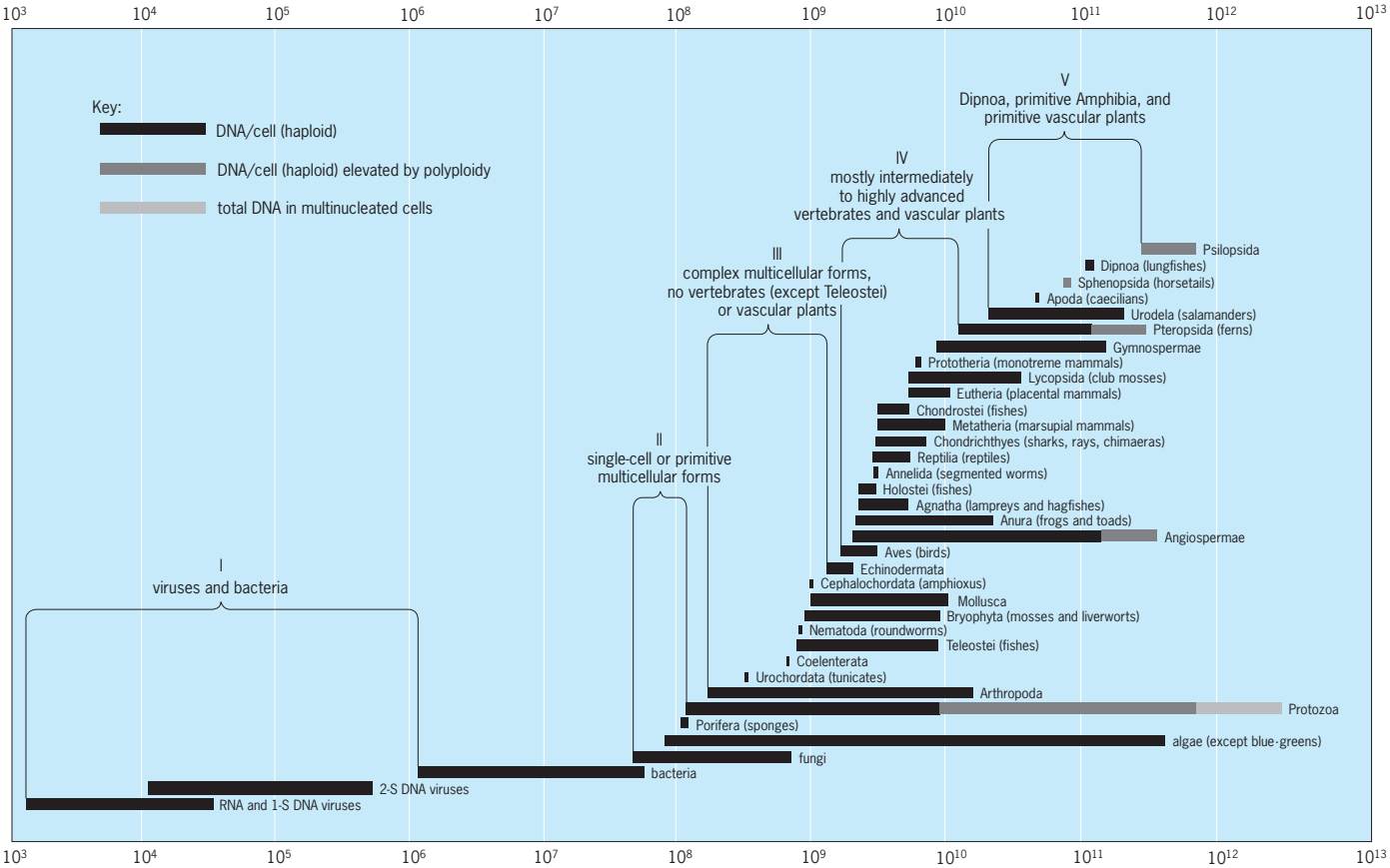

Fig. 6. Haploid DNA amounts (RNA in some viruses). Amounts are expressed as numbers of nucleotides. One gram of DNA = 2.01×10^{21} nucleotides. (*After A. J. Sparrow, H. J. Price, and A. G. Underbrink, A survey of DNA content per cell and per chromosome of prokaryotic and eukaryotic organisms: Some evolutionary considerations, in the Brookhaven Symposia in Biology, no. 23, Evolution of Genetic Systems, pp. 451-494, 1972*)

ability to determine the sequence of the bases in DNA. There are two general approaches for determining DNA sequence: either chemical reactions are carried out which specifically cleave the sugar–phosphate bond at sites which contain a certain base, or DNA is synthesized in the presence of modified bases that cause termination of DNA synthesis after the incorporation of a certain base. A series of four reactions, one for each base, are carried out, and the lengths of the DNA fragments produced identify the sequence of the bases. These methods can now be automated so that it is practical to determine the DNA sequences of the entire genome of an organism. The complete sequences of humans and a number of experimentally important organisms have been determined. *See* HUMAN GENOME PROJECT.

Another example of recombinant DNA technology is copy DNAs (cDNAs), which are made when an mRNA is converted to DNA using reverse transcriptase (a specialized DNA polymerase that can use RNA as a template). The resulting DNA can then be cloned, made in large quantity, and its sequence determined. Because it is much easier to manipulate DNA than RNA, cDNAs allow any RNA-based species to be characterized. *See* GENETIC ENGINEERING; MOLECULAR BIOLOGY; REVERSE TRANSCRIPTASE.

Marc S. Wold; Daniel L. Weeks

In the cell. The full genome of DNA must be substantially compacted to fit into a cell. For example, virtually every cell in the human body contains a full genome of DNA, which in turn has a total length of about 3 m (10 ft). However, this DNA must fit into a nucleus with a diameter of 10^{-5} m. This immense reduction in length is accomplished in eukaryotes via multiple levels of compaction in a nucleoprotein structure termed chromatin. To summarize, the first level involves spooling about 200 base pairs of DNA onto a complex of basic proteins called histones to form a nucleosome. Nucleosomes are connected like beads on a string (**Fig. 7**) to form a 10-nanometer-diameter fiber, and this is further coiled to form a 30-nm fiber. The 30-nm fibers are further coiled and organized into loops formed by periodic attachments to a protein scaffold. This scaffold organizes the complex into the shape of the metaphase chromosome seen at mitosis.

The nucleosome is the fundamental structural unit of DNA in all eukaryotes. Virtually all of the DNA in each cell of an organism is packaged into nucleosomes. Nucleosomes are squat cylinders (6 nm × 11 nm × 11 nm) arranged as repeating subunits along the DNA at approximately 200-base-pair intervals. Pairs of four different histone proteins (H2A, H2B, H3, and H4) constitute an octamer complex,

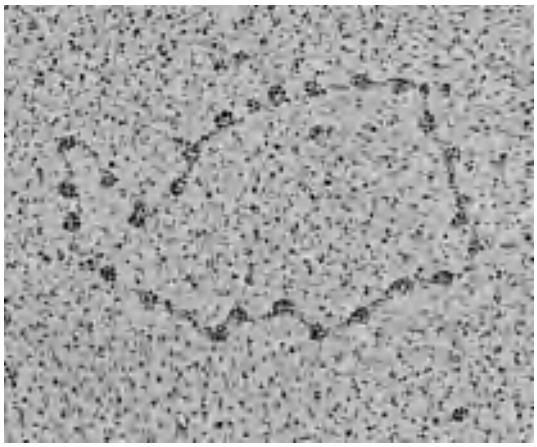

Fig. 7. Electron micrograph of the minichromosome of a virus that infects monkey cells. The circular viral DNA is organized in nucleosome "beads." In order to improve the ability to visualize individual nucleosomes, the conditions of microscopy in this figure have been adjusted to stretch the minichromosome slightly. In the cell, successive nucleosomes are in direct contact with one another, resulting in a circle of a relatively uniform fiber of about 11 nm diameter. (*From L. Lutter et al., Mol. Cell. Biol., 12:5004–5014, 1992*)

on the outside of which is wrapped 146 base pairs of DNA to form the nucleosome core. The DNA in the core is wrapped in $1^{3}/_{4}$ left-handed coils, resulting in approximately sixfold compaction of the DNA. The detailed structure of the nucleosome core (**Fig. 8**) has been determined by x-ray crystallography. Nucleosome cores are connected to one another by about 55 base pairs of linker DNA. The 146-base-pair length of DNA in the nucleosome core is invariant among all species studied, but the length of linker DNA can range from 20 to 100 base pairs among different species and different tissues. The reason for this variation is unknown. The linker DNA binds a fifth histone, H1, that plays a role in condensation of the 10-nm fiber into the higher-order structures.

Nucleosomes reduce the accessibility of DNA to DNA-binding proteins such as polymerases and other protein factors essential for transcription and replication. One level of regulation of gene expression occurs via posttranslational modification of histones. For example, specific protein factors activate transcription by adding acetyl groups to specific amino acid residues in the histones, while other factors with the opposite transcription effect remove those residues. Malfunctions in these regulatory processes have been associated with human diseases. There are several other types of histone posttranslational modifications with regulatory effects. Thus chromatin structure not only compacts the genome but also affords an additional level of control of gene expression. *See* CHROMOSOME; NUCLEOPROTEIN; NUCLEOSOME.

Leonard C. Lutter

Bibliography. B. Alberts et al., *Molecular Biology of the Cell*, 4th ed., Garland, New York, 2002; G. M. Cooper, *The Cell: A Molecular Approach*, 2d ed., Sinauer Associates, Sunderland, MA, 2000; J. Darnell et al., *Molecular Cell Biology*, 1986; T. M. Devlin, *Textbook of Biochemistry with Clinical Correlations*, Wiley-Liss, New York, 2002; D. M. Glover, *Gene Cloning: The Mechanics of DNA Manipulation*, 1985; P. C. Hanawalt, *Molecules to Living*

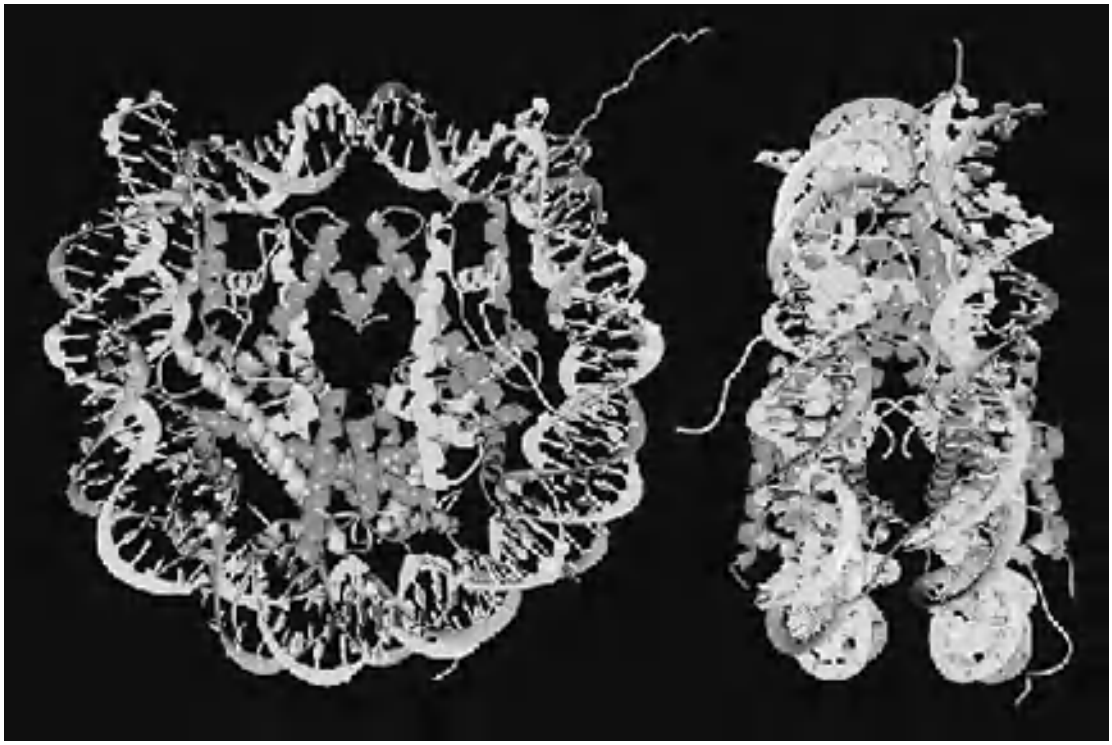

Fig. 8. Structure of the nucleosome core particle as determined by x-ray crystallography. On the left is a view down the superhelix axis, showing the DNA wrapped on the outside of the histone octamer. On the right is a view from the side, showing the two gyres of superhelix of 146 base pairs of DNA. (*From K. Luger et al., Nature, 389:251-260, 1997*)

Cells, 1985; H. F. Judson, *The Eighth Day of Creation: Makers of the Revolution in Biology*, 1996; A. Kornberg and T. A. Baker, *DNA Replication*, 2d ed., 1992; H. Lodish et al., *Molecular Cell Biology*, 5th ed., W. H. Freeman, 2003; R. McMacken and T. J. Kelly (eds.), *DNA Replication and Recombination*, 1987; J. D. Watson, *Double Helix: Being a Personal Account of the Discovery of the Structure of DNA*, 1968; J. D. Watson et al., *Recombinant DNA*, 2d ed., 1992; J. D. Watson and J. Tooze, *The DNA Story; A Documentary History of Gene Cloning*, 1983.

Depositional systems and environments

Depositional systems are descriptions of the interrelationships of form and the physical, chemical, or biological processes involved in the development of stratigraphic sequences. Depositional environments are the locations where accumulations of sediment have been deposited by either mechanical or chemical processes.

Depositional Systems

A basic premise in geological investigations is the concept of uniformitarianism—the present is the key to the past. This historical premise has been extended thoroughly in the fields of stratigraphy, sedimentology, and basin analysis to the point where ancient strata are commonly characterized by their genesis. Analysis of modern depositional systems and environments has long been the subject of geologic investigations, with studies increasing vastly in scope and volume since 1950. A greatly increased understanding of modern environments of deposition has generated a corresponding, extensive application of modern models to ancient strata that are inferred to be the analogs of modern depositional environments. *See* STRATIGRAPHY.

Traditional stratigraphic analysis, which emphasized the physically descriptive aspects, has thus changed; a critical genetic dimension has been added. Where once, for example, a formation may have been described in physical terms as a fine- to medium-grained quartzose sandstone, overlying a thick dark-gray shale sequence and underlying a coal-bearing sequence with discontinuous sands, the same sandstone may now be recognized as a delta-front sandstone, the underlying shale as a prodelta facies, and the overlying coal-bearing formation as the product of deposition on a delta plain. In this interpretation, the three distinct stratigraphic units become part of a genetically related sequence, each a component facies of a prograding delta system.

Modern systems. In a modern setting, a particular system of deposition is directly observable and known, whether it is a major delta, an alluvial fan, a meandering river, a barrier bar, a carbonate platform, or the like. Through specific observation, description, and delineation, it may be determined that a variety of depositional processes are active. At the terminus of rivers, sands may be deposited as bars in river mouths, creating the delta front; muds may be carried by suspension into the oceanic waters and deposited through flocculation, creating the prodelta; on the delta plain, a distributary channel may be depositing bed-load sands and, during flooding, may be carrying suspended muds to flanking flood basins or breaching the natural levee to

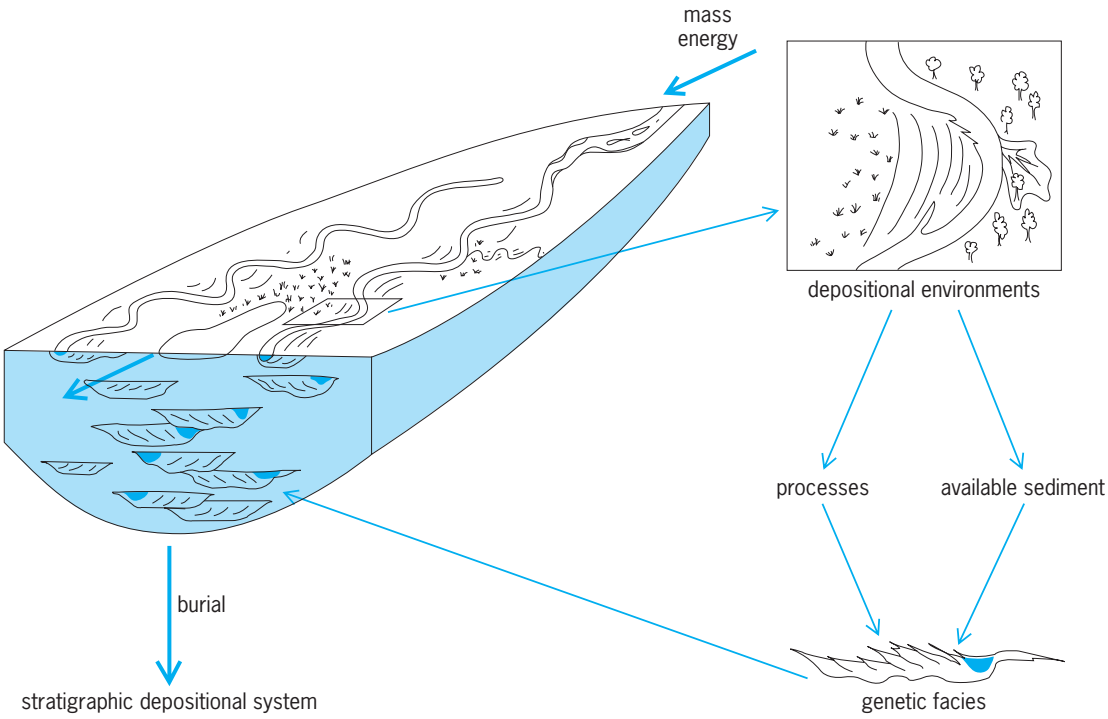

Fig. 1. An active depositional system consisting of genetically related environments. (*After W. E. Galloway, C. W. Kreitler, and J. H. McGowen, Depositional and Ground-Water Flow Systems in the Exploration for Uranium, University of Texas at Austin, Bureau of Economic Geology Special Publication, 1979*)

form crevasse splays. Each process or combination of processes gives rise to distinct, specific environments of deposition, with each resulting in a deposit which can be characterized by such features as lithologic composition, texture, sedimentary structures, geometry, size, and relationship to other deposits. *See* DELTA; FLOODPLAIN; RIVER.

Distinct physical, or in some cases biologic or chemical, products of deposition can be related directly to definable, operative processes. Such data permit the development of models of modern deposition in which processes and resulting deposits or facies are linked. By recognizing comparable physical, chemical, and biologic attributes of ancient strata, modern depositional analogs can be applied, and the original processes forming the ancient deposit can be inferred. Such an ancient deposit is called a genetic facies (**Fig. 1**). It contains the sedimentary record and constitutes a three-dimensional stratigraphic (ancient) depositional system. *See* FACIES (GEOLOGY).

Ancient record. As modern or active depositional systems consist of a complex of genetically related environments, the ancient sedimentary record likewise consists of a variety of genetic facies. When ancient genetic facies are defined and when several facies constitute a genetically related assemblage, they make up a depositional system (**Fig. 2**). Such systems, considered as three-dimensional entities, constitute specific units of entire sedimentary basin fills. A natural system may be defined as a structured set of objects, or attributes, which display recognizable interrelationships and operate together as a complex whole. The interdependence of parts is an inherent and diagnostic property of all natural systems.

An ancient depositional system is a three-dimensional, genetically defined, physical stratigraphic unit that consists of a contiguous set of process-related sedimentary facies. Several corollaries have evolved from the application of this concept. Depositional systems, such as delta, fluvial, and shelf systems, (1) are the stratigraphic equivalents of major physical geographic units; (2) form the principal building blocks of the sedimentary basin fill; and (3) can be applied where principal boundaries of the systems are preserved and where the geometry of the framework facies can be mapped.

Process-response systems. Depositional systems are process-response systems; that is, they describe the interrelationships of form and physical, chemical, or biologic processes. Such a process-response system provides a tool for prediction because it allows extrapolation beyond the data and can be used to predict properties not included in the original definition of the system. For example, the recognition of

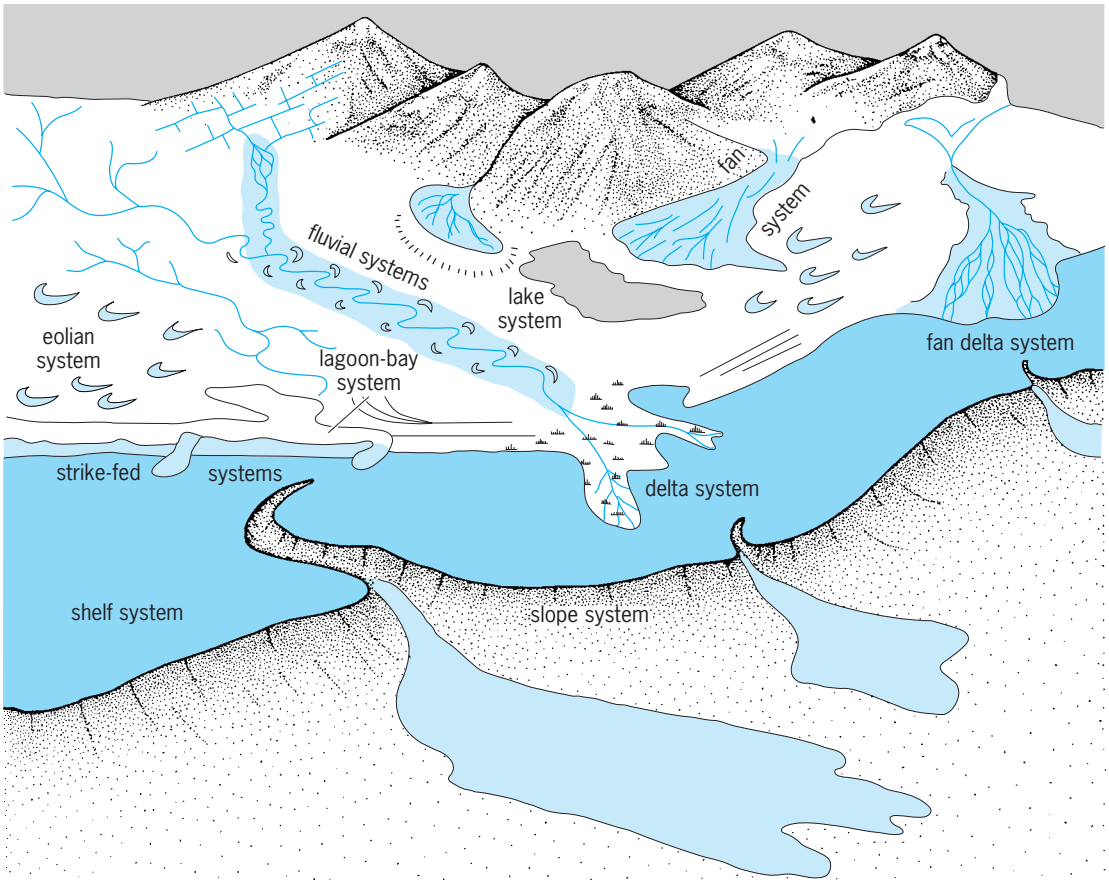

Fig. 2. Schematic setting of major depositional systems. (*After L. F. Brown, Jr., A. W. Cleaves III, and A. W. Erxleben, Pennsylvania Depositional Systems in North-Central Texas, University of Texas at Austin, Bureau of Economic Geology Guidebook 14, 1973*)

a delta-plain facies permits the prediction of the basinward occurrence of delta-front and prodelta facies, the logical component facies of a delta system.

The genetic facies and depositional system approach to analyzing ancient deposits is not a substitute for the traditional approaches to basin analysis. Rather, it utilizes all the traditional tools of analysis, particularly from the standpoint of description, but adds, through an analogy with modern processes and models, the genetic dimension.

Realms of deposition. The major realms of deposition may be classed broadly as terrigenous clastic depositional systems and biogenic-chemical depositional systems (**Figs. 3** and **4**). Each of these major systems is subdivided according to particular systems of deposition, and within each of the

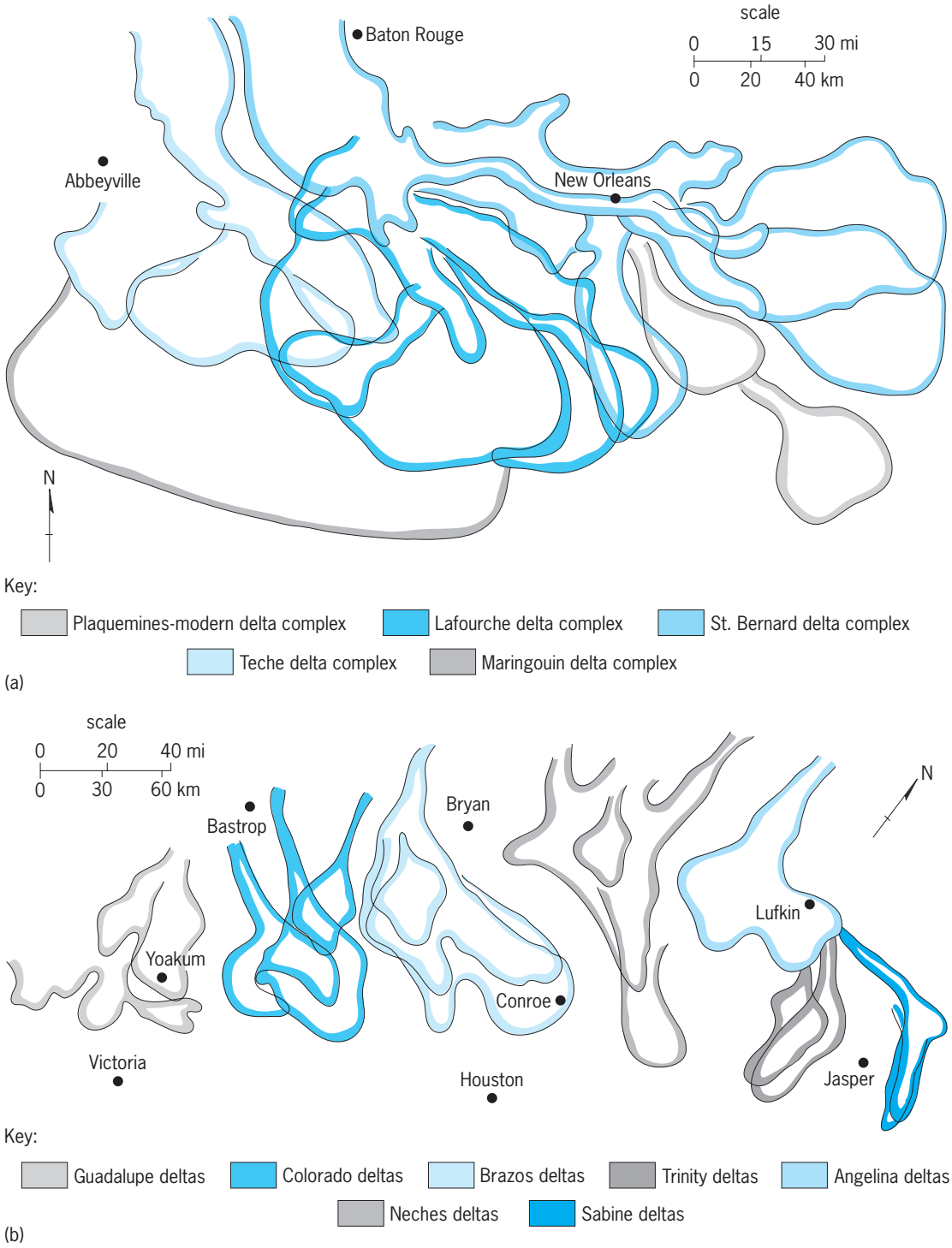

Fig. 3. Comparison of size, distribution, and arrangement of principal delta lobes of (*a*) the Holocene Mississippi delta system and (*b*) the Eocene Rockdale delta system of Texas. (*After W. L. Fisher and J. H. McGowen, Depositional systems in the Wilcox group of Texas and their relationship to the occurrence of oil and gas, Gulf Coast Ass. Geol. Soc. Trans., 17:105–125, 1967*)

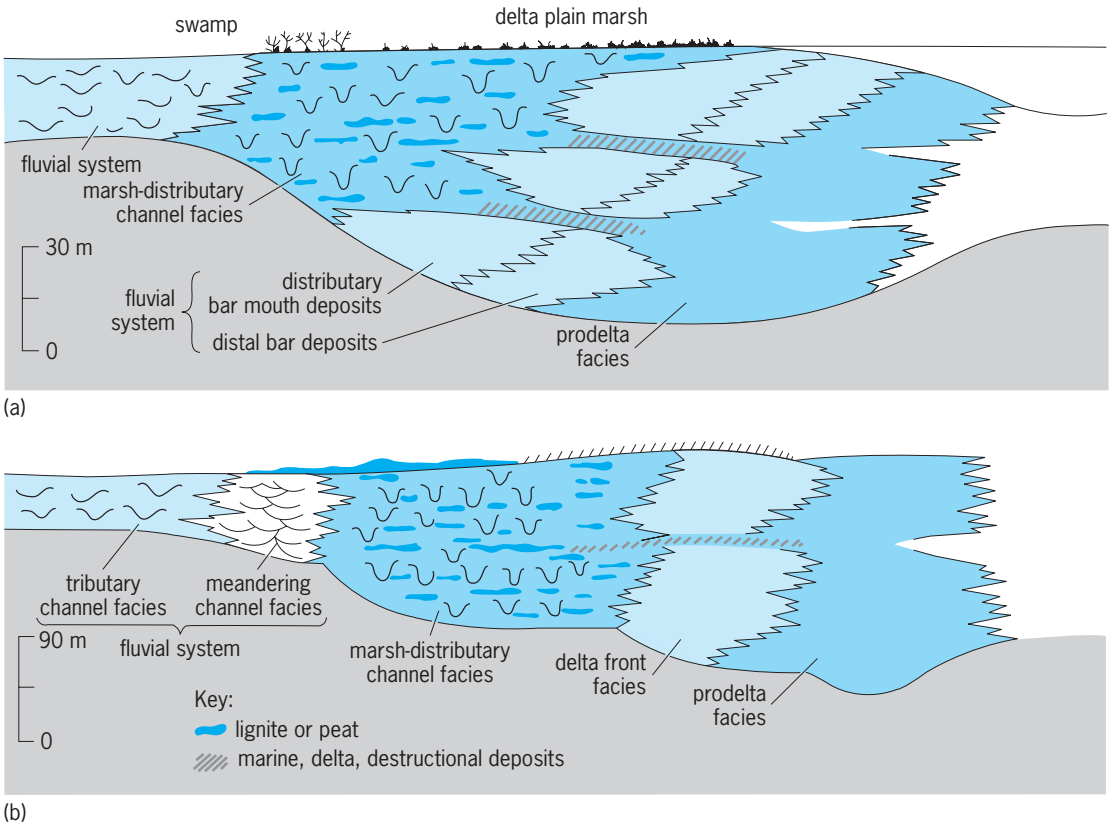

Fig. 4. Comparison of component facies of (*a*) the Holocene Mississippi delta system and (*b*) the Rockdale delta system of Texas. 1 m = 3.3 ft. (*After W. L. Fisher and J. H. McGowen, Depositional systems in the Wilcox group of Texas and their relationship to the occurrence of oil and gas, Gulf Coast Ass. Geol. Soc. Trans, 17:105–125, 1967*)

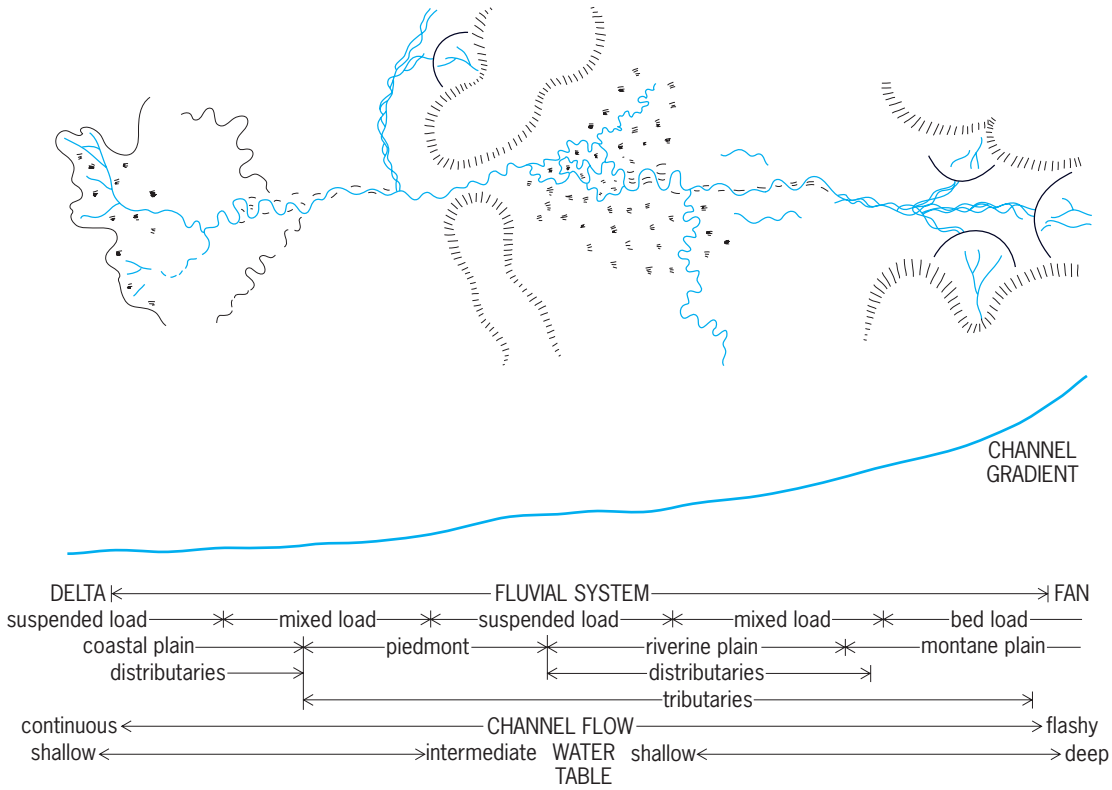

Fig. 5. Hypothetical fluvial systems. (*After W. E. Galloway, C. W. Kreitler, and J. H. McGowen, Depositional and Ground-Water Flow Systems in the Exploration for Uranium, University of Texas at Austin, Bureau of Economic Geology Special Publication, 1979*)

subdivisions is an assemblage of genetic facies, which are the fundamental units of depositional systems.

Terrigenous clastic systems, composed chiefly of sands and shales, embrace eight major systems: (1) fluvial or river systems, with braided stream systems, fine- and coarse-grained meandering stream systems, valley-fill stream systems, and straight distributary stream systems; (2) delta systems, with lobate and elongate river-dominated systems; (3) strike coastal systems, with barrier bar systems, strand-plain systems, bay and lagoon systems, and tidal flat systems; (4) fan or clastic wedge systems, with humid area and arid area alluvial fan systems and fan delta systems; (5) lacustrine systems, with open lake systems and closed or stratified lake systems; (6) continental eolian systems; (7) shelf systems, with wave- and tide-dominated systems; and (8) slope and abyssal systems, with tectonic up-lap slope systems, progradational off-lap slope systems, and destructive on-lap slope systems.

The biogenic-chemical systems consist of three major systems: (1) carbonate systems, with open and restricted carbonate platform systems, reef, bank, and atoll systems, winnowed shelf-edge systems, deep shelf systems, tidal flat and lagoonal systems, and slope systems; (2) glauconitic and authigenic shelf systems; and (3) evaporite systems, with supratidal (sebkha) systems and shallow shelf or platform and deep-water subaqueous systems.

By way of example, a hypothetical fluvial system may be considered (**Fig. 5**). A trunk channel or river evolves through bed-load (braided), mixed-load (meandering), and suspended-load phases in response to variations in gradient, sediment load, and discharge within the system. These phases, with resulting deposited genetic facies, may be characterized by differing attributes, recognizable or inferred, in both modern and ancient systems. Comparable kinds of variations, giving rise to distinct genetic facies, are recognizable in all major depositional systems.

W. L. Fisher

Depositional Environments

Depositional environments may be distinguished from erosional environments, in which erosion of the Earth's surface is taking place. Both depositional and erosional environments are of interest to geomorphologists. However, most attention to depositional environments has come from sedimentologists, particularly in order to understand the origin of sedimentary rocks. *See* EROSION.

Geophysical controls. The principal geophysical factors that control depositional environments are tectonics, climate, and topography.

Tectonics. Sediment is derived mainly from source areas that are actively undergoing uplift and erosion, and is deposited mainly in areas that are undergoing subsidence. Location of the source and basin of deposition is mainly controlled by large-scale geophysical processes acting within the Earth's mantle, so a major factor affecting the nature and distribution of sedimentary environments is the overall structural development, or tectonics, of the area. *See* PLATE TECTONICS.

Tectonics determines the major geological structure or setting of an environment of deposition, including the location and nature of the main areas undergoing uplift or subsidence. Areas with high relief, such as mountains and volcanoes, suffer rapid erosion and supply much more sediment to basins of deposition than larger areas of low relief. One investigation, for example, found that 82% of the suspended solids (mud) discharged by the Amazon River were supplied by the 12% of the drainage basin located within the Andes Mountains.

The major topographic slopes of continental surfaces are also largely determined by tectonic factors and tend to persist for long periods of geological time. The Mississippi River alluvial plain is an example. Large rivers have flowed south along the position of the modern Mississippi River plain for at least the last 200,000,000 years, and there have been many ancient precursors of the modern Mississippi delta and northern Gulf of Mexico coastal plain.

Tectonics is also one of the major factors controlling the rate of sedimentation in sedimentary environments. Local, short-term rates of accumulation of sediment may be very high. The modern Mississippi River delivers more than 1,100,000 tons (1,000,000 metric tons) of mud and sand to the delta every day; thicknesses of several meters of sediment may accumulate locally over periods of a few months. In the long term, however, the thickness of sediment that can accumulate in a shallow-water environment is determined by the regional rate of subsidence, which is determined by tectonic factors. A maximum of about 11 mi (18 km) of mainly shallow-water sediments has accumulated in the Mississippi region since the beginning of the Cenozoic, some 65,000,000 years ago. The long-term rate of sedimentation, therefore, has been only about 0.8 in. (20 mm) per year, which is still a very high rate of sedimentation averaged over long periods of geological time compared with most other areas.

The fact that the local, short-term supply of sediment is generally larger than the long-term rate of subsidence leads to rapid shifting of the centers of sedimentary deposition and therefore of sedimentary environments. In the Mississippi region, the position of the delta itself has switched abruptly several times in the last 5000 years. Sediment tends to fill in topographically low areas, not only in deltas but also in river and coastal plains and other environments supplied with large quantities of sediment. In one local area, rapid filling of a topographic low may be followed by a long period of nondeposition, when sediment is being deposited elsewhere. Sedimentation in the area starts again when subsidence (or a rise in sea level) once more produces a favorable environment for deposition. A thick sequence of sedimentary rocks, therefore, is frequently found to be composed of parts deposited rapidly during depositional episodes, separated by bounding surfaces or thin sedimentary units resulting from nondeposition

or very slow deposition over long periods of geological time.

Climate. A second important major control is climate. This includes the average temperature, the range of temperature variation, the aridity or humidity (ratio of evaporation to precipitation), and the magnitude and frequency of floods and storms. Climate in turn has an important influence on such physical factors as the salinity and energy of the environment (wind and water speeds and degree of turbulence, for example), as well as on the abundance and types of plants and animals.

Topography. In areas of subsidence and sedimentation, topography results from and controls sedimentary environments. Along a coastline of low relief, for example, spits and barrier islands are produced by waves generated in the open sea. Shallow lagoons on the landward side of barrier islands are protected from wave action by the islands themselves. The distinctive features of the lagoon environment are a result of a topography which has been produced by the accumulation of sediment in another sedimentary environment (the barrier island). *See* BARRIER ISLANDS.

Classification. Sedimentary environments can be classified into three categories: terrestrial, including alluvial fans, fluvial plains, sandy deserts, lakes, and glacial regions; mixed (shore-related), including deltas, estuaries, barrier island complexes, and glacial marine environments; and marine, including terrigenous shelves or shallow seas, carbonate shelves or platforms, continental slopes, continental rises, basin plains, ocean ridges, and ocean trenches.

Although the importance of tectonics and climate in controlling sedimentary environments is widely recognized, most classifications are based mainly on topography. Almost all distinguish terrestrial (subaerial or fresh-water) from marine environments, and also recognize an important group of mixed or shore-related environments. The above classification is just an outline: sedimentary environments are subject to almost endless subdivision, and many more complex classifications have been proposed.

Environmental controls. A number of major processes operate within environments and determine the types of sediment deposited in the environment.

Water depth. An important control is water depth, particularly in lakes and seas. Sediment deposited in shallow water is subject to constant agitation by wave action unless waves are excluded by some physical barrier. At extremely shallow depths, the sediment may be periodically exposed by changes in water level produced by tides, weather, or climatic fluctuations. Light easily penetrates to the bottom in very shallow water (unless the water is extremely turbid) so that aquatic plants, as well as animals, may flourish. Light is important not only for plants but for some organisms that live in close association with them. An example is hermatypic corals whose life cycle is so closely related to algae that typical coral reefs can develop only in water shallower than a few tens of meters. Criteria for shallow water depth in sedimentary deposits therefore include structures, such as ripple marks and cross-bedding, formed by strong currents, and particularly structures formed by wave action; structures produced by periodic exposure to the atmosphere, such as mud cracks; fossils or trace fossils (trails, burrows, and so on) indicating intense biological activity, and particularly algal growth (stromatolites, for example) and organically produced buildups of calcium carbonate (reefs and similar phenomena). Relatively deep water may be inferred from the absence of such phenomena, though there may be close resemblances between sediments formed in protected, quiet, shallow water and those deposited in deep water. *See* REEF; STROMATOLITE.

Energy of the environment. The degree of physical agitation by waves and currents, known as the energy of the environment, is frequently closely related to water depth. High-energy sedimentary environments include those in which the bottom sediment is constantly in motion, such as eolian sand dunes, the bottoms of river channels, beaches, and shallow marine environments characterized by strong waves or tidal currents. Fine sediment is maintained in suspension by the high level of turbulence in such environments and cannot accumulate unless there is an extremely high rate of supply. Typically the bottom sediment consists of well-sorted sand or gravel, with only a sparse growth of flora or fauna. Most organisms cannot flourish on a constantly moving sediment substrate. *See* DUNE; RIVER; TIDE.

Temperature. The temperature of an environment depends on the climate and water depth. Modern deep-sea environments maintain a steady temperature of a few degrees above 32°F (0°C), no matter what the latitude or climate at the surface. Temperature controls the solubility of calcium carbonate. Shallow seawater is supersaturated in calcium carbonate in warm climates but undersaturated in cool climates, so extensive deposits of limestone tend to be formed only in equatorial regions. Seawater also becomes undersaturated at depths of several thousands of meters in the oceans (below the lysocline), partly because of decreasing temperature but mainly because of an increased supply of carbon dioxide from the decay of organic detritus settling down from the surface layers of the oceans. As a consequence, sediments in the deepest parts of the ocean basins do not contain calcium carbonate, but consist of fine clay and siliceous planktonic skeletons (such as radiolarians). Environments sufficiently arid to evaporate large volumes of seawater are found only in hot climates, so temperature plays an important role in the formation of salt and gypsum deposits. At the other extreme, ice sheets and their associated sedimentary environments (moraines, glacial outwash fans, and so on) are found only in cold climates.

Salinity. The salinity of an environment depends on the aridity of the climate and on a source for the dissolved salts. Outside of restricted supersaline environments (playa lakes, coastal lagoons, and supratidal flats or sebkhas), most aqueous environments can be categorized as marine (with a salinity of 35 parts per thousand), fresh (with a salinity of less than

0.2 part per thousand), or brackish. Brackish waters are most frequently formed by the mixing of fresh water and seawater, for example, in estuaries, coastal lagoons, or deltas. Salinity in these environments frequently fluctuates widely, a condition very unfavorable to most forms of animals and plants. Therefore, brackish water environments tend to be characterized by their own specialized flora and fauna. Salinity is also important in determining the relative density of water masses. River water is less dense than seawater and tends to spread out over it as a surface layer when rivers enter estuaries or emerge from deltaic distributaries. A consequence is that river and seawater do not mix easily, and that a distinctive estuarine pattern of circulation is set up, with saline waters moving landward under surface fresh or brackish waters moving offshore. Within the ocean basins themselves much smaller differences in both salinity and temperature lead to stratification of water, with relatively little mixing between layers. The denser water masses generally move slowly along the bottom from the poles toward the equatorial regions and become progressively enriched in nutrients such as phosphates. Where the deeper waters are forced to the surface (zones of upwelling, which are generally found along the western equatorial margins of the continents), they may support prolific organic growth and produce sedimentary deposits rich in phosphates or organic matter. *See* UPWELLING.

Chemical factors. There are other chemical aspects of the environment that are of great importance for determining the mineralogy of the sediment. The two most commonly measured and cited indexes are the pH (hydrogen ion concentration, or relative acidity or alkalinity) and Eh (oxidation-reduction potential). Sea-water is mildly alkaline and oxidizing; river waters are generally mildly acidic and oxidizing. Highly acidic waters may be found in swamps and marshes, and highly alkaline waters in some shallow lakes in arid climates. Reducing waters can be found only in environments cut off from the atmosphere—for example, in the lower levels of some stratified water masses, such as in some lakes and estuaries. Reducing and acid conditions are, however, very commonly developed just below the sediment-water interface in muds deposited in both terrestrial and marine environments. Whatever the original color of such muds, they generally become black or gray below the surface as a result of the reduction of iron minerals to fine-grained iron sulfides. Conversely, sediment deposited on alluvial fans in arid regions remains in contact with air or oxygenated pore waters (in the vadose zone, above the water table) for long periods of time so that detrital iron-bearing minerals are oxidized to produce fine-grained red iron oxides. Thus the chemistry of the original depositional environment can be inferred from the sedimentary deposit only after the nature of the postdepositional (diagenetic) changes has been determined.

Biological factors. Besides the major environmental controls listed above, biological factors exert a very strong influence on the nature of many depositional environments. The compositions of the atmosphere and oceans are themselves strongly influenced by biological activity. There would be little or no oxygen in the atmosphere if it were not for the photosynthetic activity of plants. Deposition of calcium carbonate and silica in lakes and the oceans takes place largely through the action of plants and animals, and organic matter deposited along with mineral particles is largely responsible for the development of reducing conditions within sediments after deposition.

Vegetable material accumulates in swamps to form peat and coal, and fine organic detritus settles with marine muds and is the ultimate source of oil and gas. Both terrestrial and aquatic plants exert a trapping and binding action that tends to immobilize sedimentary particles, as, for example, when coastal sand dunes or tidal flats become stabilized by the growth of salt-tolerant grasses. Terrestrial vegetation plays an important role in rock weathering, and there is some evidence that less clay was produced by weathering in the period (pre-Silurian) before the land was colonized by plants. Mangrove swamps, coral reefs, and tropical rainforests are just a few examples of environments strongly influenced by a combination of a suitable climate and flora and fauna. The abundance of plants and animals in many environments is such that sediment deposited there may be strongly modified before ultimate burial. Sediment may be eaten for its organic content by deposit- or filter-feeding organisms, or may be disturbed by plant roots or burrowing animals and insects.

Ancient environments. In looking back over the geologic record of past environments, sedimentologists see some signs that ancient environments were not always the same as those on the modern Earth. There seem to have been periods of time (for example, Jurassic-Cretaceous) in which there were not large ice sheets, and thus no extensive glacial environments. Probably the bottom waters of the oceans were not as cold then as they are now, and there is some evidence that ocean circulation was less vigorous, so that stagnant, anoxic conditions could develop over large parts of some deep-sea basins.

Calcium carbonate is widely deposited in the ocean basins today (except in the deepest waters below the lysocline) largely because of the abundance of carbonate-secreting plankton, such as foraminifera. Such pelagic limestones are also common in some Cenozoic and Mesozoic sediments formed in deep water, but are rare in earlier deposits because such organisms did not evolve until Jurassic times. *See* FORAMINIFERIDA.

Conversely, some deposits abundant in the geologic record appear to be only sparsely represented by modern environments. Examples are the very extensive shallow-water limestones of the Paleozoic and marine evaporites. Some marine evaporites may have been formed by geologically unusual events, such as the complete desiccation of the Mediterranean Sea which has been suggested by some

geologists to explain the Miocene evaporites in that region.

Although the nature of carbonate-precipitating animals and plants, and therefore the nature of organic reefs and other environments of carbonate accumulation, has changed throughout geologic time, there remain some strong general resemblances between ancient and modern environments. The proportions of different environments have changed, but the main types of environments have remained the same. Only in Precambrian times is there evidence that sedimentary environments differed radically from those on the modern Earth. In the earliest Precambrian, volcanism seems to have been more active than at present and stable shallow marine environments less common. At the beginning of the Proterozoic, some 2 billion years ago, when shallow marine environments first became abundant, they were the site of extensive deposition of formations rich in iron. Modern analogs of these iron formations are hard to find, and many geologists believe that their formation is connected to the presence within the atmosphere of the time of only a very low abundance of oxygen. Such interpretations remain controversial. The guiding rule for most geologists attempting an environmental interpretation of ancient sedimentary rocks is actualism; that is, the present is the key to the past. *See* MARINE SEDIMENTS. Gerard V. Middleton

Bibliography. H. Blatt, G. Middleton, and R. Murray, *Origin of Sedimentary Rocks*, 2d ed., 1980; R. J. Chorley et al., *Geomorphology*, 1985; R. A. Davis, Jr., *Depositional Systems*, 2d ed., 1991; W. F. Fritz and J. N. Moore, *Basics of Physical Stratigraphy and Sedimentology*, 1989; H. E. Reineck and I. B. Singh, *Depositional Sedimentary Environments—with Reference to Terrigenous Clastics*, 1992.

Dermaptera

An order of insects commonly known as earwigs, in reference to an unfounded superstition that they crawl into the ears of sleeping persons. Earwigs are elongate and flattened, with a pair of forcepslike appendages, or cerci, at the end of the abdomen (see **illus.**). They are 0.4–1.2 in. (10–30 mm) long, and have chewing mouthparts, gradual metamorphosis, long legs, and shortened, leathery forewings. Some species are flightless, while others are active fliers with well-developed hindwings that fold elaborately under the forewings when the insect is at rest. Most earwigs are brown or black.

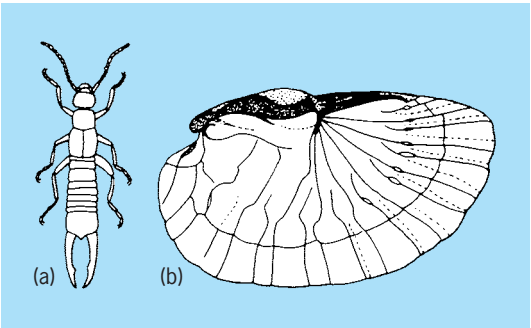

Some features of Dermaptera. (*a*) Adult of Labidura. (*b*) Hindwing of *Forficula*. (*After C. P. Brues, A. L. Melander, and F. M. Carpenter, Classification of Insects, Harvard University Press, 1954*)

Earwigs are usually nocturnal and spend the daylight hours under bark or stones, in moist cracks and crevices, or in leaf litter at the soil surface. Some are attracted to lights. Most species scavenge on decaying animal or vegetable matter. A few species are predators on smaller insects. The widely introduced European earwig, *Forficula auricularia*, is an occasional pest of buildings, and crop plants, especially seedlings and floral crops.

Some earwig species display complex mating behavior in which the male grasps the female with his cerci. Females of some species show parental care, guarding eggs and nymphs against predators.

Worldwide, there are about 1100 species of earwigs, and most of these species are tropical. Of the 22 species known from North America, half are introduced. Some apparently arrived in sailing ships' ballast over 200 years ago. Fossil earwigs are known from the Jurassic Period. *See* INSECTA. David J. Horn

Bibliography. R. H. Arnett, Jr., *American Insects: A Handbook of the Insects of America North of Mexico*, 2d ed., 2000; D. J. Borror, C. A. Triplehorn, and N. F. Johnson, *An Introduction to the Study of Insects*, 6th ed., 1997.

Dermatitis

Inflammation of the skin with redness and scaling, or if acute, with blisters, edema, and formation of a crust; also known as eczema.

Human. Of the several distinct types of dermatitis, each has unique etiology, clinical characteristics, pathology, and treatment.

Atopic dermatitis. Atopic dermatitis usually begins in infancy but may continue into adult life. The eruption is characterized by red patches accompanied by intense itching. The patches become leathery and abraded with continued rubbing and scratching. The lesions often become secondarily infected, leading to moist discharge and crusting. Although any anatomic site may be affected, the classic locations in infants are the extremities, face, and scalp. In older children and adults, the inside of the elbows, the back of the knees, wrists, eyelids, and neck are most often involved.

Often familial, atopic dermatitis may be associated with asthma or allergic rhinitis. As with asthma and allergic rhinitis, about 80% of individuals with atopic dermatitis have elevated levels of immunoglobulin E (IgE). A wide range of other immunologic abnormalities have been found.

The disease has a genetic predisposition, but its expression is modified by environmental factors. The eruption may be made worse by low temperatures,

Some common causes of allergic contact dermatitis

Chemical or material	Product	Anatomic site
Nickel	Jewelry	Earlobes, neck, wrist
Paraphenylenediamine	Hair dye	Scalp, ears, face
Glyceryl monothioglycolate	Permanent-wave substances	Scalp, ears, face
Balsam of Peru	Fragrance	Eyelids, neck
Cinnamic alcohol	Fragrance	Eyelids, neck
Thimerosal	Contact lens solution	Eyelids
Quaternium-15	Preservative (cosmetics)	Face, hands
Imidazolidinyl urea	Preservative (cosmetics)	Face, hands
Thiuram, mercapto	Rubber	Hands, feet
Potassium dichromate	Cement, leather, metals	Hands, feet, generalized
Formaldehyde	Nail polish, clothing	Eyelids, generalized
Acrylates	Artificial nails	Eyelids
Cinnamic aldehyde	Toothpaste	Perioral
Colophony	Adhesives	Feet, hands
Neomycin, cortisones, benzocaine	Topical medications	Wherever applied

decreased humidity, contact with wool, immunizations, and upper respiratory infections. Eggs, peanuts, milk, wheat, fish, soy, and chicken may accelerate the disease. Nickel, either in direct contact with the skin or via dietary intake, may play a role in the disease.

Individuals with atopic dermatitis have an increased susceptibility to certain infections. Over 90% of eczematous lesions contain *Staphylococcus aureus*, and many individuals are colonized with the organism. Widespread fungal infections and warts occur. Secondary infection with *Herpes simplex* virus, known as eczema herpeticum, is perhaps the most serious complication of atopic dermatitis. *See* HERPES; STAPHYLOCOCCUS.

A skin biopsy in atopic dermatitis shows epidermal and dermal changes. In earlier lesions, there is mild thickening of the epidermis with intercellular edema, and superficial dermal inflammation. With chronic rubbing, the edema becomes less prominent and thickening of the epidermis is more marked.

Treatment of atopic dermatitis includes oral antihistamines, emollients, topical corticosteroids (particularly ointments), and for more severe cases, systemic corticosteroids. Avoidance of irritants (wool, chemicals, harsh soaps and detergents, perfumes), extremes in temperature and humidity, overbathing, and certain foods is imperative. Mild cleansers are preferred. For weeping and crusted lesions, warm compresses with tap water or aluminum acetate solution (Burow's solution) are helpful. When secondary infection by *S. aureus* exists, oral antibiotics or topical mupirocin, an antibacterial ointment, should be used. *See* ANTIHISTAMINE; STEROID.

Contact dermatitis. Contact dermatitis is a reaction that causes acute itching, redness, swelling, large blisters, and in chronic cases, red, scaly papules (raised bumps) and plaques (abnormal flat areas). The two forms are irritant contact dermatitis and allergic contact dermatitis. In irritant contact dermatitis the eruption is caused by a nonallergic reaction resulting from exposure to an irritating substance. Allergic contact dermatitis is an immunologic reaction in persons who have been previously exposed (sensitized) to the allergen, the classic example being poison ivy. Linear streaks and geometric shapes suggest contact dermatitis. The lesions occur 24–48 h after exposure to an irritant or allergen.

The anatomic distribution of the skin lesions is important in determining the etiologic agent. Common irritants include solvents, cutting oils, petroleum products, plastics and resins, food products, and soaps and cleaners (see **table**). The diagnosis of contact dermatitis is made by a careful history, physical examination, and in some cases, patch testing to identify the allergen.

In acute contact dermatitis the histopathology shows edema with formation of blisters and sparse dermal inflammation. Over time, there is thickening of the epidermis and greater inflammation. Increased exposure to irritants may also cause localized death of skin tissue and ulceration.

The key to treatment is removal of the offending agent. Antihistamines, emollients, and topical and systemic corticosteroids are helpful. For treatment of cases of acute allergic contact dermatitis (for example, poison ivy), most individuals require oral prednisone.

Stasis dermatitis. Scaly, reddened, ill-defined plaques on the lower legs, often accompanied by swelling, hyperpigmentation, loss of hair, or ulceration, are common symptoms of stasis dermatitis. The changes are associated with continued poor venous circulation. Thickening of the outermost layer of the epidermis (hyperkeratosis), edema, and thickening of the epidermis occur. Within the outermost layer of the dermis, there is inflammation, vascular changes, hemorrhage, and iron pigment. Treatment consists of improving circulation, reducing edema (by means of elevation, compression, and diuretics), and administering topical corticosteroids.

Seborrheic dermatitis. Seborrheic dermatitis is characterized by reddish to yellow greasy scale that frequently forms on the face (eyebrows, nasolabial folds, forehead), ears, scalp, upper back, central chest, and anogential region. In contrast to atopic and contact dermatitis, itching is uncommon.

The etiology of seborrheic dermatitis is multifactorial. Genetic and hormonal factors, emotional stress, and the fungus *Pityrosporum ovale* have all been

proposed. Skin biopsies of individuals with seborrheic dermatitis reveal thickening of the epidermis, mild edema, the presence of neutrophils within the outermost layer of the epidermis, and superficial inflammation of the dermis. *See* FUNGI.

Topical corticosteroids are the mainstay of treatment for seborrheic dermatitis. In cases in which *P. ovale* is involved, topical antifungal agents (for example, ketoconazole cream) are useful. Selenium sulfide, tar, zinc pyrithionate, ketoconazole, and resorcin shampoos are useful therapies for scalp involvement. Topical or oral antibiotics are required in cases with secondary bacterial infection. In generalized or recurring cases, systemic corticosteroids may be necessary.

Exfoliative erythroderma. Exfoliative erythroderma is characterized by widespread, warm redness and scaling. Nail degeneration and loss, hair loss, fever, chills, and enlargement of the lymph nodes may also occur. The eruption may be due to an underlying primary skin disorder, such as dermatitis (seborrheic, contact, or atopic) or psoriasis; drug allergy; or leukemia or lymphoma, particularly cutaneous T-cell lymphoma. A careful history, physical examination, and skin biopsy are helpful in determining the etiology. *See* LEUKEMIA; LYMPHATIC SYSTEM; LYMPHOMA.

Individuals with exfoliative erythroderma are best managed in a hospital setting. Because heat, water, and protein losses occur through damaged skin, careful monitoring of fluid balance, protein losses, and body temperature is imperative. The application of wet dressings or topical corticosteroids and administration of oral antihistamines are important aspects of treatment. Systemic corticosteroids are useful if dermatitis or a drug-induced eruption is involved, but they may exacerbate psoriasis. Systemic antibiotics should be used in cases where a secondary bacterial infection has developed. All other medication should be discontinued if possible. *See* INFECTION.

Lisa M. Cohen; Kenneth A. Arndt

Flea allergy dermatitis. Flea allergy dermatitis, sometimes called flea-bite dermatitis, affects the skin of small pets (dogs, cats, and ferrets), and results from an allergic reaction to protein substances deposited on (or under) the surface of the skin at the time of flea feeding. It is by far the most common cause of allergic dermatitis in pet animals and a major cause of itchy skin in geographic regions where fleas are found.

Pathogenesis. The most common species of flea that parasitizes dogs and cats is *Ctenocephalides felis felis.* An allergic reaction may occur in dogs of any age following the exposure to the flea saliva. This sensitization does not require a genetic predilection as do other types of allergies, but may be antibody induced (immunoglobulin E), cell mediated (delayed reaction), or brought about by other mechanisms. Dogs only several weeks old can develop an allergic reaction to fleas, while other dogs or cats can live with fleas for many years before developing the allergy. *See* ANTIBODY.

The flea remains on the host, leaving only to find a new, more desirable host; it rarely returns to the environment. Since the adult flea is geared primarily for perpetuation of the species, the female consumes quantities of blood to sustain nourishment for egg production, producing 1000–1500 eggs over a period of 100 days.

Clinical signs. Continuous or intermittent exposure to fleas perpetuates the clinical signs, the most prominent being itchy skin. Even a limited flea infestation may elicit profound clinical signs. On dogs, red, scaly skin over the top of the lower back by the tailhead appears with a typical pattern of self-inflicted hair loss as well as lesions on the abdomen and rear legs. Secondary *Staphylococcus* infections are common in chronic cases that do not receive appropriate treatment. Scales and crusts are often observed associated with a variable rash. Cats may have a more varied array of clinical symptoms. Focal crusts (miliary dermatitis) develop on the lower back, tailhead region, and thighs. Comparable lesions may be found around the neck. Often, hair loss with minimal dermatitis affecting the abdomen is caused by excessive licking and hair pulling in that area. The disease may progress to a generalized pattern, but usually spares the head, feet, and forelegs. Self-inflicted trauma to the skin is produced by removal of the fleas through scratching, biting, licking, and chewing.

Diagnosis. The diagnosis of a flea allergy is based on the history of the animal and on clinical examination. Confirmation is made by obtaining a positive reaction to a commercial flea antigen injected in the skin, although this may be observed in only 90% of the cases. Response to therapy effective for flea control provides a tentative diagnosis. Finding evidence of fleas may not be possible. Identification and removal of fleas is best accomplished by using a fine-toothed comb, with 24–26 teeth per inch, for a short- or medium-coated pet. Small specks of black material present on the skin surface may represent flea dirt (feces), which will turn red (representing the whole blood ingested by the flea) if applied to a water-moistened cottonball and allowed to stand for 5–8 min.

Treatment. Treatment of flea allergy is directed at three areas: flea control, which includes removal of the flea from the infested pet, treatment of associated animals, and aggressive treatment of the animals' habitat; symptomatic treatment for the itching; and treatment for secondary complicating factors such as bacterial skin infection (pyoderma) or dry skin.

A variety of products are commercially prepared for animal treatment: topical products include shampoos, pour-on insecticides, sprays, powders, and spot treatments and collars; systemic treatments include oral medications and spot treatments, where percutaneous absorption provides systemic concentration of the insecticide, which the flea ingests during feeding. Most topical treatments have a short duration of activity with poor residual killing or repellent effectiveness. Systemic products also have limited activity for the flea-allergic animal because

the allergic reaction is induced by flea feeding. Some products (organophosphates and carbamates) may be toxic with repeated applications. Safer ingredients include the botanicals derived from plant sources. Examples include pyrethrins (chrysanthemum), rotenone (derris plant), or *d*-limonene or linalool from citrus produce (orange peel).

Insect growth regulators are successful because they interrupt the flea life cycle in preadult forms (eggs and larvae). Although they are principally used for premises treatment, some products have been developed for application on the animal or systematic administration.

Adjunctive treatment is directed toward topical therapy to reduce the dryness, systemic and topical drugs to decrease the itching, and specific drugs for treating concurrent diseases such as bacterial skin infections. Systemic anti-inflammatory steroid treatment may be necessary if the itching is intense and aggressive biting, licking, chewing, and scratching are present. Long-term therapy with steroids requires proper selection and interval of therapy to avoid the many potential side effects of these drugs. Antihistamines may provide some relief for itchy skin, but have a less predictable effect than steroids. However, neither antihistamines nor steroids will be effective if flea infestation on the flea-allergic pet continues. *See* ALLERGY; SIPHONAPTERA; SKIN DISORDERS.

J. M. MacDonald

Bibliography. AAD Committee on Guidelines of Care, Guidelines of care for atopic dermatitis, *J. Amer. Acad. Dermatol.*, 26:485–488, 1992; B. L. Blagburn et al., Flea control, *North American Veterinary Conference Proceedings*, Orlando, Florida, January 1994; A. A. Fischer, *Contact Dermatitis*, 3d ed., 1986; I. M. Freedburg, A. Z. Eisen, and K. Wolff, *Fitzpatrick's Dermatology in General Medicine*, 5th ed., 1998; C. E. Griffin, K. W. Kwochka, and J. M. MacDonald, *Current Veterinary Dermatology*, 1993; G. H. Muller, R. W. Kirk, and D. W. Scott, *Muller & Kirk's Small Animal Dermatology*, 6th ed., 2000.

Dermoptera

An order of mammals comprising one family (Cynocephalidae), one genus (*Cynocephalus*), and two species—*C. variegatus* (Malayan colugo) and *C. volans* (Philippine colugo). These mammals inhabit the rainforests of Southeast Asia from the Malay Peninsula and Thailand to the Philippine Islands, Java, and Borneo.

General characteristics. Colugos are about the size of a large squirrel or domestic cat. Their most distinctive feature is a gliding membrane (patagium) that attaches to the neck and the sides of the body. It extends from the tips of the fingers and toes to the tip of the tail. It gives the appearance of an animal wearing a fur coat that reaches to its toes. When at rest, the furry membrane hangs in loose, voluminous folds. No other volant mammal possesses such an extensive flight membrane. Although colugos are also known as flying lemurs, they cannot fly. (Bats, order Chiroptera, are the only flying mammals.) By spreading the patagium, the colugo is able to perform controlled glides of 100 m (330 ft) or more with little loss of height. The head is broad with a blunt muzzle and short, rounded ears. The arms, legs, and tail are long and slender. The limbs terminate in strong, sharp, recurved claws. These nocturnal mammals have large, prominent eyes and stereoscopic vision. The single pair of mammae in females are located at the sides of the body, almost in the armpits. The flight membrane and the dorsal surface are mottled brownish-gray with males being more brownish. Scattered white spots are present on the dorsal surfaces of the Malayan colugo. The mottled coat blends with the bark of trees. The underparts are paler. The dental formula is I 2/3, C1/1, PM 2/2, M 3/3 × 2 for a total of 34 teeth. The lower incisors have been modified into "comb teeth" with as many as 20 prongs radiating from one root. These specialized teeth are unlike those in any other mammal and may serve as food strainers, scrapers, and/or as an aid in grooming. Adults have head-body length of 340–420 mm (13–16.5 in.), a tail of 175–270 mm (7–11 in.), and a weight of 1.0–1.75 kg (2–4 lb). The span of the patagium is 70–120 cm (2.3–4.0 ft).

Malayan colugo (*Cynocephalus variegatus*), also known as the Malayan flying lemur. (*Photo by Ampika Korcharernkit; courtesy of Suwit Punnadee, Director of the Gibbon Rehabilitation Project, Wild Animal Rescue Foundation of Thailand*)

Habitat. Colugos are found in both lowland and mountainous areas in Southeast Asia, where they inhabit primary and secondary forests as well as coconut groves and rubber plantations. They are totally arboreal and seldom, if ever, descend to the ground. They are unable to stand erect when on the ground and are clumsy climbers. When feeding, roosting, or moving about in the trees, they do so in an upside-down position like a sloth.

Behavior. These nocturnal mammals spend the daylight hours either in a hole or cavity of a tree or hidden in dense foliage. The diet of these herbivores consists primarily of foliage such as leaves, shoots, buds, and flowers. Since leaves have a relatively low nutrient content, large quantities must be consumed.

Soft fruits may occasionally be eaten. Sufficient water is obtained through food and by licking wet leaves. Adults are normally solitary but have broadly overlapping home ranges. They are highly territorial, not only with respect to their sleeping place, but also with respect to their foraging trees.

Reproduction. Breeding may occur throughout the year with gestation being approximately 60 days. Most litters consist of a single young. The newborn is very undeveloped, similar to a marsupial, and is carried on the mother's belly, even when she is gliding. Near the tail, the mother's patagium can be folded to form a pocket for the young. Weaning occurs at about 6 months of age. It may require 3 years for colugos to reach adult size.

The slow rate of reproduction and maturation along with loss of forest habitat through logging and conversion to agricultural use has led to the Philippine colugo being classified as vulnerable by the International Union for the Conservation of Nature (IUCN). This species is also hunted for its meat and fur. The major natural predator is the Philippine monkey-eating eagle (*Pithecophaga jefferyi*).

Donald W. Linzey

Bibliography. D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford Limited, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Derrick

A hoisting machine consisting usually of a vertical mast, a slanted boom, and associated tackle (see **illus**.). Derricks have a wide variety of forms. The mast may be no more than a base for the boom; it may be a tripod, an A-frame, a fixed column, and so on. Fixed stays may guy it in place. The boom may be fixed, it may pivot at the base of the mast, it may swing horizontally from near the top of the mast, or it may be omitted. The derrick may be permanently fixed, temporarily erected, or mobile on a cart or truck.

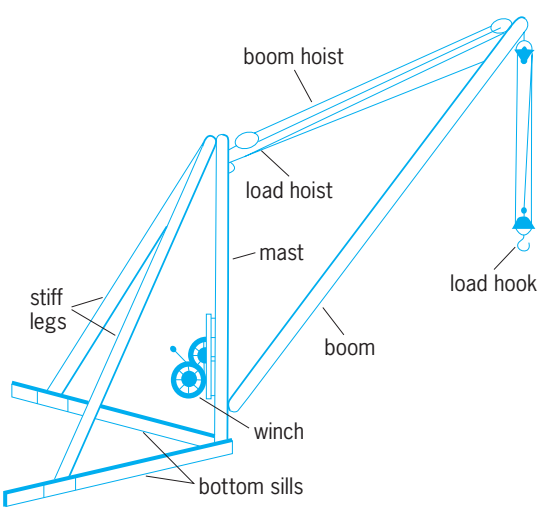

A boom derrick equipped with a swinging mast and anchored legs.

Derricks are widely used in construction, in cargo handling, and in shops. Their lifting action is intermittent compared to bucket conveyors, and their coverage is limited by the reach of the boom. A manual or powered winch provides the lifting action by coiling in the running tackle, the load swinging free as it rises. *See* BULK-HANDLING MACHINES; HOISTING MACHINES; OIL AND GAS WELL DRILLING.

David O. Haynes; Frank H. Rockett

Bibliography. W. E. Rossnagel, R. L. Higgins, and J. MacDonald, *Handbook of Rigging: For Construction and Industrial Operations*, 4th ed., 1988; H. I. Shapiro, J. P. Shapiro, and L. K. Shapiro, *Cranes and Derricks*, 3d ed., 1999.

Derris

A genus of tropical shrubs belonging to the legume family (Leguminosae). These plants, with their long branches climbing over other vegetation, occur as members of the jungle undergrowth in Malaysia. Extracts of the roots of *Derris elliptica* have long been used by the natives as an arrow poison and to stupefy fish so they can be caught more easily. Derris root is an excellent insecticide, being harmful to both chewing and sucking insects, but not poisonous to human beings. The insecticidal ingredient of derris root is a white, crystalline substance, which is called rotenone. *See* ROSALES.

Perry D. Strausbaugh; Earl L. Core

Descriptive geometry

A mathematical-graphical procedure that has for its purpose the visualization of structures and their exact representation in drawings. After analysis of the

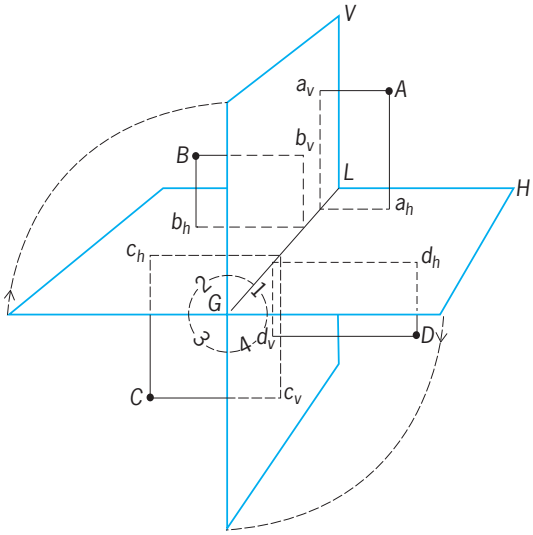

Fig. 1. Planes of projection. (*After G. J. Hood and A. S. Palmerlee, Geometry of Engineering Drawing, 4th ed., McGraw-Hill, 1958*)

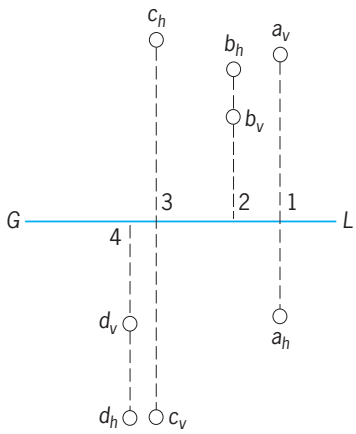

Fig. 2. Projection of points onto projection planes. (*After G. J. Hood and A. S. Palmerlee, Geometry of Engineering Drawing, 4th ed., McGraw-Hill, 1958*)

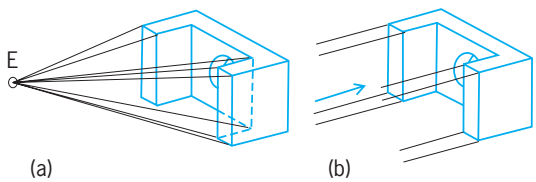

Fig. 3. Methods of viewing objects. (*a*) Perspective viewpoint (converging rays). (*b*) Orthographic viewpoint (parallel rays). (*After G. J. Hood and A. S. Palmerlee, Geometry of Engineering Drawing, 4th ed., McGraw-Hill, 1958*)

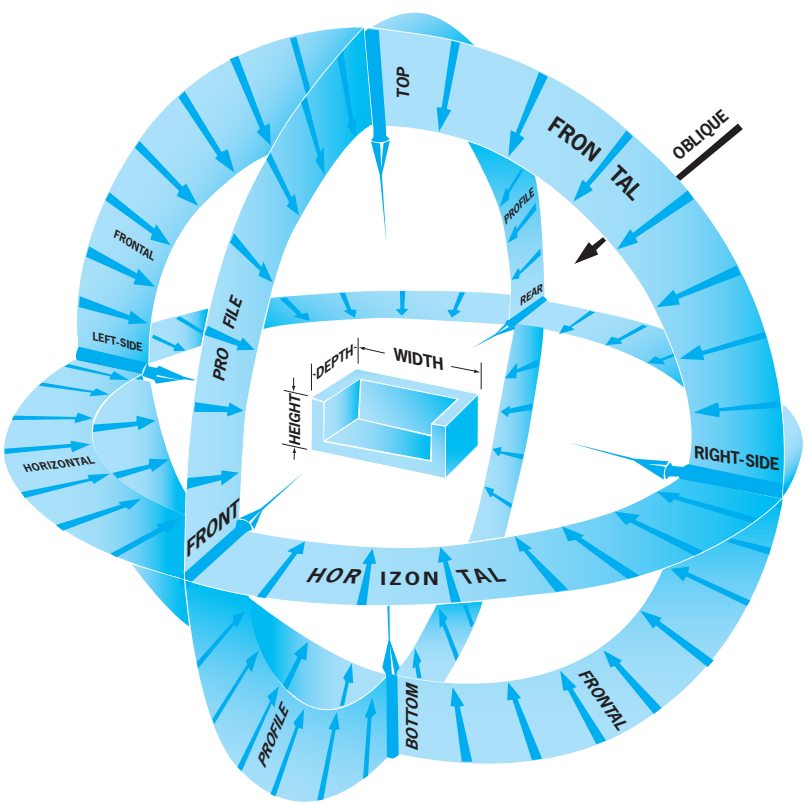

Fig. 4. A series of viewing positions for auxiliary views. (*After G. J. Hood and A. S. Palmerlee, Geometry of Engineering Drawing, 4th ed., McGraw-Hill, 1958*)

structure, each element is shown in the drawing in its exact geometrical relation to the other elements.

There are two basic methods of descriptive geometry: the projection method and the direct method. The two methods differ as regards the atttitude of mind toward the structure and toward the drawing that represents the structure.

Projection method. Gaspard Monge (1756–1818), a French mathematician, originated the projection method of descriptive geometry. While working as a designer for the French government, he was given the job of making plans for a proposed fortress. This was a tedious process and involved long calculations. He invented graphical solutions and completed the plans in such a short time that at first the commandant refused to accept them. For a long time the graphical process was kept a state secret, finally being revealed about 1795.

In the projection method the horizontal projection plane *H* and the vertical projection plane *V* intersect in the line *GL*, which is called the ground line (**Fig. 1**). These two projection planes divide space into four quadrants, or angles, as numbered in Fig. 1.

Point *A*, in the first quadrant, is projected onto the horizontal plane at a_h by means of a projection line perpendicular to the *H* plane, and onto the vertical plane at a_v by means of a projection line perpendicular to the *V* plane. The projections of the points *B, C,* and *D* in the other quadrants are located in a similar manner. If desired, a profile projection plane, perpendicular to the *H* and *V* planes, may be introduced. Right-angle projection, as described above, is called orthographic projection.

To represent horizontal and vertical projections on a flat sheet of paper, the planes are conceived as being hinged along the ground line and brought together by closing the second and fourth quadrants. Projections of *A, B, C,* and *D* then appear in a single plane (**Fig. 2**). The *H* and *V* projections of a point are always in the same perpendicular to the ground line. The usual custom in the United States is to draw objects as if they had been projected from a position in the third quadrant; this is called third-angle projection. First-angle projection is standard in some countries and professions.

There are two general types of views, perspective and orthographic (**Fig. 3**). A perspective view of an object is observed from a fixed station point, or point of view, by means of converging rays of light that meet at the eye of the observer. An orthographic view of an object is observed in a chosen direction by means of parallel rays of light. As shown in Fig. 3*b*, the observer may be imagined at a distance considered to be infinity.

Direct method. In the direct method the attention is focused on the visualized structure or object. Each view of the object is obtained by looking at the object in a definite direction. The view is orthographic. A view never is considered as two-dimensional or as projected or drawn on a plane. This is the way the engineer who makes and reads the drawings thinks of views.

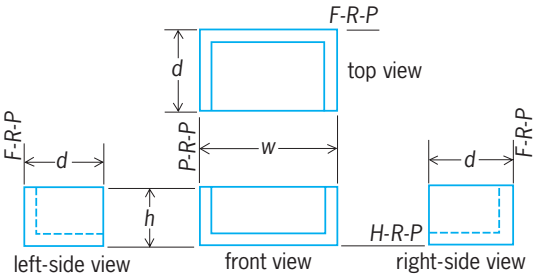

Fig. 5. Principal views of an object and some reference planes. (*After G. J. Hood and A. S. Palmerlee, Geometry of Engineering Drawing, 4th ed., McGraw-Hill, 1958*)

Orthographic views may be classified into three types: principal views, auxiliary views, and oblique views. The object can be viewed from any direction around three rings—horizontal, frontal, and profile (**Fig. 4**). The rings represent three mutually perpendicular planes. The intersections of the rings define three mutually perpendicular directions from which six principal views are observed: front and rear, top and bottom, right and left sides.

An auxiliary view can be observed around any ring in a direction perpendicular to one, and only one, of the directions in which principal views are observed.

All views other than principal or auxiliary views are oblique. In Fig. 4 a single arrow, marked oblique, indicates one of the infinite number of directions in which oblique views are observed.

Dimensions of structures. All structures occupy space and have three dimensions: width, height, and depth. These are measured in three mutually perpendicular directions (Fig. 4).

Grouping of views. In **Fig. 5** are shown four views of the object pictured in Fig. 4. Each view is placed in relation to its adjacent view in the position from which it was viewed. Edge views of three reference planes are indicated in Fig. 5. For example, *H-R-P* is the edge view of a horizontal reference plane from which all vertical height dimensions of the object may be measured. Similarly, all depth dimensions are measured from the frontal reference plane *F-R-P*, the edge view of which is seen in the two side views and in the top view.

Reference planes. **Figure 6** shows the front and top views and four auxiliary elevations of a bearing. In the front view a horizontal reference plane is taken in the auxiliary elevations. The height dimension, or elevation, of each point in the front view, above or below the reference plane, is measured in the front view and then is transferred to each of the auxiliary elevations.

In practice, views are drawn from orientations that show the true sizes of selected lines or surfaces; that is, views are drawn perpendicular to such lines or surfaces. Thus, a triangular pyramid might be viewed from four directions, each of which shows one face in its true shape.

Reading the drawing. The object in a drawing is regarded as stationary. As the reader looks at first one view and then another, he gradually obtains a detailed mental picture of the object. The reader imagines that he moves around the object, viewing it successively in each direction indicated by the view.

Geometrical elements of structures. Points, lines, and surfaces are the geometrical elements of

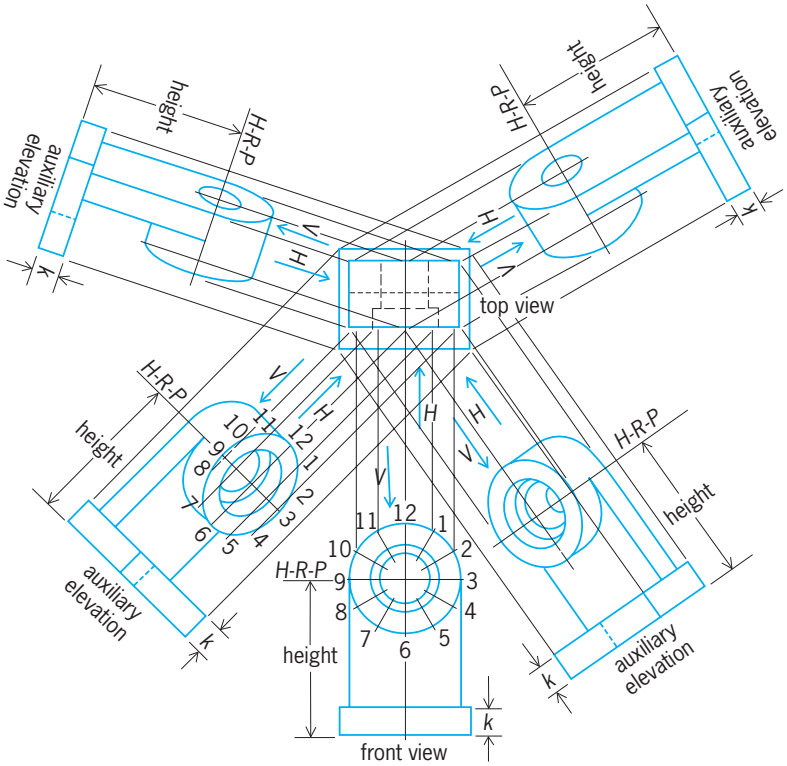

Fig. 6. Auxiliary views of a bearing. (*After G. J. Hood and A. S. Palmerlee, Geometry of Engineering Drawing, 4th ed., McGraw-Hill, 1958*)

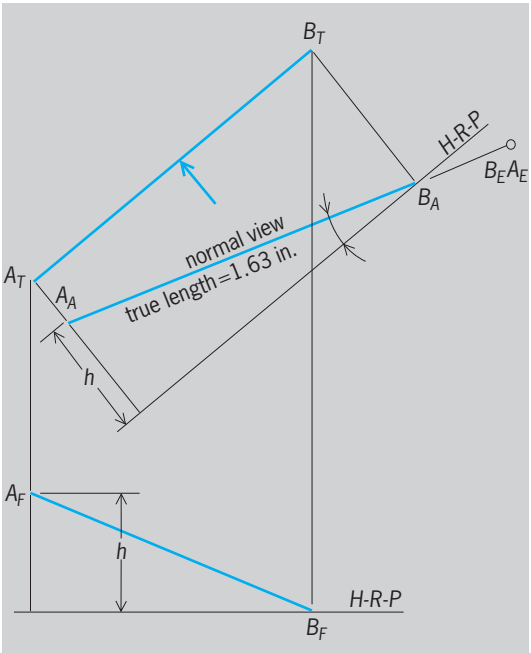

Fig. 7. True length of a line, seen in a normal view. (*After G. J. Hood and A. S. Palmerlee, Geometry of Engineering Drawing, 4th ed., McGraw-Hill, 1958*)

Geometrical elements of structures

- *Points*
- *Lines*
  - *Straight*
  - *Curved*
    - *Single-curved*: Circle, Ellipse, Parabola, Hyperbola, Trochoid, Spiral, Involute, Cycloid, Epicycloid, Hypocycloid, Sinusoid
    - *Double-curved*: General type, Helix
- *Surfaces*
  - *Ruled*
    - *Plane*: Triangle, Quadrilateral, Polygon, Prism, Wedge, Pyramid, Regular polyhedron
    - *Single-curved*: Cylinder, Cone, Convolute
    - *Warped*: General type, Hyperbolic paraboloid, Conoid, Helicoid, Hyperboloid of one sheet, Cylindroid
  - *Double-curved*
    - General type, Sphere, Ellipsoid, Paraboloid, Hyperboloid of two sheets, Torus, Surfaces of revolution, Serpentine

structures. Points are elements of lines, and lines are elements of surfaces. Lines are generated by a moving point; surfaces are generated by a moving line. The law that governs the motion of the moving point or line determines the nature of the generated line or surface. In the **table** of geometrical elements of structures, various kinds of lines and surfaces are named.

The first step in preparing a descriptive diagram is to decide what views are necessary to show the geometrical relations and locations of all elements and parts of the object. To draw the views, it is necessary to understand the measurements that must be made in each view; the principles involved in showing parallel, perpendicular, and angular relations between elements; the methods of showing the true lengths of lines and the true shapes of plane surfaces; the intersections of surfaces; and the development of surfaces.

True length of line. A normal view of a line or plane is taken in a direction perpendicular to the line or plane (**Fig. 7**). The true length of the line *AB* is seen in the normal view, here taken in a horizontal direction that is perpendicular to the line *B* in the top view. Dimension *b* is obtained from the front view. The inclination of line *AB* to the horizontal also is indicated.

Parallel lines. Parallel lines appear parallel in every view, as is illustrated by a parallelogram (**Fig. 8**).

Perpendicular lines. Perpendicular lines appear perpendicular only in a view that is normal to one or both of the lines (**Fig. 9**). In Fig. 9, *RC* is a given line, but only the top view of the line *RB* is given. The front view of *RB* is to be found. The numerals indicate the consecutive order in which the lines are drawn. A normal auxiliary elevation is drawn

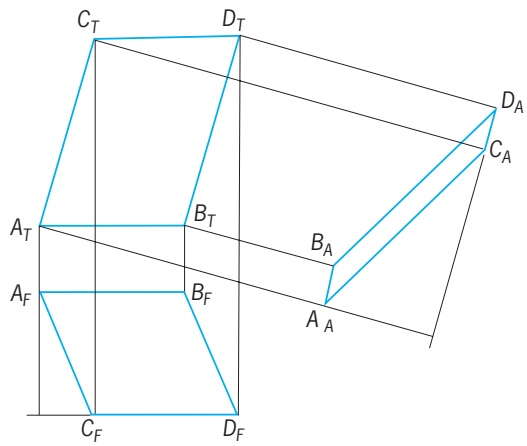

Fig. 8. Parallel lines in front, top, and auxiliary views. (*After G. J. Hood and A. S. Palmerlee, Geometry of Engineering Drawing, 4th ed., McGraw-Hill, 1958*)

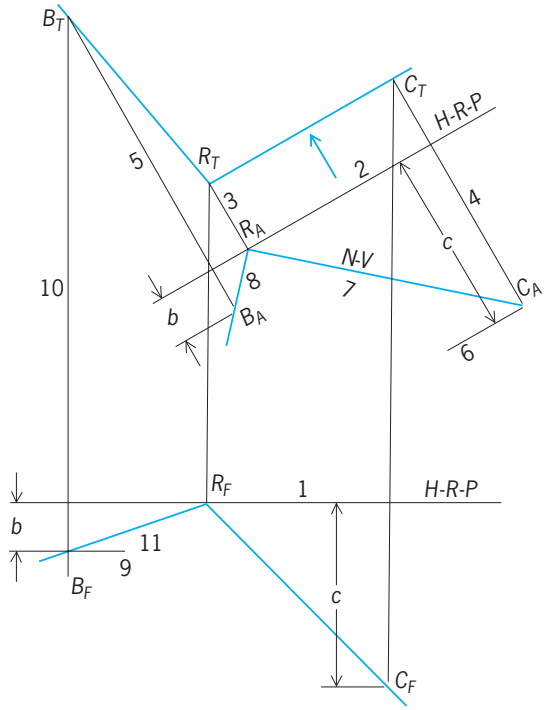

Fig. 9. Perpendicular oblique lines. (*After G. J. Hood and A. S. Palmerlee, Geometry of Engineering Drawing, 4th ed., McGraw-Hill, 1958*)

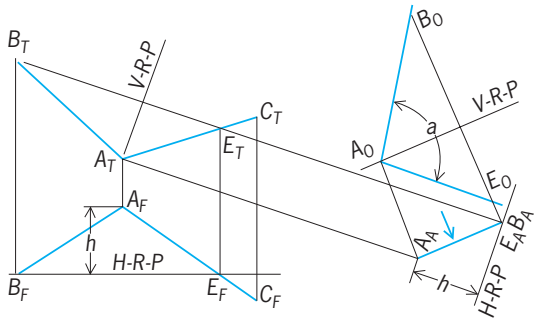

Fig. 10. Angle between two skew lines. (*After G. J. Hood and A. S. Palmerlee, Geometry of Engineering Drawing, 4th ed., McGraw-Hill, 1958*)

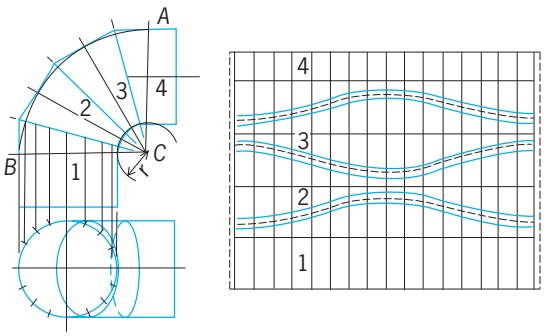

Fig. 11. Layout and top of front views of a four-piece elbow. (*After G. J. Hood and A. S. Palmerlee, Geometry of Engineering Drawing, 4th ed., McGraw-Hill, 1958*)

perpendicular to the line *RC* in the top view. In this normal view, line *RC* is drawn and *RB* is drawn perpendicular to *RC*. This view now shows a normal view of *RC*, but not of *RB*. The front view of *RB* may now be drawn.

Angle between lines. The angle between lines *AB* and *AC* is to be found (**Fig. 10**). An auxiliary elevation taken in the direction of horizontal line *BE* is drawn. This is an edge view of the plane containing given lines *AB* and *AC*. Next, a view taken in a direction perpendicular to this plane shows a normal view of the two given lines and the true size of the angle between them.

Intersection and development of surfaces. A structure is limited by the surfaces that bound it. Each surface of the structure is in turn limited by the straight or curved lines in which it intersects adjacent surfaces. These lines of intersection determine the edges and joints of the structure and are shown in drawings to describe the structure. In addition to the lines of intersection, the outlines, or contours, of curved surfaces also are shown in drawings. All intersections must be accurately located to make accurate developments.

For example, sheet metal bends or folds along a straight line; hence only plane-faced and single-curved surfaces may be developed. The single-curved surfaces are cylinders, cones, and convolutes. Warped surfaces and double-curved surfaces cannot be developed but must be formed by stretching the metal. Each type is useful in engineering designs.

A four-piece elbow is constructed from parts of right circular cylinders of uniform cross section (**Fig. 11**). At the right the development of each section is outlined by the solid lines. Allowance for seams is indicated by broken lines. No material is wasted.

The warped surfaces, such as the hyperbolic paraboloid, the concoid, the helicoid, hyperboloid of revolution, and others, have definite uses and are relatively easy to construct if the geometry of the surfaces is understood.

Topographical problems. Engineers encounter topographical problems dealing with land contours, underground and surface workings and surveys, highways, bridge piers, dams, foundations, retaining walls, excavations, underground and surface water supplies, drainage, irrigation, geological formations, rock strata, oil wells, mines, tunnels, veins of ore, dumps, layout of grounds, landscaping, and structures built into the ground and on the surface of the ground (**Fig. 12**).

The upper half of Fig. 12 illustrates the map of a section of mountainous country. The section has been prospected, and outcrops or traces of veins of ore have been located. Items such as the strike, dip, thickness, line of outcrop, and overburden of the veins can be determined and the location and length

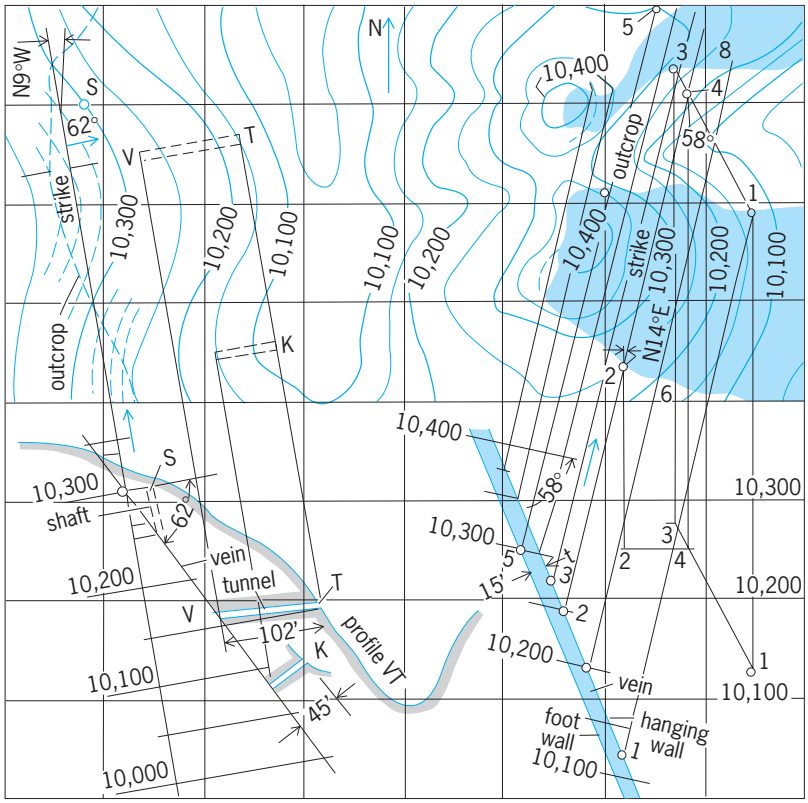

Fig. 12. Veins and outcrops of a mountainous section shown in top half. The bottom half provides additional data. (*After G. J. Hood and A. S. Palmerlee, Geometry of Engineering Drawing, 4th ed., McGraw-Hill, 1958*)

of tunnels and shafts established. *See* DRAFTING; ENGINEERING DRAWING.

Albert S. Palmerlee; Charles J. Baer

Bibliography. J. H. Earle, *Geometry of Engineers*, 1984; V. Guillemin and S. Sternberg, *Geometric Asymptotics*, 1991; E. G. Pare et al., *Descriptive Geometry*, 9th ed., 1996; S. A. Stewart, *Applied Descriptive Geometry*, 1986.

Desert

No precise definition of a desert exists. From an ecological viewpoint the scarcity of rainfall is all important, as it directly affects plant productivity which in turn affects the abundance, diversity, and activity of animals. It has become customary to describe deserts as extremely arid where the mean precipitation is less than 2.5–4 in. (60–100 mm), arid where it is 2.5–4 to 6–10 in. (60–100 to 150–250 mm), and semiarid where it is 6–10 to 10–20 in. (150–250 to 250–500 mm). However, mean figures tend to distort the true state of affairs because precipitation in deserts is unreliable and variable. In some areas, such as the Atacama in Chile and the Arabian Desert, there may be no rainfall for several years. It is the biological effectiveness of rainfall that matters and this may vary with wind and temperature, which affect evaporation rates. The vegetation cover also alters the evaporation rate and increases the effectiveness of rainfall. Rainfall, then, is the chief limiting factor to biological processes but intense solar radiation, high temperatures, and a paucity of nutrients (especially of nitrogen) may also limit plant productivity, and hence animal abundance. *See* PRECIPITATION (METEOROLOGY).

The main desert regions of the world are shown in the **illustration**. Most lie within the tropics and hence are hot as well as arid. The Namib and Atacama coastal deserts are kept cool by the Benguela and Humboldt ocean currents, and many desert areas of central Asia are cool because of high latitude and altitude.

Plant production. As a consequence of unreliable rainfall, plant productivity is much more variable than in most ecosystems. It may vary from none to 880 dry lb/acre · yr (1000 dry kg/ha · yr) although in most places there is usually some

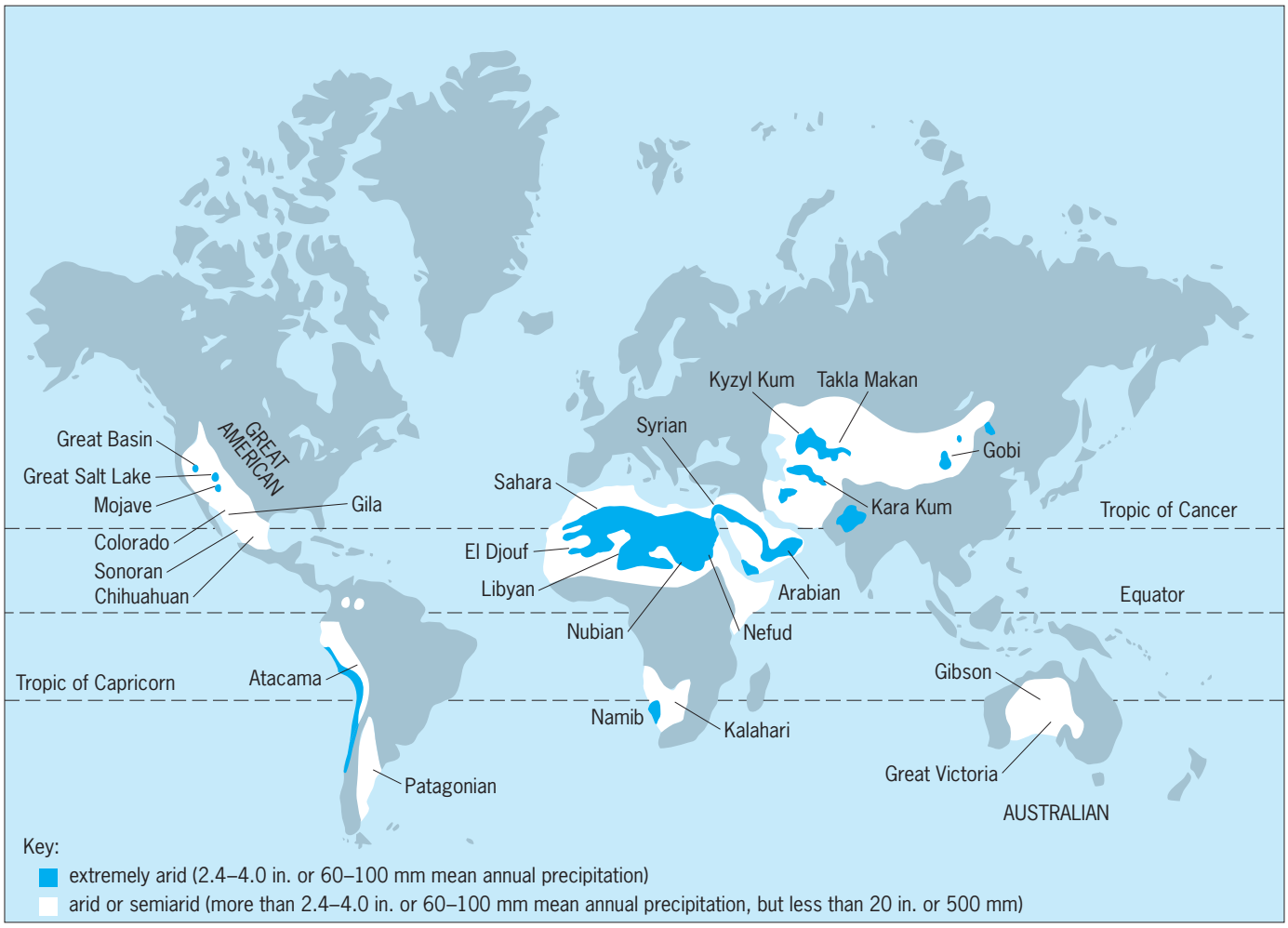

The deserts of the world; those within the tropics are hot as well as arid.

productivity even when there is no rainfall. The average biomass is generally low at 0.004–0.144 dry lb/ft^2 (0.02–0.7 dry kg/m^2) compared to 9.25 dry lb/ft^2 (45 dry kg/m^2) in tropical forest and 6.17 dry lb/ft^2 (30 dry kg/m^2) in temperate forest. Another feature of desert vegetation is the low percentage of green, photosynthetic plant biomass. There is also about three times as much dead as living plant material, a high figure compared to most other ecosystems. *See* BIOLOGICAL PRODUCTIVITY; BIOMASS.

Between 10 and 20% of the living plant material is eaten by consumers, a figure typical of terrestrial ecosystems, where the level of consumption may be as low as 5% (compared to over 90% in ocean ecosystems). Much desert vegetation tends to be hard and prickly and hence unpalatable to many consumers. Dead plant material does not accumulate and is utilized by the decomposers, particularly wood-eating termites. Seeds are abundant and can survive a long time without germinating. Many animals, including rodents, ants, and birds, depend heavily upon them for food. There is evidence of both intra- and interspecific competition for seeds, the production of which varies markedly with variations in rainfall.

A common viewpoint is that desert plants and animals live in a harsh environment and as a consequence have evolved many morphological and behavioral adaptations that enable them to "escape" the rigors of their surroundings. Certainly extreme aridity coupled with extreme heat seems a stressful environment, but whether desert organisms really possess more escape adaptations than, say, rainforest organisms is a matter for conjecture. Nevertheless, specific adaptations to drought and heat can be readily identified.

Adaptations in plants. In many deserts, annual rather than perennial plants form the bulk of the climax vegetation, quite unlike the situation in most terrestrial ecosystems where perennial plants dominate. Deserts, especially those within the tropics, periodically "bloom" with flowering annuals soon after a significant rainfall. In general, the more unreliable the rainfall in a desert, the greater the abundance and diversity of annual plants. Estimates vary, but these annuals account for up to half the primary production in a year, although for a given site they may fail altogether if there is no rain. Many annuals last for only a very short time. They tend to have the C_4 photosynthetic pathway, which means they grow quickly and opportunistically. It is these plants in particular that produce the vast quantities of seed which under extreme aridity can last for years, providing a "seed bank" until conditions are right for germination. Most ephemeral annuals do not have special drought-resistant or drought-tolerant adaptations: they are simply opportunists, much like the annuals of wetter communities which, however, tend to be successional rather than climax species. *See* PHOTOSYNTHESIS.

Succulent plants occur in all deserts but nowhere as conspicuously as in the Sonoran Desert of North America. These fleshy plants store water in their tissues, and tend to be especially prickly, a defense against browsers which need water as well as food. In American deserts the obvious succulents are Cactaceae; in Africa where cactus is almost absent, many species of Euphorbiaceae have the same types of growth form and occupy essentially the same ecological niches. *See* CACTUS; CARYOPHYLLALES.

Other perennial plants include trees and shrubs with long tap roots that can reach underground water, making them independent of scarce and unreliable rainfall. Some perennials burst into leaf only after rain, and are hence leafless most of the time. Yet others retain leaves throughout dry periods; the leaves tend to be small and narrow and require relatively little water to photosynthesize. These plants, mostly small shrubs and grasses, invariably have an extensive root system which enables them to maximize their acquisition of water. *See* PHYSIOLOGICAL ECOLOGY (PLANT); PLANT-WATER RELATIONS; PLANTS, LIFE FORMS OF.

Adaptations in animals. A majority of terrestrial invertebrates are cryptic in coloration, matching almost exactly components of the background on which they normally live. Many invertebrates and small vertebrates, such as rodents and snakes, spend the day in holes and burrows, and become active at night, when it is cooler. In all environments there are many small animals that escape by hiding, in most places to escape from predators; but in deserts escape from high daytime temperatures and radiation appears to be the chief adaptive response. The burrow constructed by the New World kangaroo rat, *Dipodomys*, protects the animal from daytime heat and radiation and also from predatory snakes and birds. The entrance of the burrow is sealed with soil which allows a high (30%) relative humidity to develop which helps the animal maintain its water budget. Kangaroo rats do not drink but exist on water derived from seeds they eat; their urine is the most concentrated of all mammals, another antidesiccation adaptation. Desert birds also shelter and avoid the heat of the day. In the Kalahari, the social weaverbird, *Philetarius socius*, builds a gigantic communal nest in an acacia or other suitable tree. The insulation in the nest protects the birds from the heat of the day and the cold of the night. *See* EXCRETION; PROTECTIVE COLORATION.

The largest of all birds, the ostrich, *Struthio camelus*, is an inhabitant of arid regions of Africa. The birds are too large to retreat into hiding during the day, although they may seek the shade of a big tree. With a high ambient temperature (95°F or 35°C or more) and no wind an ostrich loses heat by panting and by erecting the sparsely distributed feathers on its back. If there is wind, panting ceases and heat is lost by convection across the erect back feathers. At night when it is cool or even cold (below 64°F or 18°C), the back feathers are flattened and an insulating layer of air is trapped which reduces heat loss.

Soon after heavy rainfall, holes and depressions in sand or rock fill with water and within a few hours teem with microscopic life, chiefly algae, bacteria, and protozoa. After a few days countless small crustaceans appear, such as fairy and tadpole shrimps. These grow and reproduce with great rapidity. They can tolerate high water temperatures and also the high salinity which often builds up in temporary desert pools. The pools soon disappear and the land may then remain dry for months, even years, before the next downpour. The shrimps and other organisms diapause in the soil, usually as eggs which are remarkably resistant to high temperatures and extreme desiccation. *See* PHYSIOLOGICAL ECOLOGY (ANIMAL).

Aestivation. The term aestivation is used to describe a lengthy period of dormancy during which metabolism is much reduced. Many desert animals are capable of prolonged aestivation which enables them to survive during periods of food and water scarcity. In insects such as butterflies, which have a complete metamorphosis (clear-cut egg, larval, pupal, and adult stages), aestivation can occur at any stage of the life cycle but is almost invariably confined to a particular stage for a given species. Before entering aestivation, most animals seek out a secluded place that provides the best protection from temperature fluctuations and solar radiation, as well as from predators. Many species, ranging from lungfish to insect larvae, construct some form of protective cocoon in which to aestivate. *See* HIBERNATION AND ESTIVATION.

Migration and movement. An alternative to aestivation is migration or other movement away in search of better feeding and breeding conditions. Some of the African antelopes, such as the oryx, *Oryx gazella*, undertake long-distance movements which are correlated with changes in the quality and quantity of grass upon which they feed. These grass characteristics in turn are determined by rainfall, so there may be regular movements to and from areas where the rainfall is predictable and regular, or much more erratic movements in areas of unpredictable rainfall.

The semiregular migrations of several species of African weaverbirds (Ploceidae) are strongly associated with the seeding of wild grasses. The birds may form flocks of thousands or even tens of thousands of individuals and descend on and devour ripening seeds before moving on. Some of these birds, including the black-faced dioch, *Quelea quelea*, have become serious pests of millet and other cereals grown on land irrigated from underground water. *See* MIGRATORY BEHAVIOR.

The blooming of desert annuals and the leafing out of trees and shrubs soon after rainfall often result in a dramatic increase in abundance of leaf-eating insects. Eggs that have remained dormant for months produce larvae which can be so numerous as to defoliate vegetation. In Africa, white butterflies, *Belenois*, may migrate in vast numbers from areas where the larval food plants have been defoliated to areas where there is fresh growth. Pastoral people and their livestock undertake similar movements in response to changes in the availability of graze and browse for their animals, and to some extent, in response to changes in the risk of diseases such as trypanosomiasis, which affects both humans and domestic animals.

One of the most successful groups of desert animals is the grasshoppers. In many deserts there is a considerable variety of species, some of which are abundant. A few species are able to build up suddenly in numbers, often in response to unseasonal rain which has promoted a rapid growth of vegetation. They then undertake long-distance irruptive movements, followed by further breeding until they reach such numbers that they invade higher-rainfall areas and devastate crops. One of these grasshoppers, the desert locust, *Schistocerca gregaria*, can invade and seriously affect the vegetation of an area extending through 110° of longitude from West Africa to Assam, north to Turkey and south to southern Tanzania. Plagues of desert locusts may last for years and then suddenly stop, the locust reverting to a normal desert grasshopper.

Species diversity. The diversity of species of animals in a desert is generally correlated with the diversity of plant species, which to a considerable degree is correlated with the predictability and amount of rainfall. There is a rather weak latitudinal gradient of diversity with relatively more species nearer the Equator than at higher latitudes. This gradient is much more conspicuous in wetter ecosystems, such as forests, and in deserts appears to be overridden by the manifold effects of rainfall. Animals, too, may affect plant diversity: the burrowing activities of rodents create niches for plants which could not otherwise survive, and mound-building termites tend to concentrate decomposition and hence nutrients, which provide opportunities for plants to colonize.

Convergent evolution. Each desert has its own community of species, and these communities are repeated in different parts of the world. Very often the organisms that occupy similar niches in different deserts belong to unrelated taxa. The overall structural similarity between American cactus species and African euphorbias is an example of convergent evolution, in which separate and unrelated groups have evolved almost identical adaptations under similar environmental conditions in widely separated parts of the world. Convergent structural modification occurs in many organisms in all environments, but is especially noticeable in deserts where possibly the small number of ecological niches has necessitated greater specialization and restriction of way of life. The face and especially the large ears of desert foxes of the Sahara and of North America are remarkably similar, and there is an extraordinary resemblance between North American sidewinding rattlesnakes and Namib sidewinding adders. *See* EUPHORBIALES; ORGANIC EVOLUTION.

Desert community. Ecological change in deserts seems to occur slowly. Plant and animal succession is much less obvious than in other communities, although there is some evidence of cyclical change

in which species replace one another. In the Chihuahuan Desert of North America, the shrub *Larrea tridentata* reduces the wind speed enough to cause accumulation of soil and organic nutrients around the bottom of the trunk. This enables other plants to establish themselves, including the cactus *Opuntia leptocaulis*, whose seeds are introduced into the accumulated soil via the feces of rodents and birds which eat the fruit. The cactus grows and its roots overrun and outcompete those of the *Larrea* shrub, which then dies and falls over. Wind and water erosion then dislodge the rather shallow roots of the cactus, which itself dies and topples over. The vacant space is eventually filled by another *Larrea*, and the cycle starts again. No one has witnessed the complete cycle, which is believed to take many years to complete, but all stages can be identified and the presumed sequence of events drawn together. *See* ECOLOGICAL COMMUNITIES; ECOLOGICAL SUCCESSION; ECOLOGY; ECOSYSTEM. Denis F. Owen

Bibliography. G. N. Louw and M. K. Seely, *Ecology of Desert Organisms*, 1982; G. M. O. Maloiy (ed.), *Comparative Physiology of Desert Animals*, 1972; K. Schmidt-Nielsen, *Desert Animals: Physiological Problems of Heat and Water*, 1964; F. H. Wagner, *Wildlife of the Deserts*, 1980.

Desertification

Land degradation in low-rainfall and seasonally dry areas of the Earth. It can be viewed as both a process and the resulting condition. Desertification involves the impoverishment of vegetation and soil resources. Key characteristics include the degradation of natural vegetation cover and undesirable changes in the composition of forage species, deterioration in soil quality, decreasing water availability, and increased soil erosion from wind and water. Desertification is a global problem. Various stages of desertification can be seen in most of the world's drylands. In rare cases, desertification leads to abandoned, desertlike landscapes.

Causes and consequences. Although some authorities believe that climate change may be a causal factor, it is generally agreed that human activities, particularly excessive resource use and abusive land-use practices, are the primary cause of desertification. Specific activities leading to desertification include clearing and cultivation of low-rainfall areas where such cultivation is not sustainable, overgrazing of rangelands, clearing of woody plant species for fuelwood and building materials, and mismanagement of irrigated cropland leading to the buildup of mineral salts in the soil (salinization). Drought is often cited as a basic cause of desertification; however, it merely accelerates or accentuates land degradation processes already under way. *See* DROUGHT.

Consequences of desertification include reduced biological productivity, reduction of biodiversity, a gradual loss of agricultural potential and resource value, loss of food security, reduced carrying capacity for humans and livestock, increased risks from drought and flooding, and in extreme cases, barren lands that are effectively beyond restoration. Paleostudies, supported by model simulations, have shown that the intensity of Northern Hemisphere desert conditions has waxed and waned over the past 9000 years in response to the precession of the Earth's orbit about the Sun. Thus, it may be that the causal factors of desertification, whether climate change or human activities, depend on the time scale being addressed.

Concept. Desertification is an age-old problem in the world's drylands. For example, substantial areas in the region formerly known as Mesopotamia (the lower Tigris and Euphrates river basins of modern-day Iraq) have never recovered from desertification processes occurring over 4000 years ago. Likewise, much of the Mediterranean littoral has experienced varying degrees of desertification from excessive resource use beginning in the Roman era. Elsewhere in the world, on a lesser scale, desertification is known to have occurred centuries ago.

Beginning in the 1920s, there were warnings of the expansion or encroachment of the Sahara Desert into agrarian areas in the Sahel to the south. Massive soil erosion in the dust bowl in the United States during the 1930s pointed to the dangers of land degradation in drylands from unwise resource use. *See* SAHEL.

The term "desertification" was coined in 1949 by A. Aubreville, a French forester, who described the creation of desertlike conditions resulting from deforestation in the seasonally dry subhumid Sudanian belt south of the Sahel in Africa. It was not until the 1970s, however, that the term came into widespread international use. The reason was the drought in the African Sahel which began in 1968 and which by 1973 had resulted in at least 100,000 deaths and many times more refugees. It became obvious that the famine was caused by more than drought: an environmental calamity had been developing for many years due to unsustainable patterns of stockraising and cultivation that were incompatible with the region's environmental limitations. Drought served as a catalyst to reveal the damage caused by overgrazing and overcultivation. The resulting desertlike conditions made it seem that the Sahara was expanding. Similar land degradation was occurring in many other drylands, arousing fears of climate change and the specter of a global crisis.

As a result of these events, the United Nations (UN) convened the first international conference on desertification (UNCOD) in 1977. The conference aspired to provide a comprehensive understanding of desertification and its consequences for development, as well as to develop a plan of action to combat this global threat in its various settings. In 1978 the United Nations Environment Programme (UNEP) estimated that over 1.15×10^7 mi^2 (3×10^7 km^2) were severely desertified. In 1984 UNEP estimated that 35% of the world's land surface was at risk and that nearly 5.2×10^7 acres (2.1×10^7 hectares) of productive land were being reduced to near or complete uselessness each year. This agency

actively promoted the theory that the Sahara Desert was expanding southward at an average of almost 3.7 mi (6 km) per year. In response, a number of "green belt" schemes were implemented to try to halt the desert's expansion with rows of trees. In 1991 UNEP estimated that 73% of the world's rangeland, 47% of its rain-fed cropland, and 30% of its irrigated cropland was at least moderately desertified. This agency also asserted that the area irreversibly lost to desert was equivalent to all of the Earth's presently cultivated land and that desertification jeopardized the livelihoods of over a billion people in more than 100 countries. As a result of such claims, desertification became regarded as one of the world's most pressing environmental problems. During the 1992 UN "Earth Summit" Conference on Environment and Development in Rio de Janeiro, world leaders agreed to initiate negotiations on an international convention to address global desertification. The resulting Convention to Combat Desertification (CCD) was adopted in 1994. Following ratification by 50 countries, it became international law at the end of 1996. The convention's objective is to combat desertification and mitigate the effects of drought, particularly in Africa and other developing regions.

Controversies. Unfortunately, no universally accepted definition of desertification emerged from the 1977 UNCOD forum, and by the early 1980s there were more than a hundred different definitions of this phenomenon describing land degradation in ecosystems ranging from near-deserts to moist tropical rainforests. The lack of precision in defining the concept of desertification generated a series of controversies which gradually have called the validity of the term, and the concept itself, into question. It is argued that the term "desertification" should be avoided altogether in preference for the more general term "land degradation." The latter is preferred because it can be applied to all terrestrial ecosystems and avoids the misconceptions evoked by the term "desertification." There is also a basic disagreement over which ecosystems should be included in the purview of desertification research—drylands only, or all ecosystems in which degradation leads to drier conditions.

The struggle to define and determine the scope of desertification has become politicized. Primarily for political reasons—to focus international assistance efforts on poorer drought-prone developing countries, especially in the African Sahel—world leaders at the 1992 "Earth Summit" agreed to define desertification as "land degradation in arid, semi-arid, and dry subhumid areas resulting from various factors, including climatic variations and human activities." This definition was adopted by the CCD. However, the term "desertification" continues to be commonly applied in humid biomes as well—for example, for the degradation of tropical rainforest into savanna grassland.

There is also a fundamental controversy over the global extent of desertification. Most existing definitions of desertification are too general to be truly useful for estimating the extent or severity of this phenomenon. Global estimates are based on informed opinion rather than field research or remotely sensed data, and the extent of the problem may be grossly overstated. In the case of the Sahara Desert, the official wisdom during the late 1970s and 1980s was that the desert was expanding southward at a substantial rate of 3.4–5.6 mi (5.5–9 km) per year, but critics charged that this claim was based on a flawed 1975 study. Subsequent studies based on satellite image interpretation have conclusively demonstrated that there has been no permanent shift of the Sahara's southern boundary. These studies show that the Sahara boundary shifts from year to year in response to fluctuations in precipitation. According to these studies, rainfall fluctuations also adequately explain all variations in vegetative land cover from year to year. These studies seriously challenge the assertions by UNEP and other organizations that there has been a massive loss of land productivity in the Sahel from desertification. They also challenge estimates that tens of thousands of square miles of productive land are being lost to desertification each year.

There is a fundamental question about whether desertification—in the sense of extensive new desertlike environments being created through human activities and possibly climate change—is even occurring. By the mid-1990s, a revisionist critique of the conventional concept of desertification was firmly established. Key elements of this critique include the following: (1) Much of the data supporting the conventional concept of desertification was flawed, because it was collected during drought periods. (2) The concept relied too heavily on premature assumptions (for example, desert advance) and minimal field data. (3) It is exceedingly difficult to distinguish between the temporary effects of drought and the more permanent condition of desertification. This difficulty raises serious doubts about the validity of many traditional desertification studies and most estimates of the global extent of desertification. (4) Much of the literature on desertification ignores the fact that most of the world's drylands and desert areas are natural ecosystems. For example, the 35% of the Earth's land surface that is commonly claimed to be at risk of desertification includes the Sahara Desert and the world's other natural deserts—areas that are hardly at risk. (5) The traditional vision of desertification as the spread of desertlike conditions fails to recognize that land degradation in dryland areas often involves less obvious processes such as the gradual replacement of desirable forage plants by unpalatable species, with no net loss of primary productivity. (6) The imprecise and often-conflicting definitions of desertification have prevented a precise understanding of the actual problem and therefore have prevented effective treatment.

By the late-1990s, a major paradigm shift in drylands research was well established. Most specialists now acknowledge the difficulty of distinguishing between drought effects and human-induced degradation. Researchers increasingly focus on the regenerative capacity of dryland areas or on the complex localized interactions among climate variability, human

activities, and the land surface. Few specialists would deny that extensive degradation of vegetation and soil resources is occurring in many low-rainfall and seasonally dry environments. However, they increasingly concur that these are variations of land degradation processes also occurring in more humid environments, rather than a unique problem called desertification. *See* CLIMATE MODIFICATION; DESERT.

Will Swearingen

Bibliography. R. C. Balling and M. A. J. Williams, *Interactions of Desertification and Climate*, John Wiley, New York, 1996; M. Mortimore, *Roots in the African Dust: Sustaining the Sub-Saharan Drylands*, Cambridge University Press, New York, 1998; S. E. Nicholson, C. J. Tucker, and M. B. Ba, Desertification, drought, and surface vegetation: An example from the West African Sahel, *Bull. Amer. Meteorol. Soc.*, 79:815–829, 1998; D. S. G. Thomas and N. J. Middleton, *Desertification: Exploding the Myth*, John Wiley, New York, 1994.

Desiccant

A substance used to withdraw moisture from other materials. Although the removal of large quantities of water is done by evaporation, aided by moving air currents and by elevated temperature, the last traces of moisture are often held very tightly and do not evaporate readily. Furthermore, evaporation ceases when the moisture content of the material is reduced to that of the drying-air current. For final drying, a desiccant is used. This is a substance with a high affinity for water, that is, it is hygroscopic. It may retain water through capillarity or adsorption or by reacting chemically. A substance, such as calcium chloride, which retains enough water to become wet and dissolve is said to be deliquescent. The drying agent is placed directly into the gas or liquid to be dried; solid materials are placed in a desiccator, a closed vessel in which moisture escapes by vapor pressure and diffuses to the desiccant through the dry desiccator atmosphere. A desiccant loses potency as it takes on water; often it can be renewed by heating. Desiccants which form hydrates can be selected to maintain certain levels of low humidity in a closed vessel. *See* DELIQUESCENCE.

Allen L. Hanson

Important types. Silica gel possesses a high adsorptive power because of its extreme capillarity, the capillary pores occupying approximately 50% of its specific volume. The capillaries are probably spine-shaped, and the average pore diameter has been estimated to be 4×10^{7} cm, which is only about 10 times the diameter of one molecule of adsorbate. The drying efficiency of silica gel depends upon the concentration of water in the gas mixture, the temperature of the gel and gas, the properties of the condensed liquid, its wettability, and the state of the gel itself. Silica gel adsorbs water vapor preferentially in the presence of other vapors. It is readily capable of drying air to a dew point that is below −70°C (−94°F). *See* DEW POINT.

Activated alumina is prepared from aluminum trihydrate. It is a granular porous adsorbent with properties similar to those of silica gel. Alumina gel has many applications in addition to gas drying. It is used to adsorb gases and vapors from gaseous mixtures and to dry liquids.

Anhydrous calcium sulfate is prepared from a high grade of gypsum, $CaSO_4 \cdot 2H_2O$, which is dried, crushed, sized, and heated to 230–260°C (450–500°F) for 2 h. This leaves a granular porous form of anhydrous calcium sulfate, with sufficient mechanical strength to support its own weight.

Magnesium perchlorate is the equal of any desiccant from the standpoint of drying efficiency. The adsorption rate is rapid, and the first hydrate does not lose water until 135°C (275°F) is reached, thereby permitting its use for drying gases at higher temperatures than most commercial desiccants. It passes through three hydrate stages, namely, di-, tri-, and hexahydrate. The last, $Mg(ClO_4)_2 \cdot 6H_2O$, represents saturation after adsorption of 48.6% of the dry weight of water, a high capacity. Although drying efficiency tends to decrease after each hydrate formation, even the trihydrate is superior to solid sodium hydroxide (NaOH) and calcium chloride ($CaCl_2$).

Other types. Other solid desiccants that have been used or studied for gas drying are oxides, such as barium and calcium oxide, and activated carbon. Barium oxide maintains a high drying activity up to 540°C (1000°F). Barium oxide also appears to have marked possibilities for the drying of gases at high temperatures.

Calcium oxide (CaO) has long been used as a desiccant because of its low cost. Although it has high drying efficiency, its capacity is low because of the formation of carbonates on its surface from the carbon dioxide of the air.

Activated carbon is old historically and probably has been as intensively studied as any single adsorbent. Although capable of adsorbing large amounts of water vapor, activated carbon finds its major use in solvent recovery, in odor and taste removal, and as a catalyst and catalyst carrier. Instead of the preferential adsorption of water vapor, as in the case of alumina and silica gel, organic vapors tend to displace any water present on the carbon. *See* ADSORPTION; DRYING.

William R. Marshall, Jr.

Design standards

Specifications of materials, physical measurement processes, performance of products, and characteristics of services rendered. Design standards may be established by individual manufacturers, trade associations, and national or international standards organizations. They may also be de facto standards; that is, they may be informally but universally accepted by those using them. The general purpose is to realize operational and manufacturing economies, to increase the interchangeability of products, and to promote uniformity of definitions of product characteristics.

Individual firms. Individual firms often maintain extensive and detailed standards of parts that are available for use in their product designs. Usually the standards have the effect of restricting the variety of parts to certain sizes and materials. For many firms, therefore, standardization has also meant simplification, defined as reducing a standard to the minimum of different sizes or other specifications that will cover user requirements. For instance, light bulbs and industrial electric motors have thus been limited to relatively few sizes. The practice has proved highly effective for many firms, especially when they had to reduce the "birthrate" of new-line items that threatened to clog even the most sophisticated inventory control system. A further advantage is that, provided the standards can be made sufficiently versatile, it is often possible to buy generic parts; for example, a certain bolt or screw can be bought as a standard hardware product rather than being a special job based on where it is installed.

A further development of this design approach may lead to the modulization of the entire product line, by reducing it to certain major subassemblies that are common to as many products as possible. Special jobs then typically require only a few added features, and cost savings may be realized.

A possible disadvantage is that the extensive use of general-purpose parts may jeopardize the spare parts business, especially where outside manufacturers can skim off the market for the more commonly used and profitable spare parts once the original patents, if any, have expired, and then leave the more complex and slow-moving spares to the original manufacturer. However, it is precisely this aspect of standardization which is often welcomed by the users of the product. For example, transfer machines, such as are used in automobile and other large-scale manufacture of piece parts, consist of many work heads joined by a conveyor, thus automating the materials-handling part of the process. Under pressure from the major auto makers, the mounting dimensions of the work heads have had to be standardized so that the entire machine can be reassembled out of the modules of several manufacturers instead of starting anew at each model change. Major manufacturers of diesel locomotives have likewise modulized and standardized their products, with great cost reductions relative to the former, highly individualistic steam locomotives, an action which helped displace them after 1945.

Standardization also determines the nature of design practice. Especially when the specifications also give data on strength and performance as well as the usual dimensions, it is only necessary to compute loads approximately and then select the nearest standard sizes. Much design effort is thereby saved, especially on detail drawings, bills of material, and so forth. This approach also simplifies programming when computer-aided design is used. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Trade associations. Trade associations are the principal sources of American industrial standards. These involve standardization over an entire product line. In general, their scope is considerably less than that within firms with extensive standardization programs, but the technical and policy considerations in the two levels of standardization are quite similar. Trade standards are primarily concerned with specifying overall dimensions, so that products of different manufacturers may be used interchangeably; with performance, so that customers know what they are buying; and with certain design features, such as major materials in order to assure proper function. In some cases, dimensional standards particularly must be related to standards in other industries; for instance, an American butter dish must accommodate the standard 4-oz (113-g) sticks in which butter is packed. Like national standards to which they are closely related, trade association standards should be established on the basis of as broad a consensus as possible within the industry. If standards were established such that any required burden of retooling and product change would fall in a discriminatory fashion upon only certain members of the industry, legal remedy would certainly be sought under the American antitrust laws.

In large part for this reason, a major problem in the development of standards by trade associations is the resistance to simplification which often arises. A line of least resistance then is simply to take everybody's prestandardization product lines, put them together, and call the result a standard. Historically, this has resulted in some standards having so many different products as scarcely to merit the name. For instance, there are dozens of steel compositions bearing standard designations but virtually identical to several others; in spite of these redundancies, they remain as standards.

In structural steel, some shapes were eliminated soon after the major standards were set about 1905, but simplification such as that in the equivalent British standards, for instance, never took place. Thus, the "nominal sizes" of steel beams are turned out in several weights per foot. However, economic analysis of weight-bearing capacity indicates that the lightest version of each size is usually best, except in the relatively unusual cases where there is a dimensional restriction such as a need for increased clearance.

National standards organizations. The principal industrial countries have official agencies that approve, consolidate, and in some cases establish standards. Among them are the British Standards Institution, German Institute for Norms, and the American National Standards Institute, which issue the BSI, DIN, and ANSI standards, respectively. ANSI is a federation of 900 companies, large and small, and about 200 trade, technical, professional, labor, and consumer organizations. It does not develop standards, but coordinates and promotes the voluntary development of national standards by industries provided that these have been established according to detailed rules for achieving a consensus among the producing industries, consumers (through ANSI's Consumer Council), relevant government agencies, and other interested parties.

There are about 10,000 approved ANSI standards, dealing with dimensions, ratings, terminology and symbols, test methods, and performance and safety specifications for equipment, components, and products. Major applications are in construction, electrical, electronic, mechanical, and nuclear products and processes, piping and welding, heating, air conditioning and refrigeration, information systems, medical devices, physical distribution, photography and motion pictures, and textiles. Most activities in the private and public sector use standardized products at some stage and specify them routinely in their purchases. ANSI standards have been embodied in building codes and many other government regulations, notably those of the Occupational Health and Safety Administration (OSHA). Related activities by government agencies also establish standards, as in agricultural grading.

A problem in setting standards in general, but particularly national ones, is the frequent need to anticipate future technologies and needs at an early stage of product development, knowing that later any substantial changes would be very difficult. For instance, domestic electrical supply is standardized at 110 V ac, 60 Hz, in the United States, and 210–250 V ac, 50 Hz, in Europe. These standards were set before 1900 and were, in part, the result of perceived limits on the mechanical capabilities of electrical machinery. Today, it would be better to have 400-Hz current such as used in aircraft systems because it would allow the use of smaller, faster, more efficient, and less expensive motors, as well as better functioning fluorescent lights. But any change would require costly conversions of transmission systems and consuming equipment.

Another example is that of railroad gages, which were set at 4 ft $8^1/_2$ in. (1435 mm) in most of Europe and North America and narrower in much of the rest of the world. Yet a wider gage of 7 ft (2 m) would allow for more efficient use and higher speeds, albeit at higher construction costs. Such changes have never been made, but others in established gages have often proved expensive, requiring new construction and inconvenient changes for passengers and transshipments of freight.

International standards. The national standardization agencies are members of a wide variety of international groupings and United Nations agencies. The principal ones are the International Organization of Standardization (ISO) and the International Electrotechnical Commission (IEC). These attempt to coordinate national activities and promote cooperation in the area of standardization. Several of the more than 50 organizations deal with weights and measures, which have always been the subject of extensive international coordination and agreement. Others engage in transnational or international activities in the standardization of many products or cover specific regional issues and requirements. Some, like the European Economic Community (EEC), the International Telecommunications Union (ITU), the International Civil Aviation Organization (ICAO), or the administration of the General Agreement on Tariffs and Trade (GATT) mainly have other political, economic, and scientific concerns but must necessarily take note of standardization as part of their work. *See* ENGINEERING; ENGINEERING AND ARCHITECTURAL CONTRACTS; ENGINEERING DESIGN; PHYSICAL MEASUREMENT; SCIENTIFIC METHODS; STANDARD. John E. Ullman

Desmodorida

An order of nematodes in which the variable stoma may or may not be armed with a dorsal tooth which may or may not be opposed by subventral denticles. The variable amphids range from reniform-to-elongate loops to simple or multiple spirals. The cephalic sensilla are generally in three whorls; however, the second and third whorls may be combined. Distinguishing the order is the cephalic capsule or helmet and the conspicuous somatic annuli. In some groups anterior and posterior adhesion tubes are utilized in locomotion.

There are five desmodorid superfamilies: Ceramonematoidea, Desmodoroidea, Draconematoidea, Epsilonematoidea, and Monoposthioidea. Ceramonematoidea comprises marine forms with distinctive cuticular ornamentation that takes the form of crested annuli. Nothing is known of the feeding habits. Most species of Desmodoroidea are found in the marine environment. Draconematoidea is an unusual superfamily of marine forms; when relaxed, the body is dorsally and then ventrally arched into a shallow sigmoid shape. Epsilonematoidea and Monoposthioidea both comprise marine forms. *See* NEMATA. Armand R. Maggenti

Desmoscolecida

An order of nematodes in which the taxa are readily recognized by their conspicuous body annulation. The annuli may be covered with concretion rings, or the cuticle may be ornamented with scales, warts, or bristles. The cephalic sensilla reportedly are reduced in number. The internal whorl is absent, the second whorl is papilliform, and the four sensilla of the third circlet are setiform. The vesiculate amphids are oval to circular and occupy much of the cephalic region. The somatic setae are tubular, and the open distal ends are often elaborate. The posterior esophagus is only slightly expanded. When pigment spots or ocelli are present, they are just posterior to the esophagus. Females are amphidelphic, and the ovaries are generally outstretched. The male spicules are generally accompanied by a gubernaculum. Both sexes have three caudal glands.

There are two desmoscolecid superfamilies, Desmoscolecoidea and Greeffielloidea. Most species of Desmoscolecoidea are found in the marine environment, although some species have also been collected from brackish waters. Greeffielloidea are free-living nematodes found primarily in marine or, rarely, estuarine habitats. *See* NEMATA. Armand R. Maggenti

Desmostylia

An extinct order of large, hippopotamuslike, amphibious, gravigrade, shellfish-eating mammals, nearly 12 ft (4 m) long, that frequented shallow bays and coastal margins during the Oligocene and Miocene. Their main distribution was northern trans-Pacific, but discoveries in Florida indicate a much broader distribution.

There are two families, the Desmotylidae (*Desmostylus*, *Cornwallius*, and *Vanderhoofia*) and the Paleoparadoxidae (*Paleoparadoxia*). The many cusped molars are composed of a cluster of heavy enameled cylindrical columns (see **illus.**), resembling the pavement teeth in drumfish. The cheek teeth are antero-vertically replaced as in the Proboscidea and Sirenia. There are one to four pairs of procumbent tusks. These animals probably descended from an ancestral stock that also gave rise to the Proboscidea and Sirenia. *See* MAMMALIA; PROBOSCIDEA; SIRENIA.

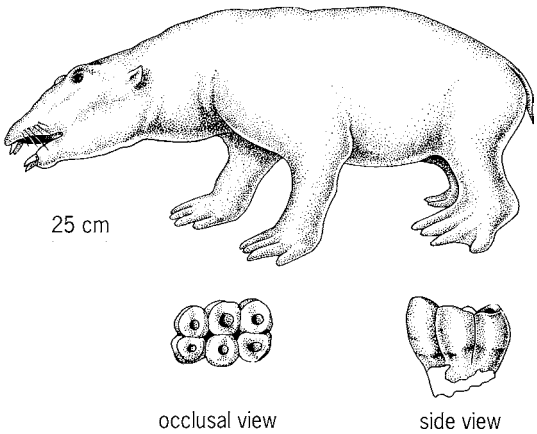

A cheek tooth and restoration of *Desmostylus* from the Miocene of California.

Gideon T. James

Bibliography. M. J. Benton, *Vertebrate Paleontology,* 1991; R. L. Carroll, *Vertebrate Paleontology and Evolution,* 1988.

Desmothoracida

An order of Heliozoia. These ameboid organisms live in a stalked perforated organic or inorganic lorica. Stiffened pseudopodia extend from the perforations to capture food which collides with them. The best-known species is *Clathrulina elegans,* found in fresh-water environments (see **illus.**). This species has a spherical organic test. Unlike most Heliozoia it has a life cycle which involves motile flagellated cells which arise from the sessile form and then may settle, transforming into amebas with thin pseudopodia. These ameboid forms secrete the stalk and lorica to give rise to the typical sessile (trophic) organism. Under unfavorable conditions the organism encysts within the lorica. *See* HELIOZOIA; PROTOZOA; SARCODINA.

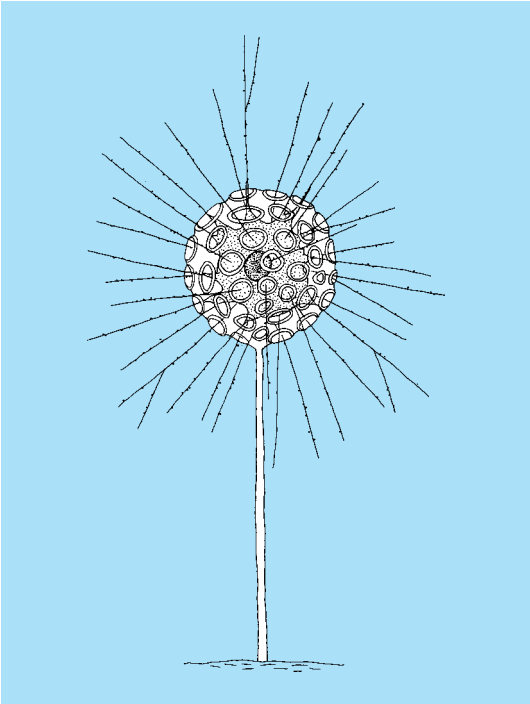

Sessile form of *Clathrulina elegans*. (*After C. F. Bardele, Cell cycle, morphogenesis and ultrastructures in the pseudoheliozoian Clathrulina elegans, Z. Zellforsch., 150:219–242, 1972*)

David J. Patterson

Bibliography. C. F. Bardele, Cell cycle, morphogenesis and ultrastructure in the pseudoheliozoian *Clathrulina elegans, Z. Zellforsch.,* 130:219–242, 1972; J. J. Lee, S. H. Hutner, and E. C. Bovee, *Illustrated Guide to the Protozoa,* Society of Protozoologists, 1984.

Desorption

A process in which atomic and molecular species residing on the surface of a solid leave the surface and enter the surrounding gas or vacuum. In stimulated desorption studies, species residing on a surface are made to desorb by incident electrons or photons. Measurements of these species provide insight into the ways that radiation affects matter, and are useful analytical probes of surface physics and chemistry. In thermal desorption studies, adsorbed surface species are caused to desorb as the sample is heated under controlled conditions. These measurements can provide information on surface-bond energies, the species present on the surface and their coverage, the order of the desorption process, and the number of bonding states or sites.

Stimulated desorption. Stimulated desorption from surfaces is initiated by electronic excitation of the surface bond by incident electrons or photons. The classical model of desorption is an adaptation of the theory of gas-phase dissociation, in which desorption results from excitation from a bonding state to an antibonding state. In this model the excitations

for gas-phase dissociation and desorption from surfaces are identical, but the surface, which can absorb energy from the excited adsorbate before it can desorb, dramatically reduces the overall dissociation yield from surfaces. Thus, processes which are strong in the gas phase can be quenched on the surface, leading to a marked difference in the relative importance of competing mechanisms in the two environments. This model applies predominantly to neutral desorption of adsorbed molecular species.

Another model, proposed by M. L. Knotek and P. J. Feibelman, which is more applicable to the phenomenon of ion desorption, was first observed in studies of the desorption of positively ionized oxygen (O^+) from the surface of titanium(IV) oxide (TiO_2). Here it is found that O^+ is desorbed not by valence level excitation, but by ionization of the titanium and oxygen core levels. These levels, of course, have little to do with bonding. Furthermore, the fact that the oxygen is desorbed as an O^+ ion (whereas it is nominally at O^{2-} on the surface) implies a large (three-electron) charge-transfer preceding desorption.

Titanium(IV) oxide is an ionic solid, and more importantly it is a maximal valency ionic compound. Maximal valency means that in the solid the titanium (Ti) is ionized to the noble-gas configuration Ti^{4+}, so that there are effectively no valence electrons on the Ti ion in the solid; its highest occupied orbital is $3p$ at about 30 eV below the Fermi level. The important thing about maximal valency ionic materials is that if a core level is ionized in either the cation or the anion, higher-lying valence electrons will decay into the deep hole left in the core level and give up the energy gained by emitting multiple-valence electrons from the solid (multiple Auger emission). Since valence electrons are only on the anion, the anion is selectively stripped in the process. In TiO_2 one electron is lost from O^{2-} by decay into the core hole; in about 10% of the events two Auger electrons are emitted, transforming the O^{2-} to an O^+ ion. An O^+ ion in a highly repulsive potential for positive ions results, which leads to O^+ desorption. In the general case of an ionic bond M^+A^- (**Fig. 1**), the incident radiation ionizes the core level of the M^+A^- pair. Auger decay of the core hole selectively removes charge from the A^- ion and results in its being transformed into an A^+, and the resulting repulsion causes the A^+ to desorb.

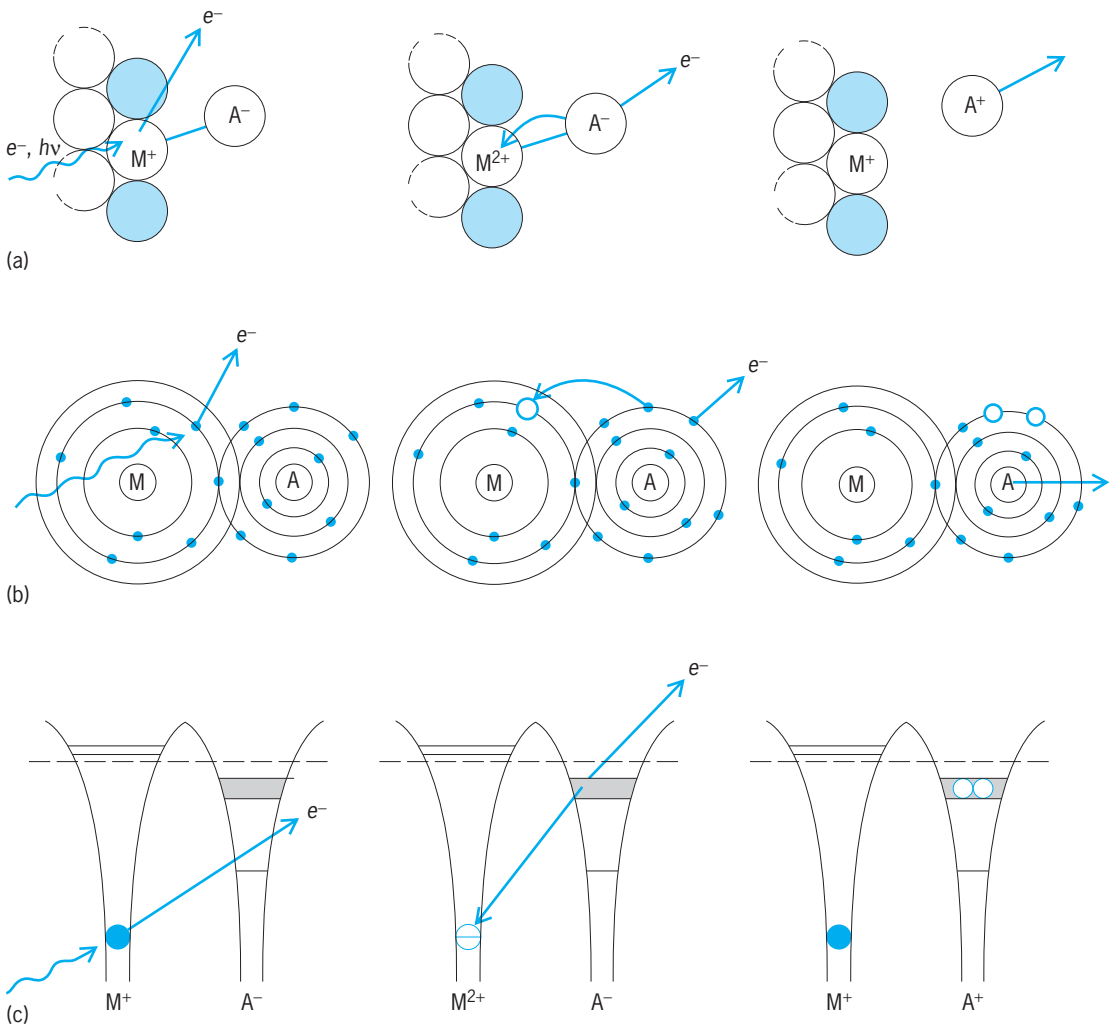

Fig. 1. Sequence of events leading to stimulated desorption by an electron (e^-) or photon ($h\nu$): (*a*) in the structural model, (*b*) in terms of orbitals, and (*c*) in terms of energy levels.

This model applies to a wide range of ionic materials, and explains why maximal valency materials like TiO_2, vanadium(V) oxide (V_2O_5), tungsten(VI) oxide (WO_3), molybdenum(VI) oxide (MoO_3), sodium fluoride (NaF), lithium fluoride (LiF), sodium chloride (NaCl), and so forth, are decomposed by irradiation. *See* AUGER EFFECT.

This mechanism for desorption can also be effective for covalently bonded surface species. Covalently bonded structures differ from ionically bonded structures in that the electronic charge which bonds the atoms together is distributed more evenly through the structure. The electronic interaction between atoms is strong so that an electronic excitation produced on one atom can quickly move to a neighbor. Thus, a simple electronic excitation can quickly move off the site where it was created before dissociation can occur. It typically takes 10^{-12} s for bond dissociation to occur, whereas the motion of simple electronic excitations can occur in 10^{-16} s. However, an important attribute of states where two or more electrons are removed from the atom is that such multiply excited states are much more difficult to move to a neighbor and hence are localized for times long enough for desorption to occur. Several features of a covalent bond can enhance this tendency to localize on the atom—for example, reducing the number of neighbors (and hence bonds), or reducing the bond's ability to transport charge to the desorbing species to neutralize the excited state.

Applications. Stimulated desorption studies are finding wide use. First, they can show the ways in which radiation affects the structure of solids. This will have important applications in the areas of radiation-induced damage and chemistry. Second, as an analytical tool, they offer a unique new way to study the physics and chemistry of atoms on surfaces which, when combined with the many other surface techniques based largely on electron spectroscopy, can provide new insight. Finally, models of the surface bond are put to a much sterner test in attempting to explain desorption phenomena. *See* RADIATION DAMAGE TO MATERIALS.

Analysis of surface structure. The techniques of electron- and photon-stimulated desorption can be very useful in the study of how atoms and molecules bond to surfaces. Since ions are desorbed only from the topmost layers of the solid, these techniques study only that outermost layer. By mass-analyzing the desorbed atomic and molecular species, the nature of the chemical species on the surface can be deduced. More importantly, an adsorbed species will be desorbed when the core level of its bonding-site atom is ionized. Since the energy of the electron or photon necessary to ionize a given atom's core level is well known and characteristic of that atom, the bonding sites of each of the desorbing species can be determined. Thus surface-, site-, and adsorbate-specific information can be obtained. Elaborate methods for deducing how the bonding site and adsorbate atoms are electronically and structurally configured can be utilized with this technique. The method is almost the only spectroscopy capable of detecting hydrogen, a classic problem in this field.

Ion angular distributions. An additional important discovery is that ion angular distributions from stimulated desorption are not isotropic, but show that ions are emitted in relatively narrow cones which project along the nominal ground-state bond directions. An angular display shows an azimuthal pattern which reflects the symmetry of the bonding site, and a polar angular distribution which reflects the bonding angle relative to the surface normal. Thus this technique provides a direct display of the surface-bonding geometry.

There are a number of ways to measure the angular distributions of desorbed ions. One of the simplest employs an image intensifier. Individual ion events result in light pulses at a phosphor screen, and time-averaged photographs of the screen provide visual images of the distributions and thus allow a direct view of the arrangement of atoms and molecules on the surface (**Fig. 2**). The directness and simplicity of the technique make it particularly attractive to measure complicated surface adsorbate structures or chemical configurations which do not contain the long-range order necessary in many other techniques for measuring structure. *See* LIGHT AMPLIFIER.

Thermal desorption. Thermal desorption mass spectroscopy is possibly the oldest technique for the study of adsorbates on surfaces. Three primary forms of the thermal desorption experiment involve measurement of (1) the rate of desorption from a surface during controlled heating (temperature-

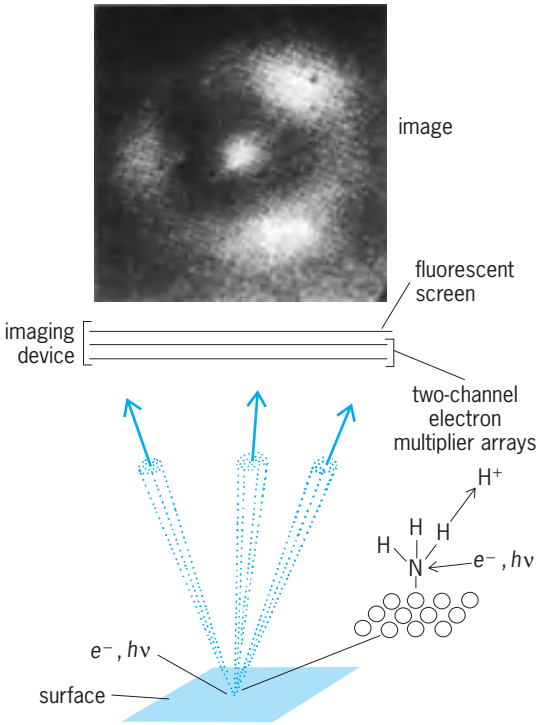

Fig. 2. Measurement of the angular distribution of ions leaving a surface in stimulated desorption.

programmed thermal desorption), (2) the rate of desorption at constant temperature (isothermal desorption), and (3) surface lifetimes and diffusion under exposure to a pulsed beam of adsorbates (molecular-beam experiments). Of the three, temperature-programmed thermal desorption is by far the most widely applied, and the discussion below is confined to it. *See* MOLECULAR BEAMS.

In thermal desorption measurements, a surface is first dosed with an adsorbate gas, and then the sample is heated at a controlled rate β while the rate of desorption of the adsorbed gas is measured with a mass-filter gas analyzer. The most straightforward information provided is the nature of the desorbed species from mass analysis, and a determination of the absolute coverage by the adsorbate, which is very difficult to obtain with other techniques. Coverage is determined by measuring the total amount of gas desorbed. The technique can also provide several important kinetic parameters of the desorption process as follows.

Since adsorption reactions are usually nonactivated, the desorption activation energy E, which thermal desorption can provide, is a good approximation to the differential heat of adsorption. Thus these studies provide information on surface-bond energies. *See* ADSORPTION.

Desorption kinetics can be quite complex, especially when there are strong interactions between adsorbates. The general form of the mth-order desorption rate equation is given by Eq. (1), where $R(t)$

$$R(t) = \nu_0^{(m)} n^m \exp(-E/kT_s) \qquad (1)$$

is the rate of desorption as a function of time, T_s is the substrate temperature, $\nu_0^{(m)}$ is the preexponential factor of the desorption-rate coefficient for order m, and n is the two-dimensional adsorbate concentration. Desorption has been observed to obey zero-, first-, and second-order kinetics. Zero-order kinetics is observed for multilayers where removal of an adsorbed species does not lower the effective coverage feeding the desorption process. First-order kinetics is observed for simple desorption of an atomic or molecular species from a submonolayer of adsorbate. Second-order kinetics is often observed in systems where dissociative adsorption has occurred and recombination reactions occur before the desorption.

Desorption spectra can determine the number of binding states of an adsorbate on a surface, from the number of desorption peaks observed, and, coupled with structural models of the surface, can help determine both the nature of the bonding sites and the bonding geometries.

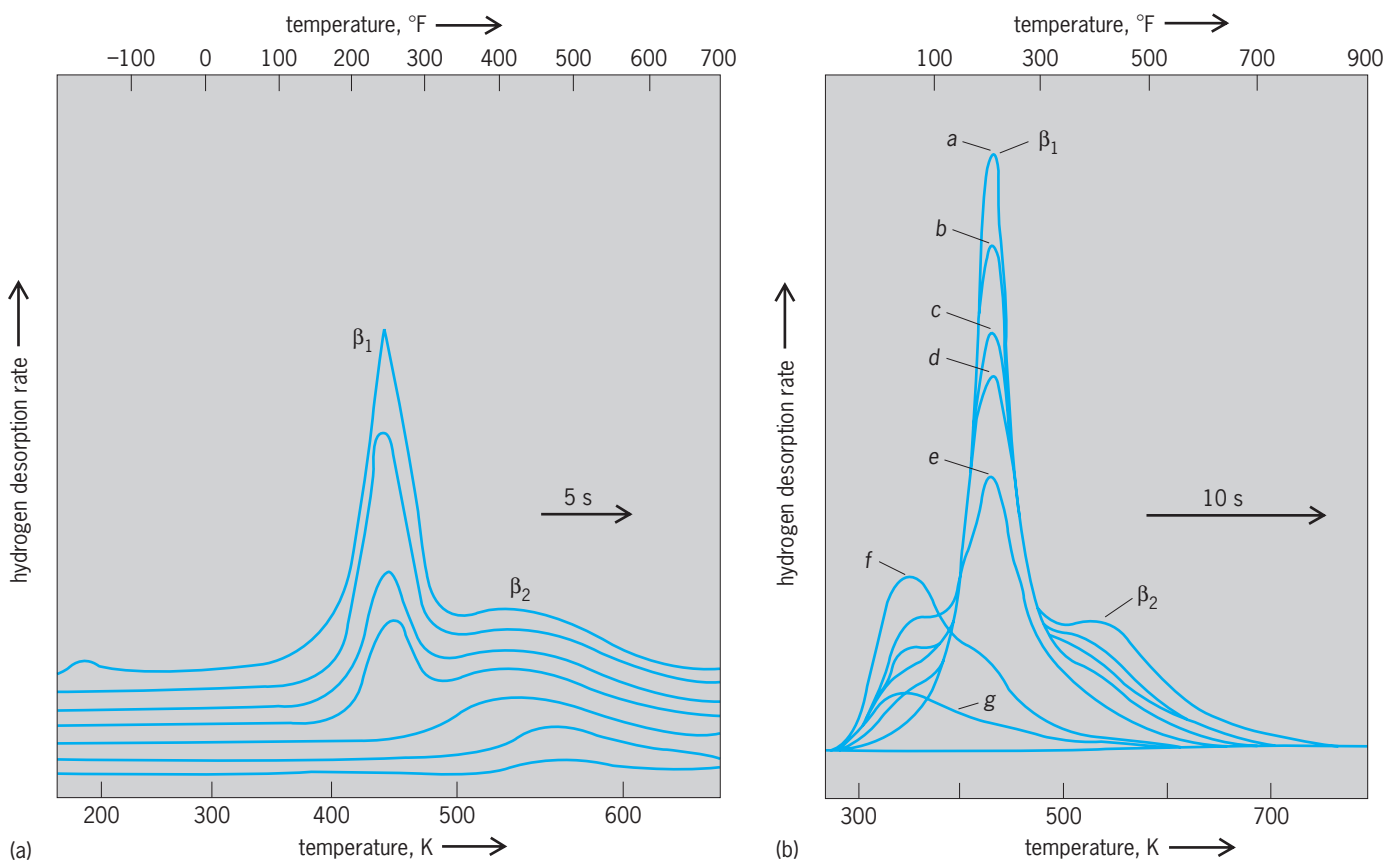

Fig. 3. Hydrogen desorption rate as a function of sample temperature for a heating rate of approximately 10°C/s. (*a*) Successive curves from bottom to top represent increasing hydrogen coverage. (*b*) Desorption when carbon monoxide (CO) is added to the hydrogen-preadsorbed surface. Successive curves from *a* to *g* represent increasing carbon monoxide coverage. (*After T. E. Madey and J. T. Yates, Desorption method A: Probes of kinetics and bonding at surfaces, Surf. Sci., 63:203–231, 1977*)

The preexponential factor of the desorption-rate coefficient, equivalent to an attempt frequency, the rate at which the desorbing particle attempts to leave the surface, not only is necessary to understand desorption kinetics, but helps in understanding the general problem of surface chemical dynamics.

The existence of indirect and direct interactions between coadsorbed species can be determined from thermal desorption. Knowledge of such interactions is directly applicable to such important problems as catalysis and corrosion. *See* CATALYSIS; CORROSION.

There are several methods to derive the kinetic parameters of Eq. (1) experimentally. If a value is assumed for the preexponential factor ν_0 (usually of the order of 10^{13}), the desorption activation energy is given by Eq. (2) for first-order kinetics, and by Eq. (3)

$$E/kT_p = (\nu_0^{(1)}/\beta)\exp(-E/kT_p) \qquad (2)$$

$$E_kT_p^2 = (\nu_0^{(2)}\, n_0/\beta)\exp(-E/kT_p) \qquad (3)$$

for second-order kinetics. Here β is the heating rate, n_0 is the initial adsorbate concentration, and T_p is the temperature of peak desorption rate. Alternatively, β can be varied and E can be derived from Eq. (4) for

$$E/kT_p = d(\ln\beta)/d(\ln T_p) \qquad (4)$$

first order; for second order, E/k is determined from the slope of a plot of $\ln(n_0T_p^2)$ versus $1/T_p$. It is also possible to derive the desorption energy from the width of the desorption peak.

One of the simplest of surface systems, hydrogen adsorbed on the (100) surface of tungsten (W), provides graphic examples of how the above types of information manifest themselves and are analyzed. **Figure 3** shows molecular desorption of hydrogen as a function of sample temperature for heating rates β as shown in the figure. At low hydrogen coverages a single β_2 state is desorbed with second-order kinetics. At coverages above 0.2 monolayer of hydrogen, the β_1 state desorbs with first-order kinetics. When carbon monoxide (CO) is added to the preadsorbed hydrogen layer, it causes the depopulation of both the β_1 and β_2 states and introduces a weak state of lower binding energy. The β_2 state is now known to be hydrogen bound at reconstructed sites on the W(100) surface, while the β_1 state is hydrogen at unreconstructed sites. *See* CRYSTAL STRUCTURE.

While the thermal desorption techniques are among the simplest of surface probes, they remain indispensable because of their directness and the variety of information they convey. Thus while surface science moves to detailed methods involving extremely sophisticated apparatus, the simple thermal desorption methods remain an important part of the overall picture. *See* SURFACE PHYSICS. M. L. Knotek

Bibliography. A. B. Burns et al. (eds.), *Proceedings of the 5th International Workshop: Desorption-Induced Electronic Transitions,* DIET V, 1993; N. Nogar (ed.), *Laser Photoionization and Desorption Surface Analysis Techniques*, 1990.

Destructive distillation

The primary chemical processing of materials such as wood, coal, oil shale, and some residual oils from refining of petroleum. It consists in heating material in an inert atmosphere at a temperature high enough for chemical decomposition. The principal products are (1) gases containing carbon monoxide, hydrogen, hydrogen sulfide, and ammonia, (2) oils, and (3) water solutions of organic acids, alcohols, and ammonium salts. For a discussion of the products of the destructive distillation, or coking, of coal *see* COAL CHEMICALS

Crude shale oil may be subjected to a destructive, or coking, distillation to reduce its viscosity and increase its hydrogen content. Subsequent catalytic hydrogenation (cobalt molybdate alumina catalyst, about 750°F or 400°C, 100–1500 lb/in.^{2} or 700–10,000 kilopascals pressure) lowers the nitrogen and sulfur contents so that the oil can then be refined by normal petroleum refinery operations. Residual oils from petroleum refinery operations are subjected to coking-distillation to reduce the carbon content. The coke is used for the manufacture of electrode carbon and the oil is returned to the feed for normal petroleum refining.

Prior to about 1920, destructive distillation of wood was an important source of methanol, acetic acid, and acetone. These chemicals are now produced at a lower cost by other methods; for example, methanol is produced by hydrogenation of carbon monoxide, acetaldehyde and acetic acid from acetylene, and acetone by fermentation processes and by decomposition of cumene hydroperoxide. The main product of the destructive distillation of wood is 40–45% charcoal used in metallurgical processes in which the low content of ash, sulfur, and phosphorus is important. Special chars made by the destructive distillation of the shells of nuts have very large surface areas and are used as adsorbents in gas masks and in some chemical processes. Char from the destructive distillation of bones is used for decolorizing solutions of raw sugar, and the oil obtained during the decomposition of bones contains recoverable amounts of pyridine, pyrrole, quinolines, and other nitrogen compounds. *See* CHARCOAL; COKE; COKING (PETROLEUM); OIL SHALE; PYROLYSIS; WOOD CHEMICALS. H. H. Storch; H. W. Wainwright

Detergent

A substance used to enhance the cleansing action of water. A detergent is an emulsifier, which penetrates and breaks up the oily film that binds dirt particles, and a wetting agent, which helps them to float off. Emulsifier molecules have an oillike nonpolar

portion which is drawn into the oil, and a polar group that is water-soluble; by bridging the oil-water interface, they break the oil into dispersible droplets (emulsion). As a surfactant, a detergent decreases the surface tension of water to help it penetrate soil. *See* EMULSION.

Soap, the sodium salt of long-chain acids, was the principal detergent until superseded in 1954 by synthetic detergents (syndets) which, unlike soap, do not form insoluble products with the calcium in hard water. Most syndets are of the anionic type, that is, sodium salts of alkyl sulfates or sulfonates which contain a chain of 7–18 carbon atoms; an example is sodium lauryl sulfate, $CH_3(CH_2)_{11}OSO_3^-,Na^+$. Alkyl benzene sulfonates (ABS) with branched carbon chains were found to persist in wastewater and have been replaced by linear alkyl benzene sulfonates (LAS), which are biodegradable by bacterial action. Anionic detergents are best for water-absorbing fibers such as cotton, wool, and silk. Nonionic detergents are polyethers made by combining ethylene oxide with a 12-carbon lauryl alcohol obtained from tall oil, a waste product from the paper industry. They are used for water-repelling "permanent press" fabrics, and their low-foaming property is desirable for automatic washers. Cationic syndets are quaternary base compounds of type $(NR_4)^+,X^-$. They are more expensive, but some are germicidal; some are used as fabric softeners and as good metal cleaners. *See* SOAP.

Detergents must contain alkaline "builders" to bind dissolved metal ions and support emulsification. Sodium pyrophosphate or polyphosphate is preferred because of low cost and high cleansing effectiveness. However, when discharged with laundry wastewater, these compounds supply nutrient to phosphate-deficient lakes and streams and thus lead to eutrophication, and their use is now banned by law. For a brief time an effective builder designated NTA was promoted as a phosphate substitute, until it was found that it solubilizes metal and may lead to metal poisoning. Less harmful, but less effective, builders such as sodium carbonate are now widely used in detergents. *See* EUTROPHICATION.

Many additives are used in detergents to provide scent, brightening (usually through fluorescent action), or bleaching action. For a time, enzymes which cleave protein molecules (proteolytic) or fats (lipolytic) were widely used as additives, but they do not add significantly to cleansing action. Biodegradability is essential for detergents; it ensures that components of detergents will be broken down by bacterial action before undesirable aftereffects can occur. Nonbiodegradable detergents can prevent effective bacterial action in septic tanks and sewage treatment plants, and can cause undesirable persistent foaming in rivers.

Allen L. Hanson

Bibliography. M. Ash and I. Ash, *Encyclopedia of Surfactants*, vol. 1, pts. A–F, 1980; A. S. Davidsohn, *Synthetic Detergents*, 7th ed., 1987; S. E. Manahan, *Environmental Chemistry*, 7th ed., 1999.

Determinant

The concept of a determinant can best be explained with reference to a matrix A. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

be an n by n square array of numbers. Such an array is called an n by n square matrix, and the numbers a_{ij} in the array are called elements of the matrix. Determinants are valuable in the solution of sets of linear equations. The determinant of the matrix A, as indicated by Eq. (1), is a number which

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \tag{1}$$

is the value of a certain function of the elements of A.

Before defining the determinant $|A|$, it may be noted that the theory of determinants had its origin in the solution of linear systems of equations and that determinants occur in virtually all branches of mathematics and related fields. *See* LINEAR SYSTEMS OF EQUATIONS; MATRIX THEORY; POLYNOMIAL SYSTEMS OF EQUATIONS.

In the linear system of equations shown as Eqs. (2), successive elimination of the unknowns x and y yields a simpler equivalent system in Eqs. (3). The

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2 \end{aligned} \tag{2}$$

$$\begin{aligned} (a_{11}a_{22} - a_{12}a_{21})x &= b_1a_{22} - a_{12}b_2 \\ (a_{11}a_{22} - a_{12}a_{21})y &= a_{11}b_2 - b_1a_{21} \end{aligned} \tag{3}$$

coefficients of Eqs. (3) are determinants of 2 by 2 matrices

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

$$\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} = b_1a_{22} - a_{12}b_2$$

$$\begin{vmatrix} a_{11} & b_1 \\ a_{2;} & b_2 \end{vmatrix} = a_{11}b_2 - b_1a_{21}$$

with elements which are the coefficients of Eq. (2).

If

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \neq 0$$

then the solution of Eqs. (2) is readily given by

ratios of the above determinants. Similarly in the system

$$a_{11}x + a_{12}y + a_{13}z = b_1$$

$$a_{21}x + a_{22}y + a_{23}z = b_2$$

$$a_{31}x + a_{32}y + a_{33}z = b_3$$

the determinant shown in Eq. (4) enters into the solution of the system.

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \quad (4)$$

The determinant of Eq. (1) is called a determinant of order n, and it can be defined by induction on n. For explanatory purposes, determinants of orders $n = 1$, 2, and 3 are defined as follows:

For $n = 1$,

$$|a_{11}| = a_{11}$$

For $n = 2$,

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\,|a_{22}| - a_{12}\,|a_{21}| = a_{11}a_{22} - a_{12}a_{21}$$

For $n = 3$,

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

This definition is seen to agree with the definitions for determinants of order 2 and 3 and, by the principle of mathematical induction, defines a determinant of order n for every n. There are many other definitions of a determinant of order n which can be proved to be equivalent to the definition given here. The determinants of order 2 and 3 are defined in terms of determinants of order 1 and 2, respectively. Suppose now that a determinant of order $n - 1$ has been defined. In Eq. (1), denote by A_{1j}, $j = 1, 2, \ldots, n$, the determinant of order $n - 1$ formed by deleting from Eq. (1) the first row and the jth column. Equation (5) then, by definition, holds.

$$|A| = a_{11}A_{11} - a_{12}A_{12} + a_{13}A_{13} + \cdots + (-1)^{n-2}a_{1n-1}A_{1n-1} + (-1)^{n-1}a_{1n}A_{1n} \quad (5)$$

As an example of Eq. (5), the relations shown in Eqs. (6) are formed.

$$|A| = \begin{vmatrix} -2 & 0 & 1 & 3 \\ 5 & 2 & -6 & 1 \\ 0 & 4 & -1 & 7 \\ 10 & 2 & 3 & 1 \end{vmatrix}$$

$$= (-2)\begin{vmatrix} 2 & -6 & 1 \\ 4 & -1 & 7 \\ 2 & 3 & 1 \end{vmatrix} - (0)\begin{vmatrix} 5 & -6 & 1 \\ 0 & -1 & 7 \\ 10 & 3 & 1 \end{vmatrix} + (1)\begin{vmatrix} 5 & 2 & 1 \\ 0 & 4 & 7 \\ 10 & 2 & 1 \end{vmatrix} - (3)\begin{vmatrix} 5 & 2 & -6 \\ 0 & 4 & -1 \\ 10 & 2 & 3 \end{vmatrix} \quad (6a)$$

$$\begin{vmatrix} 2 & -6 & 1 \\ 4 & -1 & 7 \\ 2 & 3 & 1 \end{vmatrix}$$

$$= (2)\begin{vmatrix} -1 & 7 \\ 3 & 1 \end{vmatrix} - (-6)\begin{vmatrix} 4 & 7 \\ 2 & 1 \end{vmatrix} + (1)\begin{vmatrix} 4 & -1 \\ 2 & 3 \end{vmatrix}$$

$$= 2(-1 - 21) + 6(4 - 14) + (12 + 2) = -90 \quad (6b)$$

$$\begin{vmatrix} 5 & 2 & 1 \\ 0 & 4 & 7 \\ 10 & 2 & 1 \end{vmatrix}$$

$$= (5)\begin{vmatrix} 4 & 7 \\ 2 & 1 \end{vmatrix} - (2)\begin{vmatrix} 0 & 7 \\ 10 & 1 \end{vmatrix} + (1)\begin{vmatrix} 0 & 4 \\ 10 & 2 \end{vmatrix}$$

$$= 5(4 - 14) - 2(0 - 70) + (0 - 40) = 50 \quad (6c)$$

$$\begin{vmatrix} 5 & 2 & -6 \\ 0 & 4 & -1 \\ 10 & 2 & 3 \end{vmatrix}$$

$$= (5)\begin{vmatrix} 4 & -1 \\ 2 & 3 \end{vmatrix} - (2)\begin{vmatrix} 0 & -1 \\ 10 & 3 \end{vmatrix} + (-6)\begin{vmatrix} 0 & 4 \\ 10 & 2 \end{vmatrix}$$

$$= 5(12 + 2) - 2(0 + 10) - 6(0 - 40) = 290 \quad (6d)$$

Hence

$$|A| = (-2)(-90) + (1)(50) - (3)(290) = -640$$

It can be proved from Eq. (5) that the determinant $|A|$ is the sum of all terms of the form $(-1)^j a_{1j_1}, a_{2j_2}, \ldots, a_{nj_n}$ for all possible orderings $j_1, j_2, \ldots, j_n$ of the second subscripts $1, 2, \ldots, n$, and the number j is the number of interchanges of two digits required to carry the ordering $j_1, j_2, \ldots, j_n$ into the natural ordering $1, 2, \ldots, n$. Thus $|A|$ is the sum of all products of n elements, one from each row and one from each column, with a certain sign affixed to each product. Since there are $n! = 1 \cdot 2 \cdot 3 \ldots (n - 1) \cdot n$ orderings of the numbers $1, 2, \ldots, n$, there are $n!$ terms in the sum.

In the definition of a determinant in Eq. (5), A_{1j}, $j = 1, 2, \ldots, n$, which is a determinant of order $n - 1$, is called the complementary minor of the element a_{1j} of the first row. In general, the complementary minor of any element a_{ij} is the determinant of order $n - 1$, A_{ij}, which is formed from A by deleting the ith row and jth column. The cofactor of the element a_{ij} is $(-1)^{i+j}A_{ij}$. Thus in Eq. (5), the value of $|A|$ is given as the sum of the products of each element of the first row and its cofactor. It can be shown that for any row i, $|A| = a_{i1}(-1)^{i+1} A_{i1} + a_{i2}(-1)^{i+2} A_{i2} + \cdots + a_{in}(-1)^{i+n} A_{in}$. This expansion of $|A|$ is called the expansion of $|A|$ according to the elements of the ith row. Thus in Eq. (6),

$$
\begin{aligned}
|A| = (0)\,(-1)^{3+1}&\begin{vmatrix} 0 & 1 & 3 \\ 2 & -6 & 1 \\ 2 & 3 & 1 \end{vmatrix} \\
+ (4)(-1)^{3+2}&\begin{vmatrix} -2 & 1 & 3 \\ 5 & -6 & 1 \\ 10 & 3 & 1 \end{vmatrix} \\
+ (-1)(-1)^{3+3}&\begin{vmatrix} -2 & 0 & 3 \\ 5 & 2 & 1 \\ 10 & 2 & 1 \end{vmatrix} \\
+ (7)(-1)^{3+4}&\begin{vmatrix} -2 & 0 & 1 \\ 5 & 2 & -6 \\ 10 & 2 & 3 \end{vmatrix} = -640
\end{aligned}
$$

is the expansion of $|A|$ according to the third row.

The transpose of the square matrix A is the matrix, denoted by A', which has as its ith row the ith column of A and as its jth column the jth row of A for all $i, -j$. It can be proved that $|A'| = |A|$. It follows from this property that any theorem concerning the rows of a determinant implies a corresponding result concerning the columns, and vice versa. For example, a determinant can be evaluated by an expansion according to the jth column, by which $|A|$ is the sum of the products of each element of the jth column and its cofactor.

The following further properties of determinants follow from Eq. (5):

(i) If B is a matrix obtained from A by interchanging two rows (columns) of A, then $|B| = -|A|$.

(ii) If two rows (columns) of A are identical, then $|A| = 0$.

(iii) If B is a matrix obtained from A by multiplying every element of one row (column) of A by the number m, then $|B| = m|A|$.

(iv) If B is a matrix obtained from A by adding to each element of a row (column) of A a constant multiple of the corresponding element of another row (column) of A, then $|B| = |A|$.

(v) If B is a matrix identical with A except possibly for the ith row $b_{i1}, b_{i2}, \ldots, b_{in}$, and if C is a matrix identical with A except that the ith row of C is $a_{i1} + b_{i1}, a_{i2} + b_{i2}, \ldots, a_{in} + b_{in}$, then $|A| + |B| = |C|$. (There is the corresponding property for columns.)

These properties are used to shorten computation in finding the value of a determinant. In Eqs. (6),

$$
\begin{aligned}
|A| &= \begin{vmatrix} -2 & 0 & 1 & 3 \\ 5 & 2 & -6 & 1 \\ 0 & 4 & -1 & 7 \\ 10 & 2 & 3 & 1 \end{vmatrix} \\
&= 2\begin{vmatrix} -2 & 0 & 1 & 3 \\ 5 & 1 & -6 & 1 \\ 0 & 2 & -1 & 7 \\ 10 & 1 & 3 & 1 \end{vmatrix} \text{ by (iii) applied to the second column} \\
&= 2\begin{vmatrix} -2 & 0 & 1 & 3 \\ 5 & 1 & -6 & 1 \\ -10 & 0 & 11 & 5 \\ 10 & 1 & 3 & 1 \end{vmatrix} \text{ by (iv) multiplying the second row by } -2 \text{ and adding to the third row} \\
&= 2\begin{vmatrix} -2 & 0 & 1 & 3 \\ 5 & 1 & -6 & 1 \\ -10 & 0 & 11 & 5 \\ 5 & 0 & 9 & 0 \end{vmatrix} \text{ by (iv) multiplying the second row by } -1 \text{ and adding to the fourth row} \\
&= 2(1)(-1)^{2+2}\begin{vmatrix} -2 & 1 & 3 \\ -10 & 11 & 5 \\ 5 & 9 & 0 \end{vmatrix} \text{ by the expansion according to the elements of the second column} \\
&= 2\begin{vmatrix} 0 & 1 & 0 \\ 12 & 11 & -28 \\ 23 & 9 & 27 \end{vmatrix} = 2(1)(-1)^{1+2}\begin{vmatrix} 12 & -28 \\ 23 & -27 \end{vmatrix} \\
&= -8\begin{vmatrix} 3 & -7 \\ 23 & -27 \end{vmatrix} \\
&= -8(-81 + 161) \\
&= -640 \text{ by (iv), (iii), and Eq. (5)}
\end{aligned}
$$

If A and B are square matrices of order n, then the product AB is a square matrix of order n, and the element in the ith row and jth column, for all i, j, is obtained by multiplying the n elements in the ith row of A into the n elements in the j column of B, term by term, and adding these products. A most useful property of determinants is the fact that the determinant of the product AB is equal to the product of the determinant of A and the determinant of B; that is, $|AB| = |A|\,|B|$.

In the matrix A, select any r rows and r columns ($r \leq n$). The elements common to these rows and columns form an r by r matrix M, and the determinant $|M|$ is called an r-rowed minor of A. The determinant of the $n - r$ by $n - r$ matrix of elements common to the remaining $n - r$ rows and columns of A is called the complementary minor of $|M|$. When

$r = 1$ and row i and column j are selected, then $|M| = a_{ij}$ and the complementary minor is A_{ij}, which was defined earlier. The Laplace expansion of $|A|$ gives the value of $|A|$ as a sum, with certain signs, of all possible minors formed from a fixed set of r rows multiplied by their complementary minors. If $i_1, i_2, \ldots, i_r$ are the fixed rows, then the term in the sum, which is the minor formed from the columns $j_1, j_2, \ldots, j_r$ multiplied by its complementary minor, has the sign $(-1)^{I+J}$, where $I = i_1 + i_2 + \cdots + i_r$ and $J = j_1 + j_2 + \cdots + j_r$. In Eqs. (6), choosing the second and third rows as the fixed rows gives

$$
\begin{aligned}
|A| = {} & (-1)^{2+3+1+2}\begin{vmatrix} 5 & 2 \\ 0 & 4 \end{vmatrix}\begin{vmatrix} 1 & 3 \\ 3 & 1 \end{vmatrix} \\
& + (-1)^{2+3+1+3}\begin{vmatrix} 5 & -6 \\ 0 & -1 \end{vmatrix}\begin{vmatrix} 0 & 3 \\ 2 & 1 \end{vmatrix} \\
& + (-1)^{2+3+1+4}\begin{vmatrix} 5 & 1 \\ 0 & 7 \end{vmatrix}\begin{vmatrix} 0 & 1 \\ 2 & 3 \end{vmatrix} \\
& + (-1)^{2+3+2+3}\begin{vmatrix} 2 & -6 \\ 4 & -1 \end{vmatrix}\begin{vmatrix} -2 & 3 \\ 10 & 1 \end{vmatrix} \\
& + (-1)^{2+3+2+4}\begin{vmatrix} 2 & 1 \\ 4 & 7 \end{vmatrix}\begin{vmatrix} -2 & 1 \\ 10 & 3 \end{vmatrix} \\
& + (-1)^{2+3+3+4}\begin{vmatrix} -6 & 1 \\ -1 & 7 \end{vmatrix}\begin{vmatrix} -2 & 0 \\ 10 & 2 \end{vmatrix} \\
= {} & (20)(-8) - (-5)(-6) + (35)(-2) \\
& + (22)(-32) - (10)(-16) + (-41)(-4) \\
= {} & -640
\end{aligned}
$$

The Laplace expansion is particularly useful in evaluating the determinants of matrices which have blocks of zeros, for example, using the first three rows as the fixed rows,

$$
\begin{aligned}
& \begin{vmatrix} -5 & 4 & -1 & 0 & 0 & 0 \\ 2 & 6 & 3 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & -1 & 0 \\ 0 & 0 & 0 & 2 & 7 & 1 \\ 0 & 0 & 0 & 0 & 5 & 3 \end{vmatrix} \\
& \qquad = \begin{vmatrix} -5 & 4 & -1 \\ 2 & 6 & 3 \\ 0 & 1 & -1 \end{vmatrix}\begin{vmatrix} 4 & -1 & 0 \\ 2 & 7 & 1 \\ 0 & 5 & 3 \end{vmatrix} \\
& \qquad = (51)(70) = 3570
\end{aligned}
$$

Every minor formed from the first three rows is zero by property (ii), except the one formed from the first three columns. *See* EQUATIONS, THEORY OF.

Ross A. Beaumont

Bibliography. R. A. Barnett and M. R. Ziegler, *College Algebra*, 6th ed., 1998; M. D. Grady, I. Drooyan, and E. F. Beckenbach, *College Algebra*, 9th ed., 1994; P. K. Rees, F. W. Sparks, and C. S. Rees, *College Algebra*, 10th ed., 1990.

Determinism

The principle that nature follows exact laws, so that what will happen in the future is a necessary consequence of the state of the world at any given moment in the past. This view, if fully adopted, implies that events which seem to occur by chance would be fully understood if more was known about them, and that apparently free thoughts and choices are explainable and in principle predictable in terms of neuroscience. In a looser sense, determinism refers to claims that mental freedom is much more restricted than is generally supposed. An immense philosophical literature has arisen in response to this challenge to the sense of humanity, most of it defending the freedom of a person's choices and the necessity that the individual be held responsible for them.

The question of determinism in physical science cannot be considered apart from the philosophical problem; this gives it added importance and forces its consideration in a very critical spirit.

Prequantum ideas. The idea that the world is composed of atoms moving under the influence of certain forces according to certain laws can be traced back to the Greek philosopher Leucippus, who wrote in the fifth century B.C., "Nothing happens at random but everything for a reason and by necessity." In subsequent centuries this idea of necessity was never entirely lost. It was prominent in the work of the seventeenth-century thinker René Descartes, and became widely known through his influence. Isaac Newton carried out a large part of the cartesian scientific program. His theory explained so many natural processes that it began to appear that the universe since the time of Creation might actually have run its course in a deterministic fashion like a machine, without divine intervention. Newton sought to avoid this conclusion by suggesting that gravitation may be an act of God, and that at any rate He must continually intervene to keep the starry world from collapsing under the gravitational force. The negative implications for free will were widely considered during the seventeenth century, and Descartes was among those who made suggestions for avoiding them.

A century after Newton, Pierre Simon de Laplace argued that an Omniscient Calculator, provided with exact knowledge of the state of the universe at present, would be able to predict the entire future. *See* CLASSICAL MECHANICS; NEWTON'S LAWS OF MOTION.

Quantum physics. In the early 1900s a reaction began against an exclusively mechanical explanation of nature. It took the form of a largely unjustified condemnation of atomic theory on the grounds that the proper concern of science is with things which can be observed and which do not have to be imagined. The great successes of atomic theory soon put an end to this skepticism, but enough of it remained to make easier the later renunciation of determinism on the atomic scale.

The quantum mechanics of the 1920s opened the way to a reassessment of determinism in nature when it introduced the paradox of particles which are, at

the same time, waves. A wave of length λ was supposed to accompany, or describe, a particle of momentum $p = h/\lambda$, where h is Planck's constant. But a wave with only the single Fourier component λ extends over all space from $-\infty$ to $+\infty$ and says nothing about the location of the particle. The solution to this problem was given by Fourier's superposition of waves and Max Born's probabilistic interpretation of quantum mechanics: express the probability of finding a free particle by a (complex) wave packet (**Fig. 1**). This can be represented by a superposition of waves with wavelengths in the neighborhood of λ, and according to Born the particle has appreciable probability of being found only where the wave function ψ (or more precisely, $|\psi|^2$) is large; that is, within a region of size roughly Δx on each side of x_0 in Fig. 1. It is convenient to introduce the quantity $k = 2\pi/\lambda$, in terms of which the relation $p = h/\lambda$ becomes $p = \hbar k$, where $\hbar = h/2\pi$. A theorem of Fourier analysis shows that Δk, the range of k's present in the wave packet, is related to Δx by relation (1). Each side of relation (1) can be multiplied by $\hbar$ to get relation (2). This states that the more

$$\Delta x \, \Delta k \geq 1 \tag{1}$$

$$\Delta x \, \Delta p \geq \hbar \tag{2}$$

precisely the specification of a particle's position x is attempted by means of a localized wave function (the only means available in quantum mechanics), the less precisely can its momentum be specified. *See* FOURIER SERIES AND TRANSFORMS.

Relation (2), a trivial consequence of Fourier's theory, was written down by Werner Heisenberg in 1927. It immediately raised the question of whether the inability of the theory to make more precise specifications was a blemish on the theory or whether it reflected some previously unsuspected property of nature. Heisenberg broke the three-century-old tradition of Descartes by choosing the second alternative. His principle of uncertainty, or indeterminacy, states that it is impossible to measure, and therefore to know, x and p any more accurately than is allowed by relation (2), and there are also other pairs of dynamical variables similarly related.

Heisenberg and Niels Bohr worked out an example to illustrate the uncertainty principle (**Fig. 2**). The least possible amount of light is a single quantum, a photon. Imagine an idealized optical arrangement in which this is enough to locate the particle P. The

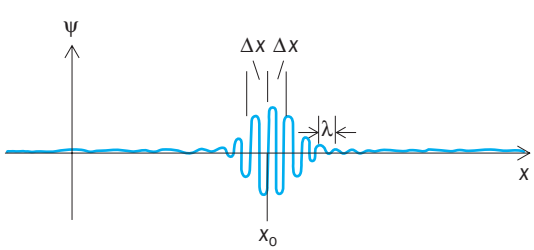

Fig. 1. A wave packet—of wavelength approximately λ, and produced perhaps by opening and closing a shutter—traveling through space.

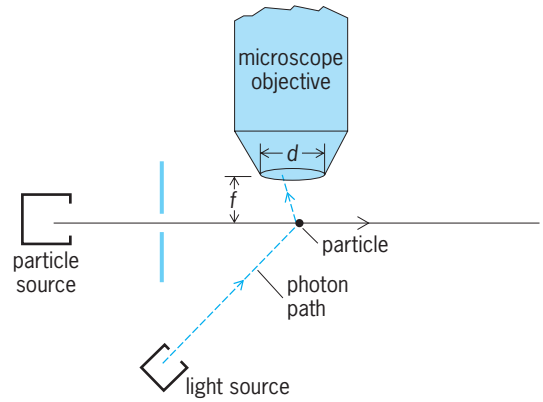

Fig. 2. Example of the uncertainty principle.

quantum emerges from the light source and bounces off P and into the lens of the microscope. In doing so, it changes the particle's momentum. If the quantum could be located after it hit P, it would be possible to allow for this recoil; but it is only known that the quantum entered the lens. The uncertainty in the recoil produces an uncertainty Δp in the particle's subsequent momentum given by relation (3), where

$$\Delta p \geq p\frac{d}{f} = \frac{hd}{\lambda f} \tag{3}$$

p is the momentum of the photon, d is the diameter of the lens, f is the distance from the particle to the lens, and λ is the wavelength of the light emitted by the light source. But by the principles of optics a microscope cannot locate a particle perfectly; the finite wavelength of the light gives a fuzziness in the image which leads to an uncertainty in the particle's measured position given by relation (4).

$$\Delta x \geq \frac{\lambda f}{2\pi d} \tag{4}$$

Multiplying these relations together gives relation (2), independently of any choices made in designing and using the microscope. Many physicists have tried to circumvent relation (2) by inventing clever devices, but it has always turned out that there is a flaw in the reasoning. *See* UNCERTAINTY PRINCIPLE.

There remains the question as to whether Heisenberg's principle is merely an unfortunate limitation on an experimenter's ability to know or whether it goes deeper. The general opinion of physicists is that of Bohr: the principle expresses a limitation of the precision with which concepts such as position and momentum can be applied at all. Lacking these concepts, Laplace's Omniscient Calculator cannot predict the future. (Nevertheless, the laws of quantum mechanics are supposed to be exact. They refer to probabilities, however, and so would not help the Calculator.)

The fall of determinism produced a profound effect among scientists and philosophers. Some welcomed it; others denied it. Einstein made the much quoted remark, "I do not believe that God plays dice," but in general, scientific opinion accepts the new situation. Whether the human mind is freed

from the bonds of determinism by Heisenberg's principle is not known. Present understanding of mental functions is far too primitive to make any pronouncements on this question. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Newtonian physics: modern ideas. Perceptive mathematicians have warned that determinism is not as obvious a consequence of newtonian physics as it might appear. Henri Poincaré and Emile Borel, and, since a famous address by Nikolai Kolmogorov, a distinguished school of Russian mathematicians have proved results which must be stated mathematically, but whose general effect is that for the vast majority of dynamical systems any error in the initial conditions, however small, will be amplified, in general exponentially, and so quickly that the predicted result will soon bear no relation to reality. Thus unless it is assumed that initial conditions are known, and computation is performed, with perfect accuracy, the Omniscient Calculator will wind up getting everything wrong.

An example of this behavior (although not exponential) involves a ball that is dropped onto a knife edge. If the ball is dropped from a point Q a little to the left of the edge, it will end up on the left side of the knife. If Q is slowly shifted toward the right, a location will be reached at which the ball begins to fall to the right. Whether or not there exists an intermediate place Q_0 from which the ball will bounce up and down on the knife edge and finally come to rest there forever, it is apparent that very small departures on either side of the real or hypothetical Q_0 will cause the ball to end up in very different places. One result of the mathematical work is to show that similar instabilities occur in most dynamical systems even if there are no knife edges. *See* CATASTROPHE THEORY.

Very few people now think that all events in the natural world are exactly determined. Experiments suggest that some human and animal behavior can reliably be predicted and controlled, but nobody knows the limits within which this can be done. David Park

Bibliography. E. G. D. Cohen (ed.), *Fundamental Problems in Statistical Mechanics III*, 1975; W. Heisenberg, *The Physical Principles of the Quantum Theory*, 1930; R. Penrose, *The Emperor's New Mind*, 1989; D. Ruelle, *Chance and Chaos*, 1993.

Deuterium

The isotope of the element hydrogen with atomic weight 2.0144 and symbols ^2H or D. Considerations of nuclear stability and a discrepancy between the chemical and physical atomic weights of hydrogen led to the prediction of a stable isotope of hydrogen of mass 2. A successful search for this isotope, deuterium, was made by H. C. Urey, F. G. Brickwedde, and G. M. Murphy in 1931. The terrestrial natural abundance of deuterium is 1 in 6700 parts of ordinary hydrogen (protium), of atomic weight 1.0078. Small variations in natural sources are found as a result of fractionation by geological processes. Industrial hydrogen, particularly that generated by electrolysis of water, may contain significantly less deuterium.

Deuterium is used mainly in the form of heavy water. In the uncombined state it finds uses as a research tool. Liquid deuterium is used in bubble chambers to study the reactions of elementary particles with the deuterium nucleus, the deuteron. Deuterons are frequently accelerated in cyclotrons to study their reactions with other nuclei and also to produce radioactive nuclides. Deuterium gas is used in the direct synthesis of organic compounds for tracer studies. Illustrative of such procedures are exchange reactions with hydrogen-containing substances in the presence of hydrogenation catalysts. If controlled thermonuclear fusion can be achieved, deuterium gas would become an exceedingly important source of power. *See* NUCLEAR FUSION; NUCLEAR REACTION.

Deuterium, D_2, is a gas at room temperature. It is prepared from heavy water, D_2O, either by electrolysis or by reaction of D_2O with metals such as zinc, iron, calcium, and uranium. It is also prepared directly by the fractional distillation of liquid hydrogen. In this process it is necessary to catalyze the disproportionation (unsymmetrical dissociation and recombination) of HD into H_2 and D_2 in order to obtain D_2.

Ordinary deuterium is a molecular mixture of two-thirds *ortho*- and one-third *para*-deuterium. At 20 K the equilibrium composition is 97.8% *o*-deuterium. Deuterium molecules obey the Bose-Einstein statistics, and the ortho species have even rotational quantum numbers, whereas the para species have odd rotational quantum numbers. The analysis of the ortho-para composition of deuterium is most conveniently made by measurement of the thermal conductivity of the gas at 77 K. The physical and chemical properties of *o*- and *p*-deuterium are very similar. In most of its physical and chemical properties, deuterium resembles protium. In most cases, deuterium is slightly less reactive than protium. An intercomparison of some of the physical properties of D_2 with those of H_2 is given in the **table**.

Deuterium gas is usually slightly contaminated with HD. The analysis of the gas for protium is most conveniently carried out by a mass spectrometer. Alternative methods of analysis are by thermal conductivity and the rate of effusion of the gas through an orifice. The gas can be converted to water, and the following properties have been used as a basis for analysis: infrared spectrum, density, index of refraction, and nuclear magnetic resonance.

The chemical reactivity of deuterium is less than that of hydrogen because of its lower zero-point energy and smaller collision frequency. At 1000 K deuterium is 32% less dissociated than protium. At room temperature deuterium atoms are electrolyzed out of water in the form of hydrogen gas, at one-eighth the rate of protium atoms. *See* DEUTERON; HEAVY WATER; HYDROGEN; TRITIUM. Jacob Bigeleisen

Selected values of some physical properties of deuterium

Property	n-H_2	n-D_2	97.8% o-D_2
Triple point	13.96 K	18.72 K	18.63 K
Normal boiling point	20.4 K	23.6 K	23.5 K
Critical temperature	33.24 K	38.35 K	
Critical pressure	12.8 atm	16.4 atm	
Heat of fusion	28.0 cal/mole	47.0 cal/mole	
Heat of vaporization (at normal boiling point)	216 cal/mole	293 cal/mole	
Molar volumes of liquid at 20 K	28.3 ml	23.5 ml	

Bibliography. G. Choppin et al., *Radiochemistry and Nuclear Chemistry*: *Theory and Applicaions*, 2d ed., 1995; G. Friedlander et al., *Nuclear and Radiochemistry*, 3d ed., 1981; S. E. Jones, F. Scaramuzzi, and D. Worledge (eds.), *Anomalous Nuclear Effects in Deuterium*, 1991.

Deuteromycota

A heterogeneous group of anamorphic (asexual or imperfect) fungi in which sporulation may occur on separate hyphae or composite fruit bodies (conidiomata). These may or may not be differentiated into distinct conidiophores and/or conidiogenous cells (conidium-bearing hyphae or cells) which occur separately or are situated over the surface of or inside a conidioma. Some lack conidiomata and do not produce conidia but form other types of propagules. Conidium-, propagule-, or conidioma-bearing septate hyphae, or individual cells, constitute the vegetative thallus. The diagnostic feature of the group is the lack of a teleomorphic (sexual or perfect) state. There are more than 2500 accepted genera containing 21,000 species.

Even though teleomorphs are absent or have not been correlated in many members of the group, it is possible to determine the teleomorphic affinities of some deuteromycetes by the presence or absence of clamp connections, differences in septal structure, and varying proportions and nature of cell wall polysaccharides. More recently the use of molecular techniques in documenting genomes has facilitated the recognition of monophyletic groups. Most have affinities with Ascomycotina and a smaller number with Basidiomycotina. The anamorphs of Zygomycotina, Erysiphales, and Urediniomycetes are readily recognizable for various reasons and are assignable to these teleomorphic groups. It has been postulated that the Deuteromycotina are a taxon of temporary nature that would gradually disappear as the missing teleomorphs are discovered and correlated. However, some anamorphic fungi have with time become isolated geographically, genetically, spatially, and nutritionally from ancestral teleomorphs. The divergent evolution of such anamorphs and the ancestral teleomorphs has resulted in a permanent residue of organisms distinct from their teleomorphic precursors but which may be correlated with them by using molecular characters. *See* ASCOMYCOTA; BASIDIOMYCOTA; ZYGOMYCOTA.

Reproduction. The Deuteromycotina are also artificial inasmuch as the group consists of fungi or states of fungi reproducing vegetatively by mitosis, not meiosis. Some organisms, such as *Aspergillus, Penicillium*, and *Fusarium* have a parasexual process in which reciprocal exchange of chromatid segments between homologous chromosomes during mitotic propagation of the diploid nucleus occurs. Instead of the regular sequence of events which obtain in the meiotic cycle, the various processes may all be occurring at one time in one mycelium. For deuteromycetes, parasexuality brings the advantage of genetic recombination. *See* MITOSIS; RECOMBINATION (GENETICS).

Classification. Traditionally, Deuteromycotina are separated into two classes: Hyphomycetes, which are mycelial forms bearing conidia on separate hyphae or aggregations of hyphae, always external to the supporting substratum not inside discrete conidiomata; and Coelomycetes, which are forms that produce conidia in pycnidial, pycnothyrial, acervular, cupulate, or stromatic conidiomata. A third class is sometimes recognized, the Agonomycetes (Mycelia Sterilia); these fail to produce conidia and thus lack conidiomata, although they do form somatic survival propagules. Currently, no ranks are distinguished within the Deuteromycotina, and the subdivisional name itself is no longer in use, the preferred term being Mitotic Fungi or Fungi Anamorphici. The current trend which will obtain for the forseeable future is for deuteromycetes to be incorporated into teleomorph classifications by using traditional morphological approaches combined with molecular systematics. *See* AGONOMYCETES; COELOMYCETES; HYPHOMYCETES.

Asporogenous (imperfect) yeasts have sometimes been separated into the Blastomycetes, which are divided into two orders—the Cryptococcales, which have both ascomycete and basidiomycete affinity and reproduce by budding only, and the Sporobolomycetales, which have a basidiomycete affinity and reproduce by both budding and ballistoconidia. *See* BLASTOMYCETES.

Commercial utilization. Deuteromycetes are ubiquitous and occupy most conceivable ecological niches, whether they are aquatic, terrestrial, or artificial. Deuteromycetes have been used extensively in manufacturing processes of different kinds. Fermentation by yeasts and enzyme production by filamentous fungi such as *Aspergillus* is the basis for the alcoholic beverage industry and many Asian

specialty food products. Various *Penicillium* species are used in the production of blue cheeses. Many deuteromycetes produce secondary metabolites and are used in antibiotic and organic acid production processes. Fungal protein is manufactured by using species of *Fusarium, Trichoderma*, and *Verticillium;* other fungal biomass, from yeasts such as *Candida* and filamentous fungi such as *Paecilomyces* and *Sporotrichum*, is produced by utilization of paper mill liquid wastes. Plant wastes or residues, such as cereal straw, can be treated by cellulytic deuteromycetes, such as *Trichoderma viride*, resulting in liberated glucose. The product is then extracted and dried, and the unhydrolyzed residues are used either as adjuncts to animal feed, for fuel, or for soil amendments in agriculture and horticulture. *Penicillium, Aspergillus*, and *Saccharomyces* are sometimes used for the extraction and recovery of metals from natural sources and industrial effluents. In natural habitats the deuteromycetes are primarily responsible for the conditioning or mineralization of plant litter. Increasingly important is the use of deuteromycetes in biological control of insects, nematodes, pathogenic fungi, and weeds. *See* DISTILLED SPIRITS; FOOD MANUFACTURING.

Medical mycology. All fungal infections of humans are opportunistic, and mycoses are almost always the result of abnormal host immunity. Even before the association of the HIV virus with acquired immune deficiency syndrome (AIDS), criteria used in defining the syndrome relied heavily on the existence of certain opportunistic infections in the absence of any of the established predisposing conditions. A significant number of these deep-seated infections are caused by deuteromycetes such as *Aspergillus, Candida, Cryptococcus*, and *Histoplasma*. Superficial mycoses due to dermatophytes such as *Trichophyton* are commoner in patients with AIDS. Deuteromycetes also cause numerous diseases of plants; these are often of considerable economic importance, as are the losses caused in stored natural and manufactured products and processed food. In addition, mycotoxins may be produced directly by growth of *Fusarium, Penicillium*, and *Aspergillus* and many others on animal feed or human food. Such mycotoxins may pass through the food chain to animal products such as milk or meat, and illnesses arising from such sources are termed secondary mycotoxicoses. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); MEDICAL MYCOLOGY; MYCOTOXIN; OPPORTUNISTIC INFECTIONS; PLANT PATHOLOGY.

B. C. Sutton

Bibliography. D. Allsopp and K. J. Seal, *Introduction to Biodeterioration*, 1986; D. L. Hawksworth, *Ascomycete Systematics*, 1994; D. L. Hawksworth et al. (eds.), *Ainsworth & Bisby's Dictionary of the Fungi*, 8th ed., 1995; W. B. Kendrick, *The Whole Fungus*, 1979; D. R. Reynolds and J. W. Taylor (eds.), *The Fungal Holomorph*, 1993; J. E. Smith and M. O. Moss, *Mycotoxins*, 1985; J. M. B. Smith, *Opportunistic Mycoses of Man and Other Animals*, 1989; B. C. Sutton, *A Century of Mycology*, 1996.

Deuteron

The nucleus of the atom of heavy hydrogen, H^2 (deuterium). The deuteron *d* is composed of a proton and a neutron. As the simplest multinucleon nucleus, the deuteron has been the subject of extensive study. Its binding energy is 2.227 MeV; that is, this is the amount of energy which must be added to a deuteron for it to dissociate into a proton and a neutron. The accurate determination of this dissociation energy provides the means of calculating the mass of the neutron, the mass of the deuteron (2.014187 amu) and proton being known from other experiments.

The intrinsic angular momenta, or spins, of the proton and neutron combine to produce a deuteron spin of unity; hence, the deuteron obeys the type of quantum statistics that is known as Bose-Einstein statistics. The deuteron possesses a magnetic moment (0.857407 nuclear magneton) and an electric quadrupole moment (2.738×10^{-27} cm^2).

Deuterons are much used as projectiles in nuclear bombardment experiments, especially to produce (d,p), (d,n), and (d,α) reactions. In the first two reactions, because of the low binding energy of the deuteron, the neutron *n* or proton *p* is stripped from it and captured by the target nucleus. Meanwhile, the other half of the deuteron (that is, the proton or neutron) carries away the excess energy. The H^1/H^2 abundance ratio in nature is 6700. *See* DEUTERIUM; NUCLEAR REACTION.

Henry E. Duckworth

Deuterostomia

A major division of the animal kingdom comprising the phyla Hemichordata, Echinodermata, and Chordata. Hemichordata consists of acorn worms (Enteropneusta) and a group of small-bodied, tentaculate forms (Pterobranchia). Echinodermata contains the starfish, sea urchins, crinoids (sea lilies), and their allies. The phylum Chordata includes the Urochordata (sea squirts and larvaceans), Cephalochordata (*Branchiostoma*, the sand lancelet, formerly *Amphioxus*), and vertebrates (see **illustration**). At one time, up to 10 phyla were included in Deuterostomia on the basis of shared morphological and developmental features; however, comparison of deoxyribonucleic acid (DNA) sequences indicates that most of those phyla are more closely allied to other groups. *See* ANIMAL KINGDOM.

General characteristics. A number of morphological and developmental features are widely shared among deuterostomes; for example, the adult mouth always forms at some distance from the blastopore (the opening into the embryonic gut), a condition termed deuterostomy, which gave the group its name. In many (but not all) deuterostomes, early embryos show radial cleavage and form coelomic body cavities (which house the internal organs) from outpockets of the embryonic gut. However, none of those features is exclusive to the group, which means that they were either present before deuterostomes evolved (and thus were inherited by other

groups as well) or have evolved in other groups independently. One feature that is exclusive to deuterostomes, though not present universally among them, is pharyngotremy, characterized by a perforation of the pharynx that permits an outflow of water that has been ingested through the mouth, usually bathing gills; it is likely a primitive character (heritable trait) of this crown group (that is, the last common ancestor of all living deuterostomes, plus all of its descendants).

Hemichordata. Molecular phylogenetic studies indicate that hemichordates and echinoderms are more closely related to each other than either is to chordates. Among hemichordates, the benthic, vermiform enteropneust body plans are the least derived (advanced): their coeloms are divided into three successive anteroposterior regions (the proboscis, collar, and trunk), and the nervous system consists of a subepidermal plexus that evidently provides most of the integration and coordination, with supplementary dorsal and ventral nerve cords. *See* HEMICHORDATA.

Echinodermata. The echinoderms are highly derived, and while three anteroposterior coelomic regions are also present during early development, they become modified, partly conjoined, and partly lost as the developing body plan produces the adult pentameral symmetry found in starfish and most other living classes. The echinoderm nervous system is also a subepidermal plexus. Echinoderm bodies are partly supported and protected by a carbonate endoskeleton of spicules or, commonly, fused plates. Living echinoderm groups do not show pharyngotremy, but extinct primitive groups, which lacked pentameral symmetry, may have displayed this characteristic. *See* ECHINODERMATA.

Chordata. The chordates evidently share a common vermiform ancestor with enteropneusts, but chordate ancestors evolved adaptations for swimming that led to an innovative body plan. Chordates evolved a springy dorsal structure, the notochord, which antagonized lateral swimming motions that were produced by anteroposteriorly arranged muscle blocks (myotomes), as well as a lateral flattening of the body, unusual in organisms living in benthic habitats but common among active swimmers. Coordination of functions was consolidated in the dorsal nerve cord, presumably positioned in association with the notochord so as to innervate the series of muscle blocks, forming a segmental system.

Acraniata and Craniata. Several important clades (groups of related organisms descended from a common ancestor) radiated from early chordates. One is Acraniata, containing the cephalochordates, particle feeders that lack axial skeletons. Another major clade is Craniata, which include agnathan vertebrate groups, jawless forms with cartilaginous skeletons represented today by hagfish and lampreys. The earliest of known Cambrian vertebrates, from approximately 520 million years ago (mya), was evidently an agnathan, as probably were the euconodonts, extinct eel-like forms having cranial cartilage and furnished with a feeding apparatus that included the first signs of bone, dating from nearly 510 mya or slightly later. Making up the bulk of Craniata today are the gnathostomes, jawed vertebrates that are often divided into seven Linnean classes that include the well-known groups of living vertebrates.

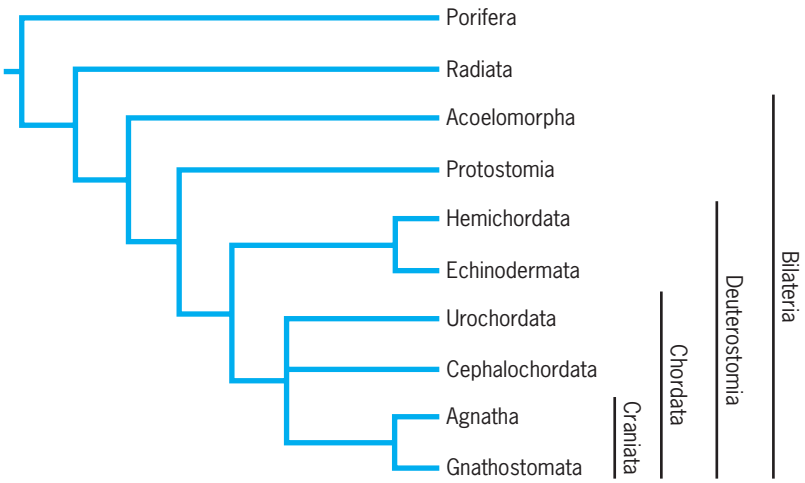

Position of Deuterostomia and some of its major divisions within the animal kingdom.

Urochordata. Another major deuterostome clade, with a considerable diversity of living marine species (about 1250), is Urochordata. It is often regarded as a phylum in its own right, but more recently has been considered a subphylum of Chordata, implying that it was part of an early chordate radiation, but as a separate branch from Acraniata and Craniata. Living members of Urochordata appear to be so highly modified and probably simplified from their ancestral condition that the ancestry of urochordates and their place in the evolution of other chordate clades has been obscured. The most obvious deuterostome characteristic of urochordates is a large, highly perforate pharynx, used as a feeding basket for filtering food items from seawater. The relation of urochordates to the chordates is indicated chiefly by the presence of a notochord in the larvae. *See* CHORDATA; VERTEBRATA.

Evolution. Molecular clock evidence (an estimate of the time of divergence of a lineage based on the rate of genetic change) suggests that the last common ancestor of deuterostomes lived over 560 mya, with the trace fossil record suggesting that this ancestor was a small bilaterian worm. The earliest undisputed deuterostome body fossils, probably dating back 525 mya, are probably extinct classes of echinoderms, easily preserved because of their durable skeletons. Spectacular fossil deposits from China, probably dating nearly 520 mya, have yielded abundant remains of soft-bodied animals, and among them are urochordates and even vertebrates, which are represented by cartilaginous forms that are probably allied to living agnathans.

Molecular studies of deuterostome genes and genomes has produced evidence of important stages in their evolution. For example, during development the genes of the *Hox* cluster help mediate the pattern of anteroposterior differentiation in all bilaterians

investigated. *Hox* genes are transcriptional regulators with homeobox motifs (DNA sequences that regulate somatic development). The last common ancestor of crown deuterostomes had at least 7 *Hox* genes. Echinoderms have 10 *Hox* genes, while cephalochordates have 14, possibly the primitive chordate number. Vertebrates have multiple clusters that have arisen by cluster duplications, 4 for mammals and 7 for some fish, but each cluster has lost many genes so that the total number of Hox genes in mammals is 39 rather than 60—each cluster averages about 10 genes. Thus in general, in the deuterostome genome a history of increasing body plan complexity is reflected by increasing numbers of *Hox* genes. However, the known genome of urochordates includes only 9 *Hox* genes; either it is primitive with respect to cephalochordates or has been simplified during evolution. *See* ANIMAL EVOLUTION; DEVELOPMENTAL GENETICS; HOMEOSIS. James W. Valentine

Bibliography. R. L. Carroll, *Patterns and Processes of Vertebrate Evolution*, Cambridge University Press, 1997; J. Cracraft and M. J. Donoghue (eds), *Assembling the Tree of Life*, Oxford University Press, 2004; H. Gee, *Before the Backbone: Views on the Origin of the Vertebrates*, Chapman and Hall, London, 1996; J. W. Valentine, *On the Origin of Phyla*, University of Chicago Press, 2004.

Developmental biology

A large field of investigation that includes the study of all changes associated with an organism as it progresses through the life cycle. The life cycles of all multicellular organisms exhibit many similarities. That is, as an organism progresses from one generation to the next there is a series of common processes: for example, gametogenesis, fertilization, embryogenesis, cell differentiation, tissue differentiation, organogenesis, maturation, growth, reproduction, senescence, and death. *See* ANIMAL GROWTH; CELL CYCLE; CELL SENESCENCE AND DEATH; EMBRYOGENESIS; FERTILIZATION; GAMETOGENESIS; PLANT GROWTH.

History

In the early nineteenth century, when professional science was emerging, the term biology was first used to include botany and zoology with agricultural and medical applications. Embryology, the first subdiscipline devoted to development, was a branch of anatomy, which at the time was concerned with organs and tissues, with pathology and teratology as sidelines. Embryology concerned itself with the observation and delineation of the stages by which an embryo was shaped into adult form: progressive changes in its size and in its organs, tissues, and cells. At the same time, the nucleate cell was being recognized as the basic unit in living plants and animals. The tiny and elusive mammalian ovum was discovered in this period.

In the second half of the century, biological discoveries abounded. Cell studies flourished, encouraged by improvements not only in optics but in other laboratory equipment and in chemicals, especially dyes. Marine organisms furnished ideal living material for the study of cellular stages in animal development. Laboratories for efficient work became organized. By the 1880s an experimental branch of embryology was an established fact: its concern was the mechanics of development.

By the end of the nineteenth century, the structural aspects of cell division (mitosis) had been observed in plant and animal tissues, including the accurate division of the chromosomes by the spindle apparatus. Fertilization of egg by sperm, with the union of the two nuclei, was recognized. Finally, meiosis, whereby the double (diploid) chromosome set was reduced to a single (haploid) condition, was elucidated. By the turn of the century, the species specificity of the chromosome sets was clear, and also the individuality of each chromosome. Almost from the time of its discovery, the chromatin apparatus of the nucleus was seen to have something to do with heredity; the rediscovery of Mendel's neglected experiments offered an obvious corroborative pattern. *See* CHROMOSOME; MEIOSIS; MITOSIS.

In the meantime the experimentalists were investigating the egg and the early stages of its development, not only in detailing cell lineages in the cleavage stages that followed insemination, but in manipulations such as hybrid fertilization, and in observing the developmental potentialities of parts of the egg or of separated blastomeres, and the direction of the cleavage pattern by pressure. External factors also were investigated: physical or chemical activation of the egg in the absence of sperm, and various developmental roles of calcium in seawater. In inland laboratories, microsurgery of vertebrate embryos was explored.

The twentieth century brought continuous expansion of experimental techniques. Microsurgery, particularly of the amphibian embryo, became a fine art. The centrifuge was used for rearranging cellular inclusions; micromanipulators were devised for surgery and injection under very high magnification; physiological apparatus was adapted for very small objects. Photomicrography was continually improved, as well as time-lapse cinematography, a great convenience in condensing a long life history for review in limited time. Optical study of living cells was greatly advanced. Dark-field and polarizing optics had long been known; interference and phase-contrast optics were developed, and electron microscopy was considered. Tissue culture was invented as a device for following cellular differentiation outside the body, and later was adapted for observation of organ and embryo development. Only much later were methods found to culture dissociated cells, and to induce cell-to-cell fusion and even hybridization of somatic cells. *See* MICROMANIPULATION; MICROSCOPE; TISSUE CULTURE.

Genetics as a science dates from about 1900. The close association between embryology, cell biology, and genetics was obvious, but functional amalgamation of these disciplines did not occur at once.

The pioneer geneticists were preoccupied with hybridization and chromosome study, and were intent on establishing a sound theory of the gene. By the 1920s, however, several distinct branches of genetics could be identified: breeding, usually associated with chromosome cytology; population genetics, associated with evolutionary, ecological, and sociopolitical orientation; and physiological genetics, which soon became developmental genetics as information accumulated relevant to the mechanism of expression of known genes. Lethal genes, as an extreme form of teratological factors, or "inborn errors of metabolism," are particularly useful for investigation of developmental processes. *See* CELL BIOLOGY; EMBRYOLOGY; GENETICS; METABOLIC DISORDERS.

In the 1930s a new trend emerged: microbial genetics. At first, eukaryote (nucleate) species with sexual cycles (for example, molds and yeasts) were shown to possess orthodox genetic systems controlling fundamental biochemical properties such as nutritional requirements and enzyme production. Eventually the prokaryote bacteria, whose possession of genes had been in question because no formed nucleus or chromosomes had been recognized, were shown to be capable of genetic exchange; they now furnish indispensable laboratory material. *See* BACTERIAL GENETICS; GENETIC ENGINEERING.

Biochemistry in this period, profiting from the almost revolutionary surge of atomic physics, was learning a great deal about macromolecules (proteins, nucleic acids, polysaccharides, lipids), to the study of which crystallographic methods were being applied. Fundamental metabolic reaction systems were beginning to be understood as serial or cyclic arrangements of substrates and enzymes. Radioactive isotopes were becoming available, to label reagents and eventually to locate reaction products within cells. Ingenious methods were devised for the separation of cellular constituents and identification of substances in very small samples. The search for the genetic material, narrowed to either protein or nucleic acid, was decided in favor of the latter. All these advances were well under way before World War II. *See* BIOCHEMISTRY.

By the end of that war, machinery had been devised for rapid analytic procedures that previously had been hopelessly tedious and time-consuming. The electron microsope was being developed and refined even so far as to resolve macromolecules, thus opening an entire new world to visual inspection. Microbiological and cell-culture methods provided experimental material supplementing large organisms, and served as excellent analytic tools. The two nucleic acids and their subunit composition were located: deoxyribonucleic acid (DNA), mainly confined to the nucleus of the cell, and ribonucleic acid (RNA), principally in the cytoplasm. The calculation of an acceptable model for DNA structure and subsequently the understanding of how the sequence of that structure could furnish a code, whereby RNA could designate protein structure (translation), constituted a spectacular step for molecular biology. The principal features of the DNA double-helix model were soon verified, and modified, by designed experiments on prokaryote material. Bacterial nucleic acid systems proved highly rewarding. Their code, designating the amino acid sequence in the primary protein assembly, was delineated, and their several functions ascertained. DNA, as predicted, acts as a template for its own replication, preceding cell division; RNA, transcribed from DNA, operates in at least three distinct forms to assemble amino acid sequences according to the transcribed code. Within a few years, protein synthesis was reconstructed in test tubes. Virus particles, with no synthesizing system of their own, were recognized as parasites consisting of a core of either DNA or RNA covered by a protein coat and capable of transferring their templates to the host cell. Bacterial as well as other DNA strands proved to break and rejoin in surprising combinations. *See* DEOXYRIBONUCLEIC ACID (DNA); ELECTRON MICROSCOPE; GENETIC CODE; MOLECULAR BIOLOGY; PROTEIN; RIBONUCLEIC ACID (RNA); VIRUS.

Dorothea Rudnick

Developmental View

As can be seen in the preceding discussion, developmental biology is a broad-ranging discipline which includes many areas of biological investigation. Thus analysis of all of the events associated with an organism as it progresses through its life cycle employs a multiplicity of approaches. In general the analysis has been reductionist in nature. Tremendous strides have been made in describing at the molecular level the developmental process of cell differentiation. However, the molecular control mechanisms which regulate cell differentiation are not known. Tissue and organ differentiation, as well as morphogenesis, are processes which have been described in detail for many situations, but little is known about the physical and chemical nature of the mechanisms involved. A complete understanding of the development of an organism will require an appreciation and comprehension of the changes which occur at all levels of organization as an organism traverses its life cycle.

The major unifying theme in biology is evolution. Not only has evolution led to the wide variety of organisms now present on Earth, but also evolution has modified the initial processes and patterns of development to the diversity of types currently encountered. This evolution of developmental parameters in multicellular organisms began as single-celled organisms became multicellular. The development of a multicellular organism entails a host of problems not faced by a single-celled organism. For example, cells in one part of the aggregate must coordinate their activities with cells in other parts, nutrients and oxygen must be provided to all cells, and water balance must be maintained. In the two major groups of multicellular organisms, plants and animals, there appear to be two differences which may account for their different solutions to the common problems associated with being multicellular. First, plants are autotrophic because they can fix carbon directly via

photosynthesis. Thus, they can remain in one spot and provide for their needs as long as light energy is available. Second, their cells are encased in a box, the cell wall, which prevents movement of cells in relation to each other and severely restricts overall movement of the organism and its parts. Thus, shape and form result primarily from cell division and expansion, and rarely from cell death. Animals, on the other hand, are usually motile and heterotrophic, obtaining food from their environment. In many instances animal cells are motile, and this makes possible additional mechanisms for attaining shape and form. With this viewpoint, one can understand many of the similarities and differences which have arisen in plants and animals as they evolved. *See* ANIMAL EVOLUTION; CELL DIVISION; CELL WALLS (PLANT); PHOTOSYNTHESIS; PLANT EVOLUTION.

Developmental biologists have focused on two central areas: the processes and associated mechanisms by which cells become different, that is, cell differentiation; and the processes and associated mechanisms by which patterns are created, that is, morphogenesis.

Totipotency. Current theories state that cells become different by expressing different genes. Thus, a liver cell is different from a muscle cell, not because it contains different genes or genetic information, but because it expresses different sets of genes. This explanation of cell differences is based upon the results of three types of experimental analysis. (1) Some plant cells are totipotent; that is, for tobacco, carrot, and a few other plant species, it has been demonstrated that a single cell (not a gamete) can divide and undergo morphogenesis to form a fertile plant. (2) Nuclei from some differentiated animal cells are totipotent. That is, a nucleus from a differentiated cell (for example, a larval intestinal cell of a frog) can be injected into a mature egg which has had its nucleus removed or destroyed, and the injected nucleus can direct normal development of a tadpole and subsequently of a fertile frog. (3) The sequences of nucleotides in the DNA of all cells in an organism appear to be the same; that is, DNA-DNA hybridization of DNA from different cell types indicates that the different cell types do not have unique DNA base sequences. Since these results indicate that all cells contain the complete genome for an organism, different cell types appear to arise as a result of the expression of unique sets of genes in each cell type. *See* CELL DIFFERENTIATION; DEVELOPMENTAL GENETICS; GENE ACTION; SOMATIC CELL GENETICS.

Determination. Initially the cells of a developing embryo are not restricted in their developmental potential or fate (that is, what type of cell they can become), but as embryogenesis proceeds, a cell's developmental potential becomes restricted or fixed. For example, after several divisions a cell from a frog embryo can develop into most frog cell types. After gastrulation, however, a surface cell in the neural plate area is restricted to develop as a neural cell type, and eventually its developmental potential will be for a specific type of neural cell. Restriction of developmental fate is called determination. Determination is an operationally defined term. In general, a cell, tissue, or organ is determined when its developmental fate is the same in its normal position in the organism, when removed from the organism and grown in isolation, and when grafted and grown at a new place on the organism. Thus, in order for determination or restriction of developmental potential to be observed, the determined state must be stable enough to withstand considerable experimental manipulation. The determined condition may however, be reversible, and the degree of stability observed varies from system to system. Determination is observed at several levels of organization (cell, tissue, organ) and occurs in both plants and animals. *See* FATE MAPS (EMBRYOLOGY).

Regulative and mosaic development. Early embryologists observed that some animal embryos divided to produce two or four cells which, when individually isolated, were capable of developing into small but normal larvae. In other species these cells were not capable of producing complete larvae but only abnormal structures which contained some of the normal cell types. The former type of embryos exhibited regulative development, while the latter exhibited mosaic development. For some time embryologists believed that these two types represented fundamentally different mechanisms of regulating development. Animals that exhibit regulative development, however, eventually exhibit mosaic development. For example, each cell of a sea urchin embryo at the four-cell stage can produce a larva, but at the eight-cell stage no single cell can undergo normal embryogenesis. Developmental biologists now believe that it is the timing of determination events which differs, and not fundamental mechanisms of development. *See* CLEAVAGE (EMBRYOLOGY); MOSAICISM.

Cytoplasmic determinants and induction. Two mechanisms have been identified that bring about determination. The first involves the presence of unique factors, called cytoplasmic determinants, which are products of the maternal genome and are located in specific areas of some animal eggs. The cells which come to contain these determinants differentiate along specific pathways. An extensively studied example is the cytoplasmic determinants in pole plasm of fruit fly (*Drosophila melanogaster*) eggs which are necessary but not sufficient for the cells to differentiate into germ cells. The second mechanism is induction, a process by which two tissues interact so that one or both differentiate along specific pathways. A classic example of induction is the action of mesoderm on the overlaying ectoderm in the frog embryo at the time of gastrulation. The mesoderm acts on the ectoderm, causing it to form the neural plate. Only ectoderm of a certain developmental age is capable of responding to the mesoderm, and this ectoderm is said to be competent. Competence or the ability to respond to an inductive signal is often a transient state. The chemical or physical base for induction and competence and the chemical identity of cytoplasmic determinants are yet to be established. *See* EMBRYONIC INDUCTION.

Molecular mechanisms. Developmental biologists have gained substantial insights into the molecular bases for determination in model organisms such as *Drosiphila*. At least three sets of cytoplasmic determinants (maternal gene products) are present in the fly egg: determinants for germ-cell formation, determinants controlling dorsal-ventral polarity, and determinants for the anterior-posterior polarity. Some of these determinants are messenger ribonucleic acids (mRNAs) coding for proteins which are transcriptional regulators (that is, proteins that regulate gene activity).

Of great interest among the genes that are activated in *Drosophila* are the homoeotic genes that specify the characteristic structures (for example, leg, wing, antenna) that are found in a particular segment of the insect. The homoeotic genes have a common, evolutionarily conserved region called the homoeobox which codes for the portion of the protein that binds to DNA and, by doing so, appears to be involved in the determination of cell fate. Homoeobox-containing genes have also been identified in plants and vertebrates. Moreover, the vertebrate *Hox* genes (vertebrate homoeobox genes) are expressed during embryogenesis appropriate for specifying cell fates. Although tremendous progress has been made in understanding cell fating in several model organisms such as *Drosophila*, there is no evidence to indicate that mammalian and plant eggs contain cytoplasmic determinants. Apparently, the initial events associated with cell fating in these groups of organisms are different from those in *Drosophila*.

Morphogenesis. Morphogenesis involves the production of form and structure by integrating the differentiation of many different cells and cell types into specific spatial patterns. This higher level of organization has been difficult to investigate in terms of establishing mechanisms. The processes of determination, competence, and induction are involved. One of the greatest challenges faced by developmental biologists is to bridge the gap between genes and patterns. It is clear that patterns are a result of gene activity, but the relationship between genes and patterns in most organisms is not well understood. *See* ANIMAL MORPHOGENESIS; PLANT MORPHOGENESIS.

Carl N. McDaniel

Bibliography. S. F. Gilbert, *Developmental Biology*, 6th ed., 2000; B. Lewin, *Genes V*, 1995; G. M. Malacinski (ed.), *Pattern Formation*: *Primer in Developmental Biology*, 1984; W. McGinnis and M. Kuziora, The molecular architects of body design, *Sci. Amer.*, 270:58–66, 1994; T. A. Steeves and I. M. Sussex, *Patterns in Plant Development*, 2d ed., 1989.

Developmental genetics

The branch of genetics which studies how genes control embryonic development. Advances in the field have emphasized the degree of conservation of the genes controlling development throughout evolution. Thus, genes which are important in such distant organisms as insects with an exoskeleton, highly segmented organization, and ventral nervous cord (*Drosophila* is the type experimental species), and vertebrates with their endoskeleton, lack of markedly segmented organization, and dorsal nervous system (the mouse is the type experimental species) share a number of highly homologous (similar) genes controlling early development. These genes are also found, and sometimes have similar roles, in descendant worms of the common ancestral annelid (although the most frequently studied descendant is a nematode, *Caenorhabditis elegans*), and some are also shared with plants.

Maternal Inheritance. Maternal inheritance is found in nematodes, mammals, and insects. For instance, many of the genes responsible for the determination of the cell's fate in the larva of *C. elegans* are expressed from messenger RNA already present in the egg; that is, they are maternal-effect genes. In this case, it is not the genotype of the zygote that influences development but that of the mother. Thus, homozygosity for a recessive mutation in the mother leads to altered development, even though the sperm is from a homozygous wild-type male and the resulting zygote was heterozygous (that is, if the organism could survive and was a female, it would lay eggs which would lead to normal development). The percentage of maternal-effect genes is also high in *Drosophila*. *See* GENE ACTION.

Imprinting. Imprinting, a developmentally important phenomenon that was first discovered in insects, is also important for mammalian development and human disease. In imprinting, genes transmitted through the testis sometimes function differently from those transmitted through the ovary. This explains why parthenogenesis does not work in mammals: the mammalian embryo requires a pronucleus derived from each sex's imprinting. Many portions of the genome have been found to be imprinted, including the reciprocal imprinting of insulinlike growth factor and its receptor. Some major human diseases occur when both a paternal and a maternal copy of a gene are not present. The Prader-Willi syndrome, a disorder of mental retardation, poor appetite regulation, and mild dysmorphic features, is an example. Advances have strongly implicated gametogenesis-specific methylation of key controlling regions in the imprinting process. Such imprints seem to be erased from the migrating germ cells en route to the developing gonad, and then are established differentially during ovigenesis and spermatogenesis, presumably by proteins uniquely expressed in the two different gonads and with specificity for the particular deoxyribonucleic acid (DNA) sequences. The actual expression of imprinting differences frequently involves competition between cis-linked genetic elements and frequently involves a nontranslated ribonucleic acid (RNA) species. *See* DEVELOPMENTAL BIOLOGY; GENETICS.

Induction. Another general phenomenon under genetic control during development is induction: the action of one cell or tissue on other cells to determine altered gene expression in them. For many years, the classic example of induction was the capacity of the dorsal blastopore lip of an amphibian embryo

to induce a second body axis after transplantation to the ventral side of another embryo; a positive-acting substance was sought. Ironically, it turns out that the signals from this organizer promote dorsal development by repressing a ventralizing signal apparently produced by a bone morphogenic protein (BMP4). Another significant inductive event, that of the notochord in patterning the neural tube, is due to the sonic hedgehog protein, one of three vertebrate hedgehog paracrine factors (that is, they affect only cells that are nearby the cell that released them).

Development control genes. Homeobox genes (*Hox* genes) are used in both insects and mammals to provide information for anterior-posterior positioning. Forkhead (forkhead box, *FOX*) genes are another highly conserved family of genes, first identified for their important developmental roles in *Drosophila*.

HOX genes. Homoeotic mutations change one paired structure to another of a more anterior or posterior compartment (for example, a leg to an antenna). The study of their structure and function has provided a paradigm for the role of genes in conveying positional information during development. In *Drosophila*, seven homoeotic genes are grouped in two complexes. Their role in establishing segmental identities is well defined, and the DNA sequence of the genes shows a highly conserved element called the homeobox. This segment codes for 60 amino acids and is highly conserved in species as distant as mammals. Importantly, this conserved sequence is also found in some pair-rule and polarity genes, and the search for genes homologous to these led to the identification of other genes that are highly conserved in animal evolution. The degree of stability of their function in evolution is reflected in the fact that the deficiencies of early development caused by mutations in a particular homeobox gene can be corrected by providing the human gene so that its product can substitute for the *Drosophila* gene.

Although *Drosophila* uses one set of homeobox genes (separated into two clusters on two different chromosomes), mammals have amplified the set of genes to a minimum of four clusters, each the size of the single cluster in *Drosophila*. These genes maintain the same patterns of expression in both mammals and *Drosophila*. They are expressed 5′ to 3′ in order of transcription, and the 5′ to 3′ order in the cluster is also reflected in the posterior-to-anterior limits of expression of the gene products.

Most mutations in *Drosophila* homeobox genes are recessive, and embryonic stem-cell knockouts have disclosed that because there is sufficient redundancy in the mammalian homeobox clusters, the homozygous absence of one homeobox gene does not always result in an apparent phenotype. In humans, dominantly inherited mutations of "gene cluster" *HOX* genes have been associated with limb and other abnormalities, whereas many dominantly inherited disorders are associated with nonclustered *HOX* genes.

Other conserved genes: FOX and Wnt. The members of the *Fox* gene family now number nearly 100. They are involved in many developmental processes. Mutations in *FOX* genes frequently cause dominantly inherited birth defects in mammals, including eye abnormalities and/or lymphedema (tissue swelling caused by dysfunction of the lymphatic vessels). A gene involved in these developmental processes, *FOXC2*, is later involved in fat cell function.

Some of these developmental control genes have a role in cancer and perform signaling functions at several developmental steps. For instance, the protein encoded by the gene *Wnt11* (a contraction of *Wingless* in *Drosophila* and *int-1*, a breast cancer gene integration site) is important in cardiac development via intracellular signaling transduction molecules (protein kinase C and *Jun* amino-terminal kinase). In contrast, another Wnt protein is involved in hair-follicle formation, in collaboration with a BMP, via an intracellular signaling cascade. The gene sonic hedgehog encodes a protein that is not only involved in neural tube induction but also in airway (lung) epithelial differentiation and in a subset of small-cell lung cancers.

Growth factors. The realization that there are many more hormonelike molecules involved in biology than was previously thought has provided for an exciting advance in organismic biology. Many of these newly discovered factors are proteins and perform roles somewhat analogous to the classic hormones of endocrinology. However, many of these molecules function in a paracrine or autocrine manner; that is, they function locally (paracrine) or on the cell (autocrine) that secretes them. The known cytokines, lymphokines, and peptide growth factors number in the many tens. Some growth-factor–like molecules remain anchored in the membrane and, therefore, must function only on neighboring cells. Other peptide growth factors adhere strongly to heparin or heparin sulfate, a common component of the extracellular matrix, and thus are limited in their diffusion and act on nearby cells. In both cases, a paracrine function can occur, whereas an autocrine function is unlikely in the former case. Nerve growth factor function is best known in the late embryo, and many of the lymphokines and cytokines are known for their roles in the adult immune and hematopoietic (blood-forming) systems. *See* CYTOKINESIS.

One stimulus to the study of the role of various growth factors in mammalian embryo development came from experiences with the culture of mammalian embryos. Fertilized embryos were early cultured to the blastocyst stage and later to the limb bud stage in the laboratory. Although the success rate was very low, this degree of success was achieved with daily changes of the media, including variations in the concentration and kind of serum supplementation, the oxygen concentration, and the degree of culture agitation. The documented need for changes in sources of serum, varying from cord blood to fetal blood, and from different species, implies that differing growth factors, which would vary between the sources of serum or plasma, might be important for embryonic development. This conclusion has been amply confirmed; and insulin, insulinlike growth factors, fibroblast and platelet-derived growth factors,

transforming growth factors, and epidermal growth factors among others, have been found to have roles in early development.

Developmental studies. Studies of embryonic development in organisms such as *C. elegans, Drosophila*, mice, and the zebrafish have shed light on the role of developmental genes and have revealed that many of the genes controlling development have been conserved throughout evolution. For instance, formation of the dorsoventral axis involves members of the transforming growth factor (TGF)-beta family of proteins [known as *decapentaplegic* in flies and BMPs in mammals], which play an important role in tissue development, cell differentiation, embryonic development, and the gene encoding the enzyme metalloproteinase (*tolloid* in *Drosophila* and BMP1 in mammals), a protein-cleaving enzyme requiring a metal ion. Other examples of evolutionary conservation include the *Pax6* gene, a member of the paired box family, which initiates the very different eye formation of *Drosophila*, of mammals, and even of cephalopods which are thought to have independently evolved a vertebrate-like eye. Although most genes involved in sex determination have not been conserved, a gene has been cloned in *C. elegans* that is highly homologous to a gene involved in the *Drosophila* sex determination cascade and to a sex determination gene in mammals (whose role has yet to be fully elucidated). These findings have resulted in the emergence of a new field of investigation known as evolutionary development, or Evo-Devo.

Caenorhabditis elegans. *Caenorhabditis elegans*, a small transparent nematode with rapid development, is highly determined, that is, a cell and its descendants are committed to particular differentiated states. In fact, each of the larva's 558 cells has an exact fate which has been carefully mapped. This fate map has been confirmed by removal of single cells, resulting in deletion of the expected descendant cells in the larva. Although the development of *C. elegans* is highly determined, there are a few examples of differentiative events that involve cell interactions, that is, where the phenotype requires cell-to-cell communication and is not a predetermined fate laid out in the egg. *See* CELL CONSTANCY; FATE MAPS (EMBRYOLOGY).

Drosophila. The complete sequence of each *Drosophila* chromosome has been obtained. Thus, almost any *Drosophila* gene that can be identified by a mutation affecting development subsequently can be cloned. In addition, the generation of new mutations by transposons (large discrete segments of DNA that are capable of moving from one chromosomal site to another in the same organism or in a different organism) generates an alternative strategy for cloning the gene. In other words, a probe for the transposon can be used to "walk" to the DNA whose expression has been altered by the transposon insertion. *See* MUTATION; TRANSPOSONS.

A powerful tool for studying *Drosophila* development involves somatic recombination techniques to create patches of homozygous mutations in an otherwise homozygous animal. These mosaics enable the study of the effects of homozygosity on a gene at a period later in development than would be possible if the mutation were homozygous lethal early in development. These studies help to readily determine if the gene effect is cell autonomous or can be corrected by diffusible factors from nearby cells; they also show that wings, legs, antennae, and eyes of the adult fly do not derive from single cells but are polyclonal in origin. In mammals, mosaicism can be studied by using embryo/embryo or embryo/embryonic stem cell chimeras. *See* CHIMERA; MOSAICISM.

Mammals. In mammals, development to the blastocyst stage occurs before implantation. The trophectoderm, primitive entoderm, and inner-cell mass have already appeared as separate lineages. In this zygote of about 100 cells, the inner-cell mass is nearly totipotent; in the mouse, as few as three inner-cell-mass cells may give rise to the adult organism. An important point about mammalian development is that it is highly regulative. In other words, the embryo has the ability to adjust its size. Bisected embryos or multiple embryos fused together lead to normal-sized fetuses. This has been confirmed for humans, where in vitro fertilization resulted in a three-eighths embryo which successfully implanted and led to a normal-sized newborn.

The long generation times of most mammals prevent them from being easily manipulated organisms for genetics. Mice, with 3 months' generation time, are the experimental organisms of choice. Humans provide much ancillary data, since physicians can screen for many natural mutations, and the large population guarantees the occurrence of many interesting natural genetic crosses. It is possible to create null mutations [changes in a gene that blocks its transcription into ribonucleic acid (RNA) and/or translation into a functional protein product] for almost any cloned gene by knocking it out (inactivating the gene) in embryonic stem cells. Alternative techniques for disrupting gene expression include the use of antisense techniques, oligonucleotides, or other stretches of nucleic acid that copy the antisense DNA strand, as well as the use of short interfering RNA (siRNA), a short sequence of RNA that finds its complementary RNA and inactivates its messenger RNA by a variety of mechanisms. *See* OLIGONUCLEOTIDE.

Zebrafish. *Danio rerio*, the zebrafish, provides a clear embryo in which the events of its early development (which mostly occur in 24 h) can be easily observed. A single female can potentially provide thousands of offspring, and homozygosity can easily be induced, which is important for studying homozygous recessive developmental genes. The generation time is not as short as that of mice, but the cost of maintaining large numbers of zebrafish is much more reasonable than that for large numbers of mice. Because of these attractive features, saturation mutagenesis for developmental pathways (that is, induction and recovery of large numbers of developmental mutations to identify all the genes affecting this function) has been undertaken in the zebrafish,

as it had previously been done in *Drosophila*. Analysis of newly found mutations has already allowed confirmation of the involvement of TGF-beta family members related to *decapentaplegic* in the dorsal ventral patterning of the zebrafish. Zebrafish have also allowed extensive genetic analysis of heart development, frequently confirming observations made in the chicken where the developing heart is more accessible. Robert P. Erickson

Bibliography. C. J. Epstein, R. P. Erickson, and B. Wynshaw (eds.), *Inborn Errors of Development: The Molecular Basis of Clinical Disorders of Morphogenesis*, Oxford University Press, New York, 2004; L. Wolpert et al., *Principles of Development*, Oxford University Press, 1998.

Developmental psychology

The study of age-related changes in behavior from birth to death. Developmental psychologists attempt to determine the causes of such changes. Most research has concentrated on the development of children, but there is increasing interest in the elderly, and to a lesser extent in other age groups. Also, most research has focused on changes in behavior at certain ages, but there has also been work on individual differences that are constant across all ages. For example, 4-year-olds resemble each other and differ from 7-year-olds in their understanding of the world, but some individuals are more intelligent than their peers at all ages. Similarly, 1-year-olds differ from 3-year-olds in their response to strangers, but some individuals are more shy of strangers than others regardless of age. Finally, although most developmental work examines humans, there has been some work on primates and other species that would be considered unethical with human beings. Thus the sensory deprivation of kittens and the separation of monkeys from their mothers have provided information about abnormal perceptual and emotional development, respectively.

Method. Developmental psychologists who study children rely more upon careful observation in natural settings than upon laboratory experiments for several reasons. First, children react strongly to unfamiliar situations; the language capacity of young children can be judged more accurately at home with their parents than in a laboratory with strangers. Second, there are ethical reasons: it is not permissible to subject newborns to pain for research purposes, but circumcisions may be observed. Third, common sense guides the psychologists: it is not possible to require parents to train their children primarily by physical punishment or by reasoning, but families with different habits can be observed. Finally, age is the primary variable studied by developmentalists, and aging cannot be administered, but only observed.

Under these circumstances, only partial conclusions can be drawn about the causes of development. The field has been dominated by descriptive research, with increasing attempts to explain developmental phenomena by the use of animal experiments or by statistical methods.

In longitudinal research, a group of individuals is studied at regular intervals over a relatively long period of time. This contrasts with cross-sectional research, where individuals of different ages are studied at the same time. Conclusions from the two types of research may differ. For example, if intelligence is studied in different groups of people at different ages, intellectual declines may be noted as early as 20 years of age, whereas lifetime studies of a group show no decline until the age of 50 and certain functions may even improve up to that time. This is due to the fact that generations differ in average education: younger people are generally better educated than older ones. These are called cohort effects. Longitudinal research eliminates cohort effects masquerading as age-related changes. Moreover, it permits an examination of the effect of early experiences on later development in a quasicausal fashion.

Finally, case studies, that is, close and extensive observations of a few subjects, have been relied upon by important developmental theorists such as S. Freud and J. Piaget.

Stages. Most developmental psychologists believe that development occurs in stages, with dramatic differences between ages. This matter, however, is debatable. Perhaps emphasis should be placed not on stages but upon the sequences through which individuals pass at various rates. Perhaps development is so gradual that neither stages nor sequences are accurate terms; instead individuals become more and more skilled with age to a point where decline begins.

Nature–nurture issue. An explanation of developmental changes requires a judgment as to the relative importance of genetically programmed maturation and environmental influences. Although most developmentalists believe that genetic endowment and environmental experience interact to account for behavior, the degree to which either affects a particular behavior is still often debated. This issue has important implications for the success of environmental intervention in the face of genetic constraints. For example, the influence on children of parental speech versus genetic programming in language acquisition is much debated, as is the origin of gender differences in behavior. *See* BEHAVIOR GENETICS.

Theories. Developmental psychology is divided roughly between those who study personal–social (emotional) development and those who study intellectual and linguistic development, although there is a small but growing interest in the overlap between these two aspects of personality, known as social cognition. The study of personal–social development in childhood is dominated by the theory of attachment formulated by J. Bowlby and extended by M. Ainsworth. In adolescence and adulthood, E. Erikson's theory of psychosocial development is prominent. The study of intellectual development at all ages is dominated by Piaget's theory of cognitive constructivism.

Bowlby's attachment theory. Ainsworth defines attachment as "an affectional tie that one person forms to another specific person, binding them together in space, and enduring over time . . . [It] is discriminating and specific." It is not present at birth, but is developed. In a word, attachment means love. Attachment behaviors such as crying, smiling, physical contact, and vocalizing are the means by which attachment is forged but are not to be equated with the more abstract, underlying construct of attachment. Attachment theory is strongly based on ethological notions. Thus, attachment is seen as serving a biological function, that is, the protection of infants by ensuring their proximity to (attached) adults. The common goal of attached individuals is proximity. Bowlby was influenced by Freud's psychoanalytic theory of development, but argues that there is a primary biological need to become attached to at least one adult, whereas Freud argued that love for a mother was secondary to her satisfaction of an infant's hunger.

Most of the work on attachment theory has explored the progressive attachment of infants to mothers, but attachment to fathers has been examined as well. Children are capable of becoming attached to other individuals, but theorists emphasize that children cared for by too many people, and no one in particular, may become attached to no one. It is also thought that the first 2 or 3 years of life are a sensitive period for the normal development of attachment. Children who do not become attached to anyone during these years are likely to experience emotional problems for the rest of their lives.

Four major phases in the development of the capacity for attachment are distinguished in early childhood (**Table 1**).

Development of specific behaviors leads to the larger changes; for example, smiling seems to be biologically programmed, since even blind infants smile. Infants smile at stimuli most likely to come from mothers, such as female voices and smiling faces. From about 3 months, infants smile more to mothers than to female strangers.

Crying is of particular interest. In the first half of the twentieth century, social learning theorists, such as J. Watson, viewed crying as a bad habit and recommended that parents ignore it. Attachment theorists, however, view crying as a signal, the infant's only means of calling the mother closer or drawing attention to a need for food. They recommend a prompt response to crying for the development of secure attachment. In one of the most important studies in the field, this dispute was resolved in favor of attachment theory: mother–infant interactions were observed over the first year of life, and it was found that the more a mother ignored her child and the longer she took to respond, the more and the longer the child cried later on.

A parent's response to a child's signals positively affects the quality of the attachment relationship between child and parent, although children become attached in some fashion even to abusive parents if there has been ample interaction. Ainsworth developed a procedure for evaluating the attachment between young children and their mothers. She observed children in an unfamiliar room responding to the presence or absence of mothers and of strangers. One group of children were clearly securely attached to their mothers, using them as a base from which to explore strange toys and people, and checking back occasionally; in their mother's absence these children protested, often trying to follow her out. Such children greeted returning mothers and sought contact with them. Other children, classified as insecurely attached, avoided contact with mothers, even turning their back upon reunion. A third group of children seemed insecure in a different way, clinging to mothers when present but refusing to be comforted upon reunion, sometimes even expressing anger. This procedure has been used in hundreds of studies, sometimes including fathers as well as mothers. Some children are securely attached to both

TABLE 1. Bowlby's theory of the development of attachment

Typical age range	Phases	Emotional characteristics, new developments
0–3 months	Undiscriminating social responsiveness	The child watches, listens to human faces; sucks and grasps to maintain contact; smiles and vocalizes more to humans than to nonhuman things; cries to attract adults and stops crying upon response.
4–6 months	Discriminating social responsiveness	The child smiles and vocalizes more to attachment figures than other humans; stops crying more readily when attended by attachment figures.
7 months–3 years	Active initiative in seeking proximity and contact	When able to crawl and walk, the child seeks out attachment figure; greets, talks to, follows, approaches, clings to attachment figure; protests figure's departure; may begin to fear strangers.
3 years on	Goal-corrected partnership	Language and cognitive development make it possible for the child to understand attachment figure's plans and to coordinate needs for proximity and contact with attachment figure's other goals.

parents, others only to one parent or to neither.

Children react strongly to prolonged separations from those to whom they have become attached. Until the age of about 3, children seem unable to understand that a parent may have good reasons for a lengthy absence. At first, they protest the parent's departure by crying, searching for them, and rejecting offers of comfort. If the separation continues, the child enters a phase of despair, crying quietly, withdrawing, and acting listless. Finally, the child recovers, seeming even cheerful, and accepting comfort from others. However, it is obvious that all is not well when the parent returns and is ignored by the child in a phase of detachment. After prolonged separations, reestablishment of trust may be slow.

Erikson's theory of psychosocial development. Like Bowlby, Erikson was a revisionist follower of Freud. Erikson deemphasized the infantile sexuality central to Freud's theory and defined psychosocial development in terms of eight crises (**Table 2**). A crisis is a conflict between two antithetical feelings. Erikson's theory agrees with Bowlby's in the sense that for him the first psychosocial crisis is the establishment of trust as against mistrust. Each age is dominated by particular psychosocial crises, and successes or failures in meeting them have lifelong consequences.

Family studies. Ever since Freud, it has been believed that early experiences, especially with the mother, are crucial in emotional development. Children have been increasingly studied in family settings, with examination of nontraditional family arrangements. Studies of the effects of maternal employment have afforded no simple conclusions, although it seems clear that day care away from the home need not cause any cognitive or emotional disturbance in children. In cases of divorce, a year of emotional crisis often follows, but subsequent disturbances are largely due to continuing parental conflict over the children.

Piaget's theory of cognitive development. Piaget was not interested in individual differences in intelligence, but rather in formulating a universal scheme of intellectual development. For Piaget, intelligence is defined as the ability to adapt to the environment, an ability that depends upon physical and psychological (cognitive) organization. Cognitive structures are termed schemata (or schemes). The adaptation process has two complementary components, assimilation and accommodation. Assimilation refers to the tendency to process new information, sometimes with distortion, in terms of existing cognitive structures. Accommodation refers to the opposite process, that is, the modification of existing cognitive structures in response to new information. In other words, whereas assimilation refers to the limits of learning, accommodation is Piaget's view of learning itself: learning occurs directly, from manipulation of materials, without external instruction or reinforcement. An infant who sucks a new rattle is assimilating it to the schema of sucking while simultaneously accommodating the sucking to the rattle: the infant has probably never sucked anything of that shape before so that the sucking must change. An individual strives for equilibrium between assimilation and accommodation, with thought being neither unrealistic (excessive assimilation) nor excessively realistic and hence disorganized (excessive accommodation). *See* COGNITION; INTELLIGENCE.

For Piaget (**Table 3**), cognition gradually becomes abstracted from perception over the course of 12 years. Cognition during at least the first 18 months is directly observable. Infants begin cognitive exploration by actively perceiving and reflexively manipulating objects, giving the name sensorimotor period to the first phase of intellectual development. Perception is a key form of early cognitive activity, especially with newborns. The newborn infant can see, hear, smell, taste, and feel much better than previously thought, though sensitivity in these areas improves throughout the first year of life.

As far as manipulation goes, Piaget termed the

TABLE 2. Erikson's theory of psychosocial development

Typical age range	Dominant crises	Outcomes of favorable ratio of positive to negative feelings
0–1 year	Basic trust vs. basic mistrust	Drive and hope; confidence that adults will meet basic needs despite temporary frustrations.
1–2 years	Autonomy vs. shame and doubt	Self-control and willpower; balanced restraint and expression despite shame and self-doubt; displays of defiance of parents.
3–5 years	Initiative vs. guilt	Direction and purpose; undertaking, planning, and executing goals despite inner guilt about some goals contemplated.
6 years–puberty	Industry vs. inferiority	Method and competence; systematic instruction produces intellectually and socially competent individual despite fears of inadequacy.
Adolescence	Identity vs. role confusion	Devotion and fidelity; a sense of self despite questions about sexual or occupational identity.
Young adulthood	Intimacy vs. isolation	Affiliation and love; commitment to partner and career despite concerns about loss of identity in this process.
Middle adulthood	Generativity vs. stagnation	Production and care; concern to establish and guide next generation at home and at work despite more self-centered concerns.
Older adulthood	Integrity vs. despair	Renunciation and wisdom; acceptance of life and death and contented reflection on a life well lived, which overcomes fear of death.

TABLE 3. Piaget's theory of cognitive development

Period and typical age range	Stages	Cognitive characteristics
Sensorimotor period (0–24 months)		
0–1 month	Stage 1: use of reflexes	Unlearned responses (reflexes) become smoother.
1–4 months	Stage 2: primary circular reactions	Repetitions of accidentally discovered sensorimotor responses that interest child; interest in sucking for sake of sucking.
4–8 months	Stage 3: secondary circular reactions	An interest in effects of actions upon objects; plays with toys.
8–12 months	Stage 4: object permanence	An ability to remember that things exist even when child does not see them; searches for things seen to disappear; appearance of genuine intentionality.
12–18 months	Stage 5: tertiary circular reactions	Experiments on objects and an interest in novelty; searches for objects seen to disappear no matter where they were found previously.
18–24 months	Stage 6: symbolic representation	Can solve problems by thinking about them rather than simply by trial and error; full object permanence in the sense that child will search for objects even when they were not seen to disappear; first abstract mental representations of objects.
Preoperational period (2–7 years)		
2–5 years	Stage 1	Deficient causal reasoning; concentration on only one dimension of reality at a time; more or less random arrangement of things in serial order; no one-to-one correspondence in (rote) counting of objects; inability to take others' perspective (egocentrism).
5–7 years	Stage 2 (transitional)	Physical causal reasoning present but primitive; vacillation about correct responses on conservation tasks; correct but laborious arrangement of things in serial order; one-to-one correspondence in counting, but length of displayed objects still misleads child; some successful role-taking but egocentric errors still occur; instruction in conservation, counting, and so on, is fruitful at this stage.
Concrete operational period (7–11 years)		Child reasons correctly; can conserve quantity, then weight, then volume, among other things because of ability to focus on several dimensions of a problem simultaneously; arranges things in serial order according to an efficient overall plan; full mastery of counting with no deception about physical distance between things counted; egocentrism has waned.
Formal operational period (11 years on)		Awareness of concrete reality as a subset of possibility; reasoning becomes more flexible and systematic; thinking is logical and in terms of abstract propositions systematically combined; mental hypotheses tested.

young child's inclination to repeat interesting activities "circular reactions." During the sensorimotor period, the child is at first interested only in its own repetitious activity (primary circular reactions), although this is an advance over simply performing reflex actions. Then the child becomes interested in actions because of the effect they have had in the past upon objects (secondary circular reactions). Finally, the child becomes interested in novel outcomes of actions on objects (tertiary circular reactions). Between the ages of 18 and 24 months, infants become capable of symbolic representation, occasionally solving problems just by thinking about them.

The major accomplishment of the sensorimotor period is object permanence, the realization that objects continue to exist even when not observable. At 8 months, the infant first demonstrates some such capacity, but it is not until the end of this period that an infant can find an object that was not observed to disappear.

During the next 5 years, sometimes termed the preoperational period, children work on concrete operations such as classifying objects into categories (animals, furniture), arranging things in serial order, figuring out causes and effects, or understanding a one-to-one correspondence of numbers to objects counted. They also eventually manipulate reality enough to overcome perceptual illusions such as that

an amount of water changes when it is poured from a short wide glass into a tall narrow glass. In other words, they grasp the notion of conservation of quantity. From 7 to 11 years, children further consolidate their concrete mental operations, grasping conservation of volume, for example, only at the age of 11. The phenomenon that conservation of different aspects of reality is understood at different ages is called horizontal decalage.

At about 12 years, many adolescents enter the final stage of intellectual development: formal operations. They become capable of abstract, logical thought. They understand reality as a subset of possible worlds, and are able to form multiple, systematic hypotheses, involving all possible combinations of relevant variables, in order to explain things.

An important point for Piaget is that children actively construct reality for themselves; they need not, indeed cannot, be instructed in many basic developments. Another of his concerns was the formulation of a universal scheme of intellectual development. Many quarrel with his age assessments of children. He underestimated the capabilities of young children because he used tasks and materials unfamiliar at the age tested. Others argue that his stages are not as coherent in real children as his theory suggests. But most people accept his sequence of stages as useful for classifying children.

The most important criticism of Piaget is directed at his concept of egocentrism. He argued that infants are born incapable of distinguishing themselves from reality. Even preoperational children (2 to 7 years) were said to be so egocentric as to be incapable of realizing that others view the world differently from themselves. Egocentrism was a cornerstone of Piaget's theory, the cause of all other deficiencies of thought in the young child. As evidence of egocentrism, Piaget pointed to the following: children often do not specify pronouns properly, misuse causal language, and talk in front of others without successful communication. However, many investigators have found that children as young as 3 can grasp the perspective of others. Children may appear egocentric, but this is a result, rather than a cause, of a lack of language skill, for example, they use a pronoun when a noun is unknown. Furthermore, it has been argued that speaking out loud without regard for communicating with anyone often serves the function of self-regulation ("I want to draw something. First I need a crayon and a piece of paper."). Children plan their actions by thinking out loud, a finding that has been supported by much subsequent research. In short, language is instrumental in developing thought. Vygotsky's realization that intrapersonal processes (such as thought) are transformed interpersonal relations (such as conversation) has reestablished a meaningful role for the instruction of children, which was overemphasized by behaviorists and subsequently underemphasized by Piaget.

Language development. There is less common ground than there perhaps should be between students of developing thought and those of developing language. The study of developing language began in a serious way when N. Chomsky rejected Skinner's behavioristic notions that language was learned in the same way that rats learn to press a bar. Chomsky argued that children are born with an innate ability to acquire language. His theory of grammar paved the way for descriptions of how language structures develop. Vocalizations at birth are limited to crying, but are followed by cooing within a few months. At 6 months, most babies begin babbling, uttering their first word at about 1 year. They speak one word at a time until the latter half of the second year, when vocabulary explodes and words are combined two at a time. These early combinations are components of full sentences. Over the next few years, children increase the length, coherence, and complexity of their utterances, telling a good personal narrative by the age of 6. In the early school years, children learn to read, decoding words and simple stories and becoming more expert. Finally, from later elementary school on, children develop writing ability. Although these are the central language accomplishments at various ages, the development of each component starts earlier and ends later than the age range specified. Thus, preschool children engage in pretend reading and invent spellings for words, and elementary school children continue to correct their pronunciations.

Children acquire language in the context of language games played with parents, other adults, and children. Adults are more repetitious and use simpler sentences when talking with children than with adults, so that children pay more attention. *See* PSYCHOLINGUISTICS.

Memory development. Language and thought are fairly well developed by age 12. Once people have formal operations and are capable of writing coherently, they continue to operate at this level throughout life. But the issue of memory deterioration with age has considerable importance. There are many types of memory, but the most important developmental change in old age recalls the most important developmental change in childhood, that is, the use of memory strategies. As children progress through elementary school, they repeat to themselves things they wish to remember and organize information they wish to learn, perhaps forming an image to link such items as a state and its capital; such strategies improve memory. In old age, the reverse seems to happen: memory strategies are used less and memory suffers. Instruction in memory strategies helps children and the elderly. *See* MEMORY.

Kohlberg's theory of moral development. L. Kohlberg's work on moral development spans the chasm between intellectual and emotional development. He studied reasoning about hypothetical moral dilemmas, such as whether a person should steal an unaffordable drug in order to save someone's life. He classified such reasoning in six stages, two at each of three levels of moral development. At birth children are considered to be premoral. By the age of 7, most children are in stage 1, chiefly characterized by the belief that people should act in certain ways in order to avoid physical or other punishment. In 2 or 3 years, children reason primarily in terms of doing things for rewards; this is stage 2. These first stages

are termed preconventional morality; rules are not yet internalized.

At the conventional stage of moral reasoning, rules are internalized. Stage 3 of this period is often termed good-boy morality, and involves reasoning focused less on rewards than on maintaining the approval of others. Stage 4 involves reasoning that unquestioningly accepts conventional rules. Actions are judged by a rigid set of regulations, religious, legal, or both. Most individuals do not develop past this point.

A few, however, do reach postconventional moral reasoning, stage 5. These individuals think in terms of moral principles. Law is viewed as a social contract, and if it violates the rights of even a few, it is considered unfair and in need of change. Rarely, a step higher to stage 6 is reached, governed by original abstract moral principles such as articulation of the golden rule.

Moral development occurs through exposure to advanced thinking. Kohlberg argued that moral development is progressive, without regression to earlier stages. Although he interviewed only boys, he considered his scheme as universal and received some support. It has, however, been argued that women are too often assigned to stage 3 for life because of an overriding concern with seeking approval and not offending others. According to this analysis, female moral development is oriented toward an equally compelling ethic of care that is misunderstood and underestimated in male standards of justice. However, in most studies using Kohlberg's system, there are no significant differences between large numbers of women and men.

Gender differences and sex roles. Only a few differences in behavior have consistently been found between boys and girls. These are small and apply only to the "average" boy and girl. Individual boys differ from each other much more than from girls. On the average, boys are more aggressive than girls and slightly better at mathematics and spatial relationships. Girls, however, are more verbal than boys.

Most of the differences between boys and girls lie in the sex roles they construct for themselves. Before they are 3, children know their own sex, but it takes another year or two to realize that sex is fixed for life. Before the age of 5, boys may not realize that their sex does not change even if they played with dolls. Preschool children prefer toys traditional for their sex. The causes of the observed differences have been strongly debated.

Other differences. Some cultures apparently express formal operations in a way very different from that proposed by Piaget, or not at all. Generational differences extend to education. Individual differences, as in temperament, are relatively stable over a lifetime. It is thus debatable whether any universal scheme of development is possible.

Developmental psychopathology. Some individual differences are acceptable (for example, high versus low activity), but others are not. Traditionally, child clinical psychology (abnormal development) and the study of normal development were separate. However, effort is being made to integrate them. Normal developmental milestones define abnormal development; abnormal development is informative about normal processes. The serious disorders of childhood include autism, attention-deficit disorder with hyperactivity, and depression. Viewed another way, abnormal children are either overcontrolled (obsessive-compulsive) or undercontrolled (impulsive, aggressive). Developmental psychopathologists, however, are interested not just in disordered development in childhood, but in abnormal individuals over their lifetime. Such studies can shed light on the effectiveness of treatments and on the way in which disorders such as hyperactivity may be displayed differently at different ages.

Applied developmental psychology. The best work in developmental psychology has always had practical implications. Nevertheless, there is increasing emphasis on making research findings available to the public and on taking an active role in developing policy. Further applied research can benefit many groups.

Allyssa McCabe

Bibliography. J. Bowlby, *Attachment and Loss*, vols. 1–3, 1980; E. Erikson, *Childhood and Society*, 2d ed., 1963; H. Ginsburg and S. Opper, *Piaget's Theory of Intellectual Development*, 2d ed., 1979; J. Gleason (ed.), *The Development of Language*, 4th ed., 1999; J. H. Markman and E. M. Flavell, *Handbook of Child Psychology*, 4th ed., 1983; L. Vygotsky, *Thought and Language*, 2d ed., 1988.

Devonian

The fourth period of the Paleozoic Era, encompassing an interval of geologic time between 418 and 362 million years before present based on radiometric data. The Devonian System encompasses all sedimentary rocks deposited, and all igneous and metamorphic rocks formed, during the Devonian Period. It is conventional that recognition of Devonian time

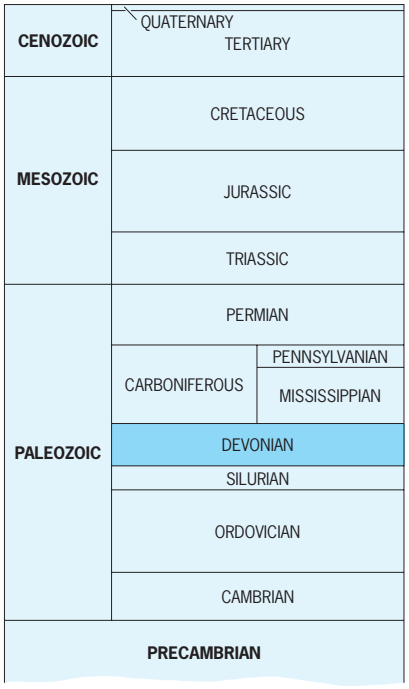

is determined by the definition of Devonian rocks. *See* PALEOZOIC.

The base of the Devonian System has been fixed, by international agreement, at an actual outcrop of sedimentary rocks at Klonk in the Czech Republic, where it corresponds to the base of the *Monograptus uniformis* graptolite zone. The top of the Devonian System, corresponding to the base of the Carboniferous System, was similarly fixed at LaSerre in southern France, recognized by the base of the *Siphonodella sulcata* conodont zone.

In 1839 Adam Sedgwick and R. I. Murchison proposed the Devonian System to encompass the marine sedimentary sequence between the Silurian and Carboniferous in the counties of Cornwall and Devon in southwestern England. In early-studied locales such as Wales and Scotland, only terrestrial Old Red Sandstone lies above the Silurian. When fossil corals found in marine rocks in Devonshire were considered by William Lonsdale to be intermediate in character between those of the better-known Silurian and Carboniferous marine deposits, these Devonshire rocks and the marine fossils they included were determined to be in the stratigraphic position of the Old Red Sandstone. This stage-of-evolution judgment was made in 1837, before Darwin's theory of natural selection had appeared. Because the rocks and faunas of the structurally complex type region were poorly known, Murchison and Sedgwick traveled to Germany in 1839 to see if their new system could be identified in the richly fossiliferous sequences below the Carboniferous in the Rhine Valley. Subsequently the German and Belgian Devonian sequences have served as worldwide standards of comparison.

About the same time, work in America by the paleontologists James Hall, John Clarke, and others was making known the superb physical and faunal development of Devonian rocks in New York, where there are few structural complications. If the Devonian had not been defined until 1842, it might now be called "Erian" after the New York deposits, which were described in much detail in the early reports of the New York State Museum of Natural History.

In modern times, it has become evident that even the Rhenish and New York sequences represent the Devonian System less than adequately. These sequences have been supplemented by the complete Lower Devonian sequence in the Barrandian region of the Czech Republic, and by other Devonian sequences in Iowa, Nevada, Australia, Canada, China, Morocco, and Siberia.

Subdivisions. The Devonian is customarily divided into Lower, Middle, and Upper series and their corresponding epochs. These, in turn, have been divided into stages and their corresponding ages (**Fig. 1**). The base of the Middle Devonian is at the base of the German Eifelian, or base of the *partitus* conodont zone. The base of the Upper Devonian is near the base of the Belgian Frasnian, or base of the former lower *asymmetricus* conodont zone (MN Zone 1), defined by the first occurrence of *Ancyrodella rotundiloba*. *See* CONODONT.

Determination of time equivalency (correlation) and age dating in the Devonian are usually accomplished biostratigraphically, by utilizing zones based on the evolutionary successions in individual fossil groups, or on a composite zonal framework based on several fossil groups.

A standard biochronology for the Devonian was first developed by utilizing ammonoid cephalopods, and this was most successful in the Upper Devonian. During the 1960s and 1970s, a standard biochronology was developed for most of the Lower Devonian, based on graptolites. Many local correlations within the Devonian have depended upon the faunal succession of other animal groups (for example, brachiopods, corals, ostracods, trilobites, dacryoconarid tentaculites, fishes) or on plant fossils. A microfossil zonal biochronology has been developed for conodonts to a point where it is superior to any other for the Devonian, in terms of both precision and wide applicability (Fig. 1).

Lithofacies and paleogeography. Because certain sedimentary rock types form under limited ranges of climatic conditions, mapping the nature and distribution of these sedimentary rock types provides a view of ancient climatic distribution. Climate is controlled by intensity of solar insolation and its effect on atmospheric circulation, which is related directly to latitude and direction of Earth's rotation as modified by position of landmasses. The abundance of tropical-climate rocks in present northern temperate continents, combined with lack of these tropical-climate rocks in large portions of southern tropical continents during much of the Paleozoic, has constituted one of the lines of evidence supporting the concept of continental drift. This is particularly obvious in the Devonian, and a recent reconstruction of continental positions has accommodated these climatic inconsistencies in an actualistic view of the Devonian world (**Fig. 2**), which also explains the heretofore enigmatic faunal distribution patterns. *See* PALEOGEOGRAPHY.

North America. Devonian rocks across most of the interior of North America, from Hudson Bay to the Ohio Valley across to Nevada and north to the Mackenzie Basin, consist of a predominantly marine carbonate suite, including both limestone and dolomite, commonly with abundant fossils, but containing relatively minor amounts of sandstone and shale. Interbedded with the carbonates in several areas are evaporites ranging from least soluble sulfates, the most common and widespread, through halite, common and widespread in western Canada and present in the Hudson Bay area and Michigan, to the most soluble potash salts, which are rare among evaporites of all ages, but which form a thick local sequence that is mined in the Canadian province of Saskatchewan. This carbonate-evaporite suite indicates a dry tropical climate for central North America in the Devonian. Shale- and sandstone-rich sequences characterize two belts in the North American Devonian. One extends along the Pacific coast from California through the Yukon to the Canadian Arctic islands, and probably represents the equatorial rainfall belt. The other extends along the Appalachians, and probably represents the southern warm temperate rainfall belt. This belt in

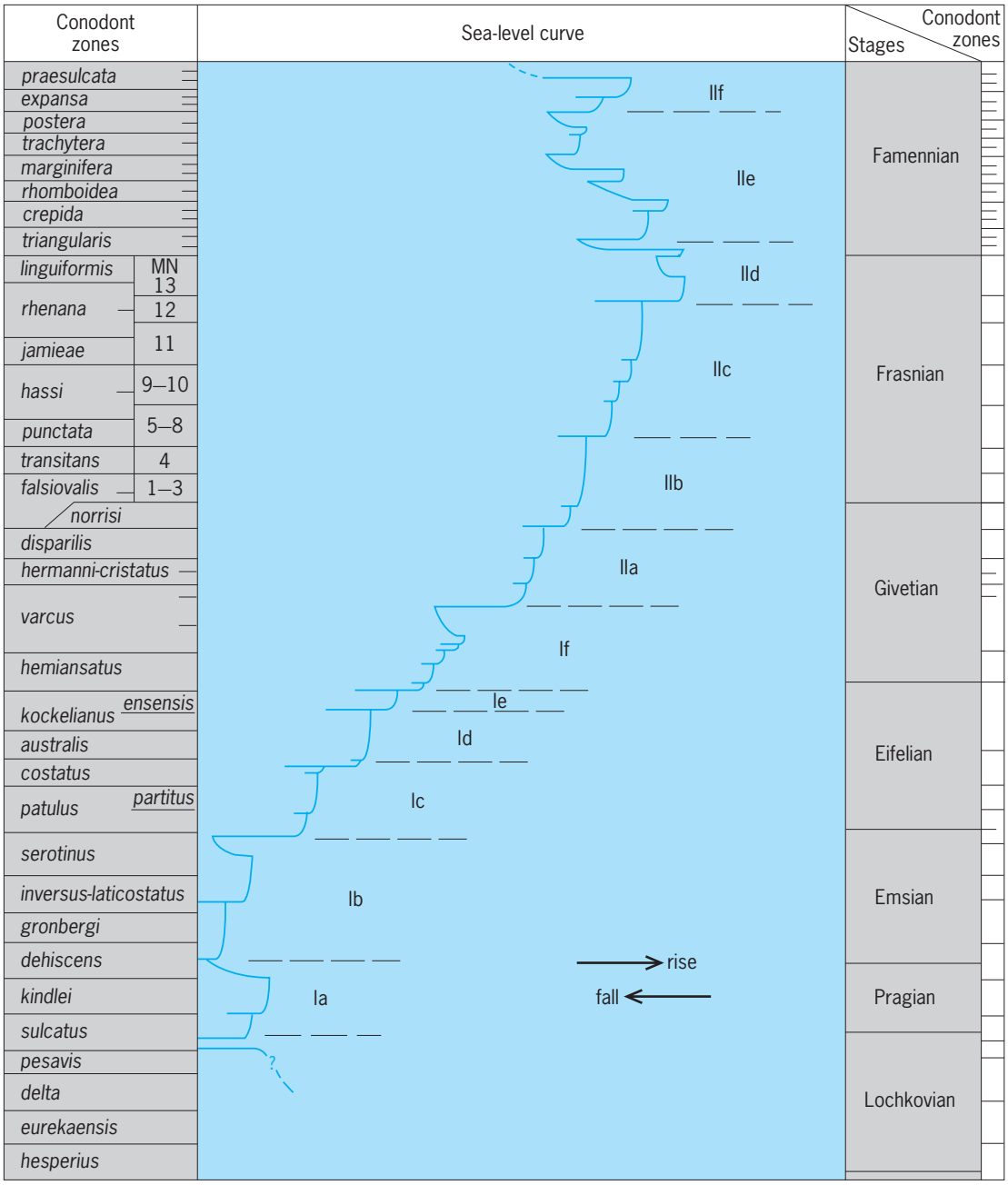

Fig. 1. Biostratigraphic zones based on conodonts (left) plotted to show approximate correspondence to the chronostratigraphic stages and series which compose the Devonian System (right) and relative sea-level changes (center) showing major fluctuations during two depositional phases (roman numerals).

New York and Pennsylvania includes widespread black shales and the famous Catskill deltaic complex (clastic wedge) that figured strongly in the early history of Devonian investigation in North America.

Europe. During the Devonian, Europe was apparently joined to North America as a result of the Silurian Caledonian orogeny to form a larger continent known variously as Euramerica, Laurussia, or the Old Red Continent. The last name derives from the Old Red Sandstone, among the first Devonian formations described, which is the coarse nonmarine detrital deposit occurring throughout most of Britain, Scandinavia, Spitsbergen, and eastern Greenland and giving evidence of the Caledonian mountains along the west. A dry tropical carbonate-evaporite sequence covers much of Russia, with equatorial bauxite-bearing thick carbonate and detrital sequences along the Ural mountain chain, the site of a seaway during the Devonian. Sequences of detrital rocks and tropical reefy carbonates characterize central Europe, particularly Germany and Belgium, with possibly deeper-water deposits in parts of southern Europe.

Asia. The largest modern continent apparently consisted of as many as 11 microcontinents during the Devonian. The largest of these was Siberia, which has dry tropical carbonate-evaporite sequences and redbeds across much of its area and more humid detrital and reefy carbonates along certain edges. The other Asian fragments also have warm-climate carbonate and detrital sequences, with equatorial

bauxites in southern China and Kazakhstan, and evaporites in Iran and other parts of Kazakhstan.

Gondwana. During much of the Paleozoic this giant continent consisted of Africa, South America, Antarctica, Australia, India, and Arabia fitted together, with smaller fragments such as Madagascar, like pieces of a jigsaw puzzle. Over most of Gondwana, including all of South America, sub-Saharan Africa, and Antarctica, Devonian rocks consist entirely of nonred sandstone and shale sequences. These indicate a humid climate, and the absence of any warm-climate rocks such as carbonates, evaporites, or bauxites strongly suggests a cool climate. Thus Gondwana was apparently centered near the South Pole in the Devonian, with most of it in the cold temperate humid belt. Only around the outer fringes in the warm temperate belt do carbonate rocks appear, in northwestern Africa, the northern Indian subcontinent, and most of Australia. Evaporites in parts of Australia indicate that this part of Gondwana extended farthest from the pole, into the dry tropical zone.

Biogeography and ocean currents. Devonian fossils are found to be distributed in three realms. Within each realm there is taxonomic similarity, which indicates that there was reproductive interchange among members of the same phyletic groups, but between each two realms there are various degrees of taxonomic dissimilarity, which indicates that there were various degrees of reproductive isolation among members of the same phylogenetic groups.

The largest realm covered Australia, Asia, Europe, western North America, and the Morocco-India fringe of Gondwana, and is termed the Old World Realm. It was unified by relatively free flowage of the warm equatorial currents and their immediate branches among the continental masses throughout this tropical to subtropical region (Fig. 2).

The Appalachian Realm covered most of eastern North America and the Colombia-Venezuela-Amazon part of northern South American Gondwana, which was adjacent to Appalachian North America during the Devonian. This region was bathed by the temperate southern west-wind current, which crossed a sufficiently broad stretch of ocean so that many Old World larvae could not make the journey, allowing endemic Appalachian forms to develop locally.

The Malvinokaffric Realm covered central Gondwana, including southern South America, southern Africa, and Antarctica. Absence of certain carbonate-secreting groups such as stromatoporoids and green algae, and great reduction of others such as corals suggest that this was a cold-water fauna, which accords with the position of this part of Gondwana suggested by the exclusively detrital rock suite. This region was bathed by a small subpolar ocean current, which was derived from the temperate west-wind

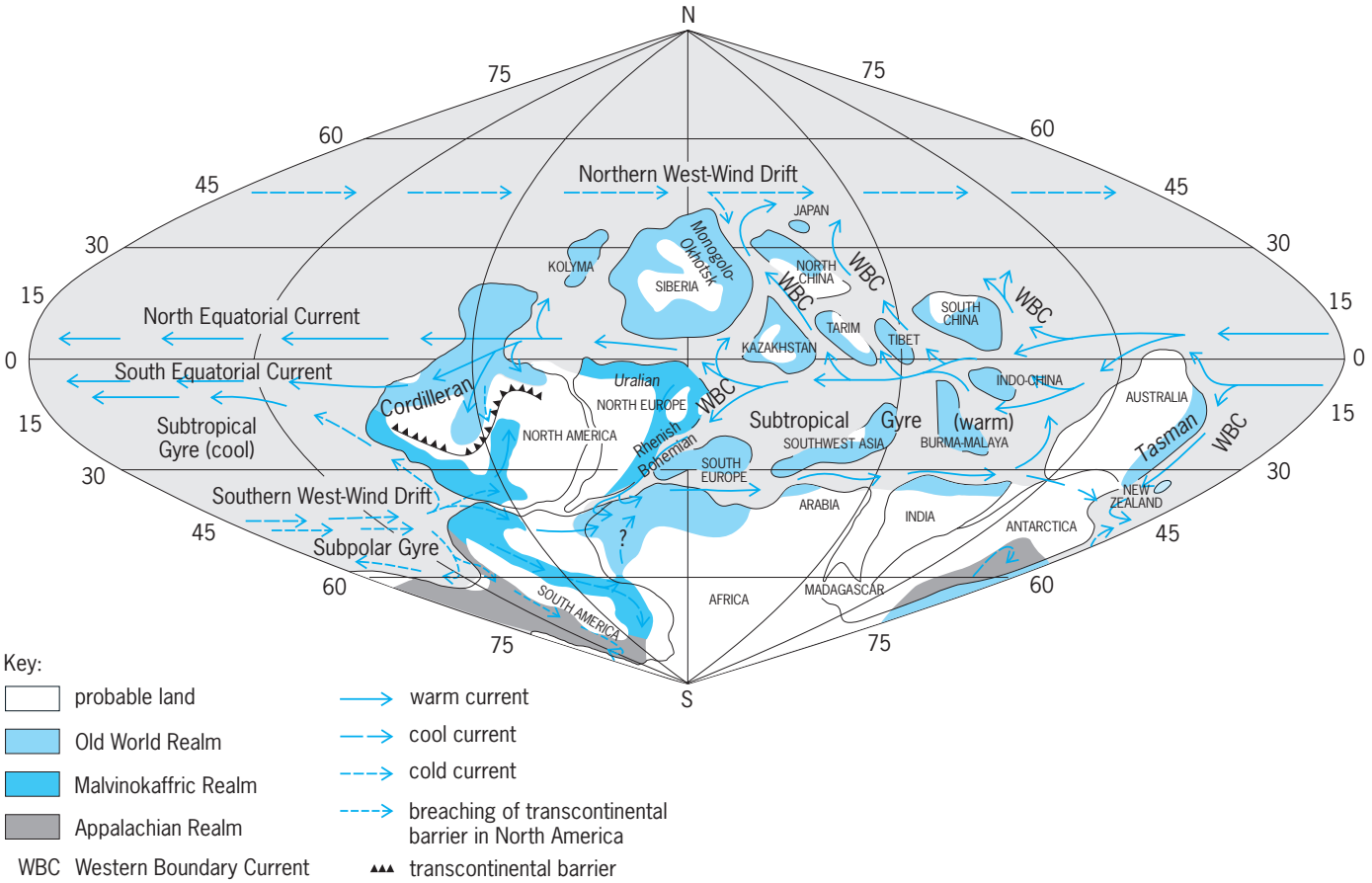

Fig. 2. Reconstruction of Middle Devonian continental positions based on distribution of lithologic climatic indicators. Probable ocean currents derived from this paleogeography explain the biogeographic realms. Kolyma and Tarim are microcontinents of the devonian. (*After M. R. Hourse, C. T. Scrutton, and M. G. Bassett, eds., The Devonian System, Palaeontol. Ass. Int. Symp. Spec. Pap. 23, pp. 99–123, 1979*)

current, but in which colder temperatures provided a barrier to Appalachian organisms.

Tectonics. Devonian mountain building was particularly noticeable along the margins of Euramerica. The Acadian orogeny formed mountainous highlands accompanied by a chain of granitic intrusions from Nova Scotia to Pennsylvania during much of Devonian time. These mountains formed a barrier that prevented mixing between organisms of the Appalachian Realm and those of the Old World Realm at the same latitude in central Europe. Erosion from the Acadian mountains produced the thick Catskill deltaic complex of New York and Pennsylvania, which spread its fine-grained sediments far into the interior of eastern North America during later Devonian time. Simultaneously, along the Arctic margin of Canada, the Ellesmerian orogeny was producing folded mountains whose erosional products formed a clastic wedge that was a mirror image of the Catskill deposits.

During latest Devonian time, the Roberts Mountains thrust, of the Antler orogeny in Nevada and Idaho, formed at the top of a subduction zone along which continental crust and overlying sediments were descending beneath oceanic sediments. A Late Devonian orogeny also affected eastern Australia. *See* OROGENY; PLATE TECTONICS.

Sea-level changes. Times of active plate movement, such as the Middle and Late Devonian, were times during which the oceanic rise systems formed very large submarine mountain chains, like the Mid-Atlantic Ridge of today. Expansion of these rise systems reduced the volumetric capacity of the ocean basins, so that marine waters rose eustatically and spread as transgressions across low-lying continental platforms, forming broad epeiric seas. This not only affected greatly the physical surface of the Devonian world, but was a major environmental factor in the evolution of plants and animals. During the earliest Devonian, the world's oceans were eustatically low, so that epeiric seas were absent from continental interiors, a continuation from the regression during the Late Silurian. Offshore marine organisms of the earliest Devonian had much in common with their predecessors.

Epeiric seas then expanded sporadically during the remainder of the Early Devonian. This transgressive episode affected Gondwana, in addition to the Euramerican platforms. The general transgressive trend greatly increased near the beginning of the Middle Devonian (transgressive part of transgressive-regressive cycle Ic, Fig. 1), and culminated, after a series of transgressive-regressive cycles, at the end of the Frasnian Age of the Late Devonian. During latest Devonian time, epeiric seas were extensive, but less evenly continuous than before orogenic movements had begun to modify the crustal surface. The interior of Gondwana did not undergo transgression after the early Middle Devonian.

As the Devonian transgressions progressed, many offshore marine animals adapted themselves to the expanding habitats of epeiric seas and then migrated widely as barriers were inundated. This led to loss of isolation, rise of competition, lowering of overall diversity, and loss of the separate realms, with replacement by a cosmopolitan Frasnian marine fauna derived from the Old World Realm. At the end of the Frasnian, still unexplained extinctions of many marine groups further reduced the organic diversity of the Devonian world.

Life. Among the marine invertebrates, trilobites (Arthropoda) were much less abundant than during the Cambrian. The planktic members of the extinct graptolites died out during the Early Devonian, at about the same time as the pelagic ammonoid cephalopods first evolved. The externally two-shelled brachiopods were at their greatest diversity, being represented by more than 900 genera. Lime-secreting corals and stromatoporoids were important and widespread in warm-water environments, and formed reefs during the Middle and Late Devonian. The extinct microfossil group known as conodonts was abundant, widespread, and rapidly evolving during the Devonian, so that conodont fossils are now regarded as the principal tools to be used for international correlation and relative age determination (Fig. 1). *See* BRACHIOPODA; CONODONT; GRAPTOLITHINA; MICROPALEONTOLOGY; STROMATOPOROIDEA; TRILOBITA.

The great diversification and radiation of fish in the Devonian has led to the term "Age of Fishes" for the period. Placoderm fish, among the most primitive of the jawed vertebrates, were successful predators in Devonian waters, and some grew to lengths up to 8 m (25 ft) just before their extinction at the end of the Devonian. The sharks, with a cartilaginous skeleton but lacking a swimbladder, may have evolved from an early placoderm. *See* PLACODERMI.

Bony fishes or Osteichthyes, a class that includes all modern fish other than sharks and agnathans, were represented in the Devonian by the primitive acanthodians, but more modern groups of bony fishes appeared in the Early Devonian. Lobe-finned bony fishes, or sarcopterygians, include both the lungfish (Dipnoi) and crossopterygians in the Devonian. The oldest known amphibians, including *Acanthostega* and *Ichthyostega*, which evolved from the rhipidistian crossopterygians, occur in strata thought to be high Upper Devonian. *See* DIPNOI; OSTEICHTHYES; SARCOPTERYGII.

Many Lower and Middle Devonian fish fossils are now known from rocks deposited in open marine environments, indicating that their habitat was marine as well as estuarine or fresh water. The vertebrates appear to have made the complete transition from ocean to dry land within the Devonian, perhaps as a result of the evolution of land plants during that period.

Land plants began to flourish near the beginning of Devonian time, and were exemplified by the vascular genus *Psilophyton* of the phylum Psilopsida. The latter gave rise in the Devonian to the Lycopsida (scale trees) and Pteropsida (true ferns). The pteropsids remain important in the world flora of today. *See* LYCOPHYTA; PALEOBOTANY; PSILOTOPHYTA; PTEROPSIDA.

J. G. Johnson; P. H. Heckel; D. J. Over

Bibliography. M. R. House, C. T. Scrutton, and M. G. Bassett (eds.), *The Devonian System*, Palaeontol.

Ass. Int. Symp. Spec. Pap. 23, 1979; J. G. Johnson, G. Klapper, and C. A. Sandberg, *Devonian Eustatic Fluctuations in Euramerica*, Geol. Soc. Amer. Bull. 96, 1985; G. R. McGhee, Jr., *The Late Devonian Mass Extinction*, Columbia Press, 1996; N. J. McMillan, A. F. Embry, and D. J. Glass (eds.), *Devonian of the World*, Canadian Soc. Petrol. Geol. Mem. 14, 3 vols., 1988; M. A. Murphy, W. B. N. Berry, and C. A. Sandberg (eds.), *Western North America: Devonian*, Univ. Calif. Riverside Campus Mus. Contrib. 4, 1977; J. B. Roen and R. C. Kepferle (eds.), *Petroleum Geology of the Devonian and Mississippian Black Shale of North America*, 1994.

Dew

The deposit of liquid water resulting from condensation of atmospheric water vapor to exposed surfaces that cool during the night. Dewfall is noticeable in the early morning after a calm, cool, clear night, usually as beads of liquid water on the outside and upward-facing surfaces of trees, buildings, and so forth. If the ground is moist, some of the condensed water can be the result of distillation of surface moisture. Dew forms when the surface temperature drops sufficiently to saturate air in contact with the surface (that is, when the surface drops to below atmospheric dew-point temperature); when the surface cools to below freezing temperature, frost occurs. *See* DEW POINT; FROST.

When water is plentiful, vegetation sometimes exudes liquid water through openings in the skin of leaves. This phenomenon, known as guttation, also contributes to the accumulation of liquid water on foliage, but is unrelated to the condensation mechanisms that cause dewfall. Guttation is often recognizable as droplets of water at the tips of leaves.

Dewfall is difficult to measure because of the small amounts of water involved, but absorbent paper is sometimes used to soak up and weigh the condensation on sample surfaces. In some studies, sensors have been attached to foliage to detect the formation of liquid water. Dewfall rates have been computed by using the surface heat energy budget: the latent heat associated with the condensation is measured, and the average rate of liquid water deposition is then calculated. The rate at which dew is deposited is limited by the surface energy budget. Nighttime cooling rates are greatest when the atmosphere is driest because radiative cooling is more rapid. Also, since clouds can limit infrared radiation loss from the surface, dewfall tends to be a clear-sky phenomenon.

Hygroscopic particles on surfaces can act as sites for condensation at temperatures higher than the atmospheric dew-point temperature. Thus, if a surface is not clean, the first deposit of moisture from the air can occur well before the surface cools to dew-point temperature. For some chemicals, such as common salt, condensation can start to occur when the local relative humidity reaches 80%; the humidity must be 100% in the case of a clean surface. Some desert plants exude hygroscopic salts from the interior of leaves which provide preferred sites for condensation and thereby create a supply of water for the plant. *See* HUMIDITY; HYDROLOGY; PRECIPITATION (METEOROLOGY); VAPOR PRESSURE. Bruce Hicks

Bibliography. B. B. Hicks, A study of dewfall in an arid region: An analysis of Wangara data, *Quart. J. Roy. Meteorol. Soc.*, 109:900–904, 1983; H. A. Mooney et al., Atmospheric water uptake by an Atacama Desert shrub, *Science*, 209:693–694, 1980; A. Zangvil and P. Druian, Measurement of dew at a desert site in southern Israel, *Geog. Res. Forum*, 2:26–34, 1980.

Dew point

The temperature at which air becomes saturated when cooled without addition of moisture or change of pressure. Any further cooling causes condensation; fog and dew are formed in this way.

Frost point is the corresponding temperature of saturation with respect to ice. At temperatures below freezing, both frost point and dew point may be defined because water is often liquid (especially in clouds) at temperatures well below freezing; at freezing (more exactly, at the triple point, +.01°C or 32.02°F) they are the same, but below freezing the frost point is higher. For example, if the dew point is −9°C (16°F), the frost point is −8°C (17.5°F). Both dew point and frost point are single-valued functions of vapor pressure.

Determination of dew point (or frost point) can be made directly by cooling a flat polished metal surface until it becomes clouded with a film of water or ice; the dew point is the temperature at which the film appears. In practice, the dew point is usually computed from simultaneous readings of wet- and dry-bulb thermometers. *See* DEW; EVAPORATION; FOG; HUMIDITY; HYGROMETER; PSYCHROMETER; VAPOR PRESSURE. J. R. Fulks

Dewar flask

A vessel having double walls, the space between being evacuated and the surfaces facing the vacuum being heat reflective. It was invented in 1892 by James Dewar as a container for liquid oxygen.

Dewar's original flasks were made of glass with a coating of mirror silver; this type is still used in the laboratory. But for shipment and storage of liquid gases, metal vacuum vessels are used (see **illus.**). Metal vessels with a capacity of 13 gals (50 liters) can preserve liquid oxygen with an evaporation loss of only 4% per day. Evaporation rates in 29,000-gal (110,000-liter) vessels designed for transport by railway are approximately 0.1% per day for oxygen and 0.8% per day for hydrogen.

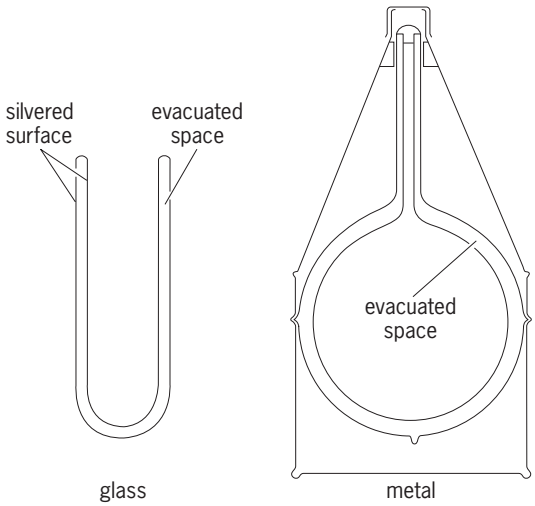

Two typical Dewar containers.

Thermos Bottle is a trademark for a Dewar vessel for hot and cold foods. H. W. Russell; G. R. Harrison

Dewaxing of petroleum

The process of separating hydrocarbons which solidify readily (waxes) from petroleum fractions. Removal of wax is usually necessary to produce lubricating oil which will remain fluid down to the lowest temperature of use. It is therefore an important step in the manufacture of lubricating oils. The wax removed may be purified further to produce commercial paraffin or microcrystalline waxes.

Most commercial dewaxing processes utilize solvent dilution, chilling to crystallize the wax, and filtration. The MEK process (methyl ethyl ketone–toluene solvent) is widely used. Wax crystals are formed by chilling through the walls of scraped surface chillers, and wax is separated from the resultant wax-oil-solvent slurry by using fully enclosed rotary vacuum filters. In one process modification, most of the chilling is accomplished by multistage injection of very cold solvent into the waxy oil with vigorous agitation, resulting in more uniform and compact wax crystals which filter faster.

In the propane process, part of the propane diluent is allowed to evaporate by reducing pressure, so as to chill the slurry to the desired filtration temperature, and rotary pressure filters are employed.

Other solvents in commercial use for dewaxing include MEK-MIBK (methyl isobutyl ketone), acetone-benzene, dichloroethane-methylene dichloride, and propylene-acetone.

Older dewaxing processes are centrifugal dewaxing, applicable only to heavy residual stocks, utilizing naphtha dilution, indirect chilling, and centrifugal separation; and cold pressing, applicable only to low-viscosity light lube fractions, in which the crystallized wax is separated from the chilled, undiluted oil in plate-and-frame-type pressure filters.

Complex dewaxing requires no refrigeration, but depends upon the formation of a solid urea-*n*-paraffin complex which is separated by filtration and then decomposed. This process is used, to a limited extent, to make low-viscosity lubricants which must remain fluid at very low temperatures (refrigeration, transformer, and hydraulic oils). Similar use is anticipated for the catalytic dewaxing process, which is based on selective hydrocracking of the normal paraffins; it uses a molecular sieve-based catalyst in which the active hydrocracking sites are accessible only to the paraffin molecules. *See* PETROLEUM; PETROLEUM PROCESSING AND REFINING; WAX, PETROLEUM.

Stephen F. Perry

Bibliography. J. H. Gary and G. E. Handwerk, *Petroleum Refining: Technology and Economics*, 1993; G. D. Hobson and W. Pohl (eds.), *Modern Petroleum Technology*, pts. 1 and 2, 5th ed., 1984; J. G. Speight, *The Chemistry and Technology of Petroleum*, 3d ed., 1999.

Dextran

A polyglucose biopolymer characterized by preponderance of α-1,6 linkage, and generally produced by enzymes from certain strains of *Leuconostoc* or *Streptococcus*. The chemical properties may be modified chemically for specific usages. While formerly its principal utility was as blood plasma substitute, dextran is also employed as packing material in column chromatography and as pharmaceutical agents; its average molecular weight determines usage to a great extent. Dextran's chemical and physical properties depend upon the strain of microorganism employed and the environmental conditions imposed upon the bacterium during growth, or the reaction conditions where an enzymatic method of dextran production is employed. *Leuconostoc* and *Streptococcus* species primarily convert sucrose to dextran and fructose. *Acetobacter* species convert dextrin to dextran; the α-1,4 linkage is converted to an α-1,6 linkage. *See* CHROMATOGRAPHY; POLYSACCHARIDE.

The conversion of sucrose to dextran and its byproduct fructose is a transglucosylation reaction. This transformation is mediated by the enzyme dextransucrase. It is readily obtained extracellularly from suitable strains of *L. mesenteroides* which are propagated under appropriate conditions. *Leuconostoc mesenteroides* NRRL B-512F has been adopted by practically all of the Western nations for the production of dextran.

Dextran, like glycogen and amylopectin, is a branched glucose polymer but differs from amylopectin in that the principal linkage between the anhydroglucose units is of the α-1,6 type; the others are of the 1,4 or 1,3 type. The enzyme system from *L. mesenteroides* NRRL B-512F forms a polymer with about 95% α-1,6 type of linkage. The ratio between the 1,6 and the non-1,6 linkages may vary from 20:1 to almost 1:1, depending on the bacterial strain, and is characteristic of the enzyme obtained from it.

Mechanism. Formation of dextran from sucrose involves the reaction shown in the **illustration**.

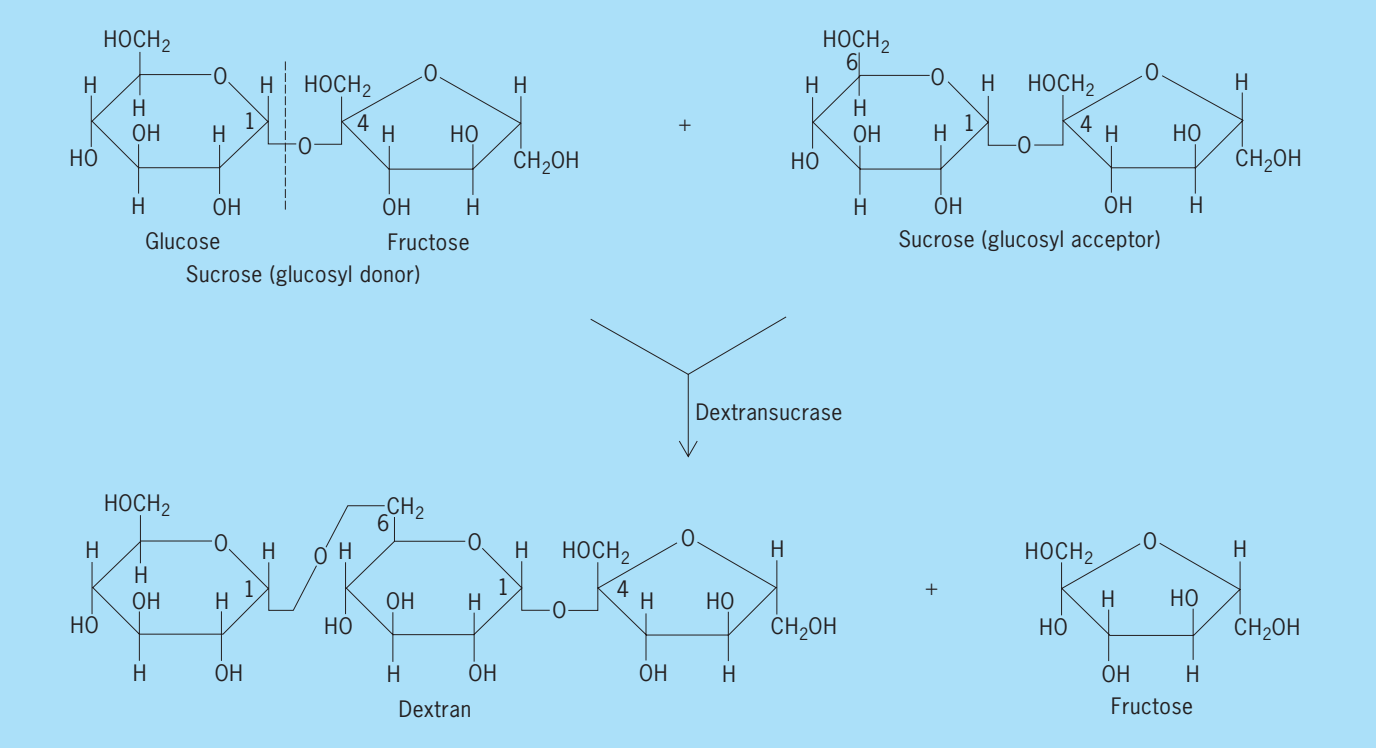

Reaction for formation of dextran from sucrose.

The reaction requires two substrates. Dextransucrase is more specific in its requirement for the glucosyl donor substrate (sucrose) than for its acceptor substrates. The acceptor substrate may be sucrose, maltose, isomaltose, α-methylglucoside, and low-molecular-weight dextran. The average molecular weight of the dextran polymerized is controlled by both the type and quantity of the acceptor substrate incorporated either in the fermentation medium or in the enzymatic reaction.

The kinetic constants can be readily varied since dextran is formed from a system where double substrate kinetics apply. Neither the V nor the Michaelis constant for the polymerization reaction is fixed; they are varied by alteration in the concentration of either cosubstrate (glucosyl donor and glucosyl acceptor). In some respects, the mechanism of the polymerization is similar to that of the condensation type in that water is eliminated. It also displays, however, certain characteristics of the chain-reaction type in that the molecular weight of the dextran produced is essentially that which is desired and requires no further modification. *See* ENZYME.

Production processes. The average molecular weight of dextran produced by most of these organisms is generally on the order of several to hundreds of millions. One strain of *Streptococcus* produces, however, a dextran with a molecular weight of about 100,000 or slightly lower. The high-molecular-weight dextran, as ordinarily produced by *L. mesenteroides,* is hydrolyzed to a product with an average molecular weight of about 75,000 for use as a blood plasma volume expander. More information is available on the utility of dextran from *L. mesenteroides* NRRL B-512F for use as blood plasma substitute than on the dextran from the *Streptococcus* sp. For use as plasma extender, the high-molecular-weight dextran is acid hydrolyzed, fractionated (as to its molecular weight distribution), and purified. The specifications on the dextran for use as a plasma volume extender vary in different countries; the physical, chemical, and biological specifications it must meet are extremely rigid. *See* INDUSTRIAL MICROBIOLOGY; POLYMERIZATION.

Henry M. Tsuchiya

Diabetes

A condition in which excessive amounts of some substances are excreted from the body. The term may refer to either of two unrelated diseases, diabetes mellitus and diabetes insipidus. The word diabetes derives from the Greek for siphon, a reference to the copious urine excretion that characterizes this affliction. In common usage, the term diabetes is synonymous with diabetes mellitus.

Diabetes Mellitus

Diabetes mellitus is a disease of abnormal carbohydrate metabolism in which glucose cannot enter the body's cells and be utilized, and therefore remains in the blood in high concentrations. Mellitus is the Latin word meaning honey, and is applicable because the urine of patients with diabetes mellitus has a sweet taste, distinguishing this disease from diabetes insipidus, in which large volumes of dilute, almost colorless and tasteless urine are passed. In diabetes mellitus the excess sugar in the blood (hyperglycemia) leads to the excretion of sugar in the urine (glycosuria), a cardinal diagnostic symptom. Glycosuria in

turn causes the excretion of large amounts of urine (polyuria), which results in dehydration and intense thirst (polydipsia). Although blood glucose is high, it cannot enter the appetite-regulating cells of the hypothalamus; hunger is therefore great, and the diabetic person tends to eat constantly (polyphagia). But because glucose cannot enter and nourish the cells, body tissues are subjected to the equivalent of starvation; rapid weight loss occurs, part of which is due to the excretion of water in urine. *See* CARBOHYDRATE METABOLISM.

Varieties. Diabetes mellitus appears in two varieties, each with its own cause: diabetes mellitus type I (formerly known as juvenile onset diabetes), caused by deficiency of the pancreatic hormone insulin (whose chief function is to promote the entry of glucose into cells); and diabetes mellitus type II (formerly known as maturity onset diabetes), in which insulin is available but cannot be properly utilized.

Diabetes mellitus type I. Of the two forms of diabetes, type I is initially the more serious and less common, afflicting approximately 1 in every 600 children. The incidence varies among countries. The highest incidence is seen in Finland and Sweden, and some of the lowest in Korea and Mexico. The United States and Canada are included among the highest. Although it usually appears before the age of 20, type I diabetes can strike at any age. It is caused by a significant shortage or complete lack of insulin secretion. The endocrine cells normally responsible for insulin production are found in the pancreas, but have nothing to do with the ordinary digestive functions of that organ. They occur as islands of endocrine tissue, called islets of Langerhans, scattered throughout the substance of the pancreas. The islets consist of at least three cell types, each of which secretes a different hormone important in regulating carbohydrate metabolism. The alpha cells produce glucagon, which elevates blood glucose; the beta cells produce insulin; and the delta cells produce somatostatin, a hormone that appears to participate in controlling the activity of the alpha and beta cells.

In type I diabetes the beta cells are destroyed, possibly by the attack of the body's own immune system or by virus infection; victims of type I require daily insulin injections to survive. Insulin must be injected rather than taken orally because it is inactivated by the enzymes of the digestive system. Without insulin therapy, a patient with type I diabetes will inevitably deteriorate into a condition of ketosis (presence in the blood of ketone bodies, the metabolic intermediates of fat metabolism) and acidosis (dangerously acidic blood pH). Ketoacidosis, as the combined condition is called, leads rapidly to coma and death in the untreated diabetic.

When Frederick G. Banting and Charles H. Best succeeded in purifying insulin from animal pancreatic tissue in 1921, type I diabetics were given their first chance at living a normal life-span. Today type I diabetics are routinely treated by injections of insulin. Human insulin became increasingly available in the 1980s through the use of recombinant deoxyribonucleic acid (DNA) technology. In this technique the gene specifying human insulin is spliced into the chromosome of a bacterium, and the billions of progeny bacteria then produce insulin as though it were one of their own bacterial proteins, thus serving as minifactories for the hormone. The bacterium used to produce insulin is *Escherichia coli*. Another DNA technique for producing insulin uses yeast instead of *E. coli*. Up to the early 1980s, insulin was extracted from cattle and pigs. This "foreign" insulin is still used, but has some unwanted side effects not seen or rarely seen with human insulin, therefore making human insulin the treatment of choice. One of the most important characteristics of insulin produced by genetic engineering, other than its being identical to human pancreatic insulin, is its limitless availability. *See* GENETIC ENGINEERING; INSULIN; PANCREAS.

Diabetes mellitus type II. Type II is the more common form of diabetes mellitus, accounting for 8 out of every 10 cases. It usually appears after the age of 30, although an uncommon variant called type II diabetes of the young occasionally occurs in obese children and teenagers. The initial symptoms of type II diabetes are much less noticeable than those of type I, and the characteristic triad of polyuria, polydipsia, and polyphagia may even be entirely absent. For this reason, type II diabetes can exist undetected for dangerously long periods. Victims are not prone to the ketoacidosis and coma that threaten type I patients, and their disease can usually be managed without insulin injections. Thus two other names for type II diabetes are nonketosis-prone diabetes and non-insulin-dependent diabetes.

Whereas type I diabetes is a disease of insulin shortage, victims of type II diabetes usually have insulin in their bloodstream. In fact, insulin levels in type II diabetics are sometimes even higher than those in nondiabetic individuals. However, since the cells of a type II patient do not respond to insulin by taking in and utilizing blood glucose as normal cells would, the type II diabetic may have hyperglycemia in spite of high insulin levels. The probable reason for this defect is that the cells have specific receptor molecules on their surface membranes whose job is to recognize insulin and trigger the biochemical steps leading to glucose uptake and utilization. Interestingly, the number of insulin receptors on cells can be changed by gaining or losing body weight; the more body fat a person carries, the fewer insulin receptors there will be on the cells. This may explain why type II diabetes is mainly a disease of the obese, and why weight reduction is such effective therapy. High glucose levels also decrease the number of insulin receptors; normalizing glucose levels is thus of utmost importance.

The most important treatment for most cases of type II diabetes consists of exercise and weight reduction, which will return glucose metabolism to normal in 9 out of 10 cases. Because the cells of type II diabetics lack responsiveness to their own insulin, insulin injection therapy is used for some of these cases. Approximately 10% of all type II patients must use insulin. This percentage appears to more than double in Black, Latino, Native American, or other minority patients living in the United States. In

addition, a family of drugs known as oral hypoglycemic agents is often used in treating the disease. These drugs, including tolbutamide, chlorpropamide, acetohexamide, and tolazamide, lower blood glucose by stimulating the pancreatic cells to release insulin, and by augmenting insulin's effect on the cells of the body. All of these drugs are known as first-generation sulfonylureas. However, there are other drugs with the same basic molecule but with chemical changes in their structure; these are called second-generation sulfonylureas. This group of drugs is safe if properly used and is effective in many cases if used in conjunction with proper diet and exercise programs.

Amylin. Amylin is a 37-amino-acid peptide hormone that occurs naturally in humans, and has been shown to be in excess in individuals with non-insulin-dependent diabetes mellitus. Amylin has been shown to be deposited in the islets of Langerhans, and may decrease insulin release and insulin production in the pancreas, thus causing a deterioration in glucose control in non-insulin-dependent diabetes.

In insulin-dependent diabetes, amylin appears to slowly decrease with the progression of the disease, and when none is found, individuals may be more prone to have hypoglycemia unawareness (low blood glucose levels without symptoms).

Complications. Both type I and type II diabetes tend to produce a number of debilitating and life-threatening complications despite the use of insulin or oral hypoglycemic agents. It is primarily these complications, not the risk of ketoacidosis, that account for the heavy burden this disease inflicts on society, and make diabetes the third leading cause of death in the United States.

Much evidence indicates that most of the complications of diabetes are caused by the elevated blood glucose levels common in this disease. Even diabetics who take regular insulin injections exhibit wide fluctuations in their blood glucose levels, in marked contrast to the tightly controlled blood glucose levels observed in nondiabetics. According to the prevalent hypothesis, the more diabetics succeed in controlling their blood glucose to near-normal levels, the less risk there will be of developing any complications of the disease. Thus, careful attention to diet, exercise, regular medication, and monitoring of blood glucose is vitally important for any diabetic.

The most common complications of diabetes are the following.

Eye damage. Diabetes is the leading cause of nontraumatic blindness in young adults. The blood vessels of the diabetic eye are prone to microaneurysms (small outpouching of the vessel wall), excessive vascular permeability (leakage from the vessel), growth of newer, more fragile vessels, and ultimately hemorrhage (bleeding). These events can lead to scarring, fibrosis, and subsequent loss of vision.

Kidney damage. In long-standing diabetes, the smallest blood vessels of nephrons (the functional filtration units of the kidney) frequently suffer the same type of damage as the blood vessels of the retina; this condition is called diabetic nephropathy. Severe diabetic nephropathy commonly leads to kidney failure, and such patients require dialysis in order to survive. The prognosis for individuals with diabetes on dialysis is poor, with only 30% surviving beyond 5 years.

Premature atherosclerosis. Atherosclerosis (often called arteriosclerosis or hardening of the arteries) is a condition in which arteries become progressively occluded by accumulations of cholesterol-containing plaque. Although atherosclerosis is common in older nondiabetics, diabetics tend to develop it earlier and in a more severe form. The most common consequence of atherosclerosis is coronary artery disease and stroke, which result when the blood supply to the heart or the brain is obstructed by a plaque within an artery. Coronary artery disease is the leading cause of death in long-term diabetics, accounting for 75% of fatalities. Also, because of atherosclerosis and restricted blood flow, diabetics are often subject to gangrene, especially of the extremities. *See* ARTERIOSCLEROSIS.

Nervous system abnormalities. Patients with diabetes may experience a variety of nervous system disturbances, collectively referred to as diabetic neuropathy. Symptoms may include muscle weakness, pain in the extremities, local paralysis, urinary incontinence, sexual impotence, and sensations of cold, heat, or tingling in various parts of the body. There may also be episodes of diarrhea, frequently nocturnal, which is a characteristic of the "diabetic diarrhea" problems involving the upper intestinal tract. Nervous system involvement of the esophagus and stomach may cause nausea and even vomiting.

Causes. Three factors that are believed to be important in the causation of this disease are heredity, viral infections, and immunological damage to the pancreas.

Heredity. Although diabetes is popularly considered a genetic disease, it is not inherited in the sense that one can inherit curly hair. Heredity does operate in determining one's risk of diabetes, but only as a predisposing factor, not as an absolute determinant. The two types of diabetes mellitus have entirely different patterns of inheritance, and the presence of one type in a family does not influence one's risk of developing the other. Genetic factors play a much more important role in type II diabetes than in type I diabetes. For example, the identical twin of a person who develops type II diabetes after age 40 is almost 100% certain to develop the disease, but someone whose identical twin has type I diabetes runs only a 52% risk of becoming a diabetic. Parents with type I diabetes face only a 2% risk of having a child with the disease, whereas the risk to the offspring of a type II diabetic is about 10%. Genetic predisposition is apparently stronger in the case of type II diabetes of the young, with a 50% chance that the child of such a diabetic will also develop the disease. The environment obviously plays a significant role in type II diabetes. *See* HUMAN GENETICS.

Virus infections. Virus infection is strongly implicated as one causative factor in the beta-cell destruction that characterizes type I diabetes, but it is not involved in type II diabetes. Some of the viruses that have been implicated are the agents of mumps and

rubella (German measles), whose action seems to be delayed for months or years, and coxsackievirus B, which is more likely to precipitate diabetes shortly after the person recovers from the influenzalike infection. Apparently, several factors operate to determine whether or not diabetes will develop in a person infected with one of these extremely common viruses. They include the organ affinity of the particular virus strain, the person's susceptibility to viral invasion of the pancreatic tissue, and the individual's capacity to regenerate islet tissue that has been damaged by a viral infection. *See* ANIMAL VIRUS.

Immunological damage to pancreas. Autoimmune reactions, wherein the body's immune defense system attacks its own pancreatic tissue as though it were a foreign substance, are also suggested as a cause for the beta-cell destruction of type I diabetes. An autoimmune response may be conducted by T lymphocytes or by soluble antibodies in the blood or by both. These reactions may arise spontaneously or may be secondary to a viral infection of the pancreas that leaves the islet tissue modified in some way that marks it for attack. *See* AUTOIMMUNITY.

Diet. Dietary management has been part of diabetic therapy since preinsulin days, when starvation diets were used to prolong the life of type I diabetics. More recently, weight reduction for victims of type II diabetes and complex carbohydrate liberalization for those of both types of diabetes mellitus have been the cornerstone of therapy.

For obese type II diabetics, weight reduction is certainly appropriate, since it tends to increase the number of insulin receptors on the cells of the body, usually making it possible to normalize blood glucose without resorting to the use of insulin injections or oral drugs.

Some other dietary measures besides reducing caloric intake are now believed beneficial for type II diabetics. The ingestion of foods high in fiber content, for example, has been shown to reduce hyperglycemia in type II diabetics.

For type I diabetics, the traditional diet is high in protein and quite restricted in carbohydrates, with only about 45% of the total daily caloric intake being carbohydrate-derived. Since most type I diabetics are thin, weight reduction is not part of their program. Presently, the recommended carbohydrate intake has been liberalized to 55–60% of the diet, plus 15–20% protein and 20–30% fat. There has been a trend toward relaxing the carbohydrate restriction for type I diabetics as evidence accumulates that diets containing even as much as 80% carbohydrates can be compatible with excellent control of diabetes, as long as the carbohydrates are complex (that is, whole grains, legumes, potatoes, bread) rather than refined (that is, sugars, candy, soft drinks, desserts, and so on). Artificially sweetened food can be very beneficial for weight control; however, it is not a necessary component of treatment.

Prospects. The ideal solution is to be able to mimic insulin secretion, with intensive insulin therapy used with either multiple insulin injections or insulin infusion pump therapy. The insulin pump continuously delivers a small amount of insulin, and the individual can increase the amount before meals based on his or her glucose level. Still, this setup is not ideal: in a nondiabetic, insulin is delivered directly into the portal blood system, while therapy delivers insulin under the skin, where it can be absorbed into the bloodstream. Hooking a glucose-monitoring device to the insulin pump would allow the system to close, permitting the individual to have an artificial pancreas implanted (several models are available).

Healthy pancreatic tissue can be transplanted from a donor into a type I diabetic, but there are many problems related to this method, such as the typing of the pancreatic tissue from the donor, the possibility of rejection, and the possibility of infections.

Diabetes Insipidus

Diabetes insipidus is caused by a deficiency of, or resistance to the action of, vasopressin, the antidiuretic hormone produced by the posterior lobe of the pituitary gland. If the pituitary fails to produce vasopressin, the condition is called central diabetes insipidus. If the kidneys do not respond to the vasopressin and fail to concentrate urine, the condition is labeled nephrogenic diabetes insipidus. Symptoms include increased thirst and frequent urination. Permanent damage to the kidneys can result if the condition is not treated.

A low level of vasopressin can be corrected with desmopressin acetate, a synthetic analog of 8-argi-ninevasopressin, which can be administered in the form of a nasal spray. Desmopressin acetate is ineffective in the treatment of nephrogenic diabetes insipidus. *See* NEUROHYPOPHYSIS HORMONE.

Jaime Davidson; Peter E. Bressler

Bibliography. S. Baba, M. K. Gould, and P. Pimmet (eds.), *Diabetes Mellitus*: *Recent Knowledge on Aetiology, Complications and Treatment*, 1985; W. E. Benson et al., *Diabetes and Its Ocular Complications*, 1988; M. Brenner, *Management of the Diabetic Foot*, 1987; P. Czernichow and A. D. Robinson (eds.), *Diabetes Insipidus in Man*, 1985; E. Friedman and C. M. Peterson (eds.), *Diabetic Nephropathy*: *Strategy for Therapy*, 1985.

Diagenesis

All the chemical, biochemical, and physical changes that sediments undergo from the time of deposition until the stage of metamorphism is reached. Diagenetic changes are gradational with metamorphism at elevated temperatures or pressures, and with atmospheric weathering effects when sedimentary rocks are exposed on the Earth's surface. Diagenesis commonly transforms sediments into sedimentary rocks and results in significant chemical and mineralogical changes. With increased burial, sands undergo grain rearrangement and compaction, mineral dissolution or replacement, and mineral precipitation as cements. Clays formed during weathering processes are transformed into illite, chlorite, and muscovite. Porous carbonate sediments consisting of unstable aragonite or magnesium-calcite are transformed to nonporous limestones composed of calcite. In spite

of these profound changes, many features such as macroscopic sedimentary structures may be preserved in both detrital (sandstones and shale) and carbonate rocks. Diagenetic processes and products vary with original sediment composition and depositional environment, depth of burial (pressure and temperature), burial time, nature and movement of interstitial fluids, and physical, chemical, and biochemical processes.

While all of the variables mentioned above are important in controlling diagenesis, the one that received the most attention in the 1990s was interstitial fluid flow. At shallow subsurface depths, the sediment-water system is open, and diagenetic processes are driven by the chemistry of pore waters that migrate through the sedimentary units. The majority of these waters are meteoric or derived from compacting shale units. At greater depths, permeabilities are reduced, and the system becomes progressively more closed. Under these conditions, the composition of framework grains may be the dominant control on diagenesis. The system can also be reopened at depth as a result of fluids introduced by clay transformations, by hydrocarbon maturation, or along fault and fracture systems (**Fig. 1**).

Processes. Diagenetic processes include purely physical ones that involve rearrangement of the sediments, such as compaction, slumping, bioturbation by organisms, infiltration, and soft sediment deformation; biochemical or organic processes such as particle accretion, boring, and decomposition; and physiochemical processes such as cementation, authigenesis (formation of new minerals), inversion, recrystallization, grain growth, replacement, and dissolution.

Diagenetic processes are also affected by microbial activity, primarily of aerobic and anaerobic bacteria, that can influence the pH and the oxidation potential (Eh) of sediment pore water and the accumulation of chemical reactants. Anaerobic bacteria may survive in sedimentary basins to depths of several thousands of meters below the Earth's surface. Sulfate-reducing bacteria, for example, may interact with the sulfate radicals in dissolved formation waters, producing hydrogen sulfide, calcite, native sulfur, and pyrite, among other products important in diagenetic reactions.

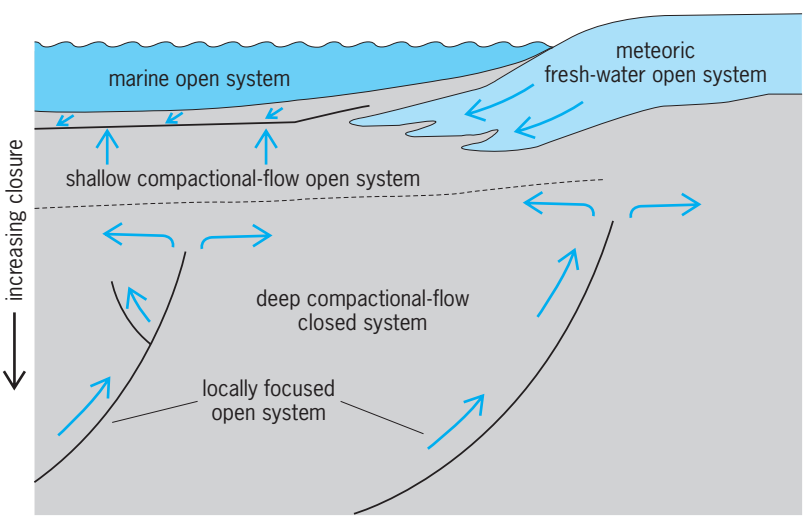

Fig. 1. Subsurface flow systems. Diagenetic processes occur in open systems at shallow subsurface depths (meteoric fresh-water open systems and shallow compactional-flow open systems). At greater depths, decreasing flow rates promote closure below approximately 1 km (3300 ft) [deep compactional-flow closed systems]. Deep-burial open-system conditions are restricted to areas associated with fault or fracture systems. (*From M. D. Wilson, ed., Reservoir Quality Assessment and Prediction in Clastic Rocks, SEPM Short Course, no. 30, 1994*)

Most diagenetic processes involve a reduction in porosity and permeability. An understanding of these two sediment properties is critical to understanding the migration of subsurface fluids and the accumulation of oil and gas and certain types of mineral deposits in subsurface rock units. However, one diagenetic process, dissolution, is of major importance in creating secondary porosity. Only a few of the more important processes will be reviewed here.

Although compaction and cementation probably constitute the principal processes responsible for loss of porosity in sediments, several processes that operate penecontemporaneously with deposition may lead to loss of porosity, even before compaction and cementation begin. These processes include bioturbation and mechanical infiltration. Bioturbation is the disturbance of the original sediment fabric by the activity of animals and plants. The principal activity is burrowing by invertebrate or small vertebrate organisms. Burrowing usually results in a worsening of sorting from the mixing of sand and clay layers and a corresponding reduction in porosity in the sand layers.

Mechanical infiltration involves the downward movement of clay-size particles carried by percolating meteoric waters. This process is particularly important in coarse-grained desert deposits where water tables are low and streams are influent during rare periods of runoff. Infiltrated clays are recognized by their clastic texture and the fact that they tend to accumulate on the upper surface of grains in geopetal fabric.

As sediments are buried, the weight of overburden or tectonic stress results in compaction. During compaction, grains move closer together, pore space is reduced, and pore waters are expelled. The amount of compaction achieved is a function of the nature of the sediment, the depth of burial, and the degree of cementation. Compaction of sediments is usually achieved by one or more of the following processes: rotation and slippage of grains; plastic or ductile, flexible, and brittle grain deformation; and pressure solution (**Figs. 2** and **3**).

Cementation involves the precipitation of authigenic minerals from aqueous solutions in the pore spaces of sediment. Geochemical calculations suggest that large volumes of pore fluids must pass through the rock to account for the observed authigenic minerals. The most commonly documented authigenic cements are carbonates, silica (quartz), clays, and zeolites. *See* AUTHIGENIC MINERALS.

Three closely related processes important in the diagenesis of limestone are inversion, recrystallization, and grain growth. Inversion is the process by

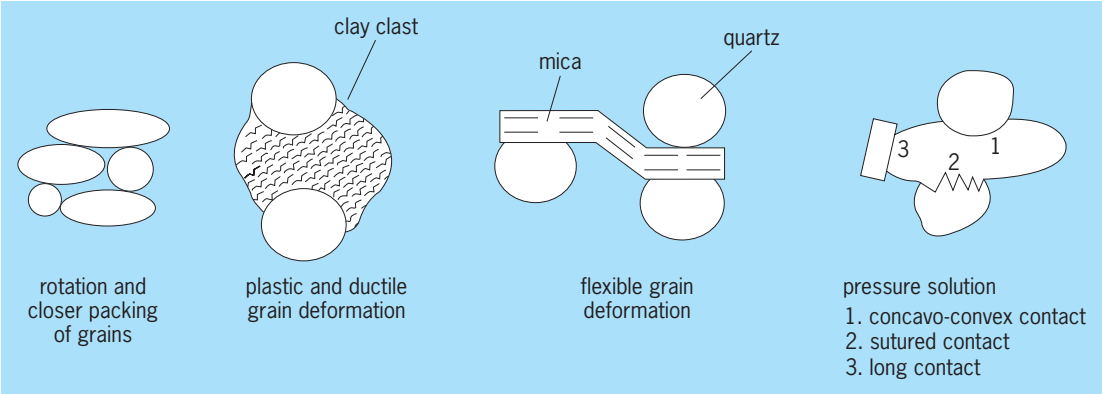

Fig. 2. Some methods of porosity loss resulting from compaction, including rotation and closer packing of grains; plastic and ductile grain deformation; flexible grain deformation; and pressure solution. (*From M. D. Wilson, ed., Reservoir Quality Assessment and Prediction in Clastic Rocks, SEPM Short Course, no. 30, 1994; and J. C. Wilson and E. F. McBride, Compaction and porosity of Pliocene sandstones, Ventura basin, California, Amer. Ass. Petrol. Geol. Bull., 72:663–681, 1988*)

which unstable minerals are converted to a more stable form with no change in chemical composition but with a different crystal structure. The most commonly documented example is the inversion of aragonite to calcite. Recrystallization and grain growth refer to processes that result in a change in crystal or grain size without chemical alteration. The distinction of which process was responsible for the resulting texture seen in many ancient limestones is not always possible. Neomorphism includes all three processes. *See* ARAGONITE; CALCITE.

Interstratal solution involves the dissolution of mineral grains and cements (Fig. 3) from strata by migrating pore fluids or ionic diffusion. This process is locally important in the formation of many oil and gas reservoirs in both limestones and sandstones. Clues to the recognition of secondary porosity in sandstones include oversized pores, partially dissolved and corroded grains, enlarged pore throats, and floating grains.

Sandstones. Three major groups of sandstones have been recognized: quartz arenites, arkosic arenites, and lithic arenites. Most quartz-rich sandstones (quartz arenites) are partially or completely cemented by silica and calcite. Secondary silica commonly occurs as overgrowths in optical continuity with original grains. Considerable controversy remains over the source of authigenic silica. Possible sources include pressure solution, illitization of smectitic clays, siliceous microfossils, dissolution of volcanic glass, feldspar alteration and dissolution, and carbonate replacement of silicate minerals. Figure 3 suggests that quartz overgrowths occur at relative low rates and over a very wide range of depths.

Carbonate cements, including calcite, dolomite, siderite, and ankerite, are the most common and diverse cement in sandstones. Several different generations of carbonate cement can be found in one sandstone formation, and they may display a considerable range of crystal forms, sizes, and distributions. The abundance of calcite cement results, in part, from the fact that connate water in shallow marine sediments is oversaturated with respect to calcium carbonate. Other sources of carbonate ions might be dissolution of fossils, clay membrane filtration, and pressure solution of interbedded limestones. *See* SANDSTONE; SILICATE PHASE EQUILIBRIA.

Clay minerals and zeolite cements may be abundant in mineralogically immature arkosic and lithic sandstones that are rich in feldspar and volcanic or other highly unstable rock detritus. These authigenic clays have delicate, crystalline morphologies and occur as rim cements on grains, as occluding cements between grains, and as fracture or vug fillings. Common authigenic clay minerals include

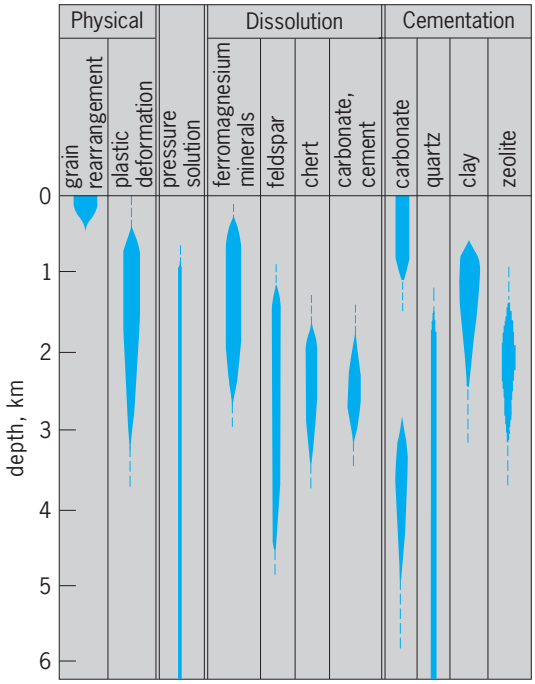

Fig. 3. General depth regimes for major diagenetic processes in detrital sediments and sedimentary rocks. Physical processes tend to dominate at shallow depths, and chemical processes at greater depths. Width of bars indicates the relative rate at which processes operate. (*From M. D. Wilson, ed., Reservoir Quality Assessment and Prediction in Clastic Rocks, SEPM Short Course, no. 30, 1994*)

chlorite, illite, smectite, kaolinite-dickite, and mixed-layer illite-smectite. Zeolite pore-fill cements may accompany these diagenetic clays in volcanic sandstones. The principal control on the development of diagenetic clays and zeolites in mineralogically immature sandstones is probably temperature, which is controlled by depth of burial. Not all immature sandstones contain abundant authigenic clay minerals or zeolites. Some sandstones contain lithic fragments, which deform plastically to produce a pseudomatrix, thus reducing permeability, restricting fluid movement, and retarding development of authigenic clay minerals and zeolites.

Three other diagenetic processes are important in felspathic sandstones. In fine-grained feldspathic sandstones, with little clay matrix, diagenetic potassium feldspars form as overgrowths on feldspar grains and as euhedral to subhedral crystals filling pore spaces. Potassium- and calcic-rich feldspathic sandstones can be altered by albitization (formation of authigenic sodium-rich albite). The principal reason that albite forms is that most pore waters are rich in sodium ions. The last process of feldspar alteration is calcification, which involves the replacement of feldspars by calcite.

The red color in some arkosic or lithic sandstones is provided by hematite, which is of diagenetic origin. The sediments are not red when deposited, but redden with time. Iron-bearing minerals in these immature sandstones alter by hydrolysis and release iron, which precipitates as hematite or as a precursor oxide which converts to hematite upon aging. *See* HEMATITE; REDBEDS.

Secondary porosity, caused by dissolution of primary or secondary minerals, is found in sandstones of all ages and depths. In spite of this fact, there remains considerable controversy over the origin, importance, and effect of secondary porosity on hydrocarbon-reservoir and ground-water-reservoir quality. Meteoric water influx is the most important leaching mechanism, and porosity produced by this mechanism may persist during deeper burial, provided that the sandstones do not have high percentages of ductile grains. Mechanisms for the production of secondary porosity at depth include silicate hydrolysis and organic acids. The quantitative importance of these various mechanisms is unclear.

Shales. Diagenetic changes in shales involve color changes, compaction, decomposition of unstable minerals and recrystallization of others, and development of micas of restricted composition. Clay-mineral components of fine-grained sediments show evidence of diagenesis, starting with ionic substitution in clay mineral lattices, and ending with complete conversion from one clay mineral type to another, as in the case of the conversion of smectite to illite. One theory that explains the mineral alteration of smectite to illite is known as the crystal growth theory, and suggests that this conversion not only involves the addition of potassium to the smectite crystals but also requires that the crystals grow in size.

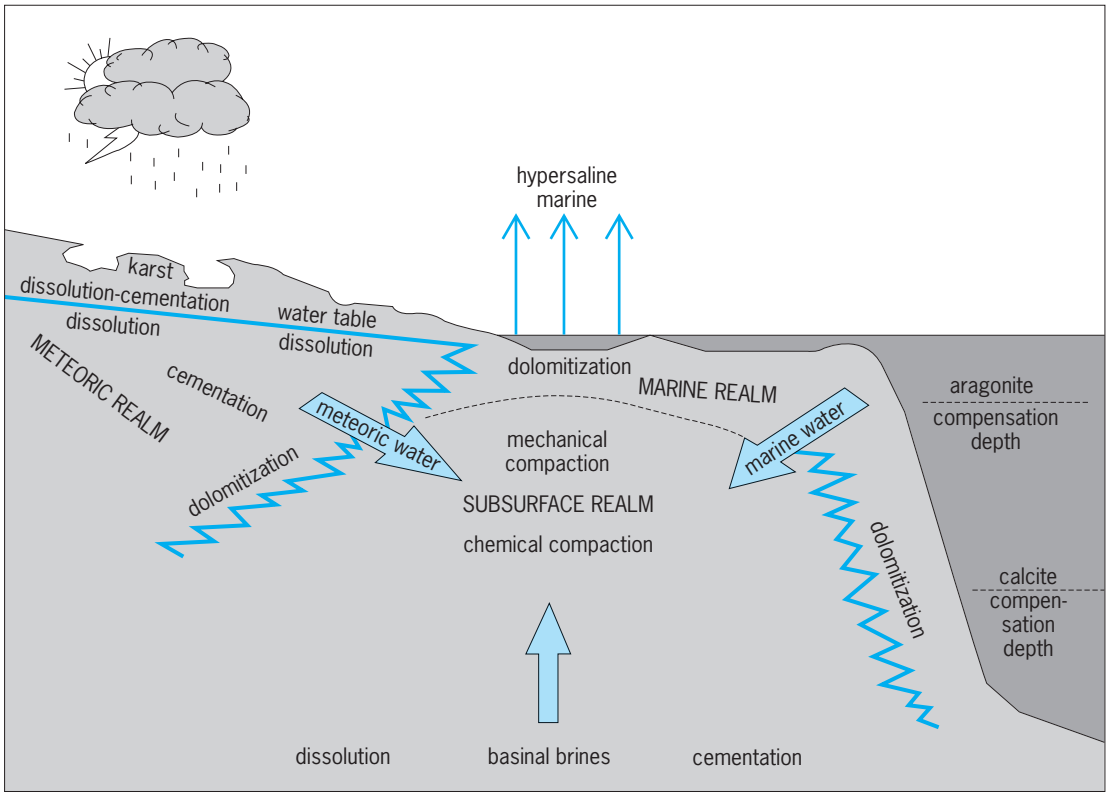

Fig. 4. Principal subsurface realms in which diagenesis of carbonate sediments occur. Major diagenetic processes that operate in each of the realms (meteoric, marine, and subsurface) are indicated. (*From C. H. Moore, Carbonate Diagenesis and Porosity, Elsevier, Amsterdam, 1989*)

Compaction is another important process in shales. Diagenetic changes accompanying compaction and mineral alteration with increased depth of burial and corresponding increases in pressure, temperature, and time result in a reduction in porosity and permeability, and the expulsion of both water and adsorbed hydrocarbons from the shales. The processes are therefore considered important to the origin of petroleum. Sediments with high percentages of clays usually contain large quantities of water at the time of deposition. Pore waters are usually expelled first during the early stage of compaction as platy clay particles pack together to form a less open framework. During the initial stage, original sediment volume is reduced by 50% or more. After this degree of compaction is achieved, further compaction is hindered by the small size of the pore diameters, and by the fact that most of the remaining waters are attached to the clay surfaces. Final expulsion of the remaining waters and any accompanying hydrocarbons may be accomplished only after the final mineral alteration of smectite to illite at burial between 6000 and 8000 ft (1800 and 2400 m). In some thick shale sequences, however, compaction may be delayed because an impermeable seal prevents the escape of pore waters. In these sequences, abnormally high pressures may build up. *See* CLAY MINERALS; ILLITE; SHALE.

Carbonates. Carbonate sediments are recognized as being especially susceptible to diagenetic modification. Carbonate diagenesis occurs in three regions or realms (meteoric, marine, and subsurface; **Fig. 4**). Marine carbonate sediments can be brought into contact with meteoric water (fresh water) by a sea-level fall or tectonic uplift. Meteoric waters are highly aggressive and readily dissolve aragonite and high-magnesium calcite. Dissolution of these two minerals may saturate pore waters with respect to calcium carbonate, resulting in precipitation of calcite cement. Karst topography is commonly generated on subaerially exposed marine carbonates by dissolution. *See* KARST TOPOGRAPHY.

On the sea floor and in the shallow-marine subsurface (the marine real), the principal diagenetic processes are bioturbation, modification of shells by boring organisms, dissolution and cementation of grains, and penecontemporaneous dolomitization. This last process is characteristic of nearshore, hypersaline environments, and takes place while the host sediment is in its original depositional setting (Fig. 4).

As carbonate sediments and rocks are buried, pressure and temperature increases and the pore fluids change. Eventually the carbonates reach the subsurface realm and undergo physical and chemical compaction and additional dissolution, cementation (Fig. 4), and replacement as dictated by the specific conditions (pore-fluid composition, temperature, pH, and Eh) of the subsurface realm. Dolomite rock (dolostone) is a common and important replacement product that occurs in this realm. Beneath modern carbonate islands, this type of dolomite is more common and pervasive than penecontemporaneous dolomites, and may provide an analog for the extensive dolostones found in the geologic record. Important hydrocarbon reservoirs are found in many of these dolostones where replacement of calcite by dolomite has resulted in the formation of intercrystalline porosity. *See* CARBONATE MINERALS; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; DOLOMITE ROCK; SEDIMENTARY ROCKS; SEDIMENTOLOGY.

Frank G. Ethridge

Bibliography. H. Blatt and R. J. Tracy, *Petrology, Igneous, Sedimentary, and Metamorphic*, W. H. Freeman, New York, 1996; S. Boggs, Jr., *Principles of Sedimentology and Stratigraphy*, Prentice Hall, 1995; D. A. Budd, Cenozoic dolomites of carbonate islands: Their attributes and origin, *Earth Sci. Rev.*, 42:1–47, 1997; J. A. Kupecz, J. G. Gluyas, and S. Bloch (eds.), Reservoir Quality Prediction in Sandstones and Carbonates, *Amer. Ass. Petrol. Geol. Mem.*, no. 69, 1997; L. S. Land, Mass transfer of large-scale burial diagenesis in the Gulf of Mexico sedimentary basin: An overview, in I. P. Montañez, J. M. Gregg, and K. L. Shelton (eds.), Basin-Wide Diagenetic Patterns: Integrated Petrologic, Geochemical, and Hydrologic Considerations, *SEPM Spec. Publ.*, no. 57, 1997; C. H. Moore, *Carbonate Diagenesis and Porosity*, Elsevier, Amsterdam, 1989; J. C. Wilson and E. F. McBride, Compaction and porosity of Pliocene sandstones, Ventura basin, California, *Amer. Ass. Petrol. Geol. Bull.*, 72:663–681, 1988; M. D. Wilson (ed.), Reservoir Quality Assessment and Prediction in Clastic Rocks, *SEPM Short Course*, no. 30, 1994.

Dialysis

A process of selective diffusion through a membrane by dissolved solutes in liquid solution. As dialysis is usually carried out, the membrane permits the diffusion of low-molecular-weight solutes (crystalloids) but prevents the passage of colloidal and high-molecular-weight solutes (macromolecules). Membranes suitable for this purpose include vegetable parchment, animal parchment, goldbeater's skin (peritoneal membranes of cattle), fish bladders, dialyzing cellophane (Visking sausage casing), and collodion (nitrocellulose deposited from alcohol-ether solution).

The solution is contained within such a membrane. The low-molecular-weight solutes are removed by placing pure solvent outside the membrane. This solvent is changed periodically or continuously until the concentration of diffusible solutes in the solution is reduced to near zero. The technique is used extensively in separating and purifying macromolecules of biological origin (see **illus.**).

Dialysis rates of ionic low-molecular-weight solutes may be increased greatly by applying an electric field to achieve an electrophoretic movement of ions through the dialytic membrane. This combined process of dialysis and electrophoretic transport of solutes through a membrane is known as electrodialysis. *See* ELECTROPHORESIS.

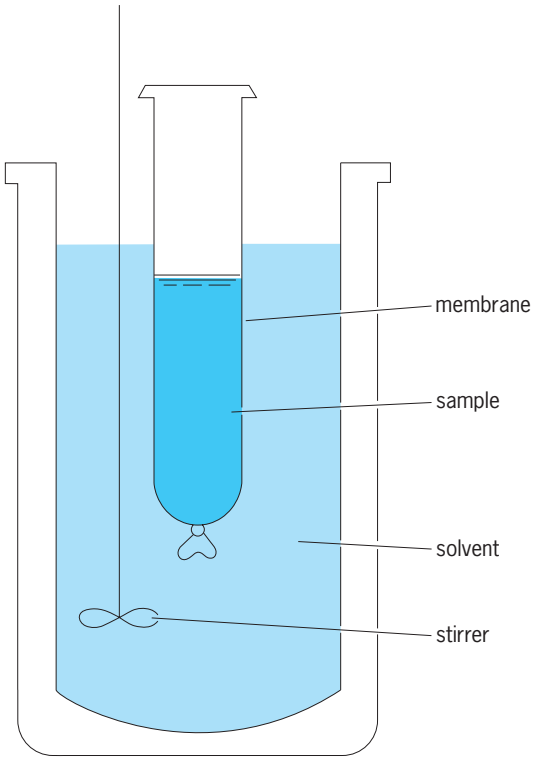

Equipment for dialysis.

Ions do not migrate readily through membranes that carry in their pores charges of the same sign as the ion; ions of charge opposite to that of the membrane are not prevented from passing through. Such a selective permeability to ions by charged membranes has been used to improve the efficiency of electrodialysis. Membranes of collodion, vegetable parchment, and cellophane carry negative charges when in contact with aqueous solutions. Animal membranes show positive charges at low pH and negative charges at high pH.

Membranes of ion-exchange resins are commercially available in cationic and anionic forms. The ion-exchange membranes show a high selectivity for ions of one charge type. *See* ION-SELECTIVE MEMBRANES AND ELECTRODES.

For desalting small quantities of solutions for chromatography and radioisotope tracer studies, micro- and semimicroelectrodialyzers have proved valuable. On a larger scale, electrodialysis combined with ion-exchange membranes has been used to remove salt from seawater. *See* COLLOID; WATER DESALINATION.

Quentin Van Winkle

Diamagnetism

That branch of magnetism which treats of diamagnetic phenomena and of the properties of diamagnetic bodies. Diamagnetism is a property exhibited by substances with a negative magnetic susceptibility, that is, by substances which magnetize in a direction opposite to that of an applied magnetic field. A diamagnetic substance has a magnetic permeability less than 1, and is repelled when placed near a magnet. The magnetization of diamagnetic substances is associated with the currents induced on application of a magnetic field. According to Lenz's law, the flow of an induced current is in such a direction as to oppose the change of flux of the inducing field; this accounts for the negative susceptibility. The diamagnetic susceptibility is invariably small, of the order of -10^{-5} cm^3/mole. *See* LENZ'S LAW; MAGNETIC SUSCEPTIBILITY.

All matter responds to applied fields in this diamagnetic fashion. However, some substances also have net electronic orbital or spin magnetic moments, or both, which can be aligned by an applied magnetic field in direction along (not opposite to) the field; this property is called paramagnetism. For these substances, as shown in Eq. (1), the observed susceptibility χ is the sum of diamagnetic and paramagnetic terms. Under ordinary conditions χ_d is temperature-independent. Hence, if χ_p follows the inverse temperature dependence of Curie's law, one can experimentally determine the separate contributions of χ_d and χ_p by measuring χ as a function of temperature. However, if χ_p is also temperature-independent, as in metals and in Van Vleck paramagnetism, the problem of determining the separate contributions becomes quite difficult. *See* PARAMAGNETISM.

$$\chi = \chi_d + \chi_p \tag{1}$$

Although all matter exhibits diamagnetism, only those substances in which $\chi_p = 0$ are referred to as diamagnetic. This is because paramagnetism, if present, usually predominates, and the gross magnetic response of the material is paramagnetic. Important exceptions are the alkali and alkaline earth metals, where χ_p is unusually small and of the order of χ_d, and salts or solutions containing only a small fraction of paramagnetic atoms. The condition for $\chi_p = 0$, and hence pure diamagnetism, is that all electronic spins be paired and all orbital moments either be zero or effectively cancel one another. Nearly all molecules with an even number of electrons satisfy this condition; an important exception is O_2. The condition is also satisfied by most nonmetallic solids, except compounds containing atoms with incomplete inner-shell electron groups, such as the transition, rare-earth, and actinide elements. *See* ELECTRON SPIN.

As stated previously, the diamagnetic response of a substance is small; only a very small fraction of the applied magnetic field is shielded from the interior of the substance by the induced diamagnetic currents. There is one case, however, in which the inducing field is completely shielded (except for small surface effects). This is the perfect diamagnetism exhibited by superconductors, and is known as the Meissner effect. *See* MEISSNER EFFECT; SUPERCONDUCTIVITY.

Langevin theory. The Langevin theory of diamagnetism is based on the idea of inducing an electronic current inside an atom. The theory employs the Larmor theorem. *See* LARMOR PRECESSION.

In a magnetic field H, the precession of the Z electrons within the atom is equivalent to a current equal to $-Z(e/c)(\omega_L/2\pi)$ in electromagnetic units. Here e/c is the magnitude of the electronic charge in emu, and ω_L is the angular Larmor frequency as shown in Eq. (2), where m is the electronic mass. The mag-

$$\omega_L = \frac{-eH}{2mc} \tag{2}$$

netic moment μ arising from this induced current is equal to the product of the current and the area of the current loop, as in Eq. (3), where $\overline{\rho^2}$ is the sta-

$$\mu = \frac{-Ze}{c}\frac{\omega_L}{2\pi}\overline{\rho^2} \tag{3}$$

tistical average, over a large number of atoms, of the square of the perpendicular distance of an electron from the field axis. This average is equivalent to $\overline{x^2} + \overline{y^2}$ if H is along z. For a random assembly of atoms, since $\overline{x^2} = \overline{y^2} = \overline{z^2}$, one may write Eq. (4), where $\overline{r^2}$ is the mean-square distance of the electron from the

$$\overline{\rho^2} = 2/3\left(\overline{x^2} + \overline{y^2} + \overline{z^2}\right) = 2/3\,\overline{r^2} \tag{4}$$

nucleus. Thus, the diamagnetic susceptibility of N atoms is given by Eq. (5). This is P. Langevin's result,

$$\chi_d = \frac{N\mu}{H} = -\frac{Ze^2N}{6mc^2}\overline{r^2} \tag{5}$$

as corrected by W. Pauli. The molar susceptibility χ_M is obtained by replacing N in Eq. (5) by Avogadro's number. Numerically, it may be expressed as Eq. (6),

$$\chi_M = (-2.83 \times 10^{10}\ \text{cm/mole})\ \Sigma\overline{r^2} \tag{6}$$

where the summation is to be taken over all the electron orbits in the atom. Since $\overline{r^2}$ is of the order of 10^{-15} cm^2, this gives χ_M of about -10^{-5} cm^3/mole.

Langevin's formula is not modified by quantum mechanics, and the problem becomes that of determining $\overline{r^2}$ of the electronic wave function. The calculation for many-electron atoms is quite involved, and experimental values of $\overline{r^2}$ as determined by using Eq. (5) or (6) give a very useful check of the nature of the wave function for large r. This complements x-ray and electron diffraction data which give information, for the most part, only for small r.

Ionic crystals. In the case of diamagnetism in ionic crystals, the Larmor theorem holds for the individual ions. The diamagnetic susceptibility may be computed with reasonable accuracy from the sum of the individual ion susceptibilities.

Diamagnetism in rare gases and in rare-gas configurations of ions in ionic crystals is shown in the **table**.

Molar diamagnetic susceptibilities (all $\times 10^{-6}$ cm^3/mole)

He −2.0	Li −.07	Mg −4.3	F −9.4
Ne −7.0	Na −6.1	Ca −10.7	Cl −24.2
Ar −19.2	K −14.6	Sr −18.0	Br −34.5
Kr −28	Rb −22.0	Ba −29.0	I −50.6
Xe −43	Cs −35.0		

Molecules. In most molecules the electrons are not moving in a single field of force, and the Larmor theorem breaks down. The total susceptibility of H_2 is given by the sum of Eq. (5) and the Van Vleck paramagnetism, the latter corresponding to the presence of a mean-square magnetic moment, although the mean moment vanishes. Calculations for more complicated molecules are exceedingly difficult. In aromatic ring molecules, such as benzene, with H normal to the ring, the electrons can precess around the ring, or at least in partial ringlike orbits about many nuclei. This gives rise to a much larger susceptibility than is possible when H is parallel to the ring. Crystals with layerlike structures, such as antimony, bismuth, and graphite, also exhibit large anisotropies in diamagnetic susceptibilities.

Bohr–van Leeuwen theorem. This theorem proves the complete absence of magnetism in classical theory. For bound electrons this comes from a cancellation of paramagnetic and diamagnetic susceptibilities, providing one does not (as did Langevin) implicitly quantize the paramagnetic moments by setting them all equal to μ. For a system of free electrons confined to a box, the induced diamagnetic currents in the interior of the box are just canceled by the currents from electrons which bounce in cuspoidal paths off the walls. Thus, magnetism is inexplicable in classical physics and is a quantum phenomenon.

Free electrons. The diamagnetism of free electrons, which vanishes classically, was calculated quantum-mechanically by L. Landau. For particles obeying Fermi-Dirac statistics, such as the electrons in a metal, the numerical value of the Landau diamagnetism is exactly one-third of the Pauli spin paramagnetism.

Bound electrons. The diamagnetism of bound electrons in other than ionic crystals is difficult to calculate. In metals the diamagnetism is a sum of contributions from the nonconducting core electrons, for which the Langevin theory is adequate, and from the conduction electrons, for which the Landau theory must be modified to take account of the periodic potential from the ion cores. Metals in the bismuth group have unusually high diamagnetic susceptibilities (about -10^{-4} cm^3/mole) coming from conduction electrons. These metals also show the de Haas–van Alphen effect, a quasiperiodic variation of susceptibility when plotted aginst $1/H$ at low temperatures. The susceptibility may even oscillate between diamagnetism and paramagnetism.

Elihu Abrahams; Frederic Keffer

Bibliography. D. Craik, *Magnetism*, 1994; A. G. Morrish, *The Physical Principles of Magnetism*, 1965, reprint 1980; J. H. Van Vleck, *The Theory of Electric and Magnetic Susceptibilities*, 1932.

Diamond

A mineral composed entirely of carbon; the hardest substance known. Diamond is a polymorph of carbon; lonsdaleite, another polymorph, is sometimes referred to as hexagonal diamond. Diamond is found

on all continents except Antarctica, which has not yet been explored for it. It occurs in nature as single crystals of gem or industrial quality, and as polycrystalline masses referred to as boart, framesite, or carbonado. It has also been found as minute black grains in some meteorites. Diamond can be synthesized in the laboratory and is produced commercially in large amounts for industrial uses. *See* CARBON; GRAPHITE.

Physical and chemical properties. Diamond has a cubic (isometric) crystal structure in which all carbon atoms have covalent (sp^3) bonds. It is this strong bonding that makes diamond hard. Nevertheless, if diamond is struck in specific directions it will readily cleave—a property utilized in the preparation of polished gem diamonds. The combination of refractive index and dispersion gives diamond its brilliance and so-called fire when cut and polished. The thermal conductivity of diamond is the highest of any material (five times that of copper). This property, plus hardness, makes diamond an ideal material for use as a cutting tool in industry and also as a heat sink in electronics. *See* CRYSTAL STRUCTURE.

Although diamond consists of carbon, at least 58 other elements have been found (for example, aluminum, 10 parts per million; hydrogen, 1000 ppm; silicon, 80 ppm) as impurities in natural diamond. However, only two, nitrogen and boron, replace carbon atoms in the diamond lattice. Nitrogen is the major impurity and may substitute for carbon in a number of ways, commonly as either isolated or paired nitrogen atoms, and as discrete platelets of nitrogen within the diamond structure. The presence or absence of nitrogen and the manner of its substitution leads to different physical properties, such as thermal conductivity, electrical resistivity, and infrared spectra (see **table**). Consequently, diamond has been classified into four main types:

Type Ia diamonds contain appreciable nitrogen, often as planar defects (plates) in the crystal structure. They exhibit strong infrared absorption at 1282 cm^{-1} and no paramagnetism. Probably 98% of all natural diamonds are type Ia.

Type Ib diamonds contain less nitrogen than type Ia, and nitrogen occurs as substitutional atoms. Type Ib diamonds are paramagnetic. While type Ib diamonds are rare in nature, most synthetic diamonds are of this type.

Type IIa diamonds contain virtually no detectable nitrogen. These exhibit the highest conductivity of all diamond types. They are very rare in nature.

Type IIb diamonds contain no nitrogen, but boron is present. They are semiconducting, extremely rare in nature, and sometimes blue in color.

Diamond is resistant to chemical attack, other than by strong oxidizing agents. In vacuum or an inert atmosphere, a clear, colorless gem diamond transforms to a gray-black mass of graphite at about 1500°C (2700°F). In air, diamond oxidizes (burns) to carbon dioxide at and above 800°C (1500°F). At high temperature, some metals (for example, tungsten, titanium, and tantalum) react with diamond to form metal carbides. Metals, such as iron, nickel, cobalt, and platinum, in the molten state are solvents for carbon and dissolve diamond; this phenomenon is used as a basis for the synthesis of diamond.

Selected physical properties of diamond

Property	Value
Crystal structure	Cubic (isometric)
Space group	Fd3m
Unit cell size	$a = 3.57$ A (0.357 nm)
Density	3.52 g/cm^3(2.03 oz/in.^{3})
Hardness	10 (Mohs scale)
Refractive index	2.417 at 589 nm
Thermal expansion	0.8×10^{-6} at 20°C (68°F)
Thermal conductivity (20°C or 68°F)	
Type Ia	800 $W \cdot m^{-1} \cdot K^{-1}$
Type IIa	2100 $W \cdot m^{-1} \cdot K^{-1}$
Resistivity (20°C or 68°F)	
Type Ia and IIa	$>10^{14}$ Ωm
Type IIb	0.1–10 Ωm

Genesis and occurrence. Most natural diamond, apart from that in meteorites, crystallizes at depths of approximately 110 mi (180 km) in the Earth's upper mantle at temperatures in the range 900–1200°C (1650–2200°F). The host rock in which diamond forms is either a magnesian-rich silica-deficient ultramafic (peridotitic) rock or an ultrabasic eclogitic rock. Minerals that constitute the ultramafic type of diamond-host rock are magnesian rich and include varieties of olivine, pyroxene, and pyrope garnet. The eclogitic rock consists of sodium-bearing pyroxene and an almandine garnet. These various constituent minerals may also occur as inclusions in diamond and result in the host being identified as either an ultramafic (peridotitic) or eclogitic diamond. Diamonds in each group have formed in a distinct and different geochemical environment in the upper mantle. *See* ECLOGITE; LITHOSPHERE; METEORITE; PERIDOTITE.

Dating. Some inclusions in diamond can be characterized as having crystallized simultaneously with the host, and are said to be syngenetic with the diamond. Isotopic analysis of such inclusions enables the age of crystallization of the inclusion, and hence host diamond, to be determined. Some diamonds formed 3.2 billion years ago (the Earth is 4.6 billion years in age), whereas the youngest diamond studied by this method has been dated as 628 million years old.

Diamond is eventually transported to the Earth's surface by unique types of volcanic eruption in which gases play a major role. The eruptions drill narrow (much less than 3000 ft or 1000 m) explosive vents or pipes through the crust of the Earth. Two different rock types, each containing diamond, may result and infill the volcanic neck or pipe. The first type is known as kimberlite, and the second as lamproite. The most productive mine in the world based on the number of diamonds produced per unit of host rock is based on lamproite—the Argyle mine in Western Australia. In general, diamonds are considerably older than the volcanic eruption that transported them to the surface. Thus diamonds that are 3.2 billion years old reached the surface

only 85 million years ago. Diamond-bearing kimberlites occur in South Africa, Botswana, Angola, Sierra Leone, Guinea, Tanzania, Brazil, Venezuela, the United States, Canada, Russia, Siberia, China, India and Australia. Only lamproites in Western Australia and one in Arkansas in the United States are diamond bearing.

Alluvial deposits. Kimberlite and lamproite are weathered and eroded by normal geological processes, and the diamonds are released. Because diamond is hard, chemically inert, and dense, it is carried into and along the river drainage systems. Diamonds, together with other minerals, eventually become trapped in alluvial or placer deposits. The diamonds in these secondary deposits may have traveled hundreds of miles, over millions of years, from the primary-source kimberlite. Typical alluvial diamond-bearing deposits occur along the Vaal and Orange rivers in South Africa, and in Angola, Zaire, the Central African Republic, Ghana, Sierra Leone, and Guinea. In South America diamonds are mined from alluvial deposits in Brazil, Guyana, and Venezuela. The historic gem diamonds recorded between the twelfth and seventeenth centuries were obtained from alluvial deposits in India. China, Kalimantan, Russia, and Western Australia are other sources for alluvial diamonds.

A unique type of alluvial deposit is the diamond-bearing marine gravel off the west coast of South Africa and Namibia. The diamonds in these deposits were transported by the Orange River from the interior of South Africa to the coast, where they have been redistributed by wave action and other marine processes. Generally, transport in a high-energy environment, such as that in a river or surf zone, results in the diamonds with serious flaws, defects, and cracks being destroyed, with only the better-quality diamonds surviving. The ultimate manifestations of this process are the marine deposits where 95% of the diamonds are of gem quality. In contrast, in a primary-source kimberlite or lamproite, the percentage of gem diamonds is much less; for example, in the Argyle mine, Western Australia, only about 10% of the diamonds are of gem quality.

Meteorites. Minute blackish crystals and small irregular polycrystalline masses (less than 0.04 in. or 1 mm) of diamond have been recovered from certain meteorites, for example, the Canyon Diablo iron meteorite which formed Meteor Crater, Arizona. There are varying opinions as to whether or not the diamonds were formed as a result of collision of asteroid-size bodies in space or as a result of impact with the Earth's surface. The latter idea is generally held. The ultrahigh-pressure form of carbon, lonsdaleite (hexagonal diamond), is also found associated with meteorites and impact craters. Lonsdaleite is often intergrown with polycrystalline masses of diamond (carbonado or framesite).

Mining and recovery. The major production of diamond is from the primary sources in South Africa, Botswana, Zaire, Australia, and Siberia, with minor amounts coming from Tanzania and China. The diamond mines in all these countries are based on kimberlite, except for the Argyle mine in Australia, where the source rock is lamproite. Although Botswana and South Africa produce the most gem diamonds, as well as Russia whose production is difficult to assess, the major worldwide production is from the Argyle mine in Australia, albeit mostly industrial diamonds. Diamond production from secondary (alluvial or placer) deposits, apart from the extensive mining of the marine gravels off the west coast of southern Africa, is relatively small compared to the output from mines based on kimberlite and lamproite. Alluvial deposits in Guinea, Ghana, Russia, and Australia are mined by large companies or government agencies. All other alluvial diamond deposits are worked by small local groups or individual miners. *See* PLACER MINING.

In small alluvial deposits diamonds are recovered by manual methods, which include sieving, jigging, and then searching the residue by eye. In the larger, more mechanized alluvial deposits, techniques similar to those of the major primary-source diamond mines are used. Diamond mines may be either open-pit or underground, although normally a new primary deposit is first mined as an open pit and later by underground development. Only the mines in South Africa (the Premier, Finsch, and Kimberley mines) are underground mines, although some mines in Siberia have been converted from open-pit to underground activity.

On average, diamond is 1 part per 20 million parts of waste rock. In diamond mining a value denoting the number of carats per 100 metric tons of rock is used to indicate the grade. At Kimberley the grade is about 80 carats/100 metric tons; at Mbuji Mayi, Zaire, the grade is roughly 200 carats/100 metric tons; at the Argyle mine in Western Australia, grades up to 400 carats/100 metric tons have been recorded. Nevertheless, the economics of diamond mining does not rest solely upon grade; factors such as the size and quality of the diamonds are equally important. *See* CARAT.

The mined rock is first crushed and screened to various sizes and then passed through heavy-media separation equipment. The intention is to separate the light fraction (less than 1.6 oz/in.^{3} or 2.7 g/cm^3), which is waste, from the heavy minerals, including diamond. This heavy fraction is referred to as the concentrate. Prior to the 1970s, much of this primary separation was done in rotary washing pans. In these pans a thick muddy liquid is kept at a density of about 0.41 oz/in.^{3} (1.25 g/cm^3). The lighter material floats to the surface, whereas diamond and other heavy minerals concentrate at the bottom of the pan and are periodically removed. Modern heavy-media separation equipment works on a similar principle, but it has more efficient control of the separation media.

The concentrate then flows with water over a vibrating table, or conveyor belt, covered with grease. Diamond has an affinity for grease and sticks to it, whereas other minerals pass over the grease with the water and are rejected. Although grease tables are used in some mines or as secondary recovery

methods, most modern plants use a method based on x-rays to separate diamonds from the waste material in the concentrate. When diamonds are irradiated by x-rays they luminesce, whereas most other minerals in the concentrate do not. The concentrate, in a narrow stream, passes through a beam of x-rays. A light-sensitive module senses the luminescence of the diamond and sends an impulse, usually to an air jet, to blow the diamond out of the stream and into a sealed container. This method is over 97% reliable and is referred to as the sortex method. Both heavy-media separation and sortex recovery may be aided by rotary pans, heavy-media cyclones, grease tables, electrostatic separators, and skin flotation. The last method is usually used to recover diamonds smaller than 0.04 in. (1 mm) in diameter. After recovery the rough diamonds are cleaned in acid and sorted for size, shape, quality, and color.

Cutting and polishing. Rough diamonds occur in a variety of shapes, including octahedra, dodecahedra (**Fig. 1**), twinned octahedra (macle), and broken or cleavage fragments. The largest diamond found, the Cullinan, was a cleavage fragment. After the rough is sorted into cuttable (gem and near gem) and industrial stones, the decision is made as to how a specific diamond will be shaped and made into a polished gem. The cutting and polishing process can result in the loss of as much as 60% of the original diamond.

The initial step is to alter the shape of the rough diamond so that it approximates more closely that of the finished gem. This may be done by cleaving or sawing. Prior to either of these actions being taken, the diamond cutter will determine the grain of the stone. The grain reflects specific crystallographic directions in the diamond along which cleaving and polishing can be more readily accomplished. The direction of the grain is often observed from natural lines or patterns on various crystal faces. In cleaving a stone, the cleaver will follow the octahedral grain. A small nick, known as the kerf, is made along the direction of cleavage in the diamond by scratching another diamond along the direction to be cleaved. A thin steel blade is then placed in the kerf and tapped; if done correctly the diamond cleaves perfectly.

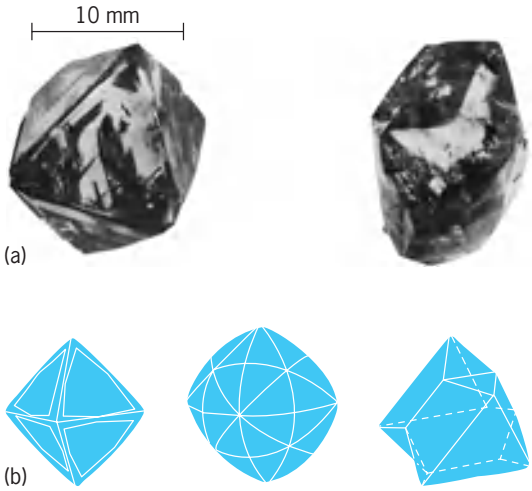

Fig. 1. Diamond. (*a*) Crystals from Kimberley, South Africa (*American Museum of Natural History specimens*). (*b*) Crystal habits (*after C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*).

Instead of cleaving, it may be necessary to saw a diamond. The saw consists of a thin phosphor bronze disk whose edge is charged with paste of fine diamond powder and olive oil. The diamond is mounted in a holder and placed on the upper part of the revolving cutting disk. Depending upon the size and nature of the diamond, it may take hours or even days to complete the sawing. Instead of sawing, in modern practice many diamonds are sliced with a laser beam. This technique has the advantage of speed, but more importantly it allows the cutting of stones with irregular areas of crystal growth (naats or knots) that are difficult or impossible to saw.

The next step in the process is bruting. The bruter (also cutler or cutter) uses a second diamond as a tool in order to round off the corners and edges of a rough stone or of a cleaved or sawn diamond. The aim is to prepare the stone to a rounded shape upon which the facets of the final gem can be ground.

The last stages are grinding and polishing facets on the diamond. These stages are done on a scaife, a horizontal cast-iron wheel dressed with a fine diamond powder and olive oil. The diamond is placed in a dop, a holder that can be adjusted to enable different faces to be ground and polished. Grinding is the initial coarse preparation of the facets, which are then polished with a finer grain size of diamond paste. When completed, a brilliant-cut diamond, the most common shape, consists of 58 symmetrical facets of different sizes and relative angles. Automation is increasingly used in the polishing industry, and there are a number of machines available that can be programmed to automatically grind and polish the facets simultaneously on several diamonds. For the most part, the automation is used on smaller stones; the larger and more important are polished by hand. Prior to about 1940, most cutting and polishing was done in Amsterdam and Antwerp. Since then diamond polishing centers have opened in London, New York, and Tel Aviv and in Russia, India, southern Africa, and China and other localities in the Far East.

Polished diamonds are graded on the basis of the 4 C's—carat, cut, clarity, and color. The carat is the unit of weight in the diamond industry and is standardized as 0.2 gram (0.0071 oz or 200 milligrams) and is divided into 100 points. Thus a 10-point diamond weighs 0.1 ct (0.00071 oz or 20 mg). The largest diamond, the Cullinan, weighed about 3000 ct (4.27 oz or 600 g) and was the size of an average human fist. The grading cut is based on how well the facets and the shape of a polished diamond compare to a standard model.

In some instances, in order to save weight or to hide blemishes, the diamond may be cut and polished to be thinner or thicker than the standard proportions. The clarity of a polished diamond depends upon the presence or absence of defects such as inclusions, cleavages, and other natural flaws.

A diamond with no such defects and having top clarity is referred to as being flawless.

Diamonds, although commonly considered to be mostly colorless, actually exist in all the colors of the rainbow. Colored stones are known as fancies, and if of excellent uniform color they are most desirable. Perhaps the most famous colored diamond is the deep-blue Hope diamond (44.5 ct), displayed in the museum of the Smithsonian Institution. The color of diamonds can be artificially changed by various types of irradiation; for example, neutrons, protons, electrons, alpha particles, gamma particles, and radioactive isotopes. Irradiation, followed by heat treatment, can turn a pale yellow, relatively worthless diamond into a pale to dark green stone, which if it were a natural color would be a prized stone. Unfortunately, it is not always possible to distinguish between an enhanced and naturally colored diamond. *See* GEM.

Synthesis. Diamonds were first synthesized in Sweden in 1953. However, the information was kept secret until after the General Electric Company in the United States was issued a patent. These early experiments used the principle that carbon dissolves in the transition elements of groups 8–10 (such as iron or nickel). At high pressures (50–60 kilobars or 5–6 gigapascals) and temperatures (1500°C or 2700°F), the dissolved carbon nucleates and crystallizes as diamond. Direct conversion of graphite to diamond was achieved in 1961 in shock-wave experiments in which transient high pressures in excess of 300 kb (30 GPa) and temperatures of about 1100°C (2000°F) existed. A similar natural phenomenon is believed to be the origin of diamonds in meteorites.

Unlike imitation diamonds, which consist of nondiamond material such as glass, cubic zirconia, or strontium titanate, synthetic diamond is identical with the natural mineral. The properties (for example, crystal structure or density) are similar, although minor differences do exist because of the different processes of synthetic and natural growth. Thus synthetic diamonds often have a pale yellow color due to the incorporation of atmospheric nitrogen in the synthesis chamber; consequently most synthetic diamonds are type Ib. Synthetic diamonds may also contain metallic inclusions that represent the flux from which they were grown; common are tiny inclusions of nickel. Furthermore, although diamonds can be synthesized with a variety of shapes, most are cubo-octahedral, a form virtually absent in natural stones. Synthetic diamonds generally are not large; most are produced in sizes below 0.004 in. (0.1 mm). These are used extensively as grit for industrial grinding purposes. Colorless gem-quality diamond can also be synthesized, but the cost of synthesis has proved to be greater than the cost of the natural product. Synthetic gem crystals greater than 11 carats have been manufactured. It is possible to differentiate between such synthetic and natural diamonds.

In contrast to growing diamond at high pressures and temperatures, diamond can also be synthesized at pressures below 1 atm (10^5 pascals) and at temperatures of around 1000°C (1900°F). In this process, carbon is deposited from a carbon-rich vapor as thin polycrystalline films of diamond on a substrate of diamond, silicon, glass, or some other suitable material. The initial experiments for this type of synthesis date from about 1962, but detailed work in the Soviet Union in the late 1960s led to intensive research in Japan and then worldwide. Diamond films can be grown in several ways. For example, diamond crystals up to 0.02 in. (0.5 mm) in size can be formed from a mixture of methane and hydrogen at about 50 torr (6.7 kPa) pressure and 1000°C (1900°F) on a silicon substrate. This method is known as thermally induced chemical vapor deposition, but other techniques may be used, including plasma chemical vapor ion deposition and electron-beam deposition. The diamond films display properties similar to those of natural diamond and have similar hardness and thermal conductivity, both significant properties for the uses of diamond films.

Applications. Diamond, apart from its use as a gem, has numerous applications in industry, and it is designated a strategic mineral. Many of the uses of natural and synthetic diamond are equivalent. Originally, natural diamond, including boart, carbonado, and framesite, was crushed to various sizes of powder and used as grinding and polishing agents for glasses, ceramics, and nonferrous metals. Diamonds, as single crystals or powders, are also bonded in metal drills and bits. Small drills are used in applications such as dental work; large drills are used in drilling for oil and other minerals. Diamond-impregnated wheels are used for cutting many hard materials, including concrete and dimension stone for architectural purposes. Synthetic diamond is sometimes preferred for various uses, as it is grown to the specific grain size rather than crushed as in the case of natural diamond. Crushing results in inherent cracks or fractures in the individual grains that decrease the life use of the material.

Diamonds are used in eye surgery, and also as heat sinks and semiconductors in the electronics industry. Diamond films (**Fig. 2**) have potential uses as scratchproof coatings on optical lenses, compact

Fig. 2. Scanning electron micrograph of a diamond film deposited by microwave-induced plasma, showing the crystalline morphology.

discs, and even on nondiamond jewelry; bearings in machines; heat sinks and semiconductors in electronics; and general inert coatings or surfaces in areas of high chemical corrosion. Natural diamond has also been used as optical windows in spacecraft.

Henry O. A. Meyer

Bibliography. E. Bruton, *Diamonds*, 2d ed., 1978; G. Davies, *Diamond*, 1984; J. E. Field (ed.), *The Properties of Natural and Synthetic Diamond*, 1992; G. E. Harlow (ed.), *The Nature of Diamonds*, 1997; J. Wilks and E. M. Wilks, *Properties and Applications of Diamonds*, 3d ed., 1991.

Diamond thin films

Very thin coatings that are applied to other materials to provide them with special physical or chemical properties unique to diamond. While crystalline diamond is extremely expensive, many of the useful properties of diamond can be obtained in a cost-effective manner by coating a substrate of interest with a polycrystalline film consisting of tiny diamond crystals that form a continuous film, often 1 micrometer or less in thickness. Diamond thin films can be grown on a variety of thicker materials that provide mechanical support and possibly electrical or thermal contact.

Fabrication. Diamond and graphite are different forms (allotropes) of pure carbon. Because graphite is the more thermodynamically stable form, growth of diamond requires special procedures. The formation of diamond thin films consists of two distinct steps: nucleation and growth. Nucleation is the initial formation of very small diamond crystallites. Growth refers to the set of chemical and physical processes that make these crystallites grow larger and ultimately form a continuous film consisting of many tiny, interlocked crystals (**Fig. 1**). The initial nucleation of diamond must usually be enhanced. Commonly used methods include "seeding" the growth mechanically by abrading the substrate with very fine diamond powder or by ion bombardment. *See* DIAMOND; GRAPHITE.

Growth of diamond films is typically performed via a chemical vapor deposition process in which methyl radicals ($CH_3\cdot$) are first created by decomposing methane (CH_4) gas, removing one hydrogen atom and its associated electron ($H\cdot$) from a molecule of methane via $CH_4 \rightarrow CH_3\cdot + H\cdot$. The methyl radicals then deposit on the solid surface. While both diamond and graphite can be formed, the selective growth of diamond is achieved by also having a large excess of hydrogen present; hydrogen atoms preferentially "etch away" any graphite that forms, leaving behind a pure diamond film. The energy needed to break apart the CH_4 and H_2 molecules can come from several different sources. One common way is to use microwaves to ignite a plasma in a mixture of methane and hydrogen (typically ~5% methane in 95% hydrogen). The deposition of methyl radicals onto the surface leads to growth of tiny diamond crystals, which grow in size and ultimately form a continuous film consisting of a tightly interlocked network of diamond crystals. The boundaries between these crystals have some influence on the properties, and the size and shape of the crystals can be controlled during growth to produce films in which specific properties are optimized (**Fig. 2**). Diamond growth rates are slow, typically 1–10 μm/h for high-quality films; faster growth rates lead to higher incorporation of graphite into the film and a degradation of most properties. *See* CHEMICAL VAPOR DEPOSITION; PLASMA (PHYSICS).

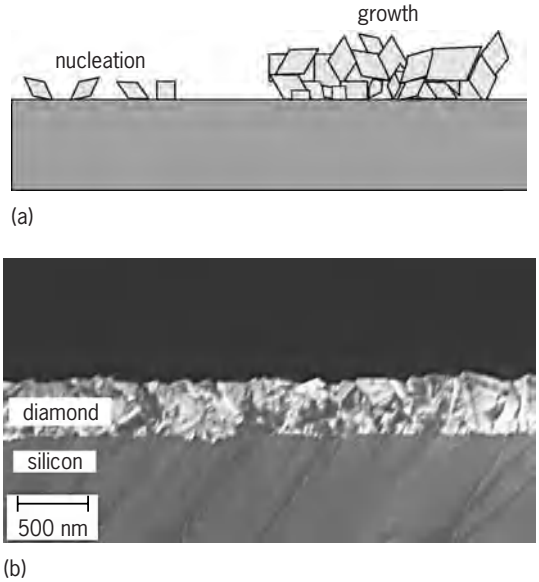

Fig. 1. Diamond thin films. (*a*) Nucleation and growth of a diamond thin film. (*b*) Scanning electron microscope side view image of a 400-nm-thick diamond thin film grown on a silicon substrate. (*Wensha Yang, University of Wisconsin-Madison*)

Properties. Much of the interest in diamond stems from its many extraordinary physical and chemical properties.

Extraordinary hardness. Diamond is the hardest known naturally occurring material.

Low friction. Diamond surfaces exhibit quite low friction.

High thermal conductivity. At room temperature, diamond conducts heat five times better than copper.

Optical transparency. Diamond transmits a wide range of wavelengths, including microwave, infrared, and ultraviolet, making it useful for specialized optics for satellites, spacecraft, and specialized measurements.

High index of refraction. The high index of refraction of 2.41 enhances reflections and contributes to diamond's sparkle; for example, in jewelry.

Electrical properties. Diamond is an insulator but can be "doped" (for example, with boron atoms) during the growth process to give it semiconducting properties.

Chemical inertness. Diamond is stable in aggressive chemical and physical environments.

Biocompatibility. Diamond and diamondlike films are used as coatings on biomedical implants.

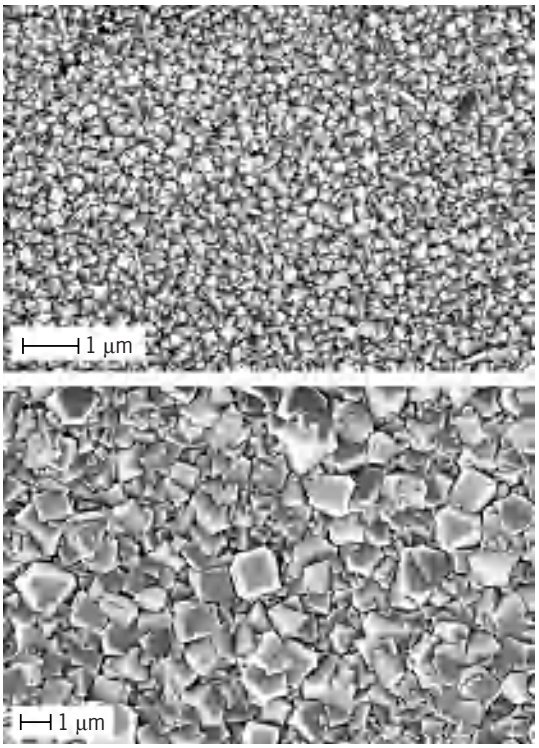

Fig. 2. Scanning electron microscope images of two diamond films, showing the individual diamond crystallites that agglomerate into a continuous film. (*a*) Very small diamond grains. (*b*) Film with larger diamond crystals. (*Bin Sun, University of Wisconsin-Madison*)

Applications. The extreme mechanical, optical, and chemical properties of diamond make it an attractive material for use in a variety of technological applications. The ability to grow diamond as a thin film provides a way to take advantage of these properties in a practical, cost-effective manner. Diamond is very useful as a thin coating on machine tools such as drills, saws, and other cutting tools. Like silicon, diamond's electrical conductivity can be changed at will, making it potentially useful for specialized applications in electronics. Diamond's high thermal conductivity makes it useful for carrying heat away from high-power electrical devices such as diode lasers. In addition, diamond has been receiving much attention because of its chemical stability, making it an attractive substrate for applications such as environmental monitoring and biological sensing.

Robert J. Hamers

Bibliography. R. E. Clausing et al., *Diamond and Diamond-like Films and Coatings*, 1991; A. H. Lettington and J. W. Steed (eds.), *Thin Film Diamond*, 1994; C. Nebel (ed.), *Diamond Thin Film II*, 2004; C. Nebel and J. Ristein (eds.), *Diamond Thin Films I*, 2003.

Diapensiales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Dilleniidae of the class Magnoliopsida (dicotyledons). The order has only a single family, the Diapensiaceae, with about 18 species of herbs and dwarf shrubs in temperate and arctic regions of the Northern Hemisphere. Within its subclass the order is marked by sympetalous flowers (flowers with the petals joined to each other by their margins, at least toward the base, forming a basal tube, cup, or saucer), separate pollen grains, and ovules that have a single integument (unitegmic) and are tenuinucellate (with the nucellus consisting of a single layer of cells covering the micropylar end of the embryo sac). *Shortia* and *Galax* genera are sometimes cultivated. *See* DILLENIIDAE; FLOWER; MAGNOLIOPHYTA; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Diapir

A buoyant mass of ductile rock or sediment that has pierced, or appears to have pierced, overlying rock, known as overburden. The overburden can yield by ductile processes or by brittle faulting (see **illus.**). Diapirs form by lateral and vertical intrusion of buoyant or nonbuoyant rock. Emplacement of a diapir, known as diapirism, involves piercement of overlying rocks.

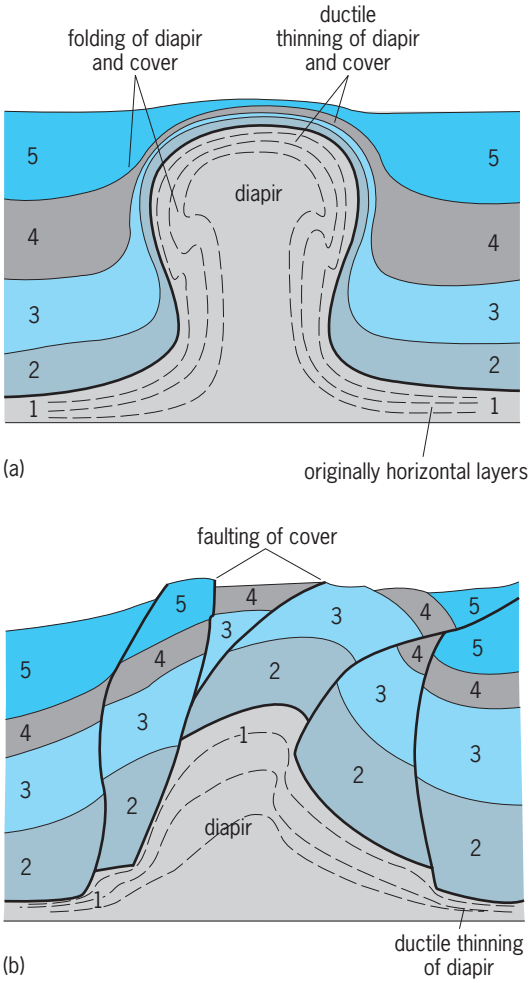

Vertical sections showing diapirism by (*a*) ductile folding of cover and (*b*) faulting of cover. Number 1 indicates the bedding of the diapir, and 2–5 identify cover layers.

Diapirs are composed of salt, gypsum, shale, mud, sand, peat, coal, limestone, ice, serpentinite, granite, gneiss, and migmatite. Salt diapirs are typically several miles wide and high, but sheetlike varieties that have spread or coalesced laterally can be as much as 200 mi (300 km) wide. This type of diapir is economically important. Gneiss and migmatite diapirs are typically 6–13 mi (10–20 km) wide but locally approach 60 mi (100 km) in width. Diapirs doming the sea floor form large islands of salt in the Persian Gulf and islands of shale in the Caspian and Banda seas. *See* GNEISS; GRANITE; MIGMATITE.

Over geologic time, solid rock can flow slowly by creep. Like glacial ice, salt can flow at the surface, but for gneiss to creep requires burial to a depth of at least 9 mi (15 km), where temperatures of 1100°F (600°C) have softened the rock. Diapirism is driven by gravitational potential energy released by the sinking of denser overburden around the diapir. Either the overburden becomes denser by compaction or metamorphism, or the diapiric rock becomes less dense by heating or infusion of water. Hot diapirs rise as part of convection systems. Even diapirs denser than their overburden can rise in grabens or in the cores of folds formed by forces stretching or compressing the crust because of plate motions or the effects of gravity on sloping continental margins. Linear loads can cause diapirs to be shaped like irregular walls, instead of plugs, which are more common. Where updomed roof strata are eroded, salt diapirs can extrude through the eroded orifice. Such diapirs grow higher by remaining near the sedimentation surface while their bases sink with the subsiding basin floor, a process known as downbuilding. *See* GRABEN; SALT DOME. M. P. A. Jackson

Bibliography. M. P. A. Jackson et al., Salt Diapirs of the Great Kavir, Central Iran, *Geol. Soc. Amer. Mem.*, no. 177, 1990; M. P. A. Jackson, D. G. Roberts, and S. Snelson (eds.), Salt Tectonics: A Global Perspective, *Amer. Ass. Petrol. Geol. Mem.*, no. 65, 1995; H. Ramberg, *Gravity, Deformation and the Earth's Crust*, 2d ed., 1981; J. G. Ramsay, Emplacement kinematics of a granite diapir: The Chindamora Batholith, Zimbabwe, *J. Struct. Geol.*, 11:191–209, 1989.

Diapsida

A group of reptiles characterized by two pairs of lateral skull openings (fenestrae) and a suborbital fenestra in the palate below the eye. This group includes living and extinct lizards, snakes, tuatara, and crocodiles, as well as a range of extinct Late Paleozoic and Mesozoic lineages including rhynchosaurs, protorosaurs, younginiforms, choristoderes, dinosaurs, and pterosaurs. In addition, most workers would now include the so-called parapsid or euryapsid reptiles (defined as possessing a single upper-skull fenestra; **Fig. 1***c*), with the view that the ancestral lower temporal fenestra was secondarily reduced or closed in the aquatic nothosaurs, plesiosaurs, pliosaurs, ichthyosaurs, and placodonts. Perhaps more surprisingly, some authors also consider the anapsid (no skull fenestrae) turtles to be diapsid derivatives, and this idea, while still contentious, has received support from many molecular phylogenetic studies, which use deoxyribonucleic acid (DNA) sequence data to determine evolutionary relationships. Birds evolved from bipedal theropod dinosaurs and are also part of the diapsid radiation, although older, traditional classifications place them in a separate class, Aves. *See* ANAPSIDA; ANIMAL EVOLUTION; AVES; DINOSAUR.

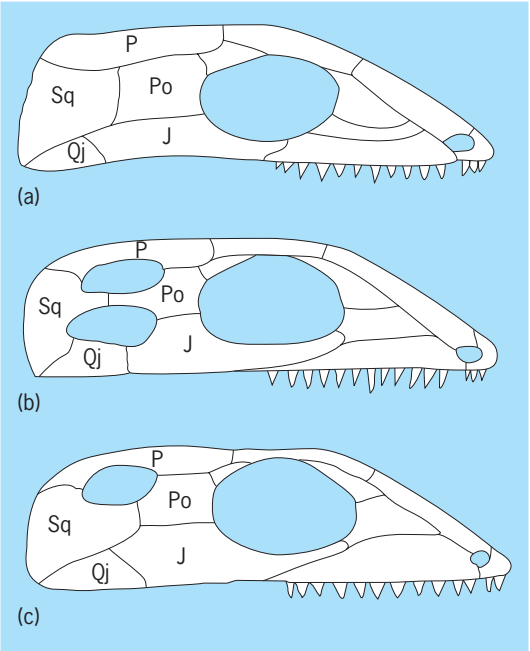

Fig. 1. Series of three skulls in lateral view: (*a*) anapsid skull condition, (*b*) diapsid skull condition, and (*c*) parapsid/euryapsid skull condition. J, jugal; P, parietal; Po, postorbital; Qj, quadratojugal; Sq, squamosal.

Skull morphology. The upper and lower fenestrae that characterize diapsids lie in the temporal region of the skull between the orbit and the ear region. This region houses the main jaw adductor muscles. The two openings arise in the position of three- and four-point junctions between skull bones in the ancestral anapsid skull (Fig. 1*a*), rather like the fontanelles in the neonatal human skull. Thus the upper fenestra usually lies between the parietal, postorbital, and squamosal bones (Fig. 1*b*), and the lower fenestra lies between the postorbital, jugal, quadratojugal (where present), and squamosal bones (Fig. 1*b*). This leaves a postorbital bar between the fenestrae and the orbit, an upper temporal bar between the upper and lower fenestrae, and a lower temporal bar below the lower fenestra. Classically the lower temporal bar is complete, but it is missing or incomplete in many diapsid lineages, either through multiple losses (for example, squamates, rhynchocephalians, protorosaurs) or as a primitive state, with early loss being followed by the independent redevelopment of a complete bar on several occasions (Tuatara, archosaurs). In snakes and some lizards, the upper temporal bar and postorbital bars are also lost as an adaptation to skull

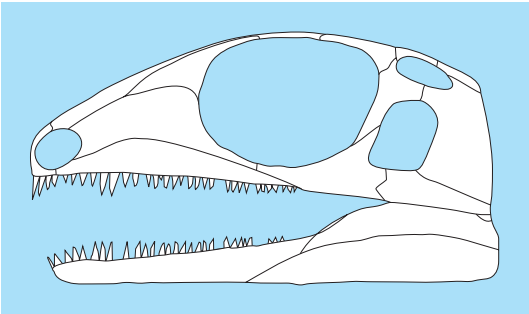

Fig. 2. Lateral view of the skull of the basal diapsid *Petrolacosaurus*. (*After R. R. Reisz, A diapsid reptile from the Pennsylvanian of Kansas, Univ. Kansas Publ. Mus. Nat. Hist., 7:1–74, 1981*)

kinesis (the ability of parts of the skull to move or flex relative to each other).

Evolution. The earliest adequately documented diapsid is *Petrolacosaurus* from the Upper Carboniferous (Pennsylvanian) of Kansas [305 million years ago (mya); **Fig. 2**]. It was lightly built and about 8 in. (20 cm) in length, with a relatively long neck and long distal limb elements. Apart from these differences in proportions and the presence of the diagnostic temporal openings, the skeleton of *Petrolacosaurus* strongly resembles that of more primitive amniotes such as *Hylonomus* (Upper Carboniferous, Canada; 310 mya). The Lower Permian *Araeoscelis*, also from North America (Texas and Oklahoma, 280 mya), is now considered to be a close relative of *Petrolacosaurus*, although it was originally classified as a parapsid due to the absence of a lower temporal fenestra (probably secondarily closed).

Diapsids remain poorly documented throughout the Late Paleozoic, an era dominated by the synapsid ancestors of mammals. Nonetheless, Upper Permian and Lower Triassic rocks chart the early stages of a major radiation. With the decline of the nonmammalian synapsids beginning at the end of the Permian and accelerating through the Triassic, diapsids diversified and came to dominate Mesozoic ecosystems—land, sea, and air. They remain highly successful today although they have been somewhat eclipsed by the second major synapsid radiation (of modern mammals). Some Permian and early Mesozoic lineages, including coelurosauravids (Upper Permian gliders from Europe and Madagascar, 255 mya), younginiforms (small terrestrial and aquatic reptiles from the Late Permian and Triassic of Africa and Madagascar, 255–245 mya), and thalattosaurs (long-snouted marine reptiles from the Triassic of North America, Europe, and China, 235–225 mya) may have branched off the diapsid stem in the early stages of the radiation. Others are more difficult to place. Specialization toward an aquatic lifestyle affects characters of the limbs, girdles, and vertebral column that would otherwise be important for classification. As a result, the phylogenetic positions of groups such as the placodonts, plesiosaurs, choristoderes, and ichthyosaurs remain problematic. *See* SYNAPSIDA.

The remaining advanced diapsids are grouped into one of two major clades, Lepidosauromorpha and Archosauromorpha, represented in the modern fauna by the lepidosaurs (lizards, snakes, amphisbaenians, tuatara) on the one hand and by the archosaurs (crocodiles and birds) on the other (**Fig. 3**). The two major diapsid groups are distinguished by skull morphology and by divergent specializations of the vertebral column, limbs, and girdles affecting posture and locomotion. *See* ICHTHYOPTERYGIA; PLACODONTIA; PLESIOSAURIA.

Lepidosauromorpha. Most lepidosauromorphs have a lightly built skull with an incomplete lower temporal bar, large eyes, and an ear region structured to support a large eardrum. They retain an essentially sprawling posture and rely on lateral undulation of the trunk in locomotion, but they developed specialized vertebral joints for added stability, and more sophisticated knee, ankle, and foot joints to improve movement over a variety of substrates. For small lepidosauromorphs, the retention of a sprawling posture was not a serious handicap, and most can run and climb on a variety of surfaces. Large-bodied lepidosauromorphs are rare and most have either been aquatic (for example, the Late Cretaceous marine mosasaurs, 93–65 mya) or secretive and slow moving (for example, heavy constricting snakes such as the python and anaconda). The exceptions are the monitor lizards of the family Varanidae, a terrestrial lineage that has repeatedly achieved large size (for example, the living Indonesian Komodo dragon and the extinct Australian *Megalania*, about 1.6 mya–40,000 years ago), but in regions devoid of large mammalian predators.

Basal lepidosauromorphs probably separated from the ancestors of archosaurs in the Early or Middle Permian, but they remain poorly represented in Early Mesozoic deposits. The record of basal lepidosauromorphs (that is, those classified outside Lepidosauria) includes the fragmentary Early Triassic *Paliguana* from South Africa (250 mya), as yet unnamed fossils from the Early Triassic of Poland, and the highly specialized long-ribbed gliding kuehneosaurs (*Kuehneosaurus*, *Icarosaurus*, 210–200 mya) from the Upper Triassic of Britain and North America. In the past, both *Paliguana* and kuehneosaurs were considered to be early lizards, but this was based on superficial features such as small size and a lightly built skull. True lepidosaurs are linked by a large suite of characters. The ancestors of lizards and snakes (Squamata) apparently diverged from those of tuatara (Rhynchocephalia) early in the Triassic but they do not appear in the fossil record until later. The first rhynchocephalians are recorded from the beginning of the Late Triassic (Germany, 220 mya), but the derived skull and dental characters of these early representatives show that they had a long unrecorded history before this. Similarly, although squamates are first known from the Early/Middle Jurassic (India, Britain, Central Asia, 160 mya), their level of diversity at that time is indicative of a pre-Jurassic history. *See* LEPIDOSAURIA; SQUAMATA.

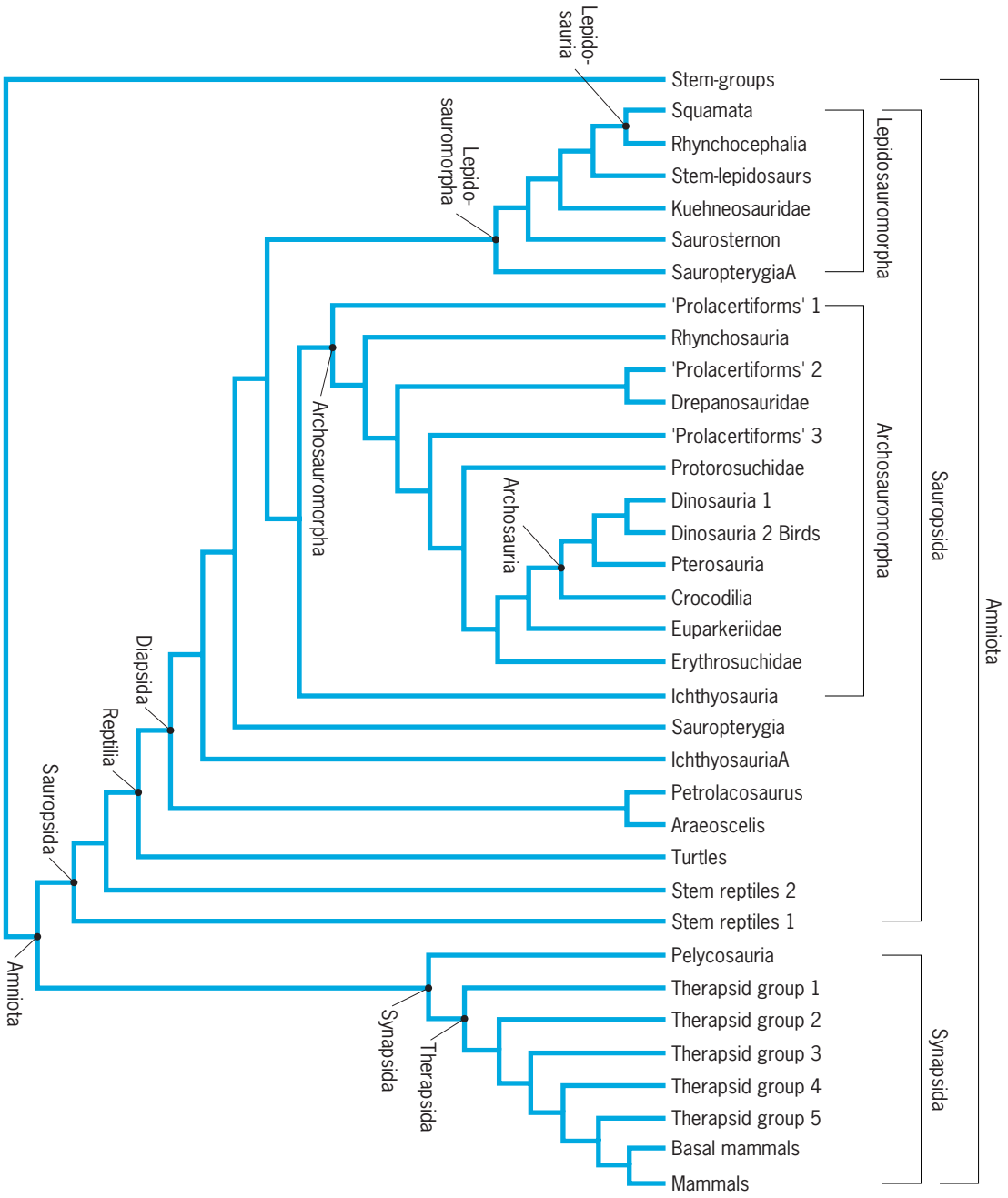

Fig. 3. Phylogenetic tree showing relationships of the main groups of amniotes, including the primary division into Sauropsida and Synapsida, and the division of Diapsida into Lepidosauromorpha and Archosauromorpha. Sauropterygia and Ichthyosauria each appear in two places in the tree because there is some debate as to their relationships. Therapsid numbers 1, 2, etc., denote different therapsid groups; however, these are purely arbitrary numbers, as the tree is really much more complex with many more branches at each level. The same applies to stem reptiles 1 and 2.

Archosauromorpha. Primitive archosauromorphs were widely distributed during the Upper Permian and Triassic and include the long-necked protorosaurs (also known as prolacertiforms), the apparently herbivorous rhynchosaurs and trilophosaurs, the arboreal or swimming drepanosaurs (Late Triassic, Italy and North America, 225 mya), and perhaps the gliding *Sharovipteryx* (Upper Triassic, Kyrgyzstan, 225 mya). Primitive archosauromorphs such as the protorosaurs and rhynchosaurs were essentially improved sprawlers, but derived genera brought the limbs under the body and coupled a more rigid trunk region with the development of a parasagittal gait. This reached its fullest expression in dinosaurs and birds that resemble mammals in having an energetically efficient upright stance and a vertebral column that flexes primarily in the dorsoventral plane. Inevitably, this allowed an increase in body size, as seen in some of the massive herbivorous dinosaurs of the Jurassic and Cretaceous. *See* ARCHOSAURIA; PROTOROSAURIA; RHYNCHOSAURIA.

More derived, archosaur relatives (grouped with

Archosauria in the more inclusive group Archosauriformes) include the primitive protorosuchians (Upper Permian to Lower Triassic, 255–245 mya, South Africa, Australia, Russia), the large erythrosuchids (Early to Middle Triassic, 250–230 mya, South Africa, China, Russia), and the more gracile euparkeriids (Early Triassic, 250 mya, South Africa, Russia, Poland).

Within Archosauria, the lineage leading to crocodiles was distinct by the Middle Triassic, as were the immediate ancestors of the dinosaurs and the pterosaurs (Fig. 3). Dinosaurs began to radiate during the Upper Triassic, and by the Upper Jurassic most of the major groups (including birds) had appeared. At the end of the Mesozoic, all the remaining nonavian dinosaurs became extinct, along with the pterosaurs, mosasaurs, and plesiosaurs. Crocodiles, most lizards, snakes, tuatara and birds all survived (as did turtles, amphibians, and many mammals). *See* REPTILIA.

Susan E. Evans

Bibliography. M. J. Benton, *Vertebrate Palaeontology*, 3d ed., Blackwell, Oxford, 2004; M. J. Benton (ed.), *The Phylogeny and Classification of the Tetrapods*, vol. 1: *Amphibians, Reptiles and Birds*, Clarendon Press, Oxford, 1988; R. L. Carroll, *Vertebrate Paleontology and Evolution*, Freeman, New York, 1997; A. S. Romer, *Osteology of the Reptiles*, University of Chicago Press, 1956.

Diarrhea

The passage of loose or watery stools (greater than 8 oz or 200 g per 24 h), usually at more frequent than normal intervals. The two principal mechanisms responsible for diarrhea are the rise in osmotic pressure of fecal contents, diminishing the absorption of water by the intestine (osmotic diarrhea); and the increase in electrolyte secretion into the intestinal lumen by its epithelial cells, leading to an increased flow of water from the intestinal mucosa into the lumen (secretory diarrhea).

Pathophysiology. Ingestion of salts containing magnesium, sulfate, or phosphate ions can bring about osmotic diarrhea since these ions are poorly absorbed by epithelial cells of the intestinal lumen; consequently, water absorption is diminished. Similarly, diarrhea may also result from the impaired absorption of disaccharides or monosaccharides. For example, individuals with a deficiency of the intestinal disaccharidase enzyme lactase have an impaired ability to split dietary lactose into glucose and galactose. As a result, such individuals can have watery diarrhea after ingesting milk, since lactose molecules remain in the intestinal lumen and inhibit water absorption.

Just as the increased transport of electrolytes and water into the intestinal lumen can bring about secretory diarrhea, the diminished absorption of electrolytes and water can also cause it and contribute to clinical manifestations in patients with secretory diarrhea. In some diseases that are characterized by secretory diarrhea (for example, cholera), fecal water can exceed several quarts per 24 h. Cholera is caused by the bacterium *Vibrio cholerae*, which colonizes the small intestine and produces a protein toxin responsible for secretory diarrhea. These toxic molecules stimulate the production of the enzyme adenylate cyclase in the epithelial cells of the small intestine, which leads to the conversion of adenosine triphosphate (ATP) to cyclic adenosine monophosphate (AMP). Cyclic AMP induces secretion of chloride ions and inhibits absorption of sodium ions by epithelial cells of the small intestine. These events lead to water secretion, impaired water absorption, and profuse watery diarrhea. Strains of *Escherichia coli* that cause the diarrhea common to travelers produce a toxin having a similar mode of action to that of cholera toxin. *See* CHOLERA.

Bacillary dysentery, which also is characterized by secretory diarrhea, is caused by bacteria of the genus *Shigella*, which invade epithelial cells of the colon and release a toxin that kills the cells by blocking their ability to synthesize proteins. Individuals with bacillary dysentery have ulcerations of the colon and excrete blood and mucus by way of the rectum. *See* BACILLARY DYSENTERY; FOOD POISONING.

Agents responsible for viral gastroenteritis include rotavirus and Norwalk virus. These viruses cause watery diarrhea, probably by colonizing and damaging epithelial cells of the small intestine, impairing their ability to absorb electrolytes and water. Protozoan parasites, including *Cryptosporidium* species and *Giardia lamblia*, cause diarrhea by mechanisms that are not well understood. *Cryptosporidium* infestation leads to profuse watery diarrhea in patients with acquired immune deficiency syndrome (AIDS). *See* MEDICAL PARASITOLOGY.

Impaired absorption of bile salts or long-chain fatty acids also causes secretory diarrhea. Nonabsorbed bile salts are deconjugated to free bile acids by bacteria in the colon, and colonic bacteria also hydroxylate nonabsorbed long-chain fatty acids. The bile acids and hydroxy fatty acids stimulate colonic secretion of electrolytes and water, which accounts for the secretory diarrhea.

Crohn's disease and ulcerative colitis are characterized by intestinal inflammation and diarrhea, with or without rectal bleeding. In these conditions diarrhea is due, at least partly, to impaired absorption of electrolytes and water, as a result of intestinal epithelial cell damage; however, the underlying cause is unknown. Diarrhea is also a feature of irritable bowel syndrome, which is a condition in which intestinal motility is increased and often appears to be a physical manifestation of anxiety or depression.

Identification, management, and treatment. If a bacterial infection or parasitic infestation is suspected, fecal microscopy and culture should be performed. Microscopy may reveal parasites (for example, *G. lamblia* cysts or *Cryptosporidium* oocysts), and culture may identify a bacterial pathogen. Where there is chronic diarrhea of unknown cause, some of the following tests may be informative: fecal fat

measurement (to test for malabsorption of fat in the intestine); endoscopic examination of the colon (to explore for a tumor or ulcerative colitis); barium radiography of the colon or small intestine, and mucosal biopsy of the small intestine (to identify diseases causing malabsorption); and examination of a fecal specimen for evidence of laxatives that a patient may surreptitiously be ingesting (for example, magnesium ions or phenolphthalein, present in many laxatives).

Management of diarrhea includes rehydration, and treatment of the causative disease. If a bacterial or protozoan pathogen can be identified as the cause, the person may benefit from treatment with an antimicrobial agent. However, not all bacterial infections of the intestine are eliminated by antimicrobial treatment, even if the bacterium is sensitive to the antimicrobial agent in the laboratory. Bismuth subsalicylate and drugs that reduce intestinal motility (for example, opiates and loperamide) are useful agents for the symptomatic treatment of diarrhea. *See* GASTROINTESTINAL TRACT DISORDERS.

Martin F. Heyworth

Bibliography. M. Field et al. (eds.), *Secretory Diarrhea*, 1980; P. Miskovitz and A. Rochwarger, *Evaluation and Treatment of Patients with Diarrhea*, 1993.

Diastem

A break in the stratigraphic record produced by local erosion or nondeposition and representing a short interval of geologic time. The breaks in deposition are those that would occur within a particular sedimentary environment, rather than those associated with a major change in environment. *See* STRATIGRAPHY; UNCONFORMITY.

Diastems are not distinctive or even identifiable features. Their presence was deduced by J. Barrell in 1917. He recognized that deposition in modern environments is not continuous, so that many short intervals of geologic time must go unrecorded. Although the rate of deposition of an individual bed may be rapid, the presence of diastems results in a slow rate of accumulation for the stratigraphic sequence as a whole. Thus the aggregate time represented by strata in a given sequence may be much less than the time encompassed by the sequence.

Nondeposition can result either from excessive turbulence in the environment or from a lack in sediment supply. Discontinuity in sedimentation occurs on many scales. Short breaks, from seconds to days, are associated with migration of bedforms, variation in wave or current energy, and tidal cyclicity. Seasonal deposition occurs due to floods, storms, and plankton blooms. Sedimentation by major floods or hurricanes may occur on a scale of decades to centuries. In the deep ocean, the interval between successive turbidity current flows can be thousands of years. *See* MARINE SEDIMENTS; SEDIMENTOLOGY; TURBIDITY CURRENT.

Charles W. Byers

Diazotization

The process by which an aromatic primary amine is converted to a diazonium compound. The preparation and reactions of diazonium salts were discovered in 1858 and were the basis of the synthetic dye industry and the development of other industrial chemistry in Europe. In diazotization, sodium nitrite is added to a solution of the amine in aqueous acid solution at 0–5°C (32–41°F). Reaction of the amine with nitrous acid gives a nitrosamine. Tautomerization and loss of water lead to the diazonium ion, which is stabilized by delocalization of the positive charge at the ortho and para carbon atoms of the ring [reaction (1)].

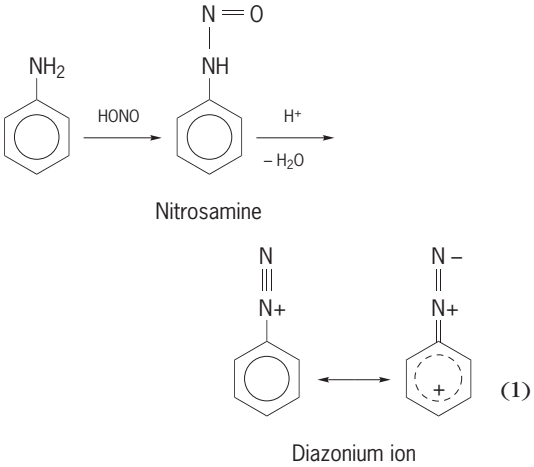

(1)

See AMINE; AROMATIC HYDROCARBON; DELOCALIZATION; TAUTOMERISM.

The overall reaction is simple and very general. Substituents of all types—alkyl, halogen, nitro, hydroxyl, sulfonic acid—can be present at any position. Heterocyclic amines such as aminothiazole or aminopyridines can also be diazotized. Aromatic diamines are converted to *bis*-diazonium compounds. Diazonium salts are generally used and handled in aqueous solution; they are explosive if isolated and dried.

Coupling reactions. Diazonium ions are mild and selective electrophiles, similar in chemical properties to the nitrosonium ion (NO^+), and they readily and reversibly form covalent bonds to nucleophiles. Reactions with amines lead to diazoamino compounds (triazenes) [reaction (2*a*), where Ar represents the

$$\mathrm{Ar{\equiv}N^+ + RNH_2 \rightleftharpoons \underset{\text{Triazene}}{ArN{=}N{-}NHR} + H^+} \qquad (2a)$$

$$\mathrm{ArN{\equiv}N^+ + OH^- \rightleftharpoons \underset{\text{Diazohydroxide}}{[ArN{=}N{-}OH]} \xrightarrow{OH^-} \underset{\text{Diazotate}}{ArN{=}N{-}O^-}} \qquad (2b)$$

aryl group]. With cyanide ion, diazocyanides are

formed. Addition of an azide salt to a diazonium solution at −20°C (−4°F) gives a diazoazide, which cyclizes to the very unstable arylpentazole; the latter decomposes on warming to an aryl azide plus nitrogen. Reaction with hydroxide ion occurs when a diazonium solution is made basic. Since the diazohydroxide is a stronger acid (in the Brönsted sense) than the diazonium ion (Lewis acid), two equivalents of hydroxide are consumed and the product is the diazotate [reaction (2*b*)]. *See* ACID AND BASE; ELECTROPHILIC AND NUCLEOPHILIC REAGENTS.

Diazonium salts also react with enols or enolizable carbonyl compounds. The products are hydrazones, which are formed by tautomerization of the azo intermediates [reaction (3)]

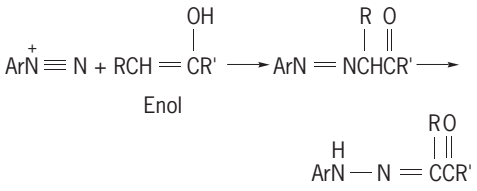

(3)

See REACTIVE INTERMEDIATES.

The great importance of diazonium compounds in dye technology lies in the coupling reactions that occur with an activated aromatic ring, such as that in phenols or aromatic amines. Coupling, or electrophilic substitution by ArN_2^+, gives compounds with an arylazo group at the position para or ortho to —OH or —NH_2. Reaction with amines occurs in weak acid solution. With phenols the phenoxide ion is the reactive species, and slightly basic solution is used. These coupling reactions occur concurrently with the equilibria of reactions (2*a*) and (2*b*). A complex balance of reaction rates, acidities, and concentrations is required, but it is possible to prepare compounds in which two different azo groups are introduced, such as the naphthol dye, with the structure below.

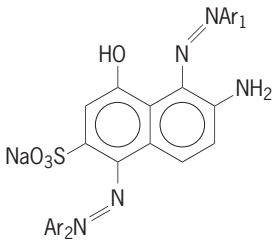

See CHEMICAL EQUILIBRIUM.

The azo dyes obtained in these coupling reactions are one of the important types of synthetic dyes. The color of the dye can be varied widely by choice of diazonium and coupling components. The coupling reaction lends itself to an important method of applying the dye to fabrics. In this process the coupling reagent, such as a naphtholsulfonic acid, is absorbed onto the fiber, and the coupling reaction is then carried out directly on the fiber by passing the fabric through a bath of the diazonium solution. *See* DYE; DYEING; TEXTILE CHEMISTRY.

Substitution reactions. Diazonium salts are central to the synthesis of a number of aromatic compounds. When a solution of a diazonium salt in water is heated, nitrogen is evolved and the resulting aryl cation is converted to the corresponding phenol. Similarly, aromatic halides and nitriles can be synthesized by treatment of the diazonium salt with an appropriate source of halide and cyanide ions (Sandmeyer reaction). Displacement of nitrogen by hydrogen can be accomplished with hypophosphorous acid. *See* PHENOL.

Diazonium salts are also useful for introducing an aryl group by substitution. Biaryls are formed in a free-radical reaction by heating a neutralized diazonium solution with an aromatic hydrocarbon. Alkenes with an electron-withdrawing group can also be arylated. *See* FREE RADICAL; ORGANIC SYNTHESIS; SUBSTITUTION REACTION.

Diazo compounds. In contrast to the aromatic series, diazotization of an alkyl amine gives an extremely unstable and short-lived diazonium ion. Because there is no stabilizing delocalization of positive charge, decomposition occurs to give a carbenium ion and a nitrogen molecule; the reaction products result from the carbocation. However, the diazo group can be retained in an aliphatic or alicyclic molecule if a proton is removed to give the system $R_2C{=}N_2$. These diazo compounds are useful and versatile reagents, and several methods for their preparation are available.

Simple diazoalkanes are prepared by alkaline hydrolysis of a *N*-alkyl-*N*-nitroso amide, sulfonamide, or urea [reaction (4*a*), where X = OR or NH_2]. Another method for introducing a diazo group is direct transfer of the group from a sulfonyl azide to an enolate, resulting in a diazoketone [reaction (4*b*)]; this re-

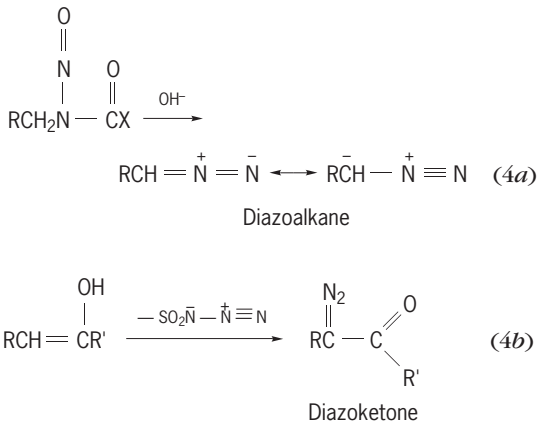

(4*a*)

(4*b*)

action brings out the similarity of the isoelectronic azido and diazoalkyl groups.

Diazoalkanes are good nucleophiles and react readily with electrophiles such as a proton or carbonyl group. With a carboxylic acid, diazomethane is protonated and the diazonium ion immediately reacts with the anion to give the methyl ester

[reaction (5)]. With a cyclic ketone the major reac-

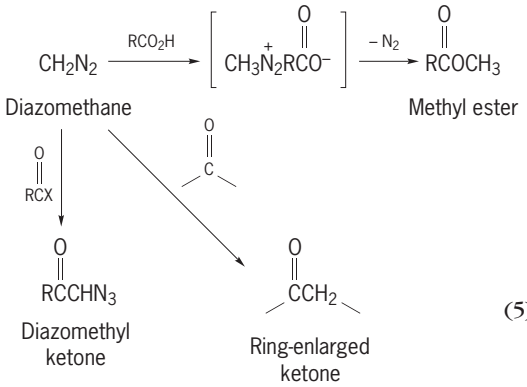

(5)

tion is ring enlargement. Acylation of diazomethane gives the diazomethyl ketone.

Diazo compounds and azides on heating or photolysis lose nitrogen to give carbenes and nitrenes, respectively. These are neutral, electron-deficient, highly reactive intermediates that can add to double bonds and also attack C-C or C-H bonds with formation of insertion products. Loss of nitrogen from the acyl derivatives, namely, diazoketones and acyl azides, is accompanied by rearrangement, probably by a concerted pathway rather than a carbene or nitrene, to ketene [reaction (6)] or isocyanate, respectively.

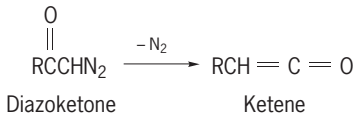

(6)

A very useful reaction of diazoketones or diazoesters involves metal-catalyzed C-C bond formation. Rhodium carboxylates such as dirhodium tetraacetate [$Rh_2(OAc)_4$] are particularly effective catalysts. In these processes the reactivity of the acyl carbene is attenuated by coordination with the metal, and stereospecific cyclizations are possible [reaction (7)].

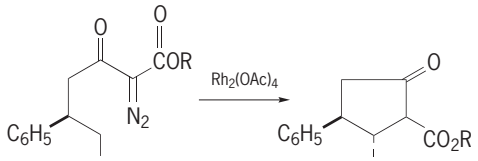

See COORDINATION CHEMISTRY; REACTIVE INTERMEDIATES.

James A. Moore

Bibliography. *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., vol. 3, 1992; S. Patai (ed.), *Chemistry of Diazonium and Diazo Groups*, 1978; M. Regitz and G. Maas, *Diazo Compounds: Properties and Syntheses,* 1986.

Dichroism

In certain anisotropic materials, the property of having different absorption coefficients for light polarized in different directions.

There are few natural materials which exhibit strong dichroism. One of the first to be discovered was tourmaline. Light transmitted by thin plates of dark forms of tourmaline is almost completely polarized. *See* POLARIZED LIGHT.

In isotropic optical materials the optical density is defined as in Eq. (1), where I_0 is the intensity of the

$$d = \log \frac{I_0}{I} \tag{1}$$

incident light, and I that of the transmitted light. In anisotropic materials that are dichroic, the value of d can vary as a function of the vibration direction of the electric vector of the light wave. Just as the index ellipsoid is used to define the birefringence of a material, a density surface can be used to define the dichroism. *See* CRYSTAL OPTICS.

Compared to the literature on birefringence and optical activity, there has been relatively little material on dichroism. This is partly because of the difficulty in making measurements. The Kramers-Kronig relationship shows that any material whose refractive index is different from unity and varies as a function of wavelength will absorb radiation at some wavelength. From the Kramers-Kronig relationship it is apparent that all optically anisotropic materials should be dichroic. From the values of the refractive index at different wavelengths, the spectral positions and intensity of the absorption can be calculated. In a linear birefringent material the refractive index depends on the polarizing direction or electric vector of the radiation. For each direction of propagation there are two perpendicular vibration directions with different refractive indices. It can be inferred that, at some wavelength, the absorption for light vibrating in these two directions will be different. For transparent materials this wavelength is in the ultraviolet and the absorption is thus difficult to measure. The absorption difference is also frequently so small that it cannot be detected. In other words, the absorption is apparently the same in each direction. Furthermore, the dichroic band may coincide with a region of isotropic absorption.

If the absorption in a dichroic material is different for different linear states of polarization, the material is termed linear dichroic. If it is different for right and left circularly polarized light, it is termed circular dichroic. Similarly, there can be elliptically dichroic crystals. In biaxial crystals there are three different refractive indices corresponding to an electric vector lying along each of the three orthogonal axes of the index ellipsoid. Such a crystal has dichroism which is different for light traveling along each of these three principal axes.

Most dichroic materials exist in the form of relatively thin sheets. Here one is dealing with a section of the dichroic surface. Associated with the sheet of material will be a direction of maximum absorption and one of minimum absorption. The density equation can be rewritten as Eqs. (2), where the densities

$$d_{\parallel} = \log \frac{I_{0\parallel}}{I_{\parallel}} \qquad d_{\perp} = \log \frac{I_{0\perp}}{I_{\perp}} \tag{2}$$

are for light vibrating parallel or perpendicular to the axis of the section of dichroic surface. *See* TRICHROISM. Bruce H. Billings

Dichroism (biology)

In absorbing, anisotropic objects the refractive index and the absorption vary with direction. As with double refraction (birefringence), there are two main directions of anisotropy, which are perpendicular to each other. This phenomenon is called dichroism and may be studied in the polarizing microscope with one polarizer. By rotating the microscope stage the object is brought in different directions to the vibration plane of the polarized light, and the change of light absorption is observed. Study of dichroism allows conclusions as to the submicroscopic fine structure of cells.

Use of stains. In visible light only a few cellular components, such as chloroplasts, show absorption. An absorption can, however, be produced by staining. The dichroic staining of plant fibers is especially simple. The walls of such cells consist of cellulose microfibrils 3–30 nanometers thick. The elongate stain particles of benzidine dyes, for example, congo red, are deposited in an oriented manner in the spaces between the microfibrils and produce an intrinsic dichroism of the fiber: colored for a vibration plane parallel, colorless for a plane perpendicular to the stain particles and fibrils. Therefore, the direction of strongest absorption indicates the course (parallel or helical) of the microfibrils in the fiber. Similar results are obtained by staining with iodine by means of the chlorzinc iodine reagent. Also, some metals, such as silver and gold, give dichroic staining; they are deposited as isotropic crystals. The dichroism is form dichroism (analogous to form birefringence), that is, produced only by the distribution of the crystals. Dichroic staining may also be carried out with animal tissue. Such studies have been made mainly with fibrous protein structures (collagenous fibers, myofibrils, neurofibrils) and with anisotropic nuclei of certain sperms.

Ultraviolet dichroism. In ultraviolet light many cell components show absorption. Ultraviolet dichroism gives direct information as to the orientation of the absorbing molecules or molecular groups in cell structures. The method has been especially helpful for studies of orientation of deoxyribonucleic acid in nuclei and chromosomes. The maximum absorption here is perpendicular to the molecular direction. Dichroism in elongate nuclei of certain sperms shows corresponding behavior, which means that in these objects the deoxyribonucleic acid molecules must be parallel to the long axis. In normal nuclei and in chromosomes there is no or only a very weak ultraviolet dichroism; therefore the deoxyribonucleic acid molecules can have no appreciable orientation.

Lignified plant cell walls show ultraviolet dichroism. It is pure form dichroism, a fact which eliminates the possibility that lignin is in an anisotropic state in the wall. *See* CELL WALLS (PLANT).

By irradiation with ultraviolet light, various compounds of the cell are caused to fluoresce. The fluorescent light is polarized if the object is anisotropic. This phenomenon, called difluorescence, is observable in lignified cell walls, and leads to the same conclusions as to lignin deposition as emerge from dichroism studies. *See* DICHROISM. Fritz Ruch

Bibliography. S. F. Mason, *Molecular Optical Activity and the Chiral Dimensions*, 1982; K. Nakanish, N. Berova, and R. Woody, *Circular Dichroism: Principles and Applications*, 1994; F. Ruch, Birefringence and dichroism of cells and tissues, in A. W. Pollister (ed.), *Physical Techniques in Biological Research*, vol. 3A, 1966.

Dicotyledons

A large group of flowering plants (angiosperms) that for many years has been considered one of the two main categories of plants, the other being monocotyledons. Dicotyledons have two seedling leaves as opposed to the single one in most monocotyledons. Several deoxyribonucleic acid (DNA) sequence studies subsequently demonstrated that there are two groups of angiosperms, but these correspond not to the number of seed leaves but to the two major pollen types. Thus, the term "dicotyledon" is no longer meaningful because some plants of this type are more closely related to monocotyledons. The group of former dicotyledons, which have pollen with a single aperture, includes magnolia, avocado, black pepper, and pipeworts; they are now termed magnoliids and include monocotyledons. The other category of dicotyledons, those with three (and often more) apertures in their pollen, are called eudicotyledons (true dicotyledons). *See* AVOCADO; EUDICOTYLEDONS; FLOWER; MAGNOLIOPHYTA; MONOCOTYLEDONS; PLANT KINGDOM. Mark W. Chase

Dicranales

A large and diverse order of the true mosses (Bryopsida), containing 13 families. Plants in the Dicranales generally have long, narrow leaves and deeply forked peristome teeth. The plants are erect and simple or merely forked and often woolly because of a dense covering of rhizoids among the leaves. The leaves, inserted in numerous rows, or rarely only two or three are often turned to one side or, when dry, crisped and curled. They are long-tapered, with a long, single costa and short or elongate cells which are sometimes bulging or papillose. The cells at the basal angles are often conspicuously differentiated. The capsules, terminal and generally exserted, may be erect and symmetric or inclined and asymmetric and often furrowed. An operculum is generally differentiated. The peristome (rarely lacking) commonly consists of 16 papillose or vertically pitted-striate teeth. The spores are spherical or rarely tetrahedral. The

calyptra is cucullate or less commonly mitrate. Chromosome numbers 12 and 14 are common, but multiples of 10, 11, 13, and 15 are also known. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA. Howard Crum

Dicyemida (Rhombozoa)

A phylum (also known as Rhombozoa) comprising approximately 100 described species of microscopic parasites that live in the renal sacs (kidneys) of many species of octopuses, some cuttlefishes, and a few squids.

Life cycle. The dicyemid life cycle is complex and consists of asexual and sexual stages (**Fig. 1**).

Nematogen. When a renal sac of a parasitized cephalopod is opened, or when fluid is withdrawn from the sac by suction, the organisms that emerge are in the asexual nematogen stage. Nematogens are slender and usually about 1 or 2 mm long with an outer layer of up to about 40 ciliated epidermal cells enclosing a large axial cell (**Fig. 2**). The two tiers of ciliated cells at the anterior end form the calotte (Fig. 2). This is inserted into a tubule or pocket in the renal sac, and the short cilia of the calotte cells interweave with microvilli of the epithelium. Nutrients are absorbed from the urine. The free surfaces of most epidermal cells have convoluted ridges that increase the absorptive area, and there are also pinocytotic vesicles where uptake of nutrients may take place.

The axial cell encloses numerous cells called axoblasts (Figs. 2 and **Fig. 3***a*, *b*) which are progeny of one or two cells that entered the axial cell early in the development of the nematogen. The axoblasts develop into daughter nematogens (Figs. 2, 3*a*, *b*), which escape from the parent nematogen and find a vacant site in the renal sac to attach themselves and asexually produce more nematogens.

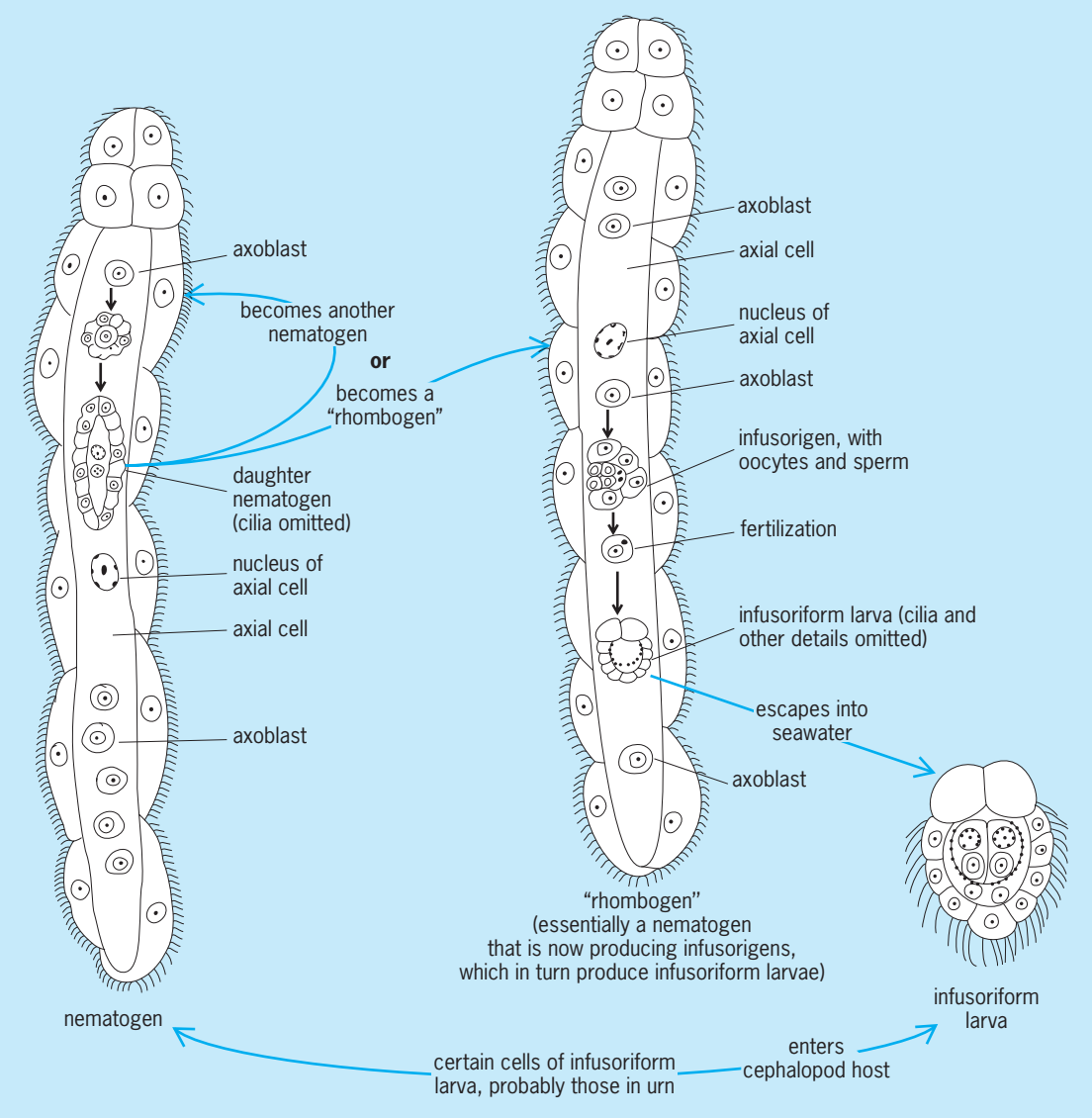

Fig. 1. Diagram of the life cycle of a dicyemid.

Infusorigen. After a nematogen has given rise to many daughter nematogens, this phase of reproduction slows down, and many of the axoblasts begin to develop into infusorigens (Figs. 3*c* and **Fig. 4***a*). These are structurally simple sexual individuals that consist, when fully developed, of a few oocytes and cells that give rise to sperm (which are much smaller than the oocytes and are atypical in that they lack a flagellum). After an oocyte has been entered by a sperm, it undergoes meiosis, casting off polar bodies, and becomes a fertilizable egg, or ovum. The egg nucleus then unites with the sperm nucleus.

Infusoriform larva. A fertilized egg develops into an infusoriform larva (Figs. 3*c* and 4*b*), so called because it resembles an "infusorian" (ciliated protozoan.) In spite of its small size (usually about 50 μm long), the organism at this stage is rather complex. It contains about 37 or 39 cells, two of which are proportionately large and filled mostly by a refringent mass of sodium inositol hexaphosphate. These cells are covered by a thin cap consisting of two cells that lack cilia but may extend posteriorly for a short distance behind the large cells they enclose. The rest of the outer covering of the infusoriform larva consists of ciliated cells. These enclose a few cells, some of which are organized into a structure called the urn, the cavity of which is partly lined by ciliated cells and partly by cells that are not ciliated; certain of the latter contain granules of glycogen. Within the cavity of the urn are four cells, each of which contains one or two nuclei and another complete cell.

A nematogen that has begun to produce infusoriform larvae is often called a rhombogen (hence the alternative name for the phylum), because the larvae have a somewhat rhomboid shape. However, a rhombogen is just an older nematogen, and it does not directly produce infusoriform larvae; these are products of sexual reproduction by an infusorigen. Infusorigens are sometimes regarded as bisexual gonads of a nematogen, but because they originate from axoblast cells they are, like daughter nematogens, essentially separate individuals.

Infusoriform larvae, after escaping from a nematogen in which they have developed, are active swimmers and will survive outside a host for several days. They are believed to represent the infective stage, but it is not known what part of an infusoriform larva becomes transformed into a nematogen. The cells within the urn, each of which contains one or two nuclei as well as another complete nucleated cell, seem like good prospects; the inner cell could conceivably develop into an axial cell and become covered by ciliated epidermal cells derived from the outer portion. This idea, however, is purely speculative.

Classification. There are two families of dicyemids, Dicyemidae and Conocyemidae. Nearly all known species and the two largest genera (*Dicyema* and *Dicyemennea*) belong to the Dicyemidae. In Dicyemidae, the epidermis of the nematogen is cellular and ciliated and the calotte is typically well developed, whereas in Conocyemidae, the epidermis (at least in part) is not divided into cells and not ciliated, and the calotte is not well developed.

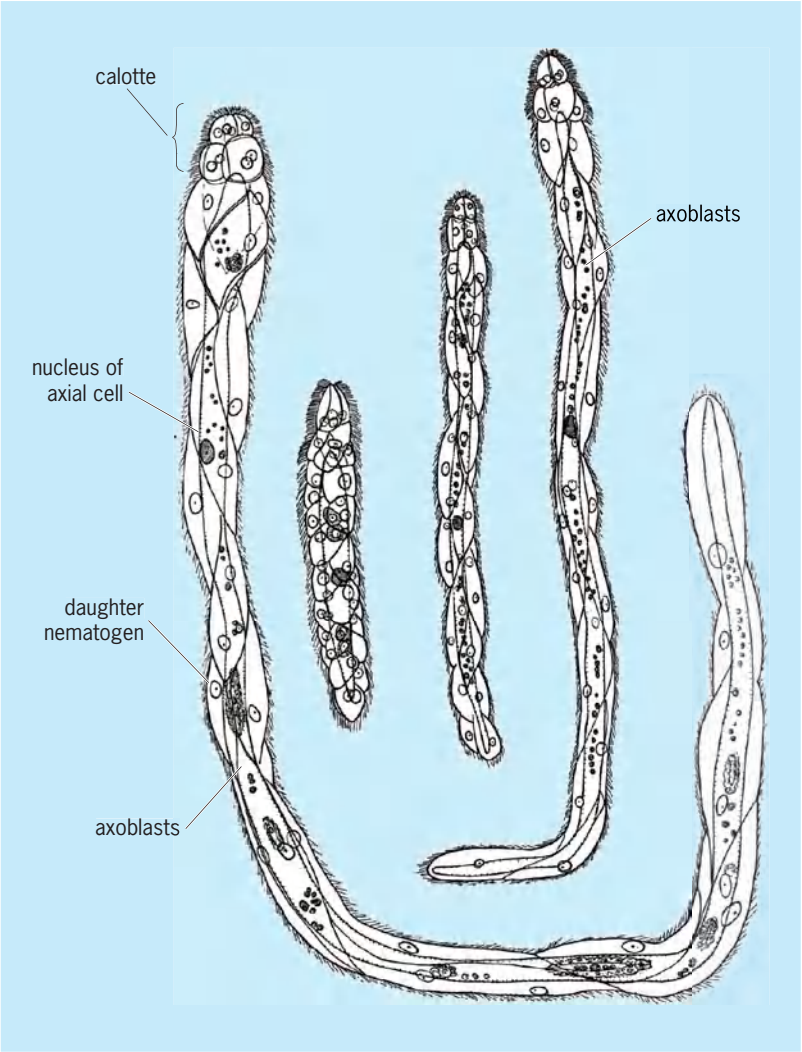

Fig. 2. ***Dicyema*, nematogen stage. The largest specimen contains daughter nematogens. (*From B. McConnaughey, Dicyema sullivani, a new mesozoan from Lower California, J. Parasitol., 35:124, 1949*)**

Dicyemids were discovered before the middle of the nineteenth century. The earliest serious research on them was published in 1876 by E. van Beneden, who placed them in a category called "Mesozoa" because of his belief that they were intermediate in structure between certain protozoans (especially ciliates) and other "less structurally advanced" invertebrates, such as flatworms. Orthonectida, first reported by Giard in 1877, are very different from dicyemids in structure and life history, but their relative simplicity led to their also being called mesozoans. For a long time, in fact, Mesozoa was a phylum name applied to the two groups. This view should be abandoned, however, because orthonectids and dicyemids are probably not closely related. Either or both groups could conceivably have evolved from free-living ancestral invertebrates or from parasitic groups. Both possibilities seem to be more

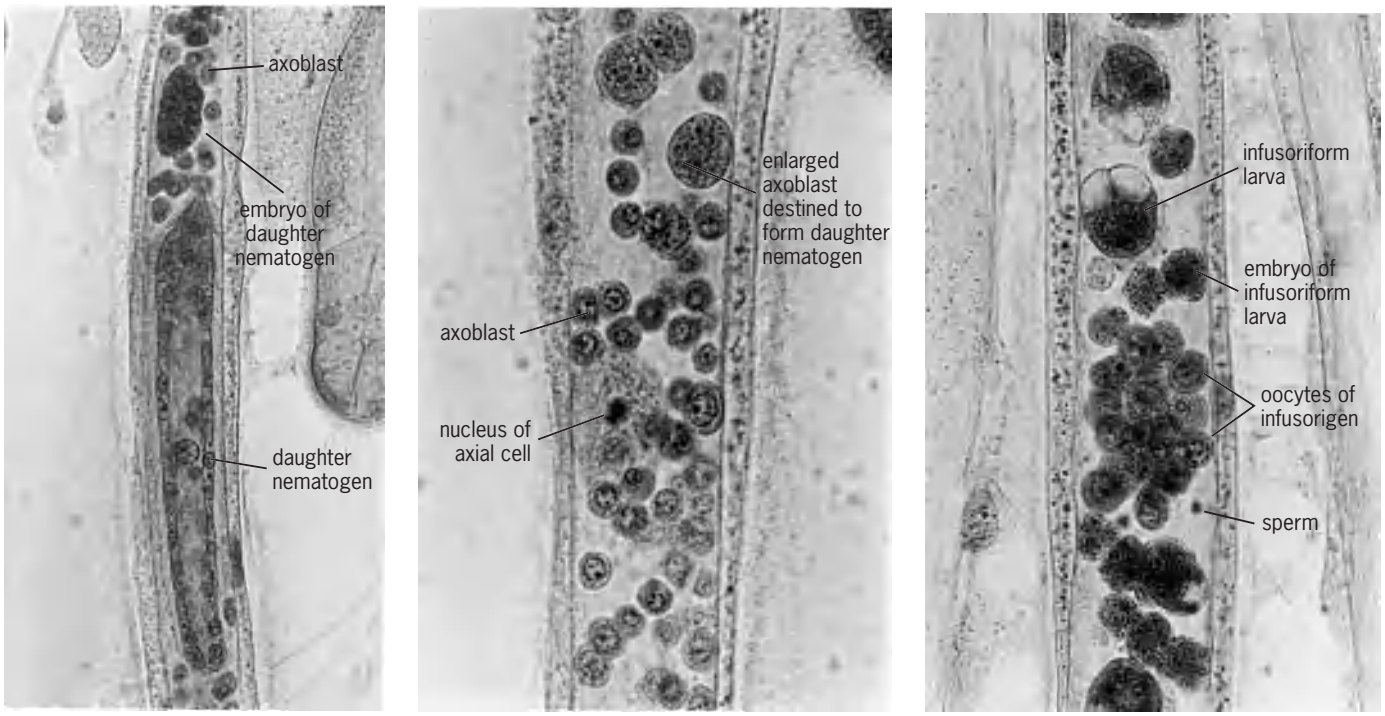

(a) (b) (c)

Fig. 3. Photomicrographs of *Dicyemennea*. (*a*) Portion of a nematogen with a daughter nematogen, already ciliated, in its axial cell. (*b*) Portion of an older nematogen whose enlarged axoblasts are destined to develop into infusorigens. (*c*) An infusorigen and stages in the development of an infusoriform larva.

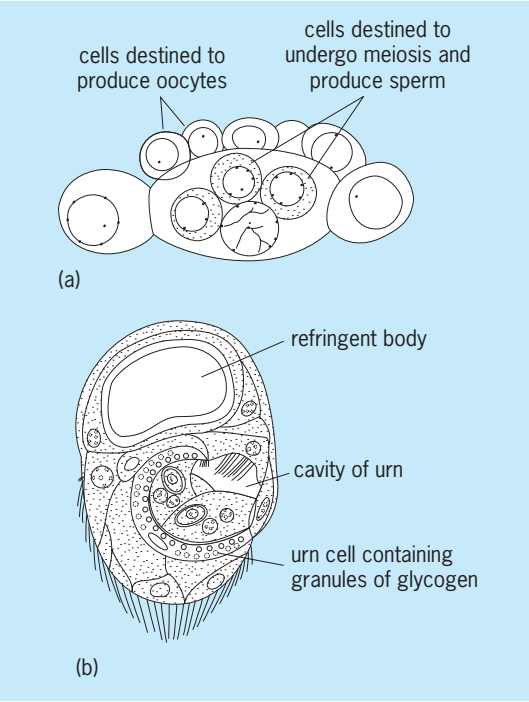

Fig. 4. Early life cycle forms. (*a*) *Dicyema*; young infusorigen with cells destined to produce oocytes and those destined to produce sperm. (*b*) *Dicyemennea*; infusoriform larva, as seen in a sagittal section close to the midline; drawing based on electron micrographs of sections.

popular with specialists than the idea that either dicyemids or orthonectids are derived from protozoa such as ciliates. *See* MESOZOA; ORTHONECTIDA.

Eugene N. Kozloff

Bibliography. H. Furuya, Phylum Dicyemida and Orthonectida, in C. Young (ed.), *Atlas of Marine Invertebrate Larvae*, pp. 149–161, Academic Press, London, 2002; H. Furuya and K. Tsuneki, Biology of dicyemid mesozoans, *Zool. Sci.*, 20:519–532, 2003; F. G. Hochberg, The parasites of cephalopods: A review, *Mem. Nat. Mus. Victoria*, 44:109–145, 1983; B. H. McConnaughey, The life cycle of the dicyemid Mesozoa, *Univ. Calif. Publ. Zool.*, 55:295–336, 1951.

Didymelales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Hamamelidae of the class Magnoliopsida (dicotyledons). The order contains a single family with but one genus, *Didymeles*, with two species. These dioecious, evergreen trees are restricted to Madagascar. The wood has vessels with scalariform perforations, which is a putatively primitive feature. The leaves are alternate, simple, and entire. The flowers are small; the staminate ones are in open clusters, without perianth, with two stamens, and the pollen grains have three germinal furrows. The pistillate flowers are commonly paired in spikelike clusters; they are hypogynous and sometimes have 1–4 scalelike sepals but no petals. The pistil has but one carpel. This curious genus has often been included in the Hamamelidales; however, the primitive nature of the wood and of the pistil make it anomalous there. *See* HAMAMELIDAE; MAGNOLIOPSIDA; PLANT KINGDOM.

T. M. Barkley

Dielectric heating

The heating of a nominally electrical insulating material due to its own electrical (dielectric) losses, when the material is placed in a varying electrostatic field.

The material to be heated is placed between two electrodes (which act as capacitor plates) and forms the dielectric component of a capacitor, as in the **illustration**. The electrodes are connected to a high-voltage source of 2–90-MHz power that is produced by a high-frequency vacuum-tube oscillator.

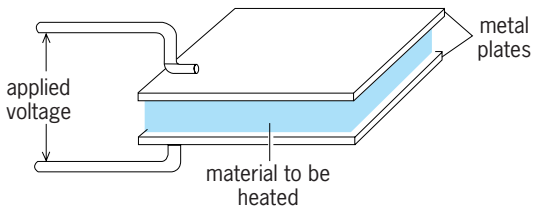

Basic assembly for dielectric heating

The resultant heat is generated within the material, and in homogeneous materials is uniform throughout. Dielectric heating is a rapid method of heating and is not limited by the relatively slow rate of heat diffusion present in conventional heating. *See* DIELECTRIC MATERIALS.

Applications. This technique is widely employed industrially for preheating in the molding of plastics, for quick heating of thermosetting glues in cabinet and furniture making, for accelerated jelling and drying of foam rubber, in foundry core baking, and for drying of paper and textile products. Its advantages over conventional methods are the speed and uniformity of heating, which offset the higher equipment costs. Because of the absence of high thermal gradients, an improved end-product quality is usually obtained.

Process. The heating rate obtainable, in watts, is shown by the equation below.

$$P = \frac{1.413AfE^2\kappa' \tan\delta \times 10^{-12}}{d} = \frac{5.56A'fE^2\kappa' \tan\delta \times 10^{-11}}{d'}$$

where P = heating rate, W
A = material area, in.^{2}
A' = material area, m^2
d = thickness of material, or electrode spacing, in.
d' = thickness of material, or electrode spacing, m
f = frequency, Hz
E = voltage across the electrodes, rms
κ' = dielectric constant of material
δ = loss angle (power factor)
$\tan\delta$ = dissipation factor

The factor $\kappa' \tan\delta$ is referred to as the loss factor, and serves to indicate the relative rates of heating for different materials at the frequency to be employed in high-frequency fields of the same intensity. *See* DIELECTRIC MEASUREMENTS.

The loss factor is affected by both temperature and frequency. Materials with a high loss factor heat rapidly, while those with a very low loss factor do not heat at all in high-frequency fields. The high loss factor of water (much higher than most insulators) means that materials which are difficult to heat by dielectric heating when dry are often efficiently heated when moisture is present.

The controllable variables in the process are voltage E and frequency f. Voltage can be raised only within limits determined by corona or dielectric breakdown, generally not exceeding 15,000–20,000 V. Voltage gradient across the work is generally the principal determinant, and this ranges between 1500 and 5000 V per inch (600 and 2000 V per centimeter), depending upon the porosity of the work. *See* CORONA DISCHARGE.

Frequencies used are as high as practicable, so that voltages can be kept within the limits given. Most standard equipment uses frequencies from 5 to 40 MHz, with power outputs up to 125 kW. Power rating decreases for the higher frequencies.

The work electrodes and the material being heated function like a capacitor. They are usually connected as a part of a resonant circuit, which is tuned to the oscillator frequency to obtain maximum transfer of power.

The electrodes are made from nonmagnetic materials, such as aluminum, copper, and stainless steel. The material to be heated is often conveyed through the electrodes on belts made from a low-loss insulating substance, such as coated glass cloth. *See* ELECTRODE.

Radiation shielding. These are certain frequencies, such as 13.56, 27.12, and 40.68 MHz, specifically allocated by international agreement for industrial use. In addition, to prevent interference with authorized radio services, such as established radio communication channels, industrial radio-frequency (rf) heating installations operating above 10,000 Hz must be properly shielded.

The usual method is to enclose completely all rf circuits and electrodes in a metal cabinet, using wire-mesh shielding over ventilation openings and good electrical contacting or overlapping surfaces on access doors. Where inlet and exit openings must exist, as in conveyorized installations, a metallic shield in the form of a vestibule or throat is used. The length and area of the shield are designed to attenuate rf radiation from the electrode area.

George F. Bobart

Bibliography. J. C. Anderson, *Dielectrics*, 1964; G. H. Brown, C. N. Hoyler, and R. A. Bierwirth, *Theory and Application of Radio Frequency Heating*, 1947; J. W. Cable, *Induction and Dielectric Heating*, 1954; *Recommended Practice for Minimization of Interference from Radio Frequency Heating Equipment*, IEE Stand. 951, 1950; Strayfield Ltd., *Introduction to Dielectric Heating*, 1975.

Dielectric materials

Materials which are electrical insulators or in which an electric field can be sustained with a minimal dissipation of power. A solid is a dielectric if its valence band is full and is separated from the conduction band by at least 3 eV.

Dielectrics are employed as insulation for wires, cables, and electrical equipment, as polarizable media for capacitors, in apparatus used for the propagation or reflection of electromagnetic waves, and for a variety of artifacts, such as rectifiers and semiconductor devices, piezoelectric transducers, dielectric amplifiers, and memory elements.

The ideal dielectric material does not exhibit electrical conductivity when an electric field is applied. The term dielectric, though it may be used for all phases of matter, is usually applied to solids and liquids. In practice, all dielectrics do have some conductivity, which generally increases with increase in temperature and applied field. For a good dielectric, such as polytetrafluoroethylene, the low-field direct-current conductivity at room temperature may be lower than $10^{-16}\ \Omega^{-1} \cdot m^{-1}$, whereas the corresponding figure for some specimens of plasticized poly(vinyl chloride) may be as high as $10^{-4}\ \Omega^{-1} \cdot m^{-1}$, similar to that of the low-conductivity semiconductors.

Breakdown. If the applied field is increased to some critical magnitude, the material abruptly becomes conducting, a large current flows (often accompanied by a visible spark), and local destruction occurs to an extent dependent upon the amount of energy which the source supplies to the low-conductivity path. This critical field depends on the geometry of the specimen, the shape and material of the electrodes, the nature of the medium surrounding the dielectric, the time variation of the applied field, and other factors. Temperature instability can occur because of the heat generated through conductivity or dielectric losses, causing thermal breakdown.

Solids may contain gaseous voids, or such voids may appear and grow in service. Since the gas will have a lower electric strength than the dielectric medium, gas discharge is likely in the voids, with accompanying destruction of the adjacent material and field intensification nearby. Thus the void grows and initiates a tree, and ultimately breakdown takes place. The same consequence may follow from the growth of water trees caused by the absorption of moisture into the bulk of the dielectric. Surface tracking is a common cause of breakdown, when a conducting path appears and elongates on the surface between the dielectric and the ambient medium, due to destruction of the material by localized discharges. After long exposure to the atmosphere, chemical degradation of solid dielectrics can occur, and so breakdown can be encouraged. In liquids, particles can be drawn into the high-field region, causing partial bridging of the high-potential gap, and so precipitate breakdown.

Thus, breakdown can be brought about by a variety of different causes, sometimes by a number of them acting simultaneously. Nevertheless, under carefully specified and controlled experimental conditions, it is possible to measure a critical field which is dependent only on the inherent insulating properties of the material itself in those conditions. This field is called the intrinsic electric strength of the dielectric. *See* ELECTRICAL BREAKDOWN.

Solids. The theory of dielectric breakdown was developed by the use of alkali halide crystals as a model, since their simple cubic lattice structure aids theoretical analysis. Although the gap between the valence and conduction bands is substantial (say 7 eV), crystal imperfections will result in isolated levels with a ground state (say 1 eV) and excited states (perhaps 0.1 eV) below the conduction band. Thus, thermal ionization ensures that there are always some electrons in the conduction band available to be accelerated by an applied field. The energy they gain is given up to the lattice by collisional excitation of vibrational modes. For the intrinsic breakdown field E^*, the power absorbed by the lattice is just less than the power gained from the field by electrons in a certain critical energy range; these electrons then accelerate up to ionization energy, and electron multiplication by an avalanche process ensues over a sufficient number of generations for local material destruction to occur. Since the electron scattering by the lattice increases with the lattice's deviation from periodicity, it would be expected that the intrinsic strength would increase with temperature, which is confirmed experimentally. *See* BAND THEORY OF SOLIDS; IONIC CRYSTALS.

However, if the temperature is increased further, the number of electrons in excited levels of the isolated states and in the conduction band ultimately becomes so high that electron-electron collisions predominate, and since the transfer of energy to the lattice is rather slow, the electron temperature rises above the lattice temperature. Provided the field is not too high, equilibrium can be attained, but there exists a critical field beyond which this is not so, and this defines the intrinsic strength. Analysis shows that this decreases with increasing temperature, and again, this is confirmed experimentally. Alkali halide crystals therefore exhibit a maximum electric strength, usually near room temperature (see **illus.**). Breakdown caused by these intrinsic processes is very fast, of the order of nanoseconds or less.

For materials with a disordered structure, in the low-temperature region the electric strength would be expected to be more and the slope less, possibly effectively zero. The transition to the high-temperature region tends to occur at a lower temperature, as the number of ionizable centers is large. The decline in strength is then quite sharp. This behavior is typical for polymers (see illus.). Although this behavior is in accordance with the theory outlined above, other explanations have been put forward, and the mechanism of intrinsic breakdown cannot be considered to be definitely

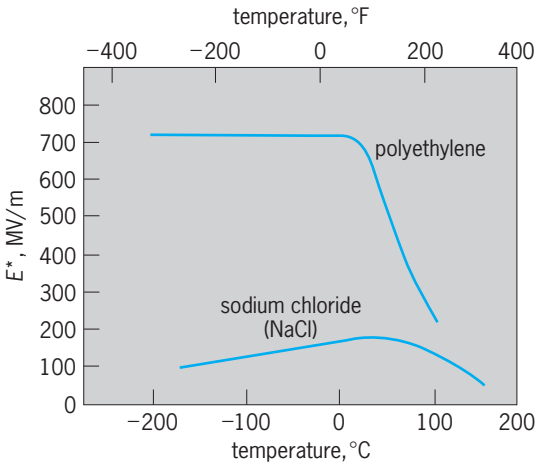

Variation of electric strength with temperature for sodium chloride and polyethylene.

established, particularly in the high-temperature region.

Another mechanism causing breakdown is thermal runaway caused by Joule heating. Because the material has some conductivity, and dielectric loss if the applied voltage is alternating, heat is generated. As the temperature rises, so does the conductivity, and if the generated heat cannot reach equilibrium with the heat loss at a sufficiently low temperature, then the material is destroyed. This is a much slower process than the electronic processes discussed above, and for most practical cases, the time for it to occur would rarely be less than about 1 millisecond, and often very much longer. Thermal breakdown often occurs in polymers, especially when the ambient is significantly above room temperature. *See* JOULE'S LAW.

Liquids. Less is known about the mechanism of liquid than of solid breakdown. Nevertheless, the development of very sophisticated photographic techniques has made it possible to study the critical processes which take place a few nanoseconds before breakdown occurs. These processes are complicated, and not necessarily the same for all liquids and conditions. In general, light detectable by image-intensifier techniques appears in the gap up to 1 microsecond or as little as 20 nanoseconds before breakdown, depending on electrode shape. This stage is probably associated with streamer formation and the growth of filamentary space-charge channels through the liquid. It is followed by the appearance of low-density disturbances, in some cases perhaps bubbles or cavities, near the cathode. These grow until the gap is bridged and breakdown occurs. The dielectric strength tends to increase with pressure and diminish with increasing temperature. Generally, liquids have a lower strength than solids; for example, to compare two good insulants at room temperature, polyethylene has a strength of 500–700 MV/m while transformer oil has a strength of 110–130 MV/m.

In large power devices, dielectric liquids provide not only insulation but also circulatory cooling. They also perform the important function of filling voids in solid insulation, for example, cables or high-voltage capacitors. It is an advantage for liquids used for impregnation to have a high E^* and a high ϵ_r' (real part of the complex relative permittivity), because this reduces the field in the filled cavity. Formerly, the liquids used for this purpose were the chlorinated biphenyls, which have these advantages and additionally have a low viscosity and are fire resistant. Unfortunately, they were found to be dangerous biochemically. However, satisfactory performance has since been achieved by the use of various esters. *See* POLYCHLORINATED BIPHENYLS.

Industrial dielectric materials. Many of the traditional materials are still in common use, and they compete well in some applications with newer materials regarding their electrical and mechanical properties, reliability, and cost. For example, oil-impregnated paper is still used for high-voltage cables, usually paper with values of ϵ_r' about 3–3.5 and tan δ (the loss tangent; the negative of the ratio of the imaginary part to the real part of the complex relative permittivity) about 2×10^{-3} at room temperature, rising to about 3×10^{-3} at 100°C (212°F); various types of pressboard and mica, often as components of composite materials, are also in use. Elastomers and press-molded resins are also of considerable industrial significance. However, synthetic polymers such as polyethylene, polypropylene, polystyrene, polytetrafluoroethylene, polyvinyl chloride, poly(methyl methacrylate), polyamide, and polyimide have become important, as has polycarbonate because it can be fabricated into very thin films.

Properties of polymers. Generally, these have crystalline and amorphous regions, increasing crystallinity causing increased density, hardness, and resistance to chemical attack, but often producing brittleness. Many commercial plastics are amorphous copolymers, and often additives are incorporated in polymers to achieve certain characteristics or to improve their workability. Additives include inorganic fillers, antioxidants, stabilizers, and pigments.

Polyethylene. Polyethylene (C_nH_{2n+2}) is the most common and most investigated polymer. It is relatively inexpensive and is easily worked. Polyethylene readily oxidizes, and so its working temperature is normally not above 70°C (158°F), although it may be worked at higher temperatures (say up to 90°C or 194°F) in suitable circumstances; it normally contains antioxidants. The molecular chains form planar zigzags, each chain being roughly perpendicular to those faces of the crystal plate containing it. Since the chain length is generally much longer than the distance between the faces of the plate (about 10 nanometers), the chain is generally folded back when it reaches the plate surface, although it can pass out into amorphous regions or other crystal plates. In its low-density form, polyethylene has a few branches along the molecular chain and is

usually 50–60% crystalline. *See* POLYOLEFIN RESINS.

High-density polyethylene has very few branches, is about 90% crystalline, and generally has better mechanical properties than those of the low-density form; its electrical properties are similar. Polyethylene is used for communications cables, and cross-linked polyethylene is used for high-voltage power cables. Here, voids may lead to treeing, and in a very moist environment a destructive phenomenon, water treeing, may occur. Polyethylene is also used for a multitude of small, inexpensive components. At room temperature, E^* is 500–700 MV/m; and over a wide frequency range, for commercial polyethylene ϵ'_r is 2.3 and tan δ is usually about 2.5×10^{-4} but may be more or less by a factor of 2.

Poly(ethylene terephthalate). Poly(ethylene terephthalate) is a useful material which has good resistance to acids and alkalies provided that they are dilute. It is insoluble in many common solvents, although it does dissolve in aromatic and chlorinated hydrocarbons, especially when warm. It softens at about 250°C (480°F).

Polypropylene. Polypropylene is similar in structure to polyethylene, but every other carbon atom has one of its hydrogen atoms replaced by a CH_3 group. Though electrically similar to polyethylene, polypropylene can be made in thinner films, say 5 micrometers as against about 25 μm for polyethylene. These films replace paper for impregnated capacitors, with reduced loss.

Polystyrene. Polystyrene is brittle at room temperature, becomes soft at 80°C (176°F), and is often modified by copolymerization. Traditionally, it is used in film form for capacitors, and it remains competitive for this application. Polystyrene is also used for coaxial-cable insulation, but in wound strip or bead form, since the solid is not very flexible. With very low water absorption, tan δ can be around 10^{-4} up to 1 MHz. For commercial polystyrene, ϵ'_r is about 2.6 and E^* about 400–600 MV/m. *See* POLYSTYRENE RESIN.

Polytetrafluoroethylene. Polytetrafluoroethylene is similar to polyethylene, with all the hydrogen atoms replaced by fluorine atoms. Highly crystalline (about 95%), it resists twisting and bending and is mechanically tough, with a very low coefficient of friction, 0.06. It is highly resistant to chemical attack and is useful in hostile environmental conditions. The loss is very low, with tan δ less than 2.5×10^{-4} at frequencies up to 1 GHz, and tan δ less than 10^{-4} at frequencies below 1 MHz; E^* is in the range 400–500 MV/m, and ϵ'_r is almost constant at 2.1 over a wide frequency range. *See* POLYFLUOROOLEFIN RESINS.

Poly(vinyl chloride). Poly(vinyl chloride) has a structure similar to that of polyethylene, but every other carbon atom has one of its hydrogen atoms replaced by a chlorine atom. It is only about 10% crystalline and is compatible with a large number of other polymers. It is easily decomposed by heat and is not suitable for continuous use above 70°C (158°F). Poly(vinyl chloride) is not well characterized, and depending on admixtures, may be rigid or flexible. Though suitable compositions may be molded, its main electrical use is for coating wires. Water absorption, especially by rigid poly(vinyl chloride), is high, but absorption may be kept down by suitable fillers. Its ϵ'_r is usually between 3.5 and 12, and tan δ may be as large as 0.2. It does not withstand high electrical stresses. *See* POLYVINYL RESINS.

Poly(methyl methacrylate). Poly(methyl methacrylate) is a clear solid, hard but easily scratched. Its dielectric properties are only moderate, and its use is restricted to undemanding conditions at normal ambient temperatures, usually in a molded form, where appearance is important. Its ϵ'_r is about 3.6, with tan δ about 0.05. *See* POLYACRYLATE RESIN.

Polyamide. Polyamide may be in linear form, known as nylon, or in aromatic form. Neither of these is of much value as a dielectric, but nylons, because of their toughness and resistance to solvents, are used to form coatings to protect insulation. The aromatic form can also make yarns, but it is most often used in impregnated board form in which it gives low-voltage insulation capable of being used continuously at temperatures of 150°C (302°F), or even higher. Its ϵ'_r is about 3.4, with tan δ about 0.03. *See* POLYAMIDE RESINS.

Polyimide. Polyimide is related to aromatic polyamide, but with additional aromatic groups. Expensive to produce, it can be made in films down to 25 μm thick, and in varnish form for wire coating. It is very tough, does not burn but chars above 800°C (1500°F), and remains flexible at liquid helium temperatures; it can be used continuously up to 250°C (482°F). Polyimide resists organic solvents and acids but not strong alkalies, has a high water absorption rate, and is subject to surface tracking. Its ϵ'_r is about 3.4, with a tan δ of 0.005. *See* HETEROCYCLIC POLYMER.

Resins. The resins are important members of the family of cross-linked polymers. Epoxy resins have a high mechanical strength, absorb very little water, and bond easily to most materials but not to polyethylene. They are used for bonding and encapsulation, and their properties can be modified by the curing process used and by the use of fillers and hardeners. The ϵ'_r of epoxy resins is about 3.6–4.0. and tan δ is 0.01–0.03 at room temperature at frequencies below 1 MHz. *See* ELECTRICAL INSULATION; PERMITTIVITY; POLYETHER RESINS; POLYMER. James H. Calderwood

Bibliography. R. Bartnikas (ed.), *Engineering Dielectrics*, vol. 3: *Electrical Insulating Liquids*, 1993; R. Bartnikas and R. M. Eichorn (eds.), *Engineering Dielectrics*, vol. 2A: *Electrical Properties of Solid Insulating Materials: Molecular Structure and Electrical Behavior*, 1983; A. Bradwell et al. (eds.), *Electrical Insulation*, IEE Electrical and Electronic Materials and Devices, no. 2, 1983; K. C. Kao, *Dielectric Phenomena in Solids*, 2004; J. I. Kroschwitz (ed.), *Electrical and Electronic Properties of Polymers*, 1988; C. Mazzetti, M. Pompili, and R. Bartnikas (eds.), Special issue on dielectric liquids, *IEEE Trans. Dielec. Elec. Insulation*, 5(3):305-407, 1998.

Dielectric measurements

Measurements of the dielectric properties of a material, which are characterized by its complex relative permittivity ϵ_r. For all materials except ferroelectrics, this quantity does not depend on applied field: the general behavior is linear, and so voltage of any convenient magnitude can be used for measurement. *See* FERROELECTRICS; PERMITTIVITY.

Bridge methods. The most commonly used apparatus for measuring ϵ_r is the alternating-current (ac) bridge. These bridges are readily available in the operating range 10–10^6 Hz, and sometimes outside it; ultralow-frequency bridges can go as low as 10^{-3} Hz. Most specimen holders for solids are essentially parallel-plate capacitors with the specimen filling all the space between the circular plates; for liquids, a test cell with cylindrical electrodes is usually employed. Guard electrodes are incorporated to avoid errors due to field fringing and surface conduction, although at the higher frequencies guard electrodes as well as the lead impedances cause difficulties, and it is usual to employ specially designed two-terminal specimen holders.

Derivation of dielectric properties. The measured impedance may be expressed in terms of an equivalent capacitance C_p and resistance R_p in parallel. From these, at an angular frequency ω, the real part of the complex relative permittivity, ϵ_r', is given by Eq. (1); the negative of the imaginary part, ϵ_r'', is given by Eq. (2); and the loss angle δ is given by Eq. (3), where C_0 is the capacitance of the empty specimen holder.

$$\epsilon_r' = \frac{C_p}{C_0} \tag{1}$$

$$\epsilon_r'' = \frac{1}{\omega C_0 R_p} \tag{2}$$

$$\tan\delta \equiv \frac{\epsilon_r''}{\epsilon_r'} = \frac{1}{\omega C_p R_p} \tag{3}$$

Alternatively, the measured impedance may be expressed in terms of an equivalent capacitance C_s and resistance R_s connected in series. In terms of these, ϵ_r', ϵ_r'', and $\tan\delta$ are given by Eqs. (4), (5), and (6).

$$\epsilon_r' = \frac{C_s}{C_0\,(1+\tan^2\delta)} \tag{4}$$

$$\epsilon_r'' = \frac{C_s \tan\delta}{C_0\,(1+\tan^2\delta)} \tag{5}$$

$$\tan\delta = \omega C_s R_s \tag{6}$$

If the material contains free charge carriers resulting in an ohmic conductivity σ, then this causes an apparent ϵ_r'' given by Eq. (7). This quantity has to be

$$\epsilon_r'' = \frac{\sigma}{\epsilon_0 \omega} \tag{7}$$

subtracted from the measured ϵ_r'' to give the value of ϵ_r'' attributable to dielectric loss processes.

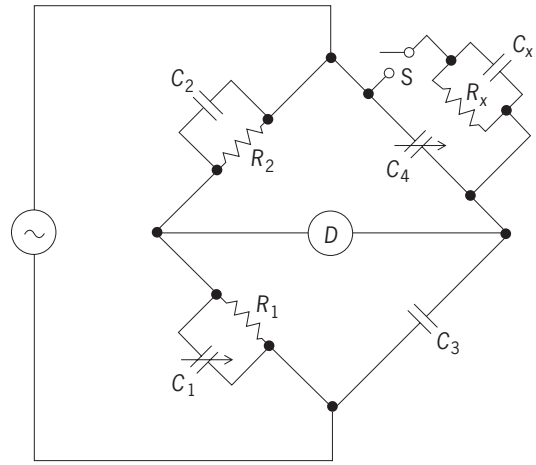

Fig. 1. High-frequency form of the Schering bridge. *D* = detector; *S* = switch.

Schering bridge. The bridges most commonly used are of the Wheatstone type, the most versatile for dielectric measurements being the Schering bridge. This is commonly used in its conjugate form arranged for substitution measurement (**Fig. 1**). The dielectric specimen is represented by the elements C_x and R_x in parallel. The bridge is first balanced (that is, adjusted until the voltage across the detector is zero) with a switch between the specimen and the bridge circuit open, and the values of adjustable capacitors C_1 and C_4 on opposite arms of the bridge noted. The switch is then closed so that the specimen is in parallel with C_4, and balance is regained by changing C_1 and C_4 by amounts ΔC_1 and ΔC_4, respectively. It may be shown that, provided $\tan\delta_x$ is less than 10^{-2}, the capacitance and loss angle of the specimen are given to a good approximation by Eqs. (8) and (9). By the use of standard Wagner grounding

$$C_x = \Delta C_4 \tag{8}$$

$$\tan\delta_x = \omega R_1 C_4 \frac{\Delta C_1}{\Delta C_4} \tag{9}$$

techniques and grounded screens around all bridge components, high-precision measurements of ϵ_r' and $\tan\delta$ may be made up to a frequency 10^6 Hz. *See* BRIDGE CIRCUIT; WHEATSTONE BRIDGE.

The Schering bridge may also be used in a different form for high-voltage measurements at 60 Hz (**Fig. 2**). R_3 and C_1 are adjusted until balance is achieved, and the capacitance and loss angle of the specimen are then given by Eqs. (10) and (11). Here, R_1 and C_1

$$C_x = \frac{C_2 R_1}{R_3\,[1+(\omega C_1 R_1)^2]} \tag{10}$$

$$\tan\delta_x = \omega C_1 R_1 \tag{11}$$

are a parallel resistance and capacitance on the arm of the bridge opposite the specimen, and C_2 and R_3 are a capacitance and a resistance on the arms adjacent to the specimen. Because the test voltage is

high, the high-potential components are surrounded by a grounded cage. The voltage appearing outside the cage is small; for example, if the test voltage is 100 kV and C_2 is 100 picofarads, then the potential difference across the variable bridge arms is about 1 V. Because the frequency is also low, R_3 and C_1 may be of the decade-box type. However, a protective device must be incorporated across the low-voltage bridge arms in case the specimen should break down and a high voltage appear outside the cage and cause damage to the bridge elements and create a hazard for the operator.

Other bridge types. The are several other four-armed types of bridge, such as the bridged-T network, which has the advantage that the source and detector have a common terminal which may be grounded. Operating on a different principle, bridges with inductively coupled ratio arms compare the impedance to be measured directly with standard components. The balance condition of such a transformer bridge (**Fig. 3**) is given by Eq. (12).

$$\frac{Z_1}{Z_2} = \frac{N_2 N_4}{N_1 N_3} \tag{12}$$

The advantages of this system are that both the source and detector may be grounded at one point and that three-terminal devices cause no difficulties, as the guard ring is connected to point G. The stray admittances associated with the guard ring then do not affect the balance.

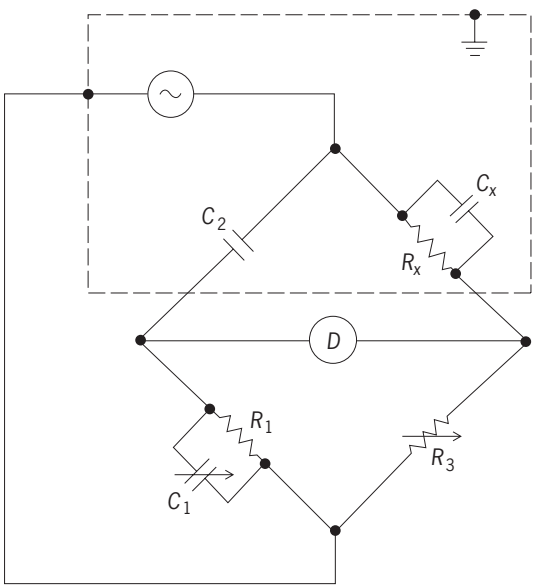

Fig. 2. High-voltage form of the Schering bridge.

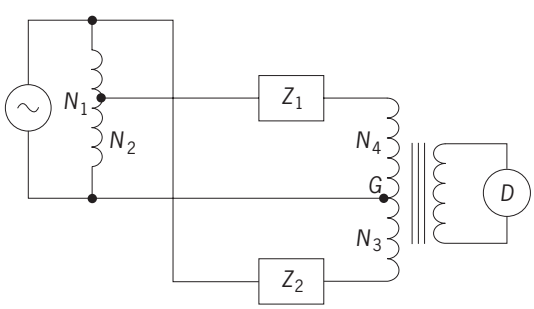

Fig. 3. Transformer bridge.

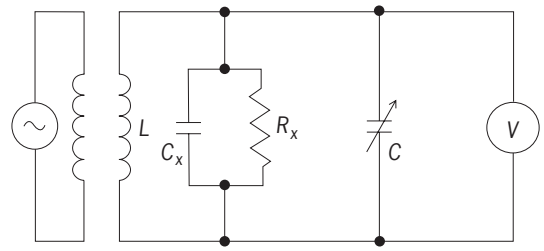

Fig. 4. Hartshorn and Ward basic circuit.

Resonance methods. Resonance methods, useful for frequencies greater than 1 MHz, involve the injection of voltage or current by one of several methods into an LC (inductance-capacitance) resonant circuit. The classical circuit arrangement is that of L. Hartshorn and W. H. Ward (**Fig. 4**). Two circular metal plates are in contact with the specimen in the shape of a disk of the same radius to form a capacitor with plate separation x, represented by elements C_x and R_x in parallel, the distance between the plates being controllable by a micrometer screw; a variable capacitor C is in parallel with the specimen. The circuit is energized by the employment of a loosely coupled oscillator, the frequency of which is adjusted for resonance as indicated by the maximum voltage V_i registered by a voltmeter V. The specimen is now removed, and resonance is restored by reducing x by an amount δx by the micrometer screw, when the capacitance between the plates is C_0, and the new maximum voltage V_0 is noted. The auxiliary variable C is now used to measure the change in capacitance δC necessary to cause a shift between the two half-power points on either side of resonance. It is evident that ϵ_r' is given by Eq. (13), and it can be shown that tan δ is given by Eq. (14).

$$\epsilon_r' = \frac{x}{x - \delta x} \tag{13}$$

$$\tan\delta = \frac{[(V_0/V_i) - 1]\,\delta C}{2C_0} \tag{14}$$

See RADIO-FREQUENCY IMPEDANCE MEASUREMENTS; RESONANCE (ALTERNATING-CURRENT CIRCUITS).

Measurements over a range of frequencies may be made by using coils with different inductance values, but ultimately the inductance required becomes impracticably small, and in the range 10^8–10^9 Hz reentrant cavities are often used. These are hybrid devices in which the plates holding the specimen still form a lumped capacitor, but the inductance and capacitances are distributed along a coaxial line. At higher frequencies, the wavelength is comparable to the dimensions of the apparatus, and transmission methods in coaxial lines and waveguides can be used.

Resonance methods can also be used at microwave frequencies, for example, by using resonant cavities

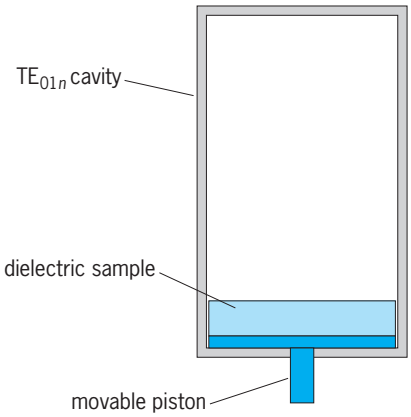

Fig. 5. TE_{01n} resonator. The resonator can be brought into resonance by moving the piston or by changing the frequency.

where the permittivity of the sample can be deduced from the change in the resonant frequency or length and Q of the cavity. TE_{01n} cavities (**Fig. 5**) need disk-shaped samples and TM_{01n} cavities need thin cylindrical samples. Resonance methods are usually limited to measurements on low-loss materials but can measure higher-loss materials by employing perturbation techniques. *See* CAVITY RESONATOR; Q (ELECTRICITY).

Open resonators measure low-loss materials. In a typical arrangement using plane and curved metal mirrors (**Fig. 6**), a Gaussian beam is generated within the resonator and the sample is placed at the beam waist on plane mirror. The microwaves are usually introduced into the cavity by two holes in the center of the curved mirror. Alternative arrangements involve two curved mirrors. Because of the small beam size, open resonators are used for frequencies above 10 GHz.

Other more specialized techniques include dielectric resonators, split-post dielectric resonators, and stripline resonators. *See* MICROWAVE.

Transmission and reflection methods. Coaxial lines are used in the frequency range 300 MHz–3 GHz, and waveguides in the range 3–30 GHz. The transmission characteristics are determined by the complex permittivity of the material filling the line or guide. Many different measurement techniques have been devised, but all derive values of the complex relative permittivity ϵ_r from its relationship to the complex propagation factor γ, that is, Eq. (15), where λ_0 is the

$$\epsilon_r = -\gamma^2 \left(\frac{\lambda_0}{2\pi}\right)^2 \tag{15}$$

free-space wavelength for the frequency employed. The propagation factor γ is obtained by measuring its real and imaginary components α and β, where α gives the attenuation, and β the wavelength λ in the material. For a traveling wave in a coaxial line, it may be shown that ϵ_r' and ϵ_r'' are given by Eqs. (16) and (17).

$$\epsilon_r' = \left(1 - \frac{\alpha^2}{\beta^2}\right)\left(\frac{\lambda_0}{\lambda}\right)^2 \tag{16}$$

$$\epsilon_r'' = \frac{2\alpha}{\beta}\left(\frac{\lambda_0}{\lambda}\right)^2 \tag{17}$$

For transmission in a waveguide limited to the single transverse TE_{01} mode, these equations are somewhat modified, since ϵ_r is given by Eq. (18), where γ_g

$$\epsilon_r = \frac{\left(\frac{1}{\lambda_c}\right)^2 + \left(\frac{\gamma_g}{2\pi}\right)^2}{\left(\frac{1}{\lambda_c}\right) + \left(\frac{1}{\lambda_{0g}}\right)^2} \tag{18}$$

is the propagation factor of this mode in the material-filled guide, λ_{0g} is the mode wavelength in vacuum, and λ_c is the cutoff wavelength of the guide. *See* TRANSMISSION LINES; WAVEGUIDE.

In practice, traveling waves are rarely used as the basis of measurement, except for high-loss materials. Usually, reflections from terminations set up standing waves, the amplitude of which in the case of a liquid-filled line can be measured by a suitable probe. The ratios of the field magnitudes at adjacent maxima, and the distance between them, give the information required. For solid, ϵ_r can be determined by probe measurements in an air-filled section of line between the signal source and a dielectric-filled section. Other methods, more often used for waveguide measurements, involve the use of a movable short-circuiting piston. *See* ATTENUATION (ELECTRICITY); MICROWAVE IMPEDANCE MEASUREMENT; WAVELENGTH MEASUREMENT.

Dielectric properties can also be measured with open-ended waveguide or coaxial probes placed directly on the surface of the material (**Fig. 7**). Coaxial probes are more convenient than waveguides because of their greater frequency range and generally smaller size. These devices are widely applicable to noninvasive dielectric measurements, including measurements on solids, liquids, and biological tissue. The traveling wave propagates along the coaxial line to its end, where fringing fields are launched

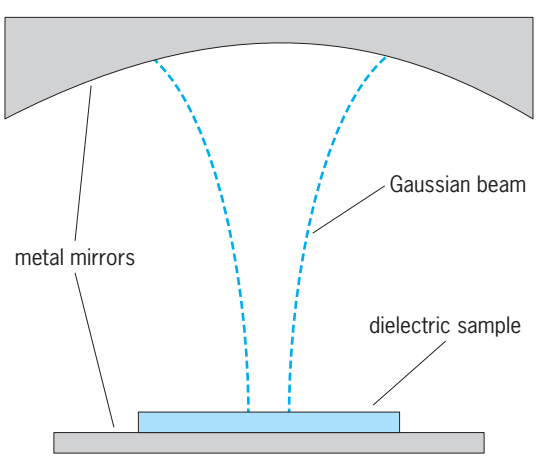

Fig. 6. Open resonator. The microwave radiation is usually introduced via waveguides through holes in the center of the curved mirror.

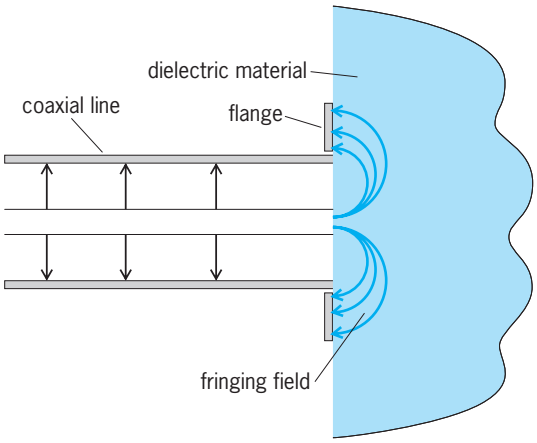

Fig. 7. Coaxial probe.

into the material. The reflection coefficient of the probe is a function of the dielectric properties of the material. The probes are often available with tested software to relate the measured reflection to the material properties. The permittivity of the material can be related to the reflection coefficient by modal analysis of the fields in the coaxial line and an analysis of the fields in the dielectric where the probe is treated as an antenna. In some cases at lower frequencies a capacitive model can be used for the fringing fields. *See* ANTENNA (ELECTROMAGNETISM).

At higher frequencies, free-space or optical-methods can also be used, in which the microwave beam is launched by a horn and can be collected by another horn. The sample is placed in the beam between the horns and its dielectric properties can be deduced from its effect on the beam. This method can be used in transmission or reflection. If corrugated horns are used in conjunction with lenses, a well-shaped Gaussian beam can be produced and this leads to more accurate measurements. Both this method and the open resonator method described above can investigate anisoptropic dielectrics.

Submillimeter measurements. Dielectric measurements are difficult to carry out in the frequency range 30–300 GHz, for which λ_0 is in the range 1 cm–1 mm, but for λ_0 less than 1 mm, methods related to infrared spectroscopy are used. Broadband continuous spectra result from Fourier transform spectroscopy, which in its simplest form is equivalent to normal infrared spectroscopy, with the specimen in one of the two passive arms of the interferometer, between the beam divider and either the source or the detector. In the more sophisticated dispersive Fourier transform spectroscopy, which is more difficult experimentally but offers some procedural and data-handling advantages, the specimen is in one of the active arms, that is, between the beam divider and either mirror (**Fig. 8**). Broadband spectroscopy is only one possible measurement technique; discrete-point spectra also may be obtained by the use of a Mach-Zehnder interferometer and a laser source. By using interferometric techniques, the frequency range can be extended up to about 5 THz. Above that frequency is the region of resonant processes, the domain of the established methods of infrared and optical measurements. *See* INFRARED SPECTROSCOPY; INTERFEROMETRY; SPECTROSCOPY; SUBMILLIMETER-WAVE TECHNOLOGY.

Time-domain methods. If a constant direct-current (dc) voltage is suddenly applied to a dielectric specimen, in principle the charging current is related through the Fourier integral transformation to the steady-state ac current which would flow if the applied voltage were sinusoidal at any particular frequency. If the dc voltage is suddenly removed, a similar relationship holds between the discharge current and ac current. Thus the variation of complex permittivity with frequency can in principle be derived from a transient signal in the time domain. Because of various limitations, the method is not capable of giving results of an accuracy at all frequencies comparable to those obtainable from a single frequency measurement. Nevertheless, with the aid of computer analysis, the response over a large frequency range can be obtained much more quickly than would be possible by using point-by-point measurement methods.

Originally, the dc step-response technique was used to measure complex permittivity at very low frequencies, down to 10^{-4} Hz. Later, it was developed for use in the range 10^8–10^{10} Hz, the material being placed in a waveguide and subjected to a train of short fast-rise-time pulses. By observing the transmitted and reflected pulses with a sampling oscilloscope, the complex permittivity may be deduced. This technique is known as time-domain spectroscopy. Because of the way in which the reflections from the various interfaces are separated in time, a gating method can be used to isolate the required reflections, that is, those from the air-dielectric interfaces.

Automated instrumentation. The bridges described above are usually operated manually, but several automated systems have become available where the appropriate electrical properties of the cell

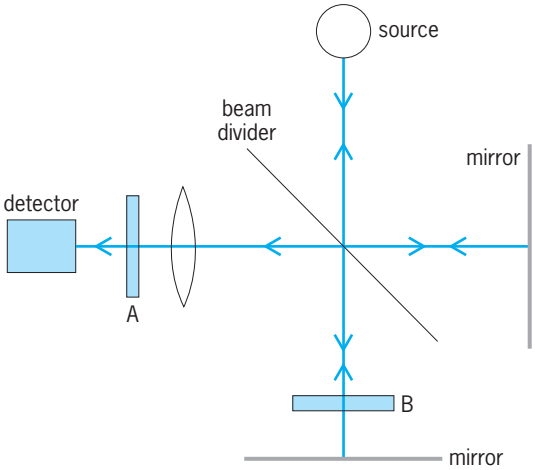

Fig. 8. Interferometer arrangement with specimen at A for conventional Fourier transform spectroscopy and at B for dispersive Fourier transform spectroscopy.

containing the dielectric material are measured automatically.

Automatic network analyzers (ANAs), developed originally for measuring the electrical properties of radio-frequency and microwave components, also measure a cell containing dielectric material, and the dielectric properties can be deduced from the difference in the electrical properties of the cell when empty and when containing the dielectric material.

Impedance analyzers can be used at low and radio frequencies. In general they are more accurate than ANAs for frequencies up to 200 MHz. In some cases the instruments are supplied with an external dielectric measurement cell and are often called materials analyzers.

Frequency response analyzers are instruments which are designed for dielectric measurements below 10 MHz down to as low as a few microhertz. The material is placed in an admittance cell which is in the form of a parallel-plate capacitor. The admittance is calculated from measurements of the in-phase and quadrature components of the voltage using an automated comparator circuit. The real and imaginary components of the permittivity of the material are obtained from the dimensions of the sample and the measured capacitance and conductance. These instruments have an impedance range from 0.01 ohm to 100 teraohms and a loss-angle resolution of less than 0.1 mrad. *See* ADMITTANCE; CAPACITANCE MEASUREMENT.

James H. Calderwood; R. Gareth Jones

Bibliography. R. Bartnikas (ed.), *Engineering Dielectrics*, vol. 2B: *Electrical Properties of Solid Insulating Materials*: *Measurement Techniques*, 1987; J. R. Birch and T. J. Parker, Dispersive Fourier transform spectroscopy, in K. J. Button (ed.), *Infrared and Millimeter Waves*, vol. 2: *Instrumentation*, pp. 137–271, Academic Press, 1979; A. R. Blythe, *Electrical Properties of Polymers*, 1979; A. Bradwell et al. (eds.), *Electrical Insulation*, IEE Electrical and Electronic Materials and Devices, no. 2, 1983; R. N. Clarke et al., *A Guide to the Characterisation of Dielectric Materials at RF and Microwave Frequencies*, Institute of Measurement and Control/National Physical Laboratory, 2003; I. Kemp et al. (eds.), Special issue on measurement techniques, *IEE Proc. Dielectric Materials, Measurements and Applications*, vol. 7, 1996; B. K. P. Scaife, *Complex Permittivity: Theory and Measurement*, 1971; A. von Hippel (ed.), *Dielectric Materials and Applications*, Wiley, 1954, reprint, Artech House, 1995.

Diels-Alder reaction

The 1,4-addition of an alkene (the dienophile) to a conjugated diene. The reaction, also known as the diene synthesis, is one of the most valuable and versatile methods for the preparation of compounds containing a six-membered ring, and proceeds most rapidly when the dienophile is substituted by electron-attracting groups, such as —C═O or —C≡N [reactions (1) and (2)].

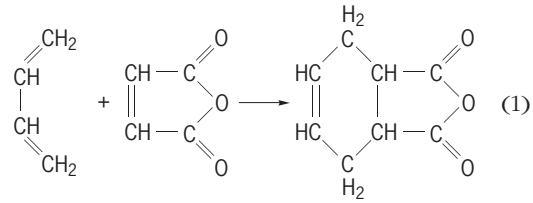

(1)

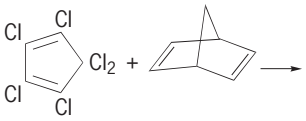

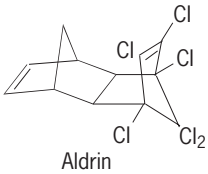

(2)

Diene component. With the exception of a few highly substituted diolefins which interfere with the reaction because of steric effects, most alkyl and aryl homologs of butadiene react readily. Alicyclic dienes are especially reactive, except in those cases where the adduct formed would contain a bridged ring having a double bond at the bridgehead. Aromatic compounds such as styrene, in which part of the conjugation is in the ring, react in a normal manner, but heterocyclic dienes may show anomalous behavior. Furan reacts normally to give a six-membered ring containing a bridged oxygen, whereas thiophene and pyrrole do not react in this manner.

Dienophile component. Most commonly, dienophiles consist of compounds containing structures

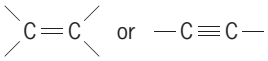

However, dienophiles are not restricted to unsaturated carbon compounds, and adducts form with

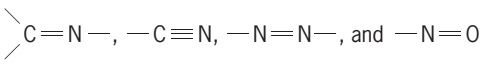

as the dienophile component. Quinones are especially reactive with dienes, requiring only mild conditions.

Reaction conditions. The Diels-Alder or diene-synthesis reaction does not require a catalyst, nor is the reaction retarded by the presence of oxidation inhibitors with which dienes are commonly treated to prevent formation of peroxides. For reactive dienophiles such as maleic anhydride, it is sufficient to mix the reactants in molar proportion, usually in a solvent such as benzene, and reaction takes

place at room temperature—often with the evolution of heat. On the other hand, the reaction of ethylene with cyclopentadiene to form bicyclo-[2.2.1]-2-heptene [reaction (3)] requires temperatures in the

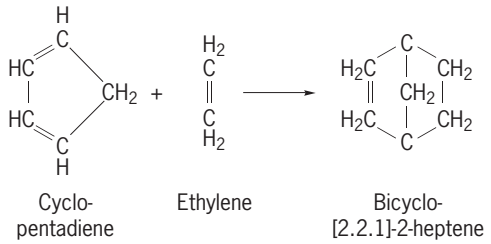

(3)

range of 190–220°C (370–430°F) and pressures in the range of 20–80 atm (2–8 megapascals).

Instances are known where the diene synthesis is a reversible reaction. For example, cyclohexene, the adduct of 1,3-butadiene and ethylene, is quantitatively transformed into its components when subjected to high temperature for a short contact time at low pressure. Similarly, dicyclopentadiene (**1**) yields

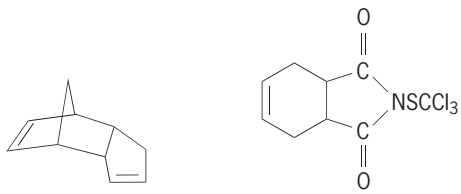

(**1**) (**2**)

two moles of cyclopentadiene when maintained at its boiling point of 170°C (338°F).

Industrially the diene synthesis is used in the production of the insecticides aldrin and dieldrin. The adduct of butadiene and maleic anhydride is used in the synthesis of the fungicide captan (**2**).

Paul E. Fanta

Bibliography. R. T. Morrison and R. N. Boyd, *Organic Chemistry*, 6th ed., 1992; A. Wassermann, *Diels-Alder Reactions*, 1965.

Diesel cycle

An internal combustion engine cycle in which the heat of compression ignites the fuel. Compression-ignition engines, or diesel engines, are thermodynamically similar to spark-ignition engines. The sequence of processes for both types is intake, compression, addition of heat, expansion, and exhaust. Ignition and power control in the compression-ignition engine are, however, very different from those in the spark-ignition engine.

Usually, a full unthrottled charge of air is drawn in during the intake stroke of a diesel engine. A compression ratio between 12 and 20 is used, in contrast to a ratio of 4 to 10 for the Otto spark-ignition engine. This high compression ratio of the diesel raises the temperature of the air during the compression stroke. Just before top center on the compression stroke, fuel is sprayed into the combustion chamber. The high temperature of the air ignites the

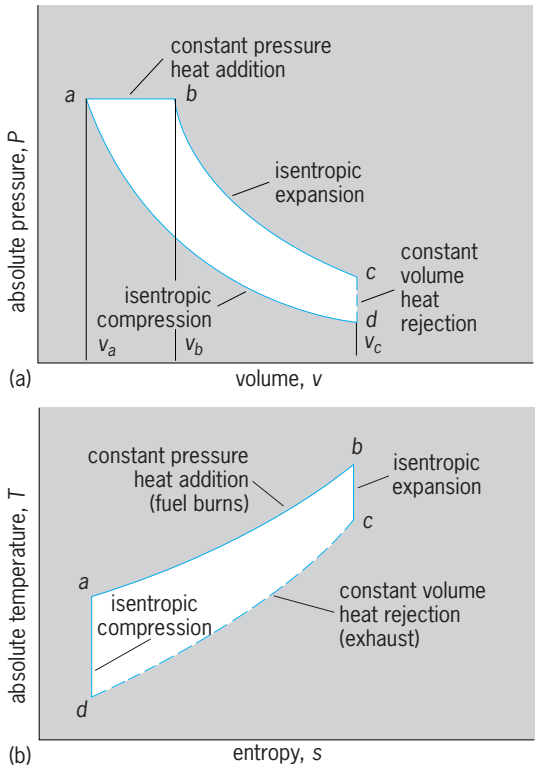

Fig. 1. Ideal diesel cycle, with (*a*) pressure-volume and (*b*) temperature-entropy bases.

fuel, which burns almost as soon as it is introduced, adding heat. The combustion products expand to produce power, and exhaust to complete the cycle.

Performance of a diesel engine is anticipated by analyzing the action of an air-standard diesel cycle. An insulated cylinder equipped with a frictionless piston contains a unit air mass. The metal cylinder head is alternately insulated and then uncovered for heat transfer.

Air is compressed until the piston reaches the top of the stroke. Then the air receives heat through the cylinder head and expands at constant pressure along path *a-b* as in **Fig. 1**, moving the piston part way down through the cylinder. Then the cylinder head is insulated, and the air completes its expansion along path *b-c* at constant entropy. The cylinder head is uncovered, and with the piston at the bottom of its stroke, a constant-volume heat rejection takes place on path *c-d*. The insulation is replaced, and the cycle is completed with an isentropic compression on path *d-a*.

An increase in compression ratio $r = v_d/v_a$ increases efficiency η, the increase becoming less at higher compression ratios. Another characteristic of the diesel cycle is the ratio of volumes at the end and at the start of the constant-pressure heat-addition process. This cutoff ratio $r_c = v_b/v_a$ measures the interval during which fuel is injected. For an engine to develop greater power output, the cutoff ratio is increased and heat continues to be added further into the expansion stroke. The air-standard cycle shows that, with less travel remaining during which to expend the additional heat energy as mechanical

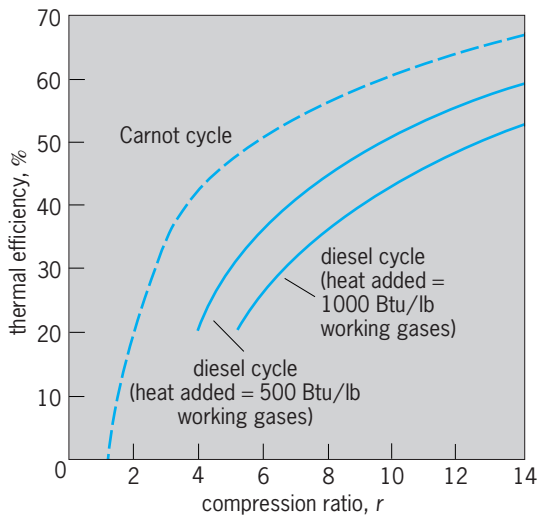

Fig. 2. Thermal efficiency of an ideal diesel cycle.

energy, the efficiency of the engine is reduced. Conversely, efficiency increases as the cutoff ratio decreases, so that a diesel engine is most efficient at light loads. Specifically, theequation below may be

$$\eta = 1 - \frac{1}{r^{k-1}}\left[\frac{r_c^k - 1}{k(r_c - 1)}\right]$$

written, where $k = c_p/c_v$, the ratio of specific heat of the working substance at constant pressure to its specific heat at constant volume (**Fig. 2**). In the limiting case when cutoff ratio r_c approaches unity, diesel cycle efficiency approaches Otto cycle efficiency for cycles of the same compression ratio.

In an actual engine with a given compression ratio, the Otto engine has the higher efficiency. However, fuel requirements limit the Otto engine to a compression ratio of about 10, whereas a diesel engine can operate at a compression ratio of about 15 and consequently at a higher efficiency.

In addition, heat can be added earlier in the cycle by injecting fuel during the latter part of the compression process *d-a*. This mode of operation is the dual-combustion or semidiesel cycle. With most of the heat added near peak compression, semidiesel efficiency approaches Otto cycle efficiency at a given compression ratio. *See* DIESEL ENGINE; INTERNAL COMBUSTION ENGINE. Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; L. Lichty, *Combustion Engine Processes*, 7th ed., 1967; D. S. Williams (ed.), *The Modern Diesel: Development and Design*, 14th ed., 1977.

Diesel engine

An internal combustion engine operating on a thermodynamic cycle in which the ratio of compression ($R_v = 15\pm$) of the air charge is sufficiently high to ignite the fuel subsequently injected into the combustion chamber. Compared to an engine operating on the Otto cycle, the diesel engine utilizes a wider variety of fuels with a higher thermal efficiency and consequent economic advantage under many service applications (**Fig. 1**). *See* DIESEL CYCLE; OTTO CYCLE.

Features. The true diesel engine, as represented in most low-speed engines, uses a fuel-injection system where the injection rate is delayed and controlled to maintain constant pressure during combustion. Adaptation of this injection principle to higher engine speeds has necessitated departure from the constant-pressure specification, because the time available for fuel injection is so short (often 2 ms or less). Nonvolatile (distillate) fuels are burned to advantage in these engines, which cannot be rigorously identified as true diesels but properly should be called commercial diesels. However, all such engines are ordinarily classified as diesels. Performance data for selected diesel engines are given in **Table 1**.

Diesel engines give high intrinsic and actual thermal efficiency (20–40%); a sample comparative heat balance is shown in **Table 2**. The engine output is controlled by regulation of the fuel supplied but without variation of the air supply (100±% excess air at full load). Supercharging (10–15 lb/in.^{2} or 70–100 kilopascals) increases power output for a given cylinder size and engine speed. With two-cycle constructions, scavenging air (approximately 5 lb/in.^{2} or 35 kPa) is delivered by crankcase compression, front-end compression, or a separate rotary, reciprocating, or centrifugal blower. The engine cylinder may be without valves and with complete control of admission of scavenging air and release of spent gases in a two-port construction, the piston covering and uncovering the ports; or the cylinder may have a single port (for admission or release) uncovered by

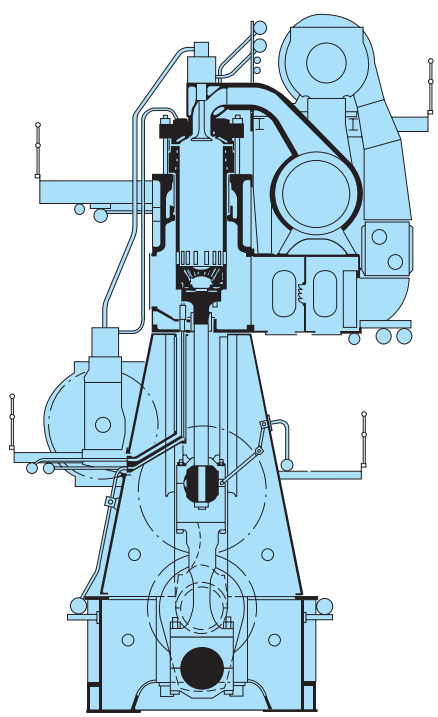

Fig. 1. Cross section of a slow-speed two-stroke diesel engine. (*Sulzer Brothers, Inc.*)

TABLE 1. Performance of selected diesel engines*

Type	Shaft horse-power (shp)	Ratio of compression, R_v	Brake mean effective pressure, lb/in.^{2}	Piston speed, ft/min	Mass, lb/hp	Overall thermal efficiency, %
Air injection engine	300–5000	12–15	50–75	600–1000	25–200	30–35
Solid injection, compression ignition						
Automotive	20–300	12–21	75–100	800–1800	7–25	25–30
Railroad	200–2500	12–15	60–90	800–1800	10–40	30–35
Stationary						
Unsupercharged	50–2500	12–15	70–80	600–1500	10–100	30–35
Supercharged	60–4000	10–13	110–125	600–1500	7.5–75	32–40
Dual fuel, stationary						
Unsupercharged	50–2500	12–15	80–90	600–1500	10–100	30–35
Supercharged	60–4000	10–13	120–135	600–1500	7.5–75	32–40

*1 hp = 746 W; 1 lb/in.^{2} = 6.9 kPa; 1 ft/min = 0.305 m/min ; 1 lb/hp = 6.08×10^{-4} kg/W.

the main piston at the outer end of its stroke and a cam-operated poppet valve in the cylinder head. The objective is to replace spent gases with fresh air by guided flow and high turbulence.

The four-cycle engine, with the intake and exhaust valves in the cylinder head, is most effective in scavenging. But the sacrifice of one power stroke out of every two is a frequent deterrent to its selection for use. Cylinder heads become complicated structures because of valve porting, jacketing, and injection-nozzle locations and the accommodation of these to effective combustion, heat transfer, and internal bursting pressures. The valve-in-head construction allows the camshaft to be located in the cylinder head or cylinder block.

Small-size (less than 200 horsepower or 150 kW) engines are conveniently started by an electric motor and storage battery. Larger engines use compressed air (about 200 lb/in.^{2} or 1380 kPa) introduced through valves in the cylinder head. Starting engine-driven generator sets may be accomplished by using the generator as a motor to crank the engine.

Cooling systems use water at 120–180°F (49–82°C) with radiators, cooling towers, and cooling ponds employed for conservation and reclamation. An antifreeze, such as ethylene glycol, and other additives are usually added to the water to prevent freezing and corrosion. Engine mountings and foundations must be designed to handle stress loadings and to reduce vibration. Exhaust systems should include wave-trap silencers or mufflers. Filters for air, fuel, lubricating oil, and (on some engines) cooling-system water help ensure engine reliability. *See* ANTIFREEZE MIXTURE; INTERNAL COMBUSTION ENGINE.

TABLE 2. Approximate allocation of losses in internal combustion engines

Type of loss	Otto engines, %	Diesel engines, %
Output	20	33
Exhaust losses	40	33
Cooling system losses	40	33
Other	<1	1
Total (input)	100	100

Passenger-car engines. Diesel engines have been installed in production automobiles in Europe since 1936. However, their acceptance in the United States was limited until after the oil shortages of 1973 and 1979, when some American manufacturers introduced diesel-powered automobiles to provide an option for improved fuel economy—generally about 25% better than that of a gasoline engine of similar displacement. Easing of the oil shortage in the early 1980s along with an increase in the price of diesel fuel contributed to a decline in public acceptance of the diesel engine.

The diesel engine in the automobile is usually a four-stroke-cycle engine with indirect injection into an auxiliary combustion chamber (**Fig. 2**). Most automobile diesel engines use a distributor-type injection pump. The fuel system often includes a fuel-conditioner assembly, which combines a water-in-fuel detector, water-fuel separator, fuel filter, fuel heater, and hand-priming pump in a single unit. *See* COMBUSTION CHAMBER.

The passenger-car diesel engine may be naturally aspirated, turbocharged, or turbocharged and intercooled. Some automobile diesel engines operate at speeds of up to 5000 rpm. Although the compression ratio with indirect injection ranges from about 20:1 to 23:1, cold starting requires additional heat from an electric glow plug located in each auxiliary chamber. *See* TURBOCHARGER.

The automotive diesel engine usually has four to six cylinders arranged in-line, or six or eight cylinders arranged in a V. Engine displacement ranges from about 1.5 to 7 liters. Some automotive four-cycle gasoline engines have been converted to diesel operation but, in general, these engines have not been as successful in the marketplace as an engine designed as a diesel.

Truck and bus engines. Diesel engines in trucks and buses are usually larger and operate at lower speeds than diesel engines in passenger cars. The typical truck and bus engine ranges in size from about 5 to 19 liters displacement, has a maximum speed of 3000 rpm or less, has direct injection into an open combustion chamber, and may be naturally aspirated, turbocharged, or turbocharged and aftercooled.

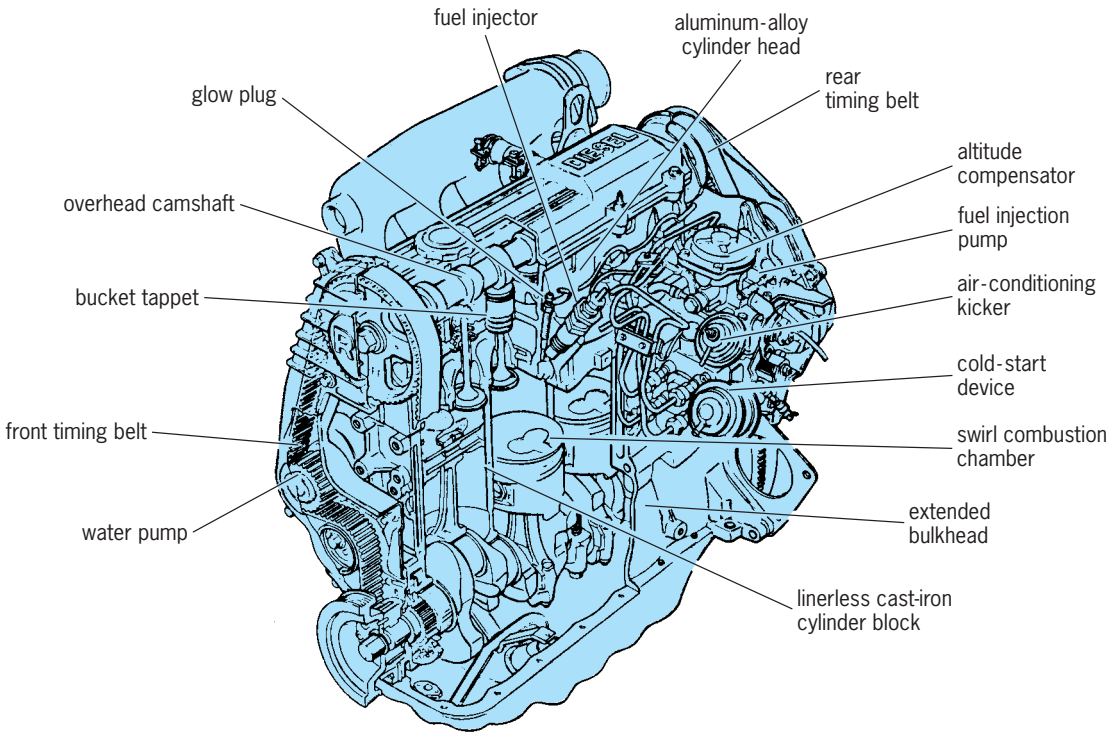

Fig. 2. A 2.0-liter, four-cylinder, four-stroke-cycle passenger-car diesel engine which has a distributor-type injection pump and indirect injection. (*Ford Motor Co.*)

Most truck diesel engines operate on the four-stroke cycle, although many buses and some trucks have two-stroke-cycle engines. These usually have intake ports in the cylinder and exhaust valves in the cylinder head, with scavenging air provided by a crankshaft-driven blower mounted on the crankcase. A unit fuel injector operated by the engine camshaft meters and injects the fuel into the combustion chamber at high pressure at the proper time.

Emissions. In addition to a noticeable odor, the exhaust gas from the diesel engine contains gaseous and particulate emissions which contribute to air pollution. The particles, or soot, may be removed by a trap oxidizer that consists of a filter and a regeneration system, which burns the trapped particles and cleans the filter. Gaseous emissions of unburned hydrocarbons (HC), carbon monoxide (CO), and oxides of nitrogen (NO_x) may be controlled through the use of electronically controlled fuel injection, exhaust-gas recirculation, and charge-air cooling. Operating the engine on low-sulfur fuel reduces sulfur and particulate emissions. *See* AIR POLLUTION; DIESEL FUEL; FUEL INJECTION. Donald L. Anglin

Bibliography. *Bosch Automotive Handbook*, 1986; L. C. R. Lilly (ed.), *Diesel Engine Reference Book*, 1984.

Diesel fuel

A broad mixture of hydrocarbons produced as distillates, as residual materials, or as blends of the two during the refining of crude petroleum. Diesel fuel usually has a distillation range of 390–715°F (200–380°C) and a specific gravity range of 0.760–0.935 [equivalent to 61.2–19.8° on the American Petroleum Institute (API) scale]. In addition to these properties, diesel fuel must have <1 weight % sulfur, < 0.1 weight % ash, < 0.5 volume % water and sediment, and a high flash point (greater than 131°F or 55°C).

Cetane number. Diesel fuel quality is defined by the cetane number, which usually falls into the range 30–60. A high cetane number indicates the potential for easy starting and smooth operation of the engine. The cetane number is the analog of the automobile engine octane number, with cetane (*n*-hexadecane, $C_{16}H_{34}$) having the arbitrarily assigned number of 100. At the other end of the scale, heptamethylnonane, an isomer of cetane, has the assigned cetane number of 0. The cetane number of a diesel fuel is determined by comparison with blends of cetane and heptamethylnonane. It corresponds to the number of parts by volume of cetane in a cetane-heptamethylnonane blend which has the same ignition quality as the fuel. *See* CETANE NUMBER.

For general comparison, the cetane number of paraffinic hydrocarbons, particularly those with straight rather than branched chains, is high, whereas the cetane numbers of aromatic and hydroaromatic (naphthenic) hydrocarbons is low. A fuel having a high octane number tends to have a low cetane number, and vice-versa. Diesel fuels usually contain 60–80 volume % paraffinic materials and 20–40 volume % aromatic-naphthenic materials. Diesel fuels having low cetane numbers (that is, around 60 volume % paraffinics) can be improved by the addition of chemicals such as amyl nitrate ($C_5H_{11}NO_3$) or hexyl nitrate ($C_6H_{13}NO_3$).

Categories. The American Society for Testing and Materials (ASTM) has categorized diesel fuels into three general groups. The need to categorize these fuels results from the varied uses of diesel engines, which are designed to operate efficiently on one of the standard diesel fuels. Use of a fuel which does not meet the designated specifications could have a detrimental effect on engine efficiency and life, and could cause increased emissions and noise (low cetane number), engine wear, and corrosion (high sulfur, high ash, and high sediment). *See* DIESEL ENGINE.

No. 1-D. This is a light distillate, similar to kerosine, for engines where frequent load changes and speed changes (truck, tractor engines) are essential. This fuel has a flash point greater than 100°F (38°C), with a minimum cetane number of 40. This fuel is believed to be particularly suitable for cold-weather operation. *See* KEROSINE.

No. 2-D. A medium distillate fuel with a lower volatility and higher density than No. 1-D, this fuel finds use in heavier-duty engines, for example, railroad engines, which operate at uniform speeds but with heavier loads than encountered during the use of No. 1-D. The flash point is greater than 125°F (52°C) and the minimum cetane number is 40.

No. 3-D. This is a heavy distillate fuel with the highest density and lowest volatility of the three diesel fuels. It finds use in low- and medium-speed engines such as marine engines and electric power generation engines, which operate under sustained loads. The flash point is greater than 130°F (54°C) and the minimum cetane rating is 30.

Other specifications. The result of so-called energy crises was the use of diesel engines in automobiles and the increased use of diesel engines in public transportation vehicles such as buses. These engines require diesel fuel with the highest possible cetane number (60 or more) and could even result in the identification and classification of an additional grade of diesel fuel. *See* PETROLEUM PRODUCTS.

James G. Speight

Bibliography. W. Francis and M. C. Peters, *Fuels and Fuel Technology*, 2d ed., 1980; J. G. Speight, *The Chemistry and Technology of Petroleum*, 3d ed., 1999.

Difference equation

A relationship between one or more independent variables, one or more dependent variables, and differences of those variables. Difference equations arise in the analysis of discrete systems (for example, a string loaded along its length with small masses), in the solution of differential equations by means of digital computers, in the implementation of digital filters, and in the discrete-time control of systems. An ordinary difference equation expresses a relationship between an independent variable t and one or more dependent variables, $y(t)$, $w(t)$, and so forth, and any successive differences of y, w. . . . The first forward difference of y, relative to the increment h, is defined by Eq. (1). The second forward difference

$$\Delta y(t) = y(t+h) - y(t) \tag{1}$$

can be obtained by applying the above definition to itself, as in Eq. (2).

$$\begin{aligned}\Delta^2 y(t) &= \Delta[\Delta y(t)] \\ &= \Delta y(t+h) - \Delta y(t) \\ &= [y(t+2h) - y(t+h)] \\ &\quad -[y(t+h) - y(t)] \\ &= y(t+2h) - 2y(t+h) + y(t)\end{aligned} \tag{2}$$

A difference equation is linear if it is of the first degree with respect all the quantities $\Delta y(t)$, $\Delta^2 y(t)$, $\Delta^3 y(t)$, ... The general solution of difference equations is possible only for the case where they are linear with constant coefficients. For this reason and because of their importance, attention will be initially limited to linear, constant-coefficient difference equations.

Linear constant-coefficient equations. The independent variable will usually have discrete values. Then the notation can be simplified by making the substitution in Eq. (3). In this case, the arguments of

$$\begin{aligned} t_k &= t_0 + kh \\ k &= 0, 1, 2, \ldots \end{aligned} \tag{3}$$

the independent variables, if more than one, can be replaced by the subscript k. Then, the forward difference operators given by Eqs. (1) and (2) can be written as Eqs. (4), where the shifting operator E has

$$\begin{aligned} \Delta y_k &= y_{k+1} - y_k = Ey_k - y_k \\ \Delta^2 y_k &= E^2 y_k - 2Ey_k + y_k \end{aligned} \tag{4}$$

been defined, incrementing the index on the dependent variable by 1 each time it is applied.

An illustration is given by the first-order difference equation (5) with initial condition $y_0 = 1$. It can be

$$y_{k+1} - ay_k = 1 \text{ or } Ey_k - ay_k = 1 \tag{5}$$

$$k \geq 0, a \text{ constant}$$

solved iteratively by putting the term ay_k on the right-hand side and finding the value of y_{k+1} in terms of y_k. Proceeding in this fashion, Eqs. (6) are obtained, or, in general, Eq. (7).

$$\begin{aligned} y_0 &= 1 \\ y_1 &= 1 + a \\ y_2 &= 1 + a + a^2 \\ &\ldots \end{aligned} \tag{6}$$

$$y_k = 1 + a + a^2 + \cdots + a^k = \frac{1 - a^{k+1}}{1 - a} \tag{7}$$

$$a \neq 1$$

The general homogeneous difference equation of order n can be written in terms of the shifting

operator as Eq. (8). From the solution of linear,

$$E^n y_k + A_1 E^{n-1} y_k + \cdots + A_{n-1} E y_k + A_n y_k = 0 \quad (8)$$
$$k \geq 0$$

constant-coefficient differential equations, the solution is assumed to be of the form $e^{rk} = \beta^k$, where r is a suitably chosen constant and $\beta = e^r$. From the definition of the shifting operator, Eqs. (9) follow, as well

$$\begin{aligned} E\beta^k &= \beta^{k+1} = \beta(\beta^k) \\ E^2\beta^k &= \beta^{k+2} = \beta^2(\beta^k) \end{aligned} \quad (9)$$

as similar equations for heigher powers of E. When these equations are substituted into Eq. (8), the result is Eq. (10). Hence, $y_k = \beta^k$ will satisfy Eq. (8) if β is a

$$\beta^n + A_1\beta^{n-1} + \cdots + A_{n-1}\beta + A_n = 0 \quad (10)$$

root of Eq. (10). If all roots are real and distinct, the most general expression satisfying Eq. (8) is Eq. (11),

$$y_k = c_1\beta_1^k + c_2\beta_2^k + \cdots + c_n\beta_n^k \quad (11)$$
$$k \geq 0$$

where $\beta_1, \beta_2, \ldots, \beta_n$ are the roots to Eq. (10) and c_1, $c_2, \ldots, c_n$ are constants to be determined by the initial conditions. *See* ALGEBRA.

An example is given by the second-order linear difference equation (12). In this case Eq. (10) becomes

$$E^2 y_k - 3Ey_k + 2y_k = 0 \quad (12)$$

Eq. (13), with solutions $\beta = 2$ and $\beta = 1$. Thus, the

$$\beta^2 - 3\beta + 2 = 0 \text{ or } (\beta - 2)(\beta - 1) = 0 \quad (13)$$

solution to Eq. (12) is Eq. (14), which follows since $1^k = 1$.

$$y_k = c_1 2^k + c_2 \quad (14)$$
$$k \geq 0$$

If a root is repeated, for example, β_1, then the two solutions corresponding to this repeated root are of the form β_1^k and $k\beta_1^{k-1}$. In the case of the difference equation (15), the corresponding roots of Eq. (10)

$$E^3 y_k - 3E^2 y_k + 4y_k = 0 \quad (15)$$

are -1, 2, and 2, and the solution is then Eq. (16).

$$y_k = c_1(-1)^k + (c_2 + c_3 k)2^k \quad (16)$$
$$k \geq 0$$

Complex roots for Eq. (10) will always occur in complex conjugate pairs for a difference equation with real coefficients. When this occurs, these complex conjugate roots, for example, β_1 and β_2, can be written in polar form as $\rho e^{j\phi}$ and $\rho e^{-j\phi}$ (where $j = \sqrt{-1}$), and the corresponding solution is equivalent to Eq. (17).

$$y_k = \rho^k(c_1 \cos k\phi + c_2 \sin k\phi) \quad (17)$$
$$k \geq 0$$

See COMPLEX NUMBERS AND COMPLEX VARIABLES.

Nonhomogeneous linear equations. The general solution of a nonhomogeneous difference equation, given by Eq. (18), is of the form given by Eq. (19),

$$E^n y_k + A_1 E^{n-1} y_k + A_{n-1} E y_k + A_n y_k = x_k \quad (18)$$
$$k \geq 0$$

$$y_k = y_{k,b} + y_{k,p} \quad (19)$$

where $y_{k,b}$ is the solution of the homogeneous equation [$x_k = 0$ in Eq. (18)], and $y_{k,p}$ is the particular solution of Eq. (18). With constant coefficients on Eq. (18), the method of undetermined coefficients can be used to find the particular solution. In this procedure, three steps are followed in constructing the form of the particular solution:

1. The family of a term represented on the right-hand side of the difference equation (called the forcing function) is constructed. (The family of a term f_k is defined to be the set of all functions of which f_k and all operations $E^m f_k$ are linear combinations. For example, for the forcing function $x_k = e^{bk}$, b a constant, these operations are $E^m e^{bk} = (e^{bm})e^{bk} =$ constant $\times\, e^{bk}$.
2. If that family has no representatives in the homogeneous solution, $y_{k,p}$ is assumed to be a linear combination of the members of that family, and constants of the combination are determined so that the difference equation is identically satisfied.
3. If that family has a representative is the homogeneous solution, each member of the family is multiplied by the smallest integer power of k for which all such representatives are removed, and the particular solution is assumed to be a linear combination of the members of the modified family.

For example, the difference equation (20) corre-

$$E^2 y_k - 3Ey_k + 2y_k = 1 + \alpha^k \quad (20)$$
$$k \geq 0$$

sponds to the homogeneous equation (12) with the homogeneous solution given by Eq. (14). The particular solution is assumed to have the form given by Eq. (21), since 1 is a solution of the homogeneous

$$y_{k,p} = Ak + B\alpha^k \quad (21)$$

equation (and is therefore multiplied by k). Substitution into Eq. (20) gives conditions for the undetermined coefficients A and B which result in Eqs. (22).

$$\begin{aligned} A &= -1 \\ B &= \frac{\alpha}{(\alpha - 1)(\alpha - 2)} \end{aligned} \quad (22)$$

When these equations are substituted into Eq. (21) and the particular solution is used in Eq. (19) along with the homogeneous solution found earlier, the total solution is found to be given by Eq. (23).

$$Y_k = c_1 2^k + c_2 - k + \frac{\alpha^{k+1}}{(\alpha - 1)(\alpha - 2)} \quad (23)$$
$$k \geq 0$$

The solution of constant-coefficient, linear difference equations is facilitated greatly by use of the z transform. *See* TRANSFORM.

Nonlinear difference equations. Nonlinear difference equations arise naturally in the solution of nonlinear differential equations. For simplicity, a first-order ordinary differential equation of the form given by Eq. (24) will be considered. If it appears impossible to integrate Eq. (24), a numerical solution may be attempted by starting with the value y_0 at t_0 and computing the approximate solution y_1 at time $t_0 + h$, the approximate value y_2 at time $t_2 = t_0 + 2h$ and so on. By a Taylor series expansion, Eq. (25), holds,

$$\frac{dy(t)}{dt} = f(t, y) \tag{24}$$

$$y(t_0) = y_0$$

$$\begin{aligned} y(t+h) &= y(t) + hy'(t) + \frac{h^2}{2}y''(t) + \cdots \\ &= y(t) + hf(t, y) + \frac{h^2}{2}f'(t, y) + \cdots \end{aligned} \tag{25}$$

where the prime indicates differentiation with respect to t and Eq. (24) has been used to get the second equation. For small values of h, a crude approximation is given by the first two terms of Eq. (25), as in Eq. (26). By using the index notation introduced earlier, the solution to Eq. (24) can be approximated by the numerical algorithm given by Eq. (27). This is called the Euler method for solving the differential Eq. (24). Because the expansion of Eq. (25) for $y(t + h)$ is truncated at the term linear in h, the Euler method referred to as first order. For small h, the error induced by this truncation is of the order of h^2. *See* DIFFERENTIATION.

$$y(t+h) = y(t) + hf \tag{26}$$

$$y_{k+1} = y_k + hf(k, y_k) \tag{27}$$

$$k = 0, 1, 2, \ldots$$

An improvement in the accuracy of the Euler method is provided by the modified Euler method, which is expressed by Eqs. (28). This is called the improved Euler method, and can be viewed as approximating the integration by a straight line through the point (t_k, y_k) with slope $f(t_k, y_k)$ in the first half interval and by a straight line of slope $f(t_{k+1}, y^*_{k+1})$ in the second half interval. Thus, it is called a predictor-corrector method, and has a truncation error of the order of h^3.

$$\begin{aligned} y^*_{k+1} &= y_k + hf(t_k, y_k) \\ y_{k+1} &= y_k + {}^1\!/_2 h[f(t_k, y_k) + f(t_{k+1}, y^*_{k+1})] \end{aligned} \tag{28}$$

The Runge-Kutta method is a fourth-order method and has truncation error of the order h^5. It is implemented by Eqs. (29).

$$\begin{aligned} k_1 &= hf(t_k, y_k) \\ k_2 &= hf(t_k + {}^1\!/_2 h, y_k + {}^1\!/_2 k_1) \\ k_3 &= hf(t_k + {}^1\!/_2 h, y_k + {}^1\!/_2 k_2) \\ k_4 &= hf(t_k + h, y_k + k_3) \\ y_{k+1} &= y_k + {}^1\!/_6(k_1 + 2k_2 + 2k_3 + k_4) \end{aligned} \tag{29}$$

The above integration methods are one-step in that the integration is based on values of $f(t, y)$ in a single interval. Multistep methods provide even greater accuracy but require values for y_k from more than one interval. Therefore, one of the challenges in starting them is to provide the necessary initial values, which is typically done with a one-step method.

Methods also exist for integrating nonlinear differential equations of order higher than one. For example, for the second-order differential equation (30), an algorithm may be based on Eqs. (31), where the prime denotes differentiation with respect to time and the value for y_k can be substituted from Eq. (30). A more accurate method is provided by a generalization of the Runge-Kutta method.

$$y''(t) = f(t, y, y') \tag{30}$$

$$\begin{aligned} y_{k+1} &= y_k + hy'_k + \frac{h^2}{2}y''_k \\ y'_{k+1} &= y'_k + hy''_k \end{aligned} \tag{31}$$

Solutions of equations for nonlinear systems based on the above integration algorithms are ill-posed in that an incorrect steady-state value may be reached. Alternative schemes based on state variables avoid this problem and are faster.

Approximating continuous-time systems. Methods exist for approximating governing equations for linear continuous-time systems by difference equations. For example, a continuous-time second-order linear system represented by the differential

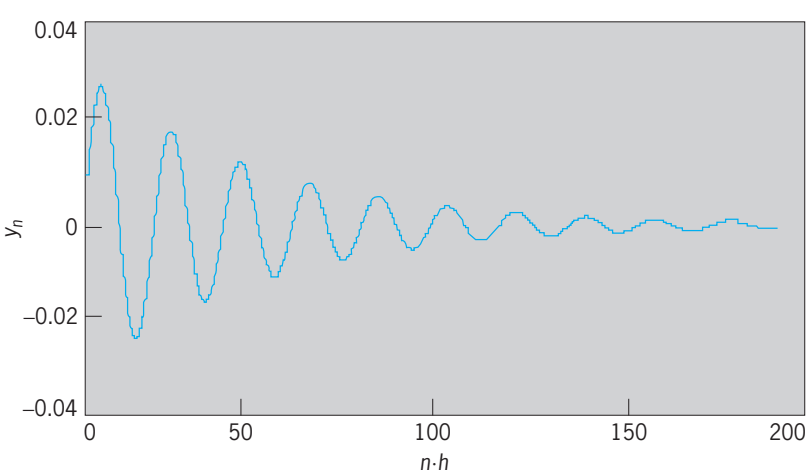

Numerical solution of Eq. (32) for $a = 0.05$, $b = 0.1$, $c = 0.01$, and a step input. The solution uses the state-variable representation of Eq. (33) and the difference-equation approximation of Eq. (36) with $h = 0.1$.

equation (32) can be recast in state-variable form as Eqs. (33). The state equations [the first two equations

$$\frac{d^2y(t)}{dt^2} + a\frac{dy(t)}{dt} + by(t) = c\frac{dx(t)}{dt} \tag{32}$$

$$\begin{aligned} \frac{dq_1(t)}{dt} &= -aq_1(t) - q_2(t) + cx(t) \\ \frac{dq_2(t)}{dt} &= bq_1(t) \\ y(t) &= q_1(t) \end{aligned} \tag{33}$$

of Eqs. (33)] can be written in matrix form as Eq. (34), where the matrices **A** and **B** are given by Eqs. (35).

$$\dot{\mathbf{Q}}(t) = \mathbf{AQ}(t) + \mathbf{Bx}(t) \tag{34}$$

$$\mathbf{A} = \begin{bmatrix} -a & -1 \\ b & 0 \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} c \\ 0 \end{bmatrix} \tag{35}$$

Approximation of these equations using Euler integration results in Eqs. (36), which can be used to obtain a numerical solution (see **illus.**).

$$\begin{aligned} q_{1,k+1} &= (1 - ha)q_{1,k} - hq_{2,k} + hcx_k \\ q_{2,k+1} &= jhbq_{1,k} + q_{2,k} \\ y_k &= q_{1,k} \end{aligned} \tag{36}$$

Partial difference equations. The numerical solution of a partial differential equation leads to a partial difference equation. The procedure is similar to that used for numerical solution of ordinary differential equations. The Laplace equation (37) is an example of a simple partial differential equation. By using Taylor's expansion, the first partial derivatives of u with respect to x and y can be approximated by Eqs. (38). Similarly, the second derivatives can be

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{37}$$

$$\begin{aligned} \frac{\partial u(x,y)}{\partial x} &\approx \frac{1}{2h}[u(x+h,y) - u(x-h,y)] \\ \frac{\partial u(x,y)}{\partial y} &\approx \frac{1}{2k}[u(x,y+k) - u(x,y-k)] \end{aligned} \tag{38}$$

approximated by Eqs. (39). Substitution into the

$$\begin{aligned} \frac{\partial^2 u(x,y)}{\partial x^2} &\approx \frac{1}{h^2}[u(x+h,y) - 2u(x,y) \\ &\qquad + u(x-h,y)] \\ \frac{\partial^2 u(x,y)}{\partial y^2} &\approx \frac{1}{k^2}[u(x,y+k) - 2u(x,y) \\ &\qquad + u(x,y-k)] \end{aligned} \tag{39}$$

Laplace equation yields the difference Eq. (40). In

$$\begin{aligned} u(x+h,y) + u(x,y+h) + u(x-h,y) \\ + u(x,y-h) - 4u(x,y) = 0 \end{aligned} \tag{40}$$

the case of the Poisson equation, which is Laplace's equation with the right-hand-side 0 replaced by the known function $f(x, y)$, the difference equation (40) has the right-hand side replaced by $h^2 f(x, y)$. In Eq. (40), h is called the mesh size, and the equation itself relates $u(x, y)$ to the values of u at the four nearest-neighbor positions. In fact, for the Laplace equation, u at (x, y) equals the mean of the values at the four neighboring positions. *See* DIFFERENTIAL EQUATION; DIGITAL FILTER; INTERPOLATION; LAPLACE'S DIFFERENTIAL EQUATION; LINEAR SYSTEM ANALYSIS; NUMERICAL ANALYSIS. Rodger E. Ziemer

Bibliography. S. Goldberg, *Introduction to Difference Equations*, 1958, reprint 1986; F. B. Hildebrand, *Methods of Applied Mathematics*, 2d ed., 1965, reprint 1992; F. B. Hildebrand, *Introduction to Numerical Analysis*, 2d ed., 1974, reprint 1987; W. G. Kelley and A. C. Peterson, *Difference Equations: An Introduction with Applications*, 2d ed., 2000; E. Kreyszig, *Advanced Engineering Mathematics*, 8th ed., 1998; J. M. Smith, *Mathematical Modeling and Digital Simulation for Engineers and Scientists*, 2d ed., 1987; R. E. Ziemer, W. H. Tranter, and D. R. Fannin, *Signals and Systems: Continuous and Discrete*, 4th ed., 1998.

Differential

A mechanism which permits a rear axle to turn corners with one wheel rolling faster than the other. An automobile differential is located in the case carrying the rear-axle drive gear (see **illus.**).

The differential gears consist of the two side gears carrying the inner ends of the axle shafts, meshing with two pinions mounted on a common pin located in the differential case. The case carries a ring gear driven by a pinion at the end of the drive shaft. This arrangement permits the drive to be carried to both wheels, but at the same time as the outer wheel on a turn overruns the differential case, the inner wheel lags by a like amount.

Special differentials (Dana, Borg Warner, Eaton) permit one wheel to drive the car by a predetermined amount even though the opposite wheel is on slippery pavement; they have been used on racing cars for years and are now used by a number of car manufacturers.

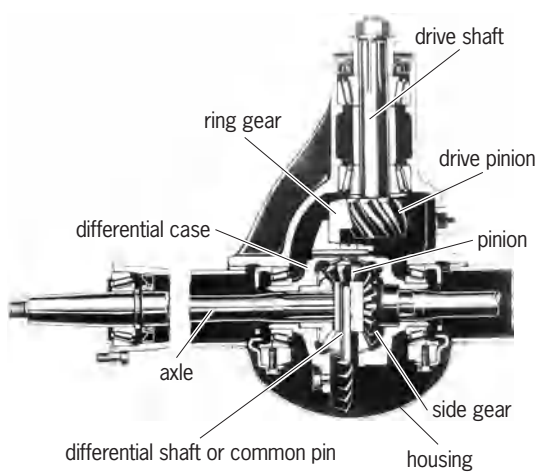

Commonly used rear-axle differential. (*Chrysler Corp.*)

In operation, engine torque applied to the differential case causes the angular contacts on the case to bear on corresponding angles on the differential pinion pins. Two contacts 180° apart push one pin to the right; the other two contacts, 180° apart, spaced 90° from the first, push the other pin to the left. On one of the pins are two pinions meshing with the right-side gear; on the other are two pinions meshing with the left-side gear. Diameters on the pinions concentric with the teeth bear against the side gear rings splined to the inner end of the axle shafts and force the rings to apply clutches which connect the differential case directly to the rings. The pinions also load the side gears axially, the latter contributing their thrusts to the clutches. The drive thus passes from the differential case to the rings splined to the respective rear-axle shafts, providing a frictional drive to one wheel, even though the other is on ice or slippery pavement. This nullifies the conventional differential action by the predetermined amount. *See* AUTOMOTIVE TRANSMISSION; GEAR. Harold Fischer

Bibliography. D. W. Dudley (ed.), *Handbook of Practical Gear Design*, 1994; B. A. Shtipelman, *Design and Manufacture of Hypoid Gears*, 1978.

Differential amplifier

An electronic circuit that is designed to amplify the difference between two voltages measured with respect to a common reference, usually designated as ground. By convention, the net difference of two voltages measured with respect to a common reference is called the differential-mode voltage, while the sum of the voltages, usually divided by two to give an average value, is called the common-mode voltage.

An ideal differential amplifier thus has exactly the same gain from each input to its output, and the amplifier produces an output that is directly proportional to its differential-mode voltage. The amplifier delivers zero output in response to common-mode voltages. If these gains are not exactly equal, then equal (common-mode) voltages applied at each input terminal will not be equal at the amplifier output and their difference will not cancel completely. The common-mode gain, the ratio of the output response of a real differential amplifier to the input signal applied equally to each input terminal, is a measure of this gain mismatch.

Differential amplification is very useful when the signal to be amplified exists in an electrically noisy environment, since the noise voltage is usually a common component of both input voltages and, hence, will cancel when the difference of the amplifier inputs is taken.

Operation. For a physical differential amplifier to work properly, the electrical paths of each input signal through the amplifier must be nearly identical. Thus, the most important requirement for a differential amplifier is that it be constructed with transistors with closely matched electrical characteristics. Integrated circuits with amplifier transistors physically close to each other meet the required close matching requirement and are ideally suited for the production of differential amplifiers. *See* INTEGRATED CIRCUITS.

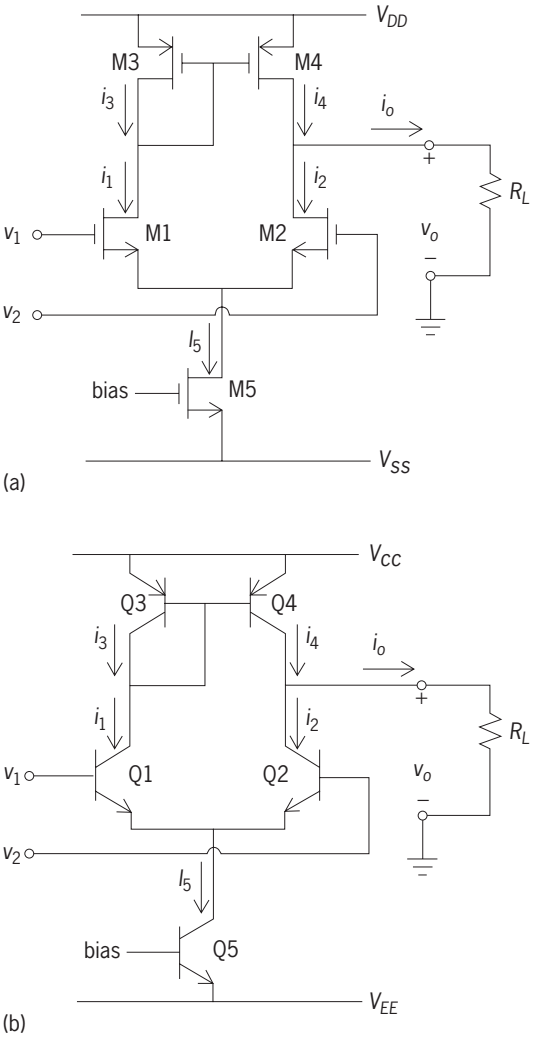

Differential amplifier using (*a*) metal-oxide-semiconductor field-effect transistors (MOSFETs), M1–M5, and (*b*) bipolar junction transistors (BJTs), Q1–Q5. V_{DD}, V_{CC} = positive power supplies; V_{SS}, V_{EE} = negative power supplies; v_1, v_2 = input voltages; M1(Q1), M2(Q2) = input transistors with drain (collector) currents i_1, i_2; M3(Q3), M4(Q4) = load transistors forming current mirror with equal source (emitter) currents, i_3 and i_4; transistor M5(Q5) functions as current sink with value I_5; I_O = output current; v_O = output voltage; R_L = load resistance.

Differential-amplifier circuits that are suitable for integrated-circuit fabrication can use either metal-oxide-semiconductor field-effect transistors (MOSFETs; **illus.** *a*) or bipolar junction transistors (BJTs; illus. *b*). The input transistor pair must be matched closely. For best performance, the two load transistors also should be matched. *See* TRANSISTOR.

A fifth transistor functions as a dc current sink whose value, I_5, is determined by the voltage applied to the gate (base) of this transistor. The positive and negative power supplies are normally equal in magnitude. When the input voltages are equal, the drain (collector) currents of the input transistors are

identical and equal to exactly half of the dc current sink. Since the source (emitter) current of the left load transistor equals the drain (collector) current of the left input transistor, it follows that if the former current is half of the dc current sink, the latter current has this same value. The function of the load transistor pair is to provide two equal currents [the source (emitter) currents of these transistors]. This circuit is called a current mirror because it forces one of these currents to equal the other. Under the above conditions, both the drain (collector) current of the right input transistor and the source (emitter) current of the corresponding member of the current-mirror pair equal half the dc current sink. These two currents leave and enter the output terminal, respectively, and therefore no current flows through the output (and the output voltage is zero). *See* CURRENT SOURCES AND MIRRORS.

If the left input voltage is increased and the right remains constant, part of the increase in the left input voltage will cause an increase in the drain (collector) current of the left input transistor, to which this voltage is applied. The other part will cause a decrease in the drain (collector) current of the right input transistor. The mirroring function of the load transistor pair forces the source (emitter) current of the right load transistor to equal that of the left, which in turn equals the increased drain (collector) current of the left input transistor. Summing currents at the output terminal of the differential amplifier shows that the output current must equal the sum of the increase in the drain (collector) current of the left input transistor and the decrease in the corresponding current of the right input transistor. If the left input voltage is decreased or the right input voltage is increased, the changes become negative and the output current becomes equal to the negative of this sum.

Gain. The output current flows through the resistance seen at the output terminal, R_1, to create the output voltage of the differential amplifier. This output voltage is referenced to ground and is called the single-ended output voltage. The maximum output resistance is equivalent to the parallel combination of the small-signal drain-source resistances of the right input transistor or the collector-emitter resistances of the corresponding current-mirror transistor. Typical voltage gains from the differential input to the signal-ended output range from 50 to 200 for the MOSFET differential amplifier and 500 to 2000 for the BJT differential amplifier. The BJT differential amplifier has a larger gain because the transconductance of the BJT is approximately 10 times greater than the transconductance of the MOSFET. The gain of the BJT differential amplifier can be increased by using cascode transistors between the right input transistor and the corresponding current-mirror transistor and the output. *See* CASCODE AMPLIFIER; GAIN.

Frequency response. The frequency response of the differential amplifier is dependent on the reciprocal product of the output resistance and the capacitance in parallel with this output resistance. Typical values of output resistance range from 50 to 500 kΩ. For a 10-picofarad capacitor connected to the output, the frequency where the gain has decreased by −3dB from its low-frequency value ranges from 32 to 320 kHz. As the load capacitor increases, this frequency will decrease. *See* RESPONSE.

Limitations. The above considerations are appropriate when the current changes discussed above are small compared to the dc current sink. When large changes are applied to either or both of the input voltages, all the dc current sink will flow through one side of the differential amplifier only. For example, if the left input voltage is much larger than the right, the drain (collector) current of the left input transistor is equal to the dc current sink and the corresponding current of the right input transistor is equal to zero. Therefore, the current mirror forces the output current to equal the dc current sink. Similarly, if the left input voltage is much less than the right, the output current becomes the negative of the dc current sink. Thus, the maximum sinking or sourcing current of the differential amplifier output is $+I_5$. When a load capacitor, C_L, is attached to the output, the voltage change rate will be limited to a value of $+I_5/C_L$. This rate limitation is called slew rate and depends on the values of I_5 and C_L.

An important characteristic of the differential amplifier is the input common-mode voltage range. This is the common-mode voltage over which the differential amplifier can still function as a differential amplifier without any degradation. The upper limit is the common-mode voltage at which the current-mirror transistors no longer operate as a current mirror because the voltage across them becomes too small. The lower limit depends on the bias setting of the transistor that functions as a dc current sink for the same reason. A differential amplifier will have an input common-mode voltage range of about 2–3V less than the difference between the positive and negative power-supply voltages.

If the output changes when the same signal is applied to the two input voltages, the common-mode gain of the differential amplifier is not zero. The causes of common-mode gain are mismatches between the input transistors, imperfect current mirroring, and a finite value of resistance of the dc current sink implemented by the fifth transistor. If the output is not zero when a dc voltage is applied to both the input voltages, the differential amplifier has a dc offset. This dc offset at the output is divided by the differential to single-ended voltage gain and reflected to the input of the differential amplifier, where it is called the input offset voltage. The input offset voltage can be as large as 15 mV for MOSFET differential amplifiers and 5 mV for BJT differential amplifiers. *See* AMPLIFIER; OPERATIONAL AMPLIFIER.

Phillip E. Allen

Bibliography. R. L. Geiger, P. E. Allen, and N. R. Strader, *VLSI Design Techniques for Analog and Digital Circuits*, 1990; P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3d ed., 1993; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Differential equation

A relationship between a function and its derivatives.

If there is one independent variable, the differential equation is called an ordinary differential equation. The general form of such an equation is shown in Eq. (1), where t is the independent variable; u

$$F(t, u, u', u'', \ldots, u^{(n)}) = 0 \tag{1}$$

is a function of t; $u' = du/dt$, $u'' = d^2u/dt^2, \ldots,$ $u^n = d^nu/dt^n$ are the derivatives of u; and $F(t, u_0, u_1, \ldots, u_n)$ is a given function of the $n + 2$ variables $t, u_0, \ldots, u_n$. The positive integer n is called the order of the differential equation; that is, the order of the differential equation is the order of the highest derivative that occurs in the equation. As an example, consider Eq. (2), which is an or-

$$u'' + u = 0 \tag{2}$$

dinary differential equation of order 2. A function u is said to be a solution to Eq. (1) if, when u and its derivatives up to order n are substituted into F, the identity expressed in Eq. (1) is valid for t in some interval $a < t < b$. In Eq. (2), $u = \cos t$ and $u = \sin t$ are solutions for $-\infty < t < \infty$, as can be immediately verified.

If there are two or more independent variables, the equation is called a partial differential equation. Again, the order of the equation is the order of the highest partial derivative that occurs in the equation. Thus, the general form of a partial differential equation of order 2, with two independent variables, is given by Eq. (3), where x and y are the indepen-

$$F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) = 0 \tag{3}$$

dent variables; $u = u(x,y)$ is a function: $u_x = \partial u/\partial x$, $u_y = \partial u/\partial y$, $u_{xx} = \partial^2 u/\partial x^2$, $u_{xy} = \partial^2 u/\partial x \partial y$, and $u_{yy} = \partial^2 u/\partial y^2$ are the derivatives of u; and $F(x,y,u,p,q,r,s,t)$ is a given function of eight variables. For example, Eq. (4) is a partial differential equation of

$$u_{xy} = 0 \tag{4}$$

order 2. It is easily verified that $u = h(x) + g(y)$ is a solution of Eq. (4) for any choice of functions h and g.

A differential equation is linear if the function u and its derivatives appear linearly in the equation, and is nonlinear otherwise. Thus, Eqs. (2) and (4) are linear, whereas Eqs. (5), (6), and (7) are nonlinear.

$$u' = \sin u \tag{5}$$

$$(1 + u_y^2)\, u_{xx} - 2u_x u_y u_{xy} + (1 + u_x^2) u_{yy} = 0 \tag{6}$$

$$u_t + uu_x + u_{xxx} = 0 \tag{7}$$

Partial differential equations occur with more than two independent variables; see Eqs. (33), (35), and (36). Another possibility is a system of differential equations. This consists of two or more equations involving one or more unknown functions which are to be solved simultaneously. For example, Eqs. (8) are a first-order linear system of ordinary dif-

$$\begin{aligned} u' &= v \\ v' &= -u \end{aligned} \tag{8}$$

ferential equations. The Cauchy-Riemann equations, Eqs. (9), are a first-order linear system of partial dif-

$$\begin{aligned} u_x - v_y &= 0 \\ u_y + v_x &= 0 \end{aligned} \tag{9}$$

ferential equations. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

Ordinary differential equations. The process of solving differential equations is called integration, and is analogous to the solution of the simple equation $u'(t) = h(t)$, which is actually the indefinite integral of h. The value of an indefinite integral is unique except for an arbitrary constant of integration. In general, the solution of an ordinary differential equation of order n involves n arbitrary constants of integration. For example, the general solution of Eq. (2) is $u(t) = A \cos t + B \sin t$, where A and B are arbitrary constants; and the solution of Eq. (5) is $u(t) = 2 \tan^{-1} (ke^t)$, where k is an arbitrary constant. *See* INTEGRATION.

In the full generality of Eq. (1), it is not possible to prove that solutions exist. For example, Eq. (10) has

$$1 + \left| \frac{du}{dt} \right|^2 = 0 \tag{10}$$

no solution. However, if an equation is of the type of Eq. (11), and if $f(t,u_0, \ldots, u_{n-1})$ has continuous

$$u^{(n)} = f(t, u, \ldots, u^{(n-1)}) \tag{11}$$

first derivatives on the set defined by $|t - t_0| < d$, $|u_0 - a_0| < d$, $|u_1 - a_1| < d, \ldots, |u_{n-1} - a_{n-1}| < d$, where $a_1, \ldots, a_n$, and d are constants, then there is one and only one function $u(t)$ defined and continuously differentiable on the interval $|t - t_0| < \epsilon$, for some $\epsilon > 0$, which is a solution of Eq. (11) and which satisfies the initial conditions $u(t_0) = a_0$, $u'(t_0) = a_1, \ldots, u^{(n-1)}(t_0) = a_{n-1}$. This result is purely an existence result. That is, it says a solution exists but does not say how to find it. Indeed, it may easily be impossible to find a solution which is expressible in terms of elementary functions.

Equations of first order. It is usual in the case of equations of first order to use x as the independent variable and to denote the unknown function by $y = y(x)$. Then Eq. (1) reduces to $dy/dx = f(x,y)$. This equation is usually written in differential form, Eq. (12), where $f = -P/Q$. In this notation, Eq. (5)

$$P(x, y)dx + Q(x, y)dy = 0 \tag{12}$$

would be rewritten as Eq. (13). For a given func-

$$\sin y dx - dy = 0 \tag{13}$$

tion $v(x,y)$ the differential of v is defined by $dv = v_x dx + v_y dy$. The first-order equation is said to be exact if there is a function v such that $dv = Pdx + Qdy$,

that is, if $\partial\upsilon/\partial x = P$ and $\partial\upsilon/\partial y = Q$. In this case the solution to Eq. (12) is given implicitly by the equation $\upsilon(x,y) = c$, with c an arbitrary constant. For example, Eq. (14) is exact, since $d(xy) = ydx + xdy$, and

$$ydx + xdy = 0 \tag{14}$$

the solution is $xy = c$. A general example of exact equations is given by the case when, in Eq. (12), P is a function of x only and Q is a function of y only. Then, if $\upsilon(x,y) = \int P(x)dx + \int Q(y)\,dy$, the solution is $\upsilon(x,y) = c$. For example, Eq. (15) is exact and has

$$dx - \frac{dy}{\sin y} = 0 \tag{15}$$

the solution $x - \log\tan(y/2) = c$.

When Eq. (12) is not exact, it may be possible to find an integrating factor, that is, a function $\mu(x,y)$ such that the equation $\mu Pdx + \mu Qdy$ is exact. If $\mu Pdx + \mu Qdy = dv$, then, once more, the solution is given by $\upsilon(x,y) = c$. Theoretically it is always possible to find an integrating factor, provided P and Q are not both equal to zero at some point. In practice, however, it may be difficult. For Eq. (16), which is not exact, an integrating factor is

$$ydx - xdy = 0 \tag{16}$$

$1/y^2$, since $d(x/y) = (ydx - xdy)/y^2$, and the solution is given by $x/y + c$. A general example is provided by the case of separable variables when, in Eq. (12), $P(x,y) = A(x)B(y)$ and $Q(x,y) = C(x)D(y)$. Then $\mu = 1/C(x)B(y)$ is an integrating factor, since $[A(x)/C(x)]\,dx + [D(y)/B(y)]\,dy = 0$ is exact. Using this method, Eq. (13), and hence Eq. (5), can be reduced to Eq. (15). As another example, consider Eq. (17), a linear first-order differential equation, or in differential form, Eq. (18). It is not exact, but $\mu(x) =$

$$\frac{dy}{dx} + a(x)y = z(x) \tag{17}$$

$$dy + [a(x)y - z(x)]dx = 0 \tag{18}$$

$\exp\left[\int a(x)dx\right]$ is an integrating factor. After multiplication by μ, Eq. (18) becomes $dv = 0$, where

$$v(x,y) = y\exp\left[\int a(x)dx\right] - \int z(x)\exp\left[\int a(x)dx\right]dx$$

Thus the general solution of Eq. (17) is given by Eq. (19), where c is an arbitrary constant.

$$y(x) = \exp\left[-\int a(x)dx\right]\cdot\left\{\int z(x)\exp\left[\int a(x)dx\right]dx + c\right\} \tag{19}$$

Second-order linear equations. The study of oscillation in mechanical systems and of simple electric circuits leads to linear ordinary differential equations of second order, typified by Eq. (20). This equation is said

$$u'' + a(t)u' + b(t)u = \upsilon \tag{20}$$

to be homogeneous when υ is identically equal to zero; otherwise, it is nonhomogeneous. The linearity of the equation has important implications for the sets of solutions to Eq. (20). If u_1 and u_2 are two linearly independent solutions to the homogeneous equation, then $Au_1 + Bu_2$ is also a solution for any choice of constants A and B. Furthermore, any solution of the homogeneous equation is of this form. Since the difference of two solutions to the nonhomogeneous equation is a solution to the homogeneous equation, it follows that the general solution to the nonhomogeneous equation has the form of expression (21), where u_p is a particular solution to the

$$u_p + Au_1 + Bu_2 \tag{21}$$

nonhomogeneous equation, u_1 and u_2 are linearly independent solutions to the homogeneous equation, and A and B are arbitrary constants. As an example, consider Eq. (22). The corresponding homogeneous

$$u'' + u = t \tag{22}$$

equation is Eq. (2). The functions $u_1(t) = \cos t$ and $u_2(t) = \sin t$ are linearly independent solutions to Eq. (2); thus the general solution to Eq. (2) is $u(t) = A\cos t + B\sin t$. It is easy to find a particular solution to Eq. (22), namely, $u_p(t) = t$. Hence the general solution to Eq. (22) is $u(t) = t + A\cos t + B\sin t$.

When the coefficients $a(t)$ and $b(t)$ in Eq. (20) are constants, solutions to the homogeneous equation exist of the form $u(t) = e^{rt}$, where r is a constant which may be complex. Substitution into Eq. (20) [with $\upsilon = 0$] yields the quadratic equation $r^2 + ar + b = 0$, which may be solved for r. If $a^2 - 4b > 0$, there are two distinct real roots, and thus two linearly independent solutions to the homogeneous equation exist. If $a^2 - 4b = 0$, there is only one root. A second solution in this case is $u(t) = te^{rt}$. If $a^2 - 4b < 0$, the roots are complex conjugate, $r_1 = \alpha + i\beta$, $r_2 = \alpha - i\beta$. Then two real linearly independent solutions to the homogeneous equation are $e^{\alpha t}\cos\beta t$ and $e^{\alpha t}\sin\beta t$. *See* DAMPING; ELECTRIC TRANSIENT; FORCED OSCILLATION; HARMONIC MOTION.

If the coefficients of Eq. (20) are nonconstant, it is sometimes difficult to find solutions to the homogeneous equation. However, if two linearly independent solutions u_1 and u_2 to the homogeneous equation are known, the method of variation of parameters yields solutions to the nonhomogeneous equation. The idea is to look for a solution to Eq. (20) of the form $A(t)u_1(t) + B(t)u_2(t)$, subject to the condition that $A'u_1 + B'u_2 = 0$. After solving for A and B, formula (23) is obtained for a particular solution to Eq. (20), where the kernel $k(t,s)$ is given by Eq. (24).

$$u_p(t) = \int_{t_0}^{t} k(t,s)v(s)ds \tag{23}$$

$$k(t,s) = \frac{u_2(t)u_1(s) - u_1(t)u_2(s)}{u_1(s)u_2'(s) - u_2(s)u_1'(s)} \tag{24}$$

If the coefficients $a(t)$ and $b(t)$ of Eq. (20) are

analytic functions (that is, if a and b have power series expansions

$$a(t) = \sum_{n=0}^{\infty} a_n t^n, \quad b(t) = \sum_{n=0}^{\infty} b_n t^n$$

which converge for t near zero), then power series solutions to the homogeneous equation can be found. The idea is to set

$$u(t) = \sum_{n=0}^{\infty} u_n t^n$$

Then

$$u'(t) = \sum_{n=0}^{\infty} n u_n t^{n-1} \text{ and } u''(t) = \sum_{n=0}^{\infty} n(n-1) u_n t^{n-2}$$

The series for a, b, u, u', and u'' are substituted into the homogeneous equation, resulting in a power series which is identically equal to zero. Setting the coefficients equal to zero results in a sequence of recursion relations for the coefficients u_n. This method can be extended to the case for which not $a(t)$ and $b(t)$ themselves but $ta(t)$ and $t^2b(t)$ are analytic functions. This extension is essential since many important equations have this feature. For example, the solution of Bessel's equation, Eq. (25), is Bessel's function of order n, shown in Eq. (26), where Γ de-

$$t^2u'' + tu' + (t^2 - n^2)u = 0 \tag{25}$$

$$J_n(t) = \left(\frac{t}{2}\right)^n \sum_{k=0}^{\infty} \frac{(-1)^k (t/2)^{2k}}{\Gamma(k+1)\Gamma(n+k+1)} \tag{26}$$

notes the gamma function of Euler. Another example is the hypergeometric equation of Gauss, Eq. (27). If

$$t(1-t)u'' + [c - (a+b+1)t]u' - abu = 0 \tag{27}$$

c is not zero or a negative integer, a solution is the hypergeometric function shown in Eq. (28).

$$F(a,b,c,t) = 1 + \frac{a \cdot b}{1 \cdot c} t + \frac{a(a+1)b(b+1)}{1 \cdot 2 \cdot c(c+1)} t^2 + \cdots \tag{28}$$

See BESSEL FUNCTIONS; GAMMA FUNCTION.

Partial differential equations. The analysis of many physical problems leads to linear partial differential equations of second order. When there are two independent variables x and y, the most general equation of this sort is Eq. (29), where the coefficients $a, \ldots,$

$$au_{xx} + 2bu_{xy} + cu_{xx} + du_x + eu_y + fu = g \tag{29}$$

g are given functions of x and y. It will be assumed that $a^2 + b^2 + c^2 \neq 0$. Then Eq. (29) is called elliptic if $b^2 - ac < 0$, parabolic if $b^2 - ac = 0$, and hyperbolic if $b^2 - ac > 0$. For each of these types there are certain typical boundary value problems which arise in physical applications and which are of mathematical interest. When there are more than two independent variables, or when the order of the equation is larger than two, there are equations which have the same qualitative properties, as do these three types, and for which similar boundary value problems exist. These equations are accordingly called elliptic, parabolic, or hyperbolic. However, some equations do not belong to any of these three categories.

Elliptic equations. The prototype is Laplace's equation, Eq. (30), for two independent variables, and Eq. (31) for three independent variables. The solu-

$$u_{xx} + u_{yy} = 0 \tag{30}$$

$$u_{xx} + u_{yy} + u_{zz} = 0 \tag{31}$$

tions are called harmonic functions. A typical boundary value problem is the Dirichlet problem. It consists of finding a harmonic function $u(x,y)$ in a region in the plane [say, the disk $D = \{(x,y) | x^2 + y^2 < 1\}$] with prescribed values on the boundary of the region [$u(x,y) = g(x,y)$ for $x^2 + y^2 = 1$, where g is a given function]. This corresponds to the physical problem of finding the steady-state temperature distribution $u(x,y)$ in a plate when the temperature distribution on the edge of the plate is known. *See* LAPLACE'S DIFFERENTIAL EQUATION.

Parabolic equations. The prototype is the heat equation, shown in one space variable x in Eq. (32), and in two space variables x and y in Eq. (33). The vari-

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0 \tag{32}$$

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 0 \tag{33}$$

able t represents time. If $u(t,x)$ is the temperature in a rod modeled by the interval $a \leq x \leq b$, and if there is no source of heat in the rod, then u is a solution to Eq. (32). A typical boundary value problem is the initial value problem, which consists of finding a solution to Eq. (32) for $0 < t < \infty$ and $-\infty < x < \infty$ with $u(0,x) = f(x)$ given. The physical interpretation consists of finding the temperature distribution in an infinite rod for all time when the initial temperature distribution is known. *See* CONDUCTION (HEAT).

Hyperbolic equations. The prototype is the wave equation, shown for one, two, and three space variables in Eqs. (34), (35), and (36), respectively. Again, t rep-

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 0 \tag{34}$$

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 0 \tag{35}$$

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial z^2} = 0 \tag{36}$$

resents time. Equation (34) is also called the equation of the vibrating string; Eq. (35) is called the equation of the vibrating membrane. As these names indicate, a large variety of oscillatory phenomena lead to hyperbolic equations. A typical boundary value problem is the Cauchy problem, which involves finding a solution $u(t, x)$ to Eq. (34) for $0 < t < \infty$ and $-\infty$

$< x < \infty$, with $u(0, x) = f(x)$ and $u_t(0, x) = g(x)$ given. If $u(t, x)$ is interpreted as the displacement of an infinite string, then the Cauchy problem consists of deriving $u(t, x)$ for all time when the initial displacement $u(0,x) = f(x)$ and the initial velocity $u_t(0, x) = g(x)$ are both known. *See* VIBRATION; WAVE EQUATION; WAVE MOTION.

Other equations. An example of a partial differential equation which is not of the types previously considered is the Korteweg-deVries (KdV) equation, Eq. (7), which describes shallow-water waves in a one-dimensional channel. The KdV equation is a nonlinear equation of the third order. Another nonlinear equation of interest is Eq. (6), which is called the minimal surface equation; its solution represents surfaces of minimal surface area (for example, soap films spanning wire frames have minimal surface area). *See* CALCULUS; DIFFERENTIATION. John C. Polking

Bibliography. D. Bleecker and G. Csordas, *Basic Partial Differential Equations*, 1997; R. Courant and D. Hilbert, *Methods of Mathematical Physics*, vol. 1, 1953, paper 1989 and vol. 2, 1962, paper 1989; S. J. Farlow, *An Introduction to Differential Equations and Their Applications*, 1994; F. John, *Partial Differential Equations*, 4th ed., 1982, reprint 1991; M. Krusemeyer, *Differential Equations*, 1994; G. F. Simmons, *Differential Equations, with Applications and Historical Notes*, 2d ed., 1991; W. A. Strauss, *Partial Differential Equations: An Introduction*, 1994.

Differential geometry

A branch of mathematics that deals with intrinsic properties of curves and surfaces in three-dimensional euclidean space. The intrinsic properties are those which are independent of the geometrical object's orientation or location in space. The subject is also concerned with nets of curves and families of surfaces, these having wide application in the arts.

The space is referred to a rectangular cartesian coordinate system (x, y, z). A space curve may be defined by a pair of independent equations $f(x, y, z) = 0$ and $g(x, y, z) = 0$ or, more meaningfully, by parametric equations: $x = x(t)$, $y = y(t)$, $z = z(t)$. In this case the arc length between two points t_0 and t is given by Eq. (1).

$$s = \int_{t_0}^{t} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dv}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\, dt \qquad (1)$$

Obviously, if s is chosen as a parameter, Eq. (2) holds. In this case dx/ds, dy/ds, dz/ds at a

$$\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 + \left(\frac{dz}{ds}\right)^2 = 1 \qquad (2)$$

point P are the direction cosines of the tangent to the curve at P.

Plane curves. Consider three nearby points P, P_1, P_2. Through them, in general, one plane may be constructed. The limiting position of this plane as P_1 and P_2 approach P is the osculating plane of the curve at P. The limiting position of the circle through P, P_1, P_2 as P_1 and P_2 approach P is the osculating circle. Its center is the center of curvature and its radius is the radius of curvature ρ. The reciprocal of the radius of curvature κ is the curvature; its value is shown in expression (3). The perpendicular to the tangent at

$$\sqrt{\left(\frac{d^2x}{ds^2}\right)^2 + \left(\frac{d^2y}{ds^2}\right)^2 + \left(\frac{d^2z}{ds^2}\right)^2} \qquad (3)$$

P in the osculating plane is the principal normal and the perpendicular at P to the osculating plane is the binormal. Another way of defining curvature is this: Let P' be a point near P, arc $PP' = \Delta s$. Let $\Delta\theta$ be the angle between the tangents at P and P'. Then Eq. (4) holds.

$$\frac{1}{\rho} = \lim_{P' \to P} \frac{\Delta\theta}{\Delta s} \qquad (4)$$

Similarly, if $\Delta\phi$ is the angle between the binormals at P and P', expression (5) is the torsion $1/\tau$ of the

$$\lim_{P' \to P} \frac{\Delta\phi}{\Delta s} \qquad (5)$$

curve at P, where τ is the radius of torsion. The torsion at P is the rate at which the curve is twisted out of the osculating plane so that the torsion of a plane curve is zero. The values of the curvature and torsion in terms of the arc determine the curve up to its position in space. The determination of the parametric equations of the curve, as well as of the solution of most problems dealing with space curves, is accomplished by means of the Frenet equations. These equations give the values of the derivatives of the direction cosines of the tangent, normal, and binormal in terms of these cosines and the curvature and torsion of the curve.

Plane curves are, of course, special cases of space curves ($\tau = 0$). One of the problems treated extensively in classical differential geometry is that of nets of curves. Such a net is given by two independent equations, Eqs. (6). For fixed values of v, Eqs. (6)

$$x = x(u, v) \qquad y = y(u, v) \qquad (6)$$

define a family of curves, the u lines, whereas fixed values of u give the v lines. Through each point there is one u line and one v line so that (u,v) may be regarded as curvilinear coordinates in the plane.

If this system is to be orthogonal, it is necessary that Eq. (7) hold.

$$\frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \frac{\partial y}{\partial u}\frac{\partial y}{\partial v} = 0 \qquad (7)$$

Equations (6) may also be interpreted as the mapping of the uv plane on the xy plane. In applications one seeks a mapping under which certain properties of figures are preserved. If arc length is to be preserved, the mapping must be a motion; if area is to be preserved, it is necessary that Eq. (8) hold.

$$\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v} = 1 \qquad (8)$$

Probably the most important mappings are the conformal ones. Under them the angles between lines are preserved; such a mapping is given by $x + iy = f(u + iv)$, where f is any analytic function of the complex variable $u + iv$ $(i = -1)$.

Surfaces. A surface in three-dimensional euclidean space may be given by $f(x, y, z) = 0$ or, more conveniently, by Eqs. (9). For a fixed value of v, Eqs. (9)

$$x = x(u, v) \qquad y = y(u, v) \qquad z = z(u, v) \tag{9}$$

describe a curve in the surface, a u line, and similarly for a fixed value of u. Thus (u,v) are curvilinear coordinates of a point in the surface. The most important quantity in the study of surfaces is the arc length of a curve. This is given by Eq. (10), where E, F, $G$are

$$ds^2 = E\,du^2 + 2F\,du\,dv + G\,dv^2 \tag{10}$$

functions of partial derivatives of (x, y, z). In the real domain E, G, and $EG - F^2$ are nonnegative and may be zero only at singular points of the surface, or at points where the matrix of the partial derivatives of (x, y, z) is of rank less than two. The right-hand side of Eq. (10) is called the first fundamental form of the surface, where E, F, G are the first fundamental quantities of the surface. All applicable surfaces have the same first fundamental form. Thus any cylinder or cone which is developable (applicable to a plane) has the fundamental form $du^2 + dv^2$, where (u, v) are rectangular cartesian coordinates. Under an arbitrary transformation of the surface coordinates (u, v) the fundamental quantities E, F, G transform linearly so that many problems in the study of surfaces reduce to the question of whether a coordinate system exists on the surface in which E, F, G satisfy desired conditions.

In order to distinguish between various surfaces having the same first fundamental form, one considers the plane tangent to the surface at a point M of coordinates (u, v) and the perpendicular distance p from the point M' of coordinates $(u + du, v + dv)$ to this tangent plane. The principal part of $2p$ is then expression (11), where D, D', D'' are functions of

$$D\,du^2 + 2D'\,du\,dv + D''\,dv^2 \tag{11}$$

the first and second partial derivatives of (x, y, z) and are therefore expressible as functions of u and v. Expression (11) is called the second fundamental form of the surface. E, F, G and D, D', D'' are not independent but must satisfy the so-called Gauss-Codazzi equations. If both forms are given, the surface is determined up to its position in space.

At each nonsingular point P of a surface, there is one tangent plane and one surface normal. A plane in a given direction containing this normal cuts out a normal section of the surface. The curvature of this curve at P is the normal curvature of the surface in the direction of the plane. As the normal plane rotates about the surface normal as an axis, the normal curvature usually changes; it takes on its maximum and its minimum values for two directions, and these are at right angles to each other. These are the principal directions at P. If all normal curvatures at P have the same value, P is called an umbilical point. Thus every point on a sphere is umbilical, and a general ellipsoid has four umbilical points. The curves on the surface whose tangents have a principal direction at every point are the lines of curvature of the surface. They are often chosen as the coordinate lines, for in that case $F = 0$, $D' = 0$. Denoting by ρ_1 and ρ_2 the radii of curvature of the principal normal sections at P, Eq. (12) is called the mean curvature of the surface

$$K_m = \frac{1}{\rho_1} + \frac{1}{\rho_2} \tag{12}$$

at P and $K = 1/\rho_1\rho_2$ is the total or gaussian curvature of the surface at P. The value of K is given by Eq. (13). Asthus defined, K_m and K depend on the

$$K = \frac{DD'' - D'^2}{EG - F^2} \tag{13}$$

enveloping space. However K. F. Gauss showed that K is expressible in terms of E, F, G and their first and second partial derivatives. Thus K is an intrinsic invariant of the surface; it has the same value for all applicable surfaces. For a plane $K = 0$, for a sphere $K = 1/a^2$, and for a pseudosphere $K = -(1/a^2)$.

Through each point of a surface in a given direction there is one curve whose osculating plane coincides with the normal plane in that direction. These curves are the geodesics of the surface. In a properly restricted neighborhood of a point there is just one geodesic through two points in that neighborhood; it is the shortest distance between the two points. From this it follows that the geodesics are tangent to the curve at the points s and $s + \Delta s$. This curvature κ_g is measured in the surface and is in general different from the curvature of the curve as measured in space. Obviously κ_g is zero for any geodesic, so that geodesics play the role of straight lines in the plane.

n-Dimensional applications. The idea of obtaining all intrinsic metric properties of a surface from the first fundamental form was extended by G. F. B. Riemann to a space of n dimensions. In this type of space a point has the coordinates $(u_1, u_2, \ldots, u_n)$ and the arc length is given by Eq. (14). Such a space is called

$$ds^2 = \sum_{i,j=1}^{n} g_{ij}(u)du_i\,du_j \tag{14}$$

riemannian and the functions $g_{ij}(u)$ are the fundamental quantities for the space. All the intrinsic properties of a surface have been generalized to a riemannian space though not in a trivial or obvious manner. It was the knowledge of this geometry that enabled Albert Einstein to work out the general theory of relativity, for according to that theory, physical space is four dimensional, the $g_{ij}(u)$ being determined by the matter and energy in the space. *See* CONFORMAL MAPPING; COORDINATE SYSTEMS; PROJECTIVE GEOMETRY; RELATIVITY; RIEMANNIAN GEOMETRY; TENSOR ANALYSIS.

Morris S. Knebelman

Bibliography. J. Beem and K. L. Duggal (eds.), *Differential Geometry and Mathematical Physics*, 1994; R. Caddeo and F. Triceri, *Differential Geometry and*

Topology, 1993; A. Helfer, *Introduction to Modern Differential Geometry*, 1991; S. P. Novikov and A. T. Fomenko, *Basic Elements of Differential Geometry and Topology*, 1991; J. J. Stoker, *Differential Geometry*, 1989; A. Visconti and A. Visconti, *Introduction to Differential Geometry for Physicists*, 1992.

Differential transformer

An iron-core transformer with movable core. A differential transformer produces an electrical output voltage proportional to the displacement of the core. It is used to measure motion and to sense displacements. It is also used in measuring devices for force, pressure, and acceleration that are based on the conversion of the measured variable to a displacement.

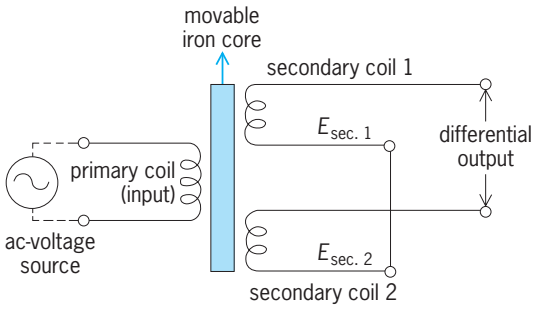

Fig. 1. Basic circuit of differential transformer.

Various available configurations, some translational and others rotational, all employ the basic circuit of **Fig. 1**: a primary winding, two secondary windings, and a movable core. The primary winding is energized with alternating voltage. The two secondary windings are connected in series opposition, so that the transformer output is the difference of the two secondary voltages. When the core is centered, the two secondary voltages are equal and the transformer output is zero. This is the balance of null position. When the core is displaced from the null point, the two secondary voltages are no longer alike and the transformer produces an output voltage. With proper design, the output voltage varies linearly with core position over a small range. Motion of the core in the opposite direction produces a similar effect with 180° phase reversal of the alternating output voltage.

The amplitude of the ac output voltage forms a V-shaped curve when plotted against core position; the phase angle abruptly reverses by 180° at the null point. When the bottom of the V-curve is examined in closer detail, it is seen that the output voltage at balance is not exactly zero. The small residual null voltage consists of higher harmonics of the input frequency, as well as a fundamental frequency component 90 degrees out of phase (called the quadrature component).

Electronic signal conditioning is commonly employed to eliminate the residual and to make the transducer usable with standard dc instrumentation.

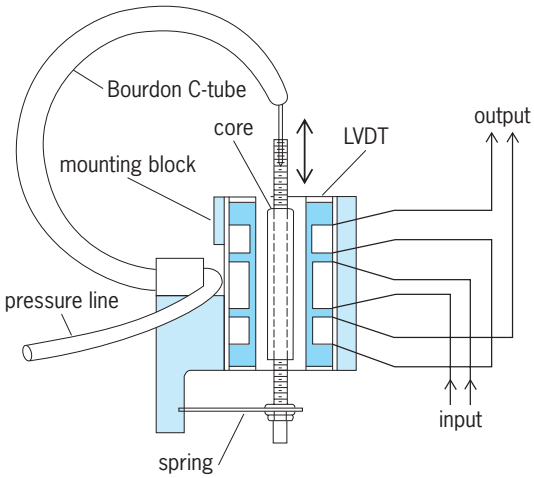

Fig. 2. Bourdon-tube pressure transducer employing a linear variable differential transformer (LVDT). (*After E. E. Herceg, Handbook of Measurement and Control, Schaevitz Engineering, 1976*)

The electronic circuit consists of an ac oscillator (carrier generator) to drive the input windings, plus a demodulator and an amplifier to convert the output into dc. The microelectronics can be built into the transformer housing, and the resulting package is sold as a dc-LVDT.

The principal advantages of the differential transformer, which are shared with capacitance transducers, over other displacement transducers, such as the resistance potentiometer, are absence of contacts and infinite resolution. No friction is introduced by the measurement, and movement smaller than a microinch (25 nanometers) can be sensed. The separation between coil and core makes the differential transformer useful in difficult and dangerous environments. Stability of the null makes it ideal as a null sensor in self-balancing devices and servomechanisms. Typical applications are machine tool inspection and gaging, pressure measurement (**Fig. 2**), liquid level control, load cells, and gyroscopic instruments.

The linear variable differential transformer (LVDT; **Fig. 3**) is the commercially prevalent form. A rotary version is also manufactured. The E-pickoff (**Fig. 4**) is an older device. Its linearity is not as good, and its principal use is as a null sensor. Both translational

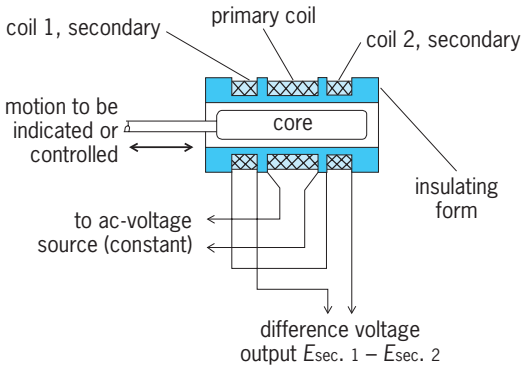

Fig. 3. Linear variable differential transformer (LVDT).

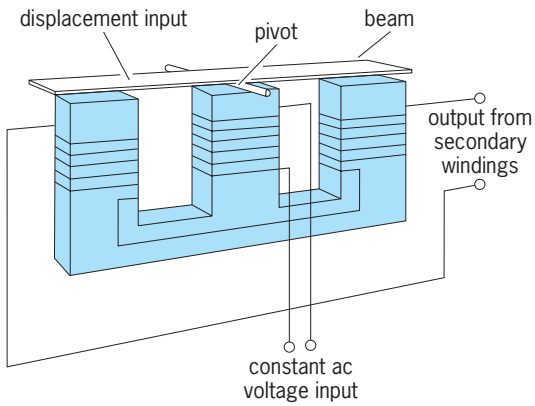

Fig. 4. E-shaped differential transformer. (*After P. J. O'Higgins, Basic Instrumentation, McGraw-Hill, 1966*)

and rotational E-pickoffs are made. *See* TRANSDUCER; TRANSFORMER. Gerald Weiss

Bibliography. J. Carstens, *Electrical Sensors and Transducers*, 1997; E. O. Doebelin, *Measurement Systems*, 5th ed., 2003.

Differentiation

A mathematical operation performed on a function to determine the effect of a change in the value of the independent variable. Functions of one variable are considered in this article. For differentiation of functions of several variables *see* PARTIAL DIFFERENTIATION.

If f is a function of x, defined on an interval containing x_0, the derivative at x_0 is, by definition, that shown in Eq. (1). For generalities about derivatives

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \tag{1}$$

and calculus *see* CALCULUS.

In the quotient on the right, in the definition of $f'(x_0)$, x is restricted to the interval on which f is defined. If $y = f(x)$, $f'(x_0)$ is also denoted by expression(2). The limit defining $f'(x_0)$ may not exist. If it

$$\left(\frac{dy}{dx}\right)_{x=x_0} \tag{2}$$

does, f is called differentiable at x_0.

The derivative of $f'(x)$, called the second derivative, is denoted by $f''(x)$ or d^2y/dx^2.

A function f is called continuous at x_0 if x_0 is in the domain of f and if Eq. (3) holds. The precise

$$\lim_{x \to x_0} f(x) = f(x_0) \tag{3}$$

formulation of the limit concept used here is the following: Let g be a function and let A be a number. Then Eq. (4) means that to each positive number

$$\lim_{x \to x_0} g(x) = A \tag{4}$$

ϵ corresponds to some positive number δ such that $|g(x) - A| < \epsilon$ whenever x is a number in the domain of g such that $x \neq x_0$ and $|x - x_0| < \delta$. It is required of g that its domain shall contain numbers x as close to x_0 as desired, but different from x_0. The domain of g may also contain x_0, but this is irrelevant.

It is a theorem that, if f is differentiable at x_0, then f is continuous at x_0. However, a function can be continuous at x_0, but not differentiable there. An example is $f(x) = |x|$ at $x_0 = 0$. It is even possible for a function to be continuous on an interval and yet not differentiable at any point of this interval.

The chief elementary applications of differentiation are (1) in expressing rates of change (velocity, acceleration), and in solving problems where, through functional relationship, the rate of change of one variable is calculated when the rate of change of another variable is known; (2) in studying graphs of functions and, more generally, in studying curves in the plane or in space of three dimensions; (3) in expressing scientific laws or principles in the form of differential equations; and (4) in the expression of various extensions and applications of the law of the mean, including such topics as l'Hospital's rule and Taylor's formula or series.

Principles. The general technique of differentiation is built upon the rules for differentiating combinations of differentiable functions. If $u = f(x)$ and $v = g(x)$ are functions differentiable for the same values of x, then $u + v$ and uv are differentiable and so is u/v if $v \neq 0$. The formulas are shown in Eqs. (5). If

$$\begin{aligned} \frac{d}{dx}(u+v) &= \frac{du}{dx} + \frac{dv}{dx} \\ \frac{d}{dx}(uv) &= u\frac{dv}{dx} + v\frac{du}{dx} \\ \frac{d}{dx}\left(\frac{u}{v}\right) &= \frac{v\dfrac{du}{dx} - u\dfrac{dv}{dx}}{v^2} \end{aligned} \tag{5}$$

u is constant in value, then $du/dx = 0$. A very powerful instrument of technique is furnished in the chain rule for composite functions. If y is a differentiable function of u and u is a differentiable function of x, then y is a differentiable function of x, and Eq. (6)

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} \tag{6}$$

results. In functional notation, if $y = f(u)$ and $u = h(x)$, then $y = F(x)$, where $F(x) = f[h(x)]$, and then $f'(x) = f'[h(x)]h'(x)$.

The technique of differentiation also leans upon facts about inverse functions. If $y = f(x)$, where f is a differentiable function and $f'(x)$ is either always positive or always negative on an interval, for example, when $a \leqq x \leqq b$, then for each y from $f(a)$ to $f(b)$ inclusive there is just one x such that $a \leqq x \leqq b$ and $y = f(x)$. Thus there is defined a function g such that $x = g(y)$ is equivalent to $y = f(x)$ with x and y restricted as indicated. This function g is differentiable, and Eq. (7) results. An example: $y = \sin x$,

$$g'(y) = \frac{1}{f'(x)} = \frac{1}{f'[g(y)]} \tag{7}$$

$x = \sin^{-1} y$ (sometimes written $x = \arcsin y$), where $-\pi/2 \leqq x \leqq \pi/2$ and $-1 \leqq y \leqq 1$. Here $f'(x) = \cos x$, and Eq. (8) gives the derivative of the inverse function.

$$\frac{d}{dy}(\sin^{-1} y) = \frac{1}{\cos x} = \frac{1}{\sqrt{1-\sin^2 x}} = \frac{1}{\sqrt{1-y^2}} \tag{8}$$

Algebraic and transcendental functions. The functions studied in elementary calculus are of two kinds, algebraic and transcendental. The basic differentiation formula for algebraic functions is $dy/dx = nx^{n-1}$ if $y = x^n$, where n is any rational number. This rule may be combined with the chain rule and used in connection with the rules for dealing with sums, products, and quotients. The differentiation of algebraic functions in general may require use of implicit function theorems. The elementary transcendental functions are the trigonometric functions, the logarithm functions, and their inverses. For a discussion of implicit function theorems *see* PARTIAL DIFFERENTIATION.

Trigonometric functions. In calculus the trigonometric functions are defined on the assumption that angles are measured in radians, so that $\sin x$ means the sine of x radians. Differentiation of $f(x) = \sin x$ is based on the fact that $\sin x/x \to 1$ as $x \to 0$. This is equivalent to $f'(0) = 1$. By combining this result with trigonometric identities, the two basic formulas shown in Eqs. (9) are derived. Then the derivatives of the other

$$\frac{d}{dx}\sin x = \cos x \qquad \frac{d}{dx}\cos x = -\sin x \tag{9}$$

functions are worked out by using the rule for quotients. The results are shown in Eqs. (10). These formulas may also be combined with the chain rule; an example is given in Eq. (11).

$$\begin{aligned} \frac{d}{dx}\tan x &= \sec^2 x \qquad \frac{d}{dx}\cot x = -\csc^2 x \\ \frac{d}{dx}\sec x &= \sec x \tan x \qquad \frac{d}{dx}\csc x = -\csc x \cot x \end{aligned} \tag{10}$$

$$\frac{d}{dx}\sin u = \cos u \frac{du}{dx} \tag{11}$$

The inverse trigonometric functions are differentiated by the methods for inverse functions, as explained earlier. The definitions are as follows:

$y = \sin^{-1} x$ means $x = \sin y$ and $-\frac{\pi}{2} \leqq y \leqq \frac{\pi}{2}$
$y = \cos^{-1} x$ means $x = \cos y$ and $0 \leqq y \leqq \pi$
$y = \tan^{-1} x$ means $x = \tan y$ and $-\frac{\pi}{2} < y < \frac{\pi}{2}$
$y = \cot^{-1} x$ means $x = \cot y$ and $0 < y < \pi$

The differentiation formulas are shown in Eqs. (12).

$$\begin{aligned} \frac{d}{dx}\sin^{-1} x &= \frac{1}{\sqrt{1-x^2}} \\ \frac{d}{dx}\cos^{-1} x &= \frac{-1}{\sqrt{1-x^2}} \\ \frac{d}{dx}\tan^{-1} x &= \frac{1}{1+x^2} \\ \frac{d}{dx}\operatorname{ctn}^{-1} x &= \frac{-1}{1+x^2} \end{aligned} \tag{12}$$

The inverses of the secant and cosecant functions are little used, and there is no standard usage about the definitions needed to make them single-valued.

Exponentials and logarithms. These two functions go together, a logarithm function being the inverse of an exponential function, or vice versa. The traditional treatment of these functions in differential calculus was for many years as follows: The nature of exponentials was assumed as known from algebra, and the definition $y = \log_a x$ if $x = a_y$ (where $a > 0$, $a \neq 1$, and $x > 0$), as well as the algebraic properties of logarithms, was also assumed as known. As a first step toward the differentiation of a logarithm function, it was traditional to show that $(1 + t)^{1/t}$ approaches a limit, denoted by e, as t approaches 0. Moreover, $2 < e < 3$, and $e = 2.718$, approximately. It can then be shown that Eq. (13) is valid. This formula suggests the advantage of choosing $a = e$ as the base for logarithms. The base e logarithm of x, called the natural logarithm of x, is denoted by $\log x$ with no subscript, or by $\ln x$. (The latter notation is favored by engineers and many physical scientists.) Then $y = \ln x$ is equivalent to $x = e^y$, and $y = e^x$ is equivalent to $x = \ln y$. The derivative of e^x can be worked out by the rule for inverse functions. The simple formulas are given in Eqs. (14). For an arbitrary base a, where

$$\frac{d}{dx}\log_a x = \frac{1}{x}\log_a e \tag{13}$$

$$\frac{d}{dx}\ln x = \frac{1}{x} \qquad \frac{d}{dx}e^x = e^x \tag{14}$$

$a \neq 1$ and $0 < a$, the formulas can be written as Eqs. (15). It is often convenient to know that $\log_a x =$

$$\frac{d}{dx}\log_a x = \frac{\log_a e}{x} \qquad \frac{d}{dx}a^x = a^x \log_e a \tag{15}$$

$(\log_a b)(\log_b x)$. In particular, with $a = 1$, one obtains $\log_a b = (\log_b a)^{-1}$.

The modern trend is to recognize that neither exponentials nor logarithms are adequately defined and studied by students before they come to calculus. In any case, satisfactory discussion of these functions requires the theory of limits. Hence it is reasonable to use calculus itself to define logarithms and develop their properties. When this is done, the development proceeds as follows: The "natural logarithm function" is defined when $x > 0$ by Eq. (16).

$$\ln x = \int_1^x \frac{dt}{t} \tag{16}$$

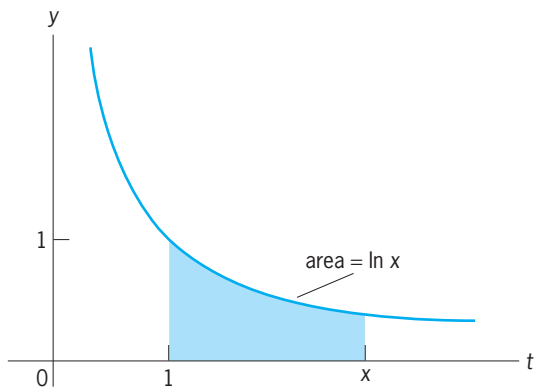

The natural logarithm function.

This means that $\ln x$ is the area between the curve $y = 1/t$ and the t axis from $t = 1$ to $t = x$, reckoned as positive if $x > 1$ and as negative if $0 < x < 1$ (see **illus.**).

By use of the fundamental properties of integrals and the relation between differentiation and integration, it can be shown quite directly that $\ln 1 = 0$, $d/dx \ln x = 1/x$, and that $\ln (AB) = \ln A + \ln B$. Moreover, $\ln x$ is a continuous function of x which increases as x increases and is such that $\ln x \to +\infty$ as $x \to +\infty$ and $\ln x \to -\infty$ as $x \to 0$. There is then a well-defined inverse function (call it the E function) such that $y = \ln x$ is equivalent to $x = E(y)$ if $x > 0$. The number $E(1)$ is denoted by e; that is, e is the unique positive number such that $\ln e = 1$. The rules about inverse functions show that $E'(x) = E(x)$. If $a > 0$ and x is an integer, it is possible to show easily that $a^x = E(x \ln a)$. This formula serves as the general definition of a^x when x is not an integer, and the results are consistent with what is expected, so that the usual exponent laws are valid. The general definition of logarithms, from this point of view, is Eq. (17). In

$$\log_a x = \frac{\ln x}{\ln a} \tag{17}$$

particular, $\log_e x = \ln x$, because $\ln e = 1$.

In this method of developing properties of exponentials and logarithms by calculus, the fact that $(1 + t)^{1/t} \to e$ as $t \to 0$ is not needed, but it can be proved, more easily than in the older traditional development, after everything else has been worked out.

The number e enters naturally into the concept of continuous compounding of interest. If the sum $\$P$ is placed at interest at the nominal rate of $100r\%$, compounded n times a year, the accumulated sum $\$S$ after t years is shown in Eq. (18). If now n is increased indefinitely, so that interest is compounded more and more frequently, the fact that Eq. (19) holds

$$S = P\left(1 + \frac{r}{n}\right)^{nt} \tag{18}$$

$$\lim_{n\to\infty}\left(1 + \frac{r}{n}\right)^{n/r} = e \tag{19}$$

shows that in the limit of continuously compounded interest, the formula for the accumulated sum after t years is $S = Pe^{rt}$. This formula is characterized by differential equation (20), which expresses what is

$$\frac{dS}{dt} = rS \tag{20}$$

sometimes called the law of natural growth. In general, if y depends on x in such a way that $dy/dx = ky$, where k is a nonzero constant, then $y = y_0 e^{kx}$, where $y = y_0$ when $x = 0$. This situation occurs in radioactive decay and in many types of growth and diminution processes in chemistry and natural science.

Applications. Several important applications of differentiation are discussed in this section.

Velocity and acceleration. If a point moves on the x axis, with coordinate x at time t, its velocity is dx/dt and its acceleration is d^2x/dt^2. When the point with coordinates (x, y) moves in the xy plane, its velocity and acceleration are vectors, the x and y components of velocity are dx/dt and dy/dt, respectively, and the corresponding components of acceleration are d^2x/dt^2, d^2y/dt^2. These results are extended naturally for a point (x, y, z) moving in three-dimensional space.

For the plane case, if the point has polar coordinates (r, θ), the velocity and acceleration can be resolved into components in the r and θ directions. The r direction at (r, θ) is the direction directly away from the origin (r increasing and θ constant). The θ direction is 90° from the r direction in the counterclockwise sense. The r and θ components of velocity are, respectively, dr/dt and $r(d\theta/dt)$, The corresponding components of acceleration are expressions (21) and (22). These are useful in discussing the motion of a

$$\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2 \tag{21}$$

$$r\frac{d^2\theta}{dt^2} + 2\frac{dr}{dt}\frac{d\theta}{dt} \tag{22}$$

particle under the influence of a central force, as in the case of a single mass moving in the gravitational field of a fixed center of attraction according to the inverse-square law.

Still another useful way of resolving acceleration into components involves the tangential and normal directions to the path. Here one needs to deal with arc length and curvature. If s is arc length measured along a plane curve in a preassigned positive sense and if K is curvature, then the velocity is a vector whose component along the curve is ds/dt, the component at right angles to the curve being zero. The component of acceleration along the curve in the positive sense is d^2s/dt^2, and at right angles to the curve (90° counterclockwise from the positive sense along the curve) is v^2K, where $v = ds/dt$. The curvature of a circle is the reciprocal of its radius R, and so a point moving around a circle of radius R with speed v experiences an acceleration toward the center of amount v^2/R. This is called centripetal acceleration. If the speed is not uniform, there is also an acceleration along the curve. *See* PARAMETRIC EQUATION.

Curve tracing. If $f'(x) > 0$, $f(x)$ increases as x increases, whereas $f(x)$ decreases as x increases when $f'(x) < 0$. Points where $f'(x) = 0$ are called critical points. With x and y axes in the usual position (x positive to the right, y positive toward the top of the page), the graph of $y = f(x)$ is called concave upward over an interval of x values if the tangent line turns counterclockwise as x increases. If the tangent line turns clockwise as x increases, the curve is called concave downward. If $f'(x) > 0$, the curve is concave upward, and it is concave downward if $f''(x) < 0$. It follows that y is at a relative maximum if $f'(x) = 0$ and $f'(x) < 0$, at a relative minimum if $f'(x) = 0$ and $f'(x) > 0$. A point where the concavity changes from one sense to the other is called a point of inflection. Such a point is a point of relative maximum or minimum for $f'(x)$. A sufficient condition for a point of inflection is that $f''(x) = 0$, $f'''(x) = 0$.

Simple harmonic motion. A point moving on the x axis in such a way that $d^2x/dt^2 = -\omega^2 x$, where $\omega > 0$, is executing what is called simple harmonic motion. If a point travels around a circle with constant speed, its orthogonal projection on a diameter executes simple harmonic motion with the center of the circle as the origin of coordinates on the diameter. The methods of calculus enable one to infer from the foregoing differential equation that x depends on t by the formula $x = A \cos(\omega t + \alpha)$, where A and α are constants. The moving point completes one cycle of its motion in time $2\pi/\omega$.

Law of the mean. For a discussion of the simple case of this law *see* CALCULUS.

An extended version, often called Cauchy's formula, is the following: Suppose F and G are continuous when $a \leqq x \leqq b$, differentiable when $a < x < b$, and suppose $G(b) \neq G(a)$. Suppose also that $f'(x)$ and $G'(x)$ are never both zero together. Then there is some number X such that $a < X < b$ and Eq. (23) is satisfied.

$$\frac{F(b) - F(a)}{G(b) - G(a)} = \frac{F'(X)}{G'(X)} \tag{23}$$

Rule of l'Hospital. This rule for finding the limit of a quotient of two functions in certain circumstances is stated as follows: Subject to certain general conditions of f and g, if either $f(x) \to 0$ and as $x \to a$ or $g(x) \to +\infty$ or $g(x) \to -\infty$ a, then Eq. (24) is valid,

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)} \tag{24}$$

provided the limit on the right exists as a finite limit or as either $+\infty$ or $-\infty$. The general conditions are that f and g are differentiable as $x \to a$ and neither $g(x)$ nor $g'(x)$ is 0 as $x \to a$.

Taylor's formula. This formula, Eq. (25), is a finite sum

$$f(x) = f(a) + \frac{f'(a)}{1!}(x - a) + \cdots + \frac{f^{(n)}(a)}{n!}(x - a)^n + R_n \tag{25}$$

expression for $f(x)$, where R_n is called the remainder. The two most useful formulas for R_n are Eq. (26), integral form, and Eq. (27), Lagrange's form. In the

$$R_n = \frac{1}{n!} \int_a^x f^{(n+1)}(t)(x - t)^n \, dt \tag{26}$$

$$R_n = \frac{f^{(n+1)}(X)}{(n + 1)!}(x - a)^{n+1} \tag{27}$$

latter form, X is some number between x and a. If, for fixed a and x, $R_n \to 0$ as $n \to \infty$ $f(x)$ is said to be represented by Taylor's infinite series expansion. *See* INTEGRATION; OPERATOR THEORY; SERIES.

Angus E. Taylor

Bibliography. M. H. Protter and P. E. Protter, *Calculus and Analytical Geometry*, 4th ed., 1988; S. K. Stein and A. Barcellos, *Calculus and Analytic Geometry*, 5th ed., 1992.

Diffraction

The bending of light, or other waves, into the region of the geometrical shadow of an obstacle. More exactly, diffraction refers to any redistribution in space of the intensity of waves that results from the presence of an object that causes variations of either the amplitude or phase of the waves. Most diffraction gratings cause a periodic modulation of the phase across the wavefront rather than a modulation of the amplitude. Although diffraction is an effect exhibited by all types of wave motion, this article will deal only with electromagnetic waves, especially those of visible light. Some important differences that occur with microwaves will also be mentioned. For discussion of the phenomenon as encountered in other types of waves *see* ELECTRON DIFFRACTION; NEUTRON DIFFRACTION; SOUND. *See also* ELECTROMAGNETIC WAVE.

Diffraction is a phenomenon of all electromagnetic radiation, including radio waves; microwaves; infrared, visible, and ultraviolet light; and x-rays. The effects for light are important in connection with the resolving power of optical instruments. *See* RADIOWAVE PROPAGATION; X-RAY DIFFRACTION.

There are two main classes of diffraction, which are known as Fraunhofer diffraction and Fresnel diffraction. The former concerns beams of parallel light, and is distinguished by the simplicity of the mathematical treatment required and also by its practical importance. The latter class includes the effects in divergent light, and is the simplest to observe experimentally. A complete explanation of Fresnel diffraction has challenged the most able physicists, although a satisfactory approximate account of its main features was given by A. Fresnel in 1814. At that time, it played an important part in establishing the wave theory of light.

To illustrate the difference between methods of observation of the two types of diffraction, **Fig. 1** shows the experimental arrangements required to observe them for a circular hole in a screen s. The light originates at a very small source O, which can conveniently be a pinhole illuminated by sunlight. In

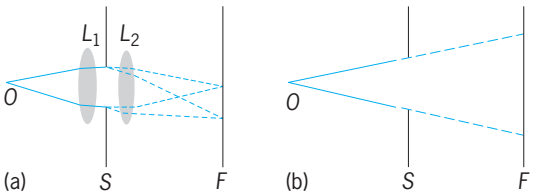

Fig. 1. Observation of the two principal types of diffraction, in the case of a circular aperture. (*a*) Fraunhofer and (*b*) Fresnel diffraction.

Fraunhofer diffraction, the source lies at the principal focus of a lens L_1, which renders the light parallel as it falls on the aperture. A second lens L_2 focuses parallel diffracted beams on the observing screen F, situated in the principal focal plane of L_2. In Fresnel diffraction, no lenses intervene. The diffraction effects occur chiefly near the borders of the geometrical shadow, indicated by the broken lines. An alternative way of distinguishing the two classes, therefore, is to say that Fraunhofer diffraction concerns the effects near the focal point of a lens or mirror, while Fresnel diffraction concerns those effects near the edges of shadow. Photographs of some diffraction patterns of each class are shown in **Fig. 2**. All of these may be demonstrated especially well by using the light beam from a neon-helium laser. *See* LASER.

Fraunhofer Diffraction

This class of diffraction is characterized by a linear variation of the phases of the Huygens secondary waves with distance across the wavefront, as they arrive at a given point on the observing screen. At the instant that the incident plane wave occupies the plane of the diffracting screen, it may be regarded as sending out, from each element of its surface, a multitude of secondary waves, the joint effect of which is to be evaluated in the focal plane of the lens L_2. The analysis of these secondary waves involves taking account of both their amplitudes and their phases. The simplest way to do this is to use a graphical method, the method of the so-called vibration curve, which can readily be extended to cases of Fresnel diffraction. *See* HUYGENS' PRINCIPLE.

Vibration curve. The basis of the graphical method is the representation of the amplitude and phase of a wave arriving at any point by a vector, the length of which gives the magnitude of the amplitude, and the slope of which gives the value of the phase. In **Fig. 3** are shown two vectors of amplitudes a_1 and a_2, pertaining to two waves having a phase difference δ of 60°. That is, the waves differ in phase by one-sixth of a complete vibration. The resultant amplitude A and phase θ (relative to the phase of the first wave) are then found from the vector sum of a_1 and a_2, as indicated. A mathematical proof shows that this proposition is rigorously correct and that it may be extended to cover the addition of any number of waves.

The vibration curve results from the addition of a large (really infinite) number of infinitesimal vectors, each representing the contribution of the Huygens secondary waves from an element of surface of the wavefront. If these elements are assumed to be of equal area, the magnitudes of the amplitudes to be added will all be equal. They will, however, generally differ in phase, so that if the elements were small but finite each would be drawn at a small angle with the

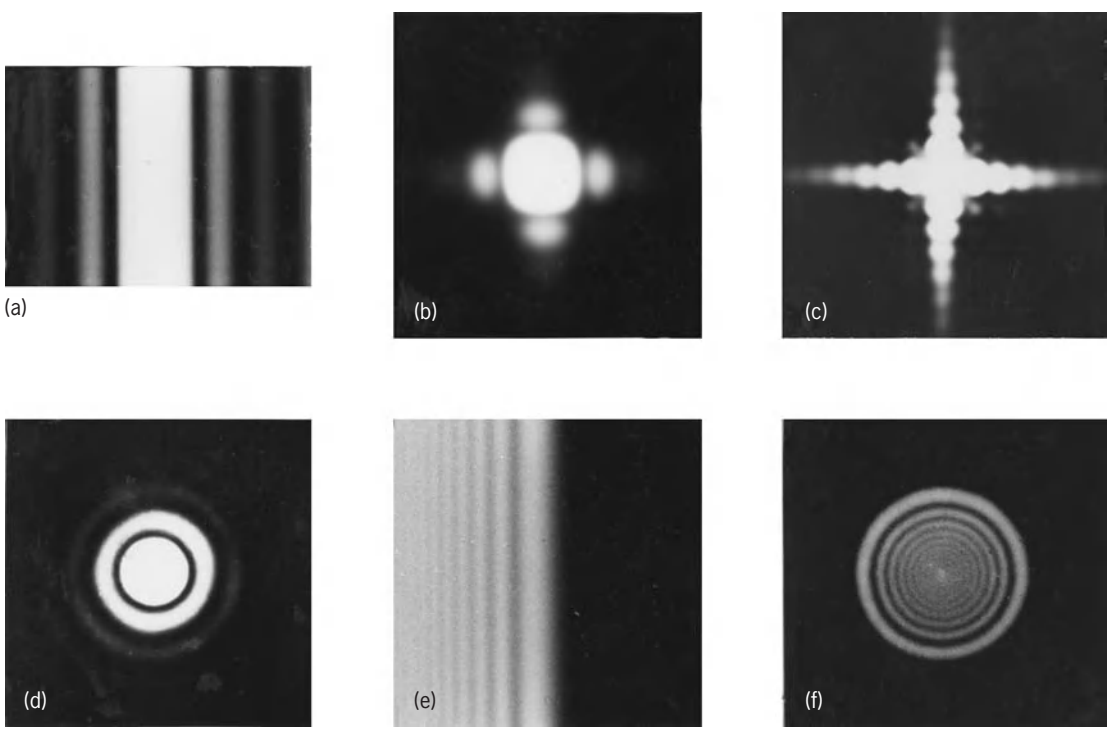

Fig. 2. Diffraction patterns, photographed with visible light. (*a*) Fraunhofer pattern, for a slit; (*b*) Fraunhofer pattern, square aperture, short exposure; (*c*) Fraunhofer pattern, square aperture, long exposure (*courtesy of F. S. Harris*); (*d*) Fraunhofer pattern, circular aperture (*courtesy of R. W. Ditchburn*); (*e*) Fresnel pattern, straight edge; (*f*) Fresnel pattern, circular aperture.

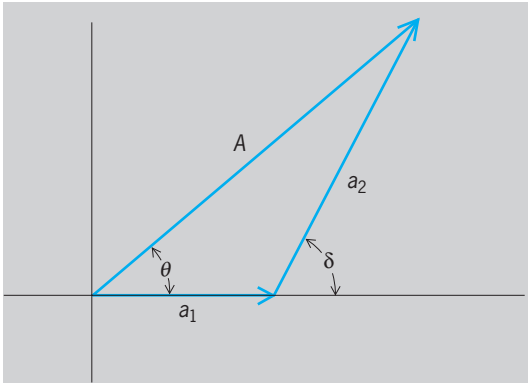

Fig. 3. Graphical addition of two amplitudes.

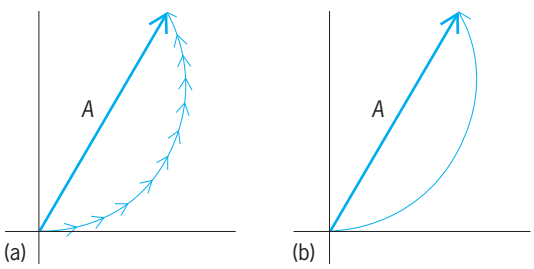

Fig. 4. Vibration curves. (*a*) Addition of many equal amplitudes differing in phase by equal amounts. (*b*) Equivalent curve, when amplitudes and phase differences become infinitesimal.

preceding one, as shown in **Fig. 4*a***. The resultant of all elements would be the vector A. When the individual vectors represent the contributions from infinitesimal surface elements (as they must for the Huygens wavelets), the diagram becomes a smooth curve, the vibration curve, shown in Fig. 4*b*. The intensity on the screen is then proportional to the square of this resultant amplitude. In this way, the distribution of the intensity of light in any Fraunhofer diffraction pattern may be determined.

The vibration curve for Fraunhofer diffraction by screens having slits with parallel, straight edges is a circle. Consider, for example, the case of a slit of width b illustrated in **Fig. 5**. The edges of the slit extend perpendicular to the plane of the figure, and the slit is illuminated by plane waves of light coming from the left. If s is the distance from O of a surface element ds (ds actually being a strip extending perpendicular to the figure), the extra distance wavelet from ds must travel in reaching a point on the screen lying at the angle θ from the center is $s \sin \theta$. Since this extra distance determines the phase difference, the latter varies linearly with s. This condition necessitates that the vibration curve be a circle.

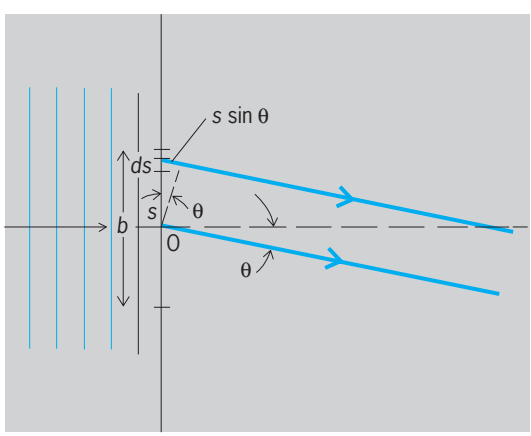

Fig. 5. Analysis of Fraunhofer diffraction by a slit.

The intensity distribution for Fraunhofer diffraction by a slit as a function of the angle θ may be simply calculated as follows. The extra distances traveled by the wavelets from the upper and lower edges of the slit, as compared with those from the center, are $+(b/2) \sin \theta$ and $-(b/2) \sin \theta$. The corresponding phase differences are $2\pi/\lambda$ times these quantities, λ being the wavelength of the light. Using the symbol β for $(\pi b \sin \theta)/\lambda$, it is seen that the end points of the effective part of the vibration curve must differ in slope by $\pm\beta$ from the slope at its center, where it is taken as zero. **Figure 6** shows the form of the

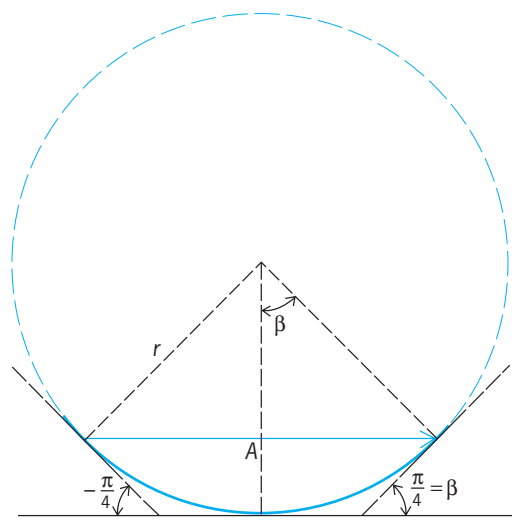

Fig. 6. Vibration curve and resultant amplitude for a particular point in Fraunhofer diffraction by a slit.

vibration curve for $\beta = \pi/4$, that is for $\sin \theta = {}^1/_4 \lambda b$. The resultant is $A = 2r \sin \beta$, where r is the radius of the arc. The amplitude A_0 that would be obtained if all the secondary waves were in phase at the center of the diffraction pattern, where $\theta = 0$, is the length of the arc. Thus, Eq. (1) holds. The intensity at any

$$\frac{A}{A_0} = \frac{\text{chord}}{\text{arc}} = \frac{2r \sin \beta}{2r\beta} = \frac{\sin \beta}{\beta} \qquad (1)$$

angle is given by Eq. (2), where I_0 is the intensity at the center of the pattern. **Figure 7** shows a graph

$$I = I_0 \frac{\sin^2 \beta}{\beta^2} \qquad (2)$$

of this function. The central maximum is twice as wide as the subsidiary ones, and is about 21 times as intense as the strongest of these. A photograph of this pattern is shown in Fig. 2*a*.

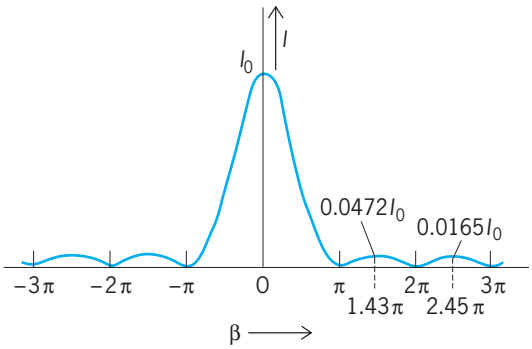

Fig. 7. Intensity distribution curve for Fraunhofer diffraction by a slit.

The dimensions of the pattern are important, since they determine the angular spread of the light behind the slit. The first zeros occur at values $\beta = (\pi b \sin \theta)/\lambda = \pm\pi$. In most cases the angle θ is extremely small, so that Eq. (3) holds. For a slit 1 mm

$$\sin\theta \approx \theta = \pm\frac{\lambda}{b} \tag{3}$$

wide, for example, and green light of wavelength 5×10^{-5} cm, Eq. (3) gives the angle as only 0.0005 radian, or 1.72 minutes of arc. The slit would have to be much narrower than this, or the wavelength much longer, for the approximation to cease to be valid.

The main features of Fraunhofer diffraction patterns of other shapes can be understood with the aid of the vibration curve. Thus for a rectangular or square aperture, the wavefront may be subdivided into elements parallel to either of two adjacent sides, giving an intensity distribution which follows the curve of Fig. 7 in the directions parallel to the two sides. Photographs of such patterns appear in Fig. 2*b* and *c*. In Fig. 2*c* it will be seen that there are also faint subsidiary maxima lying off the two principal directions. These have intensities proportional to the products of the intensities of the side maxima in the slit pattern. The fact that these subsidiary intensities are extremely low compared with that of the central maximum has an important application to the apodization of lenses, to be discussed later.

Diffraction grating. An idealized diffraction grating consists of a large number of similar slits, equally spaced. Equal segments of the vibration curve are therefore effective, as shown in **Fig. 8***a*. The resultants *a* of each segment are then to be added to give *A*, the amplitude due to the whole grating, as shown in Fig. 8*b*. The phase difference between successive elements is here assumed to be very small. As it is increased, by going to a larger angle θ, the resultant *A* first goes to zero at an angle corresponding to λ/W, where *W* is the total width of the grating. After going through numerous low-intensity maxima, *A* again rises to a high value when the phase difference between the successive vectors for the individual slits approaches a whole vibration. These small vectors *a* are then all lined up again, as they were at the center of the pattern ($\theta = 0$). The resulting strong maximum represents the "first-order spectrum," since its position depends on the wavelength. A similar condition occurs when the phase difference becomes two, three, or more whole vibrations, giving the higher-order spectra. By means of this diagram, it is possible not only to predict the intensities of successive orders for an ideal grating, but also to find the sharpness of the maxima which represent the spectrum lines. *See* DIFFRACTION GRATING.

Determination of resolving power. Fraunhofer diffraction by a circular aperture determines the resolving power of instruments such as telescopes, cameras, and microscopes, in which the width of the light beam is usually limited by the rim of one of the lenses. The method of the vibration curve may be extended to find the angular width of the central diffraction maximum for this case. **Figure 9** compares the treatments of square and circular apertures by showing, above, the elements of equal phase difference into which the wavefront may be divided, and, below, the corresponding vibration curves. For the square aperture shown in Fig. 9*a*, the areas of the surface elements are equal, and the curve forms a complete circle at the first zero of intensity. In Fig. 9*b* these areas, and hence the lengths of the successive vectors, are not equal, but increase as the center of the curve is approached,

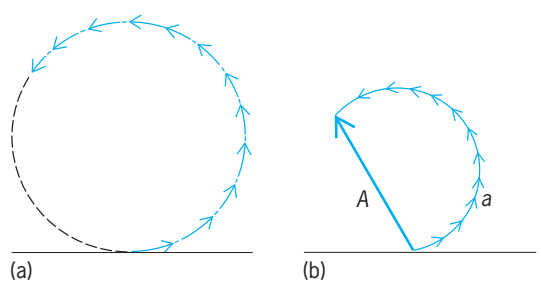

Fig. 8. Graphical analysis of diffraction grating. (*a*) Vibration curve for grating of 12 slits. (*b*) Resultant amplitude *A* formed by adding amplitudes *a* from individual slits. Each *a* represents the chord of one of the short arcs in part *a*.

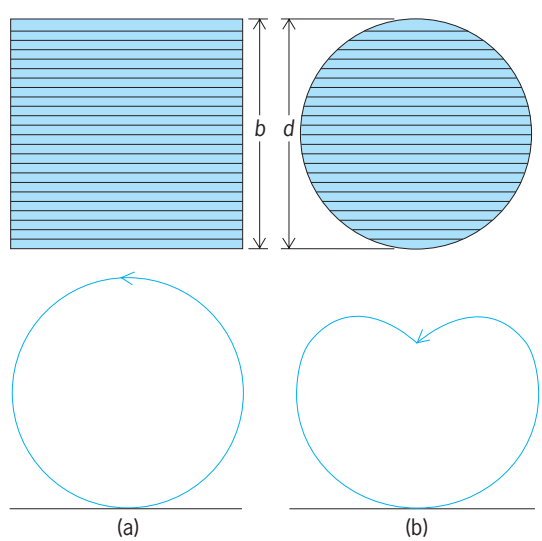

Fig. 9. Condition of vibration curves at first minimum for (*a*) a square aperture and (*b*) a circular aperture.

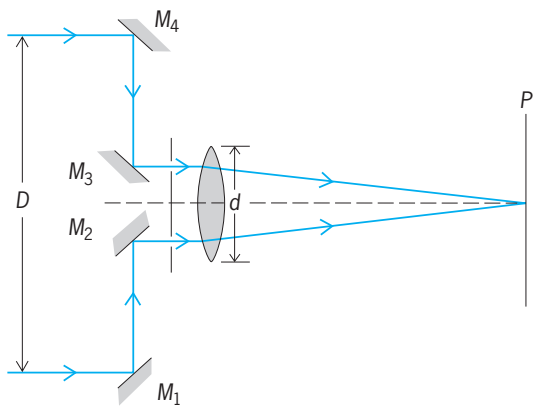

Fig. 10. Michelson stellar interferometer, which is used for measuring the diameters of stars.

and then decrease again. The result is that the curve must show somewhat greater phase differences at its extremes in order to form a closed figure. An exact construction of the curve or, better, a mathematical calculation shows that the extreme phase differences required are $\pm 1.220\pi$, yielding Eq. (4)

$$\sin\theta \approx \theta = \pm\frac{1.220\lambda}{d} \tag{4}$$

for the angle θ at the first zero of intensity. Here d is the diameter of the circular aperture. When this result is compared with that of Eq. (3), it is apparent that, relative to the pattern for a rectangle of side $b = d$, this pattern is spread out by 22%. Obviously it now has circular symmetry and consists of a diffuse central disk, called the Airy disk, surrounded by faint rings (Fig. 2*d*). The angular radius of the disk, given by Eq. (4), may be extremely small for an actual optical instrument, but it sets the ultimate limit to the sharpness of the image, that is, to the resolving power. *See* RESOLVING POWER (OPTICS).

Other applications. Among the applications of Fraunhofer diffraction, other than the calculation of resolving power, are its use in (1) the theory of certain types of interferometers in which the interfering light beams are brought together by diffraction, (2) the theory of microscopic imaging, and (3) the design of apodizing screens for lenses.

Michelson's stellar interferometer. This instrument was devised by A. A. Michelson to overcome the limitation expressed by Eq. (4). In front of a telescope are placed two fixed mirrors, shown at M_2 and M_3 in **Fig. 10**, and two others, M_1 and M_4, movable so as to vary their separation D. The light of a star is directed by these mirrors into a telescope as shown, and the diffraction patterns, of size corresponding to the aperture of the mirrors, are superimposed in the focal plane P of the lens. When D is small the resulting pattern is crossed by interference fringes. If the light source has a finite width, increasing D causes these fringes to become indistinct and eventually to disappear. For a circular disk source, such as a star, the angle subtended by the disk must be $1.22\lambda/D$ for the disappearance of the fringes. The resolving power of the telescope has thus effectively been increased in the ratio D/d. In this way angular diameters as small as 0.02 second of arc have been measured, corresponding to $D = 7$ m or 24 ft. *See* INTERFERENCE OF WAVES; INTERFEROMETRY.

Microscopic imaging. E. Abbe's theory of the microscope evaluates the resolving power of this instrument considering not only the diffraction caused by the limited aperture of the objective, but also that caused by the object itself. If this object is illuminated by coherent light, such as a parallel beam from a point source, its Fraunhofer diffraction pattern is formed in the rear focal plane of the microscope objective. Abbe took as an object a diffraction grating, and in this case the pattern consists of a series of sharp maxima representing the various orders $m = 0, \pm 1, \pm 2, \ldots$. If the light of all these orders is collected by the objective, a perfect image of the grating can be formed where the light is reunited in the image plane. In practice, however, the objective can include only a limited number of them, as is indicated in **Fig. 11**. Here only the orders $+2$ to -2 are shown entering the objective. The higher orders, involving greater values of θ, would miss the lens.

The final image must be produced by interference in the image plane of the Huygens secondary waves coming from the various orders. To obtain a periodic variation of intensity in that plane, at least two orders must be accepted by the objective. The angle θ_1 at which the first order occurs is given by $\sin\theta_1 = \lambda/d$,

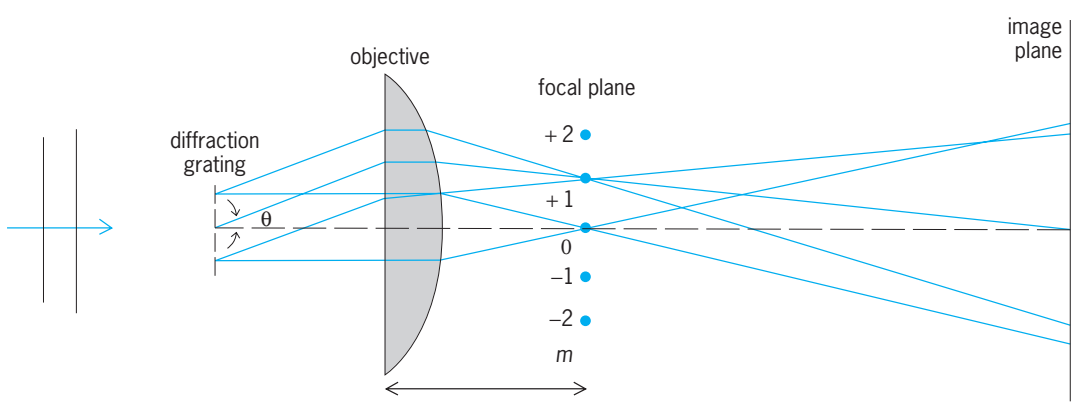

Fig. 11. Abbe's method of treating the resolving power of a microscope.

where d is the spacing of the lines in the grating. The limit of resolution of the microscope, that is, the smallest value of d that will produce an indication of separated lines in the image, may thus be found from the angular aperture of the objective. In order for it to accept only the orders 0, ± 1, the lens aperture 2α must equal at least $2\theta_1$, giving $d = \lambda/\sin\alpha$. This resolving limit may be decreased by illuminating the grating from one side, so that the zero-order light falls at one edge of the lens, and that of one of the first orders at the other edge. Then $\theta_1 = 2\alpha$, and the limit of resolution is approximately $0.5\lambda/\sin\alpha$. *See* OPTICAL MICROSCOPE.

Apodization. This is the name given to a procedure by which the effect of subsidiary maxima (such as those shown in Fig. 7) may be partially suppressed. Such a suppression is desirable when one wishes to observe the image of a very faint object adjacent to a strong one. If the fainter object has, for example, only $^1/_{1000}$ of the intensity of the stronger one, the two images will have to be far enough apart so that the principal maximum for the fainter one is at least comparable in intensity to the secondary maxima for the stronger object at that point. In the pattern of a rectangular aperture it is not until the tenth secondary maximum that the intensity of the secondary maxima falls below $^1/_{1000}$ of the intensity of the principal maximum. Here it has been assumed, however, that the fainter image lies along one of the two principal directions of diffraction of the rectangular aperture, perpendicular to two of its adjacent sides. At 45° to these directions the subsidiary maxima are much fainter (Fig. 2*c*), and even the second one has an intensity of only $^1/_{3700}$, the square of the value of the second subsidiary maximum indicated in Fig. 7.

The simplest apodizing screen is a square aperture placed over a lens, the diagonal of the square being equal to the diameter of the lens. If the lens is the objective of an astronomical telescope, for example, the presence of a fainter companion in a double-star system can often be detected by turning the square aperture until the image of the companion star lies along its diagonal. Apodizing screens may be of various shapes, depending on the purpose to be achieved. It has been found that a screen of graded density, which shades the lens from complete opacity at the rim to complete transparency at a small distance inward, is effective in suppressing the circular diffraction rings surrounding the Airy disk. In all types of apodization there is some sacrifice of true resolving power, so that it would not be used if the two images to be resolved were of equal intensity.

Fresnel Diffraction

The diffraction effects obtained when the source of light or the observing screen are at a finite distance from the diffracting aperture or obstacle come under the classification of Fresnel diffraction. This type of diffraction requires for its observation only a point source, a diffracting screen of some sort, and an observing screen. The latter is often advantageously replaced by a magnifier or a low-power microscope. The observed diffraction patterns generally differ according to the radius of curvature of the wave and the distance of the point of observation behind the screen. If the diffracting screen has circular symmetry, such as that of an opaque disk or a round hole, a point source of light must be used. If it has straight, parallel edges, it is desirable from the standpoint of brightness to use an illuminated slit parallel to these edges. In the latter case, it is possible to regard the wave emanating from the slit as a cylindrical one. For the purpose of deriving the vibration curve, the appropriate way of dividing the wavefront into infinitesimal elements is to use annular rings in the first case, and strips parallel to the axis of the cylinder in the second case.

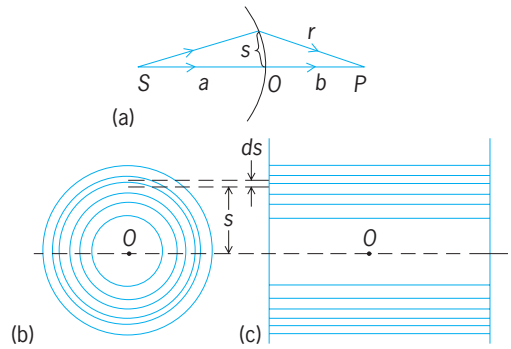

Fig. 12. Division of wavefront for constructing vibration curve for Fresnel diffraction. (*a*) Section of wave diverging from *S* and paths to screen of two secondary wavelets. (*b*) Annular zones. (*c*) Strips.

Figure 12 illustrates the way in which the radii of the rings, or the distances to the edges of the strips, must be chosen in order that the phase difference may increase by an equal amount from one element to the next. Figure 12*a* shows a section of the wavefront diverging from the source S, and the paths to the screen of two secondary wavelets. The shortest possible path is b, while r is that for another wavelet originating at a distance s above the "pole" O. Since all points on the wavefront are at the same distance a from S, the path difference between the two routes from S to P is $r - b$. When this is evaluated, to terms of the first order in s/a and s/b, the phase difference is given by Eq. (5). The phase difference across an elementary zone of radius s and width ds then becomes

$$\delta = \frac{2\pi}{\lambda}(r - b) \approx \frac{\pi(a+b)}{ab\lambda}s^2 = Cs^2 \qquad (5)$$

$d\delta = 2Csds$, so that for equal increments of δ the increment of s must be proportional to $1/s$. The annular zones and the strips drawn in this way on the spherical or cylindrical wave, respectively, looking toward the pole from the direction of P, are shown in Fig. 12*b* and *c*.

For the annular zonal elements, the areas $2\pi sds$ are all equal, and hence the amplitude elements of the vibration curve should have the same magnitude. Actually, they must be regarded as falling off slowly, due to the influence of the "obliquity factor" of Huygens' principle. The resulting vibration curve is nearly, but not quite, circular, and is illustrated in **Fig. 13***a*. It

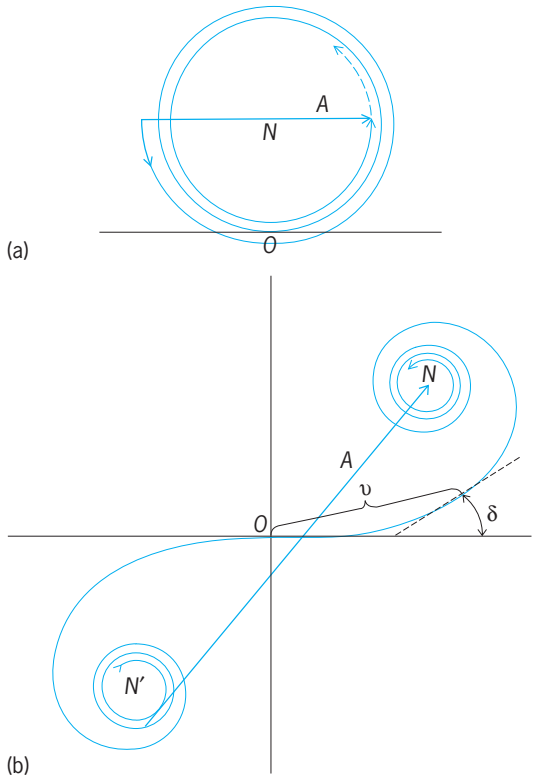

Fig. 13. Vibration curves for Fresnel diffraction patterns. (*a*) Circular division of wavefront. (*b*) Strip division (Cornu's spiral).

spirals in toward the center C, at a rate that has been considerably exaggerated in the figure. The intensity at any point P on the axis of a circular screen centered on O can now be determined as the square of the resultant amplitude A for the appropriate part or parts of this curve. The curve shown in Fig. 13*a* is for a circular aperture which exposes five Fresnel zones, each one represented by a half-turn of the spiral. The resultant amplitude is almost twice as great (and the intensity four times as great) as it would be if the whole wave were exposed, in which case the vector would terminate at C. The diffraction by other circular screens may be determined in this same manner.

Zone plate. This is a special screen designed to block off the light from every other half-period zone, and represents an interesting application of Fresnel diffraction. The Fresnel half-period zones are drawn, with radii proportional to the square roots of whole numbers, and alternate ones are blackened. The drawing is then photographed on a reduced scale. When light from a point source is sent through the negative, an intense point image is produced, much like that formed by a lens. The zone plate has the effect of removing alternate half-turns of the spiral, the resultants of the others all adding in the same phase. By putting $\delta = \pi$ and $a = \infty$ in Eq. (5), it is found that the "focal length" b of a zone plate is $s_1{}^2/\lambda$, where s_1 is the radius of the first zone.

Cornu's spiral. This is the vibration curve for a cylindrical wavefront, and is illustrated in Fig. 13*b*. The areas of the elementary zones, and hence the magnitudes of the component vectors of the vibration curve, decrease rapidly, being proportional to ds, and hence to $1/s$ (Fig. 12*c*). The definition of Cornu's spiral requires that its slope δ at any point be proportional to the square of the corresponding distance s measured up the wavefront [Fig. 12*a* and Eq. (5)]. The length of the spiral from the origin is proportional to s, but it is usually drawn in terms of the dimensionless variable υ, defined by $\upsilon = s\sqrt{2C/\pi}$ in the notation of Eq. (5). The coordinates of any point on the curve may then be found from tables of Fresnel's integrals.

As an example of the application of Cornu's spiral, consider the diffraction of an opaque straight edge, such as a razor edge, illuminated by light from a narrow slit parallel to the edge. At some point outside the edge of the shadow, say one which exposes three half-period strips beyond the pole, the resultant amplitude will be that labeled A in Fig. 13*b*. The intensity will be greater than that given by the amplitude NN', which represents the amplitude for the whole (unobstructed) wave. On going further away from the edge of the shadow, the tail of the vector A will move along the spiral inward toward N', and the intensity will pass through maxima and minima. At the edge of the geometrical shadow the amplitude is ON, and the intensity is just one-fourth of that due to the unobstructed wave. Further into the shadow the intensity approaches zero regularly, without fluctuations, as the tail of the vector moves up toward N. A photograph of the straight-edge pattern is shown in Fig. 2*e*.

Babinet's principle. This states that the diffraction patterns produced by complementary screens are identical. Two screens are said to be complementary when the opaque parts of one correspond to the transparent parts of the other, and vice versa. Babinet's principle is not very useful in dealing with Fresnel diffraction, except that it may furnish a shortcut method in obtaining the pattern for a particular screen from that of its complement. The principle has an important application for Fraunhofer diffraction, in parts of the field where there is zero intensity without any screen. Under this condition the amplitudes produced by the complementary screens must be equal and opposite, since the sum of the effects of their exposed parts gives no light. The intensities, being proportional to the squares of the amplitudes, must therefore be equal. In Fraunhofer diffraction the pattern due to a disk is the same as that due to a circular hole of the same size.

Diffraction of Microwaves

The diffraction of microwaves, which have wavelengths in the range of millimeters to centimeters, has been intensively studied since World War II because of its importance in radar work. Many of the characteristics of optical diffraction can be strikingly demonstrated by the use of microwaves. Microwave diffraction shows certain features, however, that are not in agreement with the Huygens-Fresnel theory, because the approximations made in that theory are no longer valid. Most of these

approximations, for example, that made in deriving Eq. (5), depend for their validity on the assumption that the wavelength is small compared to the dimensions of the apparatus. Furthermore, it is not legitimate to postulate that the wave has a constant amplitude across an opening, and zero intensity behind the opaque parts, except for the very minute waves of light.

As an example of the failure of classical diffraction theory when applied to microwaves, the results of the diffraction by a circular hole in a metal screen may be mentioned. The observed patterns begin to show deviations from the Fresnel theory, and even from the more rigorous Kirchhoff theory, when the point of observation is within a few wavelengths' distance from the plane of the aperture. Even in this plane itself there are detection effects that could not have been predicted from the earlier theories, which treat light as a scalar, rather than a vector, wave motion. An exact vector theory of diffraction, developed by A. Sommerfeld, has been applied only in a few simple cases, but the measurements at microwave frequencies agree with it wherever it has been tested. *See* LIGHT; MICROWAVE OPTICS.

F. A. Jenkins; W. W. Watson

Bibliography. M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; J. M. Cowley, *Diffraction Physics*, 3d revised ed., 1995; J. W. Goodman, *Introduction to Fourier Optics*, 3d ed., 2005; B. D. Guenther, *Modern Optics*, 1990; F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976, paper 2001; J. Meyer-Arendt, *Introduction to Classical and Modern Optics*, 4th ed., 1994.

Diffraction grating

An optical device that exhibits a spatial periodic variation in one or more of the following properties: transmittance, surface profile, or refractive index. Typically, though not necessarily, the variation is one dimensional (only in one direction in the plane of the grating), and the period varies from about one-half to about one hundred times the wavelength of interest. When the variation is two dimensional, the component is normally called a diffractive optical element (DOE).

A periodic arrangement of parallel narrow slits is the simplest example of a grating. Some basic characteristics of all grating types can be understood with reference to this simple form. More commonly, gratings are periodic corrugated (grooved) surface structures, and can be used in transmission or reflection, in the latter case by coating the surface with a suitable reflective coating. The same basic principles apply in all cases.

Diffraction by periodic structures. If a parallel beam of monochromatic light of wavelength λ is incident on a periodic slit assembly of period d and slit width s (**Fig. 1***a*), then diffraction theory shows that the beam will be separated into several beams traveling in different directions. Far from the grating, the diffracted irradiance is of the form given by Eq. (1),

$$I = I_0 \left(\frac{\sin \beta}{\beta} \right)^2 \left[\frac{\sin(N\gamma)}{N \sin \gamma} \right]^2 \qquad (1)$$

where I_0 is the incident irradiance, N is the number of slits, $\beta = \pi Ls/\lambda$, $\gamma = \pi Ld/\lambda$, and L is the cosine of the angle that the beam direction forms with the x axis. This function is plotted in Fig. 1*b*. The sharp main peaks, occurring along the directions indicated in Fig. 1(*a*), are called diffraction orders. *See* DIFFRACTION.

For the case of parallel slits, the distribution of power between orders depends on the ratio s/d; for the general case, it depends on the groove characteristics. In all cases, the principal peaks become sharper and the secondary oscillations weaker as the number of grooves N increases. The sharpness of the peaks can be understood if the grating is thought of as a multiple-beam interference device, where each slit or groove contributes one interfering beam. The interference picture is also helpful in understanding the position of the peaks: they will be formed in the direction where constructive interference occurs. This direction is given by the grating equation (2), where a is the angle of incidence, b is

$$\sin a \pm \sin b = m\lambda/d \qquad (2)$$

the angle of diffraction, d is the grating period, and m is an integer representing the order of diffraction. If a Cartesian sign convention for angles is adopted (positive if counterclockwise from the axis of the incident beam), the negative sign applies to transmission gratings and the positive sign to reflection gratings. *See* INTERFERENCE OF WAVES.

Blazing. For a given angle of incidence, the angle of diffraction increases with $1/d$. Since the absolute value of $m\lambda/d$ is less than 2 [as follows from Eq. (2) and the fact that the absolute values of $\sin a$ and $\sin b$ are both less than 1], as d approaches the wavelength the grating equation is satisfied for fewer orders and eventually only one order. In such a case, the diffracted energy can be concentrated mostly in a single order, which is generally desirable. With coarser gratings, many orders are allowed. However, even in this case an appropriate groove profile can also concentrate the light in a desired order. A grating with such a profile is called blazed. The principle of blazing is explained in **Fig. 2**. Light incident on a facet tends to be refracted or reflected according to Snell's law or the law of reflection, respectively. If the direction of refraction or reflection coincides with an allowed direction of diffraction, the energy is concentrated in that direction. The facet angle is called the blaze angle θ; for reflection gratings, it is related to the angles of incidence and diffraction through $\theta = (a + b)/2$. Since the diffraction angle changes with wavelength, only one wavelength (the blaze wavelength) will diffract at the desired angle once the angle of incidence is fixed. *See* REFLECTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES.

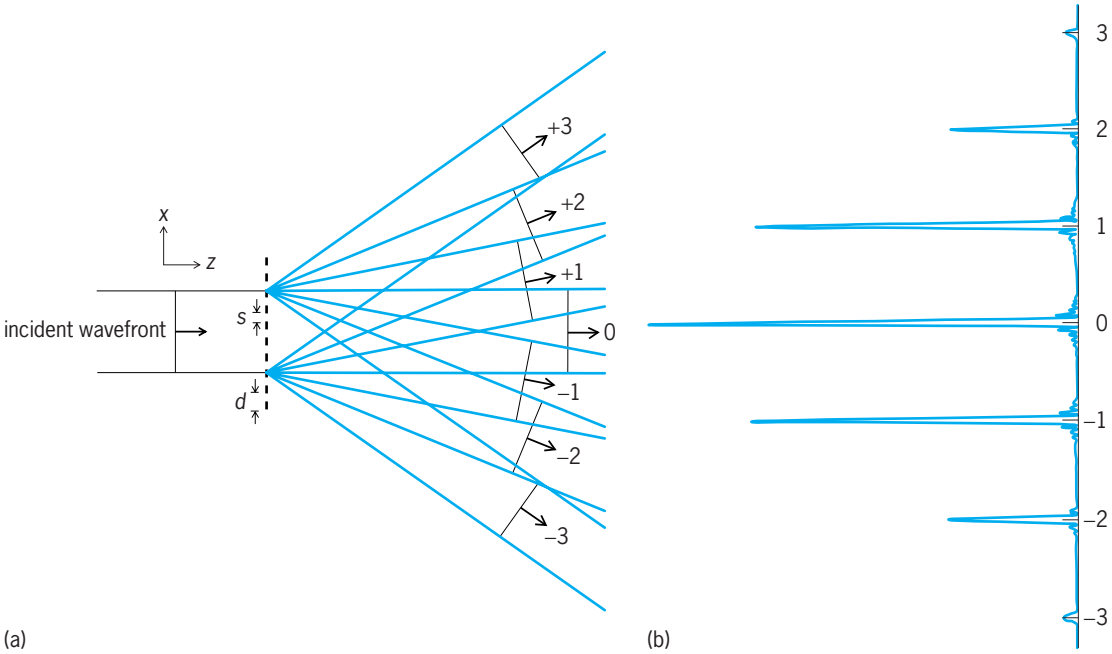

Fig. 1. Behavior of a light beam incident on a periodic slit pattern (grating). (*a*) Splitting of the beam into several beams traveling along different directions. (*b*) Far from the grating, these beams correspond to sharp peaks in the light distribution, called orders of diffraction. The distribution shown here is obtained for $N = 20$ slits and a ratio $d/s = 3.5$.

Dispersion and related properties. Though diffraction gratings can be used with monochromatic radiation, by far their most common use is to place them in the path of a light beam with the aim of dispersing the light, that is, separating it into its component wavelengths. The dispersive properties of a grating are characterized by its angular dispersion D, which can be expressed by Eq. (3), where the negative signs

$$D = \frac{db}{d\lambda} = \frac{\sin b \pm \sin a}{\lambda \cos b} = \pm \frac{m}{d \cos b} \qquad (3)$$

apply to transmission gratings. However, only the absolute value of D is significant. This equation shows that the same dispersion can be achieved with either a coarse grating used in high order or a fine grating used in low order.

The free spectral range (FSR) of a grating is the portion of the spectrum over which there is no overlap of wavelengths from adjacent orders, and is given as λ/m, where λ is here the lowest wavelength in the range of interest. Larger free spectral range is therefore achieved with a lower order m. This reduces the need for order-sorting filters or cross-dispersing elements to separate the orders.

The resolving power R of a grating is defined as the mean wavelength λ divided by the minimum resolvable wavelength difference $\Delta\lambda$; $R \equiv \lambda/\Delta\lambda$. It is given by $R = mN$, where N is the total number of illuminated grooves. This is independent of groove period, but insofar as there is an overall beam size limitation, a finer-pitch grating will provide more grooves and hence higher resolution.

Diffraction efficiency. Diffraction efficiency is defined as the percentage of the incident energy that is concentrated in the desired order. It can reach values above 90%, or even above 99% in special cases, but it is typically strongly wavelength-dependent. The interaction of light with coarse, shallow gratings can be described through scalar diffraction theory. Such gratings therefore show very little difference in efficiency between p (parallel to the grooves) and s (perpendicular to the grooves) polarizations. In this

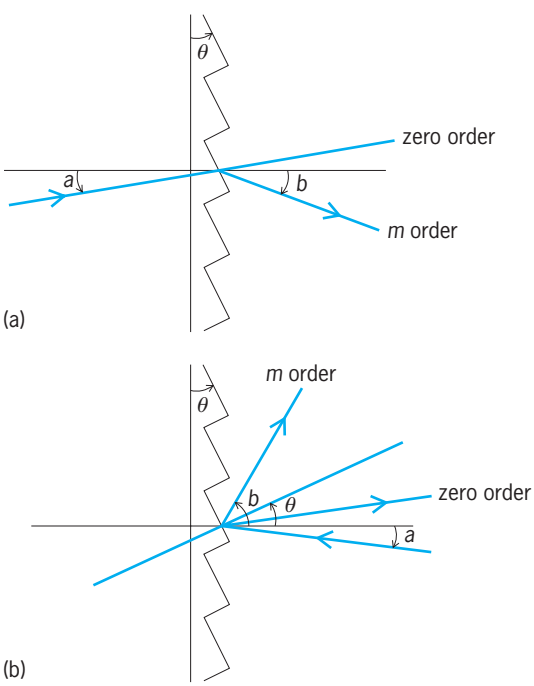

Fig. 2. Geometry of diffraction from a (*a*) transmission grating, (*b*) reflection grating.

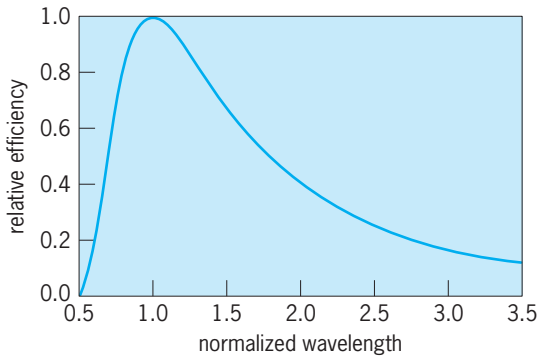

Fig. 3. Relative spectral efficiency of a coarse blazed grating as a function of normalized wavelength λ/λ_0.

case, the relative spectral efficiency takes the functional form given by Eq. (4), where λ_0 the blaze wave-

$$I/I_0 = \left\{ \frac{\sin\left[\pi\left(\frac{\lambda_0}{\lambda} - m\right)\right]}{\pi\left(\frac{\lambda_0}{\lambda} - m\right)} \right\}^2 \qquad (4)$$

length. This function is plotted in **Fig. 3** for $m = 1$. For finer gratings, the *p*-polarization relative efficiency generally retains a similar shape, while the *s*-polarization shows sharp peaks and troughs called Wood's anomalies. The efficiency curve shape can be significantly different for nonblazed gratings and can also be tailored to some extent through modification of the groove shape. *See* POLARIZED LIGHT.

Parasitic light. All light that is not absorbed or diffracted in the direction of a diffraction order is parasitic. There are two main types of parasitic light: ghosts and satellites, and scatter. Ghosts and satellites are concentrations of energy in directions other than the main diffraction orders; the distinction between the two is that satellites are formed very close to a main order. Both types arise from periodic errors in grating spacing or any other periodicity introduced by the fabrication method. Diffuse or homogeneous scatter is a continuous background light, spread more or less uniformly across the angle. Its main source is the microroughness of the groove surface. In-plane scatter is light scattered in the plane of dispersion, caused mostly by random variations in groove period and depth. In several spectroscopic applications, the absence or reduction of parasitic light is of paramount importance and dictates the choice of grating.

Types of gratings. Gratings can be classified according to the mode of use as transmission or reflection. They can also be classified as surface or volume gratings, depending on where diffraction takes place. In volume gratings, the effective grooves are formed by variation of the refractive index inside the material. Volume gratings are further classified as thin or thick, with a thick (also called Bragg) grating being one that has a thickness greater than about 10 grating periods. Although diffraction is the operating principle of all these types, the term "diffraction gratings" most often refers to surface gratings used for dispersing the light. The efficiency of volume Bragg gratings can approach 100% and the width of the spectral efficiency curve can be quite narrow. These characteristics make them desirable as filters or highly efficient, wavelength-selective mirrors. Gratings are also classified according to the fabrication technique. *See* INTERFERENCE FILTER.

Fabrication techniques. Starting in the twentieth century, gratings were first fabricated by ruling, which involves repetitively running a diamond cutting tool along the surface of a metal or glass substrate. Extremely high precision is required in the motion of the tool to avoid artifacts, such as caused by sideways motion or periodic errors in the translation mechanism. An original ruled grating, called a master, is often replicated by the use of a resin that conforms to the master grating surface before hardening. Replicas can be superior to the master at least in terms of possessing smoother grooves, leading to a reduction in scattered light. The groove shape of ruled gratings is triangular (blazed). When gratings must be formed on a curved substrate, it is very difficult to maintain a constant blaze angle, as the cutting tool angle varies relative to the substrate. This results in significant variation of diffraction efficiency across the face of the grating. Partial mitigation of this effect is achieved through multipanel gratings, with different tool angles between panels. It is difficult, however, to maintain phase coherence between the panels, and the result is typically a grating that has the resolving power corresponding to the number of rulings in only one panel.

The development of lasers made possible the fabrication of holographic (or, more properly, interference) gratings, which are made by recording the interference fringe pattern from two coherent light beams onto a suitable recording medium such as a photoresist, followed by development as necessary to form a surface pattern. The groove shape of holographic gratings is typically sinusoidal; however, development techniques can change the groove depth to advantage, and ion-milling techniques can change the groove shape. At high groove densities, holographic gratings are advantageous in terms of their potential for producing minimal scatter, as well as the flexibility to record on an arbitrary (nonplanar) surface shape. Variation of efficiency across the grating face is not observed; however, the grooves are not blazed. In addition, holographic gratings can be formed with nonstraight grooves if the interfering beams are not planar. In that case they can be used to focus a beam as well as to correct the aberrations of the optical system. So-called aberration-correcting gratings which present a slow variation in a second direction blur the distinction between diffraction gratings and diffractive optical elements, but are best classified as gratings. *See* ABERRATION (OPTICS); HOLOGRAPHY; LASER.

The development of directed-energy beam tools in addition to lasers has resulted in new fabrication techniques that provide greater flexibility by controlling the groove depth, shape, and period not only to conform to desired patterns but also to vary along the extent of the grating. Electron-beam lithography is the most advanced of these alternative techniques

at present. Like holographic gratings, these gratings can also be recorded on a variety of nonplanar substrates as well as have nonstraight grooves, while providing greater flexibility in patterning the groove shape. The inherent accuracy of lithographic tools permits the fabrication of fully coherent multipanel gratings on curved substrates. However, these are typically limited in size. Where large-area gratings are needed, holographic techniques are preferred. *See* MICROLITHOGRAPHY.

Use of gratings. A monochromator is a device that produces monochromatic (or narrow-band) radiation of any desired mean wavelength within a certain range. This is accomplished by dispersing the light of a broadband source and placing a slit or narrow aperture to select a small portion of the spectrum. By rotating the grating (changing the angle of incidence), the wavelength is swept across the stationary slit. A typical grating mounting for a monochromator is the Czerny-Turner arrangement, comprising a collimating and focusing mirror (**Fig. 4**).

A spectrometer is a device that records the spectrum of an object by dispersing the light from the object and placing a photodetector array at the output. Thus the spectrometer uses no output slit but must retain the input slit or pinhole to maintain a unique angle of incidence on the grating. When the image of a monochromatic point source is sufficiently free from aberrations, it is possible to obtain simultaneously many independent spectra of points along the input slit by using a two-dimensional array detector at the output. The detector then records the spectral information in the perpendicular direction to the slit and spatial information in the parallel direction to the slit. Such an arrangement is called an imaging spectrograph or spectrometer.

Several alternative spectrometer grating mounts are possible, in addition to the Czerny-Turner one. Concave grating mounts make use of the imaging property of a grating formed on a concave spherical surface of radius R. The entrance slit is placed on a circle of radius $R/2$ tangent to the grating at the apex, in which case the image is found on the same circle (the Rowland circle). However, the aberrations of this arrangement are severe, and modifications of the classical form using aberration-correcting holographic gratings are advantageous. For maximum resolution the Littrow mounting is needed, in which

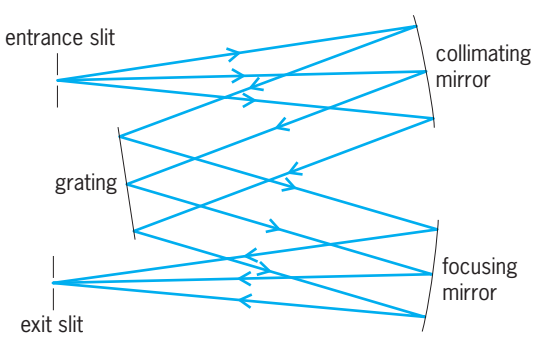

Fig. 4. Typical Czerny-Turner grating monochromator mount. The grating rotates about an axis perpendicular to the paper.

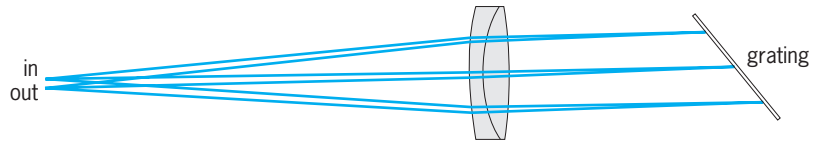

Fig. 5. Example of a grating mounted in Littrow configuration, where the angles of incidence and diffraction are approximately equal. Input and output beams are shown slightly separated in the plane of the paper, though they may also be placed above and below that plane.

the angles of incidence and diffraction are approximately equal (**Fig. 5**). A coarse grating used in the Littrow mode and in a high order of diffraction is called an echelle. *See* ASTRONOMICAL SPECTROSCOPY.

The monochromatic image of a straight slit is inherently curved in flat grating and Rowland circle–type instruments and therefore is not a good match for array detectors. Such curvature, as well as other aberrations, is well controlled in the Offner spectrograph, which uses a convex grating and concentric optics to achieve a high degree of aberration correction with low slit distortion. This makes possible the use of long slits. *See* SPECTROGRAPH; SPECTROSCOPY.

Modern developments and applications. Fabrication techniques continue to improve and new techniques are still appearing, most notably ion etching and lithography. In addition, innovative uses of gratings and new applications are constantly emerging. Imaging spectroscopy is becoming a mainstay technique in Earth and planetary remote sensing. Volume phase gratings are beginning to be considered for astronomical telescope spectrographs. Large-area gratings, assembled from several smaller ones, are needed for laser fusion applications, necessitating techniques for extremely precise assembly. Grating-based wavelength multiplexing in optical communications is an active area of development. Tunable laser cavities can utilize transmission or reflection gratings. *See* NUCLEAR FUSION; OPTICAL COMMUNICATIONS.

P. Mouroulis

Bibliography. M. C. Hutley, *Diffraction Gratings*, Academic Press, 1982; J. M. Lerner and A. Thevenon, *Diffraction Gratings: Ruled and Holographic*, Jobin Yvon Inc. (internal document); E. G. Loewen and E. Popov, *Diffraction Gratings and Applications*, Marcel Dekker, 1997; C. Palmer (ed.), *Diffraction Grating Handbook*, 5th ed., Spectra Physics (Richardson Grating Laboratory), Rochester, 2002.

Diffuser

A device in which the kinetic energy of a flow is recovered, thus giving a static pressure rise. A simple example is a diverging passage (see **illus.**) with flow entering the small end and leaving the large end. For an incompressible flow (density, ρ, is constant) the mass conservation principle, Eq. (1), indicates

$$m = \rho V A \tag{1}$$

that the velocity (V) will decrease through the diffuser as the area (A) increases, since the mass flow rate (m) is constant. It follows from the Bernoulli

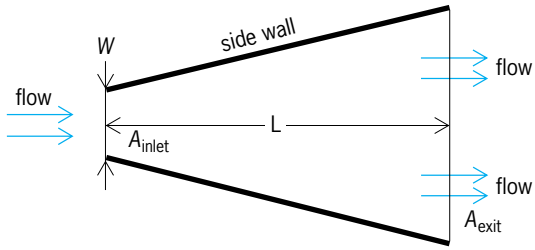

Standard diffuser geometry.

principle, Eq. (2), and from the fact that the total

$$p_0 = p + \tfrac{1}{2}\rho V^2 \tag{2}$$

pressure (p_0) changes but slightly, that the static pressure (p) rises in the diffuser. All subsonic diffusers, whether incompressible or not, experience streamtube area increase as kinetic energy is converted to static pressure rise. Supersonic diffusers operate differently since changes in density become very significant. Many different geometries are possible for diffusers. *See* BERNOULLI'S THEOREM; CONSERVATION OF MASS; SUPERSONIC DIFFUSER.

Applications. Diffusers serve many vital roles. Many fluid-dynamic processes exhibit excessive kinetic energy at some point in their operating cycle. This kinetic energy can be recovered as a useful static pressure rise by the implementation of a diffuser. For example, hydraulic turbines installed for hydroelectric plant operation invariably use a draft tube, that is, an exhaust diffuser. Likewise, ground-based gas turbines for electric power production or general-purpose expansion turbines for waste-gas expansion generally employ an exhaust diffuser for extra pressure recovery. Compressors and pumps invariably generate excessive kinetic energy as work is done by their rotors, and diffusers are required to recover the kinetic energy as a static pressure rise. The basic elements of fluid transfer include pipes, elbows, valves, and diffusers. *See* COMPRESSOR; GAS TURBINE; HYDRAULIC TURBINE; PUMP.

Performance. The performance of a diffuser is described by a static pressure recovery coefficient, C_p. In the case of incompressible flow, this coefficient is equal to the fraction of the inlet kinetic energy that is recovered or converted to a static pressure rise in the diffuser. Values of C_p cannot exceed unity (1.0) for a diffuser but can reach zero or negative numbers for diffusers whose design is so poor that the flow must be pushed through, that is, accelerated.

Diffuser performance is frequently plotted on a performance map, on which contours of recovery (curves connecting points with equal values of C_p) are shown in terms of important geometric parameters, including the area ratio (AR, the ratio of exit area to inlet area, A_{exit}/A_{inlet}) on the ordinate and the length-to-width parameter (the ratio of diffuser length to inlet width, L/W) on the abscissa. Inlet aerodynamic blockage is an important variable which establishes the effective flow-area blockage at the diffuser inlet resulting from low-momentum shear flow near the walls. Many such maps exist from which good estimates of performance can be established.

Ideal limit. It is instructive to consider the ideal limit of a diffuser. In the ideal case, the total pressure is unchanged between the inlet and the outlet. If the Bernoulli equation (2) is used to compute the static pressure rise, then Eq. (3) is obtained for the ideal re-

$$C_{p,\text{ideal}} = 1 - \frac{1}{(AR)^2} \tag{3}$$

covery coefficient after the mass-conservation equation is employed. Hence, it is clear that $C_p = 1.0$ is achievable only for both ideal flow and an infinite area ratio. As examples, for $AR = 4$, $C_{p,\text{ideal}} = 0.9375$, and for $AR = 2.0$, $C_{p,\text{ideal}}$=0.75. Of course, the ideal case can be approached only in the limit.

Effectiveness. Diffuser effectiveness (η) is the ratio of the actual static pressure recovery coefficient to the ideal recovery coefficient. Levels of η in the range 0.5–0.8 are common in industrial practice if sensible guidelines are followed. Geometry is important, and diffuser divergence angles must be limited to small angles of about 7–9°. Lengths of about 12–18 times the inlet height are needed to get good diffusion without significant flow separation (reversal). Inlets with the least flow distortion possible are desired to achieve high levels of C_p and η. Inlet swirl may help or hinder performance; moderate turbulence can help diffusers of larger divergence angle. Curvature against the axis of a diffuser will invariably degrade performance. When sensible practice is employed, effectiveness levels of about 0.7–0.8 are common; higher values are possible for unusual designs, and lower values are common for poor designs as may result when space is limited and compromises are necessitated. *See* TURBULENT FLOW.

Assessment. The above guidelines permit tentative assessment of expected diffuser performance. For a given diffuser, the area ratio may be calculated first; and then the ideal static pressure recovery coefficient, using Eq. (3) for rectilinear diffusers and alternative equations for other cases, may be computed next. Effectiveness can be estimated from the guidelines given above subject to diffuser geometry and inlet fluid-dynamic conditions.

Losses. Losses in the diffuser can be related to the pressure recovery. In the case of incompressible flow, the loss coefficient can be defined as the ratio of the total pressure drop through the diffuser to the inlet kinetic energy. An approximate relationship can be developed from the principles given above whereby the loss coefficient equals the difference between the ideal static pressure recovery coefficient and the actual recovery coefficient. By using the basic concepts presented above concerning pressure recovery, including the ideal pressure recovery, a very sensible estimate of losses within the diffuser, that is, the degradation of total pressure, can be estimated. *See* FLUID FLOW; FLUID-FLOW PRINCIPLES.

David Japikse

Bibliography. D. Japikse, *Turbomachinery Diffuser Design Technology*, 1984, 1995; D. S. Miller, *Internal Flow Systems*, 2d ed., 1990.

Diffusion

The transport of matter from one point to another by random molecular motions. It occurs in gases, liquids, and solids.

Diffusion plays a key role in processes as diverse as permeation through membranes, evaporation of liquids, dyeing textile fibers, drying timber, doping silicon wafers to make semiconductors, and transporting of thermal neutrons in nuclear power reactors. Rates of important chemical reactions are limited by how fast diffusion can bring reactants together or deliver them to reaction sites on enzymes or catalysts. The forces between molecules and molecular sizes and shapes can be studied by making diffusion measurements. *See* CELL PERMEABILITY; EVAPORATION; INTEGRATED CIRCUITS; SEMICONDUCTOR; THERMAL NEUTRONS.

Fluids

Molecules in fluids (gases and liquids) are constantly moving. Even in still air, for example, nitrogen and oxygen molecules ricochet off each other at bullet speeds. Molecular diffusion is easily demonstrated by pouring a layer of water over a layer of ink in a narrow glass tube. The boundary between the ink and water is sharp at first, but it slowly blurs as the ink diffuses upward into the clear water. Eventually, the ink spreads evenly along the tube without any help from stirring.

Diffusion equations. Adolph Fick discovered the mathematical laws of diffusion in 1855. He found that the rate of diffusion of a substance is proportional to the gradient in its concentration (C). This is expressed in his first law, Eq. (1), where J is

$$J = -D\frac{\partial C}{\partial z} \tag{1}$$

the flux density, the amount of substance diffusing parallel to the z coordinate axis per unit time per unit area. The minus sign reflects the natural tendency of substances to diffuse down concentration gradients, from higher to lower concentrations. Similar equations can be written for diffusion along the x and y coordinates in three-dimensional space.

The quantity D is called the diffusion coefficient or the diffusivity. It measures the flux produced by a given concentration gradient. If the concentration is expressed in units of moles per cubic centimeter and the flux in moles per square centimeter per second, the units of D are square centimeters per second. Mass concentrations (grams per cubic centimeter) and mass fluxes (grams per square centimeter per second) are also used. Since J is the rate of diffusion per unit area, the amount of substance diffusing per second through area A is AJ.

Fick's law for the flow of matter in a concentration gradient is analogous to Fourier's law for heat conduction in a temperature gradient. *See* CONDUCTION (HEAT).

Fick's second law. Diffusion changes the distribution of molecules as time passes. Each second, $AJ(z)$ moles of substance diffuse into a slab of solution of area A and thickness Δz (**Fig. 1**) across its left face. Simultaneously, $AJ(z + \Delta z)$ moles per second diffuse out of the slab across its right face. The net rate of diffusion into the slab is $AJ(z) - AJ(z + \Delta z)$.

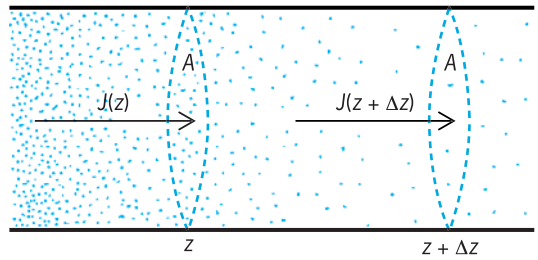

Fig. 1. Diffusion in a tube of cross-sectional area *A*. The net rate of diffusion into the fluid between *z* and *z* + Δ*z* is *AJ*(*z*) − *AJ*(*z* + Δ*z*).

Dividing the rate of diffusion into the slab by its volume $A\Delta z$ gives the rate of change of concentration, as in Eq. (2), and combining Eqs. (1) and (2) gives Fick's second law of diffusion, Eq. (3). As

$$\frac{\Delta C}{\Delta t} = -\frac{AJ(z + \Delta z) - AJ(z)}{A\Delta z} \tag{2}$$

$$\frac{\partial C}{\partial t} = -\frac{\partial J}{\partial z} = \frac{\partial}{\partial z}\left[D\frac{\partial C}{\partial z}\right] \tag{3}$$

time passes, the concentration changes in proportion to the change in the concentration gradient along the diffusion path. This illustrates diffusion's fundamental role: smoothing out gradients in composition.

Diffusion coefficients can be treated as constants if the concentration gradients are not too large. The simpler version of Fick's second law, Eq. (4), then applies.

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial z^2} \tag{4}$$

This equation is used to predict the concentrations of diffusing substances in hundreds of different applications. Slightly different equations are used for diffusion in cylindrical or spherical geometries. *See* DIFFERENTIAL EQUATION.

Onsager's equation. Just as positive charge flows from higher to lower electric potential, molecules diffuse from regions of higher to lower chemical potential (a thermodynamic property). Gradients in the chemical potential, not concentration gradients, are the true driving forces for diffusion. A more fundamental description of diffusion states that the flux of a diffusing substance is proportional to the gradient in its chemical potential, as in Eq. (5).

$$J = -L\frac{\partial \mu}{\partial z} \tag{5}$$

The thermodynamic diffusion coefficient L is proportional to the mobility of a diffusing substance, the ease with which it moves through a fluid. Small molecules in low-viscosity fluids have the largest mobilities. Fick's diffusion coefficient D equals $L\partial\mu/\partial C$.

This result shows that diffusion coefficients are not purely kinetic quantities because they depend on thermodynamic properties too. *See* CHEMICAL THERMODYNAMICS.

Diffusion processes. These comprise chemical interdiffusion, self-diffusion, and intradiffusion.

Chemical interdiffusion. By far the most frequently studied diffusion process is the molecular mixing of two different chemical substances. An example is the interdiffusion of oxygen and nitrogen. At first glance, there are two diffusion coefficients to consider, one for each substance. However, there is no bulk movement of fluid during diffusion. This restriction leads to the surprising result that the interdiffusion of two substances is described by a single diffusion coefficient, D: $J_1 = -D\partial C_1/\partial z$ and $J_2 = -D\partial C_2/\partial z$. The quantity D is called the interdiffusion coefficient or the mutual diffusion coefficient.

Gas molecules travel over distances of many molecular diameters before being deflected by collisions with other molecules. This gives gas mixtures the largest interdiffusion coefficients, approximately 0.1 $cm^2\ s^{-1}$ at room temperature and atmospheric pressure. A gas molecule diffuses about 0.1–0.5 cm in 1 s. Diffusion in liquids is about 10,000 times slower because the molecules are packed more closely. Solid-state diffusion is even slower because the molecules are locked in position by their neighbors, except for infrequent jumps.

Self-diffusion. Diffusion occurs in pure substances too, and is called self-diffusion. It can be studied by tagging some of the diffusing molecules with isotopes. For example, the self-diffusion coefficient of water is measured by releasing trace amounts of water labeled with radioactive tritium ($^3H^1HO$) at one end of a tube filled with ordinary water (^1H_2O) and then monitoring the spread of radioactivity along the tube. Magnetized molecules also serve as tagged species. *See* ISOTOPE; NUCLEAR MAGNETIC RESONANCE (NMR).

Intradiffusion. The term intradiffusion refers to the interchange of tagged and untagged species in systems of uniform chemical composition. To measure the intradiffusion coefficient of benzene in a benzene-heptane mixture containing 1 mole of benzene per liter, a 1-mole-per-liter solution of benzene in heptane is brought into contact with another 1-mole-per-liter benzene solution in which a portion of the total benzene molecules are tagged with radioactive ^{14}C.

Gases. A number of techniques are used to measure diffusion in gases. In a two-bulb experiment, two vessels of gas are connected by a narrow tube through which diffusion occurs. Diffusion is followed by measuring the subsequent changes in the composition of gas in each vessel. Excellent results are also obtained by placing a lighter gas mixture on top of a denser gas mixture in a vertical tube and then measuring the composition along the tube after a timed interval.

Rates of diffusion in gases increase with the temperature (T) approximately as $T^{3/2}$ and are inversely proportional to the pressure. The interdiffusion coefficients of gas mixtures are almost independent of the composition.

Kinetic theory shows that the self-diffusion coefficient of a pure gas is inversely proportional to both the square root of the molecular weight and the square of the molecular diameter. Interdiffusion coefficients for pairs of gases can be estimated by taking averages of the molecular weights and collision diameters. Kinetic-theory predictions are accurate to about 5% at pressures up to 10 atm (1 megapascal). Theories which take into account the forces between molecules are more accurate, especially for dense gases. *See* KINETIC THEORY OF MATTER.

Liquids. The most accurate diffusion measurements on liquids are made by layering a solution over a denser solution and then using optical methods to follow the changes in refractive index along the column of solution. Excellent results are also obtained with cells in which diffusion occurs between two solution compartments through a porous diaphragm. Many other reliable experimental techniques have been devised.

Room-temperature liquids usually have diffusion coefficients in the range $0.5\text{–}5 \times 10^{-5}\ cm^2\ s^{-1}$. Diffusion in liquids, unlike diffusion in gases, is sensitive to changes in composition but relatively insensitive to changes in pressure. Diffusion of high-viscosity, syrupy liquids and macromolecules is slower. The diffusion coefficient of aqueous serum albumin, a protein of molecular weight 60,000 atomic mass units, is only $0.06 \times 10^{-5}\ cm^2\ s^{-1}$ at 25°C (77°F).

When solute molecules diffuse through a solution, solvent molecules must be pushed out of the way. For this reason, liquid-phase interdiffusion coefficients are inversely proportional to both the viscosity of the solvent and the effective radius of the solute molecules. Accurate theories of diffusion in liquids are still under development. *See* VISCOSITY.

Applications. Diffusion is involved in a great diversity of processes.

Tracer diffusion. If tracer molecules are released in a fluid in which there is no convection, such as stirring, Fick's second equation can be integrated to predict how rapidly the tracer molecules spread out by random thermal motions.

If tracer molecules are released at position $z = 0$ in a tube of cross-sectional area A, they will cluster near the origin initially, and slowly disperse (**Fig. 2***a*). The bell-shaped concentration profiles are identical to the gaussian distribution of random errors employed in statistical sciences. This is not accidental, but a direct consequence of the random motions of diffusing molecules. In fact, the standard deviation of the tracer profile, a measure of its width, is $\sqrt{2Dt}$ after diffusion for time t. *See* STATISTICS.

Mixing without convection. Two fluids of different composition can be rapidly mixed by stirring. Without convection, the rate of mixing is diffusion controlled and much slower. As time passes, interdiffusion

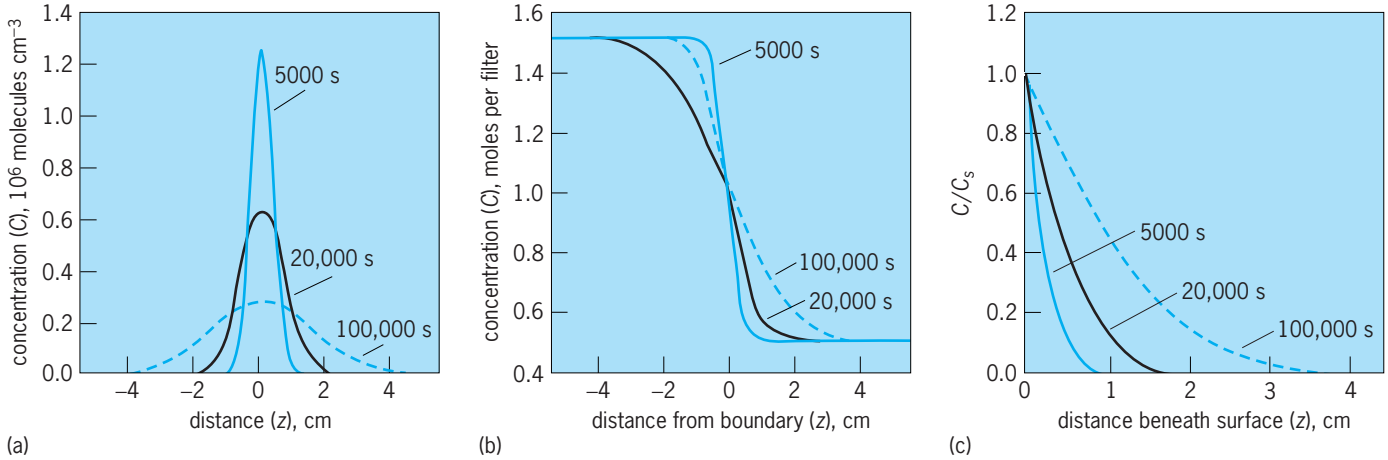

Fig. 2. Concentration profiles. (*a*) Diffusion of 10^6 tracer molecules in a tube of 1-cm^2 cross-sectional area. The molecules are initially located at position $z = 0$. $D = 1 \times 10^{-5}$ cm^2 s^{-1}. (*b*) Interdiffusion of a solution of concentration 1.5 moles per liter into a solution of concentration 0.5 mole per liter across a sharp initial boundary. (*c*) Absorption of a gas into a convection-free liquid. Concentration (*C*) of dissolved gas is given in terms of saturation concentration at liquid surface.

gradually smooths out a sharp initial boundary between two solutions (Fig. 2*b*).

Gas absorption. Concentration profiles can also be measured for the absorption and diffusion of a soluble gas into a convection-free liquid (Fig. 2*c*). The surface of the liquid is saturated with dissolved gas. After exposure for time t, the dissolved gas penetrates the absorbent to an average depth of about $\sqrt{Dt}$.

The diffusion flux at the surface of the liquid gives the rate of dissolution of gas per unit surface area of liquid. This rate is proportional to the solubility of the gas and the square root of the diffusion coefficient. The adsorption process slows down as time passes as $t^{-1/2}$. Identical results are obtained for the dissolution of a solid into a stagnant liquid.

Brownian motion. Colloidal particles suspended in a liquid move by a series of small, chaotic jumps. This kind of diffusion is called brownian motion. It was first observed by looking at suspensions of pollen grains in water under a microscope.

A. Einstein analyzed the motion of a randomly jumping particle. He showed that a particle making v jumps per second of length ℓ has diffusion coefficient $v\ell^2/2$. This result provides a useful physical model of the diffusion process. It also accounts for the slowness of diffusion in liquids relative to gases: The distance between collisions, ℓ, is relatively small for liquids. Diffusion in solids is even slower because both the step size and jump frequency are small.

If a particle is allowed to diffuse freely in one dimension, its root-mean-square displacement after time t is $\sqrt{2Dt}$. This result is frequently used to estimate the distances covered by diffusing molecules. A molecule with diffusion coefficient 1×10^{-5} cm^2 s^{-1} diffuses about 1 cm in a day but only about 25 cm in a year. The root-mean-square displacement for three-dimensional diffusion is $\sqrt{6Dt}$ (**Fig. 3**). *See* BROWNIAN MOVEMENT; COLLOID.

Porous materials. Diffusion through a porous solid is slower than diffusion through an equal volume of homogeneous fluid for two reasons. First, the diffusing molecules must travel greater distances to move around solid obstructions. Second, diffusion can occur only through the pore fraction of the total cross-sectional area.

A practical approach to this problem is to treat the porous medium as a homogeneous phase with an effective diffusion coefficient, D_{eff}. Values of D_{eff} are typically 3–10 times smaller than the diffusion coefficient D for the pore fluid. The D_{eff}/D ratio depends on the void fraction and details of the pore network, such as the number of interconnections and dead ends.

Effusion. At low pressures, gas molecules can escape through a pinhole in the wall of a container without colliding with other molecules. This process is called effusion. It is used to measure the vapor pressure of solids and liquids at high temperatures. The rate of effusion is inversely proportional to the

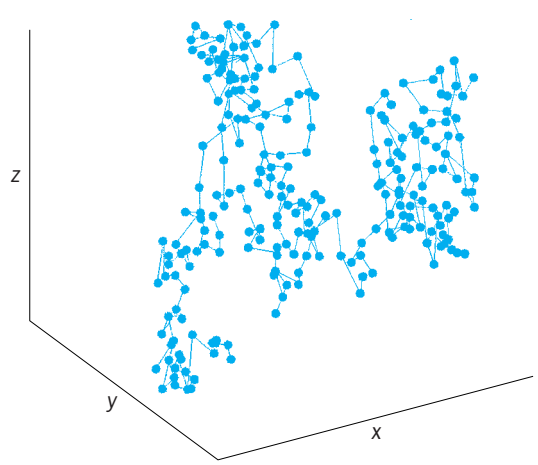

Fig. 3. Brownian motion of a particle in three dimensions.

square root of the molecular weight of the gas. Gases such as $^{235}UF_6$ and $^{238}UF_6$ can therefore be separated by effusion through porous membranes. *See* ISOTOPE SEPARATION.

Electrolyte solutions. Fick's equations also apply to diffusion in electrolyte solutions. The simplest case is a single salt dissolved in a solvent, such as aqueous sodium chloride. The negative and positive ions must diffuse at the same speed to keep all parts of the solution electrically neutral. This means that only one flux is independent and a single diffusion coefficient is sufficient to describe interdiffusion of dissolved salt and solvent.

Interdiffusion coefficients of electrolytes are sensitive to changes in concentration. For a dilute solution of a 1:1 salt, such as NaCl, the interdiffusion is the harmonic average of the diffusion coefficients of the ions: $D = 2D_+ D_-/(D_+ + D_-)$. *See* ELECTROLYTE; ELECTROLYTIC CONDUCTANCE.

Multicomponent interdiffusion. Mixtures of three or more components are of considerable practical importance. Diffusion in these systems differs in important ways from diffusion in simpler two-component mixtures. Adding a new component to an existing mixture changes the rates of diffusion of all the components in the mixture. Also, a new feature appears: coupled diffusion, where the gradient in concentration of each substance causes diffusion of the other substances in the mixture. Coupled diffusion can drive substances from regions of lower to higher concentration, opposite to the expected direction. Diffusion in multicomponent mixtures is described by adding coupled-diffusion terms to Fick's equations.

Coupled diffusion is important in concentrated solutions and in mixed electrolytes because the forces between the diffusing materials are significant. For example, in aqueous solutions of potassium chloride (KCl) and hydrogen chloride (HCl), owing to electrostatic forces, the flux of potassium chloride caused by the gradient in hydrogen chloride can be larger than the flux of potassium chloride caused by its own gradient. Hydrogen chloride diffuses much more rapidly in water-KCl-HCl mixtures than in binary water-HCl mixtures.

Diffusion with convection. In moving fluids, substances are transported by diffusion as well as convection. If the concentration of a substance is C and the solution moves with velocity U_z in the z direction, the flux of substance due to the bulk movement of the solution is CU_z. The net flux is the sum of the contributions from diffusion and convection, as in Eq. (6).

$$J = -D\frac{\partial C}{\partial z} + CU_z \tag{6}$$

See CONVECTION (HEAT).

Dispersion in pipelines. The dispersion of a tracer in a liquid carrier stream is an interesting example of combined diffusive and convective transport. The liquid near the center of the tube moves more quickly than the liquid near the tube wall. This causes an injected tracer to spread out along the axis of the tube. If the flow rate is low enough to prevent turbulence, the combined actions of axial convection and radial diffusion shape the dispersed tracer into a bell-shaped plume. Dispersion is minimized by using short, narrow tubes. Diffusion coefficients of liquids and gases are determined by measuring dispersion profiles. *See* FLUID FLOW.

Ionic diffusion in electric fields. When a salt solution is placed between charged electrodes, electrical forces draw the positive ions toward the negative electrode and the negative ions toward the positive electrode. The flux of an ion caused by an electric field is proportional to the field strength and to the ion's charge, diffusion coefficient, and concentration.

Dividing the drift velocity of a migrating ion by the electric field gives the ion's electric mobility. The electric mobilities of aqueous ions are usually in the range 0.0003–0.001 $cm^2\ s^{-1}\ V^{-1}$. Thus, a 10 V cm^{-1} electric field generates ionic drift velocities of about 0.003–0.01 cm s^{-1}.

Thermal diffusion. If one end of a convection-free column of solution is heated and the other end cooled, solute spontaneously migrates to the cooler regions and solvent migrates to the warmer regions, or vice versa. Diffusion caused by temperature gradients is called thermal diffusion.

Diffusion in nonisothermal solutions is described by adding the term $-CD_T\partial T/\partial z$ to the right-hand side of Eq. (1) to allow for the flux caused by the temperature gradient, giving Eq. (7). The quantity D_T is

$$J = -D\frac{\partial C}{\partial z} - CD_T\frac{\partial T}{\partial z} \tag{7}$$

the thermal diffusion coefficient. Positive values of D_T indicate migration to the cooler parts of a mixture. Substances with negative D_T values migrate to warmer regions.

The accumulation of material at the ends of the solution column produces concentration gradients. Eventually, thermal migration is balanced by ordinary diffusion down the concentration gradient, and a steady state is reached in which there are no further changes in concentration. The ratio of the thermal diffusion coefficient to the interdiffusion coefficient gives the fractional change in solute concentration per degree at steady state, as in Eq. (8). Typical D_T/D

$$\frac{D_T}{D} = -\frac{1}{C}\left(\frac{dC}{dT}\right)_{\text{steady state}} \tag{8}$$

values are about 0.001 K^{-1}, which corresponds to a 1% concentration difference for a 10 K (18°F) temperature difference.

Thermal diffusion is used to purify isotopes and to fractionate polymers, but for most solutions the separations are too small to compete with other techniques. A few solutions have D_T/D values as large as 0.03 K^{-1}. Thus, a 10 K (18°F) temperature difference

can produce a 30% change in concentration, which is not negligible. Thermal diffusion effects can be significant if large temperature gradients are imposed on a solution. Derek G. Leaist

Solids

Diffusion in solids is an important topic of physical metallurgy and materials science since diffusion processes are ubiquitous in solid matter at elevated temperatures. They play a key role in the kinetics of many microstructural changes that occur during the processing of metals, alloys, ceramics, semiconductors, glasses, and polymers. Typical examples of such changes include nucleation of new phases, diffusive phase transformations, precipitation and dissolution of a second phase, recrystallization, high-temperature creep, and thermal oxidation. Direct technological applications concern diffusion doping during the fabrication of microelectronic devices, solid electrolytes for battery and fuel cells, surface hardening of steels through carburization or nitridation, diffusion bonding, and sintering. *See* CREEP (MATERIALS); FUEL CELL; HEAT TREATMENT (METALLURGY); METAL, MECHANICAL PROPERTIES OF; PHASE TRANSITIONS; PLASTIC DEFORMATION OF METAL; SINTERING; SOLID-STATE BATTERY; SURFACE HARDENING OF STEEL.

The atomic mechanisms of diffusion are closely connected with defects in solids. Point defects such as vacancies and interstitials are the simplest defects and often mediate diffusion in an otherwise perfect crystal. Dislocations, grain boundaries, phase boundaries, and free surfaces are other types of defects in a crystalline solid. They can act as diffusion short circuits because the mobility of atoms along such defects is usually much higher than in the lattice. *See* CRYSTAL DEFECTS.

Diffusion coefficients. The diffusion of atoms through a solid can be described by Fick's equations (1) and (3). For a constant (concentration-independent) diffusivity, solutions of these equations exist for a wide variety of initial and boundary conditions and permit a determination of the diffusion coefficient (D) either from measurements of the concentration distribution as function of the position and time or from measurements of diffusion fluxes and concentration gradients. For example, if a diffusing species is initially concentrated in a thin layer, then after a time t it will have a bell-shaped concentration profile whose width depends on D and t in the same manner as in tracer diffusion.

For diffusion in a material with a chemical composition gradient, solutions of Eq. (3) in closed form are usually not possible. Then numerical methods are often used to deduce so-called interdiffusion coefficients.

Any diffusion coefficient is a material parameter, which depends on the composition, slightly on pressure, and strongly on temperature T. Very often the temperature dependence for a given material is described by an Arrhenius equation (9), where Q is the

$$D = D_0 \exp\left(-\frac{Q}{RT}\right) \tag{9}$$

activation enthalpy, D_0 is a preexponential factor, and R is the gas constant. *See* ENTHALPY.

The most fundamental diffusion process in a solid is self-diffusion and refers to diffusion of A atoms in a solid element A or in an A-B alloy or compound. Following Eq. (9), the self-diffusion coefficients of many metals increase by 7 to 10 orders of magnitude as the absolute temperature is increased from half the melting temperature to the melting temperature. Diffusion of a solute B in a very dilute A-B alloy with the solvent A is denoted as impurity diffusion. Self-diffusion and impurity diffusion in a homogeneous solid can be studied, for example, by utilizing small amounts of radioactive trace atoms. Diffusion in an inhomogeneous solid with a composition gradient is denoted as interdiffusion or sometimes as chemical diffusion.

Crystalline solids differ from gases and liquids in that the symmetry of the atomic arrangement can give rise to different values of D in different directions of the crystal lattice. In isotropic materials such as glasses and polycrystals with random orientation of the individual grains and in cubic single crystals, only one D value characterizes the material. Two diffusion coefficients are necessary for uniaxial (hexagonal, tetragonal, and trigonal) crystals, and three D values characterize diffusion in single crystals of lower symmetry. For example, the D value for diffusion in the basal plane of a hexagonal single crystal is different from the D value perpendicular to this plane. *See* CRYSTAL STRUCTURE.

Mechanisms. Diffusion in crystalline matter takes place by a series of jumps of individual atoms from site to site throughout the crystal. Because of the lattice vibrations, atoms in a crystal oscillate around their equilibrium (substitutional or interstitial) positions with frequencies, ν_0, of the order of the Debye frequency (typically 10^{12}–10^{13} Hz). Usually an atom is confined to a certain site by a potential barrier with barrier height G^M, which corresponds to the Gibbs free-energy difference between the configuration with the jumping atom at the saddlepoint and with that at its equilibrium position. Diffusion is thermally activated, meaning that a fluctuation of thermal energy pushes the atom over the energy barrier which exists between two neighboring sites. From reaction rate theory, the jump frequency ω at which an atomic jump will occur into an empty neighboring site may be written in the form of Eq. (10).

$$\omega = \nu_0 \exp\left(-\frac{G^M}{RT}\right) \tag{10}$$

See BOLTZMANN STATISTICS; FREE ENERGY; LATTICE VIBRATIONS; SPECIFIC HEAT OF SOLIDS; STATISTICAL MECHANICS.

At an atomic level, the diffusivity of (for example, radioactive) tagged atoms in a cubic crystal can be described by Eq. (11), where d is the jump distance

$$D = \frac{1}{6} f d^2 \omega p \tag{11}$$

(some fraction of the lattice parameter), p is the probability that the neighboring site of the jumping atom is empty, and f is the correlation factor, a term which expresses the degree of randomness of successive jumps. Various atomic mechanisms have been identified for bulk diffusion in crystals.

Interstitial mechanism. Solute atoms which are much smaller than the lattice atoms, as for example, hydrogen, carbon, nitrogen, and oxygen in metals, are incorporated in interstitial sites. They diffuse by jumping from interstitial site to interstitial site (**Fig. 4***a*). For a dilute interstitial alloy, the probability p in Eq. (10) is unity. The magnitude of the barrier height G^M is also relatively small, with the result that the diffusion coefficients for interstitial diffusion can be many orders of magnitude larger than those for the self-diffusion of lattice atoms. Interstitial diffusivities near the melting temperature of the solvent can be as high as diffusivities in liquids.

Vacancy mechanism. Self-atoms or substitutional solute atoms migrate by jumping into a neighboring vacant site. Self-diffusion in metals and alloys, and in many ionic crystals and ceramic materials, occurs via the vacancy mechanism (Fig. 4*b*). Attractive or repulsive interactions between the vacancy and substitutional solute atoms may lead to higher or lower diffusivities compared with self-diffusion of the pure solvent. *See* IONIC CRYSTALS.

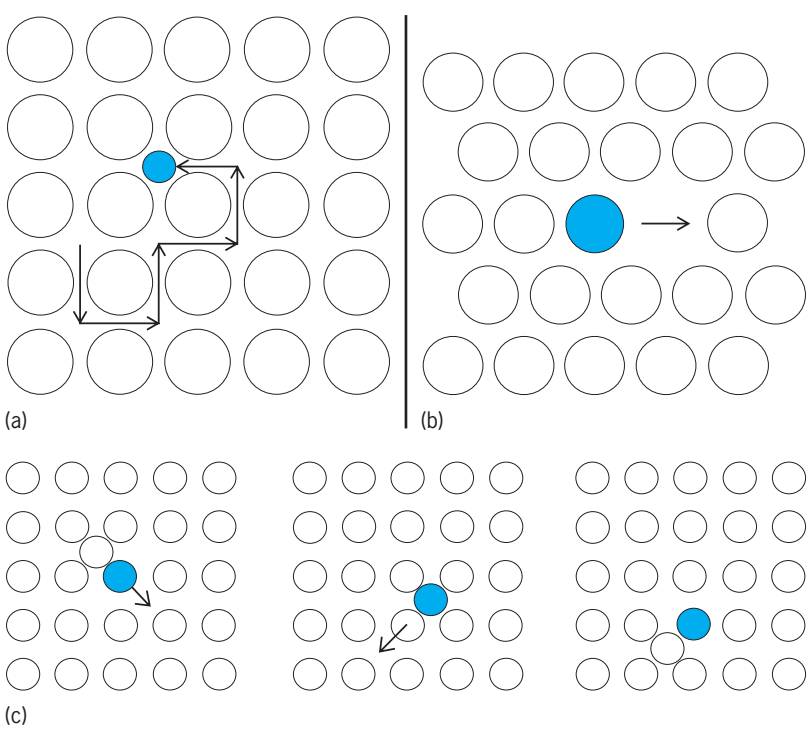

Fig. 4. Diffusion mechanisms. (*a*) Direct interstitial mechanism. (*b*) Vacancy mechanism. (*c*) Self-interstitial mechanism.

In crystals with ionic bonding characteristics such as alkali halides, in ceramics, and in some intermetallic compounds, atoms do not jump between the sublattices of the constituents. This implies that the diffusion coefficients of the various atomic species can differ significantly.

Self-interstitial mechanism. In this case, self-interstitials—extra atoms located between lattice sites—act as diffusion vehicles. A self-interstitial replaces an atom on a substitutional site which then replaces again a neighboring lattice atom (Fig. 4*c*). Self-interstitials are responsible for diffusion in the silver sublattice of silver halides. In silicon, the base material of microelectronic devices, the interstitialcy mechanism dominates self-diffusion and plays a prominent role in the diffusion of many solute atoms, including the most important doping elements. In the close-packed metals, this mechanism is unimportant under thermal equilibrium conditions because of the fairly high formation enthalpy of self-interstitials. It is, however, important for radiation-damage-induced diffusion, where self-interstitials and vacancies are created athermally by irradiation of a crystal with energetic particles. *See* RADIATION DAMAGE TO MATERIALS.

Interstitial-substitutional exchange mechanisms. Some solute atoms (B) may be dissolved on interstitial (B_i) and substitutional sites (B_s) of a solvent crystal (A). For so-called hybrid solutes the diffusivity of B_i is much higher than the diffusivity of B_s, whereas the opposite is true for the solubilities. Under such conditions, the incorporation of B atoms can occur by the fast diffusion of B_i and the subsequent change-over to B_s.

If the change-over involves vacancies (V) according to reaction (12), the mechanism is denoted as the

$$B_i + V \rightleftharpoons B_s \tag{12}$$

dissociative mechanism (sometimes also known as the Frank-Turnbull mechanism or the Longini mechanism). The rapid diffusion of copper in germanium and of some foreign metallic elements in polyvalent metals like lead, titanium, and zirconium has been attributed to this mechanism.

If the change-over involves a self-interstitial of the solvent (A_i) according to reaction (13), the mechanism is denoted as the kick-out mechanism. Examples of this mechanism have been established for some rapidly diffusing foreign elements (for example, gold, platinum, and zinc) in silicon.

$$B_i \rightleftharpoons B_s + A_i \tag{13}$$

Short-circuit diffusion. Any real crystalline material usually contains extended defects such as dislocations, grain boundaries in polycrystals, interphase boundaries in polyphase materials, and free surfaces. These regions are more open structures than the adjoining crystal lattice. The jump frequency of atoms in these regions is much higher and the activation

enthalpy of diffusion lower than in the perfect crystal. Because of the lower activation enthalpy, the mobility of atoms along extended defects decreases less rapidly with decreasing temperature than in the bulk. The result is that at low temperatures (below about half of the melting temperature) diffusion through the lattice becomes slower than diffusion along the extended defects. Then these defects constitute paths of high diffusivity which at low enough temperatures are said to short circuit lattice diffusion.

The rapid diffusion along high-diffusivity paths is often important in surface- and thin-film technologies. In thin-film microelectronic devices, shortcircuit diffusion among various components of the multilayered structures is one of the principal modes of device failure. The fast diffusion character of grain-phase and interphase boundaries also results in the fact that many solid-state reactions at lower temperatures, such as discontinuous precipitation and eutectoid transformations, are controlled by diffusion along boundaries.

Superionic conductors. Superionic conductors, sometimes referred to as fast-ion conductors or solid electrolytes, are solids which have an exceptionally high ionic conductivity over a reasonable temperature range. In some cases the conductivity approaches the magnitude found in molten salts and aqueous solutions of strong electrolytes. Typical values of the conductivity of a superionic conductor are in the range of 0.001 to 10 $(\Omega\ \mathrm{cm})^{-1}$. The conductivity of copper near room temperature is $6 \times 10^5\ (\Omega\ \mathrm{cm})^{-1}$; a typical ionic crystal has a conductivity of $10^{-4}\ (\Omega\mathrm{cm})^{-1}$ close to the melting temperature; an aqueous solution of 0.1 N NaCl has a conductivity of $10^{-2}\ (\Omega\ \mathrm{cm})^{-1}$. Diffusivity D and ionic conductivity σ are related via the Nernst-Einstein equation (14), where k_B denotes Boltzmann's constant,

$$D = \frac{\sigma k_B T}{q^2 N} \tag{14}$$

T absolute temperature, q the ionic charge, and N the volume density of mobile ions. Superionic conduction implies fast diffusion of the ions. Typically a conductivity of 1 $(\Omega\ \mathrm{cm})^{-1}$ corresponds to a diffusivity of about $10^{-5}\ \mathrm{cm}^2\ \mathrm{s}^{-1}$, which is in the same order of magnitude as diffusion in liquids. Because of their high conductivity, superionic conductors are thought of as solid electrolytes with many applications, both realized and potential, for example, in batteries, fuel cells, and sensors. *See* BATTERY; FUEL CELL.

The high ionic conductivity, depending on the material, can be of either cationic or anionic character. As a general rule, superionic conductors are materials with an open structure, which allows the rapid motion of relatively small ions. The most important classes of superionic conductors are the following:

1. Many superionic conductors have a fixed anion lattice and a three-dimensionally disordered cation lattice. Example are those of the silver iodide type (reported as early as 1913), which display a first-order phase transition between a superionic phase at high temperatures and a normal conducting phase at low temperatures. In the case of pure silver iodide (AgI), the high-temperature phase α-AgI exists above 146°C (295°F) with a body-centered cubic sublattice of fixed I^- ions. Each unit cell presents 42 interstitial sites over which the two Ag^+ ions can distribute themselves (**Fig. 5**). Since there are many more sites than Ag^+ ions, the latter can migrate easily. Materials derived from α-AgI with the formula MAg_4I_5, where M = K, Rb, Cs, and NH_4, are superionic at ambient temperature.

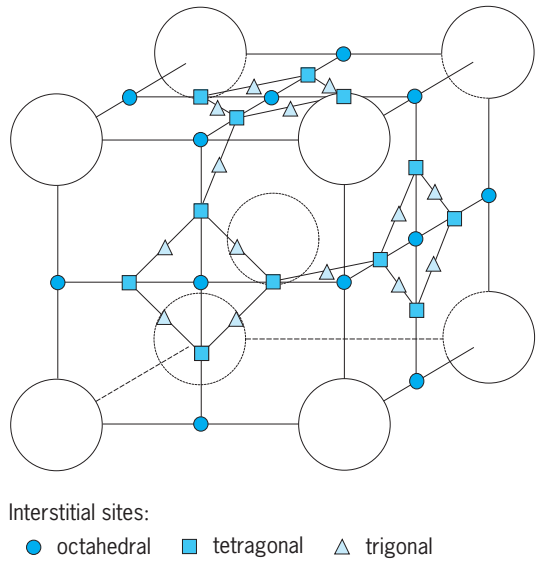

Fig. 5. Crystal structure of α-silver iodide.

2. Some fast-ion conductors have a hexagonal structure in which small cations can move either two-dimensionally in layers or one-dimensionally in channels. Materials in which two-dimensional diffusion occurs are the β-aluminas, which are nonstoichiometric aluminates with the formula $M_2O_nAl_2O_3$, where M = Na, Li, Ag, ... and n ranges from 5 to 11. A material with one-dimensional diffusion is β-eucryptite, an aluminosilicate ($LiAlSiO_4$), in which the lithium ions are located in channels parallel to the hexagonal axis.

3. Zeolites are natural or synthetic hydrated aluminosilicates with an open three-dimensional structure, in which ions are incorporated in cavities of the lattice. *See* ZEOLITE.

4. Superionic conductors are also found among compounds with the fluorite structure, such as some halides ($SrCl_2$, CaF_2, PbF_2) and oxides like ZrO_2 (zirconia), and exhibit fast-ion conductivity. Zirconia is stabilized in the cubic structure by additions of CaO and Y_2O_3.

5. A number of glasses, including lithium borate glass, have been examined as fast-ion conductors.

6. Solutions of ionic salts in certain polymers (polymer electrolytes) can exhibit conductivities in the

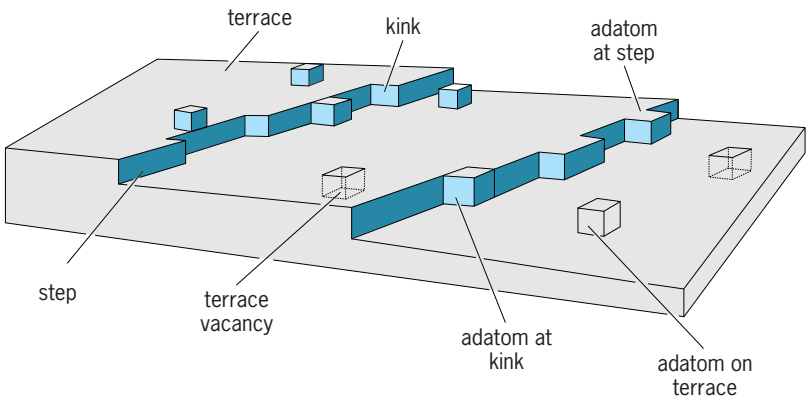

Fig. 6. Model of the surface of a single crystal.

intermediate range. The fact that these materials can be cast as thin, flexible foils led to applications as battery electrolytes.

Surface diffusion. Surface diffusion is the motion of atoms or molecules on top of the surface of a solid material. The diffusivity of an atom on the surface is usually much higher than the diffusivity of the same atom in the interior of the solid, and its activation enthalpy is lower than in the bulk material.

The surface of a crystalline solid has a complicated but well-defined crystallographic structure, which can be described in terms of terraces (low-indexed, vicinal planes), steps, and kinks (**Fig. 6**). Adatoms on the terraces and at steps and surface vacancies occur in thermal equilibrium. More complex defects, such as dislocation emerging points, impurity clusters, high-index facets, and hillocks, may be present. The structure of a real surface will give rise to a mobility of atoms that depends on their position. It will be different, for example, on terraces and on steps. Often an anisotropy of diffusion is also observed, depending on the orientation of vicinal surfaces.

Surface diffusion at temperatures that are not too high occurs via elementary jumps of surface vacancies and adatoms over one atomic distance. When the temperature rises, multiple jumps of adatoms over several atomic distances are possible. Near the melting temperature, even pseudo-liquid behavior has been observed in molecular-dynamic simulations.

Various methods can be used to study surface diffusion. Among them are observations of the migration of single adatoms by means of field-ion microscopy and scanning tunneling microscopy; observations of the redistribution of radiotracer atoms initially deposited on one part of a surface; and topographic methods that monitor the diffusion-induced change in the shape of a solid, which is driven by surface tension.

Radiation-enhanced diffusion. The diffusion coefficient of atoms in solids is dominated by the product of the equilibrium concentration of the diffusion-relevant defects times their mobility. Equilibrium concentration and defect mobility drop steeply with decreasing temperature. As a consequence of this strong temperature dependence, hardly any atomic diffusivity will remain below about 40% of the melting temperature.

In radiation-enhanced diffusion, additional point defects are produced by the incident radiation. High-energy electrons (available, for example, in a high-voltage electron microscope) colliding with the lattice atoms can produce pairs of vacancies and self-interstitials (Frenkel pairs). Irradiation with heavier particles, such as protons, fast neutrons, alpha particles, and heavy ions, causes radiation damage to the crystal lattice which introduces many Frenkel pairs per incident particle. Therefore, not only will the usual diffusion path be enhanced, but also new channels can be opened by way of defects (such as self-interstitials in the case of metals) which are not available in normal thermally activated diffusion. Radiation-enhanced diffusion can occur at relatively low temperatures provided that thermal activation permits sufficient mobility of at least one of the defect species produced during irradiation. Radiation-enhanced diffusion is of importance in materials science since it affects the properties of materials exposed to particle irradiation. Helmut Mehrer

Bibliography. E. L. Cussler, *Diffusion: Mass Transfer in Fluid Systems*, 2d ed., 1997; B. Gebhart, *Heat Conduction and Mass Diffusion*, 1993; M. E. Glicksman, *Diffusion in Solids*, John Wiley, 1999; H. Mehrer (ed.), *Diffusion in Solid Metals and Alloys*, 1990; H. Mehrer, C. Herzig, and N. A. Stolwijk (eds.), *Diffusion in Materials*, Scitec Publications, 1997; G. Murch (ed.), *Diffusion in Solids: Unsolved Problems*, 1992; J. S. Newman, *Electrochemical Systems*, 2d ed., 1991; J. Philibert, *Atom Movements: Diffusion and Mass Transport in Solids*, 1991; B. W. Rossiter and J. F. Hamilton (eds.), *Physical Methods of Chemistry*, vol. 6, 2d ed., 1992; P. Shewmon, *Diffusion in Solids*, 1989.

Digenea

A group of parasitic flatworms or flukes constituting a subclass or order of the class Trematoda, in the phylum Platyhelminthes. The name Digenea refers to the two types of generations in the life cycle: (1) the germinal sacs which parasitize the intermediate host (a mollusk or rarely an annelid) and reproduce asexually; and (2) the adult which is primarily endoparasitic in vertebrates and reproduces sexually. The adult usually is hermaphroditic, but many of the blood flukes and a few others are dioecious. Vertebrates of all classes, except the Cyclostomata, serve as definitive hosts. Those feeding on aquatic plants and animals harbor the greatest variety of digenetic trematodes, but several species occur in strictly terrestrial hosts. Effects on the vertebrate vary from no apparent harm to severe and even fatal injury. Control of flukes is by preventive measures, including interruption of their life cycles and treatment with drugs. *See* TREMATODA.

Life histories of digenetic trematodes. The life histories are complex and unlike those of other trematodes in which the egg gives rise to a single adult (**Fig. 1**). Instead, the miracidium, which develops from the egg, enters the intermediate host and becomes a germinal sac, the miracidial or mother sporocyst. Its germinal cells give rise to a second, or daughter, sporocyst generation or to one of rediae, which differ from sporocysts only in having a pharynx and saclike gut. The redia, which may or may not produce a second generation of rediae, may be present or absent in the life history. Germinal cells of the definitive sporocyst or redial generation develop into cercariae. The miracidial sporocyst rarely produces cercariae directly, and instances are known in which a redia or sporocyst may give rise to both its own kind and also to cercariae, thus maintaining the infection indefinitely in the intermediate host after the production of cercariae begins. The mode of reproduction in sporocysts and rediae has been disputed. It has been interpreted as diploid parthenogenesis in recent studies. The number of cercariae produced by the germinal sacs resulting from a single miracidium varies from a few dozen to thousands or even millions.

The cercaria usually leaves the intermediate host and then encysts, either on vegetation or after penetrating a more or less specific second intermediate, or vector, host. The encysted larva, called the metacercaria, is eaten by the vertebrate and becomes the adult fluke in that host. Cercariae of blood flukes enter the vertebrate directly by penetrating the skin. Swimmer's itch in the United States results when such larvae of avian blood flukes penetrate the human skin and die because humans are unsuitable hosts. Fish acquire certain trematodes by ingesting cercariae, which are large and attract attention. Cercariae rarely remain in the molluscan host, with or without encysting, and are eaten with the host by the vertebrate. *See* CERCARIA.

Morphology. The adult differs from other trematodes in several respects, the most obvious being the absence of a posterior adhesive organ (opisthaptor) with sclerotized hooks, plates, or multiple suctorial structures. Usually the mouth is anterior and opens into an oral sucker, but in the gasterostomes it is ventral and without a sucker (**Fig. 2**). A second sucker, the ventral one, or acetabulum, is at the posterior end of amphistomes, on the ventral surface of distomes, or absent in monostomes. The entire ventral surface may be concave and serve for adhesion, with or without the aid of a ventral sucker and glandular structures, or a portion of that surface may form a complex adhesive organ, as in the holostomes. Adult digenetic trematodes vary from about 0.2 mm (0.008 in.) to well over 1 m (3.28 ft) in length, but usually are less than 2.5 cm (1 in.) long. They occur in any part of the vertebrate that provides egress for their eggs and also in the circulatory system, from which eggs work through the tissues and escape in the feces or urine.

Classification. There is no general agreement concerning the classification of the Digenea. When few life histories were known, the group was divided into orders and families on the basis of such adult features as the number and position of suckers and components and arrangement of the reproductive system. Instances of divergent and convergent evolution with respect to those features have been revealed by life history studies which accordingly have prompted drastic revision of taxonomy of the Digenea, especially at the superfamily and order levels. The scheme of G. La Rue proposed two superorders, the Anepitheliocystidia and Epitheliocystidia, based on the presence or absence of an epithelial wall of the excretory bladder as revealed by studies on cercariae. Each superorder is divided into orders on the basis of the cercarial tail, the relationship of the excretory system to it, and the structure of the miracidium. The cercarial type is the basis of the superfamily, with that category being divided into families largely according to adult structure. In that respect, trematodes of the subclass or order Aspidobothrea

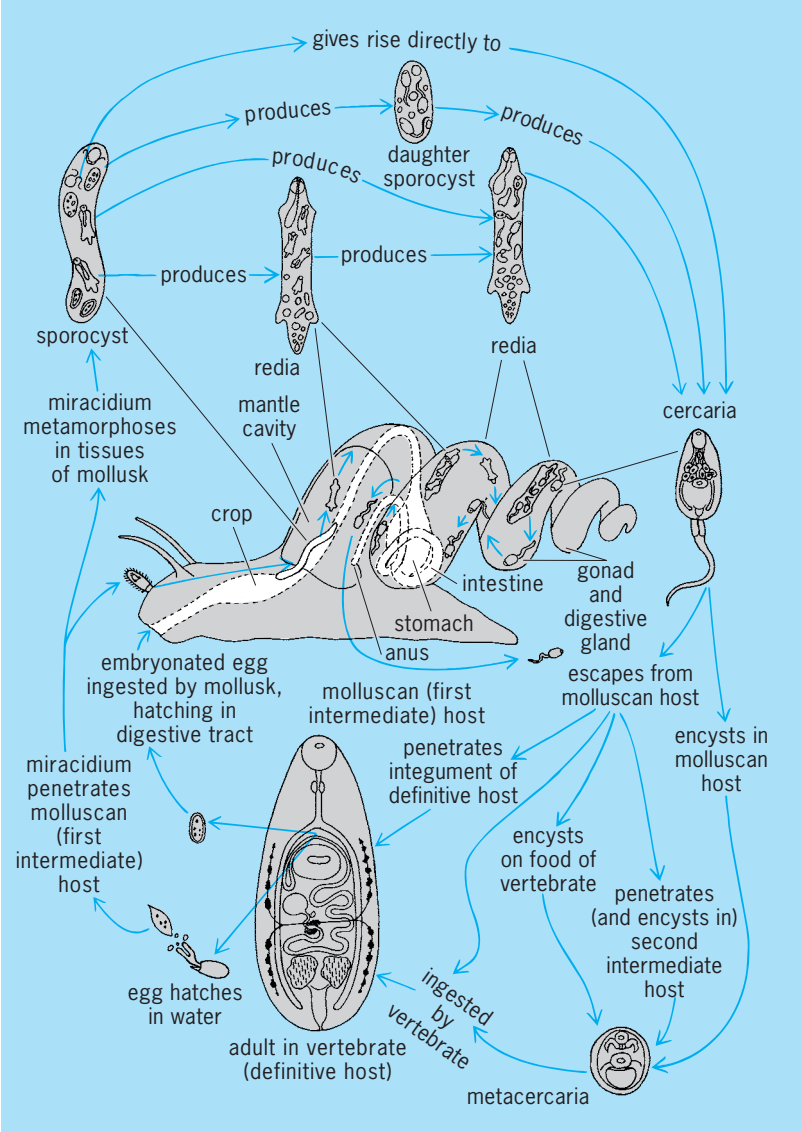

Fig. 1. Digenetic trematode cycle. (*After R. M. Cable, An Illustrated Laboratory Manual of Parasitology, Burgess, 1940*)

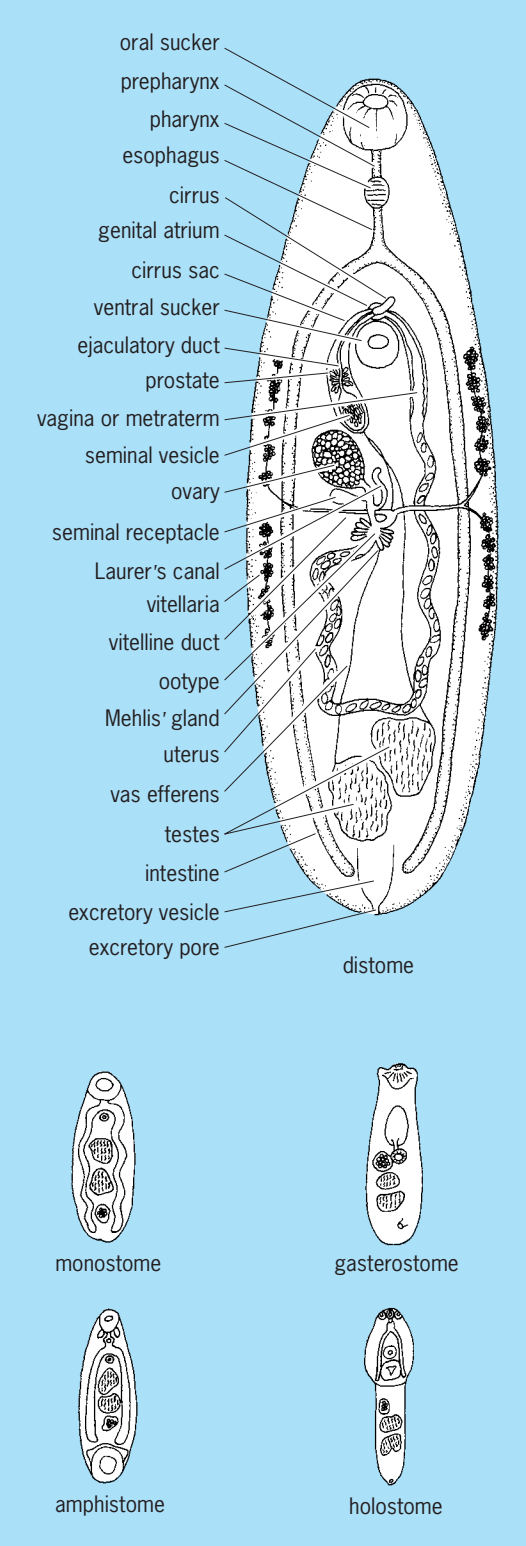

Fig. 2. Digenetic trematodes. Diagrams of various adult types. (*After R. M. Cable, An Illustrated Laboratory Manual of Parasitology, Burgess, 1940*)

(Aspidogastrea) resemble the Digenea more than the Monogenea, but so far as is known do not have germinal sacs and larval reproduction in an intermediate host. *See* ASPIDOGASTREA; PLATYHELMINTHES.

Raymond M. Cable

Digestion (invertebrate)

The process by which the complex organic molecules in foods are broken down into simpler units which can be utilized by cells for energy or incorporated into living protoplasm. Frequently, there is an essential mechanical pretreatment of the food, either before or after it is taken into the body. Mechanical treatment is especially needed in herbivorous animals that depend on the utilization of cellulose and other plant materials, or where the prey has a hard exoskeleton. Many animals have devices for crushing, grinding, or mixing.

The final processes of digestion are always chemical. Reactions in chemical digestion commonly involve hydrolysis, the accelerated addition of water to the food molecules resulting in their cleavage to simpler units. The water brings about the decomposition, and the hydrolytic digestive enzymes accelerate the process.

Enzymes. Digestive enzymes are highly specific in their action, each acting upon a particular type of molecule or even upon a particular type of bond within a molecule. Enzymes have a hydrogen ion concentration at which activity is maximal (pH optimum). The pH values in invertebrate digestive chambers generally fall in the range of pH 5 to 8, where invertebrate digestive enzymes are active. A more acid or alkaline environment is produced in special cases.

Few invertebrates have been examined thoroughly as to the number and chemical nature of enzymes involved in digestion. A complete digestive enzyme complement includes proteinases, carbohydrases, esterases including lipase, and nucleases. Proteinases hydrolyze proteins by breaking peptide bonds between amino acids. Some proteinases (enopeptidases) hydrolyze bonds within the chain of amino acids, while others (exopeptidases) act only on the bond joining the terminal amino acid to the main chain. Both categories include several enzymes specific for particular amino acids. By their combined actions, the protein is digested to individual amino acids. Similarly, a carbohydrase such as amylase hydrolyzes polysaccarides such as starch or glycogen to maltose (a disaccharide) and glucose (a monosaccharide). Various disaccharidases have been shown to be widely distributed in vertebrates. Some esterases are nonspecific, whereas lipase hydrolyzes neutral fats and oils with long-chain fatty acids. Nucleases hydrolyze DNA and RNA to nucleotides. The complement of digestive hydrolases may also include phosphatases and sulfatases.

The presence of some enzymes is related to specific diets. Phytophagous animals may have cellulase or pectinase or both for the digestion of components in the plant wall. Chitinase activity is often reported in invertebrates that ingest fungi or animals with chitinous integuments. Activities of these enzymes may often be due to enteric bacteria or protozoan symbionts. However, the production of collagenase for the digestion of vertebrate connective tissue by some maggots seems well established, and the larvae

of some intestinal parasitic nematodes may also produce it. *See* ENZYME.

Intracellular and extracellular. Digestion takes place within a chamber which is sometimes a cytoplasmic vacuole formed by endocytosis at the cell surface. This is intracellular digestion (**Fig. 1**). Single-unit-membrane bodies called lysosomes are formed in the Golgi region of the cell. They contain a variety of hydrolases operating most effectively in an acid medium. The lysosomes fuse with the food vacuoles. This process is considered to be a primitive method of digestion and is commonly found in protozoa and sponges. But it is also prevalent in the coelenterates, platyhelminthes, annelids, mollusks, and minor phyla in which immobile or mobile phagocytic cells may be involved. *See* ENDOCYTOSIS; LYSOSOME.

In most metazoans it is also possible for digestion to take place outside cells in a chamber such as a gastrovascular cavity or an alimentary canal (digestive tract). This is called luminal extracellular digestion. Species are found in most phyla that demonstrate either partially or wholly luminal extracellular digestion. Some animals demonstrate extracorporeal extracellular digestion, in which digestive enzymes are released to the outside of the animal. The digested material is taken up orally or through the surface of the organism. Even some protozoa may demonstrate this phenomenon. In metazoans demonstrating luminal extracellular digestion, it is often difficult to determine whether the digestive enzymes are exogenous (food or bacterial) or endogenous (glandular) in origin. Further, the enzymes may originate from intracellular food vacuoles at the time of egestion of indigestible residues.

Some animals have adapted to a highly specialized feeding habit and have abandoned the biosynthesis of some digestive hydrolases. Endoparasites often represent the extreme of such adaptations. Frequently a digestive tract is absent, and no evidence for the use of extracorporeal digestive enzymes can be found.

Digestive tract. The typical invertebrate digestive tract consists of an ectodermal foregut and hindgut and an endodermal midgut which is usually the major site of digestion. Particular regions of the foregut may be recognized and given special names, such as buccal cavity, pharynx, and esophagus; and the esophagus may be further differentiated into a crop or gizzard. The midgut may contain a stomach along with the intestine; and the hindgut consists of a colon or a rectum or both. It should be noted that the above regions, as designated in invertebrate digestive tracts, do not necessarily correspond in structure or function to the regions in vertebrate tracts with the same names.

Extracellular digestive enzymes are secreted by unicellular glands in the epithelium or by glandular diverticula. Mucous cells are widely distributed for lubrication. The midgut epithelium in some animals secretes a membrane between it and the contents of the gut. This peritrophic membrane is protective against abrasion and may act as a filter. The presence

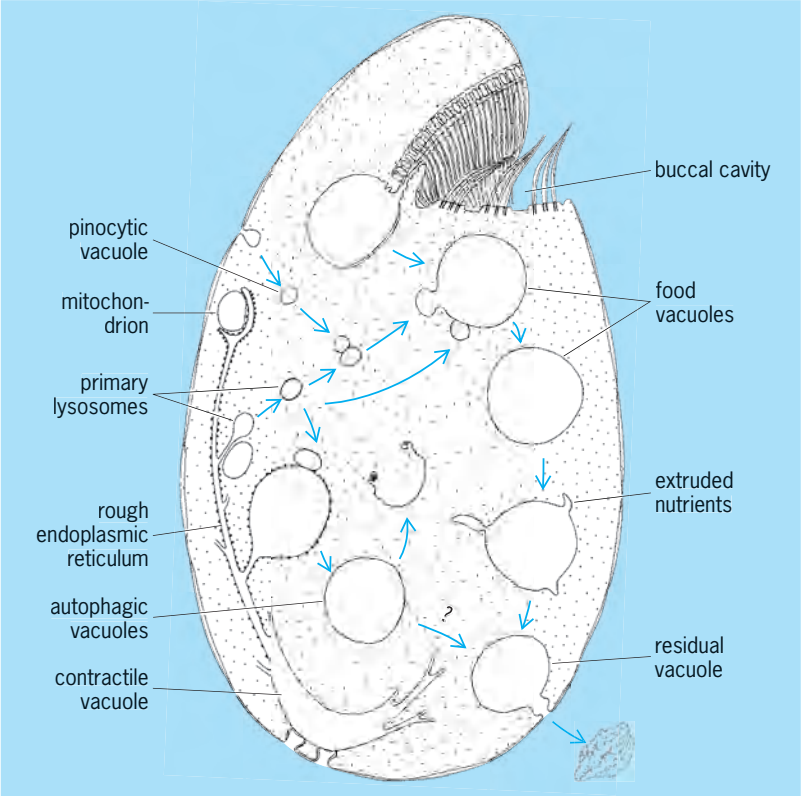

Fig. 1. Possible structural pathways of intracellular digestion in *Tetrahymena pyriformis*. (*After A. M. Elliott and G. L. Clemmons, J. Protozool., 13(2):311–323, 1966*)

of a rotating strand or rod for food movement and sorting is common in microphagous metazoans, especially ciliary feeders.

Protozoa. In most protozoa, phagocytic food vacuoles and pinocytotic vesicles are involved in food intake (Fig. 1). The organic components of the foods are digested by the activities of the acid hydrolases contributed by the coalescing lysosomes. The vacuole becomes alkaline, and bacterial enzymes as well as those in the engulfed food may also contribute to the digestive process. Many protozoan species, including free-living and parasitic forms, release hydrolases into the medium. The intracellular and extracellular enzymes may exhibit different properties, and the origin of the latter is still an open question. The properties of protozoan extracellular enzymes may differ from luminal enzymes of metazoan animals. The significance of enzyme secretion by free-living forms is unclear. In the parasitic forms, these enzymes may enhance the invasiveness of the organisms. *See* PROTOZOA.

Porifera. Digestion in sponges appears to be entirely intracellular. Most of the cells in contact with the medium are able to engulf food particles, but the flagellated collar cells (choanocytes) are the protagonists. Bacteria and other food particles captured by the choanocytes are transferred to amebocytes, where they are digested. *See* PORIFERA.

Cnidaria. In cnidarians (coelenterates), luminal digestion occurs in a gastrovascular cavity in both polyp and medusa types. Intracellular digestion occurs by phagocytic activity in gastrodermal cells.

Corals remove amino acids directly from the external medium. In sea anemones, enzymes are secreted from the gastric filaments. The saclike mass of mesenteric filaments adheres closely to the surface of the food. One of the proteinases has properties remarkably similar to vertebrate chymotrypsin and may be produced as an inactive zymogen. Tryptic activity is also present. These enzymes are also found in the sea pansy. These findings suggest that divergence of the common ancestors of these enzymes took place before the evolution of cnidarians. In sea walnuts (Ctenophora), luminal digestion begins in the pharynx, and intracellular digestion is completed in the stomach and canal system. Indigestible wastes are passed out through an anal pore. *See* CNIDARIA.

Platyhelminthes. Among turbellarian flatworms, acoels employ a syncytium of cells to ingest food. Vacuoles formed by the syncytium act as temporary gut lumina, and digestion proceeds rapidly within them. In rhabdocoels and probably alloeocoels, club-shaped gland cells secrete endopeptidases into an acid intestine prior to phagocytosis. In triclads, pharyngeal gland cells and intestinal sphere cells secrete endopeptidases prior to intracellular digestion. Digestion may be entirely luminal in polyclads. In flukes (Trematoda), proteolytic secretions produced in the glandular portion of the modified pharynx may partially digest host tissue extracorporeally before luminal digestion proceeds. There is some evidence that intracellular digestion follows. In blood feeders, gland cells in the oral sucker may secrete enzymes which cause histolysis of host tissue and initiate luminal digestion when food finally reaches the gut. Luminal enzymes may be bound to the membranes of absorptive cells. An intracellular proteinase digests globin at pH 3.9. There may also be direct uptake of nutrients through the body surface. Tapeworms (Cestoda) have no alimentary system and must use direct uptake. *See* PLATYHELMINTHES.

Rhynchocoela. The development of an anus in the digestive system of ribbon worms marks a significant departure from the digestive system of flatworms. This system is a much more efficient arrangement and one that is present in almost all higher invertebrates. In some nemerteans there is an initial extracorporeal phase of digestion, and epidermal enzymes may have a role in the uptake of nutrients from the enviroment. In others there is an initial acid-secreting mechanism in the acidophilic gland cells of the foregut which kills the prey. The intestine is richly supplied with gland cells. Digestion is initially luminal but is concluded intracellularly in phagocytic cells. *See* RHYNCHOCOELA.

Pseudocoelomates. Limited information is available concerning digestion among the pseudocoelomate animals (rotifers, gastrotrichs, kinorhynchs, gnathostomulids, entoprocts). Glandular areas of the digestive tract have been identified, and digestion is believed to be primarily if not completely luminal. The free-living adults of the hairworms (Nematomorpha) lack a digestive tract and do not feed. The parasitic juveniles absorb food directly. The spiny-headed worms (Acanthocephala) have no digestive tract and absorb nutrients directly through the body surface. The enzymes from the pharyngeal glands of roundworms (Nemata) may contribute to extracorporeal digestion or combine with additional luminal enzymes secreted from the anterior intestine. *See* ENTOPROCTA; GASTROTRICHA; GNATHOSTOMULIDA; KINORHYNCHA; ROTIFERA.

Mollusca. The molluscan alimentary canal is a highly organized ciliated surface designed for continuous processing of material. The stomach and associated diverticula have complex grooves and ridges that act as ciliary sorting tracts. Primitive gastropods and protobranchiate bivalves have salivary glands which secrete mucus that entangles the ingested food particles (**Fig. 2***a*). The anterior stomach is lined with chitin except for the ciliated sorting region which receives the ducts of the digestive glands. The posterior stomach consists of a ciliated style sac in which ciliary rotation of the ingested mucous mass produces a mucous food string drawn through the esophagus and stomach. Small food particles from the rotating mucous mass (protostyle) are sorted and conveyed to the digestive glands for intracellular digestion. Large particles are carried to the intestine for fecal pellet formation. Luminal digestion occurs in the stomach from enzymes secreted by the digestive glands. A crystalline style in a style sac completely separated from the stomach is characteristic of many bivalves and filter-feeding snails (Fig. 2*b*). The crystalline style is a long compact rod involved in chemical and mechanical digestion. The style sac secretes additions to the style base as abrasion of the protruding anterior style on the chitinous gastric shield releases style enzymes, including amylase and possibly lipase and cellulase. The presence of mucoprotein in the crystalline style precludes the extracellular digestion of proteins. A fold (typhlosole) regulates flow into and out of the main ducts of the digestive glands.

The squid shows the greatest modification of the basic pattern of the molluscan digestive tract (Fig. 2*c*). The midgut gland is divided into a large distal liver and a small proximal pancreas. A large thin-walled cecum, originating in the proximal stomach, receives the main duct of the midgut and opens into the intestine. The adaptation facilitates rapid digestion. The digestive enzymes are produced exclusively in the digestive gland. The pancreatic region secretes enzymes into the cecum, and the contractions of the cecal sac drive them into the stomach. When the contents of the stomach become liquid, the cecal valve opens, and the fluid is passed into the cecum. The hepatic portion of the digestive gland then secretes enzymes into the cecum. Digested material is absorbed in both the cecum and the intestine. *See* MOLLUSCA.

Arthropoda. Extracorporeal secretion of enzymes is a common initial phase of digestion in many arachnids, including scorpions, mites, and spiders. The epithelium of the stomach diverticula may contain zymogen cells for luminal digestion and cells for intracellular digestion in the final stages. Many arachnids survive long periods (1–2 years) without feeding.

The stomach of decapod crustaceans is the site of mechanical trituration and filtration of food particles. Enzymes pass from the digestive gland, and luminal digestion takes place in the stomach. Soluble products flow back into the digestive gland.

The chemical natures of the digestive enzyme secretions of decapod crustaceans and insects have been elucidated more than enzymes from any other group of invertebrates. Most insects have a pair of salivary glands which are involved in luminal carbohydrate digestion, but the midgut (ventriculus or stomach) is the main site of enzyme production and secretion. A proventriculus may act as a macerating gizzard (for solid food) or act as a valve in fluid feeders. Numerous symbiotic microorganisms in the midgut ceca may be major contributors to the digestive process. This is probably the case in the digestion of beeswax by a larval moth. In contrast, a silverfish that feeds on bark has been shown unequivocally to produce its own (endogenous) cellulase. Another special adaptation in insects is the high pH (10) obtained in some moths, beetles, and lice, which alters the keratin of fur, wool, and feathers for subsequent hydrolysis by the normal complement of proteinases. The stable fly and housefly have a midgut proteinase with a very low pH optimum (1.5–2.5) for an invertebrate digestive enzyme. Some predatory flies and true bugs use their salivary secretions for extracorporeal digestion. The walking worms (Onychophora) have similar extracorporeal digestion mechanisms. *See* ARTHROPODA; ONYCHOPHORA.

Annelida. In general, no specialized digestive glands are used by wandering marine polychaete worms (Annelida). The pharynx, esophagus, or anterior intestine may secrete enzymes, and coelomocytes and amebocytes may take part in intracellular digestion. In most sedentary polychaetes, luminal digestion and some absorption occur in the stomach. Polychaetes are able to accumulate amino acids, fatty acids, and glucose through the body wall directly from dilute solutions. Proteinases produced externally by stationary tubiculous polychaetes not only may act in keeping the tube free of attaching organisms but may have a role in extracorporeal breakdown of proteins.

The alimentary canals of terrestrial oligochaetes differ from those of aquatic worms. A broad range of enzymes are secreted for luminal digestion, but the origin of specific enzymes is often unknown. In bloodsucking leeches, unicellular salivary glands secrete a proteinase, hyaluronidase, and hirudin (histamine) extracorporeally. Sucked blood is stored in large crop diverticula for up to 200 days, and the mechanism of luminal digestion is still unclear. In contrast, digestion is rapid in predatory species. *See* ANNELIDA.

Sipuncula, Echiurida, and Priapulida. In peanut worms (Sipuncula) the descending coil of a long intestine is the apparent source of enzymes, but luminal digestion predominates in the loop area. In echiurids the siphonal intestine contains a digestive fluid with a complement of hydrolases. They have also been shown to accumulate amino acids directly from dilute solutions. In priapulids the intestinal fluid is alkaline (pH 8.3) and contains tryptic and chymotryptic activities. *See* ECHIURIDA; PRIAPULIDA; SIPUNCULA.

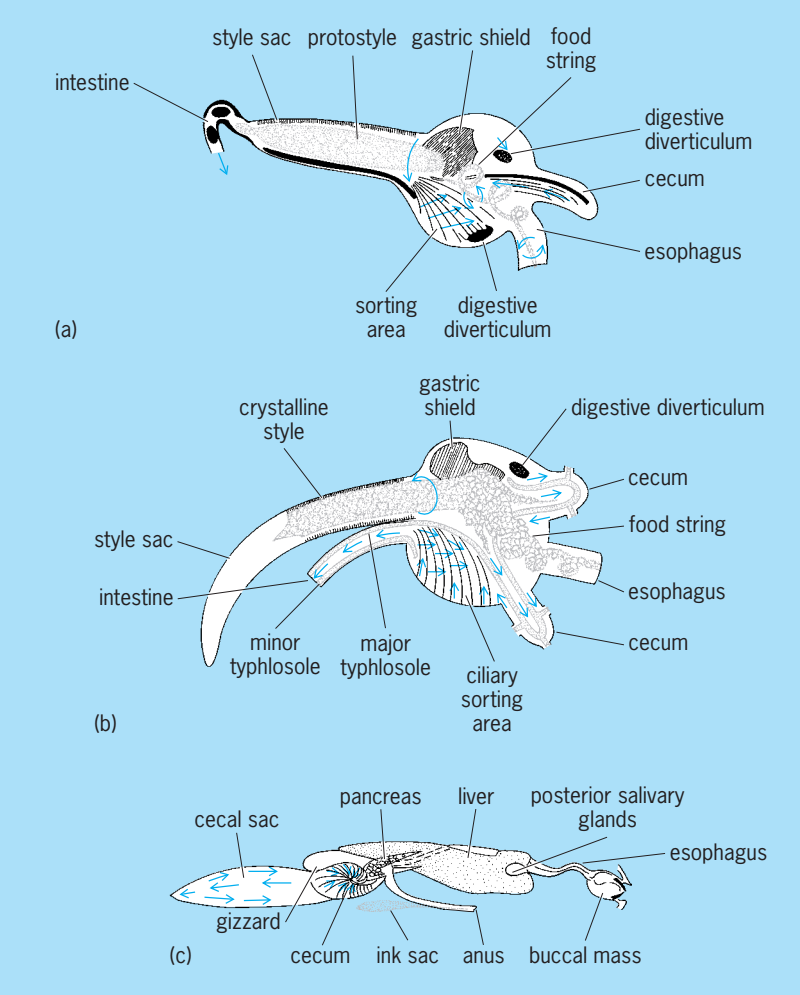

Fig. 2. Molluscan digestive systems. (*a*) Idealized form of the stomach in an early prosobranch. (*b*) A eulamellibranch stomach. (*c*) *Loligo* alimentary canal. (*After J. E. Morton, Molluscs, 4th ed., Hutchinson, 1967*)

Ectoprocta and Phoronida. Limited information indicates most digestion is extracellular in ectoprocts. Fat is digested intracellularly in some marine species. Digestion by brachiopods is chiefly intracellular within the digestive gland, but carbohydrate digestion may be extracellular. Phoronid digestion is probably intracellular within the stomach. *See* PHORONIDA.

Echinodermata. Among the Echinodermata, little is known about digestion in the crinoids and in basket and brittle stars. Coelomocytes appear to play a major role in sea cucumbers; they may invade the gut lumen to release enzymes. The esophagus and intestine are probably the digestive sites in sea urchins, and the cecum may secrete the enzymes. Coelomocytes migrate to the external surface of the urchin to predigest a variety of materials which are then ingested. The pyloric ceca of the starfish secrete enzymes in the cardiac stomach, where luminal digestion occurs. In some, the stomach may be

everted so that digestion begins outside the animal. Endopeptidases from the ceca, such as trypsin and carboxypeptidase, have been comprehensively studied. *See* ECHINODERMATA.

Chordata. Digestion is probably extracellular in sea squirts (Urochordata) and within the stomach, which contains abundant secretory cells. In the lancelet (Cephalochordata), mechanical digestion is similar to that seen in many mollusks. The mixing and sorting take place in the midgut by ciliary action. Small particles from the rotating mass are digested intracellularly in a blind digestive cecum. *See* CHORDATA; FEEDING MECHANISMS (INVERTEBRATE).

Edward J. Devillez

Bibliography. E. J. Barrington, *Invertebrate Structure and Function*, 2d ed., 1979; M. Florkin and B. Scheer, *Chemical Zoology*, vols. 1–11, 1967–1979; S. N. Prasad, *Life of Invertebrates*, 1980; C. L. Prosser, *Comparative Animal Physiology*, 4th ed., 1991.

Digestive system

The vertebrate digestive system consists of the digestive tract and ancillary organs that serve for the acquisition of food and assimilation of nutrients required for energy, growth, maintenance, and reproduction. Food is ingested, reduced to particles, mixed with digestive fluids and enzymes, and propelled through the digestive tract. Enzymes produced by the host animal and microbes indigenous to the digestive tract destroy harmful agents and convert food into a limited number of nutrients, which are selectively absorbed. The digestive systems of vertebrates show numerous structural and functional adaptations to their diet, habitat, and other characteristics. Carnivores, which feed exclusively on other animals, and species that feed on plant concentrates (seeds, fruit, nectar, pollen) tend to have the shortest and simplest digestive tract. The digestive tract tends to be more complex in omnivores, which feed on both plants and animals, and most complex in herbivores, which feed principally on the fibrous portions of plants.

Gut structure and function can also vary with the habitat and other physiological characteristics of a species. The digestive tract of fish has adaptations for a marine or fresh-water environment. The hindgut of terrestrial vertebrates plays an important role in the conservation of the fluids, electrolytes, and nitrogen in the digestive secretions, and in the urinary excretions of reptiles and birds. Fish, amphibians, and reptiles are ectotherms, whose body temperature, metabolic rate, and nutritional requirements vary with their ambient temperature. The endothermic birds and mammals require a higher and more continuous input of nutrients for the maintenance of a constant body temperature. The basal metabolic rate per gram of body weight increases with a decrease in the body mass. Therefore, small animals must process larger amounts of food per gram of body weight, thus limiting their maximum gut capacity and digesta retention time.

Anatomy. Because of wide species variations, the digestive system of vertebrates is best described in terms of the headgut, foregut, midgut, pancreas, biliary system, and hindgut.

Headgut. The headgut consists of the mouthparts and pharynx, which serve for the procurement and the initial preparation and swallowing of food. Movable, articulated jaws are found in all vertebrates except the cyclostomes (lampreys and hagfish), which are primitive, parasitic fish. Teeth, used for grasping, cutting, or tearing of food, are located in the jaws or other mouthparts of most fish. However, in some species, such as the chub, they are located in the pharynx and used as a grinding mill (**Fig. 1**), while the mouthparts of basking sharks and paddlefish are used for the filtering or sorting of food. Complex mechanisms for the microfiltration of ingested food are seen also in some species of larval amphibians, as well as flamingos and baleen whales. Some larval amphibians have horny teeth that are used for procuring prey or grazing on plants, but all adult amphibians are carnivores with a weak dentition that is used only for grasping and positioning prey. Most reptiles have teeth that are used for the procurement, cutting, tearing, or crushing of food, but turtles and birds use a beak for these purposes.

A few mammals have no teeth. The ant and termite eaters in five mammalian orders have simple teeth and weak jaws. However, the teeth of most mammals consist of incisors and canines for grasping, cutting, and tearing, and premolars and molars with uneven occluding surfaces. The articulation and musculature of the mammalian jaws allows both the vertical movement of the lower jaw and either its lateral movement, as seen in most species, or the forward and backward movement, as seen in rodents and elephants. Thus, mammals have the unique ability to use their premolar and molar teeth for the grinding as well as the crushing of food. A movable tongue aids in the procurement of food by frogs, toads, woodpeckers, and mammalian anteaters, and for placement of food between the molars of mammals. Oral glands secrete adhesive materials in frogs, toads, woodpeckers, and mammalian anteaters, toxins in some reptiles, and fluids that aid in the deglutition or swallowing of food in reptiles, birds, and mammals. *See* ORAL GLANDS; TONGUE.

Foregut. The foregut consists of an esophagus for the swallowing of food and, in most species, a stomach that serves for its storage and initial stages of digestion. The esophagus of most vertebrates is lined with a multilayer of cells that are impermeable to absorption. In most birds it contains the crop, an outpocketing of its wall that provides for the temporary storage of food. A stomach is present in all but the cyclostomes and some species of advanced fish (Fig. 1) and in the larval amphibians. In most vertebrates it consists of a dilated segment of the gut that is separated from the esophagus and midgut by muscular sphincters or valves. This is often referred to as a simple stomach. However, in birds these functions are carried out by the crop (storage), proventriculus

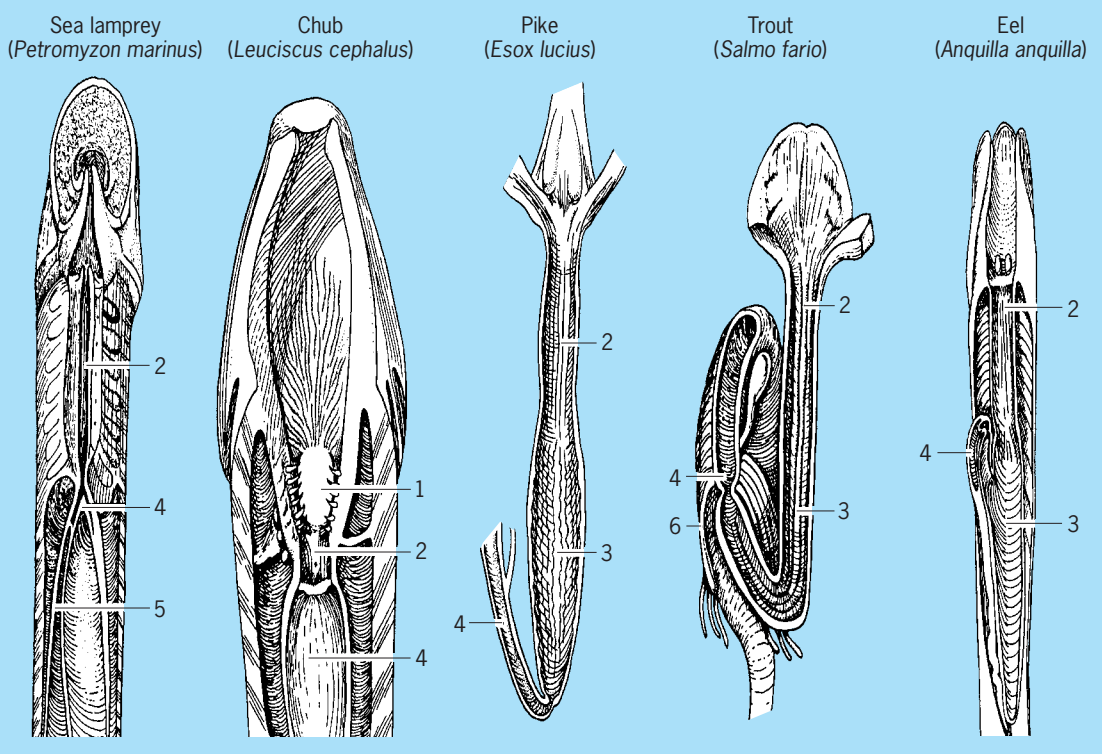

Fig. 1. Major variations in the foregut and midgut of fish. The teeth of most fish are located on the jaws or other mouthparts, but their location in the pharynx (1) of some species such as the chub allows mastication of food without the loss of fine particles through the gills. The esophagus (2) varies in length. The stomach (3) is absent in cyclostomes, such as the lamprey, and in some advanced species, such as the chub. Where present, it may be straight as in the pike, U-shaped like that of the trout, or Y-shaped with a gastric cecum, as seen in the eel. The intestine (4) tends to be relatively short with little distinction between the midgut and hindgut in most fish, but the lumen surface area and digesta retention time of the midgut are increased by a spiral valve (5) or pyloric ceca (6) in some species. (*Modified from W. Harder, Anatomy of Fishes, Part II, Schweizerbart'sche Verlagsbuchhandlung, 1975*)

(secretion), and gizzard (grinding or mastication). *See* ESOPHAGUS; STOMACH.

In most vertebrates, a major portion of the stomach is lined with a proper gastric mucosa (epithelium), which secretes mucus, hydrochloric acid (HCl), and pepsinogen. The distal (pyloric) part of the stomach secretes mucus and bicarbonate ions (HCO^-_3), and its muscular contractions help reduce the size of food particles and transfer partially digested food into the midgut. The stomach of reptiles and most mammals has an additional area of cardiac mucosa near its entrance, which also secretes mucus and bicarbonate ions.

The stomach of some mammals differs from that of other vertebrates with respect to its epithelial lining, structure, or size (**Fig. 2**). In some of the species belonging to 10 of the 20 mammalian orders, the cardiac mucosa is preceded by a region of stratified squamous epithelium, similar to that of the esophagus. This can vary from a narrow region, as seen in the pig, to one-half of the stomach, like that of the horse and rat, to a major fraction of the stomach, as seen in the ox and llama, to the entire stomach of monotremes (spiny anteaters and platypuses). The presence of this region in the stomach of anteaters and many herbivores suggests that it serves as protection against physical damage. Eight mammalian orders include species with an enlarged stomach, which may be divided into permanent compartments, as seen in the rat, ox, and llama.

Midgut. The midgut or small intestine is the principal site for the digestion of food and the absorption of nutrients. It is lined with a single layer of cells that secrete mucus and fluids, contain enzymes that aid in the final stages of carbohydrates and protein digestion, and absorb nutrients from the lumen into the circulatory system. The surface area of the lumen can be increased by a variety of means.

The lumen surface of the midgut of fish is increased by folds that form a spiral valve in some species and by 1–1000 pyloric ceca (blind sacs) in many other species (Fig. 1). In higher vertebrates the lumen surface is increased by the presence of villi, which are macroscopic projections of the epithelial and subepithelial tissue (**Fig. 3***a*). A crypt at the base of each villus contains cells that secrete fluid into the gut and hormones that help regulate the functions of the digestive system. The crypts also contain cells that serve as precursors for the generation of absorptive and mucus-producing cells, which migrate up the villus surface and are eventually sloughed off into the lumen.

The lumen surface is also expanded by a brush border of microvilli on the lumen-facing (apical) surface of the midgut absorptive cells in all vertebrates (Fig. 3*b*). The brush border membranes contain

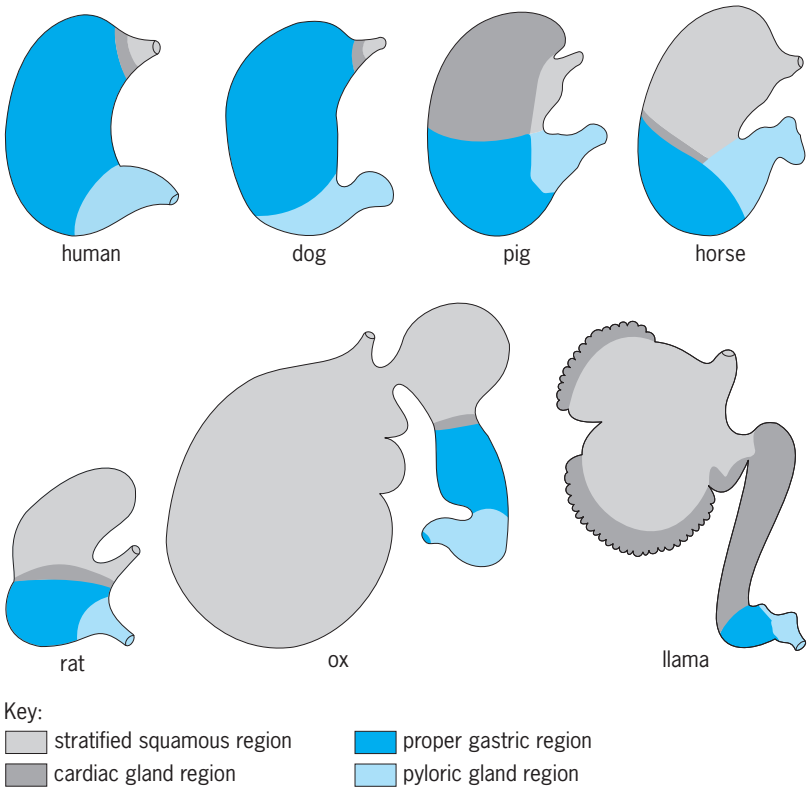

Fig. 2. Variations in the distribution of gastric epithelium and the degree of compartmentalization in the stomach of mammals. Stomachs are not drawn to the same scale; the stomach contents of the ox are approximately 70 times the volume and 14 times the weight of human stomach contents. (*After C. E. Stevens and I. D. Hume, Comparative Physiology of the Vertebrate Digestive System, 2d ed., Cambridge University Press, 1995*)

enzymes that aid in the final digestion of food and mechanisms that provide for the selective absorption of nutrients. The lumenal surface area of the human small intestine is increased 10-fold by the presence of villi and an additional 20-fold by the microvilli, resulting in a total surface area of 310,000 in.^{2} (2,000,000 cm^2), which is larger than a doubles tennis court.

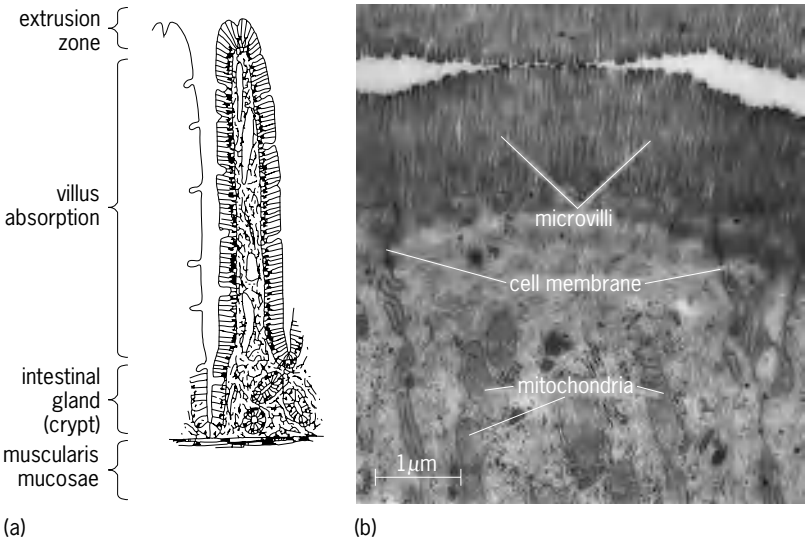

Fig. 3. Section of the wall of the small intestine. (*a*) Intestinal villus and crypt (*after R. A. Argenzio et al., Villous atrophy, crypt hyperplasia, cellular infiltration, and impaired glucose-Na absorption in enteric cryptosporidiosis in pigs, Gastroenterology, 98:1129–1140, 1990*). (*b*) Electron micrograph of microvilli on the free surface of an absorbing cell from the small intestine of a rat (*courtesy of M. L. Watson*).

Pancreas and biliary system. Digestion in the midgut is aided by secretions of digestive enzymes and fluid by pancreatic tissue, and secretion of bile by the liver. Pancreatic tissue is represented by primitive ceca along the midgut of cyclostomes and is distributed along the intestinal wall, and even into the liver, of some species in the more advanced classes of fish. However, the pancreas is a compact organ in sharks, skates, rays, many teleosts, and all other vertebrates. *See* PANCREAS.

The liver is a compact organ in all vertebrates. One of its many functions is the secretion of bile. In most vertebrates, the bile is stored in the gallbladder and released into the intestine as needed, but a gallbladder is absent in some species of fish and mammals. Bile salts serve to emulsify lipids and increase their surface area available for digestion by the water-soluble lipase. *See* GALLBLADDER; LIVER.

Hindgut. The hindgut is the final site of digestion and absorption prior to defecation or evacuation of waste products. The hindgut of fish, amphibian larvae, and a few mammals is short and difficult to distinguish from the midgut by either its diameter or epithelial lining. However, the hindgut of adult amphibians and reptiles, birds, and most mammals is a distinct segment, which is separated from the midgut by a muscular sphincter or valve. It also tends to be larger in diameter. Thus, the midgut and hindgut of these animals are often referred to as the small intestine and the large intestine. *See* INTESTINE.

The hindgut of some reptiles and many mammals includes a blind sac or cecum near its junction with the midgut. A pair of ceca are present in the hindgut of many birds and a few mammalian species. The remainder of the hindgut consists of the colon and a short, straight, terminal segment, which is called the rectum in mammals. The digestive and urinary tracts exit separately from the body of most species of fish and mammals. However, in adult amphibians and the reptiles, birds, and some mammals, this segment terminates in a chamber called the cloaca, which also serves as an exit for the urinary and reproductive systems. The hindgut or, where present, the cloaca terminates in the anus. *See* COLON; URINARY SYSTEM.

The hindgut is similarly lined with a single layer of absorptive and mucus-secreting cells that originate in crypts. However, it lacks villi, and (with the exception of the cecum of birds) its absorptive cells lack digestive enzymes and the ability to absorb most nutrients. One major function of the hindgut is to reabsorb the fluids secreted into the upper digestive tract and (in animals that have a cloaca) excreted in the urine. It also serves as the principal site for the microbial production of nutrients in the herbivorous reptiles and birds and in most herbivorous mammals. Thus, the hindgut tends to be longest in animals that need to conserve water in an arid environment, and has a larger capacity in most herbivores.

Musculature. The digestion of food, absorption of nutrients, and excretion of waste products require

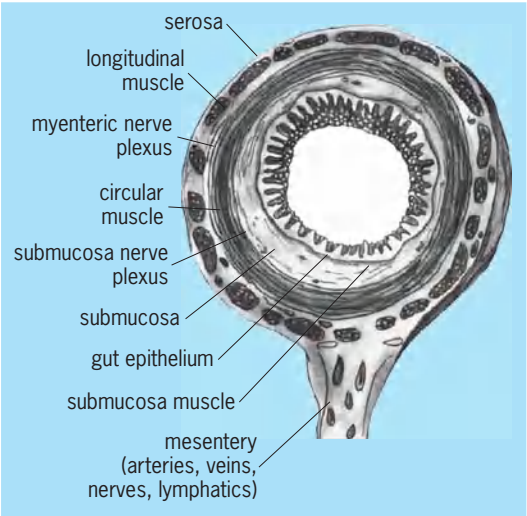

Fig. 4. Cross section of the layers of the gastrointestinal tract and the mesentery, which supplies it with blood vessels, lymphatic vessels, and nerves.

the mixing of ingesta with digestive enzymes and the transit of ingesta and digesta through the digestive tract. The digestive tract of cyclostomes is enveloped with a single layer of muscle, and digesta transit is aided by cilia in the epithelial lining of the tract. However, in other vertebrates the contents are mixed and moved by an inner layer of circular muscle and an outer layer of muscle that runs longitudinally along the tract (**Fig. 4**).

The initial act of deglutition and the final act by which waste products are defecated from the digestive tract are effected by striated muscle. This type of muscle is characterized by rapid contraction and is controlled by extrinsic nerves. The esophagus of fish and varying lengths of the mammalian esophagus are enveloped with striated muscle. However, the esophagus of amphibians, reptiles, and birds, and the entire gastrointestinal tract of all vertebrates are enveloped by smooth muscle. This smooth muscle contracts more slowly, and its rate of contraction is partly independent of external stimulation. *See* MUSCLE.

Contraction of the circular muscles aids in the mixing and propulsion of digesta. At some sites along the digestive tract, the circular muscle is thicker and under almost continuous contraction, forming sphincters or valves that delay the movement of lumen contents in either direction. The anus has an additional sphincter, which is composed of striated muscle.

Contractions of the longitudinal muscles shorten the digestive tract. In the large intestine of some

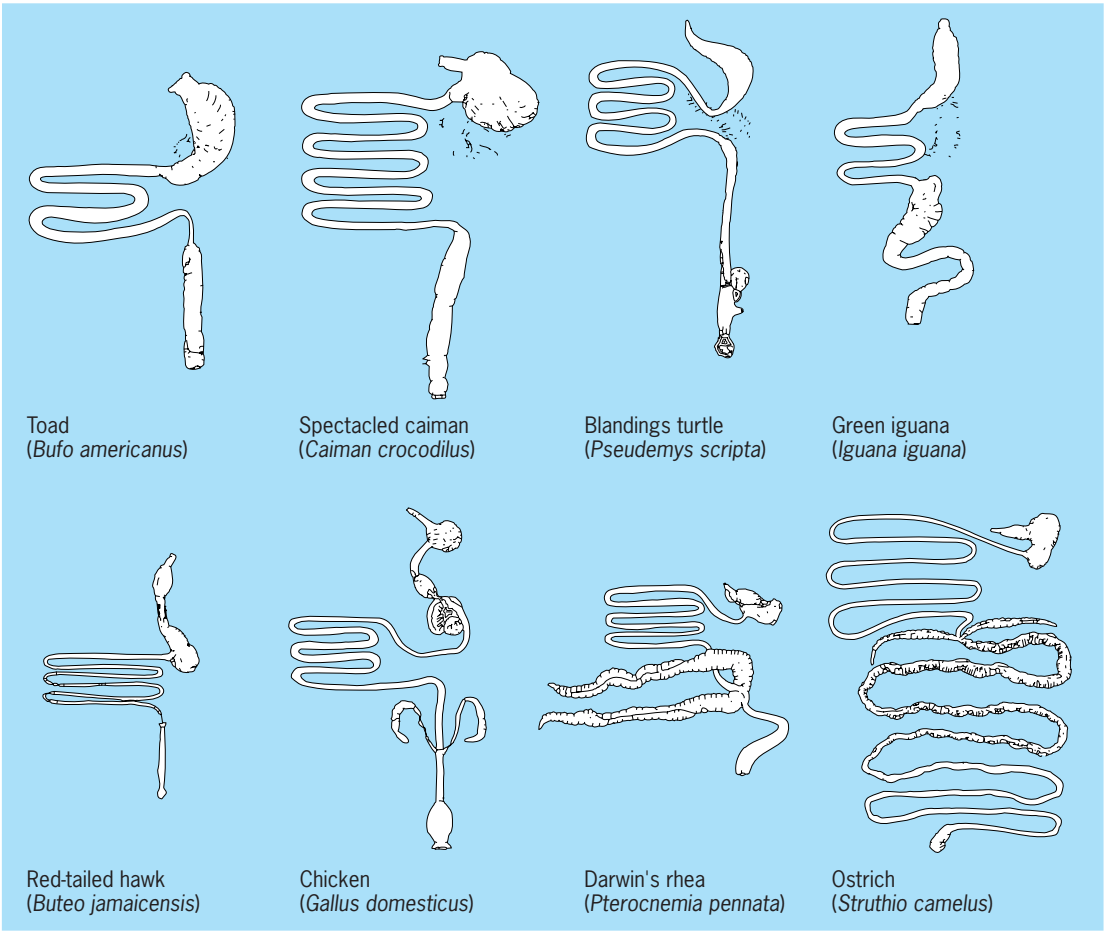

Fig. 5. Gastrointestinal tracts of amphibians, reptiles, and birds. (*After C. E. Stevens and I. D. Hume, Comparative Physiology of the Vertebrate Digestive System, 2d ed., Cambridge University Press, 1995*)

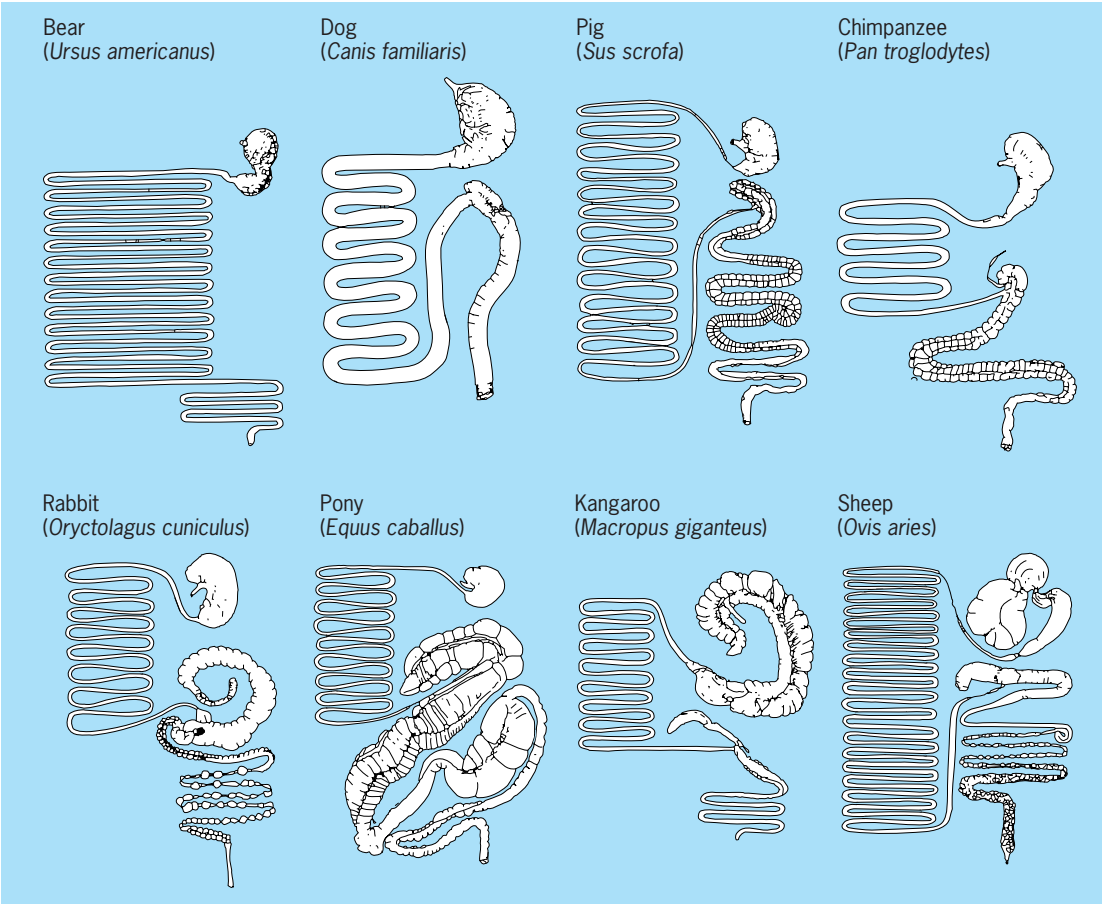

Fig. 6. Gastrointestinal tracts of carnivorous, omnivorous, and herbivorous mammals. (*Bear and chimpanzee after C. E. Stevens and I. D. Hume, Comparative Physiology of the Vertebrate Digestive System, 2d ed., Cambridge University Press, 1995; pig after R. A. Argenzio and M. Southworth, Sites of organic acid production and absorption in the gastrointestinal tract of the pig, Amer. J. Physiol., 118:454–460, 1974; dog, rabbit, pony, kangaroo, and sheep after C. E. Stevens, Comparative physiology of the digestive system, in Dukes' Physiology of Domestic Animals, ed. by M. J. Swenson, 9th ed., pp. 216–232, Cornell University Press*)

mammals and the stomach of a few, the longitudinal muscle is concentrated into three bands (taenia). Contraction of these longitudinal bands pulls these segments of the gut into a series of saccules. These saccules change their position along the gut with time and serve to both mix the contents and propel them through the gut.

Nerve and endocrine tissue. The initial act of deglutition and final act of defecation are under the voluntary control of the central nervous system. However, the remainder of the digestive system is subject to the involuntary control of nerves which release a variety of neurotransmitting or neuromodulating agents that either stimulate or inhibit muscular contractions and the secretions of glands and cells. The motor and secretory activities of the digestive system are also under the control of a wide range of other substances produced by endocrine cells that are released either distant from (hormones) or adjacent to (paracrine agents) their site of action. Although there are some major variations in the complement and activities of the neurotransmitters, neuromodulators, hormones, and paracrine agents, their basic patterns of control are similar.

Variations. The gastrointestinal tracts of adult amphibians and reptiles consist of a simple stomach, a midgut, and a hindgut that is generally expanded in diameter and referred to as the large intestine (**Fig. 5**). The hindgut is relatively short in adult amphibians, such as the toad, and in reptilian carnivores, such as the caiman. It tends to be longer in omnivorous reptiles, such as Blanding's turtle, and the proximal hindgut of most herbivorous reptiles, such as the green iguana, is expanded and includes a cecum at its juncture with the midgut. The gastrointestinal tract of most birds includes a crop, proventriculus, gizzard, paired ceca, and a short straight section of hindgut that is generally referred to as the rectum. However, ceca are absent in some birds, small in others, such as the hawk, and more developed in omnivores, such as chickens. The largest ceca are found in most of the herbivores, such as grouse and rheas. The long colon of the herbivorous ostrich is unique among avian species.

Mammals. The stomach and hindgut of mammals show an even wider range of structural variation (**Fig. 6**). Most mammals have a simple stomach, but the forestomach of many herbivorous mammals is large and divided into numerous compartments or

saccules, as seen in the sheep and kangaroo. The same is true for other advanced ruminants (bovids, goats, antelope, deer, elk, giraffe), camelids, hippopotamuses, as well as the sloths, colobid monkeys, and cetaceans (dolphins, porpoise, and whales), hyrax, and some herbivorous rodents. Although present-day cetaceans feed on invertebrates, fish, or small mammals, the multicompartmental stomach is believed to have been inherited from herbivorous ancestors. The stomach is also expanded in blood-feeding vampire bats, reaching twice the body length of some species.

The mammalian hindgut shows a similar range of variation. It is very short and indistinct in some insectivores, cetaceans, and the omnivorous bears, and relatively simple in most carnivores, such as the dog. The hindgut of omnivores tends to be more highly developed, and includes a sacculated cecum and colon in some omnivores, such as the chimpanzee and domesticated pig. The hindgut of mammalian herbivores with a simple stomach is dominated by either a large, sacculated cecum, as seen in the rabbit, or a large sacculated colon, such as that of the pony. However, regardless of diet, the hindgut tends to be short and simple in aquatic and semiaquatic species, such as cetaceans and hippos, and longest in desert species, such as camels.

Human. The anatomy of the human digestive system is similar to that of other mammalian omnivores (**Fig. 7**). The teeth and salivary glands are those of a mammalian omnivore, and the initial two-thirds of the esophagus is enveloped by striated muscle. A simple stomach is followed by an intestine, whose length consists of approximately two-thirds small bowel and one-third large bowel. The structures of the pancreas and biliary system show no major differences from those of other mammals. During early fetal development, a distinct, conical cecum is present and continues to grow until the sixth month of gestation.

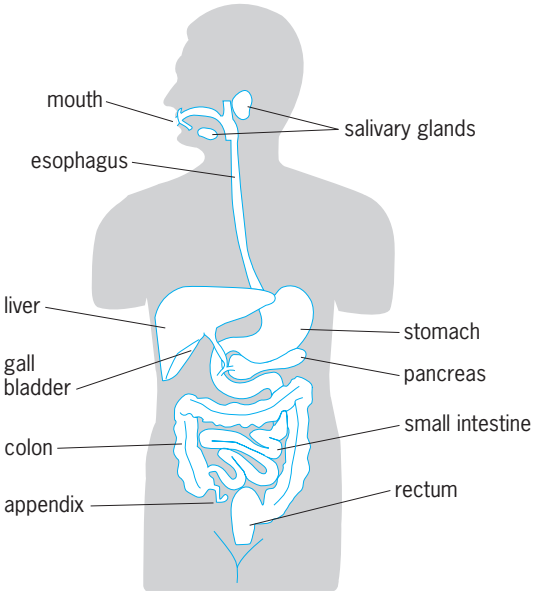

Fig. 7. Human digestive system. (*After A. C. Guyton, Textbook of Medical Physiology, 8th ed., Saunders, 1991*)

However, unlike other primates, the cecum recedes to become little more than a bulge in the proximal colon by the time of birth. The colon continues to lengthen after the birth and is sacculated throughout its length like that of the apes and a few monkeys but few other mammals.

Physiology

The major physiological activities of the digestive system are motility, secretion, digestion, and absorption. Each activity can be affected by diet and, in the cold-blooded species, is reduced with a decrease in body temperature.

Motility. The mastication of food and the movement of ingesta and digesta through the digestive tract are controlled by the motor activity of muscular contractions. Pressure of food against the palate and back of the mouth stimulates a nerve reflex that passes through a deglutition center in the brain. This reflex closes the entrance into the respiratory system and stops respiration, to prevent the inspiration of food into the lungs, and initiates muscular contractions that pass food into the esophagus. The food (bolus) is then passed down the esophagus and into the stomach by a moving wave of muscular contractions (peristalsis) accompanied by inhibition of the esophageal sphincters.

The multicompartmental forestomach of ruminants undergoes a continuous series of complex, repetitive contractions that are controlled by the central nervous system. However, the gastric motility of most species and the intestinal motility of all vertebrates are controlled partially by the intrinsic characteristics of their smooth muscle cells. The result is production of either stationary (mixing) contractions of the stomach and intestine or a series of peristaltic contractions that carry digesta on through the tract.

Gastric peristalsis is initiated by cells in the middle of the stomach, which act as pacemakers for the contractions that travel over the distal half of the stomach. Pacemakers in the midgut and hindgut can also initiate waves of peristaltic contraction. However, pacemakers located near the cloaca of reptiles and birds or the midcolon of most mammals generate periodic waves of antiperistaltic contraction, which reflux digesta along the colon and into the cecum. These antiperistaltic waves increase the retention time needed for the absorption of electrolytes and water, and the multiplication of hindgut bacteria. Peristaltic and segmental contractions eventually move the digesta into the terminal hindgut, where it initiates the relaxation of sphincters and defecation.

Digestion. The diet of vertebrates contains carbohydrates, lipids, and protein, which provide the energy and building blocks for growth, maintenance, and reproduction, and the vitamins that are required in small amounts for the regulation of metabolic processes. The lipids, proteins, and most of the carbohydrates require digestion for the release of nutrients that can be absorbed. Digestion is accomplished by enzymes produced by the digestive

Major dietary substrate end products and endogenous digestive enzymes*

Substrate	Extracellular enzymes	Intestinal mucosal enzymes	End products
Amylose Amylopectin Glycogen	α Amylase	Maltase Isomaltase	Glucose
Chitin	Chitinase	Chitobiase	Glucosamine
Sucrose		Sucrase	Glucosamine
Lactose		Lactase	Glucose Galactose
Trehalose		Trehalase	Glucose
Triglycerides	Lipase Colipase		β-Monoglyceride Fatty acids
Waxes	Lipase Esterase		Monohydric alcohol Fatty acids
Protein	Pepsin Trypsin Chymotrypsin Elastases Carboxypeptidase A Carboxypeptidase B	Aminopeptidase Tripeptidase Dipeptidase	Amino acids

*Modified from C. E. Stevens and I. D. Hume, *Comparative Physiology of the Vertebrate Digestive System,* 2d ed., Cambridge University Press, 1995.

system (endogenous enzymes) or by bacteria that are normal residents of the digestive tract.

Dietary substrates. The carbohydrates consist of monosaccharides, disaccharides, and polysaccharides. The major monosaccharides in plants and animals are glucose or fructose, and the major disaccharides are the transport carbohydrates: sucrose (glucose-fructose), lactose (glucose-galactose), and trehalose (glucose-glucose). Sucrose is the principal transport carbohydrate in plants, trehalose is the blood sugar of insects, and lactose is the milk sugar of most mammals. Polysaccharides serve as either storage or structural carbohydrates. The main storage carbohydrates are the starches amylose and amylopectin in plants and glycogen in animals. The principal structural carbohydrates are the cellulose, hemicellulose, and pectin in the cell walls of plants, and chitin, which reinforces the exoskeleton cell walls of insects and many marine invertebrates. Starch and cellulose consist of long chains of glucose molecules. Hemicellulose, pectin, and chitin are made up of a variety of monosaccharides.

The major dietary lipids (fats and oils) are triglycerides and waxes that consist of esters of alcohols and fatty acids. Triglycerides are the major form of lipid storage in most animals and the seeds of plants. Wax esters constitute most of the lipids stored in many of the invertebrates in the food chain of marine fish and mammals. Protein consists of long chains of amino acids, arranged in various combinations and permutations that must be broken down to small peptides or amino acids for intestinal absorption. The cells of plants, animals, and bacteria also contain deoxyribonucleic acid (DNA), which provides genetic information, and ribonucleic acid (RNA), which controls the activities of the cells. The vitamins required by vertebrates consist of a group of fat-soluble (A, D, E, and K) and water-soluble (B-complex and C) organic compounds.

Endogenous enzymes. Plant and animal starches are converted to oligosaccharides (short-chain structures) and disaccharides by amylase, which is secreted by the salivary glands of some species and the pancreas of all vertebrates (see **table**). The end products of starch digestion, plus the dietary disaccharides, are converted to monosaccharides by enzymes in the brush border of the absorptive epithelial cells lining the small intestine. Chitin can be digested by the enzyme chitinase, which has been demonstrated in the gastric mucosa and pancreas of many mammals, birds, reptiles, and adult amphibians. However, vertebrates do not produce enzymes capable of digesting the structural polysaccharides of plants.

Lipids are digested into alcohols, monoglycerides, and fatty acids by lipases and esterases, which are secreted predominantly by the pancreas. However, the lipases are water-soluble enzymes that can attack their substrate only at a lipid-water interface. Therefore, the lipids must be emulsified in order to provide the surface area required for efficient digestion. Emulsification is accomplished by the release of bile salts secreted by the liver and released into the midgut. The bile salts help form micelles, small spherical structures that contain the end products of lipid digestion and aid in their transfer to the lumen membrane of the epithelial cells for absorption into the cell. Once this process is complete, the bile salts are released and passed on to the distal midgut, where they are absorbed and circulated back to the liver.

Dietary protein is first broken down into long chains of amino acids (polypeptides) by gastric pepsin and pancreatic trypsin. The polypeptides are then attacked by other pancreatic proteases (chymotrypsin, carboxypeptidase, elastase) to form tripeptides, dipeptides, and amino acids. All of these enzymes are secreted in an inactive form to prevent the self-digestion of the secretory cells prior to their release. Pepsin is activated by the acidity resulting from the secretion of hydrochloric acid (HCl) into the stomach, and trypsin is activated by an enzyme (enterokinase) that is secreted by intestinal epithelium. The trypsin initially released by enterokinase activates the release of additional trypsin and the other pancreatic, proteolytic enzymes. Tri- and dipeptides are digested into amino acids by enzymes in the brush border and contents of midgut absorptive cells. Nucleic acids are digested by pancreatic ribonucleases into pentose sugars, purines, and pyrimidines.

Some vitamins must be released from their combination with other substances by digestive enzymes before they can be absorbed. Vitamin B_{12}, which is especially subject to destruction by digestive enzymes, is released from other compounds and combined with a substance called intrinsic factor secreted by the stomach. This combination protects it from digestion until it is released for absorption in the distal midgut. Therefore, a B_{12} deficiency can result from dysfunctions of gastric secretion in humans.

The levels of various digestive enzymes tend to correlate with the levels of carbohydrate, lipid, and protein in the normal diet of a species. Herbivores and omnivores tend to produce higher levels of amylase, carnivores tend to produce higher levels of proteases, and the high concentrations of ribonucleases found in ruminants and kangaroos correlate with the high levels of microbial nucleic acid produced in their forestomach. The composition of digestive enzymes also varies with time in neonate mammals. All neonate mammals feed on milk, which is high in lipids and lactose in most species, and after weaning their diet that is usually low in fat and higher in other carbohydrates. The production of digestive enzymes adjusts to these changes in the diet. For example, the newborn of most species produce high levels of lactase but low levels of the other carbohydrases. However, the latter appear well before weaning and reach their full complement at the time of birth, with a subsequent decrease in lactase production. These changes in enzyme composition anticipate the changes in diet and appear to be genetically programmed or "hard-wired" into the system. The levels of pancreatic amylase, lipase, and proteases (brush border enzymes) can also increase following an increase in the carbohydrate, lipid, or protein in the diet of adult animals.

Indigenous bacteria. Soon after birth or hatching the lumen surface and contents of the gastrointestinal tract become colonized by bacteria. Substantial numbers of bacteria can be found in all segments of the tract, but the highest numbers are present in those segments in which digesta are retained for prolonged periods of time at a relatively neutral pH. Bacteria are found in concentrations of 10^4 to 10^6 microorganisms per gram of fluid digesta in the midgut of fish and higher vertebrates. However, the development of the hindgut of terrestrial vertebrates into a major site of digesta retention was accompanied by its colonization with much higher numbers (10^{10} to 10^{11} per gram). In the ruminants and a few other species of mammalian herbivore, a large stomach serves as an additional site for colonization with similar numbers of bacteria.

Indigenous bacteria help protect the animal from pathogenic microorganisms by stimulating immunity and competing for substrates. They also convert dietary and endogenous substances that are not digested by endogenous enzymes into absorbable nutrients. Many species of indigenous bacteria can ferment sugars, starches, and structural carbohydrates into short-chain fatty acids. The short-chain fatty acids, which are predominantly acetic, propionic, and butyric acids, are readily absorbed and serve as an additional source of energy. These bacteria also synthesize microbial protein and the B-complex vitamins that may be useful to their host. *See* BACTERIAL PHYSIOLOGY AND METABOLISM.

The nutritional contributions of hindgut bacteria to mammals are illustrated in **Fig. 8**. Hindgut bacteria produce short-chain fatty acids by the fermentation of dietary carbohydrates that have escaped digestion and absorption in the midgut, and endogenous carbohydrate, such as that present in mucus. They also utilize dietary and endogenous compounds, such as digestive enzymes and urea, for the production of ammonia and microbial protein. Although most of the urea, which is the waste product of protein metabolism by the mammalian liver, is excreted by the kidneys, substantial amounts enter the gut by diffusion across its epithelial lining. The hindgut bacteria of birds, reptiles, and adult amphibians perform similar functions, except that uric acid is the waste product of hepatic protein metabolism and it enters the hindgut via the cloaca rather than by diffusion. The bacteria in the forestomach of mammalian herbivores also produce short-chain fatty acids from dietary carbohydrates, and ammonia and microbial protein from dietary protein and from urea, which reaches the forestomach via the saliva and diffusion across its epithelial lining.

Absorption of short-chain fatty acids provides the major source of the energy required by hindgut and forestomach epithelial cells and a portion of the maintenance energy required by the species. Some of the ammonia is used for the synthesis of bacterial protein, but most of the remainder is absorbed and returned to the liver for the synthesis of protein and urea. This aids in the conservation of nitrogen and reduces the amount of urea (and water) that must be excreted in the urine. The absorption of short-chain fatty acids and ammonia also plays a critical role in the hindgut absorption of water. The microbial protein and B-complex vitamins synthesized by these bacteria also can be utilized as nutrients by coprophagic and forestomach fermenting species.

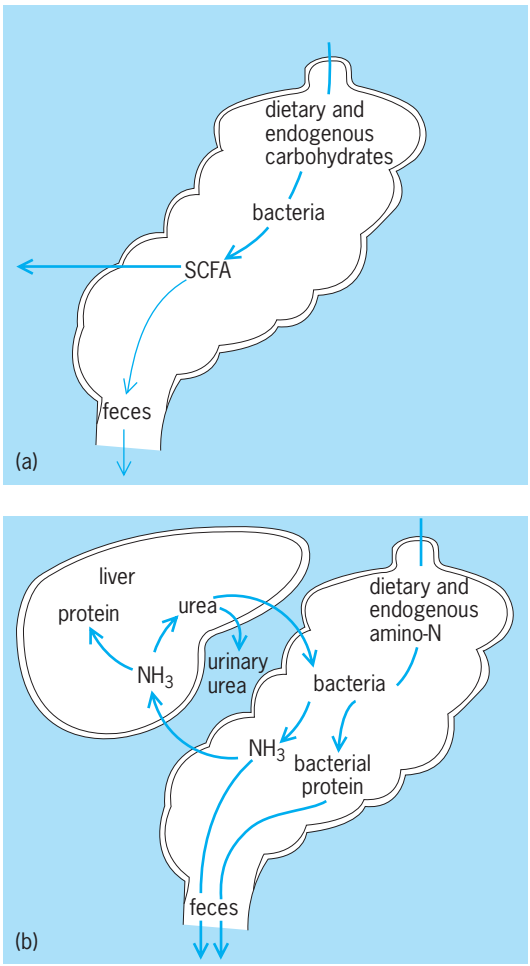

Fig. 8. Bacterial fermentation of carbohydrates and metabolism of nitrogenous compounds in the hindgut of mammals. (*a*) Fermentation of dietary and endogenous carbohydrates to short-chain fatty acids (SCFA), which are largely absorbed. (*b*) Bacterial production of ammonia and microbial protein from dietary and endogenous nitrogenous compounds. The endogenous compounds include digestive enzymes released in the upper digestive tract and the urea that passively diffuses into the gut. Ammonia is absorbed and returned to the liver for incorporation into protein and urea, which serves to conserve nitrogen and reduce the amounts of urea and water excreted by the kidneys. Bacteria in the hindgut of reptiles and birds serve similar functions, except that the uric acid that is produced instead of urea enters the hindgut via the cloaca. (*Modified from A. M. Wrong and A. J. Vince, Urea and ammonia metabolism in the human large intestine, Proc. Nutrit. Soc., 43:77–86, 1984*)

The contributions of indigenous bacteria to the production and conservation of nutrients are greatest in herbivores. Although it has been estimated that short-chain fatty acid absorption provides 4% of total maintenance energy requirement by dogs and 6–10% of the maintenance energy required by humans, they can account for 30% of the maintenance energy of rabbits and up to 70% of maintenance energy of horses and ruminants. The dietary protein and water requirements of rabbits, horses, and ruminants can be greatly reduced by the recycling of urea nitrogen; and the dietary requirements for both protein and B-vitamins are substantially reduced by coprophagy in rabbits and foregut fermentation in ruminants.

Herbivores. The appearance of herbivores that could subsist largely on the leaves, petioles, and stems of plants played a significant role in the expansion and distribution of vertebrates. Herbivory requires the ingestion of large quantities of plant material, its reduction to small particle size, and either the rapid passage of this material through the gut or its prolonged storage for bacterial fermentation. Herbivores are relatively rare among fish, larval amphibians, reptiles, and birds, and all adult amphibians are carnivores. Most herbivorous fish reduce plant material to particle size by the use of a pharyngeal mill or gizzardlike stomach and pass it rapidly through the intestine, using their midgut as the principal site of microbial fermentation. The midgut is also the principal site in larval amphibians and the Florida red-bellied turtle, emu, and panda, but the colon is the principal site in most herbivorous reptiles.

The rarity of herbivorous reptiles has been attributed to the lack of an efficient masticatory apparatus and small body size of most reptiles. The articulation of the reptilian jaws precludes the grinding action that can reduce plant material to a fine size, and the increase in mass-specific metabolic rate with a decrease in body mass limits the gut capacity and, thus, minimum body size of most reptilian herbivores to about 300 g (11 oz). This rarity of herbivores among present-day reptiles has raised questions about why herbivory was so successful for dinosaurs. The sauropods included many herbivores and the largest species of dinosaurs, and all ornithopods are believed to have been herbivores. Their success has been attributed partly to a gizzardlike stomach in the sauropods and an efficient masticatory apparatus in the ornithopods. It has also been attributed to the possibility that many dinosaurs were endothermic. However, studies of present-day reptiles suggest that the lower rates of metabolism and digesta passage that accompany a reduction in the ambient temperature may compensate for the reduction in the rates of digestion and absorption. Therefore, the rarity of present-day herbivores may be due to lack of an efficient masticatory apparatus and the susceptibility of larger species to predation by endotherms.

Although the gizzard of birds provides an efficient apparatus for the grinding of plant material, their gut capacity is limited by the requirement for flight. The herbivorous birds are confined to grouse, partridge, and the hoatzin, which fly only short distances, and the flightless rheas, emus, and ostriches. The midgut is the main site of microbial fermentation in the emu, and the colon is the principal site in the ostrich. However, among the smaller herbivores, the ceca are the principal sites in the rhea, grouse, and partridge, and the foregut (crop and distal esophagus) is the principal site in the hoatzin. The cecal fermenters include snow partridge and snow cocks, which inhabit altitudes 5000 m (16,400 ft) above sea level, and the arctic ptarmigan.

In contrast to the fish, amphibians, reptiles, and birds, herbivory was adopted by a high percentage of mammals. Herbivorous species are found in 11

of the 20 mammalian orders, and 5 of these orders contain only herbivores. The herbivorous mammals constitute the largest number of species, include the largest terrestrial species, show the widest range of ecological distribution, and provide over 90% of the mammalian biomass. Their success is due to an efficient masticatory apparatus and the development of a hindgut or foregut that can store large amounts of plant material for long periods of time.

The principal site of microbial fermentation in the herbivorous panda is a long midgut, similar to that of the bear (Fig. 6), but other mammalian herbivores have adopted the strategies of colon, cecal, or foregut fermentation. Many herbivores over 30 kg (66 lb) in body weight, including the wombats, orangutans, gorillas, tapirs, equids, rhinos, and elephants, have adopted the colon as principal site of microbial fermentation. However, most herbivores less than 30 kg in body weight are cecal fermenters, and most herbivores of intermediate body weight are foregut fermenters.

As with the birds, most small mammalian herbivores overcame the limitations of body weight on gut capacity by the rapid passage of large digesta particles through the gut and selective retention of fluid, bacteria, and small particles in an enlarged cecum. Included are the lagomorphs (rabbits, hares, and pica), folivorous marsupials (koalas, greater gliders, cuscuses, and some possums), and the herbivorous rodents. Many of these species practice coprophagy, which recovers microbial protein and B-vitamins from the feces; and some of these species ingest only the nutritionally rich feces that are derived from a periodic release of cecal contents. Lagomorphs and herbivorous rodents are distributed throughout all regions of the world. The forestomach is the principal site of microbial fermentation in a few small herbivores (rat kangaroos, sloths, colobid monkeys, and the smallest ruminants) and most intermediate-sized herbivores (kangaroos, wallabies, camelids, and advanced ruminants).

Colon and cecal fermenters have the advantages of recovering nutrients from the readily digestible components of the diet from their midgut, prior to the less efficient process of microbial digestion, and increasing their food intake, digesta passage, and recovery of these nutrients on high-fiber diets. However, foregut and coprophagic cecal fermenters make use of the protein and B vitamins synthesized by gut microbes. Foregut fermentation sacrifices the advantage of recovering readily available nutrients prior to microbial fermentation. However, it allows the microbial digestion of plant toxins and provides the most efficient system for the recovery of large amounts of the electrolytes and water secreted by herbivores.

Ruminants have the additional ability of selectively retaining large particles of plant fiber in their forestomach. This retention allows a more complete fermentation of low-quality forage and accounts for the ruminants' adaptation to desert, high altitudes, and the arctic tundra. Therefore, the adoption of cecal and foregut fermentation expanded herbivory into smaller birds and mammals, and extended their distribution into regions where forage and water are sparse.

Absorption. The cells lining the small intestine of many newborn mammals are capable of absorbing intestinal contents by a process called phagocytosis. Intestinal contents are engulfed by the lumen-facing membrane into cell vesicles. Cellular enzymes digest most of the vesicular contents, but the immunoglobulins in the mother's milk are transported intact and provide the newborn with a passive immunity to many diseases. Phagocytosis is absent in the newborn of some species and short-termed in others, such as humans. Although it can continue for longer periods in a few species, such as the rat, it disappears in all mammals by the time of weaning.

The epithelial cells that line the gastrointestinal tract are closely attached to one another at their lumen-facing border by tight junctions, which are relatively impermeable to most substances other than water. Therefore, the major restriction for the absorption of most substances from the lumen into the blood is the apical and basolateral membranes of these cells. The epithelial cell membranes consist of a lipid-protein barrier that is permeated with a limited number of small, water-filled pores. Thus, they are relatively permeable to lipid-soluble substances and water but relatively impermeable to the passive diffusion of most water-soluble substances. *See* CELL PERMEABILITY.

Lipid-soluble substances. Lipid-soluble substances can be transported across the apical cell membranes by passive diffusion down their concentration gradient. The short- and medium-chain fatty acids that result from lipid digestion in the small intestine pass directly into the blood. However, the monoglycerides and long-chain fatty acids are resynthesized into triglycerides by the epithelial cells in the midgut and incorporated into small spheres (chylomicrons), which are transported across the basolateral membrane into the lymphatic system. Fat-soluble vitamins, long-chain alcohols, and other lipids also appear to be incorporated into chylomicrons and to enter the lymphatic system.

Water-soluble compounds. The intestinal cell membranes are relatively impermeable to the passive diffusion of water-soluble monosaccharides, amino acids, vitamins, and minerals that constitute a major portion of the required nutrients. These nutrients are selectively transferred across the intestinal cell membranes by carrier-mediated transport. Membrane carriers combine with the nutrient at one membrane surface and pass it across the membrane for release at the opposing surface. Some simply facilitate the diffusion of a substance down its concentration gradient; others are capable of transporting a nutrient against its concentration gradient, which requires either a direct or indirect investment of cellular energy. *See* CELL MEMBRANES.

Fructose is transported by a carrier-mediated process that facilitates its diffusion across the apical membrane, but glucose and galactose are transported into the epithelial cells of the small intestine

against their respective concentration gradients. Their absorption into the cell is dependent upon the simultaneous absorption of sodium ions (Na^+) down the concentration gradient provided by transport of sodium by another carrier against its concentration gradient across the basolateral membrane. This relationship between glucose and sodium absorption was first suggested by the finding that cola beverages, which contain high levels of glucose and sodium, were much more effective than glucose solutions when administered orally for treatment of the diarrhea associated with cholera. The amino acids, dipeptides, and B-complex vitamins also are transported across the apical membrane of these intestinal cells by a variety of Na^+-dependent, carrier-mediated transport systems. The transport of these substances across the basolateral membranes of the cell is accomplished by sodium-independent carriers.

Short-chain fatty acids and ammonia. The short-chain fatty acids and ammonia derived from bacterial fermentation are present in both undissociated and dissociated (ionized) form. Cell membranes are relatively permeable to the passive diffusion of the lipid soluble undissociated short-chain fatty acids and ammonia, but impermeable to the diffusion of the water-soluble ionized forms that predominate at the normal pH of hindgut and forestomach contents. Therefore, a reduction in digesta pH tends to increase the rate of short-chain fatty acid absorption and decrease the rate of ammonia absorption. However, there is also evidence for the absorption of short-chain fatty acid ions in exchange for bicarbonate (HCO_3^-) and ammonia ions (NH_4^+) in exchange for hydrogen (H^+).

Minerals. The metabolic processes of the body require a number of different minerals. Some such as iron, calcium, sodium, and chloride are required in relatively large quantities. Others such as manganese and zinc are labeled trace minerals because they are required in only minute amounts. The absorption of iron and calcium involves membrane carriers. Absorption of sodium and chloride ions is discussed below.

Diet adaptations. The rates of nutrient absorption can be affected by the diet. The ratio of sugar to amino acid absorption appears to be highest in the herbivores and lowest in the carnivores of all classes of vertebrates. A change from a high-carbohydrate and low-protein diet to a high-protein and low-carbohydrate diet also can increase the rate of glucose absorption and decrease the rate of amino acid absorption by individual animals. An increase in dietary calcium, iron, copper, phosphorus, or B-complex vitamins, which can be toxic at high concentrations, appears to decrease their rates of absorption.

Electrolytes and water. The nutrient that is required in largest quantity for digestion, absorption, metabolism, and excretion of waste products is water. Because it readily diffuses across cell membranes down its concentration gradient, the net secretion or absorption of water is determined by the net secretion or absorption of all other substances. Sodium, chloride, and bicarbonate are the principal ions that are present in the extracellular fluids that bathe the body cells of all vertebrates and that are transported across cell membranes. Therefore, the transport of these electrolytes is the major driving force for the secretion or absorption of water.

Secretion. Solubilization, digestion, and the transit of food and digesta by the alimentary tract require the secretion of large volumes of fluid. Secretions by the salivary glands, pancreas, biliary system, and intestinal cells contain high concentrations of sodium, chloride, and bicarbonate ions, and gastric secretions contain high levels of hydrogen and chloride ions. The secretion of hydrochloric acid reduces the pH of gastric contents to the levels required for activation and activity of pepsin. However, the midgut requires a more alkaline pH for protection of its epithelium and the activation of the pancreatic enzymes. This is accomplished by the bicarbonate ions secreted by the pancreas, biliary system, and intestinal mucosa. The short-chain fatty acids produced by microbial fermentation can also reduce the pH of gut contents, with damaging results. Short-chain fatty acids produced in the forestomach are neutralized by bicarbonate secreted from the salivary glands and the forestomach epithelium, and those produced in the hindgut are neutralized by the bicarbonate secreted into the midgut and hindgut. *See* PH REGULATION (BIOLOGY).

Under normal conditions, an adult human consumes 2 liters (1.8 qt) of water per day, and the digestive system secretes an additional 7 liters (6.4 qt) of fluid per day; resulting in a total of 9 liters (8.2 qt) of fluid. This is equivalent to 40% of the water in the extracellular fluid space. An uncompensated loss of 15% of the extracellular fluid volume can produce clinical signs of dehydration, and a loss of 35% can result in death. This is prevented by absorption of 7 liters of this fluid by the small intestine and an additional 1.4 liters (1.3 qt) by the large intestine, leaving only 0.1 liter (0.09 qt) for excretion into the feces. Although the large intestine normally absorbs only 20% of this fluid, this percentage is critical, and its absorptive capacity can be doubled if necessary. The daily secretions of the digestive system are much higher in herbivores, reaching levels equivalent to two or three times the extracellular fluid in the sheep and horse, but with a similar 95–98% recovery prior to excretion.

Absorption. The absorption of water from the midgut is accomplished mainly by the absorption of sodium, chloride, vitamins, and the end products of carbohydrate, lipid, and protein digestion by the endogenous enzymes. Some of the sodium is absorbed in conjunction with monosaccharide, amino acid, and vitamin absorption, but some of the sodium and most of the chloride are absorbed by a mechanism that exchanges sodium for the hydrogen and chloride for the bicarbonate in the epithelial cells. The hydrogen and bicarbonate entering the gut combine to form carbon dioxide (CO_2) and water.

The hindgut appears to have evolved initially for conservation of the electrolytes, water, and nitrogen in both the digestive secretions and urinary excretions of terrestrial vertebrates. The kidney is the

principal organ for excretion of metabolic waste products, which it accomplishes by continuously filtering blood through the glomerulus of the nephron and reabsorbing water and nutrients from the renal tubule. The earliest vertebrates are believed to have been fresh-water fish, whose major problem was the rejection or elimination of excess water. Fresh-water fish do not intentionally drink water, and excess water is excreted by the kidneys.

Marine fish face the opposite problem: the potential loss of water and gain in sodium chloride (NaCl) due to the high salinity and osmotic activity of seawater. Although marine fish drink seawater, its osmotic effect on the digestive system is reduced by the absorption of sodium chloride from their esophagus. Hagfish, sharks, and rays have resolved the osmotic problem by having body fluids that are isotonic or hypertonic to seawater. However, this problem is resolved in most marine fish by the absence of glomerular filtration by the kidney and active secretion of sodium chloride by the gills and other tissues such as the rectal glands of sharks. *See* KIDNEY.

Terrestrial animals must conserve both their water and electrolytes against loss to the environment. Their kidneys, like those of fresh-water fish, filter blood and reabsorb electrolytes and water, but the urine of amphibians, reptiles, and birds is excreted into a cloaca. Antiperistaltic contractions then reflux urine along the length of the hindgut of many reptiles and birds, allowing the cloaca and hindgut to aid the kidney in the recovery of electrolytes and water from both the digestive secretions and urinary excretions.

In most mammals, the digestive and urinary systems exit the body separately, and the nephron of the kidney is lengthened by the loop of Henle, which provides for more efficient recovery of water. The hindgut continues to function as the final site for recovery of digestive fluids, and tends to be longer than that of lower vertebrates and longest in those that inhabit arid environments. Pacemaker tissue located in the midcolon generates antiperistaltic waves of contraction, which continue to aid retention of digesta in the cecum and proximal colon.

The absorption of water from the hindgut is due principally to the absorption of sodium, short-chain fatty acids, and ammonia. Sodium is absorbed in exchange for hydrogen by a mechanism that partly depends on short-chain fatty acid absorption and by a mechanism that transports sodium alone. Thus, prolonged retention of digesta by the hindgut of terrestrial vertebrates increased both the time available for absorption of electrolytes and water, and the multiplication of endogenous bacteria that contribute to this process.

Neural and endocrine control. The motor and secretory activities of the digestive system are controlled and integrated by the nervous system, hormones, and paracrine secretions. Many of these activities are controlled by nerve reflexes, consisting of sensory neurons with receptors in the digestive tract that pass information, via other neurons in the central nervous system and gut wall, back to the digestive tract. The latter neurons release neurotransmitters or modulators that stimulate or inhibit muscular contraction or cellular secretions. Reflexes responsible for the initial stage of deglutition and final stage of defecation are monitored through brain centers, allowing for voluntary control of these activities. However, most of the motor and secretory events in the digestive system are under the involuntary and sometimes interactive control of a large number of neurotransmitters, neuromodulators, hormones, and paracrine agents.

Although much of the neuroendocrine control provides a direct response to conditions in a given segment of the gut, some of these responses prepare a following segment for what is to come, or tune down the activities in a previous segment. The sight or smell of food can stimulate gastric secretions, and the act of deglutition produces a receptive relaxation of the stomach. Distention of the stomach can stimulate contractions of the colon, which accounts for the relationship between meals and defecation. Similarly, the entrance of partially digested food into the small intestine not only stimulates the secretion of pancreatic fluid and the release of bile but also decreases the rates of gastric secretion and motility.

For example, the sight or smell of food and its presence in the stomach result in neural and chemical stimulation of the release of the hormone gastrin from the stomach wall. Gastrin stimulates the secretion of hydrochloric acid and pepsinogen for the digestion of protein. As the digestive process progresses, the passage of hydrochloric acid and the end products of gastric digestion into the small intestine stimulates the release of the hormones secretin and cholecystokinin from the intestinal wall. These hormones act to inhibit the secretion of gastrin and stimulate the release of fluid and enzymes from the pancreas and the release of bile from the gallbladder. The high concentration of bicarbonate ions in these secretions neutralizes the hydrochloric acid in the gastric effluent, and the pancreatic enzymes and bile aid in the digestion of starch, proteins, and lipids.

Adaptations

Some characteristics of the vertebrate digestive system are common to all classes of vertebrates and appear to have been conserved in their evolution. Included are most of the enzymes responsible for digestion of carbohydrates, proteins, and lipids, and the mechanisms for absorption of nutrients by the midgut or small intestine. However, the headgut and hindgut show a variety of adaptations to the diet, the environment, or physiological characteristics of the animal.

Many adaptations of the mouthparts and feeding practices to the diet are seen among species in all classes of vertebrates. Prolonged retention of digesta by the hindgut or large intestine for the recovery of urinary and digestive fluids appears to have first developed in the adult amphibians. This resulted in the production of large populations of indigenous bacteria, whose digestive end products aid in the resorption of fluids and can provide additional nutrients.

Further adaptations of the hindgut led to the evolution of herbivores, which have access to a much wider range of food. The major advances in the digestive system of mammals were the development of an extremely efficient masticatory apparatus and, in some species, a more complex and voluminous large intestine or stomach. These advances account for the marked success of the mammalian herbivores and suggest that the success of the herbivorous dinosaurs may have been based on similar developments.

The digestive systems of different species are designed for different diets. The digestive system of most carnivores is designed for the intermittent consumption of high-protein diets, often at intervals of 12–18 h in many mammalian carnivores and longer periods of time for cold-blooded species. The herbivore digestive system is designed for the almost continuous consumption of large volumes of plant material. When left to their normal diet, horses and elephants may spend 18 h a day foraging for food, and ruminants spend 8 h grazing and an additional 8 h remasticating their cud.

A number of the digestive diseases of herbivores can be attributed to the feeding of high-concentrate, low-fiber diets. High levels of rapidly fermentable starch can result in a fulminating production of short-chain fatty acids and the gases CO_2 and methane (CH_4) in the forestomach of cattle and the large intestine of horses. The short-chain fatty acids lower the pH of gut contents, resulting in the growth of lactobacilli and the production of lactic acid. The high levels of short-chain and lactic acids can produce a hypertonic digesta and systemic dehydration, and their absorption can damage the gut epithelium, inhibit gut motility, and result in systemic acidosis. Excessive gas production can distend and displace the gut and inhibit its motility. This syndrome is responsible for a number of diseases involving the forestomach of ruminants and for the high incidence of torsion, impaction, and colic in the large intestine of horses on high-concentrate diets. Although humans are omnivores, the much greater prevalence of cancer and other diseases of the colon on the high-protein, high-carbohydrate, high-lipid, low-fiber diets of affluent Western societies suggests that the human large intestine is subject to similar problems.

The nutritional requirements of vertebrates are provided by an extremely complex system with a variety of different and often exotic adaptations to the diet, habitat, and other physiological characteristics of a species. Studies of these adaptations have provided important contributions to understanding of basic physiological mechanisms. They also provide information needed for the maintenance of domesticated and captive (zoo) animals and survival of endangered species. The extinction of vertebrates at a rate of approximately one species each year could be prevented or at least delayed by a better understanding of these species and their contributions to the beauty and balance of nature. C. Edward Stevens

Bibliography. A. C. Guyton, *Textbook of Medical Physiology*, 8th ed., 1991; M. Hildebrand et al., *Functional Vertebrate Morphology*, 1985; L. R. Johnson et al., *Physiology of the Gastrointestinal Tract*, 2d ed., 1987; C. E. Stevens and I. D. Hume, *Comparative Physiology of the Vertebrate Digestive System*, 2d ed., 1995; C. E. Stevens and I. D. Hume, Contributions of microbes in the vertebrate gastrointestinal tract to the production and conservation of nutrients, *Physiol. Rev.*, 78:393–427, 1998; J. F. V. Vincent and P. J. Lillford, *Feeding and the Texture of Food*, 1991; H. J. Vonk and J. R. H. Western, *Comparative Biochemistry and Physiology of Enzymatic Digestion*, 1984.

Digital computer

A device that processes numerical information; more generally, any device that manipulates symbolic information according to specified computational procedures. The term digital computer—or simply, computer—embraces calculators, computer workstations, control computers (controllers) for applications such as domestic appliances and industrial processes, data-processing systems, microcomputers, microcontrollers, multiprocessors, parallel computers, personal computers, network servers, and supercomputers. *See* CALCULATORS; DIGITAL CONTROL; MICROCOMPUTER; PROGRAMMABLE CONTROLLERS; SUPERCOMPUTER.

A digital computer is an electronic computing machine that uses the binary digits (bits) 0 and 1 to represent all forms of information internally in digital form. Every computer has a set of instructions that define the basic functions it can perform. Sequences of these instructions constitute machine-language programs that can be stored in the computer and used to tailor it to an essentially unlimited number of specialized applications. Calculators are small computers specialized for mathematical computations. General-purpose computers range from pocket-sized personal digital assistants (notepad computers), to medium-sized desktop computers (personal computers and workstations), to large, powerful computers that are shared by many users via a computer network. The vast majority of digital computers now in use are inexpensive, special-purpose microcontrollers that are embedded, often invisibly, in such devices as toys, consumer electronic equipment, and automobiles. *See* BIT; COMPUTER PROGRAMMING; EMBEDDED SYSTEMS.

The main data-processing elements of a computer reside in a small number of electronic integrated circuits (ICs) that form a microprocessor or central processing unit (CPU). Electronic technology allows a basic instruction such as "add two numbers" to be executed many millions of times per second. Other electronic devices are used for program and data storage (memory circuits) and for communication with external devices and human users (input-output circuits). Nonelectronic (magnetic, optical, and mechanical) devices also appear in computers. They are used to construct input-output devices such as keyboards, monitors (video screens), secondary

memories, printers, sensors, and mechanical actuators.

Digital System Basics

The two fundamental ways of representing information are analog and digital. Analog variables take on continuous values that are not separated by gaps, for example, distance measurements on a scale. Digital variables employ a fixed set of noncontinuous or discrete symbols (digits). Analog computers process continuous quantities; a slide rule is an early example. Such devices have been superseded by digital computers, which process information encoded in digital form; the terms computer and digital computer are now largely synonymous. *See* ANALOG COMPUTER; COMPUTER.

Binary codes. The decimal number system preferred by humans employs 10 different digits. The simpler binary or two-valued number system is better suited to computers (see **table**). The bits 0 and 1 have various physical realizations, for example, the position of a switch (off = 0 and on = 1) or an electrical voltage level ($V_{low} = 0$ and $V_{high} = 1$). Modern computers employ both of these concepts by using transistors as high-speed electronic switches to control binary voltage signals. *See* NUMBERING SYSTEMS; TRANSISTOR.

Information is stored and processed by computers in fixed-sized units called words. Common word sizes are 8, 16, 32, and 64 bits. Four-bit words can be used to encode the first 16 integers (see table). By increasing the word size, the number of different items that can be represented and their precision can be made as large as desired. A common word size in personal computers is 32 bits, which allows $2^{32} = 4{,}294{,}967{,}296$ distinct numbers to be represented.

Computer words can represent many different forms of information, not just numbers. For example, 8-bit words called characters or bytes are used to encode text symbols (the 10 decimal digits, the 52 upper- and lowercase letters of the English alphabet, and punctuation marks). A widely used code of this type is ASCII (American Standard Code for Information Interchange). Visual information can be reduced to black and white dots (pixels) corresponding to 0's and 1's. Audio information can be digitized by mapping a small element of sound into a binary word; for example, a compact disk (CD) uses several million 16-bit words to store an audio recording. Logical quantities encountered in reasoning or decision making can be captured by associating 1 with true and 0 with false. Hence, most forms of information are readily reduced to a common, numberlike binary format suitable for processing by computer. *See* COMPACT DISK; IMAGE PROCESSING; LOGIC.

Integers from 0 to 15 in decimal and binary form

Number	Decimal code	Binary code
Zero	00	0000
One	01	0001
Two	02	0010
Three	03	0011
Four	04	0100
Five	05	0101
Six	06	0110
Seven	07	0111
Eight	08	1000
Nine	09	1001
Ten	10	1010
Eleven	11	1011
Twelve	12	1100
Thirteen	13	1101
Fourteen	14	1110
Fifteen	15	1111

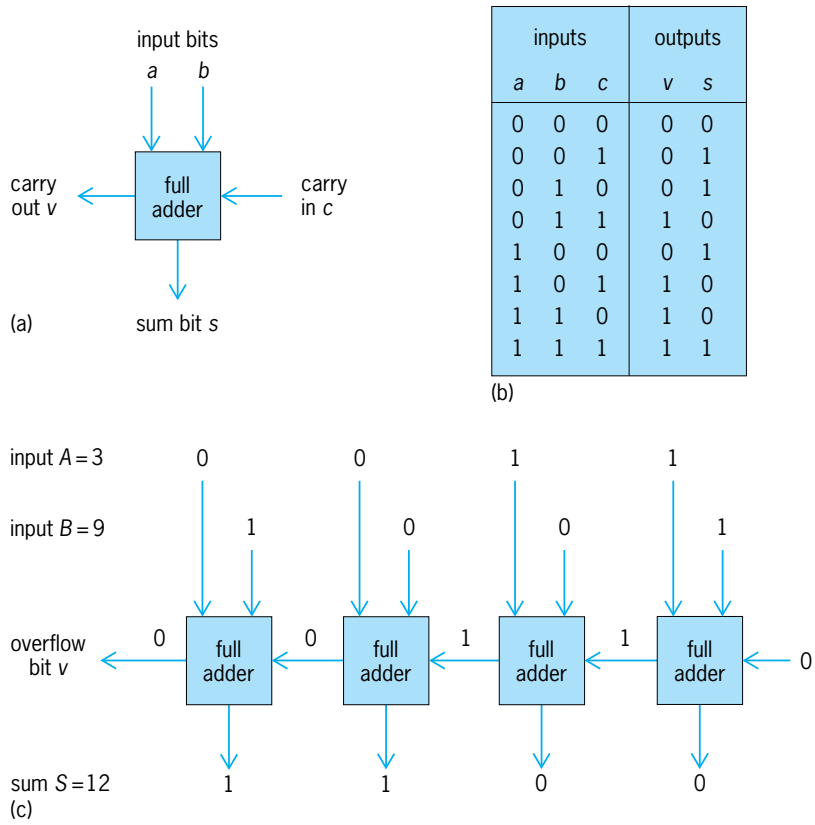

Fig. 1. Logic components. (*a*) Full adder. (*b*) Behavior of a full adder. (*c*) Four-bit adder circuit composed of four full adders.

Logic components. The operation of a digital computer can be viewed at various levels of abstraction, which are characterized by components of different complexity. These levels range from the low, transistor level seen by an electronic circuit designer to the high, system level seen by a computer user. A useful intermediate level is the logic level, where the basic components process individual bits. An example of such a small logic component is the full adder (**Fig. 1***a*), which computes the sum of three bits, a, b, and c. For example, if the input values are $a = b = c = 1$, the output of the full adder is $vs = 11$, denoting three. The behavior of a full adder is specified by a truth table (Fig. 1*b*). By linking four copies of a full adder (Fig. 1*c*), a 4-bit adder is obtained, that is, a circuit that computes the sum of two 4-bit numbers. This adder serves as a component of a larger logic circuit, and is represented by a single box symbol in the diagram of that circuit (**Fig. 2**). The adder's input-output lines, A, B, and S, are grouped into word-transmitting lines called buses. By using

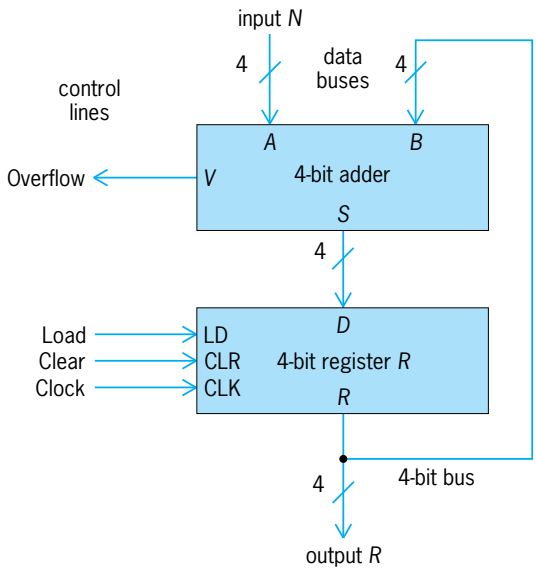

Fig. 2. Four-bit accumulator circuit.

other basic components called gates, logic circuits can be constructed to perform many useful operations. *See* LOGIC CIRCUITS.

A register is a logic component intended to store a word of information. The current contents or state of a register, R, for storing 4-bit words (Fig. 2) appear on its 4-bit output bus, also called R. The register has a 4-bit input bus, D, on which new data are placed for entry into R. It also has several control lines, two of which, Load and Clear, are used to change the state of R. The third control line, Clock, is for timing purposes, and synchronizes the actions of R with those of other registers. Setting the Load control line to 1 causes the current data on D to be entered (loaded) into R and stored there; if Load $= 0$, the load operation does not occur. Setting the Clear control line to 1 changes the state of R to zero. The precise times at which the register responds to Load and Clear are determined by the signals applied to the Clock line.

An accumulator is a circuit that combines an adder and a register (Fig. 2). It supports a basic type of add instruction found in almost all computers. This instruction is represented symbolically by ADD R, N, which is interpreted as follows: Add the external data, N, to the current contents of the accumulator register, R, and make the resulting sum the new contents of R. Another common instruction, INC R, meaning increment the contents of register R by one, can be implemented by the accumulator by first making $N = 1$ and then executing ADD R, N.This causes the accumulator to behave as a binary counter. *See* DIGITAL COUNTER.

Instructions that make decisions can also be realized by logic circuits. For example, if an answer is needed to the question "Is $R = 0$?" it can be obtained from the 4-bit accumulator by setting $N = 15$, the largest permissible 4-bit number, and then executing the instruction ADD R, N. If $R = 0$, the new value $R + N$ in the register is $0 + 15 = 15$, and the adder's output signal Overflow is 0. If R is greater than 0, then $R + N$ is a 5-bit number between 16 and 30, so the adder overflows, making Overflow $=$ 1. Hence, the Overflow signal, in conjunction with execution of the appropriate form of the ADD instruction, answers the given question.

System organization. An accumulator is a digital system that constitutes a simple processor capable of executing a few instructions. By introducing more data-processing circuits and registers, as well as control circuits for a larger set of instructions, a practical, general-purpose processor can be constructed. Such a processor forms the "brain" of every computer, and is referred to as its central processing unit. A CPU implemented on a single integrated-circuit chip is called a microprocessor. Central processing units have evolved steadily from the simple accumulatorlike processors used in the first generation of electronic computers in the 1940s, to complex microprocessors containing millions of transistors and capable of executing several hundred different instruction types. *See* MICROPROCESSOR.

A typical computer program is too large to store in the CPU, so another component called the main memory is used to store a program's instructions and associated data while they are being executed (**Fig. 3**). Main memory consists of high-speed integrated circuits designed to allow storage and retrieval of information one word at a time. All words in main memory can be accessed with equal ease; hence this is also called a random-access memory (RAM).

A computer program is processed by loading it into main memory and then transferring its instructions and data one word (or a few words) at a time to the CPU for processing. Hence, there is a continual flow of instructions and data words between the CPU

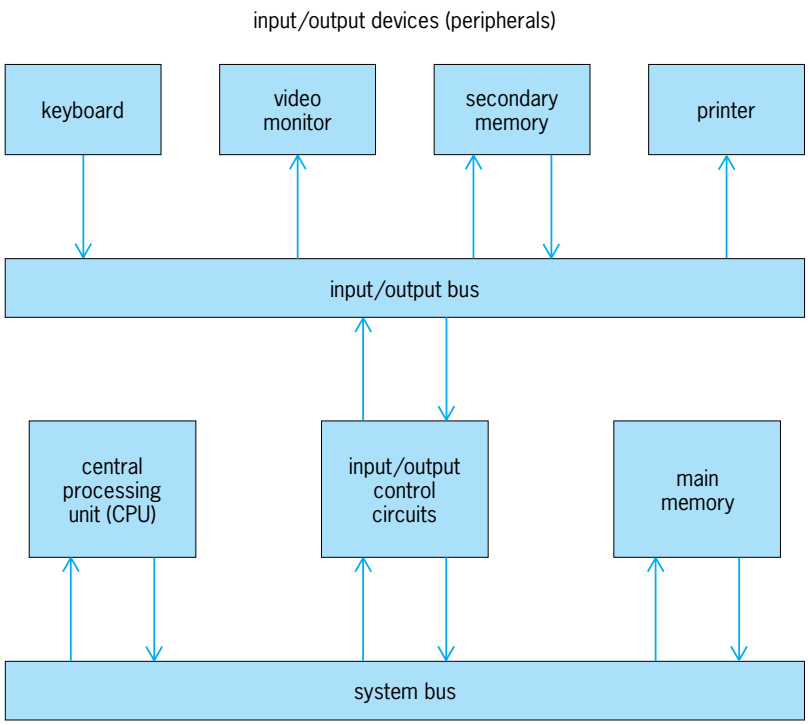

Fig. 3. General organization of a computer.

and its main memory. As millions of words must be transferred per second, a high-speed communication link is needed between the CPU and main memory. The system bus (Fig. 3) fills this role.

A computer has input-output (I/O) control circuits and buses to connect it to external input-output devices (also called peripherals). Typical input-output devices are a keyboard, which is an input device, and a printer, which is an output device. Because most computers need more storage space than main memory can supply, they also employ secondary memory units which form part of the computer's input-output subsystem. Common secondary memory devices are hard disk drives, flexible (floppy) disk drives, and magnetic tape units. Compared to main memory, secondary memories employ storage media (magnetic disks and tapes) that have higher capacity and lower cost. However, secondary memories are also significantly slower than main memory. *See* COMPUTER PERIPHERAL DEVICES; COMPUTER STORAGE TECHNOLOGY.

Central Processing Unit

The CPU contains the circuits needed to fetch instructions from main memory and execute them. Instructions are stored in consecutive memory locations or addresses, in approximately the order in which the instructions must be executed. Hence, all the CPU needs to keep track of instructions is a register that can store instruction addresses and can have its contents incremented. This register is called the program counter. Each time the CPU fetches an instruction word from memory, it increments the program counter to obtain the address of the next instruction it needs (except when branching occurs, as discussed below).

The basic operations that a computer performs are ultimately determined by the machine instruction set that is built into its CPU. All the computer's programs have to be translated into sequences of instructions drawn from this instruction set before the programs can be executed. Most programs are written in high-level programming languages such as C, Cobol, Fortran, and Pascal. The translation of such programs into executable machine-language form is carried out automatically by programs called compilers. *See* PROGRAMMING LANGUAGES.

Every possible computational task can be programmed by using a small number of different types of machine instructions, because a complex operation can be broken down into sequences of simple operations. For example, multiplication can be implemented by a sequence of additions. Thus, while some inexpensive computers have no multiply instructions, that is, they lack the logic circuits to execute multiply in one step, all can be programmed to execute a multistep multiplication algorithm. Multiplication is said to be implemented in software rather than hardware in such computers.

Instruction and data formats. Machine instructions are formatted as binary words with several parts (fields) that define the actions to be taken by the CPU when the instruction is executed. The key field is the opcode (operation code), OP, which specifies the main operation to be performed: for example, "load a 32-bit word from memory to a CPU register," or "add two 64-bit integers." The opcode field is followed by several operand address fields (two are typical), such as OP ADR1,. . . , ADR*k*. (This symbolic, nonbinary format for instructions is known as assembly language, and its format varies from CPU to CPU.) The operand addresses ADR1,. . . ,ADR*k* specify data items stored in various locations: in the instruction itself, in CPU registers, in main memory locations, or in registers associated with input-output control circuits. For example, an instruction of the form SUB R7, #13 specifies that the constant 13 is to be subtracted from the contents of the register named R7. Register R7 serves as both a data source and data destination in this two-address instruction format. The one-address instruction LOAD 102678 transfers the contents of the memory location with the indicated address 102678 to an implicit register in the CPU intended for such operations. Programs can be written in assembly languages of the above type; but as noted above, most programming is now done with higher-level languages that are more expressive and easier to use.

Numerical operands have two main formats: fixed-point to represent integers and floating-point to represent real numbers. Fixed-point numbers take the form $b_{n-1}b_{n-2}\cdots b_2b_1b_0$, where b_i is 0 or 1 and has the numerical weight 2^i (see table). The leftmost position of a number word can be reserved to represent the number's sign, with 0 denoting plus and 1 denoting minus. A floating-point number consists of a pair of fixed-point numbers M and E, which typically denote the number $M \times 2^E$. Nonnumerical data usually take the form of variable-length character strings encoded in a standard code such as ASCII.

Instruction types. The choice of the instructions to include in an instruction set is governed by several factors: the usefulness of the instructions, the cost of the hardware to implement them, and their compatibility with the instructions of other computers. Because their data are physically closer to the CPU's data-processing circuits, instructions that act on CPU registers execute faster than instructions of the same type whose operands are in memory.

The operations performed by the most common instruction types are summarized below:

1. *Load.* Copy an operand from a main memory location to a CPU register.
2. *Store.* Copy an operand to a main memory location from a CPU register.
3. *Move register.* Copy an operand from one CPU register to another.
4. *Add.* Compute the sum of two operands stored in CPU registers.
5. *Subtract.* Compute the difference of two operands stored in CPU registers.
6. *Not.* Compute the logical complement of an operand in a CPU register by changing its 1's to 0's and its 0's to 1's.
7. *Clear.* Change every bit of a CPU register to 0.
8. *Branch.* Replace the contents of the program

counter by the instruction's address field. (This is an unconditional branch instruction.)

9. *Branch on zero.* If the contents of a specified CPU register are zero, replace the contents of the program counter by the instruction's address field. (This is a conditional branch instruction.)

No explicit instructions are needed for input-output operations if input-output devices share with main memory the available memory addresses. This is known as memory-mapped input-output, and allows load and store instructions to be used to transfer data between the CPU and input-output devices.

In general, a computer's instruction set should include a selection of instructions of the following three types: (1) Data-transfer instructions that move data unchanged between the CPU, main memory, and input-output devices. (2) Data-processing instructions that perform numerical operations such as add, subtract, multiply, and divide, as well as nonnumerical (logical) operations, such as not, and, exclusive-or, and shift. (3) Program-control instructions that can change the order in which instructions are executed, for example branch, branch-on-zero, call procedure, and return from procedure.

Internal organization. The instruction unit (I unit) of a CPU (**Fig. 4**), also called the program control unit, is responsible for fetching instructions from main memory, using the program counter as the instruction address register. The opcode of a newly fetched instruction *I* is placed in the instruction register. The opcode is then decoded to determine the sequence of actions required to execute *I*. These may include the loading or storing of data assigned to main memory, in which case the I unit computes all needed addresses and issues all needed control signals to the CPU and the system bus. Data are processed in the CPU's execution unit (E unit), also called the datapath, which contains a set of registers used for temporary storage of data operands, and an arithmetic logic unit (ALU), which contains the main data-processing circuits.

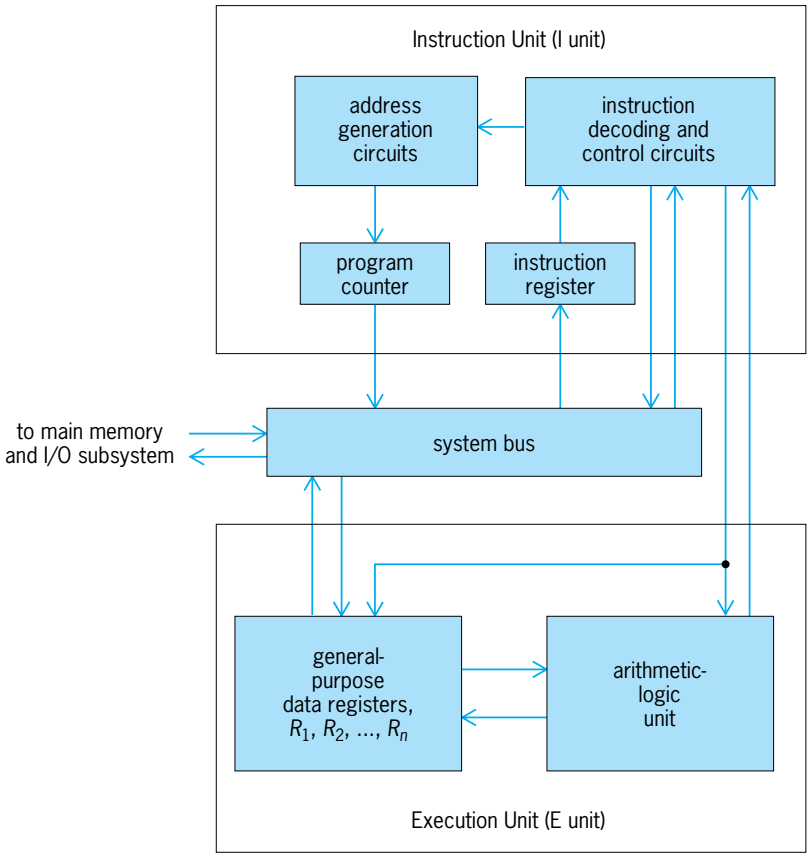

Fig. 4. Internal organization of a CPU.

Computer Spectrum

A major theme in the evolution of digital computers has been to find ways to increase the computing speed or performance of a computer at low cost. Computers have undergone a long period of development, with major advances stimulated by new technologies for manufacturing processors and memories.

Development. The modern digital computer has its origins in the work of C. Babbage, who attempted unsuccessfully to build a general-purpose, programmable computer in the mid-1800s. Around 1940, such computers were successfully built in the United States and Europe by using electronic vacuum tubes as the basic switching devices. These led to the first generation of computers. The invention of the transistor in 1947, which replaced the vacuum tube, gave rise to a second generation of smaller, faster, and more reliable computers. Most early computers were expensive, room-sized mainframe machines. *See* VACUUM TUBE.

In the mid-1960s, the first series of mainframe computers to distinguish a computer's architecture from its implementation was introduced. The architecture of a computer series is defined as a specification covering the instruction set, and the high-level structure and behavior of all members of the series. Different models in the series differ in their low-level (implementation) details. Another innovation of the 1960s was the minicomputer, a small (desk-sized) and relatively inexpensive computer with limited hardware and software facilities. Integrated transistor circuits also began to displace discrete transistor circuits, further reducing computer size and cost. A milestone occurred in 1971 with the introduction of the microprocessor. The density and speed of integrated-circuit technology grew rapidly, eventually making it possible to incorporate into a single chip most of the features of mainframe computers. As the manufacturing cost per integrated circuit is fairly constant, these developments have meant a steady improvement in the performance/cost ratio of computers. *See* COMPUTER SYSTEMS ARCHITECTURE; INTEGRATED CIRCUITS.

By 1980, single-chip microprocessors and microcomputers were being mass-produced in several varieties. The simplest microcomputers were being used as microcontrollers for dedicated applications. The more powerful microprocessors began to appear in inexpensive general-purpose computers, a development that produced the personal computer. By 1990, all computers could be placed in a few families

defined by certain microprocessor and operating-system standards. *See* OPERATING SYSTEM.

Performance measures. A simple indicator of a CPU's performance is the frequency f of its central timing signal (clock), measured in millions of clock signals issued per second or megahertz (MHz). The clock frequency depends on the integrated-circuit technology used; frequencies of several hundred megahertz are achievable with current technology. Each clock signal triggers execution of a basic instruction such as a fixed-point addition; hence, the time required to execute such an instruction (the clock cycle time) is $1/f$ microseconds. Complex instructions like multiplication or operations on floating-point numbers require several clock cycles to complete their execution. Another measure of CPU performance is the (average) instruction execution rate, measured in millions of instructions per second (MIPS).

Instruction execution time is strongly affected by the time to move instructions or data between the CPU and main memory. The time required by the CPU to access a word in main memory is typically about five times longer than the CPU's clock cycle time. This disparity in speed has existed since the earliest computers despite efforts to develop memory circuits that would be fast enough to keep up with the fastest CPUs. Maximum performance requires the CPU to be supplied with a steady flow of instructions that need to be executed. This flow is disrupted by branch instructions, which account for 20% or more of the instructions in a typical program.

To deal with the foregoing issues, various performance-enhancing features have been incorporated into the design of computers. The communication bottleneck between the CPU and main memory is reduced by means of a cache, which is a special memory unit inserted between the two units. The cache is smaller than main memory but can be accessed more rapidly, and is often placed on the same integrated-circuit chip as the CPU. Its effect is to reduce the average time required by the CPU to send information to or receive information from the memory subsystem. Special logic circuits support the complex flow of information among main memory, the cache, and the registers of the CPU. However, the cache is largely invisible to the programs being executed.

The instruction execution rate can be increased by executing several instructions concurrently. One approach is to employ several E units that are tailored to different instruction types. Examples are an integer unit designed to execute fixed-point instructions and a floating-point unit designed for floating-point instructions. The CPU can then execute a fixed-point instruction and a floating-point instruction at the same time. Processors that execute several instructions in parallel in this way are called superscalar. *See* CONCURRENT PROCESSING.

Another speedup technique called pipelining allows several instructions to be processed simultaneously in special circuits called pipelines. Execution of an instruction is broken into several consecutive steps, each of which can be assigned to a separate stage of the pipeline. This makes it possible for an n-stage E unit to overlap the execution of up to n different instructions. A pipeline processing circuit resembles an assembly line on which many products are in various stages of manufacture at the same time. The ability of a CPU to execute several instructions at the same time by using multiple or pipelined E units is highly dependent on the availability of instructions of the right type at the right time in the program being executed. A useful measure of the performance of a CPU that employs internal parallelism is the average number of clock cycles per instruction (CPI) needed to execute a representative set of programs.

CISCs and RISCs. A software implementation of a complex operation like multiply is slower than the corresponding hardware implementation. Consequently, as advances in IC technology lowered the cost of hardware circuits, instruction sets tended to increase in size and complexity. By the mid-1980s, many microprocessors had instructions of several hundred different types, characterized by diverse formats, memory addressing modes, and execution times. The heterogeneous instruction sets of these complex instruction set computers (CISCs) have some disadvantages. Complex instructions require more processing circuits, which tend to make CISCs large and expensive. Moreover, the decoding and execution of complex instruction can slow down the processing of simple instructions.

To address the defects of CISCs, a new class of fast computers referred to as reduced instruction set computers (RISCs) was introduced. RISCs are characterized by fast, efficient—but not necessarily small—instruction sets. The following features are common to most RISCs: (1) All instructions are of fixed length and have just a few opcode formats and addressing modes. (2) The only instructions that address memory are load and store instructions; all other instructions require their operands to be placed in CPU registers. (3) The fetching and processing of most instructions is overlapped in pipelined fashion.

An example of a RISC is the PowerPC microprocessor (**Fig. 5**), developed in the early 1990s. It is a family of single-chip, 32-bit microprocessors which share the same RISC architecture. The PowerPC has a set of 32 general-purpose registers for storing integer operands, and 32 more registers for floating-point operands. There are more than 200 distinct instruction types, all one word in length. The only instructions that address memory are load and store. Extensive use is made of pipelines to ensure that most instructions can be executed in one clock cycle.

The PowerPC incorporates several speedup features which are not unique to RISCs. It contains a cache whose size and organization varies with the model; the cache of the model 601 (Fig. 5) has a capacity of 32 kilobytes (2^{15} bytes) and stores both data and instructions. The PowerPC has several E units and so falls into the superscalar category. The model 601 has three E units: an integer execution unit, a floating-point unit, and a branch processing

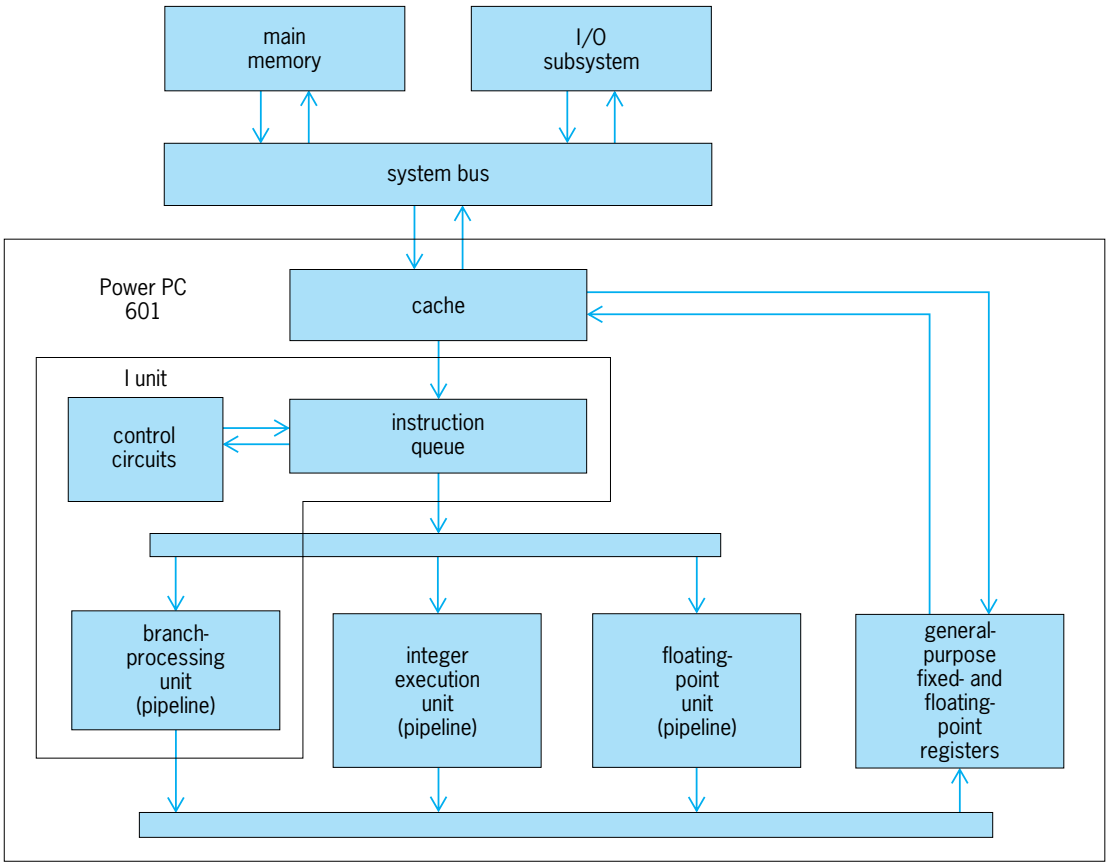

Fig. 5. Organization of the PowerPC computer, model 601.

unit, which allow up to three instructions to complete execution in one clock cycle. The integer unit executes all fixed-point numerical and logical operations, including those associated with load and store instructions. Although part of the CPU's I unit, the branch processing unit is considered an E unit for branch instructions. To keep the three E units supplied with instructions, the I unit can fetch up to eight instructions simultaneously from the cache.

Parallel processing. The degree of parallelism possible with a pipeline is small, less than 10. A parallel processing technique with the potential of achieving very great parallelism is the use of multiple independent computers or CPUs operating in unison. For example, a set of computers in a computer network can be programmed to work concurrently on different parts of the same problem. Such a loosely coupled computer system is useful for computing tasks that can easily be partitioned into independent subtasks, with infrequent communication of results among the subtasks.

Some large-scale computations can be divided into subtasks that require more frequent and rapid exchange of results among the subtasks than is possible with a computer network. To address this interprocessor communication problem, computers have been built whose CPUs are tightly coupled so they can access one another's data rapidly; these are called multiprocessors. The task of writing programs for multiprocessors is not as well understood as the corresponding problem for a single-processor computer. Nevertheless, powerful multiprocessors employing many inexpensive microprocessors as their CPUs are available commercially. They are increasingly used in applications that have extremely high performance requirements. *See* MULTIPROCESSING.

John P. Hayes

Bibliography. T. C. Bartee, *Computer Architecture and Logic Design*, 1991; J. L. Hennessy and D. A. Patterson, *Computer Organization and Design*, 2d ed., 1997; K. Hwang, *Advanced Computer Architecture*, 1993; D. Tabak, *Advanced Microprocessors*, 2d ed., 1994; S. Weiss, *IBM POWER and PowerPC*, 1994.

Digital control

The use of digital or discrete technology to maintain conditions in operating systems as close as possible to desired values despite changes in the operating environment. Traditionally, control systems have utilized analog components, that is, controllers which generate time-continuous outputs (volts, pressure, and so forth) to manipulate process inputs and which operate on continuous signals from instrumentation measuring process variables (position, temperature, and so forth). In the 1970s, the use of discrete or logical control elements, such as fluidic components, and the use of programmable

logic controllers to automate machining, manufacturing, and production facilities became widespread. In parallel with these developments was the accelerating use of digital computers in industrial and commercial applications areas, both for logic-level control and for replacing analog control systems. The development of inexpensive mini- and microcomputers with arithmetic and logical capability orders of magnitude beyond that obtainable with analog and discrete digital control elements resulted in the rapid substitution of conventional control systems by digital computer–based ones. With the introduction of microcomputer-based control systems into major consumer products areas (such as automobiles and video and audio electronics), it became clear that the digital computer would be widely used to control objects ranging from small, personal appliances and games up to large, commercial manufacturing and production facilities. Hence the focus of this discussion will be on computer-based control systems. *See* MICROCOMPUTER; MICROPROCESSOR; PROGRAMMABLE CONTROLLERS.

Computer/process interface. The object that is controlled is usually called a device or, more inclusively, a process. A characteristic of any digital control system is the need for a process interface to mate the digital computer and process, to permit them to pass information back and forth (**Fig. 1**).

Digital control information. Measurements of the state of the process often are obtained naturally as one of two switch states; for example, a part to be machined is in position (or not), or a temperature is above (or below) the desired temperature. Control signals sent to the process often are expressed as one of two states as well; for example, a motor is turned on (or off), or a valve is opened (or closed). Such binary information can be communicated naturally to and from the computer, where it is manipulated in binary form. For this reason the binary or digital computer–process interface usually is quite simple: a set of signal-conditioning circuits for each measured or controlled signal and a set of registers to transfer the bits of digital information in each direction. Each register usually contains the same number of bits as would be manipulated and stored within the digital computer.

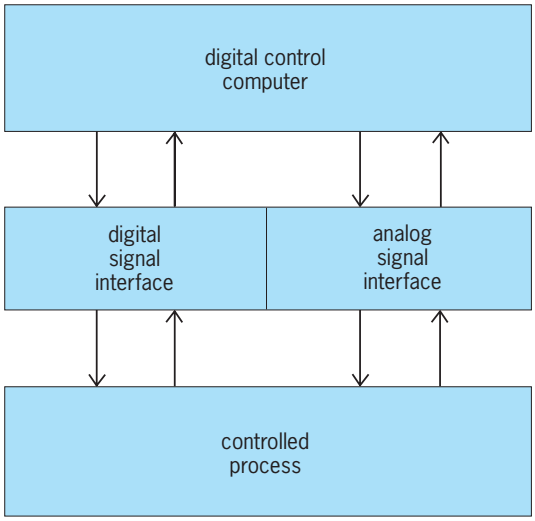

Fig. 1. Generalized digital computer control system.

Analog control information. Process information also must be dealt with in analog form; for example, a variable such as temperature can take on any value within its measured range or, looked at conceptually, it can be measured to any number of significant figures by a suitable instrument. Furthermore, analog variables generally change continuously in time. Digital computers are not suited to handle arbitrarily precise or continuously changing information; hence, analog process signals must be reduced to a digital representation (discretized), in terms of both magnitude and time, to put them into a useful digital form.

The magnitude discretization problem most often is handled by transducing and scaling each measured variable to a common range, then using a single conversion device—the analog-to-digital converter—to put the measured value into digital form. An analog-to-digital converter suitable for measurement and control purposes typically will convert signals in the range −10 to +10 V direct current, yielding an output with 12 to 16 bits of precision in 10 to 50 microseconds. A multiplexer often is used to allow a number of analog inputs to be switched into a single analog-to-digital converter. High-level signals (on the order of volts) can be switched by a solid-state multiplexer; low-level signals (on the order of millivolts, from strain gages or thermocouples) require mechanical relays followed by an amplifier to boost the signal to an acceptable input level for the analog-to-digital converter. Microcomputers are now manufactured that contain integral analog conversion circuitry for several channels on the processor chip. *See* ANALOG-TO-DIGITAL CONVERTER.

Discretization in time requires the computer to sample the signal periodically, storing the results in memory. This sequence of discrete values yields a "staircase" approximation to the original signal (**Fig. 2**), on which control of the process must be based. Obviously, the accuracy of the representation can be improved by sampling more often, and many digital systems simply have incorporated traditional analog control algorithms along with rapid sampling. However, newer control techniques make fundamental use of the discrete nature of computer input and output signals. Analog outputs from a computer most often are obtained from a digital-to-analog converter, a device which accepts a digital output from the computer, converts it to a voltage in several microseconds, and latches (holds) the value until the next output is converted. Usually a single digital-to-analog converter is used for each output signal. *See* DIGITAL-TO-ANALOG CONVERTER.

Real-time computing. In order to be used as the heart of a control system, a digital computer must be capable of operating in real time. Except for very simple microcomputer applications, this feature implies that the machine must be capable of handling interrupts, that is, inputs to the computer's internal

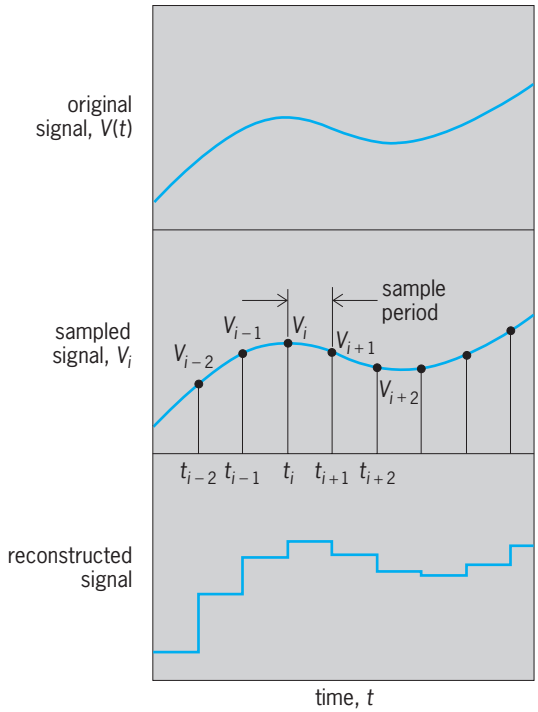

Fig. 2. Discretization in time of an analog signal.

control unit which, on change of state, cause the computer to stop executing some section of program code and begin executing some other section. Using its extraordinary computational abilities (on the order of a million instructions per second), the computer, which basically is a sequential or serial processor, can be made to appear to perform operations in parallel by proper design of its hardware and executive software. By attaching a very accurate oscillator (a so-called real-time clock) to the interrupt line, the computer can be programmed to keep track of the passage of real time and, consequently, to schedule process sampling and control calculations on a periodic basis.

The requirements of real-time computing also imply that the computer must respond to interrupts from the process. Thus a key process variable may be used to trigger an interrupt when it exceeds preset limits. The computer would be programmed to service such an interrupt immediately, taking whatever control action would be necessary to bring the variable back within limits. The ability to initiate operations on schedule and to respond to process interrupts in a timely fashion is the very basis of real-time computing; this feature must be available in any digital control system.

Programming considerations. Much of the programming of computer control systems is done in a high-level language such as BASIC, FORTRAN, C, or Ada; however, many microcomputer applications are carried out in machine (assembly) language. In either case the programmer must utilize program routines which access the devices in the process interface, for example, fetch the contents of a particular digital input register so as to check the status of some process digital element, or write out a digital result to a particular digital-to-analog converter channel. Additionally, the programmer will have the capability to schedule operations periodically or at particular times. *See* PROGRAMMING LANGUAGES.

Computer control systems for large or complex processes may involve complicated programs with many thousands of computer instructions. Several routes have been taken to mitigate the difficulty of programming control computers. One approach is to develop a single program which utilizes data supplied by the user to specify both the actions to be performed on the individual process elements and the schedule to be followed. Such an executive program, supplied by a control system vendor, might be utilized by a variety of users; and this approach often is taken for relatively standardized operations such as machining and sequential processing (manufacturing).

Another approach is to develop a rather sophisticated operating system to supervise the execution of user programs, scheduling individual program elements for execution as specified by the user or needed by the process. The multiple program elements, called tasks when used with a multitasking operating system, or called programs with a multiprogramming system, will be scheduled individually for execution by the operating system as specified by the user (programmer) or needed by the process. A simple example is given in **Fig. 3** for three tasks—a data logging task which might be scheduled every 30 s, a process alarm task which would be executed whenever a process variable exceeds limits and causes an interrupt, and an operator communication task which would be executed whenever the operator strikes a key on the console. The operating system scheduler would resolve conflicts, caused by two or three tasks needing to execute simultaneously, on the basis of user-supplied priorities; presumably the alarm task would have the highest priority here. *See* COMPUTER PROGRAMMING; SOFTWARE ENGINEERING.

Control algorithms. Many applications, particularly machining, manufacturing, and batch processing, involve large or complex operating schedules. Invariably, these can be broken down into simple logical sequences, for example, in a high-speed bottling operation the bottle must be in position before the filling line is opened. Hence the control program reduces to a set of interlocked sequential operations. Some applications—in the chemical process industries, in power generation, and in aerospace areas—require the use of traditional automatic control algorithms.

Automatic control algorithms fall into two major categories: feedforward techniques, where process disturbances which would affect the controlled variable are measured and their effect canceled out or compensated by appropriate manipulation of a process input variable; and feedback techniques. In feedback control, which makes up the vast majority of applications, the controlled variable itself is measured and subtracted from a reference (set-point) variable equal to its desired value. The resulting deviation is

the input to the controller, which then manipulates a process input variable through the final control element. If the controller is designed correctly, it should maintain the deviation quite small despite operator changes to the set point or despite environmental changes which cause disturbances to enter the control loop (**Fig. 4*a***).

Digital implementation of traditional analog algorithms. Traditional algorithms for analog control of processes have been developed over many years. One very important example is the three-mode controller, so called because the algorithm for the controller output U operates proportionally on the deviation ϵ, the time integral of ϵ, and the time derivative of ϵ as indicated in Eq. (1), where U_0 is the constant value that the

$$U = U_0 + K_c\left(\epsilon + \frac{1}{\tau_I}\int_0^t \epsilon dt + \tau_D\frac{d\epsilon}{dt}\right) \qquad (1)$$

controller output assumes when the controller first is turned on, and K_c, τ_I, and τ_D are the controller proportional gain, and integral and derivative time parameters, respectively. These controller parameters are chosen by the system designer appropriately for each control application.

In the digital implementation of a single-loop feedback control system (Fig. 4*b*), the digital computer implements the controller algorithm by using process information sampled through the analog-to-digital converter and a set point supplied by the operator. The symbolic switches and the subscript on the controller input ϵ_k (obtained from R_k and M_k) in Fig. 4*b* are intended to indicate that the information is obtained by the computer only at discrete sampling times, T time units apart. Similarly, the controller generates an output U_k at each sampling time which it transmits through the digital-to-analog converter to the process's final control element. Implementation of the three-mode algorithm in the digital (discrete) environment is straightforward. In the "velocity form" of the proportional-integral-derivative algorithm, so called because the change in U from its previous value is calculated directly [Eq. (2)], the

$$U_k = U_{k-1} + K_c\left[(\epsilon_k - \epsilon_{k-1}) + \frac{T}{\tau_I}\epsilon_k + \frac{\tau_D}{T}(\epsilon_k - 2\epsilon_{k-1} + \epsilon_{k-2})\right] \qquad (2)$$

computer need only store the controller output and two previous values of ϵ in order to calculate U at the next sampling time. In actual application, digital versions of continuous control algorithms often utilize rapid sampling (T is small) so that they respond in much the same way as analog systems do; hence the operational advantages and disadvantages of the two methods are equivalent.

Purely digital algorithms. One disadvantage of the classical feedback algorithm of Eq. (1) or (2) is that it does not function well when the dynamics of the process contain significant time delay, that is, when a change in the process input has no effect on the controlled variable for a period of time known as the dead time. So-called time-delay compensation algorithms have been available in analog form; however, only with the digital computer as controller has the implementation become so inexpensive that such applications can be widely justified. Hence the use of the three-mode algorithm in conjunction with a time-delay compensator is a common digital control approach, particularly for control of

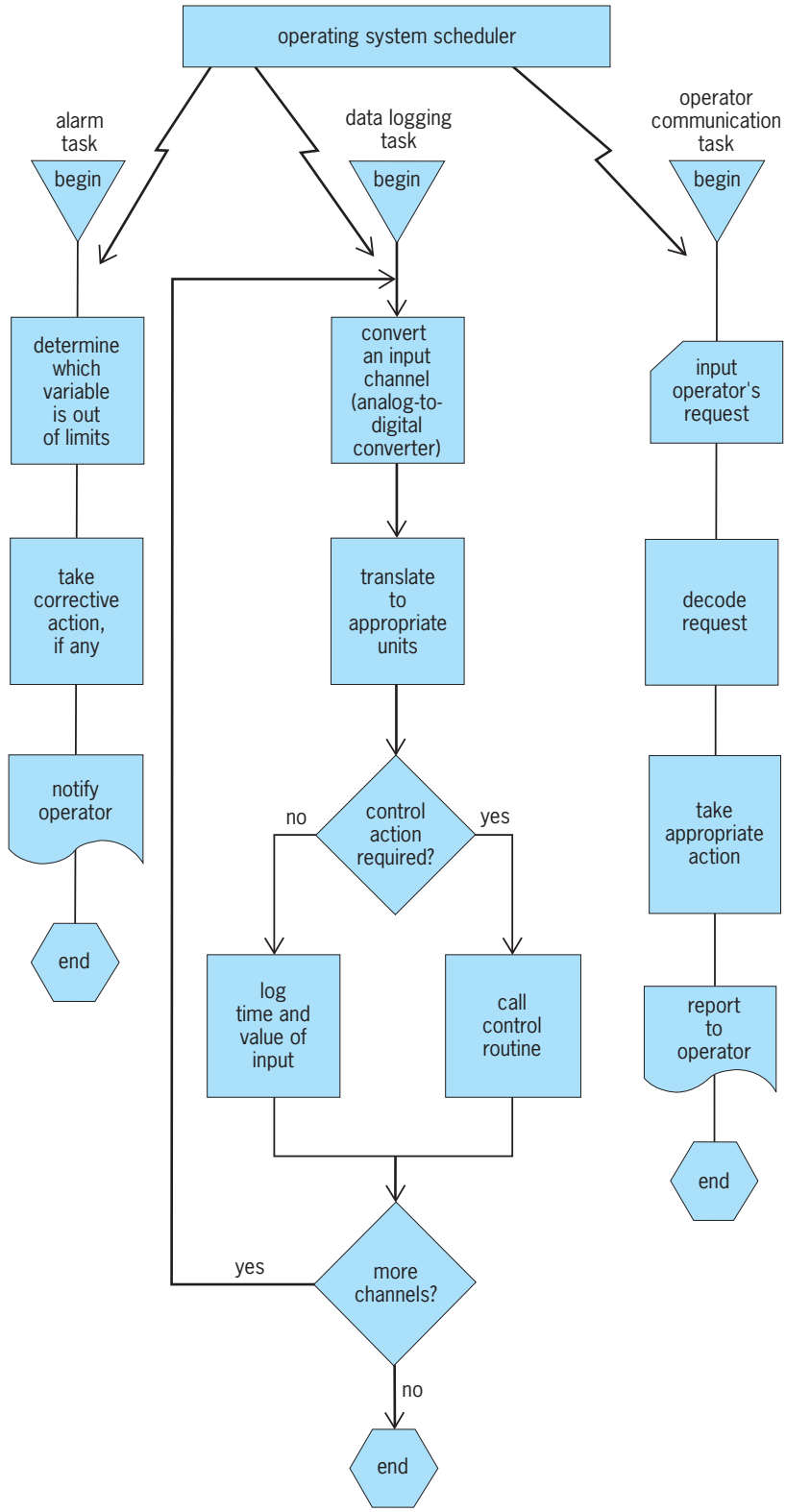

Fig. 3. Example of a multitasked control program.

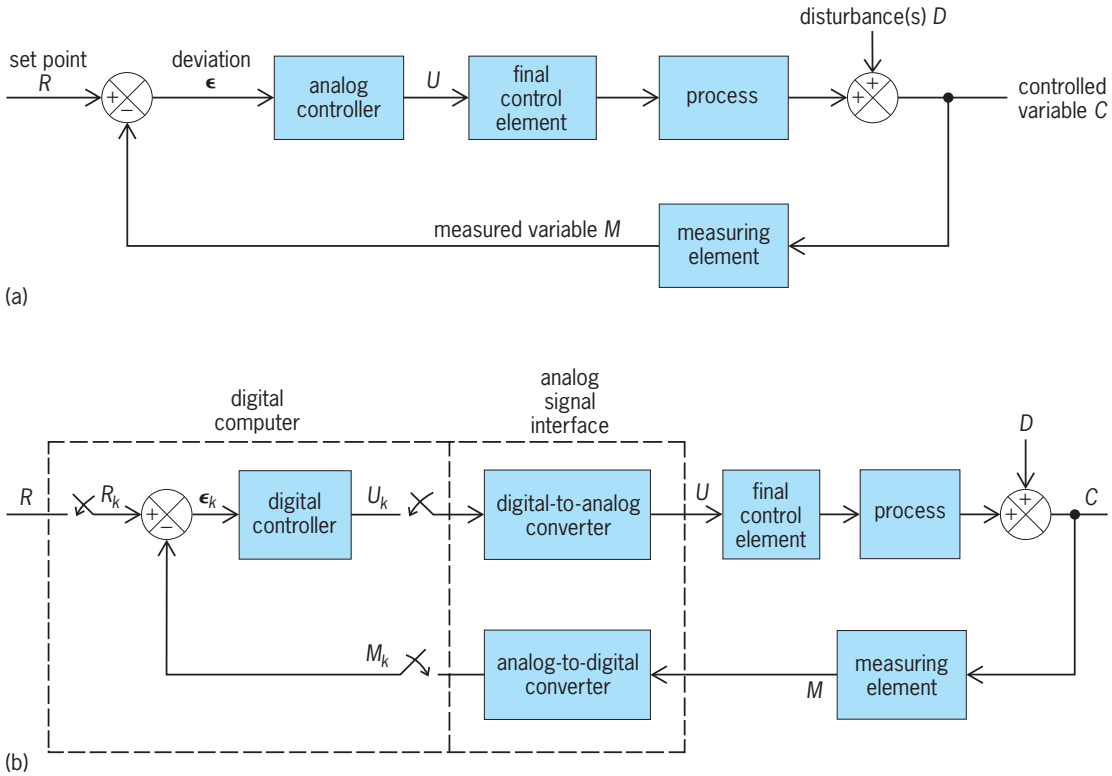

Fig. 4. Single-loop feedback control system. (*a*) Traditional analog implementation. (*b*) Digital implementation.

processes (where time delay is an important component of many, if not most, systems).

Attempts to expand the digital control medium through development of strictly digital control algorithms is an important and continuing trend. Such algorithms typically attempt to exploit the sampled nature of process inputs and outputs, significantly decreasing the sampling requirements of the algorithm (that is, making T as large as possible). For example, in the "deadbeat" algorithm, Eq. (3), the controller attempts to follow set-point changes in minimum time. The use of this algorithm presupposes that the process can be modeled approximately as a first-order plus time-delay system, where K and τ are the process gain and time constant and NT is the time delay. Hence a stepwise change in set point will cause the controlled variable to obtain its new, desired value in $N + 1$ sampling instants, the absolute minimum.

$$U_k = U_{k-1} + \frac{1}{K(N+1)(1-e^{-T/\tau})}(\epsilon_k - e^{-T/\tau}\,\epsilon_{k-1}) \tag{3}$$

The deadbeat algorithm has a number of disadvantages, including its sensitivity to errors in the assumed process model. These often result in undesirable oscillations (ringing). Dahlin's algorithm, Eq. (4), is one of a number of attempts to eliminate the undesirable features of optimal, deadbeat control. In this instance, the use of an empirical tuning parameter, λ, relaxes the requirement of minimal response time, resulting in less ringing in the system response, particularly when the process dynamics change slightly from those assumed in the first-order model.

$$U_k = U_{k-1} + \frac{1}{\left\{\begin{array}{c} K[N(1-e^{\lambda T})+1] \\ \times\,(1-e^{-T/\tau})(\epsilon_k - e^{-T/\tau}\epsilon_{k-1}) \end{array}\right\}} \tag{4}$$

Dahlin's and the deadbeat algorithms, and other similar algorithms which are designed on the basis of set-point response, often respond more poorly to disturbances than the three-mode algorithm. Hence there is a continuing search for the "best" algorithm—one which is simple, responds well to both set-point changes and disturbances, requires little knowledge of process dynamics and is insensitive to changes in process dynamics, tolerates process constraints and nonlinearities, can be extended to multi-input–multioutput situations, is easy to implement and program, and so forth. Such a universal digital control algorithm likely will not be found, but the search has led in a number of interesting directions, including the development of self-tuning algorithms and the application of time convolution techniques for predictive control strategies. These approaches would not be possible without the digital computer and the ability to design and implement computer control systems. Additional demands are made as well in the analysis of digital control systems: discrete algebraic (difference) equations substitute for differential equations used traditionally to describe processes and analog controllers; and discrete transform techniques (the Z transform) often replace the Laplace transform approach used in

analyzing continuous systems. *See* ALGORITHM; CONTROL SYSTEMS; DIGITAL COMPUTER; INTEGRAL TRANSFORM.

Duncan A. Mellichamp

Bibliography. W. Forsythe and R. Goodall, *Digital Control: Fundamentals, Theory, and Practice*, 1992; C. H. Houpis and G. B. Lamont, *Digital Control Systems: Theory, Hardware, and Software*, 2d ed., 1992; B. C. Kuo, *Digital Control Systems*, 2d ed., 1995; R. Leigh, *Applied Digital Control*, 2d ed., 1992; R. J. Vaccaro, *Digital Control: A State-Space Approach*, 1995; D. Williamson, *Digital Control and Implementation*, 1992.

Digital counter

A device which, in its simplest form, provides an output that corresponds to the number of pulses applied to its input.

Types. Counters may take many forms: geared mechanisms (water meters, odometers), relays (early telephone switch gear), vacuum tubes (early test instruments), and solid-state semiconductor circuits. This article will stress simple solid-state electronic counter circuits.

Most digital counters operate in the binary number system, since such a system is easily implemented with electronic circuitry. The number of input pulses is represented as a series of binary digits, or bits, where each bit is either a 0, or a 1 (a voltage of 0 or 5 V, usually). *See* NUMBERING SYSTEMS.

Figure 1 shows a four-bit binary counter that will count from 0 to 15; the sixteenth count input causes the counter to return to the 0 state and generate a carry pulse. This action to return to the 0 count with a carry output every sixteenth pulse makes the counter a modulo 16 counter. The four binary outputs Q_D, Q_C, Q_B, and Q_A represent the states of four flip-flops irrespective of whether the count input is 0 or 1. The counter is, therefore, a four-state-variable Moore state machine. Outputs Q_D, Q_C, Q_B, and Q_A are said to to have an 8-4-2-1 "weighting," because if Q_D through Q_A are all 1's, the binary counter output is $1111_2 = 1 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 8 + 4 + 2 + 1 = 15_{10}$, where the subscripts indicate the base of the number system. In Fig. 1*a*, the counter state-flow diagram is shown. The word "state" refers to the number, the datum, in the counter; this datum is stored in four flip-flop devices, each of which is a memory bit that stores the current value of one binary digit. Each possible state is represented by the numerical output of that state inside a circle. Upon receiving a count pulse, the counter must change state by following an arrow from the present state to the next state. In Fig. 1*b*, a table shows the counter output after a given number of input pulses, assuming that the counter always starts from the 0 state. The counter output is listed in binary, octal, decimal, and hexadecimal. Figure 1*c* shows a block diagram of the counter built with *T* flip-flops, and Fig. 1*d* shows the counter waveforms through time, with a periodic count input. The *T* flip-flop is a device that has either a 0 or a 1 on its *Q* output at all times. When the

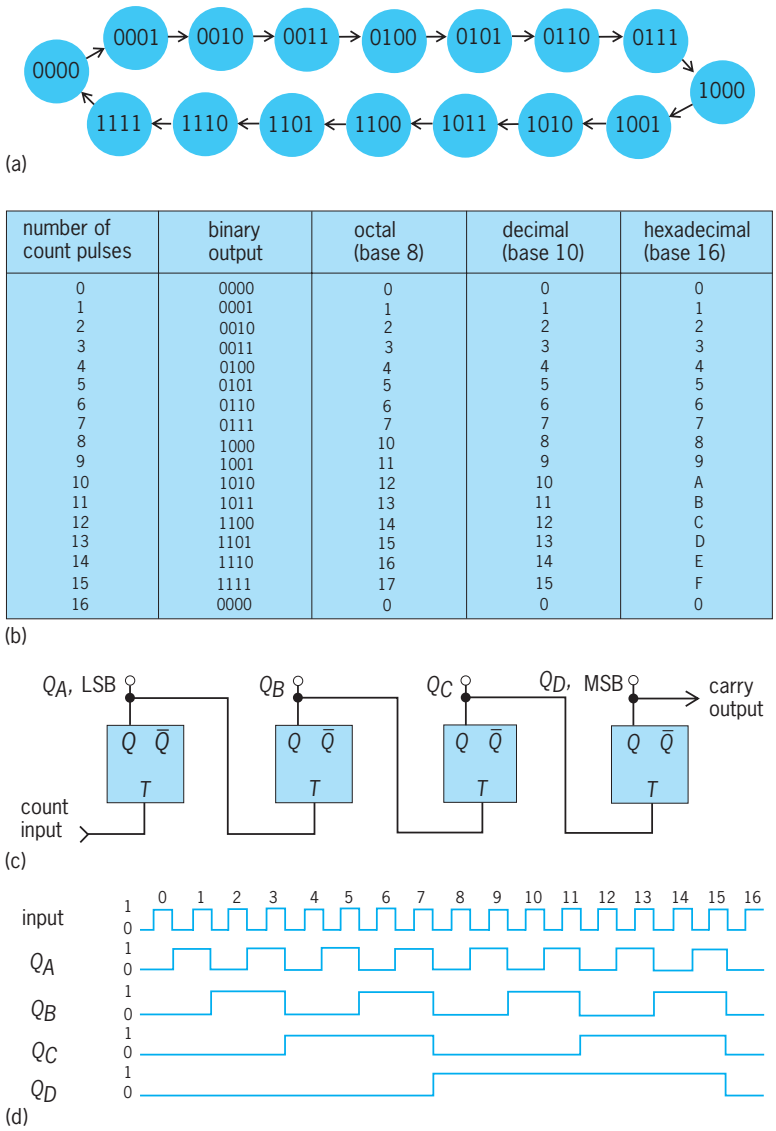

number of count pulses	binary output	octal (base 8)	decimal (base 10)	hexadecimal (base 16)
0	0000	0	0	0
1	0001	1	1	1
2	0010	2	2	2
3	0011	3	3	3
4	0100	4	4	4
5	0101	5	5	5
6	0110	6	6	6
7	0111	7	7	7
8	1000	10	8	8
9	1001	11	9	9
10	1010	12	10	A
11	1011	13	11	B
12	1100	14	12	C
13	1101	15	13	D
14	1110	16	14	E
15	1111	17	15	F
16	0000	0	0	0

Fig. 1. Four-bit binary counter using trigger flip-flops; this is a Moore machine. (*a*) State-flow diagram. (*b*) Table of the counter output in various number systems. (*c*) Circuit block diagram. LSB = least significant bit; MSB = most significant bit. (*d*) Output waveforms.

count input *T* moves from the 1 state to the 0 state, the flip-flop output must change state, from a 0 to a 1 or from a 1 to a 0. The carry output produces a 1-to-0 transition on every sixteenth count input, and therefore a divide-by-16 function.

The four bits of the counter of Fig. 1 can be grouped together and used to represent a single hexadecimal digit; in Fig. 1*b*, each counter output state represents one hexadecimal digit. A two-digit hexadecimal counter requires two sets of four-bit binary counters, the carry output from the first set of counters driving the count input of the second set of counters.

A decimal counter built from four binary counters is shown in **Fig. 2**. Let four bits of data from the binary counter represent one decimal digit. The counter will work in the same way as the counter of Fig. 1, except that all the flip-flops are reset to the 0 state when the counter moves from the $1001_2 = 9$

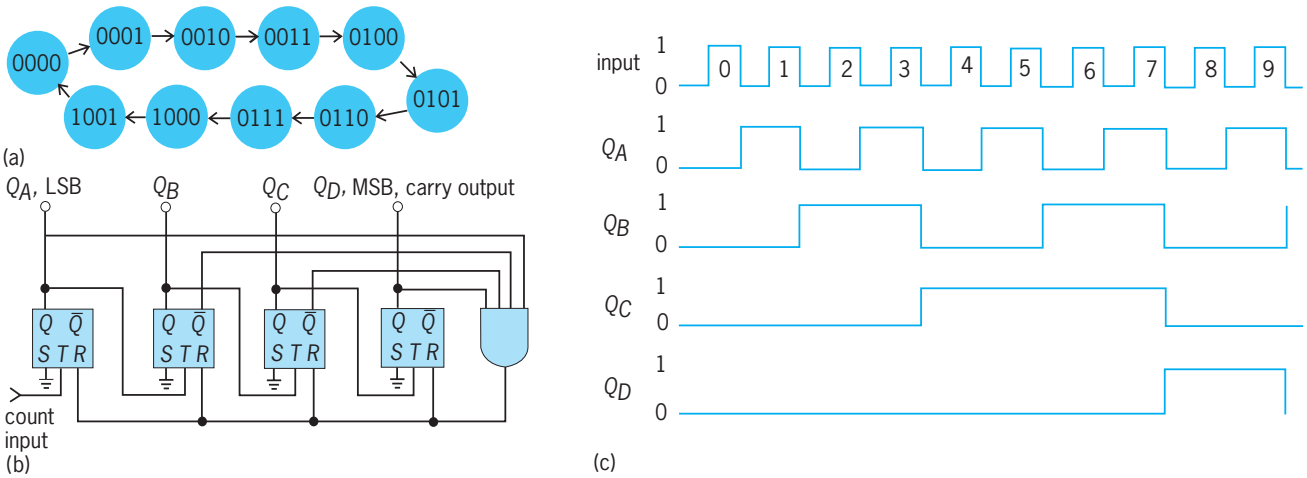

Fig. 2. Four-bit binary counter modified to be a decimal counter. (*a*) State-flow diagram. (*b*) Circuit block diagram. LSB = least significant bit; MSB = most significant bit. (*c*) Output waveforms.

state, instead of advancing to the $1010_2 = 10$ state. Besides the AND gate that is now used to detect the 1001 state of the counter and enable the resets, the circuit block diagram shows a new type of flip-flop. The *"SR"* flip-flop acts like a *T* flip-flop with an additional input that forces the *Q* output to a 1 state when the *S* (set) input is high and the *T* input has had a 1-to-0 transition applied. An *R* (reset) input acts as the *S* input does, except that the *Q* output goes to 0. This example decimal counter has an 8-4-2-1 weighted output that is known as binary coded decimal (B.C.D.). A seven-segment display is easily interfaced to the binary-coded-decimal counter using binary-coded-decimal-to-seven-segment decoder/driver circuits that are widely available.

Applications. Digital counters are found in much modern electronic equipment, especially equipment that is digitally controlled or has digital numeric displays. A frequency counter, as a test instrument or a channel frequency display on a radio tuner, consists simply of a string of decade counters that count the pulses of an input signal for a known period of time, and display that count on a seven-segment display. A digital voltmeter operates by using nearly the same idea, except that the counter counts a known frequency for a period of time proportional to the input voltage. *See* ANALOG-TO-DIGITAL CONVERTER.

A digital watch contains numerous counter/dividers in its large-scale integration (LSI) chip, usually implemented with complementary metal oxide semiconductor (CMOS) technology. A typical watch generates a 32,768-Hz crystal oscillator signal that is divided by 2^{15} (15 *T* flip-flop binary counters) to produce a 1-Hz signal that is counted as seconds by using two decimal counters that reset and produce a carry at every sixtieth count. That carry output is counted as a minutes display, generating another modulo 60 carry pulse for the hours display. The hours counters produce a carry at every twenty-fourth count to run the days counter and display; that counter generates a carry pulse every 28 to 31 days, depending on the month. *See* INTEGRATED CIRCUITS; WATCH.

Digital computers may contain counters in the form of programmable interval timers or may be programmed to operate as a counter. *See* DIGITAL COMPUTER.

Counter specifications. Counters have progressed from relays to light-wavelength-geometry very large-scale-integrated circuits. There are several technologies for building individual digital counters. Single counters are available as integrated circuit chips in emitter-coupled logic (ECL), transistor-transistor logic (TTL), and CMOS. The three technologies are listed in the order of decreasing speed and decreasing power dissipation. Commercially available integrated circuits will operate at frequencies up to 4 GHz, with experimental Josephson junction counters capable of 10 times faster operation. *See* LOGIC CIRCUITS; SUPERCONDUCTING DEVICES.

Michael E. Wright

Bibliography. L. E. Frenzel, *Digital Counter Handbook*, 1981; C. H. Roth, *Fundamentals of Logic Design*, 4th ed., 1992.

Digital evidence

Text, numerals, sound, or images (both moving and still) that are either stored or transmitted in a binary form and have potential use as evidence to prove a circumstance or event in a court of law. Binary refers to the electronic code made up of ones and zeroes that defines all the digital information. A unit of information is assigned one of two states, either 1 or 0 (that is, on and off). Thes two digits used in groups of 8 with varying sequences, for example, define the alphabet. *See* CRIMINALISTICS.

At a crime scene, investigators look for clues and motives among many items, including papers and documents such as daily appointment calendars, address books, letters, business records, checkbooks, and notes. Today, investigators are likely to find this information recorded on electronic media. Just like

paper records, digital information found on computers, personal organizers, answering machines, and cellular phones represents potentially latent (hidden) evidence that is important in investigations. For example, latent fingerprints are made visible after chemically treating surfaces. Likewise, latent evidence on computers and other digital devices is identified using computer hardware and software. Digital evidence on a cell phone may include an address book with names and telephone numbers, as well as a log of calls received, missed, made, and the times of these calls. On a personal computer, the log of Internet sites visited and e-mail sent and received might be important evidence in an investigation. *See* DATA COMMUNICATIONS; DIGITAL COMPUTER; ELECTRONIC MAIL; INTERNET; SOFTWARE.

Evidence collection. Computer forensics has come to mean the identification, preservation, extraction, documentation, and interpretation of digital evidence on computers or other electronic media for evidentiary purposes.

Forensic scientists are trained to consider evidence from both a scientific and legal perspective. When a scene is processed for evidence, several legal principles must be followed. For instance, the potential evidence must be identified, collected, transported, examined, and stored in a manner that does not change it. At the time the evidence is collected, it must be marked to identify who collected it, the date and time of collection, and where it was collected to preserve the chain of custody for acceptance in a court of law.

The law requires that the "best evidence" be used for legal purposes. In the case of a paper document, the original is preferred, although a copy can be used if the original is not available. Exact copies of digital evidence can be made from the original evidence. At a crime scene, the data on hard drives, disks, or other electronic media are copied and examined to find any evidence they might contain. The Scientific Working Group on Digital Evidence (SWGDE) has drafted the following definitions to describe the various attributes of digital evidence. *See* COMPUTER STORAGE TECHNOLOGY.

Acquisition of digital evidence: the process that begins when information or physical items are collected or stored for examination purposes. Applying the term "evidence" to these items implies that the manner in which they were collected is recognized by the courts; that is, the collection process was appropriate according to the rules of evidence in a given locality. A data object or physical item becomes evidence only when so deemed by a law enforcement official or designee.

Data objects: objects or information associated with physical items. Data objects may occur in different formats without altering the original information. Data objects include logical files, slack (unused) space, or perhaps a physical sector located on the physical item.

Physical item: an item on which data objects or information is stored or through which data objects are transferred.

Original digital evidence: physical items and the data objects associated with such items at the time of acquisition or seizure.

Duplicate digital evidence: an accurate digital reproduction of all data objects contained on an original physical item.

Evidence examination. Generally, the copy of a suspect's hard drive taken at the crime scene is the best evidence, and is stored in the evidence room as "the copy." Duplicates are made from this copy for forensic scientists and law enforcement to work from, with all work done on a dedicated forensic computer that is not connected to the Internet. This computer workstation is dedicated for forensic examination and is used in a controlled environment. The task for the forensic examiner is to locate documents or phrases that will assist the criminal investigator. For example, if a bomb was sent through the mail using a computer-printed address label, the copy of the suspect's hard drive would be examined for evidence of the address. In this case, the forensic examiner would do a keyword search and look for the words in the address. Forensic tools are software specifically designed for forensic examinations such as imaging (copying the hard drive) and restoring deleted files. Comprehensive suites of software tools are used for most forensic examinations.

Hidden data. Digital evidence may be obvious to an investigator or hidden. For example, obvious evidence might be a letter stored as a document on a hard drive, while less obvious evidence might be a word or fragment of a word that was deleted from a document but still exists on the unused space of the hard drive. Special software tools, some public and others restricted, assist forensic examiners to find this evidence. There are also forensic tools for determining when data have been deliberately hidden.

Encryption. Encryption is one way of actively hiding data. In encryption, data are encoded using an algorithm (mathematical formula), obscuring it from anyone who does not have a "secret key" to unscramble it. Once data have been encrypted, a single- or double-key decryption program is needed to read it, or else the code must be "cracked." Single-key decryption is similar to using one key to lock and unlock your home. Double-key decryption requires both a public key and a private key, where one or more people may have the public key, but only one person has the private key. Code cracking ranges from a very easy process to one that is computationally infeasible depending on the sophistication of the encryption. *See* ALGORITHM; CRYPTOGRAPHY.

Steganography. Hiding data in text, image, or audio files is called steganography. For example, a text file can be hidden in a digital picture. There are statistical analysis techniques to find such data, but they take time and effort. Commercial tools are available to detect these types of files.

Evidence authentication. A mathematical method known as hashing is used to verify the authenticity of data copied from a hard drive, disk, or file. A hash is a small mathematical summary of a data file or message. Changing one character in a file will change

the numerical value of the hash. If there has been any change in the file during transmission or from tampering, the hash number will change. A hash of the hard drive is done before copying it. The hash value of a hard drive and the copy should be exactly the same number.

To ensure chain of custody, forensic scientists must sign their name to the evidence. In the case of digital evidence, an authentic digital signature or a regular physical signature is required. Creating a digital signature is a simple process for verifying the integrity and originator of a message, whereby the signer calculates the hash value of the message, and then encrypts the hash value with his or her own private key. The resultant message "digest" is the digital signature. Appending it to the original message qualifies the object as being digitally signed.

As technology changes, the nature and handling of digital evidence will remain a challenge for forensic scientists. The wireless Web will allow for "virtual evidence rooms" where evidence is stored and secured through encryption. The chain of custody will be maintained through digital signatures, and the evidence will be checked out of the virtual evidence room with a time stamp from a digital clock synchronized with the atomic clock run by the National Institute of Standards and Technology (NIST). Carrie Morgan Whitcomb

Bibliography. E. Casey, *Digital Evidence and Computer Crime*, Academic Press, 2000; W. G. Kruse II and J. G. Heiser, *Computer Forensics: Incident Response Essentials*, Addison-Wesley, 2001; M. Pollitt, *Digital Evidence Standards and Principles*, Forensic Science Communication, 2000; T. Sammes, B. Jenkinson, and A. J. Sammes, *Forensic Computing: A Practitioner's Guide*, Springer, 2000; R. E. Smith, *Authentication: From Passwords to Public Keys*, Addison-Wesley, 2002.

Digital filter

Any digital computing means that accepts as its input a set of one or more digital signals from which it generates as its output a second set of digital signals. While being strictly correct, this definition is too broad to be of any practical use, but it does demonstrate the possible extent of application of digital-filter concepts and terminology.

Capabilities. Digital filters can be used in any signal-manipulating application where analog or continuous filters can be used. The classic applications are low-pass, high-pass, band-pass, band-stop, and narrow-band notch filters. Classic filter forms are Butterworth, Chebyshev, inverse Chebyshev, elliptic, and Bessel. Digital filters can provide these response characteristics and much more. They can be used for such applications as emphasis or attenuation of specified frequencies for spectrum shaping and equalization, for time-domain pulse shaping in data transmission, for signal detection by matched filtering, and for noise or interference removal. Because of their utterly predictable performance, they can be used in exacting applications where analog filters fail because of time- or other parameter-dependent coefficient drift in continuous systems. Because of the ease and precision of setting the filter coefficients, adaptive and learning digital filters are comparatively simple and particularly effective to implement. In many applications, digital filters provide the least costly solutions. As digital technology becomes more ubiquitous, digital filters are increasingly acknowledged as the most versatile and cost-effective solutions to filtering problems.

The number of functions that can be performed by a digital filter far exceeds that which can be performed by an analog, or continuous, filter. By controlling the accuracy of the calculations within the filter (that is, the arithmetic word length), it is possible to produce filters whose performance comes arbitrarily close to the performance expected of the perfect models. For example, theoretical designs that require perfect cancellation can be implemented with great fidelity by digital filters.

Although the independent variable of the input to the digital filter could be almost any parameter, for the sake of clarity and simplicity time will be chosen as the independent variable, unless otherwise specifically stated.

Digital signals. A digital signal, in the sense used here, is a signal that has been sampled or quantized in time (that is, it is a discrete-time signal, quantized in amplitude, and finally encoded, so that its value can be expressed in a finite set of characters, such as binary words or symbols (such as numbers) composed of bits (1's and 0's). Examples of signals are the set of temperature measurements taken at specified times in a chemical-processing vat, the outputs from navigation sensors (angle, velocity, or position as a function of time), the sensed microvoltage on the antenna in a communications system (which contains the voice message to be extracted), the operational engine parameters (temperatures, pressures, and flow rates as functions of time) in the space shuttle, and the video intermediate-frequency voltage in a television receiver. *See* ANALOG-TO-DIGITAL CONVERTER; PULSE MODULATION.

Two more simplifying assumptions can be made. Although a digital filter could accept many inputs and generate many outputs simultaneously, this discussion is restricted to the most common case, the so-called single-input, single-output (SISO) case, as opposed to the multivariable case. Furthermore, although the time variable could be sampled in an infinite number of ways, the sampling is assumed to be uniform (which is not necessary, but is the usual and simplest case), so that while $x(t)$ is a function of the continuous variable, t, only the equally spaced sample values $x(nT) = x_n$ are considered. The sampling period, $T = 1/f_s$, is fixed. The sampling frequency is f_s and is assumed to be constant; therefore, the much-preferred and simpler subscripted form (that is, x_n) of the variables is unambiguous.

Linear difference equation. The digital filter can be a program run on a personal computer, a circuit card containing many electronic components,

an application-specific integrated circuit, or a programmable microprocessor device. Although many embodiments are possible, they all fit the following description. The digital filter accepts as its input signals numerical values called input samples and produces as its output signal numerical values called output samples. The string of input samples is sometimes referred to as the input vector, and the string of output samples as the output vector. Each output sample at any particular sampling instant is a weighted sum of present and past input samples, and past output samples. If the sequence of input samples is x_n, x_{n+1}, x_{n+2}, . . ., then the corresponding sequence of output samples would be $y_n, y_{n+1}, y_{n+2}, \ldots$. The linear digital filter is uniquely characterized by its difference equation (1), which

$$y_n = a_0 x_n + a_1 x_{n-1} + a_2 x_{n-2} + \cdots + a_M x_{n-M} + b_1 y_{n-1} + b_2 y_{n-2} + \cdots + b_N y_{n-N} \quad (1)$$

tells exactly how to determine the value of each output sample. *See* DIFFERENCE EQUATION.

From this simple time-domain expression, a considerable number of definitions can be constructed. If the filter coefficients (the a's and the b's) are independent of the x's and y's, this digital filter is a linear filter. If the a's and b's are fixed, this is a linear time-invariant (LTI) filter. The order of the filter is given by the largest of the subscripts among the a's and b's, that is, the larger of M and N. If the b's are all zero (that is, if the output is the weighted sum of present and past input samples only), the digital filter is referred to as a nonrecursive (having no feedback) or finite impulse response (FIR) filter because the response of the filter to an impulse (actually a unit pulse) input is simply the sequence of the a coefficients. If any value of b is nonzero, the filter is recursive (having feedback) and is generally (although there are pathological exceptions) an infinite impulse response (IIR) filter. Sometimes the set of a coefficients is considered as the input- (sample-) weighting coefficient vector, and the set of b coefficients is considered as the output- (sample-) weighting-coefficient vector. Alternatively, the a and b coefficients can be strung together to form one large coefficient vector.

The difference equation of any linear, time-invariant, single-input, single-output digital filter looks the same. In order to design such a digital filter, all that remains is to determine the numerical values of the coefficient vector. The design, of course, flows from the functional requirements.

Although the time-domain difference equation is a useful description of a filter, as in the continuous-domain filter case, a powerful alternative form is the transfer function. The information content of the transfer function is the same as that of the difference equation as long as a linear, time-invariant system is under consideration. If the digital filter under consideration is not a linear, time-invariant filter, the transfer function cannot be used.

Transfer functions. A difference equation is converted to transfer-function form by use of the z transform. The z transform is simply the Laplace transform adapted for sampled systems with some shorthand notation introduced. The only operations on signals that are available in continuous, linear, time-invariant systems are addition (or subtraction), scaling (multiplication by a constant), and differentiation (or integration). Storage, or memory, in continuous systems is embodied through integration. The Laplace variable is s, and so is the Laplace operator for differentiation; the Laplace operator for integration is $1/s$. Similarly, the only allowable operations on signals in discrete-time, linear, time-invariant systems are addition (or subtraction), scaling (multiplication by a constant), and delay. Storage, or memory, in discrete-time systems is implemented directly. The Laplace operator for T-second storage is $\exp(-sT)$. *See* LAPLACE TRANSFORM; Z TRANSFORM.

The discrete-time version of the Laplace variable is simply the Laplace operator for storage. Because the notation $\exp(-sT)$ is so unwieldy, the shorthand time-shift variable $z = \exp(sT)$ is used, and has become a complex variable of interest in its own right. There is then an analogy: The continuous-time memory (storage) operator is $1/s$; the discrete-time memory (storage) operator is $1/z$. The z transform of x_n is $X(z)$; the z transform of x_n, scaled by α and delayed p samples (αx_{n-p}), is $\alpha X(z) z^{-p}$. The z transform of Eq. (1) is Eq. (2). The terms in this equation can be

$$\begin{aligned} Y(z) &= \sum_{m=0}^{M} a_m \cdot X(z) \cdot z^{-m} + \sum_{n=1}^{N} b_n \cdot Y(z) \cdot z^{-n} \\ &= X(z) \cdot \sum_{m=0}^{M} a_m \cdot z^{-m} + Y(z) \cdot \sum_{n=1}^{N} b_n \cdot z^{-n} \end{aligned} \quad (2)$$

rearranged to yield Eq. (3), and from this follows an expression for a rational polynomial transfer function, Eq. (4).

$$Y(z) \cdot \left[1 - \sum_{n=1}^{N} b_n \cdot z^{-n} \right] = X(z) \cdot \sum_{m=0}^{M} a_m \cdot z^{m} \quad (3)$$

$$\frac{Y(z)}{X(z)} = \frac{\sum_{m=0}^{M} a_m \cdot z^{m}}{1 - \sum_{n=1}^{N} b_n \cdot z^{-n}} \quad (4)$$

Since $z = \exp(sT)$, the complex (amplitude and phase) frequency response of the filter can be evaluated by setting $s = j\omega$, or $z = \exp(j\omega T)$. The design of the transfer function to meet a frequency-response specification is a well-established engineering procedure. *See* CONTROL SYSTEMS; LINEAR SYSTEM ANALYSIS.

FIR-filter design procedures. If $N = 0$ in Eqs. (1)–(4), an FIR filter results. The most common FIR-filter design approaches in computing the values of a_m are direct design using a procedure called the Remez-exchange algorithm (or the McClellan-Parks design procedure) and a Fourier-series design. Implementation of the filter is generally mechanized by either the equivalent of a tapped delay line (**Fig. 1***a*) or an injected delay line (Fig. 1*b*). The specific implementation choice is usually determined by characteristics

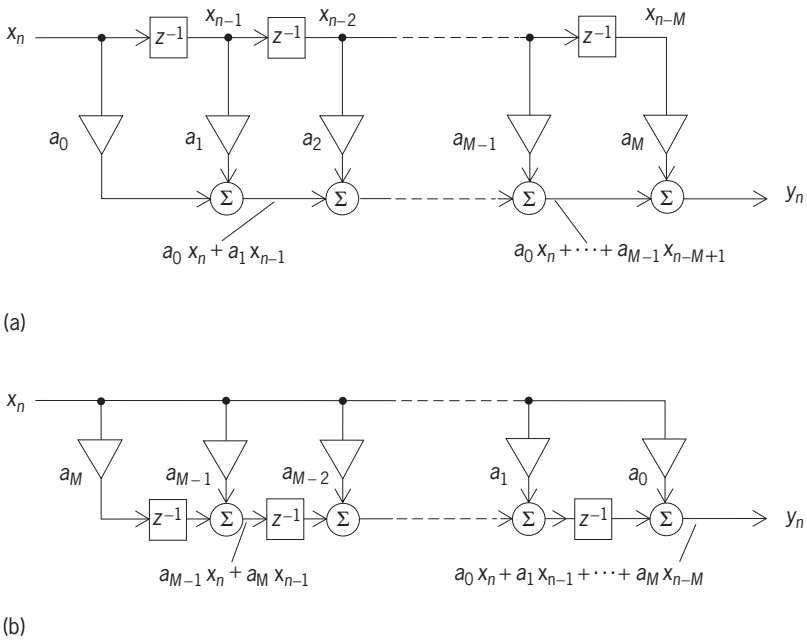

Fig. 1. Finite impulse response (FIR) filters. (*a*) Tapped delay-line filter. (*b*) Injected delay-line filter.

of the hardware used to build the filter. The delay-line approaches are replaced by fast convolution when the filter lengths become very long. Fast convolution consists of taking the Fourier transform of the input signal, multiplying the frequency-domain values by the Fourier transform of the desired filter response, then taking the inverse Fourier transform of the product. Fast convolution is very attractive and can have a mechanization-efficiency advantage over the delay-line approaches for FIR-filter lengths as short as $M = 15$. *See* FOURIER SERIES AND TRANSFORMS.

IIR-filter design procedures. For an IIR filter, a direct design approach such as one called the Deczky method can be used to determine the a_m and b_n values. Alternatively, one of the numerous cataloged continuous-time designs that have resulted from decades of work on continuous-time filters can be digitized by using one of the well-established transformation algorithms such as the bilinear-z transformation (usually the best), the matched-z transformation, or a transient-invariant z transformation (often better for certain control-systems applications). It should be emphasized that these are not z transforms; they are transformations.

Although obtaining the transfer function of the filter is (generally) unambiguously simple, the best mechanization of the IIR-filter transfer function is not correspondingly simple because the filters are mechanized with finite-resolution arithmetic, giving rise to three potential problems. The first problem is that the pole and zero locations are slightly in error because the coefficients of the transfer functions are inexact approximations to usually transcendental numbers. The second problem is that the multiplication and addition are performed with finite-precision accuracy, causing round-off quantization noise. The third problem is that the filter and signal must be scaled so that the amplitude of the signal can be adequately accommodated over the expected maximum and minimum values at all points within the filter.

These problems can be minimized by mechanizing the filter with the smallest-order sections possible. The numerator and denominator of the transfer function have real coefficients; therefore, each may

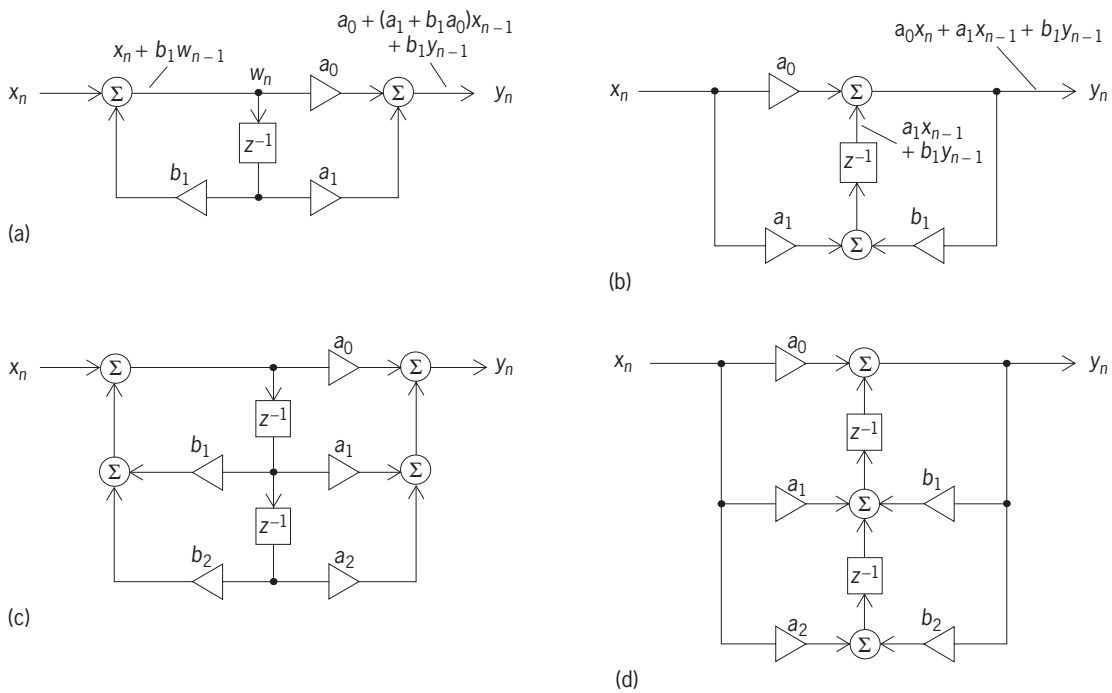

Fig. 2. Elemental building blocks of infinite impulse response (IIR) filters. (*a*) First-order building block; direct form. $w_n = b_1 w_{n-1} + x_n$, $y_n = a_0 w_n + a_1 w_{n-1}$. (*b*) First-order building block; transpose form. (*c*) Biquadratic building block; direct form. (*d*) Biquadratic building block; transpose form.

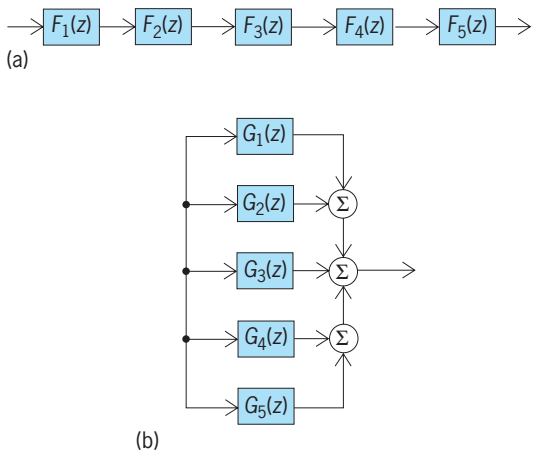

Fig. 3. Realizations of higher-order filters from elemental building blocks. (*a*) Cascade realization. (*b*) Partial-fraction-expansion realization.

be decomposed into a product of first-order factors with real roots (a numerator root is called a zero, a denominator root is called a pole) and second-order factors with complex-conjugate roots. The polynomial numerator and factored denominator may be further decomposed into alternate forms: a partial-fraction expansion, a cascade of first- and second-order building blocks, a parallel expansion of all-pass filters, a tapped cascade of all-pass filters, and the related families of lattice, ladder, and wave structures. Almost all practical linear, time-invariant, IIR digital filters are made of first- and second-order building blocks. *See* INTEGRATION.

Both first-order pole and zero IIR digital-filter building blocks and biquadratic (biquad) second-order building blocks may be realized in either direct or transpose configurations (**Fig. 2**). An intermediate state variable, w, is used as the storage variable in the direct form of the first-order building block (Fig. 2*a*). The building blocks in the partial-fraction realization of a high-order digital filter are similar to the first- and second-order building blocks in the cascade realizations (**Fig. 3**): the poles are the same as in the cascade realization, but the zeros are of order one less than the poles in each block.

All-pass filter sections are frequently used as building blocks (**Fig. 4**) because they have low coefficient sensitivity and are remarkably easy to scale. Their first- and second-order transfer functions are given by Eqs. (5). The magnitude-frequency responses of

$$H_1(z) = \frac{z^{-1} - b_1}{1 - b_1 \cdot z^{-1}}$$
$$H_2(z) = \frac{z^{-2} - b_1 \cdot z^{-1} + b_2}{1 - b_1 \cdot z^{-1} + b_2 \cdot z^{-2}} \tag{5}$$

all-pass filters are unity everywhere. These filter sections control phase alone.

Adaptive filters. So far only LTI filters have been discussed. An important class of variable-parameter filters change their coefficients to minimize an error criterion. These filters are called adaptive because they adapt their parameters in response to changes in the operating environment. An example is the FIR digital filter given by Eq. (6), where the a_m

$$y_n = \sum_{m=0}^{M} a_m \cdot x_{n-m} \tag{6}$$

are continually adjusted so that the output, y_n, will track a reference signal, r_n, with minimum error, ϵ. The performance criterion will be the minimization of some function of the error; that is, a function $J(\epsilon)$ will be minimized, where the rate of adjustment of each filter parameter is proportional to its contribution to $J(\epsilon)$, as in Eq. (7). This ad-

$$\frac{da_m}{dt} = -\mu \frac{\partial J(\varepsilon)}{\partial a_m} \tag{7}$$

justment method is known as the steepest-descent parameter-adjustment method. The convergence parameter is μ. If the convergence parameter is too

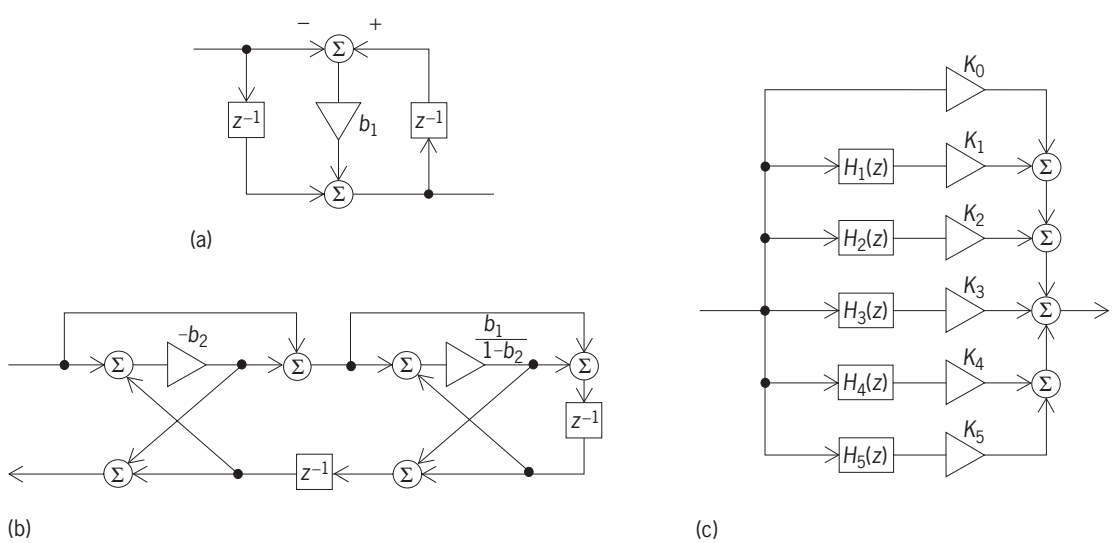

Fig. 4. Digital all-pass filters. (*a*) First-order filter. (*b*) Second-order lattice filter. (*c*) Higher-order filter assembled from a parallel bank of all-pass sections.

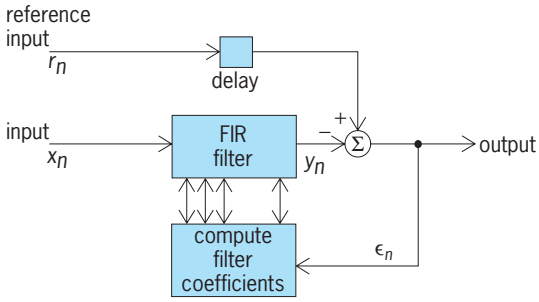

Fig. 5. Typical adaptive finite impulse response (FIR) filter.

small, convergence is very slow. If the convergence parameter is too large, the solution will oscillate around the correct value, never settling down. The error functional, $J(\epsilon)$, is generally of the form of Eq. (8), where κ is an integer. If $\kappa = 1$, the crite-

$$J(\varepsilon) = |\varepsilon|^{\kappa} \tag{8}$$

rion is least-absolute value; if $\kappa = 2$, the criterion is minimum-squared error (which is an unbiased estimate of the minimum mean-squared error, or MMSE, solution); and so forth (**Fig. 5**).

If the reference input to the filter (r_n) is signal plus additive noise and the "regular" input (x_n) is a linear-filtered version of the noise, then the adaptive filter will attempt to remove the noise from the output. This method is used in active noise-canceling systems. If both inputs are tied together, the filter will try to whiten the output; that is, the filter will attempt to remove correlated components such as echos and steady-state tones. By changing the effective time-constant of the adaption loops (making the integrators that generate the filter coefficients into low-pass filters by adding a low-gain "forgetting" loop), the correlation time is changed. Accordingly, the filter will emphasize (or deemphasize) signals with different correlation times, providing a powerful separation tool for signals such as speech. *See* ACTIVE SOUND CONTROL; SPEECH RECOGNITION.

There are also digital recursive adaptive filters (RAFs), which are used for more complicated problems. Although the mathematics is more involved because the error-criterion surface is multimodal (lumpy, rather than smooth), the basic idea is the same. *See* ADAPTIVE CONTROL; ELECTRIC FILTER; EQUALIZER.

Stanley A. White

Bibliography. S. K. Mitra and J. F. Kaiser (eds.), *Handbook for Digital Signal Processing*, 1993; A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing,* 2d ed., 1998; P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, 1993; B. Widrow and S. D. Stearns, *Adaptive Signal Processing,* 1985.

Digital-to-analog converter

A device for converting information in the form of combinations of discrete (usually binary) states or a signal to information in the form of the value or magnitude of some characteristics of a signal, in relation to a standard or reference. Most often, it is a device which has electrical inputs representing a parallel binary number, and an output in the form of voltage or current.

Figure 1 shows the structure of a typical digital-to-analog converter. The essential elements, found even in the simplest devices, are enclosed within the dashed rectangle. The digital inputs, labeled u_i, $i = 1, 2, \ldots, n$, are equal to 1 or 0. The output voltage E_o is given by Eq. (1), where V_{REF} is an analog reference

$$E_o = KV_{\text{REF}} \frac{u_1 2^{n-1} + u_2 2^{n-2} + \cdots + u_n 2^0}{2^n} \tag{1}$$

voltage and K is a constant. Thus, for a 5-bit binary converter with latched input code 10110, the output is given by Eq. (2).

$$\begin{aligned} E_o &= \left(\frac{16}{32} + \frac{0}{32} + \frac{4}{32} + \frac{2}{32} + \frac{0}{32}\right) KV_{\text{REF}} \\ &= \frac{11}{16} KV_{\text{REF}} \end{aligned} \tag{2}$$

Bit 1 is the "most significant bit" (MSB), with a weight of $\frac{1}{2}$; bit n is the "least significant bit" (LSB), with a weight of 2^{-n}. The number of bits n characterizes the resolution. [When the converter is used with a microprocessor, the bits are usually numbered from $n - 1$ (MSB) to 0 (LSB), that is, from 4 to 0 in the example of Eq. (2).] *See* ANALOG-TO-DIGITAL CONVERTER; NUMBERING SYSTEMS.

Uses. Digital-to-analog (D/A) converters (sometimes called DACs) are used to present the results of digital computation, storage, or transmission, typically for graphical display or for the control of devices that operate with continuously varying quantities. D/A converter circuits are also used in the design of analog-to-digital converters that employ feedback techniques, such as successive-approximation and counter-comparator types. In such applications, the D/A converter may not necessarily appear as a separately identifiable entity.

Circuitry. The fundamental circuit of most D/A converters involves a voltage or current reference; a resistive "ladder network" that derives weighted currents or voltages, usually as discrete fractions of the reference; and a set of switches, operated by the digital input, that determines which currents or voltages will be summed to constitute the output.

Input and output of converter in Fig. 2

Digital input code			Analog output	
u_1	u_2	u_3	$-E_o$, volts	$\frac{E_o}{V_{\text{REF}}}$
0	0	0	0	0
0	0	1	1.25	1/8
0	1	0	2.5	2/8
0	1	1	3.75	3/8
1	0	0	5.0	4/8
1	0	1	6.25	5/8
1	1	0	7.5	6/8
1	1	1	8.75	7/8

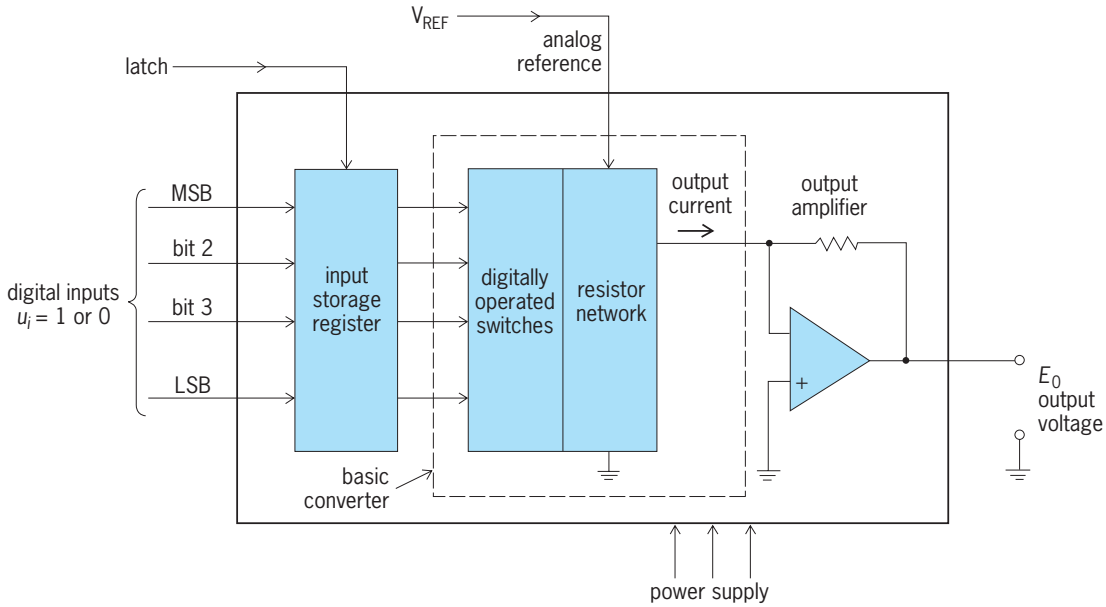

Fig. 1. Structure of a typical digital-to-analog converter.

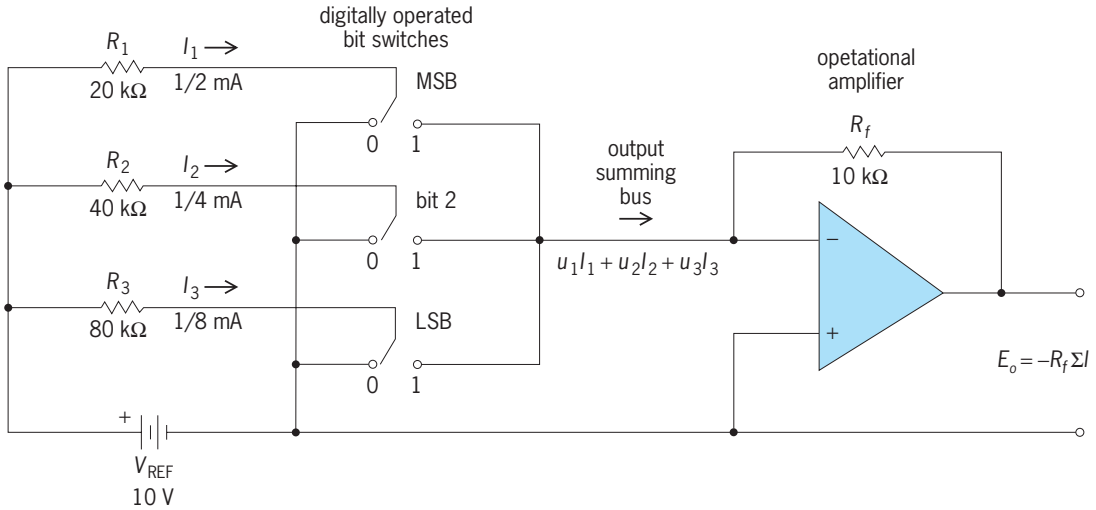

Fig. 2. Circuit of elementary 3-bit digital-to-analog converter.

An elementary 3-bit D/A converter is shown in **Fig. 2**. Binary-weighted currents developed in R_1, R_2, R_3 by V_{REF} are switched either directly to ground or to the output summing bus (which is held at zero volts by the operational-amplifier circuit). The sum of the currents develops an output voltage of polarity opposite to that of the reference across the feedback resistor R_f. The **table** shows the binary relationship between the input code and the output, both as a voltage and as a fraction of the reference.

The output of the D/A converter is proportional to the product of the digital input value and the reference. In many applications, the reference is fixed, and the output bears a fixed proportion to the digital input. In other applications, the reference, as well as the digital input, can vary; a D/A converter that is used in these applications is thus called a multiplying D/A converter. It is principally used for imparting a digitally controlled scale factor, or "gain," to an analog input signal applied at the reference terminal. *See* AMPLIFIER; ANALOG COMPUTER.

Construction. Except for the highest resolutions (beyond 18 bits), commercially available D/A converters are generally manufactured in the form of integrated circuits, using bipolar, MOS, and hybrid technologies. A single chip may include just the resistor network and switches; it may also include a reference circuit, output amplifier, and one or more sets of registers (with control logic suitable for direct microprocessor interfacing). *See* INTEGRATED CIRCUITS.

Daniel H. Sheingold

Bibliography. Analog Devices, Inc. engineering staff and D. Sheingold (ed.), *Analog-Digital Conversion Handbook*, 3d ed., 1997; D. J. Dooley (ed.), *Data Conversion Integrated Circuits*, 1980; D. F. Hoeschele, Jr., *Analog-to-Digital and Digital-to-Analog Conversion Techniques*, 2d ed., 1994.

Digitalis

A genus of the figwort family (Scrophulariaceae) ranging from the Canary Islands to central Asia. Foxglove (*Digitalis purpurea*), a native of western Europe, is the source of the important drug digitalis, much used in the treatment of heart disorders. Fresh, mature leaves are carefully selected, quickly dried, and stored in airtight containers. The active ingredient of digitalis is the glucoside digitalin. This slows and regulates the heartbeat, improving the tone and rhythm, and making the contractions more effective. The plant is also prized as a herbaceous perennial. *See* SCROPHULARIALES.

Perry D. Strausbaugh; Earl L. Core

Dill

An annual culinary herb of the carrot family, Apiaceae (Umbelliferae). It supplies four products: dill seed oil, dill leaf oil, dill seed, and dill leaf. Although all of the dill grown in the United States is one species, *Anethum graveolens,* a slightly different species, *A. sowa,* is grown in Asia. In the United States, two cultivars are grown, Mammoth and Bouquet. Bouquet is a tetraploid cultivar that was developed by treating Mammoth seed with colchicine and then selecting for a shorter (2–5 ft or 0.6–1.5 m) and more compact growth habit. Mammoth is a taller, more robust plant growing 4–8 ft (1.2–2.4 m) tall. Other varieties exist, such as those developed in India and Russia, but are not extensively grown here. Dill can be grown in most soil types in all areas of the United States, but most of the production is centered in Oregon, Washington, Florida, and California.

Flowering of dill usually occurs 60–80 days after planting, with hermaphrodite and later homogamous flowers being produced. Pollination is effected by insects—primarily bees and flies—but many other flying insects, including wasps, aid in fertilization. *See* POLLINATION.

Dill seed oil is a pale yellow oil which is steam-extracted just prior to seed shatter, about 105 to 120 days after planting. Dill grown for seed oil production is either planted in rows and furrow-irrigated or flat-planted and sprinkle-irrigated. Seeding rates vary from 4 to 20 lb/acre (3.5 to 18 kg/hectare). Most of this production provides oil used in the pickle industry for flavoring dill pickles. Dill leaf oil is little used, and relatively little dill is grown for this purpose. Dill seed is used as a spice, and large quantities of it are imported into the United States. India is the primary producer of dill seed for culinary use. *See* FAT AND OIL (FOOD).

Dill grown for leaf, often called dill weed, is cultivated in a slightly different manner. Seeding rates are 18–22 lb/acre (16–19 kg/ha), and it is planted in the fall or early spring to retard flowering. Dill leaf is harvested 50–70 days after planting. In the United States, dehydration is accomplished with forced-air dryers.

Weed control in commercial plantings is accomplished by use of herbicides, such as Treflan or Prometryn. Stoddards solvent or "carrot oil" at 40–60 gal/acre (380–570 liters/ha) has been extensively used in the past and is considered safe. *See* APIALES; SPICE AND FLAVORING.

Seth Kirby

Bibliography. S. Arctander, *Perfume and Flavor Materials of Natural Origin,* 1960; L. H. Bailey, *Manual of Cultivated Plants,* rev. ed. 1975; J. C. Uphof, *Dictionary of Economic Plants,* 1968.

Dilleniales

An order of flowering plants, division Magnoliophyta (Angiospermae), subclass Dilleniidae, class Magnoliopsida (dicotyledons). It consists of 21 families: the Dilleniaceae, with about 350 species; and the Paeoniaceae, with a single genus of some 30 species. Within its subclass the order is marked by separate carpels; numerous stamens; bitegmic (having two integuments), crassinucellate (having a nucellus that is more than one cell thick) ovules; and seeds that have a well-developed endosperm and are provided with an aril (a partial or complete covering or appendage, usually derived from the funiculus of the ovule). The garden peony *(Paeonia lactiflora)* is a familiar member of the order (see **illus.**). *See* DILLENIIDAE; MAGNOLIOPHYTA; PLANT KINGDOM.

The garden peony (*Paeonia lactiflora*), of the family Paeoniaceae. (*F. E. Westlake, National Audubon Society*)

Arthur Cronquist; T. M. Barkley

Dilleniidae

A large subclass of the class Magnoliopsida (dicotyledons) of the division Magnoliophyta (Angiospermae), the flowering plants, consisting of 13 orders, 78 families, and nearly 25,000 species. The subclass is morphologically ill-defined, but most of the orders and species have the carpels united to form

a compound pistil. The petals are either separate or joined into a sympetalous corolla, but the sympetalous members generally lack the advanced features of the subclass Asteridae. That is, they do not have a single set of stamens alternate with the corolla lobes, ovules with a massive single integument, or a reduced nucellus. The stamens, when numerous, are initiated in centrifugal sequence. Many of the species have numerous ovules on parietal placentas; that is, the placentas are borne along the walls of an ovary which is usually with a single chamber. Many have various chemical repellents and are often tanniferous. There are mustard oils or iridoid compounds in some orders, but mostly they are poor in alkaloids and without betalains.

The largest orders included in the subclass are the Violales (about 5000 species), Capparales (about 4000 species), Malvales (about 3000 species), Theales (about 3500 species), and Ericales (about 4000 species). Other orders include the Dilleniales, Lecythidales, Nepenthales, Salicales, Diapensiales, Ebenales, Primulales, and Batales. See separate articles on each order. *See* ASTERIDAE; MAGNOLIOPHYTA; PLANT KINGDOM. Arthur Cronquist; T. M. Barkley

Dimensional analysis

A technique that involves the study of dimensions of physical quantities. Dimensional analysis is used primarily as a tool for obtaining information about physical systems too complicated for full mathematical solutions to be feasible. It enables one to predict the behavior of large systems from a study of small-scale models. It affords a convenient means of checking mathematical equations. Finally, dimensional formulas provide a useful cataloging system for physical quantities.

Theory. All the commonly used systems of units in physical science have the property that the number representing the magnitude of any quantity (other than purely numerical ratios) varies inversely with the size of the unit chosen. Thus, if the length of a given piece of land is 300 ft, its length in yards is 100. The ratio of the magnitude of 1 yd to the magnitude of 1 ft is the same as that of any length in feet to the same length in yards, that is, 3. The ratio of two different lengths measured in yards is the same as the ratio of the same two lengths measured in feet, inches, miles, or any other length units. This universal property of unit systems, often known as the absolute significance of relative magnitude, determines the structure of all dimensional formulas. *See* UNITS OF MEASUREMENT.

In defining a system of units for a branch of science such as mechanics or electricity, certain quantities are chosen as fundamental and others as secondary, or derived. The choice of the fundamental units is always arbitrary and is usually made on the basis of convenience in maintaining standards. In mechanics the fundamental units most often chosen are mass, length, and time. Standards of mass (the standard kilogram) and of length (the standard meter) are readily manufactured and preserved, while the rotation of the Earth gives a sufficiently reproducible standard of time. Secondary quantities such as velocity, force, and momentum are obtained from the primary set of quantities according to a definite set of rules.

Assume that there are three primary, or fundamental, quantities α, β, and γ (the following discussion, however, is not limited to there being exactly three fundamental quantities). Consider a particular secondary quantity, expressed in terms of the primaries as $F(\alpha,\beta,\gamma)$, where F represents some mathematical function. For example, if α = mass, β = length, and γ = time, the derived quantity velocity would be $F(\alpha,\beta,\gamma) = \beta/\gamma$.

Now, if it is assumed that the sizes of the units measuring α, β, and γ are changed in the proportions $1/x$, $1/y$, and $1/z$, respectively, then the numbers measuring the primary quantities become $x\alpha$, $y\beta$, and $z\gamma$, and the secondary quantity in question becomes $F(x\alpha, y\beta, z\gamma)$. Merely changing the sizes of the units must not change the rule for obtaining a particular secondary quantity.

Consider two separate values of the secondary quantity $F(\alpha_1,\beta_1,\gamma_1)$ and $F(\alpha_2,\beta_2,\gamma_2)$. Then, according to the principle of the absolute significance of relative magnitude, Eqs. (1*a*) and (1*b*) hold.

$$\frac{F(x\alpha_1, y\beta_1, z\gamma_1)}{F(x\alpha_2, y\beta_2, z\gamma_2)} = \frac{F(\alpha_1, \beta_1, \gamma_1)}{F(\alpha_2, \beta_2, \gamma_2)} \tag{1a}$$

$$F(x\alpha_1, y\beta_1, z\gamma_1) = F(x\alpha_2, y\beta_2, z\gamma_2)\frac{F(\alpha_1, \beta_1, \gamma_1)}{F(\alpha_2, \beta_2, \gamma_2)} \tag{1b}$$

Differentiating partially with respect to x, holding y and z constant, gives Eq. (2), where F' represents

$$\alpha_1 F'(x\alpha_1, y\beta_1, z\gamma_1) = \alpha_2 F'(x\alpha_2, y\beta_2, z\gamma_2)\frac{F(\alpha_1, \beta_1, \gamma_1)}{F(\alpha_2, \beta_2, \gamma_2)} \tag{2}$$

the total derivative of the function F with respect to its first argument.

Next set the coefficients $x, y, z = 1$. This gives Eq. (3). This relation must hold for all values of the

$$\frac{\alpha_1 F'(\alpha_1, \beta_1, \gamma_1)}{F(\alpha_1, \beta_1, \gamma_1)} = \frac{\alpha_2 F'(\alpha_2, \beta_2, \gamma_2)}{F(\alpha_2, \beta_2, \gamma_2)} \tag{3}$$

arguments α, β, and γ and hence is equal to a constant. The subscripts can now be dropped, giving Eq. (4). The general solution of this differential equa-

$$\frac{\alpha}{F}\frac{dF}{d\alpha} = \text{constant} \tag{4}$$

tion is $F = C_1\alpha^a$, where a is a constant and C_1 is in general a function of β and γ.

The above analysis can now be repeated for the parameters β and γ, leading to the results given in Eqs. (5*a*) and (5*b*). These solutions are consistent

$$F = C_2(\alpha, \gamma)\beta^b \tag{5a}$$

$$F = C_3(\alpha, \beta)\gamma^c \tag{5b}$$

only if $F = C\alpha^a\beta^b\gamma^c$ where C, a, b, and c are constants. Thus every secondary quantity which satisfies

the condition of the absolute significance of relative magnitude is expressible as a product of powers of the primary quantities. Such an expression is known as the dimensional formula of the secondary quantity. There is no requirement that the exponents *a, b, c* be integral.

Examples of dimensional formulas. **Table 1** gives the dimensional formulas of a number of mechanical quantities in terms of their mass M, length L, and time T.

In order to extend this list to include the dimensional formulas of quantities from other branches of physics, such as electricity and magnetism, one may do either of the following:

1. Obtain the dimensional formulas in terms of a particular unit system without introducing any new fundamental quantities. Thus, if one uses Coulomb's law in the centimeter-gram-second electrostatic system in empty space (dielectric constant = 1), one obtains definitions (6*a*) and (6*b*).

$$\text{Force} = \frac{(\text{charge})^2}{(\text{distance})^2} \tag{6a}$$

$$\text{Charge} = \text{distance} \times \text{square root of force} \tag{6b}$$

Thus, the dimensions of charge are $M^{1/2}L^{3/2}T^{-1}$. The dimensional formulas of other electrical quantities in the electrostatic system follow directly from definition (7) and so forth.

$$\begin{aligned} \text{Potential} &= \text{energy/charge} = M^{1/2}L^{1/2}T^{-1} \\ \text{Current} &= \text{charge/time} = M^{1/2}L^{3/2}T^{-2} \\ \text{Electric field} &= \text{force/charge} = M^{1/2}L^{-1/2}T^{-1} \end{aligned} \tag{7}$$

2. Introduce an additional fundamental quantity to take account of the fact that electricity and magnetism encompass phenomena not treated in mechanics. All electrical quantities can be defined in terms of M, L, T, and one other, without resorting to fractional exponents and without the artificial assumption of unit (and dimensionless) dielectric constant as in the electrostatic system. The usual choice for the fourth fundamental quantity is charge A, even though charge is not a preservable electrical standard. Then definitions (8) and so forth hold.

$$\begin{aligned} \text{Potential} &= \text{energy/charge} = ML^2T^{-2}Q^{-1} \\ \text{Current} &= \text{charge/time} = QT^{-1} \end{aligned} \tag{8}$$

See ELECTRICAL UNITS AND STANDARDS.

It must be realized that the choice of fundamental quantities is entirely arbitrary. For example, a system of units for mechanics has been proposed in which the velocity of light is taken as a dimensionless quantity, equal to unity in free space. All velocities are then dimensionless. All mechanical quantities can then be specified in terms of just two fundamental quantities, mass and time. This is analogous to the reduction of the number of fundamental quantities needed for electricity from four to three by taking the dielectric constant (permittivity) of empty space as unity. *See* DIMENSIONLESS GROUPS.

Furthermore, one could increase the number of fundamental quantities in mechanics from three to four by adding force F to the list and rewriting Newton's second law as $F = Kma$, where K is a constant of dimensions $M^{-1}L^{-1}T^2F$. One could then define a system of units for which K was not numerically equal to 1.

In the past there has been considerable controversy as to the absolute significance, if any, of dimensional formulas. The significant fact is that dimensional formulas always consist of products of powers of the fundamental quantities.

Applications. The important uses of the technique of dimensional analysis are considered in the following sections.

Checking of equations. It is intuitively obvious that only terms whose dimensions are the same can be equated. The equation 10 kg = 10 m/s, for example, makes no sense. A necessary condition for the correctness of any equation is that the two sides have the same dimensions. This is often a help in the verification of complicated analytic expressions. Of course, an equation can be correct dimensionally and still be wrong by a purely numerical factor.

A corollary of this is that one can add or subtract only quantities which have the same dimensions (except for the trivial case which arises when two different equations are added together). Furthermore, the arguments of trigonometric and logarithmic functions must be dimensionless; otherwise their power series expansions would involve sums of terms with different dimensions. There is no restriction on the multiplication and division of terms whose dimensions are different.

π Theorem. The application of dimensional analysis to the derivation of unknown relations depends upon the concept of completeness of equations. An expression which remains formally true no matter how the sizes of the fundamental units are changed

TABLE 1. Dimensional formulas of common quantities

Quantity	Symbol	Dimensional formula
Mass	Fundamental	M
Length	Fundamental	L
Time	Fundamental	T
Velocity	Distance/time	LT^{-1}
Acceleration	Velocity/time	LT^{-2}
Force	Mass × acceleration	MLT^{-2}
Momentum	Mass × velocity	MLT^{-1}
Energy	Force × distance	ML^2T^{-2}
Angle	Arc/radius	0
Angular velocity	Angle/time	T^{-1}
Angular acceleration	Angular velocity/time	T^{-2}
Torque	Force × lever arm	ML^2T^{-2}
Angular momentum	Momentum × lever arm	ML^2T^{-1}
Moment of inertia	Mass × radius squared	ML^2
Area	Length squared	L^2
Volume	Length cubed	L^3
Density	Mass/volume	ML^{-3}
Pressure	Force/area	$ML^{-1}T^{-2}$
Action	Energy × time	ML^2T^{-1}
Viscosity	Force per unit area per unit velocity gradient	$ML^{-1}T^{-1}$

is said to be complete. If changing the units makes the expression wrong, it is incomplete. For a body starting from rest and falling freely under gravity, $s = 16t^2$ is a correct expression only so long as the distance fallen s is measured in feet and the time t in seconds. If s is in meters and t in minutes, the equation is wrong. Thus, $s = 16t^2$ is an incomplete equation. The constant in the equation depends upon the units chosen. To make the expression complete, the numerical factor must be replaced by a dimensional constant, the acceleration of gravity g. Then $s = 1/2gt^2$ is valid no matter how the units of length and time are changed, since the numerical value of g can be changed accordingly.

Assume a group of n physical quantities $x_1, x_2, \ldots, x_n$, for which there exists one and only one complete mathematical expression connecting them, namely, $\phi(x_1, x_2, \ldots, x_n) = 0$. Some of the quantities $x_1, x_2, \ldots, x_n$ may be constants. Assume further that the dimensional formulas of the n quantities are expressed in terms of m fundamental quantities $\alpha, \beta, \gamma, \ldots$. Then it will always be found that this single relation ϕ can be expressed in terms of some arbitrary function F of $n - m$ independent dimensionless products $\pi_1, \pi_2, \ldots, \pi_{n-m}$, made up from among the n variables, as in Eq. (9). This is known as the π theorem.

$$F(\pi_1, \pi_2, \ldots, \pi_{n-m}) = 0 \tag{9}$$

It was first rigorously proved by E. Buckingham. The proof is straightforward but long. The only restriction on the π's is that they be independent (no one expressible as products of powers of the others). The number m of fundamental quantities chosen in a particular case is immaterial so long as they also are dimensionally independent. Increasing m by 1 always increases the number of dimensional constants (and hence n) by 1, leaving $n - m$ the same.

The main usefulness of the π theorem is in the deduction of the form of unknown relations. The successful application of the theorem to a particular problem requires a certain amount of shrewd guesswork as to which variables $x_1, x_2, \ldots, x_n$ are significant and which are not. If $\phi(x_1, x_2, \ldots, x_n)$ is not known, one can often still deduce the structure of $F(\pi_1, \pi_2, \ldots, \pi_{n-m}) = 0$ and so obtain useful information about the system in question.

An example is the swinging of a simple pendulum. Assume that the analytic expression for its period of vibration is unknown. Choose mass, length, and time as the fundamental quantities. Thus $m = 3$. One must make a list of all the parameters pertaining to the pendulum which might be significant. If the list is incomplete, no useful information will be obtained. If too many quantities are included, the derived information is less specific than it might otherwise be. The quantities given in **Table 2** would appear to be adequate.

The microscopic properties of the bob and string are not considered, and air resistance and the mass of the string are likewise neglected. These would be expected to have only a small effect on the period. Thus $n = 5$, $n - m = 2$, and therefore one expects to find two independent dimensionless products. These can always be found by trial and error. There may be more than one possible set of independent π's. In this case one π is simply θ, the angular amplitude. Another is $l/\tau^2 g$. Since no other of the n variables contains M in its dimensional formula, the mass of the bob m cannot occur in any dimensionless product. The π theorem gives $F(\theta, l/\tau^2 g) = 0$. Therefore the period of vibration does not depend upon the mass of the bob.

TABLE 2. Dimensional formulas for a simple pendulum

Quantity	Symbol	Dimensional formula
Mass of bob	m	M
Length of string	l	L
Acceleration of gravity	g	LT^{-2}
Period of swing	τ	T
Angular amplitude	θ	0

Because τ appears in only one of the dimensionless products, this expression can be explicitly solved for τ to give $\tau = G(\theta)\sqrt{l/g}$, where $G(\theta)$ is an arbitrary function of θ. Now make the further assumption that θ is small enough to be neglected. Then $n = 4$, $n - m = 1$, and $F(l/\tau^2 g) = 0$. Thus $l/\tau^2 g =$ constant, or $\tau \propto \sqrt{l/g}$. The π theorem thus leads to the conclusion that τ varies directly as the square root of the length of the pendulum and inversely as the square root of the acceleration of gravity. The magnitude of the dimensionless constant (actually it is 2π) cannot be obtained from dimensional analysis.

In more complicated cases, where a direct solution is not feasible, this method can give information on how certain variables enter a particular problem even where $F(\pi_1, \pi_2, \ldots, \pi_{n-m}) = 0$ cannot be explicitly solved for one variable. The procedure is particularly useful in hydraulics and aeronautical engineering, where detailed solutions are often extremely complicated.

Model theory. A further application of dimensional analysis is in model design. Often the behavior of large complex systems can be deduced from studies of small-scale models at a great saving in cost. In the model each parameter is reduced in the same proportion relative to its value in the original system. *See* MODEL THEORY.

Once again the case of the simple pendulum is a good example. It was found from the π theorem that $F(\theta, l/\tau^2 g) = 0$. If the magnitudes of θ, l, τ, and g are now changed in such a way that neither argument of F is changed in numerical value, the system will behave exactly as the original system and is said to be physically similar. Evidently θ cannot be changed without altering any of the arguments of F, but l, g, and τ can be varied. Suppose that it was desired to build a very large and expensive pendulum which was to swing with finite amplitude. One could build a small model of, say, $^1/_{100}$ the length and time its swing for an amplitude equal to that for the desired pendulum. The acceleration of gravity g would be the same for the model as for the large pendulum.

The period for the model would then be just $^1/_{10}$ that for the large pendulum. Thus the period of the large pendulum could be deduced before the pendulum was ever built. In practice one would never bother with the π theorem in cases as simple as this where a full analytic solution is possible. In many situations where such a solution is not feasible, models are built and extensively studied before the full-scale device is constructed. This technique is standard in wind tunnel studies of aircraft design. *See* DYNAMIC SIMILARITY.

Cataloging of physical quantities. Dimensional formulas provide a convenient shorthand notation for representing the definitions of secondary quantities. These definitions depend upon the choice of primary quantities. The π theorem is applicable no matter what the choice of primary quantities is.

Changing units. Dimensional formulas are helpful in changing units from one system to another. For example, the acceleration of gravity in the centimeter-gram-second system of units is 980 cm/s^2. The dimensional formula for acceleration is LT^{-2}. To find find the magnitude of g in mi/h^2, one would proceed as in Eq. (10).

$$\begin{aligned} 980\frac{\text{cm}}{s^2} \times \frac{(\text{conversion factor for length})}{(\text{conversion factor for time})^2} \\ = 980\frac{\text{cm}}{\text{s}^2} \times \frac{\dfrac{1}{1.609 \times 10^5}\dfrac{\text{mi}}{\text{cm}}}{(1/3600)^2(\text{h}^2/\text{s}^2)} \\ = 7.89 \times 10^5 \text{mi/h}^2 \end{aligned} \tag{10}$$

In the past the subject of dimensions has been quite controversial. For years unsuccessful attempts were made to find "ultimate rational quantities" in terms of which to express all dimensional formulas. It is now universally agreed that there is no one "absolute" set of dimensional formulas. Some systems are more symmetrical than others and for this reason are perhaps preferable. The representation of electrical quantities in terms of M, L, and T alone through the electrostatic form of Coulomb's law leads to somewhat awkward fractional exponents, but nevertheless is just as "correct" as a representation in which charge is used as a fourth fundamental unit.

A highly symmetrical pattern results if energy, linear displacement, and linear momentum are chosen as the fundamental quantities in mechanics. In electricity one can use energy, charge, and magnetic flux. The corresponding quantities for a vibrating mass on a spring and the analogous alternating-current circuit with inductance and capacitance have similar dimensional formulas. In this analogy energy is invariant, charge corresponds to displacement, and magnetic flux corresponds to linear momentum. This correspondence is not displayed in conventional dimensional formulas. John W. Stewart

Bibliography. G. I. Barenblatt, *Dimensional Analysis*, 1987; P. W. Bridgman, *Dimensional Analysis*, 1931, reprint 1976; L. I. Sedov, *Similarity and Dimensional Methods in Mechanics*, 10th ed., 1993; T. Szirtes, *Applied Dimensional Analysis Modeling*, 1998.

Dimensionless groups

A dimensionless group is any combination of dimensional or dimensionless quantities possessing zero overall dimensions. Dimensionless groups are frequently encountered in engineering studies of complicated processes or as similarity criteria in model studies. A typical dimensionless group is the Reynolds number (a dynamic similarity criterion), $\text{Re} = vl\rho/\mu$. Since the dimensions of the quantities involved are velocity v: $[L/\theta]$; characteristic length l: $[L]$; density ρ: $[M/L^3]$; and viscosity μ: $[M/LT]$ (with M, L, and T as the fundamental units of mass, length, and time), the Reynolds number reduces to a dimensionless group and can be represented by a pure number in any coherent system of units. *See* DYNAMIC SIMILARITY.

History of named groups. The importance of dimensionless groups has long been known; as early as 1850 G. G. Stokes found the group now known as the Reynolds number and demonstrated its importance as a dynamic similarity criterion. In 1873 H. von Helmholtz showed the importance of the groups now known as the Froude and Mach numbers. The first attempt to attach names to important dimensionless groups appears to have been made in 1919 by M. G. Weber, who first named the Froude, Reynolds, and Cauchy numbers. It has become customary to name dimensionless groups after the outstanding or pioneer workers in the field. Groups are often denoted by the first two letters of the name; thus, the Reynolds number is usually denoted as Re.

Sources and uses. Many important problems in applied science and engineering are too complicated to permit completely theoretical solutions to be found. However, the number of interrelated variables involved can be reduced by carrying out a dimensional analysis to group the variables as dimensionless groups. *See* DIMENSIONAL ANALYSIS.

If almost nothing is known about the problem, it is necessary to decide which physical parameters will affect the process being studied. A dimensional analysis can then be carried out by the well-known π-theorem method of E. Buckingham or by the indicial method proposed by Lord Rayleigh. The success of these methods depends on identifying all of the important parameters for the problem. On the other hand, if the differential equations describing the phenomenon are known, inspectional analysis of the equations will yield all the important dimensionless groups even when the differential equations cannot be solved. Various procedures have been developed to determine the groups which correspond to the assumed parameters or variables. The form of the dimensionless equations linking these groups has to be determined by the use of experimental data in all except the simplest cases. Often there is no better procedure available in the case of complicated phenomena.

The advantages of using dimensionless groups in studying complicated phenomena include:

1. A significant reduction in the number of "variables" to be investigated; that is, each dimensionless

group, containing several physical variables, may be treated as a single compound "variable," thereby reducing the number of experiments needed as well as the time required to correlate and interpret the experimental data.

2. Predicting the effect of changing one of the individual variables in a process (which it may be impossible to vary much in available equipment) by determining the effect of varying the dimensionless group containing this parameter (this must be done with some caution, however).

3. Making the results independent of the scale of the system and of the system of units being used.

4. Simplifying the scaling-up or scaling-down of results obtained with models of systems by generalizing the conditions which must exist for similarity between a system and its model.

Deducing variation in importance of mechanisms in a process from the numerical values of the dimensionless groups involved; for instance, an increase in the Reynolds number in a flow process indicates that molecular (viscous) transfer mechanisms will be less important relative to transfer by bulk flow (inertia effects), since the Reynolds number is known to represent a measure of the ratio of inertia forces to viscous forces. *See* UNITS OF MEASUREMENT.

Classification. A large number of independent dimensionless groups may arise from a dimensional analysis of a complex problem; this number usually equals the number of variables in the problem minus the number of fundamental units in the system of units used. Furthermore, since any product or quotient of powers of dimensionless groups is also dimensionless, it follows that an almost limitless number of alternative dimensionless groups can be generated. Several hundred groups are known, many of them named (see **table**). It is therefore useful to be able to specify the conditions for obtaining a complete statement of a problem in terms of dimensionless groups and to classify the large number of groups in some way.

Complete sets. A set of dimensionless groups of given variables is complete if each group in the set is independent of the others and if every other dimensionless combination of the variables is a product of powers of these groups. For example, a complete set of dimensionless groups for a problem involving the variables length (l), velocity (v), density (ρ), viscosity (μ), the acceleration of gravity (g), the velocity of sound (v_s), and surface tension (γ) would be the Reynolds number, $\mathrm{Re} \equiv vl\rho/\mu$; the Froude number, $\mathrm{Fr} = v/(gl)^{1/2}$; the Mach number, $\mathrm{Ma} = v/v_s$, and the Weber number, $\mathrm{We} = v(\rho l/g)$. None of these groups is a product of powers of the others, since μ occurs only in Re, g only in Fr, v_s only in in Ma, and γ only in We. Many other

TABLE 1. List of frequently occurring dimensionless groups*

Symbol and name	Definition	Notation	Physical significance
Ar; Archimedes number	$d_p^3 g\rho_f(\rho_s - \rho_f)/\mu^2$		(Inertia force) (gravity force)/(Viscous force)^2
Bi; Biot number	hl/k_s	k_s = thermal conductivity of solid, $ML/T^3\theta$	(Internal thermal resistance)/(Surface thermal resistance)
Bi_m; mass transport Biot number	$k_m L/D_i$	L = Layer thickness, L D_i = interface diffusivity, L^2/T	(Mass transport conductivity at solid/fluid interface)/(Internal transport conductivity of solid wall of thickness L)
Bo; Bond number (also Eo, Eötvös number)	$d^2g(\rho - \rho_f)/\gamma$	d = bubble or droplet diameter, L	(Gravity force)/(Surface tension force)
Bq; Boussinesq number	$v/(2gm)^{1/2}$	m = mean hydraulic depth of open channel, L	(Inertia force)$^{1/2}$/(Gravity force)$^{1/2}$
Dn; Dean number	$(d/2R)^{1/2}(\mathrm{Re})$	d = pipe diameter, L R = radius of curvature of channel centerline, L Re = Reynolds number	Effect of centrifugal force on flow in a curved pipe
Ec; Eckert number	$v^2/C_p\Delta\theta$	$\Delta\theta$ = temperature difference, θ	Used in study of variable-temperature flow
Ek; Ekman number	$(\nu/2\omega l^2)^{1/2}$		(Viscous force)$^{1/2}$/(Coriolis force)$^{1/2}$
f; Fanning friction factor	$2\tau_w/\rho v^2 = D\,\Delta p/2L\rho v^2$	τ_w = wall shear stress, M/LT^2 D = pipe diameter, L L = pipe length, L	(Wall shear stress)/(Velocity head)
Fo_f; Fourier flow number	$\nu t/l^2$	—	Used in undimensionalization
Fo; Fourier flow number	$\alpha t/l^2$	—	Indicates the extent of thermal penetration in unsteady-state heat transport
Fo_m; mass transport Fourier number	$k_m t/l$	—	Indicates the extent of substance penetration in unstedy-state mass transport
Fr; Froude number	$v/(gl)^{1/2}$	—	(Inertia force)$^{1/2}$/(Gravity force)$^{1/2}$
Ga; Galileo number	l^3g/ν^2	—	(Inertia force)(gravity force)/(Viscous force)^2
Gz; Graetz number	$\dot{m}C_p/k_f l$	$\dot{m}$ = mass flow rate, M/T C_p = specific heat capacity (constant pressure), $L^2/T^2\theta$	(Fluid thermal capacity)/(Thermal energy transferred by conduction)
Gr; Grashof number	$l^3g\Delta\rho/\rho_f\nu^2$	$\Delta\rho$ = density driving force, M/L^3	(Inertia force)(buoyancy force)/(Viscous force)^2
Gr_m; mass transport Grashof number	$l^3g\beta_c\Delta c l\nu^2$	β_c = volumetric expansion coefficient Δc = concentration driving force	
Ha; Hartmann number (M)	$lB(\sigma/\mu)^{1/2}$	B = maganetic flux density, M/QT σ = electrical conductivity, Q^2T/ML^3	(Magnetically induced stress)$^{1/2}$/(Viscous shear stress)$^{1/2}$
Kn; Knudsen number	λ/l	λ = length of mean free path	Used in study of low-pressure gas flow
Le; Lewis number	D/α	—	(Molecular diffusivity)/(Thermal diffusivity)
Lu; Luikov number	$k_m l/\alpha$	—	(Mass diffusivity)/(Thermal diffusivity)
Ly; Lykoudis number	—	—	(Hartman number)^2/(Grashof number)$^{1/2}$
Ma; Mach number	v/v_s	v_s = velocity of sound in fluid, L/T	(Linear velocity)/(Velocity of sound)

(*continued*)

TABLE 1. List of frequently occurring dimensionless groups* (*cont.*)

Symbol and name	Definition	Notation	Physical significance
Ne; Newton number	$F/\rho_f v^2 l^2$	F = hydrodynamic drag force, ML/T^2	(Resistance force)/(Inertia force)
Nu; Nusselt number	hl/k_f	—	(Thermal energy transport in forced convection)/(Thermal energy transport if it occurred by conduction)
Pe; Peclet number	lv/α	—	(Reynolds number)(Prandtl number); (Bulk thermal energy transport in forced convection)/(Thermal energy transport by conduction)
Pe_m; mass transport Peclet number	lv/D	—	(Bulk mass transport)/(Diffusional mass transport)
Ps; Poiseuille number	$v\mu/gd_p^2(\rho_s \quad \rho_f)$	—	(Viscous force)/(Gravity force)
Pr; Prandtl number	ν/α	—	(Momentum diffusivity)/(Thermal diffusivity)
Ra; Rayleigh number			(Grashof number) (Prandtl number)
Ra_m; mass transport Rayleigh number			(Mass transport Grashof number) (Schmidt number)
Re; Reynolds number	vl/ν	—	
Re_R; rotational Reynolds number	l^2N/ν	N = rate of rotation (revolution per time), $1/T$	(Inertia force)/(Viscous force)
Sc; Schmidt number	ν/D	—	(Momentum diffusivity)/(Molecular diffusivity)
Sh; Sherwood number	$k_m l/D$	—	(Mass diffusivity)/(Molecular diffusivity)
St; Stanton number	$h/C_p\rho v$	—	(Nusselt number)/(Reynolds number) (Prandtl number); (Thermal energy transferred)/(Fluid thermal capacity)
St_m; mass transport Station number	k_m/v	—	(Sherwood number)/(Reynolds number) (Schmidt number)
Sk; Stokes number	$l^2\Delta p/\mu\nu$	—	(Pressure force)/(Viscous force)
Sr; Strouhal number	fl/v	f = oscillation frequency, $1/T$	Used in study of unsteady flow
Su; Suratnam number	$\rho l\gamma/\mu^2$	—	(Inertia force)(surface tension force)/(Viscous force)^2
Ta; Taylor number	$\omega^2\gamma\bar{r}(r_o \quad r_i)^3/\nu^2$	r_o = outer radius, L r_i = inner radius, L $\bar{r}$ = mean radius, L	Criterion for Taylor vortex stability in rotating concentric cylinder systems
We; Weber number	$v(\rho_f l/\gamma)^{1/2}$	—	(Inertia force)$^{1/2}$ /(Surface tension force)$^{1/2}$

*In many cases, roots of, or power functions of, the groups listed here may be designated by the same names in the technical literature.

Names of quantities are followed by their dimensions. Fundamental dimensions are taken to be length [L], mass [M], time [T], temperature [θ], and electrical charge [Q]. Notation for quantities that appear only once is given in the table. The following quantities appear more than once:

C_p = specific heat at constant pressure, $L^2/T^2\theta$
d_p = particle diameter, L
D = molecular diffusivity, L^2/T
g = acceleration due to gravity = weight/mass ratio, L/T^2
h = heat transfer coefficient, $M/T^3\theta$
k_f = thermal conductivity of fluid, $ML/T^3\theta$
k_m = mass transfer coefficient, L/T
l = characteristic length, L
Δp = pressure drop, M/LT^2
t = time, T
v = charcteristic velocity, L/T
α = thermal diffusivity = $k_f/\rho_f C_p$, L^2/T
γ = surface tension, M/T^2
μ = dynamic or absolute viscosity, M/LT
ν = kinematic viscosity $=\mu/\rho_f$, L^2/T
ρ = bubble or droplet density, M/L^3
ρ_f = fluid density, M/L^3
ρ_s = solid density, M/L^3
ω = angular velocity of fluid, $1/T$

SOURCE: After N. P. Cheremisinoff (ed.), *Encyclopedia of Fluid Mechanics*, vol. 1: *Flow Phenomena and Measurement*, Gulf Publishing Co., Houston, copyright © 1986. Used with permission. All rights reserved.

dimensionless groups could be formed from these same variables, such as $v^3\rho/\mu g$, but these are not independent groups because, for instance, $v^3\rho/\mu g = \mathrm{Re} \cdot \mathrm{Fr}^2$. Many of these compound groups have acquired names, however, and a separate existence of their own, since any of the combined groups can be used as alternative for any one of the component groups.

F. V. A. Engel proposed that the groups forming a complete set (as defined above) for a given problem be termed true groups for that problem, and that other groups involving only independent variables in the particular problem which may be formed as combinations of the true groups be termed compound groups. He gave rules for obtaining complete sets of groups for heat transfer and flow problems, and the arguments can be readily extended to other types of problems.

Simplexes and complexes. Sometimes more than one variable of the same physical significance is important in a problem. For instance, in fluid flow through a straight rough pipe, three length parameters (pipe diameter D, pipe length L, and surface roughness ϵ) are required to define the geometry of the system. In such a case, one of these quantities (usually D in the pipe-flow problem) is used as the characteristic quantity in forming the main dimensionless groups (the Reynolds number for pipe flow), and the others occur in the form of ratios to the first (L/D, ϵ/D). These straightforward dimensionless ratios of multiple parameters arising naturally in a problem are often referred to as simplexes, while all other groups are termed complexes. Many simplexes are unnamed, being referred to simply as "geometric ratios" or the like, but a number of named simplexes exist, for example, the Knudsen and Mach numbers. Simplexes and complexes have also been termed parametric and similarity criteria, respectively.

Independent and dependent groups. Groups formed entirely from parameters which are independent variables in a particular problem have been termed independent or characteristic groups, while those which also involve dependent variables for the given problem are known as dependent or unknown groups.

Obviously, some groups which are independent in one problem may become dependent in another, according to which parameters are specified and which are unknown in each case.

In many practical cases, experimental data are correlated to give a relationship between a dependent group and one or more independent groups. Although there are certain philosophical problems involved, this procedure has led to many very useful engineering correlations. Examples are the Moody friction factor chart, which correlates the friction factor f for pipe flow in terms of the Reynolds number and the relative pipe roughness given by Eq. (1)

$$f = f(\mathrm{Re}, \epsilon/D) \tag{1}$$

and the familiar correlations for heat and mass transfer given by Eqs. (2) and (3). This type of power

$$\mathrm{Nu} = A \cdot \mathrm{Re}^a \cdot \mathrm{Pr}^b \tag{2}$$

$$\mathrm{Sh} = A' \cdot \mathrm{Re}^{a'} \cdot \mathrm{Sc}^{b'} \tag{3}$$

relationship is suggested by the Rayleigh indicial procedure for dimensional analysis. The constants A, a, and b are determined by fitting straight lines to suitable logarithmic plots of experimental data.

A special subclass of independent dimensionless groups contains only the parameters expressing the physical properties of the system being studied. This category includes such groups as the Prandtl, Schmidt, and Lewis numbers. *See* BALLISTICS; CAVITATION; CONVECTION (HEAT); DIFFUSION; FROUDE NUMBER; GAS DYNAMICS; KNUDSEN NUMBER; MACH NUMBER; PIPE FLOW; REYNOLDS NUMBER.

John Catchpole; George D. Fulford

Bibliography. G. I. Barenblatt, *Dimensional Analysis*, 1987; N. P. Cheremisinoff (ed.), *Encyclopedia of Fluid Mechanics*, vol. 1: *Flow Phenomena and Measurement*, 1986; F. T. Farago and M. Curtis, *Handbook of Dimensional Measurement*, 3d ed., 1994; R. V. Giles, C. Liu, and J. B. Evett, *Schaum's Outline of Fluid Mechanics and Hydraulics*, 3d ed., 1994; L. I. Sedov, *Similarity and Dimensional Methods in Mechanics*, 10th ed., 1993; M. Zlokarnik, *Dimensional Analysis and Scale-up in Chemical Engineering*, 1991.

Dimensions (mechanics)

Length, mass, time, or combinations of these quantities serving as an indication of the nature of a physical quantity. Quantities with the same dimensions can be expressed in the same units. For example, although speed can be expressed in various units such as miles/hour, feet/second, and meters/second, all these speed units involve the ratio of a length unit to a time unit; hence, the dimensions of speed are the ratio of length L to time T, usually stated as LT^{-1}. The dimensions of all mechanical quantities can be expressed in terms of L, T, and mass M. The validity of algebraic equations involving physical quantities can be tested by a process called dimensional analysis; the terms on the two sides of any valid equation must have the same dimensions. *See* DIMENSIONAL ANALYSIS; UNITS OF MEASUREMENT.

Dudley Williams

Dimethyl sulfoxide

A versatile solvent, abbreviated DMSO, used industrially and in chemical laboratories as a medium for carrying out chemical reactions. Its uses have been extended to that of a chemical reagent with DMSO as a reactant undergoing a chemical change. DMSO has also gained popularity as a medicinal compound; however, its efficacy as a curative agent is not well established.

DMSO is the simplest member of a class of organic compounds which are typified by the polar sulfur-oxygen bond represented in the resonance hybrid shown in reaction (1). The molecule is pyramidal in

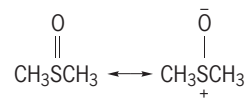

(1)

shape with the oxygen and the carbons at the corners, as in the structure below. *See* ORGANOSULFUR COMPOUND.

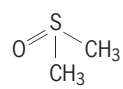

Because of the polarity of the sulfur-oxygen bond, DMSO has a high dielectric constant ($\epsilon^{20} = 48.9$) and is slightly basic. It forms crystalline salts with strong protic acids, reaction (2), and it coordi-

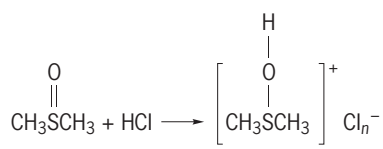

(2)

nates with Lewis acids. It is a colorless, odorless (when pure), and very hygroscopic stable liquid (bp 189°C or 372°F, mp 18.5°C or 65.3°F). It is manufactured commercially by reacting the black liquor, from digestion in the kraft pulp process, with molten sulfur to form dimethyl sulfide, CH_3SCH_3. The dimethyl sulfide is oxidized with nitrogen tetroxide to yield DMSO. This highly polar aprotic solvent is water- and alcohol-miscible and will dissolve most polar organic compounds and many inorganic salts.

Solvent. As an aprotic solvent, DMSO strongly solvates cations, leaving a highly reactive anion, reaction (3). Thus in DMSO, basicity and nucleophilicity

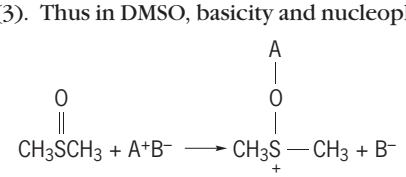

(3)

are enhanced, and it is a superior solvent for many

elimination, nucleophilic substitution, and solvolysis reactions in which nucleophile and base strength are important. Nucleophilic substitution reactions in which halogens or sulfonate esters are displaced by anions such as cyanide, alkoxide, thiocyanate, azide, and others are accelerated 1000 to 10,000 times in DMSO over the reaction in aqueous alcohol. Reaction (4) illustrates nucleophilic substitution with the

$$(CH_3)_3CCH_2Cl + NaCN \xrightarrow{DMSO} (CH_3)_3CCH_2CN + NaCl \quad (4)$$

cyanide anion. Dimethyl sulfoxide is also superior to protic solvents as a medium for elimination reactions. Reaction (5) illustrates the use of DMSO

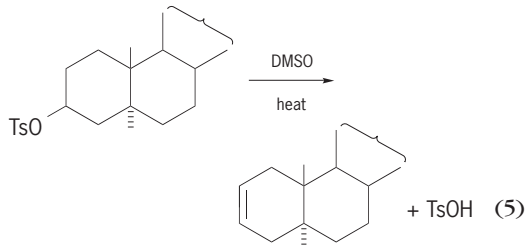

(5)

in the elimination of a tosyl group from a steroid derivative.

Eliminations can be accomplished by heating secondary halides and sulfonate esters in 100% DMSO. Other eliminations, often unsuccessful in protic solvents, are accomplished by the use of potassium *t*-butoxide [$K^{+-}OC(CH_3)_3$] in DMSO.

The elimination of tertiary and secondary alcohols has been accomplished by refluxing in 100% DMSO. Reactions (6) and (7) illustrate this elimination from a tertiary and secondary alcohol respectively.

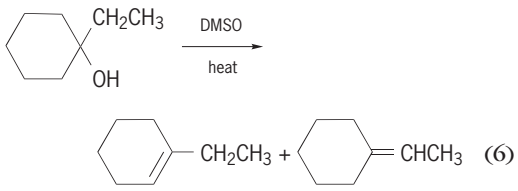

(6)

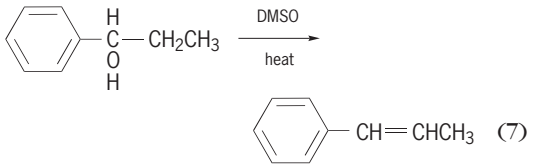

(7)

The distinction between solvent and reactant in this and previous eliminations has not been clarified. Dimethyl sulfoxide has found use as a suitable solvent for thermal decarboxylations and base-catalyzed double-bond isomerizations. A double-bond isomerization utilizing DMSO and potassium *t*-butoxide is illustrated by reaction (8). The high dielectric constant

$$CH_2{=}CHCH_2OR \xrightarrow[DMSO]{K^{+-}OC(CH_3)_3} CH_3CH{=}CH{-}OR \quad (8)$$

makes DMSO a useful solvent for dissolving resins, polymers, and carbohydrates. It is employed commercially as a spinning solvent in the manufacture of synthetic fibers. *See* SOLVENT.

Chemical reactions. The compound is capable of taking part in several types of chemical reactions. These reactions are described below.

Coordination compounds. Dimethyl sulfoxide acts as an electron donor to form coordination compounds with a large number of inorganic ions. In most complexes the DMSO ligand is attached through the oxygen atoms, and the bond strength is roughly that of the corresponding aquo (H_2O) derivative. In addition to the typical metal-ion complexes, DMSO also forms some complexes with nonmetallic compounds. For example, BF_3 forms a one-to-one complex which is highly hygroscopic and low melting. *See* COORDINATION CHEMISTRY; COORDINATION COMPLEXES.

Synthetic intermediate. Acting as an acid, DMSO reacts as indicated by reaction (9) with sodium hydride to form the strong base methylsulfinyl carbanion. Methylsulfinyl carbanion, an effective base catalyst in many condensation reactions, itself condenses with carbonyl compounds, particularly esters, reaction (10), to give β-keto sulfoxides in high yield. The

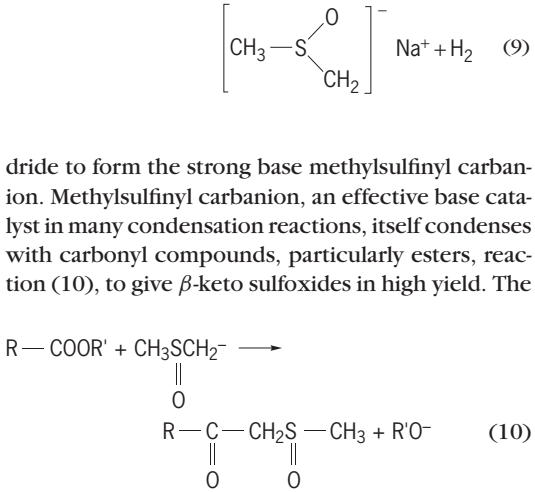

(9)

(10)

synthetic importance of the easily prepared β-keto-sulfoxides lies in the fact that they can be easily reduced by aluminum amalgam to methyl ketones in a large variety of cases, reaction (11).

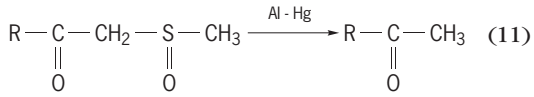

(11)

See CARBONYL; ORGANIC SYNTHESIS; REACTIVE INTERMEDIATES.

Oxidizing agent. Under the proper experimental conditions, many compounds (primary or secondary alcohols, halides, or sulfonates) can be oxidized to the corresponding carbonyl compound by DMSO. Most of the DMSO oxidations appear to involve the same dimethylalkoxysulfonium salt intermediate, which subsequently reacts with a base to give the observed carbonyl product and dimethyl sulfide. Through pathways A and B in reaction (12)

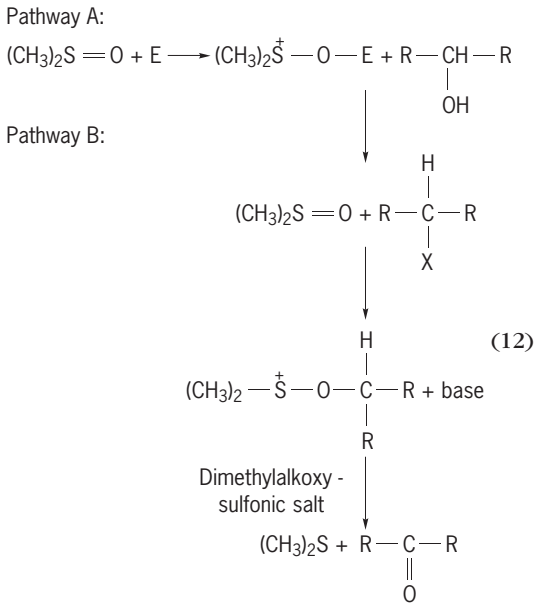

(12)

a compound may be converted into the dimethylalkoxysulfonium salt intermediate. The reaction pathway is determined by the structure of the substrate.

Pathway A involves reaction of DMSO with an intermediate activating electrophilic species E, which is subsequently displaced by the compound to be oxidized to form the dimethylalkoxysulfonium salt.

Pathway B involves a leaving group X ($X = Cl^-$, Br^-, I^-, or sulfonate) being displaced by DMSO acting as a nucleophile, and resulting directly in the dimethylalkoxysulfonium salt.

The most common technique involves use of dicyclohexylcarbodiimide as the E group and pyridine as the base and is applicable to primary or secondary alcohol groups in an almost unlimited variety of compounds including alkaloids, steroids, carbohydrates, and other complex substances. The simple experimental conditions with which most oxidations can be done have made this an important oxidation method. *See* ORGANIC REACTION MECHANISM.

Pharmacology and toxicity. The apparent low toxicity and high skin permeability have led to extensive studies in many biological systems. DMSO facilitates the transport of inorganic salts or small-molecular-weight organic compounds across skin membrane, indicating a potential hazard in commercial use.

DMSO is used as a topical anti-inflammatory agent; however, its efficacy for this purpose is not well established. Its use as such is not recommended without medical supervision because of the limited information about adverse side effects which may accompany repeated or long-term application. In some cases skin contact causes irritation and can result in lesions.

DMSO has become a readily available chemical, and many claims for its curative properties are poorly or totally unsubstantiated. Some studies failed to show any benefit of topically applied DMSO in the promotion of wound (dermal and epidermal) healing. DMSO is widely used alone or in combination with other agents to induce cell differentiation under culture conditions. Induced cellular differentiations appear to be related to DMSO-mediated alterations in cell-surface characteristics and enzyme activities. These findings suggest that DMSO should not be taken internally, taken by injection, or used topically without medical supervision.

DMSO has proven useful as a cryoprotective agent for a large number of mammalian and nonmammalian cells, probably because of its influence on membrane permeability and ability to decrease ice nucleation temperatures. Numerous physiological effects of DMSO have been studied, but its extended and routine use should be discouraged until definite clinical efficacy is demonstrated. *See* CRYOBIOLOGY.

W. W. Epstein; F. W. Sweat

Bibliography. N. A. David, The pharmacology of dimethyl sulfoxide, *Annu. Rev. Pharmacol.,* 12:353–374, 1972; W. W. Epstein and F. W. Sweat, Dimethyl sulfoxide oxidations, *Chem. Rev.,* 67:247–260, 1967; L. F. Fieser and M. Fieser, *Reagents for Organic Synthesis,* vol. 6, pp. 225–230, 1977; S. W. Jacob, E. E. Rosenbaum, and D. C. Wood (eds.), *Dimethyl Sulfoxide,* 1971.

Dinocerata

An extinct order of large herbivorous mammals, often called uintatheres, from early Cenozoic deposits of North America and northern Asia. Members of this group have semigraviportal limbs, that is, adapted to bearing considerable weights, with hoofed, five-toed feet. The dentition is somewhat reduced in all forms, later ones losing all the upper incisors and in one case (*Gobiatherium*, from the late Eocene of Mongolia) even the upper canine. The cusps of the upper molars and premolars form V-shaped patterns. The lower molars and premolars possess V-shaped crests, followed by a low shelf. A saberlike canine tooth and protective lower jaw flange are present in all forms except the aberrant *Gobiatherium*, which must have relied on other means for defense. Horns are absent or very small on the most primitive forms, but by early Eocene time North American genera had begun to develop them. Middle and late Eocene uintatheres in North America developed an imposing array of six horns. One pair was on the tips of the nasal bones, another was above the root of the saberlike canine tooth, and a third pair above the ear region. *See* DENTITION; TOOTH.

The order Dinocerata consists of one family, the Uintatheriidae, which is divided into three subfamilies: Prodinoceratinae, Uintatheriinae, and Gobiatheriinae. The Prodinoceratinae includes two Mongolian genera from the late Paleocene and early Eocene, an American genus from deposits of the same ages, and a second American genus from the late Paleocene. The Mongolian genus *Mongolotherium* is the best known of these and demonstrates that the most primitive uintatheres possessed a carnivorelike body of moderate size.

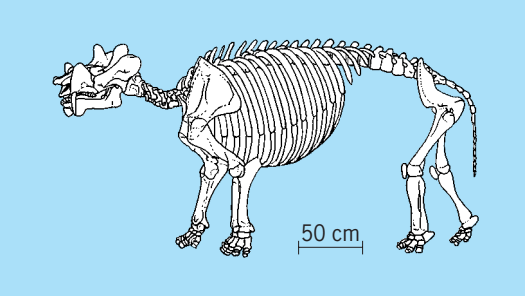

Skeleton of *Uintatherium*, a middle Eocene member of the Dinocerata.

The subfamily Uintatheriinae, or horned forms, includes all other American members of the order (see **illus.**). Uintatheres increased to rhinoceros size in the Eocene, were especially common in the middle Eocene, and died out before the Oligocene began.

The Gobiatheriinae, an aberrant Asiatic side branch of the early Prodinoceratinae, is sufficiently distinct to merit subfamily rank. This subfamily, known from one late Eocene genus, is characterized by extreme reduction of the anterior dentition, including the canine, and by lack of horns. The skull is remarkably low and flat.

The Dinocerata left no descendants. They are believed to have arisen from the arctocyonid condylarths but from a different subfamily than that which gave rise to the order Pantodonta. *See* ARCHAIC UNGULATE; MAMMALIA; PANTODONTA.

Malcolm C. McKenna

Bibliography. M. J. Benton, *Vertebrate Paleontology,* 1991; R. L. Carroll, *Vertebrate Paleontology and Evolution*; 1988.

Dinoflagellida

An order of the class Phytamastigophorea, also known as the Dinoflagellata. Many botanists consider this group to be a member of the class Dinophyceae. *Noctiluca scintillans*, one of the largest species, may measure 0.06 in. (1.5 mm), whereas the flagellate state of some species may be as small as 10 micrometers. Although primarily marine, some dinoflagellates occur in fresh water. Some possess brown chromatophores, some are variously colored, and others are colorless. Masking pigments are frequent, occurring in the cytoplasm or in chromatophores. All types of nutrition exist, and some species are parasitic. Two flagella emerge laterally from a longitudinal depression or sulcus. An encircling flagellum in a groove or girdle divides the body into epicone and hypocone; the other, extending backward, propels the organism forward. The nucleus is very large. *See* CILIA AND FLAGELLA.

Dinoflagellate species, often of bizarre form, have fixed shapes determined by thick covering plates, except in Gymnodinia which are naked or have thin pellicles. Chain-formation, ameboid, and palmella aggregates occur. *Ceratium hirundinella* (**Fig. 1**) blooms in fresh water, causing tastes and odors. Marine blooms of *Gymnodinium breve* (**Fig. 2**) produce the fish-killing red tides which occur along the Gulf Coast of the southern United States. *See* DINOPHYCEAE; PHYTAMASTIGOPHOREA; PROTOZOA.

James B. Lackey

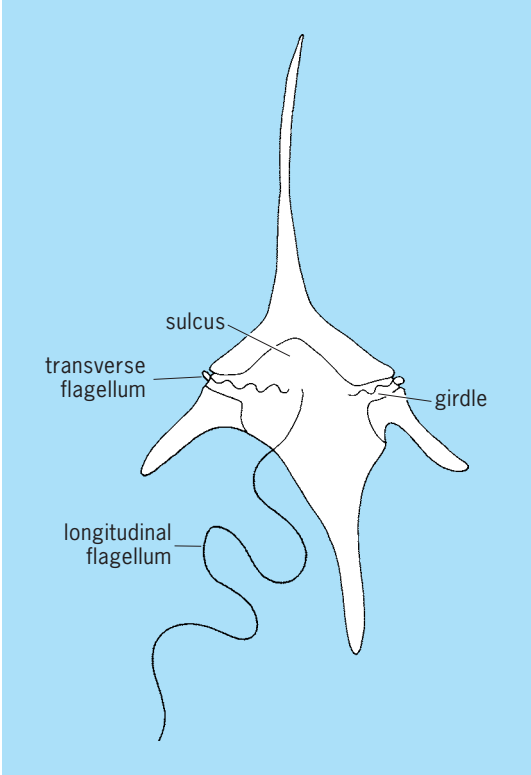

Fig. 1. *Ceratium hirundinella*, size 95–200 μm.

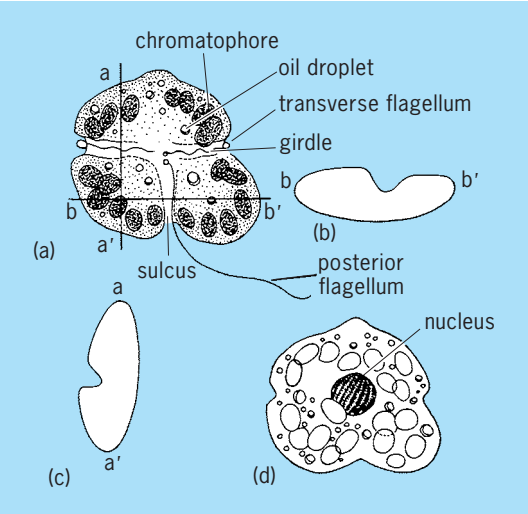

Fig. 2. *Gymnodinium breve*, which varies in size from 28 to 45 μm. (*a*) Overall view. (*b*, *c*) Transverse sections. (*d*) Fixed organism, showing central nucleus.

Dinophyceae

A large and extremely diverse class of biflagellate algae (dinoflagellates) in the chlorophyll *a–c* phyletic line (Chromophycota). In protozoological

classification, these organisms constitute an order, Dinoflagellida, of the class Phytomastigophora. Many taxonomists emphasize the distinctness of dinoflagellates by placing them in a separate division (Pyrrophyta or Pyrrhophyta) or even in a separate kingdom (Mesokaryota). More than 1200 species are known, classified into 18 orders and 54 families. Most are microscopic, but a few reach a diameter of 2 mm (0.08 in.). Cell shape is highly variable, with many planktonic species having elaborately modified surfaces. Dinoflagellates occur in marine, brackish, and fresh waters, frequently producing algal blooms. They may be benthic as well as planktonic, and a few are colonial. Ameboid, palmelloid, coccoid (with or without a gelatinous sheath), and filamentous forms are also known. *See* DINOFLAGELLIDA.

Structure. Most dinoflagellates at some stage in their life history bear a pair of flagella with one of two types of insertion. In one type, a posteriorly directed flagellum is located in a longitudinal groove, or sulcus, and a ribbonlike flagellum encircles the cell in a transverse groove, or cingulum (see **illus.**). The transverse flagellum, which provides the main propulsive force, comprises a striated paraflagellar rod and an undulating axoneme in a common sheath. In the other type of insertion, the flagella are borne apically or subapically on a cell that is roughly bilaterally symmetrical. One flagellum is directed forward, while the second (which has the structure of the transverse flagellum in the first type of insertion) encircles the first and beats in a perpendicular plane. Dinoflagellates frequently swim in a whirling motion, giving rise to their name (from the Greek *dinein*, to whirl).

Most dinoflagellates have an elaborate cell covering, the amphiesma, composed of several layers of concentric membranes between which may be found a cellulosic layer comprising separate plates (thecal plates). The number and arrangement of these plates have been accorded taxonomic significance. In some species the outer surface is covered by small scales.

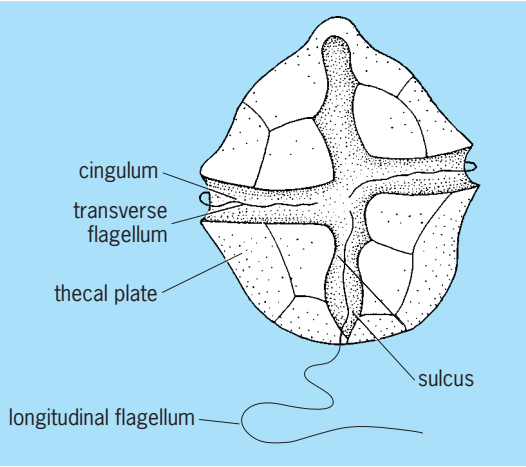

***Gonyaulax*, a dinoflagellate. The two flagella emerge from adjacent pores in the cingulum. (*After H. J. Fuller and O. Tippo, College Botany, rev. ed., Holt, 1954*)**

Ecology. Dinoflagellates span the spectrum of nutritional diversity. About half of the species are photosynthetic, and some of these are facultatively osmotrophic (absorbing nutrients) or phagotrophic (engulfing food). Of the colorless species, some are osmotrophs while others are predatory phagotrophs. Chloroplasts are characteristically yellowish- or greenish-brown. Like those of other algae of the chlorophyll *a–c* line, they are bounded by three membranes, but differ in that the outer membrane is not part of the endoplasmic reticulum and lacks ribosomes. Chloroplast deoxyribonucleic acid (DNA) is distributed in discrete granules rather than being organized in a peripheral ring. Although photosynthetic lamellae are composed of three thylakoids, a girdle lamella is absent. The main accessory pigment is peridinin, a carotenoid unique to dinoflagellates. In some species of a few genera (*Gymnodinium*, *Peridinium*), chloroplasts are derived by an unknown process from another alga (chrysophyte or cryptophyte) and thus have uncharacteristic features. In photosynthetic dinoflagellates the main food reserve is starch, which accumulates outside the chloroplast, but storage of lipids is also common. *See* CAROTENOID; CHLOROPHYLL; PHOTOSYNTHESIS.

Members of several unrelated groups of marine dinoflagellates are parasites. The hosts are varied, comprising algae (including other dinoflagellates), protozoans, coelenterates, annelids, crustaceans, mollusks, appendicularians, and fishes. Parasitic species are highly specialized or structurally reduced, and their dinophycean affinity is often shown only by features of their motile reproductive cells and of their nucleus. Both endoparasites and ectoparasites are known.

Symbiosis involving dinoflagellates is of common occurrence in marine environments. Occasionally, colorless dinoflagellates are host to endosymbiotic algae. More commonly, dinoflagellates of the family Zooxanthellaceae are endosymbiotic in animals, including foraminiferans, radiolarians, ciliates, turbellarians, acantharians, coelenterates, and mollusks. Symbiosis is most important in coral reefs, where up to half of the carbonate in calcified structures is derived by way of photosynthesis carried out by endosymbiotic dinoflagellates.

Cell biology. The dinoflagellate nucleus is large in relation to the volume of the cell. The DNA is surrounded by a basic protein that differs from the histones of other organisms. In nearly all dinoflagellates the chromatin remains condensed, enabling the chromosomes to be seen throughout the cell cycle. During mitosis the nuclear membrane remains intact. The spindle is external, with microtubules passing through channels or tunnels in the nucleus to attach to the chromosomes. The nucleus has been termed dinokaryotic in deference to its uniqueness, or mesokaryotic with reference to its postulated intermediacy between prokaryotic and eukaryotic nuclei.

A variety of light-sensitive organelles are found in dinoflagellates. Some eyespots are free in the

cytoplasm while others are bound by membranes. In the family Warnowiaceae, complex eyelike structures called ocelli provide for the projection of an image on a retinoid. Ejectile organelles called trichocysts occur in most dinoflagellates. They are discharged as long threads when the cell is stimulated by touch, heat, or chemicals. A few members of the Polykrikaceae and Warnowiaceae have larger ejectile organelles called nematocysts because of their superficial resemblance to the nematocysts of coelenterates. A multivesicular organelle called a pusule is found near the base of the flagella of most dinoflagellates. This structure may serve an osmoregulatory or secretory function.

Bioluminescence is well known in several genera of marine dinoflagellates (*Noctiluca*, *Pyrocystis*, *Pyrodinium*, *Gonyaulax*), but its functional significance is not understood.

Reproduction. Dinoflagellates commonly reproduce by binary fission. Sexual reproduction is isogamous or anisogamous and results in the production of an encysted resting stage. The cysts are represented in the fossil record abundantly as far back as the Triassic.

Toxicity. Several dinoflagellates contain substances that are toxic to other organisms. Blooms of some of these dinoflagellates cause "red tides" that are lethal to fishes or invertebrates. Blooms of other species result in the harmless accumulation of toxic substances in the tissues of organisms that prey on the dinoflagellates. When these predators are ingested by humans, however, poisoning results. Paralytic shellfish poisoning is caused by clams or mussels that have accumulated saxitoxin produced by certain species of *Gonyaulax* or *Protogonyaulax*. Ciguatera is a poisoning caused by eating fish that have accumulated ciguatoxin derived from any of several benthic dinoflagellates. *See* ALGAE. Paul C. Silva; Richard L. Moe

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; F. J. R. Taylor, On dinoflagellate evolution, *BioSystems*, 13:65–108, 1980.

Dinosauria

The term, meaning awesome reptiles, that was coined by the British comparative anatomist R. Owen in 1842 to represent three partly known, impressively large fossil reptiles from the English countryside: the great carnivore *Megalosaurus*, the plant-eating *Iguanodon*, and the armored *Hylaeosaurus*. They were distinct, Owen said, not only because they were so large but because they were terrestrial (unlike mosasaurs, plesiosaurs, and ichthyosaurs); they had five vertebrae in their hips (instead of two or three like other reptiles); and their hips and hindlimbs were like those of large mammals, structured so that they had to stand upright (they could not sprawl like living reptiles). Owen's diagnosis was strong enough to be generally valid, with some modification, 150 years later. His intention in erecting Dinosauria, though, as A. J. Desmond noted, seems to have been more than just the recognition of a new group: By showing that certain extinct reptiles were more "advanced" in structure (that is, more similar to mammals and birds) than living reptiles, he was able to discredit contemporary evolutionary ideas of the transmutation of species through time into ever-more-advanced forms (progressivism).

Over the next several decades, dinosaurs were discovered in many other countries of Europe, for example, *Iguanodon* in Belgium and *Plateosaurus* in Germany, but rarely in great abundance. The first dinosaur discoveries in the United States were from New Jersey as early as the 1850s [*Trachodon*, 1856; *Hadrosaurus*, 1858; *Laelaps* (=*Dryptosaurus*), 1868]. The genus *Troodon*, based on a tooth from Montana, was also described in 1856, but at the time it was thought to be a lizard. Spectacular discoveries of dinosaurs from the western United States and Canada began in a great rush in the 1870s and 1880s and continued into the early 1900s. From 1911 to 1914, expeditions into German East Africa discovered some of the largest dinosaurs ever collected: *Brachiosaurus* and *Tornieria*, among many others. In the 1920s, expeditions into the Gobi desert of Mongolia brought back new and unusual dinosaurs such as *Protoceratops*, and the first eggs immediately recognizable as dinosaurian. For various reasons, such grand expeditions dwindled until the late 1960s, when renewed activities in the western United States and Canada, Argentina, southern Africa, India, Australia, China, Mongolia, and even Antarctica uncovered dozens of new dinosaurs. The rate of new discoveries reached an estimate of one new find every 6 weeks (intriguing new ones include *Microraptor*, *Mononykus*, *Eoraptor*, and *Brachytrachelopan*) during the mid-1990s, a pace of discovery that has continued (or possibly been exceeded) in the early twenty-first century. The finds since the 1960s may represent up to 50% of all dinosaurs currently known to have existed.

Hip structure. As dinosaurs became better known, their taxonomy and classification developed, as well as their diversity. In 1887, H. G. Seeley recognized two quite different hip structures in dinosaurs and grouped them accordingly. Saurischia, including the carnivorous Theropoda and the mainly herbivorous, long-necked Sauropodomorpha, retained the generalized reptilian hip structure in which the pubis points down and forward and the ischium points down and backward. The remaining dinosaurs have a pubis that has rotated to point down and backward, thus extending parallel to the ischium; this reminded Seeley of the configuration in birds, and so he named this group Ornithischia. However, the ornithischian pubis is only superficially similar to that of birds, which are descended from, and are thus formally grouped within, Saurischia. Seeley's discovery, in fact, only recognized the distinctiveness of Ornithischia, and he concluded that Saurischia and Ornithischia were not particularly closely related. Even within Saurischia, there were general doubts that Sauropoda and Theropoda had any close

relationship; eventually, the word "dinosaur" was used mainly informally by paleontologists.

However, this situation has been reversed. In 1974, Robert Bakker and Peter Galton argued that there were a great many unique features, including warm-bloodedness, which characterized the dinosaurs as a natural group (also including their descendants, the birds). Although this scheme was debatable in some particulars, it spurred renewed studies anchored in the new methodology of cladistic analysis of phylogenies (evolutionary relationships). A 1986 analysis listed nine uniquely derived features (synapomorphies) of the skull, shoulder, hand, hip, and hindlimb that unite Dinosauria as a natural group; this analysis has been since modified and improved, and today Dinosauria is universally accepted as a natural group, divided into the two monophyletic clades, Ornithischia and Saurischia. *See* ORNITHISCHIA; SAURISCHIA.

Earliest dinosaurs. Dinosaurs are archosaurs, a group that comprises living crocodiles, birds, and all of the living and extinct descendants of their most recent common ancestor (**Fig. 1**). The closest relatives of dinosaurs, which evolved with them in the Middle and Late Triassic (about 240–225 million years ago), include the flying pterosaurs and agile, rabbit-sized forms such as *Lagosuchus* and *Lagerpeton*. The common ancestor of all these forms was small, lightly built, bipedal, and probably an active carnivore or omnivore. Somewhat larger, with skulls ranging 15–30 cm (6–12 in.) in length, were *Eoraptor* and *Herrerasaurus* from the Late Triassic of Argentina and *Staurikosaurus* from the early Late Triassic of Brazil (**Fig. 2**). When the latter two genera were first described in the 1960s, they were thought to be primitive saurischian dinosaurs. The relationships of these animals are contentious, however, with some authors placing them within Theropoda, others positioning them at the base of Saurischia, and a further group of specialists suggesting that they were outside the group formed by Saurischia plus Ornithischia. These Late Triassic genera testify to a burst of evolutionary change at this interesting time in vertebrate history, and show that a variety of taxa very close to the origin of dinosaurs appeared during this interval. The first definite ornithischians and saurischians appear at almost the same time, though dinosaurs remained generally rare and not very diverse components of terrestrial faunas until the beginning of the Jurassic Period (about 200 million years ago). *See* JURASSIC; TRIASSIC.

An area of great interest is how the dinosaurs and their closest relatives differ from their contemporaries. Their posture and gait hold some important clues. Like *Lagosuchus*, *Lagerpeton*, and their other close relatives, the first dinosaurs stood upright on their back legs. The head of the thigh bone (femur) angled sharply inward to the hip socket, which was slightly open (not backed by a solid sheet of bone as in other reptiles). The femur moved like a bird's, in a nearly horizontal plane; the shin bone (tibia) swung back and forth in a wide arc, and the fibula (the normally straplike bone alongside the tibia) was reduced

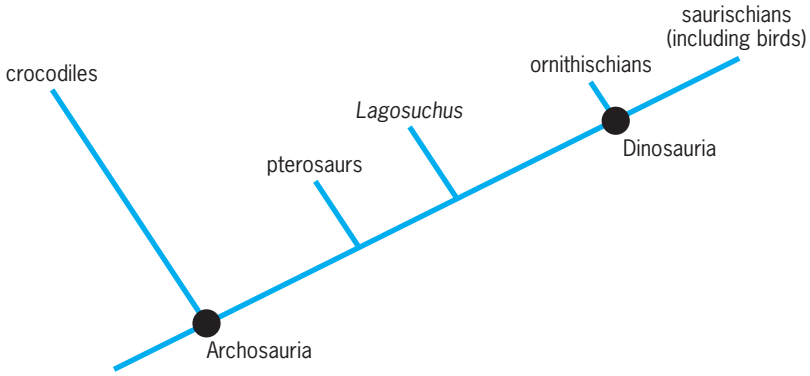

Fig. 1. Simplified phylogenetic tree of the archosaurian reptiles.

because the lower leg did not rotate about the knee, as a crocodile's or lizard's does. The ankle, too, had limited mobility: it formed a hinge joint connecting the leg to long metatarsals (sole bones), which were raised off the ground (in a stance called digitigrady, or "toe-walking"). All of these features can be seen today in birds, the living descendants of Mesozoic dinosaurs. Because the first dinosaurs were bipedal, their hands were free for grasping prey and other items, and the long fingers bore sharp, curved claws. The neck was long and S-shaped, the eyes large, and the bones lightly built and relatively thin-walled.

Ornithischia

Ornithischians (**Fig. 3**) are a well-defined group characterized by several unique evolutionary features; the entire group was analyzed cladistically in 1986, and the resulting phylogeny has been the basis of all later work. Ornithischians have a predentary bone, a toothless, beaklike addition to the front of the lower jaw that, like the front of the upper jaw, probably had a horny covering in life (**Fig. 4**). This appears to have been an adaptation for plant eating. The jaw joint was set below the occlusal plane, a nutcracker-like arrangement interpreted as increasing leverage for crushing plant material. The teeth were set in from the side of the jaw, suggesting the presence of fleshy cheeks to help sustain chewing. The cheek teeth were broad, closely set, and leaf-shaped, and were often ground down to a shearing surface. In the hip, the pubis pointed backward. In all but the most generalized ornithischians, a new prong on the pubis

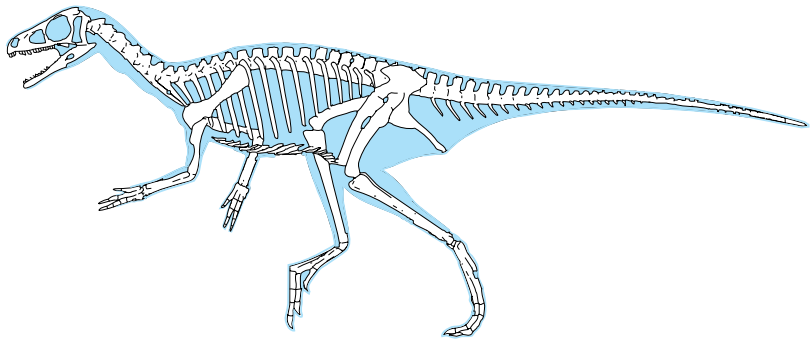

Fig. 2. *Eoraptor*, an early dinosaur or dinosaurian relative from the Late Triassic of Argentina.

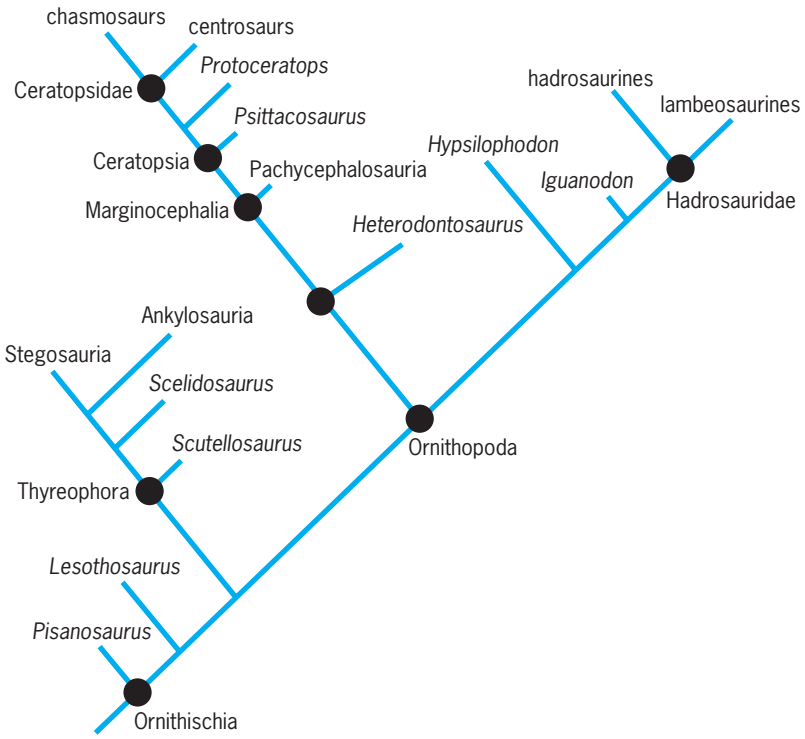

Fig. 3. Relationships of the ornithischian dinosaurs. The base of this diagram is linked to Fig. 12. (*Modified from P. C. Sereno, The evolution of dinosaurs, Science, 284:2137–2147, 1999*)

was developed from the hip socket, upward, forward, and outward. This may have provided a framework to support the guts or to anchor the hindlimb muscles.

The most generalized ornithischians known are the fragmentary *Pisanosaurus* from the Late Triassic of Argentina and the small *Lesothosaurus* from the Early Jurassic of South Africa. In the major ornithischian radiation, Thyreophora branch off first, and Cerapoda are divided into Ornithopoda and Marginocephalia. All known ornithischians are either omnivores or herbivores. *See* ORNITHISCHIA.

Thyreophora. Stegosaurs and ankylosaurs were distinguished by their elaborate armor of modified dermal bone. The earliest thyreophorans, such as the small, bipedal *Scutellosaurus* and the slightly larger, probably quadrupedal *Scelidosaurus* (both from the Early Jurassic), had such scutes all over their bodies, but not as elaborate or distinctive as those of the larger, fully quadrupedal stegosaurs and ankylosaurs.

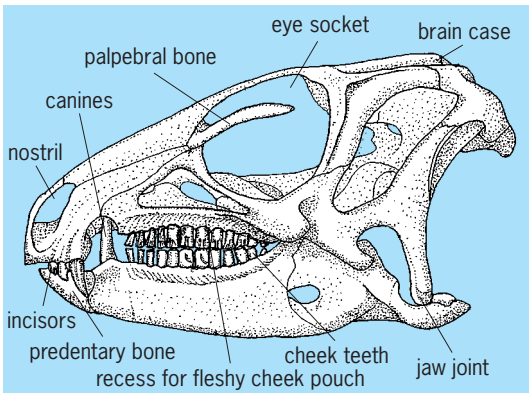

Fig. 4. Skull of *Heterodontosaurus*, showing typical ornithischian features.

Stegosaurs. Stegosaurs first appear in the Middle Jurassic of China and Europe, and were undoubtedly widespread by then, though the group never seems to have been particularly diverse: 14 genera are known, mostly from the Late Jurassic and Early Cretaceous. They had the reduced ancestral armor pattern of Thyreophora, losing the scutes on the sides of the body and retaining only a row along either side of the vertebral column (**Fig. 5**). Originally this paravertebral armor was spiky or conelike (*Huayangosaurus*), but in some forms (*Kentrosaurus*) the armor forward of the mid-dorsal region became platelike polygons, and in *Stegosaurus* the now much-enlarged, subtriangular plates occupied all but the last two terminal tail positions, which were still spiky. All well-known stegosaurs (except *Stegosaurus* itself) seem to have had a long spine projecting upward and backward from the shoulder as well. *Stegosaurus* also had a complement of bony ossicles in its throat region. In all stegosaurs the heads were relatively small and the teeth few and diamond-shaped. Stegosaurs had small brains, ranging from a walnut to a billiard ball in size. An expansion of the neural canal in the hip region was once interpreted as a "second brain" to control the hindlimbs and tail, but it is only a cavity to accommodate the sacral nerve plexus and a large glycogen ("fat") storage organ like those of birds and some other reptiles. *See* JURASSIC; CRETACEOUS.

The distinctive armor of stegosaurs has long

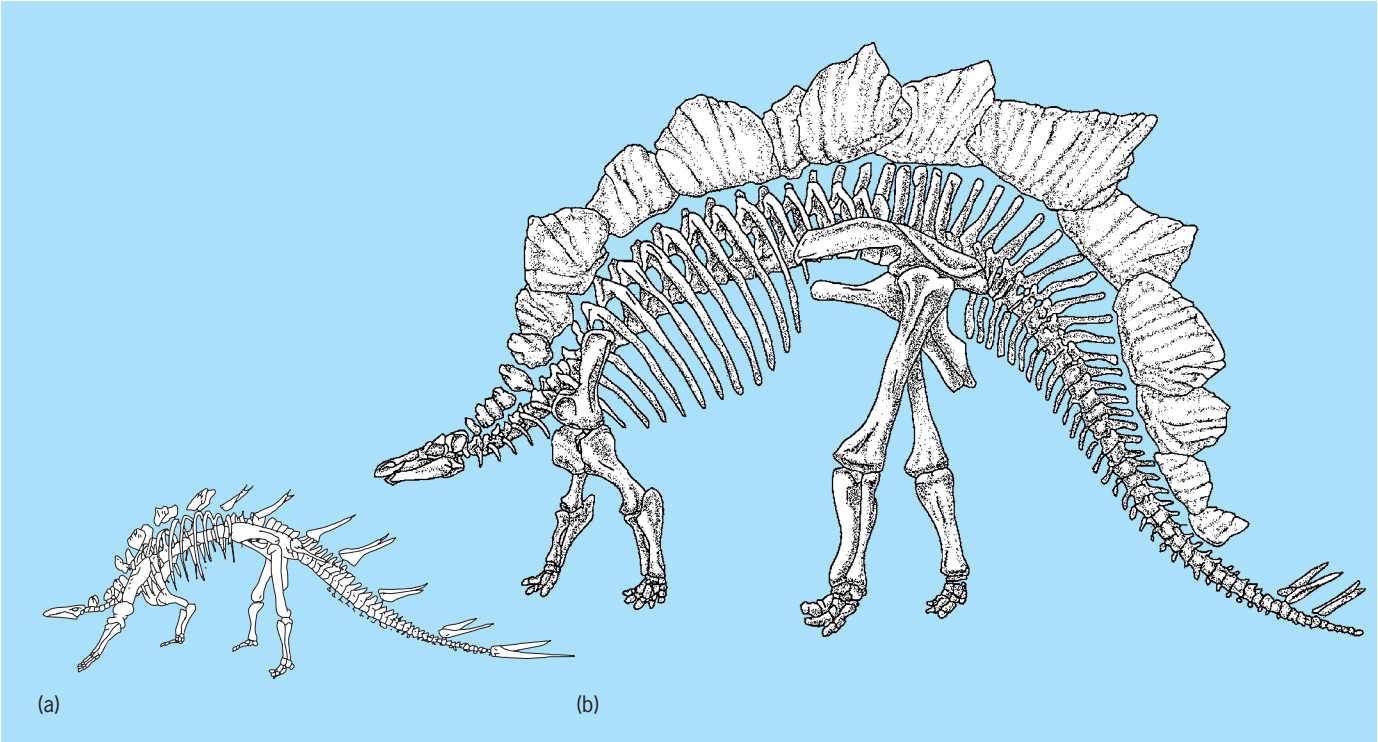

Fig. 5. Stegosaurs. (*a*) *Kentrosaurus*. (*b*) *Stegosaurus*.

fueled speculation about its position, arrangement, and function. The spikes appear to have been primitively paired, though the 17 plates of *Stegosaurus* form a single alternating, nearly medial row. Sharpey's fibers, which reflect the direction of muscle attachment to bone, show that these plates stood upright, not flat, and could not have been moved up and down. Defense is a possible function, although it must be asked why stegosaurs lost all but their paravertebral row of armor. Stegosaurs were not fast runners: the lower limbs were long and the feet and upper limbs were short, so they would have had to stand and deliver. The spikes and plates, like the shoulder spines, could have had a passive function in defense, and the terminal tail spikes, which dragged along the ground, might have made a formidable weapon. However, the plates are relatively weak "sandwiches" of latticelike bone internally, not optimal for defense. In some large stegosaurs, the broad surface area of the plates may have contributed to thermoregulation, though this hypothesis has been challenged. However, it is most likely that they were used in species recognition: no two stegosaurs have identical armor.

Ankylosaurs. Ankylosaurs are distinctive dinosaurs (**Fig. 6**). Their skulls are short, low, and flat, always

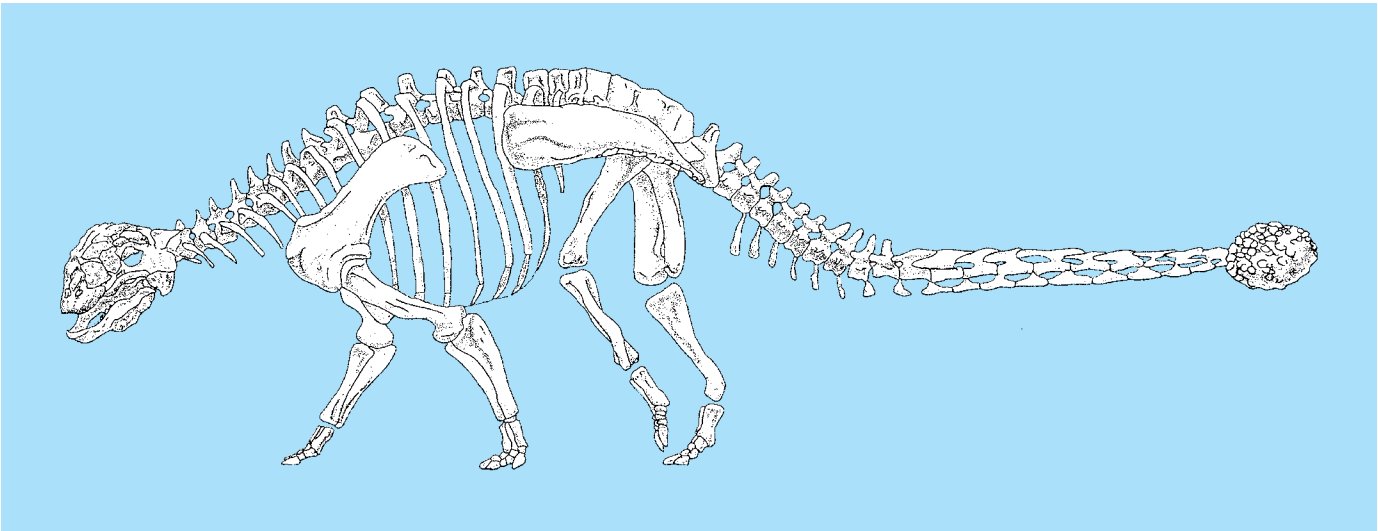

Fig. 6. *Euoplocephalus*, an ankylosaur.

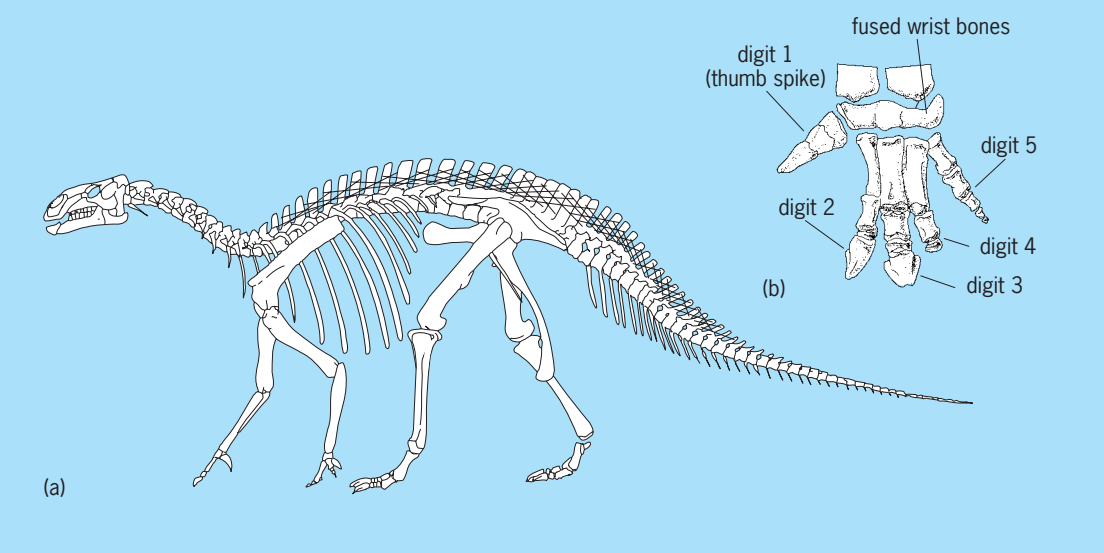

Fig. 7. ***Iguanodon*****, an ornithopod. (*a*) Skeleton. (*b*) Hand.**

wider than high in rear view, and a complement of dermal ossicles covers the skull and closes the antorbital and mandibular fenestrae. There is an S-shaped row of small, uncomplicated teeth, which sometimes display wear. The body is broad and squat, and the limbs short. Armor covers much of the body, as in Thyreophora primitively, and appears in several shapes, including keeled or spined plates, polygonal pustules, spikes, and rows of symmetrical ovals or rounded rectangles.

Ankylosauria is divided into Nodosauridae and Ankylosauridae, based on a number of distinctive but subtle features of the skull and its plating, plus several postcranial characteristics. Several authors now recognize a third group, Polacanthidae. There are not many gross differences, though ankylosaurids have a tail club (four armor plates envelop the last series of tail vertebrae, which are deeply internested and partly fused). Both groups are primarily Cretaceous, though several genera are now known from

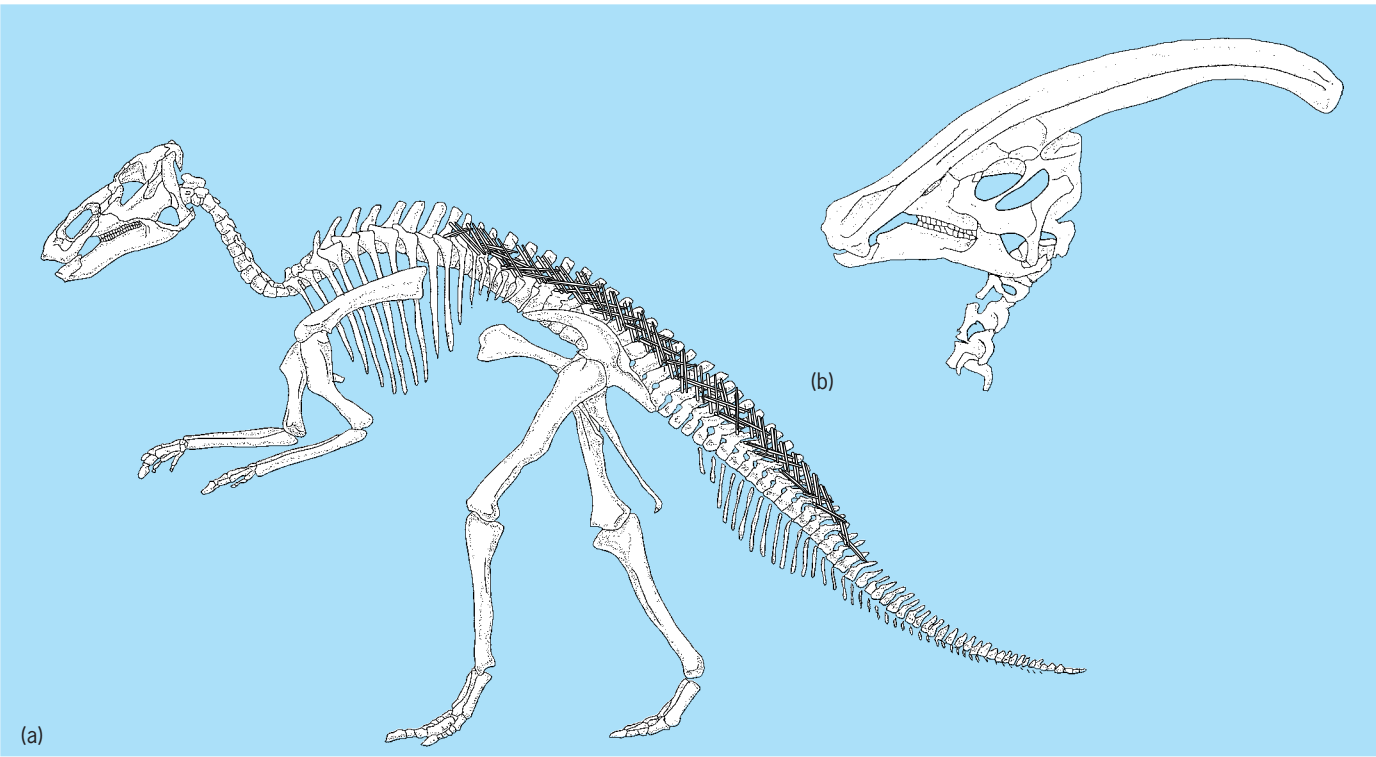

Fig. 8. Hadrosaurs. (*a*) Skeleton of the hadrosaurine *Edmontosaurus*. (*b*) Side view of skull of the lambeosaurine *Parasaurolophus*.

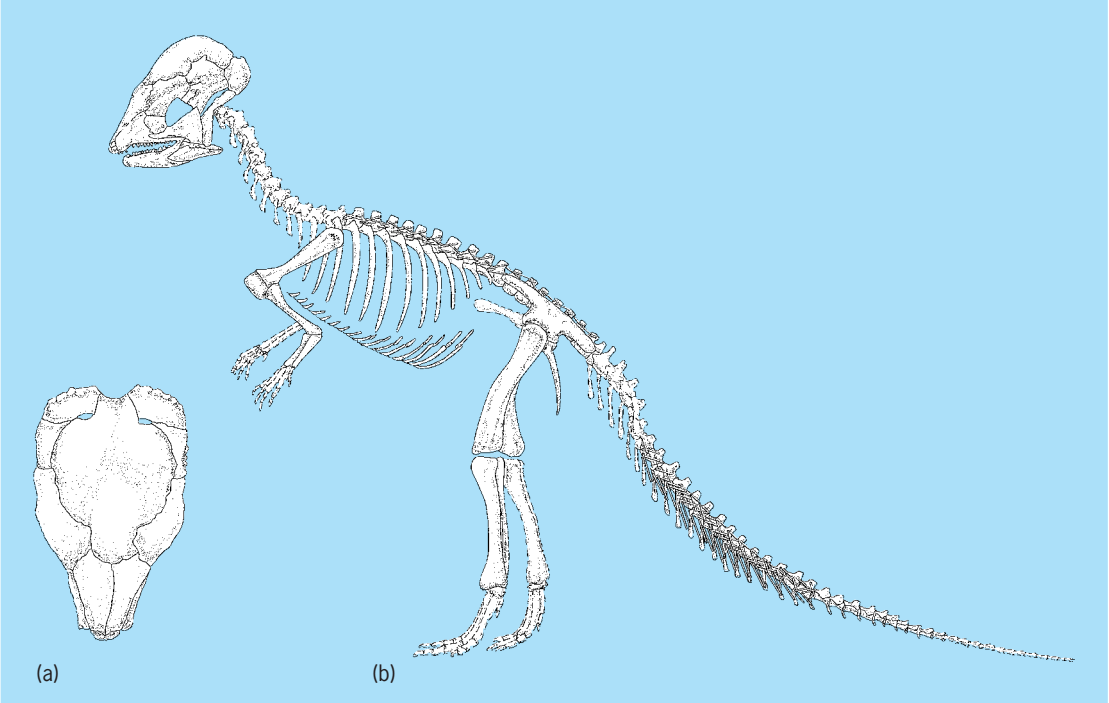

Fig. 9. ***Stegoceras*, a pachycephalosaur. (*a*) Top view of skull. (*b*) Skeleton.**

the Jurassic. Ankylosaurs, like stegosaurs, seem never to have been very diverse, and they may have been solitary or have traveled in small family groups. They could not have been fast runners, and may have depended on camouflage, squatting, or slashing with lateral spines or tail club to repel predators.

Cerapoda: Ornithopoda. In many ways, Ornithopoda includes the most generalized ornithischians, which retain the bipedal posture of ancestral dinosaurs while lacking the horns, frills, spikes, and armor of other ornithischians. Nonetheless, they have their own specializations, and through the Mesozoic Era a general progression in size and adaptation can be seen from heterodontosaurs through hypsilophodonts to iguanodonts and hadrosaurs. In ornithopods, the upper front (premaxillary) teeth are set below the cheek (maxillary) teeth, and the jaw joint is set substantially below the occlusal plane of the teeth. *Heterodontosaurus* (Fig. 4), a fox-sized, lightly built form from the Early Jurassic of South Africa, shows these general characters and a few unusual ones, including protruding "canine" teeth presumably for defense or display.

The remaining three ornithopod groups are somewhat larger in size; they lose progressively all the front (premaxillary) teeth, reduce and/or lose the antorbital and maxillary fenestrae, and reduce the third finger (the fourth and fifth are already reduced in ornithischians). *Hypsilophodon* is a generalized member of this group. It still has five pairs of premaxillary teeth, and there is a gap between its front and cheek teeth, but the cheek teeth are now closely set and their crowns are typically worn flat. It was once thought that *Hypsilophodon* lived in trees, but this was based on the presumption that the first toe was reversed (as in birds) to serve as a perching adaptation. Actually, the anatomy of the foot is normal, and there are no additional reasons to put *Hypsilophodon* in trees.

More derived (advanced) ornithopods, such as *Dryosaurus* and *Tenontosaurus*, approach the condition of iguanodonts and hadrosaurs in many respects: the front teeth are completely lost; the front

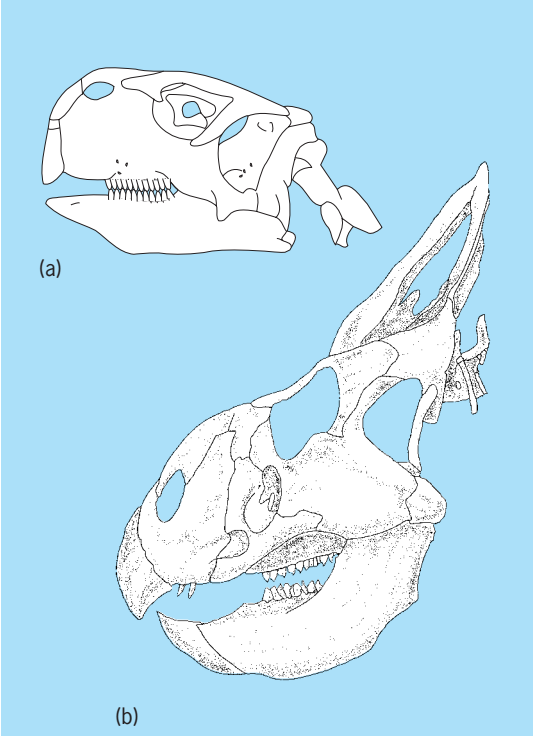

Fig. 10. Ceratopsia. (*a*) Skulls of *Psittacosaurus* and (*b*) *Protoceratops*.

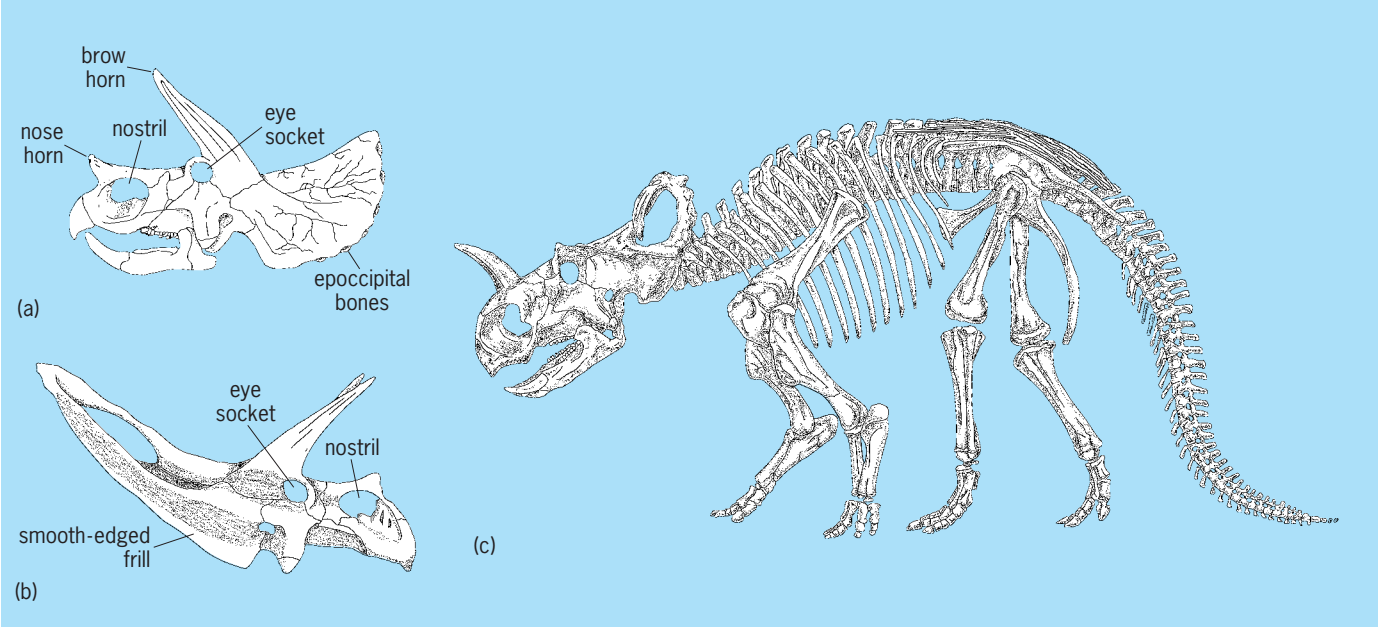

Fig. 11. Ceratopsia. (*a*) Skulls of *Triceratops* and (*b*) *Torosaurus*. (*c*) Reconstruction of *Centrosaurus*.

of the beak begins to flare; the cheek teeth become larger and more closely set, creating a single grinding battery of teeth; and the "bridge" of the nose (the area between the eyes and nostrils along the skull midline) becomes more highly arched. The forward prong of the pubis begins to grow outward until it reaches farther forward than the front of the ilium. As the cheek teeth begin to unify into opposing dental batteries, the internal kinesis (mobility) of the skull bones against each other also evolves to resist and redirect the forces of chewing.

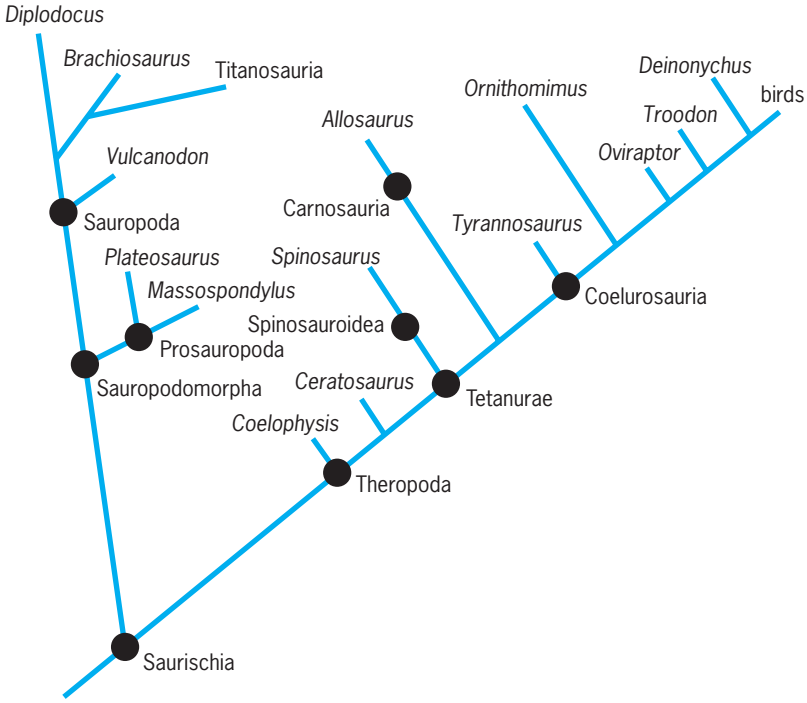

Fig. 12. Relationships of the saurischian dinosaurs. The base of this diagram is linked to Fig. 4. (*After theropod and sauropodomorph phylogenies, presented by various authors, in D. B. Weishampel et al., eds., The Dinosauria, 2d ed., University of California Press, 2004*)

The iguanodonts (**Fig. 7**) and hadrosaurs (**Fig. 8**) are the most familiar ornithopods, and were among the first dinosaurs discovered; they are still among the best represented in the fossil record. These two groups continue the trends already seen in the hypsilophodonts and dryosaurs, gaining larger size, losing the front teeth, flaring the snout further, and arranging the cheek teeth in a long, straight row for efficient crushing and slicing. In iguanodonts and hadrosaurs, the pubis changes somewhat: its original (backward) prong is reduced to a splint, and the forward prong enlarges and flares. This may reflect support for the intestinal tract, which is thought to have needed an extensive system of fermentation for the tough, low-nutrient plants that the animals ate. The claws on the fingers and toes are flattened and broadened into spoon-shaped hooves, and the middle three fingers, as with the middle three toes, may have functioned as a unit in walking on four legs. The fifth finger in these dinosaurs is quite distinctive: it points to the side, much like a thumb, and was similarly opposable to the other fingers, probably for manipulating food. The thumb is extraordinary in iguanodonts: the thumb phalanx, its terminal ungual, and its supporting metacarpal bone are fused together into a conelike spike, and this in turn is articulated tightly to the blocky, nearly immobile wrist bones. Whatever function this had in scratching, procuring food, or even defense, it had no function at all in hadrosaurs, which lost the thumb altogether.

Hadrosaurs (known as duckbills) dominated most of the Late Cretaceous of the Northern Hemisphere in numbers and diversity; their bones are found by the thousands in some deposits, suggesting mass deaths, perhaps during seasonal migrations or droughts. Their jaws were formidable food

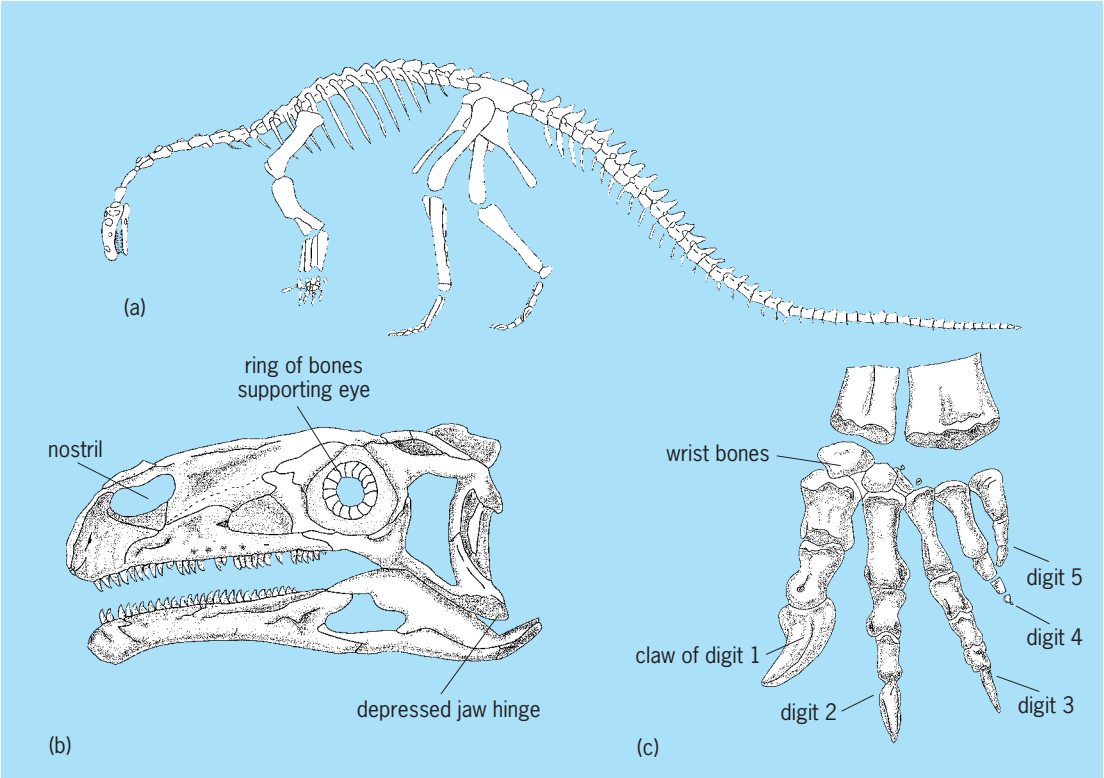

Fig. 13. Sauropodomorpha. (*a*) Skeleton of *Anchisaurus*. (*b*) Skull and (*c*) hand of *Plateosaurus*.

processors. Four or five rows of replacement teeth accompany each cheek tooth, and all of them are closely set, so as the teeth wore down they formed a single cutting and crushing surface of hundreds of teeth in each quadrant of the jaws. As in all cerapodans, the enamel was thicker on the oral side of the lower teeth and the cheek side of the upper teeth, so the teeth wore down unevenly: the cutting surface of the dental batteries was not horizontal, but beveled for better slicing as well as crushing.

Hadrosaurs are generally divided into two groups (Fig. 8): the more conservative hadrosaurines and the more derived lambeosaurines, which are noted for their profusion of skull crests and ornaments. Some crests were solid and spikelike (*Saurolophus*), while others were partly hollow and resembled bizarre hats or helmets (*Lambeosaurus*, *Corythosaurus*) or even snorkels (*Parasaurolophus*). The partial hollowness in the crests, which were formed of the premaxillary and nasal bones, was to accommodate the nasal passages that normally run through these bones. Many explanations have been proposed for this great variety in skull structure, including functions in head butting, underwater feeding, or improvement in sense of smell, but most have little support. Because lambeosaurines vary little in other skull and postcranial features, a plausible explanation is that the crests were for display and species recognition, like the horns of present-day hoofed mammals. Experimental analysis and acoustical modeling of hollow-crested forms suggested that the chambers in the crests could have served as vocal resonators to create distinctive and far-reaching sounds for species recognition and communication. This idea has contributed to the picture of sophisticated social organization that is developing for many groups of dinosaurs.

Cerapoda: Marginocephalia. Marginocephalia includes the dome-headed Pachycephalosauria (**Fig. 9**) and the frilled, usually horned Ceratopsia (**Fig. 10**). Both groups are almost exclusively Cretaceous, though a few Late Jurassic forms are now known. Marginocephalia are recognized by a small shelf or incipient frill that overhangs the back of the head; also, their hips are wider than in most other forms, though ankylosaurs have extremely wide (but very different) hips. The small shelf on the skull was the basis for the elaboration of very distinct features and functions in both Pachycephalosauria and Ceratopsia. Known from a partial postcranial skeleton from the earliest Cretaceous of Germany, *Stenopelix* is possibly an early marginocephalian.

Pachycephalosaurs (Fig. 9) are distinguished by their thick skulls, which are often ornamented with knobs and spikes. Originally, the roof of the skull was rather flat but sloped upward posteriorly and was only slightly thickened (*Homalocephale*); the two upper temporal openings on top of the skull were still prominent. In more derived forms (*Prenocephale*, *Stegoceras*, *Pachycephalosaurus*), the skull roof became rounded into a dome about the shape of a human kneecap but much thicker, and the temporal openings were covered by bone. Rows and clusters of barnacle-shaped bony knobs adorn the nose, the temporal region, and the posterior shelf in most forms, and these can be developed into long

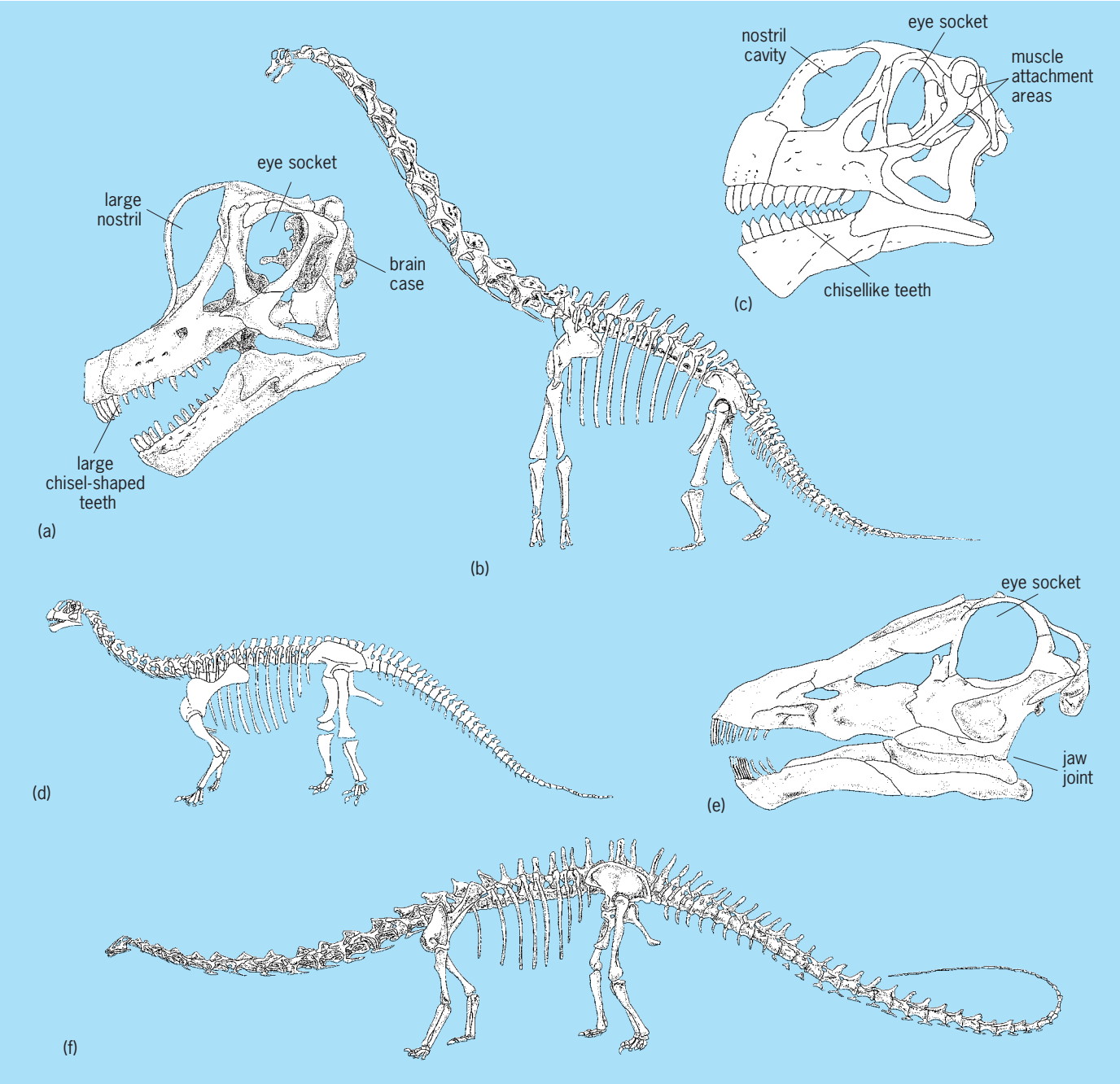

Fig. 14. Skulls and skeletal outlines of sauropods. (*a*, *b*) *Brachiosaurus*. (*c*, *d*) *Camarasaurus*. (*e*, *f*) *Diplodocus*.

spikes surrounding the dome (*Stygimoloch*). The thick, columnar bone that makes up the dome has often been interpreted as a shock absorber for head butting between competing males. However, more recent studies suggest that head-butting was quite unlikely in most forms, because the domes were so poorly designed for it. The rounded shape would cause strong torque on the neck; the dome may not have been mechanically strong enough for direct head impact; and early pachycephalosaurs had flat skull roofs that appear to have been more suited for pushing than ramming. The shape of the dome varied considerably in pachycephalosaurs, and so did its ornamentation; one pachycephalosaur had a dome with a sharp medial ridge and large spikes protruding from the sides of the head. Generally strong construction of the skull, the reinforced vertebral column (equipped with pencil-thick ossified tail tendons for support), and the strengthened pelvis suggest highly developed intraspecific (and possibly interspecific) use, perhaps mainly in merely displaying skulls to avoid direct combat.

Ceratopsia. Ceratopsia (Fig. 10), along with the hadrosaurs, were the most successful Late Cretaceous ornithischians. They are distinguished by the presence of a rostral bone, a unique, toothless, neomorphic bone that caps the front of the upper beak much as the predentary caps the lower beak in

ornithischians. Their earliest members, the basal (primitive) ceratopsians and psittacosaurs of the Late Jurassic and Early Cretaceous, were bipedal, slightly over 1 or 2 m in length (3–6 ft), and had a parrotlike beak; they were presumably browsers on low plants. The protoceratopsids initiated a trend to slightly larger size (2–3 m or 6–9 ft), quadrupedality, larger skulls, a more pronounced beak, a bony bump above the nose, and an enlarged frill at the back of the skull that usually appears as an upward-sloping arch formed by the squamosal and supporting bones.

Ceratopsidae, the great ceratopsians of the latest Cretaceous, include the centrosaurs and the chasmosaurs (**Fig. 11**). The centrosaurs generally had prominent nose horns but lacked horns above the eyes; the chasmosaurs emphasized the eye horns at the expense of the nose horns. Both groups evolved dental batteries of multiple rows of beveled teeth strikingly similar to the dentitions of hadrosaurs, and as in hadrosaurs the claws were flattened into spoon-shaped hooves. The frills were often ornamented with knobs and spikes and, as in hadrosaurs, are the most distinctive and differentiated parts of the body, the postcrania being rather conservative and not highly varied among ceratopsids. This also speaks to a function in display, rather than defense, for the great frills, which (with the notable exception of the short-frilled *Triceratops*, perhaps the best-known ceratopsian) were not solid shields but merely bony arches—hardly a suitable defense against a marauding *Tyrannosaurus* or even another ceratopsid's horns. Species recognition is a more likely use.

Saurischia

Saurischia (**Fig. 12**) includes Sauropodomorpha and Theropoda. For many years it was doubted that Saurischia formed a natural group because its two subgroups are so different, but in 1986 J. Gauthier demonstrated its validity cladistically. Uniquely derived features of Saurischia include the long neck vertebrae; the long asymmetrical hand in which the second digit is the longest; the slightly offset thumb with its short basal metacarpal, robust form, and large claw; and several other features of the skull and vertebrae. Unlike ornithischians, saurischians are fairly well represented in Late Triassic faunas as well as in the later Mesozoic. *See* SAURISCHIA; MESOZOIC.

Sauropodomorpha. This group includes the largest land animals of all time. The first sauropodomorphs (**Fig. 13**) elaborated several basic saurischian features: they had rather small skulls, elongated leaf-shaped, coarsely serrated teeth, long necks (with 10 vertebrae), robust limbs with a particularly robust thumb and claw, and a trunk that is longer than the hindlimb. *Thecodontosaurus* and *Massospondylus* typify this evolutionary stage and were probably mostly bipedal, with lengths of 2–6 m (6–18 ft). Soon, larger forms such as *Plateosaurus* appeared, and the trend toward larger size continued with the melanorosaurids of the Late Triassic and Early Jurassic. The necks became progressively longer, the bodies more massive (up to 10 m or 30 ft), and the limbs more robust, until with forms like *Vulcanodon* the condition seen in the great Sauropoda proper was attained. The relationships of these early forms, collectively called Prosauropoda, are controversial—some authors regard them as a natural group, others suggest that they are ancestral to sauropods.

Sauropoda (**Fig. 14**) includes over 100 valid genera, which are split into a number of groups. A variety of primitive forms lived during the Late Triassic to Middle Jurassic; from the Middle Jurassic onward, the more advanced neosauropods appeared. Sauropods are, in general, much larger animals than prosauropods and have at least 12 neck

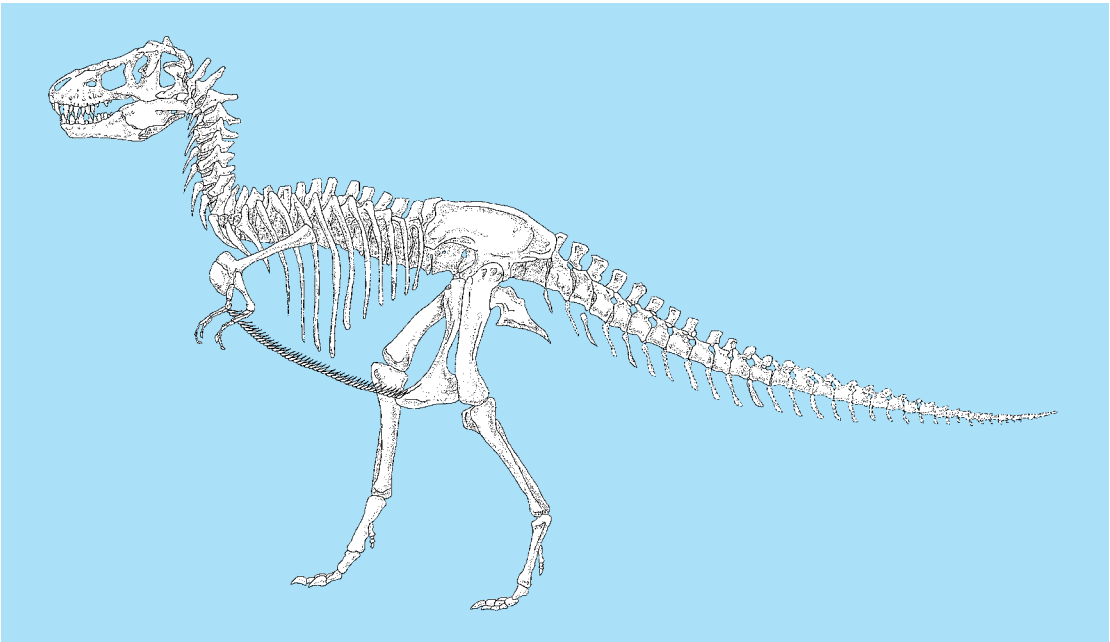

Fig. 15. ***Tyrannosaurus rex*, one of the largest known theropods.**

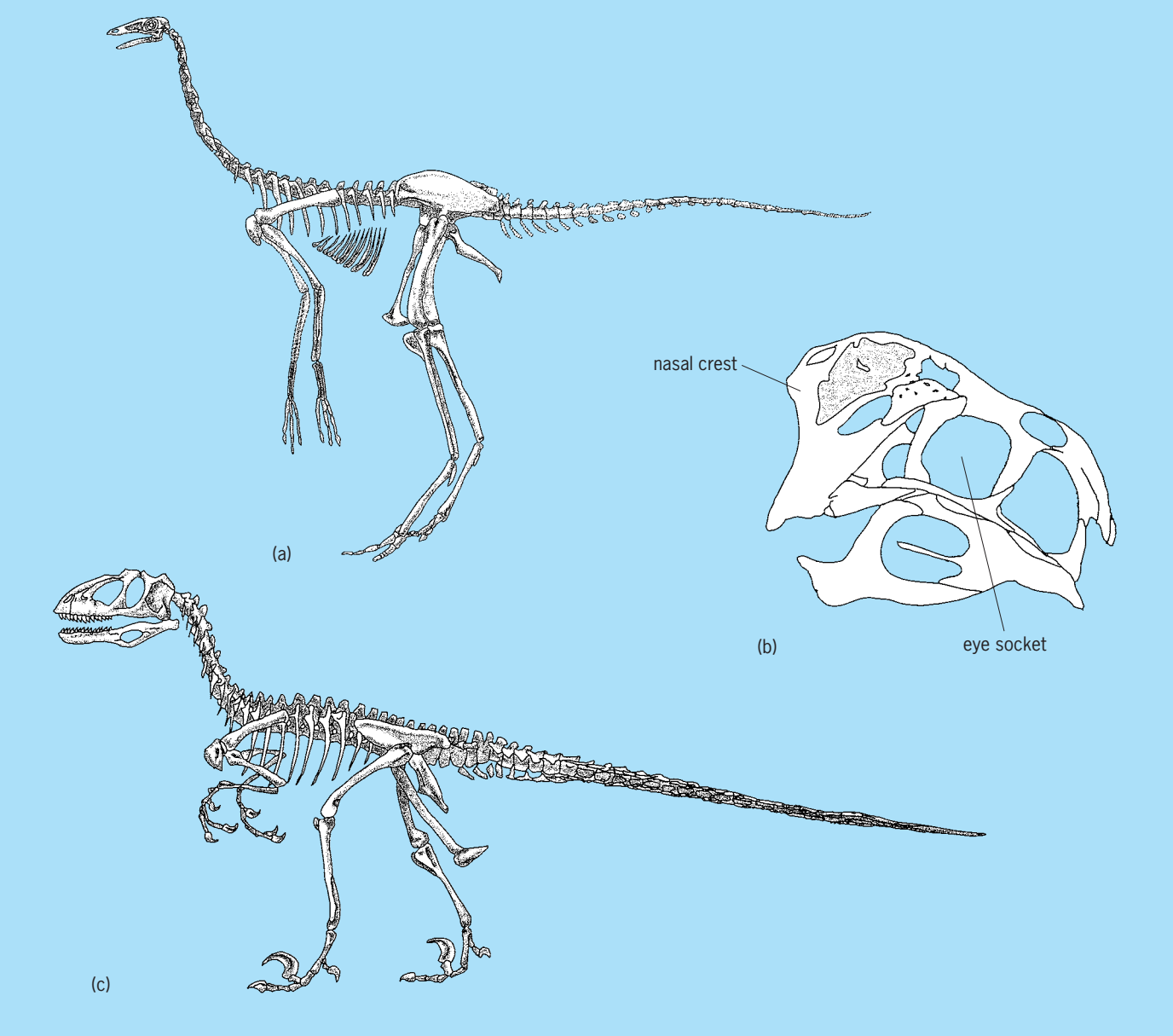

Fig. 16. Coelurosauria. (*a*) Ostrich dinosaur, *Ornithomimus*. (*b*) Skull of *Oviraptor*. (*c*) Skeleton of the dromaeosaur *Deinonychus*.

vertebrae (with a complex series of bony laminae and struts) and stout, columnar limbs. Most of the familiar sauropod groups (for example, diplodocids, brachiosaurids, and titanosaurs) are neosauropods and can be distinguished from their more primitive relatives by a number of features, including an absence of tooth serrations and a reduction in the number of wrist bones. Each sauropod group has distinctive features. For example, the diplodocids had forward-sloping teeth similar to the prongs of an iron rake, very long necks, and long, whiplike tails. About halfway along the tail, the tail chevron bones beneath the vertebrae flatten into a canoelike shape that has been interpreted as a possible adaptation for rearing up on the back legs (tripodlike) to feed in higher branches. The brachiosaurs, by contrast, had tails that were shorter than their necks, and the tail lacked the canoe-shaped chevrons of diplodocids; but their forelimbs were longer than their hindlimbs, which perhaps compensated for this lack and elevated the chest region, allowing habitual high browsing. Some sauropods, notably the titanosaur *Argentinosaurus*, could reach an estimated 70 tonnes (or 77 tons) in weight; *Seismosaurus*, a diplodocid, may have been up to 30 m (90 ft) or more in length.

Theropoda. Although meat eating seems to have been the primitive habit of dinosaurs and their close relatives, the theropods are the only group of dinosaurs to remain carnivorous. (A very small number of theropods became herbivorous.) Historically, they were divided into large carnosaurs and small coelurosaurs, virtually for convenience; but in 1986, Gauthier showed that Theropoda is a natural group united by many features of the skull and vertebrae.

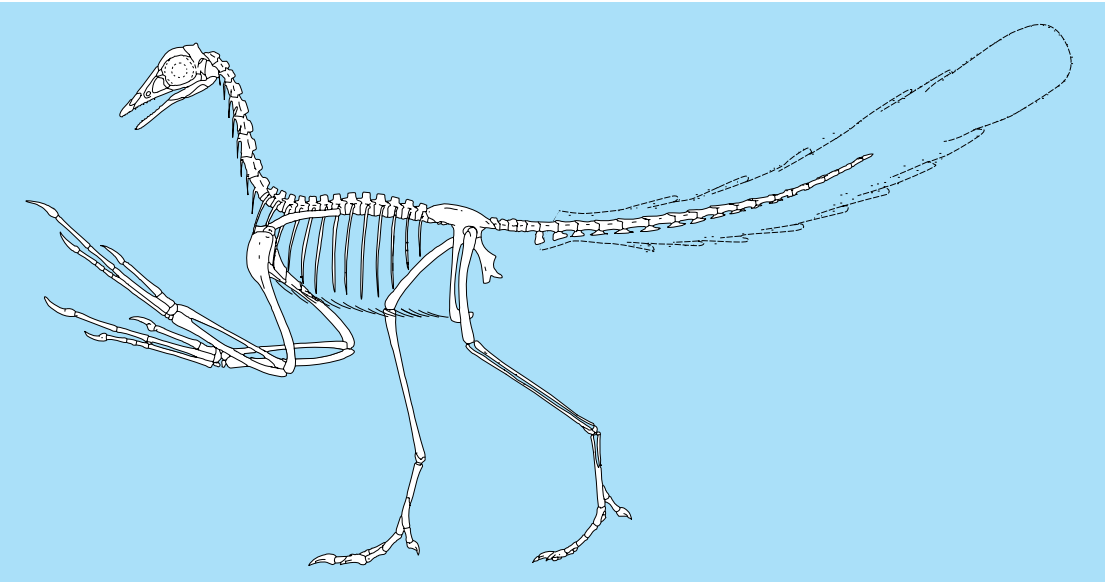

Fig. 17. Skeletal reconstruction of *Archaeopteryx*, the earliest known bird, a saurischian theropod.

In addition, the fourth and fifth fingers are lost or nearly so; the claws and teeth are sharp, and the teeth are recurved and serrated; the foot is reduced to the three middle toes; and the bones are lightly built and thin-walled.

Early in its history, Theropoda split into several lineages of primitive theropods (including coelophysoids and ceratosaurs) and a more advanced group, the Tetanurae. Primitive theropods, like the Late Triassic *Coelophysis*, were not large, but some, like the Early Jurassic *Dilophosaurus*, soon evolved to lengths of 5 m (15 ft) or more. Such ceratosaurs were the principal carnivores until the Late Jurassic. They were replaced by a great diversification of tetanuran theropods, which themselves split into three major groups: the Carnosauria (including *Allosaurus* and *Charcharodontosaurus*), the Spinosauroidea (such as *Spinosaurus* and *Baryonyx*), and the Coelurosauria (including *Ornithomimus*, *Tyrannosaurus*, *Oviraptor*, *Deinonychus*, and the birds; **Figs. 15** and **16**).

Carnosaurs had large heads with enormous, daggerlike teeth. As in nearly all theropods, the teeth were laterally compressed, with keels fore and aft for slicing, and along each keel were fine serrations that improved tearing. There is some controversy about whether large theropods were predators or scavengers. Nearly any carnivore is opportunistic enough to scavenge, and no predator takes on the strongest members of a prey species: they prefer to prey on the old, the young, and the disabled. Hence it would seem very difficult to tell the difference between adaptations for predation and those for scavenging. At any rate, few animals would want to take their chances with *Allosaurus*. Spinosauroids were similar to carnosaurs in many respects, though they possessed a number of unique specializations. For example, many forms had elongate, crocodile-like snouts, and *Spinosaurus* had tall spines on its vertebrae that might have supported a large "sail" of skin or other tissue during life.

The coelurosaurs were highly varied dinosaurs. The ornithomimids, or "ostrich dinosaurs," had reduced (or no) teeth, very long legs, and a general build reminiscent of the largest birds of today. Oviraptorosaurs had bizarrely fenestrated skulls and toothless beaks, yet the hands had sharp claws. Troodontids had long snouts with small, sharp teeth, their brains were large, and their legs and feet long. Tyrannosaurids were gigantic superpredators with massive skulls and strongly reduced forelimbs (with only two fingers per hand). The dromaeosaurs, which included the well-known *Deinonychus* and the "raptors" of the film *Jurassic Park*, had relatively large skulls and teeth, long prehensile hands and clawed fingers, a stiffened tail, and one other feature also found in some troodontids: the second (inside) toe on the foot was held off the ground and sported a trenchant, enlarged, recurved claw. In 1969 it was hypothesized that these animals would have used their long arms to grab prey, their tails to help them balance, and their large claws to rip open the bodies of their quarry. Only a few years later the Polish-Mongolian expeditions returned from the field with a pair of specimens that supported the 1969 prediction: a *Velociraptor* with its hands on the frill of a *Protoceratops*, and its feet at the prey's belly.

Dromaeosaurs were also the closest known forms to birds. The origin of birds from within dinosaurs is old: T. H. Huxley noted the strong resemblances in the 1860s. A century later J. Ostrom recognized the many unique similarities shared by *Archaeopteryx*, the earliest known bird (**Fig. 17**), and coelurosaurs, particularly forms such as *Compsognathus* and *Deinonychus*. The pubis had begun to project backward; the arms were long; the hands and fingers very long; and their digits correspond to those in the bird hand. A bony sternum had been found in some forms, and also the fused clavicles that in birds form the wishbone. The foot was

three-toed, and the first toe rotated backward as in birds. The problem was analyzed cladistically, and it was concluded that nearly 200 shared derived characters placed birds squarely within the coelurosaurian dinosaurs. In evolutionary terms, then, birds are dinosaurs, so at least one lineage of dinosaurs is not extinct. *See* ARCHAEOPTERYX.

Paleobiology

Dinosaurs laid eggs, and many nests have been found (**Fig. 18**), but matching nests and eggs to their makers is difficult unless embryos are preserved. Even then, juveniles usually lack many diagnostic characteristics of adults, so precise identification is difficult. Embryos of the hadrosaur *Maiasaura*, found in Late Cretaceous deposits from Montana, had uncalcified bones with cartilaginous joint ends, and the hatchlings were clearly helpless. Their eggshells were crushed into fragments, and the skeletons of the hatchlings found in the nests were clearly too large to have recently emerged. It appears that the young stayed around the nest and were fed by the parents until their bones fully calcified and they could fend for themselves. It is not clear how widespread either of these behaviors (or others) was among dinosaurs, but birds (living dinosaurs) are known for their extended parental care, and crocodiles, their closest living relatives, also care for their young, so it may be a general behavior among archosaurs.

Growth. If their bones are any indication, dinosaurs grew rapidly. Microscopic studies of their long bones show that the tissue was well vascularized as in birds and mammals, and that it was rapidly replaced. The texture of the bone is predominantly fibrolamellar, as in birds and mammals, and juvenile bone also shows a highly woven pattern reflecting rapid growth. Rest lines of regular occurrence have been reported in most dinosaur bones, as in those of many birds and mammals, and like these they appear to be annual. The amount of bone tissue and its many large blood vessels indicate that a lot of bone could be deposited annually, especially in juveniles and in large species. Tyrannosaurs, for example, grew twice as fast as African elephants to the same size. When dinosaurs reached full size and sexual maturity (perhaps not simultaneously), several bones co-ossified (for example, the bones that make up the braincase), as in birds and mammals. However, most dinosaur skeletons do not show these features, so they may not be fully adult.

Fig. 18. Nest of ornithischian dinosaur eggs from Montana.

Social organization. Footprints provide strong evidence that many kinds of dinosaurs traveled in groups, at least occasionally, and this is supported by records of mass burials in *Coelophysis*, *Plateosaurus*, *Allosaurus*, *Centrosaurus*, and various hadrosaurs. Social organization has the benefits of protection for the young and weak and the communication of information about food and other resources. Most major groups of vertebrates and even some invertebrates have some complex forms of social organization. Whether traveling in family groups or (at least occasionally) in large herds, all dinosaurs of any appreciable size would have had to migrate to exploit food sources successfully, if only in a local area. Annual migrations may have followed seasonal patterns of weather and vegetation.

A general picture of dinosaurian social behavior—no doubt highly variable among groups—is drawn from inferences about mass remains of trackways and bones, parental care, and the great variety of skeletal features conspicuously related to intraspecific interactions. The diversity of horns, crests, domes, knobs, and frills in many dinosaur groups contrasts starkly with the relative uniformity of their postcranial skeletons. Selection for these ornaments cannot be explained in purely functional or random terms, but strong analogies can be found in the rather similar features of living birds and mammals. Communicating, recognizing members of the same species, attracting mates, and repelling rivals, as well as delivering similar signals to members of different species, are all behavioral functions of such structures in living animals. *See* SOCIOBIOLOGY.

Metabolism. Dinosaurs evolved from within Reptilia (and so, therefore, did birds), but they are as unlike living reptiles as bats and whales are unlike horses among living mammals (and they may have been just as diverse metabolically). The physiology of extinct animals can be assessed only indirectly. No one doubts that a fossil horse was a warm-blooded, endothermic homeotherm, but such statements cannot be made about the earliest mammals and their ancestors. In the same way, the evolution of thermal strategies in dinosaurs was probably mosaic, depending on the adaptations of individual groups, and should not be considered an all-or-nothing proposition of hot-bloodedness. Many of the lines of evidence discussed above suggest that Mesozoic dinosaurs were similar in behavior and activity to mammals and birds; no evidence seems to ally them physiologically to crocodiles and lizards. But dinosaurs should be taken on their own terms, not shoehorned into models of reptile or mammal physiology based on available living analogs. The recent discoveries of nonavian dinosaurs with feathers and featherlike coverings from Early Cretaceous deposits in China indicate that the thermoregulation of at least some theropods may have been birdlike, as bone growth also shows. *See* AVES; REPTILIA.

Diversity. For 160 million years during the Mesozoic Era, dinosaurs diversified into between 500 and 1000 known genera, though many of these are based on single (often partial) specimens. Through time, dinosaurian diversity increased, but the difference between preserved diversity and some estimates of projected diversity is traced in part to varying availability of the rock record plus differential exploration. Evolution in dinosaurs was rapid: few dinosaurian genera survive more than the temporal span of a typical geologic formation (a few million years), but close relatives are often observed in succeeding formations. At the beginning of the Age of Dinosaurs, the continents were just beginning to drift apart. Dinosaurs are known from every continent and are often used to suggest land connections, such as between North America and Asia in the Late Cretaceous, or the isolation of South America during much of the Cretaceous. Some groups, such as the Ceratopsidae, apparently radiated in isolation on one or another continent.

Extinction. Many causes have been proposed for dinosaur extinction at the end of the Mesozoic Era, but most (for example, reproductive sterility, plant toxin poisoning, cataracts, supernova explosion, glaciation) have no supporting evidence. The apparent sudden decline of dinosaurs is surprising in view of their long dominance. However, the Late Cretaceous was globally unsettled in many ways. Great volcanoes were exploding in India; the Rocky Mountains were beginning to form; successive waves of regression and extinction had rocked the world's oceans; and the shallow seas were draining from many continents, including North America and Europe. These changes made the climate on land less equable and more seasonal, paving the way for destabilization of ecosystems, and perhaps creating patchy environments in which new plants and animal associations flourished. In the last few million years of the Cretaceous, known dinosaur diversity, based mainly on excellent exposures from the United States western interior, declined precipitously until, in the meters of sediment just below the Cretaceous-Tertiary boundary, only *Triceratops* and *Tyrannosaurus* survive.

An explanation is needed for this drop in diversity. Macroevolutionary patterns are based on the origination of new species and the extinction of existing ones. Dinosaur taxa were of short duration and rapid turnover; thus extinction was ubiquitous and quick throughout dinosaurian history. Toward the end of the Cretaceous, dinosaurian originations were at first burgeoning, but then they dropped; extinctions were proceeding at a near-normal pace and eventually overtook the origination rate until diversity fell to only a few taxa. In 1980, it was proposed that a giant asteroid had struck the Earth at the end of the Cretaceous, sending up a cloud of dust and vapor that blocked the sunlight for a few months or years, first lowering and then raising temperatures and causing widespread marine and terrestrial extinctions. These effects were later augmented by scenarios of global wildfires, intensely acidic rain, and other proposed environmental poisonings. As valid as the proposal of a giant asteroid impact now appears, some of the proposed biotic effects are clearly overdrawn. Dinosaurs were already declining in diversity by that time. The birds survived the impact's effects, and so did many groups of fishes, sharks, amphibians, lizards, snakes, crocodiles, and mammals. Many of these groups are notoriously sensitive to environmental disturbances. Hence any catastrophic scenarios for the terrestrial biota at the end of the Mesozoic must account both for the latest Cretaceous decline in dinosaurian diversity and the survival of any proposed environmental catastrophes by other terrestrial animals and plants. *See* EXTINCTION (BIOLOGY).

Kevin Padian; Paul M. Barrett

Bibliography. R. T. Bakker and P. M. Galton, Dinosaur monophyly and a new class of vertebrates, *Nature*, 248:168–172, 1974; P. J. Currie and K. Padian (eds.), *The Encyclopedia of Dinosaurs*, Academic Press, 1997; J. O. Farlow and M. K. Brett-Surman (eds.), *The Complete Dinosaur*, Indiana University Press, 1997; J. A. Gauthier, Saurischian monophyly and the origin of birds, *Mem. Calif. Acad. Sci.*, 8:1–55, 1986; K. F. Hirsch, K. Carpenter, and J. R. Horner (eds.), *Dinosaur Eggs and Babies*, Cambridge University Press, 1994; P. C. Sereno, The evolution of dinosaurs, *Science*, 284:2137–2147, 1999; D. B. Weishampel, P. Dodson, and H. Osmolska (eds.), *The Dinosauria*, 2d ed., University of California Press, 2004.

Diode

A two-terminal electron device exhibiting a nonlinear current-voltage characteristic. Although diodes are usually classified with respect to the physical phenomena that give rise to their useful properties, in this article they are more conveniently classified according to the functions of the circuits in which they are used. This classification includes rectifier diodes, negative-resistance diodes, constant-voltage diodes, light-sensitive diodes, light-emitting diodes, capacitor diodes, and diode lasers. *See* LASER.

Rectifier diodes. A circuit element is said to rectify if voltage increments of equal magnitude but opposite sign applied to the element produce unequal current increments. An ideal rectifier diode is one that conducts fully in one direction (forward) and not at all in the opposite direction (reverse). This property is approximated in junction and thermionic diodes. Processes that make use of rectifier diodes including power rectification, detection, modulation, and switching.

In power rectification, a rectifier diode is connected in series with an alternating-voltage supply and a load. The current through the load consists of unidirectional pulses, which can be converted into essentially constant direct current by means of a filter network that removes the varying component of current. In the detection of amplitude-modulated voltage, a similar process transforms an alternating voltage of varying amplitude into a direct voltage, the magnitude of which is proportional to the amplitude

of the alternating input voltage. In amplitude modulation, diode rectification is used to vary the amplitude of a carrier voltage in response to a signal voltage. *See* AMPLITUDE-MODULATION DETECTOR; AMPLITUDE MODULATOR; ELECTRIC FILTER; ELECTRONIC POWER SUPPLY; RECTIFIER; SEMICONDUCTOR RECTIFIER.

Diode switching circuits include circuits in which output is obtained only in the presence or absence of one or more impressed control voltages. This is accomplished by the use of diodes that are normally maintained in the nonconducting (or conducting) state and are made conducting (or nonconducting) by the control voltages. Thus the diodes either provide a connection between the input and the output terminals or short-circuit the output terminals. Diodes can also be used to connect or disconnect two or more resistances in series or parallel between two terminals at selected values of input voltage and thus to synthesize a circuit that has a desired nonlinear current-voltage characteristic. *See* SWITCHING CIRCUIT.

Negative-resistance diodes. Negative-resistance diodes, which include tunnel and Gunn diodes, are used as the basis of pulse generators, bistable counting and storage circuits, and oscillators. *See* NEGATIVE-RESISTANCE CIRCUITS; OSCILLATOR; TUNNEL DIODE.

Constant-voltage diodes. Breakdown-diode current increases very rapidly with voltage above the breakdown voltage; that is, the voltage is nearly independent of the current. In series with resistance to limit the current to a nondestructive value, breakdown diodes can therefore be used as a means of obtaining a nearly constant reference voltage or of maintaining a constant potential difference between two circuit points, such as the emitter and the base of a transistor. Breakdown diodes (or reverse-biased ordinary junction diodes) can be used between two circuit points in order to limit alternating-voltage amplitude or to clip voltage peaks. Severe clipping of a sinusoidal voltage wave produces an approximately rectangular voltage wave. *See* LIMITER CIRCUIT.

Light-sensitive and light-emitting diodes. Light-sensitive diodes, which include phototubes, photovoltaic cells, photodiodes, and photoconductive cells, are used in the measurement of illumination, in the control of lights or other electrical devices by incident light, and in the conversion of radiant energy into electrical energy. Light-emitting diodes (LEDs) are used in the display of letters, numbers, and other symbols in calculators, watches, clocks, and other electronic units. *See* LIGHT-EMITTING DIODE; PHOTOELECTRIC DEVICES; PHOTOTUBE; PHOTOVOLTAIC CELL.

Capacitor diodes. Semiconductor diodes designed to have strongly voltage-dependent shunt capacitance between the terminals are called varactors. The applications of varactors include the tuning and the frequency stabilization of radio-frequency oscillators. *See* JUNCTION DIODE; MICROWAVE SOLID-STATE DEVICES; SEMICONDUCTOR DIODE; VARACTOR.

Herbert J. Reich

Bibliography. T. F. Bogart, Jr., *Electronic Devices and Circuits*, 4th ed., 1996; R. L. Boylestad and L. Nashelsky, *Electronic Devices and Circuit Theory*, 7th ed., 1998; J. J. Brophy, *Basic Electronics for Scientists*, 5th ed., 1990; J. Millman and A. Grabel, *Microelectronics*, 2d ed., 1987; D. Schilling et al., *Electronic Circuits*, 3d ed., 1989; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Diopside

The monoclinic pyroxene mineral which in pure form has the formula $CaMgSi_2O_6$. Its space group is C2/c. Pure diopside melts congruently at 1391°C (2536°F) at atmospheric pressure; the effect of pressure up to 50 kilobars (5 gigapascals) on its melting temperature is given by Eq. (1), where P is in kilobars

$$P = 23.3\left[\left(\frac{T}{1665}\right)^{4.46} - 1\right] \qquad (1)$$

and T is in degrees Celsius. Diopside has no known polymorphs. Its structure consists of chains of SiO_4 tetrahedrons in which each silicon ion shares an oxygen with each of its two nearest silicon neighbors. These chains are linked together by calcium (Ca) and magnesium (Mg) ions in octahedral coordination.

Diopside forms gray to white, short, stubby, prismatic, often equidimensional, crystals with (110) cleavages intersecting at 87° (see **illus.**). Small amounts of iron impart a greenish color to the mineral. The indices of refraction increase with increasing iron; for pure diopside they are $n\alpha = 1.664$, $n\beta = 1.671$, and $n\gamma = 1.694$. Pure diopside is common and occurs as a metamorphic alteration of impure dolomites in medium and high grades of metamorphism, as shown by reaction (2).

$$CaMg(CO_3)_2 + 2SiO_2 \rightarrow CaMg(SiO_3)_2 + 2CO_2 \qquad (2)$$

Diopside shows extensive solid solution with a variety of other pyroxenes. The most important are $CaFeSi_2O_6$ (hedenbergite), $FeSiO_3$ (ferrosilite), $MgSiO_3$ (enstatite), $NaAlSi_2O_6$ (jadeite), $NaFeSi_2O_6$ (acmite), and $CaAl_2SiO_6$ (a synthetic pyroxene). Natural diopsidic pyroxenes also commonly contain significant amounts of chromium, titanium, and manganese, $CaMgSi_2O_6$ combined with variable amounts

Diopside crystals from St. Lawrence County, New York. (*Specimen from Department of Geology, Bryn Mawr College*)

of $CaFeSi_2O_6$, $MgSiO_3$, and $FeSiO_3$ form the common Ca-rich pyroxenes of basaltic and gabbroic rocks. A few natural Fe-poor diopsides containing up to 40% $MgSiO_3$ have been found in nodules in kimberlite. Natural diopsidic pyroxenes which show extensive solid solution with jadeite and to a lesser extent with acmite are called omphacite. Omphacite is a principal constituent of eclogites, rocks of basaltic composition which have formed at high pressure. *See* DOLOMITE; ECLOGITE; PYROXENE. F. R. Boyd

Diopter

A measure of the power of a lens or a prism. The diopter (also called dioptrie) is usually abbreviated D. Its dimension is a reciprocal length, and its unit is the reciprocal of 1 m (39.4 in.). Thus a thin lens of κ diopters has a focal length of $1000/\kappa$ mm or $39.4/\kappa$ in. The lens is collecting for positive κ, diverging for negative κ, and afocal for $\kappa = 0$. *See* FOCAL LENGTH; LENS (OPTICS).

One can speak of the power of a single surface. The power of a thin lens is then the sum of the powers of its surfaces in diopters. Analogously, the power of a group of (thin) lenses in contact is the sum of the powers of the single lenses.

For a lens which is not rotation-symmetric (having toric or cylindrical surfaces, for instance), two powers—one maximal and one minimal—must be assigned. These correspond to the powers in two perpendicular planes.

The dioptric power of a prism is defined as the measure of the deviation of a ray going through a prism measured at the distance of 1 m (39.4 in.). A prism that deviates a ray by 1 cm (0.394 in.) in a distance of 1 m (39.4 in.) is said to have a power of one prism diopter. *See* OPTICAL PRISM.

Spectacle lenses in general consist of thin lenses, which are either spherical, to correct the focus of the eye for near and far distances, or cylindrical or toric, to correct the astigmatism of the eye. An added prism corrects a deviation of the visual axis. The diopter thus gives a simple method for prescribing the necessary spectacles for the human eye. *See* EYEGLASSES AND CONTACT LENSES. Max Herzberger

Diorite

A phaneritic (visibly crystallized) plutonic rock having intermediate SiO_2 content (53–66%), composed mainly of plagioclase (oligoclase or andesine) and one or more ferromagnesian minerals (hornblende, biotite, or pyroxene), and having a granular texture. Diorite is the plutonic equivalent of andesite (a volcanic rock). This dark gray rock is used occasionally as a building stone and is known commercially as black granite. For a general discussion of textural, structural, and compositional differences *see* IGNEOUS ROCKS

Mineralogy. Gray or white plagioclase feldspar is the dominant mineral. The composition of the plagioclase averages in the oligoclase-andesine range (10–50% $NaAlSi_3O_8$). The plagioclase is frequently zoned, most often with a calcic core and sodic rims. Rocks with more calcic plagioclases and more abundant ferromagnesian minerals are gabbros. Subordinate light-colored minerals occurring in diorites are alkali feldspar (less than 5% of the rock). Intergrowths of alkali feldspar and quartz are common. Rocks with greater proportions of alkali feldspar are called monzonite. Those with more quartz are quartz-diorite or tonalite.

Ferromagnesian minerals in diorites consist usually of hornblende or biotite or both, but augite-, olivine-, and hypersthene-bearing diorites also occur. Biotite is more important to quartz diorites.

Accessory minerals in diorites usually include apatite, magnetite, and ilmenite, and sphene and zircon occur occasionally. Nepheline (or another feldspathoid) occurs rarely; where nephelene becomes abundant, the rock is named nepheline syenite.

Textures and structures. The texture of diorites is notably variable. Most often diorites are equigranular, with coarse, partly or mostly anhedral plagioclase and hornblende crystals, subordinate biotite, and interstitial quartz and orthoclase. Some diorites are porphyritic (diorite porphyry), with euhedral plagioclase and hornblende phenocrysts set in a granular groundmass. Some diorites contain conspicuous clots of light and dark minerals; more rarely there is a definite layering.

Occurrence. Diorite is found as isolated small bodies such as dikes, sills, and stocks, but it is also found in association with other plutonic rocks in batholithic bodies. It is closely associated with convergent plate boundaries where calc-alkalic magmatism and mountain building are taking place.

Origin. Many diorites are hybrid rocks that are formed by reaction of magma with wallrocks, or by mixing of two magmas. Most diorites crystallized from a magma which was dioritic or andesitic, perhaps differentiated from a basaltic melt. *See* MAGMA. William Ingersoll Rose, Jr.

Dioscoreales

A widespread order of lilioid monocotyledons composed of four or five families, the most important of which is Dioscoreaceae (650 species). These families contain some of the most peculiar plants among the monocots. Two families, Thismiaceae and Burmanniaceae (with 35 and 140 species, respectively), are largely composed of nonphotosynthetic herbaceous species that parasitize fungi, whereas other members, such as the bat plant (*Tacca*, Taccaceae) and the yams (*Dioscorea*) have net-veined leaves similar to dicots. Most of the 600 species of *Dioscorea* are vines, and all contain diosgenin, a precursor of progesterone and cortisone, leading to their being commercially collected in some parts of the world. *See* MAGNOLIOPHYTA. Mark W. Chase

Diphtheria

An acute infectious disease of humans caused by *Corynebacterium diphtheriae*. Classically, the disease is characterized by low-grade fever, sore throat, and a pseudomembrane covering the tonsils and pharynx. Complications such as inflammation of the heart, paralysis, and even death may occur due to exotoxins elaborated by toxigenic strains of the bacteria.

Etiologic agent. *Corynebacterium diphtheriae* is an obligate aerobic gram-positive bacillus that is shaped like a club. Whether a strain produces toxin or not determines the severity of the disease.

Toxigenicity is mediated by both genetic and nutritional factors. The gene coding for toxin production, *tox*$^+$, resides in several bacterial viruses (phages). When nontoxigenic *C. diphtheriae* are infected (lysogenized) by these phages, they are converted to toxin-producing organisms. In vitro, a low iron concentration in media is required for toxin production; the clinical significance of this nutritional requirement is unknown. *See* TOXIN.

Pathogenesis of diphtheria is mediated not only by the bacterial toxin but also by its cell-wall antigens and other enzymes. Nontoxigenic strains, therefore, can also produce localized disease, albeit a milder form.

Epidemiology. Humans are the only natural hosts of *C. diphtheriae.* Transmission is person to person, mostly by intimate respiratory and physical contact; the latter is more important in cutaneous diphtheria. Immune persons can become asymptomatic carriers of toxigenic *C. diphtheriae.* In the prevaccine era, respiratory diphtheria ranked among the leading causes of morbidity and mortality among children, presumably because of their high contact rates. Peaks in incidence occurred approximately every 10 years, with marked winter seasonality in temperate zones. The sexes were affected equally. In the tropics, cutaneous diphtheria predominates and is not seasonal. In unvaccinated populations, 75% or more of the population is likely to develop immunity to diphtheria by adulthood, presumably from mild or subclinical infections.

Where childhood vaccination against diphtheria is routine, the disease has practically disappeared. In North America, diphtheria (mostly cutaneous) is occasionally seen among Native Americans and indigent urban dwellers. The incidence has declined disproportionately to the fraction of the population with protective antitoxin titers. The absence of this disease may be due to an indirect effect of vaccination in reducing the circulation of toxigenic strains. Vaccine-induced immunity will wane in the absence of revaccination or exposure to toxigenic strains, leading to increased susceptibility among adults. When susceptibility increases and toxigenic strains are reintroduced, however, diphtheria outbreaks can recur. Adult susceptibility, reintroduction of toxigenic strains, and poor socioeconomic conditions led to a massive epidemic of diphtheria in the former Soviet Union, with more than 150,000 reported cases and over 5000 deaths from diphtheria from 1990 through 1998. Most reported patients in this outbreak have been 15 to 50 years old.

Signs and symptoms. Patients with pharyngotonsillar diphtheria experience sore throat, difficulty in swallowing, and low-grade fever. Laryngo-tracheal diphtheria results in hoarseness, croupy cough, difficulty in breathing, and pallor. Persons with severe cases may develop stupor, coma, and death from circulatory or respiratory failure within 10 days. Nervous system effects, principally motor, may appear 3–7 weeks after the onset of disease.

Diagnosis. Even if the classic diphtheritic membrane is present in the throat, clinical diagnosis is difficult, as a pseudomembrane also occurs in infectious mononucleosis, Group A streptococcal pharyngitis, and candidiasis. Specific diagnosis requires culture isolation of *C. diphtheriae* from the diseased site. Growth generally requires 8–12 h if no antibiotics are administered. Follow-up testing for toxigenicity is critical if *C. diphtheriae* is isolated.

Pathogenesis. The upper respiratory tract is the most common portal of entry for *C. diphtheriae*. It can also invade the skin and, more rarely, the genitalia, eye, or middle ear. The disease has an insidious onset after a usual incubation period of 2–5 days. Clinical severity is determined by the individual's preexisting immunity against diphtheria toxin, the virulence and toxigenicity of the invading *C. diphtheriae*, other coexisting disease, and the anatomic location of the infection. Anterior nasal and cutaneous diphtheria tends to be mild. Tonsillar diphtheria may extend either upward to involve the soft palate and other parts of the pharynx, or downward into the larynx or bronchi.

Toxin. Diphtheria toxin is a 60-kilodalton polypeptide that inhibits protein synthesis by blocking adenosine diphosphate (ADP) ribosylation and causes apoptotic cell death by a separate but related mechanism. The toxin is released by the multiplying *C. diphtheriae* and is absorbed by mucosal epithelial cells; their death elicits an inflammatory response. When the fluid that is released from the tissues clots, it enmeshes fibrin, dead epithelial cells, leukocytes, and bacterial organisms to form a classic adherent diphtheric pseudomembrane 1–3 mm thick. The membrane has a leatherlike appearance, and bleeding occurs when it is removed. It may be bluish, white, gray, or black, depending on the degree of hemorrhage. The membrane becomes a fertile site for further bacterial growth and toxin production. Its extension creates a larger surface for toxin absorption and portends a grave prognosis. Absorbed toxin is highly toxic to cardiac and nerve cells, leading to the common complications of myocarditis and polyneuritis.

Complications. Persons with severe pharyngeal diphtheria develop marked edema of the anterior neck resulting in a "bullneck" appearance. In laryngeal or bronchial diphtheria, the membrane provides

a large surface for toxin absorption and can suddenly occlude the airway. Absorbed toxin can damage any human tissue, but affects the heart and nervous system most severely. The resultant fatty degeneration of cardiac muscle can lead to arrhythmias, dilatation, and heart failure. Nerve damage, usually delayed and transient, can result in paralysis of the soft palate, posterior pharyngeal wall, cranial nerves, and respiratory muscles. Fatalities for diphtheria before the advent of antitoxin therapy occurred in 35% of cases overall, and were as high as 90% for laryngeal diphtheria; even with antitoxin therapy, the fatality rate is still approximately 5–10%.

Treatment. The only specific therapy is diphtheria antitoxin, administered in doses proportional to the severity of the disease. Antitoxin is produced by hyperimmunizing horses with diphtheria toxoid and toxin. It is effective only if administered prior to the binding of circulating toxin to target cells. Therefore, early use of antitoxin is advised in suspected cases with an inadequate or unclear history of immunization. Antibiotics do not alter the course, the incidence of complications, or the outcome of diphtheria, but are used to eliminate the organism from the patient. This helps to prevent further toxin production and transmission of the organism. Erythromycin or penicillin is preferred. Patients should be isolated until two successive cultures, taken at 24-h intervals, are negative. The patients should be carefully observed for cardiac and respiratory complications.

Prevention. Persons with protective antitoxin titers may become infected with diphtheria but do not develop severe disease. Since the 1920s, active immunization with diphtheria toxoid has proved safe and effective in preventing diphtheria in many countries. Diphtheria toxoid is produced by incubating the toxin with formalin. It is usually administered together with tetanus toxoid and pertussis vaccine as DTP. Active immunization requires a primary series of four doses, usually at 2, 4, 6, and 18 months of age, followed by a booster at school entry. Older persons should receive adult tetanus and diphtheria toxoids to complete their primary series, and boosters every 10 years. *See* IMMUNITY.

Control efforts following diagnosis of an initial case should be directed toward the early diagnosis, isolation, and treatment of cases; full immunization of children and adults; and antibiotic treatment of close contacts. *See* MEDICAL BACTERIOLOGY; VACCINATION. Robert Chen; Charles Vitek

Diplogasterida

An order of nematodes in which the labia are seldom well developed; however, a hexaradiate symmetry is distinct. The external circle of labial sensilla may appear setose, but they are always short, never long or hairlike. The stoma may be slender and elongate, or spacious, or any gradation between. The stoma may be armed or unarmed; the armature may be movable teeth, fossores, or a pseudostylet. The corpus is always muscled and distinct from the postcorpus, which is divisible into an isthmus and glandular posterior bulb. The metacorpus is almost always valved. The female reproductive system may have one or two ovaries, and males may or may not have caudal alae; however, a gubernaculum is always present. The male tail commonly has nine pairs of caudal papillae; three are preanal and six are caudal.

There are two superfamilies in this order, Diplogasteroidea and Cylindrocorporoidea. The members of Diplogasteroidea are predators, bacterial feeders, and omnivores; they are often found in association with insects or the fecal matter of herbivores. The Cylindrocorporoidea includes both free-living forms and intestinal parasites of amphibians, reptiles, and certain mammals. *See* NEMATA. Armand R. Maggenti

Diplomonadida

An order of the class Zoomastigophorea, phylum Protozoa. These are small, colorless flagellates, some of which are free-living and some parasitic. The body of this protozoan is bilaterally symmetrical, composed of two mirror halves, each with a nucleus and a full set of organelles. There are four flagella to a side, not all of the same length. In division of the cell the two nuclei each form a spindle, so that each daughter cell receives two nuclear complexes.

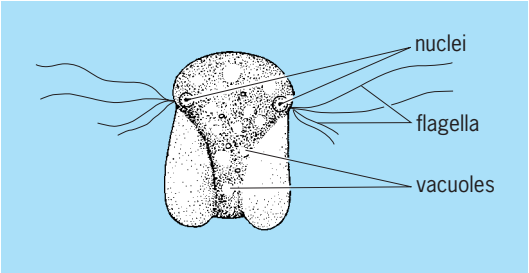

A diplomonad, *Trepomonas rotans*.

Trepomonas rotans is probably the most common of the free-living species (see **illus.**), and usually occurs in water of high organic, low (or no) oxygen content, such as sewage. It also occurs in ocean water. This protozoan may have two or four long lateral flagella and six or four short ones, according to the species. *Trepomonas* swims slowly with jerky movements, revolving on its axis. This genus also is found as a parasite or symbiont in other animals, such as amphibians. *Hexamita* is another genus containing both free-living and parasitic representatives. *Spironucleus*, named for its spiral nuclear arrangement, is parasitic. *Giardia* has several species which parasitize vertebrates. In humans giardiasis results in severe dysentery; however, adequate medication is available. *See* GIARDIASIS; PROTOZOA; SARCOMASTIGOPHORA; ZOOMASTIGOPHOREA. James B. Lackey

Diplopoda

A class of terrestrial, tracheate, oviparous arthropods. Diplopods (millipeds) are largely cryptic in habits, are saprophytic feeders, and are characterized by the development of a compact head with a pair of short, simple, eight-jointed antennae, powerful mandibles, and a subbuccal gnathochilarium formed from embryonic maxillary elements. The body is not differentiated into thorax and abdomen, but is composed of a variable number of similar, cylindrical diplosomites, each of which (except the first two or three) bears two pairs of walking legs (see **illus.**). The body wall is chitinous, with a thick impregnation of calcium carbonate in the majority of species. Most millipeds are variously adapted for rolling into a closed spiral or nearly perfect sphere when threatened. Typically, each abdominal segment is provided with a pair of glands which effuse a volatile poisonous fluid. Respiration occurs through profuse fine tracheae opening through stigmata near the bases of the legs. The digestive tract is a straight tube, except in oniscomorph species in which it is somewhat coiled. The anal opening is in the last segment and is closed by two tightly fitting anal valves. Ducts from reproductive organs pass through or behind the second pair of legs to external openings.

Reproduction. The sexes are separate and fertilization is internal, following prolonged clasping behavior. Spermatic masses are extruded beforehand from the male seminal opening onto the gonopods (modified legs of the seventh segment) from which sperm material is transferred into seminal receptacles in the cyphopods (specialized structures terminating the outer ends of the oviducts). In oniscomorph millipeds the gonopods are not developed, and the male transfers spermatophores with his mouthparts. Eggs vary greatly in size and number. They may be laid in a cluster and "brooded" by the mother, scattered singly in the humus environment, or enclosed in an igloo-shaped mud nest built by the mother. Postembryonic development is anamorphic, and the individual passes through seven growth stages, with segments and legs added at each molting period. Development is gradual, without major changes in appearance, and may require a year or more for completion. Mating usually takes place soon after the molt into sexual maturity.

Classification and distribution. More than 8000 species have been described, although so far only the fauna of Europe is well known. Existing classifications are still unsatisfactory but, in general, about 11 orders and more than 111 families are recognized. Probably as many as 25,000 species will eventually be recognized. Classification is based to a large extent upon shape of the male gonopods, which are quite constant and characteristic for each species. For higher groups the forms of segments, mouthparts, and legs are useful.

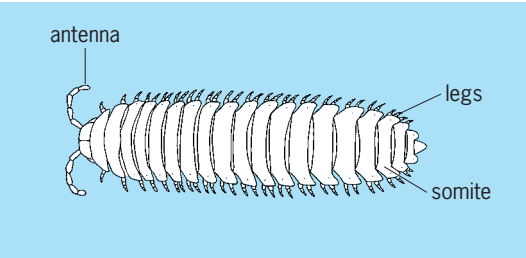

A diplopod. (*After R. E. Snodgrass, A Textbook of Arthropod Anatomy, Cornell University Press, 1952*)

Most species are local in distribution, because millipeds generally remain close to the parental habitat, and some may be restricted to a few square miles. Even genera are limited in distribution, and only a few occur on more than one continent. This endemism, plus dependence on a continuously moist and undisturbed habitat, makes diplopods well suited for studies of zoogeography and evolution. *See* ZOOGEOGRAPHY.

The diplopods have a long geological history; they arose in the Early Devonian and were well developed by the Late Pennsylvanian. There has been little adaptative radiation in the Diplopoda, and less external modification of body form than in perhaps any other large class of arthropods. *See* ARTHROPODA.

Richard L. Hoffman

Diplura

An order of ancestrally wingless Hexapoda comprising about 800 species and variously considered as apterygotan insects or as their independent sister group. These insects are equipped with three pairs of thoracic legs and a pair of antennae on the head. They are slender and white, but unlike typical insects they have intrinsic muscles in the antennal segments beyond the basal one. All diplurans lack eyes, and have paired mandibles and maxillae largely enclosed by fused folds of the head so that only the tips are exposed for retrieving food. The integument is thin and unpigmented, and the posterior end of the abdomen is furnished with a pair of cerci, which may be multisegmented (for example, in the family Campodeidae) or single segmented and modified into a pair of stout, pigmented jaws or forceps (in the family Japygidae and allied families) resembling those of adult earwigs and used for catching prey or for defense. The dipluran body size is usually small, under 0.4 in. (1 cm) long, but some Australian *Heterojapyx* species reach 2 in. (5 cm).

Their habitat is normally cryptic in soil, under stones, and in rotten wood, where Campodeidae feed on vegetable matter, while other families apparently are predaceous. Campodeidae are found worldwide and are common in temperate and tropical regions, while the other eight families are mainly restricted to warmer climates and are often discontinuous or local in distribution. *See* INSECTA.

William L. Brown, Jr.

Bibliography. D. J. Borrer, C. A. Triplehorn, and N. F. Johnson, *An Introduction to the Study of Insects*, 6th ed., 1989; Commonwealth Scientific and

Industrial Research Organization Staff, *The Insects of Australia*, 2d ed., 1991.

Dipnoi

The infraclass, also called Dipneusti, consisting of the lungfishes. The Dipnoi plus the fossil infraclass Porolepimorpha make up the subclass Dipnotetrapodomorpha, one of five subclasses of the class Sarcopterygii. The extant lungfishes are in two orders as follows:

Order Ceratodontiformes
  Family Ceratodontidae (Australian lungfishes)
Order Lepidosireniformes
  Family Lepidosirenidae (South American lungfishes)
  Protopteridae (African lungfishes)

In comparison with Devonian lungfishes, extant species have a reduced number of median fins, different tail shape, and decreased ossification of the braincase and other endochondral bones. For example, in Devonian lungfishes the caudal fin is heterocercal (that is, the dorsal lobe is usually longer) or diphycercal (symmetrical), whereas it is always diphycercal in extant species; the dorsal and anal fins are free from the caudal fin, whereas in extant species the dorsal and anal fins are confluent with the caudal fin; there are zero to three branchiostegal rays, whereas in extant species branchiostegal rays are always absent; and a gular plate is present, whereas in extant species it is absent. Also, cell size has apparently increased since the Devonian, and the living species have very large genomes. These changes suggest that pedomorphosis played a role in dipnoan evolution. *See* HETEROCHRONY.

Recent forms. Lepidosireniforms have eel-like bodies, filamentous paired fins, two lungs, tooth plates with three cutting blades, a greatly reduced skull roof, and an elongated ceratohyal important in suction feeding and respiration (**Fig. 1**). Lepidosirenids are obligate air breathers, and the heart, arterial tree, and gills are highly specialized. To breathe air, the fish protrudes its snout from the water surface, gulps air into the mouth, and then forces it into the lung by closing the mouth and raising the floor of the oral cavity. Male lepidosirenids construct and defend breeding nests containing thousands of young. *Lepidosiren paradoxa* occurs in the Amazon and Parana river systems in South America. Young South American lungfish have yellow spots, and are imported to the United States for the tropical fish trade.

African lungfishes of the genus *Protopterus* are widely distributed in major lake and river systems of sub-Saharan Africa. Four species are known: *P. annectens*, *P. dolloi*, *P. aethiopicus*, and *P. amphibius*. Some reach nearly 6 ft (2 m) in length and are important in local fishing economies. African lungfish survive periods of drought by burrowing in mud and encasing themselves in cocoons of hardened mucus.

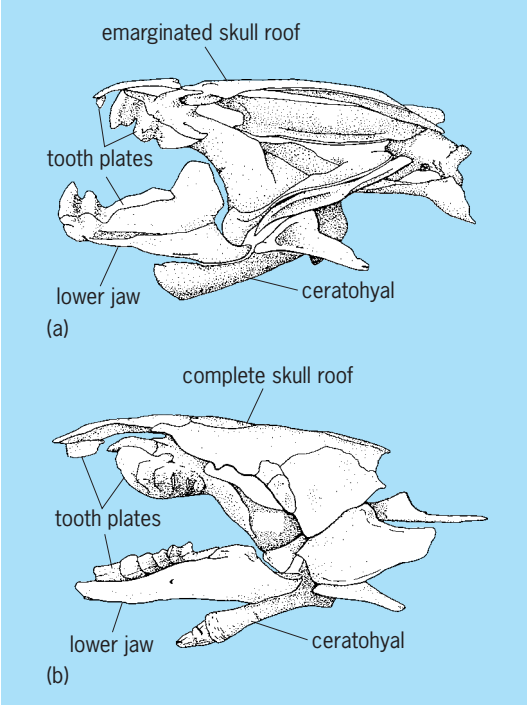

Fig. 1. Lateral views of the skulls of (*a*) African lungfish, *Protopterus aethiopicus*, and (*b*) Australian lungfish, *Neoceratodus forsteri*.

The family Ceratodontidae, represented by the single species *Neoceratodus forsteri* (**Fig. 2**), has a stout body and large scales, paddlelike paired fins, and tooth plates with flattened crushing surfaces. These fishes resemble typical Paleozoic lungfishes more closely than do the lepidosirenids. Adults reach more than 3 ft (1 m) in length. The single lung is poorly subdivided, and the gills are essential for gas exchange. During development, tooth plates are formed by the fusion of separate denticles. The diet includes plants, crustaceans, and soft-bodied invertebrates. Fossils referred to *N. forsteri* occur in the interior of Australia, but today the species is restricted to shallow, coastal rivers in Queensland, including the Burnett, Mary, and Brisbane rivers. *Neoceratodus forsteri* is completely protected by Australian law.

Some authorities ally lepidosirenids with Late Paleozoic genera such as *Gnathorhiza*, while others believe that the split between lepidosirenids and ceratodontids occurred more recently.

Phylogeny. Lungfishes show many characters seemingly intermediate to fishes and amphibians, and this has resulted in much controversy about the phylogenetic relationships of the group since the discovery of lepidosirenids in the 1830s. Most workers continue to regard rhipidistian fishes as more closely related to tetrapods than are lungfishes or coelacanths. William Bemis; Herbert Boschung

Fossils. Most Devonian dipnoans had two dorsal fins, a heterocercal tail, and cosmine (tissue composed of dentine, a pore-canal system, and a thin enamel layer) on scales and bones. Later forms lost the cosmine; their scales have a network of lines which separate each scale into plates. The caudal fin

Fig. 2. Australian lungfish (*Neoceratodus forsteri*). (*Photo © www.jjphoto.dk*)

became diphycercal and continuous with the dorsal and anal fins. The paired fins have a long, segmented axis with diverging radials, an archipterygium.

There are three types of dentition: rasping, deciduous denticles in the Early Devonian *Uranolophus* and the Late Devonian rhynchodipterids; smooth, crushing dental plates in a few Devonian forms; and crushing tooth plates with radiating ridges, which are the best-preserved fossil remains of the group, characteristic of most Devonian and all post-Devonian forms except *Conchopoma*.

Early dipnoans were marine; freshwater adaptations occur in Paleozoic and most post-Paleozoic dipnoans. Burrowing habits developed independently in Late Paleozoic gnathorhizids and extant lepidosirenids. Dipnoans were common worldwide from the Early Devonian to the end of the Triassic, but are rare since. They are restricted to the southern continents in the Cenozoic. *See* OSTEICHTHYES.

Hans-Peter Schultze

Bibliography. W. E. Bemis, W. W. Burggren, and N. E. Kemp (eds.), *The Biology and Evolution of Lungfishes*, 1987; R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1987; E. Jarvik, *Basic Structure and Evolution of Vertebrates*, vol. 1, 1980; A. Kemp, The embryological development of the Queensland lungfish, *Neoceratodus forsteri* (Krefft), *Mem. Queensland Mus.*, 20:553–597, 1982; D. E. Rosen, Lungfishes, tetrapods, paleontology and plesiomorphy, *Bull. Amer. Mus. Nat. Hist.*, 167:159–276, 1981.

Dipole

Any object or system that is equally and oppositely charged at two points or poles, such as a magnet, a polar molecule, or an antenna element. The properties of a dipole are determined by its dipole moment, that is, the product of one of the charges and their separation. *See* DIPOLE MOMENT.

An electric dipole consists of two electric charges of equal magnitude but opposite polarity, separated by a short distance (see **illus.**); or more generally, a localized distribution of positive and negative electricity without net charge whose mean positions of positive and negative charge do not coincide.

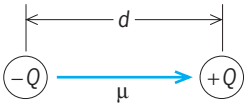

Electric dipole with moment $\mu = Qd$.

Molecular dipoles which exist in the absence of an applied field are called permanent dipoles, while those produced by the action of a field are called induced dipoles. *See* POLAR MOLECULE.

The term magnetic dipole originally referred to the fact that a magnet has two poles and, because of these two poles, experiences a torque in a magnetic field if its axis is not along a magnetic flux line of the field. It is now generalized to include electric circuits which, because of the current, also experience torques in magnetic fields. *See* MAGNET.

The term dipole is also used to describe a commonly used element of radio-frequency and microwave antennas. An electric dipole whose moment oscillates sinusoidally radiates electromagnetic waves and is known as a hertzian dipole; it is of interest in developing the theory of electromagnetic radiation. For practical purposes, a half-wave dipole, consisting of two collinear conducting rods, fed at the center, whose combined length equals half the wavelength of the radiation to be transmitted or received, is often used as an antenna element, either by itself or in an array, or as a feed for a reflector. Other shapes of dipoles (using, for example, thick cylinders or double cones) may be used to achieve a greater bandwidth. *See* ANTENNA (ELECTROMAGNETISM); ELECTROMAGNETIC RADIATION; ELECTROMAGNETIC WAVE TRANSMISSION.

A. Earle Bailey

Dipole-dipole interaction

The interaction of two atoms, molecules, or nuclei by means of their electric or magnetic dipole moments. This is the first term of the multipole-multipole series of invariants. More precisely, the interaction occurs when one dipole is placed in the field of another dipole. The interaction energy depends on the strength and relative orientation of the two dipoles, as well as on the distance between the centers and the orientation of the radius vector connecting the centers with respect to the dipole vectors. The electric dipole-dipole interaction and magnetic dipole-dipole interaction must be distinguished.

Electric dipole-dipole interaction. The center of the negative charge distribution of a molecule may fail to coincide with its center of gravity, thus creating a dipole moment. An example is the water molecule. If such molecules are close together, there will be a

dipole-dipole interaction between them. Atoms do not have permanent dipole moments, but a dipole moment may be induced by the presence of another atom nearby; this is called the induced dipole-dipole interaction. Induced dipole-dipole forces between atoms and molecules are known by many different names: van der Waals forces, originally so named because they represent the attractive term in the van der Waals equation of state; London forces or dispersion forces, because F. London explained this interaction by using the picture of optical dispersion, a picture that was in turn based on the behavior of damped harmonic oscillators. The fact that the phases of the electrons in two different molecules are correlated in the dispersion formula led to the term dynamically induced dipole-dipole interaction. The proper way to treat the interaction is by means of a quantum-mechanical derivation using perturbation theory; the lowering of the energy resulting from the presence of one molecule in the neighborhood of the other is found to have the same features as the induced dipole-dipole interaction. These induced dipole-dipole forces are also responsible for cohesion and surface tension in liquids. They also act between unlike molecules, resulting in the adsorption of atoms on macroscopic objects. Moreover, using sophisticated experimental techniques, it can be shown that the same force exists between two macroscopic objects at very short distances. *See* ADSORPTION; CHEMICAL BONDING; COHESION (PHYSICS); DIELECTRIC MATERIALS; INTERMOLECULAR FORCES; SOLUTION; SURFACE TENSION; VAN DER WAALS EQUATION.

Magnetic dipole-dipole interaction. The magnetic dipole-dipole interaction is found both on a macroscopic and on a microscopic scale. The classical magnetic dipole is produced by a current loop. Its orientation is determined by the orientation of the surface element enclosed by the current loop. Another example is the permanent magnetic dipole of a compass needle; two of these within reasonable proximity of each other illustrate clearly the influence of the dipole-dipole interaction. In quantum mechanics, the magnetic moment is partially due to a current arising from the motion of the electrons in their orbits, and partially due to the intrinsic moment of the spin. The same interaction exists between nuclear spins. *See* ELECTRON SPIN; MAGNET; NUCLEAR MOMENTS.

Magnetic dipole-dipole forces are particularly important in low-temperature solid-state physics, the interaction between the spins of the ions in paramagnetic salts being a crucial element in the use of such salts as thermometers and as cooling substances. Usually the interaction is a mixture of exchange interaction and magnetic dipole-dipole interaction. In many magnetic rare-earth salts, the interaction is purely of the latter type. Since the dipole-dipole interaction is well understood, it is possible to make a calculation from first principles of the heat capacity and the susceptibility (deviation from the Curie law) of a salt. These determine its thermometric behavior. *See* ADIABATIC DEMAGNETIZATION; EXCHANGE INTERACTION; LOW-TEMPERATURE THERMOMETRY; MAGNETIC THERMOMETER.

The so-called high-temperature expansion of the partition function, used in the calculations of the interaction, leads to the need for lattice sums. The lattice sums for dipolar interactions are conditionally convergent; that is, the result of the summation depends on the arrangement of the terms. In terms of physical properties, this means that the magnetization depends on the shape of the sample. This is described by a shape-dependent demagnetization factor. *See* STATISTICAL MECHANICS.

In magnetic cooling and low-temperature thermometry, the dipole-dipole interaction provides the coupling between the spins, needed to make the system go to equilibrium. Its influence shows up at higher temperatures as a small deviation from the Curie law. *See* CURIE-WEISS LAW; DIPOLE; DIPOLE MOMENT.

Paul H. E. Meijer

Bibliography. D. S. Betts, *Introductory Statistical Mechanics*, 1992; C. J. F. Bottcher, *Theory of Electric Polarization*, 2d ed., vol. 1, 1972, and vol. 2, 1978; J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular Theory of Gases and Liquids*, 1964; R. P. Hudson, *Principles and Application of Magnetic Cooling*, 1972.

Dipole moment

A mathematical quantity characteristic of a dipole unit equal to the product of one of its charges times the vector distance separating the charges. The dipole moment $\boldsymbol{\mu}$ associated with a distribution of electric charges q_i is given by Eq. (1), where $\mathbf{r}_i$ is

$$\mu = \sum_i q_i \mathbf{r}_i \tag{1}$$

the vector to the charge q_i. For systems with a net charge (for example, positive), the origin is taken at the mean position of the positive charges (and vice versa). Dipole moments have the dimensions coulomb-meters. Molecular dipole moments were previously expressed in debye units, where 1 debye $= 3.336 \times 10^{-30}$ C · m. *See* DIPOLE.

The electric potential Φ of a dipole at a long distance $\mathbf{R}$ from the dipole is given by Eq. (2), where θ

$$\Phi = \gamma \frac{|\mu| \cos\theta}{4\pi\epsilon_0 |\mathbf{R}|^2} \tag{2}$$

is the angle between $\boldsymbol{\mu}$ and $\mathbf{R}$, ϵ_0 is the permittivity of vacuum, and γ is a geometrical factor. ($\epsilon_0 = 8.854 \times 10^{-12}$ farad/m and $\gamma = 1$ in SI units.)

The potential energy U of a dipole in a uniform electric field $\mathbf{E}$ is $U = -|\boldsymbol{\mu}|\,|\mathbf{E}| \cos\phi$, where ϕ is the angle between $\boldsymbol{\mu}$ and $\mathbf{E}$.

The induced dipole moment $\boldsymbol{\mu}_i$ of a molecule may be expressed in terms of molecular parameters by the equation $\boldsymbol{\mu}_i = \alpha \mathbf{E}_L$, where α is the polarizability and $\mathbf{E}_L$ is the local field strength acting at the molecular site. This relation permits the interpretation of the macroscopic polarization and hence

dipole moments in terms of molecular processes. *See* PERMITTIVITY; POLARIZATION OF DIELECTRICS.

Robert D. Waldron

Dipsacales

An order of flowering plants (angiosperms) in the asterid I group of the eudicotyledons consisting of around 6 families and almost 1000 species. Like many asterids of the asterid II group, Dipsacales have an inferior ovary and opposite leaves. They also have fewer stamens than petals (or petal lobes), and the flowers are often bilaterally symmetric or of irregular symmetry. Family limits have changed recently as a result of deoxyribonucleic acid (DNA) sequence studies, but the species content of the order is still similar to previous systems of classification. The flowers are typically arranged in heads, similar to those of their close relatives in Asterales and Apiales, but the ovary often contains more than one seed (reduced to a single seed in Dipsacaceae and Valerianaceae). Familiar members of Dipsacales include elderberry (*Sambucus*, Adoxaceae), honeysuckle (*Lonicera*, Caprifoliaceae), and teasel (*Dipsacus*, Dipsacaceae). *Viburnum* (Adoxaceae) and *Abelia* (Linnaeaceae) are commonly planted ornamental shrubs. *See* ASTERALES; MAGNOLIOPHYTA; PLANT KINGDOM.

Mark Chase

Diptera

An order of the class Insecta known as the true flies and so named because they possess only two wings. This characteristic, together with a pair of balancers or halteres, distinguishes flies from all other orders of the Insecta. The three names for members of the order, known commonly as flies, gnats, or midges, form a part of the common names of most families, genera, and species of the order. The term fly also forms a part of the compound names of the insects in many other orders, such as butterfly, Mayfly, and chalcid fly, but when used alone it is correctly applied only to the members of the Diptera. The forms most commonly known as maggots are actually dipterous larvae, and keds (Hippoboscidae) are parasitic forms of flies that have lost their wings. The Diptera are the most important group of insects considered in medical entomology. Many, especially the mosquitoes, are the vectors for numerous parasites and diseases.

In number of species, it is the third largest order; only the Coleoptera and Lepidoptera are larger. There are approximately 100,000 kinds of flies in the world, of which over 18,500 inhabit America north of Mexico.

The order is ubiquitous and is more widespread than any other of the insects. Diptera are found on every continent and most islands of the world, except for the coldest parts of the Arctic and Antarctic, and at the tops of extremely high mountain ranges. They occur in almost all available ecological niches. Specimens have been taken at altitudes of many thousands of feet by specially constructed traps attached to airplanes.

Taxonomy

In the early 1970s the classification of the Diptera underwent a general reappraisal. For convenience, the families of Diptera may be divided into the Nematocera and the Brachycera, and the latter group further divided into the Orthorrhapha and the Cyclorrhapha. Adults of the Nematocera are characterized by their slow flight and long antennae, the flagellum of which is divided into 10–65 segments. These are the more primitive Diptera, and the major groups are represented in fossil deposits as old as Jurassic and Cretaceous age.

The Brachycera are generally swift fliers, and their antennae are no longer than the head. The flagellum is usually reduced to a single segment with a thread-like process called an arista. In some families of Orthorrhapha, there is no arista and the flagellum may be superficially annulate. In the Orthorrhapha, the adult fly emerges from the pupa through a T-shaped anterior opening, while in the Cyclorrhapha it emerges from a circular opening at the anterior end of the puparium.

Following is a list of scientific and vernacular names of the more common families of flies.

- Nematocera
  - Trichoceridae (winter crane flies)
  - Ptychopteridae (phantom crane flies)
  - Tipulidae (crane flies)
  - Psychodidae (moth flies and sand flies)
  - Blepharoceridae (net-winged midges)
  - Chironomidae (midges)
  - Ceratopogonidae (biting midges)
  - Simuliidae (black flies)
  - Chaoboridae (phantom midges)
  - Culicidae (mosquitoes)
  - Dixidae (dixid midges)
  - Anisopodidae (wood gnats)
  - Bibionidae (march flies)
  - Scatopidae (scatopsid flies)
  - Mycetophilidae (fungus gnats)
  - Sciaridae (black fungus gnats)
  - Cecidomyiidae (gall midges)
- Brachycera—Orthorrhapha
  - Stratiomyidae (soldier flies)
  - Rhagionidae (snipe flies)
  - Tabanidae (horse flies and deer flies)
  - Bombyliidae (bee flies)
  - Asilidae (robber flies)
  - Empididae (dance flies)
  - Dolichopodidae (long-legged flies)
- Brachycera—Cyclorrhapha
  - Lonchopteridae (spear-winged flies)
  - Phoridae (hump-backed flies)
  - Platypezidae (flat-footed flies)
  - Syrphidae (syrphid flies)
  - Pipunculidae (big-headed flies)
  - Lonchaeidae (lonchaeid flies)
  - Lauxaniidae (lauxaniid flies)
  - Drosophilidae (pomace or fruit flies)

Brachycera—Cyclorrhapha (*cont.*)
- Ephydridae (shore flies)
- Anthomyiidae (anthomyiid flies)
- Muscidae (house flies)
- Calliphoridae (blow flies)
- Sarcophagidae (flesh flies)
- Oestridae (bot flies)
- Tachinidae (tachinid flies)
- Hippoboscidae (louse flies)
- Micropezidae (stilt-legged flies)
- Diopsidae (stalk-eyed flies)
- Sciomyzidae (marsh or snail-killing flies)
- Sepsidae (black scavenger flies)
- Heleomyzidae (heleomyzid flies)
- Sphaeoceridae (small dung flies)
- Agromyzidae (leaf-miner flies)
- Tethinidae (tethinid flies)
- Chloropidae (frit flies)
- Conopidae (thick-headed flies)
- Tephritidae (fruit flies)
- Piophilidae (skipper flies)

Morphology

Adult flies vary from somewhat less than 1 mm to over an inch in body length. The head is vertical, usually with three ocelli at the dorsal vertex and the mouthparts ventral. In some cases, however, the head is distinctly longer than high when viewed laterally. The compound eyes are usually large and prominent, and are either holoptic (touching at the top of the head) or dichoptic (separated at the vertex by part of the head capsule) (**Fig. 1**).

Mouthparts. The mouthparts are modified for either lapping or piercing, but never with the mandibles apposable and capable of chewing. The structure of the mouthparts varies greatly throughout the order so that their homologies are often rather difficult to determine. In the mosquitoes and some other Nematocera, the labrum, mandibles, hypopharynx, and maxillae are greatly elongated and form a piercing tube enclosed by the labium or labellum. Throughout the Orthorrhapha there tends to be a retention of mandibles for piercing, especially in those flies that require blood meals. In the Cyclorrhapha, on the other hand, mandibles tend to be reduced or lost entirely, and with few exceptions the work of the mouthparts is largely sucking, with the maxillae and labium taking over the large share of work. In these flies the labellum is often enlarged and flattened and provided with minute rasping spines and a specialized mechanism for taking up fluids. Since all adults flies imbibe fluids only, the presence of a pharyngeal "pump" to transfer fluids from the external substrate to the gut is characteristic.

Antennae. Each antenna typically has two basal segments and one or more additional segments called the flagellum. In the Nematocera the flagellum consists of a variable number of quite similar segments. In all remaining flies there are three distinct antennal segments. The third segment, representing the flagellum, is usually longer than the two basal ones and sometimes complex. It is either annulated or produced in various shapes, and often bears an arista which is bare or haired.

In the Cyclorrhapha, a small crescent-shaped sclerite, the lunule, is situated just above the antennae. The flies in which this sclerite is separated from the rest of the frontal portion by a suture have an internal distensible sac, or ptilinum, which can be pushed through this suture before the fly is fully developed. It is used to break open the end of the puparium to allow the fly to escape. Once hardened and fully adult, the fly loses all external trace of this sac.

Thorax. The prothorax and metathorax are small and closely united with the large mesothorax which contains the musculature for the single pair of wings. The legs are usually alike, except in species in which the prothoracic pair are raptorial and species in which the fore, mid, or hind legs of the males are modified by secondary sexual characters. The coxae are usually rather long, and there are from two to five tarsomeres. Claws are present in almost all species. Beneath each claw in many species there is a membranous pad, or pulvillus. In addition, some flies carry an empodium between each pair of pulvillae.

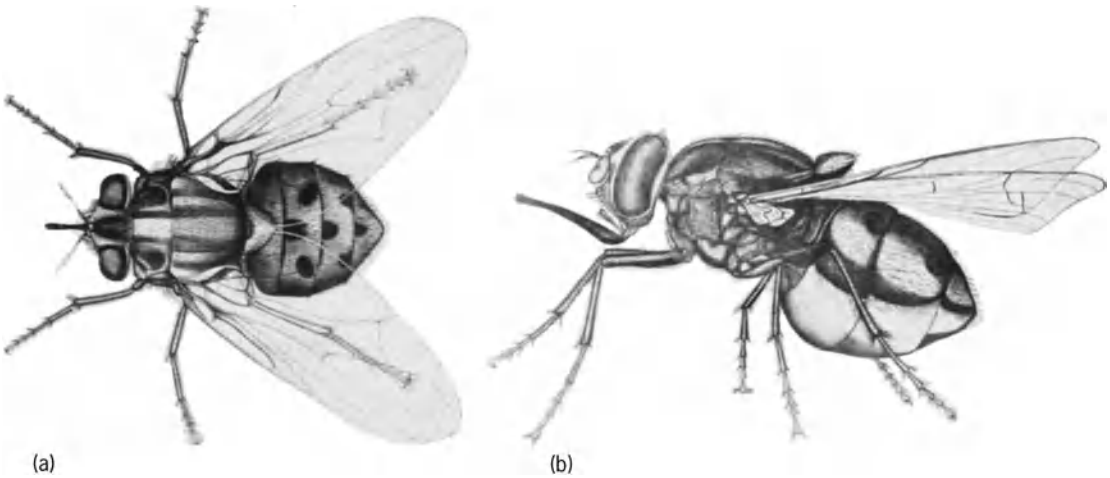

Fig. 1. Stable fly, *Stomoxys calcitrans*. (*a*) Adult female. (*b*) Female engorged with blood. (*USDA*)

This may be bristlelike or may take the form of membranous pad like the pulvillae in shape and size.

Wings. Only the first pair of wings, the mesothoracic pair, is developed. The second pair is reduced to a pair of club-shaped organs, or halteres, which serve as balancing organs in flight. They are present in most species, even when the mesothoracic wings are absent. In many species, especially those of the Nematocera, abundant scalelike setae clothe the veins and sometimes the margins of the wings. The wings of some of the higher Diptera bear yellow, brown, or black markings that aid in species identification. The venation of the primitive dipteran wing follows quite closely that of the primitive insect wing, except that vein radius (R) tends to be reduced, and only the important cross veins are present. In the higher Diptera the branches of some of the veins coalesce with others, thus further reducing the number of apparent veins present. In some Cyclorrhapha a pair of membranous lobes, the squamae or calypters, are found at the extreme base of the wing. Each of these occurs in pairs, the upper half fastened to the wing and the lower half to the thorax.

Setae. In certain families of Diptera, some of the setae covering the head and body are greatly enlarged and are used for identification. The location and relative size of these bristles are often characteristic of species, genera, or familes.

Eggs, larvae, and pupae. The eggs are usually elongated and taper at both ends. They are relatively smooth-surfaced, but the integument is often characteristically sculptured in designs that are visible only under high magnification (**Fig. 2**). The eggs of some aquatic species are provided with air sacs, of characteristic design, which aid in keeping them afloat. When laid, most eggs are whitish but darken gradually as the embryo grows within.

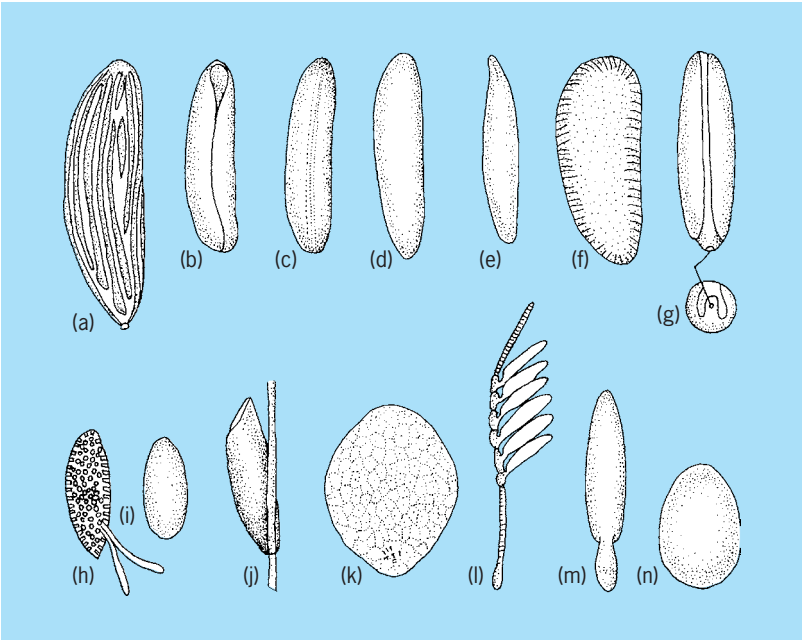

Fig. 2. Eggs of Diptera. (*a*) *Hippelates pusio*. (*b*) *Stomoxys calcitrans*. (*c*) *Cochliomyia macellaria*. (*d*) *Musca domestica*. (*e*) *Tabanus punctifer*. (*f*) *Cryptochaetum iceryae*. (*g*) *Cochliomyia hominovorax*. (*h*) *Drosophila melanogaster*. (*i*) *Tipula infuscata*. (*j*) *Gasterophilus nasalis*. (*k*) *Zenillia labatrix*. (*l*) *Lypoderma lineata* attached to a hair and (*m*) dorsal aspect of single egg. (*n*) *Simulium simile*. (*After E. O. Essig, College Entomology, Macmillan, 1942*)

Larvae, unlike adults, have elongate bodies. Mouthparts of this feeding stage are usually sclerotized structures. They are variously and, many times, intricately designed for rasping or chewing, but sometimes they are reduced or apparently entirely absent. The head of the larva of most lower Diptera is large, sclerotized, and complicated in internal structure, while in many species of higher Diptera it is reduced in size. The mouthparts are always well developed to enable the feeding stage to reach maturity readily. Although dipterous larvae never have true segmented appendages, anterior or posterior prolegs may be present. Nearly all types of spiracular arrangement are found in this order, but many larvae have prominent prothoracic and posterior spiracles with characteristic openings.

Pupae have appendages which are rather adherent to the body. In some pupae the developing adult is free, but in others it is encased in a hardened last larval skin resembling a seed or capsule and presenting no particular external features except spiracles and segmentation.

Life Cycle and Biology

The adult stage in the life history may be regarded as a reproductive and dispersal phase, during which these insects increase both in number and often in geographic range. Females select suitable oviposition sites and lay a variable number of eggs, after which they usually die. The eggs, after a period of incubation of varying length, give rise to larvae, which represent a growth phase. During this period, the insect does almost nothing but nourish itself, thus providing the tissues necessary for a later transformation into an adult. This growth period is divided into stages, or instars, each of which terminates with the molting of the larval skin to allow increase in size for the stage that follows. In the Diptera there are four larval instars. In most of the Orthorrhapha, especially those with aquatic immature forms, all four instars are active; but in the Cyclorrhapha, at the third larval molt, a hard puparium is formed inside which a quiescent fourth larval stage occurs. At the conclusion of the final larval instar, there is a dramatic reorganization in tissue structure initiated by the action of appropriate hormones on the so-called "imaginal buds," a process that eventually results in the formation of an adult insect. During this time, the insect is called a pupa because its external form is now altered and any external activity virtually ceases.

When the adult tissues are almost fully formed, the fly emerges from its pupal case and spends some time in drying, hardening, and expanding its wings, attaining color, and reaching sexual maturity. It is then ready to start its cycle once again (**Figs. 3** and **4**).

Adults. The adults, most of which are able to fly, represent the reproductive and dispersal phase of a fly's developmental cycle. After extricating itself from the puparium, the adult requires some period

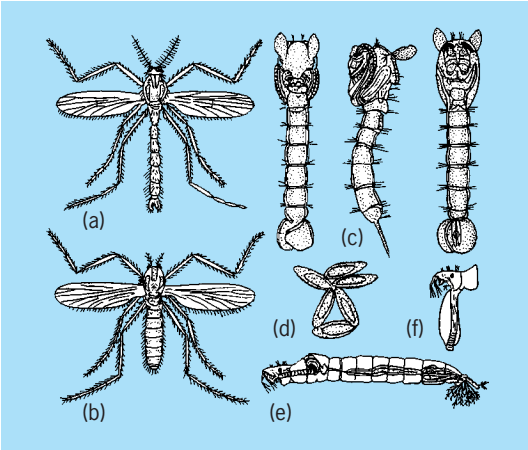

Fig. 3. Clear Lake gnat, *Chaoborus astictopus*. (*a*) Male; (*b*) female; (*c*) pupae; (*d*) eggs; (*e*) larva (phantom larva); (*f*) eversible pharyngeal sac extended from mouth. (*After E. O. Essig, College Entomology, Macmillan, 1942*)

of exposure to the air, after which it flies about to find food and to procreate.

Nutrition. Food requirements of females apparently differ from those of males because of their need for nutrient materials for developing and maturing eggs. In many of the nematocerous Diptera there is quite a different arrangement of mouthparts in the female, and nowhere is this more true than in the mosquitoes, of which most females require blood meals before they are able to develop eggs. Many female mosquitoes have decided preferences for certain hosts. Some few species, generally regarded as the most important disease vectors, are attracted principally to humans; others prefer a wide variety of mammals, and still others, birds.

The families of flies whose females require blood meals are scattered throughout the order but are concentrated mainly in two areas of the classification listing given earlier. The families Psychodidae, Ceratopogonidae, Culicidae, and Simuliidae have this habit. Females of the family Tabanidae also seek humans and other animals for blood. Among the calypterates, some Muscidae require blood, and many other species scattered among various families are attracted by exudates from the bodies of animals and humans.

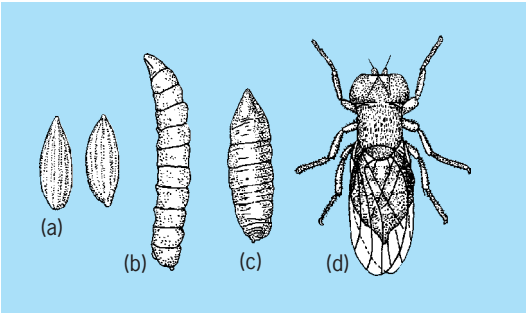

Fig. 4. Coachella eye gnat. *Hippelates pusio*. (*a*) Eggs; (*b*) larva; (*c*) pupa; and (*d*) adult. (*After E. O. Essig, College Entomology, Macmillan, 1942*)

The adults of a fairly large number of species are predators and feed on smaller anthropods. This habit is scattered throughout the order to such an extent that a listing of all families would be too lengthy. Whether such food materials are required for development of their eggs is not known with certainty.

Those flies whose larvae live and develop in plant materials produce adults with a great variety of food habits. It is known that adult Tephritidae of many species are especially attracted to exudates from scale insects, which are so numerous on many fruit trees that their honeydew covers large areas of foliage. It has been shown that certain components of this material are essential for the development of fruit fly eggs.

In the large majority of flies the food habits of adults have never been studied, and nothing is known except as the result of an occasional naturalist's observation.

Reproduction. Most female flies require a period of time between their emergence from the puparium and the time they are ready to lay eggs. The length of this period is sometimes characteristic of a species. Thus certain fruit flies are able to lay eggs 3 or 4 days after emergence while others require as long as 30 days. Females of many species will not accept males for copulation until a certain amount of time has expired; some species require fertilization after every egg-laying period, and some need fertilization only once during their lifetimes. Males usually emerge before females, presumably to assure that they will be sexually mature in time to fertilize the first eggs ready for oviposition.

Fertilization takes place in many ways throughout the order. Males of many of the nematocerous flies, and representatives of other families as well, form mating swarms. These are usually organized over an elevation in the surrounding terrain such as a mound of earth, bush, tree stump, or fence post. The swarm remains stationary over the object while the flies mill about in the swarm in every possible direction. Although such swarms are composed principally of males, some females are also present; the purpose of such swarms is to accomplish fertilization of those females.

The females of other species, especially many of the Tephritidae and perhaps of many other families, are fertilized at about the time they lay their eggs. Males of some of these species lie in wait on the surface of a fruit or some other substrate where they apparently know a female will alight to lay eggs. The male attacks just before, during, or immediately after the act of oviposition. Males have been seen to pull the ovipositor of a female from the pulp of a fruit to deposit the sperm.

Eggs. Study of fly eggs has been neglected by biologists, and most of their observations have been confined to the various ways in which females deposit them. Almost every conceivable kind of behavior accompanies egg laying. Some flies broadcast their small eggs while on the wing, spreading them about on the surrounding foliage. Eggs of some of these flies are eaten by caterpillars where they hatch

into larvae which feed within their hosts. Females of species whose larvae are parasitic cement their eggs on the surface of their host, or deposit them inside the bodies of the larvae, pupae, or adults that serve as their victims. Some females are able to deposit their eggs while their hosts are in full flight. Perhaps the most interesting habit is that of the female of *Dermatobia hominis*, which captures a passing mosquito upon which to fasten her eggs. As soon as this carrier has come to rest on an animal to feed, the *D. hominis* eggs quickly hatch and the larvae enter the mosquito's host.

Many flies whose larvae attack plant material have a sharply pointed, inflexible ovipositor to insert their eggs within the plant tissues, thus affording newly hatched larvae easy access to the succulent parts of the plant. Others without such an ovipositor fasten their eggs on the plant surface, and the newly hatched larvae immediately burrow into the plant tissue or seek a suitable place to gain entrance. The eggs of Sarcophagidae and some other flies hatch within the abdomen of the female and grow to various stages of development before escaping to continue a life of their own. In several genera of Cecidomyiidae, or gall midges, the reproductive organs develop prematurely with certain larvae and produce eggs which hatch internally so that only the new, or daughter, larvae escape from the body of the mother larva.

Larvae. It has been estimated that the immature stages of about one-half the known species of flies occur in association with water. Dipterous larvae are found in almost every conceivable habitat within this aquatic environment. Some occupy extreme situations such as intertidal wave-swept rocks on the seacoasts (Chironomidae), thermal inland waters as hot as 120°F or 49°C (Stratiomyidae), saturated brine pools (Ephydridae), and even natural petroleum seeping from the ground (Ephydridae). Adaptations such as these are unusual and occur only in a few species of the families mentioned. They are usually accounted for by special and unusual morphological adaptations.

Most aquatic larvae live in all kinds of standing water or in running water, from the smallest rivulet to the large slowly flowing rivers. The species that are found as aquatics in the ordinary sense, that is, swimming under and at the surface, are often provided with posterior spiracles especially adapted for obtaining atmospheric oxygen. These may be placed at the apex of a more or less elongated "air tube," such as that in many mosquito larvae, or on other specially adapted organs. They are almost always accompanied by a special arrangement of hairs and by a water-repellent coating so that water will not be "inhaled" at the surface. The larvae of some species may never come to the surface, but obtain oxygen by absorption through membranes in gills or through body surface membranes.

The remaining half of the order have larvae that inhabit nonaquatic habitats. Of these, a very large number of species are found in association with living plants. A large share of the Cecidomyiidae cause galls to form in plant tissue or live in galls produced by other species. Occasionally galls may be found underground in roots and stems, and often in aerial roots of epiphytes. More often they attack growing stems, leaves, fruits, and seeds, but no part of a plant can be mentioned that is not subject to a dipterous attack of one kind or another. Gall midges and certain tephritids characteristically make a large number of different kinds of galls. They are found as inquilines in galls made by other species and may even be found in masses of pitch made to exude from evergreens by their presence.

The fruit flies, Tephritidae, are so called because of their habit of feeding on ripe or nearly ripe fruit, or because they develop in the ovaries and seeds of those plants that do not bear succulent fruit. This occurs especially in tropical and subtropical areas of the world. Larvae of other species feed directly upon plant tissue, while certain Muscidae attack only roots. The Agromyzidae and certain other small acalypterates feed between layers of plant tissue in stems and leaves, causing mines that are often characteristic of the species making them.

Another large group of larvae, including those of another part of the Cecidomyiidae, are found in association with dead and decaying vegetation lying on the ground surface, and a number of tephritids and near relatives are strict saprophytes.

Many fly larvae are predaceous and are sometimes provided with specially adapted mouthparts for preying upon other species. True parasitism occurs here and there throughout the order. Larvae of some species of Tachinidae, Asilidaae, Bombyliidae, and others are parasitic in the bodies of a large number of insects. Sciomyzidae larvae feed principally on the soft bodies of snails; Pyrgotidae, on June beetles; Cryptochaetidae, on scale insects; and so forth.

Some of the more advanced flies of the calypterate group require the blood or other fluids of mammals for larval development. Most of these affect domestic livestock. Thus the primary screwworm causes wounds in the undamaged skin of large domestic animals, and the botflies (**Fig. 5**) have a complicated development during which their larvae travel into many parts of the bodies of their hosts.

The Congo floor maggot requires human or other animal blood and commonly lives in association with

Fig. 5. Botfly larvae, of a species of *Gasterophilus*, attached to inner lining of stomach of horse. (*Southwestern Biological Supply Co., Dallas, Texas*)

humans, feeding through the skin, but is not permanently attached. Larvae of the Hippoboscidae are carried full grown in the ovaries of the mother, while the mother, in turn, is a permanent resident of sheep and feeds on sheep's blood.

Pupae. As stated in the beginning section of this article, pupae may be of two kinds. In most, if not all, of the Orthorrhapha, all four larval instars are active feeding stages. They are either aquatic or phytophagous, and the skin of the last larval stage is cast free and lost at the formation of the pupa. Some pupae, especially of mosquitoes, are free-swimming just like the larvae, while others are quiescent forms. In the Cyclorrhapha, the skin of the third larval stage is retained and transformed into a barrel-shaped puparium inside which the fourth larval stage is quiescent and in which the pupa reaches full development. The various stages of this process have been described in detail for the apple maggot (*Rhagoletis pomonella*). Most cyclorrhaphous pupae are adapted for life underground or in other well protected parts of their environment, and larvae commonly burrow into the soil or these other locations to pupate.

No matter what the pupal environment, however, provision is always made for the escape of the adult. In many of the aquatic Diptera, pupation takes place at the air-water interface; mosquito pupae, for instance, simply float at the water surface until the adult is free. Adults of the Simuliidae rise to the water surface from pupae at the bottom surrounded by a bubble of air. Pupae of many species are provided with spines or hairs to help them move about.

Economic Importance

Diptera have probably more economic importance than any other insect order. This importance comes from the relatively few species that affect humans, domestic animals, and plants. The vast majority of forms have no direct importance although a number are beneficial in one way or another. Many flies visit flowers and are efficient plant pollinators. Adults may prey upon harmful arthropods or lay eggs in their bodies so that the developing larvae can use them as a source of food. Aquatic larvae and pupae under certain circumstances serve as abundant food for fish and other aquatic life, while the larvae of a number of terrestrial species act as scavengers, assisting bacteria to destroy all kinds of decaying animal and vegetable matter.

Human. Only a relatively few dipteran species cause severe economic loss to humans. Perhaps of most concern is the role played by them in disease transmission. About 70 species of *Anopheles* mosquitoes transmit an estimated 500,000 cases of malaria each year. This disease has serious debilitating effects and occurs on every continent and many islands of the world. Yellow fever, once a threat to completion of work on the Panama Canal, is transmitted principally by a single mosquito, *Aedes aegypti*. This disease is largely under control, yet the same virus is found in monkeys living in tropical rainforests of Central and South America and Africa, and is transmitted by *Aedes* and *Haemagogus* species to people who work out of doors in the jungle. Dengue, or breakbone fever, is a usually nonfatal disease of worldwide distribution that leaves its victims debilitated for several weeks. It is transmitted by *Aedes aegypti* and *A. albopictus*. Filariasis, primarily a disease of peoples of Africa, Asia, and the Pacific islands, is caused by a minute roundworm whose larvae are transmitted by a few species of *Anopheles*, *Mansonia*, *Culex*, and *Aedes*. Filiariasis affected a large number of American troops during World War II. Longstanding cases in natives of areas where the disease is endemic may result in an enlargement of the extremities called elephantiasis. With modern methods used by bacteriologists and virologists, a large number of virus diseases known to be transmitted by mosquitoes have been found in humans. Many of these, such as Sindbis virus of Egypt, may not produce clinical symptoms of disease, but others, such as the equine encephalitides, which include western, eastern, St. Louis, Japanese, and others, may be quickly fatal in humans. With the exception of the viruses, none of the diseases mentioned above are transmitted by any insects other than members of the order Diptera. *See* MALARIA; YELLOW FEVER.

A few less well-known diseases are transmitted by other flies. Black flies transmit *Onchocerca volvulus*, the causative agent of onchocerciasis, a disease affecting the eyes of natives of Central and parts of South America. Sandflies of the genus *Phlebotomus* transmit organisms that cause kala azar, Oriental sore, pappataci fever, and Oroya fever. Even today large parts of Africa remain underdeveloped because of the presence of sleeping sickness, a fatal disease transmitted by tsetse flies. Species of the genus *Chrysops* (Tabanidae) are instrumental in carrying tularemia, primarily a disease of rodents. *See* TULAREMIA.

Domestic animals. Domestic animals, hides, meat, and dairy products are affected by disease transmission or by direct attack by flies. Anthrax, tularemia, botulism, many virus diseases, and nagana, a form of sleeping sickness, are some of the diseases transmitted by members of the Diptera that take an annual toll amounting to millions of dollars.

Some flies wreak their damage by direct attack. The primary screwworm fly deposits eggs on hides of animals, and the larvae, upon hatching, burrow through the skin and into the flesh. The secondary screwworm gains entrance through holes, often infected, already present on the skin surface. *Hypoderma lineata* and *H. bovis* are botflies of special importance. Eggs of both species are laid on hairs of cattle, and the hatching larvae bore through the skin into the connective tissue. During their development they wander through the tissues of the animal, and when mature they escape through holes which they make along the spine of their host and drop to the ground to pupate. Horses are afflicted by a species of *Gasterophilus*. They lick the eggs from their bodies, and the larvae settle and dwell in their stomachs (Fig. 5), often in large numbers. Sheep are victims of the sheep botfly, whose larvae live in their nasal passages and sinuses. All of these insects

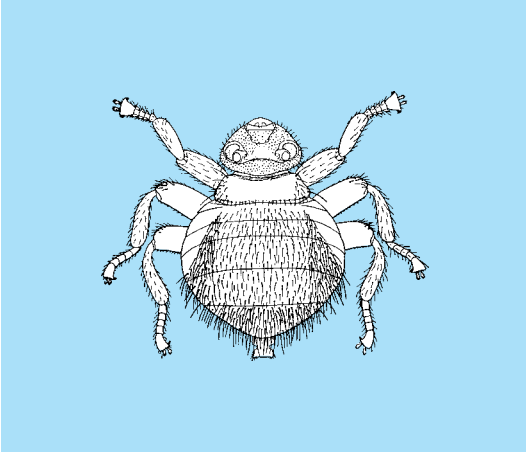

Fig. 6. Bee louse, *Braula coeca*. (*After E. O. Essig, College Entomology, Macmillan, 1942*)

affect the meat, milk, and wool production of animals through debilitation and irritation, and hides may be rendered completely useless by their entrance and exit holes. An interesting dipteran is the bee louse, which is a wingless, ectoparasite of the honeybee (**Fig. 6**).

Many species of flies are attracted by unsanitary conditions about garbage cans and rendering plants, by undressed wounds, and by seepages from infected eyes and mucous membranes of living animals. By coming into contact with these products, flies can mechanically carry on their feet and body hairs the organisms that cause typhoid, dysentery, diarrhea, cholera, yaws, and trachoma, in addition to certain parasitic worms. *See* MYIASIS.

Plants. Economic losses as a result of damage by flies to crops are perhaps not as great as those caused by other insects, yet their presence has necessitated the expenditure of large sums of money for control. Among these, perhaps the most important are members of the family Tephritidae, the larvae of which feed upon and ruin the succulent flesh and seeds of their hosts. The most important of these are the European and American cherry fruit flies; their relatives in the genus *Rhagoletis* that attack walnuts in North America; six or seven species of the genus *Anastrepha* which attack many kinds of fruit in the New World; the Mediterranean fruit fly which is a limiting factor in the production of many fruits, especially citrus, in many parts of Africa, Central and South America, and the Mediterranean Basin; and several species of the genus *Dacus*, which attack olives, citrus fruits, many kinds of vegetables, and other edible plants and plant parts in Europe, Asia, and Pacific islands.

Larvae of some flies mine the leaves of ornamentals thereby defacing them, reducing their growth potentials, and affecting their production and sale by nurseries. The Hessian fly, a gall midge, is a serious pest of wheat and other grains grown for human consumption in Europe, northern Asia and North America, and causes untold losses to growers in those areas. Several other species of Itonididae affect the growth of rice and other plants grown for human consumption.

The larvae of some species of Muscidae are known as root maggots and burrow into the underground parts of plants with considerable loss to growers of truck crops over the world. *See* ARTHROPODA; ENTOMOLOGY, ECONOMIC; INSECTA; PLANT PATHOLOGY.

Raymond J. Gagne

Bibliography. D. J. Borror et al., *An Introduction to the Study of Insects*, 6th ed., 1997; D. J. Borror and R. E. White, *A Field Guide to the Insects of America North of Mexico*, 1970; G. C. D. Griffiths, *The Phylogenetic Classification of Diptera Cyclorrhapha*, 1972; H. Oldroyd, *The Natural History of Flies*, 1964; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; H. H. Ross et al., *A Textbook of Entomology*, 4th ed., 1982, reprint, 1991.

Direct-coupled amplifier

A device for amplifying signals with direct-current components. There are many different situations where it is necessary to amplify signals having a frequency spectrum which extends to zero. Some typical examples are amplifiers in electronic differential analyzers (analog computers), certain types of feedback control systems, medical instruments such as the electrocardiograph, and instrumentation amplifiers. Amplifiers which have capacitor coupling between stages are not usable in these cases, because the gain at zero frequency is zero. Therefore, a special form of amplifier, called a dc (direct-current) or direct-coupled amplifier, is necessary. These amplifiers will also amplify alternating-current (ac) signals. *See* AMPLIFIER; ANALOG COMPUTER; BIOMEDICAL ENGINEERING; CONTROL SYSTEMS; INSTRUMENTATION AMPLIFIER.

Some type of coupling circuit must be used between successive amplifier stages to prevent the relatively large supply voltage of one stage from appearing at the input of the following stage. These circuits must pass dc signals with the least possible amount of attenuation.

Interstage direct-coupling in transistor dc amplifiers must be implemented with special care. The use of both *npn* and *pnp* transistors is a possible solution. However, the *pnp* transistors available in monolithic form have relatively poor current-gain and frequency-response characteristics. If a dc amplifier is formed by a cascade of *npn* stages, there is a positive dc level buildup toward the positive supply voltage. This voltage buildup limits the linearity and amplitude of the available output swing. The problem can be overcome by using a level-shift stage between each stage to shift the output dc level toward the negative supply with minimum attenuation of the amplified signal. Practical dc level-shift stages suitable for monolithic circuit applications can use Zener diodes, (**Fig. 1***a*), a series of diodes (Fig. 1*b*), or a V_{BE}, multiplier circuit (Fig. 1*c*). For example, the Zener diode level-shift circuit (Fig. 1*a*) provides

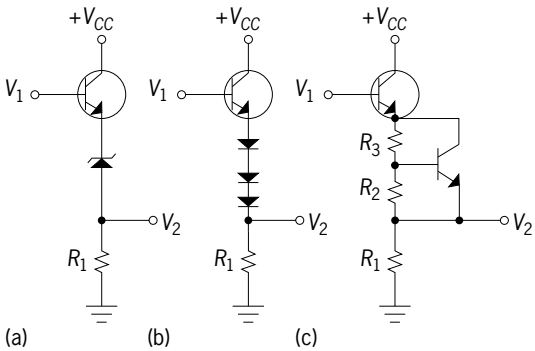

Fig. 1. Level-shift stages using (*a*) Zener diodes, (*b*) series of diodes, and (*c*) V_{BE} multiplier circuit. V_1 = input voltage, V_2 = output voltage.

a means of shifting the dc level between the input and the output by an amount equal to the sum of the base-to-emitter voltage and the Zener voltage. If the bulk resistance of the Zener diode is negligible compared to the resistance between the output terminal and ground (R_1), the voltage gain for the stage will be approximately unity. In the V_{BE} multiplier circuit (Fig. 1*c*), the output (V_2) is related to the input (V_1) by the equation below, where V_{BE} is the base-

$$V_1 = V_2 + V_{BE}\left(1 + \frac{R_3}{R_2}\right)$$

to-emitter voltage and R_2 and R_3 are two resistances in the circuit. *See* VOLTAGE-MULTIPLIER CIRCUIT.

Differential dc amplifier. It is generally recognized that the differential amplifier (**Fig. 2**) is the most stable dc amplifier circuit available. This is true because in this circuit the performance depends on the difference of the device parameters, and transistors can be manufactured using the planar epitaxial technique with very close matching of their parameters.

Carrier amplifier. A method of amplifying dc (or slowly varying) signals by means of ac amplifiers is to modulate a carrier signal by the signal to be amplified, amplifying the modulated signal, and demodulating at the output. (In some applications, such as instrument servomechanisms, output in the form of a modulated carrier is required and no demodulation is necessary.)

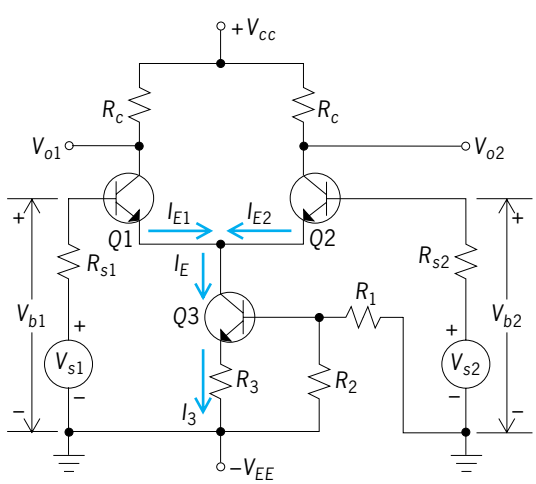

Fig. 2. Differential dc amplifier with constant-current stage in the emitter circuit.

An analysis of an actual circuit would show that in order to have an output free from harmonics introduced by the modulation, the output low-pass filter cutoff frequency must be small compared to the modulation frequency. This limits the bandwidth of the input signal to a small fraction of the bandwidth of signals which can be amplified by the types previously discussed. This is a disadvantage in most cases, but there are cases, such as the output voltage from a thermocouple, where the required bandwidth is small. Various types of choppers (bipolar transistors, field-effect transistors) are used as modulators. An additional disadvantage is that the chopper must be carefully designed to reduce to a minimum the noise and the offset voltage and leakage current of a transistor. However, this amplifier provides stable amplification of the input signal, and it is used in many industrial recording instruments. *See* MODULATION; VOLTAGE AMPLIFIER.

Chopper-stabilized amplifiers. The offset voltage of matched transistor pairs of differential amplifiers can be a source of serious problems in precision analog dc amplifier applications. Typically the offset voltage of matched metal oxide semiconductor (MOS) transistor pairs can be reduced to within ±20 mV by careful processing. However, even this low offset voltage in many applications is unacceptable. It is possible to reduce the effective input offset voltage to below ±1 mV by using chopper-stabilized amplifiers employing offset-nulling or auto-zero techniques. These techniques are essentially sampled-data methods and are based on the concept of measuring periodically the offset voltage and subsequently storing it as a voltage across a holding capacitor and then subtracting it from the signal plus the offset. Christos C. Halkias

Bibliography. D. L. Feucht, *Handbook of Analog Circuit Design,* 1990; A. B. Grebene, *Bipolar and MOS Analog Integrated Circuits Design*, 1984.

Direct current

Electric current which flows in one direction only through a circuit or equipment. The associated direct voltages, in contrast to alternating voltages, are of unchanging polarity. Direct current corresponds to a drift or displacement of electric charge in one unvarying direction around the closed loop or loops of an electric circuit. Direct currents and voltages may be of constant magnitude or may vary with time.

Batteries and rotating generators produce direct voltages of nominally constant magnitude (**illus.** *a*). Direct voltages of time-varying magnitude are produced by rectifiers, which convert alternating voltage to pulsating direct voltage (illus. *b–d*). *See* BATTERY; GENERATOR; RECTIFIER.

Direct current is used extensively to power adjustable-speed motor drives in industry and in transportation. Very large amounts of power are used in electrochemical processes for the refining and plating of metals and for the production

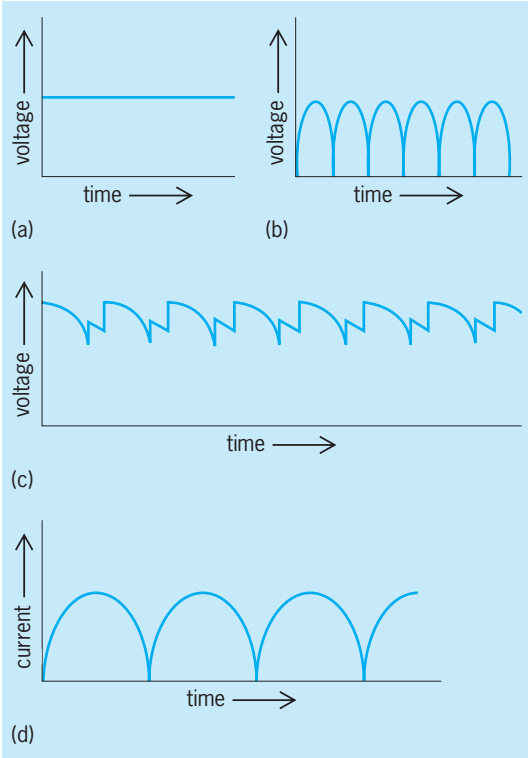

Typical direct currents and voltages. (*a*) Battery voltage; (*b*) full-wave rectified voltage; (*c*) dc voltage of the rectifier station of a high-voltage transmission link (*after E. W. Kimbark, Direct Current Transmission, vol. 1, Wiley-Interscience, 1971*). (*d*) Direct current in a resistance supplied by a full-wave rectifier.

of numerous basic chemicals. *See* DIRECT-CURRENT MOTOR; ELECTROCHEMICAL PROCESS; ELECTROPLATING OF METALS.

Direct current ordinarily is not widely distributed for general use by electric utility customers. Instead, direct-current (dc) power is obtained at the site where it is needed by the rectification of commercially available alternating current (ac) power to dc power. Solid-state rectifiers ordinarily are employed to supply dc equipment from ac supply lines. Rectifier dc supplies range from tiny devices in small electronic equipment to high-voltage dc transmission links of several thousand megawatts capacity. *See* ELECTRIC POWER SYSTEMS; SEMICONDUCTOR RECTIFIER.

Many high-voltage dc transmission systems have been constructed throughout the world. Very large amounts of power, generated as ac and ultimately used as ac, are transmitted as dc power. Rectifiers supply the sending end of the dc link; inverters then supply the receiving-end ac power system from the link. For long transmission distances, the ac-dc-ac system often is more economical than ac transmission, especially in the case of underground or underwater cable circuits. *See* ALTERNATING CURRENT; ELECTRIC CURRENT; TRANSMISSION LINES. D. D. Robb

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., McGraw-Hill, 2000; E. W. Kimbark, *Direct Current Transmission*, vol. 1, Wiley-Interscience, 1971.

Direct-current circuit theory

An analysis of relationships within a dc circuit. Any combination of direct-current (dc) voltage or current sources, such as generators and batteries, in conjunction with transmission lines, resistors, inductors, capacitors, and power converters such as motors is termed a dc circuit. Historically the dc circuit was the first to be studied and analyzed mathematically. *See* CIRCUIT (ELECTRICITY).

Classification. Circuits may be identified and classified into simple series and parallel circuits. More complicated circuits may be developed as combinations of these basic circuits.

Series circuit. A series circuit is illustrated in **Fig. 1**. It consists of a battery of E volts and three resistors of resistances R_1, R_2, and R_3, respectively. The conventional current flows from the positive battery terminal through the external circuit and back to the negative battery terminal. It passes through each resistor in turn; therefore the resistors are in series with the battery.

Parallel circuit. The parallel circuit, shown in **Fig. 2**, consists of a battery paralleled by three resistors. In this case the current leaving the positive terminal of the battery splits into three components, one component flowing through each resistor, then recombining into the original current and returning to the negative terminal of the battery.

Physical laws of circuit analysis. The operation of the basic series and parallel circuits must obey certain fundamental laws of physics. These laws are referred to as Ohm's law and Kirchhoff's laws in honor of their originators.

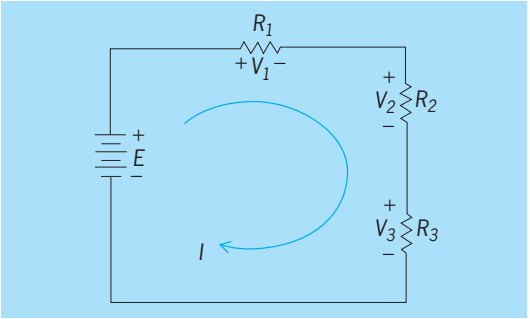

Fig. 1. Simple series circuit.

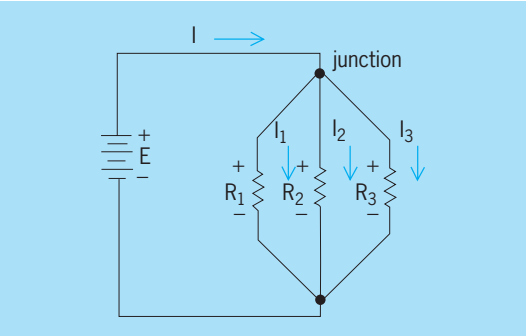

Fig. 2. Simple parallel circuit.

Voltage drops. When an electric current flows through a resistor, a voltage drop appears across the resistor, the polarity being such that the voltage is positive at the end where the conventional current enters the resistor. This voltage drop is directly proportional to the product of the current in amperes and the resistance in ohms. This is Ohm's law, expressed mathematically in Eq. (1). Thus in Fig. 1 the

$$V = IR \tag{1}$$

drop across R_1 is V_1 and has the polarity shown. *See* OHM'S LAW.

Summation of voltages. The algebraic sum of all voltage sources (rises) and voltage drops must add up to zero around any closed path in any circuit. This is Kirchhoff's first law. In Fig. 1, the sum of the voltages about this closed circuit is as given by Eqs. (2),

$$E - V_1 - V_2 - V_3 = 0 \tag{2a}$$

$$E = V_1 + V_2 + V_3 \tag{2b}$$

where the minus signs indicate a voltage drop. Written in terms of current and resistance, this becomes Eq. (3). From this results the important conclusion

$$E = I(R_1 + R_2 + R_3) = IR_{\text{total}} \tag{3}$$

that resistors in series may be added to obtain the equivalent total resistance, as shown in Eq. (4).

$$R_{\text{total}} = R_1 + R_2 + R_3 + \cdots \tag{4}$$

See KIRCHHOFF'S LAWS OF ELECTRIC CIRCUITS.

Summation of currents. The algebraic sum of all currents flowing into a circuit junction must equal the algebraic sum of all currents flowing out of the junction. In the circuit shown in Fig. 2, the current flowing into the junction is I amperes while that flowing out is the sum of I_1 plus I_2 plus I_3. Therefore the currents are related by Eq. (5). This is Kirchhoff's

$$I = I_1 + I_2 + I_3 \tag{5}$$

second law. In this case the same voltage appears across each resistor. If the currents are expressed in terms of this voltage and values of the individual resistors by means of Ohm's law, Eq. (5) becomes Eq. (6). The equivalent resistance R_{eq} that can re-

$$I = \frac{E}{R_1} + \frac{E}{R_2} + \frac{E}{R_3} = \frac{E}{R_{\text{eq}}} \tag{6}$$

place the resistances in parallel can be obtained by solving Eq. (6) in the form of Eq. (7). Therefore, re-

$$R_{\text{eq}} = \left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots\right)^{-1} \tag{7}$$

sistances in parallel are added by computing the corresponding conductances (reciprocal of resistance) and adding to obtain the equivalent conductance. The reciprocal of the equivalent conductance is the equivalent resistance of the parallel combination. *See* CONDUCTANCE.

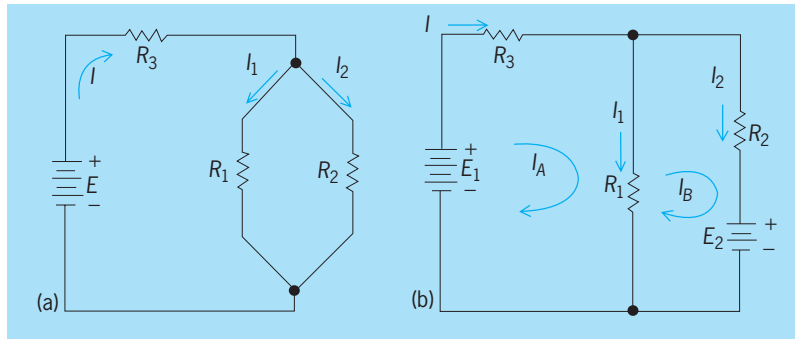

Fig. 3. Series-parallel circuits. (*a*) Single source. (*b*) Multiple sources.

Sources. Sources such as batteries and generators may be connected in series or parallel. Series connections serve to increase the voltage; the net voltage is the algebraic sum of the individual source voltages. *See* BATTERY; GENERATOR.

Sources in parallel provide the practical function of increasing the net current rating over the rating of the individual sources; the net current rating is the sum of the individual current ratings.

Series-parallel circuits. More complicated circuits are nothing more than combinations of simple series and parallel circuits as illustrated in **Fig. 3**.

Single source. Circuits that contain only a single source are readily reduced to a simple series circuit. In the circuit of Fig. 3*a* the parallel combination of R_1 and R_2 is computed and used to replace the parallel combination. The resultant circuit is now a simple series circuit consisting of R_3 and R_{eq} readily be solved for the series current if the voltage is known.

Multiple sources. Circuits that contain two or more sources located in various branches cannot be reduced to a simple series circuit (Fig. 3*b*). The three basic laws of circuit theory still hold and may be directly applied to provide a simultaneous solution to the loop currents I_A and I_B flowing in each basic series circuit present in the overall network.

In this example the summations of voltages around the individual series circuits or loops are given by Eqs. (8) and (9), which may be solved for the mathe-

$$E_1 = (R_1 + R_3)I_A - R_1 I_B \tag{8}$$

$$-E_2 = -R_1 I_A + (R_1 + R_2)I_B \tag{9}$$

matical loop currents I_A and I_B. These loop currents can in turn be identified by reference to the circuit where I_A is identical to I, and therefore I_B, as stated by Eq. (10).

$$I_A - I_B = I_1 \tag{10}$$

This method may be used to solve any complicated combination of simple circuits. Other methods are also available to the circuit analyst. *See* NETWORK THEORY.

Power. The electric power converted to heat in any resistance is equal to the product of the voltage drop across the resistance times the current through the resistance, as stated by Eq. (11*a*). By means of Ohm's law, Eq. (11*a*) may also be written as

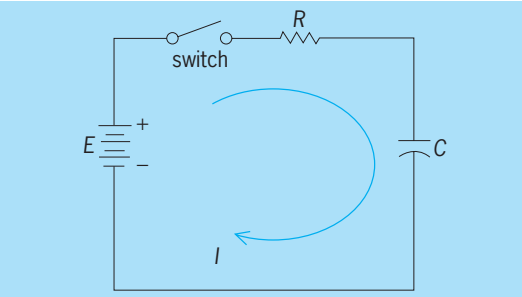

Fig. 4. A series *RC* circuit.

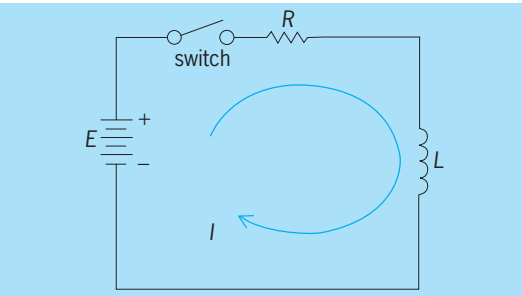

Fig. 5. A series *RL* circuit.

Eq. (11*b*). The total power dissipated in a circuit is

$$P = VI \tag{11a}$$

$$P = VI = \frac{V^2}{R} = I^2R \tag{11b}$$

the arithmetic sum of the power dissipated in each resistance.

Circuit response. In the circuits mentioned thus far, the circuit responds in an identical manner from the moment the circuit is excited (switches closed) through any extended period of time. This is not true of circuits typified by those of **Figs. 4** and **5**.

For instance, when the switch of Fig. 4 is first closed, a momentary current limited only by resistance R flows. As time passes, the capacitor, with capacitance C, charges and the voltage across it increases, eventually reaching a value equal to the applied voltage, at which time all flow of current ceases. The circuit current, given by Eq. (12), is in

$$i = \frac{E}{R}\epsilon^{-t/RC} \tag{12}$$

amperes. The product RC is known as the time constant of the circuit. The energy W in joules stored in a capacitance at any time is given by Eq. (13).

$$W = \frac{CE^2}{2} \tag{13}$$

See TIME CONSTANT.

For the circuit of Fig. 5, the initial current upon closing the switch is zero, since any attempt to cause a rate of change of current through the coil whose inductance is L induces a counter emf across the coil or inductor. Eventually this counter emf disappears and a steady-state current E/R flows indefinitely in the circuit.

The current at any time after closing the switch is given by Eq. (14) in amperes. The factor R/L is the

$$i = \frac{E}{R}(1 - \epsilon^{-Rt/L}) \tag{14}$$

time constant of the circuit. The energy W stored in an inductance at any time is given by Eq. (15).

$$W = \frac{LI^2}{2} \tag{15}$$

For a complete discussion of transient phenomena *see* ELECTRIC TRANSIENT

Robert L. Ramey

Bibliography. R. C. Dorf and J. A. Svoboda, *Introduction to Electric Circuits*, 6th ed., 2003; F. Hughes, *Fundamental DC-AC Circuits: Concepts and Experimentation*, 1990; J. D. Irwin and R. M. Nelms, *Basic Engineering Circuit Analysis*, 8th ed., 2005; D. E. Johnson et al., *Electric Circuit Analysis*, 3d ed., 1999; J. W. Nilsson and S. A. Riedel, *Electric Circuits*, 6th ed., 1999.

Direct-current generator

A rotating electric machine which delivers a unidirectional voltage and current. An armature winding mounted on the rotor supplies the electric power output. One or more field windings mounted on the stator establish the magnetic flux in the air gap. A voltage is induced in the armature coils as a result of the relative motion between the coils and the air gap flux. Faraday's law states that the voltage induced is determined by the time rate of change of flux linkages with the winding. Since these induced voltages are alternating, a means of rectification is necessary to deliver direct current at the generator terminals. Rectification is done by a commutator mounted on the rotor shaft. *See* COMMUTATION; WINDINGS IN ELECTRIC MACHINERY.

Carbon brushes, insulated from the machine frame and secured in brush holders, transfer the armature current from the rotating commutator to the external circuit. Brushes are held against the commutator under a pressure of 2–2.5 lb/in.^{2} (14–17 kilopascals). Armature current passes from the brush to brush holder through a flexible copper lead. In multipolar machines all positive brush studs are connected together, as are all negative studs, to form the positive and negative generator terminals. In most dc generators the number of brush studs is the same as the number of main poles. In modern machines brushes are located in the neutral position where the voltage induced in a short-circuited coil by the main pole flux is zero. The brushes continuously pick up a fixed, instantaneous value of the voltage generated in the armature winding. *See* COMMUTATOR.

The generated voltage is dependent upon speed n in revolutions per minute, number of poles p, flux per pole Φ in webers, number of armature conductors z, and the number of armature paths a. The

equation for the average voltage generated is

$$E_g = \frac{np\Phi z}{60a} \text{ volts}$$

The field windings of dc generators require a direct current to produce a magnetomotive force (mmf) and establish a magnetic flux path across the air gap and through the armature. Generators are classified as series, shunt, compound, or separately excited, according to the manner of supplying the field excitation current. In the separately excited generator, the field winding is connected to an independent external source. Using the armature as a source of supply for the field current, dc generators are also capable of self-excitation. Residual magnetism in the field poles is necessary for self-excitation. Series, shunt, and compound-wound generators are self-excited, and each produces different voltage characteristics.

When operated under load, the terminal voltage changes with change of load because of armature resistance drop, change in field current, and armature reaction. Interpoles and compensating, or pole-face, windings are employed in modern generators in order to improve commutation and to compensate for armature reaction. *See* ARMATURE REACTION.

Series generator. The armature winding and field winding of this generator are connected in series, as shown in **Fig. 1**. Terminals T_1 and T_2 are connected to the external load. The field mmf aids the residual magnetism in the poles, permitting the generator to build up voltage. The field winding is wound on the pole core with a comparatively few turns of wire of large cross section capable of carrying rated load current. The magnetic flux and consequently the generated emf and terminal voltage increase with increasing load current. **Figure 2** shows the external characteristic or variation of terminal voltage with load current at constant speed. Series generators are suitable for special purposes only, such as a booster in a constant voltage system, and are therefore seldom employed.

Shunt generator. The field winding of a shunt generator is connected in parallel with the armature winding, as shown in **Fig. 3**. The armature supplies both load current I_t and field current I_f. The field current is 1–5% of the rated armature current I_a, the higher value applying to small machines. The field winding resistance is fairly high since the field consists of many turns of small-cross-section wire. For voltage buildup the total field-circuit resistance must be below a critical value; above this value the generator voltage cannot build up. The no-load voltage to which the generator builds up is varied by means of a rheostat in the field circuit. The external voltage characteristic (Fig. 2) shows a reduction of voltage with increases in load current, but voltage regulation is fairly good in large generators. The output voltage may be kept constant for varying load current conditions by manual or automatic control of the rheostat in the field circuit. A shunt generator will not maintain a large current in a short circuit in the external circuit, since the field current at short circuit is zero.

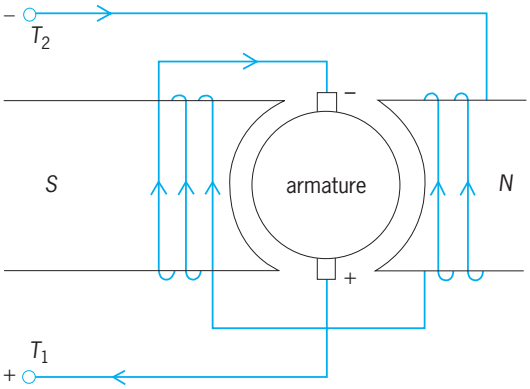

Fig. 1. Series generator.

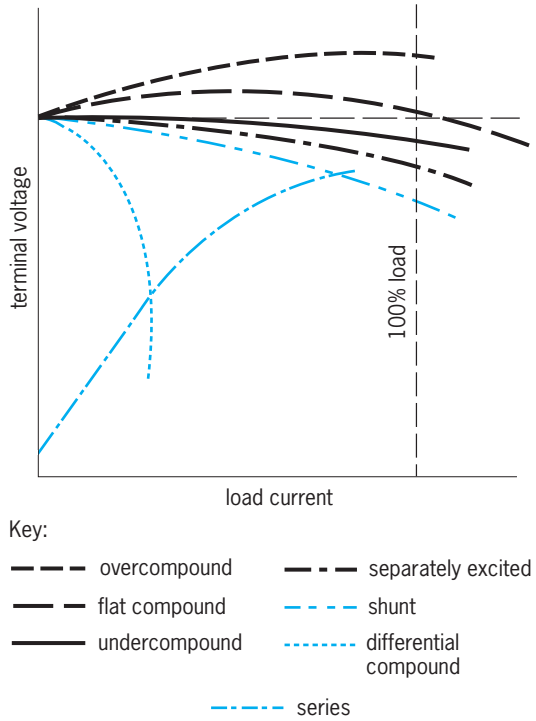

Fig. 2. External characteristics of dc generators.;

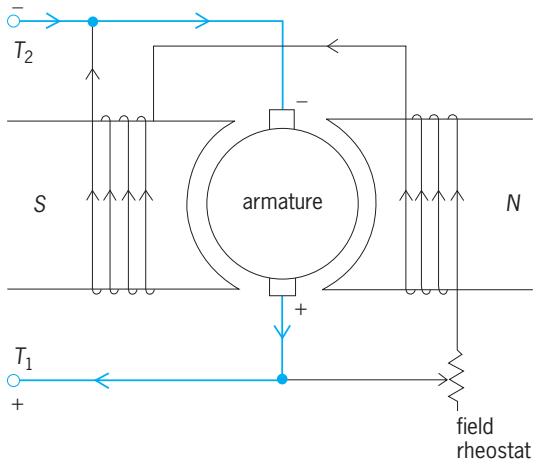

Fig. 3. Shunt generator.

The shunt generator is suitable for fairly constant voltage applications, such as an exciter for ac generator fields, battery charging, and for electrolytic work requiring low-voltage and high-current capacity. Prior to the use of the alternating-current generator and solid-stage rectifying devices in automobiles, a shunt generator, in conjunction with automatic regulating devices, was used to charge the battery and supply power to the electrical system. Shunt-wound generators are well adapted to stable operation in parallel.

Compound generator. This generator has both a series field winding and a shunt field winding. Both windings are on the main poles with the series winding on the outside. The shunt winding furnishes the major part of the mmf. The series winding produces a variable mmf, dependent upon the load current, and offers a means of compensating for voltage drop. **Figure 4** shows a cumulative-compound connection with series and shunt fields aiding. A diverter resistance across the series field is used to adjust the series field mmf and vary the degree of compounding. By proper adjustment a nearly flat output voltage characteristic is possible. Cumulative-compound generators are overcompounded, flat-compounded, or undercompounded, as shown by the external characteristics in Fig. 2. The shunt winding is connected across the armature (short-shunt connection) or across the output terminals (long-shunt connection). Figure 4 shows the long-shunt connection.

Voltage is controllable over a limited range by a rheostat in the shunt field circuit. Compound generators are used for applications requiring constant voltage, such as lighting and motor loads. Generators used for this service are rated at 125 or 250 V and are flat or overcompounded to give a regulation of about 2%. An important application is in steel mills which have a large dc motor load. Cumulative-compound generators are capable of stable operation in parallel if the series fields are connected in parallel on the armature side by means of a low-resistance cable or bus referred to as an equalizer bus. By paralleling the series fields of all compound generators, the equalizer bus maintains the same voltage across all series fields.

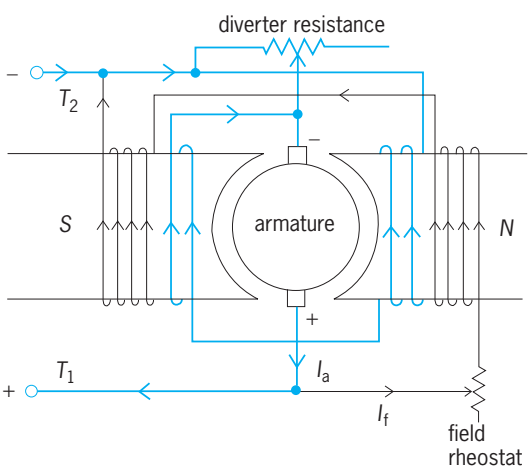

Fig. 4. Cumulative compound generator. The long-shunt connection is seen in this example.

In the differentially compounded generator the series field is connected to oppose the shunt field mmf. Increasing load current causes a large voltage drop due to the demagnetizing effect of the series field. The differentially compounded generator has only a few applications, such as arc-welding generators and special generators for electrically operated shovels.

Separately excited generator. The field winding of this type of generator is connected to an independent dc source. The field winding is similar to that in the shunt generator. Separately excited generators are among the most common of dc generators, for they permit stable operation over a very wide range of output voltages. The slightly drooping voltage characteristic (Fig. 2) may be corrected by rheostatic control in the field circuit. Applications are found in special regulating sets and in laboratory and commercial test sets.

Special types. Besides the common dc generators discussed in this article there are a number of special types, including the homopolar, third-brush, diverter-pole, and Rosenberg generators. For discussion of the Amplidyne, Regulex, and Rototrol *see* DIRECT-CURRENT MOTOR

Commutator ripple. The voltage at the brushes of dc generators is not absolutely constant. A slight high-frequency variation exists, which is superimposed upon the average voltage output. This is called commutator ripple and is caused by the cyclic change in the number of commutator bars contacting the brushes as the machine rotates. The ripple decreases as the number of commutator bars is increased and is usually ignored. In servomechanisms employing a dc tachometer for velocity feedback, the ripple frequency is kept as high as possible. *See* ELECTRIC ROTATING MACHINERY; GENERATOR.

Robert T. Weil, Jr.

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; A. E. Fitzgerald, C. Kingsley, and S. D. Umans, *Electric Machinery*, 6th ed., 2003; S. A. Nasar, *Handbook of Electric Machines*, 1987; P. F. Ryff, *Electric Machinery*, 2d ed., 1994.

Direct-current motor

An electric rotating machine energized by direct current and used to convert electric energy to mechanical energy. It is characterized by its relative ease of speed control and, in the case of the series-connected motor, by an ability to produce large torque under load without taking excessive current. Output of this motor is given in horsepower, the unit of mechanical power. Normal full-load values of voltage, current, and speed are generally given.

Direct-current motors are manufactured in several horsepower-rating classifications: (1) subfractional, approximately 1–35 millihorsepower (mhp) or 0.75–25 W; (2) fractional, $^1/_{40}$–1 horsepower (hp)

or 20–750 W; and (3) integral, $^1/_2$ to several hundred horsepower or 0.4 to several hundred kilowatts.

The standard line voltages applied to dc motors are 6, 12, 27, 32, 115, 230, and 550 V. Occasionally they reach higher values.

Normal full-load speeds are 850, 1140, 1725, and 3450 rpm. Variable-speed motors may have limiting rpm values stated.

Protection of the motor is afforded by several types of enclosures, such as splash-proof, drip-proof, dust-explosion-proof, dust-ignition-proof, and immersion-proof enclosures. Some motors are totally enclosed.

The principal parts of a dc motor are the frame, the armature, the field poles and windings, and the commutator and brush assemblies (**Fig. 1**). The frame consists of a steel yoke of open cylindrical shape mounted on a base. Salient field poles of sheet-steel laminations are fastened to the inside of the yoke. Field windings placed on the field poles are interconnected to form the complete field winding circuit. The armature consists of a cylindrical core of sheet-steel disks punched with peripheral slots, air ducts, and shaft hole. These punchings are aligned on a steel shaft on which is also mounted the commutator. The commutator, made of hard-drawn copper segments, is insulated from the shaft. The segments are insulated from each other by mica. Stationary carbon brushes in brush holders make contact with the segments of the commutator. Copper conductors which are placed in the insulated armature slots are interconnected to form a reentrant lap or wave style of winding. *See* COMMUTATION; WINDINGS IN ELECTRIC MACHINERY.

Principles. Rotation of a dc motor is produced by an electromagnetic force exerted upon current-carrying conductors in a magnetic field. For basic principles of motor action *see* MOTOR

In **Fig. 2**, forces act on conductors on the left path of the armature to produce clockwise rotation. Those conductors on the right path, whose current direction is reversed, also will have forces to produce clockwise rotation. The action of the commutator allows the current direction to be reversed as a conductor passes a brush.

The net force from all conductors acting over an average radial length to the shaft center produces a torque T given by Eq. (1), where K_t is a conversion

$$T = K_t \Phi I_a \tag{1}$$

and machine constant, Φ is net flux per pole, and I_a is the total armature current.

The voltage E, induced as a counter electromotive force (emf) by generator action in the parallel paths of the armature, plus the voltage drop $I_a R_a$ through the armature due to armature current I_a and armature resistance R_a, must be overcome by the total impressed voltage V from the line. Voltage relations can be expressed by Eq. (2).

$$V = E + I_a R_a \tag{2}$$

The counter emf and motor speed n are related by Eq. (3), where K is a conversion and machine constant.

$$n = \frac{E}{K\Phi} \tag{3}$$

Mechanical power output can be expressed by Eq. (4), where n is the motor speed in rpm, T is the

$$\text{Horsepower} = \frac{2\pi nT}{33,000} = \frac{2\pi nT'}{44,742} \tag{4}$$

torque developed in pound-feet, and T' is the torque developed in newton-meters.

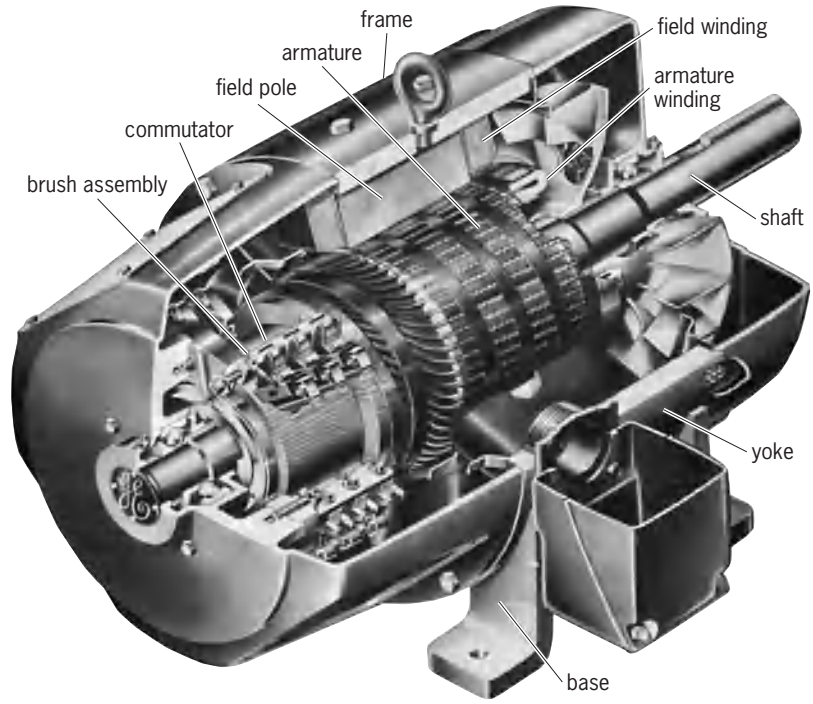

Fig. 1. Cutaway view of typical dc motor. (*General Electric Co.*)

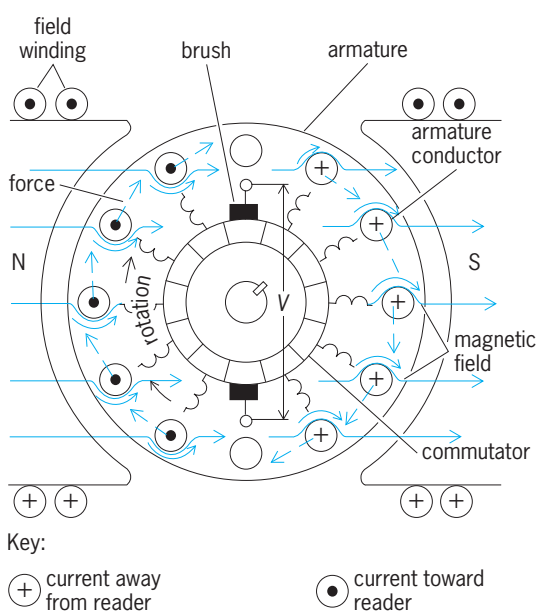

Fig. 2. Rotation in a dc motor.

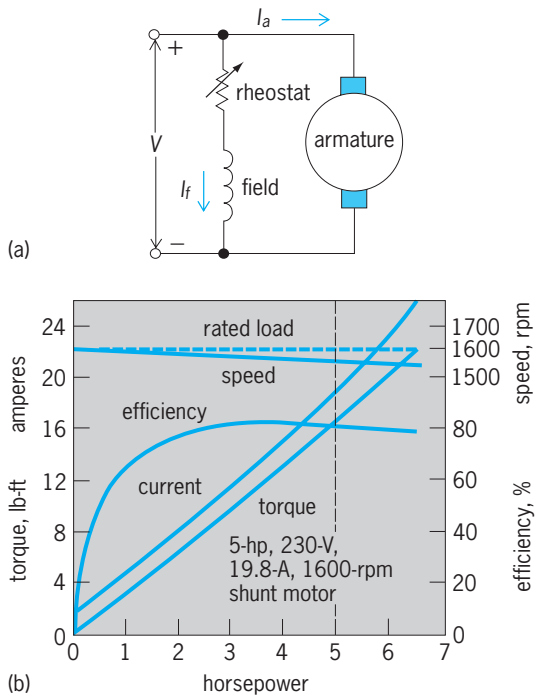

Fig. 3. Shunt motor. (*a*) Connections. (*b*) Typical operating characteristics.

By use of these four equations, the steady-state operation of the dc motor may be determined.

Types. Direct-current motors may be categorized as shunt, series, compound, or separately excited.

Shunt motor. The field circuit and the armature circuit of a dc shunt motor are connected in parallel (**Fig. 3***a*). The field windings consist of many turns of fine wire. The entire field resistance, including a series-connected field rheostat, is relatively large. The field current and pole flux are essentially constant and independent of the armature requirements. The torque is therefore essentially proportional to the armature current.

In operation an increased motor torque will be produced by a nearly equal increase in armature current, Eq. (1), since K_t and Φ are constant. Increased I_a produces an increase in the small voltage I_aR_a, Eq. (2). Since V is constant, E must decrease by the same small amount resulting in a small decrease in speed n, Eq. (3). The speed-load curve is practically flat, resulting in the term "constant speed" for the shunt motor. Typical characteristics of the motor are shown in Fig. 3*b*.

Typical applications are for load conditions of fairly constant speed, such as machine tools, blowers, centrifugal pumps, fans, conveyors, wood- and metalworking machines, steel, paper, and cement mills, and coal or coke plant drives.

Series motor. The field circuit and the armature circuit of a dc series motor are connected in series (**Fig. 4***a*). The field winding has relatively few turns per pole. The wire must be large enough to carry the armature current. The flux Φ of a series motor is nearly proportional to the armature current I_a which produces it. Therefore, the torque, Eq. (1), of a series motor is proportional to the square of the armature current, neglecting the effects of core saturation and armature reaction. An increase in torque may be produced by a relatively small increase in armature current.

In operation the increased armature current, which produces increased torque, also produces increased flux. Therefore, speed must decrease to produce the required counter emf to satisfy Eqs. (1) and (3). This produces a variable speed characteristic. At light loads the flux is weak because of the small value of armature current, and the speed may be excessive. For this reason series motors are generally connected permanently to their loads through gearing.

The characteristics of the series motor are shown in Fig. 4*b*. Typical applications of the series motor are to loads requiring high starting torques and variable speeds, for example, cranes, hoists, gates, bridges, car dumpers, traction drives, and automobile starters.

Compound motor. A compound motor has two separate field windings. One, generally the predominant field, is connected in parallel with the armature circuit; the other is connected in series with the armature circuit (**Fig. 5**).

The field windings may be connected in long or short shunt without radically changing the operation of the motor. They may also be cumulative or differential in compounding action. With both field windings, this motor combines the effects of the shunt and series types to an extent dependent upon the degree of compounding. Applications of this motor are to loads requiring high starting torques and somewhat variable speeds, such as pulsating loads, shears, bending rolls, plunger pumps, conveyors, elevators, and crushers. *See* DIRECT-CURRENT GENERATOR.

Separately excited motor. The field winding of this motor is energized from a source different from that of the armature winding. The field winding may be of either the shunt or series type, and adjustment of the applied voltage sources produces a wide range of speed and torque characteristics. Small dc motors may have permanent-magnet fields with armature excitation only. Such motors are used with fans,

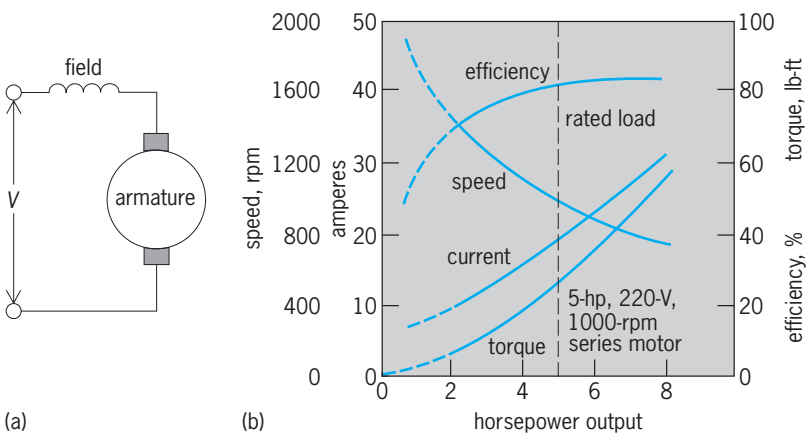

Fig. 4. Series motor. (*a*) Connection. (*b*) Typical operating characteristics.

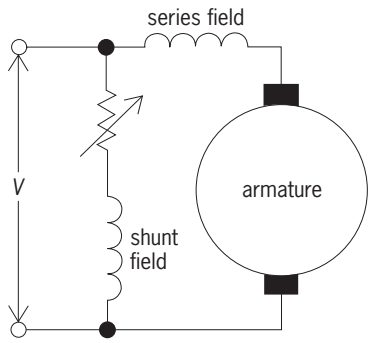

Fig. 5. Connection of a compound motor.

blowers, rapid-transfer switches, electromechanical activators, and programming devices. *See* ELECTRIC ROTATING MACHINERY. L. F. Cleveland

Bibliography. S. J. Chapman, *Electric Machinery Fundamentals*, 4th ed., 2005; A. E. Fitzgerald, C. Kingsley, and S. D. Umans, *Electric Machinery*, 6th ed., 2003; I. L. Kosow, *Electric Machinery and Transformers*, 2d ed., 1991; G. McPherson and R. D. Laramore, *An Introduction to Electrical Machines and Transformers*, 2d ed., 1990; P. F. Ryff, *Electrical Machinery*, 2d ed., 1994.

Direct-current transmission

The conveyance of electric power by conductors carrying unidirectional currents. *See* DIRECT CURRENT.

Alternating current (ac) is universally accepted for the generation of electric power and for its distribution to consumers. In order to deliver power from generating stations to load centers and to facilitate interchange of power between different power companies voltages are stepped up to transmission levels. Transmission lines typically have one or two sets of three-phase ac conductors suspended from towers. Underground or submarine cables are used less frequently, in cases where the additional cost is considered justifiable for technical or environmental reasons. They are terminated at substations which house transformers to adjust the voltage and circuit breakers to establish appropriate connections. *See* ALTERNATING CURRENT; ELECTRIC POWER TRANSMISSION; TRANSMISSION LINES.

Power companies face increasing pressure to reconcile increasing load demands with financial, environmental, and resource constraints. This emphasizes the need to operate economically and effectively while preserving reliability. Consequently, it is often preferable to interconnect systems into larger power pools, even crossing national boundaries, rather than seek total self-reliance on exclusive generation capability within each company. *See* ELECTRIC POWER SYSTEMS.

Within this established pattern of ac power systems, application of transmission equipment operating with direct current (dc) has increased since the late 1950s. Although the adoption of dc over a century ago by Thomas Edison for an entire system from generation to load has long been superseded by ac practice, special circumstances arise where dc transmission has economic or technical advantages. Thus dc becomes a medium for interconnecting points in ac systems. Special converters within converter stations change the ac to dc (rectification) and dc back to ac (inversion) to complete the power interchange.

Rationale. A dc line with two conductors is cheaper to construct and often has lower power losses than a three-phase ac line rated for the same power. Moreover, the same dc line is often considered as equal in reliability of service to a double-circuit three-phase line. The economic advantages are proportional to the line length but are offset by the substantial cost of the converting equipment. On this criterion alone, there would be a break-even distance beyond which dc would be cheaper: several hundred kilometers for overhead lines and much less for more costly underground and submarine cables. Indeed, this has been an important factor in a small minority of projects. The break-even distance is specific to the conditions and financing of each project, and it obviously contradicts the zero distance of back-to-back installations. However, several other factors influence the selection of dc.

If the ac frequencies at the converting stations are nominally the same but controlled separately, their frequency independence is maintained by the dc link. In other words, the dc system is an asynchronous link. This is the justification of many back-to-back schemes such as the ties between regions in the United States and between European countries.

Although North America operates at 60 Hz and most other parts of the world operate at 50 Hz, on occasion there is a need to interconnect ac systems having different nominal frequencies. The asynchronous nature of dc serves as a frequency changer with the control action of each converter synchronized to its local ac frequency. The Sakuma frequency changer in Japan is a back-to-back scheme, connecting two regions which, for historical reasons, have frequency standards of 50 and 60 Hz.

The electrical shunt capacitance of cables is charged and discharged at the frequency of the voltage. Since the capacitance is proportional to distance, an ac cable longer than a few tens of miles is loaded to its thermal rating by a capacitor charging current, with no power being conveyed to the remote termination. Unless the cable can be sectioned for intermediate compensating measures, dc is obligatory for many cable applications. This is especially the case for submarine links where overhead lines are not an option and economic considerations of break-even distance may or may not have favored dc.

All parts of an ac system function at the same nominal frequency. A system is designed to ensure that the generators do not lose synchronism despite load variations and large fault disturbances. The system is then considered to be dynamically stable. This becomes more of a challenge when the generating stations are geographically dispersed and remote from load centers, as opposed to the relatively tightly knit

systems found in Europe. Should an ac connection to another system be contemplated, the combined system is intended to operate in synchronism, although the dynamic stability in either system may deteriorate below acceptable security. In comparison, a dc interconnection maintains the dynamic independence of each system. For example, remote hydroelectric generation at Churchill Falls in Labrador and on rivers entering James Bay prevents Hydro-Quebec from establishing synchronous connections with neighboring power companies for reasons of potential instability. Mutual interties and export contracts for electric power to the New York power pool and New England utilities have been implemented by several dc links. The question then turns to the selection of power rating and the location of the dc terminal stations.

The automatic control circuits at the converting stations permit the dc power to be accurately set at a value determined by the system control center. Furthermore, the dc power is maintained during dynamic ac frequency disturbances and can be rapidly changed (modulated) in as little as a few milliseconds, on demand. The same control permits the direction of dc power to be reversed equally quickly. This ability to ride through disturbances and to permit precise power scheduling and modulation responsive to dynamic needs has become of increasing value in the operation of power systems. The power flow in an individual ac line cannot be independently controlled to the degree offered by dc transmission.

Applications. In most instances, manufacturers bid internationally to supply equipment adapted to the needs and specification of each project. If the project is not contingent on building generating stations, a pressing need may reduce the interval between conception and operation to as little as 2 years, although, as with many large power projects, the process can take several years and the transmission capacity may be scheduled in stages. Systems in operation around the world are of varying lengths, operating voltages, transmission capacities (megawatts), and applications. There are 77 systems in service and another 16 either under construction or under active consideration. Of the total, 25 are in the United States and Canada; the remaining 68 involve 27 other countries. Compared with the first system established in 1954, which supplies the island of Gotland with 30 MW of power from the Swedish mainland, the total nominal power rating of dc transmission from the Nelson River hydroelectric generation in Manitoba, Canada, is 3420 MW.

Categories of application are (1) long-distance overhead dc lines conveying bulk power from remote generating stations to load centers, such as Itaipu, Brazil; (2) long-distance seasonal or daily power interchange, such as the Pacific DC Intertie linking Oregon with southern California; (3) submarine cables, such as for Cross Channel 2 between Great Britain and France; (4) zero-distance connections (back-to-back schemes) between adjacent systems, such as Miles City, Montana; and (5) underground dc cables in urban areas, such as in Kingsnorth, London.

Equipment. A typical dc converter station contains conventional ac equipment in its ac switchyard supplemented by equipment specific to the ac-dc conversion. Solid-state converters are connected on the dc side with a center neutral point which is usually connected to a remote ground electrode. Balanced dc currents are circulated on each pole at plus and minus dc voltages with respect to ground (**Fig. 1**). *See* ELECTRIC POWER SUBSTATION.

Each converter is connected to the ac system through transformers which, in conjunction with the control circuits, determine the dc pole voltages. Filters remove ac harmonics arising from the conversion process, and the dc current is smoothed by a reactor which may be supplemented by more dc filter components. Filtering attenuates potential radio and telecommunication interference as well as avoiding deterioration of the ac waveforms. *See* ELECTRIC FILTER.

Although mercury-arc converters were applied until 1974, it is now the practice to use solid-state devices. The converters are made up of groups of Graetz-connected bridges, each containing six valves. Each valve uses a matrix of thyristors which are simultaneously gated into conduction through fiber optics in the firing sequence indicated by the numbers in the inset of the illustration. Each thyristor can withstand several kilovolts in its nonconducting state and several thousand amperes when conducting. Although the current rating is often sufficient without connecting thyristors in parallel, many must be connected in series to achieve the voltage

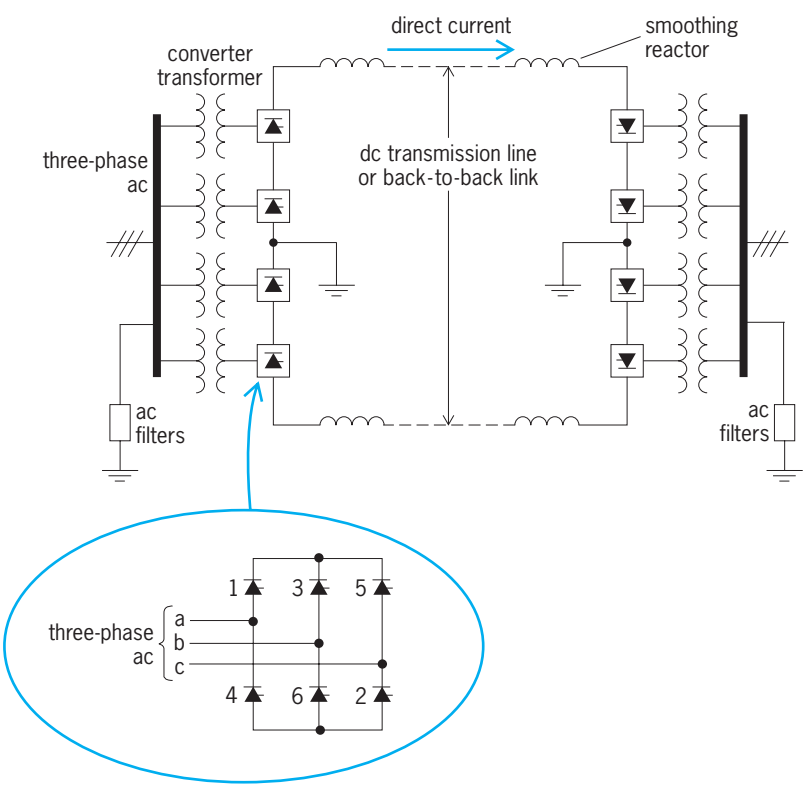

Fig. 1. Diagram of typical dc system.

withstand requirements. The result is a unidirectional valve which is controlled to turn on once per ac cycle and which turns off when its voltage reverses. Depending on the points in the cycle at which the valves are sequentially fired, the dc voltage is controlled from rectification through to inversion when the power flow is reversed and the dc converted back to ac. The six-valve bridges are often connected as series pairs in 12-pulse groups with a staggered firing sequence in order to produce a smoother dc waveform and reduced ac harmonics. *See* COMMUTATION; CONVERTER; SEMICONDUCTOR RECTIFIER.

Generally, the converters are indoors, with the connections to outdoor converter transformers made via insulated conductors (bushings) through the walls of the valve hall. The bushings are often integral parts of the transformer structures. *See* TRANSFORMER.

Except in back-to-back schemes, the converters are connected through the dc switchyard to the lines interconnecting the converter stations. For overhead lines, conductors for the two poles of opposite polarity of voltage are typically carried on the same tower. Balanced pole currents avoid ground current. In principle, ground can serve as a return-current conductor. In practice, local restrictions on the size and duration of ground current due to concerns for effects on buried pipelines often limit monopolar operation to short emergency periods. In order to achieve an effective ground connection capable of carrying, even temporarily, large currents, a separate insulated overhead line or cable is run from the converting station to a suitable ground-electrode site. The possibility of rapidly discontinuing dc current by converter control avoids the need for dc circuit breakers to deenergize dc line faults.

Trends. As a highly innovative technology, dc transmission has broadened its scope of application along with parallel development of equipment. Valves apply state-of-the art power semiconductors with advanced cooling and gating methods. Control circuits operate in real time and are now computer-based. They are highly responsive to various control strategies within the systemwide control, extending back to the system control center. The dc substations are typically interlinked by microwave communication. Circuit breakers are available to interrupt large dc currents at high voltages encountered in dc systems and can be used for special purposes. The ratings of individual thyristors have increased, resulting in lower converter sizes and costs. Large power thyristors can be triggered by a light pulse directly into the thyrstor, resulting in simplification of the electronic gate firing circuits. *See* CIRCUIT BREAKER.

A significant advance is the extension of point-to-point two-terminal practice to the construction of dc networks containing three or more converter stations for a coordinated power interchange in a so-called multiterminal system. It has been achieved by tapping the existing Italy-Sardinia connection to feed power to Corsica. A five-terminal network takes power by dc from James Bay through Quebec into New England.

China, India, and developing countries in southern Africa are actively exploring the possibility of increasing the dc voltage rating from the present maximum of ±600 kV (the Itaipu project in Brazil) to ±800 kV. This would allow power flow on one transmission line to approach 6000 MW, enough power to supply a population of a few million. However, building a transmission line of such large capacity faces challenges, including maintenance of service if the line should be unexpectedly removed from service due to adverse environmental conditions.

The blackouts in the northeastern United States and Canada (August 15, 2003) and Italy (September 28, 2003) have reawakened the idea that a blackout could be blocked from spreading by dividing the power system into segments and inserting back-to-back dc transmission units in the ac power lines wherever they cross-segment boundaries. Quebec interfaces to the northeastern U.S. power system through dc transmission, which prevented the blackout of August 15, 2003 from affecting it.

As the price of fossil fuels increases, causing electric power prices to rise as well, there are incentives to search for renewable energy. Hydroelectricity, wind, solar, and geothermal energy as well as less expensive energy sources such as some forms of coal will spur development of long-distance electric power transmission between energy sources and load centers. It is expected that high-voltage dc transmission will play a vital role in helping deliver less expensive energy to where it is most needed. *See* ELECTRIC POWER GENERATION; ENERGY SOURCES.

VSC transmission. Advanced high-voltage dc transmission facilities are being built using voltage sourced transmission or voltage source converter (VSC) transmission (**Fig. 2**). These new developments apply the VSC converter bridge, which uses gate turnoff thyristors (GTOs) and insulated gate bipole transistors (IGBTs). Its special properties include the ability to independently control real and reactive power at the connection bus to the ac system. Reactive power can be either capacitive or inductive and can be controlled to quickly change from one to the other.

A voltage source converter that functions as an inverter does not require an active ac voltage source to commutate into as does the conventional line commutated converter. The VSC inverter can generate an ac three-phase voltage and supply electricity to a load as the only source of power. It does require harmonic filtering, harmonic cancellation, or

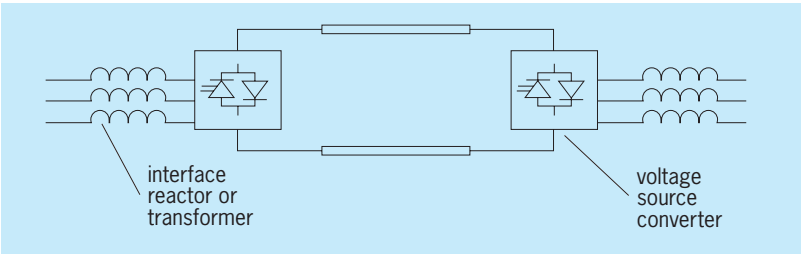

Fig. 2. Simplified diagram of VSC transmission.

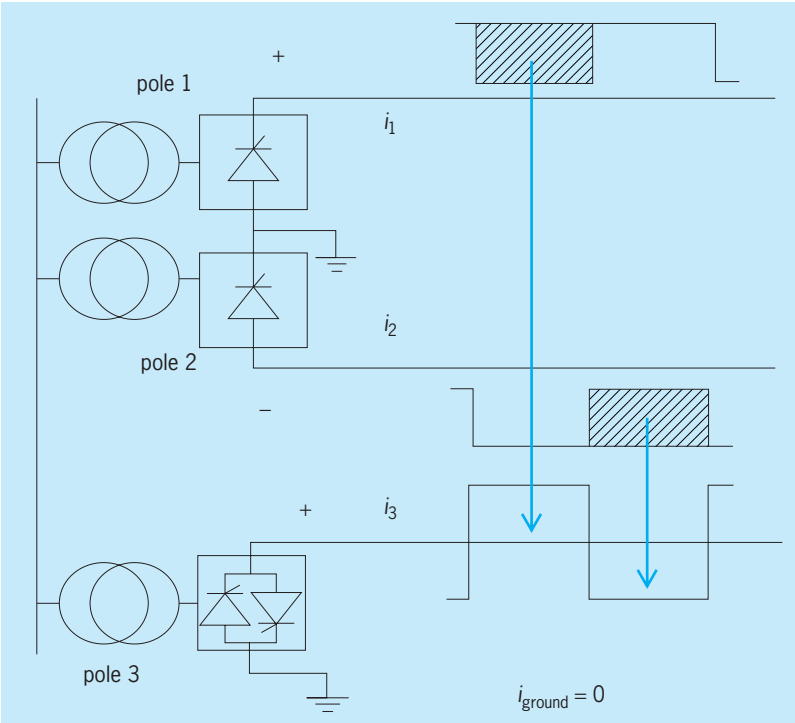

Fig. 3. Configuration of a tripole showing how current flows on each pole conductor.

pulse-width modulation to provide an acceptable ac voltage waveshape.

The first VSC transmission system went into operation in 1997 in Sweden. VSC converter bridges can be applied in back-to-back configuration and for major transmission systems. The back-to-back VSC link is the ultimate transmission and power flow controller. It can control and reverse power flow easily, and control reactive power independently on each side. With a suitable control system, it can control power to enhance and preserve ac system synchronism, and act as a rapid phase-angle power flow regulator with 360° range of control. The first North American back-to-back VSC unit was the Eagle Pass 36-MW interconnection between Mexico and Texas commissioned in 2000. For point-to-point transmission systems, power levels reaching 1100 MW and greater are possible. A 180-MW interconnection using underground cable between Queensland and New South Wales in Australia was completed in 2000. Other installations include a 330-MW under water interconnection between Long Island and Connecticut.

There is considerable flexibility in the configuration of VSC converter bridges. Many two-level converter bridges can be assembled with appropriate harmonic cancellation properties to generate acceptable ac system voltage waveshapes. Another option is to use multilevel converter bridges to provide harmonic cancellation. Both two-level and multilevel converter bridges can utilize pulse-width modulation to eliminate low-order harmonics. However, with pulse-width modulation, high-pass filters may still be required since higher-order harmonics are then generated.

The great benefit derived from inexpensive power transmission cable is environmental, although relative ease of licensing is also an advantage, since licensing of less expensive, overhead transmission lines is difficult and time consuming. Cables for VSC transmission are cheaper to bury than conventional power cables, and the lead time for licensing can be significantly reduced, allowing the transmission to be brought into service sooner than is possible for overhead transmission lines. For some locations it may be impossible to build an overhead transmission line, but buried cable may be an acceptable, although relatively expensive, option.

The maximum ratings for VSC transmission in one cable transmission circuit so far are ±300 kV and 1850 amps. This should increase in the future.

Bidirectional valves. Current flows through conventional high-voltage dc thyristor converter valves in one direction only. It is possible to construct a valve with either bidirectional thyristors, or two thyristor strings in parallel but with opposite current-conducting directions. Under controlled operation, converter bridges with bidirectional valves provide additional benefits, the most significant being three-pole (tripole) operation without a neutral return path. Such operation can be applied to achieve the maximum power in the three phases of an ac line if it is converted to dc. Tripole dc transmission is a new concept that is being actively studied.

The unique operation of tripole is based on having two poles in bipole configuration (**Fig. 3**). The positive pole raises its current to an overload value for several minutes while the negative pole current is less than rated current. The current in the third pole makes up the difference with its voltage polarity maintaining the direction of desired power flow. After several minutes, the current in the first pole reduces to allow its conductor to cool, and the current in the second pole picks up to the overload value at which the first pole was operating. The transition of current between overload and light load levels in the first and second poles requires that the current in the third pole reverse both its current direction and voltage polarity in order that power flow in the three poles remains steady. Hence, the valves in the third pole must be bidirectional.

The converters have no overload capacity above their rating and so the valves of the first and second poles must be rated to the overload level. The transmission or cable conductors, on the other hand, can easily withstand brief current overloads. If the period of overload and light load in these poles is of the order of several minutes, conductor temperature variations are minimal, of the order of several degrees, allowing the cyclic overload operation to continue. The current level in the third pole does not need to overload, but it does change polarity. The transition period is over a short period of time of several seconds under controlled conditions that maintain steady power transmission.

Significantly more power can be made to flow on a long ac transmission line when it is converted to dc in this way than is possible if the transmission

line remains ac. Consequently, maximum use can be made of existing transmission lines.

John Reeve; Dennis Woodford

Bibliography. J. Arillaga, *High Voltage Direct Current Transmission*, 2d ed., 1998; D. P. Carroll, *HVDC Transmission Systems*, 2001; Electric Power Research Institute, *Methodology for Integration of HVDC Links in Large AC Systems—Phase 1: Reference Manual*, Rep. EL-3004, 1983; E. W. Kimbark, *Direct Current Transmission*, vol. 1, 1971; E. Uhlmann, *Power Transmission by Direct Current*, 1975.

Direction-finding equipment

Equipment used for determining the direction from which a received signal is being radiated. Direction finders serve a variety of applications. Direction finders can be used for navigation purposes to determine the heading of a vehicle relative to a transmitter at a known location, and to determine the location of the vehicle when the range to that transmitter is known, as in the case of VOR/DME instrumentation. Direction finding may also be used to determine vehicle heading for guidance along a desired route toward a given destination. Surveillance uses of direction finders include discovering and locating covert transmitters used for espionage, locating the emitter in search-and-rescue missions, and tracking crewless objects in scientific missions. A direction finder may be implemented either as a ground station used to track a moving source or on a moving vehicle. Bearing information obtained by the ground station can also be passed to a moving vehicle that is equipped with a transmitter that can be tracked by the ground station. As mentioned above, a direction finder may also be found on a moving vehicle, such as an aircraft or a ship, to determine the vehicle's heading relative to a transmitter or the transmitter's bearing relative to the known track of the vehicle. *See* DISTANCE-MEASURING EQUIPMENT; VOR (VHF OMNIDIRECTIONAL RANGE).

Principles of direction finding. Most direction-finding systems are based on one or the other of two principles, where one exploits antenna gain characteristics and the other exploits phase and frequency relationships observed by a moving or a multiple-element fixed antenna.

Antenna gain. Direction-finding systems based on antenna gain take advantage of variations of the antenna gain pattern, whereby the received signal amplitude varies from a maximum to a minimum according to its angle of arrival. Signal-minimum systems exploit the distinct null found in an antenna's gain pattern. Some of the earliest direction-finding equipment using loop antennas was conceived on this principle. The advent of higher-frequency systems and highly directive antennas, such as phased arrays and parabolic dishes, provided the basis for signal-maximum systems, which can pinpoint the direction in which the signal is strongest. *See* ANTENNA (ELECTROMAGNETISM).

Induced Doppler frequency modulation. The principle of induced Doppler frequency modulation requires movement of the antenna to cause additional Doppler shift to the original signal. The antenna motion is generated in such a way that the resulting shift in received frequency will provide information about the direction of the incoming signal with respect to that of the known antenna motion. The frequency of the signal received by the moving antenna will vary in relation to that received by a fixed antenna, the latter serving as a reference signal to extract the content induced by the moving antenna. The antenna motion most suited for omnidirectional purposes is circular. For this, the induced Doppler is zero at the instant when the antenna motion is perpendicular to the bearing of the received signal. *See* DOPPLER EFFECT.

Airborne automatic direction finder. Airborne automatic direction finders (ADFs) are constrained by size and weight. For their simplicity, signal-minimum systems with loop antennas are widely used. Although there are antenna configurations that use a continuously rotating or steerable antenna to optimize the reception of a signal from any direction, a direction finder with orthogonal fixed-loop antennas has the advantage of not having any mechanically moving parts in the antenna assembly, reducing maintenance requirements and improving reliability. The signals from the loop antennas carry bearing information in their relative amplitudes as determined by the loop-antenna gain pattern (**Fig. 1**).

The patterns in Fig. 1 illustrate the antenna response to a plane-wave signal that arrives at a bearing of β relative to vertical. The loop antenna cannot determine whether the signal arrives at β or at $\beta + 180^\circ$. Therefore, an additional (omnidirectional) antenna is usually provided to fully resolve the bearing measurement. To extract this information, the signals from the loop antennas are amplitude modulated separately with modulation signals that are mutually orthogonal in phase. When the amplitude-modulated signals are combined, the resulting signal has a specific phase relationship with respect to the original modulation signal. This relationship is proportional to the bearing β of the received signal, with an ambiguity of 180°. However, the signal

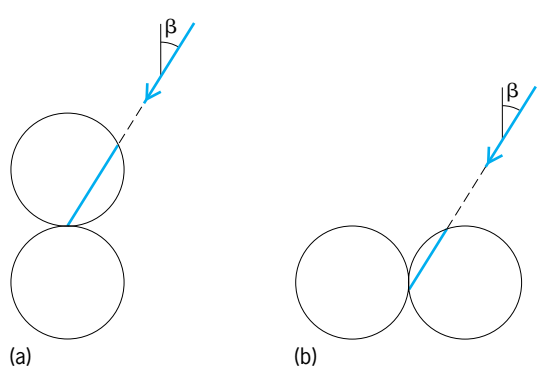

Fig. 1. Signal amplitude differences between orthogonal loop-antenna patterns. (*a*) Signal amplitude = *A* cos β. (*b*) Signal amplitude = *A* sin β.

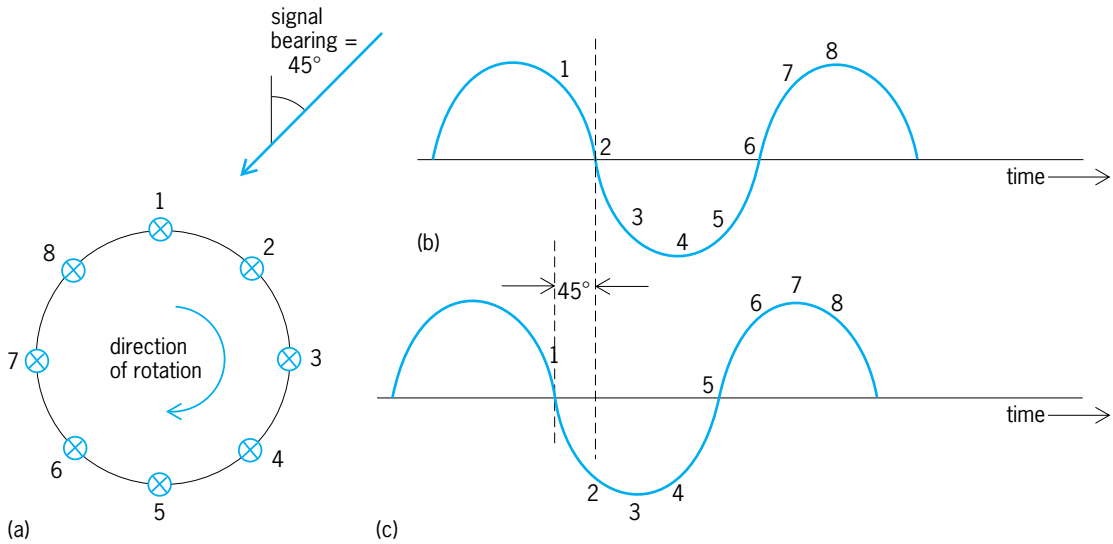

Fig. 2. Induced Doppler direction finder using circular antenna motion. (a) Diagram of equipment with sequential switching of antennas 1–8. (*b*) Extracted Doppler profile, which goes to zero at locations 2 and 6, where the antenna is moving orthogonally to an incoming signal with bearing of 45°. (*c*) Reference profile based on antenna motion.

from the omnidirectional antenna is also added to the amplitude-modulated signals from the loop antennas, contributing what is known as the sense term to the combined signal.

After synchronous demodulation to move the time dependence of the received signal, a phase comparison of the baseband signal with the original modulation signal will yield the value of β. Without the sense term, the arbitrary value of the residual phase error would produce a 180° phase ambiguity to the solution of β. *See* AMPLITUDE-MODULATION DETECTOR.

Doppler direction finder. A Doppler direction finder using circular antenna motion induces a sinusoidal frequency variation to the original frequency content of the received signal (**Fig. 2***a*). To obtain the bearing information, the signal first needs to be demodulated with a frequency-modulation (FM) discriminator. The discriminator output is a sinusoid (Fig. 2*b*) whose phase is shifted with respect to the reference signal (Fig. 2*c*) by an amount that corresponds to the bearing of the received signal. A post-FM demodulator is needed to extract that phase angle.

In most implementations of this type of direction finder, the antenna motion is synthetically generated by sequentially switching through a series of antennas arranged in a circle (Fig. 2*a*). The measurements made from each antenna provide information that can be implicitly reconstructed according to principles of signal sampling.

Accuracy factors. Direction finders based on antenna gain depend on the fidelity of the gain pattern. Pattern imperfections will affect the accuracy of the derived bearing. All direction finders are susceptible to contamination of the true signal with a spurious reflection, causing measurement error. Low- and medium-frequency direction finders are particularly prone to sky-wave contamination. *See* AIR NAVIGATION; ELECTRONIC NAVIGATION SYSTEMS.

Patrick Hwang; Richard L. Greenspan

Bibliography. M. Kayton and W. R. Fried (eds.), *Avionics Navigation Systems*, 2d ed., 1997; G. J. Sonnenberg, *Radar and Electronic Navigation*, 6th ed., 1988.

Directional coupler

A four-port waveguide device (**Fig. 1**) in which an incoming wave at any one port (for example, A) appears at two others (C and D) but not at the fourth (B). This device finds numerous applications in waveguide networks such as microwave waveguides, integrated optics, and optical fibers. It is the equivalent of the hybrid induction coil used in conventional wire circuit telephony to provide side-tone balance. *See* TELEPHONE.

In Fig. 1, P_A is the input power, and P_B, P_C, and P_D are the powers of the output waves at ports B, C, and D respectively. Ideally, $P_B = 0$, and there is

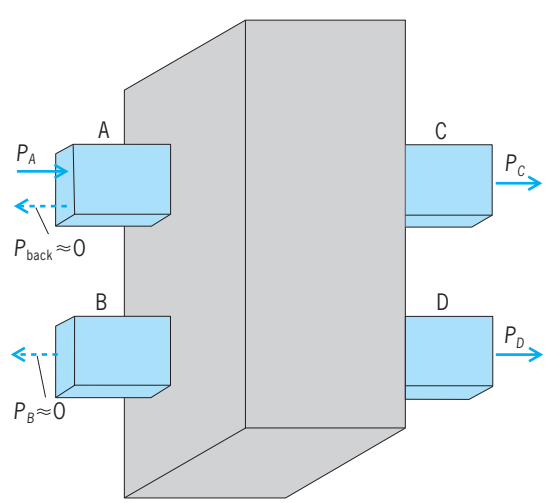

Fig. 1. Four-port directional waveguide coupler.

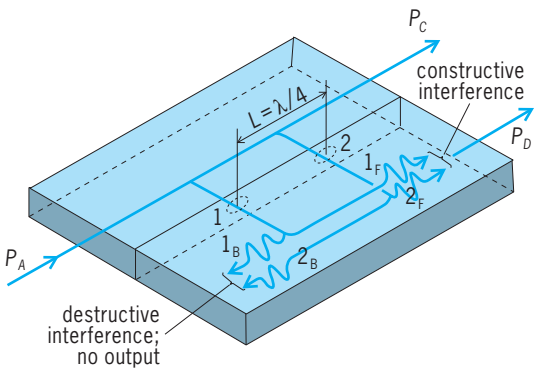

Fig. 2. Two-hole directional coupler.

no spurious reflection from the input port: $P_{\text{back}} = 0$. As a system component, the performance of a directional coupler is quoted in terms of (1) its coupling factor, $10 \log_{10} (P_D/P_A)$, in decibels (dB); (2) directivity, $10 \log_{10} (P_D/P_B)$, in dB; (3) insertion loss, $10 \log_{10} [(P_C + P_D)/P_A]$; (4) frequency band of operation; and (5) its matching impedance as measured by any spurious backward power flow from input port A. *See* ATTENUATION (ELECTRICITY); DECIBEL; ELECTRICAL IMPEDANCE; IMPEDANCE MATCHING.

Two-hole coupler. Directional couplers are usually made by introducing some form of lumped or continuous coupling among several guides so that the coupled waves interfere constructively in some directions while canceling one another in other directions. Consider two identical microwave waveguides running adjacent to each other with two small holes of the same size and shape coupling between them (**Fig. 2**). These two holes will divert a small portion of the incoming wave at port A from the left-hand waveguide into the right-hand waveguide, thereby establishing four propagating components. The two forward-propagating components 1_F and 2_F have traveled different paths but equal distances from port A to port D. They are, therefore, always in phase and will combine constructively to produce the output wave at port D. As for the backward-propagating components, 1_B leads 2_B by a phase of $4\pi L/\lambda$ radians, where L is the distance between the two holes and λ is the wavelength of the propagating radiation. Evidently, if $L = \lambda/4$, 1_B and 2_B are 180° out of phase, and since they have the same amplitudes (due to the small-hole assumption), they will cancel each other and send no power to port B. *See* INTERFERENCE OF WAVES.

Multipath and slot couplers. The relation $L = \lambda/4$ and the smallness of the holes impose two limitations on the performance of the simple two-hole directional coupler: a narrow-frequency band over which the directivity is high, and a small coupling factor which is on the order of 10–30 dB; that is, $(P_D/P_A) = 10^{-1}$ to 10^{-3}. To alleviate these problems, multipath couplers are used, where the backward cancellation is obtained through destructive interference among waves excited by three or more coupling elements (such as holes), usually spaced a quarter-wavelength apart. The coupling holes are not necessarily equal. Their coupling is frequently arranged to be in a binomial series (for example, five holes with couplings proportional to 1:4:6:4:1) or according to some other tapering law. A natural extension to the multipath coupler is the slot coupler, where the coupling element is a continuous slot that is several wavelengths long and often tapered, and the coupling is continuously distributed along the length of the interaction. *See* BINOMIAL THEOREM.

Couplers for open waveguides. Holes and slots are appropriate coupling elements for closed metallic waveguides, where the propagating field is completely enclosed within the walls of the guide. Other design alternatives exist for open waveguides such as microwave striplines, diffused dielectric optical guides used in integrated optics, and optical fibers. In these guides, the propagating field extends beyond the physical dimensions of the guiding structure (for example, the metallic plates of the stripline or the core of the fiber), and proximity coupling becomes possible.

Figure 3 shows a possible arrangement for a single-mode fiber-optic directional coupler. In a single- mode cylindrical fiber, optical power is carried in a very thin (5–10-micrometer diameter) core whose index of refraction is slightly higher than that of the surrounding cladding material. Similar to other dielectric waveguides, the electromagnetic field is not entirely confined within the core, and an exponentially decreasing, evanescent, nonradiating tail exists in the first few micrometers of the cladding. Fusion, etching, and mechanical polishing techniques may be used to bring two fiber cores close enough to each other to make possible the transfer of power from one core to the other through the interaction of the corresponding evanescent fields. The long interaction length (with respect to an optical wavelength) ensures an extremely high directivity (greater than 60 dB), and the gradual approach of the cores to each other virtually eliminates back reflections. As the separation between the cores decreases, a larger and larger percentage of the input power is coupled to port D, and a complete power transfer ($P_D = P_A$ and $P_C = 0$) is easily achieved. For even stronger coupling (smaller separation or longer interaction length), the power which was coupled to

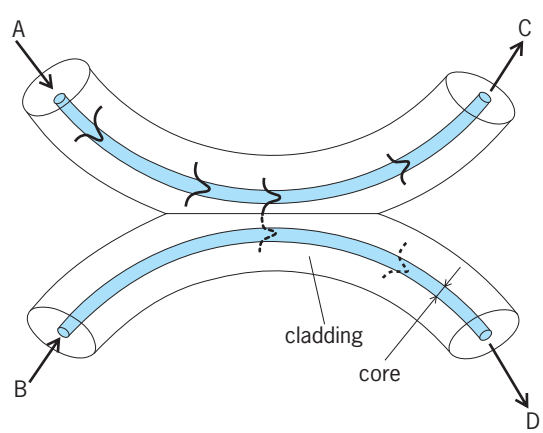

Fig. 3. Single-mode fiber-optic directional coupler.

the lower fiber may couple back to the upper fiber, a condition known as overcoupling.

Relation of output to input. Assume now that input waves are present at both port A and port B. Since directional couplers are linear devices, the output fields are linear combinations of the input fields. For simplicity, only the scalar case is considered where the input as well as the output waves have the same state of polarization. Thus, if E_A, E_B and E_C, E_D are, respectively, the amplitudes of the input and output fields, they are related to each other through Eqs. (1), or, in a matrix form, through Eq. (2). (The $\{M_i\}$ are, in general,complex numbers.)

$$\begin{aligned} E_C &= M_1 \cdot E_A + M_2 \cdot E_B \\ E_D &= M_3 \cdot E_A + M_4 \cdot E_B \end{aligned} \tag{1}$$

$$\begin{bmatrix} E_C \\ E_D \end{bmatrix} = \begin{bmatrix} M_1 & M_2 \\ M_3 & M_4 \end{bmatrix} \cdot \begin{bmatrix} E_A \\ E_B \end{bmatrix} \tag{2}$$

For a lossless coupler, the matrix M is unitary; that is, the inverse of M is also the complex conjugate of M. When the coupler is also symmetric, as in Figs. 2 and 3, M has a much simpler form, given by Eq. (3),

$$\begin{bmatrix} E_C \\ E_D \end{bmatrix} = \begin{bmatrix} \sqrt{1-K} & j\sqrt{K} \\ j\sqrt{K} & \sqrt{1-K} \end{bmatrix} \cdot \begin{bmatrix} E_A \\ E_B \end{bmatrix} \tag{3}$$

where K is the linear coupling factor and $j = \sqrt{-1}$. There is a phase difference of of 90° between the coupled portion of the input wave (M_3, E_A) and the uncoupled component (M_1, E_A). This property, which is also shared by bulk optic beam splitters, the waveguide magic tee, and the hybrid induction coil, is crucial for the understanding of the energy balance in various interferometers. *See* INTERFEROMETRY; MATRIX THEORY.

Applications. Directional couplers, whether fixed or variable, are extensively used in microwave and optical circuits. Weak couplers can selectively monitor either the forward or backward power flow in waveguides, and medium-to-strong couplers serve as amplitude combiners and splitters in various resonators and sensors. All of these applications require high directivity, low insertion loss, and negligible back reflections. Many other useful junctions are also based on the same physical principles. *See* INTEGRATED OPTICS; MICROWAVE MEASUREMENTS; MICROWAVE POWER MEASUREMENT; OPTICAL COMMUNICATIONS; OPTICAL FIBERS; WAVEGUIDE. Moshe Tur

Bibliography. S. Liao, *Microwave Devices and Circuits*, 3d ed., 1996; G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters: Impedance-Matching Networks, and Coupling Structures*, 1964, reprint 1980; A. Yariv, *Introduction to Optical Electronics*, 4th ed., 1991.

Directivity

The general property of directional discrimination displayed by systems that receive or emit waves. Thus, loudspeakers, microphones, radio antennas, underwater sound projectors, hydrophones, and even telescopes all have the common property that their effectiveness depends upon the direction from which the wave is either emitted or received. The manner in which a sender or receiver is directional depends largely upon its geometrical shape and, in particular, upon its dimensions compared to those of the wavelength involved (λ).

Directivity is a desirable property of a receiver because it permits the identification of the direction from which a signal comes and because noise from other directions is eliminated. It is desirable in a sender because the available energy can be concentrated in a given direction. A simple example of directivity is furnished by the megaphone, which effectively increases the size of the emitting area, thus increasing directivity.

Directivity is specified in terms of plane waves of a given wavelength in the presence of isotropic noise. The directional property of a projecting system is described quantitatively by a directivity factor D, which is the ratio of the intensity of the signal at a point on a designated axis to that at the same point produced by a nondirectional source radiating the same total power. The directivity index DI, given in decibels, is 10 log 10D. The directivity of some sources can be calculated, but complicated ones must be measured. Most transducers have the same directivity factor whether they are used as projectors or receivers. *See* DECIBEL.

In the case of receiver arrays, directivity arises because the various elements respond to the incoming wave in different phase. As an example, consider a receiver consisting of two elements, each small compared to λ (and therefore omnidirectional by themselves), which are separated by a distance d. A plane wave incident so that the wavefront is parallel to the axis joining the receivers has the same phase at both elements, regardless of the value of λ, and the response is always a maximum. The response for a wave coming along the axis depends on the ratio of d to λ. If $d = \lambda/2$, the response is zero because the two elements receive signals which are out of phase. The polar directivity pattern for this case is shown in **illus.** *a*.

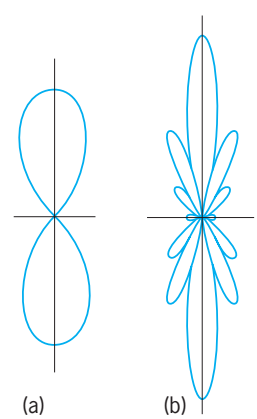

Polar directivity patterns (*a*) for a dipole with the spacing $\lambda/2$; (*b*) for linear array of seven elements spaced $\lambda/2$ apart.

If more elements are added, all spaced $\lambda/2$ apart, directivity is increased and side (or minor) lobes appear. Side lobes are directions of preferential response of lower sensitivity than that of the central direction. The multiple-element case is shown in illus. *b*. Advanced array design theory provides methods of reducing the side lobes still further by shading, or varying the amplitude and phase of the elements in a complicated way. Increased directivity can also be obtained by using acoustic lenses or parabolic reflectors. Arrays can normally be steered, either by physically rotating the array, or by introducing an appropriate phasing network for each element and steering the unit electrically.

The cases just described are samples of linear arrays in which directivity is obtained only with respect to one angular direction. Directivity in both angles (that is, where a beamlike pattern is formed) is obtained by using a receiver having extension in two dimensions. A common example of this is a plane circular piston. If the ratio of piston diameter to λ is increased, the main response is narrowed and the side lobes are reduced. For a discussion of the directivity of electromagnetic transducers *see* ANTENNA (ELECTROMAGNETISM); PARAMETRIC ARRAYS; SOUND; UNDERWATER SOUND.

William S. Cramer

Bibliography. M. Gayford (ed.), *Microphone Engineering Handbook*, 1994; R. C. Johnson, *Antenna Engineering Handbook*, 3d ed., 1993; R. J. Urick, *Principles of Underwater Sound*, 3d ed., 1983, reprint, 1996.

Disasteroida

A paraphyletic grouping of extinct irregular echinoids, members of which gave rise to the Holasteroida and Spatangoida in the Lower Cretaceous. Disasteroids have a split apical disc; the posterior two ocular plates and their associated ambulacra are separated from the remainder of the apical disc (and the anterior three ambulacra) by intercalated interambulacral plates. The apical disc became split in the early evolution of this group as a consequence of periproct migration out of the apical disc. Disasteroids have a strong bilateral symmetry. The peristome, which is small and oval and lacks buccal notches, lies toward the anterior of the lower surface. Plastron plating is usually protosternous but in late members can be meridosternous. Tuberculation is fine, dense, and uniform, without regional differentiation. Larger pore pairs, arranged in phyllodes around the mouth, suggest that disasteroids were deposit feeders with penicillate feeding tube feet. Most species probably lived infaunally, though not burrowing deeply.

Five families are included: Acrolusiidae, Collyritidae, Disasteridae, Pygorhytidae, and Tithoniidae, separated principally on apical disc plating. Approximately 75 species have been named, arranged in 17 genera. The oldest species is late Bajocian (Middle Jurassic) and the youngest Albian (Lower Cretaceous). The group achieved its greatest diversity during the Upper Jurassic. *See* ECHINODERMATA.

Andrew B. Smith

Discriminant

For a polynomial $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$, the discriminant is given by the expression

$$D = a_n^{2n-1}(x_1 - x_2)^2(x_1 - x_3)^2 \cdots (x_1 - x_n)^2(x_2 - x_3)^2 \cdots (x_2 - x_n)^2 \cdots (x_{n-1} - x_n)^2$$

where $x_1, x_2, \ldots, x_n$ are the roots of the equation $f(x) = 0$. There are $n(n-1)/2$ terms $(x_i - x_j)^2$ in the product corresponding to all possible selections of two indices i less than j from the numbers $1, 2, \ldots, n$.

The importance of the discriminant D lies in the fact that D vanishes if, and only if, the equation $f(x) = 0$ has equal roots. Since the value of the discriminant is unchanged if any two letters x_i and x_j are interchanged, it is a symmetric function of the roots, and can be expressed in terms of the coefficients of $f(x)$. Such an expression for D is most easily obtained, using the result $D = [(-1)^{n(n-1)/2}/a_n] \times R_x(f, f')$, where $f' = f'(x)$ is the derivative of $f(x)$ and $R_x(f, f')$ is the resultant of $f(x)$. For example, with $f(x) = a_2x^2 + a_1x + a_0$, and $f'(x) = 2a_2x + a_1$, notation (1) is obtained, and $D = a_1{}^2 - 4a_2a_0$, which

$$R_x(f, f') = \begin{vmatrix} a_2 & a_1 & a_0 \\ 2a_2 & a_1 & 0 \\ 0 & 2a_2 & a_1 \end{vmatrix} = -a_2a_1{}^2 + 4a_2{}^2a_0 \quad (1)$$

is the discriminant of a quadratic polynomial. The discriminant of a cubic polynomial $f(x) = a_3x^3 + a_2x^2 + a_1x + a_0$ is shown in notation (2).

$$D = 18a_3a_2a_1a_0 - 4a_2{}^3a_0 + a_2{}^2a_1{}^2 - 4a_3a_1{}^3 - 27a_3{}^2a_0{}^2 \quad (2)$$

See EQUATIONS, THEORY OF; POLYNOMIAL SYSTEMS OF EQUATIONS.

Ross A. Beaumont

Disease

A deleterious set of responses which occurs at the subcellular level, stimulated by some injury, and which is often manifested in altered structure or functioning of the affected organism. The concept of disease has been changing as understanding of biological systems has increased. For centuries disease was described in terms of gross changes. Then a major revolution in biology occurred between 1840 and 1860, based largely on the improvement and application of the light microscope. It included the works of Jacob Shleiden and Theodore Schwann, who developed the concept referred to as the cell theory. This theory states that the basic structural and functional unit of all plant and animal life is the cell. R. C. Virchow was the central figure who applied the cell theory to the pathogenesis of human

disease. He recognized that diseases were the product of the alteration of cell structure and function. *See* CELL (BIOLOGY).

A second revolution in biology started in the late 1930s. B. von Borris and E. Ruska in Germany developed the electron microscope, T. Svedberg constructed the ultracentrifuge, and L. Pauling worked on x-ray crystallography. G. W. Beadle and E. Tatum established the gene-enzyme relationship, and J. D. Watson and F. H. Crick described the three-dimensional structure of deoxyribonucleic acid (DNA). With advances in understanding and the development of sensitive probes, it became clear that the fundamental causes of diseases are based on biochemical and biophysical responses within the cell. These responses are still being categorized and, slowly, the mechanisms are being understood. *See* DEOXYRIBONUCLEIC ACID (DNA); ELECTRON MICROSCOPE; ULTRACENTRIFUGE; X-RAY CRYSTALLOGRAPHY.

Homeostasis. Cells are in a constant state of adjustment in response to fluctuating demands and stresses in order to maintain homeostasis. Homeostasis refers to functional equilibrium in an organism and to the processes that maintain it. There is a range of responses that is considered normal. If cells are pushed to respond beyond these limits, there may be an increase, a decrease, or a loss of normal structure or function. These changes may be reversible or irreversible. If irreversible, the cells may die. Thus, subcellular changes may be reflected in altered tissues, organs, and consequently organisms, and results in a condition described as diseased. *See* HOMEOSTASIS.

Manifestations. Lesions are the chemical and structural manifestations of disease. Some may be seen with the naked eye, while others may require microscopic or biochemical examination in order to be recognized. Subjective manifestations of a disease process such as weakness, pain, and fatigue are called symptoms. The objective measurable manifestations such as temperature, blood pressure, and respiratory rate changes are called signs or physical findings. Changes in the chemical or cellular makeup of an organ, tissue, or fluid of the body or its excretory products are called laboratory findings. To make a diagnosis is to determine the nature of the pathologic process by synthesizing information from these sources evaluated in the light of the patient's history and compared with known patterns of signs and symptoms. In common usage, the term disease indicates a constellation of specific signs and symptoms attributable to altered reactions in the individual which are produced by agents that affect the body or its parts. *See* CLINICAL PATHOLOGY.

Etiology. Etiology is the study of the cause or causes of a disease process. Although a disease may have one principal etiologic agent, it is becoming increasingly apparent that there are several factors involved in the initiation of a disease process. Susceptibility of the individual is an ever-present variable.

Endogenous. The etiologic factors can conveniently be divided into two categories. One group consists of endogenous (internal; within the body) factors, and may originate from errors in the genetic material. An example of a disease with an endogenous etiologic factor is sickle-cell anemia. This inherited disorder, occurring almost exclusively in blacks, results in abnormal hemoglobin. There is a mutation, a change in the normal structure of DNA, such that there is a substitution of valine molecule for glutamic acid in one of the polypeptide chains which makes up hemoglobin. It is thought that this results in altered configuration of the hemoglobin molecules, which is reflected in the altered morphology of the red blood cells at low O_2 tension. This change in morphology of red blood cells is referred to as sickling, and when cells assume this shape they tend to have mechanical difficulty in moving through the smaller blood vessels. They also tend to be quite fragile. Sickling may result in thrombosis, anoxia, and anemia. *See* HEMOGLOBIN; HUMAN GENETICS; MUTATION; SICKLE CELL DISEASE.

Exogenous. The other category of etiologic factors is exogenous (environmental). These account for the majority of disease reactions. Exogenous factors include physical, chemical, and biotic agents. The accompanying **table** lists various causative agents and examples of the diseases they produce.

Pathogenesis. Pathogenesis refers to the mechanisms by which the cell, and consequently the body, responds to an etiologic agent. It involves biochemical and physiological responses which are reflected in ultrastructural, microscopic, or gross anatomic lesions. There are a limited number of ways in which cells respond to injury. Thus, lesions produced by several agents may be morphologically similar. The agents may stimulate similar pathogenetic mechanisms. The cellular responses include adaptation, death, disturbances in growth and specialization, mutation, and neoplasia. The nature of the response is modified by the nature of the agent, dose, portal of entry, and duration of exposure, as well as many host factors such as age, sex, nutritional state, and species and individual susceptibility.

Environment. The diseases which are important in causing human death have changed since 1900. At that time, six of the ten leading causes of death in the United States were infectious (biotic) agents. These included tuberculosis, pneumonia, diarrhea, and enteritis and nephritis. During the early phase of the industrial revolution, chemical and physical agents were rarely the cause of disease and were usually limited to a small group of industrial workers. Therefore, control of infectious diseases was the major health care problem. Developments in microbiology fostered better sanitation and public health measures for the control of transmissible organisms. The development of antibiotics and immunization procedures further enabled the control of these agents. At present, only one of the ten leading causes of death in the United States, influenza and pneumonia, is due to biotic agents. *See* ANTIBIOTIC; INDUSTRIAL HEALTH AND SAFETY; INFECTIOUS DISEASE; VACCINATION.

While most of the biotic causes of diseases were being brought under control, continued population

Common exogenous and endogenous causes of disease

Causative agent	Disease
Exogenous factors	
Physical	
Mechanical injury	Abrasion, laceration, fracture
Nonionizing energy	Thermal burns, electric shock, frostbite, sunburn
Ionizing radiation	Radiation syndrome
Chemical	
Metallic poisons	Intoxication from methanol, ethanol, glycol
Nonmetallic inorganic poisons	Intoxication from phosphorus, borate, nitrogen dioxide
Alcohols	Intoxication from methanol, ethanol, glycol
Asphyxiants	Intoxication from carbon monoxide, cyanide
Corrosives	Burns from acids, alkalies, phenols
Pesticides	Poisoning
Medicinals	Barbiturism, salicylism
Warfare agents	Burns from phosgene, mustard gas
Hydrocarbons (some)	Cancer
Nutritional deficiency	
Metals (iron, copper, zinc)	Some anemias
Nonmetals (iodine, fluorine)	Goiter, dental caries
Protein	Kwashiorkor
Vitamins:	
A	Epithelial metaplasia
D	Rickets, osteomalacia
K	Hemorrhage
Thiamine	Beriberi
Niacin	Pellagra
Folic acid	Macrocytic anemia
B_{12}	Pernicious anemia
Ascorbic acid	Scurvy
Biological	
Plants (mushroom, fava beans, marijuana, poison ivy, tobacco opium)	Contact dermatitis, systemic toxins, cancer, hemorrhage
Bacteria	Abscess, scarlet fever, pneumonia, meningitis, typhoid, gonorrhea, food poisoning, cholera, whooping cough, undulant fever, plague, tuberculosis, leprosy, diphtheria, gas gangrene, botulism, anthrax
Spirochetes	Syphilis, yaws, relapsing fever, rat bite fever
Virus	Warts, measles, German measles, smallpox, chickenpox, herpes, roseola, influenza, psittacosis, mumps, viral hepatitis, poliomyelitis, rabies, encephalitis, trachoma
Rickettsia	Spotted fever, typhus
Fungus	Ringworm, thrush, actinomycosis, histoplasmosis, coccidiomycosis
Parasites (animal)	
Protozoa	Amebic dysentery, malaria, toxoplasmosis, trichomonas vaginitis
Helminths (worms)	Hookworm, trichinosis, tapeworm, filariasis, ascariasis
Endogenous factors	
Hereditary	Phenylketonuria, alcaptonuria, glycogen storage disease, Down syndrome (trisomy 21), Turner's syndrome, Klinefelter's syndrome, diabetes, familial polyposis
Hypersensitivity	Asthma, serum sickness, eczema, drug idiosyncrasy

growth (in large part, a consequence of the control of infectious disease) and the remarkable growth of industrialization have been associated with an increased prevalence of diseases caused by physical and chemical agents. Eight of the leading causes of death in the United States have environmental agents as the known or strongly suspected major etiologic factor. These include cancer, cirrhosis, and cardiovascular disease. *See* CANCER (MEDICINE); CIRCULATION DISORDERS; CIRRHOSIS.

General principles of the organism's response to toxic substances, some of which occur naturally in the environment, have evolved from a great number of investigations of agent-host interaction. These principles are: (1) All substances entering the organism are toxic; none is harmless. Dose rate of exposure and route of entry into the body determine whether a toxic response will occur or not. (2) All agents evoke multiple responses. (3) Most of the biological responses are undesirable, leading to the development of pathological changes. (4) A given dose of an agent does not produce the same degree of response in all individuals. Thus, when disease is viewed as interaction between the environment and the individual, the control of disease is largely the management of the environmental causes of disease. *See* CLINICAL MICROBIOLOGY; DEATH; ENVIRONMENTAL TOXICOLOGY; MEDICAL BACTERIOLOGY; PATHOLOGY; TOXICOLOGY; TUMOR.

N. Karle Mottet; Carol Quaife

Geochemical aspects. It is difficult to draw a definite link between geochemical environment and disease except in the few cases where high levels of toxic substances of geochemical origin have contacted organisms. Thus the effects of long-term, low-level exposure are not clearly understood. Only since the 1960s has any systematic and concerted research been carried out to explore the relationship between

geochemical environment and disease. A closely related problem is occupational health as affected by geochemical substances and environmental pollution. *See* INDUSTRIAL HEALTH AND SAFETY.

In 1831 a link was established between the severe disabilities (called miner's disease) shown by miners and the lead metal ores being mined. The Industrial Revolution, and particularly the large demand for metals generated by World Wars I and II, focused great attention on geochemical toxic hazards. The subsequent upsurge in the use of rarer elements in metallurgical and chemical industries has produced a serious health hazard. Very little is known about their environmental and clinical properties.

Of particular interest historically is the effect of geochemical dusts on the lungs of humans. The mining industry was plagued for many years by diseases related to the exposure of miners to high dust levels. Angular dust particles irritate the lining of the lungs and bronchial tubes and, if exposure is high and continuous, destroy the lung tissues, resulting in loss of respiration function. The most common disease of this type is silicosis. The problems caused by siliceous dusts can be largely minimized by using breathing masks, which filter out the offending particles. *See* RESPIRATOR.

Asbestos dusts contain sharp needlelike particles which can easily penetrate lung tissue. Some evidence suggests an abnormally high incidence of lung cancer among workers in the asbestos industry. *See* ASBESTOS.

The toxicity of a metal depends on its physical, chemical, and biochemical properties, the nature of its compounds, its mode of entry into the body, and the individual's susceptibility. The toxic properties of an element are related to its chemical form. For example, swallowing a small drop of metallic mercury could have minimal toxic effect, but a similar quantity of an organomercury compound, such as dimethyl mercury, could be fatal. Also, in the case of chromium, the trivalent form is an essential micronutrient, yet the hexavalent form is carcinogenic. The mode of intake of an element or compound will determine the effectiveness of its assimilation. The factor, together with the rate of excretion, will determine the severity and duration of the harmful effects. Generally, inhalation is a very effective way for ingesting soluble compounds, because these materials transfer directly through the pulmonary tissue into the bloodstream. Insoluble material can also be transported by the cilia of epithelial cells through the pharynx into the intestinal tract. Material injested orally appears in the stomach and is attacked by the digestive acids and enzymes. If it is solubilized in this manner, it can be transported through intestinal cells into body fluids and tissue. Absorption through the skin due to prolonged contact can also result in ingestion into the body. Generally, this is not a very efficient method of assimilation.

The effectiveness of assimilation, and indeed the degree of sensitivity to a toxic material, is dependent on a number of physiological and other factors which may differ from individual to individual. These factors include condition of health, age, body mass, sex, level of stress, and nutritional state. J. C. Van Loon

Disease ecology

The interaction of the behavior and ecology of hosts with the biology of pathogens as relating to the impact of diseases on populations. Although human diseases have been studied from an ecological perspective for centuries, diseases in plant and animal populations have only recently received much attention. The reason that ecologists have generally overlooked diseases is that diseases tend to be "invisible" compared to predators or competitors. The toll that diseases can take on all organisms (even bacteria are attacked by diseases) is enormous, and the ecology of diseases is one of the most vigorous areas of research in modern biology. *See* DISEASE; ECOLOGY.

Threshold theorem. For a disease to spread, on average it must be successfully transmitted to a new host before its current host dies or recovers. This simple observation lies at the core of the most important idea in epidemiology: the threshold theorem. The threshold theorem states that if the density of susceptible hosts is below some critical value, then on average the transmission of a disease will not occur rapidly enough to cause the number of infected individuals to increase. In other words, the reproductive rate of a disease must be greater than 1 for there to be an epidemic, with the reproductive rate being defined as the average number of new infections created per infected individual.

Human immunization programs are based on applying the threshold theorem of epidemiology to public health; specifically, if enough individuals in a population can be vaccinated, then the density of susceptible individuals will be sufficiently lowered that epidemics are prevented. Diseases such as smallpox have been eradicated because vaccination rates were high enough worldwide (70–80% immunization rates) to preclude the possibility of a smallpox outbreak. More generally, because diseases spread rapidly when densities of hosts are high, and less effectively (or not at all) when host densities are low, diseases can profoundly influence the population dynamics of their hosts—either by causing fluctuations in host density or by regulating host densities to low levels in conjunction with other factors. Even human populations seem to have been dramatically influenced by diseases. For example, between the fifth and eleventh centuries, European populations were ravaged by cycles of the plague. The initial plague epidemics killed 25–40% of the urban populations in Europe. Later cycles of plague were less deadly because the populations had evolved resistance and individuals had acquired immunity. *See* EPIDEMIOLOGY; VACCINATION.

Diseases have profoundly altered the course of human history. For example, the conquest of the Aztecs by a few hundred Spanish soldiers was due primarily to the fact that the Spaniards brought to the Western Hemisphere new diseases that decimated

the Aztec populations. In particular, the population of Mexico fell from about 20 million to 3 million between 1518 and 1568. The deaths were from successive epidemics of smallpox, measles, and typhus—diseases that had never before been seen in the New World, but to which the Europeans had built up a resistance (after centuries of exposure). Even today, new diseases can have a major impact on humans. The HIV epidemic in Africa is depleting the population of healthy workers so severely in cities that the economies of many nations are hampered. Almost 20 million people are infected with HIV in sub-Sahara Africa, with many of them being babies.

In general, the rate of reproduction for diseases is proportional to their transmissibility and to the length of time that an individual is infectious. For this reason, extremely deadly diseases that kill their hosts too rapidly may require extremely high densities of hosts before they can spread. All diseases do not behave as simply as hypothesized by the threshold theorem, the most notable exceptions being sexually transmitted diseases. Because organisms actively seek reproduction, the rate at which a sexually transmitted disease is passed among hosts is generally much less dependent on host density. For this reason, sexually transmitted diseases are especially hard to control.

Population effects. Cycles in many animal populations are thought to be driven by diseases. For example, the fluctuations of larch bud moths in Europe are hypothesized to be driven by a virus that infects and kills the caterpillars of this species. Cycles of red grouse in northern England are also thought to be driven by disease, in this case by parasitic nematodes. The interaction between the red grouse and parasitic worms is subtle; the presence of a worm infection does not necessarily mean that a grouse is harmed. Instead, it is only when grouse are laden with heavy worm burdens that effects are seen, and those effects take the form of reduced breeding success or higher mortality during the winter. This example highlights a common feature of diseases: their effects may be obvious only when their hosts are assaulted by other stresses as well (such as harsh winters and starvation). The most severe human plagues in the Middle Ages were associated with times of famine, presumably because malnourished victims were less able to withstand the plague.

Animals. The introduction of novel diseases to wild populations has created massive disruptions of natural ecosystems. For example, the introduction of rinderpest virus into African buffalo and wildebeest populations decimated them in the Serengeti. African wild ungulates have recovered in recent years only because a massive vaccination program eliminated rinderpest from the primary reservoir for the disease, domestic cattle. But the consequences of the rinderpest epidemic among wild ungulates extended well beyond the ungulate populations. For example, human sleeping sickness increased following the rinderpest epidemic because the tsetse flies that transmit sleeping sickness suffered a shortage of game animals (the normal hosts for tsetse flies) and increasingly switched to humans to obtain meals.

Plants. It is widely appreciated that crop plants are attacked by a tremendous diversity of diseases, some of which may ruin an entire year's production. The emigration of Irish people to the United States was prompted by a famine caused by a potato disease. A corn leaf blight destroyed approximately one-third of the United States corn crop in 1970 at a cost of almost $1 billion. Diseases are equally prevalent among wild populations of plants, but their toll seems to be reduced because natural plant populations are so genetically variable that it is unlikely that any given pathogen strain can sweep through and kill all of the plants—there are always some resistant genotypes. But when agronomists have bred plants for uniformity, they have often depleted genetic diversity and created a situation in which a plant pathogen that evolves to attack the crop encounters plants with no resistance (all the plants are the same) and sweeps through the fields with drastic consequences. For example, when leaf blight devastated the United States corn crop, 70% of that crop shared genetically identical cytoplasm, and the genetic uniformity of the host exacerbated the severity of the epidemic. In fact, hundreds of agricultural experiments have been performed in which the severity of plant diseases is contrasted between fields planted with a mixture of varieties and fields planted with only one variety. The results reveal that, on average, genetic variety reduces disease incidence by 50% compared to genetically homogeneous (only one variety) plantings. This is one reason why biodiversity and genetic diversity are viewed as a natural resource; such diversity is likely to harbor resistance traits, which can be bred (or transferred by genetic engineering) into food crops. *See* PLANT PATHOLOGY.

Disease emergence. Humans are dramatically altering habitats and ecosystems. Sometimes these changes can influence disease interactions in surprising ways. Lyme disease in the eastern United States provides a good example of the interplay of human habitat modifications and diseases. Lyme disease involves a spirochete bacterium transmitted to humans by ticks. However, humans are not the normal hosts for this disease; instead, both the ticks and the bacterium are maintained primarily on deer and mouse populations. Human activities influence both deer and mice populations, and in turn tick populations, affecting potential exposure of humans to the disease. The pathways for this influence are numerous and hard to predict. But one well-documented human effect has arisen from the control of natural predators of deer and the clearing of forest habitat; both changes have favored the buildup of large deer populations. As deer populations increase, so do populations of deer ticks and the opportunities for transfer of Lyme disease to humans. Lyme disease could never be maintained in human populations, but excessive deer populations and proximity of deer to human outdoor activities have created a significant health hazard in many northeastern states.

Much less certain are the impacts of anticipated global warming (caused by the burning of fossil

fuels) on diseases. There is clearly some cause for concern about the expansion of tropical diseases into what are now temperate regions in those cases where temperature sets limits to the activity or distribution of major disease vectors (such as the species of mosquito that vector malaria).

The role of natural ecosystems in diseases, and the possibility that human activities can profoundly alter the ecology of diseases, are very significant. *See* LYME DISEASE; POPULATION ECOLOGY. Peter Kareiva

Bibliography. J. Burdon, *Diseases and Plant Population Biology*, Cambridge University Press, Cambridge, 1987; J. Diamond, *Guns, Germs and Steel*, W. W. Norton, New York, 1999; B.T. Grenfell and A. Dobson (eds.), *Ecology of Infectious Diseases in Natural Populations*, Cambridge University Press, Cambridge, 1995; W. H. McNeill, *Plagues and People*, Anchor Press/Doubleday, New York, 1976; R. S. Ostfeld, The ecology of lyme-disease risk, *Amer. Scientist*, 85:338–346, 1997.

Dispersion (radiation)

The separation, by refraction, interference, scattering, or diffraction, of acoustic and electromagnetic radiation or energy into its constituent wavelengths or frequencies. For a refracting, transparent substance, such as a prism of glass, the dispersion is characterized by the variation of refractive index with change in wavelength of the radiation. Refractive index (n) is defined as the ratio of the velocity of the radiation in free space (air at standard temperature and pressure for sound, and a vacuum for electromagnetic radiation) to the velocity in the substance in question. I. Newton used a small hole in a window shade and a glass prism (see **illus.**) to disperse sunlight into a visible spectrum, from violet through red. Using a second prism, he showed that no further decomposition of any of the spectral colors could be achieved. *See* OPTICAL PRISM; REFRACTION OF WAVES.

Acoustic radiation can be explained by pure mechanics, but a full understanding of the dispersion of electromagnetic radiation requires atomic theory. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; SOUND.

Formulas and behavior. The most widely used formula for dispersion is Sellmeier's equation (1), where

$$n^2 - 1 = \sum_{i=0}^{m} \frac{A_i \lambda^2}{\lambda^2 - \lambda_i^2} \tag{1}$$

n is the refractive index at wavelength λ, and A_i, and λ_i, are constants that approximate the asymptotic behavior, bounding but not valid (since n cannot actually go to infinity) in regions of high or total absorption. Sellmeier's equation can be expanded with the binomial theorem, and the dominant terms are those of the older, simpler, but less accurate Cauchy equation (2).

$$n = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4} \tag{2}$$

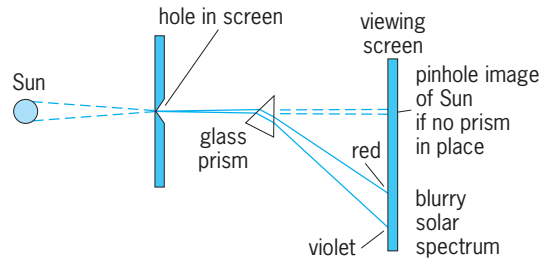

Prismatic dispersion of sunlight. (*After J. P. C. Southall, Mirrors, Prisms, and Lenses, Macmillan, 1923*)

Differentiation shows that blue light disperses more rapidly than red light, as the cube of wavelength. Different optical materials have widely varying dispersive power, which is defined by the Abbe number, given by Eq. (3), where d, F, and C refer

$$\nu_d = \frac{n_d - 1}{n_F - n_C} \tag{3}$$

to Fraunhofer absorption lines in the green, blue, and red parts of the solar spectrum. Different substances disperse wavelengths in slightly different ratios, an optical disadvantage termed irrationality of dispersion, which causes achromatic lenses to have small amounts of residual image color. *See* CHROMATIC ABERRATION; FRAUNHOFER LINES; OPTICAL MATERIALS.

The condition where the refractive index decreases as wavelength increases is termed normal dispersion. The opposite condition is termed anomalous dispersion, and almost always occurs in regions outside the range of visible wavelengths.

Gratings and prisms. Diffraction gratings are flat or curved plates, either reflective or transmissive, with finely ruled lines, that distribute light according to the grating equation (4), where $N =$ the order num-

$$N \cdot \lambda = b(\sin\theta_i + \sin\theta_d) \tag{4}$$

ber, λ is the wavelength, b is the separation between lines ruled onto a plate, and θ_i and θ_d are the angles of incidence and diffraction, relative to the grating normal. Differentiation of θ_d versus λ shows that angular dispersion is almost linear with wavelength. Reflective gratings are usable in virtually any part of the spectrum.

Diffraction-based dispersers are normally preferred for measuring the content of the radiation; prism dispersers are generally selected to study material properties of the prism or its contents. Occasionally, gratings and prisms are combined, as in an echelle spectrograph. *See* ASTRONOMICAL SPECTROSCOPY; DIFFRACTION GRATING; SPECTROGRAPH; SPECTROSCOPY. Richard A. Buchroeder

Bibliography. M. Born and E. Wolf, *Principles of Optics*, 6th ed., 1980, reprint 1987; M. Minnaert, *Light and Colour of the Outdoors*, rev. ed., 1993; S. F. Ray, *Applied Photographic Optics*, 2d ed., 1996; B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, 1991.

Dispersion relations

Relations between the real and imaginary parts of a response function. The term dispersion refers to the fact that the index of refraction of a medium is a function of frequency. In 1926 H. A. Kramers and R. Kronig showed that the imaginary part of an index of refraction (that is, the absorptivity) determines the real part (that is, the refractivity); this is called the Kramers-Kronig relation. The term dispersion relation is now used for the analogous relation between the real and imaginary parts of any response function, such as Eq. (14) below.

Response function. Consider a system in which a cause $C(t)$ [for example, a force] and its effect $E(t)$ [for example, a displacement] are related by Eq. (1),

$$E(t) = \int_{-\infty}^{t} dt'\ G(t-t')C(t') \tag{1}$$

where $G(t)$ is called the response function. Because this relation is linear in both $C(t)$ and $E(t)$, the response is said to be linear: The superposition of two causes results in the sum of their effects. Because the response function G is a function only of $t - t'$, the response is said to be time-independent: The effect at time t of a cause at time t' depends only on the time difference $t - t'$. Because the upper limit of integration is t [or equivalently, $G(t)$ can be said to vanish for negative argument, Eq. (2)], the response

$$G(t) = 0 \qquad \text{if } t < 0 \tag{2}$$

is said to be causal: the cause has no effect at earlier times. *See* CAUSALITY; LINEARITY.

Many examples can be given. A few pairs of cause and effect are: force and spatial displacement, electric field and polarization (G is electrical susceptibility), and incident wave and scattered wave (G is scattering amplitude).

Pure-tone response. The Fourier transform $g(\omega)$ of the response function $G(t)$ is defined in Eq. (3) and the inverse relation in Eq. (4). The function $g(\omega)$

$$g(\omega) = \int_{0}^{\infty} dt\ e^{i\omega t} G(t) \tag{3}$$

$$G(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega\ e^{-i\omega t} g(\omega) \tag{4}$$

is complex-valued; however, if $G(t)$ is real [which it must be if $C(t)$ and $E(t)$ of Eq. (1) are real], the crossing relation, Eq. (5), is valid, where an asterisk

$$g^*(\omega) = \int_{0}^{\infty} dt\ e^{-i\omega^* t} G(t) = g(-\omega^*) \tag{5}$$

indicates complex conjugation. The function $g(\omega)$ is called the frequency-dependent response; it describes the response to a simple harmonic (pure tone) cause: if the cause is given by Eq. (6), then

$$C(t) = \cos(\omega t - \eta),\ \omega \text{ real} \tag{6}$$

Eq. (1) implies that its effect is given by Eq. (7), where c.c. means complex conjugate and $|g|$ and δ are given by Eq. (8). One sees that the magnitude $|g|$

$$\begin{aligned} E(t) &= \int_{-\infty}^{\infty} dt'\ G(t-t')\cos(\omega t' - \eta) \\ &= {}^1\!/_2 \int_{-\infty}^{\infty} dt'\ G(t-t') e^{-i(\omega t' - \eta)} + \text{c.c.} \\ &= {}^1\!/_2 e^{-i(\omega t - \eta)} \int_{-\infty}^{\infty} dt'\ G(t-t') e^{i\omega(t-t')} + \text{c.c.} \\ &= {}^1\!/_2 e^{-i(\omega t - \eta)} g(\omega) + \text{c.c.} \\ &= {}^1\!/_2 \left|g(\omega)\right| e^{-i(\omega t - \eta - \delta(\omega))} + \text{c.c.} \\ &= \left|g(\omega)\right| \cos[\omega t - \eta - \delta(\omega)] \end{aligned} \tag{7}$$

$$g(\omega) = \left|g(\omega)\right| e^{i\delta(\omega)} \tag{8}$$

and phase δ of $g(\omega)$ are the amplitude and phase shift, respectively, of the simple harmonic effect relative to the cause. *See* FOURIER SERIES AND TRANSFORMS; INTEGRAL TRANSFORM.

Causality. Causality, Eq. (2), has an important consequence for $g(\omega)$, namely the dispersion relations given in Eq. (14) below. Their derivation will be outlined, using Cauchy's formula and other properties of analytic functions of a complex variable. Let ω take on complex values, $\omega = \text{Re}\ \omega + i\ \text{Im}\ \omega$. Then $g(\omega)$ as given by Eq. (3) is an analytic function of ω as long as the integral, and its derivative with respect to ω, exist (that is, are finite). For any stable system, $G(t)$ is bounded in magnitude for all values of t. It follows that the integral in Eq. (3) exists for all ω such that $\text{Im}\ \omega > 0$ ("in the upper half ω-plane") because then the magnitude of the factor $e^{i\omega t}$ of the integrand falls exponentially with increasing t according to Eq. (9).

$$\left|e^{i\omega t}\right| = e^{-(\text{Im}\ \omega)t} \tag{9}$$

Since, on the other hand, for decreasing t this factor rises exponentially, the existence of the integral depends on causality, Eq. (2). So, causality implies that $g(\omega)$ is an analytic function in the upper half ω-plane. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

Consequently one may write Cauchy's formula for $g(\omega)$, Eq. (10), where the integration contour C is

$$g(\omega) = \frac{1}{2\pi i} \int_{C} d\omega'\ \frac{g(\omega')}{\omega' - \omega} \tag{10}$$

any simple loop, traversed counterclockwise, which surrounds the point ω and which lies in the upper half ω'-plane [the region where $g(\omega)$ is an analytic function]. Letting the point ω approach the contour C, the Cauchy formula comes to the form Eq. (11),

$$g(\omega) = \frac{1}{2\pi i} \left[\text{P} \int_{C} \frac{d\omega'}{\omega' - \omega} g(\omega') + i\pi g(\omega) \right] \tag{11}$$

where P means principal value. [Roughly speaking, it means: omit from the range of integration the interval of C which lies within a distance ϵ on either side of the singular point $\omega' = \omega$, and then take the limit

$\epsilon \to 0$.] Solving for $g(\omega)$ gives Eq. (12), where HT

$$g(\omega) = \frac{-i}{\pi} \mathrm{P} \int_C \frac{d\omega'}{\omega' - \omega} g(\omega) \equiv -i\mathrm{HT}[g] \qquad (12)$$

means Hilbert transform (using the contour C). Now let the contour C be expanded until it encloses the entire upper half ω' plane. If $g(\omega)$ vanishes at large ω, the only part of the contour which contributes to the Hilbert transform integral runs along the real ω' axis. Taking also ω to be real, the real and imaginary parts of Eq. (12) give Eqs. (13), where Re g and

$$\mathrm{Re}\, g = \mathrm{HT}[\mathrm{Im}\, g] \qquad \mathrm{Im}\, g = -\mathrm{HT}[\mathrm{Re}\, g] \qquad (13)$$

Im g are the real and imaginary parts of g. Finally, using crossing, Eq. (5), the integral can be expressed as an integral over just positive ω', Eqs. (14).

$$\begin{aligned} \mathrm{Re}\, g(\omega) &= \frac{2}{\pi} \mathrm{P} \int_0^\infty d\omega' \, \frac{\omega' \, \mathrm{Im}\, g(\omega')}{\omega'^2 - \omega^2} \\ \mathrm{Im}\, g(\omega) &= \frac{-2\omega}{\pi} \mathrm{P} \int_0^\infty d\omega' \, \frac{\mathrm{Re}\, g(\omega')}{\omega'^2 - \omega^2} \end{aligned} \qquad (14)$$

These are the dispersion relations for the frequency-dependent response function $g(\omega)$. They say that Re g and Im g are not independent functions; given one of them, the other is determined.

Resonance. Although a causal $g(\omega)$ is analytic in the upper ω plane, it can, and usually does, have singularities in the lower ω-plane. For instance, $g(\omega)$ may have a pole at $\omega = \omega_0$ [and to satisfy Eq. (5) likewise at $\omega = -\omega_0{}^*$]; causality requires that Im $\omega_0 < 0$. In Eq. (15) such a pole term of $g(\omega)$ is

$$\begin{aligned} g(w) &= \frac{A}{\omega - \omega_0} + \frac{A}{-\omega - \omega_0^*} + \cdots \\ &= \frac{\mathrm{Re}\, A + i\, \mathrm{Im}\, A}{\omega - \omega_R + i\, {}^1\!/_2\gamma} + \frac{\mathrm{Re}\, A - i\, \mathrm{Im}\, A}{-\omega - \omega_R - i\, {}^1\!/_2\gamma} + \cdots \end{aligned} \qquad (15)$$

written out, with the notation Re $\omega_0 = \omega_R$, Im $\omega_0 = -{}^1\!/_2\gamma$. If γ is small, then $|g(\omega)|^2$ as a function of real ω peaks sharply at $\omega = \omega_R$ with a width (defined as the interval of ω between the half height points, where $|g(\omega)|^2$ is half as big as the peak) equal to γ. This is called a resonance of the response function $g(\omega)$; ω_R is its position and γ is its width. *See* RESONANCE (ACOUSTICS AND MECHANICS); RESONANCE (ALTERNATING-CURRENT CIRCUITS).

The consequence of a resonance for the effect $E(t)$ is found by using Eq. (15) in Eqs. (1) and (4). If the cause $C(t)$ turns off rapidly enough, then at large t Eq. (16) is obtained, where the ω integration was

$$\begin{aligned} E(t) &= \int dt' \, C(t') \frac{1}{2\pi} \\ &\quad \cdot \int d\omega \, e^{-i\omega(t-t')} \frac{A}{\omega - \omega_0} + \text{c.c.} + \cdots \xrightarrow{t\to\infty} \\ &\quad -iA \int dt' \, C(t') e^{-i\omega_0(t-t')} + \text{c.c.} + \cdots \\ &= \left[-iAe^{-i\omega_R t} \int dt' \, C(t') e^{i\omega_0 t'} + \text{c.c.} \right] e^{\gamma t/2} + \cdots \\ &= C_1 \cos(\omega_R t + C_2) \, e^{-\gamma t/2} + \cdots \qquad (16) \end{aligned}$$

(where C_1, C_2 are constants)

done by residues. The result is that the effect $E(t)$ has an exponentially decaying term whose time constant (for E^2) is γ^{-1}; that is, the width of the resonance peak of $|g(\omega)|^2$ and the decay constant of $E(t)^2$ are reciprocals of one another. In more physical terms, the decaying oscillation of $E(t)$ reflects a (decaying) normal mode of oscillation, a "ringing" of the system.

If Im ω_0 were positive, contrary to causality, the calculation of Eqs. (16) would show the effect $E(t)$ to have an exponentially growing term at an early time, preceding the cause. As for $g(\omega)$, the qualitative effect of causality is that at the resonance, where the magnitude $|g(\omega)|$ has a narrow peak, the phase of $g(\omega)$ [$\delta(\omega)$ in Eq. (7)] rises rapidly; if Im ω_0 were positive, contrary to causality, the phase would instead fall rapidly.

Dielectric constant and refractive index. The response of a macroscopically homogeneous dielectric (polarizable) medium to an electromagnetic plane wave of frequency ω is described by the dielectric constants $\epsilon_\parallel(\omega)$ and $\epsilon_\perp(\omega)$. The subscripts distinguish whether the electric field is parallel or perpendicular to the direction of propagation of the plane wave. The discussion will be limited to $\epsilon_\perp(\omega)$, which will be referred to as simply $\epsilon(\omega)$. The refractive index $n(\omega)$ is simply related: $n^2 = \epsilon$. But $n(\omega)$ itself will not in general satisfy dispersion relations because if $\epsilon(\omega)$ has a zero in the upper ω plane, $\sqrt{\epsilon} = n$ will have a square root singularity. *See* PERMITTIVITY.

At $\omega = \infty$, $\epsilon(\omega)$ has the value 1 (the vacuum value). At $\omega = 0$, $\epsilon(\omega)$ is singular if the medium is conductive (that is, contains charges which are free to move), according to Eq. (17), where σ_0 is the static conductiv-

$$\epsilon(\omega) \xrightarrow{\omega\to 0} \frac{4\pi i\sigma_0}{\omega} - \frac{\omega_0^2}{\omega^2}, \quad \omega_0^2 = \frac{n_v 4\pi e^2}{m} \qquad (17)$$

ity and ω_0 is the plasma frequency; n_v is the number density of free charges of charge e and mass m. Thus the function given in Eq. (18) has the properties as-

$$g(\omega) = \epsilon(\omega) - 1 - \frac{4\pi i\sigma_0}{\omega} + \frac{\omega_0}{\omega^2} \qquad (18)$$

sumed for $g(\omega)$ above and can be substituted for it in Eq. (14), resulting in the dispersion relations, Eqs. (19). In Eqs. (19), terms have been added to the in-

$$\begin{aligned} \mathrm{Re}\, \epsilon(\omega) &= 1 - \frac{\omega_0^2}{\omega^2} + \frac{2}{\pi} \int_0^\infty d\omega' \\ &\quad \cdot \frac{\omega' \, \mathrm{Im}\, \epsilon(\omega') - \omega \, \mathrm{Im}\, \epsilon(\omega)}{\omega'^2 - \omega^2} \qquad (19a) \\ \mathrm{Im}\, \epsilon(\omega) &= \frac{4\pi\sigma_0}{\omega} - \frac{2}{\pi}\omega \int_0^\infty d\omega' \\ &\quad \cdot \frac{\mathrm{Re}\, \epsilon(\omega') + \omega_0^2/\omega'^2 - \mathrm{Re}\, \epsilon(\omega) - \omega_0^2/\omega^2}{\omega'^2 - \omega^2} \qquad (19b) \end{aligned}$$

tegrands in order to make them nonsingular; these terms do not alter the integrals because $\mathrm{P} \int d\omega'/(\omega'^2 - \omega^2) = 0$.

By setting ω equal to special values, interesting results are obtained; for simplicity these will be given for the case of a nonconductor, where $\sigma_0 = \omega_0 = 0$. Setting $\omega = 0$ in Eq. (19a) results in an expression

for the static dielectric constant, Eq. (20). which are

$$\epsilon(0) = 1 + \frac{2}{\pi}\int_0^\infty d\omega\, \frac{\operatorname{Im}\epsilon(\omega)}{\omega} \tag{20}$$

$$\operatorname{Re}\epsilon(\omega) \to 1 - \frac{\omega_\infty^2}{\omega^2} \tag{21a}$$

where $$\omega_\infty^2 = \frac{2}{\pi}\int_0^\infty d\omega\, \omega \operatorname{Im}\epsilon(\omega)$$

$$\operatorname{Im}\epsilon(\omega) \to \frac{4\pi\sigma_\infty}{\omega} \tag{21b}$$

where $$4\pi\sigma_\infty = \frac{2}{\pi}\int_0^\infty d\omega\, [\operatorname{Re}\epsilon(\omega) - 1]$$

sum rules. The first of these relates an integral over all frequencies of the absorptivity, Im ϵ, to the quantity ω_∞^2, for which a direct calculation gives the value $\omega_\infty^2 = \Sigma\, n_v 4\pi e^2/m$, where the sum is over all kinds of charges in the medium, in contrast to the formula for the plasma frequency ω_0, Eq. (17); at infinite frequency all particles of the medium behave as though free. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION.

Forward-scattering amplitude. It can be shown by arguments like those used above that the forward-scattering amplitude $f(\omega)$ of light incident on an arbitrary body satisfies the dispersion relation given in Eq. (22) [the less interesting relation giving Im f

$$\operatorname{Re} f(\omega) = f(0) + \frac{\omega^2}{2\pi^2 c}\int_0^\infty d\omega'\, \frac{\sigma_T(\omega') - \sigma_T(\omega)}{\omega'^2 - \omega^2} \tag{22}$$

where $$f(0) = \frac{-Q^2}{Mc^2}$$

in terms of Re f has been omitted], where Q is the charge, M is the mass of the body, and c is the speed of light. In Eq. (22) Im $f(\omega)$ has been expressed in terms of the total cross section σ_T by the optical theorem, Eq. (23); this formula is true for the scattering

$$\operatorname{Im} f(\omega) = \frac{k}{4\pi}\sigma_T(\omega) \tag{23}$$

where k = wave number

of waves of any sort.

Equations (22) and (14) can be related by using the "optical potential" formula, Eq. (24), which relates

$$\frac{k^2(\text{in medium})}{k^2(\text{in vacuum})} = n^2 - 1 + \frac{4\pi n_v f(\omega)}{k^2} \tag{24}$$

the forward scattering amplitude $f(\omega)$, of a wave (of any sort) of frequency ω and wave number k incident on a body, to the index of refraction n of that wave in a medium consisting of a random arrangement of such bodies with mean number density n_v. However, Eq. (24) is valid only for a dilute medium, whereas Eqs. (14) and (22) are both exact if causality is true.

An interesting consequence of Eq. (22) is that a system consisting of the electromagnetic field and point charges, obeying classical mechanics, cannot be causal. For such a system it can be shown that $f(\omega)$ vanishes for $\omega = \infty$; using this in Eq. (22) gives Eq. (25), which is untrue since both terms on the

$$0 = f(0) - (2\pi^2 c)^{-1}\int_0^\infty d\omega\sigma_T \tag{25}$$

right-hand side are negative. In the quantum theory, by contrast, $\sigma_T > 0(1/\omega)$ and so no formula like Eq. (25) follows from causality. *See* NONRELATIVISTIC QUANTUM THEORY.

Elementary particle theory. In the quantum theory the scattering amplitudes of matter waves satisfy dispersion relations similar to Eq. (22) if causality holds. This gives the possibility of experimentally testing causality up to very high frequencies (corresponding to very short time intervals) by using, for instance, protons of energy up to 1000 GeV, incident on protons in a target, at Fermilab. Since the right-hand side of Eq. (22) requires knowledge of σ_T at all ω, one can strictly say only that the observed $f(\omega)$ is consistent with causality and a reasonable extrapolation of σ_T beyond observed energies.

Not only is there no experimental evidence against causality, but causality has always been a property of relativistic quantum field theory (elementary particle theory). The following is essentially equivalent to causality: Considered as functions of the Lorentz invariant scalars formed from the 4-momenta of the interacting particles (for instance, the total energy and the momentum transfer in the center-of-mass frame, in the case of a two-body scattering amplitude), the S-matrix elements (scattering amplitudes) of quantum field theory are analytic functions. More precisely, they are analytic except for singularities on Landau surfaces in the space of the (complex) arguments, the Lorentz scalars. These singularities are branch points, and if the amplitude is analytically continued around one, the amplitude becomes changed by a certain amount.

On the one hand, knowing this change allows the amplitude itself to be determined, by use of Cauchy's formula [a dispersion relation like Eq. (22) is a simple example of this]. On the other hand, the change is given in terms of other S-matrix elements. [The optical theorem, Eq. (23), is a simple example of this; the point is that σ_T is a sum (and integral) over products of two amplitudes.] These S-matrix equations, resulting from the combination of the analyticity and the unitarity of the S-matrix, are the subject of S-matrix theory. *See* RELATIVITY.

These equations can to some extent replace the equations of motion of quantum field theory. They have the advantage over the only known systematic solution of the equations of motion, namely perturbation theory (Feynman diagrams), in that they involve only S-matrix elements, that is, the amplitudes of real, observable processes. However, there is no known way to find the solution of the S-matrix equations which corresponds to a given set of field theory equations of motion.

It was the original hope of many workers in S-matrix theory that the equations of motion were in fact irrelevant and that the S-matrix was determined

by a few properties of low-mass particles (in fact perhaps by none, in the extreme view termed bootstrapping). However, the successes of quantum chromodynamics (QCD) in describing and predicting many phenomena in high-energy physics seems to show that a specific field theory, whose elementary quanta (quarks and gluons) are not observable as particles, underlies the S-matrix of "elementary" particle interactions. The S-matrix equations are true in any case, and will remain an important tool in the description and correlation of experimental data. *See* ELEMENTARY PARTICLE; QUANTUM CHROMODYNAMICS; QUANTUM FIELD THEORY; SCATTERING EXPERIMENTS (NUCLEI); SCATTERING MATRIX. Charles J. Goebel

Bibliography. J. W. Brown and R. V. Churchill, *Complex Variables and Applications*, 6th ed., 1996; G. F. Chew, *The Analytic S Matrix*, 1966; R. J. Eden, *High Energy Collisions of Elementary Particles*, 1967; R. J. Eden et al., *The Analytic S Matrix*, 1966; R. Good and T. Nelson, *Classsical Theory of Electric and Magnetic Fields*, 1971; A. Martin and T. Spearman, *Elementary Particle Theory*, 1970; P. Roman, *Advanced Quantum Theory*, 1965.

Displacement current

The name given by J. C. Maxwell to the term $\partial\mathbf{D}/\partial t$ which must be added to the current density $\mathbf{i}$ to extend to time-varying fields A. M. Ampère's magnetostatic result that $\mathbf{i}$ equals the curl of the magnetic intensity $\mathbf{H}$. In integral form this result is given by Eq. (1), where the unit vector $\mathbf{n}$ is perpendicular to

$$\oint \mathbf{H}\cdot d\mathbf{s} = \int_S \left(\mathbf{i} + \frac{\partial \mathbf{D}}{\partial t}\right)\cdot \mathbf{n}\, dS \qquad (1)$$

the surface dS. The concept of displacement current has important consequences for insulators and for free space where $\mathbf{i}$ vanishes. For conductors, however, the difference between Eq. (1) and Ampère's result is negligible. *See* AMPÈRE'S LAW; MAXWELL'S EQUATIONS.

In order to show that the displacement term is essential, consider a parallel-plate capacitor charged by a circuit carrying an alternating current. Let a closed curve s encircle one of the charging wires and be the boundary of two surfaces, S_1 which passes through the capacitor gap and S_2 which cuts the charging wire. By Gauss' electric flux theorem, the charge Q on the plate and wire between S_1 and S_2 is given by Eq. (2), where the normal is taken in the direction

$$Q = \int_{S_1} \mathbf{D}\cdot\mathbf{n}\, dS - \int_{S_2} \mathbf{D}\cdot\mathbf{n}\, dS \qquad (2)$$

of current flow in both S_1 and S_2. The current $\mathbf{I}$ or $\int \mathbf{i}\cdot\mathbf{n}dS$ equals $\partial Q/\partial t$ so that Eq. (3) holds. Thus, the

$$\int_{S_1} \frac{\partial \mathbf{D}}{\partial t}\cdot\mathbf{n}\, dS = \int_{S_2}\left(\mathbf{i} + \frac{\partial \mathbf{D}}{\partial t}\right)\cdot\mathbf{n}\, dS$$

$$= \oint \mathbf{H}\cdot d\mathbf{s} \qquad (3)$$

inclusion of the displacement current is needed to make Eq. (1) valid for any surface S bounded by s.

If one defines current as a transport of charge, the term displacement current is a misnomer when applied to a vacuum where no charges exist. If, however, current is defined in terms of the magnetic fields it produces, the expression is legitimate. In a dielectric, where an electric field produces a displacement of the negative charges with respect to the positive ones, the name has meaning. Maxwell had this sort of picture, even for a vacuum, where he postulated a polarizable ether. *See* ETHER HYPOTHESIS. William R. Smythe

Displacement pump

A pump that develops its action through the alternate filling and emptying of an enclosed volume.

Reciprocating types. Positive-displacement reciprocating pumps have cylinders and plungers or pistons with an inlet valve, which opens the cylinder to the inlet pipe during the suction stroke, and an outlet valve, which opens to the discharge pipe during the discharge stroke. Reciprocating pumps may be power-driven through a crank and connecting rod or equivalent mechanism, or direct-acting, driven by steam or compressed air or gas.

Figure 1 shows a small high-speed plunger-type power pump for high-pressure service. The three-throw crankshaft is carried in roller bearings at each end. The manifolds below the suction valves and above the discharge valves connect the three pumping cylinders to the suction and discharge piping.

Power pumps are frequently built with one or two throw cranks and double-acting liquid ends, or with five, seven, or even nine cranks where smoother flow

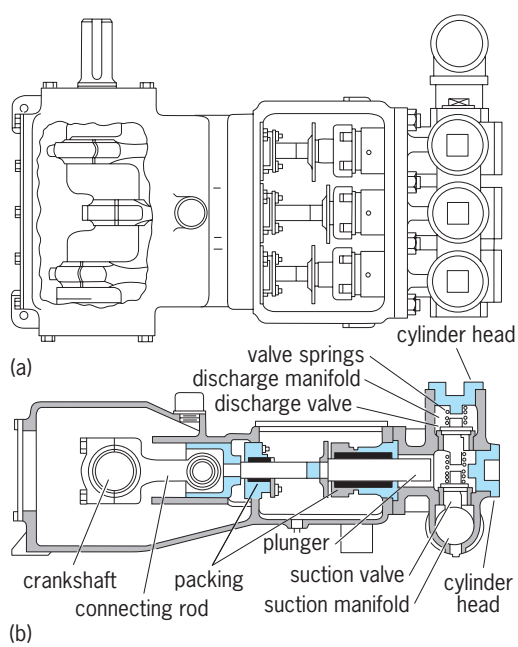

Fig. 1. Horizontal triplex power pump of reciprocating type. (*a*) Plan. (*b*) Elevation.

is desirable. Power-driven reciprocating pumps are highly efficient over a wide range of discharge pressures. Except for some special designs with continuously variable stroke, reciprocating power pumps deliver essentially constant capacity over their entire pressure range when driven at constant speed. In some applications this is an advantage, but in others it complicates the controls required.

Direct-acting steam types. A reciprocating pump is readily driven by a reciprocating engine; a steam or power piston at one end connects directly to a fluid piston or plunger at the other end. Direct-acting reciprocating pumps are simple, flexible, low-speed machines which are low in efficiency unless the heat in the exhaust steam can be used for heating. Steam pumps can be built for a wide range of pressure and capacity by varying the relative size of the steam piston and the liquid piston or plunger. The delivery of a steam pump may be varied at will from zero to maximum simply by throttling the motive steam, either manually or by automatic control. Direct-acting pumps are built as simplex, having one steam and one liquid cylinder; and duplex, having two steam and two liquid cylinders side by side. As indicated by **Fig. 2**, each steam valve of a duplex pump is positively driven by the motion of the opposite piston rod by means of cranks and links. In the case of a simplex pump, to avoid stalling at low speed, the valve linkage operates a small pilot valve which in turn controls a piston-thrown main steam valve.

Reciprocating pumps are used for low to medium capacities and medium to highest pressures. They are useful for low- to medium-viscosity fluids, or high-viscosity fluids at materially reduced speeds. Specially fitted reciprocating pumps are used to pump fluids containing the more abrasive solids.

Rotary types. Another form of displacement pump consists of a fixed casing containing gears, cams, screws, vanes, plungers, or similar elements actuated by rotation of the drive shaft. Most forms of rotary pumps are valveless and develop an almost steady flow rather than the pulsating flow of a reciprocating pump. Three of the many types of rotary pumps are shown in **Fig. 3**.

Rotary pumps require very close clearances between rubbing surfaces for their continued volumetric efficiency. Consequently they are most useful for

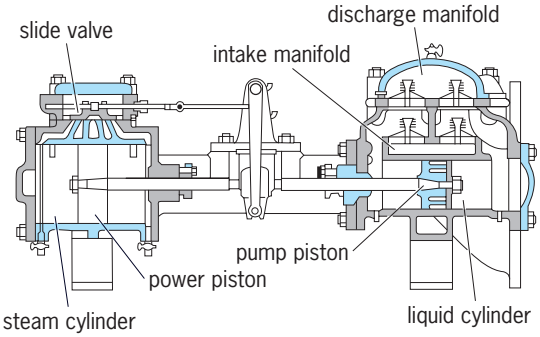

Fig. 2. Duplex type of direct-acting steam-driven feedwater pump.

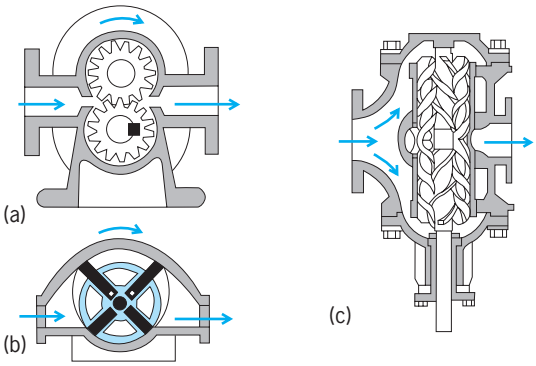

Fig. 3. Rotary pumps. (*a*) Gear. (*b*) Guided vane. (*c*) Screw. (*After E. A. Avallone and T. Baumeister III, eds., Marks' Standard Handbook for Mechanical Engineers, 9th ed., McGraw-Hill, 1987*)

pumping clean oils or other fluids having lubricating value and sufficient viscosity to prevent excessive leakage. On petroleum oils of suitable viscosity, rotary pumps are highly efficient at moderate pressure and speed, while at reduced speed they can pump, with lower efficiency, the most viscous materials. The increasing use of hydraulic actuation of machine tools and mechanisms, such as power steering of automobiles, has extended the use of rotary pumps and similar hydraulic motors.

Vacuum types. Although vacuum pumps actually function as compressors, displacement pumps are used for certain vacuum pump applications. Simplex steam pumps with submerged piston pattern fluid ends similar to the fluid end in Fig. 2 are used as wet vacuum pumps in steam heating and condensing systems. Sufficient liquid remains in the cylinder to fill the clearance volume and drive the air or gas out ahead of the liquid. Certain types of rotary pumps are arranged to retain a quantity of sealing oil when operating as vacuum pumps. *See* VACUUM PUMP.

Air-lift types. In handling abrasive or corrosive waters or sludges, where low efficiency is of secondary importance, air-lift pumps are used. The pump consists of a drop pipe in a well with its lower end submerged and a second pipe which introduces compressed air near the bottom of the drop pipe. The required submergence varies from about four times the distance from the water level to the surface for a low lift to an equal distance for a relatively high lift. The mixture of air and water in the drop pipe is lighter than the water surrounding the pipe. As a result, the mixture of air and water is forced to the surface by the pressure of submergence. *See* COMPRESSOR; PUMP; PUMPING MACHINERY. Elliott E. Wright

Bibliography. I. J. Karassik et al., *Pump Handbook*, 3d ed., 2000.

Distance-measuring equipment

An internationally standardized navigation system which allows an aircraft to measure its distance from a selected ground-based beacon. Such beacons are used throughout the world by all airliners, most of

the military aircraft of the West, and a large number of general-aviation aircraft. The range of service is line of sight up to 300 mi (480 km), and system accuracy is usually 0.1 mi (0.16 km), but precision equipment, intended for use during landing, has accuracy of 20 ft (7 m).

The airborne equipment, called an interrogator, transmits pulses of 1 kW peak power on 1 of 126 frequencies. These are in the 1025–1150-MHz band and are spaced 1 MHz apart. Each pulse has a duration of 3.5 microseconds and is paired with another, spaced 12 or 36 μs later. The combination of frequencies and pulse spacings therefore provides 252 operating channels.

The beacon on the ground, called a transponder, receives these pulses, delays them by 50 μs, and then retransmits them, usually with a power of 1 kW, on 252 frequencies lying between 962 and 1213 MHz. The pulse-pair spacing is 12 μs on those frequencies not used by the interrogator, and 30 μs on those frequencies shared with the interrogator. The transponder transmission is called the reply. The frequency difference between interrogation and reply is always 63 MHz. This arrangement allows each transmitter frequency to act as the local oscillator for its associated superheterodyne receiver, the intermediate frequency of which is 63 MHz. For landing purposes, some transponders have powers as low as 100 W (see **illus.**).

The airborne transmission has a pseudorandom pulse repetition rate ranging from about 15 pulse pairs per second in the newest equipment to as high as 150 pulse pairs per second in older equipment. This reduction is effected to accommodate a larger number of aircraft per beacon. With an average rate of 30 interrogations per second, a typical beacon can reply to 100 aircraft.

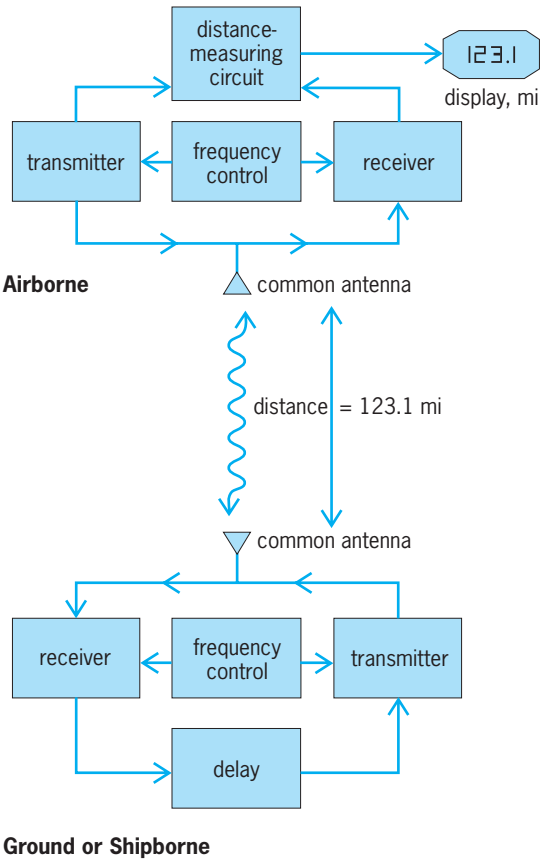

Block diagram of distance-measuring equipment system.

In the aircraft the replies to its own interrogations are recognized by their phase coherence with their own transmissions, and by the elapsed time measured between transmissions and reception (minus the 50-μs transponder delay), usually by means of a crystal clock. This elapsed time is about 12 μs for each nautical mile (7 μs for each kilometer), and is displayed in the cockpit on a digital meter, which is usually calibrated in miles and tenths of miles.

Transmission, both airborne and ground, is vertically polarized and employs the same antenna as for reception. The airborne antenna is typically a 5-in. (13-cm) blade (half wavelength) protruding from the belly of the aircraft. The ground antenna usually employs several half-wave sections, totaling about 5 ft (1.5 m) in height.

Much airborne equipment, particularly that used by general aviation, has small adapter units which compute radial speed with respect to the beacon or time-to-go with respect to the beacon, or both. By comparing these numbers with the indicated airspeed, the magnitude of headwind may be gaged. This factor is of greater importance to small aircraft than to airliners.

Typical airliner equipment weighs 12 lb (5.4 kg), while general-aviation types are about 5 lb (2.3 kg). All are entirely solid-state.

Of the 252 two-way channels, 200 are permanently paired on a one-for-one basis with the 200 channels used by the very high-frequency omnidirectional range (VOR) and its associated instrument landing system (ILS) localizer. Thus, in all civil aircraft, a common single channel selector is used for distance-measuring equipment (DME), VOR, and ILS. This rho-theta navigation system is the backbone of the airways throughout the world. It is fully implemented in the United States, Europe, and Japan, with somewhat less implementation elsewhere. A notable latter exception is Australia, where the DME, while also used with VOR, operates in the 200-MHz band. Tacan is a military alternative rho-theta system based on DME.

Jet-powered airliners, typically flying above 15,000 ft (4500 m), are frequently within range, simultaneously, of many DME beacons, whose geographical spacings have been based on service to lower-flying aircraft. This has led to the airborne scanning-DME which automatically scans all beacons within range and then selects two or more to provide a rho-rho (DME/DME) fix of somewhat greater accuracy than that obtained from rho-theta (DME/VOR).

A modification called DME-P, for precision, has been developed for use with the microwave landing system (MLS). This system can provide an accuracy of 20 ft (7 m) within 7 mi (11 km) and 100 ft (30 m) within 20 mi (12 km), using the same frequencies as the standard system (now called

DME-N), but with steeper pulse rise times, lower power, and new pulse-pair codes. *See* AIR NAVIGATION; ELECTRONIC NAVIGATION SYSTEMS; INSTRUMENT LANDING SYSTEM (ILS); MICROWAVE LANDING SYSTEM (MLS); RHO-THETA SYSTEM; TACAN; VOR (VHF OMNIDIRECTIONAL RANGE). Sven H. Dodington

Bibliography. M. Kayton (ed.), *Navigation*: *Land, Sea, Air, and Space*, 1990; R. J. Kelly and D. R. Cusick, *DME and Its Evolving Role in Aviation*, 1987; G. W. Mixter and H. Headley (eds.), *Primer of Navigation*, 7th ed., 1995; Radio Technical Commission for Aeronautics, *Minimum Standards for DME*, DO-189, 1985.

Distillation

A method for separating homogeneous mixtures based upon equilibration of liquid and vapor phases. Substances that differ in volatility appear in different proportions in vapor and liquid phases at equilibrium with one another. Thus, vaporizing part of a volatile liquid produces vapor and liquid products that differ in composition. This outcome constitutes a separation among the components in the original liquid. Through appropriate configurations of repeated vapor-liquid contactings, the degree of separation among components differing in volatility can be increased manyfold. *See* PHASE EQUILIBRIUM.

Distillation is by far the most common method of separation in the petroleum, natural gas, and petrochemical industries. Its many applications in other industries include air fractionation, solvent recovery and recycling, separation of light isotopes such as hydrogen and deuterium, and production of alcoholic beverages, flavors, fatty acids, and food oils.

Simple distillations. The two most elementary forms of distillation are a continuous equilibrium distillation and a simple batch distillation (**Fig. 1**).

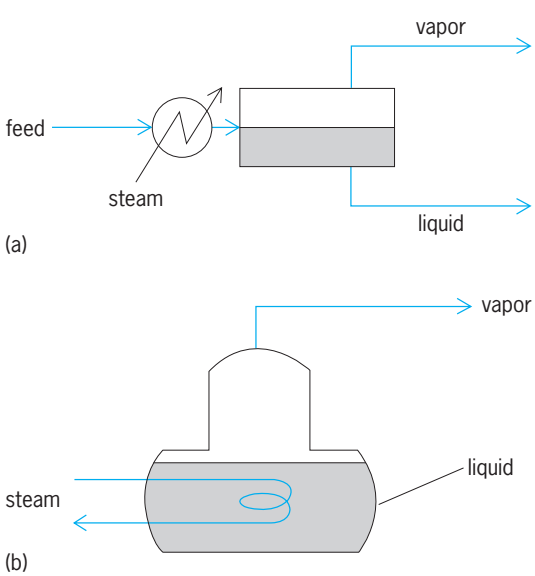

Fig. 1. Simple distillations. (*a*) Continuous equilibrium distillation. (*b*) Simple batch distillation.

In a continuous equilibrium distillation, a continuously flowing liquid feed is heated or reduced in pressure (flashed) so as to cause partial vaporization. The vapor and liquid disengage while flowing through an open drum, and the products emerge as vapor and liquid streams. The vapor product can be condensed to form a liquid distillate. It is also possible to use a vapor feed, subjected to cooling and thereby partial condensation, again followed by disengagement of the resultant vapor and liquid in an open drum. *See* VAPOR CONDENSER.

In a simple batch distillation, an entire batch of liquid is initially charged to a vessel and is then heated, typically by condensation of steam inside a metal coil within the vessel. Vapor is thereby continuously generated, and may be condensed to form a liquid distillate, which is collected. In the batch distillation, increments of vapor are formed in equilibrium with all liquid compositions ranging from the original to the final, whereas the continuous equilibrium distillation gives vapor in equilibrium with only the final liquid composition. Since the distillate consists primarily of the more volatile components and the feed liquid contains more of these substances than does the final liquid, the simple batch distillation gives a more enriched distillate than does the continuous equilibrium distillation.

Fractional distillation. Unless the vapor pressures of the species being separated are very dissimilar, a simple distillation does not produce highly purified products. Product purities can be increased by repeated partial vaporizations and condensations (**Fig. 2***a*). The liquid from an initial continuous equilibrium distillation (L_0') can be partially vaporized by additional heating. The remaining liquid (L_1') can again be heated and partially vaporized, forming another liquid (L_2'); and so forth. Each liquid is progressively enriched in the less volatile substances. Similarly, successive partial condensations of the vapor fraction (V) from the initial continuous equilibrium distillation produce vapor products successively enriched in the more volatile components. *See* VAPOR PRESSURE.

The process involved in successive partial vaporizations and condensations (Fig. 2*a*) would lead to only very small amounts of the most enriched products, along with numerous streams having intermediate compositions. A logical step is to recycle each of these intermediate streams to the prior vessel in the sequence of contactors (Fig. 2*b*). Recycling of the intermediate vapors and liquids has another highly beneficial effect in that it negates the need for intermediate heaters and coolers. The resultant process (Fig. 2*b*) is known as continuous fractional distillation. It is usually carried out in a distillation column (**Fig. 3**), which is a simpler, more compact form of equipment than the cascade of vessels used in the process of recycling intermediate vapors and liquids. However, the processes are functionally equivalent.

A distillation column consists of a series of stages, also called trays or plates, on which vapor and liquid are contacted. The liquid and vapor pass through the

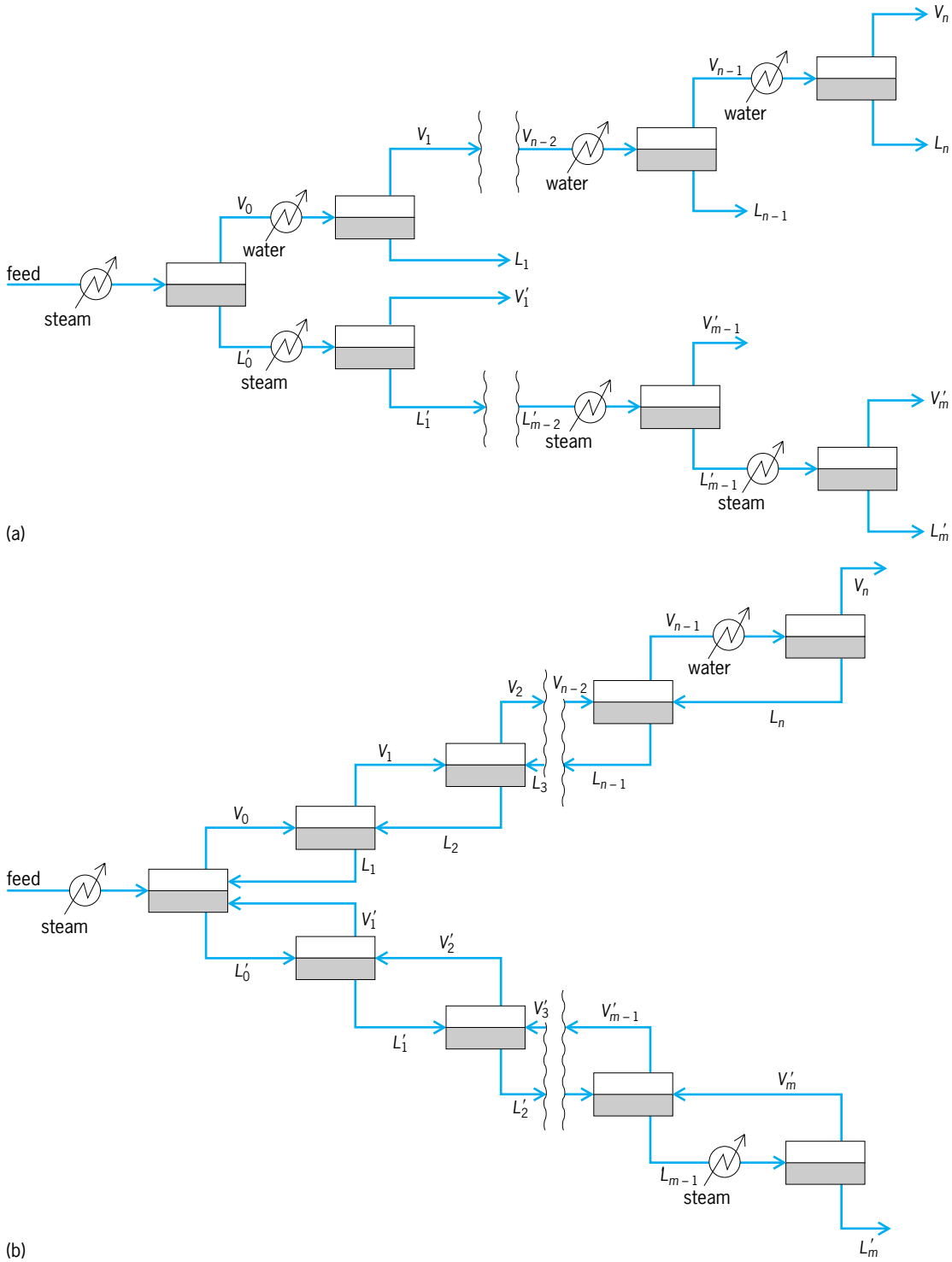

Fig. 2. Distillation by repeated vapor-liquid equilibrations. (*a*) Increasing product purities by successive partial vaporizations and condensations. (*b*) Recycling intermediate products and elimination of intermediate heating and cooling steps; equivalent to continuous fractional distillation. Symbols are explained in the text.

column in opposite directions, the liquid downward and the vapor upward, a pattern known as countercurrent flow. Liquid is held onto the trays by weirs over which the liquid exits the tray, passing into a vertical downcomer that conveys the liquid to the tray below. Vapor rises up the column through the trays successively.

Various tray designs are used. Sieve trays contain a large number of small holes, of the order of 4 mm (0.16 in.) in diameter, and the velocity of the vapor through these holes is sufficient to prevent leakage of liquid through them. Valve-cap trays contain much larger holes, of the order of 3 cm (1.2 in.), each fitted with a free-floating bonnet. Bubble-cap

trays are common in older columns and consist of a fixed bonnet, which forces the vapor out horizontally through slots. Sometimes distillation columns are constructed without discrete stages and simply have vapor and liquid flowing countercurrently through a bed of divided solid material, called packing. The form of the packing is usually chosen to give high surface area per unit bed volume, as well as a high porosity. *See* DISTILLATION COLUMN.

It is necessary for the operation of continuous fractional distillations (Figs. 2*b* and 3) that there be a cooler at the top and a heater at the bottom of the cascade of countercurrent stages. The overhead cooler is known as the condenser, and the liquid that it generates is known as reflux. The heater is termed a reboiler. The stages above the point at which the feed is supplied are known as the rectifying, or enriching, section, and they serve to remove the less volatile components from the more volatile components that become the overhead product. The stages below the feed point are known as the stripping section, since they serve to strip the more volatile components out of the less volatile components that compose the bottom product. The feed can be either vapor or liquid, or a mixture of the two. There can be multiple feeds, which may enter the column at different locations. There can also be intermediate products, or sidestreams, withdrawn.

A batch distillation can also be carried out in staged equipment. The vertical section of the batch still (Fig. 1*b*) is then equipped with trays and serves as a rectifying section, and an overhead condenser serves to generate liquid reflux, which is returned to the top of the column. Batch distillation with a rectifying section serves to remove the less volatile components to a greater extent from the overhead product; however, the absence of a stripping section means that there cannot be a similar added purification of the remaining liquid.

Vapor-liquid equilibria. The separation accomplished in a distillation relates to the difference in composition of vapor and liquid phases at equilibrium. These relationships are the subject matter of phase-equilibrium thermodynamics. *See* CHEMICAL THERMODYNAMICS; GAS; LIQUID.

Compositions of phases in equilibrium can be determined experimentally. A common device for measuring vapor-liquid equilibria is the Othmer still, in which a liquid is heated to form equilibrium vapor, which is then condensed and delivered to a distillate receiver, from which distillate liquid drains back to mix with the original liquid. Sampling and analyzing the distillate and the main liquid gives the compositions of a vapor and a liquid, respectively, in equilibrium with one another.

Powerful methods have been developed for predicting vapor-liquid equilibrium data. Vapor pressures are obtained from measurements of pure components or from any of several methods for predicting vapor pressures based on deductive reasoning. Activity coefficients are obtained from any of a variety of predictive methods; one of the more commonly used is known as the UNIFAC method.

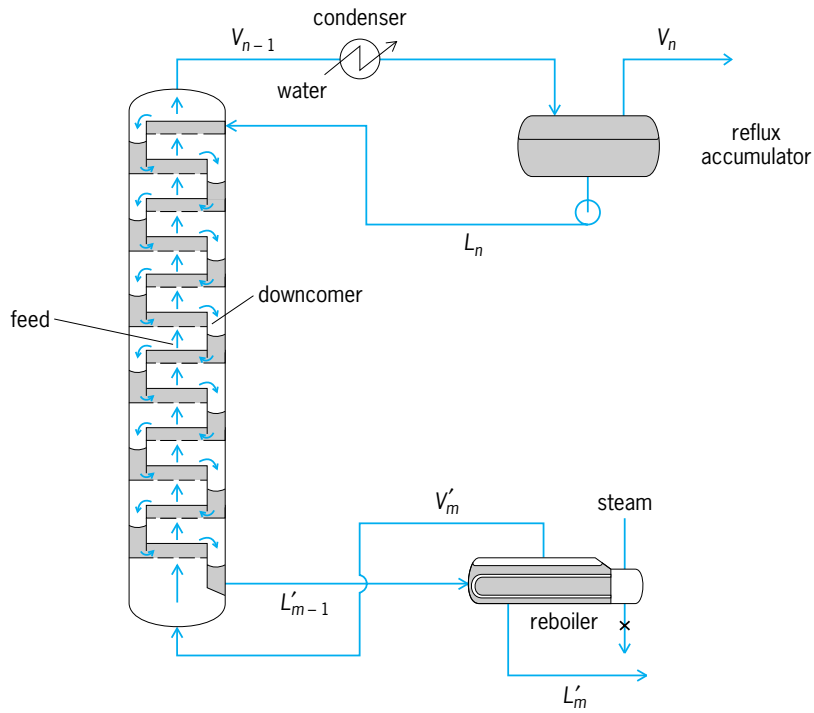

Fig. 3. Distillation column, carrying out continuous fractional distillation.

The dependence of liquid-phase activity coefficients upon liquid composition is complex; consequently vapor-liquid equilibrium relationships can have quite different characteristics. For example, the two-component benzene-toluene system (**Fig. 4**, curve A) is a relatively ideal system, for which the activity coefficients are close to 1.0 over the liquid composition range. The mole fraction of benzene in the vapor is therefore always greater than that in the

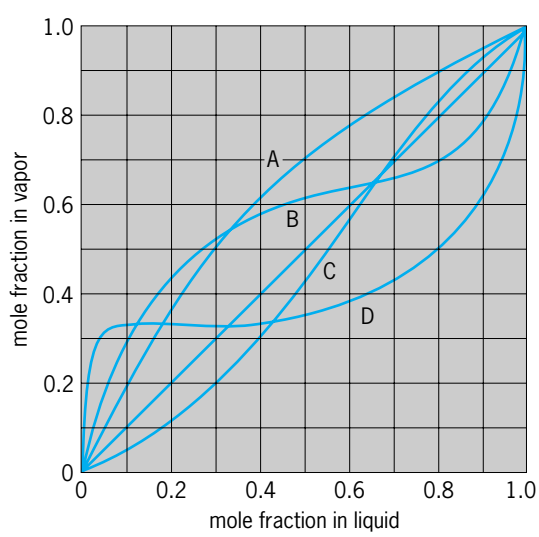

Fig. 4. Types of binary vapor-liquid equilibrium relationship; equilibrium vapor mole fraction of the more volatile (first named) component as a function of the liquid mole fraction of the more volatile component. The pressure is constant at 1 atm (101 kilopascals) for all systems. Curve A, benzene-toluene; B, acetone-carbon disulfide; C, acetone-chloroform; D, isobutyl alcohol-water.

liquid, and the relative volatility has a nearly constant value of 2.4. Such a system can be separated readily into products of high purity by continuous fractional distillation.

In the acetone-carbon disulfide system (Fig. 4, curve B), there are positive deviations from ideality, meaning that the activity coefficients are greater than 1.0 and increase for either component as that component becomes more dilute in the mixture. The result is that the relative volatility is greatest at low concentrations of the more volatile component (acetone) and decreases at higher concentrations. In this case, the curve crosses the line at which the vapor mole fraction (y_A) equals the liquid mole fraction (x_A) at 0.67, and the relative volatility (α_{AB}) becomes less than 1.0 at higher values of x_A and y_A. There is therefore a minimum-boiling azeotrope with a composition $x_A = 0.67$. Continuous fractional distillation cannot produce two products of high purity for such a system. A composition approaching that of the azeotrope is created as distillate, and depending upon the feed composition, either relatively pure A or relatively pure B is produced as bottoms. *See* AZEOTROPIC MIXTURE.

Because of hydrogen bonding between the two components, the acetone-chloroform system (Fig. 4, curve C) exhibits negative deviations from ideality, meaning that the activity coefficients are less than 1.0. Consequently, the relative volatility is less than 1.0 at low x_A and increases to greater than 1.0 at high x_A. A maximum-boiling azeotrope occurs at $x_A = 0.66$. Continuous fractional distillation cannot cross the azeotropic composition. It produces either relatively pure A or relatively pure B as distillate and a composition approaching that of the azeotrope as bottoms.

The isobutanol-water system (Fig. 4, curve D) exhibits sufficiently extreme positive deviations from ideality so that two immiscible liquid phases form over much of the composition range. Over the range of x_A where two immiscible liquids are formed, the equilibrium y_A is constant, and the curve is flat. The y_A-x_A equilibrium relationship crosses the $y_A = x_A$ diagonal within this region. The system, therefore, forms what is known as a heterogeneous azeotrope at $y_A = 0.34$, where y_A is equal to the hypothetical x_A that would obtain if the two immiscible liquids were to have formed a single liquid phase.

Distillation pressure. The pressure under which a distillation is performed is a matter of choice, although operation at pressures more removed from atmospheric becomes more costly because of the needs for pressure vessels at high pressure and for large column diameters and means of producing vacuum at subambient pressures. Often the column pressure is dictated by the pressure at which the feed is available or at which the products are to be utilized. Altering the pressure of a distillation can serve to alter the vapor-liquid equilibrium relationship, a feature that can be used to advantage. For example, the minimum-boiling azeotrope in the ethanol-water system disappears at a low-enough column pressure.

Because of the relationship between vapor pressure and temperature, the temperatures within a distillation column are lower for lower pressures. Vacuum distillation is an effective means of maintaining lower temperatures for separations involving heat-sensitive materials.

Steam distillation is an alternative to vacuum distillation for separations of organic substances. In this process, steam is fed directly to the bottom of a column and passes upward, composing a substantial fraction of the vapor phase. The combined partial pressures of the organic substances being distilled are thereby lessened, giving the lower temperatures characteristic of a vacuum distillation.

Molecular distillation. An extreme of vacuum distillation is molecular distillation, in which a liquid is vaporized from a pool or film, with a chilled condenser surface only a small distance away from the surface of the liquid. The process is carried out at such low pressure that the distance between the liquid surface and the condenser surface is less than the mean free path of the vapor, meaning that the molecules do not collide with one another or with any other molecule on their way from the liquid to the condenser surface. Typical conditions are a pressure of 0.001 mmHg (0.13 pascal) absolute and a gap of about 0.2 cm (0.08 in.). One molecular distillation apparatus, the Hickman still, uses centrifugal force to spread a thin film of liquid across a heating surface.

Molecular distillation enables a distillative separation to occur at minimum temperature and minimum time of exposure to that temperature. It is therefore well suited for heat-sensitive materials having relatively low volatilities, typically substances in the molecular weight range of 400–1200.

Extractive and azeotropic distillation. Another way to alter relative volatilities and thereby overcome limitations stemming from azeotropes and to facilitate distillations in other ways is to add another component in an amount sufficient to alter the relative volatility of the components being separated. In

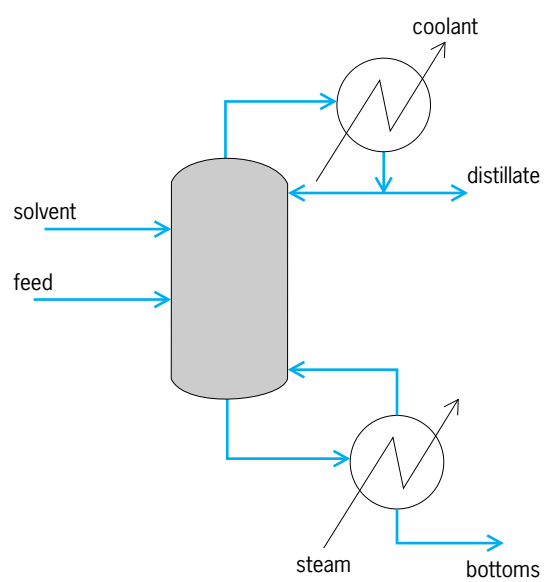

Fig. 5. Column used for extractive distillation.

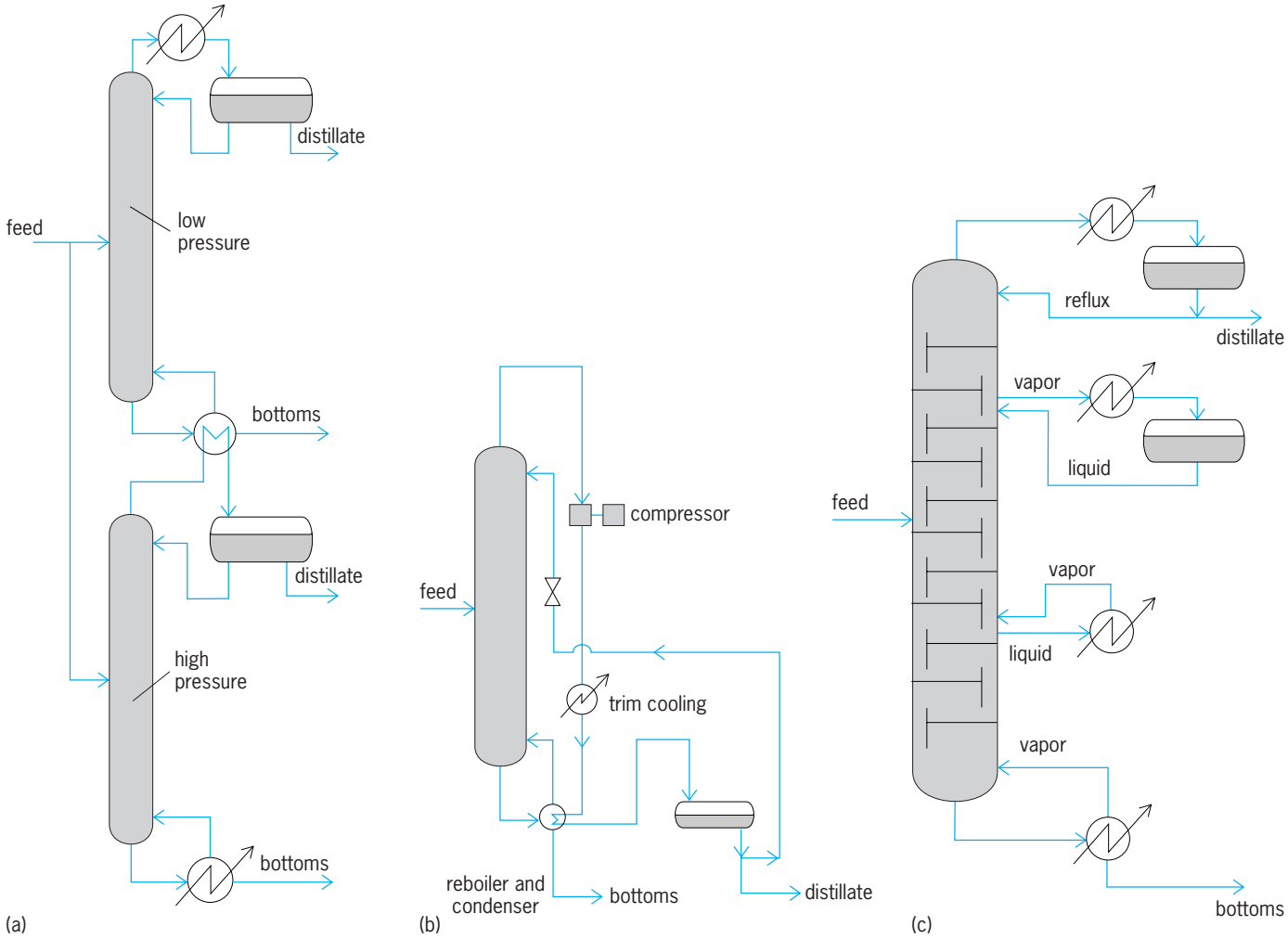

Fig. 6. Methods of reducing energy consumption in distillation. (*a*) Multieffect columns. (*b*) Vapor recompression. (*c*) Intermediate condensers and reboilers.

extractive distillation (**Fig. 5**), a solvent is added to a stage above the feed in a continuous fractional distillation. The solvent forms a substantial portion of the liquid on all stages below the addition point and alters the activity coefficients of the feed components in the liquids on those stages. If the solvent interacts with the less volatile component preferentially, the relative volatility between the feed components becomes greater, and fewer stages are needed to separate the feed components. The stages between the feed and solvent-addition points serve as the rectifying section, reducing the concentration of the less volatile component; and the stages above the solvent-addition point serve to remove solvent from the overhead distillate. A subsequent separation is needed to recover the less volatile component from the solvent.

A related process is azeotropic distillation, where an added volatile component comes overhead in the distillate. The design of the column is specified so that the added component exists at a high-enough concentration in the liquid on a number of trays to improve the relative volatility of the feed components on those stages. *See* AZEOTROPIC DISTILLATION.

Enhancing energy efficiency. Distillation consumes heat energy in the reboiler or refrigeration energy in the condenser of a column operating at subambient temperature. Often the amount of energy needed is large, and there is incentive to reduce it. Three methods have been employed to lessen energy consumption in distillation (**Fig. 6**). In one method, a distillation is carried out in two different columns, operating at different pressures (Fig. 6*a*). The difference between column pressures is such that the condenser of the high-pressure column can serve as the reboiler for the low-pressure column. Thus, the requirement for steam or other heating medium is approximately halved, although the temperature span between the high-pressure reboiler and the low-temperature condenser is approximately doubled. This approach is used in the production of oxygen and nitrogen by low-temperature distillation of air.

In a vapor-recompression system (Fig. 6*b*), the overhead vapor is compressed to a high-enough pressure that it will condense in the reboiler, thereby replacing reboiler energy with mechanical compression. This approach is most attractive for distillations of very close boiling mixtures, such as propylene

and propane, where the temperature span across the distillation is not large.

The third method uses intermediate condensers and reboilers (Fig. 6*c*). The reflux and boil-up requirements of a continuous fractional distillation are typically set by the sizes of the liquid and vapor flows needed in the vicinity of the feed stage. A less cold medium can be used in an intermediate condenser, to the extent that some of the condensation can be accomplished there rather than entirely in the overhead condenser. Similarly, a heating medium at a lower temperature can be used in an intermediate reboiler, to the extent that some of the vaporization can be accomplished there rather than entirely in the reboiler at the bottom of the column.

Computational methods. If the number of equilibrium stages, the column pressure, the descriptors of the feeds (composition, flow rate, phase condition and location), the distillate flow rate, the condenser or reboiler heat duty, and the location, phase condition, and flow rates of any sidestreams are specified, the computation of the product compositions and the compositions of liquid and vapor on all stages in a continuous fractional distillation becomes a problem of solving the so-called MESH equations. M refers to mass balances for each component for each stage; E, to vapor-liquid equilibrium relationships for each component on each stage; S, to summation of individual component flows to produce the total liquid and vapor flows leaving each stage; and H, to enthalpy balances for each stage. *See* ENTHALPY.

Historically, a number of assumptions (constant relative volatilities, constant molar liquid and vapor flows from stage to stage, and others) have often been made to simplify calculation methods for continuous fractional distillations. However, as more capable computers have increased computational power, it has become common to solve the equations in full, provided that adequate data or predictive methods are available for vapor-liquid equilibria and enthalpies. Numerous computer routines are commercially available for this purpose.

Stage efficiency. The stages, or trays, of a distillation column generally do not provide equilibrium between the vapor and liquid exiting the stage. There are two primary reasons. First, the residence times of the vapor and liquid and the intensity of contacting are usually not sufficient for equilibrium to be achieved. This effect reduces the separation as compared with simple equilibrium. Second, the liquid typically flows across the tray, and the vapor equilibrates against liquids of all compositions from inlet to outlet. This effect enhances the separation as compared with simple equilibration. Prediction of stage efficiencies is complex and is typically the area of most uncertainty in the design of distillation columns. To a large extent, it is still based upon experience, although research is continually providing insight into the flow, mixing, and mass-transfer phenomena involved. *See* FIRED HEATER. C. Judson King

Bibliography. R. Billet, *Evaporation Technology and Its Technical Applications*, 1989; E. J. Henley and J. D. Seader, *Equilibrium-Stage Separation Operations in Chemical Engineering*, 1981; C. D. Holland, *Fundamentals of Multicomponent Distillation*, 1981; C. J. King, *Separation Processes*, 2d ed., 1980; H. Z. Kister, *Distillation Design*, 1992; P. C. Wankat, *Equilibrium Staged Separations*, 1988.

Distillation column

An apparatus used widely for countercurrent contacting of vapor and liquid to effect separations by distillation or absorption. In general, the apparatus consists of a cylindrical vessel with internals designed to obtain multiple contacting of ascending vapor and descending liquid, together with means for introducing or generating liquid at the top and vapor at the bottom.

In a column that can be applied to distillation (**Fig. 1**), a vapor condenser is used to produce liquid (reflux) which is returned to the top, and a liquid heater (reboiler) is used to generate vapor for introduction at the bottom. In a simple absorber, the absorption oil is the top liquid and the feed gas is the bottom vapor. In all cases, changes in composition produce heat effects and volume changes, so that there is a temperature gradient and a variation in vapor, and liquid flows from top to bottom of the column. These changes affect the internal flow rates from point to point throughout the column and must be considered in its design.

Distillation columns used in industrial plants range in diameter from a few inches to 40 ft (12 m) and in height from 10 to 200 ft (3 to 60 m). They operate at pressures as low as a few millimeters of mercury and

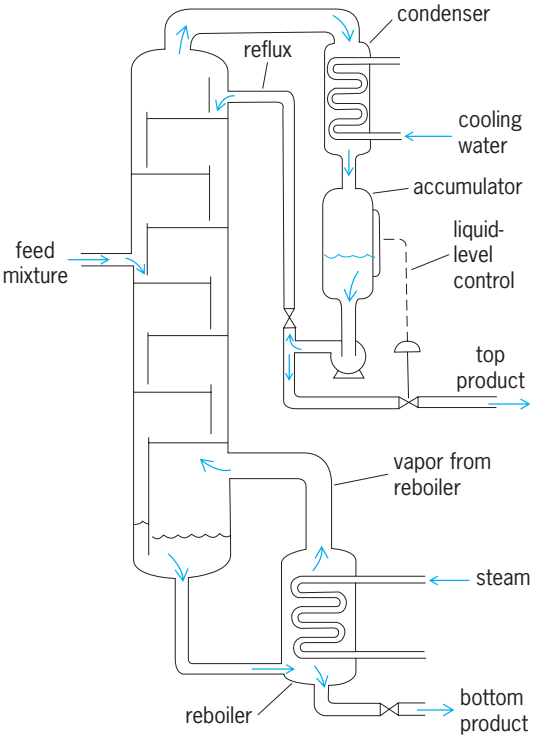

Fig. 1. Elements of a distillation column.

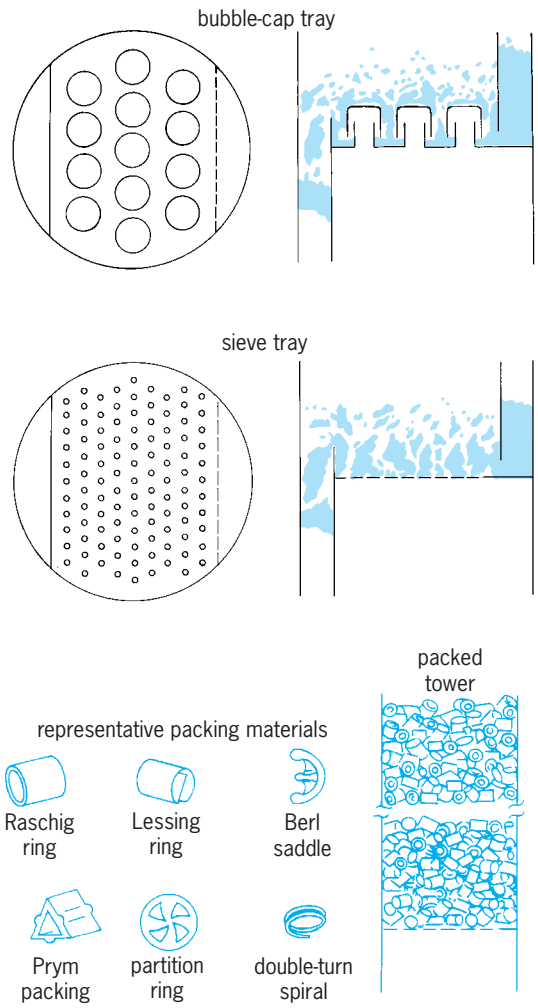

Fig. 2. Contacting devices for fractionating column

as high as 3000 lb/in.^{2} (2 megapascals) at temperatures from −300 to 700°F (−180 to 370°C). They are made of steel and other metals, of ceramics and glass, and even of such materials as bonded carbon and plastics.

A variety of internal devices have been used to obtain more efficient contacting of vapor and liquid. The most widely used devices are the bubble-cap plate, the perforated or sieve plate, and the packed column (**Fig. 2**).

Plates. The bubble-cap plate is a horizontal deck with a large number of chimneys over which circular or rectangular caps are mounted to channel and distribute the vapor through the liquid. Liquid flows by gravity downward from plate to plate through separate passages known as downcomers. The number of caps and the size of downcomers are designed to fit the expected internal flow rates of liquid and vapor.

The bubble-cap plate is the most common type of contacting device used in distillation columns. It can handle a wide range of vapor and liquid flows, is not obstructed by small amounts of rust and other solids, and is manufactured by a large number of suppliers.

The perforated or sieve plate is a horizontal deck with a multiplicity of round holes or rectangular slots for distribution of vapor through the liquid. The sieve plate can be designed with downcomers similar to those used for bubble-cap trays, or it can be made without downcomers so that both liquid and vapor flow through the perforations in the deck. The simple sieve tray has the disadvantage of a restricted operating range, so that operation at vapor flows much less than the design rate results in poor contacting.

A number of variations and improvements have been made in the sieve tray to improve the operating range. The most successful of these is the valve tray, in which a check valve is placed over each perforation to adjust for the variation in vapor flow. The valve usually consists of a disk over each perforation which is held within a cage attached to the deck. The disk is free to move up and down within the limits imposed by the cage, and thus adjusts to the vapor flow.

Although it is older than the bubble-cap plate, the sieve tray did not obtain wide use until stainless steel sheets became readily available to overcome difficulties with corrosion and fouling. With the improvements, such as the valve tray, this type of plate may supplant the bubble-cap plate in many services.

Packings. The packed column is a bed or succession of beds made up of small solid shapes over which liquid and vapor flow in tortuous countercurrent paths. Expanded metal or woven mats are also used as packing. The packed column is used without downcomers, but in larger sizes usually has horizontal redistribution decks to collect and redistribute the liquid over the bed at successive intervals of height. The packed column is widely used in laboratories. It is often used in small industrial plants, especially where corrosion is severe and ceramic or glass materials must be chosen.

Other contacting means sometimes used in fractionating columns are the disk-and-doughnut tray for very viscous or dirty liquids, the spiral ribbon for very small columns, the spray column, and the welled-wall column. *See* GAS ABSORPTION OPERATIONS; DISTILLATION.

Mott Souders

Bibliography. R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; A. H. Skelland, *Diffusional Mass Transfer*, 1985.

Distilled spirits

Potable alcoholic beverages, each obtained by distilling an alcohol-containing liquid and further treating the distillate to obtain a beverage of specific character.

Classification. The various distilled spirits are classified according to (1) the raw material which has been fermented and subsequently distilled, such as grain, molasses, and fruit; and (2) the further treatment given the distillate to add specific flavors

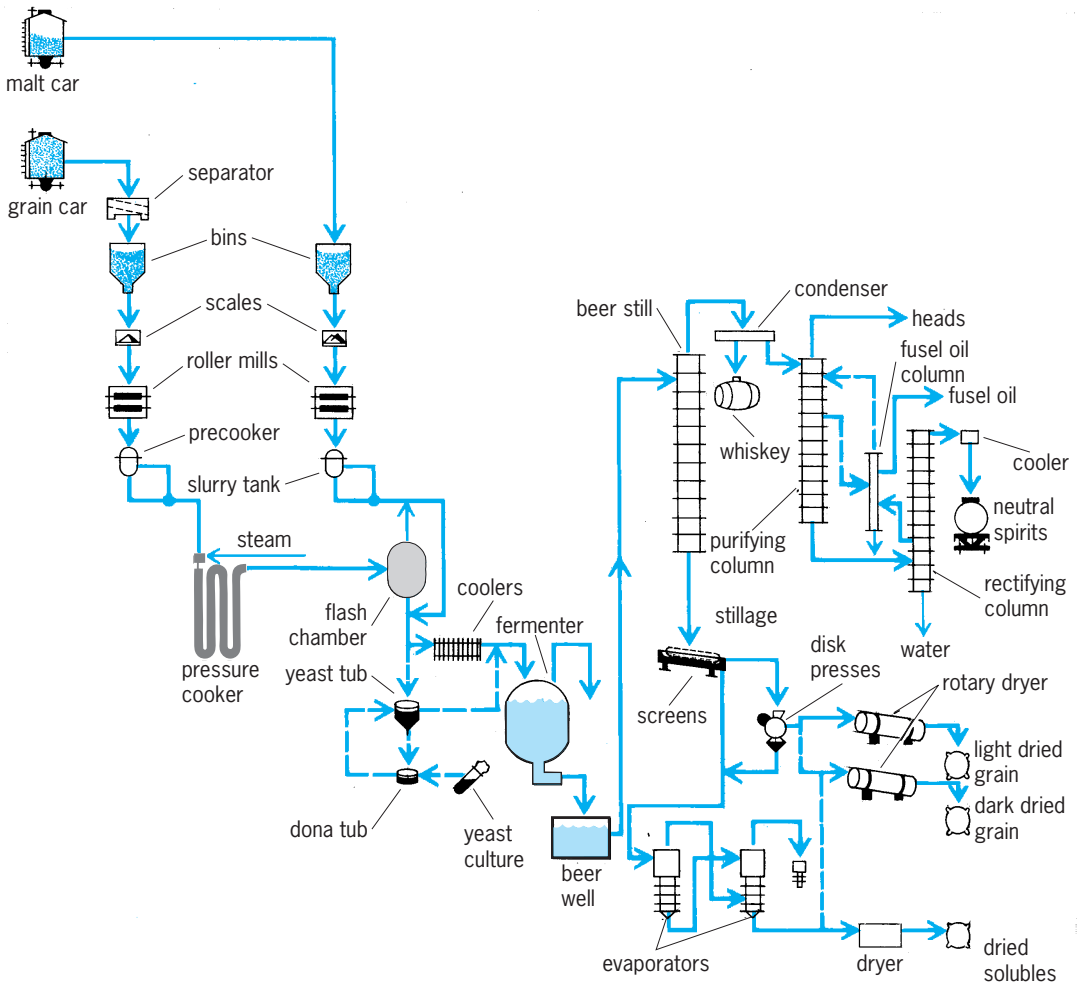

Flow diagram of whiskey production. (*J. E. Seagram and Sons*)

and aroma. These are referred to as compounded or flavored spirits. Examples of beverages made from grain are whiskeys, vodka, and gin. Some types of vodka and gin are made from potatoes. Molasses is used for the manufacture of rum, and sugarcane juice for the Brazilian cachaça or pinga. Fruits are used for the preparation of brandies, and agave juice for Mexican tequila. The so-called compounded or flavored spirits are made from distilled spirits by the addition of sugar, essential oils, color, herbs, bitters, or other ingredients. Examples are English and Dutch gins, various liqueurs or cordials, absinthe, and aquavit.

Fermentation. When starchy materials are to be fermented, it is necessary to convert the starch to fermentable sugars first. This is done by the addition of amylolytic or diastatic enzymes from malt or certain fungi. Once the starch has been converted to sugar, mostly maltose and glucose, it can be fermented by yeast to alcohol and carbon dioxide. When wine is distilled the fermentation is carried out by various strains of the wine yeast, *Saccharomyces cerevisiae* or its variety *ellipsoideus*. With other sugar-containing substrates, special distillery yeasts may be employed, usually also strains of *S. cerevisiae*, or a natural fermentation by uncontrolled species of yeast is used. *See* FERMENTATION; WINE.

Distillation process. This is done in stills of various design. In a simple pot still very little fractionation, or rectification, is accomplished. When a still is equipped with a rectifying column, fractions of much higher purity can be collected (see **illus.**). The batch-type distillation is becoming replaced more and more by continuous stills, especially in the production of grain alcohol or neutral spirits. Normally, three types of fractions are separated. A fraction with a low boiling point, the "heads," is rich in aldehydes and esters. The main fraction is ethyl alcohol and a higher-boiling fraction is rich in so-called fusel oils or higher alcohols. The latter consist of D-amyl- and isoamylalcohol, butyl- and isobutylalcohol, propanol, and others. Isoamylalcohol and amylalcohol are the principal ingredients of fusel oil. The strength of the ethyl alcohol is expressed by the term "proof." In the United States 100 proof spirit is a spirit containing 50% alcohol by volume at a temperature of 60°F (16°C).

Types. The various types of distilled spirits are discussed below.

Brandies. Brandy is a spirit obtained from the distillation of wine or a fermented fruit juice, usually after aging of the wine in wooden casks. Cognac is a brandy distilled from wines made of grapes grown within the legal limits of Charente and Charente Inférieure Departments, the Cognac region of France. Armagnac is brandy made in the Department of Gers, southeast of Bordeaux. Both types of brandy are aged for many years in oak before bottling. Before bottling, various lots are blended, the alcohol content is adjusted to 42–44% and if necessary, coloring matter such as caramel is added. Brandies distilled from grape pomace of the wine press are called *eau de vie de marc* in France and *grappa* in Italy. Spanish brandies are usually distilled from sherry wines and have a distinctive flavor, quite different from cognac or armagnac. American brandies are primarily products of California and have a flavor different from the European brandies. Whereas in Europe pot stills are most common, continuous stills are preferred in California. Apple brandy is called applejack in the United States and is called calvados in France. It is distilled from completely fermented apple juice and aged in oak barrels for 5–10 years; it has a distinct apple flavor. Other fruits from which brandy is made include black, wild cherries (kirsch or kirschwasser), plums (slivovitz from Hungary or Rumania, and quetsch or mirabelle from France), blackberries, and apricots. When stone fruits are used, some of the stones are broken or crushed and a small amount of the oil is distilled over with the spirit, giving the brandy a more or less pronounced bitter almond flavor.

Whiskeys. These distilled spirits are made by distilling fermented grain mashes and aging the distillate in wood, usually oak. Examples are Scotch whiskey, Irish whiskey, Canadian whiskey, rye whiskey and bourbon whiskey.

Scotch whiskey is made primarily from barley. The barley is converted to malt by allowing it to sprout after which it is dried in a kiln over a peat fire. The malt absorbs some of the smoke aroma which is carried over later with the spirit distilled from it. The drying and roasting of the malt is to a large extent responsible for the flavor of the whiskey. After the malt is made into a mash, it is fermented, distilled, and aged in oak casks. These are often casks in which sherry has been shipped. A Scotch blended whiskey may contain a certain percentage of grain whiskey besides malt whiskey.

Irish whiskey is made from malt, unmalted barley, and other grains such as wheat, rye, and oats. The malt is not smoke-cured as in Scotland, and as a result the flavor of the resulting whiskey is different from that of Scotch whiskey.

Many types of whiskey are made in the United States. Depending on the principal source of grain, whiskeys are classified as bourbon (corn or grain), rye whiskey (rye), and others. Blended whiskeys are differentiated from straight whiskeys by a certain content of neutral spirits. American whiskeys are aged for a number of years in new charred white-oak barrels. In all types of whiskeys discussed above, the type and quality of the water used in the plant is an important factor in the quality of the finished product.

Gins. These consists essentially of a pure grade of alcohol with the addition of an extract of the juniper berry as the chief flavoring agent. The flavor is imparted by distillation of herbs (distilled gin) or by adding essential oils (a compounded gin). There are two principal types: English, or London dry, and Dutch (jenever).

English gin is made from pure neutral spirits, which are redistilled in the presence of juniper berries and small amounts of other ingredients such as coriander seed, cardamon seed, orange peel, anise seed, cassia bark, fennel, and others. English gin is not aged.

Dutch gin is made of malt. The flavoring ingredients are mixed directly in the mash, after which it is distilled at a rather low proof. Dutch gins are heavier in body and contain more fusel alcohols and other volatile compounds (congenerics) besides ethyl alcohol than does English gin. Some types of Dutch gin are aged. Dutch gin has a pronounced flavor of its own and is consumed straight, whereas English gin is usually used as an ingredient in cocktails.

Rum. The alcoholic distillate from fermented sugar cane juice or molasses is known as rum. Cuban rum is light-bodied and light-colored. The middle fraction of the distillation, known as aguardiente, is used for making rum. The aguardiente is aged in uncharred oak barrels; it is then decolorized, filtered, and supplied with some caramel to give it the proper color. Occasionally some fruit aroma is added. It is then aged for several more years. Jamaica rum is heavier-bodied and contains more congenerics (fusel alcohol, esters, and aldehydes) than Cuban rum. Arak is a rum that comes from the island of Java. It is a dry, highly aromatic rum. A natural fermentation of the molasses by various species of yeast also contributes to the flavor of this drink.

Vodka. A product originally produced in Russia but now popular in many countries, vodka is usually made from wheat. It is highly rectified during distillation and thus is a very pure neutral spirit without a pronounced taste. It is not aged.

Cachaça or pinga. This is Brazilian spirit made by distilling naturally fermented sugarcane juice in pot stills. It is high in congenerics and sold at various degrees of proof. It is usually not aged.

Pulque and tequila. Pulque and tequila are of Mexican origin. A sweet sap is obtained from agave (century plant and American aloe) by removing the growing bud from about 3-year-old plants. The sap (aguamiel or honey water) is collected and fermented by natural fermentation and is then called pulque. Distilled pulque is tequila.

Liqueurs. Liqueurs or cordials are alcoholic beverages prepared by combining a spirit, usually brandy, with certain flavorings and sugar. In fruit liqueurs, the color and flavor are obtained by an infusion process using the specified fruit and spirit. Sugar is added after the extraction is complete. Plant liqueurs

are made by maceration of plant leaves, seeds, or roots with spirit and then distilling the product. Sugar and coloring matter are added after distillation. A large variety of different liqueurs are marketed. *See* YEAST.

Emil M. Mrak; Herman J. Phaff

Bibliography. M. A. Amerine (ed.), *Wine Production Technology in the United States*, 1981; B. H. Gump and D. J. Pruett, *Beer and Wine Production: Analysis, Characterization, and Technological Advances*, 1994; J. S. Hough et al., *Malting and Brewing Science,* 2 vols., 1982; R. Hunderfund, *Wines, Spirits, and Fermentations*, 1983; H. J. Peppler and D. Perlman (eds.), *Microbial Technology*, vol. 2: *Fermentation*, 2d ed., 1979.

Distortion (electronic circuits)

The behavior of an electrical device or communications system whose output is not identical in form to the input signal. In a distortionless communications system, freedom from distortion implies that the output must be proportional to a delayed version of the input, requiring a constant-amplitude response and a phase characteristic that is a linear function of frequency. *See* RESPONSE.

In practice, all electrical systems will produce some degree of distortion. The art of design is to see that such distortion is maintained within acceptable bounds, while the signal is otherwise modified in the desired fashion. In analog systems operating in the audio band, major emphasis has been placed on satisfying the constant-amplitude characteristic since the human ear is insensitive to changes in phase. However, in digital systems, phase distortion can lead to unacceptable timing problems. In general, distortion can be grouped into four forms: amplitude (nonlinear), frequency, phase, and cross modulation.

Amplitude distortion. All electronic systems are inherently nonlinear unless the input signal is maintained at an incrementally small level. Once the signal level is increased, the effects of device nonlinearities become apparent as distorted output waveforms. Such distortion reduces the output voltage capability of operational amplifiers and limits the power available from power amplifiers. *See* AMPLIFIER; OPERATIONAL AMPLIFIER.

The load-line characteristic for a simple transistor amplifier is extremely helpful in predicting the output waveform for resistive loads if the input is sinusoidal in form. Excursions of the input signal about the quiescent operating point of the device are readily translated to output excursions. In this way the maximum input signal level can be determined for the avoidance of undesirable output distortion.

When the load is not resistive, however, or when more precise information is required, power series representation of the device is utilized. Thus, for example, the relationship between a transistor collector current i_c and its base current i_b may be expressed by Eq. (1). Thus, if the first two terms are used, and

$$i_c = G_1 i_b + G_2 i_b^2 + \cdots + G_n i_b^n + \cdots \qquad (1)$$

the input is sinusoidal, the output contains both the fundamental and a second harmonic. The degree of distortion can then be determined by comparing the relative amplitudes of the harmonic and fundamental terms. *See* HARMONIC (PERIODIC PHENOMENA); TRANSISTOR.

Thus, if the fundamental term has amplitude A_1 and the ith harmonic has amplitude A_i, harmonic distortion may be defined by Eq. (2). A measure of

$$D_i = \frac{A_i}{A_1} \qquad i \neq 1 \qquad (2)$$

the total harmonic distortion (THD) is then given by Eq. (3). This quantity is most commonly associated

$$D = \sqrt{D_2^2 + D_3^2 + \cdots} \qquad (3)$$

with an output voltage signal.

A common form of distortion in operational amplifier systems results from the slew-rate limitation of the amplifier. To conserve quiescent current, many operational amplifiers are designed to have a class B output stage capable of driving as much as 50 times its quiescent current. This puts demands on the intermediate drive stage, which is required to slew rapidly through a one- or two-diode voltage drop during transition from the push to pull mode, and vice versa. For slew rates beyond its capacity (usually arising because of an inability to drive the output junction capacitance), the output exhibits slew-rate limiting. This takes the form of straight-line segments in sinusoidal outputs, and can result in triangularization of the waveform if severe slew-rate limiting is encountered. For square waves, slew-rate limiting begins with finite transition rates between low and high states, with a triangular waveform occurring in extreme cases. For commercially available operational amplifiers, slew rates are typically in the range 0.5–20 V/microsecond and are a function of amplifier design geometry and fabrication method.

Frequency distortion. No practical device or system is capable of providing constant gain over an infinite frequency band. Hence, any nonsinusoidal input signal will encounter distortion since its various sinusoidal components will undergo unequal degrees of amplification.

The effects of such distortion can be minimized by designing transmission systems with a limited region of constant gain. Thus, in high-fidelity systems, the amplifier response is made wide enough to capture all the harmonic components to which the human ear is sensitive. In a communications system, the channel filters are designed to have a very sharp transition between regions of signal capture and signal rejection so that there is no interference between adjacent carrier bands.

Phase distortion. Since the time of propagation through a system varies with frequency, the output may differ in form from the input signal, even though the same frequency components exist. This can easily be demonstrated by noting the difference between the addition of two in-phase sine waves and

two whose phase relationship differs by several degrees.

In digital systems, such changes can be significant enough to cause timing problems. Hence, the phase–frequency response must be made linear to obtain distortionless transmission. In practice, this can be achieved by the introduction of delay equalizers whose characteristics are designed to compensate for phase distortion in other parts of the system without affecting the magnitude characteristic of the total system. *See* EQUALIZER.

Cross modulation. Sometimes referred to as intermodulation, this occurs because of the nonlinear nature of device characteristics. Thus, if two or more sinusoidal inputs are applied to a transistor, the output will contain not only the fundamentals but also signal harmonics, sums and differences of harmonics, and various sum or difference components of fundamental and harmonic components. While these effects are generally undesirable, they may be utilized to advantage in amplitude modulation and diode detection (demodulation). *See* AMPLITUDE-MODULATION DETECTOR; AMPLITUDE MODULATOR.

Reduction of distortion by negative feedback. Amplitude distortion may be reduced in amplifier stages by the application of negative feedback. If D is the distortion voltage output with feedback, and d is the distortion voltage without feedback, it can easily be shown that D is less than d by a factor of $(1 + A\beta)$, where A is the amplifier open-loop gain and β is the feedback factor. *See* FEEDBACK CIRCUIT.

F. William Stephenson

Bibliography. G. M. Glasford, *Analog Electric Circuits*, 1988; J. Millman and A. Grabel, *Microelectronics*, 2 ed., 1987; M. Schwartz, *Information Transmission, Modulation, and Noise*, 1990; F. W. Stephenson (ed.), *RC Active Filter Design Handbook*, 1985.

Distributed systems (computers)

A distributed system consists of a collection of autonomous computers linked by a computer network and equipped with distributed system software. This software enables computers to coordinate their activities and to share the resources of the system hardware, software, and data. Users of a distributed system should perceive a single, integrated computing facility even though it may be implemented by many computers in different locations. This is in contrast to a network, where the user is aware that there are several machines whose locations, storage replications, load balancing, and functionality are not transparent. Benefits of distributed systems include bridging geographic distances, improving performance and availability, maintaining autonomy, reducing cost, and allowing for interaction. *See* LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

Key Characteristics

The characteristics that are primarily responsible for the usefulness of distributed systems are resource sharing, openness, concurrency, scalability, fault tolerance, and transparency. System and application software must be carefully designed to ensure that these characteristics are attained.

Resource sharing. Resources include hardware such as disks and printers, and software such as files, databases, and compilers. Sharing reduces cost and is necessary to support groups of users working cooperatively. Resources are physically encapsulated within one or more of the computers in a distributed system and can be accessed from other computers only by communications. Each set of resources of a particular type must be managed by a program called a resource manager which offers a communication interface to enable accessing, manipulating, and updating the resource. The resource manager is also responsible for providing a naming scheme for each type of resource, for enabling individual resources to be accessed from any location, for mapping of resource names to communication addresses, and for synchronizing concurrent accesses to ensure consistency. *See* COMPUTER PERIPHERAL DEVICES; SOFTWARE.

Openness. This characteristic determines whether the distributed system can be extended in various ways without disruption or duplication of existing services. Openness is achieved by specifying and documenting the key software interfaces of a system and publishing them, thus making them available to software developers. UNIX is an example of a relatively open system design. The development of a standard interprocess communication (IPC) mechanism in BSD (Berkeley Software Distribution) UNIX and other operating systems and of standard communication protocols increased the scope for achieving openness; it became possible for processes on any computer and any operating system to make their resources accessible to other processes on other computers in the network.

Concurrency and parallelism. These characteristics arise in distributed systems due to the separate activities of users, the independence of resources, or the location of processes in separate computers. If there are M computers in a distributed system with one central processing unit each, then up to M processes can run in parallel (execute simultaneously). If several processes access the same resource concurrently, their actions must be synchronized to ensure that they do not conflict. *See* CONCURRENT PROCESSING.

Scalability. Distributed systems must operate effectively and efficiently at many different scales: from a workstation and a file server, to a local-area network with several hundred workstations and many file servers, to several local-area networks connected to form an internetwork. System and application software should not need to change when the scale of the system increases. Also, the work involved in processing any single request to access a shared resource should be independent of the size of the network. The design philosophy of distributed systems is that no single resource, hardware or software, is assumed to be in restricted supply; as demand for a resource

grows, it should be possible to extend the system to meet it.

Fault tolerance. When faults occur in hardware or software, programs may produce incorrect results or they may stop before they have completed the intended computation. In a distributed system, hardware that is essential to the continued operation of critical applications can be replicated. This hardware is exploited for noncritical activities when no faults are present. Software recovery involves the design of software so that permanent data (files and other material stored on permanent storage) can be recovered when a fault is detected. *See* FAULT-TOLERANT SYSTEMS.

Transparency. This is the concealment from the user and the application programmer of the separation of components in a distributed system so that the system is perceived as a whole rather than a collection of independent components. Some examples of transparency include access transparency, which enables local and remote information objects to be accessed using identical operations; and location transparency, which enables information objects to be accessed without knowledge of their location. Access and location transparencies are referred to together as network transparency.

Design Issues

Several important design issues arise specifically from the distributed nature of systems, including naming, communication, software structure, workload allocation, and consistency maintenance. There are also design issues unrelated to a system's distribution, including interface design and software engineering.

Naming. Distributed systems are based on the sharing of resources and on the transparency of resource distribution. Names assigned to resources must have global meanings that are independent of location, and must be supported by a name interpretation system that can translate names to enable programs to access the resources. The design issue here is to design a naming scheme that will scale and translate names efficiently to meet appropriate performance goals. A process that requires access to a resource that it does not manage must possess a name or an identifier for it; in this context, "name" refers to names that can be interpreted by users or programs, and "identifier" refers to names that are interpreted or used only by programs. An appropriate name space must be chosen for each resource type. A hierarchical name space, for example, "/etc/passwd," is potentially infinite and enables a system to grow, while a flat name space is usually finite, since its size is determined by the maximum permissible length of names.

To resolve names, a name service is needed. The name service accepts requests to translate names in one name space to names in another name space. It manages a database, which provides a view of the names existing in the distributed system and their translations in other name services. A resource name is resolved when it is translated into a form in which it can be used to invoke an action on the resource to which the name refers. Names are resolved relative to some context. For example, the name John Smith does not have much meaning unless some contextual information is provided, such as the name of the organization in which he works, or the street where he lives. In file systems, each directory represents a context. For example, "/etc/passwd" is a name with two parts. The first part, "etc," is resolved relative to the context "/" or root. The name "/oldetc/passwd" can have a different meaning because its second part is resolved in a different context.

In a distributed system, a resolved resource name is a communication identifier. For example, a communication identifier on the Internet consists of an IP address (for example, 192.135.231.4), which identifies the host, and a port number, which identifies a particular communication. The communication identifier must be acceptable to the communications subsystem and is used to transmit a request to a resource manager. The communications subsystem may have to perform further translations to produce network addresses and routing information acceptable to lower-level software layers. *See* INTERNET.

Communication. Since the components of a distributed system are both logically and physically separate, they must communicate in order to interact. Distributed systems and applications are composed of separate software components that interact to perform tasks. It will be assumed that components that require or provide access to resources are implemented as processes; that is, each component is active and has its own address space. Communication between a pair of processes involves the transfer of data from the environment of the sending process to the environment of the receiving process. In some communication operations, the synchronization of the receiving process with the sending process is required, so that the sending or receiving process is prevented from continuing until the other process makes an action that frees it.

The sending and receiving processes together perform message-passing actions. This involves the transmission of a set of data values (message) through a specified communication mechanism (port or channel) and the acceptance of the message by the receiving process. The mechanism may be synchronous, wherein the sender waits after transmitting a message until the receiver has performed a receive operation; or asynchronous, wherein the message is placed in a queue of messages waiting for the receiver to accept them and the sender can proceed immediately. In distributed systems, the main patterns of communication are client-server and group multicast.

Client-server. This form of communication is oriented toward service provision. An exchange consists of (1) transmission of a request from client to server; (2) execution of the request by the server; (3) transmission of the reply to the client. The server must become aware of the request sent as soon as it arrives, and the client must wait until the reply has been received. In its lifetime, a server serves many

clients and it need have no prior knowledge of them. Each request contains a communication identifier that is used to transmit the reply to the client. In an open system, clients cannot be initialized to hold the communication identifiers of all the servers they want to access. Typically, when a server starts up, it registers itself with a name service, stating its network address and a well-known name for the service it provides. Clients obtain the server's address by interrogating the name service using the well-known name. *See* CLIENT-SERVER SYSTEM.

Group multicast. The target of this type of message is a group of processes. To a single group send operation, there corresponds a receive performed by each member of the group. Tasks performed by group multicast include:

1. Locating an object. A client multicasts a message containing the name of a file directory to a group of file servers. Only the one that holds the relevant directory replies.

2. Fault tolerance. A client multicasts a request to a group of servers, all of which process the request identically. As long as at least one of the servers does not fail, the service will be provided.

3. Multiple update. An event such as time can be multicast to a group.

Software structure. Each category of software corresponds to a different level of abstraction of the distributed system, and each offers a different degree of transparency to the user. A programmer may write in a distributed programming language, and use the underlying services provided. Besides applications, the main categories of software in distributed systems are the following.

Operating system kernel. This encompasses services that provide basic resource management, mainly memory allocation and protection, process creation, processor scheduling, interprocess communication, and peripheral device handling. It presents a set of universal problem-oriented programming abstractions, such as processes, to applications, shielding them from the underlying details of the processor and other hardware.

Open services. These bring the programming facilities of a distributed system up to the level required for application programming. New services can be developed and installed as required, enabling distributed systems to be adapted to the needs of many different types of applications. Services range from a basic file service to an interactive multimedia conferencing service.

Distributed programming support. Such software provides run-time support for language facilities, such as remote procedure calls (RPC) and group multicast communication, that allow programs written in conventional languages to work together in a distributed system environment.

Workload allocation. This addresses how work is allocated among resources in a distributed system.

Workstation-server model. The advantage of this model is putting the processor cycles near the user, especially for highly interactive applications. Processor performance and memory capacity of the workstation determine the size of the largest task that can be performed on behalf of the user; a single user with a large computing task is not able to obtain additional resources.

Processor pool model. This is a modification of the workstation-server model to include processors that can be allocated to users' tasks dynamically. A processor pool consists of a collection of inexpensive computers, each consisting of a processor, memory, and network interface. Each processor in the pool has an independent network connection. Processors do not have to be homogeneous; a system might include chip sets from different manufacturers in the pool, to broaden the range of software that can be run. Processors are allocated to processes for their lifetimes.

Use of idle workstations. This is another variation on the workstation-server model. At any one time, and especially overnight, a significant proportion of workstations on a network may be unused or be used for lightweight activities such as document editing. These stations have spare processing capacity, and can be used to run jobs for users who are logged on at other stations and do not have sufficient capacity at their machines.

Shared-memory multiprocessors. These contain several independent processors, each able to execute a separate program. Each processor interfaces to the shared memory via cache memory. A large number of central processing units can be deployed. This processing capacity is particularly suited to applications in areas such as virtual reality and multimedia. *See* MULTIMEDIA TECHNOLOGY; MULTIPROCESSING; VIRTUAL REALITY.

Consistency maintenance. This encompasses issues of consistency with regard to updates, replication, caches, failures, and user interfaces.

Update consistency. This issue arises when several processes access and update data concurrently. The activity of changing a related set of data values cannot be performed instantaneously. The updates should appear to be atomic; that is, a related set of changes made by a given process should appear to all other processes as though it was instantaneous. This is a significant issue in distributed systems because many users are likely to access shared data. Moreover, the operation of the system itself depends on the consistency of file directories managed by file services, naming databases, and so forth. *See* DATABASE MANAGEMENT SYSTEM.

Replication consistency. If data derived from a single source have been copied to several computers and subsequently modified at one or more of them, the possibility of inconsistencies arises between the values of data items at different computers. The consistency of replicas is generally ensured by periodic updates.

Cache consistency. This refers to the problem that arises when data values that have been cached by one client are updated by another. For example, the same blocks of a file stored in a file server are requested and cached at several workstations. When a client updates a block at one of the workstations,

the version held in the file server and those held in cache at other workstations are said to be stale. The client may send the updated block to the file server responsible for the file, but the other workstations will retain the out-of-date version unless the server notifies them.

Failure consistency. When one component in a distributed system fails, the others will normally continue to run. But if they depend on the continued operation of the failed computer or program, they may fail later. Each of the cooperating components involved in a failed computation may have reached a different point in its program. To ensure that the permanent data stored by all of the components remains consistent, recovery procedures are needed to cause the data to be "rolled back" to a known state after a failure.

User interface consistency. Whenever a user performs an input action such as a key depression or a mouse click in an interactive program, the screen becomes temporarily inconsistent with the user's model of the program's state. The program then processes the user's input and performs the necessary operations to update the screen. Unless the input is processed and the changes are transmitted to the screen quickly enough to give the user the impression of an instantaneous change, the user becomes aware of the screen inconsistency. In distributed systems, the transfer rates and the latency characteristics of networks have an impact on user interface consistency. *See* HUMAN-COMPUTER INTERACTION.

Object-Oriented Distributed Systems

The object-oriented model for a distributed system is based on the model supported by object-oriented programming languages. Distributed object systems generally provide remote method invocation (RMI) in an object-oriented programming language together with operating systems support for object sharing and persistence. Remote procedure calls, which are used in client-server communication, are replaced by remote method invocation in distributed object systems. *See* OBJECT-ORIENTED PROGRAMMING.

The state of an object consists of the values of its instance variables. In the object-oriented paradigm, the state of a program is partitioned into separate parts, each of which is associated with an object. Since object-based programs are logically partitioned, the physical distribution of objects into different processes or computers in a distributed system is a natural extension. The Object Management Group's Common Object Request Broker (CORBA) is a widely used standard for distributed object systems. Other object management systems include the Open Software Foundation's Distributed Computing Environment (DCE) and Microsoft's Distributed Common Object Manager (DCOM).

Common Object Request Broker. CORBA specifies a system that provides interoperability among objects in a heterogeneous, distributed environment in a way that is transparent to the programmer. Its design is based on the Object Management Group's object model.

This model defines common object semantics for specifying the externally visible characteristics of objects in a standard and implementation-independent way. In this model, clients request services from objects (which will also be called servers) through a well-defined interface. This interface is specified in Object Management Group Interface Definition Language (IDL). The request is an event, and it carries information including an operation, the object reference of the service provider, and actual parameters (if any). The object reference is a name that defines an object reliably.

The central component of CORBA is the object request broker (ORB). It encompasses the entire communication infrastructure necessary to identify and locate objects, handle connection management, and deliver data. In general, the object request broker is not required to be a single component; it is simply defined by its interfaces. The core is the most crucial part of the object request broker; it is responsible for communication of requests.

The basic functionality provided by the object request broker consists of passing the requests from clients to the object implementations on which they are invoked. In order to make a request, the client can communicate with the ORB core through the Interface Definition Language stub or through the dynamic invocation interface (DII). The stub represents the mapping between the language of implementation of the client and the ORB core. Thus the client can be written in any language as long as the implementation of the object request broker supports this mapping. The ORB core then transfers the request to the object implementation which receives the request as an up-call through either an Interface Definition Language (IDL) skeleton (which represents the object interface at the server side and works with the client stub) or a dynamic skeleton (a skeleton with multiple interfaces).

Architectural components. The communication between the object implementation and the ORB core is effected by the Object Adapter (OA). It handles services such as generation and interpretation of object references, method invocation, security of interactions, object and implementation activation and deactivation, mapping references corresponding to object implementations, and registration of implementations. It is expected that there will be many different special-purpose object adapters to fulfill the needs of specific systems (for example, databases).

The Object Management Group specifies four policies in which the object adapter may handle object implementation activation: shared server policy, in which multiple objects may be implemented in the same program; unshared server policy; server-per-method policy, in which a new server is started each time a request is received; and persistent server policy. Only in the persistent server policy is the object's implementation supposed to be constantly active (if it is not, a system exception results). If a request is

invoked under any other policy, the object will be activated by the OA in a policy-specific way. In order to be able to do that, the OA needs to have access to information about the object's location and operating environment. The database containing this information is called the implementation repository and is a standard component of the CORBA architecture. The information is obtained from there by the OA at the time of object activation. The implementation repository may also contain other information pertaining to the implementation of servers, such as debugging, version, and administrative information.

The interfaces to objects can be specified in two ways: either in IDL or through their addition to the interface repository, another component of the architecture. The dynamic invocation interface allows the client to specify requests to objects whose definition and interface are unknown at the client's compile time. In order to use the dynamic invocation interface, the client has to compose a request (in a way common to all object request brokers) including the object reference, the operation, and a list of parameters. These specifications—of objects and services they provide—are retrieved from the interface repository, a database that provides persistent storage of object interface definitions. The interface repository also contains information about types of parameters, certain debugging information, and so forth.

A server-side analogous to the dynamic invocation interface is the dynamic skeleton interface (DSI); with the use of this interface the operation is no longer accessed through an operation-specific skeleton, generated from an IDL specification. Instead it is reached through an interface that provides access to the operation name and parameters. (As in the dynamic invocation interface, the information can be retrieved from the interface repository). Thus the DSI is a way to deliver requests from the object request broker to an object implementation that does not have compile-time knowledge of the object it is implementing.

Interoperability. Many different ORB products are currently available; this diversity is very wholesome since it allows the vendors to gear their products toward the specific needs of their operational environment. It also creates the need for different object request brokers to interoperate. Furthermore, there are distributed and client-server systems that are not CORBA-compliant, and there is a growing need to provide interoperability between those systems and CORBA. In order to answer these needs, the Object Management Group has formulated the ORB interoperability architecture.

Implementation differences are not the only barrier that separates objects; other reasons for such a separation might include strict enforcement of security, or providing a protected testing environment for a product under development. In order to provide a fully interoperable environment, all those differences have to be taken into account. CORBA 2.0 introduces the higher-level concept of a domain, which roughly denotes a set of objects that, for some reason, be it implementation or administrative, are separated from some or all other objects. Thus, objects from different domains need some bridging mechanism (mapping between domains) in order to interact. Furthermore, this bridging mechanism should be flexible enough to accommodate both the scenarios where very little or no translation is needed (as in crossing different administrative domains within the same object request broker) and scenarios which can be less efficient but where there is the need to provide general access to object request brokers.

The interoperability approaches can be divided into mediated and immediate bridging. With mediated bridging, interacting elements of one domain are transformed at the boundary of each domain between the internal form specific to this domain and some other form mutually agreed on by the domains. This common form could be either standard (specified by the Object Management Group, for example, Internet Inter-ORB Protocol or IIOP), or a private agreement between the two parties. With immediate bridging, elements of interaction are transformed directly between the internal form of one domain and the other. The second solution has the potential to be much faster, but is the less general one; it therefore should be possible to use both. Furthermore, if the mediation is internal to one execution environment (for example, TCP/IP), it is known as a full bridge; otherwise, if the execution environment of one object request broker is different from the common protocol, each object request broker is said to be a half bridge.

Implementations of CORBA. Many implementations of CORBA are available. They vary in the degree of CORBA compliance, quality of support, portability, and availability of additional features.

Mohamed Y. Eltoweissy

Bibliography. G. Blair and J. B. Stefani, *Open Distributed Processing and Multimedia*, Addison-Wesley, 1998; G. Coulouris, J. Dollimore, and T. Kindberg, *Distributed Systems: Concepts and Design*, 3d ed., Addison-Wesley, 2000; S. Mullender (ed.), *Distributed Systems*, 2d ed., Addison-Wesley, 1993; O. Spaniol, C. Linhoff-Popien, and B. Meyer (eds.), *Trends in Distributed Systems: CORBA and Beyond*, Springer, 1996.

Distributed systems (control systems)

Collections of modules, each with its own specific function, interconnected to carry out integrated data acquisition and control in a critical environment.

Evolution of industrial control systems. Industrial control systems have evolved from totally analog systems through centralized digital computer-based systems to multilevel, distributed systems.

Analog and supervisory control. Originally, industrial control systems were entirely analog, with each individual process variable controlled by a single feedback

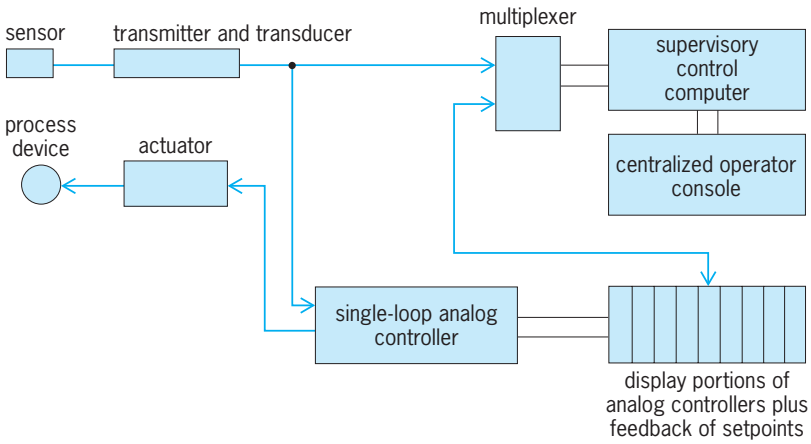

Fig. 1. Supervisory control system combining analog control with digital data acquisition and feedback of parameters to controllers.

controller. Although analog control systems were simple and reliable, they lacked integrated information displays for the process operator.

In supervisory control (**Fig. 1**), the analog portion of the system is implemented in a traditional manner (including analog display in the central operating room), but a digital computer is added which periodically scans, digitizes, and inputs process variables to the computer. The computer is used to filter the data, compute trends, generate specialized displays, plot curves, and compute unmeasurable quantities of interest such as efficiency or quality measures. Once such data are available, optimal operation of the process may be computed and implemented by using the computer to output set-point values to the analog controllers. This mode of control is called supervisory control because the computer itself is not directly involved in the dynamic feedback. *See* DIGITAL COMPUTER.

Direct digital control. Direct digital control replaces the analog control with a periodically executed equivalent digital control algorithm carried out in the central digital computer (**Fig. 2**). A direct digital control system periodically scans and digitizes process variables and calculates the change required in the manipulated variable to reduce the difference between the set point and the process variable to zero. *See* ALGORITHM.

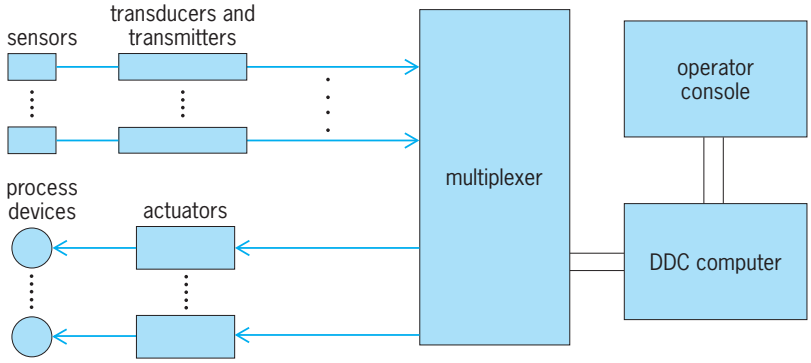

Fig. 2. Direct digital control configuration.

The advantages of direct digital control are the ease with which complex dynamic control functions can be carried out (compensation for time delay, nonlinear functions, interacting control loops) and the elimination of the cost of the analog controllers themselves. To maintain the attractive display associated with analog control systems, the display portion of the analog controller is usually provided (set-point stations). Thus operation of the process is identical to operation of digital supervisory control systems with analog controllers, except that "tuning" of the controllers (setting of gains) can be done through operator consoles.

The cost reduction which resulted from the introduction of direct digital control was offset by a number of disadvantages. The most notable of these were the decrease in reliability and the total loss of graceful degradation. Failure of a sensor or transmitter had the same effect as before, but failure of the computer itself threw the entire control system into manual operation. Hence it was necessary to provide analog controllers to back up certain critical loops which had to function even when the computer was down. As a result of this need for extra equipment, the industry tended to employ a combination of direct digital control and supervisory control, using the former only where the digital computer could perform control actions which could not be easily handled by analog controllers or where the failure of the digital computer could not jeopardize operation of the process.

Increasing demand for ever-higher levels of supervisory control, coupled with increasing installation cost, highlighted two disadvantages of centralized digital computer control of processes. First, process signals were still being transmitted from the process sensor to the central control room in analog form, meaning that separate wires had to be installed for every signal going to or from the computer. Cost of installing this cable is very high. Second, the digital computer system itself evolved into a very complex unit because of the number of devices attached to the computer and because of the variety of different programs needed to carry out the myriad control and management functions. The latter resulted in the need for an elaborate real-time operating system for the computer which could handle resources, achieve desired response time for each task, and be responsible for error detection and error recovery in a highly dynamic real-time environment. Design coding, installation, and checkout of centralized digital control systems was so costly and time-consuming that application of centralized digital control was limited.

Distributed control. Low-cost electronic hardware utilizing large-scale integrated circuits provided the technology to solve both of these problems while retaining the advantages of centralized direct digital and supervisory control. The solution, distributed control (**Fig. 3**), involved distributing control functions among hardware modules to eliminate the critical central computer. *See* INTEGRATED CIRCUITS.

Each module utilizes microprocessors in order to

carry out its function. Coupling the modules together is a digital communication system, making it feasible to geographically distribute the modules around the plant in order to minimize installation costs. This is one of the major motivations for installing such systems, since those modules which must input or output process signals may be located in the vicinity of the sensors and actuators associated with those signals and hence require minimum length of analog signal lines. Coupling modules via the digital communication system means that only one (or at most a few) communication lines need be installed around the plant. Communication links are often called "data highways" to emphasize their specialized nature and high speed. The distributed control still permits centralized collection, display, and interaction with the control system and its application through placement of operator consoles in central control rooms. *See* MICROCOMPUTER; MICROPROCESSOR.

A second motivation is the reduced design complexity of control systems, especially that of the real-time operating systems and tasks which carry out the application. The mixture of high- and low-priority tasks in one computer is replaced by tasks in separate computers or microprocessors.

A third motivation is the return of the characteristic of graceful degradation in the face of component failures. In a distributed system, loss of any module can be limited to loss of the function supplied by that module. If that function is not critical, the system merely continues without the function in a degraded performance mode. If that function is critical, it may be backed up by a duplicate module or by dynamically assigning its function to another module when the failure is detected. In any case, there is no longer the dependence on any one module for successful operation of the system, nor the need for complex and expensive backup of that module.

Multihighway distributed systems. The success of distributed control led to the demand for ever larger control systems which take the form of tightly interconnected distributed systems (**Fig. 4**). A large system by modern standards might include 10,000 or more individual sensors and as many as 30,000 computed variables. The need for timely communication of data in such large systems makes the data communication system (the data highways) very critical. Individual highways are called local highways and generally are associated with a specific unit or area of the application. The supervisory highway is usually associated with overall management of the application or plant. Its nodes are commonly operator consoles and computer interfaces, and are often in a central control room associated with the overall plant.

Module functions. A distributed control system (Fig. 3) is made up of tightly interconnected modules, each with its own specific function.

Data acquisition. A data acquisition module accepts some number of analog input signals, scans and digitizes them at arbitrary rates, performs alarming and smoothing calculations, and transmits the reduced digital data to other modules in the system for direct or supervisory control purposes. Binary (on/off) signals may also be acquired.

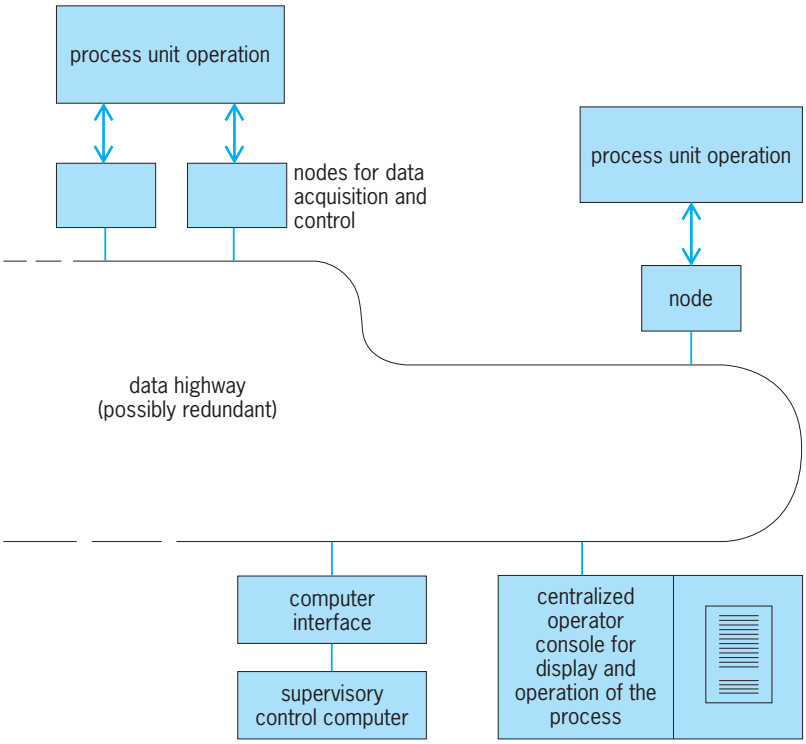

Fig. 3. Distributed data acquisition and control system with geographical distribution of nodes.

The timing of acquisition of a signal's value is often very critical to the application. For example, the signal might have to be acquired periodically with very small variation of the periods between signal value acquisitions, as in direct digital control.

Direct digital controllers. A direct digital controller performs exactly the same functions on one or a few loops as those carried out in a direct digital control central computer. These include digitizing, smoothing, alarming, control calculations, output of actuator signals, handling of transitions from manual to feedback control, and error recovery. Typical digital

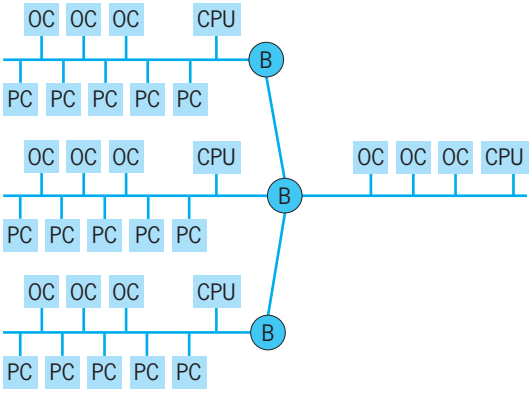

Fig. 4. Multihighway distributed system. PC = process or programmable controller module; OC = operator console; CPU = computer (central processing unit); B = bridge.

controllers permit the realization of more complex control algorithms than those available in comparable analog controllers, thereby retaining the advantage of the central direct digital control computer. The controller periodically transmits set-point, process signal, and controller status information to other modules for supervisory control purposes.

The success of direct digital control is dependent upon the periodic acquisition of the controlled variable, and this implies that the communication protocol must be able to provide this near-exact periodicity. Two protocols are used. First, the direct-digital-control application program may run periodically and, each time that data are needed, send a request message to the data acquisition devices on the network. These devices in turn acquire the data and return the value. In order for this to happen fast enough that the periodic control task in not disturbed, proprietary protocols are often used to tight-couple the devices and the control computer. For example, the request may hold the connecting bus until the device puts its data on the bus, thereby avoiding the overhead of a separate return message.

Second, a more recent protocol called Fieldbus uses a local microprocessor to handle a small number of data acquisition devices (100 or less) but allows the control computer to subscribe to the signal values at prespecified periodic intervals. The microprocessor then compensates for the communication delays by anticipating the need to send the data, and thereby initiates the data message without being requested each interval. This requires the control program to be carefully timed so that it is ready to receive and accept the data.

Operator consoles. An operator console usually has televisionlike screen displays and special functionally oriented keyboards which permit the process operator both to input commands to the control system which are equivalent to those input to integrated control systems, and to display values of process variables, groups of process variables, process status, and a variety of reports. In effect, the operator consoles function identically to those used with centralized digital control systems, except that any data required for a display or the result of any operator command must be transmitted to or from the appropriate modules in the distributed system.

Distributed control systems have expanded in size and scope. Operator consoles have undergone the greatest change. Operators must have available overviews of the entire plant and processes within the plant and be able to focus on sections of the processes to respond to alarms and other events. As a consequence, operator consoles are multiscreen-based devices, and the trend has been toward flexible organization of the screens and less dedication of individual screens to specific functions. A group of screens might be organized so that they behave like a single very large screen for overview purposes. This wall-like display facilitates cooperation among operators and managers. When the need arises to concentrate on an individual section of the process, one or more of the screens may be controlled by an operator to display this more specific information. *See* ELECTRONIC DISPLAY.

Operator consoles may display very large numbers of process variables, and hence the communication protocol usually cannot involve requesting values one at a time. Rather, the control usually subscribes to groups of data which are then passed to the console as a single large message. This group-subscription protocol is very effective in minimizing loading of the networks and thereby allowing timely acquisition of the data values.

Data logging, trending, and archiving. The display of a plot of a process variable over a previous time interval (minutes, hours, or days) is an important capability for a human operator. A data logging and trending module collects data for a number of points and stores it in a form which can be called up by an operator for display at any time. It is important to be able to archive such data for future uses, including planning operations, demonstrating correct operation of the process in defense of a legal claim, and operator training. The large volume of data to be collected makes it useful to assign this function to a separate module, often called a historian. Operator consoles must be able to recall archived data from the historian and display it in a flexible manner. Supervisory and other general-purpose computers also need access to the archived data.

Discrete programmable controllers. Discrete controls are appropriate in sequencing and batch process applications. The programmable controller is a device designed specifically for such applications and is most widely used in discrete manufacturing applications. The trend toward larger process control systems has led to the inclusion of discrete functions and the need for programmable controller functions. These may be implemented in a separate module or included as building-block algorithms in controller modules. Communication to this type of control module usually involves a block of data. Hence, communication is essentially the exchange of messages without critical timing requirements such as in periodic data acquisition. *See* PROGRAMMABLE CONTROLLERS.

Structure of nodes. In order to achieve the advantages of standard modules (that is, standard functions, standard programs, graceful degradation, and high reliability), it is convenient to modularize as much as possible and to restrict the functions carried out by each node to standard functions.

In a typical node (**Fig. 5**), a local communication system (usually a parallel bus) is provided which is independent of the serial communication links that tie nodes in the overall system together. Such a bus is very high speed (because it is physically short and because digital data are transferred in parallel) and separate so that another level of reliability can be built into the system. That is, individual modules tie to this bus, which in turn links to the external communication network by separate communication controllers. External links may be made redundant simply by providing more than one link and more than one communication controller per node.

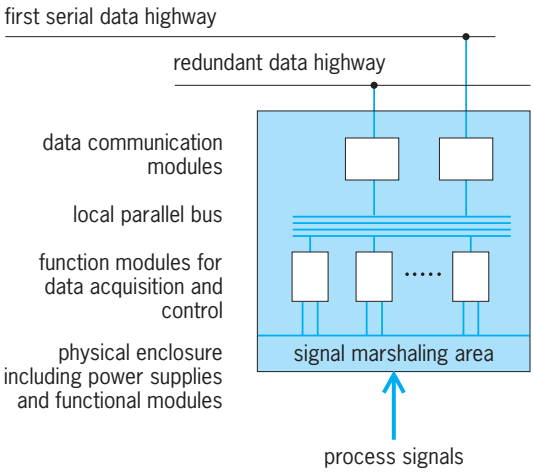

Fig. 5. Structure of one node in a distributed system showing internal functional modules.

Internally, functional modules are manufactured as separate printed circuit cards, and multiple cards are used to realize the required number of functions in a node. Hence failures on one printed circuit board do not necessarily destroy the function supplied by the node, nor does failure of one external communication link destroy the integrated system capability. Failure of all external communication links still leaves the functions within the node operating. Therefore graceful degradation can be extended to several levels. This capacity is extended even further by providing redundant power sources and power supplies.

Some manufacturers have even designed systems in which modules within a node are totally independent of one another from a functional point of view and interact only in that they use the internal local bus to communicate. An example might be a node which supports 32 control loops and is realized by four separate controller modules, each with its own input/output and power supply, and each communicating with the rest of the system via the local bus and the data communication modules, which themselves are independent and tie to the same local bus. Hence, many combinations of failures are possible which merely degrade performance rather than cause catastrophic breakdown.

Error detection and recovery within a controller module is often provided so that critical functions are not lost because of a single component failure. Error recovery requires switching to a backup module of the same type. It is necessary, however, that the current data being used by the control algorithms be available. Consequently, a node may provide a backup memory into which the process database in that node is periodically copied. After failure detection, the database is updated from the backup memory and the backup controller module is switched to replace the failed module. The process of snapshotting the process database and switching to backup hardware components, called fail-over, takes approximately 0.25 s from time of failure detection.

Serial data communication systems. The replacement of an integrated system by a set of interconnected modules does place one stringent requirement on a distributed system, namely that the communication network be fast enough to support all the demands of all modules for communication. For this reason, redundant communication links are often provided.

Data highways may be on the order of a mile (1.6 km) in length. Consequently, they are susceptible to errors produced by noise. Although the error rate is low, undetected errors cannot be tolerated because of the critical nature of the application. Hence all communication involves an exchange of messages specially formulated with check fields to ensure that all errors are detected. Errors are corrected by retransmission of the messages. The set of communication procedures and formats is termed a protocol.

Structure. The major distinguishing characteristic among communication protocols is the handling of multiple modules attempting to use the data highway simultaneously. The most commonly used protocols are ethernet (arbitrary initiation of message transmission when a quiet highway is detected, but stopping transmission when simultaneous transmission by two nodes is detected); active repeaters at each node (ring networks) so that simultaneous transmission may take place between each pair of nodes; and voluntary passing of mastership from node to node (token passing) in a logical ring structure. Although some smaller systems allow modules to communicate only in response to an invitation from a central computer (polling), this protocol does not provide adequate throughput and robustness for large systems. These differing philosophies lead to different response times to critical events.

Manufacturing application protocol. One communication protocol, the manufacturing application protocol (MAP), is being widely supported for control systems in discrete manufacturing. This standard attempts to make nodes and computers from different manufacturers compatible.

Multilevel systems. At the same time, there is increasing demand for very high levels of management in control applications. Examples include scheduling of generation loads, optimization of yield, and minimization of energy used for processing. These needs imply larger computers cooperating in the control process, and they are causing the structure of the modern control system to develop into a complex multilevel system (**Fig. 6**). Here the system consists of a set of processing units, each of which must be carefully monitored so that the overall plant control may be integrated. Distributed control systems are associated with individual process units. Each involves operator communication, supervisory control, and so forth.

A network of computers carrying out the overall plant high-level control couples the distributed systems together. The major distinction in the communication systems is that the higher-level computers do not acquire data or control physical process variables directly. Rather they communicate with the

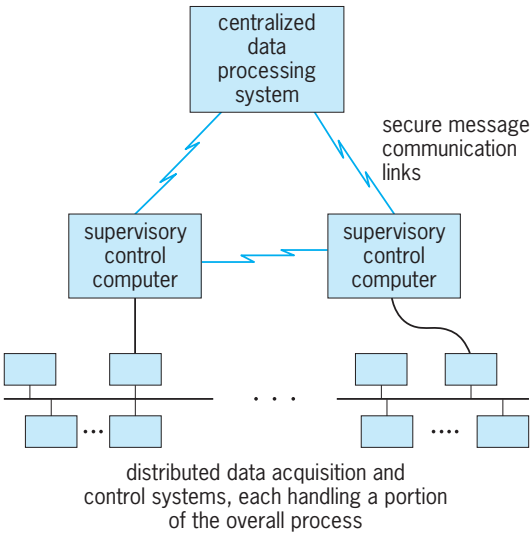

Fig. 6. Multilevel system with computers interconnected in a distributed network and using distributed data highway systems for direct communication with the process.

supervisory computers in the individual distributed control systems, exchanging messages to obtain data, parameter values, operating sequences, and so forth. *See* LOCAL-AREA NETWORKS.

Requirements for communication among the higher-level computers are very different from those for the data highways. Messages tend to be much longer (carrying a large amount of data) but require correspondingly slower response time. Hence such a network communication protocol is designed to be very general and error-free but not necessarily high-speed. This permits standards to be established which then result in low-cost hardware and software.

However, higher-level business-oriented applications need access to the real-time data collected and computed within the distributed control systems. Furthermore, large plants often have several different distributed control systems for different processes within the plants. The higher-level applications not only must access real-time data from different distributed control systems but also must combine the data in a higher-level application whose results are then used to manage the lower-level control systems. As a consequence, the preference is to require distributed control systems to be "open" rather than proprietary. Furthermore, the demand for real-time data necessitates a significant increase in data communication rates of distributed systems, requiring the use of new technology. Open standards for communication among the levels such as DCOM (distributed component object model) and CORBA (common object request broker architecture) are emerging.

Object-oriented concepts. Successful applications demand integration of data from a variety of sources and processing of the data in a flexible manner. These requirements impose severe demands upon software used in distributed control systems. In particular, this software must often be created specifically for the applications of a given plant or process. Object-oriented technology carries the promise of a building-block approach to the assembly of even very complex software. Objects are software artifacts specifically designed to encapsulate all the data and processing associated with a real-world entity. Control systems, process units, alarms, and other physically meaningful entities can be modeled as objects in software. Higher-level application software then uses these objects without being required to do the extensive software programming required to create them. This approach makes complex applications technically feasible to create and economical.

Standard interfaces. The success of object-oriented technology in distributed control systems is dependent upon standards for objects. These standard interfaces permit application software to use objects no matter where they reside in the system (physical distribution) or in what computer language they were programmed. They promise to make access to real-time data from distributed control systems as easily available as common applications on personal computers.

Object-based communication protocols. Object-based communication has emerged as a basis for communication between separate software modules either in the same computer or in separate computers. The DCOM and CORBA technologies allow the remote access of objects, called components. Each object contains some data, and some functions for manipulating that data. The overall system can be constructed by combining a given set of components, and may be updated and expanded by the addition of new components as the system needs them. The communication protocol views the module using a component no differently from an internal function. In fact, the data involved are packed into messages and transmitted across the underlying network in a standard manner which is opaque to the communicating modules.

A widely used example of object communication is the OPC (OLE for Process Control) specification, which defines a set of objects and their methods for performing data acquisition. In order to make data within a proprietary control system available to higher-level enterprise applications, a manufacturer may implement the OPC objects. The enterprise application then simply uses object-based communication to acquire needed data without any regard for communication protocols or the details of the implementation of the objects, which remains proprietary.

The combination of reliable, responsive distributed control and general-purpose communication networks leads to a system which can be adapted to critical control applications in a very flexible manner, with potential for increased productivity in plants, increased safety, and decreased energy consumption. Technology for higher-speed computation, data communication, and object-oriented software organization allows the integration of distributed control systems into plant-wide and enterprise-wise systems. *See* CONTROL SYSTEMS;

DIGITAL CONTROL; MULTILEVEL CONTROL THEORY; OBJECT-ORIENTED PROGRAMMING.

James D. Schoeffler

Bibliography. D. Chappell, *Understanding ActiveX and OLE*, Microsoft Press, 1996; M. P. Lukas, *Distributed Control Systems: Their Evaluation and Design*, Van Nostrand Reinhold, New York, 1986; *OLE for Process Control*, OPC Foundation, Austin, 1997; *OPC Data Access Automation Interface Specification*, OPC Foundation, Austin, TX, 1999; R. Orfali, D. Harkey, and J. Edwards, *The Essential Distributed Objects Survival Guide*, Wiley, New York, 1996; C. Ottino, User interface in the 1990's, *ISA Trans.*, 32:95–121, 1993; R. Probst, A software model for the 21st century, *Object Mag.*, 3:65–67, 1993.

Distribution (probability)

The results of a series of independent trials, random variables, or errors often occur in fairly regular and predictable patterns. These patterns can be expressed mathematically, and the most important are called the binomial, normal, and Poisson distributions.

Binomial distribution. Consider n independent trials, each of which results in success S or failure F, with corresponding probabilities p and $q = 1 - p$. Denote by S_n the number of successes. Because there are possible ways to select k places for S and $n - k$ places for F, the probability distribution of the random variable S_n is given by

$$\binom{n}{k}$$

where $k = 0, 1, \ldots, n$. This is the binomial distribution. Its expectation is np, its variance npq. *See* PROBABILITY.

If one lets a random variable X_k equal 1 or 0 according as the kth trial results in S or F, then $S_n = X_1 + \cdots + X_n$. The binomial distribution can therefore be approximated by the normal distribution in accordance with the central limit theorem. This special case is known as the DeMoivre-Laplace theorem; setting

$$x_k = (k - np)(npq)^{-1/2}$$

it asserts that

$$P\{S_n = k\} \approx (2\pi)^{-1/2} \exp(-x_k{}^2/2)$$

$$P\{a < S_n < b\} \approx (2\pi)^{-1/2} \int_{x_{a-1/2}}^{x_{b+1/2}} \exp(-t^2/2)\, dt$$

with a percentage error tending to 0 as $n \to \infty$ provided x_a and x_b are restricted to a fixed interval. (With $p = q = 1/2$ these formulas are useful for the evaluation of binomial coefficients and their sums.) *See* BINOMIAL THEOREM.

Normal distribution. The standard normal density is defined for $-\infty < x < \infty$ by

$$\phi(x) = (2\pi)^{-1/2} e^{-x^2/2}$$

The standard normal distribution function or error function $\Phi(x)$ is its integral from $-\infty$ to x. The normal distribution function with mean m and variance s^2 is

$$\Phi[(x - m)/s]$$

and its density is

$$(2\pi)^{1/2} s^{-1} e^{-(x-m)^2/(2s^2)}$$

As x goes from $-\infty$ to ∞, the function increases from 0 to 1. It plays an important role in many fields. In particular,

$$u(t, x; \xi) = 2\pi^{-1/2} t^{-1/2}\, e^{-(x-\xi)^2/4t}$$

is the fundamental solution of the heat (or diffusion) equation $u_t = u_{xx}$ and represents the heat distribution on the x axis at time t caused by a unit heat source initially concentrated at $x = \xi$. Probabilistically, this represents the transition probabilities in the Wiener process.

A random variable whose distribution is normal is called normal or gaussian. The sum of two independent gaussian variables is again gaussian; hereby the means and variances add. Analytically this means that the convolution of two normal distributions is again normal. This property characterizes the normal distributions among distributions with finite variances.

Central limit theorem and error theory. These are the best-known and most important applications of the normal distribution. The nature of the central limit theorem is best seen from the simplest special case: If $X_1, X_2, \ldots$ are independent random variables having a common distribution with mean m and variance s^2, their sum $S_n = X_1 + \cdots + X_n$ has mean $\mu = nm$ and variance $\sigma^2 = ns^2$; the corresponding standardized variable $S_n^* = (S_n - \mu)\sigma^{-1}$ has a probability distribution $P\{S_n^* \le x\}$ tending to $\Phi(x)$ as $n \to \infty$; in other words, the distribution of S_n gets close to $\Phi[(x - \mu)\sigma^{-1}]$.

The striking feature of this result is that an essential property of a sum of many independent components is independent of the nature of these components. The general central limit theorem shows this to be true under much wider conditions: The distribution of the sum S_n tends to normality even if the distributions of the components X_k vary with k, provided only that the components are likely to be of the same order of magnitude (so that no component has individually a noticeable effect on the sum). It is not even necessary that the X_k be independent.

Many empirical quantities (for example, the amount of water in a reservoir, certain inherited characteristics such as height, and the experimental error of physical measurements) represent the cumulative effect of many small components, and the statistical fluctuations of such quantities may be expected to follow the normal distribution. In particular, under the authority of K. F. Gauss, it has been assumed for a long time that experimental errors are approximately normally distributed. Modern research has shown the limitations of this assertion. Many distributions appear to the untrained eye as

nearly normal, but refined statistical tests prove that even the finest physical measurements depart considerably from normality, and that assignable causes (that is, large-contribution components) can be discovered statistically. This discovery shows the questionable character of the classical methods to predict the probable experimental error. In industrial quality control the departures from normality are used efficiently to discover assignable causes and thus to spot coming trouble at an early stage. *See* QUALITY CONTROL.

Misinterpretation of theory. Much harm has been caused by the widespread misinterpretation of the limit theorems in probability and of the meaning of statistical equilibrium in stochastic processes. The situation may be explained in terms of a coin-tossing game in which H and T count 1 and -1, respectively. Here $X_k = \pm 1$ with probability 1/2, and $m = 0$, $s^2 = 1$. The operational meaning of the central limit theorem in this case is as follows. Fix a large n. Repeat the same game of n tossings independently many times. One would then expect that in about 50% of the cases $S_n > 0$, in about 25% of the cases $S_n > 0.67n^{1/2}$, in about 16% of the cases $S_n > 2n^{1/2}$, and so on. What the central limit does not say is that in one game about half of the sums $S_1, \ldots, S_n$ will be positive. In fact, the arcsine law shows the opposite to be true: It is much more probable that all $S_j > 0$ than that there be equally many positive and negative ones.

Multivariate normal distribution. The previous theory carries over without essential changes to n dimensions. The n-dimensional normal density is defined by $(2\pi)^{-n/2}D^{1/2}e^{-Q(x_1,\ldots,x_n)/2}$, where Q is a positive definite quadratic form with determinant D; the matrix of variances and covariances is the reciprocal of the matrix of Q. If the n-dimensional joint distribution of the random variables $X_1, \ldots, X_n$ is normal, then each X_j is normal. The converse assertion is false, even though found in textbooks. The multivariate normal distribution is important for stationary stochastic processes. *See* STOCHASTIC PROCESS.

Poisson distribution. The Poisson distribution with parameter λ is the probability distribution of a random variable assuming the values $0, 1, 2, \ldots$ with probabilities $p_k(\lambda) = e^{-\lambda}\lambda^k/k!$. Both its expectation and variance equal λ. This is one of the most important distributions; it plays a basic role in the theory of stochastic processes and in many applications. A full understanding of its character can be obtained only from a postulational derivation and from the consideration of its many generalizations. However, much can be gained by the following elementary approach starting from the binomial distribution.

Consider a large number n of independent trials, each of which results in success or failure with probabilities p and $q = 1 - p$. Ordinarily, interest is restricted to the case where p is very small, but the expected number of successes $np = \lambda$ is of moderate size. Typical examples may be obtained by considering centenarians, color-blind people, or triplets in a large population; the defectives in an allotment of screws or fuses; or the wrong calls among all calls arriving during a day at a busy telephone. The number of successes in the n trials is a random variable with the binomial distribution, but under the present circumstances the binomial is close to the Poisson distribution and may be replaced by it. In fact, the probability of no success is $q^n = (1 - \lambda/n)^n$, which is close to $e^{-\lambda}$, the first term of the Poisson distribution. Now the ratio of the kth to the $(k - 1)$th term in the binomial distribution approaches that of the Poisson distribution.

This reasoning explains why the statistical fluctuations in the phenomena cited above follow approximately the Poisson distribution. In other circumstances the Poisson distribution appears, not as an approximation, but as the exact expression of a law of nature. This is true in particular of processes where certain events, such as radioactive disintegrations, mutations, power failures, and accidents, occur in time in such a way that (1) the probability that an event occurs during any given time interval of length dt is, asymptotically, λdt; and (2) there is no interaction or aftereffect between nonoverlapping time intervals. Under these assumptions a time interval of length t can be divided into $n = t/dt$ subintervals, each representing a trial in which the probability of success is close to λdt; a change to the limit $dt \to 0$ gives the Poisson expression $p_k(\lambda t)$ as exact probability for the occurrence of k events during a time interval of length t. A similar argument applies to random distributions of points in space, with t interpreted as volume; typical examples are stars, flaws of material, raisins in a cake, and animal litters in a field. Thus "perfect randomness" of a chance aggregate of points in space or time usually means that the fluctuations obey the Poisson law. *See* ANALYSIS OF VARIANCE; STATISTICS.

William Feller

Bibliography. E. H. Callaway, Jr., D. A. Hill, and J. V. Scholle, *Probability Distributions*, 1993; W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1, 3d ed., 1968, reprint 1991; vol. 2, 2d ed., 1971; H. J. Larson, *Introduction to Probability Theory*, 5th ed., 1995; W. Mendenhall and R. J. Beaver, *Introduction to Probability and Statistics*, 10th ed., 1998; S. M. Ross, *Introduction to Probability Theory*, 5th ed., 1994.

District heating

The supply of heat, either in the form of steam or hot water, from a central source to a group of buildings. As an industry, district heating began in the early 1900s with distribution of exhaust steam, previously wasted to the atmosphere, from power plants located close to business and industrial districts. Advent of condensing-type electrical generating plants and development of long-distance electrical transmission led to concentration of electrical generation in large plants remote from business districts. Most district heating systems in the United States rely on separate steam generation facilities close to load centers. In some cities, notably New York, high-pressure district steam (over 120 lb/in.^{2} or 830 kilopascals) is used extensively to feed turbines that drive pumps and refrigerant compressors. Although some district heating plants serve detached residences, the cost

of underground piping and the small quantities of heating service required make this service generally unfeasible.

District heating, apart from utility-supplied systems in downtown areas of cities, is accepted as efficient practice in many prominent colleges, universities, and government complexes. Utility systems charge their customers by metering either the steam supply or the condensed steam returned from the heating surfaces.

District heating has grown faster in Europe than in the United States. A substantial part of the heat distribution is by hot water from plants combining power and heat generation. Heat from turbine exhaust or direct boiler steam is used to generate high-temperature (270–375°F or 132–191°C) water, often in cascade heat convertors, where live steam is mixed with return hot water. Hot water distribution systems have the advantage of high thermal mass (hence, storage for peak periods) and freedom from problems of condensate return. On the other hand, pumping costs are high and two mains are required, supply and return. (The cost of condensate return piping and pumping is so high that many steam distributors in the United States make no provision for condensate return at all.)

Reykjavik, Iceland, is completely heated from a central hot water system supplied from hot springs. Average water temperature is 170°F (77°C).

The most extensive European developments in district heating have been in Russia, mainly in the form of combination steam-electric plants serving new towns. The Russians developed a single-pipe high-temperature water district heating system. At each building served, superheated water from the main is diluted with return water from the building system. Excess water is used in kitchens, laundries, bathrooms, and swimming pools.

An important extension of the principle of district heating was the development of district heating and cooling plants. Usually owned and operated by the natural gas utilities, central plants distribute chilled water for air conditioning in parallel conduits with hot water or steam for heating. The system serving Hartford, Connecticut, is the pioneer installation. Because the costs of these four-pipe distribution systems are extremely high, development is slow and restricted to small areas of high-density usage within a few blocks of the plant. *See* COGENERATION; COMFORT HEATING; HOT-WATER HEATING SYSTEM.

Richard L. Koral

Bibliography. American Society of Heating, Refrigerating, and Air Conditioning Engineers, *Applications Handbook*, 1993; American Society of Heating, Refrigerating, and Air Conditioning Engineers, *Systems Handbook*, 1992.

Diving

Skin diving, or diving without the aid of special equipment, is a technique to investigate underwater environments and to gather shellfish or other items of commercial importance. In breath-hold diving, fins and faceplate are usually employed to facilitate the diver's activities. For longer periods of time, divers can function by using air pumped to them from the surface. This restricts the area that can be explored since movement of the diver is limited by the length of the air hose. Diving with scuba (self-contained underwater breathing apparatus) provides an almost limitless amount of freedom. With this equipment, the amount of time the diver can stay beneath the surface and the distance traveled are limited only by the volume of compressed air in the diver's tanks or aqua-lung.

Techniques

Humans have used compressed-air diving as a tool in marine research, salvage, and construction for over two and a half centuries. The history of diving is filled with crippling accidents and death, which are frequently the result of excessive pressure and abnormal gas ratios. One of the earliest records of diving and its problems is recorded on a plaque at the entrance to Old Port Royal, Jamaica, West Indies: "Lucas Barrett, F.G.S; L.L.S.—1837–1862—Died while conducting underwater research off Port Royal on December 19, 1862—25 years old—Geological Society of Jamaica."

Fortunately, efficient training programs, a much better understanding of diving physiology, and reliable equipment have removed many of the early dangers of diving. Diving is now an effective and safe means of conducting marine research. Although some scientific diving was accomplished prior to 1940, it was not until scuba (such as J. Y. Cousteau's aqualung) became commercially available in 1950 as inexpensive equipment for the sportsperson that diving became a widespread means of conducting marine research.

All modern methods of diving deliver a breathable gas mixture to the diver's lungs and body cavities at the same pressure as that exerted by the overlying water column. Thus there is no difference in pressure between the interior and exterior parts of the diver's body, and the diver has no feeling of pressure from the overlying water, regardless of depth. The method of supply and the duration of submergence varies with the type of equipment. Equipment can be divided into two basic types: free diving (scuba), where the diver carries a supply of compressed gas on the back; and tethered diving (hard hat), with a gas supply furnished to the diver via a hose from the surface.

Scuba diving. Scuba is extensively used by trained personnel as a tool for direct observation in marine research and underwater engineering. This equipment is designed to deliver through a demand-type regulator a breathable gas mixture at the same pressure as that exerted on the diver by the overlying water column. The gas which is breathed is carried in high-pressure cylinders (at starting pressures of 2000–3000 lb/in.^{2} or 14–21 megapascals) worn on the back. The scuba diver, free from surface tending, has much more maneuverability and freedom than the conventional hard-hat diver.

Fig. 1. A diver wearing a mixed-gas scuba, exposure suit, and face mask. Depth 300 ft (90 m) on the upper lip of Scripps Submarine Canyon, La Jolla, California. (*U.S. Navy photograph, Naval Electronics Laboratory Center*)

Fig. 2. Underwater scientific investigation by scuba diving, using compressed air. (*U.S. Navy photograph, Naval Electronics Laboratory Center*)

Scuba can be divided into three types: closed-circuit, semiclosed-circuit, and open-circuit. In the first two, which use pure oxygen or various combinations of oxygen, helium, and nitrogen, exhaled gas is retained and passed through a canister containing a carbon dioxide absorbent for purification. It is then recirculated to a bag worn by the diver (**Fig. 1**). During inhalation additional gas is supplied to the bag by various automatic devices from the high-pressure cylinders. These two types of equipment are much more efficient than the open-circuit system which does not take advantage of the unused oxygen in the exhaled gas. The exhaled gas is discharged directly into the water after breathing. The closed- and semiclosed-circuit types are, however, more complicated and can be dangerous if they malfunction or are used by inexperienced divers. Most open-circuit systems use compressed air because it is relatively inexpensive and easy to obtain. Although open-circuit scuba is not as efficient as the other types, it is preferred by most scientists and engineers who work underwater because of its safety, the ease in learning its use, and its relatively low cost.

Free diving or scuba diving (**Fig. 2**) using compressed air as the breathing gas is used to some extent in most current marine scientific and engineering investigations in all environments, including the Arctic. For physiological reasons the underwater observer is limited to about 150 ft (50 m) of water depth when obtaining observations which require thinking clearly and solving problems while submerged. Below this depth when using compressed air as a breathing gas, the diver is limited, not by equipment, but by the complex temporary changes which take place in the body chemistry while breathing gas (air) under high pressure. The practical depth limit to diving with air in clear, warm water is about 240 ft (80 m).

Tables containing permissible rates of ascent and time allowable at different depths are available from the U.S. Navy Experimental Diving Unit, Washington, D.C., and from the Manned Undersea Science and Technology Office of the National Oceanic and Atmospheric Administration (NOAA), Rockville, Maryland. Tables are also available through most commercial diving supply outlets. Breathing a mixture of helium, instead of nitrogen, with oxygen can increase the diver's ability to remain safely underwater for longer periods of time, but because of the expense and difficulties of handling the specialized diving gear and breathing mixtures needed for this type of deep diving, it is not practical for most scuba use.

By using scuba, scientists are no longer restricted to the sea surface but are able to investigate underwater problems at first hand. Free diving or scuba diving is used as a tool in offshore mapping of geologic formations, biological studies in ecology, underwater engineering, and detailed oceanographic measurements.

Saturation diving. Saturation diving permits long periods of submergence (1–2 weeks). It allows the diver to take advantage of the fact that at a given

Fig. 3. Sealab II, an underwater station for deep-sea divers; 57 ft (17 m) long, 12 ft (3.7 m) in diameter. (*U.S. Navy*)

depth the body will become fully saturated with the breathing gas and then, no matter how long the submergence period, the decompression time needed to return to the surface will not be increased. Using this method, the diver can live on the bottom and make detailed measurements and observations, and work with no ill effects. This type of diving requires longer periods of decompression in specially designed chambers to free the diver's body of the high concentration of breathing gas. Decompression times of days or weeks (the time increases with depth) are common on deep dives of over 180 ft (60 m). However, when compared with diving from the surface, this type of diving has been shown to be much more efficient since it permits approximately three times the amount of work to be done for the amount of time spent on a project. Furthermore, the marine fauna become accustomed to seeing the divers living among them, and behavioral studies are greatly improved. While saturation diving is an exciting method of diving that greatly extends the time spent underwater, the long periods of time needed to decompress and the limited number of underwater laboratories capable of supporting saturated divers have kept it from becoming a popular tool for the average scientist and engineer. Its primary use has been in the servicing of offshore oil-drilling rigs, in maintenance and repair of dams, and in scientific studies in research projects, such as Hydro-Lab, Tektite, and Sea Lab (**Fig. 3**). Factors such as cold water at depth, lack of light, the requirement for bulky and expensive equipment, and the presence of strong bottom currents place a practical limit to this type of diving at about 640 ft (195 m). *See* UNDERWATER VEHICLES.

Robert F. Dill

Physiology

Environmental effects on the submerged diver are quite different from those experienced at sea level. Two elements are very evident during the dive. As the depth of surrounding water increases, pressure on the air the diver breathes also increases. In addition, as the pressure increases, the solubility of the gases in the diver's tissues increases. The tissues, therefore, accumulate certain gases which are not metabolized. The increased presence of certain gases causes specific and often dangerous physiological effects.

Pressure. The total pressure of the atmosphere at sea level is approximately 760 mmHg or 30.4 in. Hg.

Pressure increases underwater at the rate of 1 atm (10 kilopascals) for each 33 ft (10 m) that the diver descends. The total pressure applied to the body and to the breathing gas increases proportionately with depth. As pressure of the gas increases, the amount of gas that is absorbed by the body increases. This is particularly evident if the diver is breathing air within a caisson since the percentage of nitrogen in the breathing gas increases proportionately to the amount of oxygen that is removed. Likewise, the amount of carbon dioxide in the body increases during the dive, particularly if the exhaled air is not separated from the inhaled air.

According to Dalton's law, the pressure exerted by an individual gas in a mixture (the partial pressure) is independent of the pressure of the other gases. The percentage or proportion of individual gases in air is always the same at sea level, during a dive, or at high altitudes. In other words, if the percentage of oxygen is 20.8% at sea level, it will also be 20.8% at any depth in the dive. However, if the diver is breathing air at a partial pressure of 100 mmHg (4 in. Hg) and descends 132 ft (40 m), which is approximately 4 atm of pressure, the partial pressure of oxygen increases to over 400 mmHg (16 in. Hg). The increase in pressure thus causes an increase in the effects of specific gases as well as exertion of a direct force on the body of the organism. In cases where the diver descends while wearing a helmet or scuba gear, the pressure within the diver's lungs must be maintained at a high level in order to prevent the pressure of the surrounding water from collapsing the thorax. In a caisson, the pressure exerted within the lungs is increased with increased depth. This is due to the increased pressure created by the surrounding water.

Effects of specific gases. One of the more obvious effects of gases on divers is caused by nitrogen. This gas makes up about 78% of the air that is normally breathed, and its solubility in the tissues increases as atmospheric pressure increases. When nitrogen is dissolved in the body, more than 50% is contained in the fatty tissues; this includes the myelin sheaths which surround many nerve cells. When divers undergo increased pressure, amounts of nitrogen in nerve tissue increase and lead to nitrogen narcosis or "rapture of the deep." Nitrogen narcosis can occur at 415 ft (130 m) or 5 atm (500 kPa), and increases in severity as the diver descends below this depth. The irrational behavior and euphoria often seen in nitrogen narcosis can result in serious, even deadly mistakes during a dive. The maximum time that divers can remain underwater without showing symptoms decreases with increasing depth. For instance, at 130 ft (40 m) the maximum time a diver can safely remain submerged in scuba gear is 15 min; in more shallow dives, the time becomes much longer. Obviously, nitrogen narcosis is one of the dangers of diving simply because the irrational behavior associated with it cannot be tolerated underwater.

One of the earlier recognized problems associated with human diving is known as decompression sickness, the bends, or caisson disease. If a diver is allowed to stay beneath the surface for long periods of time, the volume of dissolved gases in the tissues will increase. This is particularly true of nitrogen in the case of air breathing. When nitrogen accumulates in the tissues, it remains in solution as long as the pressure remains constant. However, when pressure decreases during the ascent, bubbles form in the tissues. Nitrogen bubbles can occur in nerves or muscles and cause pain, or they can occur in the spinal cord or brain and result in paralysis, dizziness, blindness, or even unconsciousness. Bubbles forming in the circulatory system result in air embolism. If the embolism occurs in the circulation of the lungs, a condition known as the chokes occurs. The first symptom noticed by the individual is pain beneath the sternum, followed by rapid shallow breathing; this ultimately results in pulmonary edema and can cause death. This condition does not occur if the nitrogen in the breathed air is replaced with hydrogen or helium. Because of its explosive qualities, hydrogen is avoided, but helium can be safely employed. Because of its increased mobility, helium moves out of the tissues rapidly and does not readily build up in them as bubbles. *See* DECOMPRESSION ILLNESS.

A method of prevention of decompression sickness was suggested by J. S. Haldane in 1907. Haldane introduced the method of stage decompression, in which the diver is allowed to ascend a few feet and then remain at this level until the gases in the tissues have been allowed to reequilibrate at the new pressure. This stepwise ascent is continued until the diver finally reaches the surface. A modern variation of this method consists of placing the diver in a decompression chamber after the surface is reached, to allow for periods of decompression which simulate ocean depths.

Oxygen toxicity is caused by increased concentrations of oxygen in the breathing gas. It can cause hallucinations, mental confusion, muscle twitching, nausea, and vertigo. When breathing pure oxygen, a diver is usually safe at 40 ft (12 m) for only 23 min and at 20 ft (6 m) for 1.5 h. The cause of oxygen toxicity is unknown. However, it leads to convulsions and coma, which are lethal if the diver is not brought to the surface immediately. The problem can be reversed once the diver reaches the surface and is given an adequate breathing mixture.

Carbon dioxide intoxication has not been a significant problem. However, in scuba diving it can occur occasionally. Carbon dioxide intoxication causes headaches, vertigo, and increased muscular activity. It has also been known to cause increased breathing rates and body warmth. Severe carbon dioxide intoxication can cause a diver to black out or lose consciousness even in shallow water; the depressant effects occur in the central nervous system without any previous warning signs. *See* RESPIRATION.

J. Homer Ferguson

Bibliography. J. Lippman, *The Essentials of Deeper Sport Diving*, 1992.

Diving animals

Various species of diving, air-breathing animals that can be found in every class of terrestrial vertebrates, having almost complete freedom in the aquatic habitat through several specialized physiological adaptations. Modern technology has allowed the investigation of many mechanisms that permit diving animals to spend long periods of time submerged. These investigations include studies to determine the physiological changes that occur during the dive. Truly aquatic animals have no need for a specialized diving response, because they are able to obtain oxygen directly from the water and thus sustain themselves indefinitely in the water. Terrestrial animals, on the other hand, have in many cases gone back in varying degrees to the aquatic mode of living. In all of these, the problem of supplying oxygen and nutrients to the tissues while submerged with only the physical attributes of a terrestrial animal has been a major specialization.

Physiological characteristics. Diving animals have been observed by scientists for many years. In 1870, Paul Bert measured the blood volume in birds and mammals relative to their breath-holding capabilities. He found that the blood volume in the bodies of these diving species was much greater than that found in nondiving animals. The larger proportion of blood to body volume gives these animals the ability to store larger amounts of oxygen. More oxygen allows the diving animals to remain submerged for longer periods of time than the nondiving species. The times and depths of dives of some species are shown in the **table**.

In many cases, animals undergo pronounced physiological changes when they dive. Above the surface of the water, they have all of the physiological characteristics of truly air-breathing organisms. When they dive, certain changes occur in their bodies which allow them to remain submerged for several minutes to an hour or more. One physiological change noted early was reduction in the rate at which the heart beats. This reduction, or bradycardia, is dramatic in the species which are well adapted to diving (see **illus.**). In the beaver, which has a normal resting heart rate of 140 beats per minute, the rate falls to less than 20 beats per minute after submergence.

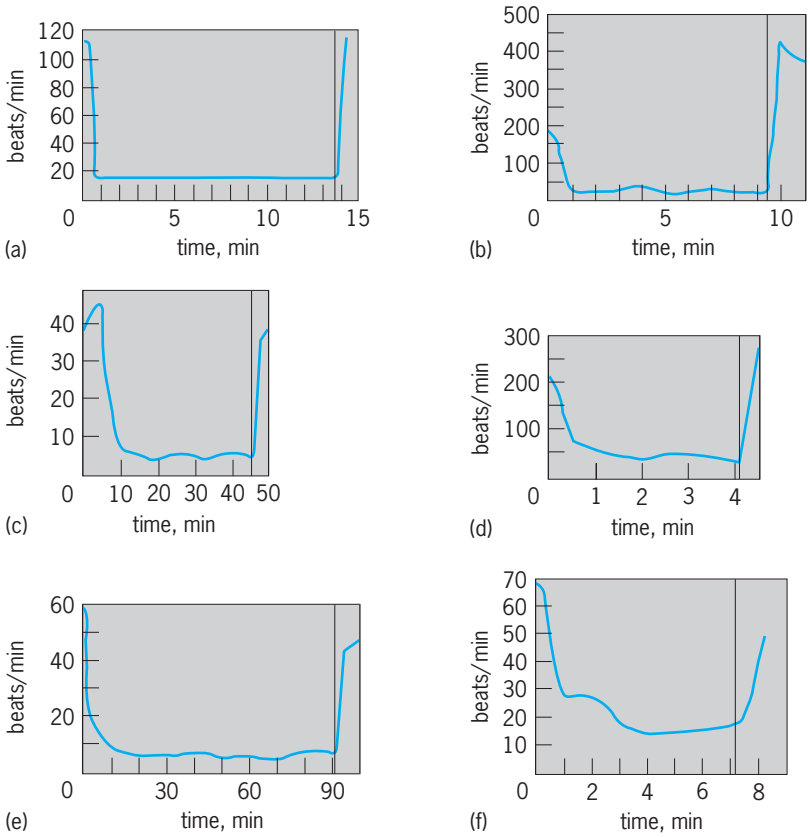

Changes in heart rate in several species during diving: (*a*) hair seal; (*b*) duck; (*c*) alligator; (*d*) muskrat; (*e*) common water snake; and (*f*) common garter snake.

Bradycardia is the result of a specific autonomic reflex which controls heart rate as well as the flow of blood to the muscles and internal organs. During the dive, blood vessels that supply skeletal muscles, kidneys, viscera, and the periphery of the body are constricted. Blood flows occurs primarily to the brain, lungs, and heart muscle while the diving animal is submerged, and blood flow to the peripheral organs is reduced or completely restricted. In this condition the animals derive energy solely from stored metabolites and oxygen. This reflex prevents lactic acid (lactate) from entering the circulation. Lactic acid is a metabolic by-product of the Embden-Meyerhof pathway, or anaerobic glycolysis. Since lactate remains in the tissues, the likelihood of its stimulating chemoreceptors found in the larger blood vessels is reduced. *See* CARBOHYDRATE METABOLISM; CHEMORECEPTION.

Another characteristic of diving animals is the capacity to store large quantities of glycogen, adenosine triphosphate (ATP), and oxygen in their muscles. Much of the oxygen used by the peripheral organs of the body is stored by means of a compound

Duration and depth of diving reported for various species

Species	Duration, min	Depth, ft (m)
Mink *(Mustela vision)*	3	—
Muskrat *(Ondatra zibethicus)*	12	—
Walrus *(Odobenus rosmarus)*	10	—
Gray seal *(Halichoerus grypus)*	20	330 (100)
Sperm whale *(Physeter catodon)*	75	2950 (900)
Common garter snake *(Thamnophis sirtalis):* forced dive	7	—
Common water snake *(Natrix sipedon):* forced dive	90	—

called myoglobin. Myoglobin is higher in concentration in the muscles of diving animals than in those of the nondiving ones. *See* MUSCLE PROTEINS; MUSCULAR SYSTEM.

Some diving animals have an enhanced ability to generate energy via anaerobic glycolysis. The enzymes involved in anaerobic glycolysis, such as lactic acid dehydrogenase and phosphokinase, are in greater concentration in the organs of diving animals than in the organs of nondiving species. Since they have greater amounts of these enzymes in their tissues, diving animals can carry out anaerobic glycolysis much more rapidly and efficiently than the nondiving species.

Oxygen debt develops when an animal utilizes glucose anaerobically. When this occurs, glucose is broken down to pyruvic acid through several intermediate stages which yield energy in the form of ATP. Pyruvic acid (pyruvate) is converted to lactic acid in the absence of oxygen. Lactate can then be converted back to pyruvate when oxygen is available. The accumulation of lactic acid in the tissues is associated with oxygen debt. When oxygen is obtainable again through breathing, the animal is said to repay its oxygen debt. Lactic acid dehydrogenase is an enzyme which is essential for the conversion of pyruvic acid to lactic acid in the absence of oxygen. It is also capable of catalyzing the reverse of this reaction when oxygen is supplied to the tissues. The presence of elevated concentrations of lactic acid dehydrogenase in the organs of diving animals allows them to repay oxygen debt more rapidly after surfacing.

All of the chemical reactions carried out in an organism constitute its metabolism. Where metabolism can be reduced, the dependence upon oxygen can also be reduced. It has been shown that in diving animals the metabolic rate is slowed when they are submerged. This is indicated by a decrease in body temperature which is 4–6°F (2–3°C) below normal during the dive. Such a drop in body temperature would result from a 25% decrese in total metabolism.

During the dive, insulation of the body core is increased as the result of a reduction in blood flow to the periphery. This effectively reduces the amount of heat exchanged between the core of the body and the environment. This is essential in reducing total metabolism. *See* CARDIOVASCULAR SYSTEM.

Diving species. The diving response has been shown in amphibians, reptiles, birds, and mammals and consists of several complex physiological changes which allow the animals to remain submerged for long periods of time.

Mammals. Among the mammals, the mink and otters, members of the family Mustelidae, are excellent divers but normally do not remain beneath the surface for more than a few minutes. The muskrat (*Ondatra zibethicus*) has been known to remain submerged for up to 12 min; however, the depth to which it dives is not definitely known. Another excellent and well-known fresh-water diver among the mammals is the beaver (*Castor canadensis*). The better divers among mammals are found in the seals, which are usually located in the colder waters of the Arctic and Antarctic. The Weddell seal, which is found in the Antarctic, is a champion among divers, going to depths of 2000 ft (600 m) and remaining submerged for up to 43 min. This is quite remarkable since the pressure of the water increases up to 883 lb/in.^{2} (609 kilopascals) at this depth. In order to sustain a dive at this depth, the organism must be able to withstand temperatures as low as 37°F (2°C). The seals are especially useful in diving experiments. In addition to their excellent diving capabilities, they are readily adaptable to captivity and can be trained to dive at the investigator's direction.

Of all the aquatic and marine mammals, whales are probably the best adapted to diving. The sperm whale (*Physeter catadon*) has been known to remain submerged for as long as 75 min at a depth of 2950 ft (900 m). The blue whale *(Sibbaldus musculus)* can dive for 49 min and reaches a depth of 330 ft (100 m). One of the factors which contributes to the depths to which whales can dive is related to their ability to collapse their lungs. Whales force the air expelled from the lungs into the trachea, which is expanded in these species and which is impermeable enough to prevent gaseous exchange with the tissues. This eliminates the possibility of absorbing high concentrations of nitrogen from the air in the lungs during the dive. Whales also have phenomenal thermoregulatory mechanisms which allow them to exist in both arctic and temperate-zone waters. *See* MAMMALIA.

Birds. A wide variety of birds are known to be excellent divers. This includes birds such as the dipper and the osprey, which are submerged for only a short period of time while feeding. On the other hand, many bird species such as the penguin can remain submerged for a long period of time. They have lost the ability to fly and are well adapted to the aquatic environment. Almost all of the ducks and geese are divers and can remain submerged from a few minutes to nearly 20 min. This group includes the grebes and loons, which are web-footed and have short wings that enable them to be excellent swimmers. *See* AVES.

Reptiles. Many reptiles are also excellent divers. During the voyage of the *Beagle*, Darwin observed the diving iguanas which live on the Galápagos Islands, off the coast of Ecuador. These lizards spend much of their lives in the aquatic habitat. Diving reptiles are also found among snake species. These include the completely marine snake (*Hydrophis belcheri*) and those which are only partially aquatic, such as the common garter snake *(Thamnophis sirtalis)* and the common water snake *(Natrix sipedon)*. Other divers among the reptiles are the alligator, crocodile, and cayman. These animals can remain submerged for prolonged periods of time partially because of their lower metabolic rate and thus lower demand for respiratory oxygen. Other good examples of divers in the class Reptilia include the turtles, both aquatic and marine. These animals may stay submerged for hours or even days depending upon the season. *See* REPTILIA.

Amphibians. The amphibians are also excellent divers and can remain submerged for long periods. All amphibians should be included in the category of divers, since they are all partially aquatic and must become submerged from time to time in their life cycle. *See* AMPHIBIA; DIVING. J. Homer Ferguson

Bibliography. P. J. Butler and D. R. Jones, The comparative physiology of diving invertebrates, *Adv. Comp. Physiol. Biochem.*, 8:179–364, 1982; G. K. Snyder, Respiratory adaptations in diving mammals, *Respir. Physiol.*, 54:269–294, 1983.

Division

In arithmetic and algebra, the process of finding one of two factors of a number (or polynomial) when their product and one of the factors are given. The symbol ÷ now used mostly in elementary English and American arithmetics to denote division first appeared in print in an algebra by J. H. Rahn published in Zurich in 1659. Division is more often symbolized by the double dot :, the bar —, or the solidus /; thus $x:y$, or x/y indicates division of a number x by a number y. Considered as an operation inverse to multiplication, x/y is a symbol denoting a number whose product with y is x. Another way to base division upon multiplication is provided by the concept of the reciprocal of a number. If y is any number (real or complex) other than 0, there is a number, denoted by $1/y$ and called the reciprocal of y, whose product with y is 1. Then x/y is the symbol for the product of x and $1/y$. This view of division furnishes a means of extending the concept to objects other than real or complex numbers. A whole number is divisible by 2 if its last digit is so divisible, and by 4 if the number formed by the last two digits is so divisible. It is divisible by 3 or 9 according to whether the sum of its digits is thus divisible, respectively, and is divisible by 11 if the difference between the sum of the digits in the odd and the even places can be so divided. *See* ADDITION; ALGEBRA; MULTIPLICATION; NUMBERING SYSTEMS; NUMBER THEORY; SUBTRACTION. Leonard M. Blumenthal

DNA helicase

Any of a ubiquitous class of enzymes that catalyze the unwinding of double-stranded deoxyribonucleic acid (dsDNA) into single-stranded DNA (ssDNA). In all organisms (other than some viruses), genetic information is locked within a double helix formed by two antiparallel DNA strands. Although dsDNA is suitable for secure information storage, hydrogen bonds formed between complementary bases (Watson-Crick base pairing) impair readout of this information by the cellular machineries that frequently require ssDNA as the template. The unwinding of dsDNA into ssDNA that is catalyzed by DNA helicases is a function critical for virtually every aspect of cellular DNA metabolism including replication, recombination, and repair.

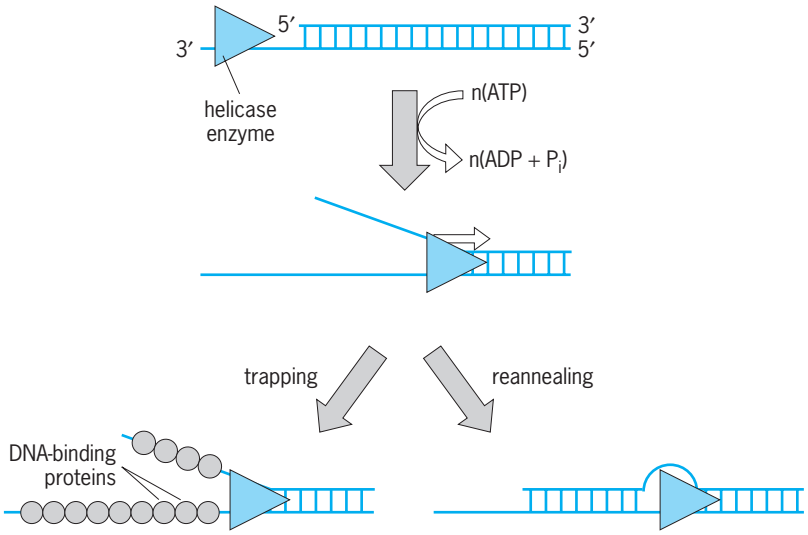

Fig. 1. Schematic representation of DNA helicase action. The helicase enzyme translocates along the DNA molecule and separates the strands. Energy for this unfavorable reaction is coupled to the hydrolysis of adenosine triphosphate (ATP) to adenosine diphosphate (ADP) and inorganic phosphate (P_i). In the presence of a single-stranded DNA-binding protein the nascent ssDNA strands are trapped and reannealing of the DNA duplex is prevented. The helicase depicted here displays a 3′ → 5′ polarity, tracking unidirectionally along the lower of the two DNA strands in the duplex (the loading strand).

First identified in the 1970s, DNA helicases are motor proteins that convert chemical energy into mechanical work. Chemical energy is derived from the hydrolysis of adenosine triphosphate (ATP) or other nucleoside triphosphates, and is coupled to mechanical work during at least two important steps within the helicase reaction cycle (**Fig. 1**): (1) the unidirectional translocation (movement) of the enzyme along the DNA molecule, and (2) the separation (unwinding) of the DNA duplex (that is, the disruption of hydrogen bonds between base pairs). *See* ADENOSINE TRIPHOSPHATE (ATP); DEOXYRIBONUCLEIC ACID (DNA).

Classifications. On the basis of amino acid sequence comparisons, helicases are divided into three large superfamilies and two smaller families that are illustrated schematically in **Fig. 2**. Moreover, these analyses identified several protein signature motifs (short conserved regions in a protein sequence) that are indicative of nucleic acid strand-separation activity and/or the capacity to translocate along DNA. Biochemical and structural data show that helicases can function as monomers, dimers, and multimers (predominantly hexamers) and that they can be classified on the basis of their substrate requirement for dsDNA, double-stranded ribonucleic acid (dsRNA), or DNA-RNA hybrids. Helicases are further classified as either 3′ → 5′ or 5′ → 3′ enzymes based on their polarity of unwinding (a region of ssDNA of the given polarity must be adjacent to the dsDNA to be unwound, to act as a loading site for the helicase).

Superfamilies 1 and 2. Superfamilies 1 (SF1) and 2 (SF2) include thousands of DNA and RNA helicases. SF1 and SF2 enzymes are primarily monomeric and dimeric helicases that unwind DNA with either 3 → 5′ or 5′ → 3′ polarity. Domains 1A and 2A in SF1 helicases and domains 1 and 2 in SF2 helicases contain

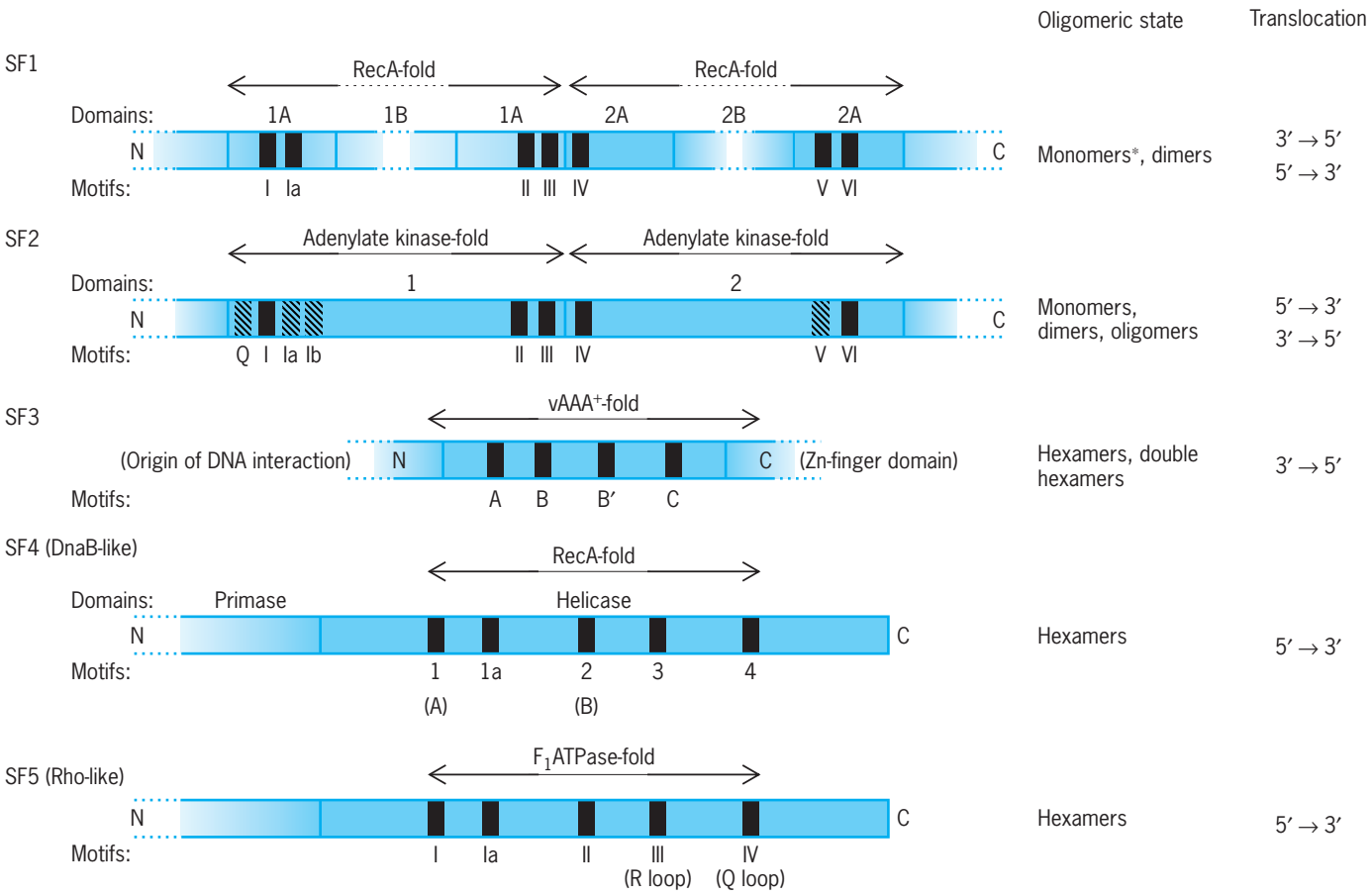

Fig. 2. Five major superfamilies of DNA helicases. The conserved signature amino acid sequence motifs are shown relative to the modular domains (shown as solid blocks) of helicases that belong to five major families. The motifs that are characteristic of all helicases within a particular superfamily are shown in black, while the motifs characteristic only to subsets of the SF2 helicases are represented by a hashed block. Alternative designations for some signature motifs are given in parentheses. In addition to the core domains containing the signature motifs, helicases frequently contain domains either adjacent to or inserted between the conserved motifs of the core domain. Examples of such domains are shown as discontinuous light color blocks. The areas containing large insertions that define additional enzymatic activities are indicated by broken lines. The oligomeric state and translocation polarity for each superfamily is indicated to the right of the block diagram.

similar sets of seven helicase signature motifs that comprise the basic structural core of these enzymes (Fig. 2). This core is responsible for both nucleotide binding and hydrolysis, and it includes a repeat of either of two structural domains: the RecA-fold in SF1 helicases or the related adenylate kinase-fold in SF2 helicases. Variation at the N and C termini, as well as the insertions between the core domains, give rise to differences in the substrate specificity or to additional enzyme activities. Motifs I and II [Walker A (GxGKT) and B (DExo/H) motifs, respectively] are the most conserved among all helicases.

Superfamily 3. Superfamily 3 (SF3) helicases are found in small DNA and RNA viruses. In contrast to SF1 and SF2 helicases, each subunit of these enzymes is composed of a single *vAAA*$^+$-fold that is responsible for nucleotide binding and hydrolysis.

Superfamilies 4 and 5. The smaller superfamilies 4 and 5 consist primarily of 5′ → 3′ hexameric helicases related either to the *Escherichia coli* DnaB helicase (SF4) or to the transcription termination factor Rho (SF5). In addition to Walker A and B motifs (characteristic of the RecA-fold), SF4 helicases contain three distinct conserved motifs (designated as motifs 1a, 3, and 4). All known SF4 enzymes are involved in the replication of bacteria and bacteriophages. These helicases are functionally and physically associated with DNA primases (enzymes that initiate the synthesis of each Okazaki fragment, short segments of newly synthesized DNA, by first synthesizing a short segment of RNA) either as two distinct but interacting proteins (for example, bacterial DnaB helicase and DnaG primase), or as two domains of one polypeptide, such as gp4 helicase-primase of bacteriophages T7 and T3 (as shown in Fig. 2). The nucleotide-binding and hydrolysis core of F5 helicases possesses the highest degree of similarity to the F1-ATPase fold, which is characteristic of proton-translocating ATPases. *See* ENZYME.

Directional translocation and DNA unwinding. It is now widely accepted that the unwinding polarity of helicases relates to a directional bias in translocation on ssDNA. For example, the enzyme depicted in Fig. 1 is a 3′ → 5′ helicase. Upon binding to the ssDNA, it starts moving toward the 5′ end of the loading strand, which brings the enzyme to the ssDNA-dsDNA junction; subsequent translocation along the ssDNA, coupled to either physical or electrostatic

destabilization of the dsDNA, results in strand separation in the duplex portion of the substrate. Conversely, a 5′ → 3′ helicase translocates toward the 3′ end of the loading strand and, therefore, requires a 5′-terminated region of ssDNA to unwind a DNA duplex.

Interestingly, high-resolution structural data suggest that the helicase signature motifs are not essential for the duplex DNA separation per se, but are for the ATP-dependent unidirectional motion of the helicases on the DNA lattice. Consequently, it is proposed that the helicase signature motifs define a modular structure for a DNA motor, while the additional domains, which may vary from one protein to another, are responsible for modulating DNA unwinding, mediating protein-protein interactions, or determining structure specificity. Although the classical substrate for a DNA helicase is duplex DNA adjacent to a region of ssDNA (as depicted in Fig. 1), some enzymes preferentially unwind more complicated structures, such as flaps, forks, bubbles, or three- and four-way junctions.

DNA translocation can also be characterized in terms of processivity, which is defined as the probability of forward translocation at each base pair opening cycle. For a highly processive enzyme, stepping forward is much more likely than dissociation. Consequently, an enzyme that never dissociates after each step has a processivity equal to one (which is a physically unrealistic situation in which the enzyme is infinitely processive), whereas an enzyme that dissociates after each step has a processivity equal to zero (and is referred to as being dispersive or distributive). The average number of base pairs unwound (or translocated) by an enzyme in a single binding event (N) relates to the processivity (P) as $N = 1/(1 - P)$. For example, the UvrD helicase, whose processivity is 0.98, unwinds on average 50 base pairs. On the other hand, the highly processive RecBCD, a multifunctional enzyme involved in the repair of dsDNA breaks by homologous recombinations, has a processivity of 0.99997 and unwinds on average 30,000 base pairs before dissociation.

Step size. Though seemingly simple, one of the more poorly defined parameters for DNA helicase behavior is the "step size." The step size has been defined in at least three different ways: (1) as the number of base pairs that the enzyme unwinds upon hydrolyzing one ATP molecule under its most efficient reaction conditions; (2) as a physical step size, which is a measure of the helicase's ability to "step over" a discontinuity in the DNA; and (3) as a distance associated with a kinetically rate-limiting step. The parameters obtained using these three definitions are not necessarily identical. For example, the RecBCD enzyme hydrolyzes about two molecules of ATP per base pair unwound, but has a kinetic step size of six base pairs. The physical step size of the related RecBC enzyme was determined to be 23 base pairs.

Accessory factors. Once a helicase has separated the strands of dsDNA, spontaneous reannealing of the duplex can be avoided if the nascent ssDNA strands are trapped by an ssDNA-binding protein or by other factors that pass the unwound nucleic acid intermediate to the next step of a reaction pathway (Fig. 1). While ssDNA-binding proteins have frequently been shown to stimulate helicase activity in vitro, helicase activity is also stimulated by other accessory factors that increase the rate or processivity of unwinding (such as processivity factors, clamp loaders, and polymerases). The primary replicative helicase of *E. coli*, DnaB, is a good example of a helicase that acts poorly in isolation but, as part of the replisome (the DNA synthesis machinery of the cell), efficiently separates the DNA strands at the replication fork. The rate of movement of the replication machinery at the fork is coordinated by an interaction between DnaB and DNA polymerase (the enzyme that synthesizes the complementary strand of DNA) that enhances its translocation rate by almost 30-fold to 1000 base pairs per second.

The physical arrangement and interactions between proteins comprising a replisome modulate their respective activities to ensure complete and coordinated synthesis of both DNA strands. DnaB is a hexameric helicase that translocates in a 5′ → 3′ direction on the DNA strand used as a template for lagging-strand synthesis. For its function within a replisome, DnaB helicase must be loaded on the lagging-strand template. This task is accomplished by the primasome assembly proteins, which coordinate recruitment of DnaB during establishment of the replication fork. Besides DNA polymerase and primasome assembly proteins, DnaB interacts with DnaG primase.

Single-molecule translocation visualization. Until recently, virtually all biochemical data on helicases were derived from conventional bulk-phase techniques, which observe the population-averaged properties of molecular ensembles. Direct observation of dsDNA unwinding by a single helicase enzyme has the potential to uncover valuable information on protein and DNA dynamics that may be obscured in bulk-phase assays. Two of these techniques were used to visualize translocation of a single molecule of the exceptionally fast and processive helicase RecBCD, which can unwind, on average, 30,000 bp of dsDNA per binding event, at a rate of 1000 bp/s, while simultaneously degrading the ssDNA products of its helicase activity.

In one approach, a device called an optical trap is used to manipulate individual, fluorescently labeled DNA molecules and to visualize their unwinding and degradation by RecBCD enzyme. In the presence of ATP, RecBCD enzyme unwinds the dsDNA, which is observed as a progressive shortening of the fluorescently labeled DNA molecule. An alternative single-molecule approach is the tethered particle motion experiment, which uses light microscopy to follow translocation of a biotin-tagged RecBCD enzyme bound to a streptavidin-coated polystyrene bead. The two single-molecule experiments are different yet complementary: the tethered particle motion experiment directly measures translocation,

whereas the optical trap method measures dsDNA unwinding. Together, these studies provide additional strong evidence for the coupling of DNA strand separation with movement of the helicase protein on its substrate lattice. Both single-molecule visualization methods show that RecBCD translocates unidirectionally and processively on dsDNA, with each molecule moving at a constant rate (within the limit of experimental detection). Although the average translocation rate is similar to that derived from bulk measurements, considerable variation is observed in the translocation rate of individual RecBCD enzymes.

These and other new techniques of single molecule detection will play an important role in determining the molecular mechanisms of DNA helicases. The next big challenges in this field are to understand how these DNA motors are incorporated into and function within the context of macromolecular machines to orchestrate complex DNA processing events. *See* MOLECULAR BIOLOGY; NUCLEOPROTEIN.

Maria Spies; Mark S. Dillingham;
Stephen C. Kowalczykowski

Bibliography. B. Alberts et al., *Molecular Biology of the Cell*, 3d ed., vol. II, chap. 6, Garland Publishing, New York, 1994; H. Lodish et al., *Molecular Cell Biology*, 4th ed., Chap. 12, Freeman, New York, 2000; P. Soultanas and D. B. Wigley, Unwinding the "Gordian Knot" of helicase action, *Trends Biochem. Sci.*, 26(1):47–54, 2001.

DNA methylation

The addition of a methyl group (—CH_3) to certain nucleotides in genomic deoxyribonucleic acid (DNA), a process which has important effects on gene expression. Besides adenine, guanine, cytosine, and thymidine, the four bases which spell out the primary sequence of genomic DNA, methylation produces a fifth base called methylcytosine. Methylcytosine is a covalent modification of DNA that is replicated when the cell divides. The enzymes that can add a methyl group to the cytosine ring to form methylcytosine are called DNA methyltransferases (DNMTs) [see reaction below]. In humans and other mammals, this

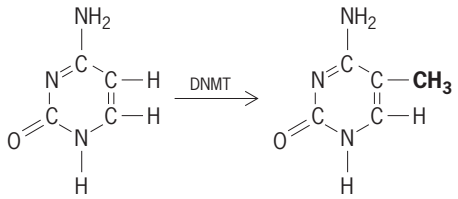

modification is imposed only on cytosines that precede a guanosine in the DNA sequence (the CpG nucleotide). In simple organisms, such as yeast and *Drosophila*, there is little or no DNA methylation. *See* DEOXYRIBONUCLEIC ACID (DNA).

The bulk of methylation in human DNA occurs in the large portion of the DNA that does not encode genes. This heavily methylated DNA replicates later than nonmethylated DNA. Late replication is associated with the formation of inactive chromatin (the threadlike strands of DNA complexed with histone proteins that form eukaryotic chromosomes), which facilitates transcriptional silencing of noncoding regions such as inserted viral sequences, repetitive DNA elements, and transposons (DNA sequences that move from their usual location into a new region of the genome). *See* NUCLEIC ACID.

Effect on gene expression. Studies in mammals have linked patterns of DNA methylation to gene expression. CpG dinucleotides are often clustered in small stretches of DNA termed CpG islands. Almost half the genes in the human genome have CpG-rich promoter regions, stretches of DNA where transcription of DNA into ribonucleic acid (RNA) begins. Methylation in a gene promoter region generally correlates with a silenced gene.

DNA methylation by itself does not directly repress transcription. Only after proteins are added to methylated sites in DNA, thereby forming chromatin, does transcriptional repression occur. Nucleosomes are structures composed of a core of histone proteins around which DNA winds. Long sequences of nucleosomes coil to form chromatin. Normally, at sites of active genes with nonmethylated CpG islands in promoter regions, the nucleosomes are widely and irregularly spaced in a manner favorable to the access of proteins that activate transcription. In contrast, when CpG islands are heavily methylated, the nucleosomes are tightly compacted and regularly spaced. This configuration excludes proteins that activate gene transcription or prevents them from acting as gene activators. Another regulatory feature of nucleosomes is contained in certain amino acid residues of histone tails. The covalent attachment of acetyl groups to lysine residues (acetylated state) in histone tails decreases the interactions between nucleosomes within the chromatin fiber, resulting in greater access of transcription factors to the DNA. These tails are kept in an acetylated state in the case of transcribed genes but in a deacetylated state in the case of hypermethylated, silenced genes. *See* GENE; GENE ACTION; NUCLEOSOME.

Role in cancer. Changes in the pattern of DNA methylation are commonly seen in human tumors. Both genomewide hypomethylation (insufficient methylation) and region-specific hypermethylation (excessive methylation) have been suggested to play a role in carcinogenesis.

DNA hypomethylation. The global level of DNA methylation is generally lower in tumor cells than in normal cells. This decrease in the overall level of DNA methylation is curious in light of the normal to high levels of DNA methyltransferase expression usually seen in tumor cells. The biological significance of DNA hypomethylation in cancer is not well understood. Because both genomic instability and hypomethylation are observed early in tumorigenesis, it has been tempting to speculate that genomic hypomethylation provides the incipient cancer cells with a way of acquiring more mutations by destabilizing the

genome. Defects in DNA methylation might contribute to chromosomal instability observed in aneuploid human colorectal cancers (that is, colorectal cancer resulting from the addition or deletion of whole chromosomes). Also, DNA hypomethylation of very discrete locations within the genome has been associated with abnormal chromosomal structures as observed in cells from patients with ICF (immunodeficiency, centromeric instability, and facial abnormalities) syndrome. This rare recessive disease is caused by mutations in the catalytic domain of the gene encoding one of the three biologically active DNMTs. *See* CHROMOSOME; HUMAN GENETICS; MUTATION; TUMOR.

DNA hypermethylation. DNA hypermethylation of CpG islands in gene promoter regions is associated with the inappropriate transcriptional silencing of tumor-suppressor genes, which indicates the important and pervasive role of DNA hypermethylation in oncogenesis. Because of its gene-silencing effect, abnormal promoter hypermethylation can have the same effect as a coding-region mutation in a copy of the gene. Examples are the *VHL* gene in renal cancer, the cell-cycle-control gene *p16* in several cancer types, and the mismatch-repair gene *hMLH1* in a subtype of colorectal cancers.

DNMTs. There are three known biologically active DNMTs in mammalian cells that catalyze the methylation process: DNMT1, DNMT3a, and DNMT3b. It is certain that each of these proteins is vital for embryonic development because inactivation of the corresponding genes in mice causes embryonic or early postnatal death. But much remains to be determined about how each participates in establishing normal DNA methylation patterns and the aberrant patterns in cancer. Though DNMT1 accounts for most methylation in mouse cells, human colorectal cancer cells lacking the *Dnmt1* gene retain significant genomic methylation and associated gene silencing. And disruption of the human *Dnmt3b* gene has been shown to reduce global DNA methylation by less than 3%. Surprisingly, genetic inactivation of both *Dnmt1* and *Dnmt3b* was found to nearly eliminate methyltransferase activity and reduce genomic DNA methylation by greater than 95%. These marked changes resulted in demethylation of repeated sequences, loss of insulinlike growth factor II (IGF2) imprinting (the differential expression of a gene depending on whether it was maternally or paternally inherited), abrogation or silencing of the tumor suppressor gene *p16INK4a*, and growth suppression. (IGF2 is thought to be an important growth factor in tumorigenesis; however, the specific physiologic effects of IGF2 in cancer are not known.) These results provide compelling evidence that the two enzymes cooperatively maintain DNA methylation and gene silencing in human colorectal cancer, and that such methylation is essential for optimal neoplastic proliferation. *See* ENZYME.

Diagnostic and therapeutic implications. The potentially reversible epigenetic changes in neoplasia present new opportunities for the clinical management of cancer. Strategies could potentially be developed that reverse gene silencing for the purpose of preventing and treating tumors. Moreover, the molecular changes associated with gene silencing may serve as markers for risk assessment, diagnosis, and prognosis. *See* CANCER (MEDICINE). Christoph Lengauer

Bibliography. A. Bird, DNA methylation patterns and epigenetic memory, *Genes Dev.*, 16:6–21, 2002; J. G. Herman and S. B. Baylin, Gene silencing in cancer in association with promoter hypermethylation, *New Engl. J. Med.*, 349:2042–2054, 2003; P. A. Jones and S. B. Baylin, The fundamental role of epigenetic events in cancer, *Nat. Rev. Genet.*, 3:415–428, 2002; C. Lengauer, Cancer—an unstable liaison, *Science*, 300:442–443, 2003; B. D. Strahl and C. D. Allis, The language of covalent histone modifications, *Nature*, 403:41–45, 2000.

Docodonta

A group of early mammals that currently includes at least eight genera from Laurasia (an ancient northern supercontinent comprising the landmasses that are now North America, Europe, Asia, Greenland, and Iceland) ranging in age from Middle Jurassic (159–180 mya) to Early Cretaceous (94–144 mya). Genera from the Late Triassic of Europe, Early Jurassic of India, and Late Cretaceous of South America have been tentatively referred to the Docodonta, but their evolutionary relationships are uncertain. All known docodonts were small. The skull of the most completely known docodont, the Late Jurassic genus *Haldanodon,* is approximately 30 mm (1.2 in) in length (**Fig. 1**). A reconstruction of its skeleton indicates the animal was equivalent in size to a modern mole. Other docodonts, known from isolated teeth or, at best, fragments of jaws and skulls, were of a similar size. *See* CRETACEOUS; JURASSIC.

The molariform cheek teeth of docodonts are transversely widened with two anteroposteriorly oriented rows of cusps linked by transverse crests (**Fig. 2**). Occlusion of the molariforms produced both a puncturing, or cutting, action and facilitated crushing and possibly grinding of their food, which probably included insects and other invertebrates. In

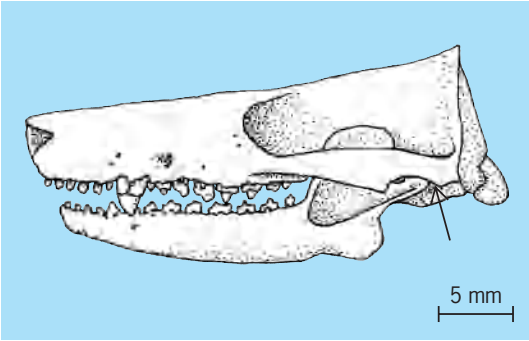

Fig. 1. Restoration of the skull of the Late Jurassic docodont *Haldanodon* in lateral view. Arrow points to dentary-squamosal articulation; the articular-quadrate articulation lies directly behind it (toward the midline of the skull). (*Illustration by Christine L. Huffard*)

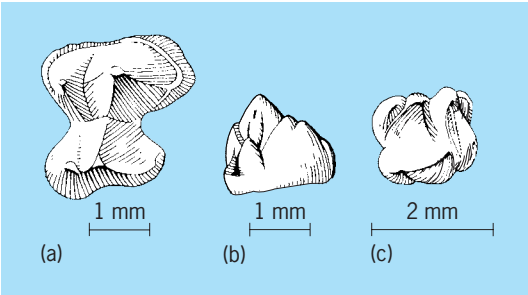

Fig. 2. Upper and lower molars of *Docodon*, a Late Jurassic docodont. (*a*) Occlusal view of an upper molar. (*b*) Internal and (*c*) occlusal views of a lower molar.

contrast to the highly derived (advanced) morphology and function of its dentition, the articulation of the lower jaws and skull of *Haldanodon* (and possibly other docodonts) exhibits a mixture of primitive and derived characters. Contact is retained between the articular bone of the lower jaw and the quadrate bone of the skull, which was the sole jaw articulation of their ancestors—the basal synapsids. The middle ear contained only one bone, the stapes (unlike the condition in other mammals, in which three bones are present). In addition to the primitive articular-quadrate articulation, a second jaw articulation, between the dentary bone of the lower jaw and the squamosal bone of the skull, had evolved. The latter articulation is present in all modern mammals in which the articular and quadrate bones form the malleus and incus of the middle ear. *See* EAR (VERTEBRATE); DENTITION.

Elements of the postcranial skeleton are known only from the Late Jurassic genus *Haldanodon*. The large size and sturdy processes for muscle attachments on the limb bones indicate the forearms were adapted for burrowing and swimming. *Haldanodon's* morphology and the depositional environment in which it was found (a coal swamp) suggest this docodont resembled modern semiaquatic desman moles and was an active swimmer that dug burrows in the banks of the ancient swamp. The nature of the sediments from which the fragmentary remains of other docodonts have been recovered suggest that they might not have been semiaquatic, but were probably small animals that scrambled through drier terrestrial environments.

The docodonts were one of several distinct lineages of basal mammals that differentiated during an evolutionary radiation that probably began in the Triassic. During the Jurassic docodonts diversified into two major groups, one limited to Asia and the other to Europe and North America. During the Cretaceous, docodonts became extinct: they were not ancestral to any other mammalian lineages. *See* ANIMAL EVOLUTION; MAMMALIA. William A. Clemens

Bibliography. M. Benton, *Vertebrate Palaeontology*, 3d ed., Blackwell, Malden, MA, 2005; T. Kemp, *The Origin and Evolution of Mammals*, Oxford University Press, 2005; Z. Kielan-Jaworowska, R. L. Cifelli, and Z.-X. Luo, *Mammals from the Age of Dinosaurs*, Columbia University Press, 2004.

Dogs

All breeds of domestic dogs, the wild dogs, and related species belong to the family Canidae. Despite the various breeds of domestic dogs that are known, the scientific name for all is *Canis familiaris*. The origin of domestic dogs is obscure, but they seem to be most closely related to the wolf. The dog has been associated with humans since Neolithic times, at least 10,000 years ago, and there are indications that as early as 4000 years ago several varieties of dogs had been developed by humans. In many respects the dog is structurally primitive and shows a genetic plasticity which accounts for the many varieties. There are more than 100 breeds of domestic dogs, and their classification is based principally on their uses.

The dog is adapted to running, is digitigrade (walking on tiptoe), and has four toes on each foot with a dewclaw on the front feet of many varieties that represent the remnant of the fifth claw. The blunt claws are nonretractible and may be used for digging. The jaw is elongate, with 42 teeth, and the dental formula is I 3/3, C 1/1, Pm 4/4, M 2/3. Dogs are mature toward the end of the first year, after which the female comes into heat twice yearly, the period lasting about 2 weeks. Like many of the vertebrates, the male has an os priapi, or penis bone. Gestation is about 8 weeks, and the size of the litter varies from 2 to 16. The pups are blind and helpless at birth but grow rapidly, opening their eyes at 10 days and walking at 20. At the age of 1 month all of the milk, or deciduous, teeth have erupted. Permanent dentition appears at about 5 months.

The wild species of the family, numbering about 36 and having a wide distribution, include several wild dogs, wolves, coyotes, foxes, and jackals.

Domestic breeds. The domestic breeds are the sporting breeds, hounds, terriers, working breeds, toy breeds, and the nonsporting group.

Sporting breeds. These animals are trained and employed for retrieving or finding game. The largest group in this class is the spaniels, of which there are about 10 types varying in size from 18 lb (8 kg) in the cocker spaniel to almost 50 lb (22 kg) in the field spaniel. The spaniels have been trained to retrieve and to flush game, and they hunt both birds and fur-bearing animals. The Brittany spaniel is the only member of the group which points.

Retrievers are specially trained to locate and return game to hunters and are used most commonly in hunting waterfowl. They are usually large dogs and excellent swimmers. There are four varieties: the curly, the Chesapeake Bay, the golden retriever, and the Labrador. With the exception of the golden retriever, the Labrador is the origin of the others by crossbreeding with spaniels and setters.

Pointers and setters, used for hunting upland game birds, range ahead of the hunter, point the game until the hunter arrives, and retrieve the fowl after it has been flushed and shot. Among the varieties that are used in these pursuits are the Weimeraner, the English and Irish setters, and the German short-haired

pointer. The setter probably is derived from pointer and spaniel ancestry.

Hounds. This group includes the basset, the bloodhound, the whippet (produced by crossing a terrier with an Italian greyhound and then backbreeding to the English greyhound), the dachshund, the wolfhound, and the beagle (**Fig. 1**). The greyhound, one of the oldest breeds, is built for speed with its thin body and long legs. The bloodhounds and foxhounds (smaller, stockier dogs) are used for hunting, mostly by scent. The beagle, now more a pet than a hunter, can also follow a scent and is easier to follow on foot.

Terriers. The terriers originally were bred for hunting burrowing animals, such as the badger and fox. The Boston terrier, probably derived from the bulldog and bull terrier, is the only breed to have originated in the United States. The fox terrier, typical of the group, may have been derived from the foxhound. It was originally bred for fox hunting but is an excellent ratter. Other terriers are the Airedale, largest of the group, and the Scottish and Skye terriers.

Fig. 1. The beagle, a popular domestic breed.

Working breeds. Most of these breeds are large animals used as draught animals, for police work, for herding, and as guide dogs for the blind. The St. Bernard is a draught animal in the Low Countries, and is used for milk delivery and other work. Other draught breeds are the Alaskan malamute, the Eskimo, and the Samoyed.

Among the animals used as guard dogs and for police work are the Doberman, the German shepherd, and the Great Dane. All these animals are specially trained for sentry duty and guard work, but the Great Dane, one of the most powerful of dogs, is also used to herd cattle and is still employed as a draught animal in France and Belgium. The collies, Belgian sheep dogs, and English sheep dogs are outstanding sheep herders.

The bulldog is now more of a pet and house dog than a guard dog, but it is pugnacious if set on an intruder and will not release the person. The poodle, said to be the most intelligent dog, can be trained as a gun dog and was originally used for duck shooting.

Toy breeds. Some of these breeds have been known for centuries and are principally household pets which may develop an instinct for protecting the premises of the owner. They are all quite small, some being miniatures of the larger breeds. The chihuahua is the smallest, weighing as little as 1 lb (0.45 kg). Some of the more popular varieties are the Pomeranian, Pekingese, and pug.

Nonsporting breeds. Many of the dogs included in the foregoing groups are included in this category as a matter of opinion. The Dalmatian may be included in this group as well as the poodle. The chow has been raised for centuries in China, where it was used as a hunting dog as well as for its meat and fur. It is now used as a guard dog and a household pet, and is unusual in that its tongue is blue.

Wild species. There are a number of wild dogs which have never been domesticated, unlike the dingo of Australia (*Canis dingo*) that is believed to have been a domestic dog introduced into Australia during prehistoric times and then reverted to its wild state. It is the only wild carnivore in Australia. This animal has been domesticated by the natives and is used to hunt kangaroos and other game. In the wild state it occurs singly rather than in packs and is destructive to livestock.

The Asiatic wild dog occurs throughout Asia, Java, and Sumatra and is considered to be three distinct species by some authorities while others regard it as two subspecies of *Cuon alpinus*, which is also known as the Siberian wild dog. In India this dog is known as the dhole and has been given the name *C. dukhunensis*. It is slightly larger than the Siberian variety, which is about 3 ft (1 m) long, and lives in forested areas. The Malay variety (*C. javanicus*) is the smallest of the three, and the fur is a brighter color. These are running hunters, hunting in packs which consist of a family or a number of families that will attack any animal they encounter.

The Cape hunting dog (*Lycaon pictus*) ranges throughout the grasslands of eastern and southern Africa but has become reduced in numbers. This animal runs in packs which may contain up to 60 individuals and usually rests during the day and hunts when it is cooler. They differ from other members of the family in having four toes on the front foot. The bush dog (*Speothos venaticus*), the maned wolf (*Chrysocyon brachyurus*), and the Cordillera fox (*Cerdocyon magellanicus*), all indigenous to South America, are considered to be wild dogs despite the common names. The bush dog is about 2 ft (0.6 m) long with short legs and tail. The maned wolf, more than 4 ft (1.2 m) long with a 16-in. (40-cm) tail, resembles a giant fox and has extremely long legs. The Cordillera fox is about 3 ft (1 m) in length, excluding the tail.

Wolves. The common European wolf (*Canis lupus*) is the species that once ranged throughout the temperate forested regions of Europe, Asia, and

North America. This species has been exterminated in the British Isles and almost so in France, but they do occur in other European countries such as Italy, Spain, and the Balkans and are still plentiful in the Scandinavian countries. The coat is usually a brown-gray color, and its thickness depends upon the environment. Wolves howl; they do not bark as do dogs. The adult stands about $2^1/_2$ ft (0.75 m) tall and weighs more than 100 lb (45 kg). They inhabit hills and plains areas in addition to forested areas. The female has one litter of from three to nine cubs each year. Mating occurs at the beginning of the year, and the gestation period is about 2 months. The female usually prepares a den near some stream, lines it with vegetation, and then cares for the cubs until they are weaned at about 6 weeks. The father remains with the family and will raise the cubs if anything happens to the female. The offspring are mature at the age of 2. Wolves hunt in packs and are extremely cunning. Due to recovery efforts, the gray wolf or timber wolf, recently restricted to Alaska and the subarctic regions of Canada, has now been restored to many areas of its historic North American range—including portions of the southwest United States, the Rocky Mountains, and the western Great Lakes region.

The coyote (*Canis latrans*) is sometimes called the prairie wolf and is a close relative of the true wolf, although it is smaller than the wolves, the adult being 2 ft (0.6 m) high and weighing about 40 lb (18 kg). Coyotes have historically inhabited the prairies, open plains, and desert areas of North America; however, since the early twentieth century they have spread eastward as far as New England and are now ubiquitous throughout the contiguous United States, being found in many types of habitats. Food consists of insects, prairie dogs, and large game when they hunt in packs, but usually this animal hunts alone. It has keen vision and an excellent sense of smell and is extremely fast. The den is frequently one which has been abandoned by a fox, badger, or skunk, and it is enlarged and cleaned thoroughly. In the spring or early summer the litter of five to seven cubs is born.

Fig. 2. Red fox (*Vulpes vulpes*). (*Photo by Gerald and Buff Corsi; © 1999 California Academy of Science*)

Jackals. These animals are scavengers as well as menaces to domestic poultry. They are nocturnal in their activities and hunt in packs. Both the oriental jackal (*Canis aureus*), the most widely distributed jackal, and the black-backed jackal (*C. mesomelas*) can be easily tamed. *Canis aureus* has spread from southeastern Europe and northern Africa through Asia as far south as Burma. It prefers higher elevations in contrast to *C. mesomelas*, which is found in the grasslands of eastern and southern Africa. It is often a solitary hunter; however, several may form a pack and follow the trail of a predator such as a lion to finish the remains of a kill. The simenian jackal (*C. simensis*), a large species found in Abyssinia, is rare, and little is known of its natural history. The most common enemy of the jackal is the leopard.

Foxes. These members of the dog family have relatively short legs and long bodies, big erect ears, pointed snouts, and long bushy tails (**Fig. 2**). The red fox (*Vulpes vulpes*), found throughout Europe and the North America as well as parts of Asia, North Africa, and Australia, actually has undergone many color phases and mutations and includes other varieties such as the silver fox and the cross fox. *Vulpes vulpes* stands a few inches over 1 ft (0.3 m) and weighs from 15 to 20 lb (6.5 to 9 kg). Scent glands are present in the anal region, which account for the characteristic odor of these animals.

Foxes may use an abandoned den of another animal, dig their own den in a protected location, or live aboveground in a secluded hollow under rocks or in tree hollows. These latter abodes are referred to as kennels. The fox is active nocturnally and preys upon small rodents, birds, and rabbits; if it lives near a stream, it also eats crayfish and fish.

The male may hunt alone or may pair with the female, known as a vixen, to chase the quarry. Mating occurs in the winter, and the gestation period is 8 weeks. The litter varies from three to eight cubs, which are born blind. The cubs are weaned at the age of 2 months, and the male assists in feeding them thereafter. They remain with the parents until autumn, learning to hunt and fend for themselves, and mature for the following year. Near farm areas they may be a nuisance to poultry breeders since they may raid coops.

Most species of foxes provide good fur, and with the destruction of natural wildlife environments and the increased demand for fur, many fox farms were started in Canada and the Scandinavian countries as well as in the northern areas of the United States. These animals breed well in captivity. *See* CARNIVORA; MAMMALIA; SCENT GLAND. Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, 6th ed., vols. I and II, Johns Hopkins

University Press, Baltimore, 1999; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, Washington, DC, 1999; J. O. Whitaker and W. J. Hamilton, *Mammals of the Eastern United States*, 3d ed., Comstock, 1998.

Dogwood

A tree, *Cornus florida*, also known as flowering dogwood, which may reach a height of 40 ft (12 m) and is found in the eastern half of the United States and in southern Ontario, Canada. It has opposite, simple, deciduous leaves with entire margins (see **illus.**).

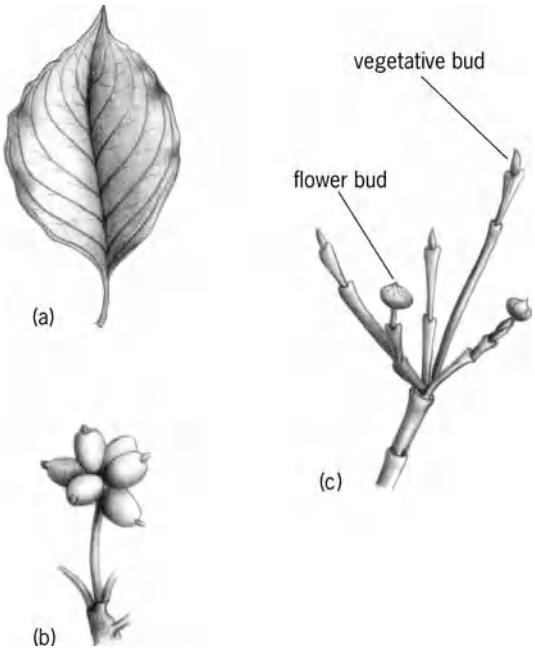

Flowering dogwood (*Cornus florida*), with characteristic (*a*) leaf, (*b*) fruit cluster, and (*c*) branch.

When this tree is in full flower, the four large, white, notched bracts or petallike growths surrounding the small head of flowers give an ornamental effect that is unequaled by any native tree. Pink, rose, and cream-colored varieties are commonly planted. The tree is tolerant of shade and at blossoming time, usually in early May, the patches of white reveal its presence, even in dense woods. The wood is very hard and is used for roller skates, carpenters' planes, and other articles in which hardness is desired. As a shade tree, it is especially desirable for the modern ranch-type house where small size is appropriate. The Pacific dogwood (*C. nuttalli*), which grows in Idaho and from southwestern British Columbia to southern California, is similar to the eastern dogwood, but has rounded bracts. The Japanese dogwood (*C. kousa*) is a similar small tree with pointed bracts and blooms in June. Other shrubby species of dogwood are used as ornamentals. *See* CORNALES; FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Dolerite

A fine-textured, dark-gray to black igneous rock composed mostly of plagioclase feldspar (labradorite) and pyroxene and exhibiting ophitic texture. It is commonly used for crushed stone. Its resistance to weathering and its general appearance make it a first-class material for monuments. *See* STONE AND STONE PRODUCTS.

The most diagnostic feature is the ophitic texture, in which small rectangular plagioclase crystals are enclosed or partially wrapped by large crystals of pyroxene. As the quantity of pyroxene decreases, the mineral becomes more interstitial to feldspar. The rock is closely allied chemically and mineralogically with basalt and gabbro. As grain size increases, the rock passes into gabbro; as it decreases, diabase passes into basalt. *See* BASALT; GABBRO.

Diabase forms by relatively rapid crystallization of basaltic magma (rock melt). It is a common and extremely widespread rock type. It forms dikes, sills, sheets, and other small intrusive bodies. The Palisades of the Hudson, near New York City, are formed of a thick horizontal sheet of diabase. In the lower part of this sheet is a layer rich in the mineral olivine. This concentration is attributed by some investigators to settling of heavy olivine crystals through the molten diabase and by others to movement of early crystals away from the walls of the passageway along which the melt flowed upward from depth, before it spread horizontally to form the sill. *See* MAGMA.

As defined, diabase is equivalent to the British term dolerite. The British term diabase is an altered diabase in the sense defined here. *See* BASALT; GABBRO; IGNEOUS ROCKS.

Carleton A. Chapman

Dolomite

The carbonate mineral $CaMg(CO_3)_2$. Often small amounts of iron, manganese, or excess calcium replace some of the magnesium; cobalt, zinc, lead, and barium are more rarely found. Dolomite merges into ankerite with increasing substitution of iron for magnesium, and into kutnahorite with increasing manganese substitution. *See* ANKERITE.

Dolomite has hexagonal (rhombohedral) symmetry and a structure similar to that of calcite, but with alternate layers of calcium ions being completely replaced by magnesium. Thus each layer of CO_3^{2-} ions lies between a layer consisting entirely of Ca^{2+} ions and one consisting entirely of Mg^{2+} ions. This ordered arrangement of the cations distinguishes the structure of dolomite from that of a randomly substituted calcite. *See* CALCITE.

Dolomite is a very common mineral, occurring in a variety of geologic settings. It is often found in ultrabasic igneous rocks, notably in carbonatites and serpentinites, in metamorphosed carbonate sediments, where it may recrystallize to form dolomite marbles, and in hydrothermal veins. The primary occurrence of dolomite is in sedimentary deposits, where it constitutes the major component of dolomite rock

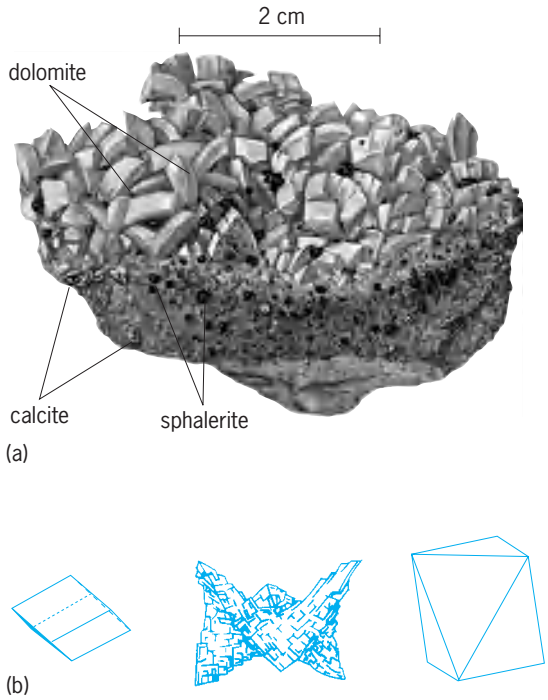

Dolomite. (*a*) Crystals with calcite and sphalerite on limestone, which were found in Joplin, Missouri (*specimen from Department of Geology, Bryn Mawr College*). (*b*) Crystal habits (*after C. S. Hurlbut, Jr., Dana's Manual of Mineralogy, 17th ed., John Wiley and Sons, 1959*).

and is often present in limestones (see **illus.**). *See* DOLOMITE ROCK; LIMESTONE; SEDIMENTARY ROCKS.

Dolomite is normally white or colorless with a specific gravity of 2.9 and a hardness of 3.5–4 on Mohs scale. It can be distinguished from calcite by its extremely slow reaction with cold dilute acid. *See* CARBONATE MINERALS.

Alan M. Gaines

Bibliography. L. L. Y. Chang, R. A. Howie, J. Zussman, *Rock-Forming Minerals, vol. 5B: Non-Silicates: Sulphates, Carbonates, Phosphates, Halides*, 1995.

Dolomite rock

Sedimentary rock containing more than 50% by weight of the mineral dolomite [$CaMg(CO_3)_2$]. The term dolostone is used synonymously. A range of the Alps in northeastern Italy is named the Dolomites because it is composed predominantly of dolomite rock. *See* DOLOMITE.

Characteristic properties. Since dolomites usually form by replacement of preexisting limestones, it is often possible to see relict grains, fossils, and sedimentary structures preserved in dolomite rocks. More often, however, original textures are obliterated. *See* LIMESTONE.

There is a very strong tendency for dolomite to form euhedral rhombs. This is useful for distinguishing dolomite from the usually anhedral calcite. Dolomite rhombs vary in size from micrometers (on edge) to millimeters. These rhombs often contain cloudy, inclusion-filled cores and clearer rims. Trace-element zonation is also common, as can be seen in stained thin sections or by cathodoluminescence. *See* PETROGRAPHY.

Field differentiation of dolomite from calcite is most easily accomplished by application of dilute hydrochloric acid. Calcite is strongly effervescent in acid, whereas dolomite is usually very weakly effervescent unless scratched or ground into a fine powder. Dolomite also often weathers to a brown color on outcrop because of the common substitution of iron for magnesium in the dolomite structure. X-ray diffraction analysis is the most reliable method for the differentiation of dolomite from calcite, and also the best method for characterization of the degree of ordering of a particular dolomite sample. *See* X-RAY DIFFRACTION.

Dolomite rocks are predominantly monomineralic. The most common noncarbonate minerals are quartz (either authigenic or detrital), clay minerals, pyrite, and glauconite. Evaporite minerals or their replacements are also common.

Occurrence and origin. Geologists have been unable to decipher the exact conditions of dolomite formation. This so-called dolomite problem revolves around several questions relating to the stoichiometry of the reaction in which dolomite is formed, the fact that dolomite is not common in young marine sediments, and the type of geological setting in which ancient dolomites formed.

If dolomite formed in nature as a primary or direct precipitate from constituent ions dissolved in aqueous solutions, dolomitization could be written as reaction (1). Alternatively, if dolomite formed as a

$$Ca^{2+} + Mg^{2+} + 2CO_3{}^{2-} \longrightarrow CaMg(CO_3)_2 \qquad (1)$$

secondary mineral by replacement of precursor aragonite or calcite (two different minerals with the same chemical composition: $CaCO_3$), the dolomitization reaction could be represented as (2). A third, intermediate option is given in reaction (3). There is no

$$2CaCO_3 + Mg^{2+} \longrightarrow CaMg(CO_3)_2 + Ca^{2+} \qquad (2)$$

$$CaCO_3 + Mg^{2+} + CO_3{}^{2-} \longrightarrow CaMg(CO_3)_2 \qquad (3)$$

evidence for dolomite forming as a primary precipitate within ocean (or lake) water by reaction (1). Hydrothermal syntheses of dolomite at temperatures of 100°C (212°F) and greater have proven that dolomite formation can take place by the replacement of calcium carbonate [reaction (2)]. Reaction (3) is almost certainly the most important in nature: most of the calcium in dolomite originates in a solid calcium carbonate precursor, but no extra calcium is released to solution. *See* ARAGONITE; CALCITE.

Examination of reaction (2) indicates that the higher the Mg^{2+}/Ca^{2+} ratio in solution, the greater the tendency for dolomite to replace calcium carbonate. If the ratio is too high, however, magnesite ($MgCO_3$) may replace dolomite [reaction (4)].

$$CaMg(CO_3)_2 + Mg^{2+} \longrightarrow 2MgCO_3 + Ca^{2+} \qquad (4)$$

The calculated range of stability of dolomite over

calcite and magnesite (at 25°C or 77°F) is between Mg^{2+}/Ca^{2+} ratios of about 0.7 and 5.5, respectively. Since many natural waters, including seawater, have Mg^{2+}/Ca^{2+} ratios within this range, dolomite should be the stable carbonate mineral in nature. Surprisingly, dolomite is very rare in young carbonate sediments, which are instead dominated by aragonite and calcite.

The relative propositions of dolomite and calcite have changed through geologic time. Dolomitic rocks are dominant in the Precambrian and early Paleozoic, whereas calcitic rocks become dominant in the late Mesozoic and continue to dominate through the present. Several alternative hypotheses have been offered to explain this. Perhaps the composition of Precambrian oceans was different from that of present-day oceans (either in their dissolved CO_2 content, their Mg^{2+}/Ca^{2+} ratio, or even their dissolved sulfate content). Perhaps just the increased chance of exposure of calcitic precursor rocks to dolomitizing solutions ensured that older rocks would be more dolomitic. Perhaps the mineralogy of potential dolomitic precursors has changed through geologic time. Aragonite dolomitizes much more rapidly than calcite; hence a shift in the dominant mineralogy of carbonate sediments from aragonite to calcite in the Mesozoic could favor the dominance of calcitic rocks over dolomitic ones.

In a few geological settings, dolomite is forming within the sediments at the present time. One general setting includes a wide variety of supratidal ponds, lagoons, tidal flats, and sabkhas such as the Solar Lake in Israel, the Coorong Lagoon in Australia, the bank tops of Andros Island in the Bahamas, and the Sabkha of Abu Dhabi. Each of these environments is hot, more saline than seawater, and rich in organic matter. A second general setting of modern dolomite formation is in continental margin sediments underlying productive coastal oceanic upwelling zones. Examples include the Peru–Chile margin, the southern California borderlands, the Gulf of California, and the Walvis Ridge west of Namibia (southwestern Africa). This second set of geological environments is characterized by lower water temperature and normal marine salinity. Like the first set, these environments are rich in organic matter. The association of dolomite formation with abundant organic carbon seems to be the one common thread linking each dolomite occurrence and possibly controlling dolomite precipitation.

Previously, researchers had focused a great deal of attention on so-called dolomitization models, which were largely concerned with how to get sufficient reactants to the site of dolomite formation. Significant amounts of dolomite formation require efficient removal of Mg^{2+} from very large volumes of seawater. A successful model therefore requires a mechanism to move large amounts of seawater over buried calcium carbonate sediments. For some reason, ordinary seawater does not permit the formation of dolomite. It is now believed that the dolomitization reaction is very slow, but can be accelerated by the removal from seawater of dissolved sulfate (which occurs during the postburial breakdown of sedimentary organic matter) and by high temperatures.

Economic implication. Dolomites are extremely important oil reservoir rocks. This is partly a result of the high porosity of many dolomite rocks and partly the result of the association between dolomite formation and the presence of sedimentary organic matter.

Dolomites are also the main host rocks for lead and zinc ore deposits. Rocks of this type are known as Mississippi Valley ore deposits. Neither the origin of the ore-forming solutions nor the association between the lead-zinc ores and the dolomite rocks has been satisfactorily explained. *See* PETROLEUM GEOLOGY; SEDIMENTARY ROCKS. Paul A. Baker

Bibliography. J. R. Allan and W. D. Wiggins, *Dolomite Reservoirs: Geochemical Techniques for Evaluating Origin and Distribution*, 1993; P. A. Baker and M. Kastner, Constraints on the formation of sedimentary dolomite, *Science*, 213(4504):214–216, 1981; D. H. Zenger, J. B. Dunham, and R. L. Ethington (eds.), *Concepts and Models of Dolomitization*, Soc. Econ. Paleontol. Mineral. Spec. Publ. 28, 1980.

Domain (electricity and magnetism)

A region in a solid within which elementary atomic or molecular magnetic or electric moments are uniformly aligned.

Ferromagnetic domains are regions of parallel-aligned magnetic moments. Each domain may be thought of as a tiny magnet pointing in a certain direction. The relatively thin boundary region between two domains is called a domain wall. Within a wall the magnetic moments rotate from the direction of one of the domains to the direction in the adjacent domain.

A ferromagnet generally consists of a large number of domains. For example, a sample of pure iron at room temperature contains many domains whose directions are distributed randomly, making the sample appear to be unmagnetized as a whole. Iron is called magnetically soft since the domain walls move easily if a magnetic field is applied. In a magnetically hard or permanent magnet material (for example, alnico or neodymium-iron-boron) a net macroscopic magnetization is introduced by exposure to a large external magnetic field, but thereafter domain walls are difficult to either form or move, and the material retains its overall magnetization.

Antiferromagnetic domains are regions of antiparallel-aligned magnetic moments. They are associated with the presence of grain boundaries, twinning, and other crystal inhomogeneities. *See* ANTIFERROMAGNETISM; TWINNING (CRYSTALLOGRAPHY).

Ferroelectric domains are electrical analogs of ferromagnetic domains. *See* FERROELECTRICS; FERROMAGNETISM; MAGNETIZATION. J. F. Herbst

Domestication (anthropology)

The adaptation of an animal or plant through breeding in captivity to a life intimately associated with and advantageous to humans. Domesticated plants and animals differ genetically and morphologically from their wild counterparts, often in ways that make them more useful to humans. For example, wild sheep are hairy but domesticated sheep are woolly. Some of the earliest domesticated animals include sheep, goats, cattle, pigs, and horses, which were domesticated in the Old World, and llamas and alpacas, which were domesticated in the Americas. Dogs are known from both the Eastern and Western hemispheres. Important early domesticated plants include cereals—wheat, barley, rice, sorghum, and millet in the Old World and maize in the Americas—and tubers such as yams and taro in the Old World and potatoes in the Americas. Other important domesticated plants include pulses, olives, cotton, flax, and fruits. Domesticated plants and animals provide humans with a variety of useful products, including food and fibers, and serve important functions, such as animal traction (that is, the use of animals to pull carts and plows).

Agricultural revolution. The beginnings of plant and animal domestication is sometimes referred to as the agricultural or Neolithic revolution. (The term "Neolithic" means new stone age and refers to the appearance of ground and polished stone axes which are commonly found on early agricultural sites.) Analogous to the industrial revolution, the agricultural revolution represented a change in subsistence technology that affected many aspects of human societies. The beginning of agriculture and animal husbandry was a significant turning point in human history. Until the end of the Ice Age, about 11,500 years ago, all humans throughout the world lived by foraging, that is, hunting animals and gathering plants. At the time of Columbus's voyages, about 500 years ago, it is estimated that about 15% of the world's population lived by hunting and gathering. Today almost no hunters and gatherers remain, and the bulk of the world's food supply comes from agricultural produce.

The study of the beginnings of agriculture has intrigued archeologists and anthropologists for over a century because agriculture provided the economic basis for the development of complex urban societies. Maize was a staple crop throughout the Americas in later pre-Columbian times, and cereals such as wheat, barley, and rice were critical to the development of complex societies in the Eastern Hemisphere. Farming provides one major advantage over hunting and gathering—a higher yield per acre. The surplus food provided by early farmers could be used to support craftsmen and other nonfood producers and to finance the construction of monumental public works such as temples and pyramids.

The reasons why human societies began to cultivate crops and herd animals are not well understood; many possible explanations for the origins of farming have been put forward. Since the early part of the twentieth century, some archeologists have suggested that the climatic changes at the end of the Ice Age either encouraged or forced early human populations to experiment with agriculture because they affected the distribution of wild plants and animals. Other scholars have suggested that population growth may have led to the adoption of farming. Still others have focused on social factors, suggesting that the social demands of competitive feasting and exchange may have encouraged innovations that led to higher food yields. Today most archeologists reject a monocausal explanation for the beginnings of farming. Environmental, demographic, and social factors undoubtedly played very different roles in the beginnings of farming in various parts of the world. *See* AGRICULTURAL SCIENCE (ANIMAL); AGRICULTURAL SCIENCE (PLANT); AGRICULTURE; NEOLITHIC.

Methods of study. Studies of plant and animal domestication are most often carried out by zooarcheologists and paleoethnobotanists, archeologists with specialized expertise in animal bones and plant remains. In order to study the origins and spread of plant and animal domestication, archeologists rely on plant and animals remains that are recovered from archeological sites. Animal bones are some of the most common finds on archeological excavations. They are generally recovered using fine screens so that the bones of small animals are not lost. Evidence for animal domestication can include the discovery of an animal outside its wild range, morphological changes such as size reduction, and changes to the demographic structure of the animal herd. Plant remains are generally recovered using the flotation process, in which archeological soils are poured into moving water. Light organic materials, such as charred seeds, are held in suspension momentarily so that they can be poured off into a fine-mesh sieve. In cereal crops, plant domestication can lead to changes in the rachis, the axis of the plant that holds the grains. Wild plants have a brittle rachis, allowing the plants to reseed themselves when the grains ripen, whereas domestication can produce a nonbrittle rachis. *See* ARCHEOLOGY; PALEOBOTANY.

Centers of domestication in the Middle East. Studies of agricultural origins have focused on the Near and Middle East because the ancient Near East was home to some of the world's earliest complex societies, including the Sumerians and Akkadians. In addition, archeological research has shown that the Near East is home to some of the world's earliest farmers. The Natufian culture of the southern Levant, the region along the east coast of the Mediterranean, is seen as pivotal in our understanding of the transition from foraging to farming.

Environmental studies have shown that climatic amelioration at the end of the Ice Age, beginning about 14,000 years ago, led to an expansion of Mediterranean vegetation throughout the region. The Mediterranean vegetation zone is characterized as an oak-pistachio parkland and includes the wild ancestors of both emmer wheat (*Triticum dicoccoides*) and barley (*Hordeum spontaneum*)

[**illus.** *a* and *c*]. Natufian foragers appear to have collected wild cereals and other plant foods such as acorns and to have hunted gazelles and other large mammals between about 14,000 and 11,500 years ago. Although the Natufians made their living by hunting and gathering, they did keep domestic dogs (*Canis familiaris*). *See* DOGS; WHEAT.

In parts of the southern Levant, at sites such as Mallaha/Eynan in northern Israel and Wadi Hammeh 27 in Jordan, Natufians inhabited small villages on a year-round basis. Colder and drier conditions toward the end of the Ice Age, about 12,000 years ago, may have disrupted the Natufian way of life, especially in the drier, more marginal parts of the southern Levant. Some archeologists have suggested that these environmental changes may have encouraged experimentation with cereal cultivation. *See* ICE AGE.

Plant cultivation. Shortly after the end of the Ice Age, foragers throughout the southern Levant appear to have started experimenting with cereal cultivation. Sites such as Netiv Hagdud in the Jordan Valley have provided archeological evidence for the earliest stages of plant cultivation. The cereal grains recovered from this site, dated from about 9000 B.C., are morphologically wild. They have a brittle rachis, which allows the barley to reseed itself. The case for cultivation is based on the high proportion of barley grains recovered at Netiv Hagdud (about 90% of the grasses at the site). Possible domestic barley grains appear at the site of Tell Aswad in Syria, dated to about 7700 B.C., but the earliest well-dated barley was recovered from Ain Ghazal in Jordan and dates to about 7000–6500 B.C. Domesticated emmer wheat was also widespread in the Near East by 7300 B.C. Barley and emmer wheat were among the earliest domesticated plants (illus. *a* and *c*). Deoxyribonucleic acid (DNA) and archeological evidence suggest that a second species of wheat, einkorn (*T. monococcum*), was initially domesticated in Anatolia (illus. *b*). During the early Neolithic period, shortly after the domestication of cereals, pulses such as lentils and bitter vetch were also domesticated in the Middle East.

The archeological evidence suggests that plant domestication preceded animal domestication in the Middle East. Early sites such as Netiv Hagdud provide evidence for early plant cultivation combined with the hunting of wild animals. *See* BARLEY; CEREAL.

Animal domestication. The earliest evidence for animal domestication in the Near East appears in the early eighth millennium B.C. Archeological evidence for domestic goats first appears at the site of Ganj Dareh in Iran (illus. *d*). In general, early herders treated male and female goats somewhat differently. Since females are the productive members of the herd, providing both milk and offspring, female animals typically survive into adulthood. Surplus male animals are often slaughtered as juveniles. The kill patterns for early Neolithic goats from Ganj Dareh reflect this differential treatment of males and females. During the seventh millennium B.C., goats from a number of sites throughout the Near East show morphological changes, such as twisting of the horn cores (the bony portion of the skull that lies underneath the horn), that are hallmarks of animal domestication. Other early Near Eastern animal domesticates include sheep, followed by cattle and pigs (illus. *e*). *See* GOAT PRODUCTION.

Primary Near Eastern domesticates. (*a*) Emmer wheat, *Triticum dicoccum* (wild from *T. dicoccoides*). (*b*) Einkorn wheat, *T. monococcum* (wild form *T. boeoticum*). (*c*) Barley, *Hordeum vulgaris* (wild form *H. spontaneum*). (*d*) Goats, *Capra hircus* (wild form *C. aegagrus*, the Bezoar goat). (*e*) Sheep, *Ovis aries* (wild form *O. orientalis*, the Asiatic mouflon). (*Reprinted with permission from P. J. Crabtree and D. V. Campana, Archaeology and Prehistory, McGraw-Hill, 2001*)

Other early centers of domestication. While the Near East appears to be home to the earliest farmers in the Eastern Hemisphere, a number of other early centers of plant and animal domestication have been identified. Rice was initially domesticated in southeast Asia, while broomcorn millet (*Panicum miliaceum*) was first domesticated in north China. Recent archeological research has suggested that New Guinea farmers may have begun to plant taro and bananas as early as 5000 B.C. A variety of plants, including sorghum, pearl millet, yams, and watermelon, were initially domesticated in sub-Saharan west Africa, possibly beginning about 4500 B.C. In addition, recent DNA studies have suggested that Africa just south of the Sahara may have been an independent center of cattle domestication as early as 10,000 years ago. Humped cattle (*Bos indicus*) were first domesticated in South Asia. Later Eurasian domesticates include horses, which were probably domesticated in the Russian/Ukrainian steppe

zone about 3500 B.C., and camels, which were domesticated in both eastern Iran (Bactrian, or two-hump, camels) and Arabia (dromedary, or one-hump, camels) about 2500 B.C. The domestication of the olive and the vine transformed Mediterranean agriculture, since olives and grapes can be grown on land that is unsuited to cereal crops. Biochemical analyses have identified wine residues from an early pottery vessel from Iran dated to 5400 B.C. Widespread finds of transport amphorae, large ceramic vessels used to transport liquids, indicate that wine and olive oil were widely traded throughout the ancient world. *See* BEEF CATTLE PRODUCTION; CAMEL; OLIVE.

Centers of domestication in the Americas. Archeologists working in the Americas have been especially interested in studying the domestication of maize, as this crop was a staple food for the Aztec, Inca, Maya, and other complex societies of the New World.

Mesoamerica. During the 1960s, the Tehuacán Valley project was designed to study the origins of plant domestication in highland southern Mexico. The initial results suggested that domesticated maize first appeared in that region about 3500 B.C. However, redating of the early domesticated corn cobs from the Tehuacán Valley using the accelerator mass spectrometer (AMS) process has shown that these cobs are nearly 1000 years younger than was originally thought. Today the earliest known maize cobs come from the site of Guilá Naquitz in Oaxaca and date to about 4300 B.C. This cave site has also produced the earliest known domesticated squash (*Cucurbita pepo*) from the region. This squash is dated to 8000–6000 B.C., making it several millennia older than the earliest domesticated maize.

Recent archeological research on maize domestication has focused on the Balsas River Valley in the Pacific highlands of southern Mexico, as DNA research has shown that the teosinte (*Zea mays* ssp. *parviglumis*, the wild ancestor of corn) from this region is most closely related to domesticated corn. Since the corn cobs from Guilá Naquitz are unquestionably domesticated, the initial cultivation of maize must have begun appreciably earlier than 4300 B.C. *See* CORN.

Other plants that were domesticated in Mesoamerica, the geographical region that includes both southern Mexico and Central America, include beans, peppers, and avocados. While a wide range of plants were cultivated in Mesoamerica, domesticated animals played only a limited role in the pre-Columbian economies of Central America. The most important animal domesticates were the dog and the turkey.

Highland South America. Highland South America was a separate center of plant and animal domestication in prehistoric times. Plants that were first domesticated in this region include quinoa (*Chenopodium quinoa*), a small-seeded cereallike plant, and potatoes. Potatoes are an especially useful crop because they can grow on relatively poor soils. Both the llama (*Lama glama*) and the alpaca (*Lama pacos*) were initially domesticated in the Andean region. The llama is a multipurpose animal that was used for meat and for bearing burdens. The alpaca was raised primarily for its fine wool. In later prehistoric times, alpaca wool was combined with cotton, which was grown in the arid regions along the Peruvian coast, to produce fine textiles. *See* ALPACA; LLAMA.

Eastern South America and eastern North America. Two other centers of plant domestication have been identified in the Americas. Manioc (*Manihot escuelenta*), a starchy root crop, was initially domesticated in the tropical rainforest region of eastern South America. A number of small-seeded plants were domesticated in eastern North America, including sunflowers, sumpweed (*Iva annua*), and goosefoot (*Chenopodium berlandieri*). Native American farmers began to domesticate these seeds around 2500 B.C., two millennia before maize was introduced to the eastern United States.

Conclusion. The beginning of farming was a true revolution in human subsistence and ways of life. Most of the foods we eat today would have been unknown to hunters and gatherers at the end of the Ice Age. Archeological research conducted since world War II has shown that plants and animals were domesticated in many different regions of the Eastern and Western hemispheres. Ongoing archeological research will continue to shed light on the agricultural revolution and its consequences. Pam J. Crabtree

Bibliography. D. G. Bradley, Genetic hoofprints: The DNA train leading back to the origin of today's cattle, *Nat. Hist.*, 112(1):36–41, 2003; J. Clutton-Brock, *Domesticated Animals from Early Times*, University of Texas Press, Austin, 1981; J. Clutton-Brock, *A Natural History of Domesticated Mammals*, 2d ed., Cambridge University Press, 2000; J. Diamond, Evolution, consequences and future of plant and animal domestication, *Nature*, 418:700–707, Aug. 8, 2002; J. Diamond, *Guns, Germs, and Steel: The Fates of Human Societies.*, W. W. Norton, New York, 1997; B. D. Smith, *The Emergence of Agriculture*, Scientific American Library, New York, 1998.

Dominance

The expression of a trait in both the homozygous and the heterozygous condition. In experiments with the garden pea, the Austrian botanist Gregor Mendel crossed plants from true-breeding strains containing contrasting sets of characters. For seed shape, round and wrinkled strains were used. When plants with round seeds were crossed to plants with wrinkled seeds (P_1 generation), all offspring had round seeds. When the offspring (F_1 generation) were self-crossed, 5474 of the resulting F_2 offspring were round and 1850 were wrinkled. Thus, the round trait is expressed in both the F_1 and F_2 generations, while the wrinkled trait is not expressed in the F_1 but is reexpressed in the F_2 in about one-fourth of the offspring. In reporting these results in a paper published in 1866, Mendel called the trait which is expressed in the F_1 generation a dominant trait, while the trait which is unexpressed in the F_1 but reappears

in the F_2 generation was called a recessive trait. *See* MENDELISM.

Traits such as round or wrinkled are visible expressions of genes. This visible expression of a gene is known as the phenotype, while the genetic constitution of an individual is known as its genotype. The alternate forms of a single gene such as round or wrinkled seed shape are known as alleles. In the P_1 round plants, both alleles are identical (since the plant is true-breeding), and the individual is said to be homozygous for this trait. The F_1 round plants are not true-breeding, since they give rise to both round and wrinkled offspring, and are said to be heterozygous. In this case, then, the round allele is dominant to the wrinkled, since it is expressed in both the homozygous and heterozygous condition. Dominance is not an inherent property of a gene or an allele, but instead is a term used to describe the relationship between phenotype and genotype. *See* ALLELE; GENE ACTION.

Incomplete dominance. The production of phenotypes which are intermediate between those of the parents is an example of partial or incomplete dominance. For example, in *Mirabilis,* the four-o'clock, plants with red flowers crossed to plants with white flowers produce pink offspring. Since in this case only enough red pigment is made to give a pink color, dominance is said to be partial or incomplete. Because of this lack of dominance, the phenotype of the heterozygote can be distinguished from that of either of the homozygous parents. In other words, the phenotypic ratio is equal to the genotypic ratio in both the F_1 and F_2 generations.

The phenomenon of incomplete dominance which results in a clear-cut intermediate phenotype is relatively rare. However, even in cases where dominance appears to be complete, there is often evidence for intermediate gene expression. For example, in the human biochemical disorder phenylketonuria, homozygous recessive individuals are severely affected, while heterozygotes are phenotypically normal. Homozygous recessive individuals have little or no activity of the enzyme phenylalanine hydroxylase. Heterozygous individuals have only 50% of the normal level of enzyme activity, yet are unaffected by the disease. *See* PHENYLKETONURIA.

Codominance. The separate and distinct expression of both alleles of a gene is an example of codominance. This is a situation unlike that of incomplete dominance or complete dominance. In humans, the MN blood group is characterized by the presence of molecules called glycoproteins on the surface of red blood cells. These molecules or antigens contribute to the immunological identity of an individual. In the MN blood system, persons belong to blood groups M, MN, or N. These phenotypes are produced by two alleles, M and N, each of which controls the synthesis of a variant glycoprotein. In the heterozygote MN, there is separate and complete expression of each allele. This is in contrast to incomplete dominance, where there is an intermediate or blending effect in heterozygotes. Codominance usually results in the production of gene products of both alleles. *See* BLOOD GROUPS.

Overdominance. Individuals in which the phenotype of the heterozygote is more extreme than in either of the parents are said to exhibit overdominance. In *Drosophila*, for example, the gene for white eye (*w*) when in the heterozygous condition causes an increase in the production of certain fluorescent pigments resulting in levels that are greater than either the wild type or the homozygous recessive genotypes. Examples of overdominance at the level of enzymes are also available. In the land snail *Cepaea hortilis*, homozygotes for alleles of the enzyme isocitrate dehydrogenase display four to five times less activity than do heterozygotes. In corn, alleles for the enzyme alcohol dehydrogenase when homozygous can produce forms that are stable and have low activity, or forms that are unstable with high activity. Heterozygotes between an unstable and a stable inactive form exhibit both stability and high activity. The concept of overdominance is also important in understanding the genetic structure of populations and is usually related to characteristics associated with fitness, such as size and viability.

Heterosis. The production of superior hybrid offspring by crossing two different strains of an organism is known as heterosis. The hybrid superiority may take the form of increased resistance to disease or greater yield in grain production. Heterosis in crop plants, such as corn, has been intensively studied, and the use of hybrid corn in the United States has more than doubled the average yield of bushels per acre. This has been accomplished by first establishing many highly inbred lines and then crossing them in all combinations to yield hybrids. The resulting hybrids are screened for increased productivity. Those with the highest yield are then used for crop production. *See* BREEDING (PLANT).

The mechanism which results in heterosis has been widely debated but is still unknown. Two theories have been advanced to explain the origin of hybrid superiority. One is based on the concept of overdominance. It is clear that heterozygotes for single gene differences may be superior to either of the homozygotes. At the molecular level, this overdominance may be caused by the production of metabolic products or enzymes which have combined the advantages of substances produced by either homozygote. Using this as a background, some researchers have argued that hybrids heterozygous at a large number of loci are able to thrive in environmental conditions too extreme for either homozygote. In other words, highly heterozygous genotypes are superior to homozygous ones, and this idea is called the overdominance theory of heterosis. This theory was first proposed in the early 1900s and is based on work with corn and other plants.

The alternate explanation for heterosis, known as the dominance hypothesis, is based on the idea that most recessive alleles are deleterious, while the dominant alleles of that gene are not. In this view, inbreeding produces a situation in which individuals become highly homozygous for alleles which include

many deleterious recessive traits. When these highly inbred lines are crossed, these deleterious alleles are masked by dominant alleles contributed by the opposite strain. In this way, then, the hybrid becomes superior to either homozygote.

It is difficult to determine which hypothesis is correct, and it may be that either or both mechanisms may operate in individual circumstances. In some systems, dominance can explain heterosis, while in other cases, overdominant heterozygotes are superior to either parental strain. *See* HETEROSIS.

Dominance hierarchies. In some cases, a population may carry a large number of alleles for a given gene. Although an individual carries only two alleles, within other members of a population many alternate forms of the same gene may exist. In systems involving multiple alleles, there is often a hierarchy or ordering of dominance relationships. The phenotypes for color in the peppered moth, *Biston betularia*, are a classic example of a dominance hierarchy. In *Biston*, there are three color forms, with a dominance ordering of black, gray, and white. In the land snail *Cepaea*, shell color involves a dominance hierarchy for a large number of alleles in the following order: brown, dark pink, pink, pale pink, dark yellow, yellow, pale yellow, and white. In such a series of multiple alleles which generate distinct phenotypes, each is dominant to all others below it in the hierarchy and is recessive to all phenotypes above it.

Dominance and haploids. Since dominance depends on the relationship between two alleles, it is a phenomenon usually limited to diploid organisms, or to those that have a diploid phase in their life cycle. In haploid organisms such as bacteria, however, a transitory diploid condition can be achieved. Genetic recombination in bacteria through events such as transformation, transduction, or conjugation involve transfer of all, or part of, a bacterial chromosome from one cell to another. The recipient cell is thus diploid for the transferred genes. Such cells, called merozygotes or partial zygotes, do not remain diploid. Genetic recombination takes place, and some of the transferred genetic material is incorporated into the host chromosome. All nonintegrated sequences are then lost from the cell, and a haploid recombinant cell results. *See* BACTERIAL GENETICS; CHROMOSOME; RECOMBINATION (GENETICS).

Partial diploids have been used in bacteria to explore the mechanisms of gene regulation. In the *lac* operon of *Escherichia coli,* the operator, which is a DNA sequence that regulates the transcription of three adjacent genes, can effectively regulate only the genes which are adjacent to it in the same chromosome. When a functional operator (without adjacent structural genes) is introduced into a cell which has a nonfunctional operator adjacent to three wild-type structural genes, the transferred operator is not able to exert control over the host's structural genes. As a consequence, the operator is said to be cis-dominant, because it can regulate only the genes which are adjacent to it on the same DNA molecule. *See* OPERON.

Dominant mutations. In general, mutations of a gene can cause a phenotypic change in two ways: the gene may be deleted or have a reduced or absent function (loss of function); or the gene product may function in an abnormal way or take on a new function (gain of function).

Most often, loss-of-function mutations result in a recessive phenotype, because in heterozygotes a 50% reduction in the amount of a gene product has no effect. However, in cases where half the normal level of gene product is not sufficient, a loss-of-function mutation in one allele can produce a dominant phenotype, called a dominant dosage effect. Several dominantly inherited human disorders, including Marfan syndrome, aniridia, and hypercholesterolemia, are examples of such mutations. For example, in hypercholesterolemia heterozygotes, half of the cell surface receptors that remove cholesterol from the bloodstream are defective, raising blood cholesterol levels, which in turn leads to atherosclerosis and heart disease.

Most dominant gain-of-function mutations produce a new phenotype when an allele functions abnormally (wrong time, wrong cell type, wrong amount). Very few dominant mutations produce a phenotype by taking on a new function. Human disorders classified as dominant gain-of-function mutations include McCune-Albright syndrome, where a mutant receptor is switched on at the wrong time, resulting in precocious puberty; Charcot-Marie-Tooth disease, caused by overexpression of a gene product (wrong amount); and Huntington's disease, one of the few cases where a mutant allele may have a new function.

Some gain-of-function mutations produce an abnormal phenotype when the mutant allele is not functional, and also prevent the product of the normal allele from functioning. This situation is called a dominant-negative effect. These effects are usually seen when the gene product is a protein that is composed of subunits and that functions in the form of a dimer or a multimer. Some forms of osteogenesis imperfecta, a dominantly inherited connective tissue disorder in humans, are caused by abnormalities in multimeric proteins.

Theories. The question of the origin and evolution of dominance was first examined by R. A. Fisher in 1928. He reasoned that most mutations are recessive and deleterious. To bring the phenotype of the heterozygote to a state which is identical to that of the dominant homozygote, he proposed that modifiers at other genetic loci are selected for, and that these modifiers act on the expression of the heterozygous phenotype, making it resemble that of the wild-type or homozygous dominant individual. This proposal was criticized by J. B. S. Haldane and others because of the degree of selection that would be required.

A physiological explanation of dominance was put forward by S. Wright in 1934. This theory attempted to account for dominance at a biochemical level by considering the kinetics of enzyme reactions and the existence of metabolic pathways. Wright argued that

variations in metabolic activity brought about by the heterozygous condition are likely to have little effect on the phenotype because enzymes are linked together in pathways so that the substrate of one enzyme is the product of another. Recessive mutations, when homozygous, may halt the activity of one enzyme and thus bring the entire pathway to a halt, producing a mutant phenotype. Heterozygotes, on the other hand, are likely to have only a reduction in activity of one enzyme which will be averaged out over the entire metabolic pathway, producing little phenotypic effect.

Molecular studies of dominance have extended Wright's ideas by exploring the kinetic structure of metabolic pathways and enzyme systems. The results obtained thus far tend to support the thrust of his hypothesis, and have established that the dominant phenotype seen in heterozygotes for a recessive allele can be explained without the need to invoke the existence of modifiers. Other work on organisms ranging from *Drosophila* to humans has determined the relationship between a mutant recessive allele, detectable changes in enzyme activity, and the phenotype. Reduction in enzyme activity is common to most recessive mutations in the heterozygous condition, with only negligible effects on the phenotype. Even reductions to 50% of normal enzyme activity in the heterozygote are not usually detectable at the phenotypic level. The example of phenylketonuria heterozygotes with 50% reduction in enzyme activity but normal phenotype is typical for human metabolic mutations. All such mutations are classified as recessive, leading to the speculation that the widespread occurrence of recessive mutations is a consequence of the kinetic organization and structure of enzyme networks and enzyme-directed pathways. *See* GENETICS; MOLECULAR BIOLOGY; MUTATION. Michael R. Cummings

Bibliography. B. Charlesworth, Adaptive evolution: The struggle for dominance, *Curr. Biol.*, 8:R502–R504, 1998; A. Cornish-Bowden, Dominance is not inevitable, *J. Theor. Biol.*, 125:333–338, 1987; J. L. Crosby, The evolution and nature of dominance, *J. Theor. Biol.*, 5:35–51, 1963; H. Kacser and J. A. Burns, The molecular basis of dominance, *Genetics*, 97:639–666, 1981; H. A. Orr, A test of Fisher's theory of dominance, *Proc. Nat. Acad. Sci. USA*, 88:11413–11415, 1991; T. Strachan and A. P. Read, *Human Molecular Genetics*, 1996; A. O. Wilkie, The molecular basis of genetic dominance, *J. Med. Genet.*, 31:89–98, 1994.

Donkeys and allies

Donkeys, asses, and mules are included in the family Equidae along with the horses and zebras. These are odd-toed ungulates and therefore belong to the mammalian order Perissodactyla.

Donkey. These animals, *Equus asinus* (see **illus.**), originated in Africa, and their only relatives in the wild state are found in Ethiopia and northwestern Somalia (Somaliland). They have been domesticated

The donkey, *Equus asinus*, sometimes referred to as a burro.

and are used principally as pack animals throughout Asia and Europe. The donkey was brought to the United States in 1868. They are hardy animals but better suited to hot climates since they are sensitive to cold. No race can be defined since they are the result of many generations of crossbreeding, but the European donkeys have certain markings which resemble the wild ass.

Ass. The wild ass of Asia (*E. hemionus*) is known by a variety of common names, and because of its extensive distribution it has been divided into a number of races or subspecies. In the central regions of Mongolia is the kulan (*E. hemionus*), the onager (*E. h. onager*) occurs in Persia and India, while the kiang (*E. kiang*), largest of all the races, is restricted to Tibet and the Sikkim. The Syrian wild ass (*E. h. hemippus*) is extremely rare and is believed to be extinct by many authorities. All of these animals are about 4 ft (1.2 m) high at the withers. They are the most abundant of the wild Equidae after the zebras.

The African wild ass (*E. asinus*) differs from the wild horse in having longer ears, a short tail which is only a tuft of hair, narrower hooves, and a shorter, more upright mane. This species is also becoming rare in its natural habitat, being found only between the Sudan and Somaliland.

Mule. This animal is a hybrid resulting from the cross of the male ass (*E. asinus*) and the mare or female horse (*E. caballus*). The mule is a surefooted animal, noted for its endurance, unexcitability, and ability to withstand hot weather. When the female ass is bred to a stallion, the cross is called a hinny. These crosses do not occur naturally and both mule and hinny are usually sterile. *See* MULE PRODUCTION; PERISSODACTYLA; ZEBRA. Charles B. Curtin

Donnan equilibrium

The distribution of diffusible ions on either side of a semipermeable membrane in the presence of macroions that are too large to pass through the membrane. The Donnan equilibrium thus is the result of (1) external constraints (boundary conditions) that

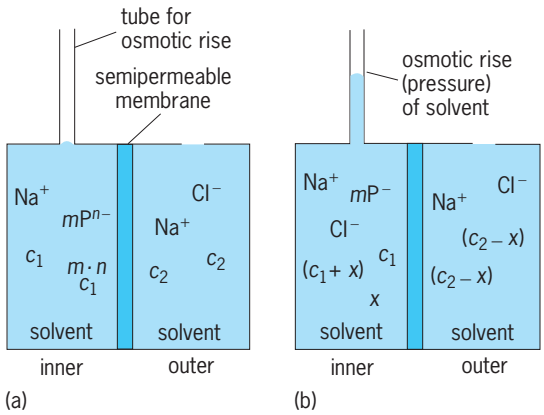

Schematic showing two compartments separated by a semipermeable membrane (*a*) before equilibrium and (*b*) at equilibrium. Originally, the left compartment contains *m* macroelectrolyte ions (P^{-n}) and its counterions (Na^+), in aqueous solution. The right compartment contains only NaCl in water. The membrane is permeable for Na^+ and Cl^- ions (and water) only.

enforce an unequal distribution of mobile ions and (2) a corresponding electrical potential on the membrane as a balance. This equilibrium is named for F. G. Donnan, the first to describe it and formulate the theory.

Typically, within a solution system that consists of two communicating compartments, the polyions of a macromolecular polyelectrolyte (a colloidal electrolyte) are confined to one compartment (for example, the inner one) by the semipermeable membrane, which allows the exchange of solvent and low-molecular-weight electrolytes by diffusion between the inner and outer compartment. Since electrical neutrality must obtain in either compartment, and since the macro-ions are confined to one side, the concentration of the diffusible ions in the two compartments cannot be the same (see **illus.**). The concentrations of the three ionic components (Na^+, Cl^-, P^{-n}) at the beginning of the experiment (illus. *a*) are: m, for P^{-n}, and $m \times n = c_1$ for Na^+ on the inside, and $Na^+ = Cl^- = c_2$ on the outside. After osmotic equilibrium has been reached and the membrane potential has been set up (illus. *b*), the concentrations are $(c_1 + x)$ for Na^+ on the inside, m unchanged for P^{-n}, and x for the Cl^-. On the outside, $Na^+ = Cl^- = (c_2 - x)$. The equilibrium condition is $(c_1 + x)x = (c_2 - x)^2$. At thermodynamic equilibrium, assuming for simplicity an electrolyte of which both ions are small, carry a charge of 1, and diffuse readily throughout the system in which they are dissolved, the chemical potential (activities) for the outer ionic pair must equal that of the diffusible ion pair on the inside. This condition is fulfilled when the product of the activities are equal, as shown in Eq. (1),

$$m_o^+ m_o^- \gamma_o^\pm = m_i^+ m_i^- \gamma_i^\pm \qquad (1)$$

where m_o and m_i are the molalities and γ_o and γ_i are the activity coefficients within the outer and inner compartments, respectively. In dilute solution, the molalities are approximately equal to the molarities, and the activity coefficients are practically equal to 1. The equilibrium of the solvent (whether the nondiffusible species is ionized or not) follows upon flux of the solvent into the (closed) cell that holds the macromolecular solute, until osmotic pressure equalizes the activities of the solvent in both cells. *See* ACTIVITY (THERMODYNAMICS).

Donnan membrane potential. The inequality of the individual ionic concentrations on either side of the membrane means that, at equilibrium, the membrane must carry a charge whose free energy is equal in magnitude and opposite in sign to the free energy that results from the unequal individual activities of the diffusible ions. This is shown in Eq. (2) for a negatively charged nondiffusible polyelectrolyte, where E is the potential difference. In other words, the presence of the nondiffusing species on one side of the membrane and the resulting differences in salt concentration enforce a polarization of the membrane (which is independent of the nature of the diffusible ions). Because of the equilibrium, the direct measurement of the electromotive force (membrane potential) by a pair of ion-reversible electrodes in each compartment would yield zero. However, the membrane potential can be measured by inserting two calomel reference electrodes into the inner and outer compartments while short-circuiting the two reversible electrodes, effectively bypassing the membrane. *See* ELECTROMOTIVE FORCE (EMF); FREE ENERGY; REFERENCE ELECTRODE.

$$E = \frac{RT}{F} \ln \frac{[Na_i^+]}{[Na_o^-]} = \frac{RT}{F} \ln \frac{[Cl_o^-]}{[Cl_i^-]} \qquad (2)$$

Observations. Equation (1) permits the determination of the concentration (activity) of the ions of the same charge as the macro-ions within the inside cell, which is equal to the activity of the ions that have diffused in from the outer cell (see illus.). It can also be seen that the osmotic pressure of a closed inner cell must be larger than the osmotic pressure that would be caused by the macro-ions alone, by an amount that corresponds to the number of ions that compensate the charges on the macro-ions, plus that of the ions that diffused in; that is, the osmotic pressure depends on the average molecular weight of the nonsolvent components of the inner cell. Therefore, if the salt concentration of the outer cell, and along with it the concentration in the inner cell, is raised, the contribution to the osmotic pressure by the macro-ions becomes relatively smaller and eventually becomes constant, amounting to the osmotic pressure of the macro-ions in their uncharged state. At the same time, the concentrations of the diffusible ions inside and outside approach one another. The same situation obtains in the absence of a membrane, when microscopic gel particles possess covalently bound ionized groups in their interior. Suspended microscopic gel particles, or dense polyelectrolyte coils, may thus carry a charge due to the effect of the Donnan equilibrium. *See* OSMOSIS.

The Donnan equilibrium is a frequent contributor to membrane potentials, but differential adsorption or, as in living systems, differences in the rates of passive or active ion transport through biological membranes are usually the main sources of membrane potentials of living cells.

If the concentration of the diffusible ions on the outside is kept near zero, as in the case of dialysis, the ions that were the counterions of the macroelectrolyte eventually diffuse out and become replaced by hydrogen ions (H^+), or hydroxide ions (OH^-); the original macro-micro-ion-salt (a salt containing ions of different sizes) has been hydrolyzed by the removal of the diffusible ions. *See* COLLOID; DIALYSIS; HYDROLYSIS; ION-SELECTIVE MEMBRANES AND ELECTRODES. F. R. Eirich

Bibliography. F. G. Donnan, The theory of membrane equilibria, *Chem. Rev.*, 1(1):73–90, 1924; P. L. Hiemenz, *Principles of Colloid and Surface Chemistry*, 3d ed., 1986; W. Hoppe et al. (eds), *Biophysics*, 1983.

Donor atom

An impurity atom in a semiconductor which can contribute or donate one or more conduction electrons to the crystal by becoming ionized and positively charged. For example, an atom of column 15 of the periodic table, such as phosphorus, arsenic, or antimony, substituting for a regular atom of a germanium or silicon crystal is a donor because it has one or more valence electrons which can be detached and added to the conduction band of the crystal (see **illus.**). Donor atoms thus tend to increase the number of conduction electrons in the semiconductor. The ionization energy of a donor atom is the energy required to dissociate the electron from the atom and put it in the conduction band of the crystal. *See* ACCEPTOR ATOM; SEMICONDUCTOR. H. Y. Fan

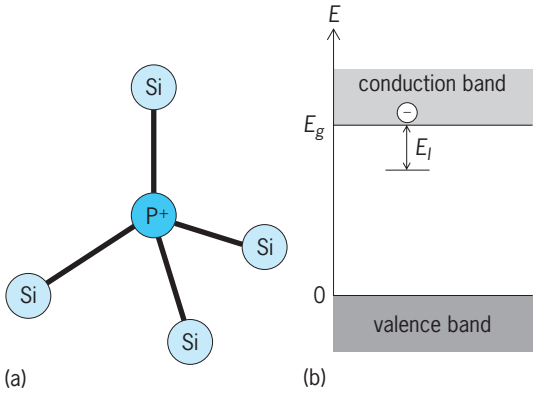

Pentavalent donor atom, phosphorus (P), in the elemental semiconductor silicon (Si). (*a*) Phosphorus atom in a substitutional position, that is, replacing silicon, a tetravalent host atom, by completing the four tetrahedral covalent bonds with its nearest neighbor silicon atoms. This results in an extra electron not needed in this bonding scheme, making phosphorus positively charged. (*b*) Energy diagram showing the conduction band lying at an energy E_g above the valence bond. The fifth electron from subtitutional phosphorus is donated to the empty conduction band. The donor electron is bound to phosphorus via Coulomb attraction with an ionization energy E_I.

Bibliography. C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; A. K. Ramdas and S. Rodriguez, Spectroscopy of the solid-state analogues of the hydrogen atom: donors and acceptors in semiconductors, *Rep. Prog. Phys.*, 44:1297–1387, 1981.

Dopamine

A catecholamine neurotransmitter that is synthesized by certain neurons in the brain and interacts with specific receptor sites on target neurons.

Biosynthesis. Catecholamine neurotransmitters—such as dopamine, norepinephrine, and epinephrine—are synthesized inside certain neurons from the common amino acid L-tyrosine. The biosynthetic pathway for catecholamine neuroransmitters has four steps (see **illustration**). (1) The enzyme tyrosine hydroxylase converts L-tyrosine to 3,4-dihydroxyphenylalanine (L-DOPA); this reaction also converts tetrahydrobiopterin and molecular oxygen to dihydrobiopterin and water. Tyrosine hydroxylase is found in all cells containing catecholamine neurotransmitters and is the regulatory, or rate-limiting, step in catecholamine biosynthesis. (2) The enzyme aromatic amimo acid decarboxylase converts L-DOPA to dopamine and carbon dioxide. In dopamine-producing cells, the dopamine is then stored in vesicles within the nerve terminal, which can fuse with the cell membrane to release dopamine into the synapse. (3) In neurons that produce norepinephrine and epinephrine, the enzyme dopamine beta-hydroxylase converts dopamine to norepinephrine; this reaction also converts ascorbate and molecular oxygen to dehydroascorbate and water. (4) In neurons that produce epinephrine (chromaffin cells), the enzyme phenylethanolamino *N*-methyltransferase converts norepinephrine to epinephrine; this reaction also converts *S*-adenosylmethionine to *S*-adenosylhomocysteine. *See* EPINEPHRINE.

Release from neuron. The release of neurotransmitter is controlled by a variety of factors, including the firing rate of the dopamine nerve cell (termed impulse-dependent release) and the release- and synthesis-modulating presynaptic dopamine receptors located on the dopamine nerve terminals. Since presynaptic dopamine receptors are sensitive to the cell's own neurotransmitter, they are called dopamine autoreceptors. Once released, dopamine also acts at postsynaptic receptors to influence behavior. The actions of dopamine in the synapse are terminated primarily by the reuptake of neurotransmitter into the presynaptic terminal by means of an active dopamine transporter. Dopamine may then be either repackaged into synaptic vesicles for rerelease or degraded by the enzyme monoamine oxidase. In addition, there is some extracellular inactivation. The dopamine transporter is an important

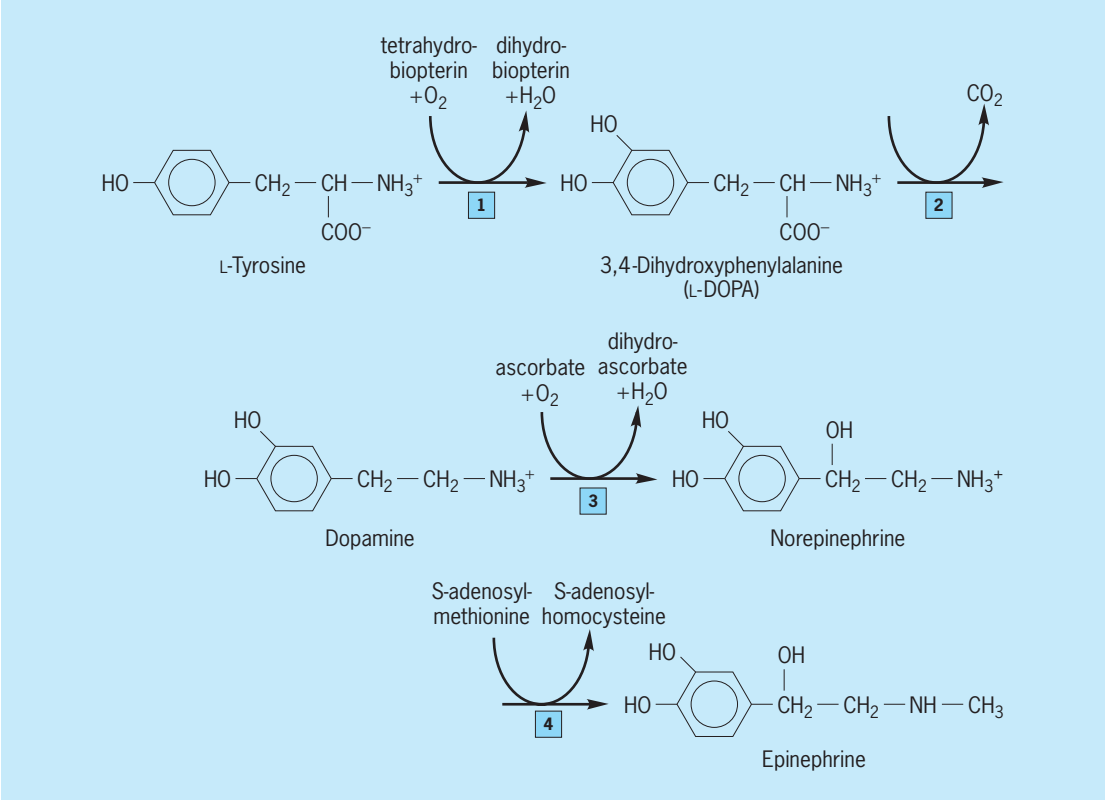

Biosynthetic pathway for catecholamine neurotransmitters.

site of action of the drugs cocaine and amphetamine. *See* AMPHETAMINE; COCAINE; MONOAMINE OXIDASE.

Dopamine neuronal systems. Although it was first thought that dopamine occurred only as an intermediate product formed in the biosynthesis of two other catecholamine neurotransmitters, norepinephrine and epinephrine, dopamine is now recognized as a neurotransmitter in its own right. Several distinct dopamine neuronal systems have been identified in the brain. These include systems within the hypothalamus and the pituitary gland; systems within the midbrain that project to a variety of cortical and limbic regions and basal ganglia; the retinal system; and the olfactory system. *See* BRAIN; NORADRENERGIC SYSTEM.

Role in behavior. The midbrain dopamine neurons which project to a variety of forebrain structures are critically involved in normal behavioral attention and arousal; abnormalities in the normal functioning of these systems have been implicated in a variety of disorders. For example, Parkinson's disease involves a degeneration of the midbrain dopamine neurons. This condition is often successfully treated by providing affected individuals with L-dopa, which is readily converted to dopamine in the brain. Attention deficit disorder, which is usually first diagnosed in childhood, is thought to involve dopamine systems, because the treatment of choice, methylphenidate, binds to the dopamine transporter and alters dopamine levels in the synapse. *See* PARKINSON'S DISEASE.

Drugs used to treat the major symptoms of schizophrenia are potent dopamine receptor antagonists. It is possible that certain types of schizophrenia are the result of increased activity in dopamine neuronal systems. In particular, it has been suggested that increased activity of either mesocortical or mesolimbic systems is likely to be involved. Although there is no doubt that neurobiological alterations underlie these disorders, a clear alteration in dopamine systems in the postmortem brain of schizophrenic individuals has not been conclusively demonstrated. A similar involvement of midbrain dopamine systems has been implicated in the multiple tic disorder Tourette's syndrome, which is treated, often successfully, with dopamine receptor antagonists. *See* NEUROBIOLOGY; SCHIZOPHRENIA.

Louis A. Chiodo; M. Todd Washington

Bibliography. J. R. Cooper, F. E. Bloom, and R. H. Roth, *The Biochemical Basis of Neuropharmacology*, 7th ed., 1996; G. J. Siegel et al. (eds.), *Basic Neurochemistry*, 6th ed., 1998; D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 4th ed., 2004; D. Voet and J. G. Voet, *Biochemistry*, 3d ed., 2004.

Doppler effect

The change in the frequency of a wave observed at a receiver whenever the source or the receiver is moving relative to each other or to the carrier of the wave (the medium). The effect was predicted in 1842 by C. Doppler, and first verified for sound waves

by C. H. D. Buys-Ballot in 1845 from experiments conducted on a moving train.

Acoustical Doppler effect. The Doppler effect for sound waves is now a commonplace experience: If one is passed by a fast car or a plane, the pitch of its noise is considerably higher in approaching than in parting. The same phenomenon is observed if the source is at rest and the receiver is passing it.

Quantitative relations are obtained from a galilean transformation between the system of the receiver, moving with velocity ($\vec{v}$) and that of the source, moving with velocity ($\vec{w}$) with respect to the medium, which may be air taken at rest. In the rest frame the wave propagates with the proper sound velocity given by Eq. (1), where λ and ν are the wavelength

$$|\vec{c}| = c = \lambda\nu \approx 333 \text{ m/s} = 1015 \text{ ft/s} \qquad (1)$$

and frequency measured in that system. Consider propagation of the wave in the direction of the vector ($\vec{c}$), forming angles α and β with the velocities ($\vec{v}$) and ($\vec{w}$) (**Fig. 1**). The frequency ν_R received is by definition the number of wavefronts (wave crests) passing the receiver per unit of time. This is given by calculating the difference between the sound velocity c and the component of the receiver velocity parallel to it $v_{\uparrow\uparrow}$ divided by the wavelength as in Eq. (2).

$$\nu_R = \frac{1}{\lambda}(c - v_{\uparrow\uparrow}) = \nu\left(1 - \frac{v_{\uparrow\uparrow}}{c}\right) = \nu\left(1 - \frac{v\cos\alpha}{c}\right) \qquad (2)$$

A moving source acts differently since it effectively changes the wavelength in the medium, as seen from the following argument: Let the source emit a wave crest at time $t = 0$, traveling during one period of the source oscillation, $T_s = 1/\nu_s$, a distance $c \cdot T_s$. The source follows the wave crest with its velocity component $w_{\uparrow\uparrow}$, parallel to ($\vec{c}$). Therefore it emits the next wave crest at a distance from the first wave crest given by Eq. (3), which is by definition the wave-

$$s = (c - w_{\uparrow\uparrow})T_s = \frac{c - w_{\uparrow\uparrow}}{\nu_s} = \lambda = \frac{c}{\nu} \qquad (3)$$

length λ in the medium. Hence (4) is obtained for the

$$\nu = \frac{\nu_s}{1 - (w_{\uparrow\uparrow}/c)} = \frac{\nu_s}{1 - [(w\cos\beta)/c]} \qquad (4)$$

frequency ν observed in the frame of the medium.

In both cases the transverse velocity components $v_{\uparrow\rightarrow}$ and $w_{\uparrow\rightarrow}$ have no effect since they run parallel to the wavefront. In the supersonic regime where w is greater than c, Eq. (4) holds only inside the Mach cone. *See* SUPERSONIC FLIGHT.

Combining Eqs. (2) and (4) relates the received frequency to the emitted frequency by the general relation of Eq. (5), which may be expanded up to second

$$\nu_R = \nu_s \frac{1 - (v_{\uparrow\uparrow}/c)}{1 - (w_{\uparrow\uparrow}/c)} \qquad (5)$$

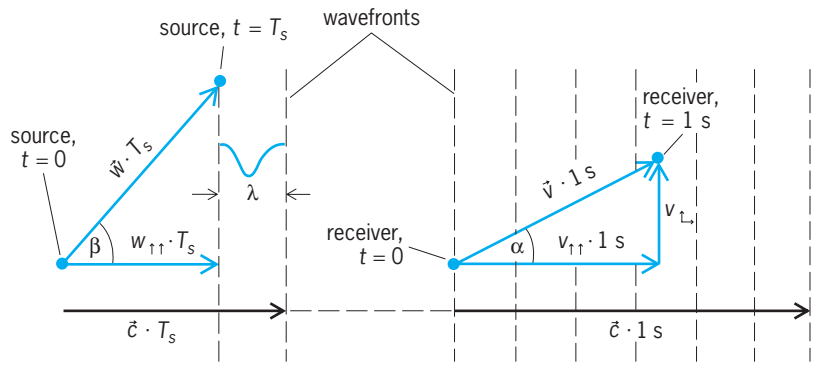

Fig. 1. Motions of source, receiver, and propagating sound that produce the acoustical Doppler effect.

order in the velocities, as in Eq. (6). The first-order

$$\nu_R = \nu_s\left(1 + \frac{w_{\uparrow\uparrow} - v_{\uparrow\uparrow}}{c} + \frac{(w_{\uparrow\uparrow} - v_{\uparrow\uparrow})w_{\uparrow\uparrow}}{c^2} + \cdots\right) \qquad (6)$$

term, the so-called linear Doppler effect, depends only on the relative motion between source and receiver, that is, on Eq. (7), where ($\vec{v}$) is the relative

$$u_{\uparrow\uparrow} = w_{\uparrow\uparrow} - v_{\uparrow\uparrow} \qquad (7)$$

velocity of the source with respect to the receiver, whereas the higher-order terms depend explicitly on the motion of the source relative to the medium. *See* SOUND.

Optical Doppler effect. The linear optical Doppler effect was first observed by J. Stark in 1905 from a shift of spectral lines emitted by a beam of fast ions (canal rays) emerging from a hole in the cathode of a gas discharge tube run at high voltage. Still, their velocity was several orders of magnitude below that of light in vacuum, c_0. This precluded a test of the second-order effect which would be sensitive to any kind of medium carrying light waves, the so-called light ether, if it existed. However, the precise interferometric experiments of A. A. Michelson and E. W. Morley (1887) which are closely related to the above considerations were sensitive to second-order effects. They showed clearly that the velocity of light is not bound to any ether, but is measured to be the same in any moving system. This result was a crucial check for A. Einstein's theory of special relativity (1905), which also makes a clear prediction for the optical Doppler effect with the help of a Lorentz transformation between source and receiver instead of a galilean transformation. This prediction is given by Eq. (8), where ($\vec{u}$) is the relative velocity, and ϑ

$$\nu_R = \nu_s \frac{\sqrt{1 - u^2/c_0^2}}{1 - [(u\cos\vartheta)/c_0]} \qquad (8)$$

is the angle between ($\vec{v}$) and the light propagation velocity ($\vec{c}$), both measured in the system of the receiver (**Fig. 2**). Equation (8) agrees to first order with Eq. (6), but differs in higher orders since it does not

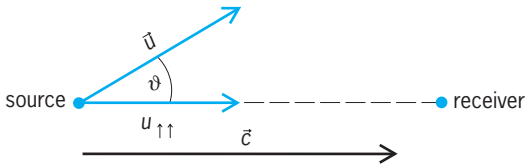

Fig. 2. Motions of source and propagating light, in the system of the receiver, that produce the optical Doppler effect.

contain the source velocity with respect to any ether, but depends entirely on the relative velocity ($\vec{v}$). *See* INTERFEROMETRY.

In contrast to the classical Doppler formula of Eq. (5), the exact relativistic optical Doppler effect of Eq. (8) also exhibits a shift for transverse motion ($\vartheta = 90^\circ$), namely by $\sqrt{1 - u^2/c_0^2}$. This transverse Doppler effect results in a lowering of the received frequency (a redshift), which is equivalent to the relativistic time dilatation in the moving system (a clock runs slower). A verification of the higher-order Doppler effect (and hence of special relativity) to a precision of 4×10^{-5} was obtained recently by two-photon spectroscopy (see below) on a beam of fast neon atoms. Since the first-order Doppler effect cancels in this method, the experiment yielded directly the redshift of the atomic transition in question. *See* RELATIVITY.

Applications. The Doppler effect has important applications in remote-sensing, high-energy physics, astrophysics, and spectroscopy.

Remote sensing. Let a wave from a sound source or radar source, or from a laser, be reflected from a moving object back to the source, which may itself move as well. Then a frequency shift $\Delta\nu$ is observed by a receiver connected to the source, which is given to first order by Eq. (9), where $u_{\uparrow\uparrow}$ is given by Eq. (7) and

$$\Delta\nu = \frac{2\nu_s u_{\uparrow\uparrow}}{c_0} \tag{9}$$

c_0 is the propagation speed of sound or electromagnetic radiation. The factor 2 is due to reflection. The measurement of $\Delta\nu$ is easily carried out with high precision, and provides an excellent means for the remote sensing of velocities of any kind of object, including cars, ships, planes, satellites, flows of fluids, or winds. *See* DOPPLER RADAR; REMOTE SENSING.

High-energy physics. Consider a head-on collision of a laser beam, with wavelength λ of approximately 6×10^{-7} m, with a beam of electrons of very high energy, say 5×10^{10} eV, which is therefore extremely relativistic with v/c approximately equal to $1 - 0.5 \times 10^{-10}$. Then the same considerations as in the case of remote sensing, but calculated relativistically, lead to the conclusion that the light, backscattered by the Compton effect, is changed from visible photons into high-energy gamma rays with a wavelength of approximately 4×10^{-17} m and a quantum energy $h\nu$ (where h is Planck's constant) of approximately 3.2×10^{10} eV. The electron has thus transferred most of its kinetic energy to the light quantum. Such experiments have been performed at high-energy accelerators. *See* COMPTON EFFECT.

Astrophysics. The light from distant stars and galaxies shows a strong Doppler shift to the red, indicating that the universe is rapidly expanding. However, this effect can be mixed up with the gravitational redshift that results from the energy loss which a light quantum suffers when it emerges from a strong gravitational field. *See* COSMOLOGY; GRAVITATIONAL REDSHIFT; REDSHIFT.

The Doppler width and Doppler shift of spectral lines in sunlight (Fraunhofer lines) are important diagnostic tools for the dynamics of the Sun's atmosphere, indicating its temperature and turbulence. *See* ASTRONOMICAL SPECTROSCOPY; SUN.

Doppler-free spectroscopy. Spectral lines from atomic or molecular gases show a Doppler broadening, due to the statistical Maxwell distribution of their velocities. Hence the spectral profile, that is, the intensity I as a function of frequency ν, is a gaussian centered around ν_0, given by Eq. (10), with full halfwidth given by Eq. (11), where k is the Boltzmann constant, T

$$I(\nu) = I(\nu_0) \exp - \left(\frac{\nu - \nu_0}{\delta\nu}\right)^2 \tag{10}$$

$$\delta_{1/2}(\nu) = (2 \ln 2)\delta\nu = (2 \ln 2)\frac{\nu_0}{c_0}\sqrt{\frac{2kT}{m}} \tag{11}$$

is the thermodynamic temperature in kelvins, and m is the mass of a molecule. The Doppler width in Eq. (11) is of order $10^{-6}\ \nu_0$ and exceeds the natural linewidth, given by Eq. (12), by orders of magnitude.

$$\delta\nu_{\text{nat}} = \frac{1}{2\pi\tau} \tag{12}$$

Equation (12) is the ultimate limit of resolution, set by the decay time τ of the quantum states involved in the emission of the line. Thus, precision spectroscopy is seriously hindered by the Doppler width. *See* LINEWIDTH; SPECTROSCOPY.

Around 1970, tunable, monochromatic, and powerful lasers came into use in spectroscopy and resulted in several different methods for completely overcoming the Doppler-width problem, collectively known as Doppler-free spectroscopy. Lasers also improved the spectroscopic sensitivity by many orders of magnitude, ultimately enabling the spectroscopy of single atoms, and both improvements revolutionized spectroscopy and its application in all fields of science. The most common Doppler-free methods are discussed below.

1. *Atomic beams.* An atomic beam, emerging from an oven, well collimated and intersecting a laser beam at right angles, absorbs resonant laser light with the Doppler width suppressed by a factor of $\sin\vartheta$, where ϑ is the collimation angle (**Fig. 3**). This straightforward method was already used before the advent of lasers with conventional light sources. *See* MOLECULAR BEAMS.

2. *Collinear laser spectroscopy on fast beams.* Another way to reduce the velocity spread in one direction is by acceleration. Consider an ion emerging from a source with a thermal energy spread δE, of the order of the quantity kT, or approximately 0.1 eV,

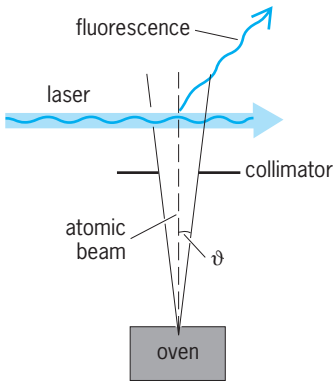

Fig. 3. Doppler-free atomic-beam spectroscopy on thermal atomic beams.

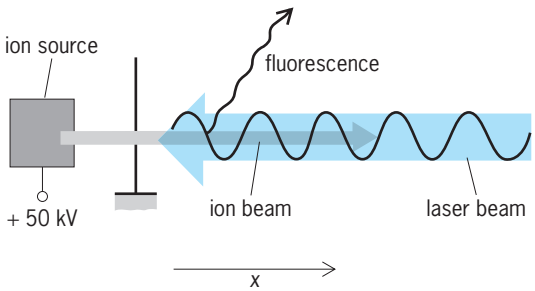

Fig. 4. Collinear laser spectroscopy on accelerated beams.

and accelerated by a potential of, say, 50 kV along the x direction (**Fig. 4**). The kinetic-energy spread in this direction remains constant, however, and may be expanded as in Eq. (13), where the overbars represent

$$\delta E \approx kT \approx \delta\left(\frac{m}{2}\,v_x^2\right) \approx m\overline{v}_x \cdot \delta v_x$$

$$\approx \frac{mc_0^2}{\nu_0^2} \cdot \overline{\Delta\nu} \cdot \delta\nu = \text{constant} \qquad (13)$$

average values. Since the average velocity $(\overline{\vec{u}_x})_x$ has been increased by a factor of 700 in the example given, its spread δv_x, and hence the Doppler width with respect to a collinear laser beam, is reduced by the same factor. Consequently, the product of the Doppler shift $(\overline{\Delta\nu})$ of the fast beam and its Doppler width $\delta\nu$ is a constant of the motion. This method is very versatile and extremely sensitive.

3. *Laser cooling.* The Doppler effect has been used to cool down the thermal motion of atoms or ions by absorption of laser light. For instance, an ion may be trapped in a vacuum by electromagnetic fields (**Fig. 5**). The radial motion is stabilized by a magnetic field B, causing small cyclotron orbits of the ion, whereas axial motion is stabilized by applying a repelling electric potential to the electrodes on the axis. A laser excites the ion at the low-frequency side of the Doppler profile of a resonance line, thus selecting excitations at which the ion counterpropagates to the light. Therefore the ion is slowed down, since it absorbs the photon momentum $h\nu/c$ (where h is Planck's constant) which is opposite to its own. After the ion emits fluorescent radiation, the process is repeated until the ion comes almost to rest.

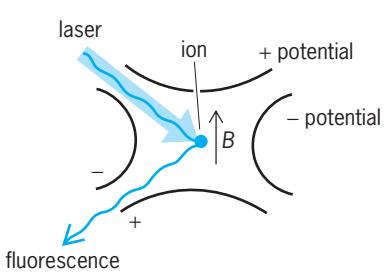

Fig. 5. Laser cooling of a trapped ion.

The repelling force of the cooling process has also been used to confine a cloud of free atoms in vacuum at the intersection volume of three mutually orthogonal pairs of counterpropagating lasers. Thermodynamic temperatures of the order of 10^{-4} kelvin have been reached in such systems. *See* LASER COOLING.

4. *Saturation spectroscopy.* Whereas the velocity distribution of the atoms is manipulated in the above examples, in the following ones this distribution is left untouched, but Doppler broadening is suppressed through the nonlinear optical interaction of the photons with the atoms. The field of nonlinear optics can be roughly characterized by the condition that the atomic system interacts with more than one light quantum during its characteristic decay time τ. In saturation spectroscopy, two almost counterpropagating laser beams of the same frequency ν are used, a strong pumping beam and a weak one (**Fig. 6**). The

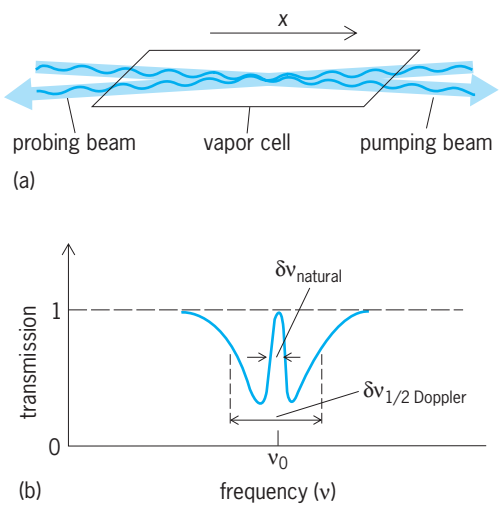

Fig. 6. Saturation spectroscopy. (*a*) Experimental configuration. (*b*) Transmission of probing beam.

pumping beam is strong enough to saturate the excitation of those atoms in the vapor which match the resonance condition, that is, which compensate the frequency offset from the center $\nu - \nu_0$ by the proper velocity component, given by Eq. (14). Saturation means that the excitation is fast as compared to the decay time, leading to an equal population of

$$v_x = c\,(\nu - \nu_0) \qquad (14)$$

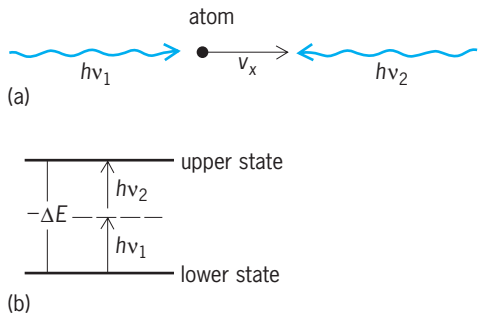

Fig. 7. Two-photon spectroscopy. (*a*) Relation of atom's motion to photons from counterpropagating laser beams. (*b*) Energy-level diagram.

the lower and the upper atomic states. Such a system is completely transparent. The probing beam feeds on atoms with exactly the opposite velocity component given by Eq. (15). It is too weak to sat-

$$-v_x = c(\nu - \nu_0) \tag{15}$$

urate, and is absorbed in the vapor except at the center frequency $\nu = \nu_0$, where it matches with the saturated part of the velocity spectrum. Resolutions of $\delta\nu/\nu_0 < 10^{-10}$ obtained in this way make possible improved values of fundamental constants and of the standards of length and time. *See* FUNDAMENTAL CONSTANTS; LENGTH; LIGHT; PHYSICAL MEASUREMENT; RYDBERG CONSTANT.

5. *Two-photon spectroscopy.* In this method, two counterpropagating laser beams are operated at half the transition frequency ν_0 of the atomic system. Then Bohr's quantum condition, that the transition energy ΔE equals $h\nu_0$, can be fulfilled by simultaneous absorption of two photons, say, one from each beam (**Fig. 7**). Including the first-order Doppler effect, the energy balance is then given by Eq. (16).

$$\begin{aligned} h\nu_1 + h\nu_2 &= h\nu\left(1 + \frac{v_x}{c}\right) + h\nu\left(1 - \frac{v_x}{c}\right) \\ &= 2h\nu = h\nu_0 \end{aligned} \tag{16}$$

Thus, the first-order Doppler effect cancels completely, and the two-photon resonance occurs sharply at half the transition frequency. *See* ATOMIC STRUCTURE AND SPECTRA; LASER SPECTROSCOPY; NONLINEAR OPTICS.

Ernst W. Otten

Bibliography. W. Demtroder, *Laser Spectroscopy: Basic Concepts and Instrumentation*, 2d ed., 1996; A. Einstein, *The Principle of Relativity*, 1923; D. Halliday, R. Resnick, and K. Krane, *Physics*, 4th ed., 1992; K. Shimoda, *Introduction to Laser Physics*, 2d ed., 1986; H. J. Smith, *Basic Doppler Physics*, 1991.

Doppler radar

A radar system used to measure the relative velocity of the system and the radar target. The operation of these systems is based on the fact that the Doppler frequency shift in the target echo is proportional to the radial component of target velocity. *See* DOPPLER EFFECT.

Airborne vehicular systems. Airborne systems are used to determine the velocity of the vehicle relative to the Earth for such purposes as navigation, bombing, and aerial mapping, or relative to another vehicle for fire control or other purposes. Ground or ship equipment is used to determine the velocity of vehicular targets for fire control, remote guidance, intercept control, traffic control, and other uses.

The Doppler frequency shift Δf is an extremely small fraction of the transmitter frequency f. It is given by the equation shown below,where V is the

$$\Delta f = \frac{2Vf}{C}\cos\gamma$$

relative speed, C is the speed of signal propagation, and γ is the angle between the velocity and the direction of propagation (**Fig. 1**). The only practical way to measure the frequency shift is by adding the echo signal to a reference signal derived from the transmitter and observing the difference, or beat frequency. Some means of obtaining coherent detection is required.

Practical techniques have been devised for obtaining the requisite coherence in continuous-wave (CW), pulsed, and frequency-modulated transmission systems. For a discussion of techniques similar to those in Doppler radar *see* CONTINUOUS-WAVE RADAR; MOVING-TARGET INDICATION.

Doppler navigation radar is a type of airborne Doppler radar system for determining aircraft velocity relative to the Earth's surface. Such a Doppler velocity sensor (**Fig. 2**) consists of at least the following elements: transmitter, antenna, receiver, Doppler frequency measuring device, and output signal generators or displays. It is generally used with a navigation computer. *See* AIRBORNE RADAR.

The signal from a single beam can provide only the velocity component in the direction of that beam. Complete velocity determination requires, therefore, the use of at least three beams. Most systems use four beams for symmetry (**Fig. 3**).

To relate the beam directions, and hence the

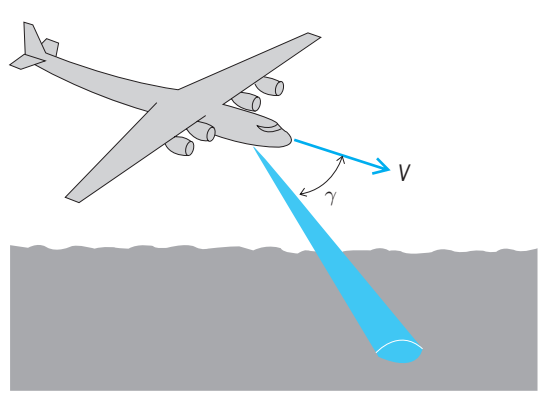

Fig. 1. Doppler frequency measurement geometry.

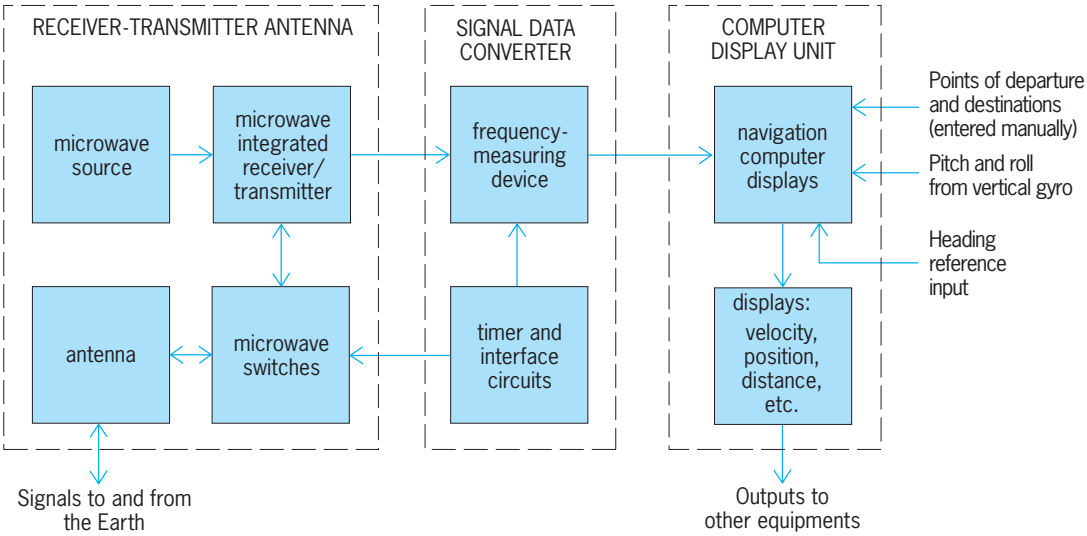

Fig. 2. Block diagram of aircraft Doppler navigation system, the AN/ASN-128.

measured velocity to an Earth-oriented coordinate system, a vertical reference must be provided. In newer systems the antennas are fixed to the aircraft. The Doppler frequencies and the vertical data, such as roll and pitch, are fed to a computer. Its outputs are electrical signals or displays representing the components of velocity. The components commonly displayed are those along-heading, across-heading, and vertical.

Various types of antennas have found use in Doppler navigation radar. These include paraboloids, microwave lenses, and linear and planar arrays. Since both volume and radome cutout area should be small, various techniques are employed to enable each antenna to form (simultaneously or sequentially) more than one beam. Pencil beams of 3–5° width are used. The beams are directed 15–25° from the vertical. Larger values result in insufficient echo power over water. *See* ANTENNA (ELECTROMAGNETISM).

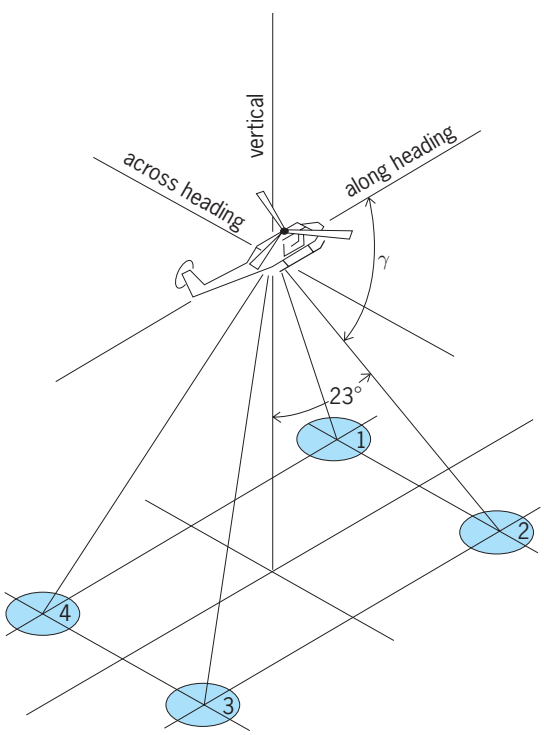

Fig. 3. Typical antenna beam arrangement.

Continuous-wave systems are coherent and are theoretically the most efficient. The chief difficulty is control of leakage of spurious signals from transmitter to receiver.

Pulsing enables the receiver to be rendered insensitive during transmission, thereby avoiding leakage signals. Coherence is achieved either by driving a transmitting power amplifier from a CW oscillator or by mixing at the detector a pair of pulse echoes received over different propagation paths.

A preferred technique is to employ sinusoidal frequency modulation. A sideband of the detected beat between echo and transmitter signal is used. Modulation index and rate and the sideband order are chosen such that echoes from nearby objects are rejected, while those from distant objects are accepted. Leakage noise is reduced at the expense of lowered efficiency.

An example of a system (**Fig. 4**) is the U.S. Army's standard airborne Doppler navigator which bears the nomenclature AN/ASN-128. It weighs 28 lb (12.7 kg), requires 89 W of power, and operates at 13.325-GHz frequency. In addition to velocities, outputs include present position, distance to go, desired track and bearing to destination, and steering signals. For helicopter use a steering hover indicator can also be provided. The system uses all-solid-state electronics and a printed grid planar array antenna. The velocity accuracy specifications are around a quarter of a percent of ground speed (and 0.2 knot or 0.1 m/s when hovering). The navigation accuracy specification is 1.3% of distance traveled when using a 1° accuracy heading reference.

France B. Berger

Fig. 4. AN/ASN-128 Doppler navigation system. (*a*) Receiver-transmitter antenna. (*b*) Computer display unit. (*c*) Steering hover indicator. (*d*) Signal data converter. (*Singer Co., Kearfott Division*)

Meteorological systems. Doppler radars measure the radial velocity of targets by processing the Doppler frequency shift associated with this velocity, and therefore they are well suited for the observation of the motion of raindrops in precipitation systems. However, precipitation has the form of a distributed target, that is, numerous independent scatterers distributed in space. The backscattered signal, selected by a sampling circuit called the range gate, is thus composed of a large number of separate scattering amplitudes, each having a Doppler frequency shift associated with the radial velocity of a particular scatterer. Therefore, a spectrum of frequency shifts (Doppler spectrum) is observed.

To obtain the Doppler spectrum, the radar signal sampled at each range gate must be submitted to a Fourier transform operation. This is typically done by digitizing that signal and implementing fast Fourier transform algorithms. Spectral moments such as mean Doppler velocity and spectrum variance are also calculated. By using minicomputers equipped with specialized cards or hardware systems based on integrated circuits, fast Fourier transforms can be performed continuously at a large number of range gates whose spacing essentially matches the radar range resolution (a few hundred meters). However, much simpler algorithms are widely used to obtain the spectral moments by direct processing of time signals. These algorithms, called pulse-pair algorithms, proceed by computing the signal autocovariance whose argument is the mean Doppler frequency normalized to the frequency sampling domain established by the radar pulse repetition frequency. *See* ALGORITHM; HARMONIC ANALYZER.

Design. By using suitable phase coherence and signal-processing techniques, the Doppler shift can be extracted from the radar signal and interpreted in terms of the target's radial velocity. The sampling of the backscattered signal at each range gate is done at the radar pulse repetition frequency, and therefore unambiguous Doppler frequency observations are made only in frequency range allowed by the sampling rate. Since the Doppler frequency shift for a given radial velocity is inversely proportional to the radar wavelength, the unambiguous velocity domain for a given pulse repetition frequency increases in proportion to this wavelength. For example, a 1-kHz pulse repetition frequency (corresponding to an unambiguous range of about 300 km or 186 mi) yields an unambiguous velocity domain of ± 25 m/s (56 mi/h) at a wavelength of 10 cm compared to ± 8 m/s (18 mi/h) at a 3.2-cm wavelength. For long-range precipitation monitoring, it is thus advantageous to use a long radar wavelength. However, the radar beamwidth for a given antenna size increases with radar wavelength, so that, for an antenna of practical size, the choice of wavelength is limited to 10 cm or less. A typical 10-cm-wavelength Doppler radar such as the one proposed in the NEXRAD (next generation radar) program is equipped with a 10-m (33-ft) dish antenna, producing a 0.7° beamwidth. For that radar, a 2-kHz pulse repetition frequency unambiguous range of 150 km or 93 mi) yields a ± 50 m/s (112 mi/h) unambiguous velocity domain. By using velocity dealiasing algorithms that resolve first- and second-order velocity ambiguities, the effective velocity domain can be increased by ± 100 m/s (224 mi/h), which practically covers all the possible wind values encountered in precipitation systems.

Single radar. Conventional meteorological radars are being gradually replaced by Doppler radars, such as those of the NEXRAD program implemented by the U.S. Weather Service. In the new radars, signal processing, data treatment, and data display are performed by a computer dedicated to the system. An example of the usefulness of a single Doppler radar is the measurement of mean wind field as a function of altitude, assuming uniform wind conditions in the area scanned by the radar beam, from processing of radial velocity-range profiles acquired by scanning the radar beam continuously in azimuth. The NEXRAD network will also aid in the detection and monitoring of severe storms and tornadoes and the tracking of hurricanes, and in specifying conditions for wind shear or microburst occurrence in convective storms in an effort to improve air traffic safety.

Multiple radars. The joint operation of two or more Doppler radars simultaneously observing the same convective storm enables the measurement of the three components of the velocity of precipitation particles. While the horizontal motion of the

precipitation particles is identified with horizontal air motion, particle vertical velocity is the sum of two terms: fall speed of the particle in still air (terminal velocity) and vertical air velocity. If an estimate of particle terminal velocity is made (from radar reflectivity, for instance), then the three components of air velocity can be derived. The assumption that air compressibility is negligible also offers a means of deriving vertical air velocity from the divergence of the horizontal velocity field. Multiple Doppler radar networks have been used extensively in convective storm research and have contributed to rapid advances in the understanding of convective storm development. *See* CALCULUS OF VECTORS; FLUID-FLOW PRINCIPLES.

Airborne radars. Meteorological Doppler radars have been installed aboard airplanes for the purpose of measuring wind fields in storm systems which cannot be reached by ground-based radars. The simplest method of operation relies on the radar beam being perpendicular to the airplane course and scanning slightly in the vertical plane. One of the most striking examples of this method is the probing of the wind field inside a hurricane. The hurricane evolves slowly and thus can be probed from different directions by the radar in a time interval during which its evolution is negligible. Several radial-velocity "snapshots" of the same motion field can then be combined to obtain the horizontal motion field of different altitude levels. This method has been used primarily as a research tool and has produced unique observations of the kinematics of hurricanes which will ultimately help to improve their modeling and trajectory prediction. *See* HURRICANE.

Millimeter-wave radars. Centimeter-wave radars are excellent tools for the probing of precipitation, but they fail to detect and observe clouds, in which particle size is below approximately 100 micrometers. Since the amount of energy returned (backscattering) by cloud droplets increases with the inverse fourth power of the radar wavelength, millimeter-wave radars can overcome this weakness. For example, the intensity of the echo returned by the same cloud droplet at a 3.2-mm wavelength is 10,000 times greater than that at a 3.2-cm wavelength. Modern development of millimeter-wave technology has produced Doppler radars operating at very short wavelengths (2–3 mm). These radars are excellent tools for the probing of cloud kinematics and liquid water, and have already produced results applicable to modeling of small cumulus clouds, for instance. They are also useful for continuous monitoring of cloud layers by providing observations of cloud water and vertical velocity. Since they produce very narrow beams with small antenna size and because the enhanced backscattering drastically reduces the need for transmitted power (and thereby power supply), millimeter-wave Doppler radars are excellent candidates for installation aboard satellites in an effort to observe motion in precipitation systems at a global scale. *See* CLOUD PHYSICS.

Doppler lidars. Doppler techniques have been applied to optical lidars (light detection and ranging). These systems operate in a manner similar to pulse Doppler radars by sending bursts of optical energy. The Doppler lidars now being used for meteorological research operate at a 10.6-micrometer wavelength. Because of the strong attenuation of infrared radiation by hydrometeors, their use is restricted to observing cloud boundaries or aerosol and pollution layers. Instead of relying on the pulse-to-pulse coherence type of operation common to pulse Doppler radars, the Doppler frequency is extracted in time intervals equal to the transmitted pulse duration and sampled from the received signal associated with one transmitted light pulse. This processing is repeated independently for several light pulses and associated echoes, with the average of these estimates calculated so as to improve the accuracy of the measurements. Only mean Doppler frequency can be extracted by using time-domain algorithms similar to those used in the pulse-pair technique discussed above, but designed for the computation of signal autocovariance within a time interval equal to a transmitted pulse duration.

Doppler lidars were initially implemented as ground-based systems operated in a vertically pointing beam mode, but they were later installed aboard airplanes. This method of operation has produced wind field measurements in the lower troposphere based on observing aerosol motion. Their use aboard satellites for measurement of the wind field has been proposed in regions where aerosol density is expected to produce sufficient backscattering intensity. *See* LIDAR; METEOROLOGICAL RADAR; RADAR; RADAR METEOROLOGY.

Roger M. Lhermitte

Bibliography. R. J. Doviak and D. S. Zrnic, *Doppler Radar and Weather Observations*, 2d ed., 1993; M. Kayton (ed.), *Navigation*: *Land, Sea, Air, and Space*, 1990; G. V. Morris and L. Harkness, *Airborne Pulsed Doppler Radar*, 2d ed., 1996; H. Sauvageot, *Radar Meteorology*, 1992; D. C. Schleher, *MTI and Pulsed Doppler Radar*, 1991.

Doppler VOR

A very high-frequency omnidirectional radio range (VOR) employing the Doppler principle. The bearing accuracy achieved by users when VOR beacons are installed in the vicinity of obstructions, or when aircraft using VOR signals fly over mountainous terrain, is deteriorated by reflections (site and enroute errors). The Doppler VOR solves this problem by two fundamental principles maintaining full compatibility of the radiated information with existing airborne receivers. It uses a wide-base antenna array (**Fig. 1**) for suppressing the effects of multipath propagation and the Doppler principle for determination of bearing. According to the Federal Radionavigation Plan of 2001, approximately half of the existing VOR stations in the United States will be phased out beginning in 2010. However, Doppler VOR will be installed

Fig. 1. Collocation of Doppler VOR and TACAN beacons installed on 100-foot high (30-m) tower with a counterpoise of 119-ft (36-m) diameter in forest near Frankfurt, Germany. Circular array consists of 39 loop antennas fed by solid-state commutator. Center carrier radiator is covered with plastic dome, which supports TACAN antenna.

in those remaining VOR sites where performance is degraded by multipath propagation. *See* VOR (VHF OMNIDIRECTIONAL RANGE).

Disturbances of VOR signals received by aircraft that are caused by multipath propagation can be reduced by wide-aperture systems because of integration or averaging between the information transferred over the direct and the reflected propagation path. In other words, when the reflections are weaker than the direct path signal, their effect tends to average out over intervals that are long compared to the multipath delay. Wide-base systems, however, have the drawback that their directional information becomes ambiguous as soon as the aperture is larger than half a wavelength, approximately.

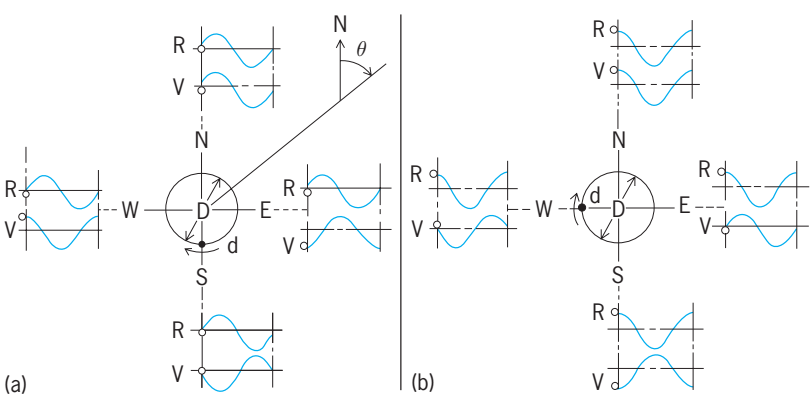

Fig. 2. Principle of Doppler VOR beacon. Reference signal R and bearing signal V which is generated by dipole d are shown as they would be measured in aircraft to the north, east, south, and west of the beacon. (*a*) Dipole d at 180°. (*b*) Dipole d at 270°. Phase difference between R and V gives bearing θ.

By applying the Doppler principle, one can overcome these limitations. A dipole rotating eccentrically on an orbit periodically changes the distance between transmitter and receiver. This periodic change of distance modulates the frequency of the radiated carrier signal by a deviation f_D according to the equation below. Here D is the diameter of the array and

$$f_D = \frac{\text{velocity}}{\text{wavelength}} = \frac{\pi D f_{\text{rot.}}}{\lambda}$$

$f_{\text{rot.}}$ is the number of rotations, or revolutions, per second. The resulting Doppler effect produces a sinusoidal frequency modulation (FM) of the carrier frequency; the phase of this FM contains the bearing information (variable phase).

In 1958 the U.S. Civil Aeronautics Agency constructed the first experimental Doppler VOR station in Indianapolis, Indiana. A nondirectional antenna radiates a radio-frequency (rf) carrier, amplitude-modulated with the rotational frequency, thus transmitting the reference information (**Fig. 2**). Another frequency, spaced by 9960 Hz from the carrier, corresponding to the subcarrier frequency of the conventional VOR, is radiated by an antenna which is in the form of a circle around the carrier antenna located in the center. The pattern generated by this antenna appears to rotate. Actually this circular motion is simulated by sequential feeding of radiators (Alford loops) arranged in a circular array. In contrast to the conventional VOR, the roles of the reference information (R) and the bearing information (variable signal V) are exchanged. This does not affect the existing airborne receiver functions.

For a bearing of 90° (aircraft east of ground station), for example, the phase difference is always 90°. When the reference signal for this bearing in Fig. 2*a* passes through zero, the variable signal is at its minimum because the dipole moves away from the observer in the east. When the dipole has moved by 90° on the circle, Fig. 2*b*, the variable signal reaches zero in the east because the motion is perpendicular to the east direction. Meanwhile the reference signal has increased to its maximum value. This mode of operation corresponds to single-sideband (SSB) modulation where the carrier is amplitude-modulated with the reference signal and its sideband is frequency-modulated with the variable signal which is determined by the instantaneous radiator position.

Certain disadvantages of this arrangement can be eliminated by the double-sideband (DSB) Doppler VOR. Here, the upper and lower sideband, spaced by 9960 Hz from the carrier, respectively, feed two radiators which are located diametrically opposite on the circular antenna array. Both sideband radiations are moved in steps from antenna to antenna along the circle. The signals radiated are now almost identical with those from conventional VOR ground stations. The carrier is radiated from the central antenna, as in the case of the SSB system.

The alternating-sideband (ASB) method is a variant

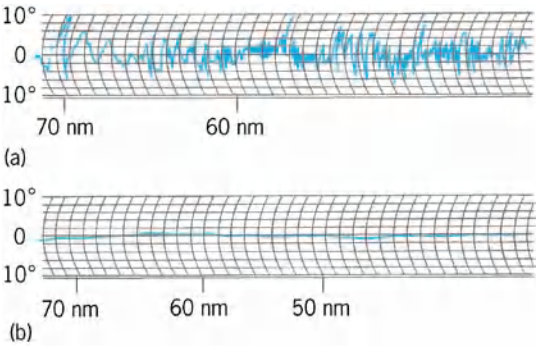

Fig. 3. Comparison of radial flight recordings. (*a*) From VOR installation. (*b*) From DVOR installation. Both experimental beacons were located at same unfavorable site. Plane is flying constant course at bearing of 133° from beacons at 4000-ft (1.2-km) altitude. Improvement of Doppler VOR is obvious. Radial flight recordings revealed for conventional VOR course some fluctuations which exceeded ICAO limits; maximum deviation of DVOR was approximately 0.5°.

of the double-sideband principle. This beacon is provided with an odd number of antennas on a circle; therefore no antenna is diametrically opposite to another. Upper and lower sideband interrupted carrier-wave trains are alternately fed to the nearly opposite antennas so that the energy travels at a constant speed around the circle. This system uses only half as many antennas as the array of a DSB system, and some difficulties are avoided concerning mutual coupling between adjacent antennas. In contrast to an SSB Doppler VOR which uses an antenna spacing of less than half a wavelength, the ASB array employs elements spaced by two-thirds a wavelength or more.

The Doppler VOR presents two essential advantages over conventional VOR: the improvement of absolute accuracy in good and poor terrain, and the spectacular radial stability and smoothness even over mountains (**Fig. 3**).

Ernst Kramar; Richard L. Greenspan

Bibliography. M. Kayton (ed.), *Navigation: Land, Sea, Air, and Space*, 1990; G. W. Mixter, *Primer of Navigation*, ed. by H. Headley, 7th ed., 1995; G. J. Sonnenberg, *Radar and Electronic Navigation*, 6th ed., 1988.

Dormancy

In the broadest sense, the state in which a living plant organ (seed, bud, tuber, bulb) fails to exhibit growth, even when environmental conditions are considered favorable. In a stricter context, dormancy pertains to a condition where the inhibition of growth is internally controlled by factors restricting water and nutrient absorption, gas exchange, cell division, and other metabolic processes necessary for growth. By utilizing the latter definition, dormancy can be distinguished from other terms such as rest and quiescence which reflect states of inhibited development due to an unfavorable environment.

This article deals with dormancy in a wide array of plant structures. However, the main focus centers on the regulation of seed and bud dormancy. This does not preclude the fact that the plant kingdom consists of a vast collection of other specialized plant parts whose dormancy may be governed by more refined systems than those that fall within the scope of this article.

Physical influences. Physically induced dormancy can be separated into two distinct classes, based on external conditions imposed by the environment (light, temperature, photoperiod) and restraints induced by structural morphology (seed-coat composition and embryo development).

External conditions. The physical environment plays a key role in dormancy induction, maintenance, and release in several plant species.

1. *Temperature.* The onset of dormancy in many temperate-zone woody species coincides with decreasing temperature in the fall. However, it is the chilling temperature of the oncoming winter which is more crucial, particularly in regard to spring budbreak. The range of chilling requirements varies considerably, not only between different species, but also within a given species. In peach (*Prunus persica*), for example, chilling requirements among various cultivars range from 350 to 1200 h of exposure to temperatures below 45°F (7°C). Different portions of a plant can also differ in their particular chilling requirements, as in the case of fruit crops (for example, peach, apple, and cherry) where floral buds break much earlier than those which give rise to the new foliage.

Seeds of many diverse plant species, such as birch (*Betula* spp.), ash (*Fraxinus excelsior*), and grape (*Vitis vinifera*), also require a certain period of chilling, often coupled with adequate moisture to alleviate a dormant state. This process, which utilizes both cold temperature and moisture, is termed stratification and relates to the manipulation of endogenous factors affecting dormancy.

2. *Light duration and quality.* Possibly the single most important environmental variable affecting dormancy is day length or photoperiod. In plants capable of dormant bud production, such as temperate-zone woody trees, long photoperiods exceeding 12 h favor a continuation of growth, whereas short day lengths usually trigger dormancy. The same photoperiodic response occurs in tuberous crops, such as the Jerusalem artichoke (*Helianthus tuberosus*) and the potato (*Solanum tuberosum*), where short days induce tuber formation which eventually results in a dormant condition.

In other plants (for example, onion bulbs) long, hot days during summer droughts actually induce dormancy. Light quality and intensity also play an important role in dormancy in addition to day length. In certain lettuce cultivars there is a specific requirement for red light to break dormancy and induce germination. *See* PHOTOPERIODISM.

3. *Water and nutrient status.* Dormancy is affected by the availability of water and nutrients as demonstrated by many grasses, desert species, and subtropical fruits which go into dormancy when

confronted by drought or lack of soil fertility. Withholding of mineral fertilization is one practice utilized to induce tuberization and dormancy in tuberous crops. Conversely, continuous fertilization can postpone dormancy in several ornamental shrubs. In the case of dormant peach seeds, nitrogen and sulfur-containing compounds such as thiourea induce germination. *See* FERTILIZER; PLANT MINERAL NUTRITION; PLANT-WATER RELATIONS.

4. *Environmental interactions.* Several of the factors previously discussed do not simply act independently, but combine to influence dormancy. The temperature range for budbreak and seed germination, for instance, is regulated by day length. In certain situations long days actually replace the need for low temperatures in inducing seed germination. In other situations optimal temperatures or lighting regimes remove dormancy in seeds even under unfavorable moisture conditions.

Physical constraints. Examples of dormancy imposed by physical restrictions are most evident in the structural morphology of dormant seeds. These restrictions specifically pertain to the physical properties of the seed coat and developmental status of the embryo.

1. *Seed-coat factors.* The seed-coat material surrounding embryos of many plants consists of several layers of tissue, termed integuments, which are infiltrated with waxes and oils. In effect these waterproofing agents enable the seed coat to inhibit water absorption by the embryo. This results in a type of seed dormancy very characteristic of legume crops (clover and alfalfa). In the cocklebur (*Xanthium strumarium*), seed-coat-induced dormancy is imposed by oxygen impermeability. The environment itself can break this type of seed-coat dormancy through alternating temperature extremes of freezing and thawing. The extreme heat induced by forest fires is especially effective.

Fungal and bacterial infection of the seed coat can also lessen seed-coat-induced dormancy by reducing impermeability to oxygen, carbon dioxide, and water. Seed-coat dormancy can be artificially overcome through scarification, a process in which physical abrasion of the seed coat is accomplished through mechanical (sand or carborundum powder) or chemical (sulfuric acid) means.

Seed-coat-induced dormancy can also result from mechanical resistance due to extremely hard, rigid integuments commonly found in conifer seeds and other tree species with hard nuts (for example, macadamias and filberts). This dormancy may persist even though the seed coat may be hydrated. Once the seed coat dries out and is subsequently rehydrated, germination will usually follow. Other mechanisms previously described can also alleviate this type of seed-coat-induced dormancy.

2. *Embryonic factors.* The morphological state of the embryo is yet another physical factor affecting dormancy. Often the embryo is in a rudimentary stage when the seed is shed from the maternal plant, as exemplified by members of the orchid family and certain deciduous trees, such as ginkgo and European ash. Dormancy will usually cease in these plants as the embryos reach an adequate state of maturation.

Embryos can be physiologically mature, yet still remain dormant within the seed due to the lack of a metabolite essential for growth. These embryos are referred to as physiologically immature, and examples exist throughout the Rosaceae family and temperate-zone woody species. Dormancy in this instance can be broken by exposure to temperatures just above freezing or through commercially applied stratification. Such processes permit afterripening of the embryo, which leads to subsequent germination.

Endogenous regulation of dormancy. Studies dealing with dormancy have resulted in searches for endogenous plant hormones which regulate the process. In fact, the discovery of several plant growth regulators evolved from the close scrutiny of endogenous compounds derived from dormant structures.

Dormancy inducers. Early studies involving dormant buds of ash (*Fraxinus americana*) and birch (*Betula pubescens*) revealed the presence of high concentrations of a growth inhibitor or dormancy-inducing and -maintaining compound. This compound was later identified as abscisic acid. As buds of these trees began to grow and elongate, the levels of abscisic acid fell appreciably, supporting the role for abscisic acid in the regulation of dormancy.

Abscisic acid is also important in the regulation of seed dormancy, as exemplified by seeds of ash in which abscisic acid levels are high during the phase of growth inhibition, but then decline rapidly during stratification, resulting in germination. *See* ABSCISIC ACID.

Dormancy-releasing agents. Evidence from buds and seeds of other plant species, including apple (*Malus*), black currant (*Ribes nigrum*), and willow (*Salix fragilis*), show that abscisic acid levels drop sharply as dormancy is broken. In conjunction with decreased levels of abscisic acid, the endogenous supply of many growth promoters, such as gibberellins, cytokinins, and auxins, have been reported to rise during budbreak in sycamore (*Acer pseudoplatanus*) as well as in Douglas fir (*Pseudotsuga menziesii*). Levels of these dormancy-releasing compounds also correlate well with the breaking of seed dormancy. Dormant filbert seeds (*Corylus avellana*) usually require a period of chilling to break dormancy. Endogenous levels of gibberellic acid in these seeds show a marked increase in direct proportion to the degree of chilling. In fact, exogenous application of gibberellic acid can substitute for the chilling requirement. *See* AUXIN; CYTOKININS; GIBBERELLIN.

Cytokinins (substituted adenyl compounds) and auxins (for example, indole acetic acid, or IAA) are rather dualistic in nature, making it difficult to categorize them either as dormancy inducers or inhibitors. Exogenous cytokinins, for example, can overcome the inhibition of axillary bud dormancy. The breaking of bud dormancy in aspen (*Populus*) coincides with an increase in cytokinin activity. Although cytokinins are usually regarded as dormancy inhibitors, they can also induce tuberization which

eventually results in the production of a dormant structure (potato tuber). Auxins are also difficult to evaluate in terms of dormancy since they retard growth in several cases (for example, apical dominance in cereal grasses and in conifers) and promote it in others (tissue culture of embryos from various plant species). *See* APICAL DOMINANCE.

Exogenous regulation of dormancy. In addition to endogenous hormones, there are a variety of compounds that can break dormancy in plant species when they are applied exogenously. Many of these substances are synthetic derivatives or analogs of naturally occurring, dormancy-releasing agents. For example, the synthetic cytokinin 6-benzyladenine can stimulate growth of quiescent citrus (*Citrus aurantifolia*) buds by means of external application, and foliar sprays of growth inhibitors, such as the sterols paclobutrazol and fluripimidol, can advance vegetative and floral budbreak in the peach (*Prunus persica*). The effect of the application of exogenous hormones is, however, very dependent on the genotype.

Synthetic chemicals can also be used in conjunction with naturally occurring compounds to enhance or break dormancy. A combination of 6-benzyladenine and the gibberellins GA_4 and GA_7 is effective in releasing dormant apple buds, whereas either chemical alone is relatively ineffective. Such chemical treatments are important in enhancing budbreak that can be greatly reduced when mild winters result in a lack of adequate chilling to trigger the production of endogenous dormancy-releasing hormones.

Mineral salts can also be beneficial in partially substituting for the natural chilling requirement to break dormancy. This rest substitution effect has been noted for foliar applications of potassium nitrate and thiourea, which selectively force apple flower and vegetative buds, respectively.

Interaction of dormancy-related compounds. The hormonal regulation of dormancy can best be perceived as a balance between dormancy inducers or maintainers and dormancy-releasing agents. This has already been discussed in studies involving fluctuating levels of abscisic and gibberellic acids in buds from several tree species. The same situation prevails in seeds where growth is closely regulated by the balance between different hormones and not necessarily by their absolute concentrations.

In pea (*Pisum sativum*) growth of the seed correlates well with hormonal fluctuations of abscisic, gibberellic, and indole acetic acids (**Fig. 1**). In this instance gibberellic acid and indole acetic acid levels rise with an increase in fresh weight and nutrient uptake by the embryo. The decline in seed growth coincides with a sharp rise in abscisic acid content and concomitant drop in gibberellic and indole acetic acids. *See* PLANT HORMONES.

Environmental versus internal factors. The physical environment exerts a marked influence on dormancy. The plant, however, needs a receptor system to perceive changes in the environment so it can translate them into physiological responses which in most cases are under hormonal control. In the case of changing day length or photoperiod, phytochrome may serve as a receptor pigment. Phytochrome essentially favors the production of either abscisic acid (short days) or gibberellic acid (long days; **Fig. 2**). Stress conditions, such as limited water or nutrient availability, favor the production of

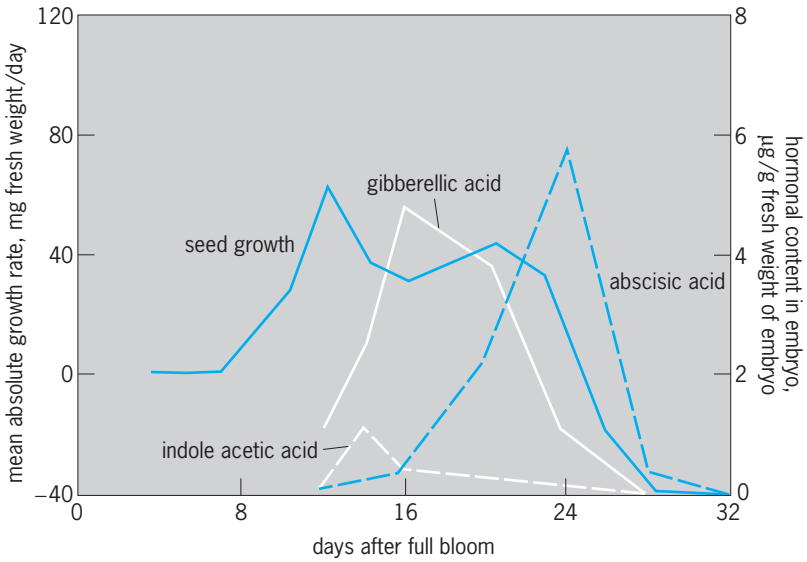

Fig. 1. Seed growth and hormonal levels in an embryo of *Pisum sativum*. Relative to indole acetic acid content, abscisic acid and gibberellic acid levels are 5 and 100 times less, respectively. (*After C. J. Eeuwens and W. W. Schwabe, Seed and pod wall development in Pisum sativum in relation to extracted and applied hormones, J. Exp. Bot., 26:1–14, 1975*)

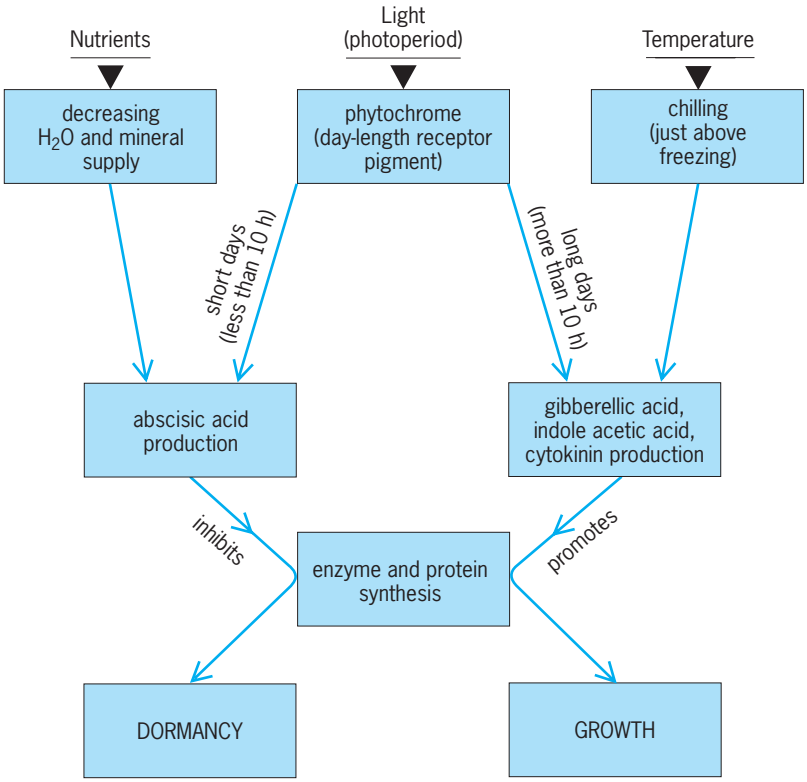

Fig. 2. Generalized scheme depicting the influence of the environment on endogenous pools of dormancy inducers and maintainers (abscisic acid) and dormancy releasers (gibberellic acid, indole acetic acid, and cytokinins), which in turn regulate dormancy via an effect on enzyme and protein synthesis in general.

abscisic acid, whereas a period of chilling often promotes synthesis of gibberellic acid and other compounds generally considered as growth promoters. The hypothesis of a receptor system such as phytochrome controlling the balance between levels of abscisic acid and gibberellic acid is intriguing, especially since the synthesis of these physiologically interactive hormones is probably associated with a common metabolic pathway originating with mevalonic acid. In this manner a definitive link between the environment and physiological changes (mediated via pools of endogenous hormones) could be established. *See* PHYTOCHROME.

Such endogenous pools are environmentally influenced and could in turn control dormancy physiologically. However, the mode of action of these endogenous growth regulators can only be postulated at this time. Whatever the specific mechanism, it probably involves the regulation of gene action at the level of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), which subsequently controls protein synthesis. In this framework, abscisic acid is believed to repress the functioning of nucleic acids responsible for triggering enzyme and protein synthesis needed for growth (Fig. 2). Gibberellic acid, on the other hand, promotes synthesis of enzymes essential for germination as in the case of α-amylase production that is crucial for barley seed growth. *See* BUD; NUCLEIC ACID; PLANT GROWTH; SEED.

Craighton S. Mauk

Bibliography. J. D. Bewley and M. Black, *Seeds: Physiology of Development and Germination*, 2d ed., 1994; R. P. Marini, Growth and cropping of Redhaven peach trees following foliar applications of fluriprimidol and paclobutrazol, *J. Amer. Soc. Hort. Sci.*, 111(6):849–853, 1986; J. D. Metzger, Role of endogenous plant growth regulators in seed dormancy of *Avena fatua*, *Plant Physiol.*, 73:791–795, 1983; T. C. Moore, *Biochemistry and Physiology of Plant Hormones*, 2d ed., 1989; E. M. Nauer and S. B. Boswell, Stimulating growth of quiescent citrus buds with 6-benzylamino purine, *HortScience*, 16:162–163, 1981; A. D. Shaltout and C. R. Unrath, Effect of some growth regulators and nutritional compounds as substitutes for chilling of Delicious apple leaf and flower buds, *J. Amer. Soc. Hort. Sci.*, 108(6):898–901, 1983.

Dormouse

The name applied to the 31 species of Old World rodents of the family Gliridae. These animals are intermediate between squirrels and rats; they are vegetarians and arboreal like the former, but most have a general appearance like the latter.

The fat dormouse (*Glis glis*; see **illus.**), the largest member of the family, is found throughout Europe and western Asia. It resembles the squirrel because of its bushy tail, grows to a length of 8 in. (20 cm), including a 3-in. (7.5-cm) tail, and like other members of the family has 20 teeth, with a dental formula I 1/1 C 0/0 Pm 1/1 M 3/3. The fat dormouse prefers deciduous or mixed forests as a habitat; it hides in the forest during the day and forages during the night for fruits. It eats large quantities of food all summer and grows fat in preparation for winter, when it hibernates and sleeps until spring. These animals were bred in captivity by the Romans as a food delicacy.

The fat dormouse, *Glis glis*.

The only species of dormouse native to Britain is the common dormouse (*Muscardinus avellanarius*), which is about the size of a house mouse. Its front legs are much shorter than the hindlegs, the thumbs are vestigial, and the big toes are reduced in size or rudimentary. Another species, found in central and southern Europe, is the garden dormouse (*Eliomys quercinus*), which is smaller than the fat dormouse but has similar habits. Both species are still eaten by humans in many places. *See* RODENTIA.

Charles B. Curtin

Dorylaimida

An order of nematodes in which the labia are generally well developed; however, many taxa exhibit a smoothly rounded anterior. The labial region is often set off from the general body contour by a constriction. The cephalic sensilla are all located on the labial region. When there is no constriction, the labial region is defined as that region anterior to the amphids. The amphidial pouch is shaped like an inverted stirrup, and the aperture is ellipsoidal or a transverse slit. The stoma is armed with a movable mural tooth or a hollow axial spear. The anterior portion of the tooth or spear is produced by a special cell in the anterior esophagus. The esophagus is divided into a slender, muscular anterior region and an elongated or pyriform glandular/muscular posterior

region. There are generally five esophageal glands with orifices posterior to the nerve ring. In some taxa there are three glands, and in others seven have been reported. The esophagointestinal valve is well developed. The mesenteron is often clearly divided into an anterior intestine and a prerectum. Females have one or two reflexed ovaries; when there is only one, the vulva may shift anteriorly. Males have paired equal spicules that are rarely accompanied by a gubernaculum. The males often have the ventromedial preanal supplements preceded by paired adanal supplements.

There are seven dorylaimid superfamilies: Actinolaimoidea, Belondiroidea, Diphtherophoroidea, Dorylaimoidea, Encholaimoidea, Nygolaimoidea, and Trichodoroidea.

Actinolaimoidea contains some of the most distinctive species with remarkable elaborations of the stoma which, however, always bears the characteristic axial spear. Feeding habits among the group vary from algae feeding to predation.

Belondiroidea is a diverse superfamily whose principal common characteristic is a thick sheath of presumed spiral muscle around the basal swollen portion of the esophagus. Nothing is known about the feeding habits of this group of nematodes, frequently encountered in soil near plant roots.

Diphtherophoroidea is presumed ancestral to the plant-parasitic superfamily Trichodoroidea. All known forms are soil inhabitants normally found in association with plant roots.

Dorylaimoidea includes representatives (family Longidoridae) that are important plant parasites and vectors of plant viruses throughout the world. The remaining families are algal feeders in soil and freshwater environments.

Encholaimoidea is a superfamily that is commonly found in soil near plant roots.

Nygolaimoidea is a superfamily of predaceous nematodes. The nygolaimoids have probably had the most varied taxonomic history among the Dorylaimida, having been treated as genera in Dorylaimoidea, a family in Leptonchoidea, and a superfamily in Dorylaimina.

The plant-parasitic Trichodoridea causes a condition called stubby root, in which root growth is arrested with little or no noticeable necrosis. Some members of this superfamily are known vectors of plant viruses. *See* NEMATA.

Armand R. Maggenti

Dosimeter

A device used to measure the dose of ionizing radiation received by an individual. There are many types of dosimeters with varying characteristics and capabilities appropriate for different applications. They differ in sensitivity, energy range, and species of radiation to which they respond. Some can be read out directly by the wearer, while others must be sent to a specially equipped facility to determine the dose.

Air ionization chamber dosimeter. This dosimeter is usually about the size and shape of a fountain pen with a clip suitable for carrying in a pocket, and is sometimes referred to as a pen meter or pocket dosimeter. The outside housing is usually made of plastic. Inside, a conductive cylinder and an insulated central electrode form an air-filled capacitor which is charged by connecting to a charging voltage supply. X-rays or gamma rays absorbed in the air space ionize the air, discharging the capacitor in proportion to the dose received. *See* CAPACITOR; GAMMA-RAY DETECTORS; X-RAYS.

The dose is read by connecting the dosimeter to a quartz fiber or other type of electrometer which reads the charge remaining on the capacitor. Some dosimeters contain an electrometer which can be viewed using a built-in lens and calibrated scale. Some also contain a built-in charging device. *See* ELECTROMETER.

The useful dose range is about 5–300 millirad (50–3000 microgray), depending on the capacitance. Due to insulator leakage, these dosimeters discharge slowly, a few millirads per week or more. *See* IONIZATION CHAMBER.

Film badge dosimeter. The badge housing is of plastic or metal, typically 50 mm square (2 in.) by 12 mm (0.5 in.) thick and equipped with a clip for attaching to clothing. It contains one or more film packets, usually standard dental x-ray film packets. The film is developed in a standardized procedure, and the optical density is measured and correlated to exposure, greater exposure being darker. *See* RADIOGRAPHY.

The side of the badge housing is equipped with one or more filters of lead, aluminum, copper, silver, or cadmium which cover different parts of the film and leave part uncovered. These filters modify the response of different areas of the film and thereby provide more information to correct the response to approximate that of tissue. For example, the uncovered portion responds down to low-energy x-rays and moderate-energy beta particles, while the lead filter area responds only to the more deeply penetrating higher energies. The high thermal neutron cross section of cadmium can be used to enhance the thermal neutron response via the gamma ray emitted on neutron capture. *See* BETA PARTICLES; NEURON; THERMAL NEUTRONS.

The effective dose range for x-rays and gamma rays is from about 30 mrad (300 μGy) to thousands of rads (tens of grays). Film badge dosimeters are fairly stable when stored under cool, dry conditions, with lifetimes suitable for readout at monthly but not longer intervals.

Thermoluminescent dosimeter. The physical packaging can be similar to a film badge with filters incorporated, but the film is replaced by pieces of thermolumiscent materials. Absorbed radiation energy excites electrons of the thermoluminescent material, creating excited metastable states. These states remain excited until the material is subsequently heated to a sufficient temperature, when they

deexcite, emitting visible light. The intensity of the emitted light is proportional to the absorbed radiation energy. The exposed material is placed in an automated instrument which heats it in a controlled manner, and the light is measured with a photomultiplier tube and the resulting signal is converted to a dose. If the exposed material is heated sufficiently to anneal all excited states, the dosimeter can be reused. *See* PHOTOMULTIPLIER.

Thermoluminescence has been observed in many materials, but four are commonly used in dosimeters. These are lithium fluoride (LiF), lithium borate ($Li_2B_4O_7$), calcium fluoride (CaF_2), and calcium sulfate ($CaSO_4$), usually activated with dopants such as manganese (Mn), titanium (Ti), and dysprosium (Dy), which enhance the light output. The light intensity versus temperature response of the material is termed its glow curve. To be useful, the glow curve must have a peak at a temperature sufficiently above ambient and storage temperatures to avoid fading, and also low enough to be conveniently read out in the reader. The details of the glow curve depend on the impurities and the fabrication history of the material.

Lithium fluoride is extensively used for dosimeters, activated with manganese and titanium. It has an average atomic number close to that of soft tissue, so it has tissue equivalent dose response without correction for x-rays and gamma rays.

Because of the high thermal neutron capture cross section of lithium-6 (^{6}Li), lithium fluoride also responds to thermal neutrons. By using a combination of ^{6}LiF and ^{7}LiF, the neutron contribution can be separated by subtracting one reading from the other. This response is strongly skewed toward low energy. To get a reading of neutrons incident upon and reflected from a person, an albedo configuration can be used. In one type, a three-layer sandwich is made with a ^{6}LiF/^{7}LiF pair on each side of a layer of thermal neutron absorber such as cadmium- or boron-loaded plastic. It is worn with one side close against the body. The side away from the body sees only incident thermal neutrons, while the body side sees primarily neutrons (of various energies) reflected and moderated by the body. The albedo response is obtained by subtracting a correction from the incident side. *See* THERMOLUMINESCENCE.

Electronic dosimeter. Usually rectangular in shape, this dosimeter ranges in size from a transistor radio to a pocket pager, though some devices resemble a fountain pen. It contains as sensing element a Geiger-Müller tube along with high-voltage supply, counting and control electronics, and a digital display, all powered by a battery. Battery life depends on features and usage but can be up to a year. *See* GEIGER-MÜLLER COUNTER.

A wide variety of features are available. For example, an audible dose rate indication can be provided by a chirp whenever a preset increment of dose is accumulated. A dose alarm can sound when a preset accumulated dose is reached. A dose-rate alarm can sound when a preset dose rate is exceeded, warning the wearer that he or she has wandered into a radiation area or a radiation source has been turned on. An electrical connection can be made with a computer for dose recordkeeping.

A less common feature is two-way wireless telemetry, permitting the dosimeter to be read and thresholds set remotely. A possible use of this capability is for a battlefield commander to assess the condition of his troops when tactical nuclear weapons have been employed.

Track etch dosimeter. The active element is a piece of CR-39 plastic (from Columbia resin 39), made by polymerization of allyl diglycol carbonate. When a high-energy charged particle, such as a proton recoiling from a fast neutron, traverses the material, it gives up energy and leaves a trail of chemical damage. The material is subsequently exposed to a suitable etchant such as sodium hydroxide (NaOH), and damage sites are preferentially etched, leaving pits which can be observed and counted under a microscope.

This dosimeter is sensitive only to high-energy neutrons, so it is usually used in conjunction with other dosimeters; for example, it is incorporated into a badge-type housing with thermoluminescent x- and gamma-ray and albedo dosimeters. The useful dose range is down to about 30 millirem (300 microsieverts). *See* PARTICLE TRACK ETCHING.

Bubble detector dosimeter. The operation of this dosimeter is based on the change to gas bubbles of superheated liquid (at a temperature above its boiling point) triggered by fast neutron interactions. A commercial bubble detector personal neutron dosimeter consists of a clear plastic tube with an aluminum fitting at one end, about the size of a large pen. It holds about 8 cm^3 (0.5 $in.^3$) of clear, solid elastic polymer containing about 10,000 superheated liquid droplets. On exposure to fast neutrons, small uniform-sized bubbles form. The bubbles are counted visually and a fast neutron dose calculated using a calibration factor determined for that dosimeter.

The dosimeter can be reset by screwing in a knob on the metal fitting which applies about 1.4 megapascals: (200 $lbf/in.^2$) pressure to the polymer, raising the boiling point and reliquefying the bubbles. Unscrewing the knob then relieves the pressure so that the droplets are again superheated, ready for reuse as a dosimeter. The dosimeter is temperature-sensitive and is available with a compensation feature if it needs to be used over varying temperatures. The useful dose range is down to less than 1 mrem (10 μSv), lower than any other fast neutron personal dosimeter. *See* HEALTH PHYSICS.

Peter Ryge

Bibliography. G. F. Knoll, *Radiation Detection and Measurement*, 3d ed., John Wiley, New York, 1999; National Council on Radiation Protection and Measurements, Instrumentation and Monitoring Methods for Radiation Protection, *NCRP Rep.*, no. 57, 1978; A. E. Profio, *Radiation Shielding and Dosimetry*, Wiley-Interscience, New York, 1979; J. Shapiro, *Radiation Protection: A Guide for Scientists and Physicians*, 3d ed., Harvard University Press, Cambridge, 1990.

Double diffusion

A type of convective transport in fluids that depends on the difference in diffusion rates of at least two density-affecting components. This phenomenon was discovered in 1960 in an oceanographic context, where the two components are heat and dissolved salts. (Double diffusion can also occur in other fluid media, such as magmas, stellar and planetary interiors, and liquid metals.) Besides different diffusivities, it is necessary to have an unstable or top-heavy distribution of one component.

In the oceanographic context, if the unstable component is the slower-diffusing (salt), with the overall gravitational stability maintained by the faster-diffusing component (heat), then "salt fingers" will form. Since warm, salty tropical waters generally overlie colder, fresher waters from polar regions, this is a very common stratification in the mid- to low-latitude ocean. Salt fingers arise spontaneously when small parcels of warm, salty water are displaced into the underlying cold, fresh water. Thermal conduction then removes the temperature difference much quicker than salt diffusion can take effect. The resulting cold, salty water parcel continues to sink because of its greater density. Conversely, a parcel of cold, fresh water displaced upward gains heat but not salt, becoming buoyant and continuing to rise (see **illus.**). The fully developed flow has intermingled columns of up- and downgoing fluid, with lateral exchange of heat but not salt, carrying advective vertical fluxes of salt and to a lesser extent heat. In the ocean, the width of individual salt fingers is 2–3 cm, the distance over which thermal diffusion is effective. Their vertical extent can vary from a fraction of a meter to several meters, often occurring within interfaces separating well-mixed layers of 5–50-m vertical extent. When a series of layers and interfaces is found, it is termed a thermohaline staircase, because of the steppy appearance of the temperature and salt profiles. Staircase structures are maintained by the falling out of salt, which is made possible by the greater lateral diffusion of heat. They are quite common in certain oceanic regions, such as the western tropical North Atlantic and the Mediterranean Sea.

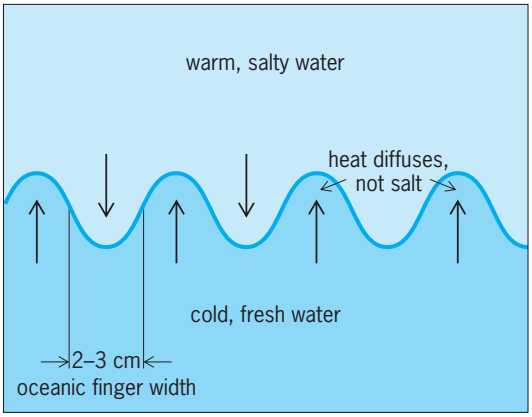

Initial growth of salt fingers arises when parcels of warm, salty water intrude into the cold, fresh water below. Heat escapes from these parcels at a greater rate than salt, because of its much larger diffusivity. The resulting cold, salty water is dense and thus continues to sink. Similarly, cold, fresh water displaced upward gains heat but not salt and becomes buoyant. The resulting network of up- and downgoing columns of fluid exchange heat laterally and carry salt downward. Such salt fingers are a few centimeters wide in the ocean, representing the scale over which thermal conduction can effectively release the gravitational potential energy stored in the unstable salt distribution.

The other form of double-diffusive convection occurs when the faster-diffusing component has an unstable distribution. In the ocean, this happens when cold, fresh water sits above warmer, saltier and denser water. Such stratifications are common in polar regions and in local areas above hot springs at the bottom of the deep sea. In this case, the greater diffusion of heat through a thin, high-gradient interface drives convection in adjacent mixed layers. Relatively little diffusion of salt occurs so that the effect is to transport heat but not salt vertically. Thermohaline staircases are formed in this instance as well, with the net heat flux through the system depending on the sharpness of the interfacial gradients. Because of the importance of diffusion across the thin interface, this type of double diffusion is termed diffusive convection. Both salt fingers and diffusive convection lead to separate vertical transport rates for heat and salt, and a flux of density which is up-gradient. Up-gradient density fluxes are necessary to maintain the thermohaline staircases; the action of turbulence or ordinary diffusion would give a down-gradient density transport and the smoothing away of steppy structure.

Recent measurements indicate that potentially disruptive turbulence is rare in the ocean interior, and the appropriate heat and salt stratifications quite common, making double diffusion a factor in many regions of the world ocean. The importance of double diffusion lies in its ability to affect water mass structure with its differential transport rates for heat and salt. This is believed to play a significant role in producing the "central water" masses of the subtropical gyres with well-defined relationships between temperature and salinity. *See* OCEAN CIRCULATION; OCEANOGRAPHY; SEAWATER. Raymond W. Schmitt

Bibliography. R. W. Schmitt, Double diffusion in oceanography, *Annu. Rev. Fluid Mech.*, 26:255–285, 1994; R. W. Schmitt, The ocean's salt fingers, *Sci. Amer.*, 272(5):50–55, 1995; M. E. Stern, *Ocean Circulation Physics*, Academic Press, New York, 1975; J. S. Turner, *Buoyancy Effects in Fluids*, Cambridge University Press, Cambridge, U.K., 1973.

Douglas-fir

A large coniferous tree, *Pseudotsuga menziesii* (formerly *P. taxifolia*), known also as red fir, belonging to the pine family (Pinaceae). It is one of the most widespread and most valuable tree species of western North America and ranks among the world's

Cone on a branch of Douglas-fir (*Pseudotsuga menziesii*).

most important. In the United States, this species is first in total stand volume, lumber production, and veneer for plywood. It is the most common Christmas tree in the West and is the state tree of Oregon. The name honors David Douglas, the Scottish botanical explorer who introduced the seeds to Europe. *See* PINALES.

Douglas-fir is a large to very large evergreen tree with narrow pointed crown of slightly drooping branches; at maturity it reaches a height of 80–200 ft (24–60 m) and a trunk diameter of 2–5 ft (0.6–1.5 m). In rare instances heights of 300 ft (90 m) are reported. The bark is dark or reddish brown, very thick, deeply furrowed, and often corky. The needlelike leaves spreading mostly in two rows are 0.75–1.25 in. (2–3 cm) long, flat, and flexible, with a very short leafstalk. Buds are distinctive, conical pointed, scaly, and dark red. The cones are elliptical, 2–3.5 in. (5–9 cm) long and light brown, with many thin rounded cone-scales, each above a longer distinctive three-pointed bract (see **illus.**). Seeds are paired and long-winged.

The natural distribution of Douglas-fir extends from southwestern Canada (British Columbia and Alberta) south through the western United States (Washington, Oregon, the Sierra Nevada, and the Rocky Mountains), and south to central Mexico. This species has been introduced into the eastern United States, Europe, and elsewhere.

Two varieties are distinguished: coast Douglas-fir of the Pacific region and Rocky Mountain or inland Douglas-fir. The latter has shorter cones with bracts bent backward. Though not so large, it is hardier and grows better in the East.

The trees form vast forests of pure even-aged stands, also with other conifers, growing rapidly and often attaining an age of 500 years or more. The largest commercial forests are in western Washington and western Oregon.

The wood is reddish, sometimes yellowish brown. Varying in weight and strength, it is moderately lightweight and moderately resistant to decay. Principal uses are building construction as lumber, timbers, piling, and plywood. The lumber is manufactured into millwork, railroad cars, boxes and crates, flooring, furniture, ships and boats, ladders, and so forth. Other products are railroad crossties, cooperage, mine timbers, fencing, and pulpwood. Douglas-fir is planted also for ornament, shade, shelterbelts, and Christmas trees.

Bigcone Douglas-fir (*P. macrocarpa*), the second native species, is local in mountains of southern California. It is a medium-sized tree with larger cones 4–6 in. (10–15 cm) long. Four or five other species are found in China and Japan. *See* FOREST AND FORESTRY; TREE.

Elbert L. Little, Jr.

Down syndrome

The set of physical, mental, and functional abnormalities that result from trisomy 21, the presence in the genome of three, rather than the normal two, chromosomes 21. Down syndrome is one of the most frequent genetic causes of mental retardation. The physical abnormalities that together give rise to the distinctive facial appearance associated with this condition include upslanting eyes with epicanthic folds (folds of skin over the inner corners), flatness of the bridge of the nose and of the middle portion of the face, and a tendency to protrude the tongue, especially when very young. Many other minor, functionally inconsequential abnormalities of the ears, hands, and feet may also be present, and stature is generally reduced.

Approximately 40% of affected individuals are born with congenital heart disease, usually an atrioventricular canal or ventricular septal defect, and obstruction of the intestinal tract (duodenal atresia or stenosis) also occasionally occurs during development. Both the heart and the intestinal anomalies are amenable to surgical correction. Only about a third of embryos and fetuses are actually liveborn. Individuals with Down syndrome are more than normally susceptible to infections, probably because of abnormalities of the lymphocyte system that result from abnormal development of the thymus, and hypothyroidism is relatively common. They also have an increased, although still low, likelihood of developing childhood leukemia. Despite these predispositions, their mean life expectancy, if they survive the first year of life, is now estimated to be between 55 and 60 years. *See* CONGENITAL ANOMALIES.

Nervous system abnormalities. Down syndrome is characterized by three salient clinical abnormalities of the nervous system. The first, recognizable immediately after birth and prominent in the first years of life, is a significant degree of hypotonia that produces a sense of floppiness or looseness. The second is mental retardation, which is generally moderate but may vary from mild to severe. The rate of acquisition of skills and capabilities is often normal during the first few months of life. However, there is then a decrease in rate so that the development of more advanced physical and intellectual skills, such as the ability to construct sentences, to walk stairs alone, and to dress, may be delayed by a year or more. Persons with Down syndrome have learned complex intellectual tasks such as reading, writing,

and arithmetic calculations, but this does not occur in all instances. The acquisition of speech appears to be more affected than other areas of cognitive development. *See* COGNITION.

There is evidence that the intellectual achievements of affected individuals reared in enriched environments, such as the home, are greater than was previously observed in individuals confined to institutions for the mentally retarded. However, even with children with Down syndrome reared at home, there is a progressive and virtually linear decline in developmental quotient (DQ) or intelligence quotient (IQ) starting within the first year. The same appears to be true for social quotients (SQ). Although a wide variety of treatments and therapeutic agents have been tried, including megavitamins and various drugs, there is not at present any accepted pharmacological treatment for Down syndrome. *See* INTELLIGENCE; MENTAL RETARDATION.

The third clinical abnormality of the nervous system, recognizable in 25 to 50% of individuals with Down syndrome living into the fifth decade, is the progressive loss of mental function resulting from the superimposition of a dementing process on the preexisting mental retardation. Associated with the development of these clinical features of dementia has been the appearance, in virtually all persons with Down syndrome older than 35 to 40 years of age, of the pathological and neurochemical changes indistinguishable from those found in the brains of individuals with Alzheimer's disease. Persons with Down syndrome thus constitute a unique population at very high risk of developing Alzheimer's disease, both pathologically and, in many instances, clinically. However, the nature of the association between the two conditions is unknown. Although Alzheimer's disease may be a direct consequence of genetic imbalance produced by trisomy 21, with increased expression of *APP* (the gene for the amyloid precursor protein), it is also possible that the trisomic condition serves to make individuals with Down syndrome more susceptible to those factors that lead to the development of Alzheimer's disease in a chromosomally normal aging population. *See* ALZHEIMER'S DISEASE.

Maternal age. The risk of having a child with Down syndrome increases with maternal age. The distribution of maternal age in the population of women having children is, therefore, the primary determinant of the overall incidence of Down syndrome. The overall incidence of Down syndrome in the newborn population is about 1 per 1000, and the maternal age-specific rate per 1000 live births increases from 0.6 at 20 years of age to 1 at 30 years, 2.6 at 35 years, 9.1 at 40 years, and 41.2 at 46 years. Analyses using deoxyribonucleic acid (DNA) markers have shown that nondisjunction, the failure of paired chromosomes 21 to separate properly during meiosis in the egg, accounts for 92% of all cases of trisomy 21 and is associated with either a decreased or an increased rate of meiotic recombination. The reason for the marked increase in nondisjunction with maternal age is not known, and there is no means for preventing it. *See* MEIOSIS.

Prenatal diagnosis. A reduction in the incidence of Down syndrome has been brought about by the use of prenatal diagnosis and selective termination of affected pregnancies. Prenatal diagnosis is performed by either chorionic villus sampling (obtaining a small piece of placental tissue) at the end of the first trimester or, more commonly, by amniocentesis (withdrawing amniotic fluid) in the second trimester. In either instance, the complement of chromosomes in the fetal cells is analyzed to determine whether there are any alterations in chromosome number or structure. Because these invasive forms of prenatal diagnosis are generally carried out in women 35 years and older, maternal serum screening has been introduced to identify pregnancies at increased risk, irrespective of the mother's age. This form of screening is based on the detection of altered levels of proteins and hormones in the maternal serum, and it must be followed up by amniocentesis to establish a definite diagnosis. Examination of the fetus by ultrasound is also being used as a means of screening pregnancies, but it is not diagnostic. *See* PREGNANCY; PRENATAL DIAGNOSIS.

Trisomy 21 and phenotype. The relationship between the presence of an extra chromosome 21 and the several phenotypic features of Down syndrome is unknown. Virtually complete sequencing of human chromosome 21 has revealed that there are about 240 known and predicted genes on the long arm (21q). Cases in which only part of chromosome 21 is triplicated have been intensively studied in an attempt to correlate particular phenotypic features of Down syndrome with specific regions or loci on the chromosome. Many of these features appear to be associated with imbalance in the so-called Down syndrome region surrounding and distal to the DNA marker *D21S55*. However, there is also evidence for contributions of genes outside this region, especially in the proximal part of 21q, and the region responsible for congenital heart disease seems to be confined to the distal part of 21q. Impaired cognition appears to be the result of the imbalance of genes located in several regions of chromosome 21. *See* CHROMOSOME; CHROMOSOME ABERRATION; GENE; HUMAN GENETICS.

Gene dosage effects. The immediate consequence of having an extra chromosome is a gene dosage effect (that is, an increase in expression of genes located on chromosome 21) for each of the loci present on this chromosome, and such gene dosage effects have been reported for several chromosome 21 loci. With possibly a few exceptions, the measured increases in activities or concentrations of products coded for by chromosome 21 in trisomic cells are close to the theoretically expected value of 1.5. However, changes in the synthesis of proteins coded for by genes elsewhere in the genome have also been observed, presumably as the result of genomic dysregulation resulting from the increased expression of chromosome 21 genes. This dysregulation is limited and does

not affect the vast majority of expressed genes. On the basis of studies in transgenic animal models, decreased platelet serotonin uptake and prostaglandin synthesis and abnormal myoneural junctions, all of which occur in Down syndrome, have been ascribed to increased activity of the enzyme CuZn-superoxide dismutase (SOD1). With this exception, and possibly the relationship between *APP* and Alzheimer's disease, it has not been possible to relate any component of the Down syndrome phenotype to overexpression of specific loci. To permit this problem to be attacked systematically, mouse models of Down syndrome have been made. *See* GENE ACTION.

Mouse models. It has been shown, by comparative gene mapping and genomic studies, that orthologs (similar genes in different species) of about half of the genes on human chromosome 21q are present on the distal part of mouse chromosome 16. The genes distal to *MX1* on human chromosome 21 are represented on mouse chromosomes 10 and 17. On the basis of these findings, several strains of mice segmentally trisomic for portions of the region of mouse chromosome 16 that is orthologous to human 21q have been generated to serve as genetic models of Down syndrome. These animals, which reproduce several of the phenotypic features of Down syndrome, including the craniofacial dysmorphology, male sterility, delayed cognitive maturation, impaired learning, and structural alterations of the cerebellum and certain neurons, are being used to investigate the contributions of specific genes to the phenotype of Down syndrome. Charles J. Epstein

Bibliography. G. T. Capone, Down syndrome: Advances in molecular biology and the neurosciences, *J. Dev. Behav. Pediatr.*, 22:40–59, 2001; W. L. Cohen, L. Nadel, and M. E. Madnick (eds.), *Down Syndrome. Visions for the 21st Century*, 2002; C. J. Epstein, Down syndrome (trisomy 21), in C. R. Scriver et al. (eds.), *The Metabolic Basis of Inherited Disease*, 8th ed., pp. 1123–1256, 2001; A. Milunsky (ed.), *Genetic Disorders and the Fetus: Diagnosis, Prevention, and Treatment*, 1998; N. J. Roizen and D. Patterson, Down's syndrome, *Lancet*, 361:1281–1289, 2003.

Drafting

The making of drawings of objects, structures, or systems that have been visualized by engineers, scientists, or others. Such drawings may be executed in the following ways: manually with drawing instruments and other aids such as templates and appliqués, freehand with pencil on paper, or with automated devices, described below.

Drafting is done by persons with varied backgrounds. Engineers often draft their own designs to determine whether they are workable, structurally sound, and economical. However, much routine drafting is done under the supervision of engineers by technicians specifically trained as drafters. *See* COMPUTER GRAPHICS; DESCRIPTIVE GEOMETRY; ENGINEERING DRAWING. Charles J. Baer

Templates. As the complexity of designs has increased and the use of standard parts has become widespread, drawings have become increasingly stylized. Graphic symbols have replaced pictorial representations. The next step in the evolution of drafting techniques has been the introduction of templates that carry frequently used symbols, from which the draftsman quickly traces the symbols in the required positions on the drawing (**Fig. 1**).

Automated drafting. Where the design procedures from which drawings are developed are repetitive, computers can be programmed to perform the design and to produce their outputs as instructions to automatic drafting equipment. Essentially, automated drafting is a method for creating an engineering drawing or similar document consisting of line delineation either in combination with, or expressed entirely by, alphanumeric characters.

The computer receives as input a comparatively simplified definition of the product design in a form that establishes a mathematical or digital definition of the object to be described graphically. The computer then applies programmed computations, standards, and formatting to direct the graphics-producing device. This method provides for close-tolerance accuracy of delineation and produces at speeds much greater than possible by manual drafting. In addition, the computer can be programmed to check the design information for accuracy, completeness, and continuity during the processing cycle.

Transmittal of the design definition information from the engineer to the computer is in algebraic terms that can be understood and processed by the computer program. This may be accomplished in two ways: (1) by the engineer trained in computer-understandable language, or (2) an intermediary who translates the engineer's notes and sketches into computer language for processing by the computer to drive a graphics output device.

In increasing use is the cathode-ray tube (CRT) console on-line with the computer to provide direct

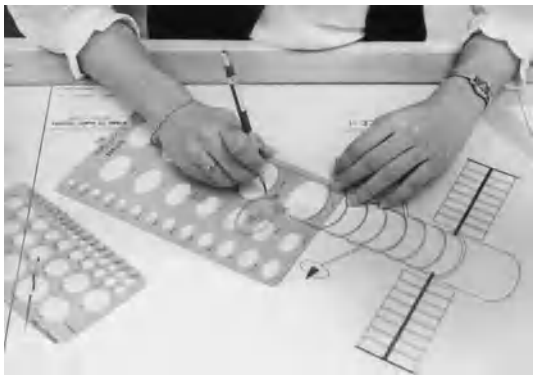

Fig. 1. Use of template in drafting.

interaction between a designer and a computer to develop the definition of a design (**Fig. 2**). The console contains a CRT on which the design is displayed as it is developed by the engineer writing with an electronic light pen to create line delineation and position characters; control keys to initiate the action to be performed on the CRT; function controls to call out program-stored information; and a keyboard, or typewriter, to apply the alphanumeric information required.

The engineering designer uses the light pen in a freehand mode to sketch the drawing and to position lines and program-stored symbols and characters. After a point is located or digitized by the light-sensing pen (which is a digitizer), the designer may type in a code for a symbol on the keyboard or may touch a symbol on the menu with the pen. A menu is a coded chart that shows graphic symbols. The symbols are displayed on the face of the CRT and are moved in sequential position by direction of the light pen. When the required symbols are in place, interconnecting lines are applied by direction with the light pen. During the development of the drawing, graphical images can be repositioned, or erased, instantaneously. The designer can observe the drawing as it progresses and call on the computer to provide computational assistance in developing the design. The information displayed on the CRT is stored in the computer memory, available to drive a graphics output device to produce a conventional engineering drawing when the design is complete (**Fig. 3**).

Graphics-producing devices include XY coordinate plotters, CRT plotters, and photocomposing units, each driven either on-line or off-line by a computer direction. XY coordinate plotters are available in models that produce highly precise drawings rapidly. Alphanumerics and symbols are either drawn by movement of the beam and drawing head, or the equipment may include a character-generator unit to imprint characters or symbols in a single stroke at the position directed by the computer.

CRT plotters produce drawings by exposing an image on sensitized film, usually 35 mm or smaller. By computer direction, the drawing image is displayed on a precision CRT face and recorded on film by a microfilm camera. There is also available a CRT plotter that produces full-size drawings up to 40 × 60 in. (100 × 150 cm) by sweeping the CRT head several times across the surface of a large sheet of sensitized film.

Photocomposing units are used primarily to produce drawings consisting of straight lines, symbols, and alphanumeric characters. Through a method of digitizing that identifies characters, lines, and their position in relation to each other, the drawing information is supplied to the computer where it is formatted to drive the photocomposing unit. The photo unit then projects a light beam successively through masks of the required symbols onto sensitized film.

The computer program required for processing design information to drive the graphics devices is unique for each type of equipment. Through these automated drafting methods, corrections and revisions can be made by providing the computer only

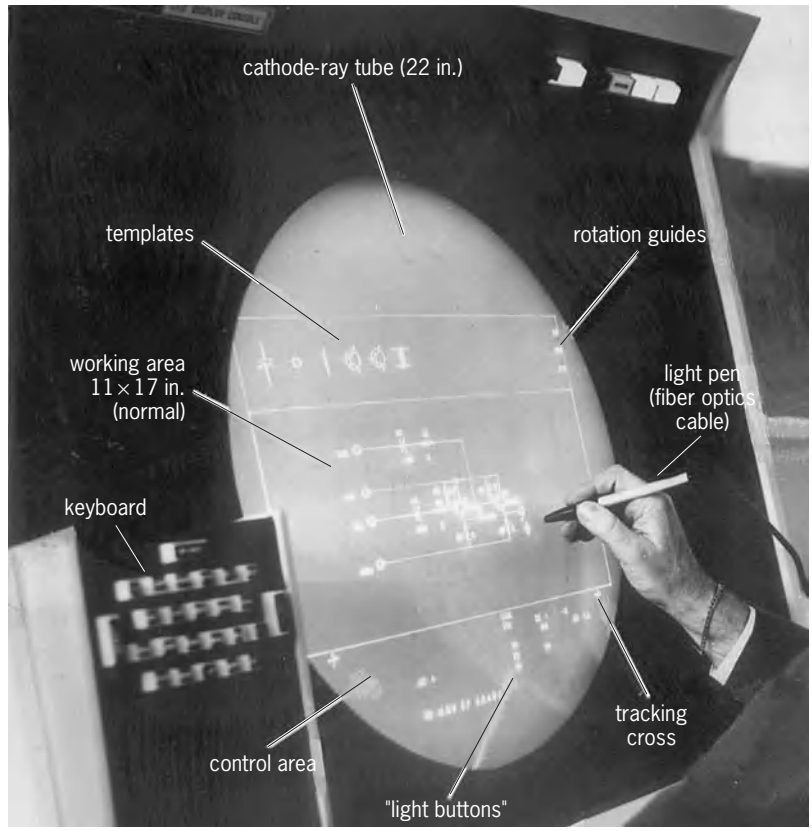

Fig. 2. Electronic light pen is used by designer on a cathode-ray tube to instruct computer in the drafting of a schematic. 1 in. = 2.54 cm. (*Lockheed*)

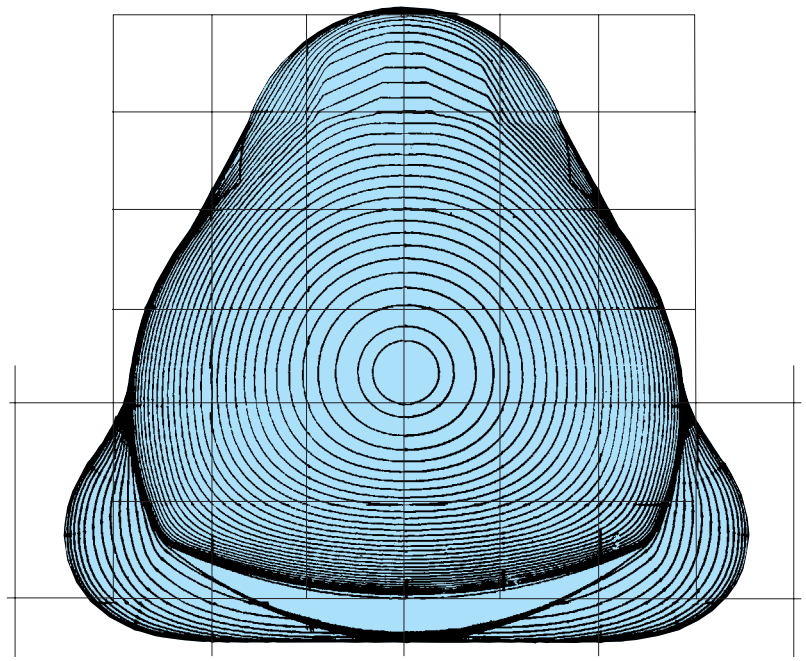

Fig. 3. Loft drawing of body stack for airplane fuselage was produced from design data by equipment such as that shown in Fig. 2. (*Lockheed*)

with the change information to revise the instructions for driving the output device. Tram C. Pritchard

Bibliography. V. Paige, *Computer Aided Drafting and Design*, 1987; D. L. Taylor, *Computer Aided Design*, 1992; M. Van der Hoogt, *The Complete Handbook of Drafting*, 1993.

Drawing of metal

An operation wherein the workpiece is pulled through a die, resulting in a reduction in outside dimensions. This article deals only with bar and wire drawing and tube drawing. For deep drawing and other processes, as performed on sheet metal, *see* SHEET-METAL FORMING

Wire and bar drawing. Among the variables involved in the drawing of wires and bars are properties of the original material, percent reduction of cross-sectional area, die angle and geometry, speed of drawing, and lubrication. The operation usually consists of swaging the end of a round rod to reduce the cross-sectional area so that it can be fed into the die; the material is then pulled through the die at speeds as high as 8000 ft/min. Short lengths are drawn on a draw bench, while long lengths (coils) are drawn on bull blocks. Most wire drawing involves several dies in tandem to reduce the diameter to the desired dimension.

Die angles usually range from 6 to 15°, the actual angle depending on die and workpiece materials. Reductions in cross-sectional area vary 10–45% in one pass, although theoretically the maximum reduction per pass for a perfectly plastic material is 63%.

Die materials are usually alloy steels, carbides, and diamond. Diamond dies are used for drawing fine wires. The purpose of the die land is to maintain dimensional accuracy (see **illus.**).

The force required to pull the workpiece through the die is a function of the strength of the material, die angle, coefficient of friction, and reduction in cross-sectional area. The work applied to the process comprises three components: ideal work of deformation, friction work, and redundant work due to nonuniform deformation within the material. Depending on a number of factors, there is an optimum die angle for which the drawing force is a minimum. In cold-drawing, the strength of the material increases due to work hardening. The terminology for the increasing hardness is shown in the **table**.

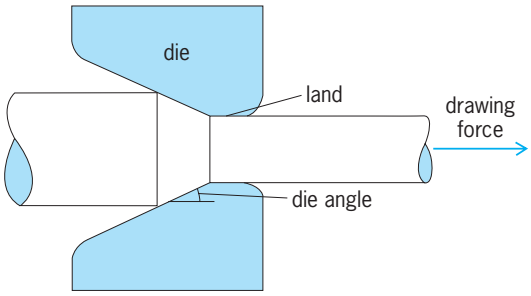

Cross section of drawing die.

Hardness terminology for drawn wire

Terminology	Reduction in area in drawing, %
1/8 hard	11
1/4 hard	21
1/2 hard	37
Hard	60
Extra hard	75
Spring	84
Extra spring	90
Special spring	94

Temperature rise in drawing is important because of its effect on die life, lubrication, and residual stresses. Also, a defect in drawn rods is the rupturing of the core, called cuppy core. The tendency for such internal rupturing increases with increasing die angle, friction, and inclusions in the original material, and with decreasing reduction per pass.

The magnitude of residual stresses in a drawn material depends on the die geometry and reduction. The surface residual stresses are generally compressive for light reductions and tensile for intermediate or heavy reductions.

Extensive study has been made of lubrication in rod and wire drawing. The most common lubricants are various oils containing fatty or chlorinated additives, chemical compounds, soap solutions, and sulfate and oxalate coatings. The original rod to be drawn is usually surface-treated by pickling to remove scale, which can be abrasive and thus considerably reduce die life. For drawing of steel, chemically deposited copper coatings are also used. If the lubricant is applied to the wire surface, it is called dry drawings; if the dies and blocks are completely immersed in the lubricant, the process is called wet drawing. *See* LUBRICANT.

Tube drawing. Tubes are also drawn through dies to reduce the outside diameter and to control the wall thickness. The thickness can be reduced and the inside surface finish can be controlled by using an internal mandrel (plug). Various arrangements and techniques have been developed in drawing tubes of many materials and a variety of cross sections. Dies for tube drawing are made of essentially the same materials as those used in rod drawing.

Serope Kalpakjian

Bibliography. W. F. Hosford, *Metal Forming: Mechanics and Metallurgy*, 2d ed., 1993; E. Lloyd, *Metal Forming: An Introduction to Some Theory, Principles, and Practice*, 1989; J. E. Neely, *Practical Metallurgy and Materials of Industry*, 5th ed., 1999.

Drier (paint)

A material added to paint formulations to facilitate the oxidation of oils. Linseed oil in its pure state normally requires about 3 days to dry to a hard film; linseed oil with driers added dries to a hard film in about 4 h.

The mechanism of drier reaction is not entirely clear. It is believed that the oil oxidation process produces peroxides that subsequently are destroyed, as radicals suitable for polymerization are formed, and that catalytic amounts of drier added to the oil accelerate peroxide destruction and the formation of free radicals. There is evidence that this mechanism may be altered when mixtures of driers are added to the oil, however, for mixtures accelerate the drying process faster than can any one drier alone.

Driers are salts of metals. Acids used to render metals soluble in oil, and therefore useful as driers, may be fatty acids, rosin, or naphthenic or octoic acids derived from petroleum. In commercially available driers, the percentages of salts dissolved in given amounts of mineral spirits are clearly defined so that additions can be accurately measured. Mineral spirit solutions containing mixtures of several metals are available when a special drying effect is desired.

Cobalt is the most reactive of drier metals, and normally is used in quantities amounting to less than 0.1% of the oil. Cobalt generally is regarded as a surface drier, and it is widely used as the only additive in thin-film paint formulations. In heavy-film paint formulations, another metal may be used to dry most of the film with cobalt as an auxiliary metal to dry the film surface and free it of tack.

Lead is less reactive than cobalt and is used in amounts of about 0.5% of the oil. Ecological requirements that paints be free from metallic contamination by lead may restrict the use of lead salts as a viable drier material, however.

Zirconium, in the form of zirconyl soap, has also been used as a commercial drier. Normal use ranges for zirconium as an auxiliary drier are between 0.1% and 0.4% in metal, based on vehicle solids. At these levels zirconium provides a thorough drying capability very close to that of lead. Its advantages are that it does not sulfide-stain, it is in the nontoxic category, and, because of excellent solubility, it does not impair gloss or gloss retention.

Manganese is somewhat less reactive than cobalt. It discolors certain oils, and for this reason is used chiefly in exterior paints subject to the bleaching action of air and sunlight.

Iron is not an effective drier at ambient temperatures, but it is often used in baking finishes. The color of iron salts usually prohibits their use in white paints.

Numerous other metals, including cerium and vanadium, have been used occasionally, and are effective driers. Their use normally is not economically justifiable, however. Zinc and calcium are employed only as auxiliary driers along with other metals.

Certain organic compounds such as 1,10-phenanthroline also catalyze the drying of oils and have been used for this purpose when freedom from all metallic contamination is required.

In general, small amounts of drier are essential to the formation of a satisfactory paint film within a reasonable time. Addition of large amounts of drier, however, can lead to premature embrittlement and failure of the paint film. *See* DRYING OIL; PAINT AND COATINGS.

C. R. Martinson; C. W. Sisler

Bibliography. W. M. Morgans, *Outlines of Paint Technology*, 3d ed., 1990; G. P. Turner, *Introduction to Paint Chemistry and Principles of Paint Technology*, 3d ed., 1980.

Drilling, geotechnical

The drilling of holes for gathering and evaluating earth materials in order to design and monitor construction projects. Geotechnical data are required for the accurate, safe, and efficient design and construction of buildings, bridges, highways, dams, and mining sites. The data are derived from analysis of soil and rock samples obtained by drilling. Soil sampling usually is done with a split-spoon sampler, a tube that opens lengthwise to remove the sample. Rock cores are obtained by diamond drilling, using a hollow, diamond-embedded drill bit to cut an intact rock sample. The samples are tested in laboratories for compressive and shear strength, grain size, weathering properties, moisture content, and consolidation projections. The results are used to determine the materials' supporting characteristics, ability to resist transmission of fluids as with dams, and ability to stand without lateral support as in highway cuts or quarries. In addition, tunneling and mining are accomplished safely only with an understanding of the properties of the soil and rock above and below the projected opening. In-hole or in-situ testing through boreholes drilled in a formation is done to determine the material's permeability, water yield, movement (with inclinometers), and settlement over time (with settlement indicators). *See* ENGINEERING GEOLOGY; ROCK MECHANICS; SOIL MECHANICS.

Elements of drilling. The drilling process requires that soil and rock particles be separated from the formation and removed from the hole. Penetration is achieved by crushing the rock as with a roller bit, by scraping, or by shearing the bottom of the hole with a drag or auger bit. Diamond drilling uses the extreme hardness of diamond to shear particles from the rock. Removal of particles from the hole is accomplished by mechanical means as with an auger, or with fluid (for example, water flush), airflow, or drill mud. In soft materials, a means (such as a steel casing) must be used to prevent the formation from collapsing into the hole.

All drilling requires the application of energy to the formation in the form of pressure, rotation, impact, or fluid flow. The action occurs by the use of tubular drill rods or augers with extensions.

Principal methods. Five commonly used drilling methods are (1) auger drilling of soils and soft rock using a helical screw with hardened bit, (2) rotary drilling with a roller bit and applied weight, (3) core drilling by rotary drilling in a circular groove or ring leaving a central core, (4) percussion drilling by the linear impact of the tool on the rock, and (5) sonic drilling, that is, fluidization of the formation by high-frequency vibration. Most often, a combination of

Fig. 1. Auger-core drill. (*Pennsylvania Drilling Co.*)

auger, rotary, and core drilling techniques is used for geotechnical drilling projects.

Auger drilling and soil sampling. Auger drilling is very common in the geotechnical soil sampling process (**Fig. 1**). Augers can have a solid stem with continuous helical flight (cutting edge) that conveys soils to the surface, or can have a hollow stem in combination with continuous flights enabling the taking of samples through the center of the augers (**Fig. 2**). Augers use hard, tungsten carbide bits to shear particles from the formation, which are conveyed to the surface by the auger flights that run to the top of the drill string. Augers are added usually in 5-ft (1.5-m) lengths, as needed.

Split-spoon samples are taken with the augers removed from the hole or through the hollow stem, continuously or at intervals. The use of a standard sampler and a 140-lb (64-kg) drive hammer is known as the standard penetration test. The sampler is driven into the ground by dropping the hammer 30 in. (760 mm) repeatedly and counting the blows to penetrate 6-in. (150-mm) intervals. This gives the geotechnical engineer an indication of the supporting characteristics of the soils. Samples from the inside of a split spoon are described and logged by the driller to give a continuous description of the formation drilled.

Rotary drilling. In this type of drilling, a rotary motion is induced in the drill rods by the drill rig and transmitted to the drill bit. The cutting action is accomplished by using a roller bit which has rolling cutters with hard projecting teeth or a drag bit which has fixed hardened cutting edges. The roller bit penetrates the rock by the crushing action of its teeth. Weight is placed on the bit by using heavy drill rods or collars or by pushing the rods from the surface with the drill rig while rotating the tools. The drag bit cuts with a scraping action at the bottom of the hole, lifting the rock particles away from the rock.

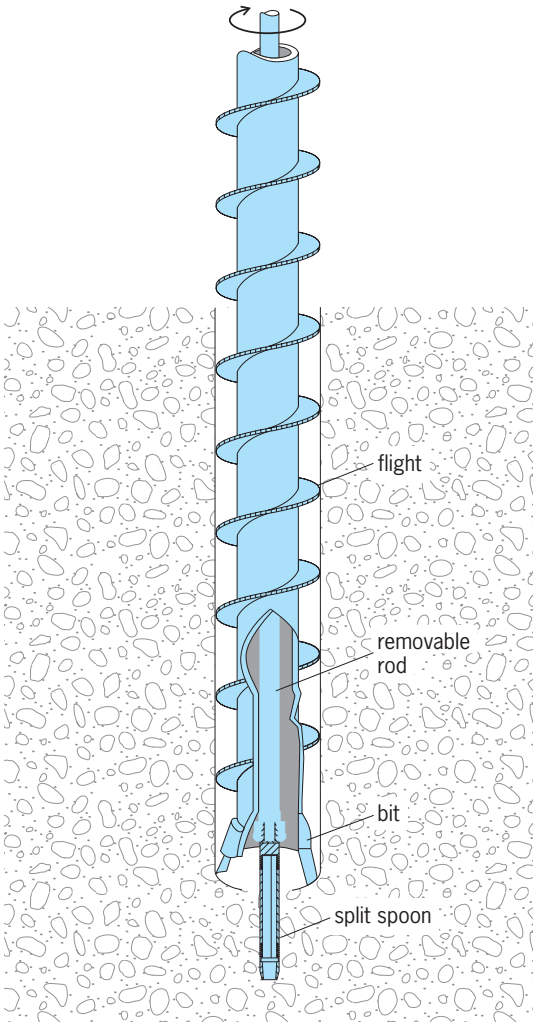

Fig. 2. Hollow-stem auger drill taking a soil sample.

Typical diamond bit core sizes.

Hole diameter, in. (mm)	Core diameter, in. (mm)
1.890 (48.0)	1.062 (27.0)
2.360 (60.0)	1.433 (36.4)
2.980 (75.7)	1.995 (50.6)
3.782 (96.0)	2.500 (63.5)
4.827 (122.6)	3.345 (85.0)

Removal of the cuttings is done by injecting air, fluid, or foam into the hole through a swivel that remains stationary as the drilling rods are rotated. The air or fluid is transmitted down the rods under pressure to the bit, where it cools the cutting surfaces and flushes the rock particles from the hole between the drill rods and the wall of the hole. Commonly, rotary drilling is used for drilling water wells using air as the flushing medium, and gas and oil wells using drilling mud. Drilling mud is a bentonite clay in combination with additives to change viscosity, density, and other properties. The mud has three functions in the deep gas and oil well drilling: it stabilizes the walls of the hole, cools the cutting edges, and flushes the cuttings from the hole.

Diamond drilling. A variation of rotary drilling is diamond drilling, which yields a continuous rock core for testing. This is also the main method for exploration of mineral deposits. In this type of drilling, the bit has diamond particles as the extremely hard cutting medium. Diamond bits are designed to cut an annular space and leave an intact rock core (see **table**). The core bit uses either a surface-set diamond pattern or an impregnated diamond matrix that wears away as it is used to expose new cutting surfaces. The surface set bits require larger diamonds (6–60 stones per carat weight) which are natural diamond stones. The impregnated diamond bits use synthetic diamonds that are very small (0.1 micrometer–1 millimeter). The core bits are used with core barrels that have an inner barrel liner, and a core lifter to enable the rock core to be pulled from the hole. For deeper holes, a retriever on a wire line cable is used to withdraw the inner barrel and core from the hole without pulling all the tools from the hole. *See* PROSPECTING.

Percussion drilling. A variation on rotary drilling is percussion drilling, where penetration is achieved by pneumatically powered impact. This method is often used for drilling blast holes for excavation as in highway cuts and quarries. The drill unit has a power unit similar to a rotating jackhammer that provides rapid impact blows to the drill string. Air is used to flush the rock fragments from the hole. Holes using this method are usually 4 in. (100 mm) in diameter or less.

Another percussion drilling method is downhole hammer drilling, where the impact action is generated at the bottom of the hole and not by the drill rig at the surface. The downhole hammer is powered by high-pressure air which actuates a piston in the hammer to act against the hammer bit. Rapid penetration rates can be achieved without using drill collars or using a heavy drill rig to push on the rods. The bits use tungsten carbide buttons (bumps) to fracture the formation.

Sonic drilling. This method is used mainly for environmental projects. The sonic drill has a special drill head that induces a powerful, high-frequency resonance in the drill string causing the formation to fluidize and allowing the tools to penetrate it. Energy is generated using rapidly rotating weights in the drill head to impart linear high-frequency vibrational energy to the drill string. Practical applications are for drilling of difficult formations such as glacial tills and boulders. Samples are taken using a dual-tube casing system, where the outer tube holds up the hole and the inner tube collects a continuous core.

Other drilling methods. Additional drilling methods are used to meet different geology and hole size and depth requirements.

One of the oldest methods is cable tool drilling, where a chisel-type bit is raised and lowered rapidly with a cable. Water is added, if needed, and the cuttings are extracted by lowering a tubular bailer and pulling out the water and rock or soil particles. This method is still commonly used for drilling water wells.

Auger mining uses the conveyor action of continuous-flight augers for actual mining operations in large-diameter horizontal holes. Drilling proceeds horizontally into a hillside or highwall of a coal mine to recover coal beyond the limits of strip mining.

Auger stem, drilled pier, and caisson drilling uses a short helix run on a solid or telescoping stem. The helix has a cutting bit to gouge out the formation and feed the material to the spiral helix. The bit is withdrawn and the soil and rock fragments are spun off by centrifugal force. This method is commonly used for building concrete piers to support buildings or other structures. It is suitable for holes 12–72 in. (304–1830 mm) or larger in diameter.

Raise drilling is used for constructing large airshafts and access ways for mines. The term "raise" comes from the drilling done from the bottom up. A small, usually less than 15-in. (380-mm) hole is drilled from the surface to the mine entry. A large-diameter bit is attached to drill rods at the bottom, and the bit is pulled to the surface while rotating. Cuttings are allowed to fall to the mine floor, where they are loaded on cars that are brought to the surface for disposal. Because of the tremendous forces that can be applied in the drilling by raise boring equipment, holes over 20 ft (6 m) in diameter can be drilled.

Reverse-circulation drilling is another variation on rotary drilling. Pressurized air or fluid is applied to the annulus between the drill rods and the hole wall and the cuttings come out through the center of the drill rods. Larger holes can be drilled with this method with less air supply. An air-assist line is usually used to help cuttings rapidly get to the surface.

Shaft drilling is shaft sinking by drilling a hole the size of the shaft, as contrasted with conventional shaft sinking where excavation is done by drilling and blasting. The smooth arch of the hole

circumference often makes lining of the shaft unnecessary in hard, massive rock. The drilling is done with larger versions of rotary drilling or reverse circulation drilling.

Directional drilling is a technique used to change the direction of a hole while drilling. This is used for hitting, from one location, multiple producing zones in offshore oil and gas drilling. Directional drilling is widely used in pipeline and utility construction to avoid interference with surface structures or lines. It has proven to be a good method for crossing under streams and rivers, avoiding damage to the channel and conflicts with recreation and commercial activity. Turbodrills are used in which the drill rods do not turn and a fluid-driven turbine motor at the end of the drill string turns a bit. A slightly offset rod behind the motor serves to push the hole in the desired direction. Complex electronics or sensors from the surface monitor the location and direction of the hole. *See* TURBODRILL.

Site evaluation. Test drilling to investigate the supporting characteristics of underlying soil and rock at proposed construction sites is done in a variety of ways. Most common is the American Society for Testing and Materials Standard Penetration Test for soil properties. In this test, a standard sampler is driven into the soil with a specific weight falling a set distance. Many other tests are available both in place and in the soils laboratory to measure moisture content, plasticity, grain size, shear strength, permeability, axial strength, deformation over time, and other properties. Some of these tests are performed on disturbed samples obtained with the Standard Penetration Test, and some are performed on larger samples obtained with test pits or undisturbed sampling with Shelby tubes. In Shelby tube sampling, a thin-walled metal tube with a formed, sharpened edge is pressed into the soil with the hydraulic feed apparatus on the drill rig.

Rock core drilling provides rock samples for visual examination and laboratory testing. The cores are tested in the laboratory for strength, permeability, freeze-thaw cycle deterioration, weathering characteristics, deformation under load over time, and other characteristics.

In-place borehole testing for engineering properties also includes pressure testing, permeability testing, pressure meter testing, pump testing for aquifer information, and cross-hole seismic tests.

Purchasers of contaminated real estate can incur tremendous cleanup costs. As a result, site evaluation for chemical properties of the underlying soil and rock has become the standard for property transfer because of the prevalence of subsurface contamination. Drilling and sampling are done, and monitoring wells of appropriate materials are installed to determine the extent, if any, of contamination generated by prior illegal dumping or substandard manufacturing processes. Thomas B. Sturges III

Bibliography. J. N. Cernica, *Geotechnical Engineering: Soil Mechanics*, Wiley, January 1995; K. Terzaghi et al., *Soil Mechanics in Engineering Practice*, 3d ed., Interscience, 1996.

Drone

An aircraft or helicopter that is flown without an on-board human pilot. These vehicles are now customarily termed "unmanned aerial vehicles" (UAVs). Until the late twentieth century, they were typically called drones or sometimes remotely piloted vehicles (RPVs) when a human operator controlled the vehicle remotely. Although guided missiles and cruise missiles have the same root heritage in the "aerial torpedoes" of the 1910s, missiles are not considered to be UAVs. Crewless versions of nonpowered aerial vehicles such as gliders and balloons can perform some of the same missions as UAVs, but are not considered to be UAV-class vehicles. *See* MISSILE.

Types. There are three basic types of UAVs: the preprogrammed, the smart, and the remotely piloted UAV (**Fig. 1**).

A preprogrammed UAV (Fig. 1*a*) responds to an on-board timer or scheduler and has no sensor contact with the ground. The UAV follows a set routine of maneuvers, altitude changes, speed changes, and course changes that are programmed through an autopilot to the UAV's control surfaces and engine throttle. The UAV can be recovered by a parachute at the end of the mission.

A smart UAV (Fig. 1*b*) carries various sensors and is equipped with an on-board computer. The ability of a smart UAV to make decisions governing course and altitude changes is limited only by its computer and sensor capacity. For example, a smart UAV could probably take off on its own from a given airport, navigate a circuitous route, make decisions en route based on weather or enemy radar action, fly to a second airport, and make a safe landing. *See* GUIDANCE SYSTEMS; NAVIGATION.

The remotely piloted UAV (Fig. 1*c*), probably the most common type of UAV, is under the constant control of an operator or pilot through radio links. The pilot or pilots can be located on the ground, in other aircraft, or on ships. A remotely piloted UAV system includes the aerial vehicle, launch and recovery systems, and ground or airborne stations for controlling the vehicle's flight and processing information from the vehicle. Typical missions for remotely piloted UAVs include reconnaissance or surveillance, jamming enemy communications, serving as decoys, and acquiring or attacking enemy targets. A UAV can acquire and maintain a lock on an enemy target using a laser designator. A decoy provides a false electronic signature to make it appear like a crewed aircraft to enemy radar. Reconnaissance or surveillance UAVs can be equipped with radar, low-light-level television, or infrared sensors that allow over-the-horizon reconnaissance imagery to be transmitted to ground commanders in real time. Rather than performing a single mission, UAVs which perform a multitude of different missions during a flight are becoming more common. *See* INFRARED IMAGING DEVICES.

Evolution. The first military experiments with UAVs began during World War I with the U.S. Navy–Sperry flying bomb and the U.S. Army–Kettering Bug. These were simply biplanes typical of that

era, controlled by a preprogrammed autopilot system. The British also experimented with a variety of drones in the 1920s, including the Larynx and DeHavilland Queen Bee, which were remotely piloted through radio control. World War II saw the debut of the combat version of UAVs, called UCAVs, although they were referred to as drones at the time. In the Pacific theatre, the Navy used TDR-1 Assault Drones to drop 2000 lb (900 kg) of bombs on Japanese ground targets. While the TDR-1s conducted their attack through antiaircraft fire, they were controlled a safe distance away by a circling crewed aircraft through television and telemetry links. Sixty-two percent of the 50 drones successfully damaged their targets while none of the orbiting crewed aircraft were damaged. In the European theatre, Project Aphrodite and Anvil used crewless B-17s under remote control to attack hardened targets in occupied France and Germany. These aircraft were loaded with as much as 24,000 lb (11,000 kg) of explosives. Take-off for these large aircraft was more difficult to control so a two-person crew initially flew the plane from England and then bailed out as a trailing crewed aircraft took over the controls via television and telemetry links. The lessons learned from these World War II experiences were applied by the Navy to send radio-controlled crewless Grumman Hellcats off of aircraft carriers to attack North Korean targets during the Korean War.

The next generation of remotely piloted UAVs emerged during the 1960s. These high-performance, crewless aircraft were flown during the Vietnam War on surveillance and reconnaissance missions. Crewed C-130 aircraft were used to air-launch, control, and then recover the reconnaissance UAVs. More than 3400 missions were flown over Southeast Asia, primarily by modified Ryan BMQ-34 Firebees (**Fig. 2**). These missions produced significant intelligence and targeting information without risking aircrews, but use of the intelligence was delayed until the UAV could be recovered and the film could be processed and interpreted. Extensive use of the modified BQM-34s in Southeast Asia revealed not only the remotely piloted UAV's promise but also its overall operational reliability and low vulnerability in combat. The most frequently used version of the BQM-34 series was the AQM-34L (model 147SC). A total of 1651 SC UAVs was launched combat, and 87.2% returned.

During the 1970s and 1980s, crewless aircraft were also effectively used as countermeasures against air defense systems. The Israelis flew Northrop MQM-74 Chukar-IIs against Egyptian air defenses during the war in 1973. In 1982, the Israelis destroyed most of the Syrian air defense radar warning and tracking system in just 2 h. Small UAVs simulating crewed attack aircraft decoyed Syrian radars into revealing their locations. Still more unpiloted and remotely piloted UAVs swarmed in to overwhelm the Syrian air defense system with false targets. These precursor tactics allowed Israeli crewed aircraft to rapidly destroy Syrian radars, as well as surface-to-air missile sites and interceptor aircraft.

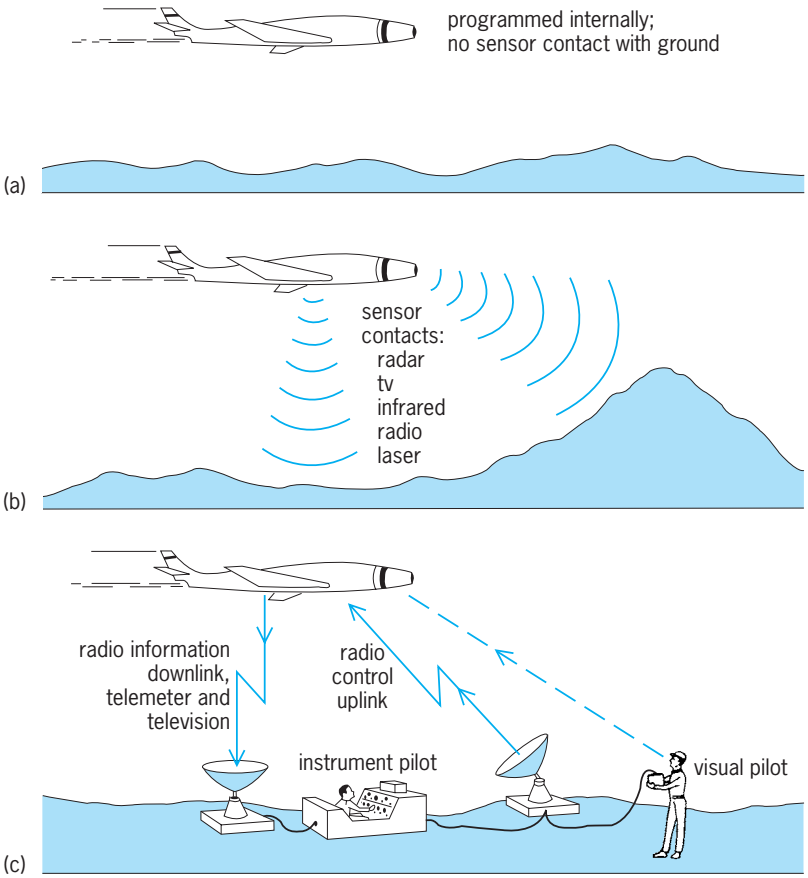

Fig. 1. UAV control schemes. (*a*) Preprogrammed UAV. (*b*) Smart UAV. (*c*) Remotely piloted UAV.

Fig. 2. Modified Ryan reconnaissance vehicles. (*a*) High-altitude version, with wing span of about 27 ft (8 m). It photographed from altitudes of 60,000 ft (18,000 m). (*b*) Low-altitude version with stretched fuselage, large wing tips, and end plates on horizontal stabilizer. It made photo runs at 1500 ft (450 m). (*Teledyne-Ryan*)

Fig. 3. General Atomics Predator UAV. (*General Atomics Aeronautical Systems, Inc.*)

Ship-borne Israeli Aircraft Industries (IAI) Pioneer remotely piloted UAVs were used as gunfire spotters during the Persian Gulf War of 1991. The Pioneers provided television imagery of the downrange effectiveness of naval gunfire to offshore U.S. battleships. The psychological effect of the heavy bombardment was so great that on one occasion, Iraqi soldiers surrendered to an orbiting Pioneer before naval gunfire commenced. UAVs such as the IAI/TRW Hunter and General Atomics Predator (**Fig. 3**) provided real-time reconnaissance imagery during the Balkans War of the mid- to late-1990s in Bosnia and Kosvo. However, there was a time delay between target identification by the UAV ground operator and a request for tactical air support up the chain of command. This led to the escape of several hundred Serbian vehicles before the arrival of attack aircraft. Based on this experience, some Predator UAVs were modified to carry AGM-114 Hellfire antitank missiles and a laser target designator. These "low-flying" modified Predator UAVs together with "high-flying" Northrop Grumman Global Hawk reconnaissance UAVs (**Fig. 4**) flying at altitudes up to 60,000 ft (18,000 m) have been used effectively in the Afghanistan and Iraq wars.

Fig. 4. Northrop Grumman Global Hawk. (*Northrop Grumman*)

Status. The U.S. Air Force and the U.S. Navy have a fly-off competition for a new-generation of UCAVs. The Boeing X-45 and the Northrop Grumman X-47 prototypes both have a stealthy design to evade enemy radar in order to attack high-value targets which may be too dangerous for crewed aircraft to attack. A capability to enable UCAVs to refuel in the air using crewed tanker aircraft is also being considered. Although large UAVs provide excellent real-time reconnaissance imagery, their relative high cost and small numbers make it difficult to provide this information to units smaller than the battalion level. In order to resolve this issue, lower-cost, less capable small UAVs with a size on the order of radio controlled model airplanes are being introduced in the U.S. Marine Corps. Even smaller Micro Air Vehicles (MAVs) are also undergoing current research and development.

About half the countries in the world operate UAVs of one form or another. The United States spends more than half of the world's annual outlays on UAVs and UAV-related research due to the large expense associated with highly capable military UAVs. However, more than half of the UAVs in the world are radio-controlled helicopter-type UAVs for agricultural purposes in Japan. The Federal Aviation Administration has begun to consider how to integrate UAVs into the civil airspace alongside private planes and passenger jets. As the performance and reliability of autopilots continue to improve, passenger-carrying aircraft may, themselves, someday become a form of a UAV by eliminating the on-board human pilot. *See* AUTOPILOT.

Roy Myose; Robert Strohl

Bibliography. P. G. Fahlstrom and T. J. Gleason, *Introduction to UAV Systems*, UAV Systems, Inc., Columbia, MD, 1998; General Accounting Office, *Aquila Remotely Piloted Vehicles*, 1987; P. Jackson (ed.), *Jane's All the Worlds Aircraft*, annually; B. Miller, USAF widens unmanned aircraft effort, *Aviat. Week Space Technol.*, 93(19):46–52, November 9, 1970; L. R. Newcome, *Unmanned Aviation: A Brief History of Unmanned Aerial Vehicles*, American Institute of Aeronautics and Astronautics, Reston, VA, 2004; J. R. Wilson, UAV Worldwide Roundup—2005, *Aerospace America*, September 2005.

Drought

A general term implying a deficiency of precipitation of sufficient magnitude to interfere with some phase of the economy. Agricultural drought, occurring when crops are threatened by lack of rain, is the most common. Hydrologic drought, when reservoirs are depleted, is another common form. The Palmer index has been used by agriculturalists to express the intensity of drought as a function of rainfall and hydrologic variables.

The meteorological causes of drought are usually associated with slow, prevailing, subsiding motions of air masses from continental source regions. These descending air motions, of the order of 600

or 900 ft/day (200 or 300 m/day), result in compressional warming of the air and therefore reduction in the relative humidity. Since the air usually starts out dry, and the relative humidity declines as the air descends, cloud formation is inhibited—or if clouds are formed, they are soon dissipated. In the United States, the area over which such subsidence prevails may involve several states, as in the 1977 drought over much of the Far West, the 1962–1966 Northeast drought, or the dust bowl of the 1930s over the Central Plains.

The atmospheric circulations which lead to this subsidence are the so-called centers of action, like the Bermuda High, which are linked to the planetary waves of the upper-level westerlies. If these centers are displaced from their normal positions or are abnormally developed, they frequently introduce anomalously moist or dry air masses into regions of the temperate latitudes. More important, these long waves interact with the cyclones along the Polar Front in such a way as to form and steer their course into or away from certain areas. In the areas relatively invulnerable to the cyclones, the air descends, and if this process repeats time after time, a deficiency of rainfall and drought may occur. In other areas where moist air is frequently forced to ascend, heavy rains occur. Therefore, drought in one area, say the northeastern United States, is usually associated with abundant precipitation elsewhere, like over the Central Plains. *See* CYCLONE.

After drought has been established in an area, there seems to be a tendency for it to persist and expand into adjacent areas. Although little is known about the physical mechanisms involved in this expansion and persistence, some circumstantial evidence suggests that numerous "feedback" processes are set in motion which aggravate the situation. Among these are large-scale interactions between ocean and atmosphere in which variations in ocean-surface temperature are produced by abnormal wind systems, and these in turn encourage further development of the same type of abnormal circulation. Atmospheric interactions called teleconnections can then establish drought-producing upper-level high-pressure areas over North America. Then again, if an area, such as the Central Plains, is subject to dryness and heat in spring, the parched soil appears to influence subsequent air circulations and rainfall in a drought-extending sense.

A theory supported by numerical modeling studies suggests that over very dry areas, particularly deserts, the loss of heat by radiation relative to the surroundings creates a slow descending motion. This results in compressional warming, lowered relative humidity in the descending air, and inhibition of rain. These processes are probably partly responsible for the tropical droughts observed in various parts of the Sahel (sub-Saharan) region during the 1970s. Of course, interactions with other parts of the atmosphere's prevailing circulation as indicated earlier also play a role in tropical droughts.

Finally, it should be pointed out that some of the most extensive droughts, like those of the 1930s dust bowl era, require compatibly placed centers of action over both the Atlantic and the Pacific. *See* CLIMATE HISTORY.

Another type of drought came to light in Indonesia and the surrounding area, sometimes extending as far westward as India. These droughts may last a couple of years and are associated with the wind systems which create El Niño, an abnormally warm mass of water near the Equator, which can have an east-west extent of hundreds, sometimes thousands, of miles. This water, as much as a few degrees above normal temperature, interacts with the overlying atmosphere to produce rising air motion with more clouds and rainfall in the eastern Pacific (from Peru westward) but sinking air motion with inhibition of rainfall over the western Pacific including Indonesia. While occurring on average every five or so years, El Niño is not periodic and is difficult to predict.

The Sahel (sub-Saharan) droughts are the most long-lasting, sometimes persisting for more than a decade. These droughts impact on agriculture and cattle production to the extent that hundreds of thousands of people starve to death in countries such as Ethiopia. These Sahel droughts are associated with abnormal displacements of the Intertropical Convergence Zone, a west-to-east-oriented area where the trade winds of both hemispheres converge. This is a zone of enhanced rainfall, so that its variable latitudinal position often determines how much rain occurs in parts of the sub-Saharan region. Once desiccation sets in, air-land feedbacks help to perpetuate the drought. *See* SAHEL; TROPICAL METEOROLOGY.

In view of the immense scale and complexity of drought-producing systems, it will be difficult to devise methods of eliminating or ameliorating them.

Jerome Namias

Bibliography. J. G. Charney, Dynamics of deserts and drought in the Sahel, *Quart. J. Roy. Meteorol. Soc.,* 101(428):193–202, 1975; C. K. Folland, T. N. Palmer, and D. E. Parker, Sahel rainfall and worldwide sea temperatures, 1901–85, *Nature*, 320:602–607, 1986; J. Namias, *Factors in the Initiation, Perpetuation and Termination of Drought,* Int. Union Geod. Geophys. Ass. Sci. Hydrol. Publ. 51, 1960; J. Namias, Multiple causes of the North American abnormal winter 1976–77, *Mon. Weath. Rev.,* 106(3):279–295, March 1978; W. C. Palmer, *Meterological Drought,* U. S. Weather Bur. Res. Pap. 45, 1965; N. J. Rosenberg (ed.), *Drought in the Great Plains: Research on Impact and Strategies,* 1980; UNESCO-WMO, *Hydrological Aspects of Drought*, Studies Rep. Hydrol. 39, 1985.

Drug delivery systems

Systems that deliver controlled amounts of a therapeutic agent over a prolonged period or in a desirable temporal fashion, thereby maintaining required plasma or tissue drug levels. Controlled drug delivery systems can take a variety of forms, such as mechanical pumps, polymer matrices or microparticulates, and externally applied transdermal patches. The

development of systems requires the coordinated effort of engineers, physicians, surgeons, chemists, and biologists.

Controlled drug delivery has long been used to treat certain diseases. Individuals with serious blood-borne infections are treated by continuous intravenous infusion of antibiotics. By monitoring the antibiotic concentration in the individual's blood and adjusting the rate of infusion, the physician precisely controls the delivery of the agent. Because this treatment is expensive and requires hospitalization, its use is limited to acute, life-threatening illnesses. In addition, routes of drug administration other than intravenous infusion are frequently required. This drug delivery concept has been applied more widely in the development of a variety of novel strategies.

Pumps. The most obvious approach for controlled drug delivery involves miniaturization of the familiar infusion system. Mechanical pumps, either totally implantable or requiring percutaneous catheters, have been used to deliver insulin, anticoagulants, analgesics, and cancer chemotherapy. Although the technology is mature and is accepted by many individuals requiring treatment, this approach has certain disadvantages. Drugs are maintained in a liquid reservoir prior to delivery, so only agents that are stable in solution at body temperature can be used. The devices are bulky, expensive, and suitable only for highly motivated individuals who can visit their physicians regularly. Because pumps deliver precisely controlled amounts of therapeutic agents, they may find additional important applications, particularly when coupled with implanted biosensors (electronics designed to continuously measure the concentration of a disease-specific analyte in tissue) for feedback control. *See* BIOELECTRONICS.

Polymers. Many disadvantages of controlled drug delivery via pumps can be avoided by using controlled-release polymers as delivery vehicles. In the early 1960s, the first methods for releasing drugs from biocompatible polymers were developed, and many approaches to polymeric controlled release have been tested. As a result, techniques for the production of polymer implants that slowly release drug molecules have become available. Nondegradable polymers, biodegradable polymers, swellable polymers, and biopolymers have been used as the basis of controlled drug delivery systems. For a specific therapeutic agent, the rate, pattern, and duration of drug release can be modified by selecting the appropriate biocompatible polymer and method of device fabrication. *See* POLYMER.

Nondegradable polymers. The first controlled-release polymer systems were based on semipermeable silicone elastomers. Hollow silicone tubes filled with a liquid suspension of a drug are immersed in water, whereupon the drug is released from the system by diffusion through the silicone wall. This technology permits the controlled release of any agent that can dissolve and diffuse through either silicone or poly(ethylene-co-vinyl acetate), the two most commonly used nondegradable, biocompatible polymers. The best example of this technology is the use of cylindrical silicone-based subdermal implants that release contraceptive steroids at a constant rate for 5 years following implantation and provide economical, reliable, long-term protection against unwanted pregnancy. An intrauterine device with a controlled-release poly(ethylene-co-vinyl acetate) coating delivers contraceptive steroids directly to the reproductive tract; local release of certain agents can reduce the overall drug dose and, therefore, reduce or eliminate side effects produced by drug toxicity in other tissues. *See* BIRTH CONTROL.

Transdermal controlled-delivery devices are similar in design to implantable reservoir systems. A single flat polymer membrane is placed between a drug reservoir and the skin. The polymer membrane is designed to control the rate of drug permeation into the skin over a period of several hours to 1 day. Transdermal delivery systems have been produced for agents that can penetrate the skin easily, such as nitroglycerin, scopolamine, clonidine, estradiol, and nicotine.

Solid matrices of nondegradable polymers can also be used for long-term drug release. In comparison to reservoir systems, these devices are simpler (since they are homogeneous and, hence, easier to produce) and potentially safer (since a mechanical defect in a reservoir device can lead to dose dumping). However, it is more difficult to achieve constant rates of drug release with nondegradable matrix systems. Constant release can sometimes be achieved by adding rate-limiting membranes to homogeneous matrices, yielding devices in which a core of polymer-drug matrix serves as the reservoir. In other cases, water-soluble, cross-linked polymers can be used as matrices; release is then activated by the swelling of the polymer matrix in water.

Biodegradable polymers. When nondegradable polymers are used for implanted delivery systems, the polymer must be surgically removed at the end of therapy. This procedure can be eliminated by the use of biodegradable polymers, which dissolve following implantation. Biodegradable polymers, first developed as absorbable sutures in the 1960s, are usually employed as matrix devices. Many chemically distinct biodegradable polymers have been developed for use as drug delivery devices. These polymers must be biocompatible (that is, provoking no undesirable or harmful tissue response) and the degradation products must be nontoxic. To provide reproducible drug delivery, these polymers must also degrade in a controlled manner. Of the polymers tested, poly(lactide-co-glycolide) has been used most frequently. Most importantly, small microspheres of biodegradable polymers can be produced as injectable or ingestible delivery systems; injectable poly(lactide-co-glycolide) microspheres that release a peptide are used for the treatment of prostate cancer in humans. The chemistry of the biodegradable polymers can be exploited in other ways to achieve controlled delivery, for example, by linking drugs to the polymer matrix through degradable covalent bonds.

Proteins and other macromolecules. Recombinant proteins and polypeptides are produced in large

quantities by the biotechnology industry. Many of these molecules, such as the recombinant tissue plasminogen activator, human insulin, and human growth hormone, have important applications in human health, although they are generally less stable than conventional drugs and more difficult to deliver. Protein drugs are difficult to use in humans, because either they are eliminated very quickly from the body or are toxic when delivered systemically at the doses required to achieve a local effect. Controlled-release polymers may overcome these problems by slowly releasing the protein into the blood stream over a long period or by releasing protein into a local tissue site, thus sparing systemic exposure. *See* GENE ACTION; GROWTH FACTOR; HORMONE; PHARMACOLOGY; TOXICOLOGY. W. Mark Saltzman

Liposomes. Since the mid-1970s, liposomes (artificial phospholipid vesicles) have been considered promising delivery systems for biologically active substances because they are biologically inert and biocompatible; very rarely produce toxic or antigenic reactions; protect the drugs enclosed within them from the destructive action of the external media by sterically preventing their contact with inactivating factors in the blood and tissues; can deliver their content inside cells and even inside different cell compartments; and water-soluble drugs can be captured by the inner aqueous space of liposomes, while lipophilic compounds can be incorporated into the liposomal membrane.

Many different methods have been suggested to prepare liposomes of different sizes (from 50 nm to more than 1 μm in diameter) and properties. The most frequently used techniques are ultrasonication (the application of high-frequency sound waves to a lipid suspension in water, resulting in the formation of closed vesicles), reverse phase evaporation (the vacuum evaporation of a lipid-containing water-organic mixture, leaving the liposomal suspension in the water), and detergent removal from mixed lipid-detergent micelles by dialysis or gel filtration (upon removal of the detergent, micelles spontaneously convert into bilayered structures, liposomes).

To increase liposome accumulation in the desired area, the use of targeted liposomes has been suggested. Liposomes with a specific affinity for an affected organ or tissue were believed to increase the efficacy of liposomal pharmaceutical agents. Immunoglobulins (antibodies), primarily of the IgG class, and their fragments (which are attached to the surface of the liposome) are the most promising and widely used targeting moieties for various drugs and drug carriers including liposomes, since they can specifically recognize and bind certain characteristic components of target organs and tissues. Different methods have also been suggested to achieve longer circulation times of liposomes in vivo, including coating the liposome surface with inert, biocompatible polymers, such as polyethylene glycol (PEG), which form a protective layer over the liposome surface and slows down liposome recognition by opsonins (substances in blood serum that accelerate phagocytosis of foreign cells to which antibodies are attached) and subsequent clearance. Liposomes with increased longevity and targetability have also been prepared by adding both the protecting polymer and the targeting moiety to the liposome surface. *See* ANTIBODY; IMMUNOGLOBULIN; OPSONIN.

Most liposomes are internalized by cells via endocytosis and destined to lysosomes for degradation. To achieve liposome-mediated drug delivery inside cells, particularly into a cytoplasmic compartment, pH-sensitive liposomes are frequently used. Lipids used for their preparation are able to undergo certain structural transitions at pH values slightly below normal physiological levels, enabling the destabilization of the liposomal and endosomal membranes and drug release.

The most successful application of liposomes for drug delivery is in cancer therapy, for example, clinical application of doxorubicin (a widely used anticancer drug for both solid tumors and blood or bone marrow tumors) trapped within long-circulating liposomes demonstrated decreased toxicity and increased tumor accumulation. Liposomes are also frequently used as pharmaceutical delivery systems for various protein and peptide drugs; deoxyribonucleic acid (DNA), for example, for gene therapy or enhanced production of therapeutic proteins (liposomes containing positively charged lipids are especially effective in binding DNA and delivering it into cells); and short interfering ribonucleic acid (siRNA) for inhibiting translation of the disease-causing protein. *See* CELL PERMEABILITY; LIPID; LIPOSOMES; PHOSPHOLIPID.

Micelles. Micelles are colloidal particles, often spherical, with the size in a nanometer range, into which many amphiphilic molecules (which have a hydrophilic, or water soluble, head attached to a long hydrophobic, or water insoluble, tail) self-assemble spontaneously. In an aqueous environment, hydrophobic fragments of amphiphilic molecules form the core of a micelle, which is segregated from the environment by hydrophilic parts of the molecules forming the micelle corona. Micelles possess a number of significant advantages as potential drug delivery systems for poorly soluble pharmaceuticals, since the hydrophobic core of micelles may be used as a cargo space for encapsulation of a variety of sparingly soluble therapeutic and diagnostic agents. Such encapsulation substantially increases the bioavailability of pharmaceuticals, protects them from destructive factors upon parenteral administration (via injection), and beneficially modifies their pharmacokinetics (the way that drugs move through the body after they are administered) and biodistribution.

Polymeric micelles prepared from certain amphiphilic co-polymers consisting of both hydrophilic and hydrophobic monomer units represent a separate class of pharmaceutical micelles. The majority of micellar preparations, especially those with anticancer drugs, are still at the preclinical development stage. *See* COLLOID; MICELLE; POLYMER.

Novel systems. There are a number of novel drug delivery systems at different stages of preclinical

development and early clinical trials. These systems include microemulsions, lipoprotein-based drug carriers, lipid nanocapsules, lipid-coated microbubbles, dendrimers (polymers having a well-defined branched structure), cells and cell ghosts (such as red blood cell ghosts, that is, cells in which part of the inner content is replaced with a drug), aerosols, magnetic nanoparticles that can be concentrated in the required region in the body under the action of an external magnetic field, and niosomes (liposome-like structures prepared of nonionic surfactants). Attempts have also been made to make drug delivery systems that are stimuli-sensitive, that is, capable of releasing DNA or an entrapped drug in response to certain physiological changes, such as decreased pH values and/or increased temperature in pathological areas. The application of nanosized drug delivery systems represents an important area of nanomedicine. *See* NANOTECHNOLOGY. Vladimir P. Torchilin

Bibliography. R. Langer, New methods of drug delivery, *Science*, 249:1527–1533, 1990; D. A. LaVan, T. McGuire, and R. Langer, Small-scale systems for in vivo drug delivery, *Nat. Biotechnol.*, 21(10):1184–1191, 2003; K. Park and R. J. Mrsny (eds.), *Controlled Drug Delivery*, American Chemical Society, Washington, DC, 2000; W. M. Saltzman and W. L. Olbricht, Building drug delivery into tissue engineering, *Nat. Rev. Drug Discov.*, 1(3):177–186, 2002; S. Svenson (ed.), *Carrier-Based Drug Delivery*, American Chemical Society, Washington DC, 2004; V. P. Torchilin, Recent advances with liposomes as pharmaceutical carriers, *Nat. Rev. Drug Discov.*, 4(2):145–160, 2005; V. P. Torchilin, Structure and design of polymeric surfactant-based drug delivery systems, *J. Controlled Release*, 73(2–3):137–172, 2001; T. L. Wyatt and W. M. Saltzman, Protein delivery from nondegradable polymer matrices, *Pharm. Biotechnol.*, 10:119–137, 1997.

Drug resistance

The ability of an organism to resist the action of an inhibitory molecule or compound. Examples of drug resistance include disease-causing bacteria evading the activity of antibiotics, the human immunodeficiency virus resisting antiviral agents, and human cancer cells replicating despite the presence of chemotherapy agents. There are many ways in which cells or organisms become resistant to drugs, and some organisms have developed many resistance mechanisms, each specific to a different drug. Drug resistance is best understood as it applies to bacteria, and the increasing resistance of many common disease-causing bacteria to antibiotics is a global crisis.

Genetic basis. Some organisms or cells are innately or inherently resistant to the action of specific drugs. In other cases, the development of drug resistance involves a change in the genetic makeup of the organism. This change can be either a mutation in a chromosomal gene or the acquisition of new genetic material from another cell or the environment.

Mutation. Any mutation that results in resistance to an antibiotic will provide a selective advantage to the bacterial strain in the presence of the antibiotic. Susceptible organisms are inhibited, while resistant ones replicate to become the predominant population. Single mutations that result in drug resistance are usually specific to a single drug or class of drugs. In order to prevent the emergence of resistance of a bacterial pathogen to a single antibiotic, physicians may prescribe a combination of antibiotics that have different mechanisms of action. In this way, a single genetic mutation conferring resistance to one class of drugs does not result in a selective advantage, since the other antibiotic may still be effective against that bacterial strain. The likelihood of a single organism developing resistance to two unrelated drugs by mutational events is much lower than to only one drug.

Gene acquisition. Organisms may acquire deoxyribonucleic acid (DNA) that codes for drug resistance by a number of mechanisms. Transformation involves the uptake of DNA from the environment. Bacteria that are able to undergo transformation are said to be competent; common species of bacteria that share this characteristic include *Haemophilus*, *Streptococcus*, and *Neisseria* spp. Once DNA is taken up into the bacterial cell, it can recombine with the recipient organism's chromosomal DNA. This process plays a role in the development and spread of antibiotic resistance, which can occur both within and between species.

Transduction, another mechanism by which new DNA is acquired by bacteria, is mediated by viruses that infect bacteria (bacteriophages). Bacteriophages can integrate their DNA into the bacterial chromosome. During excision of viral DNA for packaging into new virions, the virus may package not only its own DNA but parts of the bacterial DNA as well. This genetic material is then carried to the next bacterium infected by the bacteriophage.

Conjugation is the most common mechanism of acquisition and spread of resistance genes among bacteria. This process, which requires cell-to-cell contact, involves direct transfer of DNA from the donor cell to a recipient cell. While conjugation can involve cell-to-cell transfer of chromosomal genes, bacterial resistance genes are more commonly transferred on nonchromosomal genetic elements known as plasmids or transposons. Plasmids are extrachromosomal, self-replicating genetic elements which exist independently within the bacterial cell. In contrast, transposons do not exist independently within the cell except when transferred; during transference they are incorporated into the chromosomal DNA of the cell or into the DNA of a plasmid. Transposons are sometimes referred to as jumping genes. *See* DEOXYRIBONUCLEIC ACID (DNA).

Plasmids and transposons play an extremely important role in allowing bacteria to adapt to environmental change by accommodating the efficient exchange of genetic information both within and between bacterial species. Plasmids and transposons that carry genes coding for resistance to antibiotics are sometimes called R-factors or R-determinants. A

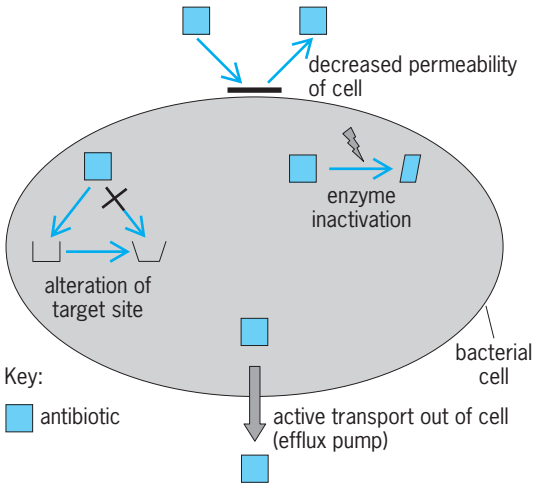

Fig. 1. Four common mechanisms of antibiotic resistance.

single plasmid or transposon often carries multiple resistance genes, and therefore provides a powerful means by which antibiotic resistance genes can disseminate. *See* PLASMID.

Mechanisms of resistance. The four most important antibiotic resistance mechanisms are alteration of the target site of the antibiotic, enzyme inactivation of the antibiotic, active transport of the antibiotic out of the bacterial cell, and decreased permeability of the bacterial cell wall to the antibiotic (**Fig. 1**).

Alteration of target site. By altering the target site to which an antibiotic must bind, an organism may decrease or eliminate the activity of the antibiotic. Alteration of the target site is the mechanism for one of the most problematic antibiotic resistances worldwide, methicillin resistance among *Staphylococcus aureus*. The most common human bacterial pathogen, *S. aureus* has acquired resistance to many different antibiotics. Against many hospital-acquired strains of methicillin-resistant *S. aureus*, only one currently approved antibiotic, vancomycin, remains active. If resistance to vancomycin becomes widespread among *S. aureus*, there exists the possibility of a postantibiotic era with respect to this pathogen, with no antibiotics available to effectively treat the bacteria.

Methicillin resistance among *Staphylococcus* species is mediated by a chromosomal gene (*mecA*) which codes for one of the penicillin-binding proteins (PBP2a) found in the cytoplasmic membrane of these bacteria. An alteration in this penicillin-binding protein decreases the affinity of the protein for all penicillin-related antibiotics. Since methicillin and related antibiotics cannot bind to this protein, they cannot interfere with cell wall synthesis. Methicillin-susceptible organisms do not carry the *mecA* gene.

Another problematic form of antibiotic resistance, penicillin resistance among *Streptococcus pneumoniae*, is also mediated by an alteration in the penicillin-binding proteins. There was a dramatic increase in resistance to penicillin among *S. pneumoniae* in the United States during the 1980s-1990s (**Fig. 2**). Furthermore, because all antibiotics that are based upon the penicillin ring structure (called the β-lactam ring) act by binding to penicillin-binding proteins, resistance to penicillin among *S. pneumoniae* is accompanied by a decrease in susceptibility to many other agents (such as other penicillins and cephalosporins). Since *S. pneumoniae* is the most common bacterial cause of pneumonia and meningitis in adults, this increasing resistance to penicillins and cephalosporins may have a tremendous impact on therapy for these serious infections.

Erythromycin resistance among some species of bacteria has also been found to be due to the alteration of a target site. In this case, a plasmid or transposon carries a gene encoding a ribosomal methylase. This enzyme causes a change in the ribosomal subunit that is targeted by erythromycin, thus causing decreased affinity and resistance to the drug. *See* BACTERIAL GENETICS.

Enzyme inactivation. The most common mechanism by which bacteria are resistant to antibiotics is by producing enzymes that inactivate the drugs. β-Lactam antibiotics (penicillins and cephalosporins) can be inactivated by enzymes known as β-lactamases. Hundreds of β-lactamases have been described; they can be both plasmid or chromosomally encoded, and have varying degrees of activity against the different β-lactam antibiotics. Many bacteria produce multiple β-lactamases. In response to the proliferation and spread of β-lactamases, the pharmaceutical industry has developed some β-lactam antibiotics that are more resistant to hydrolysis by these enzymes. In addition, some combination drugs have been produced which contain both a β-lactam antibiotic and a β-lactamase inhibitor; the inhibitor has high affinity for the β-lactamase enzyme, irreversibly binds to it, and thereby preserves the activity of the β-lactam antibiotic.

Bacterial resistance to aminoglycoside antibiotics is mediated by aminoglycoside-modifying enzymes. Acetyltransferases, phosphotransferases, and adenyltransferases have been described as modifying the structure of these drugs, rendering them less able to reach their target site. This is a common resistance mechanism employed by a broad range of bacterial species.

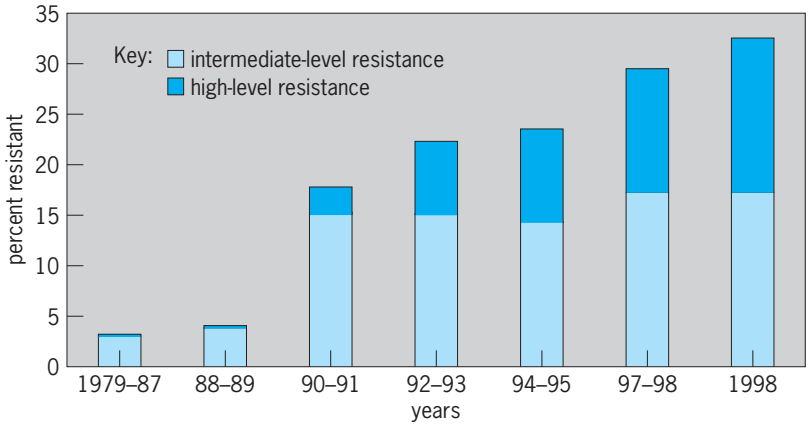

Fig. 2. Increasing penicillin-resistance rates among *Streptococcus pneumoniae* from United States medical centers since 1979.

Active transport of the drug. Active transport systems (efflux pumps) have been described for the removal of some antibiotics (such as tetracyclines, macrolides, and quinolones) from bacterial cells. In these situations, even though the drug can enter the bacterial cell, active efflux of the agent prevents it from accumulating and interfering with bacterial metabolism or replication.

Decreased permeability of the cell. Bacteria are intrinsically resistant to many drugs based solely on the fact that the drugs cannot penetrate the bacterial cell wall or cell membrane. In addition, bacteria can acquire resistance to a drug by an alteration in the porin proteins that form channels in the cell membrane. The resistance that *Pseudomonas aeruginosa* exhibits to a variety of penicillins and cephalosporins is mediated by an alteration in porin proteins.

Promoters. In the hospital environment, many factors combine to promote the development of drug resistance among bacteria. Increasing use of powerful new antibiotics gives selective advantage to the most resistant bacteria. In addition, advances in medical technology allow for the survival of sicker patients who undergo frequent invasive procedures. These factors allow bacteria to survive in sites to which antibiotic drugs do not penetrate well. When bacteria replicate in the presence of low levels of drug (levels too low to kill the bacteria or prevent replication), resistant mutants are more likely to emerge. Finally, poor infection control practices in hospitals allow for the unchecked spread of already resistant strains of bacteria. Some studies have demonstrated that only about half of physicians appropriately wash their hands before examining patients.

Outside the hospital environment, other important factors promote antibiotic resistance. The overuse of antibiotics in outpatient medicine and the use of antibiotics in agriculture exert selective pressure for the emergence of resistant bacterial strains. The spread of these resistant strains is facilitated by increasing numbers of children in close contact at day care centers, and by more national and international travel. *Streptococcus pneumoniae* develops resistance almost exclusively outside the hospital environment. Spread of penicillin-resistant strains of *S. pneumoniae* have been documented from one continent to another, and the highest prevalence of penicillin resistance in the United States was initially found in day care centers.

Impact. The most important impact of drug resistance is an increase in disease and death from infections due to drug-resistant strains of viruses or bacteria and from cancers due to chemotherapy-resistant neoplasms. The human immunodeficiency virus (HIV), which is estimated to have infected over 40 million people worldwide, has the ability to develop resistance to every antiviral agent yet approved for treatment. When HIV does become resistant to antiviral agents, treatment with these drugs predictably fails. *Staphylococcus aureus*, the most common human bacterial pathogen, is increasingly resistant to available antibiotics, and recently strains have been reported which are resistant to vancomycin, the last line of defense.

Increasing rates of drug resistance result in increased use of more expensive, and often more toxic, drugs. Prevention of transmission of resistant bacteria in hospitals requires more frequent isolation of patients, which is also expensive. Overall, the cost of antibiotic resistance has been estimated to exceed $30 billion annually in the United States.

Control. A multifaceted worldwide effort will be required to control drug resistance among disease-causing microorganisms. The first step is to perform good surveillance studies to define the extent of the problem. Ongoing programs to decrease the use of antibiotics, both in the clinics and in agriculture, will be necessary. The increased use of vaccines to prevent infection can help limit the need for antibiotics. Finally, the development of novel classes of antibiotics to fight emerging resistant bacteria will be required. Unfortunately, there is no guarantee that new antibiotics can be developed quickly enough to remain one step ahead of the most adaptable bacterial foes. *See* ANTIBIOTIC; BACTERIA. Daniel J. Diekema

Bibliography. H. S. Gold and R. C. Moellering, Antimicrobial-drug resistance, *N. Engl. J. Med.*, 335:1445–1453, 1996; R. N. Jones and M. A. Pfaller, Bacterial resistance: A worldwide problem, *Diag. Microbiol. Infect. Dis.*, 31:379–388, 1998; R. Quintiliani and P. Courvalin, Mechanisms of resistance to antimicrobial agents, in P. R. Murray et al. (eds.), *Manual of Clinical Microbiology*, 6th ed., American Society for Microbiology, Washington, DC, 1995.

Dry-docking

The removal of a ship from its water environment, which needs to be done periodically to repaint the underwater portion of the hull, repair components exposed to the water, and conduct repairs and modernizations requiring cuts in the hull. At first, ships were just beached at low tide. This practice gave way to use of a marine railway where ships were hauled up the beach on rails.

The construction of the relatively low-capacity marine railways shifted to construction of high-capacity dry docks dug out of the ground (called graving docks). These "stone docks," which were constructed of granite stone, were large enough to accommodate the dry-docking of contemporary ships displacing 2000 long tons (2032 metric tons) and more in Europe and North America in the 1700s and 1800s. In the mid-1800s, the construction of these graving docks was followed by the construction of floating dry docks, which were also of relatively high capacity. These floating dry docks were portable and allowed ship repair to expand to the far reaches of the globe by the early 1900s.

As warship size increased and the shift from wood to steel occurred in the late 1800s, an ever-expanding complement of graving docks, floating docks, and marine railways emerged, especially during World War II, and this expansion continued until late in

the twentieth century. New methods of dry-docking ships emerged in the late twentieth century. These included heavylift ships, vertical lifts, and computer-controlled floating dry docks.

Reasons for dry-docking. Ships require dry-docking for many reasons. If antifouling paint has worn off from the ship's hull, the hull will foul with grass and other marine life such as barnacles, especially in warm salt water environments. The excessive accumulation of marine growth will eventually slow the ship down, causing it to be much less fuel efficient. Divers can clean the hull; however, the frequency of diver cleaning will continue to increase unless the antifouling coating is refurbished, which requires dry-docking. Many underwater repairs can also be accomplished by divers such as propeller replacement or placing on cofferdams for localized hull welding, painting, or sea valve replacement. When large-scale welding or overhhaul of numerous sea valves is required, again dry-docking is the most economical option. About every 8 to 12 years, a ship's propulsion shafting must be pulled and refurbished, requiring the ship to be in dry dock. Major repairs or modernizations below the water line of the hull that require large sections of the hull to be removed also require dry-docking. These include major machinery repairs or replacements where access cuts through the underwater hull are required.

Some ships are modified to increase their capacity. This is usually done by cutting the ship in half at midships, pulling the two halves apart forward and aft, and then inserting what is called a parallel midbody to lengthen the ship. Of course, this must be done in a dry dock.

Some large ships such as very large crude carriers (VLCC), enormous cruise ships, and aircraft carriers are built in relatively shallow graving docks called building docks. Ship erection units that had been constructed elsewhere in the shipyard are placed in the building dock and welded to the other units.

Some ships are constructed on what are called land-level transfer facilities. Large areas of a new construction shipyard are dedicated to such a facility. Ship units are erected on a carriage assembly and welded together. After the ship is assembled, outfitted, and painted, it is moved via the carriage on a rail track system onto an adjacent floating dry dock or vertical lift for launching.

The floating dry dock adjacent to a land-level transfer facility is typically landed on a concrete grid system so that a ship can be moved onto the dock via the carriage on rails. A rail bridge is placed across the short boundary between the dock and the land portion of the rail system. Once the ship is on the floating dry dock in the correct location, the floating dry dock is deballasted so that it floats off the concrete grid system. The floating dry dock has its own mooring system using a system of anchors, chains, and motor-driven winches to move itself over a dredged hole in the harbor (or river). The floating dry dock is then ballasted down and the ship floats off the blocks. The ship is then pulled out of the dry dock.

Frequency of dry-docking. The technology of underwater paint coating systems and increased ship husbandary (underwater ship cleaning) capabilities, maintained through diving operations, have allowed the intervals between dry-dockings to increase. The U.S. Navy has lengthened the period between dry-dockings from 3 years (around 1980) to as much as 12 years, depending on vessel class.

Care must be taken to properly examine shaft and strut bearings, shaft coating systems, and hull cathotic protection systems while a ship is in dry dock to assure that deficiencies are not later identified between these extended dry-docking periods that would necessitate unscheduled and usually very costly dry-dockings.

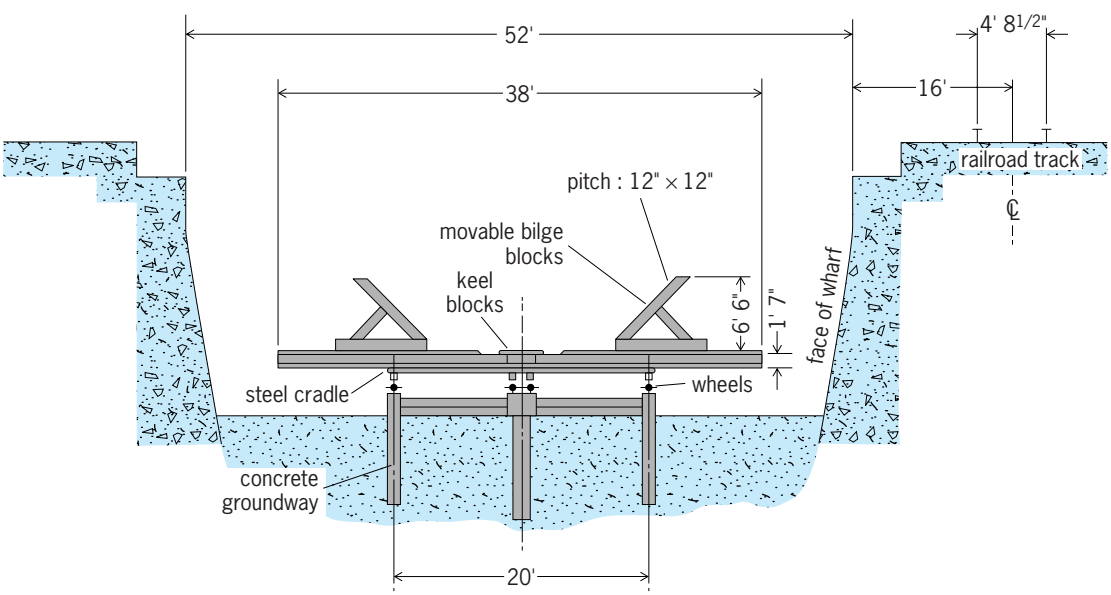

Fig. 1. End-view drawing (section looking inshore) of a marine railway circa 1955 at the Washington Navy Yard in Washington DC. This was the site of the first marine railway in United States. (*U.S. Navy*)

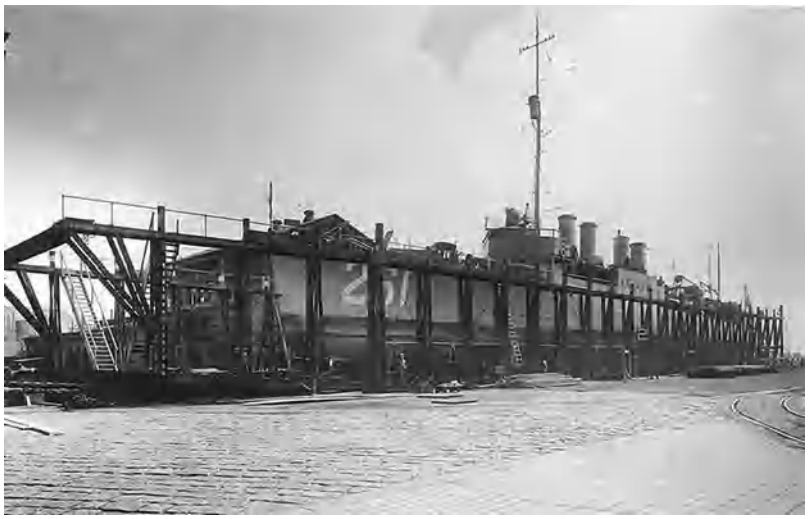

Fig. 2. USS *Delphy* (DD 261), a U.S. Navy World War I destroyer, on a marine railway after being hauled out of the water up the railway at the Boston Navy Yard (since closed) in 1919. (*U.S. Navy*)

Types of dry docks. The four types of dry docks are known as marine railways, floating dry docks, graving docks, and vertical lift docks. The size of the ship usually determines which type is used.

Marine railway. This consists of a cradle of wood or steel with rollers on which the ship may be hauled out of the water along a fixed inclined track leading up the bank of a body of water (**Figs. 1** and **2**). The advantages of a marine railway lie in the economy of the original construction and the relative low cost of maintenance. A marine railway is ideal for ships up to 5000 long tons (5080 metric tons).

Floating dry dock. This type of dry dock may be constructed of timber, steel, or concrete. Few timber floating dry docks remain. Some larger floating dry docks are built in sections called sectional floating docks which are secured together by steel plates (**Fig. 3**). These docks have the capability of self-docking their own sections for maintenance purposes. A floating dry dock is submerged enough to provide the required depth of water over the keel blocks, by filling its ballast tanks with water. The ship to be dry-docked is then positioned in the dock, the ballast tanks of the dock are pumped out by powerful electric pumps located within the dock's wingwalls, the ship lands on the keel blocks, the side blocks are hauled, and the ship is lifted out of the water.

The floating dry dock's ballast tanks are pumped out according to a carefully engineered plan. This plan is designed so that the dry-dock tank pumping sequence does not put undo stress on the dock structure and keeps the pontoon deck as straight as possible while the ship lands on the blocks and is raised out of the water.

The ship is positioned over the blocking system using lines or cables run from the dry-dock wing walls to the ship. Some of these lines accurately position the ship forward and aft, while others position the ship side to side. These lines also help move the ship in and out of the dry dock. Some ships have such deep projections below the keel (like propellers or transducers) that the ship's centerline must be kept to one side of the keel blocks while entering or leaving the dry dock.

Once the ship has fully landed on the keel blocks, the side blocks are then hauled in so that they are in contact with the hull. These side blocks are hauled in using chains and pulleys operated from the dock's wing walls. It is critical that this is done while the ship is coming out of the water prior to the ship reaching the draft of unstability. If not, the ship could roll off the keel blocks and overturn in the dry dock.

Fig. 3. USS *Los Alamos* (AFDB 7), a U.S. Navy floating sectional dry dock, at Holy Loch, Scotland (where it was located for many years), with a submarine docked in it. (*U.S. Navy*)

Fig. 4. British Royal Navy ship HMS *Halifax* in a graving dock in 1889 at Halifax, Nova Scotia, Canada. (*Courtesy of Irving Shipbuilding Inc.*)

An advantage of a floating dry dock is that it can be moved from shipyard to shipyard or towed across the sea to remote locations where dry docks are needed. The first forward-positioned floating dry dock of the U.S. Navy was the *Dewey*, which was towed from Chesapeake Bay on the United States east coast to Subic Bay, Philippines, in 1906 over a period of $6^1/_2$ months, by way of the Atlantic, the Mediterranean, the Suez Canal, the Indian Ocean, and the South China Sea.

A floating dry dock does need a place in the harbor to be submerged. This "hole" is usually an area dredged so that the floating dry dock can fully ballast down and the ship to be docked can enter the dry dock and be positioned over the keel blocks.

Graving dock. This is a dry dock dug out of the ground, usually with a thick concrete base supported by pilings and surrounded on three sides by earth held back by concrete walls anchored back into the surrounding soil. Sometimes graving dock walls are constructed of steel sheet piling structures. The seaward entrance usually has a gate which can be raised or lowered or a floatable vessel called a caisson which is ballasted down at the entrance into a sill to seal off the dock from the adjacent waterway. Large electrically driven pumps then pump out the graving dock. The dock is flooded via flooding tunnels by opening large gate-type valves.

When the dock is flooded and the gate is lowered or the caisson is removed, a ship is brought into the graving dock using boats and docking lines. The ship is precisely positioned over the blocking systems which is on the floor of the dock, and the gate is then raised or the caisson is ballasted down into the sill. The graving dock's dewatering pumps are then started and slowly the water in the dock drops. The ship lands on the blocking system, and the water continues to be pumped out until the dock floor is dry.

In **Fig. 4**, the HMS *Halifax* is seen on the blocks in a partially flooded graving dock in 1889, supported

Fig. 5. USS *Missouri* (BB 63) in Dry Dock #1, a graving dock at Long Beach Naval Shipyard in California (since closed), in 1984. (*U.S. Navy*)

Fig. 6. Graving dock being flooded at Pearl Harbor Naval Shipyard. (*U.S. Navy*).

on either side by what are called wale shores. These shores, which look like telephone poles on either side of the ship, are wedged into place to keep the ship from overturning in the dock. They are used instead of or in addition to side blocks.

Figure 5 shows the USS *Missouri* (BB 63) in a graving dock in 1984. The dry dock was built to handle aircraft carriers, but a battleship does a nice job of filling it. One advantage of a graving dock, seen here, is that the main deck of the ship is almost even with the ground level of the shipyard, making it very easy for shipyard workers to get on and off the ship. Also, the dockside portal cranes have easy reach of the entire ship.

Figure 6 shows a graving dock being flooded and a frigate's dry dock blocking system being covered by the water entering the dock. Soon the water will be high enough in the dock to float the ship off the blocks.

Graving dock dry-dockings are also inherently safer operations as no pumping plan is required. The structural limit of the graving dock still must be considered, however, when determining if the floor capacity of the dock can handle the weight of the ship and blocking system.

Vertical lift dock. This type of dry dock is somewhat similar to an elevator. The platform, supported by winch cables, is lowered into the water. A ship is floated in over the platform; then the platform is lifted using electrically powered winches all synchronized at the same speed. The ship lands on the blocking system and the platform is raised up out of the water.

The blocking system is on a cradle supported by wheels. The platform has a track system whereby the ship can be towed off the platform and moved to a work area in the shipyard (**Fig. 7**). This frees up the platform for another docking evolution.

Alternating-current synchronous induction motors are used to power the cable winches that raise and lower the ship platform. This enables the winches to all operate at the same lift speed regardless of load. The strain or load on each cable is monitored. The load on the cables is actually a measurement of the combined load distribution of the ship, the blocking and carriage system, and the vertical lift platform. Off-center weights can also be detected if a level platform shows different loads on the cables on either side of the lift.

Fig. 7. USS *Elliot* (DD 967) on a mechanical or vertical lift at the Todd Pacific Shipyard at San Pedro, California, around 1984. When the shipyard closed in the early 1990s, the vertical lift system was disassembled and moved to Malaysia, where it is now operating (as of 2005). (*Rolls-Royce Naval Marine Inc.*)

The advantage of this type of dock is that the ships, once docked, can be moved to another location, so that the vertical lift is freed up to conduct other dockings. This arrangement greatly increases the capacity of a shipyard. Sand blasting operations do require barriers to be erected to reduce the spread of the grit, and access to the ship by shipyard workers is via stairways or elevators, making it more difficult.

Hull preservation and maintenance. One of the most important maintenance actions while a ship is in dry dock is painting the hull. Usually the ship's hull is blasted with abrasives down to white metal and covered with one or more primer coats, a layer of anticorrosive paint, and two coats of antifouling paint. It is critical throughout this process that the temperature and humidity be monitored properly and that the required drying times are followed to prevent premature and catastrophic paint coating system failures later on.

Other critical work typically includes propeller replacement, pulling shafts and replacing journal bearing surfaces, refurbishing sea valves, cleaning the sea chests, and replacing transducers on the bottom. Another critical maintenance action in dry dock is the blasting and painting of all tanks that touch the underwater hull. It is very difficult to get a coating system to properly adhere to the hull plating when cool water is on the outside because the sweating of the hull inside surface prevents proper bonding and drying of the coating system. Richard D. Hepburn; Lemuel C. Robertson

Bibliography. R. D. Hepburn, *History of American Naval Dry Docks: A Key Ingredient to a Maritime Power*, Noesis, Inc., 2003; B. K. Mazurkiewicz, *Design and Construction of Dry Docks*, 1980.

Dry ice

A solid form of carbon dioxide, CO_2, which finds its largest application as a cooling agent in the transportation of perishables. It is nontoxic and noncorrosive and sublimes directly from a solid to a gas, leaving no residue. At atmospheric pressure it sublimes at −109.6°F (−78.7°C), absorbing its latent heat of 246.4 Btu/lb (573.1 kilojoules/kg). Including sensible heat absorption, the cooling effect per pound (kilogram) of dry ice is approximately 270 Btu (628.0 kJ) at storage temperatures above 15°F (−9°C) and 250 Btu (581.5 kJ) at lower temperatures. Slabs of dry ice can easily be cut and used in shipping containers for frozen foods, in refrigerated trucks, and as a supplemental cooling agent in refrigerator cars.

The manufacture of carbon dioxide gas is a chemical process. The gas is liquefied by compressing it to 900–1000 lb/in.^{2} (6.2–6.9 megapascals) gage in three stages of reciprocating compressors and then condensing it in water-cooled condensers. The liquid is expanded to atmospheric pressure where its temperature is below the triple point (−69.9°F or −56.6°C). In the expansion there is a flash separation (at −109.6°F or −78.7°C) of some solid carbon dioxide "snow" which is very porous. The snow is removed from the expansion chamber and mechanically compressed into standard 50-lb (23-kg) blocks which measure 10 by 10 by 10 in. (25 by 25 by 25 cm). *See* CARBON DIOXIDE. Carl F. Kayan

Drying

An operation in which a liquid, usually water, is removed from a wet solid in equipment termed dryers. The use of heat to remove liquids distinguishes drying from mechanical dewatering methods such as centrifugation, decantation or sedimentation, and filtration, in which no change in phase from liquid to vapor is experienced. Drying is preferred to the term dehydration, which is sometimes used in connection with the drying of foods. Dehydration usually implies removal of water accompanied by a chemical change. Drying is a widespread operation in the chemical process industries. It is used for chemicals of all types, pharmaceuticals, biological materials, foods, detergents, wood, minerals, and industrial wastes. Drying processes may evaporate liquids at rates varying from only a few ounces per hour to 10 tons per hour in a single dryer. Drying temperatures may be as high as 1400°F (760°C), or as low −40°F (−40°C) in freeze drying. Dryers range in size from small cabinets to spray dryers with steel towers 100 ft (30 m) high and 30 ft (9 m) in diameter. The materials dried may be in the form of thin solutions, suspensions, slurries, pastes, granular materials, bulk objects, fibers, or sheets. Drying may be accomplished by convective heat transfer, by conduction from heated surfaces, by radiation, and by dielectric heating. In general, the removal of moisture from liquids (that is, the drying of liquids) and the drying of gases are classified as distillation processes and adsorption processes, respectively, and they are performed in special equipment usually termed distillation columns (for liquids) and adsorbers (for gases and liquids). Gases also may be dried by compression. *See* ADSORPTION; DISTILLATION.

Drying of solids. In the drying of solids, the desirable end product is in solid form. Thus, even though the solid is initially in solution, the problem of producing this solid in dry form is classed under this heading. Final moisture contents of dry solids are usually less than 10%, and in many instances, less than 1%.

The mechanism of the drying of solids is reasonably simple in concept. When drying is done with heated gases, in the most general case, a wet solid begins to dry as though the water were present alone without any solid, and hence evaporation proceeds as it would from a so-called free water surface, that is, as water standing in an open pan. The period or stage of drying during this initial phase, therefore, is commonly referred to as the constant-rate period because evaporation occurs at a constant rate and is independent of the solid present. The presence of any dissolved salts will cause the evaporation rate to be less than that of pure water. Nevertheless, this

lower rate can still be constant during the first stages of drying.

A fundamental theory of drying depends on a knowledge of the forces governing the flow of liquids inside solids. Attempts have been made to develop a general theory of drying on the basis that liquids move inside solids by a diffusional process. However, this is not true in all cases. In fact, only in a limited number of types of solids does true diffusion of liquids occur. In most cases, the internal flow mechanism results from a combination of forces which may include capillarity, internal pressure gradients caused by shrinkage, a vapor-liquid flow sequence caused by temperature gradients, diffusion, and osmosis. Because of the complexities of the internal flow mechanism, it has not been possible to evolve a generalized theory of drying applicable to all materials. Only in the drying of certain bulk objects such as wood, ceramics, and soap has a significant understanding of the internal mechanism been gained which permits control of product quality.

Most investigations of drying have been made from the so-called external viewpoint, wherein the effects of the external drying medium such as air velocity, humidity, temperature, and wet material shape and subdivision are studied with respect to their influence on the drying rate. The results of such investigations are usually presented as drying rate curves, and the natures of these curves are used to interpret the drying mechanism. **Figure 1** shows a series of typical drying-rate curves.

The constant-rate period of drying when heat is supplied by convection is susceptible to theoretical and analytical treatment because it is essentially independent of the solid material. When drying is accomplished by heat transfer from hot gases, which also remove the evolved vapors, the constant rate may be expressed in terms of heat-transfer rates or mass-transfer rates.

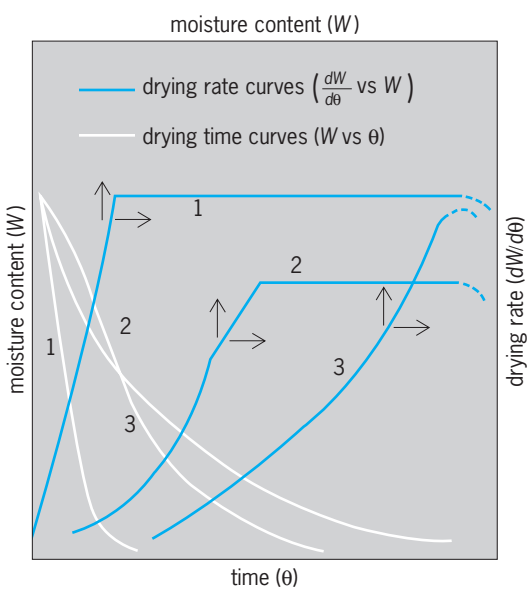

Fig. 1. Drying-time and drying-rate curves illustrating the general problem of drying: (1) Curves typical of a layer of thin material with most of the drying in the constant rate. (2) A more general case in which two stages in the falling-rate period occur. Typical of granular materials. (3) A case in which no constant rate occurs. Typical of homogeneous and colloidal materials such as soap, gelatin, and viscous solutions.

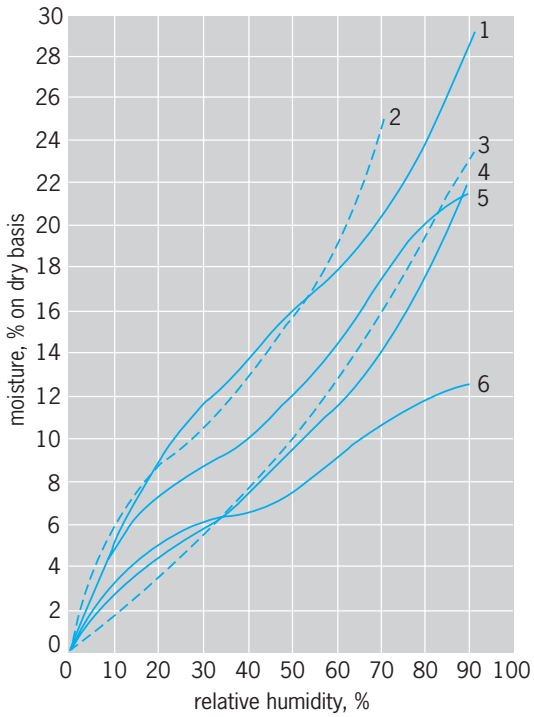

Fig. 2. Equilibrium moisture content of miscellaneous organic materials at 70°F (21°C): (1) leather; (2) tobacco; (3) soap; (4) wood; (5) catgut; (6) glue.

A constant rate of evaporation at the surface of the solid maintains the surface at a constant temperature, which, in the absence of other heat effects, is very nearly the wet-bulb temperature of the air. This temperature may range from 70 to 130°F (21 to 54°C) for convection drying, depending on the temperature and humidity of the air and on radiation. This so-called wet-bulb cooling effect is one reason why heat-sensitive solids can be dried in air at temperatures well above the decomposition temperature of the solid.

The magnitude of the constant rate can vary widely, depending on the degree of subdivision of the material, that is, the manner in which the material is exposed to the drying air. Thus, the rate of drying in spray dryers can be several hundred-thousand-fold greater than the rates in tray dryers.

A number of empirical expressions based on experimental studies have been developed for estimating the constant rate for different physical configurations of the wet material.

When materials are dried in contact with hot surfaces, termed indirect drying, the air humidity and air velocity may no longer be significant factors controlling the rate. The "goodness" of the contact between the wet material and the heated surfaces, plus the surface temperature, will be controlling. This may involve agitation of the wet material in some cases.

The falling-rate period is not as amenable to treatment as the constant-rate period because the falling

rate depends largely on the internal structure of the solid and the mechanism of moisture flow therein. In falling-rate processes, the rate of drying decreases gradually until the moisture content of the material approximates the equilibrium value. The equilibrium moisture content of a material is that moisture content to which a given material can be dried under specific conditions of air temperature and humidity. A typical equilibrium moisture curve is shown in **Fig. 2**. A unique characteristic of hygroscopic materials is that they hold or retain water at a vapor pressure less than water at the same temperature. Moisture so retained is termed bound moisture. Materials in which water exerts its normal vapor pressure at all moisture contents are termed nonhygroscopic, and in general, are easier to dry.

Classification of dryers. Drying equipment for solids may be conveniently grouped into three classes on the basis of the method of transferring heat for evaporation. The first class is termed direct dryers; the second class, indirect dryers; and the third class, radiant heat dryers. In the chart in **Fig. 3**, each class is subdivided into batch and continuous. Batch dryers are restricted to low capacities and long drying times. Most industrial drying operations are performed in continuous dryers. The large numbers of different types of dryers reflect the efforts to handle the larger numbers of wet materials in ways which result in the most efficient contacting with the drying medium. Thus, filter cakes, pastes, and similar materials, when preformed in small pieces, can be dried many times faster in continuous through-circulation dryers than in batch tray dryers. Similarly, materials which are sprayed to form small drops, as in spray drying, dry much faster than in through-circulation drying.

Direct dryers. The general operating characteristics of direct dryers are the following. (1) Drying is accomplished by convection heat transfer between the wet solid and a hot gas, the latter removing the vaporized liquid as well as supplying the heat needed for evaporation. (2) The heating medium may be steam-heated air, gases of combustion, a heated inert atmosphere such as nitrogen, or a superheated vapor such as steam. (3) Drying temperatures may range from prevailing atmospheric temperatures to 1400°F (760°C). (4) At drying temperatures below the boiling point of the liquid, increasing amounts of vapor of this liquid in the drying gas will decrease the rate of drying and increase the final liquid content of the solid. (5) When the drying temperatures are above the boiling point throughout the process, an increase in the vapor content of the gas or air, in general, will have no retarding effect on the drying rate, and no effect on the final moisture content. (6) For low-temperature drying, dehumidification of the drying gas is often required when high atmospheric humidities prevail. (7) The efficiency of direct dryers will increase with an increase in the inlet temperature of the drying gas at a fixed exhaust temperature.

Indirect dryers. The general operating characteristics of indirect dryers are as follows. (1) Drying by the transfer of heat by conduction and some radiation to the wet material; conduction usually occurs through a metallic retaining wall. The source of heat is generally condensing steam, but may also be hot water, gases of combustion, molten heat-transfer salts, hot oil, or electric heat. (2) The drying temperature of the surface of contact may range from below freezing to 1000°F (537°C). (3) Indirect dryers are especially suited to drying under reduced pressures and with inert atmospheres, and are therefore well adapted to the recovery of solvents. (4) Indirect dryers using condensing steam usually have a high efficiency because heat is supplied according to the demand, but, as in all dryers, the efficiency falls off markedly when very low final moisture contents are required. (5) Indirect dryers can handle dusty materials more readily than direct dryers. (6) The operation of indirect dryers is frequently characterized by some method of agitation to improve the contact between the hot metal surface and wet material. The nature of this contact determines the drying rate of indirect dryers. Therefore, heavy, granular materials generally give higher heat-transfer coefficients of contact than fluffy, bulky solids.

Radiant-energy dryers. These operate by the transfer of heat from a radiant source to the wet material being dried. The temperature of the radiant source may range from hot water or steam temperatures, 200–350°F (93–176°C), to the temperatures of incandescent surfaces, 1500–2500°F (815–1371°C). The generating medium may be steam, hot liquids, gas flames, or electricity, depending on the temperature desired and the equipment design.

Special types. Dielectric-heat dryers do not fall in any of the above classes inasmuch as their operation depends on the generation of heat by high-frequency fields inside the material being dried so that heat will actually flow out of the interior of the solid. These dryers are used to dry large bulky objects which have a long internal path for moisture flow.

Direct batch dryers. In operation, heated air circulates over the wet material being dried. The wet solid is supported according to its physical form. Lumber, ceramics, and similar massive objects are stacked in piles or on racks; textile skeins, painted objects, and hides are suspended on hangers; and granular materials, pastes, slurries, and liquids are placed in trays, which may be supported on stationary or movable racks. Good performance by this type of dryer depends on uniform, equal air velocities across all the wet material.

In batch–through-circulation dryers, heated air is blown through the wet material on screen-bottom trays instead of across. The material to be dried must be permeable to air flow.

Dryers of this type, but of binlike design, are used extensively in the explosives industry to dry gunpowder, and in food processing to dry and condition certain foodstuffs, such as grains and corn.

Direct continuous dryers. Tunnel and continuous tray dryers usually consist of long, enclosed housing or tunnels through which wet material is moved on trucks. Hot air is blown through the trucks. Air

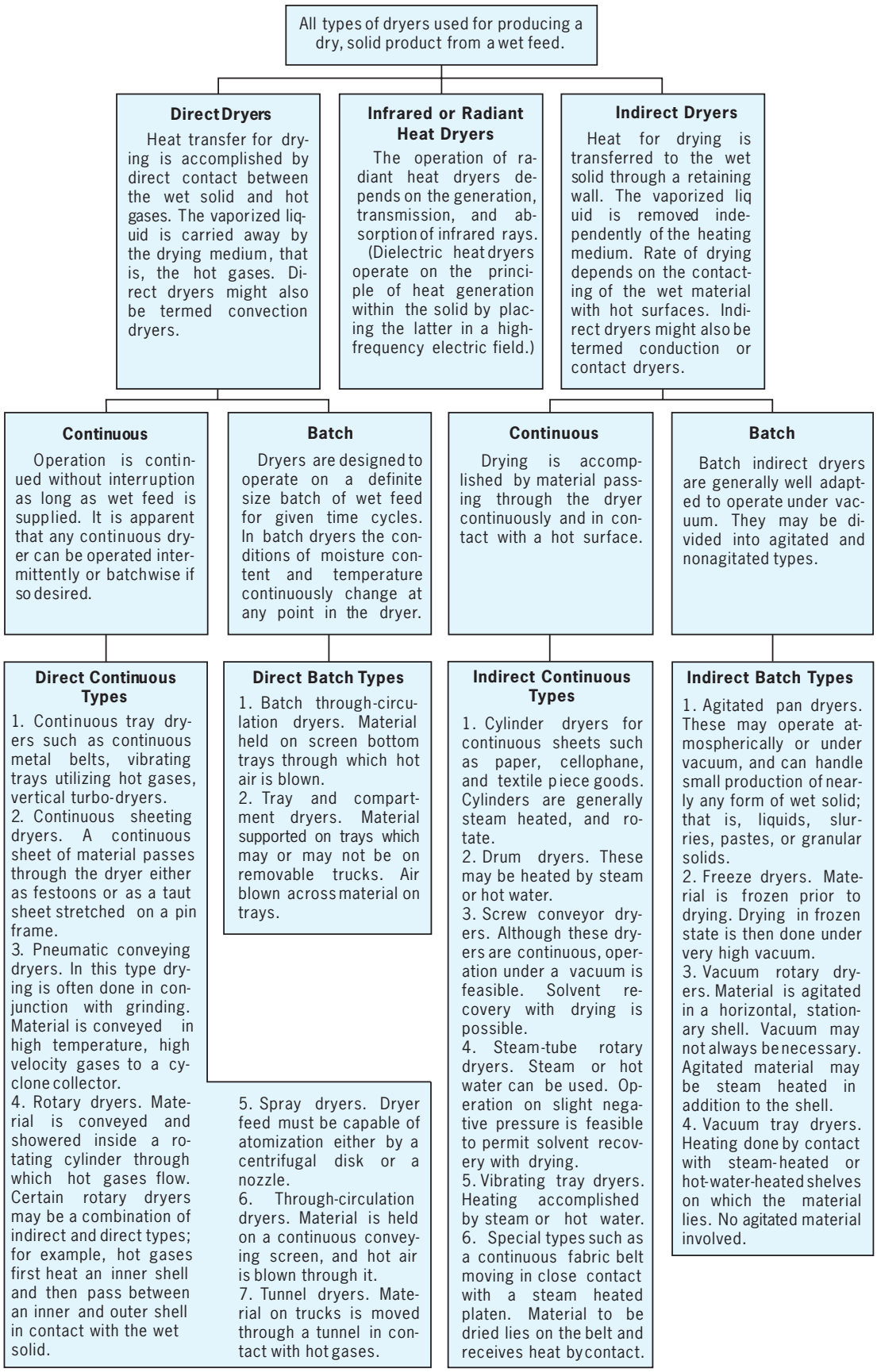

Fig. 3. Classification of dryers, based on methods of heat transfer.

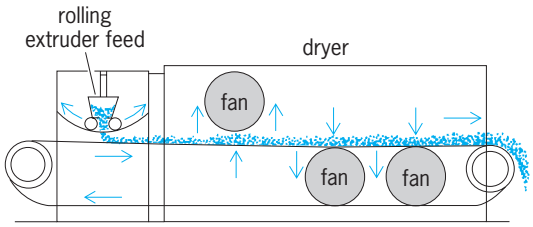

Fig. 4. Diagram of path of travel of permeable bed of wet material through a three-unit through-circulation dryer. (*Proctor and Schwartz, Inc.*)

flow may be parallel, counter flow, or at right angles (cross flow) to the movement of the trucks. The trucks may move continuously or semicontinuously through the tunnel. Tunnel dryers may be operated adiabatically, that is, without the addition of heat in the tunnel, or the air may be reheated periodically during its passage through the tunnel.

Wet granular materials are held on the trays of trucks; foodstuffs, rayon cakes, pottery, and large ceramic objects are held on racks; textile skeins are draped over rods; and hides are pasted on glass plates or hung on frames.

In continuous through-circulation dryers, heated air is blown through a permeable bed of wet material as it passes continuously through the dryer. Drying rates are much higher than in the usual tray or tunnel dryers because of the large surface area exposed per unit weight of material and because the smaller particles offer less resistance to internal moisture flow.

The operation of this type of dryer depends on whether or not the wet material is in a state of subdivision suitable for through circulation of the hot air. Some materials are already in such a permeable state. Many materials require special preliminary treatment, termed preforming, to form them into permeable beds. Preforming may include the processes of scoring on rotary filters, granulating, extruding, briqueting, flaking, and predrying on finned drums.

One type of through-circulation dryer consists of a horizontal conveying screen, which moves through a tunnellike housing. A permeable bed of the wet material is supported and conveyed on the screen, and hot air or gas is blown vertically up or down through the bed (**Fig. 4**). Through-circulation drying may also be performed in rotary-type dryers, which convey the material by a tumbling action imparted by the rotation of the dryer shell.

Conveying-screen through-circulation dryers are widely used in chemical plants to dry fibrous, flaky, and granular materials, such as cotton linters, rayon staple, cellulose acetate, silica gel, sawdust, and minerals, materials that are well suited to through circulation of air without prior treatment. They are also used extensively to dry materials, such as starch, pigments, calcium carbonate, insecticides, dyes, and intermediates, which must be preformed by one of the methods noted above.

Direct rotary dryers are used extensively in the chemical industry. A rotary dryer consists of a cylinder slightly inclined to the horizontal and rotated on suitable bearings. The rotary action of the cylinder serves to convey wet material from one end to the other while passing hot gases axially through the dryer shell. The contact of the solids and gases is further improved by means of flights arranged within the dryer shell so as to shower the wet material through the hot gas stream. A rotary dryer without such lifting flights is usually called a rotary kiln.

Rotary dryers may operate with air flow either parallel or countercurrent to the flow of the wet solid (**Fig. 5**).

Pneumatic conveying dryers, sometimes termed flash or dispersion dryers, operate on the basis of simultaneously conveying and drying a wet solid in a high-velocity stream of hot gas. Temperatures up to 1400°F (760°C) are used in these dryers. The short contact times involved permit usage gas temperatures above the decomposition temperature of the material. The gas stream acts as both the conveying and heating medium. Gas velocities on the order of 75 ft/s (23 m/s) are used. Frequently, the wet material is in such a form that some disintegrating action is required before it can be conveyed. A schematic diagram of this type of dryer is shown in **Fig. 6**. It is applicable to granular, free-flowing materials, such as coal, whey, and sodium chloride, and to sludges, filter-press cakes, and similar nongranular solids which require disintegration for proper dispersion. It is common practice to recycle dry product

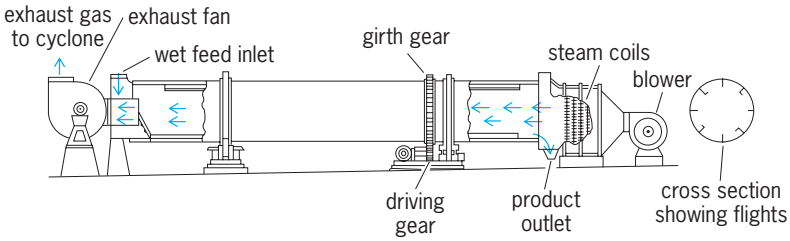

Fig. 5. Single-shell direct rotary using steam-heated air and balanced pressure by means of a blower and exhauster. (*Hardinge Co.*)

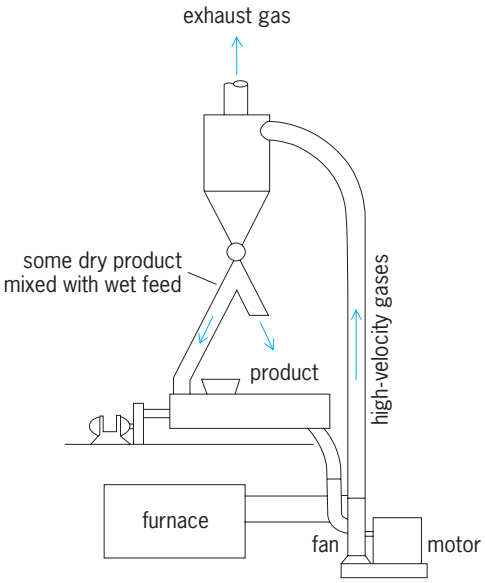

Fig. 6. Dispersion dryer.

into the wet feed in order to facilitate dispersion and handling.

Spray dryers operate on the principle of creating a highly dispersed liquid state in a high-temperature (up to 1400°F or 760°C) gas zone. The heart of a spray-drying process is the creation of small liquid droplets by spraying. This may be accomplished by means of (1) high-pressure nozzles, (2) pneumatic nozzles, and (3) high-speed rotating disks. Almost any pumpable liquid, from a thin clear liquid to a pasty sludge, can be atomized sufficiently for spray drying. Liquids above a viscosity of 1500 centipoises, however, are very difficult to atomize and spray dry. Generally, fine atomization will not produce a large percentage of droplets less than 5 micrometers in diameter. The particle size under conditions of so-called coarse atomization will be on the order of 200–600 μm. Because of the high surface-volume ratio of small drops, the actual drying time in spray dryers may be considerably less than 1 s for high-temperature operation.

Spray dryers are made in a multitude of designs. The drying gases may flow cocurrent with or countercurrent to the spray, and the spray may be directed vertically up or down, or horizontally. Various types of spray dryers are shown in **Fig. 7**. They are used extensively in the chemical, pharmaceutical, and food industries.

Fluidized-bed dryers operate by having heated air pass upward at sufficient velocity through a column or layer of granular, wet material to cause it to fluidize and become mixed by turbulent action. Wet feed may be introduced at the bottom of the column, and dry product removed from the top, as is shown in **Fig. 8**.

In direct continuous-sheeting dryers, heated air is circulated over or through a continuously moving wet sheet, which is supported by a variety of methods.

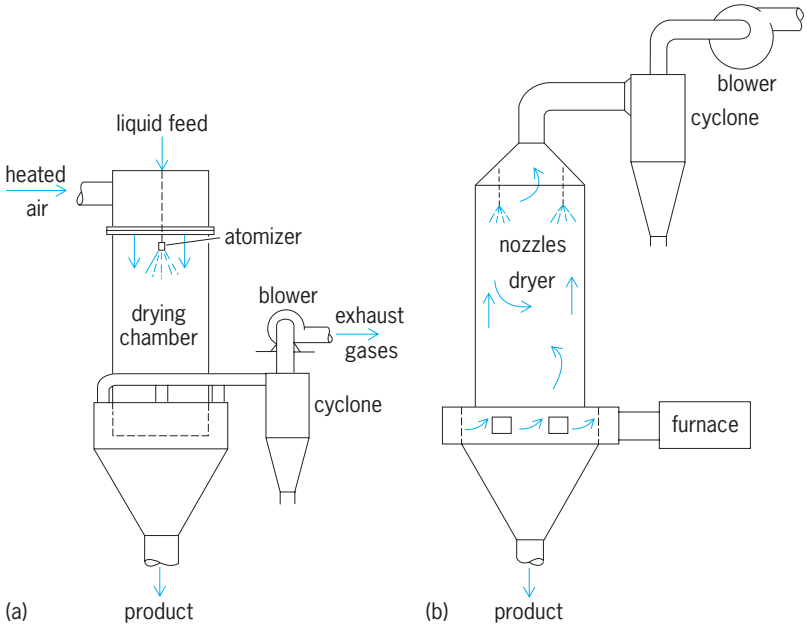

Fig. 7. Spray dryers. (*a*) Cocurrent. (*b*) Countercurrent.

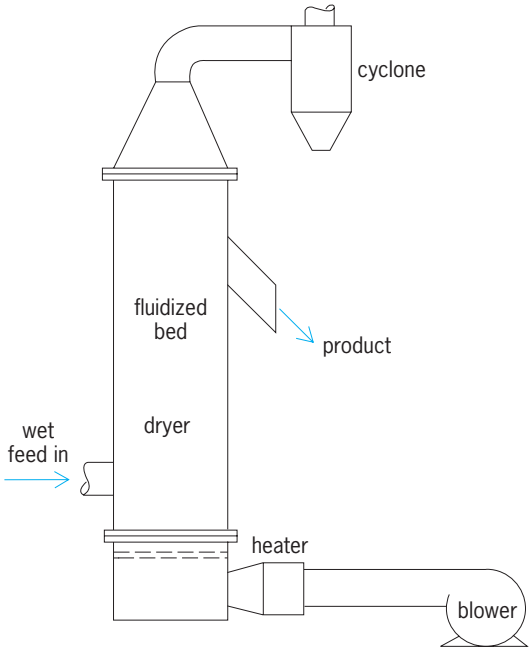

Fig. 8. Fluidized-bed dryer.

Direct continuous-sheeting dryers occur in a variety of types. The festoon or loop dryer permits drying of continuous sheet material in a relaxed state by festooning or looping it over rolls, which in turn are conveyed through the dryer. Tenter dryers are used to dry a continuous sheet of material under tension. Heated air is usually blown perpendicularly from slots or nozzles against both sides of the sheet.

Direct continuous-sheeting dryers are used to dry sheet materials that are sufficiently strong mechanically in the wet and dry states to support their own weight or withstand tension as required. Their widest use is in the textile field, in the manufacture of coated and impregnated fabrics, and in the preparation of films and coated papers.

Indirect batch dryers. Vacuum shelf dryers are indirect batch dryers which generally consist of a vacuum-tight cubical or cylindrical chamber of cast-iron or steel plate, heated supporting shelves inside the chamber, a vacuum source, and a condenser. In operation, heat transfer takes place by conduction through metal surfaces and interfaces to the wet material held on the shelves.

Vacuum shelf dryers are used extensively for drying pharmaceuticals, temperature-sensitive or easily oxidizable materials, and small batches of high-cost products where any product loss must be avoided. This type is particularly useful for recovery of valuable solvents or vapors.

Vacuum freeze dryers are used principally for drying materials that would be destroyed by the loss of volatile ingredients or by drying temperatures above the freezing point. The material is dried in the frozen state, so that a process of sublimation is involved, that is, ice sublimes directly to water vapor. Because the material dries in a rigid frozen condition, shrinkage is minimized, and the resulting structure of the dry solid is usually porous and readily soluble. This led

to the term lyophilization in the early development of freeze drying.

The equipment required for this method of drying consists of a heated drying chamber where sublimation occurs, a piping system for the transport of vapors, and a vapor-removal system for condensable and noncondensable gases. Removal of the condensable vapors may be accomplished by freezing on cold surfaces, absorption in liquid desiccants, or adsorption on solid desiccants.

One application of freeze drying (the Cryochem process) involves conduction heat transfer to the frozen solid held on a metallic surface. However, should the metal surface temperature rise above the freezing point of the solid, melting will occur. A second method of freeze drying utilizes heat transfer by radiation.

Agitated-pan dryers consist of a bowl- or pan-shaped receptacle in which wet material is stirred or agitated in contact with hot surfaces. The operation may be atmospheric or vacuum. Agitated-pan dryers are used to handle small batches of materials that can withstand agitation during drying. They are suitable for pastes and slurries, and for materials containing valuable solvents which must be recovered.

The vacuum rotary dryer is another type of batch indirect dryer with an agitator. It consists of a horizontal shell in which agitator blades attached to a horizontal rotating shaft revolve and agitate the material being dried (**Fig. 9**). Heat is supplied by condensing steam in a jacket surrounding the shell, by hot water, or by other heat-transfer fluids. Vacuum rotary dryers are used for large batches of materials that must be dried in the absence of air or where the recovery of solvents is required.

Indirect continuous dryers. Indirect rotary dryers are similar mechanically and in appearance to the direct rotary dryers discussed above. They differ primarily in that heat is transferred to the material through the metal shell or from steam tubes located around the dryer shell, rather than from hot gases as in the case of direct rotary dryers.

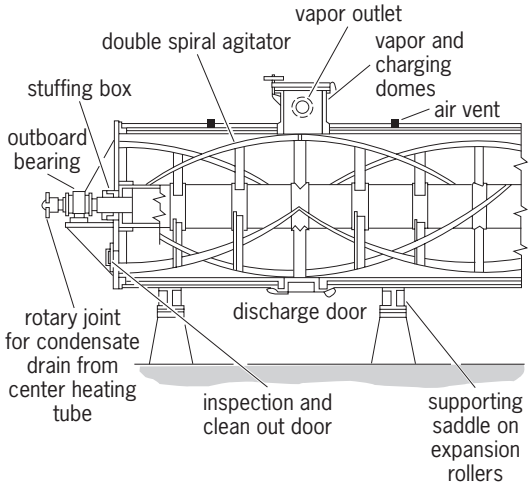

Fig. 9. Typical vacuum rotary dryer. (*Blaw-Knox Co.*)

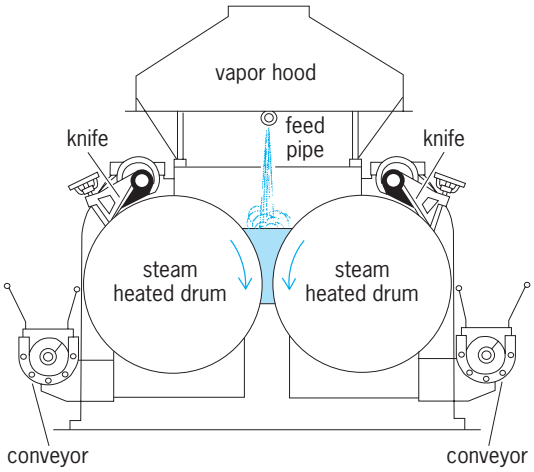

Fig. 10. Double-drum dryer with pipe feed. (*Buflovak Equipment Division, Blaw-Knox Co.*)

A screw-conveyor dryer is essentially a jacketed conveyor in which material is heated and dried as it is conveyed. In one type, the jacket may extend only to the top of the conveyor, which is left open to the atmosphere. This is termed a trough dryer. When the jacket encloses the conveyor completely, a slight negative pressure is required to sweep out the evaporated moisture.

Vibrating-conveyor dryers consist of a vibrating, heated, solid deck over which the wet solid moves. This deck can be heated by hot gases, steam, methanol vapors, or Dowtherm vapors, which pass through a jacket fastened to the deck upon which the material is conveyed. Direct gas flames can also be used as the source of heat. A hood equipped with an exhaust fan and placed over the deck removes the evaporated liquid. Infrared lamps may be mounted above the deck to increase drying rates.

Drum drying consists of applying a liquid material, solution, slurry, or paste to a revolving heated metal drum, which conducts heat to the wet film to evaporate the water during a partial revolution of the drum (**Fig. 10**). The dry material is scraped from the drum by a stationary knife. Drum dryers may be designated by type as atmospheric double-drum, atmospheric single-drum, atmospheric twin-drum, vacuum single-drum, and vacuum double-drum dryer.

In drum drying, the product is exposed to heat for only short periods of time. This has the advantage that, although the product may approach the temperature of the drum surface, there usually is no adverse effect from overheating.

Cylinder dryers, sometimes called can dryers or drying rolls, differ from drum dryers in that they are used for materials in a continuous-sheet form. Cylinder dryers may consist of one large cylindrical drum, such as the so-called Yankee dryer; more often they comprise a number of drums arranged so that a continuous sheet of material can pass over them in series. Typical of this arrangement are Fourdrinier paper-machine dryers, cellophane dryers, and slashers for textile piece goods and fibers.

The size of commercial cylinder dryers covers a wide range. The individual rolls may be 2–6 ft (0.6–1.8 m) in diameter and up to 20 ft (6 m) in width. In some cases, the width of the rolls decreases throughout the dryer in order to conform to the shrinkage of the sheet.

Drying of gases. The removal of 95–100% of the water vapor in air or other gases is frequently necessary. Gases having a dew point of −40°F (−40°C) are considered commercially dry. The more important reasons for the removal of water vapor from air are (1) comfort, as in air conditioning; (2) control of the humidity of manufacturing atmospheres; (3) protection of electrical equipment against corrosion, short circuits, and electrostatic discharges; (4) requirement of dry air for use in chemical processes where moisture present in air adversely affects the economy of the process; (5) prevention of water adsorption in pneumatic conveying; and (6) as a prerequisite to liquefaction.

Gases may be dried by the following processes: (1) absorption by use of spray chambers with such organic liquids as glycerin, or aqueous solutions of salts such as lithium chloride, and by use of packed columns with countercurrent flow of sulfuric acid, phosphoric acid, or organic liquids; (2) adsorption by use of solid adsorbents such as activated alumina, silica gel, or molecular sieves; (3) compression to a partial pressure of water vapor greater than the saturation pressure to effect condensation of liquid water; (4) cooling below dew point of the gas with surface condensers or coldwater sprays; and (5) compression and cooling, in which liquid desiccants are used in continuous processes in spray chambers and packed towers—solid desiccants are generally used in an intermittent operation that requires periodic interruption for regeneration of the spent desiccant.

Desiccants are classified as solid adsorbents, which remove water vapor by the phenomena of surface adsorption and capillary condensation (silica gel and activated alumina); solid absorbents, which remove water vapor by chemical reaction (fused anhydrous calcium sulfate, lime, and magnesium perchlorate); deliquescent absorbents, which remove water vapor by chemical reaction and dissolution (calcium chloride and potassium hydroxide); or liquid absorbents, which remove water vapor by absorption (sulfuric acid, lithium chloride solutions, and ethylene glycol).

The mechanical methods of drying gases, compression and cooling and refrigeration, are used in large-scale operations, and generally are more expensive methods than those using desiccants. Such mechanical methods are used when compression or cooling of the gas is required.

Liquid desiccants (concentrated acids and organic liquids) are generally liquid at all stages of a drying process. Soluble desiccants (calcium chloride and sodium hydroxide) include those solids which are deliquescent in the presence of high concentrations of water vapor.

Deliquescent salts and hydrates are generally used as concentrated solutions because of the practical difficulties in handling, replacing, and regenerating the wet corrosive solids. The degree of drying possible with solutions is much less than with corresponding solids; but, where only moderately low humidities are required and large volumes of air are dried, solutions are satisfactory. *See* DESICCANT; EVAPORATION; FILTRATION; HEAT TRANSFER; HUMIDIFICATION; UNIT OPERATIONS; VAPOR PRESSURE.

William R. Marshall, Jr.

Bibliography. T. Kudra and A. S. Mujumdar, *Advanced Drying Technologies*, 2001; A. S. Mujumdar, *Handbook of Industrial Drying*, 2d ed., 1995; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997.

Drying oil

A relatively highly unsaturated oil which, on exposure to air, is oxidized and polymerized to form a hard, dry film. Vegetable and fish oils are classified as nondrying, semidrying, or drying, according to their ease of autoxidation and polymerization. Due to their film-forming properties, drying oils are used in paints, inks, and varnishes. The reactivity of drying oils with oxygen results from the presence of diallylic groups (that is, two double bonds separated by methylene groups, —CH═CHCH$_2$CH═CH—) or conjugated double bonds (two carbon-carbon double bonds separated by a single bond). The number of diallylic groups *fn* per molecule is correlated with the oils' drying nature. Oils with *fn* greater than 2.2 are drying oils; those below 2.2 are semidrying to nondrying. The **table** shows the percentage of saturated and unsaturated fatty acids and the *fn* for some drying oils. These oils vary considerably by source and harvest time (see "Sunflower" in the table). *See* ALKENE; CONJUGATION AND HYPERCONJUGATION; FAT AND OIL; INK; VARNISH.

Studies of the reactions of ethyl linoleate with oxygen and driers (materials added to facilitate oxidation of oils) showed that cross-linking reactions mostly formed ether and peroxy cross-links with only about 5% of the cross-links being new C-C bonds. Polymerization of drying oils is accelerated at the surface with cobalt driers which accelerate the decomposition of hydroperoxides formed by oxidation. Zirconium is used as a through-film drier. The rearrangement and cleavage of hydroperoxides to aldehydes and ketones, during drying, causes much of the characteristic odor of oil and alkyd paints. *See* DRIER (PAINT); PAINT.

Drying oils are seldom used alone and primarily are raw materials for other binders such as oil-resin varnishes, heat-bodied oils (heated in an inert atmosphere at 300–320°C or 572–608°F), blown oils (heated in air at 150°C or 302°F), alkyd (polyester) resins, epoxy esters, and uralkyds (urethanes). These materials can be considered synthetic drying oils since *fn* has been raised above 2.2 by a chemical reaction. Alkyds are polyester resins formed by reacting vegetable oil or fatty acid with an aromatic diacid such as phthalic acid and a polyol such as

Percent fatty acid composition of selected oils

Name	Saturated[a]	Oleate, etc.	Linoleate	Linolenate	Other	*fn*
Castor	3	7	3		87[b]	0
Dehydrated castor	5	10	85			Conjugated
Coconut	91	7	2			0
Cottonseed	25	40	35			1
Fish oil (cod)					38[c]	6.9
Linseed	10	18	17	55		3.8
Oiticica	10	6	10		74[d]	4.8
Perilla	7	14	16	63		4.3
Safflower	11	13	75	1		2.3
Soybean	14	26	52	8		2.1
Sunflower (Minn.)[g]	13	26	61	Trace		1.8
Sunflower (Texas)	11	51	38	Trace		1.2
Tall oil fatty acids	8	46	41	3	2[e]	1.4
Tung	5	7	3		85[f]	Conjugated

[a]Saturated fatty acids are mainly mixtures of stearic (C_{18}) and palmitic (C_{16}) acids; coconut oil also contains C_8, C_{10}, C_{12}, and C_{14} saturated fatty acids.
[b]Ricinoleic acid.
[c]Clupanodonic (C22:6).
[d]Licanic.
[e]Rosin.
[f]α-Eleostearic acid (conjugated).

glycerin. Polyunsaturated acids dimerize by heat treatment with an acid catalyst. Dimer acids are predominantly C_{36} dicarboxylic acids. They are used in making polyesters and polyamides. *See* POLYAMIDE RESINS; POLYESTER RESINS. John L. Massingill, Jr.

Bibliography. M. Bockisch, *Fats and Oils Handbook*, AOCS Press, Champaign, IL, 1998; W. J. Muizebelt et al., *J. Coat. Technol.*, 70(876):83, 1998; Z. W. Wicks, Jr., Drying oils, *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., 8:519–532, 1993; Z. Wicks, Jr., F. Jones, and S. Pappas, *Organic Coatings: Science and Technology*, 2d ed., Wiley, 1999.

Dryolestida

The Dryolestida (formerly known as the Pantotheria or Eupantotheria) are an extinct order of mammals related to the living marsupials and eutherian (placental) mammals. Along with the closely related symmetrodonts, they were the most diverse group of mammals during the Jurassic. They are often found in some of the same quarries that yield the giant sauropod dinosaurs. Most were about the size and shape of a shrew or mouse, with teeth adapted for chopping up a diet of insects, and they apparently were adapted for living in trees like squirrels. *See* EUTHERIA; MARSUPIALIA; SYMMETRODONTA.

The earliest known dryolestoid is *Amphitherium*, first discovered in the Middle Jurassic Stonesfield Slate of England in 1838. This was an important specimen for historic reasons, proving that mammals once coexisted with the dinosaurs (at a time when it was thought that dinosaurs had died out before mammals appeared). By the Late Jurassic, dryolestoids were the most common mammals in both North America and Europe. They had seven genera of the family Dryolestidae in Europe and six in North America, as well as nine genera of the family Paurodontidae (mostly in North America). By the Early Cretaceous, dryolestids had been replaced in the Northern Hemisphere by the earliest relatives of marsupials and placentals, but they persisted until the Late Cretaceous of South America.

Of the 26 currently known genera of dryolestoids, all but one (*Henkelotherium*) are known only from remnants of teeth and jaws (see **illus.**). The teeth are very distinctive, with a number of unique specializations. The upper and lower molar teeth were made of tall triangles of three cusps, connected by crests between them. These molars were very compressed in a front-to-back direction, allowing as many as seven to nine molars in the cheek tooth row (marsupials typically have four molars; placentals have only three). The upper molar triangles (trigons) interlocked with the lower molar triangles (trigonids) like the teeth on a zipper, so that shear developed between the molars when the jaw closed. This type of dentition is similar to that found in living insectivorous mammals, and strongly suggests that dryolestoids were insectivorous as well. The upper molars have a large cusp on the exterior edge, which is called the stylocone, and another large cusp at the internal apex of the triangle, the paracone. In more advanced (tribosphenic) mammals such as marsupials and placentals, the paracone shifts to the outer edge and replaces the stylocone, and a new cusp, the protocone, develops in its place. The lower molar trigonids have a small ledge, the talonid, on the back end of the tooth, which occludes with the basin of the upper molar trigon. The paurodonts had a very similar dentition, but the cusps were lower and blunter. The teeth were less compressed in a front-to-back direction, and there were only four or five molars.

For over a century, dryolestoids were known only from dental remains. However, in 1991 a complete articulated skeleton of the paurodont *Henkelotherium* from the Late Jurassic beds of the Guimarota coal mine in Portugal was described. This skeleton was

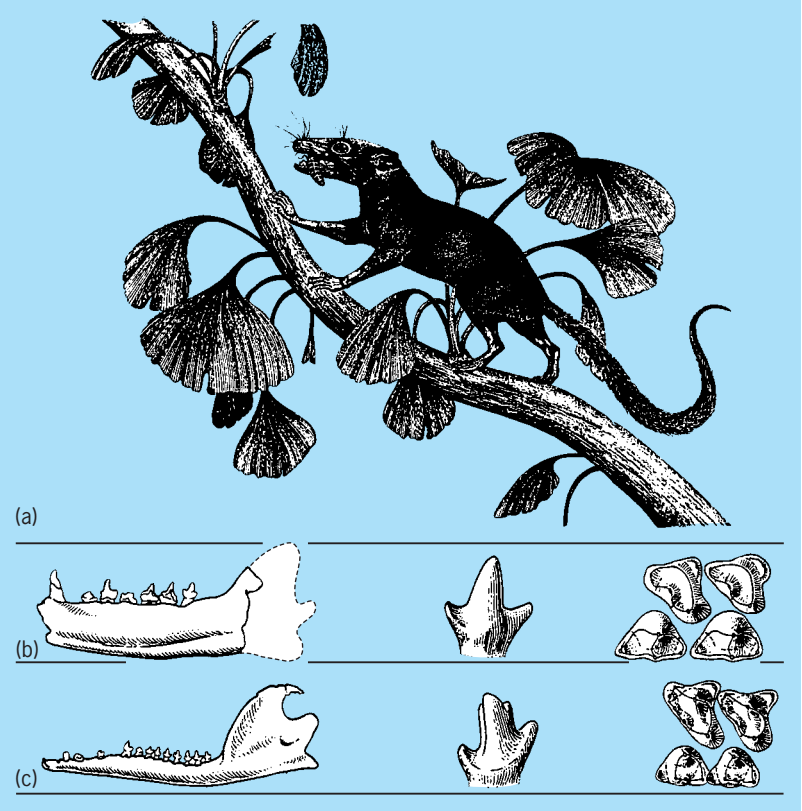

Dryolestids. (*a*) Reconstruction of *Henkelotherium* (*after Krebs, 1991*). (*b*) Side view of jaw, side view of teeth, and crown view of upper and lower molars of the paurodonts *Paurodon* and *Pelicopsis*; (*c*) side views for the dryolestids *Laolestes* and *Melanodon* (*after Prothero, 1981*).

about the size of a small mouse, but was clearly adapted for living in trees, having long clawed fingers for gripping bark and a long, slender tail for balance. The hip and shoulder bones were remarkably modern and like those of most living mammals. Many of the reptilian bones of the shoulder girdle (still found in primitive mammals such as the egg-laying platypus, and some other Mesozoic mammals) were lost. However, the marsupial bones, which project forward from the pelvis in marsupials and help support the pouch, were retained. In addition, rudimentary vestiges of some of the reptilian jaw bones (such as the coronoids and splenials) were still present, although most of the lower jaw was still made of a single bone, the dentary, as in most other mammals. *See* DENTITION.

Dryolestoid diversity peaked in the Late Jurassic, and only one Early Cretaceous dryolestid (*Crusafontia* from Spain) was known. Then, in 1991, a peculiar form, *Donodon*, was described from the Early Cretaceous of Morocco. Between 1986 and 1992, a number of new dryolestoids from the Late Cretaceous Los Alamitos Formation of Argentina were described. At least four different families are represented, and they show a much greater variety of types of teeth than are known from the Jurassic. They include normal-looking dryolestids (*Leonardus* and *Groebertherium*) like those known from the Late Jurassic, as well as the much larger, blunt-cusped *Mesungulatum* (originally mistaken for an ancient hoofed placental mammal) and *Reigitherium*, with its flat-crowned teeth. These specimens show that dryolestoids lingered on in South America long after the northern continents were dominated with marsupials and placentals, and long after the dryolestid heyday in North America and Europe in the Late Jurassic and Early Cretaceous had passed.

For over a century, dryolestoids were placed in the order Pantotheria, a taxonomic wastebasket for Mesozoic mammals that were not members of other groups. In the older literature, one can still find statements that the "pantotheres" were ancestral to marsupials and placentals. Since 1981, the taxon Pantotheria has been abandoned, because it was a mixture of unrelated animals, and also because there was a tendency to use it uncritically as an ancestral group without detailed analysis of its members. Recent research has shown that the dryolestoids were primitive in many features but highly advanced in others. Their complex mosaic of features showed the transition from archaic mammals such as the platypus and some extinct Mesozoic mammals (which still retain reptilian characteristics) to the marsupials and placentals. In addition, their unique specializations (such as seven to nine highly unusual molars) show that they were not ancestral to marsupials or placentals, but an early side branch of mammalian evolution that lived before and then contemporaneously with the earliest marsupials and placentals. *See* THERIA.

Donald R. Prothero

Bibliography. J. F. Bonaparte, New Late Cretaceous mammals from the Los Alamitos Formation, Northern Patagonia, *Nat. Geog. Res.*, 6(1):63–93, 1990; R.*nbsp;L. Carroll, *Vertebrate Paleontology and Evolution*, W. H. Freeman, 1988; D. R. Prothero, New Jurassic Mammals from Como Bluff, Wyoming, and the Interrelationships of the Non-Tribosphenic Theria, *Bull. Amer. Mus. Nat. Hist.*, 167(5):281–325, 1981.

Duality (physics)

The state of having two natures, which is often applied in physics. The classic example is wave-particle duality. The elementary constituents of nature—electrons, quarks, photons, gravitons, and so on—behave in some respects like particles and in others like waves.

Duality is often used in a more precise sense. It indicates that two, seemingly different, theoretical descriptions of a physical system are actually mathematically equivalent. Such an occurrence is very useful. Various properties and phenomena are clearer in one or the other of the descriptions, and calculations that are difficult or impossible in one description may be simple in the other. In the case of wave-particle duality, the wave description corresponds to a theory of quantized fields, where the field variables are governed by an uncertainty principle. The particle description corresponds to a Feynman integral over all particle paths in spacetime. The quantized-field and path integral theories sound very different but

are mathematically equivalent, making identical predictions. *See* FEYNMAN INTEGRAL; NONRELATIVISTIC QUANTUM THEORY; QUANTUM FIELD THEORY; QUANTUM MECHANICS; UNCERTAINTY PRINCIPLE; WAVE MECHANICS.

Wave-particle duality is a very general principle, but more specific dualities arise in particular systems. These were once believed to occur primarily in rather simple or model systems, but they have been discovered in the very theories that are of interest in formulating a unified description of the particle interactions, gravity, and quantum mechanics. These new dualities have led to many surprising insights.

Weak-strong duality. The strengths of the interactions in any system are governed by one or more coupling constants. A typical coupling constant may be denoted by the symbol g. When g is small, the interactions are weak. It is then possible to make the approximation that in any process the interaction acts only once, with a small correction from two interactions, and so on: this is known as perturbation theory. However, when g is large, perturbation theory is inaccurate, and it can miss important effects. For example, such phenomena as quark confinement and superconductivity cannot be studied with perturbation theory. *See* PERTURBATION (QUANTUM MECHANICS); QUARKS; SUPERCONDUCTIVITY.

In some systems, there is weak-strong duality, meaning that when the coupling constant g of the original description is large that of the dual description, g', is small; for example $g' = 1/g$. When g is large, so the interactions in the original description are strong and the perturbation theory in this description is highly inaccurate, then perturbation theory in the dual description gives a very accurate description.

Electric-magnetic duality. A special form of weak-strong duality is electric-magnetic duality. Maxwell's equations, governing the behavior of the electromagnetic field, have a striking symmetry between the electric and magnetic fields, except that electric field lines can end on electric charges while magnetic field lines cannot end—there are no magnetic charges. The suggestion is that magnetic charges (also called magnetic monopoles) should exist, thus making the equations more symmetric. In fact, almost all theories that unify electromagnetism with the nuclear or gravitational interactions predict that such magnetic charges can exist. Generally the magnetic charges are superheavy, explaining why they have not been seen in particle accelerators. They might also have been produced in the big bang; to explain why they have not been seen was one of the motivations for inflationary cosmology. *See* BIG BANG THEORY; INFLATIONARY UNIVERSE COSMOLOGY; MAGNETIC MONOPOLES; MAXWELL'S EQUATIONS.

The electric and magnetic charges in these unified theories are rather different from one another. The electric charges are described as points, like the electron. The magnetic charges, however, arise as solitons—fuzzy lumps of field. Moreover, the electric charges are light and weakly coupled, while the magnetic charges are heavy and strongly coupled. To be precise, this is the situation when the coupling g is small. The Montonen-Olive conjecture is that when the coupling g becomes large, then in certain theories the picture reverses: the electric charges become fuzzy, heavy, and strongly coupled, while the magnetic charges become pointlike, light, and weakly coupled. Thus there is a dual description, with magnetically rather than electrically charged fields. The solitons of one description are the pointlike particles of the other. *See* SOLITON.

There is no proof of the Montonen-Olive conjecture, because adequate methods to analyze the large-g theory still do not exist. However, there is compelling circumstantial evidence that it is true. An important point is that the conjecture applies specifically to unified theories with supersymmetry. This is a symmetry between fermionic and bosonic particles, and is believed to be a necessary part of a fully unified theory. Supersymmetry has an interesting mathematical structure, which allows certain properties (some of the masses and interactions) of the large-g theory to be obtained exactly. This does not give a complete solution of the theory, but in combination with other physical input it gives a fairly detailed picture of the physics, and in particular gives strong evidence for Montonen-Olive duality. *See* SUPERSYMMETRY.

Duality in superstring theory. It is believed that a complete theory of all particles and interactions must be based on quantization of one-dimensional objects (loops) rather than points: this is superstring theory. In superstring theory there is again the problem that perturbation theory is the main tool, giving an incomplete description of the physics. The situation has greatly improved with the discovery that weak-strong duality is a general property of string theory. In fact, there are five known string theories, and all are dual to one another. A notable feature in string theory is that in addition to strings and solitons, duality requires certain other objects as well: D-branes, which are local disturbances to which strings become fixed.

The current understanding (*see* **illus.**) is that there is a single theory, with several coupling constants (which are actually spacetime fields). In different limits the physics is described by one of the weakly coupled string theories, or by something new—a

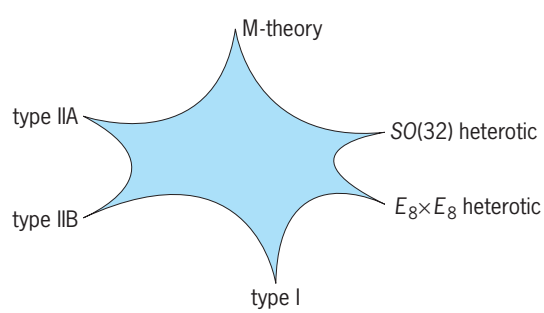

The five string theories and M-theory as limits of a single theory. The shaded region is parametrized by two coupling constants.

highly symmetric but only partially understood theory known as M-theory. The situation is analogous to the phase diagram for water (H_2O), where the different phases (solid, liquid, and gas) appear under different conditions, in analogy to the different string theories. The methods based on supersymmetry and duality have allowed an accurate mapping of the phase diagram of string theory, and have led to the discovery of new phases such as M-theory as well. *See* THERMODYNAMIC PROCESSES.

Remarkably, the same methods have also been used to solve some long-standing problems regarding the quantum mechanics of black holes. These solutions include a microscopic understanding of the analogy between black hole mechanics and thermodynamics, and progress toward the resolution of the information problem (an apparent conflict between quantum mechanics and general relativity). A further development is Maldacena duality, between black holes in string theory and ordinary quantum field theories with many fields. This has led to new understandings in field theory, string theory, and quantum gravity. *See* BLACK HOLE; QUANTUM GRAVITATION; SUPERSTRING THEORY. Joseph Polchinski

Bibliography. M. J. Duff, The theory formerly known as strings, *Sci. Amer.*, 278(2):64–69, February 1998; M. Mukerjee, Explaining everything, *Sci. Amer.*, 274(1):88–94, January 1996; J. Polchinski, String duality, *Rev. Mod. Phys.*, 68:1245–1258, October, 1996; E. Witten, Duality, spacetime, and quantum mechanics, *Phys. Today*, 50(5):28–33, May 1997.

Dubnium

A chemical element, symbol Db, atomic number 105. It was synthesized and identified unambiguously in March 1970 at the heavy-ion linear accelerator (HILAC) at the Lawrence Radiation Laboratory, Berkeley, University of California. The discovery team consisted of A. Ghiorso and colleagues. *See* PERIODIC TABLE.

The dubnium isotope, with a half-life of 1.6 s, decayed by emitting alpha particles with energies of 9.06 (55%), 9.10 (25%), and 9.14 (20%) MeV. It was shown to be of mass 260 by identifying lawrencium-256 as its daughter by two different methods.

Previous work on dubnium was reported in 1968 by G. N. Flerov and colleagues at Dubna Laboratories in Russia. They claimed to have discovered two isotopes of dubnium produced by the bombardment of ^{243}Am by ^{22}Ne ions. However, the Lawrence Radiation Laboratory work did not confirm these findings (due to energy and decay differences). *See* TRANSURANIUM ELEMENTS. Albert Ghiarso

Bibliography. S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; L. R. Norss and J. Fuger (eds.), *Transuranium Elements: A Half Century*, 1992; G. T. Seaborg and W. D. Loveland, *The Elements Beyond Uranium*, 1990.

1												13	14	15	16	17	18
1 **H**	2																2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Ducted fan

A propeller or multibladed fan inside a coaxial duct or cowling, also called a ducted propeller or a shrouded propeller, although in a shrouded propeller the ring is usually attached to the propeller tips and rotates. The duct serves to protect the fan blades from adjacent objects and to protect objects from the revolving blades, but more importantly, the duct prevents radial flow of the fluid at the blade tips. Fan efficiency remains high over a wider speed range with a properly shaped duct than without. However, fan efficiency is sensitive to duct shape at off-center design conditions. Without a well-rounded inlet lip and a variable-area exit, off-center performance may be worse than without the duct. *See* VENTILATION.

With a duct, static thrust for a given power input is higher than without one (**Fig. 1**). For this reason, propellers of vertical takeoff and landing aircraft may be ducted. At low speeds, a stator of radial airfoils downstream from the propeller or an oppositely turning coaxial propeller improve efficiency by converting slipstream rotation into axial velocity. The duct may also form a nozzle to further increase exit jet velocity. Airflow past the outer contour of the duct influences overall performance (**Fig. 2**). *See* VERTICAL TAKEOFF

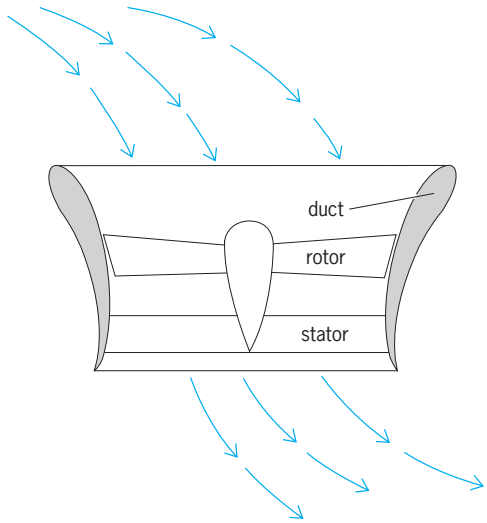

Fig. 1. Static thrust and operating speed range of fan is improved by a duct.

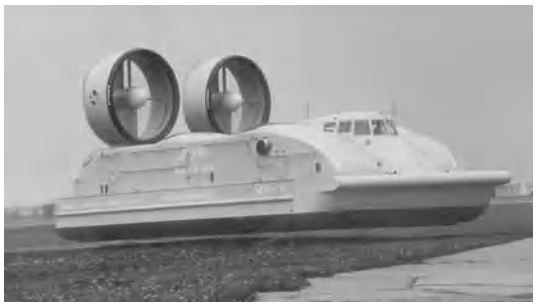

Fig. 2. Two ducted fans provide propulsion for Bell-Navy Hydroskimmer. (*Bell Aerosystems Co.*)

AND LANDING (VTOL).

Ducted fans are used in axial-flow blowers or compressors of several stages for turbine engines. In such applications, solidities are higher than for usual propellers, and stators or countervanes are usual in each stage. A ducted fan engine is a gas turbine arranged to move a larger mass of air than passes through the turbine, the additional air leaving at lower exit velocity and hence higher jet propulsion efficiency for moderate-speed aircraft than obtainable with a simple turbojet. *See* AIRCRAFT PROPULSION; CONVERTIPLANE; FAN; GAS TURBINE; TURBOFAN; TURBOJET.

Frank H. Rockett

Dumortierite

A nesosilicate mineral having the composition $Al(BO)_3(SiO_4)_3O_3$. Dumortierite crystallizes in the orthorhombic system but well-formed crystals are rare; the mineral usually occurs in parallel or radiating fibrous aggregates. There is one direction of poor cleavage. The hardness is 7 on Mohs scale, and the specific gravity is 3.26–3.36. The mineral has a vitreous luster and a color that varies not only from one locality to another but in a single specimen. It may be pink, green, blue, or violet. Dumortierite is found in schists and gneisses and more rarely in pegmatites. In the United States it occurs at Dehesa, California, and at Rochester, Nevada, where it has been mined for the manufacture of high-grade porcelain. *See* SILICATE MINERALS.

Cornelius S. Hurlbut, Jr.

Dune

Mobile accumulation of sand-sized material that occurs along shorelines and in deserts because of wind action. Dunes are typically located in areas where winds decelerate and undergo decreases in sand-carrying capacity. Dunefields are composed of rhythmically spaced mounds of sand that range from about 3 ft (1 m) to more than 650 ft (200 m) in height and may be spaced as much as 5000 ft (1.5 km) apart. Smaller accumulations of windblown sand, typically ranging in height from $^1/_4$ to $^1/_2$ in. (5 to 15 mm) and in wavelength from 3 to 5 in. (7 to 12 cm), are known as wind ripples. Dunes and ripples are two distinctly different features. The lack of intermediate forms shows that ripples do not grow into dunes. Ripples commonly are superimposed upon dunes, typically covering the entire upwind (stoss) surface and much of the downwind (leeward) surface as well.

Grain size and shape. Sand, which includes any material between 0.002 and 0.08 in. ($^1/_{16}$ and 2 mm) in diameter, moves across and between dunes primarily by the process of saltation (**Fig. 1**). Individual saltating sand grains move at nearly the same velocity as the wind and rarely reach more than 6 ft (2 m) above ground surface. When a fast-moving grain "splashes" into the underlying sand, it ejects other grains into the wind, thus perpetuating the transport process. The high-velocity, grain-to-grain impacts that take place during saltation transport produce well-rounded sand particles with distinctive surface pits that can be observed with the scanning electron microscope. Particles ranging from 0.08 to 0.2 in. (2 to 5 mm) in diameter creep slowly downwind as they are hit by saltating grains; these grains make up a small proportion of some dunes. Larger grains cannot be moved by normal winds and are therefore generally lacking in dune deposits. When particles finer than sand size are set into motion by the wind, they travel in suspension within high dust clouds. For this reason, silt and clay particles are not abundant in dunes but are deposited far downwind as broad sheets of loess. *See* LOESS; SAND.

Materials. Virtually any kind of sand-sized material can accumulate as dunes. The majority of dunes are composed of quartz, an abundant and durable mineral released during weathering of granite or sandstone. Dunes along subtropical shorelines, however, are commonly composed of grains of calcium carbonate derived in part from the breakdown of shells and coral. Along the margins of seasonally dry lakes, dunes may be composed of gypsum (White Sands, New Mexico) or sand-sized aggregates of clay minerals (Laguna Madre, Texas). *See* CLAY MINERALS; GYPSUM; QUARTZ.

Dynamics. The leeward side of most dunes is partly composed of a slip face, that is, a slope at the angle of repose. For dry sand, this angle is approximately 33°.

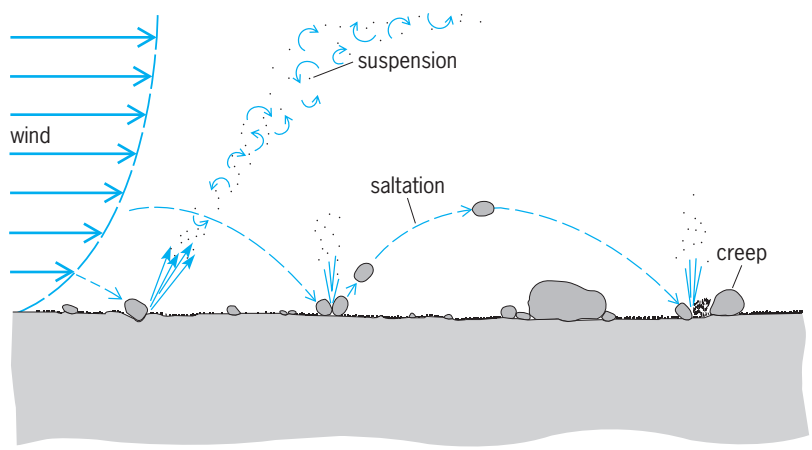

Fig. 1. Wind moves sand-sized particles by the process of saltation. Coarser material may creep or be immobile; finer material travels in suspension. Dunes are relatively free of materials coarser or finer than sand. (*After T. A. Mutch et al., Geology of Mars, Princeton University Press, 1976*)

Fig. 2. Tongue-shaped sand avalanches on the slip face of a barchanoid ridge dune on the eastern coast of Australia. (*Courtesy of D. B. Loope*)

When additional sand is deposited at the top of such a slope, tonguelike masses of sand avalanche to the base of the slope (**Fig. 2**). The dune migrates downwind as material is removed from the gently sloping stoss side of the dune and deposited by avalanches along the slip face. Much of the sand on the leeward side of the dune is later reworked by side winds into wind ripple deposits. Because the coarsest grains preferentially accumulate at the crests of wind ripples, the layering in wind ripple deposits is distinctive and relatively easily recognized—each thin layer is coarser at its top than at its base.

Sand deposited by avalanching is very loosely packed; up to 50% of its volume is occupied by air space. Wind ripple deposits are much more closely packed. A person with a knowledge of dune dynamics can choose a path which goes over firm wind ripple deposits and avoids loose avalanche deposits. Those less knowledgeable are likely to become bogged when their feet or tires sink deeply into the avalanche deposits.

Types. Dunes can be classified on the basis of their overall shape and number of slip faces. Three kinds of dunes exist with a single slip face; each forms in areas with a single dominant wind direction. Barchans are crescent-shaped dunes; their arms point down-wind. They develop in areas in which sand is in small supply. If more sand is available, barchans coalesce to form sinuous-crested dunes called barchanoid ridges. Transverse dunes with straight crests develop in areas of abundant sand supply. The axis of each of these dune types is oriented at right angles to the dominant wind, and the dunes migrate rapidly relative to other dune types. The migration rate of individual dunes is quite variable, but in general, the larger the dune, the slower the migration rate. In the Mojave Desert of southern California, barchans having slip faces 30 ft (10 m) long migrate about 50 ft (15 m) per year.

Dunes having more than one slip face develop in areas with more complex wind regimes. Linear dunes, sometimes called longitudinal or self dunes, possess two slip faces which meet along a greatly elongated, sharp crest (**Fig. 3**). Some linear dunes in Saudi Arabia reach lengths of 120 mi (190 km). Recent experimental evidence has shown that linear dunes are the result of bidirectional winds that differ in direction by more than 90°. The trend of these dunes is controlled by wind direction, strength, and duration, but the nature of the wind regime cannot be deduced from a knowledge of dune trends. Star dunes bear many slip faces and consist of a central, peaked mound from which several ridges radiate. Because they do not migrate appreciably, they grow in height as sand is delivered to them, some reaching 1000 ft (300 m).

Plant growth appears to be important to the growth and maintenance of two types of dunes. Coppice dunes are small mounds of sand that are formed by the wind-baffling and sand-trapping action of desert plants. The crescentic shape of parabolic dunes gives them a superficial resemblance to barchan dunes, but their arms point upward. Plants commonly colonize and anchor only the edges of a dune, leaving the body of the dune free to migrate. The retarded migration rate of the dune margin leads to the formation of the trailing arms of a parabolic dune. *See* DUNE VEGETATION.

Other dune types are dependent on special topographic situations for their formation. Climbing dunes develop on the upwind side of mountains or cliffs; falling dunes are formed at the sheltered, downward margin of similar features.

Distribution in space and time. Land plants are sufficiently abundant in most areas of the Earth to protect soil or sediment particles from wind erosion. It is only in areas that lack vegetation, such as along beaches and in deserts, that sufficient sand is moved to produce dunes. Unlike Earth, Mars lacks a vegetative cover, and dunes are abundant over a large portion of its surface.

Although all deserts of the world possess dunes, deserts occupy less than 1% of the arid portions of the southwestern United States. In the deserts of northern Africa, Saudi Arabia, and Australia, dunes cover approximately 30% of the ground surface. Coastal dunes are widespread except in the tropics, where heavy rainfall, dense vegetation, and relatively light winds severely limit their formation.

Deeply weathered, tree-covered dunes in South America and Africa indicate that during the last ice age the tropics were much more arid than today. Cooler global temperatures during this period

Fig. 3. Linear dune, Imperial County, California.

apparently decreased rainfall and increased wind strength, yielding conditions favorable for the expansion of dune fields.

Before the evolution and diversification of land plants on Earth, dunes may have been more widespread than today. Sandstones representing lithified, ancient dunefields have been identified on nearly all continents and are especially numerous and thick in the southwestern United States. In southern Utah, the Navajo Sandstone of Early Jurassic age (approximately 200 million years before present) reaches a thickness of over 2000 ft (600 m). The long, steeply inclined layers, or cross-beds, that were deposited during dune migration are characteristic features of these sandstones. Because dunelike mounds of sand can also form and migrate underwater, the most significant diagnostic criterion for recognizing eolian (windblown) sandstones is the presence of the distinctive wind ripple deposits. Eolian sandstones commonly contain animal trackways and traces of plant roots, and they are devoid of marine fossils larger than sand size. Typically they are found interbedded with ancient coastal or river deposits and may form important reservoirs for oil, natural gas, and fresh water. *See* EOLIAN LANDFORMS; SANDSTONE.

D. B. Loope

Bibliography. A. D. Abrahams and A. J. Parsons (eds.), *Geomorphology of Desert Environments*, 1993; R. A. Bagnold, *The Physics of Blown Sand and Desert Dunes*, 1971; R. Greeley and J. D. Iverson, *Wind as a Geological Process*, 1985; N. Lancaster, *Geomorphology of Desert Dunes*, 1995; E. D. McKee (ed.), *A Study of Global Sand Seas*, USGS Prof. Pap. 1052, 1979; K. Pye and H. Tsoar, *Aeolian Sand and Sand Dunes*, 1990.

Dune vegetation

Plants occupying sand dunes and the slacks, or swales, and flats between them. The density and diversity of dune vegetation are greater on coastal dunes than on desert dunes. *See* DUNE.

Zonation. A zonation pattern is evident in the vegetation of the coastal dunes (**Fig. 1**). A wrack, or debris, line occurs at the upper limit of the beach. Seeds caught in decaying plant material and other debris washed in on the high tides germinate here and trap windblown sand, initiating the formation of a dune. Small mounds formed in this way may eventually become foredunes as more sand accumulates and as the foredune grasses invade these mounds. Decomposing plant material (detritus) washed in on the tides provides additional nourishment for the young plants.

The foredunes, also called the primary dunes, are those closest to the water and lie behind the wrack line. The plants on these dunes, mostly grasses, are tolerant to sea spray, high winds, and sand accretion. By a system of underground stems called rhizomes, they overcome burial by sand and spread throughout the dune with new shoots arising from buds on the rhizomes. The roots and rhizomes of these dune grasses are important in stabilizing the dune sand and preventing wind erosion. The foredune plants participate in dune formation; by slowing the wind, they favor sand deposition.

Behind the primary dunes are the secondary dunes, sometimes called the dune field. In this more favorable environment, the vegetation is denser and more diverse; the foredunes block sea spray and reduce wind velocity. Here are found plants adapted to dry land and those tolerant of flooding, which often occurs in the dune slacks. The dune slacks are the low areas between dunes and are frequently a result of a blowout, an area where sand has been blown away down to where the sand is moist and close to the water table. Plants typical of wetlands often vegetate these areas. Shrub communities also inhabit the dune field and often form dense patches of vegetation.

A maritime forest may be found behind the secondary dunes. In coastal barrier beach or island locations, a salt marsh adjacent to a bay or sound may lie behind the forest.

Plant communities. Not all the plant species occupying the sand dunes of the United States can be described here. Although the names of the plants vary from coast to coast, their functions and adaptations are essentially the same. Sea rocket (*Cakile maritima*), probably the most widely distributed wrack-line beach plant, can be found in many coastal regions. It is an annual plant that initiates dune formation by trapping sand and debris.

The foredune species are primarily responsible for forming and stabilizing the seaward-facing row of

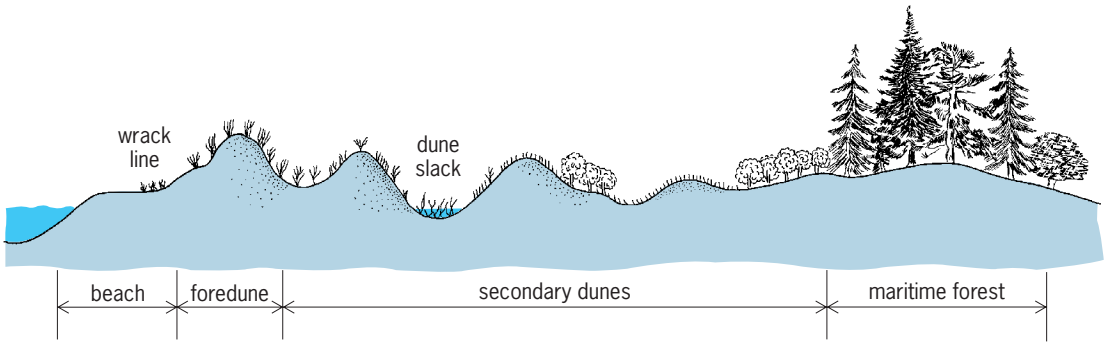

Fig. 1. A coastal dune profile.

Fig. 2. Sea oats are abundant on the foredunes of the Atlantic coast of the southeastern United States. Wrack, decomposing plant material, can be seen on the beach in the zone seaward of the sea oats.

dunes. American beachgrass (*Ammophila breviligulata*) is the dominant pioneer plant on the dunes of the North Atlantic and mid-Atlantic coast. This plant and prairie sandreed (*Calamovilla longifolia*) dominate the dunes along the shores of the Great Lakes. Sea oats (*Uniola paniculata*) dominates the foredunes of the South Atlantic coast (**Fig. 2**) and, with bitter panicum (*Panicum amarum*), is the pioneer species on the Gulf of Mexico dunes. On the northern west coast, European beachgrass (*Ammophila arenaria*) and American dunegrass (*Elymus mollis*) dominate. On the southern west coast, species of *Carpobrotus, Artemisia, Ambrosia,* and *Abrona* are most abundant.

In the secondary dunes the diversity increases. On the Atlantic coast, shrub communities include bayberry (*Myrica pensylvanica*), wax myrtle (*Myrica cerifera*), beach plum (*Prunus maritima*), and poison ivy (*Rhus radicans*). In the south, saw palmetto (*Serenoa repens*) and yucca (*Yucca filamentosa*) are common. On the Pacific coast, the shrub community includes salal (*Gaultheria shallon*), evergreen huckleberry (*Vaccinium ovatum*), chaparral broom (*Baccharis pilularis*), California figwort (*Schrophularia californica*), hooker willow (*Salix hookeriana*), and Pacific wax myrtle (*Myrica californica*). Seaside goldenrod (*Solidago sempervirens*), running beach grass (*Panicum amarum*), sand bur (*Cenchrus tribuloides*), beach heather (*Hudsonia tomentosa*), and beach pea (*Lathyrus maritimus*) are common on Atlantic and Gulf coast dunes. On the Pacific coast, common herbaceous plants are red fescue (*Festuca rubra*), dune goldenrod (*Solidago spathulata*), beach sagewort (*Artemisia pycnocephala*), and seashore lupine (*Lupinus littoralis*). On the dunes of the Great Lakes, blue joint (*Calamagrostis canadensis*), beach bea (*Lathyrus maritimus*), tansy (*Tanacetom huronense*), scouring rush (*Equisetum hyemale*), and beach heather (*Hudsonia tomentosa*) are common. Dune slacks of the Atlantic coast contain sedges such as American three-square (*Scirpus americanus*), rushes (*Juncus*), and grasses such as salt grass (*Distichlis spicata*). In the Pacific dune slacks, Douglas's spirea (*Spiraea douglasii*), slough sedge (*Carex obnupta*), Pacific silverweed (*Potentilla pacifica*), and common cattail (*Typha latifolia*) are common.

Forest species of the Atlantic and Gulf coast dunes include pitch pine (*Pinus rigida*), loblolly pine (*Pinus taeda*), sassafras (*Sassafras albidum*), wild black cherry (*Prunus serotina*), live oak (*Quercus virginiana*), and red cedar (*Juniperus virginiana*). Pacific dune forests include lodgepole pine (*Pinus contorta*), Douglas-fir (*Pseudotsuga menziesii*), western hemlock (*Tsuga heterophylla*), and sitka spruce (*Picia sitchensis*). Balsam poplar (*Populus balsamifera*) is the most common forest species of the Great Lakes dunes.

Ecology. Plants growing on sand dunes are adapted to the environment. The plants closest to the sea are usually the most tolerant of salt spray. The plants in the wrack line must be tolerant to salinity, wind, and burial by sand. Succulence helps sea rocket to tolerate the salty environment of the beach. It stores water in its fleshy stems and leaves, helping to dilute the salt taken up by its roots. Its tough cuticle withstands sand abrasion. *See* PLANTS, SALINE ENVIRONMENTS OF.

The foredune plants must be tolerant of sand burial, sea spray, and a nutrient-poor substrate. American beachgrass, for example, grows more vigorously in a location of accreting sand than in a more stable area. Its rhizomes, like those of other foredune grasses, can grow and spread rapidly as new sand is added to a dune. Once buried, these plants send a vertical rhizome to the surface, where a new shoot develops. Furthermore, some of these plants, such as American beachgrass, have specialized bacterium named *Azotobacter* associated with their roots, which fix atmospheric nitrogen into a form usable by the plant. These bacteria enable the plants to grow in an otherwise nitrogen-poor environment. Young roots secrete carbohydrates which may nourish the nitrogen-fixing bacteria. *See* NITROGEN FIXATION.

The plants in the dune slacks have morphological and physiological adaptations for growth in flooded areas. For example, the sedge American three-square contains large air spaces (aerenchyma) in its stems and roots, which provide an oxygen pathway from the shoots above the water to the oxygen-deprived roots in the flooded soil. When these plants are flooded, they increase their production of the hormone ethylene, which may stimulate the production of aerenchymatous tissue.

Management. Sand dunes form a natural defense against the sea; vegetation holds the sand in place and favors its accretion. Although dune vegetation is very tolerant of wind, windblown sand, and salt, it is sensitive to many human activities. Coastal areas are highly populated, especially in summer. The protection of these valuable ecosystems requires good management practices. Access to the beach through

Fig. 3. After one growing season, this planting of American beachgrass is already helping to stabilize this dune. Snow fence was used to trap and hold sand prior to planting.

the dunes is controlled, and damaged areas are restored with the help of porous fencing, such as snow fence, to trap and accumulate sand in eroded areas. Such fencing can increase the height and width of dunes. The new sand is planted with dune vegetation (**Fig. 3**) growing nearby or in commercial nurseries. These practices have been very successful.

Human disturbances include pedestrian and vehicular traffic, which can cause serious damage. In one typical area that was walked over less than 200 times in a year, more than half of the vegetation disappeared, leaving barren ground.

Off-road vehicular traffic is another serious threat. At the wrack line, it prevents the formation of new dunes by destroying organic matter and causing the loss of nutrients, water, and the seedlings themselves. Traffic behind the foredunes is equally destructive. Plants such as beach heather can be completely destroyed, needing years to recover. Tire tracks may remain visible for many seasons. Laws help but may be difficult to enforce. The complete abolition of beach vehicles may become necessary.

The protection of vegetation on the remaining dunes is important for several reasons. Dunes not stabilized by vegetation are highly susceptible to wind and water erosion and threatened with destruction. Their disappearance means a loss of nesting areas for many seabirds, other wildlife habitat, and valuable recreational areas. It would also leave settled areas, now protected by coastal dunes, open to erosion by wind and sea. *See* EROSION. Denise M. Seliskar

Bibliography. V. J. Chapman, *Coastal Vegetation*, 2d ed., 1976; P. J. Godfrey and M. M. Godfrey, Ecological effects of off-road vehicles on Cape Cod,*Oceanus*, 23:56–67, 1980/1981; R. R. Lewis III (ed.), *Creation and Restoration of Coastal Plant Communities*, 1982; D. S. Ranwell, *Ecology of Salt Marshes and Sand Dunes*, 1972; D. S. Ranwell and R. Boar, *Coast Dune Management Guide*, 1986; D. M. Seliskar, Waterlogging stress and ethylene production in the dune slack plant, *Scirpus americanus*, *J. Exper. Bot.*, 39:1639–1648, 1988; G. M. Silberhorn, *Common Plants of the Mid-Atlantic Coast*, 1982, revised, 1999; A. M. Wiedemann, *The Ecology of Pacific Northwest Coastal Sand Dunes: A Community Profile*, 1984.

Dunite

An ultramafic igneous rock composed of at least 90% olivine. Important accessory minerals (abundances usually less than 1%) found in different occurrences of dunite include chromian spinel, low- or high-calcium pyroxenes (enstatite and diopside), and plagioclase. If these minerals constitute greater than 10% of the rock, it is called chromitite, peridotite, or troctolite, respectively. Low-temperature alteration (less than 750°F or 400°C) causes hydration of olivine to the mineral serpentine and, where extensive, may transform dunite to the metamorphic rock serpentinite. Dunite was first described at Dun Mountain, New Zealand. The dun color is a characteristic feature of the weathered rock. *See* CHROMITE; OLIVINE; PERIDOTITE; PYROXENE; SERPENTINITE.

Composition. Dunite is ultrabasic in composition, meaning that it is low in silica compared to most crustal rocks (<45% silica) and is generally very high in magnesium (up to 54%). In some occurrences, notably in continental layered intrusions, however, the olivine may be very iron rich (variety hortonolite), and the rock can contain as much as 40% iron. Dunite is notably poor in alumina, soda, and lime; and while it may be relatively nickel rich, many other critical trace elements are nearly missing. Thus, weathering of dunite forms soils to which most terrestrial plants are poorly adapted, and such soils often host unusual plant species such as the carnivorous cobra lily and the miniature rhododendron *Kalmiopsis leachiana* found in the coast ranges of Oregon and California.

Texture and structure. Olivine crystals in some dunites can often be seen originally to have had equant shapes where the original crystal faces are in contact with intergranular plagioclase and pyroxene. Commonly, however, the olivine crystals have grown together during crystallization from a magma, or have undergone annealing recrystallization following high-temperature deformation (as commonly occurs during solid-state mantle flow) to form granular aggregates of irregularly shaped equidimensional crystals. *See* MAGMA.

Origin. In layered and southeast Alaskan-type intrusions, dunite is formed by fractional crystallization of basic magmas, whereby relatively dense olivine crystallizes and settles out of the melt, often with minor chromite, to form nearly pure pockets and layers. Dunite also forms during late upward transport by porous flow of melt in pipes through the nearly crystalline intrusions, whereby the original minerals are replaced by olivine because of reaction with the melt. Metasomatic reaction between late hydrous fluids and peridotites may also occur to produce some crosscutting dunites. Elsewhere, dunites may have formed by flow differentiation of olivine crystals in

basic dikes during melt transport along cracks and fissures in the Earth. *See* METASOMATISM; PLUTON.

In the mantle, dunite may form by precipitation of olivine from magmas migrating upward through the Earth, or by reaction of melts formed at high pressures with the mantle during upward transport along cracks or during focused porous flow where the melt is moving too fast to maintain simple equilibrium with the mantle. Thus, the occurrence of dunite is believed to mark the location of conduits where melt was transported through the Earth. *See* EARTH.

Occurrence. Dunite is found in continental associations as layers near the base of the large continental layered intrusions (fossil magma chambers crystallized deep within the crust), such as the Bushveld Complex of South Africa. It also occurs as crosscutting vertical pipes in these and in the concentric basic intrusions of the southeast Alaskan type found associated with the basement of volcanic arc terrains.

The most important occurrence of dunite is in its association with rocks believed to have come from the Earth's mantle (principally peridotite) where it marks the location of former conduits through which magmas have been transported out of the Earth to the crust. In the oceans, it is infrequently dredged from the great oceanic transform faults cutting the ocean ridges, exposed in tectonic windows in disrupted and uplifted ocean crust in association with serpentinized mantle peridotite. In 1993 and 1994, however, important dunite bodies crosscutting tectonically exposed mantle sections on the walls and floors of rift valleys have been found in the central Atlantic and eastern Pacific oceans far from transform faults. The suggestion is that the flow of magma out of the mantle upwelling beneath ocean ridges is not uniform, and is therefore important in controlling the formation and structure of the two-thirds of the Earth's crust that is formed in the oceans. *See* MID-OCEANIC RIDGE; RIFT VALLEY; TRANSFORM FAULT.

Dunite is also found as lenses, pods, and veins in alpine-type orogenic belts, where it is found crosscutting subcontinental mantle peridotites, in xenoliths of broken rock carried up to the surface by magmas erupting from the mantle, and in mantle rocks underlying ophiolite complexes. The latter are fossil sections of ocean crust and mantle, tectonically emplaced on-land during plate-tectonic collisions along continental margins and island arcs, and believed to have originally formed in island arcs and marginal seas. *See* CONTINENTAL MARGIN; EARTH CRUST; OPHIOLITE; PLATE TECTONICS; XENOLITH.

Uses. Dunite has been extensively quarried as a building stone because of its unusual pale water-green color, but is little used now because of its high susceptibility to acid rain. Dunite is also used for refractory materials for furnaces. Olivine in dunite is occasionally of gem quality, as on Zabargad Island in the Red Sea. Dunite is also the principal host of deposits of the mineral chromite. The hortonolite-dunite pipes found in the Bushveld Complex of South Africa have also been mined for platinum-group minerals. *See* GEM; IGNEOUS ROCKS; PETROLOGY; REFRACTORY.

Henry J. B. Dick

Bibliography. R. G. Coleman, *Ophiolites*, 1977; P. B. Kelemen et al., Reaction between ultramafic rock and fractionating basaltic magma, II: Experimental investigation of reaction between olivine tholeiite and harzburgite at 1150–1050°C and 5 kb, *J. Petrol.*, 31:99–134, 1990; M. J. O'Hara, Are oceanic basalts primary magma?, *Nature*, 220:683–686, 1968; L. R. Wager and G. M. Brown, *Layered Igneous Rocks*, 1967.

Dust and mist collection

The physical separation and removal of particles, either solid or liquid, from a gas in which they are suspended. Such separation is required for one or more of the following purposes: (1) to collect a product which has been processed or handled in gas suspension, as in spray-drying or pneumatic conveying; (2) to recover a valuable product inadvertently mixed with processing gases, as in kiln or smelter exhausts; (3) to eliminate a nuisance, as in fly-ash removal; (4) to reduce equipment maintenance, as in engine intake air filters; (5) to eliminate a health, fire, explosion, or safety hazard, as in bagging operations or nuclear separations plant ventilation air; and (6) to improve product quality, as in cleaning of air used in processing pharmaceutical or photographic products. Achievement of these objectives involves primarily gas-handling equipment, but the design must be concerned with the properties and relative amounts of the suspended particles as well as with those of the gas being handled.

All particle collection systems depend upon subjecting the suspended particles to some force which will drive them mechanically to a collecting surface. The known mechanisms by which such deposition can occur may be classed as gravitational, inertial, physical or barrier, electrostatic, molecular or diffusional, and thermal or radiant. There are also mechanisms which can be used to modify the properties of the particles or the gas to increase the effectiveness of the deposition mechanisms. For example, the effective size of particles may be increased by condensing water vapor upon them or by flocculating particles through the action of a sonic vibration. Usually, larger particles simplify the control problem. To function successfully, any collection device must have an adequate means for continuously or periodically removing collected material from the equipment.

Devices for control of particulate material may be considered, by structural or application similarities, in eight categories, which follow.

Gravity settling chamber. In this, the simplest type of device but not necessarily the least expensive, the velocity of the gas is reduced to permit particles to settle out under the action of gravity. Normally, settling chambers are useful for removing particles larger than 50 micrometers in diameter, although

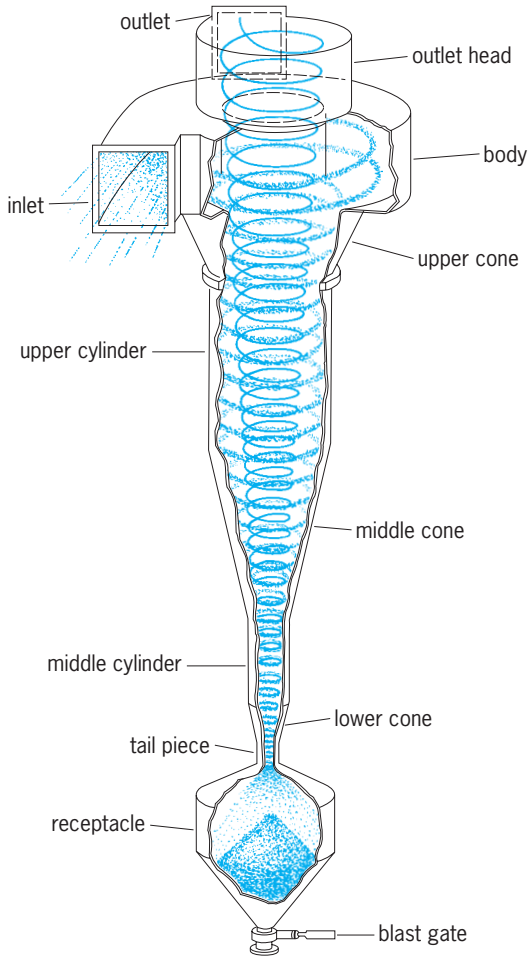

Fig. 1. Cyclone dust separator, an inertial device. (*American Standard Industrial Division*)

with special configurations they may be used to remove particles as small as 10 μm.

Inertial device. The basis of this type of device is that the particles have greater inertia than the gas. The cyclone separator, typical of this type of equipment, is one of the most widely used and least expensive types of dust collector. In a cyclone, the gas usually enters a conical or cylindrical chamber tangentially and leaves axially. Because of the change of direction, the particles are flung to the outer wall, from which they slide into a receiving bin or into a conveyor, while the gases whirl around to the central exit port (**Fig. 1**). A large variety of configurations is available. For large air-handling capacities, an arrangement of multiple small-diameter units in parallel is often used to attain high collection efficiencies and to permit lower headroom requirements than a single unit would.

Mechanical inertial units are similar to cyclones except that the rotational motion of the gas is induced by the action of a rotating member. Some such units are designed to act as fans in addition to their dust-collecting function. There are also a wide variety of other units; many are called impingement separators. Most separators that are used to remove entrained liquids from steam or compressed air fall into this category.

Packed bed. A particle-laden gas stream may be cleaned by passing it through a bed or layer of packing composed of granular materials such as sand, coke, gravel, and ceramic rings, or fibrous materials such as glass wool, steel wool, and textile staples. Depending on the application, the bed depth may range from a fraction of an inch to several feet. Coarse packings, which are used at relatively high throughput rates (1–15 ft/s or 0.3–4.5 m/s superficial velocity) to remove large particles, rely primarily on the inertial mechanism for their separating action. Fine packings, operated at lower throughput rates (1–5 ft/min or 0.3–15 m/min superficial velocity) to remove relatively small suspended particles, usually depend on a variety of deposition mechanisms for their separating effect. Packed beds, because of a gradual plugging caused by particle accumulation, are usually limited in use to collecting particles present in the gas at low concentration, unless some provision is made for removing the dust—for example, by periodic or continuous withdrawal of part of the packing for cleaning. Depending on the application and design, the collection efficiencies of packed beds range widely (50–99.999%).

Cloth collector. In such a collector, also known as a bag filter, the dust-laden gas is passed through a woven or felted fabric upon which the gradual deposition of dust forms a precoat, which then serves as a filter for the subsequent dust. These units are analogous to those used in liquid filtration and represent a special type of packed bed. Because the dust accumulates continuously, the resistance to gas flow gradually increases. The cloth must, therefore, be periodically vibrated or flexed, or back-flushed with a stream or pulse of air to dislodge accumulated dust (**Fig. 2**). A wide variety of filter media is

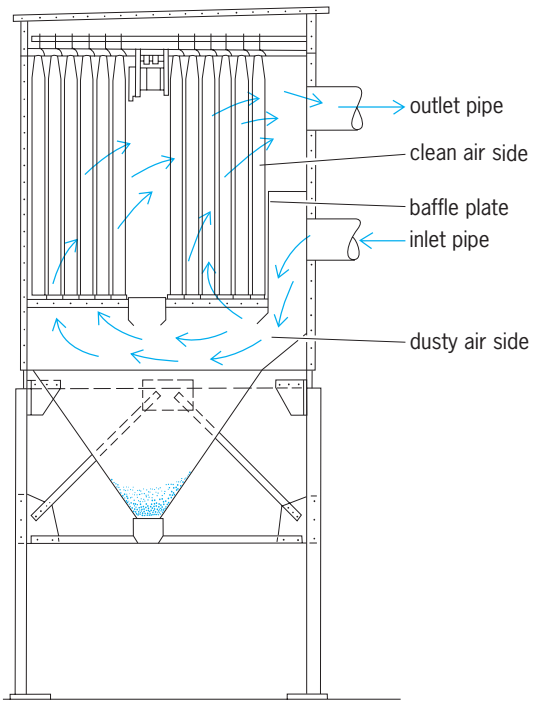

Fig. 2. Cloth collector. (*Wheelabrator Corp.*)

available. Cotton or wool sateen or felts are usually used for temperatures below 212°F (100°C). Some of the synthetic fibers may be used at temperatures up to 500°F (260°C). Glass and asbestos or combinations thereof have been used for temperatures up to 650°F (343°C). For special high-temperature applications, metallic screens and porous ceramics or stainless steel have been employed. Collection efficiencies of over 98% are attained readily with cloth collectors, even with very fine dusts.

Scrubber. A scrubber uses a liquid, usually water, to assist in the particulate collection process. An extremely wide variety of equipment is available, ranging from simple modifications of corresponding dry units to permit liquid addition, to devices specifically designed for wet operation only (**Fig. 3**). When properly designed for a given application, scrubbers can give very high collection efficiency, although the mere addition of water to a gas stream is not necessarily effective. For a given application, collection efficiency is primarily a function of the amount of power supplied to the gas stream, in the form of gas pressure drop, water pressure, or mechanical energy. With scrubbers it is important that proper attention be given to liquid entrainment separation in order to avoid a spray nuisance. Consideration must also be given to the liquid-sludge disposal problem.

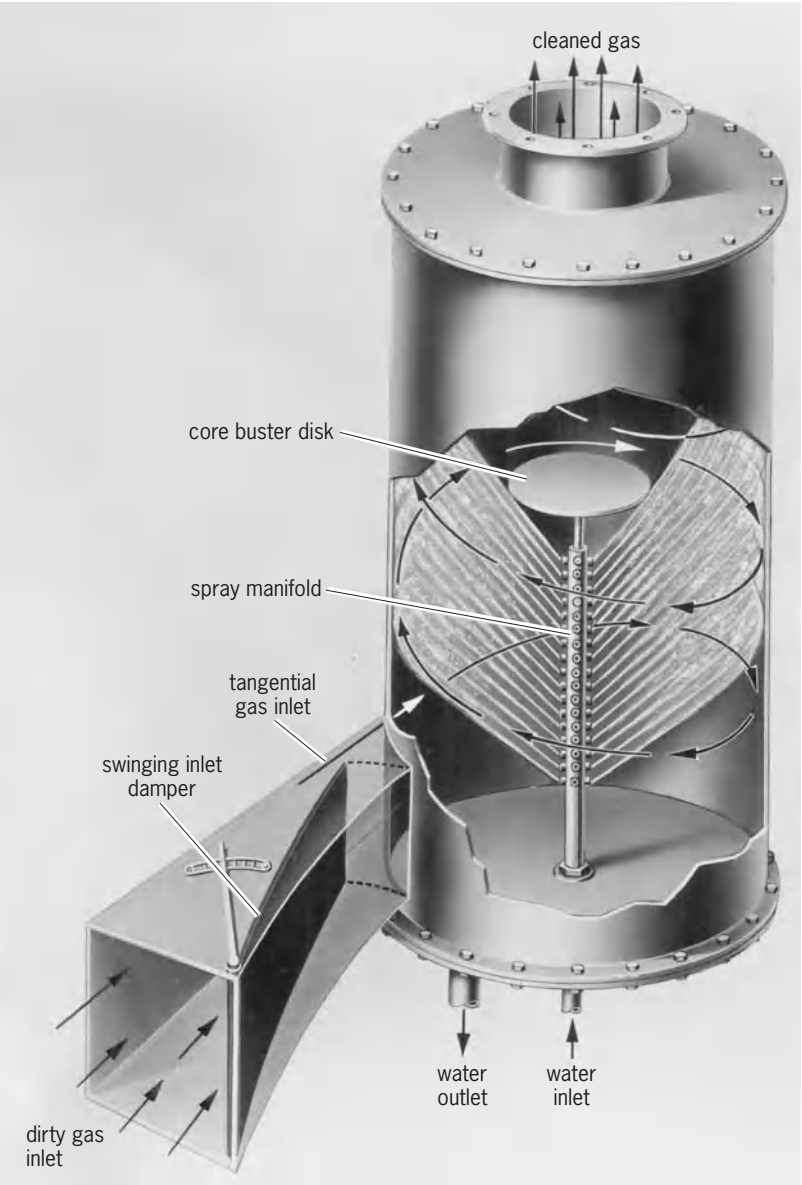

Fig. 3. Cyclonic liquid scrubber. (*Chemical Construction Corp.*)

Electrostatic precipitator. Particles may be charged electrically by a corona discharge and caused to migrate to a collecting surface. The single-stage unit, which is commonly known as a Cottrell precipitator, and in which the charging and collecting proceed simultaneously, is the type generally used for industrial or process applications. These units normally employ direct current at voltages ranging from 30,000 to 100,000 V. The two-stage unit, in which charging and collection are carried out successively, is commonly used for air conditioning applications. These units also employ direct current, ranging from 5000 to 13,000 V, and involve close internal clearances (0.25–0.5 in. or 6–12 mm). Electrical precipitators are capable of high collection efficiency of fine particles.

In dry precipitators, collected dust is dislodged by intermittent rapping or vibration of the collection plates with possible dust reentrainment. Another problem arises from high dust resistivity, which causes collected dust to be reentrained by back-ionization because the dust particles are not discharged electrically upon collection. The reentrainment of collected material as flocs in the exhaust gas, a phenomenon known as snowing, must be avoided to prevent a possible accentuated nuisance or vegetation damage problem.

In "wet" electrostatic precipitators, a film of water is allowed to flow over the collecting surface. This serves to minimize problems of dust reentrainment either by mechanical means or by the process of back-ionization. It also provides a means of removing collected dust. *See* ELECTROSTATIC PRECIPITATOR.

Air filter. This is a unit used to eliminate very small quantities of dust from large quantities of air, as in air conditioning applications. Although units in this class actually fall into one of the previous classes, they are given a special category because of wide usage and common special features. In this category are viscous-coated fiber-mat filters and dry filters. These are actually a form of packed bed and are frequently known as unit filters. The domestic furnace filter is an example of a viscous-coated unit filter. Automatic filters provided with continuous and automatic cleaning arrangements are available in both the viscous-coated and dry forms, as well as with electrostatic provisions. *See* AIR FILTER.

Miscellaneous equipment. Electrostatic fields have been used with packed beds, scrubbers, and air filters to improve collection efficiency, and with cloth collectors to dislodge dust from the cloth. Acoustic or sonic vibrations imparted to a gas stream cause particulates to collide and flocculate, forming larger particles that are more readily collected in conventional apparatus. This principle has been employed,

but has had extremely limited application because of economic and other practical considerations. In thermal precipitation, suspended particles are caused to migrate toward a cold surface or away from a heated surface by the action of a temperature gradient in the gas stream. This principle has found extensive use in atmospheric sampling work. *See* AIR POLLUTION; CHEMICAL SEPARATION TECHNIQUES; MECHANICAL SEPARATION TECHNIQUES; SMOKE; UNIT OPERATIONS.

Charles E. Lapple

Bibliography. R. H. Perry and D. Green (eds.), *Perry's Chemical Engineer's Handbook*, 7th ed., 1997; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., 1999; F. Loeffler, *Dust Collection with Bag Filters and Envelope Filters*, 1988; W. Strauss and S. J. Mainwaring, *Air Pollution*, 1984; L. Theodore and A. J. Buonicore, *Air Pollution Control Equipment*, 2 vols., 1988.

Dust storm

A strong, turbulent wind bearing large clouds of dust. The suspended dust consists predominantly of mineral particles with diameters less than 100 micrometers. Sand storms differ from dust storms in that the particles are much larger and have high settling speeds. When the wind and turbulence weaken, the larger particles settle out rapidly, leaving behind a cloud of dust with diameters mainly under 20 μm. Wind systems can carry these clouds many thousands of miles, in some cases across oceans.

Sources of dust. Dust storms can occur where the soils are loosely consolidated and poorly protected by vegetation. Soil moisture is critically important, since water binds the soil particles together and supports the growth of plants. Consequently, storms occur most frequently in arid and semiarid regions (**Fig. 1**), with the maximum frequency in regions having an annual rainfall of 4–8 $\times$ 10^{-3} in. (0.1–0.2 mm). Hyperarid areas such as deep deserts are not particularly good sources, because the surface is often covered with rock and sand, neither of which contains an abundance of fine particles.

Marginal agricultural lands are especially susceptible to storms in the spring when the soils, loosened by plowing and overgrazing, are easily blown away. Extended periods of drought have the same effect, for example, the periodic dust bowl conditions in the Great Plains of the United States. *See* DROUGHT.

Most of the global dust storm activity is concentrated in a belt of arid lands that extends from the west coast of North Africa, through the Middle East and across Central Asia, almost to the Pacific coast. Many locations in this region experience dust storms 30 to 60 days a year.

Soil factors. In typical soils, dust generation begins when winds become strong enough to dislodge particles with diameters of several millimeters and larger (**Fig. 2**). These particles roll across the surface and bounce into the air, where they pick up speed from the wind. When they fall back to the surface, they strike other particles, blasting them into the air. This process is known as saltation. These new particles continue the process, producing an avalanche effect that rapidly sets the entire surface of the soil into motion. Particles larger than about 0.2 in. (5 mm) move in a shallow layer within a few feet of the ground. Smaller particles will be blasted into the air with the large particles; because of their smaller settling velocity, they will not fall back to the surface so rapidly. In this manner, a cloud of dust particles rapidly fills the turbulent atmospheric boundary layer. The depth of this boundary layer can vary greatly, depending on meteorological conditions.

Dust generation can be impeded in a number of ways. Vegetation reduces the force of the wind on the surface, and roots tend to hold the soil together. Soil cohesiveness is increased if the soil surface is minimally disturbed. Improper tilling and overgrazing

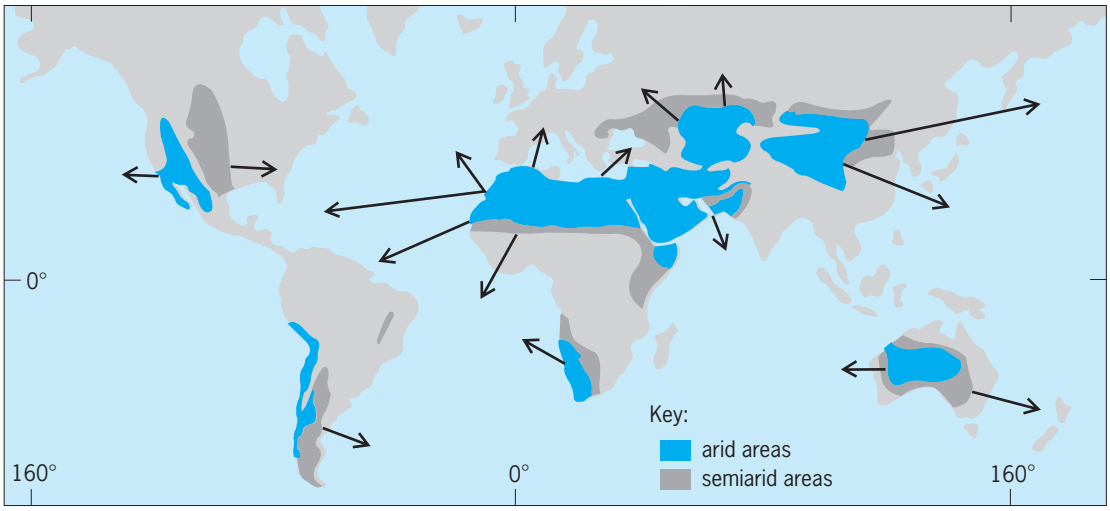

Fig. 1. Distribution of areas with high dust storm activity. The arrows indicate the major trajectories that dust storms follow. (*After K. Pye, Aeolian Dust and Dust Deposits, Academic Press, 1987*)

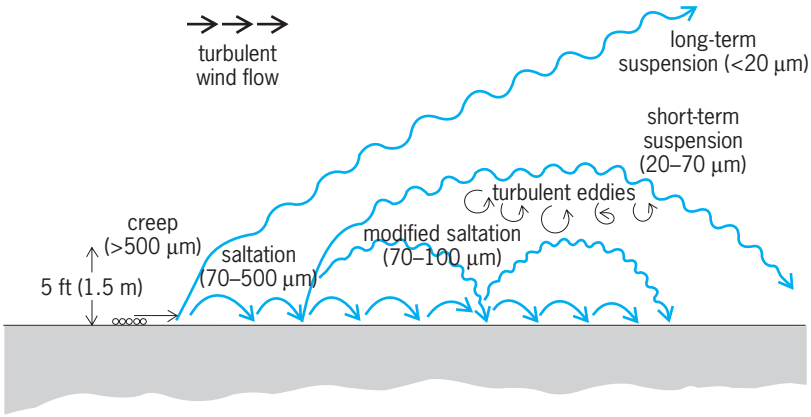

Fig. 2. Dust generation and transport near the soil surface. The bouncing dislodges both large particles, which continue the saltation process, and small particles, which can go into suspension. (*After K. Pye, Aeolian Dust and Dust Deposits, Academic Press, 1987*)

are prime factors in wind erosion. *See* AGRICULTURAL SOIL AND CROP PRACTICES; COVER CROPS; EROSION.

Meteorological factors. The wind speed required to generate dust varies widely according to soil characteristics. Typical values are 9–29 mi/h (4–13 m/s) for poorly sorted river sediments; 11–36 mi/h (5–16 m/s) for various desert environments including the deep Sahara; and 25–36 mi/h (11–16 m/s) for areas in Colorado and Arizona.

The area covered by the dust cloud and its height will depend on the meteorological conditions that generate the winds. Dust devils normally have a diameter of a few meters at most, and the height of the vortex is usually in the range of 10–300 ft (3–100 m). Although dust devils are small, they are commonplace, and in many areas they are probably responsible for much of the low-level dust generated under normal wind conditions.

In some areas, dust is commonly generated by cold downdrafts from cumulonimbus clouds. Known as haboobs, these generally last for 30 min to an hour, and they typically carry dust to heights of several hundred yards, at times to several thousand yards. The dust-generating area is confined to the region of the cloud and usually covers an area of tens to a few hundred square miles.

Major dust storms are associated with large-scale atmospheric features. The close positioning of high- and low-pressure centers produces a tight pressure gradient and strong winds. This situation is responsible for many dust events in the Middle East and for the Santa Ana in the western United States. Dust generated under these conditions is usually confined to the lower mile of the atmosphere. *See* AIR PRESSURE; WIND.

Cold fronts are responsible for truly large dust events that can be carried great distances. The vigorous winds associated with fronts are effective in generating dust over large areas. Moreover, the stronger atmospheric lifting ahead of the front can carry the dust to great altitudes, 10,000–23,000 ft (3–7 km). The high winds found at these altitudes can transport the dust thousands of miles in a few days. In China, storms of this type produce heavy dust over thousands of square miles and subsequently carry the dust over large areas of the North Pacific Ocean. The movement of these large dust clouds can be readily observed in satellite imagery. *See* AIR MASS; ATMOSPHERE; FRONT; SATELLITE METEOROLOGY.

Effects of dust. Dust plays a further role in weather in that mineral particles serve as condensation and freezing nuclei which are essential for the formation of clouds. Dust clouds typically contain particle concentrations of 10^{-7} to 10^{-4} oz/ft^3 (10^2 to 10^5 micrograms/m^3). These amounts can severely reduce visibility, causing problems for air, ship, and ground traffic over large areas. *See* CLOUD PHYSICS.

Dust can alter the distribution of solar energy in the Earth–atmosphere system, reducing the amount reaching the Earth's surface and producing heating in the troposphere. These effects can alter the dynamics of the atmosphere. Large amounts of dust as hypothesized in the nuclear winter scenarios would dramatically alter climate. *See* ATMOSPHERIC GENERAL CIRCULATION; CLIMATE MODIFICATION.

The impact of windblown dust on surfaces can produce large electrostatic charges that can affect the operation of electronic equipment. The inhalation of large quantities of dust can result in a variety of lung damage, the degree of damage depending on the composition of the dust (for example, asbestos). Organisms associated with the soil particles can also cause diseases, some of them quite serious.

J. M. Prospero

Bibliography. R. A. Bagnold, *The Physics of Blown Sand and Desert Dunes*, 1941; O. E. Barndorff-Nielsen and B. B. Willets, *Aeolian Grain Transport*, vol. 1: *Mechanics*, vol. 2: *The Erosional Environment*, 1991; T. L. Pewe, *Desert Dust: Origin, Characteristics, and Effect on Man*, Geol. Soc. Amer. Spec. Pap. 186, 1981; K. Pye (ed.), *The Dynamics and Environmental Context of Aeolian Sedimentary Systems*, 1994.

Dwarf star

A star whose state of evolution resembles that of the Sun. The term "dwarf star" derives from the early work of Ejnar Hertzsprung and Henry Norris Russell, who distinguished two kinds of stars, large ones called giants (or supergiants) and smaller ones called dwarfs. "Dwarf" is synonymous with "main sequence star" [Morgan-Keenan-Kellman (MKK) luminosity class V, where the luminosity classes are supergiants = I, bright giants = II, giants = III, subgiants = IV, and dwarfs = V], and implies not size but evolutionary condition. Dwarfs are stars that, like the Sun, fuse hydrogen into helium in their cores, the thermonuclear reactions providing energy and support. *See* GIANT STAR; HERTZSPRUNG-RUSSELL DIAGRAM.

Exact characteristics depend on chemical composition. Dwarfs of roughly solar composition (those in the galactic disk) range over the entire spectral sequence, while low-metal stars (the "subdwarfs" of the Galactic halo) occupy only the lower

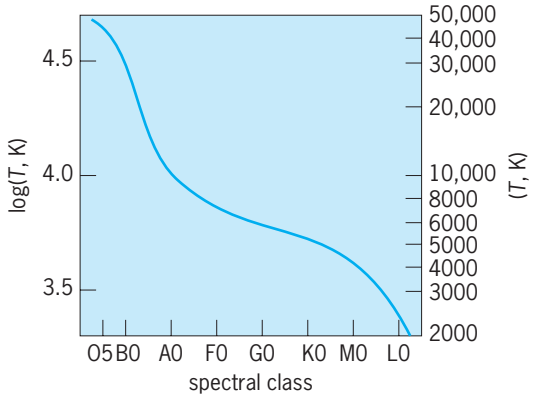

Absolute temperatures of dwarf (main sequence) stars plotted against their spectral class. The absolute temperature *T*, measured in kelvins (K), is plotted on a logarithmic scale. Temperatures drop, following a sinuous curve, from a high of about 48,000 K (85,940°F) at the hot end to 2000 K (3140°F) at the low end. (*After J. B. Kaler, Cambridge Encyclopedia of Stars, Cambridge University Press, 2005*)

end of the main sequence from spectral class G on down (low abundances of metals, defined by astronomers as all elements more massive than helium, rendering the classes problematic). At the cool limit, which lies near the warm end of class L (about L2), dwarfs of solar composition have effective temperatures around 2000 K (3140°F), absolute visual magnitudes of +21, and bolometric luminosities somewhat under 10^{-4} that of the solar luminosity. At the hot limit (now extended to class O2 from the previous O3 limit), the values are, respectively, the order of 48,000 K (85,940°F), −6, and over 10^6 solar. The Sun, a G2 dwarf (that is, G2 V) falls in the middle at 5780 K (9945°F) with an absolute visual magnitude of +4.83 (see **illustration**). These properties are produced by a mass range from just under 0.08 solar masses for warm L dwarfs, below which full hydrogen fusion cannot be turned on (and the stars become brown dwarfs), to around 120 solar masses at classes O2–O3. On average, dwarf luminosities increase approximately as the 3.5 power of the mass, though the exponent varies from 2.7 at low mass, to 4 for solar mass stars, and back to 3 for stars of higher mass. The coolest dwarfs are small, with diameters at minimum under 0.1 solar, while O dwarfs can be quite large, over 20 times solar. The numbers of objects in the different spectral classes drop dramatically as one moves upward along the main sequence. From M through O, over 70% fall into class M, whereas 10% are in class G, and a mere 0.00004% are in class O. *See* BROWN DWARF; MAGNITUDE (ASTRONOMY); SUN.

Lower-mass dwarfs, below about 1.25 solar masses and spectral class F5, get their energy dominantly from the proton-proton chain, have low rotation speeds, and have convective envelopes with radiative cores (becoming entirely convective at the lowest masses). Higher-mass dwarfs get their energy dominantly from the carbon cycle, can be fast rotators, and have radiative envelopes and convective cores. At the high-mass end, above about 10 solar masses, the dwarf stage lasts 2–20 million years, and O stars turn into supergiants and core-collapse type II supernovae. At roughly class G8 and a mass of 0.85 solar, the lifetime equals the age of the Milky Way Galaxy; dwarfs between this limit and 10 solar masses become white dwarfs. *See* CARBON-NITROGEN-OXYGEN CYCLES; PROTON-PROTON CHAIN; SPECTRAL TYPE; STAR; STELLAR EVOLUTION; STELLAR ROTATION; SUPERGIANT STAR; SUPERNOVA; WHITE DWARF STAR. James B. Kaler

Dwarfism and gigantism

Underdevelopment and overdevelopment of the skeleton, respectively. Skeletal growth is a complex process and can be disturbed in many ways. For example, overstimulation by excessive growth-hormone production during childhood can produce gigantism. This is usually due to a pituitary tumor. Insufficient stimulation of skeletal growth, resulting from hormonal, metabolic, or nutritional disturbances, leads to reduced height with normal body proportions. The shortness of stature depends on the degree of the disturbance; the designation proportionate dwarfism is used in severe cases. Proportionate dwarfism may or may not be genetic. *See* PITUITARY GLAND.

Commonly, the term dwarfism is applied to short individuals with abnormal body proportions resulting from disturbances of bone growth itself. The bones of either limbs or trunk can be most affected. The disorders, called chondrodysplasias, number well over 100 and show a wide variation in clinical features: some appear at birth, others become evident in later childhood. The severity ranges from conditions in which infants die at birth to conditions compatible with normal life. In most instances, however, the abnormal skeleton creates medical problems. Joint deformities and osteoarthritis are common. Sometimes the skeletal abnormalities affect other organ systems, as in compression of the spinal cord at the base of the skull or in the lower back. Short ribs may reduce the size of the thoracic cavity, producing breathing difficulties. Other organ systems may also be affected, as in dysfunction of the kidneys or immune system. Intelligence is usually average. The specific condition is diagnosed on the basis of clinical features and skeletal x-rays.

Almost all of the chondrodsyplasias are inherited as mendelian or single-gene traits. However, because of reduced fertility and a high incidence of new mutations for certain traits, most individuals with these disorders are born to parents of average stature. Although it is known that the gene mutations disrupt endochondral ossification, which is responsible for skeletal development and linear bone growth, the mechanisms by which this occurs are poorly understood. *See* HUMAN GENETICS; SKELETAL SYSTEM DISORDERS. William A. Horton

Bibliography. R. Dee, E. Mango, and L. C. Hurst, *Principles of Orthopaedic Practice*, 2d ed., 1997; V. C. Kelley (ed.), *Practice of Pediatrics*, 1984.

Dyadic

A mathematical abstraction corresponding to an expression of the type $\boldsymbol{\beta\gamma} + \boldsymbol{\delta\epsilon} + \cdots$, in which the elements (dyad symbols) consist of two vector symbols in juxtaposition without the intervention of either the dot ($\cdot$) or cross ($\times$). Essentially a dyad is an ordered pair of vectors subject to certain rules of operation. The first symbolic factor in a dyad (β in $\beta\gamma$, for example) is called the antecedent and the second the consequent.

The concept dyadic (J. W. Gibbs) is inherent in the formal structure of certain vectorial expressions such as $(\boldsymbol{\alpha} \cdot \mathrm{i})\mathrm{i} + (\boldsymbol{\alpha} \cdot \mathrm{j})\mathrm{j} + (\boldsymbol{\alpha} \cdot \mathfrak{k})\mathfrak{k}$ and $(\boldsymbol{\alpha} \cdot \boldsymbol{\beta})\boldsymbol{\gamma} + (\boldsymbol{\alpha} \cdot \boldsymbol{\delta})\boldsymbol{\epsilon}$. To reduce $(\boldsymbol{\alpha} \cdot \boldsymbol{\beta})\boldsymbol{\gamma} + (\boldsymbol{\alpha} \cdot \boldsymbol{\delta})\boldsymbol{\epsilon}$ to the form $\boldsymbol{\alpha} \cdot (\boldsymbol{\beta\gamma} + \boldsymbol{\delta\epsilon})$ it is sufficient to accept $\boldsymbol{\beta\gamma} + \boldsymbol{\delta\epsilon}$ as a mathematical entity and to agree that, whenever $\boldsymbol{\alpha} \cdot$ is applied from the left to $\boldsymbol{\beta\gamma} + \boldsymbol{\delta}\epsilon$, it is to be distributed to the antecedents, while the symbol $+$ is to become the plus of vector addition. Similarly, $(\boldsymbol{\beta\gamma} + \boldsymbol{\delta\epsilon}) \cdot \boldsymbol{\alpha} = \boldsymbol{\beta}(\boldsymbol{\gamma} \cdot \boldsymbol{\alpha}) + \boldsymbol{\delta}(\boldsymbol{\epsilon} \cdot \boldsymbol{\alpha})$ by distribution to the consequents.

Two dyadics $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ are by definition equal if and only if $\boldsymbol{\Phi} \cdot \boldsymbol{\alpha} = \boldsymbol{\Psi} \cdot \boldsymbol{\alpha}$ identically in $\boldsymbol{\alpha}$. In particular, if $\boldsymbol{\beta} \neq 0$, $\boldsymbol{\gamma} \neq 0$, and $\boldsymbol{\beta\gamma} = \boldsymbol{\gamma\beta}$, then $\boldsymbol{\beta}(\boldsymbol{\gamma} \cdot \boldsymbol{\alpha}) = \boldsymbol{\gamma}(\boldsymbol{\beta} \cdot \boldsymbol{\alpha})$ and $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are parallel.

If the $\mathfrak{r}_a$ are the base vectors $\partial\mathfrak{r}/\partial x^a$ of a space defined by $\mathfrak{r} = \mathfrak{r}(x^1, \ldots, x^n)$ and $\mathfrak{r}^b = g^{bc}\mathfrak{r}_c$ and the quantities T^a_b are the components of a tensor of the type $(1, 1, w)$, then $T_b{}^a\mathfrak{r}_a\mathfrak{r}^b$ is an invariant. Consequently, tensors can be related to dyadics. *See* TENSOR ANALYSIS.

Homer V. Craig

Dye

A colored substance which imparts more or less permanent color to other materials. *See* DYEING.

Not all colored substances are dyes, however. If red iron rust is ground with white sugar, the resulting mixture has an overall reddish appearance. The sugar has been colored by pigmentation, and the iron rust has been used as a pigment. Examination of the mixture under a microscope shows distinct white and red particles, and a separation can be made by dissolving the sugar out of the mixture with water. If the red iron rust is added to white cloth in water, some of the red particles may cling to the cloth, but they can be removed by rubbing or by washing with soap. Some colored substances (usually organic chemical compounds) may be added to cloth in water, and after a period of soaking, usually accompanied by heat and agitation, the cloth will be colored; no separate colored particles can be seen under ordinary microscopic examination, and the color cannot be removed by washing, even with soap. The cloth has been dyed, and the colored substance is a dye (also called a dyestuff).

Customarily, colored water-insoluble substances are called pigments. Dyes are generally water-soluble, although some are soluble only during application, after which they become insoluble. *See* PIGMENT (MATERIAL).

The mechanism by which soluble colored substances enter the internal structure of fibers and there become fixed has been variously explained in terms of the physical and chemical concepts of the times when the explanations were given. It is said to be an adsorption phenomenon, a salt formation, a quasi-chemical union caused by hydrogen bonding, or an ether linkage, and in some cases it is considered to be a true solution effect. The end result, however, is that the dye has imparted a color (not necessarily that of the solid dye itself) to the fiber which is more or less resistant to washing or removal by similar mechanical operations. The dye is said to be fixed on and to have affinity, or substantivity, for the material it has colored. The material is designated as the substrate. If the color is quite resistant to washing and light, it is called a fast color; if the color is easily removed or fades quickly, it is a fugitive dye.

Because not all water-soluble colored substances are dyes, various attempts have been made to relate chemical constitution with color and substantivity. One of the earlier and still very useful explanations was given by O. N. Witt, who stated in 1876 that all colored organic compounds (called chromogens) contain certain unsaturated chromophoric groups which are responsible for the color, and if these compounds also contain certain auxochromic groups, they possess dyeing properties. Examples of chromophores are the groups —NO_2, —N═N—, and ═CO, and of auxochromes, —NH_2, —OH. The auxochromes also influence hue according to their nature, number, and position on the chromogen molecule. An elaboration of this theory in 1888 by H. E. Armstrong regarded all chromophores as being quinoid (═R═) in structure. Many later studies have added explanations of the nature and variation of color in organic compounds, but the Witt theory provides a frame of reference for most dyes which is simple, practical, and satisfactory for all but the specialist in dye chemistry.

Dyestuffs may be classified in various ways, according to color (blue, red, and so on), origin (natural—from vegetable and animal matter—or synthetic), chemical structure, kinds of material to which they are applied, and method of application.

Color or hue. Commercial dyes are usually named to indicate the hue imparted to the dyed article. This color is not necessarily the same as that of the solid dyestuff, nor is it the same as the color of the dye solution from which the dyeing is made.

Origin. The dyes used earliest were natural dyestuffs (saffron, henna, cochineal, and logwood) derived from plants or animals. Most natural dyes are of the mordant type that requires a fixing agent. With the advent of the synthetic dyes, which are far more varied in shades and properties, and generally more brilliant, the production of the natural colors has decreased to a very small portion of the total market. Their greatest use is in the dyeing of leather.

Synthetic dyes were an early result of the development of chemistry as a science, and the first commercial production in 1857 in England of a synthetic organic chemical was that of the first commercially produced synthetic dye, mauve, discovered in 1856 by William Henry Perkin. The early history of organic

chemistry was primarily a record of investigations of material and methods for making synthetic dyes. Because one of the early and important materials used in these syntheses was aniline (derived from coal tar), the synthetic dyes have been known as coal-tar dyes and aniline dyes.

The manufacture of dyes proceeds from simple raw materials, mostly aromatic hydrocarbons such as benzene (benzol), toluene, and naphthalene, with introduction of other chemical groups, such as nitro, amino, halogen, and sulfonic acid. These intermediate compounds are then further processed by many special operations in organic chemistry such as diazotization, coupling, condensation, and fusion to give the final dyestuff. Following the chemical manufacture, the dye may be treated to give it special physical properties, and is standardized for strength. It may be sold as a dry powder, as a paste, or in solution.

Of the many thousands of dyes which have been synthesized in laboratories, about 3500 have had actual commercial use. These have been indexed in several compilations. The Society of Dyers and Colourists (England) and the American Association of Textile Chemists and Colorists give each dye an individual five-digit number as identification regardless of the specific manufacturer. Although some of the earliest dyes have common names with general usage, commercial dyes are sold under a large number of names, many of which relate to the same material.

Chemical structure. The most precise and scientific classification of dyes is based upon their chemical structure. This is the classification of interest to the research chemist and the manufacturer of dyes.

Utilization. This classification is based on the materials to which the dyes are applied.

Cloth. Coloring cotton, wool, linen, and silk is by far the largest use for dyes.

Manufactured fibers include regenerated cellulose (viscose), cellulose acetate, nylon, polyacrylonitrile, and polyester. Regenerated cellulose fiber can be dyed with dyes for cotton; dyeing other manufactured fibers with the colors commonly used for natural fibers is often less satisfactory or impossible. Special classes of dyestuffs were created for these polymeric materials. *See* TEXTILE.

Paper. This is colored both by dyeing and by pigmentation with finely divided colored materials. The color is usually added to the raw stock in the beater before the sheets are formed, or on the calender. Dyes are also applied in the coatings used on paper, such as the wax-dye coating used for duplicating (carbon) paper. *See* PAPER.

Leather. One of the earliest materials to be colored, leather has retained the use of natural dyes to a greater extent than most other materials, but there are many manufactured textile and specific dyes for leather. Because of the difficulty of penetrating the closely knit structure of leather, surface coloring by spraying or brushing is widely practiced. *See* LEATHER AND FUR PROCESSING.

Wood. Dyeing or tinting of wood is done with dyes in water, alcohol, or other solvents which evaporate after the wood has been painted with or soaked in the dye solution.

Pigments. Many soluble dyes are converted into pigments by forming insoluble salts (that is, replacing sodium in a dye salt with calcium) for use in lacquers, paints, and printing inks. This insoluble salt, called a toner, may also be deposited on inorganic fillers, such as aluminum hydroxide, to form a lake.

Food. The appearance of many foods is enhanced by artificial coloring. Butter and margarine are colored yellow, fruits and sauces which are often dulled by the process of preserving are brightened by addition of dye, and large amounts of dyestuffs are used in soft drinks and candies. In most countries, only certain colors are permitted to be used in this way.

In the United States, food must be colored by a natural dye or by a synthetic chosen from a list of certified food colors, of which the harmlessness to the human system is under constant check and review by a federal agency. Elaborate studies of the toxicity and pharmacology of any new color must be provided by the manufacturer before it is admitted to the list. *See* FOOD MANUFACTURING.

Oils. Lubricating oils and gasolines are colored for improving in appearance; for identifying grade or type of use; or as a trademark of the manufacturer. Waxes, shoe polishes, and candles are also colored.

Plastics and rubbers. Coloration of resins, plastics, and elastomers may be done by solution of a dyestuff or by dispersion of a pigment in the substrate. In this coloring operation, there is no need for substantivity, and choice of color is determined solely by the solubility, shade, and fastness properties desired.

Biological samples. The dyeing or staining of tissues of animal and vegetable matter is an important technique in physiological and medical studies. Microorganisms and cell structure are made more visible and differentiated under the microscope by selective staining with dyes of varying affinity for protoplasm.

Photography. Certain dyes (cyanines, of the polymethine class) are added to photographic emulsions to increase sensitivity to light in special regions of the spectrum. Color photographs are produced by formation of dyes from their components within the emulsion layer. *See* PHOTOGRAPHY.

Indicators. Some property of the environment may bring about a change in the hue in certain dyes, thus giving information to the observer about this property. Indicators usually respond to acidity, alkalinity, and oxidation or reduction, but may respond to heat, humidity, electricity, and water hardness. *See* ACID-BASE INDICATOR.

Miscellaneous. Soap, synthetic detergents, cosmetics, ink, hair, fur, metals, anodized aluminum, and many other materials are colored with dyes of various types. Dyes are also used to produce colored smokes, particularly for military identification. *See* INK; SMOKE.

Methods of application. This classification is used most frequently by the practical dyer.

Acid dyes. These are salts of organic acids (sulfonic and carboxylic) and are usually marketed as the sodium salts. The acid groups confer water solubility on the dyestuff molecule. When dissolved, the dye

ionizes (separates into particles with opposite electric charges), with the dye structure in the anionic (carrying the negative charge) part.

These dyes are used principally for wool, natural silk, synthetic fibers of polyamide and polyacrylic nature, leather, and paper. They are normally applied to the fiber in a solution containing some sulfuric or acetic acid, although some acid dyes will be fixed on the fiber from a neutral bath.

Because of improved fastness resulting from treatment of the dyed material with metal salts, especially chromium, a large number of acid dyes are available which contain metal atoms as part of the anion, as distinguished from the cationic metal atoms which form the salts of the organic acid. These anionically bound metal atoms do not exhibit the usual reactions of metal ions and are said to be chelated. These metalized (or premetalized) dyes have markedly superior fastness, approaching the vat dyes in this property. *See* CHELATION.

Some organic chemical compounds (mainly stilbene derivatives) have substantivity for fibers, but do not absorb light in the visible spectrum and hence show no color. They do transform some of the ultraviolet light which they absorb into visible light, thereby increasing the amount of white light reflected from them. This gives a bluing effect to yellowed materials and makes them appear whiter. Such products, used on cloth and paper, are called whites dyes, optical bleaches, or optical brighteners.

Basic dyes. These, too, are salts, but of organic bases containing amino and imino groups. The colored base is combined with a colorless acid, such as hydrochloric or sulfuric, and in solution, the dye structure is in the cation (carrying the positive charge). These dyes have exceptional brightness, but generally have only fair to poor fastness except when applied to acrylic fibers. They have wide utility for coloring wool, silk, leather, acrylic fibers, and paper.

Mordant dyes. Dyes which have little or no affinity for certain substrates may yet be fixed thereon if a mordant has first been applied. The fixation of the color is principally the result of reaction with the mordant material. These colors are also called adjective colors, as compared to the direct-dyeing substantive colors. Basic dyes which are applied to mordanted cotton are commonly still called basic colors, so that the designation mordant practically covers only acid-dye types and alizarin. The most common mordants are chromium salts for chrome dyeing processes. Often the chromium compound is applied during the dyeing, along with the dye, or it can be applied after the dyeing. These treatments may not only change the shade of the original dyeing (the self-shade), but they also may improve fastness of the dyeing to both water and light. *See* MORDANT.

Direct dyes. These dyes are normally sodium salts of sulfonic acids, and the colored part of the molecule is the anion. They differ from the acid dyes in that they are so substantive to cotton or other cellulosic fibers that they are fixed on the fiber from an aqueous solution with the assistance only of additions of common salt or sodium sulfate to the dyebath. Some authorities limit the designation substantive to these direct colors because of their outstanding affinity. These colors are also of importance in coloring paper, leather, and silk and have many miscellaneous uses. Because of the ease of application, they constitute the bulk of the package dyes used by the homemaker. This is an important class of dyestuffs, exceeded in numbers only by the acid dyes and in quantity used (in the United States) only by vat dyes.

Many direct dyeings can be improved in wetfastness by an aftertreatment such as with copper or chromium salts, formaldehyde, resins, and cationic fixing agents. It is possible to modify dyeings made by dyes of certain structures by a further chemical treatment (diazotization of a free amino group in the dye molecule, followed by coupling with a developer) of the dye on the fiber. Dyes suitable for such treatments are developed dyes. These treatments may or may not change the self-shade. *See* DIAZOTIZATION.

Ingrain dyes. These dyes are formed directly on the substrate by some type of chemical action. The principal subclasses are the azoic dyes and the oxidation dyes. In a few instances, dyes of the phthalocyanine type are developed on the fiber by special treatments.

Azoic dyes are water-insoluble azo compounds which have been formed within the substrate by chemical reaction of the intermediate components. Generally, the dyeing operation proceeds by dipping the cloth into a solution of one of the components (a hydroxy or amino component), drying the cloth without rinsing, and then treating it with a solution of the other component (usually a diazo component, which must be kept cold to prevent decomposition; hence, these colors are often called ice colors). Because the product of the reaction, which will then be deposited throughout the fiber, is a water-insoluble color, it is actually a pigment instead of a dyestuff. These colors are used for cotton and rayon, and in great quantity in printing.

Oxidation dyes are produced directly in the fiber or other substrates by a chemical oxidation following the impregnation of the substrate with certain aromatic amines. The final product is probably a pigment instead of a dyestuff. These colors are used principally for the dyeing of hair and fur.

Disperse dyes. These colors were originally developed for use on cellulose acetate, a synthetic fiber. This fiber has little affinity for the older known classes of dyestuffs. Some very slightly water-soluble colored materials transfer to this fiber from a water suspension if the colors are extremely finely divided particles. This results in a solution of the dye in the solid fiber.

Vat dyes. This class of colors is distinguished by the special method of application needed—a vatting operation, wherein the water-insoluble color is made soluble by a chemical reduction of the chromophore, the ketonic $=C=O$ group, to the $\equiv C-OH$ group, the leuco compound, which in the presence of alkali forms the water-soluble leuco salt, $\equiv C-OAlk$. This vat solution, which is often a different color from the original insoluble material, has affinity for

cotton, dyeing it with the shade of the vat solution. Upon oxidation with air or oxidizing agents, the reaction reverses to form the original water-insoluble color, leaving it deposited in the fiber as a pigment.

Sulfur dyes. Dyes of this class are also applied by a vatting technique which makes the solubilized color substantive to cellulosic fibers. These colors are made by treating a wide variety of organic compounds with sulfur and sodium sulfides. With a few exceptions, the final reaction products are not well-identified chemical compounds with known structures. These water-insoluble dyes are dissolved in alkaline sodium sulfide solution, which serves both as reducing agent and as source of alkali. After the dye is applied in the soluble leuco form, oxidation produces the insoluble dye on the fiber.

These colors are mostly used on cotton and viscose rayon. They have moderate all-around fastness and are relatively inexpensive. Hence, they are used in large quantities.

Solvent dyes. These colors are soluble in organic solvents such as benzene, gasoline, alcohol, acetone, oils, fats, and waxes. Solvent dyes color merely by solution of the dye in the substrate. They may be subclassified as spirit-soluble colors with principal solubility in alcohol and as oil-soluble colors which are soluble in benzene and vegetable and mineral oils. Uses include wood stains and varnishes, lacquers, printing and writing inks, butter and margarine, and plastics and resins. Wesley Minnis

Reactive dyes. Dyes that during the process of dyeing form a covalent bond with the substrate are known as reactive dyes. They are of particular interest on cellulosic fibers, where good washfastness is obtained without resorting to the production of large insoluble agglomerates within the fiber such as occur with vat and azoic dyes.

Structure. A reactive dye (see **illus.**) is a soluble color that contains a group capable of reacting with the hydroxyl or amino groups in cellulose, wool, silk, and other substrates. The reaction must be rapid and be carried out in the presence of water. Generally, an acid acceptor and heat are used to promote the reaction. Reaction efficiency varies from 60 to 95%, and unreacted dye must be removed by scouring to ensure washfastness.

Two types of reactive groups have been found successful in dyes which combine chemically with cellulose: active-halogen compounds which form ester or amide links [reaction (1)], and activated vinyl groups which form ether links [reaction (2)].

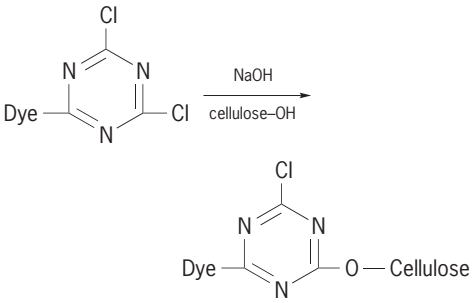

(1)

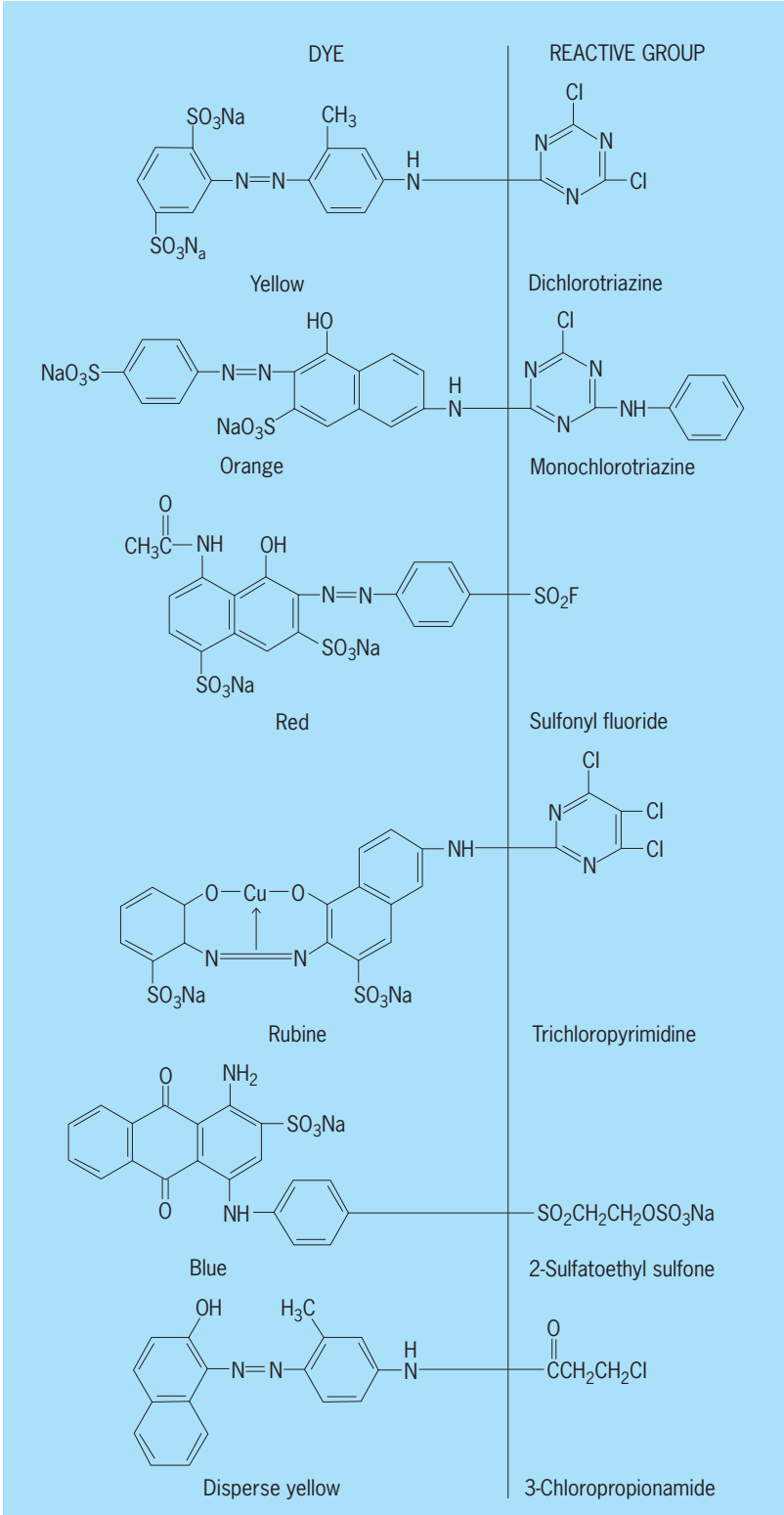

Structural formulas for typical reactive dyes.

$$\text{Dye}-SO_2CH_2CH_2OSO_2Na \xrightarrow{NaOH} \text{Dye}-SO_2CH{=}CH_2 \xrightarrow{\text{cellulose-OH}} \text{Dye}-SO_2CH_2CH_2O-\text{Cellulose} \qquad (2)$$

Application. Reactive dyes can be applied by either continuous or batch techniques. The most popular continuous methods are the pad, dry chemical pad, steam and scour, or pad thermofix and scour procedures. Batch application involves exhausting the dye with salt and then introducing the acid acceptor to promote reaction with the fiber. The method of application depends on the type of reactive groups and on the material being dyed. Nitrogenous fibers generally react at neutral or slightly acid conditions, and hydroxy-containing fibers require an alkali to promote reaction. When the reactive group is stable in the presence of the acid binder at room temperature, the process, becomes simply pad, heat, and scour.

Use and potential. The major penetration of reactive dyes has been into the cellulosic fiber dye market, as they offer improved washfastness over direct dyes, improved brightness over vat dyes, and improved rubbing fastness and shade range over azoic dyes. Reactive dyes are one of the more important single types of cellulosic dye. To a lesser degree, they are used for polyamide fibers.

The simplicity of continuous application and the brightness of reactive dyes make them useful in printing as well as dyeing. Both prints and dyed fabrics may be resin treated without significantly affecting the dye properties when selected resins are used. *See* HETEROCYCLIC COMPOUNDS; QUINONE.

Donald R. Baer; David H. Abrahams

Bibliography. S. V. Kulkarni et al., *Textile Dyeing Operations: Chemistry, Equipment, Procedures, and Environmental Aspects*, 1986; H. A. Lubs (ed.), *The Chemistry of Synthetic Dyes and Pigments*, 1955, reprint 1972; E. R. Trotman, *Dyeing and Chemical Technology of Textile Fibers*, 6th ed., 1991.

Dyeing

The application of color-producing agents to material, usually fibrous or film, in order to impart a degree of color permanence demanded by the projected end use. True dyeing covers mechanisms in which molecules of material to be dyed become involved by various means with the molecules of the coloring matter, or small aggregates thereof. There is some overlapping between true dyeing and other methods of coloring, which are called dyeing in the industry. Products which are commonly dyed include textile fibers, plastic films, anodized aluminum, fur, wood, paper, leather, and some foodstuffs.

The broad term affinity is used to describe the various types of attraction between the material to be colored and the dye. Affinity may be caused by attraction between charged dye particles and oppositely charged dye sites on the material, by various types of chemical attraction, and by formation of solid solutions of dye in the material.

When affinity is involved, dyeing is an exothermic process which may be simply stated as follows: dye in solution + undyed material → dyed material + heat.

This simplified equation explains the universally known fact that dye has a greater tendency to bleed into hot water than into cold water. It also points out that more dye will ultimately be absorbed by a fiber or film at low temperatures than at higher temperatures when equilibrium is finally reached. However, the rate of dyeing increases geometrically, whereas maximum dye absorption decreases only arithmetically with rise in temperature. Maximum dye absorption is rarely necessary or desirable; therefore, the trend in modern practice is toward dyeing at high temperatures and short times.

Dyeing is accomplished by dissolving or dispersing the colorant in a suitable vehicle (usually water) and bringing this system into contact with the material to be dyed. Although many dye molecules or aggregates may adhere to the material surface when they meet, dyeing does not occur until the adhering dye particles migrate within the fibers of films. All dyeing processes are designed to accomplish ultimately penetration of the undyed substance by the colorant.

Assistants. These are materials which do not impart color to the product to be dyed but, as the name suggests, promote or retard dyeing. Usually, but not always, they affect the dye molecule.

Swelling agents are assistants which open up the structure of the fiber temporarily so that dye molecules or aggregates may enter more freely and reach otherwise inaccessible dye sites.

Carriers are agents (often solvents of low water solubility) which accelerate dyeing by breaking up or dissolving dye aggregates and bringing them to the fiber-water interface in a size small enough to be absorbed by the material. Frequently, a carrier may exert a swelling action on the fiber in addition to its normal function.

Dye retarders are a class of dyeing assistants, usually inorganic or organic salts, which slow up the dyeing process by forming evanescent compounds with the dye, by buffering or depressing the ionization of an acid assistant, or by temporarily occupying the more active or more accessible dye sites on the fiber, later to be dislodged therefrom by the dye.

Aftertreating agents are salts, resins, or other products (more frequently applied to cellulosic fibers) to render the colored fabric more resistant to the effects of washing, perspiration, or fading by ozone or combustion gases. More often than not, their application causes a loss in lightfastness of the dyed material. Aftertreatment with copper salts, for example, is normally in order to increase lightfastness, whereas application of formaldehyde-urea or melamine resins to increase wetfastness normally affects sunlight resistance adversely.

Cellulose fibers. Cotton and rayon are most commonly dyed by immersion of the fibers in a solution of direct dyes using an electrolyte such as common salt as assistant and then boiling this dyebath. The affinity in this case may be a result of hydrogen-bonding of areas in the dye molecule to hydroxyl groups in the cellulose, more simply stated as adsorption. Such dyeings usually exhibit only commercial

(minimum) resistance to washing. Treatment of the properly dyed fibers with resins and copper, for example, increases the resistance to washing with minimum loss of light resistance. Some of the direct dyes have chemical groups (amino) which enable them to be converted after dyeing to large, more insoluble molecules by treatment with nitrous acid (diazotization) and a phenol or an amine (developing). Such dyeings are quite resistant to degradation by washing, but are usually characterized by poor resistance to fading by sunlight. Naphthol dyeing is a special form of developed dyeing wherein the naphthol is a selected type of substantive developer and is applied to the fiber first. Then a so-called diazo salt (not a dye) is added and an insoluble pigment forms within the fibers. *See* NATURAL FIBER.

Cellulose fibers may be dyed by vat dyes. These dyes are normally insoluble pigments. However, under the influence of alkali and reducing agents, they become water soluble and exhibit affinity for various fibers. After dyeing, the vat dyes revert to their insoluble form upon exposure to air or to oxygen-supplying chemicals. Indigo is a vat dye which does not possess affinity. Its alkaline-reduced solution, therefore, can be impregnated into the fiber only by successive dips and oxidations. Most vat dyes when properly applied are extremely resistant to degradation by such agents as washing, bleaching, sunlight, and perspiration.

Cellulose may be dyed by so-called fiber-reactive dyes which form tenacious bonds to cellulose and actually convert the cellulose to a different compound. The connecting links between dye and fiber are considered to be ether or ester linkages.

Conventional sulfur dyes are applied to cellulose from a sodium sulfide solution and are characterized by subdued hues, good washfastness, and low bleach resistance. They are widely employed for the production of deep shades on cottons destined for use in work and play clothes. Sulfur dyes made soluble by sodium thiosulfate are now available. These dyes are applied continuously to cellulose in the oxidized state (no affinity at this point), then reduced in a steamer thereby regaining "affinity" and thus dyeing the goods. Some of the newer "thiosulfate" or "condense" dyes produce hues far brighter than those obtainable with conventional sulfur dyes.

Basic dyes can be dyed upon cellulose which has previously been mordanted with synthetic tannins or tannic acid and tartar emetic. Dyeing is accomplished by the formation of a "lake," or pigment of the dye and mordant; the function of the cellulose is merely that of a substrate in this type of dyeing. Such dyeings are brilliant in hue but low in light resistance, and, for the latter reason, the use of basic dyes on cotton has become obsolete.

Animal fibers. The dyeing of wool, silk, and fur (felt) involves the formation of salts by reactive positively charged groups on the fiber and negatively charged color groups on anionic dyes.

Level-dyeing (acid or anionic) dyes, which require a relatively large amount of strong acid to force the dye onto the fiber, are used for carpet yarns, felt hats, and wherever even dyeing and penetration instead of washing fastness is a paramount requirement. Milling dyes, which demand less acid, are faster to washing and are used for blankets and sweaters. Where a very high resistance to light and washing is required, acid-dyeing or neutral-dyeing metallized dyes are employed.

Chrome (or mordant) dyes are applied to wool where subdued hues of excellent fastness are needed. In the metachrome method, chrome (sodium bichromate) is added to the dyebath at the start of the dyeing process. If the wool is treated with the chrome mordant before dyeing, the term chrome-bottom is used; if the chrome is added at the end of the dyeing cycle, the term top-chrome or after-chrome dyeing is applied. If silk-effect threads (such as in pin-striped cloth) are present, the after-chroming is performed in a separate fresh bath, a technique called the silk-white process.

Vat dyes and vat esters are also dyed on wool when bright shades of extreme fastness are required. The leuco vat esters are developed by acid and oxidizing agents, and the vat dyes are dyed as on cellulose but at lower temperature, lower alkalinity, and with a greater amount of reducing agent (sodium hydrosulfite). Fiber-reactive dyes for wool have been developed.

Manufactured fibers. Cellulose acetates and triacetate (Arnel) are dyed in a supension of solvent-soluble dyes by immersion (disperse-acetate). These dyes are considered to form solid solutions in cellulose acetate by passing from the water phase to the water-fiber interface and migrating from the fiber surface inward. In a technique similar to the diazotizing and developing technique described under cellulose dyeing, water-insoluble dye precursors are applied to the cellulose acetate; diazotization is accomplished with sodium nitrite and acid; β-hydroxynaphthoic acid or hydroxanil are added to develop an insoluble colorant within the fiber. Cellulose acetate may also be dyed by immersion in alcohol-water solutions or in formic acid-water solutions of certain water-soluble acid dyes.

Polyamide fibers are dyed like wool with acid, metallized acid, neutral metallized, fiber-reactive mordant dyes, azoics, and selected direct dyes from an acid bath. Vat dyes are applied by methods used for cotton dyeing, but higher temperature and swelling agents such as *o*-phenylphenol are used. Disperse dyes are dyed on nylons when maximum freedom from streaks (hosiery and flat woven goods, for example) is a prerequisite and resistance to washing and light is less important.

Acrylic fibers such as Orlon and Acrilan-16 are dyed in light shades with disperse (acetate) dyes at high temperatures. Being negatively charged, the acrylics are most successfully dyed to heavy hues with cationic (positively charged) colorants. By simultaneous use of copper salts and reducing salts, acrylics may be dyed with anionic dyes (the so-called cuprous-ion method). The positive cuprous ion forms a link between the negative dye molecule and the negative fiber molecule.

Modified acrylics (Acrilan, Zefran, Creslan, and Verel, for example) are dyed with disperse and basic dyes, as well as with dyes normally applied to natural fibers. The modification of the fibers involves the introduction of basic materials which form dye sites for anionic dyes.

Polyester fibers (Dacron, Terylene, Kodel, and Fortrel) are dyed with disperse dyes at high temperatures or by the use of swelling agents and carriers such as salicylic acid or chlorinated solvents. These fibers may also be dyed under pressure at high temperatures by leuco vat esters and disperse dyes or by applying these same dyes, drying, and passing through a hot flue or over heated rollers at about 400°F (204°C; Thermosol method).

Polypropylene fibers have been dyed in medium shades with leuco vat esters and sulfur dyes from aqueous baths at high temperatures by using thiourea dioxide as a reducing agent and diammonium phosphate and dispersing agents as dye assistants. Azoics from an acid bath and vat acids are also applicable. Solvent vapors have been used to dye polypropylenes with solvent soluble, oil, disperse, and related dyestuffs.

Polypropylene has been altered by grafting or with copolymers which permit dyeing the modified fiber with anionic or disperse dyes. A unique coloring method involves the use of dyes which can chelate onto nickel. (Some polypropylene brands contain nickel stabilizers.) Such chelate dyeings and prints have extremely high sunlight fastness. *See* MANUFACTURED FIBER.

Pigment dyeing. All fibers may be colored by application of pigments dispersed in a bonding agent. Affinity or true dyeing is not involved. Good fastness of pale shades to most normal color-destroying effects and of heavy shades to all but rubbing-off is characteristic; however, the bonding agent may stiffen sheer goods.

Dope or spin dyeing. Melt spun fibers such as polypropylene, nylon, Dynel, or Dacron may be colored in the mass. The colored plastic is melted and then extruded through spinnerets (refined shower heads), and the congealed streams are collected and processed to yarn quality.

Fibers such as cellulose acetate, acrylics, and viscose rayon are prepared by extruding their solutions into hot air or a coagulating water bath. These so-called dopes are colored with pigments before extrusion, and the coagulated fibers collected and processed into yarn.

Textile-dyeing equipment. There are two fundamental types of dyeing processes—the dye liquor is pumped through stationary material, or the material to be dyed moves through the dye liquor.

Kiers are tanks which hold loose fiber or packaged yarn and the dye liquor is pumped through the fiber at temperatures up to the boiling point or (if the fiber can withstand it) under pressure at temperatures up to 270°F (132°C). In one machine, goods are wound upon a perforated roller (or beam), and the whole placed in a pressure chamber. Dye liquor under pressure above 212°F (100°C) is pumped through the goods. Several other machines which function by pumping dye liquor through stationary material are used in modern mills.

In continuous dyeing, cloth is impregnated with dye and then passed through a series of developing, washing, and drying zones to a final take-up roll. One machine uses a passage through molten Wood's metal to develop the dye-saturated goods, whereas another uses hot oil for this purpose.

The pad-stream system consists of rollers which apply dye and necessary chemicals to the goods; development takes place on continuous passage through a steam chamber.

The pad-roll system uses an insulated oven in which a huge roll of goods, previously saturated with dye solution, slowly turns until the dye has become fixed to the fiber.

A jig promotes dyeing by winding goods through the dye bath, back and forth from one roll to another.

A beck consists of an elliptical reel which draws goods which have been sewn in an endless chain from the dye bath and plaits it back into the bath repeatedly until dyeing is complete. The use of a machine equipped with disked roller-pressure locks permits continuous dyeing at high pressures.

A relatively new machine applies dye continuously to carpet faces by means of a roller which picks up dye solution and then delivers this film of solution to be scraped off by a fluorocarbon "doctor." The resultant cascade of dye solution falls onto the carpet face, and the carpet moves forward into a steamer. Gum thickeners and a foam generator (possibly ethyl carbonate) hold the dye solution evenly throughout the pile height enabling even and nonfrosty (light pile tips) dyeing. A somewhat similar machine employing a padder and steamer is used for tow or a loose web of fibers. The latter's practical use predates the carpet dyeing machine. It employs the principle of coacervation to effect rapid dye fixation. Equipment which continuously develops disperse dyes on polyesters by dry heat (thermofixation) is wisely used. Dyeing continuously in solvent solutions or in solvent vapors is being investigated to speed up production and to minimize pollution.

Nontextile materials. Dyeing anodized aluminum was formerly believed to be accomplished by the formation of aluminum salts of dyes by the electrically deposited aluminum oxide coating. However, many dyes having no salt-forming groups were found to dye this film, and it is postulated that, at least in these cases, conical craters receive the colorant, and subsequent sealing by boiling in water or in solutions of salts in water converts the oxide to larger hydrated molecules (for example) and imprisons the dye aggregates in the sealed-off cones. At any rate, anodized aluminum is readily dyed by many textile dyes. Light and weather resistance undreamed of in textile applications of some of these same dyes is achieved.

Paper. Paper pulp is usually dyed in the paper beater by dyes normally employed for cotton; on occasion, it is tinted by wool dyes, and it is frequently tinted by addition of pigments to the beater. Finished

paper is also colored by passing it over rollers which supply dye or colored coatings to its surface (calender staining). *See* PAPER.

Wood. Wood is normally stained with solutions of dyes or dispersions of pigments in water, in solvents, or in lacquers. Wood, however, is frequently dyed by the application of water solutions of dyes at high temperature under pressures of 100 lb/in.^{2} (1700 kilopascals) or more. Freshly felled trees have been dyed by force-pumping dye solution through the length of the log. The dye solution replaces the natural moisture and sap. An inflated tubular gasket seals the log-end to the pressure cylinder in this method. Wood, destined to be cemented to corks for liquor bottle closures, has been dyed with sulfur dyes to ensure resistance to alcohol.

Monomeric substances, such as methyl methacrylate or styrene, are colored with solvent-soluble dyes. These colored solutions are impregnated into wood, and the corresponding polymer is formed therein by radiation catalysis. Such wood is extremely durable and marproof since every fibril throughout the wood is "varnished" by this process.

Leather. Leather is dyed at low temperatures with the classes of dyes normally used for wool and cotton, excepting vats and naphthols. Formic acid is normally used to exhaust the dye. For dress gloves, leather is usually colored by applying the dye on the grain surface, leaving the flesh side undyed. Leather is also dyed with natural dyes such as logwood, fustic, and quercitron. Leather fresh from tanning and containing considerable moisture is dyed in Europe by tumbling with dry water-soluble dye. Greater penetration by the dye is claimed by this process as opposed to normal dyeing in water solutions of the dyes. *See* LEATHER AND FUR PROCESSING.

Food products. Most food products which are artificially colored are not actually dyed. Maraschino cherries, however, are dyed for several hours with food dyes, then washed and placed in flavored syrup. Glazed citron, citrus, or watermelon rinds are colored by long immersion followed by drainings in a series of food-dye-colored, increasingly strong, sugar syrups. Pistachio nuts in the shell are dyed by applying food-dye solutions with or without salt during roasting operations.

Fur. Fur is usually colored by impregnations with synthetic organic intermediate compounds (amines) which are then oxidized to colored pigments (azines) with peroxides, for example. When not fully developed, some of the products produce objectionable physiological reactions so that great care is taken to assure the absence of unreacted amine in the dyed fur.

Human hair. It is not generally realized that human hair is, in the final analysis, a class of textile fiber. Cosmetic companies, appreciating this fact, have hired textile experts to do research on dyeing, waving, and washing hair. The dyeing of hair is of course limited to processes employing relatively low temperatures. Food dyes (also carbon black) mixed with citric or malic acid are used for so-called rinses. Other rinses employing basic dyes and salt or textile dyes diluted with salts, detergents, and malic, citric, or tartaric acids are on the market. Products selected, as a rule, from among the safer fur-dye intermediates are used for more permanent and intense coloring techniques; these employ hydrogen peroxide for developing. Natural dyes, such as the age-old henna and some metallic salts, are still encountered in some hair dyeing establishments. Rinses which are used to mask the yellowish tinge in gray or white hair frequently use blue or violet food or textile dyes and contain optical bleaches—substantially colorless compounds which receive invisible ultraviolet light and reflect it to the eye as a blue or violet fluorescence.

Plastic films or products. Many plastic materials may be dyed by processes similar to those for textiles. Nylon, cellulose acetate, polyethylene, polypropylene, and polyester resins are dyeable with dyes which color these materials in yarn form. Solvents such as pyridine, dioxane, alcohol, and trichlorobenzene, when applied in solution or emulsion form, swell these and other resins as an aid in applying solvent-soluble dyes. Research is under way to develop and improve reactive dyes for polycarbonate and other new resins. *See* DYE; PLASTICS PROCESSING; TEXTILE CHEMISTRY.

James E. Loughlin; David H. Abrahams

Bibliography. Bibliographic section, *Technical Manual of the American Association of Textile Chemists and Colorists*, annually; S. V. Kulkarni et al., *Textile Dyeing Operations: Chemistry, Equipment, Procedures, and Environment Aspects*, 1986; R. H. Peters, *Textile Chemistry*, vol. 3: *The Physical Chemistry of Dyeing*, 1975; E. R. Trotman, *Dyeing and the Chemical Technology of Textile Fibres*, 6th ed., 1985.

Dynamic braking

A technique for braking in which mechanical energy is converted to heat or electrical energy in order to slow or stop motion. An all-mechanical dynamic brake consists of rotating vanes that circulate a viscous fluid in a manner that generates heat. This is one way that the power of the wind is harnessed for space heating. An electric dynamic brake consists of an electric dynamo in which the mechanical energy is converted to electric form, and either converted to heat in a resistor or returned to the supply lines. Typically, electric braking is accomplished with the same machine that serves as the drive motor. Electric dynamic braking is employed in electric vehicles, elevators, and other electrically driven devices that start and stop frequently. *See* WIND POWER.

The most common type of dynamic braking will be explained for a direct-current (dc) motor. To accomplish braking action, the supply voltage is removed from the armature of the motor but not from the field. The armature is then connected across a resistor. The electromotive force generated by the machine, now acting as a generator driven by the mechanical system, forces current in the reverse direction through the armature. Thus a torque is produced to oppose

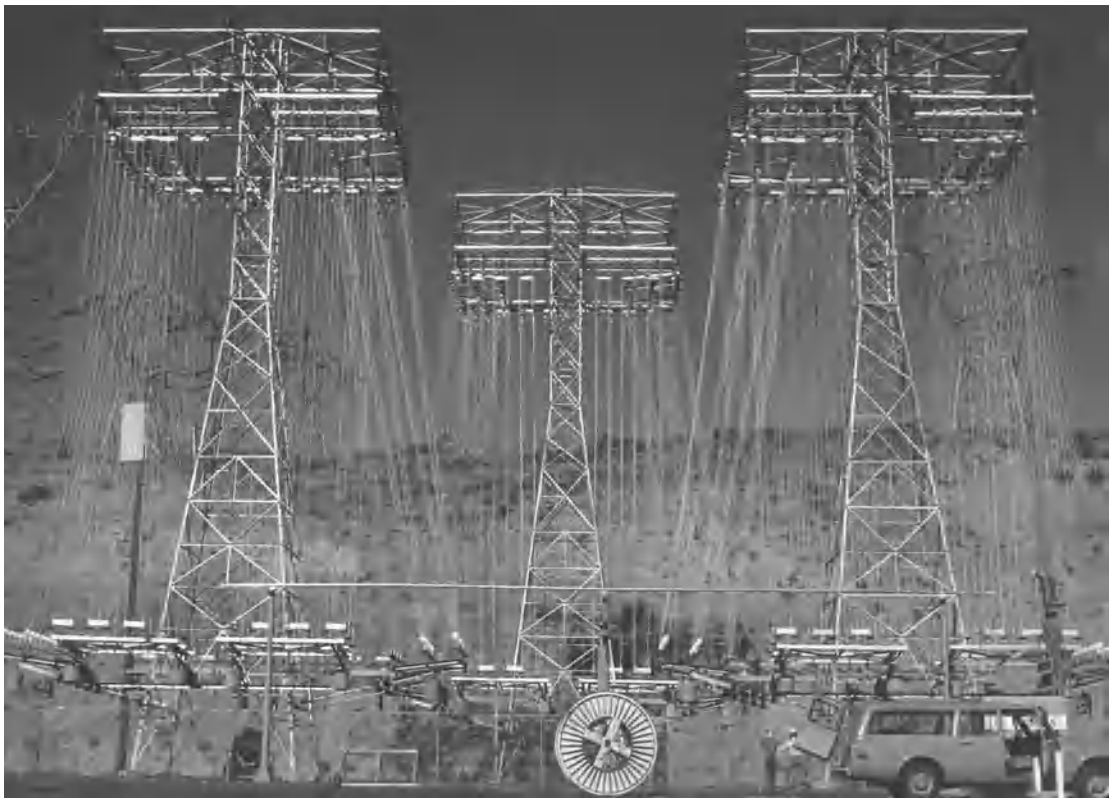

Dynamic braking resistor with a capacity of 1400 MW, installed in the Bonneville Power Administration's Chief Joseph Substation. (*From M. L. Shelton et al., Bonneville Power Administration 1400-MW braking resistor, IEEE Trans. Power Apparatus Sys., PAS-94:602–611, Institute of Electrical and Electronics Engineers,1975*)

rotation, and the load decelerates as its energy is dissipated, mostly in the external resistor, but to some extent in core and copper losses of the machine. *See* DIRECT-CURRENT MOTOR.

Electric braking can also be accomplished by causing the energy of the rotating system to be converted in the armature to electrical energy and then returned to the supply lines. This mode of operation, called regenerative braking, occurs when the counterelectromotive force exceeds the supply voltage.

Interchanging two of the lines supplying a three-phase alternating-current (ac) induction motor also produces braking. In this case, called plugging, the direction of the electromagnetic torque on the rotor is reversed to cause deceleration. Both the energy of the system and the energy drawn from the supply lines are expended in copper and core losses in the machine as heat. The power lines must be disconnected when the rotor comes to rest. *See* INDUCTION MOTOR.

Regenerative braking is a system of dynamic braking in which the electric drive motors are used as generators and return the kinetic energy of the motor armature and load to the electric supply system. This method is employed when the load is losing a large amount of potential energy, as in the case of an electric train descending a long grade or a pumped storage reservoir being emptied. The potential energy of the load is first converted to kinetic energy in the motor armature and load and then to electrical energy. The method is applicable to dc motors, to induction motors (with negative slip), and to synchronous motors. *See* PUMPED STORAGE; SYNCHRONOUS MOTOR.

Rheostatic braking is a system of dynamic braking in which direct-current drive motors are used as generators and convert the kinetic energy of the motor rotor and connected load to electrical energy, which in turn is dissipated as heat in a braking rheostat connected to the armature. This is accomplished by disconnecting the motor armature from the supply lines and connecting them to a braking rheostat, while the field circuit remains connected to the supply line. The polarity of the armature remains unchanged, but the direction of the armature current reverses, resulting in a negative torque. The torque can be controlled by means of the braking rheostat.

Sometimes the term dynamic braking is applied only when mechanical energy is dissipated in an external resistor, but its more general interpretation includes regenerative braking and plugging.

Dynamic braking is also applied in electrical utilities as an aid in stablizing operation of interconnected systems. A fault on an electric line has the effect of reducing both the electric power out of a generator and the power transmitted through an interconnection tie line. The prime mover governors, particularly for hydroelectric generators, cannot respond fast enough to avoid acceleration of the speed of one interconnected generator relative to another. The machines no longer operate synchronously, and separation of the interconnection points results.

Dynamic braking resistors can be rapidly connected to a generator bus as a dummy load until the fault is cleared and the normal power flow through the interconnection is reestablished. In this way the acceleration of the generator may be limited so that the synchronizing power of the tie line can maintain stability. The resistors employed for this purpose are huge. The dynamic braking resistor installed in Bonneville Power Administration's Chief Joseph Substation has a capacity of 1400 MW (see **illus.**). *See* ALTERNATING-CURRENT GENERATOR; ELECTRIC POWER GENERATION; ELECTRIC ROTATING MACHINERY; TRANSMISSION LINES. Arthur R. Eckels

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000.

Dynamic instability

A state of fluid flow in which the distribution of mass and momentum is unstable. Fluid flows are subject to a variety of instabilities which generally cause the flow to become more complex and which often lead to turbulent, chaotic flow. Instabilities are responsible for a variety of phenomena in natural flows, including cyclones, hurricanes, and thunderstorms in the atmosphere; mantle convection in the Earth's interior; and granules and supergranules in stellar atmospheres. A great deal of research in the geophysical and astrophysical sciences has focused on flow instabilities; and instability plays an important role in engineering problems ranging from naval engineering and aeronautics to the design of efficient heating and cooling systems. *See* FLUID FLOW; TURBULENT FLOW.

Fluid instabilities may be broadly divided into two classes: convective and dynamic. While convective instabilities can be understood rather easily in terms of forces acting on displaced parcels of fluid, dynamic instabilities are more varied and more challenging to understand. This broad class of fluid instabilities is responsible for the cyclones and anticyclones that dominate the weather in middle and high latitudes, as well as meanders and rings in ocean currents such as the Gulf Stream. These instabilities are ultimately responsible for the lack of predictability of complex flows such as are found in the atmosphere. *See* CONVECTIVE INSTABILITY; CYCLONE; WIND.

Vorticity and invertibility. Many, but not all, dynamic instabilities can be understood with the aid of a vorticity principle and an invertibility principle. The vorticity of a fluid can be thought of as the rate of rotation of a rigid paddle wheel embedded in the flow, as illustrated in **Fig. 1**. (Mathematically, it is the curl of the vector velocity field.) The rotation can be brought about by shear (a change in the speed of the flow in the direction across the flow) and by the curvature of the flow. Vorticity is important for two reasons. First, in two-dimensional flows, vorticity is conserved; that is, if the paddle wheel is followed around it is observed that its rate of rotation remains

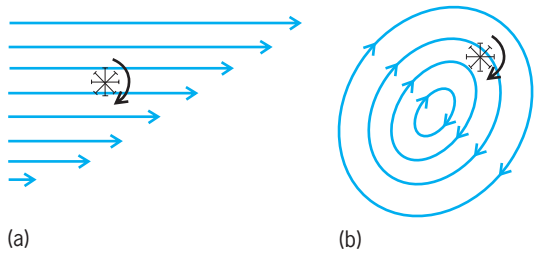

Fig. 1. The vorticity of a fluid is related to the rate at which a rigid paddle wheel would rotate if placed in the flow. (*a*) The paddle wheel rotates because the velocity of the fluid varies in a direction perpendicular to the flow. (b) Rotation is generated by the curvature of the flow.

constant. Second, in such flows, it is possible to work backward from knowledge of the vorticity at every point in the fluid to find the flow field at every point. This is called the invertibility principle. Basically, a clockwise vorticity at a point in the fluid will induce a clockwise rotation of the fluid nearby; the total velocity field in the fluid can then be found by adding the velocities induced by all the vorticities at each point. *See* CALCULUS OF VECTORS.

Barotropic instability. A shear flow is shown in **Fig. 2***a*. All the shear in this case is concentrated on a line, and thus all the vorticity is concentrated on the same line. Therefore, this flow can be perceived as being composed of a line of vortices (Fig. 2*b*). However, this flow turns out to be unstable, in the sense. If two of the vortices are displaced in the manner shown in Fig. 2*b*, the vorticity is conserved so that the rate of rotation of the vortices does not change. However, the flow induced by each of the two displaced vortices moves the other vortex in the same direction as it was displaced in originally. Thus the slightest perturbation of the original line of vortices will amplify rapidly, changing the whole flow. This is called barotropic instability, and is responsible for such phenomena as meanders in the Gulf Stream. Such instabilities draw on the kinetic energy of the original shear flow. A necessary condition for barotropic instability is that the vorticity has a maximum or minimum within the flow. Generally, an extremum strung out in a line such as in the example in Fig. 2 is unstable, while circular patches of high or low vorticity can often by stable. *See* BAROTROPIC FIELD.

Baroclinic instability. The flow of the Earth's atmosphere is three-dimensional. Under this

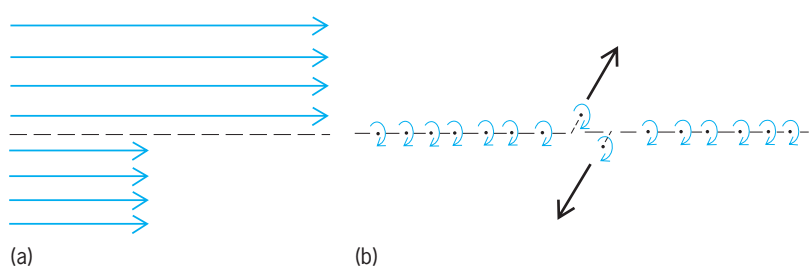

Fig. 2. Diagram showing barotropic instability. (*a*) All of the vorticity of an idealized shear flow shown concentrated on a line. (*b*) Line envisaged as a row of individual vortices. If two vortices are displaced in the manner illustrated in *b*, the flow they "induce" at each other's position enhances the original displacement.

circumstance, vorticity is not conserved because fluid converging on or diverging from a point in the fluid causes it to spin faster or slower. (The principle at work here is the conservation of angular momentum.) Fortunately, the Earth's atmosphere is usually stably stratified; that is, a parcel displaced upward will be colder than its environment and will accelerate downward toward its original position. In this sense the atmosphere is convectively stable. It is possible to identify a quantity that reflects the air's density and at the same time is conserved as long as no heat is added to or subtracted from the air. This quantity is called the potential temperature (θ) and is related to temperature and pressure by the expression

$$\theta = T\left(\frac{1000}{p}\right)^{0.287}$$

where T is the temperature in kelvins and p is the pressure in millibars. Generally, θ increases upward in the atmosphere: slowly in the troposphere (a layer of air that extends upward to roughly 6 mi or 10 km) and more rapidly in the stratosphere (a layer of air between about 6 and 18 mi, or 10 and 30 km). *See* STRATOSPHERE; TROPOSPHERE.

Because θ is conserved, a layer of air between two different surfaces of constant θ (which are nearly horizontal surfaces in the atmosphere) is always trapped between those surfaces. (If air were to cross θ surfaces, this would imply that θ is changing as the air is followed around.) If it is observed that two θ surfaces are moving farther apart, the implication is that air must be converging horizontally to fill the gap. If a floating paddle wheel could be placed on a θ surface in this circumstance, it would be observed that its spin increases with time, because of the convergence of the flow. The farther apart the surfaces are, the greater the rate of spin, or vorticity. In fact, the vorticity is directly proportional to the distance between two θ surfaces. This can be used to define a conserved quantity called potential vorticity, which is simply the ratio of the vorticity to the distance between θ surfaces. (Mathematically, it is the dot product of the curl of the vector velocity field with the gradient of θ, all divided by density.) An interesting aspect of potential vorticity is that for sufficiently large-scale flows, such as the low- and high-pressure systems that affect daily weather in middle and high latitudes, it also has an invertibility principle that is quite analogous to the invertibility principle for vorticity in two-dimensional flows. In the case of three-dimensional flows, a blob of high-potential vorticity will induce a counterclockwise circulation of fluid around it and above and below it (in the Northern Hemisphere).

The criterion for the dynamic instability of large-scale three-dimensional flows is also analogous to barotropic instability of two-dimensional flows. The necessary condition for this kind of instability is that potential vorticity must be a maximum or minimum somewhere on a surface of constant θ; once again, extrema of potential vorticity arranged along a line are almost always unstable, whereas circular blobs are often stable. The type of instability that results from maxima or minima of potential vorticity in a fluid is called internal baroclinic instability. This type of instability can draw on both the kinetic energy of the original flow and the potential energy that occurs when there are horizontal variations of density.

Observations of the Earth's atmosphere show that the potential vorticity of the troposphere in middle and high latitudes is nearly uniform. But when a gently sloping θ surface is followed into the stratosphere, a large increase of potential vorticity occurs. Even so, there is no localized maximum or minimum of potential vorticity, but just a large gradient of the quantity, concentrated near the boundary between the troposphere and stratosphere. (Many meteorologists define the stratosphere as a region of large potential vorticity.) For this reason, internal baroclinic instability seldom occurs in the atmosphere.

In order to understand the dynamic instability that causes cyclones and anticyclones (the familiar low- and high-pressure weather systems on weather maps), it is necessary to introduce a corollary to the invertibility principle, which recognizes an important and unique role of solid boundaries (such as the Earth's surface) within a fluid. This corollary states that the temperature of the fluid at the boundary behaves just as the potential vorticity does away from boundaries. That is, a patch of high temperature at the Earth's surface induces a counterclockwise rotation of the air around it, just as though this air were a blob of high-potential-vorticity fluid. Because of this corollary, it turns out that a condition for dynamic instability in the Earth's atmosphere is that the temperature gradient at the Earth's surface have the opposite sign of the gradient of potential vorticity on θ surfaces. Since surface temperature generally decreases poleward, while potential vorticity on θ surfaces increases poleward as the gently sloping θ surfaces are followed into the stratosphere, this condition is usually met in middle and high latitudes. The resulting instability is referred to simply as baroclinic instability.

The mechanism of baroclinic instability can be understood with the aid of **Fig. 3**, which shows the formation of a surface low-pressure system in the Northern Hemisphere. First, a tongue of high-potential-vorticity air downward and southward from the stratosphere. If this tongue is big enough, the counterclockwise circulation it induces extends all the way to the surface, where the tongue advects warmer air northward along the ground to its east. Once a tongue of warm air is formed at the surface, it also induces a counterclockwise circulation around it (by the corollary to the invertibility principle) which extends upward to the stratosphere. This counterclockwise circulation pulls down more high-potential-vorticity air from the stratosphere, reinforcing the original tongue, and so on. This positive feedback mechanism results in an amplification of the whole system; this is baroclinic instability. Also, the surface low-pressure center (the site of maximum counterclockwise circulation) is just west of

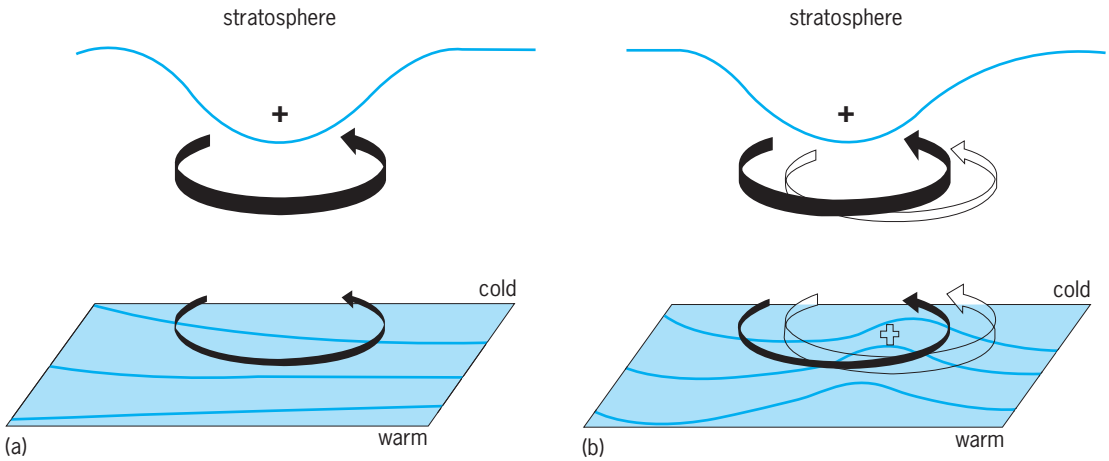

Fig. 3. Schematic diagram showing the formation of a cyclone in the Northern Hemisphere. (*a*) A tongue of high-potential-vorticity air from the stratosphere extends downward and southward into the troposphere; this tongue induces a counterclockwise circulation at the surface, where temperature decreases northward. (*b*) The induced circulation advects warm air northward, and the warm tongue induces a counterclockwise circulation of its own. This acts to pull more stratospheric air southward and downward, reinforcing the original perturbation. (*After P. J. Hoskins, M. E. McIntyre, and A. W. Robertson, On the use and significance of isentropic potential vorticity maps, Quart. J. Roy. Meteorol. Soc., 111:877–946, 1985*)

the tongue of warm air at the surface, which is where the circulations due to the warm-surface tongue and the high-potential-vorticity tongue aloft reinforce to the maximum extent. High-pressure systems form analogously, with cold tongues at the surface and low-potential-vorticity tongues aloft. *See* BAROCLINIC FIELD.

Other instabilities. Not all dynamic instabilities can be understood in the same terms in which barotropic and baroclinic stabilities may be comprehended. One such phenomenon is the Kelvin-Helmhotz instability, which occurs in stratified shear flows, that is, flows which have horizontal wind changing with height and which are at the same time convectively stable (θ increasing with height). If the wind shear is stronger than a critical value that depends on the degree of stratification, instability can occur. The resulting motions take the form of waves that often break, leading to turbulence. The Kelvin-Helmholtz instability is thought to be responsible for much of the clear-air turbulence encountered by aircraft. *See* CLEAR-AIR TURBULENCE.

There are still other forms of dynamic instability about which little is known. An example is fluid flow between two rigid, parallel plates moving at a constant speed with respect to each other. If the plates are moving slowly, the speed of the fluid flow varies smoothly from one plate to the other; this is known as Couette flow. Such a fluid has constant vorticity and is not stratified, so that it is not susceptible to any of the instabilities discussed previously. Yet when the plates are made to move fast enough, the fluid becomes unstable and quickly breaks down into a highly turbulent flow. Theoretical and laboratory work reported in 1986 has shown that flow within a rotating ellipse is unstable, even though it too has constant vorticity. Since perturbations to Couette flow have constant vorticity, the shape of the perturbation is usually elliptical and thus is unstable. This complex class of instabilities has begun to be explored, and the work may lead to important advances in the understanding of dynamic instabilities. *See* ATMOSPHERE; MICROMETEOROLOGY; NEWTONIAN FLUID.

Kerry A. Emanuel

Bibliography. A. E. Gill, *Atmosphere-Ocean Dynamics*, 1982; P. J. Hoskins, M. E. McIntyre, and A. W. Robertson, On the use and significance of isentropic potential vorticity maps, *Quart. J. Roy. Meteorol. Soc.*, 111:877–946, 1985; R. Pierrehumbert, Universal short-wave instability of two-dimensional eddies in an inviscid fluid, *Phys. Rev. Lett.*, 57:2157–2159, 1986.

Dynamic meteorology

The study of those motions of the atmosphere that are associated with weather and climate. Atmospheric motions span an enormous range of spatial and temporal scales; dynamic meteorology concentrates mainly on large-scale and mesoscale motions. Large-scale motions are those with horizontal scales in excess of a few hundred kilometers and time scales longer than a day. Such motions are strongly influenced by the rotation of the Earth and by the vertical thermal stratification of the atmosphere. Mesoscale motions have horizontal scales in the range of a few kilometers to a few hundred kilometers; they are often associated with convective clouds and precipitation.

The mechanics and thermodynamics of the unsaturated atmosphere are governed by three fundamental conservation laws: (1) conservation of mass, expressed by the mass continuity equation; (2) conservation of momentum, expressed by Newton's second law of motion; and (3) conservation of thermodynamic energy, expressed by the first law of thermodynamics. For saturated conditions, conservation

of water substance must also be considered. *See* CONSERVATION OF MASS; CONSERVATION OF MOMENTUM; NEWTON'S LAWS OF MOTION; THERMODYNAMIC PRINCIPLES.

Statics. The vertical structure of the atmosphere is determined by the equation of state for an ideal gas and by the hydrostatic relationship. The former expresses the relationship among pressure, density, and temperature at any point; the latter expresses the balance between the upward-directed component of the pressure-gradient force (associated with the approximate exponential decrease of pressure with height) and the downward-directed gravity force. The equation of state and hydrostatic equation may be combined to form the hypsometric equation (1),

$$Z_2 - Z_1 = \frac{R}{g}\int_{p_2}^{p_1} T\, d\ln p \tag{1}$$

which relates the geopotential height differences ($Z_2 - Z_1$) between two pressure surfaces p_2 and p_1 to the mean temperature T in the layer between the two surfaces; and where R (the gas constant for dry air) $= 287\ \mathrm{J \cdot kg^{-1} \cdot K^{-1}}$ and g (the acceleration of gravity) $= 9.81\ \mathrm{m \cdot s^{-2}}$. Thus, pressure decreases more rapidly with height in cold air than in warm air. *See* GAS; HYDROSTATICS.

Except in regions of active precipitation, where the heating rate due to latent heat of condensation is large, temperature changes following the motion of individual parcels of air are controlled primarily by adiabatic expansion and compression as the air parcels move to lower or higher pressure. The thermodynamic state of such parcels can be characterized by the potential temperature θ, as in Eq. (2), where p_0 [$=10^5$ pascals (1000 millibars)] is

$$\theta = T\left(\frac{p_0}{p}\right)^{R/C_p} \tag{2}$$

a reference pressure and c_p ($= 1004\ \mathrm{J \cdot kg^{-1} \cdot K^{-1}}$) is the specific heat at constant pressure. The potential temperature is the temperature that a parcel of air at pressure p and temperature T would acquire if it were moved adiabatically to pressure p_0. When all diabatic heat sources can be neglected, θ remains constant in time for each air parcel. Normally, θ increases with altitude in the atmosphere, so that an air parcel displaced upward (downward) has a value of θ less (greater) than that of its environment, and hence experiences a net buoyancy force that tends to return it to its equilibrium level. The atmosphere is then said to be statically stable. When θ is constant with height, a condition that occurs when the temperature decreases with height at a rate $10^\circ\mathrm{C \cdot km^{-1}}$ ($30^\circ\mathrm{F \cdot mi^{-1}}$), the atmosphere is said to be neutrally stable; if θ decreases with height, the atmosphere is absolutely unstable and convective motion develops spontaneously.

If an air parcel is saturated, upward displacement causes water vapor to condense and release its latent heat of condensation; the potential temperature is then no longer conserved. It is possible, however, to define an equivalent potential temperature θ_e, as in Eq. (3), where L_c is the latent heat of condensa-

$$\theta_e = \theta \exp\frac{L_c q_s}{c_p T} \tag{3}$$

tion ($=2.5 \times 10^6\ \mathrm{J \cdot kg^{-1}}$ at 0°C or 32°F), and q_s is the saturation water-vapor mixing ratio (that is, the ratio of the mass of water vapor to mass of dry air for a saturated parcel). The equivalent potential temperature is conserved for both saturated and unsaturated motions provided that the temperature used in Eq. (3) is the temperature that the parcel would have if brought to saturation by adiabatic expansion. Because q_s decreases rapidly as temperature decreases, θ_e for saturated conditions generally decreases with height even when θ itself increases with height. The atmosphere is then said to be conditionally unstable since the static stability is positive for unsaturated parcel displacements but negative for saturated parcel displacements.

The hydrostatic relationship implies that pressure decreases monotonically with altitude. Pressure may thus be substituted for height as the independent vertical coordinate. If the atmosphere is everywhere statically stable so that θ increases monotonically with height, potential temperature can also be used as a vertical coordinate. Potential temperature coordinates are useful for analysis of adiabatic motions, since in that reference frame prediction of adiabatic flow is reduced to a two-dimensional problem of following the motion on θ surfaces. Isobaric coordinates, in which pressure is the vertical coordinate, have the advantage of eliminating any explicit reference to the density field. These are the most commonly used vertical coordinates in dynamic meteorology. *See* HYDROSTATICS; ISOBAR (METEOROLOGY).

Conservation laws in isobaric coordinates. The horizontal momentum equation, the mass continuity equation, and the first law of thermodynamics are expressed in isobaric coordinates as shown in Eqs. (4)–(6), where ∇ denotes the horizontal gradi-

$$\frac{d\mathbf{V}}{dt} = -f\mathbf{k} \times \mathbf{V} - g\nabla Z + \mathbf{Fr} \tag{4}$$

$$\nabla \cdot \mathbf{V} + \frac{\partial \omega}{\partial p} = 0 \tag{5}$$

$$\frac{dT}{dt} = S_p\omega + \frac{\dot{q}}{c_p} \tag{6}$$

ent evaluated at constant pressure, t denotes time, $\mathbf{V}$ is the horizontal velocity component, $f = 2\Omega \sin\phi$ is the Coriolis parameter (where $\Omega = 7.292 \times 10^{-5}\ \mathrm{s^{-1}}$ is the angular velocity of rotation of the Earth and ϕ is latitude), $\mathbf{k}$ is the vertical unit vector, Z is the height of a surface of constant pressure, $\mathbf{Fr}$ is the force of friction, $\omega = dp/dt$ is the vertical motion in isobaric coordinates, $S_p = -T\,\partial \ln\theta/\partial p$ is the static stability, and $\dot{q}$ is the heating rate.

Equations (4) and (6) are expressed in lagrangian form; that is, the notation d/dt expresses the rate of change with respect to time following the motion of an individual parcel. For analysis and prediction purposes, it is usually more convenient to use the

eulerian form in which time derivatives are computed at fixed points in space (as partial derivatives). The differential operators for the lagrangian and eulerian forms are related by Eq. (7), where the second

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \mathbf{V} \cdot \nabla \tag{7}$$

term on the right side expresses the change due to advection by the horizontal wind.

Equation (4) states that the rate of change of the horizontal velocity following the motion of an air parcel is due to the net imbalance among three forces: the Coriolis force due to the rotation of the Earth, the horizontal pressure-gradient force, and the force due to frictional dissipation (mostly at the surface of the Earth). Equation (5) states that the divergence of the horizontal velocity, which tends to stretch or shrink the horizontal cross section of an air parcel, must be balanced by vertical shrinking or stretching due to the change in vertical mass flow with altitude. Equation (6) states that the rate of change of temperature following the motion is due to the sum of the adiabatic expansion or compression due to vertical motion, plus the net heating due to such heat sources as condensation and solar and thermal radiation (generally referred to as diabatic heating). Equation (2) is an integral of Eq. (6) in the case where diabatic heating vanishes.

Balanced flow. When the diabatic heating rate $\dot{q}$ and the frictional force $\mathbf{Fr}$ can be neglected, Eqs. (4)–(6), together with the differential form of the hypsometric equation (1), form a closed set for prediction of $\mathbf{V}$, ω, Z, and T subject to suitable initial and boundary conditions. This is a complicated set of nonlinear partial differential equations that in general must be solved by numerical methods. However, for large-scale atmospheric motions certain approximate solutions can provide considerable insight.

If friction can be ignored, which is often approximately the case above the lowest kilometer of the atmosphere, large-scale extratropical motions are approximately in geostrophic balance; that is, the horizontal pressure-gradient force and the Coriolis force are nearly equal and opposite. The wind velocity for which this balance is exact is referred to as the geostrophic wind, and is determined by Eq. (8), a

$$f\mathbf{V}_g = g\mathbf{k} \times \nabla Z \tag{8}$$

special case of Eq. (5). The force balance for the geostrophic wind is illustrated in **Fig. 1**. The Coriolis force acts perpendicular to the wind direction (to the right of the wind in the Northern Hemisphere), while the pressure-gradient force is directed down the gradient of geopotential height. Thus, geostrophic balance requires that the wind be parallel to the height contours with speed proportional to the height gradient at constant pressure. *See* CORIOLIS ACCELERATION; GEOSTROPHIC WIND.

Equations (8) and (1) may be combined to yield the thermal wind equation (9), where $\mathbf{V}_T =$

$$f\mathbf{V}_T = g\mathbf{k} \times \nabla(Z_2 - Z_1) = R\int_{p2}^{p_1} (\mathbf{k} \times \nabla T) d\ln p \tag{9}$$

$\mathbf{V}g(p_2) - \mathbf{V}g(p_1)$. Equation (9) states that the vector difference in the geostrophic wind velocity between two pressure surfaces is proportional to the horizontal gradient of the mean temperature in the layer between the two surfaces. This relationship, which is a consequence of geostrophic and hydrostatic balance, provides a strong constraint on the horizontal velocity and temperature distributions. For motions that are nearly geostrophic and nearly hydrostatic, the wind and temperature fields cannot vary independently, but they are closely coupled through their mutual relationship to the temperature (or density) distribution. *See* THERMAL WIND.

Vorticity and potential vorticity. Equation (4) indicates that momentum is not conserved on a fluid parcel, even when the flow is adiabatic and frictionless; its rate of change depends on the small difference between the Coriolis force and the pressure-gradient force. A dynamical quantity that is conserved following the motion under such conditions is the potential vorticity. Potential vorticity can be expressed in its simplest form by using potential temperature coordinates, as is shown in Eq. (10), where $\zeta = \mathbf{k} \cdot (\nabla \times \mathbf{V})$

$$P = -(\zeta + f)\frac{\partial \theta}{\partial p} \tag{10}$$

is the vertical component of the relative vorticity and f, the Coriolis parameter, is the vertical component of the Earth's vorticity.

Equation (10) indicates that the product of the absolute vorticity $(\zeta + f)$, which is a measure of the rotation of a fluid particle about a local vertical axis in inertial space, and the stability parameter $-\partial\theta/\partial p$, which is a measure of the vertical extension of the fluid parcel, is conserved following the motion. Potential vorticity conservation is the fluid dynamic analog of angular momentum conservation in solid body mechanics. As an air parcel moves from a region where it has weak static stability (that is, where

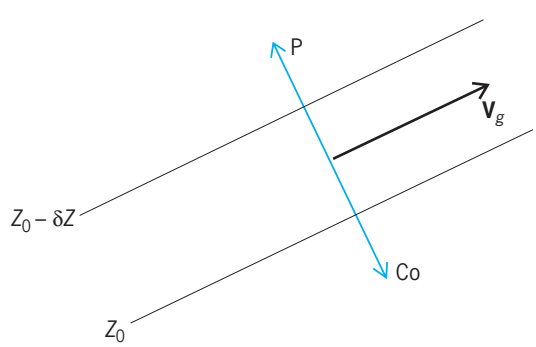

Fig. 1. Force balance for the geostrophic wind. P = pressure-gradient force. Co = Coriolis force. Z_0 and $Z_0 - \delta Z$ are constant-height contours on an isobaric surface. $\mathbf{V}_g$ = geostrophic wind.

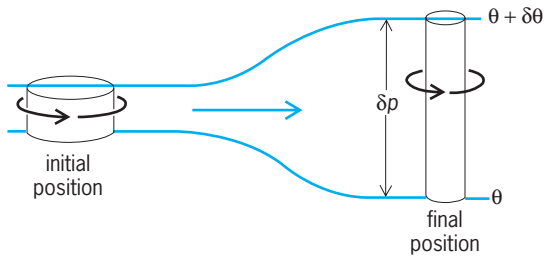

Fig. 2. Cylindrical column of air moving adiabatically, conserving potential vorticity. θ = potential temperature. p = pressure. (*After J. R. Holton, An Introduction to Dynamic Meteorology, 3d ed., Academic, 1992*)

$-\partial\theta/\partial p$ is small) to a region where its static stability is strong, the parcel must shrink in vertical extent and increase in horizontal area, with a resulting decrease in absolute vorticity (**Fig. 2**). It turns out that with specifications of suitable boundary conditions, the distribution of the large-scale wind and temperature fields in three-dimensional space can be uniquely determined from the distribution of potential vorticity. Since potential vorticity tends to be conserved following the motion, the evolution of the distribution of potential vorticity in time can be easily predicted and utilized to infer the evolution of wind and temperature. *See* DYNAMIC INSTABILITY.

Planetary boundary layer. Within the lowest kilometer of the atmosphere, there is a boundary layer in which the frictional force due to small-scale turbulent motions represents an important contribution to the force balance. The structure of the atmospheric boundary layer depends strongly on the vertical stratification in the layer. In regions of strong static stability (where potential temperature increases rapidly with altitude), turbulence is suppressed and the frictional force is weak except in the lowest several meters above ground, where strong vertical shear of the wind provides a mechanical source for turbulence. When there is strong surface heating (as over deserts) the potential temperature may actually decrease with height near the ground. Boundary-layer turbulence is then convectively driven, and the frictional force may be strong throughout at least the lowest kilometer of the atmosphere. Over vast areas of the Earth's surface, the boundary layer is typically near neutral stability. In that case, it is permissible to approximate the frictional force as a drag proportional to the wind velocity: $\mathbf{Fr} = -k\mathbf{V}$, where k^{-1} is the time scale for friction to deplete the velocity by a factor of e^{-1}. Typically, $k \simeq 10^{-5}$ per second, and the force balance in the boundary layer is a three-way balance among the friction force, Coriolis force, and pressure-gradient force.

This balance is shown schematically in **Fig. 3**. Since the Coriolis force always acts to the right of the wind direction in the Northern Hemisphere, and the frictional force is opposite to the wind direction, a balance can be achieved only if the wind has a component directed across the pressure field toward lower pressure. This component produces a net frictional convergence into surface low-pressure systems, and divergence out of surface highs. *See* CONVECTIVE INSTABILITY.

Atmospheric waves. The dynamics equations (4)–(6) are capable of describing a wide variety of systems of atmospheric motion. In order to classify and analyze these, it is useful to divide the horizontal velocity into rotational and irrotational parts. The circulation in large-scale weather systems is dominated by the rotational part of the wind; outside the tropics this can be approximated by the geostrophic wind. The irrotational part of the velocity field dominates in small-scale atmospheric motions in which the divergence of the horizontal wind plays a crucial role. Although many of the most important features are inherently nonlinear, there are several types of wave solutions that can be analyzed by using linearized versions of the dynamics.

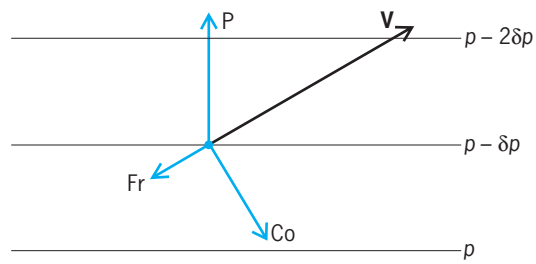

Fig. 3. Balance of forces in a well-mixed planetary boundary layer. P = pressure-gradient force. Co = Coriolis force. Fr = frictional force. p = pressure. (*After J. R. Holton, An Introduction to Dynamic Meteorology, 3d ed., Academic, 1992*)

Buoyancy and inertial waves. Waves in fluids are motions in which information can be transferred (usually by pressure perturbations) without material transport of air parcels. In all cases, atmospheric waves result from a balance between the fluid inertia and some restoring force. For example, the vertical stratification of the atmosphere causes a parcel displaced vertically to experience a restoring force due to its buoyancy; the resulting oscillation, which is dominated by the divergent part of the wind field, is referred to as a buoyancy (or internal gravity) wave. In a somewhat analogous fashion, the lateral deflection due to the Coriolis force causes a parcel displaced horizontally to experience a lateral restoring force; the resulting oscillation is known as an inertial wave. Waves in which both buoyancy and Coriolis forces are important are known as inertia-gravity waves.

The longitude–height structure of an eastward propagating buoyancy wave is shown in **Fig. 4**. Phase lines tilt toward the east with height, so that as the wave passes a particular location there is an apparent downward propagation in time. However, the energy flux in this case is actually directed upward, parallel to lines of constant phase. Buoyancy waves provide an important mechanism for transferring momentum from the lower atmosphere into the upper atmosphere. The buoyancy wave frequency ν is given by Eq. (11), where $N = (g\partial \ln \theta/\partial Z)^{1/2}$ and

$$\nu = N \cos\alpha \tag{11}$$

α is the angle of the phase lines to the local vertical. The maximum possible frequency N thus occurs for vertically oriented phase lines. In the lower atmosphere, N typically corresponds to a period of about 10 min. Only those waves in which the ratio of horizontal to vertical scale is very large have frequencies sufficiently low to be of much meteorological importance.

An important class of low-frequency buoyancy waves comprises lee waves, formed when air is forced to flow over a mountain barrier. The oscillations that are excited by such forcing may continue many wavelengths downstream of the mountain barrier. When sufficient moisture is present, condensation may occur in the updraft portions of the waves and regularly spaced bands of cloud may form. In this case, the waves are stationary with respect to the ground, but propagate relative to an observer moving with the mean wind. *See* MOUNTAIN METEOROLOGY.

Planetary waves. The most important class of large-scale atmospheric waves comprises planetary waves. They are characterized by oscillations in the rotational part of the horizontal wind. In these waves the latitudinal gradient of the mean atmospheric potential vorticity provides the restoring force that acts to oppose meridional parcel displacements. The simplest example of a planetary wave occurs in a barotropic atmosphere (an atmosphere in which potential temperature is a function of pressure alone). In a barotropic atmosphere, conservation of potential vorticity [see Eq. (10)] reduces to conservation of absolute vorticity following the motion described in Eq. (12).

$$\frac{d(\zeta + f)}{dt} = 0 \tag{12}$$

If a chain of air parcels extending around a latitude circle with $\zeta = 0$ at time t_0 is given a small meridional displacement δy, then from application of the conservation law in Eq. (12), at time t_1 the relative vorticity will be as shown in Eq. (13), where $\beta = df/dy$,

$$\zeta_{t_1} = f_{t_1} - f_{t_0} = -\beta\ \delta y \tag{13}$$

the rate of increase of the Coriolis parameter with latitude. From Eq. (13) it is evident that the relative vorticity perturbation will be positive for a southward displacement and negative for a northward displacement. Since positive (negative) relative vorticity corresponds to counterclockwise (clockwise) rotation, the induced perturbation velocities will tend to move the displacement pattern toward the west, as illustrated in **Fig. 5**. Thus, the longitudinal propagation of planetary waves is always westward relative to the mean wind. Such waves are transverse in the sense that the fluid parcel oscillations, which are in the north-south direction, are perpendicular to the direction of wave propagation.

Quantitative analysis reveals that planetary waves are highly dispersive; that is, their longitudinal phase speeds are strongly dependent on wavelength. In cartesian coordinates this dependence, in the presence of a mean zonal flow $\overline{u}$, has the form shown in Eq. (14), where $c_x - \overline{u}$ is the phase speed relative

$$c_x - \overline{u} = -\frac{\beta}{4\pi^2}\left[\frac{1}{L_x^2} + \frac{1}{L_y^2}\right]^{-1} \tag{14}$$

to the mean flow, and L_x and L_y are the wavelengths in longitude and latitude, respectively. Equation (14) shows that planetary waves always propagate westward relative to the mean flow, with phase speed proportional to the square of the wavelength. For waves corresponding to typical extratropical storms with wavelengths of a few thousand kilometers, the planetary wave phase speed is typically less than 10 m · s^{-1} (33 ft · s^{-1}), which is less than the mean wind speed. Thus, relative to the ground, such

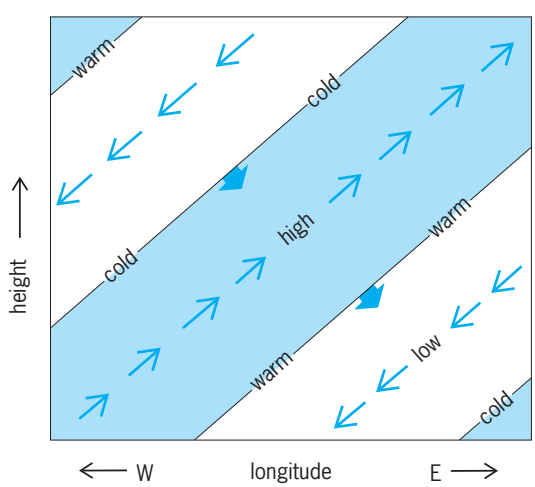

Fig. 4. Longitude–height structure of an eastward-propagating buoyancy wave. Thin arrows indicate the perturbation velocity field; heavy arrows indicate the phase velocity. The labels "high" and "low" and "warm" and "cold" indicate the axes of maximum perturbation pressure and temperature, respectively.

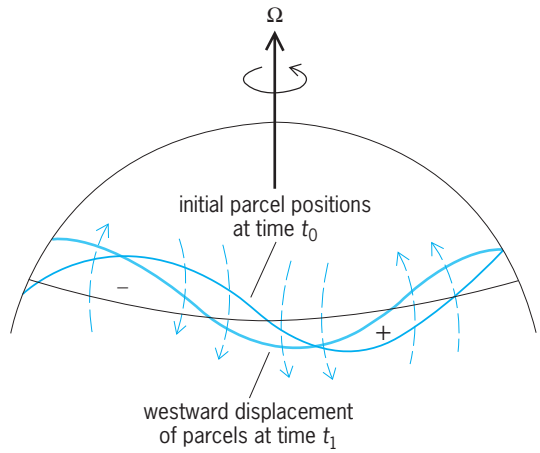

Fig. 5. Propagation mechanism for a planetary wave. Perturbation vorticity is shown by + and −, induced velocity field is shown by broken arrows. Ω = angular velocity of rotation of the Earth. (*After J. R. Holton, An Introduction to Dynamic Meteorology, 3d ed., Academic, 1992*)

disturbances tend to move eastward, but at a speed less than the average eastward wind speed.

Equation (14) also indicates that planetary waves that are stationary relative to the ground ($c_x = 0$) can exist only when the mean zonal flow is positive (that is, from the west). Thus, stationary waves are excited in midlatitudes when westerly winds encounter mountain barriers. It is this process that accounts for the lee side trough observed to the east of the Rocky Mountains.

Baroclinic instability. The barotropic planetary waves described above cannot represent any motions that involve conversion of energy between potential and kinetic forms. Such energy conversion can occur only in the presence of baroclinity, that is, variations of potential temperature on isobaric surfaces. Baroclinic energy conversion processes are responsible for the growth and maintenance of most large-scale weather disturbances.

In midlatitudes in the troposphere, potential temperature normally decreases from the Equator to the pole on isobaric surfaces. This decrease does not occur uniformly, but tends to be concentrated in the jet stream, a narrow band of strong westerly winds in the upper troposphere that encircles the globe in midlatitudes. The association of the jet stream with a strong meridional temperature gradient is a consequence of the thermal wind relationship, Eq. (9), which states that westerly winds must increase rapidly with height where the temperature decreases rapidly in latitude. *See* JET STREAM.

When the shear of the zonal wind is sufficiently strong that the meridional gradient of potential vorticity on a constant potential temperature surface is locally reversed, or when there is a nonvanishing gradient of potential temperature at the surface of the Earth, the linearized equations have solutions in the form of exponentially growing disturbances. These baroclinically unstable modes have growth rates, structures, and scales similar to those observed in developing extratropical cyclones. They are quasigeostrophic in the sense that the geostrophically balanced rotational component of the velocity field strongly dominates over the divergent component. The latter is, however, crucial in the energy cycle through which the disturbances convert potential energy stored in the meridional gradient of potential temperature into disturbance kinetic energy. Baroclinic instability provides a mode whereby infinitesimal disturbances may amplify into large-amplitude storms. However, in many cases it appears that storms in the atmosphere grow through nonlinear interactions involving preexisting disturbances. Such processes must generally be studied by numerical simulations on high-speed computers.

Mesoscale convective systems. For horizontal scales less than several hundred kilometers, the major energy source is not baroclinic instability; it is latent heat release by cumulonimbus clouds. The convective storms associated with such clouds can occur only when the atmosphere is conditionally unstable [Eq. (2)], sufficient moisture is present, and there is an initial disturbance strong enough to lift air parcels high enough to release the conditional instability. Mesoscale convective systems take a variety of forms. Among these are hurricanes, squall lines, and mesoscale convective complexes. In all of these systems the release of latent heat by convective clouds is the primary energy source, but the character of the large-scale environmental flow is generally important for determining the mode of organization for the convection. There is no single theory, analogous to the theory of baroclinic instability, that can explain the origin and growth of mesoscale convective systems. Most studies of such systems have involved numerical modeling of particular cases. *See* HURRICANE; MESOMETEOROLOGY; SQUALL LINE.

Numerical weather prediction. The dependent variables in the dynamics equations (4)–(6) are a set of fields that are continuous functions of space and time. In order to solve these equations numerically, it is necessary to approximate the various fields, such as velocity, temperature, and pressure, either by their values on a finite set of grid points in three-dimensional space (the finite difference method) or by truncated series of orthogonal functions (the spectral method). In current global weather prediction models, a mixture of techniques is often employed. Nearly all models use finite difference representations for the vertical coordinate and time; the horizontal variation is generally represented either by a network of grid points at uniform intervals of latitude and longitude or by a finite set of spherical harmonics. In the latter representation, a field, such as temperature, at a given pressure and time is given by a summation of the form shown in Eq. (15),

$$T(\lambda, \phi, p, t) = \sum_{m=-J}^{+J} \sum_{n=|m|}^{|m|+J} T_n^m(p, t) P_n^m(\sin\phi) \exp(im\lambda) \qquad (15)$$

where λ and ϕ denote longitude and latitude, respectively, and $P_n^m(\sin\phi)$ designates the associated Legendre function of the first kind. For a given spectral mode, the integer index m represents the number of waves around a latitude circle, while $n - m$ is the number of waves between the north and south poles, and J is the total number of modes retained in the expansion. The spectral method generally provides higher accuracy, especially in predicting the speed of motion of short wave features, than can be obtained with a finite difference representation of similar resolution. *See* LEGENDRE FUNCTIONS; SPHERICAL HARMONICS.

Because the lower boundary of the atmosphere does not correspond to an isobaric surface, pressure is not a very convenient vertical coordinate for numerical prediction models. In practice, therefore, pressure is usually replaced by the coordinate $\sigma = p/p_s$, in which pressure is normalized by the local surface pressure p_s. Typically, about 10–15 levels are specified in the interval $\sigma = 1$ to $\sigma = 0$, and the horizontal resolution is of order 3–5° in longitude and latitude.

To predict the weather dynamically, it is necessary to solve the dynamics equations by integrating in time, starting from an initial state of the atmospheric variables determined from observations. In practice, observational errors and poor data coverage, particularly over the oceans, make it impossible to exactly determine the initial state of the atmosphere. Moreover, the state determined from observations generally does not properly represent the true dynamical balance among the pressure and wind fields. This imbalance introduces noise that is interpreted as high-frequency inertia-gravity waves. If a prediction is attempted from such an initial state, the noise rapidly dominates the true, slowly evolving weather disturbances. In order to prevent the growth of such spurious noise, the analysis must be initialized by processing it in a manner that assures a dynamical balance between the pressure and wind fields while preserving the actual observations as closely as possible.

However, even if the initial state were known perfectly, there still would be a limit beyond which errors would dominate the forecast. This predictability limit arises because the atmosphere has a continuous spectrum of scales of motion. There are inevitably scales too small to be resolved by the model, and through nonlinear interactions these would eventually come to dominate the forecast. Thus, the theoretical length of useful forecasts for transient weather disturbances is about 3 weeks. The predictive skill of present models is still far short of this limit. *See* WEATHER FORECASTING AND PREDICTION.

Global climate modeling. In addition to their role in weather prediction, dynamical models can also be used to simulate global climate. Climate is the study of the average state of the atmosphere and its seasonal and interannual variability. Climate is determined by the joint influence of energy sources and sinks at the Earth's surface, and the transformation and transport of energy in the atmosphere and the oceans. Models that simulate these processes are usually referred to as general circulation models. Such models must contain accurate representations of all the important physical processes that influence the circulation.

Among these processes are heating and cooling by solar and infrared radiation; vertical transfer of heat and momentum by convective clouds; latent heating due to evaporation, condensation, and freezing of water; and drag and mixing due to small-scale turbulent motions in the boundary layer. In most cases, these processes involve scales of motion smaller than those explicitly resolved by the global model. They must then be parametrized in terms of the resolved variables. It is often not clear how to accomplish this objective. Thus, parametrization of subgrid-scale fields is an area of active research for global modeling. *See* ATMOSPHERE; ATMOSPHERIC GENERAL CIRCULATION; CLIMATE MODELING; FLUID-FLOW PRINCIPLES; METEOROLOGY; WIND. James Holton

Bibliography. A. E. Gill, *Atmosphere-Ocean Dynamics*, 1982; J. R. Holton, *An Introduction to Dynamic Meteorology*, 3d ed., 1992; J. T. Houghton, *The Physics of Atmospheres*, 2d ed., 1986; F. Kahutgens and E. J. Tarbuck, *Atmosphere: An Introduction to Meteorology*, 5th ed., 1991; J. Pedlosky, *Geophysical Fluid Dynamics*, 2d ed., 1993; J. M. Wallace and P. V. Hobbs, *Atmospheric Sciences: An Introductory Survey*, 1977.

Dynamic nuclear polarization

The creation of assemblies of nuclei whose spin axes are not oriented at random, and which are in a steady state that is not a state of thermal equilibrium. Under commonly occurring conditions, the spin axes of nuclei (with nonzero spin) are oriented at random; where this is not so, the nuclei are said to be polarized. Assemblies of polarized nuclei are not in a state of thermal equilibrium except under rather extreme conditions (for example, temperatures below 10 millikelvins or 0.02°F above absolute zero, and magnetic fields greater than several teslas), and schemes have been devised to produce polarized assemblies, in a steady state which is not a state of thermal equilibrium, under less extreme conditions. Such schemes constitute dynamic nuclear polarization.

Among the many applications of polarized nuclei are the following. Nuclear forces are spin-dependent, and although the spin-dependent part can be found by using unpolarized assemblies, the experiments are simpler and their interpretation is clearer if polarized nuclei are used. Assemblies of polarized nuclei have a lower geometrical symmetry than assemblies of randomly oriented nuclei, and so these have been used to investigate the fundamental symmetries of nature. Polarized nuclei have been used to enhance the signal in free precession magnetometers and similar instruments, and the use of an assembly of polarized nuclei as a gyroscope has also been suggested. *See* MAGNETOMETER; NUCLEAR ORIENTATION; PARITY (QUANTUM MECHANICS); SPIN (QUANTUM MECHANICS).

The simplest and most common parameter used to define the state of polarization is called the "polarization" or vector polarization $\mathbf{P}$. It is a vector whose component along any given axis (for example, the $z$axis) is the average value of the projection on that axis of a unit vector parallel to the nuclear spin axis. Often, nuclei lie with their spins pointing along the $z$axis in equal numbers in both directions, but none lies in the x-y plane. Such a system is not random, but $\mathbf{P} = 0$. The parameter used to quantify this arrangement is the average value of $(^3/_2 \cos^2 \Theta - {}^1/_2)$ and is one component of the so-called tensor polarization. In general, the average value of any zonal harmonic is a measure of a kind of polarization. It is a property of nuclear spin operators that harmonics of order greater than $2J$ vanish identically, where J is the nuclear spin. In particular, spin 0 nuclei cannot be polarized and spin $^1/_2$ nuclei can have a vector polarization only. *See* ANGULAR MOMENTUM; SPHERICAL HARMONICS.

There are, in general, two common features of all dynamic nuclear polarization schemes. There is

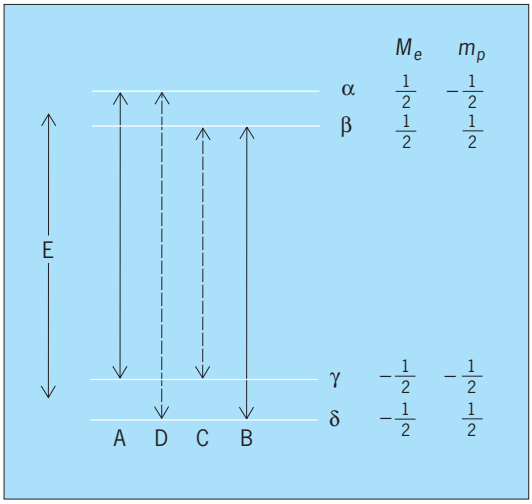

Fig. 1. Energy levels of an electron and a proton in a magnetic field.

some external system or interaction, for example, an oscillating magnetic field, which disturbs the system; this is called the pumping mechanism. There are also processes which tend to restore the system to thermal equilibrium; these are called relaxation processes. However, the functions of these two processes are not easily distinguished, for, in many schemes, the relaxation mechanisms actually play a part in setting up the nuclear polarization, this being one of the stages back to thermal equilibrium.

Microwave pumping. Consider an electron spin and a nucleus of spin $^1/_2$ (for example, a proton) in close association. Typically, the electron spin is part of a paramagnetic ion, and the proton is a hydrogen nucleus in a nearby water molecule. In a magnetic field of about 1 T, the energy levels of this combination are as shown in **Fig. 1**. M_e and m_p are, respectively, the electron and proton spin magnetic quantum numbers which specify whether the electron and proton spins are oriented in the same direction as the magnetic field or in the opposite direction. An oscillating magnetic field is applied to induce transitions between pairs of levels. The two transitions A and B, in which only an electron spin is reversed, constitute the "allowed" transition. The two weaker transitions, C and D, in which a nuclear spin is reversed, occurring one at a higher and the other at a lower field (or frequency) than the allowed transition, are called forbidden transitions. *See* MAGNETIC RESONANCE; SELECTION RULES (PHYSICS).

In the most common and most successful mechanism of this type for polarizing protons, one of the forbidden transitions (C, for instance) is excited strongly enough to saturate it. Then the populations of levels β and γ are equal, and are kept equal by the oscillating magnetic field. The strongest relaxation processes, which try to bring the system into thermal equilibrium, are associated with the strongest transitions A and B. As a result of these processes, fewer systems occupy level α than level γ, and similarly, fewer systems occupy level β than level δ. Thus more proton spins are oriented in one direction ($m_p = +^1/_2$) than in the opposite direction ($m_p = -^1/_2$), and a nuclear polarization results. By saturating the other forbidden transition, an equal and opposite polarization results. *See* MAGNETIC RELAXATION.

In another method the allowed transition is saturated, and so the populations of levels α and γ are equal, and similarly the populations of levels β and δ are equal. The strong relaxation processes, which involve only the inversion of an electron spin, are ineffective since they are swamped by the radiofrequency field. The "cross relaxation" processes, which connect level α with level δ (called the flip-flop relaxation), and which connect level β with level γ (called the flip-flip relaxation), are of different strengths. This is because in the flip-flop transition, spin angular momentum is conserved, but not so in the flip-flip transition. The relative populations of the levels are thus determined by the relative strengths of the two processes. If, as in the case of metals, the flip-flip transition can be neglected, the flip-flop process results in fewer systems occupying level α than level δ, and a nuclear polarization results. If the electron spin and the nuclear spin are coupled by magnetic dipole interaction, and direct nuclear spin lattice relaxation can be neglected, the flip-flip process is twice as strong as the flip-flop process, and a smaller nuclear polarization of the opposite sign results.

Although this explanation describes very effectively what happens, it is oversimplified. Nuclei do not interact in turn with individual electron spins, and the latter are not independent. Especially when protons are polarized in organic materials, the effect of the oscillating magnetic field is to cool the assembly of interacting electron spins, which then cool the nuclei by thermal contact. These processes are known by the acronym DONKEY effects. *See* PARAMAGNETISM.

In order to produce large polarizations, the external magnetic field should be of the order of 2 T, the frequency of the oscillating field should be about 70 GHz, and the temperature should be 1 K (2°F above absolute zero) or less. Values of polarization of 0.8 or more (corresponding to 90% of the spins oriented in one direction and only 10% in the opposite direction) are regularly obtainable for protons polarized in $La_2Mg_3(NO_3)_{12} \cdot 24H_2O$ in which 1% of the lanthanum, La, is replaced by neodymium, Nd; in an empirically determined mixture consisting mainly of butanol, C_4H_9OH, and about 1% of porphyrexide, a free radical; or in ethylene glycol or propanendiol with about 1% of an organic Cr^V complex. Deuteron polarizations of 0.4 have been obtained by using C_4D_9OD and porphyrexide at temperatures of 0.3 K (0.5°F above absolute zero). *See* COORDINATION CHEMISTRY; COORDINATION COMPLEXES.

Butanol and diol systems are favored to make targets of polarized protons for experiments in high-energy physics. The beam passing through the target causes radiation damage which adversely affects thepolarization. To a first approximation, the

polarization P falls off as an exponential function of the dose ϕ, given by the equation below, where

$$P = P_0 \exp(-\phi/\phi_0)$$

P_0 is the initial polarization and ϕ_0 is approximately 2.5×10^{15} minimum ionizing particles per square inch (4×10^{14} minimum ionizing particles per square centimeter); but this levels off to a constant value of approximately $0.7P_0$. This effect is caused by the production of free radicals, and free electrons which are trapped to form color centers, both of which are paramagnetic. Most of the radiation damage can be annealed out by raising the temperature of the target material to about −208°F (140 K) for butanol, or −136°F (180 K) for diols, and keeping it there for about 15 min. *See* COLOR CENTERS; FREE RADICAL.

Polarization by rotation. This method was most spectacularly demonstrated by rotating a crystal of yttrium ethylsulfate, $Y(C_2H_5SO_4)_3 \cdot 9H_2O$, in which a few percent of yttrium, Y, was replaced by ytterbium, Yb, in a magnetic field of 1 T at a temperature of 1.4 K (2.5°F above absolute zero), at a rate of 60 revolutions per second. The proton polarization was found to be 0.19 (corresponding to 59.5% of the spins oriented in one direction and 40.5% in the other direction); higher values can be obtained by modifying the experimental conditions. The splitting of the two energy levels of the protons in the external magnetic field is independent of the orientation of the crystal, but the splitting of the two energy levels of the Yb^{3+} ion depends strongly on the orientation of the magnetic field relative to the c axis of the crystal, being quite large when these are parallel and almost zero when they are perpendicular. The Yb^{3+} ions are strongly coupled to the lattice when this angle is 45°, but almost isolated when it is 0 or 90°. Thus, as the crystal is rotated, when it reaches the 45° position, the Yb^{3+} ions quickly come to equilibrium at the temperature of the lattice. Then, as the crystal is rotated to the 90° position, the energy levels of the Yb^{3+} ions move together; this is equivalent to adiabatic demagnetization, and the spin temperature of the Yb^{3+} ions drops. Near the 90° position, the Yb^{3+}ions can exchange energy with the protons, because the two splittings are the same, and a flip-flop transition involving a Yb^{3+} ion and a proton conserves energy. The protons and the Yb^{3+} ions thus come to thermal equilibrium at the temperature of the Yb^{3+}ions after demagnetization. Continuous rotation repeats the cycle, and the ideal steady state is one in which the proton polarization is the same as the polarization of the Yb^{3+} spins at the 45° position. *See* ADIABATIC DEMAGNETIZATION; MAGNETIC SUSCEPTIBILITY.

Optical pumping. This method, pioneered by Alfred Kastler and Jean Brossel, makes use of the fact that circularly polarized light carries angular momentum, and when it is absorbed, that angular momentum is given to the absorber. Typically circularly polarized resonance radiation (such as sodium-D radiation) is incident on, for example, sodium vapor at room temperature (at a pressure of about 10^{-6} torr or 10^{-8} lb/in.^{2} or 10^{-4} pascal) in a small magnetic "guide" field directed along the direction of the light beam. It excites only the σ^+ transitions in which the magnetic quantum number of the absorbing atom, which specifies its angular momentum along the direction of the magnetic field, increases by 1. When the excited state decays, on the average there is no change in the magnetic quantum number. Thus an equilibrium is set up in which the ground-state atom spins are polarized. Polarizations of almost 1.0 can easily be obtained for the alkali metals. *See* POLARIZED LIGHT.

The absorbing atoms are usually mixed with a gas such as argon, called a buffer gas, at a pressure of 1–100 torr (0.01–1 lb/in.^{2} or 10^2–10^4 Pa) to delay their diffusion to the walls of the container where they may be depolarized. If some other gas is present instead of, or in addition to, the buffer gas, the angular momentum of the absorber is transferred in collisions to the atoms of this gas and can end up as a nuclear polarization.

This method has been very successfully applied to^{3}He, which acts as a buffer gas. The absorbing atoms are metastable 2^3S_1 ^{3}He atoms produced by striking a weak electrodeless discharge in the ^{3}He gas. Polarizations of up to 0.4 have been achieved; under suitable conditions, when the polarizing mechanism is turned off, the polarization decays with a time constant of the order of several days. *See* OPTICAL PUMPING. James M. Daniels

Laser-induced polarization. Nuclear orientation through optical pumping can also be achieved by using lasers. This method has been refined with the availability of high-power tunable continuous-wave dye lasers. The notion of a preferred direction in space is passed on to an ensemble of randomly oriented nuclei through atomic transitions induced by a polarized laser beam.

The energy levels in an atom exhibit fine and hyperfine structure. The hyperfine structure arises through the interaction of the magnetic and quadrupole moment of the nucleus with the atomic electrons. The hyperfine energy levels are classified in terms of the total angular momentum of the electron-nuclear system $\vec{F} \cdot \vec{F} = \vec{I} + \vec{J}$, where $\vec{I}$ and $\vec{J}$ are the electron and nuclear angular momenta respectively. $\vec{F}$ can take the smaller of $2I + 1$ or $2J + 1$ values. The projection of $\vec{F}$ on an arbitrary (quantization) axis, labeled M_F, can take on the values in the range $-F$ to F in integral steps. In the absence of electromagnetic fields, the various M_F levels associated with a given$F$ have the same energy and, under normal conditions in a large ensemble of atoms, are equally likely to be populated. Optical transitions can be induced by using a polarized laser to alter the random distribution of population among the M_F states; this constitutes nuclear orientation. The transitions are governed by selection rules. For electric dipole transitions between states of opposite parity, F can change either by $+1$, -1, or 0 (transitions between a state $F = 0$ and another $F' = 0$ are not allowed). M_F can change by $+1$, -1, or 0. The M_F selection rule

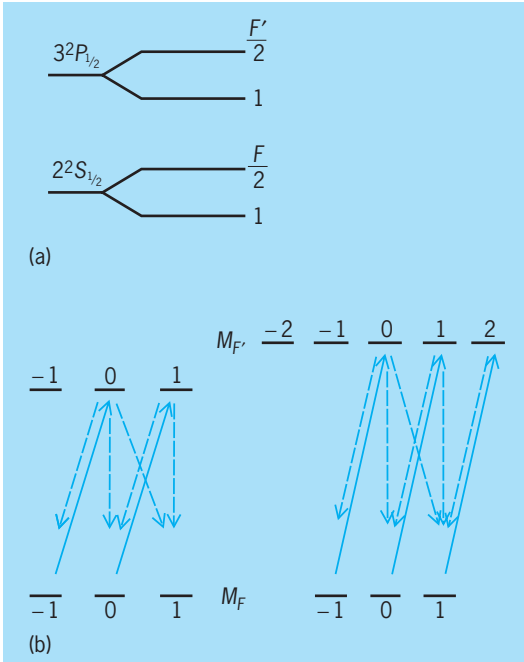

Fig. 2. Optical pumping of sodium-23. (*a*) Fine and hyperfine levels in the ground and first excited state of ^{23}Na. (*b*) An optical pumping cycle showing the allowed transitions, for excitation with a σ^+ polarized laser beam, from and to the $F = 1$ hyperfine level in the ground state. For the excitation, M_F can change by only $+1$, and for the decay, M_F can change by $+1$, -1, or 0.

depends further on the polarization of the laser light. *See* HYPERFINE STRUCTURE; NUCLEAR MOMENTS.

As an example, consider the optical pumping of a sodium-23 (^{23}Na) atom using σ^+ (circularly) polarized laser light. The quantization axis is chosen as the direction of propagation of the laser beam. For the ground state of ^{23}Na, $I = {}^3/_2$, $J = {}^1/_2$, and $F = 1, 2$. For the first excited state of opposite parity, $I = {}^3/_2, J = {}^1/_2$, and $F' = 1, 2$. These levels are shown in **Fig. 2***a*. The allowed transitions ($\Delta M_F = 1$) from the $F = 1$ ground state to the $F' = 1, 2$ excited states are shown in Fig. 2*b* as solid arrows. Transitions from and to the $F = 2$ level have been omitted for clarity. The spontaneous decay paths from the excited states have been shown as broken arrows. This process of excitation and decay constitutes an optical pumping cycle. A cursory inspection of Fig. 2*b* reveals that transitions populating the $M_F = 1$ level far outnumber those depleting it. After relatively few such pumping cycles the population in the atomic ensemble is found predominantly in the higher M_F levels. Under these conditions the spins of most of the nuclei are oriented in the direction of the laser beam. Such a technique has been used successfully to produce several species of nuclei with essentially all of the nuclear spins oriented in the same direction. Nuclear orientation can also be achieved by using a linearly polarized laser beam.

The extent to which nuclei can be oriented in this manner is limited by several factors. Some of the factors limiting the efficiency of an optical pumping cycle are: how well the laser bandwidth is matched to the absorption profile of the transitions being pumped; the oscillator strengths of the atomic transitions; the length of the laser-atom interaction time; and excited and ground-state relaxation through atom-atom and atom-wall collisions and through precession in stray magnetic fields. The thrust of research in the field of laser-induced nuclear orientation has been in overcoming these difficulties. *See* ATOMIC STRUCTURE AND SPECTRA; LASER; ZEEMAN EFFECT.

B. Shivakumar

Bibliography. A. O. Barnt et al., *Polarization Dynamics in Nuclear and Particle Physics*, 1993; C. E. Bemis, Jr., and H. K. Carter (eds.), *Lasers in Nuclear Physics*, Nuclear Science Research Conference Series, vol. 3, 1982; D. E. Murnick and M. S. Feld, Applications of lasers to nuclear physics, *Annu. Rev. Nucl. Part. Sci.*, 29:411–454, 1979; J. Sowinski and S. E. Vigdor (eds.), *Physics with Polarized Beams on Polarized Targets*, 1990; W. J. Thompson and T. B. Clegg, Physics with polarized nuclei, *Phys. Today*, 32(2):32–39, 1979.

Dynamic similarity

A relationship existing between two fluid flows when they have identical types of forces that are parallel at all corresponding points, with magnitudes related by a constant scale factor. Dynamic similarity makes it possible to scale results from model tests to predict corresponding results for the full-scale prototype.

Dynamic similarity requires faithful reproduction of detail on the model (geometric similarity); the same flow pattern, including boundary shapes (kinematic similarity); and test conditions that match relevant dimensionless ratios between model and prototype. Dynamically similar flows are said to be homologous.

A rigorous method to find the dimensionless ratios is to nondimensionalize the governing equations and boundary conditions. The dimensionless coefficients are duplicated to obtain dynamic similarity. A practical method is to list the parameters and form dimensionless ratios by dimensional analysis.

Dimensionless parameters simplify presentation of experimental results. An example is the pressure coefficient given by the equation below, where p,

$$C_p = \frac{(p - p_{\text{ref}})}{{}^1/_2 \rho V^2}$$

p_{ref}, ρ, and V are local pressure, reference pressure, density, and velocity, respectively. Corresponding points in dynamically similar flows have equal pressure coefficients even when fluid density or speed differs. Hundreds of dimensionless parameters have been derived and named. A few parameters are in especially frequent use (see **table**), and they may be combined to obtain other parameters.

It may not be possible in a practical test to match all dimensionless parameters. It is most important to

Dynamic similarity parameters

Nondimensional parameter*	Name	Physical effect
$\rho VL/\mu$	Reynolds number	Viscosity
V/c	Mach number	Compressibility
V^2/Lg	Froude number	Gravity
x_m/L	Knudsen number	Continuum
$\rho V^2 L/\sigma$	Weber number	Surface tension
$c_p\mu/\kappa$	Prandtl number	Heat conduction
$\beta TgL^3\rho^2/\mu^2$	Grashof number	Free convection

*The reference variables are defined as follows; L, length; μ, coefficient of viscosity; c, speed of sound; g, acceleration of gravity; σ, surface tension; κ, coefficient of thermal conductivity; β, coefficient of thermal expansion; T, temperature; x_m, mean free path of molecule; and c_p, specific heat at constant pressure.

match parameters that represent the dominant physical effects. Thus, correct simulation of viscous effects requires that Reynolds number be matched; Mach number may be ignored if compressibility effects are not important. In ship model tests, Froude number must be matched to duplicate wave patterns; the effect of Reynolds number on viscous drag may be predicted analytically. *See* DIMENSIONAL ANALYSIS; DIMENSIONLESS GROUPS; FLUID MECHANICS; FROUDE NUMBER; MACH NUMBER; MODEL THEORY; REYNOLDS NUMBER.

Alan T. McDonald

Bibliography. R. W. Fox and A. T. McDonald, *Introduction to Fluid Mechanics*, 4th ed., 1992; S. J. Kline, *Similitude and Approximation Theory*, 1965, reprint 1986; L. I. Sedov, *Similarity and Dimensional Methods in Mechanics*, 10th ed., 1993.

Dynamical time

The independent variable in the theories and ephemerides of solar system bodies. Dynamical time scales are appropriate for coordinate systems in the solar system. *See* EPHEMERIS.

Types of time scales. There are different types of time scales in existence. Dynamical time scales are the appropriate coordinate time scales for the barycenter (center of mass) of the solar system, the geocenter (center of mass of the Earth), or planetary coordinate times. The different dynamical time scales are related by equations based on the gravitational potentials and the theory of relativity.

The most accurate time scales are based on physics, such as atomic time scales (with a second defined as the duration of 9,192,631,770 periods of radiation of a hyperfine transition of cesium-133), and International Atomic Time (TAI), an average atomic time coordinated by the Bureau International des Poids et Mesures in France. TAI is an atomic time for the Earth geiod, a surface of constant gravitational potential that is an extension of the surface of mean sea level of the Earth's oceans, and is the basis for many time scales. Universal Time is based on the rotation of the Earth and the motion of the Sun. Coordinated Universal Time (UTC) is based on TAI and differs from UT1, the observed Earth rotation time, by less that 0.9 s. The difference is kept less than 0.9 s by inserting leap seconds when necessary at the end of June or December. Civil time scales differ from UTC according to time zones mostly by hour increments. *See* ATOMIC TIME.

Terrestrial Time (TT) is a theoretically ideal proper time. A practical realization is TT = TAI + 32.184 s. The quantity $\Delta T = \mathrm{TT} - \mathrm{UT1}$ is the difference between this ideal time scale and the rotation of the Earth. Before 1991, TT was known as Terrestrial Dynamical Time (TDT), a successor of prerelativistic Ephemeris Time (ET).

Geocentric Coordinate Time (TCG) is coordinate time for the geocenter. It is related to TT by Eq. (1),

$$\mathrm{TCG} = \mathrm{TT} + L_G \times (\mathrm{JD} - 2{,}443{,}144.5) \times 86{,}400\,\mathrm{s} \tag{1}$$

where JD is the Julian date, a widely accepted way of counting days consecutively, and $L_G \equiv 6.969290134 \times 10^{-10}$, a scale constant accounting for the Earth's gravitational and rotational potential affecting the rate of clocks operating at the geoid versus the geocenter. *See* CALENDAR; YEAR.

Barycentric Coordinate Time (TCB) is the coordinate time for the barycenter of the solar system. It differs from TCG by both secular and periodic terms. The complicated relationships between TT, TCG, and TCB are based on relativistic Lorentz transformations, although an approximation of (TCB − TCG) can be expressed as Eq. (2). Here TT_0 corresponds to JD 2,443,144.5 (1977 January 1, 0 h) TAI, the values of L_C and L_B are standard constants, $\mathbf{v}_e$ is the coordinate velocity of the Earth, $\mathbf{x}_e$ is the coordinate of the Earth's center of mass, and $\mathbf{x}$ is the coordinate of an event in barycentric reference coordinates. Periodic terms, denoted by $P(\mathrm{TT})$, have a maximum amplitude of around 1.6 m. *See* LORENTZ TRANSFORMATIONS; RELATIVITY.

$$(\mathrm{TCB} - \mathrm{TCG}) = \frac{L_C \times (\mathrm{TT} - \mathrm{TT}_0) + P(\mathrm{TT}) - P(\mathrm{TT}_0)}{(1 - L_B)} + \frac{\mathbf{v}_e \cdot (\mathbf{x} - \mathbf{x}_e)}{c^2} \tag{2}$$

Barycentric Dynamical Time (TDB) was originally intended to serve as an independent argument of barycentric ephemerides and equations of motion, and was defined by an International Astronomical Union 1976 resolution to differ from TT only by periodic terms. Later, it became clear that this condition cannot be rigorously satisfied. Operationally, each ephemeris realization defines its own version of TDB; the linear drift between TDB and TCB is usually chosen so that the rate of TDB matches TT as closely as possible for the time span covered by the particular ephemeris. TDB is sometimes designated as Barycentric Ephemeris Time ($\mathrm{T_{eph}}$) when used as the time scale of the Jet Propulsion Laboratory Development Ephemerides. When TDB is used, the units of length and mass are slightly different from the geocentric reference systems, and the astronomical unit

needs to be rescaled, depending upon the frame of reference and the accuracy required.

History. It was recognized that the rotation of the Earth was irregular. Thus, in 1952 the International Astronomical Union adopted Ephemeris Time, based on Simon Newcomb's theory of the Earth's motion as a uniform time scale. It became recognized that Ephemeris Time could be determined better from observations of the Moon than of the Sun. Also, in 1955 the more accurate atomic time became available, and it became the basis for Ephemeris Time. With improvements in accuracy, the relativistic distinctions between coordinate times had to be introduced. Hence, TDT and TDB were introduced in 1984. Difficulties with the names, definitions, and accuracies led to the introduction of TT in 1991 and of TCB and TCG in 2000 to replace TDT and TDB. *See* TIME.

P. Kenneth Seidelmann

Bibliography. *Astronomical Almanac*, annually; G. H. Kaplan, *The IAU Resolutions on Astronomical References Systems, Time Scales, and Earth Rotation Models*, USNO Circ. 179, 2005; J. Kovalevsky and P. K. Seidelmann, *Fundamentals of Astrometry*, 2004; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, rev. ed., 1992; P. K. Seidelmann and T. F. Fukushima, Why new timescales?, *Astron. Astrophys.*, 265:833–838, 1992; M. Soffel et al., The IAU 2000 resolutions for astrometry, celestial mechanics and metrology in the relativistic framework: Explanatory supplement, *Astron. J.*, 126:2687–2706, 2003.

Dynamics

That part of classical mechanics which deals with the relation between the motions of material bodies and the forces acting upon them.

Basic concepts. Dynamics proceeds by adopting certain intuitively acceptable concepts which are associated with measurable quantities. These essential concepts and the measurable quantities used for their specification are as follows:

1. Space configuration refers to the positions and orientations of bodies in a reference frame adopted by the observer. It is expressed quantitatively by an arbitrarily chosen set of space coordinates, of which cartesian and polar coordinates are examples. All space coordinates rest on the notion of distance measurement. *See* FRAME OF REFERENCE.

2. Duration is expressed quantitatively by time measured by a clock or comparable mechanism. *See* ACCELERATION.

3. Motion refers to change of configuration with time and is expressed by time rates of coordinate change called velocities and time rates of velocity change called accelerations. The classical assumption that coordinates behave as analytic functions of time permits representation of velocities and accelerations as first and second derivatives, respectively, of the space coordinates with respect to time. *See* VELOCITY.

4. Inertia is an attribute of bodies implying their capacity to resist changes of motion. A body's inertia with respect to linear motion is denoted by its mass. *See* INERTIA.

5. Momentum is an attribute proportional to both the mass and velocity of a body. Momentum of linear motion is expressed as the product of mass and linear velocity. *See* MOMENTUM.

6. Force serves to designate the influence exercised upon the motion of a particular body by other bodies, not necessarily specified. A quantitative connection between the motion of a body and the force applied to it is expressed by Newton's second law of motion, which is discussed later. *See* FORCE.

Distance, time, and mass are commonly regarded as fundamental, all other dynamical quantities being definable in terms of them. *See* MASS; TIME.

Newton's second law. A primary objective of classical dynamics is the prediction of the behavior of bodies which are subject to known forces when only initial values of the coordinates and momenta are available. This is accomplished by use of a principle first recognized by Isaac Newton. Newton's statement of the principle was restricted to the linear motion of an idealized body called a mass particle, having negligible extension in space.

The basic dynamical law set forth by Newton and known as his second law states that the time rate of change of a particle's linear momentum is proportional to and in the direction of the force applied to the particle. This statement, although special in form, serves as a basis for more comprehensive statements of the principle which have since appeared.

Stated analytically, Newton's second law becomes the differential equation (1), in which m represents

$$\frac{d(mv)}{dt} = F \tag{1}$$

the particle's mass, v its velocity, F the applied force, and t the time. Equation (1) provides a definition of force and of its units if units of mass, distance, and time have previously been adopted. The classical assumption of constancy of mass permits Eq. (1) to be expressed as Eq. (2), where a represents the lin-

$$ma = F \tag{2}$$

ear acceleration. A particle in physical space requires three cartesian coordinates, x, y, and z, to specify its position. Its linear acceleration is a vector with three cartesian components, the second time derivatives of x, y, and z. Equation (2) therefore equates two vectors, requiring equality of their components expressed in detail by Eqs. (3). These are the newtonian

$$m\frac{d^2x}{dt^2} = F_x \tag{3a}$$

$$m\frac{d^2y}{dt^2} = F_y \tag{3b}$$

$$m\frac{d^2z}{dt^2} = F_z \tag{3c}$$

equations of motion of an unconstrained particle in space. If the three force components F_x, F_y, and F_z

are expressed functions of the coordinates and time, the dependence of each coordinate upon the time is implied and can in favorable cases be found as solutions of the equations of motion in the form of Eqs. (4). The primary objective of dynamics is

$$x = x(t) \quad y = y(t) \quad z = z(t) \tag{4}$$

achieved in the discovery of such functions.

One-dimensional particle motion. The motion of a particle which remains on the x axis, either because of constraints or initial conditions, is determined by Eq. (3*a*) alone, whose solution is simplified by the absence of y and z. Such one-dimensional dynamical problems provide an attractively simple introduction to the subject. Examples are the motion of a body falling vertically, subject to gravitational force, and linear harmonic motion.

Two-dimensional particle motion. The motion of a particle remaining in the plane of the x and y axes is determined by Eqs. (3*a*) and (3*b*), from which z is absent. Two-dimensional problems are reasonably tractable and include many of physical interest such as the motion of a projectile (exterior ballistics), and of a body attracted toward a central point, as in planetary motion. Solution of a two-dimensional problem is frequently simplified by change of variables which reduces it to a pair of one-dimensional problems.

Three-dimensional particle motion. All three equations of motion, Eqs. (3), apply to an unconstrained particle in space. Complete solutions are possible only when the functions expressing force components are relatively simple in character. Fortunately, many of the solvable cases correspond to important physical examples in which simplicity of the forces allows separation into one- and two-dimensional motions. Three-dimensional projectile motion without friction is an example.

Newton's third law. The behavior of systems composed of two or more interacting particles is treated by newtonian dynamics augmented by Newton's third law of motion which states that when two bodies interact, the forces they exert on one another are equal and oppositely directed. The important laws of momentum and energy conservation are derivable for such systems (the latter only for forces of special type) and useful in solution of problems. The equations of motion for systems of more than two interacting particles in space are mathematically intractable in the absence of geometrical constraints or special initial conditions, but assumptions approximating the physical situation permit solution of many problems of physical interest. The principles of particle dynamics are transferred to extended bodies by regarding them as systems of particles subject to specified mutual constraints and mutual forces. *See* BALLISTICS; GRAVITATION; HARMONIC MOTION; ORBITAL MOTION; RIGID-BODY DYNAMICS.

Russell A. Fisher

Bibliography. H. Goldstein, C. E. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; F. J. Keller, W. E. Gettys, and M. J. Skove, *Physics*, 2d ed., 1993; C. Kittel, W. D. Knight, and M. A. Ruderman, *Mechanics*, Berkeley Physics Course, vol. 1, 2d ed., 1973; K. R. Symon, *Mechanics*, 3d ed., 1971.

Dynamometer

A device for measuring the torque, force, or power available from a rotating shaft. The shaft speed is measured with a tachometer, while the turning force or torque of the shaft is measured with a scale or by another method. Horsepower (hp) may be read from the instrumentation or calculated from Eq. (1),

$$\text{hp} = \frac{2\pi NT}{33,000} \tag{1}$$

where N is the shaft speed in rpm and T is the torque in pound-feet. Power may also be found from Eq. (2) or Eq. (3), where kW represents kilowatts.

$$\text{hp} = \text{torque (lbft)} \times \frac{\text{rpm (shaft speed)}}{5252} \tag{2}$$

$$\text{kW} = \text{torque (newton-meter)} \times \frac{\text{rpm (shaft speed)}}{9549} \tag{3}$$

See TACHOMETER.

Types. The two types are the transmission dynamometer and the absorption dynamometer. The transmission dynamometer transmits the force while measuring the elastic twist of the output shaft. An absorption dynamometer absorbs the power and dissipates it as heat by restraining the output shaft mechanically with a friction brake, hydraulically with a water brake, or electrically with an electromagnetic force. Since the restraining element tends to rotate with the output shaft, the force of the shaft can be determined by measuring the force required to arrest the rotation of the restraining element. Torque is then calculated by multiplying the force times the length of the lever arm, or the distance through which the force acts. The oldest type of absorption dynamometer is the prony brake, which mechanically applies a friction load to the output shaft by means of wood blocks, flexible bands, or other friction surface. *See* TORQUE.

One type of electric dynamometer consists of a direct-current (dc) machine with the stator cradle-mounted in antifriction bearings. The rotor is connected to the shaft of the machine under test. The field current is introduced through flexible leads. The stator is constrained from rotating by a radial arm of known length to which is attached a scale for measuring the force required to prevent rotation. The torque of the connected machine is found from the product of the lever arm length and the scale reading, after correcting the scale reading by the amount of the zero torque reading.

Uses. The most common use of the dynamometer is in determining the power of an electric motor or engine of a car, truck, or other vehicle. A dynamometer that connects to the engine crankshaft is an engine dynamometer. One that has rollers turned by the vehicle drive wheels is a chassis dynamometer;

this type is widely used in the automotive industry for mileage accumulation, emissions, fuel economy, and performance testing of cars and trucks. For fuel-economy testing, dynamometer inertia simulation may be provided with either mechanical-flywheel equivalent weights or electrical load.

If the machine under test is a motor, the dynamometer will act as a generator. Dynamometer output is then absorbed by a loading resistance or by feeding it into a direct-current line. The amount of output is adjusted by changing the loading resistance or the field excitation. If the machine under test is a generator or mechanical load, the dynamometer will act as a motor. Speed is adjusted by changing the armature voltage or the field excitation. *See* DIRECT-CURRENT GENERATOR; DIRECT-CURRENT MOTOR; ELECTRIC ROTATING MACHINERY; GENERATOR; MOTOR.

Donald L. Anglin

Dysprosium

A metallic rare-earth element, Dy, atomic number 66 and atomic weight 162.50. The naturally occurring element is composed of seven stable isotopes. Dysprosium forms a white oxide, Dy_2O_3, which dissolves in acid to give a yellowish-green solution. *See* PERIODIC TABLE; RARE-EARTH ELEMENTS.

The metal is attacked readily by air at high temperatures, but at room temperatures, in massive blocks, it is fairly stable in the atmosphere and remains shiny for long periods of time. Dysprosium is paramagnetic, but as the temperature is lowered, it becomes antiferromagnetic at the Néel point (178 K or −139°F) and ferromagnetic at the Curie point (85 K or −306.4°F). At very low temperatures, the metal shows strong anisotropic magnetic properties.

Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner, Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

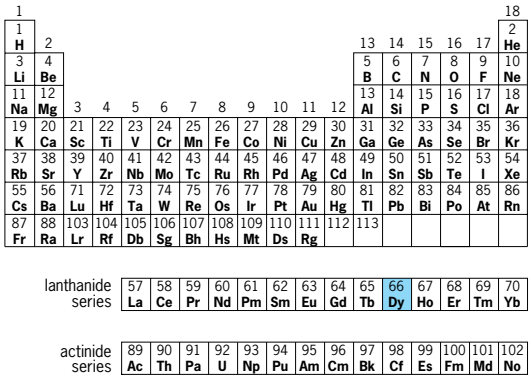

E

e (mathematics) — Eating disorders

e (mathematics)

The number e usually defined as the limit approached by the expression

$$\left(1+\frac{1}{n}\right)^n$$

as n approaches infinity. If the given expression is expanded by the binomial theorem, and if one uses the theorem that the limit of the quotient of two polynomials of equal degree as the variable tends to infinity is equal to the ratio of the coefficients of highest degree, one obtains the expansion in relation (1).

$$e = 1 + \frac{1}{1} + \frac{1}{1\cdot 2} + \frac{1}{1\cdot 2\cdot 3} + \cdots + \frac{1}{1\cdot 2\cdot 3\cdot n} + \cdots \quad (1)$$

Clearly e is larger than 2; it may be easily shown that it is smaller than 3. It can also be shown by elementary methods that e is irrational; that is, it cannot be represented as the quotient of two integers. Furthermore, e is transcendental; it does not satisfy any algebraic equation with integral coefficients. The transcendentality of e was proved by the French mathematician C. Hermite in 1873; the proof constitutes an important milestone in the history of mathematics.

By the method outlined above, it may be shown that the limit of

$$\left(1+\frac{x}{n}\right)^n$$

as n tends to infinity is e^x, and moreover that relation (2) holds. The function e^x is of great impor-

$$e^x = 1 + \frac{x}{1} + \frac{x^2}{1\cdot 2} + \cdots + \frac{x^n}{1\cdot 2\cdot 3\cdots n} + \cdots \quad (2)$$

tance in mathematical analysis and is encountered in numerous problems in applied mathematics. For this reason it has been extensively tabulated. The most recent tables of exponentials for both positive and negative values of x were computed by the Computation Laboratory of the National Bureau of Standards. One of the most important formulas in mathematics involving e is Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$. This is readily obtained from relation (2) by replacing x by $i\theta$. The real part of the expansion is recognized as the expansion of $\cos\theta$, while the coefficient of i is recognized as the expansion of $\sin\theta$. An immediate consequence of Euler's formula is de Moivre's formula $(\cos\theta + i\sin\theta)^m = \cos m\theta + i\sin m\theta$. If the first member of the latter formula is expanded by the binomial formula, and the real and imaginary parts of both members are equated, one obtains two important formulas which permit the evaluation of $\cos m\theta$ and $\sin m\theta$ in terms of $\cos\theta$ and $\sin\theta$. *See* BINOMIAL THEOREM; CALCULUS; LOGARITHM.

Arnold N. Lowan; Salomon Bochner

Ear (vertebrate)

The organ which sends information about sound to the brain, constituting the sense of hearing, as well as vestibular information about the orientation of the head in space. The vertebrate ear is generally divided into three regions that have discrete functions: The inner ear is found in all vertebrates, and it subsumes both hearing and balance (functions). The external ear and the middle ear, not found in all vertebrates, enhance hearing. *See* HEARING (VERTEBRATE).

Ear structure. The inner ear is embedded in the ear (or otic) capsule and has a common embryological development in all vertebrate groups. As the presumptive brain forms in the early embryo, an ectodermal thickening, the auditory placode, appears on each side of the head surface. This balls up, sinks inward, and is then pinched off to form a hollow sac of epithelial tissue (called an otocyst) within what will become the otic bones just lateral to the posterior part of the hindbrain. The embryonic cells of the otocyst will differentiate to form all the epithelial portions of the inner ear. Surrounding mesenchymal

tissue forms the connective tissue of the inner ear. As differentiation proceeds, tissue forms and develops the nerve (cochlear-vestibular, or eighth cranial, nerve) that carries information from the inner ear to the brain. In this manner, the otocyst and immediately surrounding embryonic tissues form all parts of the inner ear, auditory and vestibular.

The middle ear and external ear are not found in the fishes (including jawless fishes, sharks, and bony fishes). All tetrapods (amphibians, reptiles, birds, and mammals) have a middle ear with a tympanic membrane. Reptiles, birds, and mammals also have an external auditory meatus (or canal) which extends from the tympanic membrane to the external surface of the head. Mammals generally have an external structure, the pinna, that helps "collect" and carry the sound to the ear canal and then to the tympanum. The major difference in the middle ear among tetrapods is that it has a single ear bone, or ossicle (often called the columella or stapes), in amphibians, reptiles, and birds, while mammals have three middle-ear bones (malleus, incus, and stapes).

In comparing the inner ears of different vertebrates, from sharks and fishes to mammals, the major structural differences are associated with the auditory part of the ear. With few exceptions (notably in some fishes that detect very low or very high frequency sounds), the vestibular portion of the inner ear is developmentally, structurally, and functionally nearly the same in all vertebrates.

Sensory hair cells. The basic sensory unit in the inner ear is the sensory hair cell (**Fig. 1**). These specialized cells are morphologically similar in all of the epithelial structures of the ear in all vertebrates (and in the lateral line of fishes and amphibians), but they may have either auditory or vestibular functions depending upon the associated superstructure. The superstructure serves to facilitate the transmission of vibrations from the environment to the hair cells. For the vestibular apparatus, the superstructure blocks external vibratory energy, but sensitizes the sensory hair cells to the pull of gravity and to acceleratory and deceleratory movements of the head. *See* LATERAL LINE SYSTEM.

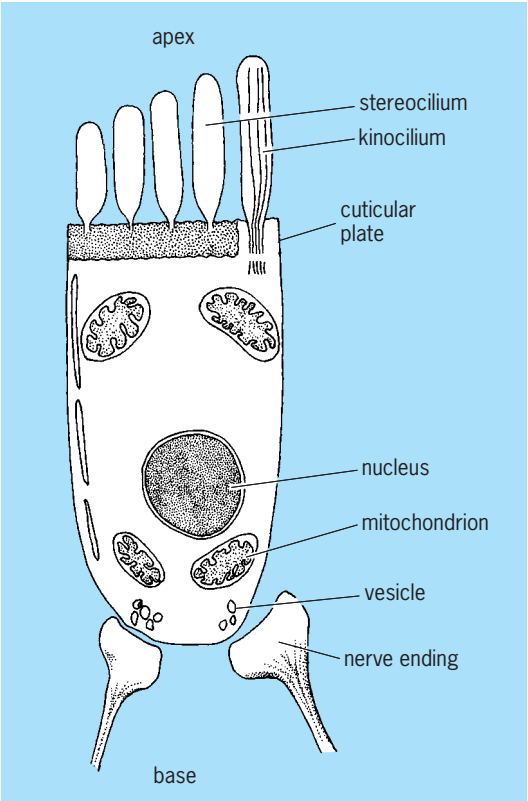

Fig. 1. Generalized diagram of a sensory hair cell found in the ear of all vertebrates.

The sensory hair cell is a columnar, polarized structure from whose apex extend thin cilia that resemble hairs. Each hair cell has many such cilia, making up a ciliary bundle which bends in response to motional energy. The cilia in each bundle include many stereocilia and a single, eccentrically positioned kinocilium. The stereocilia are a homogeneous structure within a simple membrane covering, while the kinocilium contains one pair of microfibrils in the center and nine pairs around the periphery. The stereocilia are arranged in an orderly pattern, and the kinocilium is located at one end of the stereociliary bundle. The cilia extend into an extracellular fluid-filled space, with their tips embedded in a gelatinous membrane. *See* CILIA AND FLAGELLA.

The body of the sensory hair cell lies below the ciliary bundle. Just below the cell membrane, at the apex of the cell, lies a relatively homogeneous electron-dense region, the cuticular plate, from which the stereocilia emerge (Fig. 1). The kinocilium also arises from the cellular apex, but from a pair of basal bodies that are located to the side of the cuticular plate. Immediately below the cuticular plate lies a group of mitochondria, the energy-producing organelles found in all cells. Nearer the basal part of the cell is the nucleus, and below that, another group of mitochondria and a group of small, vesiclelike structures similar to the synaptic vesicles of nerve endings. These structures contain neurotransmitters that are similar to those found in other neuronlike cells in the nervous system. In the sensory hair cell, the vesicles release transmitters into the synaptic space between the cell and the neurons of the eighth cranial nerve that innervate the cells. These transmitters excite the neuron, which then sends on information to the brain. *See* MITOCHONDRIA; NEUROBIOLOGY; SYNAPTIC TRANSMISSION.

The sensory hair cell is the detector of motion, either produced by compression and rarefaction of molecules due to sound waves, or imparted by movement of the head against gravity. This motion produces bending of the ciliary bundles, and this in turn results in a change in configuration of the membrane overlying the stereocilia and opening of channels in the membrane. It is generally thought that these channels admit calcium into the cell, and this in turn interacts with other components of the cell. Ultimately, the energy generated by these interactions causes release of neurotransmitter at the base of the cell, and results in stimulation of afferent neurons which contact the cell.

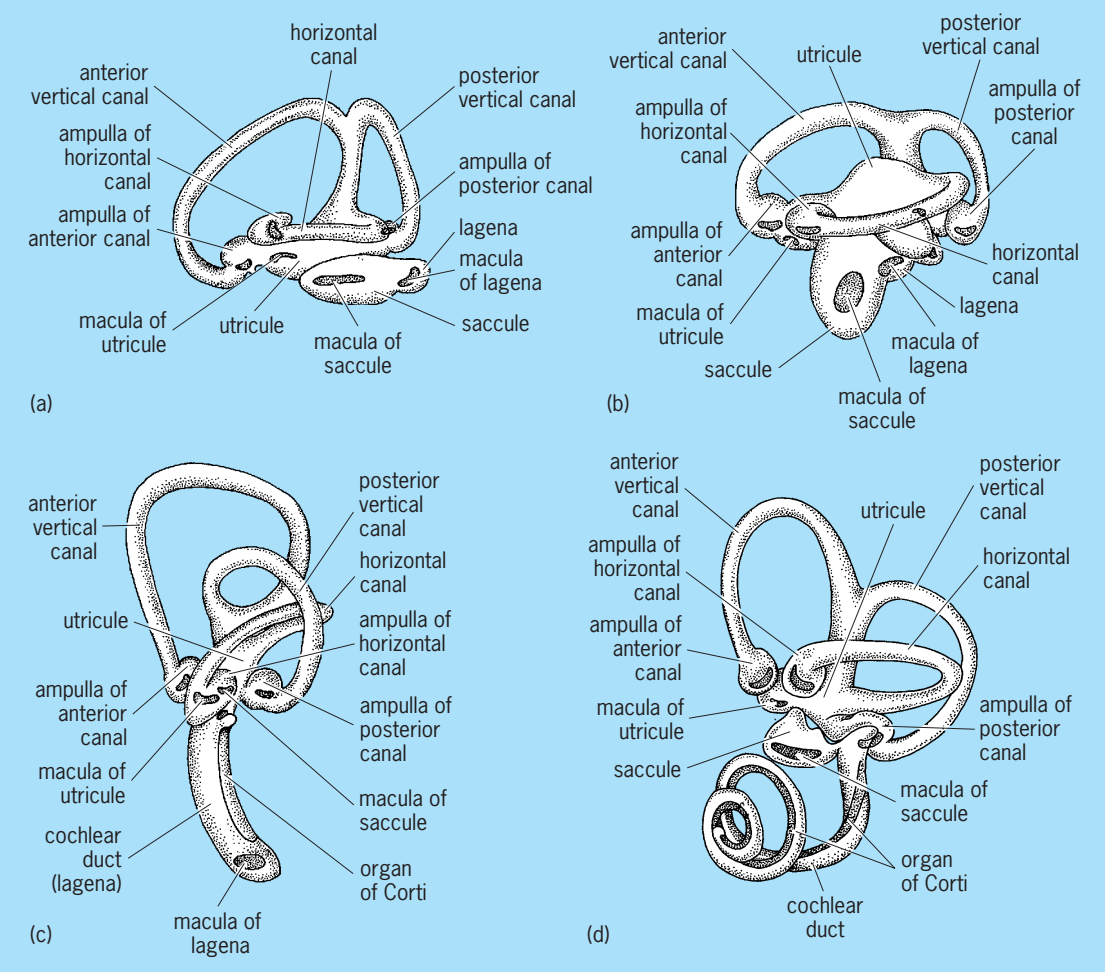

Fig. 2. Left-ear external view of the membranous labyrinths of (*a*) teleost, (*b*) frog, (*c*) bird, and (*d*) mammal. (*After A. S. Romer, The Vertebrate Body, 3d ed., Saunders, 1962*)

Regeneration of sensory hair cells. Once it was widely believed that all sensory hair cells found in a vertebrate are present at birth, and that the only change in number of these cells is a decrease due to cell death as a result of aging or damage produced by chemicals or loud sounds. However, recent findings have shown that both ideas are incorrect, at least in several vertebrate groups. It is now known that elasmobranchs (sharks and rays) and fishes continue to produce additional sensory hair cells throughout their lives (as long as they grow). As a result, the number of sensory hair cells in a large shark or fish is much greater than in a younger animal. Amphibians also add sensory hair cells after birth, but in smaller numbers than in fishes, and there is some evidence that such addition occurs in the vestibular part of the ear in birds. However, mammals do not add sensory hair cells after birth.

It has also been shown in fishes and birds that hair cells will regenerate after trauma induced by certain drugs (such as streptomycin and related drugs) and loud sounds. Both stimuli kill sensory hair cells, but after several days or weeks these cells regenerate, and it now appears that full function returns to the ear, at least in birds. Again, there is no evidence to indicate that destroyed hair cells are replaced in mammals.

Fishes. While the basic structure of the inner ear and of the sensory hair cells is the same in all vertebrates, a number of interesting differences reflect the evolution of the ear and the ways in which different animals detect, process, and use sound (**Fig. 2**).

In elasmobranchs and bony fishes the inner ear is located in the brain (cranial) cavity somewhat behind the eye (**Fig. 3**). The inner ear has several regions, including three semicircular canals and otolith organs. Other than the very primitive jawless fishes (agnathans: lamprey and hagfish) which have one or two semicircular canals, all other vertebrates have three canals. All fishes, amphibians, reptiles, and birds have three otolith organs—the saccule, utricle, and lagena—while mammals do not have the lagena. It has been suggested by some investigators that the lagena gave rise to the cochlea, which is the main auditory organ of mammals.

Each of the three semicircular canals lies on one plane of space. The horizontal canal is in the horizontal plane relative to the animal's normal head position; the anterior and posterior canals are vertical

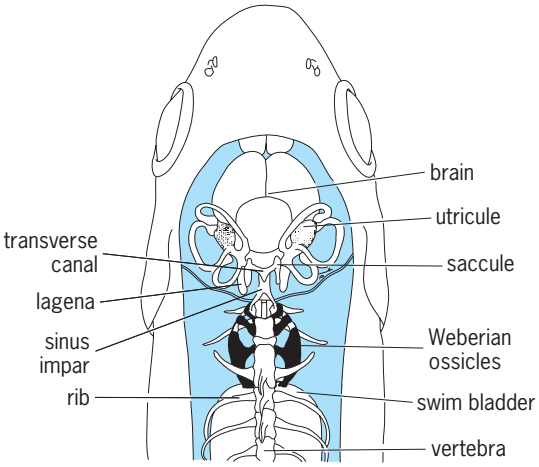

Fig. 3. Dorsal view of the anterior part of the head and body of an otophysan fish showing the ear and its position relative to the brain. Also shown are the swim bladder and the Weberian ossicles. (*From A. N. Popper and S. Coombs, Auditory Mechanisms in Teleost Fishes, Amer. Sci., 68:429–440, 1980, redrawn from K. von Frisch, Über den Gehörsinnder Fische, Biol. Rev., 11:210–246, 1936*)

and at right angles to each other as well as to the horizontal canal (Fig. 2*a*). While the names of the canals differ in humans, due to the upright position of the body, the canals in humans are identical in form and function to those in all other vertebrates.

At one end of each endolymph-filled semicircular canal is a widened area, the ampulla, which has a sensory area called the crista (or crista ampullaris). The crista contains large numbers of sensory hair cells, as well as other cells which provide support for the hair cells. At the base of the hair cells are nerve endings from the vestibular branch of the eighth cranial nerve.

The long kinocilia and stereocilia of the hair cells of the crista extend into a gelatinous structure, the cupula. The cupula blocks the semicircular canal at this point and prevents the endolymph from free movement into the ampulla. Upon acceleration or deceleration of the animal's head, movement of the fluid in the semicircular canal of the corresponding plane lags slightly behind the movement of the membranous canal and of the head itself. When the fluid catches up, it puts pressure upon the cupula, bending the hair cells and initiating nerve impulses. Since more than one semicircular canal can react at a time, depending upon the precise movement of the head, the brain receives very precise information about head movement.

Each of the otolith organs also has a sensory area, called a macula, that contains hair cells and supporting cells. The cilia of the otolithic organs are not nearly as long as those found in the ampullae of the semicircular canals. They are, however, embedded in a thin gelatinous membrane that also contains very dense calcium carbonate crystals. In elasmobranchs, as in primitive fishes and all tetrapods, these crystals are called otoconia. In most bony fishes the crystals are fused into a single mass in each otolith organ called the otolith. The otoliths generally have very species-specific shapes and grow as the fish grows by adding daily and yearly rings. By proper preparation of the otoliths, and particularly those of the saccule, it is possible to tell the age of a fish, as well as the quality of diet and other physiological characteristics about the animal. In some ways, this is analogous to dating a tree by the number of growth rings.

Hearing. Fishes are able to detect a wide range of sound using their inner ear. Most fishes detect sounds from 30 to 800 or 1000 Hz, with best hearing from 200 to 500 Hz. However, some fishes, called hearing specialists (such as goldfish and catfish), have evolved special mechanisms to enhance hearing to 3000 or 4000 Hz. A few species in the herring family can detect sounds to almost 200,000 Hz, and it is thought that this capability has evolved to enable these species to detect echolocating signals of their major predators, dolphins, and avoid being eaten. Significantly, fish not only can detect the presence of sound but also are able to discriminate between sounds, determine the direction of a sound source, and detect sounds in the presence of noise, as can mammals.

Many species of fish also produce sounds and use sound for communication, as in reproduction and territorial behavior. In these species, hearing is needed for communication with other animals. Other species, such as elasmobranchs, detect prey by the sounds they make. However, many species of fish hear well and yet are not known to produce sounds or use sounds to find food (such as the goldfish). It is thought that hearing evolved so the animals can gain general information about their environment from the noises that are constantly found in the sea. In many ways, this is similar to the kind of information that a person gets from sound upon entering a dark room. Thus, it is likely that the use of sound for communication came after the sense of hearing evolved.

Tetrapods detect sounds that impinge on the tympanic membrane and then are carried by the middle-ear bones to the inner ear, where the sounds set the fluids of the ear into motion and thus stimulate the sensory hair cells. In fishes, however, this kind of pathway is not needed since sound is already traveling through water. Indeed, since the fish's body is the same density as that of water, sound would travel right through the fish were it not for the otoconia or otoliths. Since these structures are much denser than the fish's body and the water, they stay still while the fish's body and the attached sensory hair cells move with the sound field. Since the stereocilia are attached to both the top of the hair cells and to the otoconia or otolith, they are bent as their base moves with the macula and their tops stand still with the otoliths. This bending sends signals to the nerves and then to the brain, indicating the presence of a sound.

Swim bladder. Some species of fish, the hearing specialists, use a secondary structure, the swim bladder, to enhance hearing capabilities. The swim bladder is a bubble of gas found in the abdominal cavity of most

bony fishes (but not elasmobranchs), and it is used primarily for buoyancy control, though it may also be used in sound production in some species. Since the swim bladder is filled with gas, its density is different from that of the rest of the fish, and in a sound field the walls of the swim bladder are set into vibration and act as a small sound source to send sounds to the ear. This method enhances hearing only in the hearing specialist species, where the front end of the swim bladder comes close to the saccule or utricle, or there is some bony connection that serves in a manner analogous to the tetrapod middle-ear bones. The best-known example is found in the group of fishes called the Otophysi (catfish, goldfish, carp), where a series of bones, the Weberian ossicles, transmit the movement of the swim bladder walls to the saccule of the inner ear (Fig. 3). These hearing specialists can hear a wider range of frequencies than nonspecialists, and their hearing sensitivity at each frequency is also better. *See* SWIM BLADDER.

Investigators who first studied fish hearing in the early part of the twentieth century thought that the saccule and lagena were the auditory parts of the ear in fishes (pars inferior) while the utricle served vestibular functions along with the semicircular canals (pars superior). However, more recent studies strongly suggest that all three otolith organs in fishes are involved in both the vestibular and the hearing senses.

Tetrapods. Many structural and functional features of the fish inner ear are also found in the tetrapod ear. The inner ear of tetrapods is embedded in the otic bones of the skull, with the membranous labyrinth attached to the bony labyrinth by connective tissue but suspended in perilymphatic fluid. There are three semicircular canals, with cristae, and, except in mammals which do not have a lagena, the three otolithic organs (Fig. 2*b*, *c*, *d*). In their morphology and physiology the vestibular parts of fish and tetrapod ears are nearly the same. For the most part, the tetrapod otolithic organs function only as vestibular organs rather than playing an auditory role as they do in fishes.

Adaptations for aerial habitat. Audition in land vertebrates poses special problems not met by fishes. These problems arise from the aerial (as opposed to the aquatic) environment and from the different specific gravities of environment and organism.

The auditory inner ear is in a fluid environment and must receive sound vibrations from an aerial environment (unlike in fishes, where the sound is carried in basically the same fluid as found in the inner ear). Because air is much less dense and more elastic than water, almost all sound energy in air is reflected from a fluid surface rather than being transmitted to the fluid. In physical terms, this is called an impedance mismatch between air and fluid. In order for the sound energy in air to be transmitted to fluids, a mechanism must be present to match the impedance of air with that of the fluids. A combination of increasing the sound pressure and decreasing the particle velocity at the interface between air and fluid can match their impedances. This is the primary function of the tetrapod middle-ear apparatus. *See* SOUND.

With the emergence of vertebrates on land, gilled respiration was lost. The gill slits became obliterated except for the most anterior, which, highly modified, became the middle-ear cavity, retaining its connection to the mouth cavity in the structure called the eustachian tube. The cavity does not quite open to the outside, but is covered by a very thin layer of epithelium, called the tympanic membrane.

Amphibians. The tympanic membrane of frogs and toads is located on the lateral surface of the head. Attached to its inner aspect is a small rodlike bone, the stapes, or columella, which runs through the air space of the middle-ear cavity and plugs a small hole, the oval window, beyond which are the inner-ear fluids. The frog's tympanic membrane collects sound energy and transmits it through the columella to the inner-ear fluids. The surface area of the tympanic membrane is quite large, while the surface area of the columella footplate, where it meets the inner ear fluids, is much smaller. As a consequence, although the amount of energy collected by the tympanic membrane is the same as that received by the inner-ear fluids, the energy has been concentrated into a much smaller area by the time it reaches the fluids. Thus, the force per unit area (sound pressure) at the inner ear is much greater than it is at the tympanic membrane. This is the impedance-matching mechanism of the middle ear of amphibians. *See* AMPHIBIA.

Sensory areas. In the lagenar portion of the amphibian's membranous labyrinth are two areas of hair cells, the amphibian and basilar papillae, that are found in no other vertebrate group. The basilar papilla lies on the posterior wall of the saccule between the oval window and the round window, another membrane-covered opening between middle ear and inner ear. The ciliary bundles on the sensory hair cells of the basilar papilla reach into the endolymph and are embedded in the gelatinous tectorium, or tectorial membrane. This tectorium is thought to correspond to the structure of the same name in the ears of amniotes (tetrapods other than amphibians), but differs in that it not only receives cilia from the hair cells but also sends down strands of its own between the hair cells to attach to the supporting cells.

The amphibian papilla has a structure similar, although not identical, to the basilar papilla. Movements of the inner-ear fluids cause bending of cilia of the hair cells in both papillae, but they respond differentially to different frequencies. The precise frequency range that each responds to differs in different species, and it is related to the frequencies of the sounds used by a particular species for communication.

Sound reception. Vibratory energy enters the inner ear at the oval window, passes through the basilar and amphibian papillae causing them to vibrate, and then dissipates at the round window. Physiological experiments on frogs and toads have shown that the basilar and amphibian papillae are the primary organs receiving sound and transducing its physical

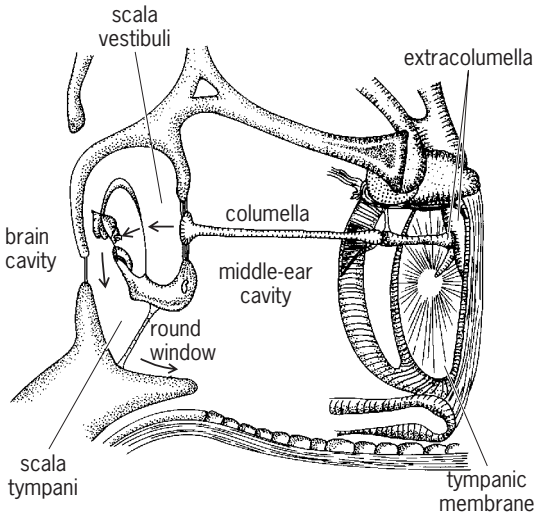

Fig. 4. Diagrammatic view of the middle and inner ears of a lizard. (*After D. Webster and M. Webster, Comparative Vertebrate Morphology, Academic Press, 1974*)

vibrations into nerve impulses. However, some data suggest that the saccule may also serve an auditory function in frogs and toads, particularly for the detection of very low frequency vibrations, and even seismic motion.

Birds and reptiles. In most reptiles and birds the tympanic membrane lies not on the surface of the head but internally, at the end of the tube called the external auditory meatus. A middle-ear cavity (with its eustachian tube to the mouth) lies medial to the tympanic membrane. A single ossicle, the columella, crosses this cavity from the tympanic membrane to the oval window at the inner ear. As in amphibians, the surface area of the tympanic membrane is much greater than that of the footplate of the columella. As a result, the force per unit area is thus greatly increased, and aerial vibrations cause fluid vibrations in the inner ear (**Fig. 4**). *See* AVES; REPTILIA.

Auditory end organ structure. While both birds and reptiles have saccule, utricle, and lagena, as well as semicircular canals, they also have a newly evolved end organ, the basilar papilla, which is the part of the ear used for hearing in both groups of animals. (The avian and reptilian basilar papilla is thought to be a totally different structure, in terms of evolution and embryonic origin, than the basilar papilla found in amphibians). The lagena lies at the end of the elongate basilar papilla (Fig. 2*c*), and so it has been suggested by some investigators that the basilar papilla is derived from the lagena. However, there are few data to support this suggestion. The basilar papilla in birds and reptiles is often also called the cochlea, and there is some evidence to suggest that this end organ is directly related to the mammalian cochlea.

The basilar papilla in reptiles is generally somewhat shorter than that found in birds, and there is considerable variation in the specific structure of this end organ in different species. In particular, the basilar papilla in lizards shows wide variation. However, very little is known about the functional significance of this variation, and in fact, very little is known about hearing capabilities of reptiles since they are hard to test in behavioral experiments. At the same time, since the basilar papilla is so well developed in many species, it is likely that these animals are capable of detecting and discriminating a wide range of sounds. While there is less variability in the basilar papillae of birds, it is known that many of these species hear very well (some species to over 7000 Hz), and sound is a very important part of the avian communication system.

The basilar papilla contains sensory hair cells that are supported by sustentacular cells and arranged in rows, and each has a kinocilium and several stereocilia. In birds, the basilar papilla sensory hair cells are differentiated into short and tall hair cells, which may have different functions in hearing.

Efficiency of transduction. Above the basilar papilla the cilia of the hair cells are embedded in a gelatinous rectorial membrane whose precise shape differs markedly in different species. It is attached not to the supporting cells, as in amphibians, but to the spiral limbus, a ridge of epithelial cells firmly adhering to the bony labyrinth on one side of the basilar papilla. Vibrations in the inner ear cause the basilar membrane and its papilla to move, while the tectorial membrane, anchored on bone, does not move. The hair cells extending from the papilla and embedded in the tectorial membrane are thus bent, and more efficiently so than are those of amphibians, in which the membrane, the hair cells, and the supporting cells are attached only to each other.

Mammals. The mammalian ear consists of three parts: the external ear which receives the sound waves; the middle ear which transmits the vibrations by a series of three small bones; and the inner, or internal, ear, a complex bony chamber placed deep in the skull (**Fig. 5**). The external auditory meatus plus the newly evolved pinna, a cartilaginous structure

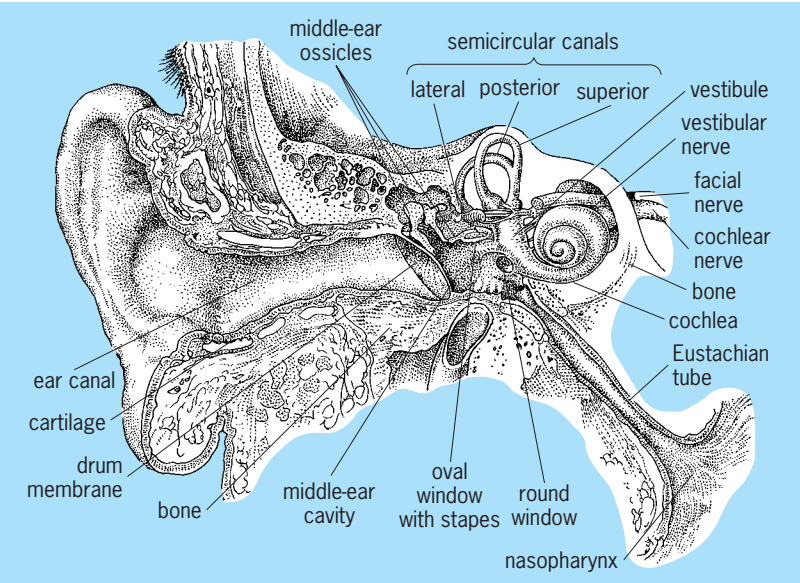

Fig. 5. Schematic drawing of the human ear. (*After M. Brödel, Three Unpublished Drawings of the Anatomy of the Human Ear, Saunders*)

projecting from the ear, compose the external ear. The shape and size of the pinna vary greatly. In elephants it is a huge, flapping structure, while in whales and moles it is completely absent. The size of the pinna in other species falls between the two extremes. The auditory function of the pinna varies widely in different species. In some species the pinna is moved in the direction of a sound source and helps the animal focus sound to the external auditory meatus and then down the ear canal. In other species, such as humans, the pinna may have a lesser function, but even in humans the pinna helps to discriminate between sounds coming from the front and back of the head so that the person can better tell the direction of a sound source. *See* MAMMALIA.

Middle ear. As in other tetrapods, the first gill slit is modified as a middle-ear cavity, communicating with the pharynx by way of the eustachian tube (**Fig. 6**). In other tetrapods this tube is permanently open, while in mammals it is usually closed. In order to equalize the pressure between the inside of the middle-ear cavity and the environment, the tube must be opened by yawning or swallowing. This is the "popping" in the ears which is experienced when one drives rapidly up a mountain or rides in an elevator.

Instead of the single columella of other tetrapods, the mammalian middle ear has three bones, closely articulated with one another. The innermost is the stapes, which fits into the oval window of the inner ear and is homologous with the columella. Attached to the tympanic membrane is the malleus, and lying between the malleus and stapes is the incus. These outer two bones are derived from the quadrate and articular bones, which in other tetrapods and most fishes are part of the jaw joint.

The geometry and arrangement of these three ossicles constitute a lever system which produces both an amplification of pressure and a decrease in particle velocity. However, the amplification of force per unit area, from large tympanic membrane to small stapes footplate, also occurs here. The adaptive value of three ossicles rather than one appears to be an added flexibility for hearing loud as well as soft sounds without endangering the delicate inner ear. In spite of having additional bones, the mammalian middle ear functions basically as do those of amphibians, reptiles, and birds in transforming aerial vibrations into fluid vibrations within the inner ear.

Inner ear. In the mammalian inner ear the vestibular apparatus is much like that of other tetrapods. The auditory portion, however, is considerably different (**Fig. 7**). It is elongated and coiled into a snail shape. This structure is called the cochlea. The epithelium of the basilar papilla is more differentiated in mammals than in other tetrapods, and it is called the organ of Corti. The number of turns in the cochlea varies from less than one in the generalized monotremes, to two and a half in humans, to five in the guinea pig. At the base of the cochlea is the oval window, which carries sound energy into the inner ear, and the round window, where this energy is dissipated after traveling in the cochlea.

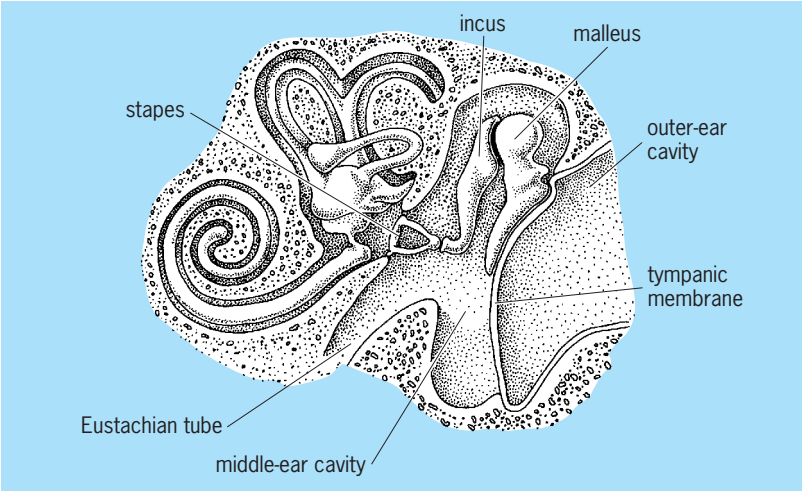

Fig. 6. Otic region of the heah of a mammal (showing the ear region only). (*A. S. Romer, The Vertebrate Body, 3d ed., Saunders, 1962*)

Running the length of the coiled cochlea are three channels; the uppermost, the scala vestibule, and the lowest, the scala tympani, communicate with each other at the helicotrema, an opening at the apex of the cochlea. These channels are filled with perilymph. In the center is the scala media, or cochlear duct. The cochlear duct is filled with endolymph, and it is separated from the scala vestibule above by the thin Reissner's membrane and from the scala tympani below by the basilar membrane.

The basilar membrane is suspended on both sides by ligaments or bone. The basilar membrane varies regularly in width, being narrow at the base (where it is most responsive to high frequencies) and wide at the apex (where it is most responsive to low frequencies).

Resting upon the basilar membrane is the organ of Corti. The organ of Corti contains several cell types in addition to the auditory hair cells. The hair cells are of two types, one type on either side of the pillar cells. The hair cells lying on the internal side of the pillar cells are called the inner hair cells, and those lying on the external side are called the outer hair cells. Morphologically, the inner hair cells are flask-shaped, whereas the outer hair cells are cylindrical. The external cell membrane of the outer hair cells is also lined by molecules which appear to act as molecular motors which can contract and cause a change in the length of the sensory cells. These motors are not found in inner hair cells, or the vestibular hair cells of mammals, or any hair cells in other vertebrates.

In both inner and outer hair cells of many mammals, including humans, the kinocilium is completely missing after birth of the animal, and only stereocilia are present. Typically, the single row of inner hair cells and three rows of outer hair cells run the length of the basilar membrane. Outside the hair cells and supporting cells are lines of border cells.

There may be up to 20,000 sensory hair cells in a cochlea of a normal young human, although the number of hair cells declines with age as a result of normal cell death, damage due to some medications,

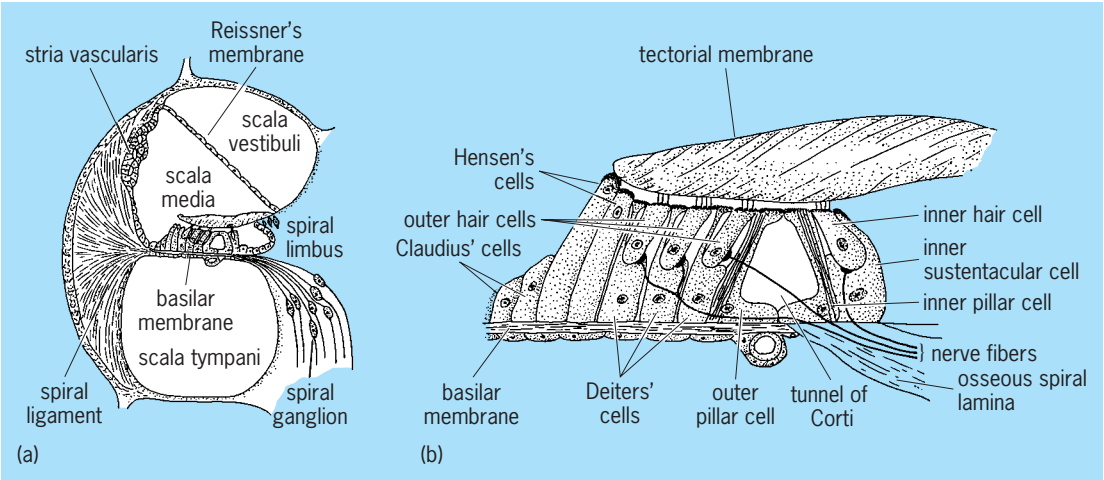

Fig. 7. Inner ear of a mammal. (*a*) Cross section of a single turn of the cochlea. (*b*) Organ of Corti. (*After D. Webster and M. Webster, Comparitive Vertebrate Morphology, Academic Press, 1974*)

and trauma caused by loud sounds. Indeed, hearing abilities of humans (and presumably other mammals) decline with age. A healthy teenager may hear sounds from below 20 Hz to upward of 20,000 Hz, while an adult 40 or 50 years old may hear sounds only to 14,000 Hz (or even less). This loss of hearing is associated with death of sensory hair cells.

Lining the external, or outer, wall of the scala media is a highly vascularized portion known as the stria vascularis. This area is metabolically very active and supplies the total blood and nutrients for the organ of Corti. Along the inner wall of the scala media (which is attached to the bone of the modiolus) is the area known as the spiral limbus. This is attached firmly to the gelatinous rectorial membrane which extends over the organ of Corti. In turn, the stereocilia of the sensory hair cells of the organ of Corti are embedded in the tectorial membrane. As in birds and reptiles, the tectorial membrane is held steady while the basilar membrane vibrates; the embedded hairs are bent, and transduction of physical energy into nerve impulses occurs (Fig. 5).

Sounds entering the mammalian inner ear at the oval window travel along the basilar membrane from basal to apical ends, causing vibrations of the membrane. The largest motion of the basilar membrane to a sound depends on the frequency components of the sound. Different frequencies maximally excite different regions of the basilar membrane, with high frequencies causing maximum motion of the basal end (the end near the oval window) and low frequencies maximally exciting regions closer to the apical end (near the helicotrema). This produces a "place mechanism" of hearing. When complex sounds, such as speech, enter the ear, different regions of the basilar papilla respond to different frequencies making up the sound. In effect, the basilar membrane is capable of breaking down a sound into its component frequencies.

The basis for responding to different frequencies at different regions of the basilar membrane is the stiffness of the membrane itself. The stiffness is a function of many factors, including the width and thickness of the membrane which varies along the length of the cochlea. The response of the different regions of the organ of Corti to specific frequencies is also thought to be enhanced by the sensory hair cells themselves. Whereas early investigations suggested that both inner and outer hair cells were involved in detection of sound per se, recent evidence suggests that the inner hair cells have the major role in hearing, while the outer hair cells modify the function of the ear and help to enhance the sensitivity of the inner hair cells.

Each of the outer hair cells is innervated by nerves which bring signals to the cells from the brain, while the inner hair cells have nerves which carry signals from the cells to the brain. The signals to the outer hair cells appear to tell the cells to contract or lengthen, in response to the nature of the sound being detected, and this change in length affects the stiffness of the membrane, and thus the sensitivity of the inner hair cells. *See* HEARING (HUMAN).

Arthur N. Popper; Douglas B. Webster

Bibliography. R. J. Dooling, R. R. Fay, and A. N. Popper (eds.), *Comparative Hearing: Birds and Reptiles*, Springer-Verlag, New York, 2000; R. R. Fay and A. N. Popper (eds.), *Comparative Hearing: Fish and Amphibians*, Springer-Verlag, New York, 1999; R. R. Fay and A. N. Popper (eds.), *Comparative Hearing: Mammals*, Springer-Verlag, New York, 1994; A. N. Popper and R. R. Fay, Evolution of the ear and hearing: Issues and questions, *Brain Behav. Evol.*, 50:213–221, 1997; N. B. Slepecky, Structure of the mammalian cochlea, in P. Dallos, A. N. Popper, and R. R. Fay (eds.), *The Cochlea*, Springer-Verlag, New York, 1996.

Ear protectors

Objects or devices which insert into the ear canal or cover the outer ear and which have a variety of applications. They may protect the ears from

exposure to caustic materials, extremely cold temperatures, or excessive noise; serve as a barrier against solid particles collecting on the surface of the outer ear (auricle); prevent accidental injuries from flying objects in industrial activities, from contact during physical activities, or from excessive heat or fire; prevent moisture or solid particles from entering the ear canal; reduce noise to acceptable levels so signals are heard better through earphones; and protect the outer ear from direct sun rays. The purpose for wearing ear protectors determines the characteristics of the device to be used.

Hearing protectors. The most common use of ear protectors is to protect hearing from excessive noise; therefore "hearing protector" is a most appropriate term. Dry cotton has been used for many years as a form of hearing protection, but unfortunately this material provides very limited protection against the intrusion of excessive noises. Developments in plastics have greatly improved the amounts of protection that may be achieved. Comfort is one of the major factors in the successful use of all devices. The amount by which a device reduces or attenuates the magnitudes of noises reaching the structures of the inner ear is also an important consideration. Advances in materials used to construct hearing protection devices have led to significant improvements in both attenuation and comfort for users. The use of hearing protectors can prevent both temporary (auditory fatigue) and permanent noise-induced hearing losses, and may actually enhance a person's ability to hear in the presence of excessive noise. *See* HEARING (HUMAN).

Hearing protectors that insert into the opening of the ear canal are usually referred to as earplugs. Those that enclose the structures of the outer ear are usually called earmuffs. The manner in which each is constructed and used may differ considerably. A variety is found among earplugs: premolded plugs designed to accommodate the anatomic characteristics of most human ears; custom-molded devices that mirror the anatomic contours and features of an individual ear; expandable material which is compressed into a cylinder and inserted into the opening of the canal, where it expands to make a comfortable seal; fibrous material partially formed by the user and inserted into the opening of the ear canal; and earcaps that are held in place against the opening of the ear canal by pressure from a light-weight headband. Earmuffs usually consist of a molded cup completely encasing the outer ear and fitted with a compliant cushion that rests on the head in the area surrounding the ear. Earmuffs are held firmly in place by pressure from a headband which may be placed over the top of the head, under the chin, or behind the neck. Earmuffs may also be attached to a hardhat or helmet.

Attenuation measurement. Various schemes have been devised to measure and express the noise reduction or attenuation provided by different types of hearing protectors. The most common method used for expressing amounts of attenuation compares the sound reduction that occurs when the device is worn with values obtained when the ears are not protected. Studies that have compared attenuation under laboratory conditions with that achieved in actual usage consistently reveal that users do not typically obtain the amounts of protection measured in controlled laboratory situations. The most common method for expressing values of attenuation is the noise reduction rating (NRR). Generally, NRR values range from 0 to approximately 30, with higher numbers indicating greater amounts of noise reduction (see **illus.**). *See* LOUDNESS.

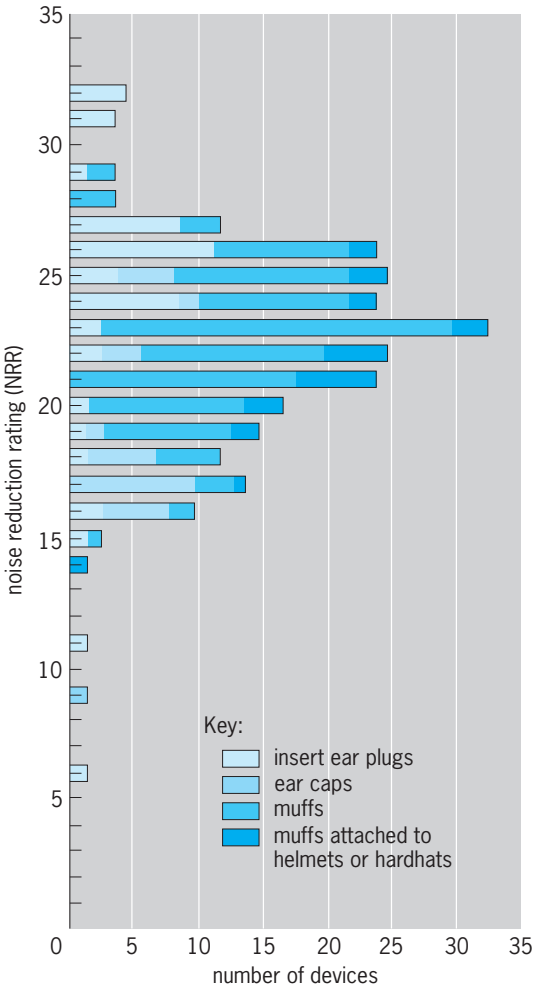

Noise reduction ratings (NRR) for various types of personal hearing-protection devices. A total of 242 devices were evaluated. (*After B. L. Lempert, Compendium of hearing protection devices, Sound Vibrat., 18(5):26–39, 1984*)

Protection from water. People frequently wear earplugs to prevent water from entering the ears while taking a shower or swimming. While few earplugs totally exclude water from entering the ear canals, especially if the head is completely immersed, some, if properly fitted, can significantly impede the passage of water. The pressure exerted by the water on the outer surface of the plugs increases dramatically the deeper one goes below the surface, making it even more difficult to exclude the water totally.

Noise identification. Millions of people possess some degree of noise-induced loss which could have

been prevented by wearing hearing protection during encounters with excessive noise. Controlling such excessive exposures requires that hazardous noises be identified. A fairly simple approach is to consider the following rule: if the noise encountered would cause one to use a loud voice to communicate at a distance of approximately 1 ft (0.3 m), the noise should be avoided or the levels reaching the inner ears should be reduced by the use of hearing protection. *See* ACOUSTIC NOISE. Donald C. Gasaway

Bibliography. E. H. Berger, Using the NRR to estimate the real world performance of hearing protectors, *Sound and Vibration*, 17(1):12–18, 1983; C. M. Harris (ed.), *Handbook of Noise Control*, 3d ed., 1991; A. H. Suster and P. L. Johnson, Progress in controlling occupational noise exposure, *Noise Control Engineering Journal*, 44(3):121–126, 1996.

Early modern humans

The earliest representatives of people anatomically similar to living humans evolved from more archaic humans approximately 100,000 years ago. The process by which they emerged from and eventually replaced those late archaic humans remains unclear. It is likely that they evolved locally from such predecessors in northeastern Africa. Over the succeeding 50,000 years, their range expanded and contracted with changing global climatic cycles to include at times the Mediterranean Near East and portions of northern Africa. Early modern humans and their biology and way of life, therefore, initially had little advantage over late archaic humans. *See* NEANDERTALS.

Subsequently, starting around 50,000 years ago, the descendants of these earliest modern humans were able to expand their geographical range. They spread into regions occupied by geographical groups of late archaic humans, in some cases admixing with those local populations and in other regions displacing them. The spread of early modern humans appears to have taken place initially across northern Africa to Morocco and southward through eastern Africa to southern Africa. At the same time, prior to 40,000 years ago, they appear to have spread eastward across southern Asia, through Indonesia, and across open water into greater Australia. Between 40,000 and 30,000 years ago, these early modern humans dispersed northward across Eurasia from the Mediterranean to the Pacific, and westward across Europe, reaching the Atlantic peripheries of Europe as late as 30,000 years ago. The last of these groups, the early modern humans of western Europe, were previously referred to as the Cro-Magnon people, after the site in France where their remains were first discovered in 1868.

This scenario for the emergence and spread of modern humans is based on evidence from the human fossil record (documenting the earliest modern humans in different regions), living human genetic diversity (similarities among humans, indicating a generally recent shared ancestry), and recent human anatomical diversity (requiring considerable time to become established). Taken together, these lines of evidence highlight the complex and geographically variable nature of this evolutionary period, the emergence of modern humans.

Early modern humans were biologically the same as modern peoples and would blend in with living peoples. They differed from living people primarily in their tendency to have a rugged, athletic build. This is evident in the prominent muscle attachments and the structural reinforcements of the limb bones. Given their mobile hunting and gathering existence, their legs and feet were especially strongly built. In addition, they had slightly larger teeth than those of recent humans, providing greater resistance to the abrasion of eating and cooking without utensils.

These populations were generally tall, males being 175–180 cm (5 ft 10 in.) and females 160–165 cm (5 ft 4 in.) on average. From about 20,000 years ago until the twentieth century, few human populations achieved such large statures. As a result of these tall bodies and muscularity, their brains were relatively large, averaging about 1500 cm^3 (90 $in.^3$) as opposed to averages of about 1350 cm^3 (80 $in.^3$) common for recent humans. Yet, when their brain sizes are scaled against estimated body weights, their brains were relatively the same size as those of living humans.

Early modern humans were successful hunters and gatherers, occupying most of the inhabitable regions of the Old World. Although frequently portrayed as big game hunters, they lived by hunting small to medium-size animals, especially antelopes, deer, goats, and occasional horses and cattle, and by gathering wild plants, fish, and shellfish for food. They were sufficiently successful that the levels of biological stress from famine or injury, as indicated by the lesions on their skeletons and teeth, are among the lowest known for any prehistoric human group.

Their effectiveness as hunters and gatherers was due in part to their technology. They developed elaborate stone tool technologies, producing long blades that became blanks for tools with replaceable cutting edges and points. This advance was made possible by the use of bone, antler, and wood for carefully made hafts and handles for the sharp stone edges. They were also the first to fire clay into ceramics, and they wove carrying bags with a variety of techniques. Yet, their ability to live effectively as hunters and gatherers depended upon their extensive knowledge of the environment, so that they could harvest wild game and plants rather than seek them opportunistically. This knowledge was communicated through the first elaborate symbolic systems known, which consisted of a variety of geometric notational systems and the first representational art. They were also the first humans to commonly wear jewelry, and hence to modify their personal social images, suggesting more complex social roles than were previously known.

Although these behavioral advances are associated with early modern humans, most of them appear

only after about 50,000 years ago and hence are associated with the dispersal of modern humans. What behavioral advances, if any, were associated with the earliest modern humans remains unclear. However, their fossil remains suggest that even they were less stressed and more efficient as hunter-gatherers than their archaic predecessors and contemporaries. *See* FOSSIL HUMANS. Erik Trinkaus

Bibliography. M. H. Nitecki and D. V. Nitecki (eds.), *Origins of Anatomically Modern Humans*, Plenum, 1994; C. Stringer and C. Gamble, *In Search of the Neanderthals*, Thames & Hudson, 1993; J. Svoboda, V. Ložek, and E. Vlček, *Hunters between East and West*, Plenum, 1996; M. H. Wolpoff, *Human Evolution*, McGraw-Hill, 1996.

Earphones

A class of energy transducers capable of receiving alternating current and generating acoustic waves resembling very closely the characteristics of that current. The movement of an element (diaphragm) is accomplished by magnetic attraction, electrostatic attraction, or the piezoelectric effect (the expansion or contraction of certain crystalline substances in response to electric charges). Earphone systems include the driver element with its diaphragm and arrangements for magnetic flux, or electrostatic or direct electric charge, plus a casing, one or more acoustic cavities and ports, acoustic damping and insulation, and some arrangement for coupling the driver to the human ear. The wiring connecting the precedent amplifier to the driver, which may be incorporated into the earphone system, may in modern systems be a complicated circuit which feeds, for each of two stereophonic channels, some part of the current to each ear. The time delays and energy ratios at the two earphones (each of which may contain two drivers) can be appropriately adjusted so that the listener is given the illusion that the sound sources are not "within the head" but are externalized appropriately in all three planes of space. The various types of earphones are described below. *See* TRANSDUCER.

Magnetic. In this type, a permanently magnetized diaphragm is moved in and out by an electromagnet energized by alternating current. Early models had two poles attracting and repelling a relatively heavy metal diaphragm with complex vibrational modes, the frequency response being limited to about 300–2500 Hz. In one development, the heavy metal diaphragm was replaced by a thin sheet of aluminum held by a ring of magnetic metal. The coil through which the signal current traveled had to move only the relatively low-weight ring, a greatly increased efficiency and with improved frequency response.

Dynamic. These earphones are actually small dynamic loudspeakers. In some, a small coil fed by the sound source is bonded to the membranous diaphragm (**Fig. 1**). Alternating current in the coil thus drives the diaphragm toward and away from a permanent magnet in the rear of the casing. In another configuration, the coil is relatively large and is attached to the diaphragm only at its edge; the diaphragm is a molded dome with a rolled edge and moves freely in the magnetic field. Such earphones are standard for high-fidelity communications.

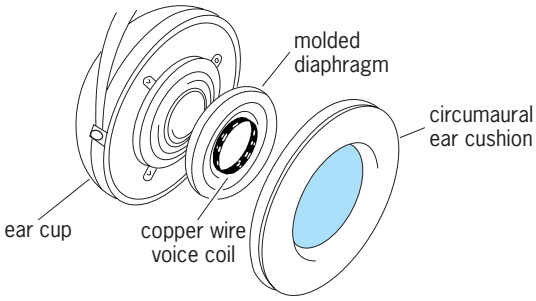

Fig. 1. Exploded view of conventional dynamic earphone. (*After P. Milton, Headphones: History and Measurement, Audio, 26(5):90–94, May 1978*)

Miniature magnetic. A common earphone is a small unit of which the output port fits snugly into a plastic olive in the ear canal (**Fig. 2**). These are widely used with small radios, with the more powerful hearing aids, by television commentators, in business transcription devices, and in many other communications situations. A coil fed by the signal current is wrapped around a pole and imparts in-and-out motion to the diaphragm proportional to the alternating current flow.

Another configuration (**Fig. 3**) is incorporated inside the case of most ear-level and all-in-the-ear hearing aids. A metal armature has a free end between the two poles of a magnet. When alternating current is flowing in one direction, the armature at the free end is turned into a south pole and moves toward the north pole of the permanent magnet. As the current changes direction, the free end is turned into a north pole and the armature moves in the other direction. The armature moves a diaphragm, creating sound which is ported into the ear canal. *See* HEARING AID.

The diaphragm can be of any light material, not necessarily magnetic, and can be formed into an efficient shape for producing and directing an acoustic wave. Such miniature earphones are now made so efficient as to compete favorably with much larger types with respect to acceptably flat frequency

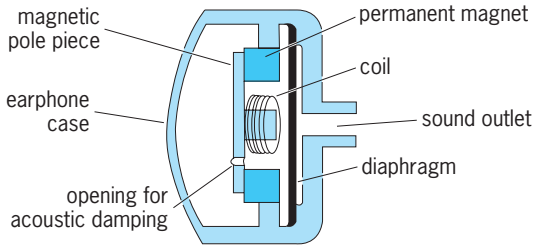

Fig. 2. Cross section of external-type hearing aid earphone. (*After W. R. Hodgson and P. H. Skinner, eds., Hearing Aid Assessment and Use in Audiologic Habilitation, Williams and Wilkins, 1977*)

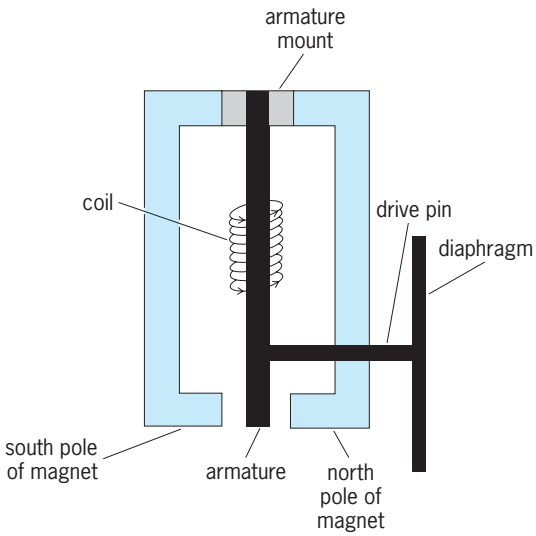

Fig. 3. Cross section of internal-type hearing aid earphone. (*After W. R. Hodgson and P. H. Skinner, eds., Hearing Aid Assessment and Use in Audiologic Assessment, Williams and Wilkins, 1977*)

response in the octave above 10 kHz and with respect to sound power levels in the ear canal with acceptably low total harmonic distortion.

Electrostatic. Efficiency is increased if the mass of the diaphragm is reduced to a minimum. A thin (2.5–12 micrometers) metallized plastic film can be used, on which a large constant electrostatic charge is placed by an auxiliary unit, and the motion of the diaphragm is controlled by the audio signal impressed on perforated wire mesh plates on either side (the push-pull arrangement reduces second-harmonic distortion). The assembly is mounted in a relatively large cavity and coupled to the head by a circumaural cushion (that is, one which completely surrounds the auricle). Such earphones (**Fig. 4**) are light and comfortable and have excellent frequency-response and transient-distortion characteristics.

Dynamic-electrostatic. The diaphragm of a dynamic-electrostatic (or orthodynamic) earphone (**Fig. 5**) is a permanently polarized electret of fluorocarbon. Consequently the need for an added source of polarization voltage (a drawback inherent to electrostatic earphones) is eliminated. In one development, certain problems of such a system are avoided by separating the functions of the diaphragm and the electret. The diaphragm can be of the thinnest (2.5 micrometers) metallized polyester, stretched between two electrets in push-pull. The diaphragm is matched in acoustic impedance very closely to the surrounding air, and the audio signal can be led directly to the electrets. *See* ELECTRET TRANSDUCER.

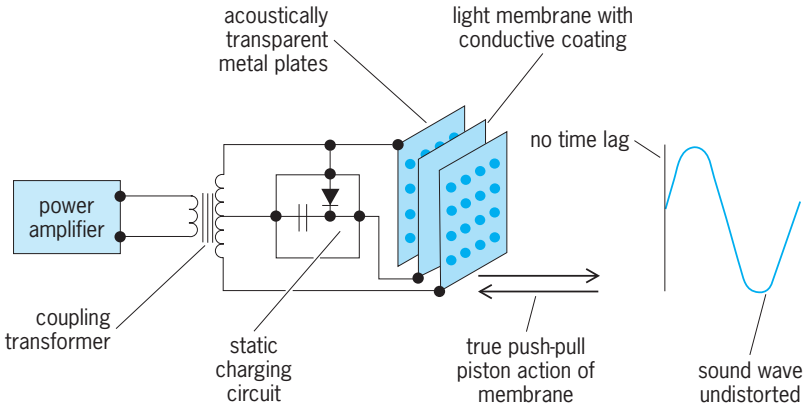

Fig. 4. Block diagram of electrostatic earphone. (*After H. Souther, An Adventure in Headphone Design: The Model ESP-6 Electrostatic Headphone, Koss Electronics Inc., Cat. 111, 1969*)

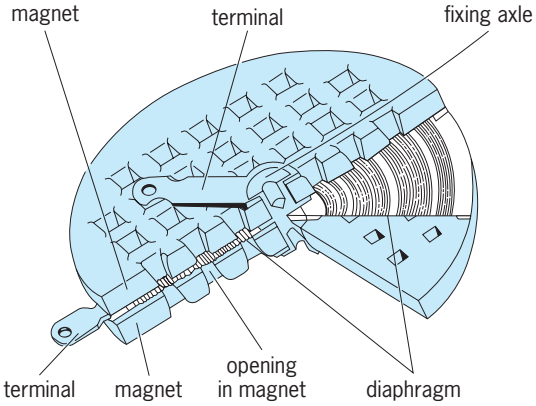

Fig. 5. Construction of orthodynamic earphone driver. (*After P. Milton, Headphones: History and Measurement, Audio, 26*(5):*90–94, May 1978*)

Piezoelectric. Certain crystalline substances expand and contract when alternating voltage is applied. In some piezoelectric earphones, a crystal element is coupled mechanically to the center of a small (about 2.5 cm diameter) cone. Such earphones can be lightweight and cheap and may be acceptable for speech communication.

Substances other than quartz crystals, notably high-polymer films such as vinylidene fluoride, will expand and contract when an alternating current is applied. A thin sheet of such material can be stretched by a factor of 4, and aluminum vapor deposited over it. If a high dc voltage is now applied, and removed, the sheet will become and remain piezoelectric. Two sheets are bonded together back to back, and when alternating current is applied between the metallized faces the sheet will expand and contract. If a rectangular sheet is slightly curved and mounted rigidly top and bottom, much as a playing card might be grasped between the fingers, the movement in response to alternating current will be in and out in proportion and an acoustic wave will be created (**Fig. 6**). Earphones on this principle have desirable characteristics of stability, acoustic response, and cost, but require more amplification than an electret or other earphone to operate as efficiently. *See* PIEZOELECTRICITY.

Real-ear response. An earphone system contains components which are in some respects as important as the driver. The configuration of the acoustic ports, for example, or the acoustic or other damping built into the system will help determine the

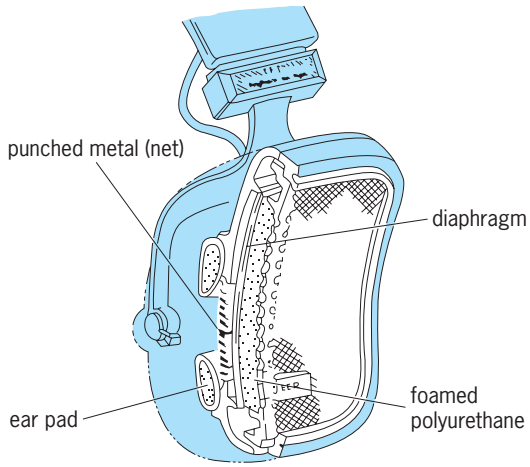

Fig. 6. Cutaway view of rear of high-polymer piezoelectric earphone. (*After P. Milton, Headphones: History and Measurement, Audio*, *26*(5):90–94, *May 1978*)

eveness of the frequency response, while the cushions, whether circumaural or supraaural (fitting against the auricle), help determine the low-frequency real-ear response.

The effects of changes in these elements as reflected in acoustic differences at the eardrums of actual persons are not easy to determine. It is not difficult to couple an earphone system to a microphone by means of a standard cavity, perhaps incorporating a flat plate for the larger circumaural systems, but such a practice is useful primarily for the manufacturer's quality control, and has limited relevance to the performance of the system on even the average human head. The reference equivalent threshold sound-pressure level of a certain earphone system, expressed in voltage applied as frequency changes, has to be determined by having a panel of normal-hearing listeners loudness-balance that system at each frequency against another (arbitrarily chosen) reference system.

The acoustic impedance of the several elements of the external ear, head, and torso have been measured, and artificial heads incorporating the results of such measurements are available. These devices allow fairly successful predictions as to how an earphone system will perform on an anthropometrically average person. *See* ACOUSTIC IMPEDANCE.

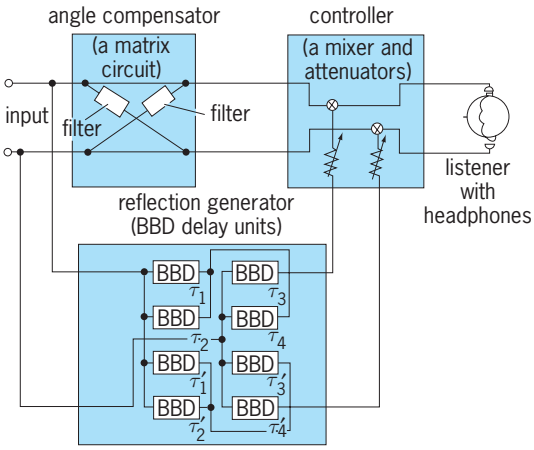

Fig. 7. Block diagram of projected sound localization earphone system for stereophonic sources. BBD-bucket brigade device. (*After N. Sakamoto, T. Gotoh, and Y. Kimura, On "out-of-head" localization in headphone listening, J. Audio Eng. Soc., 24:710–716, 1976*)

Realistic simulation. Any single-channel recording of a real acoustic event will, when played back to an earphone in one ear, sound "in the head." Even stereo recordings from two channels played back to earphones on two ears, while furnishing the illusion of movement, still are not externally localized by the listener. Externalization is improved if some of the signal from the left channel is time-delayed and applied also to the right earphone (the same, of course, for the right channel to the left earphone), and it is also improved if a realistic ratio is achieved between the acoustic energy density ratio from the "direct" versus the "indirect" sound sources (as from reflective walls). Stereophonic earphone systems have been built which incorporate these time delays and ratios, and which furthermore feed the signal from the right channel of an artificial-head stereophonic recording to the left earphone (and the left channel to the right earphone) through frequency filters which simulate the differential acoustic effects at various frequencies of the head and external ears in the original recording situation. Thus the eardrums under earphones are presented with the exact acoustic conditions generated by a loudspeaker or other sound source in an actual room, and such earphone systems very materially advance the important psychoacoustic feature of externalization of sound and of acoustic realism generally (**Fig. 7**). *See* LOUDSPEAKER; MICROPHONE; SOUND-REPRODUCING SYSTEMS; STEREOPHONIC SOUND.

J. Donald Harris

Earth

The third planet from the Sun and the largest of the four inner, or terrestrial, planets. The Sun is an average-sized, middle-aged star situated toward the outer edge of one of the spiral arms of the Milky Way Galaxy. So far as is known, Earth is unique in the solar system in having life. Whether life exists in the universe beyond the solar system is unknown. However, recent discoveries have proven that there are planets in orbit around many nearby stars, so there are possibly billions of Earth-like planets in the universe where life may yet be discovered. *See* GALAXY, EXTERNAL; MILKY WAY GALAXY; PLANET; SOLAR SYSTEM; STAR; SUN; UNIVERSE.

Earth has one natural satellite, the Moon. Otherwise, Earth's nearest neighbors in space are Venus, which is about 108×10^6 km (67×10^6 mi) from the Sun, and Mars, about 228×10^6 km (141×10^6 mi) from the Sun. Earth is about 150×10^6 km (93×10^6 mi) from the Sun. Distances between planets are so immense that it is difficult to comprehend them. To put the solar system into perspective, think of the Sun as a basketball. The planet closest to the Sun, Mercury, would be the size of a grain of sand slightly

less than 1 mm in diameter and 12 m from the basketball. Venus, the second planet out, would be a granule 2.4 mm in diameter, 22 m away, and Earth would be a granule 2.5 mm in diameter, 30 m from the basketball. *See* MARS; MERCURY (PLANET); MOON; VENUS.

Motions. A person standing on the surface of Earth undergoes a complex motion through space. The motion has two principal components, one arising from the orbital motion of Earth around the Sun, the other from the rotation of Earth around its axis.

Orbital motion. Earth completes an orbit around the Sun in 365 days, 5 h, 48 min, 46 s; the orbit defines the length of the year. The length of the day is determined by the period of Earth's rotation about its axis. The fact that the year is not a whole number of days has affected the development of the calendar. *See* CALENDAR.

The orbit is close to circular but is actually elliptical with the Sun at one focus. Earth's closest approach to the Sun (perihelion), which happens about January 3, brings Earth to within 147×10^6 km (91.5×10^6 mi); Earth is at aphelion (farthest from the Sun) about July 4, when Earth is 152×10^6 km (94.5×10^6 mi) away. The average velocity of Earth in its orbit is 107,000 km/h (66,600 mi/h), or 29.6 km/s (18.5 mi/s). Because the orbit is elliptical, Earth's velocity varies systematically, being greatest at perihelion and least at aphelion. The changing velocity affects measurements of solar time. *See* CELESTIAL MECHANICS; ORBITAL MOTION.

Rotation. Earth rotates on its axis once each day. The axis of rotation is perpendicular to the Equator, and the Equator is inclined at about 23.5° to the plane of Earth's orbit around the Sun. As Earth moves in its orbit, the north spin axis, or north geographic pole, points in the direction of the star Polaris, making it the North Star or polestar. One result of the tilt of the Equator relative to the orbital plane is that different parts of Earth receive differing amounts of sunlight through the year; this is the primary cause of seasons. *See* EQUATOR; POLARIS; ROTATIONAL MOTION.

Earth's spin axis has not always pointed in the direction of Polaris. Earth is not a sphere; it bulges at the Equator so that the equatorial radius is 21 km (13 mi) greater than the polar radius. The gravitational attractions of the Sun and Moon on the equatorial bulge cause the spin axis to slowly trace out a cone in space—that is, as time passes, the north spin axis will no longer point toward Polaris. This slow conical motion of the spin axis is called precession, and a complete cycle requires 25,800 years. Superimposed on the precession is a much smaller fluctuation, called nutation, that results from the movement of the Sun twice a year, and the Moon twice a month, across the plane of the celestial equator. *See* EARTH, GRAVITY FIELD OF; PRECESSION OF EQUINOXES.

Earth's period of rotation, the day, is used to define the second, which is the basic unit of time. The length of the day varies irregularly owing to tidal friction, motions in Earth's core, and seasonal atmospheric circulation. For this reason, the second was defined as 1/86,400 of a mean solar day averaged over a year. To avoid the obvious problems in determining the length of the mean solar day, the second is now defined atomically. *See* DAY; TIME.

Earth's speed of rotation is a maximum in late July and early August, and a minimum in April; the difference in the length of the day is about 0.0012 s. Since about 1900, Earth's rotation has been slowing at a rate of about 1.7 s per year. In the geologic past, Earth's rotational period was much faster. Daily, monthly, tidal, and annual growth rings on fossil marine organisms reveal that about 350×10^6 years ago (Middle Devonian Period) the year had 400–410 days, and 280×10^6 years ago (Pennsylvanian Period) the year had 390 days. These observations are very close to the calculated values of the effect of tidal friction. The slowing of Earth's rotation is caused by tidal friction between the sea floor and the ocean water. The Moon is the main cause of tides, and the total rotation energy in the Earth-Moon system is conserved, so that the energy lost by Earth is gained by the Moon. This causes the Moon to move farther from Earth, and this in turn lengthens the period of the Moon's revolution. *See* DEVONIAN; EARTH ROTATION AND ORBITAL MOTION; PENNSYLVANIAN; TIDE.

Satellites. The Moon is Earth's only natural satellite. The Moon's mass is 1/81.3 of Earth's, and its average distance from Earth is 383,403 km (238,247 mi). The center of mass of the Earth-Moon system is within Earth, about 4645 km (2886 mi) from Earth's center, and it is about this point that the Earth-Moon system revolves.

The Moon's period of revolution and rotation is 27 days, 7 h, 43 min, 11.5 s. The orbital plane of the Earth-Moon system is inclined to the orbital plan of the Earth-Sun system at an angle of 5°8′33″. At times the Sun, Earth, and Moon are in a line, and an eclipse of the Sun or the Moon occurs. *See* ECLIPSE; SATELLITE (ASTRONOMY).

Size, shape, mass, and density. Earth is an oblate spheroid. The mean equatorial radius is 6378.139 km (3963.37 mi), and the polar radius is 6356.779 km (3950.10 mi), the difference being 21.360 km (13.27 mi).

Earth's mass is 5.976×10^{27} g (0.2108×10^{27} oz), being the sum of 5.974×10^{27} g (0.2107×10^{27} oz) for solid Earth, 1.4×10^{24} g (0.049×10^{24} oz) for the ocean, and 5.1×10^{21} g (0.18×10^{21} oz) for the atmosphere. Earth's average density is 5.518 g/cm^3, which is just about double the density of the common rocks that form at Earth's surface, indicating that Earth's interior is more dense than the surface. Seismic studies have confirmed that Earth is layered both compositionally and mechanically (**Fig. 1**). *See* ATMOSPHERE; DENSITY; EARTH INTERIOR; OCEANOGRAPHY; SEISMOLOGY.

Internal structure. The deepest compositional layer is the core, which is divided into a solid inner core and a liquid outer core. Both the inner and outer core have the same composition, believed to be nickel-iron plus a small amount of lighter elements such as sulfur and silicon. Electric currents moving in the molten metal outer core are believed to be the origin

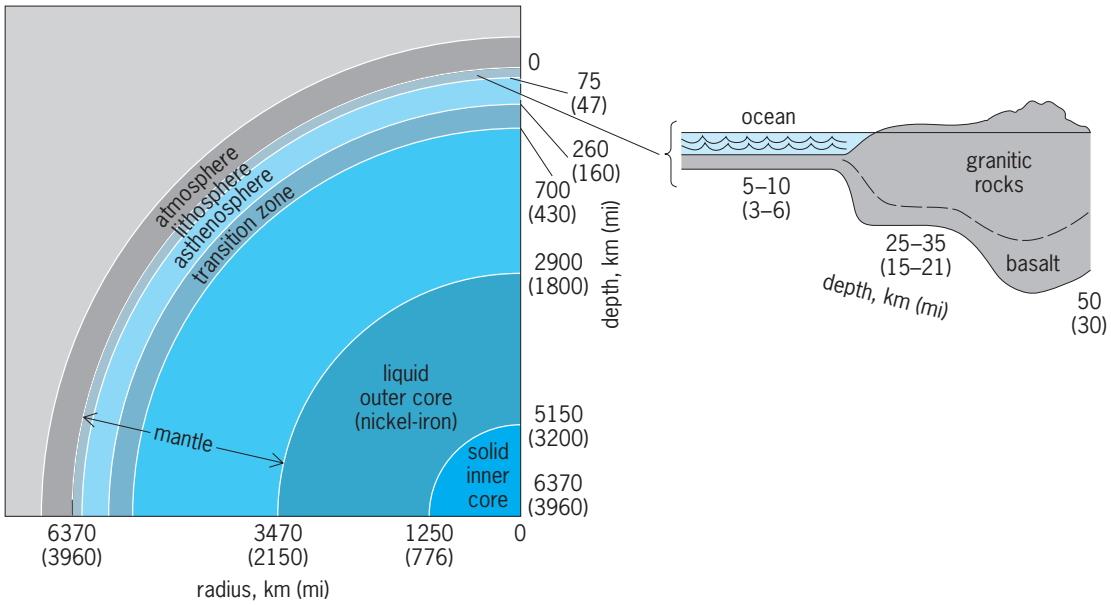

Fig. 1. Principal layers of Earth.

of Earth's magnetic field. Above the core is the mantle which, on the basis of density of rare rock samples brought up from deep in the mantle in kimberlite pipes, and other evidence, is believed to be composed of silicate minerals, and in particular olivine and pyroxene. A rock composed largely of olivine and pyroxene is called a peridotite. *See* OLIVINE; PERIDOTITE; PYROXENE; SILICATE MINERALS.

Within the mantle there are two prominent seismic discontinuities, one at a depth of about 400 km (248 mi), the other at a depth of about 670 km (416 mi). The seismic discontinuities are marked by pronounced changes in the velocities of seismic waves. The 400-km discontinuity is presumed to be due to the breakdown of pyroxene to garnet plus stishovite, and the transformation of olivine to a spinel structure with a consequent increase in density of about 5%. The 670-km seismic discontinuity is thought to be due to the breakdown of olivine spinel to perovskite and periclase, and the transformation of garnet to a perovskite structure. The lower mantle, from 670 km to the core-mantle boundary at a depth of 2885 km (1793 mi), does not exhibit any pronounced seismic discontinuities. *See* GARNET; PEROVSKITE; SPINEL; STISHOVITE.

Above the mantle is Earth's crust, and between the crust and the mantle there is a pronounced seismic discontinuity known as the Mohorovičić discontinuity, or Moho. The crust is of two kinds, both of which are less dense and compositionally different from the peridotitic mantle below. Beneath the ocean the crust is basaltic in composition and about 8 km (5 mi) thick. The crust beneath the continents is granitic in composition and averages 35 km (21.7 mi) in thickness but ranges up to 80 km (49.7 mi), as beneath Tibet. The oceanic crust is geologically young because it is continually created and destroyed through the process of plate tectonics. No part of the oceanic crust that is older than about 180×10^6 years has yet been discovered. The continental crust is much older than the oceanic crust. Continental rocks as old as 4×10^9 years have been discovered in Canada, and the fact that they are highly deformed indicates a long and eventful history. *See* EARTH CRUST; GRANITE; PLATE TECTONICS.

Mechanical layering. The internal temperature of Earth increases from an average of 15°C (59°F) at the surface to about 5000°C (9032°F) at the center of the core, approximately the same temperature as the surface of the Sun. Measurements show that the thermal gradient beneath the ocean is about 13°C/km (69.7°F/mi), while that in the continental crust is about 6.7°C/km (51.4°F/mi). The combined influence of temperature and pressure are such that rocks in the crust and upper mantle are mechanically strong and rigid up to a temperature of about 1300–1350°C (2372–2462°F), equivalent to a depth of about 100 km (62 mi) beneath the oceans and 200 km (124 mi) beneath the continents. The layer of mechanically strong rocks, which includes the uppermost 92 km (57 mi) of the mantle beneath the oceanic crust and the uppermost 165 km (103 mi) of the mantle beneath the continents, is called the lithosphere. Below the lithosphere, rocks lose strength and can be readily deformed because temperatures and pressure almost reach melting conditions of mantle rocks. A layer of weak rocks about 100 km thick and known as the asthenosphere underlies the lithosphere. The mechanical strength of the lithosphere is such that heat moves through it by conduction, but the strength of the asthenosphere is such that heat can move through it by convection. Beneath the asthenosphere both temperature and pressure increases with depth, but the temperature gradient is only about 0.5°C/km (1.4°F/mi) because it is determined by convection. As a result, the strength of rocks in the middle and lower mantle slowly increases, but at all depths below the

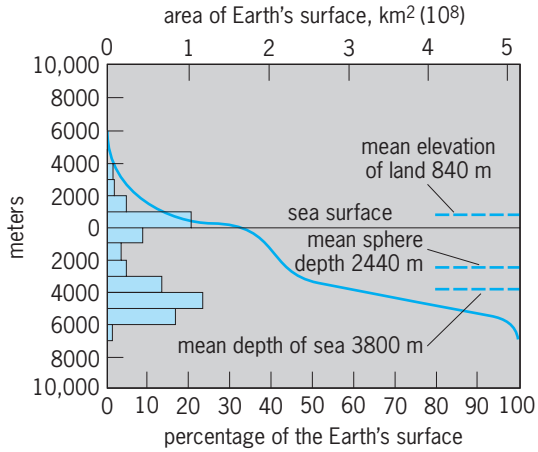

Fig. 2. Hypsographic curve of land surface elevation. 1 m = 3.3 ft; 1 km^2 = 0.4 mi^2. (*After E. Kossinna, Die Tiefen des Weltmeeres, Veröff. Inst. Meereskunde Univ. Berlin, Neue Folge A. Geogr.-naturwiss., 9:1–70, 1921*)

lithosphere mantle rocks are capable of flow so that heat moves outward from the core to the base of the lithosphere by convection. *See* ASTHENOSPHERE; CONDUCTION (HEAT); CONTINENTS, EVOLUTION OF; EARTH, CONVECTION IN; LITHOSPHERE.

Crustal structures and surface features. The surface of solid Earth has a bimodal distribution of elevations (**Fig. 2**). If the water from the ocean could be removed, it would be apparent that continents stand high (average elevation is 840 m or 2755 ft above sea level), while the ocean floor sits low (average elevation is 3800 m or 12,464 ft below sea level). This difference in elevation arises because rigid lithosphere floats on the weak asthenosphere, and because the density of oceanic lithosphere (that is, lithosphere capped by oceanic crust) is greater than the density of continental lithosphere.

On the continents, mountain belts are the most dramatic features. They range in elevation from Mount Everest, 8848 m (29,030 ft), in the Himalaya Mountains to older, deeply eroded ranges that are now barely above sea level. Granitic and metamorphic rocks are generally exposed in the cores of mountain ranges. The overlying rocks that cover most of the Earth's surface are sedimentary, mainly of shallow marine origin, that may or may not have been deformed. The deformation is the result of compression and tension that causes folding and faulting, and may be accompanied by intrusion and metamorphism. Movements and collisions of tectonic plates are the principal cause of mountain building. Mountains generally are formed over several tens of millions of years. The rocks deformed in the process are generally marine sedimentary rocks formed along the margins of continents. *See* DEEP-MARINE SEDIMENTS; MARINE SEDIMENTS; METAMORPHIC ROCKS; OROGENY; SEDIMENTARY ROCKS.

Some topographic mountains are formed by volcanoes. The most striking are long chains of andesitic stratovolcanoes, such as the Aleutian and Andean chains, formed as a result of melting in the mantle induced by sinking tectonic plates. *See* VOLCANO.

Much of the surface of the continents is covered by a thin veneer of sedimentary rocks. Where the underlying rocks of the plains and hills that make up most of the continents are exposed, the rocks and their structures are similar to those found in most mountain ranges. This leads to the theory that continents are formed by the deeply eroded remnants of earlier mountain ranges, and that today's continents have been assembled by accretion and by plate collisions.

The topographic features underlying the oceans are similarly diverse and reveal more evidence of a dynamic Earth. The continental shelf, an area covered by shallow water, generally less than 150 m (500 ft) deep, surrounds the continents at most places. Such areas are generally underlain by continental, granitic rocks, and are submerged parts of the continents. Continental slopes are the transition between the continental shelf and the ocean floors. Their tops are generally less than 150 m below sea level, and they slope down to about 4400 m (14,000 ft). They are narrow, steep features, with slopes generally between 2 and 6°, but some are up to 45°. They are generally underlain by thick accumulations of sedimentary rocks. *See* SEA-FLOOR IMAGING.

Submarine trenches and their associated volcanic island arcs are formed as a result of a tectonic plate of lithosphere sinking into the mantle beneath the edge of an overriding plate. The deepest place on Earth is in the Mariana Trench, 11,022 m (36,152 ft) below sea level.

The ocean floor is the most widespread surface feature of Earth. Beneath an average of 4.4 km (2.75 mi) of seawater are about 2.3 km (1.4 mi) of sedimentary rocks with some intercalated basalt, and below that is the oceanic crust, consisting of 4–6 km (2.5–3.7 mi) of basaltic rocks. Interrupting the ocean floor at many places are submarine mountains formed by basaltic volcanoes. Some of these volcanoes are very large and form oceanic islands such as the Hawaiian Islands. At other places, vertical movements of the ocean floor are revealed by sunken islands whose flat tops were formed by wave erosion (guyots) and coral atolls.

The ocean floor rises gradually to the mid-ocean ridges, a more or less continuous feature through all the oceans with some branches and offsets. The

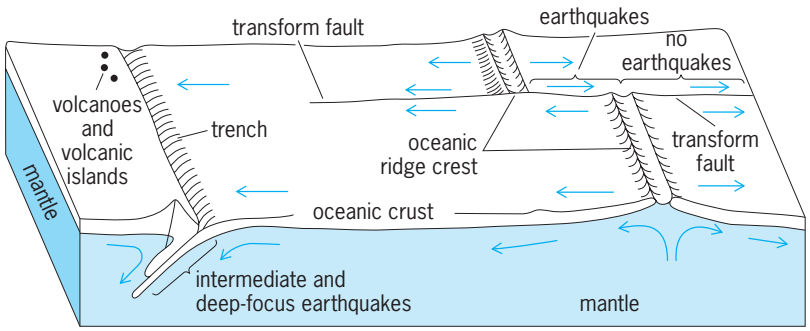

Fig. 3. Sea-floor spreading. (*After G. Gross, Oceanography, 3d ed., Charles E. Merrill, 1976*)

ridges range between 480 and 4800 km (300 to 3000 mi) wide and are much more rugged and irregular than the ocean floors. They rise about 3000 m (10,000 ft) above their base on average. Lines of parallel volcanoes, steep scarps, and a central valley mark the mid-ocean ridges. The central parts of the generally more or less symmetrical ridges are the most active volcanically and seismically. *See* MARINE GEOLOGY; MID-OCEANIC RIDGE.

The source of the heat energy that deforms the surface rocks is the crystallization of the liquid core and the decay of radioisotopes within the whole Earth. The energy released at the bottom of the mantle may be the cause of convection in the mantle, which in turn causes plate tectonics and so deforms the crust, but the actual processes are probably much more complex and are not yet clearly understood. *See* RADIOISOTOPE (GEOCHEMISTRY).

Plate tectonics. New oceanic lithosphere capped by basaltic crust is created at the mid-ocean ridges, and this newly formed plate moves away from the ridges (**Fig. 3**). The tectonic plates formed in this way may carry continents on them, and are the mechanism of continental drift. Paleomagnetic data from the continents indicate that the continents have moved relative to each other. The tectonic plates capped by basaltic crust plates are consumed at the trench–volcanic island arc areas. *See* PALEOMAGNETISM; SUBDUCTION ZONES.

As well as the ridge and the trench, a third type of plate boundary occurs where two plates slide past each other at a transform fault. Moving plates may collide in several ways (**Fig. 4**). Such collisions account for the deformed rocks found in the crust. *See* TRANSFORM FAULT.

The evidence for continental drift in the geological past includes matching of rock types, ages, fossils, climates, and structures (mountain ranges), as well as the paleomagnetic data. Evidence showing or suggesting present movements consists of shallow earthquakes along mid-ocean ridges and transform faults that offset them; deep earthquakes associated with deep-sea trench–volcanic island arc areas; direct measurement of movement; volcanic activity at mid-ocean ridges; and volcanic activity at trench–island arc areas. *See* GEODESY.

Atmosphere. Earth's temperature and gravitation are such that an atmosphere is present. The major constituents are nitrogen and oxygen. The atmosphere, especially oxygen, and the presence of water, both at the surface and in the atmosphere, make life possible. Precipitation, mainly rain, results in running water such as streams and rivers on the continents. Running water is the main cause of erosion of the continents, and most of the landscapes have been eroded by water, although some are eroded by wind or ice (glaciers). *See* EROSION; GLACIOLOGY; GRAVITY; NITROGEN; OXYGEN; WATER.

The atmosphere shields Earth from most meteorites. Meteors are caused by frictional heat when such high-velocity objects move through the atmosphere. A few of these objects pass through the atmosphere and strike Earth, causing craters. Other bodies, such as the Moon, are not protected by an atmosphere and so are struck more frequently by meteorites. The craters formed on the Moon are preserved, because without an atmosphere almost no erosion occurs. *See* METEORITE.

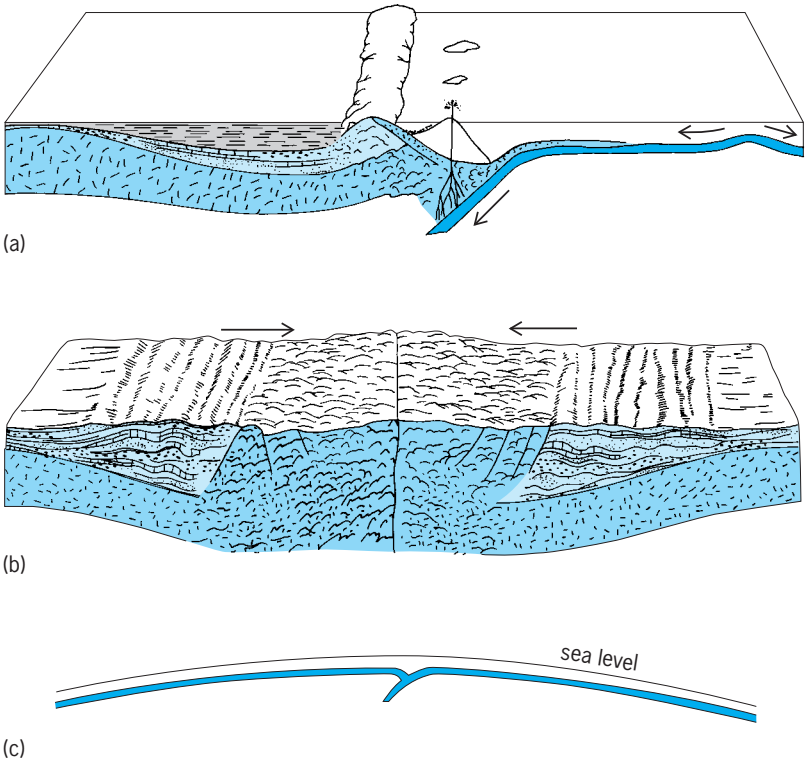

Fig. 4. Plate collisions: (*a*) continent-ocean plate collision; (*b*) continent-continent plate collision; (*c*) ocean-ocean plate collision. (*After R. J. Foster, Physical Geology, 3d ed., Charles M. Merrill, 1979*)

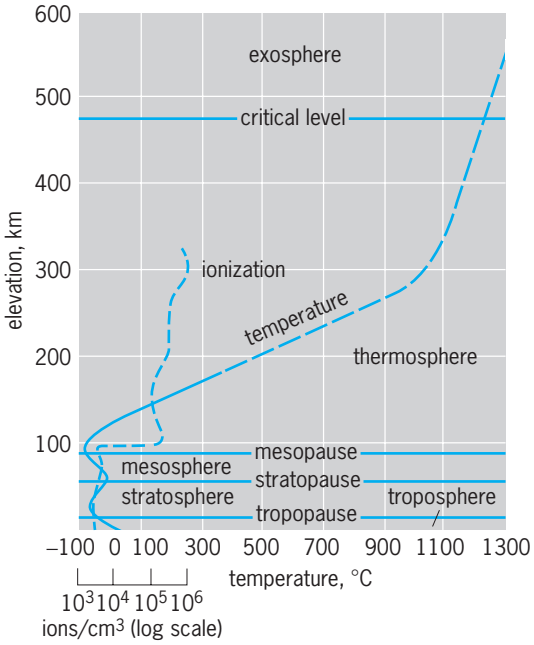

Fig. 5. Structure of the atmosphere. The log scale applies only to the ionization curve. 1 km = 0.6 mi; °F = (°C × 1.8) + 32.

A thin ozone layer in the atmosphere also shields the Earth from lethal ultraviolet radiation from the Sun. The structure and composition of the atmosphere is shown in **Fig. 5**. *See* OZONE.

Age. Earth, along with the rest of the solar system, is believed to have formed about 4.55×10^9 years ago. This age is determined by dating radioactive isotopes in meteorities. Meteorites are believed to be fragments produced by collisions among small bodies formed by the same process that created the solar system. Theoretical studies of the Sun and other studies of radioactive isotopes also suggest a similar age. *See* EARTH, AGE OF. Brian J. Skinner

Bibliography. B. W. Murck and B. J. Skinner, *Geology Today: Understanding Our Planet*, 1999; F. Press and R. Siever, *Understanding Earth*, 2d ed., 1998; B. J. Skinner and S. C. Porter, *The Dynamic Earth*, 4th ed., 2000.

Earth, age of

An estimate of the age of the planet Earth based on the lead isotope systematics. The first meaningful estimate was calculated from the average isotopic composition of terrestrial lead (Pb) available at the surface of the planet. The feasibility of the calculation stems principally from the unique circumstance that two isotopes of the element uranium (^{235}U and ^{238}U) decay to two daughters that are also isotopes of another element, lead (^{207}Pb and ^{206}Pb, respectively). This eliminates the requirement of determining the planetary budget of either of the two elements. The total number of ^{206}Pb and ^{207}Pb atoms are related to their respective parents by the equation of radioactive decay, expressed as Eq. (1) for lead-206 and Eq. (2) for lead-207, where ^{238}U and ^{235}U are the present planetary content of the two isotopes whose decay constants are λ and λ' respectively, T is the age of the planet, and subscript p denotes primordial Pb, which existed in the solar nebula prior to the formation of planets. *See* LEAD ISOTOPES (GEOCHEMISTRY).

$$^{206}\mathrm{Pb} = (^{206}\mathrm{Pb})_\mathrm{p} + {}^{238}\mathrm{U}(e^{\lambda T} - 1) \qquad (1)$$

$$^{207}\mathrm{Pb} = (^{207}\mathrm{Pb})_\mathrm{p} + {}^{235}\mathrm{U}(e^{\lambda' T} - 1) \qquad (2)$$

Of the four lead isotopes, ^{208}Pb, ^{207}Pb, ^{206}Pb, and ^{204}Pb, the last is not produced in any known process of natural radioactive decay, and is used as an index isotope. For example, Eq. (1) takes the form of Eq. (3), where $\mu \equiv (^{238}\mathrm{U}/^{204}\mathrm{Pb})_{\mathrm{today}}$, and the subscript m denotes a measured ratio of a representative sample. In the Pb-Pb method, the two radiometric clocks expressed in Eqs. (1) and (2) are combined to yield Eq. (4), where 1/137.9 is the present ratio of ^{235}U/^{238}U, a constant reflecting the nonfractionation of the two isotopes. Primitive Pb, supposedly primordial, was found in the troilite phase of the Cañon Diablo iron meteorite, and the composition ratios of such Pb are used in Eq. (4), in which T is then the only unknown. Assuming the existence of a sample representative of the Earth as a whole, it is possible to calculate the actual age of the Earth merely on the basis of the Pb isotopic composition of such a sample. Usually what is calculated is an approximate age for the Earth, because in view of the extensive global and regional differentiation which the Earth has experienced, the assumption of a representative sample is an oversimplification.

$$\left(\frac{^{206}\mathrm{Pb}}{^{204}\mathrm{Pb}}\right)_\mathrm{m} - \left(\frac{^{206}\mathrm{Pb}}{^{204}\mathrm{Pb}}\right)_\mathrm{p} = \mu(e^{\lambda T} - 1) \qquad (3)$$

$$\frac{(^{207}\mathrm{Pb}/^{204}\mathrm{Pb})_\mathrm{m} - (^{207}\mathrm{Pb}/^{204}\mathrm{Pb})_\mathrm{p}}{(^{206}\mathrm{Pb}/^{204}\mathrm{Pb})_\mathrm{m} - (^{206}\mathrm{Pb}/^{204}\mathrm{Pb})_\mathrm{p}} = \frac{1}{137.9}\left(\frac{e^{\lambda' T} - 1}{e^{\lambda T} - 1}\right) \qquad (4)$$

Single-stage model. A calculation based on Eq. (4) represents a single-stage model where terrestrial Pb is assumed to have evolved (through the decay of U) in a single stage (of duration T) from an initial Pb composition identical to the nebula's primordial Pb. It is thus implicit in the model that the Earth accreted directly, or nearly so, from the solar nebula.

In his pioneering work, C. C. Patterson chose oceanic sediments as the representative sample, to which he applied Eq. (4). Repeating that calculation but using the more newly determined values for decay constants and primordial Pb yields Earth age (T) = 4.43×10^9 years and the ratio ^{238}U/^{204}Pb (μ) = 9.3. This terrestrial μ value is quite high when compared with estimated $\mu < 1$ for the solar nebula and $\mu \sim 0.2$ inferred for the source of the Allende chondrite. If it is assumed that the Earth's average composition is chondritic ($\mu < 1$), then it follows from the model that the high terrestrial μ is the result of chemical fractionation that took place in conjunction with the formation of the planet about 4.43×10^9 years ago.

The mechanism of such fractionation is subject to a wide range of speculations. It could be argued that the high terrestrial μ value resulted from loss of Pb by volatilization, possibly during accretion. This conjecture appears in general accord with other observations inferring depletion of the Earth in the volatiles. Alternatively, the high μ may be restricted to an outer layer, the counterpart of which is, perhaps, a deep interior sulfide region which had scavenged primordial Pb from the accreting materials. Since no terrestrial Pb reflecting low μ (say 1 to 3) has ever been sampled, it follows that convection within the Earth may be ineffective in transporting material from the deep interior. *See* EARTH, CONVECTION IN; EARTH INTERIOR.

If, in contrast to the above, $\mu \sim 9.3$ is assumed to be characteristic of the primitive material prior to its accumulation to form the planet, then the calculated age would apply to the primitive material

and not necessarily to the Earth (which could have accumulated at a much later date). This possibility, however, was generally dismissed by geochronologists because it rested on the improbable premise that either the Earth accumulation and subsequent differentiation did not result in U-Pb fractionation, or chemical fractionation did occur but the resulting spectrum of μ values was averaged on the surface of the planet back to the original μ value. Not only were these implausible scenarios dismissed, but a general belief that the Earth is as old as the meteorite (4.56×10^9 years) prevailed, despite the lack of direct evidence. Subsequently, the Apollo program yielded strong evidence for the axiom of a single age for planets and meteorites: differentiated lunar rocks as old as 4.55×10^9 years were found.

Two-stage model. The axiom of 4.56×10^9 years' age applicable to planets and meteorites requires that the younger single-stage age of the Earth is an artifact resulting from the shortcomings of the model which does not take into account the length of the duration of accretion before the onsetting of U-Pb fractionation, and the multistage differentiation history of the planet.

In partial accounting for the complexities, a two-stage model may be applied and may be expressed as Eq. (5), where T_A and T_D are the times of accretion

$$\left(\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right)_{\text{m}} - \left(\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right)_{\text{p}} = \mu_1(e^{\lambda T_{\text{A}}} - e^{\lambda T_{\text{D}}}) + \mu_2(e^{\lambda T_{\text{D}}} - 1) \quad (5)$$

and differentiation respectively, and the subscripts 1 and 2 refer to the first and second stage. For simplicity, T_A may be fixed at 4.56×10^9 years. The three unknowns (μ_1, μ_2, and T_D) are solved for in conjunction with knowledge of the ages of uranium-free systems, for example, galenas. The two-stage age of a galena may be given by Eq. (6), where g denotes a galena which was deposited T_g years ago.

$$\left(\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right)_{\text{m}} - \left(\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right)_{\text{p}} = \mu_1(e^{\lambda T_{\text{A}}} - e^{\lambda T_{\text{D}}}) + \mu_2(e^{\lambda T_{\text{D}}} - e^{\lambda T_{\text{g}}}) \quad (6)$$

The intricate connection of the parameters is explained with help of the **illustration**. In this schematic presentation, 1, 2, 3, and 4 represent four meteorites falling on a straight line whose slope corresponds to 4.56×10^9 years. The farther away from primordial lead, the higher is the μ of the sample. Modern terrestrial Pb represents the average Pb composition as inferred from surface samples. The slope of the broken line corresponds to the single-stage age of the Earth as given by Eq. (4) According to the two-stage model, modern terrestrial lead did not evolve directly from primordial lead, but rather was shifted from the meteoritic line m as a result of U enrichment in a magma reservoir. The time of that differentiation event is given by the slope of the dotted line which intersects line m at X. The significance

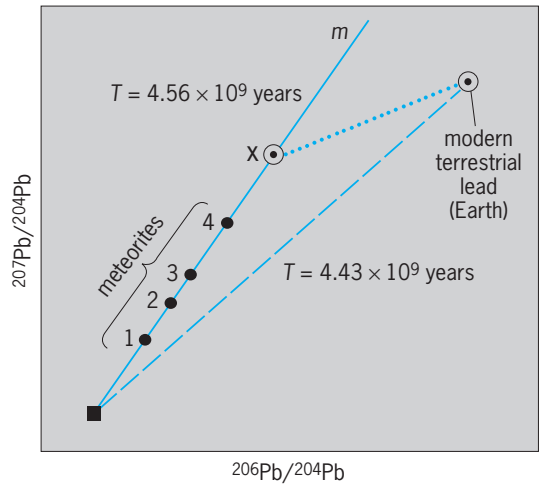

Schematic diagram of Pb isotopes showing 1–4, which define a line *m*, with a slope corresponding to 4.56 × 10⁹ years. Earth plots off the meteoritic line.

of X is that it gives the present isotopic composition of modern terrestrial lead if the differentiation event did not occur. In other words, point X would yield the correct single-stage age, given by Eq. (7).

$$\left(\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right)_{\text{X}} - \left(\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right)_{\text{p}} = \mu_{\text{X}}(e^{\lambda T} - 1) \quad (7)$$

Note that μ_{X} equals μ_1 of Eq. (5) [as it is the ratio of $^{238}\text{U}/^{204}\text{Pb}$ prior to the differentiation episode], and $T = 4.56 \times 10^9$ years.

The following steps are involved in the calculation: For an assumed value for T_D (that is for an assumed slope for the dotted line in the illustration) the corresponding composition of undifferentiated earth (point X) is obtained from the illustration and the value of μ_{X} (that is, μ_1) is calculated from Eq. (7). This is entered as μ_1 into Eq. (5) from which μ_2 is then calculated. Finally all three parameters, μ_1, μ_2 and T_D are placed in Eq. (6) to yield a calculated T'_g value. The latter is compared with the known accepted deposition age, T_g, of the galena. The proper parameters are the ones which lead to $T'_g = T_g$.

Using these relationships in conjunction with a large number of conformable galenas and feldspars, the following parameters were calculated for Earth: $T_D = 3.7 \times 10^9$ years, $\mu_1 = 7.2$, and $\mu_2 = 9.7$.

Congruency of conformable galenas. The significance of the two-stage model discussed is in revealing the possibility of a global differentiation episode long after the Earth's accretion. The method of calculation, however, suffers from two main weaknesses: the age of the Earth is assumed rather than determined, and the composition of modern terrestrial lead is not a well-defined entity. In an attempt to overcome these difficulties, the so-called congruency method was put forward. The method provides a simultaneous solution to the ages of cognate galenas as well as the age of the source from which they were derived. Application of the method to conformable galenas shows that they were derived from

two distinct ancient sources, A and B, whose single-stage ages are $T_A = 4.54 \times 10^9$ years and $T_B = 4.43 \times 10^9$ years. Oceanic sediments appear intimately associated with source *B*, which explains the young age obtained from Patterson's calculation.

Sources A and B appear to be counterparts derived from a parental source, denoted AB, through a process involving U depletion in one case ($\mu A = 7$) and its enrichment in the other ($\mu B = 9$) relative to the parental source ($\mu AB = 8$). Significantly, the indicated time of this differentiation is approximately 3.6×10^9 years ago, which appears to substantiate the first-order observation of the T_D calculation. In view of such differentiation, the single-stage ages of sources A and B are gross approximations bracketing the actual age of the primary parental source AB. The age of the latter, as obtained from Eq. (4), is 4.50×10^9 years. This suggests, as an upper limit, an accretion duration of approximately 6×10^7 years prior to the onsetting of planetary differentiation resulting in the formation of the parental source.

Conclusion. There exist about half a dozen modern calculations of the Earth's age based on the Pb isotope systematics. The results fall in the narrow range of $4.43–4.57 \times 10^9$ years, yielding an average age of 4.50×10^9 years $\pm 0.07 \times 10^9$ years. With a dispersion of a mere $\pm 1.5\%$, the age question seems to be reasonably solved, but active research in this area of geochemistry continues. The primary objective, however, is not the determination of the age with ultimate accuracy, but rather the elucidation of the evolutionary history of the planet. In particular, this type of isotopic research entails the potential of providing strict geochemical constraints which would have to be accommodated by the geophysical models. *See* GEOCHRONOMETRY; METEORITE; RADIOACTIVITY; ROCK AGE DETERMINATION.

Fouad Tera

Bibliography. G. B. Dalrymple, *Age of the Earth*, 1994; A. J. Gancarz and G. J. Wasserburg, Initial Pb of the Amitsog gneiss, West Greenland, and implications for the age of the Earth, *Geochim. Cosmochim. Acta*, 41:1283–1301, 1977; J. S. Stacey and J. D. Kramers, Approximation of terrestrial lead isotope evolution by a two-stage model, *Earth Planet.Sci. Lett.*, pp. 207–221, 1975; F. Tera, Congruency of comformable galenas: Age of the Earth, *12th Lunar and Planetary Science Conference*, pp. 1088–1090, 1981; G. Tilton and R. Steiger, Lead isotopes and the age of the Earth, *Science*, 150:1805–1807, 1965.

Earth, convection in

The notion that thermal convection currents must exist in the Earth's mantle occupies a central position in theories which have been proposed to account for the occurrence of continental drift and sea-floor spreading. Apart from the controlling influence which these currents exert on surface geological processes such as mountain building, they also govern the rate at which heat is transferred from the hot interior of the Earth to its surface and so are crucial to understanding the thermal history of the planet. Mantle convection serves as the guiding principle in the geological and geophysical sciences. *See* CONTINENTS, EVOLUTION OF.

Sea-floor spreading and continental drift. Although even the earliest geographers were impressed by the apparent fit of the coastlines of South America and Africa, it was not until the methods of paleomagnetism had been developed that a plausible case could be made that these two continents had once been joined together as part of the supercontinent Pangea over 100 million years ago. The most important advance in understanding the agency through which breakup of this supercontinent occurred was made in the early 1960s in studies of the pattern of magnetic anomalies observed over a mid-oceanic ridge. The characteristic pattern of magnetic stripes (**Fig. 1**) could be understood in terms of the sea-floor spreading hypothesis, which had been advanced a few years previously. It was suggested that hot material rising from the Earth's mantle beneath the ridge crest acquired an induced magnetization as it was cooled below the Curie point in the Earth's local magnetic field. The material was normally or reversely magnetized, depending upon the polarity of the field at the time of cooling. Although the reason for the episodic reversal of the polarity of the field is not understood in detail, the fact that such reversal occurs allows the sea floor, as it spreads horizontally from a mid-oceanic ridge, to produce a magnetic tape recording of its own large-scale horizontal motion. When the spatial pattern of reversals which is recorded on the sea floor in this way is combined with the time scale for the same reversals established on land, using radiometric dating methods, the pattern of magnetic stripes can be used to measure the horizontal spreading rate as a function of location. When this is done for the globe-encircling pattern of oceanic ridges, local spreading rates are obtained that vary from 0.4 to 4 in. (1 to 10 cm) per year. *See* DATING METHODS; MID-OCEANIC RIDGE; PALEOMAGNETISM; ROCK AGE DETERMINATION; ROCK MAGNETISM.

Paleomagnetic methods have therefore been instrumental in establishing that hot mantle-derived material is continuously rising under and spreading horizontally from the mid-oceanic ridges. That cold and relatively old material is elsewhere descending into the mantle has been demonstrated in an equally convincing fashion by using the methods of seismology. Prior to the advance of the spreading hypothesis, it was suggested that the deep ocean trenches were regions where ocean floor was being downthrust into the mantle along Benioff zones (**Fig. 2**). These zones are defined by the internal locus of deep seismicity. This process is referred to as the subduction of oceanic lithosphere. Taken together with the evidence of spreading, the occurrence of subduction constitutes strong evidence for the existence of a thermally induced circulation in the Earth's mantle. Such a circulation is by definition one which is maintained against dissipative frictional processes by the continuous conversion of potential into

kinetic energy through the action of the buoyancy force. This conversion is effected whenever hot material rises (spreading) and cold material sinks (subduction). The store of potential energy upon which the circulation continuously draws is found in the large increase of temperature between the Earth's surface and the core-mantle boundary. *See* LITHOSPHERE; SUBDUCTION ZONES.

Among the other lines of evidence which have combined to strongly reinforce this model, perhaps the most important concerns the characteristic variations of heat flow and bathymetry across the ocean floor as a function of ocean-floor age. From the many thousand heat flow measurements which have been made in boreholes on the ocean bottom, it is clear that heat flow varies inversely as the square root of the age of the sea floor. Near the mid-oceanic ridges, the heat flow achieves a maximum value which is in excess of 400 mW · m^{-2}, while the flow of heat through the oldest sea floor (age less than 150×10^6 years) is between 40 and 45 mW · m^{-2}. These numbers should be compared with the mean heat flow through the continents of about 60 mW · m^{-2}. Ocean-floor depth (bathymetry) varies in a similarly systematic fashion. The shallowest depths are found above the mid-oceanic ridges, and depth increases with the square root of the age of the ocean floor,

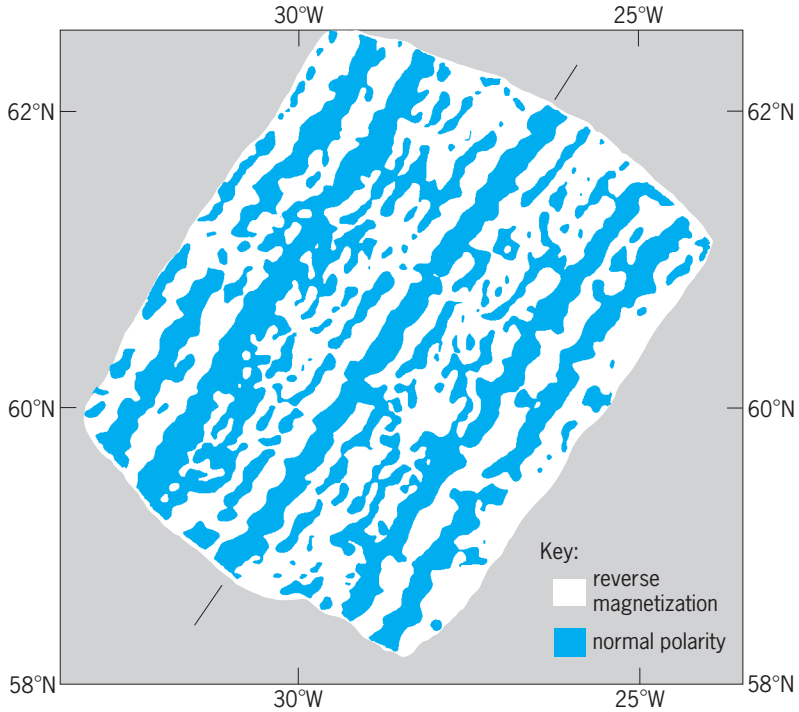

Fig. 1. Pattern of magnetic stripes over the Reykjanes Ridge southwest of Iceland. (*After J. R. Heirtzler, X. LePichon, and J. G. Baron, Magnetic anomalies over the Reykjanes Ridge, Deep Sea Res., 13:427, 1966*)

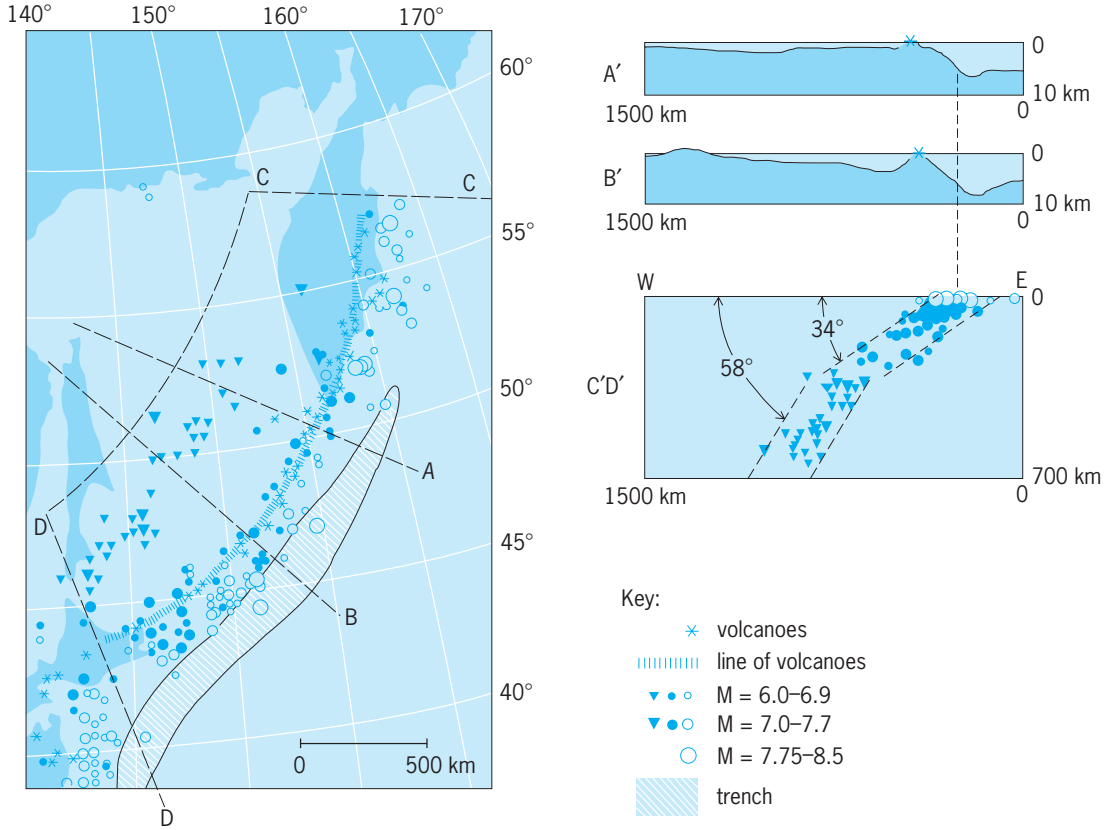

Fig. 2. Geographic setting of the Kurile Trench off the Kamchatka Peninsula, showing the deep ocean trench and the characteristic line of volcanoes in the vicinity of the subduction zone. The sections A′ and B′ illustrate the bathymetry normal to the trench axis, while the section C′D′ shows the localization of earthquake epicenters along the Benioff zone. The surface projections of all earthquake locations are shown in the geographic plan, and their magnitudes *M* are represented symbolically. 1 km = 0.6 mi. (*After H. Benioff, Orogenesis and deep crystal structure: Additional evidence from seismology, Bull. Geol. Soc. Amer., 65:385, 1954*)

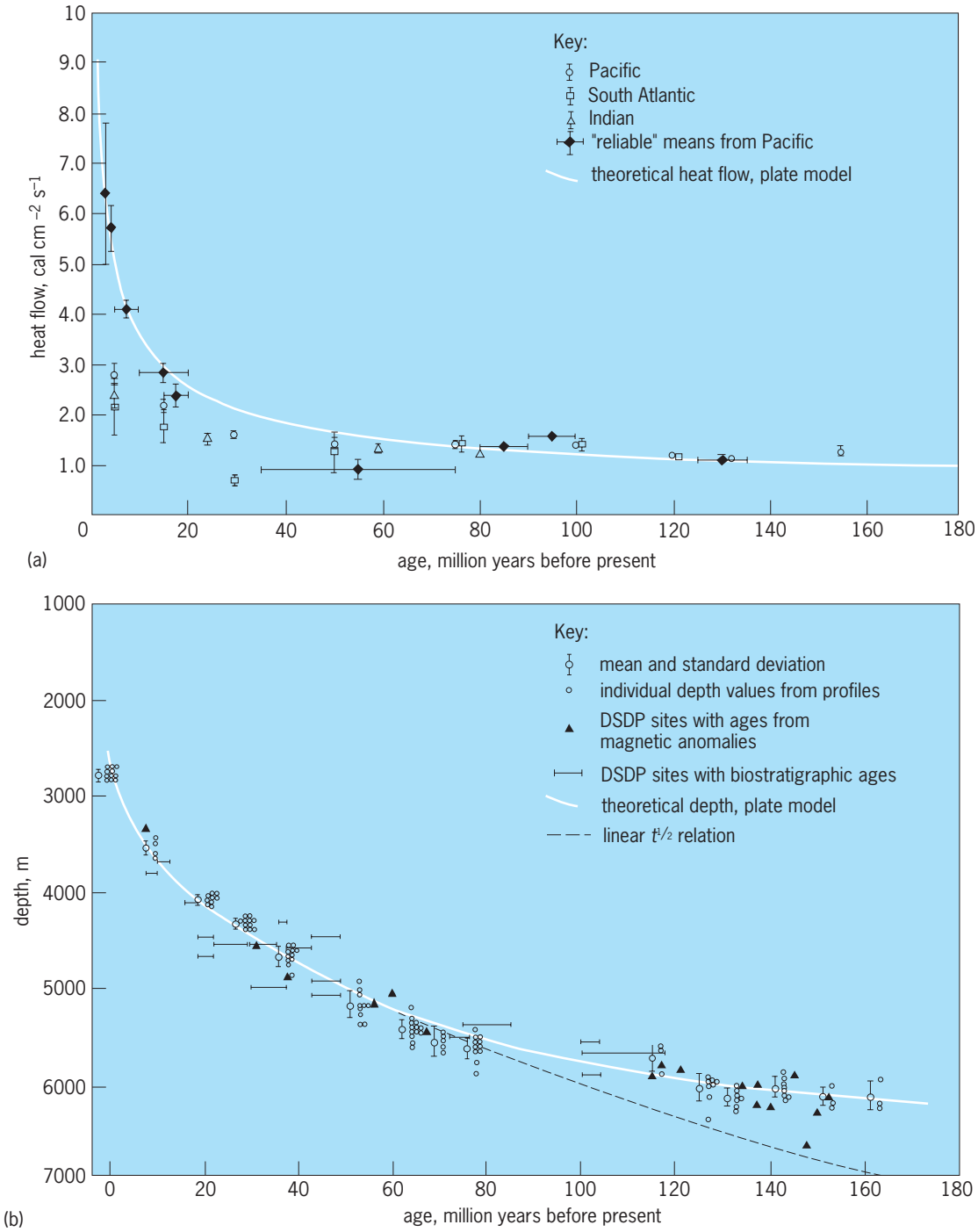

Fig. 3. Characteristic variation of heat flow and bathymetry. (*a*) Variation of surface heat flow as a function of age of the ocean floor in millions of years before present for the Pacific, South Atlantic, and Indian oceans. (*b*) Variation of sea-floor bathymetry with age of the ocean floor for the North Pacific Ocean. For ages less than approximately 80 × 10^6 years, the depth curve follows the simple $t^{1/2}$ decay which is predicted by boundary layer theory for heated-from-below convection. 1 m = 3.3 ft; 1 cal · cm^2 s^{-1} = 42 kW · m^{-2}. (*After B. Parsons and J. G. Sclater, An analysis of the variation of ocean floor heat flow and bathymetry with age, J. Geophys. Res., 82:803, 1977*)

reaching values in excess of approximately 2 mi (3 km) in the oldest oceans (**Fig. 3**). These characteristic variations are precisely the variations to be expected if the ocean floor constitutes an intrinsic part of the cold surface boundary layer of a mantle-wide convective circulation and if this circulation is forced to an important extent by heating from below due to the action of a substantial heat flux across the core-mantle boundary. *See* CONTINENTAL DRIFT; EARTH, HEAT FLOW IN.

Plate tectonics. From a global perspective, attempts to describe the surface manifestations of the underlying mantle convective circulation have given rise to a body of descriptive ideas which have collectively been called plate tectonics. In this kinematic picture, the surface of the Earth is seen to be

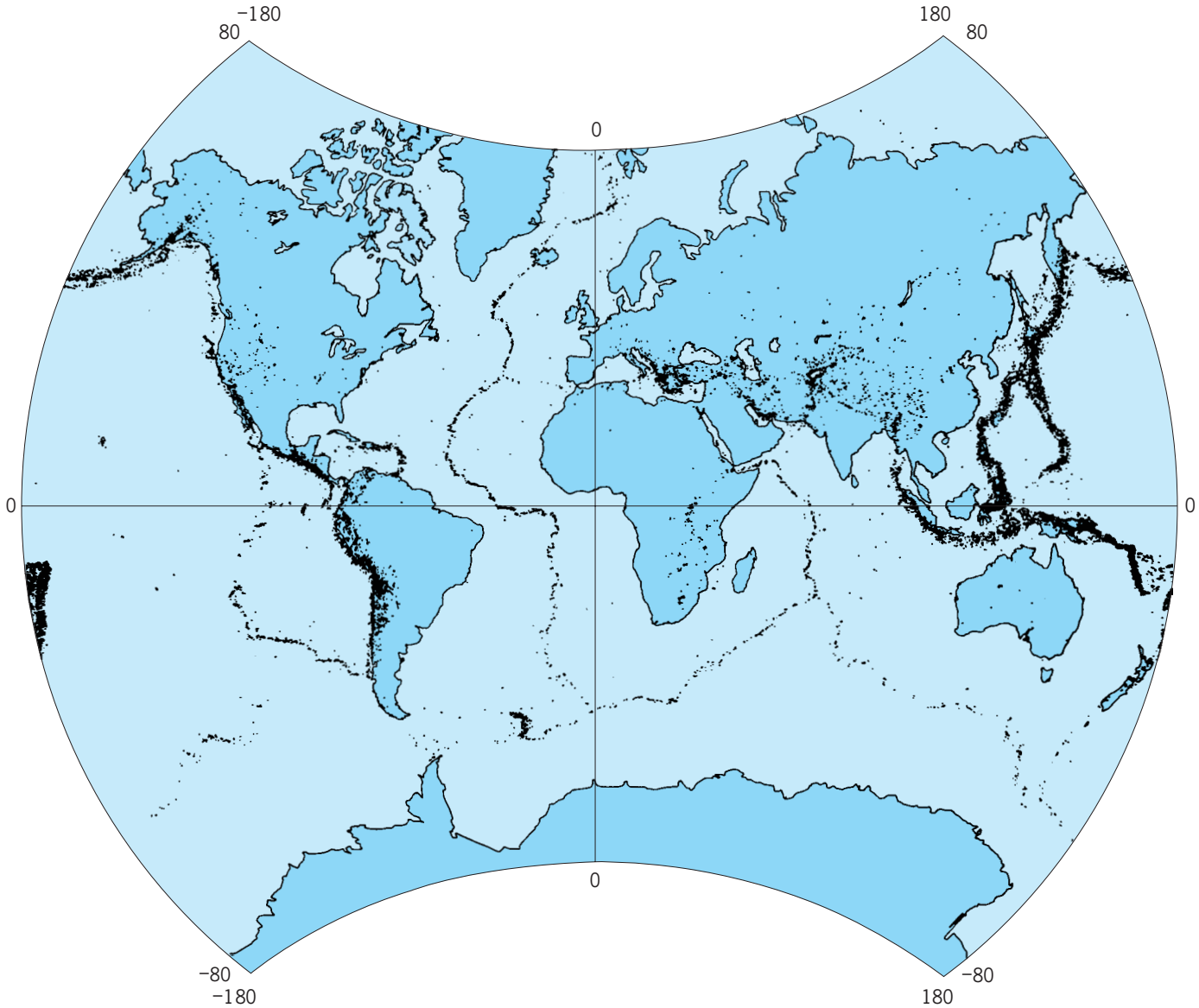

Fig. 4. Global seismicity showing surface projections of all earthquakes which occurred during 1965–1975 with magnitudes in the range 4.5–5.5. Plate boundaries are located along the seams marked by high frequency of earthquake occurrence.

divided into a relatively small number of rigid plates which interact with each other only at their boundaries. Continents are carried passively by the plates and drift with them over the Earth's surface. The boundaries separating adjacent plates are defined on the basis of the global distribution of seismicity, in terms of which it was first recognized that the system of mid-oceanic ridges, deep ocean trenches, and their important interconnections via transform faults was continuous. *See* MARINE GEOLOGY; TRANSFORM FAULT.

Volcanic activity is confined to regions of active subduction along plate margins (**Figs. 4** and **5**). Most major mountain systems are also found in the same regions (the Himalayas, Andes, Alps, and North American Rockies). The idea of plate tectonics is to describe surface geological processes such as mountain building and volcanism as consequences of the interactions among this relatively small number (12) of plates. For example, the currently growing Himalaya Mountains are a product of the collision between the Indian and Eurasian plates. Most volcanic activity is also sharply confined to plate boundaries. The molten rock involved in a volcanic eruption at a mid-ocean ridge is produced by pressure-induced melting when hot material rising in the ascending limb of a convection cell crosses the melting curve as it decompresses. In such locations the lavas are predominantly basaltic. Andesitic volcanoes, on the other hand, are found in the island arcs associated with subduction zones, and their lavas are distinctly different in composition from those found on mid-oceanic ridges. Not all volcanic activity is confined to the plate boundaries. So-called intraplate volcanism, such as that associated with the Hawaiian Islands, also exists, and the explanation of this

Fig. 5. Map of global tectonic and volcanic activity over the past million years. (*After The Physical World, National Geographic Society, Van der Grinten projection, 1975*)

phenomenon remains somewhat enigmatic. One suggestion is that such features are produced by thermal plumes of small horizontal scale, which are formed episodically in the hot lower thermal boundary layer of the circulation in consequence of the temperature dependence of mantle viscosity, and which can rise quickly and adiabatically to the base of the lithosphere, where they induce intense partial melting.

The detailed relation between the surface plates and the convective circulation in the underlying mantle of the Earth is a subject of scientific investigation about which there is no consensus. Most evidence, however, suggests that there is an approximately 1:1 relation between surface plates and deep mantle convection cells, although the three-dimensional geometry of this large-scale flow is expected to be complex. It is perhaps surprising therefore that simple physical models based upon understanding of thermal convection in liquids can provide a general outline of understanding of complex geological processes. *See* PLATE TECTONICS.

Dynamic models of mantle convection. The ability of the mantle convection hypothesis to reconcile the surface observations can be seen most clearly through an understanding of the meaning of the nondimensional parameter called the Rayleigh number, Ra, which determines the strength of the convective circulation which will be caused by a given thermal forcing across a layer of viscous fluid of depth d. This number is defined as shown in Eq. (1),

$$\mathrm{Ra} = \frac{g\alpha d^4 (dT/dz|_b + Qd)}{\kappa^2 \nu} \qquad (1)$$

where g is gravitational acceleration acting through the layer, α is the coefficient of thermal expansion, κ is the coefficient of thermal diffusivity, $dT/dz|_b$ is the

temperature gradient at the base of the layer, Q is the rate of internal heating per unit volume (due to radioactive decay processes, for example) divided by the product of density and specific heat capacity, and ν is the kinematic viscosity. Theoretical analysis shows, and laboratory observation confirms, that unless Ra exceeds a critical value Rc, which is normally of order 10^3, thermal convection is impossible.

An estimate of Ra for the whole mantle of the Earth can be made by employing the experimentally determined parameters: $g \approx 4 \times 10^2$ in. $\cdot$ s^{-2} (10^3 cm $\cdot$ s^{-2}), $\alpha \approx 1.4 \times 10^{-5}$ K^{-1}, $\kappa \approx 3.875 \times 10^{-3}$ in.$^2 \cdot$ s^{-1} (2.5×10^{-2} cm$^2 \cdot$ s^{-1}), and $d \approx 1.7 \times 10^3$ mi (2.9×10^3 km) as the mantle depth. The kinematic viscosity ν may be inferred through analysis of postglacial rebound data, which requires a value of ν between 4.65×10^{21} and about 1.55×10^{22} in.$^2 \cdot$ s^{-1} (3×10^{22} and about 10^{23} cm$^2 \cdot$ s^{-1}), with an increase from the low value to the high value across the seismic discontinuity at 402 mi (670 km) depth. Also obtained from the glacial rebound data is a measurement of the thickness $t \approx 60$ mi (100 km) of the layer near the surface, called the lithosphere, in which the effective viscosity is extremely high. The remaining parameters in Eq. (1), $dT/dz|_b$ and Q, are the least well determined, but can be estimated separately on the basis of alternate limiting assumptions. If $Q \approx 0$ in the mantle, an estimate can be made of $dT/dz|_b \approx (T_b - T_s)/d$, where T_b is the temperature at the core mantle boundary and T_s is the surface temperature. Since the outer core is mostly iron, and since it is in a liquid state, a minimum estimate of the temperature T_b may be obtained from the melting point of iron at core-mantle boundary pressures, and estimates are in the neighborhood of 7200°F (4000°C). This could be reduced substantially owing to the presence of a light alloying element, and such a component must exist in the outer core in order to explain the seismically observed mean atomic weight of the material in this region. Using $T_b \approx 3000$°C in Eq. (1) and $Q = 0$, Ra $\approx 10^7$ is obtained. This value is four orders of magnitude larger than the normal critical value of 10^3, so that the mantle is convectively unstable and thermally driven circulations must exist.

The Rayleigh number may also be estimated on the basis of the alternate assumption $dT/dz|_b = 0$, and a lower bound on Q may be deduced by assuming that the heat which leaves the Earth's surface is in equilibrium with the rate at which heat is generated internally by radioactive decay processes. This gives $\rho_m c_p Q = 3q_s/d$, where q_s is the mean surface heat flow of about 80 m W $\cdot$ m^{-2}, and ρ_m and c_p are the mean mantle density and specific heat capacity, respectively. Substituting this value of Q into Eq. (1) yields a value of Ra which is also enormously in excess of the critical value. No matter how the Rayleigh number is estimated, the conclusion that thermal convection currents are expected in the Earth's mantle is inescapable.

The question as to how these convection currents are driven, whether predominantly by heating from below through $dT/dz|_b$ or by internal heating through Q, is not well understood. Also, no consensus exists concerning the depth to which the circulation extends. The heated-from-below model is preferred from several points of view, however, and it will be employed here to demonstrate the ability of the convection hypothesis to reconcile the surface observations described previously. By using boundary layer theory for thermal convection at high Rayleigh number, simple expressions are obtained for the characteristic horizontal velocity of the circulation $\overline{u}$, the thermal boundary layer thickness δ, and the surface heat flux q_s, as shown in Eqs. (2), where the $a_i(\Delta)$ are functions of the aspect ratio of convection Δ, which is the ratio of the horizontal scale w of a convection cell to its vertical scale d. In laboratory convection, the realized aspect ratio is $w/d \approx 1$, and if this is assumed for the mantle circulation with d equal to the mantle thickness, horizontal scales of several thousand kilometers are implied, and in fact, such is the value of the mean horizontal scale of the observed surface plates. If the value Ra $\approx 10^7$ is used in Eqs. (2) and $\Delta = 1$, then $\overline{u} \approx 2.8$ in. (7 cm) per year, $\delta \approx 60$ mi (100 km), and $q_s \approx 100$ mW $\cdot$ m^{-2}. These are very close indeed to the observed surface plate speed, the observed thickness of the lithosphere, and the observed mean surface heat flow. The most important assumption in this simple picture of the convection circulation in the mantle is that the oceanic lithosphere is an intrinsic part of the cold thermal boundary layer of the flow responsible for plate creation and destruction. When this assumption is coupled with the notion that convection fills the entire mantle, the observed mean properties of the circulation are related as they would be in a laboratory convection apparatus. *See* CONVECTION (HEAT).

$$\bar{u} = a_1(\Delta)\mathrm{Ra}^{2/3}\kappa/d \qquad (2a)$$

$$\delta = a_2(\Delta)d(Rc/\mathrm{Ra})^{1/3} \qquad (2b)$$

$$q_s = a_3(\Delta)(\mathrm{Ra}/Rc)^{1/3}\rho c_p \kappa \Delta T/d \qquad (2c)$$

Mantle convection with phase transitions. In order to develop a more detailed understanding of the mantle convention process, beyond that embodied in the simple scaling relations discussed above, it is necessary to construct explicit hydrodynamic models. Most such models are based upon the conventional conservation laws for mass, momentum, and internal energy and the equation of state for a viscous fluid. The equation of state considers the incremental variations of density that exist in the mantle in consequence of any transitions of phase that may be induced by the hydrostatic compression of its constituent material. In fact, the existence of such phase transformations is well established. The most important of these from the perspective of the dynamics of the mantle convection process are the exothermic transition from olivine to the α-spinel structure that occurs near a depth of 240 mi (400 km) and the endothermic transition from β-spinel to a mixture of perovskite and magnesiowüstite that occurs near 400 mi (660 km) depth.

A series of contour maps showing the temperature field in the mantle can be developed from a modern hydrodynamic model of the mantle convection process. The individual images show the mantle in half-section (spanning a time period of about 600 million years) because the numerical model is based upon the assumption that the flow, although occurring in a three-dimensional spherical shell, may be treated as though it were independent of the longitudinal coordinate. The sequence of images demonstrates that the influence of the phase transitions, especially the endothermic transition at 400 mi (660 km) depth, on the convection process is profound. Expressed in the simplest way, the impact of this transition is such as to quite strongly inhibit the flow of mass across this horizon. The degree of inhibition is in fact so strong that the flow spends extensive periods of time in a state in which the circulation is almost perfectly layered. This state is disrupted by episodic instabilities of the thermal boundary layer that develops across the 660-km endothermic horizon when the flow is in the layered state. These disruptions take the form of intense "avalanches'' of cold material from the upper mantle into the lower mantle, during which the vigor of the convective circulation increases significantly. In the axisymmetric model, these episodes of "whole mantle'' convection appear to last for times of order 100–200 million years, while the intervening periods of layered convection last two to three times longer.

If such processes occur in the real Earth, they could be involved in a variety of different and heretofore unexplained phenomena. Principal among these is the so-called supercontinent cycle of ocean basin opening and closing. This cycle and the cycle of orogeny (mountain building) that accompanies it is well expressed in the histogram of occurrence of surface continental rocks as a function of their age. When an avalanche occurs across the 660-km transition, the intense downwelling that occurs at this location causes even extremely distant material on the Earth's surface to collect in the basin above it. This is a plausible mechanism whereby continental fragments may be collected together to form a supercontinent.

Seismic-tomography-based models. In attempting to verify the predictions of the a priori models of mantle convection, discussed above, images of the three-dimensional internal structure of the mantle density field derived on the basis of the application of seismic tomography are important. These methods, which are very similar to the CAT scan techniques employed in medicine, are able to deliver models of the internal variation of a seismic body wave speed, typically that for shear waves. On the basis of the assumption that variations in shear-wave speed may be attributed to variations of temperature alone, the three-dimensional images of shear-wave speed may be mapped directly into variations of temperature. *See* COMPUTERIZED TOMOGRAPHY.

Mathematical expressions developed from a model based on seismic tomographic imaging may be solved to predict a variety of observable characteristics of the mantle convective circulation. The most important characteristics undoubtedly include the planet's gravitational field, the velocities of the surface plates, the topography on the core-mantle boundary, the dynamic component of the topography on the Earth's surface, and the advected component of the radial heat flow. Two such comparisons can be illustrated—the first for the nonhydrostatic geoid and the second for the dynamic topography of the Earth's outer solid surface. Since the existing models of seismic tomography provide information only on the longest horizontal wavelength constituents in the internal density field, and are similarly constrained in their ability to resolve radial structure, the theory can be tested only on these largest scales. W. R. Peltier

Bibliography. G. F. Davies, *Dynamic Earth: Plates, Plumes and Mantle Convection*, 1999; J.-P. Poirier, *Introduction to the Physics of the Earth's Interior*, 2d ed., 2000; F. Press and R. Siever, *Understanding the Earth*, 3d ed., 2000.

Earth, gravity field of

The field of gravitational attraction of the Earth. Since, at the Earth's surface, the small centrifugal force due to the Earth's rotation is inseparably superimposed on the attraction, the gravity field is usually understood to include also the effect of the centrifugal force.

Concepts

The gravitational potential V at a point P is the potential energy, due to the Earth's gravitational attraction, of a unit mass situated at P. In other words, V is equal to the work done if a unit mass is brought from infinity to the point P under the influence of the Earth's gravitational field. A mathematical expression is obtained by integrating over the Earth the formula for the potential of a point mass, but this expression is almost useless for practical application. For the potential outside the Earth, an expansion into an infinite series (of spherical harmonics, discussed later in this article) is useful; the principal term of this series is GM/r. This term, in which G is the newtonian gravitational constant (6.67×10^{-8} m^3 kg^{-1} s^{-2}), M is the total mass of the Earth, and r is the distance of P from the Earth's center of mass, represents formally the attraction of a spherically symmetric Earth; the consideration of this term only is not, in general, sufficient. The gravitational force $\mathbf{F}$ on a unit mass is the gradient vector of V, $\mathbf{F} = \text{grad } V$; that is, the components of $\mathbf{F}$ are $(\partial V/\partial x, \partial V/\partial y, \partial V/\partial z)$.

Gravitation and gravity. The resultant of gravitation (pure attraction) and centrifugal force is called gravity. Gravity is the force that acts on the body at rest with respect to the Earth since the effects of attraction and of centrifugal force cannot be separated because of the equivalence of gravitational and inertial mass; thus gravity determines the weight of a body.

The gravity potential W is the sum of the gravitational potential V and the potential of centrifugal force, which is given by a simple analytical expression and may be considered as known.

A body moving with respect to the Earth is also affected by the Coriolis force. Like centrifugal force, the Coriolis force is an inertial force due to the Earth's rotation, but unlike centrifugal force, it does not possess a potential and hence cannot be easily incorporated into the gravity field. Therefore Coriolis force is not considered in the context of terrestrial gravitation. This is perfectly adequate since this force is zero for bodies at rest with respect to the Earth, and almost all measuring systems are at rest. *See* CORIOLIS ACCELERATION; GRAVITATION.

Gravity vector. The gravity vector **g** represents the force of gravity on a unit mass. It is the gradient vector of the gravity potential, $\mathbf{g} = \text{grad } W$. The magnitude of the gravity vector is the intensity of gravity, or briefly, gravity g. The dimension of g is force per unit mass, or acceleration. The SI unit is $\text{m} \cdot \text{s}^{-2}$. The cgs unit gal ($1 \text{ gal} = 1 \text{ cm} \cdot \text{s}^{-2}$), named after Galileo, is still frequently used, especially the milligal ($1 \text{ mgal} = 10^{-3} \text{ gal} = 10^{-5} \text{ m} \cdot \text{s}^{-2}$). Gravity g on the Earth's surface varies from about 978 gals at the Equator to about 983 gals at the poles. The direction of the gravity vector defines the vertical, or plumb line.

Geoid. The surfaces of constant gravity potential, $W = \text{const}$, are called equipotential surfaces or level surfaces. The surface of a quiet lake is part of a level surface. So is the surface of the oceans, after some obvious idealization; the whole level surface so defined is called the geoid. After C. F. Gauss, the geoid is considered as the mathematical surface of the Earth, as opposed to the visible topographical surface. The plumb lines intersect the level surfaces orthogonally (**Fig. 1**); they are not quite straight but very slightly curved.

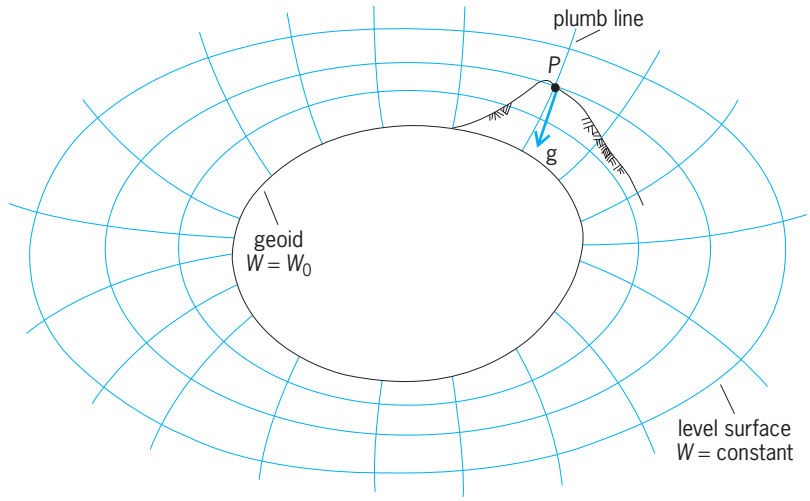

Fig. 1. Level surfaces and plumb lines. The arrow gives the direction of the gravity vector *g* at point *P*. *W* is the gravity potential.

Measurement

The quantity that is measured most commonly is the gravity g. The determination of g as such is called an absolute gravity measurement. Usually only relative gravity measurements are performed, determining the difference between, or the quotient of, the gravity values at two different points. The direction of the gravity vector, which gives the plumb line in space, is measured by astronomical methods. Differences in the gravity potential W are obtained by geodetic leveling. Finally, certain derivatives of g and similar quantities are measured by instruments such as the torsion balance.

Satellite methods have enormously improved knowledge of the gravity field. In the future, new satellite technologies will advance gravity-related measurements, to be studied along with classical theoretical measurements, which will fully retain their importance. *See* GRAVITY.

Classical gravity measurements. Absolute gravity measurements are now exclusively performed by falling-body instruments, which use the fact that the free fall of a test object is proportional to gravity g: if a body, originally at rest, is dropped in vacuum, the vertical distance s covered during time t is represented by Eq. (1), which gives Eq. (2). Since s and t can be

$$s = \frac{1}{2} g t^2 \tag{1}$$

$$g = \frac{2s}{t^2} \tag{2}$$

very accurately measured, gravity g is obtained with a relative accuracy of 10^{-8} to 10^{-9}. Both stationary and movable instruments are used routinely.

The former use of pendulum apparatuses for measuring absolute g gave only a relative accuracy of 10^{-6} and is now completely obsolete.

Relative gravity measurements are performed by spring gravimeters, which are based on Hooke's law: the extension s of the elastic spring is proportional to the force of gravity, so that the difference of gravity g between two points P_1 and P_2 is given by Eq. (3), where k is a factor of proportionality. The

$$g_2 - g_1 = k(s_2 - s_1) \tag{3}$$

same principle is used in accelerometers. The accuracy is consistent with that of modern absolute gravity measurements. *See* ACCELEROMETER; GRAVITY METER.

Since relative gravity measurements are much simpler, absolute determinations of gravity are performed only at relatively few points on the Earth's surface. At other points, gravity is determined relative to these absolute stations. Relative measurements of gravity are made by pendulums or gravimeters.

Basically, an instrument that measures gravity (a gravity sensor) is nothing but a special form of accelerometer. The principle of the spring balance is also utilized for accelerometers, and inversely vibrating-string accelerometers have been proposed as gravity sensors since the frequency of a loaded vibrating string is a function of g. Accelerometers have

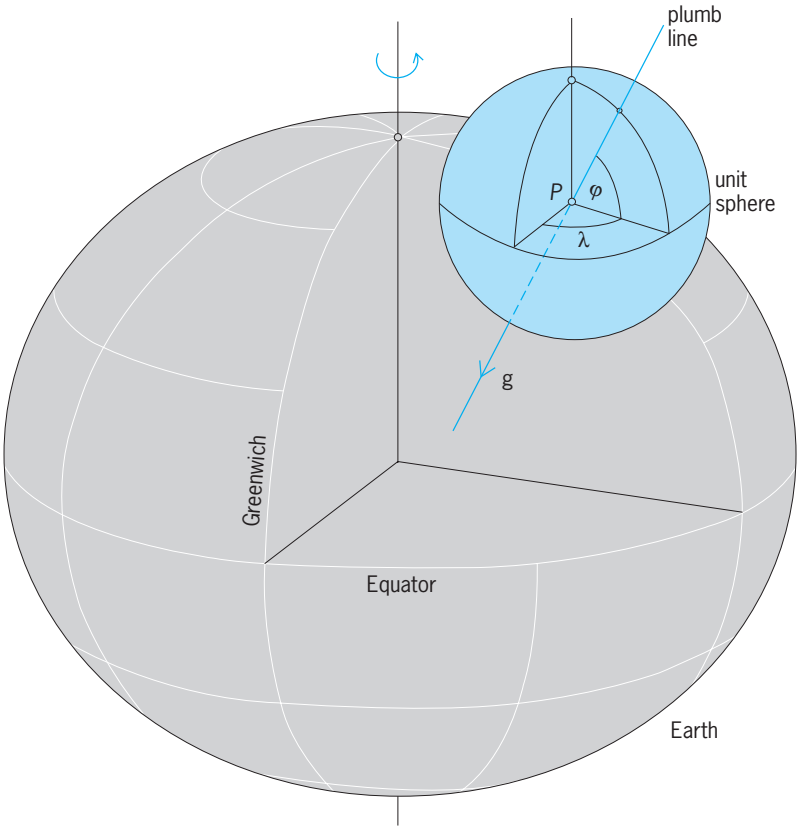

Fig. 2. Direction of plumb line is defined by the geographical coordinates φ and λ. The unit sphere is frequently used for representing spatial directions.

been especially proposed for use in gradiometers, as discussed in the section below on gravity gradients.

Direction of gravity vector. This direction, the plumb line, is defined by two angles; φ, geographical latitude, and λ, geographical longitude (**Fig. 2**). These angles are determined by astronomical methods to an accuracy of 0.1 second of arc or better. If g, φ, and λ are known, then the gravity vector $\mathbf{g}$ is completely determined.

Potential differences. Differences in the potential W can also be measured. If dn is the distance between two neighboring level surfaces of potential difference dW, then $dW = -g\ dn$, so that Eq. (4)

$$W_B - W_A = \int_A^B g\ dn \tag{4}$$

holds. The geodetic operation of leveling (spirit leveling or differential leveling) determines dn. The leveling increments dn are multipled by gravity g and summed along the leveling line that connects points A and B. In this way the difference of potential W between A and B is determined.

Gravity gradients. The gradient of gravity along a certain direction s, $\partial g/\partial s$, is the component of the vector grad g along this direction. Instruments measuring gravity gradients are called gradiometers.

Horizontal gradients (gradients along any horizontal direction) are obtained by means of the torsion balance invented by R. Eötvös about 1900. This instrument consists essentially of a vertically suspended torsion wire carrying a light horizontal beam at which two equal masses are suspended at different heights. The torsion balance gives horizontal gradients of g and, in addition, certain characteristics of the curvature of equipotential surfaces.

The practical application of the torsion balance has suffered from the competition of modern gravimeters, but the interest in gradiometers is being revived because they can be used for gravitational measurements aboard an artificial satellite (a gravimeter would in this case constantly indicate $g = 0$!).

A general gradiometer may be viewed as a combination of two gravity sensors separated by a small distance; the difference in their readings, divided by the distance, gives the gravity gradient. The use of vibrating-string sensors, mentioned above, has been proposed for this purpose. The full power of this technique will be achieved by satellite gradiometry.

Gravitational Field

The gravitational field of a reference ellipsoid of revolution which closely approximates the real Earth is called the normal field. The difference between the real field and the normal field is called the anomalous field.

Normal field. Since the Earth is very nearly an ellipsoid, the gravity field of a suitable ellipsoid of revolution is a good approximation to the gravity field of the Earth and can thus be considered as a normal terrestrial gravity field. The normal gravity potential is denoted by U. An ellipsoid of revolution—the reference ellipsoid—that is close to the geoid is defined to be an equipotential surface of this normal field, where $U = U_0 =$ constant.

Normal gravity. Gravity on such a "level ellipsoid" is called normal gravity γ; it is given by the gravity formula, Eq. (5), where γ_e is normal gravity at the equator, φ is geographical latitude, and ϵ is a very small constant (for numerical values see below).

$$\gamma = \gamma_e(1 + \beta \sin^2 \varphi - \epsilon \sin^2 2\varphi) \tag{5}$$

The other constant, β, is related to the geometric flattening of the terrestrial ellipsoid, $f = (a - b)/a$ (where $a =$ semimajor axis and $b =$ semiminor axis) by A. C. Clairaut's formula, Eq. (6), where $m =$ centrifugal force at Equator/gravity at Equator, in which terms of the second order in f are neglected. This remarkable relation permits the determination of the Earth's flattening f from gravity measurements.

$$f + \beta = \frac{5}{2}m \tag{6}$$

The level ellipsoid and its gravity field are defined by four constants, for example, the set (a, GM, J_2, ω). Here a and GM have been introduced before; J_2 is the dynamical form factor or dynamical ellipticity of the Earth, defined as $J_2 = (C - A)/Ma^2$, where C and A are the Earth's principal moments of inertia; and ω is the Earth's angular spin velocity.

The constants f, β, and γ_e are related to these constants, to the first order in f, by the formulas (7), (8), and (9).

$$f = \frac{3}{2} J_2 + \frac{1}{2} m \tag{7}$$

$$\beta = 2m - \frac{3}{2} J_2 \tag{8}$$

$$\gamma_e = \frac{GM}{a^2}\left(1 + \frac{3}{2} J_2 - m\right) \tag{9}$$

The set of values given below defines the Geodetic Reference System 1980 adopted by the International Union of Geodesy and Geophysics.

$$\begin{aligned} a &= 6,378,137 \text{ m} \\ GM &= 3,986,005 \times 10^8 \text{ m}^3 \cdot \text{s}^{-2} \\ j_2 &= 0.00108263 \\ \omega &= 0.00007292115 \text{ s}^{-1} \end{aligned}$$

The corresponding flattening is $f = 1/298.257$, and the gravity formula becomes Eq. (10).

$$\gamma = 978.0327(1 + 0.0053024 \sin^2 \varphi - 0.0000058 \sin^2 2\varphi) \text{ gal} \tag{10}$$

External and internal field. Gravity above the ellipsoid decreases according to formula (11), where γ_b is

$$\gamma_b = \gamma - (0.30877 - 0.00044 \sin^2 \varphi)b + 0.000072b^2 \tag{11}$$

gravity at elevation b and latitude φ, and γ is the gravity at the same latitude on the surface of the ellipsoid; γ_b and γ are measured in gals, and b is measured in kilometers.

As for the internal field, the internal level surfaces are approximately ellipsoids that become more and more spherical with increasing depth. The theory of the internal field is governed by a differential equation published by Clairaut in 1743. This equation relates the flattening of the level surfaces to the density under the assumption of hydrostatic equilibrium. The solution of this equation offers a possibility of computing the flattening of the Earth from the constant of astronomical precession, which is quite accurately known. Values for f obtained in this way and values obtained more directly from the analysis of satellite orbits, from J_2 [see Eq. (10)], show discrepancies which have yet to be completely explained.

Anomalies of the gravitational field. The normal ellipsoidal field incorporates the main part of the Earth's gravitational field. The difference between the real field and the normal field is called the anomalous field. The anomalous potential T is the difference between the actual gravity potential W and the normal gravity potential U. (The circumstance that the symbol T has been previously used to denote the period of oscillation of a pendulum will cause no confusion.)

Other quantities of the anomalous field are the gravity anomaly and the deflection of the plumb line. Both are more readily accessible to observation than the anomalous potential T, and they may be used to determine T.

Gravity anomalies. The gravity anomaly Δg is defined as the difference between gravity at sea level, g, and normal gravity at the reference ellipsoid, γ: $\Delta g = g - \gamma$. Since measurements are not generally made at sea level, the measured gravity value must be reduced from the Earth's surface to sea level, that is, to the geoid. Depending on the way in which this is achieved, different types of gravity anomalies are obtained.

This reduction may be done without considering the masses below the observation station; that is, g is reduced as if these masses did not exist and the station were "in free air." This is the free-air reduction. The amount of this reduction, to be added to the measured gravity, is by formula (14) approximately $0.309b$ mgals, where b is the elevation in meters.

In the process of gravity reduction, the masses above the geoid may also be removed computationally. This is the Bouguer reduction. The amount to be added to measured gravity is now $0.197b$, the units being the same as before; this figure is based on a rock density of 2.67 g/cm^3.

Instead of completely removing the masses above the geoid, they may be shifted computationally into its interior in such a way as to get a level homogeneous crust according to some theory of isostasy (see later section of this article). This is an isostatic reduction. There are also a number of other gravity reductions.

Deflections of plumb line. The deflection of the plumb line, or deflection of the vertical, is the deviation of the actual plumb line from the normal to the reference ellipsoid. It is characterized by two components, the component ξ in a northern direction and the component η in an eastern direction. If φ_n and λ_n are the normal (ellipsoidal) coordinates, as opposed to their actual counterparts φ and λ, then ξ and η are given by Eqs. (12) and (13). Astronomical observations give φ and λ, whereas φ_n and λ_n may be obtained from geodetic work such as triangulation. The ξ and η so determined are called astrogeodetic deflections of the vertical.

$$\xi = \varphi - \varphi_n \tag{12}$$

$$\eta = (\lambda - \lambda_n)\cos\varphi \tag{13}$$

Anomalous potential and the geoid. The height of the geoid above the reference ellipsoid, the geoidal height N, is intimately related to the anomalous potential T by Bruns's formula, $T = \gamma N$, where γ is normal gravity. Therefore, the determination of T is equivalent to the determination of the geoid.

A relative determination can be made by the astrogeodetic method. Differences $T_B - T_A$ or $N_B - N_A$ are obtained by integrating astrogeodetic deflections of the vertical along a line connecting the points A and B. By this method it is possible to determine the geoid in a limited region from data (plumb-line deflections) in this region only. The astrogeodetic method

furnishes very detailed and accurate geoidal maps (or maps of the potential T) in areas where good astrogeodetic data are available. However, since this method is only relative, the position of the geoid so obtained is determined only apart from an unknown shift with respect to the Earth's center of mass. The use of the astrogeodetic method is restricted to land areas because the necessary data cannot yet be obtained at sea.

An absolute determination of the anomalous potential or the geoidal height can be achieved from gravity anomalies Δg. This is done by an integral formula due to Stokes (1849). Stokes' formula determines the geoid absolutely; that is, the geoidal heights so obtained refer to an ellipsoid whose center coincides with the Earth's center of mass. It requires the gravity anomalies to be given all over the Earth; this is practicable because gravity can be measured at sea as well as on land areas. Large regions, especially the oceans, have not been satisfactorily surveyed gravimetrically. This brings about certain restrictions on the practical applicability of the gravimetric method; nevertheless it has given important results. **Figure 3** shows part of the "Columbus Geoid," a historical landmark computed at Ohio State University.

Harmonic analysis. A periodic function can be expanded into a Fourier series; this is called harmonic analysis. Harmonic analysis may be generalized to functions of more than one variable, such as T or Δg. Instead of Fourier series there is, in this particular case, series of so-called spherical harmonics; such series are applied also in other geosciences, such as geomagnetism.

A series of spherical harmonics is infinite. In practice, there is usually a restriction to series truncated after terms of lower degree. Higher degree terms cannot be determined properly and must be neglected. In this way the general broad features of the expanded function are retained, although fine details are lost.

The spherical-harmonic expansion of T is readily obtained if such an expansion of Δg is available. The practical use of this method is again impaired by the lack of gravity data on parts of the Earth's surface, but this drawback can be largely compensated by a combination of terrestrial and satellite data.

The motion of an artificial satellite in the Earth's gravitational field is most strongly affected by the harmonics of lower degree. Therefore satellite observation gives values of the lower degree spherical harmonics of the potential, which provide a good qualitative picture of the general behavior of the terrestrial gravity field and, in particular, of the large features of the geoid (**Fig. 4**). *See* SPHERICAL HARMONICS.

The practical use of spherical-harmonic expansions for representing the Earth's gravity field is usually expressed in terms of the geoid, and is of basic importance. The expansions are usually obtained by a combination of satellite and gravity data. Satellite orbital data give realistic expansions up to degree of about $N = 30$, with a geoidal accuracy of about ± 5 m (16 ft) and a resolution which can be estimated as follows. One degree of arc corresponds to roughly 100 km (61 mi) of distance of the Earth's surface. A spherical-harmonic expansion to degree of $N = 30$, Eq. (14) gives about $180^\circ/30 = 6^\circ = 600$ km (37 mi)

$$\text{Resolution} \equiv \text{half wavelength} \mathrel{\hat{=}} \frac{180^\circ}{N} \mathrel{\hat{=}} \frac{20{,}000\ \text{km}}{N} \text{ or } \frac{12{,}200\ \text{mi}}{N} \qquad (14)$$

which, together with a rms error of about $\delta N = \pm 5$ m (16 ft) is not very satisfactory. This was the state of the art at the beginning of the satellite era, around 1965.

Nowadays, earth gravity models (EGM) combine satellite data with terrestrial gravity and are computed to a spherical-harmonic degree of $N = 360$. This improves the resolution by Eq. (14) to $0.5^\circ = 50$ km (30 mi) [theoretically], the accuracy being about 1 m (3.3 ft) or better. Current models even use $N = 720$ with about the same real accuracy since the same data are used.

Even $N = 1000$ is being envisaged, but without significantly new kinds of measurements it is difficult to really improve the situation. Such new techniques are satellite-to-satellite tracking (SST) and satellite gravity gradiometry (SGG), combined with the Global Positioning System (GPS). These techniques will be outlined below (satellite altimetry has

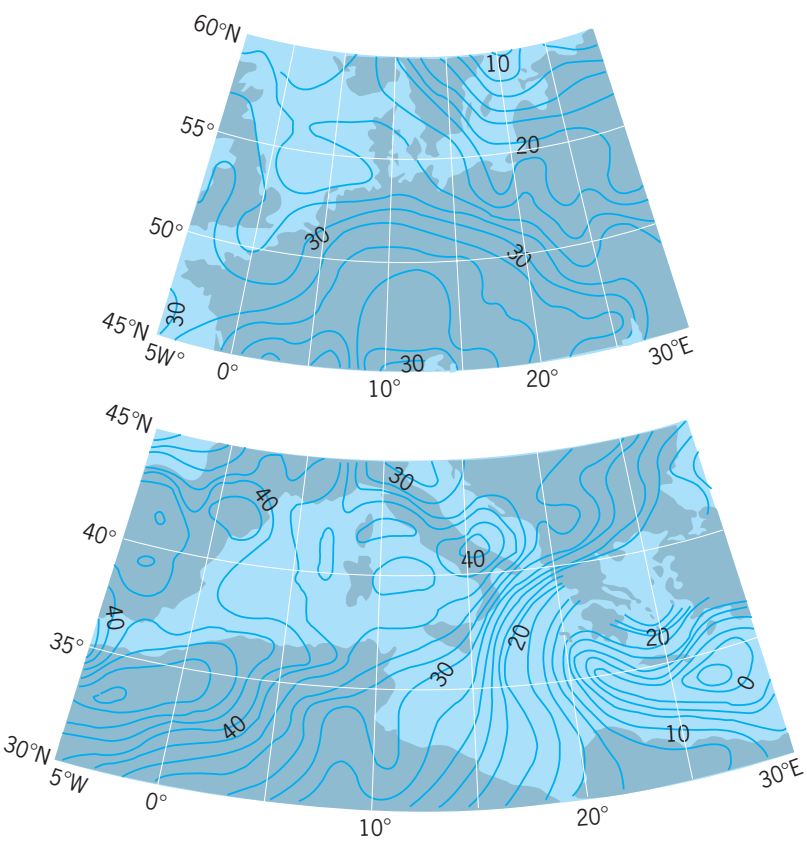

Fig. 3. Gravimetric geoid of Europe, 1957. Geoidal heights in meters (1 m = 3.3 ft), referred to an ellipsoid of flattening 1/297.

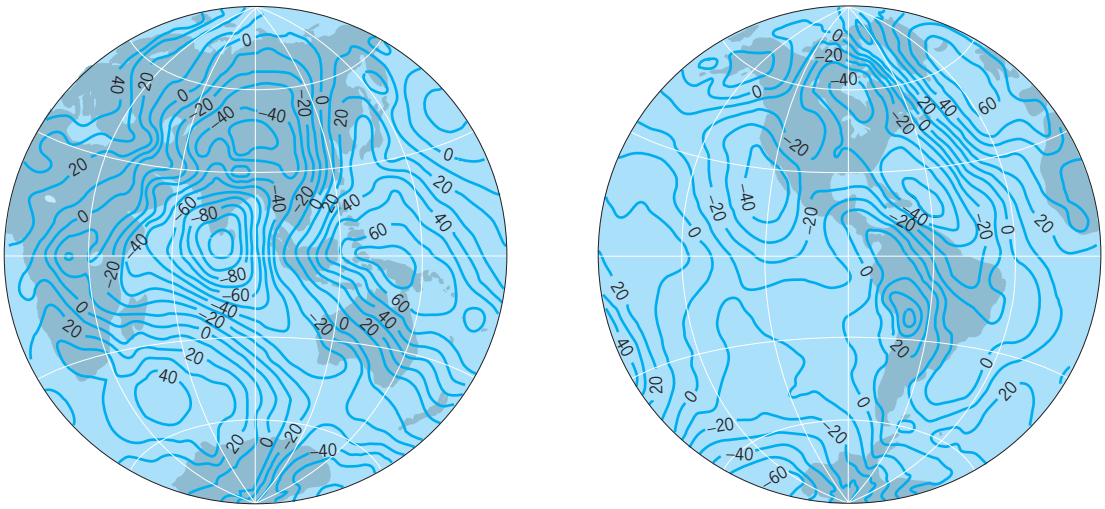

Fig. 4. Geoid from satellite observations combined with terrestrial gravity measurements (EGM96–360). Geoidal heights are in meters (1 m = 3.3 ft), referred to the Geodetic Reference System 1980.

already been taken into account in the $N = 360$ and 720 models).

Satellite altimetry. A direct determination of the oceanic geoid has become possible by measuring the height of a satellite above sea level. This is done by observing the travel time of a radar or laser signal emitted by a satellite and reflected by the ocean surface. If the satellite orbit is known with respect to the reference ellipsoid, the geoidal height follows. A breakthrough in understanding the ocean geoid was achieved by the altimetry satellites *GEOS 3*, 1975–1977, and *Seasat*, 1978.

Satellite-to-satellite tracking. The distance between two satellites changes with location in space due to the irregularities of the gravity field. If continuously and accurately monitored, information of the irregular field may be obtained. Roughly, $N = 100$ could be expected. Implementation will come soon after the year 2000.

Gravity gradiometry. This is an application of the principle of terrestrial gradiometry mentioned above to a gradiometer built into a satellite. Achievable resolutions are higher than with satellite-to-satellite tracking: $N = 250$ is being considered feasible. In combination with other satellite and terrestrial techniques, realistic $N = 720$–1000 models with accuracies on the order of a few centimeters appear possible. Practical application is to come slightly after satellite-to-satellite tracking.

Global Positioning System. GPS and its Russian counterpart, GLONASS, are purely geometrical and therefore (almost) independent of the gravity field. By furnishing precise positions, they make possible a separation of gravitational and geometric effects. In particular, GPS gives inertial accelerations (second derivatives of position) and thus separates them from the truly gravitational forces. This permits putting gravimeters and gradiometers into an airplane whose position is monitored by GPS. At least, airborne gravimetry is already a practically applied reality.

Each of these methods—astrogeodetic, GPS, gravimetric, and the different satellite techniques—has its specific merits and shortcomings. In many respects, they complement each other. By suitable combinations of these data, a very precise gravity field, formulated in terms of a "centimeter-geoid," may become a reality in the near future.

Temporal variations. For most purposes, the Earth's gravitational field may be considered invariable in time. However, it is subject to extremely small periodic variations due to tidal effects. These are caused by the attraction of the Sun and Moon. The attraction acts directly by superimposing itself onto the Earth's gravitational attraction; and it acts also indirectly by slightly deforming the Earth and shifting the waters of the oceans, so that the attracting terrestrial masses themselves are modified.

The lunar effect on gravity attains a maximum of 0.20 mgal, and the solar effect, a maximum of 0.09 mgal; both are well within the measuring accuracy of modern gravimeters. The results of stationary gravimeters recording variations of gravity may be used to draw conclusions as to the elastic behavior of the Earth under the influence of tidal stresses. *See* EARTH TIDES.

Relations to Other Research

The determination of Earth's gravitational field depends on other studies, such as the astronomical determination of the plumb line, and has application in the search for mineral deposits and so on.

Geophysics and geology. The anomalies of the terrestrial gravitational field are caused by mass irregularities. These may be the visible irregularities of topography such as mountains; or they may be invisible subsurface density anomalies. This is the reason why it is possible to use gravity measurements for investigating the underground structure of the Earth's crust. Thus analysis of gravity is applied by geophysicists and geologists for studying general features of the crust, and by exploration geophysicists for searching for shallow density irregularities that might indicate the presence of mineral deposits.

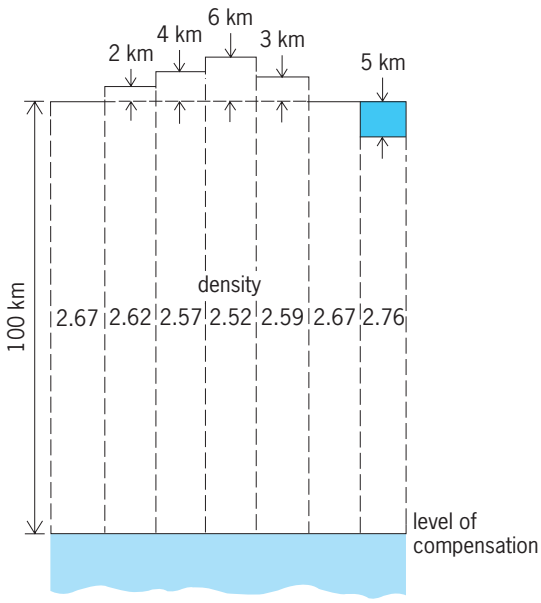

Fig. 5. Pratt isostasy model. 1 km = 0.6 mi.

Isostasy. If the mountains were simply superposed on an essentially homogeneous crust and if the ocean depressions were simply hollowed out from such a crust, the irregularities of the gravity field would be almost 10 times as large as they actually are. This indicates that the visible mass anomalies, such as mountains or ocean depressions, are to a large extent compensated: The crustal density underneath the mountains is smaller than normal, and underneath the oceans, greater than normal.

Theories of isostasy were developed by J. H. Pratt and by G. B. Airy, both about 1850. According to Pratt, the mountains have risen from the underground somewhat like a fermenting dough (**Fig. 5**). According to Airy, the mountains are floating on a fluid lava of higher density, somewhat like an iceberg on water, so that the higher the mountain the deeper it sinks (**Fig. 6**).

Pratt became aware of isostatic compensation through investigating astrogeodetic deflections of the vertical in the Himalayas. At one station he computed from the visible masses a deflection of 28 seconds of arc, whereas the observed value was only 5 seconds of arc.

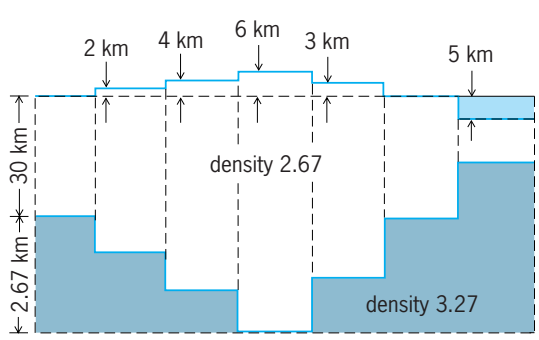

Fig. 6. Airy isostasy model. 1 km = 0.6 mi.

Isostasy is also the reason for a conspicuous behavior of the Bouguer anomalies in mountain areas. Here they are systematically negative and increase in magnitude on the average about 100 mgals/1000 m of mean elevation.

If isostatic compensation were complete, then the isostatic gravity anomalies as defined above would be zero. In general, they are small and fluctuate smoothly around zero; this property makes them particularly well suited for interpolation and extrapolation to poorly observed areas in geodetic applications. Systematically nonzero isostatic anomalies indicate deviation from isostatic equilibrium. Therefore isostatic gravity anomalies, and also Bouguer anomalies, are used to investigate the degree and the mode of isostatic compensation. It appears that the Earth's crust is compensated isostatically to about 90percnt;; the compensation seems in general to be of Airy type rather than Pratt type. *See* EARTH.

Exploration geophysics. Local mass anomalies at shallow depth, which might be related to mineral deposits, are discovered by a local very dense and accurate gravity survey by means of gravimeters. Bouguer anomalies are computed to eliminate as far as possible the effect of visible masses, and various filtering techniques are employed to isolate and localize disturbances. *See* GEOPHYSICAL EXPLORATION.

Geodesy. Geodetic instruments employ spirit levels and other devices to orient them with respect to the horizontal or, what amounts to the same thing, to the plumb line. Since the plumb line is defined by the gravitational field, it can be understood why this field enters essentially into almost all geodetic measurements, even into apparently purely geometric ones. In return, geodetic techniques are among the most efficient means for determining the gravitational field. The "mathematical figure of the Earth" for the purpose of geodesists, the geoid, is defined as a surface of constant potential W. "Heights above the sea level" are heights above the geoid; their determination is therefore a physical as well as a geometric problem. (Geodetic theories have been developed which employ only quantities referred to the Earth's topographical surface; but here the gravitational field enters in an even more complicated way.) Thus geodesy is essentially concerned with the Earth's gravitational field and its determination; the theory of the figure of the Earth is to a large extent equivalent to the theory of terrestrial gravitation. *See* GEODESY.

Astronomy and satellite dynamics. Astronomical observations are indispensable for determining the plumb line, which defines the direction of the gravity vector. Parameters of the Earth's gravitational field such as equatorial gravity γ_e are used to express in kilometers the fundamental astronomical unit, the mean distance of Sun and Earth. The Earth's equatorial radius a, which gives the scale of the geoid, is also used as a basic astronomical constant. *See* ASTRONOMICAL UNIT; EARTH ROTATION AND ORBITAL MOTION.

Terrestrial gravitation affects the motion of the Moon and of the artificial satellites. By observing

these satellites, it is therefore possible to determine essential parameters of the gravitational field, such as the geocentric gravitational constant GM, which is the product of the Newtonian gravitational constant and the Earth's mass. The dynamical ellipticity J_2 defined above, and other lower degree coefficients of the spherical-harmonic expansion of the gravitational potential, may also be determined in this way. GM and J_2 are the coefficients of degrees zero and two of this expansion. *See* SATELLITE (SPACECRAFT).

Helmut Moritz

Bibliography. B. Battrick (ed.), *European Space Agency (ESA), Gravity Field and Steady-State Ocean Circulation Mission*, Doc. SP-1233(1), Noordwijk, The Netherlands, ESTEC, 1999; A. H. Cook, *Gravity and the Earth*, 1969; G. D. Garland, *The Earth's Shape and Gravity*, 1977; W. A. Heiskanen and H. Moritz, *Physical Geodesy*, 1967; H. Jeffreys, *The Earth*, 6th ed., 1976; F. G. Lemoine et al., *The Development of the Joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96*, Doc. NASA/TP-1998-206861, Greenbelt, MD, Goddard Space Flight Center, 1998; H. Moritz, *Advanced Physical Geodesy*, 1980; H. Moritz, *The Figure of the Earth*, Wichmann, Karlsruhe, 1999; F. Press and R. Siever, *The Earth*, 4th ed., 1986.

Earth, heat flow in

The Earth's heat flow is defined as the amount of heat escaping per unit time from the interior across each unit area of the Earth's solid surface. The movement of heat within the Earth and its eventual loss through the surface are central elements in the modern theory of plate tectonics. The plate movements over the surface of the Earth are seen as one manifestation of a heat engine at work in the interior, and the heat flow at the surface as the exhaust from the engine. *See* EARTH INTERIOR; PLATE TECTONICS.

That the interior of the Earth is a great reservoir of heat is a fact known from antiquity. The eruption of molten rock from volcanoes and the egress of thermal waters and steam from subterranean regions in the form of hot springs and geysers are manifestations of the Earth's internal heat. These phenomena have always been apparent to the human inhabitants on every continent. With the development of underground mining, it became uncomfortably apparent to workers in the deeper levels that the temperature within the Earth generally increased with depth.

Because heat flows from the warmer to the cooler parts of a body, the observation that the Earth's temperature increases with depth clearly implies that heat is escaping from the interior by conduction through the rocky crust. William Thomson (Lord Kelvin) used the heat flow in estimating the age of the Earth. He argued that as the Earth cooled, following its formation and solidification from molten rock, the rate at which the temperature increased with depth (the geothermal gradient) lessened over time, and thus by measuring the geothermal gradient the length of time that the Earth had been cooling could be estimated. Such an estimate is based on the assumption that all the heat being lost is drawn from the endowment of heat that the Earth possessed at the time of its initially molten condition. This assumption was later shown to be untenable because a significant fraction of the Earth's heat flow is derived not from the initial heat but from the decay of radioactive elements present in trace amounts within the Earth. Thus the Earth's heat should be thought of not as a finite quantity acquired at the time of formation and gradually dissipated over time, but as a quantity that in large part has been continually produced, albeit in diminishing amounts throughout the history of the Earth. *See* EARTH, AGE OF; EARTH, CONVECTION IN.

Measurement. The quantitative measure of heat flow, q, is obtained as the product of the geothermal gradient, grad T, multiplied by the thermal conductivity, K, of the material through which the heat is passing: $q = K \text{ grad } T$. The determination of the gradient and conductivity differs in continental and oceanic settings. *See* CONDUCTION (HEAT); HEAT RADIATION; HEAT TRANSFER.

Continents. On continents the geothermal gradient, that is, the rate of increase of temperature downward, is ordinarily obtained by measuring the temperature at several points (or continuously) down a borehole. Most measurements are obtained in mineral exploration boreholes in the depth range 100–500 m (330–1650 ft); at shallower depths there are disturbances to the borehole temperature profile, principally due to circulating ground water, but also arising from the daily and annual temperature variations, irregular distributions of vegetation, topographic relief, erosion and deposition, climatic variations, and other causes. Holes of greater depth, while generally freer of disturbances in the deeper parts, are only infrequently available. Occasionally temperature measurements can be obtained in horizontal borings at different levels of an underground mine, and a geothermal gradient thereby determined. The thermal conductivity is measured on samples of rock obtained from the borehole as it was drilled. The conductivity is usually determined by a laboratory instrument in which a known quantity of heat is conducted through the sample and the temperature gradient established across the sample is measured. Other approaches to the determination of conductivity involve heating the surface of the rock sample with a heating wire or laser for a short time, and monitoring the rise and subsequent fall of the temperature of the rock. *See* DRILLING, GEOTECHNICAL.

Oceans. In oceanic measurements, a hollow probe several meters long equipped with temperature sensors at intervals along the length is allowed to plunge into the soft sediment of the ocean bottom. The temperature of the sediment is thus obtained at several depths, and the geothermal gradient established. As the probe plunges into the soft ocean sediment, a core is collected in the interior cavity and is brought to the surface when the probe is retrieved. The

conductivity of the soft sediment is obtained by inserting a heating needle into it and monitoring the rise of temperature with time as heat is supplied to the core. Some oceanic heat flow probes contain the heating needle internally and allow the direct measurement of conductivity in place while the probe is immersed in the ocean bottom.

The deep ocean floor is a thermally stable environment, generally free from diurnal and seasonal temperature variations that disturb equally shallow temperature measurements on continents. However, it is not free of temperature disturbances arising from the circulation of water through the fractures in the volcanic rock of the oceanic crust. This circulation, actually observed as ocean-bottom hot springs, is most active along the mid-ocean ridges, comprising the youngest areas of the oceanic crust. The hydrothermal circulation on the ridges gives rise in places to the spectacular "black smoker" venting of water at temperatures up to around 350°C (660°F), with important implications for life in the deep oceans. At the ridge crest, unusual biological communities exist at the hot water vents. Because of their relative youthfulness, the ridges have received less of a sediment cover than the older areas of ocean floor, where the ocean water is denied entry into the crustal fractures by a thick sedimentary blanket. *See* MID-OCEANIC RIDGE.

Global data set. In the first review of heat flow measurements in 1954, the total count of measurements both on continents and in oceans numbered only in a few tens. Subsequently, the data set has grown substantially, principally because of the recognition of the role of the Earth's heat in the internal dynamics and surficial tectonics of the planet. Measurements have been made at more than 20,000 sites around the globe in both continental and oceanic settings (**Fig. 1**). On a $5 \times 5^\circ$ grid, the observations cover 62% of the Earth's surface. The most intensive continental coverage has been in the United States, Europe, and Japan, with moderate coverage in Canada, Australia, India, and southern Africa. Gaps include Antarctica and much of Asia, Africa, and South America. The oceanic measurements have been concentrated in the latitude range 60°N to 30°S; the Arctic and southern oceans are only weakly represented with measurements. Despite the uneven distribution and incomplete coverage, a number of useful results bearing on the thermal and tectonic history of the continents and oceans have been obtained from studies of the heat flow data.

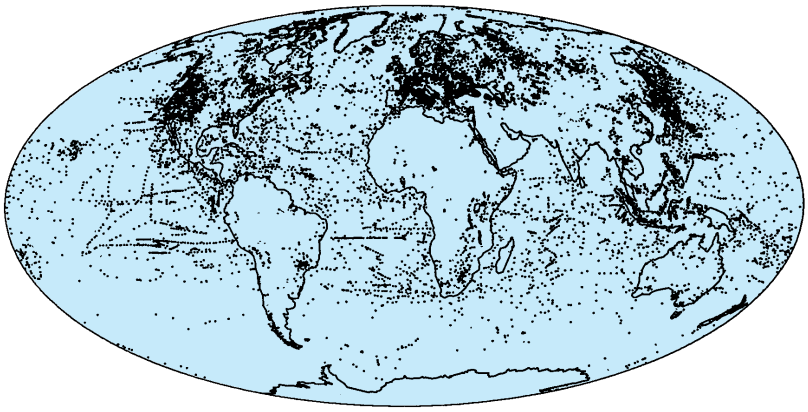

Fig. 1. Geographic distribution of heat flow measurement sites. (*After H. N. Pollack, S. J. Hurter, and J. R. Johnson, Heat flow from the Earth's interior: Analysis of the global data set, Rev. Geophys., 31(3):267–280, 1993*)

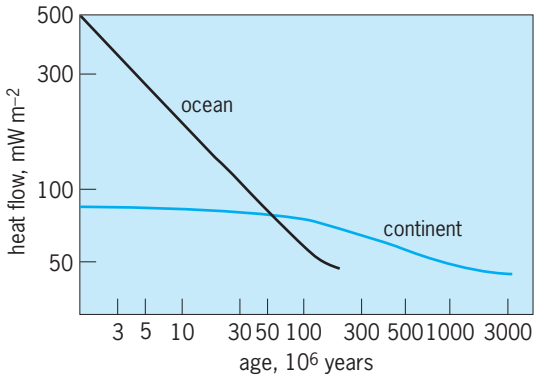

Fig. 2. Heat flow–versus–age relationships for continents and oceans. (*After D. Sprague and H. N. Pollack, Heat flow in the Mesozoic and Cenozoic, Nature, 285:393–395, 1980*)

Heat flow and age relationships. The most significant empirical relationship to emerge from the data studies is that heat flow generally decreases with increasing crustal age at the site of the measurement (**Fig. 2**). In the oceans the significant age is the time of magmatic emplacement of the basaltic crust at a mid-ocean ridge. On the continents it is usually the age of the last major tectonic, magmatic, or thermal metamorphic event to have affected the measurement site. The pattern of the decrease in heat flow with age differs from ocean to continent. While both settings show a heat flow in the older segments of about 40–45 milliwatts per square meter, the oceanic heat flow shows a greater range over a shorter time interval than does the continental heat flow.

The oceanic heat flow–versus–age curve is somewhat obscured in the younger oceanic crust because of seawater circulating through the basalt; however, measurements at sites where sediments have sealed the crust and prevented seawater penetration verify that the inverse square-root dependence on crustal age is valid in the younger crust. The decay of continental heat flow with age extends over a much longer period of time and is more complex than that of the oceans. The continental heat flow principally comprises (1) a radiogenic component derived from the enrichment of the continental crust in heat-producing radioactive isotopes, (2) a component derived from the cooling of tectonothermally mobilized lithosphere, and (3) a background heat flow of probable deep origin. *See* RADIOISOTOPE (GEOCHEMISTRY).

Radiogenic heat in Earth's crust. The heat production from radioactive decay contributes significantly to the heat flow from the Earth's interior. In fact, the entire surface heat flow on continents could arise within the crust if the surface concentrations of the important heat-producing isotopes such as are

Radioisotope concentrations and heat production in the Earth's crust and upper mantle and in chondritic meteorites*

Location	U, ppm	Th, ppm	K, %	Th/U	K/U, $\times 10^4$	Heat production, 10^{-6} W · m^{-3}*
Continental crust	1.25	4.8	1.25	3.8	1.0	0.80
Oceanic crust	0.08	0.20	1.08	2.5	1.3	0.16
Upper mantle	0.02–0.04	0.06–0.15	0.006–0.10	3.0–3.8	0.3–2.5	0.01–0.04
Chondritic meteorites	0.015	0.056	0.09	3.7	6.0	0.025

*U, uranium; Th, thorium; K, potassium; ppm, parts per million.

found in granites and gneisses persisted throughout the full thickness of the crust. However, probable lower crustal rocks such as granulites and migmatites have lesser concentrations of these isotopes, and so radiogenic heat production apparently decreases with depth in some manner. The existence of radiogenic heat and its generation throughout the history of the Earth proved to be the weak link in the Kelvin theory of the age of the Earth.

Radioisotopes. The principal heat-producing isotopes are thorium-232 (^{232}Th), uranium-238 (^{238}U), potassium-40 (^{40}K), and uranium-235 (^{235}U) with respective half-lives of 14.0, 4.47, 1.25, and 0.70×10^9 years. Other radioactive isotopes do not presently contribute significant heat because their decay chains are not sufficiently energetic, or their abundances are insignificant, or their half-lives are too short. The concentrations of uranium and thorium in rocks are generally in trace amounts measured in parts per million, while potassium is much more abundant, in the range of a few percent, of which a small but well-known fraction is ^{40}K (see **table**). A range of values for the upper mantle is characteristic of various possible mantle rocks as inferred from xenoliths brought to the surface by igneous and tectonic processes. If the entire mantle had the lesser concentrations, then crustal and mantle radiogenic heat production would be less than half the present-day heat loss; whereas if the greater concentrations were representative of the mantle as a whole, radiogenic heat production would be about equal to the present-day heat loss.

The abundances of these isotopes are commonly obtained from a multichannel gamma-ray spectrometer, which separates and counts the gamma-ray emissions on the basis of energy. Certain steps in the decay scheme of each isotope involve gamma radiation with characteristic energies, and thus the number of counts at certain levels of the energy spectrum is indicative of the isotopic abundance in the sample. Atomic absorption spectrometry can also be used to determine the amounts of the more abundant potassium, and mass spectrometry to determine the much smaller abundances of uranium and thorium. *See* ATOMIC SPECTROMETRY; MASS SPECTROMETRY.

The explanation for the relatively higher concentrations of uranium, thorium, and potassium in the continental crust resides in the relatively large ionic radii of these isotopes compared to the size of silicon, aluminum, magnesium, calcium, and iron, which combined with oxygen make up the bulk of the Earth's mantle as oxide minerals. The larger ions fit less easily into the compact crystal lattices of the mantle oxides, and in magmatic and metamorphic events they are more easily mobilized and tend to follow the magmatic products and metasomatic volatiles upward into the crust, where they are more easily incorporated in the more open crystal structures characteristic of the lower-pressure environment of the crust.

Heat flow province. The crustal radiogenic component of continental heat flow has been studied extensively. The principal observation is that, within a heat flow province, the regional variation in heat flow is well correlated with the heat production of the surface rocks. The relationship is expressed as $q_0 = q_r + bA_0$, where q_0 is the surface heat flow, A_0 is the heat production of the surface rocks, q_r is the reduced heat flow (the heat flow intercept for zero heat production), and b (the slope of the line) is a quantity with dimension of length that historically has been interpreted as a vertical scale length which characterizes the distribution of the heat sources with depth. The reduced heat flow is then interpreted as that flux that arises below the depth range defined by the empirical scale length. A heat flow province comprises a geographic area in which the heat flow and heat production are linearly related (**Fig. 3**); each heat flow province displays a characteristic reduced heat flow and scale length parameter. However, the traditional one-dimensional (vertical) interpretations have been questioned by investigators who argue that two- and three-dimensional heat transfer can also give rise to an apparent linear heat flow–heat production relationship, depending on both the horizontal and vertical scale lengths of crustal heterogeneities. In the multidimensional interpretations the linear relationship reveals more about the characteristic dimensions of the crustal heterogeneity than it does about the distribution of heat sources with depth.

Global heat loss. The heat flow–versus–age relationships (Fig. 2) that provide the critical constraints for thermal history models of the crust and upper mantle of both oceans and continents can be put to another useful purpose: estimating the heat flow in areas where no measurements have been made but for which the age of the crust is known. In a general way, the distribution of ages in the continental crust has been known for some time, and with the advent of the geomagnetic reversal time scale, large areas

Fig. 3. Geographic distribution of 17 heat flow provinces where the linear relationship between measured heat flow and radiogenic heat production from near-surface rocks has been empirically determined. (*After I. Vitorello and H. N. Pollack, On the variation of continental heat flow with age and the thermal evolution of continents, J. Geophys. Res., 85(B2):983–995, 1980*)

of the oceanic crust have been dated on the basis of their magnetic signatures. Thus it has become possible to estimate the heat flow in unsurveyed areas and map the heat flow over the entire surface of the Earth (**Fig. 4**).

The oceanic ridge system, where new oceanic lithosphere is produced, is important in delivering internal heat to the surface; fully half the Earth's current heat loss comes from oceanic lithosphere produced in the last 66 million years, representing only 31% of the Earth's surface. Other regions of above-average heat flow include the cordillera of western North America, alpine Europe, and the marginal basins of the western Pacific. Areas of low heat flow include all major continental shields and platforms and older ocean basins. Regions where lithospheric plates slide past each other, such as the San Andreas Fault of western North America, show no special heat flow signature, suggesting that friction between the plates is insufficient to generate anomalous heat. The mean heat flow through continents (including marine continental shelves) and oceans is 65 and 101 $mW \cdot m^{-2}$, respectively. The speculation that the heat flow through oceanic and continental crust is approximately equal, advanced before the relationships between heat flow and tectonic age were well established, is no longer tenable. *See* LITHOSPHERE.

The average heat flow over the entire Earth is 87 $mW \cdot m^{-2}$ and is comparatively a trickle; it is sufficient to bring a thimbleful of water to a boil in about 2 years, or if collected over the area of a football field, would be adequate to light four 100-watt incandescent bulbs. The radiant energy from the Sun intercepted by the Earth is some 4000 times greater than the geothermal flux. The absorption and reradiation of the incident solar energy are the principal processes that determine the surface temperature of the Earth. The geothermal energy is of little significance to the surface temperature but is of paramount importance in considering the Earth's internal thermal condition. *See* GEOLOGIC THERMOMETRY; GEOTHERMAL POWER.

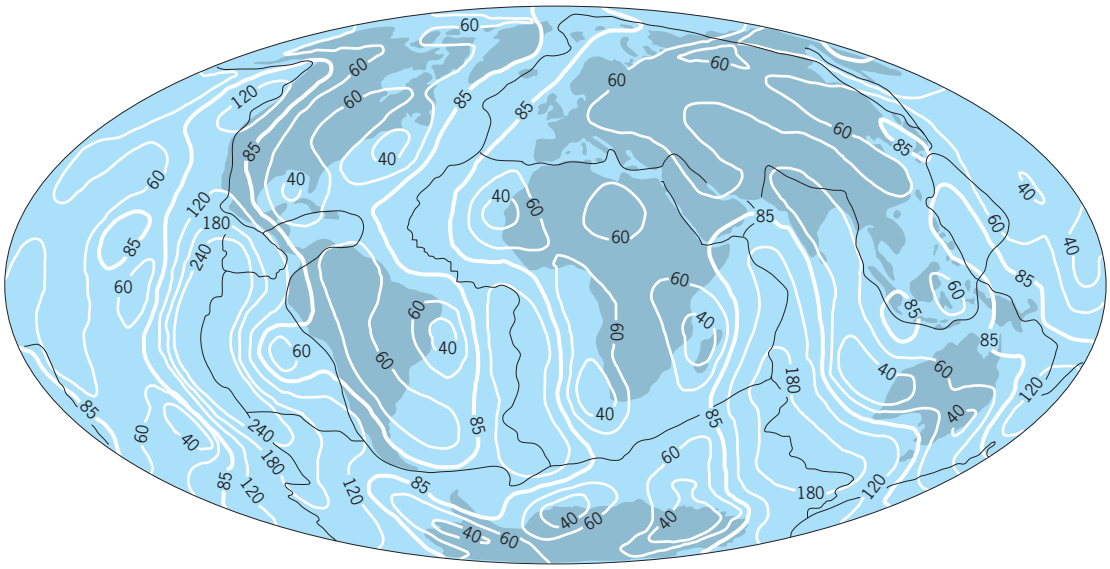

Fig. 4. Present-day global heat flow (degree 12 spherical harmonic representation). Contours are in milliwatts per square meter. (*After H. N. Pollack, S. J. Hurter, and J. R. Johnson, Heat flow from the Earth's interior: Analysis of the global data set, Rev. Geophys., 31(3):267–280, 1993*)

Heat flow in the past. The present-day heat flow from the Earth has commonly been considered a useful if not precise upper bound on the quantity of heat being produced radiogenically in the interior. However, if the heat flow has a temporal variability on a scale that is short compared to the half-lives of the principal heat-producing isotopes, the present-day heat flow must be seen as a much coarser constraint on the Earth's internal heat generation.

Just as the present-day heat loss can be determined from the heat flow–versus–age relationships and the present-day distribution of crustal ages, so can the heat loss at a past time be calculated, provided the age distribution of the crust at that time can be estimated. For the past 180 million years, the age distribution of the oceanic crust at a given time can be determined by subtracting from the present-day distribution all oceanic crust produced from that time to the present and resurrecting and reinstating crust which has been subducted in the interval from that time to the present. The age distribution of the restored lithosphere can be estimated from the present-day age distribution, and the probability that crust produced at an earlier time will have survived to the present day. Continents can be treated similarly, but with different survival probabilities which reflect the longevity of the continents. This method has yielded estimates of the mean heat flow in oceans and continents and the global flux for the past 180 million years. The global mean heat flux, presently 87 $mW \cdot m^{-2}$, has fluctuated in a range of −5% to +15% around the present-day value. The mean heat flow exceeded the present-day flux in the interval 60–100 million years before present, when the rate of production of oceanic crust was high. The mean age of the ocean floor, presently 62 million years, has ranged between 45 and 65 million years, accounting for most of the variation in the global heat loss.

For times earlier than 180 million years ago, an analysis based on the present-day age distribution and survival probabilities becomes unfeasible because in the oceans, where 70% of the global heat loss takes place, there is no oceanic crust surviving with which to estimate past rates of production of the sea floor. Moreover, in studying the Earth's heat loss over intervals of several hundred million to billions of years before the present, it is necessary to consider the half-lives and relative abundances of the principal heat-producing isotopes and the loss of primordial heat acquired during planetary accretion and core segregation. The isotopic endowment of the bulk Earth is unknown but is sometimes assumed to be similar to the chondritic meteorites (see table), which are thought to be representative of the primitive material from which the Earth differentiated. However, the rocks of the Earth's crust have a lower potassium-uranium ratio than do the chondrites. Because of the relatively short half-life of ^{40}K, the heat production 3 billion years ago from a chrondritic composition would have been more than four times the present-day heat production, while a terrestrial composition, based on observed potassium-uranium ratios of crustal rocks, would yield heat at only somewhat more than twice the present rate. *See* METEORITE.

Generalized model calculations of the thermal history of the Earth using the potassium-uranium ratio of terrestrial crustal rocks show that the present-day heat production replaces much but not all of the present-day heat loss. The inequality between heat production and heat loss arises because a finite transit time is required between the time when heat is produced within the body of the Earth and when it escapes across the outer surface, during which interval the heat sources have decayed. The present-day heat loss thus can be seen as older radiogenic heat, delayed in its escape from the Earth's interior by the heat transfer process. Because the replacement heat production is less than the heat loss, the Earth is necessarily cooling. In spite of the various uncertainties about the blend of the nuclear fuel, the thermal history calculations indicate that the heat loss from the Earth 3 billion years ago was at least twice as great as the present-day heat loss. However, the augmented heat flow may not have been distributed uniformly over the Earth. Rather, the increased heat loss may have been best accommodated by a more rapid rate of production of oceanic crust, with only a modest increase of heat flow through the continents. This latter speculation derives from crustal temperatures inferred from metamorphic rocks in ancient continental terrains, which suggest that the heat flow entering the base of the continental lithosphere was approximately the same as today. Henry N. Pollack

Bibliography. H. N. Pollack, S. J. Hurter, and J. R. Johnson, Heat flow from the Earth's interior: Analysis of the global data set, *Rev. Geophys.*, 31(3):267–280, 1993.

Earth crust

The low-density outermost layer of the Earth above the Mohorovičić discontinuity (the Moho), a global boundary that is defined as the depth in the Earth where the compressional-wave seismic velocity increases rapidly or discontinuously to a value in excess of 4.7 mi/s (7.6 km/s; the upper mantle). The crust is also the cold, upper portion of the Earth's lithosphere, which in terms of plate tectonics is the mobile, outer layer that is underlain by the hot, convecting asthenosphere. *See* ASTHENOSPHERE; LITHOSPHERE; MOHO (MOHOROVICIC DISCONTINUITY); PLATE TECTONICS.

Continental crust. The Earth's continental crust has evolved over the past 4 billion years, and is highly variable in geologic composition and internal structure. The worldwide mean thickness of continental crust is 24 mi (40 km), with a standard deviation of 5.4 mi (9 km; **Fig. 1**). The thinnest continental crust (found in the Afar Triangle, northeast Africa) is about 9 mi (15 km) thick, and the thickest (the Himalayan Mountains in China) is about 47 mi (75 km) thick. Ninety-five percent of all continental crust has a thickness within two standard deviations of the mean thickness, between 13 mi (22 km) and

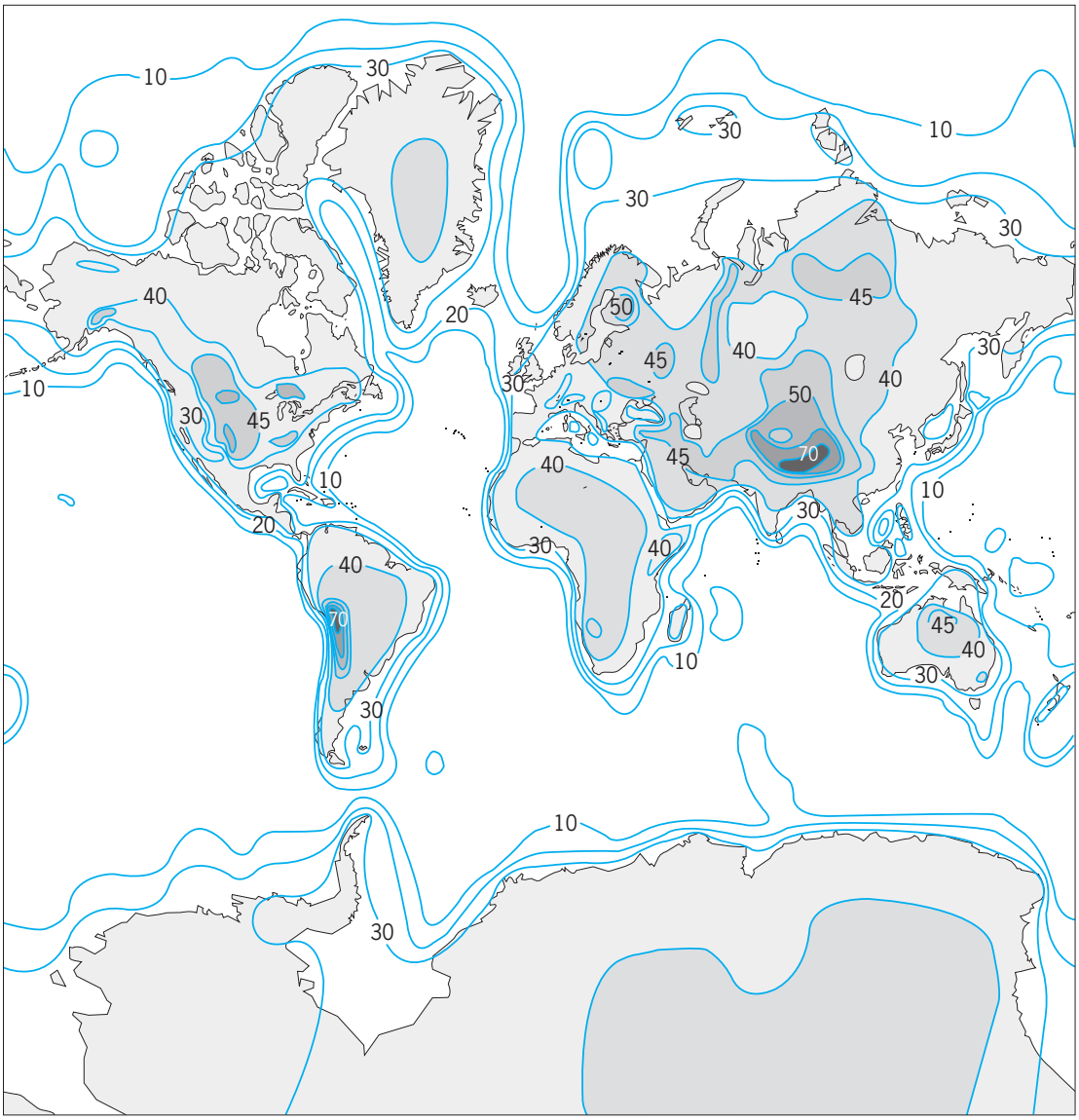

Fig. 1. Mercator projection of crustal thickness (10-km contour interval plus 45-km contour) based on seismic refraction profiles.

37 mi (58 km). The Antarctic has a crustal thickness of 24 mi (40 km) in the ancient, stable (cratonic) region of East Antarctica, and about 12 mi (20 km) in the recently stretched (extended) crust of West Antarctica. Continental margins, which mark the transition from oceanic to continental crust, range in thickness from about 9 mi (15 km) to 18 mi (30 km). *See* CONTINENTAL MARGIN.

Seismic refraction profiles. The internal structure of the crust reveals much about crustal evolution, and one effective way of describing this structure is based on the results of seismic refraction profiles that define crustal layers (that is, layers with characteristic compressional-wave seismic velocity). This seismological method for investigating the crust is widely used for two reasons: seismic refraction measurements are made on a global basis, and the seismically defined layers correspond to geologic layers. There is a relationship between the compressional-wave seismic velocity, density, and rock composition for coarse-grained igneous rocks (**Fig. 2**). Seismic reflection profiles can be used to relate seismic velocities to crustal composition and density. A composite crustal section can be deduced from these data. *See* SEISMOLOGY.

Despite its geologic complexity, the continental crust may generally be divided into four layers: an uppermost sedimentary layer, and an upper, middle, and lower crust composed of crystalline rocks (**Fig. 3**). The sedimentary cover of the continental crust is an important source of natural resources. This cover averages 0.6 mi (1 km) in thickness, and varies from 0 (for example, on shields) to more than 9 mi (15 km) in deep basins. In stable continental crust of average thickness (25 mi or 40 km), the crystalline upper crust is commonly 6–9 mi (10–15 km) thick and has an average composition equivalent to a granite. The middle crust is 3–9 mi (5–15 km) thick and has a composition equivalent to a diorite; and the lower crust is 3–12 mi (5–20 km) thick and has a composition equivalent to a gabbro. Owing to increasing temperature and pressure with depth, the

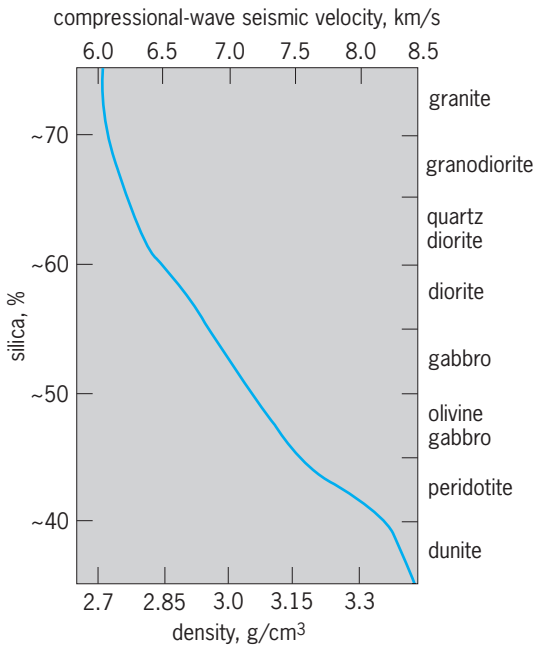

Fig. 2. Relationship between compressional-wave seismic velocity, igneous rock type, silica content, and density. Seismic refraction profiles provide a measurement of compressional-wave velocity as a function of depth in the crust.

metamorphic grade of rocks increases with depth, and the rocks within the deep continental crust generally are metamorphic rocks, even if they originated as sedimentary or igneous rocks (**Fig. 4**). *See* DIORITE; GABBRO; GRANITE; METAMORPHIC ROCKS.

Crustal properties vary systematically with geologic setting, which may be divided into seven groups: orogens (mountain belts), shields and platforms, island arcs (volcanic arcs), continental magmatic arcs, rifts, extended (stretched) crust, and forearcs (Fig. 3). Orogens are typified by thick crust [average thickness is 29 mi (46 km), but the maximum thickness is as much as 47 mi (75 km) in the Himalayas of Asia and Andes of South America]. The seismic structure of orogens is highly variable and is related to the nature of the orogen (for example, continent-continent collision versus ocean-continent collision). However, a thick upper crustal layer is common, and this is the result of thickening of the upper crustal section by wedging and stacking of the compressed upper crust. Shields and platforms, such as the Canadian Shield and the Russian Platform, commonly have an approximately 26-mi-thick (42-km) crust, including a 3–6-mi-thick (5–10-km) high-velocity (4.4–4.7-mi/s or 7.1–7.5-km/s) lower crust. In comparison with shields, island arcs (such as Japan) have thinner crusts and significantly shallower middle and lower crustal layers due to the intrusion of mafic (that is, low silica content) plutons. Continental magmatic arcs, such as the Cascades volcanoes of the northwestern United States, intrude preexisting continental crust, and therefore they are generally 3–9 mi (10–15 km) thicker than island arcs. Continental rifts, such as the East African and Rio Grande rifts, have an average crustal thickness of about 22 mi (36 km), but show great variability in seismic structure. Extended continental crust, such as the Basin and Range Province of the western United States, averages 18 mi (30 km) in thickness, and has relatively little high-velocity (greater than 4.2 mi/s or 6.8 km/s) crustal material. Forearcs are regions that were formed by the coast, near volcanic arcs, including much of the west coast of North America. They typically have thin crust,

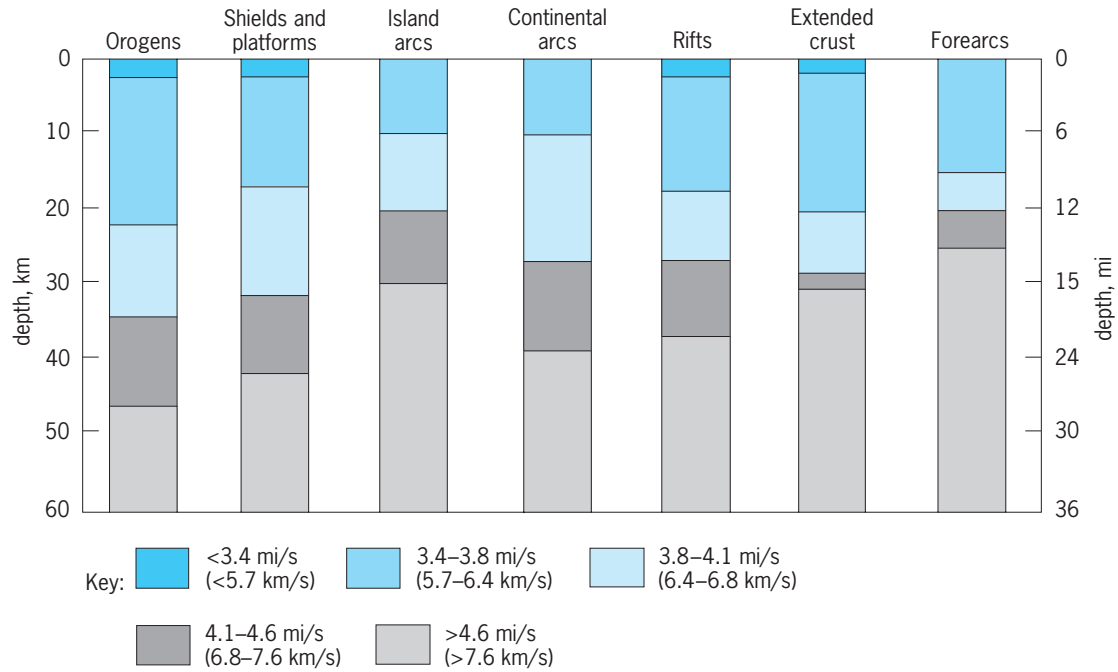

Fig. 3. Average crustal columns for seven geologic provinces, showing variations in crustal thickness (light gray areas correspond to the mantle). Compressional-wave velocity, which has been divided into five ranges, is directly related to geologic composition, as indicated in Fig. 2. (*After W. Mooney et al., JGR, 1998*)

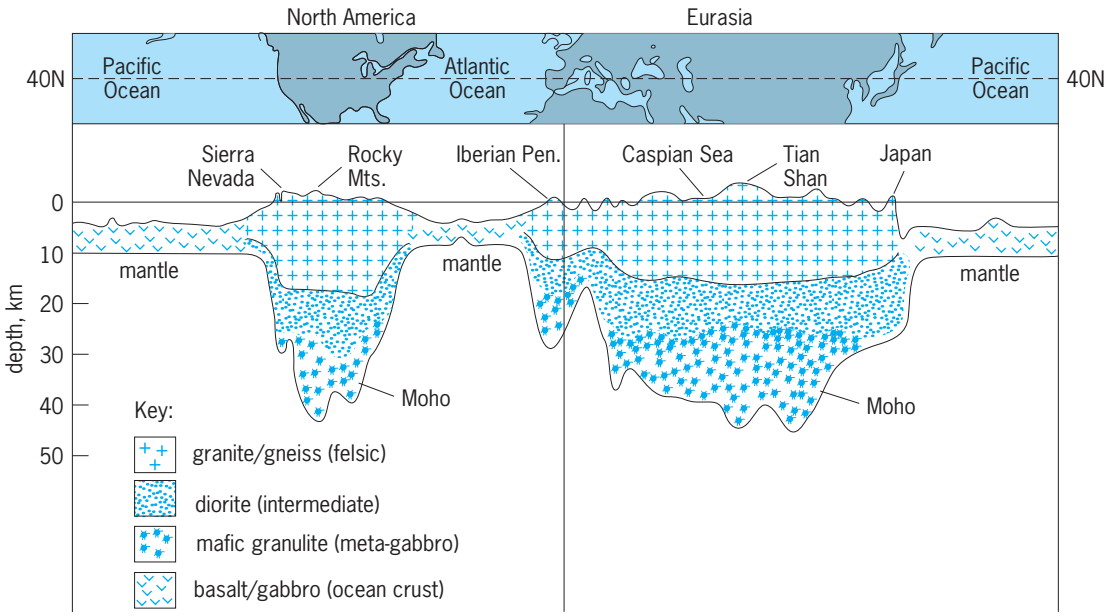

Fig. 4. Cross section of the Earth's crustal thickness at 40°N. Felsic rock is found throughout the upper continental crust above diorite and mafic granulite. The crust is thicker beneath mountain ranges and other regions of high elevation. Vertical thickness relative to horizontal distance is exaggerated 100×. (*After W. Mooney, IASPEI, 2004*)

about 15 mi (25 km), and have a thick (9 mi or 15 km) upper crustal section that consists of relatively low-density metasedimentary rocks. *See* NORTH AMERICA; OCEANIC ISLANDS; PLUTON; RIFT VALLEY; SEDIMENTARY ROCKS; VOLCANO.

Seismic reflection profiles. Seismic reflection profiles use vertically traveling waves to obtain a high-resolution image of the crust, whereas refraction profiles use horizontally traveling waves to obtain seismic velocity information. Reflection profiles provide images of the deep crust that may be interrupted in terms of lithologic, metamorphic, and porosity layering, dipping fault zones, and (more rarely) the presence of crustal melts. Distinct reflection patterns that correlate with geologic setting can also be identified. These patterns provide important information on how the crust has been formed and modified. In many regions, particularly in extended crust, the lower crust contains numerous horizontal reflections that terminate at the depth of the Moho. These reflections are most likely the result of igneous intrusions into the lower crust, combined with horizontal stretching of the warm and ductile lower crust. *See* IGNEOUS ROCKS.

Seismic reflection profiles across mountain belts (orogens) show near-horizontal sheets of rock (nappes) in zones of crustal thickening and deformation. These upper crustal nappes can be followed to midcrustal levels in the Alps, and what has been termed thin-skin tectonics [the horizontal transport of 3–6 mi-thick (5–10-km) sheets of upper crustal material over large distances] has been confirmed in the Appalachians of the eastern United States and in western Europe. *See* CORDILLERAN BELT; MOUNTAIN SYSTEMS.

Seismic reflection profiles in Precambrian crust show pronounced, widespread structural features dating from that period, indicating long-term thermal and mechanical stability. It is possible that the geometry of these structures is qualitatively consistent with plate tectonic processes operating 2 billion years ago. *See* PRECAMBRIAN; SEISMOLOGY.

Geoelectrical measurements. Measurements of electrical conductivity within the deep continental crust have revealed high conductivities at midcrustal depths that provide important information on crustal properties. High crustal conductivity can be caused by several factors, including the presence of saline fluids, partial melts, sulfide minerals, or carbon films. In regions of high heat flow, such as actively extending crust, high crustal conductivities are often found at midcrustal depths (6–12 mi or 10–20 km), and may indicate the melting of wet granitic rocks or the presence of fluids produced by metamorphic dehydration reactions. In young mountain belts, such as the Carpathian Mountains of eastern Europe and the Alaska Range of central Alaska, dipping high-conductivity zones can be correlated with large carbon-rich flysch and black shale basins that have been crushed and deeply buried during the mountain-building process. *See* BLACK SHALE; GEOELECTRICITY; OROGENY; TURBIDITE.

Continental crustal evolution. The properties of the continental crust provide important constraints on the process of crustal evolution. The continents coalesce and disperse as ocean basins open and close, and continental crust is deformed and modified during rifting and at ocean-continent and continent-continent collisions. A net increase in the total volume of continental crust requires extraction of crustal material from the mantle and the stabilization of this material within a continental block.

At least three processes provide new continental crust. The first is the accretion and consolidation of island arcs, such as Japan or the Aleutian Islands, onto a continental margin. The second process is the

tectonic underplating of oceanic crust at active subduction zones. In this process, the continental crust grows from below as oceanic crust is welded to the base of the continental margin, either when subduction stops or when subduction steps oceanward and a new trench is formed. This process has been identified in western Canada and southern Alaska. The third process is the magmatic inflation of the crust at continental arcs, rifts, and regions of crustal extension. This process has been identified in many regions. *See* GEODYNAMICS.

Although both the accretion of island arcs and tectonic underplating of oceanic crust can be documented as stages of continental growth, these processes by themselves would result in a bulk crustal composition that is more mafic (low silica content) than the actual bulk composition of continental crust. This apparent discrepancy can be resolved in at least two ways. First, the mafic lower crust of some island arcs may be delaminated (peeled off from below) and subducted back into the mantle, thereby leaving the remaining crust less mafic. Second, an originally thick mafic lower crust may undergo secondary differentiation that produces an intermediate composition melt and a mafic-ultramafic residue (restite), which ponds at the base of the crust. After cooling, the restite would have a seismic velocity of about 5 mi/s (8 km/s) and therefore would constitute a new, shallower seismic Moho with a less dense, less mafic overlying crust. Walter D. Mooney

Oceanic crust. The surface of the ocean crust, except for some locally high volcanoes and plateaus, resides some 1–3 mi (2–5 km) below sea level, approximately 3–6 km below the average level of the continents. While the continental crust consists mainly of ancient mountain belts, with an average age of 1500 Ma, oceanic crust is relatively young and dynamic, having a maximum age of only 200 Ma. Most of it was produced at mid-ocean ridges during sea-floor spreading, where ridges define the accret-

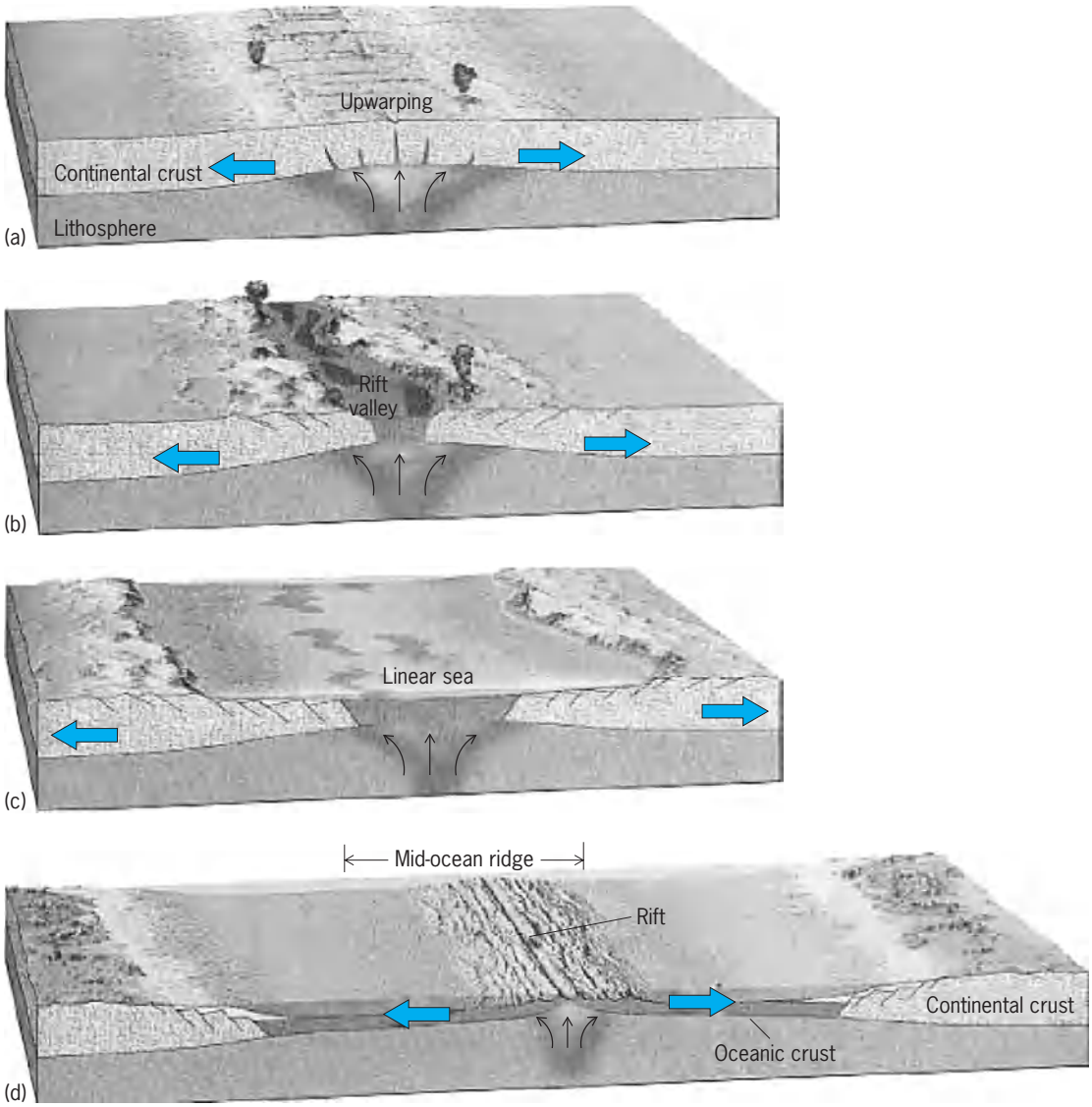

Fig. 5. The sea-floor spreading mechanism begins with continental rifting initiated by magmatic processes (*a*). Once a basin has rifted apart significantly (*b*), just as in the East African Rift Valley, it may fill with water, eventually forming a lake or sea (*c*). Over time, the extrusion of magma forms new crust and the ridge continues to spread, creating a vast ocean (*d*).

ing boundaries of major lithospheric plates moving about the surface of the Earth. Thus, the oldest rocks of the ocean crust date back no earlier than the rifting episodes that initiated the most recent phase of continental drift (the Pangaean breakup) in Late Jurassic times. *See* JURASSIC; MID-OCEANIC RIDGE.

Sea-floor spreading. The sea-floor spreading mechanism holds that the ocean basins, for example the Atlantic and Indian oceans, developed after continental rifting. During sea-floor spreading, magmatic upwelling at the central mid-ocean ridge repeatedly introduces new basaltic crust along a narrow median rift, in the center of a basin within a global system of underwater mountains. This system of mountains, which is over 12,000 mi (20,000 km) long, is interconnected through all the major ocean basins, including the Pacific and Arctic basins. It has been determined that this process of rifting, basaltic magmatism, and continued sea-floor spreading has produced the vast majority of ocean crust (**Fig. 5**). The rates of plate motions are deduced from the records of Earth's magnetic reversals preserved in the basaltic rocks surrounding ocean ridges. While magnetic domains in liquid magma will rotate according to the Earth's magnetic field, once lava solidifies, the orientation of material is no longer flexible and the basalt will "freeze," reflecting the direction of the Earth's magnetic field at its time of emplacement. The anomalies, which are mapped with magnetometers towed from ships, document the shifting patterns of plate motion in the ocean basins, as well as the history of sea-floor spreading. They demonstrate that the ocean crust is one vast, ~6-km-thick (4-mi) basaltic lava field covering nearly 70% of the Earth's surface. *See* BASALT; BASIN; CONTINENTAL DRIFT; EARTH, CONVECTION IN; LAVA; MAGMA; MAGNETOMETER; ROCK MAGNETISM.

The extent of this lava field and its fundamental basaltic composition have been documented by extensive dredge sampling programs along active ridges, the transform faults that offset them, and the fracture-zone scars that trail far off onto the flanks of the ridges. The older portions of the ocean crust have been sampled in dozens of localities by drilling through a carapace of mainly biogenic marine sediments. In all cases, the continuity of the spreading process has been confirmed. Thus, emplacement of basalt at mid-ocean ridges is the universal mechanism by which ocean crust is formed. *See* MARINE SEDIMENTS; TRANSFORM FAULT.

Characteristics. The base of the ocean crust is defined by the location of the seismic Moho. The Moho is usually an abrupt transition between rocks with average seismic velocities of about 4.2 mi/s (6.7 km/s) to rocks with velocities of about 5 mi/s (8 km/s). The former velocities agree with experimental determinations on varieties of nonporous mafic rocks including gabbros and amphibolites (rocks that have significant proportions of plagioclase and little olivine), whereas the latter velocities can represent only nearly fresh (nonserpentinized) peridotite, consisting mainly of olivine and pyroxenes with virtually no plagioclase. In the oceans, explosion seismology and seismic reflection carried out from ships are used to determine both the location of the Moho and the structure of the overlying crust. Both gabbros and serpentinized peridotites are present at various tectonic exposures in the ocean crust, such as along fracture zones, the walls of median rifts, and low-angle detachment surfaces in the floors of rifts. Thus, it is likely that gabbros are the rocks of the lower ocean crust that overlie mantle peridotites. However, there is no known place in the ocean basins where a coherent crust-mantle transition is exposed in outcrop. *See* AMPHIBOLITE; PERIDOTITE.

There are fault slices of former ocean crust on land, known as ophiolites, where nearly complete cross sections through the crust can be mapped and sampled (**Fig. 6**). These are robust indications that the ocean crust consists (from top to bottom) of submarine extrusives (usually pillow basalts), feeder dikes (often vertically sheeted), or sills, gabbros, and peridotites. There is much uncertainty, however, about the extent to which typical ophiolites, most of which formed in island-arc or backarc environments, can represent abyssal ocean crust, which is produced at the major accreting plate boundaries. Moreover, due to the extent of alteration and metamorphism, particularly in ultramafic sections, the rocks in ophiolites are sometimes not well correlated with those in ocean crust. Thus, seismic models of the ocean crust and the nature of the Moho transition therein cannot be determined directly. *See* OPHIOLITE.

Nevertheless, the general seismic model of the ocean crust and the basic ophiolite model of crustal stratigraphy have generally been confirmed both by direct observation and by drilling. Sheeted dikes, for example, have been observed from submersibles on

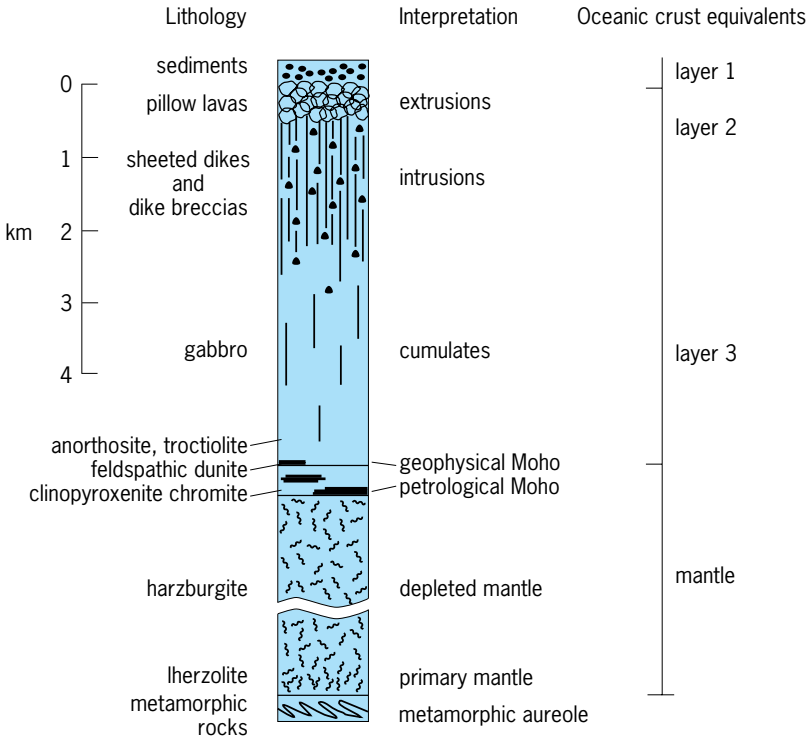

Fig. 6. Interpretation of Bay of Islands ophiolite crust as oceanic crust and mantle, with the various thicknesses estimated from geologic mapping. (*After R. G. Coleman, Ophiolites, Springer, 1977*).

faulted exposures in both the equatorial Atlantic and eastern Pacific. Seismic velocities in the upper part of the ocean crust are consistent with the presence of pillow lavas or flows with many cracks and cavities. Velocities increase systematically with depth, indicating that this porosity structure diminishes, primarily because of compaction, filling of veins with alteration minerals, and transition to compact dikes. The deepest hole drilled in ocean crust, on the flank of the Costa Rica Rift in the eastern equatorial Pacific, reflects these relationships in downhole logging measurements through 1 mi (1.8 km) of igneous material. The transition to rocks with velocities of about 4.2 mi/s (6.7 km/s) is actually within the compacted dike sequence, and does not formally correspond to the lithologic transition to gabbros, which still has not been identified. *See* PETROLOGY; STRATIGRAPHY.

Crustal formation. The composite lithology of gabbros, dikes, and extrusives forms ocean crust above the mantle in the following way. Gabbros comprise the frozen walls, roof, and floor of a spreading ridge's magma-chamber complex, which is more or less continuously supplied with magma from the upper mantle. The flux of magma, and thus the details of this process, differ according to whether the ridge is spreading rapidly or slowly, or is near an unusually hot region of mantle (that is, Iceland). The magma is supplied because of decompression melting of hot peridotite as it convects upward or rises diapirically into the region beneath the spreading ridge. Melts are driven out of the mantle rocks because they are under stress. While they are still molten, they coalesce, flow into fissures, and feed into the lower crust, where they cool and begin to crystallize. *See* DIAPIR; HOT SPOTS (GEOLOGY).

The steady-state process of sea-floor spreading, however, requires that magmas be supplied to the lower crust nearly continuously. This forces magmas already residing in the crust out toward the surface. Elastic materials will initially respond to stress by bending or stretching, but not by breaking (similar to taffy or putty), in a process called ductile deformation. If the rate of stress is fast enough, then rocks will fracture under brittle deformation. At high temperatures (perhaps 1100–1300°F or 600–700°C), rocks will tend to experience only ductile deformation. At lower temperatures, they will more likely experience brittle failure. Earthquakes, in particular, are evidence for large-scale brittle failure at some depth in the crust or upper mantle.

Rocks in the uppermost crust, chilled by hydrothermal circulation of seawater, often experience brittle failure along vertical fractures under the combined impetus of upward magma injection and sea-floor spreading. The tensional stress regime at spreading ridges determines that most fracturing will occur in a narrow zone parallel to the ridge (**Fig. 7**). Thus, dikes tend to be sheeted, and when these breach the sea floor, eruptions occur. Lava flows through dikes will be interrupted by more dikes as the ridge continues to spread, and repeatedly buried by younger flows after being carried to one side of the ridge or the other. Observations made from drilling and submersible instruments indicate

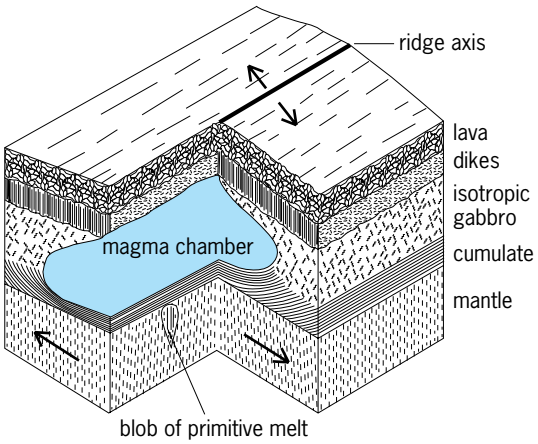

Fig. 7. Open-system magma chamber that may exist beneath all active mid-oceanic ridges, except those that are spreading extremely slowly. (*After M. Wilson, Igneous Petrogenesis, Routledge, Chapman, and Hall, 1988*).

that this process takes place perhaps 10–20 times in the course of building a given segment of crust at rifted, slowly spreading ridges such as the Mid-Atlantic Ridge. Lava flows are often thinner along the much faster spreading East Pacific Rise and, therefore, more eruptive events may be required to build the same length of crust there. This speculation, though, has not yet been confirmed by drilling.

Coarse-grained gabbros crystallized from the magma chamber complex comprise approximately two-thirds of ocean crust. They may exist as solids in the deep crust due to the hydrothermal processes acting at the top of the crust that cool the steady-state crystallizing mass to nearly 90°F (50°C) below the average temperature of magmas arriving from the mantle. This cooling mechanism will crystallize at least 50–70% of molten material at a fast-spreading ridge before it can erupt. At slowly spreading ridges, where magma is supplied more gradually, hydrothermal cooling drives crystallization in the deep crust almost to completion. At both rapidly and slowly spreading ridges, crystallization occurs by the plating out of minerals (olivine, plagioclase, and clinopyroxene) along the walls of intrusive bodies in the gabbroic mass, all within close range (a few kilometers) of the center of spreading. The crystallization forces changes in the compositions of melts and thus drives magmatic differentiation in the ocean crust. The sequence of liquid compositions produced by this process changes from olivine basalt, to titanium-rich ferrobasalt, to rarely erupted ferroandesites and dacites. The latter liquids often remain in the lower crust, crystallizing to oxide-rich gabbro and plagiogranite. Occasionally, at slowly spreading ridges, rift dynamics will allow normal faults to penetrate the bodies of crystallizing magma-chamber complexes. This produces shear zones along which melts may migrate. Thus, dredged and drilled gabbroic rocks from these environments are often lithologically complex. While the deformed rocks may be highly sheared, with gneissic or porphyroclastic textures, they are often cut by younger veins of oxide gabbro or granitic material. *See* DACITE; GABBRO; GNEISS; OLIVINE.

The fundamental melt product at mid-ocean ridges is variously termed abyssal tholeiite or mid-ocean-ridge basalt. In most places, this parental material is remarkably uniform in its fundamental petrologic and geochemical properties. Thus, it has a consistent crystallization sequence from place to place, and in most areas possesses depleted isotopic and trace-element signatures. The source of most abyssal tholeiites is a vast, globally distributed shallow mantle reservoir that previously experienced partial melting (thus losing radioactive parents to the radiogenic isotopes) at some distant time in the geological past. Currently, most tholeiites are experiencing further partially melting. There are places along spreading ridges, however, where less depleted or even enriched basaltic lavas are erupting. Most of these are in elevated, hot-spot regions such as Iceland and the Azores in the North Atlantic, Bouvet in the South Atlantic, Reunion in the Indian Ocean, and Easter Island in the Pacific. The ocean crust near these islands tends to be both shallower and thicker than average, in addition to being composed of isotopically distinctive basalts. The general processes of crustal accretion at these places, though, are approximately the same as elsewhere in the ocean crust. *See* EARTH; FAULT AND FAULT STRUCTURES; ISOSTASY; STRUCTURAL GEOLOGY; SUBDUCTION ZONES. James H. Natland; Walter D. Mooney

Bibliography. A. W. Bally and A. R. Palmer (eds.), *The Geology of North America: An Overview*, vol. A, Geological Society of America, 1989; R. G. Coleman, *Ophiolites*, 1977; A. E. J. Engel, H. L. James, and B. F. Leonard, *Petrological Studies: A Volume in Honor of A. F. Buddington*, Geological Society of America, 1962; D. M. Fountain, R. Arculus, and R. Kay (eds.), *Lower Continental Crust*, 1992; I. G. Gass, S. J. Lippard, and A. W. Shelton (eds.), *Ophiolites and Oceanic Lithosphere*, Geol. Soc. London Spec. Publ. 13, 1984; R. Meissner, *The Continental Crust: A Geophysical Approach*, 1986; W. Mooney et al., *Crustal 5.1: A Global Crustal Model at 5° by 5°*, 1998; W. Mooney, *Seismic Handbook*, International Association of Seismology and Physics of the Earth's Interior (IASPEI), 2004; A. Nicolas, *Structure of Ophiolites and Dynamics of Oceanic Lithosphere*, 1987; L. C. Pakiser and W. D. Mooney (eds.), *Geophysical Framework of the Continental United States*, Geol. Soc. Amer. Mem. 172, 1990.

Earth interior

All of the Earth beneath the land surface and the ocean bottom, including the crust, the mantle, and the core. The interior is not accessible to direct observation. Nevertheless, a rather detailed model has been constructed on the basis of measurements made at or above the surface. Measurements of gravity, the geomagnetic field, surface heat flow, and surface deformation can all be used to put constraints on the Earth model, but the most detailed information about the interior is provided by seismic measurements. To the nonspecialist, seismic methods are perhaps best known for their application to the oil exploration industry, where seismic data are used to map the subsurface structure of sedimentary basins. In the exploration of the Earth's interior, the seismic waves being analyzed are usually generated by earthquakes, and measurements are made of waves propagating through the interior of the body (body waves), waves propagating along the surface (surface waves), and standing waves bringing the whole Earth into a state of oscillation (free oscillations). Such measurements, when properly interpreted, provide information about seismic-wave velocities in the Earth. On the other hand, seismic-wave velocities can also be measured in laboratory experiments where rock samples are subjected to the high pressures and temperatures typical of conditions in the deep interior. Meteorites provide rock samples of materials that are probably abundant in the solar system. The comparison of laboratory and field measurements thus leads, by inference, to a model where the composition and temperature distribution can be specified to some extent. What emerges from these studies is a picture not only of the structure but also of the evolution and dynamics of the Earth's interior.

Earth among the planets. The Earth is a planet with a mean radius of 3959 mi (6371 km). The principal departure of the Earth from a spherical shape is an equatorial bulge caused by the centrifugal effect of the Earth's rotation. The geometrical flattening, which is defined by the ratio $(a - c)/a$, where a and c are the equatorial and polar radius, respectively, is estimated to be about 1 part in 298. Internal layers similarly exhibit a flattening, which gradually decreases toward the center of the Earth. The mean density of the Earth is 3.189 oz/in.^{3} (5.517 g/cm^3), and the moment of inertia is 0.527 times that of a body of constant density with the same size and mass as the Earth. The fact that this factor for the Earth is much smaller than unity is one of the clearest indications that the Earth has a central core of greater density than the surrounding mantle. A plausible assumption consistent with the composition of meteorites is that of a stony mantle composed mainly of minerals rich in silicate and magnesium and a core composed mainly of iron. The Earth is probably not unique in that respect. It is generally assumed that the planets and meteorites formed at the same time about 4.5 billion years ago, shortly after the origin of the Sun. From the matter that was then available, most of the lighter gaseous material was amassed by the large outermost planets, while the four smaller innermost earthlike planets and the meteorites consist mainly of heavier rocky material. The mass and moment of inertia of the earthlike planets are all consistent with the model of an iron-rich core surrounded by a more silicate-rich mantle, although the relative proportions may be different for different planets, as shown in **Fig. 1**. Venus is the planet resembling Earth most closely in size and mass, yet there are observable differences in behavior of the two planets. Of interest for comparison with the Earth's interior is the virtual absence of a magnetic field produced by Venus. *See* SOLAR SYSTEM.

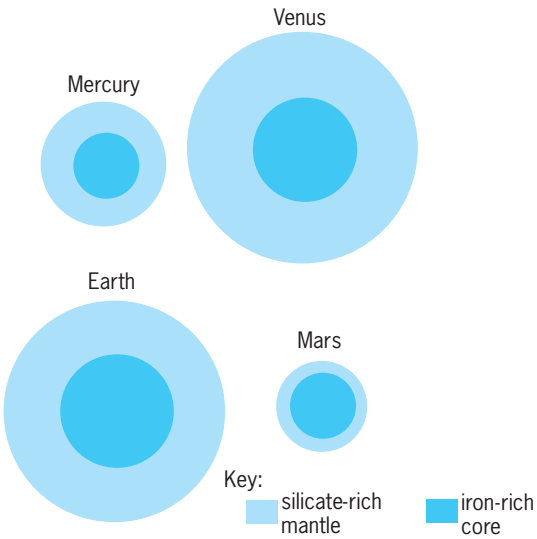

Fig. 1. Speculative models of the interiors of the terrestrial planets. All are rocky, divided into a crust (not shown here), mantle, and metallic core. (*After R. T. Merrill and M. W. McElhinny, The Earth's Magnetic Field, Academic Press, 1983*)

Earth model concept. One significant aspect of the evolution of the Earth is the differentiation of relatively heavy and light material, as evidenced in a crude way by the moment of inertia. It is then not unreasonable to envisage an idealized concept of the Earth that is spherically stratified, with the heavier material in the innermost shells. Of course, the Earth is not spherically symmetric, and this is most obvious at the surface, but a spherically symmetric model does serve the purpose of identifying the essentially layered structure of the Earth, and it serves as a reference against which to measure lateral variations of structure. Evidence for this layered structure has come first of all from seismic studies. Thus several standard Earth models have been constructed in the past. One of the more comprehensive models was constructed in 1981 by A. Dziewonski and D. Anderson, the Preliminary Reference Earth Model (PREM). The principal regions of this model are listed in **Table 1**. *See* EARTH.

Seismic structure. The seismic structure of the Earth is studied by analyzing body waves, surface waves and free oscillations, and anisotropy.

Body waves. These are generated by earthquakes or large (nuclear) explosions and are recorded by many seismic stations throughout the world. There are about 2000 stations, but any particular wave is usually recorded at only a limited subset of these. Body waves are characterized as P (primary) and S (secondary) waves. Both wave types are supported by a solid, but P waves have a higher velocity and arrive first. It is routine practice that the stations report the arrival time of the first P wave to an earthquake information center, where these data are used to locate the earthquake and determine its origin time. By subtracting the origin time from the arrival times, the travel times of waves from the earthquake to the stations are determined, and these travel times can be used to determine the seismic velocity structure of the Earth's interior. Methods to determine the three-dimensional velocity structure are collectively known as tomography, a term derived from similar applications of travel time inversion used in medicine. However, the original purpose in seismology was to determine the velocity structure of a spherically stratified Earth model. The pioneering work of H. Jeffreys and K. Bullen in 1932–1942 resulted in travel time tables for many types of P and S waves, and formed the basis for the construction of P and S velocity models of the Earth. *See* EARTHQUAKE; SEISMOLOGY.

Surface waves and free oscillations. These are also generated by earthquakes. The velocity with which surface waves propagate is a function of the frequency of the waves, that is, the surface waves are dispersive. The dispersion depends on the seismic velocity structure of the surface layers. The free oscillations of the Earth resonate only at certain discrete frequencies (the eigenfrequencies), that is, the free oscillation spectrum would be a line spectrum (in practice the lines are somewhat smeared because of dissipation of the oscillations). The eigenfrequencies depend on the velocity structure of the whole Earth. Thus surface waves and free oscillations provide an alternative data set for constructing velocity

TABLE 1. Principal regions of a standard Earth model

	Layer	Approximate depth range,* mi	(km)
(1)	Ocean layer	0–1.8	(0–3)
(2)	Upper and lower crust	1.8–15	(3–24)
(3)	Lithosphere below the crust	15–50	(24–80)
(4)	Asthenosphere (low-velocity zone)	50–140	(80–220)
(5)	Upper mantle above phase or compositional changes near 240 mi (400 km)	140–240	(220–400)
(6)	Transition region between phase or compositional changes near 240 and 416 mi (400 and 670 km)	240–416	(400–670)
(7)	Lower mantle above core–mantle boundary layer	416–1703	(670–2741)
(8)	Core–mantle boundary layer (D″)	1703–1796	(2741–2891)
(9)	Outer core	1796–3200	(2891–5150)
(10)	Inner core	3200–3959	(5150–6371)

*Depth ranges are uncertain, especially in the crust and upper mantle.

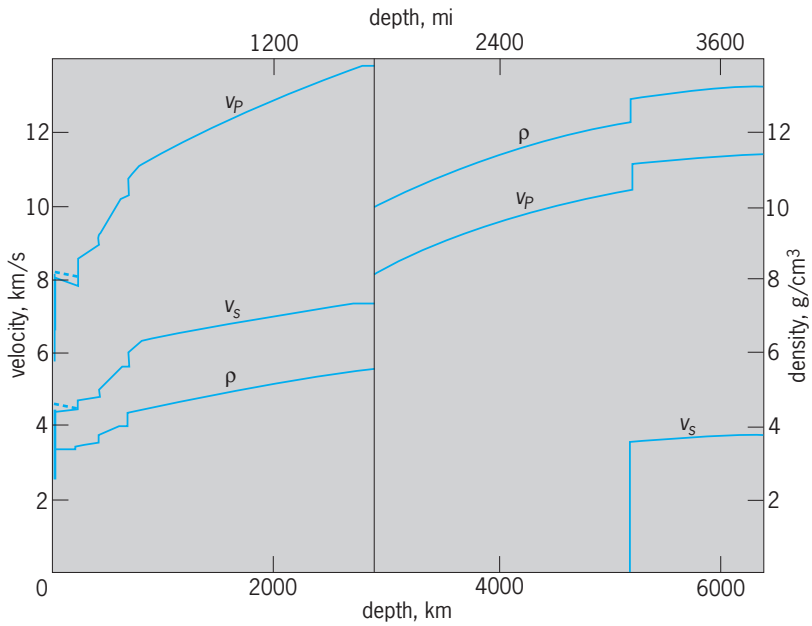

Fig. 2. Preliminary Reference Earth Model (PREM) for P- and S-wave velocities, v_p and v_s, and density ρ. In the part of the upper mantle where the model is anisotropic, velocities are given in both the vertical and horizontal direction. The broken lines represent the horizontal components of velocity. 1 km = 0.6 mi. 1 g/cm^3 = 0.5902 oz/in.^{3} (*After A. M. Dziewonski and D. L. Anderson, Phys. Earth Planet. Int., 25:297–356, 1981*)

models. After the big Chilean earthquake of 1960 provided the first observational evidence of free oscillations of the Earth, methods for inverting eigenfrequencies were developed. However, when seismic velocity models based on free oscillations became available, it was discovered that they were somewhat slower than velocity models based on body-wave travel times. It was initially thought that the travel time model was biased, since nearly all stations are located on continents. However, the effect was not sufficient to explain the difference, and the question was first resolved when the effect of anelastic damping of the waves was taken into account. PREM is based on both travel times and free-oscillation data. The model for a wave period of 1 s is shown in **Fig. 2**. *See* OSCILLATION.

The quality factor (Q) is used to quantify the damping of wave motion. The reciprocal of the quality factor (Q^{-1}) is defined as the fraction of elastic energy dissipated, per wavelength for traveling waves like body waves or surface waves, and per period for standing waves like free oscillations. Two consequences of the dissipation are a slight shift of the eigenfrequencies of standing waves and a slight change of the velocity of traveling waves. The changes depend on the period of the waves. In particular, a velocity model based on free oscillations with periods about 200 s is expected to be slower than a model based on body waves with periods about 1 s, which is precisely what the results show.

The quantity Q is also an important indicator of the physical state of the interior. The intrinsic Q at any point equals the Q of a body wave passing through. There is in fact a different Q for P and S waves; usually S waves dissipate faster than P waves, hence $Q_s < Q_p$. Q is a function of frequency such that dissipation effectively takes place within a certain frequency band, called the absorption band. The position of this band depends on temperature, pressure, and other special conditions. **Table 2** gives a rough estimate of Q for P waves at a 1-s period in different regions of the Earth. Q is strikingly low in the asthenosphere and near the top of the inner core. These are also solid regions of the Earth believed to be near the melting point. Viscosity similarly decreases with increasing temperature, and a connection between Q and viscosity has been contemplated. *See* ASTHENOSPHERE; VISCOSITY; WAVE MOTION.

Anisotropy. This refers to the directional dependence of the wave velocity. Anisotropy in the Earth's interior may be induced by flow of material containing noncubic crystals or elongated grains. Anisotropy has been observed in the lithosphere, especially under the oceans, and has been postulated for the asthenosphere. This anisotropy has been explained in terms of alignment of olivine crystals, which are thought to be the dominant material in these layers. In PREM the asthenosphere and lithosphere below the crust are anisotropic, although for practical reasons the anisotropy is of the simplest possible type such that there is no directional dependence within the horizontal plane; this is known as uniaxial anisotropy or transverse isotropy. The anisotropic regions of PREM (Fig. 2) are characterized by velocities in both the vertical and the horizontal directions. Anisotropy may occur in deeper regions of the Earth as well, but is difficult to resolve. However, to account for the differences in travel time between P waves passing through the inner core in the equatorial and polar directions, respectively, it has been proposed that the inner core is anisotropic, although the physical mechanism is not understood. *See* LITHOSPHERE; OLIVINE.

TABLE 2. Q for P waves at 1-s period

Region	Depth range, mi (km)	Q_p
Lithosphere	0–50 (3–80)	1500
Asthenosphere	50–140 (80–220)	200
Upper mantle	140–240 (220–670)	400
Lower mantle	416–1796 (670–2891)	1000
Outer core	1796–3200 (2891–5150)	10,000
Top of inner core	3200–3262 (5150–5250)	200
	3262–3387 (5250–5450)	300
	3387–3511 (5450–5650)	600
Central inner core	3511–3959 (5650–6371)	1000

Density structure. In an elastic body like the Earth, the seismic velocities v_p and v_s at any point depend on the local elastic moduli and the density. For most purposes, just two elastic moduli are needed: the incompressibility or bulk modulus, and the rigidity or shear modulus. The incompressibility measures the resistance to volume change under hydrostatic pressure, and the rigidity measures the resistance to change of shape under shear stress. Since there are just two measured quantities, v_p and v_s, the three parameters of the medium—density, bulk, and shear modulus—cannot be determined without additional assumptions and constraints. Two constraints are given by the Earth's mean density and moment of inertia. Another experimental constraint is given by laboratory measurements of the relation between v_p and density for relevant rock samples. Further assumptions are that in each of the principal layers of the Earth, where the seismic velocity profile is smooth, the composition of the material is chemically homogeneous, and phase changes are absent. The density distribution in each layer can then be computed by a method developed by L. Adams and E. Williamson, with a correction term added by F. Birch in case the temperature gradient exceeds the adiabatic value. Bullen, in the 1930s, used these principles with the assumption of an adiabatic temperature gradient in the different layers (so that there is no correction term) to construct a density structure consistent with the seismic velocity model. Later in the 1970s, this density model could be tested independently by using free-oscillation eigenfrequencies. It is necessary to distinguish between toroidal and spheroidal oscillations. Particle motion in the latter has a nonzero radial component, and the eigenfrequencies depend on the density structure. The density structure needed to satisfy the free-oscillation data is remarkably similar to Bullen's original density model. The density structure of PREM is shown together with the seismic velocities in Fig. 2. For the core and lower mantle, deviations from homogeneous and adiabatic conditions are insignificant, except for the region immediately above the core–mantle boundary.

The incompressibility, the rigidity, the gravity, and the hydrostatic pressure inside the Earth are obtained as by-products of the density calculations. The incompressibility increases steadily with depth throughout the mantle and core, and its changes across the core–mantle boundary are remarkably small. The rigidity also increases steadily throughout the mantle, but the fluid outer core has zero rigidity. In the inner core the rigidity is again nonzero but is probably less than in the lower mantle. This gives an abnormally low ratio of shear to bulk modulus for the inner core. The variation of gravity and pressure inside the Earth is illustrated in **Fig. 3**. *See* ELASTICITY.

Composition and state. Estimated abundances of rocky, refractory materials in meteorites and the Sun give a rough idea of what the Earth is composed of, although the relative proportions of metals and minerals in various portions of the Earth are not necessarily identical. To find out where and in which proportions the different materials reside in the Earth, high-pressure and high-temperature laboratory measurements are compared to the seismic and density structure. The Earth comprises a crust, a mantle, and a core, so that there is a compositional differentiation into at least three regions. Each of these regions is differentiated again, both vertically and, at least for the crust and upper part of the mantle, laterally. *See* ELEMENTS, GEOCHEMICAL DISTRIBUTION OF; HIGH-PRESSURE PHYSICS.

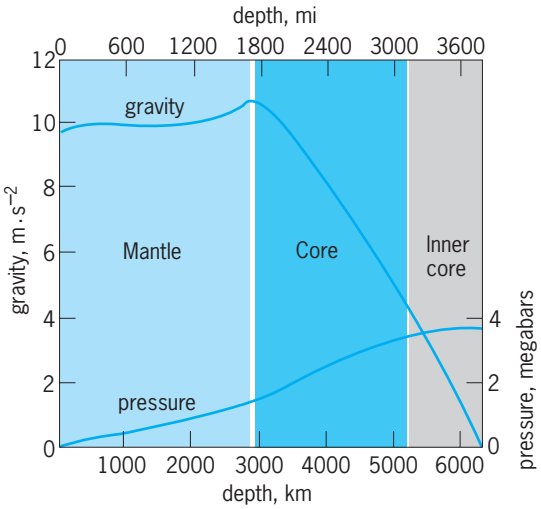

Fig. 3. Variations of gravity and pressure inside the Earth. 1 m = 3.3 ft. 1 megabar = 100 GPa. (*After P. Melchior, The Physics of the Earth's Core, Pergamon Press, 1986*)

Crust. The Earth's crust is on average about 12 mi (20 km) thick but is thinner under oceans and thicker under continents. Thickness under some high mountain ranges may be up to 36 mi (60 km; the so-called mountain roots). The boundary between the crust and the underlying mantle is called the Mohorovičić discontinuity or, more usually, the Moho. The continental and oceanic portions of the crust have different compositions. Continental crust can be divided into a lighter, granitic upper crust and a lower crust that is more basaltic, like the oceanic crust. Despite the striking heterogeneity of the crust, some generalizations can be made. Both the crust and the

TABLE 3. Chondritic and solar estimates of terrestrial composition versus upper-mantle composition (percent by mass)*

Component	Chondritic	Solar	Lherzolite†
Mantle			
MgO	37.7	32.7	42.2
SiO_2	52.5	45.0	44.2
Al_2O_3	3.8	3.2	2.1
CaO	3.0	3.4	1.9
FeO	3.3	15.7	8.3
MgO/SiO_2 ratio	0.718	0.727	0.955
Core			
Fe_2O	32.2	32.0	

*After D. L. Anderson, Where on Earth is the crust?, *Phys. Today*, 42:38–46, 1989.
†Composed mainly of olivine. Believed to be the major rock type in shallow upper mantle.

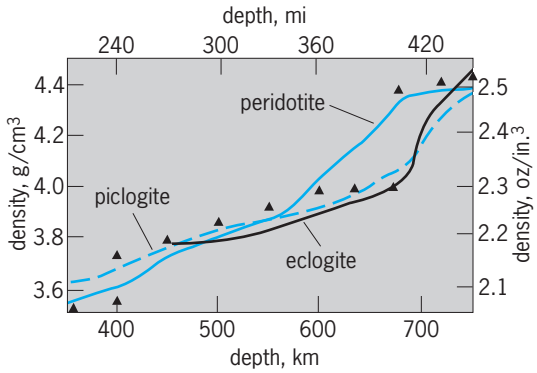

Fig. 4. Density versus depth for several mantle constituents along the 1400°C (2552°F) adiabat. Data points show the density variation of PREM. All three constituents shown here contain magnesium-rich silicates. (*After D. L. Anderson, Where on Earth is the crust?, Phys. Today, 42:38–46, 1989*)

mantle are rich in silicates, but the mantle has much more magnesium, and the crust has proportionally more of the relatively light elements such as calcium, aluminum, sodium, and potassium in the form of oxides. It has been noted that the relative abundance of these light elements in the stony meteorites leads to an expectation that much more of these elements should be on Earth than is contained in the crust. The missing part may be contained in former crust that has been subducted in the process of plate tectonics and is now on its way to being recycled throughout the mantle. Only oceanic crust is being subducted. *See* EARTH, CONVECTION IN; EARTH CRUST; METEORITE; MOHO (MOHOROVIĬĆ DISCONTINUITY); PLATE TECTONICS.

Mantle. This is the region between the crust and the core. It is common to make a distinction between the upper mantle and lower mantle, which are separated by the so-called 670-km discontinuity at a depth of 416 mi.

The upper mantle includes the lithosphere below the crust, the asthenosphere, and the region of phase transformations (the transition zone). It contains also the subducting slabs of oceanic lithosphere. The difference between continental and oceanic crust extends to at least the bottom of the lithosphere and probably to the bottom of the asthenosphere. It is believed that the stiff lithosphere and the soft asthenosphere differ mainly in mechanical strength, not in composition. The shallow part of the upper mantle is likely to be relatively depleted in the light elements typical of the crust. A comparison with estimates for stony meteorites (chondrites) and for the Sun (**Table 3**) serves to illustrate the depletion. On the other hand, the light elements do not seem to be underrepresented in estimates of the composition of the deeper parts of the upper mantle. It has been suggested that this region contains more eclogite, the high-pressure form of basalt, which is subducted with the oceanic lithosphere. It is the densest common upper-mantle mineral. Olivine is probably the most abundant upper-mantle mineral. At pressures prevailing near 240 mi (400 km) depth, it transforms to a denser structure (spinel), with a density increment of about 10%. Phase transformations of other common upper-mantle minerals are also expected at the depths spanned by the transition zone (**Fig. 4**). *See* BASALT; ECLOGITE; SPINEL.

The lower mantle, the mantle below 416 mi (670 km), appears to have a homogeneous composition down to the boundary layer above the core. Experimental results seem to indicate that mineralogy

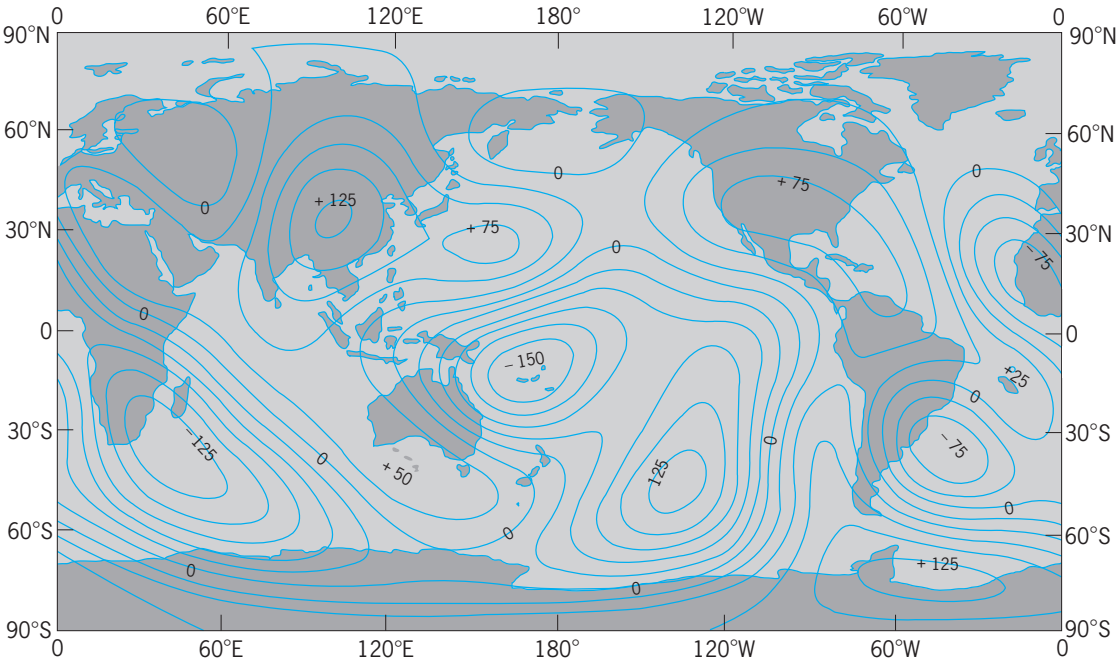
0
60°E
120°E
180°
120°W
60°W
0
90°N
60°N
30°N
0
30°S
60°S
90°S
+ 125
+ 75
0
+ 75
0
0
0

is dominated by ferromagnesium oxides (wüstite) and silicates (perovskite). The high-pressure perovskite phase of ferromagnesium silicate is thought to be the most abundant mineral of the Earth. *See* PEROVSKITE.

670-km discontinuity. This represents the transition from upper to lower mantle. There seems to be no compelling geochemical evidence against the view that the presumed dominant lower-mantle minerals are the high-pressure products of upper-mantle spinel, and there would be no chemical change across the boundary. On the other hand, such a change is not precluded by the evidence either, and an increased iron oxide content for the lower mantle is sometimes proposed. In any case, the transition represents a density increase of perhaps 10% and a viscosity increase by perhaps a factor of 10. It is not clear that subducted lithospheric slabs can penetrate this barrier. In fact, the seismic evidence suggests that some slabs do penetrate but others do not. One speculative model consistent with these results is that relatively young, thin oceanic plates are trapped at the boundary, but the upper, cool differentiated parts of thick plates may break through. It would imply a so-called leaky barrier to mantle flow.

Core–mantle boundary layer. This is often identified with the layer labeled D″ in Bullen's original seismic model. The seismic heterogeneity in this layer has a wide spectrum of scale lengths. **Figure 5** is a smoothed picture of large-scale lateral variations of P-wave velocity. Significant small-scale features are also inferred. From temperature considerations it is clear that the seismic anomalous layer is also a thermal boundary layer, and it has been proposed that this layer becomes unstable and forms convection plumes. The very big contrast in composition and material properties across the core–mantle boundary (**Table 4**) also creates favorable conditions for chemical heterogeneity above the boundary, although the observational evidence is unclear. The argument is that patches of material intrinsically denser than the bulk of the mantle will be trapped at the boundary, in much the same way as the relatively light continents are considered to be left over on the surface of the Earth. The analogy explains the wording "continents on the core–mantle boundary." However, unlike the surface of the Earth, the boundary itself is relatively smooth, with undulations of the order of a kilometer or less. This may be a consequence of low viscosity of the material at the boundary. Ironically, although it might seem that seismology would put strong constraints on topography of the boundary, the initial maps showed very large undulations. With the results of a more recent seismological study in agreement with a relatively smooth boundary but requiring variations in the boundary layer, the paradox appears to be removed. A speculative diagram synthesizing the various pieces of information is shown in **Fig. 6**.

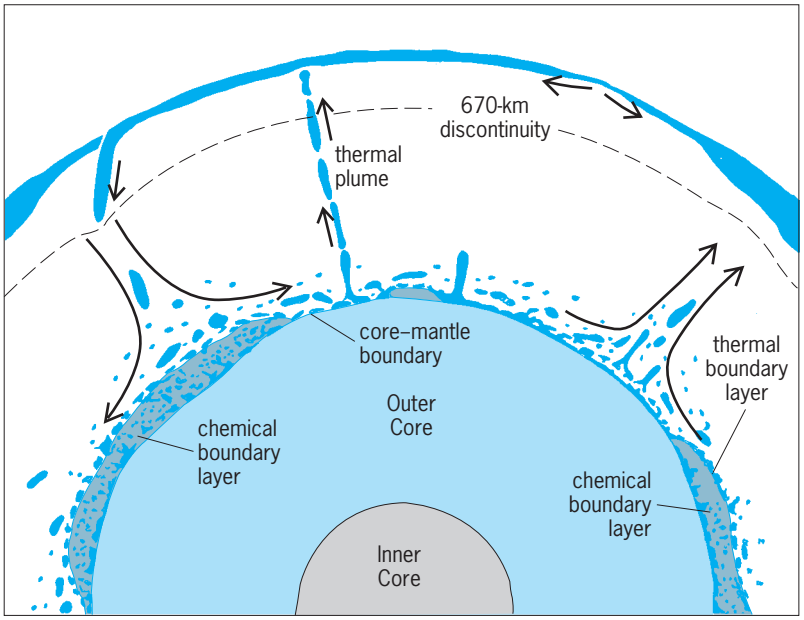

Fig. 6. Speculative, schematic model of the core–mantle boundary layer. A heterogeneous chemical boundary layer is embedded in a thermal boundary layer. Large-scale mantle circulation transports chemical heterogeneity to the base of the mantle as well as returning some by entrainment. Thermal plumes caused by boundary-layer instabilities ascend from the transition zone and disrupt the chemical boundary layer. Although the core–mantle boundary itself is relatively smooth, some topography is induced by both the large- and small-scale dynamic systems. 670 km = 400 mi. (*After T. Lay, Structure of the core–mantle transition zone: A chemical and thermal boundary layer, Eos, 70:49–59, 1989*)

Core. This is the central region of the Earth. Its seismic and density structure are well matched by a solid inner core of iron or nickel iron and a liquid outer core of iron mixed with about 10% lighter material. Candidate light elements are sulfur and oxygen, with some preference for the former. The outer core can be taken as homogeneous for all practical purposes. Seismic Q suggests that the inner core boundary is at the melting point. This is not the melting point of

TABLE 4. Contrast in material properties across the core–mantle transition zone*

Composition	Iron + 10% (O, S, Si, C)†	Magnesium-iron oxides and silicates
State	Molten alloy	Subsolidus
Temperature	3800–4700 K (6380–8000°F)	2600–3100 K (4680–5120°F)
Flow velocity	$0(10^{1})$ km/year	$0(10^{-4})$ km/year
Viscosity	$0(10^{-2}–10^{3})$ Pa · s	$10^{22}–10^{23}$ Pa · s
Density	9.90×10^{3} kg/m^3	5.57×10^{3} kg/m^3
Rigidity	0(0?) Pa	2.911×10^{11} Pa
Electrical conductivity	$10^{5}–10^{6}$ S/m	$10^{-2}–10^{2}$ S/m
Incompressibility	6.35×10^{11} Pa	6.85×10^{11} Pa

*After T. Lay, Structure of the core–mantle transition zone: A chemical and thermal boundary layer, *Eos*, 70:49–59, 1989.
†O = oxygen; S = sulfur; Si = silicon; C = carbon.

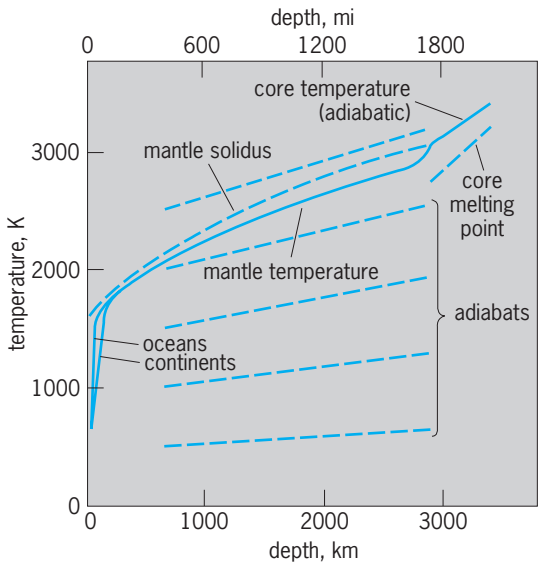

Fig. 7. Schematic temperature profile of the mantle. Subsequent estimates of core temperature are higher than suggested in the figure; the temperature jump across the thermal boundary layer at the base of the mantle will then be correspondingly higher. °F = (K × 1.8) − 459.67. (*After F. Stacey, Physics of the Earth, 2d ed., Wiley, 1977*)

pure iron; the presence of impurities in the liquid depresses the melting point considerably.

Core–mantle coupling. This is believed to be responsible for decade-long variations in the length of the day of up to about 5 milliseconds. The idea is that fluid flow in the outermost regions of the core is coupled to the mantle, so that the rotation rate of the mantle can fluctuate in response to time variations of the flow in the core. This view is supported by an observed correlation between the decade variations in the length of the day and certain variations of the geomagnetic field (the latter being associated with core flow). The coupling may be electromagnetic or topographic. It has been suggested that topographic coupling can be quite effective, provided that there is indeed topography on the core–mantle boundary (of the order of a few hundred meters).

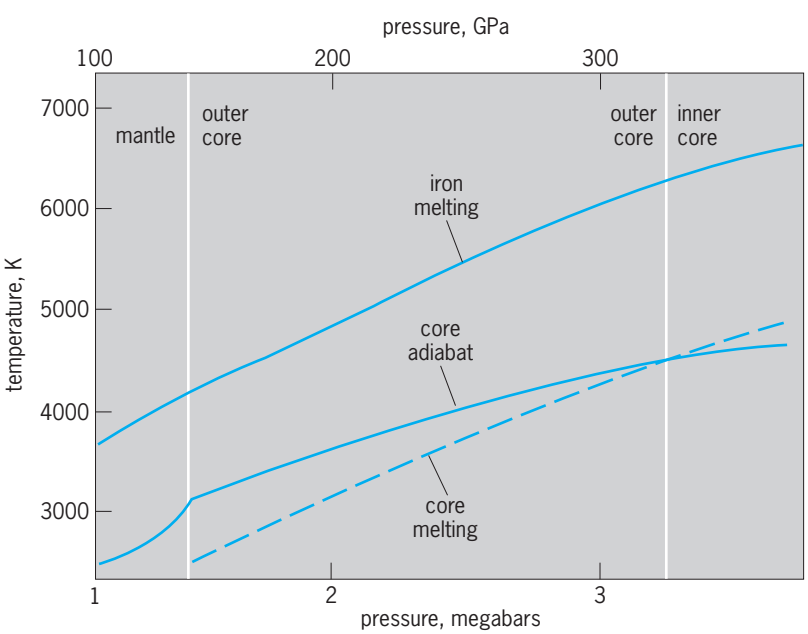

Fig. 8. Schematic temperature profile of the core. The melting-point curve for the core is substantially below the melting point of pure iron. All temperatures may have to be raised to correspond with further estimates of the melting point of iron. °F = (K × 1.8) − 459.67. (*After D. J. Stevenson, Models of the Earth's core, Science, 214:611–619, 1981*)

Temperature. Mantle temperature estimates are constrained by the following: (1) A melting-point curve for the mantle is obtained from laboratory measurements of the melting point of presumed mantle minerals under pressure. The actual temperature should not be too far below the melting point in order that material can flow. (2) Convection is supposed to be the dominant mode of heat transport below the lithosphere. In a slowly convecting system the temperature gradient is probably slightly superadiabatic, which is a gradient larger than that prevailing under adiabatic conditions. (3) The temperature gradients in the crust can be measured. Since conduction is the mode of heat transport through the lithosphere, temperatures can be estimated by extrapolating the surface temperature downward, until the melting point curve for the mantle is approached. This point is often taken as the bottom of the lithosphere, and temperatures in the mantle are estimated by extrapolating downward from this point; a schematic temperature profile is shown in **Fig. 7**. The thermal gradients on continents and under oceans differ substantially, and continental lithosphere is thicker than oceanic. At the base of the mantle a thermal boundary layer with a strongly superadiabatic gradient exists. Estimates of core temperature suggest that the temperature jump across this layer is larger than is indicated in Fig. 7. The figure does not suggest a thermal boundary layer between the upper and lower mantle (near 416 mi or 670 km depth); but this is, in fact, an unresolved problem.

Core temperature presumably follows a nearly adiabatic gradient, at least in the convecting outer core. A calibration point for the temperature profile is provided by laboratory measurements of the melting point of iron under high pressure and by assuming the inner core boundary is at the melting point. This point is about 1000 K (1800°F) below the melting point of pure iron under the same pressure, since the liquid iron (in the outer core) is not pure but has about 10% light material mixed in. A temperature profile for the core is shown in **Fig. 8**. More recent estimates of the melting point of iron under high pressure are higher than suggested in Fig. 8, and so the actual temperature profile will be correspondingly higher.

Heat sources are the radioactive elements, especially in the continental crust and the upper mantle, and the slowly cooling Earth itself. The total surface heat flux is about 3×10^{13} W, and cooling may contribute about a quarter of this value. A flux of about 3×10^{12} W through the core–mantle boundary may be due to cooling of the core and the consequent crystallization of the inner core, although the possibility of heating due to decay of radioactive potassium (^{40}K) cannot be neglected. The crystallization

process itself is a good example of ongoing differentiation of the Earth. *See* EARTH, HEAT FLOW IN; RADIOACTIVITY.

Geomagnetic dynamo. The Earth's magnetic field has been studied for centuries, but it is only rather recently that an understanding of the generation of the field has developed. In 1946 W. Elsasser proposed a model of the magnetic field generated by inductive motions in a liquid core. Evidence for a liquid core came 20 years earlier, when H. Jeffreys noted the absence of seismic S waves passing through the core. A geomagnetic dynamo theory, based on the work of W. Elsasser and later on that of E. Bullard, is probably the only theory capable of explaining the important features of the magnetic field such as reversals, secular variation, and the near alignment of the dipole and rotation axes. However, the mathematical problems are formidable, and as yet no satisfactory dynamo model exists that reproduces secular variation and reversals. An intriguing hypothesis is that these features are governed by conditions in the core–mantle boundary layer. *See* GEOMAGNETISM.

Reversals. This term refers to polarity transitions of the dipole part of the geomagnetic field. The field on the Earth's surface is about 90% dipole. Paleomagnetic results have demonstrated that the polarity of the field has reversed many times in the past, but the frequency of reversals appears to be nonstationary. One hundred million years ago the magnetic field was in the middle of a 30-million-year period of normal polarity (the Cretaceous quiet interval). This is in sharp contrast to the behavior both 150 million years ago and recently, when the mean polarity interval is of the order of several hundred thousand years. Although the connection is unclear, such long-term changes are currently believed to be related to variations in the thickness of the core–mantle boundary layer; thermal boundary-layer studies indicate that such variations can occur. It is speculated that when the layer is thick, the cooling rate and power supply are small and the field is in a quiescent state, with reversals infrequent or even absent. Conversely, when the layer is thin, the cooling rate and power supply are large and the field is in an agitated state with frequent reversals, which is the state occurring at present. *See* PALEOMAGNETISM.

Secular variation. Included in the secular variation of the magnetic field, among other factors, is the westward drift of the field and the decrease of the dipole field strength during the past few hundred years; if the decrease continues at the present rate, a reversal will occur in roughly a thousand years. The secular variation of the magnetic field at the Earth's surface has been reasonably documented during the past few hundred years. J. Bloxham and D. Gubbins have made an estimate of the field at the core–mantle boundary during this time period that implies inverting the observed data of the surface magnetic field. In **Fig. 9** the field is shown for the years 1715 and 1980. It is striking that some prominent features seem to be stationary, and the speculation is that these features are controlled by lateral temperature variations in the core–mantle boundary layer.

Energy source. The source of the energy for the geodynamo can be specified on the basis of an energy balance argument. It is estimated that the power needed for the dynamo process itself is about 4×10^{11} W. It is also estimated that in the process a heat flux equaling about 2.7×10^{12} W flows through the core–mantle boundary by conduction. An obvious heat source is radioactive decay; this requires a certain amount of radioactive potassium in the outer core. The problem with this process is the very low efficiency; only about 5.5% of the heat is becoming available to power the dynamo.

Another mechanism not suffering from an efficiency problem has been proposed. When the core cools, solid iron (or nickel-iron) freezes out on the

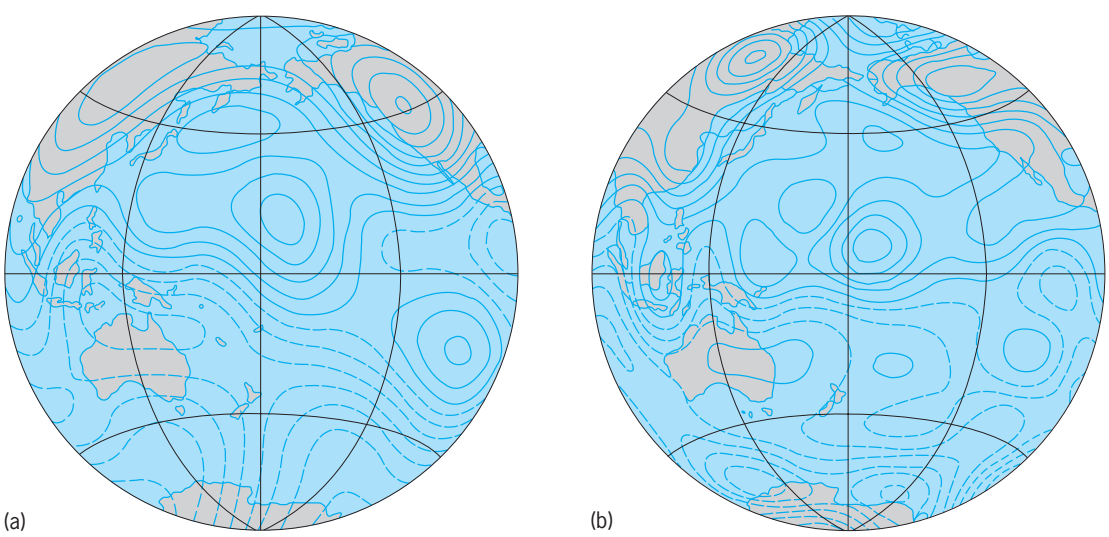

Fig. 9. Contour plots of the radial component of the magnetic field at the core–mantle boundary (*a*) for the year 1715 and (*b*) for 1980. The contour interval is 100 microteslas; solid-line contours represent flux into the core, and broken-line contours represent flux out of the core. (*After J. Bloxham and D. Gubbins, The secular variation of the Earth's core, Nature, 317:777–781, 1985*)

boundary of the inner core, leaving a lighter constituent behind in the outer core. The rising of the lighter material through the liquid core leads to so-called compositional convection, and its efficiency is very high. The additional latent heat of crystallization is less important in that respect.

Assuming the inner core has grown to its present size in about 3 billion years, the available power may exceed 4×10^{11} W if the density jump across the inner core boundary exceeds 0.302 oz/in.^{3} (0.5 g/cm^3). The density difference is thus a crucial quantity for the geodynamo. In principle, seismology can provide estimates of the density. In practice, the density jump across the inner core boundary is not well resolved. Durk J. Doornbos

Bibliography. D. L. Anderson, Where on Earth is the crust?, *Phys. Today*, 42:38–46, 1989; K. E. Bullen and B. A. Bott, *An Introduction to the Theory of Seismology*, 4th ed., 1985; T. Lay, Structure of the core–mantle transition zone, *Eos*, 70:49–59, 1989; P. Melchior, *The Physics of the Earth's Core*, 1986; J.-P. Poirier, *Introduction to the Physics of the Earth's Interior*, 2d ed., 2000.

Earth resource patterns

The physical character and distribution of natural resources at the face of the Earth. No section of the Earth is exactly like any other in its resource endowment. Combinations of land and ocean, latitudinal differences in insolation, variations in receipt of precipitation, patterns of geology, and deformation of the Earth's crust all converge to create different resources in various regions of the world. Nevertheless, there is some repetition in these natural features and forces of nature across the Earth. These similarities that exist from place to place distinguish regional patterns in the availability of resources on a global scale.

Delineation of the Earth's resource patterns begins with differentiation between continental and marine resources. Although the resources of the oceans and seas have been used by people since earliest times, the more than 6,000,000,000 on the Earth today are primarily dependent upon the resources of the land for their existence.

Five principal resources associated with the land are soils, forests, grasslands, fresh-water resources, and minerals. Although other resources such as native animal life may be of local importance, and although the very concept of "resources" has been extended to include such complexes as recreation resources, these five land resources remain of fundamental importance for the material support of human life.

Natural resources include any substance or material, produced by natural forces within the Earth system, that is used or valued by humans. Some of these resources more quickly regenerate, and are naturally

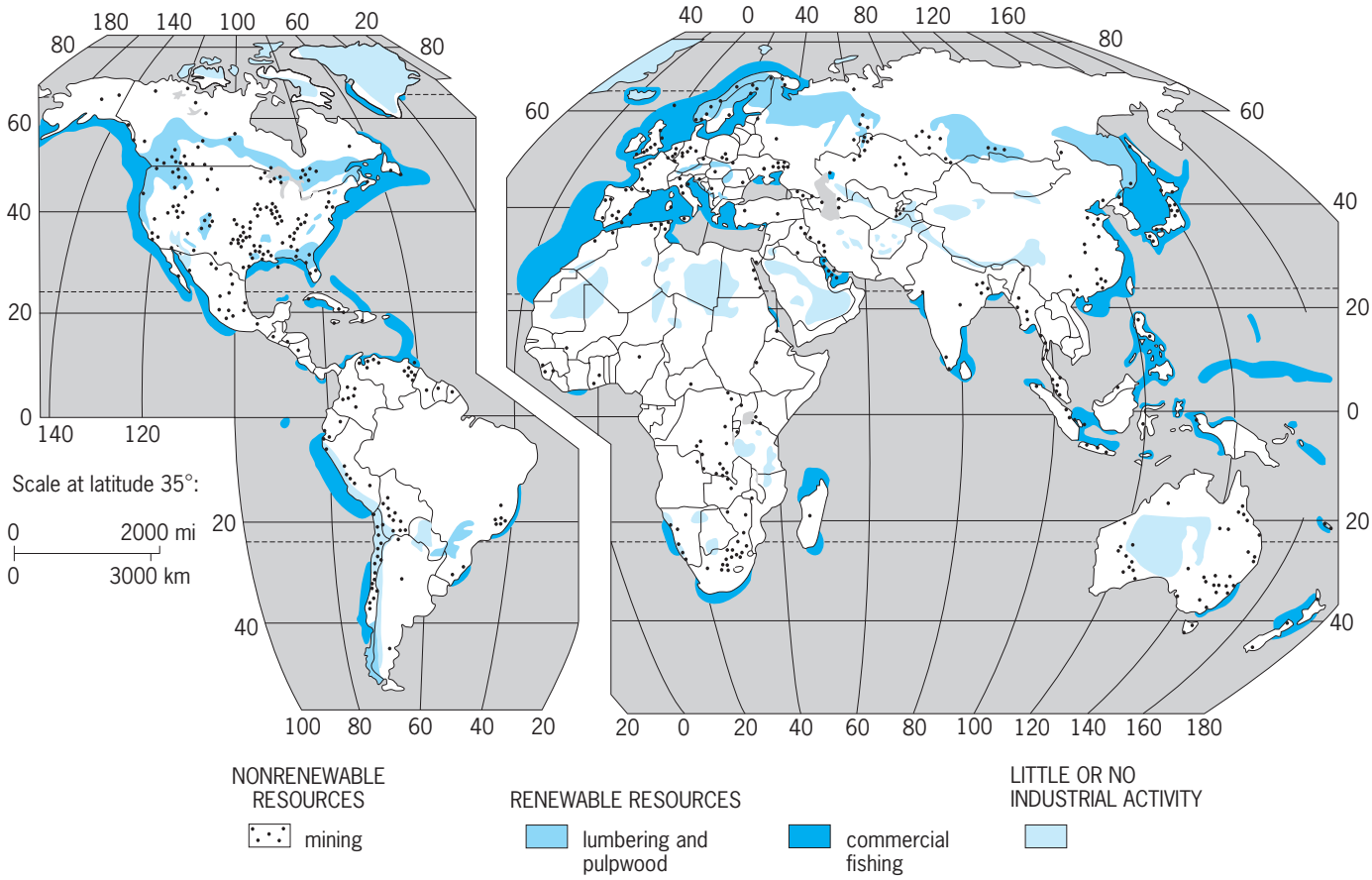

Fig. 1. Regions of selected primary and secondary industries.

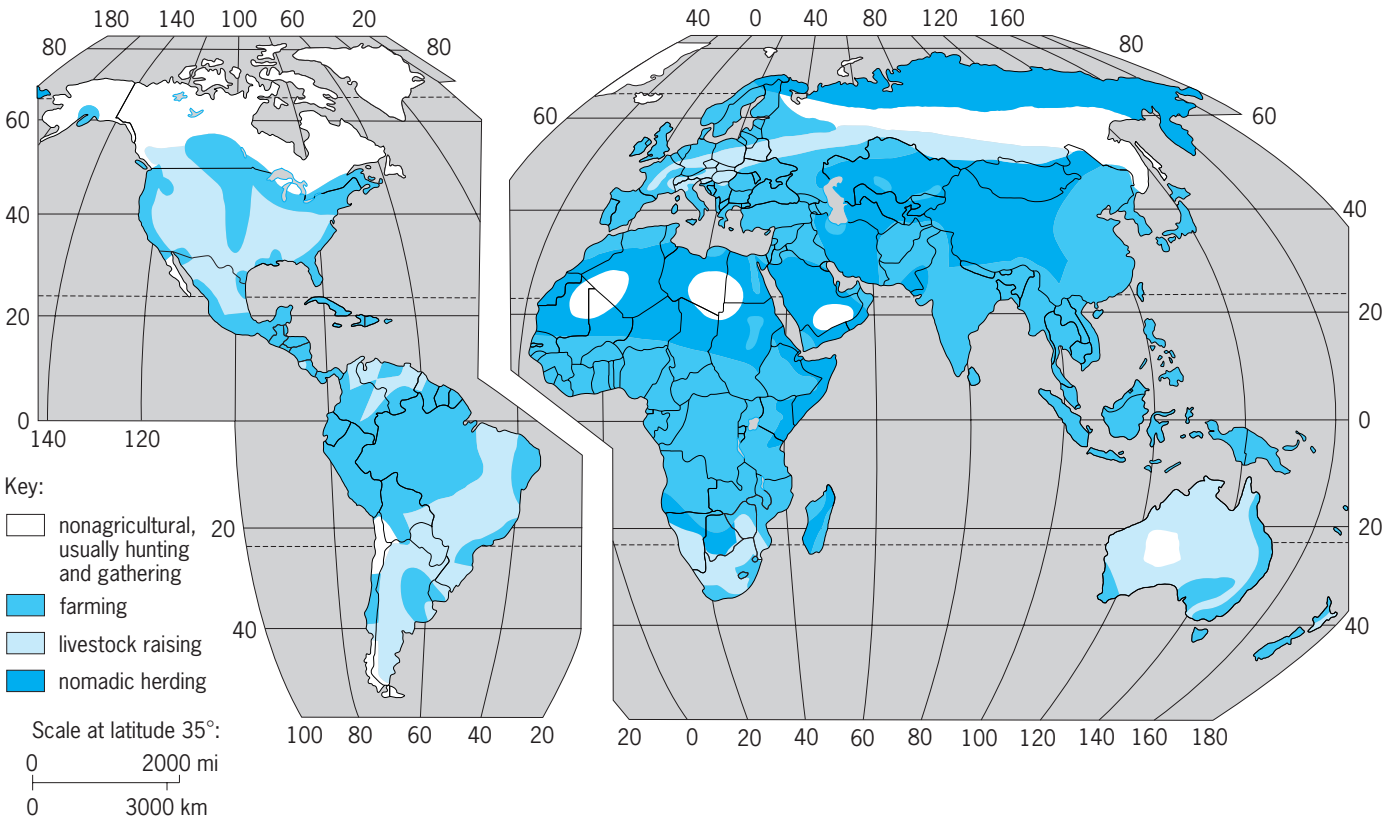

Fig. 2. Major agricultural regions.

replaced after being extracted. These resources—forests, marine life, soil, fresh water, agricultural crops—are termed renewable resources (**Figs. 1** and **2**) and may be used indefinitely if rate of use is balanced with rate of regrowth or replenishment. Other resources cannot be regenerated so quickly and are steadily used up by humans. These nonrenewable resources include oil, natural gas, coal, and metals.

Two types of natural forces have converged to create the natural materials that are used and valued by humans. One set of forces consists of the basic climatic controls, including latitude, distribution of land and water, the wind and pressure system of the rotating Earth, the major landforms of the continents, and the elevation of the land surface above sea level. A second, independent set of forces consists of the tectonic and rock-forming processes which have operated over the Earth. The climatic controls account for variations in regional climates over the continents, and these climates in turn help to shape the the unique character of forests, grasslands, and fresh-water resources, as well as some of the fundamental attributes of soils and the agriculture they support. The second set of forces may be regarded as even more fundamental since the movements of the Earth's plates and their associated continents in conjunction with the Earth's plate tectonics account not only for the position on the Earth and hence the latitudinal location of each continent, but also for the global distribution of land and water and continental landforms, all with consequences for regional climatic patterns. Moreover, rock composition and surface configuration also influence the development of forests, grasslands, water resources, and soils which alter the patterns within the climatic regions, and surface and subsurface geology are fundamental to an understanding of the global patterns of minerals on and beneath the Earth's surface.

Climatic types. Because many resources reflect the climatic regime under which they form, it is important to understand distinctions between climates, where these various climates are found, and how climates influence the natural development of resources. World climates may be broadly classified in four types, which vary according to combinations of temperature and precipitation, and vary seasonally in both phenomena due to latitude and proximity to oceans or mountains (**Fig. 3**). Distinction is made between the so-called humid climates and the water-deficient climates, with subtypes as follows:

Humid microthermal
1. Polar and ice cap
2. Tundra
3. Taiga
4. Puna

Humid mesothermal
5. Upper midlatitude
6. Humid subtropical

Humid macrothermal
7. Wet-and-dry
8. Rainforest

Water-deficient
9. Desert
10. Semiarid
11. Mediterranean

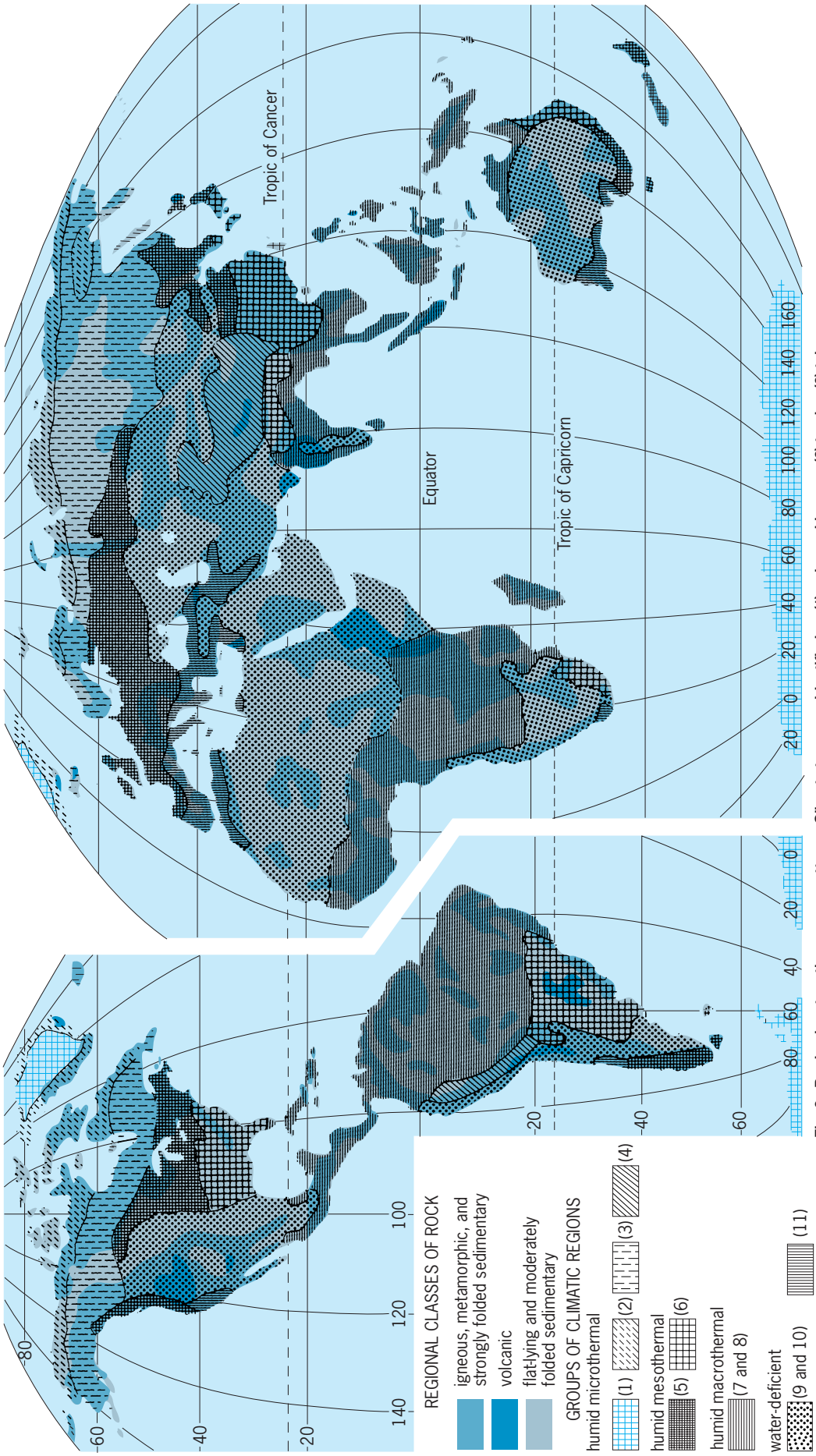

Fig. 3. Predominant earth resource patterns. Climate types are identified as (1) polar and ice cap, (2) tundra, (3) taiga, (4) puna, (5) upper midlatitude, (6) humid subtropical, (7 and 8) wet-and-dry and rainforest, (9 and 10) desert and semiarid, (11) mediterranean. The map is a flat polar quartic equal-area projection.

Humid microthermal regions. These areas of predominantly low temperature are so unfavorable to soil formation and use in agriculture that under present techniques their population-carrying capacity is low even in those regions with the warmest summers.

1. In polar and ice cap areas where soils and vegetative growth are essentially absent, available resources necessarily are dominantly marine and land animal life, on which the sparse native human settlement is almost wholly dependent. Despite the enormous size of the Antarctic, settlements are exclusively in the Arctic, except for special government-supported Antarctic stations.

2. Tundra, except for minor alpine locations, is entirely within the Northern Hemisphere. The principal renewable resources are lichens and the native animal life, such as reindeer and caribous, which can use these as food. Parts of the tundra may be considered a grazing land, as managed herds of reindeer are pastured nomadically. The natural resource significance of tundra lands for the larger world may be greatly enhanced locally where minerals are being extracted, such as the oil field adjacent to Prudhoe Bay on Alaska's North Slope. *See* TUNDRA.

3. Millions of acres of boreal coniferous forest, or taiga, located in a broad curving zone from Scandinavia across the northern, European part of Russia, and Siberia, and, east of the Bering Sea, across much of Alaska and northern Canada, has given its name to this climatic belt. Varieties of spruce, fir, and larch, which are of particular significance to the pulping industries, constitute the most valuable known renewable resource of the humid microthermal regions. The zone of taiga climate is of scant importance for agriculture; soils are predominantly thin, stony, and infertile, and development for agriculture is further discouraged by the possible incidence of frost in every month of the year and by the short growing season of 80 days or less. As in the case of the areas of tundra climate, the resource significance of particular localities within the taiga zone is locally enhanced by deposits of minerals; globally significant deposits of gold, silver, iron, lead, zinc, uranium, copper, and other metals are found in the taiga lands of Canada and Russia. Fresh-water resources are extensive, and some hydroelectric power is developed. *See* TAIGA.

4. The puna type of climate is found at much lower latitudes but at elevations generally 10,000 ft (3048 m) or more above sea level, and is characteristically cold. Certain plateaus, particularly the high plateau of Tibet and the intermontane Andean plateau (Altiplano) of southeastern Peru and western Bolivia, belong to this group. Low temperatures preclude tree growth, and the principal renewable resource is low-productivity grazing land. Thin, stony soils are limited in production to hardy small grains and root crops. The resource significance of the South American puna is greatly enhanced by metal deposits, particularly those of copper, tin, lead, and zinc. *See* PUNA.

Humid mesothermal regions. Considered in the light of present-day technology, the heart of the world's renewable natural resource base is in the humid mesothermal regions with their generally adequate precipitation and intermediate temperatures. These middle latitude regions of Earth contain a large share of the world's most productive soils which support both crop and pasture lands. Some productive coniferous and broad-leaf forests occur in these climatic regions. Concentrated surface and subsurface fresh-water supplies are relatively abundant, but owing to the high population densities and major urban and industrial complexes in these regions, the water resources are often intensively developed to the point where the quantity and quality of resources is diminished.

5. Upper midlatitude climate contains most of the lands adapted to the raising of wheat, barley, rye, and oats. In addition, certain areas within this climatic type, and particularly the North American corn belt and the Middle Danube Basin and the North Italian Plain, also produce maize. Further crops, including soybeans, are also important, and extensive forage cropping supports dairying and meat animal raising in east-central North America, western and central Europe, and the nonarid part of southern European Russia. The lands which support the agricultural types of this climatic region possess the most extensive areas of superior soils on Earth. They are relatively deep and are often moderately to highly endowed with plant nutrients and humus. Some of the best soils have developed from glacial drift; others have benefited from the deposition of fine-grained windblown loess deposits during the glacial period. Extensive preagricultural grasslands also contributed high humus content to some of the soils. Finally, the extensive plains under which these soils are deployed, and the generally favorable growing season conditions of temperature and rainfall, also underlie the present high productivity of the agriculture supported by the soils in this climatic type. It is within the upper midlatitude climate that the region of greatest surplus food production (the North American interior) exists.

6. Humid subtropical areas have an ample water supply and relatively long growing season (200 days or longer), making the best of these lands potentially very productive for crops or for forest products. Soils other than in floodplains tend to be less fertile than those of upper midlatitudes, however, owing to the effects of relatively high rainfall and temperatures on the removal of plant nutrients, and generally require fertilization for sustained cultivation over long periods. Where cropping is on alluvial soils of floodplains large and small, as in central and southern China and in southern Japan, and in the Ganges Plain of India, soils are more easily maintained at high levels of productivity. High rural population densities generally preclude the generation of large exportable food surpluses from regions with the best soils. Although the favorable combination of temperature and moisture can result in high timber growth, as on the best-managed tracts of forest in the American South, the potential is not realized in other areas of the Earth having this type of climate, owing to

previous large-scale deforestation. Water resources are generally abundant; in part of southern and eastern Asia, fresh-water resources underlie the most extensive humid land irrigation in the world—namely, for paddy rice.

Humid macrothermal regions. These low latitude, predominantly winterless regions of warm to hot temperatures are divided according to the regime of rainfall:

7. Wet-and-dry, with a pronounced dry season, characterizes the tropical savanna regions of the Earth. These tall grasslands, mixed with scattered tropical trees, offer some of the most extensive grazing lands in Africa and northern Australia. Some of the savanna region, the original home of the big game herd animals in the African plains, is now used for commercial livestock and mixed farming.

8. Rainforest, with year-round growing seasons but over extensive areas their soils are lateritic with a high iron content, and harden irreparably under use for cropping. Problems induced by fungal growth, bacterial disease, and insect abundance also handicap the use of the soil resources. Within the great alluvial valleys, flooding may also be disruptive. Owing to these adverse factors, shifting cultivation is common, and much land is not in production at any particular time. Within the rainforest regions, extensive and rapid-growing forests occur, including the largest such forest area on the Earth in the Amazon Basin, but are mostly unexploited commercially in the twentieth-century economy. Extensive clearing of the Amazon forest for agricultural development appears likely during the last two decades of the twentieth century. Sites of enormous potential hydroelectric generation are largely undeveloped. Extensive iron and aluminum deposits have become concentrated within the rainforest soils after centuries of heavy rainfall has leached away most other soil substances. Many of these deposits have been exploited by domestic and multinational mining corporations. *See* RAINFOREST.

Water-deficient regions. Receipts of moisture are scant or lacking during much of the year. Except where exotic water supplies are available for irrigation, soils inherently are less productive than those of any humid region with comparable land surface. There are three major subdivisions.

9. Deserts differ strikingly in their form and in temperature conditions, but everywhere present meager resources for agriculture. Where water is available, desert oases blossom, but vast areas contain only scrub growth, ephemerals, or virtually no vegetation. Livestock-carrying capacity of desert scrub is meager, though all but the most extreme deserts have traditionally been occupied by groups of nomadic pastoralists who typically herd goats, sheep, or camels. Sparsity of vegetation has made mineral prospecting and exploration somewhat easier than in vegetation-covered humid areas, and in this century desert occupancy often started with mineral discoveries. *See* DESERT.

10. The semiarid regions are basically grasslands, which have grazing as their characteristic resource use. Because of cyclic rainfall variability, people have converted the inherently fertile soils, as in China and the United States, to cereal growing during periods of higher rainfall. Rainfall fluctuations, however, make the soils unstable under cultivation. For this reason, these regions have suffered consistently from unsuited cultivation and overgrazing. Semiarid lands are responsive to and most productive under irrigation, but neither surface-water nor ground-water resources are adequate to irrigate more than a fraction of the area. *See* GRASSLAND ECOSYSTEM.

11. Regions of mediterranean climate, because of their winter rainfall, generally are classed as humid lands. However, the greater part of the gowing season is water-deficient, and the most productive agricultural lands depend on irrigation. Soils are the major resource since water deficiency is pronounced enough to discourage forest productivity. Major mineral deposits may complement agricultural lands in a few areas. *See* CLIMATOLOGY.

Rock composition and surface configuration. Imposed on the basic resource pattern induced by climatic differences are variations in rock composition and surface configuration which cause intraregional differences within the patterns already described. Although not exactly the same in their effects, the variations caused by these two geographical elements are often concomitant and may be treated together as follows:

1. Rock composition and structure
   a. Flat-lying and moderately folded sedimentary rocks
   b. Igneous, metamorphic, strongly folded sedimentary rocks
   c. Volcanic rocks
2. Surface configuration
   a. Floodplains and other flat or gently sloped surfaces
   b. Mountains and maturely dissected hill lands, plateau faces, of faces of cuestas

These elements of crustal variation produce the six following geographical differences in resource endowment:

First, all major agricultural lands are on flat-lying or moderately folded sediments and have gentle slopes, well exemplified in such alluvial valleys as the Mississippi, Nile, Huang, and Ganges-Brahmaputra, and such other outstanding agricultural areas as the North American corn belt or the Paris Basin.

Second, productive secondary agricultural lands, particularly in the tropics, are located on volcanic areas where soils have been formed through weathering or wind action. Examples include the Deccan Plateau in west-central peninsular India, and volcanic soils fringing much of the central mountain backbone of Java.

Third, agricultural lands are extremely limited on igneous rock areas, no matter what the surface configuration, in regions north of 40°N latitude, as illustrated in the Laurentian Shield area of Canada and

the Fenno-Scandian Shield in northern Europe. In the humid tropical and subtropical climates, however, where weathering has proceeded long enough to produce a substantial soil mantle as on the Piedmont of the southeastern United States, underlying igneous rocks do not have the same negative effect on agricultural development.

Fourth, forest lands are not limited in their extent (although limited in productivity) by either crustal rock composition or surface configuration.

Fifth, mountains are important catchment areas and sources of fresh water and the services which may be derived from water. Most hydroelectric generation, or generation potential, is associated with mountains. In arid and semiarid regions, mountains are sources of water for irrigation, domestic and industrial supply, wood products, and warm-season grazing lands.

Sixth, mineral resources have definite patterns which are associated with rock structure and composition. Major deposits of coal and lignite, petroleum, and natural gas are, with few exceptions, found in flat-lying or gently folded and faulted sedimentary rocks, as in Texan oil fields and the coal fields of the Allegheny Plateau. Sedimentary nonmetallics (the phosphates, potash, sulfur, nitrates, and limestone) as well as bauxite and uranium (carnotite) also are associated with sedimentary rocks.

Associated with the igneous and metamorphic rock areas are most metals, for example, iron (usually), lead, copper, tin, the ferroalloys, gold, and silver. Most gems and some nonmetallic minerals (mica and asbestos) are found in the same associations. Uranium (pitchblende) occurs in these rocks.

Whereas these associations are well recognized, the mineral deposits themselves have a highly erratic geographical occurrence owing to the great variety of processes involving both mineral enrichment and dispersal which have occurred. The broad global pattern of rock classes with which minerals are related is delineated on the map in the illustration.

Employed and potential resources. Resources have meaning insofar as they are placed in use or are available for future exploitation. Distinction must be made between the employed and the potential resources. In practice this distinction is complex, but here only the simple geographical distinction will be noted. Employed resources are those which are significant to the present support of humans, at least locally. In general, the denser the population and the more advanced the technical arts of an area, the greater the need for production from resources, and employed resources become more nearly synonymous with all known resources. Thus, the recognized resources of the European peninsula and of the northeastern United States are mainly employed resources. On the other hand, the natural resources of the Amazon Basin, or much of Africa, Siberia are still largely potential resources.

Marine resources. Although the physical and biotic geography of the oceans is much less fully explored than that of the continents, enough is known to indicate that both living and mineral resources extend far beyond those presently exploited. The employed resources are rather sharply localized. The principal exploitation of marine animals and vegetation occurs: over the continental shelves; in the vicinity of the mixing of warm and cold currents; near large upwellings which occur particularly off the west coast of continents in lower middle latitudes; and adjacent to densely populated countries. Thus, the North Atlantic near Europe and from New England to Newfoundland contains heavily exploited fishing grounds, as do the seas near Japan, Korea, and southern California, and also waters of the Gulf of Mexico.

Minerals are derived from three separate types of marine sources: from sedimentary deposits underlying the continental shelves; from inshore deposits on the surface of the continental shelves; and from seawater. By far the most valuable of the mineral resources exploited from marine environments is petroleum; there has been an expansion of exploration, drilling, and pumping from beneath continental shelf waters, as off the Gulf coast of Louisiana, in the middle of the North Sea between Scotland and Norway, and offshore in the Persian or Arabian Gulf. Offshore placer deposits on the surface of the continental shelves yield gold, platinum, and tin. Common salt, magnesium, and bromine are derived directly from seawater. *See* OIL AND GAS, OFFSHORE.

Potential marine resources include the population of life-forms now exploited in some parts of the world, but not in others; animal and vegetative species now unused; fresh water from desalted seawater; and minerals so far unexploited which are in solution, are precipitated to the ocean bottom, or lie within rock below the surface of the continental shelf. One of the interesting speculative resources appears in the large quantities of so-called manganese nodules that cover some sections of sea bottoms at intermediate depths.

Resources of the continents. The resource pattern of the Earth may be summarized in a brief description of that for each continent and its neighboring waters.

Eurasian continent. As the largest landmass in the world, the Eurasian continent has the largest area of agricultural land in use, a very extensive total forest land area in which the softwood coniferous forest belt from Scandinavia to eastern Siberia predominates, and a wide variety of mineral resources. Great differences mark the major sections of the continent. The most productive agricultural areas are generally near the edge of the landmass, in western and central Europe and extending eastward into much of the central and southern sections of Russia in Europe; in the Indian subcontinent; and in mainland east Asia from the Red River valley in Vietnam northward to the Great Wall of China. The aggregate mineral endowment is outstanding and includes the largest known aggregate iron ore reserves in the world within Russia, located particularly in districts adjacent to the Urals; very substantial coal deposits, including the Ruhr field and the extensive coal beds in northern China and Manchuria; and what increasingly appears to be one of the two great concentrations of petroleum fields on Earth, namely, the Persian Gulf fields

shared by a number of separate states in the Middle East. The southeastern and eastern borders of the Eurasian heartland, moreover, have some of the great, but still undeveloped, hydroelectric generation sites of the world. Off the coasts of western Europe and Japan are the two most productively employed fisheries of the world. *See* ASIA; EUROPE.

Africa and Australia. Much of the entire area of Africa north of approximately 12°N must be classed as desert or semiarid, with few exotic water sources other than the Nile. Much of the remainder of the continent has wet-and-dry or rainforest climates, with the former predominating greatly in area. Seasonal drought, soil infertility, and widespread problems of laterization handicap agricultural exploitation. The east African highlands from Ethiopia southward, the high veld in South Africa, and the loftier sections of Zimbabwe, Rhodesia, and the Nile Valley and Delta are noteworthy exceptions. Except along the Nile there are still potential agricultural land resources, but they are comparatively minor. Associated with the extensive areas of igneous and metamorphic rocks which underlie much of the continent and particularly its southern half are outstanding deposits of metalliferous minerals, such as the copper ores astride the Zaire-Zambia boundary, the chrome-bearing ores of Zimbabwe, Rhodesia, and the gold deposits of the Witwatersrand in South Africa. Other mineral resources include diamonds, uranium, and, in Nigeria and the far north of the continent (Libya, Algeria), petroleum and natural gas. The water resources of mid-Africa include the largest potential hydroelectric power on Earth.

Similar general remarks may be made about Australia, whose much smaller area is covered mostly by desert, semiarid, and tropical wet-and-dry environments. Most of the agricultural productivity is peripheral, especially in the southeast. Metallic minerals at currently exploitable levels of size and richness support a substantial number of mining operations on the continent. *See* AFRICA; AUSTRALIA.

South America. The land resource is dominated by the unbroken extent of rainforest and wet-and-dry climates stretching east of the Andes from Colombia to northern Argentina and by substantial areas of water-deficient territory along the west coast and in the south and northeast of the South American continent. Some highly fertile soils in flat humid subtropical lands both west and east of the Paraná–La Plata river system are of minor extent by comparison with the whole. The Amazon Basin contains the largest stand of tropical hardwood forest on Earth. Metallic mineral resources are abundant in three general regions: the Andes, the largely crystalline rock highlands of eastern and southeastern Brazil, and the low plateau south of the lower Orinoco River. The Caribbean coast of Venezuela has the most productive petroleum reserves on the continent. *See* SOUTH AMERICA.

North America. Large sections of North American lands benefit from the advantages which characterize midlatitude humid-land resources under present technology. Disadvantages of desert and semiarid environments in much of the western half of the North American continent are tempered somewhat by the interspersal of mountain ranges throughout these drier regions. Taiga and other northern climatic environments are in considerable part coincident with the igneous and metamorphic rocks of the Laurentian Shield; and tropical environments are of small extent. In sum, this continent may be considered to have one of the best-balanced sets of resources, considering its substantial endowment in minerals of many different kinds, extensive forest lands, large annually renewed fresh-water supplies, great and varied agricultural lands, and the productive fisheries off both Atlantic and Pacific coasts. In addition, evidence has accumulated that in the southern section of the continent there may exist a major concentration of petroleum and natural gas resources extending southeastward from eastern Mexico and the contiguous continental shelf beneath the waters of the Gulf of Mexico to the fringes of the western Caribbean, with eventual productivity on the order of magnitude of the cluster of fields in the Middle East. Finally, North America has the highest ratio of employed resources to land area of all continents. In addition, it still contains considerable potential resources, including vast hydroelectric potential in eastern Canada (Quebec), and extensive (mostly) untapped oil deposits in the shale oil beds of Colorado and Wyoming, and the tar sands (thick oil in sandstone) of Alberta. *See* NORTH AMERICA.

Summary comment. The Earth's resource pattern has certain general characteristics. (1) Minerals usable under present technology are found in every environment, although mineral types differ according to location in sedimentary or igneous and metamorphic rock areas. Mineral exploration will continue indefinitely in all land areas, but the mineral resource possibilities of North America and the European part of the Eurasian continent have been examined in greater detail than those of any other large area. Ocean basins are the least-known part of the world as to mineral possibilities. (2) Agricultural lands and forest lands usable under present technology are dominated by those lying in midlatitudes. Sections of the taiga are important as forest resources. (3) The great potential agricultural and forest resources, if some technological improvement is assumed, lie within the humid tropical environments.

Donald J. Patton; Dale R. Lightfoot

Bibliography. M. Allaby, *Earth: Our Planet and Its Resources*, 1994; D. L. Clawson and J. S. Fisher, *World Regional Geography: A Development Approach*, 1998; G. T. Miller, *Environmental Science: Sustaining the Earth*, 4th ed., 1993; J. Rees, *Natural Resources: Allocation, Economics and Policy*, 1990; World Resources Institute, *World Resources 1998–1999*, 1998.

Earth rotation and orbital motion

The rotation of the Earth about its axis is demonstrated by the classical Foucault pendulum experiment. Its revolution in its orbit around the Sun is shown by the annual parallactic displacement of

relatively nearby stars against the background of more distant stars. However, because the Earth is not truly a rigid symmetric body and because it interacts with other members of the solar system gravitationally, these motions vary with time. *See* FOUCAULT PENDULUM; ORBITAL MOTION; PARALLAX (ASTRONOMY).

Rotation of the Earth

Historically the rotation of the Earth has been an essential part of timekeeping. It was assumed that the rotational speed of the Earth was essentially constant and repeatable, and that the length of the day resulting from this constant rotational speed was naturally useful as a measure of the passage of time. Astronomical observations, however, have shown that the speed with which the Earth is rotating is not constant with time. The variations in rotational speed may be classified into three types: secular, irregular, and periodic. The secular variation of the rotational speed refers to the apparently linear increase in the length of the day due chiefly to tidal friction. This effect causes slowing of the Earth's rotational speed and lengthening of the day by about 0.0005 to 0.0035 s per century.

The irregular changes in speed appear to be correlated with physical processes occurring on or within the Earth. These have caused the length of the day to vary by as much as 0.01 s over the past 200 years. Irregular changes consist of so-called decade fluctuations with characteristic periods of 5–10 years as well as variations that occur at shorter time scales. The decade fluctuations are apparently related to processes occurring within the Earth. The higher-frequency variations are now known to be largely related to the changes in the total angular momentum of the atmosphere. *See* ATMOSPHERIC GENERAL CIRCULATION.

Periodic variations are associated with periodically repeatable physical processes affecting the Earth. Tides raised in the solid Earth by the Moon and the Sun produce periodic variations in the length of the day of the order of 0.0005 s with periods of 1 year, $^1/_2$ year, 27.55 days, and 13.66 days. Seasonal changes in global weather patterns occurring with approximately annual and semiannual periods also cause variations in the length of the day of this order. *See* EARTH TIDES.

Knowledge of the rotational speed of the Earth is required for observers on the Earth who want to know the orientation of the Earth in an inertial reference frame. This includes navigators, astronomers, and geodesists. The rotational speed of the Earth remains essentially unpredictable in nature due to the incompletely understood irregular variations. Because of this, astronomical observations continue to be made regularly with improving accuracy, and the resulting data are the subject of continuing research in the field.

Observations of rotational speed. Astronomical observations of quasars, the Moon, and artificial Earth satellites are used to create a nonuniform time scale that is based strictly on the rotation of the Earth within an inertial reference system defined by the positions and motions of the celestial objects. This Universal Time scale (UT1) is compared with time scales known to be more uniform such as those determined from atomic clocks (International Atomic Time), the motion of the Earth about the Sun (Dynamical Time), or the motion of solar system objects. Variations in the differences among these types of time scales may be used to determine variations in the rotational speed of the Earth. Astronomical observations of time are made routinely by observatories located around the world for this purpose. Differences between the astronomically determined time scale and a uniform time scale are published routinely by the International Earth Rotation and Reference System Service (IERS). One-second adjustments are made at infrequent intervals to the uniform time scale Coordinated Universal Time (UTC) so that the difference between UT1 and UTC is never greater than 0.9 s. These adjustments are called leap seconds. UTC serves as the basis for civil time in most of the countries of the world. *See* ATOMIC CLOCK; ATOMIC TIME; DYNAMICAL TIME; TIME.

The times and locations of past total solar eclipse observations have also been analyzed to provide information on the length of the day. If the Earth has changed its speed of rotation since ancient times, the path of an eclipse that occurred thousands of years ago would be displaced in longitude with respect to the path that would have occurred if the rotational speed had remained constant. Ancient records of eclipses, while not made with great accuracy, are valuable because they were made long ago. Comparison of very old observations of the longitude of the Sun with current theories of the motion of the Sun based on a uniform time scale has also been used to estimate the increase in the length of the day since ancient times. *See* ECLIPSE.

Careful analyses of this information reveal the three types of variations in the speed of rotation. These ancient observational data form the basis for estimates of the secular deceleration in the speed of rotation. The more recent information, having been obtained with higher accuracy and more regularity, has shown the changes in the acceleration causing the irregular variations in the length of the day. These data have also been used to detect the periodic variations in the length of the day (**Fig. 1**).

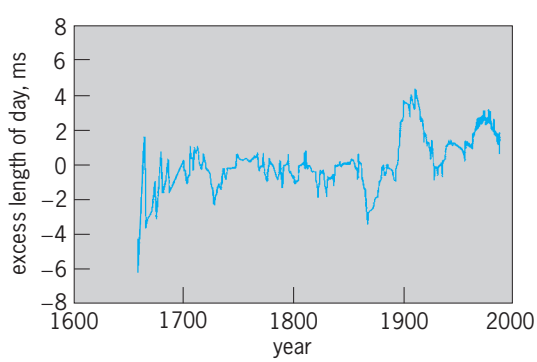

Fig. 1. Excess length of day (length of day—86,400 s), measured in milliseconds, since 1650.

Causes of variations. The conservation of angular momentum of the Earth requires that changes in the Earth's moments of inertia produce changes in the speed with which the Earth is rotating. The moments of inertia depend on the distribution of mass on and within the Earth. This includes the mass contained in the atmosphere and in the oceans. As mass is redistributed, the moments of inertia change, producing changes in the rotational speed of the Earth. An example of such an effect is the spinning ice skater who is able to change rotational speed by redistributing mass by changing the positions of the arms. *See* ANGULAR MOMENTUM; MOMENT OF INERTIA; ROTATIONAL MOTION.

The crust and the interior of the Earth are not strictly rigid. Tides generated by the Sun and the Moon change the shape of the Earth. The Earth is thought to have a fluid core with material possibly in motion. Coupling between such a core and the mantle would change the speed of rotation of the mantle from which the observations are made. The Earth may undergo changes in the moments of inertia as landmasses move or rise and fall. All of these motions cause changes in the Earth's moments of inertia and may cause changes in the rotational speed. Seasonal and short-term variations are caused by an exchange of momentum between winds and the crust of the Earth. Precise calculations of the total angular momentum using global observations of the wind speed and direction demonstrate this effect. The variations in the rotational speed are apparently caused by the interaction of the winds with the Earth's topography, and may be the leading cause for most of the irregular high-frequency variations in the rotation of the Earth.

Another known cause of variations in the rotational speed is tidal friction. The Moon raises tides in the ocean. Friction carries the maximum tide ahead of the line joining the center of the Earth and Moon (**Fig. 2**). The resulting couple diminishes the speed of rotation of the Earth, and this reacts on the Moon to increase its orbital momentum. The sum of the angular momentum of the Earth and the orbital momentum of the Moon remains constant. This produces an increase in the size of the orbit of the Moon and a reduction of its angular speed about the Earth. Tidal friction should be distinguished from actual changes in the moments of inertia of the Earth brought about by the tides. *See* TIDE.

The effect of tidal friction is to increase both the distance of the Moon and the length of the lunar month measured with a uniform time scale. Because of the change in the lunar month, the Moon is observed to have an orbital deceleration in terms of a uniform time scale. The proportional change in the distance, however, is greater than in the length of the month. Hence, in terms of Universal Time, the Moon appears to have a secular orbital acceleration, an effect discovered by E. Halley in 1693 from a study of ancient eclipses. In 1853 J. C. E. Adams found that gravitational theory could account for only about half of the acceleration found by Halley. It has become clear since then that the rotational speed is also decreased because of tidal friction.

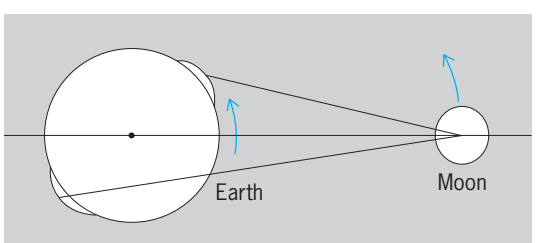

Fig. 2. Couple produced by tidal friction.

Revolution about the Sun

The motion of the Earth about the Sun is seen as an apparent annual motion of the Sun along the ecliptic. That the effect is caused by the motion of the Earth and not that of the Sun is proved by the annual parallactic displacement of nearby stars and by the aberration of light, which causes an apparent annual displacement of all stars on the celestial sphere. *See* ABERRATION (ASTRONOMY).

Orbit of the Earth. A large number of astronomical observations of the positions of the Sun and other solar system objects have been made and are being made continuously. This information is required to determine the nature of the motion of the Earth about the Sun. Observations are analyzed using the mathematical methods of celestial mechanics to provide improved estimates of the motions of the solar system objects in the future and to describe the past motions of the objects. The description of the apparent motion of the Sun in the sky provides the determination of the orbit of the Earth. The Earth moves around the Sun in a plane called the ecliptic. The orbit is characterized in part by the period of revolution, the mean radius, and its eccentricity, a quantity that describes the elliptical nature of the orbit. *See* CELESTIAL MECHANICS; ECLIPTIC; ORBITAL MOTION.

Period of revolution. The true period of the revolution of the Earth around the Sun is determined by the time interval between successive returns of the Sun to the direction of the same star. This interval is the sidereal year of 365 days 6 h 9 min 9.51 s of mean solar time or 365.25636 mean solar days. The period between successive returns to the moving vernal equinox is known as the tropical year of 365 days 5 h 48 min 45.2 s or 365.24219 days. The length of the tropical year is regarded as the length of the year in common usage for calendars. The period of time between successive passages at perihelion (the closest approach of the Earth to the Sun) is called the anomalistic year of 365 days 6 h 13 min 53.26 s or 365.25964 days. The length of each of these years depends on observationally determined astronomical quantities. Improvements in the determination of these quantities will result in slight changes in the numerical values. The lengths of the years listed above are given for the year 2000. These values vary slowly as a consequence of the long-period perturbations of the Earth's orbit by other planets. *See* CALENDAR; PERTURBATION (ASTRONOMY).

Mean radius of orbit. The mean distance from the Earth to the Sun, or the semimajor axis of the Earth's orbit, was the original definition of the astronomical unit (AU) of distance in the solar system. Its absolute value fixes the scale of the solar system and the whole universe in terms of terrestrial standards of length. The distance between the Earth and the Sun can be determined by a variety of methods; one method is measuring the solar parallax. The Sun's mean equatorial horizontal parallax p is the angle subtended by the equatorial radius r of the Earth at the mean distance a of the Sun. Mathematically, $a = r/\sin p = 206,265\ r/p$, if p is in seconds of arc. The equatorial radius of the Earth is 3963.182 mi. (6,378,136 m). Geometrical, gravitational, and physical methods have been used at various times to measure the solar parallax. *See* ASTRONOMICAL UNIT; PARALLAX (ASTRONOMY).

Geometrical methods. The geometrical methods involve the direct measurement by optical triangulation of the parallax of a nearby planet (Mars or Venus) or other nearby solar system object such as the minor planet Eros at its closest approach to the Earth. Because the relative distances in the solar system are accurately known in terms of the astronomical unit, the absolute measurement of one very accurate distance can be used to calibrate the scale of the system (**Fig. 3**).

Gravitational methods. The gravitational methods involve the determination of the ratio of the mass m of the Earth to the mass M of the Sun from the perturbations in the motion of a minor planet (such as Eros) caused by the Earth. The method also depends essentially on an accurate determination of the acceleration of gravity g and the length of the sidereal year T in seconds of mean solar time. The distance a is then found from the application of Kepler's third law, $4\pi^2a^3/T^2 = gr^2[1 + (M/m)]$. Because g and T are known with high accuracy, a determination of M/m gives a directly in terms of terrestrial standards of length. Although it is customary to express this result as a parallax angle, the Earth's radius is not needed to compute a, but it is necessary to make allowances for the mass of the Moon relative to the mass of the Earth. *See* GRAVITY.

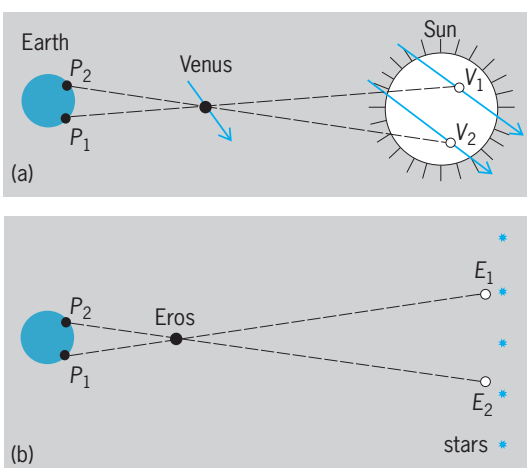

Fig. 3. Determination of solar parallax. (*a*) Observation of Venus in transit. (*b*) Observation of Eros in opposition.

Physical methods. One physical method depends on a determination of the ratio of the mean orbital velocity, $V = 2\pi a/T$, to the accurately known velocity of light. This ratio can be derived either from the annual variation of the radial velocities of ecliptic stars (or occasionally planets) determined by observations of the Doppler shift of spectral lines, or with less accuracy, from the constant of aberration. *See* DOPPLER EFFECT.

The most precise method relies on measurement of the travel time of electromagnetic signals reflected from objects in the solar system or received from artificial interplanetary probes. The distance of the planet is the ratio of half of the round-trip travel time of the electromagnetic signal to the velocity of the electromagnetic waves. Since the distance of the planet is known precisely in astronomical units from the orbital ephemerides, measuring the actual distance in meters at some given time permits the astronomical unit to be calibrated in terms of terrestrial units of length. Also the Doppler shift of radar reflections gives the radial velocity in $m \cdot s^{-1}$, which, on comparison with the known value in $AU \cdot s^{-1}$, gives the astronomical unit in meters. The currently adopted value of the astronomical unit is $1.495978706 \times 10^{11}$ m (92,955,807 mi).

Eccentricity of orbit. The eccentricity of the Earth's orbit can be determined by the variations of the apparent diameter of the Sun's disk. Determinations of higher accuracy are based on the variable speed of the Sun's apparent motion along the ecliptic and the laws of elliptic motion. The nonuniformity of the Sun's motion manifests itself in the equation of time, which is the difference between solar time determined from the actual observation of the Sun, and mean solar time which is based on the motion of a conventionally defined fictitious point having a uniform motion close to the average of the Sun's motion. The difference in the two types of time arises in part from the obliquity of the ecliptic (the angle between the equatorial plane of the Earth and the plane of the ecliptic) and in part from the eccentricity of the Earth's orbit. The adopted value of the eccentricity is 0.0167086171540. The number varies slowly with time due to perturbations from the planets. The Earth is at perihelion (closest distance to the Sun) on January 2. It reaches its greatest distance from the Sun on July 2. *See* EQUATION OF TIME.

Seasons. The fact that the Equator of the Earth is inclined in space by about 23.°5 to the orbital plane of the Earth (the ecliptic) causes the Northern Hemisphere to be exposed to the more direct rays of the Sun during part of the Earth's revolution around the Sun. The Southern Hemisphere receives the more direct rays 6 months, or a half revolution, later. This effect causes the seasons.

Other Motions

In addition to the rotation of the Earth and its orbital motion about the Sun, the Earth experiences various

small motions about its center of mass. Precession and nutation are examples, and these are caused by the gravitational attraction of the Sun and Moon on the nonspherical Earth. Because the Earth is ellipsoidal in shape, the gravitational attraction of these bodies produces a couple acting on the equatorial bulge, changing the orientation of the Earth about its center of mass. This motion can be predicted with accuracy, based on the knowledge of the shape of the Earth and of the motion of the Earth around the Sun and that of the Moon about the Earth. Precession causes the axis of angular momentum of the Earth to describe a 23.°5 cone in space with a period of about 26,000 years. Nutation causes the axis of angular momentum to "nod" slightly in space as it executes the precessional motion. The main period of the nutational motion is 18.6 years. Other periodic motions due to the gravitational attraction of the Sun and the Moon are also included in nutation. *See* NUTATION (ASTRONOMY AND MECHANICS); PRECESSION OF EQUINOXES.

Because the axis of symmetry of the Earth is not aligned precisely with the axis of rotation, the Earth also executes a motion about its center of mass known as polar motion. This motion, caused by geophysical and meteorological effects on and within the Earth, can be detected by the motion of the Earth's axis over the surface of the Earth. It is not predictable with accuracy, and must be observed continuously to provide the most precise information on the orientation of the Earth. Polar motion is characterized mainly by approximately 435-day and a 365-day periodic quasicircular motions of the axis of rotation on the surface of the Earth. The radius of the circular motion is of the order of 16 ft (5 m), but this varies. *See* EARTH; PLANET. Dennis D. McCarthy

Bibliography. J. M. A. Danby, *Fundamentals of Celestial Mechanics*, 2d ed., 1988; J. Kovalevsky, *Modern Astrometry*, 2d ed., 2002; K. Lambeck, *The Earth's Variable Rotation: Geophysical Causes and Consequences*, 1980, paper, 2005; D. D. McCarthy and W. E. Carter (eds.), *Variations in Earth Rotation*, 1990; H. Moritz and I. Mueller, *Earth Rotation, Theory and Observation*, 1987; D. Morrison, S. C. Wolff, and A. Franknoi, *Abell's Exploration of the Universe*, 7th ed., 1995; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, 1992; F. R. Stephenson, *Historical Eclipses and the Earth's Rotation*, 1997.

Earth sciences

Sciences that involve attempts to understand the nature, origin, evolution, and behavior of the Earth or of its parts and to comprehend its place in the universe, especially in the solar system. Understanding has advanced primarily through improved appreciation of the complex, usually cyclical interactions that take place among distinct parts of the Earth such as the lithosphere, atmosphere, hydrosphere, and biosphere. Geophysics is the study of the physics of the Earth, emphasizing its physical structure and dynamics. Geochemistry is the study of the chemistry of the Earth, dealing with its composition and chemical change. Geology is the study of the solid Earth and of the processes that have formed and modified it throughout its 4.5-billion-year history. *See* GEOCHEMISTRY; GEODESY; GEOLOGY; GEOPHYSICS; SOLAR SYSTEM.

Many branches of geology are considered separate sciences. Mineralogy is the study of the composition, structure, and properties of minerals. Petrology involves understanding how rocks originate and evolve, as well as rock description and classification. Specialties related to petrology include sedimentology and volcanology. Stratigraphy is the study of the origin, age, and development of layered, generally sedimentary rocks. Paleontology is the study of ancient (fossil) life. Historical geology is the study of the evolution of the Earth and its life. Geomorphology is the study of landscapes and their evolution. Seismology is the study of earthquakes and their effects. Structural geology is the study of deformed rocks. Engineering geology relates to the support of human constructions by underlying rock. *See* ENGINEERING GEOLOGY; GEOLOGY; GEOMORPHOLOGY; HYDROLOGY; MINERALOGY; PALEONTOLOGY; PETROGRAPHY; PETROLOGY; SEISMOLOGY; STRATIGRAPHY; STRUCTURAL GEOLOGY; VOLCANOLOGY.

Oceanography is the study of the oceans; limnology, the study of lakes; hydrology, the study of underground and surface water; and glaciology, the study of glaciers, ice caps, and ice sheets. These disciplines address the study of water in and on the Earth. The gaseous outer parts of the planet are the province of the atmospheric sciences, including meteorology, which is concerned with the weather and weather forecasting; climatology, which deals with longer-term and regional variations; and aeronomy which, because it deals with the outermost ionized region of the atmosphere, is much concerned with solar terrestrial interactions, including the aurora borealis and aurora australis. The biosphere embodies all life on Earth, and its study includes molecular biology, zoology, botany, and ecology. Geography, the study of all that happens at the Earth's surface, has been distinct insofar as it has encompassed not only physical and biological sciences but also the social sciences, including aspects of political science and economics. This distinction is fading rapidly as other earth sciences become more involved with social considerations.

A great deal of overlap exists among the earth sciences, and all have been subdivided into many disciplines. But progressively, especially since the International Geophysical Year (IGY) in 1957, subdivision has been complemented by integration as scientists have become more aware of the ways in which parts of the Earth system can interact. In trying to assess any possible impacts of predicted global warming, scientists have discerned interactions attributable to the burning of fossil fuels and the destruction of forests over the past 200 years. In the twenty-first century the effects of those interactions

could lead to both rising atmospheric temperatures and rising sea level. The prospect of possible adverse changes in the human environment has generated concern, and for that reason earth scientists are working with politicians to clarify what is happening and, as deemed necessary, to help initiate appropriate responses.

There are two main reasons why earth scientists have come to consider that they can help address such issues. First, Earth-observing instruments on satellites in geostationary and low Earth orbit now permit the acquisition, processing, and storage of the very large data sets that are required to show how the atmosphere, ocean, biosphere, hydrosphere, and lithosphere are interacting and changing globally on time scales from daily to decadal. Second, advanced computational capabilities now allow scientists to manipulate the large, satellite-derived data sets and also to produce models of the behavior of the interactive systems. These activities are in their infancy. It is not possible to say how they may develop as the earth sciences evolve. The success of scientists in predicting, months in advance, the advent of a severe ENSO (El Niño Southern Oscillation) is an example of the kind of activity that may become more common. One such ENSO prediction was far enough in advance to enable farmers to switch to the planting of drought-resistant crops.

Although studies in the earth sciences that are focused on the well-being of the growing human population are likely to be increasingly important, basic understanding in the earth sciences is also likely to come to depend more on very large data sets. Improved understanding of the fundamentals of the Earth system will require the acquisition of very large data sets, their manipulation, processing, and storage. This will be complemented by the improving ability to construct models of the parts of the Earth system and of the ways in which those parts interact. Models will be constructed, tested, and abandoned in favor of better models as more and better data are acquired. Quality control on both data acquisition and processing will be critical. Kevin Burke

Bibliography. Intergovernmental Panel on Climate Change (IPCC), Working Group 1, *Climate Change 1995: The Science of Climate Change*, ed. by J. T. Houghton et al., Cambridge University Press, New York, 1996; F. Press and R. Siever, *Understanding Earth*, 2d ed., W. H. Freeman, New York, 1998.

Earth tides

Cyclic motions of the Earth, sometimes over a foot or so in height, depending on latitude, caused by the same lunar and solar forces which produce tides in the sea. These forces also react on the Moon and Sun, and thus are significant in astronomy in evaluations of the dynamics of the three bodies. For example, the secular spin-down of the Earth due to lunar tidal torques is best computed from the observed acceleration of the Moon's orbital velocity. In oceanography, earth tides and ocean tides are very closely related. *See* EARTH, GRAVITY FIELD OF; GEODESY; TIDE.

By far the most widely used earth tide instruments are the tiltmeter and the gravimeter. Both instruments have the merits of portability, high potential precision, and low cost. Thus they are able to advance economically an important mission—the global mapping of earth tides and ocean tides. *See* GRAVITY METER; SEISMOGRAPHIC INSTRUMENTATION.

Tide-producing potential. Tidal theory begins simply, by evaluating the tidal forces produced on a rigid, unyielding, oceanless, spherical Earth by the Moon and Sun. The Earth, Moon, and Sun can be regarded as perfectly spherical; the results of this idealized assumption will be modified later in this text to relate earth tide observations to deformable Earth models.

In the expansion of the total gravitational potential of the satellite as a sum of solid spherical harmonics, the tide-producing potential consists only of those terms which vary within the Earth and are the source of Earth deformations. The omitted terms are the constant which is arbitrarily chosen to produce null potential at the Earth's center and the first-degree term expressing the uniform acceleration of every part of the Earth toward the satellite (thus also entailing no distortion). Accordingly, the tide-producing potential U is expressible as the sum of solid spherical harmonics of degree 2 and higher in the notation which follows.

If the mass of the satellite is denoted by m, its distance from the Earth's center by R, the distance of the observing point from the Earth's center by r, and the geocentric zenith angle measured from the line of centers by θ, then for points $r < R$, Eq. (1)

$$U = U_2 + U_3 + \cdots = GmR^{-1}\left[\left(\frac{r}{R}\right)^2 P_2(\cos\theta) + \left(\frac{r}{R}\right)^3 P_3(\cos\theta) + \cdots\right] \quad (1)$$

holds, where G is Newton's universal gravitational constant (6.67×10^{-8} cm^3 g^{-1} s^{-2}).

On an idealized oblate Earth of equatorial radius r_e and flattening constant f, a point on the surface at geocentric latitude ϕ has radial distance $r = r_e C(\phi)$, where $C(\phi) \equiv 1 - f \sin^2 \phi$. The vertical (upward) component of tidal gravity $\partial U/\partial r$ and the horizontal component $r^{-1}\partial U/\partial\theta$ in the azimuth direction away from the satellite, with $P^1{}_n(\cos\theta) = -\partial P_N(\cos\theta)/\partial\theta$, are given by Eqs. (2) and (3). Here

$$g_r = \frac{\partial U}{\partial r} = GmC\alpha^3 r_e^{-2} \sum_{n=2}^{\infty} n(\alpha C)^{n-2} P_n(\cos\theta) \quad (2)$$

$$g_\theta = r^{-1}\frac{\partial U}{\partial\theta} = -GmC\alpha^3 r_e^{-2} \sum_{n=2}^{\infty} (\alpha C)^{n-2} P_n^1(\cos\theta) \quad (3)$$

$\alpha_m \equiv r_e/R_m$ and $\alpha_S \equiv r_e/R_S$ for the Moon and Sun, respectively.

TABLE 1. Parameters for some equilibrium tides

Species	Darwin symbol	Doodson number	Period	Amplitude factor, *b*
Long period	Lunar	055565	18.6 y	0.066
	Sa	056554	1 y	0.012
	SSa	057555	$^1/_2$y	0.073
	MSm	063655	31.85 d	0.016
	Mm	065455	27.55 d	0.083
	MSf	073555	14.77 d	0.014
	Mf	075555	13.66 d	0.156
	—		13.63 d	0.065
Diurnal	O_1	145555	25.82 h	0.377
	P_1	163555	24.07 h	0.176
	K_1	165555	23.93 h	0.531
Semidiurnal	N_2	245655	12.66 h	0.174
	M_2	255555	12.42 h	0.908
	S_2	273555	12.00 h	0.423
	K_2	275555	11.97 h	0.115

Harmonic constituents. Equations (2) and (3) indicate the simple nature of the dependence of tidal gravity on the rigid Earth upon the distance R and the zenith angle θ of the satellite. But the time variations of R and θ are complex, because of the Earth's rotation and the complexity of the orbital motions. The objective in earth tide observations is assistance in determining present relevant properties of the Earth. **Table 1** lists the periods and amplitude factors b of the larger tidal harmonic constituents over the broad range of tidal periods. In classifying tidal constituents mnemonically, G. H. Darwin used alphabets in terms of pairs with respect to the Moon (M) and the Sun (S), and A. T. Doodson used argument numbers, which are the coefficients of the six astronomical, independent variables arranged in order of increasing speed. The first three argument numbers are called the constituent number, of which the first two-digit number is the group number. The first digit of the group number designates species 0, 1, and 2 for long-period, diurnal, and semidiurnal, respectively. The purely geometric complexities essential in applied tidal theory are deemphasized here and, when feasible, the simple zenith-centered satellite coordinate system used in Eqs. (1)–(3) is adopted.

Tidal torques. The orbital acceleration $\dot{n}_m$ of the Moon is the critical quantity in Eq. (4), giving the

$$N_m = -{}^1/_3[M_e M_m/(M_e + M_m)]r_m^2 \dot{n}_m \qquad (4)$$

lunar torque N_m, which slows the Earth's spin. (Here M_e and M_m are masses of the Earth and Moon respectively, and r_m is the distance to the Moon.) The value for N_m has been calculated to equal -3.9×10^{23} dyne-cm, to which corresponds the energy loss rate $\dot{E}$, shown in Eq. (5), and the relative spin-down,

$$\dot{E} = N(\omega - n_m) = 2.7 \times 10^{19} \text{ ergs/s} \qquad (5)$$

$-\dot{\omega}/\omega = N/c\omega = 0.21$ eon^{-1} (1 eon $= 10^9$ years). Here ω is the Earth's rotational velocity, and c its principal moment of inertia.

There is an observation equation for the M_2 ocean tidal constituent from which an estimate of the tidal acceleration of lunar longitude is made; the value $\dot{n}_m = -27.4'' \pm 3.0''$ century^{-2} is obtained. That means $N_m = -4.8 \times 10^{23}$ dyne-cm, to which corresponds the energy loss rate $E = 3.3 \times 10^{19}$ ergs/s, and the relative spin-down value of 0.26 eon^{-1}.

Tidal loss in the solid Earth. The gravest free mode of the Earth, period 0.9 h, has the same external form and nearly the same internal geometry as the M_2 tide bulge. Its observed $Q = 350 \pm 100$ can be used to estimate the loss rate $\dot{E}_B$ in the bodily M_2 tide (Q^{-1} is the loss rate number, and M_2 is the principal semidiurnal tide). The result, $\dot{E}_B = 7 \times 10^{17}$ ergs/s, is only 3% of the required total in Eq. (5). The rate at which the Earth dissipates the total lunar-solar tidal energy has also been estimated to be $5.7 \pm 0.5 \times 10^{19}$ ergs/s, the share attributed to the oceans is $5.0 \pm 0.3 \times 10^{19}$ ergs/s.

Instrumental dimensionless amplitude factors. The reading of an earth tide meter is altered by the fact that it is anchored to a yielding Earth. The Earth chosen is assumed to be the radially symmetric standard.

Gravimeter. Basically, a gravimeter consists of a mass generally supported by a spring. (The superconducting model used a magnetic field.) Variations in gravity are measured by the extension of the spring or in the null method by the small corrections required to restore the original configuration. This low-drift-rate superconducting gravimeter is designed to increase to several months the intervals between replenishing the helium supply. It is expected to give high precision in gravity measurements. A satellite affects the mean local value of gravity in three ways: by its direct attraction, by the tidal change in elevation of the observing station, and by the redistribution of mass in the deformed Earth.

Tiltmeters. The tiltmeter measures changes in the angle of tilt between the tiltmeter's foundation and the local vertical, both subject to tidal variations. Several types of tiltmeters are in use. In the horizontal pendulum the rotation axis is fixed at a small angle with the vertical to produce high sensitivity to tilts. In the level tube, which may be several hundred meters or more long, the difference in elevation of a fluid is measured at the two ends. In this way a sample of the tilt of a large region is obtained, but the

horizontal pendulum sometimes also samples a large volume.

Extensometer. The extensometer, or linear strain meter, measures the change in distance between two reference points. In the Benioff design these positions are connected with a quartz tube to bring the points into juxtaposition for precise measurements of relative displacement. By using laser beams, the measurements may be made over distances of kilometers, but the short-baseline strain-mechanical meters still have their use.

Indicated and true phase lags. The phase indications of a gravimeter or tiltmeter are only a fraction of the true angular lag of the tidal bulge. The tidal bulge is assumed to retain its no-loss equilibrium form, but with the axis carried forward from the Moon's zenith by the small angle ϵ required to produce the Moon's known orbital acceleration.

Theories and characteristic numbers. Three basic numbers, Love numbers h and k and Shida number l, characterize the elastic behavior of an elastically yielding Earth. Thus h represents the ratio of the tidal height of the yielding Earth to that of the equilibrium tide, k the ratio of the additional potential due to the tidal deformation of the yielding Earth to the deforming potential, and l the ratio of horizontal displacement of the yielding Earth to that of the equilibrium tide.

Static theory. Static theory of earth tides (the frequency domain) is formulated for a spherical, nonrotating, and isotropic Earth. It is based on the reduction of the problem of elastic deformation of a sphere to a system of six first-order differential equations which are subjected to numerical integration in one formulation. When these equations are integrated with different Earth models by means of the Runge-Kutta method, the second-order Love numbers that are derived are virtually independent of either the presence of the solid inner core or the presence or absence of a low-velocity layer in the upper mantle (**Table 2**). Also, the continental or oceanic crustal model is only of second order with respect to the effects of the oceans.

Dynamic theory. Theoretically, a torque-free, nearly diurnal nutation or resonance is indeed possible by the presence of the liquid core inertially coupled to the mantle, and the predicted motion can give a resonance amplification to the nearly diurnal tides. However, theoretical Earth models, while complicated, are a great simplification of the actual Earth.

Time domain. Traditionally, theories of tides for the Earth, the oceans, or the atmosphere were always in the frequency domain. However, tides as observed are typically in the time domain, that is, in time series. Although development of theories of tides in the time domain is more complicated than that in the frequency domain, for practical purposes it would be more realistic to develop theories of tides in the time domain. The Earth, the oceans, and the atmosphere eventually must be treated as a single system, responding to the complete spectrum of the tide-generating forces, wind stresses, and interactions, as well as other periodic and nonperiodic dynamic forcings.

The time-domain earth tide is solved for a spherical, nonrotating, elastic, and isotropic Earth, and also for a rotating Earth. In their formulation, both the spheroidal and toroidal components are simultaneously introduced into the general form of the linearized lagrangian equations of motion, which govern small elasto-gravitational disturbances away from equilibrium of an arbitrary, uniformly rotating, self-gravitating, elastic Earth with an arbitrary initial state of stress field. The resulting equations of motion are those of a set of integro-partial differential equations, the solutions to which satisfy a set of boundary conditions in the Earth. There are strong coupling effects due to the rotation of the Earth.

Interaction of earth and ocean tides. Neither the problem of earth tides nor that of ocean tides can be solved independently. Characterization of ocean tides in the open oceans is uncertain. Calculations of ocean tides in the open oceans made directly from Laplace's tidal equations without consideration of the elastic yielding of the Earth due to the tidal and load deformation are inadequate. The interdependency of the problems of earth and ocean tides is a problem in geodynamics.

Tidal gravity and ocean tides. A transcontinental tidal gravity profile, consisting of nine stations across the United States along the 40th parallel, show the consistent results of the dominant effect of the ocean tides in the Pacific and Atlantic oceans (**Figs. 1** and **2**).

Ocean tides are the primary cause of the observed tidal gravity variations. These perturbations essentially depend only on the amplitude and phase of the ocean tide and the distance between the

TABLE 2. Earth tide parameters for different Earth models

Model	h	k	l	k/h	δ	γ	l/h
Alsop and Kuo							
Gutenberg-Bullen	0.619	0.305	0.088	0.493	1.162	0.686	0.142
Jeffreys-Bullen B	0.618	0.304	0.088	0.492	1.162	0.686	0.142
Jeffreys-Bullen A	0.607	0.299	0.081	0.493	1.158	0.692	0.133
Gutenberg-Bullen A	0.608	0.300	0.082	0.493	1.158	0.692	0.135
Farrell							
Gutenberg-Bullen Earth model	0.3040	0.6114	0.0832	0.497	1.1554	0.6926	0.136
Oceanic mantle	0.3055	0.6149	0.0840	0.497	1.1567	0.6906	0.137
Shield mantle	0.3062	0.6169	0.0842	0.496	1.1576	0.6893	0.136

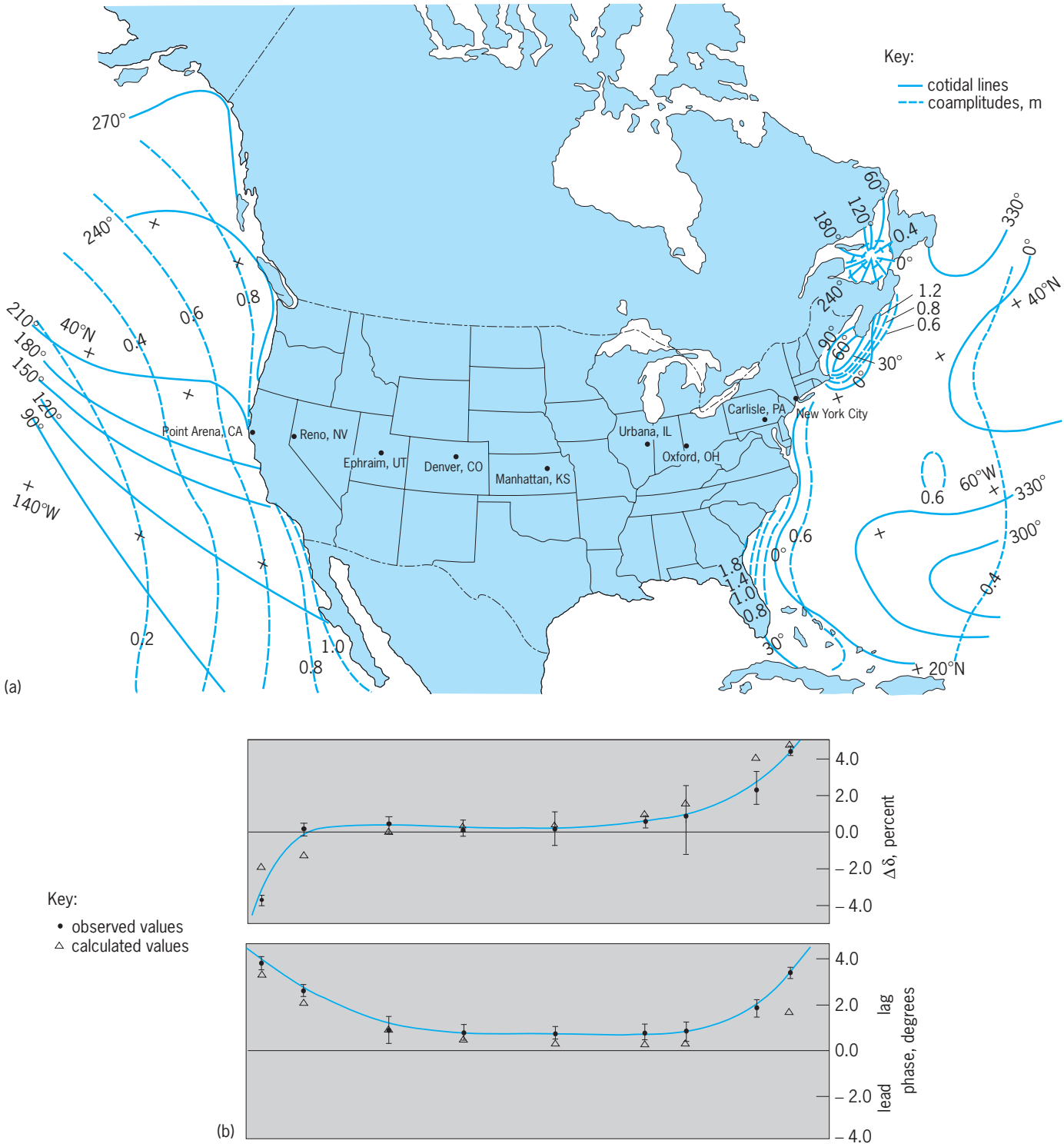

Fig. 1. Effect of the ocean tides in the Atlantic and Pacific oceans on the earth tides for the semidiurnal tidal constituent M_2. (*a*) Ocean tides. (*b*) Comparison of the observed and calculated values. (*After J. T. Kuo et al., Transcontinental gravity profile across the United States, Science, 168:968–971, 1970*)

observation station and the ocean tidal load; land-based and island-based tidal gravity measurements can provide an independent set of observation data. If tidal gravity measurements are made accurately, it is more appropriate to consider the inverse problem of mapping ocean tides in the open oceans with extended earth tidal gravity measurements on the adjacent lands, supplemented by ocean-bottom stations and the ocean tidal information derived from nearshore and island stations.

Open ocean tides can be mapped by solving the inverse problem using land- and island-based tidal gravity measurements, coupled with shore and island

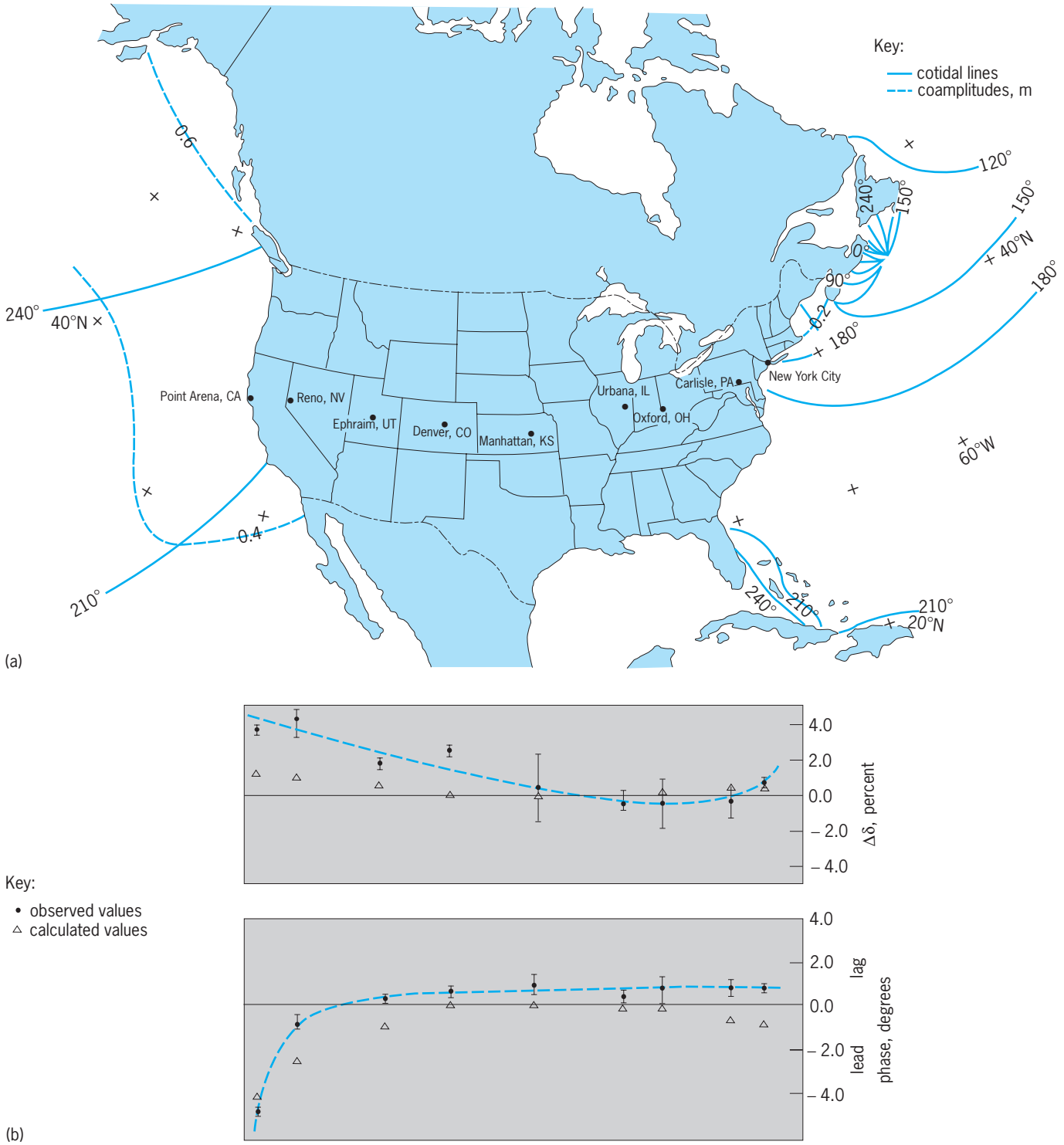

Fig. 2. Effect of the ocean tides in the Atlantic and Pacific oceans on the earth tides for the diurnal tidal constituent O_1. (*a*) Ocean tides. (*b*) Comparison of the observed and calculated values. (*After J. T. Kuo et al., Transcontinental gravity profile across the United States, Science, 168:968–971, 1970*)

ocean tidal measurements and a few ocean-bottom measurements.

Modified Laplace's tidal equations. Understanding of open ocean tides can be obtained both through numerical integration of Laplace's tidal equations and through direct measurements of tides in the deep oceans. However, the coamplitude and cophase tidal charts, principally of the tidal constituent M_2 produced by means of the numerical integration of Laplace's tidal equations, apparently are quite sensitive to the boundaries of the ocean basins and the law of friction, and generally fail to give close agreement with the tidal observations on mid-oceanic islands. Thus, the original formulation of Laplace's tidal equations, which are based on the assumption of a rigid Earth, are inadequate to serve as a theoretical basis

for calculating the ocean tides. The contribution of the tidal and loading deformation of the elastic Earth cannot be neglected in the formulation of the tidal equations.

Tidal tilt and tidal strain. Measurements of tidal tilt and tidal strain are more complicated than those of tidal gravity. Unlike tidal gravity measurements, which are virtually independent of inhomogeneities of the elastic properties in the Earth's crust or even in the upper mantle, tidal tilt and tidal strain measurements, in addition to the influence of ocean tidal loading, are very sensitive to the influences of the inhomogeneities, including the site of instrumental installation, topography, and geological structure.

It is well known that the underground openings in otherwise uniform rock mass strongly distort the elastic strain field. Tiltmeters installed in underground tunnels generally measure cross-tunnel strain-induced tilts, in addition to the regional tilt.

Moreover, the inhomogeneities in the Earth's crust, such as variable surface topography and geological discontinuities and the cavity in which the instruments are installed, are recognized as having significant influence in tidal tilt and tidal strain measurements. Tidal strain was observed from seven strain stations across the continental United States. A comparison of the tidal strain as predicted for a radially symmetrical, stratified, and homogeneous Earth model with corrections for the loading effects of the worldwide ocean tides (derived from the numerical integration of Laplace's tidal equations) indicated that the topographic and geological influences on tidal strain amount to as much as ±25% of the tidal strain for some of the strain stations, and the topographic effect on tidal strain exceeded ±50% of the tidal strain in some extreme cases.

Satellite and interferometry measurements. It is possible to estimate values for the amplitudes and phases of the M_2, K_1, and P_1 ocean tidal constituents. With improved accuracy of satellite orbital determination, long time spans of accumulated data would permit an improvement in the accuracy of ocean tidal parameters in the near future.

Very long baseline interferometry (VLBI) has also been used to determine vector separations between radio telescopes and positions of radio sources. It has a precision of under 2 in. (5 cm) for baselines from 600 to 2400 mi (1000 to 4000 km) and less than about 4 in. (subdecimeter) for intercontinental lengths of about 3600 mi (6000 km).

Nearly diurnal nutation. The motion of the nearly diurnal nutation (or resonance) describes the free oscillation of the mantle's inertia axis about the axis of the Earth's rotation. The natural period of such a resonance of the Earth depends on the internal structures and elastic properties of the Earth and is not exactly equal to a sidereal day. However, calculations indicate that only the major K_1 and the minor ψ_1 tidal constituents with a period close to a sidereal day have amplitudes changed by the resonance effect. The theoretical results of Molodensky model II have been compared with the observed values of

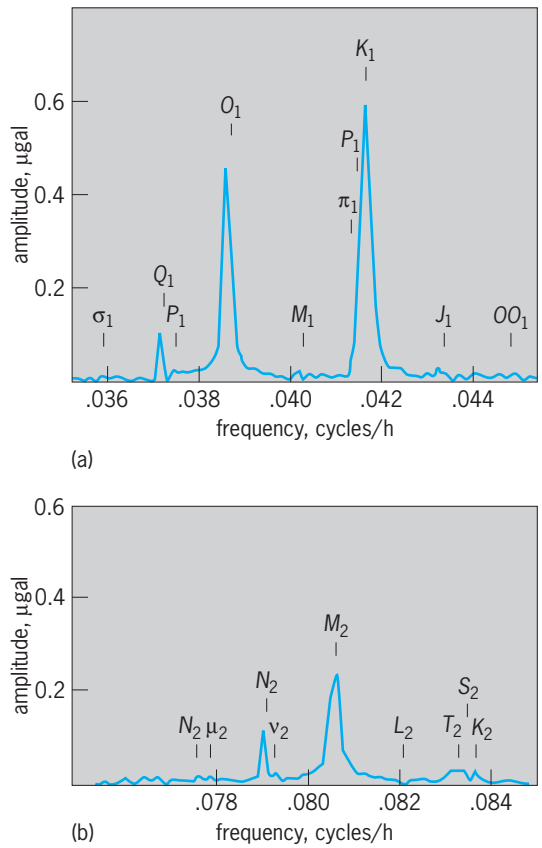

Fig. 3. Observed amplitude spectrum of (*a*) diurnal and (*b*) semidiurnal constituents at South Pole. 1 μgal = 1 μm/s^2. (*After B. V. Jackson and L. B. Slichter, The residual daily earth tides at the South Pole, J. Geophys. Res., 79(11):1711–1715, 1974*)

the tilt and gravimetric factors obtained by a number of workers at the nearly diurnal resonance of the solar waves ψ_1 and K_1 (**Fig. 3**).

By stacking qravity measurements, the complex eigenfrequency with strength, period, and damping of the nearly diurnal free wobble are determined. The estimate of the real part of the strength amplitude is in close agreement with the first experimental determination of the inner-pressure gravimetric factor.

Effects of inertia, ellipticity, and anisotropy. In an elliptical and rotating Earth, there is a coupling between the spherical and toroidal displacement fields in the core. In 1981 the Dahlen-Smith formulation of free oscillation of a rotating, elliptical Earth was adopted to calculate the gravimetric factors for M_2 and O_1 as a function of latitude, which reach a maximum at the Equator and a minimum at the poles. A uniform means of instrumental calibration may bring the observed values of later workers close to that given by the former method.

One intriguing problem in earth tidal studies is the effect of anisotropy in the Earth's mantle on the earth tide and ocean-load tide. Equations of deformation of the Earth with anisotropic regions through lagrangian mechanics are derived. The solution indicates that earth tides are virtually blind, at least to

the anisotropy of the upper mantle. However, load tides are affected by as much as 2.5%.

Earth tides and earthquakes. Volcanic earthquake swarms at Pavlof Volcano (southwestern Alaska Peninsula) correlated significantly with solid earth tidal stress rate for periods just before and just after explosive eruptions. In the 1974 minor eruption sequence, preeruptive earthquake swarms occurred during increasing earth tidal extension. Earth tides may have an effect on certain aspects of volcanic activity at Pavlof, and the polarity of the observed tidal correlation varies systematically during a volcanic eruption. *See* EARTHQUAKE.

Louis B. Slichter; John T. Kuo

Bibliography. O. Francis, and P. Mazzega, What we can learn about ocean tides from tide gauge and gravity loading measurements, *Proc. 11th Int. Symp. Earth Tides*, 1991; J. Hinderer, W. Zurn, and H. Legros, Interpretation of the strength of the "nearly diurnal free wobble" resonance from stacked gravity tide observations, *Proc. 11th Int. Symp. Earth Tides*, 1991; J. T. Kuo, Y. C. Zhang, and Y. H. Chu, Time-domain total Earth tides, *Proc. 10th Int. Symp. Earth Tides*, 1985; G. Y. Li and H. T. Hsu, Tidal modeling theory with a lateral inhomogeneous, inelastic mantle, *Proc. 11th Int. Symp. Earth Tides*, 1991; S. D. Pagistakis, Effect of anisotropy in the Earth's mantle on body and ocean load tide, *Proc. 11th Int. Symp. Earth Tides*, 1991; Y. C. Zhang and J. T. Kuo, Time-domain earth tides for a rotating Earth, *Proc. 11th Int. Symp. Earth Tides*, 1991.

Earthquake

The sudden movement of the Earth caused by the abrupt release of accumulated strain along a fault in the interior. The released energy passes through the Earth as seismic waves (low-frequency sound waves), which cause the shaking. Seismic waves continue to travel through the Earth after the fault motion has stopped. Recordings of earthquakes, called seismograms, illustrate that such motion is recorded all over the Earth for hours, and even days, after an earthquake.

Characteristics

Earthquakes vary immensely in size, from tiny events that can be detected only with the most sensitive seismographs, to great earthquakes that can cause extensive damage over widespread areas. Although thousands of earthquakes occur every day, and have for billions of years, a truly great earthquake occurs somewhere in the world only once every year. When a great earthquake occurs near a highly populated region, tremendous destruction can occur within a few seconds. In 1976, 600,000 people were killed in Tangshan, China, by a single earthquake. The city of Lisbon, one of the principal capitals of that day, was utterly destroyed, with high loss of life, in 1755. In the twentieth century such cities as Tokyo and San Francisco have been leveled by earthquakes. In these more modern cases, much of the damage was not due to the shaking of the earthquake itself, but was caused by fires originating in the gas and electrical lines which interweave modern cities, and by damage to fire-fighting capability which rendered the cities helpless to fight the conflagrations.

Cause. The locations of earthquakes which occurred between 1957 and 2000 are shown on the map in **Fig. 1**. The map shows that earthquakes are not distributed randomly over the globe but tend to occur in narrow, continuous belts of activity. Approximately 90% of all earthquakes occur in these belts, which define the boundaries of the Earth's plates. The plates are in continuous motion with respect to one another at rates on the order of centimeters per year; this plate motion is responsible for most geological activity.

Plate motion occurs because the outer cold, hard skin of the Earth, the lithosphere, overlies a hotter, soft layer known as the asthenosphere. Heat from decay of radioactive minerals in the Earth's interior sets the asthenosphere into thermal convection. This convection has broken the lithosphere into plates which move about in response to the convective motion in a manner shown schematically in **Fig. 2**. The plates move apart at oceanic ridges. Magma wells up in the void created by this motion and solidifies to form new sea floor. This process, in which new sea floor is continually created at oceanic ridges, is called sea-floor spreading. Since new lithosphere is continually being created at the oceanic ridges by sea-floor spreading, a like amount of lithosphere must be destroyed somewhere. This occurs at the oceanic trenches, where plates converge and the oceanic lithosphere is thrust back down into the asthenosphere and remelted. The melting of the lithosphere in this way supplies the magma for the volcanic arcs which occur behind the trenches. Where two continents collide, however, the greater bouyancy of the less dense continental material prevents the lithosphere from being underthrust, and the lithosphere buckles under the force of the collision, forming great mountain ranges such as the Alps and Himalayas. Where the relative motion of the plates is parallel to their common boundary, slip occurs along great faults which form that boundary, such as the San Andreas fault in California. *See* ASTHENOSPHERE; EARTH, HEAT FLOW IN; LITHOSPHERE; MAGMA; VOLCANOLOGY.

According to the theory of plate tectonics, the motion of the plates is very similar to the movement of ice floes in arctic waters. Where floes diverge, leads form and water wells up, freezing to the floes and producing new floe ice. The formation of pressure ridges where floes converge is analogous to the development of mountain ranges where plates converge. *See* OROGENY; PLATE TECTONICS.

Stick-slip friction and elastic rebound. As the plates move past each other, little of the motion at their boundaries occurs by continuous slippage; most of the motion occurs in a series of rapid jerks. Each jerk is an earthquake. This happens because, under

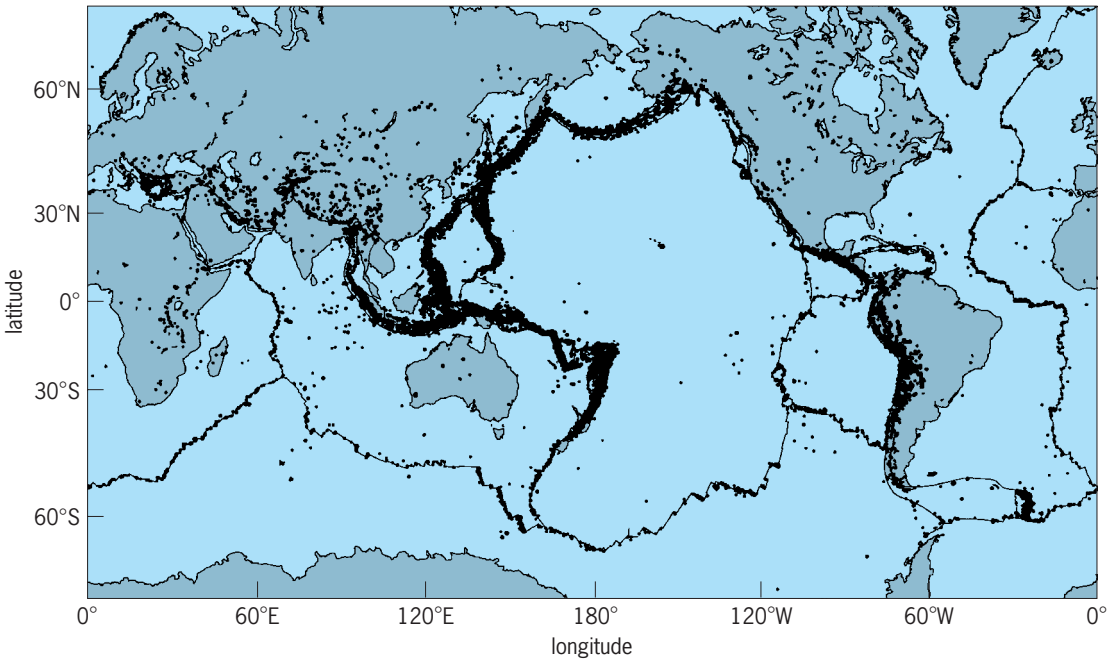

Fig. 1. Seismicity of the Earth from 1957 to 2000; depths to 700 km (435 mi). Earthquakes are plotted as circles; the plate boundaries are black. (*After E. R. Engdahl et al., Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, Bull. Seis. Soc. Amer., 88:722–743, 1998; and A. Villaseñor et al., Toward a comprehensive catalog of global historical seismicity, Eos, 78:581, 583, 588, American Geophysical Union, 1997*)

the pressure and temperature conditions of the shallow part of the Earth's lithosphere, the frictional sliding of rock exhibits a property known as stick-slip, in which frictional sliding occurs in a series of jerky movements, interspersed with periods of no motion—or sticking. In the geologic time frame, then, the lithospheric plates chatter at their boundaries, and at any one place the time between chatters may be hundreds of years.

The periods between major earthquakes is thus one during which strain slowly builds up near the plate boundary in response to the continuous movement of the plates. The strain is ultimately released by an earthquake when the frictional strength of the plate boundary is exceeded. This pattern of strain buildup and release was discovered by H. F. Reid in his study of the 1906 San Francisco earthquake. During that earthquake, a 250-mi-long (400-km) portion of the San Andreas fault, from Cape Mendocino to the town of Gilroy, south of San Francisco, slipped an average of 12 ft (3.6 m). Subsequently, the triangulation network in the San Francisco Bay area was resurveyed; it was found that the west side of the fault had moved northward with respect to the east side, but that these motions died out at distances of 20 mi (32 km) or more from the fault. Reid had noticed, however, that measurements made about 40 years prior to the 1906 earthquake had shown that points far to the west of the fault were moving northward at a slow rate. From these clues, he deduced his theory of elastic rebound, illustrated schematically in **Fig. 3**. The figure is a map view, the vertical line representing the fault separating two moving plates. The unstrained rocks in Fig. 3*a* are distorted by the slow movement of the plates in Fig. 3*b*. Slippage in an earthquake, returning the rocks to an unstrained state, occurs as in Fig. 3*c*. *See* FAULT AND FAULT STRUCTURES.

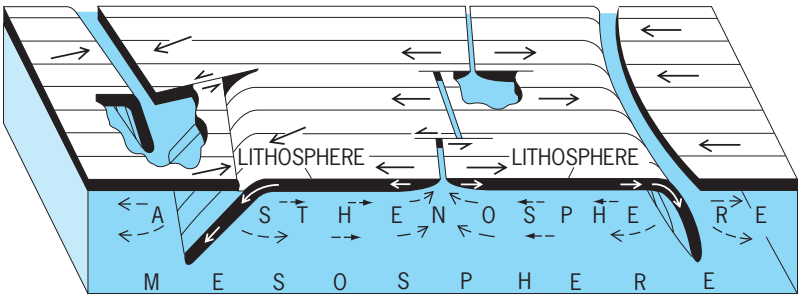

Fig. 2. Movement of the lithosphere over the more fluid asthenosphere. In the center, the lithosphere spreads away from the oceanic ridges. At the edges of the diagram, it descends again into the asthenosphere at the trenches. (*After B. Isacks, J. Oliver, and L. R. Sykes, Seismology and the new global tectonics, J. Geophys. Res., 73:5855–5899, American Geophysical Union, 1968*)

Classification. Most great earthquakes occur on the boundaries between lithospheric plates and arise directly from the motions between the plates. Although these may be called plate boundary earthquakes, there are many earthquakes, sometimes of substantial size, that cannot be related so simply to the movements of the plates.

Near many plate boundaries, earthquakes are not restricted to the plate boundary itself, but occur over a broad zone—often several hundred miles wide—adjacent to the plate boundary. These earthquakes, which may be called plate boundary-related earthquakes, do not reflect the plate motions directly, but are secondarily caused by the stresses set up at the plate boundary. In Japan, for example, the plate boundaries are in the deep ocean trenches offshore

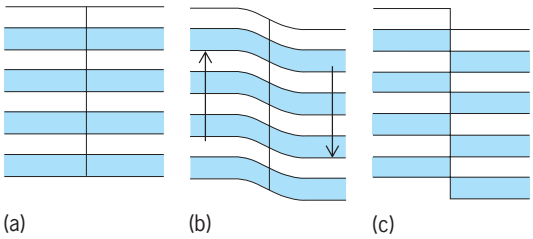

Fig. 3. Schematic of elastic rebound theory. (*a*) Unstrained rocks (*b*) are distorted by relative movement between the two plates, causing strains within the fault zone that finally become so great that (*c*) the rocks break and rebound to a new unstrained position. (*After C. R. Allen, The San Andreas Fault, Eng. Sci. Mag., Calif. Inst. Technol., pp. 1–5, May 1957*)

of the Japanese islands, and that is where the great plate boundary earthquakes occur. Many smaller events occur scattered throughout the Japanese islands, caused by the overall compression of the whole region. Although these small events are energetically minor when compared to the great offshore earthquakes, they are often more destructive, owing to their greater proximity to population centers.

Although most earthquakes occur on or near plate boundaries, some also occur, although infrequently, within plates. These earthquakes, which are not related to plate boundaries, are called intraplate earthquakes, and can sometimes be quite destructive. The immediate cause of intraplate earthquakes is not understood. Some of them can be quite large. Three of the largest earthquakes known to have occurred in the United States were part of a sequence of intraplate earthquakes which took place in the Mississippi Valley, near New Madrid, Missouri, in 1811 and 1812. Another intraplate earthquake, in 1886, caused moderate damage to Charleston, South Carolina.

In addition to the tectonic types of earthquakes described above, some earthquakes are directly associated with volcanic activity. These volcanic earthquakes result from the motion of undergound magma that leads to volcanic eruptions.

Sequences. Earthquakes often occur in well-defined sequences in time. Tectonic earthquakes are often preceded, by a few days to weeks, by several smaller shocks (foreshocks), and are nearly always followed by large numbers of aftershocks. Foreshocks and aftershocks are usually much smaller than the main shock. Volcanic earthquakes often occur in flurries of activity, with no discernible main shock. This type of sequence is called a swarm.

Size. Earthquakes range enormously in size, from tremors in which slippage of a few tenths of an inch occurs on a few feet of fault, to the greatest events, which may involve a rupture many hundreds of miles long, with tens of feet of slip. Accelerations exceeding 1 *g* (acceleration due to gravity) can occur during an earthquake. The velocity at which the two sides of the fault move during an earthquake is only 1–10 mi/h (10–100 cm/s), but the rupture front spreads along the fault at a velocity of nearly 5000 mi/h (3 km/s). The earthquake's primary damage is due to the generated seismic waves, or sound waves which travel through the Earth, excited by the rapid movement of the earthquake. The energy radiated as seismic waves during a large earthquake can be as great as 10^{12} cal (4.19×10^{12} joules), and the power emitted during the few hundred seconds of movement as great as a billion megawatts.

The size of an earthquake is given by its moment: average slip times the fault area that slipped times the elastic constant of the Earth. The units of seismic moment are dyne-centimeters. An older measure of earthquake size is magnitude, which is proportional to the logarithm of moment. Magnitude 2.0 is about the smallest tremor that can be felt. Most destructive earthquakes are greater than magnitude 6; the largest shock known was the 1960 Chile earthquake, with a moment of 10^{30} dyne-centimeters (10^{23} newton-meters) or magnitude 9.5. It involved a fault 600 mi (1000 km) long slipping 30 ft (10 m). The magnitude scale is logarithmic, so that a magnitude 7 shock is about 30 times more energetic than one of magnitude 6, and 30×30, or 900 times, more energetic than one of magnitude 5. Because of this great increase in size with magnitude, only the largest events (greater than magnitude 8) significantly contribute to plate movements. The smaller events occur much more often but are almost incidental to the process.

The intensity of an earthquake is a measure of the severity of shaking and its attendant damage at a point on the surface of the Earth. The same earthquake may therefore have different intensities at different places. The intensity usually decreases away from the epicenter (the point on the surface directly above the onset of the earthquake), but its value depends on many factors and generally increases with moment. Intensity is usually higher in areas with thick alluvial cover or landfill than in areas of shallow soil or bare rock. Poor building construction leads to high intensity ratings because the damage to structures is high. Intensity is therefore more a measure of the earthquake's effect on humans than an innate property of the earthquake.

Effects. Many different effects are produced by earthquake shaking. Although the fault motion that produced the earthquake is sometimes observed at the surface, often other earth movements, such as landslides, are triggered by earthquakes. On rare occasions the ground has been observed to undulate in a wavelike manner, and cracks and fissures often form in soil. The flow of springs and rivers may be altered, and the compression of aquifers sometimes causes water to spout from the ground in fountains. Undersea earthquakes often generate very long-wavelength water waves, which are sometimes called tidal waves but are more properly called seismic sea waves, or tsunami. These waves, almost imperceptible in the open ocean, increase in height as they approach a coast and often inflict great damage to coastal cities and ports. *See* LANDSLIDE; TSUNAMI.

Prediction

Earthquake prediction research has been going on for nearly a century. A successful prediction, specifying the time, location, and magnitude of an

earthquake, would save lives and billions of dollars in housing and infrastructure costs. Unfortunately, successful earthquake predictions are extremely rare. There are two basic categories of earthquake predictions: forecasts (months to years in advance) and short-term predictions (hours or days in advance). Forecasts are based a variety of research, including the history of earthquakes in a specific region, the identification of fault characteristics (including length, depth, and segmentation), and the identification of strain accumulation. Data from these studies are used to provide rough estimates of earthquake sizes and recurrence intervals.

An example of an earthquake forecast is the identification of seismic gaps, portions of the plate boundaries that have not ruptured in a major earthquake for a long time. These regions are most likely to experience great earthquakes in the future. **Figure 4** shows seismic gaps for the circum-Pacific region and indicates which gaps were most likely to experience a large or great earthquake in 1989–1999. Large earthquakes did occur in several of the likeliest gaps, but many large earthquakes occurred in less likely gaps as well. Earthquake probability estimates are another example of forecasting. Geologic, geodetic, and seismic information are combined to estimate the frequencies of damaging earthquakes in a specific region. Recent regional earthquake probability estimate studies have resulted in forecasts of an 80–90% probability of a magnitude 7 or larger earthquake in the southern California region before 2024, and a 70% probability of a magnitude 6.7 or larger earthquake in the San Francisco Bay region before 2030.

Short-term earthquake prediction is still entirely in the realm of ongoing research, and no method is known to be reliable. Evidence is emerging that short-term prediction may be inherently impossible due to the complex and chaotic nature of the earthquake process.

Christopher H. Scholz; Kaye M. Shedlock

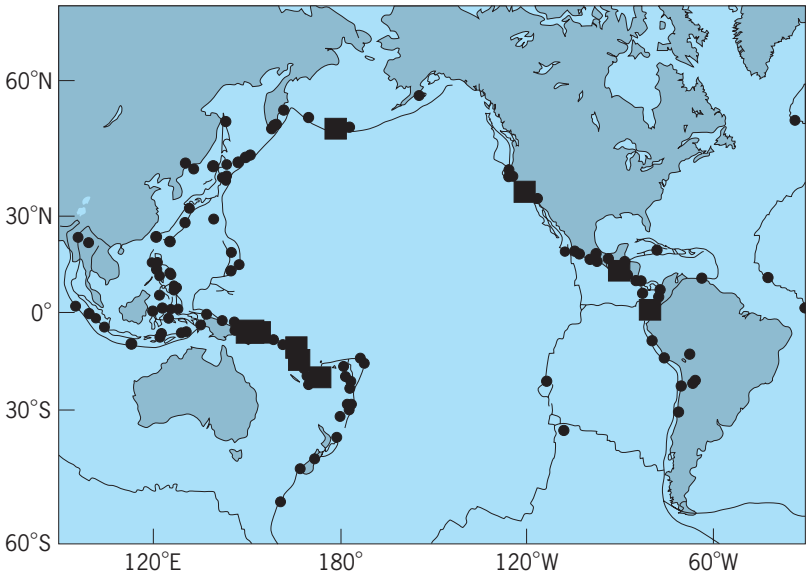

Fig. 4. Major seismic gaps, western Pacific. The plate boundaries are shown in black. Black squares mark the locations of seismic gaps with a 50% or greater chance of being filled by a large earthquake between 1989 and 1999. Black dots are magnitude 7.0 or larger earthquakes that occurred between 1989 and 1999. (*After S. P. Nishenko, Pure Appl. Geophys., 135:169–259, 1991*)

Deep Earthquakes

Most earthquakes occur at depths shallower than about 50 km (30 mi) [**Fig. 5***a*] and are usually found near plate boundaries. A few percent of all shocks occur at depths of 300–700 km (183–427 mi), depths that correspond to earth pressures of 100,000–250,00 atm (1–2.5 × 10^{10} Pa; Fig. 5*b* and **Fig. 6**). That the mantle can suddenly rupture rather than flow plastically at such conditions has elicited wonder since deep earthquakes were first discovered in the 1920s. Modern insight into these phenomena has come from scientific advances of plate tectonics, seismic tomography, and the mineral physics of the deep mantle based on very high pressure experiments on mantle minerals.

Most, if not all, deep earthquakes occur in inclined belts inside slabs, the cold, dense, and strong lithospheric plates that dive deeply into the Earth's mantle in places where plates are converging. Seismic waves have been used to image variations in the seismic wave speeds in the Earth. These anomalies in seismic tomographic images reflect differences in temperature, mineralogy, or composition. As expected, deep earthquakes occur in the high-wave-speed anomalies that mark cold slabs, anomalies that have been traced to depths as great as 2000 km (1220 mi) or more.

Curiously, earthquakes occur no deeper than 650–700 km (397–427 mi), far shallower than the maximum depths to which slabs descend. This abrupt shutoff and the gradual onset of the deep earthquake population at 300–350 km (183–214 mi) bracket approximately the transition zone of the mantle where seismic wave speeds abruptly increase (Fig. 6*a*). High-pressure experiments indicate that the mineralogy of the mantle changes at those depths and pressures from upper-mantle mineralogy (dominantly olivine and pyroxenes) to the minerals spinel, ilmenite, and majorite in the transition zone and, in turn, to the lower-mantle perovskite and oxide minerals. Slab mantle penetrating through the transition zone is expected to transform to these denser minerals.

Most deep earthquakes occur in the depth interval of the transition zone where upper-mantle slab minerals are reconstructed to their denser structural forms. Attention has therefore been drawn to the possibility that deep earthquakes are somehow caused by the mineralogical transformation of slabs as they descend into and through this region. Early speculation was that deep earthquakes represent rapid implosions that might occur when slab minerals transform suddenly to their denser, high-pressure forms. The patterns of seismic waves that radiate from deep earthquake sources indicate, however, that such disturbances represent slip on a fault, as do shallow earthquakes. If a connection exists between deep earthquakes and mantle phase changes,

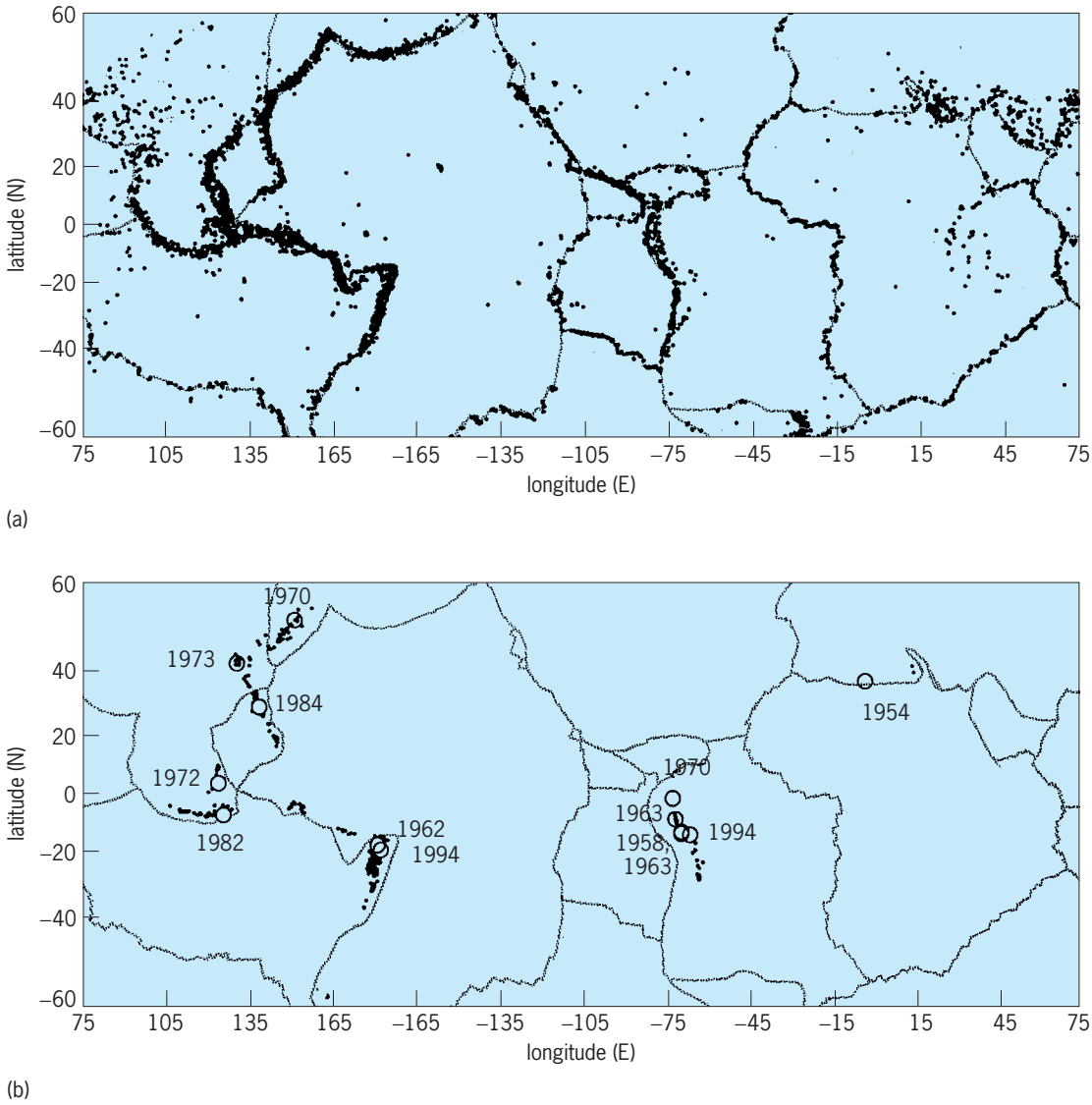

Fig. 5. Map showing the global distribution of shallow and deep earthquakes. (*a*) Earthquakes shallower than 40 km (25 mi). (*b*) Earthquakes deeper than 300 km (186 mi; open circles with dates indicate the locations of the 13 largest recorded deep shocks, including the 1994 Bolivian event).

the underlying process must facilitate failure by faulting.

A clue to the nature of this possible connection comes from the observation that deep earthquakes do not occur in all slabs, only in those that are very cold because they are descending at very fast rates. Low slab temperatures are important because such conditions favor the metastable persistence of upper-mantle minerals in the coldest interiors of slabs as they descend into the transition zone (Fig. 6*b*). Laboratory studies show that some minerals deformed under metastable conditions will rupture by an unusual shear instability in which the mineral is transformed to denser minerals in the shear zone. This shear instability, called transformational faulting, is not inhibited by high pressures and hence is an attractive candidate for the faulting mechanism of deep earthquakes. According to this theory, deep earthquakes do not occur in the lower mantle because low-density upper-mantle slab rocks are too buoyant to sink into the lower mantle.

An extraordinary demonstration of the potential scale of deep seismic faulting was demonstrated by the great deep earthquake that occurred at a depth of about 640 km (390 mi) beneath the Amazonian rainforest of Bolivia on June 9, 1994. The ground motions of this magnitude 8.2 shock, the largest deep event on record, were felt by people in places as far north as Toronto and Seattle, corresponding to surface distances from the earthquake focus exceeding 8000 km (4880 mi). Theoretical analysis of the seismograms written by this earthquake indicate that deep rupture occurred on a horizontal fault over an area of about 1600–2500 km^2 (618–965 mi^2) with a rupture duration of about 40–50 s and a shear offset of about 10–15 m (33–49 ft). This great, deep rupture occurred in the Nazca plate subducting beneath the South American plate, and was so energetic that it excited Earth's free oscillation modes and the planet rang like a low-frequency bell for many weeks after

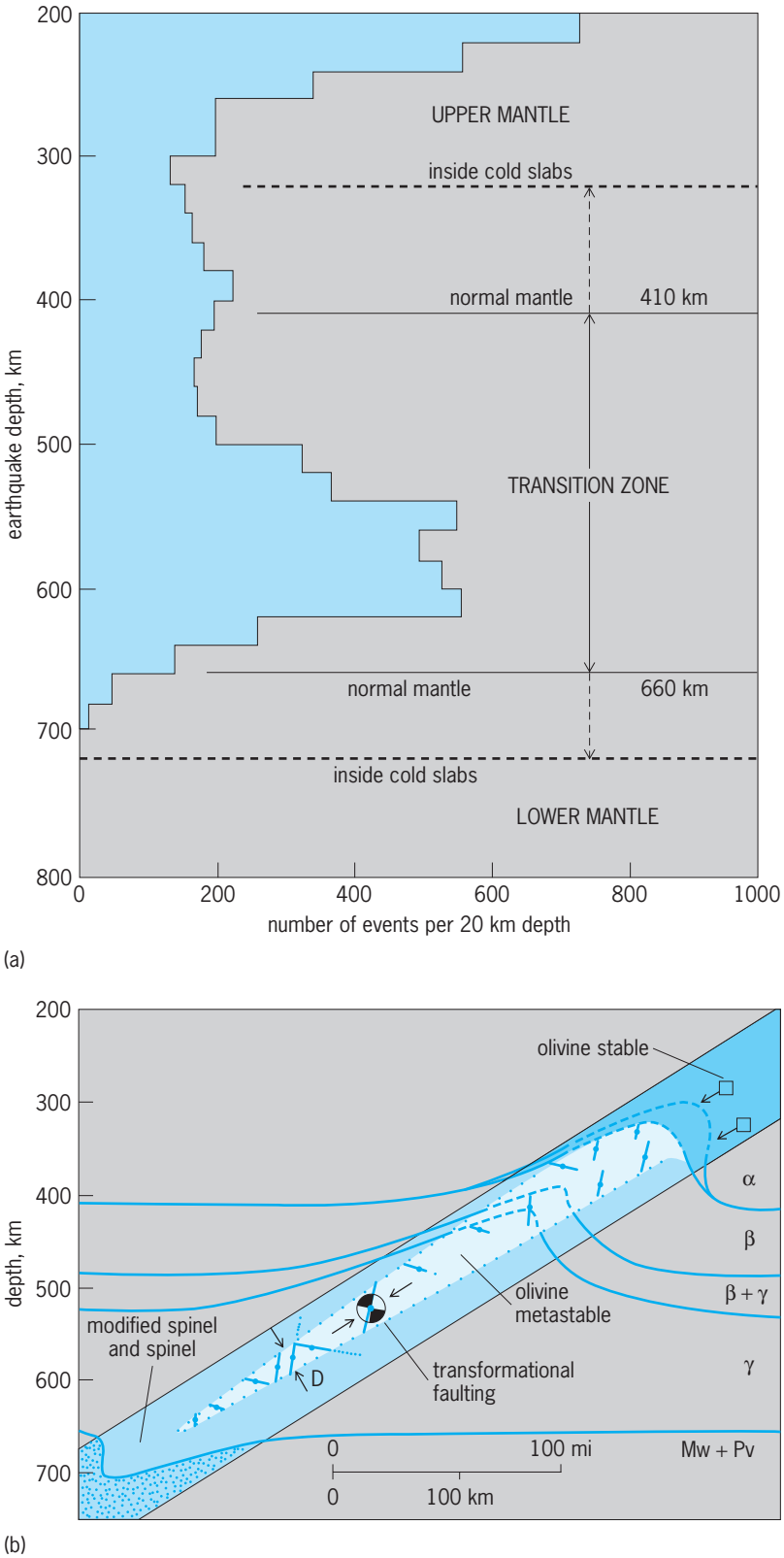

Fig. 6. Depth histogram of earthquakes compared to the mineralogical structure of the mantle. (*a*) Depth histogram of well-located earthquakes (1964–1991) in relationship to seismic-wave speed discontinuities caused by mineralogical changes in normal mantle and in cold slabs. (*b*) Hypothetical mineralogical structure of very cold slabs and normal mantle, emphasizing the phase changes associated with the olivine component [$(Mg,Fe)_2SiO_4$] of the mantle [α (olivine), β (modified spinel) and γ (spinel)]; Mw + Pv are magnesiowüstite and perovskite, the higher-pressure minerals dominating the lower mantle. Transformational faulting is a shear instability that can occur in metastable olivine under stress and can produce deep earthquakes.

the earthquake. The global occurrence of such earthquakes is the deepest direct expression of Earth's thermal convection system. *See* EARTH, CONVECTION IN.

Kaye M. Shedlock

Bibliography. G. L. Berlin, *Earthquakes and the Urban Environment*, 3 vols., 1980; B. Bolt, *Earthquakes and Geological Discovery*, 1993; *Earthquake Information Bulletin*, U.S. Geological Survey, Reston, Virginia, six times yearly; E. R. Engdahl, R. van der Hilst, and R. Buland, Global teleseismic earthquake relocation with improved travel times and procedures for depth determination, *Bull. Seis. Soc. Amer.*, 88:722–743, 1998; *Facing Geological Hazards*, U.S. Geol. Surv. Prof. Pap. 1240-B, 1981; C. Frohlich, Deep earthquakes, *Sci. Amer.*, 260(1):48–55, 1989; S. Kirby et al., Metastable mantle phase transformations and deep earthquakes in subducting oceanic lithosphere, *Rev. Geophys.*, 34:261–306, 1996; T. Lay, *Structure and Fate of Subducting Slabs*, Academic Press, San Diego, 1994; C. Lomnitz, *Fundamentals of Earthquake Prediction*, 1994; T. Rikitaki, *Earthquake Forecasting*, 1982; K. M. Shedlock and L. C. Pakiser, *Earthquakes*, U.S. Geol. Surv. Gen. Int. Publ., 1994; A. Villaseñor et al., Toward a comprehensive catalog of global historical seismicity, *Eos Trans. Amer. Geophys. Union*, 78:581, 583, 588, 1997.

Earthquake engineering

The branch of engineering concerned with reducing earthquake or seismic risk to structures. Because strong earthquakes are rare events, building codes have traditionally allowed a significant degree of damage. Even high-seismic regions, such as San Francisco or Tokyo, typically experience a strong earthquake only once in many decades. If an earthquake occurs, buildings and other structures are designed such that most will be damaged (but should not collapse), and will have costs for repairs, business interruption, and potentially casualties. While building collapse is the primary cause of loss of life in most earthquakes, other contributors to earthquake loss include equipment and contents damage, business interruption, and damage to lifelines, such as water, power, gas, communications, and transportation. To limit these losses to acceptable levels, earthquake engineering involves a process of (1) seismic hazard identification, (2) structural analysis, design, and/or retrofitting to prevent structural collapse and reduce property damage, and (3) review of equipment and operations to prevent disruption due to earthquakes—that is, an integrated, comprehensive program of facility seismic review, analysis, retrofit, emergency planning, and risk transfer, involving the expertise of mechanical engineers, operations specialists, emergency planners, and insurers, in addition to geoscientists and structural engineers.

Hazards. Most earthquakes are caused by the fracture and sliding of portions of the Earth's crust along faults, which may be hundreds of miles long, from 1 mi (1.6 km) to over 100 mi (160 km) deep, and

sometimes not readily apparent on the ground surface. Earthquakes can occur anywhere on Earth, but most occur along major tectonic plate boundaries ("Ring of Fire"), especially on the circum-Pacific plate boundary, the Caribbean, and the Trans-Alpide belt (stretching from southern France through the Mediterranean and the Middle East, along the Himalayan foothills and the Indonesian archipelago). In the United States, 39 of the 50 states are considered at moderate to high seismic risk, with major earthquake potential in the western states, Alaska, the central United States (St. Louis–Memphis region), and portions of the east coast (South Carolina, Massachusetts). *See* EARTHQUAKE; SEISMIC RISK.

Earthquakes can cause significant damage to the built environment due to fault rupture, vibratory ground motion (shaking), inundation (tsunami, seiche, dam failure), various kinds of permanent ground failure (liquefaction, landslide), and fire or hazardous materials release. In a particular event, any of these hazards can dominate, and historically each has caused major damage and great loss of life. The expected damage given a specified value of a hazard parameter is called vulnerability, and the product of the hazard and the vulnerability is the expected loss or seismic risk.

For most earthquakes, shaking is the dominant and most widespread cause of damage. Shaking near the actual earthquake rupture lasts only during the time when the fault ruptures, a process which takes seconds or at most a few minutes. The seismic waves generated by the rupture propagate long after the movement on the fault has stopped, spanning the globe in about 20 min. Typically earthquake ground motions are powerful enough to cause damage only in the near field (that is, within a few tens of kilometers from the causative fault). In a few instances, long period motions have caused significant damage at great distances to selected structures. A prime example was the 1985 Mexico City earthquake, where a magnitude 8.1 earthquake occurring at a distance of approximately 400 km from Mexico City caused numerous collapses of mid- and high-rise buildings.

The most important scales for measuring earthquake magnitude are surface-wave magnitude M_s, body-wave magnitude m_b, and moment magnitude M_W (the last is the more commonly preferred scale today). Magnitude can be related to the total energy release by $\log_{10} E = 11.8 + 1.5M_S$, where E is the total energy in ergs. Since $10^{1.5} = 31.6$, an increase of one magnitude unit is equivalent to 31.6 times more total energy release, an increase of two magnitude units ~1000 times more total energy, and so forth. Whereas magnitude is a measure of the overall size of a single earthquake, intensity is a measure of the effect, or the strength, of an earthquake hazard at a specific location. Intensity scales in use include the Modified Mercalli Intensity (MMI) in the United States, Medvedev-Sponheur-Karnik (MSK-81) in Europe, and the Japan Meteorological Agency (JMA) scale. Roman numerals are traditionally used for intensity scales to indicate the qualitative nature of the scales, which are based on subjective observations rather than instrumental records. For the MMI and MSK scales, 0 is no earthquake, VI is the initiation of damage to poor-to-average structures, and XII is total destruction. Earthquake shaking at a site is recorded on a seismometer, which digitally or optically records a time history of ground accelerations at the site.

Fig. 1. Relatively new reinforced concrete buildings that collapsed in the Marmara, Turkey, earthquake, August 17, 1999.

Statistical analysis of hundreds of such accelerograms provide attenuation regressions, which permit estimation of the maximum acceleration, velocity, or displacement that a structure will experience at a site as a function of the structure's natural period, the site-specific soil properties, and the magnitude, distance, and depth of the hypothesized earthquake. In the United States, ground motions due to all potential earthquakes have been synthesized by the U.S. Geological Service, so that any site's seismic hazard (based on an average soil type) can be quickly determined. Based on this hazard, or on building code requirements mapped at a regional scale, the seismic lateral force requirements against which a structure must be designed are determined. *See* SEISMOLOGY.

Structural analysis and design. In the developed world, existing buildings and infrastructure constitute by far the preponderance of earthquake risk, so that a major focus of earthquake engineering is on identification, analysis, and mitigation (reduction) of this risk. New construction is generally safer, but is still a focus for earthquake engineers, especially larger or unusual structures. In the developing world, even recent and new construction is a significant contributor to seismic risk, due to lack of building code enforcement (**Fig. 1**). The most seismically hazardous existing building structures are low-strength masonry (such as adobe in developing countries), unreinforced brick masonry (such as exists in older portions of United States cities), nonductile reinforced concrete (typically, concrete frames constructed prior to the 1980s), and certain kinds of precast concrete buildings. Steel and wood structures are not immune to earthquake damage, but typically are less collapse-prone. However, the tallest high-rise building to collapse, for any reason, was the 23-story Piño Suarez steel-framed building in the 1985 Mexico City earthquake. Building vulnerability is typically mitigated via structural retrofits involving strengthening wall-diaphragm connections, adding shear walls or bracing, and improving ductility of columns by steel jacketing or other methods. A diaphragm is a horizontal force-distributing element, such as a roof or floor.

Earthquake analysis and design of a structure is commonly performed assuming the structure remains elastic (that is, the design neglects the effects of damage), though inelastic analyses (which takes into account material inelasticity) are increasingly used. For ordinary design, the actual earthquake forces are reduced by a response modification factor that varies depending on the assumed inherent ductility of the structure, based on its lateral force-resisting system (moment frame, braced, shear wall) and material (wood, steel, concrete). This force reduction, along with appropriate safety factors, is combined with gravity and other loads in the design of the overall structure. Pseudo-static or linear dynamic analytical methods are used to determine structural member forces. Pseudo-static methods are based on the structure's natural period (first mode of vibration) and are appropriate only for simple structures. Linear dynamic methods account for the first and higher modes, which analyses are usually more quickly performed in the frequency domain, using techniques based on the fast Fourier transform. For larger and more important structures, nonlinear dynamic analyses are used, typically in the time domain. In a nonlinear dynamic analysis, the structure is subjected to earthquake acceleration time histories (actual or synthesized records, scaled to match the site hazard), and member response into the inelastic range is taken into account, including P-Δ effects (that is, the increase in overturning moment due to the structure's weight P times its lateral deflection Δ). Dynamic analysis has become common due to the advent of more powerful computers and specialized software—ETABS, SAP2000, ANSYS, STAAD, and LARSA are some of the structural analysis packages commonly used today. *See* BUILDINGS; LOADS, DYNAMIC; STRUCTURAL ANALYSIS.

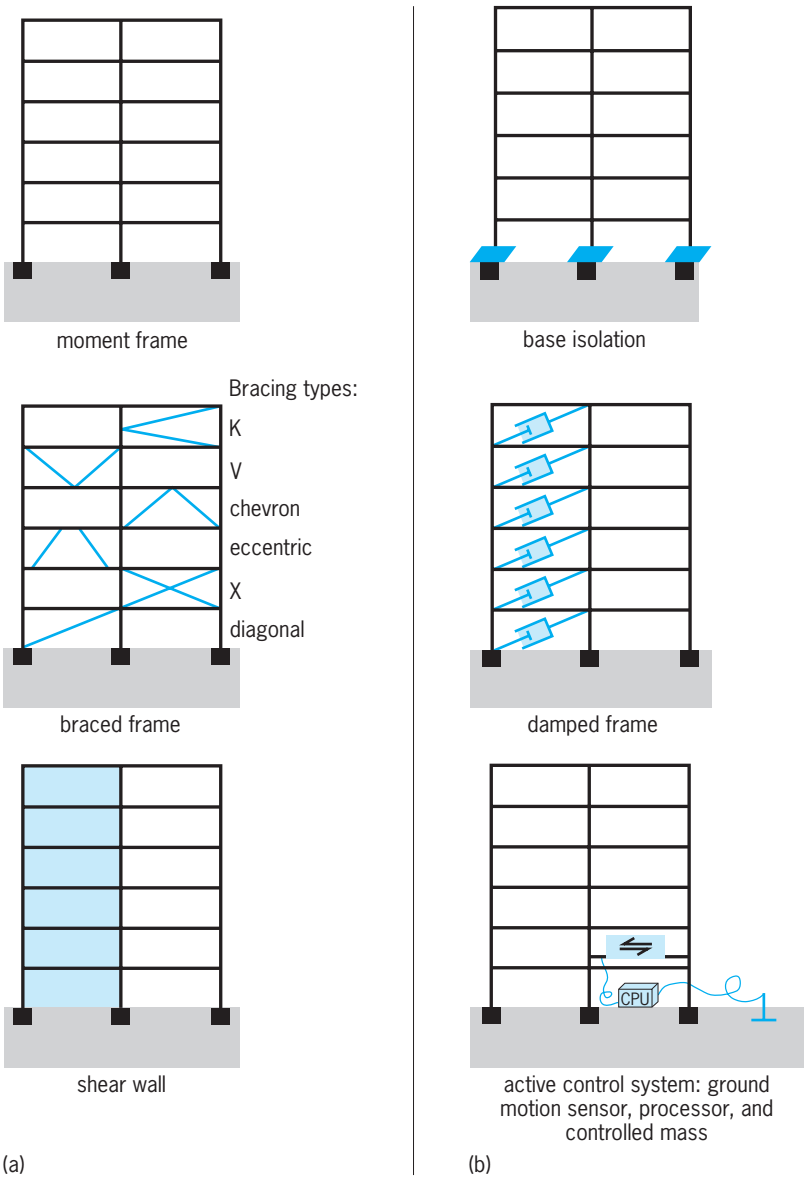

Fig. 2. Structural designs: (*a*) Traditional earthquake force-resisting systems. (*b*) Emerging technologies for earthquake force-reducing systems.

New design approaches. Traditionally, engineers have designed structures to resist earthquake forces, intending the strength of structural members to be equal to or greater than the seismic demand placed on them (**Fig. 2**). During the 1980s and 1990s, new approaches to seismic design emerged, which involved modifying the structural response to reduce earthquake loads to more tolerable levels. These included base isolation, supplemental damping, and active control. Base isolation involves placing special components, termed isolators (not always at the base, so that the technique is more properly termed structural isolation), within the structure that are relatively flexible in the lateral direction, yet can sustain the vertical load. When the earthquake causes ground motions beneath the structure, the isolators allow the structure to respond more slowly than it would without them, resulting in lower seismic demand on the structure. Isolators may be laminated steel with high-quality rubber pads, sometimes incorporating lead or other energy-absorbing materials, or parabolic dish-shaped base plates, which rely on the structure's own weight trying to "climb" the sloping sides of the "dish" to counteract the lateral force of the earthquake. Supplemental damping involves placing dampers within the structure, which retard the structural response of the normal or lightly damped structure, again resulting in lower seismic demand on the structure. Active control involves placing hydraulic rams or other actuating devices within the structure, which introduce forces counter to those caused by the earthquake, thus negating some or all of the seismic demand. To determine how much force to apply to an actuator involves a real-time solution of the dynamics of the structure. Sensors must measure the ground motion at the base of the structure, a structural analysis for these accelerations must be performed on a dedicated processor, and the optimum pattern of forces must be determined and applied via quick-response actuators during a time interval equal to or less than the dynamic response of the structure. Base isolation and supplemental damping are currently entering the mainstream of structural engineering, while active control is still largely the subject of research.

The most recent development in earthquake engineering is performance-based design, in which the expected structural damage due to the maximum expected earthquake is quantified. If the expected damage is unacceptably high, the owner and the engineer may agree on a design in excess of building code requirements. *See* ARCHITECTURAL ENGINEERING; CIVIL ENGINEERING. Charles Scawthorn

Bibliography. B. A. Bolt, *Earthquakes*, W. H. Freeman, 5th ed., 2003; W. F. Chen (ed.), *Handbook of Structural Engineering*, 2d ed., CRC Press, 2005; A. K. Chopra, *Dynamics of Structures*, 2d ed., 2000; S. L. Kramer, *Geotechnical Earthquake Engineering*, Prentice Hall, 1995; C. H. Scholz, *The Mechanics of Earthquakes and Faulting*, 2d ed., Cambridge University Press,, 2002; P. I. Yanev, *Peace of Mind in Earthquake Country*, Chronicle Books, San Francisco, 1991.

East Indies

A loosely defined region in southeast Asia broadly comprising former colonial possessions of European nations and now forming the countries of Malaysia, Brunei, and Indonesia. The islands of the East Indies extend for about 2800 mi (4500 km) from western Sumatra to New Guinea. They form part of a region of great geological and biological diversity.

Geography. Three broad geographic subdivisions can be made (see **illus.**): (1) Sundaland, which comprises the shallow marine Sunda Shelf, peninsular southeast Asia, and the islands west of the Makassar and Lombok straits; (2) northern Australia, which includes the shallow marine Sahul and Arafura shelf areas and islands surrounding northern Australia and New Guinea; and (3) Wallacea, which comprises the islands south of the Philippines between Sundaland and northern Australia. Each division is an important geomorphological and biological region. The boundaries of Sundaland and northern Australia can be drawn at the 660-ft (200-m) bathymetric contour. Deep trenches with depths averaging 5 mi (8 km) border the margins of Sundaland with the Indian Ocean, and northern Australia with the Pacific Ocean. In contrast, within the area of Wallacea there are a number of deep troughs, and there are large variations in relief with volcanic and nonvolcanic mountains, typically up to 6600–9900 ft (2000–3000 m) in height, separated by deep marine basins underlain by extended continental and oceanic crust with depths of several kilometers. *See* ASIA; AUSTRALIA.

Landscape. Sundaland is very flat with a mean elevation only slightly above sea level. A shallow shelf currently lies beneath the sea between Borneo, Java, Sumatra, and the Malay Peninsula, and has been intermittently emergent in the past, notably during Pleistocene glaciations. Highlands are found in the Malay Peninsula, Sumatra, Java, and Borneo (comprising Brunei, the Malaysian states of Sabah and Sarawak, and Indonesian Kalimantan). The central part of the Malay Peninsula has a western range of granite highlands up to 7260 ft (2200 m) which provided the once important tin deposits of Malaysia. Further east are lower peaks, again mainly underlain by intrusive granitic rocks. In Sumatra, the narrow Barisan mountain range runs the length of the island on its western side, within which is the major active Sumatran strike-slip fault zone, and on which is built a chain of volcanoes with heights up to 12,550 ft (3800 m). Most of eastern Sumatra is swampy alluvial lowland. Java is one of the richest agricultural islands of the tropics. There are more than 40 young volcanoes with heights up to almost 12,200 ft (3700 m) forming a continuous spine running the length of the island. In contrast to Sumatra and Java, Borneo has a broad low central mountain range running from southwest to northeast which becomes higher toward the northeast. The border of Kalimantan with Sarawak and Sabah essentially follows the drainage divide, and there are a dozen or so peaks above 6600 ft (2000 m), of which the highest is Mount Kinabalu

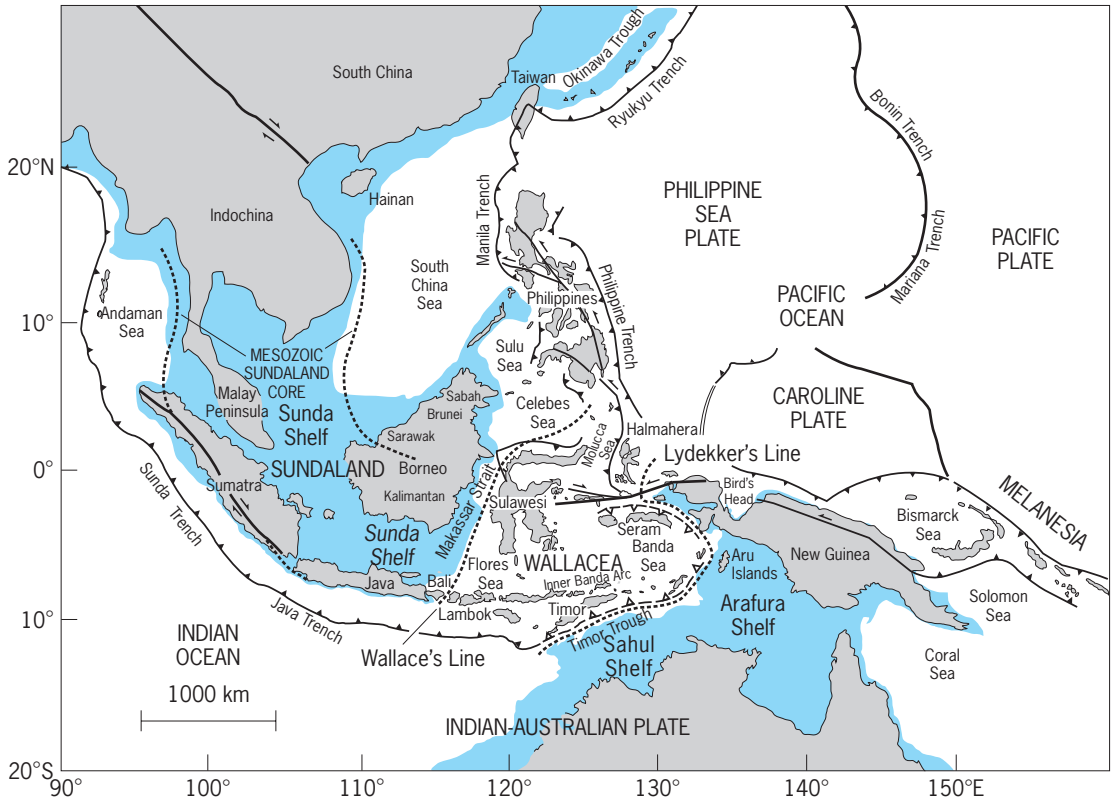

East Indies. The colored areas represent the shallow continental shelves of Sundaland and northern Australia, drawn at the 660-ft (200-m) isobath. The western boundary of Wallacea is Wallace's Line of 1863. The eastern boundary is Lydekker's Line, which is the western limit of strictly Australian fauna. Major tectonic plates are named. The Asian plate is a composite of several smaller plates.

in Sabah (13,525 ft or 4098 m). Kinabalu is formed of young granite, testifying to very rapid and recent uplift, and its glacial landscape at the summit shows that it was the site of a Pleistocene icecap. Nearly half of Borneo is lowlands under 660 ft (200 m), which are frequently swampy. In Borneo the largest rivers of Sundaland drain west and north into the Makassar Strait, and the South China, Sulu, and Celebes seas, where they deposited immense volumes of sediment in late Cenozoic deltas which are very rich in oil and gas.

Like Sundaland, northern Australia has a mean elevation close to sea level except in New Guinea, the world's second largest island, which is nearly 1350 mi (2200 km) long. The high mountains of the central part of the island are a Tertiary fold and thrust belt rising to more than 16,500 ft (5000 m). Many of them are snow-capped, even though they are within 5° of the Equator, and display the effects of Pleistocene glaciation. North of the central mountains there is a broad valley with a lower coastal range to its north. The rapid uplift of the mountains and the heavy tropical rainfall mean that there are many fast-flowing rivers. The rivers have very high sediment yields, and those of New Guinea are estimated to carry as much sediment in total as all those draining North America.

Within Wallacea are numerous small rugged islands, all of which have risen above sea level in the last few million years. The largest is Sulawesi, in which most land is above 1485 ft (450 m) and where very young central mountains rise to over 11,000 ft (3400 m). Numerous lakes occupy depressions within young strike-slip fault zones; Lake Poso is 4950 ft (1500 m) deep. Timor and Seram, although smaller, are similarly mountainous, with many peaks above 5950 ft (1800 m) and maximum heights on both islands close to 9900 ft (3000 m).

Climate and ocean circulation. At present, almost all the islands lie in a belt close to the Equator within the Intertropical Convergence Zone (ITCZ). An equatorial climate prevails, with high rainfall and, except at higher elevations, high temperatures throughout the year. Diurnal variations are greater than the difference of mean temperatures of the hottest and coldest months. High relative humidity is normal in most lowland regions. Regionally, there are significant variations in rainfall, reflecting topography and position with respect to major landmasses and oceans, and each island's climate can be different. Borneo is the only large island within which there is broadly an ever-wet tropical climate. Elsewhere, monsoons, reflecting movement of air masses between Asia, Australia, and the surrounding oceans, control seasonal changes in winds and variations in precipitation. A tropical climate with typically more than 120 in. (3000 mm) of rainfall each year prevails in Sumatra, western Java, Borneo, Sulawesi, and the Moluccas; but extended dry seasons occur in eastern Java and the Lesser Sunda Islands. Climatic and

topographic variation is reflected in a considerable variety in vegetation, from tropical rainforest to savanna, now further altered by human activity. *See* CLIMATOLOGY; TROPICAL METEOROLOGY.

The shallow seas with narrow deep-water passages of Wallacea mean that the region is particularly important for oceanic circulation. The Indonesian throughflow is a current to the Indian Ocean which transports large amounts of warm water from the Pacific, influencing sea surface temperatures, salinity, and rainfall. The magnitude and variations of this current are important controls on the thermohaline balance of the Pacific and Indian oceans, and perhaps on global thermohaline circulation. Most water passes from the North Pacific via the Celebes Sea, Makassar Strait, Flores Sea, and Banda Sea. Less important volumes of water move from the South Pacific east of Halmahera into the Banda Sea. Changes in the relative importance of these currents are probably seasonal and in the longer term may vary with El Niño. The importance of the present throughflow suggests that the convergence of Australia and Asia in the last few million years may have influenced local and possibly global climate, as the Pacific-Indian seaway has become increasingly restricted. *See* EQUATORIAL CURRENTS; OCEAN CIRCULATION.

Geological history. The East Indies are characterized by intensely active seismicity and volcanic activity. The correlation of seismicity, volcanicity, deep trenches, and strong negative gravity anomalies along the Sunda and Banda arcs was noted long before the formulation of the theory of plate tectonics; these are now known to be features characteristic of the subduction of oceanic lithosphere. The history of convergence of the Pacific, Indian-Australian, and Asian plates offers a broad explanation of the geological development and complexity of the region, but many small plates also need to be considered. However, the full details of the development of the region are still far from understood because of its size, relative inaccessibility, and the nature of the terrain. *See* CONTINENTAL DRIFT; GEODYNAMICS; PLATE TECTONICS.

In general, the region owes its origin to the pre-Cenozoic breakup of the Gondwana supercontinent, the subsequent movement of Gondwana fragments northward, and their eventual collision with Asia. Many fragments separated from Gondwana and amalgamated in southeast Asia over a considerable period of time. The process of rifting led to formation of new oceans, and the northward motion of Gondwana fragments required subduction of older oceanic crust at the edges of the growing Asian continent. By the late Mesozoic, fragments derived from Gondwana formed a composite Sundaland core surrounded by subduction zones.

Two major fragments, India and Australia, separated from Gondwana in the Cretaceous and moved northward as parts of different plates. India initially collided with the Asian continent about 50 million years ago, but continued to move north accompanied by complex internal deformation within Indochina and mainland southeast Asia, which continues to the present day. To the east of India, long-term subduction of oceanic lithosphere of the Indian plate produced the arcuate system of volcanoes and seismicity which extends through Sumatra and Java. Australia moved more slowly northward and during the last 25 million years has made an indirect collision with a composite southeast Asia which includes some of the Gondwana fragments that arrived earlier, and also the island arcs formed due to the subduction of oceanic crust north of Australia. Further east, arc-continent collisions have resulted from elimination of small oceanic basins formed above subduction zones north of Australia; this system of arcs and oceanic basins can be traced east along the Australian margin into Melanesia. Fragments of these arcs have been dispersed in New Guinea, east Indonesia, and the Philippines by the movement of the Pacific plate.

In east Indonesia the northward movement of Australia during the Cenozoic has been marked by arc-continent collision, major strike-slip motion within the north Australian margin in northern New Guinea, and collision of continental fragments derived from Australia. The Banda arc exemplifies many of the complexities of the region. There is the unusual 180° curvature of the arc. Deformation is very young and still active, and high mountains on land are separated by deep ocean basins. On its north side, Australian fragments are moving westward by strike-slip faulting. On its south side, young mountains in Timor contain deep-sea Australian margin sediments uplifted to elevations of more than 6600 ft (2000 m) within the last 3 million years following arc-continent collision. At the center of the arc is the Banda Sea, which is a young oceanic basin produced by very rapid extension, probably driven by subduction. The Banda region offers a snapshot of the processes which have been active in the entire region for many tens of millions of years. As a result, the region is geologically very complex and includes fragments of Gondwana continental origin, volcanic island arcs, and oceanic crust, all of which have been deformed by collision with the Sundaland core and then by subsequent collisions which have continued to the present day. However, despite the long-term overall convergence of the Asian, Australian, and Pacific plates, there has been frequent opening of small oceanic basins as well as extension related to strike-slip movements at plate boundaries within the region; the Banda Sea is one such basin. *See* MARINE GEOLOGY.

Plants and animals. Climatic changes during the past 20 million years have significantly influenced the distribution of many plants and animals in the region. The northward motion of Australia caused northern Australia and New Guinea to enter the tropical belt during this period. Collisions in eastern Indonesia closed ocean seaways and altered circulation between the Pacific and Indian oceans. At the same time, atmospheric circulation was affected by the rise of the Himalayas and the later rise of mountains in Borneo, Sulawesi, and New Guinea. About 20 million years ago, tropical ever-wet climates became established in much of the region, followed by a change

to monsoonal climates in many areas about 5 million years ago. More recently, during the Pleistocene, sea levels fluctuated dramatically and rapidly as a result of glacial cycles, which also caused considerable variations in rainfall and temperature with consequent effects on plant and animal life. Shallow parts of the continental shelves were exposed when sea levels were low, and land bridges formed between islands. *See* PLEISTOCENE.

Today, the waters of southeast Asia contain the highest marine faunal diversity in the world, and the islands of the East Indies contain some of the most diverse collections of plant and animal species. Tropical forests support more species than any other terrestrial ecosystem, and southeast Asia therefore includes some of the most important biological resources of the planet as well as a complex biogeography reflecting the many factors that have influenced the present distributions of plants and animals. The East Indies have been a critical region in the development of ideas concerning biogeographic patterns and their causes. Alfred Russel Wallace was the first to recognize a division between Asiatic and Australian floras and faunas within the region, and in 1863 he drew a boundary between the islands of Bali and Lombok and farther to the north between Borneo and Sulawesi, at the Makassar Strait. The site of this boundary, now known as Wallace's Line, and its significance for different organisms, has been the subject of much discussion since it was first proposed, and indeed Wallace himself proposed different positions for the boundary at different times. Later authors drew boundaries in other places; and it is now recognized that there is a biogeographic region of transition, named Wallacea, situated between areas with Asiatic and Australian floras and faunas, with elements of both but where organisms show a high degree of endemism. This region extends from east of the Makassar Strait to west of the Bird's Head of New Guinea. *See* BIOGEOGRAPHY.

The origin of the biogeographic patterns is complex. There was dispersal of species throughout the region as fragments of Gondwana collided with Asia and as new land and oceanographic connections were established when sea level changed. Vicariance (fragmentation of ancestral species by newly created physical or ecological barriers) resulted from the rise of mountains, formation of deep marine basins, changes in distribution of land and sea as sea levels changed, and modification of habitats as climates changed. The tropics are regions of immense genetic diversity, and there was also rapid speciation as new habitats were established due to emergence of new landmasses. New Guinea is one example of such a pool of diversity. Mountains have risen rapidly in northern New Guinea in the last few million years as the result of collision of Pacific island arcs with the north Australian margin. As a result, climate changed, new habitats were created, and as new barriers separated some species, new pathways developed for the dispersal of others. Many animals and plants of Australian and Pacific origin speciated rapidly in this newly emergent region.

In addition, the tectonic convergence in the region provided new connections between Australia and Asia significantly modified by global climatic changes in the last million years. Glacially related lowering of global sea level resulted in emergence of the Sunda Shelf, and temporary land bridges joined present-day Sumatra, Java, Bali, and Borneo to mainland Asia, allowing the exchange of animals and plants. Major parts of the Sahul Shelf and Arafura Shelf were also emergent, allowing dispersal from Australia into New Guinea and surrounding islands. During interglacial periods the climate warmed again, sea level rose, and islands were once again separated. This process was repeated several times. However, many islands of Wallacea remained isolated by deep water which formed an effective barrier to the migration of most plants and animals between Asia and Australia. Local fluctuations in rainfall and temperature, and consequently vegetation, enhanced the effects of oscillations of dispersal and vicariance and left pockets of isolated fauna and flora. Thus, the plants and animals of Sundaland remain closely related to those of Asia, whereas those of northern Australia, New Guinea, and surrounding islands are predominantly Australian with Pacific influences. Wallace's Line should not be expected to be a sharp single boundary for all organisms at one place. The biogeographic region of Wallacea is a moving interface between Asian and Australian flora and fauna reflecting the establishment of new migration pathways about 20 million years ago. These new pathways were due to collision, the barriers to complete intermingling due to deep marine basins, the effects of glacial climatic changes, differences in abilities of different organisms to disperse, and complications owing to possible remnants of Gondwana elements derived from both Asian and Australian sides.

Resources. The entire region is immensely rich in natural resources, in particular petroleum, minerals, and timber. Petroleum has been produced in large quantities on land in Sumatra, Borneo, and the Bird's Head of New Guinea since the mid-twentieth century. Most oil and gas provinces are in Cenozoic basins, and in the late 1990s the East Indies provided about 5% of annual world oil production. In recent years, exploration and production has moved offshore and is increasingly moving into deeper waters. Many parts of the region have considerable potential for geothermal energy production. Mineral production has been historically important, with major discoveries of tin, gold, and nickel. The late-twentieth-century discovery of major mineral deposits in the northern New Guinea margin suggests that the young island arcs of the region will continue to be targets for exploration, and large deposits are likely to be found both on and off shore. *See* EARTH RESOURCE PATTERNS.

Robert Hall

Bibliography. *The Ecology of Indonesia Series*, vols. 1–8, Dalhousie University, Canada, Periplus Editions (Hong Kong) Ltd., 1987–1997; R. Hall and D. J. Blundell (eds.), Tectonic Evolution of SE Asia, *Geol. Soc. London Spec. Publ.*, no. 106, 1996; R. Hall and J. D. Holloway (eds.), *Biogeography and Geological*

Evolution of SE Asia, Backhuys Publishers, Leiden, 1998; W. Hamilton, Tectonics of the Indonesian Region, *U.S. Geol. Surv. Prof. Pap.*, no. 1078, 1979.

Eating disorders

Eating disorders are characterized by abnormal eating behaviors and beliefs about eating, weight, and body shape. The two formal diagnoses are anorexia nervosa and bulimia nervosa. In addition, there are many cases of abnormal eating that have only some of the features required for an eating disorder diagnosis; these cases are classified as eating disorders not otherwise specified. Another eating disorder diagnosis has been proposed, termed binge eating disorder. Obesity is classified as a general medical condition and not as an eating disorder (a psychiatric condition) because it is not consistently associated with psychological or behavioral problems.

There are also three childhood eating disorders: (1) Pica is a persistent pattern of eating nonnutritive substances in infants or young children. (2) Rumination disorder involves repeated regurgitation and rechewing of food. This behavior is not the result of a gastrointestinal or medical condition; the partially digested food comes back into the mouth without any observable nausea, disgust, or attempt to vomit. (3) Feeding disorder of infancy or early childhood is the persistent failure to eat adequately, as reflected in insufficient weight gain for age. Pica and rumination disorder are thought to be uncommon and frequently associated with developmental delays and mental retardation. Perhaps half of the pediatric hospitalizations for inadequate weight gain (which constitute 1–5% of all pediatric hospitalizations) may be due to feeding disorder of infancy or early childhood.

Anorexia nervosa. Anorexia nervosa is characterized by a refusal to maintain a minimal normal body weight (defined as 15% below average weight for height), an intense fear of becoming fat, and, if female, amenorrhea for at least 3 months. The majority of cases of anorexia nervosa are classified as restricting type; these individuals achieve abnormally low weight by severely dieting, fasting, and often by exercising compulsively. In severe cases, patients refuse to eat and can die of starvation or severe medical complications. Another subtype of anorexia nervosa is binge eating/purging type; in addition to the low weight, binge eating and purging through vomiting or the misuse of laxatives, enemas, or diuretics occur. Despite being emaciated or dangerously thin, persons with anorexia nervosa perceive themselves as overweight (distorted body image), deny the seriousness of their condition, and have an intense fear of becoming fat. This fear often increases even as weight continues to drop. Self-evaluation is dominated by the importance of weight and body shape.

Epidemiology and associated features. Anorexia nervosa occurs in roughly 1% of adolescent and young adult females. Most cases (90%) are female, and the majority are Caucasian and come from middle-class or higher socioeconomic groups. This disorder is more prevalent in industrialized countries that share western views regarding thinness as an ideal. It develops most frequently during adolescence; mean age of onset is 17 years with bimodal peaks at ages 14 and 18. The prevalence of anorexia nervosa has remained constant over the past few decades; one notable change is the higher incidence in women over age 30, although this still represents a minority of cases.

Persons with anorexia nervosa frequently manifest symptoms of depression (for example, depressed mood, loss of interest, social withdrawal, insomnia, difficulty in concentrating) and anxiety. It is sometimes difficult to ascertain whether these symptoms reflect secondary effects of the starvation or the coexistence of psychiatric disorders. The restricting type of anorexia nervosa is associated with obsessionality, rigidity, perfectionism, and overcontrol, whereas the binge/purge subtype is associated with greater mood instability and impulsivity across a wide range of areas, including substance abuse.

Although some cases of anorexia nervosa show no evidence of medical problems, prolonged starvation affects most organ systems, and a wide array of medical problems tend to be present. Frequent medical findings include leukopenia (decrease in white blood cells) and anemia, dehydration, hypercholesterolemia, abnormal liver and thyroid functions, electrolyte imbalances, low estrogen levels (females), osteoporosis, and cardiovascular problems (sinus bradycardia, severe hypotension, and arrhythmias). Anorexia nervosa tends to be a chronic refractory condition and is associated with considerable morbidity and mortality. Long-term mortality from anorexia nervosa is estimated at 5–10% with most deaths resulting from starvation, cardiac events, or suicide.

Etiology and maintenance. The causes of anorexia nervosa are not yet understood but are likely to involve a complex combination of genetic, familial, psychological, and sociocultural factors. There is an increased risk of anorexia nervosa among first-degree relatives of persons with the disorder and perhaps with depression. Individuals with certain psychological styles (low self-esteem, perfectionism, obsessiveness) and from families with certain characteristics (rigidity, overinvolvement, poor communication, high achievement demands) may be particularly susceptible to western society's pervasive emphasis on thinness as the ideal. The onset of anorexia nervosa tends to follow a period of dieting and is frequently triggered by a stressful life event or transition.

Treatment. Since the starvation and weight loss can be life-threatening, initial treatment efforts need to focus on weight gain and the reestablishment of regular eating patterns. Inpatient hospitalization is frequently necessary to prevent continued weight loss, to stabilize the weight and any medical comorbidities, and to initiate the nutritional rehabilitation. The initial period of weight stabilization and nutritional rehabilitation is typically met by the patient with great resistance and emotional upset. Consistency and structure during this period are essential

for the patient's success, and family therapy is generally helpful to family members.

Although significant psychological issues tend to be present, it is generally ineffective to address these until weight has been stabilized. The effects of starvation on psychological functioning coupled with the obsessional preoccupation with food and weight preclude any meaningful psychotherapy. Once weight gain is achieved, psychotherapies can become useful. Available evidence favors the use of cognitive-behavioral therapy and family-based therapy. Effective medication treatments have not been found. Relapse rates are high, and relapse is most common during the year following treatment, although it can occur even years later. Patients who achieve normal weight during treatment and are able to maintain that weight have the best long-term prognosis. *See* ANOREXIA NERVOSA; PSYCHOTHERAPY.

Bulimia nervosa. Bulimia nervosa is characterized by recurrent episodes of binge eating (eating large amounts of food while experiencing a subjective sense of lack of control over the eating), the regular use of extreme weight compensatory methods (for example, self-induced vomiting; abuse of laxatives, diuretics, or diet pills; severe dieting or fasting; vigorous exercise), and dysfunctional beliefs about weight and shape that unduly influence self-evaluation or self-worth.

Epidemiology and associated features. Bulimia nervosa occurs in roughly 2% of adolescents and adults. It is most common in females (90% of cases), Caucasians, and middle-class or higher socioeconomic groups. The prevalence of bulimia has increased over the past few decades; persons born after 1960 have a higher rate of developing bulimia. Bulimia nervosa is also becoming more common in non-Caucasian groups.

Persons with bulimia nervosa have high rates of depression, anxiety, and substance abuse problems. The majority of persons with bulimia nervosa are of average weight, although frequent weight fluctuations are common. The preoccupation with weight and shape coupled with considerable shame and guilt regarding their eating practices contributes to frequent interpersonal and social difficulties.

Although this condition is less dangerous than anorexia nervosa, medical complications can occur in bulimia nervosa because of the extreme dieting, binge eating, and purging. Dental erosion and periodontal problems are common. Electrolyte imbalance and dehydration can result in serious medical complications, including cardiac arrhythmias. In rare cases, esophageal bleeding and gastric ruptures occur.

Etiology and maintenance. Bulimia nervosa is likely to result from a combination of genetic, familial, psychological, and sociocultural factors. A family history of bulimia and obesity is considered a risk factor for developing bulimia. In addition, negative experiences such as teasing about weight and shape may precipitate dieting; strict dieting is thought to represent a major risk factor for the development of bulimia. Although many persons have weight and diet concerns, the development of bulimia is thought to arise only in vulnerable individuals and usually after a stressful event. Bulimia nervosa is a self-maintaining vicious cycle: strict dieting and deprivation triggers binge eating, followed by purging to undo the caloric intake and emotional fear of weight gain, followed by shame and disgust, followed by strict dieting.

Treatment. Bulimia nervosa can often be treated successfully with outpatient therapies by appropriately trained professionals. Cognitive-behavioral therapy and interpersonal psychotherapy have been found to be most effective for reducing binge eating and vomiting and improving associated concerns such as depression, self-esteem, and body image. These two therapies also have the best results over the long term. Certain types of pharmacotherapy, notably antidepressant medications, are also effective, although their benefits are less certain once the medication is discontinued. Antidepressant medications appear to help people with bulimia nervosa even if they are not depressed. The combination of antidepressants and cognitive-behavioral therapy appears to be superior to medication treatment alone but not to cognitive-behavioral therapy alone.

Binge eating disorder. Binge eating disorder is characterized by recurrent episodes of binge eating but, unlike bulimia nervosa, no extreme weight control behaviors (purging, laxatives, fasting) are present. Persons with binge eating disorder have chaotic eating patterns and frequently overeat as well as binge.

Epidemiology and associated features. Although obesity is not required for the diagnosis, many people with binge eating disorder are overweight, because strict dieting and frequent purging are not present. Binge eating disorder is estimated to occur in 3% of the general population but in roughly 30% of obese persons. Binge eating disorder occurs most frequently in adulthood, affects men nearly as often as women, and occurs across different ethnic groups.

Overweight persons with binge eating disorder differ from overweight persons who do not binge. In particular, obese binge eaters are characterized by higher levels of psychiatric problems (depression, anxiety, substance abuse) and psychological problems (poor self-esteem, body image dissatisfaction) than non–binge eaters and closely resemble persons with bulimia nervosa. Overweight persons with binge eating disorder are at high risk for further weight gain and weight fluctuations and associated medical complications.

Etiology. The etiology of binge eating disorder is unknown. Adverse childhood experiences, family histories of depression and obesity, and teasing experiences about weight and shape have been identified as risk factors. In contrast to the other eating disorders, a history of dieting does not seem to predate binge eating in roughly half of the cases.

Treatment. Cognitive-behavioral therapy is effective for reducing binge eating and improving associated concerns such as depression, self-esteem, and body image, but does not seem to result in weight loss. There is some evidence that behavioral weight control treatment can reduce binge eating and facilitate

weight loss. Antidepressant medications appear to reduce binge eating but do not produce weight loss; relapse is rapid after discontinuation of the medication. *See* AFFECTIVE DISORDERS. Carlos M. Grilo

Bibliography. C. Fairburn and K. Brownell (eds.), *Eating Disorders and Obesity: A Comprehensive Handbook*, 2d ed., Guilford Press, 2002; C. Fairburn and P. Harrison, Eating disorders, *Lancet*, 361:407–416, 2003; D. Garner and P. Garfinkel (eds.), *Handbook of Treatment for Eating Disorders*, 2d ed., Guilford Press, 1997; C. Grilo et al., Report of the National Institutes of Health workshop on the development of research priorities in eating disorders, *Psychopharm. Bull.*, 33:321–333, 1997; National Collaborating Centre for Mental Health, *Eating Disorders: Core Interventions in the Treatment and Management of Anorexia Nervosa, Bulimia Nervosa, and Related Eating Disorders—NICE Guidelines*, British Psychological Society and Royal College of Psychiatrists, 2004.

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

6 EBE-EYE

www.MHEST.com

E

Ebenales — Eyepiece

Ebenales

An order of flowering plants, division Magnoliophyta (Angiospermae) in the subclass Dilleniidae of the class Magnoliopsida (dicotyledons). The order consists of 5 families and about 1750 species, the Sapotaceae (about 800 species) and Ebenaceae (about 450 species) being the largest and most familiar families. The Ebenales are woody, chiefly tropical, sympetalous plants (those with flowers have the petals joined by their margins, at least toward the base, forming a basal tube, cup, or saucer) with usually twice as many stamens (including staminodes) as corolla lobes. When there is a single set of stamens, these are generally opposite the corolla lobes. Ordinarily there are only one or two ovules in each locule and the placentation is axile (that is, the placenta forms a central axis in an ovary which is divided by longitudinal partitions into two or more chambers). In a cross section of such an ovary, the partitions radiate from the central axis (placenta) to the lateral wall of the ovary. Chicle (from *Manilkara sapota*, in the Sapotaceae) and ebony (*Diospyros ebeneum*) are obtained from members of the Ebenales. *See* CHICLE; DILLENIIDAE; EBONY; MAGNOLIOPHYTA; PLANT KINGDOM.

Arthur Cronquist

Ebola virus

Ebola viruses are a group of exotic viral agents that cause a severe hemorrhagic fever disease in humans and other primates. The four known subtypes or species of Ebola viruses are Zaire, Sudan, Reston, and Côte d'Ivoire (Ivory Coast), named for the geographic locations where these viruses were first determined to cause outbreaks of disease. Ebola viruses are very closely related to, but distinct from, Marburg viruses. Collectively, these pathogenic agents make up a family of viruses known as the Filoviridae.

Infectious agent. Filoviruses have an unusual morphology, with the virus particle, or virion, appearing as long thin rods. Virions have a diameter of 0.08 micrometer and a minimal length of 0.7–0.9 μm; this length is comparable to that of a small bacterium. However, when these viruses are grown in cell cultures, they can form long filamentous and branched forms that reach 14 μm or more. This is especially true for Ebola viruses.

A filovirus virion is composed of a single species of ribonucleic acid (RNA) molecule that is bound together with special viral proteins, and this RNA–protein complex is surrounded by a membrane derived from the outer membrane of infected cells. Infectious virions are formed when the virus buds from the surface of infected cells and is released. Spiked structures on the surface of virions are formed by three molecules of a single glycoprotein and are firmly embedded in the virion membrane. These structures project from the virion and serve to recognize and attach to specific receptor molecules on the surface of susceptible cells. This recognition and binding allows the virion to penetrate into the cell. Then, with the virion free to operate within the cytoplasm, the genetic information contained in the RNA molecule directs production of new virus particles by using the cellular machinery to drive synthesis of new viral proteins and RNA. *See* RIBONUCLEIC ACID (RNA); VIRUS.

Pathogenesis. Although much is known about the agents of Ebola hemorrhagic fever disease, the ecology of Ebola viruses remains a mystery. The natural hosts of filoviruses remain unknown, and there has been little progress at unraveling the events leading to outbreaks or identifying sources of filoviruses in the wild. Fortunately, the incidence of human disease is relatively rare and has been limited to persons living in equatorial Africa or working with the infectious viruses. However, in this time of rapid transportation over large distances, the threat of agents such as Ebola viruses spreading to remote areas is

taken very seriously by public health professionals. People infected with the virus are not as contagious as, say, persons suffering from a cold or measles, and the virus is spread primarily through close contact with the body of an infected individual, his or her body fluids, or some other source of infectious material.

Diagnosis. Ebola virus hemorrhagic fever disease in humans begins with an incubation period of 4–10 days, which is followed by abrupt onset of illness. Fever, headache, weakness, and other flulike symptoms lead to a rapid deterioration in the condition of the individual. In severe cases, bleeding and the appearance of small red spots or rashes over the body indicate that the disease has affected the integrity of the circulatory system. Contrary to popular belief, individuals with Ebola virus infections do not melt or have their organs dissolve, but die as a result of a shock syndrome that usually occurs 6–9 days after the onset of symptoms. This shock is due to the inability to control vascular functions and the massive injury to body tissues.

Ebola viruses can be found in high concentrations in tissues throughout the body and are especially evident in the liver, spleen, and skin. From studies of human cases and experimentally infected animals, it appears that the immune response is impaired and that a strong cellular immune response is key to surviving infections. This immunosuppression may also be a factor in death, especially if secondary infections by normal bacterial flora ensue. No severe human disease has been associated with Reston virus infections, and it appears that this virus may be much less pathogenic for humans than it is for monkeys. *See* IMMUNOSUPPRESSION.

Epidemiology. The first described cases of Ebola virus disease occurred in 1976, when simultaneous outbreaks of two distinct subtypes occurred in northern Zaire and southern Sudan. Many of the human infections occurred in local hospitals, where close contact with fatal cases, reuse of contaminated needles, and a low standard of medical care led to most fatalities. An Asian form of Ebola virus was discovered in a type of macaque, the cynomolgus, exported from the Philippines to the United States in late 1989. It was determined that a single monkey breeding facility in the Philippines was the source of the virus. Named Reston virus after the Virginia city in which infected monkeys were first identified, it represented a new form of Ebola virus. A repeat of this incident took place in early 1992, when monkeys from the same breeding facility were shipped to Siena, Italy.

In 1994, another new species of Ebola virus was identified, associated with chimpanzee deaths in the west African country of Côte d'Ivoire. This episode signaled the reemergence of Ebola virus in Africa, which had not been seen since 1979 when an outbreak of the Sudan subtype occurred. A large outbreak of disease was identified in early May 1995 in and around the city of Kikwit, Zaire. This outbreak of a Zaire subtype was a repeat of the 1976 episode, but this time 600 mi (1000 km) to the south. Another Ebola virus outbreak took place in 1995 in Gabon and resulted in human fatalities, but the details were slow to be revealed. This virus has been isolated, and preliminary reports indicate that it may be a variant of the Zaire subtype. This virus may have also been responsible for human fatalities in Gabon reported to have occurred in early 1996. Soon after this outbreak, yet another Reston virus outbreak occurred when infected monkeys were sent to a United States holding facility in Alice, Texas. That introduction was quickly detected and stopped by quarantine and testing measures implemented after the first Reston virus episode in 1989.

Control and prevention. Outbreaks of Ebola virus disease in humans are controlled by the identification and isolation of infected individuals, implementation of barrier nursing techniques, and rapid disinfection of contaminated material. Diagnosis of Ebola virus cases is made by detecting virus proteins or RNA in blood or tissue specimens, or by detecting antibodies to the virus in the blood. Such testing is important in tracking and controlling the movement of the virus during an outbreak.

Dilute hypochlorite solutions (bleach), 3% phenolic solutions, or simple detergents (laundry or dish soap) can be used to destroy infectious virions. No known drugs have been shown to be effective in treating Ebola virus (or Marburg virus) infections, and protective vaccines against filoviruses have not been developed. However, research is being directed at developing such treatment and should provide insights into the disease mechanisms used by filoviruses.

Evolution. Filoviruses have been shown to be genetically related to two other virus families, the Paramyxoviridae (including measles virus and mumps virus) and the Rhabdoviridae (including rabies virus and vesicular stomatitis virus). These viruses evolved from a common progenitor in the very distant past. These viruses evolved into very distinct lineages, yet have maintained a similar approach to reproducing.

The evolutionary profile, or phylogeny, for Ebola viruses and Marburg viruses has been determined. Ebola viruses have evolved into a separate and very distinct lineage within the filovirus family, and the individual species of Ebola viruses have also evolved into their own sublineages. Filoviruses show a great deal of diversity, indicating that they have likely coevolved with their natural hosts over a long period of time. Ebola viruses within a given subtype did not appear to show a great deal of change over periods of many years. This genetic stability or stasis in the wild suggests that these viruses have evolved to occupy very specific ecological niches. Anthony Sanchez

Bibliography. Centers for Disease Control, Ebola-Reston virus infection among quarantined nonhuman primates—Texas, 1996, *MMWR*, 45:314–316, 1996; B. N. Fields et al. (eds.), *Fields Virology*, 3d ed., 1996; L. G. Horowitz, *Emerging Viruses: AIDS and Ebola—Nature, Accident or Intentional*, 1996; A. Sanchez et al., Reemergence of Ebola virus in Africa, *Emerg. Infect. Dis.*, 1:96–97, 1995;

A. Sanchez et al., The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing, *Proc. Nat. Acad. Sci. USA*, 93:3602–3607, 1996.

Ebony

A genus, *Diospyros*, of the ebony family, containing more than 250 species. Some species are important for succulent fruits, such as date plum, kaki plum, and persimmon, and several for timber, particularly the heartwood, the true ebony of commerce.

Although it is popularly supposed to be a black wood, most species have a heartwood only streaked and mottled with black. The heartwood is brittle, breaks with a conchoidal fracture, and is difficult to work, but it has long been in demand. The sapwood is white, becoming bluish or reddish when cut.

The use of ebony can be traced to the early Egyptians, who probably obtained it in Ethiopia. Ebony from India was known to the Greeks before 350 B.C. At present black ebony is used for knife handles, piano keys, finger boards of violins, hairbrush backs, inlays, and marquetry. Some of the woods called ebony, however, belong to different families, especially the pulse family, Leguminosae.

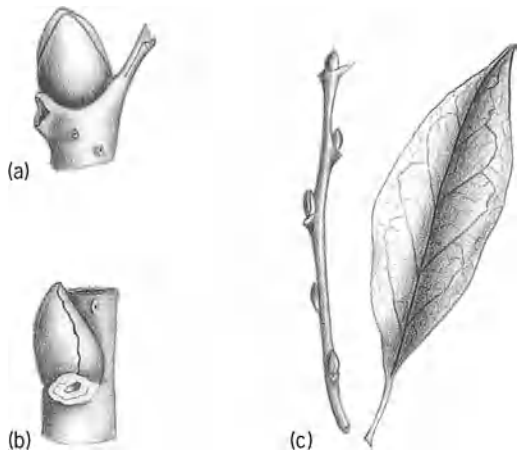

Persimmon (*Diospyros virginiana*). (*a*) Terminal bud. (*b*) Lateral bud. (*c*) Twig and leaf.

Persimmon (*D. virginiana*), of the southeastern United States, is one of numerous tropical or subtropical species (see **illus.**). Usually a medium-sized three with black chunky bark, it attains 100–125 ft (30–37 m) in height with a trunk 20–30 in. (50–75 cm) in diameter. The sapwood is in demand for the manufacture of weaving shuttles and heads of golf clubs. The fruit is sweet and edible when slightly overripe, but when immature it is extremely pungent. The species in tropical America are too small or rare to be of economic value, although several of them have black heartwood used locally for making walking sticks, inlays, and miscellaneous articles of turnery and carving. *See* FOREST AND FORESTRY; TREE. Arthur H. Graves; Kenneth P. Davis

Ebriida

An order of flagellate Protozoa, subphylum Sarcomastigophora, class Phytamastigophorea. Two genera, *Ebria* and *Hermesinum*, remain of the once numerous order, as determined by fossil remains. They possess an internal solid siliceous skeleton forming a shallow flattened or slightly arched structure that is enclosed by clear cytoplasm. Skeleton form and structure are distinctive. These phytoflagellates do not have chromatophores. A conspicuous nucleus lies anteriorly and two long and fine flagella emerge near it. Reserve materials apparently are fat. *See* CILIA AND FLAGELLA.

The organisms are common in the inshore waters of the Atlantic Ocean and the Gulf of Mexico. However, numbers sufficient to influence the general ecology have not been observed, nor has any ingestion of particulate matter. Their only occurrence seems to be marine, although somewhat lowered salinities are certainly tolerated. *See* PHYTAMASTIGOPHOREA; PROTOZOA. James B. Lackey

Ecdysone

The molting hormone of insects. It was isolated in crystalline form from the common silkworm (*Bombyx mori*) by A. Butenandt and P. Karlson in 1954, and its chemical structure was elucidated in 1965. It is a derivative of cholesterol and its structure can be described as $2\beta,3\beta,14\alpha$,22R,25-penta-hydroxy-Δ^7-5β-cholestene-6-one. Shortly thereafter the molting hormone of crustaceans (crustecdysone) was isolated and identified as 20-hydroxy-ecdysone (**Fig. 1**). The sample compound was identified as one of the active compounds in *B. mori* extracts by P. Hocks and coworkers, and H. Hoffmeister and coworkers proposed the name ecdysterone. A group of Japanese chemists, working on the isolation of steroids from plants, came across some similarities of the spectra of 20-hydroxy-ecdysone and their compounds inokosterone and isoinokosterone; the latter was identical with 20-hydroxy-ecdysone. Inokosterone carries the last hydroxy group in position 26 instead of 25, which reduces the biological activity by a factor of 10. Several other structurally related and biologically active steroids have been isolated from plant sources, including ecdysone, but nothing is known about a possible biological function of these steroids in plants. *See* STEROID.

Ecdysone is synthesized in insects from cholesterol; 7-dehydro-cholesterol may be an intermediate. It has been shown that ecdysone is also readily metabolized and inactivated by an enzyme system; the inactivation occurs in the living insect, as well as in homogenates. The activity of the inactivating system varies markedly during development. The most striking physiological activity of ecdysone is the induction of puffs (zones of gene activity) in giant chromosomes of the salivary glands and other organs of the midge *Chironomus*. It has been shown that 20-hydroxy-ecdysone also induces puffs. The

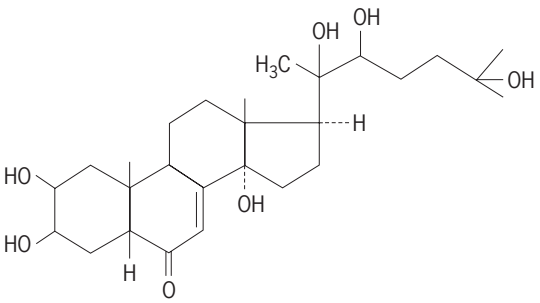

Fig. 1. Structure of 20-hydroxy-ecdysone.

induction of puffs has been visualized as the primary action of the hormone, indicating that ecdysone controls the activity of specific genes. The postulated mechanism for the action of ecdysone is shown in **Fig 2**; the primary effect is gene activation. (In the figure DNA is deoxyribonucleic acid, ATP is adenosine triphosphate, GTP is guanosine triphosphate, CTP is cytosine triphosphate, and UTP is uridine triphosphate.) The active gene produces messenger ribonucleic acid (mRNA), which in turn combines with ribosomes and becomes involved with protein synthesis. These proteins may be very specific and may be enzyme molecules. It has been shown that ecdysone stimulates the synthesis of mRNA, among which is the messenger for dopa decarboxylase. This enzyme is involved in the biosynthesis of the sclerotizing agent *N*-acetyl-dopamine. *See* CHROMOSOME; INSECT PHYSIOLOGY. Peter Karlson

Bibliography. F. Schwalm, *Insect Morphogenesis*, 1987; F. Taylor and R. Karban (eds.), *The Evolution of Insect Life Cycles*, 1986; A. Zaslavsky, *Insect Development*, 1988.

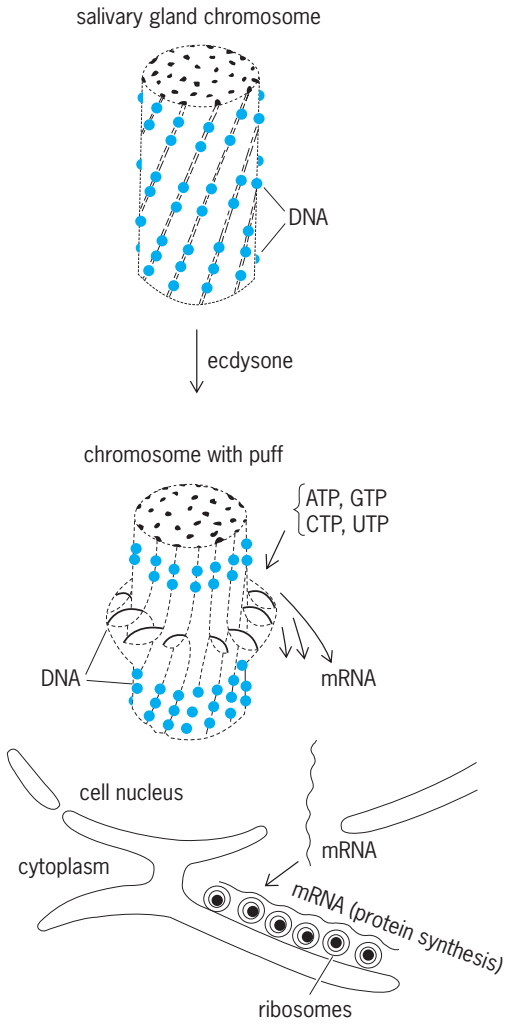

Fig. 2. Drawing of postulated mechanism of action of ecdysone in protein production in the midge *Chironomus*.

Ecdysozoa

A major division within the animal kingdom containing the majority of animal species and comprising the phyla Priapulida, Kinorhyncha, and Loricifera (no common names), Arthropoda (insects, crustaceans, spiders, etc.), Onychophora (velvet worms), Tardigrada (water bears), and Nematoda (roundworms). These forms include animals with disparate morphologies, but their close relationship is suggested by molecular DNA sequence comparisons—especially of both large and small subunit RNA molecules—and is consistent with myosin heavy-chain sequence comparisons and with a variety of uniquely shared genomic and morphological features. Ecdysozoans were named after their habit of molting (ecdysis), unique among invertebrates, and presumably all are descended from a common molting ancestor. *See* ANIMAL KINGDOM; MOLTING (ARTHROPODA).

General characteristics. Ecdysozoa is the major subdivision of the Protostomia (see **illustration**). In most protostomes the adult mouth forms on the site of the blastopore, and the secondary body cavity (the coelom) forms by splitting between tissue layers of the mesoderm. None of the ecdysozoan phyla have a eucoelom—a secondary body cavity that contains the major body organs and acts as a fluid skeleton—but may have smaller coelomic spaces, particularly during development, that function as ducts or contain specialized organs. Fluid skeletons are furnished by primary body cavities that either are pseudocoels (in the absence of a blood circulatory system) or, in arthropods and onychophorans, are hemocoels (contain blood). In this feature they differ from most deuterostomes and lophotrochozoans (see illustration). The cuticle is largely flexible in some forms (priapulids, onychophorans) but is moderately to strongly sclerotized in others and even mineralized in some arthropods, and acts to stiffen the body wall and to serve as a site for muscle attachments, forming an exoskeleton in arthropods. Growth is accommodated by periodic molting and then secretion of a larger cuticle. None of the marine ecdysozoans has primary planktonic (free-floating) larvae, although in some arthropods several early molting stages are planktonic and serve for dispersal. *See* COELOM.

Priapulida, Kinorhyncha, and Loricifera. These phyla are pseudocoelomate worms, and all have minute

species that live interstitially in marine sediments; loriciferans are attached to sand grains, while the others are free-living. The cuticles bear secreted spines or plates that are sensory or aid in burrowing. Some priapulids also grow to much larger sizes [to 20 cm (8 in.)] and form prominent burrow structures. Kinorhynchs have a segmented cuticular exoskeleton and longitudinal muscle system, but the other two phyla are unsegmented. *See* KINORHYNCHA; LORICIFERA; PRIAPULIDA.

Arthropoda, Onychophora, and Tardigrada. These phyla are more closely related to each other than to other ecdysozoans on molecular evidence. All three are segmented to some extent and possess limbs. In onychophorans and tardigrades, the cuticle is rather flexible, and the limbs (lobopods) contain pseudocoelomic cavities that provide turgor. In arthropods, however, the cuticle is stiffened to form an exoskeleton, and the limbs act as levers operated by intrinsic muscles. All these phyla were originally marine but have their chief diversity in terrestrial habitats today; the onychophora are now entirely nonmarine. *See* ARTHROPODA; ONYCHOPHORA; TARDIGRADA.

Nematoda. This phylum consists of small, roundish worms that are unsegmented and pseudocoelomate, though minute species may essentially lack any body cavity. Many nematodes are parasitic, and one highly derived parasitic group, the Nematomorpha (horsehair worms), is sometimes considered as a phylum of its own. *See* NEMATA (NEMATODA); NEMATOMORPHA.

Evolution: diversity. The earliest fossils that are very likely to be ecdysozoans are Early Cambrian trails and feeding marks left by scratching limbs, probably dating 525–520 Ma (million years ago), although burrows from as early as about 540 Ma could have been formed by priapulids. Body fossils of ecdysozoans appear around 530–525 Ma, represented then by stem groups (extinct branches) of priapulids, arthropods, and onychophorans, and by extinct arthropod allies that are segmented but evidently lack jointed appendages; some of these are quite large by Cambrian standards, up to 50 cm (20 in.) long. There is also an extinct Early Cambrian vermiform group, the paleoscolecids, which may be pseudocoelomate ecdysozoans. Tardigrades are first known from about 510–500 Ma, while the other phyla do not have body fossil records.

Ecdysozoans dominate the early animal fossil record, with a great disparity of form (mostly extinct variants of arthropod architectures) and in terms of species richness. During the Cambrian (543–500 Ma), over 70% of known animal species are trilobites, an extinct arthropod class. Ecdysozoans continued to be prominent in the seas until today, where they are now challenged for dominance only by Mollusca among invertebrate phyla. The earliest indications of animal fossils from terrestrial environments are traces, burrows in soils deposited near 410 Ma; body fossil assemblages appearing near 400 Ma and slightly later are almost entirely of arthropods, including millipedes, springtails, and rarely insects; they are dominated by detritus feeders and predators. Today, arthropod species richness, especially of insects, eclipses that of all other animal phyla in terrestrial habitats. For insects, the small body sizes and modular body plans within a secreted, segmented exoskeleton and with jointed appendages that can be specialized individually—together with the habitat heterogeneity afforded by the rise of diverse global plant communities—have produced unique opportunities for evolutionary radiations, which seem to have been fully exploited. *See* FOSSIL; TRILOBITA.

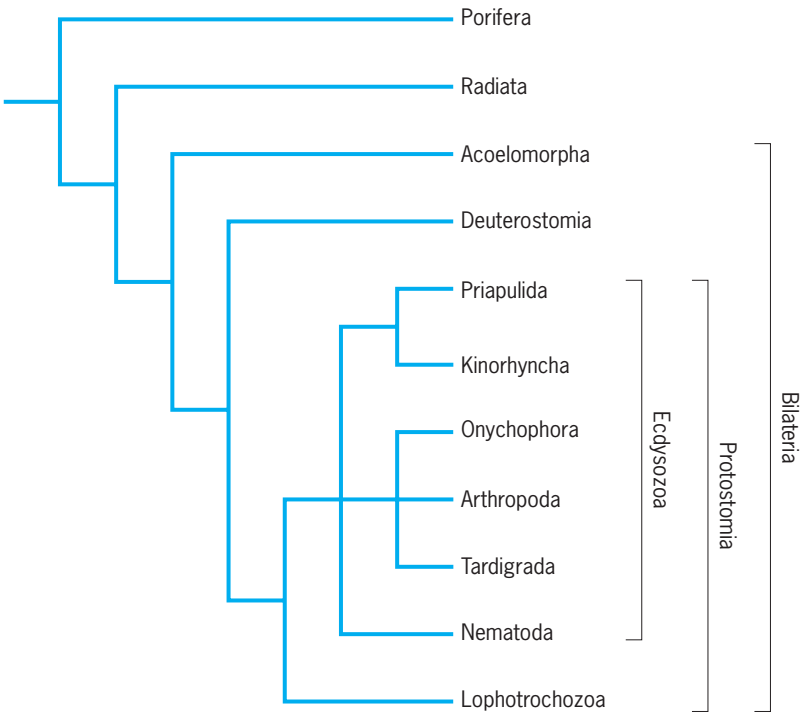

Relationships among phylum-level taxa within the Ecdysozoa, and the relation of the Ecdysozoa to other major animal groups.

Evolution: genomics. The two best-known model organisms for evolutionary genetic and developmental studies, the fly *Drosophila* (Arthropoda) and the roundworm *Caenorhabditis* (Nematoda), are ecdysozoans. Understanding of the genetic processes underpinning body plan development has been worked out largely in these forms, which have both segmented and unsegmented architectures, respectively. As the key genes regulating these processes are widely distributed among other animal groups as well, including humans, comparative studies using the ecdysozoan models are helping to clarify the evolution of animal genomes during the branchings of the tree of life. *See* ANIMAL EVOLUTION; ORGANIC EVOLUTION.

James W. Valentine

Bibliography. R. C. Brusca and G. J. Brusca, *Invertebrates*, Sinauer Associates, Sunderland, MA, 2003; J. Cracraft and M. J. Donoghue (eds.), *Assembling the Tree of Life*, Oxford University Press, 2004; R. A. Fortey and R. H. Thomas (eds.), *Arthropod Relationships*, Chapman and Hall, 1998; M. W. Kirschner and J. C. Gerhart, *The Plausibility of Life*, Yale University Press, 2005; J. W. Valentine, *On the Origin of Phyla*, University of Chicago Press, 2004.

Echinacea

A superorder of Euechinoidea, having a rigid test, the periproct within the apical system, keeled teeth, a complete perignathic girdle, and branchial slits. J. Durham and R. Melville (1957) include five orders in this group. There were formerly distributed among the Stirodonta and Camarodonta in the classification of R. Jackson (1912). *See* ECHINODERMATA; ECHINOIDA; EUECHINOIDEA; HEMICIDAROIDA; PHYMOSOMATOIDA; SALENIOIDA; TEMNOPLEUROIDA.

Howard B. Fell

Echinodermata

A phylum of exclusively marine animals with a peculiar body architecture dominated by a five-part radial symmetry. Echinodermata [from the Latin *echinus* (spine) + *dermis* (skin: "spiny skins")] include the sea stars, sea urchins, and related animals. The body wall contains an endoskeleton of numerous plates (ossicles) composed of calcium carbonate in the form of calcite and frequently supporting spines. The plates may be tightly interlocked or loosely associated. The spines may protrude through the outer epithelium and are often used for defense. The skeletal plates of the body wall, together with their closely associated connective tissues and muscles, form a tough and sometimes rigid test (hard shell) which encloses the large coelom. A unique water-vascular system is involved in locomotion, respiration, food gathering, and sensory perception. This system is evident outside the body as five rows of fluid-filled tube feet. Within the body wall lie the ducts and fluid reservoirs necessary to protract and retract the tube feet by hydrostatic pressure. The nervous system of these headless animals arises from the embryonic ectoderm and consists of a ring around the mouth with connecting nerve cords associated with the rows of tube feet. There may also be diffuse nerve plexuses, with light-sensing organs, lying below the outer epithelium. The coelom houses the alimentary canal and associated organs and, in most groups, the reproductive organs. The body may be essentially star-shaped or globoid. The five rows of tube feet define areas known as ambulacra, ambs, or radii; areas of the body between the rows of tube feet are interambulacra, interambs, or interradii.

The larvae are usually planktonic (free-floating) with a bilateral symmetry, but the adults are usually sedentary and benthic (bottom-dwelling in a marine environment). They inhabit all oceans, ranging from the shores to the greatest ocean depths.

The phylum comprises about 7000 existing species. The Echinodermata have a good fossil record with about 13,000 fossil species. They first appeared in the Early Cambrian and have been evolving, since pre-Cambrian time, for well over 600 million years. During this time several differing body plans have arisen. The surviving groups show few resemblances to the original stock. The existing representatives fall into five, possibly six, classes: Crinoidea (sea lilies and feather stars); Asteroidea (sea stars); Ophiuroidea (brittle stars); Echinoidea (sea urchins, sand dollars, and heart urchins); and Holothuroidea (sea cucumbers). The recently described class Concentricycloidea (sea daisies) has been referred to the Asteroidea by some experts and retained as a distinct class by others.

The outline of classification for the phylum is shown below. Classes with living representatives are elaborated to the level of subclass or order.

Phylum Echinodermata
"Homalozoans"
Class: Ctenocystoidea
Stylophora
Homostelea
Homoiostelea
"Crinozoans"
Class: Eocrinoidea
Paracrinoidea
Rhombifera
Diploporita
Crinoidea
Blastoidea
Coronoidea
Parablastoidea
Edrioblastoidea
Subclass: Inadunata
Camerata
Flexibilia
Articulata
Order: Millericrinida
Cyrtocrinida
Bourgeticrinida
Isocrinida
Uintacrinida
Roveacrinida
Comatulida
"Asterozoans"
Class: Somasteroidea
Asteroidea
Order: Platyasterida
Trichasteropsida
Paxillosida
Notomyotida
Spinulosida
Valvatida
Velatida
Forcipulatida
Class: Concentricycloidea
Order: Peripodida
Class: Ophiuroidea
Order: Stenurida
Oegophiurida
Phrynophiurida
Ophiurida
"Echinozoans"
Class: Helicoplacoidea
Camptostromatoidea
Edrioasteroidea
Cyclocystoidea
Ophiocistioidea
Echinoidea

Subclass: Perischoechinoidea
Order: Bothriocidaroida
Echinocystitoida
Subclass: Cidaroidea
Order: Cidaroida
Subclass: Euechinoidea
Order: Diadematoida
Echinothurioida
Pedinoida
Arbacioida
Echinoida
Phymosomatoida
Salenioida
Temnopleuroida
Clypeasteroida
Holectypoida
Cassiduloida
Spatangoida
Class: Holothuroidea
Order: Dendrochirotida
Dactylochirotida
Aspidochirotida
Elasipodida
Molpadiida
Apodida

Morphology

Despite displaying an astonishing variety of body forms, existing echinoderms share common anatomical features: symmetry, body wall, skeleton, nervous system, coelom, alimentary system, reproductive organs, and water-vascular system.

Symmetry. Adult echinoderms usually show pentamerism, a five-part radial symmetry in which there are normally five axes radiating out from a central point in the body. The mouth may lie on the underside, and the anus may lie on top of the animal. There are numerous exceptions: in the crinoids the mouth and anus lie on the same side, and in the holothuroids the mouth and anus lie at opposite ends of a cylindrical body. The side of the animal bearing the mouth is termed the oral side, and the side away from the mouth is the aboral side. The relationship of the mouth with the substratum was at one time considered important so that those echinoderms which were essentially sessile, with a mouth facing away from the substratum, were classified as subphylum Pelmatozoa, while those mobile forms in which the mouth faced the substrate were classified as Eleutherozoa. These differing life habits have little bearing upon formal echinoderm classification, but the terms "pelmatozoan" and "eleutherozoan" are useful descriptors of lifestyles in the phylum.

It has been suggested that five-part symmetry in echinoderms confers strength upon the skeleton supporting the body wall. Pentamerism was not an original feature of the echinoderms. Certain early groups such as the carpoids did not show it, and modern echinoderm larvae have bilateral symmetry.

Body wall. The body wall is cloaked in a thin epidermis, a layer of epithelial cells. The external layer is often ciliated, and the electron microscope has shown that these cells have an outer membrane that bears a great number of minute tubular extensions like microvilli. This arrangement of the epidermis is similar in appearance to the absorptive epithelium of many intestinal tissues. Below the epidermis, the dermis contains the skeletal elements. Internal to the dermis are the body-wall muscles and connective tissue. In the crinoids, ophiuroids, some asteroids, and some echinoids, the skeletal elements make up a great part of the body-wall volume. In the more flexible asteroids, the spherical echinoids, and the holothuroids, the volume of these elements is reduced. The inner surface of the body wall is also lined with epithelium and borders the coelom.

Skeleton. The skeleton is unique and consistent in all groups. Almost every plate, spine, or ossicle is composed of a single calcite crystal. During development each plate grows as a result of calcite secretion by a group of cells. The plate grows like a three-dimensional lattice (stereom) to produce a reticulate (having or resembling a network of fibers or lines) ossicle, with the tissue that secretes and maintains it (stroma) occupying the spaces within (**Fig. 1**).

Sensory and neuromuscular system. Echinoderms have simple sensory systems. The nervous system is ectodermal in origin and comprises a radial nerve cord lying along each ambulacrum. Each radial nerve connects with the circumesophageal nerve ring, and it is believed that interradial coordination takes place via this ring. The absence of a head means that sensory receptors and large areas of integrative nervous elements are not aggregated in one place. The nervous system is essentially diffuse, although specialized sense organs, such as the recently discovered "eyes" of certain brittle stars, occur. It appears that the circumoral ring and radial nerve cords take

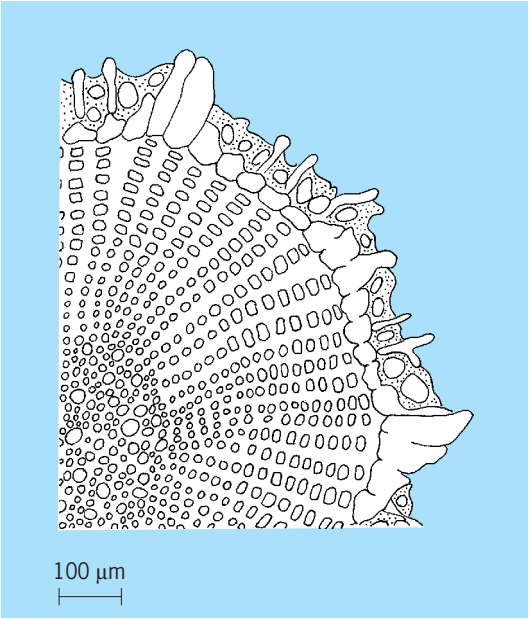

Fig. 1. Cross section through the spine of an echinoid showing the microscopic structure of the skeleton. The stereom forms a continuous mesh, and the stroma that secretes it lies in the interspaces which, in this case, are arranged in radial rows.

responsibility for gross activites such as posture, coordinated locomotion, and food collecting, but in the asteroids and the echinoids there are well-developed basiepithelial nerve networks that coordinate the numerous test appendages. Echinoderm muscles are principally smooth, but striated muscles have been detected in a few instances. *See* NERVOUS SYSTEM (INVERTEBRATE).

Coelom. The coelom is extensive and usually arises by enterocoely, or pouching, from the gut; but when there is direct development, with larval stages greatly reduced or absent, it may be schizocoelic, involving splitting in the mesoderm. The perivisceral coelom surrounds the viscera. During development, part of the coelom forms the water-vascular system, so the spaces within the tube feet and vessels are coelomic. *See* COELOM.

Alimentary system. In general the mouth is situated on the same side as the tube feet. From it arises the alimentary canal, which runs through the body to the anus (absent in the ophiuroids and some asteroids). The gut is lined by endoderm, and in some groups digestive caeea (blind pouches at the beginning of the large intestine) increase the area for absorption.

Particulate food is collected by the tube feet in many crinoids, ophiuroids, and holothuroids. Most asteroids are carnivores and engulf the prey, or evert part of the stomach over it, while some echinoids have well-developed chewing teeth for rasping at encrusting plants and animals. Many holothuroids, some echinoids, and a few asteroids ingest mud or sand, obtaining sustenance from associated organic material such as bacteria and diatoms. *See* DIGESTION (INVERTEBRATE).

Reproductive system. The reproductive organs lie within the coelom in the interradial position, except in the crinoids in which they arise on the arms. In most echinoderms there are five compact gonads, but the holothuroids have just one, and certain echinoids may have two, three, or four. The sexes are usually separate, The gonads discharge gametes by short ducts into the surrounding seawater. In some echinoderms, especially in polar regions, fertilized gametes are retained, and the young develop in special sacs in the body wall.

Water-vascular system. This is a multifunctional fluid-filled coelomic system that probably evolved first as a respiratory system but later took on increasingly important roles in food collection and locomotion (**Fig. 2**). It comprises a number of protrusible, hollow, tentacle-like tube feet, arranged externally in rows along each radial area of the body. These areas, composed of single to multiple rows of tube feet, are termed ambulacra. A small modified test ossicle, the madreporite, perforated by many irregularly shaped pores, connects with the water-vascular system via a calcified stone canal. It was thought that the water-vascular system opened directly to the surrounding seawater via the madreporite so that extra fluid could be drawn in to replenish the water-vascular fluid. The fluid volume of the water-vascular system may be more constant than was previously believed, and relatively little fluid passes in or out of the madreporite. Each tube foot is associated with a small bulbous reservoir inside the body. The feet are extended hydrostatically, and retracted by longitudinal muscles in their walls.

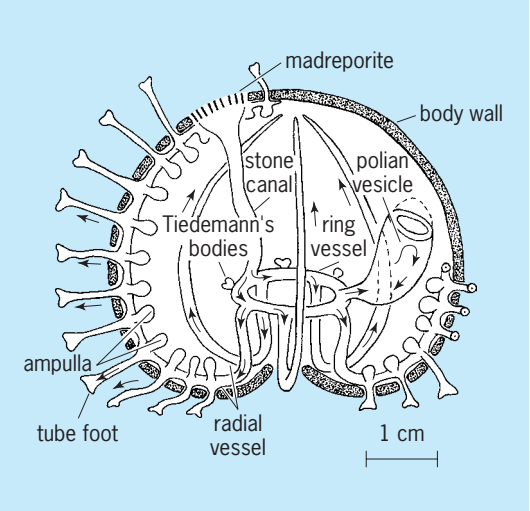

Fig. 2. Diagram of the water-vascular system in an echinoid. Arrows show direction of fluid flow.

The form and the activity of tube feet vary greatly between the groups of echinoderms. In the crinoids they are fine, tapering structures suited for food collection and having no locomotory role. In the brittle stars they lack terminal suckers and are more important for food gathering than locomotion. In some asteroids, particularly the burrowers, they are tapering and pointed, while in others they are suckered and capable of adhering to the substratum. This is also the case in regular echinoids (which have no front or back end and can move in any direction), but in irregular echinoids (which have a definite front and back and do move in a particular direction) the tube feet are modified to suit the burrowing way of life. In holothuroids the tube feet surrounding the mouth are modified as tentacles of varying form and complexity and are used for food collecting. The rest may be present as locomotor or respiratory organs. In a few groups, such as the apodous and molpadiid holothurians, tube feet are missing from the body wall. When tube feet serve an important locomotor function, postural muscles govern the stepping movements.

Each tube foot is under the control of the radial nerve cord and is connected by a short lateral branch to the radial water-vascular canal as well as to the individual ampullae. All the radial canals are linked by the circumoral ring canal so that water-vascular fluid may be withdrawn from one part of the system to be supplied to another.

Saclike structures are associated with the ring canal. These are polian vesicles, which act as reservoirs for water-vascular fluid, and Tiedemann's bodies, which act as glands in which wandering coelomocytes (coelom corpuscles) are formed.

Embryology

Most echinoderms have an indirect development with a prolonged, food-gathering larval stage (**Fig. 3**). The larva feeds, usually on microscopic phytoplankton, using a ciliary mechanism. Two well-marked larval types occur: the pluteus group, with long-armed, bilaterally symmetrical, easel-shaped forms, common to ophiuroids and echinoids; and the auricularia group, barrel-shaped forms with a winding ciliated band which may be produced into lobes. The latter group is common to asteroids and holothuroids. In most asteroids the auricularia stage is followed by a similar but more complex larva, the bipinnaria, or sometimes also by an anchored final larval stage, the brachiolaria. Surviving crinoids have essentially a direct development, sometimes with a simple yolky larva, a vitellaria, which does not feed. This is also the case for several members of other extant echinoderm groups. *See* INVERTEBRATE EMBRYOLOGY.

Phylogeny

The auricularia larva presents close and striking resemblances to the tornaria larva of some enteropneusts, and the enterocoelous development parallels that in primitive chordates. Hence echinoderms and chordates have long been regarded as related. However, the significance of similarities in the larvae of echinoderms and protochordates must be viewed in the context of molecular, morphological, and paleontological research. It seems likely that the similarities between the larvae of ophiuroids and echinoids, and asteroids and holothuroids, are due to convergent evolution and not to common evolutionary origins. The results of paleontology, molecular

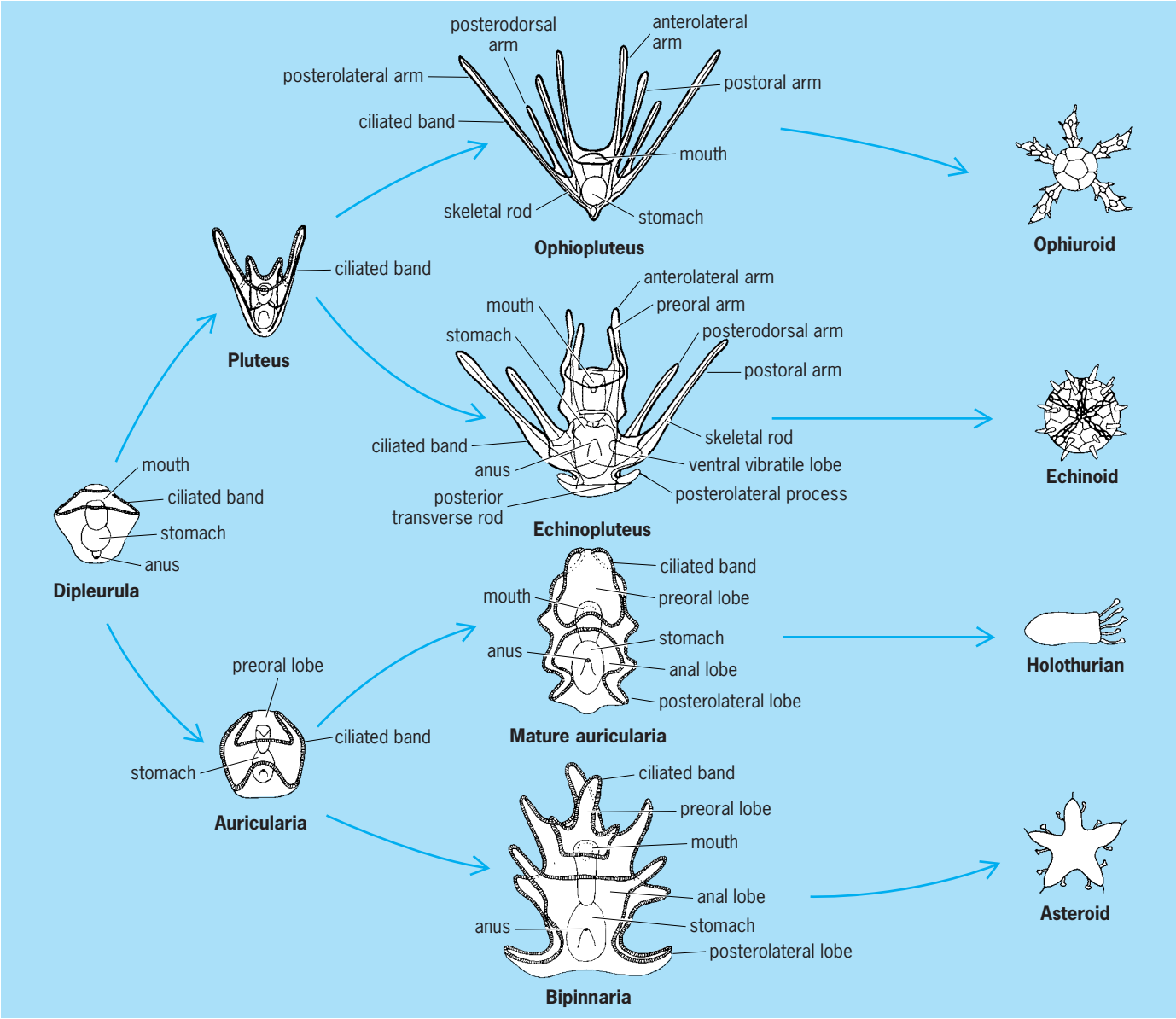

Fig. 3. General scheme indicating relationships of bilaterally symmetrical echinoderm larvae. (*After H. B. Fell*)

studies, and morphology indicate that ophiuroids and asteroids are closely related. Therefore it follows that within the phylum larval similarities do not indicate phylogenetic affinities. It is inadvisable to try to extrapolate beyond the phylum, so as to infer phylogenetic affinity between hemichordates and echinoderms solely on the ground that the auricularia resembles the tornaria. E. Marcus expressed the opinion that indirect development, with possession of pelagic larvae, must be prototypical for echinoderms and protochordates, and the asteroids and ophiuroids must be closely related, and therefore broad phylogenetic conclusions cannot be drawn on the basis of their larvae. L. Hyman grouped the extant echinoderms as their larval similarities suggest, and concluded that the arrangement adopted by paleontologists must be wrong. H. B. Fell and later authors have reaffirmed that the paleontological evidence overrules nebulous embryological considerations. David L. Pawson; Andrew C. Campbell

Ecology and Feeding

Echinoderms occur everywhere in the world's oceans, but they are usually rare in areas where the salinity of the water is greatly reduced or where pollution levels are significant. In many areas, echinoderms are dominant invertebrates in terms of numbers or biomass, and locally sea stars can be the top predators. On a rocky shore, predatory sea stars and vegetarian sea urchins may be common on and under rocks. Brittle stars are typically concealed under rocks; they may emerge at night to forage for small organisms. In deeper water, where conditions are suitable for suspension feeding, huge aggregations of brittle stars or feather stars may occur, with their arms extended into the water for feeding. Echinoderms flourish on and near coral reefs, and it is in these tropical areas that they can achieve their greatest diversity. In polar regions, especially around Antarctica, echinoderms, along with sponges, dominate in most habitats (**Fig. 4**). Sandy to muddy bottoms may be populated with large numbers of burrowing detritus-eating sea urchins, long-armed brittle stars, and sea cucumbers. In such habitats, burrowing sea cucumbers may extend their sticky tentacles into the surrounding water to capture small drifting organisms. In deeper water, below 100 m (330 ft), stalked sea lilies can be common on hard substrates, efficiently suspension-feeding with their arms deployed in the form of a parabolic bowl. In the deep sea, echinoderms can also be dominant; on abyssal (relating to great ocean depths) plains, sediment-swallowing sea cucumbers can make up more than 95% of the total biomass on the sediment surface. Most echinoderm groups occur in the deep sea to depths in excess of 9000 m (5.6 mi). David L. Pawson

Fig. 4. Echinoderms in the deep sea near Antarctica at a depth of 595 m (1950 ft). Three sea cucumbers (*Scotoplanes globosa*) are feeding on seafloor sediments. At bottom right is a brittle star, probably *Ophiomusium* species. Scattered on the seafloor but barely visible are several feather stars. (*Courtesy of the U.S. National Science Foundation*)

Fossils

Echinodermata are an especially important group of fossil invertebrates. Except for reworked fragments found in some nonmarine deposits, echinoderm remains occur exclusively in strata laid down on sea bottoms—chiefly those of shallow seas. Many of these deposits, ranging in age from Cambrian to late Tertiary, contain abundant echinoderm fossils. In addition, fossil echinoderms undoubtedly are widely distributed beneath all oceans, although remains of these organisms in deep-water sediments, even of recent origin, are virtually unknown because of their inaccessibility.

The echinoderms are well adapted to preservation as fossils owing to their abundant calcareous skeletal elements. Paleontological importance of the group is explained partly by this fitness but more by the diversity of their kinds, the generally short-lived existence and wide geographic distribution of most recognized taxonomic units, and the clearness with which evolutionary trends can be defined.

Classification and Description

For some years, fossil and living echinoderms have been arranged into four or five subphyla. As formal groupings, the subphyla have been largely abandoned, for it is recognized that the subphylum definitions and distinctions are artificial rather than truly reflecting phylogeny. However, the subphylum names are useful for grouping the various echinoderm classes, and they are used here in an informal sense: homalozoans, crinozoans, echinozoans, and asterozoans. Similarly, the terms eleutherozoan (free-moving) and pelmatozoan (sedentary or sessile), previously used in formal classifications, are useful today in categorizing the lifestyles of echinoderms. *See* ELEUTHEROZOA; PELMATOZOA.

Homalozoans. The distinctive features that separate the extinct homalozoans (or carpoids) from all other Echinodermata are a complete lack of radial

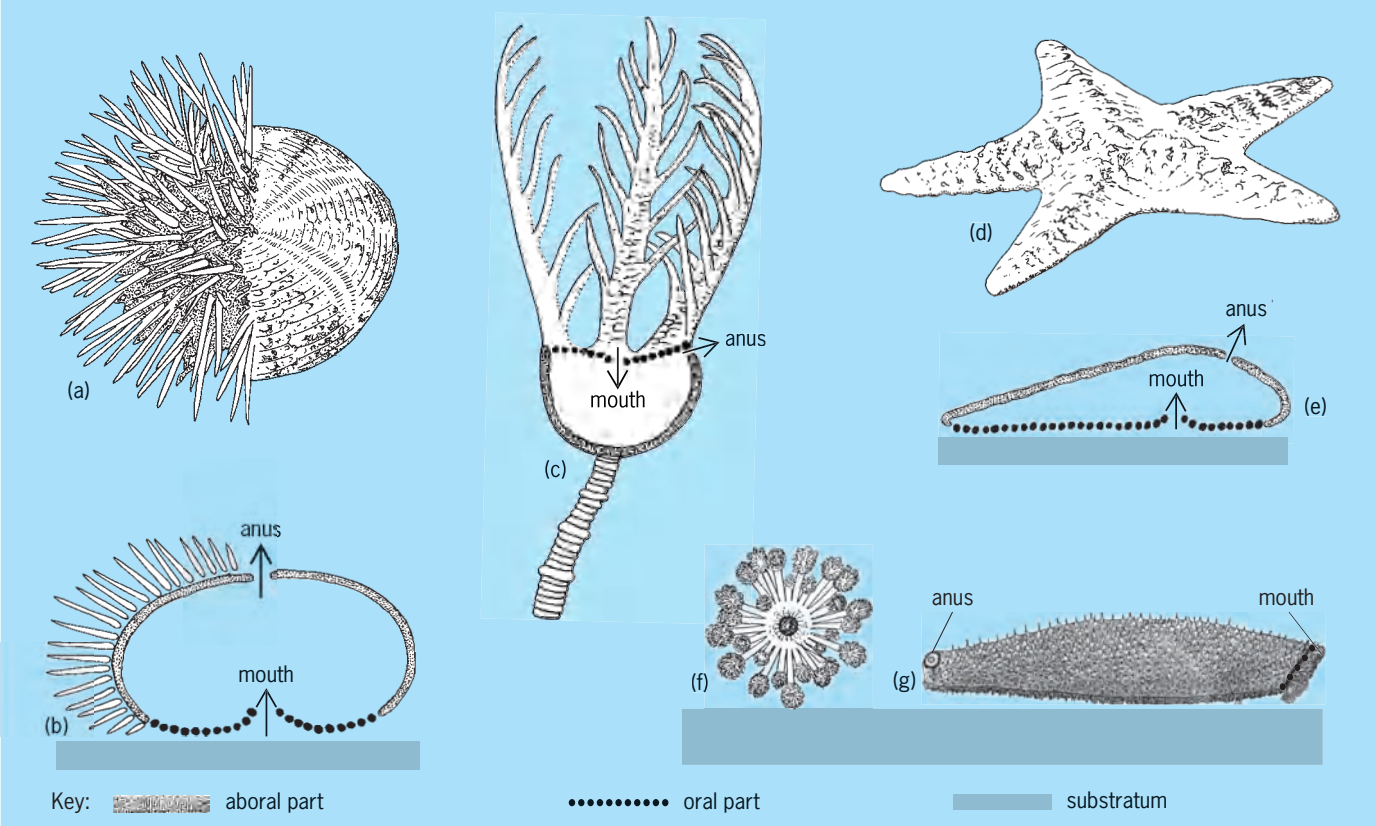

Fig. 5. Representative types of echinoderms. (*a*) Regular echinoid (*Lytechinus*, Recent), oblique aboral view, right half with spines removed, showing two ambulacra and three interambulacra. (*b*) Diagrammatic section through *a*, showing the stoutly built test. (*c*) Crinoid showing the stem and three arms, calyx sectioned. (*d*) Asteroid (*Dermasterias*, Recent), aboral oblique view. (*e*) Diagrammatic section through the ray at left in *d* and opposite the interambulacrum. (*f*) Holothuroid, oral view showing tentacles around the mouth. (*g*) Holothurian, lateral view.

symmetry in arrangement of skeletal parts and the flattened body form, from which one or more specialized slender appendages may extend. Plates enclosing the body are irregular in shape and size and vary in number between forms. There are four quite dissimilar classes of carpoids: Homostelea, Ctenocystoidea, Stylophora, and Homoiostela. Some experts, especially R. Jefferies, believe that some of the homalozoans, the Stylophora in particular, are not echinoderms at all, but are "calcichordates," groups forming the stem of the chordates. There has been considerable debate in the scientific literature about this matter, with persuasive evidence presented on both sides. Most paleontologists tend toward the idea that all homalozoans are true echinoderms. The homalozoans range in age from earliest Cambrian to Carboniferous. *See* CARPOIDS; HOMALOZOA.

Crinozoans. The crinozoans evidently include types of echinoderms least modified from the ancient progenitors of the phylum. Crinozoans range from Cambrian to Recent, but only the Crinoidea, abundant in Paleozoic formations, have survived to the present. In some classifications, the Crinoidea and Paracrinoidea are treated as the subphylum Crinozoa in the strict sense, and the other classes are grouped together in the subphylum Blastozoa. A characteristic crinozoan feature is orientation of the body with the oral (ventral) side directed upward and aboral (dorsal) side downward; thus, although commonly used, the designations ventral and dorsal are not suited to these particular echinoderms (**Figs. 5***c* and **6**). There are nine classes of crinozoans:

(1) Eocrinoidea (the crinozoan stem group) and (2) Paracrinoidea are crinozoan groups attached by a stalk. They have ovoid bodies enclosed by irregularly arranged plates that commonly lack well-developed radial symmetry and possess a combination of characters typical of the two following classes. (3) Diploporita and (4) Rhombifera (previously grouped in the polyphyletic Cystoidea, with saclike bodies, and each possibly polyphyletic) are a diverse group of primitive crinozoans distinguished especially by pores or tubular canals that penetrate irregularly arranged plates enclosing the body. Attachment to the substrate was by means of a stem composed of superposed discoid plates. (5) Crinoidea (sea lilies) include the most abundant crinozoans, most of the fossil forms anchored to the sea bottom by means of a short to very long stem and distinguished by upwardly directed arms which functioned for gathering food. Some ancient and most modern crinoids are stemless, known as feather stars, and adapted for a free-living existence. True radial symmetry and regularity of skeletal construction are attributes that distinguish crinoids. (6) Blastoidea (budlike forms) are small to

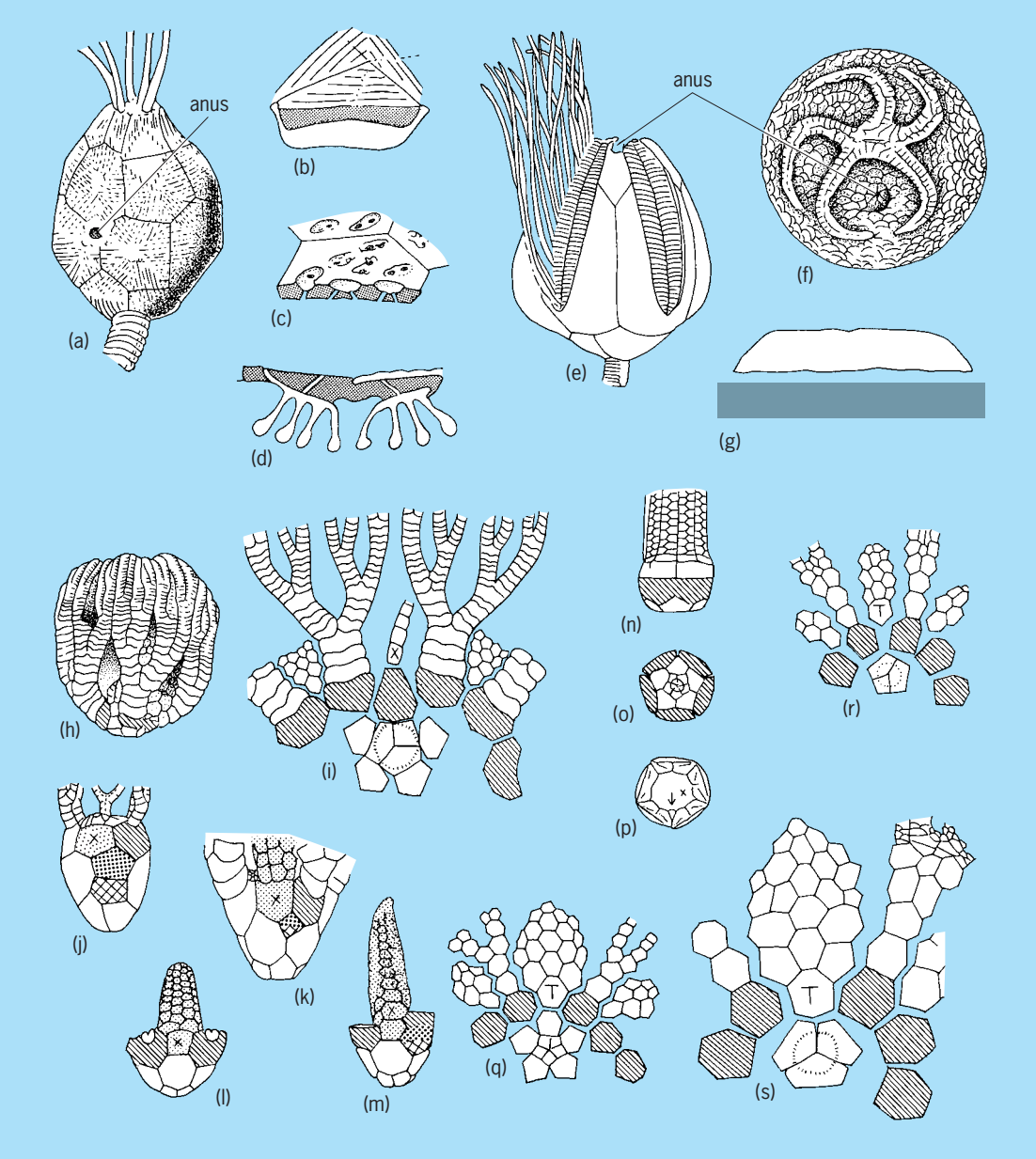

Fig. 6. Fossil pelmatozoan echinoderms. (*a*) Rhombiferan cystoid (*Echinoencrinus*, Ordovician). (*b*) Diagrammatic oblique view and section through a cystoid pore rhomb (half of one rhomb toward front bisected by a suture between plates). (*c*) Oblique view and section of a diploporitan cystoid plate. (*d*) Section through the ambulacrum of a blastoid (*Pentremites*, Mississippian) showing hydrospires beneath the lancet and side plates. (*e*) Blastoid (*Pentremites*) with brachioles restored along the side of one ambulacrum. (*f*) Edrioasteroid (*Carneyella*, Ordovician), oral view. (*g*) Section through a stalkless edrioasteroid. (*h*, *i*) Flexible crinoid (*Taxocrinus*, Mississippian), posterior side and partial plate diagram showing the right posterior plane of bilateral symmetry defined by the infrabasal circlet. (*j–p*) Inadunate crinoids, some showing anal sacs and parts of arms: (*j*) *Carabocrinus*, Ordovician; (*k*) *Botryocrinus*, Devonian; (*l*) *Cyathocrinites*, Mississippian; (*m*) *Cupulocrinus*, Ordivician; (*n–p*) *Delocrinus*, Pennsylvanian. (*q*) Diplobathrid camerate crinoid (*Ptychocrinus*, Ordovician), plate diagram. (*r*, *s*) Monobathrid camerates: (*r*) *Macrostylocrinus*, Devonian; (*s*) *Periechocrinites*, Mississippian, plate diagrams showing noteworthy distinctions in basal and radial circlets. (*Parts f, g after L. H. Hyman, The Invertebrates, vol. 4: Echinodermata, McGraw-Hill, 1955*)

medium-sized crinozoans having a highly developed radial symmetry and very regular arrangement of the few plates surrounding the body. The body was attached to the substrate by a very slender stem, which had a circular cross section. (7) Coronoidea is a small group of middle Paleozoic crinozoans, similar to blastoids and probably ancestral to them. Coronoids differ in the structure of the ambulacral plate and possess erect, as opposed to recumbent, ambulacra. (8) Parablastoidea is another small, principally Ordovician class of crinozoans with cystoid-like thecae but with uniserial ambulacra. (9) Edrioblastoidea comprises a single Ordovician genus with a theca that is blastoid-like in shape and symmetry but lacks brachioles or hydrospires (or pore rhombs). *See* CRINOZOA.

Eocrinoids. This small group of Cambrian to Silurian crinozoans combines characteristics of

cystoids and crinoids, yet differs significantly from both. They resemble cystoids in their mode of branching of the ambulacral grooves and ventrolateral location of the anus but lack thecal pores or distinct pore rhombs; they are like crinoids in plate structure and similarity of plate arrangement in the calyx (plated body). The eocrinoids have stems and unbranched or bifurcating arms. They were probably ancestral to all other crinozoans. *See* EOCRINOIDEA.

Paracrinoids. Paracrinoidea are stem-bearing echinoderms that also combine features of cystoids and crinoids, but in a manner quite unlike that of the eocrinoids. The paracrinoids, now known only from Middle Ordovician to Lower Silurian deposits, have numerous plates of the calyx that are not arranged in series and that lack a ventrally differentiated area corresponding to the tegmen of crinoids. The plates have a cystoid-like pore structure, but the arms are comparable to those of crinoids.

Cystoids. The Ordovician to Devonian cystoids, comprising the classes Diploporita and Rhombifera, are extinct crinozoans of globose, subcylindrical, or flattened ellipsoidal form that are characterized mostly by irregularity of the body plates (Fig. 6*a*) and a short, weak stem. A variable number of slender appendages (brachioles) on the upper side of the calyx, supplemented by ambulacral grooves, served for gathering food. The mouth was located centrally at the summit of the calyx, and an anus was fairly well down on one of the sides, which accordingly is defined as posterior. Other small orifices, interpreted as hydropore and gonopore, occur near the anus. In one class, Diploporita, the numerous calyx plates are perforated by numerous pairs of minute tubular openings that allow seawater to circulate through them (Fig. 6*c*). Remaining cystoids are named Rhombifera (rhomb-bearing forms) because some or all of their relatively few calyx plates are penetrated by rhomb-shaped groups of tubes running parallel to each other and to the plate surfaces (Fig. 6*a*). The two halves of any rhomb lie on adjoining plates, with the tubes crossing the plate boundaries at right angles (Fig. 6*b*). The pattern of calyx plates suggests that the blastoids and crinoids may be descendants of rhombiferan cystoids. *See* DIPLOPORITA; RHOMBIFERA.

Crinoids. In terms of abundance of fossil remains, the crinoids outrank all other echinoderms combined (Fig. 5*c*). There are about 5000 extinct species and at least 700 living kinds. Some Paleozoic deposits hundreds of feet thick in areas measured in hundreds of square miles are largely composed of fossil remains of trillions of crinoids. *See* CRINOIDEA.

Skeletal features. Although many modern and some ancient crinoids are stemless as adults, this group of pelmatozoans typically is attached to the sea bottom by a more or less elongate stem composed of superposed calcareous discs (columnals) perforated centrally by a circular, pentagonal, or pentastellate canal. At the opposite extremity of the stem, which exceptionally may be 15 m (50 ft) tall, is the crinoid body, encased in regularly arranged plates and surmounted by branched or unbranched free-moving arms. The conjoined plates below the free arms make up the so-called dorsal cup; this cup, along with the plates of the ventral surface composing the tegmen, makes up the crinoid calyx (Fig. 6*h–s*). The calyx and its attached arms are termed the crown. The mouth is located on the ventral surface. The anus lies on the tegmen, is raised above it on an anal sac (Fig. 6*k–m*), or may lie on the side of the dorsal cup. The posterior side of the crinoid is defined by the position of the anus in one of the interrays or by extra plates introduced in such a position on one side of the dorsal cup; the ray opposite to the posterior interray is defined as anterior. This establishes a plane of bilateral symmetry that more of less modifies the fundamental radial symmetry of the crinoid (**Fig. 7***a*).

The plates of the crinoid dorsal cup are arranged in a regular pattern of successive circlets, but with differences in various groups that furnish a basis for classification. Each circlet normally contains five plates. At the base of each ray is a radial plate. Beneath the circlet of radials are five basals that are interradial in position; some crinoids have a still lower circlet of infrabasals that alternate with the basals and hence occur in radial position. Crinoids with a single circlet of plates below the radials are termed monocyclic (Fig. 6*r* and *s*), and those with two circlets are dicyclic (Fig. 6*h–q*).

In some crinoids the tegmen is stoutly constructed of small, irregularly arranged plates, while in others it consists of a flexible, leathery integument that may be studded with calcareous ossicles. Five subequal oral plates larger than others of the tegmen may occur interradially around the mouth. The arms of crinoids are extremely variable in plan and construction. They may be unbranched or moderately to highly branched, with or without very numerous tiny branchlets called pinnules, and composed of a single or double series of arm plates (brachials). These characters, along with the mode of articulation among plates of the rays, are important also in classification.

Main types. Four subclasses of crinoids are recognized, three of them distributed from Ordovician to Permian; one of the three (Inadunata) persists to Middle Triassic. The fourth subclass (Articulata) ranges from the Mesozoic to present day.

The Inadunata ("not united" forms, referring to lack of incorporation of lower arm plates in dorsal cup) are crinoids with a relatively small dorsal cup containing one or two circlets of plates below the radials and having the arms entirely free above the cup (Fig. 6*j–p*). They include 1750 or more species that exhibit great variety in form and evolutionary trends. In a majority the anteroposterior plane of bilateral symmetry is well marked, without other deviation from a regular pentameral plan (Fig. 7*a*), but in one group (superfamily Homocrinacea) a surprising degree of bilateral symmetry was developed in the plane of the left anterior ray (Fig. 7*c*). Some of these crinoids have a strongly downbent crown that was hinged on the summit of the stem. In another group (superfamily Heterocrinacea) a subordinate plane of

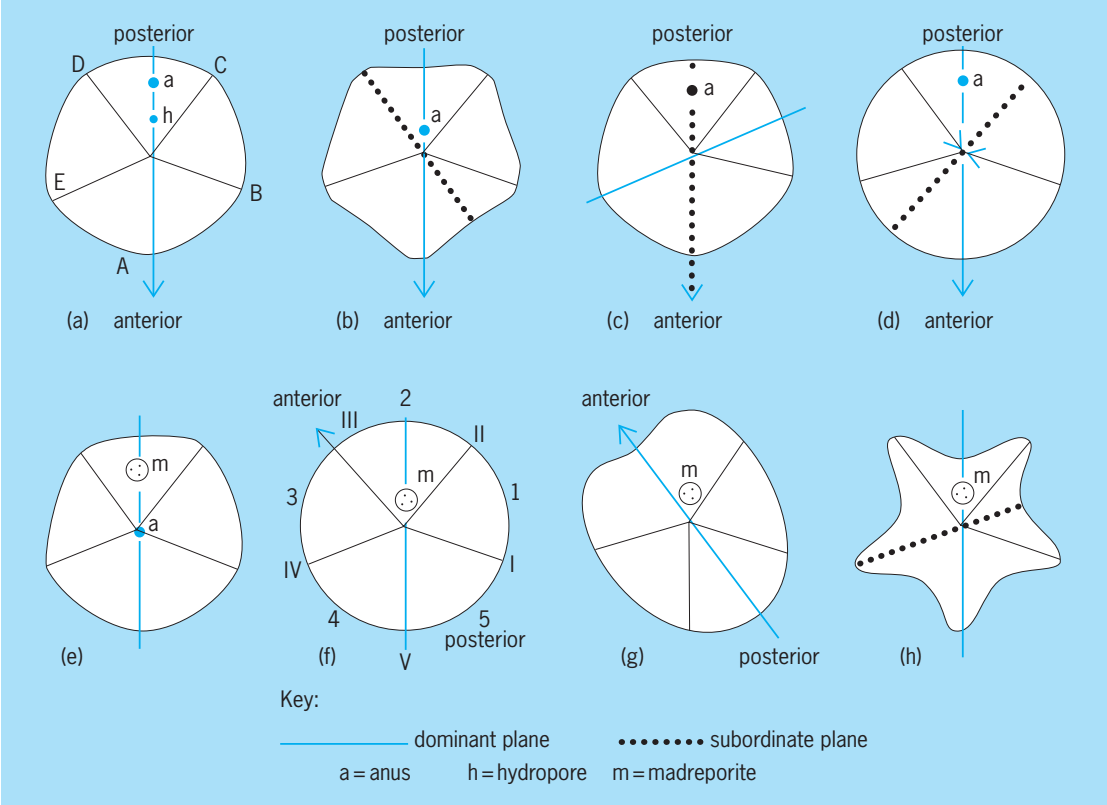

Fig. 7. Bilateral symmetry developed in various echinoderms. Anterior and posterior directions are indicated where distinguished. All diagrams represent aboral views. (*a*) Crinozoans (in general, most crinoids, cystoids, edrioasteroids); letters denote ray designations according to the Carpenter (P. H. Carpenter, 1884) system: A, anterior; B, right anterior; C, right posterior; D, left posterior; E, left anterior. (*b*) Blastoids, heterocrinoids, with subordinate bilateral symmetry in the left posterior plane. (*c*) Homocrinoids, with primary bilateral symmetry in the left anterior plane. (*d*) Glyptocrinoids, flexible crinoids, and rhombiferan cystoids, with subordinate bilateral symmetry in the right posterior plane. (*e*) Holothuroids. (*f*) Regular echinoids, rays marked according to the Lovén (S. Lovén, 1874) system, III being considered anterior. (*g*) Irregular echinoids, with prominent bilateral symmetry in the left posterior plane of crinozoans. (h) Asteroid, with subordinate bilateral symmetry in the left anterior plane.

bilateral symmetry is oriented in the left posterior plane (Fig. 7*b*). *See* INADUNATA.

The Camerata (chamber or "box" forms) are most numerous among ancient crinoids in both variety and quantity of individuals (Fig. 6*q–s*). More than 2500 species of these fossils have been described, two-thirds of which come from Mississippian rocks alone. The camerates are distinguished by the stout construction of their calyx, which incorporates lower ray plates and interradials in the dorsal cup, and subtegminal location of the mouth. Types with both one and two circlets of plates beneath the radials are common. Anteroposterior bilateral symmetry is developed almost universally in these crinoids as a modification of the dominant pentameral pattern (Fig. 7*a*); a secondary plane of bilateral symmetry directed through the right posterior ray prevails in the monobathrid suborder Glyptocrinina, as in all flexible crinoids (Fig. 7*d*). *See* CAMERATA.

The Flexibilia (flexibles) are a distinctive assemblage of exclusively dicyclic crinoids characterized by movable ligamentous union between most of the plates and several constant features in the organization of the calyx (Fig. 6*b* and *i*). They include approximately 300 described species, all of which exhibit a secondary bilateral symmetry in the right posterior plane in addition to their generally well-marked primary bilateral symmetry directed anteroposteriorly (Fig. 7*d*). *See* FLEXIBILIA.

The Articulata (forms divided into joints) comprise Mesozoic and Cenozoic crinoids, represented by about 500 fossil and about 700 living species. The stalked articulate crinoids have been classified into four orders: Millericrinida, Cyrtocrinida, Bourgeticrinida, and Isocrinida. All of these orders have fossil and extant representatives. Some experts have abandoned the order categories for the 100 species of living stalked crinoids and prefer to arrange them into 11 families. There are three orders of unstalked crinoids: Uintacrinida and Roveacrinida are extinct, while the Comatulida (feather stars) has about 600 living species. Extant sea lilies are confined to depths below about 100 m (330 ft). The stalk is usually attached to a hard substratum—a rock, or a piece of shell embedded in soft mud—by means of a terminal disc cemented to the substratum or by means of claw-like whorls of cirri. In a few species, the end of the stalk is branched and rootlike for anchoring the stalk in soft sediments. When detached from the substratum and lying on the seafloor, sea lilies are capable of slowly moving to a new site. In contrast, the feather stars are

highly mobile, often capable of active swimming by thrashing of the arms in an up-and-down motion. Feather stars range from shallow waters to great ocean depths. They are common on coral reefs, and in certain suitable habitats can reach population densities of hundreds of individuals per square meter. Sea lilies and feather stars are passive suspension feeders; they extend their arms into the water and capture small drifting organisms and particulate matter on their sticky tube feet. Captured items are conveyed to the mouth along food grooves by means of ciliary action. In extant crinoids, the gonads develop inside the pinnules near the bases of the arms, causing conspicuous swelling. At the appropriate time for breeding, the pinnules burst to release eggs or sperm into the surrounding seawater. *See* ARTICULATA (ECHINODERMATA).

Blastoids. Blastoidea, Ordovician to Permian, display a regular, fivefold radial symmetry (Fig. 6*e*), on which are superposed a first order of bilateral symmetry in the anteroposterior plane and a second order directed through the left posterior ray (Fig. 7*b*). The calyx of average-size specimens is small, with diameter of about 15 mm (0.6 in.) and height of 20 mm (0.8 in.). It is composed of 23 or 24 plates, of which five (lancet plates) are not visible externally on unweathered specimens. Within forks of the radials are a multitude of very diminutive so-called side plates that conceal the lancets; they form ambulacral areas bordered laterally by rows of threadlike free armlets (brachioles) that function as food gatherers. Particles of food are conducted to midlines of the ambulacra and thence upward to the mouth, which is at the center of the ventral surface. Beneath the ambulacra are extremely delicate, longitudinally folded, calcareous lamellae that enclose narrow troughs for circulation of water admitted through pores at the base of the brachioles; these structures, termed hydrospires, correspond to the slitlike parts of pore rhombs in the rhombiferan cystoids (Fig. 6*d*). More than 500 species have been described. *See* BLASTOIDEA.

Asterozoans. The asterozoans are star-shaped, free-moving echinoderms with well-marked radial symmetry in arrangement of their skeletal parts. They are easily recognized by their strongly developed arms and central disk (Fig. 5*d* and *e*). The arms of most asterozoans, with their tube feet functioning for locomotion and predation, are highly mobile. They range in age from Early Ordovician to Recent but are more abundant and varied in modern seas than in the fossil record. They are divided into four classes: Somasteroidea, Asteroidea, Concentricycloidea, and Ophiuroidea. The extinct Somasteroidea ranged from the Ordovician to the Devonian. Somasteroids resemble true sea stars in general shape, with broad arms, but the arms are unique in having a pinnate arrangement of rodlike plates. Asteroidea (sea stars) are distinguished by prominent flexible rays that join the central disk without clearly shown demarcation (Fig. 5*d* and *e*). Concentricycloidea are flattened and discoidal with no evident arms. The Ophiuroidea (brittle stars) have long, slender, snakelike arms which are sharply set off from the central discoid body. They are very active echinoderms, able to crawl rapidly in any direction.

Somasteroids. These rare, broad-armed sea stars are especially characterized by the featherlike arrangement of parallel, rodlike plates in their wide petaloid rays. The rods extend outward from the medially placed ambulacral plates. A food-carrying groove is contained in each ray. The aboral side of the skeleton has a rough, coarse meshwork.

Asteroids. The sea stars have a long fossil history, ranging back to the Ordovician. Sea stars are comparable to the ophiuroids in their general stellate form (Fig. 5*d* and *e*, Fig. 7*b*). They differ from the brittle stars in two main respects: lack of strongly marked separation between the body and its radially disposed hollow arms, and the open ambulacral grooves along the ventral side of the arms. The skeletal elements tend to be loosely joined; consequently, asteroids are not well suited for fossilization in a manner showing the skeletal arrangement of the entire animal. Sea stars may range in size from a few millimeters to over a meter (3.3 ft) in diameter. Ossicles along the ambulacra occur in two or four series differentiated as ambulacrals and adambulacrals. Some asteroids lack obvious arms and are essentially pentagonal with ambulacral grooves on their ventral (oral) side. Others, such as the Heliasteridae, develop as many as 44 arms as they grow, losing their fundamental pentameral symmetry but structurally resembling common types of asteroids. A single madreporite lies on the upper (aboral) surface, placed between the arms (interradially). An anus is present in the center of the aboral surface in most groups and absent in others (notably the Paxillosida).

Behavior. Asteroids are universal symbols of the ocean. They are conspicuous in a great variety of habitats, from mud to sand to rock to coral reefs. In some areas they can be present in great numbers, causing local habitat damage. The multiarmed, crown-of-thorns sea star, reaching 1 m (3.3 ft) in diameter, is common on Indo-Pacific coral reefs. It can move across reefs in large swarms, eating the soft coral tissues and causing short- to long-term devastation of reefs. Some groups of sea stars, notably the Paxillosida, have pointed rather than suckered tube feet, and they tend to swallow prey whole (such as small clams and sand dollars), ejecting the shells from the mouth when digestion is complete. The Forcipulatida are notorious predators; a forcipulatid may attach its suckered tube feet to the two shells of a clam or a mussel and then, for several hours if necessary, gently pull on the two shells, until the bivalve's muscles tire, the shells open a fraction of an inch, and the sea star drops its stomach into the bivalve, digesting the animal inside its own shell. Some deep-sea asteroids are mud-swallowers, digesting from the mud whatever organic material may be present.

Classification. The classification of the Asteroidea has been revised by many authors over the past 50 years or so. In recent years, the most commonly used classification of both fossil and extant forms is that proposed by D. B. Blake. Extinct

Paleozoic orders are Platyasterida, Pustulosida, Hemizonida, and Uractinida. Post-Paleozoic sea stars include the extinct order Trichasteropsida, and seven extant orders, all of which also have fossil representatives: Paxillosida, Notomyotida, Velatida, Valvatida, Spinulosida, Forcipulatida, and Brisingida. *See* ASTEROIDEA; FORCIPULATIDA; NOTOMYOTIDA; PAXILLOSIDA; VALVATIDA.

Concentricycloids. The sea daisies comprise three extant deep-sea species, two known from the Pacific Ocean and one from the Atlantic. Sea daisies are flattened, discoidal echinoderms up to 15 mm (0.6 in.) in diameter. The upper surface is covered with delicate overlapping plates; the lower surface has a mouth frame and a single peripheral ring of tube feet. These animals have been found only in association with pieces of wood on the deep-sea floor. Presumably, they attach to the wood substratum using their suckered tube feet. This class was discovered and diagnosed in 1986, and its formal status has been debated ever since. Several scientists have argued persuasively that the concentricycloids should be placed in the Asteroidea, while others maintain that they represent a separate class of echinoderms.

Ophiuroids. The brittle stars or serpent stars composing the Ophiuroidea are highly mobile echinoderms closely related to the sea stars. They are distinguished readily by their external form, since the small, rounded to pentagonal or scalloped disk is sharply set off from the symmetrically placed long, slender arms that extend radially outward from it. The arms may be smooth or spiny. Almost invariably they are five in number and unbranched, but in a few kinds, such as the Recent basket stars (Gorgonocephalidae), the arms are repeatedly bifurcated. All known fossil ophiuroids have simple arms. The skeleton of the central disk is composed of many regularly arranged plates, without any deviation on either the aboral or oral surfaces from perfect pentamerous symmetry. The mouth is at the center of the oral surface; an anus is lacking. A single interradial madreporite is present on a plate near the mouth; in some basket stars there are five interradial madreporites around the outer margin of the mouth skeleton. The arms are solid, not hollow as in asteroids. The arm skeleton is distinctive in being internal, consisting of "vertebrae" formed from fused ambulacral ossicles. Brittle stars are common in a great variety of habitats—living under rocks, concealed in coral reefs, or scattered in enormous numbers across great expanses of mud on the deep-sea floor. Many brittle stars are selective detritus feeders; others feed by extending their arms into the seawater and capturing small drifting organisms or particles on their sticky tube feet. Recent investigations have shown that certain brittle stars have well-developed eyes in their upper-arm skeletons, and they can respond rapidly to changing light regimes, seeking shelter in crevices as necessary.

Fossil ophiuroids are known from Lower Ordovician to Pleistocene (Recent), but they are not abundant. Fewer than 100 species have been described from Paleozoic, Mesozoic, and pre-Recent Cenozoic formations, as compared with approximately 2000 known living species. Ophiuroids are arranged in four orders: the extinct Stenurida (Ordovician to Devonian) and Oegophiurida (Ordovician to Carboniferous), and the extant Phrynophiurida and Ophiurida (Ordovician to Recent). *See* OPHIUROIDEA.

Echinozoans. These echinoderms generally have an ovoid or globose body, and they lack armlike appendages. Some echinozoans may have a low discoid body or an elongate cylindrical to almost wormlike form. All have conspicuous radial symmetry of their skeletal parts, and in almost all the symmetry is pentamerous. The skeleton of some echinozoans consists of a rigidly constructed test to which movable jaw parts and external spines are attached, whereas in others all skeletal parts either are joined together flexibly or are reduced in size and separated from one another by leathery tissue. The echinozoans comprise seven classes: Helicoplacoidea, Camptostromatoidea, Edrioasteroidea, Cyclocystoidea, Ophiocistioidea, Echinoidea, and Holothuroidea, the first five of which are long-extinct groups, while the others are represented by both fossil and living forms. (1) The Helicoplacoidea are Lower Cambrian forms with a flexible, expansible test composed of spirally arranged plates. (2) The Lower Cambrian Camptostromatoidea are conical or domal animals with plates of varying size overlapping on the lower theca. (3) Edrioasteroidea ("sessile sea stars") range from the Ordovician to the Permian. They have a many-plated test, ranging from rigid to somewhat flexible, of generally discoid shape with five distinct, straight or curved rays on the upper surface. Some were able to move about, but nearly all are found attached to some hard foreign object, such as a brachiopod or other invertebrate shell. (4) The Cyclocystoidea are Ordovician-to-Devonian echinoderms with a disk-shaped body, the upper surface of which is covered by plates arranged in concentric rings. (5) The Ophiocistioidea are Ordovician-to-Devonian echinoderms that differ from other classes in having large fine-plated tube feet, but share some skeletal features with the Holothuroidea, with which they have sometimes been aligned. (6) The Echinoidea (sea urchins) are characterized by a rigid to flexible skeleton of varied shapes, though mostly ovoid to subglobose, covered by numerous long or short movable spines (Fig. 5*a* and *b*). Pentameral symmetry is strongly developed, modified in all orders by more or less distinct bilateral symmetry and in some by differentiation of anterior and posterior regions. Very abundant in present-day seas, sea urchins range from Early Ordovician onward. (7) The Holothuroidea (sea cucumbers) are soft-bodied, leathery-skinned echinozoans with an elongate cylindrical body and skeletal parts usually consisting of discrete, generally microscopic, ossicles ("little bones") embedded in the body wall and internal tissues (Fig. 5*f* and *g*). *See* ECHINOZOA.

Edrioasteroids. A unique assemblage of mainly early Paleozoic attached echinoderms (Fig. 6*f* and *g*), Edrioasteroidea were distributed from Lower Cambrian to Upper Carboniferous. They had no stem for anchorage but were fixed by their entire base. The upper (ventral) side was covered by many

flexibly joined plates, which are sharply differentiated into rows of paired ambulacrals with adjoining adambulacrals and irregular interambulacrals. The five ambulacral tracts curve outward from the centrally located mouth, with their extremities often deflected rather strongly in a consistent way. The anus lies in one of the interrays that accordingly is regarded as posterior. *See* EDRIOASTEROIDEA.

Helicoplacoids. Numerous representatives of Helicoplacoidea have been found in the oldest fossil-bearing strata (*Olenellus Zone*) of the Lower Cambrian at localities in western Nevada and southern California. The body is ovoid to pyriform, enclosed in a test of loosely joined, spirally arranged rows of elongate plates, with columns, interpreted as ambulacral, and intervening ones, as interambulacral, originating at an apical pole and extending to the opposite oral pole. The test could be distended and retracted so as to change its shape and volume. Other echinoderms associated with these oldest echinozoans include representatives of the pelmatozoan class Eocrinoidea. *See* HELICOPLACOIDEA.

Echinoids. The class Echinoidea is the most important group of fossil eleutherozoan echinoderms (Fig. 5*a* and *b*; **Fig. 8**). They are distinguished from crinozoans by freedom from a sessile mode of life, and they share with the holothuroids a downward orientation of their oral surface. Modern species of echinoids total approximately 800, as shown by T. Mortensen's several-volume monograph (1928–1952) covering all known living kinds. In comparison, 2000 species of sea stars, 2000 species of brittle stars, and 1200 species of sea cucumbers are alive today. All eleutherozoan groups are known as far back as Ordovician time, and yet fossil echinoids are more numerous than all others of the assemblage combined. Such disparity reflects inequality in fitness of the different sorts of skeletons to be preserved. Not all parts of echinoid skeletons may be found intact. Multitudes of movable spines attached to the test during life tend to be separated and scattered, as does the echinoid masticatory apparatus, which is known as the Aristotle's lantern. Thus, knowledge of the entire skeletal structure of most fossil echinoids is incomplete. *See* ECHINOIDEA.

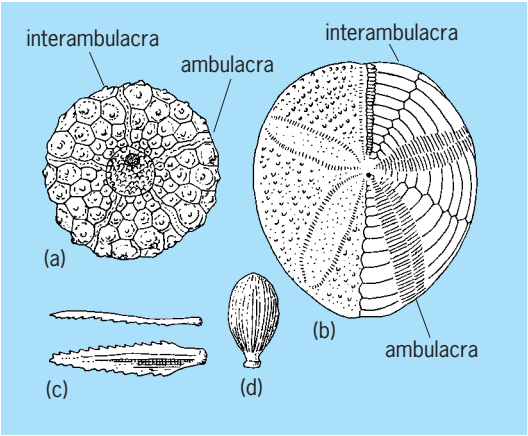

Fig. 8. Fossil echinoids. (*a*) Regular form (*Echinocrinus*, Mississippian), aboral view showing the anus and periproct, wide interambulacra and narrow ambulacra. (*b*) Irregular form (*Eupatagus*, Paleogene), aboral view, left side showing an unweathered appearance with many spine bases, right side showing the plate structure with ornamentation omitted; anterior ray directed upward. (*c*, *d*) Types of echinoid spines: (*c*) *Porocidaris*, Paleogene; (*d*) *Balanocidaris*, Jurassic.

Morphology. The echinoid test is typically globular, unevenly ovoid, or discoid in shape (Figs. 5 and 8). Mainly it is formed by 10 meridionally disposed bands of plates, of which five are ambulacral, distinguished by porelike openings between or through the plates for passage of the tube feet; the five bands of plates that alternate with the ambulacrals are interambulacral and lack perforations for tube feet. In most echinoids, both living and fossil, the plates are joined together rigidly, but in a minority the union is somewhat flexible and imbricating (having overlapping edges). Virtually all post-Paleozoic echinoids have a constant arrangement of the ambulacral and interambulacral groups of plates, each of the 10 bands comprising a double column, making 20 columns of plates in the test as a whole. The plates are widest at the equator (ambitus) of the test and increasingly narrow toward the oral and aboral poles.

The mouth, located on the underside, is surrounded by a peristome, an area of naked integument or flexibly united small plates. The anus generally is placed on the aboral side of the test opposite the mouth, but it may be shifted backward to a posterior position or even to the oral surface. The anus is surrounded by a bare or small-plated periproct. Another part of the test important for orientation and classification is the so-called apical system of 10 special plates at the aboral pole, five so-called ocular plates alternating with five genital plates. One of these genital plates is an interambulacrally placed sieve plate (madreporite) that is a landmark for identifying homologous parts, not only among diverse kinds of fossil and living echinoids but among echinoderms generally. Fossil remains of echinoids include numerous spines not associated with the test that bore them. Even so, the spines are commonly distinctive and useful as paleontological tools (Fig. 8*c* and *d*).

Orientation. Seemingly, there should be little reason for discussing the orientation of echinoids because the downward-directed oral surface clearly is ventral and, at least among irregular echinoids which are characterized by obvious bilateral symmetry, locomotion consistently is in a single direction along the path defined by the plane of this symmetry. If such direction of movement is considered to be forward, the ambulacral ray on this side of the test must be anterior and the opposite interambulacrum is posterior (Fig. 7*g*). In irregular echinoids, such as the spatangoids, the spines nearly all point backward. In many irregular genera, the mouth is found to have shifted forward and the anus rearward. These observations have led to adoption by specialists of a scheme of numerical designation of the ambulacral and interambulacral rays, introduced by S. L. Lovén (1874) and known as the Lovén system (Fig. 7*f*). It is applied to fossils as well as living echinoids and to both irregular and regular types. Orientation of

the regular echinoids seemingly should not be easy, since they move in any direction with equal ease and without detectable preference; also, except for the madreporite, the test has perfect radial symmetry. Inclusion of the regulars is possible only by use of the off-center location of the madreporite on the aboral surface as a reference point, assuming that relative position is the same in all echinoids. The interambulacrum with the madreporite is identified as number 2 of the Lovén system.

Among crinozoans, which almost certainly embrace the ancestors of echinoids, the posterior interradius is readily and positively identified. If this orientation is applied to echinoids, a discrepancy becomes evident at once, for the adopted anteroposterior plane of echinoids clearly is that coinciding with the left posterior ray of the crinozoans (Fig. 7*e* and *f*). Interambulacrum 2 (Lovén), considered as right anterior by echinoid students, is equivalent to the posterior interambulacrum of crinoids, for example. This does not mean that echinoid orientation is wrongly conceived but merely that evolution of these echinoderms has pursued a divergent path of its own, which incidentally duplicates the subordinate plane of bilateral symmetry in some crinozoans (blastoids, heterocrinoids).

Habitats and behavior. Echinoids can be extremely common from the rocky intertidal down to abyssal ocean depths. The irregular urchins, such as sand dollars, sea biscuits, and heart urchins, are always associated with soft substrates—sand, sandy mud, or mud. Sand dollars and sea biscuits live at the surface of sand or burrow just beneath the surface. Many of these echinoids ingest sand, obtaining sustenance from small organisms, such as bacteria, associated with the sand grains. Heart urchins burrow into the substratum and remain buried, feeding on particulate matter in the mud. The regular urchins are usually associated with hard substrates, typically rock, coral, or rubble, but numerous species may occur on sand or mud in quiet waters. Most of the regular urchins are vegetarians, feeding primarily on algae. Typically, echinoids reproduce by liberating eggs or sperm into the seawater from genital pores on top of the body. There are usually five pores, but in irregular urchins there may be four, three, or two pores. Fertilization of eggs takes place in the seawater, and the developing embryo becomes a planktonic larva. Within a few weeks, the urchin develops on the left side of the larva, breaks away, and sinks to the bottom as a juvenile. Some urchins have a nonfeeding yolky larva, and in others the larval stage is omitted altogether, the young being brooded by the female parent in special chambers.

Classification. Former taxonomic arrangement of the echinoids recognized two main subclasses, the regular echinoids (Regularia) and irregulars (Irregularia). Present knowledge, based on fossils as well as living forms, shows that this arrangement is artificial. The so-called regulars are a composite assemblage that on the one hand contains several extinct primitive kinds, exclusively Paleozoic, and on the other, advanced forms of modern type, all Mesozoic and Cenozoic, that include the stocks (Diadematacea) from which irregular echinoids were derived. Accordingly, classification was considerably revised (J. W. Durham and R. V. Melville, 1957; A. B. Smith, 1984; and others) so as to take account both of Mortensen's comprehensive work on living echinoids and of evidence from paleontology.

Ophiocistioids. Ophiocistioidea is an aberrant echinozoan group containing five known genera distributed from Lower Ordovician to Middle Devonian. The chief peculiarity of the group is the presence of numerous armlike appendages attached to the oral surface in an ambulacral position. They are interpreted as much-enlarged podia that are covered by abundant minute plates, a character seemingly shared with at least one early holothurian. Small ossicles in the ophiocistioid body wall closely resemble ossicles of holothurians. An early relationship with the holothuroids has been suggested.

Holothuroids. The sea cucumbers, or Holothuroidea, are a group of echinozoans that must have diverged very early from other echinoderm stocks. They clearly exhibit a fivefold radial symmetry characteristic of the phylum but otherwise differ radically from any other echinoderm assemblage. They are greatly elongated along the oral-aboral axis so as to have a subcylindrical form, and they lie on one of their sides identified as ventral because it is in contact with the substratum (Fig. 5*f* and *g*, Fig. 7*e*). The mouth, surrounded by feeding tentacles, lies at one extremity and the anus at the opposite end. The tentacles can assume a variety of forms, from richly branching (dendritic) to shield-shaped, to featherlike, to digitiform (fingerlike). Sea cucumbers are unique among living echinoderms in possessing an internal madreporite and a single gonad. In a small percentage of species, the body is enclosed in a test of conspicuous plates, but in most the body is flexible, the skeleton consisting of microscopic calcareous plates loosely distributed in the dermis. When the animal dies, decay of the uncalcified tissues liberates the skeletal parts, which almost invariably become scattered about. Less than a dozen whole-animal fossil holothuroids have been described. The fossil remains of the holothuroids usually consist only of discrete microscopic ossicles. Even so, about 200 species of these ossicles have been described from rocks ranging from Devonian to Pleistocene. Most of the fossil forms can be referred to one or the other of the orders of extant holothuroids. Sea cucumbers range from rocky shores to the greatest ocean depths. On the deep-sea floor they can be present in enormous numbers and may comprise 95% of the total weight of animals on the seafloor. Typically sea cucumbers ingest the soft sandy to muddy substrata on which they occur. Such feeding may be nonselective mud-swallowing or selective detritus ingestion. The sea cucumbers with branching tentacles are suspension feeders, extending their sticky tentacles into the water and capturing small drifting organisms or organic particles.

About 20 tropical species are prized as food, especially in Asia, and as a consequence these species are more or less fished out throughout the tropics. Six orders of living sea cucumbers are recognized,

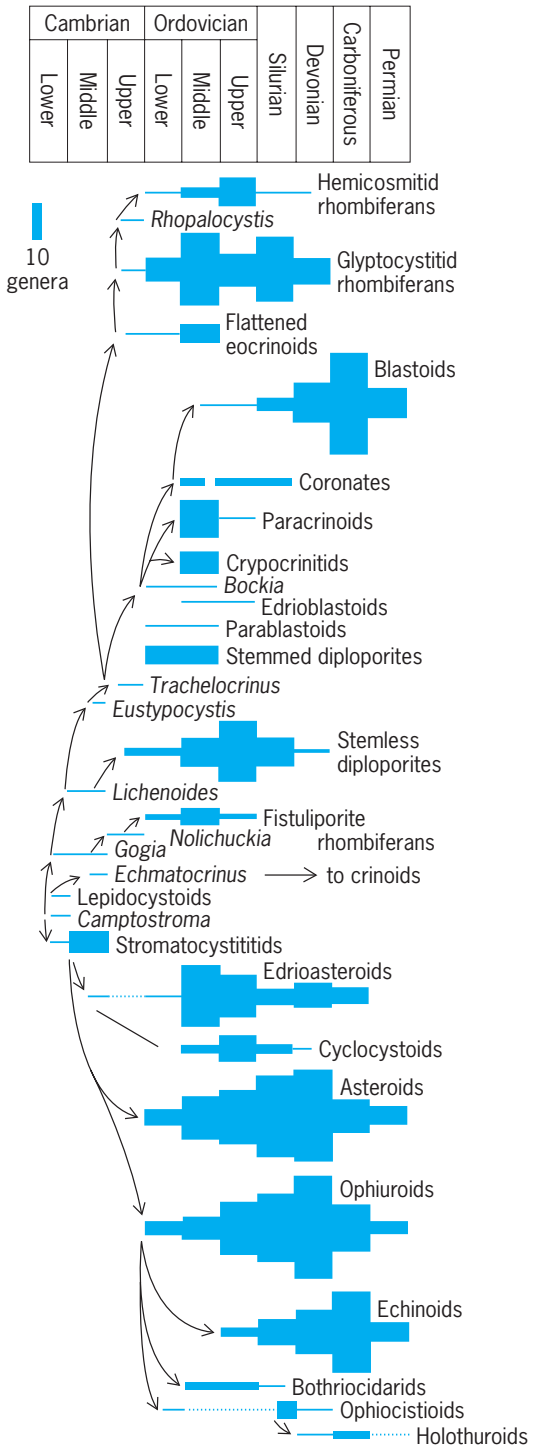

Fig. 9. Phylogeny and diversity range chart for the Paleozoic echinoderms. (*After C. R. C. Paul and A. B. Smith, The early radiation and phylogeny of echinoderms, Biol. Rev., 59:443–481, 1984*)

based upon characters of the feeding tentacles and skeleton. *See* HOLOTHUROIDEA.

Origin and Phylogeny

The evolution and relationships of the major groups of echinoderms are under active discussion, and several major issues have yet to be resolved. The extraxial/axial theory (EAT) of B. David and R. Mooi offers a new approach to discussion of skeletal homologies. Echinoderms evolved very rapidly near the beginning of the Paleozoic Era, and Lower Cambrian deposits contain such divergent branches of the phylum as homalozoans Helicoplacoidea, Edrioasteroidea, and Eocrinoidea (**Fig. 9**). These are primitive types of echinoderms. Cystoids, crinoids, and blastoids, as well as all recognized main groups of asterozoans and echinozoans (except holothurians), appear in Ordovician strata. During the Paleozoic, numerous well-marked evolutionary trends are discernible in nearly all echinoderm groups, including free-moving forms (especially echinoids) as well as crinozoans. Many small classes of echinoderms became extinct during the Paleozoic, and the surviving groups, especially the crinoids, lost many members at the Late Permian mass extinction. All groups of modern echinoderms have their origin in early Paleozoic stocks, and the lines of their phylogeny are mostly indicated by the fossil record. Echinoids predominate in Mesozoic and Cenozoic echinoderms.

David L. Pawson; Raymond C. Moore;
J. John Sepkoski, Jr.

Bibliography. B. David and R. Mooi, Embryology supports a new theory of skeletal homologies for the phylum Echinodermata, *C.R. Acad. Sci. Paris Ser. 3*, 319:577–584, 1996; L. H. Hyman, *The Invertebrates: Echinodermata*, vol. 4, 1955; D. T. J. Littlewood, Echinoderm class relationships revisited, in *Echinoderm Research*, ed. by H. Roland et al., A. A. Balkema, Rotterdam, 1995; C. G. Messing, Living comatulids, *Paleontol. Soc. Pap.*, 3:3–30, 1997; R. Mooi, Not all written in stone: Interdisciplinary syntheses in echinoderm paleontology, *Can. J. Zool.*, 79:1209–1231, 2001; C. R. C. Paul and A. B. Smith (eds.), *Echinoderm Phylogeny and Evolutionary Biology*, 1988; A. B. Smith, Echinoderm larvae and phylogeny, *Annu. Rev. Ecol. Sys.*, 28:219–241, 1997; A. B. Smith, *Echinoid Palaeobiology*, 1984.

Echinoida

An order of Echinacea with a camarodont lantern, smooth test, imperforate noncrenulate tubercles, ambulacral plates of echinoid type, and shallow

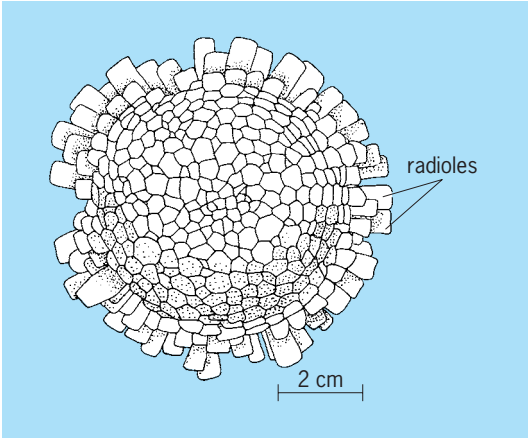

***Colobocentrotus atratus*, aboral aspect, a Pacific species adapted for life on wave-exposed coral reefs.**

branchial slits. There are numerous tropical and temperate species, some of them remarkably adapted to living on coral reefs (see **illus.**). *See* ECHINOIDEA.

The five included families are principally distinguished by characters of the pedicellariae. The Parasaleniidae are oblong forms with trigeminate ambulacral plates. They range from the Eocene to the present day. The Echinidae possess trigeminate or polyporous plates with the pores in a narrow vertical zone. Strongylocentrotidae are polyporous, with the pores in 2–4 vertical series. The Echinometridae show a variety of forms, which include polyporous types with an oblong test, and trigeminate or polyporous types with a spherical test. *See* ECHINACEA; ECHINODERMATA. Howard B. Fell

Bibliography. A. Smith, *Echinoid Palaeobiology*, 1984.

Echinoidea

A class of Echinodermata known as the sea urchins, also including sand dollars, sea biscuits, and heart urchins. In echinoids the body is enclosed in a hard shell, or test, formed from regularly arranged plates that bear movable spines (**Fig. 1**). There are no arms, but radii are represented by five double rows of tube feet arranged as meridians between the upper and lower poles of the body.

There are about 800 living species, and some 5000 fossil species have been recorded, included in 225 genera. Echinoids are classified in 15 orders, which are grouped in three subclasses. Sea urchins range in size from approximately 5 mm (0.2 in.) across the test to 20 cm (8 in.). Although many species are dull or dark in color, some are brilliant shades of purple, red, green, or orange. Others species have particolored striped spines, and deep-sea forms may be white. The earliest echinoids occur in the Late Ordovician. Morphological and molecular analyses suggest that the echinoids are most closely related to the holothuroids (sea cucumbers) within the Echinodermata. *See* ECHINODERMATA; EUECHINOIDEA; SAND DOLLAR; SEA URCHIN.

Relation to humans. Most sea urchins are harmless, although some tropical species have hollow, brittle spines that cause septic wounds if they break off after penetrating the skin. The genus *Araeosoma* carries venomous spines that can inflict dangerous wounds, and a related genus, *Asthenosoma*, is said to be able to kill an adult human. The egg masses, or roe, of sea urchins are highly prized in Asian cuisine. As a consequence, active fisheries for sea urchins and sea urchin farms exist in the United States, Chile, and other countries. *See* ECHINOTHURIOIDA.

Ecology. Sea urchins occur in all seas from low-tide level downward. *Pourtalesia* reaches a depth of more than 4.5 mi (7250 m) in the Banda Trench. The rounded sea urchins feed mainly on algae, often being partially covered by day under stones, shells, or pieces of seaweed held over the test by tube feet, and emerging at night or at high tide. Heart urchins, sand dollars, and the like live buried in mud or sand, feeding on organic matter associated with ingested sand or selectively feeding on detritus.

Sea urchins move slowly, using muscles at the bases of the spines to swing the spines like stilts. The tube feet are used in concert with spines to aid in locomotion. The suctorial tube feet can be used to ascend steep surfaces and act as anchors.

Numerous sea urchin parasites are recorded. Among these are protozoans, nematodes, and gastropods, of which several genera burrow into the test. In a few species, crabs may live in the rectum or on the test, feeding on spines and tube feet. Other animals shelter among the spines or attach themselves to exposed hard parts.

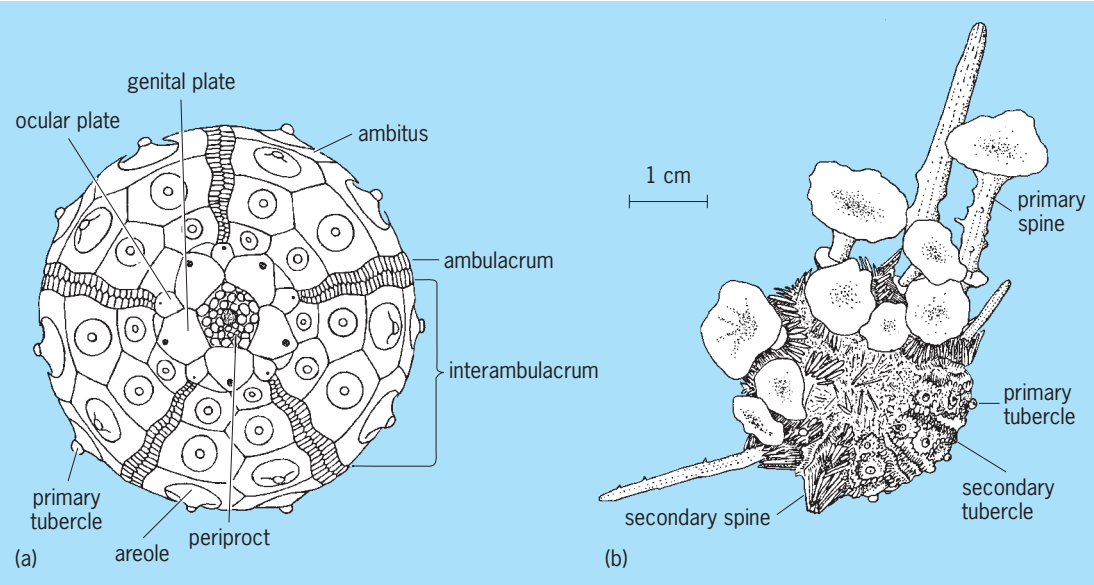

Fig. 1. Structure of the echinoid test of *Goniocidaris parasol*. (*a*) Naked test. (*b*) Test with spines.

Skeleton. The skeleton comprises the test and spines.

Test. The test is composed of regularly arranged plates that collectively form a rigid, or occasionally flexible, investing shell. In all echinoids since the Triassic, the test comprises 10 meridional areas, each composed of two vertical columns of plates. The meridians converge above and below at the upper and lower poles. The equatorial zone is termed the ambitus. Of the 10 meridional areas, five are ambulacra (or ambs); each amb consists of a double row of plates carrying tube feet. Alternating with the ambs are five meridional interambulacra (or interambs). The interambs are usually significantly wider than the ambs (**Fig. 2**).

Paleontology reveals that the test evolved only after much trial and error. The earlier Paleozoic forms showed extreme instability in the number of columns of plates. The cidaroids were the first group in which stability was achieved (in the Permian and Triassic), but a rigid, spherical shape was not adopted until the Jurassic. All of the bilaterally symmetrical irregular echinoids evolved from cidaroid ancestors. *See* PERISCHOECHINOIDEA.

Spines. Most echinoids carry spines that articulate with tubercles on the test plates and are moved by muscles at the spine bases. Large primary spines articulate with large primary tubercles, smaller secondary spines with secondary tubercles. If a ligament links the spine to the tubercle, the tubercle carries a small hole where the ligament is attached and is termed perforate. Tubercles without such a ligament are imperforate. The spine muscle may attach to an indented pattern on the edge of the boss (a protuberant part) on which the tubercle stands; if so, the tubercle is termed crenulate; if indentations are absent, the tubercle is noncrenulate. The muscle is attached to the test plate on a saucer-shaped depression around the tubercle, the areole. All these features are used in systematic diagnoses of the orders and families.

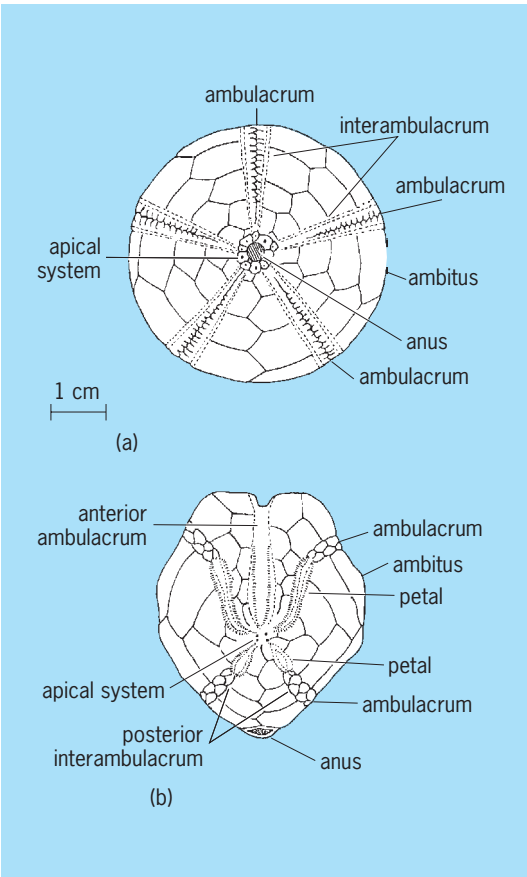

Fig. 2. Tests and related structures of (*a*) regular (endocyclic) echinoid and (*b*) irregular (exocyclic) echinoid.

At the upper pole lies a circlet of 10 plates, the apical system. Of these, five are at the top of the ambs and the five that alternate with them are at the top of the interambs. Each plate atop the interambs carries a gonopore, and is termed a genital plate. One of these also carries a madreporite (a delicately perforated sieve plate that takes in water to the vascular system). The other five are termed ocular plates, an old and misleading name for each ocular plate carries a tentacle, not an eye. The sea urchin adds new plates to its skeleton at the site of the ocular plates. Thus, the youngest plates in the test are nearer to the oculars, and plates more distant from the oculars, that is, toward the mouth, are the oldest. Individual plates grow by adding calcium carbonate to their margins. The resulting growth rings on the plates are seldom correlated with elapsed growth time.

In the spherical (regular) echinoids (Fig. 2*a*), the anus lies within the apical system on a membrane termed the periproct. These forms are sometimes termed endocyclic. In the sand dollars and heart urchins, the anus becomes displaced outside the apical system, it enters an interamb termed the posterior interamb, and the echinoid is said to have become irregular (exocyclic) [Fig. 2*b*]. Irregular echinoids tend to be secondarily bilaterally symmetrical, although a five-part radial symmetry is always evident. These features were formerly used in classification and are still valuable in determining trends of development in evolution. Irregular forms may exhibit modification of the apical system. The anus, for example, tends to obliterate the posterior genital plate (and its gonad) as it migrates backward, and the lost structures may not be replaced, or considerable distortion may ensue. *See* IRREGULARIA; REGULARIA.

Lantern. The mouth lies on the lower (oral) side of the test, surrounded by soft or plated skin, the peristome. In regular forms the mouth is central; in irregular forms it is central or may be displaced to a radius opposite the interamb that contains the anus. The radius containing the mouth is then termed anterior. In regular forms and in some irregular forms, the mouth is furnished with a ring of five powerful jaws, each with one large tooth. The jaws and teeth collectively make up the so-called lantern, first described by Aristotle. The lantern is moved by muscles that are attached to an inturned flange of the test, around the peristome, the perignathic girdle. Variations in the structure of the girdle, lantern, and teeth are used in classification. The teeth may be

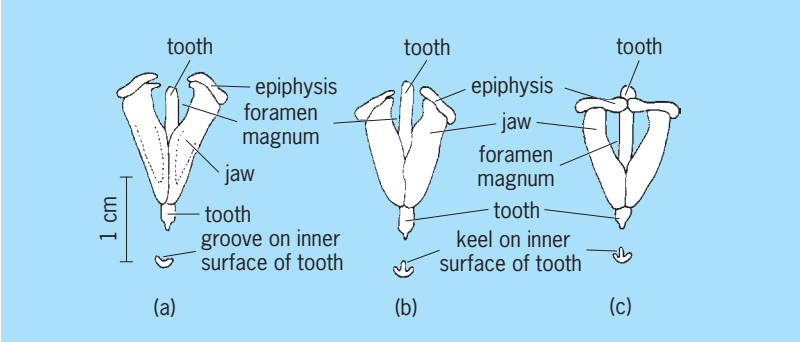

Fig. 3. Lantern structure and types of dentition for each main type. A single interradial jaw is illustrated, and below each is a cross section of the tooth. (*a*) Aulodont. (*b*) Stirodont. (*c*) Camarodont.

grooved or keeled, and the jaw may be partly or completely roofed over by epiphyses (**Fig. 3**). Three main types of dentition are distinguished: (1) aulodont, in which the teeth are grooved and epiphyses do not meet, so that there is an open foramen magnum in the jaw (Fig. 3*a*); (2) stirodont, in which the teeth are keeled within and the foramen magnum is open (Fig. 3*b*); and (3) camarodont, in which the teeth are keeled and the foramen magnum is closed by the epiphyses (Fig. 3*c*). These characters are useful in taxonomy. However, it was shown by T. Mortensen (during 1928–1951) that parallel dentitional evolution has occurred in various groups, and the three orders formerly based on dentition have been abandoned. As concise morphological descriptive terms, aulodont, stirodont, and camarodont remain valuable, and they are used in that sense in taxonomic diagnoses.

Water-vascular system. This system is highly developed and greatly influences the form of the ambulacral plates. The ring vessel rests upon the lantern and bears small reservoirs—polian vesicles. The stone canal passes upward through the coelom to the madreporite. The radial vessels and their ampullae lie on the inner surface of the ambs, within the test. The tube feet alone emerge to the exterior by way of paired pores.

Each tube foot traverses the test wall by means of two pores, termed a pore pair. One pore serves for the outward flow of hydrocoel fluid, the other for inward flow. The pore pairs lie on the amb plates (**Fig. 4**). In young stages and in the adults of Cidaroida and the irregular echinoids, each amb plate bears only one pore pair. Such amb plates are termed simple (Fig. 4*a*). In the older stages of noncidaroid regular echinoids, the amb plates tend to fuse into compound plates, which therefore carry more than one pore pair. The most common arrangement is that in which an arc of three pore pairs occurs on a plate. Such a plate is termed trigeminate or oligoporous (Fig. 4*b*, *d*, and *f*). Plates with four or more pore pairs are termed polyporous (Fig. 4*c*, *e*, and *g*). One component of a plate is usually larger than the others and is termed the primary. The position of the primary and the arrangement of the pore pairs provide characters used in taxonomy.

Diadematoid plates may be trigeminate or polyporous with the primary immediately over the lowest element (Fig. 4*b* and *c*). Arbacioid plates are similar, but the pore pairs lie in a vertical series, and the secondary elements (demiplates) are rectangular (Fig. 4*d* and *e*). Echinoid plates differ from diadematoid plates in having the primary as the lowest element (Fig. 4*f* and *g*). The ambulacra in some irregular echinoids tend to change from simple meridians into petal-shaped areas around the upper pole and around the mouth. Such ambs are described as petaloid, and the parts that surround the apical system are called the petals, whereas those around the mouth are termed phyllodes. A further

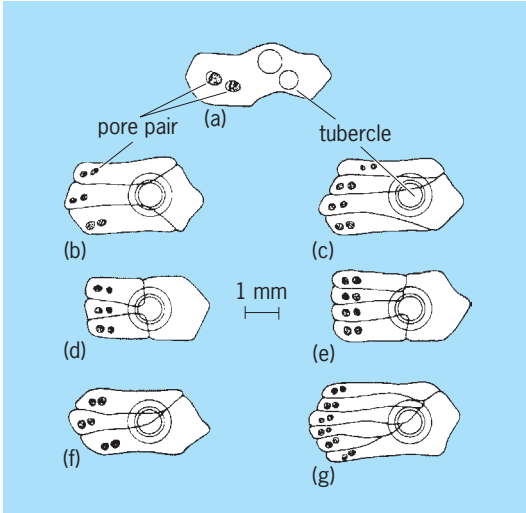

Fig. 4. Types of ambulacral plates as they are used in diagnoses of the orders. (*a*) Simple. (*b*) Trigeminate diadematoid. (*c*) Polyporous diadematoid. (*d*) Trigeminate arbacioid. (*e*) Polyporous arbacioid. (*f*) Trigeminate echinoid. (*g*) Polyporous echinoid.

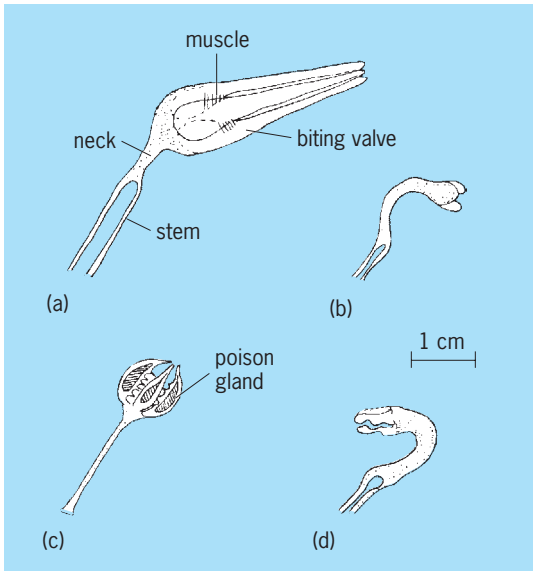

Fig. 5. Pedicellariae of echinoids. (*a*) Tridentate type. (*b*) Triphyllous or trifoliate. (*c*) Globiferous or gemmiform. (*d*) Ophicephalous.

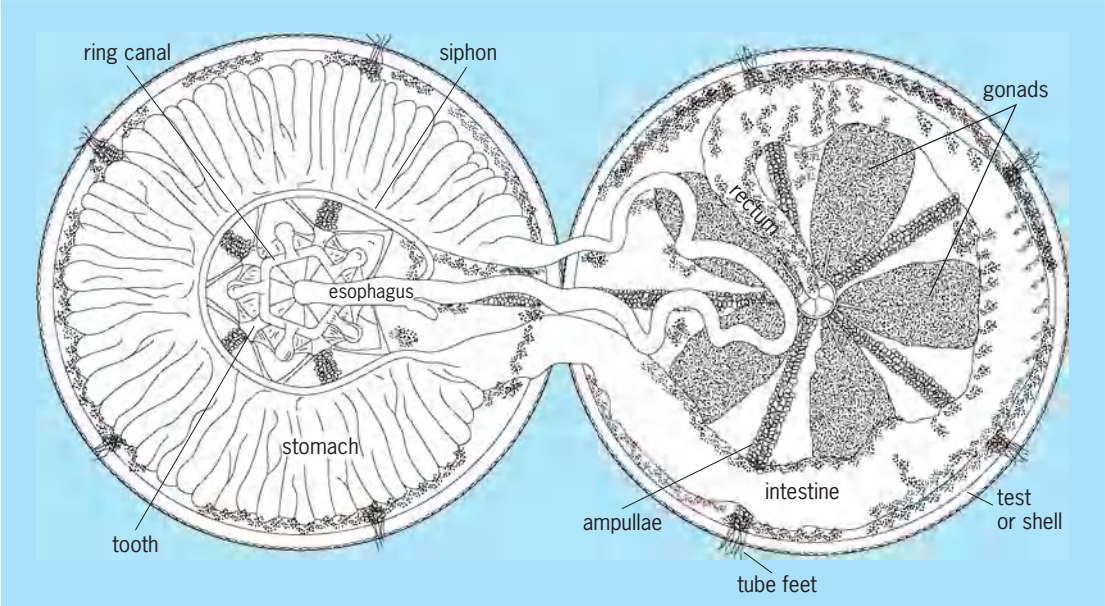

Fig. 6. Echinoid cut horizontally across the ambitus. Left is the adoral hemisphere, gonads removed. (*Adapted from W. R. Coe, Echinoderms of Connecticut, 1912; republished in T. I. Storer and R. L. Usinger, General Zoology, 3d ed., McGraw-Hill, 1957*)

development occurs in the Cassiduloida, where a flowerlike structure, the floscelle, is present.

The surface of the test plates is usually smooth but sometimes sculptured; that is, a raised pattern of ridges (epistoma) or of ridges and grooves ramifies among the tubercles. *See* TEMNOPLEUROIDA.

Some echinoids have special respiratory organs or gills attached to the peristome. The gills, if present, usually notch the margin of the peristome. The notches, termed gill slits, may vary in size and shape or may be absent. They provide characters used in taxonomy. Many irregular forms use the tube feet as respiratory organs.

Small grasping organs, pedicellariae, are well developed in echinoids, in which they take the form of a beak carried on a stalk. The beak is made up of three (rarely two or four) movable jaws, operated by muscles and sometimes provided with venom glands. They respond to tactile stimuli and seize any small organisms or particles that may touch the skin. The intrusive material is passed from one to the other until one ambital pedicellaria drops it off the urchin. The chief types are shown in **Fig. 5**.

Nervous system. The nervous system follows the same pattern as the water-vascular system. The tube feet, especially those near the mouth, may serve as tactile and taste organs. A few sea urchins have photosensitive eyespots on the upper surface of the test, scattered in the ectoderm. The so-called ocular pores are not sensitive to light. Minute spherical stalked bodies attached to the skin in some echinoids are believed to be organs of balance.

Alimentary system. The alimentary canal is tubular. It lies in the coelom, attached to the wall of the test by mesenteries. The stomach runs from the esophagus in a counterclockwise coil (as viewed from above), and the intestine retraces the route in reverse. The rectum passes upward to the anus (**Fig. 6**).

Reproduction. Like sea stars, sea urchins seem to become sexually mature after 1 year but continue to grow for several years. The life span is unknown but may average 5–6 years in medium-sized species, and some deep-sea and cold-water species may live much longer. The sexes are normally separate, although fertile hermaphrodites are known. The gonads may be reduced in number to only three or two. If a pelagic larva is present, it is an echinopluteus. In some species the young are brooded among the spines or in sunken petals, as in *Abatus*. Spines and some other organs are regenerated after injury. Autotomy (the process of self-amputation of appendages) and autoevisceration (expulsion of the digestive tract and associated organs through the anus) are unknown.

David L. Pawson; Howard B. Fell

Bibliography. J. W. Durham and R. V. Melville, A classification of echinoids, *J. Paleontol.*, 31(1):242–272, 1957; G. Hendler et al., *Sea Stars, Sea Urchins and Allies: Echinoderms of the Florida Keys and the Bahama Islands*, 1995; D. T. J. Littlewood, Echinoderm class relationships revisited, in *Echinoderm Research*, ed. by H. Roland et al., A. A. Balkema, Rotterdam, 1995; R. Mooi, Not all written in stone: Interdisciplinary syntheses in echinoderm paleontology, *Can. J. Zool.*, 79:1209–1231, 2001; T. Mortensen, *Monograph of the Echinoidea*, 5 vols., 1928–1951; D. Nichols, *Echinoderms*, 1962; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; J. S. Pearse and R. A. Cameron, Echinodermata: Echinoidea, in A. C. Giese, J. S. Pearse, and V. B. Pearse (eds.), *Reproduction of Marine Invertebrates*, vol. 6, pp. 514–662, Boxwood Press, Pacific Grove, 1991; A. B. Smith, Echinoderm larvae and phylogeny, *Annu. Rev. Ecol. Sys.*, 28:219–241, 1997; A. Smith, *Echinoid Palaeobiology*, 1984.

Echinorhiniformes

An order of sharks (bramble sharks) represented by one family, one genus, and two species. The family (Echinorhinidae) differs from all other sharks in possessing the following combination of characters: body and fins covered with thornlike denticles; two small spineless dorsal fins placed posteriorly to the pelvic fin origin; anal fin lacking; minute spiracles placed well behind the eyes; teeth similar in both jaws; and caudal fin without a subterminal notch. Maximum length is 2 m (6.6 ft).

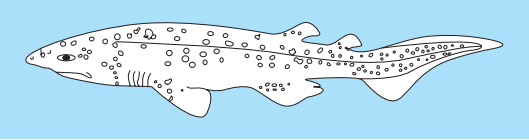

Bramble shark, *Echinorhinus brucus*. (*Courtesy of J. S. Nelson, 2006*)

Echinorhinus brucus (see **illustration**) occurs in the western North Atlantic, Indian, and western Pacific oceans, whereas *E. cookei* is limited to parts of the Pacific Ocean. *Echinorhinus brucus* has relatively few large denticles, differing from *E. cookei*, with relatively numerous small denticles. Both species inhabit temperate waters of continental and insular shelves and slopes, usually to depths greater than 200 m (660 ft) and less than 900 m (2950 ft). Their food consists of fishes, including small sharks, and crabs. Bramble sharks are ovoviviparous (producing eggs that develop internally and hatch before or soon after extrusion).

The family was previously placed in the Squaliformes. *See* CHONDRICHTHYES; ELASMOBRANCHII; SELACHII. Herbert Boschung

Bibliography. M. R. de Carvalho, Higher-level elasmobranch phylogeny, basal squaleans, and paraphyly, pp. 35–62, in M. L. J. Stiassny, L. R. Parenti, and G. D. Johnson (eds.), *Interrelationships of Fishes*, Academic Press, San Diego, 1996; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, 2006.

Echinothurioida

An order of Echinothuriacea with solid and hollow primary radioles, diademoid ambulacral plates, noncrenulate tubercles, and the anus within the apical system. The extant members of the family Echinothuriidae are all deep-water forms. The Echinothuriidae have a large flexible test which collapses into a disk at atmospheric pressure, and the middle element of the diademoid plates is much larger than the other two elements. Some species carry venomous spines. Echinothuriids range from the Late Jurassic to present day. *See* ECHINODERMATA; ECHINOIDEA; PYGASTEROIDA. Howard B. Fell

Echinozoa

A subphylum of free-living echinoderms in which the body is essentially globoid with meriodional symmetry. They lack arms, brachioles, or other appendages, and do not at any time exhibit pinnate structure. The Echinozoa range from the Early Cambrian to the present day. There are four classes that can definitely be placed here: (1) Edrioasteroidea, Lower Cambrian to lower Carboniferous echinozoans in which the mouth and anus were both directed upward and ambulacra (three to five in number) served as food-collecting areas; (2) Echinoidea, the existing and fossil sea urchins, originating in the Middle Ordovician; (3) Ophiocistioidea of the Lower Ordovician through Middle Devonian, with a domed aboral surface cover with large polygonal plates and a flat adoral side with a mouth and five radiating ambulacra; and (4) Holothuroidea, the existing and fossil sea cucumbers, which apparently first appeared in the Devonian. Two other extinct echinoderm classes can be placed here for convenience: (5) Lower Cambrian Camptostromatoidea, conical or domal animals with plates of varying size that overlapped on the lower theca; and (6) Lower Cambrian Helicoplacoidea, cylindrical animals with a spirally plated test and three ambulacra on the surface. The latter class may be the sister group of both echinozoans and crinozoans. *See* CAMPTOSTROMATOIDEA; ECHINOIDEA; HELICOPLACOIDEA; HOLOTHUROIDEA.

The oldest definite echinozoans are stromatocystitids of the Lower and Middle Cambrian. This group, which may have camptostromatids as its sister group, may have been ancestral to other edrioasteroids and, perhaps, other echinozoans. The asterozoans (asteroids and ophiuroids) may have been derived from the echinozoans, making the subphylum paraphyletic. *See* ECHINODERMATA.

Howard B. Fell; J. John Sepkoski, Jr.

Bibliography. R. C. Moore and C. Teichert (eds.), *Treatise on Invertebrate Paleontology*, pt. T, 1978; C. R. C. Paul and A. B. Smith, The early radiation and phylogeny of echinoderms, *Biol. Rev.*, 59:443–481, 1984; C. R. C. Paul and A. B. Smith (eds.), *Echinoderm Phylogeny and Evolutionary Biology*, 1988.

Echiura

A small group of unsegmented, wormlike, marine animals once linked with the Sipuncula and Priapulida under the taxon Gephyrea. Since 1940, echiurans have been regarded as a separate phylum of the animal kingdom with affinities to the segmented annelid worms. They range from the tropical to the polar seas, burrowing in sea-floor sediments from the intertidal area to depths of 30,000 ft (9140 m). *See* ANNELIDA; PRIAPULIDA; SIPUNCULA.

Morphology. The echiuran body consists of a round or elongated trunk with an anterior prostomium (see **illustration**). The mouth is located at the anterior end of the trunk where the tonguelike prostomium

is attached. The anus is at the posterior end. The body wall of the trunk may be thin enough to see red blood and internal organs within it, or it may be green and opaque as in the Bonelliidae. The muscles of the body wall may be a uniform sheet or visibly gathered into varying numbers of bundles that run from anterior to posterior (front to back). There is usually a pair of setae (hooked bristles) situated ventrally a short distance behind the mouth, and in the genera *Echiurus* and *Urechis* there are one or more rings of setae that surround the anus. Within the burrow the trunk undergoes rhythmic contractions that move water through the burrow.

The prostomium is muscular and capable of considerable extension. It is important in the identification of species but often becomes accidentally detached and lost during collection. This can make identification difficult. In most echiurans the prostomium is elongated with a rounded or blunt anterior end. In the bonellid echiurans the tip of the prostomium is forked. Living in burrows in marine sediments, echiurans extend the prostomium out onto the surface where it collects and transports nutrient-rich sediments to the mouth to be consumed. Species in the genus *Urechis* have a short stubby prostomium. These species secrete a mucus net within their burrow and then pump water through it to filter out food.

The mouth, located at the base of the prostomium, leads into the gut, which is divided into distinct regions designed to systematically process the food. Two characteristic organs, the anal vesicles, are present at the posterior end of the gut near the anus. These organs may be tubular or branched and can extend for varying distances toward the anterior end of the body cavity. Water is intermittently drawn in through the anus to fill these sacs and then is expelled. Both respiratory and excretory functions have been proposed for these organs. Nephridia (organs typically known as having an excretory function in most invertebrates) are present on the ventral body wall behind the anterior setae. In echiurans the nephridia have been modified to gather ripe gametes (eggs or sperm) from the body cavity and store them prior to spawning. A ventral nerve cord runs from the anterior to the posterior end of the trunk. *See* DIGESTION (INVERTEBRATE); EXCRETION.

The sexes are separate and indistinguishable in most echiurans. In the Bonelliidae, however, the sexes are separate and very dissimilar, the male being minute and living within the female's egg storage organ. There they fertilize the eggs as they are spawned. The sex of bonellid offspring is determined largely by environmental factors.

Classification. Traditional morphological and embryological evidence has suggested a phylogenetic (evolutionary) relationship between these animals and the annelids, but the echiurans have remained a separate phylum. Embracing this taxonomic status, A. C. Stephen and S. J. Edmonds recognized three orders—Echiuroinea, Xenopneusta, and Heteromyota—the last order with one problematic species attached. Most species are included in the Echiuroinea, which is divided into families Bonelliidae and Echiuridae.

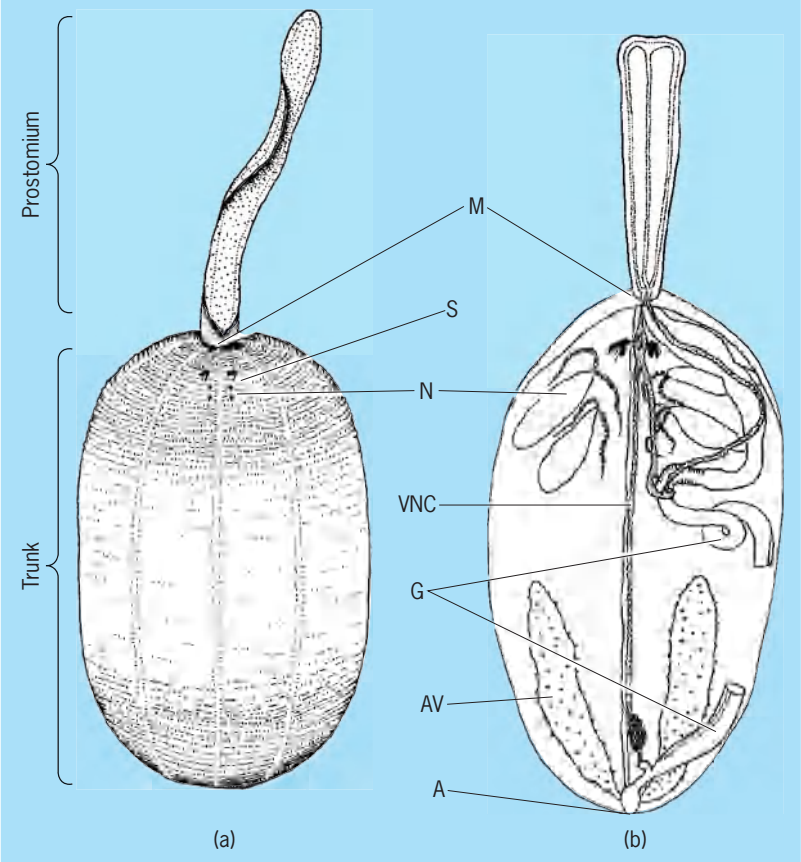

Echiuran anatomy. (*a*) External anatomy, ventral view. (*b*) Internal anatomy, dorsal view. A, anus; AV, anal vesicles; G, gut (middle section removed for clarity); M, mouth; N, nephridia (egg and sperm storage organs); S, setae; VNC, ventral nerve cord (*modified from A. C. Stephen and S. J. Edmonds, 1972*).

Contemporary evidence used to classify animals and help determine their relationships (systematics and phylogeny) now comes from molecular data (for example, DNA and RNA nucleotide sequences) and sophisticated microscopic techniques. The DNA sequence of one important gene in an echiuran has been found to be more similar to the same gene in annelid worms than it is to other closely related but nonannelid animals. This finding suggests that the echiurans may indeed be very closely related to the segmented annelid worms; however, their unsegmented body remained problematic. Recent fluorescent microscopic techniques focused on the structure of the ventral nerve cord have revealed structural repetition in the anterior to posterior axis of the organ that suggests the presence of a more subtle segmental body plan in echiurans. Taken together, these two findings have revived a hypothesis that considers echiurans to be members of a distinct class within the phylum Annelida. Classification below this taxonomic level has not been formally proposed. However, if one were, it would likely retain the orders currently recognized. John F. Pilger; Mary E. Rice

Bibliography. R. Hessling, The use of confocal laser-scanning microscopy (cLSM) and 3d reconstruction

for systematic zoology—on the phylogeny of the Echiura, *Microsc. Today*, 1(5):8–11, 2001; R. Hessling and W. Westheide, Are Echiura derived from a segmented ancestor?—Immunohistochemical analysis of the nervous system in developmental stages of *Bonellia viridis, J. Morphol.*, 252:100–113, 2002; D. McHugh, Molecular evidence that echiurans and pogonophorans are derived annelids, *Proc. Nat. Acad. Sci.*, 94:8006–8009, 1997; J. F. Pilger, Echiura, in F. W. Harrison and M. E. Rice (eds.), *Microscopic Anatomy of Invertebrates*, vol. 12, Wiley-Liss, New York, 1993; J. F. Pilger, Sipunculans and echiurans, in S. F. Gilbert and A. M. Raunio (eds.), *Embryology: Constructing the Organism*, Sinauer Associates, Sunderland, MA, 1997; A. C. Stephen and S. J. Edmonds, *The Phyla Sipuncula and Echiura*, British Museum (Natural History), 1972.

Echo

A sound wave which has been reflected or otherwise returned with sufficient magnitude and time delay to be perceived in some manner as a sound wave distinct from that directly transmitted. Multiple echo describes a succession of separately indistinguishable echoes arising from a single source. When the reflected waves occur in rapid succession, the phenomenon is often termed a flutter echo.

Echoes and flutter echoes are generally detrimental to the quality of the acoustics of rooms. They may be minimized through the proper selection of room dimensions, room shape, and distribution of sound-absorbing materials. Flutter echoes, for example, may be minimized by making the opposite walls of a room nonparallel or by making one of the walls highly sound absorptive. For a more complete discussion of the effect of room shape on echoes *see* ARCHITECTURAL ACOUSTICS

Individual echoes will generally be heard only if the time delay between the direct sound wave and a reflected sound wave is in excess of approximately 50 milliseconds. In a large auditorium where the reverberation time is of the order of seconds, many reflected waves will be present. These reflections will not be troublesome if their intensity is sufficiently below that of the initial sound. A relation between the approximate percent of listeners detecting some alteration in the acoustics of a room and the sound pressure level of the echo as a function of the delay time following the arrival of the direct sound is shown in the **illustration**. *See* REVERBERATION.

Echoes have been put to a variety of uses in measurement problems. For example, the distance between two points can be measured by timing the duration required for a direct sound originating at one location to strike an object at the other point and to return an echo to the location of the initial source. For the application of this principle to the detection of submerged objects *see* SONAR

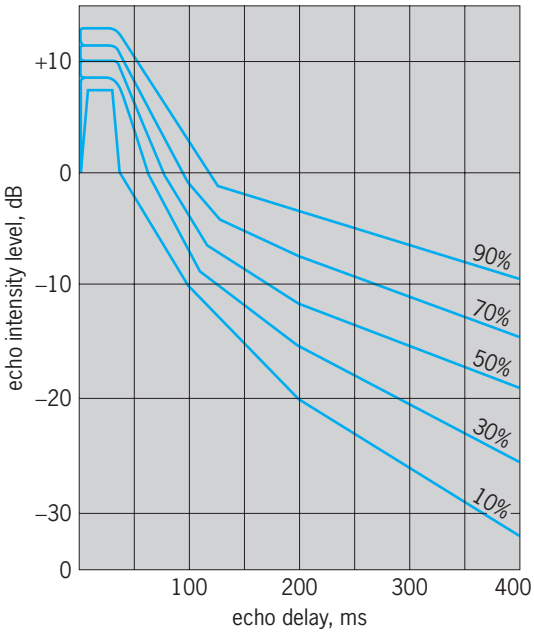

Estimated percent of listeners disturbed by the echo related to the intensity of the echo (relative to initial sound intensity) and the time delay between the inital sound and perception of the echo. (*After R. H. Bolt and P. E. Doak, J. Acousi. Soc. Amer., 22:507–509, 1950*)

Ultrasonic echo techniques have achieved considerable success in nondestructive testing of materials. When an ultrasonic wave is propagated through a metal, the presence of a crack or other flaw will cause a sound wave, or echo, to be reflected back to the initial source location. Observing the time delay between the original sound and the perception of the echo permits the location of the flaw to be determined. This technique has been found particularly useful in such problems as examining metal casting for internal defects and determining the location of cracks in pipes or welded structures. It has also been employed to locate brain tumors in humans. For further discussion of the application of echo reflection techniques *see* REFLECTION OF SOUND; SOUND; ULTRASONICS.

William J. Galloway

Echo sounder

A marine instrument used primarily for determining the depth of water by means of an acoustic echo. A pulse of sound sent from the ship is reflected from the sea bottom back to the ship, the interval of time between transmission and reception being proportional to the depth of the water.

Echo sounders, sometimes called fathometers, are used by vessels for navigational purposes, not only to avoid shoal water, but as an aid in fixing position when a good bathymetric chart of the area is available. Some sensitive instruments are used by commercial fishers or marine biologists to detect schools of fish or scattering layers of minute marine life. Oceanographic survey ships use echo sounders for charting the ocean bottom. **Figure 1** shows an

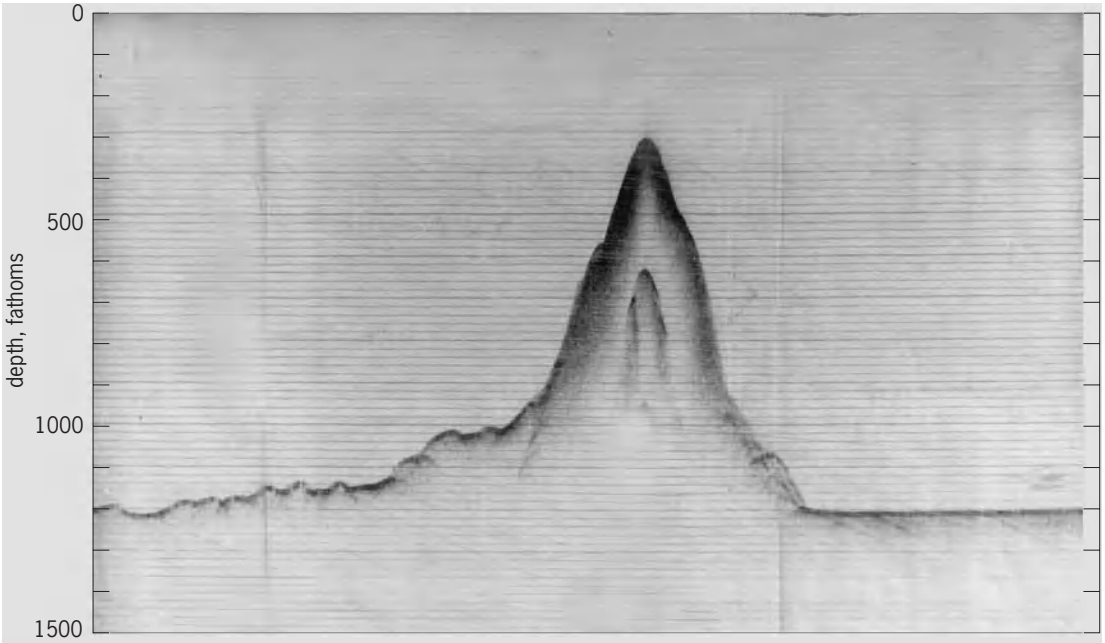

Fig. 1. Echo-sounder record of a seamount in the deep ocean. 1 fathom = 1.8 m. (*Woods Hole Oceanographic Institute*)

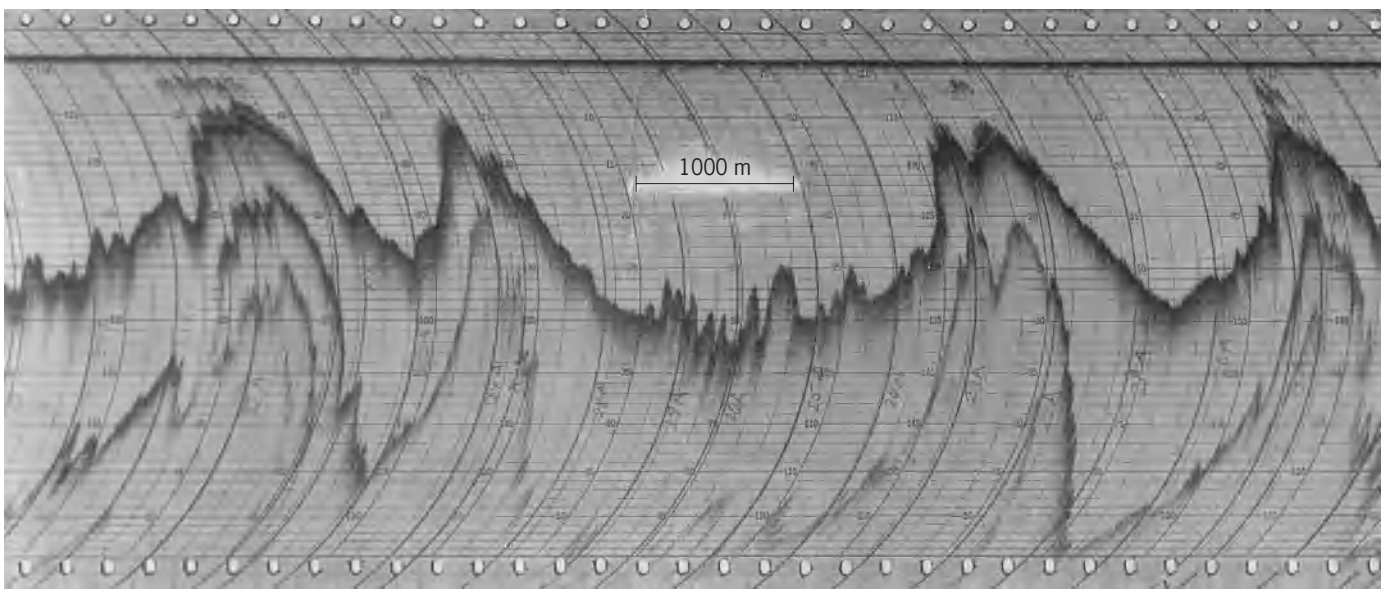

Fig. 2. Typical record of bottom profile obtained by an echo sounder. 1 m = 3.3 ft. (*USCGS*)

echo-sounder record obtained by oceanographers of a seamount (undersea mountain). *See* SCATTERING LAYER.

An echo sounder is really a type of active sonar. It consists of a transducer located near the keel of the ship which serves (in most models) as both the transmitter and receiver of the acoustic signal; the necessary oscillator, receiver, and amplifier which generate and receive the electrical impulses to and from the transducer; and a recorder or other indicator which is calibrated in terms of the depth of water. An echo sounder actually measures time differences, so some average velocity of sound must be assumed in order to determine the depth. The frequency generally employed is in the low ultrasonic range (20,000–30,000 Hz). The depth display may be given by a trace-type recorder which supplies a continuous permanent record (**Fig. 2**) or, as is the case with less expensive commercial instruments, it may be a dial-type or digital display, giving instantaneous depth of water. *See* MARINE GEOLOGY; SONAR; UNDERWATER SOUND.

Robert W. Morse

Echocardiography

A diagnostic procedure that uses ultrasound at a frequency of 2.5–10 MHz to provide an image of the heart. It is based on the principle that the interface between tissues of different acoustical impedance

causes the ultrasound to be reflected to the transducer, which spends a fraction of each second receiving these echoes. There are many interfaces between blood and the various structures in the heart that contact blood, such as the heart walls, valves, and great vessels. Also, the surface of the heart reflects ultrasound because it is surrounded by the lungs, which are filled with air.

The ultrasound apparatus can estimate the distance from the transducer to the tissue interface that is reflecting the sound, because the velocity of sound in tissue is relatively constant and the time required to send and receive sound can be measured (distance = velocity × time). A graphic representation of the locale of each interface can then be displayed on a video monitor or digital screen, and as the beam moves through the heart a composite two-dimensional image of cardiac structures can be constructed. The sampling rate of the ultrasound device is adequate to display heart motion in real time. *See* BIOMEDICAL ULTRASONICS; ULTRASONICS.

Procedure. An ultrasound image of the heart is generated on a monitor (see **illustration**). The image is generated by electronically moving the ultrasound beam repeatedly through an arc. The transducer is usually applied to the anterior chest by using a coupling gel devoid of air. This procedure is referred to as transthoracic echocardiography. Small transducers can also be attached to probes placed in the esophagus behind the heart—a procedure known as transesophageal echocardiography; during heart surgery the transducer can be placed directly on the heart—a procedure known as epicardial echocardiography; and a small transducer can be placed on a catheter to visualize the inside of vessels—a procedure known as intravascular ultrasound. The closer the transducer is to the heart, the higher the resolution of the resulting images. Axial resolution (parallel to the path of the sound beam) for all echocardiographic systems is quite high [approximately 1 mm (0.04 in.)], but lateral resolution (perpendicular to the path of the sound beam) is reduced and deteriorates the farther away the structure is from the transducer. The frame rate of the image is about 30 per second.

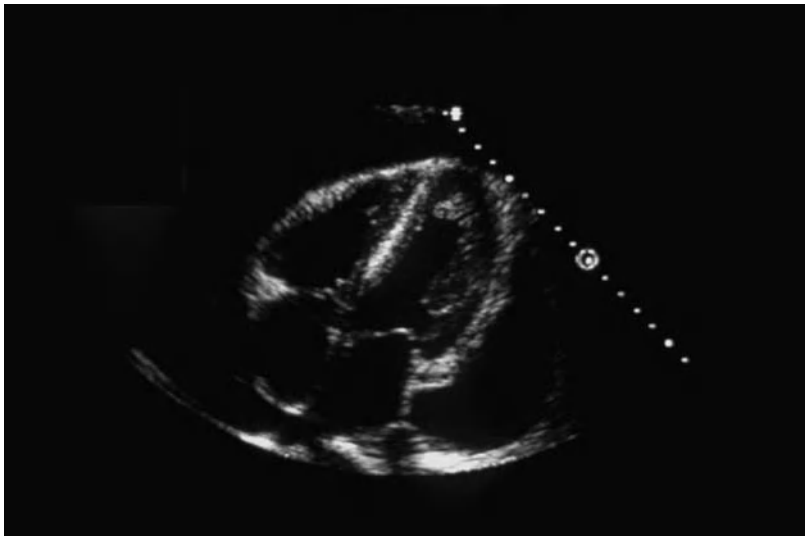

Two-dimensional echocardiogram from a position near the apex of the ventricles. The left ventricle is on the right, separated from the left atrium below by the mitral valve. The right ventricle is on the left, and the right atrium is below it with the tricuspid valve in between.

Two-dimensional echocardiography. Two-dimensional echocardiography is completely harmless and provides excellent real-time images of the heart. It is very useful for determining the anatomy and function of heart valves, detecting abnormal amounts of pericardial fluid, and defining the complex anatomy of congenital heart defects. It is effective for estimating size and function of heart chambers, thickness and mass of the heart wall, and size of the great vessels. The equipment is portable, and imaging can be accomplished during stresses, such as exercise. Thus, two-dimensional echocardiography is frequently employed to evaluate suspected or overt heart disease.

Blood flow. Echocardiography can also be used to assess blood flow in the heart by employing the Doppler principle. The movement of a structure during an examination performed by ultrasound results in a shift in the frequency of the returning sound wave (echo). The shift in frequency is proportional to the velocity of the moving object. Since the ultrasound frequency transmitted (fo), the speed of sound in tissue (C), and the angle of the sound beam to the direction of movement (θ) are known, velocity (V) can be calculated by measuring the frequency shift (Fd), using Eq. (1).

$$V = \frac{Fd \times C}{2fo \times \cos\theta} \tag{1}$$

Blood, the object usually interrogated by Doppler ultrasound of the heart, moves because of a pressure gradient (ΔP) in the heart. The pressure gradient can be calculated from the simplified Bernoulli equation (2) if velocity (V) is known.

$$\Delta P = 4V^2 \tag{2}$$

Doppler systems. Three basic types of Doppler echocardiographic systems are available on most ultrasound machines: continuous-wave, pulsed, and color-flow.

In a continuous-wave (CW) Doppler system, sound is continuously transmitted and received. The advantage of CW Doppler is that the very high velocities of severe structural heart defects are accurately quantitated. The disadvantage is range ambiguity, that is, the inability to locate the origin of the high-velocity blood flow along the sound beam.

Pulsed Doppler avoids range ambiguity by being able to range-gate the returning echoes so that only certain areas in the sound beam are sampled. However, pulsed Doppler cannot resolve very high velocities because the pulse repetition frequency limits Doppler frequency-shift sampling. Pulsed and CW Doppler are useful for detecting and localizing increased blood velocities associated with obstructions to flow. *See* DOPPLER EFFECT.

Pulsed Doppler can also be directed at the myocardium to measure the velocity of muscle contraction or relaxation. Currently, these tissue Doppler velocities are used to assess relaxation properties of the left ventricle.

The third system is color-flow Doppler. Small sample volumes [1–2 mm^3 (0.04–0.08 $in.^3$)] throughout the two-dimensional echocardiographic image are sampled for velocity. Only velocities fast enough to result from blood movement are displayed, superimposed on the two-dimensional echocardiographic image and color-encoded. Movement (velocity) toward the transducer is usually red, and movement away is usually blue. The relative velocity is indicated by the intensity of the color. Turbulent flow is a mosaic color (for example, green). The resultant images display blood moving through the heart in real time as a color on the black-and-white anatomy of the heart and its internal structures. Color-flow Doppler is excellent for displaying abnormal flow due to valve leakage or other structural defects.

M mode. The complete echocardiographic examination also includes selective sampling along certain radians in the arced two-dimensional image, and displaying of the movement of the structures crossed by the radian over time in an analog fashion. In this format, sampling rates can be increased to 1000 per second so that complex motion patterns of a structure can be elucidated. This format is called time-motion echocardiography or M-mode echocardiography. It was actually the first (1950s) commercially available form of ultrasound for examining the heart. It is adjunctive to two-dimensional and color-flow Doppler echocardiography for answering specific questions about heart structural dimensions and function. *See* HEART DISORDERS.

Michael H. Crawford

Bibliography. M. D. Cheitlin et al., ACC/AHA/ASE 2003 guideline update for the clinical application of echocardiography, *J. Amer. Coll. Cardiol.*, 42:954–970, 2003.

Echolocation

The biological sonar that bats, porpoises, and certain other animals use to navigate without the visual system. Several different groups of animals have evolved the ability to perceive objects by emitting sounds and hearing the echoes that the objects reflect to their ears. The locations and characteristics of the objects are represented by acoustic properties of the echoes, and the ears and auditory systems of these animals act as the sonar receiver. The sense of hearing is specialized for converting echo information into displays of objects, which are perceived as acoustic images that guide the animal's behavior. The best-known examples of echolocating animals are bats (Microchiroptera) and porpoises and toothed whales (Cetacea). However, several other kinds of mammals (some flying foxes, shrews, and rats) and birds (oilbirds and cave swiftlets) also can echolocate. *See* SONAR.

Echolocation sounds. Bats produce their ultrasonic sonar sounds from the larynx and broadcast them through the open mouth or through a specialized transmitting antenna formed around the nostrils. Many kinds of bats (leaf-nosed bats, horseshoe bats) have oddly shaped nasal structures for sound emission that appear bizarre unless it is realized that they guide the beam of the sonar transmissions. Porpoises produce their sonar sounds from structures located beneath the blowhole on the top of the head (through which they breathe) and project them into the water through the rounded, protuberant forehead, which contains acoustically specialized tissue. The sonar sounds of porpoises and whales are very brief impulses, or clicks, which contain a wide range of ultrasonic frequencies, all occurring at the same instant. The rate of emission of these sonar clicks depends in part upon the distance to objects that interest the animal, and can vary from several sounds per second to several hundred. Porpoises use their sonar to find fish and presumably to perceive objects beyond the relatively restricted range of vision under water.

The sonar sounds of bats vary from one family and species to another and consist of short (0.5 to about 10 ms) frequency-modulated (FM) signals that are often coupled with short or long (0.5 to over 100 ms) constant-frequency signals (**Fig. 1**). Their frequency span, or bandwidth, is important because the quality or sharpness of the acoustic images produced from FM echoes improves as the bandwidth increases. Under favorable conditions, bats that use wide-band FM signals can perceive the distance to a target with an acuity of a fraction of a millimeter, for example. Constant-frequency sonar signals are well suited for initial detection of targets because they concentrate their energy near a known frequency to which the sonar receiver can be turned, and long-duration constant-frequency signals are useful for determining the velocity of a target from the Doppler shift of echoes. Bats that emit constant-frequency signals perceive small changes in the frequency of echoes and use this ability to determine how fast they are moving relative to objects. Bats that emit only FM signals observe changes in a target's position from one echo to the next to determine relative velocity. *See* DOPPLER EFFECT; FREQUENCY MODULATION.

Most bats are insectivorous and hunt at night. They emit sonar sounds in a stereotyped pattern correlated with stages in the interception of a flying insect (Fig. 1). The entire interception, from initial detection to capture, lasts less than a second. Bats produce sonar sounds at a low emission rate while searching for prey; the rate increases progressively while approaching the target after it has been detected. During the approach, the bat tracks the target and decides whether to attempt interception. During the final stage of interception, which lasts only about a tenth of a second, the sonar sounds appear at a rate as high as 100 sounds per second. The constant-frequency or FM structure of the sounds (Fig. 1) represents the acoustic strategy that different species of

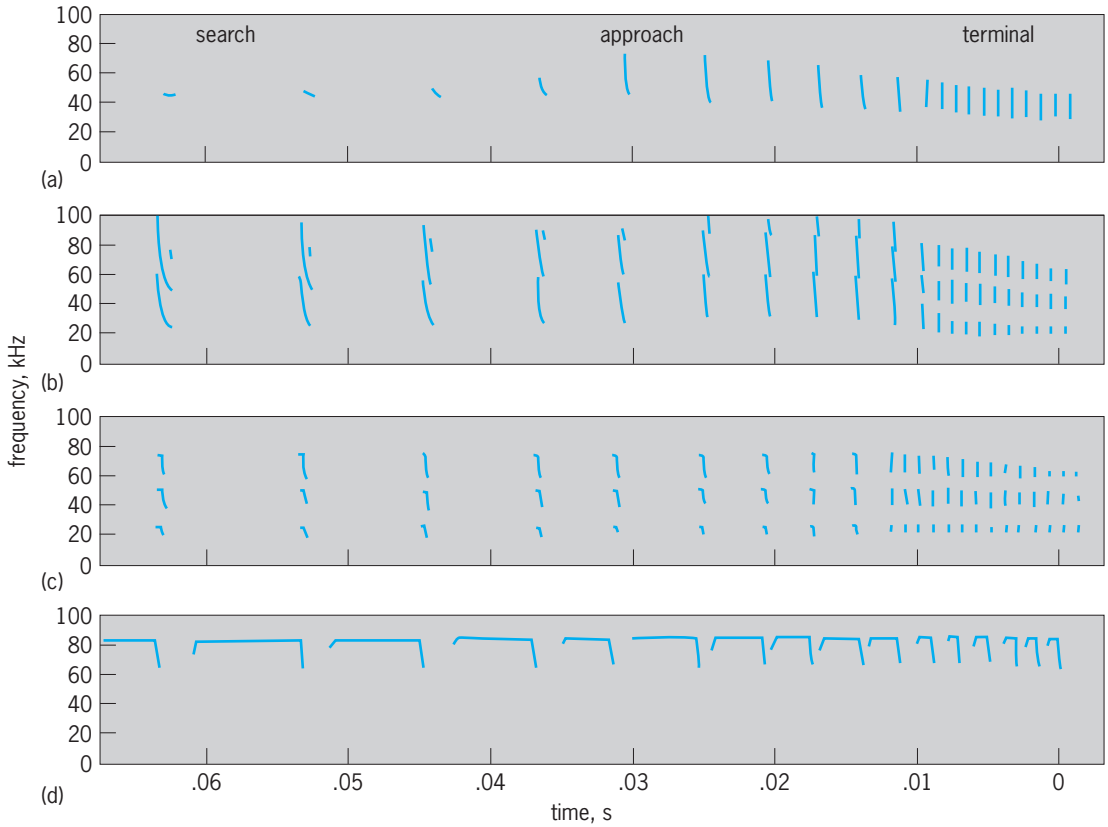

Fig. 1. Sound spectrograms representing sonar signals of (*a*) *Tadarid brasillensis*, (*b*) *Eptesicus fuscus*, (*c*) *Pteronotus personatus*, and (*d*) *Rhinolophus ferrumequinum* during search, approach, and terminal stages of pursuit of prey. Detection, identification, and tracking usually take place in less than 1 s. Spectrograms *b* and *c* use only FM signals, whereas *a* and *d* use constant-frequency signals during the search and change to FM signals during the approach and terminal buzz. (*After J. A. Simmons, F. B. Fenton, and M. J. O'Farrell, Echolocation and pursuit of prey of bats, Science, 203:16–21, 1979*)

bats have evolved to solve the problems of finding prey.

Auditory images. The images in echolocation are derived from either frequency-modulated or constant-frequency echoes.

Frequency-modulated images. The distance to a target, or target range, is represented acoustically by the delay of FM echoes. Bats first encode echo delay as the timing of nerve discharges but eventually represent delay by the location of neural responses within a discrete patch of the auditory cortex that functions as a biological map of target range. Some FM echo-locating bats combine their perception of target range with their perception of the shape of targets by perceiving the distance to different parts of the target along the same scale as the distance to the target as a whole.

The bat in **Fig. 2** emits an FM sonar sound and receives an echo from the target, here shown to be a flying moth. The target reflects echo components from parts located at different ranges (for simplicity, one echo component from its head and one from its wings). The target as a whole has a range, r, and the target's structure consists of the range separation, Δr. These target features are represented in the echo by the echo's delay, t, and the time separation of the two components within the echo, Δt, because the echo contains a separate component from each part of the target, at delay t and delay $t + \Delta t$. The bat's inner ear represents the FM emission and the echo as spectrograms analogous to those shown in Fig. 1. The spectrograms in Fig. 2 are patterns of mechanical excitation produced by the inner ear itself. Mechanical excitation then evokes neural discharges that convey the spectrograms along parallel neural channels into the auditory nervous system. The time separation of the echo components, Δt, creates notches in the spectrum of the echo as a whole separated by the frequency interval, Δf, which equals $1/\Delta t$. These spectral notches appear in the neural spectrogram of the echo as a "scalloping" pattern of the array of discharges (solid circles) caused by the lengthening of discharge latency at the lower amplitudes of the echo within the notch as compared to outside the notch. Some neural discharges near the center of the notch may even fail to occur at all (open circles) if the echo is very weak in the notch.

The delay of echoes is presented in the bat's brain by spectrogram delays, t_f, at each frequency in the emission and the echo (Fig. 2). Spectrogram delays are further processed to be encoded, frequency by frequency, in terms of the location of neural activity on the target-range map in the auditory cortex, which is the bat's auditory display of targets. Spectrogram delays are somehow integrated across frequency to determine the distance to the

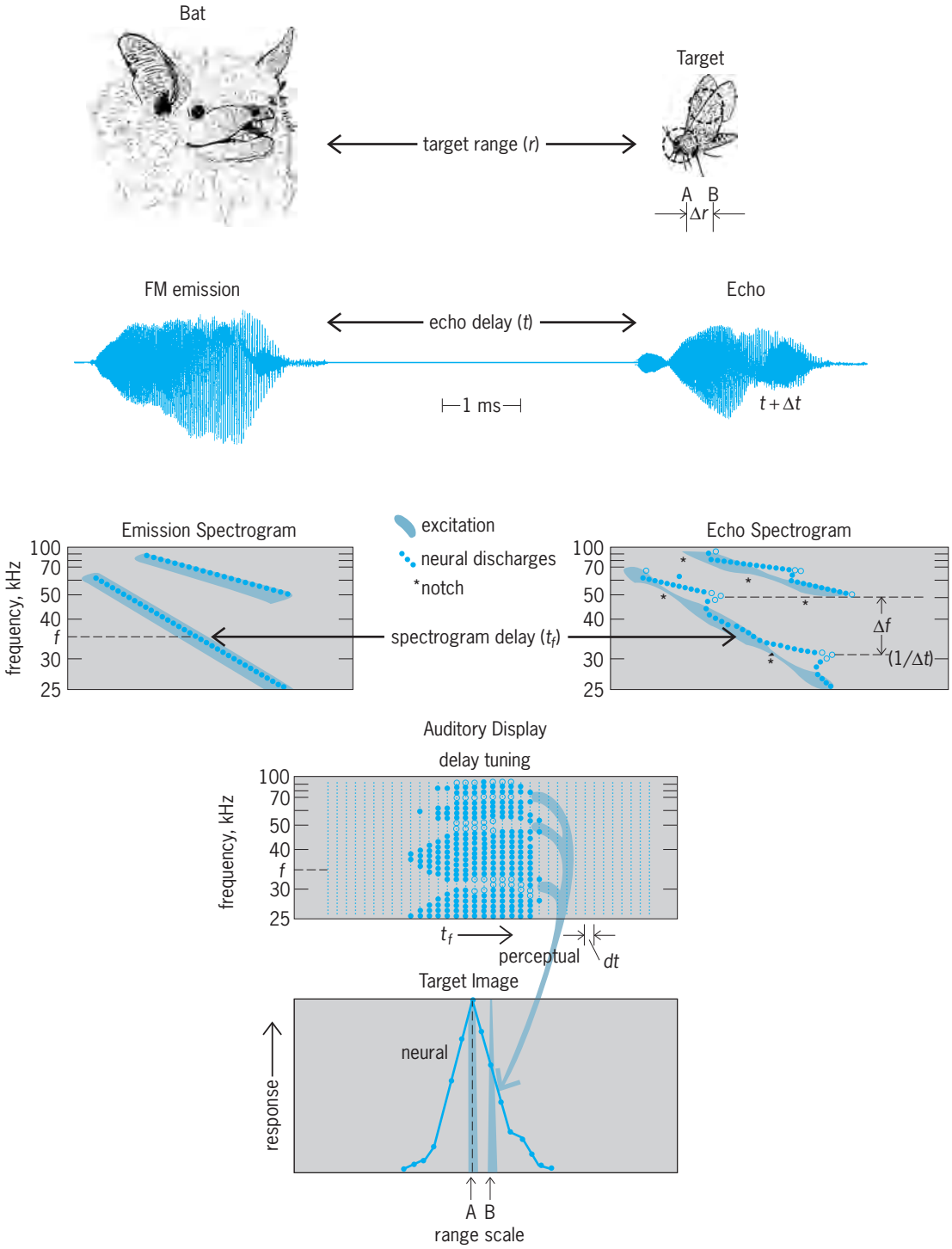

Fig. 2. Stages by which bats convert echoes of FM sonar sounds into images that incorporate the distance and the shape of the target. Details are discussed in the text. (*After J. A. Simons, C. F. Moss, and M. Ferragamo, Convergence of temporal and spectral information into acoustic images of complex sonar targets perceived by the echolocating bat, Eptesicus fuscus, J. Compar. Physiol. A, 1989*)

leading edge of the target from the earliest component of the echo, at delay t. The auditory display in Fig. 2 shows the distribution of neural activity along the range axis of the auditory cortex. The arrow leading from the auditory display to the target image shows the effect of the locations of the notches in the spectrum that produce a wavy pattern of activity in the cortex across frequencies. This wave pattern is converted into an estimate of the location of the echo's second component, at delay $t + \Delta t$. Somehow the single-peaked distribution of neural activity across the range map is converted into a two-peaked pattern to represent the target's image as two discrete components (shown as peaks labeled perceptual, A and B) located at ranges r and $r + \Delta r$. This conversion is an example of the brain creating a single, unitary psychological dimension out of neural representations that treat different features of stimuli

in different ways. It illustrates an important function of the brain that is not understood.

Constant-frequency images. The use by bats of constant-frequency echoes to determine target velocity from echo Doppler shifts is another example of neural information processing leading to perception. The bat's inner ear represents different frequencies of sound by the location of neural discharges within the population of neurons in the auditory nerve. The frequency of a constant-frequency echo activates only those neurons tuned to that particular frequency. This activity is passed along from the auditory nerve to higher centers of hearing in the brain. When the bat receives a particular constant-frequency echo frequency, neurons in a corresponding location on frequency maps in the auditory cortex respond, whereas other neurons do not. One of these cortical frequency maps becomes a map of target velocity for constant-frequency echoes. In this case, no neural computations are needed to create the frequency map because it already exists at the inner ear. However, neural computations are needed to adjust the relationship between each frequency and the size of the frequency change corresponding to each velocity, because the Doppler shifts are proportional to the frequency being shifted. Other frequency maps segregate neural responses to constant-frequency echoes to facilitate detection of the target's fluttering motions, such as the wing beats of insects, for identification of prey. *See* CHIROPTERA; PHONORECEPTION; ULTRASONICS.

James A. Simmons

Bibliography. D. R. Griffin, *Listening in the Dark*, 1986; P. Nachtigall and P. W. B. Moore, *Animal Sonar Systems: Processes and Performance*, 1988; J. A. Simmons, M. B. Fenton, and M. J. O'Farrell, Echolocation and pursuit of prey by bats, *Science*, 203:16–21, 1979; J. A. Simmons, C. F. Moss, and M. Ferragamo, Convergence of temporal and spectral information into acoustic images of complex sonar targets perceived by the echolocating bat, *Eptesicus fuscus, J. Compar. Physiol. A*, 1989.

Echovirus

Members of the Picornaviridae family, genus *Enterovirus*. The name is derived from the term enteric cytopathogenic human orphan virus. More than 34 antigenic types exist. Only certain types have been associated with human illnesses, particularly with aseptic meningitis and febrile disease, with or without rash. Their epidemiology is similar to that of other enteroviruses. *See* ENTEROVIRUS; PICORNAVIRIDAE.

Echoviruses resemble polioviruses and coxsackieviruses in size (about 28 nanometers) and in many other properties. They are nonpathogenic for newborn mice, rabbits, or monkeys, but cytopathogenic for monkey kidney and other tissue cultures (see **illus.**). *See* COXSACKIEVIRUS; TISSUE CULTURE.

Diagnosis is made by isolation and typing of the viruses in tissue culture. Antibodies form during convalescence. *See* ANIMAL VIRUS.

Joseph L. Melnick

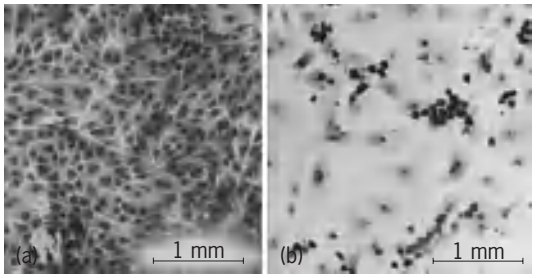

Monkey kidney tissue culture. (*a*) Normal culture. (*b*) Three days after infection with echovirus type 1. (*Courtesy of J. L. Melnick, Baylor College of Medicine*)

Bibliography. N. J. Schmidt and W. Emmons (eds.), *Diagnostic Procedures for Viral, Rickettsial, and Chlamydial Infections*, 7th ed., 1995; Z. Woldehiwet and R. Miodrag (eds.), *Rickettsial and Chlamydial Diseases of Domestic Animals*, 1993.

Eclipse

The occultation (obscuring) of one celestial body by another. Solar and lunar eclipses take place at syzygies of the Sun, Earth, and Moon, when the three bodies are in a line. At a solar eclipse, the Moon blocks the view of the Sun as seen from the Earth. At a lunar eclipse, the Earth's shadow falls on the Moon, darkening it, and can be seen from wherever on Earth the Moon is above the horizon. *See* SYZYGY.

Eclipses of the Sun could be seen from other planets as their moons are interposed between the planets and the Sun, though their superposition is not as coincidental in angular size as it is for the Earth-Moon system. Eclipses of the moons of Jupiter are well known, occurring whenever the moons pass into Jupiter's shadow. An eclipse of the Sun visible from Uranus that occurred in 2006 was imaged by the *Hubble Space Telescope*. Certain binary stars are known to eclipse each other, and the eclipses can be followed by measuring the total light from the system. *See* BINARY STAR; ECLIPSING VARIABLE STARS; JUPITER.

Related phenomena are transits, such as those of Mercury and Venus, which occur when these planets cross the face of the Sun as seen from Earth. They are much too small to hide the solar surface. Transits of the Earth will be seen from spacecraft. Occultations of stars by the Moon are commonly seen from Earth, and are studied to monitor the shape and path of the Moon; a solar eclipse is a special case of such an occultation. Occultations of stars by planets and by asteroids are now increasingly studied; the rings of Uranus were discovered from observations of such an occultation, and the atmospheres of Titan and Pluto are best known from occultation studies. *See* OCCULTATION; TRANSIT (ASTRONOMY); URANUS.

Solar Eclipses

A solar eclipse can be understood as an occultation of the Sun by the Moon or, equivalently, the Moon's shadow crossing the Earth's surface. The darkest

part of the shadow, from which the Sun is entirely hidden, is the umbra (**Fig. 1**). The outer part of the shadow, from which part of the Sun can be seen, is the penumbra.

Solar eclipses can be central, in which the Moon passes entirely onto the solar disk as seen from Earth, or partial, in which one part of the Sun always remains visible. Central eclipses can be total, in which case the Moon entirely covers the solar photosphere, making the corona visible for the period of totality, or annular, in which case the Moon's angular diameter is smaller than that of the Sun because of the positions of the Earth and Moon in their elliptical orbits. At an annular eclipse, a bright annulus of photospheric sunlight remains visible; it is normally thousands of times brighter than the corona, leaving the sky too blue for the corona to be seen.

The plane of the Moon's orbit is inclined by 5° to the plane of the Earth's orbit (the ecliptic), so the Moon's shadow commonly passes above or below the Earth each month at new moon. But two to five times each year, the Moon's shadow reaches the Earth, and a partial, annular, or total eclipse occurs. The Moon is approximately 400 times smaller than the Sun but is also approximately 400 times closer, so its angular diameter in the sky is about the same as the Sun's. Thus the Moon fits approximately exactly over the photosphere, making the phenomenon of a total eclipse especially beautiful.

Phenomena. The partial phases of a total eclipse visible from the path of totality last over an hour. In the minute or two before totality, shadow bands—low-contrast bands of light and dark caused by irregularities in the Earth's upper atmosphere—may be seen to race across the landscape. As the Moon barely covers the Sun, photospheric light shines through valleys on the edge of the Moon, making dots of light—Baily's beads—that are very bright in contrast to the background. The last Baily's bead gleams so brightly that it appears as a jewel on a ring, with the band made of the corona; this appearance is known as the diamond-ring effect (**Fig. 2**). It usually lasts for 5–10 s, and in the clearest skies for as long as 40 s.

During the diamond-ring effect, the solar chromosphere becomes visible around the limb of the Moon, glowing pinkish because most of its radiation is in the form of emission lines of hydrogen, mostly the red hydrogen-alpha line. Its emission-line spectrum apparently flashes into view for a few seconds, and is called the flash spectrum. As the advancing limb of the Moon covers the chromosphere, the corona becomes fully visible (**Colorplate** *a*). Its shape is governed by the solar magnetic field; common are equatorial streamers and polar tufts. At the maximum of the solar activity cycle, so many streamers exist that the corona appears round when it is seen in projection, as viewed from Earth. At the minimum of the solar activity cycle, only a few streamers exist so that the corona appears more elongated in projection. *See* SOLAR CORONA; SUN.

Totality (**Fig. 3**) lasts from an instant up through somewhat over 7 min, 30 s. At its end, the phenomena repeat in reverse order, including chromosphere, diamond ring, Baily's beads, shadow bands, and the partial phases.

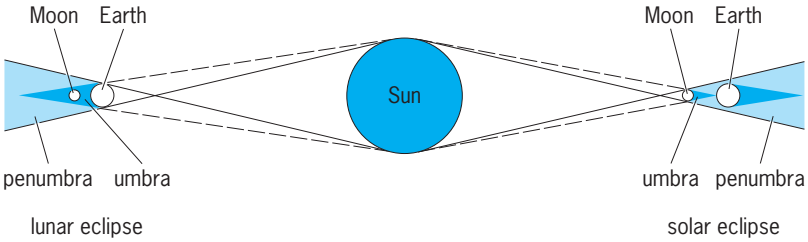

Fig. 1. Circumstances of solar and lunar eclipses (not to scale).

The corona is now monitored continuously in x-rays and in ultraviolet and visible light by orbiting spacecraft such as the *Solar and Heliospheric Observatory* (*SOHO*; ultraviolet and visible light), the *GOES Solar X-ray Images*, and the *Transition Region and Coronal Explorer* (*TRACE*; ultraviolet). The spacecraft *Solar-B*, a successor to the *Yohkoh* spacecraft, was launched in late September 2006 to continue such studies. Eclipse observations can provide ground truth for comparison with these missions on the days of eclipses. *See* SATELLITE (ASTRONOMY).

Positions and timing. The paths of the Sun and Moon in the sky intersect at two points, the ascending node and the descending node. Only when both the Sun and the Moon are near a node can an eclipse occur. Thus eclipse seasons take place each year, whenever the Sun is near enough to the node so that an eclipse is possible. Each eclipse season is 38 days long. Because the Sun's gravity causes the orientation of the Moon's elliptical orbit to change with an 18.6-year cycle, the nodes slide along the ecliptic and a cycle of two eclipse seasons—an eclipse year—has a period of 346.6 days, shorter than a solar year. *See* MOON.

Fig. 2. Diamond-ring effect at the beginning of totality during the solar eclipse of March 29, 2006, observed from Kastellorizo, Greece. (*Jay M. Pasachoff/Science Faction, Williams College Eclipse Expedition*)

Fig. 3. Solar eclipse of April 8; 2005, observed during totality from a ship in the mid-Pacific Ocean, west of the Galápagos Islands. (*Jay M. Pasachoff and Dava Sobel/Science Faction*)

There must be at least one solar eclipse each eclipse season, so there are at least two each year. There may be as many as five solar eclipses in a calendar year (technically, a year), though most of these will be partial. Some of the partials will occur near the ends of the eclipse seasons, and it is also possible for there to be three eclipse seasons during a solar year. Adding lunar eclipses (including penumbral lunar eclipses, which may not be noticeable), there may be seven eclipses in a year.

Saros. An important coincidence relates lunar months and eclipse years. A total of 223 lunar months (technically, synodic months, the period of the phases) takes 6585.32 days. A total of 19 eclipse years—passages of the Sun through the same node of the Moon's orbit—takes 6585.78 days, and 242 nodical months—passages of the Moon through the node—take 6585.36 days. (Nodical months are also called draconic months, after the ancient Chinese dragon once thought to have been devouring the Sun at a solar eclipse.) Thus eclipses appear with this period of 18 years $11^1/_3$ days (plus or minus a day, depending on leap years), a period known as the saros. Further, 239 periods of the variation of distance of the Moon from the Earth, the anomalistic month, is 6585.54 days, so the relative angular sizes of the Sun and Moon are about the same at this interval. (The anomalistic month differs from the nodical and sidereal months because the orientation of the Moon's elliptical orbit drifts around in its orbital plane.)

As a result of the saros, almost identical eclipses recur every 18 years $11^1/_3$ days. The significance of the $^1/_3$ day is that the Earth rotates one-third of the way around, and the eclipse path is shifted on the Earth's surface. Thus the June 30, 1973, 7-min eclipse in Africa was succeeded in a saros series by the July 11, 1991, eclipse in Hawaii, Mexico, and Central and South America, which reached maximum duration of 6 min 54 s in Mexico, and will lead to the 6 min 40 s eclipse of July 22, 2009. After a saros, the Sun is slightly farther west than its original position, and the Moon is slightly north or south, depending on whether it is near an ascending or a descending node, so the eclipses in a saros drift from north to south or from south to north, starting near one pole and departing from the other. A complete series takes 1244 to 1514 years.

Motion over Earth's surface. The Moon's shadow travels at approximately 2100 mi/h (3400 km/h) through space. The Earth rotates in the same direction that the shadow is traveling; at the Equator, the resulting motion of the Earth's surface makes up about 1040 mi/h (1670 km/h), making the speed of the eclipse across the Earth's surface 1060 mi/h (1730 km/h). When eclipses cross higher latitudes, the speed of motion associated with the Earth's rotation is not as high, so the eclipse speed is even higher. The supersonic Concorde took advantage of an equatorial eclipse in 1973 to keep up with totality for 74 min. Ordinary jet aircraft cannot keep up with eclipses, so the term eclipse chasing is almost never accurate.

Scientific value. Even with advances in space technology, total solar eclipses are the best way of seeing the lower corona. Coronagraphs, telescopes for which special shielding and internal occulting allow observation of the corona from a few sites in the world on many of the days of the year, are limited to the innermost corona or to use of special filters or polarization. *See* CORONAGRAPH; SOLAR CORONA.

Coronagraphs have been sent into orbit, notably aboard *Skylab*, the *Solar Maximum Mission*, and *SOHO*, but limitations in spacecraft control lead to the necessity of overocculting the photosphere, cutting out the inner corona. For example, the coronagraph on *Solar Maximum Mission* occulted 1.75 times the solar diameter. The innermost coronagraph, no longer in operation, on *SOHO*'s Large Angle Spectrographic Coronagraph (LASCO) overocculted to 1.1 solar radii, and LASCO's two remaining coronagraphs image the Sun from 1.5 to over 30 solar radii. Other types of observations from space also apply to the corona, such as imaging x-ray observations. Further, any space observation is extremely expensive, over 100 times more for *SOHO* than for a given extensive eclipse expedition. So it is useful and necessary to carry on ground-based eclipse observations, ground-based coronagraph observations, and space-based observations to get the most complete picture of the Sun. Further, independent derivations of certain basic quantities must be carried out to provide trustworthy results; there have been cases in which data reduced from space observations had to be restudied because the need for different calibrations had shown up in eclipse work.

Colorplate *b* is a merger of eclipse and *SOHO* spacecraft images, and shows how ground-based images during a total solar eclipse fill a doughnut-shaped region between the inner and outer spacecraft images that is not visible from Earth at other times. The central image from *SOHO* shows the Sun's disk at temperatures around 60,000–80,000 K through a filter showing helium gas. The outer *SOHO* image shows gas at the millions of degrees typical of the Sun's corona, with LASCO's coronagraphs hiding the bright inner corona. Filling in the doughnut ring allows coronal streamers to be traced from their roots on the Sun's surface, through the eclipse corona, and into the solar wind, the expanding outer corona. *See* SOLAR WIND.

Related eclipse studies include use of the advancing edge of the Moon to provide high spatial resolution for radio observations of the Sun and, historically, of celestial radio sources. The atmospheric effects of the removal of incident sunlight from the Earth's atmosphere have also been studied at eclipses.

Historically, the test of the deflection of starlight carried out at the eclipse of 1919 and repeated in 1922 was the first verification of Einstein's general theory of relativity. These results were restudied for the 1979 Einstein centennial, and their accuracy improved. The experiment is a very difficult one, and has been attempted at some eclipses, notably 1970 in Mexico and 1973 in Africa, without improving on the early results. But the effect has been verified to higher accuracy by studies in the radio part of the spectrum and by the observation of gravitational lenses, so optical eclipse tests are no longer necessary. *See* GRAVITATIONAL LENS; RELATIVITY.

Annular eclipses. Central eclipses in which the Moon is sufficiently far from the Earth that it does not cover the solar photosphere are annular. The Moon's umbra is about 232,000 ± 4000 mi (374,000 ± 6400 km) in length while the Moon's distance is 237,000 ± 16,000 mi (382,000 ± 25,000 km), so the umbra sometimes falls short of reaching the Earth's surface. Annular eclipses, like total eclipses, occur about every 18 months on average.

Since the corona is 1,000,000 times fainter than the photosphere, if even 1% of photosphere is showing the corona is overwhelmed by the blue sky and cannot be seen. So most annular eclipses are of limited scientific use.

The annular eclipse of 1984, visible from the southeastern United States, provided 99.8% coverage. The major scientific work carried out involved detailed timing of the Baily's beads in order to assess the size of the Sun. Such work has also been carried out at total eclipses. Some of the results, compared with solar diameters deduced from historical eclipse paths, seem to show a possible shrinking of the Sun by a measurable amount in a time of decades or centuries, which would lead to impossibly large effects on geological time scales; this has led to the suggestion that the Sun could be oscillating in size. But the question of whether any real effect is present has not been settled. The 1984 annular eclipse was so close to total that the corona could even be briefly seen and photographed, although no scientific studies of the corona were made. In the May 10, 1994, annular eclipse, widely viewed across the United States, only 94% of the Sun's diameter was covered, more typical of annular eclipses. The 1999 annular eclipse, whose path crossed Australia, provided 99% coverage, but since the Sun is about 1,000,000 times brighter than the full moon, that still left about 10,000 times more light than on a moonlit night.

TABLE 1. Total solar eclipses, 2005—2024†

Calendar date	Maximum duration of totality (min: s)	Geographic region of visibility
Apr. 8, 2005	0:42	Starts annular in western Pacific Ocean, becomes total in eastern Pacific, then becomes annular again for Costa Rica, Panama, Colombia, and Venezuela
Mar. 29, 2006	4:07	Eastern Brazil, Atlantic Ocean, Ghana, Togo, Benin, Nigeria, Niger, Chad, Libya, Egypt, Turkey, Russia
Aug. 1, 2008	2:27	Northern Canada, Greenland, Arctic Ocean, Russia, Mongolia, China
July 22, 2009	6:39	India, Nepal, Bhutan, Burma, China, Pacific Ocean
July 11, 2010	5:20	South Pacific Ocean, Easter Island, Chile, Argentina
Nov. 13, 2012	4:02	Northern Australia, south Pacific Ocean
Nov. 3, 2013*	1:40	Annular only at beginning and end of path: Atlantic Ocean, Gabon, Congo, Zaire, Uganda, Kenya, then annular in Ethiopia
Mar. 20, 2015	2:47	North Atlantic Ocean, Faeroe Islands, Arctic Ocean, Svalbard
Mar. 9, 2016	4:10	Indonesia (Sumatra, Borneo, Sulawesi, Halmahera), Pacific Ocean
Aug. 21, 2017	2:40	Pacific Ocean, United States (Oregon, Idaho, Wyoming, Nebraska, Missouri, Illinois, Kentucky, Tennessee, North Carolina, South Carolina), Atlantic Ocean
July 2, 2019	4:33	South Pacific Ocean, Chile, Argentina
Dec. 14, 2020	2:10	Pacific Ocean, Chile, Argentina, south Atlantic Ocean
Dec. 4, 2021	1:55	Antarctica
Apr. 20, 2023*	1:16	Total except at beginning and end of path: South Indian Ocean, western Australia, Indonesia, Pacific Ocean
Apr. 8, 2024	4:28	Pacific Ocean, Mexico, United States (Texas, Oklahoma, Arkansas, Missouri, Kentucky, Illinois, Indiana, Ohio, Pennsylvania, New York, Vermont, New Hampshire, Maine), southeastern Canada, Atlantic Ocean

*These are hybrid or annular-total eclipses. They are total along the middle portions of their paths but annular at the ends.

†Data courtesy of M. Littmann, K. Willcox, and K. Espenak, *Totality—Eclipses of the Sun*, Oxford University Press, 1999.

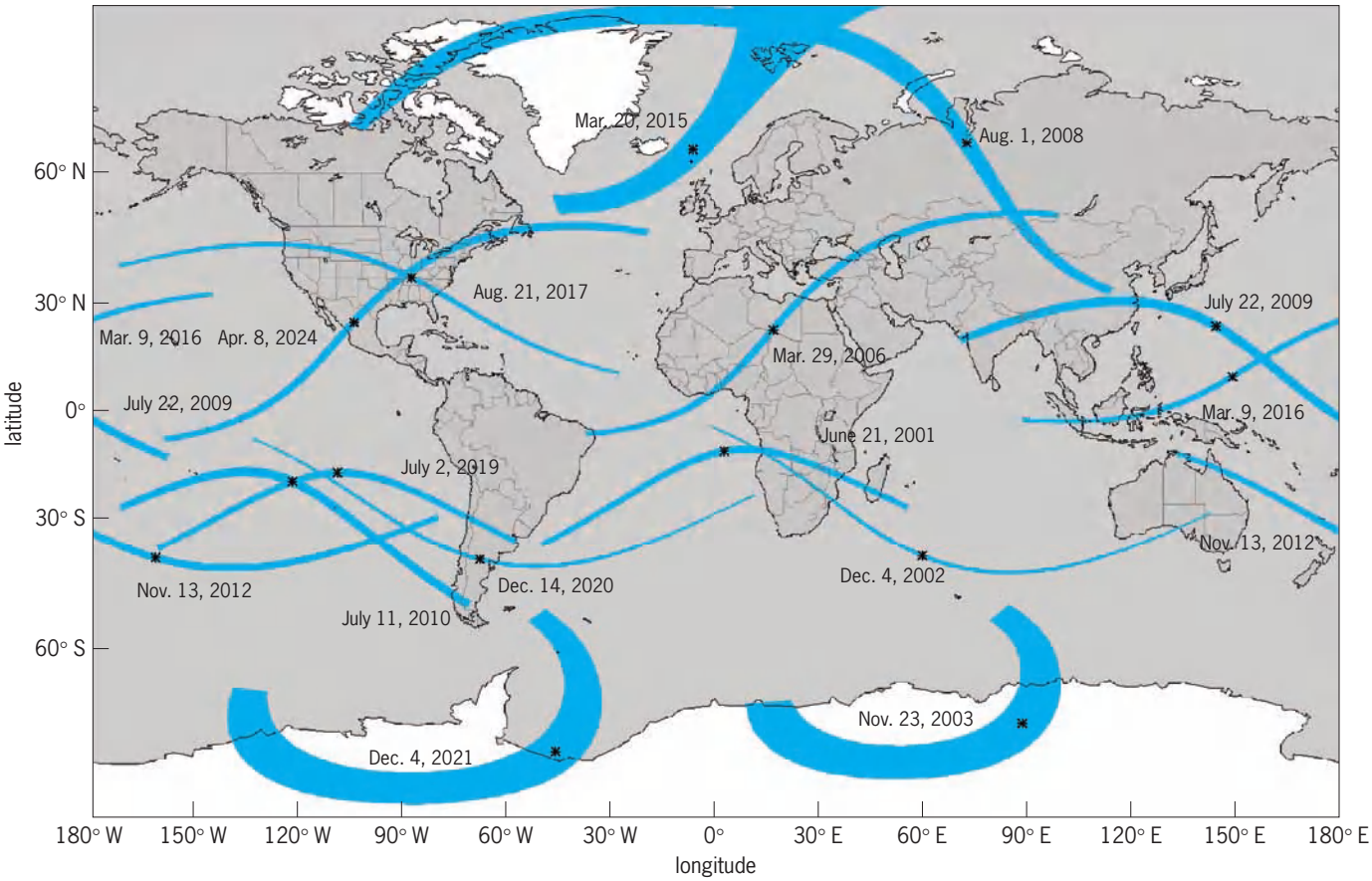

Fig. 4. Paths of all total solar eclipses from 2001 through 2025. (Paths of hybrid annular/total eclipses are not shown). The point of greatest eclipse is indicated by an asterisk on each eclipse path. (*F. Espenak, NASA's Goddard Space Flight Center, http://sunearth.gsfc.nasa.gov/eclipse/eclipse.html*)

The April 8, 2005, hybrid eclipse was annular at the ends of its path and total in the middle, over the Pacific Ocean, where the umbral cone reached the Earth's surface (Fig. 3)

Recent and future eclipses. Notable total eclipses in terms of duration and favorable weather were on June 21, 2001, in southern Africa; on December 4, 2002, in southern Africa and south Australia; and on March 29, 2006, in Africa, Turkey, and into Asia (**Table 1**; **Fig. 4**). The next total solar eclipses are on August 1, 2008, in Mongolia and China; on July 22, 2009, in China; and on July 11, 2010, on Easter Island and in southernmost South America. Major scientific expeditions carry out research on such occasions.

The next total eclipse to cross the continental United States will be in 2017, and the next eclipse to cross Canada will be in 2024.

Observing a solar eclipse. A solar eclipse is nature's most magnificent phenomenon. The false impression has grown that it is often hazardous to observe a solar eclipse, whereas, because of educational efforts on eye safety, actually few reports of eye injury exist and the incidence of lasting eye damage is even lower.

The total phase of an eclipse is completely safe to watch with the naked eye. The total brightness of the corona is only that of the full moon, so is equally safe to watch. The darkness at the disappearance of the diamond ring effect comes so abruptly that people have no trouble telling when the safe time begins, and the diamond ring at the end of totality is so relatively bright that it is clear when it is time to look away. A glance at the Sun just before or after totality is not harmful; it is only staring for an extended time (more than a few seconds with the naked eye) or looking at the not totally eclipsed Sun through binoculars or a telescope that can cause harm.

During the partial phases, it is possible to follow what is going on without any special aid by watching the ground under a tree. The spaces between the leaves make a pinhole camera, and project the solar crescent myriad times onto the ground. A pinhole camera can be made individually by punching a small hole (approximately 0.1–0.2 in. or 2–5 mm in diameter) in a piece of cardboard, and holding it up to the Sun. A crescent image is projected onto a second piece of cardboard held 8 to 40 in. (20 cm to 1 m) closer to the ground, or onto the ground itself. An observer looks at the second cardboard, facing away from the Sun.

For direct observation of the partial phases, a special solar filter must be used. Fogged and exposed black-and-white (not color) film that contains silver and is developed to full density provides suitable diminution of the solar intensity across the entire spectrum. Inexpensive commercial solar filters

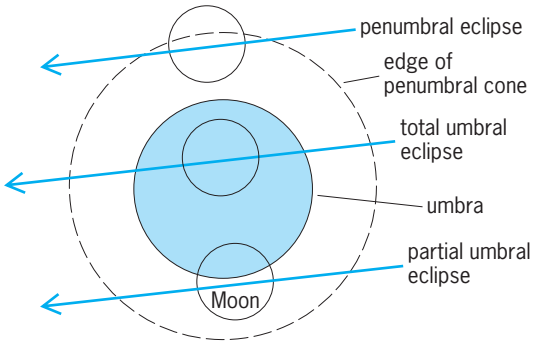

Fig. 5. Cross section through the Earth's shadow cone (umbra) at the Moon's distance, showing the motion of the Moon at three types of lunar eclipses: penumbral eclipse, partial umbral eclipse, and total umbral eclipse.

made of aluminized Mylar can also be used. Gelatin "neutral-density" filters are actually not neutral in the infrared, and so should not be used. Neutral-density filters made by depositing chromium or other metals on glass are safe if they are ND4 or ND5, and are commercially available, as is #14 welder's glass.

Jay M. Pasachoff

Lunar Eclipses

An eclipse of the Moon occurs when the Moon passes through some part of the Earth's two shadows. The inner shadow or umbra is that region where the Earth blocks all direct rays from the Sun. The outer shadow or penumbra is a zone where the Earth blocks part but not all of the Sun's rays. The two shadows result in three types of lunar eclipses (**Fig. 5**):

1. *Penumbral eclipse*. A part or all of the Moon passes through the Earth's penumbral shadow.
2. *Partial eclipse*. A part but not all of the Moon passes through the Earth's umbral shadow.
3. *Total eclipse*. The entire Moon passes through the Earth's umbral shadow.

A lunar eclipse can occur only during the full moon phase. Furthermore, the Moon must also be near one of the two nodes of its orbit. The nodes are the points where the plane of the Moon's orbit intersects the plane of the Earth's orbit. Since the Moon's orbit is tipped 5° to the Earth's orbit (the ecliptic), the Moon passes above or below the shadows during most full moons. But several times each year, full moon occurs near one of the nodes and some portion of the Moon passes through one or both of the Earth's shadows. *See* PHASE (ASTRONOMY).

All partial lunar eclipses begin and end with penumbral phases. Similarly, all total eclipses begin with a penumbral phase followed by a partial phase. Totality is then followed by a partial and a penumbral phase in that order (**Fig. 6**).

Lunar eclipses are visible from much larger geographic areas than are solar eclipses. This is due to the fact that each lunar eclipse can be viewed from the entire night hemisphere of the Earth. Unlike solar eclipses, which require special filters for eye protection, lunar eclipses are completely safe to view using the unaided eye or binoculars.

Visibility. For each type of lunar eclipse, the Moon's maximum depth into either the penumbra or umbra is characterized by a value called the eclipse magnitude. This is the fraction of the Moon's diameter immersed in the shadow. For total eclipses, the umbral eclipse magnitude is always equal to or greater than 1.0.

Penumbral eclipses can last up to 5 h. However, these events are of only academic interest since they are quite faint and difficult to see. Only the deepest penumbral eclipses (penumbral eclipse magnitude greater than 0.7) are visible to the unaided eye as a subtle shading along one edge of the Moon.

Partial eclipses are far easier to see and appear as an obvious bite or missing section along one edge of an otherwise full moon. They can last as long as $3\,^1/_2$ hours but are usually shorter. Partial eclipses are relatively colorless although the edge of the Earth's umbra may exhibit a reddish tinge when viewed through a telescope. The umbral eclipse magnitude of a partial eclipse is always greater than 0.0 and less than 1.0.

Total lunar eclipses have a maximum duration of nearly $2\,^1/_2$ hours. Their appearance is dramatically different from both penumbral and partial eclipses. When the Moon is completely immersed within the Earth's umbral shadow, it is illuminated only by indirect sunlight that has been refracted through and filtered by the Earth's atmosphere. In the process, most of the blue and green colors present in sunlight are removed. The remaining light is deep red or orange and is greatly dimmed. This gives total lunar eclipses their

Fig. 6. Stages of the total lunar eclipse of January 21, 2000. (*F. Espenak*)

Fig. 7. Total lunar eclipse of December 9, 1992. (*F. Espenak*)

characteristically blood red color, which frightened many early civilizations. If the Earth had no atmosphere, the Moon would be completely black during a total eclipse. *See* REFRACTION OF WAVES; SCATTERING OF ELECTROMAGANETIC RADIATION.

The actual appearance of the Moon during totality can vary from one eclipse to the next. The colors range from dark gray or brown, through various shades of red, to bright orange and yellow. These visual differences are directly related to the quality and quantity of dust suspended in Earth's atmosphere. Smoke from forest fires, airborne particles from dust storms, and ash from volcanic eruptions contribute to the color and darkness of the Moon during totality. When large amounts of these particulates are present, they have a strong filtering effect and produce a correspondingly darker eclipse. For example, the 1991 eruption of Mount Pinatubo in the Philippines pumped large amounts of volcanic ash 115,000 ft (35 km) high into the atmosphere where it was transported globally via the jet stream. Red sunsets were seen around the world, and the total lunar eclipse of December 1992 was so dark it rendered the Moon nearly invisible during totality (**Fig. 7**).

Frequency. The geometry and recurrence of lunar eclipses have much in common with solar eclipses. Lunar eclipses occur during eclipse seasons lasting about 39 days with a cycle of two seasons every 346.6 days. Lunar eclipses with similar characteristics repeat every 6585.32 days (about 18 years 11 days), which is the same Saros cycle described above for solar eclipses.

Every lunar eclipse is either preceded or followed by a solar eclipse within a time span of 15 days. Of course, the lunar eclipse will occur at full moon, while the solar eclipse will occur 2 weeks earlier or later at new moon. The link between lunar and solar eclipses means that there are a maximum of seven eclipses within a calendar year. They can occur in combinations of five and two, or four and three, with either solar or lunar eclipses in the majority.

In an average century, there are 243 lunar eclipses consisting of 89 penumbral, 84 partial, and 70 total eclipses. Every year has a minimum of two and a maximum of five lunar eclipses.

Future lunar eclipses. The twenty-first century will have 230 eclipses of the Moon, consisting of 87 penumbral, 58 partial, and 85 total eclipses. During this epoch, 76 years will have two lunar eclipses each, 18 years will have three eclipses, and 6 years will have four eclipses. No year contains more than four eclipses. The last time five lunar eclipses occurred during one calendar year was 1879, and the next will be 2132.

During the 22-year period 2004 through 2025, there will be 20 total eclipses of the Moon (**Table 2**). Deep eclipses have large magnitudes and long

TABLE 2. Total lunar eclipses, 2004–2025*

Date†	Umbral magnitude	Duration of totality	Geographic region of eclipse visibility
May 4, 2004	1.309	1 h 16 min	South America, Europe, Africa, Asia, Australia
Oct. 28, 2004	1.313	1 h 21 min	Americas, Europe, Africa, Central Asia
Mar. 3, 2007	1.238	1 h 14 min	Americas, Europe, Africa, Asia
Aug. 28, 2007	1.481	1 h 31 min	Eastern Asia, Australia, Pacific, Americas
Feb. 21, 2008	1.111	0 h 51 min	Central Pacific, Americas, Europe, Africa
Dec. 21, 2010	1.262	1 h 13 min	Eastern Asia, Australia, Pacific, Americas, Europe
June 15, 2011	1.705	1 h 41 min	South America, Europe, Africa, Asia, Australia
Dec. 10, 2011	1.110	0 h 52 min	Europe, Eastern Africa, Asia, Australia, Pacific, North America
Apr. 15, 2014	1.296	1 h 19 min	Australia, Pacific, Americas
Oct. 8, 2014	1.172	1 h 00 min	Asia, Australia, Pacific, Americas
Apr. 4, 2015	1.006	0 h 12 min	Asia, Australia, Pacific, Americas
Sept. 28, 2015	1.282	1 h 13 min	Eastern Pacific, Americas, Europe, Africa, western Asia
Jan. 31, 2018	1.321	1 h17 min	Asia, Australia, Pacific, western North America
July 27, 2018	1.614	1 h 44 min	South America, Europe, Africa, Asia, Australia
Jan. 21, 2019	1.201	1 h 03 min	Central Pacific, Americas, Europe, Africa
May 26, 2021	1.016	0 h 19 min	Eastern Asia, Australia, Pacific, Americas
May 16, 2022	1.419	1 h 26 min	Americas, Europe, Africa
Nov. 8, 2022	1.364	1 h 26 min	Asia, Australia, Pacific, Americas
Mar. 14, 2025	1.183	1 h 06 min	Pacific, Americas, western Europe, western Africa
Sept. 7, 2025	1.367	1 h 23 min	Europe, Africa, Asia, Australia

*Data courtesy of Fred Espenak, http://sunearth.gsfc.nasa.gov/eclipse/eclipse.html.
†Date of mid-eclipse (Greenwich Mean Time).

durations (for example, June 15, 2011, and July 27, 2018).

Fred Espenak

Bibliography. F. Espenak, *Fifty Year Canon of Lunar Eclipses: 1986–2035*, NASA RP-1178, Sky Publishing, 1989; F. Espenak, *Fifty Year Canon of Solar Eclipses: 1986–2035*, NASA RP-1216, Sky Publishing, 1987; L. Golub and J. M. Pasachoff, *Nearest Star: The Surprising Science of Our Sun*, Harvard University Press, 2001; P. Guillermier and S. Koutchmy, *Total Eclipes: Science, Observations, Myths and Legends*, Springer-Praxis, 1999; M. Littman, K. Willcox, and F. Espenak, *Totality: Eclipses of the Sun*, 2d ed., Oxford University Press, 1999; B.-L. Liu and A. D. Fiala, *Canon of Lunar Eclipses 1500 B.C.—A.D. 3000*, Willmann-Bell, 1992; J. Meeus and H. Mucke, *Canon of Lunar Eclipses: −2002 to +2526*, 3d ed., Astronomisches Büro, Vienna, 1992; H. Mucke and J. Meeus, *Canon of Solar Eclipses: −2003 to +2526*, 2d ed., Astronomisches Büro, Vienna, 1992; J. M. Pasachoff, *A Field Guide to the Stars and Planets*, 4th ed., updated, Houghton Mifflin, 2006; J. M. Pasachoff, *The Complete Idiot's Guide to the Sun*, Alpha Books, 2003; J. B. Zirker, *Total Eclipses of the Sun*, expanded ed., Princeton University Press, 1995.

Eclipsing variable stars

Double star systems in which the two components are too close to be seen separately but which reveal their duplicity by periodic changes in brightness as each star successively passes between the other and the Earth, that is, eclipses the other. Studies of the light changes and the radial velocity changes of each component permit the computation of the radii, masses, and densities of the components—important quantities that cannot be measured directly in single stars. In addition, these close double stars are useful in studies of mass loss and of stellar evolution. Since eclipsing stars are variable in light, they are included in general variable star catalogs under the same system of nomenclature. *See* BINARY STAR; VARIABLE STAR.

Periods. The periods of light variation range from less than 3 h for very close systems to over 27 years for the peculiar system Epsilon Aurigae. However, the majority of the periods lie between 0.5 and 10 days. In many cases the periods are not constant but change with time. In a few cases the variation is caused by a slow change in the orientation of the major axis of an elliptical orbit; in such cases the rate of change combined with other quantities gives information concerning the manner in which the density of the star increases from the outer layers to the center. In most cases, however, the changes are unpredictable and are probably connected with ejections of matter from one of the stars.

Velocity curves. The radial velocity (velocity of approach or recession) of each star can be determined at any time by the displacement of the spectral lines. A plot of velocities against time over one period of orbital revolution is known as a velocity curve. The maximum radial velocity of approach or recession depends on the true orbital velocity of the star and the "inclination," or the amount by which the plane of the orbit is tilted relative to the line of sight to the Earth. (Technically, the inclination is the angle between a perpendicular to the orbit plane and the line of sight; when the inclination is 90°, the eclipses are central.) If the inclination is known, the orbital velocity of each star can be calculated from the radial velocity. The orbital velocity multiplied by the period will give the circumference of the orbit and from this the radius of each star about the center of mass. From the size of the orbit and the period, by using the law of gravitation, the mass of each star can be calculated in terms of the Sun's mass. However, the inclination cannot be determined from the velocity curve alone; the quantities finally determined are $m_1 \sin^3 i$, $m_2 \sin^3 i$, and $a \sin i$, where i is the inclination, $m_1$and m_2 the stellar masses, and a the radius of a circular orbit or half the major axis of an elliptical one.

Light curve. The light curve shows the changes in brightness of the system throughout one orbital revolution. The manner in which the light changes during the eclipse of each star by the other depends very strongly on three factors: the size of each star relative to the radius of the orbit (r_1/a andr_2/a) and the inclination. The determination of precisely what relative sizes and inclination give a computed curve which will approximate satisfactorily the observations is one of the more difficult problems of modern astronomy. However, by use of complex sets of tables the problem can be solved in many cases. Numerical values for these three quantities are thus found. Several alternate methods have been developed which make use of the capabilities of electronic computers. These take into account simultaneously the various interaction effects such as distortion of the shape of the components in the earlier methods. These had to be removed from the light curve before a solution could be carried out. *See* ECLIPSE; LIGHT CURVES.

Absolute dimensions. The masses, radii, and densities of the stars, usually expressed in comparison with those of the Sun, are called the absolute dimensions of the system. Combining the results from the light and velocity curves yields this fundamental information which cannot be obtained from either approach alone. The inclination, determined from the shape of the light curve, can be substituted in the quantities $m_1 \sin^3 i$, $m_2 \sin^3 i$, and $a \sin i$ to find m_1, m_2, and a. Thus the masses of the stars in terms of the Sun's mass and the size of the orbit in terms of the Sun's radius can be found. Since the radii of the stars in fractions of the size of the orbit have been determined from the light curve, the sizes of the stars relative to the size of the Sun (either in miles or kilometers) can now be computed. **Figure 1** shows the relative sizes and separation of the stars in a typical eclipsing system.

Complications. Before the light curve can be "solved" to give the above quantities, corrections must be made for other factors which influence

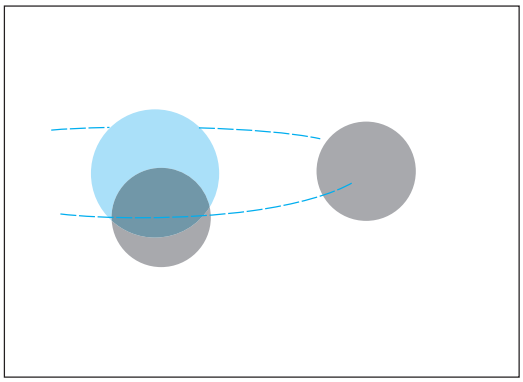

Fig. 1. Relative sizes and separation of components of R Canis Majoris. Shaded star represents cooler component. It is shown in two positions: at the middle of an eclipse and when separation would be greatest as viewed from Earth. Even at the greatest separation, all such systems appear as single stars.

the light changes. This is done by studying the light between eclipses where, were it not for these complications, the brightness of the system would not change. One of these effects is called ellipticity. The tidal attraction of each star for the other has caused distortions until the stars in extreme cases resemble footballs more than baseballs in shape. The technical term is prolate ellipsoids, although when the stars differ in size and mass they will also differ in shape. Further, the radiation of each star falling on the side of the other nearest it will cause this side to be brighter than the side turned away. The difference will be most marked for the cooler component where the effect of the intense radiation from the nearby hotter star is most strongly evident.

There are other effects, some very poorly understood, which cause light changes between eclipses, and all of these must be carefully studied before the analysis of the eclipse begins.

Evolutionary changes. Studies of single stars indicate that, when the hydrogen in the center of the star has been converted into helium, the star undergoes a relatively rapid expansion in size. The presence of a nearby companion complicates the picture considerably, but it does seem clear that much of the mass of the star must be lost to the system or possibly transferred to the other component.

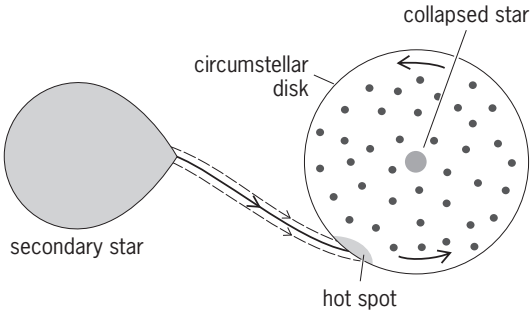

Fig. 2. Model generally accepted for certain types of eclipsing binary stars, in particular those known as dwarf novae.

The mass-losing star eventually becomes a collapsed object—a white dwarf, a neutron star, or a black hole. Each of these types has been identified in at least one binary system. Then when the secondary, originally less massive star begins its expansion, matter from it is transferred to the collapsed object, often with dramatic results. *See* BLACK HOLE; NEUTRON STAR; WHITE DWARF STAR.

It is now believed that all explosive variables (novae, recurrent novae, and so forth), with the exception of supernovae, are members of close binary systems. At least some of the x-ray sources are close binaries in this state, although the detection (by instruments carried aboard satellites) of x-radiation from Algol and other systems which are not yet in this state indicates that other physical mechanisms may also be responsible. Some systems show intermittent bursts of radiation at radio frequencies. Evidence indicates the presence of clouds of circumstellar material.

In a few of the eruptive variables, particularly those known as dwarf novae, rapid scintillation is found, presumably from a hot spot where the transferring mass collides violently with a circumstellar disk of relatively low-density material revolving around the collapsed star (**Fig. 2**); the scintillation stops periodically when the spot is eclipsed by the other component. Instruments on satellites have extended observations to the far ultraviolet, as well as the x-ray, regions of the spectrum. Thus, in addition to the classical reasons for studying eclipsing variable stars, observation of them leads into many branches of astrophysics. *See* ASTROPHYSICS, HIGH-ENERGY; CATACLYSMIC VARIABLE; NOVA; STAR; STELLAR EVOLUTION; X-RAY ASTRONOMY. Frank Bradshaw Wood

Bibliography. A. H. Batten (ed.), Algols, *Space Sci. Rev.*, 50 (1–2):384, 1989; A. H. Batten, *Binary and Multiple Systems of Stars*, 1973; E. F. Milone, *Light Curve Modeling of Eclipsing Binary Stars*, 1995; J. Sahade and F. B. Wood,*Interacting Binary Stars*, 1978.

Ecliptic

The apparent path of the Sun across the sky over the course of a year. The Earth's mean orbit defines the ecliptic plane, and an observer on Earth sees the Sun, traveling in this plane over the course of a year, follow the ecliptic. The orbits of the Moon and planets are at slight angles to the ecliptic, but these objects never appear too far in the sky from the ecliptic, which is often drawn on star maps. The ecliptic was identified by Greek and Babylonian scientists in about the fifth century B.C.

The ecliptic intersects the celestial equator at an angle of approximately $23^1/_2{}^\circ$, an angle known as the obliquity of the ecliptic. The points of intersection are the vernal equinox and the autumnal equinox. As a result of this obliquity, the Sun can be overhead at different times of year for locations on Earth between latitudes $+23^1/_2{}^\circ$ and $-23^1/_2{}^\circ$, that is, between the Tropic of Cancer and the Tropic of

Capricorn, respectively. *See* ASTRONOMICAL COORDINATE SYSTEMS; EARTH ROTATION AND ORBITAL MOTION. Jay M. Pasachoff

Bibliography. J. Mitton, *Cambridge Dictionary of Astronomy*, Cambridge University Press, 2001; J. M. Pasachoff and A. Filippenko, *The Cosmos: Astronomy in the New Millennium*, 3d ed., Brooks/Cole, 2007; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, University Science Books, 1992.

Eclogite

A very dense rock composed of red-brown garnet and the grape-green pyroxene omphacite. Eclogites possess basaltic bulk chemistry, and their garnets are rich in the components pyrope ($Mg_3Al_2Si_3O_{12}$), almandine ($Fe_3Al_2Si_3O_{12}$), and grossular ($Ca_3Al_2Si_3O_{12}$), while the pyroxenes are rich in jadeite ($NaAlSi_2O_6$) and diopside ($CaMgSi_2O_6$). Accessory phases depend on the environment of crystallization, but may include quartz, kyanite, amphibole, olivine, orthopyroxene, zoisite, mica, and rutile, but never plagioclase.

Occurrences. Eclogite occurrences may be subdivided into three broad categories:

Group A eclogites are found as layers, lenses, or boudins in schists and gneisses seemingly of the amphibolite facies. Quartz, together with zoisite or kyanite, commonly occurs in these rocks. Amphibole of barroisitic composition may also be present. These rocks originate as metamorphosed basic bodies (dykes, sills, flows) or layers of dolomitic marl intercalated with metasedimentary schists. *See* QUARTZ.

Group B eclogites are found as inclusions in kimberlites and basalts. They are frequently accompanied by xenoliths of garnet peridotite; both eclogites and peridotites can be diamond-bearing, suggesting that these rocks come from deep in the mantle. *See* BASALT.

Group C eclogites are found as blocks and lenses in schists of the glaucophane-schist facies. Such eclogites do not contain kyanite, rarely contain quartz, but bear amphibole, epidote, rutile, or sphene. The presence of pillow lavas converted to eclogite indicates that these rocks had a crustal origin.

Conditions of formation. Eclogites can form over a wide range of pressure and temperature, but are clear indicators of rather high pressure, as their very high density might suggest. Experimental work has shown that rocks of basaltic composition are converted to eclogite at high pressures (see **illus.**).

The distribution of Fe^{2+} and Mg^{2+} between garnet and omphacite has been determined as a function of temperature, and allows approximate temperature regimes to be assigned to the three categories of eclogite. Eclogites of group C formed at relatively low temperatures and probably crystallized isofacially with rocks of the glaucophaneschist facies. Group A eclogites probably formed over a wider range of temperature and pressure, but at somewhat higher temperatures than group C. The very high temperatures estimated for group B eclogites are consistent with their origin in the upper mantle. *See* GARNET; OMPHACITE.

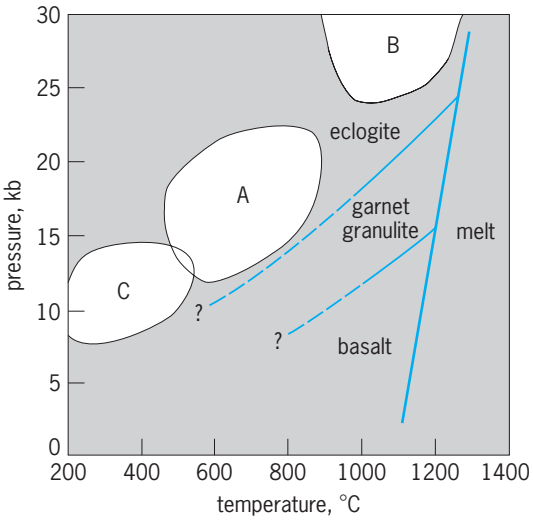

Possible pressure-temperature fields for eclogites from gneiss areas: A from kimberlites; B from blueschist terrains; C basalt-eclogite transition. 1 kilobar = 10^5 kilopascals. °F = (°C × 1.8) + 32.

Eclogite facies. Eclogite is the name given to the highest-pressure facies of metamorphism; the critical mineral assemblage defining this facies is garnet + omphacite, together with kyanite or quartz in rocks of basaltic composition. Where sedimentary and granitic rocks have been metamorphosed under eclogite facies conditions, they result in spectacular omphacite + garnet + quartz-bearing mica schists and metagranitic gneisses such as are found in the Sezia-Lanzo zone of the western Alps. *See* METAMORPHIC ROCKS.

The relationship between the eclogite and other facies is often unclear. Eclogites, for instance, can form in relatively anhydrous parts of blueschist and amphibolite facies terrains where normally their hydrated equivalents would crystallize. Some such eclogites are inferred, on textural grounds, to have formed in a dry environment. In other cases the coarse textures, the presence of veins of diagnostic eclogite facies minerals such as omphacite + kyanite + quartz, and fluid inclusions suggest that their formation was accompanied by copious fluid. Although very high pressures are required to stabilize eclogite under hydrous conditions, the presence of diluent species (for example, CO_2 or halides) in the fluid lowers the required pressures.

Eclogites and the mantle. The high density of eclogite, together with its elastic properties, makes it a candidate for upper mantle material. Large quantities of basaltic oceanic crust are returned to the mantle through the process of subduction, where prevailing high pressures convert it to eclogite. The quantity and distribution of eclogite within the mantle is not known; that it occurs is known from the nodules brought up in kimberlite pipes and in basalts. *See* HIGH-PRESSURE PROCESSES. Timothy J. B. Holland

Bibliography. M. G. Best, *Igneous and Metamorphic Petrology*, 1982; B. W. Evans and E. H. Brown (eds.), *Blueschists and Eclogites*, 1986; L. A. Raymond, *Petrology: The Study of Igneous, Sedimentary, and Metamorphic Rocks*, 1994; D. C. Smith (ed.), *Eclogites and Eclogite-Facies Rocks*, 1989; M. Suk, *Petrology of Metamorphic Rocks*, 1983.

Ecological communities

Assemblages of living organisms that occur together in an area. The nature of the forces that knit these assemblages into organized systems and those properties of assemblages that manifest this organization have been topics of intense debate among ecologists since the early years of this century. On the one hand, there are those who view a community as simply consisting of species with similar physical requirements, such as temperature, soil type, or light regime. The similarity of requirements dictates that these species be found together, but interactions between the species are of secondary importance and the level of organization is low. On the other hand, there are those who conceive of the community as a highly organized, holistic entity, with species inextricably and complexly linked to one another and to the physical environment, so that characteristic patterns recur, and properties arise that one can neither understand nor predict from a knowledge of the component species. In this view, the ecosystem (physical environment plus its community) is as well organized as a living organism, and constitutes a superorganism. Between these extremes are those who perceive some community organization but not nearly enough to invoke images of holistic superorganisms. *See* ECOSYSTEM.

Emergent and collective properties. The crux of this debate is whether communities have emergent properties, a concept that itself is sufficiently confused, so that part of the ecological debate may be semantic. Some denote by emergent property any property of a group that is not a property of the component individuals. Others call any group trait a collective property, and reserve as emergent only those collective properties that do not derive trivially from properties of component individuals or from the very definition of the group. An emergent property, in this conception, represents a new level of organization, and cannot be predicted even with complete knowledge of all properties of individuals in the group. Thus, the number of species in a community is a property that is not defined for any particular species, but once one knows all the species in a community one can easily measure the property "number of species" for that community. In the restricted view of emergent property, the number of species in a community is a collective property, but not an emergent one.

Whether any properties of communities are emergent is debatable, but all ecologists agree that communities have collective properties that are significant both biologically and practically. For example, it is not a trivial matter to understand why the number of species in a given community is S and not $2S$, and what the consequences are of having S species. Also, human exploitation of ecological communities often rests on a thorough knowledge of certain collective properties, such as biomass produced per unit time, or rate of cycling of certain nutrients. *See* BIOMASS.

Size. Every community comprises a given group of species, and their number and identities are distinguishing traits. Most communities are so large that it is not possible to enumerate all species; microorganisms and small invertebrates are especially difficult to census. However, particularly in small, well-bounded sites such as lakes or islands, one can find all the most common species and estimate their relative abundances. The number of species is known as species richness, while species diversity refers to various statistics based on the relative numbers of individuals of each species in addition to the number of species. The rationale for such a diversity measure is that some communities have many species, but most species are rare and almost all the individuals (or biomass) in such a community can be attributed to just a few species. Such a community is not diverse in the usual sense of the word. Patterns of species diversity abound in the ecological literature; for example, pollution often effects a decrease in species diversity. The most popular species diversity statistic is given in the equation below, where

$$H' = \sum_{i=1}^{s} p_i \log (p_i)$$

S is the number of species and p_i is the fraction of all the individuals (or all the biomass) in the community contributed by species i. For many sets of data on diversities of groups of communities (or parts of communities, such as all birds or all plants), values of H' are highly correlated with species richness, suggesting that H' may not be the ideal expression of biotic diversity.

The main patterns of species richness that have been detected are area and isolation effects, successional gradients, and latitudinal gradients. Larger sites tend to have more species than do small ones, and isolated communities (such as those on oceanic islands) tend to have fewer species than do less-isolated ones of equal size. Later communities in a temporal succession tend to have more species than do earlier ones, except that the last (climax) community often has fewer species than the immediately preceding one. Tropical communities tend to be very species-rich, while those in arctic climates tend to be species-poor. This observation conforms to a larger but less precise rule that communities in particularly stressful environments tend to have few species.

Composition. Species composition is important, largely in assisting in the classification of communities. If the complete species list for each community and the proportional representation of each species

in it were known, no two communities would be found to be identical. The criteria are arbitrary for how similar these lists must be before two communities are viewed as the same.

Communities are usually denoted by the presence of species, known as dominants, that contain a large fraction of the community's biomass, or account for a large fraction of a community's productivity. Dominants are usually plants. Determining whether communities at two sites are truly representatives of the "same" community requires knowledge of more than just the dominants, however. "Characteristic" species, which are always found in combination with certain other species, are useful in deciding whether two communities are of the same type, though the designation of "same" is arbitrary, just as is the designation of "dominant" or "characteristic."

Although there is no objective answer to the question of whether two communities are the same, the question bears heavily on the matter of whether communities are in fact integrated superorganisms with emergent properties. One would not expect a superorganismic entity to have an infinite variety of representations, for the same cohesive forces that bind it into a holistic unit should constrain it to a limited number of forms. Organisms for the most part do not form a continuous gradient of all conceivable phenotypes, but rather are constrained by a number of physiological and morphological relationships (such as allometry). Thus there is generally no trouble in saying that individual X is a human, individual Y is a longleaf pine, and so on. Although there is phenotypic variation in all species, it is limited and centered on recognizable modes.

An active area of research among animal ecologists has been the search for limits to how similar two species can be and still occur in the same community. That there exists some limiting similarity beyond which coexistence is impossible is a theoretical consequence of conceptions of the ecological niche; but clear evidence, in any instance, that such limiting similarity is the reason why a given species is absent from a given community is exceedingly difficult to amass. There are two main difficulties. First, it must be demonstrated that some resource is limiting so that species that are too similar will find it in short supply. Second, an appropriate index of a species' use of the resource (its niche) must be determined. Much attention has been focused on the size of an animal or of its trophic appendages as such an index, though there is much evidence that many other forces affect these sizes. For few, if any, animal communities has it been shown that size similarity restricts species composition or that patterns of size differences among species are highly predictable or regular.

Spatial distribution patterns. A related matter bearing on the aptness of the superorganismic community metaphor is that most organisms have clear spatial boundaries; communities often do not. There is no difficulty discerning where one human ends and another begins, but there are not always clear boundaries to communities. Occasionally, very sharp limits to a physical environmental condition impose similarly sharp limits on a community. For example, serpentine soils are found sharply delimited from adjacent soils in many areas, and have mineral concentrations strikingly different from those of the neighboring soils. Thus they support plant species that are very different from those found in nearby nonserpentine areas, and these different plant species support animal species partially different from those of adjacent areas. Here two different communities are sharply bounded from each other.

Usually, however, communities grade into one another more gradually, through a broad intermediate region (an ecotone) that includes elements of both of the adjacent communities, and sometimes other species as well that are not found in either adjacent community. One has little difficulty telling when one is in the center of either of the adjacent communities A and B, but exactly when one passes from A to B is not easily discerned. The reason is that, though each species in a community is adapted to its physical environment and to some extent to other species in the community, the adaptations to the physical environment are usually not identical, and most of the adaptations to one another are not obligatory.

The environment created by the dominant species, by their effects on temperature, light, humidity, and other physical factors, and by their biotic effects, such as allelopathy and competition, may entrain some other species so that these other species' spatial boundaries coincide with those of the dominants. The mangrove skipper, *Phocides pygmalion*, can feed only on red mangrove, *Rhizophora mangle*, so whatever aspects of the physical environment limit the mangrove (especially temperature), the skipper will also be limited to the same sites. However, many other species that feed on the mangrove (for example, the io moth, *Automeris io*) also feed on other plants, and their spatial boundaries do not coincide with those of the mangrove. Nor do most of the species in a community share identical physical requirements. Black mangrove (*Avicennia germinans*) co-occurs with red mangrove in most sites, and the two are normal constituents of a community often termed mangrove swamp. But *Avicennia* can tolerate much colder weather than can *Rhizophora*, and so it is found much farther north in the Northern Hemisphere. Eventually it too ceases, and the mangrove community is replaced in more northerly areas by salt marsh communities. In some intermediate regions, salt marsh grasses and black mangrove are found together. There is no clear boundary between the two communities.

This continuous intergradation of most communities argues against the superorganism concept, but there are aspects of the spatial arrangement within communities that suggest that the component species are far from independent, and indicate, if not complete holism, at least that some community properties are not easily predicted from exhaustive knowledge of the component species. One example

is stratification, the vertical arrangement of canopy layers in most forests. Individuals of the different species do not have heights that are independently and continuously distributed from the ground to the top of the tallest tree. Instead, there are a few rather distinct strata, with each species at maturity characteristically occupying one of these. Tropical forests from all parts of the world, even though they may have completely different species compositions, usually contain five fairly clear strata: a topmost layer composed of the tallest tree species, two lower layers of smaller trees, a shrub layer, and a ground layer. There are doubtless good physical reasons why the diffusion of light can explain this characteristic structure given a knowledge of evolution and plant physiology (though no completely compelling explanation has yet surfaced), so it may be that this is an elaborate collective property rather than an emergent one. In either case, there is clearly a high degree of multispecies organization in this spatial arrangement. *See* PHYSIOLOGICAL ECOLOGY (PLANT).

In addition to vertical arrangement, horizontal locations of individuals of different species are usually not random. Usually, individuals of a given species are clumped; they are found on average closer to one another than one would have predicted. Probably the major reason for this is response to habitat heterogeneity: conspecific individuals tend to favor more similar habitats than do heterospecific individuals. Individuals of different species may also be nonrandomly arranged with respect to one another. Competitive interactions may cause two species typically to be found in different microsites, while mutualistic interactions or preference for a similar physical habitat may cause two species to be associated spatially. *See* POPULATION ECOLOGY.

Succession. More or less distinct communities tend to follow one another in rather stylized order. As with recognition of spatial boundaries, recognition of temporal boundaries of adjacent communities within a sere is partly a function of the expectations that an observer brings to the endeavor. Those who view communities as superorganisms are inclined to see sharp temporal and spatial boundaries, and the perception that one community does not gradually become another community over an extended period of time confirms the impression that communities are highly organized entities, not random collections of species that happen to share physical requirements. In this superorganismic view, ecological succession of communities in a sere is analogous to the life cycle of an organism, and follows a quite deterministic course. That secondary succession following a disturbance often leads to a community resembling the original one is the analog in the superorganism to wound repair in an organism. The driving force for succession, in this conception, is that the community or its dominant species modify the environment so that these dominant species are no longer favored, and when the dominant species are replaced, the bulk of the community, complexly linked to the dominant species and to one another, disappears as well, to be replaced by the next community in the sere.

This superorganismic conception of succession has been replaced by an individualistic succession. Data on which species are present at different times during a succession show that there is not abrupt wholesale extinction of most members of a community and concurrent simultaneous colonization by most species of the next community. Rather, most species within a community colonize at different times, and as the community is replaced most species drop out at different times. Thus, though one can usually state with assurance that the extant community is of type A or type B, there are extended periods when it is difficult to assign the assemblage of species at a site to any recognizable community.

That succession is primarily an individualistic process does not mean that there are not characteristic changes in community properties as most successions proceed. Species richness usually increases through most of the succession, for example, and stratification becomes more highly organized and well defined. A number of patterns are manifest in aspects of energy flow and nutrient cycling. *See* ECOLOGICAL SUCCESSION.

Functional organization. Living organisms are characterized not only by spatial and temporal structure but by an apparent purpose or activity. In short, they are "doing something," and this activity has been termed teleonomy. In humans, for example, various physiological functions are continuously under way until death intervenes. Communities have functions analogous to physiology.

In the first place, the various species within a community have different trophic relationships with one another. One species may eat another, or be eaten by another. A species may be a decomposer, living on dead tissue of one or more other species. Some species are omnivores, eating many kinds of food; others are more specialized, eating only plants or only animals, or even just one other species. These trophic relationships certainly unite the species in a community into a common endeavor, the transmission of energy through the community. This energy flow is patently analogous to an organism's mobilization and transmission of energy from the food it eats.

One aspect of energy flow is a candidate for an emergent property: the topology of the food web. Examination of this topology for a number of webs suggests that they are highly constrained in structure. For example, the maximum number of trophic levels in a web rarely exceeds five. One reason may be that the total amount of energy that has not already been degraded by the time the energy has passed through three or four levels may not be enough to sustain a viable population of a species that would feed at still higher levels. An alternative explanation is that the population dynamics of a web with so many levels would probably confer mathematical instability on the web, so that one or more species would be eliminated. Other properties of food webs that have been discerned include a low number of

omnivore species (those feeding on more than one level), and a tendency for the number of predator species and the number of prey species to be in the ratio of 4:3. No explanation is forthcoming for the latter observation. The former observation is held to reflect mathematical instability that arises from the presence of omnivores in a web. Other workers contend that neither pattern will be maintained when much more comprehensive data are available on what paths calories actually follow as they flow through a community. If these patterns do not turn out to be artifacts of incomplete knowledge, they would appear to be emergent properties reflecting a high degree of organization. *See* FOOD WEB.

Just as energy flows through the communities, so do nutrients move. A calorie of energy does not move in the abstract from organism to organism; rather, the calorie is bound up in a molecule that moves when one organism eats (or decomposes) another. Or the calorie may be respired away as a result of the metabolism of the organism that ingests it. A calorie of energy, once respired by some member of the community, is no longer available to the community. But the molecule associated with that calorie, or a new molecule produced from it, is still present and can go through the food web again. Thus nutrients can cycle within a community, while energy flow, once the energy is transformed to heat, is one-way. Nutrient cycling is analogous to circulation in an organism, and combines with energy flow to support the superorganism metaphor. Different nutrients cycle at different rates, and for several nutrients the cycle within the community is linked to cycles in other communities. Nutrients exist in abiotic entities, as well as biotic organisms, so nutrient cycling is as much an ecosystem trait as a community one. A number of properties of nutrient cycling (such as rate, turnover times, and sizes of different pools or compartments) have been measured. Whether any of these are emergent as opposed to collective properties has yet to be determined.

Productivity. By virtue of differing rates of photosynthesis by the dominant plants, different communities have different primary productivities. Tropical forests are generally most productive, while extreme environments such as desert or alpine conditions harbor rather unproductive communities. Agricultural communities are intermediate. Algal communities in estuaries are the most productive marine communities, while open ocean communities are usually far less productive. The efficiency with which various animals ingest and assimilate the plants and the structure of the trophic web determine the secondary productivity (production of organic matter by animals) of a community. Marine secondary productivity generally exceeds that of terrestrial communities. *See* AGROECOSYSTEM; BIOLOGICAL PRODUCTIVITY.

Reproduction. A final property that any organism must have is the ability to reproduce itself. Communities may be seen as possessing this property, though the sense in which they do so does not support the superorganism metaphor. A climax community reproduces itself through time simply by virtue of the reproduction of its constituent species, and may also be seen as reproducing itself in space by virtue of the propagules that its species transmit to less mature communities. For example, when a climax forest abuts a cutover field, if no disturbance ensues, the field undergoes succession and eventually becomes a replica of the adjacent forest. Both temporally and spatially, then, community reproduction is a collective rather than an emergent property, deriving directly from the reproductive activities of the component species. *See* ALTITUDINAL VEGETATION ZONES; BOG; CHAPARRAL; DESERT; ECOLOGY; GRASSLAND ECOSYSTEM; MANGROVE; MUSKEG; PARAMO; PUNA.

Daniel Simberloff

Bibliography. T. F. H. Allen and T. B. Starr, *Hierarchy: Perspectives for Ecological Complexity*, 1988; M. Begon, M. Mortimer and D. J. Thompson, *Population Ecology: A Unified Study of Animals and Plants*, 3d rev. ed., 1996; P. G. Digby and R. A. Kempton, *Multivariate Analysis of Ecological Communities*, 1987; J. R. Trabalka and D. E. Reichle (eds.), *The Changing Carbon Cycle*, 1986.

Ecological competition

The interaction of two (or more) organisms (or species) such that, for each, the birth or growth rates are depressed and the death rate increased by the presence of the other organisms (or species). Competition is recognized as one of the more important forces structuring ecological communities, and interest in competition led to one of the first axioms of modern ecology, the competitive exclusion principle. The principle suggests that in situations where the growth and reproduction of two species are resource-limited, only one species can survive per resource.

Lotka-Volterra equations. The competitive exclusion principle was originally derived by mathematicians using the Lotka-Volterra competition equations. This model of competition predicts that if species differ substantially in competitive ability, the weaker competitor will be eliminated by the stronger competitor. However, a competitive equilibrium can occur if the negative effect of each species on itself (intraspecific competition) is greater than the negative effect of each species on the other species (interspecific competition). Because the competitive exclusion principle implies that competing species cannot coexist, it follows that high species diversity depends upon mechanisms through which species avoid competition. *See* MATHEMATICAL ECOLOGY.

Species coexistence. One way that the problem of competitive coexistence has been addressed is to extend the competitive exclusion principle to a continuum of resources through the theory of limiting similarity. Limiting similarity is defined as the maximum level of similarity in the use of a set of resources that will allow competing species to coexist. However, this approach still assumes that several species

cannot coexist on the same resources unless they use those resources differently.

A more realistic approach to understanding the coexistence of competing species is to view community structure as dynamic, in which species densities do not remain constant over time at each spatial location. Fluctuations in densities may be a result of floods, fires, and other abiotic disturbances, or of biotic interactions, such as predation, herbivory, and disease. They key to coexistence within a homogeneous patch in such a dynamic world is the prevention of competitive equilibrium and hence competitive exclusion.

In general, competitive exclusion can be prevented if the relative competitive abilities of species vary through time and space. Such variation occurs in two ways. First, dispersal rates into particular patches may fluctuate, causing fluctuations in the numerical advantage of a species in a particular patch. This may occur, for example, through differences among species in dispersal ability. Second, competitive abilities of species may be environmentally dependent and, therefore, fluctuate with local environmental changes. Different species will be favored under different sets of environmental conditions, allowing each species to have periods of strong recruitment.

Competitive exclusion can also be avoided if fluctuations in environmental factors reduce the densities of potentially competing species to levels where competition is weak and population growth is for a time insensitive to density. Therefore, intensities of competition within and between species fluctuate with time, and competing species are able to coexist.

In patchy environments, local extinctions or population reductions resulting from predators or environmental fluctuations can allow species to colonize and grow for some time in patches in the absence of competition. In this way, even inferior competitors can persist if they are, for example, relatively less susceptible to predation or are relatively better dispersers.

Coexistence is not merely a result of environmental harshness or fluctuations but also involves the critical element of niche differentiation (that is, species must differ from one another if they are to coexist). However, the focus is not how species coexist by partitioning resources, but how species can coexist on the same resources by differing sufficiently in their responses to environmental conditions and fluctuations. *See* ECOLOGICAL COMMUNITIES; ECOLOGICAL SUCCESSION.

Applications. Competition theory has been applied to human-manipulated ecosystems used to produce food, fiber, and forage crops as well as in forestry and rangeland management. Although many characteristics of agricultural systems are similar to those of natural ecosystems, agricultural communities are unique because they are often managed for single-species (sometimes multispecies) production and they are usually characterized by frequent and intense disturbance. Studies of competition in agriculture have primarily examined crop loss from weed abundance under current cropping practices, and have evaluated various weed control tactics and intercropping systems. Because both total density and species proportion influence the outcome and interpretation of competition experiments, the ability to quantify the intensity of competition and to separate the effects of competition between individuals of the same species versus competition between individuals of different species (such as weed versus crop) depends on experimental design. Experimental approaches that systematically vary total and relative plant densities provide better quantification of competition in agroecosystems. Factors that influence competition in agroecosystems include the timing of plant emergence, growth rates, spatial arrangements among neighbors, plant-plant-environment interactions, and herbivory. *See* ECOLOGY; ECOLOGY, APPLIED.

Paul C. Marino

Bibliography. E. F. Connor and S. Simberloff, Competition, scientific method, and null models in ecology, *Amer. Sci.*, 74:155–162, 1986; S. R. Reice, Nonequliibrium determinants of biological community diversity, *Amer. Sci.*, 82:424–435, 1994; J. Weins, On competition and variable environments, *Amer. Sci.*, 65:590–597, 1977.

Ecological energetics

The study of the flow of energy within an ecological system from the time the energy enters the living system until it is ultimately degraded to heat and irretrievably lost from the system. It is also referred to as production ecology, because ecologists use the word production to describe the process of energy input and storage in ecosystems.

Ecological energetics provides information on the energetic interdependence of organisms within ecological systems and the efficiency of energy transfer within and between organisms and trophic levels. Nearly all energy enters the biota by green plants' transformation of light energy into chemical energy through photosynthesis; this is referred to as primary production. This accumulation of potential energy is used by plants, and by the animals which eat them, for growth, reproduction, and the work necessary to sustain life. The energy put into growth and reproduction is termed secondary production. As energy passes along the food chain to higher trophic levels (from plants to herbivores to carnivores), the potential energy is used to do work and in the process is degraded to heat. The laws of thermodynamics require the light energy fixed by plants to equal the energy degraded to heat, assuming the system is closed with respect to matter. An energy budget quantifies the energy pools, the directions of energy flow, and the rates of energy transformations within ecological systems. *See* BIOLOGICAL PRODUCTIVITY; FOOD WEB; PHOTOSYNTHESIS.

The peak of studies in ecological energetics occurred in the 1960s and early 1970s largely because a major concern of the International Biological

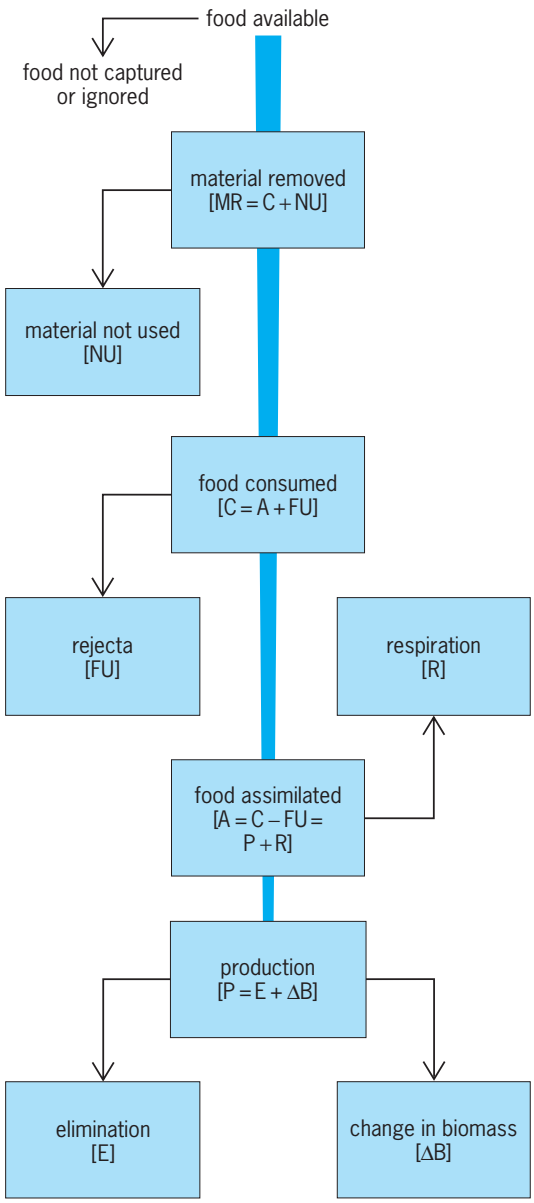

Fig. 1. Diagrammatic representation of energy flow through an ecological unit; abbreviations are explained in the text.

Program was an appraisal of the biological productivity of terrestrial and aquatic communities in relation to human welfare. Initially considered to have the potential of becoming a unifying language in ecology—an ecological Rosetta Stone—the subject has yielded little in the way of general theory.

The essentials of ecological energetics can be most readily appreciated by considering the schema (**Fig. 1**) of energy flowing through an individual; it is equally applicable to populations, communities, and ecosystems. Of the food energy available, only part is harvested (MR) in the process of foraging. Some is wasted (NU), for example, by messy eaters, and the rest consumed (C). Part of the consumed food is transformed but is not utilized by the body, leaving as fecal material (F) or as nitrogenous waste (U), the by-product of protein metabolism. The remaining energy is assimilated (A) into the body, part of which is used to sustain the life functions and to do work—this is manifest as oxygen consumption. The remainder of the assimilated energy is used to produce new tissue, either as growth of the individual or as development of offspring. Hence production is also the potential energy (proteins, fats, and carbohydrates) on which other organisms feed. Production (P) leads to an increase in biomass (B) or is eliminated (E) through death, migration, predation, or the shedding of, for example, hair, skin, and antlers.

Pathways. Energy flows through the consumer food chain (from plants to herbivores to carnivores) or through the detritus food chain. The latter is fueled by the waste products of the consumer food chain, such as feces, shed skin, cadavers, and

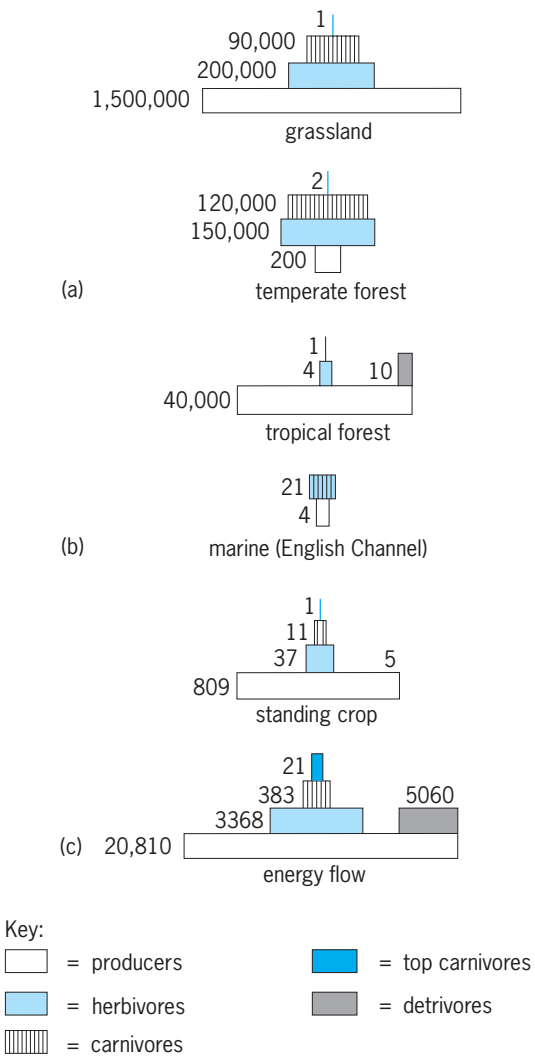

Fig. 2. Trophic levels of a number of ecosystems represented in different units. (*a*) As numbers of individuals per 1000 m^2 of grassland and temperate forest community in summer; microorganisms and soil animals excluded. (*b*) The standing crop or biomass (grams dry weight per meter squared) of terrestrial (Panamanian tropical rainforest) and marine (English Channel) communities; note the inversion of the marine pyramid. (*c*) The aquatic community of Silver Springs, Florida, represented as standing crop (kilocalorie per meter) and energy flow (kilocalories per meter per year). (*After E. P. Odum, Fundamentals of Ecology, 3d ed., W. B. Saunders, 1971*)

nitrogenous waste. Most detritus is consumed by microorganisms such as bacteria and fungi, although this food chain includes conspicuous carrion feeders like beetles and vultures. In terrestrial systems, more than 90% of all primary production may be consumed by detritus feeders. In aquatic systems, where the plants do not require tough supporting tissues, harvesting by herbivores may be efficient with little of the primary production passing to the detrivores.

Pyramids of biomass are used to depict the amount of living material, or its energetic equivalent, present at one time in the different trophic levels (**Fig. 2**). Although the energy flow cannot increase at higher trophic levels, pyramids of biomass may be inverted, especially in aquatic systems. This occurs because the index P/B is inversely related to the size of the organisms. Hence a small biomass may support a high level of production if the biomass is composed of small individuals (**Fig. 3**).

Units. Traditionally the calorie, a unit of heat energy, has been used, but this has been largely replaced by the joule. Confusion is possible, especially in the field of nutrition, because with an initial capital, Calorie may denote kilocalories. Biomass or standing crop is expressed as potential energy per unit area, but the other compartments in Fig. 1, for example *P* and *R*, are expressed in terms of energy flux or rates. The efficiency values such as P/A are dimensionless, but the ratio P/B is a rate—the inverse of the turnover time.

Measurement of energy flow. For illustrative purposes some general methods for assessing biological productivity are described here in the context of energy flow through a population. Production is measured from individual growth rates and the reproductive rate of the population to determine the turnover time. The energy equivalent of food consumed, feces, and production can be determined by measuring the heat evolved on burning a sample in an oxygen bomb calorimeter, or by chemical analysis—determining the amount of carbon or of protein, carbohydrate, and lipid and applying empirically determined caloric equivalents to the values. The latter three contain, respectively, 16.3, 23.7, and 39.2 kilojoules per gram of dry weight. Maintenance costs are usually measured indirectly as respiration (normally the oxygen consumed) in the laboratory and extrapolated to the field conditions. Error is introduced by the fact that animals have different levels of activity in the field and are subject to different temperatures, and so uncertainty has surrounded these extrapolations. Oxygen consumption has been measured in animals living in the wild by using the turnover rates of doubly labeled water (D_2O).

Levels of inquiry. Ecological energetics is concerned with several levels of inquiry: the partitioning of energy between the compartments denoted in Fig. 1; the pathways traced by the energy as it passes through the trophic levels; and the efficiency of energy transfer between trophic levels. The ratio of energy flux through one compartment in Fig. 1 to any previous compartment is referred to as an efficiency. Numerous efficiencies can be calculated both within and between trophic levels. The most common are the assimilation efficiency (A/C), namely the proportion of energy assimilated by the body from the food consumed, and the production efficiency (P/A), which denotes the proportion of energy assimilated which ends up as new tissue. These various efficiencies combine to limit the energy available to the higher trophic levels. The ratio of food consumed or ingested at one trophic level to that ingested by the next lower level is termed ecological efficiency. A value of 10% for this efficiency is often cited; consideration of the A/C and P/A efficiencies

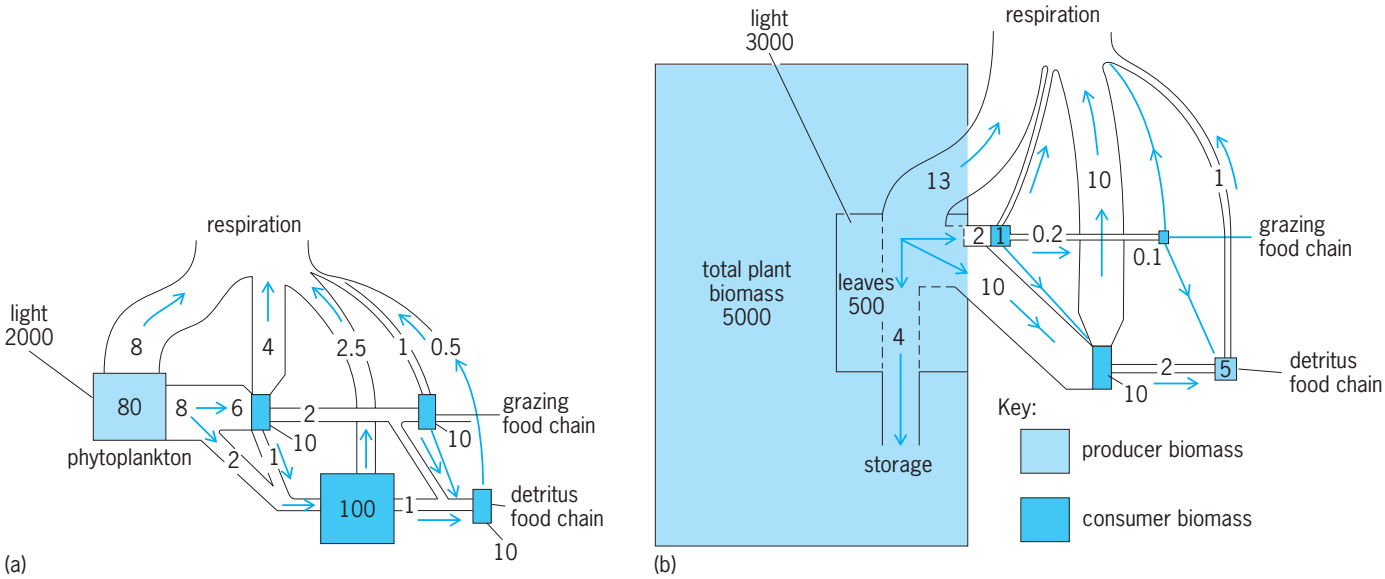

Fig. 3. Models of energy flow through two contrasting ecosystems: (*a*) a marine bay and (*b*) a forest. The boxes represent the biomass in kilocalories per meter and the flow lines show the energy flux (kilocalories per meter per day) along the various pathways. The boxes and the flux lines are scaled to indicate their relative magnitudes. (*After E. P. Odum, Relationships between structure and function in the ecosystem, Jap. J. Ecol., 12:108–118, 1962*)

Production efficiency of populations of various classes of animals living in the wild*

Animal group	Production efficiency (P/A), %
Shrews	0.9
Birds	1.3
Other mammals	3.1
Fish, ants, and termites	9.8
Invertebrates other than insects	25.0
Herbivores	20.8
Carnivores	27.6
Detrivores	36.2
Insects except ants and termites	40.7
Herbivores	38.8
Detrivores	47.0
Carnivores	55.6

*After W. F. Humphreys, Production and respiration in animal populations, *J. Anim. Ecol.*, 48:427–454, 1979.

of most organisms shows that it could seldom exceed 15–20%. However, the effect of heat losses at each trophic level in limiting the length of food chains in nature remains controversial.

Factors affecting efficiency. Respiration rate of organisms is scaled as the three-quarters power of body weight. Hence larger organisms have proportionately slower rates of respiration. This scaling factor seems to affect many rate processes in the body so that size does not influence those efficiencies which are the focus of ecological energetics. However, different types of organisms of the same size have different metabolic rates. For example, warm-blooded animals have much higher weight-specific respiration rates than cold-blooded ones. Analysis of energy budgets derived for wild-living animals shows that a number of taxonomic and trophic groups can be distinguished according to characteristic production efficiencies (see **table**). Production efficiency appears to be related to the general level of metabolic activity—animals with high rates of metabolism generally having low production efficiency.

Due to the loss of usable energy with each transformation, in an area more energy can be diverted into production by plants than by consumer populations. For humans this means that utilizing plants for food directly is energetically much more efficient than converting them to eggs or meat. *See* BIOMASS; ECOLOGICAL COMMUNITIES; ECOSYSTEM.

W. F. Humphreys

Bibliography. S. Brody, *Bioenergetics and Growth*, 1945; R. M. May, Production and respiration in animal communities, *Nature*, 282:443–444, 1979; E. P. Odum, *Fundamentals of Ecology*, 3d ed., 1971; K. Petrusewicz and A. Macfadyen, *Productivity of Terrestrial Animals: Principles and Methods*, 1970; J. Phillipson, *Ecological Energetics*, 1966; S. L. Pimm, *Food Webs*, 1982.

Ecological methods

The methods required for the measurement of features of plant and animal populations or for quantifying energy flow or other aspects of ecological communities. Ecologists are concerned with such a wide range of variables that they may use many other techniques, including in particular those for the measurement of weather conditions or for determining features and composition of air, water, and soil. *See* ECOLOGICAL COMMUNITIES.

Patterns of distribution. Ecologists often attempt to describe the way that individual plants or animals are distributed in the habitat. One pattern is the clumping of organisms. This may occur because only part of the habitat is suitable for them (as when rushes occur in the damp patches of a field) or because they represent the offspring of an original colonist and have not moved far from the parent. Another pattern is known as random distribution because each organism is independent of the other (as when thistledown settles on a plowed field). A regular pattern describes the situation when organisms are virtually equidistant from each other. It is often a reflection of competition.

Such patterns are normally determined by sampling the organisms from unit areas—generally squares known as quadrats. For vegetation sampling, long pins resembling knitting needles are dropped into the ground and plants that touch them are recorded. These pins are regarded as infinitely small quadrats and are termed point quadrats. Various statistics, such as the mean ($\overline{x}$) and the variance (s^2), may be calculated from the results of such sampling, and their magnitudes and relationships to each other will show whether the pattern is clumped, random, or regular: if the ratio $s^2/\overline{x} \sim 1$ the individuals are randomly distributed; if the ratio is greater than 1 they are clumped; but if less than 1 they are in a regular pattern. *See* STATISTICS.

Estimation of populations. For plants, individuals (such as trees) may be counted directly, and the main problem is often (as with grasses) deciding what is a plant. However, the populations of only a few animals can be estimated by direct counting; in these cases, the organism is either sessile (like barnacles) or the individuals are very large, and all the individuals in an area can be seen (or at least recorded on a photograph) at the same time. Aerial photography and remote sensing techniques have provided powerful approaches for the estimation of the numbers of large animals (such as cattle or elephants) over wide areas or the extent and type of vegetation cover. However, most animals are active and small, and only a part (sample) of the population can be seen at any one time. Alternatively, it may be tedious and time-consuming to count the animals, so only those in small parts (samples) can be counted (for example, mites per leaf rather than per tree). *See* AERIAL PHOTOGRAPHY; REMOTE SENSING.

One approach to estimating population size is to mark a number of the animals (a), perhaps with a colored dye or paint, and release these into the population. If it can be assumed that they have mixed completely (and the marking has not affected their behavior), the proportion that they constitute of the next sample will be the same as the proportion that the original number marked constitutes of the

total population (N). This is shown in the equation below, where r is the number of marked individu-

$$\frac{r}{n} = \frac{a}{N} \qquad \text{or} \qquad N = \frac{an}{r}$$

als recaptured, n is the total number in the sample after marking, a is the number originally marked, and N is the population. This calculation also assumes that there are no births or deaths, or any migration in the period between the samples. Often these assumptions are not justified, and many more complex systems, of both sampling and marking (on several occasions) and of calculation, have been devised.

An alternative approach is to take a number of samples and determine the total number of animals in them. The average number per sample is then multiplied by the total number of sample units in the habitat (for example, leaves on a tree) to obtain the population for that habitat. Special techniques are often needed to determine the total animals in a sample. For example, mites and insects in samples of soil (or litter) are usually extracted by placing the soil in a special apparatus, the Berlese-Tüllgren funnel, where a light bulb or other heat source slowly dries out the soil so that the animals move down and eventually drop through the funnel to a collecting tube. Other soil animals (for example, many types of worms) require moist conditions, and they can be driven out of wet samples by lowering the oxygen level (again by heat) or by chemicals. There is a wide variety of methods for different organisms and for various habitats. Special methods may also be necessary to take the samples: a common method of sampling insects from grassland is the D-vac sampler, basically a large, specially designed vacuum cleaner. *See* MATHEMATICAL ECOLOGY.

Population dynamics. The aim of population estimation is often to contribute toward an understanding of how the population is changing over time, such as to determine trends or the major mortalities. Methods are devised to measure the amount of mortality (for example, the remains of dead animals or plants can be found, or the level of parasitism assessed), the numbers born (natality), and the organisms moving in (immigration) and out (emigration) of the habitat. Special computational techniques are used to estimate the numbers of individuals in a population and to construct life tables. All these methods make certain assumptions, but since these are not always justified, there is great value in making population estimates by a number of different methods. *See* POPULATION ECOLOGY.

Energy flow. The contribution of organisms of different sizes to the functioning of an ecosystem may be best compared in terms of their total weight—the biomass. The biomass in a habitat is termed the standing crop. It is measured in wet weight or dry weight, the latter obtained by freeze drying. The dynamics of an ecosystem, that is, the exchanges between the different components, may be expressed in terms of energy transfers to form a picture of the energy flow. The energy content of a given biomass is often determined by burning a sample in a bomb calorimeter; the heat rise is a measure of the calorific value of the sample. Then it is possible to determine the amount of animal or plant material produced (the production) and, for animals, the amount of food consumed. *See* BIOLOGICAL PRODUCTIVITY; BIOMASS.

Two equations are particularly useful for determining the energy relationships of organisms:

$$\text{Consumption} - (\text{fecal and urinary waste}) = \text{assimilation}$$

$$\text{Assimilation} = \text{respiration and production}$$

Consumption and production can usually be measured in terms of biomass and calorific value. Respiration is often measured with special respirometers, and assimilation can sometimes be measured directly by using an inert marker in the food.

Other methods are used to measure the flow of specific nutrients in ecosystems; special computational methods are used to calculate indices to describe the diversity of animal and plant communities; many types of traps or similar devices are used to capture animals; and radioactive isotopes can be used as markers for organisms or materials. Ecologists are always devising new methods for the many and varied problems they encounter. *See* ECOLOGICAL ENERGETICS; ECOLOGY; ECOSYSTEM. T. R. E. Southwood

Bibliography. R. Clarke (ed.), *The Handbook of Ecological Monitoring*, 1986; P. A. Erickson, *Environmental Impact Assessment: Principles and Applications*, 1987; J. A. Ludwig and J. F. Reynolds, *Statistical Ecology: Primer on Methods and Computing*, 1988; R. Southwood and P. A. Henderson, *Ecological Methods*, 3d ed., 2000; H. Walter and S. W. Breckle, *Ecological Systems of the Geobiosphere*, 1985.

Ecological modeling

The use of computer simulations or mathematical equations to address questions that cannot be answered solely by experiments or observations. Ecological models have two major aims: to provide general insight into how ecological systems or ecological interactions work; and to provide specific predictions about the likely futures of particular populations, communities, or ecosystems.

Models can be used to indicate general possibilities or to forecast the most likely outcomes of particular populations or ecosystems. Models differ in whether they are "basic" or are intended to address management decisions. As ecology has grown in its sophistication, models are increasingly used as decision support tools for policy-makers. *See* POPULATION ECOLOGY.

Areas of modeling. The earliest ecological models were produced by mathematicians attempting to explain cycles of predators and prey, disease epidemics, pest outbreaks, and human population

growth. These early models are the foundations of ecology and exemplify the value of mathematics as a tool for formulating biological questions. For example, the Italian mathematician Vito Volterra was asked to explain why the fish that are eaten by sharks decreased in relative abundance when fishing halted in the Mediterranean Sea during World War I. The models developed by Volterra not only accounted for this decline in shark prey with the war-related curtailment of harvest but also indicated that insect pests might actually increase in numbers if chemical insecticides killed both their predators and the insects themselves [a phenomenon observed 40 years later after the invention of dichlorodiphenyltrichloroethane (DDT) and other chemical insecticides]. More generally, Volterra's models predicted a tendency for predators and their prey to fluctuate even if the environment stayed the same from year to year—another prediction that is generally upheld by data.

Models of human population growth stimulated the development of mathematical tools for converting schedules of birth and death into predictions about annual rates of population change. Although these demographic models originated over 50 years ago, modifications of them provide the foundation for much modern conservation biology. For instance, by inputting rates of mortality and reproduction for loggerhead sea turtles into a simple model, it was found that extinction of the turtles could best be prevented if large young adults were protected from shrimp trawlers; legislation for this was passed by the U.S. Congress.

Models of virtually every possible type of ecological interaction have been developed (competition, parasitism, disease, mutualism, plant-herbivore interactions, and so forth). The models vary in their level of detail. Some models simply keep track of the density of organisms, treating all organisms of any species as identical (mass action models). At the other extreme, the movement and fate of each individual organism may be tracked in an elaborate computer simulation (individual behavior models).

Building a model. The best way to understand how ecological modeling works is to actually build a model. One of the most important general models in ecology describes the consequences of habitat loss for species that require specific types of habitat. The type of model described below is called a metapopulation model because it keeps track of many populations in an archipelago of patches. In order to model how such species are likely to be influenced by losses of habitat, ecologists have examined models that range from simple ("pencil-and-paper") exercises that can be expressed with a few equations to intricate computer models that involve hundreds of complex mathematical relationships and lengthy computer programs.

A useful pencil-and-paper model about the consequences of habitat destruction may be devised by carefully expressing ideas in mathematical language. For example, begin by simplifying the world to patches of habitat that are either occupied by the species of interest (a fraction x) or not occupied by the species $(1 - x)$. No species will occupy habitat forever, since there will inevitably be local death or even disappearance of an entire population within some patch of forest or alpine meadow. Conversely, empty or unoccupied habitat will not remain vacant forever, as long as there is some source of colonists that can find their way to the open habitat. A simple equation (1) summarizes the change

$$\textbf{change} = \begin{pmatrix} \text{gains in } x \\ \text{due to colonization} \\ \text{of empty patches} \end{pmatrix} - \begin{pmatrix} \text{losses in } x \\ \text{due to disappearance} \\ \text{from occupied patches} \end{pmatrix} \quad (1)$$

in species abundance expected in such a world of habitat patches (occupied or unoccupied) with perpetual turnover; here **change** represents the change in the fraction of occupied patches per unit time.

This expression may be made more precise by hypothesizing that there is a constant loss or disappearance of the species from occupied patches at a rate e; this means the "minus term" of Eq. (1) becomes $-ex$. It is reasonable to assume that vacant patches of habitat $(1 - x)$ get colonized in proportion to the fraction of occupied habitat from which colonists might arise (c times x, where c indicates the dispersal or colonization ability of the species, and x is the proportion of habitat occupied as mentioned above). This is a specific model that predicts the fraction of habitat patches that should be occupied by an organism, and how that fraction changes through time [Eqs. (2)–(4)]. The species will increase as long

$$\textbf{change} = cx(1 - x) - ex \quad (2)$$

$$\textbf{change} = cx(1 - x - D) - ex \quad (3)$$

$$(1 - D - e/c) \quad (4)$$

as **change** is greater than zero. **Change** is greater than zero if x is less than $(1 - e/c)$; and **change** is less than zero if x is greater than $(1 - e/c)$. Hence x will tend to move toward the fraction given by $(1 - e/c)$, and it should be clear why $(1 - e/c)$ can be called the equilibrium fraction of habitat occupied by the species of interest. As long as colonization rates are sufficiently high (that is, c is greater than e), the species will persist forever. Habitat destruction permanently removes a fraction D of the vacant habitat; thus, instead of $(1 - x)$ fraction of habitat being available to be colonized, only $(1 - x - D)$ habitat is available to be colonized. With habitat destruction, Eq. (2) is modified to Eq. (3). With habitat loss incorporated through the D term, the equilibrium fraction of habitat occupied by a species is now (4). Thus, if habitat loss D is too large, the species will become extinct because the equilibrium will be less than zero, which means the species will decline until it has totally disappeared. This model makes several important points. For example, even though there is empty habitat (which means $1 - x$ is greater than 0), the species may still relentlessly

decline simply because so much habitat has been destroyed that the colonization of empty habitat occurs too infrequently to sustain the species. This exposes the fallacy of arguments that have been made regarding old growth forests and the endangered spotted owl species in the Pacific Northwest: the notion that there is ample old growth forest for this endangered species because some of the forest is devoid of owls is clearly wrong. The model also shows that there is a threshold for habitat destruction (for example, when habitat loss D becomes greater than $1 - e/c$), which when crossed implies inevitable gradual decline of a species to extinction.

A key feature of ecological modeling is a parameter estimation. The parameters in the model related to Eq. (4) are e, c, and D. If the value of these parameters was known, it would be possible to predict whether or not the species of interest could persist, and if it could persist, what fraction of habitat would be expected to be occupied by that species. Thus, in order to apply the model even crudely to any situation, it is necessary to estimate the extinction rate or disappearance rate (fraction lost per year or week), the colonization rate, and the fraction of habitat that has been lost altogether. This might seem like an easy task, but in most circumstances the factor that most thwarts ecological modelers is estimating the parameters. Because parameters may be difficult to estimate, there is typically a great deal of uncertainty surrounding the application of any model. For example, suppose it is known that the disappearance rate for spotted owls in old growth forest patches is 3% loss per year and the colonization rate is 5%, but it is uncertain whether or not 35% or 45% of the habitat has been destroyed. If $D = .35$ the species could persist (because $1 - .35 - .03/.05$ is greater than 0), but if $D = .45$ the species would go extinct (because $1 - .45 - .03/.05$ is less than 0). Any uncertainty about habitat loss D would yield uncertainty about the fate of the species.

Simple algebraic models such as Eqs. (1)–(4) are very useful for indicating general principles and possibilities. They were part of the testimony considered in court cases regarding the halting of logging in Washington and Oregon so that spotted owls would not cross the threshold for habitat loss, at which point the species would be doomed. But something as simple as Eq. (4) could never be practically applied to the management of any specific forest and spotted owl population. In order to be a management tool, the model must be more complicated and detailed to reflect the specific situation under examination. For example, instead of a few equations, ecologists have modeled spotted owl populations and old growth forests in Washington using a detailed computer simulation that keeps track of habitat in real maps at the scale of hectares. In these simulation models, owls are moved as individuals from one hectare to another, and their fate (survival, death, or reproduction) is recorded in the computer's memory. By tracking hundreds or even thousands of owls moving around in this computer world, different forestry practices corresponding to different logging scenarios can be examined. *See* ECOLOGY, APPLIED; MATHEMATICAL ECOLOGY; SYSTEMS ECOLOGY.

Advantages and limitations. A model is a formal way of examining the consequences of a series of assumptions about how nature works. Such models refine thinking and clarify what results are implied by any set of assumptions. As models become more complicated and specific, they can also be used as surrogate experimental systems; in other words, ecologists may use models to conduct experiments that are too expensive or impractical in the field. For example, to protect against extinction of northern spotted owls in the northwestern United States, investigators do not have the time or the resources to engage in thousand-acre experiments that entail different forestry practices (for example, logging in a way that leaves behind clusters of old growth forest versus logging randomly across the landscape). However, if enough is known about how owls move through forests, a simulation model can be devised that "does the experiment," which may suggest the best logging practice.

One danger of ecological modeling is the uncertainty of the models and the shortage of supporting data. Properly used, models allow exploration of a wide range of uncertainty, pointing out the limits of current knowledge and identifying critical information required prior to management decision making. Increasingly, the public will be asked to contribute to decisions about logging, sprawling urban development, removing dams from rivers, and so forth; critical information regarding these decisions will be provided by ecological models. It is therefore essential that citizens understand how models are formulated and applied. Conclusions based on models should not be quickly dismissed because the model may be the best line of reasoning available; however, it would not be prudent to rely solely on the output of any model. Ecological modeling is a branch of science that requires expertise in ecology, mathematics, statistics, and usually computer programming. Done well, ecological modeling can be tremendously valuable as a source of general scientific principles and as a decision-support tool in arenas of resource management and conservation. *See* ECOLOGY. Peter Kareiva

Bibliography. N. Gotelli, *A Primer of Ecology*, 2d ed., Sinauer Press, Sunderland, MA, 1998; J. Maynard Smith, *Models in Ecology*, Cambridge University Press, Cambridge, 1974; J. Roughgarden, *A Primer of Ecological Theory*, Prentice Hall, Upper Saddle River, NJ, 1998.

Ecological succession

A directional change in an ecological community. Populations of animals and plants are in a dynamic state. Through the continual turnover of individuals, a population may expand or decline depending on the success of its members in survival and reproduction. As a consequence, the species composition of communities typically does not remain static with

time. Apart from the regular fluctuations in species abundance related to seasonal changes, a community may develop progressively with time through a recognizable sequence known as the sere. Pioneer populations are replaced by successive colonists along a more or less predictable path toward a relatively stable community. This process of succession results from interactions between different species, and between species and their environment, which govern the sequence and the rate with which species replace each other. The rate at which succession proceeds depends on the time scale of species' life histories as well as on the effects species may have on each other and on the environment which supports them. *See* ECOLOGICAL COMMUNITIES; POPULATION ECOLOGY.

The course of ecological succession depends on initial environmental conditions. Primary succession occurs on novel areas such as volcanic ash, glacial deposits, or bare rock, areas which have not previously supported a community. In such harsh, unstable environments, pioneer colonizing organisms must have wide ranges of ecological tolerance to survive. In contrast, secondary succession is initiated by disturbance such as fire, which removes a previous community from an area. Pioneer species are here constrained not by the physical environment but by their ability to enter and exploit the vacant area rapidly.

As succession proceeds, many environmental factors may change through the influence of the community. Especially in primary succession, this leads to more stable, less severe environments. At the same time interactions between species of plant tend to intensify competition for basic resources such as water, light, space, and nutrients. Successional change results from the normal complex interactions between organism and environment which lead to changes in overall species composition. Whether succession is promoted by changing environmental factors or competitive interactions, species composition alters in response to availability of niches. Populations occurring in the community at a point in succession are those able to provide propagules (such as seeds) to invade the area, being sufficiently tolerant of current environmental conditions, and able to withstand competition from members of other populations present at the same stage. Species lacking these qualities either become locally extinct or are unable to enter and survive in the community.

Primary succession. In some cases, seres may take hundreds of years to complete, and direct observation at a given site is not possible. Adjacent sites may be identified as successively older stages of the same sere, if it is assumed that conditions were similar when each seral stage was initiated.

Glacier Bay. In the Glacier Bay region of Alaska glaciers have retreated, in phases, some 61 mi (98 km) since 1750, leaving a series of moraines of known ages supporting a range of seral vegetational types. Soil factors become modified by vegetation, enabling new species to become established (**Table 1**). Acidic decomposition products of alder leaves sharply reduce soil pH, whereas virtually no change occurs on bare glacial till or under other vegetation. Pioneer species must tolerate low nitrogen levels, but alder is able to fix atmospheric nitrogen (N) by the presence of microbial symbionts in root nodules and is correlated with a rapid increase in soil nitrogen by way of leaf fall. After invasion by spruce, levels of accumulated nitrogen fall as nitrogen becomes incorporated in forest biomass, and the annual additions from alder are reduced. Soil organic matter increases progressively, and influences the structural development of the soil. The mature forest remains only on well-drained slopes. In areas of poorer drainage, invasion by *Sphagnum* moss leads to replacement of trees by wet acidic bog or muskeg.

Lake Michigan. Another example of primary succession has been recorded on a sequence of dune ridges bordering Lake Michigan, ranging in age to 12,000 years. Succession is initiated on a bare sand surface either following a fall in lake level, as occurred in phases since the last glaciation, or due to wind erosion of an existing dune redepositing sand (**Table 2**). Marram grass impedes the transport of sand across bare dune surfaces and so promotes accretion of sand. At the same time, it grows and branches vigorously, thus maintaining cover on the expanding dune (**Fig. 1**).

TABLE 1. Successional changes in vegetation and soils observed on an aged series of moraines at Glacier Bay, Alaska

Years	Vegetation	Soil environment
0	Initial colonizers: mosses, fireweed, horsetail, *Dryas*, dwarf willows; later, willows form dense scrub	Initially pH 8.0–8.4 due to $CaCO_3$: soil N and organic matter lacking
50	Alder invades, forming dense thickets less than 33 ft (10 m) tall	pH falls to 5.0 in 30–50 years due to acidic products of alder leaf decomposition; marked increase in soil N due to fixation in alder root nodules; soil organic matter accumulates
170	Sitka spruce invades, forming dense forest	Reduction in soil N by incorporation in forest biomass; progressive increase in soil organic matter
250+	Western and mountain hemlock enter (climax forest on well-drained slopes)	
	Sphagnum bog with occasional pines replaces forest in poorly drained areas	Soil becomes waterlogged, deoxygenated, acidified

TABLE 2. Succession on sand dunes of Lake Michigan

Years	Vegetation	Soil environment	Ground invertebrates
0	Marram grass (*Ammophila*) invades by rhizome migration	Initial sand pH 7.6	
6		Sand accumulates, dune builds up	White tiger beetle (*Cicindela lepida*) Sand spider (*Trochosa cinerea*)
20	Marram declines in vigor	Sand surface stabilized	White grasshopper (*Trimerotropis maritima*) Longhorn grasshopper (*Psinidia fenestralis*) Burrowing spider (*Geolycosa pikei*)
50	Jack pine and white pine become established, with light-demanding understory		Digger wasps (*Bembex, Microbembex*) Bronze tiger beetle (*Cicindela scutellaris*) Ant (*Lasius niger*) Migratory locust (*Melanoplus*) Sand locusts (*Agoeneotettix, Spharagemon*)
100	Black oak becomes established, with shade-tolerant understory		Digger wasp (*Sphex*) Ant lion (*Cryptoleon*) Flatbug (*Neuroctenus*) Grasshoppers (six species) Wireworms (*Elateridae*) Snail (*Mesodon thyroides*)
150			
•			
•			
1000		Soil N builds up to 0.1%	
•			
10,000		Sand pH drops to 4.0 due to $CaCO_3$; soil matures to a nutrient-poor brown humic sand	

Succession on Lake Michigan dunes demonstrates that local factors may modify the typical pattern (**Fig. 2**). In damp depressions with impeded drainage, a grassland community develops. Sheltered pockets on lee slopes have a moister microclimate and tend to accumulate leaf litter from more exposed parts of the dunes, as well as receiving protection from frequent fires. As a result, a more nutrient-rich soil can develop, and succession proceeds via basswood to more nutrient-demanding oak-hickory and finally maple-beech woodland, typical of normal soils of the region. The black oak association appears to be stable on dune soils, since it is tolerant of low nutrients and water limitation, and tends to maintain these conditions by returning few nutrients in the leaf litter.

Climax community. Early stages of succession tend to be relatively rapid, whereas the rates of species turnover and soil changes become slower as the community matures. Eventually an approximation to the steady state is established with a relatively stable community, the nature of which has aroused considerable debate. Earlier, the so-called climax vegetation was believed to be determined ultimately by regional climate and, given sufficient time, any community in a region would attain this universal condition. This unified concept of succession, the monoclimax hypothesis, implies the ability of organisms progressively to modify their environment until it can support the climatic climax community. Although plants and animals do sometimes ameliorate environmental conditions, evidence suggests overwhelmingly that succession has a variety of stable end points. This hypothesis, known as the polyclimax hypothesis, suggests that the end point of a succession depends on a complex of environmental factors that characterize the site, such as parent material, topography, local climate, and human influences.

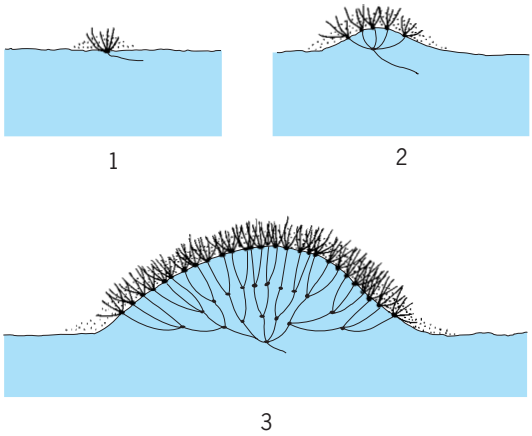

Fig. 1. Dune formation by the gradual deposition of wind-carried sand particles around aerial shoots of *Ammophila arenaria*. 1, 2, and 3 indicate successive periods of time. (*After K. A. Kershaw, Quantitative and Dynamic Ecology, 2d ed., Arnold, 1973*)

In the Lake Michigan sand dunes, the course of succession and its climax appear to be determined by physiographic conditions at the start (Fig. 2). Similarly, the transformation of glacial moraine forest to muskeg depends on local drainage. In the tropical rainforest of Moraballi Creek, Guyana, five apparently stable vegetation types have been distinguished on different soil types under the same climate. A mixed forest is present on red loam, whereas the Wallaba forest occurs on bleached sand, with the Mora forest type in areas liable to flooding.

Autogenic vs. allogenic factors. In the examples of succession discussed above, the chief agent in modifying the environment is the community itself: thus marram stabilizes the sand dune surface, and alder increases the soil nutrient status of moraine soil. These actions of the community on the environment, termed autogenic, provide an important driving force promoting successional change, and

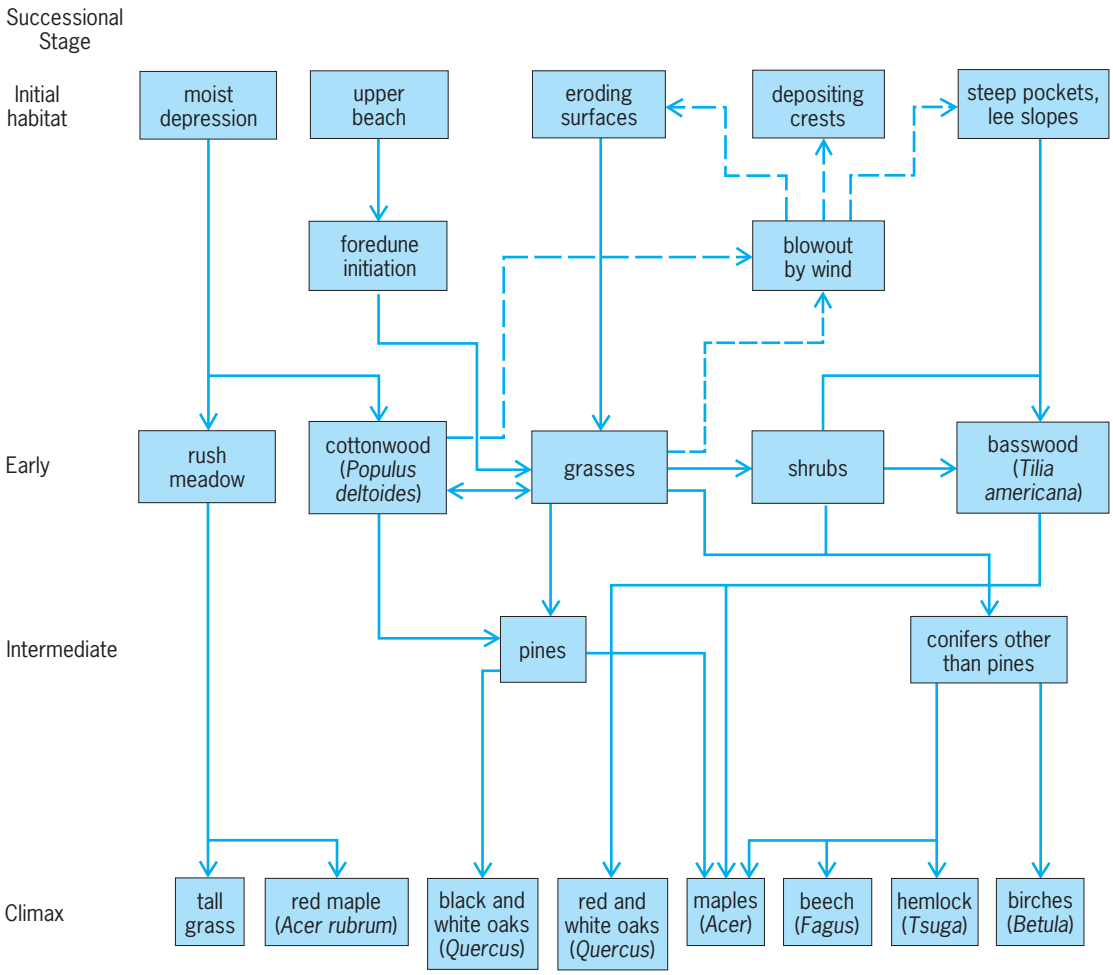

Fig. 2. Alternative patterns of primary plant succession on Lake Michigan dunes, depending on initial conditions, extrinsic variables, and colonization-invasion patterns. (*After S. J. McNaughton and L. L. Wolf, General Ecology, 2d ed., Holt, Reinhart, and Winston, 1979*)

are typical of primary succession where initial environments are inhospitable. Alternatively, changes in species composition of a community may result from influences external to the community called allogenic. For example, in aquatic ecosystems (the hydrosere) the community commonly develops from open water with submerged and floating aquatic plant species toward a swamp community in which rooted emergent plants dominate in shallower water, until finally the marsh is colonized by land plants from the surrounding area as sediment deposition continues and the soil dries out. Reduction in water depth, enabling colonization by marsh species and finally terrestrial species, occurs with input of waterborne and airborne sediments—thus the aquatic phase of the hydrosere is controlled by input of materials from outside the system. Similarly, lakes are typically subject to enrichment of nutrients from surrounding areas, resulting in increased productivity. Extremely high production occurs in culturally eutrophic lakes, which receive nutrient inputs from human activities. In aquatic systems where the influence of allogenic factors such as siltation are apparently minimal, vegetation tends to develop via a series of productive intermediate steps toward an oligotrophic community, that is, one with low productivity, dominated by *Sphagnum* moss (**Fig. 3**). *See* EUTROPHICATION.

Whereas intrinsic factors often result in progressive successional changes, that is, changes leading from simple to more complex communities, external (allogenic) forces may induce retrogressive succession, that is, toward a less mature community. For example, if a grassland is severely overgrazed by cattle, the most palatable species will disappear. As grazing continues, the grass cover is reduced, and in the open areas weeds characteristic of initial stages of succession may become established.

Heterotrophic succession. In the preceding examples of succession, the food web was based on photosynthetic organisms and there was a slow accumulation of organic matter, both living and dead. This is termed autotrophic succession. In some instances, however, addition of organic matter to an ecosystem initiates a succession of decomposer organisms which invade and degrade it. Such a succession is called heterotrophic. In an Illinois pasture, cow dung is degraded by a seral community of some 40 to 60 invertebrate species over a period of 30 days. The newly deposited cow pat is immediately visited by

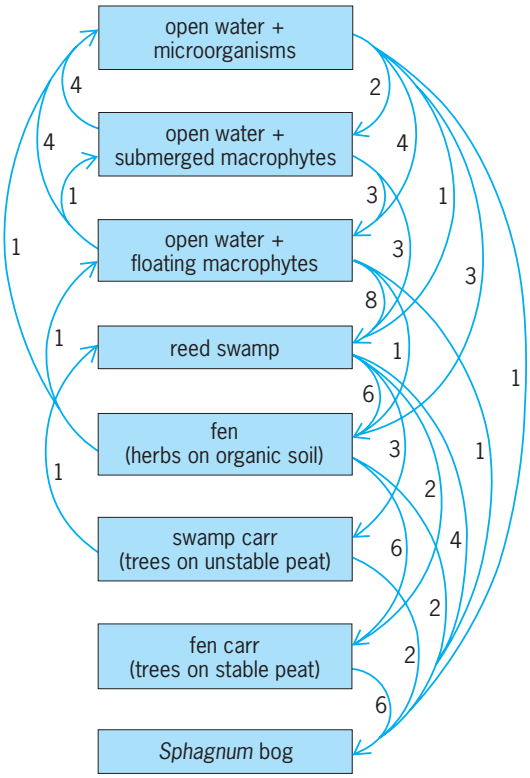

Fig. 3. Transitions between vegetation "stages" free from obvious allogenic influences derived from 20 pollen diagrams for lake sediments and peat. Figures show numbers of observed transitions.

the horn fly (*Haematobia irritans*), which quickly lays eggs and returns to the cow. It is followed by several other dung flies whose larvae are eaten by beetles such as *Sphaeridium scaraboides* which burrow through the dung and lay their eggs. A parasitic wasp (*Xyalophora quinquelinata*) deposits eggs inside maggots of *Sarcophaga* flies. As the dung ages and dries out, it is inhabited by a wider variety of species. Little is known in detail of the importance of the various saprovores in degrading the dung and in modifying the microhabitat, or of their dependence on the activities of saprophytic fungi.

Discharge of organic effluent into a river is detectable downstream by a progression in chemical factors and in the biota. Succession in time is here equivalent to the change in species composition resulting from the decline in organic effluent that corresponds to the distance from the discharge. Marked reduction in dissolved oxygen directly below an outfall results from respiration of microorganisms as they degrade organic matter. Detritivores tolerant of low oxygen concentrations, such as *Tubificidae* and *Chironomidae*, attain high population densities in the bottom sediments. Subsequently, a bloom of algae is typical, utilizing released nitrate and phosphate. As the river flows downstream, the aquatic food web progressively changes from a heterotrophic to an autotrophic basis and productivity declines to its normal level as the "clean water" community returns. *See* BIOLOGICALS; FOOD WEB; PRODUCTIVITY.

Secondary succession. Following the partial or complete destruction of an established community by disturbing events such as fire or clearfelling, and similarly on the cessation of grazing or tillage, a sequence of species invasion and replacement ensues. Compared to the slow initial progress of primary succession in which amelioration of the environment plays an important part, secondary succession is characterized initially by rapid turnover of typically opportunist species which invade relatively congenial habitats.

Piedmont. Abandoned fields in Piedmont area of North Carolina show a rapid sequence of replacement of herbaceous species, apparently related to the life histories of the plants (**Table 3**). Horseweed produces seeds in late summer which germinate immediately so that the plant overwinters as a juvenile, grows rapidly the following year, and dies after seeding in the summer. Aster seeds do not germinate until the following spring, and seedlings grow slowly due to shading by established horseweed plants. In addition, decaying horseweed roots inhibit its own growth and, to a greater extent, that of aster. Horseweed attains dominance in the first year by efficient seed dispersal and rapid establishment. Being a perennial, aster is able to outcompete horseweed in the second year despite the inhibitory effect of the latter. Seedlings of aster are present in abundance in third-year fields but are less drought-resistant than those of broomsedge, which outcompetes aster except in fields with more available water, where aster survives the competition longer. *See* ALLELOPATHY.

Broomsedge seeds are not available to colonize initially because seeds are not produced until the end of the plant's second year and require a period of cold dormancy before germination. Seedling growth is apparently enhanced by decomposition products of the previous colonists. The late establishment of broomsedge in the succession is dependent not on changes brought about by earlier colonists but on the life history of the plant.

TABLE 3. Secondary succession on abandoned fields in the Piedmont area of North Carolina*

Years after last cultivation	Dominant plant	Other common species
0 (autumn)	Crabgrass (*Digitaria sanguinalis*)	
	↓	
1	Horseweed (*Erigeron canadensis*)	Ragweed (*Ambrosia elatior*)
	↓	
2	Aster (*Aster pilosus*)	Ragweed
	↓	
3	Broomsedge (*Andropogon virginicus*)	
	↓	
5–15	Shortleaf pine (*Pinus echinata*)	Loblolly pine
	↓	
50–150	Hardwoods (oaks)	Hickory

*From C. Krebs, *Ecology*, 2d ed., Harper and Row, 1978.

After broomsedge, shortleaf pine invades the herb community. Pine seeds require mineral soil and minimal root competition to become established, and the seedlings are not shade-tolerant. Hence after about 20 years, under a dense pine canopy, reproduction of pines is almost lacking. Accumulation of litter and shade under pines causes the old-field herbs to die out.

Oak seedlings become established after about 20 years, when the depth of litter is adequate to prevent desiccation of acorns. Organic matter in the soil surface layer also increases, improving its water-holding capacity. After about 50 years, several oak species become established and gradually assume dominance as the pines fail to reproduce. Unlike pine, which is capable of germinating on bare soil, oaks and other hardwoods require changes in the soil resulting from pine litter before their seedlings can establish successfully.

While plant species are the main indicators of succession, it is important to note that animal species are also changing over time. In the Piedmont, the changes in bird species as succession proceeds have been well documented. Just a few species, such as meadowlarks and grasshopper sparrows, are found in the initial stages, while more complicated assemblages of species are common in the latter forest stages.

Nova Scotia forest. Following clearfelling in a Nova Scotia forest, the course of secondary succession involves invasion by shrubs (raspberry) followed by understory trees (pincherry, aspen), followed by shade-intolerant species (red maple, paper birch), and finally a shade-tolerant community (hard maple, yellow birch, white ash). Perhaps shade-tolerant species such as red maple, which are of low commercial value, could be inhibited if strip felling were practiced by the forestry industry as an alternative to clearfelling in large blocks. Commercially desirable shade-tolerant species such as white ash would be favored where there was greater local shading of the regenerating community.

Mechanisms of species replacement. Observed changes in the structure and function of seral communities result from natural selection of individuals within their current environment. Three mechanisms by which species may replace each other have been proposed; the relative importance of each apparently depends on the nature of the sere and stage of development.

1. The facilitation hypothesis states that invasion of later species depends on conditions created by earlier colonists. Earlier species modify the environment so as to increase the competitive ability of species which are then able to displace them. Succession thus proceeds because of the effects of species on their environment.

2. The tolerance hypothesis suggests that later successional species tolerate lower levels of resources than earlier occupants and can invade and replace them by reducing resource levels below those tolerated by earlier occupants. Succession proceeds despite the resistance of earlier colonists.

3. The inhibition hypothesis is that all species resist invasion of competitors and are displaced only by death or by damage from factors other than competition. Succession proceeds toward dominance by longer-lived species.

None of these models of succession is solely applicable in all instances; indeed most examples of succession appear to show elements of all three replacement mechanisms. In secondary succession on North Carolina croplands, stimulation of broomsedge growth by decomposition products of previous colonists, and the requirement of oak seedlings for a deep litter layer in which to germinate, exemplify facilitation. The ability of broomsedge to displace aster in competition for water suggest the tolerance mechanism, whereas the inhibition hypothesis is supported by the greater tolerance of horseweed seedlings than aster seedlings to horseweed decomposition products.

Deterministic vs. stochastic succession. Succession has traditionally been regarded as following an orderly progression of changes toward a predictable end point, the climax community, in equilibrium with the prevailing environment. This essentially deterministic view implies that succession will always follow the same course from a given starting point and will pass through a recognizable series of intermediate states (such as in Fig. 2). In contrast, a more recent view of succession is based on adaptations of independent species. It is argued that succession is disorderly and unpredictable, resulting from probabilistic processes such as invasion of propagules and survival of individuals which make up the community. Such a stochastic view reflects the inherent variability observed in nature and the uncertainty of environmental conditions. In particular, it allows for succession to take alternative pathways and end points dependent on the chance outcome of interactions among species and between species and their environment.

Consideration of community properties such as energy flow supports the view of succession as an orderly process. Early in autotrophic succession gross primary productivity (P_g) increases rapidly with community biomass (B), whereas community respiration (R) increases more slowly (**Fig. 4**). As a result, net primary productivity (P_n, where $P_n = P_g - R$) builds up early in succession, and the ratio P_g/B is at its

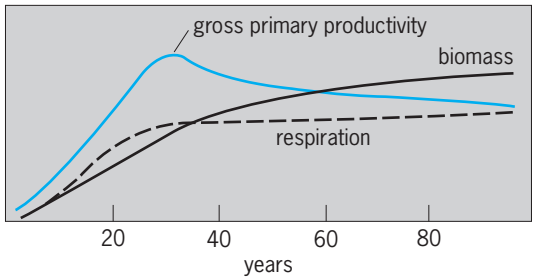

Fig. 4. The energetics of ecosystem development in a forest. The difference between gross primary productivity and respiration is the net primary productivity. (*After E. P. Odum, Ecology, 2d ed., Holt, Reinhart, and Winston, 1975*)

highest in the initial stages. As the community increases in biomass and complexity over time, more complete overall utilization of basic resources such as light limits further increase in primary productivity, whereas R continues to increase because of the increase in tissue to support. Hence zero and the biomass of a mature forest community no longer accumulates. The rate of gross primary productivity typically becomes limited also by the availability of nutrients, now incorporated within the community biomass, and declines to a level sustainable by release from decomposer organisms. Species diversity tends to rise rapidly at first as successive invasions occur, but declines again with the elimination of the pioneer species by the climax community.

Trends in community function, niche specialization, and life history strategy are summarized in **Table 4**. As the community acquires increasing maturity, P_n declines to zero, nutrients become incorporated in biotic pools, broad-niched species are replaced by those with more specific requirements, and the structural organization of the community increases.

Regeneration of an area of subtropical rainforest in Queensland was observed after the vegetation and surface litter were removed with a bulldozer. Because of small environmental differences in the 10 quadrats observed, succession took four directions after the demise of the first ubiquitous colonizing species, resulting in four apparently stable plant associations. This divergence may result merely from small-scale variation in topography within the 65 × 130 ft (20 × 40 m) experimental site and from differing efficiencies of removal of surface litter between the quadrats. Hence the different plant associations detected could be interpreted as divergent products of succession or as phases within a larger-scale vegetation unit.

Stochastic aspects of succession can be represented in the form of models which allow for transitions between a series of different "states." Such models, termed Markovian models, can apply

TABLE 4. Proposed successional trends in ecosystem structural and functional organization, species characteristics, evolutionary factors, and homeostasis*

Ecosystem property	Ecosystem stage: Successional (immature)	Ecosystem stage: Climax (mature)
	Energy Flow	
Gross productivity/respiration (P_g/R)	Autotrophic >1 Heterotrophic <1	
Biomass supported/unit energy flow (B/P_g)	Low	High
Net productivity (P_n)	High	Low
Type of food chains	Linear, grazing	Webs, detritus
	Nutrient Flow	
Mineral cycles	Open	Closed
Flow rate: organism-environment	Rapid	Slow
Role of detritus	Unimportant	Important
	Community Structure	
Total organic matter	Little	Much
Location of chemicals	Habitat pools	Biotic pools
Species richness	Low	High
Species evenness	Low	High
Biochemical diversity	Low	High
Spatial heterogeneity	Low	High
	Species Life History Characteristics	
Niche breadth	Broad	Narrow
Organism size	Small	Large
Life cycles	Short, simple	Long, complex
	Selection Pressure	
Growth form	Rapid growth (r-selection)	Feedback control (K selection)
Production	Quantity	Quality
	Overall Homestasis	
Internal symbiosis	Undeveloped	Developed
Nutrient conservation	Poor	Good
Resistance to perturbations	Poor	Good
Entropy	High	Low
Information	Low	High

*After Odum, 1976.

		from			
		A	B	C	D
to	A	+	•	•	•
	B	++	+	•	•
	C	•	++	+	•
	D	•	•	++	++

(a)

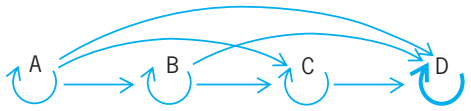

		from			
		A	B	C	D
to	A	+	•	•	•
	B	+	+	•	•
	C	+	+	+	•
	D	+	+	+	++

(b)

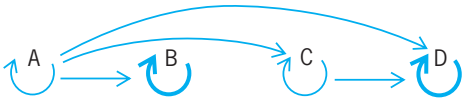

		from			
		A	B	C	D
to	A	+	•	•	•
	B	+	++	•	•
	C	+	•	+	•
	D	+	•	+	++

(c)

Fig. 5. Schematic representation of three postulated mechanisms for species replacement in succession: (*a*) facilitation, (*b*) tolerance, and (*c*) inhibition. Arrows indicate the direction in which the systems tend to move between different states, A–D. Relative probabilities of movement are indicated by thickness of arrows and by symbols in the accompanying transition matrix, where • = close to zero, + = moderate, and ++ = high.

at various levels: plant-by-plant replacement, changes in tree size categories, or transitions between whole communities. A matrix of replacement probabilities defines the direction, pathway, and likelihood of change, and the model can be used to predict the future composition of the community from its initial state. With alternative transition matrices, this simple model could represent a linear progression toward a stable end state, or a cyclical, recursive sequence of communities. The three postulated mechanisms for species replacement discussed above can be illustrated by the topology of alternative Markovian models (**Fig. 5**). The facilitation model of succession is represented as a linear sequence with greater probabilities of progression toward the final state (D) than maintenance of intermediate states. In the tolerance model, later stages may develop from earlier stages depending on the availability of propagules of subsequent stages and their competitive ability. Again, only state D has a high probability of self-replacement. In the inhibition model, there is a high probability that an intermediate state (in this case, B) will persist by strong self-replacement, thereby truncating the normal succession toward state D. A very low probability that B will change to C or D is assumed. Hence, a high degree of realism can be achieved with a simple model system, and alternative predictions of successional changes can be compared with observed data. *See* ECOLOGY; STOCHASTIC PROCESS.

Peter Randerson

Bibliography. D. H. Boucher (ed.), *The Biology of Mutualism: Ecology and Evolution,* 1985, reprint 1989; R. E. Ricklefs and G. L. Miller, *Ecology*, 4th ed., 1999; P. Yodis, *Introduction to Theoretical Ecology*, 1988.

Ecology

The subdiscipline of biology that concentrates on the relationships between organisms and their environments; it is also called environmental biology. Ecology is concerned with patterns of distribution (where organisms occur) and with patterns of abundance (how many organisms occur) in space and time. It seeks to explain the factors that determine the range of environments that organisms occupy and that determine how abundant organisms are within those ranges. It also emphasizes functional interactions between co-occurring organisms. In addition to its character as a unique component of the biological sciences, ecology is both a synthetic and an integrative science since it often draws upon information and concepts in other sciences, ranging from physiology to meteorology, to explain the complex organization of nature.

Environment is all of those factors external to an organism that affect its survival, growth, development, and reproduction. It can be subdivided into physical, or abiotic, factors, and biological, or biotic, factors. The physical components of the environment include all nonbiological constituents, such as temperature, wind, inorganic chemicals, and radiation. The biological components of the environment include the organisms. A somewhat more general term is habitat, which refers in a general way to where an organism occurs and the environmental factors present there. *See* ENVIRONMENT.

A recognition of the unitary coupling of an organism and its environment is fundamental to ecology; in fact, the definitions of organism and environment are not separate. Environment is organism-centered since the environmental properties of a habitat are determined by the requirements of the organisms that occupy that habitat. For example, the amount of inorganic nitrogen dissolved in lake water is of little immediate significance to zooplankton in the lake because they are incapable of utilizing inorganic nitrogen directly. However, because phytoplankton are capable of utilizing inorganic nitrogen directly, it is a component of their environment. Any effect of inorganic nitrogen upon the zooplankton,

then, will occur indirectly through its effect on the abundance of the phytoplankton that the zooplankton feed upon. *See* PHYTOPLANKTON; ZOOPLANKTON.

Just as the environment affects the organism, so the organism affects its environment. Growth of phytoplankton may be nitrogen-limited if the number of individuals has become so great that there is no more nitrogen available in the environment. Zooplankton, not limited by inorganic nitrogen themselves, can promote the growth of additional phytoplankton by consuming some individuals, digesting them, and returning part of the nitrogen to the environment.

Ecology is concerned with the processes involved in the interactions between organisms and their environments, with the mechanisms responsible for those processes, and with the origin, through evolution, of those mechanisms. It is distinguished from such closely related biological subdisciplines as physiology and morphology because it is not intrinsically concerned with the operation of a physiological process or the function of a structure, but with how a process or structure interacts with the environment to influence survival, growth, development, and reproduction.

Scope. There are a wide variety of approaches to ecology because of its broad, comprehensive character. Ecological studies can be characterized by the type of organisms studied, the habitat where studies take place, the level of organization that is of interest, and the methodology used. These are nonexclusive categories, and ecologists combine them in various ways while doing research. Major subdivisions by organism include plant ecology, animal ecology, and microbial ecology. Subdivisions by habitat include terrestrial ecology, the study of organisms on land; limnology, the study of freshwater organisms and habitats; and oceanography, the study of marine organisms and habitats.

The levels of organization studied range from the individual organism to the whole complex of organisms in a large area. Autecology is the study of individuals, population ecology is the study of groups of individuals of a single species or a limited number of species, synecology is the study of communities of several populations, and ecosystem, or simply systems, ecology is the study of communities of organisms and their environments in a specific time and place.

Higher levels of organization include biomes and the biosphere. Biomes are collections of ecosystems with similar organisms and environments and, therefore, similar ecological properties. All of Earth's coniferous forests are elements in the coniferous forest biome. Although united by similar dynamic relationships and structural properties, the biome itself is more abstract than a specific ecosystem. The biosphere is the most inclusive category possible, including all regions of Earth inhabited by living things. It extends from the lower reaches of the atmosphere to the depths of the oceans. *See* BIOME; BIOSPHERE.

The principal methodological approaches to ecology are descriptive, experimental, and theoretical. Much of ecology through the first half of the twentieth century was descriptive, concentrating on describing the variety of populations, communities, and habitats throughout Earth. Experimental ecology, which involves manipulating organisms or their environments to discover the underlying mechanisms governing distribution and abundance, has become of increasing importance since the mid-1960s. Theoretical ecology uses mathematical equations based on assumptions about the properties of organisms and environments to make predictions about patterns of distribution and abundance. It also has become of increasing importance. All of these approaches, however, are evident in the origins of ecology. The science was more descriptive in its early phases and has become more experimental and theoretical in recent decades.

Ecosystem. An ecosystem is the organisms and physical factors in a specific location that are interrelated through the flow of energy and chemicals to form a characteristic trophic structure (**Fig. 1**). The concept of the ecosystem is fundamental to ecology and may be applied at various levels of organization, although it commonly encompasses several different species. The trophic structure of an ecosystem characterizes organisms according to their feeding level and how those feeding relationships of species result in specific patterns of energy flow and chemical cycling. The living mass of a given population or trophic level at any given time is called biomass. A change in mass with time is referred to as net productivity. *See* BIOLOGICAL PRODUCTIVITY; BIOMASS.

Primary producers, largely green photosynthetic plants, utilize the energy of the Sun and inorganic molecules from the environment to synthesize organic molecules. Those organic molecules serve as food for higher trophic levels. Primary consumers, or herbivores, feed on the producers. Secondary consumers, or carnivores, feed on other consumers, while omnivores feed at several trophic levels. Decomposers feed on the dead tissues of other organisms.

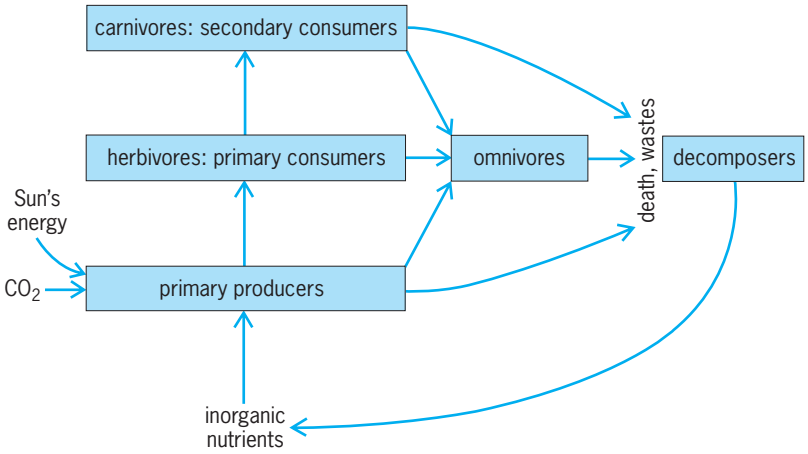

Fig. 1. Generalized trophic of ecosystems. Arrows indicate flow of energy and chemicals.

Each single pathway of energy and chemical flow is referred to as a food chain; the entire collection of food chains in an ecosystem is a food web, or trophic web. Since energy is expended to do work and chemicals are released back into the environment at each step of the trophic web, energy flow in ecosystems is unidirectional, constantly diminishing up the trophic web, while chemicals are recycled to the environment where they can be reutilized in organic syntheses. *See* FOOD WEB.

The ultimate energy source of all ecosystems is the radiant energy of photons from the Sun. Two major types of ecosystems can be distinguished, however, by their proximate energy sources. Autotrophic ecosystems have primary producers as a principal component and sunlight as the major initial energy source. Heterotrophic ecosystems depend upon preformed organic matter that is imported from autotrophic ecosystems elsewhere. The organisms in a stream, for example, may be highly dependent upon organic matter which has eroded from surrounding terrestrial ecosystems.

The ecosystem concept is flexible in application, but the way that it is applied must be clearly specified. An ecosystem study might concentrate on all the organisms and their environments in a specified forest. Alternatively, it might confine itself to the forest floor and soil, concentrating on trophic relationships within that more restricted habitat. *See* ECOSYSTEM.

Autecology

Autecology is the study of particular organisms, typically directed toward determining the traits they possess, and the traits they lack, that allow them to occupy, or be excluded from, certain environments. Early in the twentieth century, autecology placed considerable emphasis upon the morphology and anatomy of organisms, but the field has become diversified in the last few decades into a variety of different approaches. Autecology now also includes physiological ecology, which emphasizes the role of physiological processes, and chemical ecology, which emphasizes the role of biochemical traits. *See* PHYSIOLOGICAL ECOLOGY (PLANT).

One of the earliest generalizations of ecology was the principle of limiting factors, which states that organisms are limited by the factor or combination of factors that are farthest from the requirements of the organism. It was first stated by Liebig based on his studies of soil fertility. He recognized that there usually was a single chemical that would promote yield when it was added to the soil. When that initial limiting factor had been enriched, he found that addition of another nutrient might promote yield even more. The British physiologist Blackman restated the principle in its modern form based on his studies of photosynthesis revealing that the process sometimes could be limited simultaneously by more than one factor.

Though autecology is concerned with individual organisms, that information is often used to contrast the traits of organisms that occupy different environments. Thus there is a strong comparative character to autecological studies; two major habitat comparisons, discussed below, are marine vs. freshwater, and terrestrial vs. aquatic.

Marine vs. freshwater habitats. Life first evolved in the sea; the colonization of freshwater habitats that followed was accompanied by the evolution of traits associated with exploiting the more dilute solutions of freshwater habitats. Seawater is over 3.6% dissolved solids, largely sodium chloride, while freshwater contains 0.01–0.05% dissolved solids. Colonization of freshwater was coupled with the evolution of relatively impervious body surfaces, mechanisms allowing organisms to accumulate ions against a concentration gradient, and the production of large quantities of very dilute urine or other excretory waste. For example, many marine invertebrates are isotonic (equal in osmotic pressure) to seawater so they have no tendency to either lose or gain water. Elasmobranchs have high concentrations of urea in the blood, and although the concentrations of ions is lower than is seawater, the concentrations of urea plus ions is isotonic. Marine teleosts have blood osmotic concentrations well below that of seawater, and they therefore tend to lose water to the sea. They counteract this loss by drinking large quantities of seawater, absorbing it through the gut, and secreting the salt from special cells in the gills. Most freshwater animals are hypotonic (lower osmotic concentration) to the solution that bathes them so they tend to gain water. *See* OSMOREGULATORY MECHANISMS.

Terrestrial vs. aquatic habitats. The colonization of terrestrial habitats was accompanied by the evolution of a large number of traits that allowed organisms to occupy habitats where water was a comparatively rare component of the environment and the atmosphere and soils became principal media surrounding organisms.

Since the body was no longer bathed in fluid, plants and animals both evolved traits allowing them to conserve water and obtain it when it was scarce. Plants developed extensive absorptive surfaces which allowed them to grow into moist regions of the soil and obtain both water and nutrients from the soil solution. Animals often drank water directly and developed outer tissue layers that were extremely impervious to water. The evolution of vascular systems in both animals and plants allowed them to transport water from limited sites of acquisition to the rest of the body. Since the atmosphere is a much less dense and buoyant medium than water, terrestrial plants and animals also evolved traits providing rigid support, such as the skeletons of vertebrates and the strengthened woody stems of vascular plants.

Terrestrial habitats are also characterized by much greater temperature fluctuations than occur in aquatic habitats. Animals therefore evolved physiological traits that allowed them to maintain their body temperatures constant, as well as behavioral traits that involved the ability to move into habitats where favorable temperatures occurred. Physiologically,

evolution culminated in homeothermy in birds and mammals; these organisms expend energy to maintain body temperature within narrow limits. Other animals, called heterotherms, which are incapable of maintaining their body temperature over a wide range of environmental temperatures, often move into localized environmental areas providing appropriate temperatures, and spend unfavorable seasons in quiescent metabolic states. *See* THERMOREGULATION.

Finally, the colonization of terrestrial habitats was associated with many changes in reproductive methods due to the inability of organisms to effectively disperse gametes through the surrounding medium. Such adaptations include the behavioral mating systems of birds and mammals that allow mates to find one another, and the elaborate pollination mechanisms of flowering plants that may involve specific transfer of pollen between individuals of the same species by animal vectors. *See* POLLINATION; REPRODUCTIVE BEHAVIOR.

Physiological and chemical ecology. The postwar development of analytical instrumentation that was much more sensitive to many chemicals than previous analytical techniques led directly to the development of physiological and chemical ecology, which applies that instrumentation to studying organisms in their natural environments. The development of infrared gas analyzers that detect carbon dioxide and water vapor at concentrations that occur in nature led to many studies of photosynthesis, respiration, and transpiration by plants. Those studies have revealed two different types of photosynthesis that occur in plants occupying different types of environments. C_3 plants, so named because the initial stable product of carbon fixation is a three-carbon molecule, are characteristic of almost all aquatic habitats and of terrestrial environments with lower temperatures, less arid climates, and lower solar radiation intensities. C_4 plants, so named because the initial stable product of photosynthesis is a four-carbon molecule, are more common in arid, hotter environments with more intense solar radiation. Those differences are due to the fact that C_3 photosynthesis is intrinsically more efficient but is also more temperature-sensitive and generally more likely to be light-saturated well below the intensity of full sunlight. Thus, C_4 plants are common in tropical and subtropical, arid terrestrial ecosystems, while C_3 plants are more common in most other ecosystems. *See* PHOTOSYNTHESIS.

Chemical ecology studies have been largely devoted to studying rare molecules that have major biological effects. Many animals, for example, produce pheromones that are extremely dilute in the environment but can attract mates, repel predators, or kill pathogens. Many plants also produce organic molecules that are present in their tissues in very small concentrations but act to deter herbivores and pathogens. The evolution of highly toxic molecules by plants sometimes has led to the evolution of specific mechanisms of avoidance or detoxification by a limited number of animals. That type of coupled evolution in which populations change genetically in response to the properties of each other is called coevolution. For example, some animals, such as the monarch butterfly whose larvae feed on milkweeds, have even appropriated aspects of their food plant's chemistry to protect them from predators. Milkweeds produce cardiac glycosides, chemicals that poison the heart muscles of vertebrates. The larvae of the monarch sequester those chemicals in their body, which protects them from such potential predators as birds and small mammals. *See* CHEMICAL ECOLOGY.

Population Ecology

Population ecology is the study of the vital statistics of populations, and the interactions within and between populations that influence survival and reproduction. Population ecologists are concerned with the balances between births and deaths that determine the rate of change of population size through time. When births exceed deaths, of course, population size increases, and when deaths exceed births, population size declines. Vital statistics are often expressed on a per capita basis or, in human populations, on a per-10,000-persons basis. Density, the number of organisms per unit area, is a fundamental property of populations.

One of the principal concerns of population ecologists is identifying and determining the importance of factors that control population densities. Since A. van Leeuwenhoek first calculated the massive potential of flies to reproduce—that is, one pair of flies could produce 746,496 flies after only 3 months of breeding—population ecology has been concerned with the factors that limit the complete realization of this reproductive potential. There are two fundamentally different types of factors limiting population density. Density-independent factors are environmental factors that reduce reproduction or increase death rate independently of the number of organisms in the population. Weather is believed to be a major density-independent factor that often affects populations in a catastrophic way to reduce density. Severely cold winters, deep snowfalls, exceptionally wet growing seasons, and extreme drought occur with a frequency and intensity that are independent of population size. Although these factors can determine population density, they cannot regulate that density since regulation involves maintaining density within certain limits. Density-dependent factors influence the survival or reproduction of individuals in a way that is proportional to density. They can, therefore, regulate population density. Food supply, predators, disease, and such behavioral interactions as territoriality can limit survival and reproduction with increasing intensity as population density increases.

Two vital statistics of populations of principal interest to population ecology are survivorship and reproductive schedules (**Fig. 2**): It is the balance between the probability that an individual will live to reach a certain age, and the average number of

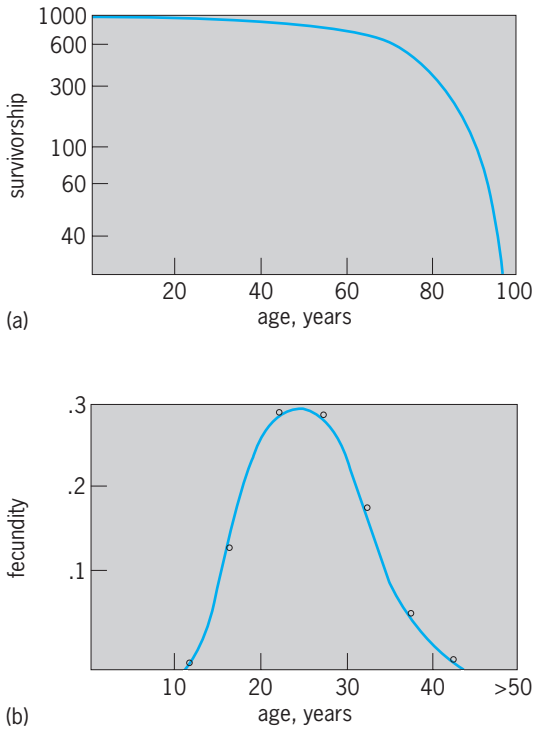

Fig. 2. Demographic statistics of females in the United States. (*a*) Survivorship, where the vertical axis is the number surviving, of 1000 born alive (on a logrithmic scale). (*b*) Reproductive schedule, where the vertical axis is the number of female offspring per female of given age, at 5-year intervals.

offspring produced by individuals of that age that determine population growth rate.

Population growth. One of the earliest quantitative expressions in ecology was a description of density-dependent population growth called the logistic equation. It was first described by P. F. Verhulst in 1838, but it was completely forgotten until its rediscovery by R. Pearl and L. Reed in 1920. Population growth rate is a consequence of population size (N) and per capita birth (b) and death (d) rates, as shown by Eq. (1), where dN/dt is rate of change

$$\frac{dN}{dt} = (b - d)N \qquad (1)$$

of population size. If the symbol r is used to represent the realized per capita reproductive rate, that is, $b - d$, then Eq. (1) can be written in the form of Eq. (2). If r is a constant, a population described

$$\frac{dN}{dt} = rN \qquad (2)$$

by Eq. (2) would grow exponentially with time. The logistic equation recognizes two factors lacking in Eq. (2) that can influence population growth: the maximum per capita reproductive potential of the population, r_m, and the carrying capacity of the environment, K. Maximum reproductive potential would only be realized in the most favorable environment when population density was low enough that individuals did not make simultaneous demands on the same resources. Carrying capacity is a measure of both the amount of resources in the environment and the efficiency with which organisms use those resources. A population is at its carrying capacity when no additional individuals can be supported by an environment. Incorporating those constants into the growth equation above gives Eq. (3), the logistic

$$\frac{dN}{dt} = r_m N\left(1 - \frac{N}{K}\right) \qquad (3)$$

equation. Maximum per capita reproductive potential is determined by the genetic ability to reproduce in an unlimited environment. A population growing according to this equation has a sigmoid, or S-shaped, pattern of population size through time (**Fig. 3**). The maximum population growth rate, which, again, is the product of population size and per capita reproductive rate, is reached when the population is halfway to K.

Dividing Eq. (3) through by N and rearranging gives Eq. (4). Plotting r against N gives a straight line

$$\frac{dN}{Ndt} = r_m - \left(\frac{r_m}{K}\right)N \qquad (4)$$

with an intercept on the y axis that is equal to r_m and an intercept on the x axis that is equal to K.

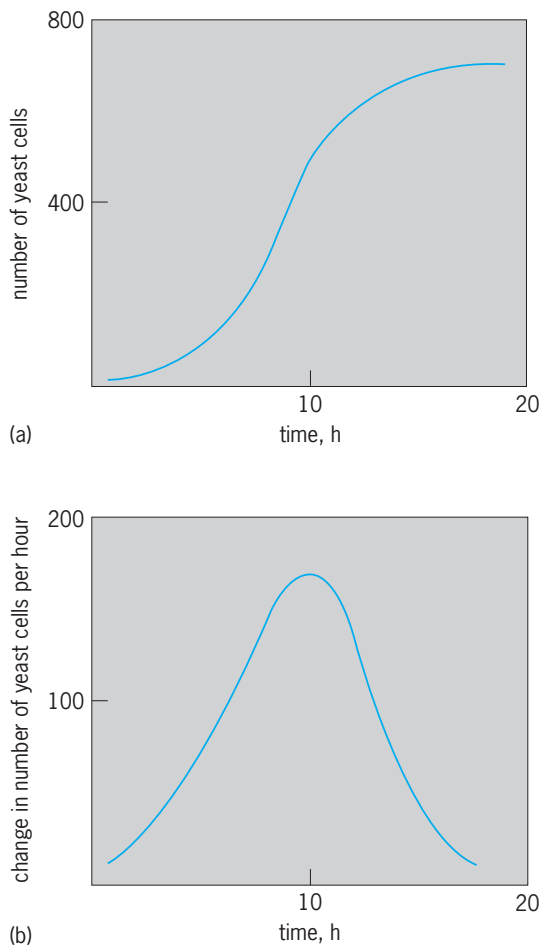

Fig. 3. A yeast population growing according to the logistic equation. (*a*) Population size. (*b*) Growth rate.

An equation as simple as the logistic equation provides merely a yardstick against which the growth of real populations can be compared since it contains a number of assumptions that will not be met in nature. Principal among those assumptions in the logistic equation are that carrying capacity is a constant; that each individual in the population is equivalent enough to all other individuals to be represented in the average population statistic, r_m; that all individuals have an equal probability of mating; and that there are no time lags so that birth and death rates are adjusted immediately for each increment of N as population size increases to K. Much of the study of the ecology of real populations involves understanding how those assumptions are violated.

Competition is inherent in the logistic equation, and in population growth in nature. Competition occurs when the resources in the environment are in short supply relative to the demands that individuals make on those resources. In the logistic equation, that competition is expressed in the decline of r as N increases. This type of competition between members of the same species population is called intraspecific competition.

Population interactions. Competition also is a principal type of interaction between the members of different species, when it is termed interspecific competition. Another important type of interaction between populations is predation, when the individuals of one population feed on the individuals of another population. Both of these interactions were described mathematically in the 1920s by two ecologists working independently, A. Lotka and V. Volterra. The Lotka-Volterra competition equations are Eq. (5) and the exact equivalent, Eq. (6), where 1 and 2 refer

$$\frac{dN_1}{dt} = r_{m1}N_1\frac{K_1 - N_1 - \alpha_{12}N_2}{K_1} \tag{5}$$

$$\frac{dN_2}{dt} = r_{m2}N_2\frac{K_2 - N_2 - \alpha_{21}N_1}{K_2} \tag{6}$$

to the different species, with α_{12} being the competitive effect of an individual of species 2 on species 1, and α_{21} being the competitive effect of an individual of species 1 on species 2. These Lotka-Volterra competition equations recognize that each species depresses the growth rate and carrying capacity of the other. Depending upon the respective competition coefficients, carrying capacities, and maximum per capita reproductive rates, there are four potential outcomes of competition: (1) species 1 may go extinct; (2) species 2 may go extinct; (3) the species may have an unstable equilibrium in which they coexist for a while until chance events allow one species to increase a bit above the equilibrium and the other species then goes extinct; or (4) there may be a stable equilibrium in which both species coexist at some population size below the carrying capacity for either species alone.

The Lotka-Volterra predation equations describe population ecology when members of a predator population, P, feed on members of a prey population, H. The prey equation is Eq. (7), where k_1 is

$$\frac{dH}{dt} = r_m H\left(1 - \frac{H}{K}\right) - k_1 PH \tag{7}$$

a constant that is the predation rate per capita of predator. The equation for the predator is Eq. (8),

$$\frac{dP}{dt} = k_2 PH - k_3 P \tag{8}$$

wherek_2 is the efficiency with which predators convert their food into more offspring and k_3 is the per capita predator death rate. Unlike the logistic and Lotka-Volterra competition equations, which predict that populations will reach some final constant size, the Lotka-Volterra predator-prey equations predict that predator and prey populations will tend to oscillate through time. At low densities of both predator and prey, the values of H/K and k_1PH in the prey population equation will both approach zero, and the prey will grow according to the unmodified logistic, approaching the maximum per capita reproductive rate. As the prey population increases, the predator population also can increase, and there is a tendency for predator density to overshoot prey density, leading to oscillations through time. Such oscillations are, in fact, often observed in nature, with a classic example being Canadian populations of snowshoe hare, the prey, and predators, such as lynx (**Fig. 4**). Whether such oscillations in real populations are due to the instability of predator-prey relations inherent in the Lotka-Volterra predation equations, or to the fluctuations of other factors such as weather or food supply available to the hares, has not yet been proven. *See* MATHEMATICAL ECOLOGY; POPULATION ECOLOGY.

Parasitism and disease are additional types of population interactions involving detrimental effects on one participant and beneficial effects on the other participant. A parasite is an organism that grows, feeds on, and is sheltered by another organism, the host. Disease is an abnormal condition of an organism, often due to infection, that impairs normal physiological activity. Parasitism and disease often are described mathematically by equations similar to the Lotka-Volterra predation equations, although they

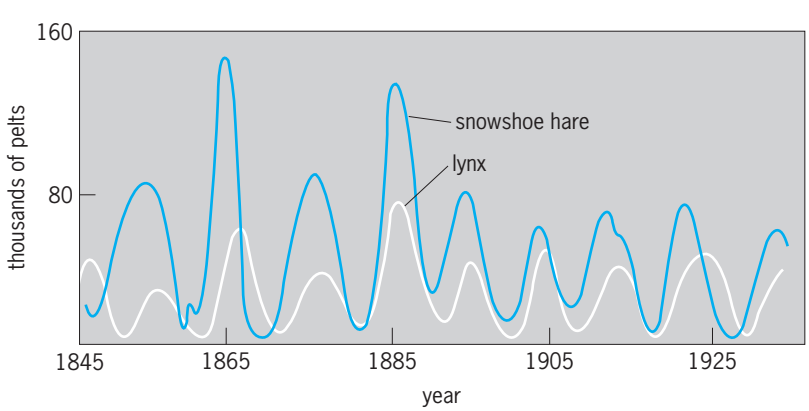

Fig. 4. Oscillations of predator and prey populations. Data are from pelt records of lynx and snowshoe hare of the Hudson's Bay Company, in Canada.

are commonly modified to account for the dispersal of parasites or pathogens between different hosts. These population interactions have been of growing interest in ecology, although they have been the subject of the entire field of epidemiology, an important aspect of medicine. The importance of parasitism and disease in natural ecosystems is evident from the extinctions of two important tree species, the American chestnut (*Castanea dentata*) and American elm (*Ulmus americana*), from the forests of eastern North America during this century. Those extinctions were caused, respectively, by chestnut blight and Dutch elm disease, both due to pathogens introduced into North America from Europe. *See* EPIDEMIOLOGY.

Symbiosis is a type of interaction between populations in which a participant is promoted by the interaction. There are two types of symbiosis: commensalism, which occurs when one participant is promoted by the interaction and the other is unaffected, and mutualism, which occurs when both participants are promoted by the interaction. Mutualism often involves organisms that are very different and have coevolved such a strong mutual interdependence that neither can exist in natural environments in the absence of the other. Symbiosis commonly is described mathematically by variations of the Lotka-Volterra competition equations in which the sign preceding the α becomes positive for one (commensalism) or both (symbiosis) populations. This transforms the competition coefficients into mutualism coefficients.

Mutualism is an extremely important type of ecological interaction since it often allows the symbionts to perform functions that have proven impossible for any single organism to accomplish with the same efficiency. Lichens are symbiotic associations between fungi and algae. They are capable of colonizing some of the most inhospitable environments on Earth, such as the bare faces of boulders and extremes of drought, cold, and heat. The fungal and algal species are so closely integrated that lichens have been classified by the binomial systems that are applied to other true species.

An example of the functional importance of symbiosis to ecology is the digestion of cellulose, which is accomplished largely by symbiotic associations. Cellulose is the most abundant organic molecule on Earth, but this vast energy source is largely inaccessible to living organisms except to symbionts. Among the most important symbiotic associations capable of digesting cellulose are ruminant animals and their gut flora, and termites and their gut protozoa. In the absence of these symbiotic associations, Earth would be covered with vast residues of undecomposed cellulose.

Synecology

Synecology is the study of groups of populations co-occurring in time and space; it is also called community ecology. Synecology is concerned with factors controlling the species composition of communities and why species composition changes in different environments and with time in the same habitat. Species composition of a community may include density, biomass, or productivity of individuals of different species.

The study of communities is one of the oldest branches of ecology, arising strongly in the studies of naturalists in the sixteenth to nineteenth century, and one of the most descriptive. Only since the mid-1960s have ecologists begun to turn significantly to experimental studies that manipulate either the environment or the species composition of communities. Those experiments are designed to understand mechanisms responsible for the species compositions of communities and the interactions that occur between different species in the community. *See* ECOLOGICAL COMMUNITIES.

Community organization. A principal concern of community ecology is the life forms of organisms, the taxonomic composition of co-occurring species groups, and the spatial arrangement of species. Plant life forms are relatively easy to characterize, and communities in widely separated geographic areas but similar environments commonly have a similar life-form composition. Most of the species in tropical rainforests, for example, are trees and shrubs, while most of the species in grasslands are small shrubs, and perennials with buds belowground. Although the species compositions of communities in different geographic areas may be quite different, the life forms are often quite similar in similar environments.

Another aspect of community organization that is a concern of synecology is spatial patterns, or dispersion. Ecologists typically distinguish between dispersal, the process of movement, and dispersion, the arrangement of individuals in space. The three types of dispersion patterns are random, regular, and clumped (aggregated). In a random distribution, the probability of an individual occupying a given location in a community is not influenced by the presence of another individual at a nearby location, which implies that no important factor, such as environment or parental proximity, has influenced the dispersion of individuals. In a regular distribution, the probability of an individual occupying a location is reduced if another individual occupies a nearby location; this distribution may be due to competition for an evenly spread environmental resource. In a clumped distribution, which may be due to proximity to seed source, or environmental heterogeneity, the probability of an individual occupying a location increases if another individual occupies a nearby location. *See* POPULATION DISPERSAL; POPULATION DISPERSION.

Temporal distributions are also a concern of synecology. By substituting the phrase "at a given time" for the phrase "at a given location" in the above definitions, temporal distributions are judged by the same criteria. Spatial and temporal patterns of individuals are important aspects of community ecology because they can provide insight into the underlying distributions of limiting factors in the environments and how species are related to those factors. Species

with distributions in which individuals are clumped in space or time often utilize resources that also are clumped in space or time. Limited dispersal also can lead to a clumped distribution due to the clustering of offspring around their parents. Regular distributions can often result from competition between individuals for a resource that is evenly distributed in space or time. Desert plants, for example, frequently have a regular distribution pattern because they compete for soil moisture, and each individual must utilize the moisture available in a certain minimum soil volume if it is to survive.

Niche. The niche concept refers to the environmental factors that control a species, and its distribution in relation to those factors. A species' niche is a consequence of the genetic properties of the individuals of that species, the environmental factors that affect it, and interactions with other species. G. E. Hutchinson described the niche as an *n*-dimensional hypervolume occupied by a population, where *n* is the number of environmental factors affecting it. He also distinguished between a fundamental niche, which is the genetic potential of the species, and a realized niche, which is the range of environmental factors occupied in the presence of other organisms.

Species are often distributed along environmental gradients as a series of bell-shaped curves (**Fig. 5**). Those curves represent the realized niches of species. Removal of other species, a type of experiment that is becoming increasingly common in synecology, would be required to reveal a species' fundamental niche. A species that is narrowly distributed has a narrow niche and is often referred to as an ecological specialist, while a widely distributed species has a broad niche and is often referred to as an ecological generalist. It is often inferred that intense interspecific competition will lead to specialization, restricting a species' realized niche to only that portion of the fundamental niche where it is a strong competitor. Conversely, intense intraspecific competition will tend to lead to a broader niche that reduces the effects of different members of the same species upon each other. However, there are as yet insufficient experimental studies to strongly support either of these inferences.

Species diversity. Species diversity is a property of a community that encompasses both the number of species present (species richness) and the equality of relative abundances of the species present (equitability of evenness). One of the principal concerns of synecology since its origin has been why communities differ so dramatically in the number of species present and the similarities of their relative abundances. Naturalists attached to exploratory expeditions during the historical colonial period were particularly overwhelmed by the variety of plant and animal life in the wet tropics. Tropical forests seemed to consist of an almost countless number of plant and animal species compared with communities in the temperate climates with which the naturalists were familiar. Indeed, there is a general gradient of increasing diversity from the poles to the Equator.

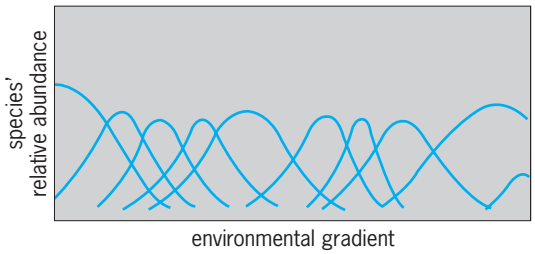

Fig. 5. A continuum of species distributions along an environmental gradient, where each line represents a different species. Note that peaks of species abundance do not overlap, and that breadth of curves differs between species.

Various explanations for the latitudinal differences in biotic diversity have been proposed, but none is wholly satisfactory and none has gained prominence. Chief among those explanations are: (1) Nontropical habitats have been subject to severe climatic fluctuations, including glaciation, that lead to repeated species extinctions. (2) The moderate climates of tropical locations allow species to become more specialized for narrow ranges of climate and other environmental factors. (3) Biological interactions, such as competition and predation, are stronger in tropical locations, leading to a proliferation of different adaptive types. (4) Productivity is higher in tropical locations, allowing more species to coexist due to greater available food. (5) Spatial variation of environments is greater in the tropics, leading to a proliferation of specialists in different areas. (6) There is greater environmental complexity in the tropics, partly because of more species, leading to greater specialization among species in a sort of positive feedback. *See* BIOGEOGRAPHY.

Community: reality vs. abstraction. One of the most contentious issues throughout the history of ecology has been the degree to which the species that co-occur in space and time represent an organized, meaningful biological unit and to what extent they are merely a fortuitous collection of individuals. Two of the principal ecologists in American ecology in the early decades of the twentieth century argued these points. F. Clements referred to communities as being like an organism, while H. A. Gleason said that communites were merely a happenstance collection of species with similar ecological tolerances. Clements's view prevailed until the middle of the twentieth century, and then Gleason's view gained ascendency, largely because of quantitative studies of community composition revealing that species distributions along environmental gradients were commonly the type of bell-shaped curves in Fig. 5. Those bell-shaped curves suggested that each species population had its own, individualistic response to the environment, and that the species that occurred at a given point on the gradient did so fortuitously.

Still, there is no gainsaying that similar communities occur in similar habitats. Although the species compositions of communities may vary considerably

in different geographic areas, the occurrence of communities with very similar life forms in geographically separated locations with similar environments suggests that a community is not merely a random collection of species. Many experiments in marine intertidal communities indicate that the removal of one species has a cascading effect that alters the abundances of many other species, even leading to the extinction of some species and invasion of the community by other species that were previously absent. These results indicate that the species in a community are interconnected through the functional processes of competition, predation, and symbiosis in ways producing a community that is not just a random collection of species.

Succession. Succession is the process of change in community composition and environmental properties over time on a site. Primary succession occurs in newly formed habitats, such as glacial moraines, river levees, or volcanic ash. Secondary succession occurs in habitats where the previous community has been destroyed or severely disturbed, such as following forest fire, abandonment of agricultural fields, or epidemic disease or pest attack.

A general consequence of succession is an amelioration of physical factors and a reduction of their importance as controlling factors, and an increase in the complexity and importance of biological factors. In newly formed ash fields or newly abandoned fields, sunlight is intense, nutrients are often in poor supply, winds are strong, and evaporative stress is severe. As such sites are colonized by plants, all of these physical factors tend to be modified and trophic webs develop that depend upon the plants. Species diversity increases as more species colonize, and increasingly complex trophic patterns develop. There often is an increase in the stature of the vegetation, which reduces sunlight, wind, and evaporative stress below the canopy. That creates new conditions that allow species to invade that were incapable of withstanding the physically harsh conditions early in succession.

Early in succession, there is directional change in both environmental factors and species composition of the community. Ultimately a relatively constant environment is reached and species composition no longer changes in a directional fashion but, instead, fluctuates about some mean, or average, community composition. That stage is called the climax.

As in the nature of the community itself, there is disagreement among ecologists about the exact nature of succession, and of the climax. Nevertheless, there is a general tendency for communities to become more complex and for biological factors to increase in importance and the effects of physical factors to be reduced as time passes. *See* ECOLOGICAL SUCCESSION.

Systems Ecology

Systems ecology is the study of dynamic relationships between the units of the ecosystem, particularly those relationships that influence rates of energy flow and chemical cycling. The ecosystem concept, as pointed out above, can be applied to many different levels of organization and varying degrees of complexity. Ecosystem science, or systems ecology, can be broadly divided into approaches concerned with relationships between trophic levels and those that employ with specific details of population dynamics of community organization within trophic levels. In both approaches the goal is to develop mechanistic explanations for the observed behavior of the system as a whole. Often this goal is accomplished through simulation modeling.

Simulation modeling involves the use of mathematical equations to characterize the state and dynamics of a system. Mathematical approaches to population dynamics, as seen earlier in this article, commonly use simplified equations, such as the logistic, to model the dynamics of one or a few populations. Simulation modeling differs principally in using a set of coupled equations to characterize the processes occurring in ecosystems. Those models often have much more detail than the simplified models of population dynamics, and require numerical solution using digital computers. A simulation model typically begins with a conceptual model or diagram consisting of boxes that represent components of system state and connecting arrows representing processes that interconnect the components. Simulation modeling uses rate equations to describe the processes and evaluate the mechanisms that are important to understanding both the dynamics and functional properties of ecosystems. Simulation models have as goals both isolating the general features of a broad range of similar ecosystems and providing sufficient detail to accurately predict how a specified system will respond to perturbations.

A system can be defined as a group of interacting and interdependent elements that form a collective entity. An ecosystem, therefore, is a collection of environmental factors and organisms that interact through specific, dynamic relationships defined by the food web of which they are members. Systems may be open or closed. A closed system is completely self-contained and does not interact with any other system; an example would be molecules of a gas enclosed in a container that was a perfect insulator. Those molecules would function independently of all other things. An open system interacts with other systems. All ecosystems are open systems. In particular, they are open to energy, whose ultimate source is the Sun. Proximally, however, energy can be imported into an ecosystem from some other system where the energy of photons from the Sun is converted into organic energy by photosynthesis.

Energy flow. Energy flow in ecosystems is approached in a way closely related to the laws of thermodynamics, that branch of physics that deals with the relationships among different forms of energy. The first law of thermodynamics states that the energy input to a system either is stored or is used to do work. It is generally given as Eq. (9), where Q

$$Q = \Delta E + W \tag{9}$$

is energy input, ΔE is a change in the energy content of the recipient system, and W is work done by that system. Ecologists restate the law as Eq. (10),

$$P_g = P_n + R \tag{10}$$

where P_g is gross productivity, P_n is net productivity, and R is respiration. The change in energy content, P_n, is therefore equal to the energy input minus the work done. That work has two components: one is maintenance respiration, the cost of maintaining the structure of the system at the time of energy input, and the other is growth respiration, the cost of synthesizing new biomass.

The second law of thermodynamics also is important to energy balance in ecosystems. That law states that the entropy of isolated systems always tends to increase. Entropy (S) is a measure of the randomness, disorder, or lack of organization of a system. In energy terms it is measured as the heat capacity of the system (joules $\cdot$ degree^{-1} $\cdot$ mole^{-1}) at a particular temperature (T). In isolated systems, entropy always increases. Because ecosystems are open systems, they are able to maintain their organization and avoid decaying to a state of maximum entropy and zero free energy. The maintenance component of respiration is the cost of maintaining present structure, that is, replacing the free energy degraded to heat. The growth component of respiration is the cost, that is, the free energy intake necessary to build structure and organization. Those costs are expressed both as an increase in the entropy (thus an increase in heat content) of the environment to which organisms in an ecosystem are coupled as well as the loss of heat produced by "friction" representing the degree of thermodynamic inefficiency of the metabolic energy transformations. *See* OPEN-SYSTEMS THERMODYNAMICS (BIOLOGY); THERMODYNAMIC PRINCIPLES.

A fundamental generalization of systems ecology is that net productivity diminishes about one order of magnitude at each level in the trophic web. If primary P_n averaged 100 g/m^2 $\cdot$ yr, secondary P_n at the herbivore level would be about 10 g/m^2 $\cdot$ yr, and at the carnivore level would be about 1 g/m^2 $\cdot$ yr. This so-called energy pyramid has two causes. First, the above-mentioned maintenance costs of respiration at each trophic level diminish the energy available to higher levels. Second, the ecological harvest of energy by any trophic level is not fully efficient. Therefore, the energy flow must diminish at each successive trophic level in an ecosystem.

In most ecosystems, certainly, the biomass of primary producers is greater than that of the herbivores which, in turn, is greater than carnivore biomass. Sometimes, however, particularly in aquatic habitats, biomass increases up the trophic web. That is possible only when lower trophic levels consist of organisms with very high rates of net productivity per unit of biomass. The ratio of energy flow to biomass is referred to as turnover time. If turnover time is short, as it is for small organisms with rapid life cycles, it is possible for the biomass pyramid to be inverted; that is probably most common in the open ocean where large, long-lived organisms are at upper trophic levels and small, short-lived organisms are at lower trophic levels. *See* ECOLOGICAL ENERGETICS.

Chemical cycling. Another fundamental concern of systems ecology is the pattern of chemical flow that is coupled to the process of energy flow. Each time that energy is used to do work, some chemicals are released back into the environment. Death also releases chemicals back into the environment.

There are two different types of trophic webs in ecosystems. The grazing food web is based on the consumption of the tissues of living organisms. The detritus food web is based on the consumption of dead organic material, called detritus in aquatic systems and sometimes referred to as litter in terrestrial ecosystems. Chemicals are recycled to the environment from the grazing food web each time that work is done, and as the excretory wastes of living organisms. Chemicals are recycled to the environment from the detritus food web as organisms utilize excretory wastes and the dead tissues of organisms.

Decomposition is the process of degrading the energy content of dead tissues and simultaneously releasing chemicals back into the environment. When those chemicals are released in inorganic forms, the process is called mineralization. Decomposition and mineralization typically involve many steps and take place in a trophic web fully as complex as those of the grazing food web. There is a size-dependent hierarchy as particles are broken into progressively smaller particles; that subdivision increases surface-volume ratios, exposing progressively more of the substance to decay. There also typically is a chemical hierarchy. Fats, proteins, and simple carbohydrates are the most easily utilized organic chemicals, and they generally are rapidly attacked. Wood, chitin, and bones are only slowly decomposed. More resistant substances accumulate in terrestrial ecosystems as soil humus and in aquatic ecosystems as sediments, or ooze, beneath the water. In geological time, those sediments were transformed into the hydrocarbon deposits that serve as sources of coal, oil, and gas for modern industrial society.

Microorganisms are exceptionally important in the processes of decomposition and mineralization. The final steps of mineralization are almost invariably accomplished by microorganisms. For example, the cycling of nitrogen in ecosystems involves microorganisms at several critical steps (**Fig. 6**). The only significant pathway introducing nitrogen into trophic webs is the process of nitrogen fixation in which gaseous, molecular nitrogen is converted into organic nitrogenous compounds by nitrogen-fixing bacteria and blue-green algae. Some of these bacteria and algae are free-living, and some participate in symbiotic associations with other organisms. The organic molecules are degraded in the detritus food web by ammonifying bacteria that release ammonium ion or ammonia. Nitrite bacteria convert the ammonia into nitrite, and nitrate bacteria oxidize the nitrogen further into nitrate; those conversions are referred to collectively as nitrification. Both

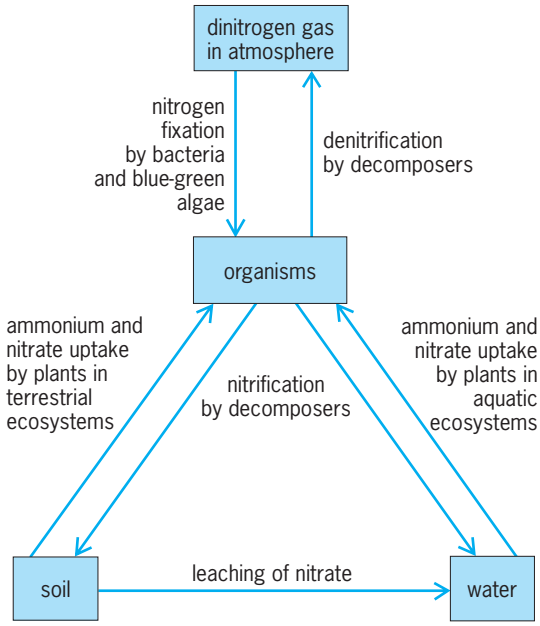

Fig. 6. Generalized nitrogen cycle. Arrows represent flow of nitrogen.

nitrate and ammonium ion can reenter the food web via uptake by plants. Denitrifying bacteria convert the nitrate into nitrogen gas, a process called denitrification that releases gaseous nitrogen back into the atmosphere. *See* BIOGEOCHEMISTRY; MATHEMATICAL ECOLOGY.

Applied Ecology

Applied ecology is that branch of ecology dealing with practical problems of immediate social importance. Ecology is often confused in the minds of laypersons with the environmental movement, a social interest group concerned with environmental degradation and problems of resource supply to human societies. Applied ecology makes a fundamental contribution to those social and political concerns by identifying environmental problems, gaging their significance, and suggesting potential solutions. Among the environmental issues to which ecology has made an important contribution have been problems of population growth and resource supply, acid rain, eutrophication, consequences of pollution, biological control of crop pests, range management, forestry, and the ecological consequences of nuclear war. From T. R. Malthus's essay on human population growth to recent scientific evaluations of nuclear war, ecology has always been concerned with problems that are important to the affairs of humans. *See* ECOLOGY, APPLIED. Samuel J. McNaughton

Bibliography. B. Glaeser (ed.), *The Green Revolution Revisited*, 1987; H. Howe and L. Westley, *Ecological Relationships of Plants and Animals*, 1988, reprint 1990; J. Keating, *Interdependence in the Natural World*, 1987; R. P. McIntosh, *The Background of Ecology: Concept and Theory*, 1985, reprint 1986; M. B. Rambler et al. (eds.), *Global Ecology: Towards a Science of the Biosphere*, 1988.

Ecology, applied

The application of ecological principles to the solution of human problems and the maintenance of a quality life. It is assumed that humans are an integral part of ecological systems and that they depend upon healthy, well-operating, and productive systems for their continued well-being. For these reasons, applied ecology is based on a knowledge of ecosystems and populations, and the principles and techniques of ecology are used to interpret and solve specific environmental problems and to plan new management systems in the biosphere. Although a variety of management fields, such as forestry, agriculture, wildlife management, environmental engineering, and environmental design, are concerned with specific parts of the environment, applied ecology is unique in taking a view of whole systems, and attempting to account for all inputs to and outputs from the systems—and all impacts. In the past, applied ecology has been considered as being synonymous with the above applied sciences. *See* SYSTEMS ECOLOGY.

Ecosystem ecology. Ecological systems, or ecosystems, are complexes of plants, animals, microorganisms, and humans, together with their environment. Environment includes all those factors, physical, chemical, biological, sociocultural, which affect the ecosystem. The complex of life and environment exists as an interacting system and is unique for each part of the Earth. The unique geological features, soils, climate, and availability of plants, animals, and microorganisms create a variety of different types of ecosystems, such as forests, fields, lakes, rivers, and oceans. Each ecological system may be composed of hundreds to thousands of biological species which interact with each other through the transfer of energy, chemical materials, and information. The interconnecting networks which characterize ecosystems are often called food webs (**Fig. 1**). It is obvious from this structural feature of interaction that a disturbance to one population within an ecosystem could potentially affect many other populations. From another point of view, ecosystems are composed of chemical elements, arranged in a variety of organic complexes. There is a continual process of loss and uptake of chemicals to and from the environment as populations are born, grow and die, and are decomposed. Ecosystems operate on energy derived from photosynthesis (called primary production) and from other energy exchanges. The functional attributes of ecosystems, such as productivity, energy flow, and cycling of chemical elements, depend upon the biological species in the ecosystem and the limiting conditions of the environment. *See* BIOLOGICAL PRODUCTIVITY; ECOLOGICAL ENERGETICS; FOOD WEB.

Ecological systems develop in accord with the regional environment. Although these systems have evolved to resist the normal expected perturbations encountered in the environment, unusual disturbances and catastrophic events can upset and even destroy the system. In this case, recovery can occur

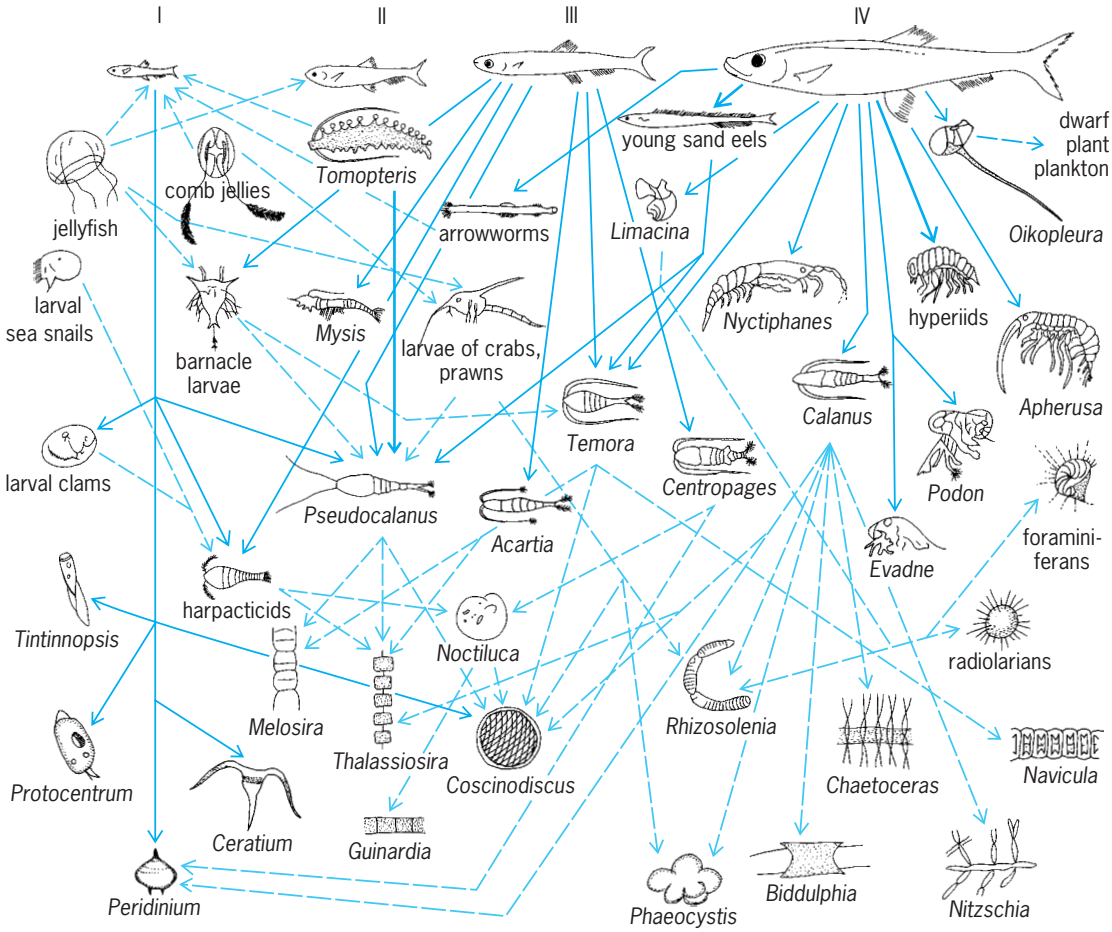

Fig. 1. Food relations of the herring at different stages in its life. Sizes of ecologists (I) 0.6 to 1.3 cm, (II) 1.3 to 4.5 cm, (III) 4.5 to 12.5 cm, and (IV) over 12.5 cm. Solid lines indicate food eaten directly by herring. 1 cm = 0.4 in.

after the disturbance stops. Recovery is termed ecological succession since it comprises a sequence of communities which succeed each other until a dynamic steady state is reestablished. *See* ECOLOGICAL SUCCESSION.

Populations in these ecosystems fill a variety of structural and functional roles within the system. Often, groups of populations coevolve, so that they form a more or less isolated subunit. For example, the pollinators of a plant species, and their predators and parasites, form such a guild. Populations continually adapt and develop through natural selection, expanding to the limit of their resources. Population growth is, therefore, due to an increase in resources or a relaxation of limiting factors. *See* ECOSYSTEM; POPULATION ECOLOGY.

Ecosystem management theory. The objective of applied ecology management is to maintain the system while altering its inputs or outputs. Often, ecology management is designed to maximize a particular output or the quantity of a specific component. Since outputs and inputs are related (**Fig. 2**), maximization of an output may not be desirable; rather, the management objective may be the optimum level. Optimization of systems can be accomplished through the use of systems ecology methods which consider all parts of the system rather than a specific set of components. In this way, a series of strategies or scenarios can be evaluated, and the strategy producing the largest gain for the least cost can be chosen for implementation.

The applied ecology management approach has been partially implemented through the U.S. National Environmental Policy Act. This act requires that an environmental-impact analysis be carried out by an interdisciplinary team of specialists

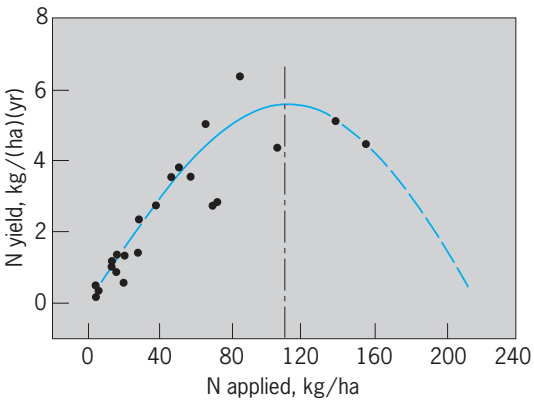

Fig. 2. Relationship between N yield and N consumption (1964–1969).

representing the subjects necessary for an ecosystem analysis. This team is required to evaluate a project on the basis of its environmental, social, and cultural features. One alternative that must be considered is that of no alteration of the system.

A variety of general environmental problems within the scope of applied ecology relate to the major components of the Earth: the atmosphere, water, land, and the biota. The ecological principles used in applied ecology are discussed elsewhere; a sequence of environmental problems of special importance to applied ecology is discussed below.

Atmospheric problems. The atmosphere is one of the most important components of the environment to consider from the viewpoint of applied ecology since it connects all portions of the Earth into one ecosystem. The atmosphere is composed of a variety of gases, of which oxygen and nitrogen make up the largest percentage. It is not uniform in its depth or its composition, but is divided into several layers or zones which differ in density and composition. Although most interaction with humans occurs in the zone nearest Earth, the most distant parts of the atmosphere are also important since they affect the heat balance of the Earth and the quality of radiant energy striking the surface. Disturbances to these portions of the atmosphere could affect the entire biosphere. *See* ATMOSPHERE.

The composition of the atmosphere varies according to location. The qualities of minor gases such as carbon dioxide, the amounts of various metallic elements, and the quantity of water vapor and dust all may differ, depending on the relative distance from land or sea. But, in addition, the atmospheric composition may change in time. For example, over the history of the planet, the percentage composition of oxygen has changed from a very oxgen-poor environment to the present atmosphere, with 20.95% oxygen by volume. *See* ATMOSPHERE, EVOLUTION OF; ATMOSPHERIC CHEMISTRY.

Human activities may introduce a variety of pollutants into the atmosphere. The principal pollutants are carbon dioxide, sulfur compounds, hydrocarbons, nitrogen oxides, solid particles (particulates), and heat. The amounts of pollutants that are produced may be quite large, especially in local areas, and have increased in amount as industrialization has become more widespread. Industrial and domestic activities have also been estimated to put 2.18×10^8 tons (1.96×10^8 metric tons) of sulfur into the atmosphere per year. In most cases, these pollutants have increased during the recent past, and in many areas have become a serious problem. *See* AIR POLLUTION.

Atmospheric problems can have a variety of impacts on humans. Numerous observers have attributed climatic change to atmospheric alteration, since any change in the gaseous envelope of the Earth could alter the heat balance and the climate. The Earth's climate is not constant, and it is difficult to establish an exact correlation between pollution and variation in temperature or solar radiation at the Earth's surface. The pollutants most likely to have an effect on climate are carbon dioxide and solid particles. Pollution may also affect the chemical balance of regions of the Earth. These effects may be extremely complicated. For example, it has been reported that sulfur oxides produced in the industrial districts of northern Europe have moved north in the atmosphere over Scandinavia. The sulfur oxides react with water to form sulfuric acid, which rains out in precipitation. The acid rain changes the acidity of the soil in this region and may depress the activity of blue-green algae, which fix nitrogen from the atmosphere. Acid rain appears to cause reduction in the growth of trees. Other pollutants may act in a similar complicated fashion through the network of interactions in the Earth ecosystem. *See* ACID RAIN.

Finally, atmospheric pollution has direct effects on plants and animals and human activities. Pollutants, like other materials, can act as limiting influences on the growth, reproduction, and survival of plants and animals. A variety of plants, such as lichens and mosses, that are extremely sensitive to pollution can be used to indicate the degree of atmospheric deterioration. In some severe cases, all vegetation and animal life may be destroyed in the vicinity of the polluting industry. Gases and solids are taken into the lungs of humans, and cause disease or discomfort. In cities, such pollutants as asbestos and lead are exceedingly dangerous to the population. And finally, the impact of pollutants on buildings, clothes, artwork, and machines is costly.

Control of atmospheric pollution requires interception of pollutants at the point of discharge. Industrial control can be achieved by the use of special filters, precipitators, and other devices. Control of pollution for automobiles also may involve special equipment, as well as redesign of engines and fuel. Reduction in dust and similar general sources of air pollution may demand a change in the operation producing the problem.

Water problems. The aquatic environment is of equal importance to applied ecology since most of the Earth's surface is covered by the oceans, and the land is connected to the oceans by streams and rivers. Thus, like the atmosphere, the waters are a connection between distant parts of the biosphere and can carry a disturbance from its origin to another region.

The composition of waters varies widely, and it is essential in evaluating aquatic health to establish the base-line conditions which are stable under the normal or undisturbed conditions. Water pollution arises from a variety of sources—industry, domestic sewage, agricultural fertilizers and feedlots, construction activities, and forest practices. Principal pollutants are sediments, organic pollutants containing nitrites, and phosphates, lead, mercury, hydrocarbons, pesticides, and various synthetic chemicals. The impact of the pollutant depends both on its chemical nature and on the quantity released. All water bodies receive quantities of chemicals and solid materials; a variety of organisms which break down and utilize these inputs have evolved. Serious problems arise when the inputs to the water body become unusually large or contain synthetic materials

which cannot be decomposed by the extant organisms. Further problems may develop through concentrations of pollutants in the food web of the aquatic system. For example, if a chemical is not metabolized by organisms, but is concentrated in their tissues, as are some pesticides, then as each organism is eaten by another, the body burden of chemicals can increase. In this way, predators may obtain very large and dangerous amounts of pollutants. The decline in population of certain fish-eating birds has been attributed to this process of transfer and concentration.

Aquatic pollution has many consequences of significance to humans. An excess of chemical materials which enrich plant and animal growth can cause rapid increase in life. This process, termed eutrophication, may entail dramatic increases in the algal, planktonic, or rooted aquatic plant populations, with the result that the water body appears green in color or becomes clogged with vegetation. Toxic chemicals released to water bodies may directly kill aquatic life or, if present in sublethal amounts, may change the species of plants and animals present. Often aquatic pollution is not a dramatic either-or proposition, with all fish or other aquatic life killed; rather, more commonly a trend toward an increase in the more resistant species is seen, with the elimination of those forms which are especially susceptible to the pollutant. Aquatic pollution also involves heat, especially that derived from industrial activities, including nuclear power and fossil fuel plants. In these instances, water is used to cool the machines or reactors and is exhausted to the environment at elevated temperatures. Since all metabolic and chemical processes are influenced by heat, thermal pollution should have a significant effect on aquatic systems, but thus far it has been difficult to prove that such an impact occurs.

Other aquatic problems of interest to applied ecology concern alteration of water channels by impoundments or channelization and irrigation. In each instance the natural pattern of water movement is altered, and deterioration of the environment may result. Impoundments limit the natural movement of sediment and chemical elements; production patterns in the water and lands below the impoundment may be altered, and other changes may occur. But, on the plus side, impoundments often provide fisheries, electrical energy, recreation, and other advantages. Irrigation problems may involve the movement of salts from depths in the soil, with deposition near the surface. Disturbance of the chemical equilibrium of the soil, in turn, interferes with plant growth. *See* WATER POLLUTION.

Terrestrial and soil problems. Terrestrial environments constantly undergo a degrading and decomposing process owing to the action of water, frost, wind, and other environmental processes on the surface which involve the linkages joining land, water, and atmosphere. Human activities may accelerate these natural processes. In addition, the use of chemical materials on the land may have effects similar to those resulting from their addition to water. Most terrestrial environmental problems are caused by agricultural, grazing, or forest practices. Probably the most serious effect concerns practices which increase the rate of surface erosion. Only a small percentage of the Earth's surface is suitable for agriculture, and the loss of soil from these areas is extremely serious. As a consequence, certain regions have been denuded and are no longer productive. Overgrazing may also remove the cover of vegetation and allow water and wind to erode the soil. Deserts have increased in extent almost everywhere because of overgrazing, and in India the increase in the Rajasthan Desert can be measured in feet per year. Dust from this desert blows as far east as Thailand. Overcutting trees and lack of reforestation programs also may increase soil erosion and nutrient losses in forest regions. These impacts are not solely the mark of modern civilization. Misuse of the land has been noted in many past civilizations and can even be a problem for present-day primitive societies. *See* DESERTIFICATION; EROSION; FOREST AND FORESTRY.

However, modern agriculture has added new problems to those of the primitive farmer. Various chemicals such as fertilizers, pesticides, and herbicides are used to increase agricultural production. These chemicals may be needed because of past misuse of the land or because of economic demands in a society that does not recognize the need to maintain and protect terrestrial resources. Organic and ecological agricultural practices seek to reestablish a pattern of land use without causing deterioration of the soil and biotic resources. Although reestablishment of the pattern may result in somewhat lower productivity, proponents of "ecoagriculture" argue that high productivity can be maintained without loss of soil through erosion and without a reduction in fertility. *See* AGRICULTURE.

Probably the most serious short-term impact arises from the use of chemicals on the land. In the most extreme cases, the health of the agriculturalist may be affected by the materials. But more commonly, the pesticide or chemical in the soil is taken up by the crop and then enters the human food chain. Many modern governments maintain agencies to advise farmers on the proper amounts of chemicals to apply so that buildup does not occur; other agencies periodically sample and analyze foodstuffs for residues. In this way, the consumer can be protected from misuse of chemicals or from excessive concentrations. Unfortunately, these agencies seldom consider the impacts of agricultural chemicals on other animal food chains. For example, the soil fauna and the natural nitrogen-fixing organisms present in the soil, as well as the terrestrial faunas living near the agricultural or forest plantations, can be significantly affected; and populations of animals, even beneficial species, may be reduced through misuse of chemical materials. However, extinction of plants and animals, which is also a serious applied ecology problem, usually is due to the destruction of their habitat. Pollution, disturbance of the land, and overhunting may provide the final cause of the destruction of a particular living species.

Nuclear energy. Industrialized societies require large quantities of energy. Energy production from nuclear reactors has been enthusiastically developed in many regions of the biosphere. However, nuclear energy also has environmental consequences that are of concern to applied ecology and that must be considered when these facilities are designed and operated, so that negative impacts on the environment do not occur.

Nuclear energy has three primary environmental consequences: the storage of radioactive products, release of radioactive material to the environment, and, as mentioned above, release of heat.

There are various kinds of radioactivity associated with the particular elements used in the reactor. In the process of the generation of energy, these fuel elements are changed into a suite of radioactive materials. Although these materials, in turn, form a new source of chemicals, the process of separation and concentraton is very costly and dangerous. In either case, however, the processes result in radioactive waste that must be stored for periods of hundreds, even thousands of years. The fact that the potential danger of these wastes will require technical attention for periods of time longer than the histories of many modern societies is a prime argument against the widespread use of nuclear energy. However, proponents of the use of nuclear energy state that certain geological structures such as salt mines can be safely used for storage indefinitely.

A second environmental problem concerns the loss of relatively small quantities of radioactive materials to the environment during chemical processing in reactors or chemical plants. If these materials enter the body, they can cause disease and death. Like pesticides, radioactive chemical materials may be concentrated in food chains and can appear in relatively large concentrations. In certain fragile environments such as the arctic tundra, the food chains are very short. Thus, radioactive chemicals derived from testing atomic weapons pass through lichens or reindeer or caribou to humans. Concentrations in certain localities may be high enough to cause concern to public health authorities. *See* NUCLEAR REACTOR; RADIOACTIVITY.

Population problems. Applied ecology also is concerned with the size of the human population, since many of the impacts of human activities on the environment are a function of the number and concentration of people. The human population on Earth has increased exponentially, and in many countries this increase poses almost insurmountable problems. Control of environmental degradation, even a concern for environment, is nonexistent when the population is undernourished, starving, ill-housed, and underemployed. Social disorder, alienation, psychological disturbances, physical illness, and other problems have been correlated with overpopulation. Population also places demands on the resources of the Earth; as the standard of living rises, these demands increase.

The human population problem is exceedingly complicated because the growth of population is controlled largely by the decisions of individual families. The family may visualize several different strategies—the number of children that is best for the family, best for their social group or tribe, or best for the human race—or have no strategy at all. Families may decide that a large family is best even under serious conditions of overpopulation. Considerable evidence indicates that the size of the family declines as the population becomes less rural and more industrial. Thus, some specialists urge economic development, regardless of environmental impact, as a means of solving the population problem. Others urge that the family be more directly influenced to reduce the number of children. Direct action might entail birth control advice, medical abortion and sterilization, and taxation. Yet others argue that no measures such as these can be significant and that the human population will be controlled by famine, war, or disease. Each of these positions leads toward a set of social policies, all of which have an environmental impact which, in turn, affects the human society. *See* HUMAN ECOLOGY.

Environmental planning and design. The foregoing discussion suggests that there is an optimum environment for the human race which is influenced by a variety of population densities and activities. Thus, although the population of the United States is relatively sparse, it has a large environmental impact. This, in turn, suggests that the human environment and society could be designed in such a way to minimize the negative impacts and provide a satisfactory productive life for the population. Environmental design considers economic and social policy, as well as the impact of designed rural, urban, transport, industrial, and other systems. It also considers the design of the individual environment of house, furniture, clothes, and so on. Considering the often violent impact modern society has had on the environment, a design revolution is required to reorganize the environment created by society so that these impacts can be reduced.

Environmental planning and design obviously have a deep political component, since the methods used to redesign society depend upon the control of individual demands. At one extreme, individual demand is allowed to express itself without limit, and education is used to create in the individual a realization that environmental constraints must be recognized. At the other extreme, the government or party controls demand through regulation. Most societies operate somewhere between these extremes.

Throughout human history, utopian designs have been developed for human societies and environments. Today these designs pay more attention to environmental features of society and are often labeled as ecological. Applied ecology, thus, considers not only the alteration of specific features of the modern industrial society to correct some environmental defect but also the fundamental reorientation of society to achieve a balance between humans and the natural world on which they depend. *See* ECOLOGY.

Frank B. Golley

Ecosystem

A functional system that includes an ecological community of organisms together with the physical environment, interacting as a unit. Ecosystems are characterized by flow of energy through food webs, production and degradation of organic matter, and transformation and cycling of nutrient elements. This production of organic molecules serves as the energy base for all biological activity within ecosystems. The consumption of plants by herbivores (organisms that consume living plants or algae) and detritivores (organisms that consume dead organic matter) serves to transfer energy stored in photosynthetically produced organic molecules to other organisms. Coupled to the production of organic matter and flow of energy is the cycling of elements. *See* ECOLOGICAL COMMUNITIES; ENVIRONMENT.

Autotrophic production. All biological activity within ecosystems is supported by the production of organic matter by autotrophs (organisms that can produce organic molecules such as glucose from inorganic carbon dioxide; **Fig. 1**). More than 99% of autotrophic production on Earth is through photosynthesis by plants, algae, and certain types of bacteria. Collectively these organisms are termed photoautotrophs (autotrophs that use energy from light to produce organic molecules). In addition to photosynthesis, some production is conducted by chemoautotrophic bacteria (autotrophs that use energy stored in the chemical bonds of inorganic molecules such as hydrogen sulfide to produce organic molecules). The organic molecules produced by autotrophs are used to support the organism's metabolism and reproduction, and to build new tissue. This new tissue is consumed by herbivores or detritivores, which in turn are ultimately consumed by predators or other detritivores.

Terrestrial ecosystems, which cover 30% of the Earth's surface, contribute a little over one-half of the total global photosynthetic production of organic matter (approximately 60×10^{15} grams of carbon per year; **Fig. 2**). The global rate of photosynthetic production of organic matter in terrestrial ecosystems is regulated by temperature, water availability, and nutrient concentrations. Correspondingly, the rate of photosynthetic production in the tropics (average rate = 1800 g C m^{-2} y^{-1}) is 25-fold greater than in deserts (average rate = 70 g C m^{-2} y^{-1}) and 13-fold greater than in tundra and alpine ecosystems (average rate = 140 g C m^{-2} y^{-1}).

Oceans, which cover 70% of the Earth's surface, produce approximately 51×10^{15} g C y^{-1} of organic matter (Fig. 2). Photosynthesis in oceans varies greatly across the Earth in response to nutrient availability (primarily nitrogen and iron). Photosynthesis is greatest along continental margins and in regions where deep, nutrient-rich water upwells into surface waters receiving light. Photosynthetic production is greatest in regions of deep upwelling (averaging 420 g C m^{-2} y^{-1}), intermediate in coastal margins (averaging 250 g C m^{-2} y^{-1}), and lowest in the open ocean (averaging 130 g C m^{-2} y^{-1}).

Whereas most ecosystems receive sunlight to support photosynthesis, photosynthetic production does occur in the deep ocean and in caves. In a few unique ecosystems, such as deep-sea thermal vents and caves, the basis of organic matter production is by chemoautotrophic bacteria. These ecosystems harbor surprisingly diverse communities that function in the complete absence of sunlight.

Heterotrophic production. The organic matter produced by autotrophic organisms is ultimately consumed by heterotrophs (organisms that obtain energy and nutrients by consuming other organisms; Fig. 1). Globally, most organic matter produced by plants and algae is consumed by detritivores. Herbivores consume only about 5% of photosynthetic production, fire destroys another 5%, and a small fraction of autotrophic production is lost through burial within the Earth's crust (to potentially become fossil fuels). The proportion of photosynthetic production consumed by herbivores can vary considerably among ecosystems. In terrestrial ecosystems, much plant tissue is structural molecules such as cellulose and lignin that are difficult or impossible for most organisms to digest. In aquatic ecosystems, in which algae dominate, a much greater proportion of photosynthetic production is consumed by herbivores. Algae have little in the way of structural molecules and contain more molecules that can be consumed by grazing organisms. *See* BIOMASS; ECOLOGICAL SUCCESSION.

Food webs. Organisms are classified based upon the number of energy transfers through a food web (Fig. 1). Photoautotrophic production of organic matter represents the first energy transfer in ecosystems and is classified as primary production. Consumption of a plant by a herbivore is the second energy transfer, and thus herbivores occupy the second trophic level, also known as secondary production. Consumer organisms that are one, two, or three transfers from photoautotrophs are classified as primary, secondary, and tertiary consumers. Moving through a food web, energy is lost during each transfer as heat, as described by the second law of thermodynamics. Consequently, the total number of

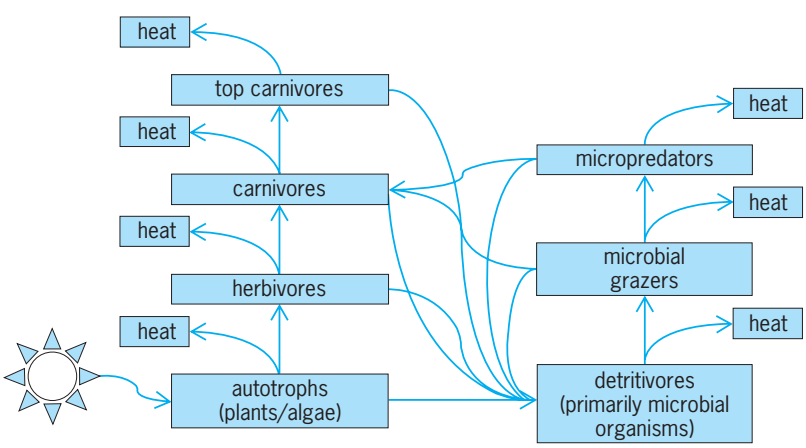

Fig. 1. General model of energy flow through ecosystems.

energy transfers rarely exceeds four or five; with energy loss during each transfer, little energy is available to support organisms at the highest levels of a food web. *See* ECOLOGICAL ENERGETICS; FOOD WEB.

Biogeochemical cycles. In contrast to energy, which is lost from ecosystems as heat, chemical elements (or nutrients) that compose molecules within organisms are not altered and may repeatedly cycle between organisms and their environment. Approximately 40 elements compose the bodies of organisms, with carbon, oxygen, hydrogen, nitrogen, and phosphorus being the most abundant. If one of these elements is in short supply in the environment, the growth of organisms can be limited, even if sufficient energy is available. In particular, nitrogen and phosphorus are the elements most commonly limiting organism growth. This limitation is illustrated by the widespread use of fertilizers, which are applied to agricultural fields to alleviate nutrient limitation. *See* BIOGEOCHEMISTRY.

Nitrogen cycle. Nitrogen commonly limits the rate of primary production in terrestrial, fresh-water, estuarine, and oceanic ecosystems (**Fig. 3**). In one turn of the biogeochemical cycle of nitrogen, (1) cyanobacteria and certain types of eubacteria transform atmospheric dinitrogen into organic molecules such as protein (nitrogen fixation); (2) organisms are consumed by other organisms and organic nitrogen is converted into ammonium as a waste product (decomposition); (3) ammonium either is assimilated by plants, algae, and bacteria into organic forms or is converted to nitrate by chemoautotrophic bacteria (nitrification); and (4) nitrate is converted back to dinitrogen by anaerobic bacteria (denitrification). Whereas nitrogen is very abundant in the atmosphere as dinitrogen (N_2), only a select group of organisms can transform dinitrogen into organic forms (amino acids), and the transformation requires a lot of metabolic energy. Humans have greatly impacted the global nitrogen cycle by doubling the amount of ammonium, nitrate, and organic nitrogen in the environment through the widespread use of fertilizers and combustion of fossil fuels. This massive change in nitrogen forms has led to problems with acid precipitation, alterations of global patterns of primary production rates, increased production of nitrous oxide (a greenhouse gas), and pollution of aquatic ecosystems.

Phosphorus cycle. Phosphorus also commonly limits primary production. In contrast to nitrogen, phosphorus lacks an atmospheric phase. The primary source of phosphorus to ecosystems is the weathering of the Earth's crust. This inorganic phosphorus is then assimilated by plants, algae, and bacteria and is incorporated into organic forms such as phospholipids and adenosine triphosphate (ATP). As organisms are consumed by other organisms, organic phosphorus is converted back to inorganic forms. Both organic and inorganic forms are slowly carried by water from plants and soils to streams and rivers, and ultimately to the world's oceans where phosphorus is buried in sediments. The global phosphorus cycle is completed over long, geologic time scales when ocean sediments are compacted into rock and this new rock is uplifted onto continents. As with the nitrogen cycle, humans have dramatically impacted the phosphorus cycle by application of fertilizers. The major source of phosphorus fertilizers is from the mining of guano. Current estimates project readily accessible sources of guano being depleted by the end of the twenty-first century.

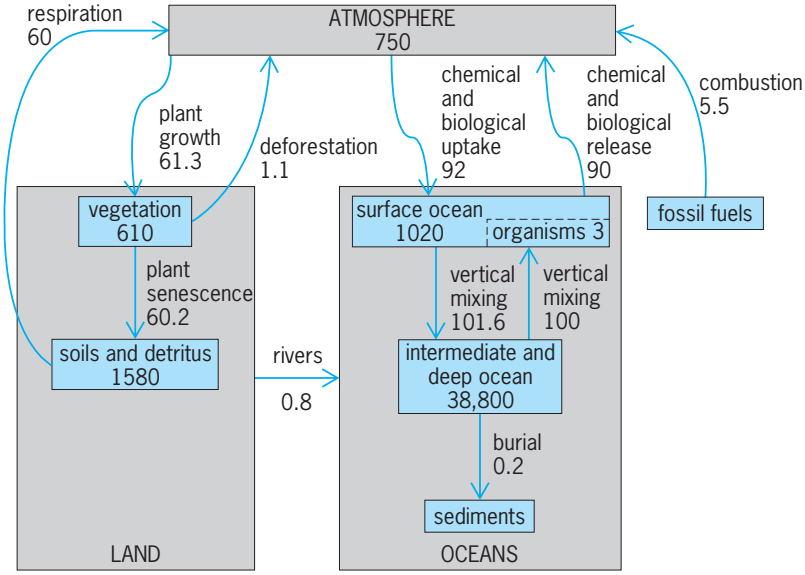

Fig. 2. Model of the global carbon cycle illustrating movement of carbon between the atmosphere and terrestrial and oceanic ecosystems. The storage of carbon in these ecosystems is in picograms (1 pg = 10^{15} g), and fluxes of carbon between boxes are in pg y^{-1}.

Carbon cycle. The carbon cycle has also received considerable research focus, especially with the increase of carbon dioxide in the atmosphere due to the combustion of fossil fuels, and the impact on global climate. Carbon cycles between the atmosphere and terrestrial and oceanic ecosystems. This cycling results, in part, from primary production and decomposition of organic matter. Rates of primary production and decomposition, in turn, are regulated

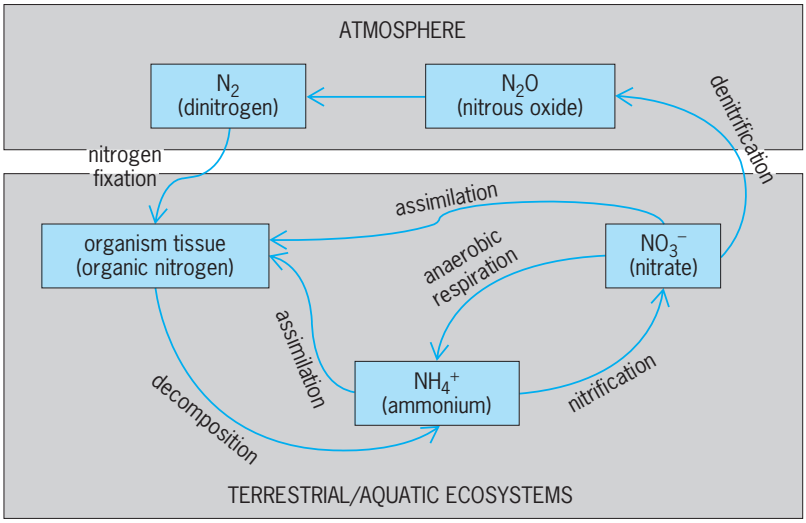

Fig. 3. General model of the nitrogen cycle in ecosystems.

by the supply of nitrogen, phosphorus, and iron. The combustion of fossil fuels is a recent change in the global cycle that releases carbon that has long been buried within the Earth's crust to the atmosphere. Carbon dioxide in the atmosphere traps heat on the Earth's surface and is a major factor regulating the climate. This alteration of the global carbon cycle along with the resulting impact on the climate is a major issue under investigation by ecosystem ecologists. *See* AIR POLLUTION; CONSERVATION OF RESOURCES; ECOLOGY, APPLIED; HUMAN ECOLOGY; WATER POLLUTION.

Jeremy B. Jones

Bibliography. F. B. Golley, *A History of the Ecosystem Concept in Ecology*, Yale University Press, 1993; J. B. Hagen, *An Entangled Bank: The Origins of Ecosystem Ecology*, Rutgers University Press, 1992; W. H. Schlesinger, *Biogeochemistry: An Analysis of Global Change*, 2d ed., Academic Press, 1991; R. H. Waring and S. W. Running, *Forest Ecosystems: An Analysis at Multiple Scales*, Academic Press, 1998.

Ecotone

A geographic boundary or transition zone between two different groups of plant or animal distributions. The term has been used to denote transitions at different spatial scales or levels of analysis, and may refer to any one of several attributes of the organisms involved. For example, an ecotone could refer to physiognomy (roughly, the morphology or appearance of the relevant organisms), such as between the boreal forest and grassland biomes; or it could refer to composition, such as between oak-hickory and maple-basswood forest associations; or it could refer to both. Ecotones are generally distinguished from other geographic transitions of biota by their relative sharpness. The ecotone between boreal forest and prairie in central Saskatchewan occurs over a hundred kilometers or so, in contrast to the transition from tropical forest to savanna in South America or Africa that is associated with increasing aridity and is dispersed over hundreds of kilometers. The "tension zone" between broadleaf deciduous forests in south-cental Michigan and mixed forests to the north is similarly sharp. Ecotones are thought to reflect concentrated long-term gradients of one or more current environmental (rather than historical or human) factors. Though often climatic, these factors can also be due to substrate materials, such as glacial sediments or soils. Regardless of their specific environmental basis, most ecotones are thought to be relatively stable.

A good example of an ecotone is the abrupt prairie-forest transition in North America that extends southeastward from south-central Alberta, Canada, through central Minnesota, and ultimately (in more diffuse form) to approximately southern Lake Michigan. Separating the grassland biome in the western Midwest and Great Plains from the boreal and deciduous forest biomes to the northeast, this ecotone is thought to reflect long-term patterns of precipitation and evaporation (warmer, drier to the southwest). These, in turn, are produced by modal (most frequent) atmospheric circulation patterns over North America as they repeatedly steer daily weather features across the region and ultimately generate environments, either directly or indirectly through patterns of disturbance, that are conducive to prairie or forest persistence in different parts of the region. Typically, specific sites within an ecotone will support communities that are more like one or the other broad groupings of biota the ecotone separates, depending on local factors including environmentally related patterns of disturbances such as fire.

Ecotones are often reflected in the distributions of many biota besides the biota used to define them. The prairie-forest ecotone, for example, is defined not only by the dominant vegetation components but also by many faunal members of the associated ecosystems, such as insects, reptiles and amphibians, mammals, and birds, that reach their geographic limits here. Across the glaciated midwestern United States, where landscapes of different geologic ages co-occur, the associated soil patterns are sometimes reflected in similarly bold ecological patterns. In central Illinois, for example, the transition from a Wisconsinan-aged glacial landscape (deposited approximately 20,000 years ago) in the north to an Illinoisan-aged landscape (several hundred thousand years older) to the south corresponds to a transition to older (better-developed) soils, different dominant tree species in the upland wooded tracts, and even contrasts in the herpetofauna, with many species in one area historically absent from the other. *See* ALPINE VEGETATION; BIOME; ECOLOGICAL COMMUNITIES; ECOSYSTEM; FOREST ECOSYSTEM; GRASSLAND ECOSYSTEM; LIFE ZONES; SAVANNA; ZOOGEOGRAPHY.

Jay R. Harman

Bibliography. J. A. Harrington and J. R. Harman, Climate and vegetation in central North America: Natural patterns and human alterations, *Great Plains Quart.*, 11:103–112, 1991; J. B. Lachavanne and R. Juge (eds.), *Biodiversity in Land/Inland Water Ecotones*, CRC Press–Parthenon, 1997; K. E. Medley and J. R. Harman, Relationships between the vegetation tension zone and soils distribution across central Lower Michigan, *Mich. Bot.*, 26:78–87, 1987.

Eddy current

An electric current induced within the body of a conductor when that conductor either moves through a nonuniform magnetic field or is in a region where there is a change in magnetic flux. It is sometimes called Foucault current. Although eddy currents can be induced in any electrical conductor, the effect is most pronounced in solid metallic conductors. Eddy currents are utilized in induction heating and to damp out oscillations in various devices.

Causes. If a solid conductor is moving through a nonuniform magnetic field, electromotive forces (emfs) are set up that are greater in that part of the conductor that is moving through the strong part of

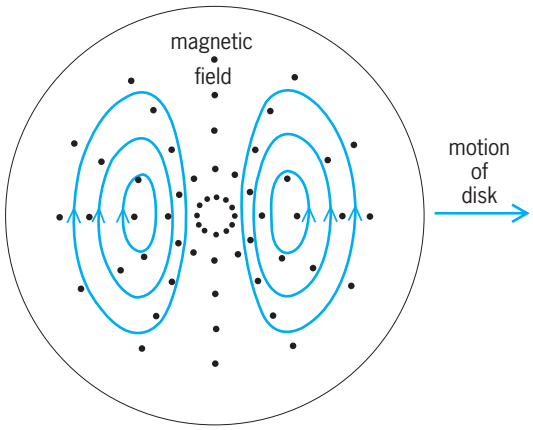

Fig. 1. Eddy currents which are induced in a disk moving through a nonuniform magnetic field.

the field than in the part moving through the weaker part of the field. Therefore, at any one time in the motion, there are many closed paths within the body of the conductor in which the net emf is not zero. There are thus induced circulatory currents that are called eddy currents (**Fig. 1**). In accordance with Lenz's law, these eddy currents circulate in such a manner as to oppose the motion of the conductor through the magnetic field. The motion is damped by the opposing force. For example, if a sheet of aluminum is dropped between the poles of an electromagnet, it does not fall freely, but is retarded by the force due to the eddy currents set up in the sheet. If an aluminum plate oscillates between the poles, it will be stopped quickly when the switch is closed and the field set up. The energy of motion of the aluminum plate is converted into heat energy in the plate. *See* ELECTROMAGNETIC INDUCTION; LENZ'S LAW.

Eddy currents are also set up within the body of material when it is in a region in which the magnetic flux is changing rapidly, as in the core of a transformer. As the alternating current changes rapidly, there is also an alternating flux that induces an emf in the secondary coil and at the same time induces emfs in the iron core. The emfs in the core cause eddy currents that are undesirable because of the heat developed in the core (which results in high energy losses) and because of an undesirable rise in temperature of the core. Another undesirable effect is the magnetic flux set up by the eddy currents. This flux is always in such a direction as to oppose the change that caused it, and thus it produces a demagnetizing effect in the core. The flux never reaches as high a value in the core as it would if there were no eddy currents present.

Laminations. Induced emfs are always present in conductors that move in magnetic fields or are present in fields that are changing. However, it is possible to reduce the eddy currents caused by these emfs by laminating the conductor, that is, by building the conductor of many thin sheets that are insulated from each other rather than making it of a single solid piece. In an iron core the thin iron sheets are insulated by oxides on the surface or by thin coats of varnish. The laminations do not reduce the induced emfs, but if they are properly oriented to cut across the paths of the eddy currents, they confine the currents largely to single laminae, where the paths are long, making higher resistance; the resulting net emf in the possible closed path is small. Bundles of iron wires or powdered iron formed into a core by high pressure are also used to break up the current paths and reduce the eddy currents. *See* CORE LOSS.

Kenneth V. Manning

Testing. The testing of metallic materials through the detecting of eddy currents induced within them is widely used. A high-frequency alternating current in a coil placed close to a metallic material induces eddy currents within the material through electromagnetic induction. The flux caused by the eddy currents threads the coil producing the original flux and alters its apparent impedance. This alteration is a measure of the strength of the eddy currents. The flow of the eddy current is impeded by the material's electric resistance, which reflects characteristics such as hardness and chemical composition. Therefore, eddy current measurements reveal various properties of materials.

The hardness of aluminum alloys used in the aerospace industry can be related to their electrical conductivity. Eddy-current conductivity meters are widely used to nondestructively monitor these alloys at various stages of aircraft production, in particular after critical heat treatment processes. The rapid measurement of electrical conductivity is one of the many checks made by coin vending machines, so the conductivity of alloys used in the manufacture of coins is often determined using an eddy-current conductivity meter to assure close quality control.

Flaws such as open cracks within materials cause eddy currents to detour around them and effectively enhance electric resistance; accordingly, eddy current testing can be used to detect flaws. Differences in thickness of foils and plating cause differences in the electric resistance; hence, eddy current testing is also convenient for estimating thickness. Stronger

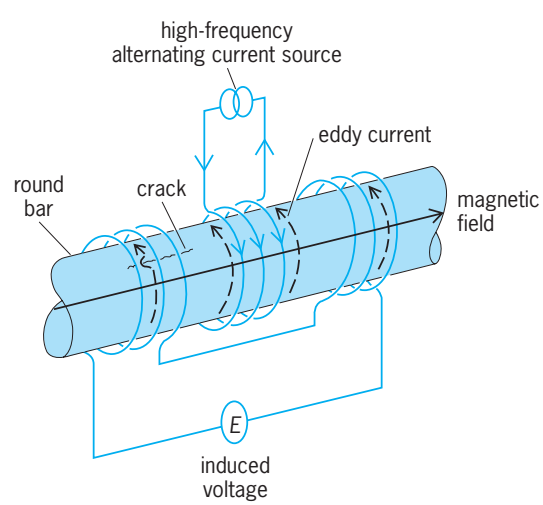

Fig. 2. The alignment of coils used to inspect a long bar.

magnetic fields are induced within ferromagnetic materials undergoing testing, so that the magnetic permeability also affects the eddy current. The evaluation of the hardness of ferritic steel through eddy current testing is based upon this effect of permeability, which is very susceptible to hardness. *See* FERROMAGNETISM; MAGNETIC MATERIALS; MAGNETISM.

Figure 2 shows the general alignment of coils to detect flaws within long bars. Annular coils encircle the round bar. The alternating current applied in the central coil generates a magnetic field inside the coil and within the bar in the direction of their axis, and thereby induces eddy currents along concentric circular paths within the bar. Simultaneously this eddy current produces a counter magnetic field which is superimposed on the primary one. The combined magnetic field gives rise to electromotive forces, that is, voltages between terminals of each coil. Two coils on either side of the central coil are wound in opposite directions, connected in series. The voltage between terminals is monitored continuously. When a flaw within a moving bar enters one of the two coils, the voltage reading corresponds to the difference between the voltages of the coils. This alignment permits the detection of flaws with high sensitivity, because it eliminates noises that originate from irregularities in diameter or physical properties of materials, whose variation is negligibly small between the closely spaced coils.

Since eddy current testing is of the noncontacting type, it can detect flaws within materials moving at high speed and even those raised to red heat in hot-rolling mills. Inspection of the latter is advantageous to exclude the influence of variation in permeability, because ferromagnetism vanishes at temperatures above the Curie point (1416°F or 769°C for iron). *See* CURIE TEMPERATURE; NONDESTRUCTIVE EVALUATION. Kazuo Watanabe; A. E. Drake

Braking and damping. The principle outlined above, whereby a conductor which is moving through a nonuniform magnetic field experiences a retarding force, is used to brake or damp the oscillations of moving objects. In these applications, the field is that produced by a permanent magnet or an electromagnet supplied with direct current. The retarding force is proportional to the product of the flux produced by the magnet ϕ and the effective eddy current I produced by it. The effective eddy current, however, results from the rate of flux cutting, that is, the flux ϕ multiplied by the velocity of movement υ. The retarding force is thus proportional to $\phi^2\upsilon$. For a constant flux, which applies with a permanent magnet or electromagnet with constant current, the retarding force is proportional to the velocity of the moving object. With an electromagnet, the current can be varied, to adjust the ratio of retarding force to velocity. This ratio is also affected by the physical dimensions and materials used, for example, the resistivity of the moving object. Brakes are often fitted to rotating objects, and in these applications the retarding torque is proportional to the angular velocity ω.

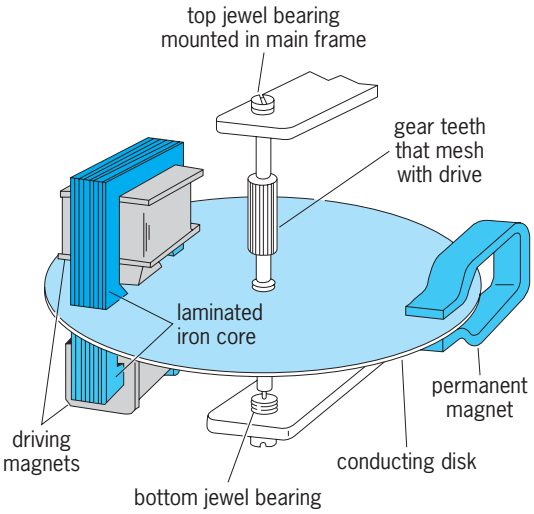

Fig. 3. Basic construction of electromechanical energy meter.

Eddy-current brakes are probably most widely used in the electromechanical energy meters which are installed in consumers' premises by electricity supply companies. In this application, the meters (**Fig. 3**) contain electromagnets which produce fluxes proprotional to the system voltage and current. These fluxes induce emf's and eddy currents in a conducting disk, which is mounted in jewel bearings so that it may rotate freely. The eddy currents react with the fluxes to produce a torque proportional to the power being supplied. Eddy-current braking, produced by a flux from a permanent magnet cutting the disk, provides a retarding torque proportional to the disk speed. Because the frictional torques are negligible, the disk runs at a speed at which the driving and retarding torques are almost equal, and as a result the speed is proportional to the power being supplied to the consumer. A geared drive to indicating dials meshes with gear teeth on the disk's spindle. *See* WATT-HOUR METER.

The same form of braking is also employed widely in the induction-type relays to provide time-graded protection. Eddy-current brakes incorporating electromagnets are used to provide controllable load torques on test equipment for small machines. *See* RELAY.

Eddy-current damping is used to suppress undesirable mechanical oscillations and is used widely in deflection-type indicating instruments. *See* DAMPING. Arthur Wright

Bibliography. Atomic Energy of Canada, Ltd, *Eddy Current Testing*, 2 vols., 1987; G. Birnbaum and G. Free (eds.), *Eddy-current Characterisation of Materials and Structures*, ASTM Spec. Publ., no. 722, American Society for Testing and Materials, 1981; B. I. Bleaney and B. Bleaney, *Electricity and Magnetism*, 2 vols., 3d ed., 1976, paper, 1989; D. Hagemaier, *Fundamentals of Eddy Current Testing*, 1990; A. E. Drake and A. C. Lynch, AC conductivity standards for the calibration of eddy-current conductivity meters, *J. Phys. E*, 120:137–139, 1987;

H. L. Libby, *Introduction to Electromagnetic Nondestructive Test Methods*, 1971, reprint, 1979; E. M. Purcell, *Electricity and Magnetism*, 2d ed., 1985; S. S. Udpa and P. O. Moore (eds.), *Eddy Current Testing*, vol. 5 of *Nondestructive Testing Handbook*, 3d ed., American Society for Nondestructive Testing, 2004.

Edema

An abnormal accumulation of fluid in the cells, tissue spaces, or cavities of the body, also known as dropsy. An excess of fluid in the pleural spaces is referred to as hydrothorax, in the pericardial sac as hydropericardium, and in the peritoneal cavity as ascites. Anasarca is a generalized subcutaneous edema.

There are three main factors in the formation of generalized edema and a fourth which plays an important role in the formation of local edema. They are (1) permeability of the capillary wall, (2) colloid osmotic pressure of the plasma proteins, (3) hydrostatic pressure in the capillaries, and (4) lymphatic obstruction.

Permeability of capillary wall. Normally the capillary walls are freely permeable to water, salts, and dissolved gases but are almost impermeable to proteins. When the vessel wall is injured by toxins, anoxia, or paralytic dilatation, the capillary endothelium becomes permeable to proteins. With a diffusion of protein into the tissues, the plasma osmotic pressure is lowered and the osmotic pressure of the tissue is increased. Under these circumstances fluid collects in the tissue spaces.

Such a condition plays an important role in inflammatory edema. It is a factor in the edema of severe infections, metabolic intoxications, asphyxia, anaphylactic reactions, secondary shock, and acute nephritis. It also contributes to the edematous conditions when there is a fall in the level of plasma proteins.

Osmotic pressure of plasma proteins. A fall in plasma proteins tends to decrease the forces tending to reabsorb and hold fluid in the vascular compartment. Albumin is the protein of greatest importance in this regard. When the plasma protein level drops below 3 g/100 ml, the colloid osmotic pressure is no longer sufficient to maintain a balance with the hydrostatic pressure of the blood, which tends to drive fluid out into the tissue spaces. Hence more fluid goes out into the tissue spaces and remains there until a new equilibrium is reached.

This form of edema is seen in association with prolonged malnutrition (nutritional edema) as during a famine or with chronic nutritional or metabolic defects. With a marked loss of albumin in the urine as in nephrotic syndrome, there follows a lowering of the plasma albumin fraction and the development of edema.

In edema of kidney disease the protein content and specific gravity of the edema fluid are low. The blood reveals a markedly elevated cholesterol and a drop in total protein level with a relatively greater drop in the albumin fraction. As the plasma osmotic pressure drops, water passes into the tissues and with it crystalloids. These substances, especially sodium chloride, are retained in the tissues. Thus as water is taken in, it passes rapidly into the tissues and is not eliminated into the urine. *See* OSMOREGULATORY MECHANISMS.

Capillary hydrostatic pressure. Under normal conditions the hydrostatic pressure in the arterial end of the capillary is sufficient to overcome the plasma osmotic pressure and drives fluid out of the capillary into the tissue spaces. During passage through the capillary the pressure drops to a level low enough to allow the osmotic pressure of the proteins to draw fluid back into the vascular compartment. An increase in hydrostatic pressure at the venous end of the capillary will upset the balance, resulting in a decreased absorption of tissue fluid by the osmotic pressure of the plasma proteins. Under these circumstances an increased amount of fluid will be returned via the lymphatics, but as the condition progresses edema will develop. Such a situation can follow venous congestion of long duration. *See* CIRCULATION.

Cardiac edema following the generalized venous congestion of cardiac failure is the commonest form of this type of edema. The fluid which collects in the tissue spaces is affected by changes in position, being more marked in the dependent portions of the body. Fluid also collects in the serous cavities. The lymphatics empty into the venous system, and therefore a rise in pressure in the venous system results in an increased pressure within the lymphatics, which also contributes to the edema formation. As the capillary walls are distended, they become more permeable. In addition, a state of chronic hypoxia may exist, causing a further insult to the capillary endothelium with a loss of protein into the tissue spaces.

Other factors seem to play a role in cardiac edema. With decreased cardiac output there is reduced renal blood flow and glomerular filtration rate, with a consequent reduced excretion of salt and water. This may be responsible for an increased volume of extracellular fluid and plasma which, in turn, is followed by a rise in venous pressure.

Pulmonary edema (edema within the lung) is usually a form of cardiac edema but may be secondary to other factors such as inflammation. *See* RESPIRATORY SYSTEM DISORDERS.

Another condition resulting from an increased hydrostatic pressure in the capillaries is postural edema. This occurs when an individual has been standing motionless for a long period of time; the fluid collects in the subcutaneous tissues of the feet and ankles.

Cirrhosis of the liver causes an impediment to the flow of blood through the portal circulation. There is a consequent rise in venous pressure and ascites forms. *See* CIRRHOSIS.

Lymphatic obstruction. A portion of the intercellular tissue fluids returns to the circulation via the lymphatics. Obstruction to this channel will contribute to local edema. Infestation by filaria is one cause of lymphatic obstruction, particularly in the

tropics. Lymphatic channels may be destroyed or obstructed by surgical procedures, resulting in localized edema. Milroy's disease is a chronic hereditary edema thought to be due to lymphatic obstruction. *See* LYMPHATIC SYSTEM.

Management of edema. The management of patients with edema is directed toward the treatment of the underlying medical condition. Diuretics are often used, most successfully in cardiac edema. Romeo A. Vidone; Irwin Nash

Bibliography. A. C. Guyton and J. E. Hall, *Textbook of Medical Physiology*, 9th ed., 1996; J. M. Kissane (ed.), *Anderson's Pathology*, 9th ed., 1989; S. L. Robbins et al., *Pathologic Basis of Disease*, 6th ed., 1999; J. B. Wyngaarden and L. H. Smith (eds.), *Cecil Textbook of Medicine*, 17th ed., 1985.

Edentata

The former name of an order that included the Xenarthra (sloths, anteaters, armadillos), Pholidota (pangolins), and Tubulidentata (aardvarks). The order was so named because of the apparent lack of some or all teeth, although most of these species are not toothless but have at least vestigial teeth. It was subsequently realized that Edentata was polyphyletic—that it contained unrelated families and was thus invalid. It is now thought that any similarities between these groups is the result of similar adaptations to a common way of life and are not indicative of actual relationships. Thus, pangolins and aardvarks are now placed in separate orders, and the order Xenarthra was established to group the remaining families (which are all related). The term edentate is now used only as a term of convenience when referring to these unusual mammals; it has no taxonomic validity. *See* AARDVARK; ANTEATER; ARMADILLO; DENTITION; MAMMALIA; PHOLIDOTA; SLOTH; TOOTH; TUBULIDENTATA. Donald W. Linzey

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Ediacaran biota

A widely distributed group of soft-bodied marine organisms that are preserved as fossils in rocks of latest Proterozoic age (600–543 million years ago, or Ma). The biota characterizes a geological period, known as the Ediacarian or the Vendian, which precedes the widespread appearance of animals with mineralized skeletons. The name Ediacara refers to an abandoned mining area about 380 mi (600 km) north of Adelaide, South Australia.

Discovery and distribution. Although Ediacaran fossils were described in the 1930s from southwest Africa (Namibia), it was the discovery in 1946 of abundant fossil "jellyfish" at Ediacara that sparked international interest in this biota. Subsequently, similar or identical fossils were found in central England (1957); southeastern Newfoundland (1967); northern Russia, the Ukraine, northern Siberia, and the Ural Mountains (1960s–1990s); North Carolina (1966); northwestern Canada (1979); southern Nevada (1997); and elsewhere. About 30 localities have been described, with the most diverse biotas found at Ediacara, Namibia, and on the coast of the White Sea in northern Russia. As some of these sites could not have been less than about 6000 mi (10,000 km) apart, no matter how the continents were arranged at the time, there is no doubt that the Ediacaran biota was a globally distributed marine biota. Ediacaran fossils have also been found in both shallow-water and deep-water facies. The biota appears to show a degree of biogeographic provinciality, but it has proved difficult to untangle the effects of geographic separation, paleoenvironmental differences, age differences, and sampling and preservational biases on the composition and preservation of the biota at each major site. A number of taxa are found in rocks formed under widely different conditions, and these, at least, may have been ecological generalists.

Age of assemblage. Recent radiometric dating has bracketed the Ediacaran biota between 600 and 543 Ma, with most sites between 565 and 543 Ma. Most known occurrences of Ediacaran organisms precede the earliest great radiation of skeletal fossils (archaeocyaths, trilobites, mollusks, brachiopods), or else they can be placed as latest Proterozoic on other evidence. Although unfossiliferous strata separate the Ediacaran biota from the earliest Cambrian fossils at many localities, Ediacaran biotas in Namibia and Nevada are found in rock sequences extending right up to the base of the Cambrian. Radiometric dates have confirmed that there was no significant time gap between the disappearance of the Ediacaran organisms and the Cambrian radiations. In fact, a few Ediacaran fossils have been found in Cambrian strata; the biota did not entirely die out before the Cambrian. *See* CAMBRIAN.

Fossil preservation. At almost all sites, the fossils are preserved as impressions in some kind of detrital sedimentary rock. Commonly, they are found on the bases of sandstone beds (South Australia), within sandstone beds (Namibia), or below volcanic ashes swept into deep water by wind and turbidity currents (Newfoundland). The organisms appear to have lived in continental shelf to slope environments and are normally preserved in sediments that were deposited under fairly quiet conditions below normal wave base. Organisms with resistant bodies are usually preserved as concave impressions in the casting medium, but other fossils appear to be the sediment fillings of either soft bodies or the cavities left by their decay. Most strata where Ediacaran fossils have been found also show evidence of extensive microbial growth that formed firm mats on the sediment surface. By stabilizing and firming the sediment, and possibly by speeding up its lithification, these mats are now thought to have been crucial to the fossilization of these soft-bodied organisms.

Taxonomic diversity. More than a hundred scientific names have been given to various kinds of Ediacaran fossils, but many of these are based on questionable material, or they are unnecessary synonyms for the same genus or species preserved in different ways. Nevertheless, at a conservative estimate there are probably 40–50 distinct genera or probable genera of Ediacaran organisms worldwide. Assigning these genera to higher taxa, however, has been controversial. Prior to 1985, many scientists who worked on the Ediacaran fossils attempted to relate them to the phyla of living invertebrates. A real problem is that few of these fossils can unequivocally be referred to living or extinct animal taxa. Because many of the fossils are simply circular structures with or without radial or concentric markings, they impart little information and are difficult to interpret. Some of these circular "medusoids" could be the remains of cnidarian jellyfish (Hydrozoa or Scyphozoa), but very few show diagnostic cnidarian characters such as fourfold symmetry, or even diagnostic metazoan characters such as definite guts. The recognition of many critical features is hampered by the nature of the preservation. In addition, many Ediacaran fossils, notably most forms from Namibia, are so unusual in shape that they cannot be placed firmly in any modern group. *See* CNIDARIA.

Many workers have placed the more unusual organisms in an extinct higher taxon of phylum grade, commonly named the Petalonamae. Others have used differences in symmetry and body organization to identify and characterize several major taxonomic groups regarded, in general, as extinct higher taxa of phylum or class grade. Beginning in the mid-1980s, much controversy and interest were sparked by proposals that all of the Ediacaran organisms were constructed on a single basic plan that was radically different from any other animal. According to this hypothesis, the Ediacarans were "quilted" organisms lacking heads, muscles, or digestive systems, and consisted of parallel sheetlike walls held together by regularly spaced internal partitions, and the whole organism was inflated by body fluids. Without digestive systems, these organisms would have lived by taking up dissolved nutrients or perhaps by photosynthesis. Proponents of this hypothesis classify the Ediacaran organisms in an extinct kingdom of multicellular life, the Vendobionta. Still others have proposed that the Ediacaran organisms belonged to extant kingdoms other than the animals; they were lichens, algae, or single-celled or colonial protists.

Controversy still reigns, but in all likelihood no hypothesis is entirely incorrect. Ediacaran fossils have been described which do show clear features diagnostic of living animal phyla. However, many of the Ediacaran organisms remain impossible to definitively link with living animals, and they may well represent extinct phyla or even kingdoms.

Distinctive genera such as the discoidal "worm" *Dickinsonia* (**Fig. 1**), the sac-shaped *Ernietta*, the paddle-shaped "sea pen" *Charnia*, and the trifoliate *Pteridinium* seem to form a coherent group of organisms with a Vendobionta-like "quilted" or "air-mattress" type of construction. Whether this common structural pattern is due to convergent evolution in response to particular environmental conditions remains uncertain, and so the Petalonamae or even the Vendobionta may be an unnatural group of unrelated taxa. There is, for example, evidence that "sea pens" such as *Charniodiscus* and *Rangea* (**Fig. 2**) lived attached to the ocean floor by

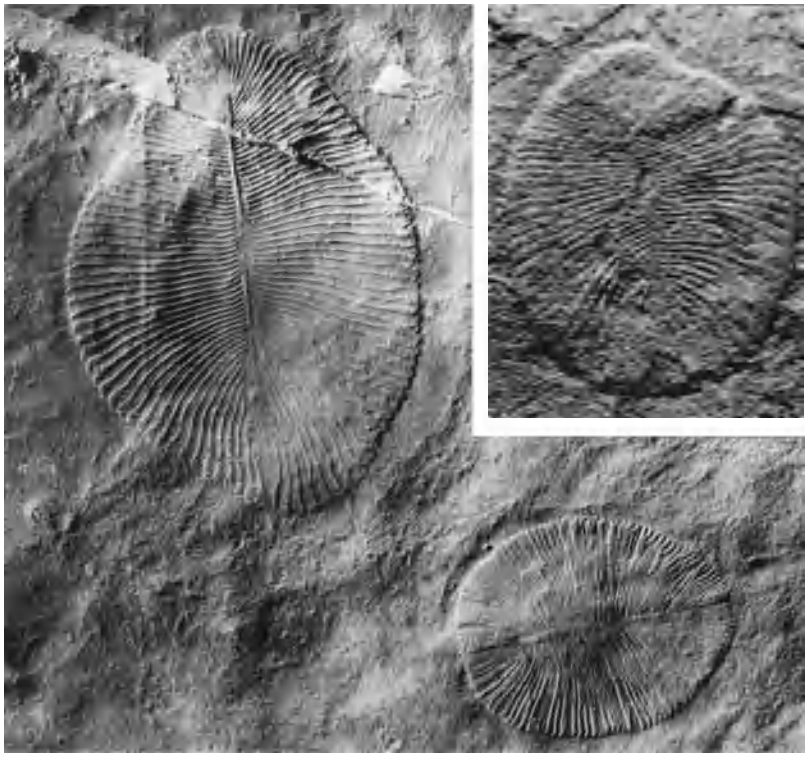

Fig. 1. Natural impressions of *Dickinsonia* from Australia and from Russia (inset). The Australian specimens, 5 and 3 in. (13 and 8 cm) in length, have about the same number of segments and may be regarded as the contracted and expanded states of a single individual.

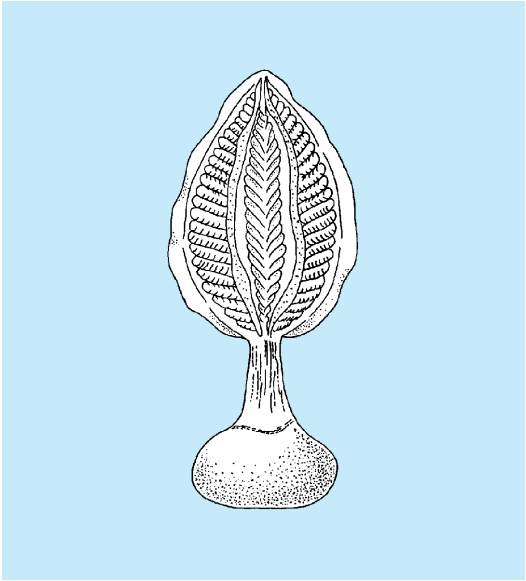

Fig. 2. Reconstruction of *Rangea*, a complex, sessile genus of "sea pens" from Namibia.

globular holdfasts, whereas *Ernietta* lived partially buried in sand with no holdfast. Unnamed spindle-shaped "sea pens" from Newfoundland also show no trace of an attachment structure. *Dickinsonia* seems to have been free-living and mobile and may have had a gut.

The members of a second group of Ediacaran animals are unified by their threefold symmetry. All have a circular outline and low conical shape reminiscent of modern jellyfish, and they have been regarded as an extinct class of the Cnidaria by some researchers. However, the morphological complexity of genera such as *Tribrachidium* and *Albumares* argues against such a view. It is plausible that the triradiate forms are members of a second extinct phylum, the Trilobozoa. True free-living cnidarians may be represented by "medusoids" such as *Ediacaria* and *Eoporpita*, which are now thought to be benthic polyplike organisms. A few forms, such as the tetraradial *Conomedusites* and the chondrophorine-like *Ovatoscutum* (**Fig. 3**), may have been planktonic cnidarians. The Australian *Palaeophragmodictyon* shows impressions of a spicule network and is almost certainly a sponge.

Although several higher invertebrate phyla such as the Annelida (segmented worms), Arthropoda, and Echinodermata have been identified from the Ediacaran biota, only a few of the known genera could be used to make a plausible case for the presence of these phyla. *Spriggina floundersi*, known from tens of specimens from Ediacara and nearby sites, had a long, bilaterally symmetrical, segmented body with a boomerang-shaped feature at one end. *Spriggina* has been placed with the annelids, but it is equally likely that it was a primitive soft-shelled arthropod. In this interpretation, the boomerang-shaped structure is considered to be homologous with the arthropod head. A few *Spriggina*-like fossils show traces of branching caecae within this head, closely resembling the anatomy of trilobites and other arthropods from Cambrian sites such as the Burgess Shale. Yet another fossil, *Kimberella quadrata*, appears to have borne a stiff shell on its dorsal side, which covered a soft muscular "foot" surrounded by numerous gill-like folds. *Kimberella* is likely to be closely related to the Mollusca, although it may not be closely related to any extant molluscan class.

Fig. 3. Discoidal fossil *Ovatoscutum* from South Australia may be a link between the Ediacaran fauna and younger biotas. It is probably an impression of the float of a chondrophorine cnidarian.

Fig. 4. Mold of *Cloudina* or a related genus, an annulated tubular fossil that was mineralized in life. *Cloudina* has been found together with soft-bodied Ediacaran fossils at several localities; this one is from Nevada.

Finally, a few organisms associated with the Ediacaran biota formed hard or even biomineralized parts or skeletons. Biomineralization did not absolutely originate in the Early Cambrian. The calcified tube *Cloudina* (**Fig. 4**) and the agglutinated tube *Archaeichnium* may have been formed by worms. Mineralized sponge spicules have been found in Ediacaran-age rocks in Mongolia. Other mineralized fossils are more difficult to decipher and include tubular, cup-shaped, conical, and hexagonal forms. Organic-walled tubes include the threadlike sabelliditids, possibly made by annelid or pogonophoran relatives; and the tetraradial *Corumbella*, which resembles the tubes made by certain scyphozoan polyps. Small, unmineralized toothlike fossils (*Redkinia*), which somewhat resemble annelid jaws, are known from Ediacaran assemblages in Russia. *See* ANNELIDA; ARTHROPODA; BURGESS SHALE; ECHINODERMATA.

Associated trace fossils. The marks left in soft sediments by otherwise unknown animals provide another source of knowledge of late Precambrian animal life. The figure-8-shaped trail in **Fig. 5** indicates that animals capable of directed, muscular, gliding motion (like that of a garden snail) coexisted with more typical members of the Ediacaran biota. In a similar fashion, strings of fecal pellets demonstrate the existence of animals with one-way guts, and closely meandering marks imply an ability for

Fig. 5. Ediacaran trace fossil *Gordia* found on the sole of a sandstone bed in South Australia gives clear evidence for the existence of mobile animals in the late Precambrian.

systematic grazing. There is little evidence for vertical burrowing in rocks of this age.

Significance. Although many Ediacaran fossils are enigmatic, there is sound evidence that sponges, cnidarians, bilaterian worms, and possibly arthropods and other phyla were present. This implies that the Animalia originated even farther back in time. The largest Ediacaran fossils, reaching up to 3 ft (1 m) in length, are flattened "fronds" that had large surfaces compared with their volumes. Proponents of the Vendobionta hypothesis have claimed that such large organisms, lacking guts or muscles, must have been photosynthetic or chemosynthetic and probably contained symbiotic, photosynthetic microorganisms. However, such a lifestyle is also found in modern reef corals and various other marine animals. Alternatively, the high ratio of surface to volume may represent adaptations to an atmosphere and hydrosphere relatively low in oxygen. *See* PALEONTOLOGY.

Bruce Runnegar; Ben Waggoner

Bibliography. J. G. Gehling, Microbial mats in terminal Proterozoic siliciclastics: Ediacaran death masks, *Palaios*, 14:40–57, 1999; J. P. Grotzinger et al., Bio-stratigraphic and geochronologic constraints on early animal evolution, *Science*, 270:598–604, 1995; G. M. Narbonne, The Ediacara biota: A terminal Neoproterozoic experiment in the evolution of life, *GSA Today*, 8(2):1–6, 1998; B. Runnegar, Vendobionta or Metazoa? Developments in understanding the Ediacara "fauna," *N. Jb. Geol. Paläontol. Abh.*, 195:303–318, 1995; A. Seilacher, Vendobionta and Psammocorallia: Lost constructions of Precambrian evolution, *J. Geol. Soc. London*, 149:607–613, 1992; R. C. Sprigg, Early Cambrian (?) jellyfishes from the Flinders Ranges, South Australia, *Trans. Roy. Soc. S. Aust.*, 71:212–224, 1947.

Eel

A fish in the order Anguilliformes, one of several orders of the superorder Elopomorpha, which share a leptocephalus (a slender, transparent larva with a long narrow head) [**Fig. 1**] larval development. The name eel also applies to other fishes similar in appearance to true eels, that is, having a serpentine body, lacking pectoral fins, and usually lacking scales. *See* ANGUILLIFORMES.

True Eels

The Anguilliformes consist of suborders Anguilloidei, Muraenoidei, and Congroidei.

Anguilloidei. This suborder comprises three families.

Freshwater eels (family Anguillidae). These eels (**Fig. 2**) are found in tropical and temperate seas except for the eastern Pacific and southern Atlantic; they consist of one genus and 15 species. They are catadromous (inhabiting freshwater and migrating to spawn in salt water), living practically all of their adult life in freshwater.

Mud eels (family Heterenchelyidae). These eels inhabit the tropical Atlantic, Mediterranean, and eastern Pacific; they consist of two genera and eight species. They typically burrow head-first into the substrate.

Spaghetti eels (family Moringuidae). These eels inhabit the tropical Indo-Pacific and western Atlantic; they consist of two genera and six species. Their bodies are extremely elongate and adapted for burrowing head-first into the substrate.

Muraenoidei. This suborder comprises three families.

False morays [family Chlopsidae (Xenocongridae)]. These eels inhabit the tropical and subtropical Atlantic, Indian, and Pacific oceans; they consist of eight genera and 18 species. Their gill openings are restricted to small rounded, laterally placed apertures; pectoral fins are absent in two of the eight genera; and lateral-line pores are restricted to the head.

Myroconger eels (family Myrocongridae). These eels inhabit the eastern tropical Atlantic and Pacific oceans; they consist of one genus and four species. Their gill

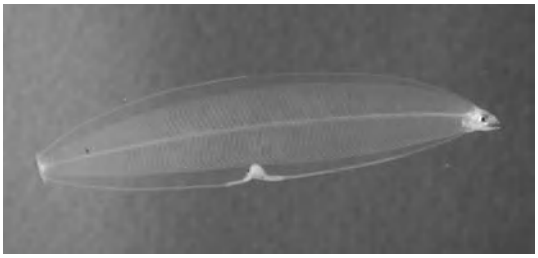

Fig. 1. Leptocephalus larva. (*Photo by Antonio J. Ferreira, © California Academy of Sciences*)

Fig. 2. *Anguilla anguilla*. (*Drawing by Robbie Cada, www.FishBase.org*)

openings are small but not greatly restricted; pectoral fins are present; their bodies are strongly compressed; and lateral-line pores are located in the branchial (pertaining to the gills) region above the pectoral fins.

Moray eels (family Muraenidae). These eels inhabit tropical and temperate seas of the world, with some species entering freshwater; they consist of 15 genera and about 185 species. Their gill openings are restricted to small rounded, laterally placed apertures; the pectoral fins are absent; and lateral-line pores are located on the head and branchial region. The giant moray (*Gymnothorax javanicus*) of the Indo-Pacific attains a length of 3 m (10 ft), making it one of the largest eels.

Congroidei. This suborder comprises nine families.

Cutthroat eels (family Synaphobranchidae). These eels inhabit the Atlantic, Indian, and Pacific oceans; they consist of 10 genera and 32 species. The gill openings are low, at or below the insertion of the pectoral fins; hence, the name cutthroat eels.

Snake eels and worm eels (family Ophichthidae). These eels inhabit tropical and warm temperate seas of the world, with some species in freshwater; they consist of 52 genera and about 290 species. The family differs from other eels in the unusual form of the branchiostegal (pertaining to the membrane covering the gills) rays, which broadly overlap those of the other side, leaving the branchiostegal region inflated or slightly bulbous. *Ophichthus rex*, known only in the Gulf of Mexico from Florida to Texas, and *O. ophis*, from the tropical and subtemperate western and eastern Atlantic, both with total lengths of about 210 cm (83 in.), are the largest ophichthids.

Shorttail eels (family Colocongridae). These eels inhabit the Atlantic, Indian, and western Pacific oceans; they consist of one genus and five species. Shorttail eels, with a blunt head and stubby body, are the least elongate of all the true eels.

Longneck eels (family Derichthyidae). These eels inhabit the Atlantic, Indian, and Pacific oceans; they consist of two genera and three species. These are mesopelagic [inhabiting the middle ocean region, from about 200 to 1000 m (650 to 3300 ft)] to bathypelagic [inhabiting the deep ocean region, from about 1000 to 4000 m (3300 to 13,000 ft)] eels, with a maximum length of about 60 cm (24 in.).

Pike congers (family Muraenesocidae). These eels inhabit the tropical Atlantic, Indian, and Pacific oceans; they consist of possibly six genera and 13 species. Pike congers have a large mouth equipped with large conspicuous teeth; their large eyes are covered with skin; and they have a conspicuous lateral line. Five species attain a total length of 200 cm (78 in.) or more, with the largest known being 250 cm (98 in.).

Snipe eels (family Nemichthyidae). These eels are mesopelagic in the tropical and temperate Atlantic, Indian, and Pacific oceans; they consist of three genera and nine species. Snipe eels are the most bizarre anguilliforms. Their bodies are extremely long and ribbonlike; one genus, *Nemichthys*, has over 750 vertebrae. The jaws are very long and narrow, and the teeth cannot be occluded (that is, the teeth do not come into contact with cusps of the opposing teeth fitting together), except in sexually mature males, which undergo an extreme shortening of the jaws and loss of teeth. The two sexes are so different that at one time some species of snipe eels were thought to be different taxa, even at the suborder level.

Conger eels (family Congridae). These eels inhabit the tropical to temperate Atlantic, Indian, and Pacific oceans; they consist of 32 genera and 160 species. Most congrid eels are dull colored, usually tan or brown on the back and silvery on the sides and belly, with only a few having spots or stripes; the lateral line is complete; the dorsal and anal fins are well developed and continuous with the caudal fin; and the pectoral fins are usually well developed, except in garden eels, in which the pectorals are reduced to a tiny fleshy flap. Garden eels live in colonies, where they partly submerge themselves in tubelike burrows. With their head bent facing the current, they sway back and forth picking plankton as it drifts by. Congrid eels vary greatly in size—the smallest, *Bathycongrus odontostomus*, of the Indo-West Pacific, is only 4.5 cm (1.8 in.) in total length, whereas *Conger conger* of the eastern Atlantic attains a total length of 300 cm (118 in.). The longest true eel in American waters is *C. oceanicus*, with a total length of 230 cm (90 in.).

Duckbill eels (family Nettastomatidae). These eels inhabit the Atlantic, Indian, and Pacific oceans; they consist of six genera and about 38 species. They have elongate heads, large mouths, and flattened snouts—hence the name duckbill; their tails are attenuated; and pectoral fins are usually absent in adults. Their maximum length is about 100 cm (39 in.).

Sawtooth eels (family Serrivomeridae). These eels are pelagic (living in the open ocean) in the tropical and temperate Atlantic, Indian, and Pacific oceans; they consist of two genera and about 10 species. Their jaws are extremely elongate and slender, but less so than in snipe eels; their gill openings are connected ventrally; the tails are attenuated; and the vomerine teeth (that is, teeth on the vomer, which is the median unpaired bone on the roof of the mouth) are lancet-like and arranged in a sawlike row. Their maximum total length is about 78 cm (31 in.).

Other "Eels"

Other "eels" are classified in orders Saccopharyngiformes, Gymnotiformes, and Siluriformes.

Order Saccopharyngiformes. This order comprises two families.

Bobtail snipe eel (family Cyematidae). These fish are bathypelagic in the Atlantic, Indian, and Pacific oceans; they consist of two monotypic genera. Their bodies are relatively short and compressed; they lack lateral-line pores; their eyes are small to vestigial; the tip of the tail is blunt; the caudal fin is present; and maxillae are present. The maximum length is about 15 cm (6 in.).

Pelican eels (family Eurypharyngidae). These eels inhabit the tropical and temperate Atlantic, Indian, and Pacific oceans; they consist of one species. They have a huge head, a pelicanlike mouth, jaws with numerous minute teeth, and minute pectoral fins. This fish

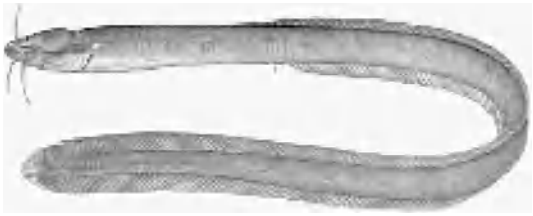

Fig. 3. Eel catfish (*Channallabes apus*). (*Drawing by P. Mertens, www.FishBase.org*)

is unique among teleosts in having five gill arches. The maximum length is about 74 cm (29 in.).

Order Gymnotiformes. This order comprises a single family.

Electric eel (family Gymnotidae). Another true-eel mimic is the electric eel (*Electrophorus electricus*), a member of the family Gymnotidae (nakedback knifefishes). It occurs in the Orinoco and Amazon rivers of northern South America and is unique among the gymnotiforms in having electric organs capable of delivering discharges up to 600 volts. The electric eel is further identified by the lack of a dorsal fin, while displaying a very long anal fin terminating at the tip of the tail. It has an oral respiratory organ for breathing air and a rounded body that lacks scales, and can continuously add vertebrae throughout life. It attains a total length of 220 cm (87 in.). *See* ELECTRIC ORGAN (BIOLOGY).

Order Siluriformes. A catfish that could easily be mistaken for an eel inhabits the Congo River Basin of Africa. It is called the eel catfish (*Channallabes apus*) [**Fig. 3**] and is the only one among about 2870 species of catfishes that bears the name eel. The eel catfish, belonging to the siluriform family Clariidae (air-breathing catfishes), has reduced eyes, lacks pectoral and pelvis fins, has long dorsal and anal fins, and has the habit of burrowing in the substrate, all characteristics of many species of true eels. *See* SILURIFORMES.

Herbert Boschung

Bibliography. R. Froese and D. Pauly (eds.), FishBase, World Wide Web electronic publication, version 05/2005; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Effective dose 50

This term is used chiefly to characterize the potency of a drug by the amount required to produce a response in 50% of the subjects to whom the drug is given. The term is also known as ED_{50} or median effective dose. At one time it was usual to try to measure the effect of a drug by noting the amount which was just sufficient to produce a particular response; but when it was realized that this amount varied greatly from subject to subject, attention was turned to measuring the effect on a group of subjects. Suppose, for example, that a drug is being used to relieve a certain type of pain; then the median effective dose is of such a size that it controls the pain in 50% of the sufferers and is insufficient to control it in the remaining 50%.

The term median effective dose is most commonly applied in connection with drugs, but it may be used when various other sources of stimuli, for example, x-rays, are under consideration. The response must be of the kind known as quantal, or all or nothing, where the investigator is simply able to report that the response either was or was not elicited; for example, convulsions did or did not occur, hemorrhage was or was not produced, pregnancy did or did not ensue, the animals did or did not survive.

The median effective dose is not a well-defined quantity until the test animal, the end response, and such factors as the route of injection of a drug and the state of nutrition of the animal are specified. In the work of an investigator who uses the appropriate controlled conditions, however, the ED_{50} is a reproducible measurement.

Determination of the ED_{50} of a drug requires the administration of at least two separate amounts of the drug, each one given to several subjects. Suppose that an animal physiologist wishes to measure the ED_{50} of a hormone, estrone, that causes estrus, or heat. He has specified that he will use spayed female rats as experimental animals, and has also laid down various conditions that he regards as important in controlling the results of his investigation: the age of the animals, perhaps, and the route by which the hormone will be administered. It is unlikely that a reliable measurement could be made with fewer than 30 animals. Suppose that a total of 60 were used, 20 at each of 3 doses, with the following results:

Dose, international units	*Animals showing estrus, %*
2	20
3	45
4	80

Inspection of the results suggests that a dose somewhat greater than 3 units will cause estrus in 50% of the animals. Methods based on the theory of probability can be used to give a more objective analysis of the data and to provide a measure of the error of the estimate, but it is sufficient to note for the present purpose that the ED_{50} is approximately 3. It is clear that in measuring the ED_{50} some choices of the dosage to be used will work out better than others. Doses giving a response in a very small or very large percentage of the subjects contribute little information to the measurement of ED_{50}. The ideal is to have doses on either side of the ED_{50}, and the closer they are to ED_{50}, the better. However, in practice a pharmacologist sometimes determines an ED_{50} as one part of a larger experiment to investigate the dose response curve of a drug, or to compare the potency of two preparations of a drug. In such a case, though he would still want to have doses on each side of the ED_{50}, he would not want them close to the ED_{50}. More information about the total curve would be given if the low dose had an effect in, say, 25% of the animals and the high dose in about 75%. *See* BIOASSAY.

Colin White

Bibliography. C. Janney and J. Timpke, *Calculation of Drug Dosages*, 4th ed., 1993; R. Wiederhold, *Dosages and Calculations*, 1991.

Effector systems

Those organ systems of the animal body which mediate overt behavior. Injury to an effector system leads to loss or to subnormal execution of behavior patterns mediated by the system, conditions termed paralysis and paresis, respectively.

Overt behavior consists of either movement or secretion. Movement results from contraction of muscle. Secretion is a function of glands. Neither muscular contraction nor glandular secretion is autonomous but is regulated by an activating mechanism which may be either neural or humoral. In neurally activated systems the effector organ, whether muscle or gland, is supplied by nerve fibers originating from cell bodies situated in the central nervous system or in peripherally located aggregates of nerve cell bodies known as ganglia. The nerve fibers make intimate contact with, but are not protoplasmically continuous with, the cells of the effector organ. Activation of the effector organ occurs when the nerve cell body is excited and generates a nerve impulse, an electrochemical alteration which is conducted along the nerve fiber to its terminations on the effector organ cells. Here the nerve impulse releases from the nerve ending a chemical transmitter that generates a similar electrochemical alteration in the effector cells, and this alteration, in turn, leads to either contraction or secretion. Such systems composed of a muscle or a gland along with their regulating nerves are termed neuromuscular or neuroglandular effector systems, respectively. Examples are the skeletal muscles together with their motor nerve supplies, and the medullae of the adrenal glands along with their innervation (splanchnic nerves). In many neurally regulated effector systems (for example, skeletal muscle and adrenal medulla), function is totally dependent on intact innervation, and denervation leads to functional paralysis. In other organs (for example, the salivary glands), denervation causes only temporary paralysis. When recovery occurs, the gland may oversecrete continuously (paralytic secretion), apparently because the denervated gland cells become unusually sensitive to certain blood-borne chemical agents (denervation supersensitivity). *See* BIOPOTENTIALS AND IONIC CURRENTS; ENDOCRINE MECHANISMS; MOTOR SYSTEMS.

In other effector systems (humeromuscular and humeroglandular), the activating agent is normally a blood-borne chemical substance produced in an organ distant from the effector organ. Uterine smooth muscle is uninfluenced by the uterine nerve activity but contracts vigorously when the blood contains oxytocin, a chemical substance elaborated by the posterior lobe of the hypophysis; sensitivity to oxytocin increases progressively during pregnancy. Similarly, secretion of pancreatic juice is independent of pancreatic innervation; the regulating agents are blood-borne substances (cholecystekinin and secretin) produced by cells in the wall of the small intestine. Generally, in such humorally regulated effector systems, activation is more delayed and more prolonged than in neurally regulated systems. *See* PANCREAS; UTERUS.

Finally, some effector systems are hybrid in the sense that both nerves and humors regulate their functions. The smooth muscles of arterioles contracts in response to either nerve stimulation or epinephrine, a substance secreted into the bloodstream by the adrenal medulla. Secretion of hydrochloric acid by the gastric mucosa is increased by activation of the vagus nerve or by the presence in the blood of histamine, a substance found in many tissues of the body. Effector systems with both neural and humoral regulation are never completely paralyzed by denervation but may be deficient in reaction patterns when the quick integrated activation provided by neural regulation is essential. For example, following extensive vascular denervation, humoral agents may maintain sufficient arteriolar constricton to sustain the blood pressure in static situations. However, the normal capacity to increase arteriolar constriction to offset the gravitational effects of rising from the prone to standing positions is permanently lost, with the result that such postural changes may lead to inadequate cerebral blood flow and consequent fainting. *See* HISTAMINE; NERVOUS SYSTEM (VERTEBRATE); PSYCHOLOGY. Theodore C. Ruch; Harry D. Patton

Efficiency

The ratio, expressed as a percentage, of the output to the input of power (energy or work per unit time). As is common in engineering, this concept is defined precisely and made measurable. Thus, a gear transmission is 97% efficient when the useful energy output is 97% of the input, the other 3% being lost as heat due to friction. A boiler is 75% efficient when its product (steam) contains 75% of the heat theoretically contained in the fuel consumed. All automobile engines have low efficiency (below 30%) because of the total energy content of fuel converted to heat; only a portion provides motive power, while a substantial amount is lost in radiator and car exhaust. *See* AUTOMOTIVE ENGINE; BOILER; POWER.

In such simple cases the value is clear. However, in some others it can be difficult to calculate exactly. For example, the efficiency of the process for converting corn to alcohol for use as automobile fuel can be computed as the ratio of the heat value of the alcohol to the heat value of all the energy used to produce it, including the fuel for the tractor to plow and harvest the cornfield and even the fuel used to create the steel and fabricate the tractor. The question is then, how much to include in the overall efficiency determination. *See* SIMPLE MACHINE. F. R. E. Crossley

Efflorescence

The spontaneous loss of water (as vapor) from hydrated crystalline solids. The thermodynamic requirement for efflorescence is that the partial

pressure of water vapor at the surface of the solid (its dissociation pressure) exceed the partial pressure of water vapor in the air.

A typical efflorescent substance is Glauber's salt, $Na_2SO_4 \cdot 10H_2O$. At 25°C (77°F) the dissociation pressure for the process in reaction (1) is 2586.4 pas-

$$Na_2SO_4 \cdot 10H_2O(s) \rightarrow Na_2SO_4(s) + 10H_2O(g) \quad (1)$$

cals (0.776 in. Hg), 81% of the saturation vapor pressure of pure water at this temperature. In a sufficiently humid atmosphere Glauber's salt also can deliquesce by the process shown in reaction (2).

$$Na_2SO_4 \cdot 10H_2O(s) + H_2O(g) \rightarrow Na_2SO_4 \text{ (saturated aqueous solution)} \quad (2)$$

The vapor pressure of the saturated solution is 2919.27 Pa (0.876 in. Hg), 92% of the vapor pressure of pure water. Thus, Glauber's salt at 25°C (77°F) is stable in atmospheres having relative humidities of 81–92%; below 81% it effloresces; above 92% it deliquesces.

The spontaneous loss of water normally requires that the crystal structure be rearranged, and consequently, efflorescent salts usually go to microcrystalline powders when they lose their water of hydration. *See* DELIQUESCENCE; PHASE EQUILIBRIUM; VAPOR PRESSURE. Robert L. Scott

Egg (fowl)

A single, large, living, female sex cell enclosed in a porous, calcareous shell through which gases may pass. Although they vary in size, shape, and color, the eggs of chickens, ducks, geese, and turkeys are essentially the same in structure and content (see **illus.**). Inward from the shell are the outer and inner shell membranes which are also permeable to gases. The membranes are constructed to prevent rapid evaporation of moisture from the egg but to allow free entry of oxygen, which is necessary for life. Air begins to penetrate the shell soon after the egg is laid, and it tends to accumulate in a space between the two membranes at the large end of the egg. *See* CELL (BIOLOGY).

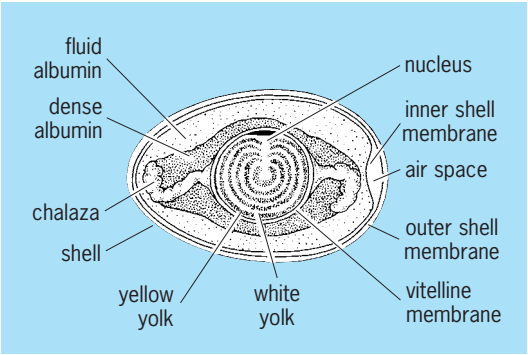

Egg of a bird. (*After L. P. Sayles, ed., Biology of the Vertebrates, 3d ed., Macmillan, 1949*)

The inner shell membrane surrounds a mass of fluid albumin which, in turn, encloses a body of dense albumin; these two types of protoplasm constitute the so-called egg white. The central part of the egg is occupied by the yolk, which contains the vital egg nucleus and its associated parts. The yolk consists of alternating layers of yellow and white yolk. The yolk, enclosed by the vitelline membrane, is held in place by the chalaza which is anchored at each end of the egg and prevents undue mechanical disturbance. *See* CELL NUCLEUS; YOLK SAC.

When a sperm nucleus fuses with the egg nucleus, the process is called fertilization. Within a few hours or less, the fertilized egg begins a series of cell divisions and differentiations which result in the formation of the embryo. The embryo then undergoes further cell modification and eventually develops into the young of the species. In time the pressure created by the growth of the embryo causes the shell to rupture, and the young hatches. As the young emerges from the shell, it carries with it a part of the food and water originally in the yolk on which it can subsist for a few days. After this initial period, the young must have access to food and water. *See* CELL DIVISION; FERTILIZATION (ANIMAL); MITOSIS.

For a discussion of the steps involved in the development of various kinds of fertilized eggs into mature embryos *see* EMBRYOLOGY; OVUM; REPRODUCTION (ANIMAL); POULTRY PRODUCTION. James F. Ferry

Egg water balance. The water vapor content of the environment is usually less than that found inside eggs, so water vapor is lost continually throughout incubation. Eggshell conductance, that is, the inverse resistance of the shell to gas movement, is directly related to egg mass, and inversely related to incubation time. The relationship between eggshell conductance and daily water loss, egg mass, and incubation time predicts that all bird eggs should lose about 15% of their initial mass as water vapor during incubation and that the vapor pressure gradient across the shell is similar for all bird species and relatively constant throughout incubation. The fractional water loss is apparently obligatory since eggs prevented from losing water show increased embryonic mortality and deformity. The obligatory nature of egg water loss may be related to the observation that the embryo inflates its lung just prior to hatching with the gas in the air space of the egg. The volume of the air space is related to the quantity of water loss since the egg shell is rigid.

Respiratory gas exchange. While the daily water loss from avian eggs during incubation appears to be relatively constant, such is not the case for the daily oxygen consumption (O_2 flux) and carbon dioxide production (CO_2 flux). The metabolic activity increases throughout incubation as the embryo develops and increases in mass, so both the O_2 and CO_2 exchange must similarly increase. Since the gas conductance of the shell is constant (all gases diffuse through the same pores), the pressure of O_2 inside the egg must fall and the pressure of CO_2 must rise if the embryo is to gain O_2 and lose metabolically

produced CO_2. The egg O_2 exchange of a variety of bird species measured just prior to hatching appears to increase in proportion to egg mass in much the same way that the eggshell gas conductance (as measured with water vapor) increases with increase in egg mass. These relationships allow one to predict that the gas pressures of O_2 and CO_2 (in the air space) are the same for all bird embryos just prior to hatching and are in fact similar to values one might expect to see in adult bird lungs. It would appear that one function of the eggshell is to prepare the embryo for transition to lung breathing at birth. *See* RESPIRATION.

Ralph A. Ackerman

Bibliography. R. W. Burley and D. V. Vadehra, *The Avian Egg: Chemistry and Biology*, 1989; B. M. Carlson, *Patten's Foundations of Embryology*, 6th ed., 2000.

Eggplant

A warm-season vegetable (*Solanum melongena*) of Asiatic orgin belonging to the plant order Polemoniales (formerly Tubiflorales). Eggplant is grown for its usually egg-shaped flesh fruit (see **illus.**). and is

Eggplant (*Solanum melongena*), cultivar Black Magic. (*Joseph Harris Co., Rochester, New York*)

eaten as a cooked vegetable. Cultural practices are similar to those used for tomatoes and peppers; however, eggplant is more sensitive to low temperatures. Popular purple-fruited varieties (cultivars) are Black Beauty and a number of hybrid varieties; fruits of other colors, including white, brown, yellow, and green, are used chiefly for ornamental purposes. Harvesting generally begins 70–80 days after planting. Florida and New Jersey are important eggplant-producing states. *See* PEPPER; SOLANALES; TOMATO.

H. John Carew

Ehrlichiosis

Ehrlichiosis is a tick-borne infection that often is asymptomatic but also can produce an illness ranging from a few mild symptoms to an overwhelming multisystem disease. Ehrlichiosis is included with those infections that are said to be emerging, either because they have been recognized only recently or because they were previously well known but now are occurring more frequently.

History. Prior to 1986, the bacteria of the genus *Ehrlichia* were of interest primarily to veterinarians. These microorganisms have a global distribution and are capable of producing a hemorrhagic disease in dogs and febrile illnesses of horses, cattle, sheep, and bison. A constant feature in all these circumstances is the predilection of the *Ehrlichia* to enter the cellular elements of the blood. The specific *Ehrlichia* species determined whether mononuclear cells, granulocytes, or platelets were invaded. Human illness associated with these organisms was first recognized in southern Japan in the 1950s. Because these illnesses seemed to be a local phenomenon, they received little attention elsewhere.

However, in 1986 a 51-year-old man, after planting some trees while vacationing in Arkansas, discovered some ticks on his neck and removed them uneventfully. Ten days later, he developed fever accompanied by a serious systemic illness that eluded specific diagnosis and required a prolonged hospitalization before his recovery. Careful microscopic examination of a slide preparation of his peripheral blood revealed distinctive inclusion bodies in his lymphocytes and monocytes. When reviewed by microbiologists and pathologists, these inclusions were identified as ehrlichiae. This illness was the first case of ehrlichiosis diagnosed in the United States. To date, over 400 cases of ehrlichiosis have been diagnosed in 30 states, and ehrlichiosis now is considered to be the most common tick-borne infection in this country.

Epidemiology. Human ehrlichiosis is caused by two distinct species: *E. chaffeensis* and an unnamed ehrlichial species. *Ehrlichia chaffeensis* infects primarily mononuclear blood cells; the disease produced by this species is referred to as human monocytic ehrlichiosis. The other ehrlichial species invades granulocytic blood cells, causing human granulocytic ehrlichiosis. The latter organism closely resembles *E. equi*, a species that infects horses.

Both of these ehrlichia species are transmitted to humans by the bite of infected ticks; however, the two species appear to have somewhat different geographic distributions. *Ehrlichia chaffeensis* occurs most commonly in the south-central and southeastern states, where it is associated primarily with the Lone Star tick (*Amblyomma americanum*); it is also transmitted by the common dog tick (*Dermacentor variabilis*). The agent of human granulocytic ehrlichiosis is found in the upper midwestern states of Wisconsin and Minnesota, as well as in several northeastern states. This agent seems to be transmitted principally by the deer tick (*Ixodes scapularis*).

Although ticks (the vector) are the mode of transmission of ehrlichial infections to humans, the ticks must acquire the ehrlichial organisms from animal sources (the reservoir hosts). The animal reservoirs for these organisms have not yet been confirmed, but white-tailed deer are a likely reservoir host of *E. chaffeensis* in the southeastern United States. *See* IXODIDES.

Clinical signs and pathogenesis. The forms of the disease caused by the two ehrlichial species are indistinguishable. Illness occurs most often during April–September, corresponding to the period when ticks are most active and humans are pursuing outdoor activities. Most individuals relate a history of tick exposure and have been engaged in work or recreation in wooded areas where they encountered ticks.

While obtaining its blood meal by biting an individual, the infected tick inoculates ehrlichiae into the skin. The microorganisms then spread throughout the body via the bloodstream. *Ehrlichia* must achieve an intracellular residence in order to survive and multiply. *Ehrlichia chaffeensis* seeks out tissue macrophages in the spleen, liver, and lymph nodes. It is suspected that the agent of human granulocytic ehrlichiosis has a tropism for granulocytic precursor cells in the bone marrow. Once the organism has entered the cell, it multiplies there for a period of time without producing symptoms (the incubation period). In ehrlichiosis, the incubation period can last from 1 to 3 weeks after exposure to the infected tick. Thereafter, individuals develop fever, chills, headache, and muscle pains. Gastrointestinal symptoms such as nausea, vomiting, and loss of appetite also are common. Laboratory abnormalities regularly include anemia, low white blood cell and platelet counts, and abnormal liver function. More severely ill individuals also may manifest abnormalities of the central nervous system, lungs, and kidneys. Elderly patients are more likely to become seriously ill than children and young adults. Fatality rates up to 4% have been documented. Because the clinical presentation is nonspecific, the diagnosis of ehrlichiosis may not be immediately apparent. Prolonged intervals between the onset of illness and the administration of appropriate therapy can lead to more severe disease symptoms and a greater risk of fatality. *See* CLINICAL MICROBIOLOGY; HEMORRHAGE; INFECTION.

Diagnosis and treatment. An important clue to the diagnosis of human granulocytic ehrlichiosis is the recognition of cytoplasmic vacuoles filled with ehrlichiae (morulae) in circulating neutrophils. Careful examination of stained smears of peripheral blood often yields such findings in human granulocytic ehrlichiosis. In contrast, similar detection of infection in circulating monocytes is a distinctly rare event in disease caused by *E. chaffeensis.* Laboratory diagnosis usually is made by detecting an increase in species-specific antibodies in serum specimens obtained during the acute and convalescent phases of the illness. However, such serologic testing is of no use in establishing the diagnosis before treatment is initiated. In an attempt to shorten this diagnostic lag, some research laboratories are offering to test blood specimens by using polymerase chain reaction (PCR) technology. Unfortunately, this method is not yet widely available. Therefore, therapy must be initiated on clinical suspicion.

Ehrlichiosis closely resembles another tick-borne illness, Rocky Mountain spotted fever, except that the rash, characteristic of spotted fever, is usually absent or modest. Hence, ehrlichiosis has been referred to as spotless fever. Fortunately, both diseases can be treated with tetracycline antibiotics. Most individuals respond to tetracycline therapy within 48–72 h. *See* RICKETTSIOSES.

Prevention. The avoidance of tick bites is fundamental to preventing ehrlichiosis. Sometimes this can best be accomplished by staying out of tick-infested areas. It has been documented that the regular use of insect repellant by individuals who spend time in grassy or wooded areas offers substantial protection against tick bites. It also is advisable to wear long-sleeved shirts and trousers that can be tucked into boots. Individuals should examine themselves for ticks periodically. Ticks attached to the skin can be removed with tweezers by grasping the tick as close to the point of attachment as possible and pulling gently and steadily.

Steven M. Standaert; William Schaffner

Bibliography. J. E. Dawson et al., The interface between research and the diagnosis of an emerging tick-borne disease, human ehrlichiosis due to *Ehrlichia chaffeensis*, *Arch. Int. Med.*, 156:137–142, 1996; J. S. Dumler and J. S. Bakken, Ehrlichial diseases of humans: Emerging tick-borne infections, *Clin. Infect. Dis.*, 20:1102–1110, 1995; S. M. Standaert et al., Ehrlichiosis in a golf-oriented retirement community, *N. Engl. J. Med.*, 333:420–425, 1995; D. H. Walker and J. S. Dumler, Emergence of the ehrlichioses as human health problems, *Emerging Infect. Dis.*, 2:18–29, 1996.

Eicosanoids

A family of naturally occurring, biologically active substances derived from 20 carbon polyunsaturated fatty acids such as arachidonic acid. Members of the family include prostaglandins, thromboxanes, leukotrienes, and epoxyeicosatrienoic acids (EETs). Functioning as local hormones, they are synthesized and exert their biological actions in the same tissue. Many have physiological and pathological effects on the cardiovascular, pulmonary, reproductive, and digestive systems.

Biosynthesis and nomenclature. Eicosanoids are formed from polyunsaturated 20-carbon essential fatty acids. These fatty acids include 8,11,14-eicosatrienoic acid, 5,8,11,14-eicosatetraenoic acid (arachidonic acid), and 5,8,11,14,17-eicosapentaenoic acid. Arachidonic acid is the most abundant precursor of the eicosanoids in humans. It is esterified to phospholipids and other complex lipids in the cell membrane; thus, the cellular concentration of free arachidonic acid is low.

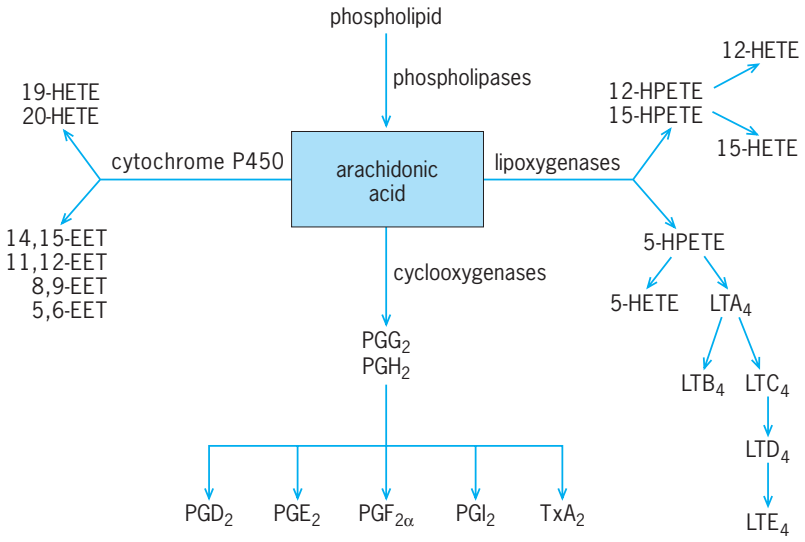

Enzymology of arachidonic acid metabolism. Structures and abbreviations for the compounds are given in Table 1.

Hormones such as histamine, bradykinin, and angiotensin II or physical or chemical stimuli increase the biosynthesis of eicosanoids by releasing arachidonic acid from the membrane lipids and making it available to the eicosanoid-synthesizing enzymes (see **illus.**). These stimuli activate phospholipase A_2 or phospholipase C either directly or by increasing intracellular calcium. Phospholipase A_2 hydrolyzes arachidonic acid from phospholipids, usually phosphatidylcholine or phosphatidylethanolamine. Phospholipase C converts phosphatidylinositol to a 1,2-diglyceride. The sequential action of diglyceride lipase and monoglyceride lipase on the diglyceride releases arachidonic acid. The free arachidonic acid is reesterified into cellular lipids or metabolized by cyclooxygenases, lipoxygenases, or cytochrome P450 monooxygenases. These enzymes add molecular oxygen or oxygens to the fatty acid.

Metabolites of cyclooxygenases. The cyclooxygenases, also called prostaglandin endoperoxide synthases, metabolize unsaturated fatty acids to the prostaglandins and thromboxanes. There are two forms of the enzyme. Cyclooxygenase-1 is a constitutive enzyme found in most cells, whereas cyclooxygenase-2 is an inducible enzyme that becomes elevated in response to mediators of inflammation such as cytokines. Both enzymes perform the same two enzymatic reactions: endoperoxide synthesis and peroxide reduction. The endoperoxide synthase activity oxygenates and cyclizes the arachidonic acid, producing the cyclic endoperoxide prostaglandin G_2 (PGG_2) [**Table 1**]. The peroxidase activity reduces PGG_2 to PGH_2. Both PGG_2 and PGH_2 are unstable intermediates that are further metabolized by isomerases to PGE_2 and PGD_2, by prostacyclin synthase to PGI_2, and by thromboxane synthase to TXA_2. While most cells contain cyclooxygenase, cells vary in their complement of endoperoxide-metabolizing enzymes. As a result, cells make different prostaglandins. For example, vascular endothelial cells contain prostacyclin synthase and not thromboxane synthase. As a result, they synthesize PGI_2 but not TXA_2. In contrast, platelets contain only thromboxane synthase and thus make predominantly TXA_2. This selective synthesis regulates the prostaglandins produced and limits their sites of action. The cyclooxygenases are inhibited by nonsteroidal antiinflammatory drugs such as aspirin, indomethacin, and ibuprofen. By inhibiting cyclooxygenases, these drugs inhibit the synthesis of all of the prostaglandins and thromboxanes and block all of their biological actions. PGE_2 and PGD_2 are inactivated enzymatically in the lung and liver. TXA_2 and PGI_2 are chemically unstable and degrade nonenzymatically to the biologically inactive TXB_2 and 6-keto $PGF_{1\alpha}$, respectively. *See* ASPIRIN.

The cyclooxygenase metabolites fall into categories based on the substituents on the cyclopentane ring (Table 1). PGG_2 and PGH_2 are cyclic endoperoxides, PGE_2 and PGD_2 are hydroxy ketones, PGI_2 has a double ring structure, and $PGF_{2\alpha}$ is a diol. TXA_2 has a six-membered oxane ring. These categories of prostaglandins are further divided by the number of double bonds in the side chain, which is indicated by a subscript 1, 2, or 3. The prostaglandins produced from arachidonic acid have the subscript 2.

Metabolites of lipoxygenases. Lipoxygenases are a family of enzymes that catalyze the addition of a single molecule of oxygen into the unsaturated fatty acid to form a lipid hydroperoxide (Table 1). They are generally constitutive enzymes that depend on the availability of free arachidonic acid. This reaction involves hydrogen abstraction from an allylic carbon adjacent to two cis-double bonds. Since arachidonic acid has four cis-double bonds, several isomeric hydroperoxyeicosatetraenoic acids (HPETEs) may be synthesized, commonly 5-, 12-, and 15-HPETEs. The lipoxygenases place the hydroperoxy group on specific carbons. Cells differ in the lipoxygenases that they contain. For example, platelets contain only 12-lipoxygenase, which catalyzes the formation of only 12-HPETE, whereas leukocytes contain both 5-lipoxygenase and 12-lipoxygenase. 5-Lipoxygenase catalyzes the formation of only 5-HPETE. Like PGG_2 and PGH_2, HPETEs are unstable intermediates that may be further metabolized. They are reduced either enzymatically or nonenzymatically to the hydroxyeicosatetraenoic acid (HETE). Of particular importance is the 5-lipoxygenase pathway which leads to the production of the leukotrienes (Table 1). LTA_4 synthase catalyzes the rearrangement of 5-HPETE to the unstable epoxide, LTA_4. There are two possible enzymatic fates for LTA_4: hydrolysis by LTA hydrolase to LTB_4, or conjugation with glutathione by LTC synthase giving LTC_4. The glutathione group of LTC_4 may be sequentially metabolized by specific peptidases. Removal of the glutamic acid from the glutathione group gives LTD_4, and further removal of the glycine group gives LTE_4. LTB_4, LTC_4, LTD_4, and LTE_4 are all biologically active.

Metabolites of cytochrome P450. Cytochrome P450 is a family of greater than 100 isozymes (any two or more enzymes chemically distinct but functionally

TABLE 1. Structures of eicosanoids

Name	Abbreviation	Structure
Arachidonic acid (5,8,11,14-eicosatetraenoic acid)	AA	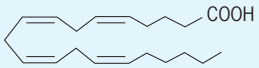
Prostaglandin G_2	PGG_2	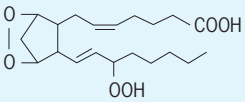
Prostaglandin H_2	PGH_2	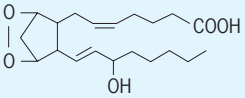
Prostaglandin E_2	PGE_2	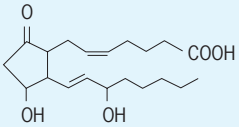
Prostaglandin D_2	PGD_2	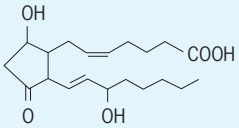
Prostaglandin $F_{2\alpha}$	$PGF_{2\alpha}$	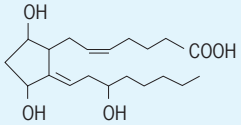
Thromboxane A_2	TXA_2	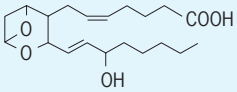
Thromboxane B_2	TXB_2	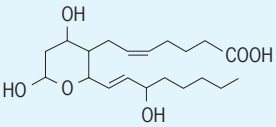
Prostaglandin I_2	PGI_2	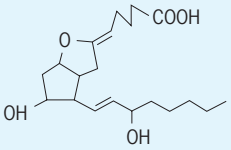
6-Ketoprostaglandin $F_{1\alpha}$	$6\text{-ketoPGF}_{1\alpha}$	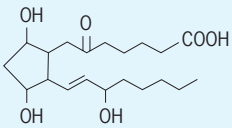
5-Hydroperoxyeicosatetraenoic acid	5-HPETE	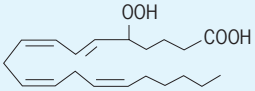
5-Hydroxyeicosatetraenoic acid	5-HETE	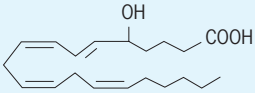
Leukotriene A_4	LTA_4	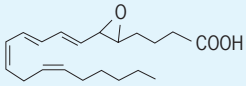
Leukotriene B_4	LTB_4	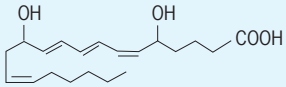
Leukotriene C_4	LTC_4	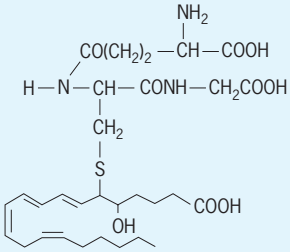
Leukotriene D_4	LTD_4	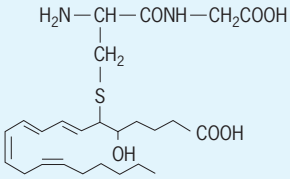
Leukotriene E_4	LTE_4	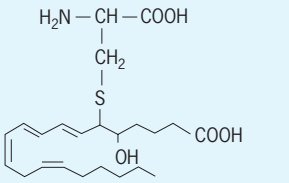
11,12-Epoxyeicosatrienoic acid	11,12-EET	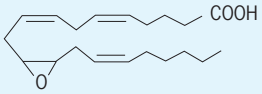
20-Hydroxyeicosatetraenoic acid	20-HETE	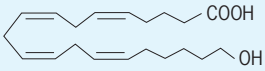
Anandamide		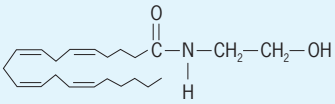

the same) that metabolize endogenous compounds such as fatty acids and cholesterol and exogenous compounds such as drugs and xenobiotics. Arachidonic acid is metabolized by ω- and ω-1-hydroxylases or monooxygenases to 19-HETE and 20-HETE and by epoxygenases to EETs (Table 1). The ω-hydroxylations differ from the lipoxygenase reaction: the HETE is formed directly without the HPETE intermediate, and the ω-hydroxylase does not require the cis-double bonds. These hydroxylations are catalyzed by CYP1A and CYP4A isozymes. The epoxygenases add an oxygen across one of the four double bonds of arachidonic acid, giving 14,15-, 11,12-, 8,9-, or 5,6-EET. Epoxygenations result from CYP2B, CYP2C, and CYP2J isozymes. These enzymes are constitutively found in cells of the liver, kidney,

pituitary gland, and blood vessels. They are regulated by the availability of free arachidonic acid.

Other pathways. The free-radical peroxidation of arachidonic acid and arachidonic acid containing phospholipids gives rise to the isoprostanes. The isoprostanes associated with phospholipids may be released by phospholipases. These eicosanoids are comparable to the prostaglandins in structure, and some also have actions similar to the prostaglandins. Since their synthesis is independent of cyclooxygenase, their formation is not inhibited by nonsteroidal anti-inflammatory drugs.

Arachidonic acid and other fatty acids are coupled to ethanolamine by an amide bond to form *N*-acyl ethanolamides. For arachidonic acid, arachidonylethanolamide or anandamide is produced (Table 1). The synthetic pathway for anandamide and related *N*-acyl ethanolamides is unclear. Two pathways are proposed: hydrolysis of anandamide from *N*-arachidonylphosphatidylethanolamine by phospholipase D or direct condensation of arachidonic acid with ethanolamine. Anandamide is degraded by a fatty acid amidohydrolase, giving arachidonic acid and ethanolamine. The amidohydrolase acting in reverse may represent the synthetic pathway. The brain and heart are major sources of these compounds.

Physiological and pathological actions. The actions of the eicosanoids are diverse, affecting most organ systems (**Table 2**). As a result, when an eicosanoid is administered systemically in pharmacological doses, many of these actions are observed. However, physiologically and pathologically, eicosanoids are local hormones. They are synthesized and exert their action in a tissue, and only certain eicosanoids are made by the cells of a tissue. As a result, the action of an eicosanoid is usually discrete and related to the physiological demands of the tissue.

Mechanism of action. The prostaglandins, leukotrienes, TXA_2, and anandamide act by binding to and activating membrane receptor proteins. These receptors are of the seven-membrane spanning, serpentine receptor subgroup and are coupled to guanine nucleotide binding proteins. Distinct receptors have been identified for many eicosanoids (PGD_2, PGE_2, $PGF_{2\alpha}$, PGI_2, TXA_2, LTB_4, LTC_4, LTD_4/LTE_4) and anandamide. In many cases, subgroups of receptors have been identified. The existence of receptor subgroups explains how some eicosanoids such as PGE_2 have multiple actions. The prostaglandin receptors are termed prostanoid receptors, and they are named after the primary prostaglandin that activates the receptor. For example, the PGI_2 prostanoid receptor is termed the IP receptor. The other receptors are named in an analogous manner. Receptor subgroups are indicated by numbered subscript. The receptors are coupled to stimulation of adenylyl cyclase in the case of IP, DP, and EP_2 receptors. Adenylyl cyclase is inhibited by anandamide via a cannabinoid receptor and PGE_2 via the EP_3 receptor. The receptors for the leukotrienes, TXA_2, $PGF_{2\alpha}$, and PGE_2 (EP_1) are coupled to activation of phospholipase C and increases in intracellular calcium. Receptors have not been identified for 20-HETE or the EETs. Their

TABLE 2. Physiological actions of arachidonic acid metabolites

Eicosanoid	Biological effect	Physiologic or pathologic role
PGE_2	Vasodilation	Protects against intense vasoconstriction
	Decreases gastric acid secretion	Protects gastric mucosa
	Pain sensitization	Pain, inflammation
	Bronchodilation	
	Contracts gastrointestinal smooth muscle	Peristalsis, diarrhea
$PGF_{2\alpha}$	Contracts uterine smooth muscle	Parturition, dysmenorrhea
PGI_2	Contracts gastrointestinal smooth muscle	Peristalsis, diarrhea
	Vasodilation	Protects against intense vasoconstriction
	Inhibits platelet aggregation	Protects against thrombosis
	Decreases gastric acid secretion	Protects gastric mucosa
	Renin release	Blood pressure regulation
TXA_2	Vasoconstriction	Hemostasis, angina pectoris
	Causes platelet aggregation	Thrombosis, angina pectoris
	Contracts most smooth muscle	
	Bronchoconstriction	Bronchial asthma
PGD_2	Inhibits platelet aggregartion	
	Vasodilation	
	Pulmonary vasoconstriction	
LTC_4/LTD_4	Contracts gastrointestinal smooth muscle	
	Constricts peripheral airways	Bronchial asthma
	Increase mucous secretion	Bronchial asthma
	Leakage in microcirculation	Inflammation
LTB_4	Neutrophil and eosinophil chemotaxis	Inflammation
	Leakage in microcirculation	Inflammation
	Causes neutrophil aggregation	Inflammation
12-HETE/12-HPETE	Neutrophil chemotaxis	Inflammation
20-HETE	Vasoconstriction	Myogenic tone
EETs	Vasodilation	Regulates tissue blood flow
	Releases pituitary hormones	
Anandamide	Hypothermia	
	Antinociception	
	Catalepsy	
	Cerebral vasodilation	

mechanisms of action have not been clearly determined.

Cyclooxygenase metabolites. The prostaglandins and TXA_2 have several actions that are necessary for maintaining normal homeostasis and organ integrity. These effects are mediated by their constitutive synthesis through cyclooxygenase-1. PGE_2 and PGI_2 are potent vasodilators. These prostaglandins are synthesized in the blood vessel in response to vasoconstrictors such as angiotensin II and antagonize the effect of the vasoconstrictor. This action protects the organ from intense vasoconstriction and maintains nutritive blood flow. In the absence of these prostaglandins, injury to the organ occurs. These prostaglandins also protect the gastric mucosa. They inhibit gastric acid secretion and may exert other cellular protective effects. In their absence, gastric ulcers develop. TXA_2 is made principally by platelets. It causes vasoconstriction and promotes platelet aggregation. These actions prevent excessive blood loss when a blood vessel is damaged or severed. PGI_2 is made by the blood vessel wall and has the opposite effects. It inhibits platelet aggregation and prevents intravascular platelet aggregates from forming and obstructing blood flow. The balance between the proaggregatory or vasoconstrictor action of TXA_2 and the antiaggregatory or vasodilator action of PGI_2 is critical to cardiovascular homeostasis. For example, an increase in the effect of TXA_2 over PGI_2 has been implicated in ischemic heart disease. $PGF_{2\alpha}$ contracts uterine smooth muscle. With menstruation and destruction of the uterine lining, arachidonic acid is released and metabolized to $PGF_{2\alpha}$. Contraction of the uterine smooth muscle by $PGF_{2\alpha}$ contributes to the pain associated with menstruation. In pregnancy, an increase in prostaglandin synthesis is a major determinant of the onset of labor.

In inflammatory processes, immune cells migrate into the site of inflammation and release cytokines. These inflammatory mediators induce the expression of cyclooxygenase-2 in the immune cells and cells of the tissue. The increased synthesis of PGE_2 and PGI_2 mediates the vasodilation and pain sensitization associated with inflammation and arthritic conditions. Inhibitors of cyclooxygenase-1 and -2 such as aspirin and ibuprofen relieve the signs and symptoms of inflammation. However, they also block the constitutive, homeostatic functions of the prostaglandins by blocking cyclooxygenase-1, which can result in renal failure and gastric ulcers. Selective cyclooxygenase-2 inhibitors such as celecoxib are anti-inflammatory without the deleterious effects on homeostatic functions of the prostaglandins.

Lipoxygenase metabolites. The leukotrienes are synthesized by leukocytes, and they also serve as inflammatory mediators. The leukocytes are immune cells and become sequestered and activated at a site of inflammation. When activated, they release leukotrienes. LTB_4 is chemotactic and promotes leukocyte aggregation. Thus, it attracts other inflammatory cells to the site of inflammation and sequesters them. LTC_4, LTD_4, and LTB_4 increase capillary permeability, causing the leakage of fluid and protein from the vasculature into the tissue. This contributes to the swelling associated with inflammation. LTC_4 and LTD_4 also cause bronchiolar constriction and increase mucus formation in the airways. These actions are thought to contribute to bronchial asthma. Inhibitors of 5-lipoxygenase and leukotriene receptor antagonists are useful in treating inflammation and bronchial asthma.

Cytochrome P450 metabolites. The EETs and 20-HETE are involved in the regulation of vascular tone and organ blood flow. In the blood vessel, EETs are synthesized by the endothelial cells in response to vasodilator hormones such as acetylcholine and bradykinin. The EETs diffuse to the adjacent smooth muscle cells and exert their action. They open potassium channels and hyperpolarize the smooth muscle cell membrane. This results in vasodilation. Thus, EETs serve as mediators of vasodilation. 20-HETE is made by smooth muscle cells in response to increases in intravascular pressure. It inhibits potassium channel activity and depolarizes the cell membrane. Depolarization promotes the influx of calcium and vasoconstriction. 20-HETE is a mediator of myogenic tone in blood vessels. In the blood vessel wall, 20-HETE and the EETs are counterregulatory mechanisms that control blood flow.

Other pathways. Anandamide is a naturally occurring agonist of the cannabinoid receptor. It has biological actions identical to Δ^9-tetrahydrocannabinol, the active ingredient in marijuana, but is shorter-acting. Anandamide is synthesized and degraded by discrete regions of the brain and causes hypothermia (decreased body temperature), antinociception (decreased pain response), catalepsy (rigidity and decreased response to external stimuli), and cerebral vasodilation. It remains to be determined if anandamide functions as a chemical mediator in the brain. *See* ARTERIOSCLEROSIS; ARTHRITIS; ASTHMA; GOUT; MARIJUANA; PAIN; ULCER. William B. Campbell

Bibliography. C. Denzlinger, Biology and pathophysiology of leukotrienes, *Crit. Rev. Oncology/Hematology*, 23:167–223, 1996; C. J. Hillard and W. B. Campbell, Biochemistry and pharmacology of arachidonylethanolamide, a putative endogenous cannabinoid, *J. Lipid Res.*, 38:2382–2398, 1997; J. C. McGiff, Cytochrome P-450 metabolism of arachidonic acid, *Annu. Rev. Pharmacol.*, 31:339–369, 1991; W. L. Smith, L. J. Marnett, and D. L. DeWitt, Prostaglandin and thromboxane biosynthesis, *Pharmacol. Ther.*, 49:153–179, 1991; C. S. Williams and R. N. Dubois, Prostaglandin endoperoxide synthase: Why two isoforms?, *Amer. J. Physiol.*, 270:G393–G400, 1996.

Eigenfunction

One of the solutions of an eigenvalue equation. A parameter-dependent equation that possesses nonvanishing solutions only for particular values (eigenvalues) of the parameter is an eigenvalue equation, the associated solutions being the eigenfunctions (sometimes eigenvectors). In older usage the

terms characteristic equation and characteristic values (functions) are common. Eigenvalue equations appear in many contexts, including the solution of systems of linear algebraic equations (matrix equations), differential or partial differential equations, and integral equations. The importance of eigenfunctions and eigenvalues in applied mathematics results from the widespread applicability of linear equations as exact or approximate descriptions of physical systems. However, the most fundamental application of these concepts is in quantum mechanics where they enter into the definition and physical interpretation of the theory. Only linear eigenvalue equations will be discussed. *See* EIGENVALUE (QUANTUM MECHANICS); ENERGY LEVEL (QUANTUM MECHANICS); NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Matrix equations. An orienting example will first be presented. A common problem, for instance in the theory of angular momentum, is to find the column vectors v which are solutions of the matrix, Eq. (1), where M is an N-by-N matrix with components M_{ij} which are assumed to be known, v is a column vector of length N with components v_i, and λ is the parameter. All quantities can in general be complex.

$$M \cdot v = \lambda v \tag{1}$$

For example, the nonvanishing solutions to Eq. (2)

$$\begin{pmatrix} -1 & 3 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \tag{2}$$

may be sought. This equation can be rewritten as $M' \cdot v = \vec{0}$ for $M' = (M - \lambda 1_2)$. (Here, 1_2 is the two-dimensional unit matrix, and $\vec{0}$ is the zero vector.) The condition for such an equation to have nontrivial solutions for the vector, v, is that the determinant, $\det M'$, of the matrix M' is zero. Otherwise the inverse of M' would exist, implying that the only allowed solution is $v = (M')^{-1}\,\vec{0} = \vec{0}$. Thus, Eq. (3) is required.

$$\det \begin{pmatrix} -1-\lambda & 3 \\ 2 & -\lambda \end{pmatrix} = (1+\lambda)\lambda - 6 = 0 \tag{3}$$

The solutions are $\lambda = 2$ and $\lambda = -3$, the eigenvalues of the matrix equation. The two associated eigenvectors $v^{(1)}$ and $v^{(2)}$ are discovered by substituting these values into Eq. (2). For the eigenvalue $\lambda = 2$, the relation $v_1^{(1)} = v_2^{(1)}$ is found. This does not set the overall scale of $v^{(1)}$. Similarly, for $\lambda = -3$ the relation $v_1^{(2)} = -\,3v_2^{(2)}/2$ is found. The scale of the eigenvectors can be fixed by imposing the additional condition that the generalized dot product $(v^{(i)})^* \cdot v^{(i)} = 1$ for $i = 1, 2$.

The general N-by-N matrix equation can similarly be rewritten as Eq. (4). The condition for the existence of nonvanishing solutions for v is Eq. (5). This

$$(M - \lambda 1_N)v = \vec{0} \tag{4}$$

$$\det\,(M - \lambda 1_N) = 0 \tag{5}$$

determinant, when expanded, gives an Nth-order algebraic equation for λ of the form of Eq. (6). By the

$$\lambda^N + a_1\lambda^{N-1} + a_2\lambda^{N-2} + \cdots + a_N = 0 \tag{6}$$

fundamental theorem of algebra, this equation has N solutions λ_i, $i = 1, \ldots, N$, the eigenvalues of M, if λ is allowed to be complex. The solutions can be degenerate; in other words, a given value, λ, can appear more than once as a solution.

A particularly important special case is when the matrix M is self-adjoint (also known as hermitian), namely, when M is equal to its complex-conjugate transpose, as in Eq. (7). The eigenvalues of M are

$$M^\dagger \equiv M^{T*} = M \tag{7}$$

then guaranteed to be real, and the eigenvectors are orthonormal, as in Eq. (8), where $\delta^{ij} = 0$ when $i \neq j$

$$(v^{(i)})^* \cdot v^{(j)} = \delta^{ij} \tag{8}$$

and $\delta^{ij} = 1$ when $i = j$. For k degenerate eigenvalues, only a k-dimensional subspace is uniquely specified, but it is always possible to choose k such orthonormal eigenvectors, one for each eigenvalue. *See* ANGULAR MOMENTUM; MATRIX CALCULUS; MATRIX THEORY; RIGID-BODY DYNAMICS.

Linear operators. The general setting for linear eigenvalue equations is the notion of a linear operator $\hat{L}$ acting in a vector space V endowed with an inner product. A complex vector space is a set of elements (vectors), $\{|u\rangle, |v\rangle, \ldots\}$, that have the operations of addition and scalar multiplication defined, such that $\alpha|u\rangle + \beta|v\rangle$ is an element of V for any elements, $|u\rangle$ and $|v\rangle$, of V, and any complex numbers, α and β. Here, the notation introduced by P. Dirac in the context of quantum mechanics is employed. The operations of addition and scalar multiplication must further satisfy properties such as associativity, commutativity, and the distributive law.

A linear operator $\hat{L}$ is a rule that takes an arbitrary vector $|u\rangle$ in V into a, generally different, vector $|Lu\rangle$ in V, and that satisfies Eq. (9) for any vectors, $|u\rangle$

$$\hat{L}(\alpha|u\rangle + \beta|v\rangle) = \alpha|Lu\rangle + \beta|Lv\rangle \tag{9}$$

and $|v\rangle$, in V, and complex numbers α and β. The usual action of a matrix on a vector by matrix multiplication satisfies this linearity property. In this general setting, eigenvalue equations again take the form $\hat{L}|v\rangle = \lambda|v\rangle$. The total set of eigenvalues of a linear operator $\hat{L}$ acting on V is called the spectrum of $\hat{L}$. For differential operators this spectrum can contain continuous parts.

An inner product on the vector space V, denoted $\langle u|v\rangle$, is a complex-valued function of two arbitrary vectors, $|u\rangle$ and $|v\rangle$, in V which satisfies Eqs. (10)

$$\langle u|u\rangle \geq 0,\ \langle u|u\rangle = 0 \text{ only for } |u\rangle = |0\rangle \tag{10a}$$

$$\langle u|v\rangle = \langle v|u\rangle^* \tag{10b}$$

$$\langle u \mid \alpha v + \beta w\rangle = \alpha\langle u|v\rangle + \beta\langle u|w\rangle \tag{10c}$$

for any vectors, $|u\rangle$, $|v\rangle$, and $|w\rangle$, in V and complex

numbers α and β. Here, $|0\rangle$ is the zero vector in V. The inner product generalizes the usual dot product for three-dimensional vectors in physical space.

A basis for an N-dimensional complex vector space V is any set of N vectors $|e^{(i)}\rangle$ in V such that any vector $|u\rangle$ in V can be uniquely written as the sum in Eq. (11), where the N coefficients can be complex.

$$|u\rangle = \sum_{i=1}^{N} \alpha_i |e^{(i)}\rangle \tag{11}$$

For a general vector space, the number of basis elements can be (countably) infinite, with $N = \infty$, or the basis elements can even be labeled by a continuous parameter, in which case the dimension of the vector space is uncountably infinite and sums, such as in Eq. (11), need to be replaced by integrals. In both these latter cases, subtle questions of the convergence of various expressions arise.

Given any such basis and an inner product, a matrix representation L_{ij} of any abstract linear operator $\hat{L}$ can be constructed, by defining the components as in Eq. (12). The notion of the adjoint $\hat{L}^\dagger$ of an operator $\hat{L}$ can be generalized by the definition in Eq. (13)

$$L_{ij} = \langle e^{(i)}|Le^{(j)}\rangle \tag{12}$$

$$\langle u|L^\dagger v\rangle = \langle u|Lu\rangle^* \tag{13}$$

for any $|u\rangle$ and $|v\rangle$ in V, and the matrix adjoint is then given by $L^\dagger{}_{ij} = L^*_{ji}$, in accordance with Eq. (7). A self-adjoint operator satisfies the property $\hat{L}^\dagger = \hat{L}$ for all elements in the vector space V. Just as in the case of a self-adjoint matrix, the eigenvalues of a such an operator are purely real, and the eigenvectors can always be chosen to be orthonormal.

Since the eigenvectors $\{|v^{(i)}\rangle\}$ of any self-adjoint operator $\hat{L}$ acting on V are a set of N orthonormal vectors in the N-dimensional space V, they form a basis for V and any vector $|w\rangle$ can be expanded as in Eq. (14), where the generally complex numbers

$$|w\rangle = \sum_{i=1}^{N} |v^{(i)}\rangle\langle v^{(i)}|w\rangle \tag{14}$$

$\langle v^{(i)}|w\rangle$ play the role of the α_i in Eq. (11). Since $|w\rangle$ is arbitrary, this equation implies the important consequence, Eq. (15), where $\hat{1}$ is the identity operator

$$\sum_{i=1}^{N} |v^{(i)}\rangle\langle v^{(i)}| = \hat{1} \tag{15}$$

on the space V (namely, the operator that takes any vector to itself). Equation (15) is known as the completeness relation for the eigenvectors of $\hat{L}$.

The relation (14), or equivalently (15), implies a very useful representation for any self-adjoint operator $\hat{M}$. If $|w\rangle$ is expanded in the set of eigenvectors $|\tilde{v}^{(i)}\rangle$ associated with $\hat{M}$, then, using the definition $\hat{M}|\tilde{v}^{(i)}\rangle = \lambda_i|\tilde{v}^{(i)}\rangle$, along with the fact that $|w\rangle$ is arbitrary, then gives Eq. (16). This is a simple,

$$\hat{M} = \sum_{i=1}^{N} \lambda_i |\tilde{v}^{(i)}\rangle\langle\tilde{v}^{i}| \tag{16}$$

finite-dimensional example of the spectral theorem, stating that the action of an operator may be represented purely in terms of its eigenvalue spectrum and eigenvectors. For matrices, this result says that any self-adjoint matrix can be put into a purely diagonal form by choosing the eigenvalues as a basis, the diagonal entries being the eigenvalues of the matrix. *See* HILBERT SPACE; LINEAR ALGEBRA; OPERATOR THEORY.

Differential eigenvalue equations. Eigenvalue problems also appear very frequently in the solution of linear differential equations. An example is the motion of a uniform vibrating string with fixed ends at $x = 0, L$. The partial differential equation that describes the displacement $y(x, t)$ of the string from its equilibrium position (taken to be $y = 0$) is Eq. (17).

$$\frac{\partial^2 y(x,t)}{\partial t^2} - c^2 \frac{\partial^2 y(x,t)}{\partial x^2} = 0 \tag{17}$$

where c is the wave velocity on the string. Along with this equation, appropriate initial and boundary conditions must be specified. The fixed end-point condition implies $y(0,t) = y(L,t) = 0$ for all t. A typical initial condition is that the string is held stationary with some particular displacement and then released.

The solution of the problem is greatly facilitated by the principle of superposition, which states that, for two solutions, $y^{(1)}(x,t)$ and $y^{(2)}(x,t)$, of a linear equation, the sum in expression (18) is also a solution

$$\alpha y^{(1)}(x,t) + \beta y^{(2)}(x,t) \tag{18}$$

for arbitrary constants α and β. The general strategy to solve such a problem is therefore to decompose $y(x, t)$ into sum of simple motions, the eigenfunctions of the problem, whose time development can straightforwardly be solved for, and then form a suitable linear combination such that the initial conditions are satisfied.

Explicitly, elementary solutions of the form $y(x, t) = h(x) \cos \omega (t - t_0)$ are studied. Substituting this into Eq. (17) results in Eq. (19). This is of the

$$\frac{d^2 h(x)}{dx^2} = -\frac{\omega^2}{c^2} h(x) \tag{19}$$

form of an eigenvalue equation with ω^2 as the parameter. The set of (differentiable) functions $\{f(x)\}$ on the interval $[0,L]$ which vanish at the end points forms a vector space V (of infinite dimension), and the differential operator d^2/dx^2 is a linear operator acting on V. An inner product on this space can be defined by Eq. (20), the adjoint operator $(d^2/dx^2)^\dagger$

$$\langle f|g\rangle \equiv \int_0^L dx (f(x))^* g(x) \tag{20}$$

can be defined with respect to this inner product by applying Eq. (13), and d^2/dx^2 can be shown to be a self-adjoint operator. The eigenvalues of this operator must therefore be real, and the associated eigenfunctions satisfy orthonormality and completeness relations.

The general solution to Eq. (19) is Eq. (21), with

$$b(x) = A\cos(\omega x/c) + B\sin(\omega x/c) \qquad (21)$$

A, B arbitrary constants. Imposing the boundary conditions at $x = 0$, L gives $A = 0$ and $B\sin(\omega L/c) = 0$. A nontrivial solution ($B \neq 0$) therefore fixes the frequency ω to be one of the discrete set given by Eq. (22), the eigenvalues of the problem. The associ-

$$\omega_n = \frac{\pi nc}{L} \quad \text{for} \quad n = 1, 2, \ldots \qquad (22)$$

ated eigenfunctions are $b_n(x) = \sin(\omega_n x/c)$. In the absence of fixed end-point boundary conditions (that is, for a string of infinite length) the possible values of ω would not be restricted, and the spectrum of eigenvalues would be continuous.

A linear combination, Eq. (23), which, by the prin-

$$y(x, t) = \sum_{n=1}^{\infty} a_n \sin(\omega_n x/c)\cos\ \omega(t - t_0) \qquad (23)$$

ciple of superposition, is still a solution to the original equation (17), is now constructed in such a way that the initial conditions are satisfied. The analog of the completeness relations, Eqs. (11) and (15), appropriate for this case imply that any continuous function $y_0(x)$ representing the initial displacement [satisfying $y_0(0) = y_0(L) = 0$] on the interval $[0,L]$ can be expanded in such a series (Fourier's theorem). The orthogonality property then leads to the formula (24) for the coefficients. Equation (24), together

$$a_n = \frac{2}{L}\int_0^L dx \sin(\pi nx/L) y_0(x) \qquad (24)$$

with Eqs. (22) and (23), constitutes a full solution of the problem. *See* DIFFERENTIAL EQUATION; FOURIER SERIES AND TRANSFORMS; INTEGRAL TRANSFORM; ORTHOGONAL POLYNOMIALS. John March-Russell

Bibliography. G. B. Arfken and H.-J. Weber (eds.), *Mathematical Methods for Physicists*, 4th ed., 1995; R. Geroch, *Mathematical Physics*, 1985; E. Merzbacher, *Quantum Mechanics*, 3d ed., 1997; R. Shankar, *Principles of Quantum Mechanics*, 2d ed., 1994.

Eigenvalue (quantum mechanics)

One of the values of the parameter in an eigenvalue equation for which the equation has a solution. A linear eigenvalue equation is a parameter-dependent equation of the form (1) that possesses nonvanish-

$$\hat{L}|v\rangle = \lambda|v\rangle \qquad (1)$$

ing solutions, $|v^{(n)}\rangle$, only for particular values, λ_n, of the parameter. These are the eigenvalues, the associated solutions being called the eigenfunctions (sometimes eigenvectors). Here $\hat{L}$ is a linear operator acting in a generalized vector space, V, with an inner product. The concepts of eigenvalue and eigenfunction are of fundamental importance in quantum mechanics. The relation between the mathematical formulation of the theory and the possible results of a physical measurement is couched in terms of the eigenvalues and eigenvectors of a privileged set, $\{\hat{L}\}$, of linear operators which represent all possible physical observables. *See* EIGENFUNCTION.

Wave functions and operators. The mathematical formulation and physical interpretation of quantum mechanics is encapsulated in a number of fundamental postulates. The first is that all possible information regarding the physical state of any system at a given time is encoded in a generalized vector of a vector space, V. In most quantum-mechanics applications these vectors are complex-valued functions, $\psi(x)$, of the coordinates, x, of the system, termed wave functions. Only this case is considered in the following. Physical observables, L, are represented by self-adjoint linear operators acting on this function space. For instance, the momentum of a particle moving in one dimension is represented by the differential operator in Eq. (2), where $\hbar$ is Planck's constant over 2π.

$$\hat{p} = -i\hbar\frac{d}{dx} \qquad (2)$$

Furthermore, any self-adjoint linear operator, $\hat{L}$, has a set of eigenfunctions, $\phi^{(n)}(x)$ (for simplicity, assumed indexed by a discrete label n), in terms of which any wave function, $\psi(x)$, in V may be expanded, as in Eq. (3). Here the coefficients a_n are

$$\psi(x) = \sum_n a_n \phi^{(n)}(x) \qquad (3)$$

complex numbers. The sense of Eq. (3) is that both sides are equivalent, at least as far as all physical measurements are concerned. This is known as the expansion postulate.

Measurement of observables. A particularly important aspect of the theory is that the only possible result of a physical measurement of an observable, L, is one of the eigenvalues, λ_n, of the operator $\hat{L}$ corresponding to L, as in Eq. (4). Since the result of a

$$\hat{L}\phi^{(n)}(x) = \lambda_n \phi^{(n)}(x) \qquad (4)$$

physical measurement is necessarily a real number, the set of operators allowed as representatives of observables must be restricted by the requirement that their eigenvalues are purely real. Self-adjoint operators have this property, and this is the reason why representatives of observables are restricted to this class. In addition, the probability of finding the result λ_n when the system is initially in an arbitrary state $\psi(x)$ is given by Eq. (5), where the brackets rep-

$$\text{Prob}(\lambda_n) = |\langle\phi^{(n)} \mid \psi\rangle|^2 = |a_n|^2 \qquad (5)$$

resent the inner product on the space V, given by Eq. (6). After measuring a particular value, λ_n, the

$$\langle\phi^{(n)} \mid \psi\rangle = \int dx(\phi^{(n)}(x))^* \psi(x) \qquad (6)$$

new state of the system is given purely by $\psi(x) = \phi^{(n)}(x)$. Only when the wave function of the system is a single eigenfunction, $\phi^{(n)}(x)$, of some observable,

$\hat{L}$, can the system be said to definitely have the property λ_n.

Eigenvalue spectra. Unlike the case of classical mechanics where dynamical quantities such as energy or angular momentum can take on continuously varying values, the spectrum of eigenvalues is very often discrete, or contains a discrete portion. The values of the associated dynamical quantities are then said to be quantized. For instance, the energy of an electron bound in an atom can be only one of a discrete set of energy eigenvalues. *See* ENERGY LEVEL (QUANTUM MECHANICS).

Interpretation of superposition. Although Eq. (3) used the fact that the superposition principle applies to quantum mechanics, its interpretation is quite different from that in classical mechanics. For example, if two functions describing the displacement of an elastic string are superposed, $a_1y_1(x) + a_2y_2(x)$, then this new string configuration has, in general, quite different physical properties from a string with either displacement $y_1(x)$ or $y_2(x)$. However, in quantum mechanics the superposition $a_1\phi^{(1)}(x) + a_2\phi^{(2)}(x)$ of two eigenfunctions has the interpretation that sometimes the system has exactly the physical properties of $\phi^{(1)}(x)$, other times that of $\phi^{(2)}(x)$, with probabilities $|a_1|^2$ and $|a_2|^2$ respectively.

Commutation of operators. If the energy operator (the hamiltonian), $\hat{H}$, of the system commutes with an operator $\hat{L}$, that is, $\hat{H}\hat{L} = \hat{L}\hat{H}$, then a wave function that is an eigenfunction $\phi^{(n)}(x)$ of $\hat{L}$ stays so as the system dynamically evolves. Thus, the associated eigenvalue, λ_n, is a constant of the motion, and can be thought of as labeling the state of the system. The term quantum number is then used for λ_n. Further, it is possible to find a wave function which is simultaneously the eigenvector of two different operators if and only if they commute. A fundamental property of quantum mechanics is that the maximal number of attributes that can be simultaneously, and precisely, known about any system is given by the eigenvalues of the largest possible commuting set of operators that also includes the hamiltonian $\hat{H}$, in other words, the maximal set of quantum numbers. For example, the operators representing momentum and position do not commute, and it is not possible to know both simultaneously with arbitrary accuracy.

Although these postulates imply a picture of the physical world quite different from that in classical physics, the consequences of these postulates, together with the rest of quantum mechanics, have been verified with extraordinary precision. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

John March-Russell

Bibliography. S. Gasiorowicz, *Quantum Physics*, 2d ed., 1995; E. Merzbacher, *Quantum Mechanics*, 3d ed., 1997; R. Shankar, *Principles of Quantum Mechanics*, 2d ed., 1994.

Einsteinium

A chemical element, Es, atomic number 99, a member of the actinide series in the periodic table. It is not found in nature but is produced by artificial nuclear transmutation of lighter elements. All isotopes of einsteinium are radioactive, decaying with half-lives ranging from a few seconds to about 1 year. *See* ACTINIDE ELEMENTS; PERIODIC TABLE; RADIOACTIVITY.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	1 H																	2 He
	3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
	11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
	19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
	37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
	55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
	87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					
lanthanide series			57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb		
actinide series			89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No		

Einsteinium is the heaviest actinide element to be isolated in weighable form. The metal is chemically reactive, is quite volatile, and melts at 860°C (1580°F); one crystal structure is known. *See* TRANSURANIUM ELEMENTS.

Stanley G. Thompson; Glenn T. Seaborg

Bibliography. S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; J. J. Katz and G. T. Seaborg, *The Chemistry of the Actinide Elements*, 2d ed., 1986; L. Morss and J. Fuger (eds.), *Transuranium Elements: A Half Century*, 1992; G. T. Seaborg, *Modern Alchemy: Selected Papers of Glenn T. Seaborg*, 1994; G. T. Seaborg and W. D. Loveland, *The Elements Beyond Uranium*, 1990.

El Niño

In general, an influx of warm water to the central and eastern equatorial Pacific Ocean off the coast of Peru and Ecuador, with a return period of 4–7 years. El Niño events come in various strengths: weak, moderate, strong, very strong, and extraordinary. The size of an event can be determined using various criteria: the amount of warming of sea-surface temperatures in the central and eastern Pacific from their average condition; the areal extent of that warm water anomaly; and the length of time that the warm water lingers before being replaced by colder-than-average sea-surface temperatures in this tropical Pacific region.

Under normal conditions the winds blow up the west coast of South America and then near the Equator turn westward toward Asia. The surface water is piled up in the western Pacific, and the sea level there is several tens of centimeters above average, while the sea level in the eastern Pacific is below average. As the water is pushed toward the west, cold water from the deeper part of the ocean along the Peruvian coast wells up to the surface to replace it. This cold water is rich with nutrients, making the coastal upwelling region along western South America among the most productive fisheries in the world.

Every 4–7 years those winds tend to die down and sometimes reverse, allowing the warm surface waters that piled up in the west to move back toward the eastern part of the Pacific Basin. With reduced westward winds, the surface water also heats up. Sea level drops in the western Pacific and increases in the eastern part of the basin. An El Niño condition can last for 12–18 months or longer, before the westward-flowing winds start to pick up again.

Limiting the impact. El Niño causes shifts in regional and local climate conditions around the globe. For many locations, forecasting El Niño provides useful information about how local weather conditions might be affected during the lifetime of the phenomenon. This lead time enables societies to prepare for the impacts of these extreme events.

The 1997–1998 and 1982–1983 events are considered the two biggest El Niños in the past 100 years. Both surprised the scientific community, but in different ways. The 1982–1983 event was the most damaging up to that time; no one had forecast it. The 1997–1998 event was larger and developed earlier and faster than expected. While El Niño's onset was not forecast, many of its worldwide impacts were, enabling societies to prepare for the onslaught. The 1982–1983 event emphasized to scientists and policy makers the importance of understanding and forecasting El Niño. The 1997–1998 event exemplified how many sectors of society (for example, agriculture, mining, construction, fisheries, hydroelectric power, and health) use El Niño forecasts to improve their decision-making processes.

El Niño is considered the second biggest climate-related influence on human activities, after the natural flow of the seasons. Although the phenomenon is at least thousands of years old, its impacts on global climate have only recently been recognized. Due to improved scientific understanding and forecasting of El Niño's interannual process, societies can prepare for and reduce its impacts considerably.

Michael H. Glantz

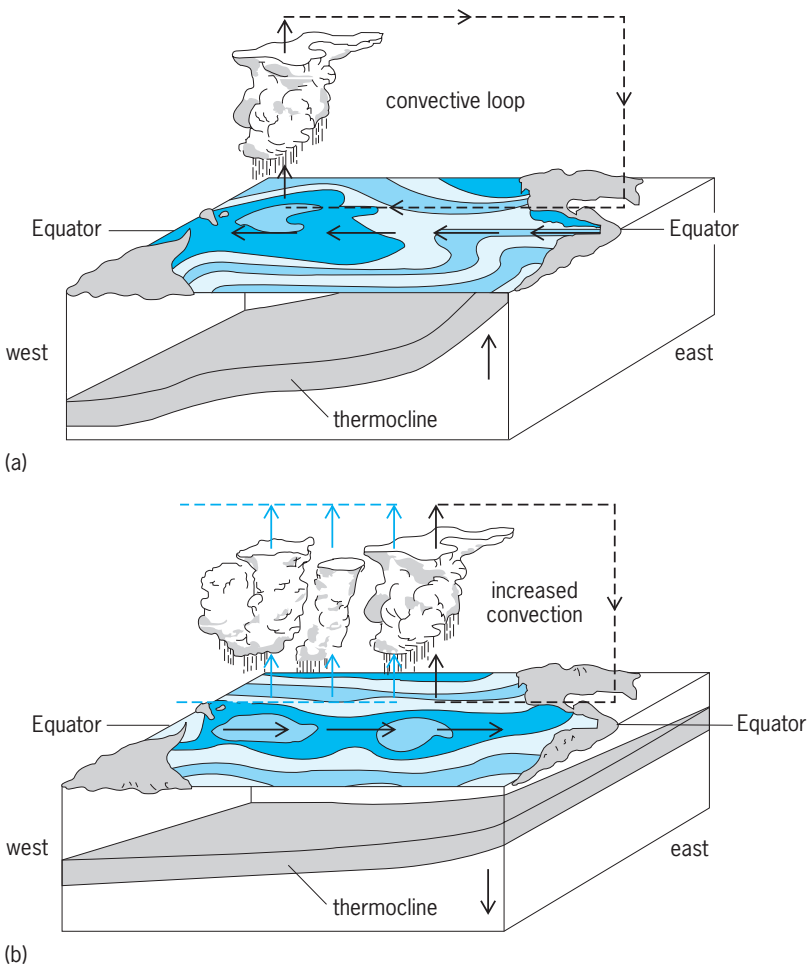

Fig. 1. Conditions of the tropical atmosphere–ocean system over the equatorial Pacific. (*a*) Normal, in which the thermocline is highly tilted and approaches the surface in the east Pacific, the eastern sea-surface temperature is cool, the western sea-surface temperature is warm, and rainfall remains in the far west. (*b*) El Niño (warm phase of ENSO) in which the thermocline flattens, the east and central Pacific warms, and the rainfall moves into the central Pacific.

Tropical Pacific Ocean. The tropical Pacific Ocean is 15,000 km (9300 mi) wide. Over the very warm water at its western end, there is heavy rainfall, low atmospheric surface pressure, and large-scale atmospheric ascent. At the eastern end the water is usually cool, there is little rainfall, atmospheric surface pressure is high, and the air is descending (**Fig. 1***a*). This east-west asymmetry of sea-surface temperature in the Pacific is quite remarkable, for the Earth rotating about its axis sees the Sun with no east-west asymmetry.

The proximate cause of this east-west difference in the sea-surface temperature is the direction of the surface wind, which in the tropics normally blows from east to west. If there were no wind, the sea-surface temperature would be uniformly warm in the east-west direction, and it would rapidly drop at deeper levels in the ocean. The region of rapid temperature change with depth is called the thermocline (Fig. 1*a*), and in the absence of winds would remain flat from east to west.

The thermocline can be thought of as marking the depth of the cold water. When the thermocline is close to the surface, cold water is near the surface; when the thermocline is deep, cold water is far from the surface. The normal westward wind tilts the thermocline, and deepens it in the west and shallows it in the east (Fig. 1*a*). The westward wind tends to bring cold water to the surface in the east (where the cold water is near the surface), and to bring warm water to the surface in the west (where the cold water is far from the surface). The west-to-east decrease of sea-surface temperature is therefore due to the cold water upwelled in the east by the westward winds. In the absence of westward winds, the temperature of the tropical Pacific would be uniform from east to west and would have about the same temperature that the western Pacific has now.

Superimposed on this east-to-west temperature difference is an annual cycle that has east Pacific sea-surface temperatures coolest in August and warmest in April, in contrast to midlatitudes, which have January and July as the extreme months. In general, the warm pool in the west moves northwest

during northern summer and southeast during northern winter. Throughout this complicated annual cycle, the basic pattern is maintained, with warm water in the west and its concomitant heavy rainfall, low pressure, and upward atmospheric motion. *See* PACIFIC OCEAN; SEAWATER; TROPICAL METEOROLOGY; UPWELLING.

Temperature, pressure, and precipitation cycles. Every 4–7 years, the temperature pattern of the tropical Pacific changes dramatically. It is as if all the processes that normally maintain the climatic conditions no longer occur and the tropical Pacific becomes uniformly warm, which would be expected in the absence of winds. This sporadic failure of the normal climate and its transition to a warmer Pacific has acquired the name El Niño (Fig. 1*b*). Occasionally the opposite also occurs, where the eastern Pacific becomes cooler than normal, rainfall decreases still more, atmospheric surface pressure increases, and the westward winds become stronger. This irregular cyclic swing of warm and cold phases in the tropical Pacific is called the El Niño Southern Oscillation (ENSO). The Southern Oscillation is the oscillation of the difference of surface pressure between the east and the west as the sea-surface temperature changes. When the eastern tropical Pacific is warm (El Niño or warm-phase conditions), the pressure in the east decreases and the pressure in the west increases, so the anomalous pressure difference (the Southern Oscillation Index) is negative. Precipitation actively participates in the cycle since regions of persistent precipitation lie over the warmest waters. *See* CLIMATOLOGY; MARITIME METEOROLOGY; PRECIPITATION (METEOROLOGY).

Observing ENSO. Prior to 1979, observations of equatorial Pacific climate were limited to surface measurements from islands and merchant ships; beginning in 1979, satellite estimates of surface and atmospheric-column quantities became available. Aside from a small number of research cruises, the interior of the ocean remained a mystery. As a result of the international Tropical Ocean Global Atmosphere (TOGA) program (1985–1995), an observing system specifically designed to reveal and monitor those features then believed most relevant to ENSO was deployed, and is maintained to this day (**Fig. 2**).

The centerpiece of the system is a set of 70 buoys moored to the bottom of the ocean by anchors (usually discarded railroad wheels) and connected to a surface torus by 4–5 km (2.5–3 mi) of line. The torus contains a suite of instruments that measure near-surface quantities (winds, humidity, and sea-surface temperature). Hung on the line beneath the surface and extending downward to about 500 m (1600 ft) is a set of instruments that measure temperature and pressure (with some measuring ocean currents) and relay the information to computers on the torus. Once a day the information is relayed to satellites and then becomes freely available by the Global Telecommunication System, whose users are the weather services of the world, and by the Internet to other users and to the public. Additional observations by ships, surface buoys drifting with the currents, and tide

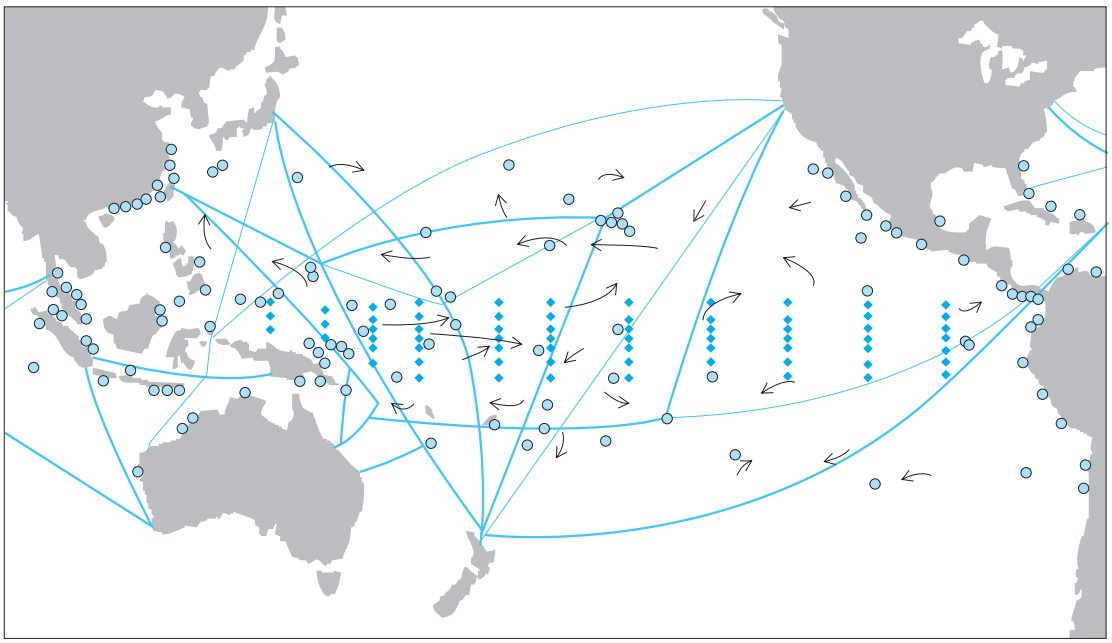

Key:

◆ moored buoys ← drifting buoys

○ tide gauge stations — volunteer observing ships

Fig. 2. ENSO Observing System. Diamonds are the moored buoys, arrows are drifting buoys, points are tide gauges, and lines are ships.

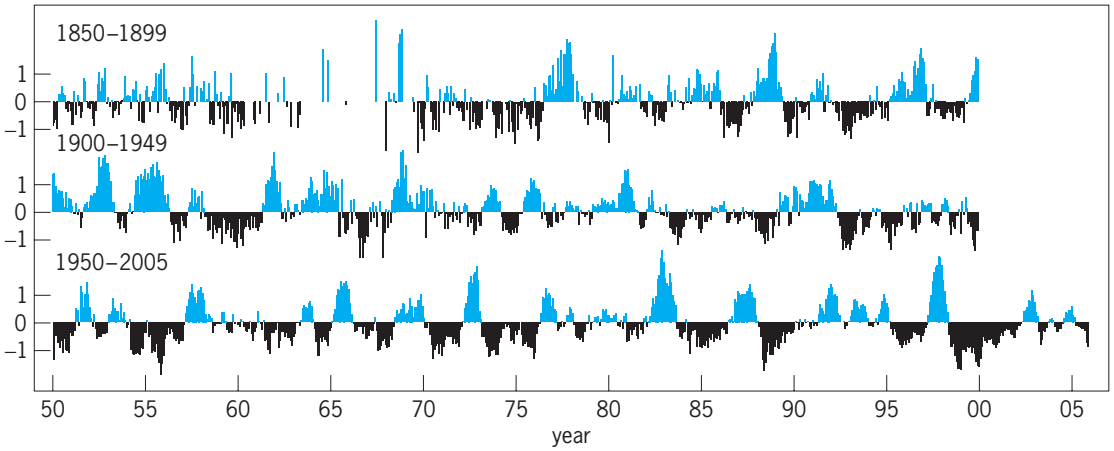

Fig. 3. Time series of an index of ENSO consisting of the monthly average temperature anomaly in the equatorial Pacific from 6°S to 6°N, 90°W to 180°W. The global mean sea-surface temperature deviation has been subtracted for each month to remove the effect of the gradual warming of the Earth's temperature, which is approximately 0.8°C (1.44°F) over the record. There are no observations for many of the months in the earlier periods, especially the 1860s, which accounts for their more ragged appearance.

gauges at islands complete the system. The ENSO observing system has been in full operation since 1995 and provides an unprecedented view of the evolution of ENSO. In particular, it was available to observe what is arguably the largest warm phase of ENSO in the last 150 years, the event of 1997–1998. *See* INSTRUMENTED BUOYS; SATELLITE METEOROLOGY.

History of ENSO. Instrumental records of surface atmospheric pressure and sea-surface temperatures document ENSO variability back to the 1870s and, with less confidence, to the 1850s. The monthly average central and eastern equatorial Pacific sea-surface temperature exhibits persistent, large-magnitude (>1°C or 1.8°F) deviations from the annual cycle (**Fig. 3**) that are the temporal signature of ENSO. The 1982–1983 and 1997–1998 warm phases of ENSO were characterized by temperature deviations in excess of 2°C (3.6°F). It also can be seen that the frequency of the warm and cold episodes varies throughout the record and, in particular, from decade to decade. The years since the 1997–1998 warm episodes have been, on average, colder than the long-term mean.

ENSO-related precipitation and temperature variations produce changes and hence proxy records of ENSO variability in many geophysical and biological systems. Occasional measurements of South American glacier masses corroborate some of the major ENSO episodes of the past 100 years. For the period back to 1500, records of sailing transit times, newspaper and literary accounts of rainfall and flooding, descriptions of mortality of marine organisms and birds, and other sources have been used to chronicle large-magnitude ENSO warm episodes. Centennial to millennial ENSO variability can be inferred from lake deposit and tree growth bands, and coral, shellfish, and glacier isotope concentrations.

The latter proxy records have been used to document ENSO back to 130,000 years ago (hereafter 130 Ka), and include epochs in which the solar forcing was different from today. ENSO variability appears to have been diminished during 6 Ka, a period when precession of the Earth's orbit enhanced the boreal summer equatorial solar forcing (output) and diminished the boreal winter solar forcing. ENSO episodes tend to develop during the summer, and the enhanced forcing is thought to have inhibited ENSO occurrences by differentially warming the western equatorial Pacific, which leads to increased precipitation in the west, and westward surface winds, stronger upwelling, and colder sea surface temperatures in the east. The global glacier extent in 6 Ka was not too different from today, at least in comparison with the epochs centered on 20 Ka (last glacial maxima) and 140 Ka, during which the ice volumes were five times the current values. The coral record suggests that ENSO variability was diminished relative to today during these icy epochs. Changes in the eccentricity of the Earth's orbit significantly changed the solar forcing during these periods, and the mechanisms by which these slow insolation changes influence ENSO are a subject of current research. *See* CLIMATE HISTORY; INSOLATION.

Teleconnections. The core of the ENSO is in the tropical Pacific, but the phenomenon influences the climate throughout the tropics and the extratropical latitudes of the winter hemisphere. Such planetary-scale atmospheric patterns are referred to as teleconnections and, in the case of ENSO, they are due to changes in the distribution of tropical Pacific precipitation.

Typical teleconnections during the boreal winter of a warm ENSO cycle are shown in **Fig. 4**. Within the tropics, warm ENSO is associated with enhanced precipitation in the central and eastern equatorial Pacific, torrential precipitation in the northern coastal desert of Peru, and diminished precipitation (and sometimes droughts) over Indonesia, northern Australia, northeast Brazil, and southern Africa. The expansion of the region of persistent precipitation into the tropical Pacific also leads to warming of the entire tropical troposphere by about

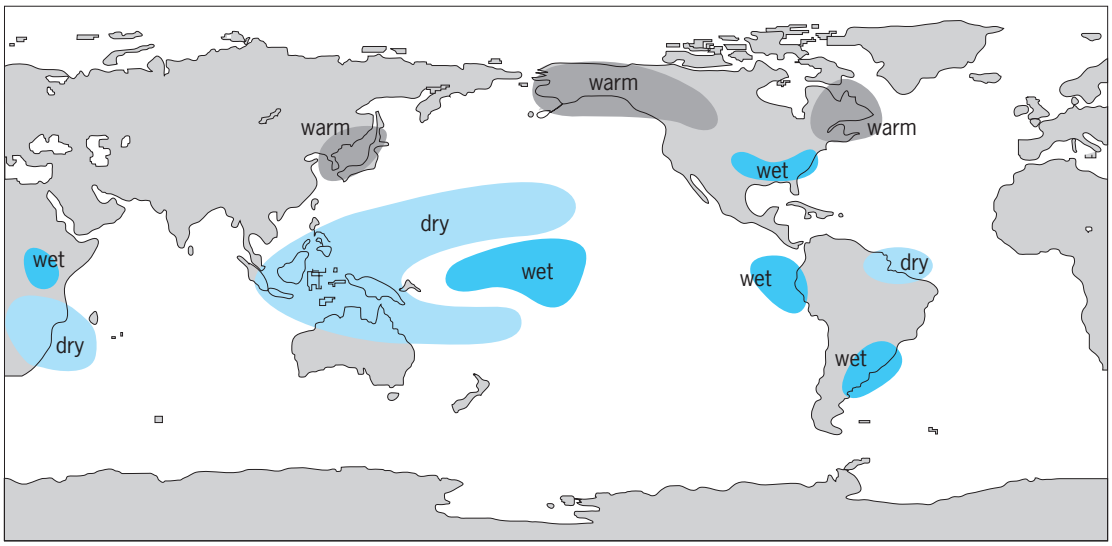

Fig. 4. Teleconnection pattern of warm phases of ENSO to the globe.

1°C (1.8°F). Precipitation deviations of opposite sign are observed during cold ENSO cycles, with the exception of the Peru coast, where it only rains only during warm ENSO. Enhanced atmospheric subsidence during warm ENSO warms the tropical Indian Ocean (**Fig. 5***b*, upper panel) and inhibits the normal February through May rains in northeast Brazil as well as the precipitation that maintains the equatorial Andean glaciers. Warm ENSO is associated with more hurricanes in the eastern Pacific and less in the Caribbean/Atlantic, and vice versa during cold ENSO cycles.

ENSO influences the higher latitudes primarily through changes in the preferred paths of storms, called storm tracks, in the winter hemisphere. Winter storms (weather) and their preferred locations are associated with the mean position of the jet streams, and these storms move heat north and south and produce precipitation in the middle latitudes. The rearrangement of storm tracks by ENSO produces deviations in the seasonal mean climate and in the distribution of extreme daily weather events over North America (Fig. 4), with diminished storminess over Alaska and Canada, and enhanced precipitation over the southern United Sates. Similar relationships exist in the Southern Hemisphere during austral winter with, for example, a southward shift of precipitation from the Andes' Altiplano at 15°S to the Chilean coast during ENSO warm episodes.

The year-to-year, decadal, and secular variability of the tropical Pacific climate all influence the teleconnections and climate impacts on the Northern Hemisphere. Tropical Pacific climate variability on both the year-to-year and longer time scales influence surface temperature and precipitation over the contiguous United States in the regions shown in Fig. 4. The extratropical temperature and precipitation deviations are different for warm and cold ENSO phases and reflect a nonlinear relationship between the storm tracks and the planetary-scale atmospheric circulation in the extratropics. Although the teleconnections from the Pacific are primarily into the winter hemisphere, there is observational and modeling evidence that cool surface-temperature deviations in the tropical Pacific can contribute to persistent summertime droughts in the central United States.

Decadal ENSO. Tropical Pacific climate variability is dominated by the year-to-year swings of the ENSO cycle that are shown in Fig. 3. There are also small-amplitude changes in sea-surface temperatures over broad regions of the Pacific, including the tropics, on time scales longer than the year-to-year variability. Figure 5*a* presents a pair of time series that documents tropical Pacific temperature variability.

Year-to-year ENSO temperature deviations (Fig. 5*a* upper) are characterized by deviations in excess of 0.5°C (0.9°F), centered on the Equator in the central and eastern Pacific, with equatorial upwelling playing a dominant role in producing these temperature changes. Deviations of comparable magnitude and like sign are found along the coasts of Ecuador and Peru, and are due to changes in upwelling. Smaller-magnitude (<0.15°C or 0.27°F) warm deviations occur in the tropical Indian Ocean, and cold deviations (<−0.15°C) are found in the western equatorial Pacific, North Pacific (30°S–50°N), and South Pacific (20°S–50°S). Longer, decadal time-scale surface-temperature deviations (Fig. 5*b* lower) are superficially similar to those for the year-to-year variability, but the extratropical and tropical Pacific deviations are comparable in magnitude (>0.2°C or 0.36°F) on longer time scales, and the central and eastern equatorial temperature deviations extend over a broader latitudinal scale than for the year-to-year variability. For comparison, the secular warming of the oceans, which has been removed from these analyses, is approximately 0.8°C (1.44°F) over the same period. The decadal variability is associated with cooler eastern Pacific conditions in the mid-1940s through late 1970s and warmer conditions in the 1980s and 1990s (Fig. 5*a* lower). This decadal oscillation is related to the Pacific Decadal Oscillation,

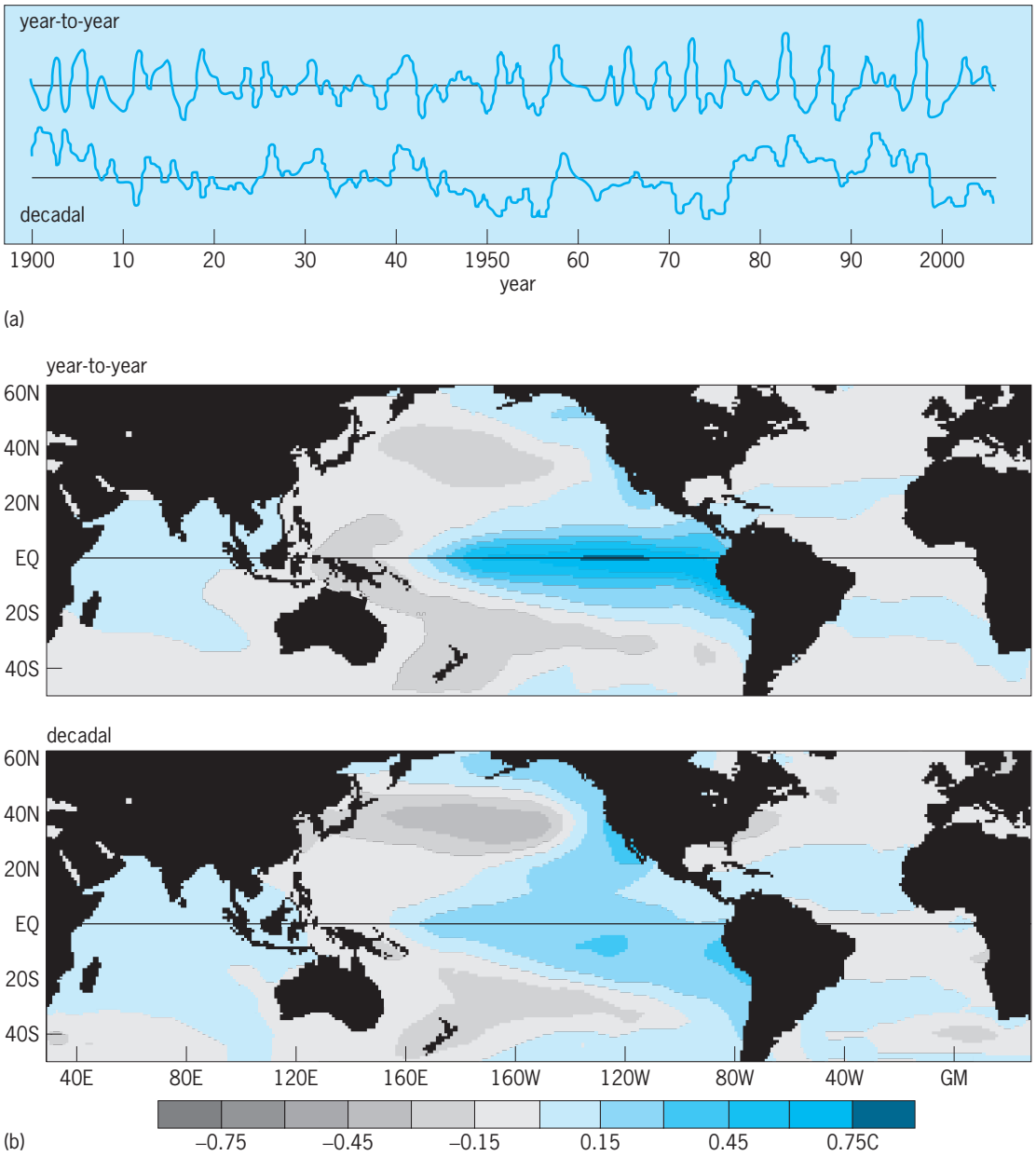

Fig. 5. Time series of (*a*) tropical Pacific ENSO and ENSO-like variations and (*b*) the characteristic global patterns that go with these series.

which has effects similar to those of ENSO on the United States.

The existence of decadal (and longer time scale, say global warming) modulations of ENSO call into question the definition of ENSO anomalies with respect to a single unchanging background climate.

Theories of ENSO. It is impossible to understand ENSO by considering the atmosphere or the ocean alone. The normal tropical Pacific sea-surface temperature (warm in the west, cold in the east) can be understood in terms of the wind blowing westward over the surface of the ocean. Similarly, the normal atmosphere over the tropical Pacific (heavy rainfall in the west, westward winds, low pressure, and ascending air) can be understood in terms of the sea-surface temperature. But to predict the future evolution of ENSO, the winds would have to be known in order to predict the sea-surface temperature, and the sea-surface temperature would have to be known in order to predict the winds.

The only way to simulate the consistent evolution of the atmosphere and ocean is to make coupled numerical models of them. A coupled model is one in which the atmospheric component sees and drives the oceanic component at the same time as the oceanic component sees and drives the atmospheric component. The way to predict the future evolution of ENSO is to start with the atmosphere–ocean system in its current state and allow the coupled system to evolve in a mutually consistent manner. *See* CLIMATE MODELING.

The first such coupled model was constructed in the mid-1980s by S. E. Zebiak and M. A. Cane. It was a simplified model in that only the upper layers

of the ocean were modeled and, more importantly, the annual cycle of sea-surface temperature and surface winds was specified. In this sense, the model was an anomaly model, calculating anomalies with respect to the specified annual cycle. Its importance lay in the fact that it was able to simulate ENSO in ways thought to be relevant. This advance led to a relatively simple explanation of ENSO and made possible the first successful long-term prediction of sea-surface temperature in the eastern Pacific by means of a numerical model of the coupled atmosphere–ocean system. *See* CLIMATIC PREDICTION.

The mechanism of ENSO that proved to be operative in the simplified Zebiak-Cane model relied on instability of the coupled system: when water in the eastern Pacific was warm, eastward winds were forced at the western edge of the warm water. These eastward winds further warm the water by a number of different mechanisms: by advecting warm water from the warmer West Pacific; by deepening the thermocline and therefore quelling the upwelled cold water; and by weakening the westward winds and therefore weakening the upwelling itself. As the instability sets on and the warm water grows, planetary (Rossby) waves are excited that contain the seeds of the reversal of the instability. They propagate westward, are reflected at the western edge of the Pacific, and return to counter the growing warming in the eastern Pacific. The combined sequence of events is called the delayed oscillator mechanism. The period and amplitude of the ENSO cycle depends on the strength of the coupling, the size of the Pacific, and the delay of the Rossby waves. This mechanism is not complete, however. It explains the regular ENSO cycle but not its irregularity as evidenced in Fig. 3. This irregularity has variously been explained by the addition of random noise to the cycle, the nonlinear interaction of the ENSO cycle with the annual cycle, the nonlinear interaction of the ENSO cycle with other growing instabilities, and by many other possibilities. *See* OCEAN WAVES.

A completely different class of theories is that the interaction between the atmosphere and the ocean is not unstable. How can a stable system exhibit ENSO? Although the theory is mathematically recondite, there is a class of dynamics (nonnormal) that can exhibit a growth of disturbances followed by decay. Since all ENSO phases first grow and then decay, this has to be allowed as an explanation of ENSO. A stable coupled atmosphere–ocean system perturbed by random noise has been shown to have properties and predictive power currently indistinguishable from those of unstable theories.

Modeling ENSO. Since ENSO is a coupled atmosphere–ocean phenomenon, the system must be composed of a numerical model of the atmosphere suitably coupled to a numerical model of the ocean. If we want also to model the effects of ENSO on the rest of the world, the models must be global and include land, ice and snow, and so on. The model must be run for a very long time, and the statistics of the annual cycle and ENSO are recorded and compared to historical data such as in Fig. 3. This has been done to very good effect for simplified models where the annual cycle is specified and only anomalies from the annual cycle are calculated. But for fully coupled numerical models where the only input is from the Sun and all else is calculated, the simulation of ENSO has not been so good. The problem is that, due to systematic biases, the simulation of the annual cycle in the tropical Pacific is generally poorly done, and some of the higher-frequency phenomena (such as the intraseasonal oscillations) are also poorly done. These issues generally are agreed to be the dominant problem with the simulation of ENSO. This makes it difficult to examine the changes of ENSO with global warming, but it does not preclude effective predictions since the predictions are initialized by observations. While the systematic biases in the model may tend to pull the prediction back to unrealistic states over time, there will generally be a time in which the predictions are pretty good. It would still be of great value for both short-term predictions (a season or so in advance) and global-warming simulations to improve the models in the manner indicated. Todd Mitchell; Edward S. Sarachik

Bibliography. K. Achuta Rao and K. R. Sperber, Simulation of the El Niño Southern Oscillation, *Clim. Dyn.*, 19:191–209, 2002; A. G. Barnston et al., Long-lead seasonal forecasts—where do we stand?, *Bull. Amer. Meteorol. Soc.*, 75:2097–2114, 1994; A. G. Barnston et al., Multimodel ensembling in seasonal climate forecasting at IRI, *Bull. Amer. Meteorol. Soc.*, 84:1783–1796, 2003; M. A. Cane, S. E. Zebiak, and S. C. Dolan, Experimental forecasts of El Niño, *Nature*, 321:827–832, 1986; H. R. Diaz and U. Markgraf (eds.), *El Niño: Historical and Paleoclimatic Aspects of the Southern Oscillation*, 1993; L. Goddard et al., Current approaches to seasonal-to-interannual climate predictions. *Int. J. Climatol.*, 21:1111–1152, 2001; T. N. Palmer et al., Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER), *Bull. Amer. Meteorol. Soc.*, 85:853–872, 2004; S. G. H. Philander, *El Niño, La Niña, and the Southern Oscillation*, 1990; C. F. Ropelewski and M. S. Halpert, North American precipitation and temperature patterns associated with the El Nino/Southern Oscillation (ENSO), *Mon. Weath. Rev.*, 114:2352–2362, 1986; J. M. Wallace et al., On the structure and evolution of ENSO-related climate variability in the tropical Pacific: Lessons from TOGA, *J. Geophys. Res. Oceans*, 103:14,241–14,260, 1998; S. E. Zebiak and M. A. Cane, A model El Niño/Southern Oscillation, *Mon. Weath. Rev.*, 115:2262–2278, 1987; Y. Zhang, J. M. Wallace, and D. S. Battisti, ENSO-like interdecadal variability 1900–93, *J. Clim.*, 10:1004–1020, 1997.

Elasmobranchii

The subclass within the Chondrichthyes (cartilaginous fishes) that includes the sharks (Euselachii) and the skates and rays (Batoidei). The other subclass within Chondrichthyes, according to traditional

classifications, is the Holocephali (chimaera, or ratfishes). It is probable that both groups arose independently during the Silurian or Early Devonian from a group of extinct armored fishes, the Placodermi. The elasmobranchs are distinguished by separate gill openings, amphistylic or hyostylic jaw suspension, and sensory ampullae (of Lorenzini) in the head region. Characters shared with the holocephalans include a variably calcified cartilaginous endoskeleton, placoid scales, urea-retention mechanism, clasper organs in the male for internal fertilization, and the absence of an air (swim) bladder. *See* CHIMAERIFORMES; CHONDRICHTHYES; PLACODERMI; SCALE (ZOOLOGY); SWIM BLADDER.

Evolutionary development. The history of the elasmobranchii can best be understood in terms of three successive evolutionary levels: cladodont, hybodont, and modern. Because of their incomplete fossil record, it is impossible to work out the detailed phylogenetic relationships of the extinct and living groups.

Cladodonts. The cladodont sharks first appeared in the Devonian; except for the specialized pleuracanths (**Fig. 1***a*), which persisted into the Triassic, they had disappeared by the end of the Pennsylvanian. The cladodonts all possess amphistylic jaw suspension. Their teeth have a pointed central cusp, two or more lateral cusps, and a flattened disklike base. The notochord was persistent and was not replaced by centra.

The basal part of the pectoral fin skeleton is variable in design but never tribasal. The radial cartilages of the pectoral and pelvic fins are not divided and extend more or less to the fin margin. The pelvic plates are separated. Claspers are usually present. The dorsal fins are preceded by spines. The caudal fin is heterocercal, nearly equilobate, and the radials of the hypochordal lobe are unsegmented. The body scales are usually multilobed, each lobe having a separate pulp cavity. Some cladodont sharks possess relatively large semicircular canals; in others these are the same size as in modern sharks. *See* COPULATORY ORGAN; EAR (VERTEBRATE).

Cladodonts were primarily pelagic predators, capable of seizing and tearing prey. The relatively stiff fins suggest that their maneuverability was more limited than in modern sharks. Representative cladodonts include the Devonian genera *Cladodus* and *Cladoselache* (Fig. 1*b*). *See* DEVONIAN.

Hybodonts. The hybodonts first appeared in the Mississippian Period. Although their skeleton resembles that of the cladodonts, they were more advanced in having only a tribasal pectoral fin skeleton, divided and shortened paired fin radials, clearly differentiated anal fin, and ribs and hemal arches along the entire length of the unreduced notochord. The acquisition of the tribasal pectoral fin, present in all later sharks, and a more flexible hypochordal lobe in the caudal fin apparently enabled these sharks to maintain their equilibrium more effectively than the cladodonts.

Some hybodonts had cladodont-like teeth and presumably fed on swimming prey, but much of the adaptive radiation at this level involved the development of crushing or grinding dentitions related mainly to bottom feeding. The late Palaeozoic genus *Ctenacanthus* is a good example of a conservative hybodont; the Mesozoic *Hybodus* (Fig. 1*c*) has a more advanced fin structure and a crushing dentition. *Ctenacanthus* lived from the Mississippian into the Permian, and *Hybodus* from the Triassic through the Cretaceous.

Modern elasmobranchs. Modern sharks arose during the Jurassic Period long before the hybodonts had become extinct in the Late Cretaceous. The transition from the hybodont to the modern level is poorly documented by fossils. However, there are three groups of aberrant (part hybodont, part modern) living sharks. Heterodontidae (Port Jackson sharks), Chlamydoselachidae (frilled sharks), and Hexanchidae (six-gill and seven-gill sharks). All except the Port Jackson sharks are rare deep-water forms, and each group has a somewhat different combination of hybodont and modern shark characters. These archaic sharks suggest the kind of evolutionary experimentation that took place during the rise of the modern sharks. *See* JURASSIC; SELACHII.

The skates and rays (batoids) are mostly bottom-dwelling elasmobranchs that possess many features distinct from the sharks. Present evidence suggests that the batoids arose during the Jurassic from some benthic hybodont group.

Modern features. Morphologically the modern shark level may be defined by the following features:

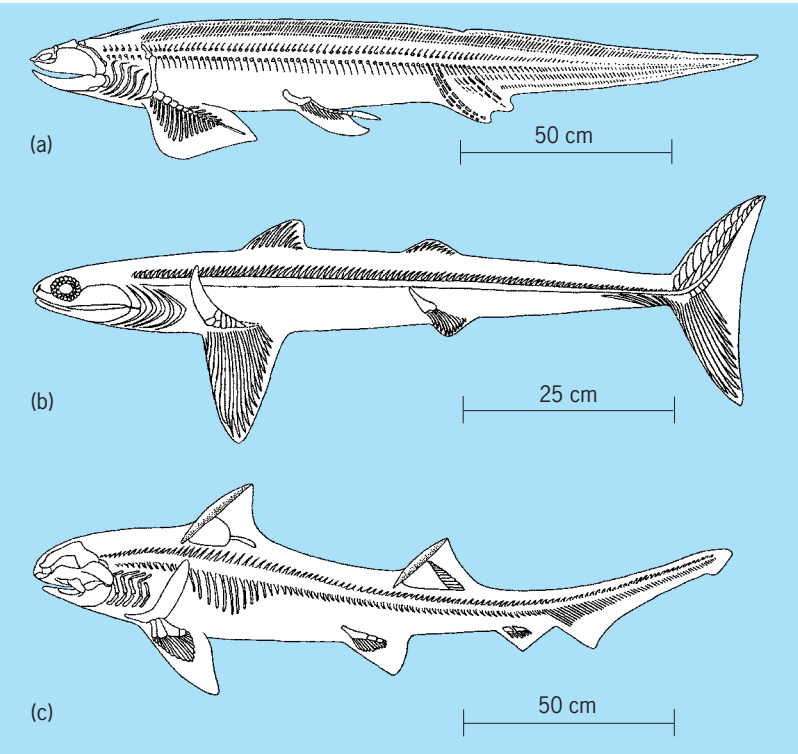

Fig. 1. Lateral views of fossil sharks. (*a*) *Pleurocanthus*, a specialized cladodont. (*b*) *Cladoselache*, a cladodont. (*c*) *Hybodus*, a hybodont.

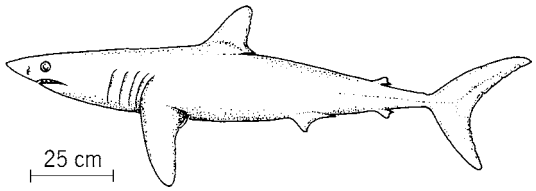

Fig. 2. Mackeral shark (*Isurus*). (*After H. B. Bigelow and W. C. Schroeder, Fishes of the Western North Atlantic, pt. 1, 1948*)

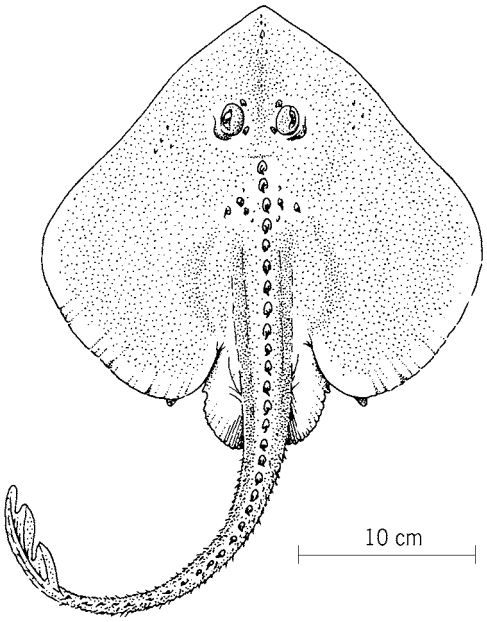

Fig. 3. Skate (*Raja*). (*After H. B. Bigelow and W. C. Schroeder, Fishes of the Western North Atlantic, pt. 2, 1953*)

hyostylic jaw suspension, shortened jaws, jaw protrusion mechanism, complete replacement of the notochord by calcified centra, fusion of the pelvic plates, and presence of only single-lobed placoid scales. Although the rostrum was short and blunt in the cladodonts, it became elongated in some hybodonts and, with a few exceptions, it projects well beyond the mouth in all modern sharks (**Fig. 2**).

Skates and rays differ from the sharks in having a depressed body, modification in the mode of gill ventilation, freer hyostylic jaw suspension, elaboration of the pavement dentition for crushing or grinding, enlargement of the pectoral fins, and disappearance of the anal fin. With the pectoral fins assuming the main role in locomotion, the posterior part of the body, including the caudal fin, became reduced to whiplike proportions (**Fig. 3**). *See* BATOIDEI; SELACHII. Bobb Schaeffer

Bibliography. H. B. Bigelow and W. C. Schroeder, *Fishes of the Western North Atlantic*, pt. 1, 1948, pt. 2, 1953; C. E. Bond, *Biology of Fishes*, 2d ed., 1997; E. S. Herald, *Living Fishes of the World*, 1961; F. H. Pough et al., *Vertebrate Life*, 5th ed., 1998; B. Schaeffer, Comments on elasmobranch evolution, in P. W. Gilbert et al. (eds.), *Sharks, Skates and Rays*, 1967.

Elasticity

The property whereby a solid material changes its shape and size under the action of opposing forces, but recovers its original configuration when the forces are removed. The theory of elasticity deals with the relations between the forces acting on a body and the resulting changes in configuration, and is important in many branches of science and technology, for instance, in the design of structures, in the theory of vibration and sound, and in the study of the forces between atoms in crystal lattices.

Elastic constants. The forces acting on a body are expressed as stresses and measured as force per unit area. Thus if a bar $ABCD$ of square cross section (**Fig. 1***a*) is fixed at one end and subjected to a force F uniformly distributed over the other end DC, the stress is $F/(DC)^2$. This stress causes the bar to become longer and thinner and to assume the shape $A'B'C'D'$. The strain is measured by the ratio (change in length)/(original length), that is, by $(B'C' - BC)/(BC)$. According to Hooke's law, stress is proportional to strain, and the ratio of stress to strain is therefore a constant, in this case Young's modulus, denoted by E, so that $E = F(BC)/(DC)^2\ (B'C' - BC)$. *See* HOOKE'S LAW; STRESS AND STRAIN; YOUNG'S MODULUS.

Poisson's ratio σ is defined as the ratio of lateral strain to longitudinal strain, so that $\sigma = BC(DC - D'C')/DC\ (B'C' - BC)$. The bar of Fig. 1*a* is in a state of tension, and the stress is tensile; if the force F were reversed in direction, the stress would be compressive. Stresses of this type are called direct or normal stresses; a second type of stress, known as tangential or shear stress, is illustrated in Fig. 1*b*. In this case, the configuration $ABCD$ becomes $ABC'D'$, with the shear forces F acting in the directions AB and CD. The shear strain is measured by the angle θ, and if the body is originally a cube, the shear stress is $F/(DC)^2$. The ratio of stress to strain, $F/(DC)^2\theta$, is the shear or rigidity modulus G, which measures the

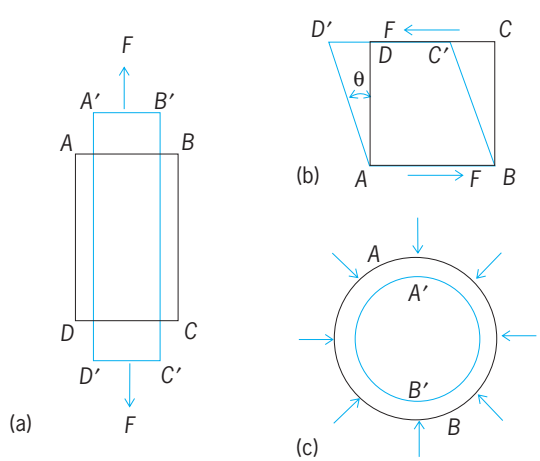

Fig. 1. Stresses on a bar. (*a*) Direct or normal stress. (*b*) Tangential or shear stress. (*c*) Change in volume with no change in shape. (All deformations are exaggerated.)

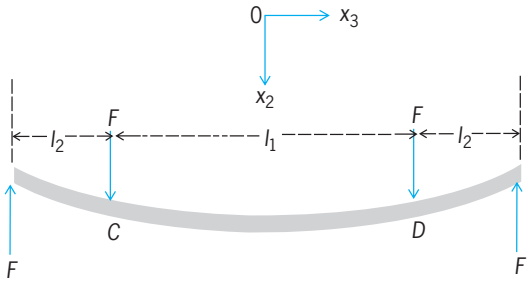

Fig. 2. Flexure and torsion in a bar.

resistance of the material to change in shape without change in volume.

A further elastic constant, the bulk modulus k, measures the resistance to change in volume without changes in shape, and is illustrated in Fig. 1*c*. The original configuration is represented by the circle AB, and under a hydrostatic (uniform) pressure P, the circle AB becomes the circle $A'B'$. The bulk modulus is then $k = P\upsilon/\Delta\upsilon$, where $\Delta\upsilon/\upsilon$ is the volumetric strain. The reciprocal of the bulk modulus is the compressibility.

Determination of values. The elastic constants may be determined directly in the way suggested by their definitions; for instance, Young's modulus can be determined by measuring the relative extension of a rod or wire subjected to a known tensile stress. Less direct methods are, however, usually more convenient and accurate. Prominent among these are the dynamic methods involving frequency of vibration and velocity of sound propagation. The elastic constants can be expressed in terms of frequency of (or velocity in) regularly shaped specimens, together with the dimensions and density, and by measuring these quantities, the elastic constants can be found. *See* ULTRASONICS.

The elastic constants can also be determined from the flexure and torsion of bars. As an illustration, consider a bar AB (**Fig. 2**) of breadth b (in the x_1 direction) and depth d (in the x_2 direction) supported by forces F at the ends, and loaded symmetrically by forces F at points C and D. Over the portion CD there is a uniform bending moment $M = Fl_2$, and the theory of bending shows that the portion CD is bent into the arc of a circle such that Eq. (1) applies, where R

$$R = \frac{EI}{M} \tag{1}$$

is the radius of curvature, E is Young's modulus, and I is the moment of inertia of cross section, equal to $bd^3/12$ for a rectangular cross section. The longitudinal stress at the lower face of the bar is tensile, and at the upper face, compressive. The middle plane of the bar is free of stress, and is the neutral axis. The stress at a distance x_2 from the neutral axis is shown in Eq. (2).

$$T = \frac{Ex_2}{R} = \frac{Mx_2}{I} \tag{2}$$

It is thus possible to determine E from Eq. (1) by measuring I, R, and M; conversely, if E is known, the stress may be determined from Eq. (2). *See* LOADS, TRANSVERSE.

Practical limitations. In practice, stress is only proportional to strain, and the strain is only completely recoverable within certain limits, called the elastic limits of the material. The stress below which the strain is completely recoverable is sometimes called the limit of perfect elasticity, and the stress up to which Hooke's law is obeyed is sometimes called the proportional limit or limit of linear elasticity. Above the elastic limits, the material is subject to time-dependent effects, and as the stress is further increased, the ultimate strength of the material is approached. *See* PLASTICITY; STRENGTH OF MATERIALS.

Theory of elasticity. In classical elasticity theory, it is assumed that the strains are always small; Hooke's law is therefore obeyed; the strains are completely recoverable and, moreover, are superposable, so that the strain produced by the joint action of two or more stresses is the sum of the strains produced by them individually.

In order to develop the theory, it is necessary to specify the stresses and strains more closely. **Figure 3** shows the stress components T_{ij} (where i, j may take the values 1, 2, or 3) acting on the faces of a cube, parallel to the coordinate axes x_1, x_2, x_3. The first suffix indicates the direction of the stress component and the second the direction of the normal to the plane under consideration. Stresses of the type T_{11} are normal stresses, and of the type T_{12}, shear stresses. The conditions for zero rotation of the cube are $T_{12} = T_{21}$, $T_{13} = T_{31}$, $T_{23} = T_{32}$, and there are therefore six independent stress components.

In addition to the stresses T_{ij}, body forces proportional to volume (for instance, forces due to the weight of the body) may also be acting. If the stresses T_{ij} vary with position, application of Newton's second law leads to Eq. (3) for the x_1 direction where ρ

$$\frac{\partial T_{11}}{\partial x_1} + \frac{\partial T_{12}}{\partial x_2} + \frac{\partial T_{13}}{\partial x_3} + X_1 = \rho f_1 \tag{3}$$

is the density, f_1 is the acceleration, and X_1 the body force component per unit volume along x_1, together with two similar equations for the x_2 and x_3 directions. If $f_1 = f_2 = f_3 = 0$, these equations become the equations of equilibrium, and if, further, $X_1 = X_2 = X_3 = 0$, they become the equations of equilibrium in the absence of body forces. The preceding

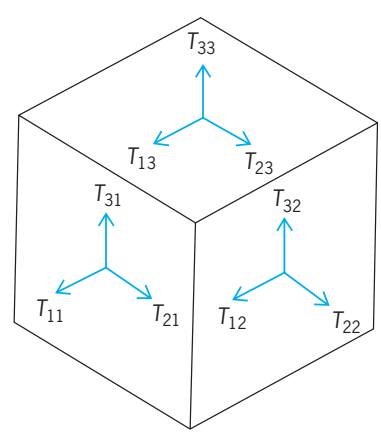

Fig. 3. Stress components acting on the faces of a cube.

equations are important in many branches of elastic theory and, for example, provide a starting point in the study of vibrating bodies and of the twisting of cylinders and prisms with cross sections of various shapes. *See* TORSION.

The components of strain are specified in a similar way to the stresses. There are six independent strain components: S_{11}, S_{22}, S_{33}, S_{23}, S_{13}, and S_{12}. If, as a result of strain, the coordinates of a point x_1, x_2, x_3 become x_1+u_1, x_2+u_2, x_3+u_3, the quantities u_1, u_2, and u_3 are the components of the displacement vector, and the strain components are as displayed in Eq. (4), so that, for example the relations shown by Eq. (5) would hold true.

$$S_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{4}$$

$$S_{11} = \frac{\partial u_1}{\partial x_1} \qquad S_{12} = \frac{1}{2}\left(\frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1}\right) \tag{5}$$

By eliminating the displacements from these equations, the so-called compatibility equations are obtained with three of the type shown in Eq. (6) and three of the type shown in Eq. (7).

$$\frac{\partial^2 S_{22}}{\partial {x_3}^2} + \frac{\partial^2 S_{33}}{\partial {x_2}^2} = 2\frac{\partial^2 S_{23}}{\partial x_2 \partial x_3} \tag{6}$$

$$\frac{\partial^2 S_{11}}{\partial x_2 \partial x_3} = \frac{\partial}{\partial x_1}\left(-\frac{\partial S_{23}}{\partial x_1} + \frac{\partial S_{13}}{\partial x_2} + \frac{\partial S_{12}}{\partial x_3}\right) \tag{7}$$

The stresses and strains have so far been denoted by two suffixes. This is essential if the methods of tensor analysis are to be applied to elasticity problems, but for many purposes a single suffix notation is adequate. The change from a two- to a one-suffix notation for the stresses is simply $T_{11}=T_1$, $T_{22}=T_2$, $T_{33}=T_3$, $T_{23}=T_4$, $T_{13}=T_5$, $T_{12}=T_6$. The change of notation for the strains is $S_{11}=S_1$, $S_{22}=S_2$, $S_{33}=S_3$, $2S_{23}=S_4$, $2S_{13}=S_5$, $2S_{12}=S_6$; the factor 2 is required to make the strains S_4, S_5, and S_6 conform with the usual definition of shear strain (Fig. 1*b*).

Hooke's law generalized. Hooke's law may be generalized to the statement that each stress component is proportional to each strain component, equivalent to the six equations (8*a*) which may be written more concisely as Eq. (8*b*), where the summation extends

$$\begin{aligned} T_1 &= c_{11}S_1 + c_{12}S_2 + c_{13}S_3 + c_{14}S_4 + c_{15}S_5 + c_{16}S_6 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ T_6 &= c_{61}S_1 + c_{62}S_2 + c_{63}S_3 + c_{64}S_4 + c_{65}S_5 + c_{66}S_6 \end{aligned} \tag{8a}$$

$$T_q = \sum_r c_{qr}S_r \tag{8b}$$

over $r = 1, 2, 3, 4, 5$, and 6. The elastic constants c_{qr} are termed the stiffnesses; there are altogether 36 of them but they are subject to the reciprocal relations $c_{qr}=c_{rq}$ imposed by thermodynamic requirements, and the number is thus reduced to 21.

Additional relations can be derived from the three assumptions that the interatomic forces act along the lines joining the centers of atoms in the lattice, that the atoms are situated at centers of symmetry, and that the lattice is initially at zero stress. These relations, called Cauchy relations, are $c_{23}=c_{44}$, $c_{13}=c_{55}$, $c_{12}=c_{66}$, $c_{14}=c_{56}$, $c_{25}=c_{46}$, $c_{45}=c_{36}$ and, if true, would reduce the number of stiffnesses to 15. Experiment shows, however, that they are not true in general; nevertheless, their investigation provides an indication of the extent to which these three assumptions hold in any particular case.

The generalized Hooke's law can also be written to express the strains in terms of the stresses given in Eq. (9), in which the quantities s_{qr} are the elastic compliances. If the six simultaneous equations of Eq. (8) are solved for the strains, the compliances are obtained in terms of the stiffnesses as in Eq. (10), where Δc is the determinant shown in Eq. (11) and

$$S_q = \sum_r s_{qr}T_r \qquad (r = 1, 2, 3, 4, 5, 6) \tag{9}$$

$$s_{qr} = \frac{\Delta c_{qr}}{\Delta c} \tag{10}$$

$$\Delta c = \begin{vmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ c_{12} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ c_{13} & c_{23} & c_{33} & c_{34} & c_{35} & c_{36} \\ c_{14} & c_{24} & c_{34} & c_{44} & c_{45} & c_{46} \\ c_{15} & c_{25} & c_{35} & c_{45} & c_{55} & c_{56} \\ c_{16} & c_{26} & c_{36} & c_{46} & c_{56} & c_{66} \end{vmatrix} \tag{11}$$

Δc_{qr} is the cofactor obtained by deleting the row and column containing c_{qr} from the determinant Δc.

The 21 stiffnesses (or compliances) of the generalized Hooke's law describe the elastic behavior of a material belonging to the triclinic crystal system. The existence of symmetry elements reduces the number of independent elastic constants in the other crystal systems to the following numbers: monoclinic, 13; orthorhombic, 9; tetragonal, 7 or 6; trigonal, 7 or 6; hexagonal, 5; and cubic, 3. Materials belonging to all of these systems are anisotropic, and the elastic properties depend upon direction within the material. If the properties are independent of direction, the material is isotropic and its elastic behavior is completely described by two independent stiffnesses (or compliances). *See* CRYSTALLOGRAPHY.

The stress-strain relations, referred to the principal axes in the orthorhombic, hexagonal, cubic, and isotropic systems, are given in the **table**. The equations involving the compliances are completely analogous, with S and T interchanged, and s_{qr} written for c_{qr}, except where $T_q = {}^1/_2\,(c_{11} - c_{12})\,S_q$, in which case $S_q = 2(s_{11} - s_{12})\,T_q$.

Rochelle salt is an example of an orthorhombic crystal; materials which, although not crystalline, possess the same symmetry and matrix of elastic constants as orthorhombic crystals are said to be orthotropic. Wood and plywood are materials of this description, and orthotropic elastic theory has also been applied to laminated plastics and reinforced concrete.

Single-crystal zinc, cobalt, magnesium, and ice are hexagonal materials; they are transversely isotropic because the properties are independent of direction

Stress-strain relations

Stress	Strain: Orthorhombic	Hexagonal	Cubic	Isotropic
$T_1 =$	$c_{11}S_1 + c_{12}S_2 + c_{13}S_3$	$c_{11}S_1 + c_{12}S_2 + c_{13}S_3$	$c_{11}S_1 + c_{12}S_2 + c_{12}S_3$	$c_{11}S_1 + c_{12}S_2 + c_{12}S_3$
$T_2 =$	$c_{12}S_1 + c_{22}S_2 + c_{23}S_3$	$c_{12}S_1 + c_{11}S_2 + c_{13}S_3$	$c_{12}S_1 + c_{11}S_2 + c_{12}S_3$	$c_{12}S_1 + c_{11}S_2 + c_{12}S_3$
$T_3 =$	$c_{13}S_1 + c_{23}S_2 + c_{33}S_3$	$c_{13}S_1 + c_{13}S_2 + c_{33}S_3$	$c_{12}S_1 + c_{12}S_2 + c_{11}S_3$	$c_{12}S_1 + c_{12}S_2 + c_{11}S_3$
$T_4 =$	$c_{44}S_4$	$c_{44}S_4$	$c_{44}S_4$	$(c_{11} - c_{12})S_4/2$
$T_5 =$	$c_{55}S_5$	$c_{44}S_5$	$c_{44}S_5$	$(c_{11} - c_{12})S_5/2$
$T_6 =$	$c_{66}S_6$	$(c_{11} - c_{12})S_6/2$	$c_{44}S_6$	$(c_{11} - c_{12})S_6/2$

in all of the planes normal to the hexagonal axis.

Single-crystal copper, gold, silver, nickel, and the alkali halides (for example, sodium chloride) are important cubic materials. The stress-strain equations are derived from those of the orthorhombic system by superimposing the condition that the three principal directions are all equivalent. This does not mean that the properties are independent of direction; for example, the compliance s_{11}' in an arbitrary direction is given by Eq. (12), where a_1, a_2, a_3 are the

$$s'_{11} = s_{11} - 2(s_{11} - s_{12} - s_{44}/2) \cdot \left(a_1^2 a_2^2 + a_2^2 a_3^2 + a_3^2 a_1^2\right) \quad (12)$$

cosines of the angles between the arbitrary direction and the cubic axes. This equation shows that s'_{11} depends on orientation unless $s_{44}/2 = s_{11} - s_{12}$. *See* PHOTOELASTICITY. R. F. S. Hearmon

Bibliography. A. P. Boresi, R. J. Schmidt, and O. M. Sidebottom, *Advanced Mechanics of Material*, 5th ed., 1993; S. F. Borg, *Fundamentals of Engineering: Elasticity*, 1990; J. E. Marsden and T. J. Hughes, *Mathematical Foundations of Elasticity,* 1983, reprint 1994; S. P. Timoshenko and J. N. Goodier, *Theory of Elasticity,* 3d ed., 1970.

Electret

A solid dielectric with a quasipermanent electric moment. Electrets may be classified as real-charge electrets and dipolar-charge electrets. Real-charge electrets of present importance are either solid dielectrics with charges of one polarity at or near one side of the dielectric and charges of opposite polarity at or near the other side, or cellular dielectrics with charges of opposite polarity on the two sides of the voids. The dipolar-charge electrets are dielectrics with aligned dipolar charges. Some dielectrics are capable of storing both real and dipolar charges. Examples of charge arrangements in electrets are shown in the **illustration**. *See* POLARIZATION OF DIELECTRICS.

Modern electrets used in research and in applications are often films ranging in thickness from the submicrometer scale for silicon-based materials to sheets of 5–50-μm (0.2–2-mil) thickness (foil electrets) for suitable polymer materials. They are frequently metallized on one or both sides, depending on the intended use. In many applications of the solid real-charge electrets, the external fields of the charged dielectric are used and the films are metallized on only one side to allow the electric field to emanate from the nonmetalized side. In applications of the cellular real-charge electrets and the dipolar electrets, the piezoelectric effect generated by such samples is used and the films are metallized on both surfaces. *See* PIEZOELECTRICITY.

Solid real-charge electrets. In these electrets, the intrinsic conductivity and carrier mobility must be small enough that excess charges can persist for long periods of time. Suitable materials are the fluorocarbons polytetrafluoroethylene (PTFE) and its copolymer fluoroethylenepropylene (FEP), as well as polypropylene (PP), polycarbonate (PC), polyimide (PI), polyethylene terephthalate (PETP), and inorganic materials such as silicon dioxide and aluminum oxide. PETP can also assume a considerable dipole polarization. *See* POLYMER.

Charging of real-charge electrets, metallized on one side, can be achieved by exposure of the nonmetallized side to a corona discharge; by application

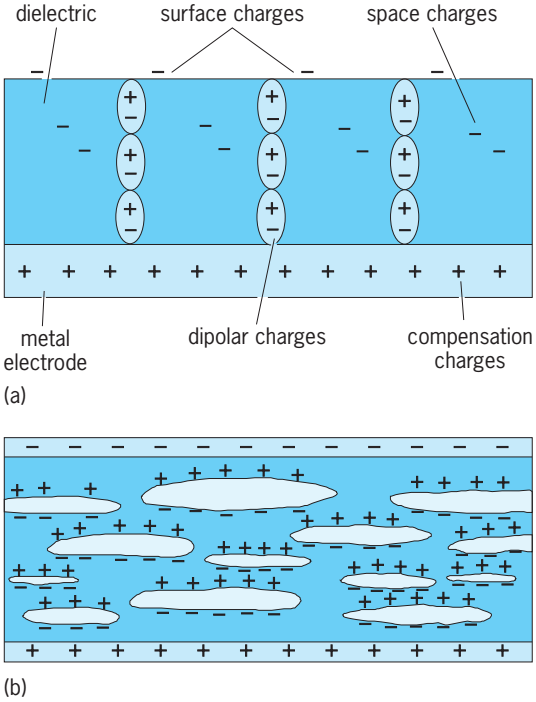

Schematic cross sections of electrets. (*a*) Solid electret disk metallized on one side with real charges (surface and space charges) and dipolar charges. (*b*) Cellular electret metallized on both sides with charges of opposite polarities on the upper and lower surfaces of the voids.

of an electric field at elevated temperatures through an electrode opposing the nonmetallized side of the dielectric; or by contacting the nonmetallized side with a biased, wet electrode. Another charging technique, allowing greater control than other methods, is electron injection with a partially penetrating electron beam. In all cases, carriers of one polarity are deposited onto or injected into the insulator and trapped, and a compensation charge of equal magnitude but opposite sign flows into the sample electrode. Thus, the charge arrangement exhibits an electric moment. Another method, applied to photoconductors metallized on both sides with transparent electrodes, consists in the application of an electric field during irradiation with ultraviolet or visible light. Carrier generation by the light and charge separation by the field yield the desired electric moment, resulting in a photoelectret. *See* CORONA DISCHARGE; PHOTOCONDUCTIVITY.

PTFE and FEP are electronegative and therefore hold negative charges better than positive ones. Real-charge densities of up to 5 millicoulombs/m^2 (3.2 microcoulombs/in.^{2}) can be stored in these polymers. Electrets are stable over long periods of time. At the customarily used charge densities of about 0.2 millicoulombs/m^2 (130 nanocoulombs/in.^{2}), corresponding to surface potentials of approximately 250 V in a 25-μm-thick (1-mil) sample, the time constant of the decay of the surface potential of PTFE and FEP is about 200 years if storage is at room temperature and low relative humidity under shielded conditions. This time constant drops to about 10 years at 50°C (122°F) and 90% relative humidity.

Silicon dioxide (SiO_2) electrets are preferably made of a thermally grown silicon dioxide layer, usually less than 1 μm (0.04 mil) thick, on a silicon substrate. Negative charges deposited onto the free oxide surface show better stability than negative volume charges or positive charges. Surface-charging techniques, such as the corona method, are therefore utilized. After charging, a compensation charge of opposite polarity is located at the silicon–silicon dioxide interface. Charge densities are generally chosen in the range 10–20 mC/m^2 (6–12 μC/in.^{2}), corresponding to surface potentials of 300–600 V on a 1-μm-thick layer. Time constants of the charge decay exceed those for PTFE and FEP.

Cellular real-charge electrets. These newer electrets with piezoelectric properties consist of a cellular dielectric, typically a polymer such as PP, with voids of the order of a few micrometers height and lateral dimensions of several tens of micrometers. Such voids can be generated by biaxially stretching a polymer film filled with small inorganic particles. During the stretching, microcracks form around the particles and develop into flat voids. These can be enlarged in the thickness direction by a pressure treatment, consisting of an increase of the pressure of the surrounding gas such that gas penetrates into the voided structure. A subsequent pressure release inflates the material. Thereafter, the sample is metalized on one of its surfaces.

Charging of such a sample is achieved by exposure of its unmetallized surface to a corona discharge. The depositied charges together with their countercharges on the metallization generate a high internal field that will cause breakdowns in the voids and thus deposition of charges of opposite sign on the internal void surfaces. Finally, the sample is metallized on the other surface. The material now exhibits a large piezoelectric d_{33} coefficient that is very stable at room temperature. Problematic is the thermal depolarization at elevated temperatures found in the PP samples. Other materials with higher depolarization temperatures are now under study. The samples also show hysteresis behavior as a function of the applied field and are therefore often referred to as ferroelectrets.

Dipolar-charge electrets. In these electrets, the relaxation time of the molecular reorientation has to be large enough that the dipole alignment persists for long periods of time. Most prominent among such materials are poly(vinylidenefluoride) [PVDF] and vinylidenefluoride-trifluoroethylene copolymers [VDF-TrFE], which exhibit piezoelectric and pyroelectric properties linked to the dipole polarization. Inorganic piezoelectric crystals and ceramics also belong to this category but are not discussed here. *See* PIEZOELECTRICITY; PYROELECTRICITY.

The preferred technique for the polarization of dipolar electrets is also the above corona method applied to one-sided metallized samples. Again, the injected charges set up a large field within the dielectric that aligns the dipoles. Another method consists in the direct application of a high field to a two-sided metallized sample at an elevated temperature. After the polarization process, the samples exhibit sizable d_{31} and d_{33} coefficients.

In the case of PVDF, polarization densities of up to 0.1 C/m^2 (65 μC/in.^{2}) can be achieved by corona charging at room temperature with fields of up to 400 MV/m (10 MV/in.) on samples prestretched three to five times their original length. The stretching makes the material, which is semicrystalline, recrystallize in a polar form and also reorients the lamellar crystals so that the dipoles are more precisely aligned by the applied field.

PVDF has a relatively stable polarization at room temperature. At 100°C (212°F), the polarization decays initially (within minutes) to half its original value and stabilizes thereafter. Similar stability is achieved for the piezoelectric activity of this material.

Comparison of piezoelectric properties. Typical d_{33} coefficients for cellular PP and solid PVDF are compared in the **table** with values for the conventional piezoelectric materials quartz and lead zirconate

Comparison of piezoelectric properties

Piezoelectric material	d_{31}, pC/N	d_{33}, pC/N
Quartz	0.73 (d_{14})	2.3 (d_{11})
PZT-5	170	370
PVDF	25	20
Cellular PP (expanded)	1	600

titanate (PZT-5) . Since the electret materials are polymers, they have certain other properties such as mechanical softness and flexibility that distinguish them from the conventional piezoelectrics. Particularly noteworthy is the huge d_{33} coefficient of cellular PP. The electrets have already found numerous commercial applications.

Applications. Important commercial applications of real-charge electrets are in electroacoustic and electromechanical transducers, air filters, electret dosimeters, and adhesive posters. Of particular significance are electret microphones, which are made in quantities of more than 2×10^9 annually and account for 90% of the microphones produced. Also of interest are biological applications based on the blood compatibility of charged polymers or on their favorable influence on wound or fracture healing. Commercial applications of dipolar electrets are in piezoelectric transducers for airborne and waterborne sound and in pyroelectric detectors. *See* AIR FILTER; DIELECTRIC MATERIALS; DOSIMETER; ELECTRET TRANSDUCER; ELECTRICAL INSULATION; MICROPHONE.

Gerhard M. Sessler

Bibliography. S. Bauer, R. Gerhard-Multhaupt, and G. M. Sessler, Ferroelectrets: Soft electroactive foams for transducers, *Phys. Today*, 57(2): 37–43, 2004; Institute of Electrical and Electronics Engineers, *IEEE 12th International Symposium on Electrets*, 2005; V. Kestelman, L. Pinchuk, and V. Goldade, *Electrets in Engineering*, Kluwer Academic, 2000; G. M. Sessler and R. Gerhard-Multhaupt (eds.), *Electrets*, vols. 1 and 2, 3d ed., Laplacian Press, 1999.

Electret transducer

A device for the conversion of acoustical or mechanical energy into electrical energy, and vice versa, which utilizes a quasipermanently charged or polarized dielectric material (electret). Examples are certain microphones, headphones, hydrophones, and ultrasonic devices. Depending on the electret material and the usage, either the external electric field of real-charge electrets or the piezoelectric activity of dipolar electrets is used. In the simplest implementation, a transducer based on real-charge electrets consists of a metal backplate (first electrode) covered by a mechanically tensioned diaphragm. The diaphragm is a film electret carrying a metal coating (second electrode) on the side facing away from the backplate. Provisions are made to maintain a shallow air gap between electret and backplate. The air gap is occupied by an electrostatic field originating from the electret charges. Upon acoustical or mechanical deflection of the diaphragm, such a device generates an electrical output signal between its two electrodes; similarly, application of an electrical signal results in diaphragm deflections. Such electret devices are therefore self-biased electrostatic or condenser transducers. They thus exhibit all the advantages of this transducer class, such as wide dynamic range and flat response over a frequency range of several decades, without requiring the external bias necessary in conventional transducers of this kind. On the other hand, a transducer based on the piezoelectric activity of a dipolar electret may consist of just the electret film covered by two electrodes. Compression or bending of the film causes again an electrical output signal between the electrodes, while the application of a voltage to the electrodes results in mechanical action. *See* ELECTRET; ELECTROSTATICS; TRANSDUCER.

Film electret. Suitable polymer films are electrically charged or polarized to produce an external electric field or piezoelectric activity, respectively. Materials capable of quasipermanently trapping charge carriers (that is, storing charges over periods of years or decades) include polytetrafluoroethylenepropylene (PTFE), its copolymer polyfluoroethylenepropylene (FEP or Teflon®), and polypropylene (PP), while substances capable of carrying a quasipermanent dipole polarization include poly(vinylidenefluoride) [PVDF] and some of its copolymers. A polymer of considerable interest is cellular PP, which can carry charges on the surfaces of its internal voids and therefore exhibits a very large piezoelectric activity. The cellular PP loses its charges at elevated temperatures and is therefore suitable only for room-temperature use. Before charging or poling, all the materials are either metallized on one side or backed up with a metal electrode.

Electret microphones. The most widely used electret transducer is the electret microphone. This is a condenser microphone consisting of two electrodes separated by an electret and an air gap. The electret generates an electric field in the air gap. If one of the electrodes is designed as a diaphragm that vibrates in response to an incident sound wave, an output signal proportional to the sound pressure is generated.

Such microphones are designed as transducers with a backplate electret or a diaphragm electret. The backplate-electret microphones, which are the more common transducers, consist of a (6–12-μm) (0.25–0.5-mil) Teflon layer adhering to a metallic or metal-coated backplate and charged to 200–400 microcoulombs/m^2 (130–260 nanocoulombs/in.^{2}), corresponding to an external bias of about 100 V. The surface of this electret is placed next to a metallic or metallized diaphragm, leaving a shallow air layer of about 20 μm (0.8 mil) thickness. The stiffness of the air layer can be decreased (and thus the sensitivity of the microphone can be improved) by connecting the air layer to a larger cavity by means of small holes through the backplate. The electrical output of the microphone is taken between the backplate, which is insulated from the outer case, and the metal part of the diaphragm. The output is fed into a high-impedance preamplifier. The diaphragm-electret microphones typically consist of a 12–25-μm-thick (0.5–1-mil) Teflon or PTFE film, metallized on one side and charged on the other side to charge densities similar to the above-mentioned values. A backplate and an intermediate air gap complete the system. These transducers are simpler and less expensive than the backplate-electret microphones but mechanically not as stable. This is attributed to stress relaxation of the tensioned Teflon diaphragm, which

causes small changes in sensitivity.

At frequencies below the resonance frequency, the acoustical properties of electret microphones are largely governed by the restoring force on the diaphragm. Since the mechanical tension of the diaphragm is generally kept at a relatively low value (about 10–50 newtons/m or 0.7–3.5 lb/ft), the restoring force is determined by the compressibility of the air layer. Controlling the restoring force by the air layer is advantageous because changes in tension due to stress relaxation thus have only a minor effect on the sensitivity, which is largely independent of transducer area.

Under open-circuit conditions, a displacement d of the diaphragm of an electret microphone causes a frequency-independent output voltage given by Eq. (1), where E is the (constant) electric field in

$$v = Ed = \frac{\sigma D_1 d}{\varepsilon_0(D_1 + \varepsilon D_2)} \tag{1}$$

the air layer of the transducer; σ is the charge density of the electret projected onto its surface; D_1 and D_2 are the thicknesses of electret film and air layer, respectively; ε_0 is the permittivity of free space; and ε is the relative dielectric constant of the electret material. As in conventional electrostatic transducers, the displacement is proportional to the applied pressure in a wide frequency band extending from a lower cutoff given by a pressure-equalization leak in the back cavity to an upper cutoff determined by the resonance frequency. The voltage response for constant sound pressure is frequency-independent in this range.

Typical electret microphones designed for the audio-frequency range have constant sensitivities of 10–50 mV/pascal (1–5 mV/microbar or 70–350 V/lb · in.$^{-2}$) in the frequency range 20 to 15,000 Hz. Nonlinear distortion is less than 1% for sound-pressure levels below 140 dB, and the impulse response is excellent, owing to the flat amplitude and phase characteristics. Other properties of electret microphones are their low sensitivity to vibration, owing to the small diaphragm mass, and their insensitivity to magnetic fields. Compared with conventional electrostatic transducers, electret microphones have the following advantages: they do not require a dc bias; they have three times higher capacitance per unit area, resulting in a better signal-to-noise ratio; and they are not subject to destructive arcing between foil and backplate in humid atmospheres and under conditions of water condensation.

Various electret microphones for operation at infrasonic and ultrasonic frequencies, covering the range from 0.001 Hz to 200 MHz, have been designed by properly positioning the upper and lower cutoff frequencies discussed above. Furthermore, transducers with directional characteristics such as cardioid, bidirectional, toroidal, and second-order unidirectional have been developed. *See* MICROPHONE.

Electret headphones. Another group of electret transducers made of real-charge electrets are headphones. The above microphone designs can also be utilized in this case. These transducers produce a certain amount of nonlinear distortion due to the partially quadratic dependence of the force F per unit area, exerted on the diaphragm, on the applied ac voltage v, as can be seen from Eq. (2). High-voltage

$$F = \frac{-\frac{1}{2}(D_1\sigma + \varepsilon\varepsilon_0 v)^2}{\varepsilon_0(D_1 + \varepsilon D_2)^2} \tag{2}$$

sensitivities have been achieved by making the air-gap thickness small. The distortion can be eliminated by the use of push-pull transducers and monocharge electrets.

For a push-pull transducer, a diaphragm consisting of two electrets, each metallized on one side and with these metal layers in contact, is sandwiched between two perforated metal electrodes, forming a symmetrical system. Application of a signal voltage $\bar{v}$ in antiphase to the two electrodes causes the electrodes to exert forces F_1 and F_2 on the diaphragm. These forces are given by Eq. (2) if $\pm\bar{v}$ is substituted for v. The net force $F = F_1 - F_2$ on the diaphragm is determined by Eq. (3), which is linear in $\bar{v}$, indicat-

$$F = \frac{-2\bar{v}\sigma\varepsilon D_1}{(D_1 + \varepsilon D_2)^2} \tag{3}$$

ing the absence of nonlinear distortion. The linearity, however, is maintained only if both air-layer thicknesses remain at the same value D_2. This is possible only if D_2 is relatively large, which makes such systems less sensitive than the above-described single-ended transducers.

This transducer can be simplified and improved by using an asymmetric system in which a single electret coated on one side with a low-conductivity layer is used. The diaphragm holds an excess charge because of the presence of an induction charge on the front electrode. Thus, application of an ac signal causes a net force between diaphragm and plates that is linear if the resistivity of the coating (or the resistor R) prevents charge equalization during a period of lowest-frequency signal applied.

An interesting modification of the above push-pull transducers can be achieved by using nonmetallized monocharge electrets, which differ from customary electrets in that they carry only a single-polarity charge and no compensation charge. It is possible to form these electrets by electron injection or by the use of wet-contact charging methods. Since a compensating charge is absent, monocharge electrets have a strong external field that is independent of electrode distance. Thus, if such electrets are used in transducer configurations, large air gaps of the order of 1 mm (0.04 in.) can be used without loss in field strength. This in turn results in highly efficient transducers that can operate at large displacement amplitudes. *See* EARPHONES.

Electromechanical transducers. Apart from their use in electroacoustic transducers, such as microphones and earphones, real-charge electrets have been applied to electromechanical transducers. Examples are touch or key transducers, impact-sensitive line transducers, relay switches, and optical display panels.

Relatively close in their design principles to the

electroacoustic transducers described above are the first two of these groups. The touch or key transducers depend on the manual deflection of an electret relative to an electrode in a microphonelike structure. Large-voltage signals can be generated with such devices. Line transducers consist essentially of a coaxial cable with polarized dielectric in which the center conductor and the shield serve as electrodes. Mechanical excitation resulting in a deformation of the shield at any point along the length of such a cable produces an electrical output signal.

Very different design principles are used in the relay switches and the optical display panels. The relay switches utilize the external fields of electrets mounted in bistable arrangements to open or close contacts. Such devices have power requirements 1000 times smaller than equivalent electromagnetic systems; thus considerable power conservation over present magnetic components is possible. A similar design is used in the optical display panels where illuminated channels are opened and closed by hinged and opaque electrets moved by electrostatic forces. Although relatively high voltages are required, the necessary currents are small and flow only during the switching operation. *See* ELECTRONIC DISPLAY; RELAY.

Hydrophones based on PVDF. Many types of hydrophones based on the piezoelectric polymer PVDF and its copolymers have been described. Among these are hydrophones for the sonic and near-ultrasonic frequency ranges based on the use of thick sheets or cylinders of the piezoelectric material, as well as hydrophones for the higher ultrasonic range, generally designed as membrane or needle hydrophones.

The lower-frequency hydrophones can be relatively simple devices consisting of an electroded film of the piezoelectric polymer that is encapsulated for waterproofing. If the dimensions of the hydrophone are small compared to the wavelength of the sound, equal sound pressure in all directions prevails and the output signal of the transducer is given by the hydrostatic piezoelectric constant. More sophisticated designs include stiffening plates to avoid resonances due to extensional modes of the polymer film, two-layer constructions to improve shielding, and cylindrical designs. Arrays or transducers with shaped elements allow directional characteristics to be obtained. Hydrophones based on these principles generally operate in the frequency range from less than 1 kHz to several MHz.

The membrane hydrophones for the higher ultrasonic range, up to 100 MHz, also consist of one or two peripherally clamped thin sheets of piezoelectric polymer with a diameter of several centimeters, poled in a tiny central section of less than 1 mm (0.04 in.) in diameter, which is the actual sensing area. These hydrophones have the advantage of being more or less transparent to sound, thus allowing sound pressure to be measured without disturbing the sound field by generating unwanted reflections. In the needle hydrophones, a metallic needle serves as the support and inner electrode for a surrounding layer of the piezoelectric polymer, which is covered on its outer side with another electrode. The outside diameter of the needle is generally less than 1 mm (0.04 in.) to 0.3 mm (0.012 in.), thus permitting again measurements without perturbation of the sound field. The frequency range of needle hydrophones can extend up to 80 MHz. *See* HYDROPHONE.

Piezoelectric cellular transducers. The very large piezoelectric thickness coefficient of charged cellular PP can be used to sense sound, force, impact, vibration, or motion. For example, cellular PP has been used in microphones of very simple design. Such transducers consist essentially of only the PP film and suitable housing and shielding. If stacked layers of film are used, the microphone sensitivity increases proportionally to the number of layers. Using this technique, microphones consisting of six layers have been built that show sensitivities of 15 mV/Pa, comparable to the sensitivity of conventioanl electret microphones. Other features of these transducers are relatively flat frequency response, low noise level, and low distortion. The major advantage is that a minute and critical air gap, as used in electret microphones, is not needed. Cellular PP can also be used for sound generation. Since the generated sound pressure increases with frequency, sound sources are particularly suitable for the ultrasonic range, below the resonance frequency of the PP film that is in the region above 100 kHz.

Important electromechanical transducers based on cellular PP are piezoelectric floor mats, keyboards, control panels, pickups for musical instruments, and biomedical sensors. The floor mats are step- or impact-sensitive and can detect the motion of patients in hospitals, of intruders in industrial plants or military installations, or of vehicles on highways. Because of the inexpensive material, large floor areas may be covered. Push-buttons of cellular PP for keyboards and control panels have the advantage of high sensitivity and mechanical flexibility. The biomedical sensors are used to monitor, for example, heartbeat, breathing, and forces on limbs in humans and animals.

Applications. Because of their favorable properties, simplicity, and low cost, electret transducers have been used in many applications as research tools and in the commercial market.

Among the research applications are microphones for use in acousto-optic spectroscopy, applied to the detection of air pollution and to the study of reaction kinetics of gases and optical absorption of solids. Because of the favorable noise performance of electret microphones, the detection threshold for air pollutants has been lowered by more than an order of magnitude. Other applications of electret microphones have been in aeronautics and shock-tube studies, in which the low vibration sensitivity of these transducers is crucial. The wide frequency range of electret transducers, discussed above, made possible their application in infrasonic atmospheric studies and in ultrasonic investigations of liquids and solids. A broad field of application is in noise

monitoring and noise abatement. Examples are all kinds of mobile and stationary noise measurements as well as active noise reduction. In addition, audio and ultrasonic arrays of electret microphones have been used in near-field acoustic holography. Yet other uses of electromechanical electret transducers are in such diverse areas as microforce detection, vibration analysis, and leak detection in space stations. For example, microforce sensors based on PVDF are used in the micromanipulation and microassembly of MEMS (micro-electro-mechanical systems) structures. In all these applications, the simplicity and reliability of electret transducers are of importance. *See* ACOUSTICAL HOLOGRAPHY; ACTIVE SOUND CONTROL; INFRASOUND; MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS); NOISE MEASUREMENT; PHOTOACOUSTIC SPECTROSCOPY; SHOCK TUBE; ULTRASONICS.

Of all commercial applications of electret devices, the electret microphone for amateur, professional, studio, and telephone use is most prominent. About 2×10^9 such microphones are produced annually, accounting for about 90% of the entire output of microphones. Almost half of these are used in stationary and mobile telephones. Other uses of electret microphones are in high-fidelity work, such as audio studios, and in MP-3 players, sound-level meters, noise dosimeters, camcorder, and toys. Miniature microphones are used in hearing aids, in telephone operator headsets, and in communication headsets, often as noise-canceling (first-order gradient) transducers. The success of the electret microphone is primarily due to its acoustic quality and low cost. It is noteworthy that, owing to their low vibration sensitivity, electret microphones were the first transducers to be built directly into widely used tape recorders. In addition to electret microphones, there are single-backplate and push-pull electret headphones, also with monocharge electrets. The latter are among the highest-quality headsets available. *See* HEARING AID; MAGNETIC RECORDING; SOUND RECORDING; TELEPHONE.

Apart from their use in electroacoustics, electret devices have found many commercial applications in the underwater, ultrasonic, and electromechanical fields. The PVDF hydrophones in the varieties discussed above for the sonic and ultrasonic frequency ranges have been commercially available for a long time and, because of their many advantages, enjoy an undisputed position as the preferred devices in the field. The same holds for the ultrasonic devices for airborne sound based on space-charge and piezoelectric polymers. In the electromechanical field, transducers based on PVDF and on cellular piezoelectric PP have been introduced commercially. In particular, step-sensitive floor mats, touch-sensitive keyboards and control panels, and biomedical force sensors are available.

Gerhard M. Sessler

Bibliography. Institute of Electrical and Electronics Engineers, *IEEE 12th International Symposium on Electrets*, 2005; G. M. Sessler and R. Gerhard-Multhaupt (eds.), *Electrets*, vols. 1 and 2, 3d ed., Laplacian Press, 1999.

Electric charge

A basic property of elementary particles of matter. One does not define charge but takes it as a basic experimental quantity and defines other quantities in terms of it. The early Greek philosophers were aware that rubbing amber with fur produced properties in each that were not possessed before the rubbing. For example, the amber attracted the fur after rubbing, but not before. These new properties were later said to be due to "charge." The amber was assigned a negative charge and the fur was assigned a positive charge.

According to modern atomic theory, the nucleus of an atom has a positive charge because of its protons, and in the normal atom there are enough extranuclear electrons to balance the nuclear charge so that the normal atom as a whole is neutral. Generally, when the word charge is used in electricity, it means the unbalanced charge (excess or deficiency of electrons), so that physically there are enough "nonnormal" atoms to account for the positive charge on a "positively charged body" or enough unneutralized electrons to account for the negative charge on a "negatively charged body."

The rubbing process mentioned "rubs" electrons off the fur onto the amber, thus giving the amber a surplus of electrons, and it leaves the fur with a deficiency of electrons. *See* ELECTROSTATICS.

In line with the previously mentioned usage, the total charge q on a body is the total unbalanced charge possessed by the body. For example, if a sphere has a negative charge of 1×10^{-10} coulomb, it has 6.24×10^8 electrons more than are needed to neutralize its atoms. The coulomb is the unit of charge in the SI system of units. *See* COULOMB'S LAW; ELECTRICAL UNITS AND STANDARDS.

The surface charge density σ on a body is the charge per unit surface area of the charged body. Generally, the charge on the surface is not uniformly distributed, so a small area ΔA which has a magnitude of charge Δq on it must be considered. Then σ at a point on the surface is defined by the equation below.

$$\sigma = \lim_{\Delta A \to 0} \frac{\Delta q}{\Delta A}$$

The subject of electrostatics concerns itself with properties of charges at rest, while circuit analysis, electromagnetism, and most of electronics concern themselves with the properties of charges in motion. *See* CAPACITANCE; ELECTRIC CURRENT; ELECTRICITY.

Ralph P. Winch

Electric contact

A part, in an electrical switching device, made of conducting material, for the purpose of closing, opening, or changing the conductive path of an electrical circuit. To open or close a circuit, an electric contact is made to come in contact with or separate from its mating part. Devices embodying contacts for these

purposes are electric switches, relays, contactors, and circuit breakers. Contacts may be actuated directly or through a linkage that is driven either manually, mechanically, electromagnetically, hydraulically, or pneumatically.

An electric contact is also an essential part of a variable resistor. In such an application the contact or wiper provides a moving connection to the resistive element. For such service the contact material is chosen for low abrasion combined with absence of a high-resistance film such as would be formed by an adherent oxide. *See* POTENTIOMETER; RHEOSTAT.

Requirements for contacts differ markedly depending on applications. For contacts in relays and low-power applications, reliability in completing a circuit may be of the utmost importance. Surface films and contaminants are not tolerable. Noble metals such as gold and platinum are sometimes sealed hermetically for this purpose. For the purpose of carrying high continuous currents, electrical resistance must be kept to a minimum. Contacts with high silver content are usually used. High contact forces, up to hundreds of pounds, may be applied when the contacts are in the closed position. For contacts that must interrupt high currents, refractory materials such as tungsten and molybdenum are often used. These materials also possess antiwelding characteristics, which are essential for contacts that must close in on short-circuit currents and later open under normal operating force.

On many occasions all the aforementioned properties are desired in one material. Special alloys such as silver cadmium oxide or refractory materials impregnated with silver have been developed for this purpose.

A related electrical part is the brush on rotating machines. Pairs of brushes provide continuity between the external circuit and slip rings or commutator segments on the rotating portion of the machine. Similar arrangements of brushes and slip rings connect to continuously rotating radar antenna mounts. Graphite, because of its inherent lubricating nature, is a usual constituent of such sliding contacts for high-current operation; corrosion-resistant metals or metal alloys are used for low-current applications. *See* CIRCUIT (ELECTRICITY); ELECTRIC ROTATING MACHINERY.

Thomas H. Lee

Bibliography. ARS Electronics, *Receiving Tube Characteristics and Socket Connection Guide*, 1981; CES Industries Inc. Staff, *Contactor Sensor Operation*, 1982; R. Holm and E. Holm, *Electric Contacts*, 4th ed., 1967; Institute of Electrical and Electronics Engineers, *1998 IEEE 44th Holm Conference on Electrical Contacts*, 1998.

Electric current

The net transfer of electric charge per unit time. It is usually measured in amperes. The passage of electric current involves a transfer of energy. Except in the case of superconductivity, a current always heats the medium through which it passes. For a discussion of the heating effect of a current *see* JOULE'S LAW.

On the other hand, a stream of electrons or ions in a vacuum, which also may be regarded as an electric current, produces no local heating. Measurable currents range in magnitude from the nearly instantaneous 10^5 or so amperes in lightning strokes to values of the order of 10^{-16} A, which occur in research applications.

All matter may be classified as conducting, semiconducting, or insulating, depending upon the ease with which electric current is transmitted through it. Most metals, electrolytic solutions, and highly ionized gases are conductors. Transition elements, such as silicon and germanium, are semiconductors, while most other substances are insulators.

Electric current may be direct or alternating. Direct current (dc) is necessarily unidirectional but may be either steady or varying in magnitude. By convention it is assumed to flow in the direction of motion of positive charges, opposite to the actual flow of electrons. Alternating current (ac) periodically reverses in direction.

Conduction current. This is defined as the transfer of charge by the actual motion of charged particles in a medium. In metals the current is carried by free electrons which migrate through the spaces between the atoms under the influence of an applied electric field. Although the propagation of energy is a very rapid process, the drift rate of the individual electrons in metals is only of the order of a few centimeters per second. In a superconducting metal or alloy the free electrons continue to flow in the absence of an electric field after once having been started. In electrolytic solutions and ionized gases the current is carried by both positive and negative ions. In semiconductors the carriers are the limited number of electrons which are free to move, and the "holes" which act as positive charges.

Displacement current. When alternating current traverses a capacitor, there is no physical flow of charge through the dielectric (insulating material, gas, or vacuum), but the effect on the rest of the circuit is as if there were a continuous flow. Energy can pass through the capacitor by means of the so-called displacement current, and this displacement current generates a surrounding magnetic flux, just as an orthodox current would. James Clerk Maxwell introduced the concept of displacement current in order to make complete his theory of electromagnetic waves. *See* ALTERNATING CURRENT; CAPACITANCE; CONDUCTION (ELECTRICITY); DIELECTRIC MATERIALS; DIRECT CURRENT; DISPLACEMENT CURRENT; ELECTRIC INSULATOR; ELECTRICAL RESISTANCE; FREE-ELECTRON THEORY OF METALS; SEMICONDUCTOR; SUPERCONDUCTIVITY.

John W. Stewart

Electric distribution systems

Systems that comprise those parts of an electric power system between the subtransmission system and the consumers' service switches. It includes distribution substations; primary distribution feeders;

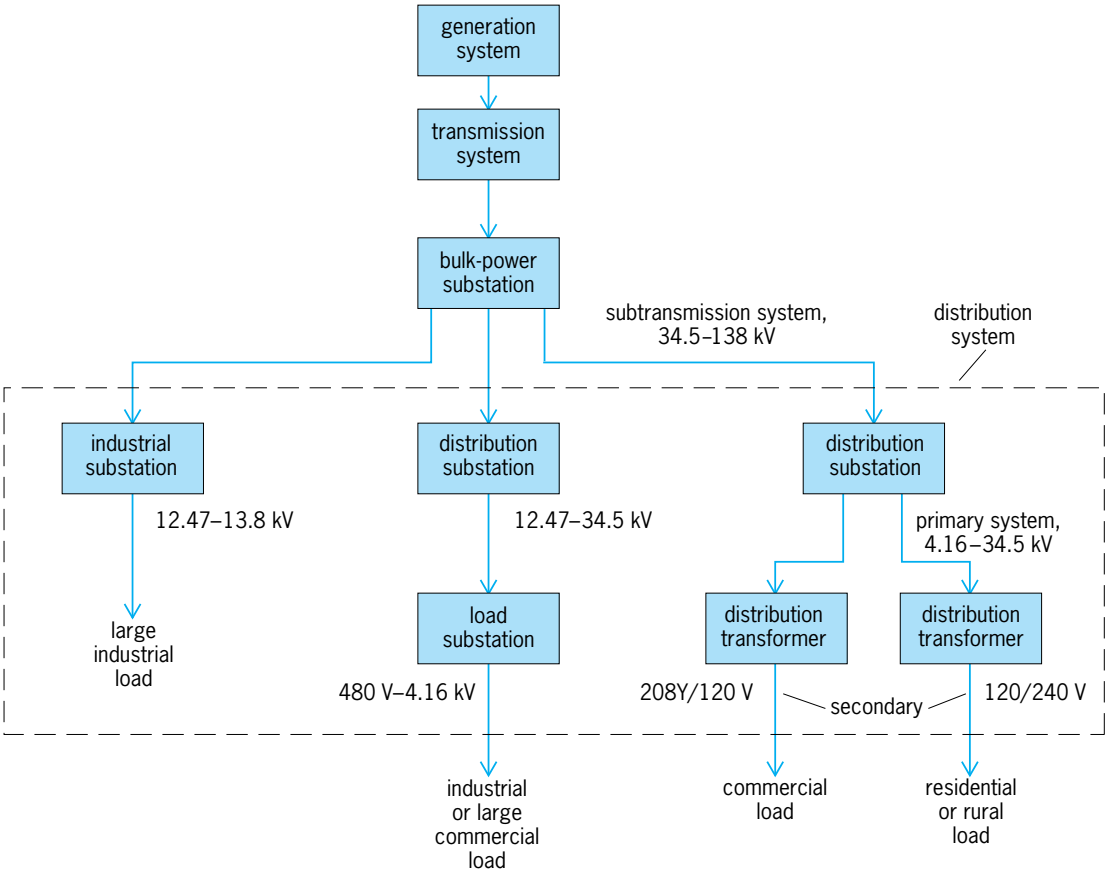

Fig. 1. Overview of the power system from generation to consumer's switch.

distribution transformers; secondary circuits, including the services to the consumer; and appropriate protective and control devices. Sometimes, the subtransmission system is also included in the definition.

The subtransmission circuits of a typical distribution system (**Fig. 1**) deliver electric power from bulk power sources to the distribution substations. The subtransmission voltage is usually between 34.5 and 138 kV. The distribution substation, which is made up of power transformers together with the necessary voltage-regulating apparatus, bus-bars, and switchgear, reduces the subtransmission voltage to a lower primary system voltage for local distribution. The three-phase primary feeder, which usually operates at voltages from 4.16 to 34.5 kV, distributes electric power from the low-voltage bus of the substation to its load center, where it branches into three-phase subfeeders and three-phase and occasionally single-phase laterals. Most of the three-phase distribution system lines consist of three-phase conductors and a common or neutral conductor, making a total of four wires. Single-phase branches (made up of two wires) supplied from the three-phase mains provide power to residences, small stores, and farms. Loads are connected in parallel to common power-supply circuits. *See* ALTERNATING CURRENT.

Subtransmission. The subtransmission system is that part of the power system which delivers power from bulk power sources, such as large transmission substations. The subtransmission may be made up of overhead open-wire construction on wood poles or underground cables. Although the voltages of these circuits can occasionally range from 12.47 to 345 kV, the majority are at the 69-, 115-, and 138-kV voltage levels. There is a trend toward the use of the higher voltages as a result of the increasing use of higher transmission voltages. The subtransmission-system designs vary from simple radial systems to a subtransmission network. The major considerations affecting the design are cost and reliability. *See* CONDUCTOR (ELECTRICITY); ELECTRIC POWER TRANSMISSION.

Substation. The distribution substation is an assemblage of equipment for the purpose of switching, regulating, and changing the supply voltage from subtransmission to primary distribution. More important substations are designed so that the failure of a piece of equipment in the substation or one of the subtransmission lines to the substation will not cause an interruption of power to the customer. The location of a substation is dictated by the voltage levels, voltage regulation considerations, subtransmission costs, substation costs, and the costs of primary feeders, mains, and distribution transformers. It is also restricted by esthetic considerations. A typical substation may have power transformers, circuit breakers, disconnecting switches, station buses and insulators, current-limiting reactors, shunt reactors, current transformers, potential transformers, capacitor voltage transformers, coupling capacitors,

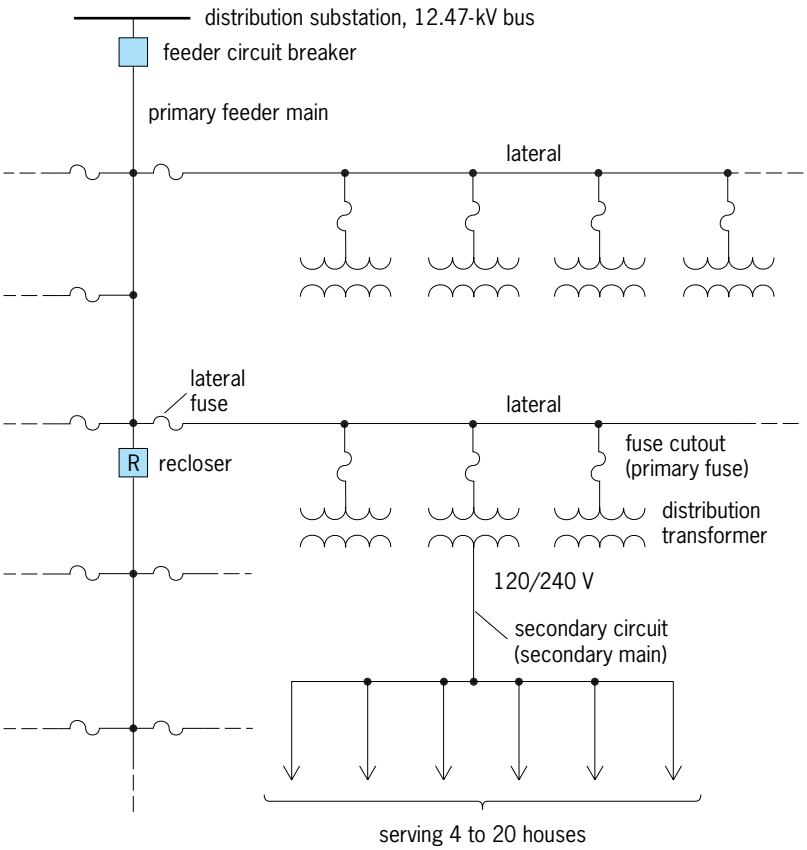

Fig. 2. Typical power distribution feeder.

protective relays, lightning arresters, station batteries, and other equipment. *See* ELECTRIC POWER SUBSTATION.

Primary system. The part of the electric utility system which is between the distribution substation and distribution transformers is called the primary system. It is made up of circuits known as primary feeders or distribution feeders. A feeder includes the main or main feeder, which usually is a three-phase four-wire circuit, and branches or laterals, which usually are single-phase or three-phase circuits, tapped off the main (**Fig. 2**). A feeder is usually sectionalized by means of reclosing devices at various locations so as to remove as little as possible of a faulted circuit and to hinder service to as few consumers as possible. This goal is achieved through the coordination of all the fuses and reclosers.

The primary system leaving the substation is most often in the 11,000–15,000-V range. A specific voltage used is 12,470-V line-to-line and 7200-V line-to-neutral (conventionally written 12,470Y/7200 V). Some utilities still use a lower voltage, such as 4160Y/2400 V. However, the use of primary distribution circuits in the 25- and 35-kV classes is increasing; all are four-wire systems. Single-phase loads are connected line-to-neutral on the four-wire systems. Congested and heavy-load locations in metropolitan areas as well as new residential areas are customarily served by underground primary feeders made up of radial three-conductor cables.

The primary system design is affected by the cost, importance of the load, and the required level of service continuity and reliability. The simplest and least expensive, and thus the most common, is the radial design. The low reliability of this design is improved by a modification called the loop-type primary design, in which the feeder loops through the feeder load area and returns to the substation bus. The most expensive and reliable design is the primary-network type, in which a system of interconnected feeders is supplied by a number of substations. Such a system supplies a load from several directions and consequently provides the maximum reliability and quality of service.

Secondary system. This is the part of the electric power system between the primary system and the consumer's property. The secondary distribution system includes distribution transformers, secondary circuits (secondary mains), consumer services (or service drops), and meters to measure consumer energy consumption. Generally, the secondary systems are designed in single phase for areas of residential customers and in three phase for areas of industrial or commercial customers with high load densities. Secondary voltages are provided by distribution transformers that are connected to the primary system, and they usually are associated with utilization voltages. Residential and most rural customers are supplied by 120/240-V single-phase three-wire systems. Commercial and small industrial customers are supplied by 208Y/120-V or 480Y/277-V three-phase four-wire systems. The secondary voltage also supplies multiple street lights.

Spot networks. These are used in downtown areas for high-rise buildings with extremely high load densities, and sometimes also for areas of industrial or commercial customers. Often, large shopping centers are supplied by spot networks. In such a network, all transformers and protecting apparatus are placed at the same location.

Good voltage. Essentially, good voltage means that the average voltage level is correct, that variations are within acceptable voltage limits, and that sudden and momentary changes in voltage level do not cause objectionable light flicker. Of course, the utilization voltage of the customer changes with changing load on the system as well as the location of the customer on the system. However, voltage variation is usually less than 5% at the customer's meter.

Good continuity. Service continuity is the provision of uninterrupted electric power to the customer. Good continuity means that such service is almost always provided.

Back-up systems. To increase the service reliability for critical loads, such as hospitals, computer centers, and crucial industrial loads, back-up systems, such as emergency generators or batteries, with automatic switching devices are provided. Computer installations are backed up by uninterruptible power supplies (UPS). *See* UNINTERRUPTIBLE POWER SUPPLY.

Overhead construction. Most existing distribution systems in residential, industrial, and rural areas are

of overhead construction. The distribution transformers are mounted near the tops of poles, and bare primary and secondary conductors are strung from pole to pole.

Underground construction. Almost all residential developments now being built are served by underground residential distribution (URD) systems using underground cables for esthetic reasons. The URD system is free from service outages and accompanying repair expenses caused by lightning, rain, ice, sleet, snow, and wind storms. Also, the probability of hazard caused by building fires and by motor vehicles and other foreign objects coming into contact with the system elements is far less than what might be expected on overhead systems. Other costs, such as the cost of tree trimming, are totally eliminated by using the URD systems. In general, underground distribution systems cost more than comparable overhead systems. However, greater public interest in the esthetic appearance of residential communities and reductions in the cost of underground equipment and installations have made the use of URD systems popular. Residential underground systems normally do not have the refinements and operating advantages found in conventional downtown commercial underground systems. Extensive duct banks, transformer vaults, access holes, and submersible equipment are not required in residential systems.

Cables. In the past the most frequently used underground cable was stranded copper with an oil-impregnated paper insulation wrapped in a lead sheath. It is expensive to buy, install, and splice, but it has a long service life. It is available in single- and three-phase cable and in voltages up to 46 kV. Practical residential underground distribution was made possible by the development of direct-buried underground cable that used a relatively inexpensive plastic insulation, typically cross-linked polyethylene (XLPE), and practical elbow connectors. The insulated cable has stranded conductors surrounding a bare concentric neutral. Other insulating plastics include high-molecular-weight polyethylene (HMWP), sometimes called high-density track-resistant polyethylene (HDPE), and ethylene propylene rubber (EPR). The use of such plastics as the protective sheath for cables has made practical their direct burial in the ground without the need for ducts and access holes. Splicing procedures for such cables are also simple and less costly. Long lengths of cable capable of being plowed directly into the ground or placed in narrow and shallow trenches naturally reduce installation and maintenance costs. Further savings may be realized if other facilities, such as telephone, cable television, gas, and water pipes, are installed simultaneously.

System design. In general, the heavy three-phase feeders are located overhead along the periphery of a residential development, and the laterals to the pad-mount transformers are buried about 40 in. (100 cm) deep. The lateral cable conductors come up from underground into the high-voltage compartment of the distribution transformer. The secondary service lines then run to the individual dwellings at a depth of about 24 in. (60 cm), and come up into the dwelling meter through a conduit. The service conductors run along easements and do not cross adjacent property lines.

Faults. The frequency of faults is lower on underground systems than on overhead systems. Also, faults are not so likely to "bunch up" due to storm conditions. However, faults are much more difficult and time-consuming to find, isolate, and repair on underground systems. Service-restoration requirements have dictated primary-system designs which operate as a normally open loop. In the event of a cable fault, such a design facilitates the location and isolation of the failure and faster service restoration to all customers on the unfaulted segment of the primary loop. A well-designed URD system should permit ease of sectionalizing in order to isolate a faulted portion, and should be flexible enough to accept load growth with a minimum number of changes. The looped primary and single-phase banked secondary system provide these requirements at the lowest cost. The banked secondary system takes better advantage of the diversity among loads and allows the use of transformers with a smaller kilovolt-ampere capacity and possibly fewer transformers. It also offers better average voltage conditions at the loads and less voltage dip due to motor starting for a given transformer rating, conductor size, and spacing between transformers than the simple-radial system offers. *See* VOLT-AMPERE.

Transformers. The transformers used for the URD system are hermetically sealed against moisture, including their bushings and terminals. The terminals may contain one or more insulated disconnecting elbows, enabling the disconnection of the transformer and the sectionalizing of the primary circuits. The transformers may be completely buried or installed on ground-level pads or semiburied pads. Pad-mounted transformers are the predominant type of transformer in URD systems. They are usually installed on concrete slabs, or pads. They may utilize "dead-front" configuration with separable insulated connectors, or elbows. Many other combinations of pad-mount construction and accessory equipment are available. Residential subsurface transformers are used much less frequently. They are installed in relatively tight-fitting vaults with the cover grating of the vault at ground level. Direct-buried transformers are not popular because of the problem of location of accessory equipment and its operation under adverse weather conditions such as snow and ice. *See* TRANSFORMER.

Dispersed storage and generation. Electric distribution systems are undergoing major changes to accommodate the development of alternative sources for generating electric energy and enhanced opportunities for small power producers and cogenerators. Usually, these generators are small (typically ranging from 10 kW to 10 MW and connectable to either side of the meter) and can be economically connected only to the distribution system. If properly planned and operated, dispersed storage and generation may provide benefits to distribution

systems by reducing capacity requirements, improving reliability, and reducing losses. Dispersed-storage-and-generation technologies include hydroelectric, diesel generators, wind-electric systems, solarelectric systems, batteries, storage space and water heaters, storage air conditioners, hydroelectric pumped storage, photovoltaics, and fuel cells. *See* ELECTRIC POWER SYSTEMS; ENERGY SOURCES.

Turan Gönen

Bibliography. American National Standards Institute, *Preferred Voltage Ratings for A-C Systems and Equipment: Guide for (EEI R-6-1949) (IEC38)*, ANSI C84.1, 1989; M. T. Bishop, A. G. Jones, and W. F. Israel, Overcurent protection alternatives for underground distribution systems, *IEEE Trans. Power Delivery*, 10(1):69–77, January 1995; T. Gönen, *Electric Power Distribution System Engineering*, 1986; A. J. Pansini, *Guide to Electrical Power Distribution Systems*, 1992; W. L. Weeks, *Transmission and Distribution of Electrical Energy*, Harper and Row, New York, 1981.

Electric field

A condition in space in the vicinity of an electrically charged body such that the forces due to the charge are detectable. An electric field (or electrostatic field) exists in a region if an electric charge at rest in the region experiences a force of electrical origin. Since an electric charge experiences a force if it is in the vicinity of a charged body, there is an electric field surrounding any charged body.

Field strength. The electric field intensity (or field strength) **E** at a point in an electric field has a magnitude given by the quotient obtained when the force acting on a test charge q' placed at that point is divided by the magnitude of the test charge q'. Thus, it is force per unit charge. A test charge q' is one whose magnitude is small enough so it negligibly alters the field in which it is placed. The direction of **E** at the point is the direction of the force **F** on a positive test charge placed at the point. Thus, **E** is a vector point function, since it has a definite magnitude and direction at every point in the field, and its defining equation is (1).

$$\mathbf{E} = \frac{\mathbf{F}}{q'} \tag{1}$$

Ralph P. Winch

Principle of superposition. The field at a distance of r meters from a point charge of q coulombs in vacuum is given by Eq. (2), where **E** is the field

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0 r^2} \tag{2}$$

strength in volts per meter and $\epsilon_0 = 8.85 \times 10^{-12}$ farad/meter is the permittivity of free space. **E** is a vector, directed along the radius vector from the point charge; positive **E** is directed away from a positive charge or toward a negative charge. For an assembly of charges, the resultant field is, by the principle of superposition, the vector sum of the field components due to the individual charges. This summation may be performed directly, but in practical cases Gauss' theorem often affords a more powerful and convenient method. *See* GAUSS' THEOREM.

A. E. Bailey

Electric displacement. Electric flux density or electric displacement **D** in a dielectric (insulating) material is related to **E** by either of the equivalent equations (3), where **P** is the polarization of the

$$\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P} \qquad \mathbf{D} = \epsilon\mathbf{E} \tag{3}$$

medium, and ϵ is the permittivity of the dielectric which is related to ϵ_0 by the equation $\epsilon = \epsilon_r\epsilon_0$, ϵ_r being the relative permittivity of the dielectric. In empty space, $\mathbf{D} = \epsilon_0\mathbf{E}$. The units of **D** are coulombs per square meter.

In addition to electrostatic fields produced by separations of electric charges, an electric field is also produced by a changing magnetic field. The relationship between the **E** produced and the rate of change of magnetic flux density $d\mathbf{B}/dt$ which produces it is given by Faraday's law of induced electromotive forces (emfs) in Eq. (4), where $d\mathbf{s}$ is a vector element

$$\oint \mathbf{E} \cdot d\mathbf{s} = -\int_A \frac{d\mathbf{B}}{dt} \cdot d\mathbf{A} \tag{4}$$

of path length directed along the path of integration in the general sense of **E**. Thus $\oint \mathbf{E} \cdot d\mathbf{s}$ is the emf induced in this closed path of integration. The area of the surface bounded by the path of integration is A, and the direction of $d\mathbf{A}$, an infinitesimal vector element of this area, is the direction of the thumb of the right hand when the fingers encircle the path of integration in the general sense of **E**. The right side of Eq. (4) is seen to be the negative of the time rate of change of the magnetic flux linking the path of integration chosen for the left side.

In an electrostatic field, $\oint \mathbf{E} \cdot d\mathbf{s}$ is always zero. *See* ELECTRIC CHARGE; ELECTROMAGNETIC INDUCTION; POTENTIALS.

Ralph P. Winch

Bibliography. D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, 7th ed., 2005; E. M. Purcell, *Electricity and Magnetism*, 2d ed., 1985; H. D. Young and R. A. Freedman, *Sears and Zemansky's University Physics*, 11th ed., 2003.

Electric filter

A transmission network used to selectively modify the components of a signal according to their frequencies. In most cases a filter is used to enhance signals of desired frequencies while suppressing signals of undesired frequencies. An ideal filter would pass only desired frequencies while completely suppressing all unwanted frequencies, without any dispersion in time of the frequencies. Unfortunately, ideal filters are impossible to achieve.

Applications. Electric filters are used in most electronic communication systems. Whether communication is over wire, free space, or optical fiber, multiple channels of information can be multiplexed on different frequency bands. Unwanted signals

and noise are introduced along the communications path. The main function of electric filters is to separate the desired signal or channel from all others and from any noise or interference. For example, an AM radio receiver may have a low-pass filter after the antenna to separate the AM frequency band from higher frequency bands and, elsewhere in the radio, a band-pass filter to select the desired station out of the AM band. *See* ELECTRICAL COMMUNICATIONS; MULTIPLEXING AND MULTIPLE ACCESS; RADIO RECEIVER.

Although electronic filters are commonly thought of as devices for conferring selectivity to communication paths, they are used in almost every part of electronic equipment, such as the damping element in phase-locked loops, cleanup devices for frequency sources, and pulse expansion and compression devices for radar. One of the simplest and most common filters is the bypass capacitor used to restrict high-frequency electronic noise. *See* ELECTRICAL NOISE; PHASE-LOCKED LOOPS; RADAR.

Transfer function. The performance of a filter is, as is any transmission network, characterized by a transfer function. A transfer function is the ratio of the output signal to the input signal. In many filter types, including those composed of lumped elements, the transfer function can be expressed as a ratio of two polynomials, shown in Eq. (1). Here H is the transfer function, a_i and b_i are the coefficients of the two polynomials, and $s = j\omega$, ω being the frequency in radians and $j = \sqrt{-1}$. The zeros of the denominator of the transfer function are called the poles or natural frequencies of the filter. The zeros of the numerator are referred to as the transmission zeros or occasionally finite attenuation poles. Attenuation is the reciprocal of the transfer function or $1/H(s)$. Both the poles and the zeros of the filter can be complex and, for stability considerations, the poles must lie in the left half of the complex s-plane. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

$$H(s) = \frac{b_0 + b_1 s + b_2 s^2 + \cdots + b_n s^n}{1 + a_1 s + a_2 s^2 + \cdots + a_n s^n} \qquad (1)$$

The characteristics of a filter can be analyzed from its transfer function. The transfer function can be expressed in a polar form, shown in Eq. (2), where

$$H(j\omega) = |H(j\omega)| e^{j\theta(\omega)} \qquad (2)$$

$|H(j\omega)|$ is the magnitude and $\theta(\omega)$ is the phase function. The square of the magnitude is equal to the ratio of the output power to maximum available power. The power gain of the filter, expressed in decibels, is given in Eq. (3). Loss would be defined as the negative of gain. For example, the gain function of a band-pass filter approximates the ideal amplitude response, wherein the gain is constant within the passband and frequencies outside the passband are completely suppressed (**Fig. 1***a*). The phase function of a typical band-pass filter decreases as the frequency is increased through the passband (Fig. 1*b*). Another important parameter is group delay (Fig. 1*c*), defined in Eq. (4). Group delay is used to characterize how different frequencies are dispersed in time. Digital data are often transmitted by using shifts in frequency or phase. If the delay variation is too great,

$$\alpha(\omega) = 20 \log_{10} |H(j\omega)| \quad \text{dB} \qquad (3)$$

$$\tau(\omega) = \frac{d\theta(\omega)}{d(\omega)} \qquad (4)$$

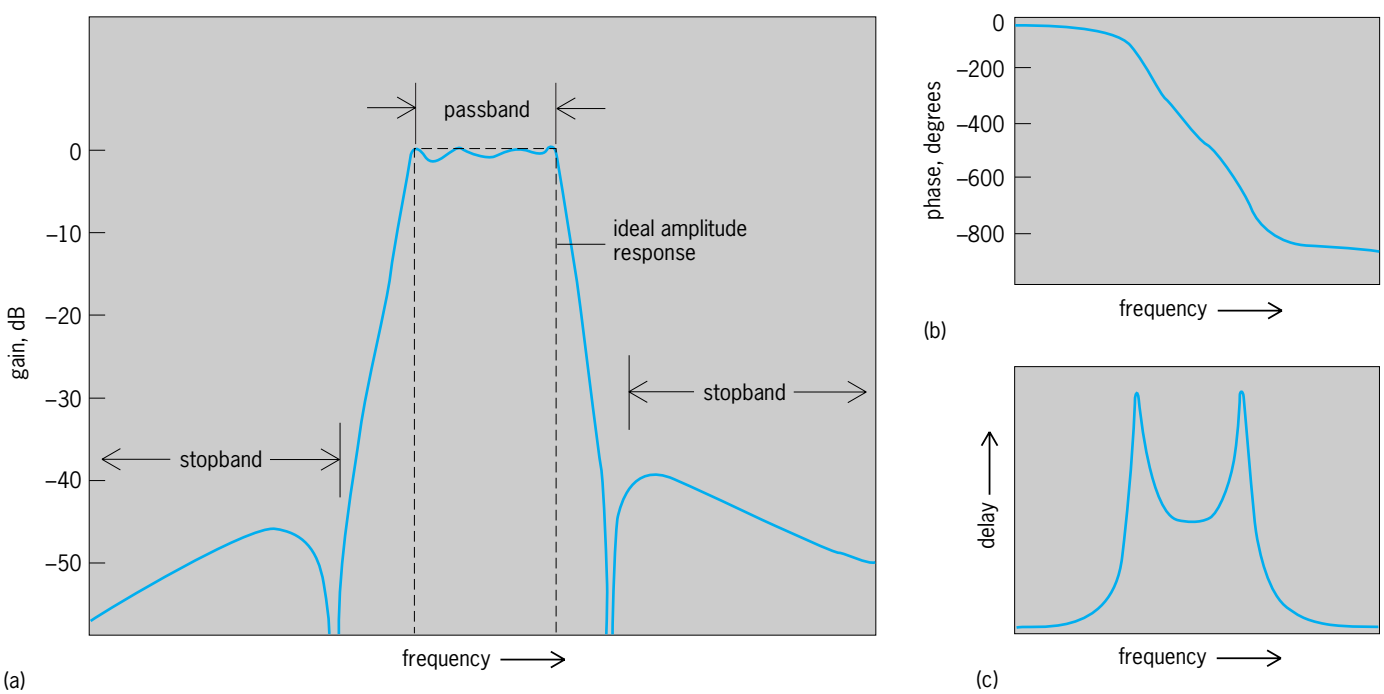

Fig. 1. Characteristics of a typical band-pass filter. (*a*) Gain characteristics, compared with those of an ideal filter. (*b*) Phase plot. (*c*) Delay plot.

these shifts will be detected incorrectly. *See* DECIBEL; GAIN; RESPONSE.

Characteristics. Filters are characterized in multiple ways. The expression low-pass, Butterworth, *LC* describes a filter. The descriptior low-pass indicates the relation of the passed to the rejected frequencies. Butterworth describes the type of polynomials in the transfer function. *LC* indicates the construction method. This filter is made of inductors (*L*'s) and capacitors (*C*'s).

Filter types. Filters are classified by the relationship of the frequencies that are selectively passed, referred to as the passband, to those which are attenuated, referred to as the stopband. An ideal low-pass filter passes all frequencies below a specified cutoff frequency and rejects those above. A high-pass filter does the opposite. An ideal band-pass filter will pass a band of frequencies while rejecting all others (Fig. 1*a*); a band-reject filter will reject a band of frequencies and pass all others.

An all-pass filter passes all frequencies but does, however, modify the time delay characteristics. It normally corrects delay distortions caused by other sections of a communication path.

All the above classifications are based on frequency-domain considerations. In addition, there are two terms that apply to the time-response characteristics of a filter. A finite impulse response (FIR) filter, when exposed to a change in input, will settle to a steady state within a finite amount of time. An infinite impulse response (IIR) filter will continue oscillating in a decaying manner forever.

A further consideration in classifying a filter is whether the frequency response is constant in time or varies. If it varies with time, as a function of the input signal, it is known as an adaptive filter. This type of filter finds use in speech and image enhancement and echo cancellation.

Approximations. Realizable filters are only approximations to the ideal. An ideal filter has instantaneous cutoff from the passband to the stopband. It passes all the desired frequencies with equal amplitude and delay at each frequency. Actual filters have a band of frequencies where the filter transitions from the passband to the stopband. The passband amplitude may not be flat but can have ripple (Fig. 1*a*). In the stopband of a filter that includes attenuation poles, the amplitude response will flare back up between those attenuation poles. The filter can also be dispersive, where different frequencies have different time delays.

The transfer function is the ratio of two polynomials. At least one of these polynomials would require an infinite number of terms to meet the instantaneous cutoff requirements of the ideal filter. However, there are several classes of polynomial that do approximate the ideal filter. These functions are used to synthesize actual filters.

The Chebyshev approximation to the ideal filter is a polynomial function optimized so that for a given stopband rejection the maximum variation in the passband is a minimum. The numerator of the transfer function of a Chebyshev filter would be 1, and it would have equal ripple in the passband; that is, all the peaks in the amplitude would be of the same height and all the dips of the same depth. An elliptic-integral filter, also known as a Cauer filter, has both equal ripple in the passband and equal minima of attenuation in the stopband; that is, the heights of the flarebacks between the attenuation poles are all the same. Both the numerator and denominator polynomials of the transfer function are of order greater than zero. The passband has the same criteria as the Chebyshev and, in addition, the minimum attenuation in the stopband is maximized. For polynomials of a given order, Chebyshev and Cauer filters provide the least ripple for the maximum stopband selectivity. Higher ripple can be traded for an increase in selectivity. However, an increase in delay variation and a degradation in stability will also occur. Increased stability and lower delay variation is obtained by allowing the ripple to decrease. By decreasing it in a Chebyshev design to zero, where the derivatives of *H(s)* at center frequency become zero, a Butterworth filter is obtained. The Butterworth polynomial is the most stable. Changes in its coefficients or the element values used to realize Butterworth response cause the least changes in ripple of all the polynomial types.

In many applications the time-domain characteristics are as important as those of the frequency domain. In radar the key information is provided by pulse peaks. In many digital systems the information is encoded in the phase of a signal. If all the frequencies passing through a filter have the same time delay, the signal shape will remain unchanged and the encoded information will be recoverable. A filter with a gaussian-function amplitude response will have a linear phase and a constant delay characteristic. The normalized transfer function for a low-pass filter is givenby Eq. (5). This function is not a polynomial

$$|H(j\omega)| = e^{-\omega^2/2} \tag{5}$$

and not physically realizable. Approximations that try to match the amplitude or delay response of this function are known collectively as gaussian filters. The Bessel filters, also known as Thomson filters, are members of this group. They have maximally flat delay in the sense that the derivatives of delay at center frequency vanish. Other members of this group try to match the gaussian function only in the passband portion of their response. This characteristic allows increased selectivity in the stopband. Others try to provide constant delay in the Chebyshev sense by having equal delay ripple. *See* DISTRIBUTION (PROBABILITY).

Filter realization. Because filters are used over wide frequency and bandwidth ranges and with such varying performance criteria, many methods have been devised for creating a filter function (see **table**).

A term used to measure the quality of components in passive filters is *Q*. The quantity *Q* is defined as the ratio of energy lost in one cycle to total energy stored. It is a measure of how much loss an element has. An inductor with an inductance of *L* henrys and

Application ranges of various filter types

Filter type	Frequency range	Bandwidth (% of center frequency)
LC	0–5 GHz	>0.5
Crystal	5 kHz–300 GHz	0.01–5
Ceramic	300 kHz–50 MHz	0.5–20
Mechanical	1 kHz–700 kHz	0.02–10
Surface-acoustic-wave (SAW) resonator	30 kHz–2 GHz	0.01–0.1
SAW transversal	30 kHz–2 GHz	0.2–50
Active	<100 kHz	>0.5
Switched capacitor	<40 kHz	>0.5
Digital	<10 MHz	>0.1
Microwave	0.5 GHz–200 GHz	>0.5
Ceramic dielectric resonator (CDR)	0.7 GHz–5 GHz	>0.1

series parasitic resistance of R ohms will have a Q equal to $\omega L/R$, where ω is the angular frequency. Filter elements with low Q values increase the loss of a filter and limit the minimum bandwidth that may be achieved. *See* Q (ELECTRICITY).

LC and RC filters. These filters are based on lumped element inductors (L's), capacitors (C's), and resistors (R's). The transfer function for a typical RC filter (**Fig. 2***a*), which contains one pole at $-1/RC$ on the real axis and no zeros, is given in Eq. (6). Because the

$$H(s) = \frac{1}{1 + sCR} \tag{6}$$

poles of RC filters can be located only on the negative real axis, only low-pass responses can be achieved. To achieve complex poles and zeros, inductors must be used (Fig. 2*b*). Inductors have relatively low Q's, and at low frequencies are bulky. The Q's of typical inductors range from ten to a few thousand, and the larger the Q, the larger the inductor. Many of the other filter types were created to avoid the problems with inductors.

Most LC filters are designed as ladder structures with resistive termination. A ladder structure consists of alternating series and shunt arms (**Fig. 3***a*). The poles and zeros of the transfer function are provided by resonant circuits within the arms. The ability to control these resonances gives the ladder structure stability in both the passband and stopband. An alternate to a ladder would be a multipath structure. To achieve high rejection in the stopband requires that the amplitude and phase be closely controlled at the points where the paths join in order that the signals from the separate circuits will cancel. Therefore components must have tighter tolerances than those needed to realize a ladder design. Therefore, ladder structures tend to predominate in LC filter design. In crystal-filter design, however, because of the closeness of the impedance pole and zero of crystals, multipath (mainly lattice) structures are common. *See* RESONANCE (ALTERNATING-CURRENT CIRCUITS).

In a ladder structure a transfer-function pole occurs if a series arm has an impedance zero or a shunt arm has an impedance pole. Transfer-function zeros occur where a series impedance arm has a pole, or a shunt arm a zero. A capacitor connected in series or parallel with an inductor would be the simplest circuit for creating either an impedance zero or a pole. An impedance structure that contains both a finite-frequency pole and a zero consists of a capacitor (C_1), in series with an inductor capacitor, (L_1) and a second capacitor (C_2), connected in parallel (Fig. 3*b*). This circuit also contains a pole at zero frequency and a zero at infinite frequency.

An LC structure is created by choosing a transfer function, of the form shown in Eq. (1), to meet the filter's specification. This function may be a Chebyshev or other standard type or one created by the designer that is optimum for the specification. By finding the zeros of the numerator and denominator polynomials of the transfer function, the numbers of transfer-function poles and zeros at zero, finite, and infinite frequency are determined. At this point an LC circuit structure can be synthesized. By factoring out poles or zeros sequentially from the transfer function, the type of arm and component values required can be determined. As the process continues, the ladder gains arms as the transfer function describing the residual portion of the circuit is reduced in complexity.

Acoustic filters. This class of filters includes crystal, ceramic, mechanical, and surface-acoustic-wave (SAW) filters. These devices convert electrical energy

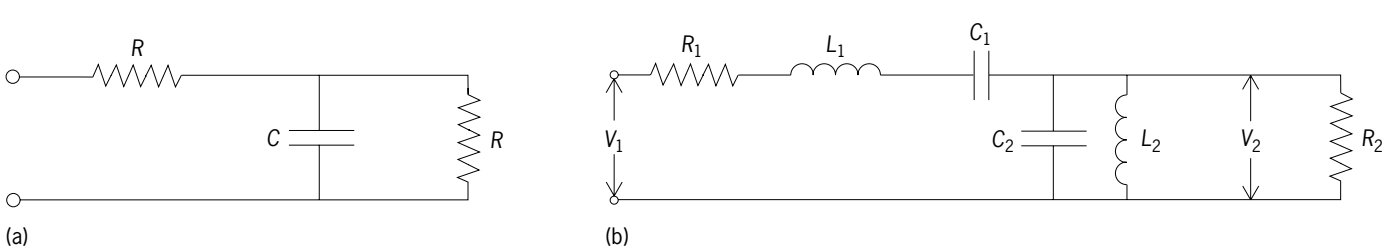

Fig. 2. Passive, lumped-element filters. (*a*) RC low-pass filter. (*b*) *LC* band-pass filter.

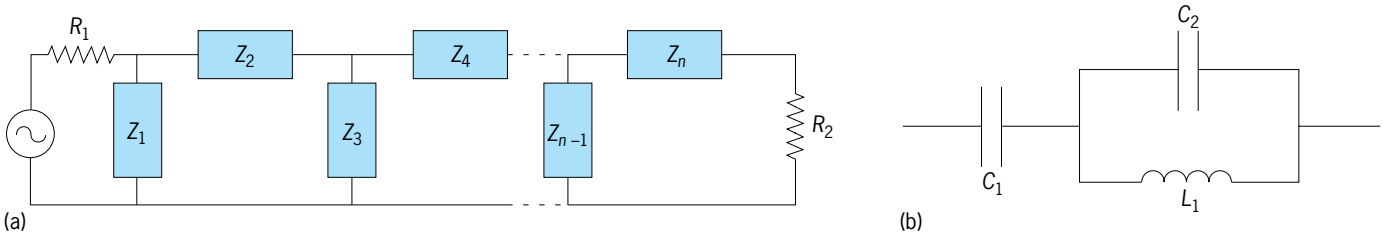

Fig. 3. Ladder structure. (*a*) Overall structure with alternating series and shunt arms. (*b*) LC circuit whose impedance contains both a finite-frequency pole and a zero. Such circuits are used in the series and shunt arms of the ladder structure.

to mechanical vibrations, process the signal acoustically, and then convert the energy back to an electrical form. The equations describing a mechanically vibrating resonator, where energy is cycled between kinetic motion and stress, match those of an inductor and capacitor (*LC*) attached in parallel, where energy is cycled between the electric field of the capacitor and the magnetic field of the inductor. However, the mechanical resonators have much higher *Q*'s and better stability than the LC circuit. With the addition of a transducer to convert electrical energy to acoustic, the *LC* circuits can be replaced with mechanical resonators. *See* RESONANCE (ACOUSTICS AND MECHANICS); TRANSDUCER.

SAW filters can be built in both transversal structures and resonators structures. Crystal, ceramic, and mechanical filters are realized only as resonator filters. Acoustic resonator filters use mechanical resonators coupled together either electrically or acoustically. Each resonator provides either a pole or a zero in the transfer function. In addition, the piezoelectric coupling of mechanical to electrical energy provides a second resonance that may be used to create additional poles and zeros. *See* HELMHOLTZ RESONATOR; PIEZOELECTRICITY.

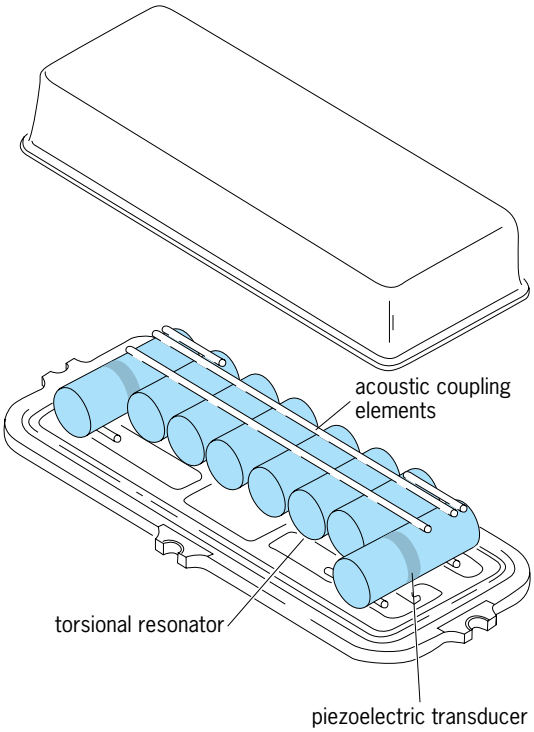

Fig. 4. Mechanical filter.

The difference between crystal and ceramic filters is mostly in the material that they use for their resonators. In crystal filters the resonators are made from quartz. It is a piezoelectric material with excellent temperature stability and high *Q*. Ceramic filters use other piezoelectric materials, mainly a lead-zirconate-titanate (PZT) ceramic, for their resonators. These materials have much higher coupling coefficients than quartz and therefore can achieve much wider bandwidths, but have lower *Q*'s and poorer temperature stability.

Mechanical filters (**Fig. 4**) use metal for their resonant elements. Coupling between elements is done mechanically. Piezoelectric or other transducer types are used at the input and output to convert electrical energy to acoustic energy. Mechanical filters have temperature stabilities and *Q*'s between those of quartz and ceramic filters.

SAW resonators trap surface acoustic waves between two reflecting structures. Interdigital fingers are used to couple electrical energy piezoelectrically into and out of the acoustic wave. Transversal SAW filters use a transmitting and receiving set of piezoelectric interdigital fingers. One or both of the finger sets are weighted to create a desired impulse response. The spacings between interdigital fingers act as time delays for the acoustic wave. These time delays and weighting of the fingers are analogous to processes used in digital filters. Because of this analogy, many of the algorithms that are used for transversal digital filters have been applied to SAW filters. *See* SURFACE-ACOUSTIC-WAVE DEVICES.

Other filter realizations. Many techniques are used to create filters. Inductors are replaced with transistor networks in active filters, discussed below, to reduce size and cost. By using an analog-to-digital converter the transfer function can be created mathematically by a digital processor. At high frequencies, transmission lines and waveguide structures replace lumped elements. *See* DIGITAL FILTER; INTEGRATED-CIRCUIT FILTER; MICROWAVE FILTER; SWITCHED CAPACITOR CIRCUIT. Donald P. Havens

Active filters. An active filter comprises resistors, capacitors, and active elements such as operational amplifiers. It is also referred to as an active-*RC* filter.

Advantages and disadvantages. Active filters can realize the same filter characteristics as passive ones comprising resistor, capacitor, and inductor elements.

They have, however, several advantages over their passive counterparts:

1. Active filters can provide gain, and are frequently used to simultaneously match filtering (frequency-determining) and gain specifications.

2. They are readily implemented in integrated-circuit technology, whereas the inductor element of passive filters is not readily realized. As a result, the active filter is inexpensive, and is attractive for its small size and weight. In addition, it is readily included with other signal-processing functions on a single integrated circuit. *See* INTEGRATED CIRCUITS.

3. The design of active filters is considerably simpler than that of passive ones. In addition, it is easy to provide for variability, which can be used to change filter characteristics by electrical input signals.

The active filter also has some disadvantages:

1. Since the active filter contains electronic components, it requires a power supply, which adds to the complexity of the realization. The electronic components also place restrictions on the level of the signals that can be applied to the filter and on the noise component that the filter may add to the filtered signal. *See* ELECTRONIC POWER SUPPLY.

2. The mathematical process by which the active filter produces filtering characteristics in general requires the use of internal feedback. When this feedback is positive, the resulting filter may be very sensitive to lack of precision in component values, and the effects of aging and environmental conditions. *See* FEEDBACK CIRCUIT.

Single-amplifier filters. One important classification of active filters is according to the number of gain elements (amplifiers) that are used in their design. Historically, the first practical filter used a single amplifier of positive low gain, in the range of 1–5. The amplifier was realized by an operational amplifier and two feedback resistors, R_A and R_B. The resulting configuration could be used for low-pass, high-pass, and band-pass applications. Other filtering functions could be obtained by additional circuit complexity. The circuit is usually referred to as a Sallen-Key filter (**Fig. 5***a*). Since the Sallen-Key filter uses positive feedback, the sensitivity problem mentioned above usually restricts the use of this filter to low-Q applications.

A more frequently utilized single-amplifier filter uses an inverting operational amplifier directly as the active element. The filter is called an infinite-gain, single-amplifier filter (Fig. 5*b*). Similar filter configurations may be used to realize high-pass and band-pass filter characteristics. Although the design equations for the infinite-gain, single-amplifier filter are more complex than those for the Sallen-Key one, the use of negative feedback which is implicit in this realization provides for a far more practical realization with considerably lower sensitivities.

The single-amplifier, second-order filter realizations described above can also be used to realize higher-order filter functions. [The order of a filter is the number n in Eq. (1), giving its transfer function.] To do this, separate filter sections are used to realize second-order factors, $T_i(s)$ of the overall transfer function, and the resulting sections are connected in a cascade configuration in which the output of one filter section is used as the input of the next one (Fig. 5*c*). There are $n/2$ sections labeled T_i, and since each is of second order, an nth-degree function results.

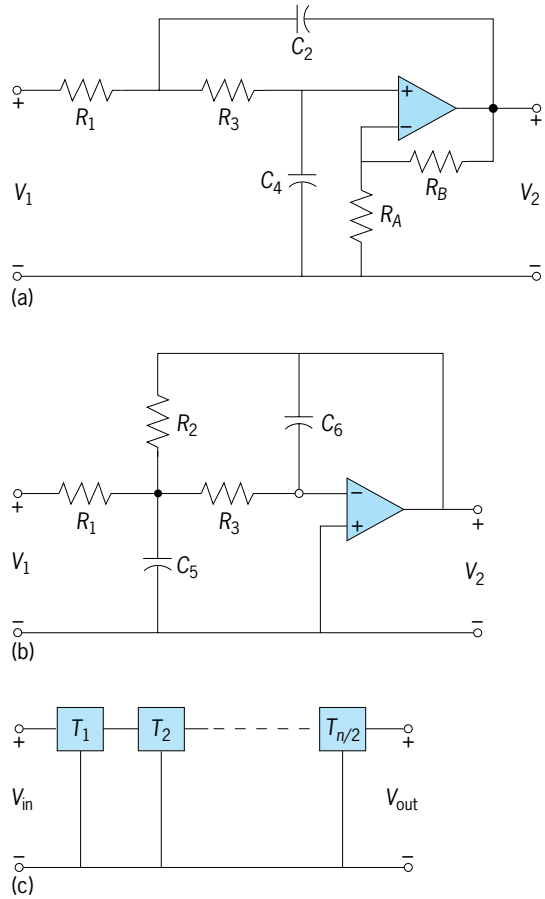

Fig. 5. Single-amplifier filters. (*a*) Low-pass Sallen-Key filter. (*b*) Low-pass, infinite-gain, single-amplifier filter. (*c*) Cascade of second-order filter realizations.

Multiple-amplifier filters. Another type of second-order filter, which can be used either by itself or as a building block in a cascade, is the multiple-amplifier filter. Such a filter uses more than one active element (operational amplifier), typically three. Although these filters have a more complex structure than the single-amplifier ones, and usually have a greater number of passive RC elements as well, they have excellent sensitivity properties and are easy to design. They are the most important general-purpose filter structures.

One of the best-known multiple-amplifier active filters is the state-variable filter, also known as the KHN (from the names of its inventors, W. J. Kerwin, L. P. Huelsman, and R. W. Newcomb) [**Fig. 6***a*]. Three outputs, V_{HP}, V_{BP}, and V_{LP}, for high-pass, band-pass, and low-pass transfer functions respectively, are available. The versatility of the state-variable filter is improved if a fourth operational amplifier is provided as a summing amplifier. By summing the high-pass, low-pass, and band-pass outputs, a biquadratic filter function can be realized. This includes notch filters,

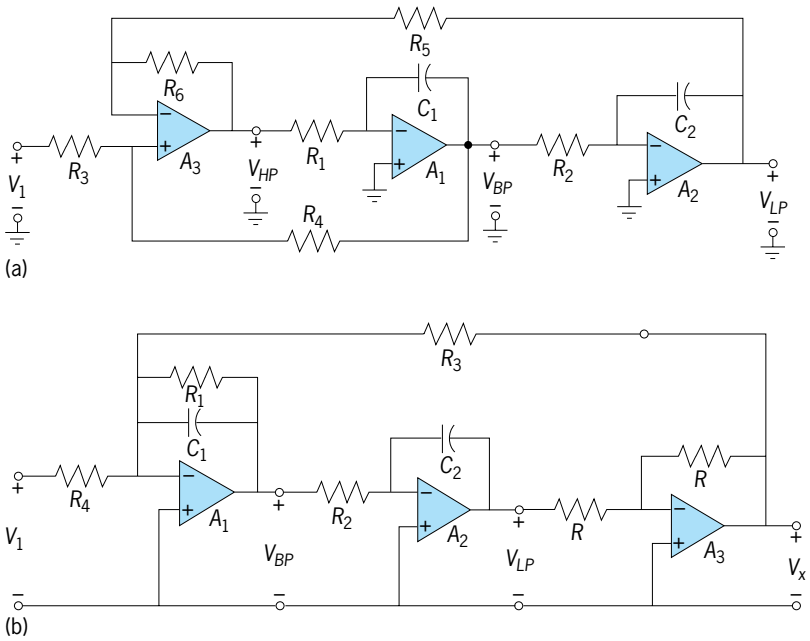

Fig. 6. Multiple-amplifier filters. (*a*) State-variable (KHN) filter. (*b*) Tow-Thomas filter.

elliptic (Cauer) filters, and all-pass filters. Realizations of higher than second order can be produced by cascading appropriately designed second-order realizations.

Another versatile multiple-amplifier filter, the Tow-Thomas filter (Fig. 6*b*), has the advantage of being even easier to design than the state-variable filter, but has the disadvantage of not providing a high-pass output in its basic configuration.

Direct realizations. A completely different class of active filters than the single- and multiple-amplifier ones described above is referred to as direct realizations. This class uses a passive *RLC* filter realization as a prototype, then directly derives an active-*RC* configuration from the prototype. Since passive filters can be of any order, the resulting class of active filters can also have any desired order. These high-order active filters are said to be directly realized in contrast to the cascade higher-order realizations described above. There are two main classes of direct realization filters.

The first type of direct realization is one in which the inductor element is replaced by a simulated inductor. The other elements in the circuit remain the same. This technique is especially applicable to the realization of high-pass filters in which the inductor elements all appear with one terminal connected to ground. This arrangement simplifies the connection of the power supplies for the active elements. An example of a simulated inductor realization is a fifth-order high-pass filter (**Fig. 7***a*). In a direct active-*RC* realization of the filter (Fig. 7*b*), each of the inductors is simulated by an active-*RC* circuit consisting of four resistors, a capacitor, and two operational amplifiers.

Another example of an element-replacement direct realization is one in which the elements are first subjected to an impedance transformation of $1/s$, in which $R \rightarrow C$, $L \rightarrow R$, and C becomes an element with an impedance $1/Cs^2$. This element is called a frequency-dependent negative resistor (FDNR), and is indicated in circuit diagrams by a four-bar symbol. Such a transformation leaves the voltage transfer function invariant but permits the realization of a low-pass filter with grounded active-*RC* elements. An example is a fifth-order *RLC* low-pass filter (**Fig. 8**), in which two active-*RC* circuits, each consisting of two capacitors, three resistors, and two operational amplifiers, realize the FDNRs.

The second type of direct realization is one in which the voltages and currents of a passive *RLC* filter are simulated by the voltages in a circuit consisting of damped and undamped inverting integrators. The resulting configuration has a series of feedback paths and is called a leapfrog filter.

Computer-aided design. The design of various active filters (as well as their passive counterparts) can be simplified by the use of various computer-aided design programs which operate on a personal computer. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Parasitics. In general, active filters are designed with the assumption that the active elements are ideal (no parasitics). For example, the operational amplifier is assumed to have infinite gain, infinite input

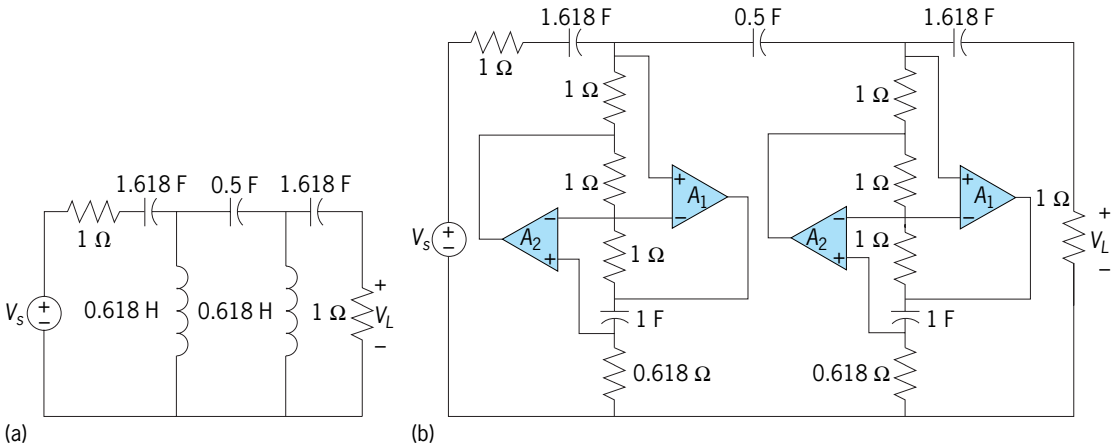

Fig. 7. Simulated-inductor realization of fifth-order high-pass filter. (*a*) *RLC* prototype, normalized for cutoff frequency of 1 rad/s (0.159 Hz). (*b*) Direct active-*RC* realization.

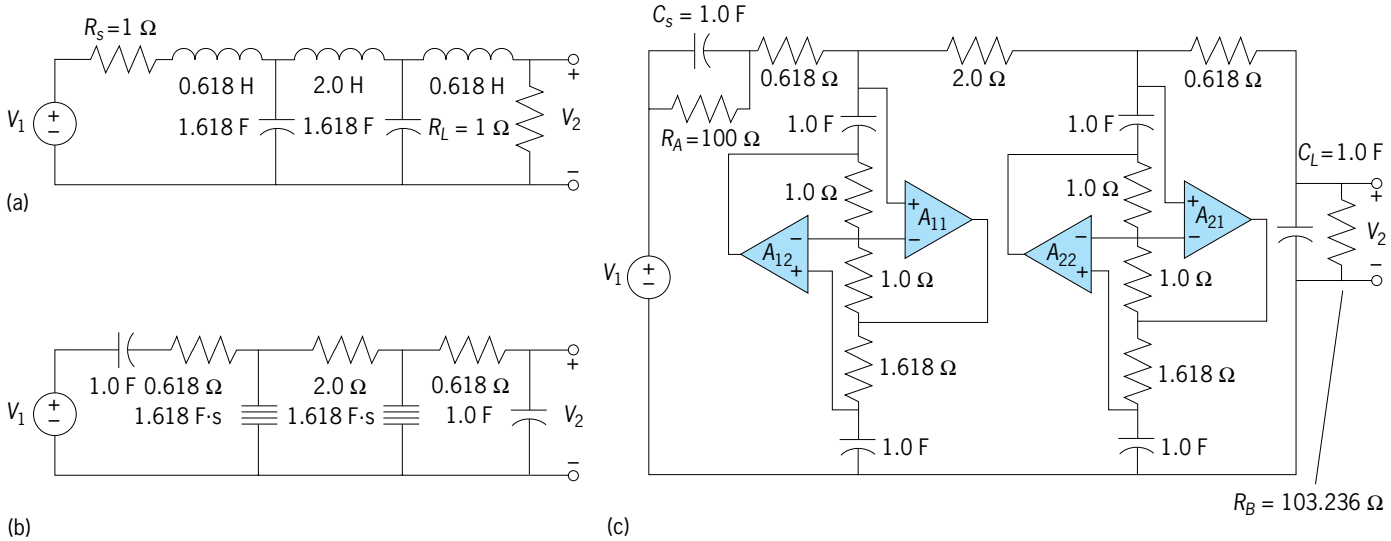

Fig. 8. FDNR realization of fifth-order low-pass filter. (*a*) *RLC* prototype, with normalized low-pass function with cutoff frequency of 1 rad/s (0.159 Hz). (*b*) Filter after impedance transformation. (*c*) Final realization, in which active-*RC* circuits realize FDNRs.

impedance, zero output impedance, and an infinite frequency range (gain-bandwidth). All practical filter realizations must be evaluated for the effect that the nonideal parasitics have on actual filter performance. One of the most troublesome of the parasitics is the operational-amplifier gain bandwidth. Typically, for a given application, tuning charts may be developed which can be used to provide compensation for such operational-amplifier limitations. *See* AMPLIFIER; CIRCUIT (ELECTRONICS); OPERATIONAL AMPLIFIER.

Lawrence P. Huelsman

Bibliography. H. Blinchikoff and A. Zverev, *Filtering in the Time and Frequency Domains*, 1976, reprint 1986; M. Hasler and J. Neirynck, *Electric Filters*, 1986; L. P. Huelsman, *Active and Passive Analog Filter Design*, 1993; S. Mitra and C. Kurth, *Miniature and Integrated Filters*, 1989; R. Schaumann and M. Van Valkenburg, *Design of Analog Filters: Passive, Active-RC, and Switched Capacitor*, 2d ed., 2000.

Electric furnace

A chamber heated to high temperature by electricity. The furnace consists of a refractory shell to resist the high temperatures attained within the chamber and minimize the heat losses to the surrounding area, plus a power source and electrical circuit to provide the heat. The operating temperatures range from a few hundred degrees Celsius to the melting points of refractory metals. The heating mechanism may be resistance, induction, microwave, or an electric arc in the largest installations. Electric furnaces are used primarily in the metallurgical industries for heat-treating, melting, and smelting operations. Resistance furnaces are used for heat treating. Induction furnaces are used to melt relatively low-melting alloys for casting, while arc furnaces are used for melting and smelting applications. The range of furnace powers, sizes, and geometries varies immensely. Powers range from a few kilowatts to 100 megavolt-amperes in a single unit. The furnace volume may be less than a cubic meter to more than 300 m^3. Both cylindrical and rectangular geometries are common. *See* ELECTRIC HEATING; ELECTROMETALLURGY; FURNACE CONSTRUCTION; HEAT TREATMENT (METALLURGY); INDUCTION HEATING; METALLURGY; REFRACTORY; RESISTANCE HEATING.

Electric-arc furnaces, in which the prime mechanism of heat transfer is from an electric arc, may be of the direct-arc or indirect-arc type. In the direct-arc furnace, an arc is struck between the charge being melted (usually scrap steel) and an electrode. In the indirect-arc type, an arc is struck between adjacent electrodes, and this arc then heats the charge. The arc may also be submerged below the surface of the charge. In this case, some of the current may heat the charge by resistance heating through the charge, while the rest of the current passes through the arc. Submerged-arc furnaces are among the largest units in operation and are used primarily for the smelting of ferroalloys and silicon metal, as well as phosphorus, calcium carbide, calcium silicon, copper, and copper-nickel mattes. *See* ARC HEATING.

Both alternating current and direct current are used to operate electric-arc furnaces. Historically, alternating-current devices with three or six carbon electrodes were used. A cylindrical geometry with the electrodes arranged in a triangle is common for three electrode systems, while the largest smelting furnaces are usually rectangular with the six electrodes installed in a line. Direct-current furnaces are cylindrical with a single axial electrode acting as cathode, while the furnace charge and bottom becomes the anode. Submerged-arc furnaces are continuously charged with feed, and there is a continuous discharge of gas from the furnace roof. Molten metal and/or slag are removed by tapping at discrete time intervals. Melting furnaces have a removable

roof for charging and often a means of tilting the furnace to remove the molten product. Although most arc furnaces use graphite electrodes, some smaller-capacity plasma furnaces use nonconsumable water-cooled metal electrodes, which allow for better control of the furnace atmosphere.

Presently one-third of the world's steel production and 50% of the United States' steel production comes from electric-arc furnaces. The largest units have graphite electrodes 60–80 cm (24–31 in.) in diameter and a tap weight of 80–100 metric tons (88–110 tons) of liquid steel. The active power may be up to 100 megawatts. A typical FeNi submerged-arc smelting furnace would have six electrodes and an operating load of 75 MVA. *See* STEEL MANUFACTURE.

Richard J. Munz

Bibliography. N. Barcza, The application of dc plasma-arc technology to extractive metallurgical processing, *High Temp. Mater. Proc.*, 1:1–15, 1997; F. Fereday, *Commentary: Submerged Arc Furnaces*, TechCommentary published by CMP, EPRI Center for Materials Production, Carnegie Mellon Research Institute, 15230-2950, 1996; J. Kunze and R. Degel, History, current status and new trends of submerged arc furnace technology for ferroalloy metals, *Steel GRIPS*, 1(3):164–169, 2003; D. Neuschütz, Metallurgical applications of high power are heating systems, *High Temp. Mater. Proc.*, 4(3):309–321, 2000.

Electric heating

Methods of converting electric energy to heat energy by resisting the free flow of electric current. Electric heating has several advantages: it can be precisely controlled to allow a uniformity of temperature within very narrow limits; it is cleaner than other methods of heating because it does not involve any combustion; it is considered safe because it is protected from overloading by automatic breakers; it is quick to use and to adjust; and it is relatively quiet. For these reasons, electric heat is widely chosen for industrial, commercial, and residential use.

Types of electric heaters. There are four major types of electric heaters: resistance, dielectric, induction, and electric-arc.

Resistance heaters. Resistance heaters produce heat by passing an electric current through a resistance—a coil, wire, or other obstacle which impedes current and causes it to give off heat. Heaters of this kind have an inherent efficiency of 100% in converting electric energy into heat. Devices such as electric ranges, ovens, hot-water heaters, sterilizers, stills, baths, furnaces, and space heaters are part of the long list of resistance heating equipment. *See* RESISTANCE HEATING.

Dielectric heaters. Dielectric heaters use currents of high frequency which generate heat by dielectric hysteresis (loss) within the body of a nominally nonconducting material. These heaters are used to warm to a moderate temperature certain materials that have low thermal conducting properties; for example, to soften plastics, to dry textiles, and to work with other materials like rubber and wood. *See* DIELECTRIC HEATING.

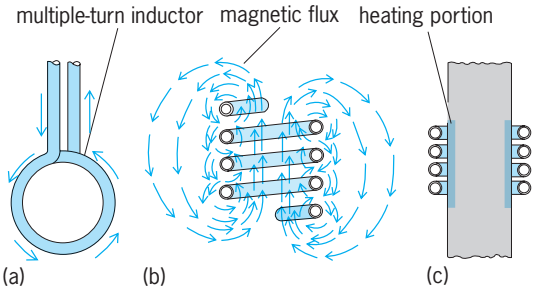

Fig. 1. Induction heater. Arrows indicate current flow. (*a*) End-on view. (*b*) Magnetic fields. (*c*) Heating pattern.

Induction heaters. Induction heaters produce heat by means of a periodically varying electromagnetic field within the body of a nominally conducting material. This method of heating is sometimes called eddy-current heating and is used to achieve temperatures below the melting point of metal. For instance, induction heating is used to temper steel, to heat metals for forging, to heat the metal elements inside glass bulbs, and to make glass-to-metal joints. **Figure 1** shows the magnetic fields and heating pattern produced by an induction heater. *See* INDUCTION HEATING.

Electric-arc heaters. Electric-arc heating is really a form of resistance heating in which a bridge of vapor and gas carries an electric current between electrodes. The arc has the property of resistance. Electric-arc heating is used mainly to melt hard metals, alloys, and some ceramic metals (**Fig. 2**). *See* ARC HEATING.

General design features. All electrical parts must be well protected from contact by operators, work materials, and moisture. Terminals must be enclosed within suitable boxes, away from the high heat zone, to protect the power supply cables. Repairs and replacements should be possible without tearing off heat insulations.

Resistance heaters are often enclosed in pipes or tubes suitable for immersion or for exposure to difficult external conditions. Indirect heating is done by circulating a heat transfer medium, such as special oil or Dowtherm (liquid or vapor), through jacketed vessels. This permits closer control of heating-surface temperature than is possible with direct heating.

Some conducting materials can be heated by

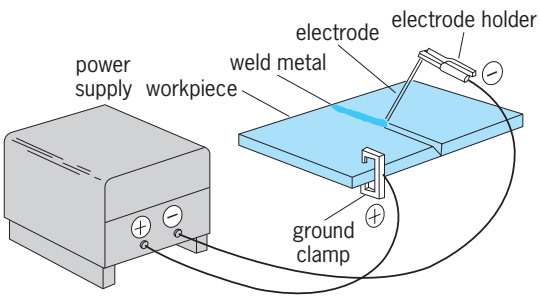

Fig. 2. Typical electric-arc heater.

passing electric current through them, as is done in the reduction of aluminum. Some conducting liquids can be heated by passing an electric current between immersed electrodes. Heat is produced by the electrical resistance of the liquid.

The supply of necessary electric power for large heating installations necessitates consultation with the utility company. The demand, the power factor of the load, and the load factor all affect the power rates. Large direct-current or single-phase alternating-current loads should be avoided. Polyphase power at 440–550 V permits lower current and reduced costs. *See* FURNACE CONSTRUCTION.

Electric heating for houses. Electricity is one choice for heating houses, but with only a 35% efficiency rate, electricity has been a less attractive option than the direct use of gas and oil for heating homes.

One type of electric heater for houses is the heat pump. The heat pump can be used alone or to supplement the output of direct resistance heating (heat strips). The heat pump works on a reversed refrigeration cycle; the pump pulls the heat from the cold outside air and forces it into the house. The efficiency of this process is directly affected by the ambient temperature and the desired interior temperature; the larger the difference in these temperatures, the lower the efficiency of the pump. Generally, however, the efficiency of the heat pump is higher than the efficiency of resistance heating, so that the efficiency of the heat pump is actually higher than 100%. *See* HEAT PUMP.

Since the heat pump performs both heating and cooling and is more efficient than other forms of electric heat, it can be an economical way to heat houses.

Common electric heating systems in houses are central heating employing an electric furnace with forced air circulation; central heating employing an electric furnace with forced water circulation; central heating using radiant cables; electrical duct heaters; space (strip) heaters which use radiation and natural convection for heat transfer; and portable space heaters. Mo-Shing Chen

Bibliography. C. J. Erikson, *Handbook of Electric Heating for Industry*, 1994; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; L. Miles, *Heat Pumps: Theory and Service*, 1993; M. Orfeuil, *Electric Process Heating*, 1987; H. J. Sauer, *Heat Pump Systems*, 1990.

Electric insulator

A material that blocks the flow of electric current across it. Insulators are to be distinguished from electrolytes, which are electronic insulators but ionic conductors. *See* ELECTROLYTE; IONIC CRYSTALS.

Electric insulators are used to insulate electric conductors from one another and to confine electric currents to specified pathways, as in the insulation of wires, electric switchgear, and electronic components. They provide an electrical, mechanical, and heat-dissipation function. The electrical function of an insulator is characterized by its resistivity, its dielectric strength (breakdown voltage), and its dielectric constant (relative permittivity). Insulators can be solid, liquid, or gaseous (see **table**).

Resistivity. The resistivity of a material is a measure of the electric current density that flows across it in response to an applied electric field. The electric current density is the sum of contributions from all species of mobile charge carriers, including electrons and ions. Each contribution is proportional to the concentration of the mobile species, its charge, and its mean (drift) velocity in response to the applied field. Solid and liquid insulators have direct-current resistivities of 10^{10} ohm-meters at room temperature as compared to 10^{-8} Ω-m for a good metal or 10^{-3} Ω-m for a fast ionic conductor. Insulators minimize the concentrations of mobile charge carriers by having a filled valence band separated from an empty conduction band by a large energy gap and by having all equivalent lattice sites either occupied or vacant of ionic species. An energy gap greater than 5 eV can be obtained in either ionically or covalently bonded crystals and molecules. Oxides may have the advantage of being inert to further oxidation, but they are more subject to short-range, if not long-range, ionic motion. Protons introduced by the absorption or adsorption of water are the most prevalent source of ionic conduction. Impurities and crystalline defects are sources of mobile electrons accessible at lower temperatures and electric fields. *See* BAND THEORY OF SOLIDS; ELECTRICAL RESISTANCE; ELECTRICAL RESISTIVITY; SEMICONDUCTOR.

Dielectric strength. A material used in the electrical industry for insulation, capacitors, or encapsulation may be exposed to large voltage gradients that it must withstand for the operating life of the system. Power lines operate at transmission voltages in excess of 1 MV, for example. Failure occurs if an electric short-circuit develops across the material. Such a failure is called dielectric breakdown. The breakdown voltage gradient, expressed in kilovolts per millimeter, is a measure of the dielectric strength. Dielectric breakdown of a solid is destructive; liquids and gases are self-healing.

Breakdown is initiated by the injection of mobile charge carriers into the conduction or valence band of the insulator. There are three basic types: intrinsic, thermal, and avalanche. Intrinsic breakdown occurs where electrons are injected across a metal-insulator interface. It depends primarily on the applied field gradient, it occurs in less than 1 microsecond, and it is independent of sample shape. Thermal breakdown is associated with thermal excitation of electrons into a conduction band; resistive losses of the excited electrons lead to thermal runaway. This mechanism depends primarily on the temperature rather than the applied field gradient, and is dependent on both sample shape and the rate of application of the field. Avalanche breakdown occurs where collisions between mobile and stationary electrons multiply

Characteristics of electric insulators*

Insulator	Breakdown voltage, kV/mm	Resistivity, Ω-m,	Relative permittivity (k)
Solid			
Paper	20	10^{10}	3–4
Nylon	16	10^{13}	3–4
Glass	20–40	10^{10}–10^{14}	5
Rubber	16–50	10^{13}–10^{14}	3
Mica	25–200	10^{11}–10^{15}	5
Poly(tetrafluoroethylene)	60	10^{14}–10^{15}	2.1
Porcelain	10–20	10^{16}	5–7
Polystyrene	25	10^{16}	2.5
Liquid			
Saturated hydrocarbons†	140–200	10^{17}–10^{18}	
Aromatic hydrocarbons†	160–280	10^{17}–10^{18}	
Transformer and silicon oils‡	15		
Gaseous§			
Air	3		
Carbon dioxide	2.7		
Sulfur hexafluoride	7.5		

*At 20°C (68°F) and atmospheric pressure (101.3 kPa)
†Physically and chemically purified.
‡Commercial high purity.
§Between two electrodes 25 cm (10 in.) in diameter.

the concentration of mobile electrons. It generally occurs in conjunction with intrinsic or thermal breakdown.

In addition, there are three pseudo-types of breakdown of great practical importance: discharge, electrochemical, and mechanical. Dielectric discharge is associated with a gaseous breakdown at the surface or within included pores of the material. Electrochemical breakdown is a result of chemical reactions that, over time, reduces the dielectric strength. Mechanical breakdown is precipitated by cracks or other defects that distort the applied field. *See* BREAKDOWN POTENTIAL; ELECTRICAL BREAKDOWN; ELECTRICAL CONDUCTION IN GASES.

Dielectric constant. An insulator is also known as a dielectric. The dielectric constant k is defined as the ratio of the capacitance of a flat-plate condenser, or capacitor, with a dielectric between the plates to that with a vacuum between the plates; this ratio is also the relative permittivity of the dielectric. The dielectric constant is a measure of the ability of the insulator to increase the stored charge on the plates of the condenser as a result of the displacement of charged species within the insulator. Displacement of a charge q through a distance δ creates an electric dipole of moment $p = q\delta$. The displacement of charge takes time, so in an alternating-current circuit $k = k'\text{-}ik''$ is a complex number; a measure of the losses introduced by the charge displacements is the loss tangent, $\tan\delta \equiv k''/k'$. Low dielectric constants are desirable for the insulation of alternating-current circuits.

In ceramics and glasses, four primary short-range motions contribute to the total polarization of the dielectric: electronic, atomic, dipolar, and interfacial. Electronic polarization is due to a shift of the valence-electron cloud relative to the atomic nucleus. It gives a resonance absorption peak at optical frequencies. Atomic polarization is due to a displacement of positive and negative ions with respect to each other. It gives a resonance in the infrared range 10^{12}–10^{13} Hz that may be quite broad where several types of ions are present. Dipolar, or orientational, polarization is a perturbation of the thermal motion of ionic or molecular dipoles to give a net dipolar orientation in the direction of the applied field. The preferred orientation of permanent dipoles (Stevels deformation polarization), which may be particularly important in liquid and gaseous insulators, occurs in the 10^{11}–10^{12}-Hz frequency range and the switching of dipoles between two equivalent equilibrium orientations in a solid occurs at lower frequencies, 10^3–10^6 Hz. Interfacial, or space-charge, polarization occurs where the motions of mobile charges are blocked by internal barriers such as two-phase interfaces. This mechanism produces a large capacitance and dielectric constant at low frequencies. Orientational and interfacial polarizations introduce losses in insulators from 10^{-3} to 10^9 Hz; atomic and electronic polarizations in glasses and ceramics are important for optical communication. *See* CAPACITANCE; DIELECTRIC MATERIALS; FERROELECTRICS; OPTICAL COMMUNICATIONS; PERMITTIVITY; POLARIZATION OF DIELECTRICS.

Mechanical properties. Successful application of solid insulating materials also depends on their mechanical properties. Insulation assemblies commonly must withstand thermal-expansion mismatch, tension, compression, flexing, or abrasion as well as a hostile chemical-thermal environment. The introduction of cracks promotes the penetration of moisture and other contaminants that promote failure, and the presence of pores may cause damaging corona discharge on the surface of a high-voltage conductor. As a result, composite materials and engineered shapes must be tailored to meet the challenges of a particular operational environment. *See* CORONA DISCHARGE.

For example, glasses and varnishes are used as sealants, and oil is used to impregnate high-voltage, paper-insulated cables to eliminate air pockets. Porcelain is a commonly used material for the suspension of high-voltage overhead lines, but it is brittle. Therefore, a hybrid insulator was developed that consists of a cylindrical porcelain interior covered by a mastic sealant and a silicone elastomer sheath heat-shrunk onto the porcelain core. The circular fins of the outer sheath serve to shed water. However, twisted-pair cables insulated with poly(tetrafluoroethylene) are used for high-speed data transmission where a small dielectric constant of the insulator material is needed to reduce signal attenuation. *See* CONDUCTOR (ELECTRICITY); ELECTRIC POWER TRANSMISSION; FLUORINE; PORCELAIN; SILICONE RESINS.

Liquid and gaseous insulators. Liquid or gas insulation provides no mechanical strength, but it may provide a cheap, flexible insulation not subject to mechanical failure. Biphenyls are used as insulating liquids in capacitors; alkyl benzenes in oil-filled cables; and polybutenes for high-pressure cables operating at alternating-current voltages as high as 525 kV. Sulfur hexafluoride (SF_6) is a nonflammable, nontoxic electron-attracting gas with a breakdown voltage at atmospheric pressure more than twice that of air. Fluorocarbons such as C_2F_6 and C_4F_8 as well as the Freons are also used, and breakdown voltages have been increased significantly in gas mixtures through a synergistic effect. Used as an insulating medium in high-voltage equipment at pressures up to 10 atm (1 megapascal), sulfur hexafluoride can reduce the size of electrical substations by a factor of 10 over that of air-insulated equipment. Enclosure of metal cable in a metallic conduit filled with sulfur hexafluoride gas has been used to shield nearby workers from exposure to high electric fields. *See* CIRCUIT BREAKER; FLUOROCARBON; POLYCHLORINATED BIPHENYLS.

Thermal conductivity. Finally, the ability to transfer heat may be an overriding consideration for the choice of an electric insulator. Electrical machines generate heat that must be dissipated. In electronic devices, the thermal conductivity of the solid substrate is a primary consideration. Where mechanical considerations permit, circulating liquid or gaseous insulation is commonly used to carry away heat. Liquids are particularly good transporters of heat, but they are subject to oxidation. In transformers, for example, the insulators are generally mineral or synthetic oils that are circulated, in some places with gaseous nitrogen to inhibit oxidation, to carry away the heat generated by the windings and magnetic core. Silicone fluids, while flammable, have a higher fire point than mineral oil and a lower dissipation factor. Circulating gaseous sulfur hexafluoride is also used for transformer insulation. Hydrogen is a gaseous insulator used to cool rotating electrical machines because it has a heat-conducting capacity over seven times that of air. *See* ELECTRIC ROTATING MACHINERY; ELECTRICAL INSULATION; HYDROGEN; TRANSFORMER.

John B. Goodenough

Bibliography. R. Bartnikas, *Electrical Insulating Liquids in Dielectrics*, vol. 3, 1994; R. Bartnikas et al. (eds.), *Engineering Dielectrics*, vol. 1, 1979, vol. 2A, 1983, vol. 3, 1987; C. Johnson, *Handbook of Electrical and Electronic Technology*, 1994; F. H. Kreuger, *Industrial High Voltage*, vol. 1: *Electric Fields, Dielectrics, Constructions*, 1991; W. T. Shugg, *Handbook of Electrical and Electronic Insulating Materials*, 2d ed., 1995.

Electric organ (biology)

An effector organ found in six different groups of fishes; output is an electric pulse (**Table 1**; **Fig. 1**). Voltages large enough to aid in prey capture or predator deterrence are produced by various strongly electric fishes. These include the electric eel (*Electrophorus electricus*) from South America; the electric catfish (*Malapterurus electricus*) from Africa; the family of electric rays, Torpedinidae, which are widely distributed in the world's oceans; and possibly the stargazer genus, *Astroscopus*, of the western Atlantic. Weakly electric fishes emit a lower voltage, the energy source for an active electrosensory system that monitors electrical impedance in the environment. These weak signals also serve in intra- and interspecific communication. There are three groups of weakly electric fishes. First, the South American knifefishes, the Gymnotiformes, are a large and diverse group of several families that also include *Electrophorus*. Second, the electrically active African Mormyriformes are composed of the numerous species of the family Mormyridae and the single species *Gymnarchus niloticus* in the family Gymnarchidae. Finally, many species of skates and rays of the family Rajidae occur in marine waters worldwide.

The strongly electric fishes are remarkable for the high voltage and power of their discharges (**Table 2**). For example the electric eel generates pulses in excess of 500 V. In contrast, a large torpedo generates a smaller voltage, about 50 V in air, but the maximum current is larger, and the peak pulse power can exceed 1 kW (or about 1 hp). The electric organs may constitute a substantial fraction of the body mass. The weakly electric fishes have much smaller organs emitting pulses of a few tenths of a volt to several volts; such amplitudes are still large compared to those recorded outside ordinary, nonelectric fishes.

Electric organ discharge is explicable in terms of the properties of ordinary excitable cells, that is, those of nerve and muscle. The single generating cells of electric organs are called electrocytes. They are modified striated muscle fibers that have lost the ability to contract, although they retain muscle filaments to varying degrees, and they produce electric signals as muscle fibers do.

Electrocyte operation. The basic operation of an electrocyte is as follows. Consider a flattened cell which contains a high concentration of potassium ions (K^+), a low concentration of sodium ions (Na^+), and an equivalent number of unspecified anions. Let

TABLE 1. Groups of electric fishes*

Common name	Family	Genera and species	Strength of organ discharge	Distribution
Skates, ordinary rays	Rajidae	*Raja*, many species, a number of other genera not known to be electric	Weak	Marine, cosmopolitan
Electric rays, torpedos	Torpedinidae	About 20 genera	Strong, up to 60 V or 1 kW, some perhaps weak	Marine, cosmopolitan
Mormyrids, elephant-nosed fish (many lack enlarged chin or snout)	Mormyridae	About 11 genera, several with many species	Weak	Fresh-water, Africa
Gymnarchus	Gymnarchidae	1 species, *Gymnarchus niloticus*	Weak	Fresh-water, Africa
Gymnotid eels, electric eel, knifefishes	Electrophoridae	1 species, *Electrophorus electricus*	Strong, more than 500 V	Fresh-water, South America
	Gymnotidae	1 species, *Gymnotus carapo*	Weak	Fresh-water, South America
	Sternopygidae	About 8 genera, a number of species	Weak	Fresh-water, South America
	Sternarchidae (Apteronotidae)	About 5 genera, a number of species	Weak	Fresh-water, South America
Electric catfish	Malapteruridae	1 species, *Malapterurus electricus*	Strong, more than 300 V	
Stargazers	Uranoscopidae	1 electric genus, *Astroscoupus,* several species	Strong (?), about 6 V in air from small animals	Marine, western Atlantic

*From M. V. L. Bennett, Electric organs, in W. S. Hoar and D. J. Randa ll (eds.), *Fish Physiology,* vol. 5, pp. 347–491, Academic Press, 1971.

it be bathed in a solution which is high in Na^+, and low in K^+. Initially its membrane is specifically permeable to potassium ions (K^+) over its entire surface, and in this condition the membrane potential is uniform and no currents flow (**Fig. 2***a*). The cell interior will be at some negative voltage, at which the tendency for K^+ to diffuse out and down its concentration gradient is just balanced by the negativity of the interior, tending to prevent the positively charged ions from diffusing out. This "equilibrium" potential is about 90 mV inside negative for most excitable cells. When the cell is activated, one of its membrane faces becomes much more permeable to Na^+. The potential at which this membrane would be at equilibrium in terms of Na^+ flow is about 50 mV inside positive, but the cell interior is negative. Thus, Na^+ flows inward through the activated face, making the interior more positive. The other membrane face that still is specifically permeable to K^+ is displaced from its equilibrium potential, and as a result K^+ flows outward through it. In effect, the two faces act as batteries oriented in series (Fig. 2*b*). All operate just this simply; apposed membranes have different permeabilities during activity, allowing ions to flow down their concentration gradients and thereby generating electric currents. The proximate sources of energy for organ discharge are the differences in ionic concentrations between interior and exterior of the electrocytes. During activity these concentrations become more alike; between periods of activity, metabolic energy is expended to restore the concentration differences. *See* BIOPOTENTIALS AND IONIC CURRENTS.

During activity, currents crossing all of the electrolyte membranes in series flow around the outside of the electric organ through the surrounding body

TABLE 2. Discharge characteristics of electric fishes

Species	Approximate discharge voltage, V	Duration, ms	Frequency	Number of electrocytes in series
Electrophorus electricus, electric eel	700	2	0–250	6000
Malapterurus electricus, electric catfish	300	1.5	0–200	?
Torpedo nobiliana, electric ray	60	5	0–150	1000
Astroscopus γ-graecum, stargazer	6	5	0–150	200
Gnathonemus petersii, mormyriform	5	0.3	2–100	100
Gymnarchus niloticus, mormyriform	1	2	250	140
Gymnotus carapo, gymnotiform knifefish	1	1.5	50–150	90
Sternarchus albifrons, sternarchid (apteronotid) knifefish	1	1	800	90
Sternopygus macrurus, gymnotiform knifefish	1	10	50 male, 150 female	90
Raja erinacea, rajid skate	2	100	0–5	200

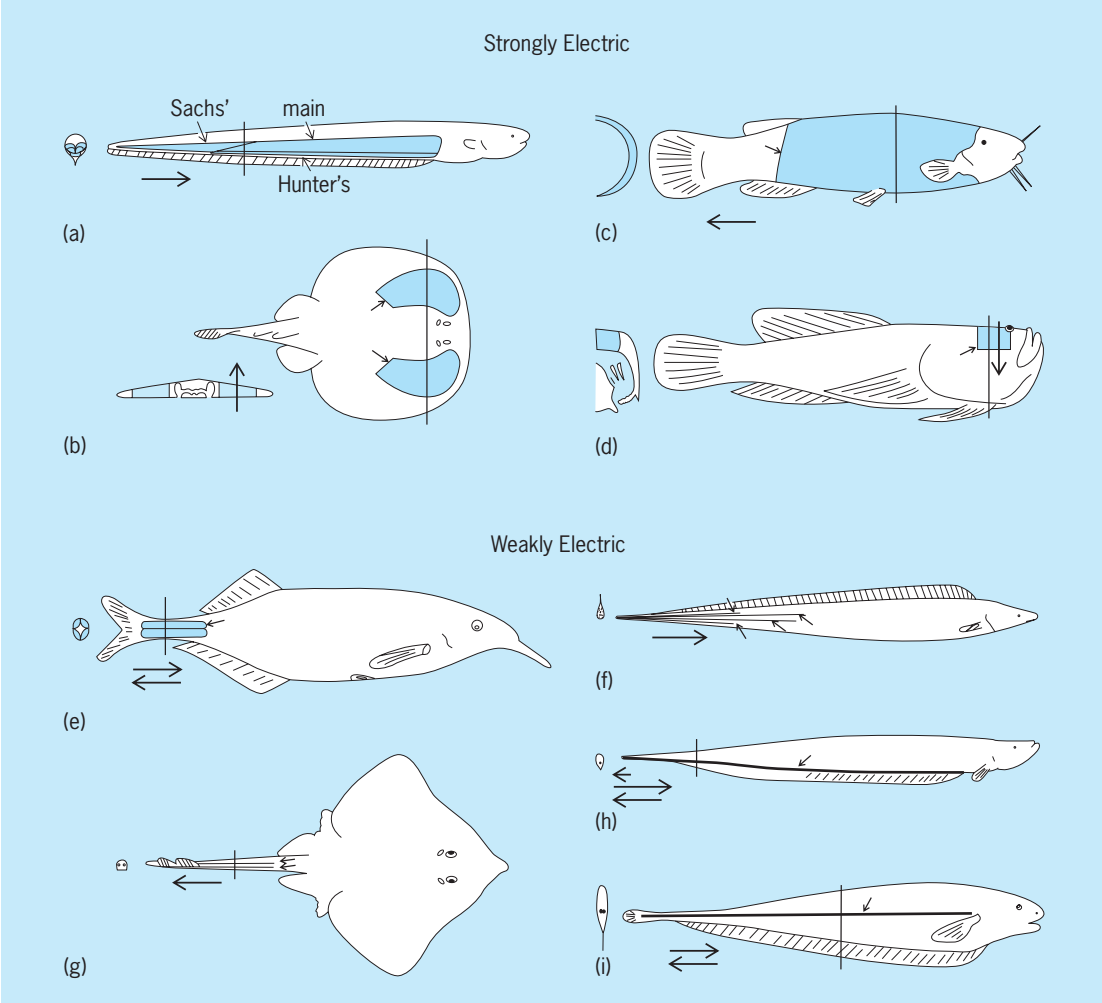

Fig. 1. Representative electric fishes: (*a*) *Electrophorus*; (*b*) *Torpedo*; (*c*) *Malapterurus*; (*d*) *Astroscopus*; (*e*) *Gnathonemus*; (*f*) *Gymnarchus*; (*g*) *Raja*; (*h*) *Gymnotus*; (*i*) *Sternarchus*. Electric organs are shaded or solid and indicated by small arrows. Cross-sectional views through the organs are also shown at the levels indicated by the solid lines. The large arrows indicate the direction of active current flow through the organs; more than one arrow indicates successive phases of activity, and the relative lengths are proportional to relative amplitudes of the phases. (*After W. S. Hoar and D. J. Randall, eds., Fish Physiology, vol. 5, Academic Press, 1971*)

and water, producing a field, similar in form to that around a bar magnet (Fig. 2*c*). A voltage drop occurs across any object in the field. If the objects differ in conductivity from water, they distort the pattern of current flow, and these distortions can be detected by the electrosensory system (see below).

Ions enter or leave a cell by passing through specific protein channels in the cell membrane. The extracellular fields generated by electrocytes are significantly larger than those generated by ordinary excitable cells, but the difference is not in the nature of the ionic permeabilities or the magnitude of the membrane voltages. Rather it is in the shape of the cells and the distribution and density of ionic channels. The maximum voltage generated by a single electrocyte is not much over 0.1 V. The extraordinarily large voltages that are generated by strongly electric organs result from the arrangement of many elements in series (Fig. 2*c*) and in parallel. For example, in the electric eel many thousands of cells occur in series (Table 2). In the torpedos there are hundreds of cells in series.

Electrocyte membranes. The arrangement of specific kinds of membrane differs in different electrocytes. In the strongly electric fishes, electrocytes are always flattened (hence the terms electroplaque and electroplax, **Fig. 3**). The simplest are those of the electric ray and stargazer. One face is densely innervated and generates only postsynaptic potentials (PSPs), as at the nerve-muscle junction. The PSPs result from a permeability increase to Na^+ and K^+, so that the membrane is at electrical equilibrium near 0 V instead of inside positive. The PSPs are produced by action of a neurotransmitter substance, acetylcholine. The transmitter is secreted by the innervating nerves. It combines briefly with receptor molecules in the electrocyte membrane, causing the ion channels to open; then it is inactivated by the enzyme acetylcholinesterase, which is found at high concentration at the synapses. The uninnervated face of these electrocytes is inexcitable; its permeability does not change when a potential is applied. At rest the uninnervated face has a much lower electric resistance than the innervated face. During

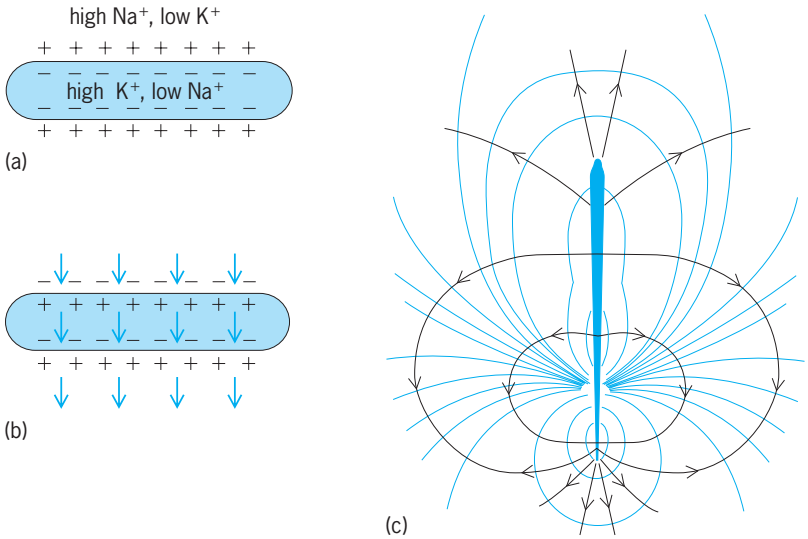

Fig. 2. Current flow through a single electrocyte and fields around an electric fish. (*a*) Cross section of a flattened cell at rest. Each face is specifically permeable to K^+, and the cell interior is at a negative potential. Because equal potentials are opposed, no current flows. (*b*) During activity that greatly increases the permeability to Na^+ in the upper face, this membrane moves toward inside positivity. As a result, two batteries act in series and current flows. (*c*) Flow of current (black lines) and isopotentials (colored lines) around a weakly electric gymnotid during a head positive phase of organ discharge. (*After W. Heiligenberg, Principles of Electrolocation and Jamming Avoidance in Electric Fish, Springer, 1977*)

PSPs, however, the resistance of the innervated face greatly decreases and the two faces have comparable resistances. These characteristics maximize the effectiveness of the cells as generating elements.

The electrocytes of the electric eel (and of several weakly electric relatives) are slightly more complex. In these cells, as in those of the electric ray and stargazer, one face is innervated and one is not, the uninnervated face being of low resistance and inexcitable. In addition to PSP channels, the innervated membrane has channels specifically permeable to Na^+; they open whenever the cell interior becomes about 10 mV or so more positive than its resting potential (depolarization). The PSPs generated by the innervated membrane cause some of these channels to open, greatly increasing the permeability to Na^+. The initial opening of Na^+ channels allows an influx of Na^+, which makes the interior yet more positive, causing a further opening of Na^+ channels, and so forth; the response is said to be regenerative. The positive feedback action continues until virtually all the Na^+ channels are open, and an impulse resembling the action potential of axons and muscle fibers is generated. In contrast to the electrically opened Na^+ channels, the PSP generation channels are little affected by membrane potential. In the electric eel the density of PSP channels is much lower than in the electric ray, and most of the change in membrane potential results from the electrically opened Na^+ channels. Two mechanisms can contribute to termination of the impulse. In the eel the Na^+ channels simply close after a brief period, and the membrane returns to the K^+ equilibrium potential. In others there are additional K^+ channels that are opened by depolarization somewhat more slowly than the Na^+ channels, and these speed the return of the membrane potential to its resting value.

The three classes of response, mediated by the PSP channels and by the Na^+ and K^+ channels opened by depolarization, can be mixed together in a number of different arrangements. In some weakly electric fishes, both innervated and uninnervated membranes generate impulses. First, current flows in one direction through the cell as the innervated membrane generates its impulse; this current then excites the uninnervated face, though with some delay, producing an impulse with current flow in the opposite direction. Thus, a biphasic external potential results. In the electric catfish, for example, one face generates a briefer impulse than the other. The resting resistances of the two faces are about equal, and the role of the briefer impulse appears to be to increase the K^+ permeability of that face, thus making the resistances of the two faces comparable during the later part of the longer lasting impulse, thereby increasing current flow. A similar phenomenon occurs in some skates, except that there are no electrically opened Na^+ channels. PSPs generated in the innervated face depolarize the uninnervated face, opening K^+ channels in it, and thus increasing current flow.

In several weakly electric fishes (for example, *Gymnarchus*) the uninnervated membrane has a very low permeability to K^+ or to any other ions. This membrane does have a large electrical capacity that permits significant current flow. Because the current flow is capacitative, the discharge has no direct current (dc) component. The absence of a dc component can be related to the functioning of the discharge in the electrosensory system.

Discharge. In all strongly electric fishes the activity of individual electrocytes is highly synchronized, thus maximizing the voltage and current produced. The discharges are monophasic pulses which are presumably more effective in stimulating prey or predator (A and A′ in **Fig. 4**). In weakly electric fishes the discharges are more diverse. There are a few with discharges like the strongly electric fishes, but others generate diphasic or triphasic pulses that need not be synchronous throughout the electric organ (B and B′ in Fig. 4). Triphasic discharges can result from out-of-phase addition of diphasic responses from two sets of electrocytes (*Gymnotus*). Alternatively a single electrocyte can generate a triphasic response (some mormyrids). The shape of electrocytes is also more variable in weakly electric fishes. The cells can be drum-shaped (*Gymnotus*, Fig. 3) or spindle-shaped (*Sternopygus*, Fig. 3). The innervation can occur on a thin process or stalk from the main body of the cell (*Malapterurus*, Fig. 3; some gymnotids). The stalks may also be multiple (mormyrids, Fig. 3; some gymnotids) and may even turn around and pass through holes in the flattened part to the opposite side before receiving their synapses (some mormyrids). Propagated activity in these stalks introduces a small component in overall organ discharge.

A number of gymnotids have a separate small "accessory" electric organ in the head region.

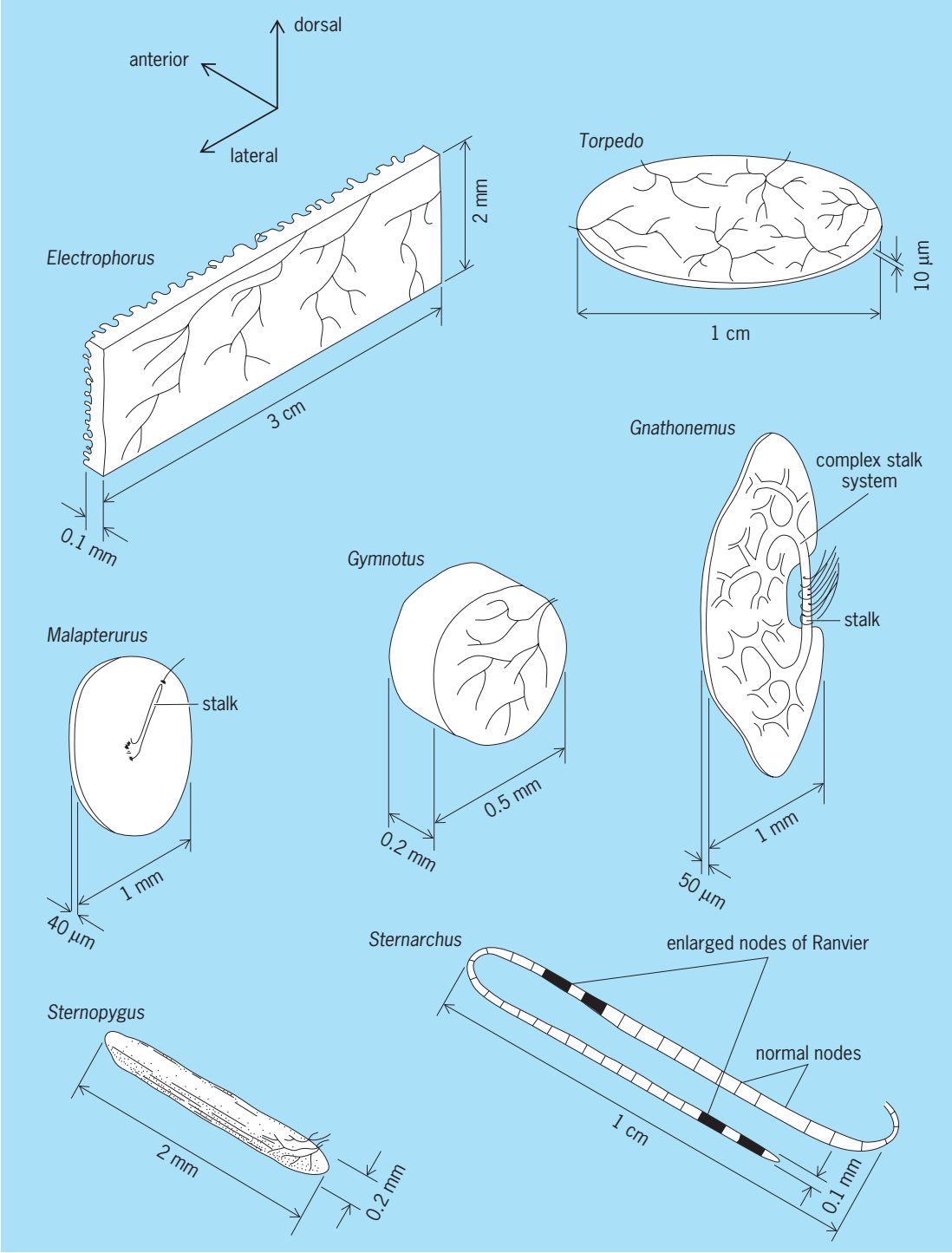

Fig. 3. Representative electrocytes. Body axes are indicated by the arrows except that the *Torpedo* electrocyte is shown ventral surface up. Innervation is indicated by branching, except for *Malapterurus* in which a single nerve fiber makes a single contact on the stalk tip. Also, the electrocyte of *Sternarchus* is a nerve fiber.

Electroreceptors are also concentrated on the head, and presumably the accessory organ functions in detecting localized objects in this region.

In strongly electric fishes the shape of the electric organ can be related to the conductivity of the environment (Fig. 1). In the marine forms, the torpedos and the stargazers, the organ is short and wide, particularly in the torpedos. In the fresh-water forms, the electric catfish and eel, the organs are long and relatively narrower. In the marine, as compared to fresh-water forms, fewer cells occur in series and more in parallel, and the voltages are lower and the currents higher in the higher-conductance medium.

In weakly electric fishes, including the marine rajids, the organs are generally long and slender. They have fewer than a dozen cells in parallel and of the order of 100 in series (Table 2).

In view of the high power that can be generated

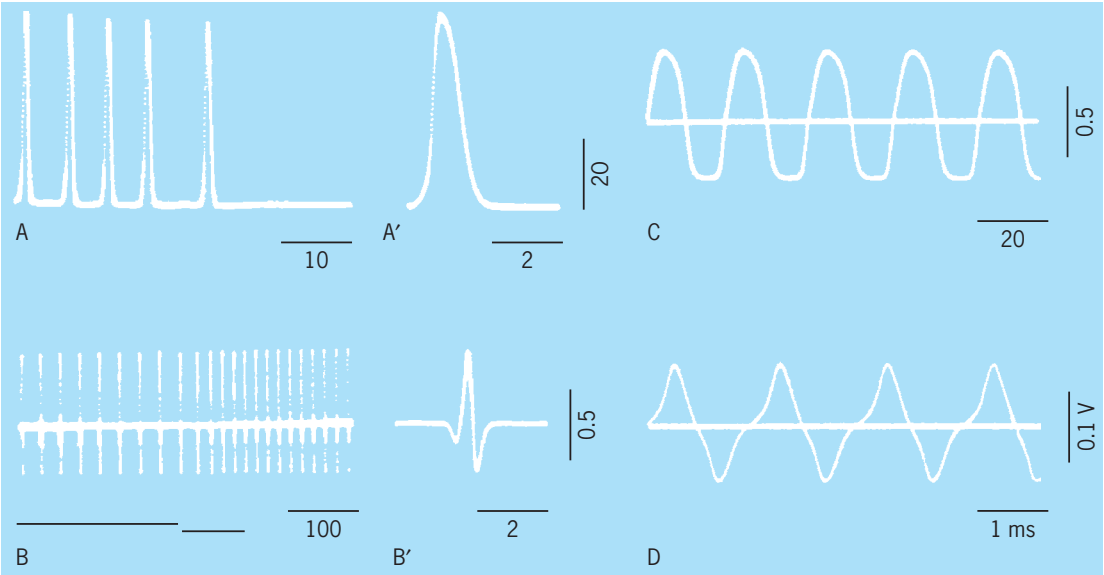

Fig. 4. Patterns of electric organ discharge. (A) An electric catfish; potentials recorded between head and tail in a small volume of water with head negativity upward. Pulses attain a maximum frequency of 190/s. (A′) Single pulse. (B–D) Weakly electric gymnotids immersed in water; discharges recorded between head and tail, head positivity upward. (B) *Gymnotus*; pulses are emitted at a basal frequency of approximately 35/s. Tapping the side of the fish at the time indicated by the downward step in the lower trace causes an acceleration up to about 65/s. (B′) Faster sweep showing the triphasic pulse shape. (C) *Sternopygus*; pulse frequency is about 55/s. The horizontal line indicates the zero potential level. (D) *Sternarchus*; pulse frequency is about 800/s. The horizontal line indicates the zero potential level. (*After W. S. Hoar and D. J. Randall, eds., Fish Physiology, vol. 5, Academic Press, 1971*)

by strongly electric fishes, it is not surprising that they discharge the organs only infrequently, and generally in response to an external stimulus (A in Fig. 4). The fresh-water weakly electric fishes, however, emit pulses continually (B–D in Fig. 4), and the range of frequency is wide, from a few pulses per second to more than a thousand. The frequencies and discharge patterns are species-specific and differ considerably. Two broad types can be defined, pulse species and continuous-wave species. The pulse species emit brief pulses separated by a much longer interval; nonspecific sensory stimulation usually causes an increase in the frequency of discharge (B and B′ and Fig. 4). Continuous-wave species emit monophasic or biphasic pulses separated by intervals of about the same duration as the pulses (C and D in Fig. 4). These species generally modify their discharge frequency very little, although all show small changes when presented with a stimulus at a frequency close to their own. These changes, termed jamming avoidance responses, are presumed to reduce interference in electrolocation.

The discharges of rajids under normal conditions have not been described. They are evoked with difficulty by handling and consist of long-lasting, repetitive, asynchronous discharges of the generating cells.

Neural control. The neural control of electric organ discharge has been studied in representative species of most groups. Generally the frequency is set in a command nucleus, a compact group of neurons in the brain stem. The command neurons integrate various neural inputs, both excitatory and inhibitory, and the group of neurons generates a highly synchronous impulse. Each synchronous discharge is the command signal that is responsible for each electric organ discharge. Synchronization is mediated by electrical synapses that interconnect the cells. In the continuously discharging fishes the command neurons are probably spontaneously active, but they still receive sensory and other inputs that modulate discharge frequency. The command signal may be relaxed by one or two intervening groups of cells, the ultimate group innervating the electrocytes. (In the sternarchids the electric organ actually consists of the axons of the final neurons; Fig. 3.) Since different parts of the electric organ are at significantly different distances from the brain, compensatory mechanisms are present to adjust for the variations in nerve conduction and thus to ensure synchrony of organ discharge. For example, nerves to nearer parts may take more devious pathways, and they may conduct impulses at lower velocities.

Electrosensory systems. A knowledge of the electrosensory systems is integral to understanding the biological significance of electric organs. A number of groups of fishes, both electric and nonelectric, have receptors specialized for the detection of electric fields. These groups include all the elasmobranchs; the primitive bony fishes, lungfishes, bichirs, sturgeons, and paddlefish; the catfishes; and the two teleostean groups of fresh-water electric fishes. Electrosensory systems can be divided into two categories: those activated by fields from extrinsic sources, or passive systems, and those activated by the fishes' electric organs, or active systems. The passive systems generally are sensitive to voltages of low frequency such as are generated by respiratory movements in fishes, muscle or skin lesions, and geomagnetic sources. They can be extraordinarily sensitive. For example, elasmobranchs can respond to

voltage gradients as low as 0.01 microvolt/cm or 1 V uniformly distributed over 1000 km. The active systems utilize the electric organ discharge to test the conductivity of the surrounding medium, thereby revealing changes or inhomogeneities. The active systems are also involved in intra- and interspecific communication.

The electroreceptors are cutaneous organs distributed over the body surface. They are innervated by cranial nerves of the acousticolateralis system. The gymnotids and mormyrids have separate sets of receptors for active and passive systems; the electric organ discharges are often of high frequency with little dc component, and the two receptor systems detect electric fields over different frequency bands. The Mormyriformes have a third set of receptors, very sensitive to high frequencies, that appears to be specialized for detecting the electric discharges of other fishes. The diversity of organ discharges in terms of both frequency and pulse shape may be partly accounted for in terms of interspecific communication and interaction.

In the skate the electric organ discharge is sufficiently long-lasting to be capable of activating the passive electrosensory system. It is unknown whether their discharges are utilized for electrolocation or for communication.

Evolution. The electrosensory systems allow an explanation of the evolution of strongly electric organs. Charles Darwin supposed that a weakly electric organ would have been an intermediate stage in the gradual development of a strongly electric organ, but could see no use for a weak organ. Later, Hans Lissman, who initially described the active electrosensory system, proposed that they provide a function for Darwin's hypothesized intermediate stage. This notion is supported by the electric eel and at least one torpedinid; both have strongly and weakly electric organs, and the eel is known to utilize the weak pulses for electrolocation in the same way as the related weakly electric gymnotids. The electric catfish has a passive electrosensory system, as do most nonelectric catfish. No weakly electric catfish are known, and the hypothetical intermediate leading to *Malapterurus* is either extinct or has not yet been discovered. The electrosensory capabilities of the stargazer have not been adequately investigated.

Electrocution. A question often asked about the strongly electric fishes is how they keep from electrocuting themselves. When immersed in water, they do not appear to be at all affected by their own discharges, although their electroreceptors are undoubtedly activated. In air, when the organ is not loaded down by the conductivity of the water, they often do twitch when they discharge their organs, but as would be expected, they are much more resistant than other fishes. One factor is that their nerves and central nervous system are well insulated by many layers of connective and fatty tissue. Whether their brains are more resistant to stunning by a given current density has not been determined; certainly the excitability of individual cells is no different from that in animals in general.

Source of data. Electric fishes have proved useful in defining extremes of excitability properties. Historically they served to validate Luigi Galvani's proposal of animal electricity, and in modern times have provided, and are continuing to provide, a wealth of biophysical data of general and comparative interest. They also supply important starting materials for the biochemical characterization of excitable membranes, for example, the acetylcholine receptor at the synapse. Because evolution has led to large volumes of tissue with high concentrations of the relevant macromolecules, the electric organs provide a rich source for isolation and analysis. *See* ACETYLCHOLINE; SYNAPTIC TRANSMISSION.

Michael V. L. Bennett

Electric power generation

The production of bulk electric power for industrial, residential, and rural use. Although limited amounts of electricity can be generated by many means, including chemical reaction (as in batteries) and engine-driven generators (as in automobiles and airplanes), electric power generation generally implies large-scale production of electric power in stationary plants designed for that purpose. The generating units in these plants convert energy from falling water, coal, natural gas, oil, and nuclear fuels to electric energy. Most electric generators are driven either by hydraulic turbines, for conversion of falling water energy; or by steam or gas turbines, for conversion of fuel energy. Limited use is being made of geothermal energy, and developmental work is progressing in the use of solar energy in its various forms. *See* GENERATOR; PRIME MOVER.

Fossil-fuel and hydroelectric plants, and various types of advanced power sources, are discussed in detail below. For a discussion of nuclear power plants *see* NUCLEAR POWER; NUCLEAR REACTOR.

General Considerations

An electric load (or demand) is the power requirement of any device or equipment that converts electric energy into light, heat, or mechanical energy, or otherwise consumes electric energy as in aluminum reduction, or the power requirement of electronic and control devices. The total load on any power system is seldom constant; rather, it varies widely with hourly, weekly, monthly, or annual changes in the requirements of the area served. The minimum system load for a given period is termed the base load or the unity load-factor component. Maximum loads, resulting usually from temporary conditions, are called peak loads, and the operation of the generating plants must be closely coordinated with fluctuations in the load. The peaks, usually being of only a few hours' duration (**Figs. 1** and **2**), are frequently served by gas or oil combustion-turbine or pumped-storage hydro generating units. The pumped-storage type utilizes the most economical off-peak (typically 10 P.M. to 7 A.M.) surplus generating capacity to

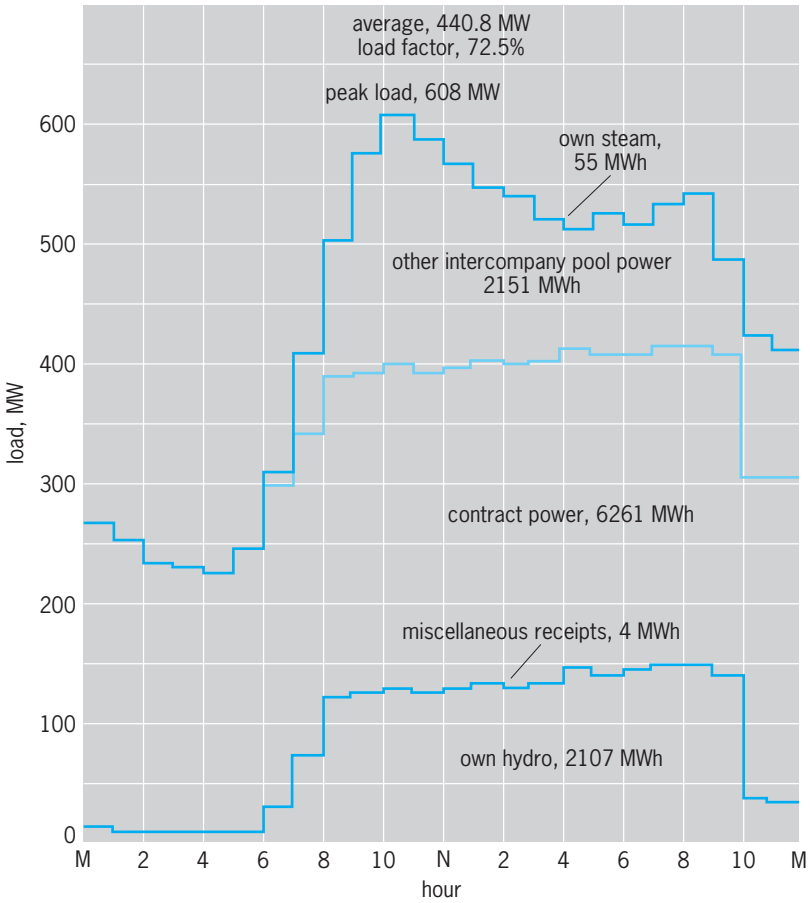

Fig. 1. Load graph indicates net system load of a metropolitan utility for typical 24-h period (midnight to midnight), totaling 10.578 MWh. Such graphs are made to forecast probable variations in power required.

released through hydraulic turbine generators during peak periods. This type of operation improves the capacity factors or relative energy outputs of base-load generating units and hence their economy of operation.

Actual variations in the load with time are recorded, and from these data load graphs are made to forecast the probable variations of load in the future. A study of hourly load graphs (Figs. 1 and 2) indicates the generation that may be required at a given hour of the day, week, or month, or under unusual weather conditions. A study of annual load graphs and forecasts indicates the rate at which new generating stations must be built; they are an inseparable part of utility operation and are the basis for decisions that profoundly affect the financial requirements and overall development of a utility.

Generating unit sizes. The size or capacity of electric utility generating units varies widely, depending upon type of unit; duty required, that is, base-, intermediate-, or peak-load service; and system size and degree of interconnection with neighboring systems. Base-load nuclear or coal-fired units may be as large as 1200 MW each, or more. Intermediate-duty generators, usually coal-, oil-, or gas-fueled steam units, are of 200 to 600 MW capacity each. Peaking units, combustion turbines or hydro, range from several tens of megawatts for the former to hundreds of megawatts for the latter. Hydro units, in both base-load and intermediate service, range in size up to 825 MW.

The total installed generating capacity of a system is typically 20 to 30% greater than the annual

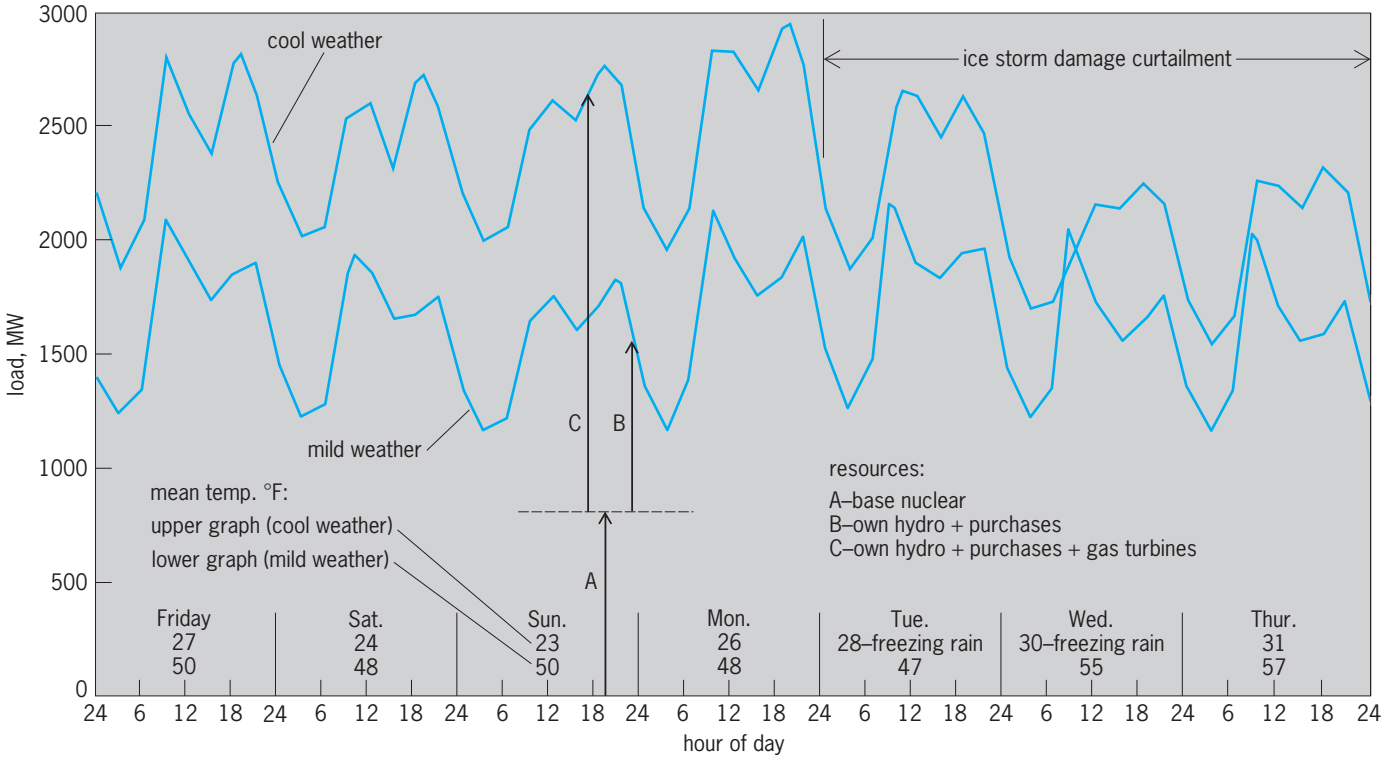

Fig. 2. Examples of northwestern United States electric utility weekly load curves showing same-year weather influence.

predicted peak load in order to provide reserves for maintenance and contingencies.

Power-plant circuits. Both main and accessory circuits in power plants can be classified as follows:

1. Main power circuits to carry the power from the generators to the step-up transformers and on to the station high-voltage terminals.
2. Auxiliary power circuits to provide power to the motors used to drive the necessary auxiliaries.
3. Control circuits for the circuit breakers and other equipment operated from the plant's control room.
4. Lighting circuits for the illumination of the plant and to provide power for portable equipment required in the upkeep and maintenance of the plant. Sometimes special circuits are installed to supply the portable power equipment.
5. Excitation circuits, which are so installed that they will receive good physical and electrical protection because reliable excitation is necessary for the operation of the plant.
6. Instrument and relay circuits to provide values of voltage, current, kilowatts, reactive kilovoltamperes, temperatures, pressures, flow rates, and so forth, and to serve the protective relays.
7. Communication circuits for both plant and system communications. Telephone, radio, transmission-line carrier, and microwave radio may be involved.

It is important that reliable power service be provided for the plant itself, and for this reason station service is usually supplied from two or more sources. To ensure adequate reliability, auxiliary power supplies are frequently provided for start-up, shutdown, and communication services.

Generator protection. Necessary devices are installed to prevent or minimize other damage in cases of equipment failure. Differential-current and ground relays detect failure of insulation, which may be due to deterioration or accidental overvoltage. Overcurrent relays detect overload currents that may lead to excessive heating; overvoltage relays prevent insulation damage. Loss-of-excitation relays may be used to warn operators of low excitation or to prevent pulling out of synchronism. Bearing and winding overheating may be detected by relays actuated by resistance devices or thermocouples. Overspeed and lubrication failure may also be detected.

Not all of these devices are used on small units or in every plant. The generator is immediately deenergized for electrical failure and shut down for any overlimit condition, all usually automatically. *See* ELECTRIC PROTECTIVE DEVICES.

Voltage regulation. This term is defined as the change in voltage for specific change in load (usually from full load to no load) expressed as percentage of normal rated voltage. The voltage of an electric generator varies with the load and power factor; consequently, some form of regulating equipment is required to maintain a reasonably constant and predetermined potential at the distribution stations or load centers. Since the inherent regulation of most alternating-current (ac) generators is rather poor (that is, high percentagewise), it is necessary to provide automatic voltage control. The rotating or magnetic amplifiers and voltage-sensitive circuits of the automatic regulators, together with the exciters, are all specially designed to respond quickly to changes in the alternator voltage and to make the necessary changes in the main exciter or excitation system output, thus providing the required adjustments in voltage. A properly designed automatic regulator acts rapidly, so that it is possible to maintain desired voltage with a rapidly fluctuating load without causing more than a momentary change in voltage even when heavy loads are thrown on or off.

In general, most modern synchronous generators have excitation systems that involve rectification of an ac output of the main or auxiliary stator windings, or other appropriate supply, using silicon controlled rectifiers or thyristors. These systems enable very precise control and high rates of response. *See* SEMICONDUCTOR RECTIFIER; VOLTAGE REGULATOR.

Generation control. Computer-assisted (or on-line controlled) load and frequency control and economic dispatch systems of generation supervision are being widely adopted, particularly for the larger new plants. Strong system interconnections greatly improve bulk power supply reliability but require special automatic controls to ensure adequate generation and transmission stability. Among the refinements found necessary in large, long-distance interconnections are special feedback controls applied to generator high-speed excitation and voltage regulator systems.

Synchronization of generators. Synchronization of a generator to a power system is the act of matching, over an appreciable period of time, the instantaneous voltage of an alternating-current generator (incoming source) to the instantaneous voltage of a power system of one or more other generators (running source), then connecting them together. In order to accomplish this ideally the following conditions must be met:

1. The effective voltage of the incoming generator must be substantially the same as that of the system.
2. In relation to each other the generator voltage and the system voltage should be essentially 180° out of phase; however, in relation to the bus to which they are connected, their voltages should be in phase.
3. The frequency of the incoming machine must be near that of the running system.
4. The voltage wave shapes should be similar.
5. The phase sequence of the incoming polyphase machine must be the same as that of the system.

Synchronizing of ac generators can be done manually or automatically. In manual synchronizing an operator controls the incoming generator while observing synchronizing lamps or meters and a synchroscope, or both. The operator closes the connecting switch or circuit breaker as the synchroscope needle slowly approaches the in-phase position.

Automatic synchronizing provides for automatically closing the breaker to connect the incoming machine to the system, after the operator has properly

adjusted voltage (field current), frequency (speed), and phasing (by lamps or synchroscope). A fully automatic synchronizer will initiate speed changes as required and may also balance voltages as required, then close the breaker at the proper time, all without attention of the operator. Automatic synchronizers can be used in unattended stations or in automatic control systems where units may be started, synchronized, and loaded on a single operator command. *See* ALTERNATING-CURRENT GENERATOR; PHASE-ANGLE MEASUREMENT; SYNCHROSCOPE.

Eugene C. Starr

Fossil-Fuel Plants

Fossil fuels are of plant or animal origin and consist of hydrogen and carbon (hydrocarbon) compounds. The most common fossil fuels are coal, oil, and natural gas. The less common ones include peat, oil shale, and biomass (wood and so forth), as well as various waste or by-products such as steel mill blast furnace gas, coke-oven gas, and refuse-derived fuels. Fossil-fuel electric power generation uses the combustion heat energy from these fuels to produce electricity. *See* COAL; COKE; FOSSIL FUEL; NATURAL GAS; OIL SHALE; PEAT; PETROLEUM.

Steam power plants. A fossil-fuel steam power plant operation essentially consists of four steps (**Fig. 3**): (1) Water is pumped at high pressure to a boiler, where (2) it is heated by fossil-fuel combustion to produce steam at high temperature and pressure. (3) This steam flows through a turbine, rotating an electric generator (connected to the turbine shaft) which converts the mechanical energy to electricity. (4) The turbine exhaust steam is condensed by using cooling water from an external source to remove the heat rejected in the condensing process. The condensed water is pumped back to the boiler to repeat the cycle. Figure 3 also shows features to increase cycle efficiency, including preheating of the boiler feedwater by using steam extracted from the turbine, and reheating the high-pressure turbine exhaust steam before it enters the intermediate-pressure turbine.

Steam cycle. Modern fossil-fuel power plants are based on a steam cycle first proposed by W. J. M. Rankine in 1908. The basic Rankine cycle underwent conceptual evolution through the 1950s that improved cycle efficiency to current levels but which also increased the complexity of building and operating a fossil-fuel power plant. Commercial development lagged behind conceptual cycle development because of the unavailability of suitable materials and fabrication techniques to accommodate the higher steam pressures and temperatures required to increase cycle efficiency. **Table 1** summarizes this steam cycle evolution with respect to turbine inlet steam pressure, turbine heat rate, and turbine cycle and overall thermal efficiencies, along with related available main steam piping materials from the 1930s through the 1950s. *See* RANKINE CYCLE; THERMODYNAMIC CYCLE; VAPOR CYCLE.

During the 1960s and 1970s, there was no significant cycle efficiency improvement, primarily due to the advent of nuclear power and the abundant supply of low-cost fossil fuel. In the 1980s, the interest in cycle efficiency improvement was renewed due to the uncertain future of nuclear energy and escalating fossil-fuel costs. However, due to the return of relative cost and price stability of fossil fuels since the late 1980s, interest in cycle efficiency

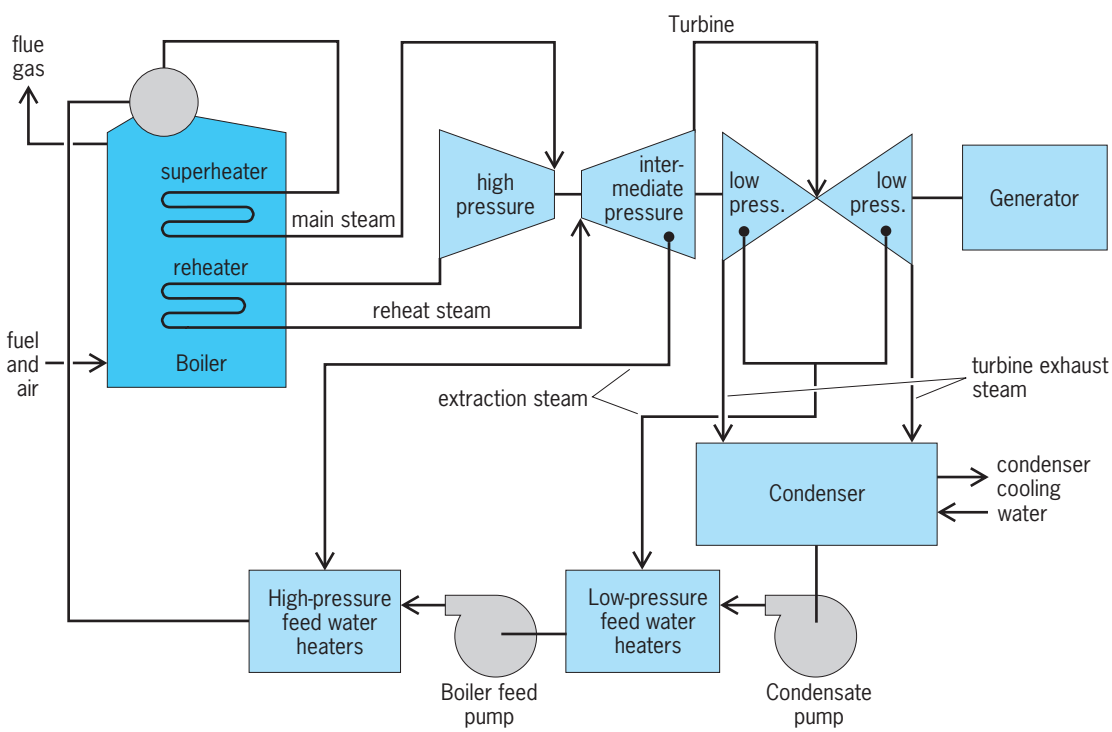

Fig. 3. Steam power plant flow diagram.

TABLE 1. Steam cycle development in fossil-fuel power plant

Years	Steam cycle, psig/°F/°F* (MPa/°C/°C)	Main steam piping material	Turbine cycle† heat rate, Btu/kWh (MJ/kWh)	Thermal efficiency (%)	
				Turbine cycle†	Overall‡
Late 1930s	1250/950 (8.7/510)	Chromium-molybdenum alloy(0.2% C, 0.5% Cr, 0.5% Mo)	9350 (9.86)	36.50	29.2
Mid 1940s	1450/1000/1000 (10.1/538/538)	Chromium-molybdenum alloy (0.2% C, 1.0% Cr, 0.5% Mo)	8150 (8.60)	41.9	33.5
Late 1940s	2000/1050/1050 (13.9/566/566)	Chromium-molybdenum alloy (0.2% C, 1.0% Cr, 0.5% Mo)	7700 (8.12)	44.3	35.5
Early 1950s	2400/1100/1050	Chromium-molybdenum	7500 (7.91)	45.5	36.4

*Main steam pressure/main steam temperature/reheat steam temperature.
†Cycle heat rate and thermal efficiency are at rated load for a condenser backpressure of 1.43 in. Hg (36.3 mmHg or 4.84 kPa) absolute and a condensate temperature of 90°F (32.2°C).
‡Based on a boiler efficiency of 85% and a power plant auxiliary power requirement of 5% of plant output.

improvement has diminished significantly. In fact, few if any large steam power plants with main steam pressure, main steam temperature, and reheat steam temperature higher than 2400 psig, 1100 °F, and 1050 °F (16.6 MPa, 593 °C, and 566 °C) respectively (the level achieved in the early 1950s) have been built since 1990. With the rapid development of Asian and Pacific rim countries including China, there is renewed interest in smaller-size 200–300-MW units, with emphasis on reducing capital cost.

Plant size and facilities. A given power plant's size is determined by the utility company's need for power as dictated by the electrical demand growth forecast. The additional generating capacity needed is matched to commercially available turbine-generator capacity sizes. The boiler is selected to suit the turbine steam flow requirements, and the remaining accessory equipment is sized to serve the needs of the resulting boiler-turbine steam cycle.

A typical large fossil-fuel power plant consists of several major facilities and equipment (**Fig. 4**), including fuel handling and processing, boiler (including furnace), turbine and electric generator, condenser and condenser heat removal system, feedwater heating and pumping system, flue gas–cleaning system, and plant controls and control system.

1. *Fuel handling*. Fuel type determines the fuel-handling system requirements. Natural gas is normally delivered by pipeline directly and fired in the furnace, needing no special handling or storage. For coal and oil, which are usually delivered by rail, ship, or pipelines, the fuel-handling system includes fuel unloading, storage, reclaiming from storage, and fuel preparation for combustion (such as coal pulverizing and pneumatic conveying to the burners, and oil heating and atomizing). Coal is usually stored in open stockpiles and oil in closed tanks. Normally, 30 to 180 days of fuel requirement is stored at the plant site to maintain operation during fluctuations in fuel delivery.

2. *Boiler*. The fuel type also determines the size, design, and operation of the furnace and boiler, which become larger and more complex (in ascending order) for gas, oil, and coal. For coal, the size and complexity also increase with the decline in coal rank and quality (in ascending order)

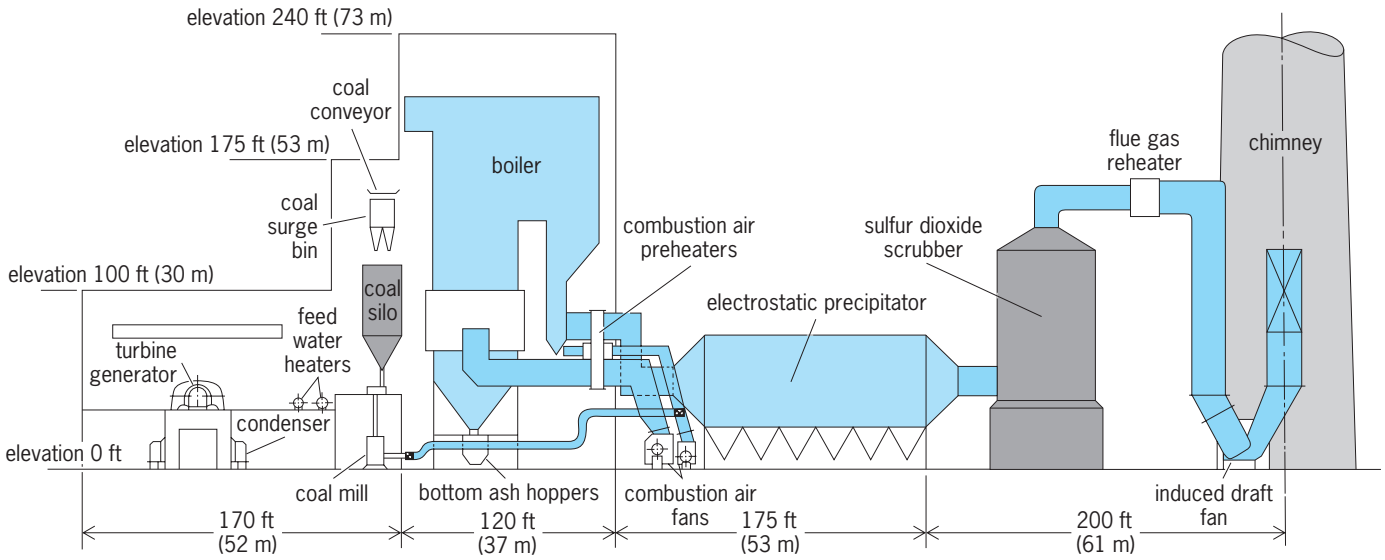

Fig. 4. Schematic cross section of a 600-MW coal-fired steam power plant.

for bituminous, subbituminous, lignite, and brown coal. Specialized boiler designs are required to handle peat, shale, and the various waste fuels. One evolving technology, fluidized-bed boilers, can effectively burn a wide range of low-quality solid fuels in the same boiler. Units of up to 200 MW are commercially proven and available. Most operating plants are in the size range of 120–180 MW. *See* BOILER; FIRE-TUBE BOILER; FLUIDIZED-BED COMBUSTION; STEAM-GENERATING FURNACE; STEAM-GENERATING UNIT; WATER-TUBE BOILER.

3. *Turbines*. Modern power plants, depending on the size and steam conditions, have single or multiple-casing turbines with high-, intermediate—, and low-pressure sections. High-pressure steam from the boiler after expanding through one or more turbines is exhausted to a condenser under vacuum, where the steam is condensed by cooling water. *See* STEAM TURBINE.

4. *Condenser*. About 50% of the fuel heat input to the boiler is ultimately rejected as waste heat in the condenser. Condenser heat-removal-system selection depends on the geography, climatology, and water availability at each plant site. There are basically two types of heat rejection systems: open (or once-through) cooling and closed cooling. The open system is the most simple and normally used where water is abundantly available. It takes water from an ocean, lake, or river and discharges the heated water to a downstream location. The closed system is used when it is determined not to return heated water to the water source due to environmental or other constraints. The closed system rejects heat to the atmosphere by using a secondary heat rejection system such as a cooling tower or pond. However, closed-system operation involves evaporating a portion of the water circulated, requiring a water supply capable of restoring the quantity of water evaporated. Dry cooling (with air condensers) is less efficient, but is used in arid areas where the water supply is limited. The temperature of the cooling medium (water or air) at the condenser inlet determines the cycle thermal efficiency by establishing the steam condensation temperature. *See* COOLING TOWER; STEAM CONDENSER; VAPOR CONDENSER.

5. *Feedwater heating*. The condensed steam is pumped back to the boiler inlet to complete the cycle, with heat from turbine extraction steam transferred along the way to increase cycle efficiency. Condensate pumps deliver feedwater from the condenser through low-pressure feedwater heaters to the boiler feed pumps, which increase the feedwater pressure and pump it through high-pressure feedwater heaters to the boiler inlet. A modern power plant typically has six feedwater heaters, which can be increased in number to improve cycle efficiency, but with offsetting increases in plant complexity and cost.

6. *Flue gas and liquid waste*. Fossil-fuel power plants produce gaseous and liquid wastes which are treated prior to discharge to minimize effects on the environment. In many countries, environmental regulations define the extent of flue gas and liquid waste treatment required.

Major pollutants in the flue gas include particulates (except for natural gas fuel), sulfur dioxide, and nitrogen oxides. Incombustible matter in coal and oil yields particulates upon combustion which are collected in devices such as electrostatic precipitators and fabric filters (baghouses). These devices are capable of removing more than 99% of the particulate matter. *See* AIR FILTER; ELECTROSTATIC PRECIPITATOR.

A variety of chemical processes is available for the removal of sulfur dioxide from the flue gas. Flue gas desulfurization systems use alkaline chemicals such as lime, limestone, sodium, or magnesium compounds in solution, slurry, or dry form to react with and remove sulfur dioxide. The choice depends on the fuel sulfur content, performance requirement, capital and operating costs, and market for the desulfurization by-products. The throw-away wet limestone slurry process is the most widely used in the United States due to the limited market for by-products. A similar process producing gypsum as the by-product is common in Europe and Japan. *See* GAS ABSORPTION OPERATIONS.

Nitrogen oxides emission can be reduced by combustion modifications which limit their formation during combustion. These techniques involve reduction of flame temperature by prolonging the combustion time (staged combustion) and reducing the combustion air quantity (low–nitrogen oxides burners).

Processes to remove nitrogen oxides after their formation involve injection of ammonia into the hot flue gas in the presence or absence of catalysts. They are referred to as the selective catalytic reduction and noncatalytic reduction processes, respectively. The selective catalytic reduction process is widely used in Japan and Europe, particularly in Germany, in both large-scale electric utilities and in small-scale industrial boilers. The noncatalytic reduction process is used in the United States in some small-scale industrial boilers. Typical nitrogen oxide reductions with the selective catalytic reduction process range from 70 to 90% and with the noncatalytic reduction process from 30 to 70%.

Waste liquid pollutants include suspended solids, toxic metal compounds, dissolved chlorine, and high-acidity or high-alkalinity, high-temperature, and oily compounds. Treatment prior to discharge or reuse usually involves one or more of the following: neutralization, sedimentation, clarification, filtration, dechlorination, and incineration. *See* CLARIFICATION; FILTRATION; SEDIMENTATION (INDUSTRY).

7. *Plant controls*. The power plant size and steam cycle pressure and temperature determine plant controls and control systems requirements. In ascending order of complexity and sophistication, these include pneumatic analog controls, electronic analog controls, hard-wired digital controls, and distributed digital control with microprocessors and minicomputers. Modern power plants typically use pneumatic local controls and electronic analog remote

TABLE 2. Typical cost distribution of major systems and equipment for a coal-fired steam power plant*

Systems and equipment	% of capital cost
Boiler (including furnace)	30
Turbine and electric generator	18
Flue gas desulfurization system	16
Particulate collection system	7
Coal-handling equipment	5
Condenser and cooling power	5
Ash-handling equipment	3
Pumps	3
Chimney	2
Water treatment	1
Other mechanical equipment	10

*Excluding land, site preparation, and building costs.

controls with electromechanical or solid-state interlocking devices and variable-speed motor drives. *See* CONTROL SYSTEMS; DIGITAL CONTROL; DISTRIBUTED SYSTEMS (CONTROL SYSTEMS); MICROCOMPUTER; MICROPROCESSOR; PROCESS CONTROL.

A typical cost distribution by major systems and equipment for a coal-fired steam power plant is given in **Table 2**.

Gas turbine plants. Power plants with gas turbine–driven electric generators are often used to meet short-term peaks in electrical demand. They are generally small (up to 150–200 MW), have a low thermal efficiency, require clean fuels such as natural gas or light oils, and consequently are more expensive to operate. However, these plants have a relatively low capital cost and can be used in areas with low power demand and limited water availability. Diesel engine–driven electric generators are also used under similar conditions. *See* DIESEL ENGINE.

Gas turbine power plants use atmospheric air as the working medium, operating on an open cycle where air is taken from and discharged to the atmosphere and is not recycled. In a simple gas turbine plant (**Fig. 5**), air is compressed and fuel is injected into the compressed air and burned in a combustion chamber. The combustion products expand through a gas turbine and exhaust to the atmosphere. Variations of this basic operation to increase cycle efficiency include regeneration (where exhaust from the turbine is used to preheat the compressed air before it enters the combustion chamber) and reheating (where the combustion gases are expanded in more than one stage and are reheated between stages).

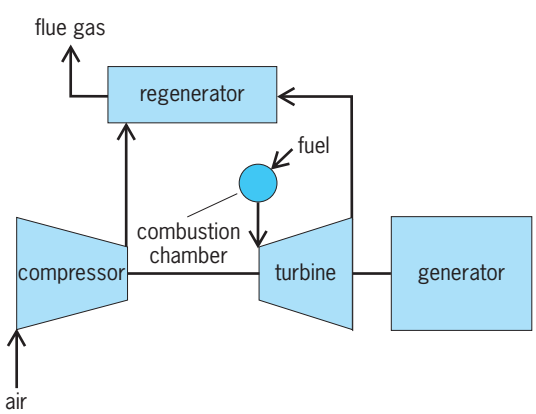

Fig. 5. Gas turbine plant flow diagram.

Combined-cycle plants use both gas and steam turbines and offer overall thermal efficiency improvement and operating economy. The gas turbine exhaust heat is used to generate steam for operating steam turbine generators. Combined-cycle plants with overall thermal efficiencies greater than 50%, where operating economies justify it, are commercially proven and available. Some very large plants in the 500–1500-MW size range are in commercial operation. *See* GAS TURBINE; STEAM TURBINE.

Cogeneration is another method of improving overall thermal efficiency, where steam exhausted from the steam turbine generator is used in industrial processes. Substantial energy is saved as heat rejected in the condenser is put to use. *See* COGENERATION.

Advanced concepts. Advanced fossil-fuel power generation concepts such as integrated coal gasification combined cycle and magnetohydrodynamics offer the promise of significant overall thermal efficiency improvement. In the integrated coal gasification combined cycle, gas produced from coal (after proper purification) is burned and expanded through a gas turbine, with the exhaust gas heat used to generate steam for a steam turbine generator. In the magnetohydrodynamic process, an electrically conducting medium such as hot flue gas from a coal combustor (seeded with alkali metal salts to increase the flue gas electrical conductivity) is passed through a magnetic field (magnetohydrodynamic channel) to produce electricity. The magnetohydrodynamic channel exhaust gas is then used to generate steam in a conventional steam power plant. The alkali salts are recovered and reused. Power generation based on these concepts can reach a thermal efficiency of 48–52%, whereas a steam turbine plant has an efficiency of 30–40% and a conventional gas turbine plant 22–28%. *See* COAL GASIFICATION; MAGNETOHYDRODYNAMIC POWER GENERATOR.

Maris T. Fravel; Natarajan Sekhar

Hydroelectric Plants

Hydroelectric generation is an attractive source of electric power because it is a renewable resource and a nonconsumptive use. In the broadest sense, hydroelectric power is a form of solar power; the resource is renewed by the solar cycle in which water is evaporated from the oceans, transported by clouds, and falls as precipitation on the landmasses, and returns through rivers to the ocean, generating power on the way. Hydroelectric power can be defined as the generation of electricity by flowing water; potential energy from the weight of water falling through a vertical distance is converted to electrical energy. The amount of electric power P that can be

generated is given by the equation below, where Q

$$P = \frac{Qh}{k} \qquad \text{kW}$$

is the volume flow of water, h is the height through which the water falls, and k is a constant equal to 11.8 when Q is given in cubic feet and h in feet (English units), and equal to 0.102 when Q is in cubic meters and h is in meters (SI units).

Development. In the United States, the first hydroelectric generating station was built in 1882 in Appleton, Wisconsin. The total power output was sufficient to operate some 250 electric lights. Rapid expansion followed, and by 1900 about 300 hydroelectric plants were in operation worldwide. This rapid expansion was a consequence of the development in the nineteenth century of the hydraulic reaction turbine. Also, by the same date, the concept of pumped storage had been introduced in Switzerland, but extensive development in the United States did not occur until the advent of the reversible pump–turbine design in the early 1950s.

Major components. A typical hydroelectric development consists of a dam to divert or store water; waterways such as a forebay, canals, tunnels, and penstocks to deliver the water to the hydraulic turbine and a draft tube, tunnel, or tailrace to return the water to the stream; hydraulic turbines and governors; generators and exciters; electrical controls to provide protection and to regulate frequency, voltage, and power flow; a powerhouse to enclose the machinery and equipment; transformers and switching equipment; and a transmission line to deliver the power to the load center for ultimate distribution (**Fig. 6**).

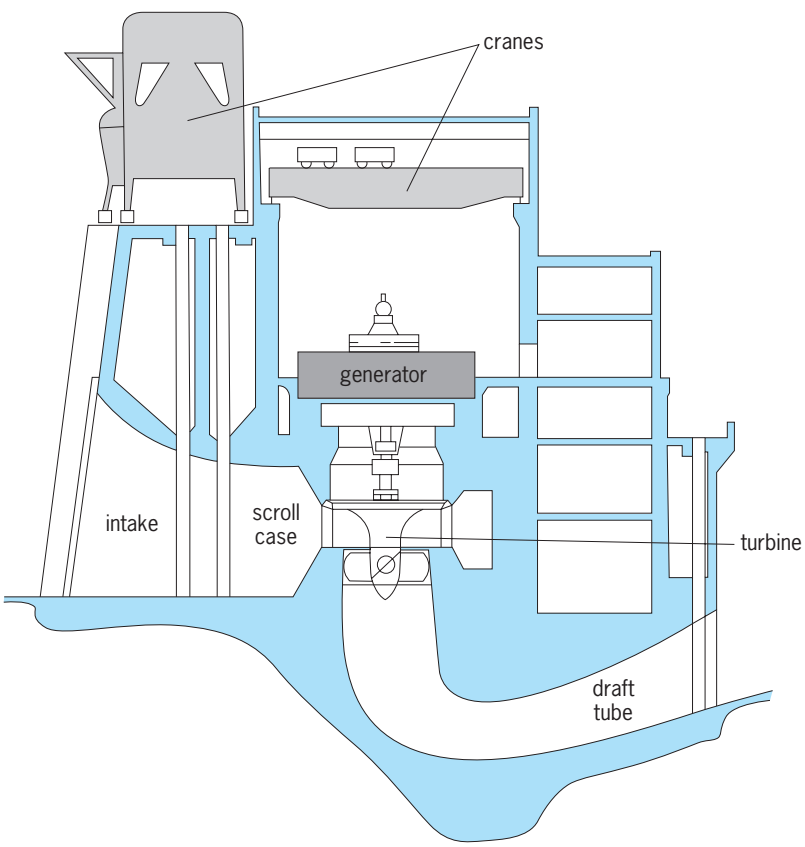

Fig. 6. Cross section of a hydroelectric powerhouse. Arrangement is typical for an adjustable propeller-type reaction turbine. The intake and powerhouse are integral; waterways consist of intake, scroll case, and draft tube. Turbine and generator are set with the axis vertical.

Dam. Dams are functionally important as they regulate the water supply to the plant and also are frequently the largest single element of cost in a hydroelectric development. Dams may be classified according to materials or type of construction, height, or use. *See* DAM.

Waterways. Waterways may be very short or as much as 10 mi (16 km) or more in length. The type of waterway is selected on the basis of length, pressure, plant arrangement, and economy. Very short waterways are required in installations where the powerhouse and the dam are integral structures; then the waterway is short and the length dictated by the plant arrangement. In other low-head plants, canals and forebays are frequently used. Conversely, very long waterways are also common in the developments where the head is great. Tunnels and penstocks are commonly used in high-head developments to deliver the water to a powerhouse some distance from the impoundment.

Some of the important aspects of the hydraulic design of waterways are provision for dewatering to facilitate inspection and maintenance, provision for water hammer resulting from surges caused by sudden start-up or shutdown, and optimization of hydraulic losses for an economic design. *See* WATER HAMMER.

An intake structure may be provided at the entrance to the waterway to reduce the hydraulic losses and to provide a closure gate for dewatering the waterway. Surge chambers are required at abrupt vertical changes in the direction of the waterways to provide relief of surges on shutdown and to prevent separation of the water column on start-up. When tunnels are used for the tailrace, surge chambers may be required for surge suppression.

Penstocks are defined as large-diameter pipelines for delivering water to the turbine or the steel-lined part of a tunnel adjacent to the powerhouse. Materials other than steel have been used in penstock construction; in the past, wood staves were common; and fiber glass is now used selectively.

Turbines. Hydraulic turbines are used to drive an electric generator, thereby converting the potential energy of the water to electrical energy. Reaction- or impulse-type turbines are used depending upon the head available, with a reaction turbine being used for heads under 1200 ft (365 m) and an impulse turbine for greater heads. Reaction turbines operate in a closed system from headwater to tailwater. For impulse turbines, only the portion to the turbine is closed and under pressure; from that point, flow is by gravity. As water passes through the reaction turbine, conversion from potential to mechanical energy takes place and, through a shaft,

drives the generator. For an impulse turbine, the potential energy of the water is first converted to kinetic energy in high-velocity jets which impinge on buckets at the circumference of the impulse wheel. There the kinetic energy is converted to mechanical energy for driving the generator. Reaction turbines are classified as Francis or propeller types, and propeller types are further categorized as fixed-blade and adjustable-blade, or Kaplan-type. Both classes of reaction turbines are normally installed with the axis vertical. *See* HYDRAULIC TURBINE; IMPULSE TURBINE; REACTION TURBINE.

Generator. In vertical installations, the generator is located above the turbine. When the turbine rotates, so does the rotor of the synchronous generator which produces electrical energy in a conventional way. An operating speed is chosen, as well as a generator with the correct number of poles to develop the desired frequency. From the generator, electricity is carried by buses to switching apparatus and then to transformers, where the voltage is stepped up for transmission. *See* ALTERNATING-CURRENT GENERATOR; HYDROELECTRIC GENERATOR.

Bulb units. Bulb units, which take their name from their shape, are an adaptation of the propeller unit and differ from the conventional unit in two significant respects: the alignment is horizontal, and the generator is totally enclosed within the bulb and mounted in the water passage. Water flows to the unit in a horizontal conduit and passes in the annular space between the bulb and the walls of the conduit. As the water passes the bulb, it drives the propeller and exits from the unit through a divergent horizontal water passage. Flow is controlled in a manner similar to the vertical units by wicket gates.

Classification of developments. Hydroelectric developments may be classified according to purpose, use, location of powerhouse, and head. The descriptive classifications commonly used are the run-of-river or storage, high- or low-head, base- or peak-load, multipurpose or single-purpose, and conventional or pumped-storage. Powerhouses may be indoor or outdoor, aboveground (surface) or underground. A number of these classifications may apply to any specific project such as conventional, run-of-river with surface, outdoor powerhouse, or pumped-storage with an underground powerhouse.

Multipurpose projects. A multipurpose project has a number of uses, some of which may be governed by competing criteria. Hydroelectric generation is generally compatible with other uses because it is nonconsumptive. Such projects normally require large reservoirs to accommodate the multiple uses. The most frequent purposes are irrigation, power, flood control, navigation, water supply, and recreation. Some adjustment must be made in either the amount of or the timing of the generation for compatibility with other uses. On major river systems in the United States, multipurpose projects have been developed in conjunction with navigation, flood control, and power. In the semiarid states, irrigation has played a dominant role in multipurpose developments.

Pumped storage. Pumped storage is a process for converting large quantities of electrical energy to potential energy by pumping water to a higher elevation where it can be stored indefinitely, then released to pass through hydraulic turbines and generate electrical energy on demand. Storage is desirable, as the consumption of electricity is highly variable according to the time of day or week, as well as seasonally. Consequently, there is excess generating capacity at night and on weekends. This excess capacity can be used to generate energy for pumping, hence storage. Normally, pumping energy can be obtained cheaply at night or on weekends, and its value will be upgraded when used for daytime peak loads.

In a typical operation, water is pumped at night or on weekends from a lower to an upper reservoir, where it is stored. The water can be retained indefinitely without deterioration or significant loss. During the daylight hours when the loads are greatest, stored water is released to flow from the upper to the lower reservoir through hydraulic turbines which generate electricity. No water is consumed in either the pumping or the generating cycles. A hydroelectric pumped-storage development is similar to a typical hydro installation, except that it has two reservoirs of essentially equal size situated to maximize their difference in elevation. Also, since it is a closed system, it need not be located directly on a large stream but must have a source for initial filling and to make up for the losses of evaporation and seepage. The second principal difference is the pump-turbine which is capable of rotation in either direction and acts as a pump in one direction and a turbine in the other. *See* ENERGY STORAGE; WATERPOWER.

Dwight L. Glasscock

Advanced Power Sources

Interest has developed in alternative or advanced power sources. Most of this attention has focused on solar or solar-related technologies, although several additional options including geothermal and fusion have been examined. Genuine interest in alternatives developed following the oil embargo in 1973.

The supplies of fossil and nuclear fuels are classified as nonrenewable and therefore are gradually being depleted. The advanced power sources of interest are those which are renewable and linked to the Sun, naturally occurring steam or hot water in the Earth's crust, or fusion. Since all solar sources suffer from intermittent availability, they require special consideration. As these various options are considered, three significant problems must be addressed. First, availability and the primary form of occurrence of the resource must be assessed. Second, a conversion system must be developed to transform the source into a usable form of energy, typically electricity. Third, the electricity production source must be effectively and efficiently appended to the existing electrical supply system.

Solar power. The solar options consisting of photovoltaics, wind, solar thermal process, and biomass will be discussed below. Other solar alternatives,

including solar ponds, passive solar, and ocean-thermal energy conversion, and other ocean power systems such as tide and wave power will not be treated. *See* TIDAL POWER.

Photovoltaics. The collection of solar energy takes many forms, but one of its most desirable configurations is direct conversion to electricity. The device configurations are modular in form, making them convenient from a mass-production viewpoint.

The heart of the photovoltaic system is a thin flat layer of semiconductor material. When the material is struck by sunlight, electrons are freed, producing an electric current. The direct-current (dc) power is passed through a dc load, into a storage battery, or converted to alternating current (ac) for general use in electric utility grids. Typically, individual solar cells are ganged together to form photovoltaic modules about 5 ft^2 (0.5 m^2) in size with a generating capacity of about 50 W. Typically, about half the cost of a solar system lies with the solar cell modules, and the remainder is directed toward power conditioning, electrical wiring, installation, and site preparation.

Typically, silicon or gallium arsenide are used to fabricate solar cells, although other semiconductor materials are being developed. Silicon technology is the most advanced because it is the least expensive, and takes many different forms including single-crystal, polycrystal, and amorphous configurations. The efficiency of these configurations ranges from about 6 to about 14% in module form. Amorphous silicon offers significant potential because of low manufacturing cost.

One way to advance the power yield from photovoltaics is to use concentrator systems. Special lenses and mirrors are used to focus sunlight on the cells, thereby raising the output of a module. Concentrator systems result in larger electric current which increases cell losses but improves overall economic efficiency of the system. Concentrator technology can raise module efficiencies to about 20%. Photovoltaic systems have potential application at the household level; however, the best near-term application is in array configurations for utility application. *See* SOLAR CELL.

Wind. The use of wind energy dates back to sailing ships and windmills. Today the interest in wind is for electricity generation by wind turbines. Wind energy conversion systems consist of several major subsystems. Among the most important is the mechanical system and electromechanical rotary converter. A wind energy conversion system is designed to rotate at either constant or variable speed as the wind varies. The variable-speed system usually offers high wind-collection efficiency; however, constant-speed units permit simpler electrical systems.

There are a multitude of aeroturbine designs available, with each striving to enhance the wind to mechanical power conversion effectiveness. A measure of this effectiveness is the power coefficient (C_p). The power coefficient is expressed as a function of the tip speed ratio, that is, the ratio of the blade tip speed to the wind speed. In variable-speed systems, C_p is tuned to its optimum value, whereas constant-speed systems cannot operate at optimum C_p over the entire wind speed range. This does not suggest that variable speed is a panacea, since it exposes the mechanical drive train and tower to potential resonance problems and requires more power conditioning of the electrical output.

The characteristics of the aeroturbine and the electric generator are not totally compatible; therefore, an element is required to effectively tie these two components. The mechanical interface has two major assignments. First, aeroturbines usually rotate at low speeds, which are not desirable for electrical production. Typically, a transmission is used to convert low-speed, high-torque mechanical power into high-speed, low-torque mechanical power for electrical conversion. Second, the mechanical interface can be used to regulate the drive train shaft stiffness, and thus shape its dynamic performance.

Ultimately the mechanical energy that has been produced by the aeroturbine and conditioned by the mechanical interface must be converted to an electrical form. In general, the type of wind system, constant- or variable-speed, determines the type of electric generator—synchronous, induction, or dc. Variable-speed systems tend to produce initially poor-quality ac or dc electric power, and then use power conditioners in the electrical interface to enhance it. Usually, constant-speed systems produce high-quality electric power at the generator terminals.

Further development of wind turbine technology is required to provide economical systems; however, wind electric systems show promise. Systems under development will produce as much 4–5 MW per machine. Typically, units larger than 200 kW are used by utilities. *See* WIND POWER.

Solar thermal. The process of collecting and concentrating solar energy at a focal point is known as solar thermal. There are two broad classes of solar thermal collection systems: power tower and point- and line-focus. The power tower is a large central power-generating system which shows significant promise, while the point- and line-focus systems are smaller and would be deployed as so-called dispersed sources. In both cases, solar energy is focused at a point or a line and heat is transferred to a working fluid.

The power tower concept has received significant attention because of its potentially large power production capability. A 10-MW demonstration plant, called Solar One, developed by the U.S. Department of Energy near Barstow, California, consists of 1818 heliostats, each containing 12 separate mirror facets designed to aim their energy at a central receiver on top of a structural steel tower. The heliostats are computer-controlled and adjust during the day to track the Sun so that optimal collection efficiency is maintained. The working fluid in this system is water, which is converted to steam and then passed through

a conventional steam turbine for eventual conversion to electricity. One significant application for the solar thermal system is as a repowering plant. Repowering is a term applied to the concept of adding a solar thermal plant to a conventional power plant to supplement steam production.

The point- and line-focus systems are smaller facilities which produce power at lower temperatures (200–600°F or 100–300°C). A typical point system is a parabolic dish which concentrates sunlight on a receiver a few feet away. Line-focus or trough systems consist of a series of cylindrical or parabolic trough lines with mirrors to collect and concentrate the Sun's radiation. Both the point- and line-focus systems could be used in smaller cogeneration facilities for commercial and industry electric power systems. *See* SOLAR ENERGY.

Biomass. Biofuel applications, involving consumption of wood products and garbage, have existed for many years. Perhaps the single largest hurdle for this application of wood and other products is the creation of effective distribution and supply systems. In general, these materials are low-density, thus causing significant transportation problems in delivering them to the chosen site for consumption.

An alternative to direct burning of biofuels is conversion to combustible liquids, gases, and solids. One process is classified as thermochemical—pyrolysis and gasification. Pyrolysis is a low-temperature process for forming gases, liquids, and solids by heating them in the absence of air, while gasification is a high-temperature process for producing gaseous fuels in the presence of an oxidant (air). A second class of processes for converting biomass to an alternative form is biochemical, which is better known as fermentation, that is, microbial transformation of organic feed materials that takes place without oxygen and produces alcohols and organic chemicals, such as methane, ethanol, acetic acid, and acetone. *See* FERMENTATION; PYROLYSIS.

The typical process of converting solar radiation to electricity by using vegetation as the collector is very inefficient (**Table 3**). It is clear that direct conversion by means other than vegetation is desirable. *See* BIOMASS.

Geothermal power. The natural emissions of steam (geysers), hot springs, and volcanoes represent potential sources of electricity production and are indicative of the steam and hot water potential embedded in the Earth's crust. Geothermal energy manifests itself in three basic types: hydrothermal, geopressured, and petrothermal. Hydrothermal is the natural production of steam when water is vaporized by coming into contact with hot rock. In geopressured systems, hot water is generated in deep reservoirs embedded in sand and shale formations within the Earth. The hot water is trapped and pressurized and is saturated with natural gas. Petrothermal systems consist of hot rock at the Earth's surface which could have a fluid injected into it and pumped out again, to extract theral energy. The amount of geothermal potential for electricity production is estimated at 220×10^{15} Btu (2.3×10^{20} joules), which is about three times the United States' energy consumption in 1980.

Presently, most geothermal power-producing facilities are limited to application of geysers or so-called dry steam sources. Since dry steam represents less than 1% of the total geothermal potential, other systems are under development to harness the Earth's natural heat. The largest near-term potential appears to be the hydrothermal resource. The present approach of converting high-temperature (above 410°F or 210°C) water is direct-flash technology. Hot water is extracted from the earth and its pressure is dropped, causing the water to vaporize (flash-boil). The steam is applied to conventional steam turbine technology. The overall efficiency of this process requires improvement since significant energy is lost when flashing the hydrothermal fluid. Since the hot water arrives from the geothermal well as a two-phase mixture of steam and water, a centrifugal process is under development to separate steam and water, capturing natural steam and optimizing conversion of the hot water to steam.

Since much of the geothermal potential occurs at lower temperatures, below 410°F (210°C), a process is being developed to capture this energy. The process is called binary since hot geothermal energy is used to vaporize a secondary fluid such as a hydrocarbon like isobutane or isopentain which boils at a lower temperature. The heat from the geothermal source is transferred to the secondary fluid through a heat exchanger, which in turn produces steam for a conventional steam cycle.

Although the various geothermal sources follow different paths to the production of steam, several common problems exist. Most hydrothermal fluids contain dissolved minerals. As the temperature of the fluid drops, the minerals precipitate out causing scaling in most flow channels and thus reducing the efficiency and effectiveness of the system. The corrosive nature of the brine hydrothermal fluids is a general problem. Perhaps the other general problem is the noncondensable gases found dissolved in the hydrothermal fluids. These gases reduce thermal efficiency, accelerate equipment wear, and cause environmental concerns. *See* GEOTHERMAL POWER.

TABLE 3. Typical energy losses in using vegetation for generation of electricity

Energy losses	% of energy in incident solar radiation
Due to leaf reflection and ineffective absorption	9
Due to invisible light, unusable for photosynthesis	55
Due to photosynthetic conversion	31
Due to plant respiration and metabolism	2
Due to conversion to electricity	2
Available as electricity	1

Fusion power. The use of nuclear energy as a power source has been confined to nuclear fission or the splitting of atomic nuclei for energy production. In the process of fusion, atomic nuclei merge by overcoming their normal electrostatic repulsion. When the nuclei fuse, energy is released as a result of the freeing of fast-moving nuclear particles. The Sun is a natural fusion reactor and is indicative of the potential from this source of energy.

Atomic mixtures used in fusion, called plasmas, consist of positively and negatively charged particles resulting from heating a gas to a high temperature. Fusion requires that the working plasma have not only extreme temperature to free the nuclei but also a sufficiently dense or compact environment to merge them. As a gas heats up, it will expand unless confined, resulting in a very low-density mixture that will not allow nuclear fusion. A successful fusion effort must therefore combine both high temperature and confinement.

The Sun accomplishes the process of fusion with its own high temperature (15×10^{6}°C) and its enormous mass which holds the plasma by gravity at proper densities. On Earth, such gravitational systems are not possible and thus higher temperatures will be necessary. The development of fusion reactors has concentrated on the two problems of temperature and confinement. Heating appears a less formidable problem than confinement. Electric currents, radio-frequency waves, and neutral-beam heaters have been effective approaches to the heating problem. Containment systems have focused on two approaches: magnetic and inertial.

The leading contender in magnetic confinement is a doughnut-shaped device called a tokamak. A magnetic field is used to control the plasma by deflecting the charged particles into a cylindrical or magnetic bottle and preventing them from striking the walls of the containment vessel. Such an occurrence would drop the temperature of the plasma, preventing the fusion process. In inertial confinement, a pulsed-energy source, called a drive, is used to compact and heat the fusion fuel in a single step. This results in rapid burning of the fuel to yield energy release. *See* ELECTRIC POWER SYSTEMS; NUCLEAR FUSION.

Thomas W. Reddoch

Bibliography. P. N. Cheremisinoff, *Air Pollution Control and Design for Industry*, 1993; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; L. L. Freris, *Principles of Wind Energy Conversion Systems*, 1990; General Electric Co., *Updating Excitation Systems*, Publ. GET-6675B, 1986; K. W. Li and A. P. Priddy, *Power Plant System Design*, 1985; P. Longrigg and E. H. Buell, *Electric Power from Renewable Energy Technologies*, Solar Energy Research Institute, May 1983; A. J. Pansini and K. D. Smalling, *Guide to Electric Power Generation*, 1994; J. G. Singer, *Combustion-Fossil Power Systems*, 4th ed., 1993; S. C. Stultz and J. B. Kitto, *Steam: Its Generation and Use*, 40th ed., 1990; R. Van Overstraeten and R. P. Mertens, *Physics, Technology and Use of Photovoltaics*, 1986; C. L. Wadhwa, *Electrical Power Systems*, 2d ed., 1992; J. Weisman and L. E. Eckart, *Modern Power Plant Engineering*, 1985; J. H. Willenbrock and H. R. Thomas (eds.), *Planning, Engineering and Construction of Electric Power Generation Facilities*, 1980; A. J. Wood and B. J. Wollenberg, *Power Generation, Operation, and Control*, 2d ed., 1996; K. E. Zweibel, *Harnessing Solar Power: The Photovoltaics Challenge*, 1990.

Electric power measurement

The measurement of the time rate at which electrical energy is being transmitted or dissipated in an electrical system. The potential difference in volts between two points is equal to the energy per unit charge (in joules/coulomb) which is required to move electric charge between the points. Since the electric current measures the charge per unit time (in coulombs/second), the electric power p is given by the product of the current i and the voltage v (in joules/second = watts), as in Eq. (1).

$$p = vi \qquad \text{(watts)} \tag{1}$$

See ELECTRIC CURRENT; ELECTRICAL UNITS AND STANDARDS; POWER.

Alternate forms of the basic definition can be obtained by using Ohm's law, which states that the voltage across a pure resistance is proportional to the current through the element. This results in Eq. (2),

$$p = i^2R = \frac{v^2}{R} \qquad \text{(watts)} \tag{2}$$

where R is the resistance of the element and i and v are the current through and voltage across the resistive element. Other commonly used units for electric power are milliwatts (1 mW = 10^{-3} W), kilowatts (1 kW = 10^3 W), megawatts (1 MW = 10^6 W), and, in electromechanical systems, horsepower (1 hp = 746 W). *See* ELECTRICAL RESISTANCE; OHM'S LAW.

These fundamental expressions yield the instantaneous power as a function of time. In the dc case where v and i are each constant, the instantaneous power is also constant. In all other cases where v or i or both are time-varying, the instantaneous power is also time-varying. When the voltage and current are periodic with the same fundamental frequency, the instantaneous power is also periodic with twice the fundamental frequency. In this case a much more significant quantity is the average power defined by Eq. (3), where T is the period of the periodic wave.

$$P = \frac{1}{T}\int_t^{t+T} vi\,dt \qquad \text{(watts)} \tag{3}$$

The average power is usually the quantity of interest since in most cases the electric power is converted to some other form such as heat or mechanical power and the rapid fluctuations of the power are smoothed by the thermal or mechanical inertia of the output

system. The remainder of this article is concerned primarily with average power in systems with periodic voltages and currents.

Sinusoidal ac waves. A number of basic concepts in electric power measurement are directly related to the behavior of systems where the voltages and currents are simple sinusoids.

Average power. In a circuit containing only resistance, the voltage and current waves have the same points of zero crossing and are said to be in phase. The power curve, obtained by directly multiplying the instantaneous values of voltage and current, is a double-frequency sinusoidal wave which is entirely positive. It is clear from the symmetry of the power wave that the average power over one cycle is one-half of the peak power.

When a circuit contains both resistance and inductance, the current lags behind the voltage wave. The power curve has the same double-frequency sinusoidal shape but is shifted downward so that over a portion of the cycle the instantaneous power is negative. During this part of the cycle, power is actually being delivered back to the source from the magnetic field of the inductance. The average of the double-frequency power wave is less than one-half of the peak because of the portion which has negative values.

Quantitative relationships describing such results can be obtained by writing the voltage and current waves as Eqs. (4) and (5), where V_m and I_m are max-

$$v = V_m \sin 2\pi f t = \sqrt{2}\, V \sin 2\pi f t \quad \text{(volts)} \quad (4)$$

$$i = I_m \sin (2\pi f t - \phi) = \sqrt{2}\, I \sin (2\pi f t - \phi) \quad \text{(amperes)} \quad (5)$$

imum values, V and I are root-mean-square (rms) or effective values, f is the frequency, and ϕ is the angle by which i lags v. The instantaneous power is then given by Eq. (6), which can be rewritten as Eq. (7), from which the average power is given by Eq. (8).

$$p = vi = 2\, VI \sin 2\pi f t \, \sin (2\pi f t - \phi) \quad \text{(watts)} \quad (6)$$

$$P = VI[\cos\phi - \cos(4\pi f t - \phi)] \quad \text{(watts)} \quad (7)$$

$$P = VI \cos\phi \quad \text{(watts)} \quad (8)$$

See ALTERNATING CURRENT; ALTERNATING-CURRENT CIRCUIT THEORY.

Power factor, apparent power, and reactive power. The average power P in Eq. (8) is also called the real or active power and is distinguished from the simple product VI by the factor $\cos\phi$. The product VI is called the apparent power (which is measured in volt-amperes), and the factor $\cos\phi$ is called the power factor so that, by definition Eq. (9*a*) holds or, in words, Eq. (9*b*).

$$\cos\phi = \frac{P}{VI} \quad (9a)$$

$$\text{Power factor} = \frac{\text{average power}}{\text{apparent power}} \quad (9b)$$

One additional quantity is introduced by rewriting Eq. (7) in the alternative form of Eq. (10), which indicates that the single double-frequency sinusoidal

$$p = VI[\cos\phi\,(1 - \cos 4\pi f t) + \sin\phi \sin 4\pi f t] \quad \text{(watts)} \quad (10)$$

instantaneous power wave of a circuit containing both resistance and inductance can be broken into two double-frequency sinusoidal waves. One of the waves is entirely above the zero line, and the average of this wave is the average power given by Eq. (8). The other wave is symmetric above and below the zero line and represents the instantaneous power flow into and out of the magnetic field of the inductor. This portion of the instantaneous power is termed the reactive power. The peak amplitude is given by Eq. (11), and is measured in volt-amperes-

$$Q = VI \sin\phi \quad \text{(vars)} \quad (11)$$

reactive or vars. The quantity Q is nominally taken as positive for an inductive load. It is evident from Eqs. (8) and (11) that the apparent power, the average power, and the reactive power are related by Eq. (12). *See* VOLT-AMPERE.

$$V^2 I^2 = P^2 + Q^2 \quad \text{(volt-amperes)} \quad (12)$$

Nonsinusoidal periodic waves. For more complex periodic waves, the voltage and current can be expressed as Fourier series summations consisting of a dc component V_0, a fundamental component V_1, and a set of harmonics V_n as in Eq. (13), with the same

$$v = V_0 + \sum_{n=1}^{\infty} \sqrt{2}\, V_n \sin (2\pi f n t + \phi) \quad \text{(volts)} \quad (13)$$

general form for the current. It is easily shown by the methods used previously that only terms with the same frequency contribute to the power. All other products contribute only reactive power. The average power can be found by repeated application of Eq. (8) for each frequency present in both v and i. *See* FOURIER SERIES AND TRANSFORMS; HARMONIC (PERIODIC PHENOMENA); NONSINUSOIDAL WAVEFORM.

Low-frequency measurements. The measurement of power in a dc circuit can be carried out by simultaneous measurements of voltage and current by using standard types of dc voltmeters and ammeters. The product of the readings typically gives a sufficiently accurate measure of dc power. If great accuracy is required, corrections for the power used by the instruments should be made. In ac circuits the phase difference between the voltage and current precludes use of the voltmeter-ammeter method unless the load is known to be purely resistive. When this method is applicable, the instrument readings lead directly to average power since ac voltmeters and ammeters are always calibrated in rms values. *See* AMMETER.

Wattmeter measurements. In power-frequency circuits the most common instrument for power measurement is the moving-coil, dynamometer wattmeter. This instrument can measure dc or ac power by carrying out the required multiplication and averaging on a continuous analog basis. The instrument has four terminals, two for current and two for voltage, and reads the average power directly. It can be built with an accuracy better than 0.25% of full-scale deflection and with frequency response up to about 1 kHz. For harmonics within its frequency range, the reading is independent of the voltage and current waveforms. *See* WATTMETER.

Digital wattmeters. Electronic wattmeters are available which give a digital indication of average power. Their primary advantage, in addition to minimizing errors in reading the instrument, is that the frequency range can be greatly extended, up to 100 kHz or more, with good accuracy. In addition, such instruments also often provide other measurements such as rms voltage and current, apparent power, reactive power, and power factor. These instruments are of great value in power measurements involving very distorted waveforms as are common at the input of many electronic power supplies. Accurate measurement of power in these situations can be accomplished only with the wide bandwidth of these electronic instruments.

Polyphase power measurements. Measurement of the total power in a polyphase system is accomplished by combinations of single-phase wattmeters or by special polyphase wattmeters which are integrated combinations of single-phase wattmeter elements. A general theorem called Blondel's theorem asserts that the total power supplied to a load over N wires can be measured by using $N - 1$ wattmeters. The theorem states that the total power in an N-wire system can be measured by taking the sum of the readings of N wattmeters so arranged that each wire contains the current coil of one wattmeter. One voltage terminal of each wattmeter is connected to the same wire as its current coil, and the second voltage terminal is connected to a common point in the circuit. If this common point is one of the N wires, one wattmeter will read zero and can be omitted.

The most common polyphase system is the three-phase system, and the two most common connections are the three-phase, three-wire system, in which the sources and loads may be Y-connected or delta (Δ)-connected, and the three-phase, four-wire system, which utilizes Y-connected sources and loads. The fourth wire is the neutral wire, which connects the neutral point of the source to the neutral point of the load.

In one arrangement for measuring the power in a four-wire system, each three-phase wattmeter directly measures the power in one of the load elements. The sum of the three readings is the total power. However, this connection will read the correct total power even if the load is more complex, including, for example, one or more Δ-connected loads in parallel with the load illustrated in the figure. With multiple loads, only the sum of the meter readings is significant. For three-wire systems, the most common connection is the two-wattmeter arrangement. With this arrangement, again only the sum of the two readings is of significance. Care must be exercised in making the connection (the positive and negative terminals must be connected as shown) since one of the wattmeters will have a negative reading when the power factor is low or the load unbalanced. When this occurs, the negative-reading wattmeter must have its current or voltage coil reversed to obtain the value of the negative meter reading. Thus, careful attention to the positive and negative terminal connections is a necessity in this method.

High-frequency measurements. At frequencies significantly above power frequencies, dynamometer wattmeters become inaccurate and cannot be used. The newer digital wattmeters have usable ranges well above audio frequencies and make accurate audio-frequency power measurements quite feasible. Generally, however, power measurements at higher frequencies are based on indirect methods which do not attempt to carry out the multiplication and integration required for direct electric power measurement.

Voltmeter-based methods. For frequencies up to a few hundred megahertz, the voltage across a standard resistance load can be measured and the power calculated from V^2/R. Instruments which combine the resistive load and voltmeter are called absorption power meters. For the lower frequency end of the range, such instruments often contain an input transformer with taps to provide a range of input resistance values. Adjusting the transformer to achieve a maximum meter reading allows matching the effective load resistance to the source output impedance to maximize the power output from the source.

Diode detector–based methods. A diode may be used as a detector for radio-frequency (rf) power measurement. This type of instrument is simple and easy to use but is less accurate than the thermally based systems described below. The diode is operated on the square-law portion of its characteristic curve, and as a result the voltage on the capacitor C is proportional to the average rf power input. There are a number of limitations to the performance of these instruments. The dc signal at the output is very small, requiring amplification to enable accurate measurement, and care must be taken to avoid errors caused by electronic and thermoelectric effects. Temperature changes affect the diode resistance, and hence the sensitivity and reflection coefficient can vary; the matching resistor, usually 50 ohms, is used to terminate the input source to reduce these temperature effects. *See* DIODE.

Thermocouple-based methods. A thermocouple consists of two dissimilar metals joined at one end. When the joined end is heated and the other end is at a lower temperature, an electric current is produced (the

thermoelectric or Seebeck effect). The current is proportional to the temperature difference between the two ends. For electric power measurements, the hot end is heated by a resistor supplied from the rf power source to be measured. Modern devices use thin-film techniques to make the thermocouple and resistor to ensure good thermal coupling. The output is a low-level dc signal as in diode detector systems. Since the thermocouple method is affected by drift in the thermocouple and associated electronic circuits, most instruments contain a source arranged to be connected in place of the rf input to calibrate the instrument. Thermocouple instruments generally have a wider range of measurement and are more rugged than most other high-frequency power measurement instruments. *See* THERMOCOUPLE.

Bolometric methods. A bolometer is basically an electric bridge circuit in which one of the bridge arms contains a temperature-sensitive resistor. In principle, the temperature is detected by the bridge circuit. Bolometric bridges use either barreters or thermistors as the temperature-sensitive resistor. A barreter is simply a short, thin wire, usually of platinum, whose resistance changes significantly for small amounts of power dissipation. Since the element must be operated at relatively high temperatures, near its burn-out level, it is easily destroyed by accidental overloads. The thermistor, which has largely replaced the barreter since it is more rugged, is a semiconductor device with a negative temperature coefficient. For rf power measurements, the thermistor is fabricated as a small bead with very short lead wires so that essentially all the resistance is in the bead. *See* BARRETTER; BOLOMETER; THERMISTOR.

In power measurements, absolute measurement of the resistance variation is not generally acceptable since the change in resistance can affect the reflection coefficient of the probe. Thus, to expand the useful power range of the instrument, comparison methods are employed. For example, the bridge can be arranged so that, without any input power, a dc or low-frequency bias is applied to the sensing element and the bridge is balanced. The input power is then applied causing an imbalance. The bias is then changed to return the bridge to balance, and the change in the bias power is used to indicate the input power. A major advantage of thermistor-based bolometric techniques is that they operate at high signal levels and shielding requirements are minimized.

Calorimeter methods. The most accurate high-frequency power measurement methods involve calorimetric techniques based on direct determination of the heat produced by the input power. The power to be measured is applied to the calorimeter, and the final equilibrium temperature rise is recorded. Then the input signal is removed, and dc power applied until the same equilibrium temperature is attained. The dc power is then the same as the signal power. Most power measurement calorimeters use a flow-through coolant rather than a static fluid coolant, and can be built to measure large amounts of power. Calorimeter methods are usually used in standards laboratories rather than in industrial applications. *See* CALORIMETRY.

Sensors and connections. Ideally the power sensor should accept all of the high-frequency power to be measured. There are, however a number of factors that reduce the calibration factor of the sensor below the ideal of 100%. For example, mismatch between the impedance of the rf or microwave source and that of the sensor will result in some reflection of the incident power. Some power may also be lost in the conducting walls of the sensor, and some may be radiated into space. Typically, the calibration factor of a commercial power sensor at various frequencies within its range is supplied by the manufacturer.

The accuracy of high-frequency power measurements is very dependent on correct setup of the instrumentation to minimize losses. Most often, directional couplers and attenuators are used to supply the power meter with the desired signal at the appropriate power level for the instrument. Care must be exercised to assure that all portions of the test circuit are properly matched to avoid reflections and that losses are minimized. *See* DIRECTIONAL COUPLER; ELECTRICAL MEASUREMENTS; IMPEDANCE MATCHING; TRANSMISSION LINES.

Donald W. Novotny

Bibliography. M. Braccio, *Basic Electrical and Electronic Tests and Measurements*, 1978; *Code for Electrical Metering*, ANSI C12, 1982; D. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; Instrument Society of America, *ISA Standards and Practices for Instrumentation*, 9th ed., 1988; P. Kantrowitz, G. Kousouri, and L. Zucker, *Electronic Measurements*, 1979; F. F. Mazda, *Electronic Instruments and Measurement Techniques*, 1987.

Electric power substation

An assembly of equipment in an electric power system through which electrical energy is passed for transmission, distribution, interconnection, transformation, conversion, or switching. *See* ELECTRIC POWER SYSTEMS.

Specifically, substations are used for some or all of the following purposes: connection of generators, transmission or distribution lines, and loads to each other; transformation of power from one voltage level to another; interconnection of alternate sources of power; switching for alternate connections and isolation of failed or overloaded lines and equipment; controlling system voltage and power flow; reactive power compensation; suppression of overvoltage; and detection of faults, monitoring, recording of information, power measurements, and remote communications. Minor distribution or transmission equipment installation is not referred to as a substation.

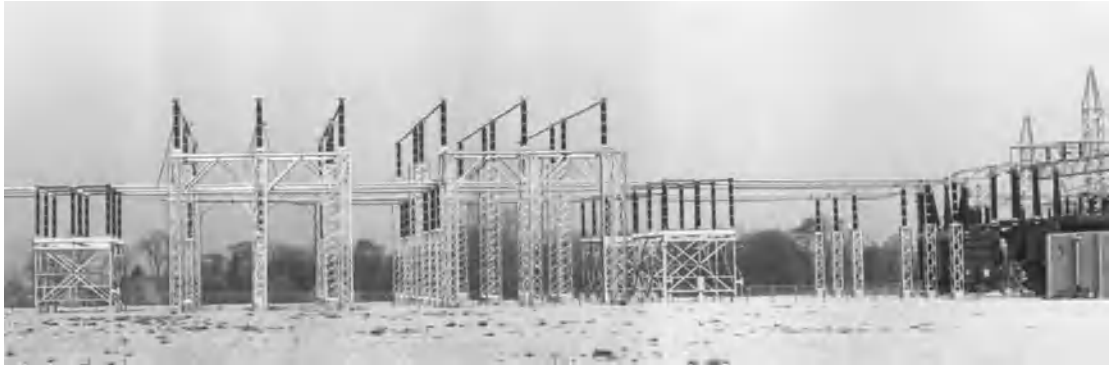

Fig. 1. A typical transmission substation (345 kV).

Classification. Substations are referred to by the main duty they perform. Broadly speaking, they are classified as transmission substations (**Fig. 1**), which are associated with high voltage levels; and distribution substations (**Fig. 2**), associated with low voltage levels. *See* ELECTRIC DISTRIBUTION SYSTEMS.

Substations are also referred to in a variety of other ways, which may be described as follows.

Transformer substations. These are substations whose equipment includes transformers.

Switching substations. These are substations whose equipment is mainly for various connections and interconnections, and does not include transformers.

Customer substations. These are usually distribution substations located on the premises of a larger customer, such as a shopping center, large office or commercial building, or industrial plant.

Converter stations. The main function of converter stations is the conversion of power from ac to dc and vice versa. Converter stations are complex substations required for high-voltage direct-current (HVDC) transmission or interconnection of two ac systems which, for a variety of reasons, cannot be connected by an ac connection. The main equipment includes converter valves usually located inside a large hall, transformers, filters, reactors, and capacitors. **Figure 3** shows a ± 250-kV 500-MW HVDC converter station together with a 245-kV ac substation. The building in the center of the converter station contains all the thyristor valves. On each side of the building against the wall are converter transformers connected to the valves inside. Further away from the building are filters and capacitors. Going away from the building is the dc line, with a typical dc line tower in the uppermost part of the figure. *See* DIRECT-CURRENT TRANSMISSION.

Air-insulated substations. Most substations are installed as air-insulated substations, implying that the busbars and equipment terminations are generally open to the air, and utilize insulation properties of ambient air for insulation to ground. Modern substations in urban areas are esthetically designed with low profiles and often within walls, or even indoors.

Metal-clad substations. These are also air-insulated, but for low voltage levels; they are housed in metal cabinets and may be indoors or outdoors (Fig. 2).

Fig. 2. View of a residential outdoor distribution substation with double-ended underground supply to transformers which are positioned at either end of the metal-clad switchgear.

Fig. 3. View of ±250-kV 500-MW HVDC converter station (upper part) and 245-kV ac substation (lower part). The large building of the converter station contains all the thyristor valves. (*General Electric Co.*)

Gas-insulated substations. Acquiring a substation site in an urban area is very difficult. Land in many cases is either unavailable or very expensive. Therefore, there has been a definite trend toward increasing use of gas-insulated substations, which occupy only 5–20% of the space occupied by the air-insulated substations. In gas-insulated substations, all live equipment and bus-bars are housed in grounded metal enclosures, which are sealed and filled with sulfur hexafluoride (SF_6) gas, which has excellent insulation properties. These substations have the appearance of a large-scaleplumbing (**Fig. 4**).

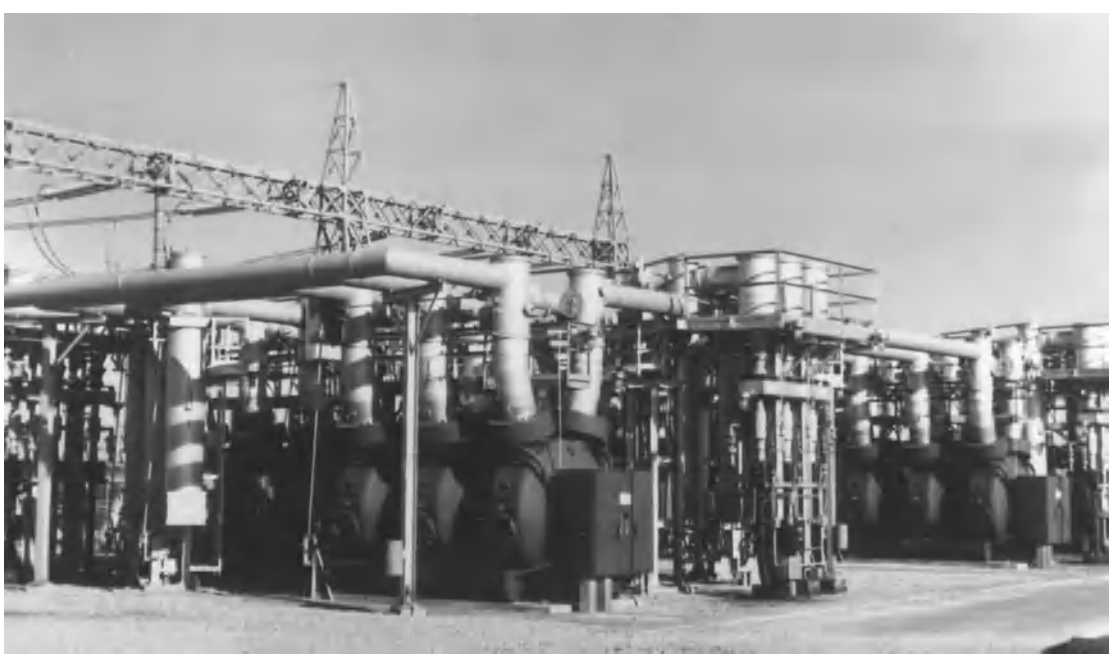

Fig. 4. A 345-kV ac gas-insulated substation. (*Gould-BBC*).

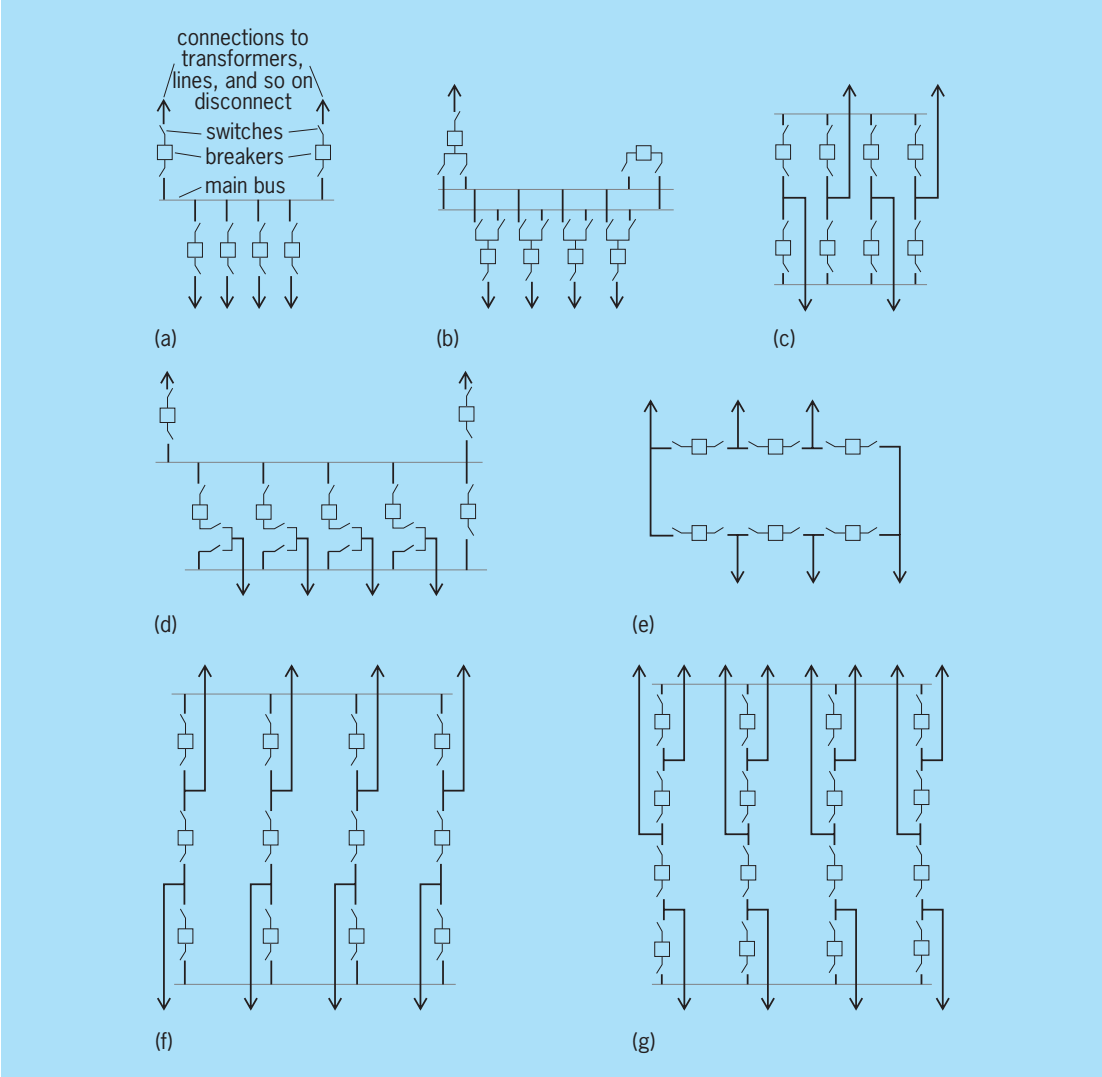

Fig. 5. One-line diagrams of substation switching arrangements. (*a*) Single bus. (*b*) Double bus, single breaker. (*c*) Double bus, double breaker. (*d*) Main and transfer bus. (*e*) Ring bus. (*f*) Breaker-and-a-half. (*g*) Breaker-and-a-third.

The disk-shaped enclosures in Fig. 4 are circuit breakers.

Mobile substations. For emergency replacement or maintenance of substation transformers, mobile substations are used by some utilities. They may range in size from 100 to 25,000 kVA, and from 2400 to 220,000 V. Most units are designed for travel on public roads, but larger substations are built for travel on railroad tracks. These substations are generally complete with a transformer, circuit breaker, disconnect switches, lightning arresters, and protective relaying.

Substation switching arrangement. An appropriate switching arrangement for "connections" of generators, transformers, lines, and other major equipment is basic to any substation design. There are seven switching arrangements commonly used (**Fig. 5**). Each breaker is usually accompanied by two disconnect switches, one on each side, for maintenance purposes. Selecting the switching arrangement involves considerations of cost, reliability, maintenance, and flexibility for expansion.

Single bus. This involves one common bus for all connections and one breaker per connection. This is the least costly arrangement, but also the least desirable for considerations of reliability and maintenance.

Double bus, single breaker. This involves two common buses, and each connection of the line equipment and so forth can be made to either bus through one or the other breaker. However, there is only one breaker per connection, and when a breaker is out for maintenance, the connected equipment or line is also removed from operation.

Double bus, double breaker. This arrangement has two common buses and two breakers per connection. It offers the most reliability, ease of maintenance, and flexibility, but is also the most expensive.

Main and transfer bus. This is like the single-bus arrangement; however, an additional transfer bus is provided so that a breaker can be taken out of service by transferring its connection to the transfer bus which is connected to the main bus through a breaker between the two buses.

Ring bus. This may consist of four, six, or more breakers connected in a closed loop, with the same number of connection points. A breaker can be taken out of service without also taking out a connection.

Breaker-and-a-half. This arrangement involves two buses, between which three breaker bays are installed. Each three-breaker bay provides two circuit connection points—thus the name breaker-and-a-half.

Breaker-and-a-third. In this arrangement, there are four breakers and three connections per bay.

Substation equipment. A substation includes a variety of equipment. The principal items are listed and briefly described below.

Transformers. These involve magnetic core and windings to transfer power from one side to the other at different voltages. Substation transformers range from small sizes of 1 MVA to large sizes of 2000 MVA. Most of the transformers and all those above a few MVA size are insulated and cooled by oil, and adequate precautions have to be taken for fire hazard. These precautions include adequate distances from other equipment, fire walls, fire extinguishing means, and pits and drains for containing leaked oil. *See* TRANSFORMER.

Circuit breakers. These are required for circuit interruption with the capability of interrupting the highest fault currents, usually 20–50 times the normal current, and withstanding high voltage surges that appear after interruption. Switches with only normal load-interrupting capability are referred to as load break switches. *See* CIRCUIT BREAKER.

Disconnect switches. These have isolation and connection capability without current interruption capability. *See* ELECTRIC SWITCH.

Bus-bars. These are connecting bars or conductors between equipment. Flexible conductor buses are stretched from insulator to insulator, whereas more common solid buses (hollow aluminum alloy tubes) are installed on insulators in air or in gas-enclosed cylindrical pipes. *See* BUS-BAR.

Shunt reactors. These are often required for compensation of the line capacitance where long lines are involved.

Shunt capacitors. These are required for compensation of inductive components of the load current.

Current and potential transformers. These are for measuring currents and voltages and providing proportionately low-level currents and voltages at ground potential for control and protection.

Control and protection. This includes (a) a variety of protective relays which can rapidly detect faults anywhere in the substation equipment and lines, determine what part of the system is faulty, and give appropriate commands for opening of circuit breakers; (b) control equipment for voltage and current control and proper selection of the system configuration; (c) fault-recording equipment; (d) metering equipment; (e) communication equipment; and (f) auxiliary power supplies. *See* ELECTRIC PROTECTIVE DEVICES; RELAY; VOLTAGE REGULATOR.

Many of the control and protection devices are solid-state electronic types, and utilize digital techniques with microprocessors. Most of the substations are fully automated locally with a provision for manual override. The minimum manual interface required, along with essential information on status, is transferred via communications channels to the dispatcher in the central office.

Other items which may be installed in a substation include: phase shifters, current-limiting reactors, dynamic brakes, wave traps, series capacitors, controlled reactive compensation, fuses, ac to dc or dc to ac converters, filters, and cooling facilities.

Substation grounding and shielding. Good substation grounding is very important for effective relaying and insulation of equipment; however, the safety of the personnel is the governing criterion in the design of substation grounding. It usually consists of a bare wire grid, laid in the ground; and all equipment grounding points, tanks, support structures, fences, shielding wires and poles, and so forth, are securely connected to it. The grounding resistance is reduced to be low enough that a fault from high voltage to ground does not create such high potential gradients on the ground, and from the structures to ground, to present a safety hazard. Good overhead shielding is also essential for outdoor substations, so as to virtually eliminate the possibility of lightning directly striking the equipment. Shielding is provided by overhead ground wires stretched across the substation or tall grounded poles. *See* GROUNDING; LIGHTNING AND SURGE PROTECTION. Narain G. Hingorani

Bibliography. J. Arrilaga, *High Voltage Direct Current Transmission*, 2d ed., 1998; W. M. Flanagan, *Handbook of Transformer Design and Applications*, 2d ed., 1993; C. H. Flurscheim (ed.), *Power Circuit Breaker Theory and Design*, rev. ed., 1982; A. C. Franklin and D. P. Franklin, *The J&P Transformer Book,* 11th ed., 1983.

Electric power systems

Complex assemblages of equipment and circuits for generating, transmitting, and distributing electrical energy. The elements of a power system (**Fig. 1**) form complex networks which require energy control centers to monitor and regulate their operation. Various sources of primary energy, such as coal, nuclear fission, hydro power, geothermal, and wind, can be used to drive the electric generator in Fig. 1.

Since electrical energy plays a central role in industrialized societies, the reliability of electric power systems is a critical factor in their planning, design, and operation, and considerable effort is directed toward quantifying performance. In addition to the provision of electrical energy, attention is given to the myriad of items that consume this energy, that is, the power system load.

Principal Elements

Electrical systems require generating stations to produce electric power, transmission systems to carry

it to areas of consumption, substations to transform it for use in industrial and residential areas, and distribution systems to carry it to customers (Fig. 1). Coordinated interconnections between power systems play a critical role in ensuring an acceptable degree of reliability.

Generation. Electricity in the large quantities required to supply electric power systems is produced in generating stations, commonly called power plants. Such generating stations, however, should be considered as conversion facilities in which the chemical or nuclear energy of fuel (coal, oil, gas, or uranium); the kinetic energy of falling water, wind, or sea waves; or light from the Sun is converted to electricity. *See* ELECTRIC POWER GENERATION; POWER PLANT.

Steam stations. Most of the electric power used in the United States is obtained from generators driven by steam turbines. Units of 650, 800, and 950 MW are commonplace for new fossil-fuel-fired stations, with 1300 MW the largest in service. The most commonly installed units in nuclear stations are 800–1100 MW. *See* STEAM TURBINE.

Coal is an important fuel for more than half of steam turbine generation. Other fuels include natural gas and heavy fuel oil; the remainder is generated from the nuclear energy of slightly enriched uranium. As nuclear units have come into commercial operation, the contribution of uranium to the electrical energy supply has risen to about 15% of the total fuel generated. However, uranium's share should decrease, because no additional nuclear plants are slated for service and instead some existing nuclear power stations are slated for closure.

Combustion of coal produces sulfur dioxide and nitrogen oxides in the stack gases. To reduce these emissions below those permitted by the Environmental Protection Agency, modern coal-fired stations use either low-sulfur coal (that is, less than 0.3% sulfur by weight) or scrubbers that react the sulfur dioxide with a reagent to permit its removal. Nitrogen oxides are controlled by controlling combustion temperatures. Research is being conducted on clean coal technologies to minimize the production of carbon dioxide. *See* AIR POLLUTION.

Nuclear steam stations in the United States are mostly of the water-cooled and moderated types in which the heat of a controlled nuclear reaction is used to convert ordinary or "light" water into steam to drive a conventional turbine generator. *See* NUCLEAR POWER; NUCLEAR REACTOR.

Hydroelectric plants. Waterpower supplies about 10% of the electric power consumed in the United States, but this share will decline because very few sites remain undeveloped where sufficient water drops far enough in a reasonable distance to drive large hydraulic turbines. Much of the very small share of the planned additional hydroelectric capability will be used at existing plants to increase their effectiveness in supplying peak power demands, and as a quickly available source of emergency power. Special incentives from the federal government have made development of small hydroelectric plants of 15 MW or less financially attractive to private developers. *See* HYDRAULIC TURBINE; HYDROELECTRIC GENERATOR.

Some hydroelectric plants actually draw power from other generating facilities during light system-load periods, to drive their turbines in a reverse mode to pump water from a river or lake into an artificial reservoir at a higher elevation. From there, the water can be released through the hydraulic turbines when the power system needs additional generation. These pumped-storage installations consume about 35% more energy than they return to the power system and, accordingly, cannot be considered as primary energy sources. Their use is justified, however, by their ability to convert surplus power that is available during low-demand periods into prime power to serve system needs during peak-demand intervals, a need that otherwise would require building more generating stations for operation during the relatively few hours of high system demand. *See* ENERGY STORAGE; WATERPOWER.

Combustion turbine plants. Gas-turbine–driven generators, now commonly called combustion turbines because of the use of light oil as fuel, have gained wide acceptance as an economical source of additional power for heavy-load periods. In addition, they offer the fastest erection time and the lowest investment cost per kilowatt of installed capability. Offsetting these advantages, however, is their relatively less efficient consumption of more costly fuel. Combustion turbine units, even in the largest rating (100 MW), offer extremely flexible operation and can be started and run up to full load in as little as 10 minutes. Thus they are useful as emergency power sources, as well as for operating during the few hours of daily load peaks.

Combustion turbines have an additional role besides applying peaking or emergency power. Some installations use their exhaust gases to heat boilers that generate steam to drive steam turbine generators. Such combined-cycle units offer fuel economy comparable to that of modern steam plants and at considerably less cost per kilowatt. In addition, because only part of the plant uses steam, the requirement for cooling water is considerably reduced. However, wide acceptance is inhibited by the government's restrictions on light fuel oil for them. This barrier will be resolved by the development of systems for fueling combustion-turbine installations with gas derived from coal, natural gas, or by direct firing with pulverized coal. *See* COAL GASIFICATION; COGENERATION; GAS TURBINE.

Internal combustion plants. Internal combustion engines of the diesel type drive generators in many small power plants. In addition, they offer the ability to start quickly for operation during peak loads or emergencies. However, their small size, commonly about 2 MW per unit (although a few approach 10 MW), has limited their use. Such installations account for less than 1% of the total power-system generating capability in the United States, and make an

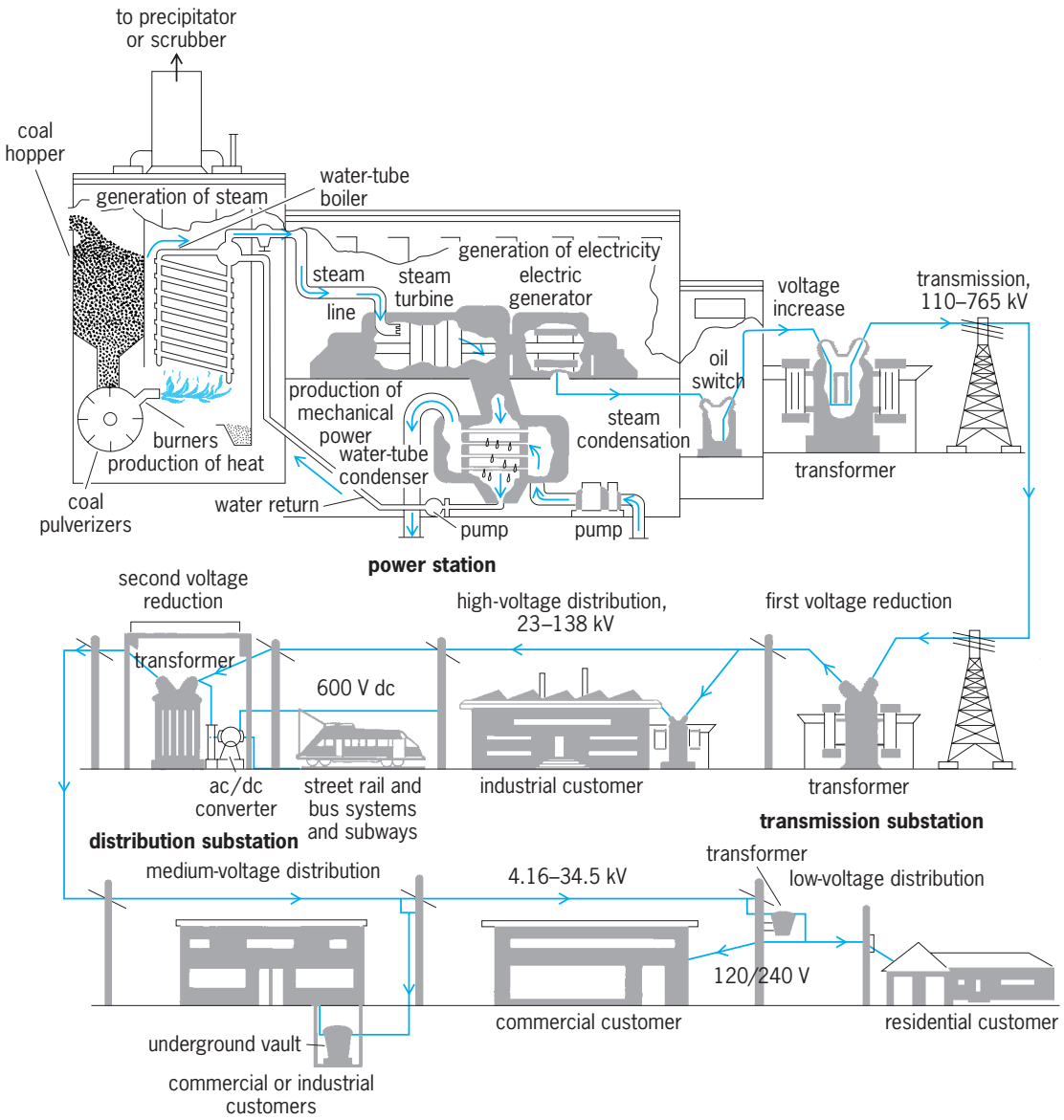

Fig. 1. Major steps in the thermal generation, transmission, and distribution of electricity.

even smaller contribution to total electric energy consumed. At present they are used primarily as standby units to provide electricity in case of emergencies and loss of electric grid. *See* DIESEL ENGINE; INTERNAL COMBUSTION ENGINE.

Renewable resources. Utilities actively seek to develop generating resources that do not consume fuel, including geothermal steam, wind, Sun, ocean waves, tides, and biomass-powered capacity. Of these, wind power is now the fastest growing segment of the power generation industry. Wind energy technology worldwide has grown 20% annually since 2000, and it is projected that by 2020 wind energy will meet 6% of the electric needs of the United States, matching the share of hydro power. *See* BIOMASS; ENERGY SOURCES; GEOTHERMAL POWER; SOLAR ENERGY; WIND POWER.

Three-phase output. Because of their simplicity and efficient use of conductors, three-phase 60-Hz alternating-current systems are used almost exclusively in the United States. Consequently, power-system generators are wound for three-phase output at a voltage usually limited by design features to a range from about 11 kV for small units to 30 kV for large ones. The output of modern generating stations is usually stepped up by transformers to the voltage level of transmission circuits used to deliver power to distant load areas. *See* TRANSFORMER.

Transmission. The transmission system carries electric power efficiently and in large amounts from generating stations to consumption areas. Such transmission is also used to interconnect adjacent power systems for mutual assistance in case of emergency and to gain for the interconnected power systems the economies possible in regional operation. Interconnections have so expanded that most of the generation east of the Rocky Mountains regularly operates in parallel and in synchronism. More than 90% of all

Power capability of typical three-phase open-wire transmission lines

Line-to-line voltage, kV	Capability, MVA
115 ac	60
138 ac	90
230 ac	250
345 ac	600
500 ac	1200
765 ac	2500
800 dc*	1500

*Bipolar line with grounded neutral.

generation in the United States and Canada, exclusive of Alaska and Hawaii, is or can be linked.

Transmission circuits are designed to operate up to 765 kV, depending on the amount of power to be carried and the distance to be traveled. The permissible power loading of a circuit depends on many factors, such as the thermal limit of the conductors and their clearances to ground, the voltage drop between the sending and receiving end and the degree to which system service reliability depends on it, and how much the circuit is needed to hold various generating stations in synchronism. A widely accepted approximation to the voltage appropriate for a transmission circuit is that the permissible load-carrying ability varies as the square of the voltage (see **table**).

Alternating-current transmission voltages have increased many times since transmission began as a distinct function around 1890 (**Fig. 2**). Lines at 500 and 765 kV were introduced in the 1960s. The very high capacity of lines at 765 kV limits their use, but 2454 mi (3950 km) are in service in the United States.

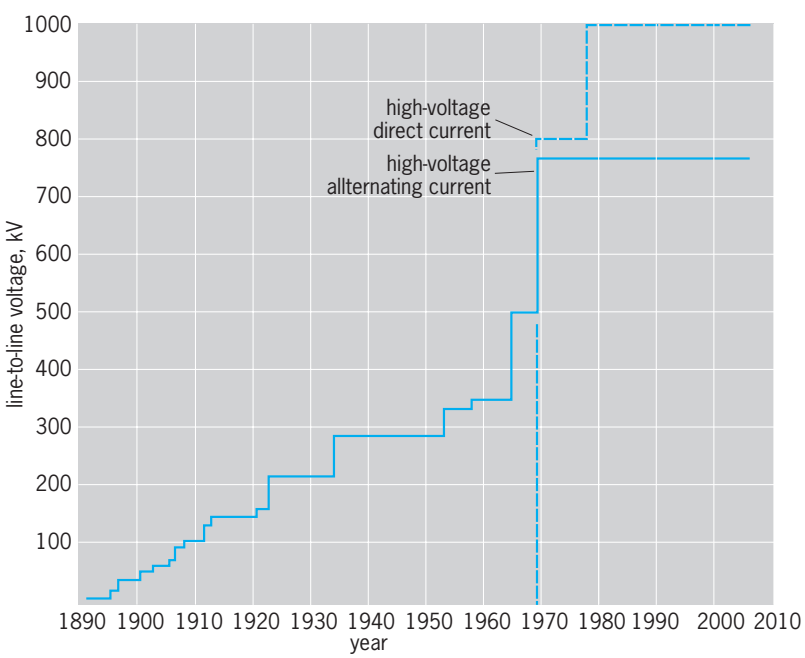

Fig. 2. Growth of alternating-current transmission voltages from 1890 onwards.

Transmission engineers have anticipated even higher voltages, of 1100 to 1500 kV, but they are fully aware that this objective may prove too costly in space requirements and funds to gain wide acceptance. Experience gained at 500 and 765 kV verifies that the prime requirement no longer is insulating the lines to withstand lightning discharges, but insulating them to tolerate voltage surges caused by the operation of circuit breakers. Audible noise levels, especially in rain or humid conditions, are high, requiring wide buffer zones. Environmental challenges have been brought on the basis of possible negative biological effects of the electrostatic field produced under EHV lines, although research has not shown any such effects. *See* ELECTRIC PROTECTIVE DEVICES; LIGHTNING AND SURGE PROTECTION.

Experience has indicated that, within about 10 years after the introduction of a new voltage level for overhead lines, it becomes necessary to begin connecting underground cable. The first 500-kV cable in the United States was placed in service in 1976.

Another approach to high-voltage long-distance transmission is high-voltage direct current (HVDC), which offers the advantages of less costly lines, lower transmission losses, and insensitivity to many system problems that restrict alternating-current systems. Its greatest disadvantage is the need for costly equipment for converting the sending-end power to direct current, and for converting the receiving-end direct-current power to alternating current for distribution to consumers. However, the development of solid-state converter valves made overhead HVDC economical for distances more than about 400 mi (650 km), but only 25–30 mi (40–50 km) in underground construction. Where systems that are out of synchronism must be interconnected, much shorter distances may be economic. An extreme example of this is at Eel River, New Brunswick, Canada, where back-to-back converters connect Canadian and United States systems. Starting in the late 1950s with a 65-mi (105-km) 100-kV system in Sweden, HVDC has been applied successfully in a series of special cases around the world, each one for a higher voltage and greater power capability. The first such installation in the United States was put into service in 1970. It operates at 800 kV line to line, and is designed to carry a power interchange of 1440 MW over an 841-mi (1354-km) overhead tie line between the winter-peaking Northwest Pacific coastal region and the summer-peaking southern California area. These HVDC lines perform functions other than power transfer, however. The Pacific Intertie is used to stabilize the parallel alternating-current transmission lines, permitting an increase in their capability; and back-to-back converters with no tie line between them are used to tie together two systems in Nebraska that otherwise could not be synchronized. The first urban installation of this technology was energized in 1979 in New York. *See* DIRECT-CURRENT TRANSMISSION.

In addition to these high-capability circuits, every large utility has many miles of lower-voltage transmission, usually operating at 110 to 345 kV, to carry bulk power to numerous cities, towns, and large industrial plants. These circuits often include extensive lengths of underground cable where they pass through densely populated areas. *See* ELECTRIC POWER TRANSMISSION; TRANSMISSION LINES.

Interconnections. As systems grow and the number and size of generating units increase, and as transmission networks expand, higher levels of bulk-power-system reliability are attained through properly coordinated interconnections among separate systems. This practice began around 1930.

Most of the electrical utilities in the contiguous United States and a large part of Canada operate as members of power pools, and these pools in turn are interconnected into one gigantic power grid, known as the North American Power Systems Interconnection. The operation of this interconnection, in turn, is coordinated by the operating committee of the North American Electric Reliability Council (NERC). Each individual utility in such pools operates independently, but has contractual arrangements with other members in respect to generation additions and scheduling of operation. Their participation in a power pool affords a higher level of service reliability and important economic advantages.

The Northeast blackout of November 9, 1965, stemmed from the unexpected trip-out of a key transmission circuit carrying emergency power into Canada and cascaded throughout the northeastern states to cut off electrical service to some 30,000,000 people. It spurred the utilities into a chain reaction affecting the planning, construction, operation, and control procedures for their interconnected systems. They soon organized regional coordination councils to cover the entire contiguous United States and four Canadian provinces. Their objective was to further improve reliability of the planning and operation of their generation and transmission facilities. Currently, there are eight regional reliability councils (**Fig. 3**).

Then, in 1968, the North American Electric Reliability Council (NERC) was established to serve as an effective body for the collection, unification, and dissemination of various reliability criteria for use by individual utilities in meeting their planning and operating responsibilities (Fig. 3).

Increased interconnection capability among power systems reduces the required generation reserve of each of the individual systems. In most utilities the loss-of-load probability (LOLP) is used to measure the reliability of electrical service, and it is based on the application of probability theory to unit-outage statistics and load forecasts. A common LOLP criterion is 1 day in 10 years when load may exceed generating capability. The LOLP decreases (that is, reliability increases) with increased interconnection between two areas until a level is reached which depends upon the amount of reserve, unit sizes, and annual load shape in each area.

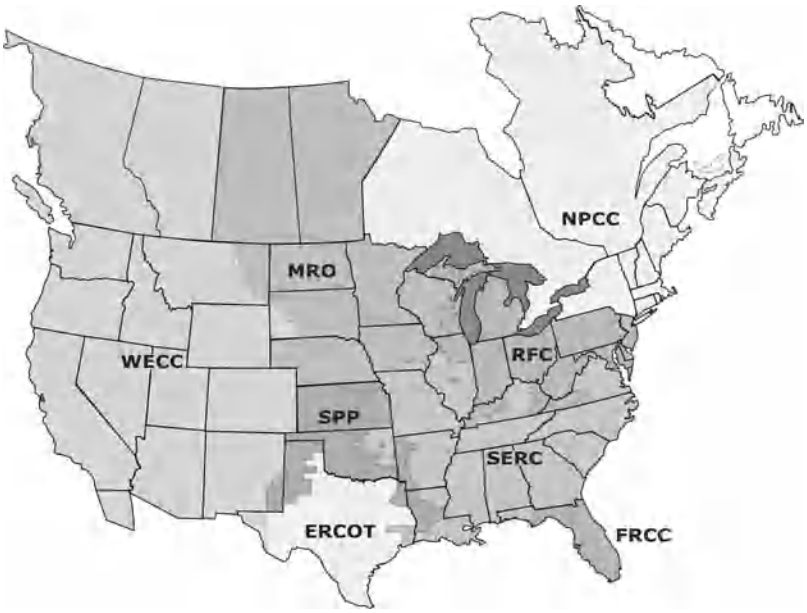

Regional Reliability Councils

ERCOT, Electric Reliability Council of Texas
FRCC, Florida Reliability Coordinating Council
MRO, Midwest Reliability Organization
NPCC, Northeast Power Coordinating Council
RFC, Reliability First Corporation
SERC, SERC Reliability Corporation
SPP, Southwest Power Pool Inc.
WECC, Western Electricity Coordinating Council

Fig. 3. Areas served by the eight regional reliability councils, coordinated by the North American Electric Reliability Council, that guide the planning, coordination, and operation of generation and transmission facilities. (*North American Electric Reliability Council, http://www.nerc.com*)

Traditionally, systems were planned to withstand all reasonably probable contingencies, and operators seldom had to worry about the possible effect of unscheduled outages. Reserve margin, which is the excess of capacity over load at maximum annual peak, while difficult for every system, is generally considered acceptable at 25%. Operators' normal security functions were to maintain adequate generation online and to ensure that such system variables as line flows and station voltages remained within the limits specified by planners. However, stronger interconnections, larger generating units, and rapid system growth spread the transient effects of sudden disturbances and increased the responsibilities of operators for system security.

System security is concerned with service continuity at standard frequency and voltage levels. The system is said to be insecure if a contingency would result in overloading some system components, in abnormal voltage levels at some stations, in change of system frequency, or in system instability, even if there is adequate capability as indicated by some reliability index. The concepts of power system

reliability and security are discussed in more detail below.

Substations. Power delivered by transmission circuits must be stepped down in facilities called substations to voltages more suitable for use in industrial and residential areas. On transmission systems, these facilities are often called bulk-power substations; at or near factories or mines, they are termed industrial substations; and where they supply residential and commercial areas, distribution substations.

Basic equipment in a substation includes circuit breakers, switches, transformers, lightning arresters and other protective devices, instrumentation, control devices, and other apparatus related to specific functions in the power system. *See* ELECTRIC POWER SUBSTATION.

Distribution. That part of the electric power system that takes power from a bulk-power substation to customers' switches, commonly about 35% of the total plant investment, is called distribution. This category includes distribution substations, subtransmission circuits that feed them, primary circuits that extend from distribution substations to every street and alley, distribution transformers, secondary lines, house service drops or loops, metering equipment, street and highway lighting, and a wide variety of associated devices.

Primary distribution circuits usually operate at 4160 to 34,500 V line to line, and supply large commercial institutional and some industrial customers directly. Since about 1950, by far the majority of lines constructed have been in the 15-kV class. Primary lines may be overhead open wire on poles, spacer or aerial cable, or underground cable. Legislation in more than a dozen states requires that all new services to developments of five or more residences be put underground. The bulk of existing lines are overhead, however, and will remain so for the forseeable future.

At conveniently located distribution transformers in residential and commercial areas, the line voltage is stepped down to supply low-voltage secondary lines, from which service drops extend to supply the customers' premises. Most such service is at 120/240 V, but other common voltages are 120/208 and 125/216 V, and for larger commercial and industrial buildings, 240/480, 265/460, or 277/480 V. These are classified as utilization voltages. *See* ELECTRIC DISTRIBUTION SYSTEMS.

Electrical utility industry. In the United States, which has the third highest per capita use of electricity in the world and more electric power capability than any other nation, electrical systems as measured by some criteria are the country's largest industry.

The utility industry in the United States is pluralistic in the nature of its ownership. Investor-owned utilities comprise about 78% of total installed capacity, and about the same percentage of customers served. Publicly owned utilities (that is, municipal, state, and power-district utilities) own and operate about 10% of the installed capacity and serve almost 14% of the customers; cooperatives own 3% of the capacity and serve about 10% of the customers; and federal agencies, while serving no measurable percentage of customers, own almost 10% of installed capacity. Electrical utilities are among the most capital-intensive of industries, primarily because of the huge investment in generating units.

Introduction of deregulation over the past few years has resulted in a significant reorganization of the electric utility industry. Instead of vertically integrated utilities being engaged in all three aspects—generation, transmission, and distribution—the industry is being organized horizontally with independent power producers (IPPs), independent system operators (ISOs), and independent distribution companies (DISCOs). The ISOs manage transmisson of power from the IPPs to the DISCOs to meet the contractual obligations of the IPPs and DISCOs. This structure is supplemented by electricity markets with bidding mechanisms for the sale and purchase of electricity, and long-term and short-term contracts between the IPPs and the DISCOs. In this reorganization, the ISO link is still regulated and has the responsibility of maintaining the reliability of electricity supply.

William C. Hayes; Om P. Malik

Power System Operation

The operation and control of the generation-transmission-distribution grid is quite complex because this large system has to operate in synchronism and because many different organizations are responsible for different portions of the grid. In North America and Europe, many public and private electric power companies are interconnected, often across national boundaries. Thus, many organizations have to coordinate to operate the grid, and this coordination can take many forms, from a loose agreement of operational principles to a strong pooling arrangement of operating together.

Power-system operations can be divided into three stages: operations planning, real-time control, and after-the-fact accounting. The main goal is to minimize operations cost while maintaining the reliability (security) of power delivery to customers. When all the utilities were vertically integrated—that is, the operator was also the owner of the generation and transmission—centrally optimizing the generation production was a main duty of the operator. Many regions of the world have deregulated generation ownership, and in those areas the market transactions determine generation production while the network operator is mainly concerned with the reliability of the grid. Operations planning requires the forecasting of load demand in the next few hours, days, weeks, or months that allows the market to determine the schedules and the operator to approve these schedules if they meet the reliability requirements. The market determines the buying and selling of power for the long and short terms, and after the markets clear (which involves matching the buy and sell bid prices) the

generating companies can schedule hydro resources, fossil fuels, and equipment maintenance over many weeks, and the commitment (startup and shutdown) of generating units over many hours. The operator has to run contingency (what-if) studies using these schedules to ensure that all reliability criteria are met or else the buy–sell transactions have to be adjusted to ensure that the reliability criteria are met.

Real-time control of the system is required to respond to the actual demand of electricity, which always deviates somewhat from the forecast, and any unforeseen contingencies (equipment outages). Maintaining reliability of the system so that a possible contingency cannot disrupt power supply is an integral part of the real-time control function.

After-the-fact accounting tracks actual delivery of energy between organizations so that billing can be generated.

For loosely coordinated operation of the grid, each balancing authority or utility takes responsibility for the operation of its own portion while exchanging all relevant information with its partners. For pool-type operations, a group of utilities sets up a pool where the operational decisions may be made centrally and then implemented by each utility. All of this requires significant data communication as well as engineering computation within a utility as well as between utilities. The use of modern computers and communications technology makes this possible, and the heart of system operations in a utility is the energy control center.

The monitoring and control of a power system from a centralized control center became desirable quite early in the development of electric power systems, when generating stations were connected together to supply the same loads. As electrical utilities interconnected and evolved into complex networks of generators, transmission lines, distribution feeders, and loads, the control center became the operations headquarters for each utility. Since the generation and delivery of electrical energy are controlled from this center, it is referred to as the energy control center or energy management system.

Although these centers have greatly advanced in technology, their basic control objective of maintaining reliability and security remains the same. In a vertically integrated utility the economic goal of minimizing the cost of supply is achieved by optimizing the dispatch and scheduling of its generators, whereas in deregulated regions the electricity markets optimize the supply costs. Increasingly, the generation is being privatized, and the generation companies have to compete in the electricity markets to sell their supply. The computerized tools needed to run the market and those used by the generation companies to determine their bid strategies are becoming more sophisticated. The rest of this section, however, describes the tools that are used to operate the grid itself in a reliable manner that avoids cascading failures and blackouts.

Evolution. In early control centers measurements of important power flows, voltages, and currents were displayed on banks of meters. Circuit-breaker and switch positions were sometimes monitored and displayed by the switching of lights, placed appropriately on a wall map of the power system. Hard-wired control switches were used to open or close breakers remotely from the control centers. These supervisory control and data acquisition (SCADA) functions became more versatile and flexible with the computerization of the control center.

Automatic generation control (AGC), which adjusts the generator outputs to follow the load changes, became absolutely necessary when utilities started to interconnect, and customized controllers were initially built for the purpose. The use of first analog and then digital computers greatly improved the capabilities of automatic generation control. For example, the problem of meeting the load with the cheapest combination of generating sources requires a multistep calculated solution, for which digital computers are ideal.

Digital computer–based SCADA-AGC control centers became common in the 1960s. They could handle more information, provide the information to the operator more efficiently, and provide monitoring and control of the power system more reliably. Also, the availability of unprecedented computational power made it possible to develop new functions for better control. Security analysis and scheduling functions based on on-line computer programs are now available to the control-center operator, providing more economic and secure control.

Configuration. The control-center computers are connected through communication channels, usually microwave links, although fiber optics are increasingly used, to remote terminal units. A remote terminal unit is placed at each substation and generating station, and it can gather all the measurements at that station and execute the control commands sent from the control center.

At the control center itself, many computers must be linked together to accomplish the various functions. For example, one workstation may handle the SCADA functions, another may be dedicated to security analysis, while several more may handle the graphics interface needed by the operator. Such distributed processing configurations (**Fig. 4**) are common, and open-system standards are used for computer operating systems and communications, making it much easier to expand and modify existing hardware configurations. *See* DISTRIBUTED SYSTEMS (COMPUTERS); DISTRIBUTED SYSTEMS (CONTROL SYSTEMS).

Redundancy has to be carefully designed into the control-center configuration so that the probability of losing the critical functions is very small. In addition to redundant computers, backup is usually provided for communication channels, remote terminal unit circuits, operator consoles, and other equipment. *See* MULTIPROCESSING.

Human-machine interface. The banks of control panels have been replaced by operator consoles which typically consist of color display screens, a keyboard with extra function buttons, logging

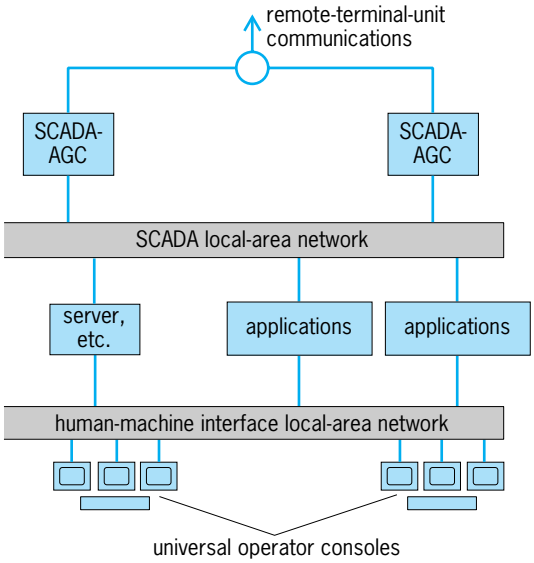

Fig. 4. Typical distributed-processing computer configuration for an energy control center.

printers, and telephones. All the measured and calculated data can be viewed at the display screens, and all control can be initiated and monitored from these consoles. A control center will usually have several consoles (**Fig. 5**), each dedicated to a particular aspect of system operation, like interchange monitoring or transmission control. Because the consoles are computer-controlled, their responsibilities can be changed at any time to accommodate changing procedures or availability of operators. The use of wall-mounted maps of the power system is quite common. Increasingly, these maps are very large liquid-crystal displays (LCDs) or plasma displays and are driven by the computers to denote breaker positions, line overloading, or out-of-limit voltages. *See* CATHODE-RAY TUBE; ELECTRONIC DISPLAY.

The control-center computers must handle thousands of displays, many of which are continually accessed and updated. This imposes a high demand on the computers, especially during emergencies, when the operator has to absorb and react to rapidly changing power-system conditions.

Control centers use display generator programs that allow the on-line creation of displays by the users, without computer-programming changes. Both the hardware and software techniques for graphics have improved dramatically over the years, and the quality of the human-machine interface is very high.

Supervisory control and data acquisition. All measurement data are scanned periodically by the computers, usually every 2 to 4 s. These data and others calculated from them are available to the operators through the displays. The data are checked for limit violations, and alarms are generated if necessary. Computerization and good database management programs have made it feasible for a control center to handle data from tens of thousands of information points. *See* DATABASE MANAGEMENT SYSTEM.

The supervisory control for switching of circuit breakers is usually done by using the displays and the function keys on the keyboard, with automatic checking and verification by the computers.

Automatic generation control. To match the load at all times by controlling generation, an automatic feedback control called load frequency control is used. The errors in frequency and scheduled power interchanges with neighboring utilities are continually checked, and control signals are sent to the generating units to adjust their generation levels. The units that can be controlled are those that the generation owners have agreed to put under control and those that have the ability to ramp their outputs up and down. The generators are compensated for this service at a contracted or market rate. (**Fig. 6**).

Fig. 5. Operator consoles at the energy control center of the New England Power Exchange (NEPEX) in Holyoke, Massachusetts. (*ESCA Corporation,* © *Richard G. Shaw*)

The generation control functions also include reserve monitoring to ensure adequate generation reserve, and interchange scheduling to automatically schedule contracted sales and purchases of energy between companies. If generation reserve is not adequate, the operator has to buy reserve capacity, and if interchange scheduling cannot be maintained at the contracted levels because of reliability considerations, the operator has to follow procedures to change the scheduled interchanges.

Security analysis functions. This set of functions (**Fig. 7**) is the direct result of the availability of powerful digital computers at control centers. A mathematical model of the power system can be stored in the computer and automatically updated in real time by using the measurements that are being continually monitored by the data-acquisition system. A program called the state estimator can track the real-time conditions of the power system with the stored model.

This real-time model can be used to provide the operator with answers to security questions. The contingency analysis program can analyze the effect of all probable contingencies on this model. If equipment failures at particular locations are determined to cause unacceptable operating conditions, the operator is alerted and can change the operating conditions so that these contingencies are no longer threatening to the system. This contingency analysis can be a static analysis (power flow), but for those systems that have stability problems a dynamic analysis is also used.

By using a study power flow program, the operator can also study this model for other purposes. If changes to the operating condition are contemplated by the operator, they can be first studied on this model to determine if they produce the desired results. An enhanced version, called the optimal power flow, can even determine the best operating strategies that should be tried to get the desired results.

Scheduling functions. For generation owners, optimizing costs is important. Steam-generating units have limits on how fast they can be started up or shut down. Hydroelectric units have limited available energy depending on the storage capacity of their reservoirs. These and other constraints affect the commitment and decommitment of generating units to meet the daily and weekly cycles of load. Unit commitment and hydrothermal coordination computer programs are available to determine the most economic scheduling pattern. With the availability of large computers these programs can be integrated with operations data so that the scheduler can update the results conveniently for changing conditions. These operations-planning functions, especially those that plan over many weeks, are often performed by engineering personnel distinct from system operators, on a separate computer system that shares data with the control-center computers.

In regions where generation owners are deregulated, the scheduling of generation depends on long-term (weeks) bilateral contracts and short-term (24 hours or less) market transactions. These markets are computerized and use some sort of auction method to pick the bids and clear the market. Although the basic energy market is straightforward, the markets for reserve capacities and other ancillary services can be rather complicated. Moreover, even after the markets have cleared, the buy–sell transactions have to be checked by the system operator for their feasibility. This is because the markets clear on bid prices without regard to whether the transactions cause congestion (transmission overloads). Thus the operator has to check for congestion and manage the congestion by altering the transactions. This can be done iteratively between the market computers and the reliability studies or in one step by using the optimal power flow.

Logging and accounting. Computers have largely automated the logging and accounting functions, resulting in more comprehensive record keeping of the actual power flows and reduced labor costs. Daily, weekly, and monthly reports are produced automatically.

Computer programs are used to determine the cost of electric energy using the contract or market prices and the actual flows which are always

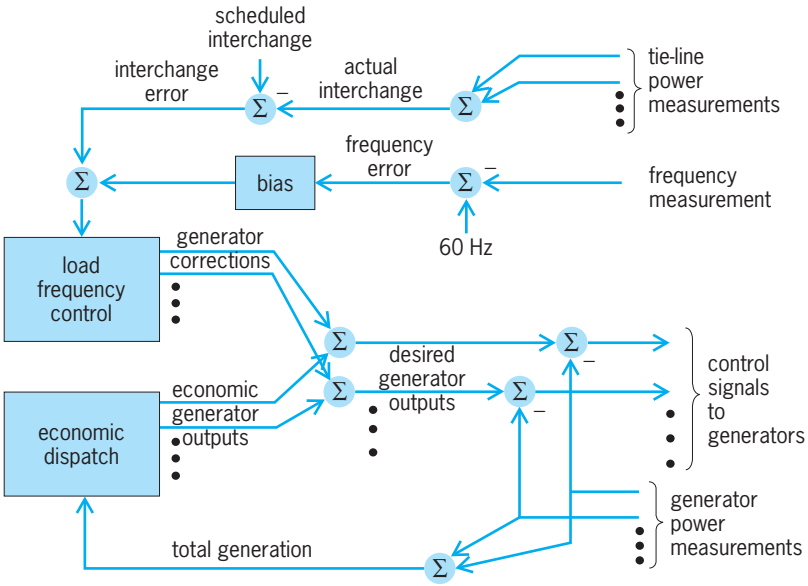

Fig. 6. Flow chart of the automatic generation control functions of an energy control center.

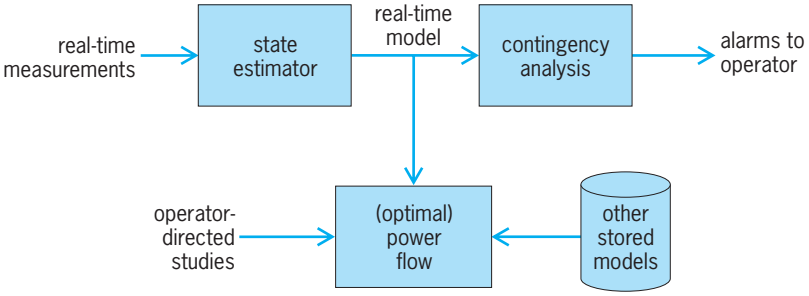

Fig. 7. Flow chart of the security analysis functions of an energy control center.

different from the contracted flows. The purpose is to automatically calculate the billing for energy that was bought and sold between generating companies and the distribution companies or loads. The same data can be used to compensate the transmission companies for their transmission service.

Operator training simulator. As the functions conducted at the control center have become more sophisticated and operating the power system more complex, the training of the operator has also become more important. The use of simulation training, like that used to train airline pilots and air-traffic controllers, has become common because there are no good alternatives for learning to handle emergency situations. The capability of computers to model the power system in real time has made this simulation training possible.

Impact of deregulation. The worldwide trend toward deregulating the electric power industry has had a significant impact on how the electric grid is operated. The presence of more organizations that has resulted from the separation of generating, transmission, and distribution companies has increased the need for new operational procedures to ensure the reliability and security of electricity supply. Here also, superior communication and computation technology play a vital role; in fact, such deregulation would not be feasible without them. Some of the trends in separating the economic functions from the reliability functions are described above, including some of the new tools like market systems, but changes are happening at different rates in different parts of the world and best practices for operating these deregulated systems are still emerging.

Anjan Bose

Power-System Planning

The two traditional objectives of power-system planning are (1) to provide electricity as cheaply as possible while (2) satisfying standards of reliability, safety, and quality. However, since about 1980, social, economic, and technical changes have altered the orientation of power-system planning everywhere. Power system planning, once done by vertically integrated utilities in isolation, has become a cooperative and even a competitive process.

Load forecasting. Planning begins with load forecasts. For decades, demand in the United States grew at about 7% per year and was correlated to the gross domestic product, but since about 1970 the rate of load growth has averaged about 2% per year. Overall demand is forecast in three ways: extrapolation of historic growth, econometric modeling, and end-use forecasting. A fourth technique, based on historic experience, is used to forecast local-area load growth. For day-to-day or hour-to-hour operations forecasting, weather and social patterns dominate.

Extrapolation. If load grows by the same percentage each year, the logarithm of load grows at a linear rate. A straight line passing through and beyond a plot of the logarithm of historical data gives an extrapolation forecast of the logarithm of load; load itself is forecast by taking the antilog. This method fails if the rate of load growth changes. *See* ALGEBRA; EXTRAPOLATION; LOGARITHM.

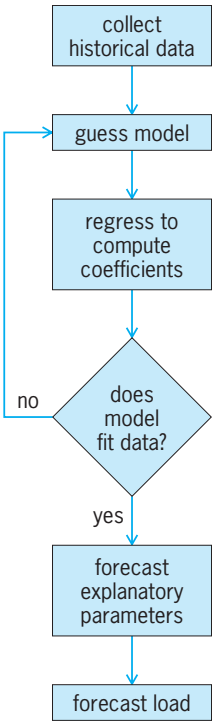

Fig. 8. Econometric load forecasting.

Econometric modeling. Mathematical models (formulas) can be developed that relate changes in load to changes in explanatory variables like population, gross domestic product, prices, and so forth (**Fig. 8**). Historical data are collected on loads and on possible explanatory factors. Often, the first formula or model tried is linear, like the equation below. Regression is

$$\text{LOAD} = k_1 + k_2 \times \text{POPULATION} + k_3 \times \text{GROSSDOMESTICPRODUCT} + \cdots$$

a mathematical procedure for computing the values of the coefficients k_1, k_2, and so forth, to make the model consistent with the historical data. Statistical tests measure just how closely the model matches the data. *See* STATISTICS.

Once the model adequately relates the historical explanatory data to past loads, the explanatory factors are forecast. Finally, the formula or model is used to compute the load forecast from the forecasts of explanatory factors.

Econometric modeling often works better than extrapolation. But both methods infer future load growth from historical data—a reasonable approach only if major changes in the economy do not occur.

End-use forecasting. Residential load forecasts can be built up from energy-use forecasts. The number of homes is forecast first. Major appliances are identified and their market penetrations are projected. The annual energy used by each type of appliance is also projected. This is multiplied by the appliance penetration and by the number of homes to give an energy forecast per appliance. The residential forecast is the sum of the appliance forecasts. This

method takes more effort than the other two and does not work well for commercial or industrial loads, but it depends less on the assumption that the past foretells the future.

Local-area forecasting. Forecasting loads in local areas is important for subtransmission and distribution system planning. Load often develops in an S-shaped pattern, with low historic loads succeeded by rapid growth as new residential or commercial development occurs, followed by low growth again when development saturates. Forecasting is based on experience with prior development patterns.

Generation planning. Centralized generation planning is done by utilities or other organizations having a statutory or implied "obligation to serve" and guaranteed recovery of costs. Generation options include building new utility-owned power plants, buying power from independent power producers or other entities, retiring old generating units or extending their lives, and repowering older units to increase their capacity or efficiency.

Which options are most attractive varies from country to country, and among regions in larger countries, depending on local conditions. For example, there are still vast unexploited hydroelectric power resources in Latin America, Asia, and Africa, while in much of Europe and the United States most of the economically attractive hydroelectric power has been developed.

Plans traditionally are evaluated in terms of two main criteria: reliability and cost (variable and fixed). They must also satisfy environmental and other requirements. Planners gage reliability in terms of reserves (installed generation capacity minus peak demand) or any of several probabilistic measures of the likelihood that enough generators will be available to meet the load.

Variable costs are roughly proportional to how heavily a plant is operated; fuel costs are an example. Variable costs include not only the cost of a particular option but also how it will affect the costs of the rest of the system. Variable costs often are calculated by using a production-cost computer program that simulates the operation of the region's plants.

Fixed costs are independent of how a particular plant is run. They include interest, depreciation, profits, taxes, and so forth. In the United States, fixed and variable costs are often added and expressed as revenue requirements, the amount that customers will have to pay if the utility selects a particular option. To choose among several options, the present worths of their revenue requirements are compared.

Planners also assess the effects of load forecast uncertainty, generator forced outages, possible variations in fuel price, uncertainty in performance of new or rebuilt units, construction cost overruns, and, particularly in less-developed countries, uncertainty in construction time.

While an independent power producer (IPP) can be a generation-planning option from the perspective of a utility, it adds uncertainty and competes with utility-owned generators. It is evaluated much like a utility-owned plant in terms of its effect on cost of electricity, system reliability, and so forth. If the IPP's cost or performance is uncertain, the utility must ensure that it will be allowed to pass cost variations to its customers through tariffs. If there is uncertainty with regard to whether the IPP will be built, the utility must have contingency plans to replace it. If the IPP is viewed as a competitor, the utility attempts to design and operate its own generators at a cost lower than the price the IPP can afford to charge.

Planning is done from an entirely different perspective by IPPs or other entities without guaranteed cost recovery and without an obligation to serve. Decisions are based on market analyses, where load growth is only one of the critical drivers. Other drivers include the costs of existing and possible future generating units with which the IPP will compete. The IPP also considers regulations and market rules, which can affect its operation and its profitability. This type of planning is different in detail but not in concept from the development planning that any commercial organization must make. Planning is complicated by the high initial cost and long economic lives of generating units. Utilities are happy to recover their investments over 40 years. Most investors want less than 10 years.

Transmission planning. A transmission system consists of current-carrying hardware (lines, transformers, structures, and so forth), control and communication equipment (computers, fuses, relays, and so forth), and practices and procedures. The transmission system is coupled to the generation system; together, they form an integrated energy conversion machine. The transmission system is not a transportation system like a gas pipeline, and it cannot be analyzed independent of the generation system.

One transmission planning option is construction of new alternating-current or direct-current lines, as discussed above. Other options include upgrading existing lines, installing transformers or compensation devices (capacitors and reactors), and introducing new operating procedures. Advances in semiconductor devices have greatly increased the variety of control techniques available. Collectively called FACTS (flexible alternating-current transmission system), these control devices allow utilities to utilize more fully the capability of the transmission system and to improve reliability. *See* STATIC VAR COMPENSATOR.

The final decision between two technically equivalent options is based on what each would cost the ratepayer, but the system must meet reliability criteria.

Reliability is assessed with computer programs that simulate network flows by using complex mathematical models of the generation and transmission system. The most sophisticated of these models simultaneously solve tens of thousands of nonlinear equations, with tens of thousands of unknowns; solution techniques are so powerful that this takes a few seconds on a personal computer.

Engineers often plan the transmission system to withstand the sudden loss of any single generating unit or network element under peak-load conditions; this is called the first-contingency criterion. Since this criterion requires the system to withstand the most severe contingency, the system is presumably more than adequate for less serious failures, and can generally withstand more than one of these.

Strategic and integrated resource planning. Strategic planning is concerned with long-range decisions affecting the fundamental nature of an enterprise. Options include acquisitions, divestitures, mergers, and diversification.

Least-cost planning or integrated resource planning (the terms are synonymous) is strategic planning for energy. It differs from tactical generation and transmission planning in three ways. First, integrated resource planning considers more options, with unbiased treatment of supply (generation) options and demand options such as conservation and load shifting. Second, integrated resource planning seeks to reconcile the objectives of a variety of stakeholders. Finally, integrated resource planning deals in a more fundamental way with risk and uncertainty.

Options. Numerous demand options are available, as discussed more fully below. Some are active; the utility can switch off loads when system conditions require it. Others are passive, relying on price signals to influence customers to reduce or shift demand. Others are both active and passive: in demand subscription service, for example, customers are paid to join a program requiring them to reduce load to their subscription level if the utility requests it.

Objectives. While generation and transmission planning concentrate on two main objectives (reliability and ratepayer cost), integrated resource planning also recognizes costs and benefits perceived by the investors, the regional or national economy, neighbors concerned with the environment, utility employees, and so forth. These costs and benefits cannot all be reduced to monetary terms, and they represent objectives that often conflict. For instance, a clean-burning premium fuel is good for the neighbors but expensive for the ratepayer.

The traditional approach for resolving conflicting objectives is to create a "utility function" by simply adding all of the objectives, each multiplied by an appropriate weighting factor. The utility function is then optimized. An extensive body of theory has been developed to support this approach. The basic difficulty is that objectives often are truly incommensurate. Even something as obviously economic as reliability cannot be expressed precisely in monetary terms.

A newer approach is to express all objectives in their natural units and to seek solutions that give the best trade-offs or compromises among these conflicting objectives.

Uncertainties and risk. An uncertainty is a parameter whose value is unknown. Uncertainties can be modeled probabilistically or as unknown-but-bounded variables, depending on the circumstances. Both types of models have their place. A common error is to insist on using a probability model, because it appears to be more elegant mathematically, in circumstances where it is not valid.

For hundreds of years mathematicians and businessmen have studied the problem of risk. Some business risks are quantified in economic terms.

Nonetheless, many risks cannot be so quantified and managed. It is not reasonable to cover risks of terrorism, for example, by simply buying an insurance policy. The risks are not purely economic. Similarly, the risks of a blackout cannot be monetarized.

The approach to risk described here does not require that risks be reduced to dollar terms.

Risk is the hazard to which an entity is exposed because of uncertainties. Measures of risk include the likelihood or conditions under which a particular decision will be regrettable, and the magnitude of the possible regret. A decision that is not likely to be regretted, or where the possible regret is low, is said to be robust.

Integrated resource planners assess risk and develop hedges to manage it. Hedges include building small plants rather than big ones, buying futures contracts to protect against price volatility, doing pilot demand-management studies before selecting a demand-side option, and designing expansion so that it can be staged.

For example, one utility considered two major planning options: building a large coal-fired plant near the load, and building a subarctic hydroelectric plant with a long transmission line. The life-cycle costs were about the same, with the hydroelectric plant perhaps a bit cheaper, but the utility opted for the coal plant to avoid the risks associated with subarctic construction and a long line.

Regulatory changes. The regulatory compact is an unwritten agreement that permits electric utilities to enjoy monopoly protection and guaranteed cost recovery and profits. In exchange they accept an obligation to serve all comers and eschew monopoly profits. This century-old compact is being changed in a variety of ways in much of the world. Some feel that the industry is now mature and does not need special protection, and that a competitive market can do better than a regulated monopoly to meet customer needs at the lowest price.

Changes in the regulatory structure of the industry create important technical options and uncertainties. One key issue concerns the handling of transmission access and power wheeling, the provision of transmission services to energy buyers and sellers who are not directly interconnected. There are difficult engineering and equity problems associated with this issue, and they have not been resolved completely.

Planning in developing countries. Electric power is less available in the developing world because (1) its cost is beyond the means of many families and businesses, and (2) capital is not available for the massive investments needed to build generation and distribution facilities. These barriers can be overcome in part by planning and building the power system differently. Reliability and other quality-of-service standards in Addis Ababa should not be the same as in Tokyo. Furthermore, economic analyses

must use high discount rates (cost of capital) to reflect unavailability of capital to utilities and consumers. This means that minimizing first cost is more important than minimizing life-cycle costs. Equipment and development should be designed to be cheap initially, even if they must be replaced sooner rather than later. Hyde M. Merrill

Power System Reliability

An electric power system serves the basic function of supplying customers, both large and small, with electrical energy as economically and as reliably as possible. The reliability associated with a power system is a measure of its ability to provide an adequate supply of electrical energy for the period of time intended under the operating conditions encountered. Modern society, because of its pattern of social and working habits, has come to expect the power supply to be continuously available on demand. This availability, however, is not physically possible in reality due to random system failures which are generally outside the control of power system engineers, operators, and planners.

The probability of customers getting disconnected can be reduced by increased investment during the planning phase, the operating phase, or both. Overinvestment can lead to excessive operating costs, which must be reflected in the tariff structure. Consequently, the economic constraints can be violated even though the system may be highly reliable. On the other hand, underinvestment can lead to the opposite situation. Design, planning, and operating criteria and techniques have been developed over the past three or four decades in an attempt to resolve and satisfy the dilemma between the economic and reliability constraints.

The criteria and techniques first used in practical applications were basically deterministic (rule-of-thumb) ones; for instance, installed generating capacity equals the expected maximum load demand plus a fixed percentage of the expected maximum demand. The essential weakness of these methods is that they do not account for the probabilistic and stochastic nature of system behavior, customer load demands, and component failures. Such aspects can be considered only through probabilistic criteria.

It is important to conjecture at this point what can be done regarding reliability assessment and why it is necessary. Failures of components, plant, and systems occur randomly; the frequency, duration, and impact of failures vary from one year to the next. Generally, all utilities record details of the events as they occur, and produce a set of performance measures, such as system availability, estimated unsupplied energy, number of incidents, and number of hours of interruption. These measures are valuable since they identify weak areas needing reinforcements and modifications, establish chronological trends in reliability performance, establish existing indices which serve as a guide for acceptable values in future reliability assessments, enable previous predictions to be compared with actual operating experience, and monitor the response to system design changes.

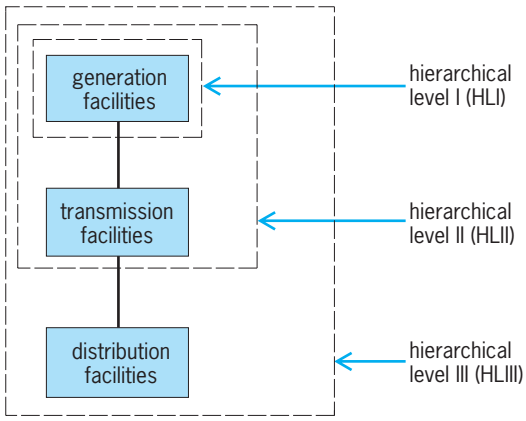

Fig. 9. Hierarchical levels in reliability evaluation.

Hierarchical levels. A modern power system is complex, highly integrated, and usually very large. Analyzing the entire system as a whole, apart from being extremely difficult, may provide results that are so vast that meaningful interpretation will be difficult, if not impossible. The power system can be categorized, for purposes of planning, organization, control, monitoring, operation, maintenance, and so forth, into three distinct zones of generation, transmission, and distribution. While this subdivision of the system may seem somewhat simplistic, most power utilities are divided into these zones or segments, and perhaps some smaller subzones such as subtransmission. The three zones perform quite different functions but must act in unison to supply the customer with reliable electric energy. These zones can also be combined to form three distinct hierarchical levels (HL) [**Fig. 9**], and reliability evaluation can be performed at one or more of these levels. There are two main approaches that can be used to evaluate the reliability of power systems regardless of the hierarchical level: analytical and simulation.

Adequacy and security. The concept of power system reliability, the overall ability of the system to satisfy the customer load requirements economically and reliably, is extremely broad. For the sake of simplicity, it can be divided into the two basic aspects of system adequacy and system security. Adequacy relates to the existence of sufficient facilities within the system to satisfy customer load demands. These include the facilities to generate power, and the associated transmission and distribution facilities required to transport the generated energy to the load points. Adequacy therefore relates to static system conditions. Security pertains to the response of the system to the perturbations or disturbances to which it is subjected. These may include conditions associated with local and widespread disturbances and loss of major generation or transmission. Most of the techniques presently available are in the domain of adequacy assessment.

Reliability cost and reliability worth. Due to the complex and integrated nature of a power system, failures in any part of the system can cause interruptions which range from inconveniencing a small number of local residents to a major and widespread

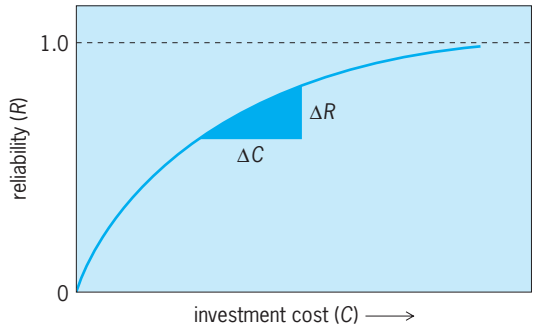

Fig. 10. Incremental cost of reliability.

catastrophic disruption of supply. The economic impact of these outages is not necessarily restricted to loss of revenue by the utility or loss of energy utilization by the customer but, in order to estimate the true costs, should also include indirect costs imposed on customers, society, and the environment due to the outage. For instance, 84% of the total costs of the 1977 New York blackout were attributed to indirect costs. In order to reduce the frequency and duration of these events, it is necessary to invest in the design phase, the operating phase, or both. This involves answering a few difficult questions—how much should be spent? is it worth spending any money? should the reliability be increased, maintained at existing levels, or allowed to degrade? who should decide, the utility, a regulator, the customer? on what basis should the decision be made?

The first step in answering the above questions is illustrated in **Fig. 10**, which shows how the reliability of a product or system is related to the investment cost; that is, increased investment is required in order to improve reliability. This figure clearly shows the general trend that the incremental cost ΔC to achieve a given increase in reliability ΔR increases as the reliability level increases. Alternatively, a given increase in investment produces a decreasing increment in reliability as the reliability is increased. In either case, high reliability is expensive to achieve.

HLI evaluation. The total problem can be divided into two conceptually different areas designated as static and operating capacity requirements. The static capacity area relates to the long-term evaluation of this overall system requirement. The operating capacity area relates to the short-term evaluation of the actual capacity required to meet a given load level.

One practice that has developed among utilities over many years is to measure the adequacy of both the planned and installed capacity in terms of the percentage reserve, which tends to compare the relative adequacy of capacity requirements provided for totally different power systems on the basis of peak loads experienced for each system over the same time period, and also does not attach any penalty to a power-generating unit because of size, unless the unit size exceeds the total capacity reserve.

During planning, it is necessary to determine how much capacity needs to be installed in order to satisfy the expected demand at some point in the future, and to provide sufficient reserve to perform corrective and preventive maintenance. The ability to move the generated energy to bulk supply points and customers is not considered at HLI.

Generating capacity reliability (HLI reliability) is defined in terms of the adequacy of the installed generating capacity to meet the system load demand. Outages of generating units in excess of the estimates could result in "loss of load," that is, the available capacity (installed capacity minus capacity on outage) being inadequate to supply the load. In general, this condition requires emergency assistance from neighboring systems and emergency operating measures such as system voltage reduction and voluntary load curtailment. Depending on the shortage of the available capacity, load shedding may be initiated as the final measure after the emergency actions.

Probabilistic criteria and indices used in HLI studies are:

1. Loss of load probability (LOLP), the probability that the load will exceed the available generation. Its weakness is that it defines the likelihood of encountering trouble (loss of load) but not the severity.

2. Loss of load expectation (LOLE), defined as the average number of days (or hours) per period on which the load is expected to exceed the available capacity.

3. Loss of energy expectation (LOEE), the expected energy that will not be supplied (in kilowatt-hours or megawatt-hours per period) due to those occasions when the load exceeds the available generation. It is presently less used than LOLE but is a more appealing index since it encompasses severity of the deficiencies as well as their likelihood.

4. Frequency and duration (F&D) indices, which measure how often a power outage occurs (frequency, in occurrences per period), and how long the outage lasts (duration, in hours per occurrence) each time it occurs.

The adequacy of the generating capacity in a power system is normally improved by interconnecting the system to another power system. The actual interconnection benefits depend on the installed capacity in each system, the total tie-line capacity, the load levels, and the type of agreement in existence between the systems. A loss-of-load situation is considered to arise in a system when the available assistance through the interconnections cannot offset a capacity deficiency arising due to capacity outages and load demands in that system.

HLII evaluation. An important element in the planning process, which is not considered in HLI studies, is the development of a suitable transmission network to transport the generated energy to the customer load points. The transmission network can be divided into the two general areas of bulk transmission and distribution facilities. The distinction between these two areas cannot be made strictly on a voltage basis but must include the function of the facility within the system.

In addition to providing the means to move the generated energy to the terminal stations, the bulk transmission facilities must be capable of

maintaining adequate voltage levels and loadings within the thermal limits of individual circuits, and also of maintaining system stability limits. The total problem of assessing the adequacy of the generation and bulk power transmission systems in regard to providing a dependable and suitable supply at the terminal stations can be designated as composite system reliability evaluation. Quantitative assessment of the adequacy of a composite system can be performed using a contingency evaluation approach. The basic procedure involves the selection and evaluation of contingencies (component failures), the classification of each contingency according to related failure criteria, and the accumulation of adequacy indices. Two types of indices exist: load point (bus) indices and system indices. Load point indices reflect the adequacy of individual bulk supply points, while the system indices reflect the overall adequacy of the entire system.

HLIII evaluation. HLIII includes all three functional zones in an actual assessment of customer load point reliability. The HLIII reliability performance evaluation is extremely complex. It can, however, be performed (under certain specified conditions) by utilizing the major load bus indices at HLII as input to the distribution networks.

Over the past few decades, distribution systems have received considerably less attention in terms of reliability modeling and evaluation as compared to the generating systems. The main reasons are that the generating stations are individually very capital intensive and that the generation inadequacy can have widespread catastrophic consequences for both society and its environment. Consequently, greater emphasis has been placed on ensuring the adequacy and meeting the needs of this part of the power system. A distribution system, on the other hand, is relatively cheap and outages have very localized effect. Therefore less effort has been devoted to quantitative assessment of various alternative designs and reinforcements. Analysis of customer failure statistics, however, indicates that for many utilities the distribution system makes the greatest contribution to the unavailability of supply to the customer.

The three basic reliability indices of average failure rate (number of failures per period of time), average outage time (duration of failure in hours per occurrence), and average annual outage time (hours per year of power interruptions) are calculated in HLIII studies. Although the three primary indices are fundamentally important, they do not always give a complete representation of the system behavior and response. For example, the same indices will be evaluated irrespective of whether 1 or 100 customers were connected to the load point or whether the average load at a load point was 10 kW or 100 MW. In order to reflect the severity of a system outage, additional reliability indices can be, and frequently are, evaluated: System average interruption frequency index SAIFI, customer average interruption frequency index (CAIFI), system average interruption duration index (SAIDI), customer average interruption duration index (CAIDI), average service availability index (ASAI), and average service unavailability index (ASUI). (ASAI indicates the percentage ratio of customer-hours of service provided to customer-hours requested.) ASUI indicates the percentage ratio of customer-hours of service not provided to customer-hours requested. *See* RELIABILITY, AVAILABILITY, AND MAINTAINABILITY. Lalit Goel

Demand-Side Management

Electricity is produced and distributed in direct response to the demands of residences, business, and industry. These demands come from the billions of electrical devices and appliances which are switched on and off. The demand which each device puts on the electrical system varies by its electrical characteristics, by the way it is used, and by the time of day. Power-system designs are based on plans in which power-system engineers try to forecast the future pattern and amount of the demand for electricity. That prediction is based on a complex set of formulas which consider a number of critical components. These considerations include the electrical demand from the appliances and devices that consume electricity, including predictions of how their population and efficiency may change over time; how new devices may become popular in the future; how people's behavior, or the way people use the devices, may change over time; how economic activity will allow business and industry to flourish and permit enough disposable income to cause consumers to purchase electricity; and the price of electricity and of competing energy forms (such as natural gas).

The role of electric power has grown steadily in both scope and importance over the past century. Developments in key technologies—including electric lighting, motors, computers, and telecommunications—have continuously reshaped daily life and increased the productivity of its commercial and industrial foundations. But the nature of end-use devices is changing dramatically. New uses are emerging from the convergence of developments in energy, telecommunications, transportation, the Internet, and electronic commerce, thus affecting the overall use of electricity.

These new devices require advanced, digital-quality service capabilities. In the United States, the electricity supply system was designed decades ago to nominally maintain 99.99% (5 sigma) reliability, while effectively ignoring outages lasting less than 5 minutes. This was generally acceptable in the twentieth century's analog, electromechanical economy. However, this traditional reliability standard is no longer relevant in today's digital world, where a power interruption or distortion of as little as a quarter of a cycle (0.004 second) can shut down a microprocessor-controlled end-use electrical device or process. This sensitivity infers a systemic (combined supply network plus end-use device) reliability requirement of the order of 10 sigma plus. By comparison, the current system performance results in an electricity unreliability cost to the U.S. economy that has grown to about a 50 surcharge on every dollar of electricity purchased, with no

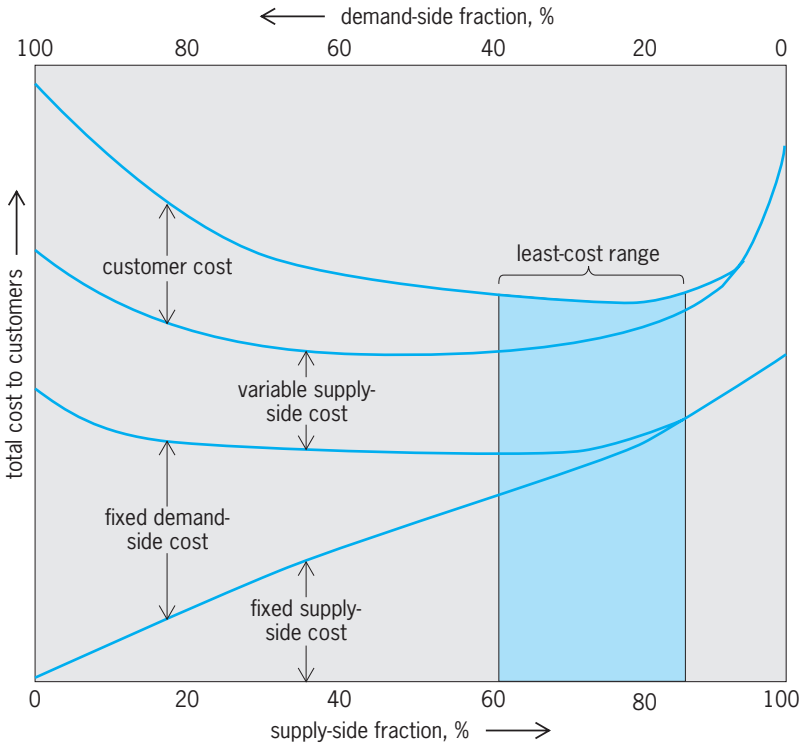

Fig. 11. Balance between supply-side and demand-side management activities.

end in sight until the system is modernized and its reliability transformed. This required performance improvement can most effectively be met by a "systems" approach involving a combination of electricity supply and end-use performance improvements.

Meeting the energy requirements of an increasingly digital society will require applying a combination of advanced technologies—including power generation (for example, central and distributed energy resources), interface devices, and less sensitive end-use equipment and circuits. In addition, the electricity infrastructure must be modernized to incorporate real-time communications, sensors, and computational capability.

The future demand for electricity, estimated based on assumptions about energy-consuming devices and their use, has traditionally been treated as a predetermined quantity by utility planners. Their responsibility has been to forecast that quantity and plan the addition of new generating units accordingly. By changing the efficiency of an appliance, or by changing the way it is used, a utility can alter the amount of future demand or alter the pattern of use in order to change the kind of new electrical generation required. However, traditionally, electrical utilities were not involved in customer programs to accomplish these objectives.

However, new technology is needed if society is to leverage the ever-expanding opportunities of networks such as the Internet and if electric utilities' natural connectivity to consumers is to be effectively utilized to revolutionize both the role of the rapidly changing electricity service industry and the way consumers may be connected to electricity markets in the future. This electricity energy service transformation will also open the door for significant efficiency improvements throughout the energy supply chain while facilitating greater use of relatively clean renewable and distributed energy resources.

In order to increase efficiency, control cost, and meet increasing demand for quality, electrical utilities have begun to address the increasingly complex needs of consumers and to shift from supply-side management, which considers only generating capacity, toward demand-side management. The latter promotes the efficient and lowest-cost utilization of facilities and augments the choices available to utilities and their customers. A compromise between supply-side and demand-side management can produce a range of low-cost resources called the least-cost range (**Fig. 11**).

Definition. Demand-side management is the planning and implementation of those utility activities designed to influence customer use of electricity in ways that will produce desired changes in the utility's load shape, that is, in the time pattern and magnitude of a utility's load. Utility programs in the category of demand-side management include load management, new uses, energy efficiency, electrification, customer generation, and adjustments in market share. Although this broad definition of demand-side management has many merits, in practice the term is generally used more narrowly to encompass increased efficiency, load management, and conservation, all with an emphasis on reducing the need for electrical energy or generation capacity.

Although there is an infinite combination of load-shape-changing possibilities, six changes in two categories can illustrate the range of possibilities: (1) load management including peak clipping, valley filling, load shifting, strategic load growth, and flexible load shape; and (2) energy efficiency. These changes are not mutually exclusive and may frequently be employed in combinations.

Load management. The first step in the shift toward demand-side management was the introduction and evolution of load management. Load-management programs focus on reducing customer use strategically at the time of high utility system loads. The goal is to avoid construction of generation facilities that would be operated for relatively few hours per year or costly power purchases when customer loads can be shifted or displaced at a lesser cost. There are several important components of load-management activities, including direct load control, pricing-based options (interruptible or curtailable rates), and thermal energy storage. The main costs to the utility for load-management programs are rate reductions or bill credits that are used to promote customer participation.

Energy efficiency. Efficiency-type demand-side management programs include various means by which a utility encourages the installation and use of high-efficiency appliances and devices. Programs often involve rebates for buying efficient appliances, bill credits, special rates, advertising, and promotion.

Programs have been focused on efficient, building insulation, heating and air-conditioning equipment installation, water heater wraps, high-efficiency lighting, and efficient refrigerators. Often appliance dealers, wholesalers, electricians, and others are involved in the utility efforts.

Consumer portal. Developments in electric power systems are stimulating the integration of a communications system throughout much of the power system. Once that system is present, connectivity to the ultimate consumers and facilitation of demand-side management can be enhanced with communications. This enhancement will allow three new areas of functionality: one that relates directly to electricity services (for example, added billing information or real-time pricing and demand-side management), one that involves services related to electricity (for example, home security or appliance monitoring); and a third that involves what are more generally thought of as communications services (for example, data services).

The integration of electric energy services and communications with consumer facilities and equipment opens a wide variety of opportunities to enhance consumer energy services, and to provide the consumer with access to a variety of new value-added services while enhancing the performance and operation of the energy system. Since consumer actions create the demand for energy that drives the operation of the energy system, communication with consumer equipment has the potential to profoundly impact the future of the energy system as well as to improve energy efficiency.

A consumer portal (CP) represents the technology, data standards, and protocols that enable the two-way communication of electronic messages with consumer-owned networks and intelligent equipment, and provide a pathway to the full development and implementation of a variety of demand-side management and energy-related services.

Energy saving through electricity. Because of the energy conversion which occurs in producing electricity, it has been argued that a unit of electricity saved is equivalent to preserving three units of primary energy. While this is often the case, concerns about electricity efficiency have sometimes extended to an assumption that electricity use is inefficient and adversely impacts the environment. However, a full understanding of energy efficiency requires a focus on total resource requirements. Electricity offers unique precision and control in its application, enhancing its efficiency. In addition, electricity offers benefits in comfort and environmental advantages. Because of these attributes, in almost every case new electric appliances and devices require less total resources than comparable natural gas or oil-fired systems. Increasing applications of electricity or electrification have been fundamental in increasing national productivity.

As originally defined, demand-side management not only is measured in terms of power-systems savings but also includes full-cycle energy efficiency (total resource use), environmental benefits, and productivity. As such demand-side programs can include not only load management and energy efficiency but new technologies which can expand electricity use.

Full-cycle energy efficiency. Reduced electricity consumption is not necessarily the goal of policy aimed at energy efficiency. A more comprehensive goal is to reduce total consumption of energy resources. For example, the total resource efficiency of an advanced high-efficiency air-source electric heat pump may be compared with the total resource efficiency of the best available gas furnace. By tapping solar energy, the heat pump delivers 3.4 or more units of heat for every unit of electricity produced. An older production and delivery system supplying the heat pump might lose 68% of the fuel in power-plant generation (0.32 factor) and 8% in transmission and distribution (0.92 factor), still yielding 100% efficiency. By contrast, gas has only an 86% efficiency, without considering the requirement for an electric fan. *See* HEAT PUMP.

Other examples where electricity appears to have an advantage in terms of total resource efficiency include ground-source heat pumps, plug-in hybrid electric vehicles, and a variety of advanced industrial processes. In these cases, utility policy that is aimed at energy efficiency may imply increased electricity use.

Environmental benefits. The key environmental benefit stems from efficiency-related demand-side management. By obtaining the same end-use service with less electricity, there is a reduction in power-plant operation and fuel usage, implying a reduction in central-station emissions. These environmental benefits are an important figure of merit in measuring the success of demand-side management. However, most electric appliances and devices are far cleaner at the point of end use. A comprehensive demand-side management strategy must examine the comparative environmental benefits of alternative devices and systems.

Productivity. Application of automation and electric technologies can have important productivity implications. The benefits come in form of faster processing, more precise process control, reduced material and process waste, and higher product quality. The net impact is often reduced process cost, implying reduced product cost and increased competitiveness. Enhanced productivity is an important measure of success, even if these productivity gains require greater electricity usage.

Demand-side management practice. System expansion can be delayed or eliminated and the use of critical resources reduced by significantly reducing energy and peak consumption through demand-side management. Indeed, demand-side management can help achieve a broad range of operational objectives by merely changing the systems load shape. Numerous industries have found that changing the pattern of the demand for their product can be profitable. For example, telephone utilities have long offered reduced evening rates to shift demand and to encourage off-peak usage, airlines offer night coach fares, and movie theaters offer reduced matinee

prices—all examples of deliberate attempts to change the demand pattern for a product or service.

In an electric utility, the physical plant is designed to serve the projected demand for electricity in the least-cost manner, given a specified level of desired quality and reliability. If the load shape is not fixed but may be altered, the cost of serving the load can be reduced still further. Cost reduction due to the changes in the load shape arise primarily from three attributes: reduction in the requirements for new assets or energy, higher utilization of facilities, and more efficient operation of facilities.

Higher utilization of existing and planned facilities can also be achieved through a program of load growth. While such programs increase total costs due to higher fuel and other operating expenses, they reduce unit costs by spreading the fixed costs (debt service and dividends) over more units of energy sold. Clark W. Gellings

Bibliography. P. M. Anderson and A. A. Fouad, *Power System Control and Stability*, 2d ed., 2003; R. Billinton and R. N. Allan, *Reliability Evaluation of Power Systems*, 2d ed., 1996; Edison Electric Institute, *Statistical Yearbook of the Electrical Utility Industry*, annually; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2001; J. D. Glover and M. S. Sarma, *Power System Analysis and Design*, 3d ed., 2001; T. Gonen, *Modern Power System Analysis*, 1988; J. J. Grainger and W. D. Stevenson, *Power System Analysis*, 1994; E. Kahn, *Electric Utility Planning and Regulation*, 2d ed., 1991; P. Kundur, *Power System Stability and Control*, 1994; R. H. Miller and J. H. Malinowski, *Power System Operation*, 3d ed., 1994; North American Electric Reliability Council, *Electricity Supply and Demand* (10-year forecast), annually; North American Electric Reliability Council, *Reliability Assessment* (10-year forecast), annually; F. Saccomanno, *Electric Power System: Analysis and Control*, 2003; *Transmission Line Reference Book, 200 kV and Above*, Rep. EL-2500, Electric Power Research Institute, 3d ed., 2005; A. J. Wood and B. F. Wollenberg, *Power Generation, Operation, and Control*, 2d ed., 1996.

Electric power transmission

The transport of generator-produced electric energy to loads. An electric power transmission system interconnects generators and loads and generally provides multiple paths among them. Multiple paths increase system reliability because the failure of one line does not cause a system failure. Most transmission lines operate with three-phase alternating current (ac). The standard frequency in North America is 60 Hz; in Europe, 50 Hz. The three-phase system has three sets of phase conductors, one to four conductors per phase, and system voltage is defined as the root-mean-square voltage between any two of the phases. Long-distance energy transmission occasionally uses high-voltage direct-current (dc) lines. *See* ALTERNATING CURRENT; DIRECT CURRENT; DIRECT-CURRENT TRANSMISSION.

Overhead transmission lines are used in open areas such as interconnections between cities. However, environmentally preferable underground cables are used in congested areas within cities. The transmission lines generate electric and magnetic fields. The electric fields produce harmless but annoying sparks. For example, a worker standing on a wooden ladder may experience mild sparks when a coworker, standing on the ground, hands over a tool. The magnetic fields have become a source of concern, as discussed below. The use of underground cables eliminates the electric fields completely and reduces the magnetic fields in most cases. The sight of a large transmission line is not pleasing for most people. A further problem is the possibility of electrocution due to traffic accidents or accidental contact. The use of cable eliminates both problems.

The electric power system can be divided into the distribution, subtransmission, and transmission systems. With operating voltages less than 34.5 kV, the distribution system carries energy from the local substation to individual households, uses both overhead and underground lines, and is generally divided into high-voltage distribution (above 1000 V) and low-voltage distribution (110–480 V). *See* ELECTRIC DISTRIBUTION SYSTEMS.

With operating voltages of 69–138 kV, the subtransmission system distributes energy within an entire district and regularly uses overhead lines. For example, it may interconnect large substations located outside a town with the distribution substations inside the town. *See* ELECTRIC POWER SUBSTATION.

With operating voltage exceeding 230 kV, the transmission system interconnects generating stations and large substations located close to load centers by using overhead lines; high-voltage underground cables are used for special purposes only. Examples are the supply of large urban substations or underwater cable connections.

Requirements

The energy transportation system must be reliable and highly efficient. Both transmission-line capacity and supplied-power quality must meet variable load requirements.

Reliability. Transmission lines must transport energy continuously to customers in spite of being subject to the environment, that is, pollution, rain, storms, and lightning. System operation can itself produce disturbances which may interrupt line operation. Present utility practices require that the number of transmission line outages be less than one per year per 100 mi (160 km). The most frequent cause of line outages is the single-phase short circuit caused by lightning strikes, pollution-initiated flashover of the insulator, or trees touching the line during a storm.

Safety. Human activity near or under transmission lines may result in short circuits or accidents. A frequent cause of accidents is contact between the line and an operating crane. Transmission lines must

meet strict safety standards. Most important are strict clearances between the conductor and ground, and between the conductor and supporting structures. The electric and magnetic fields which are in an inevitable by-product of system operation have also caused concern, as discussed below.

Capacity. The maximum power that can be transmitted through a line depends on the size of the conductor, voltage drop, and system stability. The current-carrying capacity of a conductor limits the maximum power for short lines, while voltage drop and system stability are the determining factors for long lines.

Economy. Economic operation of the line is determined by system losses, operation costs, maintenance costs, and investment expenses. The dominant losses are the resistance (I^2R) losses in the conductors.

Power quality. Transmission-line voltage depends on the load. Load current produces voltage drop on the line impedance, reducing line voltage, and line capacitance increases voltage at the line end. Voltage at the receiving end of the line is usually low at heavy load during peak hours, and high at light load during night hours. Proper operation of the electrical system requires maintenance of the line voltage within narrow limits. Typical line voltage is kept between ±5%, and regulation is achieved by switched shunt capacitors and shunt reactances.

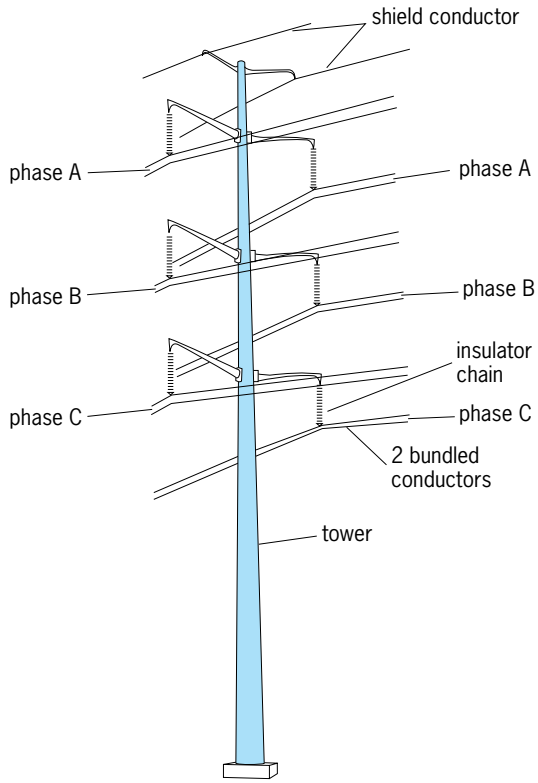

Fig. 1. Double-circuit, high-voltage transmission line.

Overhead Alternating-Current Transmission

Overhead transmission lines distribute the majority of the electric energy in the transmission system. A typical high-voltage line has three (**Fig. 1**) phase conductors to carry the current and transport the energy, and two grounded shield conductors to protect the line from direct lightning strikes. The usually bare conductors are insulated from the supporting towers by insulators attached to grounded towers or poles. The normal distance between the supporting towers is a few hundred feet.

Routes. Most low- and medium-voltage transmission lines are located along highways, roads, or streets. High-voltage lines are placed in transmission-line corridors to reduce interference with the public. Several lines usually share the same corridor.

Hardware. The major components of a transmission line are conductors, insulators, and towers or poles.

Conductors. Transmission lines use ACSR (aluminum cable, steel reinforced) and ACAR (aluminum cable, alloy reinforced) conductors. In an ACSR conductor, a stranded steel core carries the mechanical load, and layers of stranded aluminum surrounding the core carry the current. An ACAR conductor is a stranded cable made of an aluminum alloy with low resistance and high mechanical strength. ACSR conductors are usually used for high-voltage lines, and ACAR conductors for subtransmission and distribution lines. Ultrahigh-voltage (UHV) and extrahigh-voltage (EHV) lines use bundle conductors. Each phase of the line is built with two, three, or four conductors connected in parallel and separated by about 1.5 ft (0.5 m). Bundle conductors reduce corona discharge. *See* CONDUCTOR (ELECTRICITY).

Insulators. Lower-voltage lines use post insulators, while the high-voltage lines are built with insulator chains or long-rod composite insulators. A post insulator is made of either porcelain or composite material and bolted to the cross arm of the tower. The conductor is tied to the top of the insulator. A typical composite post insulator (**Fig. 2***a*) consists of a fiberglass core covered by a silicone or ethylene propylene diene monomer (EPDM) rubber weathershed. The fiberglass provides high mechanical strength, and the rubber assures long leakage distance and good performance in bad weather.

An insulator chain is typically built with ball-and-socket glazed porcelain insulators (Fig. 2*b*). Under typical operating conditions, voltage is around 8–12 kV per unit. Wind swings the vertical chains and reduces clearances to the tower, a movement that is eliminated by a V arrangement.

In a suspension-type composite insulator (Fig. 2*c*), high-strength glass fibers bonded by epoxy or polyester resin form a fiberglass rod that is protected by an injection-molded weathershed made of silicone or EPDM rubber elastomer with hydrated aluminum filler. A corrosion-resistant iron or steel end fitting is glued with epoxy and wedged, crimped, or compressed to the fiberglass rod.

Industrial pollution covers insulators with a contamination layer that becomes conductive when wet by fog or dew, reducing insulator flashover voltage. Pollution-caused flashover results in many service interruptions. In polluted environments, composite

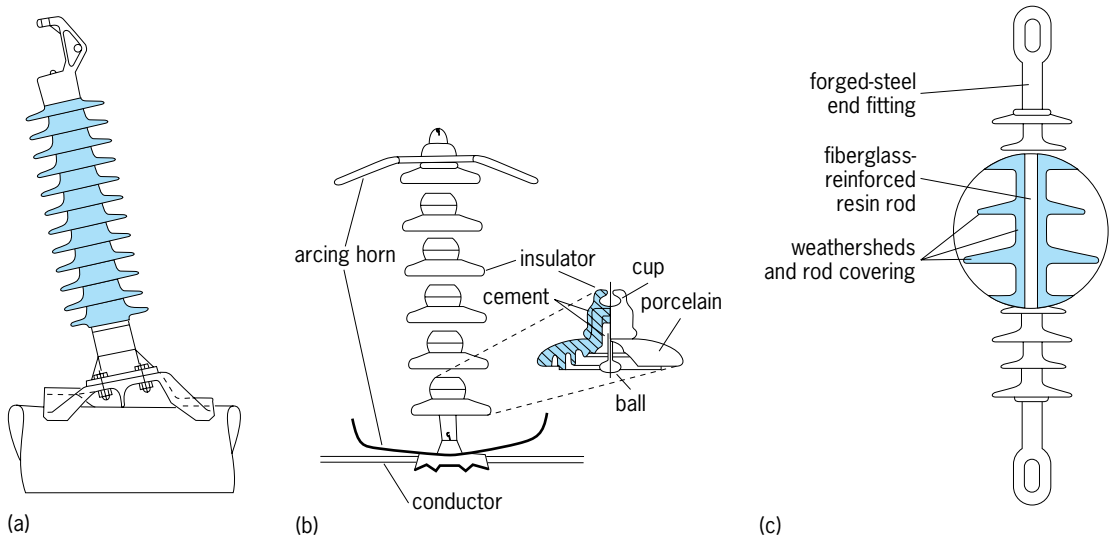

Fig. 2. Transmission-line insulators. (*a*) Composite post insulator. (*b*) Insulator chain. (*c*) Suspension-type composite insulator. (*Sediver, Inc.*)

insulators usually perform better than porcelain insulators. The performance of porcelain insulators can be improved by covering the insulators with silicone grease or a thin composite (silicone rubber) layer. In extremely severe conditions, insulators are washed regularly. Fog-type porcelain insulators are built with deep corrugation on the lower side to increase leakage distance. *See* ELECTRIC INSULATOR; ELECTRICAL INSULATION.

Towers and poles. A transmission line is supported by poles or towers. Distribution lines use wooden poles treated with chemicals (**Figs. 3***a*, *b*). In some areas, wooden poles are replaced by concrete or tapered metal structures, whose higher mechanical strength permits longer span length, reducing the number of towers required for a line. These towers also have longer life spans and better appearance.

Some double-circuit high-voltage lines use lattice

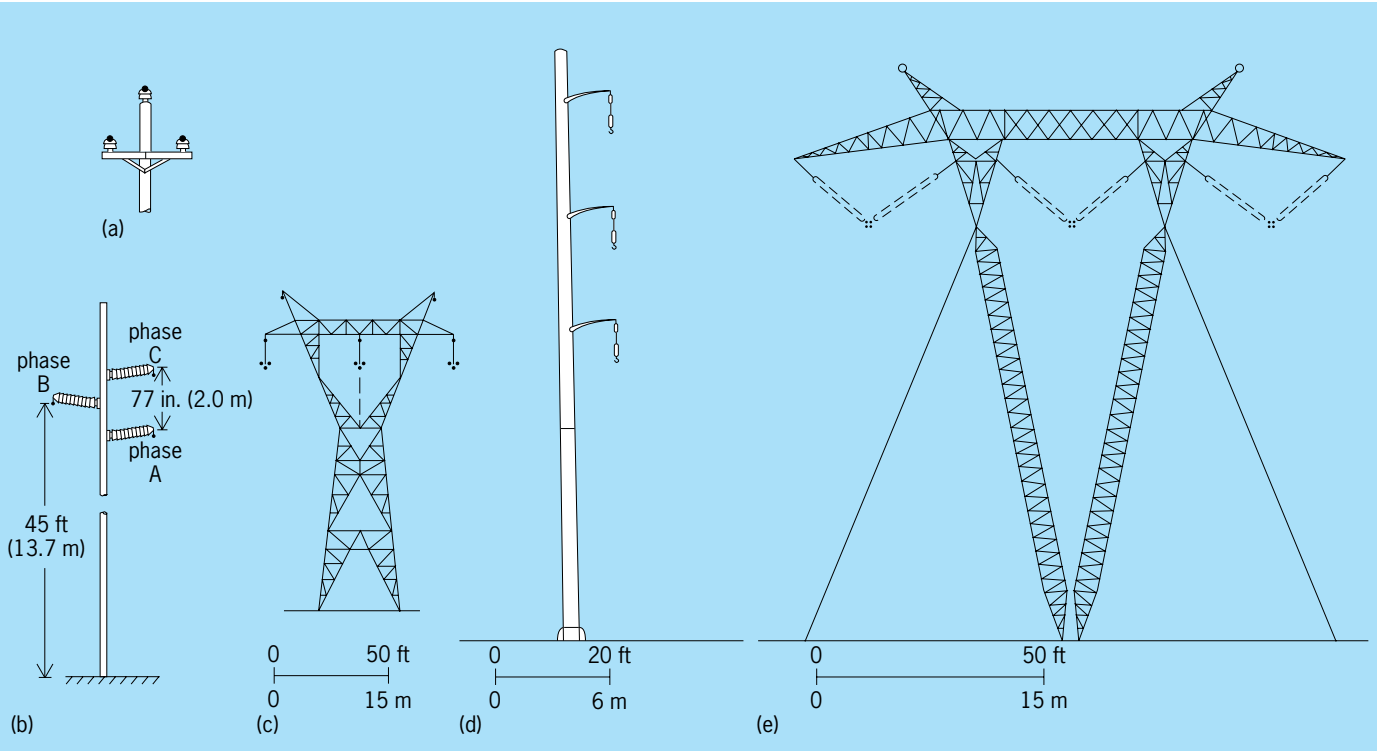

Fig. 3. Transmission-line towers and poles. (*a*) Distribution wood pole with porcelain post insulators. (*b*) Distribution wood pole with composite post insulators. (*c*) Steel tower with lattice structure and insulator chains. (*d*) Tapered, steel-tube tower with insulator chains. (*e*) Guyed V aluminum tower. (*After D. G. Fink and H. W. Beaty, eds., Standard Handbook for Electrical Engineers, 13th ed., Mc Graw-Hill, 1993*)

towers (Fig. 3*c*). Other high-voltage lines built with tubular steel poles (Fig. 3*d*) require less right-of-way and have a more pleasing appearance. Some extrahigh-voltage lines use guided towers in a V arrangement, supported by steel wires (Fig. 3*e*).

A transmission line is divided into sections, each terminated by a stronger dead-end tower with nearly horizontal insulators. From 10 to 25 suspension or tangent towers with vertical insulators are installed between the two dead-end towers, which have only transverse and vertical strength. Dividing the line assures that mechanical failures such as the breakage of conductors will affect only a small section of the line between dead ends. Stronger angle towers are used when lines change direction.

Mechanical performance. Transmission lines are subject to environmental adversities, including wide variations of temperature, high winds, and ice and snow deposits. Typically designed to withstand environmental stresses occurring once every 50–100 years, lines are intended to operate safely in adverse conditions.

Tension and sag. Conductor temperature varies with air temperature and power load, affecting sag and tension. Increased temperature reduces conductor tension and increases sag, reducing ground clearance (**Fig. 4**), while lower temperature increases tension of transmission lines.

Wind and ice loading. Extreme winter weather may coat conductors with ice, which increases conductor weight, as well as tension and sag. High wind produces a horizontal force on the conductor that increases tension and swings the conductor to an inclined position, reducing clearance between the conductor and the tower. Fortunately, high wind, ice loading, and minimum temperature do not occur simultaneously. Combined mechanical loads should not produce tension exceeding the conductor's tensile yield strength, and clearance should be greater than the minimum safety clearance.

Vibration. Nonuniform wind pressure produces conductor vibration. This is a low-amplitude, audio-frequency mechanical oscillation that bends the conductor at the point where it is clamped to the insulator. Stress from repeated bending may cause fatigue breakage of individual wires in the stranded conductor. Ultimately, complete breakage of the conductor may result in service interruptions. Dampers with weights are installed a few feet from the insulators to eliminate vibration-caused conductor breakage.

Galloping. The sudden drop of ice loading in a span causes galloping, a low-frequency, large-amplitude, mechanical oscillation that generates high mechanical stress and reduces the distance between phases in vertical conductors. The danger of galloping can be reduced by melting ice with high currents, using a low-voltage auxiliary transformer to drive the current through a short-circuited line.

Electrical performance. Variable weather affects line operation. Extreme weather reduces corona inception voltage, leading to an increase in audible noise, radio noise, and telephone interference. Load

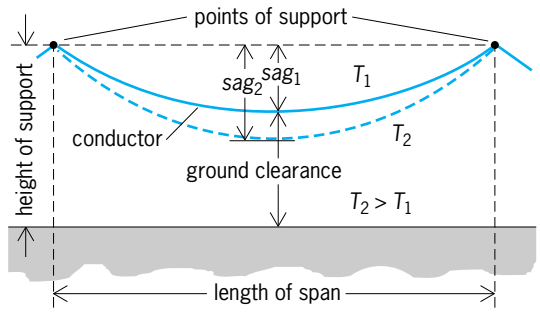

Fig. 4. Catenary curve showing the effect of temperature increase from T_1 to T_2.

variation requires regulation of line voltage. A short circuit generates large currents, overheating conductors and producing permanent damage.

Line parameters. Transmission-line parameters are conductor resistance, inductance, and capacitance. Each phase has different inductance if the distance between conductors is not equal, and different capacitance if conductor heights are different. Longer lines are transposed by cyclic changes of conductor positions, permitting the use of average inductance and capacitance.

Line inductance and capacitance of a single-circuit, three-phase line can be calculated from the distances between the phases and the radius and geometry of the phase conductors. Conductor alternating-current resistances can be obtained from tables. Alternating-current resistance differs slightly from direct-current resistance because of the skin effect, and is also dependent on temperature. Use of a bundle conductor reduces both line reactance and resistance and increases shunt capacitance. *See* CAPACITANCE; ELECTRICAL RESISTANCE; INDUCTANCE; REACTANCE; SKIN EFFECT (ELECTRICITY).

Equivalent circuit. The parameters (inductance, resistance, and capacitance) are distributed along the line. For calculation of the relationship between sending- and receiving-end voltages and currents, distributed parameters are concentrated to an equivalent resistance, inductance, and capacitance (**Fig. 5**). Equivalent values for a short line are calculated by multiplication with the line length; for a long line, more complicated equations are used. An equivalent circuit represents only one phase of a three-phase system; thus, one-third of the three-phase power flows from the sending end (supply) to the receiving end (load). Voltages in this single-phase equivalent circuit are line-to-ground. The relationship between

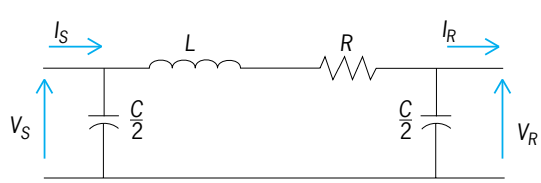

Fig. 5. Single-phase equivalent circuit for a transmission line. Line capacitance (C) is divided into two equal parts. V_S, I_S = voltage and current at sending end; V_R, I_R = voltage and current at receiving end.

sending- and receiving-end voltages and currents can be calculated by the Kirchhoff equations. *See* KIRCHHOFF'S LAWS OF ELECTRIC CIRCUITS.

Typically, in no-load or open-circuit conditions, receiving-end voltage is higher than sending-end voltage by as much as 5–25%, and can produce insulator flashover, voltage transformer fault, and so forth. The receiving-end voltage may be limited by shunt inductance, which loads the line. This inductance may be switched off when the load increases.

Increased load produces a voltage drop on the line and may reduce receiving-end voltage below sending-end voltage. This voltage drop can be compensated for by switched capacitors connected in series with the line. Capacitor reactance is negative and subtracted from line reactance, reducing the voltage drop and increasing receiving-end voltage for inductive load. Continuous control of series compensation can be achieved by shunting the series capacitor with a thyristor-controlled reactance. Delay of thyristor firing controls the time in which the inductance is connected in parallel with the capacitor in each half cycle, reducing the effect of capacitance. *See* SEMICONDUCTOR RECTIFIER.

Line power-transfer capacity. The power that a line can transport is limited by the line's electrical parameters. Voltage drop is the most important factor for distribution lines; where the line is supplied only from one end, the permitted voltage drop is about 5%.

The high-voltage line interconnects two buses in the electric system. Each of these buses is directly or indirectly connected to a voltage source. The amplitude and phase-angle difference between bus voltages regulates the direction and amount of power transferred through the line. The power transfer in a line with negligible losses is given by the equation below. Here, the bus voltages at the receiving and

$$P = \frac{V_S V_R}{X} \sin \delta$$

sending ends are V_R and V_S respectively; the phase angle between V_R and V_S is $\delta = \delta_S - \delta_R$, where δ_R and δ_S are the voltage phase angles at the receiving and sending ends; and the line reactance is X. Maximum power transfer occurs when $\delta = 90^\circ$. Stability of this operation requires that δ be not more than 40–60°. Maximum transferred power can be increased by reduction of series impedance by using series compensation. Voltage at both ends can be supported by shunt capacitors or static var compensators, which continuously generate and regulate reactive power. *See* STATIC VAR COMPENSATOR.

Thermal loading limit. Conductor temperature must be lower than the temperature which causes permanent elongation. A typical maximum steady-state value for ACSR is 212°F (100°C), but in an emergency temperatures 10–20% higher are allowed for a short period of time (10 min to 1 h). Steady-state temperature is calculated from the heat balance of the conductor. The sum of resistance (I^2R) losses and solar-radiation-generated heat is equal to the sum of convection- and radiation-produced heat losses. The line is typically designed for 104°F (40°C) ambient, 212°F (100°C) conductor temperature, and 2 ft/s (0.6 m/s) wind speed. Maximum current is reduced by increased altitude and affected by weather conditions. The typical design value for a 500-kV line with two conductors in a bundle is 2400 A.

Corona. This discharge is generated when the electric field at the surface of the conductor becomes larger than the breakdown strength of the air. This high electric field accelerates free electrons to high speed, and the collision of these electrons with air molecules produces ionization. Ionization generates visible light, and fast ion movement produces losses and an audible hissing noise. The oscillatory nature of the discharge generates high-frequency, short-duration current pulses, the source of corona-generated radio and television interference. Because the electric field decreases rapidly with distance, the corona discharge occurs only in the vicinity of the conductor. Surface irregularities such as water droplets cause local field concentration, enhancing corona generation. Thus, during bad weather, corona discharge is more intense and losses are much greater. *See* ELECTRICAL INTERFERENCE.

Corona discharge generates audible noise with two components: a broad-band, high-frequency component, which produces crackling and hissing, and a 120-Hz pure tone. The loudest noise is generated during rain, but low-level noise during fog or wet conditions is equally disturbing; fair-weather noise is inconsequential. *See* CORONA DISCHARGE.

Electric field. Transmission-line conductors are surrounded by an electric field which decreases as distance from the line increases, and depends on line voltage and geometry. Near the conductors, this field generates corona discharge. At ground level, it induces current and voltage in grounded bodies, causes corona in grounded objects, and can induce fuel ignition. Current induced in the human body results in unpleasant sensations such as hair stimulation, and the field produces a secondary spark discharge when an ungrounded person touches a grounded object under the line. Utilities limit the electric field at the perimeter of right-of-ways to about 1000 V/m. *See* ELECTRIC FIELD.

Magnetic field. Load current produces an ac magnetic field around the transmission line that decreases with distance from the line. For example, the maximum magnetic field at ground level under a line which carries 2000 A is 176 milligauss (17.6 microtesla), and decreases to 9.6 milligauss (0.96 microtesla) at 330 ft (100 m) from the line. The field depends on line geometry and current, but is independent of line voltage. The Earth produces a dc magnetic field around 500 mG (50 μT).

Lightning protection. Lightning strikes produce high voltages and traveling waves on transmission lines, causing insulator flashovers and interruption of operation. Steel grounded shield conductors at the tops of the towers (Fig. 1) significantly reduce, but do not eliminate, the probability of direct lightning strikes to phase conductors. The ground resistance of the towers is important. A direct lightning strike to the shield wires or tower drives high current

to the ground. The ohmic voltage drop on the tower's ground resistance increases the tower's potential and may cause insulator flashover, referred to as backflash. Reduction of tower footing resistance reduces the probability of backflash. Proper line design results in 0.5–1 fault per year per 100 mi (160 km). *See* LIGHTNING AND SURGE PROTECTION.

Switching surges. The operation of circuit breakers causes switching surges that can result in interruption of inductive current, energization of lines with trapped charges, and single-phase ground fault. A switching surge is a unidirectional or oscillatory impulse with a 250–500-microsecond rise time and 2–10-μs time to half value. These surges can have peak voltages 2–3.5 times the peak value of the 60-Hz voltage. A switching surge is most dangerous in extrahigh voltage lines (500 kV and above). Modern circuit breakers, operating in two steps, reduce switching surges to 1.5–2 times the 60-Hz voltage. First, the breaker inserts a resistance into the circuit. Then the breaker interrupts the already reduced current, or shorts the resistance and completes the line energization. *See* CIRCUIT BREAKER.

Telephone interference. Line current induces a disturbing voltage in telephone lines running parallel to transmission lines. Because the induced voltage depends on the mutual inductance between the two lines, disturbance can be reduced by increasing the distance between the lines and shielding the telephone lines. Telephone interference has also been reduced by the use of fiber optics and digital telephone techniques. A short circuit may induce dangerous voltages in neighboring telephone lines. *See* ELECTRICAL SHIELDING; INDUCTIVE COORDINATION.

Underground and Submarine Power Transmission

Most cities use underground cables to distribute electrical energy. These cables virtually eliminate negative environmental effects and reduce electrocution hazards. However, they entail significantly higher construction costs, typically 6–10 times those of transmission lines with similar capacity for medium-voltage distribution, and about 20 times higher for high-voltage transmission. High-voltage power cables are also used to transmit energy through wide deep rivers and bays where building transmission lines is difficult.

Cable construction. Underground cables are divided into two categories: distribution cables (less than 69 kV) and high-voltage power-transmission cables (69–500 kV).

Distribution cables. Extruded solid dielectric cables dominate in the 15–33-kV urban distribution system. In a typical arrangement (**Fig. 6***a*), the stranded copper or aluminum conductor is shielded by a semiconductor layer, which reduces the electric stress on the conductor's surface. Cross-linked polyethylene (XLPE) or ethylene-propylene-rubber (EPR) is used most frequently for solid dielectric insulation. The insulation is protected by a semiconducting metallic tape shield which terminates the electric field. A poly(vinyl chloride) [PVC] jacket provides mechanical protection and prevents moisture penetration.

The major problem with these cables is trees generated by electric fields and water penetration. Trees are channels in the dielectric produced by electric discharges. In most cases, they start at the conductor and progress slowly through the insulation, ultimately producing dielectric breakdown and cable failure.

Oil-impregnated paper-insulated cables are used for higher voltages and in older installations (Fig. 6*b*). The stranded copper conductor is shielded by a metallized paper tape, which alleviates the unevenness caused by stranding. The conductors are insulated with oil-impregnated paper tape. Each insulated conductor is covered by metallized paper tape, which produces a concentric electric field. Oil-impregnated paper fillers are used to create a round cable assembly, held together by a metallized paper binder. The assembly in protected from moisture aggression by a lead or lead-alloy shield. Cables designed for insulation in a duct are protected by an extruded PVC sheet. Cables designed for direct burial are covered by a fibrous material (jute) impregnated by bituminous compound and double steel-tape armoring.

High-voltage power-transmission cables. Cable temperatures vary with load changes, and cyclic thermal expansion and contraction may produce voids in the cable. High voltage initiates corona in the voids, gradually destroying cable insulation. Low-pressure oil-filled cable construction reduces void formation. A single-phase concentric cable has a hollow conductor with a central oil channel. Three-phase cables have three oil channels located in the filler. Channels are supplied by low-pressure oil which penetrates the paper insulation and fills the voids

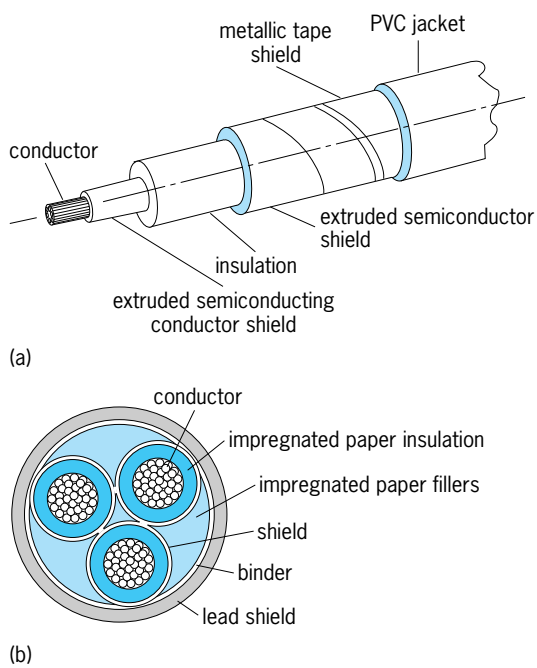

Fig. 6. Underground distribution cables. (*a*) Extruded solid dielectric cable. (*b*) Three-phase oil-impregnated paper-insulated cable. (*After R. Bartnikas Srivastava, Elements of Cable Engineering, Sandford Educational Press, 1980*)

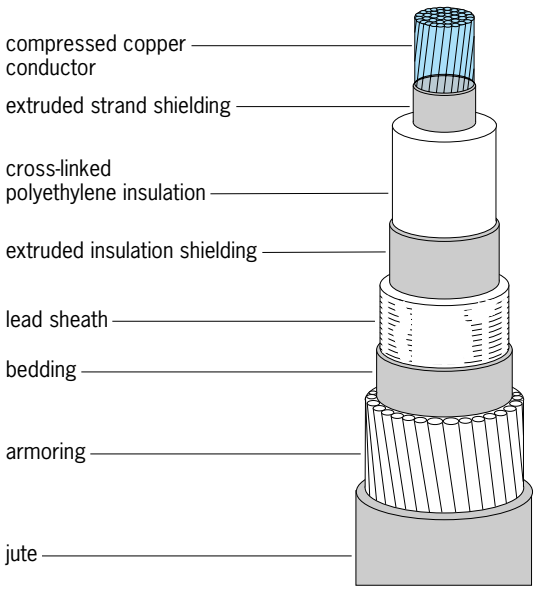

Fig. 7. Solid-dielectric submarine cable. (*ASEA Kabel*)

produced by thermal expansion. Oil is degassed before installation, and the system is hermetically sealed to avoid gas penetration during operation. These cables need complicated joints and oil tanks alongside them, which can cause difficulties. They are used between 69 and 500 kV, although 750-kV cable has been developed.

For extrahigh voltage, high-pressure oil-filled pipe cables are used. Three single-conductor shielded cables are installed in a steel pipe. The pipe is filled with high-pressure (200-lb/in.^{2} or 1.4-megapascal) low-viscosity oil that penetrates the insulation and prevents corona generation. Oil pressure and temperature are maintained by automatically operated pumping stations.

Submarine cables. High-voltage cables are frequently used for crossing large bodies of water. Water provides natural cooling, and pressure reduces the possibility of void formation. A typical submarine cable (**Fig. 7**) has cross-linked polyethylene insulation, and corrosion-resistant aluminum alloy wire armoring that provides tensile strength and permits installation in deep water.

Cable parameters. Cables have significantly larger phase-to-ground capacitance and smaller inductances than overhead lines. Large capacitance results in high charging current, which reduces the power transfer capacity. At critical cable length, charging current is equal to load current. Typical critical length is in the range of 15 mi (25 km) at 500 kV and 45 mi (70 km) at 138 kV. Charging current can be compensated by shunt reactors or static var compensators. Because of the small inductance, reactive components of the voltage drop dominate in a cable system, and the voltage drop is significantly less than in an overhead line.

Cable power-transfer capability is dependent on the temperature of the dielectric, which depends on the heat generated by resistive power loss and on the thermal resistance. In a three-phase system, the unbalanced load current induces a circulating current in the cable shield, which heats the cable and reduces power-transfer capacity.

Cable power-transfer capacity can be improved by externally cooling the cables. One method used in large substations is installing pipes cooled by circulating water near the cables.

Installation. In urban areas, cables are installed in underground concrete ducts requiring access holes every 1000–1500 ft (300–500 m). The ducts facilitate location of faulted cables and provide protection against accidental cutting of the cable during road construction.

In suburban areas, cables may be buried in trenches. Concrete slabs or bricks are often placed on top of the cable for protection and marking. The disadvantages of this method are difficulty in fault location and a lack of mechanical protection.

Cables are manufactured in sections 1000–1500 ft (300–500 m) long, with joints connecting the sections for urban cable. These joints, which are located in the access holes, frequently incur faults, and their preparation requires highly skilled workers. Bending and pulling may damage cables during installation, and proper installation requires expensive equipment and skilled labor. *See* ELECTRIC POWER SYSTEMS; TRANSMISSION LINES.

George G. Karady

Bibliography. Electric Power Research Institute, *Transmission Line Reference Book: 345 kV and Above*, EL 2500, 1987; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; J. A. Williams and P. L. Ostermann, *Underground Transmission Systems Reference Book*, Electrical Power Research Institute, 1992.

Electric protective devices

A particular type of equipment applied to electric power systems to detect abnormal and intolerable conditions and to initiate appropriate corrective actions. These devices include lightning arresters, surge protectors, fuses, and relays with associated circuit breakers, reclosers, and so forth.

From time to time, disturbances in the normal operation of a power system occur. These may be caused by natural phenomena, such as lightning, wind, or snow; by falling objects such as trees; by animal contacts or chewing; by accidental means traceable to reckless drivers, inadvertent acts by plant maintenance personnel, or other acts of humans; or by conditions produced in the system itself, such as switching surges, load swings, or equipment failures. Protective devices must therefore be installed on power systems to ensure continuity of electrical service, to limit injury to people, and to limit damage to equipment when problem situations develop.

Protective devices, like any type of insurance, are applied commensurately with the degree of protection desired or felt necessary for the particular system. Thus, application of these devices varies widely.

Protective relays. Protective relays are compact analog or digital networks, connected to various points of an electrical system, to detect abnormal conditions occurring within their assigned areas. They initiate disconnection of the trouble area by circuit breakers. These relays range from the simple overload unit on house circuit breakers to complex systems used to protect extrahigh-voltage power transmission lines. They operate on voltage, current, current direction, power factor, power, impedance, temperature, and so forth, as well as combinations of these. In all cases there must be a measurable difference between the normal or tolerable operation and the intolerable or unwanted condition. System faults for which the relays respond are generally short circuits between the phase conductors, or between the phases and grounds. Some relays operate on unbalances between the phases, such as an open or reversed phase. A fault in one part of the system affects all other parts. Therefore relays and fuses throughout the power system must be coordinated to ensure the best quality of service to the loads and to avoid operation in the nonfaulted areas unless the trouble is not adequately cleared in a specified time. *See* FUSE (ELECTRICITY); RELAY.

Zone protection. For the purpose of applying protection, the electric power system is divided into five major protection zones: generators; transformers; buses; transmission and distribution lines; and motors (**Fig. 1**). Each block represents a set of protective relays and associated equipment selected to initiate correction or isolation of that area for all anticipated intolerable conditions or trouble. The detection is done by protective relays with a circuit breaker used to physically disconnect the equipment. For other areas of protection *see* GROUNDING; UNINTERRUPTIBLE POWER SUPPLY.

Fault detection. Fault detection is accomplished by a number of techniques. Some of the common methods are the detection of changes in electric current or voltage levels, power direction, ratio of voltage to current, temperature, and comparison of the electrical quantities flowing into a protected area with the quantities flowing out. The last-mentioned is known as differential protection.

Differential protection. This is the most fundamental and widely used protection technique (**Fig. 2**). The inputs to the relays are currents from a current transformer. Current through the relay (nominally 5 or less amperes) is proportional to the high-power current of the main circuit. For load through the equipment or for faults either to the right or left of the current transformers, the secondary currents all sum to essentially zero (Fig. 2*a*). However, for internal trouble, they add up to flow through the relay (Fig. 2*b*).

A typical three-line diagram for a two-circuit differential system is shown in **Fig. 3**. For increased sensitivity, differential relays have restraint and operating windings. On through conditions (Fig. 2*a*) the restraint winding currents desensitize the operating winding, while full sensitivity is available for the internal faults (Fig. 2*b*). Thus for a fault between the current transformers, secondary current flows through the restraining windings (with reversed phase in one of the coils) and in the operating coils to operate the relay, trip the two circuit breakers, and isolate the fault or damaged equipment from the rest of the power system.

The equipment to be protected can be a generator, transformer, bus, motor, or line. In the last case, with many miles between the circuit breakers, telephone-type low-energy pilot wires, optical fibers, radio frequency on the power line, or a microwave channel is used. The quantities at the line terminals are thus differentially compared via the channel, in what is known as pilot relaying.

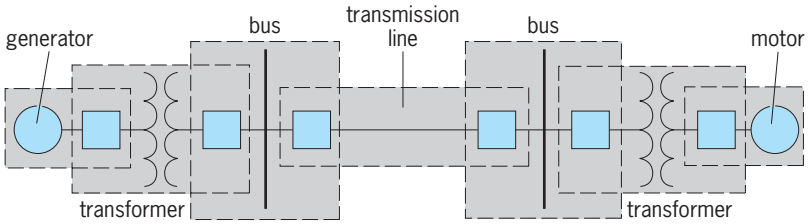

Fig. 1. Zones of protection on simple power system.

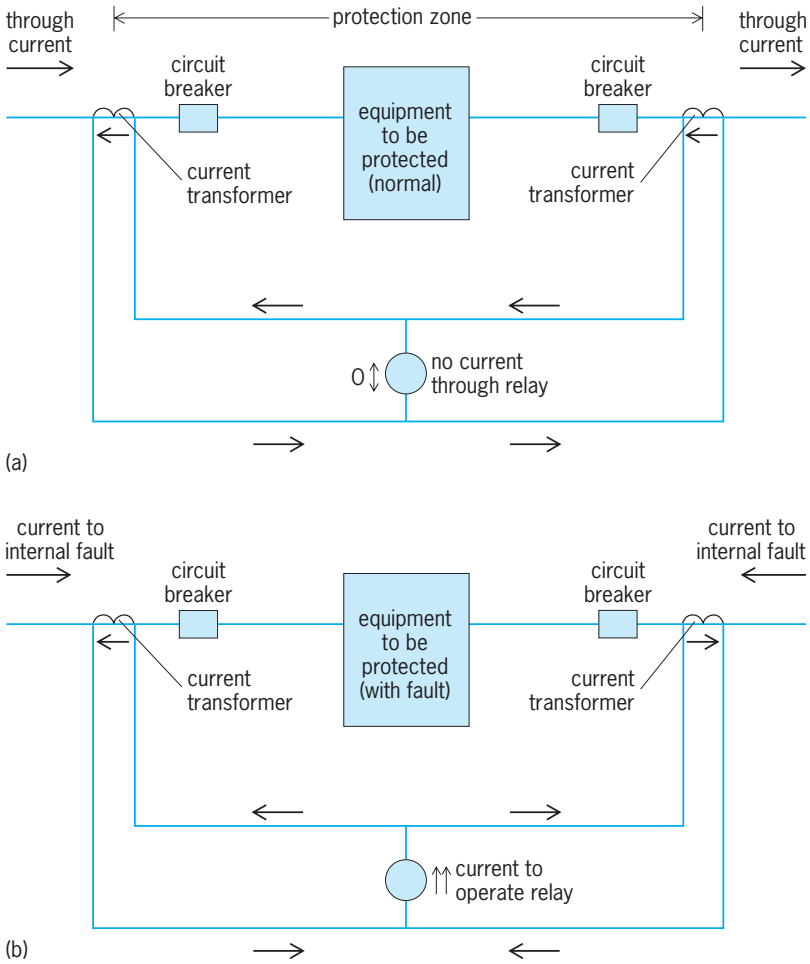

Fig. 2. Simple differential protection scheme. The system compares currents to detect faults in the protection zone. Current transformers reduce the primary current to a small secondary value. (*a*) Normal operation. (*b*) With fault.

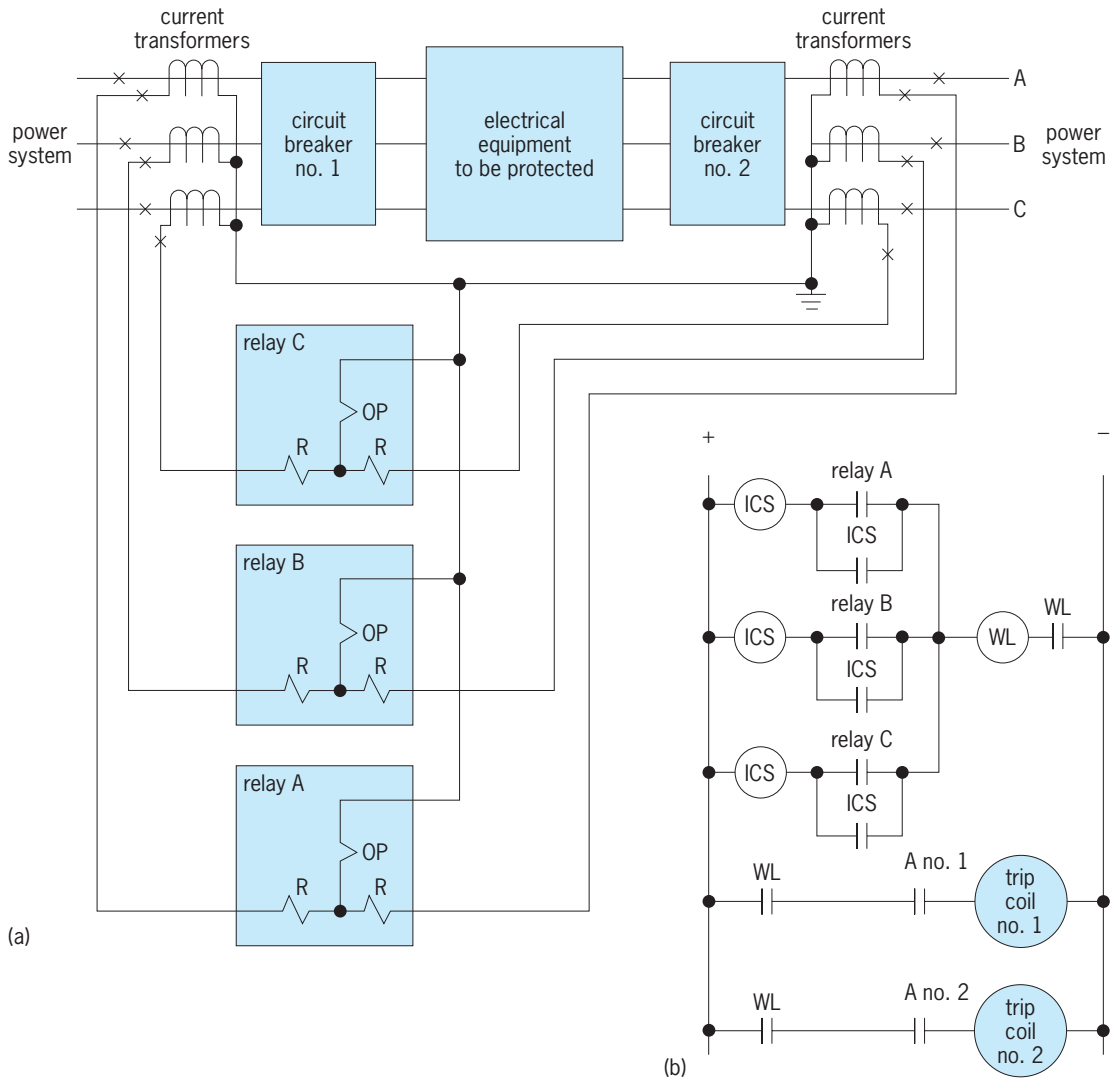

Fig. 3. Differential protection. (*a*) Typical arrangement for electrical equipment. Relays A, B, and C are differential relays. (*b*) Trip circuit. A = auxiliary out-off breaker; ICS = operation indicator and circuit seal-in auxiliary relay; OP = operating coil; R = restraining coil; WL = auxiliary hand-reset lock-out relay.

Overcurrent protection. This must be provided on all systems to prevent abnormally high currents from overheating and causing mechanical stress on equipment. Overcurrent in a power system usually indicates that current is being diverted from its normal path by a short circuit. In low-voltage, distribution-type circuits, such as those found in homes, adequate overcurrent protection can be provided by fuses that melt when current exceeds a predetermined value.

Small thermal-type circuit breakers also provide overcurrent protection for this class of circuit. As the size of circuits and systems increases, the problems associated with interruption of large fault currents dictate the use of power circuit breakers. Normally these breakers are not equipped with elements to sense fault conditions, and therefore overcurrent relays are applied to measure the current continuously. When the current has reached a predetermined value, the relay contacts close. This actuates the trip circuit of a particular breaker, causing it to open and thus isolate the fault. *See* CIRCUIT BREAKER.

Overcurrent relays are of two basic types: instantaneous and inverse-time. Instantaneous units have no intentional time delay and operate in the order of 10 to 50 ms. Inverse-time units operate quickly for high currents and with increasing time for smaller currents.

Overcurrent protection is applied in all parts of the power system. A typical application for the protection of a line between two stations R and S is shown in **Fig. 4***a*. The overcurrent relays are connected to the current transformers to the left of breaker E and to the right of breaker F to operate for all faults in the primary protection zone between the two stations R and S.

If there is a significant magnitude difference between the close-in fault F_1 and the far bus fault F_2, instantaneous overcurrent relays are set to operate for faults F_1 out to n_1 but not to operate for fault F_2. The operating zone is indicated by the horizontal line from the relays at breaker E to n_1, and for the relays at breaker F to n_2. These lines represent the

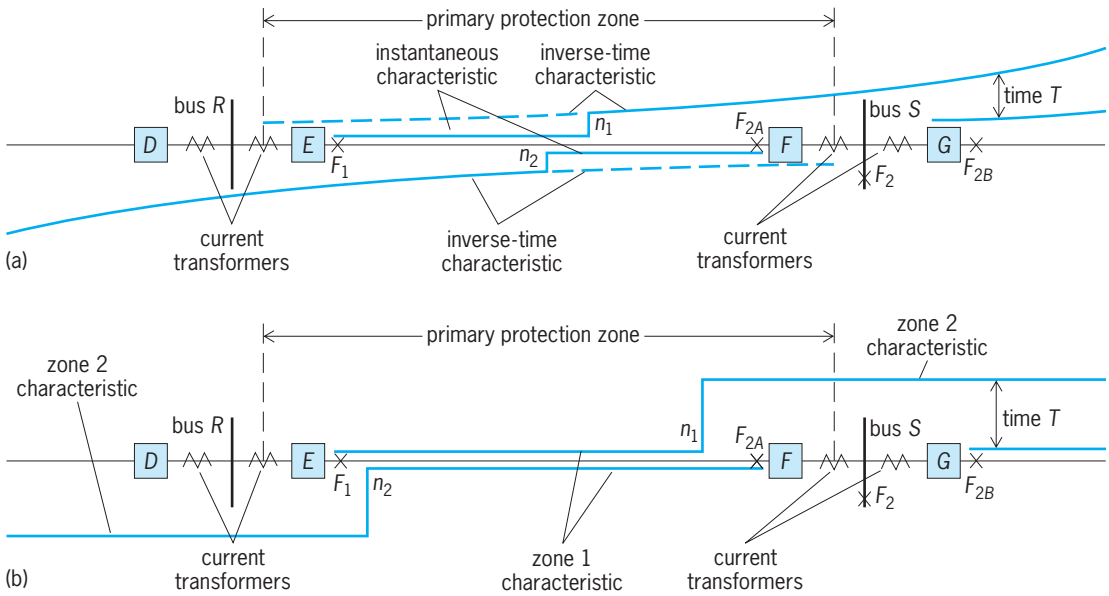

Fig. 4. Typical relay applications and settings. (*a*) Instantaneous and inverse-time overcurrent relays *D*, *E*, *F*, and *G* are circuit breakers. For the relay characteristics shown, operating times are shown vertically. Instantaneous relay time is instantaneous from the relay current transformers to points n_1 and n_2. (*b*) Distance relay. Zone 1 relay characteristic is instantaneous. Zone 2 relay characteristic is instantaneous with operation delayed by time *T*.

maximum coverage. When the fault level decreases, less of the line will be protected.

The current magnitude through the relays at breaker E for a fault at either F_2, F_{2A}, or F_{2B} is the same. Thus it is not possible for the relays at breaker E to determine if fault F_2 is at F_{2A} for which it should operate, or at F_{2B} where it should be cleared by the relays at breaker G. Inverse-time overcurrent relays with directional sensing into the line are applied at both breakers E and G. The relays at E are set to operate as fast as possible for faults at F_1 but to delay for a time interval T for the F_2 faults. Unfortunately this arrangement delays tripping for the internal fault F_{2A}, but it does provide time for the relays and breaker at G to clear the faults F_{2B} and faults to the right of G. If, however, either the relays or breaker G fails to operate for faults F_{2B}, then in time T breaker E will be opened to remove the fault current from the left side. This is backup protection by the relays and breaker at E for faults on the line to the right of breaker G. This process is known as relay coordination or selective settings.

Distance protection. Distance-type relays operate on the combination of reduced voltage and increased current occasioned by faults. They are widely applied for the protection of higher voltage lines. An application is illustrated in Fig. 4*b*. A major advantage is that the operating zone is determined by the line impedance and is almost completely independent of current magnitudes. Typical instantaneous settings n_1 and n_2 are 90% of the total line impedance.

The same difficulty of distinguishing internal faults F_{2A} from external faults F_{2B} at breaker E exists, and so a second distance unit is applied. It is set well into the adjacent lines as shown and is operated through a timer T. This permits the G relays and breaker an opportunity to clear faults F_{2B} and faults to the right before E relays and breaker can operate as backup.

Overvoltage protection. Lightning in the area near the power lines can cause very short-time overvoltages in the system and possible breakdown of the insulation. Protection for these surges consists of lightning arresters connected between the lines and ground. Normally the insulation through these arresters prevents current flow, but they momentarily pass current during the high-voltage transient to limit overvoltage. Overvoltage protection is seldom applied elsewhere except at the generators, where it is part of the voltage regulator and control system. In the distribution system, overvoltage relays are used to control taps of tap-changing transformers or to switch shunt capacitors on and off the circuits. *See* LIGHTNING AND SURGE PROTECTION.

Undervoltage protection. This must be provided on circuits supplying power to motor loads. Low-voltage conditions cause motors to draw excessive currents, which can damage the motors. If a low-voltage condition develops while the motor is running, the relay senses this condition and removes the motor from service.

Undervoltage relays can also be used effectively prior to starting large induction or synchronous motors. These types of motors will not reach their rated speeds if started under a low-voltage condition. Relays can be applied to measure terminal voltage, and if it is below a predetermined value the relay prevents starting of the motor.

Underfrequency protection. A loss or deficiency in the generation supply, the transmission lines, or other components of the system, resulting primarily from faults, can leave the system with an excess of load. Solid-state and digital-type underfrequency relays are connected at various points in the system to detect this resulting decline in the normal system

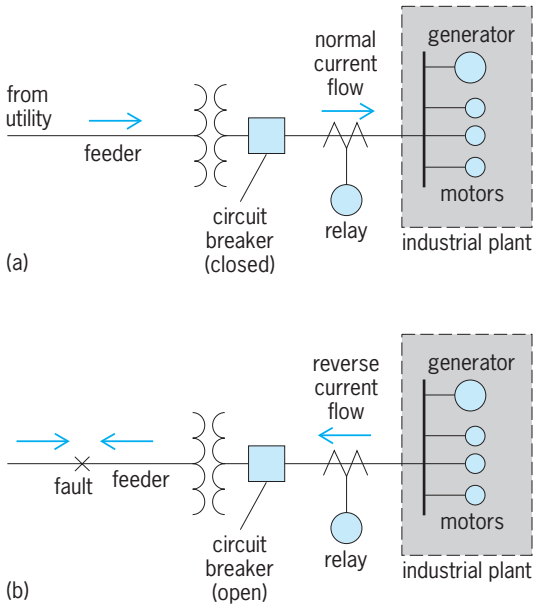

Fig. 5. Reverse-current protection. (*a*) Normal load conditions. (*b*) Internal fault condition; relay trips circuit breaker under reverse-current condition.

frequency. They operate to disconnect loads or to separate the system into areas so that the available generation equals the load until a balance is reestablished.

Reverse-current protection. This is provided when a change in the normal direction of current indicates an abnormal condition in the system. In an ac circuit, reverse current implies a phase shift of the current of nearly 180° from normal. This is actually a change in direction of power flow and can be directed by ac directional relays.

A common application of reverse-current protection is shown in **Fig. 5**. In this example, a utility supplies power to an industrial plant having some generation of its own. Under normal conditions, current flows from the utility to the plant (Fig. 5*a*). In the event of a fault occurring on the utility feeder (Fig. 5*b*) the current reverses direction and flows from the plant to the fault location. The relay operates and trips the circuit breaker, isolating the plant from the utility, thus preventing an excessive burden on the plant generator. Usually in these cases, the plant generator cannot carry the plant load, so that underfrequency relays are used to shed noncritical load. When the utility tie is restored, the shed loads then can be reconnected for full plant service.

Phase unbalance protection. This protection is used on feeders supplying motors where there is a possibility of one phase opening as a result of a fuse failure or a connector failure. One type of relay compares the current in one phase against the currents in the other phases. When the unbalance becomes too great, the relay operates. Another type monitors the three-phase bus voltages for unbalance. Reverse phases will operate this relay.

Reverse-phase-rotation protection. Where direction of rotation is important, electric motors must be protected against phase reversal. A reverse-phase-rotation relay is applied to sense the phase rotation. This relay is a miniature three-phase motor with the same desired direction of rotation as the motor it is protecting. If the direction of rotation is correct, the relay will let the motor start. If incorrect, the sensing relay will prevent the motor starter from operating.

Thermal protection. Motors and generators are particularly subject to overheating due to overloading and mechanical friction. Excessive temperatures lead to deterioration of insulation and increased losses within the machine. Temperature-sensitive elements, located inside the machine, form part of a bridge circuit used to supply current to a relay. When a predetermined temperature is reached, the relay operates, initiating opening of a circuit breaker or sounding of an alarm. J. Lewis Blackburn

Bibliography. J. L. Blackburn, *Protective Relaying: Principles and Applications*, 2d ed., 1997; W. A. Elmore (ed.), *Protective Relaying: Theories and Applications*, 1994; S. H. Horowitz, *Protective Relaying for Power Systems*, 2 vols., 1981, 1992; Institute of Electrical and Electronics Engineers, *IEEE Guides and Standards for Protective Relaying Systems*, rev. ed., 1991; National Association of Relay Manufacturers, *Engineers Relay Handbook*, 4th ed., 1992.

Electric rotating machinery

Any form of apparatus, having a rotating member, which generates, converts, transforms, or modifies electric power. Essentially all of the world's electric power is produced by rotating electrical generators, and about 70% of this energy is consumed in driving electric motors. Therefore, in a large sense, modern technology is built on the foundation of electric machinery. Electric machines are electromechanical energy converters; generators convert mechanical energy into electrical energy and motors convert electrical energy into mechanical energy. The basic laws of physics on which electric machinery is based were developed in the late nineteenth century.

Physical laws. It is possible to construct electric machines that are based on the energy stored in electric or magnetic fields. Those using electric fields are not efficient except in very small sizes, and as a result most machines are based on magnetic attraction or repulsion. The operation of electric machines can be explained in terms of two laws of physics: Ampère's force law and Faraday's law of induction. When a conductor is placed in a magnetic field, it experiences a force **F** that is proportional to the current i and to the strength of the magnetic field **B**. This force is in a direction at right angles to both the current and the field. This is known as Ampère's force law and can be expressed mathematically as Eq. (1), where **l** is the length of the conductor. Ampère's law has numerous equivalent statements, from any of which the production of torque

$$\mathbf{F} = i\mathbf{l} \times \mathbf{B} \tag{1}$$

can be explained. The second law is Faraday's law, which states that an electromotive force **E** will be induced in a circuit when that circuit experiences a change in flux linkage λ. The magnitude of **E** is proportional to the time rate of change of flux linkage. The electromotive force will be directed so that it produces currents that will oppose the change of flux linkage. Faraday's law can be written as Eq. (2).

$$\mathbf{E} = \frac{d\lambda}{dt} \tag{2}$$

See ELECTRIC FIELD; ELECTROMAGNETIC INDUCTION.

Principles of operation. An electric machine can be constructed on the principle that a magnet will attract a piece of permeable magnetic material such as iron or magnetic steel. In **Fig. 1***a*, a pole structure is shown along with a magnetic block that is allowed to rotate. The magnetic block will experience a torque tending to rotate it counterclockwise to the vertical direction. This torque, called a reluctance or saliency torque, will be in the direction to minimize the reluctance of the magnetic circuit. In Fig. 1*b*, a winding is added to the rotor (the part which is allowed to rotate). In this case there is an additional torque on the rotor in the counterclockwise direction produced by the attraction of opposite poles. This torque will be approximately proportional to the sine of the angle θ. While the magnets in Fig. 1 are electromagnets, permanent magnets could be used with the same effect. *See* MAGNET; WINDINGS IN ELECTRIC MACHINERY.

In these examples, if the rotor were allowed to move under the influence of the magnetic forces, it would eventually come to rest at an equilibrium position, $\theta = 0$. Since most applications require continuous motion and constant torque, it is necessary to keep the angle between the rotor magnetic field and the stator magnetic field constant. Thus, in the above examples, the stator magnetic field must rotate ahead of the rotor. In a polyphase machine this is accomplished by distributing the winding so as to produce a rotating magnetic field. In the case of a three-phase machine the windings are distributed

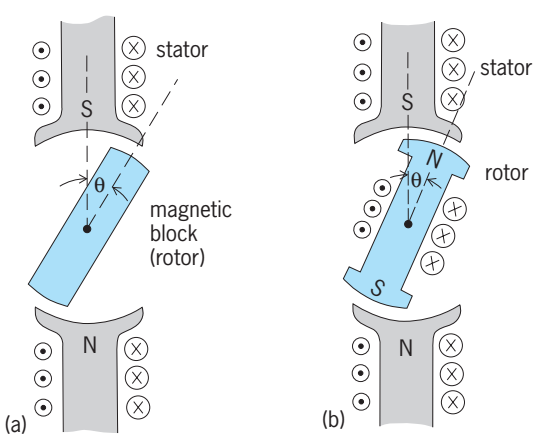

Fig. 1. Devices illustrating principles of electric machines. (*a*) Permeable rotor and stator with magnetic pole structure. (*b*) Device with magnetic pole structures on both stator and rotor.

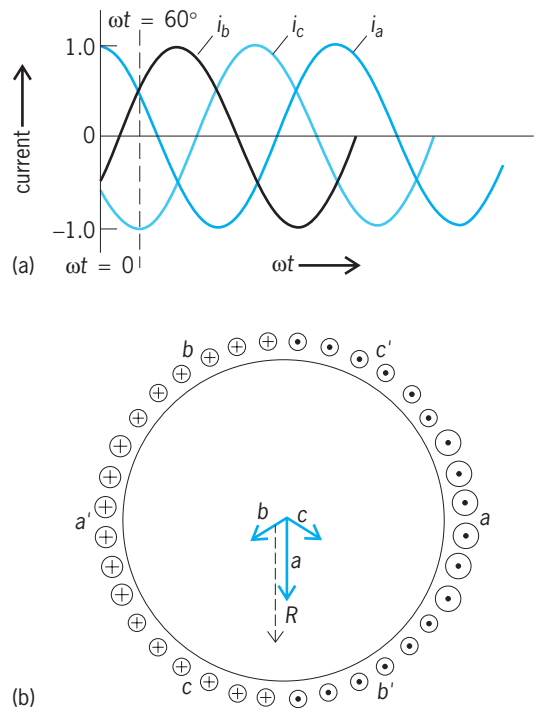

Fig. 2. Three-phase machine. (*a*) Three-phase balanced currents with angular frequency ω. (*b*) Distribution of currents and fields at $\omega t = 60°$.

in space so that their magnetic axes lie 120° apart. During normal operation the currents in the three windings (i_a, i_b and i_c in phases *a*, *b*, and *c*; **Fig. 2***a*) have the same magnitude but are 120° out of phase. (Such currents are called balanced currents.) At time $t = 0$ the current in phase *a* is maximum, and the currents in phases *b* and *c* are $-{}^1/_2$ of their peak values. The magnetic field *R* due to the combined three-phase system lies in the same direction that phase *a* current alone would produce (Fig. 2*b*). At a later time, $\omega t = 60°$, where ω is the angular frequency of the currents, the current in phase *c* is at its negative peak, while the currents in phases *a* and *b* are ${}^1/_2$ of their peak value. The resultant field is along the (negative) magnetic axis of phase *c*, the total distribution having rotated 60° in space. It can be demonstrated that the fundamental spatial component of the field has constant magnitude and rotates uniformly at angular frequency ω. In this way it is possible to keep a constant angle between the stator and rotor magnetic fields and achieve constant torque. *See* ALTERNATING CURRENT.

In single-phase machines, there is no starting torque, and a device such as a capacitor or a short-circuiting ring is used to simulate the effect of a rotating distribution at starting. The single-phase winding with sinusoidal current produces two rotating pole distributions, one rotating in the forward direction and one in the backward direction. In single-phase alternating-current machines, once the rotor is rotating (in either direction) there is a net magnetic torque in the direction of rotation for motor operation and opposite the direction of rotation for generator operation.

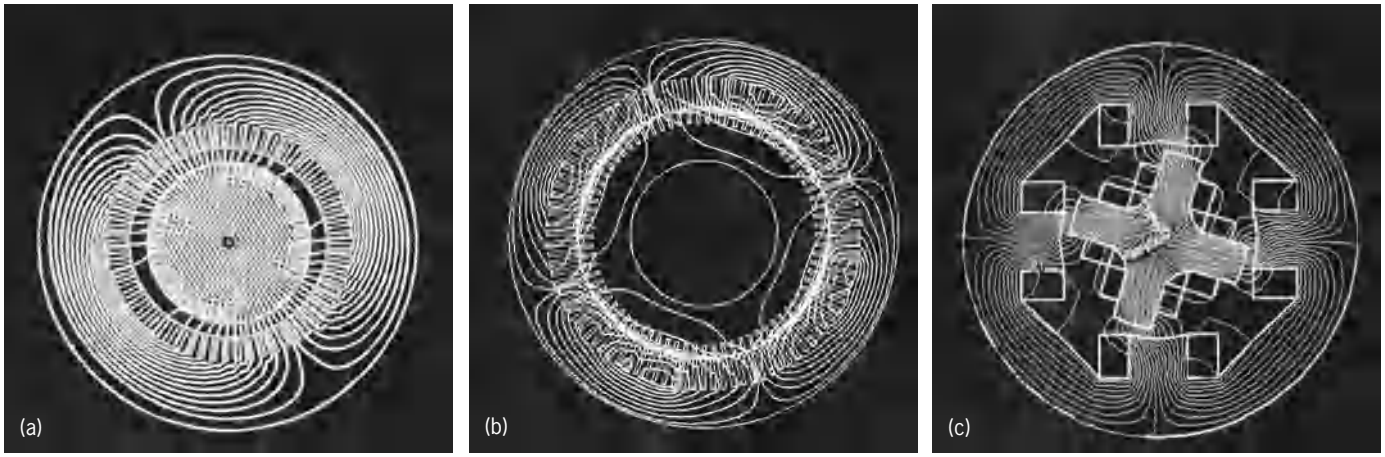

Fig. 3. Magnetic flux lines in rotating electric machines operating under load. (*a*) Three-phase, round-rotor synchronous generator (*Rensselaer Polytechnic Institute*). (*b*) Three-phase, squirrel-cage induction motor (*Rodolpho Palma*, *Rensselaer Polytechnic Institute*). (*c*) Direct-current motor (*MAGSOFT Corp.*).

Types. Although there are many variations, the three basic machine types are synchronous, induction, and direct-current machines. These machines may be used as motors or as generators, but the basic principles of operation remain the same.

Synchronous machines. The synchronous machine runs at a constant speed determined by the line frequency. There is an alternating-current winding (normally on the stator) and a direct-current winding (normally on the rotor). **Figure 3*a*** shows the magnetic field lines for a round-rotor synchronous generator operating under load. *See* ALTERNATING-CURRENT GENERATOR; ALTERNATING-CURRENT MOTOR; SYNCHRONOUS MOTOR.

Induction machines. The induction machine is another alternating-current machine which runs close to synchronous speed. The alternating-current winding of the stator is similar to that of the synchronous machine. The rotor may have an insulated winding (wound rotor) but more commonly consists of uninsulated bars embedded in a laminated structure and short-circuited at the end (squirrel cage). There is normally no voltage applied to the rotor. The voltages are produced by means of Faraday's law of induction. In an induction motor the stator-produced flux-density wave rotates slightly faster than the rotor during normal operation, and the flux linkages on the rotor therefore vary at low frequency. The rotor currents induced by these time-varying flux linkages produce a magnetic field distribution that rotates at the same speed as the stator-produced flux wave. The flux distribution of a squirrel-cage induction motor under load is shown in Fig. 3*b*. *See* INDUCTION MOTOR.

Direct-current machines. In a direct-current motor, direct current is applied to both the rotor and the stator. The stationary poles on the stator produce a stationary magnetic field distribution. Since the angle between the stator-produced poles and rotor-produced poles must remain constant, the direct-current machine uses a device known as a commutator which switches the current from one rotor circuit to another so that the resulting field is stationary. A direct-current motor operating under load is depicted in Fig. 3*c*. *See* DIRECT-CURRENT GENERATOR; DIRECT-CURRENT MOTOR.

Additional systems. Besides the magnetic system that produces the energy conversion in rotating electric machines, a number of other systems operate in these devices.

Electric circuits. The windings of the machines may be any combination of series or parallel circuits, possibly with multiple turns in each circuit using one or many phases. The geometrical layout of these circuits is an important consideration in the electrical design because it affects the efficiency and performance of the device. *See* CIRCUIT (ELECTRICITY).

Thermal system. Heat is generated in the windings because of Joule heating and in the magnetic circuit by eddy currents and hysteresis. When the rotor is turning, there are also frictional and aerodynamic losses. The heat associated with all of these losses must be removed. Ventilation of the machine to limit the temperature to an acceptable level is often the limiting factor in machine design. The most commonly used cooling medium is air. Large machines may use high-pressure hydrogen or purified water as a coolant. *See* AIR COOLING; CORE LOSS; EDDY CURRENT; ELECTRICAL RESISTANCE.

Insulation system. Machines vary in voltage levels from a fraction of a volt to many thousands of volts. A dielectric material is used to coat the windings and insulate them from each other and from other parts of the machine. Deterioration of insulation with age and with increased temperature is one factor that limits the lifetime of machines. Mechanical forces on insulation within windings and between windings and other parts are also a limit on the design. *See* DIELECTRIC MATERIALS; ELECTRICAL INSULATION.

Mechanical system. The rotor on a rotating machine must be free to turn. This means that a system of bearings (often with a lubricating system) must be provided. The rotor windings must be supported

against the centrifugal forces due to rotation. The electromagnetic torque that is applied to the rotor is also applied, in an opposite direction, to the stator, which must be suitably supported. If vibrations are not limited to acceptable levels, excessive audible noise or fatigue failures may occur. *See* ACOUSTIC NOISE; ANTIFRICTION BEARING; MECHANICAL VIBRATION; METAL, MECHANICAL PROPERTIES OF.

With these different modes of operation and design considerations, thousands of variations of motors and generators have been developed for industrial and domestic use. *See* GENERATOR; MOTOR.

Sheppard Salon

Bibliography. B. Chalmers and A. Williamson, *A. C. Machines, Electromagnetics and Design*, 1991; A. E. Fitzgerald, C. Kingsley, and S. D. Umans, *Electric Machinery*, 6th ed., 2003; I. L. Kosow, *Electric Machinery and Transformers*, 2d ed., 1991; S. A. Nasar and I. Boldea, *Electric Machines: Dynamics and Control*, 1992; M. G. Say, *Alternating Current Machines*, 5th ed., 1983; M. G. Say, *Direct Current Machines*, 2d ed., 1986.

Electric spark

A transient form of gaseous conduction. This type of discharge is difficult to define, and no universally accepted definition exists. It can perhaps best be thought of as the transition between two more or less stable forms of gaseous conduction. For example, the transitional breakdown which occurs in the transition from a glow to an arc discharge may be thought of as a spark.

Electric sparks play an important part in many physical effects. Usually these are harmful and undesirable effects, ranging from the gradual destruction of contacts in a conventional electrical switch to the large-scale havoc resulting from lightning discharges. Sometimes, however, the spark may be very useful. Examples are its function in the ignition system of an automobile, its use as an intense short-duration illumination source in high-speed photography, and its use as a source of excitation in spectroscopy. In the second case the spark may actually perform the function of the camera shutter, because its extinction renders the camera insensitive. *See* SPECTROSCOPY; STROBOSCOPIC PHOTOGRAPHY.

Mechanisms. The spark is probably the most complicated of all forms of gaseous conduction. It is exceedingly difficult to study, because it is a transient and because there are so many variables in the system. Some of these variables are the components of the gaseous medium, the gas pressure, the chemical form of the electrodes, the physical shape of the electrodes, the microscopic physical surface structure, the surface temperature, the electrode separation, the functional dependence of potential drop on time, and the presence or absence of external ionizing agents. One or more of these conditions may change from one spark to the next. Because of the great complexity, it will be impossible to do more than touch on some of the main features in this article.

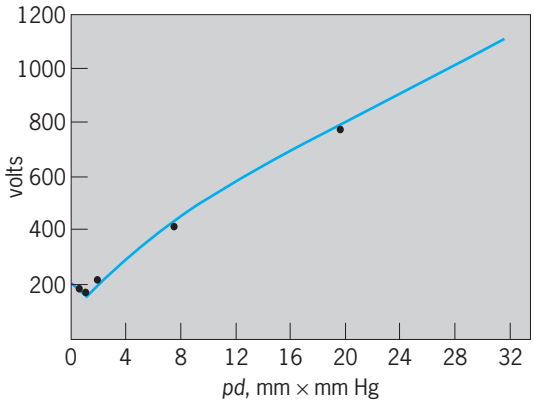

Fig. 1. Dependence of sparking potential on *pd* for a sodium cathode in hydrogen gas. (*After L. B. Loeb and J. M. Meek, The Mechanism of the Electric Spark, Stanford University Press, 1941*)

The dependence of breakdown, or sparking, potential on pressure p and electrode separation d is considered first. It was shown experimentally by F. Paschen and theoretically by J. S. Townsend that the sparking potential is a function of the product pd and not of p or d separately (**Fig. 1**). Further, there is a value of pd for which the sparking potential is a minimum. Thus, if it is desired to prevent sparking between two electrodes, the region may be evacuated or raised to a high pressure. The latter method is used in accelerators of the electrostatic generator variety. Here the entire apparatus is placed in a pressurized tank.

Qualitatively, one of the aspects of a spark is that the entire path between electrodes is ionized. It is the photon emission from recombination and decay of excited states which gives rise to the light from the spark. Further, if the spark leads to a stable conduction state, the cathode must be capable of supplying the needed secondary electrons, and the conduction state produced must permit the discharge of the interelectrode capacitance at the very minimum. *See* ARC DISCHARGE; ELECTRICAL CONDUCTION IN GASES; GLOW DISCHARGE.

In a consideration of the mechanism involved in the spark, the time required for the breakdown of the gas in a gap is an important element. L. B. Loeb pointed out that this time is often less than that required for an electron to traverse the gap completely. This implies that there must be some means of ionization present other than electron impact and that the velocity of propagation of this ionizing agent or mechanism must be much greater than the electron velocity. It seems definitely established that this additional method must be photoionization. In the intense electric field which is necessary for the spark, the initial electron will produce a heavy avalanche of cumulative ionization. Light resulting from the decay processes will produce ionization throughout the gas and electrons at the surfaces by the photoelectric

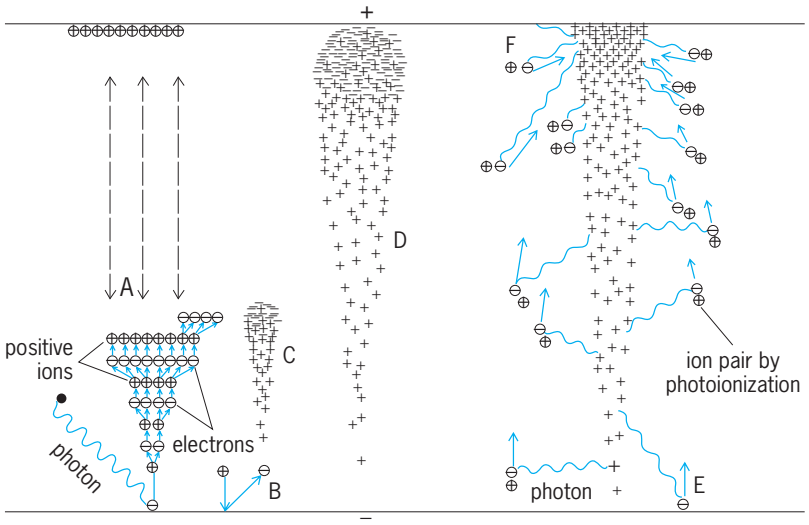

Fig. 2. Electron multiplication and avalanching during an electric spark discharge. (*After L. B. Loeb and J. M. Meek, The Mechanism of the Electric Spark, Stanford University Press, 1941*)

effect (**Fig. 2**). The electrons resulting from this will in turn produce further avalanches through the entire region, so that in a time of the order of 10^{-8} s the entire path becomes conducting. If the pressure is approximately atmospheric, the spark will be confined to a relatively narrow region, so that the conducting path, while not straight, will be a well-defined line. If the external circuit can supply the necessary current, the spark will result in an arc discharge. At lower pressure the path becomes more diffuse, and the discharge takes on either a glow or arc characteristic.

Figure 2 shows A, the electron multiplication of electrons by the cumulative ionization of a single electron liberated from the cathode by a photon; B, a secondary electron emitted from the cathode by a positively charged ion; C, the development and structure of an avalanche, with positively charged ions behind electrons at the tip; D, the avalanche crossing the gap and spreading by diffusion; and F, an older avalanche when electrons have disappeared into the anode. A positive space-charge boss appears on the cathode at F. Ion pairs out from the trail indicate the appearance of photoelectric ion pairs in the gas produced by photons from the avalanche. E shows a photoelectron from the surface of the cathode produced by the avalanche.

Theory. Mathematically, the theory of Townsend predicts that the current in a self-sustained discharge of the glow variety will follow Eq. (1), where I is

$$I = I_0 \frac{e^{\alpha x}}{1 - \gamma e^{\alpha x}} \tag{1}$$

the current with a given plate separation x, I_0 is the current when x approaches zero, and α and γ are constants associated with the Townsend coefficients. This equation represents the case where the electrode separation is varied while the ratio of electric field to pressure is held constant. The condition for a spark is that the denominator approach zero, which may be stated as in Eq. (2). Loeb indicated that this criterion must be handled carefully. Townsend's equation really represents a steady-state situation, and it is here being used to explain a transient effect. If the processes which are involved are examined more carefully, it appears that there should be a dependence on I_0 as well.

$$\gamma = e^{-\alpha x} \tag{2}$$

Glenn H. Miller

Bibliography. M. N. Hirsh and H. J. Oskam (eds.), *Gaseous Electronics*, 1978; Y. P. Raizer and J. E. Allen, *Gas Discharge Physics*, 1997.

Electric susceptibility

A dimensionless parameter measuring the ease of polarization of a dielectric. In a vacuum, the electric flux density **D** (measured in coulombs/m^2) and electric field strength **E** (volts/m) are related by Eq. (1) where ϵ_0 is the permittivity of free space,

$$\mathbf{D} = \epsilon_0 \mathbf{E} \tag{1}$$

having the approximate value 8.854×10^{-12} farad/m. In a dielectric material, polarization occurs, and the flux density has the value given by Eq. (2), where

$$\mathbf{D} = \epsilon_0 \epsilon_r \mathbf{E} = \epsilon_0 \mathbf{E} + \mathbf{P} \tag{2}$$

ϵ_r is the relative permittivity of the material and **P** is the polarization flux density. This can be written as Eq. (3), where $\chi_e = \epsilon_r - 1$ is known as the electric

$$\mathbf{P} = \epsilon_0(\epsilon_r - 1)\mathbf{E} = \epsilon_0 \chi_e \mathbf{E} \tag{3}$$

susceptibility of the dielectric material. It is a measure of that part of the relative permittivity which is due to the material itself.

The electric susceptibility can be related to the polarizability α by expressing the polarization in terms of molecular parameters. Thus Eqs. (4) hold, where

$$\mathbf{P} = N\langle\mu\rangle_{\text{avg}} = N\alpha\mathbf{E}_L \qquad \chi_e = \frac{N\alpha\mathbf{E}_L}{\epsilon_0\mathbf{E}} \tag{4}$$

N is the number of molecules per unit volume, $\langle\boldsymbol{\mu}_{\text{avg}}\rangle$ is their average dipole moment, and $\mathbf{E}_L$ is the local electric field strength at a molecular site. At low concentrations, $\mathbf{E}_L$ approaches **E**, and the susceptibility is proportional to the concentration N. For a discussion of the properties and measurement of electric susceptibility *see* PERMITTIVITY; POLARIZATION OF DIELECTRICS.

Robert D. Waldron; A. Earle Bailey

Electric switch

A device that makes, breaks, or changes the course of an electric circuit. Basically, an electric switch consists of two or more contacts mounted on an

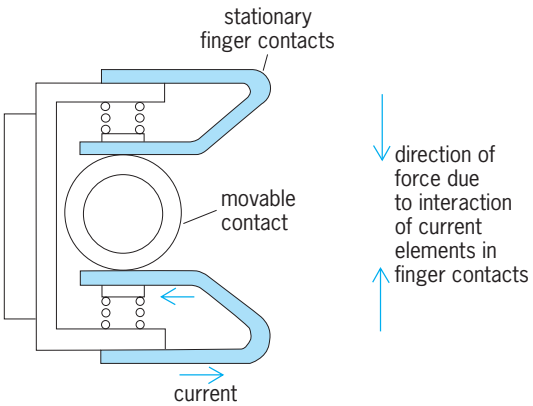

Fig. 1. With paired fingers for fixed contacts, short-circuit forces tend to hold contacts closed.

insulating structure and arranged so that they can be moved into and out of contact with each other by a suitable operating mechanism. *See* CIRCUIT (ELECTRICITY).

The term switch is usually used to denote only those devices intended to function when the circuit is energized or deenergized under normal manual operating conditions, as contrasted with circuit breakers, which have as one of their primary functions the interruption of short-circuit currents. Although there are hundreds of types of electric switches, their application can be broadly classified into two major categories: power and signal. *See* CIRCUIT BREAKER.

Power switches. In power applications, switches function to energize or deenergize an electric load. On the low end of the power scale, wall switches are used in homes and offices for turning lights on and off; dial and push-button switches control power to electric ranges, washing machines, and dishwashers. On the high end of the scale are load-break switches and disconnecting switches in power systems at the highest voltages (several hundred thousand volts).

For power applications, when closed, switches are required to carry a certain amount of continuous current without overheating, and in the open position they must provide enough insulation to isolate the circuit electrically. The latter function is particularly important in high-voltage circuits because it is the practice in the electrical industry to forbid people from working on electrical equipment unless it is isolated from the electrical supply system by a visible break in air. As an added precaution, a grounding switch is often used to connect the equipment to the ground before permitting any direct contact by a worker.

Load-break switches are required also to have the capability of interrupting the load circuit. Although this requirement is easily met in low-voltage and low-current applications, for high-voltage and high-current circuits, arc interrupters, similar to those used in circuit breakers, are needed.

In medium-voltage applications the most popular interrupter is the air magnetic type, in which the arc is driven into an arc chute by the magnetic field produced by the load current in a blowout coil. The chute may be made either of ceramic materials or of organic materials such as plexiglass. For voltages higher than 15,000 V, gaseous interrupters are sometimes used. Special gases such as sulfur hexafluoride (SF_6) have been found to have good interrupting capability for load currents without a high-speed flow. For short-circuit currents, even with SF_6, a high-speed flow is needed. Vacuum interrupters have been produced with high-voltage load-breaking duties. *See* BLOWOUT COIL; SHORT CIRCUIT.

Some load-break switches may also be required to have the capability of holding the contacts in the closed position during short-circuit conditions so that the contacts will be blown open by electromagnetic forces when the circuit breaker in the system interrupts the short-circuit current. For most contact configurations, the interaction between the short-circuit current and the magnetic field tends to force the contacts to part. Special configurations such as finger contacts are sometimes used to overcome this problem (**Fig. 1**).

Signal switches. For signal applications, switches are used to detect a specified situation that calls for some predetermined action in the electrical circuit. For example, thermostats detect termperature; when a certain limit is reached, contacts in the thermostat energize or deenergize another electrical switching device to control power flow. Centrifugal switches prevent overspeeds of motors. Limit switches prevent cranes or elevators from moving beyond preset positions. Many different types of switches perform such signaling purposes in communication systems, computers, and control systems for industrial processes. *See* ALTERNATING-CURRENT MOTOR; SWITCHING SYSTEMS (COMMUNICATIONS); THERMOSTAT.

Switches for signaling purposes are often required to have long life, high speed, and high reliability. Contaminants and dust must be prevented from interfering with the operation of the switch. For this purpose, switches are usually enclosed and are sometimes hermetically sealed.

Among the many different arrangements for contacts, the knife switch (**Fig. 2**), because of its

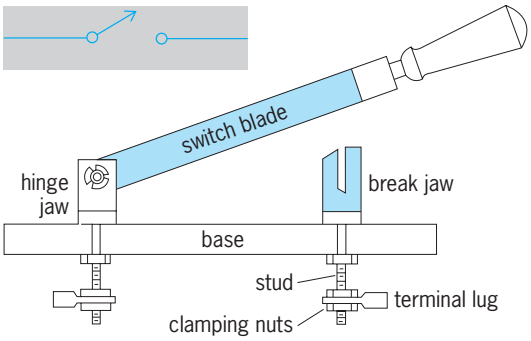

Fig. 2. Early knife switch with schematic symbol of single switch shown at top.

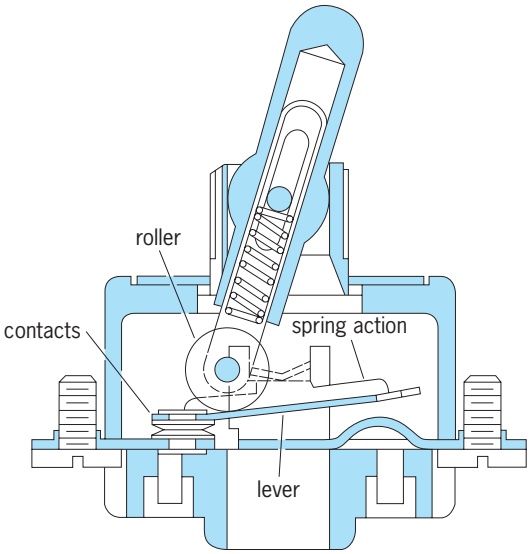

Fig. 3. Toggle mechanism actuates butt-contact switch.

early widespread use, is the basis for the elemental switch symbol. In the knife switch, a metal blade is hinged to a stationary jaw at one end and contacts a similar jaw at the other end.

In the leaf-spring switch, parallel strips of spring metal are sandwiched between blocks of insulation at one end and pushed into or out of contact at the other end. The sliding contact switch takes the form of a dial or drum with metal segments engaging contact fingers that wipe over the dial or drum surface. A butt-contact switch (**Fig. 3**) consists of one movable contact in the form of a disk or bar and one or two stationary contacts of similar form. In the mercury switch, a pool of mercury is made either to bridge the contacts or to separate from them (**Fig. 4**).

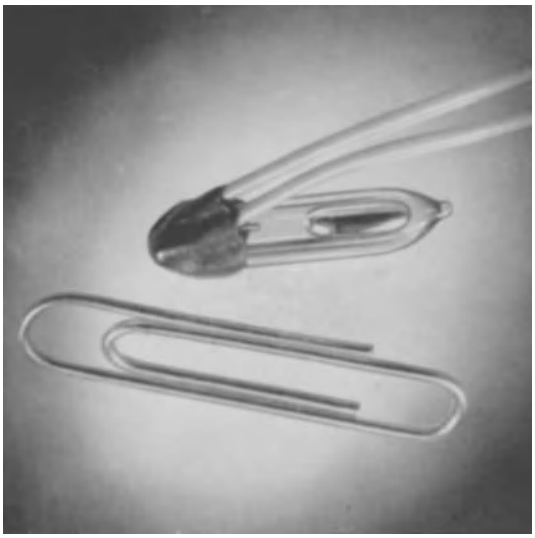

Fig. 4. Miniature mercury element may be used with a variety of tilting mechanisms to cause switch action. (*Minneapolis-Honeywell Regulator Co.*)

Switches frequently are composed of many single circuit elements, known as poles, all operated simultaneously or in a predetermined sequence by the same mechanism. Switches used in complex machines, such as computers, may have a large number of poles. Switches used in power circuits usually have from one to four poles, depending on the kind of circuit. Switches are often typed by the number of poles and referred to as single-pole or double-pole switches, and so on. It is common to express the number of possible switch positions per pole, such as a single-throw or double-throw switch.

Thomas H. Lee

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999.

Electric transient

A temporary component of current and voltage in an electric circuit which has been disturbed. In ordinary circuit problems, a stabilized condition of the circuit is assumed and steady-state values of current and voltage are sufficient. However, it often becomes important to know what occurs during the transition period following a circuit disturbance until the steady-state condition is reached. Transients occur only in circuits containing inductance or capacitance. In general, transients accompany any change in the amount or form of energy stored in the circuit. After the following Introduction, transients in direct-current (dc) and alternating-current (ac) circuits with lumped elements are discussed. Transients of a more complex nature occur on distributed-parameter circuits, such as transmission lines. Such phenomena are elaborated upon in a separate section on transients in power systems. Finally, the modeling of transients in digital simulation is explained.

Introduction

An electric circuit or system under steady-state conditions of constant, or cyclic, applied voltages or currents is in a state of equilibrium. However, the circuit conditions of voltage, current, or frequency may change or be disturbed. Also, circuit elements may be switched in or out of the circuit. Any change of circuit condition or circuit elements causes a transient readjustment of voltages and currents from the initial state of equilibrium to the final state of equilibrium. In a sense the transient may be regarded as superimposed on the final steady state, so that Eq. (1) applies.

$$\begin{pmatrix}\text{Instantaneous}\\ \text{condition}\end{pmatrix} = \begin{pmatrix}\text{final}\\ \text{condition}\end{pmatrix} + \begin{pmatrix}\text{transient}\\ \text{terms}\end{pmatrix} \quad (1)$$

Furthermore, since the instantaneous condition at the first instant of disturbance (time zero) must be

the initial condition, it may be described by Eq. (2).

$$\begin{pmatrix}\text{Initial}\\ \text{condition}\end{pmatrix} = \begin{pmatrix}\text{final}\\ \text{condition}\end{pmatrix} + \begin{pmatrix}\text{transient terms}\\ \text{at time zero}\end{pmatrix} \quad (2)$$

A great deal of information may be obtained from these two "word equations" without recourse to mathematics. For example, if the weight on the end of a vertical spring is suddenly increased, its final displacement can be determined. Since the spring-weight combination is known to be an oscillating system, the amplitude of the transient oscillation follows from Eq. (2) as the difference between the initial and final displacements.

The nature of an electric transient is determined by three things: (1) the circuit or network itself—the interconnections of its elements and the circuit parameters (resistances R, inductances L, capacitances C, mutual inductances M); (2) the initial conditions of voltages, currents, charges, and flux linkages at the start of the transient; and (3) the nature of the disturbance which initiated the transient.

The circuit, or network, is usually defined by a diagram of connections showing all of the interconnections, junctions, meshes, circuit parameters, voltage and current sources and their polarities, and switches. Corresponding to the network a differential (or integral-differential) equation, or a set of such equations, may be written. These equations may also be written in terms of operational calculus, or as Laplace transforms, or in any other suitable mathematical equivalents. They are established in accordance with Kirchhoff's laws: (I) The sum of the currents i at a junction is equal to zero: $\Sigma i = 0$. (II) The sum of the voltages e around a mesh is equal to zero: $\Sigma e = 0$. *See* KIRCHHOFF'S LAWS OF ELECTRIC CIRCUITS.

The voltage drops associated with a resistance, inductance, mutual inductance, and capacitance, respectively, are given by Eqs. (3). In applying

$$\begin{aligned} e_R &= Ri & e_L &= L\frac{di}{dt} \\ e_i &= M_{12}\frac{di_2}{dt} & e_c &= \frac{1}{C}\int i\,dt \end{aligned} \quad (3)$$

Kirchhoff's law to a closed mesh, it is merely necessary (with due regard for signs and polarities) to equate the sum of all the voltage sources (such as generators and batteries) to the sum of all the voltage drops in the mesh, as in Eq. (4), where m refers

$$\sum_m e_m = \sum_m \sum_n \left(R_{mn} i_n + L_{mn}\frac{di_n}{dt} + \frac{1}{C_{mn}}\int i_n\,dt\right) \quad (4)$$

to any branch of the mesh in question, and n to any current causing a voltage drop in that branch (a branch may carry currents from meshes other than the mesh in question, or be mutually coupled with other branches). Equation (4) formulates the differential equation of the circuit in terms of the voltages and currents. It is sometimes convenient to make use of charges q and fluxes ϕ by the substitutions of Eqs. (5).

$$i = \frac{dq}{dt} \quad \text{or} \quad q = \int i\,dt \quad \text{and} \quad N\phi = Li \quad (5)$$

Since Eq. (4) contains an integral, it is convenient to eliminate it either by differentiating once or by substituting Eq. (5). Then Eqs. (6) follow.

$$\begin{aligned} \sum_m \frac{de_m}{dt} &= \sum_m \sum_n \left(R_{mn}\frac{di_n}{dt} + L_{mn}\frac{d^2 i_n}{dt^2} + \frac{i_n}{C_{mn}}\right) \\ \sum_m e_m &= \sum_m \sum_n \left(R_{mn}\frac{dq_n}{dt} + L_{mn}\frac{d^2 q_n}{dt^2} + \frac{q_n}{C_{mn}}\right) \end{aligned} \quad (6)$$

See COUPLED CIRCUITS.

It is customary to factor out the current from Eq. (4) and write the equation in terms of a generalized impedance operator Z_{mn} defined by Eq. (7). Since an equation of this type can be

$$\begin{aligned} \sum_m e_m &= \sum_m \sum_n \left(R_{mn} + L_{mn}\frac{d}{dt} + \frac{1}{C_{mn}}\int dt\right) i_n \\ &= \sum_m \sum_n Z_{mn} i_n \end{aligned} \quad (7)$$

written for each of the N-meshes of the network, the totality of the differential equations for the entire network constitutes a system of simultaneous differential equations in the form of Eqs. (8).

$$\begin{aligned} e_1 &= Z_{11}i_1 + Z_{12}i_2 + \cdots + Z_{1N}i_N \\ e_2 &= Z_{21}i_1 + Z_{22}i_2 + \cdots + Z_{2N}i_N \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ e_N &= Z_{N1}i_1 + Z_{N2}i_2 + \cdots + Z_{NN}i_N \end{aligned} \quad (8)$$

In general, the process of solution of such a set of equations for any current leads to a differential equation of order $2N$; there will be $2N$ integration constants associated with it, and these integration constants must be determined from the initial conditions at the first instant of the disturbance.

Equation (9) is the solution of an ordinary

$$\begin{pmatrix}\text{Complete}\\ \text{solution}\end{pmatrix} = \begin{pmatrix}\text{particular}\\ \text{integral}\end{pmatrix} + \begin{pmatrix}\text{complementary}\\ \text{solution}\end{pmatrix} \quad (9)$$

differential equation with constant coefficients and is in two parts.

The particular integral depends on the form of the applied voltage, and represents the final steady-state solution.

The complementary solution is independent of the form of the applied voltages, but depends on the initial conditions in the circuit, and represents the transient terms. Thus Eq. (9) is the mathematical equivalent of Eq. (1).

The opening of a switch may be simulated by superimposing an equal and opposite current and thereby canceling the current through the switch. Likewise, the closing of a switch may be simulated by canceling the voltage across its terminals by an equal and opposite voltage. The effect of a cancellation current (or voltage) may be superimposed on the currents and voltages already existing in the system, as in Eq. (10).

$$\begin{pmatrix}\text{Resultant}\\ \text{voltages}\\ \text{and}\\ \text{currents}\end{pmatrix} = \begin{pmatrix}\text{those which}\\ \text{would exist if}\\ \text{switches were}\\ \text{not operated}\end{pmatrix} + \begin{pmatrix}\text{those due to}\\ \text{“cancellation”}\\ \text{voltages or}\\ \text{currents}\end{pmatrix} \tag{10}$$

In many practical cases transients are one of three types:

1. Single-energy transients, in which only one form of energy storage (electromagnetic or electrostatic) is present; the transient exhibits simple exponential decay from the initial to the final conditions.
2. Double-energy transients, in which both forms of energy storage are present; the transient is either aperiodic or a damped sinusoid.
3. Combination of 1 and 2.

DC Circuit with Lumped Elements

Transients are initiated in dc circuits either when a switch is closed on a dc voltage, or when a switch is opened and the detached circuit is permitted to discharge its energy sources. The assumption is made that the suddenly applied dc voltage is sustained at a constant value for an indefinite period after the switching operation, and time is counted from the instant, $t = 0$, at which the switch is operated. It is further usually assumed, where possible, that all circuit parameters (resistances R, inductances L, and capacitances C) are constant, so that the resulting differential equations are linear.

The energy sources in dc systems comprise dynamoelectric machines, mercury-arc rectifiers, vacuum tubes, and electric batteries. Energy W is stored in a circuit in inductances ($W = 1/2Li^2$) and capacitances ($W = 1/2Ce^2$). Energy is dissipated in circuit resistances ($W = Ri^2$).

Resistance and inductance in series. A lumped parameter circuit comprising a dc voltage source E, a resistance R, and an inductance L is shown in **Fig. 1**. When the switch S is closed, the voltage E is suddenly impressed on the series circuit. By Kirchhoff's second law the differential equation is given by Eq. (11).

$$E = Ri + L\frac{di}{dt} \tag{11}$$

Equating the right-hand side of Eq. (11) to zero, the complementary solution is found by assuming that Eq. (12), in which I and a are constants to be

$$i = I\epsilon^{at} \tag{12}$$

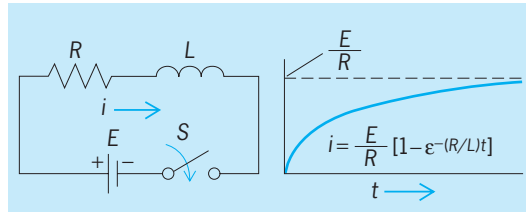

Fig. 1. Resistance and inductance in series.

determined, applies. Upon substitution in Eq. (11), Eq. (13) results, from which follows $a = -R/L$.

$$0 = RI\epsilon^{at} + aLI\epsilon^{at} = (R + aL)i \tag{13}$$

The particular integral is the final steady-state solution or $i(\infty) = E/R$. The complete solution, Eq. (14), is the sum of the complementary solu-

$$i = \frac{E}{R} + I\epsilon^{-(R/L)t} \tag{14}$$

tion and the particular integral. But this expression contains the unknown integration constant I which must be found from the initial conditions. At $t = 0$, the instant at which the switch is closed, there can be no current in the circuit, since it is impossible to store energy in an inductance instantaneously. Therefore $i(0) = 0$ and when this is put in Eq. (14), Eq. (15) is obtained.

$$i(0) = 0 = \frac{E}{R} + I \quad \text{or} \quad I = -\frac{E}{R} \tag{15}$$

The complete solution by Eq. (14) therefore is given by Eq. (16), which is in the form of Eq. (1).

$$i = \frac{E}{R} - \frac{E}{R}\epsilon^{-(R/L)t} \tag{16}$$

Here E/R is the final (steady-state) condition and $-(E/R)\epsilon^{-(R/L)t}$ is the transient term. Note also that at $t = 0$ this equation gives $i(0) = 0$, agreeing with Eq. (2). The graph of the transient is shown in Fig. 1, where the current starts at zero, rises exponentially, and approaches its steady-state value of E/R as t approaches infinity. *See* ELECTRICAL RESISTANCE; INDUCTANCE.

Resistance and capacitance in series. A resistance-capacitance series circuit is shown in **Fig. 2**. A residual charge q_0 is on the capacitor just prior to closing the switch S.

Equation (17) is the differential equation of the

$$E = Ri + \frac{1}{C}\int i\,dt \tag{17}$$

circuit. This equation may be solved as it is, or

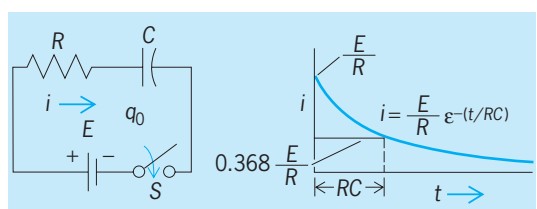

Fig. 2. Resistance and capacitance in series.

converted to a linear differential equation, Eq. (18),

$$0 = R\frac{di}{dt} + \frac{i}{C} \tag{18}$$

by differentiation or restated in terms of the instantaneous charge on the capacitor by substituting $i = dq/dt$, giving Eq. (19).

$$E = R\frac{dq}{dt} + \frac{q}{C} \tag{19}$$

The solution to this equation, following precisely the steps of the previous paragraph, is given by Eq. (20), in which Q is the integration constant. But

$$q = Q\epsilon^{-(t/RC)} + CE \tag{20}$$

initially, at $t = 0$, there was a residual charge q_0 on the capacitance; therefore the relationship given by Eq. (21) pertains, and upon the substitution of Eq. (21) in Eq. (20), Eq. (22) is obtained. This

$$q(0) = q_0 = Q + CE \text{ or } Q = q_0 - CE \tag{21}$$

$$q = CE - (CE - q_0)\epsilon^{-(t/RC)} \tag{22}$$

equation is in the form of Eq. (1), in which CE is the final (steady-state) condition and the other term is the transient. Initially, at $t = 0$, $q = q_0$ in agreement with Eq. (2). The current is given by Eq. (23).

$$\begin{aligned} i = \frac{dq}{dt} &= \frac{CE - q_0}{RC}\epsilon^{-(t/RC)} \\ &= \left(\frac{E}{R} - \frac{q_0}{RC}\right)\epsilon^{-(t/RC)} \end{aligned} \tag{23}$$

Equations (22) and (23) have been plotted in Fig. 2 for the case $q_0 = 0$. *See* CAPACITANCE.

Resistance, inductance, and capacitance in series. This circuit is shown in **Fig. 3**. The capacitor is assumed to have an initial charge q_0, or an initial voltage $V = q_0/C$. The differential equation of the circuit is given by Eq. (24).

$$E = Ri + L\frac{di}{dt} + \frac{1}{C}\int i\,dt \tag{24}$$

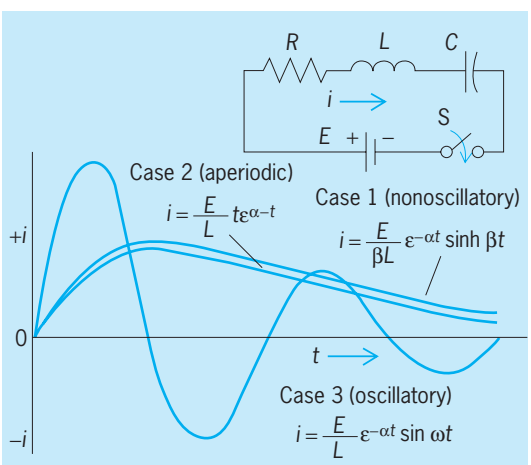

Fig. 3. Resistance, inductance, and capacitance in series for three cases.

Differentiating once to clear the integral, Eq. (25)

$$0 = R\frac{di}{dt} + L\frac{d^2i}{dt^2} + \frac{i}{C} \tag{25}$$

is obtained. Equation (25) is a second-order linear differential equation with constant coefficients; since it is equal to zero, it will possess only a complementary solution. Equation (26) is assumed, which, upon substitution in Eq. (25), yields Eq. (27).

$$i = A\epsilon^{at} \tag{26}$$

$$0 = aRA\epsilon^{at} + a^2LA\epsilon^{at} + (A/C)\epsilon^{at} \tag{27}$$

Canceling the constant A and the exponential, there results a quadratic in a whose solution gives the two possible values shown in Eqs. (28).

$$\begin{aligned} a_1 &= \frac{-RC + \sqrt{R^2C^2 - 4LC}}{2LC} \\ a_2 &= \frac{-RC - \sqrt{R^2C^2 - 4LC}}{2LC} \end{aligned} \tag{28}$$

Associating the integration constant A_1 with a_1 and A_2 with a_2, the solution takes the form of Eq. (29).

$$i = A_1\epsilon^{a_1t} + A_2\epsilon^{a_2t} \tag{29}$$

The voltage across the capacitor, Eq. (30), is obtained from Eq. (24).

$$\begin{aligned} e_C &= \frac{1}{C}\int i\,dt = E - Ri - L\frac{di}{dt} \\ &= E - (R + a_1L)A_1\epsilon^{a_1t} - (R + a_2L)A_2\epsilon^{a_2t} \end{aligned} \tag{30}$$

Initially, at $t = 0$, the current must be zero because of the inductance, and the capacitor voltage is $e_C = V$. By the first of these conditions, $i(0) = 0$, in Eq. (29) it is seen that $A_2 = -A_1$. And by the second condition, $e_C(0) = V$, in Eq. (30), Eq. (31) is obtained.

$$\begin{aligned} V &= E - (R + a_1L)A_1 - (R + a_2L)A_2 \\ &= E - (a_1 - a_2)LA_1 \\ A_1 &= \frac{E - V}{(a_1 - a_2)L} = \frac{C(E - V)}{\sqrt{R^2C^2 - 4LC}} \end{aligned} \tag{31}$$

Then Eq. (32) is the complete solution.

$$i = \frac{C(E - V)}{\sqrt{R^2C^2 - 4LC}}(\epsilon^{a_1t} - \epsilon^{a_2t}) \tag{32}$$

There are three special cases of this solution, depending on the nature of the radical in Eq. (32).

Nonoscillatory case. $R^2C > 4L$. In this case the radical is positive, and the exponents a_1 and a_2 are real and negative, $a_1 = -\alpha + \beta$, $a_2 = -\alpha - \beta$, where $\alpha = R/2L$ and $\beta = \sqrt{R^2C^2 - 4LC}/2LC$. Then i is given by Eq. (33), which is shown in Fig. 3, where q_0 is

$$\begin{aligned} i &= \frac{C(E - V)}{\sqrt{R^2C^2 - 4LC}}\epsilon^{-\alpha t}(\epsilon^{+\beta t} - \epsilon^{-\beta t}) \\ &= \frac{2C(E - V)}{\sqrt{R^2C^2 - 4LC}}\epsilon^{-\alpha t}\sinh \beta t \end{aligned} \tag{33}$$

assumed to be 0 and $V = 0$.

Aperiodic case. $R^2C = 4L$. In this case $a_1 = a_2 = -R/2L$, and the radical in the denominator of Eq. (32) is zero, as is the numerator. The indeterminant is easily evaluated from Eq. (33) upon letting $\beta \to 0$, thus giving Eq. (34). This is the aperiodic or

$$i = \frac{2C(E-V)}{2LC\beta} e^{-\alpha t} \sinh \beta t \mid_{\beta \to 0}$$

$$= \frac{E-V}{L} e^{-\alpha t} t \tag{34}$$

critical case, and is illustrated in Fig. 3.

Oscillatory case. $R^2C < 4L$. In this case the radical in Eq. (28) becomes imaginary and the exponents take the form of Eqs. (35), and Eq. (32) becomes Eq. (36). This is a damped oscillation and is illustrated in Fig. 3.

$$a_1 = \frac{-R}{2L} + j\frac{\sqrt{4LC - R^2C^2}}{2LC} = -\alpha + j\omega$$
$$a_2 = \frac{-R}{2L} - j\frac{\sqrt{4LC - R^2C^2}}{2LC} = -\alpha - j\omega \tag{35}$$

$$i = \frac{C(E-V)}{2j\omega LC} e^{-\alpha t}(\epsilon^{j\omega t} - \epsilon^{-j\omega t})$$

$$= \frac{E-V}{\omega L} \epsilon^{-\alpha t} \sin \omega t \tag{36}$$

AC Circuit with Lumped Elements

Alternating-current transients differ from direct-current transients in two important respects: (1) The final condition, or steady state, is an alternating or cyclic one, and (2) the amplitudes of the transient terms depend on the point on the ac applied voltage wave at which the transient is initiated, and can therefore have many different values, or even change sign.

In the section on dc transients, solutions were carried out for *RL*, *RC*, and *RLC* circuits switched onto a voltage source. In the present section, the solution will be carried out in detail for the *RLC* circuit only; the others will be regarded as special cases of this general solution by putting $C = \infty$ and $L = 0$, respectively.

Resistance, inductance, and capacitance in series. Consider the circuit of **Fig. 4** in which an *RLC* circuit is suddenly switched onto an ac voltage $e = E \sin(\omega t + \gamma)$ at an electrical angle γ displaced from $\omega t = 0$. Assume that a current I is flowing in the circuit and a voltage V is across the capacitance at the instant the switch is closed on the alternator. The voltage equation and the initial conditions then are given by Eqs. (37) and (38).

$$E \sin(\omega t + \gamma) = Ri + L\frac{di}{dt} + \frac{1}{C}\int i\,dt \tag{37}$$

$$i = I \quad \text{and} \quad e_c = V \quad \text{at} \quad t = 0 \tag{38}$$

Differentiating Eq. (37) once to clear it of the integral gives Eq. (39). The complementary solution

$$\omega E \cos(\omega t + \gamma) = L\frac{d^2i}{dt^2} + R\frac{di}{dt} + \frac{i}{C} \tag{39}$$

of this equation is the same as in the dc case of

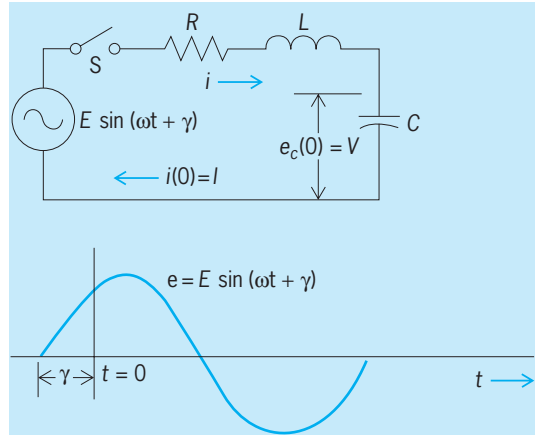

Fig. 4. ***RLC*** **circuit with alternating current source.**

Eq. (25), that is, Eq. (29), where a_1 and a_2 are given in Eq. (28).

The particular integral of Eq. (39) is its final steady-state current. This is most easily obtained as Eq. (40), the ordinary ac solution for the current, to which Eqs. (41) and (42) apply.

$$i_{ac} = \frac{E}{Z} \sin(\omega t + \gamma - \theta) \tag{40}$$

$$Z = \sqrt{R^2 + (\omega L - 1/\omega C)^2} \tag{41}$$

$$\tan\theta = \frac{\omega L - 1/\omega C}{R} = \frac{\omega^2 LC - 1}{\omega CR} \tag{42}$$

The complete solution then is given by Eq. (43),

$$i = \frac{E}{Z} \sin(\omega t + \gamma - \theta) + A_1\epsilon^{a_1 t} + A_2\epsilon^{a_2 t} \tag{43}$$

in which A_1 and A_2 are integration constants and a_1 and a_2 are given in Eq. (28).

The capacitor voltage is given by Eq. (44).

$$e_c = \frac{1}{C}\int i\,dt = -\frac{E}{\omega CZ}\cos(\omega t + \gamma - \theta) + \frac{A_1}{Ca_1}\epsilon^{a_1 t} + \frac{A_2}{Ca_2}\epsilon^{a_2 t} \tag{44}$$

Now applying the initial conditions of Eq. (38) at $t = 0$, there results from Eqs. (43) and (44), respectively, Eqs. (45) and (46).

$$i(0) = I = \frac{E}{Z}\sin(\gamma - \theta) + A_1 + A_2 \tag{45}$$

$$e_c(0) = V = \frac{-E}{\omega CZ}\cos(\gamma - \theta) + \frac{A_1}{Ca_1} + \frac{A_2}{Ca_2} \tag{46}$$

Solving Eqs. (45) and (46) simultaneously for A_1 and A_2 and using Eq. (28), Eqs. (47)–(49) result.

$$A_1 = \frac{1}{\sqrt{R^2C^2 - 4LC}}\left\{\frac{I}{a_2} - CV - \frac{E}{Z}\left[\frac{1}{\omega}\cos(\gamma - \theta) + \frac{1}{a_2}\sin(\gamma - \theta)\right]\right\} \tag{47}$$

$$A_2 = \frac{-1}{\sqrt{R^2C^2 - 4LC}} \left\{ \frac{I}{a_1} - CV - \frac{E}{Z}\left[\frac{1}{\omega}\cos(\gamma - \theta) + \frac{1}{a_1}\sin(\gamma - \theta)\right]\right\} \quad (48)$$

$$a_1 = \frac{-R}{2L} + \sqrt{\left(\frac{R}{2L}\right)^2 - \frac{1}{LC}}$$
$$a_2 = \frac{-R}{2L} - \sqrt{\left(\frac{R}{2L}\right)^2 - \frac{1}{LC}} \quad (49)$$

These values, together with Eqs. (41) and (42), inserted in Eqs. (43) and (44) give the solution in a form suitable for the critically damped case. However, if $R^2C < 4L$, then the radicals in Eqs. (47)–(49) become imaginary and these expressions become thecomplex numbers given in Eqs. (50), whereupon

$$A_1 = {}^1/_2(M + jN) \qquad A_2 = {}^1/_2(M - jN)$$
$$a_1 = -\alpha + j\omega \qquad a_2 = -\alpha - j\omega \quad (50)$$

Eq. (43) takes the damped oscillatory form given in Eq. (51).

$$i = \frac{E}{Z}\sin(\omega t + \gamma - \theta) + \epsilon^{-\alpha t}(M\cos\omega t - N\sin\omega t) \quad (51)$$

It is evident from Eq. (43) or Eq. (51) that the final steady-state value, after the transient has died out, is the alternating current of Eq. (40). It is also clear from Eqs. (47) and (48) that the amplitudes of the transient terms depend upon the angle γ on the ac applied voltage wave at $t = 0$.

Resistance and inductance in series. This may be regarded as a special case of Eq. (43), in which $C = \infty$ (the capacitance short-circuited) and $V = 0$. Consequently, the solution is given by Eq. (52). Thus

$$i = \frac{E}{Z}\sin(\omega t + \gamma - \theta) + \left[I - \frac{E}{Z}\sin(\gamma - \theta)\right]\epsilon^{-(R/L)t} \quad (52)$$

the transient starting from an initial value I decays exponentially to its final ac steady-state value.

There will be no transient if the switch is closed at the angle γ on the voltage wave such that Eq. (53) applies.

$$I = \frac{E}{Z}\sin(\gamma - \theta) \quad (53)$$

Resistance and capacitance in series. This may be regarded as a special case of Eq. (43), in which $L = 0$ and $I = 0$. Under these conditions the solution is given by Eq. (54) and the capacitor voltage is,

$$i = \frac{E}{Z}\sin(\omega t + \gamma - \theta) - \left[\frac{V}{R} + \frac{E}{\omega CRZ}\cos(\gamma - \theta)\right]\epsilon^{-t/RC} \quad (54)$$

by Eq. (44), defined by Eq. (55). Thus the transient,

$$e_c = \frac{-E}{\omega CZ}\cos(\omega t + \gamma - \theta) + \left[V + \frac{E}{\omega CZ}\cos(\gamma - \theta)\right]\epsilon^{-t/RC} \quad (55)$$

starting with a voltage V, decays exponentially to its final ac steady-state value. If the switch is closed at an angle γ on the voltage wave such that Eq. (56)

$$V = -\frac{E}{\omega CZ}\cos(\gamma - \theta) \quad (56)$$

applies, there will be no transient. Loyal V. Bewley

Power Systems with Lumped and Distributed Elements

In an electric power system, a transient is a temporary disturbance leading to changes in the amounts of energy stored in the capacitances and inductances, causing variations of power system voltage and current waveforms. Such disturbances may also be referred to as electromagnetic transients to indicate that oscillatory variations are caused through the alternating conversion of electric energy of capacitors into magnetic energy of inductors. The frequency range typically extends from several hertz to several megahertz.

Classification. A classification of electric transients by cause is given in **Fig. 5**. Lightning strikes induce currents that cause electric transients in the form of traveling waves to propagate in forward and backward directions on lines.

Equipment switching events change the network topology, that is, the network configuration, and force redistribution of stored energy. Switching can also result in traveling waves on lines or cause energization and deenergization of lumped capacitances and inductances.

Conceptual analysis. The following study of the transient recovery voltage (TRV) serves as introduction to conceptual analysis of electric transients involving both harmonic oscillations and traveling waves. **Figure 6*a*** shows a line-to-ground short circuit on a transmission line. The fault current is supplied via a transformer and needs to be interrupted to clear the fault. The opening of a circuit breaker triggers electric transients due to a change in the

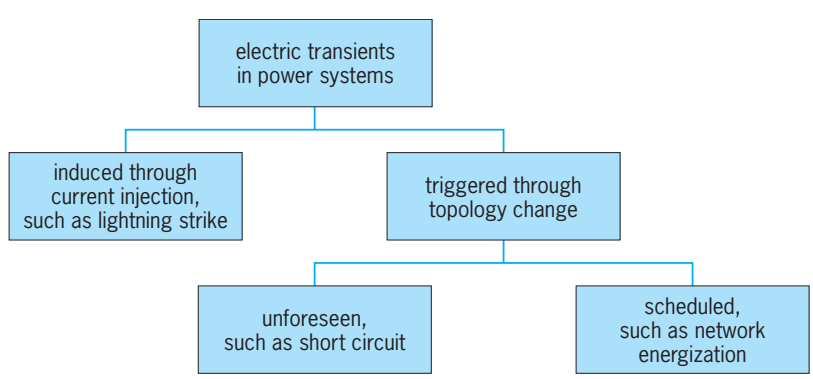

Fig. 5. Causes of electric transients in power systems.

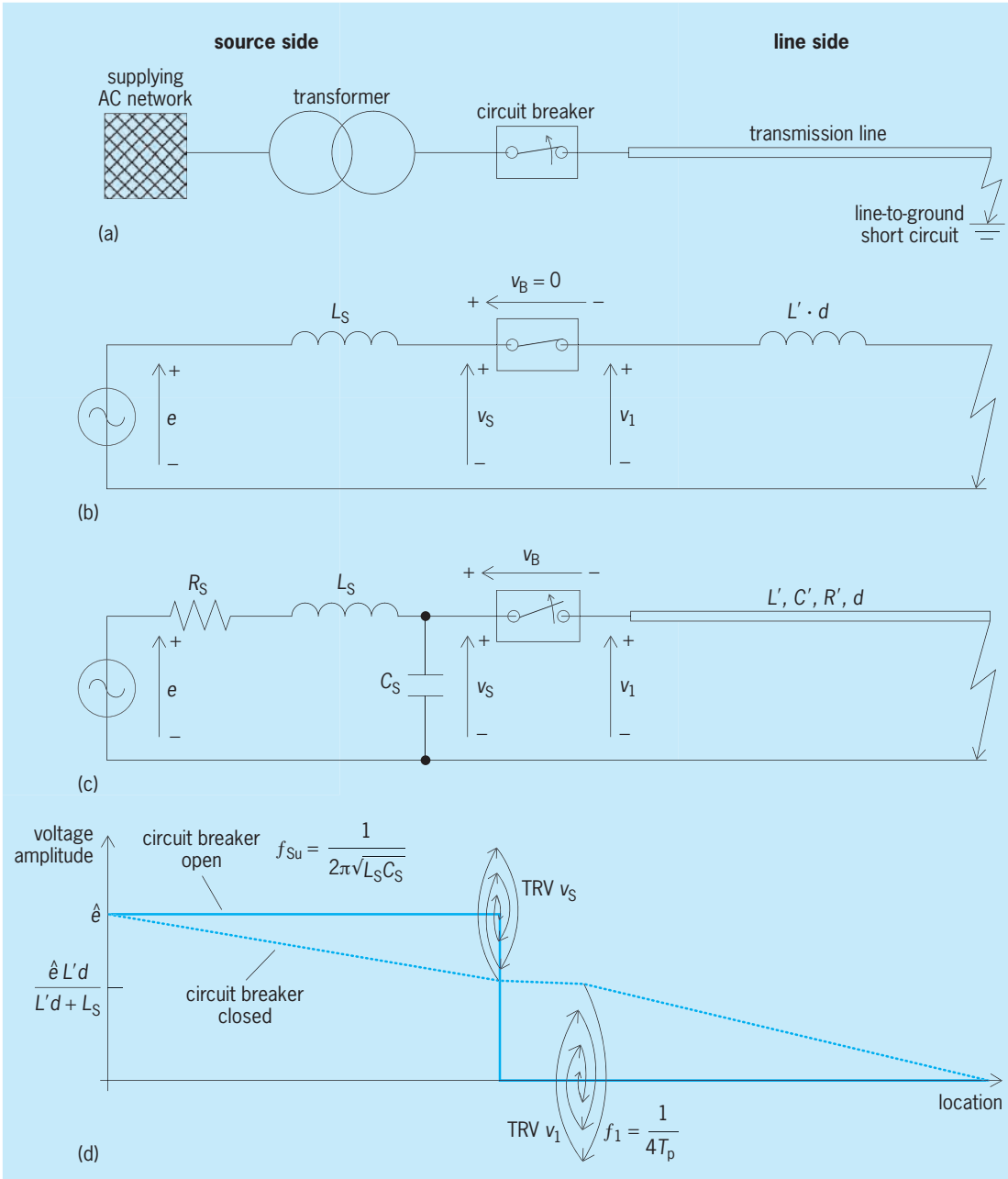

Fig. 6. Analysis of electric transients involving both harmonic oscillation and traveling waves. (*a*) Configuration of power system for study of transient recovery voltage (TRV). (*b*) Model for evaluating steady state before breaker opening. (*c*) Model for evaluating electric transients upon breaker opening. (*d*) Qualitative assessment of transients described by voltages.

topology of the system. In order to focus on the behavior of transients, only single-pole networks will be considered. (Multipole networks can be represented as single-pole equivalents.) *See* CIRCUIT BREAKER; SHORT CIRCUIT; TRANSMISSION LINES.

Before the circuit breaker opens, the voltage v_B across it is zero, as indicated in Fig. 6*b*, and hence $v_S = v_l$. The circuit of Fig. 6*b* models the configuration of Fig. 6*a* at source frequency f_e, which is 50 or 60 Hz in utility power systems depending on regional preferences. The purpose of this model is to estimate the voltage profile while the fault is still in place and the circuit breaker is closed. This allows the determination of the steady-state voltages giving initial conditions from which the transients will start. The line has inductance per unit length L', and the distance to the fault is d. The source side is modeled through an equivalent ac voltage source e behind a lumped inductive element L_S, which represents inductances of the transformer and the supplying ac network. Since the circuit breaker is to open at the instant when the fault current crosses zero to reduce stress, the voltages will reach their peak values at this instant. This is due to the inductive behavior of the line and the resulting phase shift of 90° between the fault current and the voltages v_S and v_l. The voltage profile with the circuit breaker closed is shown by a line in Fig. 6*d*, which gives the

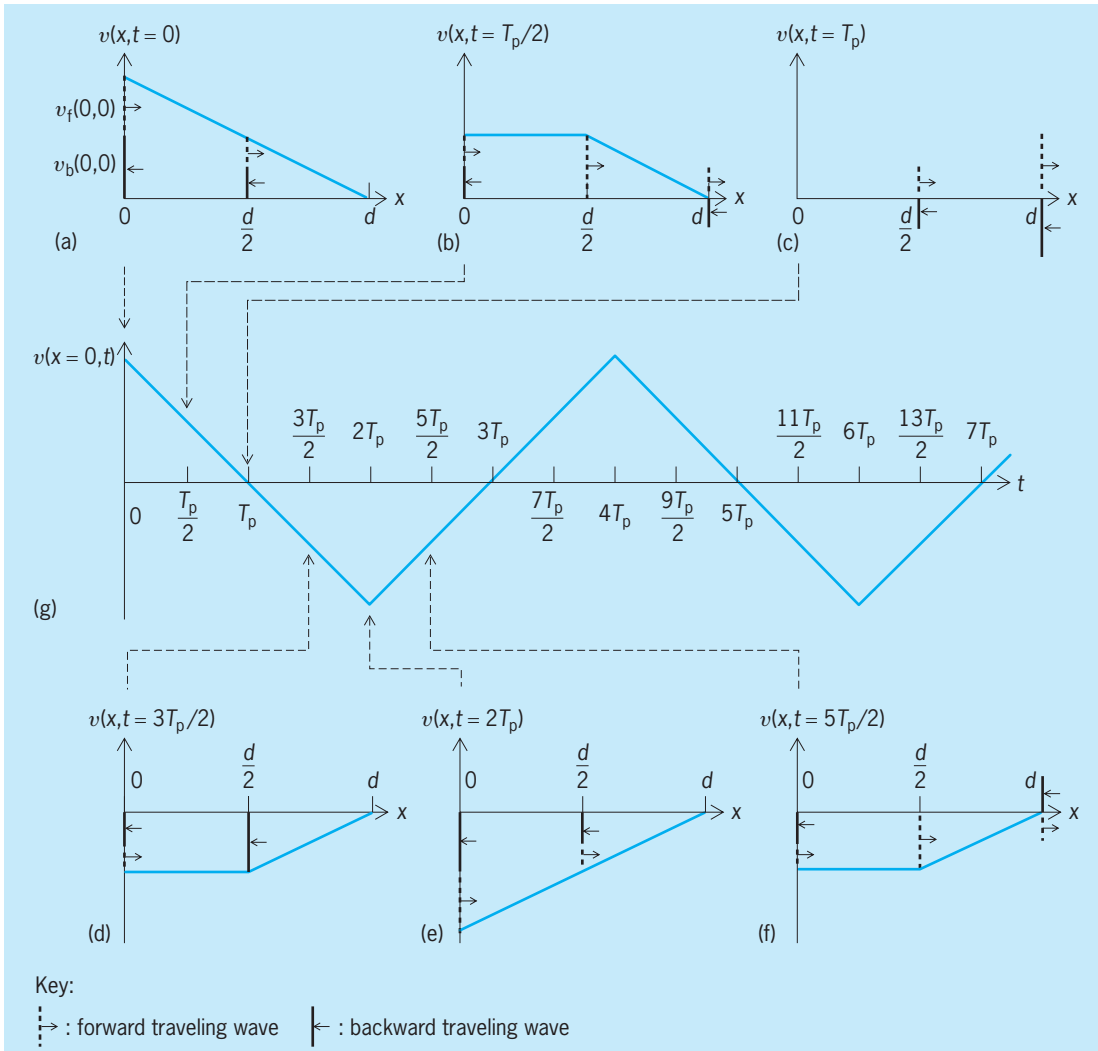

Fig. 7. Graphical interpretation of electromagnetic transients on transmission line. (*a–f*) Voltage profiles along line. (*g*) Derived voltage at beginning of line as a function of time.

amplitude as a function of location. The amplitude of the voltage source is denoted by $\hat{e}$. At the location of the short circuit, the voltage is zero. The amplitudes of v_S and v_l are obtained with the voltage divider expression given on the ordinate. Figure 6*d* also shows the steady-state voltage profile that is reached when the circuit breaker is open and the fault current is zero.

To assess the electric transients between the two steady states, the steady-state model of Fig. 6*b* is not adequate. To represent the redistribution of energy, both inductors and capacitors need to be modeled as indicated in Fig. 6*c*. The lumped capacitive element C_S represents the capacitance of transformer and bus bars on the source side, while C' gives the capacitance per unit length of the line. Losses are accounted for through R_S and R' on the source and line sides, respectively.

As the circuit breaker opens, the source and line sides become decoupled. The TRV v_S must first increase, and describes a damped harmonic oscillation around the new steady state (Fig. 6*d*). If R_S were neglected, the undamped frequency of this harmonic oscillation on the source side would be given by Eq. (57), but because of the presence of R_S the observed value is slightly lower. The exact trajectory of the source-side TRV can be calculated with the techniques developed for circuits with lumped parameters. At the entrance to the line, the TRV v_l will first decrease.

$$f_{Su} = 1/(2\pi\sqrt{L_S C_S}) \tag{57}$$

The waveform of the TRV v_l is influenced by traveling waves which are investigated by means of **Fig. 7** for the case of a lossless line. In Fig. 7*a* the initial voltage profile along the line is redrawn from Fig. 6*d*. The entrance to the line is at $x = 0$, and the circuit breaker is opened at $t = 0$. The waves propagate in all possible directions. Since the line is one-dimensional, there exist a forward-traveling wave marked with subscript f and arrow in direction of increasing x, and a backward-traveling wave marked with subscript b and arrow in direction of decreasing x. At any point on the line, the voltage is given as the sum of forward and backward traveling waves.

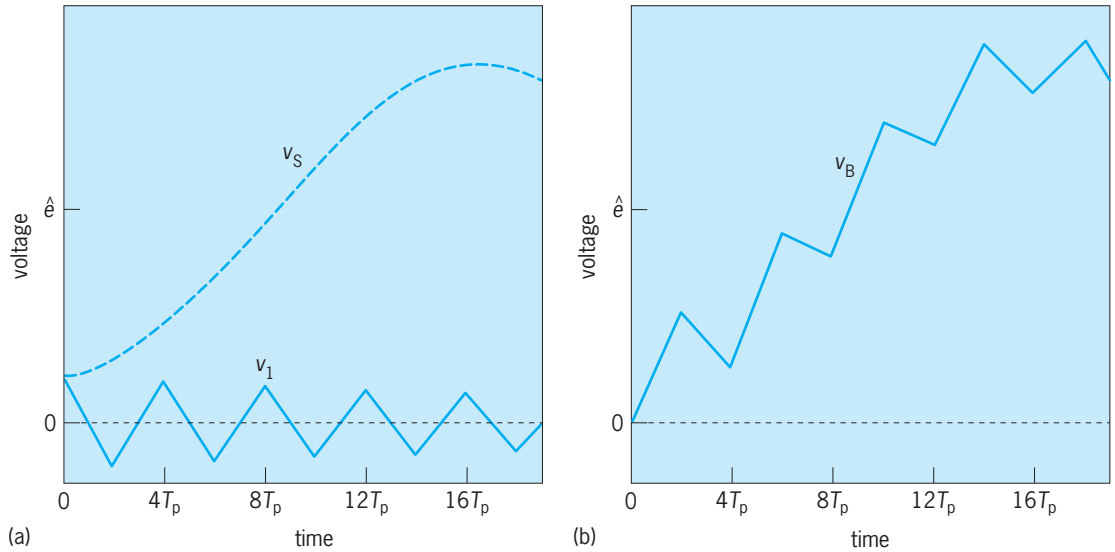

Fig. 8. Transient recovery voltages (TRVs). (*a*) TRVs on source side (v_S) and line side (v_l). (*b*) TRV across circuit breaker (v_B).

The propagation time that the waves need to travel from one end of the line to the other end is denoted by T_p.

Figure 7*b* illustrates how the original voltage profile of $t = 0$ has changed as $t = T_p/2$ is reached. The forward traveling wave of $x = 0$ at $t = 0$ traveled to $x = d/2$. The forward traveling wave of $x = d/2$ at $t = 0$ has moved by the same amount from $x = d/2$ to $x = d$. Likewise, the backward traveling wave of $x = d/2$ at $t = 0$ moved in opposite direction to reach $x = 0$. The waves are reflected at both ends. Since there is a short circuit at $x = d$, the voltage at $x = d$ must be zero at any time, that is, $v(d,t) = 0$, and therefore $v_f(d,t) = -v_b(d,t)$. At $x = 0$, there is an open circuit where the voltage reaches a maximum and $v_f(d,t) = v_b(d,t)$. Therefore, at the entrance of the line $v(0,t) = 2\, v_b(0,t)$.

Figure 7*c* to 7*f* shows the further evolution of the voltage profile in increments of $T_p/2$. Figure 7*g* shows the voltage $v_l = v(0,t)$ at the entrance to the line as a function of time derived from the changing voltage profiles. The voltage is of sawtooth shape with a period of $4T_p$, that is, the frequency of oscillation of v_l, is $f_l = 1/(4T_p)$. On overhead lines, where the insulating material is air, the waves propagate at the speed of light $c_0 = 3 \times 10^8$ m/s (1.68×10^5 mi/s). At a distance to the fault of $d = 1$ km (0.62 mi), this leads to $f_l = c_0/(4\mathrm{d}) = 75$ kHz. Such a fault is also known as a kilometric fault or short-line fault. For longer distances, the frequency is lower, but the voltage peak is higher as shown by the voltage divider in Fig. 6*d*. Thus, the stress on the circuit breaker due to the rate of voltage rise, which is directly related to the frequency, decreases, while the stress due to the reached voltage amplitude increases. For shorter distances, the opposite is the case. In the case of the kilometric fault, the combination of the stress due to rate of rise and amplitude of the voltage has been found to be particularly severe. Appropriate sizing of power system equipment depends on knowledge of the properties of electric transients.

In **Fig. 8***a*, the TRV waveforms on source and line sides are drawn in the same diagram. The source-side TRV (v_S) describes a harmonic oscillation in the form of a damped sinusoid which, as indicated in Fig. 6*d*, starts from its lowest point given by the voltage divider. In the steady state, after the transient has died out, v_S approximates the ac voltage source e. Since the frequency of the source-side TRV is much lower than the frequency of the line-side TRV, only the start of the source-side TRV is shown. On the line side, damping is taken into account (which is not done in Fig. 7), as evidenced by the decaying peak values of the sawtooth. The resulting TRV $v_B = v_S - v_l$ across the breaker is shown in Fig. 8*b*.

Digital Simulation

The conceptual analysis of the preceding section allows understanding of causes and consequences of electric transients. To the study specific cases and in particular those that involve complex circuits and systems, the use of digital simulation is required. Among the popular simulators are the Electromagnetic Transients Program (EMTP) and others of its type. Key to the success of EMTP is the companion modeling of network elements, where the differential equations characterizing a network element are approximated by a companion model consisting of a resistive circuit. A library of companion element models can so be established as shown in the **table** and derived hereafter.

Companion modeling for lumped-parameter elements. In the continuous time domain, the lumped inductance is described through differential equation (58).

$$v(t) = L\frac{\mathrm{d}i(t)}{\mathrm{d}t} \tag{58}$$

For digital simulation, Eq. (58) must be discretized. Using the trapezoidal integration yields

Library of companion element models

Element type	Representation in continuous-time domain		Companion element model in discrete-time domain	
Lumped resistance	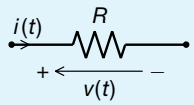	$v(t) = i(t)R$	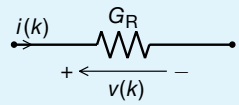	$i(k) = G_R v(k)$ $G_R = 1/R$
Lumped inductance	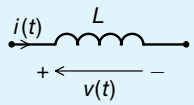	$v(t) = L\frac{di(t)}{dt}$	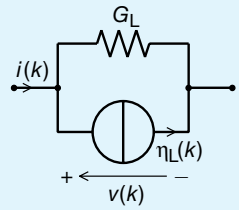	$i(k) = G_L v(k) + \eta_L(k)$ $G_L = \frac{\tau}{2L}$ $\eta_L(k) = i(k-1) + G_L v(k-1)$
Lumped capacitance	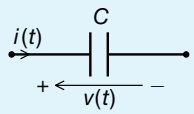	$i(t) = C\frac{dv(t)}{dt}$	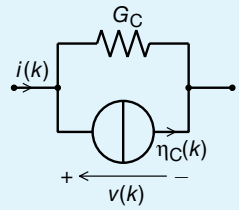	$i(k) = G_C v(k) + \eta_C(k)$ $G_C = \frac{2C}{\tau}$ $\eta_C(k) = -i(k-1) - G_C v(k-1)$
Distributed parameter, lossless line	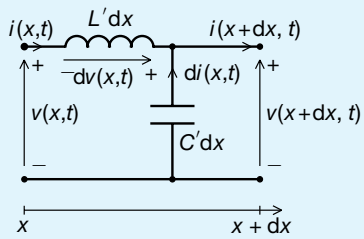	$\frac{\partial v(x,t)}{\partial x} = -L'\frac{\partial i(x,t)}{\partial t}$ $\frac{\partial i(x,t)}{\partial x} = -C'\frac{\partial v(x,t)}{\partial t}$	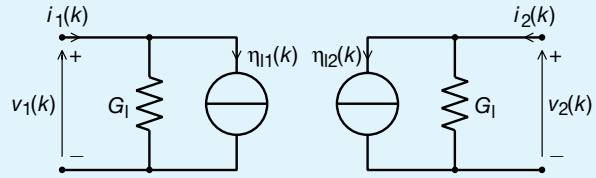	$i_1(k) = G_I v_1(k) + \eta_{I1}(k)$ $\eta_{I1}(k) = -i_2(k-\kappa) - G_I v_2(k-\kappa)$ $i_2(k) = G_I v_2(k) + \eta_{I2}(k)$ $\eta_{I2}(k) = -i_1(k-\kappa) - G_I v_1(k-\kappa)$
	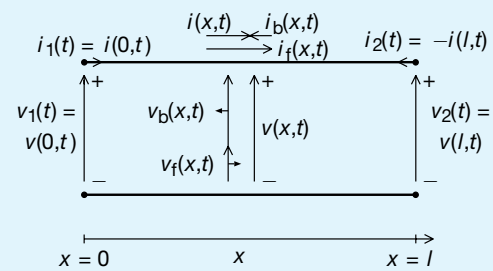			$G_I = \sqrt{C'/L'}$ $\tau = \frac{T_p}{\tau}$ $T_p = l\sqrt{L'C'}$

Eq. (59), where τ is the time step size separating successive discrete time points at which solutions are calculated and k is the time step counter. The higher the frequencies of the electric transients to be studied, the smaller must be the time step size to allow for accurate tracking of the waveforms. Rearranging Eq. (59) yields the companion model of Eq. (60) [also shown in the table], where G_L and $\eta_L(k)$ are given by Eqs. (61) and (62).

$$\tfrac{1}{2}[v(k) + v(k-1)] = L\frac{[i(k) - i(k-1)]}{\tau} \tag{59}$$

$$i(k) = G_L v(k) + \eta_L(k) \tag{60}$$

$$G_L = \frac{\tau}{2L} \tag{61}$$

$$\eta_L(k) = i(k-1) + \frac{\tau}{2L}v(k-1) \tag{62}$$

The companion model simulates the inductance through a conductance and a current source, which provides the initialization for the solution at time step k based on the results of prior time step $k-1$. The same procedure is applied to derive the companion model for the capacitance. In the case of the resistance, no discretization is necessary.

Companion modeling for distributed-parameter elements. A lossless transmission line of length l can be

considered as being composed of an infinite number of incremental segments with incremental length dx for which Eqs. (63) and (64) apply, as indicated in the table.

$$\frac{\partial v(x,t)}{\partial x} = -L' \frac{\partial i(x,t)}{\partial t} \tag{63}$$

$$\frac{\partial i(x,t)}{\partial x} = -C' \frac{\partial v(x,t)}{\partial t} \tag{64}$$

As already established, the voltage at any point is given by the sum of forward and backward traveling voltage waves: $v(x,t) = v_f(x - ct) + v_b(x + ct)$ with c, given by Eq. (65), being equal to the speed of light

$$c = 1/\sqrt{L'C'} \tag{65}$$

in the case of overhead lines. A forward traveling current wave $i_f(x - ct) = v_f(x - ct)/Z_l$ is associated with $v_f(x,t)$, and a backward traveling current wave $i_b(x + ct) = v_b(x + ct)/Z_l$ is associated with $v_b(x,t)$, where the characteristic impedance Z_1, given by Eq. (66),

$$Z_l = \sqrt{L'/C'} \tag{66}$$

relates the associated waves. Using the sign conventions given in the table, the voltages and currents at each point of the line are given by Eqs. (67) and (68).

$$v(x,t) = Z_l i_f(x - ct) + Z_l i_b(x + ct) \tag{67}$$

$$i(x,t) = i_f(x - ct) - i_b(x + ct) \tag{68}$$

From these two equations follow Eqs. (69) and (70).

$$i_f(x - ct) = \frac{1}{2}\left(i(x,t) + \frac{1}{Z_l} v(x,t)\right) \tag{69}$$

$$i_b(x - ct) = \frac{1}{2}\left(-i(x,t) + \frac{1}{Z_l} v(x,t)\right) \tag{70}$$

The argument $(x - ct)$ of i_f is zero for $x = 0$, $t = 0$ and $x = d$, $t = T_p$. Furthermore, at terminal 1, $x = 0$, $i_1(t) = i(0,t)$, and at terminal 2, $x = l$, $i_2(t) = -i(l,t)$. Therefore, from Eqs. (69) and (70) follows Eq. (71), which can be rearranged to Eq. (72),

$$i_1(t - T_p) + \frac{1}{Z_1} v(t - T_p) = -i_2(t) + \frac{1}{Z_1} v_2(t) \tag{71}$$

$$i_2(t) = \frac{1}{Z_1} v_2(t) - \left(i_1(t - T_p) + \frac{1}{Z_1} v_1(t - T_p)\right) \tag{72}$$

Similarly, Eq. (73) holds.

$$i_1(t) = \frac{1}{Z_1} v_1(t) - \left(i_2(t - T_p) + \frac{1}{Z_1} v_2(t - T_p)\right) \tag{73}$$

Discretization readily leads to the companion element model of the lossless line shown in the table.

Companion modeling for the network. The companion element models are connected to the overall companion network model in accordance with the topology of the power system under study. For the companion network model, node equations are formulated to obtain the nodal equation system. *See* NETWORK THEORY.

Solution procedure. Once the nodal equation system is known, the time step counter k is set to zero and a simulation loop is entered. The solution to the nodal equation system is calculated, and the quantities of interest are provided as output. Then, the time step counter is incremented, the loop is reentered, and the operations are repeated. The simulation stops once a termination condition is met. Simulation results can be displayed graphically, as shown in Fig. 8. *See* NUMERICAL ANALYSIS; SIMULATION. Kai Strunz

Bibliography. J. B. Aidala and L. Katz, *Transients in Electric Circuits*, Prentice Hall, 1980; H. W. Dommel, *EMTP Theory Book*, Microtran Power System Analysis Corp., Vancouver, 1992; A. P. S. Meliopoulos, *Power System Grounding and Transients*, Marcel Dekker, New York, 1988; K. Strunz, *Numerical Methods for Real Time Simulation of Electromagnetics in AC/DC Network Systems*, VDI-Verlag, Düsseldorf, 2002; L. van der Sluis, *Transients in Power Systems*, Wiley, Chichester, 2001; N. Watson and J. Arrillaga, *Power Systems Electromagnetic Transients Simulation*, IEE, London, 2003.

Electric vehicle

A ground vehicle propelled by a motor that is powered by electrical energy from rechargeable batteries or other source onboard the vehicle, or from an external source in, on, or above the roadway. Examples are the golf cart, industrial truck and tractor, automobile, delivery van and other on-highway truck, and trolley bus. In common usage, electric vehicle refers to an automotive vehicle in which the propulsion system converts electrical energy stored chemically in a battery into mechanical energy to move the vehicle. This is classed as a battery-only-powered electric vehicle. The other major class is the hybrid-electric vehicle, which has more than one power source. *See* AUTOMOBILE; BUS; TRUCK.

History. Construction of the first electric vehicle is credited to the French inventor and electrical engineer M. Gustave Trouvé, who demonstrated a motorized tricycle powered by lead-acid batteries in 1881. In the United States, Andrew L. Riker is credited with building the first electric vehicle (also a tricycle) in 1890, and by 1891 William Morrison had built the first electric four-wheeler. In France in 1899, a four-wheel electric vehicle driven by Camille Jenatzy became the first car to break 60 mi/h (96 km/h). By then, production of battery-powered vehicles for use as personal transportation, commercial trucks, and buses had already begun.

Electric vehicles, with their instant starting, quiet running, and ease of operation, peaked in their challenge to steam- and gasoline-powered cars in 1912. The limited performance, range, and speed of electric vehicles, plus the need for frequent battery charging, restricted their usefulness and dampened their popularity. By the 1920s, the piston-type internal combustion engine had prevailed as the dominant automotive powerplant. Most production and development work on electric vehicles ended during

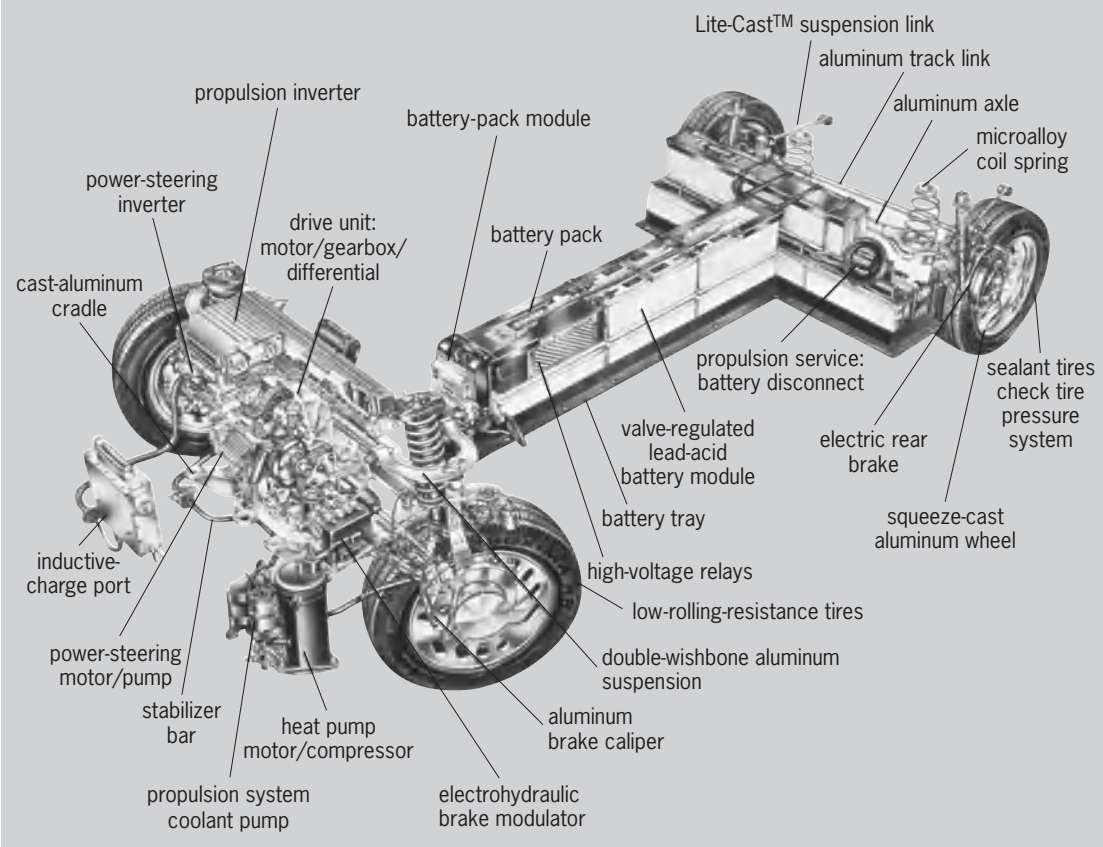

Fig. 1. EV1, a two-seat electric vehicle powered by lead-acid batteries. (*General Motors Corp.*)

the 1930s. *See* AUTOMOTIVE ENGINE; ENGINE; INTERNAL COMBUSTION ENGINE.

In the 1960s, interest revived in electric vehicles as a result of concern with diminishing petroleum reserves, rising cost of crude oil production, and air pollution from the automotive engine that burned gasoline which was refined from crude oil. Over the years, a few electric vehicles had been constructed, usually by converting small light cars and trucks into electric vehicles by removing the engine and fuel tank and installing an electric motor, controls, and batteries. However, during that time no major automotive manufacturer brought out an electric vehicle. *See* AIR POLLUTION; GASOLINE; PETROLEUM.

The Clean Air Act of 1963 and its amendments established limits on emissions from new vehicles sold in the United States. In 1990, the California Air Resources Board decided to further reduce air pollution by mandating (but later rescinding) that 2% of each automaker's sales must have zero emissions in the 1998 model year. This demand for a zero-emission vehicle (ZEV) could be met only by the electric vehicle, which typically was powered by lead-acid batteries. Used in an electric vehicle, lead-acid batteries have two major weaknesses: relatively high weight for the amount of energy stored, and reduced capacity in cold weather.

To help develop a better battery for electric vehicles, the U.S. Advanced Battery Consortium was formed in 1991. The purpose of this partnership among United States automakers and the electric utility industry was to develop advanced batteries capable of providing future generations of electric vehicles with significantly increased range and performance.

In 1996 General Motors began limited marketing of the electric vehicle EV1 (**Fig. 1**). The EV1 was the first specifically designed electric car produced by a major automaker since before World War II. Other automakers also have developed and tested electric vehicles and vehicle conversions powered by lead-acid or advanced batteries.

Battery-only power. The General Motors EV1 is a battery-only-powered vehicle containing 26 lead-acid batteries which are assembled into a T-shaped pack that provides a nominal voltage of 312 volts (**Fig. 2**). These batteries differ from 12-V automotive batteries primarily in duty cycle. In the electric vehicle, the batteries must supply all the energy needs. Except for regenerative braking, there is no onboard charging. The batteries provide the power to propel the vehicle, and to power the lights and all accessories such as air conditioning and radio. As a result, electric-vehicle batteries go through much deeper discharge cycles than an automotive battery, which seldom is discharged more than 5% of its rated capacity. *See* BATTERY.

Because of the increased duty cycle, electric-vehicle batteries can deliver 85% of their charge without damaging the batteries or shortening their useful life. Ideally, this could provide the EV1 and similar electric vehicles with a useful range of 70 mi

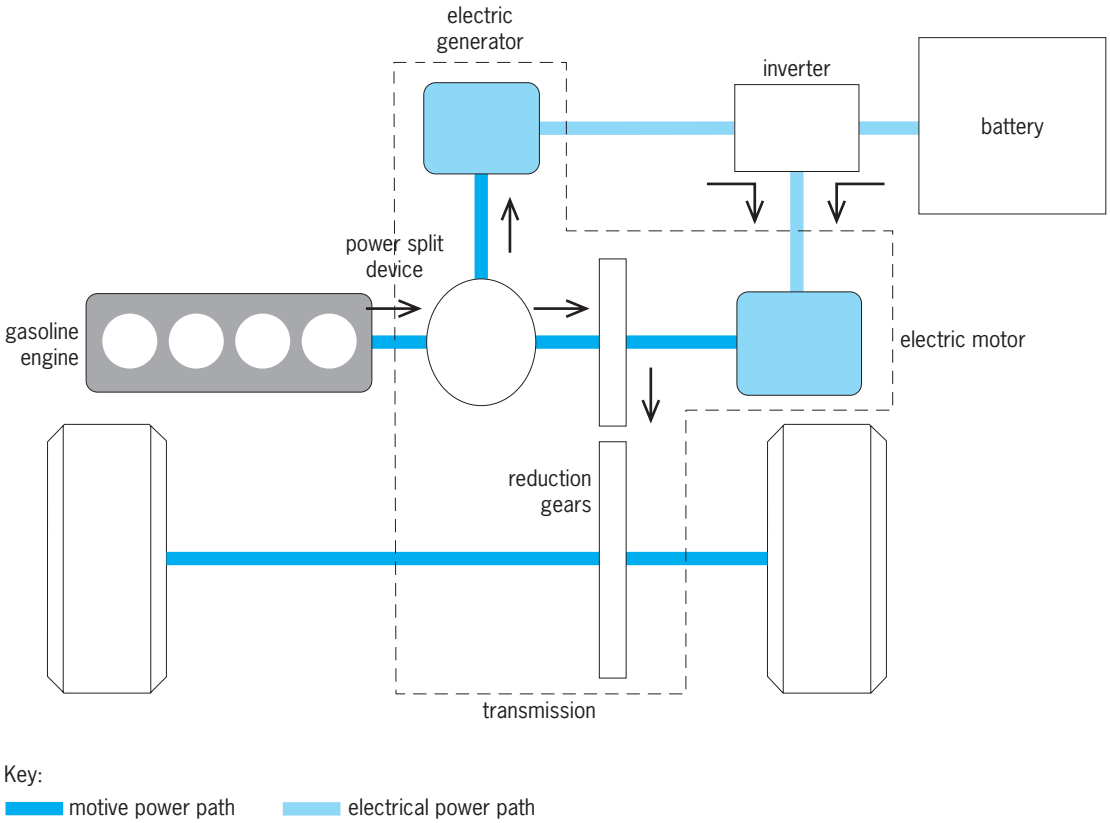

Fig. 2. Power flows through a hybrid vehicle which uses a gasoline engine to propel the vehicle and drive a generator that can operate the electric motor directly or charge the batteries, as necessary. (*Toyota Motor Sales, U.S.A., Inc.*)

(113 km) of city driving or 90 mi (145 km) of highway driving. A top speed of 80 mi/h (130 km/h) may be possible, but at a sacrifice in range, which also is shortened by hilly terrain and use of any electrical equipment on the vehicle.

In addition to lead-acid batteries, other batteries are used in electric vehicles. These include the newer nickel-metal hydride battery and the lithium-ion battery.

Hybrid power. A hybrid electric vehicle has more than one source of power. These sources can be different types of energy storage devices, power converters, and inverters. The first hybrid vehicle is credited to an Italian, Count Felix Carli. In 1894, Carli constructed an electric-powered tricycle that had a system of rubber springs which could release a short burst of additional power when needed.

Although power losses occur each time that energy is converted from one form to another, hybrid drive can be more efficient than a conventional automotive engine. However, having two or more power sources can increase the complexity, cost, and weight of a hybrid vehicle, as well as its manufacturing, safety, emissions, maintenance, and service problems. *See* AUTOMOTIVE ENGINE.

Internal combustion engine. Since the 1890s, major hybrid-vehicle research and development work has focused on adding a small internal combustion engine to an electric vehicle. Typically, the battery-powered motor drives the wheels, while the engine, usually running at constant speed, drives a generator that charges the batteries. Operating the engine at constant speed reduces fuel consumption and produces cleaner exhaust gas than if the engine were larger, operating at variable speed, and providing the sole source of vehicle power.

In some hybrid vehicles, engine power is split (**Fig. 2**). Part of the engine power propels the vehicle, while part drives the generator which can operate the motor directly or charge the batteries, as necessary. Reportedly, engine exhaust emissions of hydrocarbons, carbon monoxide, and oxides of nitrogen are about 10% that of a conventional gasoline-engine vehicle. In addition, fuel efficiency is doubled.

Fuel cell. In this electrochemical device, the reaction between a fuel, such as hydrogen, and an oxidant, such as oxygen or air, converts the chemical energy of the fuel directly into electrical energy. The fuel cell is not a battery and does not store energy, although the fuel cell also has two electrodes separated by an electrolyte. As long as fuel is supplied to one electrode of the fuel cell and oxygen or air to the other, a voltage is produced between the electrodes. When an external circuit connects the electrodes, electrons will flow through the external circuit. Since fuel-cell voltage is less than 1 V, stacks of fuel cells are connected together to provide the needed electrical energy.

When fuel cells are the primary power source in a

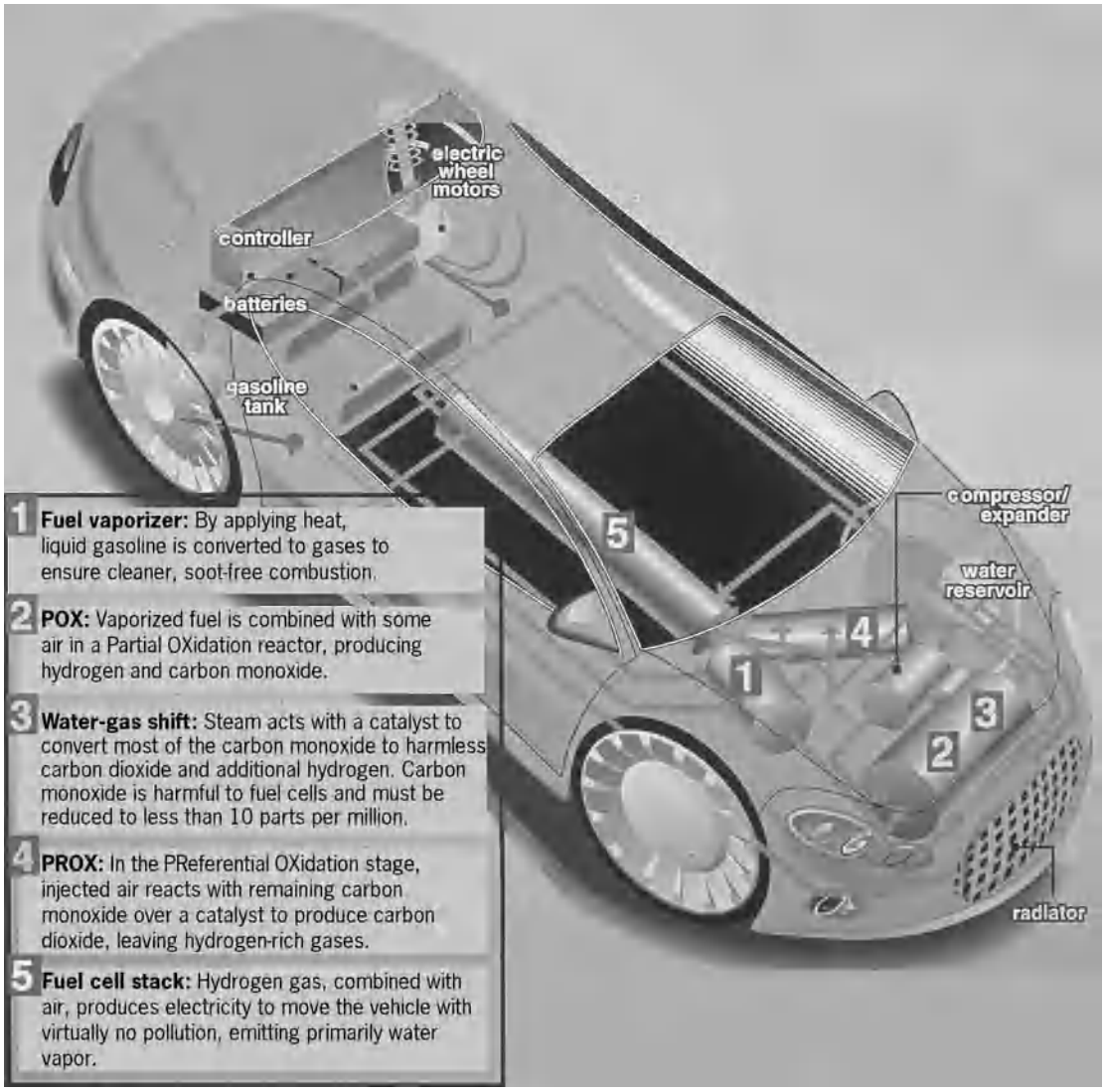

Fig. 3. Layout and five-step process that produces electricity from gasoline, which powers the fuel cells. (*Chrysler Corp.*)

hybrid vehicle, batteries provide secondary power. Fuel cells do not provide immediate output during a cold start. Until the fuel cells reach operating temperature, which may take about 5 min, a battery pack supplies the power for initial startup and vehicle movement. *See* FUEL CELL; HYDROGEN.

Three types of fuel cell under development for electric vehicles are the hydrogen-fueled, the methanol-fueled, and the gasoline-fueled. (1) A hydrogen-fueled cell runs on hydrogen gas stored in pressure tanks carried by the vehicle. Range of the vehicle is determined by the amount of compressed hydrogen that the tanks can hold. When hydrogen is used as the fuel, fuel cell operation produces no significant amounts of unwanted emissions. Water vapor and electricity are the only products. However, widespread use of hydrogen as a near-term vehicle fuel is unlikely because there exists no infrastructure of hydrogen refueling stations which are in place and accessible to the public. (2) To provide the fuel cell with hydrogen gas while avoiding the problems of hydrogen refueling, a methanol-to-hydrogen reformer onboard the vehicle can produce hydrogen gas from liquid methanol. Installed on production models, this would allow motorists to refuel vehicles in the conventional manner at existing service stations through any pump which dispensed methanol. However, use of a reformer to obtain hydrogen lowers vehicle efficiency and creates some emissions of carbon dioxide. (3) A gasoline-to-hydrogen reformer on the vehicle can extract hydrogen gas from gasoline. The hydrogen is then delivered to the fuel cell stack (**Fig. 3**). Use of gasoline could move fuel cell technology years closer to production in automotive vehicles, while reportedly improving fuel efficiency by 50% and emissions by 90%. *See* METHANOL.

Donald L. Anglin

Bibliography. Robert Bosch GmbH, *Automotive Handbook*, 6th ed., 2005; Society of Automotive Engineers, *SAE Handbook*, 3 vols., annually; M. H. Westbrook, *The Electric Car: Development and Future of Battery, Hybrid and Fuel-Cell Cars*, 2001.

Electrical breakdown

A large, usually abrupt rise in electric current in the presence of a small increase in electric voltage. Breakdown may be intentional and controlled or it may be accidental. Lightning is the most familiar example of breakdown.

Breakdown in gases and solids. In a gas, such as the atmosphere, the potential gradient may become high enough to accelerate the naturally present ions to velocities that cause further ionization upon collision with atoms. If the region of ionization does not extend between oppositely charged electrodes, the process is corona discharge. If the region of ionization bridges the gap between electrodes, thereby breaking down the insulation provided by the gas, the process is ionization discharge. When controlled by the ballast of a fluorescent lamp, for example, the process converts electric power to light. In a gas tube the process provides controlled rectification. *See* ARC DISCHARGE; BREAKDOWN POTENTIAL; CORONA DISCHARGE; ELECTRIC SPARK; ELECTRICAL CONDUCTION IN GASES; FLUORESCENT LAMP; GLOW DISCHARGE; IONIZATION.

In a solid, such as an insulator, when the electric field gradient exceeds 10^6 V/cm, valence bonds between atoms are ruptured and current flows. Such a disruptive current heats the solid abruptly; the rate of local temperature rise may fracture the insulator, the high temperature may carbonize or otherwise decompose the insulation, or as occasionally happens when lightning strikes a tree, the heat may ignite the insulator. *See* ELECTRIC INSULATOR; ELECTRICAL INSULATION.

In a semiconductor if the applied backward or reverse potential across a junction reaches a critical level, current increases rapidly with further rise in voltage. This avalanche characteristic is used for voltage regulation in the Zener diode. In a transistor the breakdown sets limits to the maximum instantaneous voltage that can safely be applied between collector and emitter. When the internal space charge extends from collector junction through the base to the emitter junction, even a voltage below the avalanche level can produce a short circuit, in which case the phenomenon is termed punch-through. *See* SEMICONDUCTOR; TRANSISTOR; ZENER DIODE.

Frank H. Rockett

Laser-induced breakdown in gases. For sufficiently high intensities, laser radiation can ionize (or breakdown) the gas atoms or molecules through which it is propagating. The laser radiation intensity required to produce ionization is dependent on the wavelength of the light and on the duration of the laser pulses as well as the gas species and pressure. In general, the breakdown by laser radiation can occur by two different mechanisms, a cascade- or collision-induced ionization, or a multiple-photon absorption process. The most often encountered mechanism is the cascade process, but multiple-photon ionization can be important under certain conditions.

Multiple-photon ionization. Multiple-photon ionization occurs when an atom absorbs directly from the laser radiation field a sufficient number of photons to ionize the atom. The number of photons required is simply the ionization potential of the atom divided by the energy of a photon. For example, the photon energy of the ruby laser at 0.69-micrometer wavelength is 1.78 eV. Therefore, the multiple-photon ionization of argon having an ionization potential of 15.8 eV requires the simultaneous absorption of nine photons. Because of the nature of this process, the ionization is strongly dependent on the laser intensity and wavelength. Experimental measurements of multiple-photon ionization have shown that the threshold for air at pressures below 100 torr (2 psi or 13 kilopascals) is 10^{18} W/m^2 (10^{17} W/ft^2) for a 50-picosecond-duration-pulse, 0.69-μm-wavelength laser. Multiple-photon ionization can be important for laser-induced breakdown at short wavelengths, short pulse durations, and low gas pressures. *See* IONIZATION POTENTIAL; PHOTON.

Cascade breakdown. For the cascade breakdown process, a free electron is accelerated in the radiation field of the laser, but only gains an increment of energy upon collision with atoms. This energy absorption continues until the electron has sufficient energy to ionize an atom by an inelastic collision, producing a second electron. The two electrons then gain additional energy from the field, producing four electrons and so forth. Under breakdown conditions this process continues until the gas is fully ionized. An electron gaining energy from the laser radiation field can also lose energy by processes such as electronic excitation of atoms, or electrons can be lost from the breakdown process by diffusion out of the laser beam or by recombination or by attachment to form negative ions. The rate of ionization necessary to overcome these loss processes then defines the breakdown threshold intensity.

Breakdown threshold. Since high-intensity lasers are usually pulsed devices, an important parameter affecting the breakdown threshold is the duration of the laser pulse. If the laser pulse is long compared to the characteristic electron-loss time, breakdown occurs when the rate of electron production just exceeds the losses; breakdown is determined by the laser intensity and is independent of the laser pulse duration. If the laser pulse is short relative to the characteristic loss times, breakdown occurs when there is sufficient fluence (energy per unit area) in the pulse to generate full ionization of the gas. For air at atmospheric pressure and large-diameter beams (such that electron diffusion is not important), the threshold is fluence-dependent for pulses less than approximately 10^{-7} s and is intensity-dependent for pulses having a longer duration. As an example, for air at atmospheric pressure and laser beams larger in diameter than 4×10^{-3} in. (10^{-2} cm), the threshold fluence for the 10.6-μm carbon dioxide (CO_2) laser wavelength is 10^5 joules/m^2 (10^4 joules/ft^2) for pulse durations less than 10^{-7} s, and for longer pulses the intensity threshold is 3×10^{13} W/m^2 (3×10^{12} W/ft^2). These theoretically predicted values have been experimentally verified.

Conditions for occurrence. Two important factors affecting the occurrence of breakdown are the source of the initiating electron and the presence of aerosol particles. The electron required for initiating the cascade process must be provided by some mechanism other than the cascade, such as photoionization from natural causes or a spark or discharge. For visible-wavelength laser-induced breakdown, the initial electron can be the result of multiple-photon absorption ionization of impurities, but for long-wavelength radiation, certainly 10.6 μm, an external source must provide it. Particulate matter, such as naturally occurring aerosols in the atmosphere, can have a dramatic effect on the breakdown threshold of gases. For particles of 0.1 μm and larger, the interaction of the laser radiation with the particle vaporizes the particle, ionizes the vapor, and subsequently ionizes the surrounding gas. The threshold for this process can be several orders of magnitude lower than the threshold for gas breakdown in the absence of aerosols. *See* ATMOSPHERIC CHEMISTRY.

Pressure dependence. The laser-induced gas breakdown has an interesting pressure dependence. The rate of laser energy absorption by free electrons increases with pressure because of the increase in the electron-atom collision frequency. This energy absorption increase continues as gas pressure increases until the collision frequency equals the oscillatory frequency of the laser radiation field. For pressures above this level, the rate of energy absorption decreases with increasing pressure, and the breakdown approaches that induced by a constant electric field in the high pressure limit. Therefore, there is a minimum in a plot of the threshold as a function of pressure that occurs when the collision frequency equals the oscillatory frequency. As an example, for atmospheric air and 10.6-μm-wavelength radiation this minimum occurs at approximately 44 atm (650 lb/in.^{2} or 4.5 megapascals).

For gas pressures lower than the minimum, the gas breakdown threshold varies as the inverse square of the wavelength; that is, the shorter the wavelength, the more difficult it is to break down the gas. This behavior has been verified for wavelengths from 0.69 to 10.6 μm. The threshold for gas breakdown at pressures above the minimum is independent of the laser wavelength.

Applications. Laser-induced gas breakdown is important as a limitation for propagating high-intensity laser radiation, and as a source for the production of a high-density, high-temperature plasma for a number of applications. Gas breakdown represents the first step in research leading to the study of laser-generated controlled fusion. The plasma produced by laser-induced breakdown can be used as a convenient source for soft x-rays and an intense source of ultraviolet or visible radiation for a fast-rise-time light source. *See* LASER; NUCLEAR FUSION; PLASMA (PHYSICS).

David C. Smith

Bibliography. G. Bekef, *Principles of Laser Plasmas*, 1976; L. A. Dossado and C. J. Fothergill (eds.), *Electrical Breakdown and Degradation in Polymers*, 1992; E. E. Kunhardt and L. H. Luessen (eds.), *Electrical Breakdown and Discharges in Gases*, 1983; Y. P. Raizer, *Laser-Induced Discharge Phenomena*, 1977; Y. P. Raizer and J. E. Allen, *Gas Discharge Physics*, 1997 reprint.

Electrical codes

Systematic bodies of rules governing the practical application, installation, and interconnection of electrically operated equipment, devices, and electrical wiring systems.

National Electrical Code. The basic code used throughout the United States, and a number of other countries throughout the world, for indoor and outdoor electrical installations to supply lighting, motors, appliances, and machines is the National Electrical Code, prepared under the direction of the National Fire Protection Association (NFPA). It is approved by the American National Standards Institute, and the 2005 edition of the code is known as NFPA 70-2005 (ANSI).

The National Electrical Code was originally drawn in 1897 as a result of the united efforts of various insurance, electrical, architectural, and allied interests. The original code was prepared by the National Conference on Standard Electrical Rules, composed of delegates from various interested national associations.

In 1911 the National Conference of Standard Electrical Rules was disbanded, and since then the NFPA has acted as the code's sponsor. Beginning with the 1920 edition, the National Electrical Code has been under the further auspices of the American National Standards Institute with the NFPA continuing in its role as administrative sponsor. Since that date, the NFPA committee that produces the code has been identified as ANSI Standards Committee C1 (formally USAS C1 or ASA C1).

The provisions of the National Electrical Code are under constant review by a number of panels whose members are selected to provide broad representation of electrical, industrial, and public interests. The code is amended in its periodic republication every 3 years or by tentative interim amendments (TIAs) which are announced by bulletins and through the technical press.

The National Electrical Code is purely advisory as far as the National Fire Protection Association is concerned, but it is very widely used for legal regulatory purposes. The code is administered by various local inspection agencies, whose decisions govern its application to individual installations. Local inspectors are largely members of the International Association of Electrical Inspectors. This organization, the National Electrical Manufacturers Association, the National Electrical Contractors Association, the Institute of Electrical and Electronic Engineers, the Edison Electric Institute, the Underwriters' Laboratories, Inc., the International Brotherhood of Electrical Workers, governmental groups, and independent experts all contribute to the

development and application of the National Electrical Code.

Compliance with the provisions of the code can effectively minimize fire and accident hazards in any electrical design. It sets forth requirements that constitute a minimum standard for the framework of electrical design. As stated in its introduction, the code is concerned with the "practical safeguarding of persons and property from hazards arising from the use of electricity." The National Electrical Code is recognized as a legal criterion of safe electrical design and installation in many jurisdictions throughout the country and the world. In some large municipalities within the United States, the National Electrical Code either is accepted with certain modifications or serves as the basis for a local electrical code. Where accepted by the local governing body, the National Electrical Code is used in court litigation and by insurance companies as a basis for insuring buildings.

Other standards. In addition to the National Electrical Code volume itself, other standards and recommended practices are made available in pamphlet form by the National Fire Protection Association, which oversees the promulgation and periodic revision of over 270 standards. These cover such special subjects as hospital operating rooms, municipal fire alarm systems, garages, aircraft hangars, and other equipment with great potential hazards due to improper design.

The National Electrical Safety Code (to be distinguished from the National Electrical Code) is published by the Institute of Electrical and Electronic Engineers, Inc. Designed as ANSI C2, this code presents basic provisions for safeguarding persons from hazards arising from the installation, operation, or maintenance of (1) conductors and equipment in electric supply stations, and (2) overhead and underground electric supply and communications lines. Basically, this code applies to the outdoor circuits of electric utility companies and to similar systems or equipment on commercial and industrial premises.

Municipal codes. The National Electrical Code is incorporated bodily or by reference in many municipal building ordinances, often with additional provisions or restrictions applicable in the particular locality. Some large cities have independent electrical codes; however, the actual provisions in most such codes tend to be basically similar to the National Electrical Code.

Testing of electrical products. Standards on the construction and assembly of many types of electrical equipment, materials, and appliances are set forth in literature issued by the National Electrical Manufacturers Association (NEMA) in conjunction with Underwriters Laboratories, Inc. (UL). The Underwriters Laboratories examines, tests, and determines the suitability of materials and equipment to be used according to code regulations. Each year, it publishes three volumes listing commercially available electrical products which have been found acceptable with reference to fire and accident hazards and which conform with the application and installation requirements of the code. The three volumes are titled *Electrical Construction Materials Directory, Electrical Appliance and Utilization Equipment Directory*, and *Hazardous Location Equipment Directory*. The Underwriters' Laboratories publishes other literature such as *Gas and Oil Equipment Directory* and *Fire Protection Equipment Directory*, dealing with special equipment involving hazard to life or property.

Testing of electrical equipment and products is permitted to be performed by any "nationally recognized testing laboratory," as stipulated by the Occupational Health and Safety Administration (OSHA). Underwriters Laboratories is one such company. Tests performed by other such companies are conducted in accordance with the applicable NEMA/UL standard.

Administration. Electrical codes are administered locally by inspectors. The inspectors, who may review plans and specifications, make periodic site visits during the construction process to visually examine electrical work during installation and after the work is completed to ensure compliance with applicable rules or ordinances.

Electrical inspection bureaus are maintained in many cities by the city itself. In other cities, a third-party inspection agency, such as the New York Board of Fire Underwriters, which provides electrical inspectors for any jurisdiction within New York State that wishes to avail itself of the Board's services, is used. Typically where third-party inspectors are used, the companies recognized to perform inspections within a given jurisdiction are identified by the local building department. In communities where codes are enforced by ordinance, inspections may be performed by municipal electrical inspectors. Utility inspectors examine the service entrance and metering installation for compliance with prevailing utility regulations.

Federal and state buildings are usually inspected by authorized government electrical inspectors. In these instances inspection includes both safety consideration and the requirements of the particular job specifications. Other specification (by underwriters or municipal inspectors) is often waived. *See* ELECTRONIC EQUIPMENT GROUNDING; GROUNDING; WIRING.

J. F. McPartland; Brian J. McPartland

Bibliography. J. F. McPartland, *Handbook of Practical Electrical Design*, 3d ed., 1999; J. F. McPartland, *National Electrical Code Handbook*, 25th ed., 2005; *National Electrical Safety Code*, NFPA 70-2005 (ANSI), 2005.

Electrical communications

That branch of electrical engineering dealing with the transmission and reception of information. Information can be transmitted over many different types of pathways, such as satellite channels, underwater acoustic channels, telephone cables, and

fiber-optic links. Characteristically, any communications link is noisy. The receiver never receives the information-bearing waveform as it was originally transmitted. Rather, what is received is, at best, the sum of what was transmitted and noise. In reality, what is more likely to be received is a distorted version of what was transmitted, with noise and perhaps interference. Consequently, the design and implementation of a communications link are dependent upon statistical signal-processing techniques in order to provide the most efficient extraction of the desired information from the received waveform. *See* COMMUNICATIONS CABLE; COMMUNICATIONS SATELLITE; DISTORTION (ELECTRONIC CIRCUITS); ELECTRICAL INTERFERENCE; ELECTRICAL NOISE; MICROWAVE; OPTICAL COMMUNICATIONS; RADIO-WAVE PROPAGATION; TELEPHONE SERVICE; TRANSMISSION LINES; UNDERWATER SOUND; WAVEGUIDE.

Broadly speaking, there are two basic classes of communication waveforms, those involving analog modulation and those involving digital modulation. The former type implies that the modulation process allows the actual information signal to modulate a high-frequency carrier for efficient transmission over a channel; this is achieved by using the continuum of amplitude values of an analog waveform. Examples of analog modulation systems include amplitude-modulation (AM) and frequency-modulation (FM) systems, as well as a variety of others such as single-sideband (SSB), double-sideband (DSB), and vestigial-sideband (VSB) systems. In digital modulation systems, the initial information-bearing waveform, assuming it is in analog form (such as voice or video), is first sampled, then quantized, and finally encoded in a digital format for carrier modulation and transmission over the channel. *See* MODULATION.

Analog modulation. Analog communication systems are often characterized as linear or nonlinear modulation systems. Linear systems include AM and single-, double-, and vestigal-sideband systems, whereas nonlinear systems include FM and phase-modulation (PM). Among the linear analog systems, AM is probably the most widely used, because its receiver structure can be implemented in the form of a simple envelope detector. However, AM is very inefficient from a power point of view, because its signal structure results in part of its total transmitted power being wasted in an unmodulated carrier instead of all its power going to the information-bearing portion of the waveform. *See* AMPLITUDE MODULATION; FREQUENCY MODULATION; PHASE MODULATION; SINGLE SIDEBAND.

This problem is remedied with either single- or double-sideband systems, but their receivers are more complex than the AM receiver. Furthermore, the single-sideband system, which has the most complex transmitter, requires only half the bandwidth of the other two, thereby giving it yet another advantage over AM. The vestigal-sideband system is essentially a compromise between the double sideband and the single sideband, for although implementation with it is somewhat easier than with a single sideband, its bandwidth is greater than that of the single sideband but less than that of the double sideband.

The disadvantage of any linear analog modulation system is that its ultimate performance, as measured by the ratio of the signal power of the final demodulated waveform to the noise power of the demodulated waveform, called the signal-to-noise ratio (SNR), is limited by this ratio at the input to the receiver. In nonlinear analog communications, this is not necessarily the case. There the output signal-to-noise ratio can be made significantly higher than that at the input, and consequently significantly higher than the output signal-to-noise ratio of a linear modulation scheme. However, the price for this improvement in performance is increased transmission bandwidth. *See* SIGNAL-TO-NOISE RATIO.

Digital modulation. In digital communications, it is necessary to convert the information into a digital format. At times, the information might originate in an appropriate digital format, as in the transfer of computer files from one computer to another (so-called computer communications). However, in many other instances, the information originates in analog form. This is the case for both voice communications and video transmission. Therefore, some means of digitizing the analog waveform is required.

The classical technique for accomplishing this is probably the use of pulse-code modulation (PCM). There the analog signal is first sampled as a minimum at what is known as the Nyquist frequency (that is, at a rate corresponding to twice the bandwidth of the waveform). The samples are quantized and encoded into a binary word. The digits of this word are used to modulate a carrier for transmission over the communications channel. *See* PULSE MODULATION.

Pulse-code modulation is one example of the general technique of source encoding. Other schemes attempt to remove as much redundancy as possible from the message in order to minimize the number of binary digits needed to reproduce it. Examples of practical source-coding schemes are delta modulation and differential pulse-code modulation. In both of these techniques, an estimate of the current sample is made, the error between the sample and its estimate is ascertained, and this error signal is encoded into binary digits to be modulated on the carrier.

Once the waveform is in digital format, many modulation techniques can be used. These consist of both binary schemes, whereby one of two possible signals is transmitted across the channel, and M-ary schemes with $M > 2$, whereby one of M signals is used to convey the information in any given signaling interval. Typically, $M = 2^n$, for some integer n, and if all M signals are equally likely (that is, the a priori probability of transmitting any one of them is $1/M$), then each transmitted symbol conveys n bits of information. Thus, a binary system conveys one bit of information each time a symbol is transmitted, a quarternary system conveys two bits of information each time a symbol is sent, and so on.

For binary systems, the most widely used modulations are phase-shift keying and frequency-shift

keying. In the former scheme, a radio-frequency carrier has its phase changed to 0 and π radians each time a binary 1 or a binary 0 is to be transmitted, respectively. In frequency-shift keying, the information is transmitted by sending one of two possible frequencies during each signaling interval.

Many modulations can be used with M-ary schemes. Perhaps conceptually the simplest are the$M$-ary versions of binary phase-shift or frequency-shift keying, for example, the transmission of a carrier which uses one of M possible phases during each signaling interval. However, other schemes make use of a combination of both amplitude and phase modulation; these techniques are termed quadrature amplitude modulation (QAM).

Communication theory. The theoretical foundations of digital communications stem from the results of statistical decision theory and information theory. By using concepts developed in the former discipline, it is possible, in principle, to design an optimal receiver for a given transmitted waveform and a given noise environment. For example, in the detection of one of M known signals, all of which have equal energy and equal probability to be transmitted a priori, in a noise environment that contains only radio-frequency receiver noise, one form of an optimum receiver consists of a parallel filter bank of M matched filters with one filter matched to each of the M possible transmitted waveforms. A matched filter has an impulse response which is the time inverse of the signal to which it is matched. *See* DECISION THEORY; ELECTRIC FILTER.

Information theory has as its key results the ultimate limitations and performance capabilities of any communication system. The most important of these is embodied in the so-called noisy channel-coding theorem, which states that, under a fairly broad set of conditions, it is possible to achieve perfectly reliable communication over a noisy (unreliable) channel. This is accomplished with the help of error-correction coding, a technique of adding a controlled amount of redundancy to the information sequence in order to correct the errors that are made when the data are transmitted over the noisy channel. This seemingly astounding result is not especially practical, since it leads to system characteristics which are not realistic, such as infinite delay before the message can be decoded. However, its tremendous significance is due to the fact that it shows the power of error-correction coding in the design of communication systems. *See* INFORMATION THEORY.

Multiple-access techniques. In many communication systems, multiple sources and multiple destinations are present, and the manner in which accessing the channel is achieved becomes important. Perhaps the most common examples of such multiple-access links are those used by commercial radio and television stations. These systems operate by assigning to each transmitted waveform a distinct frequency band which is adjacent to but nominally not overlapping with its neighboring bands. In this way, there is approximately no interference among the various users. This type of operation, known as frequency-division multiple accessing (FDMA), can be used with either analog or digital modulation formats. *See* RADIO; TELEVISION.

Because of certain problems that arise with frequency-division multiple accessing, such as the generation of intermodulation products when such a waveform is passed through a nonlinear amplifier, other types of multiple-accessing schemes have been devised. (An intermodulation product corresponds to a frequency generated at the output of a nonlinear device due to the nonlinear interaction of two or more frequencies at the input of the device.) In particular, when digital modulation is being employed, both time-division multiple access (TDMA) and code-division multiple access (CDMA) are possible alternatives.

In time-division multiple access, all users occupy the full bandwidth of the channel, but they transmit sequentially in time, and only one user has access to the channel at any given time. The system operates as follows. Each user is assigned a certain amount of channel time, and when the time slot of one user ends, the next user begins. After all the users have had an opportunity to transmit, the channel returns to the first user and the process is repeated. The time for the channel to return to a given user is referred to as the frame time. A predetermined pattern of bits, known as a frame synchronization pattern, is transmitted at the beginning of each frame. This pattern is known by each user, and it allows the users to synchronize to it so that they will know to transmit their data only during the allocated time interval, thus avoiding interference between the users. The advantage of the system over frequency-division multiple accessing is that the intermodulation problem disappears since, even if the channel is nonlinear, only one signal is on it at any time.

The technique of code-division multiple access operates by allowing multiple users to individually occupy the full bandwidth of the channel while simultaneously transmitting. This is accomplished by superimposing on the data of each user a wide-bandwidth digital waveform which identifies that user. Typically, a different identification sequence is assigned to every user. At the receiver this sequence for a given user allows the receiver to demodulate the data of that user while rejecting the waveforms of the other users.

Code-division multiple access is an application of a more general technique known as spread spectrum communication, in which the transmission bandwidth is purposely made much larger than the information bandwidth. This has the advantage that the resulting waveform is much less vulnerable to interference than a conventional waveform. Since the interference could be due to any one of a variety of sources, such as multipath or intentional jamming, this sytem has an inherent natural interference immunity which both frequency-division and time-division multiple access lack. *See* SPREAD SPECTRUM COMMUNICATION.

Other types of multiple-accessing schemes are designed to operate over a specific channel. One

example is space-division multiple access, which is used over a satellite communications link. In this scheme, spot beams on the satellite are used to illuminate different areas on the earth, so that the same frequency band can be used simultaneously by each of the spot beams without causing any interference. *See* MULTIPLEXING AND MULTIPLE ACCESS.

Random access. In many cases, the number of potential users is much greater than the number which can be simultaneously accommodated on the channel. However, if the percentage of time employed by each user is statistically very small, the users can compete for access to the channel. When a given user has a message ready for transmission, most of the time a vacant slot on the channel will be found, allowing data transmission. However, at times, all available slots on the channel are taken, and the user has to delay sending the message.

Systems that operate in this manner are referred to as random-access systems, and are typical of computer communication networks. Depending on the geographical size of such networks, they have come to have their own specialized names. For example, they are known as local-area networks (LANs) if they are concentrated over an area roughly the size of a few blocks or less. If the terminals which are furthest apart are within a few miles of one another, they are referred to as metropolitan-area networks (MANs). Networks that span still larger geographical distances are called wide-area networks (WANs). *See* DATA COMMUNICATIONS; LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

Laurence B. Milstein

Bibliography. M. Schwartz, *Information Transmission, Modulation and Noise*, 4th ed., 1990; H. Taub, *Principles of Communication Systems*, 2d ed., 1986; R. E. Ziemer and W. H. Tranter, *Principles of Communications*, 3d ed., 1990.

Electrical conduction in gases

The process by means of which a net charge is transported through a gaseous medium. It encompasses a variety of effects and modes of conduction, ranging from the Townsend discharge at one extreme to the arc discharge at the other. The current in these two cases ranges from a fraction of 1 microampere in the first to thousands of amperes in the second. It covers a pressure range from less than 10^{-4} atm (10 pascals) to greater than 1 atm (10^5 pascals). *See* ARC DISCHARGE; TOWNSEND DISCHARGE.

In general, the feature which distinguishes gaseous conduction from conduction in a solid or liquid is the active part which the medium plays in the process. Not only does the gas permit the drift of free charges from one electrode to the other, but the gas itself may be ionized to produce other charges which can interact with the electrodes to liberate additional charges. Quite apparently, the current voltage characteristic may be nonlinear and multivalued. *See* ELECTRICAL CONDUCTIVITY OF METALS; ELECTROLYTIC CONDUCTANCE; SEMICONDUCTOR.

The applications of the effects encountered in this area are of significant commercial and scientific value. A few commercial applications are thyratrons, gaseous rectifiers, ignitrons, glow tubes, and gas-filled phototubes. These tubes are used in power supplies, control circuits, pulse production, voltage regulators, and heavy-duty applications such as welders. In addition, there are gaseous conduction devices widely used in research problems. Some of these are ion sources for mass spectrometers and nuclear accelerators, ionization vacuum gages, radiation detection and measurement instruments, and thermonuclear devices for the production of power.

The discussion of this complicated process will be divided into two parts. The first will deal with the basic effects involved, including production and loss of charges within the region and the motion of charges in the gas. The second part will deal with the mechanism of conduction.

Basic effects. To produce gaseous conduction, two conditions must obtain. First, there must be a source of free charges. Second, there must be an electric field to produce a directed motion of these charges. Considering the first of these, one finds that the free charge concentration is a result of a number of processes which produce and remove charges.

Sources of free charges. In many gaseous devices, a thermionic-emission cathode is included. The process of electron emission from a heated electrode is well known. Closely related to this as a source of electrons is field emission. Here, a strong positive field at a metallic surface lowers the barrier for electron emission. Thus, the electron current from a surface at a given temperature may be significantly increased. *See* THERMIONIC EMISSION.

Both of these effects result in electron production. Another effect, photon absorption, may result only in electrons if the absorber is a solid. However, if the photon interacts with a gas molecule or atom, ionization may result and both an electron and a positive ion can be obtained. The photon may come from some external source or it may be a secondary effect of the gaseous conduction. It may have a wavelength in the visible, ultraviolet, or x-ray region.

Conduction in flames is largely a result of the thermal production of ionization. This is a specialized field which has long been of interest in chemistry and combustion studies. To produce appreciable thermal ionization, the temperature must be high, as in a flame. If the effective temperature is known, the ionization concentration may be determined from statistical mechanics. Thermal ionization is also of tremendous importance in devices for production of power by thermonuclear processes, and in ion-propulsion equipment. A special form of this is surface ionization, in which a hot surface may cause ionization of a gas atom that comes in contact with it.

Another source of ionization is particle radiation due to cosmic rays, radioactive material in the walls or in the gas, or particles produced from an external source. These particles may then produce an ionized track in the gas. An example is the ionization produced by an α-particle from a polonium source in an ionization chamber.

In most of these methods of charge production, the sources are primary. That is, the presence of other free charges is not important in the production of ionization or electrons. Other processes are secondary in origin, although they may be of prime importance. It was pointed out that photons could originate either externally or as a secondary effect. Field emission could be a secondary effect, also. Other methods of ionization, however, are generally thought of as being secondary in origin. Ionization of the gas by electron impact is such a case. Here free electrons may gain enough energy in an electric field to interact with an atomic or molecular electron to produce an ion pair.

Cumulative ionization is an extension of ionization by impact. If the original electron and its offspring gain enough energy so each may produce another electron, and if this process is repeated over and over, the result is called an avalanche, and the ionization thus produced is referred to as cumulative ionization. This is the basis for particle detection in some ionization devices. *See* GEIGER-MÜLLER COUNTER; IONIZATION CHAMBER.

Another secondary source is electron emission from either electron or positive ion bombardment of a surface. This should not be confused with thermionic emission resulting from heating under bombardment. *See* SECONDARY EMISSION.

Other sources which may be important are atomic collisions, sputtering, and collisions of the second kind. In the first case, an atom or heavy ion may collide with an atom to produce ionization. This is quite unlikely until an energy of many times the ionization energy is obtained. The second is somewhat analogous to secondary electron emission. Here the positive ions strike a surface and knock out atoms or groups of atoms. Some of these come off as ions. In the third case, an excited atom may interact with an atom or molecule which is chemically different and has an ionization potential lower than the excitation potential of the excited atom. The result may be the decay of the excited state with ionization of the struck molecule or atom. Symbolically, this is shown by reaction (1), where A* is the excited atom, B the

$$\mathrm{A}^* + \mathrm{B} \rightarrow \mathrm{A} + \mathrm{B}^+ + e^- \qquad (1)$$

struck atom or molecule, and e^- an electron. *See* EXCITATION POTENTIAL; IONIZATION POTENTIAL; SPUTTERING.

Free-charge removal. The net free-charge concentration is a balance between charge-production and charge-removal processes. Recombination is one such process. Here, an electron or heavy negative ion and a positive ion may recombine. The energy transition may appear as electromagnetic radiation or may be carried off by a third body, if one is present. There are a wide variety of conditions which may lead to recombination. Where the temperature and electric field are high, the recombination will occur predominantly at the walls.

The method of charge removal is important from the aspect of conduction, however. If the charges move to the appropriate electrodes under the influence of the field and there recombine, then they contribute to the current. If they simply diffuse to the walls and recombine there with ions of the opposite sign, or if they recombine in the gas volume, they may not appear as part of the external current.

Motion of the charges. The motion of the charges within the gas will be largely influenced by the potential function. For the usual regular geometries, this could be calculated in principle if there were no charges present. However, in a gas with free charges distributed throughout, the problem is quite different. The charges modify the charge-free potential, but the potential itself determines how the charge will move. The motion of the charges further modifies the potential and so on. Although the situation can be described physically by Poisson's equation, it is generally impractical to carry out an analysis. As a practical result, the potential function must be determined by measurements which are made with probes. This requires careful procedures to obtain significant results.

Diffusion of ions. Ion diffusion is a type of random motion which is always present and is the result of thermal or agitation energy. The randomness of the motion is brought about by the many collisions with molecules and other ions. A great difference exists in the motions of electrons and heavy ions. Because of low mass, electrons are easily deflected, so they move erratically. They diffuse badly, and follow field lines only generally. Again because the mass is small, an electron can give up appreciable energy only in an inelastic collision, in which excitation or ionization takes place. Hence, electron agitation and diffusion will be much greater in a pure inert gas than in a gas having many low-energy molecular states. Conversely, heavy ions exchange energy effectively at every collision. Diffusion is much less, so that they follow the electric field lines more closely than do electrons.

Mechanism of conduction. The ionic mobility μ relates drift velocity v to electric field X by Eq. (2). For

$$v = \mu X \qquad (2)$$

electrons, the mobility is high, and a drift velocity of 10^6 cm/s or greater may be obtained. The electronic mobility is not a true constant, but varies with field, pressure, temperature, and gas composition. For heavy ions, the mobility is much more nearly constant, but is still dependent on these quantities to some extent. Drift velocities are usually of the order of 10^3–10^4 cm/s. Thus, in a typical conduction device, an electron may move from one electrode to the other before a heavy ion is displaced appreciably.

It would appear from the foregoing that if accurate information about the important processes existed, one could predict the characteristics of the conduction process under given conditions. Unfortunately this is not the case. Generally, the situation is so complicated that the theory can yield only qualitative

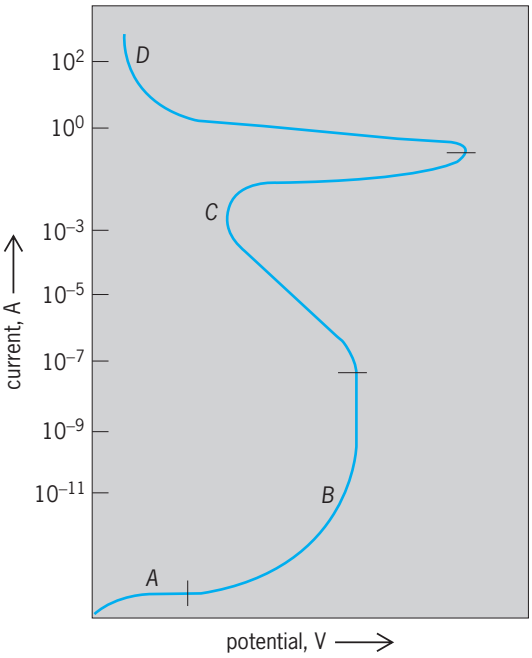

Current-potential characteristics for a two-electrode device with constant pressure.

predictions. Accordingly, most of the information concerning the various forms of gaseous conduction is empirical. In the present description, it will be possible to mention the main features of a few of these modes.

The **illustration** shows a sample voltage-current characteristic for a two-electrode device with constant pressure. It is assumed that there is a constant source of ionization which could be any of the primary sources previously discussed. In region *A*, the current first rises and then over a limited range is relatively constant as the voltage across the electrodes is increased. The initial rise is the result of the collection of charges which were either recombining or diffusing to the walls. The nearly constant current region is the result of the collection of almost all of the charges.

In region *B*, further increase in voltage produces an increase in current. Here, ionization by electron impact is occurring. The situation is described by specifying that each free electron makes α additional ion pairs in traveling 1 cm in the direction of the field. The number of ion pairs produced per second in 1 cm at a distance x from the cathode (assuming parallel plate electrodes) is given by Eq. (3),

$$n = n_0 e^{\alpha x} \tag{3}$$

where n_0 is a constant depending on the initial number of electrons. This is a form of the Townsend equation, and α is the first Townsend coefficient. In the region *B* in the illustration, the increase in current represents an increase in α. Near the end of this region, the current increases more rapidly with applied field. Here, additional effects are taking place, such as the photoelectric process and secondary emission. This situation is described by Eq. (4), where β is the second Townsend coefficient,

$$i = i_0 \frac{(\alpha - \beta)\, e^{(\alpha-\beta)x}}{\alpha - \beta\, e^{(\alpha-\beta)x}} \tag{4}$$

i_0 is the initial electron current at the cathode, and i is the anode current as a function of plate separation x; β is also a function of electric field.

At the end of the region, the slope becomes infinite, and if the external resistance is not too large, the current will jump in a discontinuous fashion. The transition is referred to as a spark, and the potential at which it occurs is the breakdown or sparking potential. The region *B* is called a Townsend discharge and is not self-sustained. Thus, if the source of primary ionization were removed, the discharge would cease. *See* BREAKDOWN POTENTIAL; ELECTRIC SPARK.

As the potential reaches the sparking potential, a transition occurs to the region *C*. This is the self-sustained glow discharge region. Over an extensive current range, the voltage drop remains substantially constant. During the current increase, a glow occurs at the cathode, and at the upper end of the range, the cathode is completely covered. At this point, a further current increase can be achieved only if the potential drop across the discharge is increased. This portion of the characteristic is known as the abnormal glow. Throughout this portion of the discharge characteristic curve, secondary effects are quite important. Particularly vital are the effects of cumulative ionization and secondary emission at the cathode. *See* GLOW DISCHARGE.

Further increase in current leads to another mode of discharge, the arc. This is shown as region *D* in the illustration. Characteristic of this mode is the low cathode potential fall and the very high current density. It is generally felt that the predominant effect in the production of the large number of electrons at the cathode necessary for the arc is thermionic emission. This is consistent with the very high temperatures known to exist either generally or locally on the cathode. Although the arc type of discharge has very great commercial value, the mechanism of its operation is not very well understood.

In addition to these general types of conduction, there are very special cases of considerable interest. Some of these are the corona discharge, radiofrequency or electrodeless discharge, hot-cathode discharge, and discharges in the presence of a magnetic field.

Glenn H. Miller

Bibliography. M. N. Hirsh and H. J. Oskan (eds.), *Gaseous Electronics*, vol. 1: *Electrical Discharges*, 1978; E. E. Kunhardt and L. H. Luessen (eds.), *Electrical Breakdown by Discharges in Gases*, 2 vols., 1983; Y. P. Raizer and J. E. Allen, *Gas Discharge Physics*, 1997.

Electrical conductivity of metals

The property of a metal that measures its ability to conduct electricity, following Ohm's law. Electrical conductivity σ (measured in $\Omega^{-1} \cdot m^{-1}$) is the

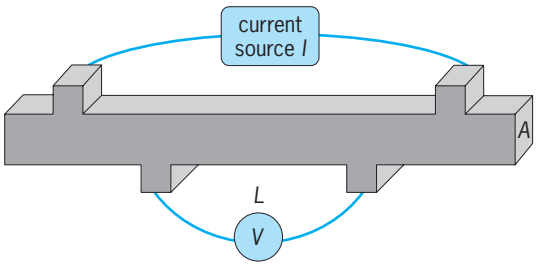

Fig. 1. Schematic of a four-probe measurement of electrical conductivity. Current *I* is fed through the outer leads, and voltage drop *V* is measured on the inner leads. In this way the contact potential drop experienced by the applied current is localized at the junctions with the outer leads. Measuring *V* with minimal current through the voltmeter minimizes the contact potential contribution to the measured conductance $G = I/V$.

reciprocal of the resistivity ρ [in $\Omega \cdot$ m, where resistance R in Ohms (Ω) is voltage drop V in volts (V) divided by current I in amperes (A)]. Ohm's law, $V = IR$, may also be written in the form of Eq. (1), where j is current density (current per

$$j = \sigma E \qquad \text{or} \qquad E = \rho j \tag{1}$$

unit area, I/A, measured in A/m^2) and $E = V/L$ is electric field or electrical potential gradient, measured in V/m, where V is the voltage drop measured across a length L of material (**Fig. 1**). The conductivity σ is an intrinsic property of a pure material, related to the measured conductance $G = I/V$ via $G = \sigma A/L$, just as resistance $R = V/I$ relates to resistivity ρ via $R = \rho L/A$. Thus a large-gauge conductor [large area (A)] of short length has a high conductance, but the physical dimensions do not affect the conductivity σ or the resistivity ρ. Positive current flows from higher to lower voltage, and σ is never negative. *See* ELECTRICAL RESISTANCE; ELECTRICAL RESISTIVITY; OHM'S LAW.

Metals versus insulators. Metals have large conductivity (typically greater than $10^5\ \Omega^{-1} \cdot \text{m}^{-1}$ but never greater at room temperature than that of silver, $0.66 \times 10^8\ \Omega^{-1} \cdot \text{m}^{-1}$.) As temperature T is lowered, conductivity almost always increases in metals (**Fig. 2**). These properties contrast with semiconductors and insulators, which have smaller conductivity and different temperature dependence. The reason concerns the spectrum of quantum excited states in matter (**Fig. 3**). Outer electrons (occupying frontier orbitals in chemical language) have a continuous range of energy levels in solids, called a band of quantum states. In metals the highest-energy occupied state and the least-energy empty state differ only infinitesimally in energy, whereas in semiconductors or insulators they are separated by a significant energy gap. In both semiconductors and metals the states that correspond to the frontier orbitals are mobile, meaning that they permit electrons to travel from atom to atom. In semiconductors the energy gap is moderate (typically 0.5–3.0 eV). If electrons are introduced into empty states above the gap, a semiconductor behaves much like a metal. In an insulator the energy gap is larger, 5 eV or more, and electrons typically get trapped on atoms or defects, rendering them comparatively immobile. *See* BAND THEORY OF SOLIDS; ELECTRIC INSULATOR; HOLE STATES IN SOLIDS; SEMICONDUCTOR.

Propagation of electrical current. To outside observers, electrical current flow in a wire resembles water flow in a pipe, but the analogy does not explain what is going on internally. In water flow, molecules have a net drift velocity, which is preserved by intermolecular collisions. Locally there is Galilean invariance, which means that an internal observer moving with the hydrodynamic current and observing individual molecules would not be able to detect the current. In a metal the electrons disengage from their parent positive ions. The ions vibrate around fixed positions, not participating in the current. When electrical current j flows, the electrons have a net drift velocity which can be detected by a local observer watching ions as well as electrons. The system does not have Galilean invariance. Collisions of electrons do not conserve j, and act to restore electrons to the $j = 0$ state of thermal equilibrium, not the moving local equilibrium of hydrodynamics. This is the source of resistance in metals. Electrical current is less "collective" than hydrodynamic current. However, at very low temperature many metals become superconductors, seen in the data of Fig. 2. This is a different state of electronic matter in which electrons behave more collectively than they do in the "normal" or higher-temperature resistive state of metallic electronic matter. The collective (and quantum-mechanically phase-coherent) superconducting state allows $j \neq 0$ with $E = 0$, thus making $\sigma = \infty$. *See* SUPERCONDUCTIVITY.

Electrons in quantum theory are waves as well as particles, and can diffract around the crystalline

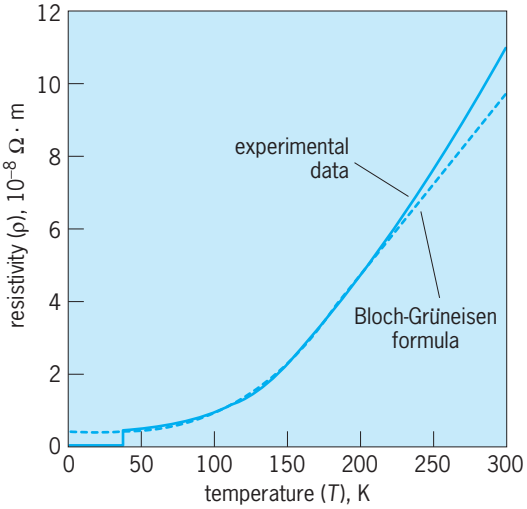

Fig. 2. Electrical conductivity of magnesium diboride (MgB_2), plotted, as is customary for metals, as resistivity (ρ) versus temperature (T). The unit $10^{-8}\ \Omega \cdot$ m is often written as 1 micro-ohm cm or 1 $\mu\Omega$-cm. Below $T = 40$ K $= -233°$C, MgB_2 is a superconductor with $\rho = 0$. A fit using Bloch-Grünisen theory, explained in the text, is also shown. (*Data from Z. X. Ye et al., Electron scattering dependence of dendritic magnetic instability in superconducting MgB_2 films, Appl. Phys. Lett., 86:5284–5286, 2004*)

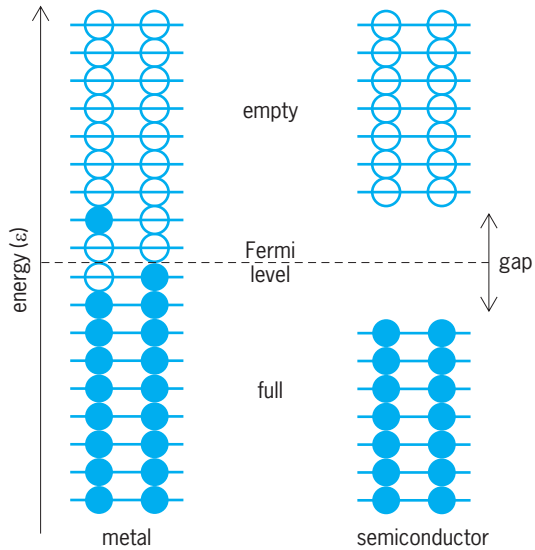

Fig. 3. Schematic energy spectrum in a metal versus a semiconductor. Filled and empty circles represent full and empty quantum states, two per level corresponding to the two spin orientations, just as in an atom. Adjacent levels are infinitesimally close except at the Fermi level of a semiconductor, which lies in the energy gap where there are no quantum states. The metal states are shown with a filled state above the Fermi level and a vacant state or hole below the Fermi level, as would happen by ordinary thermal excitation at room temperature. Semiconductors are likely to have few thermal excitations. In such a state the conductivity σ would be close to zero.

lattice of ions. In perfect crystals, if ions were not vibrating, the electron waves would remain coherent over the whole sample size, and conductivity would be infinite. The actual finite value occurs because electrons scatter from defects in the lattice, from atomic vibrations, and from each other. The last source of scattering is small, and the other sources are minimized in pure crystals at low temperature. Unlike molecules in water, which collide frequently, electrons in good metals collide infrequently and have long "mean free paths" or distances of travel between collisions. The collision frequency for water molecules at room temperature (300 K), denoted $1/\tau$, is about 6×10^{12}/s. This comes from dividing the thermal velocity, given by Eq. (2), by the distance $\sim 10^{-10}$ m between collisions. For electrons in metals the frequency of collisions with atomic vibrations at $T = 300$ K is given by Eq. (3). In these formulas, k_B is Boltzmann's constant, M is the mass of a water molecule, and $\hbar$ is Planck's constant divided by 2π. Compared to atoms, electrons have much higher mean velocities, $v_{el} \sim (2\text{–}10) \times 10^5$ m/s, and v_{el} is independent of temperature. This is a quantum effect caused by the Pauli principle. In a free-electron model, where all diffraction from atoms is ignored, the value is called the Fermi velocity and is given Eq. (4), where n is the number of electrons per unit volume, of the order of 10^{29} m^{-3}, and m is the electron mass. As a consequence of the larger velocity, the mean free path for electrons, estimated as velocity divided by collision rate, is typically 10 to 20 interatomic spacings, 50 times longer than for water molecules. *See* CRYSTAL DEFECTS; EXCLUSION PRINCIPLE; FREE-ELECTRON THEORY OF METALS; LATTICE VIBRATIONS; QUANTUM THEORY OF MATTER.

$$v_{th} = \sqrt{3k_BT/M} \approx 600 \text{ m/s} \tag{2}$$

$$1/\tau \approx (1-10)k_BT/\hbar \sim 10^{14}/\text{s} \tag{3}$$

$$v_F = \hbar(3\pi^2 n)^{1/3}/m \tag{4}$$

Bloch-Boltzmann theory. The long mean free path of electrons in metals is the key to formulating a theory. In a perfect crystal the quantum description of electrons as particle-wave entities is similar to the idea of electromagnetic waves confined in a resonant cavity. The wavelengths have to "fit" the dimensions of the cavity, which causes quantization of wavelength λ. The individual quantum states or orbitals are labeled by their wavevectors $k = 2\pi/\lambda$. The orbital labeled by k has an energy $\varepsilon(k)$, and is pictured as a wave oscillating with a space and time dependence given approximately by $\cos\,[kx - \varepsilon(k)t/\hbar]$, like a traveling electromagnetic wave with frequency $\omega(k) = \varepsilon(k)/\hbar$. A wave packet built from such orbitals moves with velocity given by Eq. (5). There is no simple formula for the function $\varepsilon(k)$. It has to be measured by photoemission or computed using band theory, and varies from metal to metal. The wave packets in real metals do not remain phase-coherent over the whole sample, only over the distance between collisions. Provided this distance is longer than a wavelength, the wavelength and wavevector remain meaningful quantities, which can be used to build a theory, as was done by Felix Bloch in 1928.

$$v(k) = [d\varepsilon(k)/dk]/\hbar \tag{5}$$

Bloch's theory starts from the recognition of Bloch orbitals with wavevectors k. It then adopts from Boltzmann the notion of the distribution $F(k)$ which gives the probability that the state k is occupied. Because of the Pauli principle, this probability cannot exceed 2 (one for spin "up" and one for spin "down"). For states of low energy [$\varepsilon(k)$ far below the Fermi level ε_F], the occupancy is very close to 2, and for states of high energy [$\varepsilon(k)$ far above the Fermi level ε_F], the occupancy is very close to 0. In thermal equilibrium the occupancy $F(k)$ is equal to the Fermi-Dirac function, given by Eq. (6). The conductivity is calculated by summing the velocities of the occupied states, yielding Eq. (7).

$$f(k) = \frac{2}{\exp\left[\dfrac{\varepsilon(k) - \varepsilon_F}{k_BT}\right] + 1} \tag{6}$$

$$\sigma = j/E = \frac{-e}{E \cdot Vol} \sum_k v(k)F(k) \tag{7}$$

When the field E is zero, the distribution $F(k)$ is $f(k)$ in Eq. (6), which has $j = 0$ because states with

positive and negative velocities $v(k)$ are equally populated. The effect of the field E is to cause the k-vectors of the occupied states to change according to Bloch's law, Eq. (8). This is the quantum version

$$\hbar dk/dt = -eE \tag{8}$$

of Newtonian acceleration by the electric force $-eE$, using de Broglie's momentum $\hbar k$ of a quantum wave. The acceleration is continuously resisted by scattering. Therefore the total shift of the typical k-vector is given by Eq. (9), where τ is the time between collisions. Then, using $F(k) \approx f(k - \Delta k)$, we can write the answer for the conductivity as Eq. (10), where $(n/m)_{\text{eff}}$ is given by Eq. (11).

$$\sigma = \left(\frac{n}{m}\right)_{\text{eff}} e^2\tau \tag{10}$$

$$\left(\frac{n}{m}\right)_{\text{eff}} = \frac{-1}{\hbar Vol}\sum_k v(k)\frac{df(k)}{dk} = \frac{1}{Vol}\sum_k\left[\frac{d^2\varepsilon}{\hbar^2 dk^2}\right]f(k) \tag{11}$$

See NONRELATIVISTIC QUANTUM THEORY.

The last part of Eq. (11) is derived by integration by parts after substituting the expression given by Eq. (5) for $v(k)$, and can be interpreted as electron density, given by Eq. (12), times reciprocal effective mass, given by Eq. (13). The Bloch theory recovers a

$$n = \left(\sum f(k)\right)\Big/ Vol \tag{12}$$

$$1/m^* = d^2\varepsilon/d(\hbar k)^2 \tag{13}$$

result proposed on classical grounds by Paul Drude shortly after the discovery of the electron, namely $\sigma = ne^2\tau/m$. However, the interpretation is quite different. A proper quantum theory for the scattering rate τ is needed, and was provided by Bloch by making a quantum generalization of the Boltzmann scattering operator. *See* BOLTZMANN TRANSPORT EQUATION.

An approximate solution of Bloch's theory, known as the Bloch-Grüneisen formula, is often very successful in fitting data, as shown in Fig. 2 for magnesium diboride (MgB_2). In this example, theory deviates from experiment at higher temperatures because high-frequency lattice vibrations, involving boron atoms, are not well represented by the form of the approximate theory. The extrapolated $T = 0$ resistivity value, $\rho_0 = 0.4 \times 10^{-8}\ \Omega \cdot \text{m}$, is called the residual resistivity, and would be zero in a perfect crystal.

Bloch's theory has failures, but they are outnumbered by the successes. One failure is that collective behavior can sometimes be seen at low temperature, the outstanding example being superconductivity. Other examples are more controversial, and include so-called Luttinger liquid behavior, which definitely happens in ideal one-dimensional theoretical models and probably is seen in careful experiments on such systems as carbon nanotubes. Another case is transport by so-called sliding charge-density waves, perhaps seen in quasi-one-dimensional metals like $NbSe_3$. A more common failure is that in metals which are very impure or which have very strong scattering, the mean free path between scattering events may become so short that the wavelength and wavevector of electron quantum states is no longer definable. This undermines the basis of Bloch's theory. *See* CHARGE-DENSITY WAVE.

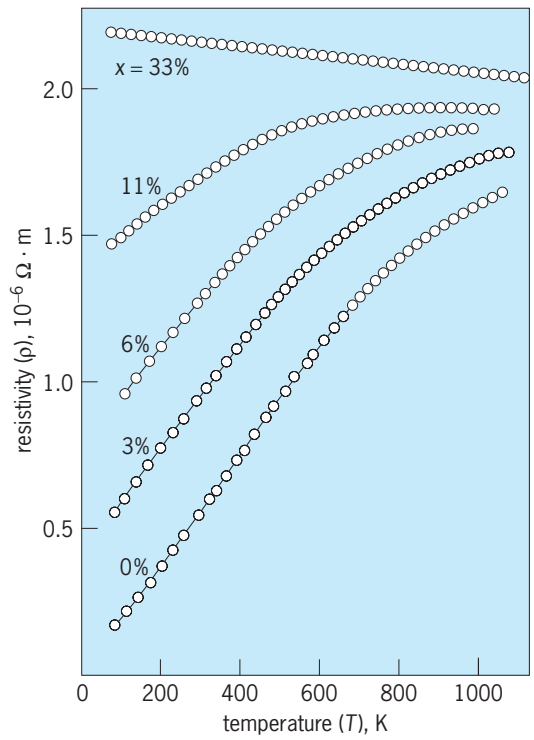

Fig. 4. Resistivity ρ of $Ti_{1-x}Al_x$ alloys as function of temperature (T) for various aluminum concentrations x. (*From J. H. Mooij, Electrical conduction in concentrated disordered transition-metal alloys, Phys. Stat. Sol. (a), 17: 521–530, 1973*)

Such a failure of the theory (technically, a failure of the quasiparticle picture) is illustrated in **Fig. 4**, showing the resistivity of $Ti_{1-x}Al_x$ alloys. At the lower aluminum concentrations, $x = 3\%$ and 6%, the $\rho(T)$ curves shift upward, obeying Mattheissen's rule, given by Eq. (14), where the residual resistivity

$$\rho = \rho_0 + \rho_{\text{pure}}(T) \tag{14}$$

ρ_0 is proportional to impurity concentration x. For the larger concentrations, a too short mean free path causes the quasiparticle picture underlying Bloch-Boltzmann theory to fail. This in turn causes failure of Mattheissen's rule. The quasiparticle picture also fails at high temperature for all concentrations x. *See* MATTHIESSEN'S RULE.

A more abstract formulation of nonequilibrium effects like conductivity exists under the name of Kubo formulas. These provide a route to improved theories.

Philip B. Allen

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Sauders, Philadelphia, 1976; J.

Bass, Electrical resistivity of pure metals and dilute alloys in *Landolt-Börnstein Tables 15a*, Springer, Berlin, 1980; O. Gunnarsson, M. Calandra, and J. E. Han, Colloquium: Saturation of electrical resistivity, *Rev. Mod. Phys.*, 75:1085–1099 2003; C. Kittel, *Introduction to Solid State Physics*, 8th ed., Wiley, New York, 2005; J. M. Ziman, *Electrons and Phonons*, Oxford, 1960, reprint 2001.

Electrical connector

A device that joins electrical conductors mechanically and electrically to other conductors and to the terminals of apparatus and equipment. The term covers a wide range of devices designed, for example, to connect small conductors employed in communication circuits, or at the other extreme, large cables and bus-bars.

Electrical connectors are applied to conductors in a variety of ways. Soldered connectors have a tube or hole of approximately the same diameter as the conductor. The conductor and connector are heated, the conductor inserted, and solder flowed into the joint until it is filled. Solderless connectors are applied by clamping the conductor or conductors in a bolted assembly or by staking or crimping under great mechanical force, usually by means of special tools designed for the purpose.

Industrial and power types. Connectors for electric power systems are commonly cast and assembled with bolts. The material most commonly used is copper alloy, which may vary in composition and properties depending upon the intended use of the connector. Where heavy currents flow through the body of the connector, an alloy of relatively high conductivity is used. Where the connector serves principally to clamp conductors together, an alloy of higher strength and lower conductivity may be used. An important consideration governing the choice of alloy is that it must have a coefficient of thermal expansion very close to that of the conductor material itself (copper or aluminum), so that the connection will remain secure through wide temperature changes.

Connectors designed for severe outdoor service and heavy current and mechanical loading are called power connectors and are widely used in power substations for cable-to-stud connections at equipment, cable-to-bus, bus-to-bus, and bus-to-line connections.

For industrial application, more compact constructions and greater versatility may be required, and connectors with such features are called industrial connectors. In practice, both classes are employed in power systems and industrial plants, the choice depending upon the actual service conditions.

Typical connector types are in-line splice couplers, T-tap connectors, terminal lugs, and stud connectors. Couplers join conductors end to end. T-tap connectors join a through conductor to another conductor at right angles to it (**Fig. 1***a*). Terminal lugs join the conductor to a drilled tongue for bolting to the terminals of equipment (Fig. 1*b*). Stud connectors join the conductor to equipment studs; the stud clamp is threaded or smooth to match the stud. Many variations of these types are made.

Split-bolt connectors are a compact construction widely used for splices and taps in building wiring. The bolt-shape casting has a wide and deep slot lengthwise. The conductors are inserted in the slot and the nut is drawn up, clamping the conductors together inside the bolt (Fig. 1*c*). Newer types of connectors are made to serve this same function.

Expansion connectors or flexible connectors allow some limited motion between the connected conductors. The clamp portions of the connector are joined by short lengths of flexible copper braid and may also be held in alignment by a telescoping guide.

Heavily tinned copper-alloy bolted connectors or aluminum-alloy-body connectors may be used to connect aluminum conductors. They are applied with an oxide-inhibiting compound. Where copper and aluminum conductors are to be clamped together, they are separated by bimetal or tinned washers to prevent galvanic corrosion.

In electronic equipment assembly, involving hookup of wires to terminals and of other components onto printed circuit boards, connections are made by soldering the wire or component to the terminal or printed circuit board. Although soldered terminals are by far the most common type of wiring connections in electronic chassis applications and in manufacture of electronic and electrical equipment, the pressure-wrap connection is a tool-applied method used in the field that speeds up connections where very many have to be made, as in wiring telephone and signal and control equipment. The tool that applies the wire wrap to the terminal stud assures an extremely tight, low-resistance, high-reliability connection by slightly elongating the wire as it is wrapped, putting a tight constrictive force on the terminal.

Separable types. These consist of matched plugs and receptacles, which may be readily separated to disconnect a conductor or group of conductors from the circuit or system. Separable connectors are commonly used for the connection of portable appliances and equipment to an electric wiring system. In electronics and communication they are used for connecting various components so that a single component can be readily removed for servicing without disturbing the circuitry.

For connections to an electric power supply, separable connectors consist of receptacles, usually fixed and permanently connected to the wiring system,

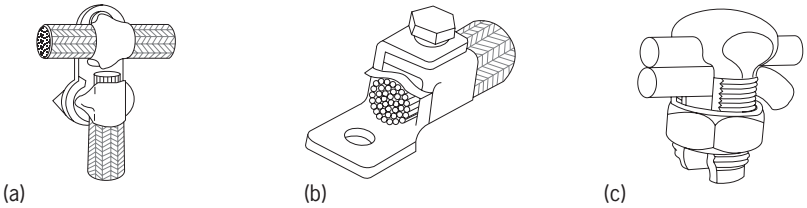

Fig. 1. Types of connectors. (*a*) T-tap connector. (*b*) Terminal lug. (*c*) Split-bolt connector.

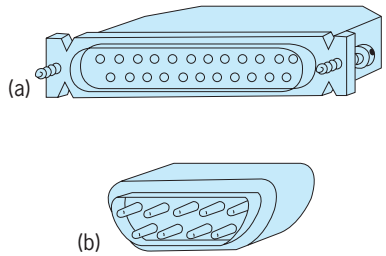

Fig. 2. Typical subminiature computer communications connectors. (*a*) 25-pin RS-232 plug. (*b*) 9-pin RS-422 jack.

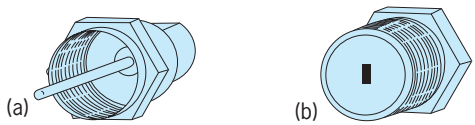

Fig. 3. Typical coaxial cable connection: 5/8 in., 24 threads per inch. (*a*) Jack. (*b*) Plug.

and attachment plugs, usually connected to a cord or cable extending to the appliance or equipment served. Other constructions provide cable-to-cable and cable-to-appliance connections.

Connectors for power systems are typically two-wire, three-wire, or four-wire, depending upon the circuit requirements. They are made in a wide range of types and sizes for applications involving various voltage, current, mechanical, and environmental requirements. They are so constructed that a particular receptacle will accept only a matching plug of identical characteristics, rating, and circuit configuration. Two-wire connectors that can be coupled in only one position are said to be polarized. Multiconductor connectors are designed, by a nonuniform arrangement of contacts or by a key in the assembly housing, so that they may be coupled in only one position.

Locking types. These are designed so that, when coupled, they may not be separated by a strain or pull on the cord or cable. In a typical construction the plug is inserted and twisted through a small arc, locking it securely in place. For hazardous areas, where the spark caused by separating a connector under load may be dangerous, connectors are designed so that the circuit is made or broken within the receptacle only in the fully engaged and locked position. Receptacles may also be interlocked mechanically with a switch, allowing insertion or removal of the plug only when the switch is in the open position.

Multiwire connectors used in electronic, communication, and control circuits contain more connections of much smaller current capacity than power connectors (**Fig. 2**). Connector constructions are designed primarily for mechanical and electrical contact reliability.

Coaxial connectors provide for connecting and disconnecting a coaxial cable from an equipment circuit (**Fig. 3**). They are so constructed as to maintain the conductor configuration, through the separable connection, and the characteristic impedance of the coaxial cable. Coaxial connectors are made of copper or aluminum with contact and current-carrying surfaces plated with silver or gold. They are made in a variety of sizes adaptable to the cables with which they are used. They are designed for cable-to-chassis and cable-to-cable connections in plug, jack, receptacle, feed-through, right-angle, and T-adapter designs. *See* COAXIAL CABLE.

Common receptacles. Plug receptacles, sometimes called convenience outlets, are a type of wiring device distributed throughout buildings for the attachment of cord plug caps from portable lamps and appliances. In residences at least one such outlet must be provided for every 12 linear feet (3.66 m) or major fraction of wall perimeter. The receptacle is conventionally two-wire, 120-V, 15-A, and supplied by a 15- or 20-A branch circuit. The slots which accept the attachment plug are parallel. The device may consist of a single receptacle, a duplex receptacle, or a triplex receptacle (for one, two, or three pairs of blade slots for plug attachment, respectively) in a single outlet box. More attachment plugs can be provided by ganging two or more devices in a common outlet box under a common wall plate. Receptacles may also be combined with other devices (as switches and pilot lights).

In split-circuit duplex receptacles, each half of the device is supplied independently through a separate terminal, with typically one part connected directly to the branch circuit conductors and the other connected through a switch. Grounding receptacles have an additional contact that accepts the third round or U-shaped prong of a grounding attachment plug (**Fig. 4**). Grounding receptacles will accept either grounding or nongrounding attachment plugs,

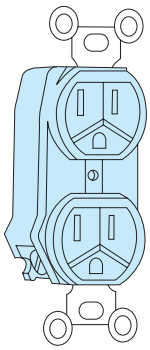

Fig. 4. Grounding-type 15-A receptacle.

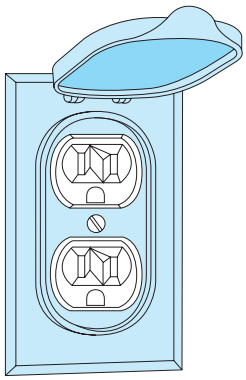

Fig. 5. Weatherproof receptacle with hinged cover.

but grounding attachment plugs can be used only in grounding receptacles. The grounding contact of the receptacle is connected internally to a supporting strap providing a ground both through the outlet box and the grounding conductor, armor, or raceway of the wiring system. *See* GROUNDING.

Multioutlet assemblies are a type of surface raceway designed to hold conductors and plug receptacles, which are usually spaced uniformly along the length of assembly.

Weatherproof receptacles (**Fig. 5**) are equipped with a weatherproof plate, gasketed around the edges, with a hinged or threaded cover over the device. *See* WIRING.

J. F. McPartland; Brian J. McPartland

Bibliography. T. Croft and W. I. Summers (eds.), *American Electricians Handbook*, 13th ed., 1996; IEEE, *National Electrical Safety Code 1997*, 1996; J. F. McPartland, *The Handbook of Practical Electrical Design*, 3d ed., 1999; J. F. McPartland, *National Electrical Code Handbook*, 23d ed., 1999.

Electrical degree

A time interval equal to 1/360 of the time required for one complete cycle of alternating current. Mechanical rotation is often measured in degrees, 360° constituting one complete revolution. In describing alternating voltages and currents, the time for one complete cycle is considered to be equivalent to 360 electrical degrees (360°) or 2π electrical radians. For example, if the frequency f is 60 cycles per second (60 Hz), 360° corresponds to 1/60 second and 1 electrical degree to 1/12,600 second.

There is a definite relationship between electrical and mechanical degrees in rotating electric generators and motors. **Figure 1** shows typical coil and angular relationships in a two-pole alternator. As the magnetic field in the machine moves relative to the coils in the armature winding, the coils are linked sequentially by the fluxes of north and south magnetic poles; two flux reversals induce one cycle of voltage in a given coil. Thus, in a two-pole machine such as in Fig. 1, 360° of electrical cycle corresponds to 360° of mechanical rotation, and an angle measured in mechanical degrees has the same value in electrical degrees. However, in a machine with more than two poles, one electrical cycle is generated per pair of poles per revolution. For example, a six-pole machine generates three cycles of voltage in each armature coil per revolution. In this case, each mechanical degree is equivalent to 3 electrical degrees. The four-pole case is illustrated in **Fig. 2**. In general, the relationship below is valid, where p is the number of magnetic poles of either the rotor or the stator. It follows that the electrical angle between the centers of succeeding poles of opposite polarity is always 180 electrical degrees.

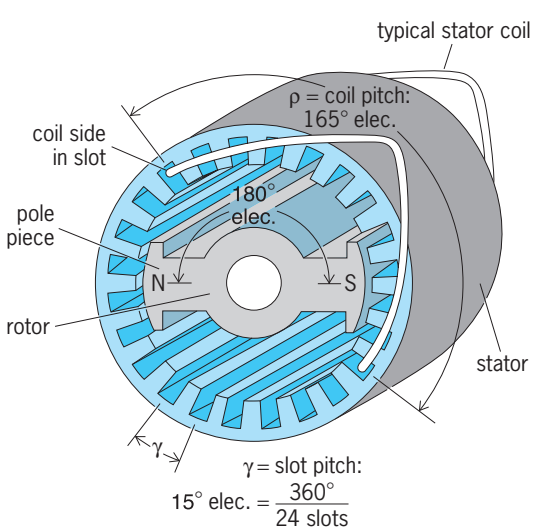

Fig. 1. Coil and angular relationships in a two-pole alternator.

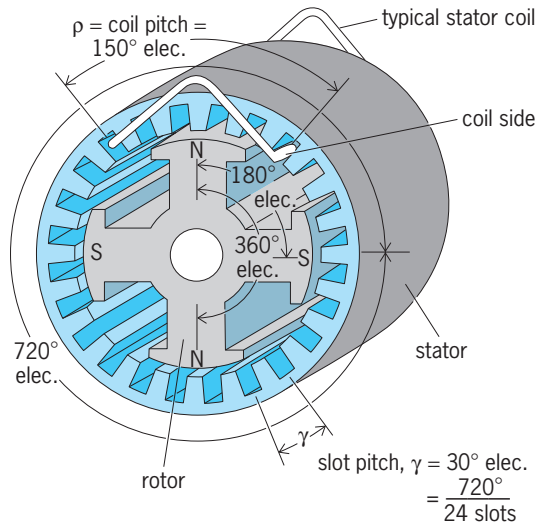

Fig. 2. Coil and angular relationships in a four-pole alternator.

$$\text{Number of electrical degrees in a given angle} = \frac{p}{2} \cdot \left(\begin{array}{l}\text{number of mechanical}\\ \text{degrees in that angle}\end{array}\right)$$

The concept of electrical degrees simplifies the analysis of multipolar machines by allowing them to be analyzed on a two-pole basis. Furthermore, it permits trigonometry to be used in solving alternating-current problems. *See* ALTERNATING CURRENT; ELECTRIC ROTATING MACHINERY; GENERATOR; MOTOR; WINDINGS IN ELECTRIC MACHINERY.

George McPherson, Jr.

Electrical energy measurement

The measurement of the integral, with respect to time, of the power in an electric circuit. The absolute unit of measurement of electrical energy is the joule, or the charge in coulombs times the potential difference in volts. The joule, however, is too small (1 watt-second) for commercial use, and the more commonly used unit is the watt-hour (3.6×10^3 joules). The most common measurement

application is in the utility field. *See* ELECTRICAL UNITS AND STANDARDS.

Electrical energy is one of the most accurately measured commodities sold to the general public. Many methods of measurement, with different degrees of accuracy, are possible. The choice depends on the requirements and complexities of the measurement problems. Basically, measurements of electrical energy may be classified into two categories, direct-current power and alternating-current power. The fundamental concepts of measurement are, however, the same for both.

Methods of measurement. There are two types of methods of measuring electrical energy: electric instruments and timing means, and electricity meters.

Electric instruments and timing means. These make use of conventional procedures for measuring electric power and time. The required accuracy of measurement dictates the type and quality of the measuring devices used (for example, portable instruments, laboratory instruments, potentiometers, stopwatches, chronographs, and electronic timers). Typical methods are listed below. *See* ELECTRIC POWER MEASUREMENT.

1. Measurement of energy on a direct-current circuit by reading the line voltage and load current at regular intervals over a measured period of time. The frequency of reading selected (such as one per second, one per 10 s) depends upon the steadiness of the load being measured, the time duration of the test, and the accuracy of measurement desired. The first dc electric energy meter was an electroplating cell in series with the load. It deposited a mass of metal on an electrode exactly proportional to the total charge transported to the load. The electrode was weighed periodically to determine the energy used. Errors were introduced if line voltage was not constant. It was replaced by more convenient instruments. *See* CURRENT MEASUREMENT; VOLTAGE MEASUREMENT.

In electrical energy measurements, the losses in the instruments must be considered. Unless negligible from a practical standpoint, they should be deducted from the total energy measured. If the voltmeter losses are included in the total energy measured, then watt-hours $= (VI - V^2R)t/3600$, where V is the average line voltage (volts), I is the average line current (amperes), R is the voltmeter resistance (ohms), and t is the time (seconds).

2. Measurement of energy on a direct-current circuit by controlling the voltage and current at constant predetermined values for a predetermined time interval. This method is common for controlling the energy used for a scientific experiment or for determining the accuracy of a watt-hour meter. For best accuracy, potentiometers and electronic timers are desirable.

3. Measurement of energy on an alternating-current circuit by reading the watts input to the load at regular intervals over a measured period of time. This method is similar to the first, except that the power input is measured by a wattmeter.

4. Measurement of energy on an alternating-current circuit by controlling the voltage, current, and watts input to the load at constant predetermined values. This method is similar to the second, except that the power input is measured by a wattmeter. A common application of this method is to determine the standard of measurement of electrical energy, the watt-hour.

5. Measurement of energy by recording the watts input to the load on a linear chart progressed uniformly with time. This method makes use of a conventional power record produced by a recording wattmeter. The area under the load record over a period of time is the energy measurement.

Electricity meters. These are the most common devices for measuring the vast quantities of electrical energy used by industry and the general public. The same fundamentals of measurement apply as for electric power measurement, but in addition the electric meter provides the time-integrating means necessary to electric energy measurement.

A single meter is sometimes used to measure the energy consumed in two or more circuits. However, multistator meters are generally required for this purpose. Totalization is also accomplished with fair accuracy, if the power is from the same voltage source, by paralleling secondaries of instrument current transformers of the same ratio at the meter. Errors can result through unbalanced loading or use of transformers with dissimilar characteristics.

Watt-hour meters are generally connected to measure the losses of their respective current circuits. These losses are extremely small compared to the total energy being measured and are present only under load conditions.

Other errors result from the ratio and phase angle errors in instrument transformers. With modern transformers these errors can generally be neglected for commercial metering. If considered of sufficient importance, they can usually be compensated for in adjusting the calibration of the watt-hour meter. For particularly accurate measurements of energy over short periods of time, portable standard watt-hour meters may be used. Errors may also arise due to integral-cycle control. This is a well-established method of power control in which a silicon-controlled rectifier circuit acts to turn the power on for several cycles and off for a different number of cycles. The main source of error in the induction meter is the difference between the mechanical time-constants for two on-off intervals, making the meter read high.

Watt-hour meters used for the billing of residential, commercial, and industrial loads are highly developed devices. Over the last years many significant improvements have been made, including improvements in bearings, insulating materials, mechanical construction, and sealing techniques which exclude dust and other foreign material. *See* WATT-HOUR METER.

Automatic remote reading. Various aspects of the energy crisis spurred active development of automatic meter-reading systems with the functional

capability of providing meter data for proposed new rate structures, for example, time-of-day pricing, and initiating control of residential loads, such as electric hot-water heaters.

Automatic meter-reading systems under development generally consist of a utility-operated, minicomputer-controlled reading center, which initiates and transmits commands over a communication system to a terminal at each residential meter. The terminal carries out the command, sending the meter reading back to the reading center, or activating the control of a residential load.

Several communication media have been proposed and tested by system developers, including radio, CATV, use of the existing subscriber phone lines, and communication over the electric distribution system itself.

Quantities other than watt-hours. Included in the field of electrical energy measurement are demand, var hours, and volt-ampere hours.

Demand. The American National Standards Institute defines the demand for an installation or system as "the load which is drawn from the source of supply at the receiving terminals, averaged over a suitable and specified interval of time. Demand is expressed in kilowatts, kilovolt-amperes, amperes, kilovars and other suitable units" (ANSI C12.1-1988).

This measurement provides the user with information as to the loading pattern or the maximum loading of equipments rather than the average loading recorded by the watt-hour meter. It is used by the utilities as a rate structure tool.

Var hour. ANSI defines the var hour (reactive volt-ampere hour) as the "unit for expressing the integral, of reactive power in vars over an interval of time expressed in hours" (ANSI C12.1-1988).

This measurement is generally made by using reactors or phase-shifting transformers to supply to conventional meters a voltage equal to, but in quadrature with, the line voltage.

Volt-ampere hour. This is the unit for expressing the integral of apparent power in volt-amperes over an interval of time expressed in hours. Measurement of this unit is more complicated than for active or reactive energy and requires greater compromises in power-factor range, accuracy, or both. Typical methods include: (1) Conventional watt-hour meters with reactors or phase-shifting transformers tapped to provide an in-phase line voltage and current relationship applied to the meter at the mean of the expected range of power-factor variation. (2) A combination of a watt-hour and a var-hour meter mechanically acting on a rotatable sphere to add vectorially watt-hours and var-hours to obtain volt-ampere hours, volt-ampere demand, or both.

Measurement of volt-ampere hours is sometimes preferred over var-hours because it is a more direct measurement and possibly gives a more accurate picture of the average system power factor. This would not necessarily be true, however, where simultaneous active and reactive demand are measured and recorded. *See* ELECTRICAL MEASUREMENTS.

William H. Hartwig

Bibliography. D. A. Bell, *Electronic Instrumentation and Measurement*, 2d ed., 1994; *Code for Electrical Metering*, ANSI C12.11988, 8th ed., 1988; D. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; Instrument Society of America, *ISA Standards and Practices for Instrumentation*, 12th ed., 1995; L. Schnell (ed.), *Technology of Electrical Measurements*, 1993; L. M. Thompson, *Electrical Measurements and Calibration*: *Fundamentals and Applications*, 2d ed., 1994.

Electrical engineering

The branch of engineering that deals with electric and magnetic forces and their effects. *See* ELECTRICITY; ENGINEERING; MAGNETISM.

Scope. Electrical engineers design computers and incorporate them into devices and systems. They design two-way communications systems such as telephones and fiber-optic systems, and one-way communications systems such as radio and television, including satellite systems. They design control systems, such as aircraft collision-avoidance systems, and a variety of systems used in medical electronics. Electrical engineers are involved with generation, control, and delivery of electric power to homes, offices, and industry. Electric power lights, heats, and cools working and living space and operates the many devices used in homes and offices. Electrical engineers analyze and interpret computer-aided tomography data (CAT scans), seismic data from earthquakes and well drilling, and data from space probes, voice synthesizers, and handwriting recognition. They design systems that educate and entertain, such as computers and computer networks, compact-disk players, and multimedia systems. *See* CHARACTER RECOGNITION; COMMUNICATIONS SATELLITE; COMPACT DISK; COMPUTER; COMPUTERIZED TOMOGRAPHY; CONTROL SYSTEMS; DIGITAL COMPUTER; ELECTRIC HEATING; ELECTRIC POWER GENERATION; ELECTRIC POWER SYSTEMS; ELECTRIC POWER TRANSMISSION; ELECTRICAL COMMUNICATIONS; ILLUMINATION; MULTIMEDIA TECHNOLOGY; OPTICAL COMMUNICATIONS; OPTICAL FIBERS; RADIO; TELEPHONE SERVICE; TELEVISION; VOICE RESPONSE.

Early development. Early experimenters developed batteries, electromagnets, and electric motors. In 1832, M. Faraday showed that electric and magnetic effects are not separate, and combined the results of various observations into a single theory of electromagnetic induction. The first major commercial application of electromagnetism was the telegraph, demonstrated in 1844. *See* BATTERY; ELECTROMAGNET; ELECTROMAGNETIC INDUCTION; MOTOR; TELEGRAPHY.

The last quarter of the nineteenth century saw rapid developments in lighting, stemming from the invention of the incandescent lamp, and in electric power generation and delivery. During this period, lighting moved to homes, offices, and schools, and motors powered new forms of transportation and began to transform factories. A major technical

debate in this period was whether to generate and distribute electric energy as direct current or alternating current. After the invention of the transformer in 1883, alternating-current generation and delivery quickly became the dominant system, because of its greater efficiency and versatility. *See* ALTERNATING CURRENT; DIRECT CURRENT; ELECTRIC ROTATING MACHINERY; INCANDESCENT LAMP; TRANSFORMER.

The latter nineteenth and early twentieth centuries saw the development of the telephone and wireless telegraphy. Telegraph wires provided communications for railroads and nearly instantaneous message capability for business, industry, and personal data, while the telephone industry grew rapidly. Wireless telegraphy linked continents with high-powered spark transmitters.

Electronics. The audion, invented in 1906 and later called the triode, was a vacuum tube (or valve) with three electrical elements. The triode was used to design and build amplifiers, a basic element in modern electronics. Amplifiers enable a weak information signal to control a local power source, yielding a more powerful output. Amplification and other applications of tubes enabled the rapid development of the electronics industry. The junction transistor, whose invention in 1948 followed that of the point-contact transistor in the previous year, became the dominant electronic device by the 1960s, replacing vacuum tubes. *See* AMPLIFIER; TRANSISTOR; VACUUM TUBE.

The integrated circuit, invented in 1959, quickly revolutionized electronics. This device or process enables designers to put many transistors and associated components into a single package. The number of electronic devices in a single integrated-circuit chip now routinely exceeds 1 million. These units have made the personal computer a reality, and are installed in home appliances, cars, offices, and factories. *See* ELECTRONICS; EMBEDDED SYSTEMS; INTEGRATED CIRCUITS; MICROCOMPUTER; MICROPROCESSOR.

Feedback. Electronic devices can produce, along with amplification, distorted output signals. An invention of the 1920s, called negative feedback, reduces this problem. A fraction of the amplifier output signal is subtracted from the input signal. While this subtraction reduces gain and signal level, it also reduces distortion. The idea illustrates the trade-offs that must be made throughout the design process, as competing requirements are balanced. While positive feedback can sometimes find application, it is often a serious problem in electronics. Positive feedback is responsible for the screeching sound in a public address system when a person inadvertently points a microphone at a loudspeaker. *See* DISTORTION (ELECTRONIC CIRCUITS); FEEDBACK CIRCUIT.

Control systems. Control systems use feedback. Common examples are thermostats that control furnaces or air conditioners. Automobile cruise controls and automatic pilots for aircraft are applications. Other examples are found in chemical processes and steel-rolling mills. In control systems, electronic circuits enable desired results to be compared with actual results, and the feedback enables the reduction of the error to nearly zero. Electrical engineers work with other engineers in the design and development of control systems. *See* AUTOPILOT; PROCESS CONTROL; THERMOSTAT.

Computers. Electrical engineers have a major role in the continuing development and application of computers, both large and small. They work with computer engineers, computer scientists, and other professionals. Electrical engineers design the data-storage systems, commonly called disks (floppy or hard), the central processing unit (CPU), and the display. The disks are magnetic storage devices, and proper operation of the disks requires a sophisticated control system. The CPU, which is the heart of the computer, includes very large scale integrated (VLSI) circuits for computations, for temporary storage during computation, and for control of printers, disk drives, and displays. The display may be a large vacuum tube with a phosphorescent screen, whose design requires knowledge of electromagnetic theory and materials, or it may be a liquid-crystal display, a technology that has developed rapidly. Computer capability continues to double every 18–24 months. *See* COMPUTER STORAGE TECHNOLOGY; ELECTRONIC DISPLAY.

The computer-communication system known as the internet or worldwide web is a linkage of computers around the world. The internet uses telephone lines, and in turn the telephone system uses satellites, microwave communications, wire communications, and radio. This system allows users to exchange data and ideas rapidly, reliably and at very low cost. *See* DATA COMMUNICATIONS.

Optics. In optical communications, another major contribution of electrical engineering, light pulses and waves carry data over very small strands of glass. The glass replaces copper wire, and the new systems are less expensive, more reliable, and in many respects safer to the user than those they replace. Fiber-optic networks provide two-way video and audio links for universities, colleges, public schools, libraries, hospitals, and court houses.

Challenges. The integration of communications equipment, control systems, computers, and other devices and processes into reliable, easily understood, and practical systems is a major challenge, which has given rise to the discipline of systems engineering. Electrical engineering must respond to numerous demands, including those for more efficient and effective lights and motors; better communications; faster and more reliable transfer of funds, orders, and inventory information in the business world; and the need of medical professionals for access to medical data and advice from all parts of the world. *See* INFORMATION SYSTEMS ENGINEERING; MEDICAL INFORMATION SYSTEMS; SYSTEMS ENGINEERING.

Edwin C. Jones, Jr.

Bibliography. J. E. Brittain (ed.), *Turning Points in American Electrical History*, 1977; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; T. P. Hughes, *Networks of Power*, 1983, reprint 1993; J. D. Irwin and

D. V. Kerns, Jr., *Introduction to Electrical Engineering*, 1995; C. R. Paul, S. A. Nasar, and L. E. Unnewehr, *Introduction to Electrical Engineering*, 2d ed., 1992; J. D. Ryder, *Engineers and Electrons*, 1984.

Electrical impedance

The measure of the opposition that an electrical circuit presents to the passage of a current when a voltage is applied. In quantitative terms, it is the complex ratio of the voltage to the current in an alternating-current (ac) circuit.

The **illustration** shows a generalized ac circuit, which may be composed of the interconnection of various types of circuit elements. If the voltage is $v = V_0 \sin(\omega t + \theta)$ and the current is $i = I_0 \sin(\omega t + \beta)$, then the circuit can be analyzed by considering a complex voltage $V = V_0 e^{j\theta}$ and a complex current $I = I_0 e^{j\beta}$. The impedance of the circuit is then given by Eq. (1), and Z is a complex number given by Eq. (2). R, the real part of the impedance,

$$Z = \frac{V}{I} \tag{1}$$

$$Z = R + jX \tag{2}$$

is the resistance of the circuit, and X, the imaginary part of the impedance, is the reactance of the circuit. The units of impedance are ohms. *See* ELECTRICAL RESISTANCE; REACTANCE.

The modulus of the impedance is given by Eq. (3),

$$|Z| = \sqrt{R^2 + X^2} \tag{3}$$

and this is the ratio of the maximum voltage to the maximum current. The phase angle of the impedance is given by Eq. (4), and this is the angle by

$$\psi = \tan^{-1} \frac{X}{R} \tag{4}$$

which the current waveform lags the voltage waveform. The power factor of the circuit is $\cos \psi$ and is given by Eq. (5). The power dissipated in the circuit is given by Eq. (6).

$$\cos \psi = \frac{R}{\sqrt{R^2 + X^2}} \tag{5}$$

$$P = {}^1/_2 \, |V| \, |I| \cos \psi \tag{6}$$

For a circuit composed of resistors, inductors, and capacitors, R is always positive, but X may be positive or negative. For example, in a circuit consisting of a resistor R and inductor L in series, the impedance at frequency ω is given by Eq. (7), so R is greater than

$$Z = R + jL\omega \tag{7}$$

0 and X is greater than 0. For a circuit consisting of a resistor R and capacitor C in series, the impedance at frequency ω is given by Eq. (8) so R is greater than

$$Z = R + \frac{1}{jC\omega} = R - j\frac{1}{C\omega} \tag{8}$$

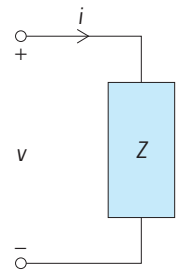

Generalized alternating-current circuit.

0 and X is less than 0. *See* ALTERNATING-CURRENT CIRCUIT THEORY.

J. O. Scanlan

Electrical instability

A condition that is generic to all physical systems that use the amplification of signals, power, or energy in various forms. Electrical circuits are perhaps the best-known example because the use of amplification is very common there. In order for instability to occur, a closed (circular) path must exist, and the net amplification around it must be sufficiently large that any random signal (for example, noise) that travels around the closed path grows uncontrollably. The condition is eventually limited by some physical aspect of the system such as nonlinearity, saturation, or power available. The energy to create the uncontrollable signal growth is obtained by conversion from some other form of energy. For example, a high-frequency electrical oscillation can draw its energy from a direct- current (dc) power supply. *See* ENERGY; LINEARITY; SATURATION.

There can be various degrees of stability and instability in an electrical circuit or system, and instability may be desirable or undesirable depending on the context. Desirable instability is displayed by a cross-coupled positive-feedback latch in a computer memory cell, or an oscillator that is used to create the high-frequency carrier wave of a radio transmitter. An undesirable instability might arise in the loss of control over a power transmission system due to surges on the generators, leading to a necessary decoupling of the system. Another example is the high-pitched whine heard when a radio-station telephone caller leaves the home radio receiver on and thus allows feedback from the radio receiver to the telephone, that is, a closed-loop signal path. *See* FEEDBACK CIRCUIT; OSCILLATOR.

Yet another example is a computer simulation program that fails because of lack of convergence when trying to solve for the signal levels in a circuit. This could be due to a numerical instability in the computer algorithm. *See* ALGORITHM; SIMULATION.

Electrical instability can occur at zero frequency (dc), for example, in the cross-coupled memory latch, but its accidental occurrences are generally at high frequencies. A configuration may be designed to be stable at low frequencies, but inadvertent phase shifts or time delays at higher frequencies may

create a positive-feedback situation. For this reason, the design of amplifiers, feedback control systems, and so forth necessarily involves much attention to the higher-frequency aspects even though the desired use may be primarily at lower frequencies. *See* AMPLIFIER; CONTROL SYSTEMS.

The intellectual basis for the design of electrical circuits and systems to ensure their stability is quite well developed, based on linear mathematics. Computer tools are extensively used in complex application areas. However, the situation is less satisfactory in the design of deliberately unstable systems, such as oscillators, because their behavior is finally explained by less tangible nonlinear limiting mechanisms. *See* CONTROL SYSTEM STABILITY; LINEAR SYSTEM ANALYSIS.

Miles A. Copeland

Electrical insulation

A nonconducting material that provides electric isolation of two parts at different voltages. To accomplish this, an insulator must meet two primary requirements: it must have an electrical resistivity and a dielectric strength sufficiently high for the given application. The secondary requirements relate to thermal and mechanical properties. Occasionally, tertiary requirements relating to dielectric loss and dielectric constant must also be observed. A complementary requirement is that the required properties not deteriorate in a given environment and desired lifetime. *See* CONDUCTOR (ELECTRICITY).

Electric insulation is generally a vital factor in both the technical and economic feasibility of complex power and electronic systems. The generation and transmission of electric power depend critically upon the performance of electric insulation, and insulation now plays an even more crucial role because of the energy shortage.

Insulation Requirements

The important requirements for good insulation will now be discussed.

Dielectric strength. The basic difference between a conductor and a dielectric is the availability of charge carriers with high mobility in a conductor, whereas free charge has little mobility in a dielectric. Dielectric strength is a measure of the electric field required to cause sudden and substantial motion of charge across the surface or through the bulk of a dielectric. Dielectric strength usually deteriorates with increased moisture content and with elevated temperature. For high-voltage (on the order of kilovolts) applications, dielectric strength is the most important single property of the insulation.

Steady-state strength. The voltage at which a sudden high-current discharge punctures an insulator is called the breakdown voltage. In a uniform field this voltage divided by the electrode gap is called the dielectric strength, the breakdown field strength, the electric strength, or the breakdown stress. In a nonuniform field it is important to differentiate between the maximum field and the average field at breakdown. *See* BREAKDOWN POTENTIAL; ELECTRICAL BREAKDOWN.

An insulator with an operational dielectric strength of 4–12 kV/mm (100–300 V/mil) is considered good. Many of the insulators listed with a dielectric strength more than an order of magnitude higher than this are ordinarily stressed at these lower values in practical operation. The reason is that imperfections in the form of voids, cracks, filamentary defects, and so forth occur whenever long lengths and thick sections of insulating material are manufactured.

Whenever two dielectrics are in series in an alternating electric field, the electric field E is higher in the medium of lower dielectric constant κ by the ratio of dielectric constants, $E_2/E_1 = \kappa_1/\kappa_2$. Therefore the electric field is higher in a cavity, which generally has a dielectric constant of 1, by a factor of 2–3, than it is in the insulator. Furthermore, the dielectric strength is much less in the defect than in the insulator (**Table 1**; **Fig. 1**). Thus, in addition to operating insulators at much lower than intrinsic dielectric strength, schemes such as oil impregnation are used

TABLE 1. Properties of various good insulators

Insulator	Resistivity, 10^{14} ohm-cm	Dielectric strength, 10^3 V/mil*	Power factor $\times 10^{-4}$	Dielectric constant	Tensile strength, 10^3 lb/in.^{2}†	Chemical resistance	Flammability
Lexan	10–1000	8–16	10–30	3	8	Good	Self-extinguishing
Kapton H	1000–10,000	3–8	20–50	3	22	High	Self-extinguishing
Kel-F	10–1000	2–6	20–40	2–3	5	High	No
Mylar	0.1–1	4–16	30–120	3	20	Good	Yes
Parylene	10–10,000	6–10	2	2–3	10	Very good	Yes
Polyethylene N	100–500	1–17	3–30	2	4–6	Good	Yes
Polytetrafluoroethylene	0.1–1000	1–7	1–4	2	4–5	High	Very low
Air (1 atm, 1 mm gap)	—	0.1	10^{-3}	1	0	Stable	No
Sulfur hexafluoride (1 atm, 1 mm gap)	—	0.2–0.3	10^{-3}	1	0	Stable	No
Vacuum ($<10^{-5}$ torr or 1.3×10^{-3} Pa, 1 mm gap)	—	2–3	0	1	0	Stable	No

*10^3 V/mil = 4×10^5 V/cm.
†10^3 lb/in.^{2} = 6.9 (MPa).

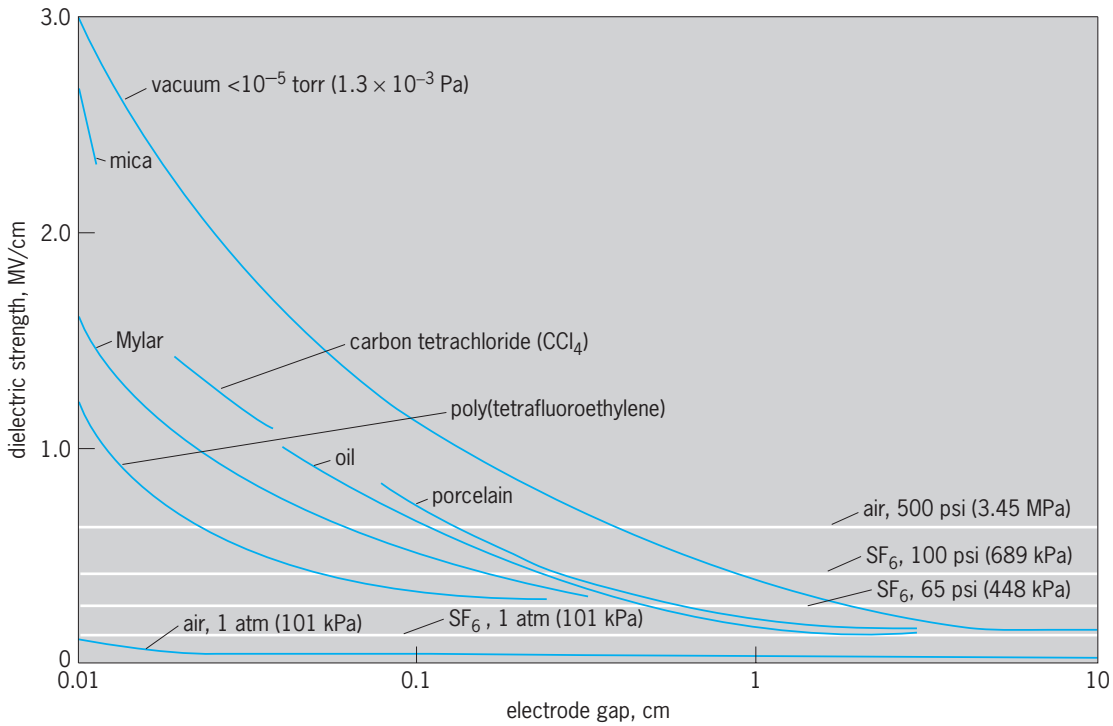

Fig. 1. Uniform field dielectric strength for solid, liquid, gas, and vacuum insulation.

when possible to fill the defects to increase both their dielectric constant and their dielectric strength.

Although solid extruded dielectric cable is now employed up to at least 420 kV system voltage and operates at average fields in the range of 15 kV/mm, oil-impregnated-paper or paper-polypropylene laminate (PPLP) wound-tape insulation can be used at even higher voltages. The probability that defects in the tape will line up in proximity is much less than in a solid piece of the same insulation material. The tape thickness also limits the size of the defect in the direction of the electric field. *See* DIELECTRIC MATERIALS.

Laboratory testing of insulation over a wide range of conditions is conducted to determine the effects of aging, high temperature, heat cycling, ultraviolet radiation, moisture, and so forth. A statistical analysis of the data based upon the weak-link Weibull statistics is then used to determine safe operating voltages and electric fields for the required lifetimes and environments. Operational withstand stress is sometimes arbitrarily defined as being at least two standard deviations below the average dielectric strength in the laboratory.

Impulse strength. Impulse strength is a measure of an insulator's ability to withstand a rapid rise and rapid fall (measured in microseconds) of electric stress. In addition to steady-state dielectric strength, an insulator must be able to meet a basic impulse level (BIL) and a basic switching level (BSL) surge (not as high a rate of voltage increase as for BIL) requirement. The impulse breakdown electric field strength is greatest for solids, followed by liquids, gases, and vacuum. For example, a development in tape insulation consisting of a combination of cellulose paper bound to polypropylene film or fibers [known as polypropylene laminated paper (PPLP)] and impregnated with oil has an impulse strength of 2.7 MV/cm, whereas the impulse strength of sulfur hexafluoride from 100 to 400 kilopascals (15 to 60 lb/in.^{2}) goes only from 0.2 to 0.45 MV/cm. Thus, in designing transmission cables, sufur hexafluoride cables are limited by their BIL, whereas most solid insulation cables are limited only by their steady-state dielectric strength. This is why solid insulation cables are so much more compact than sulfur hexafluoride cables for the same ratings, in addition to the fact that they have higher steady-state dielectric strength. However, when a solid dielectric breaks down electrically, the punctured and surrounding regions must be replaced. After sulfur hexafluoride breaks down, the gas is self-healing and recovers dielectric strength. If no damage has been incurred by the conductors, operation may resume.

The impulse strength for sulfur hexafluoride increases with increasing pressure but tends to level off above 400 kPa (60 lb/in.^{2}). The impulse strength of liquids increases less rapidly with increasing pressure, leveling off above 2 megapascals (300 lb/in.^{2}). For example, carbon tetrachloride's impulse strength goes from 2.2 to 2.6 MV/cm in going from 100 kPa to 2 MPa (15 to 300 lb/in.^{2}). This pressure dependence for liquids tends to disappear for rates of field increase greater than 4 MV/cm-μs. This is experimental evidence that a "bubble" mechanism (see below) cannot be operative in such short times.

Resistivity. Resistivity is a measure of how much current will be drained away from the conductor through the bulk (bulk resistivity) or along the surface (surface resistivity) of the dielectric.

An insulator with resistivity equal to or greater than 10^{13} ohm-cm may be considered good. While resistivity measurements of conductors are straightforward, measurements on insulators are much more difficult and are subject to a much wider range of values for a given material. A dc measurement of current is made after application of a uniform electric field for a specified period of time. A time variation from 1 s to 1 min can lead to orders-of-magnitude variation in the measured resistivity, and the "true" resistivity can be nearly impossible to determine.

Since a resistivity measurement is nondestructive, it can be used as an on-line quality and uniformity check during manufacturing processes. During actual operation in the field, a sharp decrease in resistivity usually signals an ensuing service failure. *See* ELECTRICAL RESISTIVITY; INSULATION RESISTANCE TESTING; NONDESTRUCTIVE EVALUATION.

Dielectric loss. When a dielectric is subjected to an alternating field, a time-varying polarization of the atoms and molecules in the dielectric is produced. The response time of polarization phenomena varies from nearly instantaneous for electronic polarization to seconds or even hours for some forms of orientational polarization. If the time constant for polarization results in a lag between the applied field and polarization, a loss results, of which the power factor is a measure. The importance of the power factor varies with application and tends to become more important at high temperatures, where the loss tends to increase and can cause thermal runaway, and in cryogenic systems, where the the energy required to remove the heat that results from dielectric loss can be great.

Dielectric loss is a function of both temperature and frequency and generally decreases as these two variables decrease. However, it is not a monotonic function of these parameters, and at various discrete values of temperature and frequency quasiresonances occur where the loss can be large. A sharp rise in dielectric loss, usually accompanied by a rapid temperature increase, is an indication of impending breakdown.

Neglecting these high-loss sites, **Table 2** indicates the power factor in the three temperature regions of general interest, 300 K or 80°F (conventional), 77 K or −321°F (cryoresistive), and 4 K or −452°F (superconducting). Values of less than 10^{-3} are good at 300 K, values less than 10^{-4} are good at 77 K, and values less than 10^{-5} are good at 4 K. With the increasing need to deliver large blocks of power (greater than 3000 MW) for long distances (greater than 80 km or 48 mi) from remote energy sources, advanced power transmission must have insulation capable of meeting cryogenic requirements. A power factor of less than 10^{-5} at the operating cryogenic temperature is generally desirable, since the overall power loss is amplified by the refrigeration inefficiency, which may require as much as 500 W per watt of dissipation. *See* SUPERCONDUCTING DEVICES.

TABLE 2. Power factor of insulators at three temperature

Insulator	Power factor × 10^{-4} 300 K (80°F)	77 K (−321°F)	4 K (−452°F)
Kapton H	20–50	5	0.5
Polyethylene (low density)	3–30	0.1	0.02
Polytetrafluoroethylene	1–4	0.9	0.02
Kraft paper	50	20	7
Vacuum (<10^{-5} torr or 1.3×10^{-3} Pa)	0	0	0

Ideally the power factor should be independent of electric field stress. However, owing to processes unrelated to polarization, such as electron emission and ionization, the dielectric loss tends to increase with increasing electric field, and this is reflected in an increasing power factor.

The power factor is the cosine of the phase angle ϕ between the voltage and current, or equivalently the ratio of resistance to impedance for the medium. Occasionally confusion arises regarding terminology when the dielectric loss is related to the complementary angle δ, which satisfies Eq. (1) for small δ. Thus

$$\text{Power factor} = \cos\phi = \sin\delta \doteq \tan\delta \doteq \delta \qquad (1)$$

the power factor is often used interchangeably with the terms loss tangent and loss angle.

Dielectric constant. The dielectric constant, also known as the relative permittivity or specific inductive capacity (SIC), is a measure of the ability of the dielectric to become polarized, taken as the ratio of the charge required to bring the system to the same voltage level relative to the charge required if the dielectric were vacuum. It is thus a pure number, but is in fact not a constant, and may vary with temperature, frequency, and electric-field intensity. For highly polar materials such as water, the variation with temperature and frequency is dramatic, going from 80 at 300 K (80°F) to 3 at 77 K (−321°F), or from 80 at 60 Hz to 20 at 10 kHz.

In addition to the problem of intensification of the electric field in regions of relatively lower dielectric constant, a low-dielectric-constant insulation is desirable for two more reasons. In ac transmission cables, the lower the dielectric constant, the more the current and the voltage will be in phase. Thus, more usable power will be delivered, without the need for reactive compensation. Furthermore, in reducing the charging current (which is proportional to the dielectric constant), concomitant power losses (related to the square of the charging current) are also reduced. A high dielectric constant is desirable in capacitors, since the capacitance is proportional to it. *See* CAPACITANCE; DIELECTRIC MEASUREMENTS; PERMITTIVITY.

Dielectric absorption. When high voltage is applied to a solid dielectric, charge tends to be absorbed into the dielectric. When the voltage is removed, some of this change comes back out of the dielectric over time. This property is known as dielectric absorption. For example, if high voltage is applied to a polymer film capacitor for several minutes, the

capacitor is shorted for a few seconds, and then the capacitor is left open-circuit, a voltage will appear across the capacitor terminals as a result of the absorption charge which evolves from the capacitor film. A "perfect" capacitor, such as a vacuum capacitor, would have no dielectric absorption. Many capacitors used in electrical or electronic applications, including most high-voltage capacitors, are based on polymer films and do exhibit dielectric absorption. Poly(tetraflyoroethylene) [PTFE] and polystyrene capacitor films exhibit among the lowest dielectric absorption.

Insulator Properties

Table 1 summarizes the properties of various good insulators at or near 60 Hz and ambient temperature (about 300 K or 80°F). The range of values indicates not only differences between measurements of different investigators, but differences resulting from small variations in a parameter. For example, the dielectric strength depends on the thickness and area of the sample being tested. It generally decreases with increasing thickness and area. The voltage at which electrical breakdown occurs increases less than linearly with thickness. Thus the corresponding value of electric field decreases.

All insulators may be classified as either solid or fluid. Solid insulation is further divided into flexible and rigid types.

Flexible insulation. Flexible hydrocarbon insulation is generally either thermoplastic or thermosetting. Thermosets are initially soft, and can be extruded by using only pressure. Following heat treatment, when they return to ambient temperature, they are tougher and harder. After thermosetting, nonrubber thermosets are harder, stronger, and have more dimensional stability than the thermoplastics. Thermoplastics are softened by heating, and when cool become hard again. They are heat-extruded.

Thermoplastics. The first seven insulators listed in Table 1 are thermoplastics, as are nylon, polystyrene, and poly(vinyl chloride). Those listed in Table 1 are solid polymers that are used at ambient temperature and have good potential for cryogenic applications, and their monomeric structures are shown in **Fig. 2**. Polymers come in amorphous and ordered states which have different dielectric properties. The ordered state is frequently called "crystalline," but is not the same as the crystalline state of solids. *See* POLYMER.

Lexan is a trade name for a group of polyesters formed from carbonic acid, and generically called polycarbonate (PC). Polycarbonate has good electrical and mechanical properties, good dimensional stability, good resistance to creep, high distortion temperature, and ease of molding and extrusion. *See* POLYESTER RESINS.

Kapton H is a trade name for polypyromellitimide (PPMI). PPMI has exceptional resistance to thermal degradation, because it is one of the polyimides that incorporates multiple bonds along the backbone of the chain. It has good mechanical strength and good

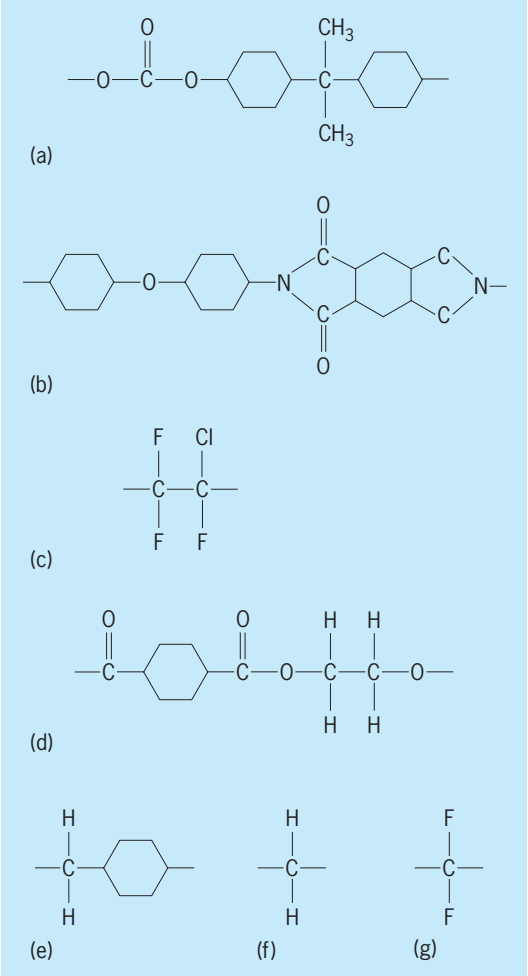

Fig. 2. Monomeric structures of polymers used as electrical insulators. (*a*) Lexan. (*b*) Kapton H. (*c*) Kel-F. (*d*) Mylar. (*e*) Parylene. (*f*) Polyethylene. (*g*) Polytetrafluoroethylene.

dimensional stability at both high and low temperatures, retaining ductility down to cryogenic temperatures. *See* HETEROCYCLIC POLYMER.

Kel-F is a trade name for poly(chlorotrifluoroethylene) [CTFE]. It is one of the best insulators in its ability to exclude moisture, in high-temperature strength, in chemical inertness, and in low coefficient of friction. *See* POLYFLUOROOLEFIN RESINS.

Mylar is a trade name for poly(ethylene terephthalate) [PET]. PET is a high-molecular-weight polyester with a stiff polymer chain and resilient interchain bonds that retain good mechanical properties beyond 150°C (300°F). Because of their toughness, PET films are used not only for electrical insulation, but also for magnetic tapes. *See* MAGNETIC RECORDING.

Parylene is a trade name for poly(*p*-xylene) [PPX]. PPX is linear and highly regular, and was developed to be highly radiation-resistant. It has good dimensional stability down to cryogenic temperatures. It has not had widespread use owing to its cost. *See* POLY-P-XYLYLENE RESINS.

Polyethylene (PE) is a widely used inexpensive electrical insulator with good mechanical and

excellent dielectric properties. Polyethylene can be cross-linked chemically or by irradiation to increase its mechanical strength and improve its high-temperature properties. Chemical cross linking leaves impurities which increase loss. Electron-beam cross linking results in lower loss. Cross-linked polyethylene (XLPE) power cables are manufactured for up to 420 kV system voltage using chemical cross linking, while XLPE high-frequency signal coaxial cables are made using electron-beam cross linking to reduce high-frequency loss. Although polyethylene has a higher dielectric strength than chemically cross-linked XLPE, its low melting point of 118°C (244°F) makes it less desirable for extruded dielectric cables than XLPE. Even after cross linking, melting of crystallites in XLPE results in a precipitous drop in yield stress above 90°C (194°F) and very large thermal expansion in the same temperature range. These characteristics make implementation of splices and terminations problematic for XLPE cables rated to 105°C (221°F). Ethylene propylene rubber (EPR) with inorganic filler is also be used in extruded solid dielectric power cables. Since the EPR polymer is nearly amorphous and the compound gains its physical properties from the inorganic filler, the mechanical and thermal properties of EPR cable dielectric are much more stable with temperature, both low and high, than XLPE, to the point that EPR cable can be rated to operate at 140°C (284°F). *See* POLYOLEFIN RESINS.

Polytetrafluoroethylene and its copolymer with hexafluoropropylene are highly ordered and orientable polymers made of chains of CF_2 with its strong C-F bond. This is why their structure involves little or no branching or cross linking. Thus this polymer has a low coefficient of friction, is stiff, inert, and insoluble, and has a high melting point of 327°C (621°F). Because of its high symmetry and tight bonding, it has one of the lowest known power factors and dielectric absorption for a solid, as well as other excellent dielectric properties. It may be processed by powder and sinter techniques as well as by extrusion or injection molding.

Polystyrene is a good insulator that is used in thin films or where it will not be stressed or bent, because it is stiff and brittle when it is thick. Fillers in the form of silica, alumina, mica, and so forth reduce the brittleness (probably by inhibiting crack propagation), improve the thermal conductivity, and reduce thermal expansion. Hydrated fillers, such as aluminum trihydrate, inhibit tracking, since the waters of hydration tend to be arc-quenching. Polystyrene capacitor film has among the lowest loss ($\sim 10^{-5}$) and dielectric absorption of all capacitor films. *See* POLYSTYRENE RESIN.

Thermosets. The various rubber compounds, neoprene, and epoxies are thermosets. Natural rubber, butyl rubber, buna S rubber, and silicone rubber are good insulators that are tough and flexible. Although chemically and mechanically tough, neoprene does not have good dielectric properties. The epoxies are not very flexible but are used at joints of flexible insulation because of their good bonding, electrical, and mechanical properties. *See* POLYETHER RESINS; RUBBER.

Cellulose paper. Cellulose paper insulation is neither thermoplastic nor thermosetting. It is widely used in cables and rotating machinery in multilayers and impregnated with oil. It has a relatively high dielectric loss that hardly decreases with decreasing temperature (Table 2), which rules it out for cryogenic applications. Oil-impregnated cellulose paper has very high dielectric strength perpendicular to the paper (through the layers of oil-impregnated paper) but relatively low dielectric strength parallel to the paper surface. Because of its high dielectric strength, the high loss has not been a deterrent to its use in conventional ambient-temperature applications. However, the high dielectric strength deteriorates quickly if moisture permeates the paper. *See* CELLULOSE; PAPER.

Rigid insulation. This second major group of solid insulators includes glass, mica, epoxies, ceramoplastics, porcelain, alumina, other ceramics, and fiber-reinforced plastic (FRP). Rather than being used to insulate wires and cables except for mica, these materials are used in equipment terminations (potheads) and as support insulators (in tension or compression) for overhead lines whose primary dielectric is air. These rigid structures must be shock-resistant, be relatively water-impervious, and be able to endure corona discharges over their surfaces. *See* CORONA DISCHARGE.

For use as termination insulators, such as for transformers, underground transmission lines, reactors, and generators, both high bulk dielectric strength and high surface dielectric strength are important. For use in connection with equipment such as overhead transmission lines, the requirement for bulk dielectric strength is less severe, and the emphasis is on the surface. Surface dielectric strength of several kilovolts per inch is adequate, and the creepage paths are usually designed for approximately 400 V/cm (1 kV/in.).

In addition to normal electromagnetic operating forces and windage stress, these insulators must be able to withstand switching surges and shocks from short-curcuit currents. These fault currents may be 10 times higher than the normal current, resulting in forces that are 100 times higher, since the force is proportional to the square of the current.

Porcelain is widely used for such applications. It is a hard, brittle material made by firing feldspar, quartz, clay, and other minerals at high temperature. Glass is a completely amorphous material containing 50–90% silica. Glass has a dielectric strength of about 2 MV/cm at 300 K (80°F). Mica is a mineral that occurs naturally in a laminated form. For thin sections of 0.002 cm thickness, mica has a dielectric strength as high as 10 MV/cm perpendicular to the direction of lamination. Fiber-reinforced plastics, often composites of glass fiber and epoxy, covered with silicone rubber are replacing glass and porcelain in many applications, including transmission-line insulators and high-voltage terminations. This structure has the advantages of much lower weight, improved

performance under polluted conditions as a result of the silicone rubber exterior insulation, and being explosion-proof when operated under pressure. *See* GLASS; MICA; PORCELAIN.

Fluid insulation. Liquids, gases, and vacuum fall in the category of fluid insulation. For all of these, the electrical structure must be such as to contain the fluid in the regions of high electric stress.

Liquids. The main types of insulating liquids are the mineral oils, silicones, chlorinated hydrocarbons, and the fluorocarbons with dielectric strengths on the order of megavolts per centimeter. Many other liquids also have good dielectric strength, such as carbon tetrachloride, toluene, hexane, benzene, chlorobenzene, alcohol, and even deionized water. However, special problems with liquids, such as chemical activity, high thermal expansion, thermal instability, low boiling point, and tracking (conducting carbon residue) after arcing, rule out many liquids. Even transformer oil in combination with the insulating paper in activated transformers produces water which must be filtered out.

Gases. Most gases have a dielectric constant of about 1, and low dielectric loss. **Table 3** compares the relative dielectric strengths of various gases. Those containing fluorine or chlorine are strongly electronegative. Of course, the temperature at which the gas condenses out of the gaseous state is an important consideration, as well as chemical stability, tracking, reactivity, and cost.

Air is used as a dielectric in a wide variety of applications, ranging from electronics to high-voltage (765-kV) and high-power (2000-MW) electric transmission lines. Dry air is a reasonably good insulator (Table 1). However, its dielectric strength decreases with increasing gap (Fig. 1).

Sulfur hexafluoride (SF_6; Table 1) has, at 1 atm (101 kPa), about three times the dielectric strength of air, carbon dioxide, or nitrogen, with this difference increasing with pressure and gap in uniform electric fields (Fig. 1). Sulfur hexafluoride is an electronegative gas which impedes electrical breakdown by capturing free electrons and forming negative ions. Not only does sulfur hexafluoride possess a high dielectric strength, but also, as a result of very rapid recombination deionization after being decomposed in an arc, it recovers its dielectric strength after arcing much more rapidly than air. This property makes it 100 times as effective as air and other gases for use in power circuit breakers. It also is used frequently in electrical utility substations as an insulating medium.

TABLE 3. Comparison of various gaseous dielectrics

Gas	Condensation temperature (at 1 atm or 101 kPa), °C (°F)	Dielectric strength relative to nitrogen
N_2	−195.8 (−320.4)	1.0
CO_2	−78 (−108)	0.9
SF_6	−63.8 (−82.8)	2.5–3
CCl_4	76 (169)	6.3
$CFCl_3$	23.8 (74.8)	3–4.5
CF_2Cl_2	−29.8 (−21.6)	2.4–2.5
CF_2ClCF_2Cl	3.6 (38.5)	2.8
CF_3CN	63 (−81)	4
C_2F_5CN	−30 (−22)	5
C_3F_7CN	1 (34)	6
C_3F_8	−36.7 (−34.1)	2–2.9
C_4F_6	−5 (23)	4
C_4F_{10}	−2 (28)	2.5
C_5F_8	25 (77)	6
C_5F_{10}	22 (72)	>4
He	−269 (−452)	<0.2

Vacuum. Vacuum (that is, pressures of less than 10^{-3} Pa or 10^{-5} torr) has one of the highest dielectric strengths in gaps ranging from 0.1 to 1 mm (0.004 to 0.04 in.; Fig. 1; Table 1). However, as the gap increases, its dielectric strength decreases rapidly. A perfect vacuum should be a perfect insulator, since there would be no charge carriers present to contribute to electrical conductance. The limitation of vacuum as a dielectric results from the effects of a high electric field at the surface of electrodes in vacuum, which can cause field emission of electrons, rather than because a perfect vacuum is far from being realized in the laboratory.

The dielectric properties of vacuum can degenerate rapidly because vacuum offers no resistance to the motion of charge carriers once they are introduced into the vacuum region. The collision mean free path is of the order of meters at pressures less than 10^{-2} Pa (10^{-4} torr), and this is below the pressure range for a gas discharge. However, avalanche processes occur at the electrodes, and vapor from the electrodes becomes the arcing medium.

Vacuum has found increasing applications as an insulator in electron tubes, photocells, high- frequency capacitors, electron microscopes, particle accelerators, distribution voltage (∼15 kV) circuit breakers, and so forth. It is used as an arcing medium with exceptional rate of recovery of dielectric strength. The realization of its full potential is strongly linked to the constitution of the electrodes.

Nature of Electrical Breakdown

Electrical breakdown in gases first began to be understood in 1889, and breakdown in solids in 1935. Breakdown in liquids is less well understood. Breakdown in vacuum (in terms of an encompassing predictive theory) is least well understood, despite experimental investigations going back to 1897. *See* ELECTRICAL BREAKDOWN.

Gases. The Townsend avalanche criterion satisfactorily accounts for the threshold for electrical breakdown in nonelectronegative gases. This theory may be modified to include the process of electron capture by electronegative gases. When the space-charge field becomes large, other mechanisms enter in, such as streamers. At very high pressures, the electric field strength for breakdown is so high that field emission from the electrodes enters into the process.

The Townsend criterion for avalanche is given by Eq. (2). Here α, the first Townsend coefficient, is

$$\gamma(e^{\alpha\delta} - 1) = 1 \qquad (2)$$

the number of ionizing collisions per unit length in

the direction of the field made by an electron; and γ, the second Townsend coefficient, is the number of secondary electrons produced at the cathode per electron produced by impact ionization in the gap, δ. Thus the dielectric strength of a gas is a function of both the gas species and the cathode material.

Although impact ionization in the gas produces equal numbers of positive ions and electrons, a space charge would develop even with equal displacement of the negative and positive charges. Since the electrons have greater mobility, this further separates the charges. In those situations when enough positive-ion space charge is produced at the original avalanche head to cause the space-charge field to be almost as large as the applied field, streamer breakdown occurs.

In highly nonuniform fields, where the electric field is very high at one electrode and low at the other, localized breakdown (called corona) which does not bridge the gap can occur. Corona can reduce the electric field around the pointed electrode, in effect rounding it off. This results in a higher breakdown voltage than otherwise, a phenomenon known as corona stabilization.

Most systems in the field become contaminated with conducting particles which greatly reduce the gaseous dielectric strength. This problem is important enough to warrant incorporating particle traps in such equipment, particularly those using electronegative gases. *See* ELECTRICAL CONDUCTION IN GASES.

Solids. For ambient temperature and pressure, the densities of solids are 10^3–10^4 times those of gases. This results in dielectric strengths 10–10^2 times higher than gases. At these higher fields, quantum-mechanical processes such as electron tunneling enter in. Additionally, through excitation of lattice vibration a new mechanism for reducing the energy of the accelerated electrons is provided. *See* LATTICE VIBRATIONS; PHONON; QUANTUM MECHANICS.

For the purpose of considering electrical breakdown, solids may be put in the categories of polar crystals, nonpolar crystals, quasicrystalline (highly ordered), and amorphous. On the basis of only density, the dielectric strength of solids and liquids may be expected to be comparable, and this is so. The alkali halides such as potassium bromide, KBr, and sodium chloride, NaCl, are examples of polar crystals. Mica is an example of a nonpolar crystal. The polymeric solids such as polyethylene can vary between highly ordered and amorphous. Glass is totally amorphous. In their absence of structure, the amorphous materials more nearly resemble and behave like highly viscous liquids. *See* AMORPHOUS SOLID; POLAR MOLECULE.

Above 0 K (-459.67°F) some electrons in the high-energy tail of the electron energy distribution have enough energy to find themselves in the conduction band of an insulator. This is why insulators have a slight conductivity. A. R. Von Hippel assumed that above a critical electric field, these thermally excited electrons would gain energy faster in the conduction band than they would lose it. Thus they would gain enough energy to ionize the solid and cause breakdown. The values of critical field, calculated from this theory, agreed reasonably well with the measured values of the breakdown fields for the alkali halides. *See* BAND THEORY OF SOLIDS.

A simplifying assumption in Von Hippel's treatment is that the electrons in the conduction band all have the same energy. A consequence of this assumption is that since there is no ionization below the critical field, the prebreakdown current can be associated only with these free electrons, implying that (1) the prebreakdown current should be extremely small, (2) the prebreakdown current should be independent of the electric field below the critical field, and (3) the ionization coefficient should be discontinuous—going from zero to a large value at the critical field. Since these predictions are at variance with experiment, new theories were developed. H. Fröhlich assumed an energy distribution for the electrons in the conduction band. C. M. Zener assumed tunneling of electrons into the conduction band because of the high electric field.

Both the Von Hippel and Fröhlich theories neglected interaction of the conduction electrons with the lattice vibrations (phonons). Introducing a further refinement, F. Seitz demonstrated that the phonon interaction is dominant in nonpolar crystals, and is nonnegligible in polar crystals.

Despite the success of the above approaches, there are solids in which other mechanisms also occur, such as thermal, electromechanical, gas-discharge, and electrochemical breakdown. In thermal breakdown, part of the solid reaches a critical temperature at the breakdown field, causing chemical deterioration or melting. Even when thermal breakdown is avoided by using thin specimens, massive electrodes, and pulsed voltages, a gas-discharge mechanism of breakdown may initiate in voids in the specimen, or between the electrodes and the solid. Such discharges may cause treeing (Lichtenberg figures) in the solid, and may chemically and mechanically degrade the solid to the point of breakdown.

Liquids. Strictly speaking, liquids do not have an intrinsic dielectric strength. It is necessary to specify the electrode composition, geometry, and the time duration of the applied field. For very long stress durations, the time is important even for solids. Thus, except for electrode composition, this distinction between liquids and solids may be only a matter of degree. In liquids, microprojections on the cathode and migration of suspended particles (due to the gradient in the field) to the high-field regions contribute to a decrease in dielectric strength. Dissociation of the liquid by field-emitted electrons, electrochemical processes, and thermal bubble formation further complicate electrical breakdown in liquids.

A conditioning effect occurs, also common to vacuum, in which the breakdown strength increases progressively with the first few breakdowns, provided that the breakdown current is limited. Other commonalities with vacuum are curvature and area effects. Contrary to expectation, up to a point the dielectric strength is higher for electrodes of smaller

radius of curvature, and higher for electrodes of smaller area.

The fact that the breakdown strength increases with pressure is indicative that a bubble is formed during the breakdown process. However, if the voltage is applied in the nanosecond range, the pressure dependence vanishes and another breakdown mechanism becomes dominant. A piece of evidence in support of the bubble mechanism of breakdown for longer stress duration is the direct relationship between the boiling point and the breakdown strength for the *n*-alkanes.

Vacuum. As pointed out above, breakdown in vacuum has a number of features in common with liquid breakdown. L. Cranberg suggested that breakdown is initiated when a charged clump of loosely adhering material is removed from one electrode surface by the electric field, strikes the opposite electrode, and thus causes a high enough temperature to produce local evaporation. A discharge then ensues in the metal vapor. (The clump here is analogous to particle-initiated gas or liquid breakdown.)

Cranberg's model predicts that $VE > K$, where V is the breakdown voltage, E is the macroscopic electric field at the electrode where the clump originates, and K is a constant characteristic of a given pair of electrodes. For a uniform field and gap d, this implies that V is proportional to $d^{1/2}$, which agrees reasonably well with experiment. However, it does not predict the area effect, nor the polarity effect where the breakdown voltage is much higher when the pointed electrode is the anode. Moreover, it incorrectly predicts the curvature effect.

L. B. Snoddy was the first to suggest anode vaporization as a result of electron bombardment to account for breakdown. A. J. Ahearn was the first to suggest local heating of the cathode as initiating breakdown. L. C. Van Atta, R. J. Van de Graaff, and H. A. Barton were the first to hypothesize a particle interchange multiplication process. The basic assumption here is that, at a critical voltage, a free charged particle upon striking an electrode produces an avalanche of charged particles by secondary emission, with photoemission also playing a role.

There have since been many variations on these models. However, none has yet correctly predicted (even qualitatively) the area and curvature effects, as well as the gap dependence of approximately $d^{1/2}$. M. Rabinowitz introduced a hypothesis that predicts at least qualitatively the known experimental results, without reference to any particular model in terms of processes that are assumed to occur. He observed that the initiation of vacuum breakdown and gap conduction processes occurs so fast (on the order of 10^{-9} s), compared with the time constants of most breakdown circuits, that only the capacitively stored energy of the electrodes and electrode supports discharges within this time. (Light travels less than 30 cm or 1 ft in 10^{-9} s). His assumption was simply that breakdown cannot occur until there is sufficient stored energy in the electric field in the gap, because only this energy (or a fraction of it) is available to break down the gap.

The energy in the field is the capacitively stored energy, $^1/_2CV^2$. The hypothesis leads to Eq. (3), where

$$W = {}^1/_2 f C V^2 \tag{3}$$

W is the critical energy needed to initiate breakdown (characteristic of the electrodes), C is the capacitance of the electrode system, and f is the fraction of the capacitively stored energy available to produce breakdown, $0 < f < 1$. For a uniform field, this predicts that the breakdown voltage V is proportional to $d^{1/2}$. The curvature and area effects are also qualitatively predicted. Mario Rabinowitz; Steven A. Boggs

Bibliography. R. Bartnikas et al. (eds.), *Engineering Dielectrics*, vol. 1, 1979, vol. 2A, 1983, vol. 2B, 1987, vol. 3, 1993; A. Bradwell et al. (eds.), *Electrical Insulation*, 1983; L. A. Dissado and J. C. Fothergill, *Electrical Degradation and Breakdown in Polymers*, 1992; H. G. Erdman (ed.), *Electrical Insulating Oils*, 1988; D. Kind and H. Karner, *High Voltage Insulation Technology*, 1985; G. G. Raju, *Dielectrics in Electric Field*, 2003.

Electrical interference

Disturbance to the normal or expected operation of electrical or electronic devices, equipment, and systems. Electrical interference is sometimes called radio-frequency interference (RFI) or electromagnetic interference (EMI). Such disturbances may range from nuisances (for example, an electric shaver jamming a nearby broadcast radio receiver) to catastrophes (for example, midair aircraft collision due to EMI navigation errors in holding patterns during bad weather in terminal areas). Electrical noise is a broader term that includes those phenomena that are generally termed electrical interference, but also includes naturally occurring currents or voltages that are more or less continuous and cannot be completely eliminated.

Electrical interference originates from one or more of the following: transmitters such as those used for broadcast, communication, radar, and navigation; artificial incidental emission sources such as from sparking of motor brushes, automotive ignition, and fluorescent lamps; and natural phenomena such as lightning and electrostatic discharges. The interference emissions may be radiated through space or conducted along the paths of electrical wires in power lines and signal cables. *See* COMMUNICATIONS CABLE; ELECTROMAGNETIC RADIATION; TRANSMISSION LINES.

Electrical interference is usually controlled by national regulations, such as those recommended by the International Special Committee on Radio Interference (CISPR), of which there are 25 member nations, including the United States. In the United States, the Federal Communications Commission, through Part 15 (low-level radiators such as computers, wireless microphones, and radio receivers) and Part 18 (industrial, scientific, and medical devices) of its regulations, attempts to mitigate

interference through radiated and conducted emission control. The military sector uses its own standard, MIL-STD-461B, to designate required control of both emission and susceptibility on all purchased equipment and subsystems having electrical or electronic elements. The intelligence agencies regulate allowable emission levels from equipment carrying classified information (known as TEMPEST) through their specifications, NACSIM-5100A.

Excessive electrical disturbance to power mains and through radiation is a form of environmental pollution. Unlike other forms of pollution, however, electrical noise cannot usually be detected with the human senses. Instrumented antennas and conduction probes are used with receivers or spectrum analyzers to detect the presence of emissions and to determine if they might create (or to diagnose) an EMI problem. They are also used for specifications and regulation certification. *See* SPECTRUM ANALYZER.

There are several means of mitigating electrical interference, including grounding, bonding, shielding, filtering, and transient control. There are two major approaches. One is to reduce radiation from boxes, cabinets, racks, and consoles by the design of multilayer printed circuit boards and backplanes, or to use metal or metallized plastic cases with EMI-protected apertures. The other involves reduction of radiation to or from the interconnecting cable by using filter pin connectors or foil-wrap shields terminated with complete coverage by the connector backshell at the respective box bulkhead. *See* BONDING; ELECTRIC FILTER; ELECTRIC TRANSIENT; ELECTRICAL NOISE; ELECTRICAL SHIELDING; ELECTROMAGNETIC COMPATIBILITY. Donald R. J. White

Bibliography. K. Javor, *Introduction to the Control of Electromagnetic Interference*, 1993; M. Mardiguian, *Electromagnetic Control in Components and Devices*, 1971; M. Mardiguian, *EMI Control Methodology and Procedures*, 1989; D. R. J. White, *EMI Control in the Design of Printed Circuit Boards*, vol. 1, 1996.

Electrical loading

The connection of a specific circuit (called the load) to another circuit. **Illustration** *a* shows a source of electrical energy (*S*), a coupling circuit (*C*), and a load (*L*). In power transmission, *S* is the generator of the power plant, *C* represents the connecting transmission lines and transformers, and *L* is, for instance, a household with its electrical appliances. The entire system operates to supply the load demand of *L*, that is, the electrical energy required by the load. In a high-fidelity sound system, *S* is the output of the phonograph cartridge or the tape playback head, *C* is the amplifier, and *L* is the coil of the loudspeaker. *See* ELECTRIC POWER SYSTEMS; SOUND-REPRODUCING SYSTEMS.

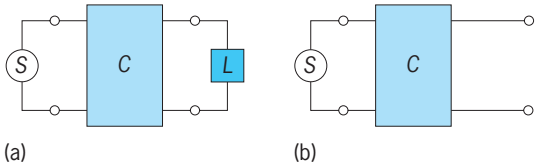

Systems (*a*) under load and (*b*) under no load.

When the load is disconnected, the system operates under no-load conditions (illus. *b*).

Loading is used in certain wire-line transmissions in order to improve their characteristics. This is often done by connecting loading inductors along telephone lines, or along certain data transmission circuits. These inductors reduce the attenuation of the line, thus improving the transmission of signals. Waveguides, special antennas, and coaxial cables are often loaded with specific circuits to achieve specific purposes. *See* ANTENNA (ELECTROMAGNETISM); COAXIAL CABLE; INDUCTOR; TRANSMISSION LINES; WAVEGUIDE. Shlomo Karni

Bibliography. D. J. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; M. Kaufman and A. H. Seidman, *Handbook for Electronics Engineering Technicians*, 2d ed., 1985; J. Markus and C. Weston, *Essential Circuits Reference Guide*, 1988.

Electrical measurements

Measurements of the many quantities by which the behavior of electricity is characterized. Such measurements are of great economic importance because of their impact upon industrial, commercial, and scientific activities. Measurements are required of all electrical quantities, possibly extending over a wide dynamic range and at frequencies ranging from 0 to 10^{12} Hz. The International System of Units (SI) is in universal use for all electrical measurements. Electrical measurements are ultimately based on comparisons with realizations, that is, reference standards, of the various SI units. These reference standards are maintained by the National Institute of Standards and Technology in the United States, and by the national standards laboratories of many other countries. Many electrical measurements are made to measure nonelectrical quantities. For example, platinum-resistance-thermometer measurements facilitate accurate temperature measurements, and eddy-current conductivity measurements are made to assess heat treatment of aluminum alloys. *See* ELECTRICAL UNITS AND STANDARDS.

Direct-current measurements. Direct-current (dc) measurements include measurements of resistance, voltage, and current in circuits in which a steady current is maintained.

Resistance measurement. Resistance is defined as the ratio of voltage to current. For many conductors this ratio is nearly constant, but depends to a varying extent on temperature, voltage, and other environmental conditions. The best standard resistors are made from wires of special alloys chosen for low dependence on temperature and for stability. They may have drift rates as low as 0.01 part per million (ppm) per year. *See* ELECTRICAL RESISTANCE; RESISTOR.

The SI unit of resistance, the ohm, is realized by means of a quantized Hall resistance standard. This is based upon the value of the ratio of fundamental constants h/e^2, where h is Planck's constant and e is the charge of the electron, and does not vary with time. Wire-wound standard resistors can be compared with this standard with an uncertainty of about 1 part in 10^8. *See* HALL EFFECT.

For resistance values below about 1 kilohm, the greatest accuracy is achieved by the use of four-terminal measurements, although this introduces considerable complication into resistive bridges.

Resistors in series produce a value equal to their sum, but resistors in parallel produce a resistance whose reciprocal, called its conductance, is equal to the sum of the conductances of the components. For this reason it is sometimes more convenient to use conductance in siemens, rather than the resistance (in ohms); for example, a resistance of 0.1 Ω is equally well expressed as a conductance of 10 S. *See* CONDUCTANCE.

The principal instruments for accurate resistance measurement are bridges derived from the basic four-arm Wheatstone bridge, and resistance boxes. In the Wheatstone bridge the same current is passed through the unknown and reference resistors, and the resistance ratio is effectively deduced from the ratio of potential differences across them. Alternatively, by using a dc current comparator, the currents through the resistors are adjusted so as to produce equal potential differences across them, when the resistance ratio is given by the ratio of currents. The achievement of constant resistance ratios is much easier than that of constant resistance, so that if a bridge is used only to measure the ratio between a standard resistor and an unknown, it need not be constructed of resistors of the high and expensive quality of the standard. When used in this way, many commercial bridges can provide considerably better accuracy than is available by using their built-in reference resistors. *See* WHEATSTONE BRIDGE.

Many multirange digital electronic instruments measure resistance potentiometrically, that is, by measuring the voltage drop across the terminals to which the resistor is connected when a known current is passed through them. The current is then defined by the voltage drop across an internal reference resistor. For high values of resistance, above a megohm, an alternative technique is to measure the integrated current into a capacitor (over a suitably defined time interval) by measuring the final capacitor voltage. Both methods are capable of considerable refinement and extension. *See* OHMMETER; RESISTANCE MEASUREMENT.

Voltage measurement. The SI unit of voltage, the volt, is realized by using arrays of Josephson junctions. This standard is based on frequency and the ratio of fundamental constants e/h, so the accuracy is limited by the measurement of frequency. The secondary standards, which may be either saturated Weston mercury-cadmium standard cells or semiconductor Zener-diode-based electronic voltage standards, can be compared with the Josephson array with uncertainties of about 1 part in 10^8. Josephson arrays can produce voltages between 200 μV and 10 V.

Weston standard cells are sensitive to temperature, vibration, and the passage of current, and have to be kept in temperature-controlled enclosures stabilized to about 1 millikelvin. Electronic standards do not have these problems, but they may be sensitive to humidity and pressure, and their short-term noise and long-term drift in value are both slightly worse than those of standard cells. Voltage standards are usually compared back to back so that the resulting small voltage difference can be measured by a sensitive digital voltmeter which need not be calibrated to such a high accuracy as if the full voltage of the standard had to be measured.

At the highest levels of accuracy, higher voltages are measured potentiometrically, by using a null detector to compare the measured voltage against the voltage drop across a tapping of a resistive divider, which is standardized (in principle) against a standard cell. Dividers intended for this use up to 2 kV are usually called volt ratio boxes, and sometimes contain built-in self-calibration facilities. Resistive dividers for the accurate measurement of still higher voltages, up to about 500 kV, need increasing care to minimize the effects of the high electric fields, which may give errors due to current leakage through corona discharges. These can be avoided by the use of guarding techniques. *See* CORONA DISCHARGE; POTENTIOMETER.

The Zener diode reference standard is the basis for most commercial voltage measuring instruments, voltage standards, and voltage calibrators. The relative insensitivity to vibration and other environmental and transportation effects makes the diodes particularly useful as transfer standards. Under favorable conditions these devices are stable to a few parts per million per year. *See* ZENER DIODE.

Most dc digital voltmeters, which are the instruments in widest use for voltage measurement, are essentially analog-to-digital converters which are standardized by reference to their built-in reference diodes. Some of these instruments may be microprocessor-controlled, in which case they may be remotely or automatically operated, and may include built-in error correction for linearity over the indicated range. The basic range in most digital voltmeters is between 1 and 10 V, near the reference voltage. Other ranges are provided by means of resistive dividers, or amplifiers in which gain is stabilized by feedback resistance ratios. In this way these instruments provide measurements over the approximate range from 10 nanovolts to 10 kV. *See* DIGITAL-TO-ANALOG CONVERTER; MICROPROCESSOR; VOLTAGE MEASUREMENT; VOLTMETER.

Current measurement. The most accurate measurements of direct currents less than about 1 A are made by measuring the voltage across the potential terminals of a resistor when the current is passed through it. It is important that the resistor has been calibrated under the conditions of use, or that allowances are made for the differences. Higher currents, up to about 50 kA, are best measured by means

of a dc current comparator, which accurately provides the ratio of the high current to a much lower one which is measured as above. At lower accuracies, resistive shunts may be used up to about 5000 A, but the effective calibration of such shunts is a difficult process. *See* CURRENT COMPARATOR; CURRENT MEASUREMENT.

Power frequency measurements. The principal needs for accurate measurement at power frequencies (usually 60 or 50 Hz, but possibly up to 35 kHz for lighting systems) arise from the sale of electrical energy. There are also widespread needs for less accurate measurement for the monitoring and control of generators, distribution systems, and plants.

Alternating-current voltage and current measurements. Alternating-current (ac) voltages are established with reference to the dc voltage standards by the use of thermal converters. These are small devices, usually in an evacuated glass envelope, in which the temperature rise of a small heater is compared by means of a thermocouple when the heater is operated sequentially by an alternating voltage and by a reference (dc) voltage. Suitable series resistors, which have been independently established to be free from variation with frequency, permit direct measurement of power frequency voltages up to about 1 kV. Greater accuracy is provided by multijunction (thermocouple) thermal converters, although these are much more difficult and expensive to make. Thermal converters essentially measure the root-mean-square voltage; for a sinusoidal wave the peak value is $(2)^{1/2}$ times the root-mean-square value. Practical wave-forms are usually not sinusoidal, and the relationship between the mean and peak values will depend on the harmonic content or form factor of the waveform. Improvements in digital electronics have led to alternative approaches to ac measurement. For example, a line frequency waveform may be analyzed by using fast sample-and-hold circuits and, in principle, be calibrated relative to a dc reference standard. Also, electronic root-mean-square detectors may now be used instead of thermal converters as the basis of measuring instruments. *See* NONSINUSOIDAL WAVEFORM; THERMAL CONVERTERS.

Voltages above a few hundred volts are usually measured by means of a voltage transformer, which is an accurately wound transformer operating under lightly loaded conditions. These transformers are usually calibrated by comparison with a high-voltage capacitative divider, which can be independently evaluated. This permits the precise measurement of the voltage ratio and phase error of the transformer under its specified conditions of use.

The principal instrument for the comparison and generation of variable alternating voltages below about 1 kV is the inductive voltage divider. In principle, this resembles a series of tapped autotransformers connected to give a multidecade arrangement. Inductive voltage dividers are very accurate and stable devices, and at best provide an uncertainty of division of less than 1 in 10^8. They are widely used as the variable elements in bridges or measurement systems. *See* INDUCTIVE VOLTAGE DIVIDER.

Alternating currents of less than a few amperes are measured by the voltage drop across a resistor, whose phase angle has been established as adequately small by bridge methods. Higher currents are usually measured through the use of current transformers, which are carefully constructed (often toroidal) transformers operating under near-short-circuited conditions. The performance of a current transformer is established by calibration against an ac current comparator, which establishes precise current ratios by the injection of compensating currents to give an exact flux balance. *See* INSTRUMENT TRANSFORMER.

Commercial instruments for measurement of ac quantities are usually dc measuring instruments, giving a reading of the voltage obtained from some form of ac-dc transducer. This may be a thermal converter, or a series of diodes arranged to have a square-law response, in which case the indication is substantially the root-mean-square value. Some lower-grade instruments measure the value of the rectified signal, which is usually more nearly related to the peak value.

Power and energy measurements. The principal complication in power measurement is that the power is the vector product of the voltage V and current I, so a wattmeter is essentially a vector multiplier of these quantities. Although electromechanical instruments such as dynamometers still exist, electronic instruments are now cheaper and more accurate, and are much more widely used, especially where integration into an automated system is required. A long-established type of energy meter consists of a revolving conducting disk driven by the interaction between induced currents generated by voltage and current coils. Such meters are remarkable for their accuracy (particularly at low energy levels), stability, and low cost. Nevertheless, modern electronic energy meters that are competitive with the rotating disk types are available; they offer the advantages of digital indication and the possibility of remote interrogation via power line communications. *See* ELECTRICAL ENERGY MEASUREMENT; WATT-HOUR METER.

The great majority of wattmeters work at levels of current and voltage less than about 10 A and 250 V, single phase. High-power levels such as the output from power stations at several hundred megawatts per phase are measured through the use of current and voltage transformers.

Several wattmeters have appeared in which the multiplication is carried out electronically. In one method a pulse height is determined by the instantaneous value of one variable and the pulse width by the instantaneous value of the other. The integrated energy of the pulse string then gives the power. A second approach is direct simultaneous sampling of the voltage and current, which are then reduced to digital values. The product of each sample pair is stored, and the accumulated total of the series is a

measure of the power. Another approach to power measurement is the use of the quarter-squares principle, using the identity below. In this method the

$$(V + I\cos\theta)^2 - (V - I\cos\theta)^2 = 4\,VI\,\cos\theta$$

outputs of two square-law devices are summed to give a direct power reading. All these approaches have been shown to be capable of producing accuracy comparable to that of the best wattmeters of more conventional design. *See* ELECTRIC POWER MEASUREMENT; WATTMETER.

Audio- and radio-frequency measurements. At the highest level of accuracy, audio-frequency (af) voltages are measured in the same way as at power frequencies, through the use of thermal converters, and employing inductive voltage dividers to provide variable ratios. Such methods provide uncertainties of about 5 ppm at 1 kHz, increasing to about 30 ppm at 1 MHz.

Both pointer and digital instruments for voltage measurement are similar to those in use at power frequencies. Current measurement is needed less often at these frequencies, and generally only small currents (less than about 1 A) are involved. These can readily be measured using a resistive transducer.

Since a typical af electrical component has both reactance and resistance, the ratio of af voltage to current depends upon the impedance of the component. A large number of bridges are available for the measurement of impedance at these frequencies, and most of these are especially suited to the measurement of particular parameters or ranges of values. The most accurate bridges are those based on transformers and inductive dividers, generally at frequencies below 100 kHz, and using coaxial systems throughout. Automatic network analyzers permit measurements to be made with great convenience within the range 10 kHz–40 GHz. Both vector and scalar instruments are available, with the latter being less expensive but not providing phase information. *See* CAPACITANCE MEASUREMENT; INDUCTANCE MEASUREMENT; RADIO-FREQUENCY IMPEDANCE MEASUREMENTS.

For power measurement, the use of dynamometer wattmeters is usually restricted to frequencies below 10 kHz. The range of electronic wattmeters, that is, those performing electronic vector multiplication, may often extend to 1 MHz, but in this frequency range the measurement of power by vector multiplication of voltage and current is superseded by the use of resistive bolometers or calorimeters. These devices provide a response when they are heated by a known fraction of the power to be measured; this approach may be used up to and beyond microwave frequencies. *See* BOLOMETER.

Microwave frequencies. At frequencies above about 300 MHz, where the circuit dimensions become an appreciable fraction of the wavelength, it becomes necessary to use totally enclosed circuits (waveguides), and the measurements of voltage, current, and lumped impedance are replaced by those of wave impedance, reflection and transmission coefficients, voltage standing-wave ratio, attenuation, and power. The measurement techniques, other than for power measurement, are drastically changed. *See* ATTENUATION (ELECTRICITY); MICROWAVE IMPEDANCE MEASUREMENT; MICROWAVE NOISE STANDARDS; MICROWAVE POWER MEASUREMENT.

Fast pulse measurements. The measurement of electrical pulses with rise times of a few picoseconds is of growing importance in the fields of communications, computing, and digital electronics. The most commonly used instrument for measuring repetitive pulse trains is the sampling oscilloscope, and this is calibrated through electrooptical techniques in which short laser pulses are used to study electrical pulses as they travel along transmission lines on electrooptic crystals. In a variant of this, a small electrooptic crystal probe carrying fast laser pulses is brought close to the transmission line carrying the electric pulse, so that the electric field of the pulse affects the optical transmission of the probe and hence modifies the laser pulses. *See* ELECTROOPTICS; OPTICAL PULSES.

Automated measurement systems. There has been a noticeable trend toward the use of automated measurement systems for electrical measurements, facilitated by the readiness with which modern digital electronic instruments may be interfaced with computers. Many of these instruments have built-in microprocessors, which improve their convenience in use, accuracy, and reliability. A built-in reference standard may be used for self-calibration, and the results of many measurements may be stored, processed, and displayed in a variety of ways. A complex measurement system may comprise several such instruments operating under computer control and automatically producing measurement reports. The overall effect is that the performance of accurate measurements becomes less labor-intensive and more cost-effective. R. Gareth Jones; Owen C. Jones

Bibliography. D. A. Bell, *Electronic Instrumentation and Measurements*, 2d ed., 1994; *Code for Electrical Metering*, ANSI C12.1-1995, 1995; C. H. Dix and A. E. Bailey, Electrical standards of measurement, *Proceedings of the IEE*, vol. 122, no. 10R, October 1975; D. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; A. Hartland, Quantum standards for electrical units, *Contemp. Phys.*, 29:477–498, 1988; Institute of Measurement and Control, London, *A Guide to Measuring Direct and Alternating Current and Voltage below 1 MHz*, 2003; Institute of Measurement and Control, London, *A Guide to Measuring Resistance and Impedance below 1 MHz*, 1999; Instrument Society of America, *ISA Standards Library for Measurement and Control*, 1999; L. D. Jones and A. F. Chin, *Electronic Instruments and Measurements*, 2d ed., 1991; B. P. Kibble and G. H. Rayner, *Coaxial AC Bridges*, 1984; L. Schnell (ed.), *Technology of Electrical Measurements*, 1993; L. M. Thompson, *Electrical Measurements and Calibration*: *Fundamentals and Applications*, 2d ed., 1994.

Electrical model

A mathematical description or electrical equivalent circuit that represents the behavior of a device or system. Models for complex systems are often represented by networks of models for simpler electrical devices such as resistors, capacitors, transistors, and transformers. By using analogies between current or voltage and other physical parameters, equivalent circuits can also be used to analyze thermal, mechanical, magnetic, and acoustic systems. A recurring issue in development and in application of models is the use of simplifying assumptions to allow a compromise between accuracy and complexity. A hierarchy of models often exists, ranging from highly accurate but complex physics-based nonlinear time-domain computer models to linear equivalent-circuit models suitable for hand calculation. The best model for a particular application is the simplest one that predicts the relevant behavior with acceptable accuracy. *See* EQUIVALENT CIRCUIT; NETWORK THEORY.

Electrical models can be divided into nonlinear, large-signal and linear, small-signal models. Each of these can be further divided into time-invariant and time- or frequency-dependent categories.

Large-signal models. These are usually derived by applying physical laws to generalized devices. Since electromagnetic effects propagate across space at the speed of light, the most detailed and accurate version of these models needs to contain a description of device behavior across space and time. Such "distributed-circuit" models take the form of partial differential equations whose independent variables consist of three-dimensional spatial coordinates and time. Electromagnetic field models of this degree of complexity are required where the dimensions of the circuit are large compared with the wavelength at which the circuit operates, or when electromagnetic wave radiation is of interest. Unfortunately, these equations are difficult to solve, even numerically by computer simulation, and usually need to be simplified. *See* MAXWELL'S EQUATIONS.

At lower operating frequencies, where the physical dimensions of a circuit are a small fraction of a wavelength, spatial variations of electrical quantities can be neglected and a system of ordinary differential equations, with time as the sole independent variable, can model the circuit. In most electrical systems modeled at this "lumped-circuit" level, the variables of interest are the voltages at certain locations (nodes) in the circuit and the currents passing through these nodes. Other variables such as electric charge or magnetic flux also may be of interest. Physical models at this level of detail, without additional approximation, are often still highly complex and require lengthy computer solutions. *See* DIFFERENTIAL EQUATION.

An alternative approach is empirical modeling, in which a simple mathematical form is assumed for the model, with key physical parameters being determined by measurements on representative devices. Such models further reduce predictive power for increased simplicity.

Successive simplifications are required to reduce circuit-model complexity to a point where pencil-and-paper calculations can produce an initial new system design. Once this preliminary design is completed, computer simulations using more detailed models can further refine physical circuit-performance predictions. Depending on the degree of simplification employed in the design process, it may be necessary to construct circuit prototypes to verify adequate circuit performance. If the agreement between circuit-model predicted performance and actual physical performance is inadequate, the simplified circuit models needs to be modified to account for these parasitic (unmodeled) effects and the entire design sequence needs to be repeated. In an attempt to avoid this expensive and time-consuming process, much effort has been expended in the continual development and enhancement of computer tools for circuit modeling. *See* COMPUTER-AIDED CIRCUIT DESIGN.

Many devices with lumped-circuit models can be represented with three controlling nodes or terminals. One of these can be selected as a reference to which the other voltages are compared. The result is a two-port network (**Fig. 1***a*), with four variables: two port voltages and two terminal currents. Such a system can be characterized by using two independent equations. Ideally these equations should be expressible in closed-form: two of the variables are expressed as functions of the remaining two. For example, with voltages v_1 and v_2 in Fig. 1*a* as the independent variables, the closed-form model equations can be expressed as Eqs. (1) and (2). For devices

$$i_1(t) = f_1[v_1(t), v_2(t)] \tag{1}$$

$$i_2(t) = f_2[v_1(t), v_2(t)] \tag{2}$$

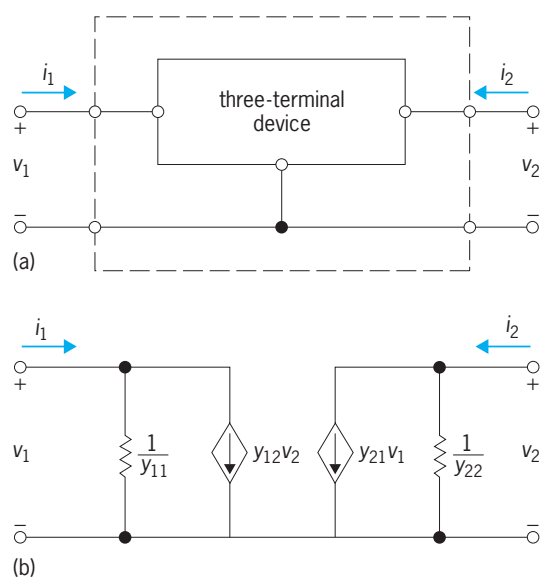

Fig. 1. Three-terminal device. (*a*) Two-port representation. (*b*) Y-parameter equivalent circuit. The diamond-shaped symbols represent voltage-controlled current sources.

controlled by more than three terminals, more port variables and more equations must be added to the model.

Once a suitable model is chosen, it can be used to determine the terminal currents and voltages that result when the device is connected in an electrical circuit. There are two main application areas for large-signal models. One is to find the response of a system to time-varying inputs. The other is to determine the voltages and currents of a circuit with zero input variations, as a preliminary to small-signal analysis.

Small-signal modeling. In many applications, current and voltage variations are small enough that the device's characteristics are approximately independent of signal size. This allows the device to be represented by a simplified linear equivalent circuit. To find the values of the equivalent circuit's elements, the input signal is replaced with its average value, and all currents and voltages are found by using the large-signal model. The results represent the quiescent operating point or Q point for the circuit. All currents and voltages can be approximated by series expansions around the Q point. For small enough signal variations, terms above first order can be neglected. For example (putting aside time derivatives for the moment), the currents in Eqs. (1) and (2) can be written as Eqs. (3) and (4), where the Δ's denote

$$i_1 = I_1 + \Delta i_1 \cong I_1 + \frac{\partial i_1}{\partial v_1} \cdot \Delta v_1 + \frac{\partial i_1}{\partial v_2} \cdot \Delta v_2 \quad (3)$$

$$i_2 = I_2 + \Delta i_2 \cong I_2 + \frac{\partial i_2}{\partial v_1} \cdot \Delta v_1 + \frac{\partial i_2}{\partial v_2} \cdot \Delta v_2 \quad (4)$$

signal-dependent variations, I_1 and I_2 are the Q-point currents, and the derivatives are evaluated at the Q point.

An equivalent circuit can then be drawn in which only the variations of the signal are modeled; all signal-independent voltage and current sources are set to zero. The nonlinear device is replaced in the circuit by its linearized equivalent, whose element values are given by the derivatives in the series expansion. Equations (3) and (4) lead to the equivalent circuit shown in Fig. 1*b*, where the parameters of Eqs. (5) are the derivatives of the nonlinear equa-

$$\begin{aligned} y_1 &= \frac{\partial i_1}{\partial v_1} y_{12} = \frac{\partial i_1}{\partial v_2} \\ y_{21} &= \frac{\partial i_2}{\partial v_1} y_{22} = \frac{\partial i_2}{\partial v_2} \end{aligned} \quad (5)$$

tions evaluated with v_1 and v_2 set equal to their Q-point values. These so-called *y*-parameters have units of conductance. Other small-signal representations result from different choices of dependent and independent variables.

Linearity has several benefits. One is the applicability of systematic solution methods for linear circuits that simplify both manual and computer solutions. Linearity also allows superposition: a linear system's response to the sum of two inputs equals the sum of its responses to the inputs applied separately. An important consequence of superposition is that the system is completely specified by its response to sinusoids of arbitrary frequency. By using Fourier analysis, any input can be written as a unique sum of sinusoids whose magnitude and phase as functions of frequency make up the spectrum of the input. The linear system's response to each sinusoid is a sinusoid of the same frequency. The gain (amplitude increase) and phase shift vary with frequency but not with input amplitude. This frequency response determines the output spectrum for arbitrary inputs. Linear circuits can be analyzed entirely in the frequency domain. *See* FOURIER SERIES AND TRANSFORMS; GAIN; LINEARITY; RESPONSE; SUPERPOSITION THEOREM (ELECTRIC NETWORKS).

Neglecting time derivatives in the small-signal circuit derivation restricts the model's use to the midband region, where the gain is independent of frequency; in the midband equivalent circuit, all inductors and capacitors are treated either as short or open circuits.

Frequency-domain equivalent circuit. At high frequencies, parasitic effects inevitably reduce the gain. This frequency dependence is usually modeled by addition of capacitors or inductors to the model. The values of these elements depend (usually nonlinearly) on the Q point. The external circuit, of course, may also contain frequency-dependent elements. In the time domain, the current and voltage in an inductor or capacitor are related through time integrals or derivatives. These elements can be represented in the frequency domain with an imaginary impedance, directly or inversely proportional to frequency. The circuit can then be analyzed by using complex algebra, which is much easier than solving the time-domain differential equations. As an example of a frequency-domain model, **Fig. 2** shows a high-frequency small-signal equivalent circuit for a bipolar transistor. *See* ALTERNATING-CURRENT CIRCUIT THEORY; TRANSISTOR.

Distortion. If the signal amplitudes are too large, the small-signal approximation is invalid, causing distortion: the generation of frequency components in the output that are not present in the input. Distortion is analyzed in the time domain by using large-signal models. *See* DISTORTION (ELECTRONIC CIRCUITS).

Noise models. Many electrical devices produce random current variations due to the discrete charge of the electron. This electrical noise limits the

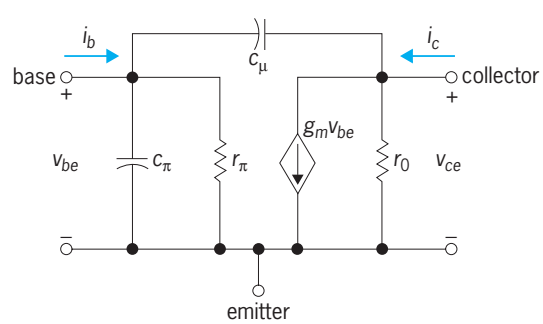

Fig. 2. High-frequency hybrid-π model for a bipolar transistor.

minimum usable signal amplitude. In the time domain, noise signals can be described only statistically. The statistical distributions can be used to represent noise in the frequency domain. The noise power at any frequency can be determined from the autocorrelation function, a measure of how closely two values of the signal measured at different times correspond as the time between the measurements is varied. Physical or empirical models are used to predict the spectrum's dependence on the Q point. *See* CIRCUIT (ELECTRONICS); ELECTRICAL NOISE.

Robert M. Fox; Philip V. Lopresti

Bibliography. G. Massobrio and P. Antognetti, *Semiconductor Device Modeling with SPICE*, 2d ed., 1993; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 5th ed., 2004; Y. P. Tsividis, *Operation and Modeling of the MOS Transistor*, 3d ed., 2006; M. E. Van Valkenburg, *Network Analysis*, 3d ed., 1974.

Electrical noise

Interfering and unwanted currents or voltages in an electrical device or system. Electrical noise, or simply noise, has a significant effect on the design and operation of almost all electrical and optical systems which are used to communicate or process information. Noise is responsible for the familiar static observed on home radio receivers, the clicking sounds on frequency-modulation (FM) radios operating in fringe (near-threshold) areas, and the "snow"-type granularity on the viewing screen of a television receiver displaying a weak signal. In general, noise provides the fundamental limitation to the range over which radio or optical signals can be transmitted and received with integrity. Noise is, therefore, of great importance to the engineers who design and operate such systems.

It is convenient to differentiate between noise which results from human activity and that which is naturally occurring. Noise which results from human activity, such as that generated by an electrical appliance or an automotive ignition, can usually be eliminated or minimized by good design practice (shielding, filtering, equipment location, and so forth). Naturally occurring noise can be further subdivided into that which is irregular or erratic in nature and that which is more or less continuous. An example of noise which is irregular or erratic is that associated with an electrical storm. This type of noise is sometimes dealt with in the system design, but since it is only occasionally present, it does not ordinarily constitute a design limitation. On the other hand, naturally occurring noise which is essentially continuous in time is responsible for the fundamental limitation cited above. The remainder of the article therefore concentrates on this type of noise. *See* ATMOSPHERIC ELECTRICITY; CROSSTALK; ELECTRIC FILTER; ELECTRICAL INTERFERENCE; ELECTRICAL SHIELDING; ELECTROMAGNETIC COMPATIBILITY; ELECTRONIC EQUIPMENT GROUNDING; GROUNDING; STATIC.

Sources. Most noise generation is a consequence of the spontaneous fluctuations which occur within matter at the microscopic level. In electrical circuits these fluctuations give rise to what are commonly referred to as thermal noise and shot noise. Thermal noise is generated by the random motion of free electrons in a resistor or any conductor with resistance. The random motion, and thus the noise generated, is proportional to the temperature of the medium. At absolute zero temperature on the Kelvin scale (−459.67°F), all motion ceases and no noise is generated. Shot noise is most commonly identified with the fluctuations in the current of a vacuum tube caused by the random emission of electrons from its heated cathode. Shot noise is also observed in semiconductor devices as random fluctuations in carrier density when an electric field is applied. There are other types of noise associated with electrical circuits, but shot noise and thermal noise are by far the most important. *See* FREE-ELECTRON THEORY OF METALS; KINETIC THEORY OF MATTER; SEMICONDUCTOR; VACUUM TUBE.

In a system in which signals are transmitted through the atmosphere [for example, amplitude-modulation (AM) or FM radio broadcast, or satellite communications], the receiving system will always receive noise as well as the desired signals. This noise is a result of thermal radiation from the Earth, planets, Sun, Moon, the galaxy (galactic noise), radio-emitting stars, and atmospheric gases. In addition, there is a small background level of thermal radiation, uniformly distributed, which is believed associated with the big bang origin of the universe. All of these noise sources, weighted by the directional characteristics of the receiving antenna, will contribute to the overall system noise. *See* COSMIC BACKGROUND RADIATION; HEAT RADIATION; TERRESTRIAL RADIATION.

In an optical communications system, a signal level is represented by a number of energy packets called photons. The mean arrival rate of the photons at the detector is proportional to the optical intensity or signal strength. At the detector (a photodiode), the photons are absorbed, each creating a hole-electron pair and thus a current in which the electrons are randomly positioned in time and in which the mean number of electrons is proportional to the optical intensity. The statistical nature of this process gives rise to fluctuations in the number of photons representative of a given level and, subsequently, the number of electrons generated to represent that level. If the detector has internal gain as in an avalanche photodiode, each hole-electron pair can create additional hole-electron pairs. This process, however, is statistical in nature, resulting in a mean value of gain but giving rise to additional fluctuations in the generated current. *See* MICROWAVE SOLID-STATE DEVICES; OPTICAL COMMUNICATIONS; OPTICAL DETECTORS; PHOTODIODE; PHOTON.

Mathematical analysis. Because of the statistical or random nature of noise, noise voltages and currents must be dealt with by using the branch of mathematics that deals with random variables, that is, probability theory. A random variable can be characterized by a probability density function $p(x)$. The definition of

$p(x)$ is such that the probability that the random variable x lies between x_1 and x_2 is given by Eq. (1). Obviously, $p(x)$ can be normalized as shown in Eq. (2).

$$\text{Prob}\ (x_1 < x < x_2) = \int_{x_1}^{x_2} p(x)\, dx \tag{1}$$

$$\int_{-\infty}^{\infty} p(x)\, dx = 1 \tag{2}$$

Investigations have shown that fluctuation noise, in general, has a probability density function which is gaussian or normal, provided that the noise can be represented as a large number of independent overlapping samples. The gaussian distribution has the form given in Eq. (3), where σ is the standard deviation, x is the random variable (noise voltage or current), and the average value is zero. *See* DISTRIBUTION (PROBABILITY); PROBABILITY.

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2} \tag{3}$$

If the average value is nonzero, the distribution will have the form illustrated by Eq. (4), where V is the average value. V might be a signal voltage, for example, with the total instantaneous voltage fluctuating about that level.

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-V)/2\sigma^2} \tag{4}$$

In the foregoing, it has been assumed that the statistical properties of the random variable are invariant with time. The statistical process associated with such a system is said to be a stationary process. In experimentally collecting data to determine the probability distribution of a random variable (for example, a noise voltage), two methods are available. First, a large number of samples can be collected from a system at a sequence of times. Second, and alternately, the outputs of a large number of identical systems can be sampled simultaneously. The latter method is called the ensemble method. If the statistics collected by either of these methods are invariant with time, the process is stationary. Furthermore, if both methods yield identical statistics, the process is ergodic. Most noise processes are both stationary and ergodic. For such processes the mean value [V in Eq. (4)] is just the time average of the noise, and the variance σ^2 is the average of the square of the fluctuations about the mean value. *See* PROBABILITY (PHYSICS); STATISTICAL MECHANICS.

Two additional functions that are useful in characterizing random noise are the autocorrelation function $R(\tau)$ and the power spectral density $G(w)$. $R(\tau)$ provides a measure of the dependence between any two values of the noise variable $x(t)$ and $x(t + \tau)$, as specified in Eq. (5). $G(w)$ gives the distribution of average power in noise as a function of frequency w. $G(w)$ and $R(\tau)$ are Fourier transform pairs; that is, they satisfy Eqs. (6), where $j = \sqrt{-1}$. This is a useful relationship because it relates the autocorrelation function of a noise voltage at the output of a linear system to a function $G(w)$ which can be expressed in terms of the transfer function of the system. For example, if white noise (uniform power spectrum) is passed through a system with transfer function $H(w)$, the autocorrelation function at the output of the system is simply the Fourier transform of $H(w)^2$. *See* FOURIER SERIES AND TRANSFORMS.

$$R(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t)\, x(t+\tau)\, dt \tag{5}$$

$$\begin{aligned} G(w) &= \int_{-\infty}^{\infty} e^{-jw\tau} R(\tau)\, d\tau \\ R(\tau) &= \int_{-\infty}^{\infty} e^{-jw^{\tau}} G(w)\, dw \end{aligned} \tag{6}$$

Many network calculations involving shot noise and thermal noise can be handled simply by considering the noise sources as ordinary signal sources but with the distinction that mean-squared values are employed. Total voltages (or currents) can be determined by adding mean-squared values.

Noise figure and noise temperature. A quantity of considerable interest in describing the characteristics of an electrical or optical signal is the ratio of signal power to noise power (S/N) measured at various points in the system. In particular, the signal power–to–noise power ratio at the output of a network, $(S/N)_{\text{out}}$, can be related to the signal power–to–noise power ratio at the input of the network, $(S/N)_{\text{in}}$, by a term called the noise figure F, as expressed by Eq. (7). F is a quality factor of the network because it provides a measure of the degradation introduced by the network to the signal-to-noise ratio. An ideal, lossless, noise-free network is one for which $F = 1$. Frequently F is expressed in decibels as shown by Eq. (8). Typical signal-to-noise ratio minimum requirements may range from 3 to 15 dB for many detection systems or may be as large as 50 dB for a high-quality voice communications circuit. *See* SIGNAL-TO-NOISE RATIO.

$$F = \frac{(S/N)_{\text{in}}}{(S/N)_{\text{out}}} \tag{7}$$

$$F_{\text{dB}} = 10\ \log F \tag{8}$$

In radio astronomy and space communications and, in general, systems which employ very low-noise first-stage amplifiers in the receiving system (for example, masers and parametric amplifiers), it has become common practice to characterize the amplifier and, indeed, the system in terms of an equivalent noise temperature. The noise figure defined previously can be expressed in terms of this equivalent noise temperature, as shown in Eq. (9),

$$F = \frac{T_e}{T_o} + 1 \tag{9}$$

where T_e is the equivalent noise temperature and T_o is the ambient temperature (usually taken as 290 K or 62°F). Then equivalent noise temperature is

expressed in terms of noise figure by Eq. (10). Temperatures are always expressed in kelvins. *See* AMPLIFIER; MASER; PARAMETRIC AMPLIFIER; SPACE COMMUNICATIONS.

$$T_e = (F - 1)T_o \quad \text{K} \tag{10}$$

If the system consists of a series of cascaded networks, each with an equivalent noise temperature T_i and gain G_i, it can be shown that the overall equivalent noise temperature referred to the input terminals of the first stage is given by Eq. (11). For this system or any system in which the overall equivalent noise temperature is known, the noise power spectral density is simply given by kT_e watts/hertz, where k is Boltzmann's constant (1.38×10^{-23} joule/K) and T_e is in kelvins. *See* BOLTZMANN CONSTANT.

$$T_e = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1 G_2} + \cdots + \frac{T_n}{G_1 G_2 \cdots G_{n-1}} \tag{11}$$

Thermal noise. Spontaneous fluctuations in a resistor at temperature T will give rise to a noise voltage at its terminals with mean-squared value given by Eq. (12), where k is Boltzmann's constant, R is the resistance in ohms, and B is the bandwidth in hertz. Equation (12) is often referred to as the Nyquist formula and the thermal noise as Nyquist noise or Johnson noise in recognition of the original investigators.

$$\overline{V_N^2} = 4kTRB \tag{12}$$

Equation (12) is actually an approximate formula, but it is valid with useful accuracy over all radio and microwave frequencies of general interest. It is not, however, valid at optical frequencies. The Nyquist formula shows that the spectral density $4kTR$ is constant. Noise with a constant or uniform distribution with frequency is referred to as white noise.

In a communications system, thermal noise is evidenced throughout the receiving system: the antenna, the transmission lines, amplifiers, and so forth. In systems where very low noise is vital to operation (such as those used in space communications), cryogenic cooling of the first stage of amplification is frequently employed. Liquid helium and liquid nitrogen have commonly been used for this purpose. Receiver noise temperatures approaching 4 K (−452°F) have been demonstrated. In these systems, antenna noise temperature becomes very important.

Antenna noise temperature is defined formally in Eq. (13), where A_e is the effective receiving area of the antenna in m^2, λ is the wavelength in m, $P_N(\theta,\phi)$ is the normalized power pattern of the antenna, and $T(\theta,\phi)$ is the temperature distribution over all space in kelvins. $T(\theta,\phi)$ includes contributions from the atmosphere, the galaxy, radio-emitting stars, and so forth, as well as thermal radiation from the Earth itself. In Eq. (13), Ω is a solid angle and the integration is carried out over 4π steradians. $P_N(\theta, \phi)$ will depend on the pointing direction of the antenna; consequently, the antenna temperature will depend on the pointing direction of the antenna. At microwave frequencies, the antenna temperature is usually a minimum when the antenna is pointed to zenith. *See* ANTENNA (ELECTROMAGNETISM).

$$T_A = \frac{A_e}{\lambda^2} \int\int T(\theta, \phi)\, P_N(\theta, \phi)\, d\Omega \tag{13}$$

Shot noise. Shot noise, like thermal noise, is thermal in origin. However, it differs from thermal noise in two significant ways: (1) for shot noise to be present in a circuit, a voltage (for example, cathode to anode) or electric field must be impressed; (2) shot noise is not characterized by a uniform, wide-band (white) spectral density, but rather has a high-frequency cutoff given approximately by the inverse transit time of an electron from cathode to anode.

The probability that N electrons will be emitted from a heated cathode in a time interval τ is given by the Poisson distribution, Eq. (14), where $\overline{n}$ is the average number of electrons emitted per second. If the electric field is sufficiently strong, all the emitted electrons are attracted to the anode, and the fluctuations in current are given by the fluctuations in emission. This is called the temperature-limited case, and the mean-squared noise current is given by the shot-effect formula of Eq. (15), where e is the charge on an electron, $\overline{I}$ is the average current, and B is the bandwidth in hertz.

$$P(N) = \frac{(\overline{n}\tau)^N e^{-\overline{n}\tau}}{N!} \tag{14}$$

$$\overline{i_s^2} = 2e\overline{I}B \tag{15}$$

If the applied electric field is not sufficiently strong to pull all the electrons to the anode, a negative space charge develops in the vicinity of the cathode which tends to inhibit shot-effect fluctuations. This is the space-charge-limited case, and the shot noise is then approximated by Eq. (16), where Γ^2 is a space-charge reduction factor varying commonly between 0.01 and 1. The value of Γ^2 depends on physical geometry, cathode temperature, and applied voltage. *See* SPACE CHARGE.

$$\overline{i_s^2} = 2e\overline{I}\Gamma^2 B \tag{16}$$

Noise in optical detectors. The reception of an optical signal involves the use of a photodetector in which the optical signal is converted to an electric current. Thus, both optical noise and electronic noise can be expected to influence the signal-to-noise ratio at the receiver output.

Optical noise relates to the random arrival of photons at the detector and subsequent generation of hole-electron pairs. The number of hole-electron pairs generated during the period t to $t + \tau$ is a random variable having the Poisson distribution given

by Eq. (17), where Λ is given by Eq. (18), where η is

$$P(n) = \frac{\Lambda^n e^{-\Lambda}}{n!} \tag{17}$$

$$\Lambda = \int_t^{t+\tau} \frac{\eta}{hf} p(t)\, dt \tag{18}$$

the quantum efficiency (fraction of light absorbed), h is Planck's constant, f is the optical frequency, and $p(t)$ is the optical power.

For the case where $p(t)$ has a constant value p_0, an expression for the signal-to-noise ratio is given as Eq. (19), where e is charge on the electron, P_N is the

$$S/N = \frac{\left(\dfrac{\eta e p_0}{hf}\right)^2}{\dfrac{2\eta e^2 p_0}{hf} B + P_N} \tag{19}$$

mean-square thermal (electronic) noise power associated with amplification of the generated current, and B is the bandwidth in hertz.

Many modern optical communications systems employ avalanche photodiodes. In the avalanche photodiode, a received photon creates a hole-electron pair, which in turn, by means of ionization collisions, creates additional hole-electron pairs. Thus the detector has internal gain. However, the process which gives rise to the gain is itself a random process, and therefore, additonal fluctuation noise is generated. The signal-to-noise ratio for the optical receiver employing an avalanche photodetector is given by Eq. (20), where G represents the interval

$$S/N = \frac{\left(\dfrac{neGp_0}{hf}\right)^2}{\dfrac{2ne^2G^2Fp_0B}{hf} + P_N} \tag{20}$$

gain and F is the excess noise factor.

If, in Eq. (19) or (20), the thermal noise term is sufficiently small compared to the optically generated noise so as to be neglected, the resulting S/N is referred to as the quantum limit. In most present-day systems, however, thermal noise is found to be the dominant term.

Noise measurement. The most common apparatus used to measure noise is a total power receiver consisting of a wide-band amplifier, a square-law detector, an integrator, and an indicator. The sensitivity of the receiver (the minimum detectable noise fluctuation) is generally improved as the front-end bandwidth and the integration time is increased. Sensitive receivers of this type are called total power radiometers and are frequently employed in radio astronomy. *See* RADIOMETRY.

The noise temperature (or noise figure) of an amplifier can be determined by what is commonly called the Y-factor method. The measurement utilizes two calibrated noise sources and a power receiver of the type just described. One of the noise sources is connected to the input of the amplifier, and the output power is measured with the receiver. This is repeated with the second noise source. The ratio of power received in the two cases is recorded and denoted Y. The amplifier noise temperature is then given by Eq. (21), where T_1 and T_2 (where T_2 is

$$T_{\text{amp}} = \frac{T_2 - YT_1}{Y - 1} \quad \text{K} \tag{21}$$

greater than T_1) are the noise temperatures of the calibrated terminations. It is assumed that the amplifier has sufficiently high gain that the noise contribution from the receiver can be neglected.

The total system noise temperature can be measured by using the noise-adding method. In this method an increment of noise is injected into the system and the total received noise power is compared to the noise power without the added increment. Then the total system noise power is given by Eq. (22), where Y is given by Eq. (23), and ΔT is the

$$T_{\text{syst}} = \frac{\Delta T}{Y - 1} \quad \text{K} \tag{22}$$

$$Y = \frac{\text{system noise power} + \text{injected noise power}}{\text{system noise power}} \tag{23}$$

noise temperature of injected noise.

Modern test equipment is available which greatly simplifies the measurement of the noise performance of circuits. Automated instrumentation which provides output directly in terms of noise figures is, for example, readily available. *See* ELECTRICAL NOISE GENERATOR.

Randall W. Kreutel

Bibliography. M. S. Gupta, *Noise in Circuits and Systems*, 1988; J. D. Kraus, *Radio Astronomy*, 2d ed., 1986; J. E. Midwinter, *Optical Fibers for Transmission*, 1979, reprint 1991; R. Morrison, *Noise and Other Interfering Signals*, 1991; R. Pettai, *Noise in Receiving Systems*, 1984; M. Schwartz, *Information, Transmission, Modulation and Noise*, 4th ed., 1990.

Electrical noise generator

A device that produces electrical noise for use in electrical measurements. Electrical noise generators are commonly employed to measure the noise figure or noise temperature of radio receivers. They are also used in various other tests in radar and communications systems. Celestial noise sources are used to calibrate large antennas.

Some standard types of noise generators are hot-wire, diode, gas-discharge tube, hot and cold loads (terminations), and radio-star. A hot-wire noise source consists of the filament of a lamp heated by direct current. Thermal noise having spectral density $4kTR$, where k is Boltzmann's constant and T and R are the temperature and resistance of the filament respectively, is generated across the terminals of the filament. A diode noise generator utilizes the temperature-limited shot effect to generate noise. At frequencies less than the reciprocal transit time of the diode, the noise spectral density is $2e\bar{I}$, where e is the charge on the electron and $\bar{I}$ is the average anode current. A gas-discharge noise generator, commonly

referred to as a noise tube, consists of a fluorescent light tube enclosed in a waveguide. Noise generation is essentially thermal. The noise tube is commonly employed at microwave frequencies. Hot and cold loads consist of well-matched terminations, either transmission line or waveguide, held at a given temperature by using an oven or by applying cryogenic refrigeration. Noise generation is thermal. Common temperatures for noise-generating terminations are nominally 80 and 300 K (−316 and 80°F)

Celestial radio sources (radio stars) are commonly employed as reference noise sources for evaluating the characteristics of very low-noise, high-gain space communications receiving antennas. There are a number of accurately calibrated sources available—the choice depending on system parameters (frequency, antenna gain, system noise temperature, elevation angle, and so forth) and physical location. The most common radio sources employed are Cassiopeia A, Taurus A, Cygnus A, and Orion A. The first three are classified as nonthermal sources in which radiation results from relativistic electrons interacting with an interstellar magnetic field. The electrons are rotated in a plane perpendicular to the magnetic field direction, and radiation is characterized by a component polarized parallel to that plane. The polarized component is small, however, and the major portion of the radiation is unpolarized. The nonthermal sources have flux densities which decrease with increasing frequency and, consequently, tend to have a cutoff frequency above which they are not usable. Orion A is a thermal source in which radiation occurs from a hot, ionized cloud. Orion A has a constant flux density at frequencies above 2 GHz. As a point of reference, a typical value of thermal noise received from Cassiopeia A by a 10-m-diameter (33-ft) antenna operating in the C band (4–6 GHz) would be about 150 K (−190°F). *See* POLARIZED LIGHT.

Other possible sources include the Sun, Moon, and the planets. The solar flux density is equal to that from blackbody radiation at 6000 K (10,300°F) at wavelengths less than 1 cm, but is greater than this at longer wavelengths. At radio wavelengths the radiation can be greatly enhanced during periods of sunspot activity. In general, at longer wavelengths the solar radio emission can be considered as the superposition of a slowly varying component dependent on the sunspot number and a rapidly varying component resulting from solar flares. The polarization of the former has a strong circularly polarized component, while that of the latter is generally random. *See* SUN.

At radio wavelengths, the Moon emits like a thermal blackbody with temperature in the range from 150 to 280 K (−190 to 44°F) depending on the lunar phase. At long wavelengths the lunar-phase dependence is minimal, with small (less than 10%) variations about a nominal temperature of approximately 190 K (−118°F).

It is important that variations in brightness over the source be corrected, that polarized components of flux be accounted for, and that attenuation characteristics of the intervening medium, the atmosphere, be included. *See* ELECTRICAL NOISE; RADIO ASTRONOMY.

Randall W. Kreutel

Bibliography. M. S. Gupta (ed.), *Electrical Noise: Fundamentals and Sources*, 1977; M. S. Gupta, *Noise in Circuits and Systems*, 1988; J. D. Kraus, *Radio Astronomy*, 2d ed., 1986; R. Morrison, *Noise and Other Interfering Signals*, 1991; A. Van der Ziel, *Noise: Sources, Characterization and Measurement,* 1970.

Electrical resistance

Opposition of a circuit to the flow of electric current. Ohm's law states that the current I flowing in a circuit is proportional to the applied potential difference V. It is obeyed very accurately for metallic and many other conductors over a wide range of temperatures. The constant of proportionality is defined as the resistance R. Hence, Eq. (1) holds. If V and I

$$V = IR \tag{1}$$

are measured in volts and amperes, respectively, R is measured in ohms. Microscopically, resistance is associated with the impedance to flow of charge carriers offered by the material. For example, in a metallic conductor the charge carriers are electrons moving in a polycrystalline material in which their journey is impeded by collisions with imperfections in the local crystal lattice, such as impurity atoms, vacancies, and dislocations. In these collisions the carriers lose energy to the crystal lattice, and thus Joule heat is liberated in the conductor, which rises in temperature. The Joule heat P is given by Eq. (2).

$$P = I^2R = IV = \frac{V^2}{R} \tag{2}$$

See CRYSTAL DEFECTS; ELECTRICAL CONDUCTIVITY OF METALS; ELECTRICAL RESISTIVITY; JOULE'S LAW; OHM'S LAW.

Peter A. Schroeder

Electrical resistivity

Opposition of a material to the flow of electricity. For a conductor of uniform cross section A (in square meters), length l (in meters), and resistance R (in ohms, Ω), the resistivity is defined by Eq. (1).

$$\rho = \frac{RA}{l} \qquad (\Omega \cdot m) \tag{1}$$

Unlike the extensive quantity R, ρ is an intensive quantity, that is, an intrinsic property of a material, independent of its amount or shape. The range of resistivity values is extremely wide (see **table**). At the extremes are conductors such as copper with a room-temperature resistivity of $1.7 \times 10^{-8}\ \Omega \cdot m$ and insulators such as quartz with a resistivity greater than $10^{16}\ \Omega \cdot m$. The range is even wider when temperature dependence (a property useful in classifying materials) is taken into account. *See* ELECTRIC INSULATOR; ELECTRICAL RESISTANCE.

Resistivities at room temperature

Material	$\Omega \cdot m$
Poly(tetrafluoroethylene)	10^{18}
Quartz	10^{16}
Diamond	10^{12}
Glass	10^{11}
Sodium chloride*	10^{2}
Silicon, pure	10^{3}
Germanium, pure	5
Metglass 2204 ($Ti_{50}Be_{40}Zr_{10}$)	3
Bismuth	1.2×10^{-6}
Antimony	4.2×10^{-7}
Nichrome	10^{-6}
Iron	10×10^{-8}
Potassium	7×10^{-8}
Copper	1.7×10^{-8}
Copper, pure†	10^{-12}

*At 1000 K (1340° F).
†At 4.2 K (–452° F).

Crystalline metals. For a pure metal such as copper (see **illus.**), the room-temperature resistivity is almost entirely caused by scattering of electrons by lattice vibrations, or, in quantum terms, phonons. As the temperature decreases, so does the phonon density, resulting in a decrease of resistivity. At the temperature of liquid helium (4.2 K or −452°F) this is so low that phonon scattering makes very little contribution to the resistivity, which is then determined by the concentration of imperfections in the crystal lattice. Thus, at 4.2 K and below, the resistivity may be $10^{-11}\ \Omega \cdot m$ or even lower. In an unstrained metal the imperfections responsible for the residual restivity at 4.2 K is often taken as a measure of purity. *See* CRYSTAL DEFECTS; LATTICE VIBRATIONS; PHONON.

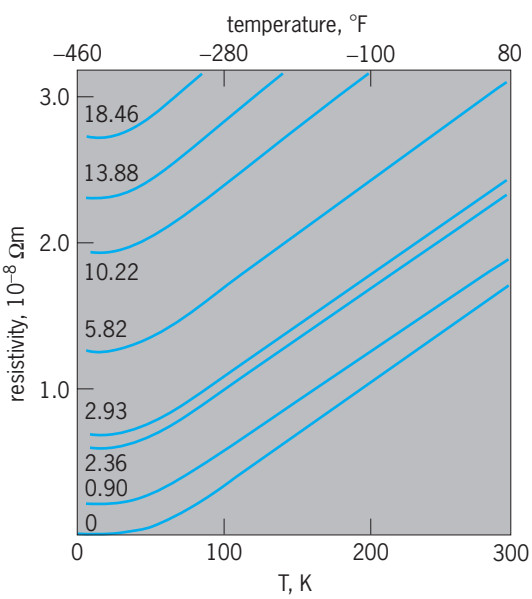

Resistivity of the disordered copper-zinc alloy system. Numbers on the curves give the concentration of zinc in atomic percent. (*After W. E. Henry and P. A. Schroeder, The low-temperature resistivities and thermopowers of α-phase copper-zinc alloys, Can. J. Phys., 41:1076–1093, 1963*)

Many other scattering mechanisms contribute to resistance. Among these is mutual scattering of the charge carriers. Inasmuch as there are about 10^{23} electrons per cubic centimeter in a typical metal, a large contribution might be anticipated. However, the number of electrons that can interact with each other is severely restricted by energy considerations, and the T^2 variation of resistivity with absolute temperature T, typical of electron–electron scattering, is seen only at temperatures of a few kelvins where the phonon contribution to the resistivity becomes negligible. *See* ELECTRICAL CONDUCTIVITY OF METALS.

Disordered crystalline alloys. These are alloys in which the component ions are randomly distributed on the crystal lattice sites. For such an alloy the resistance can be written as in Eq. (2), where ρ_r is

$$\rho = \rho_r + \rho_i \tag{2}$$

the residual resistance produced by electron scattering by the impurity ions, and ρ_i is the resistivity contribution arising from the scattering of electrons by phonons. Matthiessen's rule states that ρ_i is the same for alloys as for the pure metal. For the more dilute alloys in systems such as copper-zinc, this is nearly true, as indicated by the parallelism of the graphs of resistivity versus temperature (see illus.), but Matthiessen's rule is at best an approximation which worsens as the concentration of the alloying constituent increases. Nichrome is an alloy with a high resistivity because of the scattering from a disordered array of the constituent ions. Because of its high resistivity and stability at high temperatures, it is frequently used as a heating element. The resistivity can become very high in an amorphous metal, such as Metglass 2204 (see table), in which all long-range order is lost. But the resistivity of a concentrated alloy may be considerably reduced if the component ions form an orderly array. This occurs in intermetallic compounds, which can be represented chemically as A_pB_q, where p and q are small integers, all the A atoms occupy one orderly array of sites, and all the B atoms occupy another. For example, the gold-gallium compound $AuGa_2$ has a residual resistivity smaller than that of many metals in their purest available form. *See* INTERMETALLIC COMPOUNDS; MATTHIESSEN'S RULE; METALLIC GLASSES.

Semimetals. Bismuth, antimony, arsenic, and graphite are classified as semimetals and have higher resistivities than iron, potassium, and copper (see table), because they have a much lower concentration of mobile electrons that can participate in the conduction process. Their resistivity, like that of metals, increases with temperature.

Semiconductors. The semiconductors silicon and germanium also have a higher resistivity than metals, again because the number of charge carriers is much reduced. They differ from semimetals in that, in their pure or intrinsic form, the number of charge carriers and the conductivity increase exponentially with absolute temperature, and consequently their resistivity decreases as the temperature increases. *See* SEMICONDUCTOR.

Ferromagnetic materials. Several transport properties of ferromagnetic alloys can be explained by assuming conduction in parallel by the spin-up and spin-down electrons. Such a model is assumed to explain the resistance of magnetic multilayers. For example, in a sample consisting of alternate layers of copper and cobalt with layer thicknesses up to 10 nanometers, the magnetizations of the cobalt layers lie in the plane of the layers and, for the appropriate thicknesses of the copper layers, the magnetizations of alternate cobalt layers are in opposite directions. When the magnetizations are made to line up by applying a magnetic field, the resistance is observed to drop so dramatically that this phenomenon is frequently referred to as giant magnetoresistance. The phenomenon is explained in terms of spin-dependent scattering. *See* MAGNETORESISTANCE.

Charge-transfer salts. A metallic conducting state with conductivity decreasing with temperature is also exhibited by organic charge-transfer salts. These consist of large organic donor molecules (cations) which can stack in various ways, and smaller inorganic acceptor molecules which arrange themselves between the stacks. TMTSF (tetramethyltetraselenafulvalene) and BEDT-TTF [bis(ethylenedithiolo)tetrathiafulvalene] are typical of complicated organic donor molecules. The TMTSF stacks in chains which give rise to quasi-one-dimensional conductivity. The BEDT-TTF salts are more two-dimensional. These materials are exciting in that the available choice of anions yield materials of low dimensionality with a multitude of interesting properties, including superconductivity. *See* ORGANIC CONDUCTOR.

Ionic conductors. The resistivity of ionic conductors such as sodium chloride (NaCl; see table) decreases with temperature. However, conduction in these materials is a thermally activated process, in which the conducting ions have to pass over the potential barriers separating them from an adjacent site. *See* IONIC CRYSTALS.

Resistance thermometry. The sensitivity of the resistivity to temperature has practical applications in thermometry. Platinum resistance thermometers are used between 14 and 330 K (−434 and 134°F); rhodium-iron alloys between 1 and 800 K (−458 and 980°F); and carbon and germanium at very low temperatures (0.01 to 100 K or −459.65 to 280°F) where their sensitivity arises from the rapid decrease in charge carrier density as the temperature is lowered. *See* LOW-TEMPERATURE THERMOMETRY; THERMOMETER.

Peter A. Schroeder

Bibliography. J. S. Dugdale, *The Electrical Properties of Metals and Alloys*, 1977; K. M. Hellwege and O. Madelung (eds.), *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology*, *New Series*, *Group 3*: *Crystal and Solid State Physics*, vol. 15a: *Metals*: *Electronic Transport Phenomena*, 1982; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; K. Schröeder, *Handbook of Electrical Resistivities of Binary Metallic Alloys*, 1983.

Electrical shielding

The imposition of a metal or composite barrier between one or more sources of electrical noise and their victims with the objective of reducing or eliminating electrical interference. Examples of the barrier are the case or housing of equipment; shields covering interconnecting cables between equipment; large cabinets, racks, or consoles; shielded (screen) rooms; and entire shielded buildings or vehicles.

Shielding effectiveness. The principal measure of a shield's performance is the shielding effectiveness. It is defined by theequation below, where SE_{dB} is

$$SE_{\mathrm{dB}} = 20 \log_{10} \frac{F_b}{F_a}$$

the shielding effectiveness in decibels, F_b is the electric (or magnetic) field strength before imposition of the barrier, and F_a is the electric (or magnetic) field strength after imposition of the barrier. *See* DECIBEL.

Shielding is obtained by the combination of reflection loss and absorption loss. The former is due to the impedance mismatch between the wave impedance of an oncoming wave and the surface impedance of the interposed barrier. Absorption loss corresponds to the attenuation due to skin effect at higher frequencies, and is dependent upon the frequency, conductivity, permeability, and thickness of the barrier. Shielding effectiveness is the sum of both losses. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; ATTENUATION (ELECTRICITY); RADAR-ABSORBING MATERIALS; REFLECTION AND TRANSMISSION COEFFICIENTS; REFLECTION OF ELECTROMAGNETIC RADIATION; SKIN EFFECT (ELECTRICITY).

Metal shields. Most intentional shields are made of metal to ensure high reflection losses. Even thin metals, such as household aluminum foils (with a thickness of about 1.5 mils or 0.038 mm), offer shielding effectiveness in excess of 100 dB. At low frequencies, these foils become electromagnetically transparent (that is, they do not attenuate magnetic fields). Thus, if there is a shielding problem due to the selection and makeup of the metal barrier, it is likely to occur only for low-frequency magnetic fields.

To obtain significant shielding to magnetic fields at low frequencies, the metal barrier must be very thick or composed of a highly permeable material such as Mumetal, Supermalloy, or Hypernom. Often such shields are fabricated in two or more layers, frequently laminated, to obtain good shielding properties per unit size or weight.

Leaky apertures. The major problem encountered in achieving adequate shielding at other than low frequencies is leaky apertures, not the metal choice. To provide the integrity required of the basic shield metal, all apertures must be protected or secured, or special electrical components and gaskets should be used (see **illus.**).

Shielded rooms. Shielded enclosures, ranging in size from small rooms to entire buildings, are

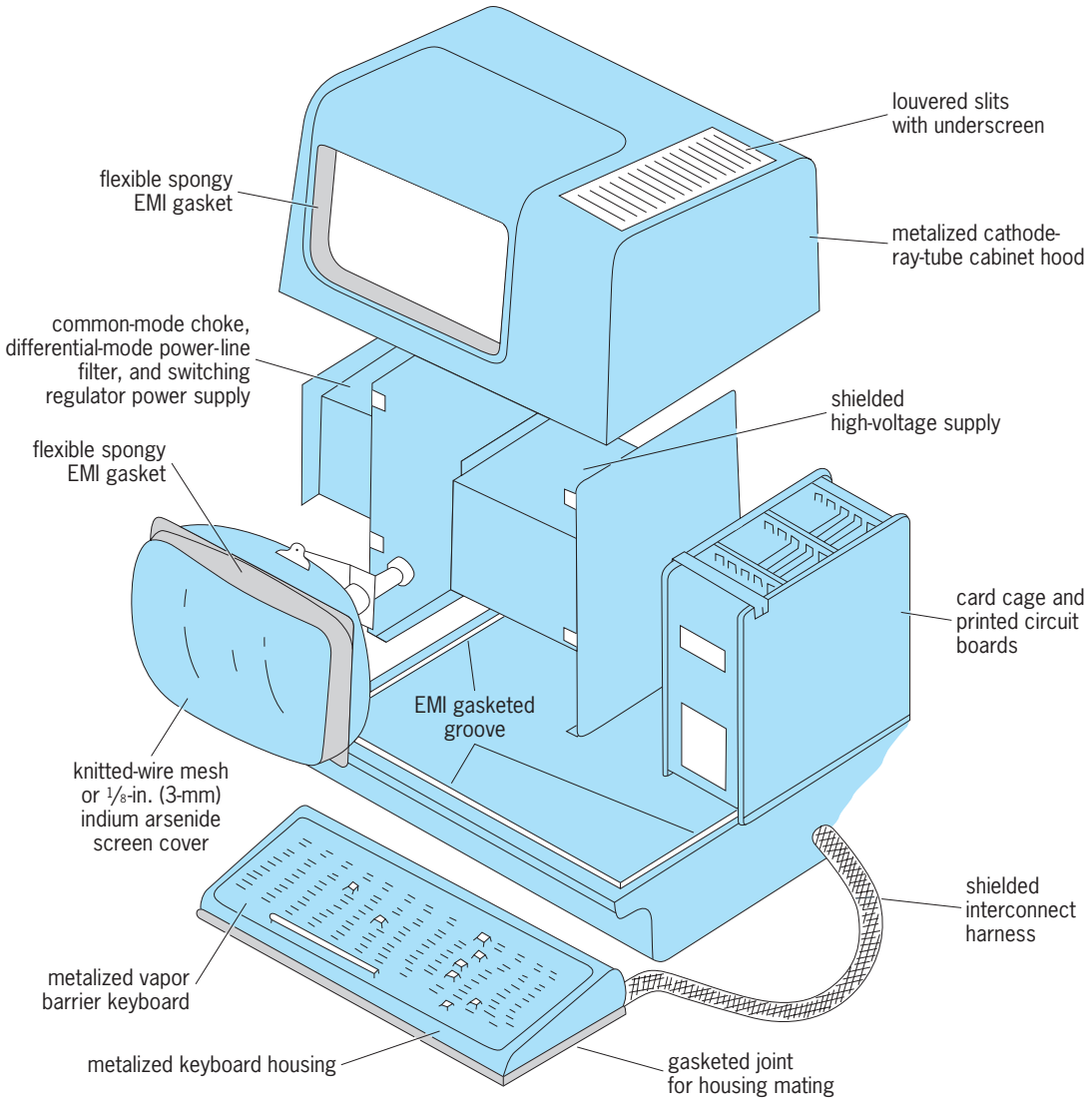

Some measures for controlling electromagnetic interference in video display terminals.

commercially available for performing as radiated emission and susceptibility test chambers, secure facility radiation containment, and protection for sensitive equipment from outside electromagnetic radiation. These enclosures typically provide in excess of 120 dB attenuation from 10 kHz to 1 GHz, and undergo reduced performance at lower and higher frequencies. Where extended performance is required, shielded rooms are usually made of solid seam-welded panels of galvanized steel. Special attention is given to leaky areas such as doors, and entrances where cables or wires such as power-line and telephone are brought in through filters. *See* ELECTRIC FILTER; ELECTROMAGNETIC PULSE (EMP).

Cable shields. Interconnecting cables between equipment often act as an "antenna farm" in which they behave as undesired pickup antennas and provide radiation escape from internal signals conducted along their length. Shielding such cables and harnesses is the dominant protection mechanism; other options include absorbers and filters. Cable shields may vary in complexity from one cover braid, through multibraids and composite braids or foils, to extruded tubes or conduits. *See* COMMUNICATIONS CABLE; ELECTRICAL INTERFERENCE; ELECTRICAL NOISE; ELECTROMAGNETIC COMPATIBILITY.

Donald R. J. White

Bibliography. L. H. Hemming, *Architectural Electromagnetic Shielding Handbook*, 1992; M. Mardiguian and D. R. J. White, *Electromagnetic Shielding*, vol 3: *Interference Control Technologies*, 1988; D. R. J. White, *Handbook on Electromagnetic Shielding Materials and Performance*, 2d ed., 1980.

Electrical units and standards

The process of measurement consists in finding out how many times the quantity to be measured contains a fixed quantity of the same kind, called a unit. The definitions of the units often involve complex physical theory and do not lend themselves readily to practical realization. The concrete representations

of units are known as measurement standards. In practice, measurements are made by using an instrument calibrated against a local reference standard, which itself has been calibrated either directly or by several links in a traceability chain against the national standard held by the national standards laboratory.

In order that measurements of similar quantities made in different countries may be compared, it is important that the values of the standards held by these laboratories should be known in relation to one another. International harmonization of national standards is the concern of the International Bureau of Weights and Measures (Bureau International des Poids et Mesures, or BIPM) that was set up at Sèvres, near Paris, under the Convention of the Meter (1875). The General Conference on Weights and Measures (CGPM) has overall control of the organization at the political level and meets every few years. Scientific questions and the running of the bureau are the responsibility of the International Committee of Weights and Measures (CIPM), which meets annually and is assisted by a number of consultative committees, including the Consultative Committee for Electricity and Magnetism (CCEM). The CCEM encourages research on electrical standards and arranges international comparisons of national standards, including standards of radio frequency and microwave quantities. *See* PHYSICAL MEASUREMENT.

Electrical and Magnetic Units

With the growth of the electric telegraph and supply industries in the nineteenth century, the need was recognized for accurate measurements of electrical quantities, based on generally accepted units and standards. The scientific community played an important part in developing these. From the start, the electrical units had a logical scientific basis, rather than deriving from ancient arbitrary practice.

The cgs system. A proposal by W. E. Weber in 1851 led to the absolute cgs system in which all units of quantities to be measured could be derived from the base units of length, mass, and time—the centimeter, gram, and second. This system was widely adopted although it had three weaknesses: the size of the units was inconvenient for practical use; it was difficult to realize the units from their definitions; and there were separate sets of units for electrostatic and electromagnetic quantities, based respectively on the inverse-square laws of force between electric charges and between magnetic poles. These inverse-square laws involved the permittivity ϵ_0 and the permeability μ_0 of free space (also known as the electric and magnetic constants). Ratios of the electrostatic and electromagnetic units for the same quantity involved the product $\epsilon_0\mu_0$, which J. C. Maxwell showed to be equal to the reciprocal of the square of the speed of light, c_0^{-2}. This was the first example of a relationship between the electrical units and a fundamental physical constant.

International units. The first weakness was resolved by international agreement in 1881 to fix the practical units—the volt, the ohm, and the ampere—at 10^8, 10^9, and 0.1 times the respective cgs electromagnetic units. The other weaknesses were avoided by the decisions of the 1908 International Congress in London, where realizations of these units in terms of easily reproducible standards were defined: the international ohm as the resistance of a column of mercury of specified dimensions; the international ampere as the current which would deposit silver in a silver voltameter at a given rate; and the international volt as their product, with an alternative definition in terms of the electromotive force of a Weston cell. The values were chosen to approximate as closely as the available information allowed to the exact multiples of the cgs units, but the international units were defined in their own right and their values were not changed when the results of later absolute determinations became available. They continued in use until officially abolished by the CGPM in 1948. *See* ELECTROMOTIVE FORCE (CELLS).

The mksa units. A more fundamental change resulted from a proposal by G. Giorgi in 1902. This led to the adoption of the mksa system of units, in which there are four base units: the meter, the kilogram, the second, and the ampere. Use of the meter and the kilogram instead of the centimeter and the gram gave units of a size more convenient for practical use, and use of the ampere as a base unit resolved the conflict between electrostatic and electromagnetic units while maintaining the magnitudes of the widely used practical units. This was a truly coherent system, in the sense that other units were derived from the base units without the need for factors of proportionality other than unity.

SI units. From the mksa system the present-day SI (Système Internationale) has developed, by the addition of further base units to include other fields of measurement. SI units were formally adopted by the CGPM in 1954.

The seven base units of SI are listed in **Table 1**. Of these, the first two, the kilogram and the second, are defined independently of the others. The kilogram is the mass of the international prototype of the kilogram held at the BIPM; the second is the duration of 9 192 631 770 periods of oscillation of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom. The definitions of the other base units involve these two units. Until October 1983, the meter was also independently defined, in terms of the wavelength of a krypton lamp, but at that time

TABLE 1. SI base units

Quantity	Unit	Symbol
Mass	kilogram	kg
Time	second	s
Length	meter	m
Electric current	ampere	A
Thermodynamic temperature	kelvin	K
Luminous intensity	candela	cd
Amount of substance	mole	mol

TABLE 2. Some derived electrical units

Quantity	Unit and symbol	Derivation
Potential difference, emf	volt, V	$W \cdot A^{-1} = m^2 \cdot kg \cdot s^{-3} \cdot A^{-1}$
Resistance	ohm, Ω	$V \cdot A^{-1} = m^2 \cdot kg \cdot s^{-3} \cdot A^{-2}$
Electric charge	coulomb, C	$s \cdot A$
Capacitance	farad, F	$C \cdot V^{-1} = m^{-2} \cdot kg^{-1} \cdot s^4 \cdot A^2$
Conductance	siemens, S	$A \cdot V^{-1} = m^{-2} \cdot kg^{-1} \cdot s^3 \cdot A^2$
Magnetic flux	weber, Wb	$V \cdot s = m^2 \cdot kg \cdot s^{-2} \cdot A^{-1}$
Inductance	henry, H	$Wb \cdot A^{-1} = m^2 \cdot kg \cdot s^{-2} \cdot A^{-2}$
Magnetic flux density	tesla, T	$Wb \cdot m^{-2} = kg \cdot s^{-2} \cdot A^{-1}$
Magnetic field strength	ampere per meter	$m^{-1} \cdot A$
Current density	ampere per square meter	$m^{-2} \cdot A$
Electric field strength	volt per meter	$V \cdot m^{-1} = m \cdot kg \cdot s^{-3} \cdot A^{-1}$
Permittivity	farad per meter	$F \cdot m^{-1} = m^{-3} \cdot kg^{-1} \cdot s^4 \cdot A^2$
Permeability	henry per meter	$H \cdot m^{-1} = m \cdot kg \cdot s^{-2} \cdot A^{-2}$

the CGPM decided to take a fixed value for the speed of light and redefined the meter as the length of the path traveled by light in vacuum in 1/299 792 458 of a second. As a result, it is now possible to realize the meter with a lower uncertainty than previously. *See* LENGTH; LIGHT.

The units of other physical quantities (derived units) are derived from the base units by simple numerical relations, since SI is a coherent system. Examples are the unit of force, the newton ($N = m \cdot kg \cdot s^{-2}$); of energy, the joule ($J = N \cdot m$); and of power, the watt ($W = J \cdot s^{-1}$). Some of the more important derived electrical and magnetic units are listed in **Table 2**.

The SI base unit for electrical measurements is the ampere (A), the unit of electric current. It is defined in terms of a hypothetical experiment as that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross section, and placed 1 meter apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newton per meter of length. According to electromagnetic theory, the force per unit length between two such conductors a distance d apart and carrying a current I is given by Eq. (1).

$$F = \frac{\mu_0 I^2}{2\pi d} \tag{1}$$

The definition of the ampere is thus equivalent to fixing the value of μ_0 as $4\pi \times 10^{-7}$ henry per meter.

The volt (V) is the unit of potential difference and of electromotive force. It is defined as the potential difference between two points of a conducting wire carrying a constant current of 1 ampere when the power dissipated between these points is equal to 1 watt. From the ampere and the volt, the ohm is derived by Ohm's law, and the other derived quantities follow in a similar manner by the application of known physical laws. *See* OHM'S LAW.

The following are the remaining units of electrical and magnetic quantities that have special names in the SI, together with their formal definitions, where appropriate.

Coulomb (C): The unit of electric charge, equal to 1 ampere-second. The coulomb is the quantity of electricity carried in 1 second by a current of 1 ampere.

Farad (F): The unit of capacitance, equal to 1 coulomb per volt. The farad is the capacitance of a capacitor between the plates of which there appears a potential difference of 1 volt when it is charged by a quantity of electricity of 1 coulomb.

Henry (H): The unit of inductance, equal to 1 weber per ampere. The henry is the inductance of a closed circuit in which an electromotive force of 1 volt is produced when the electric current in the circuit varies uniformly at the rate of 1 ampere per second.

Ohm (Ω): The unit of electrical resistance, equal to 1 volt per ampere. The ohm is defined as the resistance between two points of a conductor when a constant potential difference of 1 volt, applied to these points, produces in the conductor a current of 1 ampere, the conductor not being the seat of any electromotive force.

Siemens (S): The unit of electrical conductance (the reciprocal of resistance), equal to 1 ampere per volt. It was formerly known as the mho.

Tesla (T): The unit of magnetic flux density, equal to 1 weber per square meter.

Weber (Wb): The unit of magnetic flux, equal to 1 volt-second. The weber is the magnetic flux which, linking a circuit of one turn, would produce in it an electromotive force of 1 volt if it were reduced to zero at a uniform rate in 1 second.

The mechanical units of frequency (hertz), energy or work (joule), and power (watt) are frequently involved in expressing electrical and magnetic quantities. The cgs units, such as the gauss, gilbert, maxwell, and oersted, formerly used, are not part of the SI and are now obsolete. *See* UNITS OF MEASUREMENT.

Electrical Standards

Realization of the values of the electrical and other units from their SI definitions involves great experimental difficulties. For this reason, it is customary for national standards laboratories to maintain stable primary standards of the units against which other reference standards can be compared. From time to time, absolute determinations of the values of these

primary standards were made in terms of their definitions, using the methods described below.

For a number of years the ampere and the ohm were determined using a current balance and a calculable mutual inductance. The latter were superseded by the calculable capacitor and the moving-coil balance until the late 1980s, when higher accuracies became necessary. Other methods were explored, but by then the Josephson effect and the quantum Hall effect had made possible the standardization of the volt and the ohm by relation to fundamental physical constants. The recommendation by the CCEM of the values to be adopted for these constants in 1990 led to a complete change in primary electrical standards and the method of handling them; and all the major national laboratories, and the BIPM, now use this method.

For many years the primary standards maintained by most laboratories were the volt, in terms of the mean electromotive force of a group of Weston cells, and the ohm, using a group of standard resistors. A range of reference standards of other quantities are derived from these, including direct-current (dc) voltage and resistance at a variety of levels; alternating-current (ac) voltage, resistance, and power; capacitance and inductance; radio-frequency (rf) and microwave quantities; magnetic quantities and properties of materials; dielectric properties; and other quantities. These secondary standards are used for day-to-day measurements and for the calibration of local reference standards of other users in the national measurement system. *See* CAPACITANCE MEASUREMENT; ELECTRIC POWER MEASUREMENT; INDUCTANCE MEASUREMENT; MAGNETIC MATERIALS; MICROWAVE MEASUREMENTS; PERMITTIVITY; RESISTANCE MEASUREMENT; VOLTAGE MEASUREMENT.

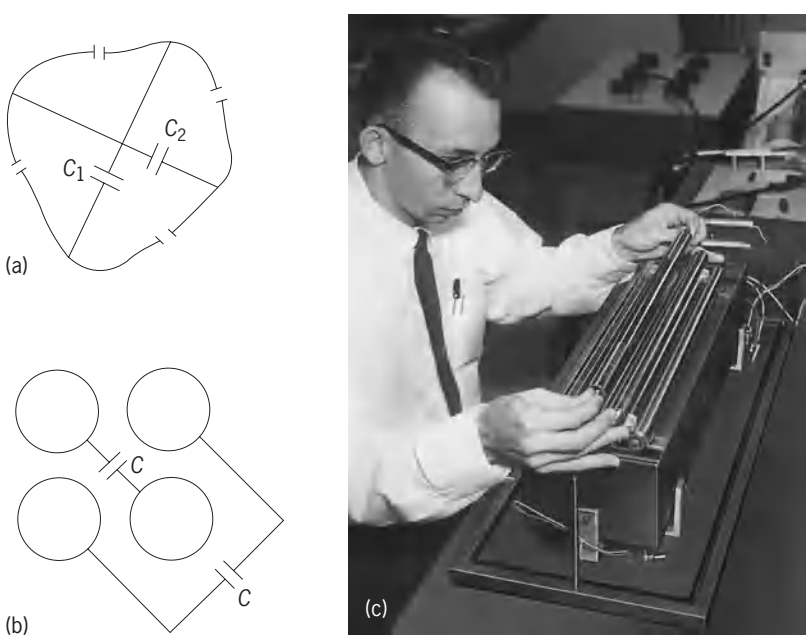

Fig. 1. Calculable capacitor. (*a*) Cross section of general cylindrical capacitor with four segments. (*b*) Cross section of symmetrical capacitor with cylindrical electrodes. (*c*) Calculable capacitor used in absolute determination of the ohm. (*U.S. National Institute of Standards and Technology*)

Absolute Determinations

It has long been the practice that national standards laboratories measure their primary electrical standards in terms of the kilogram and second, with the help of electromagnetic theory and the definition of the ampere. Two different measurements are needed to establish the values of the standards of the volt and the ohm.

For many years the current balance, which measures the force between pairs of cylindrical coils by weighing, has been a mainstay of absolute methods. The other traditional method was the determination of the ohm by comparison with a mutual inductance whose value was calculable from its dimensions. But the accuracy of both measurements has been limited to a few parts in 10^6 by difficulties in determining the dimensions of the coils involved in the experiments and the heating effects of the currents. This order of accuracy is not adequate for modern requirements, and these methods are now of historic interest only. *See* CURRENT BALANCE.

Calculable capacitor. In 1956, A. M. Thompson and D. G. Lampard published a new theorem in electrostatics which showed that if a conducting cylinder of any cross section is divided parallel to the axis into four parts (**Fig. 1***a*) the cross capacitances per unit length, C_1 and C_2, obey Eq. (2).

$$\exp\frac{-\pi C_1}{\epsilon_0} + \exp\frac{-\pi C_2}{\epsilon_0} = 1 \qquad (2)$$

This theorem has been applied as the basis for a capacitor whose change in capacitance as a guard electrode is moved can be calculated exactly from the measurement of a single length. This can be done to a very high degree of accuracy by using a laser interferometer. The problems of making many complex dimensional measurements which limited the accuracy of the current balance and the calculable mutual inductor are thus no longer a limitation. If the capacitor is made symmetrical with four cylindrical electrodes (Fig. 1*b*), then $C_1 = C_2$ and the capacitance per unit length is given by Eq. (3). This is a

$$C = \frac{\epsilon_0(\log_e 2)}{\pi} \qquad \text{farads per meter} \qquad (3)$$

little less than 2 picofarads per meter.

An early capacitor of this kind is shown in Fig. 1*c*. By means of a series of bridges, the value of a 1-ohm resistor can be measured in terms of the calculated reactance of the capacitor. In this way an absolute determination of the value of a standard ohm can be made with an uncertainty (1 standard deviation) of about 5 parts in 10^8.

Electrometer. A number of laboratories have made determinations of the volt by an electrometer method, in which the electrostatic potential difference between two electrodes causes a force which can be measured mechanically. W. G. Clothier devised an electrometer in which the lower electrode is the surface of a pool of mercury. This rises against the force of gravity when the potential is applied,

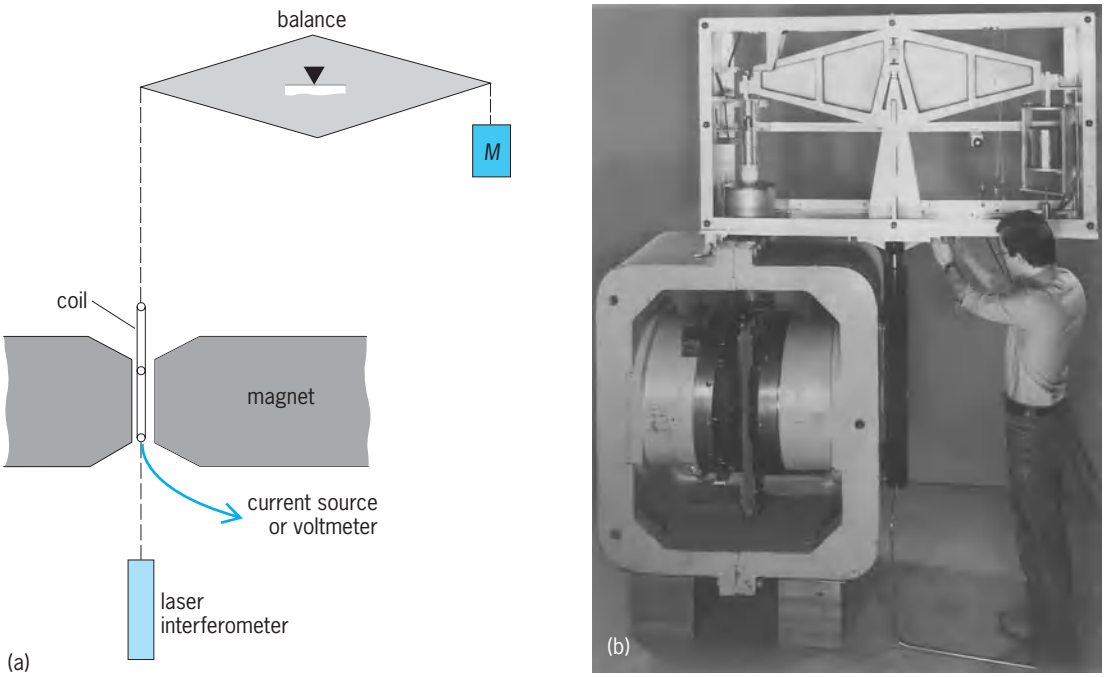

Fig. 2. Moving-coil balance. (*a*) Schematic diagram. (*b*) Equipment. (*U.K. National Physical Laboratory*)

and its movement can be measured very accurately by interferometric means. In this way an uncertainty (1 standard deviation) of about 3 parts in 10^7 has been achieved.

Moving-coil balance. B. P. Kibble developed equipment which in effect measures the electrical watt in terms of the mechanical watt (**Fig. 2**). From this the value of the volt can be determined absolutely. A coil is suspended from the arm of a balance, partly in the field of a powerful magnet. Its position in the vertical plane is measured by a laser interferometer. By using the apparatus successively in two different modes, it is not necessary to know the dimensions of the coil or the strength and distribution of the magnetic field; many errors cancel out, and the problems of accurate dimensional measurements are avoided.

In the first part of the experiment, a current I flows through the coil and the induced force is balanced by a mass M. In the second part, the coil is connected to a voltage-measuring system and made to move through the field with a velocity u, measured by the laser interferometer. The induced voltage V is measured in terms of the primary standard of the volt. Then Eq. (4) follows from the equality of the

$$VI = Mgu \tag{4}$$

electrical and mechanical units of power, where g is the acceleration due to gravity. If the current and the velocity are adjusted so that V is equal to the voltage drop when the current I flows through a resistor whose value R is known accurately in terms of the primary standard of resistance, then V is given by Eq. (5). With equipment based on this principle, an

$$V = (MguR)^{1/2} \tag{5}$$

uncertainty (1 standard deviation) of about 1 part in 10^7 has been achieved.

Josephson effect. If a Josephson junction, consisting of two superconductors separated by an insulating layer so thin that the wave functions are coupled, is irradiated with microwave energy at a frequency ν, it has a voltage–current characteristic that consists of a series of voltage steps (**Fig. 3***a*). The height of these steps is given by Eq. (6), where K_J is known as the

$$\Delta V = \frac{\nu}{K_J} \tag{6}$$

Josephson constant; its value is independent of the materials used and of the experimental conditions. For a practical irradiation frequency the voltage step for a single junction is in the millivolt range, but by using many junctions in series (Fig. 3*b*) it is possible to use the effect to provide a voltage standard in the region of 1 volt and thus to monitor the voltage of a standard cell. *See* JOSEPHSON EFFECT.

Quantum Hall effect. If an electric current I is passed through a semiconductor placed in a magnetic field of induction B_z, at right angles to the current (**Fig. 4***a*), a voltage V_H is developed across the semiconductor in the direction orthogonal to both the current and the field. This is known as the Hall effect, and the ratio R_H, given by Eq. (7), is known as the Hall resistance.

$$R_H = \frac{V_H}{I} \tag{7}$$

In 1980, K. von Klitzing demonstrated that the Hall resistance in certain semiconductors could be quantized in high magnetic fields at very low temperatures. If the semiconductor in Fig. 4*a* consists of a GaAs/AlGaAs heterostructure, in which a layer of

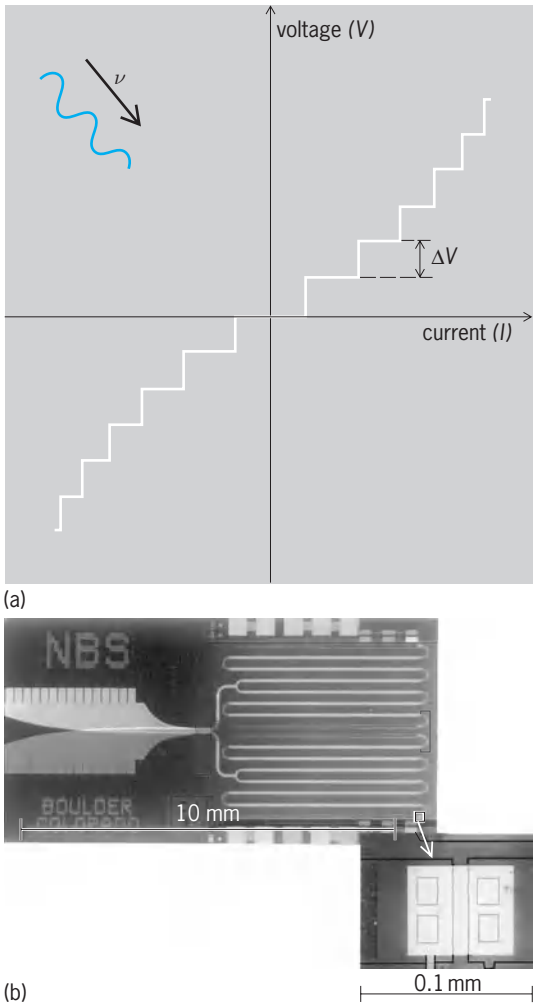

Fig. 3. Josephson effect. (*a*) Josephson junction characteristic. (*b*) Linear array of more than 2000 Josephson junctions in a folded microwave stripline, giving a total voltage step in the region of 1 volt. (*U.S. National Institute of Standards and Technology*)

aluminum gallium arsenide is deposited on a gallium arsenide substrate so that current flow is confined to the interface and is effectively two-dimensional, the Hall resistance takes the form shown in Fig. 4*b* and given by Eq. (8), where $i = 1, 2, 3, \ldots$, and R_K is a

$$R_H = \frac{R_K}{i} \tag{8}$$

constant having a value in the region of 25 kilohms and known as the von Klitzing constant.

Again, the constant is independent of experimental conditions, and it is possible to set up a system working on a selected plateau of the resistance curve of Fig. 4*b* which acts as a resistance standard and can be used to monitor the values of the primary standards of resistance held by national standards laboratories. *See* HALL EFFECT.

International comparisons. For measurements made in one country to be valid in another, it is essential that national primary standards should be extremely stable and their values accurately known in relation to corresponding standards in other countries. The voltages of Weston cells and the resistances of the best wire-wound resistors drift typically by a few parts in 10^7 per year. This is significantly greater than the precision, a few parts in 10^8, with which the values of different standards can be compared. For many years the BIPM has maintained groups of standard cells and standard resistors, similar to those of the national laboratories. It has been the practice of these laboratories to compare their standards every 3 years with those of BIPM by sending traveling standards to Sèvres. In this way it has been possible to monitor variations in national standards over a period of several decades and to adjust their values when the drift has become excessive.

The BIPM also arranges the direct intercomparison of national standards of derived quantities, under the auspices of the CCEM. This has particular value for rf and microwave quantities.

The uncertainties associated with traveling standards and the drifts in value between comparisons have left little margin between national standards and the demands of users for improved accuracy. There are clear advantages if national standards can be referred with sufficient accuracy to some invariant natural physical constant, allowing standards in different countries to be related without transporting them. This has become possible for the volt and the ohm by using the Josephson and quantum Hall effects.

Values of constants. Theory indicates that the values of the Josephson and von Klitzing constants should be equal to $2e/h$ and h/e^2, respectively, where e is the charge on the electron and h is Planck's constant. These are further examples of relations between electrical standards and fundamental physical constants. It might be thought that the effects could be used for absolute determinations of the volt and

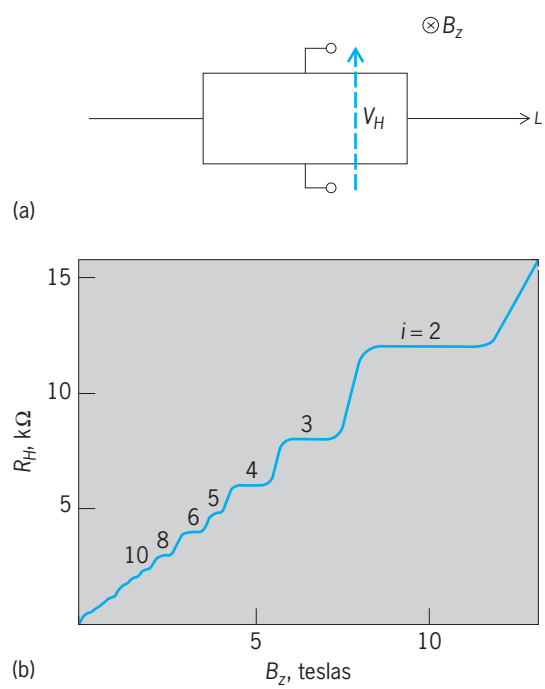

Fig. 4. Quantum Hall effect. (*a*) Geometry of semiconductor. (*b*) Hall resistance characteristic. i = quantum number.

the ohm. This is not so, because measurements of e and h refer to the electrical standards and neither is known independently with sufficient accuracy. It is rather the case that these relations provide a new route for the measurement of e and h.

The values of k_J and R_H have therefore to be determined in terms of the absolute values of the electrical standards. Much effort has been devoted to these measurements and to obtaining international agreement on the values to be used. *See* FUNDAMENTAL CONSTANTS.

International Agreement

As the basis for a generally accepted system of measurement, it is essential that two requirements are satisfied: stability of the primary standards and international agreement on the values to be assigned to them. The first of these has been met for electrical standards by use of the Josephson and quantum Hall effects. The CCEM made an attempt to meet the second requirement in 1972, when it recommended that the value 483 594.0 gigahertz per volt should be used for the Josephson constant. However, this proved to be premature: several major national laboratories continued to use their own, different values, and subsequent measurements have shown this value to be in error by about 8 parts in 10^6.

Recognizing this, the CCEM encouraged further research and in 1986 set up working groups that made critical assessments of the experimental results bearing on the values of the Josephson and von Klitzing constants. In September 1988 the CCEM recommended that from January 1, 1990, all laboratories should adopt the values of 483 597.9 GHz/V for the Josephson constant (to be designated K_{J-90}) and 25 812.807 Ω for the von Klitzing constant (to be designated R_{K-90}). It recommended that all laboratories should adjust the values assigned to their reference standards accordingly. To ensure consistency, it was recommended that these changes should take place on January 1, 1990, and not before. The CIPM subsequently endorsed the CCEM recommendations.

The CCEM estimated the uncertainties in the recommended values of the constants and expressed them by indicating that the Josephson effect together with the value of K_{J-90} can be used to establish a reference standard of emf having an uncertainty (1 standard deviation) relative to the SI volt of 4 parts in 10^7 and reproducibility significantly better; similarly, the quantum Hall effect, together with the value of R_{K-90}, can be used to establish a reference standard of resistance having an uncertainty (1 standard deviation) relative to the SI ohm of 2 parts in 10^7 and again a reproducibility significantly better.

Implementation of the CCEM recommendations required much work and expense on the part of national standards laboratories and national calibration services. The values of national standards of the volt and the ohm had to be adjusted by up to 10 parts per million for the volt and smaller amounts for the ohm. These changes, of course, also affected the secondary standards of other derived electrical quantities and caused consequential changes in the calibration values assigned to accurate instruments.

Much work is continuing at the BIPM and the national laboratories to improve further the accuracy and availability of electrical standards. Josephson arrays operating at 10 V dc are available, and the BIPM has conducted international comparisons of the volt and the ohm using traveling Josephson array standards and quantum Hall resistance standards. These have led to results differing by no more than 2 and 4 parts in 10^9. The BIPM also continues to carry out a variety of international comparisons of derived quantities. *See* ELECTRICAL MEASUREMENTS.

A. Earle Bailey

Bibliography. R. J. Bell (ed.), *SI: The International System of Units*, 6th ed., 1993; Comité Consultatif d'Electricité et Magnétisme, *Rapport de la 18e Session, 1988, 19e Session, 1992, 20e Session, 1995, 21e Session, 1997, 22e Session, 2000, 23e Session, 2002, 24e Session, 2005*, BIPM, Sèvres, France; B. P. Kibble, Present state of electrical units, *Proc. IEE*, 138A:187–197, 1991; B. W. Petley, Electrical units: The last thirty years, *Metrologia*, 32:495–502, 1995; B. N. Taylor and T. J. Witt, New international electrical reference standards based on the Josephson and quantum Hall effects, *Metrologia*, 26:47–62, 1989.

Electricity

Those physical phenomena involving electric charges, their motions, and their effects. The motion of a charge is affected by its interaction with the electric field and, for a moving charge, the magnetic field. The electric field acting on a charge arises from the presence of other charges and from a time-varying magnetic field. The magnetic field acting on a moving charge arises from the motion of other charges and from a time-varying electric field. Thus electricity and magnetism are ultimately inextricably linked. In many cases, however, one aspect may dominate, and the separation is meaningful. *See* ELECTRIC CHARGE; ELECTRIC FIELD; MAGNETISM.

Historical development. The earliest observations of electrical effects were made on naturally occurring substances. Magnetism was observed in the attraction of metallic iron by the iron ore magnetite. The natural resin amber was found to become electrified when rubbed (triboelectrification) and to attract lightweight objects. Both of these phenomena were known to Thales of Miletus (640–546 B.C.). Jerome Cardan in 1551 first clearly distinguished the difference between the attractive properties of amber and magnetite, thus presaging the division of electrical and magnetic effects. He also envisioned electricity as a type of fluid, a viewpoint that was developed more extensively in the late eighteenth and early nineteenth centuries. In 1600 W. Gilbert observed variations in the amounts of electrification

of various substances. He divided substances into two classes, according to whether they did or did not electrify by rubbing. The division actually is into poor and good conductors, respectively. A two-fluid theory was first proposed by C. F. duFay in 1733. A one-fluid theory of electricity was propounded in 1747 by Benjamin Franklin, who called an excess of the fluid positive electrification, and a deficiency of fluid negative electrification. This theory fell into disrepute, but the choice of positive and negative remains. Although fluid theories of electricity were superseded at the end of the nineteenth century, the concept of electricity as a substance persists.

The quantitative development of electricity began late in the eighteenth century. J. B. Priestley in 1767 and C. A. Coulomb in 1785 discovered independently the inverse-square law for stationary charges. This law serves as a foundation for electrostatics. *See* COULOMB'S LAW; ELECTROSTATICS.

In 1800 A. Volta constructed and experimented with the voltaic pile, the predecessor of modern batteries. It provided the first continuous source of electricity. In 1820 H. C. Oersted demonstrated magnetic effects arising from electric currents. The production of induced electric currents by changing magnetic fields was demonstrated by M. Faraday in 1831. In 1851 he also proposed giving physical reality to the concept of lines of force. This was the first step in the direction of shifting the emphasis away from the charges and onto the associated fields. *See* ELECTROMAGNETIC INDUCTION; ELECTROMAGNETISM; LINES OF FORCE.

In 1865 J. C. Maxwell presented his mathematical theory of the electromagnetic field. This theory proposed a continuous electric fluid. It remains valid today in the large realm of electromagnetic phenomena where atomic effects can be neglected. Its most radical prediction, the propagation of electromagnetic radiation, was convincingly demonstrated by H. Hertz in 1887. Thus Maxwell's theory not only synthesized a unified theory of electricity and magnetism, but also showed optics to be a branch of electromagnetism. *See* ELECTROMAGNETIC RADIATION; MAXWELL'S EQUATIONS.

The developments of theories about electricity subsequent to Maxwell have all been concerned with the microscopic realm. Faraday's experiments on electrolysis in 1833 had indicated a natural unit of electric charge, thus pointing toward a discrete rather than continuous charge. Thus, the groundwork for exceptions to Maxwell's theory of electromagnetism was laid even before the theory was developed. H. A. Lorentz began the attempt to reconcile these viewpoints with his electron theory in 1895. He postulated discrete charges, called electrons. The interactions between the electrons were to be determined by the fields as given by Maxwell's equations. The existence of electrons, negatively charged particles, was demonstrated by J. J. Thomson in 1897 using a Crookes tube. The existence of positively charged particles (protons) was shown shortly afterward (1898) by W. Wien, who observed the deflection of canal rays. Since that time, many particles have been found having charges numerically equal to that of the electron. The question of the fundamental nature of these particles remains unsolved, but the concept of a single elementary charge unit is apparently still valid. Of these many particles only two, the electron and the proton, exist in a stable condition on Earth. *See* BARYON; ELECTROLYSIS; ELECTRON; ELEMENTARY PARTICLE; HYPERON; MESON; PROTON; QUARKS.

A second departure from classical Maxwell theory was brought on by M. Planck's studies of the electromagnetic radiation emitted by "black" bodies. These studies led Planck to postulate that electromagnetic radiation was emitted in discrete amounts, called quanta. This quantum hypothesis ultimately led to the formulation of modern quantum mechanics. The most satisfactory fusion of electromagnetic theory and quantum mechanics was achieved in 1948 with the work of J. Schwinger and R. Feynman in quantum electrodynamics, which suppressed the particle aspect and emphasized the field. *See* HEAT RADIATION; QUANTUM ELECTRODYNAMICS; QUANTUM MECHANICS.

Sources. The sources of electricity in modern technology depend strongly on the application for which they are intended.

The principal use of static electricity today is in the production of high electric fields. Such fields are used in industry for testing the ability of components such as insulators and capacitors to withstand high voltages, and as accelerating fields for charged-particle accelerators. The principal source of such fields today is the Van de Graaff generator. *See* PARTICLE ACCELERATOR.

The major use of electricity today arises in devices using direct current and low-frequency alternating current. The use of alternating current, introduced by S. Z. de Ferranti in 1885–1890, allows power transmission over long distances at very high voltages with a resulting low-percentage power loss followed by highly efficient conversion to lower voltages for the consumer through the use of transformers. *See* ALTERNATING CURRENT; ELECTRIC CURRENT.

The development of materials that are superconducting at liquid-nitrogen temperatures may alter this concept. The cost of liquid nitrogen is trivial compared with the cost of the liquid helium previously required for maintaining superconductivity in materials. If the new materials can be formed into malleable wire, this economic factor could lead to the construction of lossless power lines and other applications in computers and in magnetically levitated transportation. *See* MAGNETIC LEVITATION; SUPERCONDUCTING DEVICES; SUPERCONDUCTIVITY.

Large amounts of direct current are used in the electrodeposition of metals, both in plating and in metal production, for example, in the reduction of aluminum ore. To avoid power transmission difficulties, such facilities are frequently located near sources of abundant power. *See* DIRECT CURRENT; ELECTROCHEMISTRY; ELECTROMETALLURGY; ELECTROPLATING OF METALS.

The principal sources of low-frequency electricity are generators based on the motion of a conducting medium through a magnetic field. The moving charges interact with the magnetic field to give a charge motion that is normal to both the direction of motion and the magnetic field. In the most common form, conducting wire coils rotate in an applied magnetic field. *See* ALTERNATING-CURRENT GENERATOR; DIRECT-CURRENT GENERATOR; GENERATOR.

The rotational power is derived from a water-driven turbine in the case of hydroelectric generation, or from a gas-driven turbine or reciprocating engine in other cases. In the latter situation the high-pressure gas that is required is obtained by heating a fluid or gas, using some thermal source, usually the burning of fossil fuels. Economic considerations, particularly the cost of natural gas and oil and the need to conserve them for petrochemical purposes, have pointed out the need for increased reliance on coal and nuclear reactors as heat sources. *See* ELECTRIC POWER GENERATION; GAS TURBINE; HYDROELECTRIC GENERATOR; NUCLEAR REACTOR.

An alternate method is magnetohydrodynamic generation. The requisite motion is that of a high-temperature gas flowing through the magnetic field. The gas consists of the combustion products of the fossil-fuel heat source with a small amount of potassium added to increase the conductivity. Extensive studies have been conducted, especially in the United States and Russia, to determine the feasibility of this method. Because the gas temperatures are very high, the exhaust gas can be used as a heat source for a conventional generator. It is estimated that a combined facility would be 50–55% efficient. This compares with an efficiency of 35–40% for the conventional facility. An added benefit appears to be a ready solution of the sulfur dioxide (SO_2) emission problem for coal-fueled systems. The potassium would efficiently interact with the sulfur dioxide and yield an easily removable solid residue. *See* ACID RAIN; ENERGY SOURCES; MAGNETOHYDRODYNAMIC POWER GENERATOR; MAGNETOHYDRODYNAMICS.

A more direct method of using fission or fusion reactors is the direct conversion of the energy released in the nuclear process into electricity. This has been achieved on a laboratory scale in the case of fission reactors.

Many high-frequency devices, such as communications equipment, television, and radar, involve the consumption of only moderate amounts of power, generally derived from low-frequency sources. If the power requirements are moderate and portability is needed, the use of ordinary chemical batteries is possible. Ion-permeable membrane batteries are a later development in this line. Fuel cells, particularly hydrogen-oxygen systems, are being developed. They have already found extensive application in earth satellite and other space systems. The successful use of thermoelectric generators based on the Seebeck effect in semiconductors has been reported in Russia and in the United States. In a particularly compact low-power device constructed in the United States, the heat needed for the operation of such a generator has been supplied by the energy release in the radioactive decay of suitably encapsulated isotopes produced in fission reactors. *See* BATTERY; FUEL CELL; ION-SELECTIVE MEMBRANES AND ELECTRODES; NUCLEAR BATTERY; THERMOELECTRIC POWER GENERATOR; THERMOELECTRICITY.

The solar battery, also a semiconductor device, has been used to provide charging current for storage batteries in telephone service and in communications equipment in artificial satellites. *See* SOLAR CELL.

Direct conversion of mechanical energy into electrical energy is possible by utilizing the phenomena of piezoelectricity and magnetostriction. These have some application in acoustics and stress measurements. Pyroelectricity is a thermodynamic corollary of piezoelectricity. *See* MAGNETOSTRICTION; PIEZOELECTRICITY; PYROELECTRICITY.

Some other sources of electricity are those in which charged particles are released with some energy and collected in some manner. Charged particles are suitably released in radioactive decay, in the photoelectric effect, and in thermionic emission, among other ways. The photovoltaic effect may also be in this group. *See* ELECTRON EMISSION; PHOTOVOLTAIC EFFECT; RADIOACTIVITY; THERMIONIC EMISSION.

The differences of work functions of various materials can be used for energy conversion. The contact potential difference may be used to convert heat directly to electricity or to provide improved collection for currents arising from some other source such as radioactivity. *See* WORK FUNCTION (ELECTRONICS).

Other possible sources of electricity arise from the existence of electrokinetic potentials in flowing fluids and of phase-transition potentials such as occur in the Workman-Reynolds effect. The possibilities of combining several effects also exist as exemplified in thermogalvanic potentials. It also appears that organic materials (as distinguished from the inorganic materials for which most of the work already described was done) merit investigation. A primitive type of organic solar battery has been developed. *See* ATMOSPHERIC ELECTRICITY; CIRCUIT (ELECTRICITY); CONDUCTION (ELECTRICITY); ELECTRIC POWER MEASUREMENT; ELECTRICAL ENGINEERING; ELECTRICAL UNITS AND STANDARDS; ELECTROKINETIC PHENOMENA; ELECTRONICS; GEOELECTRICITY.

Walter Aron

Bibliography. P. H. Abelson (ed.), *Energy: Use, Conservation and Supply*, American Association for the Advancement of Science, 2 vols., 1974, 1978; B. I. Bleaney and B. Bleaney, *Electricity and Magnetism*, 2 vols., 3d ed., 1976, paper, 1989; R. P. Feynman et al., *Feynman Lectures on Physics: The Definitive and Extended Edition*, 3 vols., 2005; R. J. Fowler, *Electricity: Principles and Applications*, 6th ed., 2004; E. M. Purcell, *Electricity and Magnetism*, 2d ed., 1985; P. B. Visscher, *Fields and Electrodynamics*, 1988.

Electrochemical equivalent

The weight of a substance, according to Faraday's law, produced or consumed by electrolysis with 100% current efficiency during the flow of a quantity of electricity equal to 1 faraday or 96,485.309 coulombs per mole (1 coulomb corresponds to a current of 1 ampere during 1 second). Electrochemical equivalents are essential in the calculation of the current efficiency of an electrode process.

The electrochemical equivalent of a substance is equal to the gram-atomic or gram-molecular weight of this substance divided by the number of electrons involved in the electrode reaction. For example, the electrochemical equivalent of zinc, for which two electrons are required in order to deposit one atom, is Zn/2 or 65.37/2 g. Thus, the faraday is equal to the product of the charge of the electron times the number of electrons (the Avogadro number) required to react with 1 atom- or molecule-equivalent of substance. The value of the faraday computed in this manner agrees with values obtained from electrochemical determinations. The relative error on the value 96,485.309 coulombs is smaller than ±0.01%. *See* COULOMETER; ELECTROLYSIS. Paul Delahay

Electrochemical process

The principles of electrochemistry may be adapted for use in the preparation of commercially important quantities of certain substances, both inorganic and organic in nature.

Inorganic Processes

Inorganic chemical processes can be classified as electrolytic, electrothermic, and miscellaneous processes including electric discharge through gases and separation by electrical means. In electrolytic processes, chemical and electrical energy are interchanged. Current passed through an electrolytic cell causes chemical reactions at the electrodes. Voltaic cells convert chemicals into electricity. Electrothermic processes use electricity to attain the necessary temperature for reaction. For related information *see* ELECTROCHEMISTRY; ELECTROLYSIS; ELECTROLYTIC CONDUCTANCE; ELECTROMOTIVE FORCE (CELLS).

For a discussion of equipment used to convert alternating current to direct current for electrolytic plants. *See* MOTOR-GENERATOR SET; SEMICONDUCTOR RECTIFIER; SYNCHRONOUS CONVERTER.

Voltaic cells are used for the intermittent production of small amounts of electricity. When the chemicals involved are exhausted and must be replaced, the unit is called a primary cell. A special case of the primary cell is the fuel cell in which the fuel and anoxidizer are fed continuously to the cell, converted to electricity, and the products removed. If exhausted components can be revived by passing electricity backward through the unit, it is called a secondary cell, storage battery, or accumulator. Cells may be connected in parallel or in series to form a battery.

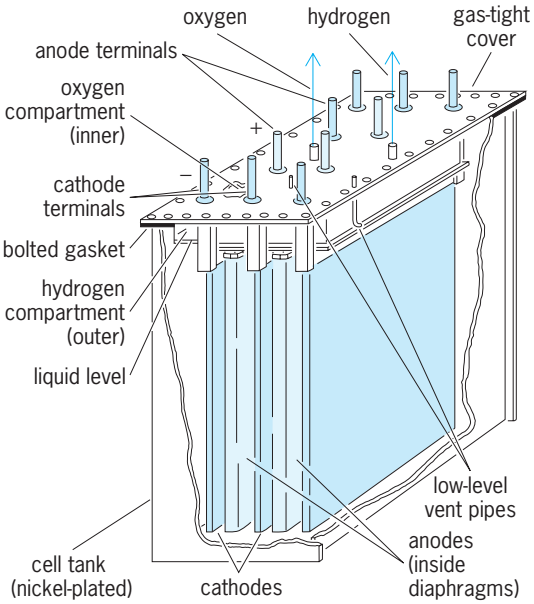

Fig. 1. Diagram of Stuart hydrogen-oxygen cell. (*Electrolyser Corp. Ltd., Toronto, Canada*)

For a discussion of the theory and description of commercial primary, secondary, and fuel cells *see* BATTERY; FUEL CELL.

Electrolysis in aqueous solutions. The electrolysis of water to form hydrogen and oxygen, according to the reaction $2H_2O \rightarrow 2H_2 + O_2$, may be considered as the simplest process for aqueous electrolytes. It does not compete with hydrogen from propane or from natural gas and with oxygen from liquid air, exceptin small installations. While simplicity, high hydrogen purity requirement, and lower capital cost (in small plants) have justified electrolytic plants, severely rising energy costs limited such applications. The electrolyte is 18–30% NaOH or KOH, the latter having a lower resistance but higher cost. The cathode is steel and the anode is nickel-plated steel separated by a diaphragm, usually of asbestos (**Fig. 1**). A cell voltage of 2.0–2.5 V is the summation of the decomposition voltage of water (1.23 V at room temperature) and the oxygen and hydrogen over voltages, plus the *IR* drop through the electrolyte, electrode contacts, and bus bars. The raw material is distilled or demineralized water. There are monopolar cells operating up to 20,000 A, bipolar or filter-press cells using 2000–5000 A, and 150 cells in series at 300–400 V overall. Cells are available which operate at 600 psi (4.1 megapascals). A. T. Kuh described 11 cells using 25–30% KOH and indicated designs operating at pressures up to 200 atm (20 MPa). *See* HYDROGEN; OXYGEN.

Heavy water, or deuterium oxide, used in moderating nuclear reactors is also a by-product of the electrolysis of water. Protium (H^1) is preferentially discharged, so that the electrolyte becomes richer in deuterium. Electrolysis must be combined with catalytic exchange or distillation processes when used

in primary or earlier stages of concentration; it is also used in final stages to produce 99–100% concentration. *See* DEUTERIUM; HEAVY WATER.

Metallurgical applications. Protective or decorative coatings on a base metal such as steel are obtained by electroplating. Plating may also be used to replace worn metal or to provide a wear-resistant surface. The final surface may require several layers of different metals or even layers of the same metal deposited under varying conditions. The metals plated are copper, cadmium, chromium, cobalt, gold, iron, lead, nickel, the platinum metals, silver, tin, andzinc, and many alloys. Electrogalvanizing is preferred over hot dipping for applying zinc to steel. Tin plate for containers is electrolytic. Factors affecting the resulting plate include pretreatment and cleaning of the metal surface, current density, concentration of metal ions, agitation, temperature, conductance of solution, pH, and addition agents. *See* ELECTROPLATING OF METALS; METAL COATINGS.

Electroforming is a method of forming or reproducing articles by electrodeposition. In contrast to electroplating, the product is removed from the base surface or mold. A nonconducting surface can be made conductive by metallizing or with graphite powder. A low-melting-point metal mold can be used, or a mold can be plated with a metal from which the final metal can be removed. The electrodeposits may be up to 0.5 in. (13 mm) thick, and after completion are removed from the mold. Phonograph records, electrotype, textile replicas, and brass instrument bells are common products of electroforming, which is also widely used in the automotive and electronic industries.

Electrodeposition of metal powders is used to produce particles in the 1–1000-micrometer range for use in powder metallurgy and metallic pigments. Powdered iron made electrolytically makes stronger parts than other types, and electrolytic powdered copper makes parts that are easier to machine and have improved wear.

In the anodizing of aluminum articles, a coating approximately 0.001 in. (25 μm) thick is applied, which can be dyed or made impervious. The article to be anodized is cleaned and made anodic in sulfuric acid solution.

Electrolytic polishing of metals is accomplished by making the article anodic in an electrolyte of mixed acids, such as phosphoric-chromic acids or phosphoric-sulfuric acids. High points on the surface apparently dissolve, to give a unique polish not achievable by mechanical polishing.

Electrolytic machining of metals is accomplished by making the metal part anodic in a suitable electrolyte. Metal dissolves at the anode and hydrogen discharges at the cathode. The cathode is contoured as a negative of the shape to be developed in the work piece. By circulating the electrolyte under very high pressure and using high current densities (30–2000 A/in.^{2} or 5–300 A/cm^2), practical machining rates are achieved. Very hard and very thin metal surfaces can be machined without changing the heat treatment.

Electrorefining is a process for purifying metals and recovering their impurities, which at times are more valuable than the original metal. Copper from its ore or scrap is purified of volatilizable impurities, cast into an anode, dissolved in an electrolyte in a cell, and deposited at a cathode as a very pure, highly conductive metal for electrical engineering purposes (**Fig. 2**). Gold, silver platinum, selenium, and tellurium are recovered as by-products. Nickel is freed of copper by using a diaphragm between anode and cathode, with cobalt and the platinum

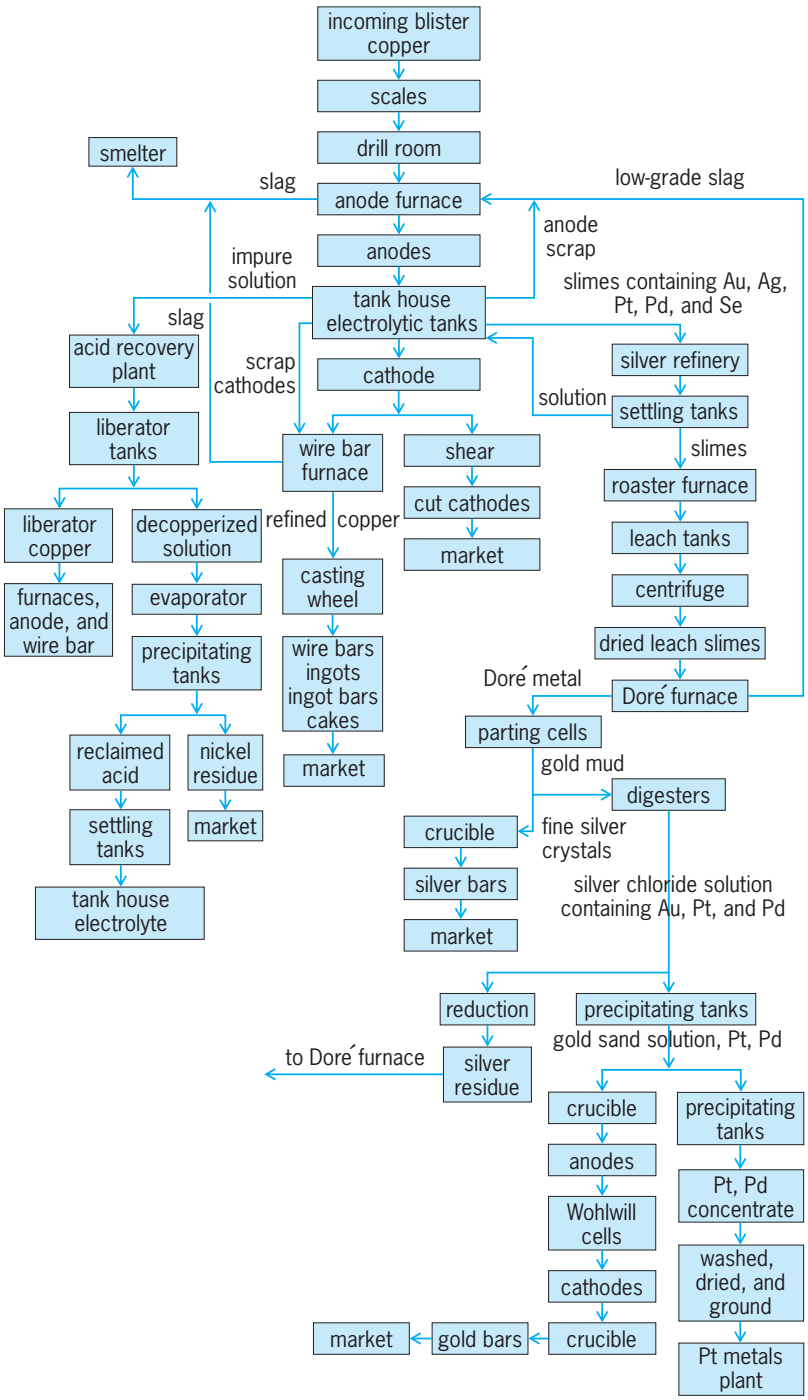

Fig. 2. Flowsheet for a copper plant and the refinery by-products resulting from a product of 99.97% conductor copper. (*Ontario Refining Co., International Nickel Co. of Canada, Ltd.*)

TABLE 1. Electrolytics for various individual metals

Metal	Source	Method	Electrolyte	Application
Antimony	Crude metal	Refining	Sulfates	Alloys, batteries, chemicals
Antimony	Antimony ores	Winning	Sulfides	
Bismuth	Lead refining slimes slimes	Refining	Chlorides	Alloying agent, medicine
Cadmium	Zinc residues and slimes	Winning	Sulfates	Electroplating, alloys
Chromium	Chromium ores	Winning	Sulfates	High-temperature, alloys
Cobalt	Complex cobalt ores	Winning	Sulfates	Alloying agent, electronics
Cobalt	Complex nickel ores	Refining	Sulfates and chlorides	Alloying agent, electronics
Copper	Copper ores, complex ores	Winning	Sulfates	Electrical conductors, wire, brass alloys
Copper	Crude metal, secondary materials, waste	Refining	Sulfates	Electrical conductors, wire, brass alloys
Copper powder	Refined metal	Plating	Sulfates	Powders for powder metallurgy, oilless bearings
Copper sheet	Purified electrolyte	Plating	Sulfates	Sheet for printed circuits, electronics
Gallium	Sodium aluminate liquors	Winning	Caustic	Low-melting-point metals and alloys
Gold	Copper refining slimes	Refining	Chlorides	Jewelry, dentistry, plating
Indium	Residues, wastes, ores	Refining	Chlorides	Silver alloys, jewelry, television
Iron	Low-carbon steel	Refining	Sulfates	Powder metallurgy
Lead	Crude lead	Refining	Fluosilicates	Separation of bismuth etc., alloys, fittings
Manganese	Manganese ores	Winning	Sulfates	Stainless steels, carbon-free metal, alloys
Nickel	Crude metal, nickel matte	Refining	Sulfate-chloride	Plating, alloys, stainless steel
Silver	Copper refinery slimes, lead residues, crude metal, ore	Refining	Nitrates	Jewelry, electrical applications, alloys
Silver	Silver alloys, photo-graphic wastes	Winning	Nitrates	Jewelry, electrical applications, alloys
Solder	Waste and crudes	Refining	Fluosilicate	Metal joining, electrinic components
Tin	Crude metal	Refining	Cresol-sulfonates	Tin plate, solder, alloys
Zinc	Zinc ores, complex ores	Winning	Sulfates	Die-casting alloys, battery cups, brass, galvanizing
Zinc	Ores and residues	Winning	Caustic	Chemicals, paint

metals as by-products. Lead is freed from bismuth; tin from lead, antimony, and bismuth with silver as a by-product; and zinc from copper and lead, with cadmium as a by-product, to make commercial a 99.99% pure metal for diecasting. The last operation is usually referred to as electrowinning. This has rapidly replaced pyrometallurgy or fire processing because it can eliminate sulfur dioxide atmospheric pollution and contamination by particulates.

Electrowinning, sometimes termed aqueous electrometallurgy, involves processing of metallic ores, usually of very low metal content but large in volume, by leaching solutions, usually sulfates, to obtain metal-containing electrolytes which can be processed with insoluble anodes and metal cathodes. Examples are found in cadmium, copper, cobalt, manganese, and zinc.

Electrorefining, electrowinning, and electroforming are summarized as to the individual metals in **Table 1**, while **Table 2** gives anodes, cathodes, and diaphragms for commercially successful operation. **Table 3** lists energy consumption of aqueous electrochemical operations. *See* ELECTROMETALLURGY.

Electrolytic corrosion of metals. This occurs because some parts of the surface of metals act as

TABLE 2. Anodes, cathodes, and diaphragms used on commercial processes

Metal	Method	Anode	Anolyte-catholyte	Diaphragm	Cathode	Voltage
Antimony	Winning	Insoluble	Same	No	Steel	2.5–3
Cadmium	Winning	Insoluble	Same	No	Aluminum	4.0
Chromium	Winning	Insoluble	More acid–Acid	Yes	Hastelloy	4.2
Cobalt	Refining	Soluble	Same	No	Cobalt-stainless steel	2.5
Copper	Refining	Soluble	Same	No	Copper	0.2–0.3
Copper	Winning	Insoluble	Same	No	Copper	2–2.1
Gold	Refining	Soluble	Same	No	Gold	0.5–2.8
Lead	Refining	Soluble	Same	No	Lead	0.35–0.45
Manganese	Winning	Insoluble	Acid–Alkaline	Yes	Stainless steel	
Nickel	Refining	Soluble	Same–Pure	Yes	Nickel	2.4
Silver	Refining	Soluble	Same	Yes	Stainless steel or carbon	1.3–5.4
Tin	Refining	Soluble	Same	No	Tin	0.3
Zinc	Winning	Insoluble	Same	No	Aluminum	3.25–3.7

anodes and corrode, whereas other parts act as cathodes and do not corrode.

Cathodic protection is provided if the whole surface is made cathodic to a separate anode and sufficient voltage is available between the two electrodes. This type of protection is used to inhibit corrosion of boilers, condensers, underground pipelines, ships, and water tanks. Sacrificial anodes of zinc, magnesium, or aluminum alloy may provide the potential, or inert anodes such as graphite, stainless steel, or platinum-plated titanium may be used with power supplied from a rectifier.

Anodic protection can be used to create a passive layer on the surface of some metals, such as steel and stainless steel in some environments. It is a practical method of controlling corrosion of tanks in the chemical industry, but is not feasible for copper or brass vessels. The tank is made anodic. An inert cathode, such as platinum-clad metal, is installed in the liquid in the tank. Current is applied so as to maintain a predetermined voltage between the anodic surface and a reference electrode, such as silver–silver chloride, in the liquid. Equipment to maintain precise potential control at high current output has made anodic protection practical. For example, a current of 0.0015 A/ft^2 (0.016 A/m^2) at +0.900 V to a

TABLE 3. Energy consumption of electrochemical products*

Industry	kWh/lb	lb/kWh	Voltage/tank, cell, or furnace	Range, A/unit	Line voltage	Range, kVA/cell or furnace	kVA/line
Electrolytic refining							
Copper: multiple system,	0.09–0.2	5–13.3	0.2–0.3	6000–15, 000	80–200	1.2–4.5	480–3000
series system	0.074	13.5	120–200				
Gold (troy lb)	0.15	6.6	0.5–2.8	150–500	3–10	0.1–1.4	0.5–5
Lead	0.07–0.09	11.1–14.3	0.35–0.5	5000–6000	100–185	1.75–3.0	500–1200
Nickel	1.1	0.9	2.4–2.6	5000–6000	220–230	12–16	1300–1500
Silver, Moebius (troy lb)	0.27–0.3	3.3–3.7	2.3–2.8	400–500	45–250	1–1.4	18–125
Silver, Thum (troy lb)	0.33–0.6	1.67–3.0	1.3–5.4	150–200	200–220	0.3–1.2	30–50
Solder	0.08	12.5	0.34				
Tin	0.085	11.8	0.3–0.35	5000	100–200	1.5–2.0	500–1000
Electrowinning							
Antimony			2.5–3	1500	300	4	450
Cadmium	0.65–0.97	1.03–1.54	2.5–3.1				
Chromium	5.0–8.4	0.12–0.2	4.2–4.3	10,000	250–300	40–45	2500–3000
Cobalt	1.2–1.56	0.64–0.83	3.7–4.5				
Copper	0.89–1.34	0.77–1.12	2.0–2.12	10,000–25,000	150–200	20–30	1500–5000
Manganese	4–4.5 dc	0.22–0.25	5–5.4	6000	600	30–40	3600
	5–5.3 dc	0.19–0.2					
Silver	0.58–1.0	1–1.73	1–1.5	300		0.3–0.5	0.3–0.5
Zinc	1.4–1.6	0.61–0.71	3–3.7	5000–10,000	600–800	15–40	3000–8000
Metal melting							
Copper	0.12–0.15	6.67–8.33	85–225	12,000–30,000		6000	
Copper alloys	0.15–0.3	3.33–6.67					
Steel: cold charge, hot	0.237–0.35	2.86–4.22	80–250	30,000–100,000		25,000–33,000	
charge	0.05–0.2	5–20	80–250	30,000–100,000		25,000–33,000	
Zinc	0.045	22.2					
Metal powders							
Copper	0.3–0.5	2–3.3	0.3–0.5	10,000–15,000			
Iron	1.20	0.83	2.5				
Nickel	1.13	0.83	1.4–1.5				
Zinc	1.37	0.73	3.4				

*1 kWh/lb = 7.937 MJ/kg; 1 lb/kWh = 0.1260 kg/MJ.

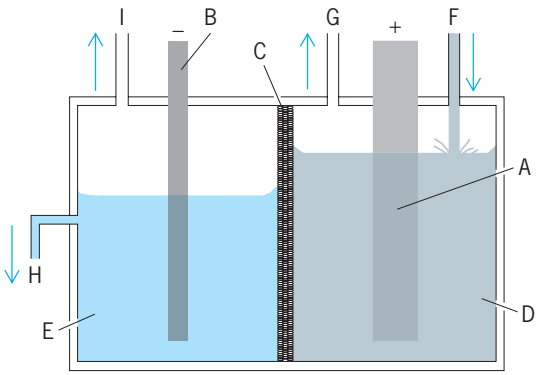

Fig. 3. Diagram of diaphragm cell for chlorine and caustic soda. A = graphite anode, B = iron screen cathode, C = asbestos diaphragm, D = anode compartment for brine and chlorine, E = cathode compartment for NaOH-NaCl cell liquor, F = brine inlet, G = chlorine outlet, H = cell liquor (caustic soda) outlet, I = hydrogen outlet.

silver–silver chloride electrode will maintain passivity of a carbon-steel tank holding 93% sulfuric acid at 27°C (80°F). *See* CORROSION.

Alkali-chlorine processes. Electrolysis of alkali-halides is the basis of the alkali-chlorine and chlorate industries. Chlorine, Cl_2, and caustic soda, NaOH (or caustic potash, KOH), are made by electrolysis of brine, a solution of sodium chloride, NaCl, in water. This is represented by reaction (1).

$$2NaCl + 2H_2O \xrightarrow{\text{electricity}} 2NaOH + Cl_2 + H_2 \qquad (1)$$

Two processes are use to prevent the products from the diaphragm cell and the mercury cell. In the diaphragm cell process (**Fig. 3**) an asbestos diaphragm is interposed between a graphite anode and an iron screen cathode. Saturated purified brine fed around the anode passes through the diaphragm to the cathode. Chlorine is formed at the anode. Hydrogen is released at the cathode, leaving NaOH as a 10–15% solution and 10–15% residual NaCl crystallizes out and is recycled. The decomposition voltage of brine to form chlorine and hydrogen is 2.3 V. At 0.75 A/in.^{2} (11.6 A/dm^2) the average voltage components of a diaphragm cell are as follows:

Anode potential	1.50 V
Cathode potential	1.25
Anolyte	0.47
Diaphragm	0.30
Conductors	0.18
Total	3.70 V

The current efficiency is 95.5–95.5%, due to some oxygen discharge at the anode and some chlorine being carried through the diaphragm. An installation is pictured in **Fig. 4**.

In the mercury cell process brine is electrolyzed between graphite anodes and a flowing mercury cathode, forming a dilute (0.2–0.4%) sodium amalgam which is decomposed in another compartment by water in contact with graphite surfaces to form H_2 and NaOH (**Figs. 5–7**). The products of the mercury cell are purer than those of the diaphragm cell. To offset the cost of mercury, a much higher current density is used in mercury cells. Typical components of voltage in a mercury cell at 5.12 A/in.^{2} (80 A/dm^2) are as follows:

Anode potential, reversible	1.34 V
Cathode potential, reversible	1.76
Decomposition voltage	3.10 V
Anode polarization	0.35 V
Cathode polarization	0.06
Electrolyte	0.60
Conductors and contacts	0.29
Total cell voltage	4.40 V

Economic factors dictate the use of higher current densities, equal to or exceeding 6.5 amperes/in.^{2} (100 A/dm^2). Cell voltage at these higher current densities can be calculated for good cell designs on the market from the equation: $V = 3.20 + 0.015\,C$, where C is the cathode current density in amperes per square decimeter. The current efficiency is approximately 95%. Inefficiency reactions are demonstrated in reactions (2) through (6). Adverse conditions can increase these inefficiencies. *See* CHLORINE.

$$2OH^- \longrightarrow {}^1/_2O_2 + H_2O + 2e \quad \text{(at anode)} \qquad (2)$$

$$H^+ + e \longrightarrow {}^1/_2H_2 \quad \text{(at cathode)} \qquad (3)$$

$$Cl_2 + 2Na(Hg) \longrightarrow 2NaCl + (Hg) \quad \text{(at cathode)} \qquad (4)$$

$$Cl_2 + H_2O \longrightarrow Cl^- + ClO^- + 2H^+ \qquad (5)$$

$$3ClO^- \longrightarrow ClO_3^- + 2Cl^- \qquad (6)$$

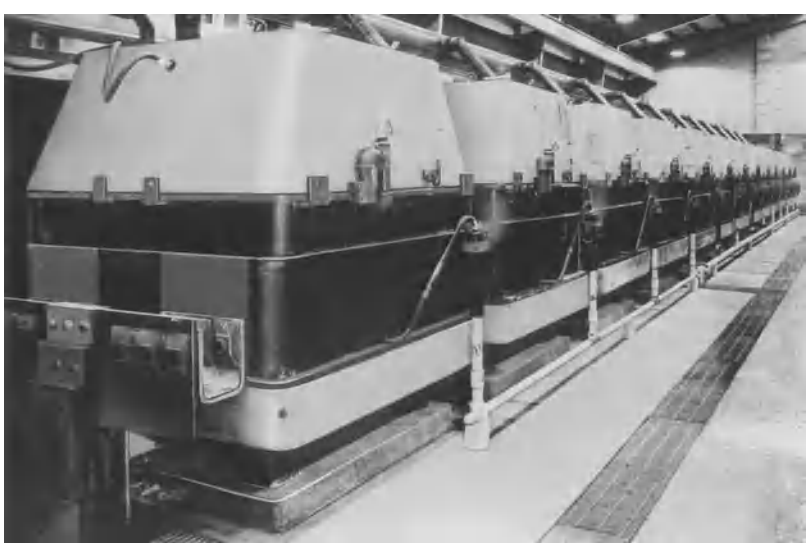

Fig. 4. Photograph of an installation of diaphragm alkali-chlorine cells. (*Hooker Chemical Corp.*)

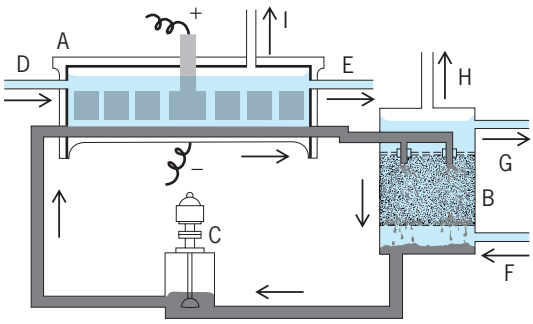

Fig. 5. Diagram of a mercury alkali-chlorine cell. A = electrolyzer, B = decomposer with graphite packing, C = mercury pump, D = feed brine, E = spent brine, F = water, G = 50% caustic soda, H = hydrogen, I = chlorine.

Two factors have been important in the development of chlorine technology: (1) The Nafion diaphragm of the synthetic resin type replacing the deposited asbestos diaphragm. (2) The dimensionally stable anode replacing the graphite anode. The dimensionally stable anode is a titanium substrate with a platinum-group coating of metals and oxides. Its use eliminates the continuously necessary voltage increase or anode-cathode spacing adustment needed because of graphite-anode wear.

In addition, there has been a swing away from mercury cells because of widespread publicity of so-called mercury poisoning by mercury discharges. These discharges have been reduced by better hydrogen cooling and recycle of metal dross, cutting mercury losses as well as permitting better "housekeeping."

Sodium hypochlorite is formed when the products of the electrolysis of brine are mixed. Electrolytic cells have been built for this purpose, but have limited or special use, such as for sterilization of swimming pools and algae control in power plant condensers. Sodium hypochlorite is usually made chemically.

Sodium chlorate is made in cells with graphite or lead peroxide anodes and steel cathodes. When mixing is encouraged, changes take place according to reactions (7)–(9). The overall reaction is (10).

$$6NaCl + 6H_2O \xrightarrow{\text{electricity}} 6NaOH + 3Cl_2 + 3H_2 \quad (7)$$

$$6NaOH + 3Cl_2 \longrightarrow 3NaClO + 3NaCl + 3H_2O \quad (8)$$

$$3NaClO \longrightarrow NaClO_3 + 2NaCl \quad (9)$$

$$NaCl + 3H_2O \xrightarrow{\text{electricity}} 3H_2 + NaClO_3 \quad (10)$$

The temperature is kept below 40°C (104°F) in cells using graphite anodes to prevent excessive attack. The optimum efficiency is at pH 6.8; hydrochloric acid is added as required. Sodium dichromate prevents reduction of chlorate at the steel cathode. The conversion of hypochlorite to chlorate is a some what slow chemical reaction, occurring partly in the cells and partly in a rundown tank. Salt is added and electrolysis is continued until the sodium chloride is down to about 100 g/liter and the chlorate has reached the desired concentration. It is then recovered by crystallization. Cells operate at 3–3.5 V, 1–3 A/dm^2, and 80–85% current efficiency. Energy consumption is about 2.5 kWh/lb (20 megajoules/kg).

Hydrochloric acid electrolysis is of interest for recovery of chlorine from HCl resulting as a byproduct from organic chlorinations. A filter-press electrolyzer is used which has 30–50 unit cells with polyvinyl-cloth diaphragms and graphite electrodes. Hydrochloric acid of 30–33% concentration is fed to the anode compartment. Weak acid is withdrawn at about 20% and reconcentrated by absorbtion of HCl gas. The graphite anode is not attacked as long as the concentration is kept at 20% HCl or higher. The current efficiency is 92–96%, the loss due to electrical leakage. Energy consumption is 1800 kWh/2000 lb (7.1 MJ/kg) chlorine (direct current). The voltage balance of a unit cell is as follows:

Anode potential	1.02 V
Cathode potential	0.28
Anode polarization	0.2
Cathode polarization	0.5
Electrolyte, diaphragm	0.3
Total	2.30 V

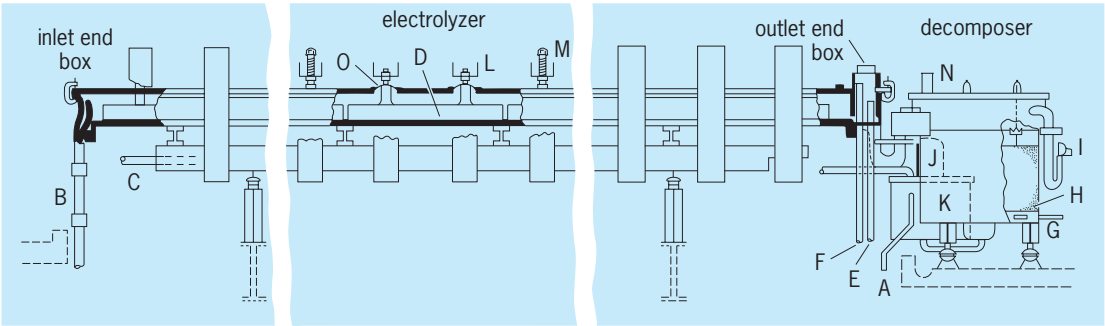

Fig. 6. Longitudinal section of Olin Mathieson E-11 mercury cell. A = dilute caustic outlet, B = brine inlet, C = mercury return, D = anode, E = brine-chlorine outlet, F = outlet end box vent, G = water inlet, H = graphite packing, I = caustic outlet, J = mercury pump, K = mercury pump sump, L = anode support bus, M = lifting screws, N = hydrogen outlet, O = anode seal. (*Olin Mathieson Chemical Corp.*)

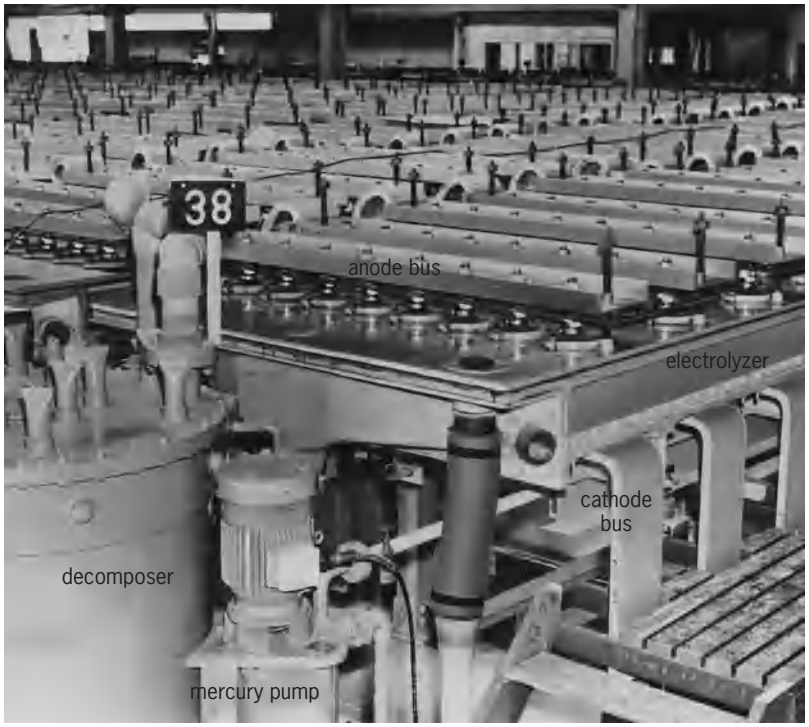

Fig. 7. Chlor-alkali mercury cell; 300,000-A capacity. (*Olin Mathieson*)

Oxidations and reductions. These reactions occur in all cells, but in a narrower sense oxidation reactions are those in which oxygen or chlorine at the anode oxidizes some material to form a new compound; reduction reactions are those in which hydrogen, liberated at the cathode, reduces a material to a new product. There are no commercial applications of inorganic electrochemical reductions by this narrow definition.

Sodium perchlorate is made by oxidation of a solution containing $NaClO_3$ at pH 6.1–6.4 by use of a platinum anode and an iron cathode, with chromate in the electrolyte. A lead dioxide anode may be used with a stainless-steel or nickel cathode, with no chromate in the electrolyte. Energy consumption is 1.4–1.6 kWh/lb (11–13 MJ/kg) $NaClO_4$ (direct current). Other perchlorates are made by metathesis with $NaClO_4$.

Persulfuric acid, $H_2S_2O_8$, is made by oxidizing sulfuric acid as an intermediate in the production of hydrogen peroxide. The reactions for the process are shown in reactions (11) and (12). Alkali persulfates

$$2HSO_4^- \xrightarrow{\text{electricity}} H_2S_2O_8 + 2e \qquad (11)$$

$$H_2S_2O_8 + 2H_2O \longrightarrow 2H_2SO_4 + H_2O_2 \qquad (12)$$

can be made in the same way. The cell has smooth platinum anodes, a porous stoneware diaphragm, and a lead cathode cooled to 30°C (86°F). The reversible potential is 2.18 V and the operating voltage 5.0–5.5 V. The energy released as heat in the cell must be removed by cooling with stoneware or glass coils in the cell. Hydrogen peroxide is recovered by distillation. *See* PEROXIDE.

Lead peroxide anodes are used in sodium chlorate, sodium perchlorate, sodium bromate, and periodate orperiodic acid regeneration cells. The material is dense and made from a lead nitrate solution. For some uses it is produced on a steel base, from which it is removed mechanically and chemically. It is also applied to tantalum, platinum-clad tantalum, or graphite.

Periodic acid is used in producing dialdehyde starch, and spent solution from oxidation can be regenerated in a cell by use of lead dioxide anode, a porous ceramic diaphragm, and an iron cathode.

Electrolytic manganese dioxide for batteries is made by electrolyzing hot $MnSO_4$ solutions at pH 6.5–7.5 by use of graphite electrodes. The MnO_2 deposited on the anode is pulverized with the graphite and separated mechanically. Energy consumption is 1 kWh/lb (8 MJ/kg) MnO_2. The quality of the MnO_2 for battery use depends on cell temperature and anode current density.

Ion-permeable membrane cells. These utilize diaphragms made of ion-exchange resins. Cation-permeable membranes permit cations to pass through but not anions, where as the reverse holds for anion-permeable membranes. **Figure 8** shows the movement of ion and water in an electric membrane stack. Purification of seawater is the most important application. Salt has been recovered from seawater that has been concentrated in this way. *See* ION-SELECTIVE MEMBRANES AND ELECTRODES; WATER DESALINATION.

Fused-salt electrolysis. Aluminum, barium, beryllium, cerium and misch metal, fluorine, lithium, magnesium, sodium, molybdenum, thorium, titanium, uranium, and zirconium are obtained by electrolysis of fused salts, because water interferes with the desired reaction. Raw materials must all be purified before addition to fuse-salt electrolytes as inaqueous electrolytes. Metallizing is a process of depositing a metal as an alloy on a substrate from a fused complex metal salt.

Aluminum is produced in steel pots, lined with carbon, graphite, or silicon carbide, containing an electrolyte of alumina dissolved in fused cryolite, $AlF_3 \cdot 3NaF$, at 950–1000°C (1740–1830°F). The pool of aluminum in the bottom of the pot is cathode. Contact is made by iron bars buried in the carbon lining or by titanium diboride contacts. Anodes are prebaked carbon blocks or the Söderberg type made from carbon paste baked in place. The arrangement of the apparatus for aluminum production is shown in **Fig. 9**. Aluminum is siphoned out periodically. Oxygen released at the carbon anode forms carbon monoxide. *See* ALUMINUM.

Barium, used in the electronics industry and in alloys, is obtained by aluminum reduction at low pressures and high temperatures.

Beryllium is made by batch electrolysis of fused salt, starting with 25% $BeCl_2$ and 75% NaCl in a chrome-iron pot, which acts as cathode to a graphite anode. Beryllium is deposited on the wall of the pot and is cleaned out and broken up when cold. Salt is

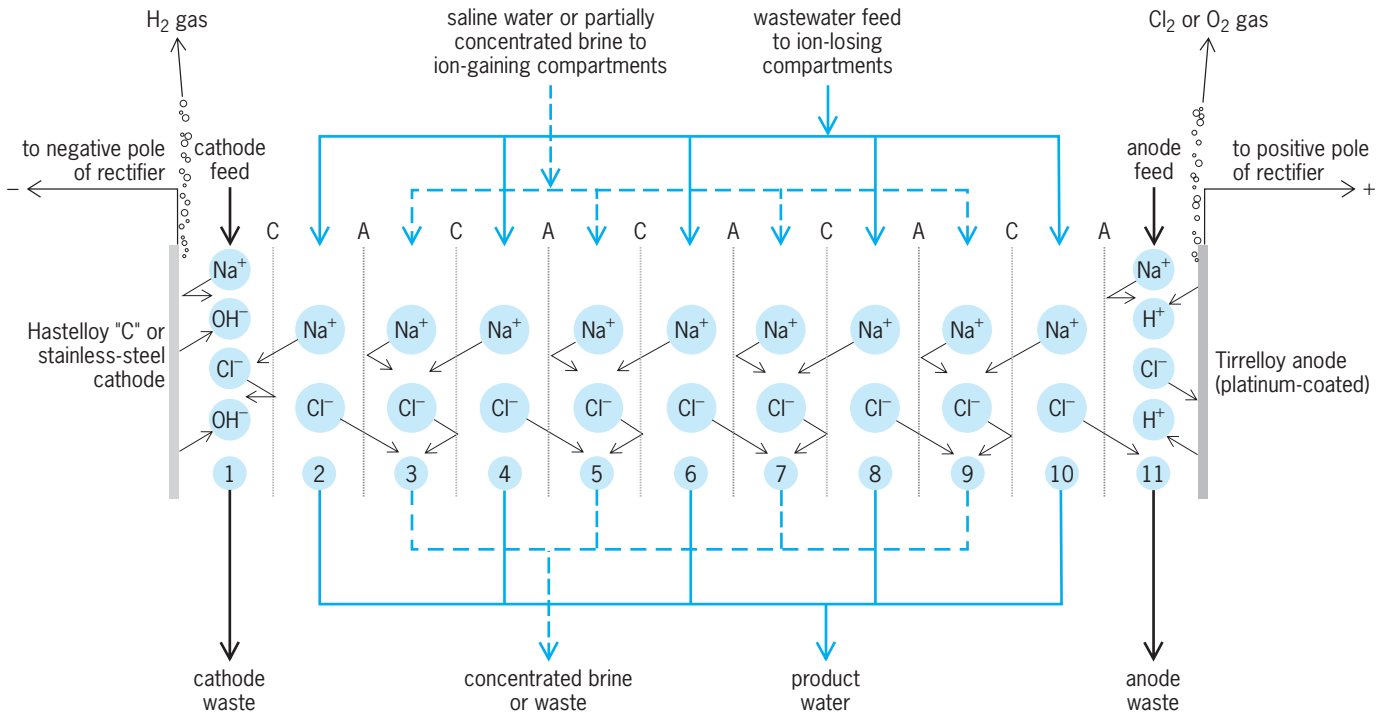

Fig. 8. Diagram of the basic ion and water flow in electric membrane stack. C = cation membrane; A = anion membrane; Na^+ = any cation, such as sodium; Cl^- = any anion, such as chloride. Numbers = compartments.

washed out with water. The metal is in the form of bright crystalline flakes. A beryllium-copper eutectic can be made by using a copper cathode. Beryllium alloys containing copper and nickel are made from BeO in arc furnaces.

Calcium has been made by electrolysis of pure fused calcium chloride about 800°C (1470°F). In this case the cathode is solid calcium, and is mechanically withdrawn from the cell as a "carrot." Calcium is now made by aluminum reduction at low pressures and high temperatures.

Cerium and misch metal are made from $CeCl_3$ or mixtures of chlorides of cerium, lanthanum, and neodymium in fused-salt electrolysis with NaCl. Misch metal is used for lighter flints.

Fluorine for separation of uranium isotopes is produced by electrolysis of 40% HF in KF between carbon anodes and steel cathodes at 88–100°C (190–212°F). A diaphragm of Monel screen keeps the products H_2 and F_2 separated. Dry HF gas bubbled continuously into the electrolyte. At a current density of 1 A/in.^{2} (15 A/dm^2) the cell operates at 9–12 V and 96% current efficiency. Energy consumption is 3.0 kWh/lb (24 MJ/kg) fluorine. The theoretical decomposition potential is 2.85 V.

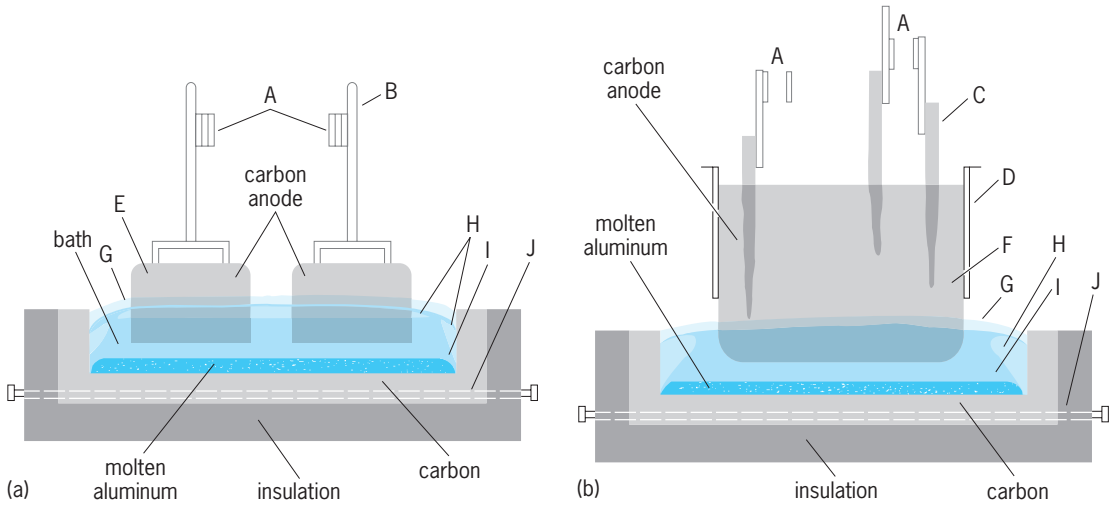

Fig. 9. Two types of aluminum cells, (*a*) utilizing prebaked carbon anode and (*b*) utilizing Söderberg carbon anode. A = anode bus, B = anode rod, C = anode stub, D = anode casing, E = prebaked carbon anode, F = Söderberg carbon anode baked in place, G = crust of frozen bath and alumina, H = frozen bath, I = bath (electrolyte), J = steel pin cathode collector.

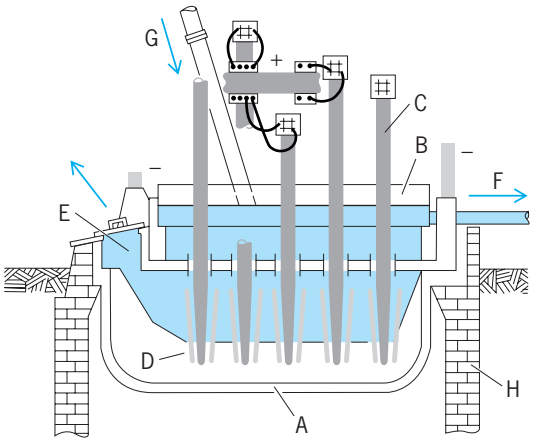

Fig. 10. Cross section of the Dow magnesium cell. A = steel container, B = ceramic cover, C = graphite anodes, D = steel cathodes, E = magnesium collecting well, F = chlorine outlet, G = magnesium chloride feed, H = furnace setting.

Lithium is made by electrolysis of fused 60% LiCl and 40% KCl at 450–500°C (840–930°F) in a cell similar to the Downs sodium cell.

Magnesium is produced by electrolysis of fused 25% $MgCl_2$ and 75% NaCl at around 700°C (1290°F). The Dow process for making magnesium from seawater uses material approximating $MgCl_2 \cdot 2H_2O$, which is fed around the graphite anodes, where dehydration occurs (**Fig. 10**). Gas from the anode compartment is wet chlorine, air, and hydrogen chloride; the latter is used to make fresh magnesium chloride from magnesium hydroxide. Magnesium metal is deposited on steel cathodes, which direct the metal to a collecting zone. The cell is a cast-steel pot in a furnace setting. Other cells uses molten anhydrous magnesium chloride feed. They have brick-lined steel bodies with graphite anodes (**Fig. 11**). Magnesium chloride, with other chlorides, is separated in vacuum crystallizers from brines, dehydrated, and melted in electric resistance cells from which molten $MgCl_2$ is tapped periodically to feed the cells. Molten magnesium is ladled from the cells

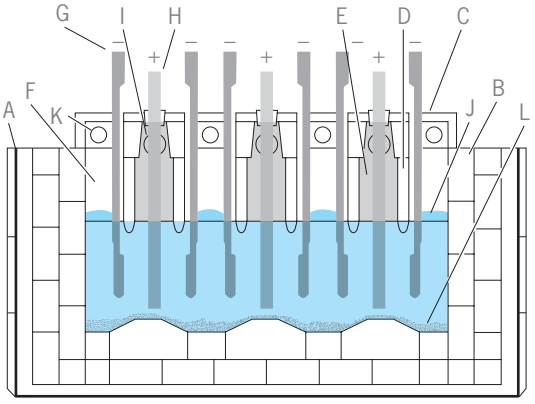

Fig. 11. Cross section of anhydrous magnesium chloride cell. A = steel box, B = ceramic lining, C = ceramic cover, D = ceramic separators, E = anode (chlorine) compartment, F = cathode (metal) compartment, G = iron cathode, H = graphite anode, I = chlorine outlet, J = magnesium, K = cathode compartment vent, L = mud.

and cast into molds. The cells operates at 6–7 V and 80–88% current efficiency and use 8–8.5 kWh/lb (63–67 MJ/kg) metal.

Sodium was once made by electrolysis of fused NaOH, but since 1929 it has been made by electrolysis of NaCl in the Downs cell. The electrolyte is 40% NaCl and 60% $CaCl_2$ at 590°C (1090°F). The cell consists of a brick-lined steel vessels. Four graphite anodes project upward from the bottom. the cathode is made of steel cylinders concentric with the anodes and supported from iron arms extending through the sides of the cell, which also conduct current. A diaphragm of 26-mesh-per-inch (10 mesh-per-centimeter) iron screen directs the sodium into an inverted trough leading to a riser pipe, which cools the metal and conducts it to a collecting tank beside the cell. Chlorine is collected in a nickel cone inverted overthe anode. Pure dry salt is fed to the cell. The reactions in the cell are Eqs. (13) and (14). The

$$2NaCl + 2e \longrightarrow 2Na + Cl_2 \tag{13}$$

$$2Na + CaCl_2 \rightleftharpoons Ca + 2NaCl \tag{14}$$

metal at cell temperature is 5% Ca, but as it cools, the second equation is reversed so that the cool metal is about 1% Ca. It is filtered just above the melting point of the sodium, and the final product is under 0.04% Ca. Cells of 38,000 amp operate at 7 V and 83% current efficiency. Energy consumption is about 4 kWh/lb (32 MJ/kg) metal.

Molybdenum, thorium, titanium, uranium, and zinconium can all be made by electrolysis of their complex halides, K_3MoCl_6, $ThF_4 \cdot KF$, K_2TiF_6, KUF_5, and K_2ZrF_6, in molten NaCl. Inert atmospheres are required, and the cell is usually a graphite crucible acting as anode with a graphite or molybdenum cathode, on which the metal deposits as crystals or powder. The batch of metal on the cathode is cooled in an inert atmosphere, then broken off, pulverized, and leached with water. The metal powder is used as such or melted in a vacuum arc furnace.

In the case of tantalum, pure K_2TaF_7 is heated to 900°C (1650°F) in a graphite pot acting as an anode witha removable metal cathode. When the cathode is loaded with deposited metal, it is removed and quickly replaced. The bath is replenished with K_2TaF_7. The cathode deposit is pulverized and washed with acid. The metal powder can be compacted by sintering.

Electrothermics. The manufacture of many products requires temperatures higher than can be obtained by combustion methods. Electric heat can usually be developed at, or close to, the point where it is required, so that it is relatively quick. It permits easy control of the atmosphere for oxidizing, reducing, or neutral conditions.

Products of the electric furnace include iron and steel; ferrallous; nonferrous metals and alloys; the exotic metals titanium, zircomium, hafnium, thorium, and uranium; and nonmetallic products such as calcium carbide, calcium cyanamide, sodium cyanide, silicon carbide, boron carbide, graphite,

fused alumina, magnesium oxide, quartz, silica, thoria, zirconia, lime, spinel, kyanite, sodium aluminate, dolomite, boric acid and borides, carbides of zirconium and titanium and related metals, graphite, phosphorus and phosphoric acid, and chlorides of magnesium, boron, zirconium, and titanium. An electric smelting process converts ilmenite into iron and a titanium slag for a pigment manufacture. Zinc metallurgy uses an electric furnace. Steam is generated in electric boilers where economically feasible.

Large amounts of United States steels, stainless steels, and special alloys are made in electric furnaces. **Table 4** gives values for electrode consumption, furnace voltage, secondary amperage (from-transformer to electrodes), and current density.

Methods of electric heating utilize resistance, arcs, or induction. Resistance furnaces (**Fig. 12**) may use the substance being heated as a resist or, or auxiliary resistors may be used. Arc furnaces (**Fig. 13**) may have a direct arc between an electrode and the material being heated, for example, steel scrap or the arc may be between two or more electrodes. The arc may be indirect or it may be submerged in the material being melted. Induction furnaces (**Fig. 14**) use the crucible or its charge as the closed secondary circuit of a transformer with low- or high-frequency alternating current. The temperature in the carbon arc is approximately 3100°C (5600°F). Graphite or amorphous carbon are used for electrodes, depending on economics. Metals such as titanium, zirconium, and halfnium may be compacted into electrodes which are consumed by resistance or arc melting or combinations of both in a nonreactive-atmosphere or vacuum furnace. Some types of furnaces are adaptable to several products, but usually the design is developed for an individual product. *See* ELECTRIC FURNACE; ELECTRIC HEATING; PYROMETALLURGY; STEEL MANUFACTURE.

Zone refining of metals for the electronics industry, such as silicon for diodes and transistors, is accomplished by induction melting of the metal in a narrow zone and slow movement of the molten in the metal ingot from one end to the other in an evacuated or inert gas-filled enclosure. Impurities move toward the end of the ingot. The operation is repeated until the desired purity is obtained. *See* ZONE REFINING.

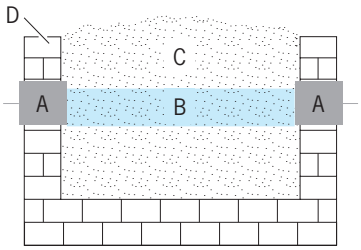

Fig. 12. Resistance furnace. A = current conductors, B = conducting core of granulated carbon, C = granular charge, D = furnace wall of brick and refractory.

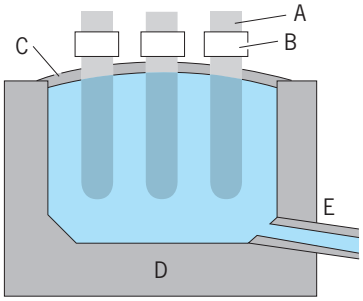

Fig. 13. Three-phase arc furnace. A = carbon or graphite electrodes, B = current clamps, C = refractory roof, D = refractory shell, E = tap hole.

High-melting point metals, carbides, oxides, and nitrides are melted with a plasma-arc torch by coating objects with molten droplets carried in a jet of inert gas passed through the torch. Raw material, such as a powder or wire, is fed through a dc arc between a tungsten cathode and water-cooled copper anode. A strong bond with the base material is obtained. *See* TORCH.

Processes in gases. Electrical discharge through gases has industrial application in ozone production and nitrogen fixation. Ozonizers consist of two metal electrodes with an air gap and a dielectric, such as gas, between them. Very dry air passed through an air gap will then contain 10–12 mg ozone per liter.

TABLE 4. Representative values for electrode consumption, voltage, current, and current density for electric furnace products

Product	Electrode consumption* Approx. lb/ton of product	Electrode consumption* Approx. lb/1000 kWh	Voltage range, phase to phase	Current range, secondary amperes	Current density, A/in.^{2}
Ferrosilicon, 50%	30–40	7	130–195	35,000–55,000	35–55
Ferrosilicon, 65%	60–90	10	125–170	35,000–50,000	35–60
Ferrosilicon, 75%	90–130	12	120–170	35,000–50,000	35–60
High-carbon ferrochrome	30–50	10	110–200	25,000–40,000	25–40
Silicomanganese	80–100	23	110–140	35,000–70,000	35–70
Ferromanganese	30–50	16	85–100	30,000–60,000	35–65
Calcium carbide	40–60	17	120–200	30,000–80,000	40–70
Phosphorus	35–70	3	250–400	13,000–30,000	20–30
Refined copper	4–6	19	80–190	12,000–30,000	80–110
Nonferrous castings	4–5	15	100–110	800–5000	110–300

*1 lb/ton = 0.5 kg/metric ton. 1 lb/k Wh = 0.1260 kg/MJ.

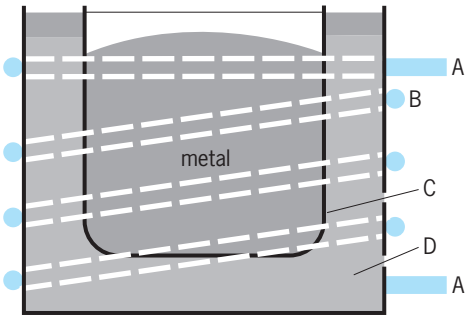

Fig. 14. Induction furnace. A = terminals of high frequency, B = copper coil, C = crucible, D = refractory.

Fixation of nitrogen by passing air through an arc furnace, thus forming oxides of nitrogen, was once practiced when power was cheap in Norway, France, and Italy, but has been replaced by conventional processes. *See* ELECTRICAL CONDUCTION IN GASES; NITROGEN FIXATION; OZONE.

Electromagnetic separation. Magnetic separation removes tramp iron from mixtures of granular solids, suspensions, and solutions. Magnetic separation is also used to separate solids of various magnetic susceptibilities, such as mineral fractions. *See* MECHANICAL SEPARATION TECHNIQUES.

Electrodialysis. This is the separation of low-molecular-weight electrolytes from aqueous solutions by migration of the electrolyte through semipermeable membranes in an electric field. It is used on an industrial scale for deashing starch hydrolyzates and whey, and in many municipalities for producing potable water from saline water. Its uses also include the concentration of liquid foods such as dairy products and citrus juices, the recovery of sulfite pulp waste and pickling acid, and the isolation of proteins. *See* COLLOID; DIALYSIS.

Electrophoretic deposition. This is the deposition of a nonconductive material in a finely divided state from a suspension in an inert medium. Electrophoresis is the migration of colloidal particles, which acquire positive or negative charges in an electric field. The process is useful in electropainting; for instance, electropainting of automobile bodies and other objects has now been adopted on a large scale. Rubber latex is an example of negatively charged colloid which can be plated on an anode. Electronic components can be coated with inorganic salts, oxides, and ceramics suspended in organic media. Bitumen can be electrodeposited out of an aqueous dispersion onto the anodic surface of steel pipe by use of an axially placed cathode. *See* ELECTROPHORESIS.

Electroendosmosis. This is the movement of a liquid with respect to an immobilized colloid in an electric field. The process is used in the dehydration of peat, dye pastes, and clay. Dies in a clay extrusion press can be "lubricated" by making the die cathodic which attracts a film of water to it from the wet clay. It is also used commercially for dewatering soils in mining, road building construction, and other civil engineering works.

Electrostatic technique. The deposition of charged particles from suspension in gases has many useful applications. The Cottrell electrostatic precipitator removes dusts and mists from gases. In one form, a fine wire is axial in a pipe and insulated from it. The pipe is grounded, and the wire is negative in a high-voltage dc circuit. Particles in the gas that is passing through the pipe become electrically charged and move to the pipe. Liquid particles drain off, while solids are periodically vibrated off. *See* ELECTROSTATIC PRECIPITATOR.

Spray painting with a high voltage between the gun and the work is particularly effective in providing an even coating with an economical use of paint or irregular and open surfaces, such as a screen. Complete coverage is achieved because an uncovered spot or pore has a higher field strength and attracts the charged droplets to it.

In xerography a sheet of plain paper is electrically sensitized in those areas corresponding to an original so that colored resin particles carrying an opposite charge are attracted and retained only on the sensitized areas, thus producing a visible image corresponding to the original.

Abrasive paper and cloth are coated with an adhesive and abrasive powers attracted to the base material in an electrostatic field. Pile fabrics can be produced in a similar manner, with the short fibers oriented by the electric field. Powdered adhesive can be applied to paper.

Plant design for electro process. The greatest changes have taken place in the electrical distribution system portion of electrolytic plants. Copper bus-bars have taken for granted for years. They show good corrosion resistance, low contact voltages, and contacts can be kept clean readily. Aluminum is not as corrosion-resistant, and contacts are poor unless specifically designed. Aluminum and copper are competitive conductors when the ratio of the cost of a given amount of aluminum to that of copper is about 1.7. Aluminum is below the equality price, copper is above. Therefore, electrolytic plants cannot afford the luxury of copper bus-bars, but still want the advantages of copper contacts—namely, the copper-surface aluminum bus-bar and the copper-shoe aluminum crossbar, achieved by plating, machined sections, explosive forming, and duplex preparation.

Electrolytic engineering encompasses cell design; electrical isolation; use of sumps instead of sewers; economic selection of bus-bars; solution circuit isolation; under-floor distribution of lines, supports, conductors, crystallizing liquors, and plastic pipe and conduits; lined and unlined cells; contacts and their inspection; and avoidance of pollution.

Early in plant development, the problems of impurities in solutions and the effects of traces of the electrolytic deposit were evident, appreciated, and solved, largely by Edisonian methods. Maximum tolerance were served up, and purification procedures were developed to a high level in copper and zinc. Crystallizing solutions were a problem for manganese. In this area, control of deposition potential,

voltage-amperage and amperage-time recording instruments gave the concept of impurity concentrations affecting the deposition potential and the current density needed to reach this potential. There was an explanation for receding or resolution of deposits. In chromium the problem was solved, in effect, by converting the ore during purification into a chemical compound purified by crystallization. Purification of electrowinning solutions of magnesium, sodium, potassium, barium, calcium, iron, aluminum, antimony, arsenic, tin, cadmium, silica, lead nickel, and cobalt by chemical methods is a common occurrence.

The reconditioning of cathodes for metal deposition and ease of product removal in the electrowinning of manganese and chromium is a repetitive, demanding, and skilled-labor operation. Lines have been designed and put in operation resembling the continuous, conditioning cleaning, pickling, and preparation lines of the electroplating phases of the automotive industry, a long-delayed cross-fertilization. When electrolytic copper sheet needed separate plants for printed electronic circuits, auxiliary electrodes were developed to control edges and ensure more uniform thickness. Perhaps manufacture of starting sheets can be mechanized by adopting continuous electroplating lines.

The concept of electrolytic cell as an insulated mechanism, supported on structures and insulated from them, with separate support for conductors, supply lines, circulating systems, all insulated and electrically broken so that they are not shunt circuits, has been well established for decades. Tanks were wood, but are now unreinforced concrete, lined or unlined, and in few cases glass fiber or cloth-reinforced plastics, such as poly(vinyl chloride) or its competitors. In general the copper industries studies have shown that the investment cost for a lead-lined concrete cell is a little more than half that of a ribbed, equivalent-strength plastic. Because of local availability, tin refining cells are poly(vinyl chloride)-lined concrete units, but chromium metal is made in plastic cells. A few foreign plants employ "painted" or thin-lined concrete.

Selenium rectifiers are used, to a greater extent than copper oxide, for small power supplies, control and signaling circuits, and for applications in which their relative immunity to voltage surges is desired.

There have been no mercury-arc rectifiers built in the United States since about 1960, and very few in Europe for electrolytic or industrial applications, except for high-voltage direct current, welding, and transmission. Monocrystalline diode rectifiers, first germanium and then silicon, have completely replaced mercury-arc rectifiers in industrial applications. Manufacturers have converted to silicon in many multianode mercury-arc rectifiers and excitron and ignition units. All new installations and expansions utilize silicon rectifiers.

The only type of power rectifiers for electrolytic loads, such as aluminum reduction, metal refining, electrowinning, chlorine and chlorate cells, and so on, are silicon rectifiers. Silicon rectifiers range in sizes up to 90,000 A and up to nearly 1000 V, with a maximum rating of 30,000 kVA.

Organic Processes

Organic electrochemistry was once regarded as a tantalizing area with many important laboratory achievements but few successes in commercial practice. This situation has changed, however, in that electroorganic processes are commercially advantageous if they can fulfill either of two conditions: (1) performance under conditions of voltage corresponding thermodynamically to the conversion of an organic group to a reduced or oxidized group, with the cell products relatively easy to isolate and purify; (2) performance of a highly selective, specific technique to make an addition at a double bond, or to split a particular bond (for example, between carbon atoms 17 and 18 of a complex molecule having 25 carbon atoms).

Selectivity and specificity are highly important in electroorganic processes for the manufacture of complicated molecules of vitamins and hormones—as well as for the medicinal products whose action on pathogenic organisms is a function of their spatial arrangement, steric forms, and resonance.

The electrolytic approach can also be competitive for some low-cost, tonnage products. Here continuous processing is important, and only a single phase should be present, that is, a solution rather than an emulsion, dispersion, or mechanical mixture. Only for fairly valuable products is it practical to find a conducting solvent and then to engineer around it.

The electrolytic oxidation and reduction of organic compounds differ from the corresponding and more familiar inorganic reactions only in that organic reactions tend to be complex and have low yields. The electrochemical principles are precisely those of inorganic reactions, while the procedures for handling the chemicals are precisely those of organic chemistry.

Most organic molecules are insoluble in the aqueous solutions that are the best electrical conductors. Unless solubility can be increased, the only other approach is to use organic solvents. These make relatively poor conductors; hence are encountered power loss, heat build-up, chemical inversion, and often, stepwise, complex reactions.

Oxidations. Commercial success in inorganic electrochemistry has come about by well-engineered combinations of organic and inorganic techniques in areas where strictly chemical methods are either impossible or inefficient, for example, in catalytic hydrogenation or oxidation.

The conventional oxidation reagents of the organic chemist are expensive. There is no market for the oxidant once it is reduced, and chemical regeneration is prohibitive in cost. Accordingly, these reagents are avoided. Electrolytic regeneration, however, can be relatively inexpensive if linked to a carefully controlled organic operation. Typical of this approach is manganese dioxide oxidation of anthraquinone with electrolytic regeneration of the oxidant. In the electrolytic oxidation of anthracene

there is a cost-efficient process that utilizes a 20% sulfuric acid suspension with a small quantity of ceric sulfate as a catalyst. Other examples include chromic acid oxidation of oleic acid to perlargonic and azelaic acids, where the oxidant is regenerated electrolytically, and the electrolytic regeneration of periodic acid (a costly oxidizing reagent) in the dialdehyde starch process. These involve savings not only in the purchase price of the oxidants but also in the disposal cost of products that cannot be marketed.

In electrolytic reactions the cell surface is the only source of reductants or oxidants, and these must be produced at highest efficiencies. The mass action effects, concentrations, temperature of reaction, reaction velocities, diffusion, and equilibria between the initial and final products of the reaction all apply to electrochemical reduction and oxidation in the same manner as do the corresponding reactions carried on outside the electrolytic cell.

Materials that are reduced absorb hydrogen at the cathode, and may be considered cathodic depolarizers; those that are oxidized absorb oxygen at the anode, and are anodic depolarizers. Oxidation reactions may involve substances other than oxygen, such as chlorine.

Anodes are selected with a high oxygen or halogen overvoltage, and cathodes are selected with a high hydrogen overvoltage. (Because the accumulation of electrolysis products at anode and cathode causes polarization, the overvoltages are needed to move the products away and keep the process going.)

Reductions. Substances that are easy to reduce may be acted on, at the interface of cathodes, with low hydrogen overvoltage. Hard-to-reduce materials may require much higher overvoltages, which are reached through either the cathode composition or the current density.

The aromatic nitrogen-containing organic compounds have been extensively studied by many researchers. As early as 1900, F. Haber and K. Elbs showed that the nitro compounds of the type RNO_2, where R is an organic radical, could either be directly reduced to the amine RNH_2 (that is, aniline), or successively reduced to the nitroso product RNO, the beta aryl hydroxylamine RNHOH, the azoxy product RN—O—NR, and the material containing an azo group RN═NR, which in turn is reduced to the hydrazoform RNH═NHR or the idene type as $H_2NR{=}RNH_2$.

Although commercial processes based on these reactions were successful, they were eventually replaced by improved nonelectrolytic processes.

Oxidation of starch. The original process for the oxidation of starch to dialdehyde starch was by periodic acid, continuously regenerated at the anode of an electrochemical cell. Dialdehyde starch is an important polymeric aldehyde used in paper manufacture and leather tanning. In this process a lead anode, coated with lead dioxide, is immersed in a suspension of starch; the cathodes are steel, enclosed in porous aluminum oxide diaphragms. One difficulty with this process is that free iodine, formed by the migration of iodine ions to the anode, tends to react with starch, despite the presence of the diaphragm.

In 1960 an improved two-stage process was devised, where just enough periodic acid is added to oxidize the starch in a straightforward chemical reaction. The iodic acid produced during the reaction is then separated from the dialdehyde starch and sent to an electrolytic cell for reconversion to periodic acid. In the electrolytic cell, this reconversion is achieved with current efficiencies of 80–90%. Materials of construction are structural plastics that are unaffected by the reagents or reactants; only the electrodes are metal. Continuous operation is achieved, with the instrumentation permitting good control of temperature, concentration, conversions per pass, gas venting or collection, external recirculation rates, and other rate factors. The wide range over which such control can be exercised allows the cell to function like a reactor or catalytic converter. In fact, the cell has many of the characteristics of the catalytic reactors, the main difference being that in the cell the variables are related to the type and surface of the electrodes, while in the catalytic reactor they are related to the surface of the catalyst and its support.

All piping is manifolded exterior to the cell (unlike the hydrogen-oxygen cell), and the feeds are valve controlled externally. The electrode spacing is prearranged, and all reaction chambers are in parallel both relative to flow and electrically. Because valve closure can shut off a fraction of the parallel streams through the cell, shut downs because of mechanical or other failures are minimized. The electrolytes are simple solutions, not emulsions or suspensions, and do not contain stabilizers or salting-in components. Gas evolution, separation, and collection are separately handled in anolyte and catholyte streams. These two streams may be blended in any proportion desired for maintenance of pH and control of concentration.

Adiponitrile production. In 1965, an electroorganic process for adiponitrile production was developed and commercialized. Power was used as a reagent in the electrolytic reductive coupling (or hydrodimerization) of acrylonitrile to adiponitrile, a component of nylon-6,6. The adiponitrile molecule is a six-carbon dimer, essentially made up of two molecules of three-carbon acrylonitrile.

In the commercial plant, each cell consists of a lead cathode and an alloy anode separated by a sulfonated polystyrene resin membrane. Spacers support the membrane and prevent it from touching the electrodes. The cells are made of polypropylene with neoprene gaskets and are grouped in banks of 24. The reactants, the catholyte and anolyte, circulate separately in an upflow direction through a bank of cells in parallel; the direct current passes through the bank in series. The catholyte contains a quanternary ammonium salt, that is, a McKee salt. This provides ions for electrical conductivity and increases the acrylonitrile's solubility above 10%; lesser concentrations would lead to the formation of proprionitrile.

The catholyte is circulated to remove the film of product at the cathode surface.

Lead alkyl production. In one electroorganic process for the electrolytic production of lead alkyls, electrochemical synthesis of tetramethyllead or tetraethyllead is achieved by employing a Grignard reagent and lead shot. The electrolyte consists of alkyl magnesium halides, which ionize in ether solutions. The solution is fed to 8000-gallon (30,280-liter) electrolytic cells, with lead pellet anodes, the walls of the cell as cathode, and a nonconducting low-permeability diaphragm. The reaction is shown in reaction (15), where R represents the methyl or ethyl groups.

$$4R^- + Pb - 4e \longrightarrow R_4Pb \qquad (15)$$

The $MgCl^+$ ions migrate to the cathode cell walls. Magnesium chloride and metallic magnesium are formed according to reaction (16).

$$4MgCl^+ + 4e \longrightarrow 2Mg + 2MgCl_2 \qquad (16)$$

In this process a variable-voltage approach is used. As the rate of reaction drops and the resistance of the cell increases, more and more voltage is applied to the cell, thus maintaining a constant rate of reaction per unit of time.

Mixed alkyl lead compounds may be made by electrolyzing ethylmagnesium chloride and adding methyl chloride in the cells; or, conversely, by starting with methylmagnesium chloride and then adding ethyl chloride.

The electrolysis cell bank is made up of 10 cells, each with an associated recirculation surge drum. In the cell the copper bus handles the low-voltage, high-amperage current loads. The lead pellet storage hopper and tubular conveyor equipment are located directly above the cell dome. The upper third of the recirculating drum is an insulated refrigerant-storage vessel that supplies coolant to carry off heat developed during the course of electrolysis. Refrigerant vapors are returned to centrifugal compressors. Charles L. Mantell

Bibliography. A. Bard and L. R. Faulkner, *Electrochemical Methods, Fundamentals and Applications*, 2d ed., 2000; D. R. Crow, *Principles and Applications of Electrochemistry*, 4th ed., 1994; A. Kuhn (ed.), *Industrial Electrochemical Processes*, 1971; H. Lund and M. M. Baizier (eds.), *Organic Electrochemistry: An Introduction and a Guide*, 3d ed., 1991; D. Pletcher and F. C. Walsh, *Industrial Electrochemistry*, 2d ed., 1993; K. Scott, *Electrochemical Reaction Engineering*, 1991.

Electrochemical series

A series in which the metals are listed in the order of their chemical reactivity, the most active at the top and the less reactive or more "noble" metals at the bottom. In a broader sense such an activity series need not be limited to the metals but may be carried on through the electronegative (nonmetallic) elements as well. See the **table** for a list of common elements.

Electrochemical series of the elements*

Element	Symbol	Element	Symbol
Lithium	Li	Zinc	Zn
Potassium	K	Chromium	Cr
Rubidium	Rb	Gallium	Ga
Cesium	Cs	Iron	Fe
Radium	Ra	Cadmium	Cd
Barium	Ba	Indium	In
Strontium	Sr	Thallium	Tl
Calcium	Ca	Cobalt	Co
Sodium	Na	Nickel	Ni
Lanthanum	La	Molybdenum	Mo
Cerium	Ce	Tin	Sn
Magnesium	Mg	Lead	Pb
Scandium	Sc	Germanium	Ge
Plutonium	Pu	Tungsten	W
Thorium	Be	**Hydrogen**	**H**
Beryllium	Th		
Uranium	U	Copper	Cu
Hafnium	Hf	Mercury	Hg
Aluminum	Al	Silver	Ag
Titanium	Ti	Gold	Au
Zirconium	Zr	Rhodium	Rh
Manganese	Mn	Platinum	Pt
Vanadium	V	Palladium	Pd
Niobium	Nb	Bromine	Br
Boron	B	Chlorine	Cl
Silicon	Si	Oxygen	O
Tantalum	Ta	Fluorine	F

*According to standard oxidation potentials E° at 25° C or 77° F.

The electrochemical series as it applies to metals was first established by laboratory experiments in which the purpose was to determine which metals would displace others from solutions of their salts. Thus a clean strip of zinc immersed in a solution of copper sulfate is soon found to be covered by a deposit of copper, while zinc in turn goes into solution from the strip as zinc ions. By definition, then, zinc is a more reactive metal than copper, since it will displace copper from a solution of Cu^{2+} ions. The reaction is readily seen to be an oxidation-reduction transfer of electrons, which can be summarized by the equation below. Similarly, copper will displace

$$Cu^{2+} + Zn = Cu + Zn^{2+}$$

silver from a solution containing Ag^+ ions, depositing crystals of metallic silver and coloring the solution with Cu^{2+} ions. From these observations an activity series may be set up in the order Zn, Cu, Ag. By exhaustive experiments with other metals it becomes possible to draw up a complete list in the order of chemical activity, in which the metals at the top of the list are those which are found to give up their electrons most readily (that is, are the most electropositive elements). Such a list is shown in the table, where lithium exhibits the most reactivity as a metal.

The ease with which an isolated atom of an element gives up an electron, known as the first ionization potential, is a precise physical quantity which can be measured by electrical experiments

on gases or vapors at low pressure. The replacement experiments which determine the order of the electrochemical series take place in a very different environment, since they involve solid phases and also aqueous solutions with their consequent hydration effects. Moreover, it might well be expected that displacement reactions in solution would depend upon the concentrations of the reagents used, and also upon the presence or absence of other dissolved substances. *See* IONIZATION POTENTIAL.

To obtain a more accurate and reproducible activity series, it is best to turn to the more exact quantity called electrode potential, or oxidation-reduction potential, which is defined as the voltage developed by a sample of pure metal immersed in a solution of one of its salts (at unit activity and at 25°C or 77°F) versus a hydrogen electrode immersed in hydrochloric or sulfuric acid of equivalent concentration. For details about this measurement *see* ELECTRODE POTENTIAL

It is evident that by confining the experimental conditions to a standard concentration and temperature the hydration and concentration effects can be kept quite constant, making possible a more exact listing of metals according to their activity. Hence any present-day electrochemical series must rely on the measurements of oxidation potential and should be in agreement with the accepted values determined from such electrochemical cells. Such reliance has the further advantage that the series need not then be confined to metals but may be extended to the nonmetals, or electronegative elements. As before, those metals which will liberate hydrogen from dilute acids (such as hydrochloric or sulfuric) will stand above hydrogen in the series, while those metals and nonmetallic elements which will not liberate hydrogen from such dilute acids will stand below hydrogen in the list. Since the oxidation potentials are also related to the equilibrium constants for reversible reactions, it becomes possible to calculate oxidation potentials from other information when direct experiments are inconvenient, as in the case of the alkali metals versus aqueous solutions of their salts. *See* ELECTROCHEMISTRY; ELECTRONEGATIVITY; OXIDATION-REDUCTION. Eugene G. Rochow

Bibliography. F. A. Cotton, G. Wilkinson, and P. L. Gaus, *Basic Inorganic Chemistry*, 3d ed., 1994; C. A. Hampel (ed.), *Encyclopedia of Electrochemistry*, 1964, reprint 1972.

Electrochemical techniques

Experimental methods developed to study the physical and chemical phenomena associated with electron transfer at the interface of an electrode and solution. The objective is to obtain either analytical or fundamental information regarding electroactive species in solution. Fundamental electrode characteristics may be investigated also.

The physical and chemical phenomena important in electrode processes generally occur very close to the electrode surface (usually within a few micrometers). Mass transfer of species involved in an electrode process to and from the bulk of solution is one important aspect. Inclusion of a large excess of inert electrolyte in most electrochemical systems eliminates electrical migration (current flow as a result of movement of ions of the substrate toward or away from the electrode) as an important means of mass transfer for electroactive species, and only convection and diffusion are considered.

Important chemical aspects of electrode processes include the oxidation or reduction occurring as a result of electron transfer, and coupled chemical reactions. Coupled reactions are initiated by production or depletion of the primary products or reactants at the electrode surface. Identification of the nature and mechanism of such coupled reactions is of particular importance in studies of electrode reactions of organic compounds, where multireaction and multielectron cascades are often found to be initiated by electron transfer to or between the electrode and the electroactive substance.

The primary experimental variables involved in electrochemical techniques are the potential E, the current I, and the time t. The potential or current at the working electrode is controlled and the other is observed as a function of time. The many ways in which either may be controlled give rise to the wide variety of controlled-potential or controlled-current techniques. In all such techniques it is necessary to specify whether mass transport of the electroactive species to the electrode is by convection or diffusion, since mathematical treatments of these two processes are quite different. Mass transport by convection is more efficient than diffusion by several orders of magnitude but is much more difficult to model mathematically.

The general scheme for electron transfer at an electrode in solution is shown in reaction (1), where O

$$O + ne^- \underset{k_b}{\overset{k_f}{\rightleftharpoons}} R \tag{1}$$

and R are the oxidized and reduced forms of the electroactive species and n is the number of electrons transferred. When k_f and k_b, the rate constants for the forward and back reaction, are very fast and O and R are not involved in preceding or following chemical reactions, the system is called reversible and the Nernst equation (2) holds. In this relation,

$$E = E^{\circ\prime} + \frac{0.059}{n} \log \frac{C_0}{C_R} \tag{2}$$

E is the electrode potential, $E^{\circ\prime}$ is the formal standard potential for the redox couple, and C_O and C_R are concentrations at the electrode surface. In the following discussions, only reversible reduction processes will be considered, although oxidations are equally applicable. *See* ELECTROCHEMICAL PROCESS; ELECTRODE; ELECTRODE POTENTIAL; ELECTROLYSIS; OXIDATION-REDUCTION.

Controlled potential. A variety of methods of this type have been developed, depending on whether the electrode potential is held constant or varied

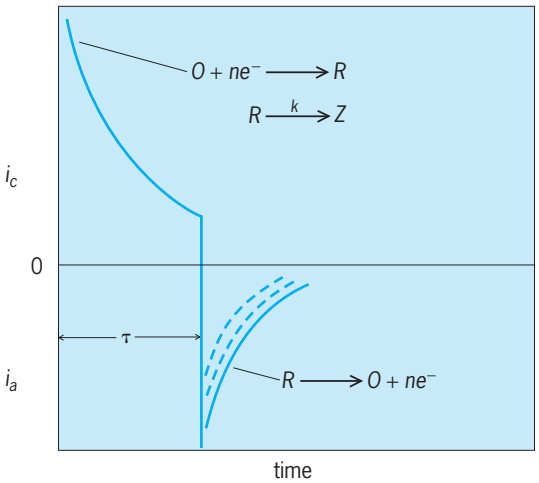

Fig. 1. Current–time behavior for double-potential-step experiment at a stationary working electrode. Broken lines in the reverse current-time curve represent cases where the rate of chemical reaction is fast compared to the time scale of the experiment. $i_a/i_c = f(k)$.

during the experiment and on whether mass transport is by convection (stirring) or diffusion.

Constant potential with convection. In this technique, known as controlled-potential electrolysis, the electrode potential is held constant as the solution is stirred or the electrode is rotated at a constant rate. The current is controlled by the concentration of electroactive substance and the stirring rate. Complete conversion of the electroactive substance takes place at a first-order rate whose constant is determined by cell geometry and stirring efficiency. With good cell geometry and efficient stirring, electrolysis is complete in a manner of minutes.

Constant potential with diffusion (chronoamperometry). When a reducing potential is imposed instantaneously on a stationary working electrode in quiescent solution, current will rise sharply and then decay as the electroactive species near the electrode is depleted by electrolysis. The magnitude of the current is proportional to the bulk concentration of electroactive species, and if the imposed potential is sufficiently negative of $E^{\circ\prime}$, the Nernst equation demands complete conversion to the reduced form. Under these conditions, the current for this chronoamperometric, or potential-step, experiment is diffusion-controlled and decays with $1/t^{1/2}$. This is shown by Eq. (3), where F is the faraday, A is the electrode

$$I = nFAC^* \left(\frac{D}{\pi t}\right)^{-1/2} \tag{3}$$

area, D is the diffusion coefficient for the electroactive species, and C^* is the bulk concentration of electroactive species.

An important variant on this experiment is double-potential-step chronoamperometry, in which a second potential step is applied. During the initial step, electrolysis occurs, depleting the oxidized form O, but producing the reduced form R in the immediate vicinity of the electrode. If the potential is instantaneously switched back to the initial value after a time τ, species R will be reoxidized, and an inverted I-t curve will be obtained (**Fig. 1**). This figure also illustrates the common convention that reduction and oxidation currents are plotted in the positive (upward) and negative (downward) directions, respectively. The value of this technique for studying coupled chemical reactions is illustrated by the broken lines in Fig. 1, which demonstrate that while coupled processes have no effect upon the first step, the current after the second step is diminished by a following chemical reaction (decomposition of R to Z). The ratio of forward and back currents, measured at $t = \tau$ and $t = 2\tau$, respectively, provides a convenient experimental measure of this additional chemical complication. The degree to which the current is diminished by coupled reactions can be used to obtain both rate constants and the molecularity of such reactions. A wide dynamic range of rate constants can be measured by varying τ. *See* ELECTROCHEMISTRY; ELECTROLYSIS.

Variable potential. Several procedures of this type have been developed.

Linear sweep voltammetry (LSV). In this technique, the potential of a stationary electrode in quiescent solution is varied by applying a linear voltage ramp to the electrode. The resulting signal is recorded as a plot of current versus potential. A peak-shaped voltammogram is observed (**Fig. 2**). The decay of current after the peak arises, as in the experiment at fixed potential, because of control of the current by diffusion; indeed, the decay after the peak has a $1/t^{1/2}$ dependence. The height of the peak is dependent upon the concentration of the electroactive species, and the peak potential E_p is related to $E^{\circ\prime}$. Analytically, the experiment is applicable in the 10^{-3} M

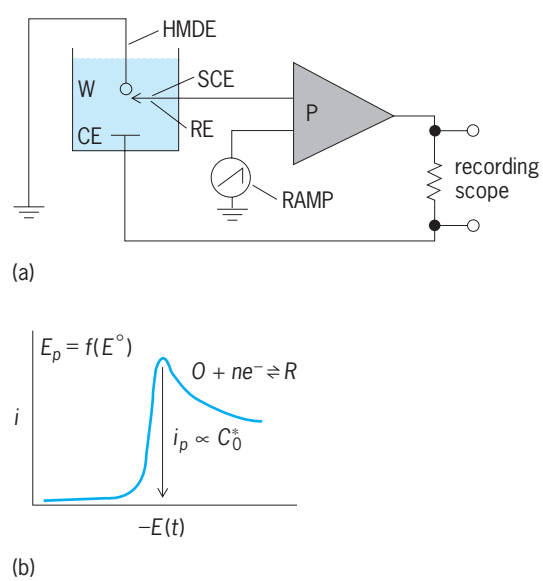

Fig. 2. Linear sweep voltammetry. (*a*) Generalized instrumentation. (*b*) Current–voltage curve obtained on the scope. HMDE = hanging mercury drop electrode; SCE = saturated calomel reference electrode; CE = auxiliary counter electrode; P = potentiostat; W = working electrode; RE = reference electrode; and RAMP = linearly increasing voltage signal.

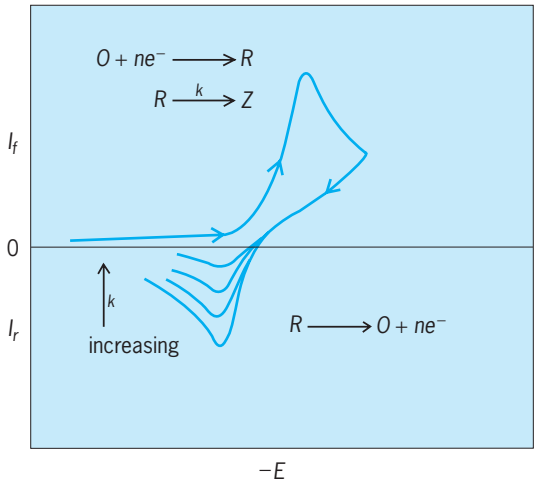

Fig. 3. Current–voltage curves in cyclic voltammetry.

range, but this can be extended to the 10^{-5} *M* range by plotting *dI/dE*, the derivative of the current with respect to potential, as a function of potential (so-called derivative LSV); to a good approximation, the background current rises constantly with potential in the region of E_p and is therefore eliminated by differentiation.

Cyclic voltammetry. The most widely used electrochemical analytical and mechanistic technique is cyclic voltammetry which bears the same relation to linear sweep voltammetry as double-potential-step chronoamperometry does to the single-step experiment. That is, at the end of the first linear ramp, the potential is swept linearly back to the initial potential. Typical cyclic voltammograms are shown in **Fig. 3** for the case of reduction of *O* to *R* and subsequent decomposition of *R* to *Z*. The figure shows a series of cyclic voltammograms corresponding to different values of the rate constant *k*. As in the double-step experiment, the appearance of the forward sweep is unaffected by the coupled chemical process. However, as *k* increases, the height of the peak on the reverse sweep due to reoxidation of *R* decreases. Sweep rates may be varied over a range of roughly 10^{-2} to 10^{2} V/s in conventional cyclic voltammetry. This may be extended considerably, as high as 10^{5} V/s or greater by using microelectrodes (electrodes with diameters in the micrometer range) and appropriate cell shielding to avoid stray currents.

Voltammetry in stirred solution. The potential in a cell configured for controlled potential electrolysis is varied linearly with time (about 5 mV/s), and the current is measured as a function of potential. The current increases sharply when the potential passes through the region of $E^{\circ\prime}$ for an electroactive species in the solution. Between reduction steps, current plateaus are established, the heights of which are proportional to the concentrations of each electroactive species. A typical voltammogram is shown in **Fig. 4*a***, where I_L is the limiting (plateau) current which is proportional to the bulk concentration of electroactive species C^*. The point designated as $E_{1/2}$ is called the half-wave potential and is approximately equal to $E^{\circ\prime}$. A small-amplitude (1–15 mV, 10–100 Hz) alternating potential (Fig. 4*b*) has been superimposed on the linear voltage ramp applied to the electrode of Fig. 4*a*. The resulting alternating-current response is peak-shaped, and the heights of the peaks are proportional to C^*. This experiment, alternating-current voltammetry, can be used to determine rates of heterogeneous electron transfer to or from the electrode to the electroactive species; in this case, it is generally done in quiescent solution and at a solid or dropping mercury electrode.

Stripping analysis. These techniques can sometimes be combined in useful ways. Analytical sensitivity is enhanced by the method known as stripping analysis. A constant reducing potential is applied to an electrode in a stirred solution containing one or more metal ions for a fixed time (about 1–60 min) sufficient to deposit substantial amounts of the metal on the electrode. After this controlled-potential electrolysis, a linear potential sweep to positive (oxidizing) potentials is applied to the electrode. Because of the preconcentration of the metal onto the electrode, the current is not limited by mass transport, and thus the stripping current is considerably larger than would have been observed in a voltammogram in stirred or quiescent solution. Sensitivity as high as 10^{-9} *M* is achieved. The technique is useful for measuring trace-metal analyses in seawater.

Square-wave polarography. This technique resembles conventional polarography at a dropping mercury electrode, except that a small-amplitude (about 1–25 mV) square wave (frequency about 225 Hz) is imposed on the controlling dc potential. The current measured is an alternating component obtained by measuring the current at times near the end of each half-cycle. This avoids the large capacitive background current associated with instantaneous electrode potential changes and allows greater

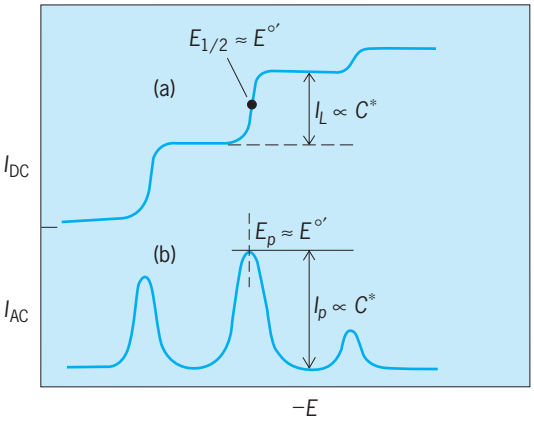

Fig. 4. Plots of the variation of current with constant applied potential in reduction processes. (*a*) Direct current versus voltage at a rotated working electrode (or stirred solution at a stationary electrode), where three successive reductions of electroactive species occur. (*b*) Alternating-current versus direct-current voltage applied at a working electrode in an alternating-current voltammetric experiment under the conditions of Fig. 4*a*.

sensitivity. Analyses over the range of 10^{-3} to 10^{-7} M are possible.

Pulse polarography. This is a variation of square-wave polarography, where a single small potential step is imposed on the controlling dc potential during the life of each mercury drop. The change in current with each step is measured and plotted as a function of the dc potential. It is convenient to make current measurements near the end of the mercury drop life, where the drop is growing slowest and the capacitative current is minimal. For this reason, the drop time is usually controlled mechanically and electronically synchronized with the current measurement. The sensitivity is similar to that of square-wave polarography. Like alternating current and square-wave polarography, a peak-shaped curve resembling Fig. 4*b* is obtained. All of these techniques may use stationary electrodes of carbon or a variety of metals in place of the dropping mercury electrode; in this case, they are referred to as voltammetry rather than polarography. *See* POLAROGRAPHIC ANALYSIS.

Controlled current. Controlled-current techniques, although easier to implement than those in which the potential is controlled, are not as selective and therefore have fallen out of favor in recent years. In chronopotentiometry, a constant current is imposed at the working electrode, and its potential is monitored with time. The electrode must assume a potential which will cause electrolysis sufficient to maintain the imposed current. Thus, the electrode initially adopts the potential of the most easily reduced species. As this species is depleted near the electrode, the potential shifts to that of the next most easily reduced species. This continues until reduction of solvent or electrolyte occurs. Characteristic curves are plotted in **Fig. 5**. The transition time τ is related to the bulk concentration of the electroactive species. However, it is difficult to establish an accurate background correction, thus rendering void the use of chronopotentiometry for analytical purposes.

Coulostatic analysis involves application of a very large, short pulse of current to the electrode, after which the cell circuit is opened and the return of the working electrode potential to its initial value

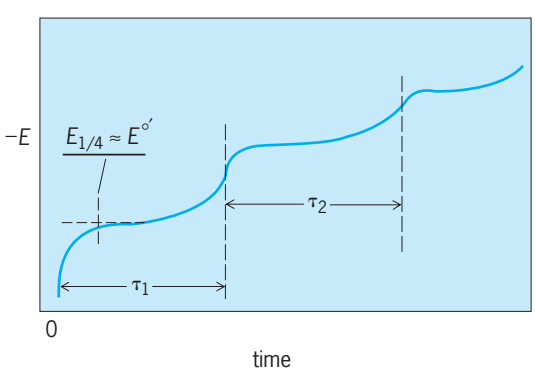

Fig. 5. Potential-time curves for chronopotentiometric experiments. Curves represent successive reduction of three different electroactive species in solution.

is monitored. The open-circuit condition requires that the current necessary to discharge the electrode interface comes from electrolysis of electroactive species in solution. The current pulse charges up the electrode–solution interface to a new potential, and the electrode discharges, and returns to its original potential, by reducing the electroactive species in solution. The time dependence of the change in electrode potential is proportional to C^*, but use of this dependence for analytical purposes requires the independent and inconvenient measurement of the electrode–solution capacitance. The change in electrode potential ΔE versus $t^{1/2}$ results in a straight-line plot, the slope of which is proportional to concentration, as in Eq. (4), where C is the electrical

$$\pm\Delta E = \frac{2nFD^{1/2}C^*}{\pi^{1/2}C}t^{1/2} \qquad (4)$$

capacitance associated with the electrode–solution interface.

Digital simulation. Computer simulation of electrode reactions involving one or more electron transfer steps together with several coupled chemical reactions is a challenging task because concentration gradients of the various chemical species involved vary as a function of both time and distance from the electrode during the experiment. Major advances in the necessary mathematical and computational techniques, as well as increases in computing power even at the desktop level, have led to corresponding major advances in the ease with which electrode reactions can be digitally simulated. Many modern electrochemical investigations, especially in the areas of organic and organometallic chemistry, rely on digital simulation to confirm mechanistic postulates and to extract kinetic and thermodynamic parameters.

Thin-layer electrochemistry. Electrochemical techniques can be applied in cells where a solution only 10–100 mm thick is electrolyzed. The principal advantage is that, for experiments lasting more than about 1 s, mass transfer limitations can be ignored. Thus, correlations between observed parameters and system characteristics are more straightforward. For example, for a chronopotentiometric experiment the transition time represents complete depletion of the electroactive species. Thus, C^* is proportional to τ for each of a sequence of reduction steps. Also, if a linear potential sweep is applied, a symmetrical current peak is obtained, the height or area of which is related to the concentration of electroactive species.

Albert J. Fry

Bibliography. A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*, 2d ed., 2000; A. J. Fry, *Synthetic Organic Electrochemistry*, 2d ed., 1989; P. T. Kissinger and W. R. Heineman (eds.), *Laboratory Techniques in Electroanalytical Chemistry*, 2d ed., 1996; K. B. Oldham and J. C. Myland, *Fundamentals of Electrochemical Science*, 1994; D. T. Sawyer, A. Sobkowiak, and J. L. Roberts, Jr., *Electrochemistry for Chemists*, 2d ed., 1995.

Electrochemistry

The science dealing with the chemical changes accompanying the passage of an electric current, or the reverse process in which a chemical reaction is used as the source of energy to produce an electric current, as in a battery or fuel cell. Electric conduction occurs through the motion of charged particles. The charged particles may be electrons (as in metals or semiconductors) or ions, which are electrically charged atoms, molecules, or molecular aggregates. Ionic conduction in electrolytes (liquid solutions, molten salts, and certain ionically conductive solids) is a phase of electrochemistry, as is corrosion. Conduction in metals, semiconductors, and gases is generally considered a portion of physics. *See* ELECTROLYTIC CONDUCTANCE.

Galvanic cells. These are better known as electric batteries. Many chemical reactions can be arranged to produce electrical energy by physically separating the reaction into half-reactions: one supplying electrons to an electrode forming the negative terminal of the cell, the other removing the electrons from the positive terminal. In the lead storage battery, electrons are supplied to the negative terminal by a half-reaction in which metallic lead is oxidized to form lead sulfate. At the positive terminal, lead dioxide is reduced to lead sulfate.

The electrons flowing in the external circuit from the negative to the positive terminal constitute the desired electric current. Charging of the lead storage battery by forcing a current to flow in the reverse direction results in the reversal of both half-reactions, and the storage of electric energy in the form of lead and lead dioxide. Such a cell is called a secondary cell, in contrast to a primary cell, such as the Leclanché cell or dry cell, which is not designed to be recharged. In the Leclanché cell, electric energy is produced by the oxidation of zinc and the reduction of manganese dioxide at a carbon electrode. The electrolyte is a moist mixture of zinc chloride, ammonium chloride, and powdered carbon. The fuel cell is designed for the continuous production of electric current through the consumption of oxidant and reductant at separate electrodes. The most common fuel cell is the hydrogen-oxygen (or air) cell with alkaline electrolyte. Many other systems, including hydrocarbon-chlorine and sodium-sulfur, have been proposed. *See* BATTERY; FUEL CELL.

Electrodeposition. The most important type of chemical reaction brought about by the passage of electric current is the deposition of a metal at a cathode from a solution of its ions. Electroplating of many metals, such as silver, cadmium, nickel, and chromium, is used for protective and decorative coatings. Electroforming is a variety of electrodeposition in which the electrically nonconductive prototype of an article to be produced is rendered conductive by spraying with a thin metallic coating, next electroplated with a metallic deposit that is then stripped from its substrate and filled with backing to reproduce the original article. Electrowinning is another application of electrodeposition used for the commercial production from molten salts of active metals such as aluminum, magnesium, and sodium, and from aqueous solution of other metals such as copper, manganese, and antimony. Electrorefining is commonly used to purify metals such as silver, lead, and copper. The impure metal is used as the anode, and purified metal is deposited at the cathode. *See* ELECTROMETALLURGY; ELECTROPLATING OF METALS.

Electrolytic processes. Many electrode reactions other than metal deposition are of commercial or scientific use. Electrolysis of brine to yield chlorine at the anode, hydrogen at the cathode, and sodium hydroxide in the electrolyte is a very important industrial process, consuming a significant fraction of the national electrical production. Many organic compounds can be prepared electrolytically. *See* ELECTROLYSIS.

Electrothermics. While not strictly electrochemical, electrothermics is generally recognized as a part of the field. It includes high-temperature processes involving electric arc or resistance furnaces. *See* ELECTRIC FURNACE.

Electroanalytical chemistry. Many electrochemical measurements are useful for analytical purposes. Electrodes are commonly used for analytical purposes through measurement of their potentials and include the glass electrode for pH measurements, and ion-selective electrodes for certain ions, such as sodium or potassium ion (special glass compositions), calcium ion (liquid membrane), and fluoride ion (doped lanthanum fluoride single crystals). Solid electrodes, especially those composed of platinum and carbon, are generally used to examine current-voltage behavior of substances in solution. Polarography involves the use of a dropping mercury electrode as one electrode of an electrolytic cell. Its use is declining except for analytical applications and those involving operation at highly negative potentials. Qualitative analysis is carried out by measurement of characteristic potentials (half-wave potentials) for electrode processes, and quantitative analysis by measurement of diffusion-controlled currents. Coulometry involves the application of Faraday's law for analytical purposes. *See* COULOMETER; ELECTRODE; ION-SELECTIVE MEMBRANES AND ELECTRODES; POLAROGRAPHIC ANALYSIS.

Several methods involving electrolysis during short periods of electrolysis permit the application of diffusion theory (in the absence of convection) to calculate mass-transport rates and to identify reactive intermediates and reaction mechanisms in multistep electrode processes. These techniques include linear sweep and cyclic voltammetry (measurement of currents with linear voltage scan) and chronopotentiometry (measurement of potential-time transients under constant current conditions). Several titration methods involve electrochemical measurements, for example, conductometric, potentiometric, and amperometric titrations. *See* ELECTROCHEMICAL TECHNIQUES; TITRATION.

Electrode kinetics. Studies of the kinetics of electrode processes are valuable not only for

understanding the mechanisms of electrode reactions but also for studying homogeneous reactions that occur in solutions preceding or following the charge-transfer step. Such studies are made by pulse or transient techniques, or by steady-state methods involving dynamic systems, such as rotating-disc electrodes or flowing solutions.

Miscellaneous phenomena. Electrochemical transport of ions through synthetic or natural membranes is important for processes such as desalination of water and electrodialysis. In biological systems, the transmittal of nerve impulses and the generation of electrical signals such as brain waves are basically of electrochemical origin. A set of related phenomena can be grouped together under electrokinetic behavior, including the motion of colloidal particles in an electric field (electrophoresis), the motion of the liquid phase relative to the stationary solid under the influence of a potential gradient (electroosmosis), and the inverse generation of a potential gradient caused by a flowing liquid (streaming potential). Alternating-current phenomena, such as dielectric behavior, double-layer charging, and faradaic rectification, may also be included in a general definition of electrochemistry. Corrosion and passivation of metals are electrochemical in nature. *See* COLLOID; CORROSION; ELECTROENCEPHALOGRAPHY; ELECTROKINETIC PHENOMENA; ELECTROPHORESIS; STREAMING POTENTIAL.

Albert J. Fry; Herbert A. Laitinen

Organic electrochemistry. Organic electrochemistry involves the study of the chemical reactions that take place when an electric current is passed through a solution containing one or more organic compounds. It is a highly interdisciplinary science because understanding and fully developing a given organic electrochemical reaction may involve not only techniques of synthesis, purification, and identification of organic products but also theory and practice of a variety of sophisticated electroanalytical techniques, surface science, cell design, electronics, engineering scaleup, and materials modification. In recent years, organic electrochemistry has been of increasing interest for industrial applications because the costs of electrical current have been rising more slowly than the costs of conventional chemical reagents and because electrochemical procedures can be environmentally less intrusive than conventional chemical processes. A number of fine chemicals are made electrochemically on scales ranging from several kilograms for expensive specialty compounds to several tons per day. A few chemicals are made on a much larger scale; the best known is adiponitrile [$NC(CH_2)_4CN$], a key intermediate in the synthesis of nylon.

Types of processes. An extremely wide range of organic chemical conversions can be carried out electrochemically. A single electrochemical cell and power supply can generate either anode potentials powerful enough to oxidize saturated alkanes or cathode potentials sufficiently negative to reduce benzene to cyclohexene. If necessary, such processes could be done simultaneously in the same electrochemical cell. An inorganic analogy illustrates this, whereby it is possible to oxidize fluoride ion to elemental fluorine at the anode of an electrochemical cell while reducing sodium ion to metallic sodium at the cathode. Of course, it is necessary to separate the two compartments so that the two products do not immediately react with each other. Almost any functional group can be modified electrochemically. *See* OXIDATION-REDUCTION.

The wide range of redox potentials that can be attained electrochemically and the precision with which this can be done represents only one aspect of organic chemical reactivity. A variety of other processes are possible in electrochemical cells. For example, it is possible to generate a chemical reagent at an electrode under carefully controlled conditions for reaction with another constituent of the medium. Reagents that have been generated in this manner include bromine, chlorine, iodine, iodine ion, hydrogen ion, hydroxide ion, solvated electrons, the powerful oxidant ruthenium oxide (RuO_4), and a variety of low- and high-valent metal ions. For many substrates, electrolysis is done at constant current, in which a fixed current is passed through the solution until the reaction is complete. In this technique, a simple and inexpensive power supply can be used. Selective electrolysis of a complex substance may require controlled-potential electrolysis, which uses a more sophisticated power source (a potentiostat) to maintain the electrode potential at a constant value. *See* ELECTROCHEMICAL PROCESS.

Methodology. The way that an electrochemical reaction is done depends upon the complexity of the process and the amount of material desired. In the laboratory, small amounts of material, from a few milligrams to a hundred grams, can be prepared in relatively simple cells or even in open beakers. Larger quantities are usually prepared in flow cells, often with many of these arranged in series or parallel. Industrial-scale electrochemical reactions are usually done in flow cells to reduce power costs and simplify operating conditions. An emerging technology is the use of microreactors for rapid (<1 s) reaction of electrochemically generated intermediates.

Electrode materials. A wide variety of substances have been used as electrodes, including the metals platinum, mercury, and iron, as well as carbon, which is proving to be the most versatile and least expensive electrode material. For many applications, the material of which the electrode is composed has little effect on the nature of the reaction taking place. The principal requirement is simply that the material be a conductor of electricity. Some applications, particularly those involving adsorbed intermediates and reagents, do take place more efficiently at electrodes made of certain materials. Electrodes made of an intimate mixture of carbon and sulfur, selenium, or tellurium have been used to introduce the latter elements into an organic substrate. Electrolysis at a platinum electrode in acidic media can effect catalytic hydrogenation of unsaturated substances. Other applications depend upon a special coating on the electrode surface to change

the nature of the electrode reactions occurring there. This coating can be a conductive polymer, a chemical redox agent, or an enzyme, where a high degree of selectivity is required. *See* ORGANIC CONDUCTOR.

Reaction mechanisms. The electrochemical behavior of thousands of organic substances has been studied. Consequently, it is possible to make generalizations about the kinds of processes likely to take place with a given substrate. Cathodic organic reactions fall into several categories: cleavage of single bonds, reduction of functional groups, and reduction of large conjugated systems such as activated alkenes and aromatic compounds. Anodic reactions are equally diverse. The oldest and best-known anodic reaction is the Kolbe reaction, in which electrochemical oxidation of carboxylate ions yields dimeric products with the evolution of carbon dioxide. A wide range of functional groups can be accommodated in this versatile reaction. Anodic oxidation of other substrates generally involves initial formation of cation radicals, which may undergo loss of a proton or alkyl group, reaction with solvent or a second component of the medium, or dimerization.

The powerful yet precisely controllable oxidizing and reducing conditions that can be achieved electrochemically are useful for generating novel organic intermediates, including carbocations, carbanions, radicals, radical ions, carbenes, nitrenes, and arynes. These reactive intermediates can frequently be trapped by other organic substances added to the medium to extend the synthetic utility of an electrode reaction. *See* ORGANIC REACTION MECHANISM; REACTIVE INTERMEDIATES.

Voltammetric methods are commonly used to provide detailed mechanistic information. In these, the electrode potential is varied in a controlled fashion in the vicinity of the redox potential of the substrate, and the current response is measured as a function of experimental variables such as scan rate, concentration, and added reagents. Other methods include identification of the products from a preparative electrolysis (often at controlled potential in mechanistic experiments); coulometry, that is, measurement of the actual amount of current passed in the electrolysis; comparison of experimental voltammetric or other responses with computer simulations of the theoretical behavior for various mechanisms; studies of changes in the structure of the electrode surface during electrolysis; and spectroscopic identification of intermediates during electrolysis. Microelectrodes, which have micrometer diameters, extend the speeds with which electrochemical measurements can be made, and thus permit measurement of even very fast rates of chemical reactions associated with electron transfers at electrodes.

Albert J. Fry

Bibliography. A. J. Fry, *Synthetic Organic Electrochemistry*, 2d ed., 1989; H. Lund and O. Hammerich (eds.), *Organic Electrochemistry*, 4th ed., 2000; T. Shono, *Electroorganic Synthesis*, 1991; S. Torii (ed.), *Novel Trends in Electroorganic Synthesis*, 1998.

Electrochromic devices

Self-contained, hermetically sealed, two-electrode electrolytic cells that change their ability to transmit (or reflect) light in response to a small bias (typically 1–2 V) applied across the two electrodes. The operation of electrochromic devices relies upon their electrochromic material content. These materials are organic or inorganic substances that are able to interconvert between two or more color states upon oxidation or reduction, that is, upon electrolytic loss or gain of electrons. The electrochromic materials that are appropriate for most practical applications are strong light absorbers in one redox state but colorless in another. Historically, the phenomenon of electrochromism based on physical phenomena, such as the Franz-Keldish and Stark effects, dates back to 1932. Since the mid-1970s, however, development has been focused on color changes carried out electrochemically, so that in modern usage the terms electrochromic devices and materials refer to devices and substances defined above. *See* STARK EFFECT.

Assembly. A typical electrochromic device is a sandwichlike structure with two glass plates and an electrolyte (**Fig. 1**). Each glass plate is coated on the inside with a transparent electrically conducting layer of indium-tin oxide, which operates as an electrode. Electrochromic mirrors include an additional reflective coating (for example, aluminum) on the outside of one of the glass plates. The electrolyte carries the ionic current inside the cell between the two electrodes, and it can be as simple as a salt (for example, sodium chloride, NaCl) dissolved in a dissociating solvent such as water. However, development has focused on gel and solid electrolytes, because they offer several advantages: they are easier to confine in the space between the electrodes; they function as

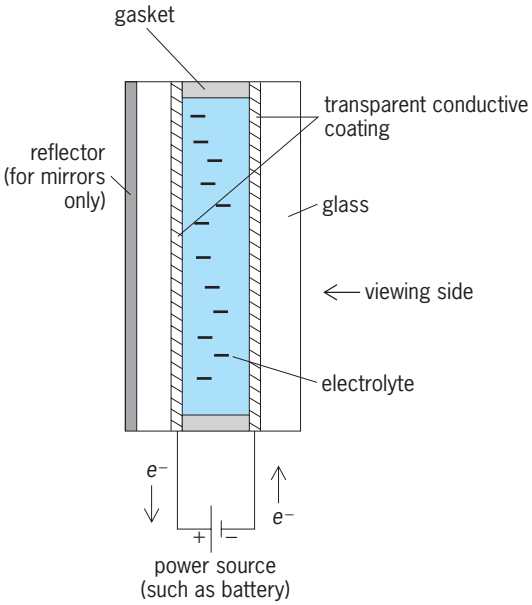

Fig. 1. Electrochromic device. The electrochromic materials can be either dissolved in the electrolyte or coated on the transparent electrodes.

laminators holding the two glass plates together; and their use minimizes the hydrostatic pressure that can cause substrate deformation and leakage problems, particularly in large-area devices such as smart windows.

State-of-the-technology electrochromic devices utilize two electrochromic materials with complementary properties: the first electrochromic material is normally reduced (ECM_1^{red}) and undergoes a colorless-to-colored transition upon oxidation (loss of electrons), while the second electrochromic material is normally oxidized (ECM_2^{ox}) and undergoes a similar transition upon reduction (gain of electrons). The electrochromic materials ECM_1^{red} and ECM_2^{ox} are selected so that they do not react with each other. The oxidation of ECM_1^{red} and the reduction of ECM_2^{ox} then are forced by the external power source (Fig. 1), which operates as an electron pump consuming energy in order to transfer electrons from one electrode to the other. Oxidation of ECM_1^{red} occurs at the positive electrode (anode) and is a source of electrons, while reduction of ECM_2^{ox} occurs at the negative electrode (cathode) and is a sink of electrons. This approach, known as complementary counterelectrode technology, has two distinct advantages. First, the long-term operating stability of the electrochromic cell is greatly enhanced, because providing both a source and a complementary sink of electrons within the same cell prevents any electrolytic decomposition of the electrolyte. Second, the reinforcing effect of two electrochromic materials changing color simultaneously enhances the contrast difference between the color states per unit charge consumed. Depending on the location of the two electrochromic materials within the electrochromic devices, three main types of such devices exist: solution, precipitation, and thin-film (**Fig. 2**). *See* ELECTRODE; ELECTROLYTE; OXIDATION-REDUCTION.

Solution type. In these devices both electrochromic materials, ECM_1^{red} and ECM_2^{ox}, are dissolved in the electrolyte, and they move to the electrodes by diffusion driven by the concentration gradients that are generated by the electrolytic depletion of ECM_1^{red} and ECM_2^{ox} in the vicinity of the electrodes (Fig. 2*a*). The main advantage of the solution-type electrochromic devices is the variety of materials that can be used; any material that can be oxidized or reduced and is electrochromic is a potential candidate. Also, unlike the other types of electrochromic devices, the maximum color intensity can be controlled easily by varying the thickness of the electrolyte layer, typically between 0.1 and 5 mm (0.004 and 0.2 in.). More than 90% of the electrochromic antiglare rearview mirrors for automobiles sold worldwide are based on solution-type devices. A typical device employs *N,N,N′,N′*-tetramethyl-*p*-phenylenediamine (TMPD) as ECM_1^{red}, and *N,N′*-diheptyl-4,4′-bipyridinium dication (HV^{2+}) as ECM_2^{ox}. Electric current forced through the electrode-solution interface is responsible for the electrolysis of the electrochromic materials to their colored forms. *See* ELECTROLYSIS.

In organic solvent-based electrolytes the blue radicals $TMPD^{\cdot+}$ and $HV^{\cdot+}$ are also soluble and diffuse away from the electrodes back into the bulk electrolyte, where they meet, react, and annihilate each other back to their respective colorless states. This property is a desirable feature that renders electrochromic rearview mirrors self-erasing, thereby providing a fail-safe system that automatically recovers the colorless, fully reflective state under power-failure conditions. However, in order for these devices to remain colored, power must be applied continuously. This property increases the energy consumption and makes these devices rather unsuitable for application in architectural glaze. Finally, another potential drawback of this approach is the speed of coloration, which is relatively slow (typically on the order of a few seconds) because it is controlled by diffusion of the electrochromic materials in the bulk electrolyte, a slow process. Implementation of gel or solid electrolytes with these devices slows the diffusion processes even further. *See* DIFFUSION.

Precipitation type. In these devices at least one of the electrochromic materials is originally dissolved in the electrolyte, but upon oxidation or reduction the colored product is electrodeposited onto the corresponding electrode (Fig. 2*b*). One of the first examples (1962) of electrolytic cells in light modulation is a precipitation device in that it employs the reversible electroplating of silver. Later (1973) it was discovered that in aqueous electrolytes the one-electron reduction product, $HV^{\cdot+}$, of the *N,N′*-diheptyl-4,4′-bipyridinium dication, HV^{2+}, forms a blue precipitate on the electrode. That discovery significantly shifted attention from physical to electrochemical electrochromism.

The switching speed of precipitation-type devices is comparable to that of solution types, because both are controlled by diffusion. But precipitation devices offer inherently higher resolution. For example, with patterned-segment electrodes they can be used potentially for high-resolution displays. However, such devices have not been commercialized because precipitation-type electrochromic materials usually suffer poor cycling lifetimes; certain recrystallization effects within the colored precipitates progressively render them electrically inaccessible by the electrode.

Thin-film type. In these devices, both electrochromic materials are immobilized and confined as thin-film coatings (0.1–5 micrometers thick) on the surfaces of the two electrodes (Fig. 2*c*). The main advantage of this arrangement is that physical separation of the two electrochromic materials prevents annihilation of the colored forms, and provides an open-circuit memory that significantly decreases the consumption of power. A typical device consumes less than 10 millicoulombs/cm^2 for full coloration. In a sense, a thin-film-type electrochromic device is a rechargeable battery, in which the color of the electrodes depends upon the state of charge. Discharging (decolorizing) can be carried out either by reversing the bias or by simply short-circuiting the two

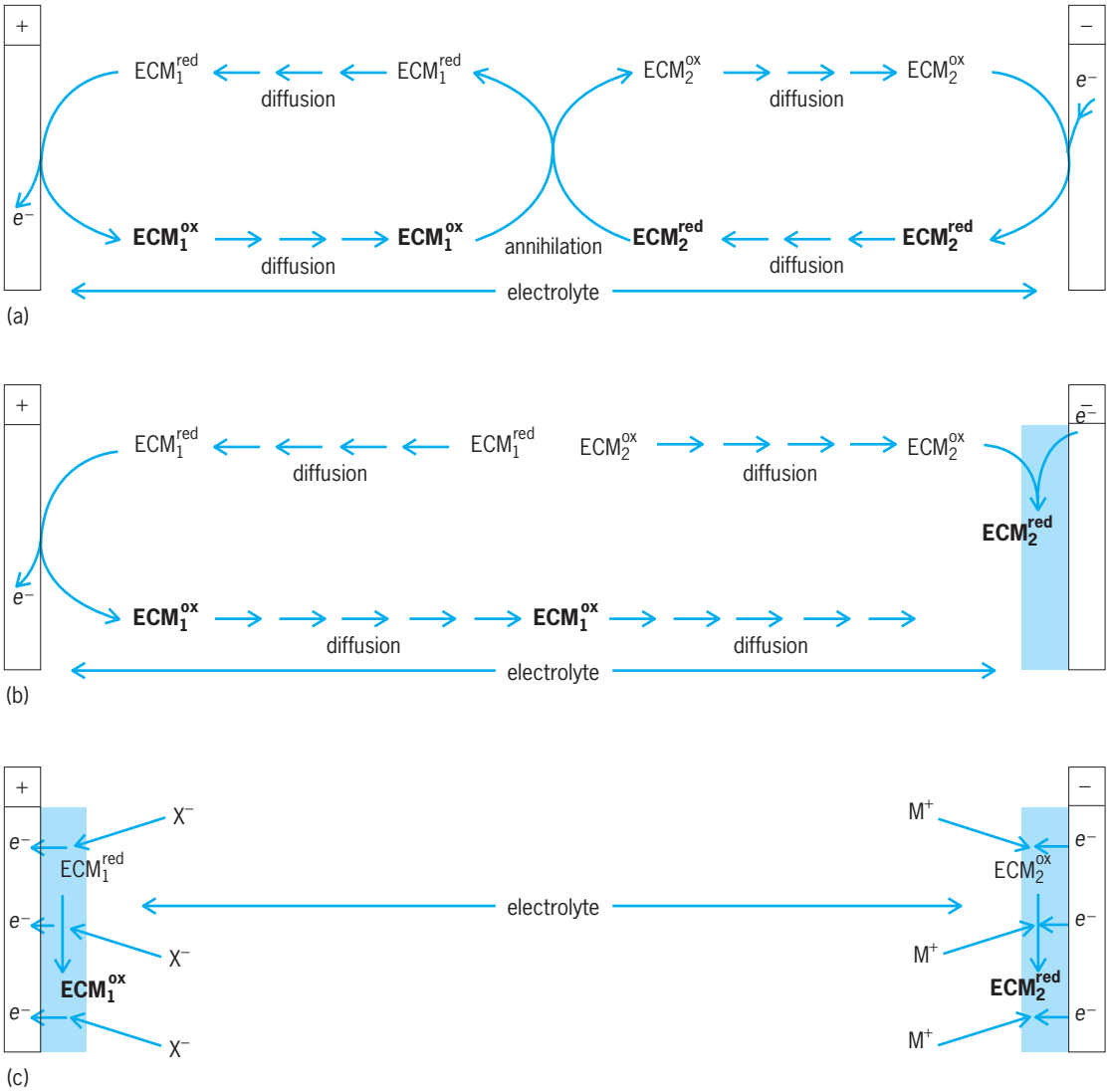

Fig. 2. Operating principles of electrochromic devices. (*a*) Solution type. (*b*) Precipitation type. (*c*) Thin-film type. The color-bearing forms of the electrochromic materials are shown in bold type.

electrodes. Furthermore, the switching speed of this type of device can be greatly improved because coloration does not depend on electrolysis of electrochromic materials dissolved in the bulk electrolyte. In practice, however, coloration depends on diffusion of ions in thin solid films, which is a much slower process than diffusion in the bulk electrolyte. Therefore, depending on the nature of the electrochromic materials, the switching speed of thin-film devices can vary from a fraction of a second to a few tens of seconds.

Several organic and inorganic materials are available for surface confinement on electrodes as thin films. A film of tungsten trioxide (WO_3), for instance, is an inorganic electrochromic material which, upon reduction, uptakes hydrogen ions (H^+) reversibly from an aqueous electrolyte [or lithium ions (Li^+) and sodium ions (Na^+) from a nonaqueous electrolyte] to form a blue tungsten bronze (M_nWO_3, where M is H, Li, or Na). Certain electrochromic rearview mirrors used in trucks employ WO_3 as ECM_2^{ox}, and a nickel hydroxide [$Ni(OH)_2$] film as ECM_1^{red}. These are manual, switch-operated devices which are preset at a level with which the driver feels comfortable. Another inorganic electrochromic material that is complementary to WO_3 and suitable for thin-film electrochromic devices is Everitt's salt, $K_4Fe_4[Fe(CN)_6]_3$, which is the colorless reduced form of Prussian blue, $Fe_4[Fe(CN)_6]_3$.

Among organic materials, conductive polymers such as poly(pyrrole), poly(aniline), and poly(thiophene) have been evaluated as thin-film electrochromic materials for oxidative coloration, and polymeric films derived from diquaternized 4,4′-bipyridine for reductive coloration. Generally, organic electrochromic polymer films are able to switch color in less than a second, and they seem to be inherently more durable than their inorganic counterparts because they better accommodate the structural changes induced upon oxidation and reduction. *See* ORGANIC CONDUCTOR.

Properties. Electrochromic devices are analogous to liquid-crystal devices in that they do not generate their own light but modulate the ambient light.

Liquid-crystal devices require use of polarizers; consequently, their viewing angle is limited, and lateral size limitations are imposed because the spacing between the electrodes (thickness) must be controlled within a few micrometers over the entire device area. Electrochromic devices do not require polarizers, thereby allowing a viewing angle approaching 180°, and contrast ratios similar to black ink on white paper (20:1 or better); moreover, control of the thickness is not important. Other desirable features of electrochromic devices include inherent color, continuous gray scale, and low average power consumption for the thin-film-type devices. Furthermore, it has been shown that electrochromic thin films can be patterned with a 2–5-μm resolution to form a large number of display elements that can be matrix-addressed. Nevertheless, even though there is no apparent intrinsic limitation, the best cycling lifetimes claimed for electrochromic materials are of the order of 10–20 million cycles, while the lifetime of liquid-crystal devices is of the order of several hundred million cycles. This long lifetime has made liquid-crystal devices a very successful technology in matrix-addressed, flat-panel displays.

The larger tolerance in thickness variation for electrochromic devices renders them better suited than liquid-crystal devices for large-area light modulation applications, such as smart windows, space dividers, and smart mirrors. Another possible application is in large-area displays that do not need frequent refreshing, such as signs and announcement boards. Reconfigurable optical recording devices (for example, disks) have been proposed as a high-resolution application that is within the presently available lifetimes of electrochromic materials. *See* ELECTROCHEMICAL PROCESS; ELECTROCHEMISTRY; ELECTRONIC DISPLAY; ELECTROOPTICS; LIQUID CRYSTALS. Nicholas Leventis

Bibliography. A. R. Kmetz and F. K. von Willisen (eds.), *Nonemissive Electrooptic Displays*, 1976; C. M. Lampert and C. G. Granqvist (eds.), *Large-Area Chromogenics: Materials and Devices for Transmittance Control*, vol. IS 4, SPIE Institute Series, 1990; N. Leventis and Y. C. Chung, New complementary electrochromic system based on polypyrrole-prussian blue composite, a benzylviologen polymer, and poly(vinylpyrrolidone)/potassium sulfate aqueous electrolyte, *Chem. Mater.*, 4:1415–1422, 1992; S. Matsumoto (ed.), *Electronic Display Devices*, 1990.

Electroconvulsive therapy

The controlled induction of a grand mal seizure under anesthesia for treating individuals suffering from certain psychiatric illnesses. Electroconvulsive therapy (ECT) was introduced as a treatment for psychiatric disorders by the Italian neurologist Ugo Cerletti in 1938. Cerletti developed a method of electrically inducing seizures in laboratory animals as a means of studying epilepsy. Aware of the observation that individuals with epilepsy and depression experienced an improvement in mood following a seizure, he postulated that inducing a seizure in a depressed individual might improve his or her condition. In 1938 he reported the first case of a psychiatric patient treated with ECT with dramatic, positive results.

Since 1938 ECT has developed considerably. The major changes in the treatment since its inception include the use of general anesthesia to modify the motor component of the seizure; the use of a pulse wave ECT apparatus, which allows careful dosing of electrical energy; the delineation of specific pathological conditions that respond to ECT; and a clearer understanding of how electrode placement and energy dosing affects both therapeutic results and side effects.

Procedure. ECT is performed under general anesthesia. While often performed on hospitalized patients, it is increasingly used in an outpatient setting. The patient is evaluated medically to determine the risk of anesthesia. The required medical workup consists of a physical examination, an electrocardiogram, and laboratory tests, usually including a blood count and determination of electrolytes concentrations. The patient is not allowed to eat or drink after midnight prior to the procedure. Both cardiovascular monitoring and electroencephalographic (EEG) monitoring are done during the procedure. An intravenous line is started, and the patient is anesthetized using a short-acting barbiturate. After the patient is fully asleep, he or she is paralyzed using a short-acting muscle relaxant. When the patient is fully asleep and paralyzed, a grand mal seizure is induced using a pulse wave ECT apparatus. The seizure typically lasts 30 to 60 s. Patients regain consciousness in 5 to 15 min. They are monitored until they are fully recovered and able to eat and drink, which usually takes an additional 30 min.

Treatment dosing. Although all ECT treatment induces a grand mal seizure, it is now known that not all seizures are alike. Both the efficacy of the treatment and the side effects are influenced by stimulus electrode placement and the quantity of energy. Bilateral and right unilateral electrode placements are the most widely used treatment placements and the most researched. Bilateral placement with a dosing slightly above seizure threshold is marginally more effective than right unilateral placement with an energy dosing 4–6 times above seizure threshold. Bilateral placement yields a slightly faster response but also slightly more cognitive side effects. Recently, bifrontal electrode placement with an energy dosing slightly above seizure threshold has been introduced with results comparable to bilateral placement. However, this placement requires more study before it is widely introduced into clinical practice.

ECT-responsive disorders. The use of ECT to treat serious psychiatric disorders has been refined over the past half century. It has shown efficacy in major depression with and without psychotic features; in both the manic and depressed phases of bipolar disorder; in the treatment of acute schizophrenia, particularly when confusion is a prominent symptom;

and in the treatment of catatonia (a psychiatric condition characterized by unresponsiveness to external stimuli) regardless of the psychiatric etiology. Due to the expansion of medication options over the last 20 years, pharmacotherapy is almost always tried before ECT is initiated. The most notable exceptions to this are in circumstances in which a fast response is needed, such as in the treatment of individuals who are dangerously suicidal or severely nutritionally impaired. ECT can show results in days rather than the weeks it typically takes with medication treatment.

The efficacy of ECT in these disorders varies. It appears to be most effective in the treatment of depressed individuals with psychotic symptoms, yielding about a 92% response rate. This compares with an approximately 62% response rate in individuals with major depression without psychotic symptoms. ECT is also extremely effective in the treatment of catatonia, with response rates of about 90%. The efficacy of medication in the treatment of mania and acute schizophrenia is so high that ECT is rarely used in these disorders currently. ECT has been used to treat obsessive-compulsive disorder (OCD) with disappointing results. However, there is a high rate of comorbid (coexisting) depression with OCD, and in this circumstance ECT can be very effective at treating the depression.

ECT has also shown efficacy in the treatment of neuroleptic malignant syndrome (NMS), a confusional state believed to result from an idiosyncratic reaction to antipsychotic medication. ECT has been shown to limit the symptoms and shorten the course of NMS, particularly when it is initiated early.

ECT also has efficacy in treating the depression of Parkinson's disease. In addition, patients with Parkinson's without depressive symptoms have shown improvement in their motor symptoms after receiving a course of ECT. *See* AFFECTIVE DISORDERS; OBSESSIVE-COMPULSIVE DISORDER; PARKINSON'S DISEASE; SCHIZOPHRENIA.

Side effects. The side effects of ECT are divided into two categories: those resulting from the induction of a grand mal seizure and those resulting from the administration of anesthesia.

Grand mal seizure. The induction of a grand mal seizure is accompanied by numerous physiologic changes that can cause side effects. When the electrical stimulus is administered, there is an immediate centrally mediated vagal nerve response, which results in sinus bradycardia (a slow heart rate). This is followed by sinus tachycardia (a fast heart rate) and increase in blood pressure during the seizure. When the seizure ends, there is a peripherally mediated vagal nerve response, which again results in bradycardia. As a consequence, the heart is subjected to rapid changes in rate and blood pressure. This can result in both ventricular and atrial arrhythmias. Typically these arrhythmias consist of several premature atrial contractions or premature ventricular contractions that are self-limited. The potential does exist for the development of significant and potentially life-threatening arrhythmias. For this reason alone, ECT is always done in the presence of an anesthesiologist who is prepared to treat any arrhythmia that might occur. *See* CARDIAC ELECTROPHYSIOLOGY.

The occurrence of a grand mal seizure is associated with short-term memory loss and confusion. The seizure itself causes both anterograde and retrograde amnesia, that is, loss of memory for events occurring both after and before the seizure, respectively. Typically this effect is cumulative so that individuals report memory problems worsening over the course of treatment. These effects are most intense during the treatment course and the several weeks surrounding it. On rare occasion individuals have reported autobiographical memory loss as far back as 6 months prior to the treatment. Evaluating the deleterious cognitive effects from ECT is confounded by the fact that depression causes significant memory problems, as do many of the medications used in its treatment. When assessing the risks and benefits of ECT in regard to memory problems, it is important to weigh the fact that tests show memory functioning—as measured by attention, concentration, and the capacity to acquire and use new information—actually improves after ECT.

The grand mal seizure also increases intracranial pressure. A minor consequence of this is headache. A potentially major consequence is that any central nervous system condition that is at risk for worsening from this increase in pressure needs to be carefully evaluated prior to the individual's undergoing treatment. For example, aneurysms can potentially rupture, or tumors could result in herniations. *See* AMNESIA; MEMORY; SEIZURE DISORDERS.

Anesthesia. Anesthesia-related side effects are similar to those associated with anesthesia used in any minor surgical procedure. The most serious risk is an idiosyncratic reaction to the anesthetic agent, which is associated with a mortality rate of approximately 1 in 100,000. More common side effects are minor headaches, muscle pain, dry mouth, and, rarely, urinary retention. *See* ANESTHESIA.

Mechanism of action. The therapeutic effects of ECT are related to the energy level, the grand mal seizure, and the electrode placement. The treatment produces several effects on neurotransmitters that can be responsible for its efficacy. There is an increase in both neurotransmitter production and release and neurotransmitter receptor sensitivity. There is also an increase in seizure threshold over the course of the treatments that is associated with a clinical response. Energy dosing appears to be important with respect to the electrode placement. Energy dosing at seizure threshold is effective with bilateral electrode placement, but the dosing needs to be four to six times above threshold when right unilateral placements are used. The reason for this is not precisely understood. There is a relative sparing of the speech and language center when right unilateral electrode placement is used, which may explain the reduction in cognitive side effects with this technique.

ECT has been shown to increase long-term potentiation (LPT) pathways in laboratory animals. LPT is

involved in the formation of neural nets related to memory formation. When this effect is blocked in laboratory animals using chemical agents, the memory side effects of ECT are reduced. Work is now under way with human subjects to see if blocking this effect can result in decreased memory side effects. Robert Ostroff

Bibliography. American Psychiatric Association Committee on Electroconvulsive Therapy, *The Practice of Electroconvulsive Therapy*, 2d ed., American Psychiatric Press, Washington, DC, 2001; S. H. Lisanby et al., The effects of electroconvulsive therapy on memory of autobiographical and public events, *Arch. Gen. Psychiat.*, 57:581–590, 2000; K. G. Rasmussen, S. M. Sampson, and T. A. Rummans, Electroconvulsive therapy and newer modalities for the treatment of medication-refractory mental illness, *Mayo Clinic Proc.*, 77:552–556, 2002; H. A. Sackeim et al., A prospective, randomized, double-blind comparison of bilateral and right unilateral electroconvulsive therapy at different stimulus intensities, *Arch. Gen. Psychiat.*, 57:425–435, 2000.

Electrode

An electrical conductor through which an electric current enters or leaves a conducting medium, whether it be an electrolytic solution, solid, molten mass, gas, or vacuum. For electrolytic solutions, many solids, and molten masses, an electrode is an electric conductor at the surface of which a change occurs from conduction by electrons to conduction by ions. For gases and vacuum, the electrodes merely serve to conduct electricity to and from the medium. *See* ELECTRODE POTENTIAL; ELECTROLYSIS; ELECTROMOTIVE FORCE (CELLS). Walter J. Hamer

Electrode potential

The equilibrium potential difference between two conducting phases in contact, most often an electronic conductor such as a metal or semiconductor on the one hand, and an ionic conductor such as an electrolyte solution (a solution containing ions) on the other. Electrode potentials are not experimentally accessible, but the differences in potential between two electronic conductors making contact with the same ionic conductor (that is, the difference between two electrode potentials) can be measured. A useful scale of electrode potentials can therefore be obtained when a particular electrode potential is set equal to zero by definition. There are several conventions, based on different definitions of the zero point on the scale of electrode potential, but all tables use the so-called standard hydrogen convention. *See* ELECTRODE; REFERENCE ELECTRODE.

Mechanism. The interfacial potential difference is usually the consequence of the transfer of some charge carriers from one conducting phase to the other. For example, when a piece of silver, which contains silver ions and free, so-called conduction electrons, is in contact with an aqueous solution of silver nitrate, the only species common to the two phases are the silver ions. Their concentration (volume density) is constant in the metal but variable (from zero to the solubility limit of the silver salt used) in the solution. When more silver ions transfer from the solution to the metal than in the opposite direction, an excess of negatively charged nitrate ions remains in the solution, which therefore acquires a negative charge. However, the metal gains more silver ions than it loses, and therefore acquires a positive charge. Such a charge separation leads to a potential difference across the boundary between the two phases. The continued buildup of such charges makes the potential of the metal more and more positive with respect to that of the solution. This effect in turn leads to electrostatic repulsion of the silver ions in the solution phase immediately adjacent to the metal; these are the very metal ions that are candidates for transfer across the boundary. Consequently, the electrostatic repulsion decreases the tendency of silver ions to move from the solution to the metal, and eventually the process reaches equilibrium, at which point the tendency of ions to transfer is precisely counterbalanced by the repulsion of the candidate ions by the existing potential difference. At that potential, there is no further net transfer of charges between the contacting phases, although individual charges can still exchange across the phase boundary, a process which gives rise to the exchange current.

With a small number of metals, especially silver and mercury, interfacial equilibrium is often established rapidly, well within milliseconds. However, with most metals, the equilibrium state is reached much more slowly, especially when the electrode process involves bond breaking or other complicated reaction mechanisms. In some cases, equilibrium is never attained, and the equilibrium potentials can be known only through calculation.

Concentration dependence. For a metal in contact with its metal ions of valence z, the potential difference E can be expressed in terms of a standard potential E° (describing the affinity of the metal for its ions) and the concentration c of these ions in solution through the Nernst equation (1), where R is the gas

$$E = E^\circ + (RT/zF) \ln c \qquad (1)$$

constant, T is the absolute temperature, and F is the Faraday.

When the metal ions in solution form a sparingly soluble salt, the solubility equilibrium can be used to convert a metal electrode responding to its own metal cations (positive ions) into an electrode responding to the concentration of anions (negative ions). Typical examples are the silver/silver chloride electrode, based on the low solubility of silver chloride (AgCl), and the calomel electrode, based on the poor solubility of calomel (Hg_2Cl_2).

An equilibrium potential difference between a metal and an electrolyte solution can also be established when the latter contains a redox couple, that is, a pair of chemical components that can be converted into each other by the addition or withdrawal of electrons, by reduction or oxidation respectively. In that case the metal often merely acts as the supplier or acceptor of electrons. When metal electrons are donated to the solution, the oxidized form of the redox couple is reduced; when the metal withdraws electrons from the redox couple, its reduced component is oxidized. Again, the buildup of a charge separation generates a potential difference, which counteracts the electrochemical charge transfer and eventually brings the process to equilibrium, a state in which the rate of oxidation is exactly equal to the rate of reduction. The dependence of the equilibrium electrode potential on the concentrations of c_{ox} and c_{red} of the oxidized and reduced forms respectively is described by a Nernst equation of the form (2), where n denotes the number of electrons

$$E = E^{\circ} + (RT/nF)\ln(c_{ox}/c_{red}) \qquad (2)$$

(e^-) transferred between the oxidized species (Ox) and the reduced species (Red) in the reactions $\mathrm{Ox} + ne^- \rightleftarrows \mathrm{Red}$. A typical example is the reduction of hydrogen ions H^+ to dissolved hydrogen molecules H_2, and vice versa, in which case the reactions are $2H^+ + 2e^- \rightleftarrows H_2$, and for which the Nernst equation is of the form (3).

$$E = E^{\circ} + (RT/2F)\ln(c^2_{H^+}/c_{H_2}) \qquad (3)$$

Redox potentials involving a gas are often established slowly, if at all. For determinations of such a redox potential, platinum is often used as the metal, because it is chemically and electrochemically stable. *See* OXIDATION-REDUCTION.

Electrode potentials can also be established at double phase boundaries, such as that between two aqueous solutions separated by a glass membrane. This glass electrode is commonly used for measurements of the pH, a measure of the acidity or basicity of solutions. The mechanism by which the glass electrode operates involves ion exchange of hydrogen ions at the two glass-solution interfaces. *See* ION EXCHANGE.

In all the above examples, the two contacting phases can have only one type of charge carrier in common. Usually, no equilibrium potential difference is established when more than one type of charge carrier can cross the interface, but often (depending on the nature of the metal and of the chemical components of the solution, and sometimes also depending on the geometry of the contact region) an apparently stable potential can still be obtained, which corresponds to zero net charge transfer. This can be a so-called mixed potential, important in metal corrosion, or a junction potential, which figures in most measurements of electrochemical potentials and usually limits the accuracy and precision of such measurements, including that of the pH. *See* CORROSION; ELECTROCHEMICAL SERIES.

Complications. In determining electrode potentials, there are several complications. In the first place, it follows from thermodynamics that the Nernst equation should be written in terms of activities rather than concentrations. The difference between these two parameters is often small but seldom completely negligible. The problem is that the activities of ions (in contrast to those of neutral molecules) are not experimentally accessible, although there are theoretical models that allow their estimation in sufficiently dilute solutions. *See* ACTIVITY (THERMODYNAMICS).

Second, measurements of potential differences always involve the potential difference between two metals rather than that between a metal and a solution. Therefore, electrode potentials as defined above cannot be measured either. The latter complication is circumvented by setting the electrochemical potential of one particular metal-solution interface equal to zero by definition. W. Nernst suggested that a normal hydrogen electrode be used as the common zero point; he defined this as a metal in contact with a solution saturated with hydrogen gas at 1 atm partial pressure and containing 1 mole of a monoprotic acid per liter. However, it was soon found that the nature of the acid used, and the presence of other components in the solution, would affect the potential of such an electrode—a reflection of the fact that the electrode responds to the activity rather than the concentration of the hydrogen ions. The problem here is that the activity of ions (including that of hydrogen ions) is not known, except at or near infinite dilution. Thus the standard hydrogen electrode, defined in terms of hydrogen ion activity rather than concentration, is not experimentally realizable. An alternative definition, in terms of an acid activity, has an anion sensitivity similar to the original definition in terms of concentrations. Moreover, use of the hydrogen electrode can lead to a high liquid junction potential, whereas use of the standard calomel electrode tends to reduce it. In practice, therefore, other electrodes are often used, such as silver–silver chloride electrode or a calomel electrode. In many areas of electrochemistry, the de facto reference electrode is the saturated calomel electrode, which consists of mercury plus calomel in a saturated aqueous solution of potassium chloride (KCl). The standard hydrogen electrode is used mostly for physico-chemical calculations.

Because these measurements involve a potential difference, there has been considerable confusion about the definition of that difference; that is, whether it is defined as the potential of the metal minus that of the solution, or the other way around. This is a matter of a sign convention. The problem is usually framed in terms of oxidation potentials versus reduction potentials. In 1953 the international community settled on the latter and simply called them electrode potentials.

Another zero point is often used to anchor the scale of electrode potentials of semiconductor electrodes. Here the reference point is the work function, that is, the electrical work necessary to remove an

electron from the interior of the phase to a far-away position in vacuum. This is, theoretically, a more meaningful scale. Unfortunately, the work function is not known with sufficient accuracy and precision to make this scale practical. *See* WORK FUNCTION (THERMODYNAMICS).

There are other factors besides charge transfer that can cause interfacial potential differences. Oriented dipolar molecules at interfaces can give rise to interfacial potential differences, even though they represent a charge separation within a layer of only molecular dimensions. Similarly, adsorbed molecules and ions can affect the interfacial potential difference when their adsorption involves a partial charge transfer. None of these phenomena affect the thermodynamic potentials, but they have effectively blocked attempts to define an absolute potential that is both practical and theoretically meaningful.

Applications. There are four main applications for measurements of electrode potentials: (1) in the establishment of the oxidative and reductive power of redox systems, the so-called electromotive series; (2) as concentration probes, such as in pH measurements; (3) as sources of chemical equilibrium data; and (4) as the primary (or independent) variable in studies of electrode reactions.

The electromotive series is a listing of redox couples in the order of their potentials. The most positive potentials correspond to the strongest oxidants, while most negative potentials identify the strongest reductants. This series therefore serves to organize redox systems according to their oxidative or reductive power. *See* ELECTROMOTIVE FORCE (CELLS).

A combination of two electrodes can be employed to determine the concentration of a particular species in a sample solution. One of the two electrodes is the indicator electrode, sensitive to a particular ionic species in the sample solution. The other, reference electrode typically comprises a separate compartment containing a metal in contact with its sparingly soluble salt and an excess of the anion of that salt (as in the calomel and silver/silver chloride electrodes), or with a redox couple such as the combination of iodide and triiodide ions. The solution in the reference electrode compartment then makes contact with the sample solution through a very constrictive contact, such that solution mixing is minimized. Moreover, the nature of the solution in the junction is chosen so that the potential difference across that liquid junction will be as small as possible. This can be achieved by using a high concentration of a salt with, as much as possible, equitransferent ions (ions that have near-equal mobilities, such as in KCl or NH_4NO_3). Under those conditions the liquid junction potential can often be neglected, and the Nernst equation can be written in terms of measurable potentials and concentrations. This is the basis of the measurement of pH using the combination of a glass electrode and a calomel, silver/silver chloride or iodide/triiodide reference electrode. Many other indicator electrodes are available, responding in a similar fashion to different cations and anions. There are also electrodes that respond to neutral compounds. Such electrodes are often based on the presence of an immobilized enzyme. For instance, when an enzyme reacts with its substrate, it may release or consume hydrogen ions. The electrode then responds to the resulting change in the concentration of hydrogen ions and, therefore, indirectly to the enzyme substrate. *See* ION-SELECTIVE MEMBRANES AND ELECTRODES.

Because equilibrium potentials can be correlated with chemical equilibrium constants, measurements of the potentials of electrochemical cells are a major source of numerical information about equilibrium constants. Measurements of electrode potentials not involving liquid junctions can yield data accurate to about 10 microvolts, from which salt activities (but not the corresponding ionic activities) can be calculated. Measurements involving liquid junctions (such as pH measurements using reference electrodes with liquid junctions) are much less reliable, because of the uncertainties in the liquid junction potentials, of the order of one or a few millivolts.

The rates of electrode reactions depend strongly (typically, exponentially) on the electrode potential, and most measurements of the rates of electrode reactions are therefore performed under tight control of this potential, often using a specialized instrument called a potentiostat or, in biophysical applications, a voltage clamp. *See* BIOPOTENTIALS AND IONIC CURRENTS; ELECTROCHEMISTRY. Robert de Levie

Bibliography. M. S. Antelman, *The Encyclopedia of Chemical Electrode Potentials*, 1982; A. J. Bard and H. Lund (eds.), *Encyclopedia of the Electrochemistry of the Elements and Compounds*, vols. 1–15, 1973–1984; A. J. Bard, R. Parsons, and J. Jordan (eds.), *Standard Potentials in Aqueous Solutions*, 1985.

Electrodermal response

A transient change in certain electrical properties of the skin, associated with the sweat gland activity and elicited by any stimulus that evokes an arousal or orienting response. Originally termed the psychogalvanic reflex, this phenomenon became known as the galvanic skin response. Electrodermal response (EDR) has replaced galvanic skin response as the collective term.

Mechanism. The skin of a relaxed person has a low electrical conductance (high resistance), and the skin surface is some 40 mV negative with respect to interior tissues. Sweat gland activity changes these electrical properties by increasing skin conductance and by changing the balance of positive and negative ions in the secreted fluid. Human skin is perforated by the ducts of 2–3 million eccrine sweat glands that are distributed most densely on the chest and forehead and, especially, on the palms and soles, where the orifices can be seen (under 10× magnification) as a row of tiny craters lining the dermatoglyphic ridges. On most of the body surface, sweat glands subserve the function of thermoregulation, but on the palms and soles, sweating varies with

psychological arousal. In many persons this "mental sweating" can spread to regions of the head or chest, especially if the person is highly aroused or if the ambient temperature is high.

Significance. Tonic skin conductance varies with psychological arousal, rising sharply when the subject awakens and rising further with activity, mental effort, or especially stress. Phasic skin conductance responses are wavelike increases in skin conductance that begin 1–2 s after stimulus onset and peak within about 5 s. The amplitude of the skin conductance response varies with the subjective impact of the eliciting stimulus, which in turn varies with the intensity of the stimulus, its novelty or unexpectedness for the subject, and its meaning or signal value. Aroused subjects display spontaneous skin conductance responses, generated apparently by mental events or other internal stimuli; their frequency, like the tonic skin conductance level, increases with the level of arousal. Skin potential phenomena are more complicated, and it is doubtful that there is any psychophysiological information present in skin potential level or skin potential responses that cannot be more easily obtained from measurements of skin conductance. Volar sweating, the parent phenomenon, increases the tactile sensitivity, the resistance to abrasion, and the adhesive qualities of these grasping surfaces; thus its function seems to be preparatory to manipulation.

Measurement. A low-impedance voltage source (0.5 V, 500 ohms) is applied to the skin through nonpolarizing electrodes attached typically to the ventral surface of the distal phalange, where sweat glands are most numerous. Skin contact is effected by means of an electrolyte paste containing 0.5% KCl or NaCl. A small (500-ohm) resistor in series with the subject generates a voltage that varies directly with changes in skin conductance. This signal is amplified and recorded on the moving paper chart of a polygraph. *See* LIE DETECTOR.

Applications. Electrodermal responses are measured in studies of emotion and stress, conditioning, habituation, and cognitive processing: that is, when it is desired to assess the differential or changing impact of a series of stimuli. *See* ELECTROENCEPHALOGRAPHY; SYMPATHETIC NERVOUS SYSTEM.

David T. Lykken

Electrodiagnosis

Diagnosis employing equipment that can be used clinically to measure the intrinsic electrical activity of the heart (electrocardiography), the brain (electroencephalography), peripheral nerves (nerve conduction), and skeletal muscles (electromyography). Electrodiagnostic procedures depend upon the difference in electrical potential between the interior and external surfaces of living cells. This potential difference is created by use of cellular metabolic energy to establish and maintain the concentration gradient of ions across the cell membrane. A discharge of the electrical potential (depolarization) by stimulation or injury occurs in a wave proceeding from cell to cell. The ensuing electrical currents are conducted through the tissues to the surface of the body where electrodiagnostic equipment amplifies and records the changes in potential.

In peripheral nerve conduction studies, signals elicited by electrical stimulation of a nerve are measured either over a muscle supplied by the nerve or along the nerve itself. In another electrodiagnostic test, termed electromyography, electrodes are inserted into a muscle to record its electrical activity. A set of diagnostic tests called evoked potentials records the electrical responses of the brain that follow visual or auditory stimuli.

Electrocardiography. The electrocardiogram (ECG) employs electrodes on the extremities and at various points on the chest wall to record voltage changes due to depolarization and repolarization of heart muscle (myocardium). The ECG is most useful for the diagnosis and treatment of disturbances in heart rhythm (arrhythmias) or in the conduction of electrical impulses within the heart. Both inadequate delivery of oxygen to the myocardium (myocardial ischemia) and localized death of myocardium (infarction) produce characteristic changes in the ECG pattern. The ECG also detects hypertrophy, or overgrowth (of the myocardium), and certain abnormalities of blood electrolytes. ECG recordings during a graded exercise test or over a 24-h period of normal activities (Holter monitoring) may detect ischemic changes or arrhythmias that are not present at rest. *See* CARDIAC ELECTROPHYSIOLOGY; INFARCTION.

Electroencephalography. The electroencephalogram (EEG) records waves of electrical activity from the surface of the brain through the placement of electrodes on the scalp in a standardized pattern. Electroencephalography is the most useful procedure for evaluating individuals with epileptic seizures. Because of their greater sensitivity, computerized tomography (CT) scanning and magnetic resonance imaging (MRI) have largely replaced the EEG for the detection and localization of structural brain abnormalities such as tumors or infarcts. The most important use of EEG is for evaluation of brain abnormalities due to metabolic disorders, drug intoxication, and brain infections that can produce changes in the EEG pattern. The total absence of electrical activity (noted by a flat line on the EEG) for 30 min is an important criterion for diagnosing brain death. *See* DEATH; ELECTROENCEPHALOGRAPHY.

Peripheral nerve conduction studies. Nerve conduction velocity is determined by supramaximally stimulating the surface over a nerve and measuring the time it takes for the signal to reach a recording electrode over the innervated muscle (for a motor nerve) or at another spot along the path of a sensory nerve. Nerve damage results in slowed conduction or decreased amplitude of electrical signals along a sensory nerve. Slowed conduction is a common feature of demyelinating familial neuropathies; reduced amplitudes of action potentials along sensory nerves

are observed in diabetic neuropathy. The delivery of rapid, repetitive shocks to a motor nerve results in a swift reduction in the amplitude of the electrical response of the muscle in individuals with myasthenia gravis. *See* MYASTHENIA GRAVIS.

Electromyography. Electromyography (EMG) assesses the function of muscles with measurements of their electrical activity both at rest and during contraction. The electrical activity of muscles is examined by inserting an electrode into them. The amplified signals are displayed on an oscilloscope screen or heard through a loudspeaker. Abnormal EMG patterns can result from damage to the muscles or their nerve supply. Normal skeletal muscles are electrically silent at rest, but damage to or interruption of the nerve supply to a muscle can lead to scattered electrical spikes caused by spontaneous depolarization of single nerve fibers. During their contraction, damaged muscles produce electrical signals that are diminished in amplitude and duration. Electromyography is useful in individuals with amyotrophic lateral sclerosis (Lou Gehrig's disease) and muscular dystrophy. *See* ELECTROMYOGRAPHY.

Evoked potentials. Visual, auditory, or other sensory stimulation produces small transient potentials in the brain. These evoked potentials require hundreds or thousands of rapidly repeated stimuli to generate the evoked potential or response that requires a computer to summate and average the signals. Frequently used for the diagnosis of optic neuritis or multiple sclerosis is a recording over the back of the skull of the visual evoked potential generated by the display of a light signal to one eye at a time. The auditory evoked potential recorded over the scalp in response to high-frequency clicks delivered by an earphone is also useful for the diagnosis of multiple sclerosis. *See* BIOPOTENTIALS AND IONIC CURRENTS.

Simeon Margolis

Bibliography. M. J. Aminoff, *Electrodiagnosis in Clinical Neurology*, 4th ed., Churchill Livingstone, New York, 1999; K. H. Chiappa (ed.), *Evoked Potentials in Clinical Medicine*, 3d ed., Lippincott-Raven, Philadelphia, 1997; A. L. Goldberger, *Clinical Electrocardiography: A Simplified Approach*, 6th ed., Mosby, St. Louis, 1999; A. H. Ropper and R. H. Brown, Electrophysiologic and laboratory aids in the diagnosis of neuromuscular disease, Chap. 45 in *Adams and Victor's Principle of Neurology*, 8th ed., McGraw-Hill, New York, 2005; A. H. Ropper and R. H. Brown, Special techniques for neurologic diagnosis, Chap. 2 in *Adams and Victor's Principle of Neurology*, 8th ed., McGraw-Hill, New York, 2005.

Electrodynamics

The study of the relations between electrical, magnetic, and mechanical phenomena. This includes considerations of the magnetic fields produced by currents, the electromotive forces induced by changing magnetic fields, the forces on currents in magnetic fields, the propagation of electromagnetic waves, and the behavior of charged particles in electric and magnetic fields. Classical electrodynamics deals with fields and charged particles in the manner first systematically described by J. C. Maxwell, whereas quantum electrodynamics applies the principles of quantum mechanics to electrical and magnetic phenomena. Relativistic electrodynamics is concerned with the behavior of charged particles and fields when the velocities of the particles approach that of light. Cosmic electrodynamics is concerned with electromagnetic phenomena occurring on celestial bodies and in space. *See* ELECTROMAGNETISM; ELECTRON MOTION IN VACUUM; MAXWELL'S EQUATIONS; QUANTUM ELECTRODYNAMICS; RADIO-WAVE PROPAGATION; RELATIVISTIC ELECTRODYNAMICS.

John W. Stewart

Electroencephalography

The biomedical technology and science of recording the minute electric currents produced by the brains of human beings and other animals. Electroencephalography (EEG) was discovered by Hans Berger in the late 1920s, and was found to have important clinical significance for the diagnosis of brain disease. The interpretation of EEG records has become a clinical specialty for neurological diagnosis.

Technology and physiology. The recording machine, the electroencephalograph, usually produces a 16-channel ink-written record of brain waves, the electroencephalogram. It is interpreted by an electroencephalographer. The placement of about 20 equally spaced electrodes pasted to the surface of the scalp is in accordance with the standard positions adopted by the International Federation of EEG, and is called the 10/20 system. Electrode positions are carefully measured so that subsequent EEGs from the same person can be compared. About 10 patterns or montages of combinations of electrode pairs are selected for transforming the spatial location from the scalp to the channels which are traced on the EEG pen writer.

The aggregate of synchronized neuronal activity from hundreds of thousands or millions of neurons acting together form the electrical patterns on the surface of the brain. The cellular basis of the EEG depends on the spontaneous fluctuations of postsynaptic membrane potentials between the inside and the outside of the dendritic processes of postsynaptic cells. The brain waves do not derive from the synchronized axonal spike discharges of cerebral cortical nerve cells.

Electrical voltage is transduced from the scalp by differential input amplifiers and amplified about a million times in order to drive the pens for the paper record. The recording usually takes 30–60 min during a relaxed waking state, and also during sleep when possible. Often, activating procedures are used, such as a flickering light stimulator and hyperventilation or overbreathing for about 3 min.

Normal EEG. EEG waves are defined by form and frequency. Various frequencies are given Greek

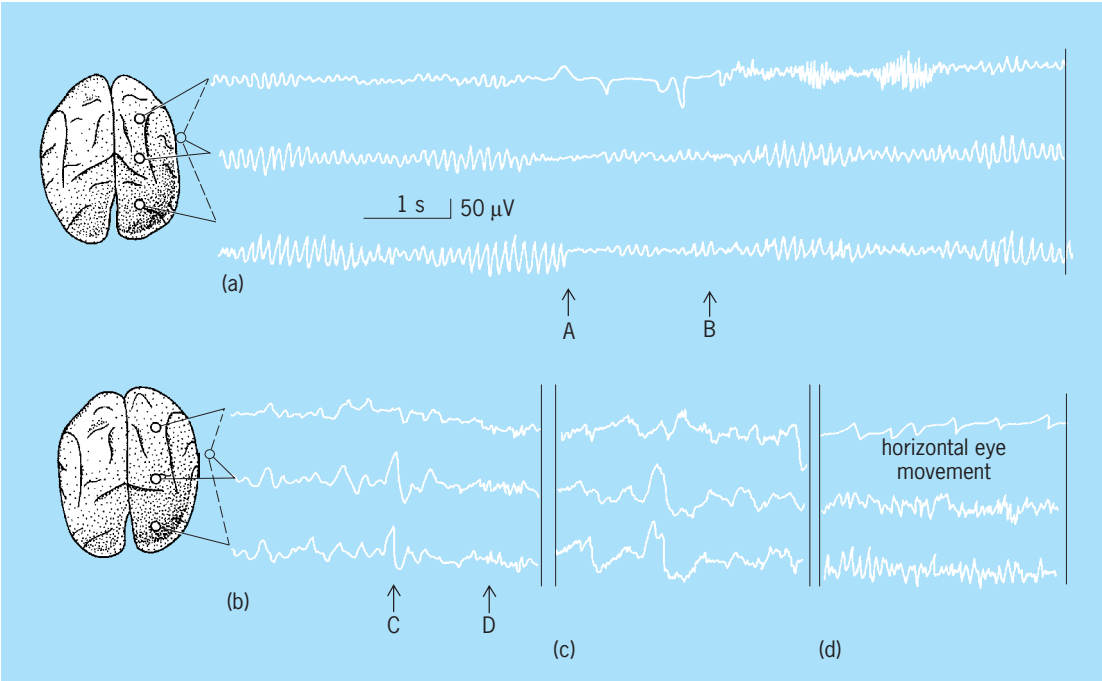

Fig. 1. Normal EEG waves. (*a*) The alpha rhythm in the waking record blocks at eye opening; eyes opened at A and closed at B. (*b*) A light-sleep EEG has theta waves and a sleep hump at C and spindles at D. (*c*) Deep sleep reveals constant, irregular slow delta waves. (*d*) In the dream state, alpha and theta activities are seen, almost like in the waking record.

letter designations. Alpha rhythm is defined as 8–12-Hz sinusoidal rhythmical waves (**Fig. 1***a*). Alpha waves are normally present during the waking and relaxed state and enhanced by closing the eyes. They are suppressed or desynchronized when the eyes are open, or when the individual is emotionally aroused or doing mental work. They may be synchronized by bright light flashes and driven over a wide range of frequencies by repetitive visual stimulation (alpha driving). They are of highest amplitude in the posterior regions of the brain. The alpha rhythm develops with age, reaching maturity by about 12 years, stabilizes, and then declines in frequency and amplitude in old age (over 65).

Beta rhythms are faster, low-voltage sinusoidal waves, usually about 14–30 Hz. They are more prominent in the frontal areas. They are often synchronized and prevalent during sedation with phenobarbital or with the use of tranquilizers and some sedative drugs.

Slower rhythms are theta and delta waves. Theta waves of 4–7 Hz usually replace the alpha rhythm during drowsiness and light sleep (Fig. 1*b*). Delta waves of 0.5–4 Hz are present during deep sleep in normal people of all ages (Fig. 1*c*), and they are the primary waves present in the records of normal infants. Delta waves are almost always pathological in the waking records of adults.

Sleep stages are objectively defined by the use of EEG records. Theta waves are present in drowsiness and light sleep (stage I). Light sleep (stage II) is characterized by negative sharp waves at the vertex (top of head), called sleep humps, and spindle bursts of 13–15-Hz waves in the central regions of the head (over the Rolandic fissures). These are intermixed with theta waves, and seen only during light sleep, sometimes called the hump-and-spindle stage of sleep.

Low-voltage, desynchronized delta waves are present during deep sleep (stage III), and higher-amplitude, slower, and more synchronized delta waves are seen in the deeper sleep stage (stage IV). Stage V sleep is called the dream stage or rapid-eye-movement (REM) stage. That is when dreams take place, accompanied by rapid horizontal eye movements, as if the sleeper were scanning a real scene. The EEG pattern is more like waking, with alpha rhythm present (Fig. 1*d*). The REM state is the one in which it is hardest to wake up the subject, so that it is sometimes called paradoxical sleep. *See* SLEEP AND DREAMING.

Clinical uses of EEG. The EEG reveals functional abnormalities of the brain, whether caused by localized structural lesions, essential paroxysmal states such as epilepsy, or toxic and abnormal metabolic conditions. The three major classes of abnormalities are asymmetries between the hemispheres, slow rhythms, and very sharp waves or spikes. Slow waves represent a depression of cerebral cortical activity or injury in the projection pathways beneath the recording electrodes. Sharp waves or spikes often indicate a hyperexcitable or irritable state of the cortex. During a full epileptic seizure attack, spikes become repetitive and synchronized over the whole surface of the brain. *See* SEIZURE DISORDERS.

The EEG is frequently used for the evaluation of comatose states. The record is slowed in all areas in coma, with delta waves predominating. If the EEG becomes isoelectric or flat for several hours, brain function is not recoverable and the coma may be

considered terminal. "Brain death" is indicated by a flat EEG, recorded at the highest gain with widely spaced electrode positions and the absence of cerebral reflexes and spontaneous respiration. *See* DEATH.

Cerebral evoked potentials. Computer advances in the analysis of EEG signals that are emitted by the brain during sensory stimulation and motor responses have led to the discovery and measurement of electrical waves known as event-related potentials or evoked potentials. These responses are averaged by a computer to enhance the small signals and increase the signal-to-noise ratio, so that they may be graphed and seen.

During the first 10 ms after a click is presented to the ear, electrical responses can be detected from the cochlea, the brainstem, and midbrain structures. Responses to 1000 clicks must be averaged to detect those small, rapid, earliest evoked responses.

This example of a far-field evoked potential is termed the brainstem auditory evoked response (BAER), usually derived from a vertex lead at the top of the scalp compared to a lead at one or the other earlobe as a reference electrode. Each ear is stimulated by clicks at a loudness level set at a certain number of decibels above the auditory threshold. The vertex electrode is far from the origin of the initial electrical activity coming from the cochlear nerve and the brainstem and the midbrain thalamic auditory nuclei, which are far-field or several inches from the active vertex electrode, which is recording field potentials generated by volume conduction rather than by cortical electrical activity from synaptic neural transmission pathways. Neuronal activity from the site of the recording electrode is termed near-field potentials. BAERs are clinically useful in diagnosing peripheral hearing loss due to damage to the structures of the ear, as distinct from hearing disorders involving the neural transmission pathways to the cortex by way of the brainstem and nuclei in the thalamus. Also, evoked potential audiometry has become an active area of clinical research for the evaluation of hearing ability in subjects who cannot respond voluntarily, such as very young children and language-impaired adults.

Somatosensory evoked responses (SERs) are far-field, fast low-voltage brainstem evoked responses derived from stimulating peripheral nerves, usually in the arms or legs, and recorded also by vertex electrodes referred to the earlobes as neutral sites. Like BAERs, the SERs are recorded to trace the somatosensory levels of neural functioning from the spinal cord to the brainstem, thalamus, and cortical areas. Abnormally long latencies or severely reduced amplitudes of the normal waves are diagnostic of neural damage at various levels of conduction of the neural impulses at various synaptic levels.

Fewer signals can be averaged to detect and measure the amplitudes and delays of the larger, slower signals that arrive at the sensory and association cortex 50–150 ms after sensory stimulation. Slower responses have been recorded from association cortex of the brain when psychological factors such as attention, decision, and surprise are indicated. In general, fast potentials relate to the normal physiological functions of direct anatomical pathways, and slower potentials relate to the elaboration of significance and meaning of the responses mediated by nonspecific pathways and association cortex.

Cerebral evoked response patterns have been determined for the major senses under a variety of stimulating conditions. Repetitive electrical stimuli to the skin activate the sensory nerves. Strong stimuli to the eyes such as repeated bright flashes or reversing light and dark squares in a checkerboard pattern, which looks like a checkerboard moving back and forth from left to right and back again, evoke visual responses from the primary and association areas of the visual cortex (VERs). Sinusoidal variations in brightness several times per second are also utilized to evoke visual responses, synchronizing occipital-area brain rhythms, the alpha rhythm, to the stimulating light frequency. The clinical applications of VERs include estimations of visual acuity in nonresponsive subjects, including young children and adults who are not capable of giving verbal or voluntary responses. Visual and other evoked responses are valuable for the diagnosis of early multiple sclerosis, often before any other definite symptoms appear. Optic nerve atrophy or lesions in the visual pathway are detected by these methods.

Evoked responses have also been discovered in the motor systems of the brain preceding voluntary movement by similar averaging methods, using muscle movement outputs as the synchronizing signals to the computer, and tracing voltage changes prior to movements or even speech.

Psychological personality types or psychiatric illnesses may be indicated by modifications of these slow components of the evoked potentials. A negative slow shift, called the contingent negative variation (CNV), is seen when the subject associates a relationship of expectancy between two sensory stimuli when a response is required. If a warning signal is given 1 s before an imperative signal to which the subject must press a button, the contingent negative variation is seen between the two stimuli. The contingent negative variation has been found to be absent or disrupted in some behavioral disorders.

Both normal EEG waves and the fast and slow evoked responses are depressed on the side of a brain injury in one hemisphere compared to the normal hemisphere. Other computer techniques involve the power spectrum analysis of the EEG, such as the Fourier transform, and also cross-correlation between equivalent channels of the two hemispheres.

Evoked potentials are quantified by computer averaging, which enhances the signal-to-noise ratio by the square root of the number of trials. Various evoked potential components are described by the electrical sign, positive or negative, and by their latency in milliseconds to peak amplitude. Factor analytic methods, such as the principal components analysis, are also used by some investigators.

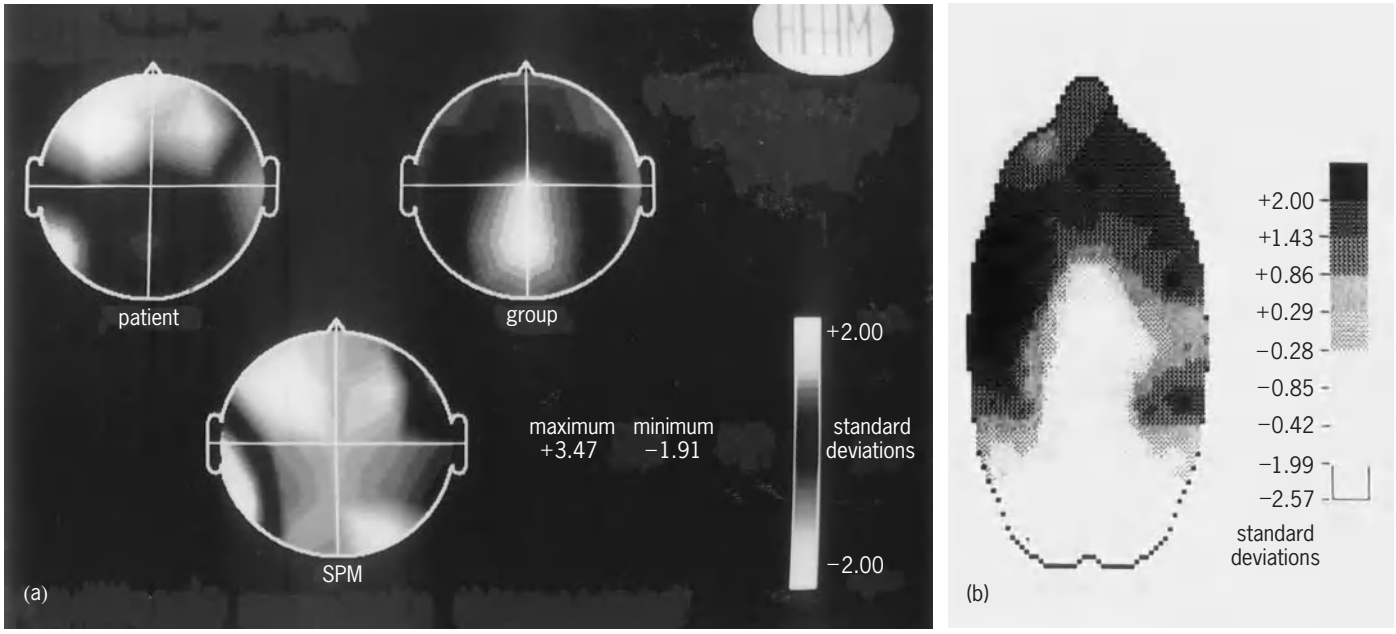

Fig. 2. Two examples of display conventions for topographic mapping. (*a*) Gray-scale technique, depicting amplitude of electrical discharge at particular points in time across the cortex. The maps display comparison of an individual's electroencephalogram with those of a control group, each normalized by subtracting the mean scores and dividing by the standard deviations based on values recorded at each electrode on 3 min of eyes-closed recording. Patient's map shows response to an auditory stimulus at time 156–192 ms after stimulus onset. Group map depicts the electrical response of a comparable control group to the same stimulus at the same point in time. Significance probability map depicts regional differences between patient and control group based upon standard deviation from mean amplitude (*courtesy of F. H. Duffy*). (*b*) Modification of the standard 10–20-electrode placement system, developed by M. Buchsbaum, using 32 electrodes and a four-point linear interpolation. Map shown is a top-down view, with a gray scale depicting brain activity within a specified frequency range.

Topographic mapping. The complexity of evoked potential and EEG analysis makes interpretation difficult in relation to where various components originate and their pattern of spread through time along the neural transmission pathways. Both of these considerations are especially important in determining the way that the brain processes cognitive information as well as sensory signals, that is, how the brain performs at work rather than at rest. In the early 1950s, scientists devised a method of displaying the distribution of EEG wave frequencies over the surface of the head. About 10 oscilloscope displays represented the various head positions and presented the dominant frequency distributions at those spatial locations. It was not until the 1980s, with the development of minicomputers and color graphics screens, that the presentation of topographic information could be analyzed in sophisticated statistical ways for research and clinical purposes by electroencephalographers and neurophysiologists. This method is best known as brain electrical activity mapping (BEAM) and is used in many research investigations of brain activity patterns in learning and language dysfunctions, psychiatric disorders, aging changes and dementia, and studies of normal and impaired child development. Difficult neurological diagnostic problems that do not show anatomical deformities by brain scan methods may often be clarified by these new electrographic proccedures.

To create topographic images of electroencephalography, numeric values are assigned to frequency bands, and these are translated into a gray scale or a rainbow scale. Data are collected from a number of electrodes, usually 20, and an interpolation algorithm is used to develop intermediate values between electrodes (**Fig. 2**).

Topographic images of evoked potentials can be made at millisecond intervals over a period of time and then viewed, like a movie or cartoon, as the computer displays these images on screen at the terminal. The spectral, temporal, and spatial analyses required in electroencephalography and evoked potential interpretation are performed by the computer. When reading the traditional electroencephalogram, the neurophysiologist may take advantage of these imaging techniques to support an interpretation. **Figure 3** provides a simple block diagram of the elements used in generating topographic images of brain electrical activity.

When viewing a topographic map, the clinician is still confronted with the critical question of whether the image (or electroencephalographic recordings associated with the image) represents the neurological activity of a normal, healthy individual or that of one with a pathological condition. A statistical technique called significance probability mapping (SPM) has been developed to aid in this discrimination. Significance probability mapping compares an individual's topographic image to an image based on neurological data of a comparable reference group, and generates a new image that depicts the deviation of the individual from the normal group.

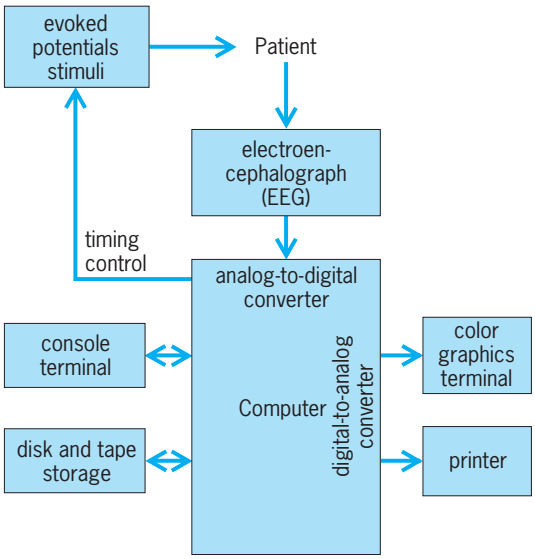

Fig. 3. Primary elements of equipment required for topographic mapping of brain electrical activity.

Such comparisons provide the final complicated step in clinical evaluation of electroencephalographic evoked potential data, and have proven extremely valuable in diagnosing neurologic abnormalities. Additionally, significance probability mapping techniques are valuable in research applications comparing two or more groups of subjects under a number of resting or activated (evoked) conditions. *See* BIOPOTENTIALS AND IONIC CURRENTS; BRAIN; NEUROBIOLOGY.

Jerome Cohen

Bibliography. F. H. Duffy, *Topographic Mapping of Brain Electrical Activity*, 1986; D. Giannitrapani and L. Murri (eds.), *The EEG of Mental Activities*, 1988; R. Johnson, Jr., et al. (eds.), *Current Trends in Event-Related Potential Research*, 1988; E. Niedermeyer, F. H. Lopes da Silva, and L. Fernando (eds.), *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 3d ed., 1993.

Electrokinetic phenomena

Phenomena associated with the movement of charged particles through a continuous medium or with the movement of a continuous medium over a charged surface. The four main electrokinetic phenomena are electrophoresis, electroosmosis, streaming potential, and sedimentation potential, or Dorn effect. These phenomena are related to one another through the zeta potential ζ of the electrical double layer that exists in the neighborhood of the charged surface.

Electrically charged layers. The distribution of electrolyte ions in the neighborhood of a negatively charged surface and the variation of potential ψ with distance from the surface are shown in **Fig. 1**. According to O. Stern, two different layers of ions are associated with the charged surface. The layer of ions immediately adjacent to the surface is called the Stern layer. The ions of this layer are held to the charged surface by a combination of electrostatic attraction and specific adsorption forces, such as short-range van der Waals interactions and chemical bonds. The thickness δ of this layer is assumed to be equal to the ionic radius of the adsorbed ion species. The second layer of ions is the Gouy layer. The boundary between the two layers is the limiting Gouy plane. The ions in the Gouy layer are acted upon only by electrostatic forces and thermal motions of the liquid environment (Brownian motion), and they form a diffuse atmosphere of opposite charge (positive charge in Fig. 1) to the net charge at the limiting Gouy plane. The net charge density of the diffuse ion atmosphere of the Gouy layer decreases exponentially with distance from the limiting Gouy plane. The Gouy layer forms the positive half of an electrical double layer, and the charged surface plus the Stern layer form the negative half. The effective distance of separation $1/\kappa$ between the two halves of the double layer is determined by the concentration of electrolyte (ionic strength). For an electrolyte of univalent ions in water at 25°C (77°F), the relationship for $1/\kappa$ from the Debye-Hückel theory is Eq. (1), where c is the concentration of electrolyte (moles/liter).

$$\frac{1}{\kappa} = \frac{3 \times 10^{-8}}{\sqrt{c}} \tag{1}$$

Variation of potential ψ with the distance x from the charged surface is shown by a solid curve in Fig. 1. Here ψ_0 represents the thermodynamic reversible electrode potential which is independent of the properties of the electrical double layer and dependent only on the activity of the ion which is in reversible electrochemical equilibrium with the substance of the charged surface. The potential ψ

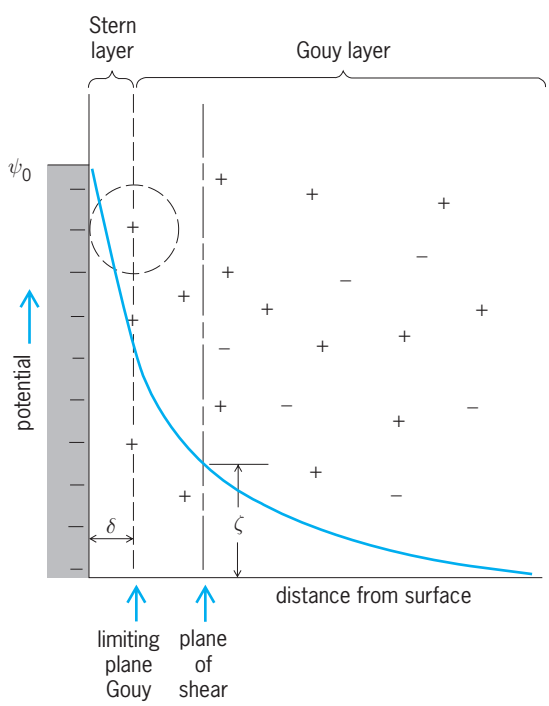

Fig. 1. Electrical double layer.

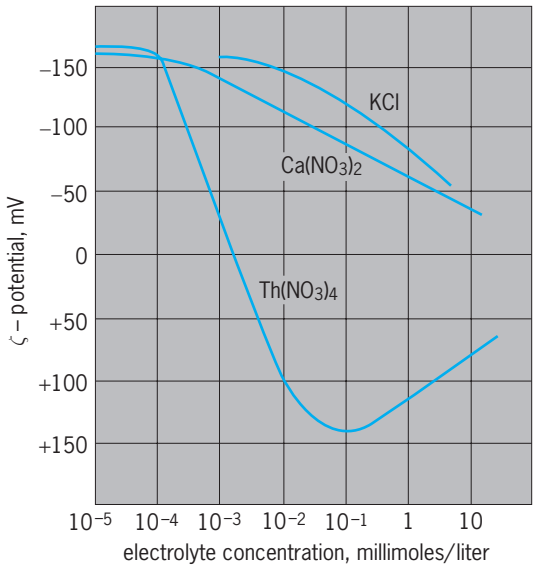

Fig. 2. Effect of electrolyte concentration on the ζ-potential of glass-water interfaces.

decreases linearly with increasing distance x in the region of the Stern layer. In the region of the Gouy layer, ψ decreases exponentially with increasing distance x.

Displacement of charged layers. In the four listed electrokinetic phenomena, a displacement occurs at some plane (plane of shear) between the charged surface and its atmosphere of ions. The position of the slipping plane in Fig. 1 is shown to be located in the Gouy layer. The potential of the plane of shear is the ζ-potential. From the theories of Gouy and Chapman, for spherical particles Eq. (2) holds. Here

$$\zeta = \frac{q}{Da}\,\frac{1}{1+\kappa a} \tag{2}$$

$1/\kappa$ is the effective thickness of the double layer, q the net charge of the particle inside the plane of shear, D the dielectric constant of the liquid, and a the particle's radius at the plane of shear. For flat surfaces

$$\zeta = \frac{4\pi e}{D\kappa} \tag{3}$$

Eq. (3) holds where e is the charge per unit area of surface. Equations (2) and (3) show that ζ-potential is determined by the net charge at the plane of shear and $1/\kappa$, the effective thickness of the ion atmosphere. In turn, ζ-potential controls the rate of transport between the charged surface and the adjacent liquid. The relationship between rate of transport υ_E and ζ-potential which is valid for all four electrokinetic phenomena is Eq. (4), where υ_E is the velocity

$$\upsilon_E = \frac{D\zeta E}{4\pi\eta} \tag{4}$$

of the liquid at a large distance from the charged surface, E is the field strength (V/cm), and η is the viscosity of the liquid. The conditions for validity of Eq. (4) are that the double layer thickness ($1/\kappa$) must be small compared to the radius of curvature of the surface; the substance of the surface must be nonconducting; and the surface conductance of the interface must be negligible. The equations which relate ζ-potential to electroosmotic flow rate and streaming potential may be obtained from Eq. (4) by use of Poiseuille's law for laminar flow through a capillary. For electrophoresis and sedimentation potential (Dorn effect), υ_E is the velocity of the particles. E is the applied field strength for electrophoresis, whereas it is the gradient of potential developed by the sedimentation of charged particles in the Dorn effect.

Electrophoresis, electroosmosis, and streaming potential experiments have been shown to yield identical ζ-potentials for several different interfaces, particularly glass-water and protein-water systems. The sedimentation potential has not been significantly studied.

The effect of electrolytes on the ζ-potential of glass-water interfaces is shown in **Figs. 2** and **3**. As shown in Fig. 2, an increase in electrolyte concentration produces a decrease in ζ-potential, and ions of high charge of opposite sign to that of the surface can completely reverse the sign of the ζ-potential. The explanations for these two effects are given in Fig. 3, where the variation in ψ with distance from the surface is shown for low concentration of electrolyte in curve 1; moderate concentration of electrolyte in curve 2; and charge reversal by adsorption of ions (Th^{4+} on glass) in curve 3. Curves 1 and 2 show that an increase in electrolyte concentration reduces ζ-potential by reducing $1/\kappa$, as indicated by Eqs. (1), (2), and (3). Curve 3 shows that reversal of charge by ion adsorption occurs in the Stern layer and that

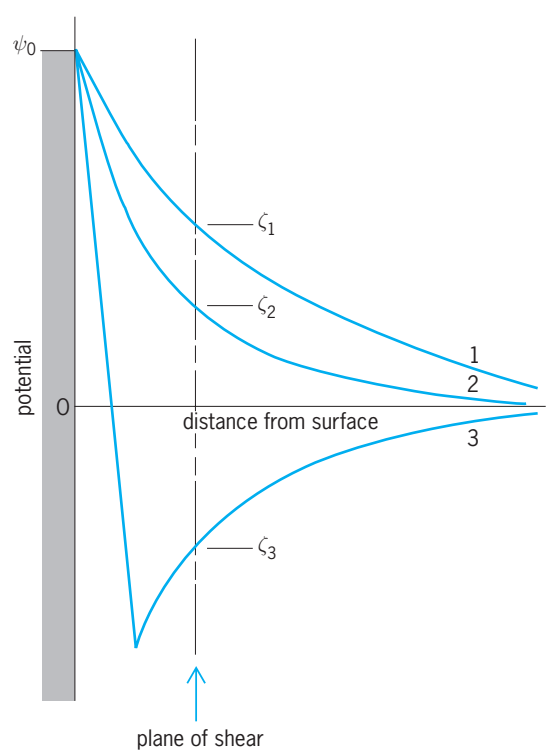

Fig. 3. Reversal of ζ-potential by ion adsorption.

this gives rise to a ζ-potential of opposite sign to the original value. *See* COLLOID; ELECTROPHORESIS; STREAMING POTENTIAL. Quentin Van Winkle

Electroless plating

A chemical reduction process which, once initiated, is autocatalytic. The process is similar to electroplating except that no outside current is needed. The metal ions are reduced by chemical agents in the plating solutions, and deposit on the substrate. Electroless plating is used for coating nonmetallic parts. Decorative electroless plates are usually further coated with electrodeposited nickel and chromium. There are also applications for electroless deposits on metallic substrates, especially when irregularly shaped objects require a uniform coating. Electroless copper is used extensively for printed circuits, which are produced either by coating the nonmetallic substrate with a very thin layer of electroless copper and electroplating to the desired thickness or by using the electroless process only. Electroless iron and cobalt have limited uses. Electroless gold is used for microcircuits and connections to solid-state components. Deeply recessed areas which are difficult to plate can be coated by the electroless process.

Nonmetallic surfaces and some metallic surfaces must be activated before electroless deposition can be initiated. Activation on nonmetals consists of the application of stannous and palladium chloride solutions. Once electroless plating is begun, it will continue to a desired thickness; that is, it is autocatalytic. The process thus differs from a displacement reaction, in which a more noble metal is deposited while a less noble one goes into solution; this ceases when the more noble deposit, if pore-free, covers the less noble substrate.

Electroless copper and gold deposits consist of very small crystals. Electroless nickel deposits are really highly supersaturated alloys containing phosphorus or boron, depending on the reducing agent used. These deposits, which are so fine-grained that they are almost amorphous, can be precipitation-hardened. Electroless nickel deposits are generally harder and more brittle than the electroplated variety. Corrosion resistance of electroless nickel depends, among other factors, on the uniformity of distribution of the phosphorus or boron. Adhesion to nonmetallic substrates is achieved primarily by mechanical means—by plating into pores which are created by selectively etching the substrate. An advantage of electroless plating over plating with current is the more uniform thickness of the surface coating. *See* ELECTROPLATING OF METALS. Rolf Weil

Electroluminescence

A general term for the luminescence excited by the application of an electric field to a system, usually in the solid state. Solid-state electroluminescent systems can be made quite thin, leading to applications in thin-panel area light sources and flat screens to replace cathode-ray tubes for electronic display and image formation. *See* LUMINESCENCE.

Destriau effect. Modern interest in electroluminescence dates from the discovery by G. Destriau in France in 1936 that when a zinc sulfide (ZnS) phosphor powder is suspended in an insulator (oil, plastic, or glass ceramic) and an intense alternating electric field is applied with capacitorlike electrodes, visible light is emitted. The phosphor, prepared from zinc sulfide by addition of a small amount of copper impurity, was later shown to contain particles of a copper sulfide (Cu_2S) phase in addition to copper in its normal role as a luminescence activator in the zinc sulfide lattice. The intensification of the applied electric field by the sharp conductive or semiconductive copper sulfide inhomogeneities is believed to underlie the mechanism of Destriau-type electroluminescence. Minority carriers are ejected from these high-field spots into the low- or moderate-field regions of the phosphor, where they recombine to excite the activator centers. This interpretation is supported by microscopic observations showing that the light is emitted from a few small spots in each phosphor grain. The structure of a Destriau-type electroluminescent cell is shown in **Fig. 1**; the light is observed through the transparent indium–tin oxide electrode.

The brightness L of lamps made with Destriau-type phosphors increases rapidly with the applied voltage V according to the equation below, where A and B

$$L = A \exp \frac{-B}{V^{1/2}}$$

are constants. The brightness increases slightly less than linearly with increase in frequency, until a saturation is reach at about 100,000 Hz. Maximum efficiency is not obtained under the same operating conditions as maximum brightness. Although flat-panel lamps were developed for general lighting application at 120 V with an active layer as thin as 1 mil (25 micrometers), the efficiency of these devices is low compared to conventional light sources; for example, electroluminescence, 3–5 lumens/watt; incandescent lamp, 16 lm/W; and fluorescent lamp,

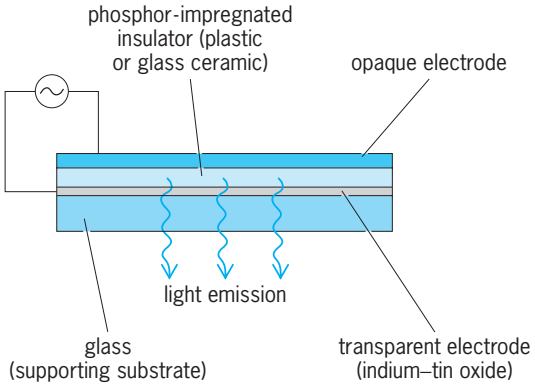

Fig. 1. Structure of powdered-phosphor (Destriau) electroluminescent cell, edge view (not to scale).

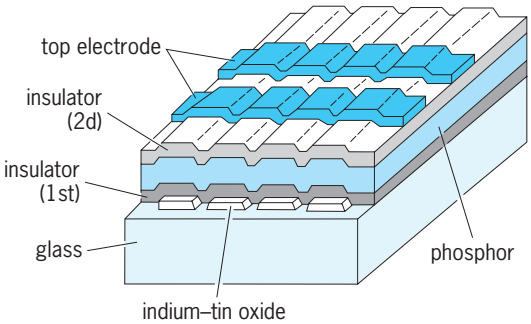

Fig. 2. Structure of a thin-film electroluminescent matrix display. (*After J. Kuwata et al., Feasibility of a 40 × 40 cm² TFEL display with 2000 × 2000 lines operating at 50 W, Society for Information Display Digest, pp. 297–300, 1988*)

80 lm/W. Electroluminescent lamps are therefore not competitive for general illumination, but they have many specialized uses. *See* LIGHT PANEL.

Thin-film electroluminescence. The application of electroluminescence to display and image formation received great impetus from work in the late 1960s and mid-1970s on thin-film electroluminescence (TFEL), giving rise to devices that are different in structure and mechanism from the Destriau conditions. The phosphor in these devices is not a powder but a thin (about 500 nanometers) continuous film prepared by sputtering or vacuum evaporation. The luminescence activators are manganese or rare-earth ions, atomic species with internal electronic transitions that lead to characteristic luminescence. The phosphor film does not contain copper sulfide or any other separate phase, and is sandwiched between two thin (about 200 nm) transparent insulating films also prepared by evaporative means. Conducting electrodes are applied to the outside of each insulating film; one of the electrodes is again a transparent coating of indium–tin oxide on glass, which serves as supporting substrate. If an imaging matrix is desired, both electrodes consist of grids of parallel lines, with the direction of the grid on one insulator (row) orthogonal to the other grid (column). By approximate circuitry the entire matrix can be scanned, applying voltage where desired to a phosphor element that is located between the intersection of a row and column electrode (**Fig. 2**).

A thin-film electroluminescent device is thus a symmetrical structure (conductor–insulator–phosphor–insulator–conductor). It acts like a pure capacitor at low applied voltage; no light is emitted until the voltage reaches a threshold value determined by the dielectric properties of the insulator and phosphor films (**Fig. 3**). Above this threshold the device is no longer purely capacitive, a dissipative current flows, and light emission occurs. The brightness increases very steeply with the applied voltage but is finally saturated. The light output, or average brightness, is roughly proportional to the frequency up to at least 5 kHz, and also depends on the waveform of the applied voltage.

Thin-film electroluminescent devices are believed to operate as follows. As voltage is applied, electrons trapped at the insulator–phosphor interface are released when the electric field reaches a critical value of about 5 MV/in. (2 MV/cm). The electrons are accelerated through the phosphor film, causing impact excitation of the activator centers, which then emit their characteristic luminescence. After the electrons complete their passage through the phosphor, they are retrapped at the second insulator–phosphor interface; on reversal of polarity the action is repeated and the electrons return to the original interface.

The best thin-film electroluminescent phosphor is manganese-activated zinc sulfide, which emits yellow light peaking at 585 nm. A brightness of approximately 10^4 footlamberts (3.5×10^4 candelas/m^2) has been achieved with this phosphor at saturation under 5 kHz excitation, and it exhibits a brightness of 25 footlamberts (85 cd/m^2) and an efficiency of 2 lm/W in a conventional thin-film electroluminescent matrix application at a frequency of 60 Hz. Activation of zinc sulfide and certain alkaline earth sulfides with different rare earths has yielded many other promising electroluminescent phosphors emitting blue, green, red, and white, and making full-color matrix-addressed thin-film electroluminescent displays possible. The light output of thin-film electroluminescent displays has been very reliable, with typically only 10% loss after tens of thousands of hours of operation.

Hysteresis in brightness versus voltage curves offers the possibility of using thin-film electroluminescence in memory devices. In conjunction with

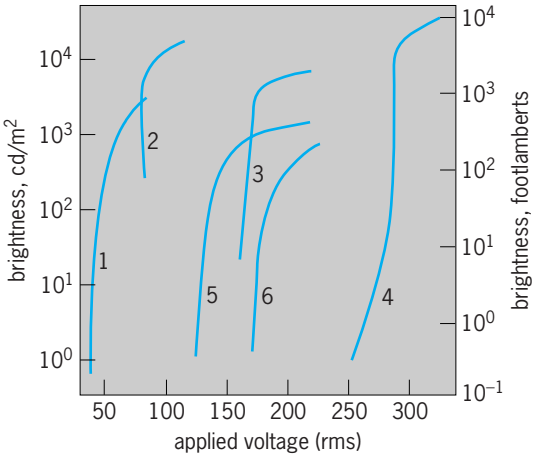

Curve	Emission color	Activator	Insulator	Frequency (sine wave)
1	Yellow	Manganese	Lead titanate	5 kHz
2	Yellow	Manganese	Aluminum oxide	5 kHz
3	Yellow	Manganese	Aluminum oxide	10 kHz
4	Yellow	Manganese	Yttrium oxide	5 kHz
5	Green	Terbium	Yttrium oxide	5 kHz
6	Red	Samarium	Yttrium oxide	5 kHz

Fig. 3. Brightness-voltage characteristics of some zinc sulfide thin-film electroluminescent devices as a function of activator, insulating film, and operating frequency. (*After R. Mach and G. O. Müller, Physical concepts of high-field thin-film electroluminescent devices, Phys. Stat. Sol. (a), 69:11–66, 1982*)

photoconductors, electroluminescence can also be employed in light amplifiers and logic circuits for computers, although semiconductor logic circuits are faster. *See* ELECTRONIC DISPLAY; LIGHT AMPLIFIER.

Injection electroluminescence. Injection electroluminescence results when a semiconductor *pn* junction or a point contact is biased in the forward direction. This type of emission was first observed from silicon carbide (SiC) in England by H. J. Round in 1907. It is the result of radiative recombination of injected minority carriers with majority carriers amaterial. Such emission has been observed in a large number of semiconductors, including silicon (Si), germanium (Ge), diamond, cadmium sulfide (CdS), zinc sulfide (ZnS), zinc selenide (ZnSe), zinc telluride (ZnTe), zinc oxide (ZnO), and many of the so-called III-V compounds, such as gallium phosphide (GaP), gallium arsenide (GaAs), indium phosphide (InSb), boron phosphide (BP), boron nitride (BN), and aluminum nitride (AlN). The wavelength of the emission corresponds to an energy equal, at most, to the forbidden band gap of the material, and hence in most of these materials the wavelength is in the infrared region of the spectrum. In some cases the efficiency is high, approaching one emitted photon. In suitable structures the excitation intensity can become so high that stimulated rather than spontaneous emission predominates, and laser action results, with spectral narrowing and coherent emission. Direct electrical excitation is more convenient than optical excitation, but the beam divergence of injection lasers is normally much greater than that of gas lasers or the optically pumped solid type. *See* JUNCTION DIODE; LASER; SEMICONDUCTOR; SEMICONDUCTOR DIODE.

Other effects. If a *pn* junction is biased in the reverse direction, so as to produce high internal electric fields, other types of emission can occur, but with very low efficiency. The presence of very energetic ("hot") carriers can result in emission at energies greater than the band gap of the material (avalanche emission). The emission in this case may be correlated with small active regions called microplasmas. Light emission may also occur when electrodes of certain metals, such as Al or Ta, are immersed in suitable electrolytes and current is passed between them. In many cases this galvanoluminescence is electroluminescence generated in a thin oxide layer formed on the electrode by electrolytic action. In addition to electroluminescence proper, other interesting effects (usually termed electrophotoluminescence) occur when electric fields are applied to a phosphor which is concurrently, or has been previously, excited by other means. These effects include a decrease or increase in steady-state photoluminescence brightness when the field is applied, or a burst of afterglow emission if the field is applied after the primary photoexcitation is removed. *See* PHOTOLUMINESCENCE.

James H. Schulman; Clifford C. Klick

Bibliography. H. K. Henisch, *Electroluminescence*, 1962; H. F. Ivey, *Electroluminescence and Related Effects*, 1963; A. H. Kitai (ed.), *Solid State Luminescence: Theory, Materials, and Devices*, 1993; J. I. Pankove (ed.), *Electroluminescence*, 1977; S. Shionoya (ed.), *Electroluminescence*, 1989.

Electrolysis

A means of producing chemical changes through reactions at electrodes in contact with an electrolyte by the passage of an electric current. Electrolysis cells, also known as electrochemical cells, generally consist of two electrodes connected to an external source of electricity (a power supply or battery) and immersed in a liquid that can conduct electricity through the movement of ions. Reactions occur at both electrode-solution interfaces because of the flow of electrons. Reduction reactions, where substances add electrons, occur at the electrode called the cathode; oxidation reactions, where species lose electrons, occur at the other electrode, the anode. In the cell shown in the **illustration,** water is reduced at the cathode to produce hydrogen gas and hydroxide ion; chloride ion is oxidized at the anode to generate chlorine gas. Electrodes are typically constructed of metals (such as platinum or steel) or carbon. Electrolytes usually consist of salts dissolved in either water or a nonaqueous solvent, or they are molten salts. *See* ELECTROCHEMISTRY; ELECTRODE; ELECTROLYTE; OXIDATION-REDUCTION.

Applications include industrial synthesis of chemicals, electroplating of metals, metallurgical extraction and refining of metals, metal finishing and electromachining, and the production of electricity in batteries. Corrosion of metals often occurs through electrolytic processes. Electrolytic cells are used in analytical chemistry and in laboratory studies of reaction mechanisms. *See* BATTERY; CORROSION.

Principles. The flow of current, measured in amperes (where 1 A equals the passage of 1 coulomb of charge per second), represents the rate of flow of

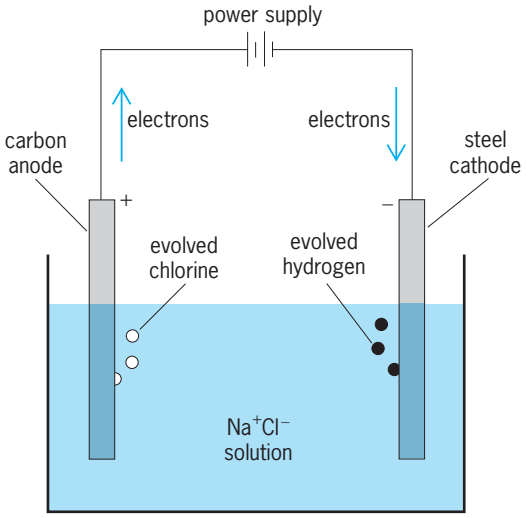

Schematic diagram of an electrolysis cell in which the electrolyte is a solution of sodium chloride.

electrical charge through the electrolysis cell. The quantity of a substance produced or consumed in an electrode reaction is proportional to the quantity of electricity (coulombs) passed during the electrolysis a constant current of 1 A passed for 1 h (equivalent to 3600 coulombs) will produce 0.018656 mole or 1.3228 g of chlorine gas. The quantity of electricity is the integral of the current over the duration of the electrolysis and can be determined with a coulometer. *See* COULOMETER; ELECTROCHEMICAL EQUIVALENT.

In bulk electrolysis, for example, in the production of chemicals, the flow of current produces appreciable changes in the concentration of species in the electrolyte. Cells for bulk electrolysis usually employ relatively large electrodes with the liquid kept in constant motion. For analytical applications or the characterization of chemical systems, the electrolysis occurs only near the surface of the electrodes. Here, unstirred solutions and small electrodes are typically used.

When only one reaction occurs at an electrode, it is said to occur with 100% current efficiency. When two or more reactions occur at the same electrode, for example, when both oxygen and chlorine are evolved at the anode of a cell, the current efficiency of each reaction is given by the fraction of the quantity of electricity consumed for each process. For large-scale electrolytic processes, the power consumption (or energy) efficiency is of interest because it is related to the cost of electrical energy needed to produce a given amount of product. The energy efficiency is the ratio of the ideal theoretical energy for the cell reaction to the actual energy (directly proportional to the cell voltage) consumed in the electrolysis. The cell voltage depends upon a number of factors, including the thermodynamic energy requirements for the electrode reactions, the energy needed to drive the reactions at a given rate, and the resistance of the electrodes and electrolyte. *See* OVERVOLTAGE.

In analytical applications the electrode reaction at only one of the electrodes is of interest, and a third electrode, called a reference electrode, is usually introduced into the cell so that the potential of that electrode can be determined with respect to a known reference potential. The measured potential is also less perturbed by solution resistance effects in a three-electrode cell. *See* REFERENCE ELECTRODE.

Applications. There are many industrial applications for the production of important inorganic chemicals. Chlorine and alkali are produced by the large-scale electrolysis of brine (the chloralkali process) in cells carrying out the same reactions as those shown in the illustration. Other chemicals produced include hydrogen and oxygen (via water electrolysis), chlorates, peroxysulfate, and permanganate. Other processes are carried out with molten salts as solvents, because the electrode reactions of interest would be masked in aqueous solutions by the electrolysis of water to produce hydrogen and oxygen. Electrowinning of aluminum, magnesium, and sodium metals is carried out by molten salt electrolysis. Fluorine is produced by electrolysis of a 2:1 mixture of anhydrous hydrofluoric acid and potassium fluoride.

The major electrolytic processes involving organic compounds are the hydrodimerization of acrylonitrile to produce adiponitrile and the production of tetraethyllead. Many other organic compounds have been studied on the laboratory scale.

Electroplating involves the electrochemical deposition of a thin layer of metal on a conductive substrate, for example, to produce a more attractive or corrosion-resistant surface. Chromium, nickel, tin, copper, zinc, cadmium, lead, silver, gold, and platinum are the most frequently electroplated metals. Metal surfaces can also be electrolytically oxidized (anodized) to form protective oxide layers. This surface-finishing technique is most widely used for aluminum but is also used for titanium, copper, and steel. Electrolytic capacitor production involves anodization of aluminum, tantalum, and niobium. Anodic dissolution of a metal is used in electrochemical machining to produce a desired structure. It is mainly applied in the machining of very hard alloys or in producing complex structures that would be difficult to make by conventional machining methods. *See* ELECTROPLATING OF METALS.

Metals can be purified by electrorefining. Here, the impure metal is used as the anode, which dissolves during the electrolysis. The metal is plated, in purer form, on the cathode. Copper, nickel, cobalt, lead, and tin are all purified by this technique. *See* ELECTROMETALLURGY.

Electroanalysis involves the use of electrolytic processes to identify and quantitate a species. Coulometric methods are based on measuring the quantity of electricity used for a desired process. Voltammetric methods allow characterization of species through an analysis of the effect of potential and electrolysis conditions on the observed currents. Interest in photo-electrolysis is growing; this involves the utilization of radiant (for example, solar) energy to produce electricity or to drive chemical reactions in electrolytic cells. In such cells, which have not yet found practical application, the irradiated electrodes are usually made of semiconductor materials, for example, titanium dioxide, gallium arsenide, or cadmium selenide. Electrolytic cells can also be used to purify waste streams, for example, by the cathodic deposition and recovery of metals and the oxidation of organic pollutants. *See* POLAROGRAPHIC ANALYSIS.

Allen J. Bard

Bibliography. A. J. Bard and H. Lund (eds.), *The Encyclopedia of the Electrochemistry of the Elements*, 1973–1986; A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*, 2d ed., 2000; D. R. Crow, *Principles and Applications of Electrochemistry*, 4th ed., 1994; H. Lund and O. Hammerich (eds.), *Organic Electrochemistry*, 4th ed., 2000; D. Pletcher and F. C. Walsh, *Industrial Electrochemistry*, 2d ed., 1990; M. Schlesinger and M. Paunovic (eds.), *Modern Electroplating*, 4th ed., 2000.

Electrolyte

A material that conducts an electric current when it is fused or dissolved in a solvent, usually water. Electrolytes are composed of positively charged species, called cations, and negatively charged species, called anions. For example, sodium chloride (NaCl) is an electrolyte composed of sodium cations (Na^+) and chlorine anions (Cl^-). The ratio of cations to anions is always such that the substance is electrically neutral. If two wires connected to a light bulb and to a power source are placed in a beaker of water, the light bulb will not glow. If an electrolyte, such as sodium chloride, is dissolved in the water, the light bulb will glow because the solution can now conduct electricity. The amount of electric current that can be carried by an electrolyte solution is proportional to the number of ions dissolved. Thus, the bulb will glow more brightly if the amount of sodium chloride in the solution is increased. *See* ELECTRIC CURRENT; ION.

Hydration. Water is a special solvent because its structure has two different sides. On one side is the oxygen atom, and on the other are two hydrogen atoms. Covalent molecules, such as water, are held together by covalent bonds, which are formed when two atoms share a single pair of electrons. However, when two different atoms form a covalent bond, the sharing of electrons is not always equal. An electron in a covalent bond between two different atoms might spend more time near one atom or the other. In water, the electrons from the O-H bonds spend more time near the oxygen atom than near the hydrogen atoms. As a result, the oxygen has somewhat more negative charge than the hydrogen atoms. This phenomenon is not the same as ionization, where the electron is completely transferred from one atom to the other; so to indicate the difference, the oxygen is said to have a partial negative charge (δ^-) and the hydrogen to have a partial positive charge (δ^+). *See* CHEMICAL BONDING.

Since water has two different sides with opposite partial charges, it is said that water is polar. When ions are dissolved in polar solvents such as water, the solvent molecules are organized around the ion to create the maximum amount of electrostatic attraction. Thus, if a cation is dissolved in water, the water molecules will organize themselves around the ion with the oxygen atoms directed toward the cation, because the partial negative charges on the oxygen atoms in the water molecules are electrostatically attracted to the positive charge on the cation. Likewise, water molecules will organize about anions with the partially positive hydrogen atoms directed toward the negatively charged anion. When water molecules surround dissolved ions, the ions are said to be hydrated. The electrostatic attraction associated with hydration provides the energetic driving force for dissolving ions. Since water is polar, it will also dissolve polar molecules that are not charged, such as ethanol (CH_3CH_2OH), which contains a polar O-H bond. Thus, water can dissolve both ionic substances and polar molecules that are not charged. *See* SOLVENT; WATER.

Strong and weak electrolytes. Any substance that produces ions when dissolved is an electrolyte. These substances include ionic materials composed of simple monatomic ions, such as sodium chloride, or substances composed of polyatomic ions, such as ammonium nitrate (NH_4NO_3). When these substances are dissolved, hydrated ions are generated, as in reactions (1) and (2).

$$NaCl \rightarrow Na^+ + Cl^- \quad (1)$$

$$NH_4NO_3 \rightarrow NH_4{}^+ + NO_3{}^- \quad (2)$$

A special type of electrolyte is an acid, in which the cation is H^+. When acids are dissolved in water, H^+ ions are produced along with an anion, which can be either a monatomic ion, as in hydrochloric acid (HCl), or a polyatomic ion, as in nitric acid (HNO_3) or acetic acid ($HC_2H_3O_2$), as in reactions (3)–(5).

$$HCl \rightarrow H^+ + Cl^- \quad (3)$$

$$HNO_3 \rightarrow H^+ + NO_3{}^- \quad (4)$$

$$HC_2H_3O_2 \rightarrow H^+ + C_2H_3O_2{}^- \quad (5)$$

Electrolytes such as sodium hydroxide (NaOH) that yield the hydroxyl (OH^-) anion when dissolved in water are called bases [reaction (6)]. Some

$$NaOH \rightarrow Na^+ + OH^- \quad (6)$$

molecules that do not contain ions, such as ammonia (NH_3), generate OH^- ions when dissolved in water as in reaction (7); these bases are also electrolytes.

$$NH_3 + H_2O \rightarrow NH_4{}^+ + OH^- \quad (7)$$

Polar covalent molecules, such as ethanol, dissolve in water but do not generate any ions and are called nonelectrolytes.

Many electrolytic substances completely dissociate into ions when they are dissolved in water. In these cases, the reactions shown above would proceed completely to the right, leaving only dissolved ions and no associated, electrically neutral molecules. For example, when sodium chloride is dissolved in water, all of the dissolved material is present as Na^+ and Cl^- ions, with no dissolved NaCl molecules. Substances that completely dissociate are called strong electrolytes, because every molecule dissolved generates ions that contribute to the electrical conductivity. Ionic substances, such as NaCl and NH_4NO_3, are strong electrolytes. Acids and bases that completely dissociate are called strong acids and strong bases, and these substances are also strong electrolytes. Hydrochloric acid (HCl) and HNO_3 are strong acids, so every molecule of HCl or HNO_3 that dissolves generates a H^+ cation and a Cl^- or $NO_3{}^-$ anion, respectively.

Some substances dissolve in water but do not dissociate completely. For example, when acetic acid is dissolved in water, some molecules dissociate to form H^+ and $C_2H_3O_2{}^-$ ions, while others remain

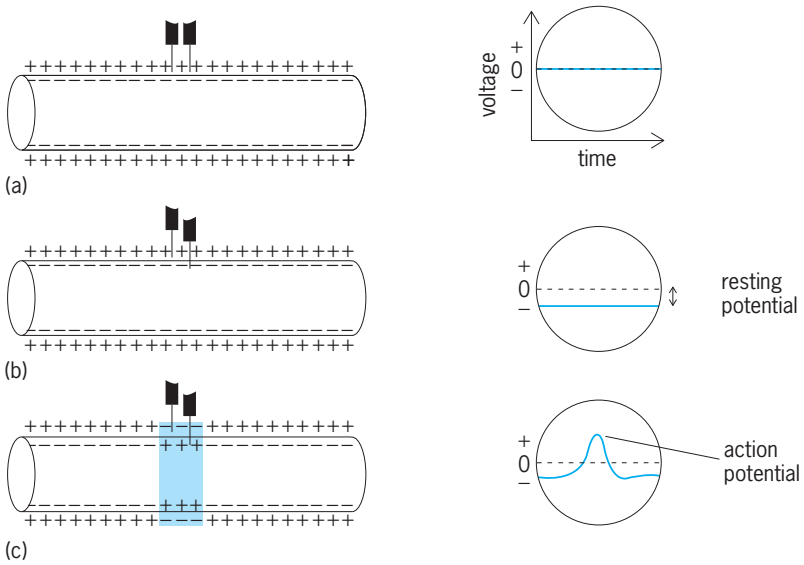

Measurement of the electric potential across a neuronal membrane by using an oscilloscope. (*a*) With both electrodes outside the membrane, no potential is measured. (*b*) With one electrode inside the membrane and the neuron in the resting state, the resting potential is measured. (*c*) When a nerve impulse travels along the neuron, the charge is reversed, and an action potential that is the opposite of the resting potential is measured for a brief period. The shaded region represents the portion where the polarity is reversed. (*After H. Curtis and N. S. Barnes, Biology, 5th ed., Worth, 1989*)

associated as $HC_2H_3O_2$ units, which contain polar O-H bonds and are therefore readily soluble in water. Acids such as acetic acid that do not dissociate completely are called weak acids. Similarly, the reaction of ammonia with water in reaction (7) proceeds to only a small extent. Thus, some molecules of ammonia react with water to generate OH^- and NH_4^+ ions, but many simply remain as NH_3. Since each dissolved molecule of ammonia does not generate an OH^- anion, ammonia is said to be a weak base. The electrical conductivity of an electrolyte solution depends on the number of ions present, so a solution of acetic acid or ammonia will not conduct as well as a solution of an equal amount of a strong electrolyte. Thus, when a light bulb is connected to a solution of acetic acid, it will not glow as brightly as a light bulb connected to a solution of sodium chloride. In fact, when 100 acetic acid molecules are dissolved in water, 99 exist as associated $HC_2H_3O_2$ molecules, while only one dissociates to form H^+ and $C_2H_3O_2^-$ ions. Therefore, in order to get an equal amount of electrical conductivity as a solution of a strong electrolyte, 100 times more acetic acid must be dissolved. For this reason, substances that dissociate somewhat but not completely are called weak electrolytes. Weak acids and weak bases are the most common types of weak electrolytes. *See* IONIC EQUILIBRIUM.

Physiological roles. In biological systems, electrolytes play important roles in regulating kidney function and the retention of water. Electrolytes are also vital for providing the electric current needed for nerve impulses in neurons. When the ratio of cations to anions is different on two sides of a cell membrane, the electric charge on either side of the membrane is different, leading to the establishment of an electric potential. The cell membranes of neurons are impermeable to anions, but they contain special proteins that maintain a higher concentration of cations outside the neuron by either actively transporting sodium (Na^+) and potassium (K^+) ions into and out of the cell or by changing the permeability of the membrane to these cations. Since anions cannot follow the cations out of the neuron, a negative charge builds up inside the neuron and a positive charge builds up outside the neuron, generating an electric potential. When no impulse is being transmitted by the neuron, it is said to be in its resting state and the resulting electric potential is called the resting potential. If two electrodes are placed outside the neuron in its resting state, no potential is measured (**illus.** *a*); however, if one electrode is inside the neuron and the other electrode is outside, the resting potential can be measured (illus. *b*). *See* ELECTRODE; OSMOREGULATORY MECHANISMS.

When a nerve impulse is generated, a localized section of the neuron suddenly becomes permeable to Na^+ cations. Since the concentration of cations is higher outside the neuron, Na^+ ions rapidly flow into the neuron because of attraction to the excess negative charge inside and the concentration gradient. An excess of positive charge is built up inside the neuron, and the electrical charge is reversed. As a result, a new action potential that is opposite of the resting potential is generated (illus. *c*). When the Na^+ ions enter the neuron, the membrane becomes permeable to K^+ cations, which rapidly flow out of the neuron, restoring the resting potential. The action potential therefore persists for only a half millisecond. This process is repeated at adjacent sites in a single direction along the neuron, propagating the nerve impulse. *See* BIOPOTENTIALS AND IONIC CURRENTS; SYNAPTIC TRANSMISSION. H. Holden Thorp

Bibliography. A. J. Bard and L. R. Faulkner, *Electrochemical Methods*: *Fundamentals and Applications*, 2000; R. Chang, *Chemistry*, 1991; H. Curtis and N. S. Barnes, *Biology*, 1989; L. Stryer, *Biochemistry*, 4th ed., 1995; G. Wulfsberg, *Principles of Descriptive Inorganic Chemistry*, 1992.

Electrolytic conductance

The transport of electric charges, under electric potential differences, by particles of atomic or larger size. This phenomenon is distinguished from metallic conductance, which is due to the movement of electrons. The charged particles that carry the electricity are called ions.

Positively charged ions are termed cations; the sodium ion, Na^+, is an example. The negatively charged chloride ion, Cl^-, is typical of anions. The negative charges are identical with those of electrons or integral multiples thereof. The unit positive charges have the same magnitude as those of electrons but are of opposite sign. Colloidal particles, which may have relatively large weights, may be ions, and may carry many positive or negative

charges. Electrolytic conductors may be solids, liquids, or gases. Semiconductors have properties that are intermediate between the metallic conductors and insulators.

Measurement. Conductances are usually reported as specific conductances κ, which are the reciprocals of the resistances of cubes of the materials, 1 cm in each dimension, placed between electrodes 1 cm square, on opposite sides. These units are sometimes called mhos, that is, ohms spelled backward. Conductances of solutions are usually measured by Friedrich Kohlrausch's method, in which a Wheatstone bridge is employed. Such a bridge is shown diagrammatically in **Fig. 1**. The resistances R_3 and R_4 (usually of the same value) form two arms of the bridge. Resistance R_2 is adjustable, and the remaining arm is the cell holding the electrolytic conductor, or as is usually stated, solution of electrolyte. Direct current and the usual galvanometers cannot be used because of an apparent failure of Ohm's law. Passage of direct current produces chemical reactions and a back electromotive force (emf) is generated by the galvanic action of the products. By using an alternating current, the electrochemical reactions occurring when the current is briefly passed in one direction may be reversed when the direction of the current is changed. When a small alternating-current input signal is used, practically all the electric charge passed during each half cycle is stored in the electric double layer, which acts as a capacitor. The electrodes are usually made of platinum and are platinized, that is, coated with finely divided platinum. The surface area, and hence the electrode capacitance, is thereby greatly increased. By making measurements at several frequencies and extrapolating to infinite frequency, the effect of electrode reactions can be eliminated. For less exact measurements, a fixed frequency of 60–1000 Hz is commonly used.

To determine the conductance C, that is, the reciprocal resistance $1/R$ of the cell of Fig. 1, the resistance R_2 of the bridge is adjusted until a minimum of sound is heard in the telephone.

Greater sensitivity may be obtained by electronic amplification of the off-balance signal and by using an oscilloscope to detect the point of balance. When the bridge is in balance, the conductance is given by the relation $C = R_4/R_2R_3$. From this the specific conductance κ may be obtained from the equation $\kappa = KC$, in which K is the cell constant. Occasionally,

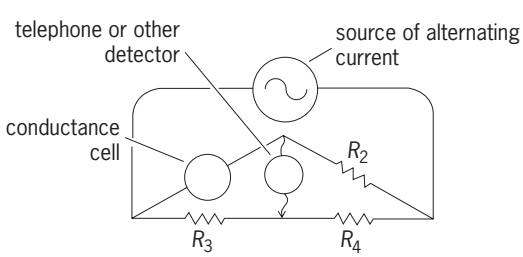

Fig. 1. Wheatstone bridge circuit for the measurement of electrolytic conductance. R_3 and R_4 are fixed resistances; R_2 is a variable resistance.

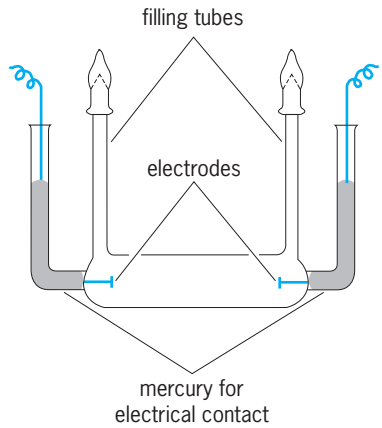

Fig. 2. Diagram of conductance cell.

this constant can be computed from the dimensions of the cell. Usually, however, it is determined by using a solution whose κ value is accurately known from measurements in such a cell, or by comparison with the specific conductance of mercury.

For precision work, care must be taken to avoid errors due to electrical reactances. This has been done in specially designed bridges. A typical, properly designed conductance cell is shown in **Fig. 2**. The cell is filled with solution through the center tubes. Electrical contact is made with the electrodes by platinum wires sealed through the glass wall. These connect the mercury in the outside tubes, which are widely spread to avoid errors due to electrical capacity.

Equivalent conductance. Although many substances and mixtures show electrolytic conductance, the greater part of the research on the subject has been on aqueous solutions of salts, acids, and bases. There have been considerable data accumulated for solutions of such electrolytes in nonaqueous solvents, such as alcohols. The data are usually given in terms of equivalent conductance Λ, which is defined by Eq. (1), in which κ is the

$$\Lambda = \frac{1000\kappa}{c} \tag{1}$$

specific conductance and c is the concentration in equivalents per liter. Values of Λ change with the concentration and, in general, increase as the solutions measured are made more dilute, that is, as c is decreased. A plot of values of the equivalent conductance λ against $\sqrt{c}$ for some typical electrolytes is shown in **Fig. 3**. Svante Arrhenius, who was the first to assume that electrolytic conductance is due to freely moving charged ions, explained the decrease of Λ with increasing c by assuming that the number of ionic carriers gets smaller as the concentration increases, and he computed a degree of dissociation α by formula (2). The term Λ_0 is obtained by

$$\alpha = \frac{\Lambda}{\Lambda_0} \tag{2}$$

determining Λ at a series of low concentrations and extrapolating to a limiting value, termed the equivalent conductance at infinite dilution. Though

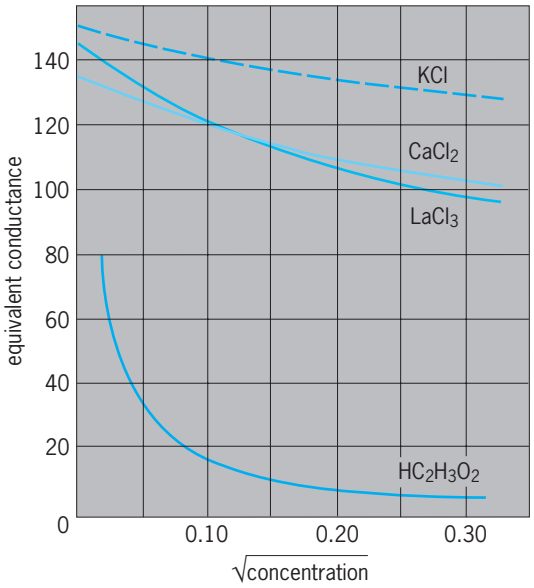

Fig. 3. Equivalent conductance at infinite dilution for some typical electrolytes.

Eq. (2) has been shown by later work to give nearly the right values of α for certain poorly conducting solutions, it is now considered to be much in error for the so-called strong electrolytes. These include most salts, such as potassium chloride, KCl, and sodium sulfate, Na_2SO_4, and inorganic acids and bases, such as hydrochloric acid, HCl, and sodium hydroxide, NaOH. For an electrolyte which yields two types of ion, it can be shown that Eq. (3) holds,

$$\Lambda = F\alpha\,(U^+ + U^-) \tag{3}$$

in which F is the faraday and U^+ and U^- are the mobilities, or speeds, under unit potential difference of the positive and negative ions, respectively. For Eq. (2) to hold, these mobilities must be constant from the equivalent concentration c at which Λ is measured to infinite dilution.

Since the advent of the Debye-Hückel theory of interionic attractions, strong electrolytes have been considered to be completely dissociated, that is, the term α of Eq. (3) is equal to unity for these substances. The decreases observed in the values of the equivalent conductances Λ, with increases in concentration, are assumed to be due to reductions in the values of the ionic mobilities U^+ and U^-. According to the theory of P. Debye and E. Hückel, the ion possesses an ionic atmosphere distributed with radial symmetry around the ion as center. This is due to the fact that interionic attractions and repulsions, together with thermal vibrations, tend to produce a slight preponderance of negative ions around a positive ion, and vice versa. The presence of this atmosphere leads to the lowering of ionic mobilities with increasing ion concentrations. The adaptation of the Debye-Hückel theory for conducting solutions is due to Lars Onsager. His equation for very dilute uni-univalent electrolytes, such as sodium chloride, is shown in Eq. (4), in which θ and σ are given for uni-univalent electrolytes by Eqs. (5) and (6), in which D

$$\Lambda = \Lambda_0 - (\theta\Lambda_0 + \sigma)\sqrt{c} \tag{4}$$

$$\theta = \frac{8.16 \times 10^5}{(DT)^{3/2}} \tag{5}$$

$$\sigma = \frac{8.28}{\eta(DT)^{1/2}} \tag{6}$$

is the dielectric constant at the absolute temperature T and η is the viscosity.

Equation (4) yields accurate values of the data for strong electrolytes up to concentrations of about 0.001 M, above which there are small deviations. Modifications of Eq. (4) for solutions of salts of higher valence types, such as calcium chloride, $CaCl_2$, and lanthanum chloride, $LaCl_3$, are available and have also been found to agree with the data for dilute solutions.

Onsager, in his derivation of Eq. (4), treated the ions as point charges. Later Raymond Fuoss and Onsager extended the theory to include the radii of the ions and also the effects of higher concentrations. The ion sizes obtained from conductance data agree closely with those calculated from activity measurements. *See* ACTIVITY (THERMODYNAMICS).

Equation (4), or its empirical and theoretical extensions, can be used to obtain values of Λ_0 from the data on equivalent conductances of dilute solutions. Some typical figures for limiting equivalent conductances Λ_0 of typical strong electrolytes in aqueous solution at 25°C (77°F) are listed below.

HCl	426.16	NH_4Cl	149.86	$MgCl_2$	129.40
NaOH	247.8	KNO_3	144.96	NaI	126.94
KBr	151.9	$CaCl_2$	135.84	NaCl	126.45
KI	150.38	Na_2SO_4	129.9	LiCl	115.03

Ionic conductances. The fraction of the current carried by a given type of ion is known as the transference number, t, which depends on the relative mobilities of the ions. This relation is shown in Eqs. (7). The ionic mobilities, and therefore the transference numbers, vary somewhat with concentration, but the limiting equivalent conductances Λ_0 for various electrolytes can be accurately represented as the sum of the limiting ionic conductances. Thus, for potassium chloride, $\Lambda_{0,KCl} = \lambda_{0,K^+} + \lambda_{0,Cl^-}$, and the value of λ_{0,Cl^-} is the same whether it is derived from measurements on HCl, NaCl, or KCl solutions. This additive relation is known as Kohlrausch's law of the independent mobility of ions. However, it is necessary to obtain the value of λ_0 for at least one ion constituent independently in order to establish the ion conductances of the other ions. The relation used is shown in Eqs. (8), in which λ_0 is the

$$t^+ = \frac{U^+}{U^+ + U^-} \qquad t^- = \frac{U^-}{U^+ + U^-} \tag{7}$$

$$t_0{}^+\Lambda_0 = \lambda_0{}^+ \quad \text{or} \quad t_0{}^-\Lambda_0 = \lambda_0{}^- \tag{8}$$

limiting equivalent conductance of an electrolyte and t_0^+ and t_0^- are the limiting transference numbers of the positive and negative ion constituents, respectively.

The same value of λ_{0,Cl^-}, within 0.02%, is obtained from precision conductance and transference measurements on solutions of hydrogen, lithium, sodium, and potassium chlorides. Values of the limiting ionic conductance at 25°C (77°F) are given below for some ions.

H^+	349.82	$^1/_2Ba^{2+}$	63.64
OH^-	198	Ag^+	61.92
SO_4^{2-}	79.8	$^1/_2Ca^{2+}$	59.50
Br^-	78.4	$^1/_2Mg^{2+}$	53.06
I^-	76.8	Na^+	50.11
Cl^-	76.34	HCO_3^-	44.48
K^+	73.52	$CH_3CO_2^-$	40.9
NH_4^-	73.4	$CH_2ClCO_2^-$	38.7
NO_3^-	71.44	Li^+	38.69

Nonaqueous systems. In addition to the study of water solutions of electrolytes, considerable study has been given to electrolytes in nonaqueous and mixed solvents. In general, the same principles as those outlined above apply to the interpretation of the results. However, fewer of the electrolytes are completely dissociated, and the degrees of dissociation of the weaker acids and bases are lower. This is due to the fact that, in general, the dielectric constants of nonaqueous solvents are smaller than those of water, so that the attractions between positive and negative ions are greater.

It will be observed that in this article discussion is confined to quite dilute solutions of electrolytes. For concentrated solutions few generalizations of any value can be given.

Molten salts exhibit a wide range of conductivities, depending upon their structures. Salts of alkali and alkaline earth metals usually are largely ionic in character and are highly conductive in the molten state, whereas heavy metal salts may be essentially covalent and exhibit little or no conductivity. Thus the conductivities of liquid arsenic chloride ($AsCl_3$) and bismuth chloride ($BiCl_3$) near their melting points are approximately 10^{-6} and 0.44 ohm^{-1} cm^{-1}, respectively, reflecting the more ionic structure of $BiCl_3$.

If, instead of using quite low potentials in the measurement of electrolytic conductances, voltages of the order of 100,000 are employed, the conductances observed are no longer constant but tend to increase with the potential used. Under these conditions Ohm's law evidently is not valid. This increase of conductance with high potentials is called the Wien effect. This effect is in accord with the interionic attraction theory. When the velocity of the ions becomes sufficiently great, the ion atmospheres do not have time to form to their full extent, so that both the electrophoretic and time of relaxation effects exert less influence on the conductance. However, a large Wien effect is also found for weak acids and bases. It would appear that the high potentials produce, temporarily, additional ionization of these substances. This explanation has been proposed and discussed theoretically by Onsager. If very high frequencies are used in the measurements, an increase in the conductance, termed the Debye-Falkenhagen effect, is observed. This can also be explained by the interionic attraction theory.

Solids. Electrical conduction in solids can range from purely electronic (in metals and semiconductors) to purely ionic (in solids that are electronic insulators but have ions with appreciable mobilities). Examples of purely ionic solid conductors are silver iodide and silver sulfide, in which all of the current is carried by silver ions. In other cases, such as metallic oxides, both ions may serve as charge carriers. Ionic conduction must lead to transport of matter and to chemical reactions at the electrodes, whereas electronic conduction leads only to a transfer of charge. *See* ELECTROCHEMISTRY; ELECTROMOTIVE FORCE (CELLS). Herbert A. Laitinen

Bibliography. J. L. Copeland, *Transport Properties of Ionic Liquids*, 1974; D. Inman and D. G. Lovering (eds.), *Ionic Liquids*, 1981; R. F. Snipes, *Statistical Mechanical Theory of the Electrolytic Transport of Non-Electrolytes*, 1974.

Electromagnet

A soft-iron core that is magnetized by passing a current through a coil of wire wound on the core. Electromagnets are used to lift heavy masses of magnetic material and to attract movable magnetic parts of electric devices, such as solenoids, relays, and clutches.

The difference between cores of an electromagnet and a permanent magnet is in the retentivity of the material used. Permanent magnets, initially magnetized by placing them in a coil through which current is passed, are made of retentive (magnetically "hard") materials which maintain the magnetic properties for a long period of time after being removed from the coil. Electromagnets are meant to be devices in which the magnetism in the cores can be turned on or off. Therefore, the core material is nonretentive (magnetically "soft") material which maintains the magnetic properties only while current flows in the coil. All magnetic materials have some retentivity, called residual magnetism; the difference is one of degree. *See* MAGNETIZATION.

A magnet, when brought near other susceptible material, induces magnetic poles in the susceptible material and so attracts it. A force is developed in the susceptible material that tends to move it in a direction to minimize the reluctance of the flux path of the magnet. The reluctance force may be expressed quantitatively in terms of the rate of change of reluctance with respect to distance. *See* MAGNETISM.

In an engineering sense the word electromagnet does not refer to the electromagnetic forces incidentally set up in all devices in which an electric current

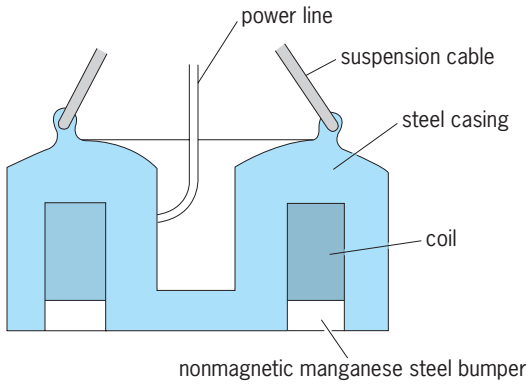

Cross section of circular lifting electromagnet.

exists, but only to those devices in which the current is primarily designed to produce this force, as in solenoids, relay coils, electromagnetic brakes and clutches, and in tractive and lifting or holding magnets and magnetic chucks.

Electromagnets may be divided into two classes: traction magnets, in which the pull is to be exerted over a distance and work is done by reducing the air gap; and lifting or holding magnets, in which the material is initially placed in contact with the magnet. For examples of the first type *see* BRAKE; CLUTCH; RELAY; SOLENOID (ELECTRICITY).

Examples of the latter type are magnetic chucks and circular lifting magnets. The **illustration** shows a typical circular lifting magnet. The outer rim makes up one pole and the inner area is the opposing pole. The coil is wound cylindrically around the center pole. Manganese steel, used as a protective cover plate for the coil, is nonmagnetic and thus forces the flux through the magnetic member being lifted.

The mechanical force between two parallel surfaces is given by Maxwell's equation, shown below,

$$F = B^2A/2\mu_0 \text{ (newtons)}$$

where B is the flux density (in webers/m^2), A is the cross-sectional area (in m^2) through which the flux passes, and μ_0 is the permeability of free space ($4\pi \times 10^{-7}$ henry/m). When two poles are active, the force produced by each is calculated to find the total force. An interesting result of this relation is that the force is not simply the result of the total flux (BA) but also of the flux density. Thus if the same flux can be forced through one-half the cross-sectional area, the net pull will be doubled. In practice it is difficult, if not impossible, to calculate the actual lifting capacity of the magnet by using Maxwell's equation since the capacity varies with the shape and kind of material lifted, how the material is stacked, and other factors. Therefore, lifting magnets are usually rated on their all-day average lifting capacity.

Since currents are large and the circuit is highly inductive, control of a lifting magnet is a problem. If the line switch were simply opened, a destructive arc would result due to the coil's inductance not allowing the current to change instantaneously to a zero value. Therefore, the controller employed with a lifting magnet usually does the following things automatically: (1) reduces magnet current after an initially high value to a lower holding value to reduce heating of the magnet, (2) introduces a shunt discharge resistor across the magnet coil before allowing the line to be opened when the operator turns the magnet off, and (3) causes a reduced current of reverse polarity to flow in the magnet coil for a short time after the operator turns the switch off. Thus the residual magnetism is canceled and scraps and small chunks that might have continued clinging to the magnet are released. Jerome Meisel

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; G. R. Jones, M. A. Laughton, and M. G. Say (eds.), *The Electrical Engineers' Reference Book*, 15th ed., 1993; E. M. Purcell, *Electricity and Magnetism*, 2d ed., 1984.

Electromagnetic compatibility

The situation in which electrical and electronic devices and systems work as intended, both within themselves and in their electromagnetic environment.

Electromagnetic interference (EMI) is said to exist when unwanted voltages or currents are present so that they adversely affect the performance of a device or system. Such voltages or currents may reach the victim circuit or device by conduction or by nonionizing radiation. In all cases, electromagnetic interference arises because of a combination of three factors: a source, a transmission path, and a response, at least one of which is unplanned. Electromagnetic interference control refers to the process of making design changes or adjustments of signal or noise levels in order to achieve electromagnetic compatibility (EMC). Examples of electromagnetic interference often experienced by the public are herringbone patterns on the television screen and buzzing and clicking sounds in the radio.

Sources of interference. Designers usually do not intend that their devices be sources of interference. However, what is a desired signal in one path may be an undesired signal, that is, noise, in a path into which it is inadvertently coupled. Interference may be unintentional, or nonfunctional, such as an arc discharge, radiation from a lightning stroke, a corona discharge from high-tension power lines, or a noise caused by a sudden change in current flow in a conductor. Functional interference, while not usually produced for interference purposes, often includes sine waves, computer clock pulses, speech or video waves, or pulses forming data trains. An example of functional interference is signal leakage from cable television systems. Fluorescent lamps, commutators, automotive ignition systems, and industrial, scientific, and medical equipment such as diathermy machines all constitute sources of interference, as does the electromagnetic pulse (EMP) which accompanies a nuclear detonation. Short-wave radio

listeners often experience interference from high-frequency radar from Russia. Interference also exists across national borders because of incomplete international agreement on the use of the radio-frequency spectrum. *See* ELECTRICAL INTERFERENCE; ELECTROMAGNETIC PULSE (EMP); RADIO SPECTRUM ALLOCATION.

Interference coupling. Electromagnetic interference always starts with a flow of current through a conductor (which may include a conducting gas), and always appears at the victim in the form of a current or a voltage. The coupling path, however, may be a conduction or a radiation path. Thus the actual paths include common wiring, capacitance between devices, mutual inductance between adjacent wiring, nonionizing radiation, or wires in an electromagnetic field. Such coupling is aided by the fact that all conductors, whether wires, printed circuit board traces, or the conductors within an integrated circuit, exhibit both resistance and inductance. Likewise, between wires, mutual inductance and capacitance exist, resulting in a variety of possible coupling paths. *See* CAPACITANCE; COUPLED CIRCUITS; INDUCTANCE.

Grounds and bonds. Grounding is the establishment of an electrically conductive path between an electrical or electronic element of a system and a reference point or plane known as the ground. Grounding thus is a circuit concept, whereas bonding refers to the physical implementation of that concept. Grounding may also refer to an electrical connection made to the earth.

Grounds for currents whose wavelength is long compared with circuit dimensions are best made directly to a single point in a system. For currents whose wavelength is comparable to, or shorter than, circuit dimensions, however, the multipoint ground must be used to avoid the situation in which the ground lead might be a significant fraction of a quarter-wavelength, and thus not serve as a low-impedance conductor. *See* ELECTRONIC EQUIPMENT GROUNDING; GROUNDING.

Shields. The purpose of shielding is to confine radiated energy to a specific region, or to prevent radiated energy from entering a specific region. The most effective shield is a solid metallic enclosure, made of a permeable metal (for example, iron or steel) if frequencies below 100 kHz are to be shielded, or of any metal if higher frequencies are to be shielded. However, the solid shield does not permit light, air, water, or other substances to be passed through it, so shields with holes, including screens, braids, and honeycomb arrangements, as well as conductive glass may be needed. The widespread use of plastic enclosures has made thin film shields vital in achieving the needed shielding effectiveness in the use of such enclosures. *See* ELECTRICAL SHIELDING.

Filters. An electrical filter offers relatively little opposition to the passage of certain frequencies or direct current (dc) while blocking the passage of other frequencies. Accordingly, filters play a significant role in reducing conducted interference to the extent that such interference has a spectral content different from that of the desired signals.

A filter may be either reflective or lossy. Reflective filters present an impedance mismatch to unwanted frequencies, thereby returning them to the input, whereas lossy filters absorb unwanted frequencies. A filter may be designed on a time-domain basis as well as on a frequency-domain basis. *See* ELECTRIC FILTER; IMPEDANCE MATCHING.

Mathematical models. The complexity of electronic circuitry no longer allows the use of simple trial-and-error methods in achieving an electromagnetically compatible system design. System behavior can be expressed in terms of equations that describe source outputs, transmission or coupling characteristics, and susceptor responses. Such models have been incorporated into system analysis programs which not only predict system compatibility but also can be used to evaluate the effect of parameter changes on overall system performance, as may be needed in analyzing a proposed subsystem specification waiver, or in generating interface specifications for use within an overall system design. *See* SIMULATION.

Standards. Numerous standards have been developed for military as well as for industrial and consumer products. For military equipment, MIL-STD-461B, *EMI Characteristics Requirements for Equipment*, defines the emission and susceptibility limits applicable to units under various test conditions; MIL-STD-462, *Measurement of EMI Characteristics*, describes how the measurements are to be made.

Such organizations as the Institute of Electrical and Electronics Engineers (IEEE), the Society of Automotive engineers (SAE), the Electronic Industries Association (EIA), and the American National Standards Institute (ANSI) are active in promulgating standards for electronic equipment, including electromagnetic compatability aspects.

The Federal Communications Commission's (FCC) *Rules and Regulations*, Part 15, as amended by FCC Docket 20780, establishes emission limits on "computing devices," defined as any electronic devices or systems that use digital techniques. A class A computing device is destined for commercial, industrial, or business use. It must meet certain radiation limits (for frequencies of 30 to 100 MHz) at a 96-ft (30-m) distance, and conduction limits (for frequencies of 0.45 to 30 MHz) with respect to its power leads. A class B computing device is destined for residential use. This device must meet certain radiation limits at a 10-ft (3-m) distance. Both its radiation limits and its conduction limits are more stringent than those of a class A device. *See* DIGITAL COMPUTER; MICROCOMPUTER.

Mutual effects. Digital systems, such as computers, tend to interfere with analog systems, such as voice and video communications, more readily than analog systems interfere with digital systems. Therefore, data streams to be transmitted over analog voice circuits are converted to a quasianalog tone form first. Computer clocks also may have to be shielded and

their output circuits may have to be filtered to prevent interference to communication equipment. In addition, personal computers must be connected to television receivers so that the video output of the personal computer does not reach the television receiving antenna, which then would radiate such signals. *See* ELECTRICAL NOISE. Bernhard E. Keiser

Bibliography. J. Goedbloed, *Electromagnetic Compatibility*, 1993; B. E. Keiser, *Principles of Electromagnetic Compatibility,* 3d ed., 1987; C. R. Paul, *Introduction to Electromagnetic Compatibility*, 1992; D. A. Weston, *Electromagnetic Compatibility Principles and Applications*, 1991.

Electromagnetic field

A changing magnetic field always produces an electric field, and conversely, a changing electric field always produces a magnetic field. This interaction of electric and magnetic forces gives rise to a condition in space known as an electromagnetic field. The characteristics of an electromagnetic field are expressed mathematically by Maxwell's equations. *See* ELECTRIC FIELD; ELECTROMAGNETIC RADIATION; ELECTROMAGNETIC WAVE; MAXWELL'S EQUATIONS. Jesse W. Beams

Electromagnetic induction

The production of an electromotive force either by motion of a conductor through a magnetic field in such a manner as to cut across the magnetic flux or by a change in the magnetic flux that threads a conductor.

Motional electromotive force. A charge moving perpendicular to a magnetic field experiences a force that is perpendicular to both the direction of the field and the direction of motion of the charge. In any metallic conductor, there are free electrons, electrons that have been temporarily detached from their parent atoms.

If a conducting bar (**Fig. 1**) moves through a magnetic field, each free electron experiences a force due to its motion through the field. If the direction of the motion is such that a component of the force on the electrons is parallel to the conductor, the electrons will move along the conductor. The electrons will move until the forces due to the motion of the conductor through the magnetic field are balanced by electrostatic forces that arise because electrons collect at one end of the conductor, leaving a deficit of electrons at the other. There is thus an electric field along the rod, and hence a potential difference between the ends of the rod while the motion continues. As soon as the motion stops, the electrostatic forces will cause the electrons to return to their normal distribution.

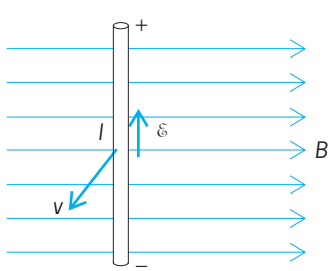

Fig. 1. Flux density *B*, motion *v*, and induced emf ℰ when a conductor of length *l* moves in a uniform field. *(After R. L. Weber, M. W. White, and K. V. Manning, Physics for Science and Engineering, McGraw-Hill, 1957)*

From the definition of magnetic induction (flux density) B, the force on a charge q due to the motion of the charge through a magnetic field is given by Eq. (1), where the force F is at right angles to a

$$F = Bqv \sin\theta \tag{1}$$

plane determined by the direction of the field, and the component $v \sin\theta$ of the velocity is perpendicular to the field. When B is in webers/m^2, q is in coulombs, and v is in m/s, the force is in newtons.

The electric field intensity E due to this force is given in magnitude and direction by the force per unit positive charge. The electric field intensity is equal to the negative of the potential gradient along the rod. In motional electromotive force (emf), the charge being considered is negative. Thus, Eqs. (2)

$$E = \frac{F}{-q} = -Bv \sin\theta = -\frac{\mathscr{E}}{l} \qquad \mathscr{E} = Blv\sin\theta \tag{2}$$

hold. Here l is the length of the conductor in a direction perpendicular to the field, and $v \sin\theta$ is the component of the velocity that is perpendicular to the field. If B is in webers/m^2, l is in meters, and v is in m/s, the emf $\mathscr{E}$ is in volts.

This emf exists in the conductor as it moves through the field whether or not there is a closed circuit. A current would not be set up unless there were a closed circuit, and then only if the rest of the circuit does not move through the field in exactly the same manner as the rod. For example, if the rod slides along stationary tracks that are connected together, there will be a current in the closed circuit. However, if the two ends of the rod were connected by a wire that moved through the field with the rod, there would be an emf induced in the wire that would be equal to that in the rod and opposite in sense in the circuit. Therefore, the net emf in the circuit would be zero, and there would be no current.

EMF due to change of flux. When a coil is in a magnetic field, there will be a flux Φ threading the coil the magnitude of which will depend upon the area of the coil and its orientation in the field. The flux is given by $\Phi = BA \cos\theta$, where A is the area of the coil and θ is the angle between the normal to the plane of the coil and the magnetic field. Whenever there is a change in the flux threading the coil, there will be an induced emf in the coil while the change is taking place. The change in flux may be caused by

a change in the magnetic induction of the field or by a motion of the coil. The magnitude of the induced emf, Eq. (3), depends upon the number of turns of

$$\mathscr{E} = -N\frac{d\Phi}{dt} \tag{3}$$

the coil N and upon the rate of change of flux. The negative sign in Eq. (3) refers to the direction of the emf in the coil; that is, it is always in such a direction as to oppose the change that causes it, as required by Lenz's law. If the change is an increase in flux, the emf would be in a direction to oppose the increase by causing a flux in a direction opposite to that of the increasing flux; if the flux is decreasing, the emf is in such a direction as to oppose the decrease, that is, to produce a flux that is in the same direction as the decreasing flux. *See* FARADAY'S LAW OF INDUCTION; LENZ'S LAW.

Consider the case of a flat coil of area A rotating with uniform angular velocity ω about an axis perpendicular to a uniform magnetic field of flux density B. For any position of the coil, the flux threading the coil is $\Phi = BA\cos\theta = BA\cos\omega t$, where the zero of time is taken when θ is zero and the normal to the plane of the coil is parallel to the field. Then the emf induced as the coil rotates is given by Eq. (4).

$$\mathscr{E} = -N\frac{d\Phi}{dt} = -NBA\frac{d(\cos\theta)}{dt} = NBA\omega\ \sin\omega t \tag{4}$$

The induced emf is sinusoidal, varying from zero when the plane of the coil is perpendicular to the field to a maximum value when the plane of the coil is parallel to the field.

Self-induction. If the flux threading a coil is produced by a current in the coil, any change in that current will cause a change in flux, and thus there will be an induced emf while the current is changing. This process is called self-induction. The emf of self-induction is proportional to the rate of change of current. The ratio of the emf of induction to the rate of change of current in the coil is called the self-inductance of the coil.

Mutual induction. The process by which an emf is induced in one circuit by a change of current in a neighboring circuit is called mutual induction. Flux produced by a current in a circuit A (**Fig. 2**) threads or links circuit B. When there is a change of current in circuit A, there is a change in the flux linking coil B, and an emf is induced in circuit B while the change is taking place. Transformers operate on the principle of mutual induction. *See* TRANSFORMER.

The mutual inductance of two circuits is defined as the ratio of the emf induced in one circuit B to the rate of change of current in the other circuit A. For a detailed discussion of self- and mutual inductance *See* INDUCTANCE.

Coupling coefficient. This refers to the fraction of the flux of one circuit that threads the second circuit. If two coils A and B having turns N_A and N_B, respectively, are so related that all the flux of either threads both coils, the respective self-inductances are given by Eqs. (5) and the mutual inductance of the pair is

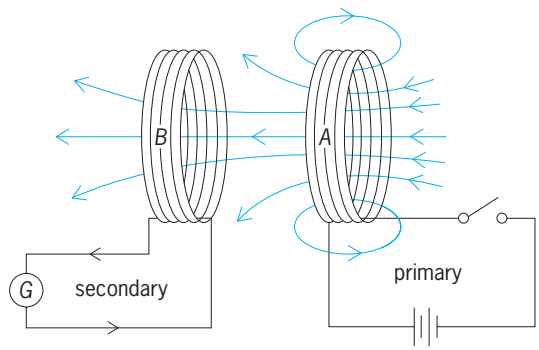

Fig. 2. Mutual induction. An emf is induced in the secondary when the current changes in the primary. (*After R. L. Weber, M. W. White, and K. V. Manning, Physics for Science and Engineering, McGraw-Hill, 1957*)

given by Eq. (6). Then Eqs. (7) hold.

$$L_A = \frac{N_A\Phi_A}{I_A} \qquad L_B = \frac{N_B\Phi_B}{I_B} \tag{5}$$

$$M = \frac{N_A\Phi_B}{I_B} = \frac{N_B\Phi_A}{I_A} \tag{6}$$

$$M^2 = \frac{N_AN_B\Phi_A\Phi_B}{I_AI_B} = \frac{N_A\Phi_A}{I_A}\frac{N_B\Phi_B}{I_B} = L_AL_B \qquad M = \sqrt{L_AL_B} \tag{7}$$

In general, not all the flux from one circuit threads the second. The fraction of the flux from circuit A that threads circuit B depends upon the distance between the two circuits, their orientation with respect to each other, and the presence of a ferromagnetic material in the neighborhood, either as a core or as a shield. It follows that for the general case that Eq. (8) holds.

$$M \le \sqrt{L_AL_B} \tag{8}$$

The ratio of the mutual inductance of the pair to the square root of the product of the individual self-inductances is called the coefficient of coupling K, given by Eq. (9). The coupling coefficient has a maximum value of unity if all the flux threads both circuits, zero if none of the flux from one circuit threads the other. For all conditions, K has a value between 0 and 1.

$$K = \frac{M}{\sqrt{L_AL_B}} \tag{9}$$

Applications. The phenomenon of electromagnetic induction has a great many important applications in modern technology. For example, *see* COUPLED CIRCUITS; GENERATOR; INDUCTION HEATING; MICROPHONE; MOTOR; SERVOMECHANISM.

Kenneth V. Manning

Bibliography. W. J. Duffun, *Electricity and Magnetism*, 4th ed., 1990; D. Halliday and R. Resnick, *Physics*, pt. 2, 4th ed., 1992; A. Mottershead, *Introduction to Electricity and Electronics*, 2000; E. M. Purcell, *Electricity and Magnetism*, 2d ed., 1984.

Electromagnetic pulse (EMP)

A transient electromagnetic signal produced by a nuclear explosion in or above the Earth's atmosphere. Though not considered dangerous to people, the electromagnetic pulse (EMP) is a potential threat to many electronic systems.

Discovery. The existence of a nuclear-generated EMP has been known for many years. Originally predicted by scientists involved with the early development of nuclear weapons, it was not considered to be a serious threat to people or equipment. Then in the early 1960s some of the high-altitude nuclear tests conducted in the Pacific led to some strange occurrences many miles from ground zero. In Hawaii, for example, some 800 mi (1300 km) from the Johnston Island test, EMP was credited with setting off burglar alarms and turning off street lights. In later tests that were conducted in Nevada, significant EMP-induced signals were coupled to cables. These experimental results gave credibility to the potential military use of EMP and encouraged investigators to more accurately describe its origin, electromagnetic characteristics, and coupling to systems, and to develop an affordable method for system protection.

Initial nuclear radiation. In a typical nuclear detonation, parts of the shell casing and other materials are rapidly reduced to a very hot, compressed gas, which upon expansion gives rise to enormous amounts of mechanical and thermal energy. At the same time the nuclear reactions release tremendous amounts of energy as initial nuclear radiation (INR). This INR is in the form of rapidly moving neutrons and high-energy electromagnetic radiation, called x-rays and prompt gamma rays. About a minute after detonation, the radioactive decay of the fission products gives rise to additional gamma rays and electrons (or beta particles), known as residual nuclear radiation (RNR). The distribution of the total explosive energy of a hypothetical fission detonation in the atmosphere below an altitude of 6 mi (10 km) is 50% blast, 35% thermal, 10% RNR, 5% INR. At higher altitudes where the air is less dense, the thermal energy increases and the blast energy decreases proportionately. *See* BETA PARTICLES; GAMMA RAYS; NUCLEAR FISSION; NUCLEAR REACTION; RADIOACTIVITY.

EMP is associated with the INR output, which is a small percentage of the total explosive energy. Nevertheless, EMP is still capable of transferring something of the order of 0.1–0.9 joule/m^2 (0.007–0.06 ft-lbf/ft^2) onto a collector, more than enough to cause upset or damage to normal semiconductor devices.

Early research first developed the physics of high-altitude EMP. Then in 1978 a consistent explanation was set forth of how EMP is related to the total gamma source strength regardless of the detonation height above ground. **Figure 1** identifies six contributions to the gamma source strength for a hypothetical 1-megaton surface burst. As the detonation height is elevated, gamma source contributions from neutrons interacting with the ground and air decrease; for a high-altitude detonation (higher than 37 mi or 60 km) the gamma source strength becomes essentially the prompt gammas from the nuclear burst. The gamma source strength wave shape then approaches a smooth curve, approximating a double exponential, and since EMP is generated from the gamma source strength as described below, it too approaches a double exponential wave shape.

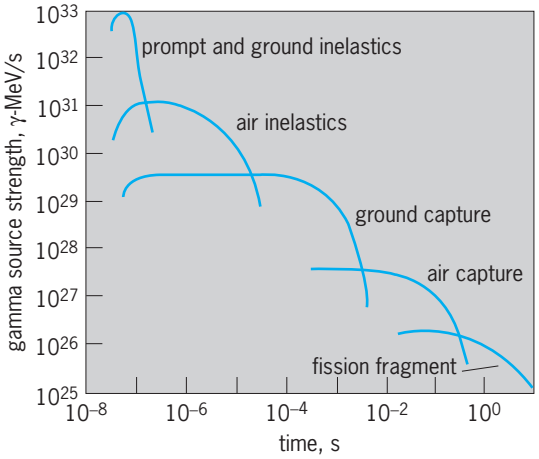

Fig. 1. Total gamma source strength versus time for nominal 1-megaton surface burst. Both horizontal and vertical scales are logarithmic. (*After C. L. Longmire, On the electromagnetic pulse produced by nuclear explosions, IEEE Trans. Antennas Propag., AP-26(1):3–13, 1978*)

EMP generation in a high-altitude burst. As the prompt gammas move away from a high-altitude nuclear detonation (**Fig. 2**), those gamma rays moving toward the Earth penetrate a more dense region of the atmosphere called the source or deposition region. In this 6-mi thick (10-km) region, approximately 15–21 mi (25–35 km) above the Earth, the highly energetic gamma rays interact with the air molecules to form Compton electrons (with energies starting at 1 MeV) and less energetic gamma rays, which then proceed in the same general direction as the original gamma rays. The fast Compton electrons eventually slow down by stripping other electrons from air molecules to form secondary electron-ion pairs. (Though these secondary electrons and ions

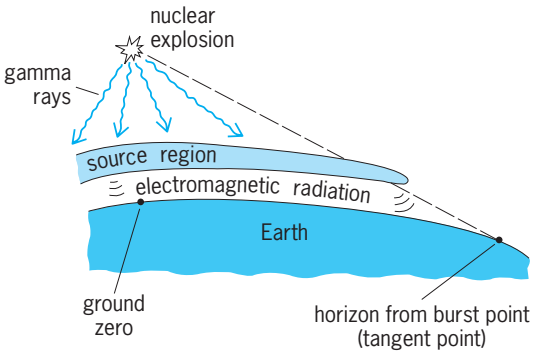

Fig. 2. Schematic representation of the EMP in a high-altitude burst. The extent of the source region varies with the altitude and the yield of the explosion. (*After S. Glasstone and P. J. Dolan, eds., The Effects of Nuclear Weapons, U.S. Department of Defense and the Energy Research and Development Administration, 3d ed., 1977*)

do not contribute to the generation of the EMP, they do cause the region to become highly conductive, and therefore play an important role in determining the EMP wave shape and amplitude.) While slowing down the very intense, short-duration flux of Compton electrons is also deflected by the Earth's geomagnetic field, according to Eq. (1), where the

$$\vec{F} = -q(\vec{v} \times \vec{B}) \tag{1}$$

deflection force $\vec{F}$ is perpendicular to the geomagnetic field $\vec{B}$ and the velocity $\vec{v}$ of a Compton electron of charge q. The Compton electrons then spiral about the geomagnetic lines, radiating electromagnetic energy in the form of EMP until they eventually recombine with local, positively charged ions. *See* COMPTON EFFECT; SYNCHROTRON RADIATION.

Characteristics for a high-altitude burst. For a high-altitude nuclear detonation, the radiated EMP observed at large distances from the source region can be represented in time t as an electric field $E(t)$ and a magnetic field $H(t)$, given by Eqs. (2) and (3), with

$$E(t) = E_0(e^{-\alpha t} - e^{-\beta t}) \qquad t \geq 0 \tag{2}$$

$$H(t) = H_0(e^{-\alpha t} - e^{-\beta t}) \qquad t \geq 0 \tag{3}$$

$E_0 = 5.2 \times 10^4$ V/m (1.6×10^4 V/ft), $H_0 = 1.4 \times 10^2$ A/m (4.2×10 A/ft), $\alpha = 4.0 \times 10^6$ s^{-1}, and $\beta = 4.76 \times 10^8$ s^{-1}. (In air, E and H are related by the impedance of free space, 377 ohms.) If the observer is directly below the detonation (ground zero), the polarization of both fields is predominantly horizontal; if the observer is at the horizon, the fields can have both horizontal and vertical components. The magnitude of these field components varies according to the latitude and longitude of the observer and the direction of the burst. *See* POLARIZATION OF WAVES.

Surface-burst EMP. Should the nuclear detonation occur closer to the Earth, the EMP generation process becomes far more complex and the electric and magnetic fields become very complicated. The most dramatic change occurs with a surface burst (**Fig. 3**).

When the observer is somewhere within the source region of a near-surface burst, where the air conductivity varies between 10^{-4} and 10^{-2} siemen/m (3×10^{-5} and 3×10^{-3} siemen/ft), the resultant electric field is predominantly vertical, and the resultant magnetic field is polarized perpendicular to the plane of the figure. The electric field polarization is due to the Compton-electron and ion-charge-separation fields which tend to become perpendicular to the conducting earth, leaving a resultant vertical electric field near the ground. As these Compton electrons move radially away from the detonation, they curve earthward and return to the detonation point through the conducting earth. This Compton current loop then gives rise to a resultant magnetic field. When the observer is far from the detonation (outside the source region), the electric and magnetic fields begin to approximate fields from a vertical dipole, decaying with distance r as $1/r$.

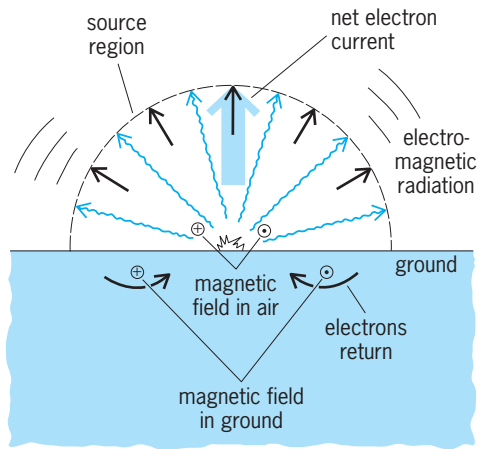

Fig. 3. Schematic representation of the EMP in a surface burst. (*After S. Glasstone and P. J. Dolan, eds., The Effects of Nuclear Weapons, U.S. Department of Defense and the Energy Research and Development Administration, 3d ed., 1977*)

Internal EMP. It is also possible for INR (both x-rays and gamma rays) to directly interact with systems, causing EMP signals internal to structures. This phenomenon has been called internal or system-generated EMP and is potentially a serious problem for satellites in orbit and for electronics in metallic enclosures on or near the ground. These forms of EMP are generated by x-rays interacting with satellites and gamma rays impinging on ground-based enclosures, producing currents of Compton electrons internally that then produce internal electromagnetic waves. They are very dependent upon the nuclear detonation, the system topology, and the relative position of one to the other.

Coupling. The coupling of these rather wide-band (10^4 to 10^8 Hz) signals to systems of different topologies can be significant. For example, it is not unusual to predict voltage and current levels of hundreds of thousands of volts and a few thousand amperes coupled by high-altitude EMP onto extended systems. The exact level of coupling depends upon the size of the system and its orientation with respect to the incident field, and upon whether or not it is near an earth ground. For ground-based systems the incident field components can actually be enhanced or degraded, depending upon the polarization. Solving Maxwell's equations at the interface between an ideal conductor and a dielectric shows that the net horizontal electric field is zero at the interface and the net vertical electric field doubles. (For the magnetic field components the opposite is true.) Since the Earth ground is not a perfect conductor, the field components do not completely cancel or quite double in amplitude. *See* MAXWELL'S EQUATIONS.

An estimate of about 1 joule (0.7 ft-lbf) of EMP-coupled energy is considered reasonable for many systems. Even if the coupling onto circuits is inefficient, as little as 10^{-13} J can upset some semiconductor devices and 10^{-6} J can cause damage. The potential for such upset and damage in critical electronic circuits has led to the incorporation of EMP protection in many system designs. This protection

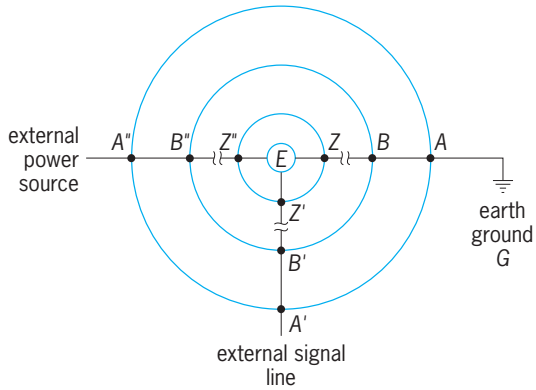

Fig. 4. Typical system protection scheme.

is most prevalent in communications systems whose disruption by EMP is considered an important civil and military vulnerability. *See* ELECTROMAGNETIC RADIATION; NUCLEAR EXPLOSION.

Protection. The most common form of protection incorporated in system designs is a combination of shielding and penetration control. **Figure 4** illustrates a protection scheme in which a system's electronics *E* is isolated from the external environment by one or more nested, shielded enclosures (often called Faraday cages). Penetration control is then maintained by minimizing the number of shield penetrations (in this case, a power line, a signal line, and a ground wire connecting *E* to earth ground *G*) and by applying terminal protection devices, such as spark gaps, high-power Zener diodes, or metal oxide varistors, at selected shield penetration points (A, A', A'', B, B', and B''; or Z, Z', and Z''; or both). In this way, system protection can be designed not only for EMP but also for other electromagnetic transients (such as near-strike lightning and electromagnetic interference). Furthermore, cost-effective, field maintainable protection can be achieved by properly selecting off-the-shelf shielding techniques and terminal protection devices and applying them to systems in accordance with such military standards as MIL-STD 461. *See* ELECTRIC PROTECTIVE DEVICES; ELECTRICAL INTERFERENCE; ELECTRICAL SHIELDING; ELECTROMAGNETIC RADIATION; LIGHTNING AND SURGE PROTECTION; NUCLEAR EXPLOSION. Robert A. Pfeffer

Bibliography. S. Glasstone and P. J. Dolan (eds.), *The Effects of Nuclear Weapons*, U.S. Department of Defense and the Energy Research and Development Administration, 3d ed., 1977, reprint 1983; K. S. Lee (ed.), *EMP Interaction: Principles, Techniques and Reference Data*, rev. ed, 1986; J. R. Pierce et al., *Evaluation of Methodologies for Estimating Vulnerability to Electromagnetic Pulse Effects,* National Research Council Rep., 1984.

Electromagnetic pump

A pump that operates on the principle that a force is exerted on a current-carrying conductor in a magnetic field. The high electrical conductivity of the liquid metals pumped (liquid metals are used as the heat-transfer media in some nuclear reactors and magnetohydrodynamic systems) allows a pumping force to be developed within the metals when they are confined in a duct or channel and subjected to a magnetic field and to an electric current. These pumps are designed principally for use in liquid-metal-cooled reactor plants where liquid lithium, sodium, potassium, or sodium-potassium alloys are pumped. Other metallic and nonmetallic liquids of sufficiently high electrical conductivity, such as mercury or molten aluminum, lead, and bismuth, may also be pumped in nonnuclear applications. The absence of moving parts within the pumped liquid eliminates the need for seals and bearings that are found in conventional mechanical pumps, thus minimizing leaks, maintenance, and repairs, and improving reliability. In liquid-metal-cooled nuclear reactor plants, electromagnetic pumps with a capacity of up to several thousand gallons per minute have operated without maintenance for decades. *See* MAGNETOHYDRODYNAMICS; NUCLEAR POWER; NUCLEAR REACTOR.

The major classifications of electromagnetic pumps, either conduction or induction, is based on the method employed to cause the electric current to flow in the pumped liquid.

Direct-current conduction pumps. In this simplest form of electromagnetic pump the liquid-containing pump duct has two heavy electrodes to provide a path for direct current to be conducted at a right angle to the direction of pumping force in the duct. A magnetic field generated by either a permanent magnet or electromagnet is impressed across the duct at the location of the electrodes in a direction at right angles to both the current flow and the pumping direction. This orthogonal arrangement of current, magnetic field, and pump duct produces a force in the liquid which is dependent on the magnitude of the current, magnetic field, and the length of the current path in the liquid. It is also possible to arrange the duct so that the current flowing in the liquid also produces the required magnetic field (that is, the self-induced pump). The physical arrangement of the conduction pump allows only a single passage of current between electrodes, resulting in a relatively short length of current path in the liquid. To compensate, the current must be very high (thousands of amperes), thus requiring special dc power supplies such as homopolar generators. However, the very low voltage allows special electrical insulation to be used, which is suitable for high-temperature (1500°F or 800°C) operation.

In addition to the externally energized dc conduction pump described above, there are two other unique pump design concepts of the dc conduction pump family in use or under development. In the thermoelectric electromagnetic pump, the current in the duct is generated by thermoelectric elements attached to the duct electrodes which are driven by the temperature difference between the duct and a heat sink. The major application of thermoelectric electromagnetic pumps is in the

liquid-metal loops of space reactors. *See* THERMOELECTRICITY.

The second design concept is the flow coupler. In this design a dc conduction pump duct in a secondary (driven) loop is directly coupled to a magnetohydrodynamic generator located in an adjacent parallel duct of a primary (driving) loop to which it is connected by appropriate buswork. Both ducts lie in a magnetic field produced by a permanent magnet or electromagnet. The initial use of this concept was to drive small auxiliary loops in experimental liquid-metal facilities. However, with the advent of liquid-metal pool reactors the concept is being studied as a means of circulating the primary coolant in the pool.

Alternating-current conduction pump. This pump uses a pump duct, magnetic field, and electrode arrangement similar to the dc conduction pump, except that the magnetic field is produced by an ac (single-phase) electromagnet and the current is produced by the transformer action of the same or separate single-phase electromagnet with a secondary (step-down) winding of usually one turn. The heavy secondary winding connects to the electrodes. With the sinusoidal magnetic field and duct current in phase with each other, a pumping force will be developed. This design can be operated directly from normal single-phase ac power equipment since it has an integral transformer to develop the necessary high current. It too can have a high-temperature application. *See* ELECTROMAGNET; TRANSFORMER.

Helical induction pump. This pump closely resembles a polyphase squirrel-cage ac induction motor in principle and operating characteristics. The fluid channel is formed by helical vanes encased in a concentric cylinder and is analogous to the rotor of a conventional motor. Arranged around the periphery of the outer cylinder is a polyphase multipole winding which is similar to the stator of an induction motor. The polyphase winding creates a rotating magnetic field which induces a current in the liquid metal, and this current in turn reacts with the magnetic field to produce a force on the fluid tending to rotate it in the duct. However, the fluid is channeled by the helical vanes, forcing it to produce an axial pressure and resultant axial flow in the duct. The helical induction pump is used in applications requiring higher pressures and lower flows than the linear induction pumps described below. *See* INDUCTION MOTOR.

Linear induction pumps. These pumps use a stationary polyphase ac winding to produce a traveling magnetic field that induces a current in the liquid. The reaction between the induced current and the traveling magnetic field causes the liquid to follow the traveling magnetic field. The force generated in the liquid is confined in the duct, producing a linear flow.

Two principal designs are the flat linear and the annular linear induction pumps. The flat linear induction pump has a rectangular duct connected to inlet and outlet nozzles. Several liquid-metal-cooled reactors use such pumps in their main heat transport systems. The annular induction pump design has a cylindrical duct formed by two concentric tubes.

These pumps are designed for loop application. They are temperature-limited with conventional electrical insulation and in most cases require a separate cooling system for the electrical windings. However, for application to pool- or tank-type liquid-metal-cooled reactors, there is considerable interest in the design and development of large submersible self-cooled (windings cooled by the liquid being pumped) linear induction pumps. Leslie R. Dahl

Bibliography. R. S. Baker and M. J. Tessier, *Handbook of Electromagnetic Pump Technology*, 1987; J. Collett et al., *Thermoelectric Electromagnetic Pump Design for Sp-100, Symposium on Space Nuclear Power Systems*, 1988; *Fast Breeder Reactors—Experience and Trends: Proceedings of a Symposium*, Lyons, 1986; *International Conference on Design and Safety of Advanced Nuclear Power Plants*, Tokyo, 1992; *International Conference on Fast Reactors and Related Fuel Cycles*, Kyoto, 1991; A. E. Walter and A. B. Reynolds, *Fast Breeder Reactors*, 1981.

Electromagnetic radiation

Energy transmitted through space or through a material medium in the form of electromagnetic waves. The term can also refer to the emission and propagation of such energy. Whenever an electric charge oscillates or is accelerated, a disturbance characterized by the existence of electric and magnetic fields propagates outward from it. This disturbance is called an electromagnetic wave. The frequency range of such waves is tremendous, as is shown by the electromagnetic spectrum in the **table**. The sources given are typical, but not mutually exclusive, as is shown by the fact that the atomic interstellar hydrogen radiation whose wavelength is 0.210614 m falls in the radar region. The other monochromatic radiation listed is that from positron-electron annihilation whose wavelength is 2.42626×10^{-12} m.

Detection of radiation. In theory, any electromagnetic radiation can be detected by its heating effect. This method has actually been used over the range from x-rays to radio. Ionization effects measured by cloud chambers, photographic emulsions, ionization chambers, and Geiger counters have been used in the gamma- and x-ray regions. Direct photography can be used from the gamma-ray to the infrared region. Fluorescence is effective in the x-ray and ultraviolet ranges. Bolometers, thermocouplers, and other heat-measuring devices are used chiefly in the infrared and microwave regions. Crystal detectors, vacuum tubes, and transistors cover the microwave and radio frequency ranges.

Free-space waves. A charge in simple harmonic (linear sinusoidal) motion in a vacuum generates a simple wave which becomes spherical at distances from the source much larger than the amplitude of the motion and so great that many oscillations have occurred before the disturbance arrives. The wave is plane when the dimensions of the area observed are very small compared with the radius of spherical

Electromagnetic spectrum

Frequency, Hz	Wavelength, m	Nomenclature	Typical source
10^{23}	3×10^{-15}	Cosmic photons	Astronomical
10^{22}	3×10^{-14}	γ-rays	Radioactive nuclei
10^{21}	3×10^{-13}	γ-rays, x-rays	
10^{20}	3×10^{-12}	X-rays	Atomic inner shell, positron-electron annihilation
10^{19}	3×10^{-11}	Soft x-rays	Electron impact on a solid
10^{18}	3×10^{-10}	Ultraviolet, x-rays	Atoms in sparks
10^{17}	3×10^{-9}	Ultraviolet	Atoms in sparks and arcs
10^{16}	3×10^{-8}	Ultraviolet	Atoms in sparks and arcs
10^{15}	3×10^{-7}	Visible spectrum	Atoms, hot bodies, molecules
10^{14}	3×10^{-6}	Infrared	Hot bodies, molecules
10^{13}	3×10^{-5}	Infrared	Hot bodies, molecules
10^{12}	3×10^{-4}	Far-infrared	Hot bodies, molecules
10^{11}	3×10^{-3}	Microwaves	Electronic devices
10^{10}	3×10^{-2}	Microwaves, radar	Electronic devices
10^{9}	3×10^{-1}	Radar	Electronic devices, interstellar hydrogen
10^{8}	3	Television, FM radio	Electronic devices
10^{7}	30	Short-wave radio	Electronic devices
10^{6}	300	AM radio	Electronic devices
10^{5}	3000	Long-wave radio	Electronic devices
10^{4}	3×10^{4}	Induction heating	Electronic devices
10^{3}	3×10^{5}		Electronic devices
100	3×10^{6}	Power	Rotating machinery
10	3×10^{7}	Power	Rotating machinery
1	3×10^{8}		Commutated direct current
0	Infinity	Direct current	Batteries

curvature. In this case the choice of the rectangular coordinates x and z as the directions of the oscillation and of the observation or field point, respectively, permits the electric intensity **E** and the magnetic flux density **B** to be written as Eq. (1). The field

$$E_x = vB_y = E_0 \cos\,[\omega(t - v^{-1}z)] \qquad (1)$$

amplitude E_0 is constant over the specified area and not dependent on z if the z range is small compared with the source distance, as in stellar radiation. The angular frequency of the source is ω radians per second, which is the frequency ν in hertz multiplied by 2π. The velocity of the wave is v, the direction of propagation z, and the time t. The wavelength λ is $2\pi v/\omega$. If t is in seconds and z is in meters, then v is in meters per second and λ is in meters. It is found that in a lossless, isotropic, homogeneous medium Eq. (2) holds; here μ is the permeability, and ϵ the

$$v = (\mu\epsilon)^{-1/2} \qquad (2)$$

capacitivity, or dielectric constant. This wave is transverse because **E** and **B** are normal to z. It is plane-polarized because E_x and B_y are parallel to fixed axes. The plane of polarization is taken as that defined by the electric vector and the direction of propagation.

Plane waves. An electromagnetic disturbance is a plane wave when the instantaneous values of any field element such as **E** and **B** are constant in phase over any plane parallel to a fixed plane. These planes are called wavefronts. In empty unbounded space, **E** and **B** lie in the wavefront normal to each other; if the wave is unpolarized, their direction fluctuates in this plane in random fashion. If the plane waves are bounded, as on transmission lines and in waveguides, the amplitudes may vary over the wavefront, and in the case of waveguides and crystals some of the elements will not in general lie in the wavefront. The equation for an undamped plane wave whose front is normal to z is Eq. (3), where F is one of the

$$F = \Phi_1(x, y)f_1(z - vt) + \Phi_2(x, y)f_2(z + vt) \qquad (3)$$

field elements such as **E** or **B**. Note that if an observer sees a certain value of $\Phi_1(x,y)$ at z and then jumps instantaneously in the z direction to a point $z + \Delta z$, the observer will, after waiting a time $\Delta z/v$, see the same value $\Phi_1(x,y)$ because Eq. (4) is valid.

$$f(z - vt) = f[z + \Delta z - v(t + \Delta z/v)] \qquad (4)$$

Thus, the first term represents a wave moving in the z direction with a velocity v. The second term represents a wave in the negative z direction. The form of $\Phi_1(x,y)$ and $\Phi_2(x,y)$ depends on the boundary conditions. *See* WAVE EQUATION.

Spherical waves. A wave is spherical when the instantaneous value of any field element such as **E** or **B** is constant in phase over a sphere. The radiation from any source of finite dimensions becomes spherical at great distances in an unbounded, isotropic, homogeneous medium. The equation for an undamped spherical wave is Eq. (5). The first term represents a

$$F = r^{-1}\Phi_1(\theta, \varphi)f(r - vt) + r^{-1}\Phi_2(\theta, \varphi)f(r + vt) \qquad (5)$$

diverging and the second a converging wave. Again, the form of $\Phi_1(\theta,\varphi)$ and $\Phi_2(\theta,\varphi)$ depends on the nature of the source and other boundary conditions.

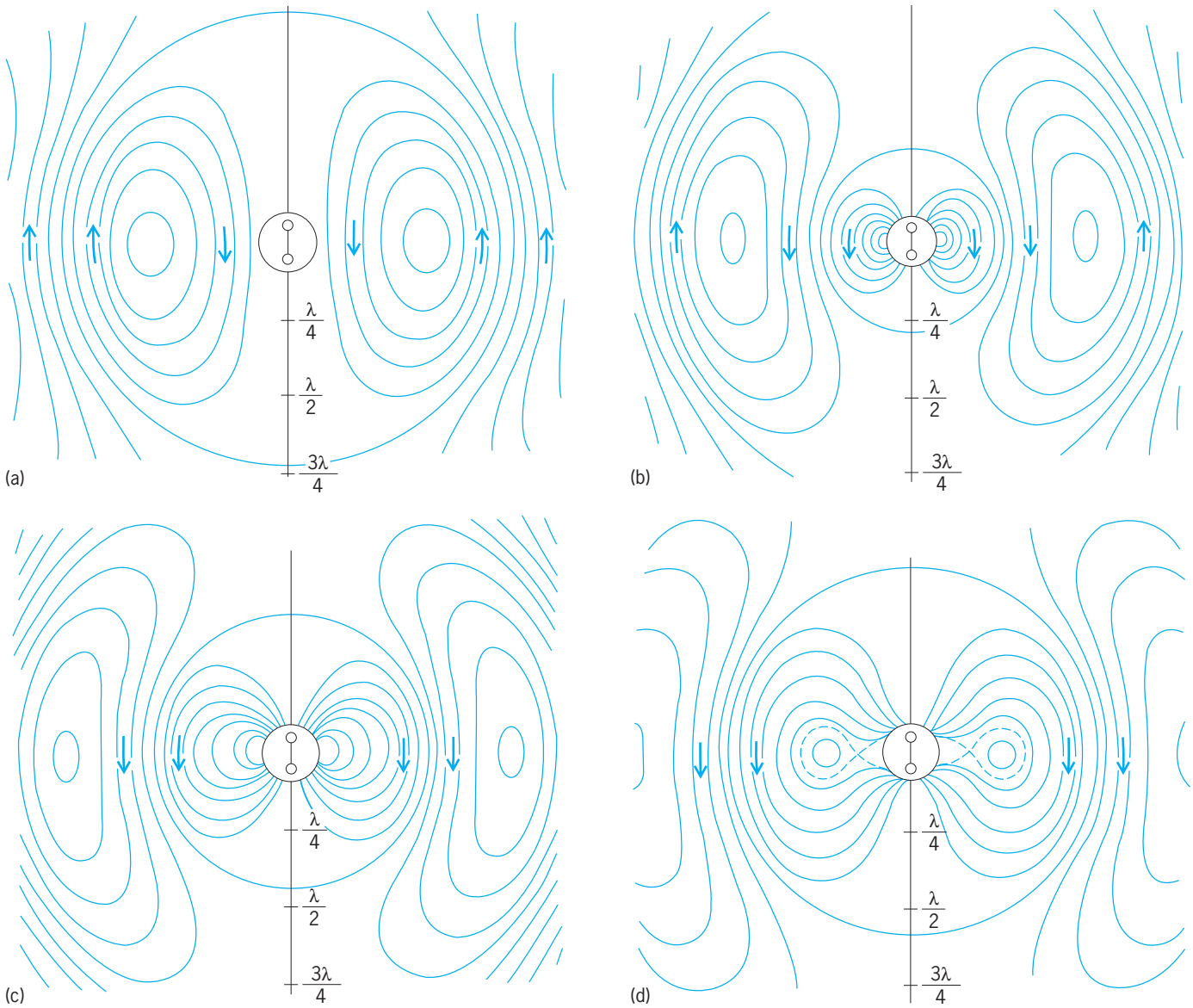

Diagrams of electric dipole. The outward moving electric field lines generated by the hertzian oscillator are shown at successive eighth-period intervals. (*a*) $t = 0$; (*b*) $t = T/8$; (*c*) $t = T/4$; (*d*) $t = 3T/8$.

Damped waves. If there are energy losses which are proportional to the square of the amplitude, as in the case of a medium of conductivity γ which obeys Ohm's law, then the wave is exponentially damped, and Eq. (1) becomes Eq. (6). The symbol α is called

$$E_x = E_0 e^{-\alpha z} \cos(\omega t - \beta z) \tag{6}$$

the attenuation constant, and β the wave number or phase constant which equals ω/υ', where υ' is the damped-wave velocity. The electric wave amplitude at the origin has been taken as E_0. The ratio of E_0 to B_0, as well as that of α to β, depends on the permeability μ, the capacitivity ϵ, and the conductivity γ of the medium. In terms of the phasor $\check{E}_x$, Eq. (6) may be written as the real part of Eq. (7).

$$E_x = \check{E}_x e^{j\omega t} = E_0 e^{-(\alpha + j\beta)z} e^{j\omega t} \tag{7}$$

This is exactly the form for the current on a transmission line. (Phasors are complex numbers of form such that, when multiplied by $e^{j\omega t}$, the real part of the product gives the amplitude, phase, and time dependence.)

Wave impedance. Those trained in transmission line theory find it useful to apply the same techniques to wave theory. Consider an isolated tubular section of the wave in Eq. (1) bounded by $x = 0$, $x = 1$, and $y = 0$, $y = 1$ as a transmission line. The potential across the line between $x = 0$ and $x = 1$ is **E**. The line integral of **B** around the $x = 0$ boundary from $y = 0$ to $y = 1$ is μI by Ampère's law and equals **B** because **B** is zero on the negative side. Thus, the impedance of the line is, making use of Eqs. (1) and (2), given by Eq. (8). This depends only on the properties of

$$\check{Z}_k = \frac{V}{I} = \frac{\mu E}{B} = \frac{E}{H} = \left(\frac{\mu}{\epsilon}\right)^{1/2} = \eta \tag{8}$$

the medium and is known as the wave impedance.

In transmission line theory the ratio μ/ϵ would be replaced by the ratio of the series impedance $\check{Z}_L = j\omega L$ to the shunt admittance $\check{Y} = j\omega C$, where L is the inductance per unit length, and C the capacitance per unit length across the line. If there is a resistance R per unit length across the line, then $1/R$ must be added to Y. This resistance is $1/\gamma$ for the tubular section. Thus, for a conducting medium, Eq. (8) becomes Eq. (9). The last term is a common

$$\check{Z}_k = \left(\frac{j\omega\mu}{\gamma + j\omega\epsilon}\right)^{1/2} = \frac{j\omega\mu}{\alpha + j\beta} \tag{9}$$

transmission line form. The reflection and refraction of plane waves at plane boundaries separating different mediums may be calculated by transmission line formulas with the aid of Eqs. (8) and (9).

Electric dipole. A charge undergoing simple harmonic motion in free space is a dipole source when the amplitude of the motion is small compared with the wavelength. The term is loosely applied to the hertzian oscillator, usually pictured as a dumbbell-shaped conductor in which the electrons oscillate from one end to the other, leaving the opposite end periodically positive. An electric dipole of moment M is defined as the product qa when two large, equal and opposite charges, $+q$ and $-q$, are placed a small distance a apart. A dipole is oscillating when M is periodic in time and is the simplest source of spherical waves. Much can be learned by a study of H. Hertz's picture of the outward moving electric field lines at successive time intervals of one-eighth period in a plane which passes through the hertzian oscillator axis, shown in the **illustration**. The most striking feature of the pictures is that, after breaking loose from the dipole, all electric field lines are closed, which means that the divergence of **E** is zero. This is true of all unbounded waves. It is also noteworthy that the waves become truly spherical with a fixed wavelength λ only in a direction perpendicular to the dipole and at a distance which greatly exceeds the dipole dimensions. This distance is beyond the edges of the picture. Lengths $\lambda/4$, $\lambda/2$, and $3\lambda/4$ are marked off on the axis for comparison. The magnetic field lines are circles coaxial with the oscillator, so they intersect the plane of the diagram normally. They are most dense where the electric lines are closely spaced. The radiant energy emitted by atoms and molecules is essentially radiation of the dipole type. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; ANTENNA (ELECTROMAGNETISM); DIFFRACTION; ELECTROMAGNETIC WAVE TRANSMISSION; GAMMA RAYS; HEAT RADIATION; INFRARED RADIATION; INTERFERENCE OF WAVES; LIGHT; MAXWELL'S EQUATIONS; MICROWAVE; POLARIZATION OF WAVES; RADIATION; RADIO-WAVE PROPAGATION; REFLECTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES; SCATTERING OF ELECTROMAGNETIC RADIATION; TRANSMISSION LINES; ULTRAVIOLET RADIATION; WAVEGUIDE; WAVE MOTION; X-RAYS.

William R. Smythe

Bibliography. M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; N. A. Dyson, *X-Rays in Atomic and Nuclear Physics*, 2d ed., 1990; B. D. Guenther, *Modern Optics*, 1990; A. Ishimaru, *Electromagnetic Wave Propagation, Radiation, and Scattering*, 1996; M. F. Iskander, *Electromagnetic Fields and Waves*, 2000; J. A. Kong, *Electromagnetic Wave Theory*, 3d ed., 2000; S. Ramo, J. R. Whinnery, and T. Van Duzer, *Fields and Waves in Communications Electronics*, 3d ed., 1993; V. V. Sarwate, *Electromagnetic Fields and Waves*, 1993; W. R. Smythe, *Static and Dynamic Electricity*, 3d ed., 1968, reprint 1989.

Electromagnetic wave

A disturbance, produced by the acceleration or oscillation of an electric charge, which has the characteristic time and spatial relations associated with progressive wave motion. A system of electric and magnetic fields moves outward from a region where electric charges are accelerated, such as an oscillating circuit or the target of an x-ray tube. The wide wavelength range over which such waves are observed is shown by the electromagnetic spectrum. The term electric wave, or hertzian wave, is often applied to electromagnetic waves in the radar and radio range. Electromagnetic waves may be confined in tubes, such as wave guides, or guided by transmission lines. They were predicted by J. C. Maxwell in 1864 and verified experimentally by H. Hertz in 1887. *See* ELECTROMAGNETIC RADIATION; MAXWELL'S EQUATIONS.

William R. Smythe

Electromagnetic wave transmission

The transmission of electrical energy by wires, the broadcasting of radio signals, and the phenomenon of visible light are all examples of the propagation of electromagnetic energy. Electromagnetic energy travels in the form of a wave. Its speed of travel is approximately 186,000 mi/s (3×10^8 m/s) in a vacuum and is somewhat slower than this in liquid and solid insulators. An electromagnetic wave does not penetrate far into an electrical conductor, and a wave that is incident on the surface of a good conductor is largely reflected.

Electromagnetic waves originate from accelerated electric charges. For example, a radio wave originates from the oscillatory acceleration of electrons in the transmitting antenna. The light that is produced within a laser originates when electrons fall from a higher energy level to a lower one. *See* LASER.

The waves emitted from a source are oscillatory in character and are described in terms of their frequency of oscillation. Local telephone lines (not using carrier systems) carry electromagnetic waves with frequencies of about 200–4000 Hz. Medium-wave radio uses frequencies of the order of 10^6 Hz, radar uses frequencies of the order of 10^{10} Hz, and a ruby laser emits light with a frequency of

4.32×10^{14} Hz. The method of generating an electromagnetic wave depends on the frequency used, as do the techniques of transmitting the energy to another location and of utilizing it when it has been received. *See* RADAR.

The communication of information to a distant point is generally accomplished through the use of electromagnetic energy as a carrier. A familiar example is the telephone, in which sound waves in the range of frequencies from a few hundred to a few thousand hertz are converted into corresponding electromagnetic waves, which are then guided to their destination by a pair of wires. Another familiar example is radio, in which the signals are caused to modify an identifiable characteristic, such as the amplitude or frequency, of an electromagnetic carrier wave. The electromagnetic wave, thus modified, or modulated, is radiated from an antenna and can be received over a considerable region. *See* MODULATION; RADIO BROADCASTING.

Features of electromagnetic waves. Figure 1 illustrates schematically some of the essential features of an electromagnetic wave. Shown in Fig. 1 are vectors that represent the electric field intensity E and the magnetic field intensity H at various points along a straight line taken in the direction of propagation of the wave. The electric field is in a vertical plane and the wave is said to be vertically polarized. The magnitude of the field, at a given instant, varies as a sinusoidal function of distance along the direction of propagation. The magnetic field intensity H lies in a plane normal to that of E and, at each point, is proportional in magnitude to E, as shown in Eq. (1),

$$\frac{E}{H} = \sqrt{\frac{\mu}{\epsilon}} \tag{1}$$

where H is the magnetic field intensity in amperes/m, E is the electric field intensity in volts/m, ϵ is the permittivity, or absolute dielectric constant, of the medium, and μ is the absolute permeability of the medium. For a vacuum, $\epsilon = 8.854 \times 10^{-12}$ farad/m and $\mu = 4\pi \times 10^{-7}$ henry/m; therefore for a vacuum the ratio E/H is approximately 377 ohms. This ratio is termed the wave impedance of the medium. *See* MAXWELL'S EQUATIONS; WAVE EQUATION.

The E and H waves travel along a straight line, as suggested in Fig. 1. Of the two possible directions along this line, the actual direction of travel can be determined by imagining a screw with a right-hand thread placed along the axis and turned from E toward H; then the longitudinal direction of travel of the screw is the direction of propagation of the energy.

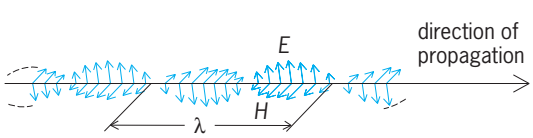

Fig. 1. Representation of an electromagnetic wave at a particular instant of time.

The velocity of travel of the wave is shown in Eq. (2). In a vacuum this is approximately 3×10^8 m/s. The velocity in air is only slightly smaller.

$$v = 1\sqrt{\mu\epsilon} \tag{2}$$

The wavelength is the distance between two successive similar points on the wave, measured along the direction of propagation. The wavelength is denoted by λ in Fig. 1.

As the wave travels past a stationary point, the values of E and H at the point vary sinusoidally with time. The time required for one cycle of this variation is termed the period, T seconds. The number of hertz is the frequency f; and $f = 1/T$. In one cycle the wave, traveling at the velocity v, moves one wavelength along the axis of propagation. Therefore, $\lambda = vT$, or may be calculated by Eq. (3). Assuming a velocity

$$\lambda = \frac{v}{f} \tag{3}$$

of 3×10^8 m/s, an electromagnetic wave having a frequency of 60 Hz has a wavelength of 5×10^6 m, or approximately 3100 mi. At a frequency of 3 MHz (3×10^6 Hz), λ is 100 m, and at 3000 MHz, λ is 10 cm. Visible light has a frequency of the order of 5×10^{14} Hz and a wavelength of approximately 6×10^{-5} cm.

The density of energy in an electric field is $\epsilon E^2/2$ joules/m^3, and that in a magnetic field is $\mu H^2/2$ J/m^3. With the aid of Eq. (1), the relationships become those shown in Eq. (4). Therefore the

$$\frac{\mu H^2}{2} = \frac{\mu(\sqrt{\epsilon/\mu}E)^2}{2} = \frac{\epsilon E^2}{2} \tag{4}$$

electric and magnetic fields carry equal energies in the electromagnetic wave. The total energy density at any point is equal to ϵE^2 J/m^3. Since this is transported with a velocity equal to $1/\sqrt{\mu\epsilon}$, the rate of flow of energy per square meter normal to the direction of propagation is $\epsilon E^2 v$ or $E^2\sqrt{\epsilon/\mu}$ watts/m^2. In radio broadcasting a field of 50 millivolts/m is considered to be strong. An electromagnetic wave with this intensity has an average energy density of 2.2×10^{-14} J/m^3, and the average rate of energy flow is 6.6×10^{-6} W/m^2.

Radiation from an antenna. Figure 2 illustrates the configuration of the electric and magnetic fields about a short vertical antenna in which flows a sinusoidal current of the form $i = I_{max} \sin 2\pi ft$ amperes. The picture applies either to an antenna in free space (in which case the illustration shows only the upper half of the fields), or to an antenna projecting above the surface of a highly conducting plane surface. In the latter case the conducting plane represents to a first approximation the surface of the Earth. The fields have symmetry about the axis through the antenna. For pictorial simplicity only selected portions of the fields are shown in Fig. 2. The magnetic field is circular about the antenna, is perpendicular at every point to the direction of the electric field, and is proportional in intensity to the magnitude of the electric field, as indicated by Eq. (1). All parts of the wave

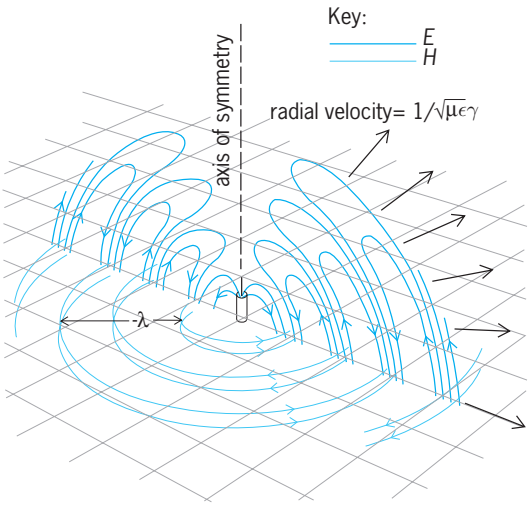

Fig. 2. Configuration of electric and magnetic fields about a short vertical antenna.

travel radially outward from the antenna with the velocity given by Eq. (2); the wave is described as spherical, with the antenna located at the center of the wave. The wavelength of the radiation is given by Eq. (3).

If a short antenna that projects above a highly conducting plane surface carries a current of $i = I_{\max} \sin 2\pi ft$ that is uniform throughout the length of the antenna, the intensity of the electric field in the radiated wave is that shown in Eq. (5). where l is the

$$E_{\max} = \sqrt{\frac{\mu}{\epsilon}}\,\frac{I_{\max}}{r}\,\frac{l}{\lambda}\cos\theta \qquad \text{V/m} \qquad (5)$$

length of the antenna, r is the radial distance from the antenna, and θ is 0, and varies inversely with the distance from the antenna.

If the rate of flow of energy per unit area $E^2\sqrt{\epsilon/\mu}$ is integrated over an imaginary spherical surface about the antenna, the average power radiated is that shown by Eq. (6). The factor $(l/\lambda)^2$ is of par-

$$P_{\text{av}} = \frac{2\pi}{3}\sqrt{\frac{\mu}{\epsilon}}\,I_{\max}^2\left(\frac{l}{\lambda}\right)^2 \qquad \text{W} \qquad (6)$$

ticular importance, for it indicates that a longer antenna is required at the longer wavelengths (lower frequencies). The radiation of appreciable energy at a very low frequency requires an impracticably long antenna.

The foregoing relations assume a uniform current throughout the length of the antenna. An approximation to this can be achieved in practice by connecting a long horizontal conductor to the top of the antenna. Where some such construction is not utilized, the current will not be uniform in the antenna and will, in fact, be zero at the tip. The results given by Eqs. (4) and (5) must be modified, but the qualitative features of the radiation remain as shown in Fig. 2. *See* ANTENNA (ELECTROMAGNETISM).

Often it is desired to concentrate the radiated energy into a narrow beam. This can be done either by the addition of more antenna elements or by placing a large reflector, generally parabolic in shape, behind the antenna. The production of a narrow beam requires an antenna array, or alternatively a reflector, that is large in width and height compared with a wavelength. The very narrow and concentrated beam that can be achieved by a laser is made possible by the extremely short wavelength of the radiation as compared with the cross-sectional dimensions of the radiating system.

Propagation over the Earth. The foregoing discussion shows some of the important features of the radiation of electromagnetic energy from an antenna, but is oversimplified insofar as communication to and from positions on or near the Earth is concerned. The ground is a reasonably good, but not perfect, conductor; hence, the actual propagation over the surface of the Earth will show a more rapid decrease of field strength than that indicated by the factor of $1/r$ in Eq. (5). Irregularities and obstructions may interfere. In long-range transmission the spherical shape of the Earth is important. Inhomogeneities in the atmosphere refract the wave somewhat. For long-range transmission, the ionized region high in the atmosphere known as the Kennelly-Heaviside layer, or ionosphere, can act as a reflector. The electric field of the wave produces oscillation of the charged particles of the region, and this causes the refractive index of the layer to be smaller than that of the atmosphere below. The result is that, if the angle of incidence is not too near the normal and if the frequency of the wave is not too high, the wave may be refracted back toward the Earth. Successive reflections between ionosphere and Earth can provide communication for long distances around the periphery of the Earth. *See* RADIO-WAVE PROPAGATION.

Hollow waveguides. When an electromagnetic wave is introduced into the interior of a hollow metallic pipe of suitably large cross-sectional dimensions, the energy is guided along the interior of the pipe with comparatively little loss. The most common cross-sectional shapes are the rectangle and the circle. The cross-sectional dimensions of the tube must be greater than a certain fraction of the wavelength; otherwise the wave will not propagate in the tube. For this reason hollow waveguides are commonly used only at wavelengths of 10 cm or less (frequencies of 3000 MHz or higher).

A single wave of the type in Fig. 1 cannot propagate longitudinally inside a tubular conductor since, at some portions of the inner surface of the conducting tube, the E vector of the wave necessarily would have a component tangential to the surface. This is impossible because an electric field cannot be established along a good conductor, such as the wall of the tube. An electromagnetic wave can propagate along the interior of the tube only by reflecting back and forth between the walls of the tube. This reflection is a comparatively simple one between the plane surfaces of a rectangular tube, but is a complex reflection in tubes of other cross-sectional shapes.

A dielectric rod can also be used as a waveguide. Such a rod, if of insufficient cross-sectional

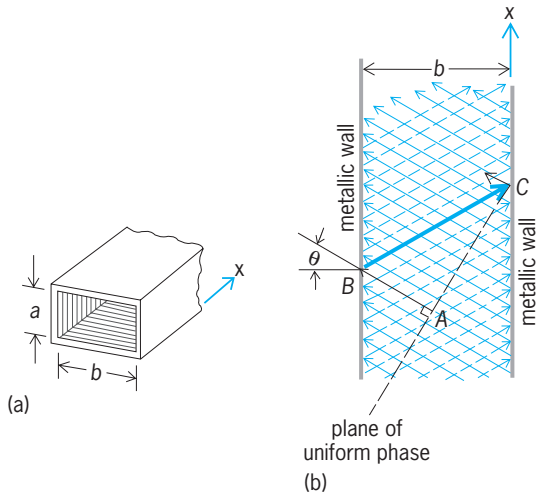

Fig. 3. Hollow metallic waveguide of rectangular cross section. (*a*) Guide. (*b*) Paths of electromagnetic energy in the simplest mode of propagation.

dimensions, can contain the electromagnetic wave by the phenomenon of total reflection at the surface.

A hollow metallic waveguide of rectangular cross section is shown in **Fig. 3***a*. The simplest mode of propagation is indicated in Fig. 3*b*. The entire space is filled with a plane electromagnetic wave which moves obliquely to the left in the direction shown by the solid arrows. This wave has its E vector normal to the paper and its H vector in the plane of the paper. Any plane normal to the direction of propagation is a plane of uniform phase (thus the name plane wave), and one such plane is indicated in the illustration by a broken line. The wave strikes the wall at an angle θ from the normal and is reflected at an equal angle. As the wave is reflected, the direction of its E vector reverses so as to make the tangential component of the electric field equal to zero at the conducting wall. The wave incident on the left wall thus is reflected to the right, where it is again reflected and moves to the left. By successive reflections the energy propagates longitudinally along the interior of the guide. As the wave incident upon the wall reflects and reverses the direction of its E vector, electric currents are caused to flow in the conducting wall. Since the wall is not a perfect conductor, some of the energy of the wave is transformed into heat. Consequently the amplitude of the wave diminishes exponentially as it passes down the guide; this phenomenon is termed attenuation. For an electromagnetic wave with a frequency of 3000 MHz (wavelength of 10 cm) propagating down the interior of a rectangular copper waveguide with cross-sectional dimensions of 1.4 by 3.2 in (or 4 by 8 cm) half the power is lost in a distance of approximately 450 ft (150 m). Hollow waveguides are used chiefly for short-distance transmission, as from a transmitter to an antenna. *See* WAVEGUIDE.

The requirements on the reflection of the wave, as outlined above, restrict the wavelength that can be propagated in a hollow guide. Consider the ray ABC in Fig. 3. The wave propagates from A to B, where it is reflected with reversal of the E vector; thereupon it propagates from B to C, where it is again reflected with another reversal of the E vector. But AC is a line of equal phase, and so the wave emerging from C must have the same phase as that at A. Thus the distance ABC must be an integral multiple of a wavelength, or $n\lambda$, where n is a positive integer. The distance ABC is $2b \cos \theta$ where b is the breadth of the guide; hence, $n\lambda = 2b \cos \theta$. The condition for propagation down the axis of the guide is that $\theta > 0$; hence $\cos \theta < 1$, and the restriction on wavelength is $\lambda < 2b/n$. The greatest ratio of wavelength to breadth of guide is obtained when $n = 1$, whence $\lambda/b < 2$. Therefore, the breadth of the waveguide must be somewhat greater than $\lambda/2$.

In the simple mode of propagation described above, the fields are independent of distance in the direction of the dimension a, and this dimension has no influence on the propagation. The net electric vector caused by the sum of the two waves is everywhere transverse to the longitudinal axis of the guide, and so the mode is described as transverse electric (TE).

If the wavelength of the radiation is small enough in comparison with the cross-sectional dimensions of the guide, more complex modes of propagation are possible, in which the wave reflects obliquely against a side wall, proceeds to the top of the guide, and reflects from there to the other side wall, then to the bottom wall, and so on. With this type of reflection it is possible to have both transverse-electric and transverse-magnetic (TM) modes. In the latter the net H vector is everywhere transverse to the axis of the guide.

When the dimensions of the guide are such that complex modes are possible, so also are the simple ones. The transmission of energy by a combination of modes introduces complications in abstracting the energy from the guide at the receiving end. Propagation in only the simplest mode is ensured by selecting the dimension b to be greater than $\lambda/2$ but not as large as λ, and also by restricting the dimension a so as to render complex modes impossible.

Waveguides of circular cross section are sometimes used. Analysis of these shows that the first TE mode is propagated if the diameter of the guide is greater than 0.586λ, and that the first TM mode is propagated if the radius is greater than 0.766λ.

Two-conductor transmission lines. Electromagnetic energy can be propagated in a simple mode along two parallel conductors. Such a waveguiding system is termed a transmission line. Three common forms are shown in **Fig. 4**. If the spacing between conductors is a small fraction of the wavelength of the transmitted energy, only one mode of propagation is possible. This corresponds to the wave of Fig. 1, with the direction of propagation taken longitudinally along the line. The E and H vectors are in the plane of the cross section, and the mode is termed transverse electromagnetic (TEM). The E vector must be at right angles to a highly conducting surface, and the oscillating H vector must be parallel with such a surface. With two separated conductors, there is for each

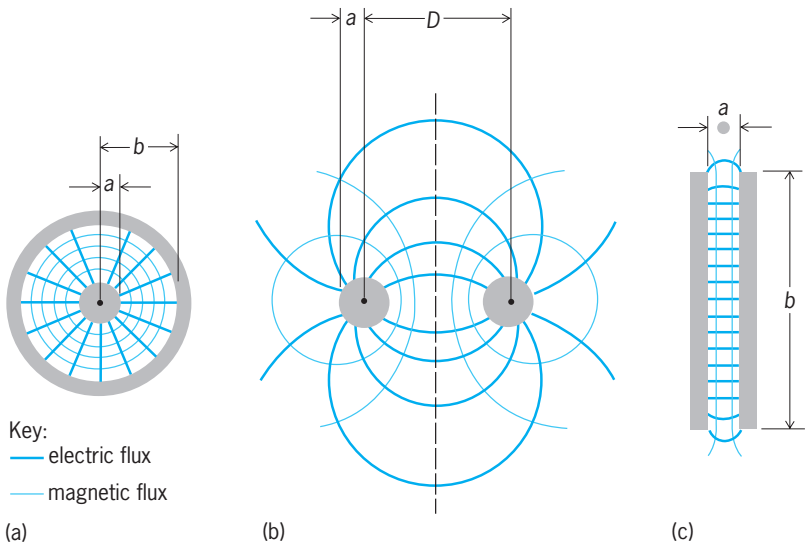

Fig. 4. Cross sections of common two-conductor transmission lines. (*a*) Coaxial cable. (*b*) Two-wire line. (*c*) Parallel-strip line.

geometrical arrangement of conductors one and only one cross-sectional field configuration which will satisfy the boundary conditions at the metal surface. The field configurations for coaxial and two-wire lines are shown in Fig. 4. At each point the ratio of E to H is as given by Eq. (1), and the velocity of propagation of the wave is as given by Eq. (2). Half of the propagated energy is contained in the electric field and half in the magnetic field. This mode of propagation is in contrast with the more complex modes required in a hollow metal pipe, where the conditions required at the boundaries can be satisfied only by means of reflections at the metal walls. As a result, the two-conductor transmission line does not have the upper limit on wavelength that was imposed on the hollow waveguide by the requirement of reflections; in fact, the two-conductor line operates completely normally at zero frequency (direct current).

At wavelengths that are small enough to be comparable with the cross-sectional dimensions of the line, more complex modes, involving reflections from the surfaces of the conductors, become possible. High-frequency energy can thus be propagated in several modes simultaneously. In a coaxial cable a rough criterion for the elimination of higher modes is that the wavelength should be greater than the average of the circumferences of the inner and outer conductors.

As the wave propagates along the line, it is accompanied by currents which flow longitudinally in the conductors. These currents can be regarded as satisfying the boundary condition for the tangential H field at the surface of the conductor. The conductors have a finite conductivity, and so these currents cause a transformation of electrical energy into heat. The energy lost comes from the stored energy of the wave, and so the wave, as it progresses, diminishes in amplitude. The conductors are necessarily supported by insulators which are imperfect and cause additional attenuation of the wave. In a typical open-wire telephone line operating at voice frequencies, half the energy is lost in a distance of perhaps 60 mi (95 km). The losses increase with frequency, and for a typical air-insulated coaxial line operating at 5 MHz, half the energy is lost in a distance of less than 1 mi (1.6 km). At a frequency of 3000 MHz ($\lambda = 10$ cm), typical distances in which half the energy is lost are, for air-insulated coaxial cable, 82 ft (25 m); for flexible coaxial cable insulated with polyethylene, 33 ft (10 m).

Noise. In a transmission line intended for the transmission of large amounts of power, such as the cross-country lines joining electrical generating stations to centers of population, the loss of an appreciable proportion of the power enroute is a serious matter. In a communication system, however, the average rate of flow of energy is rather small, and the intrinsic value of the energy itself is not of prime importance. The important characteristic of such a system is the accurate transmission of information, and the limiting factor is noise. Noise is always present in a transmission channel. Two common causes are thermal agitation and nearby electrical discharges. In a transmission system conveying information by an electromagnetic wave, the loss of energy in transmission becomes a serious matter if the wave is attenuated to the point where it is not large enough to override the noise. Amplifiers must be inserted in the transmission system at sufficiently close intervals, so that the signal never falls into the noise level, from which it could not be recovered and interpreted accurately.

Circuit analysis of transmission lines. Because the conductors of a transmission line are almost always spaced much closer together than a quarter wavelength of the electromagnetic energy that they are guiding, it is possible to analyze their performance quantitatively by circuit theory. It is then possible to deal with the voltages between the conductors and the currents flowing along the conductors, instead of with the electric and magnetic fields that exist in the insulating medium.

The waveguiding properties of the transmission line can be examined most conveniently if losses of energy are ignored. If L is defined as the inductance of the pair of conductors per unit length and C the capacitance between the conductors per unit length, field theory shows that L is μF_g and C is ϵ/F_g, where F_g is a geometrical factor that depends on the cross-sectional configuration of the conductors. For a coaxial line (Fig. 4*a*), $F_g = (^1/_2\pi)\log_\epsilon(b/a)$. For a two-wire line (Fig. 4*b*), $F_g = (1/\pi)\log_\epsilon(D/a)$. For a parallel-strip line (Fig. 4*c*), neglecting edge effects, $F_g = a/b$.

In the circuit analysis of a transmission line, the line can be visualized as being composed of a cascaded set of sections, each of short length Δx, as shown in **Fig. 5***b*. The partial differential equations which describe the voltage e and the current i are shown by Eqs. (7) and (8).

$$\frac{\partial e}{\partial x} = -L\,\frac{\partial i}{\partial t} \tag{7}$$

$$\frac{\partial i}{\partial x} = -C\,\frac{\partial e}{\partial t} \tag{8}$$

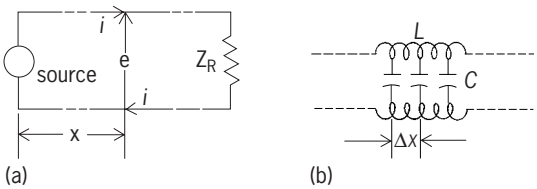

Fig. 5. Schematic representation of a transmission line. (*a*) Circuit diagram. (*b*) Visualization of *L* and *C*.

The solution of these equations is shown by Eqs. (9) and (10), where f_1 and f_2 are any finite, single-

$$e = f_1 \frac{x-t}{\sqrt{LC}} + f_2 \frac{x+t}{\sqrt{LC}} \tag{9}$$

$$i = \frac{1}{\sqrt{L/C}} \left(f_1 \frac{x-t}{\sqrt{LC}} - f_2 \frac{x+t}{\sqrt{LC}} \right) \tag{10}$$

valued functions of the arguments $x - t/\sqrt{LC}$ and $x + t/\sqrt{LC}$, respectively. These are interpreted physically as traveling waves, the first travelling in the positive x direction with the speed $1/\sqrt{LC}$ and the second traveling in the negative x direction at the same speed. Substitution of the values for L and C for any configuration of conductors yields the velocity $1/\sqrt{LC}$, which equals $1/\sqrt{\mu\epsilon}$.

The quantity $\sqrt{L/C}$ has the dimensions of ohms, termed the characteristic impedance Z_0 of the line, as shown by Eq. (11). Thus, Z_0 is a real quantity

$$Z_0 = \sqrt{L/C} = \sqrt{\mu/\epsilon}\, F_g \tag{11}$$

(a resistance) and is equal to the wave impedance of the insulating medium, $\sqrt{\mu/\epsilon}$, multiplied by the geometrical factor, F_g, characteristic of the particular configuration of conductors. For the traveling waves of voltage and current, Eqs. (9) and (10), the ratio of voltage to current of the forward-traveling wave is Z_0; that of the backward-traveling wave is $-Z_0$.

In Fig. 5*a*, a source of electrical energy is connected at one end of a transmission line and an electrical load is connected to the other. Electromagnetic energy is propagated from the sending end to the receiving end, and a portion of the energy is reflected back toward the sending end if the load impedance Z_R is different from the characteristic impedance Z_0 of the line. If Z_R equals Z_0, there is no reflection of energy at the load, and in Eqs. (9) and (10), the function f_2, representing leftward-traveling energy, is absent. This is the condition desired when the purpose of the line is to deliver energy from the source of the load. The sending-end impedance of the line is then equal to Z_0.

In addition to impedance matching to reduce reflections (echoes) along a transmission line, it is also necessary to minimize signal distortion, which consists of amplitude and phase (delay) distortion. If the line attenuation is frequency-dependent, then a signal consisting of a group of different-frequency components will undergo amplitude distortion due to the unequal attenuation of each component of the signal. Similarly, if the velocity of propagation along the line is frequency-dependent, then a delay in phase of each component will result in associated phase distortion of the signal.

Signal distortion can be minimized by the use of line loading, which is the addition of series impedances along the line and which is used to adjust the line parameters to obtain the so-called distortionless condition. Under distortionless operation the attenuation and velocity of propagation are independent of frequency. For a discussion of the distortionless line *See* TRANSMISSION LINES.

Instead of loading a line, one may employ equalizing circuits to compensate for the phase distortion along the line.

Short sections of transmission line are sometimes used to provide low-loss reactive impedances and resonant circuits at high frequencies. This is done by open-circuiting or short-circuiting the receiving end of the line to provide complete reflection of the incident energy. A short-circuited low-loss line provides the sending-end impedance shown by Eq. (12),

$$Z_s = jZ_0 \tan(2\pi f l/v) \tag{12}$$

measured in ohms. When l equals $v/4f$, the line is a quarter wavelength long, the argument $(2\pi fl/v)$ of the tangent function in Eq. (12) is $\pi/2$, and Z_s approaches an infinite value. In actual practice, losses keep Z_s to a finite value. However, at high frequencies the quarter wavelength is short and the losses are small, and such a short-circuited quarter-wave section can be used successfully as a low-loss insulator. Such a section is a resonant one and can be used as a substitute for a parallel-resonant LC circuit, for example, a tank circuit for a high-frequency oscillator. At low frequencies the required quarter wavelength is so large that losses impair the performance; also, the length becomes inconveniently great. At a frequency lower than $v/4l$, the sending-end impedance of the short-circuited line is inductive, and at frequencies between $v/4l$ and $v/2l$ the impedance is capacitive. This provides the possibility of using sections of short-circuited line as reactive elements in circuits.

Walter C. Johnson

Bibliography. F. A. Benson and T. M. Benson, *Fields, Waves, and Transmission Lines*, 1991; C. W. Davidson, *Transmission Lines for Communications,* 2d ed., 1989; C. T. A. Johnk, *Engineering Electromagnetic Fields and Waves*, 2d ed., 1987; J. A. Kong, *Electromagnetic Wave Theory*, 3d ed., 2000; P. C. Magnusson et al., *Transmission Lines and Wave Propagation*, 2000; S. Ramo, J. R. Whinnery, and T. Van Duzer, *Fields and Waves in Communication Electronics*, 3d ed., 1993; W. Sinnema, *Electronic Transmission Technology: Lines, Waves, and Antennas,* 2d ed., 1988.

Electromagnetism

The branch of science dealing with the observations and laws relating electricity to magnetism. Electromagnetism is based upon the fundamental observations that a moving electric charge produces a

magnetic field and that a charge moving in a magnetic field will experience a force.

The magnetic field produced by a current is related to the current, the shape of the conductor, and the magnetic properties of the medium around it by Ampère's law. *See* AMPÈRE'S LAW.

The magnetic field at any point is described in terms of the force it exerts on a moving charge at that point. The electrical and magnetic units are defined in terms of the ampere, which in turn is defined from the force of one current upon another.

The association of electricity and magnetism is also shown by electromagnetic induction, in which a changing magnetic field sets up an electric field within a conductor and causes the charges to move in the conductor. *See* EDDY CURRENT; ELECTRICITY; ELECTROMAGNET; ELECTROMAGNETIC INDUCTION; FARADAY'S LAW OF INDUCTION; HALL EFFECT; INDUCTANCE; LENZ'S LAW; MAGNETISM; MAXWELL'S EQUATIONS; RELUCTANCE. Kenneth V. Manning

Electrometallurgy

The branch of process metallurgy dealing with the use of electricity for smelting or refining of metals. The electrochemical effect of an electric current brings about the reduction of metallic compounds, and thereby the extraction of metals from their ores (electrowinning) or the purification of the metals (electrorefining).

In other metallurgical processes, electrically produced heat is utilized in smelting, refining, or alloy manufacturing. For a discussion of electrothermics, that is, the theory and applications of electric heating to metallurgy, *see* ELECTRIC FURNACE; ELECTRIC HEATING; ELECTROCHEMICAL PROCESS; STEEL MANUFACTURE.

Electrowinning. This metallurgical process involves the recovery of a metal, usually from its ore, by dissolving a metallic compound in a suitable electrolyte and reducing it electrochemically through passage of a direct electric current.

Following beneficiation and, sometimes, chemical pretreatment, the metal-bearing constituent of the ore is dissolved in an aqueous solution or in a molten salt. The electrolysis of the purified electrolyte with direct current yields the reduced metal at the cathode (negative electrode of an electrolytic cell), and the nonmetal is the oxidation product at the anode (positive electrode). Since the process is very selective, the purity of the electrowon metal is high, and no further refining is needed.

Water is a suitable solvent for electrolysis of metals less active than zinc and manganese, and sulfuric acid is the preferred leaching agent. The high mobility of the solvated proton provides high-conductivity solutions, and the sulfate ion is electrochemically inert; acid is regenerated by anodic oxidation and recycled for leaching.

Zinc. The process of electrowinning of zinc was developed in 1915 for treating complex ores not amenable to thermal processing, and it is presently the process of choice for most new plants. Changes in the chemistry have been minor, but steady progress has been made in the areas of process instrumentation and control, automation, and pollution-free operation.

The zinc sulfide ores are concentrated, roasted, and leached in several stages with sulfuric acid, and the electrolyte is purified. Lead and silver are insoluble through the acid leach; iron precipitates by oxidation in the neutral leach, and it coprecipitates impurities such as arsenic, antimony, and germanium. Other impurities more noble than zinc, such as copper, cadmium, cobalt, and nickel, are removed by galvanic precipitation on zinc dust.

The reduction potential of zinc is far more cathodic (−0.763 V) than that of the hydrogen ion, and its efficient electrodeposition depends upon maintaining a high hydrogen overvoltage. Metals more noble than zinc deposit on the cathode, and most promote hydrogen evolution and zinc corrosion. This reduces the yield of zinc, which is expressed by the current efficiency—the ratio of the amount of electricity (coulombs) theoretically required to yield a given quantity of metal to the amount actually consumed. Current efficiencies equal to 90% are achieved by bringing the concentration of the most harmful impurities such as germanium, arsenic, and antimony down to 0.01 mg/liter (1.3×10^{-6} oz/gal). *See* OVERVOLTAGE.

The zinc electrolyte, containing 120–160 g/liter (16–21 oz/gal) of zinc as sulfate, is circulated in lead-lined or plastic-lined [for example, poly(vinyl chloride)] concrete or wooden tanks in which the electrolysis proceeds between vertically suspended electrodes. The insoluble anodes are lead that contains small amounts of silver, and the cathodes are aluminum sheets from which the zinc deposits are stripped every 24–28 h. The submerged cathode area varies from 1.3 to 2.6 m^2. The current density varies from 300 to 600 A/m^2. From the Gibbs free-energy change for chemical reaction (1), a reversible voltage

$$ZnSO_4 + H_2O = Zn + {}^1\!/_2O_2 + H_2SO_4 \qquad (1)$$

of 2 V can be calculated. The operating cell voltage is about 3.5 V; the difference is due to the anodic and cathodic overvoltages and the ohmic voltage drop in the electrolyte and at the electrode contacts. The power consumption averages 3.2 kWh/kg zinc for modern plants. This relatively high energy cost is acceptable since special high-grade zinc is produced (99.995%). *See* ZINC METALLURGY.

Copper. Copper electrowinning is similar to zinc electrowinning, but the pyrometallurgical reduction is preferred for sulfide ores. Most of the electrowon copper comes from the direct leaching of low-grade sulfide ores, or from oxide ores. The recent development of liquid ion-exchange reagents specially for copper makes possible an upgrading, through solvent extraction of dilute solutions, from dump or in-place leaching in order to obtain a concentrated solution suitable for electrolysis. The electrowinning of copper is carried out in sulfate solutions containing

from 20 to 40 g/liter (2.7–5.4 oz/gal) of copper and about 100 g/liter (13 oz/gal) of acid at current densities from 160 to 200 A/m^2 (15–19 A/ft^2). The anodes are made of lead, alloyed with antimony, and copper is deposited on copper starting sheets.

Aluminum. Electrowinning has been the only process used for the commercial production of aluminum since it was developed independently by C. M. Hall and by P. L. T. Héroult in 1885. Alumina extracted from bauxite ores and purified by digestion in a caustic soda solution is dissolved in a mixture of cryolite ($AlF_3 \cdot 3NaF$) and other fluorides (for example, CaF_2 and LiF). The molten-salt electrolyte floats on the molten aluminum, which is the cathode, and carbon anodes dip into the bath from above.

The carbon reacts with the oxygen produced by anodic oxidation. The overall cell reaction (2) corre-

$$Al_2O_3 + {}^3\!/_2C = 2Al + {}^3\!/_2CO_2 \qquad (2)$$

sponds to a theoretical reversible voltage of 1.2 V, but the cell operates under a total voltage of about 4.2 V. The heat generated by irreversible electrochemical and ohmic losses maintains the cell temperature at about 960°C (1760°F).

Current densities are about 1 A/cm^2, with total currents of 100,000 A for a modern cell. Alumina is periodically added to the melt to maintain its concentration between 2 and 5%. The current efficiency averages 89%; the main losses are attributed to the recombination of aluminum by reduction of carbon dioxide to carbon monoxide. The energy consumption is about 14 kWh/kg aluminum. *See* ALUMINUM.

Magnesium. The electrowinning of magnesium, another well-established process, accounts for about 80% of the metal output. The electrolyte is a mixture of chlorides of potassium, sodium, calcium, and magnesium. Since the reduced metal has a lower specific gravity than the electrolyte has, it floats to the surface; the cells are designed so that it does not come in contact with air or with the chlorine gas produced at the anode.

Other metals. Cadmium, and in smaller quantities, thallium and indium, are recovered by electrolysis in sulfate solutions. Cobalt is also electrowon by a similar process, and some of the nickel production is achieved by electrolysis in sulfate-chloride solutions. Pure metallic chromium and manganese are produced by electrowinning; these metals are rather difficult to electrodeposit, and the acid produced at the anode must be kept away from the cathode by a diaphragm. Antimony is electrowon from a sodium sulfide electrolyte, and gallium is produced by electrolysis of a caustic solution.

Metals more active than manganese cannot be reduced cathodically from aqueous solutions, but most can be electrowon from mixtures of molten salts. Usually, a compound of the metal is dissolved in salts of still more active metals in order to improve the conductivity and lower the melting point of the melt. Solid metallic deposits are difficult to recover, and all the commercially successful processes yield the liquid metal. The solubility and reactivity of the reduced metal in the melt is often important, and the cells must be designed to protect the reduced metal from contact with the anodic reaction product or with air. The cell electrolyte cannot be easily recycled, and the feed must be carefully purified and dehydrated.

All the sodium metal produced today comes from electrolysis of a molten chloride mixture between a steel cathode and a graphite anode separated by a perforated steel diaphragm. Lithium is produced by a similar process. The other alkali and alkaline-earth metals, as well as the rare earths, can be electrowon from molted chlorides. Processes have been developed for electrowinning titanium, tantalum, and niobium in molten salts, and uranium, thorium, and zirconium have been obtained in the laboratory. Beryllium and boron are also produced by electrolysis.

Electrorefining. This is a purification process in which an impure metal anode is dissolved electrochemically in a solution of a salt of the metal being refined; the pure metal is recovered by electrodeposition at the cathode.

Electrorefining is the most economical method for securing the high purity required for many uses of nonferrous metals. Since the same electrochemical reaction proceeds in opposite directions at the anode and the cathode, the overall chemical change is a small change in the activity of the metal at the two electrodes, and the reversible cell voltage is practically zero. The anodic and cathodic overvoltages, required for an economical reaction rate, are small, and the operating cell voltage consists mainly of the ohmic voltage drop through the electrolyte and at the electrode contacts. The power consumption is moderate, and electrorefining is less sensitive to the cost of electric power than other electrometallurgical processes are.

Electochemical refining is a more efficient purification process than other chemical methods are because it is very selective. In particular, for metals such as copper, silver, gold, and lead, the operating electrode potential is close to the reversible potential, and a sharp separation is accomplished, both at the anode, where more noble metals do not dissolve, and at the cathode, where more active metals do not deposit.

The first commercial electrolytic copper refinery was established in 1871, and today most of the copper is electrorefined. Minor quantities of some impurities lower the electrical conductivity of copper markedly, and the refining process ensures that the metal will meet the specifications of the electrical industry. Furthermore, silver, gold, and other precious metals are removed by the electrolytic refining, and their recovery is thus an economic asset of the process.

The impure copper is cast in the shape of anodes approximately 3.0 × 3.3 ft (0.9 × 1.0 m) and 1.4–1.8 in. (3.5–4.5 cm) thick. The starting cathodes are pure copper sheets, prepared by electrodeposition. The electrodes are placed vertically in cells which are lead-lined or plastic-lined wooden or concrete tanks. The anodes are replaced at regular intervals of

20–28 days, and two successive cathodes are produced during the same period. The solution contains about 45 g/liter (6 oz/gal) of copper as copper sulfate and approximately 200 g/liter (26 oz/gal) of sulfuric acid. The temperature is maintained at 55–60°C (131–140°F) to lower the resistance of the electrolyte which is circulated through the cell. The current density is usually 200–250 A/m^2, and the current efficiency averages 95%. The cell voltage is 0.2–0.3 V, and the power consumption varies from 0.18 to 0.25 kWh/kg copper.

Metals less noble than copper, such as iron, nickel, cobalt, zinc, and manganese, dissolve from the anode. Lead is oxidized, and it precipitates as lead sulfate. Other impurities such as arsenic, antimony, and bismuth partly remain as insoluble compounds in the slimes and partly dissolve as complexes in the electrolyte. Precious metals do not dissolve at the potential of the anode, and they remain as metals in the slimes, along with insoluble selenium and tellurium compounds and some metallic copper particles that fall from the anode. Metals less noble than copper do not deposit at the cathode. In particular, nickel and arsenic accumulate in the electrolyte, and their concentration is controlled by partial withdrawal and purification of the solution.

Nickel is electrorefined in a sulfate-chloride electrolyte. Its anodic dissolution requires a significant overvoltage, and both copper and nickel dissolve at the anode potential. To prevent plating of the dissolved copper at the cathode, a diaphragm cell is used, and the anodic solution is circulated through a purification circuit before entering the cathode compartment. Cobalt is electrorefined by a similar process.

Very high-purity lead is produced by electrorefining in which a fluosilicate electrolyte is used, and a sulfate bath is used for purifying tin. Silver is electrorefined in a copper and silver nitrate electrolyte, and gold is refined by electrolysis in a chloride bath.

The electrorefining of many metals can be carried out with molten-salt electrolytes, but the processes are usually expensive, and the only industrial application is the electrorefining of aluminum by the "three-layer" process. The density of the electrolyte is adjusted so that a pure molten aluminum cathode will float on the molten salt, which in turn will float on an anode consisting of a molten mixture of impure aluminum and copper. *See* PYROMETALLURGY.

Paul Duby

Bibliography. A. T. Kuhn, *Industrial Electrochemical Processes,* 1971; A. T. Kuhn (ed.), *Techniques in Electrochemistry, Corrosion, and Metal Finishing: A Handbook*, 1987; D. Pletcher and F. C. Walsh, *Industrial Electrochemistry*, 2d ed., 1993.

Electrometer

A highly sensitive instrument which measures all or some of the following variables: current, charge, voltage, and resistance. There are two classes of electrometers, mechanical and electronic. The mechanical instruments have been largely replaced by electronic types.

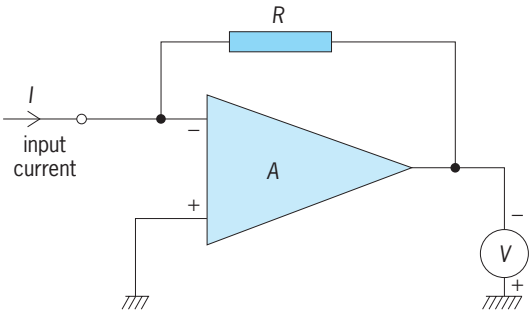

A common arrangement for an electronic electrometer.

Mechanical. These instruments, with origins in the nineteenth century and earlier, rely for their operation on the mechanical forces associated with electrostatic fields. The negligible current drawn by many electrometers contributes to their high sensitivity, particularly in the context of charge and current measurement.

In common with other electrostatic instruments, electrometers are classified as either attracted-disk or symmetrical types. Attracted-disk instruments, in which the attractive force between two plates, with a potential difference between them, is measured in terms of the fundamental units of mass and length, are sometimes termed absolute electrometers. Attracted-disk instruments are widely used as electrostatic voltmeters for measuring potentials greater than 1 kilovolt. The quadrant electrometer was a particularly sensitive instrument based on a horizontal vane suspended by a torsion fiber so that it could rotate between fixed electrodes. *See* ELECTROSCOPE; ELECTROSTATICS; VOLTMETER.

Electronic. Electronic types utilize some form of electronic amplifier, typically an operational amplifier with a field-effect-transistor input stage to minimize the input current. In the **illustration** the amplifier A has a current measuring resistor R providing negative feedback. The unknown input current I can be found from the voltmeter reading V, as $I = V/R$. Charge Q can be measured if R is replaced by a capacitor C. For this case, $Q = VC$. *See* AMPLIFIER; TRANSISTOR.

In the most sensitive applications, problems arise due to drift in the amplifier characteristics and electrical noise present in the circuit components. To obviate these effects, electrometers employing a vibrating capacitor or varactor diodes are used. The signal to be measured is converted to an alternating-current (ac) signal and subsequently amplified by an ac amplifier which is less susceptible to drift and noise. The amplified signal is finally reconverted to direct current (dc). *See* ELECTRICAL NOISE; VARACTOR.

The most sensitive electrometers commercially available can resolve currents of 0.2 femtoampere under favorable circumstances. *See* CURRENT MEASUREMENT; ELECTRICAL MEASUREMENTS; VOLTAGE MEASUREMENT.

R. W. J. Barker

Electromotive force (cells)

The voltage or electric potential difference across the terminals of a cell when no current is drawn from it. The electromotive force (emf) is the sum of the electric potential differences produced by a separation of charges (electrons or ions) that can occur at each phase boundary (or interface) in the cell. The magnitude of each potential difference depends on the chemical nature of the two contacting phases. Thus, at the interface between two different metals, some electrons will have moved from the metal with a higher free energy of electrons to the metal with a lower free energy of electrons. The resultant charge separation will produce a potential difference, just as charge separation produces a voltage across a capacitor; at equilibrium this exactly opposes further electron flow. Similarly, potential differences can be produced when electrons partition across a metal|solution interface or metal|solid interface, and when ions partition across a solution|membrane|solution interface.

Origin. How a cell emf is composed of the sum of interfacial potential differences is shown by the Daniell cell (see **illus.**) and scheme (1);

$$\begin{aligned} &\mathrm{Cu'} \underset{a}{|} \mathrm{Zn} \underset{b}{|} \text{aq soln (1)} : \\ &\qquad \mathrm{ZnSO_4} \underset{c}{\vdots} \text{aq soln (2)} : \mathrm{CuSO_4} \underset{d}{|} \mathrm{Cu} \end{aligned} \tag{1}$$

aq soln denotes an aqueous solution, a solid line indicates a phase boundary, and the broken line a porous barrier permeable to all the ions in the adjacent solutions. The copper connector to the zinc electrode is denoted Cu′. The barrier prevents physical mixing of the zinc sulfate ($ZnSO_4$) and copper sulfate ($CuSO_4$) solutions.

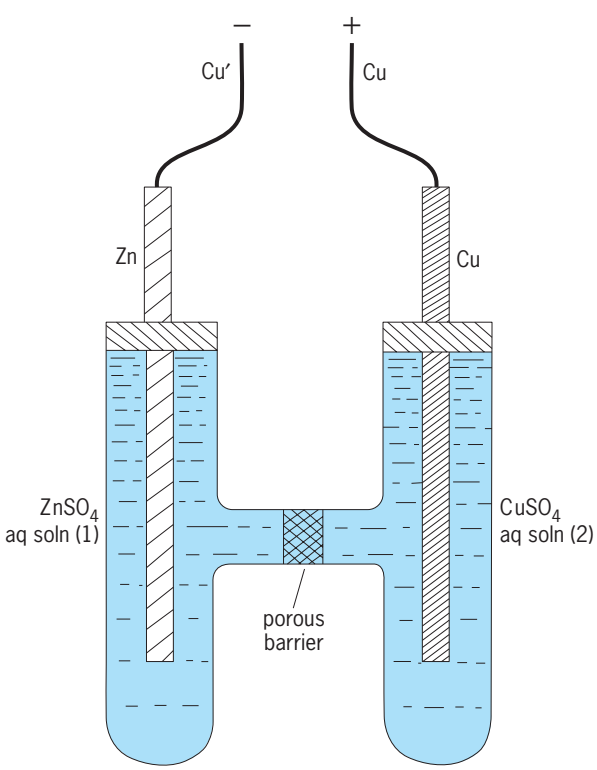

Cross-section diagram of Daniell cell. (*After W. J. Moore, Physical Chemistry, 5th ed., Longman, 1974*)

The cell emf is the open-circuit (that is, zero current) potential difference measured between the two Cu leads (any potential-measuring device must ultimately measure the potential difference between two chemically identical phases, in this example the Cu and Cu′ phases). An expression can be written for the cell emf in terms of the electrochemical potentials of the species involved in the interfacial equilibria. The electrochemical potential $\bar{\mu}_j^\alpha$ of species j in phase α is defined by Eq. (2),

$$\bar{\mu}_j^\alpha = \mu_j^\alpha + z_j F \phi^\alpha \tag{2}$$

where μ_j^α is the chemical potential of species j in phase α, z_j is the number of electronic charges (with sign) on the species, F is the Faraday constant (96485.3 coulombs for Avogadro's number of electrons), and ϕ^α is the inner potential of phase α; $z_j F \phi^\alpha$ is the electrical work required to move $z_j F$ coulombs of charge into phase α from infinite distance in a vacuum. For a species in solution, Eq. (2) can be written as Eq. (3),

$$\bar{\mu}_j^\alpha = \mu_j^{\circ\alpha} + RT \ln a_j^\alpha + z_j F \phi^\alpha \tag{3}$$

where $\mu_j^{\circ\alpha}$ is the standard chemical potential of species j in phase α, R is the gas constant (8.3145 joules/mole/degree Kelvin), T is the temperature in Kelvin, and a_j^α is the activity of species j in phase α. For a neutral species (that is, when $z_j = 0$), Eq. (4) applies.

$$\bar{\mu}_j^\alpha = \mu_j^\alpha = \mu_j^{\circ\alpha} + RT \ln a_j^\alpha \tag{4}$$

For a fully ionized solution species at low concentrations, a_j^α is well approximated by the concentration c_j^α; the activity of a pure phase (for example, zinc and copper in the Daniell cell) is defined as unity.

The value of $\bar{\mu}_j$ must be identical for any species j equilibrated between two different phases. Thus for equilibrium of species j between phases α and β, Eq. (5) applies.

$$\bar{\mu}_j^\alpha = \bar{\mu}_j^\beta \tag{5}$$

More generally, for any interfacial equilibrium between phases α and β, Eq. (6) applies; ν_j is the stoichiometric number of species j (ν_j is positive for products and negative for reactants), and $\bar{\mu}_j^{\text{phase}}$ is the electrochemical potential of species j in the phase in which it is located. Thus, for the interfacial equilibrium $Cu^{2+} + 2e^- \rightleftarrows Cu$ at interface d in scheme (1), Eq. (6) becomes Eq. (7).

$$\sum_i \nu_j \bar{\mu}_j^{\text{phase}} = 0 \tag{6}$$

$$\bar{\mu}_{Cu^{2+}}^{\text{soln(2)}} + 2\bar{\mu}_{e^-}^{\text{Cu}} = \bar{\mu}_{\text{Cu}}^{\text{Cu}} \tag{7}$$

Similar equations can be used to quantify the potential differences across each interface in

scheme (1), assuming that the potential difference across interface c is zero (a good approximation when the separator is equally permeable to all the ions in the two adjacent phases). For interfaces a, b, and d of scheme (1), Eqs. (8) can be written.

Interface a (partitioning of electrons between Cu′ and Zn phases):

$$\bar{\mu}_{e^-}^{\mathrm{Cu}'} = \bar{\mu}_{e^-}^{\mathrm{Zn}} \tag{8a}$$

Interface b (equilibrium: $Zn^{2+} + 2e^- \rightleftarrows Zn$):

$$\bar{\mu}_{\mathrm{Zn}^{2+}}^{\mathrm{soln}(1)} + 2\bar{\mu}_{e^-}^{\mathrm{Zn}} = \bar{\mu}_{\mathrm{Zn}}^{\mathrm{Zn}} \tag{8b}$$

Interface d (equilibrium: $Cu^{2+} + 2e^- \rightleftarrows Cu$)

$$\bar{\mu}_{\mathrm{Cu}^{2+}}^{\mathrm{soln}(2)} + 2\bar{\mu}_{e^-}^{\mathrm{Cu}} = \bar{\mu}_{\mathrm{Cu}}^{\mathrm{Cu}} \tag{8c}$$

Rearranging Eqs. (8) gives Eq. (9).

$$\bar{\mu}_{e^-}^{\mathrm{Cu}} - \bar{\mu}_{e^-}^{\mathrm{Cu}'} = \bar{\mu}_{e^-}^{\mathrm{Cu}} - \bar{\mu}_{e^-}^{\mathrm{Zn}} = \frac{1}{2}\left[\bar{\mu}_{\mathrm{Cu}}^{\mathrm{Cu}} - \bar{\mu}_{\mathrm{Cu}^{2+}}^{\mathrm{soln}(2)}\right] - \frac{1}{2}\left[\bar{\mu}_{\mathrm{Zn}}^{\mathrm{Zn}} - \bar{\mu}_{\mathrm{Zn}^{2+}}^{\mathrm{soln}(1)}\right] \tag{9}$$

Since $\mu_{e^-}^{\circ\mathrm{Cu}} = \mu_{e^-}^{\circ\mathrm{Cu}'}$ and for an electron $z_{e^-} = -1$, Eq. (3) can be used to write Eq. (10).

$$\bar{\mu}_{e^-}^{\mathrm{Cu}} - \bar{\mu}_{e'}^{\mathrm{Cu}'} = -F(\phi^{\mathrm{Cu}} - \phi^{\mathrm{Cu}'}) \tag{10}$$

The cell potential, E_{cell}, is simply the sum of all the equilibrium interfacial potential differences, as in Eq. (11), where it has been assumed that $\phi^{\mathrm{soln}(2)} - \phi^{\mathrm{soln}(1)} - 0$. Combining Eqs. (3), (10) and (11) (and

$$E_{\mathrm{cell}} = \left(\phi^{\mathrm{Cu}} - \phi^{\mathrm{soln}(2)}\right) + \left(\phi^{\mathrm{soln}(1)} - \phi^{\mathrm{Zn}}\right) + (\phi^{\mathrm{Zn}} - \phi^{\mathrm{Cu}'}) = (\phi^{\mathrm{Cu}} - \phi^{\mathrm{Cu}'}) \tag{11}$$

noting that the copper and zinc ions have a charge of 2+ and that the activity of the metal phases is unity) gives Eq. (12),

$$E_{\mathrm{cell}} = (\phi^{\mathrm{Cu}} - \phi^{\mathrm{Cu}'}) = -\frac{1}{F}\left[\bar{\mu}_{e^-}^{\mathrm{Cu}} - \bar{\mu}_{e^-}^{\mathrm{Cu}'}\right] = \frac{1}{2F}\left[\mu_{\mathrm{Cu}^{2+}}^{\circ\,\mathrm{soln}(2)} - \mu_{\mathrm{Cu}}^{\circ} + \mu_{\mathrm{Zn}}^{\circ} - \mu_{\mathrm{Zn}^{2+}}^{\circ\,\mathrm{soln}(1)}\right] + \frac{RT}{2F}\ln a_{\mathrm{Cu}^{2+}}^{\mathrm{soln}(2)} - \frac{RT}{2F}\ln a_{\mathrm{Zn}^{2+}}^{\mathrm{soln}(1)} \tag{12}$$

where E_{cell} is the potential difference between the two copper leads [scheme (1)] measured at open circuit (that is, when no current passes through the cell) and is therefore the emf of the cell. Dropping the superscripts denoting the solutions, Eq. (12) can be rewritten as Eq. (13),

$$E_{\mathrm{cell}} = \left[E_{\mathrm{Cu}^{2+}|\mathrm{Cu}}^{\circ} + \frac{RT}{2F}\ln a_{\mathrm{Cu}^{2+}}\right] - \left[E_{\mathrm{Zn}^{2+}|\mathrm{Zn}}^{\circ} + \frac{RT}{2F}\ln a_{\mathrm{Zn}^{2+}}\right] \tag{13}$$

where $E_{\mathrm{Cu}^{2+}|\mathrm{Cu}}^{\circ}$ and $E_{\mathrm{Zn}^{2+}|\mathrm{Zn}}^{\circ}$ are termed the standard electrode potentials of the copper and zinc half cells. Equation (13) is an example of a Nernst equation, which relates cell emf to the activities of the cell constituents. *See* CHEMICAL EQUILIBRIUM; CHEMICAL THERMODYNAMICS; ELECTROMOTIVE FORCE (EMF); THERMODYNAMIC PRINCIPLES.

Half cells. It is convenient to describe any electrochemical cell in terms of half cells. A half cell consists of an oxidant (Ox) and reductant (Red) such that $\mathrm{Ox} + ne^- \rightleftarrows \mathrm{Red}$; species Ox and Red are commonly referred to as a redox couple. The Daniell cell, for example, comprises the two half cells $Cu^{2+} + 2e^- \rightleftarrows Cu$ (redox couple is $Cu^{2+}|Cu$) and $Zn^{2+} + 2e^- \rightleftarrows Zn$ (redox couple is $Zn^{2+}|Zn$) with the half-cell potentials $E_{\mathrm{Cu}^{2+}|\mathrm{Cu}}$ and $E_{\mathrm{Zn}^{2+}|\mathrm{Zn}}$ given by Eqs. (14) and (15). Such expressions are useful only if the E°

$$E_{\mathrm{Cu}^{2+}|\mathrm{Cu}} = E_{\mathrm{Cu}^{2+}|\mathrm{Cu}}^{\circ} + \frac{RT}{2F}\ln a_{\mathrm{Cu}^{2+}} \tag{14}$$

$$E_{\mathrm{Zn}^{2+}|\mathrm{Zn}} = E_{\mathrm{Zn}^{2+}|\mathrm{Zn}}^{\circ} + \frac{RT}{2F}\ln a_{\mathrm{Zn}^{2+}} \tag{15}$$

value for each half cell is known. Values of E° can be assigned to any given half cell by arbitrarily specifying that $E^{\circ}_{\mathrm{H}^+}$ H_2 [the E° for the standard hydrogen electrode (SHE) half cell, $H^+(\mathrm{aq}, a = 1) + e^- \rightleftarrows {}^1/_2H_2(\mathrm{g}, 1\ \mathrm{atm})$] is zero. The temperature dependence $dE^{\circ}_{\mathrm{H}^+|\mathrm{H}_2}/dT$ is also specified as zero. E° values versus the SHE for selected half cells are given in the **Table 1**. In principle, any E° (and its temperature dependence) can be measured directly versus the SHE, or another half cell whose electrode potential and temperature dependence have been determined. Such an electrode is termed a reference electrode. The reference electrode is a half cell designed so that its potential is stable and reproducible, and it neither contaminates nor is contaminated by the medium in which it is immersed. Two convenient reference electrodes commonly used in aqueous systems are the saturated calomel and the silver|silver chloride electrodes (see table). *See* ELECTROCHEMISTRY; ELECTRODE POTENTIAL; OXIDATION-REDUCTION; REFERENCE ELECTRODE.

E° values and thermodynamics. The E° values associated with the two half cells that make up a cell provide fundamental thermodynamic information about the chemical reaction between the redox couples of the two half cells. Returning to the example of the Daniell cell and combining Eqs. (12) and (13) yields

Selected standard electrode potentials at 25°C

Electrode reaction	E°/V
$Li^+(aq) + e^- \rightleftarrows Li(s)$	−3.045
$Zn^{2+}(aq) + 2e^- \rightleftarrows Zn(s)$	−0.763
$2H^+(aq) + 2e^- \rightleftarrows H_2(g)$	0
$Hg_2Cl_2(s) + 2e^- \rightleftarrows 2Hg(s) + 2Cl^-$ (sat aq KCl)	0.246
$AgCl(s) + e^- \rightleftarrows Ag(s) + Cl^-(aq)$	0.222
$Cu^{2+}(aq) + 2e^- \rightleftarrows Cu(s)$	0.337
$Fe(CN)_6^{3-}(aq) + e^- \rightleftarrows Fe(CN)_6^{4-}(aq)$	0.69
$O_2(g) + 4H^+(aq) + 4e^- \rightleftarrows 2H_2O$	1.223
$F_2(g) + 2H^+ + 2e^- \rightleftarrows 2HF(aq)$	3.06

Eq. (16). The term on the left-hand side is the stan-

$$E^\circ_{Cu^{2+}|Cu} - E^\circ_{Zn^{2+}|Zn} = \frac{1}{2F}\left[\mu^{\circ\,soln(2)}_{Cu^{2+}} - \mu^\circ_{Cu} + \mu^\circ_{Zn} - \mu^{\circ\,soln(1)}_{Zn^{2+}}\right] \quad (16)$$

dard cell emf E°_{cell} (that is, the value of E_{cell} when all the cell constituents are in their standard states of unit activity), and the bracketed term on the right-hand side is $-\Delta G^\circ$, where ΔG° is the standard Gibbs energy change of the cell reaction $Cu^{2+} + Zn \rightarrow Cu + Zn^{2+}$. Hence Eq. (16) can be written as Eq. (17).

$$E^\circ_{cell} = -\frac{\Delta G^\circ}{2F} \quad (17)$$

More generally, for the two half-cell reactions $Ox_1 + n_1e^- \rightleftarrows Red_1$ and $Ox_2 + n_2e^- \rightleftarrows Red_2$, E°_{cell} is defined by Eq. (18). The cell reaction (19) involves the trans-

$$E^\circ_{cell} = E^\circ_{Ox_1|Red_1} - E^\circ_{Ox_2|Red_2} \quad (18)$$

$$n_2Ox_1 + n_1Red_2 \rightleftarrows n_2Red_1 + n_1Ox_2 \quad (19)$$

fer of n_1n_2 electrons, and Eq. (17) takes the general form (20), where $n = n_1n_2$. If the cell is not in its standard state, the relation becomes Eq. (21).

$$E^\circ_{cell} = -\frac{\Delta G^\circ}{n_1n_2F} = -\frac{\Delta G^\circ}{nF} \quad (20)$$

$$E_{cell} = -\frac{\Delta G}{nF} \quad (21)$$

Using the general thermodynamic relation $\Delta G^\circ = -RT \ln K_{eq}$, it is also possible to find the equilibrium constant K_{eq} of the cell reaction from the relation (22).

$$K_{eq} = \frac{[Ox_2]^{n_1}[Red_1]^{n_2}}{[Ox_1]^{n_2}[Red_2]^{n_1}} = \exp\left[-\frac{\Delta G^\circ}{RT}\right] = \exp\left[\frac{n_1n_2FE^\circ_{cell}}{RT}\right] \quad (22)$$

See FREE ENERGY.

Membrane potentials. When two ionic solutions of different composition are separated by a membrane, a potential difference can develop across the membrane. The complete cell requires electrical contacts with the solutions on each side of the membrane, accomplished with reference electrodes, as in scheme (23),

re(lhs) | soln(lhs) | membrane | soln(rhs) | re(rhs) (23)

where re(lhs) and re(rhs) are the reference electrodes in the left and right solutions, respectively. The cell emf will be as in Eq. (24).

$$E_{cell} = E_{re(rhs)} + \phi^{soln(rhs)} - \phi^{soln(lhs)} - E_{re(lhs)} \quad (24)$$

The potential difference across the membrane, $\phi^{soln(rhs)} - \phi^{soln(lhs)}$, is a function of the membrane properties as well as the composition of the adjacent solutions. The simplest example occurs when only one particular ion (such as the hydrogen ion, H^+) is transported across the membrane. The transported ion will tend to move from the side with high concentration to the side with low concentration; however, in the process of doing so, charge is separated across the membrane, producing a potential difference that exactly counters the ion transport. The membrane potential difference is given by Eq. (25),

$$\phi^{soln(rhs)} - \phi^{soln(lhs)} = \frac{RT}{z_jF}\ln\frac{a_j^{soln(lhs)}}{a_j^{soln(rhs)}} \approx \frac{RT}{z_jF}\ln\frac{c_j^{soln(lhs)}}{c_j^{soln(rhs)}} \quad (25)$$

where a_j and c_j are the activity and concentration of the transported ion. The cell emf is given by Eq. (26).

$$E_{cell} = E_{re(rhs)} + \frac{RT}{z_jF}\ln\frac{c_j^{soln(lhs)}}{c_j^{soln(rhs)}} - E_{re(lhs)} \quad (26)$$

If the concentration on one side, for example, $c_j^{soln(lhs)}$ is kept constant, E_{cell} will reflect any variation of the other concentration, $c_j^{soln(rhs)}$, and the system becomes a sensor for that ion.

As long as the membrane transports only a single ionic species, the potential difference is thermodynamically determined. As soon as the membrane transports more than one ionic species, irreversible thermodynamics come into play; the membrane potential difference will exhibit a complicated dependence on the concentrations and mobilities of the ions within the membrane as well as on their concentrations in the adjacent solutions. In certain circumstances, the response is still selective for one ion and the membrane electrode has analytical utility. The glass electrode used for measurement of pH, in which a thin glass membrane responds selectively to the H^+ ion, is an electrode of this type.

Potential differences known as liquid junction potentials arise where two different ionic solutions make contact through a permeable separator such as a glass frit. These interfere with thermodynamic measurements and can be minimized by interposing a salt bridge, that is, a concentrated immobilized solution of potassium chloride (KCl), or other salt in which the cation and anion have nearly identical mobilities, between the two solutions. Numerous theories based on irreversible thermodynamics have evolved to estimate the magnitude of liquid junction potentials. *See* BIOPOTENTIALS AND IONIC CURRENTS; ION-SELECTIVE MEMBRANES AND ELECTRODES.

Uses. The emf produced by a single cell or a set of cells in series (a battery) is used as a dc power source for a wide array of applications, ranging from powering wristwatches to emergency power supplies. The emf of the saturated Weston cell (scheme 27;

Hg(liquid) | Hg_2SO_4(solid) | $CdSO_4$ (sat aq) | Cd(Hg) (27)

sat aq = saturated aqueous solution) still serves as a high-level voltage reference for the National Institute of Standards and Technology. However, Josephson

arrays are now considered the most precise voltage references, and Zener diodes are used to produce reference voltages for many laboratory and field voltage measurement devices.

There are also a few photoelectrochemical cells with possible application in solar energy conversion, such as the Grätzel cell (scheme 28; nonaq = non-

$$TiO_2(\text{solid}) \mid \text{adsorbed } Ru^{2G} \text{ dye} \mid I_2,\ I^- \text{ (nonaq)} \mid C \quad (28)$$

aqueous solution), where the cell emf is produced by absorption of visible light.

The emf of a cell can also be used as an indicator of chemical composition. Devices that depend on the measurement of an open circuit-cell potential work well when the device is sensitive to a single analyte, but they are notoriously sensitive to interferences. An example is the so-called alkaline error that occurs when glass electrodes are used to measure very high pH values. For complex systems containing several species that can undergo electrochemical reaction, individual E° values can be determined using electroanalytical techniques that involve controlling the potential applied to a cell with measurement of the resultant current. Techniques such as polarography and cyclic voltammetry, for example, involve changing the potential of an indicator electrode and observing a wave or peak in the current at the redox potential of the species in solution; the height of the wave or peak indicates the concentration of that species. *See* BATTERY; ELECTROCHEMICAL TECHNIQUES; FUEL CELL; POLAROGRAPHIC ANALYSIS.

Cell voltages when current passes. E_{cell} values, which are thermodynamic quantities, do not give any information about the kinetics of the electrode reactions. When current passes through a cell, the cell voltage (that is, the potential difference between the terminals) will differ from the emf because of resistive losses within the cell (a function of cell design) and because the kinetics of the electrode reactions are not fast enough to sustain thermodynamic (nernstian) behavior at high rates of reaction. These kinetics are very dependent on temperature and the nature of the electrode material. When a cell is discharged through an external load, the cell voltage will be less than the emf; if the direction of current flow is reversed (using an external power supply), the cell voltage will be greater than the emf. In both cases the greater the current, the greater the deviation of the cell voltage from the emf. In some cases the electrode reactions are so slow that even the open-circuit cell voltage may not be a reliable measure of the emf.

For example, a cell comprising the two half cells $2H^+(aq) + 2e^- \rightleftarrows H_2(g)$ and $O_2(g) + 4H^+(aq) + 4e^- \rightleftarrows 2H_2O$ (where g = gas), both under standard conditions, has a cell emf of 1.223 V at 25°C (see table). A highly sophisticated version of this cell, the hydrogen/oxygen fuel cell, is designed to minimize internal cell resistance, and has electrodes tailored to maximize the rates of the electrode reactions; nevertheless, the voltage under load is usually less than 0.8 V, with most of the voltage loss occurring at the oxygen electrode. When an external power supply is used to electrolyze water, the cell voltage required to produce H_2 and O_2 at a reasonable rate is $\gtrsim$1.6 V; and this again depends critically on the cell design and the choice of electrodes. Mary D. Archer; Stephen W. Feldberg

Bibliography. R. G. Compton and G. H. W. Sanders, *Electrode Potentials*, Oxford University Press, 1996; D. R. Crow, *Principles and Application of Electrochemistry*, 4th ed., Blackie Academic & Professional, 1994; A. R. Denaro, *Elementary Electrochemistry*, 2d ed., 1971; D. J. G. Ives and G. J. Janz, *Reference Electrodes: Theory and Practice*, Academic Press, 1961; J. Koryta, *Ion-Selective Electrodes*, Cambridge University Press, 1975; W. J. Moore, *Physical Chemistry*, 5th ed., Longman, 1974.

Electromotive force (emf)

A measure of the strength of a source of electrical energy. The term is often shortened to emf. It is not, of course, a force in the usual mechanical sense (and for this reason has sometimes been called electromotance), but it is a conveniently descriptive term for the agency which drives current through an electric circuit. In the simple case of a direct current I (measured in amperes) flowing through a resistor R (in ohms), Ohm's law states that there will be a voltage drop (or potential difference) of $V = IR$ (in volts) across the resistor. To cause this current to flow requires a source with emf (also measured in volts) $E = V$. More generally, Kirchhoff's voltage law states that the sum of the source emf's taken around any closed path in an electric circuit is equal to the sum of the voltage drops. This is equivalent to the statement that the total emf in a closed circuit is equal to the line integral of the electric field strength around the circuit. *See* ELECTRIC CURRENT; ELECTRIC FIELD; ELECTRICAL RESISTANCE; OHM'S LAW.

A source of emf E delivering a current I supplies electrical energy at a rate of EI (in watts). Therefore, the emf in a circuit is numerically equal to the work done in carrying a charge of 1 coulomb around the circuit.

Open-circuit emf. A source such as a battery has an internal resistance which causes an internal voltage drop when a current is drawn from it. Consequently the emf at the external terminals varies with the current. By extrapolating to zero current, the open-circuit emf may be derived. It is the open-circuit emf which is usually said to be the emf of a source; together with the value of the internal resistance, it enables the performance of the source in any circuit to be calculated.

Types of emf. An emf may be steady (direct), as for a battery, or time-varying, as for a charged capacitor discharging through a resistor. A class of time-varying sources of very great practical importance is that of generators of alternating (sinusoidal) emf. Emf's may be generated by a variety of physical, chemical, and

biological processes. Some of the more important are:

1. Electrochemical reactions, as used in direct-current (dc) batteries, in which the emf results from the reactions between electrolyte and electrodes. *See* BATTERY; ELECTROCHEMISTRY; ELECTROMOTIVE FORCE (CELLS).

2. Electromagnetic induction, in which the emf results from a change in the magnetic flux linking the circuit. This finds application in alternating-current rotary generators and transformers, providing the basis for the electricity supply industry. *See* ALTERNATING-CURRENT GENERATOR; ELECTROMAGNETIC INDUCTION; FARADAY'S LAW OF INDUCTION; TRANSFORMER.

3. Thermoelectric effects, in which a temperature difference between different parts of a circuit produces an emf. The main use is for the measurement of temperature by means of thermocouples; there are some applications to electric power generation. *See* THERMOCOUPLE; THERMOELECTRIC POWER GENERATOR; THERMOELECTRICITY.

4. The photovoltaic effect, in which the absorption of light (or, more generally, electromagnetic radiation) in a semiconductor produces an emf. This is widely used for scientific purposes in radiation detectors and also, increasingly, for the generation of electric power from the Sun's radiation. *See* PHOTOVOLTAIC EFFECT; RADIOMETRY; SOLAR CELL.

5. The piezoelectric effect, in which the application of mechanical stress to certain types of crystal generates an emf. There are applications in sound recording, in ultrasonics, and in various types of measurement transducer. *See* MICROPHONE; PIEZOELECTRICITY; TRANSDUCER; ULTRASONICS. A. Earle Bailey

Electromyography

The detection and recording of electrical activity generated by muscle fibers. The basic elements of motor control in the body are the motor units which comprise motor neurons in the brainstem or spinal cord, their axons, and from ten to several hundred muscle fibers supplied by each motor neuron. Motor units vary in the size and properties of their motoneurons, the sizes and conduction velocities of their axons, the morphology of their nerve muscle junctions, and the structure and physiological properties of the muscle fibers supplied by each motor neuron. Voluntary activity recruits motor units in an orderly manner, beginning with smaller, twitch units with more slowly conducting axons recruited in weak contractions and progressing to much-harder-to-recruit, larger and faster-twitch motor units with more rapidly conducting axons.

Impulses originating in single motoneurons in response to various command signals from the central nervous system conduct to the periphery of the unit, normally causing all the muscle fibers in the unit to discharge. The electrical activity generated by the more or less synchronous discharges of all the muscle fibers in the unit may be detected by recording electrodes on the skin surface or by needles inserted into the muscle (see **illus.**). Such potentials are called motor-unit action potentials and reflect the electrical activity generated by the whole motor unit.

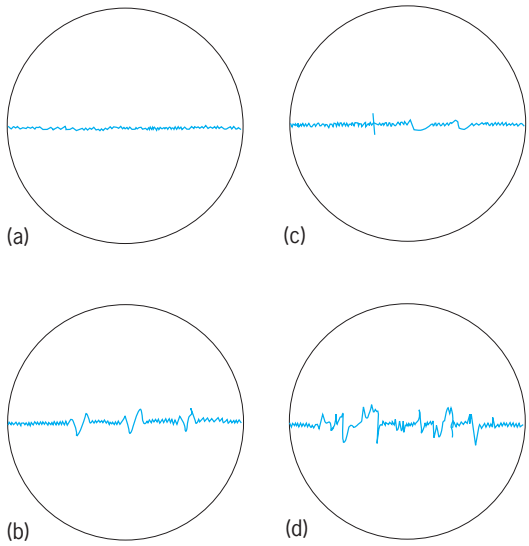

Record of synchronous and asynchronous muscle discharge. (*a*) Normal resting muscle. (*b*) Partial contraction. (*c*) Degenerating muscle. (*d*) Dystrophic muscle.

Diseases affecting motor neurons are sometimes accompanied by spontaneous discharges of the axons (fasciculations) which, if near the skin, may be visible as twitches in the muscle. Additionally, degeneration of motor axons may leave some muscle fibers deprived of their normal innervation, some of which spontaneously fire. Such single muscle-fiber discharges are called fibrillations and are readily detected for diagnostic purposes by needle electrodes inserted into the muscle.

Electromyography may also be used to study primary muscle diseases such as the muscular dystrophies, and a wide variety of other metabolic inflammatory and congenital myopathies affecting the muscle fibers rather than motor neurons or their axons. *See* BIOPOTENTIALS AND IONIC CURRENTS; ELECTRODIAGNOSIS. William F. Brown

Electron

The negatively charged constituent of ordinary matter, responsible for the chemical properties of matter, and the carrier of electricity. The electron is the lightest known particle that possesses an electric charge. Its rest mass m_e is 9.1×10^{-31} kg (2.0×10^{-30} lb), about 1/1836 of the mass of the proton or neutron, which are, respectively, the positively charged and neutral constituents of ordinary matter. The rest mass of an electron can also be expressed as 0.511 MeV/c^2, where c is the speed of light. Electrons were discovered in 1895 by J. J. Thomson

in the form of cathode rays. The electron was the first elementary particle to be identified. *See* ATOMIC STRUCTURE AND SPECTRA; CATHODE RAYS; ELECTRIC CHARGE; ELECTRONVOLT; ELEMENTARY PARTICLE; NUCLEAR STRUCTURE.

Charge. The charge of the electron is $e = 1.602 \times 10^{-19}$ coulomb. The sign of the electron's charge is negative by convention, and that of the equally charged proton is positive. This is a somewhat unfortunate convention, because the flow of electrons in a metallic conductor is thus opposite to the conventional direction of the current. *See* CONDUCTION (ELECTRICITY).

The historical method of measuring the charge of the electron is the celebrated oil drop experiment of R. A. Millikan in 1909, in which the charges of droplets of oil in air are measured by finding the electric field which balances each drop against its weight. The weight of each drop is determined by observing its rate of free fall through the air, and using Stokes' formula for the viscous drag on a slowly moving sphere. The charges thus measured are integral multiples of e. For more precise values of e and m_e, *see* FUNDAMENTAL CONSTANTS.

Electrons and matter. Electrons are emitted in radioactivity (as beta rays) and in many other decay processes; for instance, the ultimate decay products of all mesons are electrons, neutrinos, and photons, the meson's charge being carried away by the electrons. The electron itself is completely stable, according to all available evidence. Electrons contribute the volume to ordinary matter; the volume of an atom is nearly all occupied by the cloud of electrons surrounding the nucleus, which occupies only about 10^{-13} of the atom's volume. Virtually all the mass of an atom, however, is in the nucleus. The chemical properties of ordinary matter are determined by the electron cloud. *See* BETA PARTICLES; CHEMICAL BONDING; QUANTUM CHEMISTRY; RADIOACTIVITY; VALENCE.

The electron obeys the Fermi-Dirac statistics, and for this reason is often called a fermion. One of the primary attributes of matter, impenetrability, results from the fact that the electron, being a fermion, obeys the Pauli exclusion principle; the world would be completely different if the lightest charged particle were a boson, that is, a particle that obeys Bose-Einstein statistics. *See* BOSE-EINSTEIN STATISTICS; EXCLUSION PRINCIPLE; FERMI-DIRAC STATISTICS.

Spin. Every elementary particle possesses an intrinsic angular momentum called its spin. The spin of the electron has the magnitude $^1/_2\hbar$, where $\hbar$ is Planck's constant h divided by 2π. An electron thus has two spin states: spin up and spin down. To describe this, the nonrelativistic wave function is a two-component function, that is, a vector in a two-dimensional spin space; the two linearly independent vectors represent the two possible spin states. In 1928 P. A. M. Dirac derived the corresponding relativistic wave equation (Dirac equation). Here, the electron wave function must have four components; correspondingly, for a wave of given momentum, there are four internal states. In addition to the two-valued spin coordinate, there is an energy coordinate; that is, for a momentum p, the energy can be $\pm\sqrt{(m_e c^2)^2 + (pc)^2}$, where c is the velocity of light. *See* ANGULAR MOMENTUM; ELECTRON SPIN; NONRELATIVISTIC QUANTUM THEORY; RELATIVISTIC QUANTUM THEORY; SPIN (QUANTUM MECHANICS).

Positron. The negative energy states were at first an embarrassment, for they extend downward indefinitely; an electron would cascade indefinitely downward in energy, radiating photons. Electrons obey the exclusion principle, however, and this conclusion can be avoided by assuming that in empty space all the negative energy states are already occupied, and so exclude any more electrons. A new process is possible now; such a negative energy electron, by absorbing energy, can go to a positive energy state. This leaves behind a hole in the sea of negative-energy electrons; the hole has a positive energy, because it represents a missing negative energy. In fact, such a hole has all the properties of an electron except that it appears to have a positive charge (because it represents a missing negative charge). This particle is the positron, discovered in 1932 by C. D. Anderson in a cloud-chamber study of cosmic radiation. *See* POSITRON.

The electron and the positron are on an equal footing; if one started with a Dirac wave equation for the positron, identifying electrons with holes in the negative-energy positron states, one would get an equivalent theory. The apparent dissymmetry inherent in the construction of the hole theory disappears from the results when the total charge, energy, and momentum of empty space is defined to be zero; actually, the dissymmetry can be avoided at all stages in the formalism of quantum field theory. *See* ANTIMATTER; QUANTUM FIELD THEORY; SYMMETRY LAWS (PHYSICS).

Magnetic moment. The electron has magnetic properties by virtue of its orbital motion about the nucleus of its atom and its rotation (spin) about its own axis. The intrinsic magnetic moment of the electron is predicted by the Dirac equation to be the value (Bohr magneton) shown in the equation below. The actual moment μ differs from μ_D

$$\mu_D = \frac{e\hbar}{2m_e}$$

by a small amount (anomalous magnetic moment) due to electromagnetic radiative corrections: $\mu = 1.0011\ \mu_D$. This theoretical value, calculated using renormalized quantum field theory, agrees with the experimental value. *See* MAGNETON; QUANTUM ELECTRODYNAMICS.

Other leptons. The electron is the lightest of a family of elementary particles, the leptons. The other known charged leptons are the muon and the tau. These three particles differ only (as far as known) in mass; they have the same spin, charge, strong interactions (namely, none), and weak interactions. In a weak interaction a charged lepton is either unchanged (neutral weak current reaction) or changed

(charged weak current reaction) into an uncharged lepton, that is, a neutrino. In the latter case, each charged lepton (electron, muon, or tau) is seen to change only into the corresponding neutrino (ν_e, ν_μ, or ν_τ). *See* LEPTON; NEUTRINO; WEAK NUCLEAR INTERACTIONS.

Electrons are responsible for many phenomena of importance. Electrons moving through a conducting material constitute electricity. There is a magnetic field surrounding a current-carrying conductor, which makes possible electric motors and generators. Moving electrons are generally responsible for the production of light and electromagnetic radiation; examples include radio and television waves emitted by electrons moving in an antenna, microwaves generated in a magnetron in a microwave oven, and x-rays emitted by electrons passing through matter or electrons deflected in a magnetic field. Electron spin is also responsible for the behavior of magnetic materials. *See* ANTENNA (ELECTROMAGNETISM); BREMSSTRAHLUNG; ELECTRIC CURRENT; ELECTRICITY; ELECTROMAGNETIC RADIATION; GENERATOR; LIGHT; MAGNETIC MATERIALS; MAGNETISM; MAGNETRON; MOTOR; SYNCHROTRON RADIATION.

Electrons are responsible for chemical bonding in molecules. Additionally, the number of protons or electrons in an atom determines the elemental species of the atom, and thus the ranking of elements by atomic number (the number of electrons or protons in the atom) in the periodic table. *See* ATOMIC NUMBER; CHEMICAL BONDING; ELEMENT (CHEMISTRY); PERIODIC TABLE.

Electrons are a valuable probe of matter, as in an electron microscope. Electrons at high energies have a very short wavelength, and microscopes with atomic resolution can be constructed. *See* ELECTRON DIFFRACTION; ELECTRON MICROSCOPE.

Other information. It would be difficult to list all the articles wherein the electron forms an integral part of the discussion. The following articles and those referenced in the preceding discussion are intended to be merely a representative sample. *See* BAND THEORY OF SOLIDS; CHARGED PARTICLE OPTICS; COMPTON EFFECT; ELECTRON CAPTURE; ELECTRON CONFIGURATION; ELECTRON EMISSION; ELECTRON MOTION IN VACUUM; ELECTRONICS; EXCHANGE INTERACTION; FREE-ELECTRON THEORY OF METALS; MAGNETISM; PARTICLE ACCELERATOR; QUANTUM MECHANICS; RELATIVISTIC ELECTRODYNAMICS.

Charles J. Goebel; Alfred S. Schlachter

Electron affinity

The amount of energy release when an electron e at rest is captured by a species M, producing the negative ion M^-. The electron affinity of a species M can also be thought of as the ionization potential of the negative ion M^-. In these definitions, it is assumed that both M and M^- are in the lowest electronic, vibrational, and rotational state.

If the electron affinity of M is negative, the M^- ion is unstable with respect to decomposition into $M + e$. Most atoms have positive electron affinities, even though there is no net Coulomb attraction between the electron and the atom until the electron is close enough to be "a part of the atom." The simple rules of chemical valence provide a qualitative guide to the magnitude of electron affinities. Thus the noble gases, which have a filled outer electronic shell and are chemically inert, are not capable of binding an additional electron to form a negative ion. The largest electron affinities are possessed by the halogens, atoms which require only one additional electron to fill the valence shell. *See* HALOGEN ELEMENTS.

The major exception to this concept is that multiply charged negative ions—for example, O^{2-}, one of many such ions which are stable in solution—have not been observed in the gas phase. The ability to place more than one additional electron in the valence shell of a neutral atom or molecule arises from many particle electrostatic interactions with the ion's environment: either the solvent shell surrounding the ion in liquid solutions or the amorphous or crystalline region surrounding the ion in solids. *See* VALENCE.

Experimental methods. Accurate ionization potentials of the elements were known for a number of years before comparable data for electron affinities of the elements became available. In order to determine the ionization potential of an element, the element is vaporized, placed in an optical spectrometer, and observed for the onset of photoabsorption corresponding to the photoionization process, $h\nu + M \rightarrow M^+ + e$. The photon energy corresponding to the threshold wavelength for this process is the ionization energy of the species M. The analogous method for determination of an electron affinity is through observation of the threshold of the very similar photodetachment reaction, $h\nu + M^- \rightarrow M + e$. Again, the threshold wavelength for this process corresponds to the electron affinity of the species M. Unfortunately, it has not proved possible to produce the sufficiently large densities of negative ions required to utilize photoabsorption spectrometers to observe directly the threshold for the photodetachment process. Consequently, determination of accurate electron affinities lagged far behind the determination of accurate ionization potentials. However, an ion velocity modulation scheme has been devised that permits the direct observation of vibrational photoabsorption in ions; it is not generally applicable to the electron detachment process required to determine electron affinities.

The two major experimental advances that have enabled accurate electron affinity determinations have been the development of ion-beam techniques affording the detection of photoabsorption by the monitoring of photodetachment products and the availability of intense light sources in the form of lasers.

In experiments to determine electron affinities, negative ions are formed in an electrical

1 H 0.754201							2 He <0
3 Li 0.6180	4 Be <0	5 B 0.277	6 C 1.2629	7 N ≈0	8 O 1.4611215	9 F 3.399	10 Ne <0
11 Na 0.54793	12 Mg <0	13 Al 0.441	14 Si 1.385	15 P 0.7465	16 S 2.077120	17 Cl 3.61272	18 Ar <0
19 K 0.50147	20 Ca 0.043	31 Ga 0.30	32 Ge 1.233	33 As 0.81	34 Se 2.02069	35 Br 3.365	36 Kr <0
37 Rb 0.48592	38 Sr 0.10	49 In 0.30	50 Sn 1.112	51 Sb 1.07	52 Te 1.9708	53 I 3.0591	54 Xe <0
55 Cs 0.47163	56 Ba 0.18	81 Tl 0.2	82 Pb 0.364	83 Bi 0.946	84 Po 1.9	85 At 2.8	86 Rn <0

21 Sc 0.188	22 Ti 0.079	23 V 0.525	24 Cr 0.677	25 Mn <0	26 Fe 0.151	27 Co 0.662	28 Ni 1.156	29 Cu 1.235	30 Zn <0
39 Y 0.307	40 Zr 0.426	41 Nb 0.893	42 Mo 0.746	43 Tc 0.55	44 Ru 1.05	45 Rh 1.137	46 Pd 0.557	47 Ag 1.302	48 Cd <0
57 La 0.5	72 Hf ≈0	73 Ta 0.322	74 W 0.815	75 Re 0.15	76 Os 1.1	77 Ir 1.565	78 Pt 2.128	79 Au 2.30863	80 Hg <0

Periodic table showing the best values for the electron affinities of the elements. All values are reported in electronvolts. The value <0 implies that the negative ion is unstable with respect to decomposition to an electron and a neutral atom. In general, the last decimal place in each value is uncertain. (*After H. Hotop and W. C. Lineberger, Binding energies in atomic negative ions: II, J. Phys. Chem. Ref. Data, 14:731–750, 1985*)

discharge, extracted through an aperture into a high-vacuum region, formed into a negative-ion beam, mass-analyzed, and intersected by an intense laser beam. The laser-beam–negative-ion-beam intersection takes place in a high-vacuum region where no collisions are likely. The occurrence of a photodetachment event is determined by detection of the photodetached electron.

Two experimental methods evolved which produce accurate electron affinities. In the first method the laser is tunable, and a search is made for the wavelength corresponding to the threshold for the photodetachment process. In this case the electron affinity is given directly by the threshold wavelength, and experimental methods have been developed that provide accuracies of 10^{-8} eV. In the second type of experiment, called photoelectron spectroscopy, a fixed-frequency laser (of known photon energy) is employed, and the electrostatic fields are used to determine the kinetic energy of the ejected electron. From simple energy conservation arguments, the electron affinity is then given by the photon energy less the kinetic energy of the ejected electron. This latter technique is quite general, but is limited in accuracy by the resolution of the electron energy analyzer (approximately 3×10^{-3} eV for state-of-the-art analyzers). *See* LASER; LASER SPECTROSCOPY; MASS SPECTROMETRY.

Periodic trends. These laser photodetachment studies dramatically improved knowledge of the electron affinities of the elements. The **illustration** is a periodic table showing values of the electron affinities of the elements. Most of the data shown here were obtained by using laser photodetachment methods. The periodic trends in electron affinities and the qualitative effects described earlier are immediately apparent. In addition, a number of more subtle trends are observable. For example, while it might be expected that filled-shell species such as the rare gases would not be capable of binding an additional electron, the illustration shows that half-filled shells (for example, N and P) also exhibit small or negative electron affinities. Again, this effect is the result of the fact that a half-filled valence shell is spherically symmetric and behaves somewhat as though it were a filled shell. A similar situation is also seen for half-filled *d* shells, as is the case for Mn, Tc, and Re in the transition metals.

These same techniques have provided a number of accurate electron affinity determinations for molecules and free radicals, and new insight into the structural and chemical properties of ions in the gas phase. *See* IONIZATION POTENTIAL; PERIODIC TABLE.

W. C. Lineberger

Bibliography. M. T. Bowers (ed.), *Gas Phase Ion Chemistry 3*, 1984; H. Hotop and W. C. Lineberger, Binding energies in atomic negative ions: II, *J. Phys. Chem. Ref. Data*, 14:731–750, 1985; I. N. Levine, *Physical Chemistry*, 4th ed., 1994; R. J. Saykally, Infrared laser spectroscopy of molecular ions, *Science*, 239:157–161, 1988.

Electron capture

A term that refers to two processes, one atomic and the other nuclear, in both of which an electron is captured: by a charged particle (an ion) passing through matter, when an electron is captured into an atomic orbital (in the atomic case); or by a nucleus (in the nuclear case), which results in radioactive decay of the nucleus.

Atomic process. The capture and loss of electrons is an important mechanism for slowing down fast positive ions (atoms with fewer electrons than their nuclear charge) passing through matter. Electron capture is significant generally when the velocity of the ion is similar to that of the target electrons to be captured. Fast ions passing through matter thus generally lose much of their energy near the end of their range passing through matter, as their velocity becomes similar to that of the bound target electrons.

An example of electron capture is shown in **Fig. 1**, which shows the cross section (the likelihood) for electron capture by a proton passing through lithium vapor. The reaction is $H^+ + Li \rightarrow H^0 + Li^+$ (or possibly $Li^{2+} + e^-$). The cross section (the likelihood of capturing an electron) becomes very small when the proton energy (which is proportional to the square

of the velocity) becomes large. *See* CHARGED PARTICLE BEAMS.

Nuclear process. The nuclear capture of electrons (K capture) occurs by a process quite different from atomic capture and leads to a nuclear reaction. An electron is captured by the nucleus; the electron is taken from the K shell (the electron shell nearest the nucleus) of the atom. The probability of electron capture by the nucleus depends on the amount of time the electrons spend near the nucleus, that is, on the size of the electron wave function at the nuclear center. Since to a very good approximation only electrons with zero orbital angular momentum have a wave function that is finite at the center, K capture generally occurs for capture of a K-shell electron. However, L capture is also possible. *See* RADIOACTIVITY.

Electron capture (K capture) generally occurs for a nucleus with too many protons (or too few neutrons). A K-shell electron from the atom is captured by the nucleus, creating a neutron and a neutrino, resulting in a more stable nucleus. The result of essentially transforming a proton into a neutron in the nucleus is to reduce the atomic number by one, resulting in a different element (since the number of electrons in an atom is equal to the number of protons in the nucleus, and the number of electrons or protons determines the element). Atomic mass is essentially unchanged by K capture, as the mass of a proton and neutron are essentially the same. The atom is missing a K-shell electron as a result of K capture; it stabilizes by emission of an x-ray or an Auger electron: an electron from an outer shell fills the vacancy in the K shell of the atom. Thus, K capture is a simultaneous atomic and nuclear event. *See* AUGER EFFECT.

K capture is favored when there is not enough energy for the nucleus to emit a positron, which requires 1.022 MeV. An example of a nucleus without sufficient energy to emit a positron is rubidium-83, which decays to krypton-83 by K capture. *See* POSITRON.

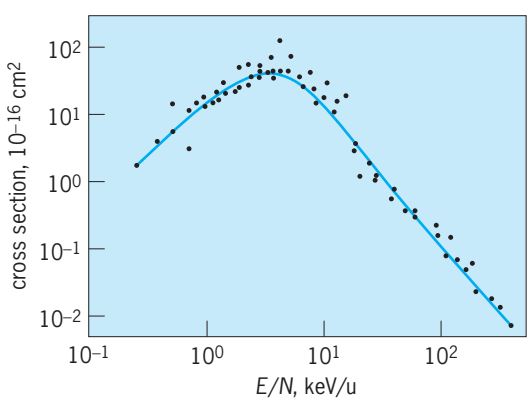

Fig. 1. Cross section (units of 10^{-16} cm^2) for capture of an electron by a proton passing through lithium vapor. Energy per nucleon (*E/N*) is in units of kiloelectronvolts per nucleon (keV/u). Data points from experiments and a curve fitted to the data are shown. (*After T. J. Morgan et al., Charge transfer of hydrogen ions and atoms in metal vapors, J. Phys. Chem. Ref. Data, 14:971–1040, 1985*).

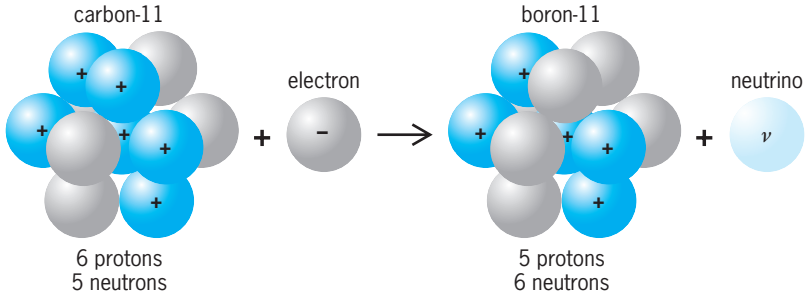

Fig. 2. Example of K capture: The reaction carbon-11 (6 protons and 5 neutrons) + electron → boron-11 (5 protons and 6 neutrons) + neutrino. (*Jefferson Lab, http://education.jlab.org/glossary/electroncapture.html*)

An example of K capture is shown in **Fig. 2**. Both carbon-11 and boron-11 have the same atomic mass (11); however, boron (atomic number 5) is one element before carbon (atomic number 6) in the periodic table.

Alfred S. Schlachter

Electron configuration

The orbital arrangement of an atom's electrons. Negatively charged electrons are attracted to a positively charged nucleus to form an atom or ion. Although such bound electrons exhibit a high degree of quantum-mechanical wavelike behavior, there still remain particle aspects to their motion. Bound electrons occupy orbitals that are somewhat concentrated in spatial shells lying at different distances from the nucleus. As the set of electron energies allowed by quantum mechanics is discrete, so is the set of mean shell radii. Both these quantized physical quantities are primarily specified by integral values of the principal, or total, quantum number n. The full electron configuration of an atom is correlated with a set of values for all the quantum numbers of each and every electron. In addition to n, another important quantum number is l, an integer representing the orbital angular momentum of an electron in units of $h/2\pi$ where h is Planck's constant. The values 1, 2, 3, 4, 5, 6, 7 for n and 0, 1, 2, 3, for l together suffice to describe the electron configurations of all known normal atoms and ions, that is, those that have their lowest possible values of total electronic energy. The first seven shells are also given the letter designations K, L, M, N, O, P, and Q, respectively. Electrons with l equal to 0, 1, 2, and 3 are designated s, p, d, and f, respectively. *See* QUANTUM MECHANICS; QUANTUM NUMBERS.

In any configuration the number of equivalent electrons (same n and l) is indicated by an integral exponent (not a quantum number) attached to the letters s, p, d, and f. According to the Pauli exclusion principle the maximum is s^2, p^6, d^{10}, and f^{14}. For example, the configuration $1s^2 2s^2 2p^6 3s$ of the

Distribution of electrons in the atoms*

Element and atomic number		K	L		M			N				O				Ground term	Ionization potential, eV
		1, 0 1*s*	2, 0 2*s*	2, 1 2*p*	3, 0 3*s*	3, 1 3*p*	3, 2 3*d*	4, 0 4*s*	4, 1 4*p*	4, 2 4*d*	4, 3 4*f*	5, 0 5*s*	5, 1 5*p*	5, 2 5*d*	5, 3 5*f*		
H	1	1	—	—	—	—	—	—	—	—	—	—	—	—	—	$^2S_{1/2}$	13.5984
He	2	2	—	—	—	—	—	—	—	—	—	—	—	—	—	$^1S_0$	24.5874
Li	3	2	1	—	—	—	—	—	—	—	—	—	—	—	—	$^2S_{1/2}$	5.3917
Be	4	2	2	—	—	—	—	—	—	—	—	—	—	—	—	$^1S_0$	9.3227
B	5	2	2	1	—	—	—	—	—	—	—	—	—	—	—	$^2P^\circ_{1/2}$	8.2980
C	6	2	2	2	—	—	—	—	—	—	—	—	—	—	—	$^3P_0$	11.2603
N	7	2	2	3	—	—	—	—	—	—	—	—	—	—	—	$^4S^\circ_{3/2}$	14.5341
O	8	2	2	4	—	—	—	—	—	—	—	—	—	—	—	$^3P_2$	13.6181
F	9	2	2	5	—	—	—	—	—	—	—	—	—	—	—	$^2P^\circ_{3/2}$	17.4228
Ne	10	2	2	6	—	—	—	—	—	—	—	—	—	—	—	$^1S_0$	21.5645
Na	11	Neon configuration			1	—	—	—	—	—	—	—	—	—	—	$^2S_{1/2}$	5.1391
Mg	12				2	—	—	—	—	—	—	—	—	—	—	$^1S_0$	7.6462
Al	13				2	1	—	—	—	—	—	—	—	—	—	$^2P^\circ_{1/2}$	5.9858
Si	14				2	2	—	—	—	—	—	—	—	—	—	$^3P_0$	8.1517
P	15				2	3	—	—	—	—	—	—	—	—	—	$^4S^\circ_{3/2}$	10.4867
S	16				2	4	—	—	—	—	—	—	—	—	—	$^3P_2$	10.3600
Cl	17				2	5	—	—	—	—	—	—	—	—	—	$^2P^\circ_{3/2}$	12.9676
Ar	18				2	6	—	—	—	—	—	—	—	—	—	$^1S_0$	15.7596
K	19	Argon configuration					—	1	—	—	—	—	—	—	—	$^2S_{1/2}$	4.3407
Ca	20						—	2	—	—	—	—	—	—	—	$^1S_0$	6.1132
Sc	21						1	2	—	—	—	—	—	—	—	$^2D_{3/2}$	6.5615
Ti	22						2	2	—	—	—	—	—	—	—	$^3F_2$	6.8281
V	23						3	2	—	—	—	—	—	—	—	$^4F_{3/2}$	6.7462
Cr	24						5	1	—	—	—	—	—	—	—	$^7S_3$	6.7665
Mn	25						5	2	—	—	—	—	—	—	—	$^6S_{5/2}$	7.4340
Fe	26						6	2	—	—	—	—	—	—	—	$^5D_4$	7.9024
Co	27						7	2	—	—	—	—	—	—	—	$^4F_{9/2}$	7.8810
Ni	28						8	2	—	—	—	—	—	—	—	$^3F_4$	7.6398
Cu	29						10	1	—	—	—	—	—	—	—	$^2S_{1/2}$	7.7264
Zn	30						10	2	—	—	—	—	—	—	—	$^1S_0$	9.3942
Ga	31						10	2	1	—	—	—	—	—	—	$^2P^\circ_{1/2}$	5.9993
Ge	32						10	2	2	—	—	—	—	—	—	$^3P_0$	7.8994
As	33						10	2	3	—	—	—	—	—	—	$^4S^\circ_{3/2}$	9.7886
Se	34						10	2	4	—	—	—	—	—	—	$^3P_2$	9.7524
Br	35						10	2	5	—	—	—	—	—	—	$^2P^\circ_{3/2}$	11.8138
Kr	36						10	2	6	—	—	—	—	—	—	$^1S_0$	13.9996
Rb	37	Krypton configuration								—	—	1	—	—	—	$^2S_{1/2}$	4.1771
Sr	38									—	—	2	—	—	—	$^1S_0$	5.6949
Y	39									1	—	2	—	—	—	$^2D_{3/2}$	6.2173
Zr	40									2	—	2	—	—	—	$^3F_2$	6.6339
Nb	41									4	—	1	—	—	—	$^6D_{1/2}$	6.7589
Mo	42									5	—	1	—	—	—	$^7S_3$	7.0924
Tc	43									5	—	2	—	—	—	$^6S_{5/2}$	7.28
Ru	44									7	—	1	—	—	—	$^5F_5$	7.3605
Rh	45									8	—	1	—	—	—	$^4F_{9/2}$	7.4589
Pd	46									10	—	—	—	—	—	$^1S_0$	8.3369

ground or lowest energy state of sodium means that there are two electrons with $n = 1$, $l = 0$; two with $n = 2$, $l = 0$; six with $n = 2$, $l = 1$; and one with $n = 3$, $l = 0$. Higher values of n and l can be achieved by excitation of normal atoms, either through photon absorption or by particle impact, temporarily moving one or more electrons to unfilled shells of larger radii and increasing the atom's electronic energy. Such excited electronic configurations are inherently unstable. Excited electrons ultimately fall back down to their normal locations closer to the nucleus, a process accompanied by the emission of one or more quanta of electromagnetic radiation (photons). While the atom is excited, the partially depleted shells are said to possess vacancies. *See* EXCLUSION PRINCIPLE.

An electron configuration is categorized as having even or odd parity, according to whether the sum of p and f electrons is even or odd. Strong spectral lines result only from transitions between configurations of unlike parity. *See* PARITY (QUANTUM MECHANICS).

Insofar as they are known from spectroscopic investigations, the electron configurations characteristic of the normal or ground states of the first 102 chemical elements are shown in the **table**. In the next-to-last column of the table, the spectral term of the energy level with lowest total electronic energy is shown. For all elements except protactinium (Pa), uranium (U), and neptunium (Np), an *LS*-coupling term is given. Here, the main part of the term symbol is a capital letter, *S*, *P*, *D*, *F*, and so on, that

Distribution of electrons in the atoms* (cont.)

Element and atomic number		Configuration of inner shells	N	O				P			Q	Ground term	Ionization potential, eV
			4, 3 4*f*	5, 0 5*s*	5, 1 5*p*	5, 2 5*d*	5, 3 5*f*	6, 0 6*s*	6, 1 6*p*	6, 2 6*d*	7, 0 7*s*		
Ag	47	Palladium configuration	—	1	—	—	—	—	—	—	—	$^2S_{1/2}$	7.5762
Cd	48		—	2	—	—	—	—	—	—	—	$^1S_0$	8.9938
In	49		—	2	1	—	—	—	—	—	—	$^2P^\circ_{1/2}$	5.7864
Sn	50		—	2	2	—	—	—	—	—	—	$^3P_0$	7.3439
Sb	51			2	3	—	—	—	—	—	—	$^4S^\circ_{3/2}$	8.6084
Te	52		—	2	4	—	—	—	—	—	—	$^3P_2$	9.0096
I	53		—	2	5	—	—	—	—	—	—	$^2P^\circ_{3/2}$	10.4513
Xe	54		—	2	6	—	—	—	—	—	—	$^1S_0$	12.1298
Cs	55	The shells 1*s* to 4*d* contain 46 electrons	—	The shells 5*s* to 5*p* contain 8 electrons		—	—	1	—	—	—	$^2S_{1/2}$	3.8939
Ba	56		—			—	—	2	—	—	—	$^1S_0$	5.2117
La	57		—			1	—	2	—	—	—	$^2D_{3/2}$	5.5769
Ce	58		1			1	—	2	—	—	—	$^1G^\circ_4$	5.5387
Pr	59		3			—	—	2	—	—	—	$^4I^\circ_{9/2}$	5.473
Nd	60		4			—	—	2	—	—	—	$^5I_4$	5.5250
Pm	61		5			—	—	2	—	—	—	$^6H^\circ_{5/2}$	5.582
Sm	62		6			—	—	2	—	—	—	$^7F_0$	5.6437
Eu	63		7			—	—	2	—	—	—	$^8S^\circ_{7/2}$	5.6704
Gd	64		7			1	—	2	—	—	—	$^9D^\circ$	6.1498
Tb	65		9			—	—	2	—	—	—	$^6H^\circ_{15/2}$	5.8638
Dy	66		10			—	—	2	—	—	—	$^5I_8$	5.9389
Ho	67		11			—	—	2	—	—	—	$^4I^\circ_{15/2}$	6.0215
Er	68		12			—	—	2	—	—	—	$^3H_6$	6.1077
Tm	69		13			—	—	2	—	—	—	$^2F^\circ_{7/2}$	6.1843
Yb	70		14			—	—	2	—	—	—	$^1S_0$	6.2542
Lu	71		14			1	—	2	—	—	—	$^2D_{3/2}$	5.4259
Hf	72	The shells 1*s* to 5*p* contain 68 electrons				2	—	2	—	—	—	$^3F_2$	6.8251
Ta	73					3	—	2	—	—	—	$^4F_{3/2}$	7.5496
W	74					4	—	2	—	—	—	$^5D_0$	7.8640
Re	75					5	—	2	—	—	—	$^6S_{5/2}$	7.8335
Os	76					6	—	2	—	—	—	$^5D_4$	8.28
Ir	77					7	—	2	—	—	—	$^4F_{9/2}$	9.02
Pt	78					9	—	1	—	—	—	$^3D_3$	8.9588
Au	79	The shells 1*s* to 5*d* contain 78 electrons					—	1	—	—	—	$^2S_{1/2}$	9.2255
Hg	80						—	2	—	—	—	$^1S_0$	10.4375
Tl	81						—	2	1	—	—	$^2P^\circ_{1/2}$	6.1082
Pb	82						—	2	2	—	—	$^3P_0$	7.4167
Bi	83						—	2	3	—	—	$^4S^\circ_{3/2}$	7.2855
Po	84						—	2	4	—	—	$^3P_2$	8.414
At	85						—	2	5	—	—	$^2P^\circ_{3/2}$	
Rn	86						—	2	6	—	—	$^1S_0$	10.7485
Fr	87						—	2	6	—	1	$^2S_{1/2}$	4.0727
Ra	88						—	2	6	—	2	$^1S_0$	5.2784
Ac	89						—	2	6	1	2	$^2D_{3/2}$	5.17
Th	90						—	2	6	2	2	$^3F_2$	6.3067
Pa	91						$2(^3H_4)$	2	6	1	2	$(4, 3/2)_{11/2}$	5.89
U	92						$3(^4I^\circ_{9/2})$	2	6	1	2	$(9/2, 3/2)^\circ_6$	6.1941
Np	93						$4(^5I_4)$	2	6	1	2	$(4, 3/2)_{11/2}$	6.2657
Pu	94						6	2	6	—	2	$^7F_0$	6.0260
Am	95						7	2	6	—	2	$^8S^\circ_{7/2}$	5.9738
Cm	96						7	2	6	1	2	$^9D^\circ_2$	5.9914
Bk	97						9	2	6	0	2	$^6H^\circ_{5/2}$	6.1979
Cf	98						10	2	6	0	2	$^5I_8$	6.2817
Es	99						11	2	6	0	2	$^4I_{15/2}$	6.42
Fm	100						12	2	6	0	2	$^3H_6$	6.50
Md	101						13	2	6	0	2	$^2F^\circ_{7/2}$	6.58
No	102						14	2	6	0	2	$^1S_0$	6.65

*Updated and revised with data from G. W. F. Drake (ed.), *Springer Handbook of Atomic, Molecular, and Optical Physics*, pp. 182–183, Springer, 2006.

represents the total electronic orbital angular momentum. Attached to this is a superior prefix, 1, 2, 3, 4, and so on, that indicates the multiplicity, and an anterior suffix, 0, $^1/_2$, 1, $^3/_2$, 2, $^5/_2$, and so on, that shows the total angular momentum, or *J* value, of the atom in the given state. The sign $^\circ$ above the *J* value signifies that the spectral term and electron configuration have odd parity. (For the elements Pa, U, and Np, a *jj* coupling term is given. Here, parentheses enclose two numbers, the first of which is the total angular momentum of the electrons in the 5*f* shell, while the second is the total angular momentum of the electron in the 6*d* shell. The suffixes following the parentheses are the same as in *LS* coupling. For these elements the *LS* coupling of the electrons in the 5*f* shell is also given in the column that lists the number of such electrons.)

The last column of the table presents the first

ionization potential of the atom when this has been derived from spectroscopic observations. In any atomic spectrum, two or more spectral lines with certain similar properties may form a series such that the reciprocal wavelengths $1/\lambda$ (number of waves per centimeter $= \sigma$) can be closely represented by a formula of the Rydberg type, $\sigma = L - R/(n + \mu)^2$, in which L is the limit of the series. R is called the Rydberg constant, and the principal quantum number n has successive integral values to which a constant fractional part μ is added. The second term vanishes when n approaches infinity, and the series limit is thus evaluated. This limit is usually coincident with the ground state of the ion, and is thus a measure (in wave-number units) of the energy required to remove from an atom its least firmly bound electron and transform a neutral atom into a singly charged ion. The energy required to ionize an atom is usually expressed in electronvolts (1 eV = 8065.48 wave numbers) and is called its first ionization potential. *See* ATOMIC STRUCTURE AND SPECTRA; IONIZATION POTENTIAL; RYDBERG ATOM; RYDBERG CONSTANT.

Molecules consist of two or more atoms that at least partially share one or more electrons with one another. Sets of quantum numbers specify molecular electron configurations in much the same way as for atoms. However, the physical meaning of some molecular quantum numbers is different than that for the atomic case, as is the term notation. *See* MOLECULAR STRUCTURE AND SPECTRA. James E. Bayfield

Electron diffraction

The phenomenon associated with interference processes that occur when electrons are scattered by atoms to form diffraction patterns. The wave character of electrons is shown most strikingly, and doubtless most conclusively, by the phenomena of interference. For this reason, the diffraction of electrons presents the most obvious confirmation of quantum mechanics. Because of the dependence of the diffraction pattern on the distances between the atoms, electron diffraction is also an important tool for the study of the structure of crystals and of free molecules, analogous to the use of x-rays for these purposes. *See* X-RAY CRYSTALLOGRAPHY; X-RAY DIFFRACTION.

According to quantum theory, any particle moving with momentum mv has a wavelength $\lambda = h/mv$, where h is Planck's constant. If the particle is an electron, and its velocity is the result of acceleration by a the potential difference V, this formula becomes Eq. (1), where m_0 and e are the rest mass

$$\lambda = \frac{h}{(2m_0eV)^{1/2}(1 + eV/2m_0c^2)^{1/2}} \qquad (1)$$

and charge of the electron, and c is the velocity of light. The last factor in the denominator represents the relativity correction, which is negligible at low voltages and amounts to only 5% at 100,000 V. For V in volts and λ in nanometers, Eq. (2) is a good

$$\lambda = \left(\frac{1.5}{V}\right)^{1/2} \qquad (2)$$

approximation to Eq. (1) at nonrelativistic energies. According to energy $E = eV$, two major techniques of structure analysis with electron beams are distinguished: low-energy electron diffraction (LEED) [$E \simeq$ 5–500 eV] and high-energy electron diffraction (HEED) [$E \simeq$ 5–500 keV]; medium-energy electron diffraction (MEED) [$E \simeq$ 500 eV–5 keV] is of little importance. In addition, electrons generated in condensed matter by incident electrons or x-ray photons are diffracted (in Auger electron diffraction and photoelectron diffraction). Unlike neutrons and x-rays, electrons penetrate matter only for a very short distance before they lose energy (by inelastic scattering) or are scattered elastically (diffracted). Due to the energy loss, the wavelength of an electron changes according to Eq. (2) so that it can no longer interfere with the other incident electrons. This loss of coherence occurs in condensed matter typically after mean free paths for inelastic scattering ranging from several tenths of nanometers at low energies ($E \simeq 50$ eV) to several tens of nanometers at high energies, which determines the application range of LEED and HEED. *See* COHERENCE; DE BROGLIE WAVELENGTH; DIFFRACTION; INTERFERENCE OF WAVES; MEAN FREE PATH; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Low-Energy Electron Diffraction

LEED is used mainly for the study of the structure of single-crystal surfaces and of processes on such surfaces that are associated with changes in the lateral periodicity of the surface. A monochromatic, nearly parallel electron beam, of 10^{-4} to 10^{-3} m (4×10^{-3} to 4×10^{-2} in.) in diameter, strikes the surface, usually at normal incidence. The elastically backscattered electrons are separated from all other electrons by a retarding field and detected with a suitable movable collector or, more frequently—after acceleration to about 5 keV energy—on a hemispherical fluorescent screen with the crystal in its center (**Fig. 1**). The intensity of the diffraction spots can be measured as a function of the energy $E = eV$ of the incident electrons to obtain so-called $I(V)$ curves (**Fig. 2**).

To obtain such results, the surface must be carefully cleaned and kept in ultrahigh vacuum (UHV, pressure $\simeq 10^{-10}$ torr $\simeq 10^{-8}$ pascal) during the experiment and must be exposed to well-defined amounts of gases or vapors at low pressures if adsorption, condensation, or corrosion is to be studied. This requirement, as well as the fact that many surfaces are changed by the electron beam—for example, due to dissociation in the case of ionic crystals or due to desorption of adsorbed gases—limits the applicability of LEED. Nevertheless, the number of surfaces and surface processes that can be studied

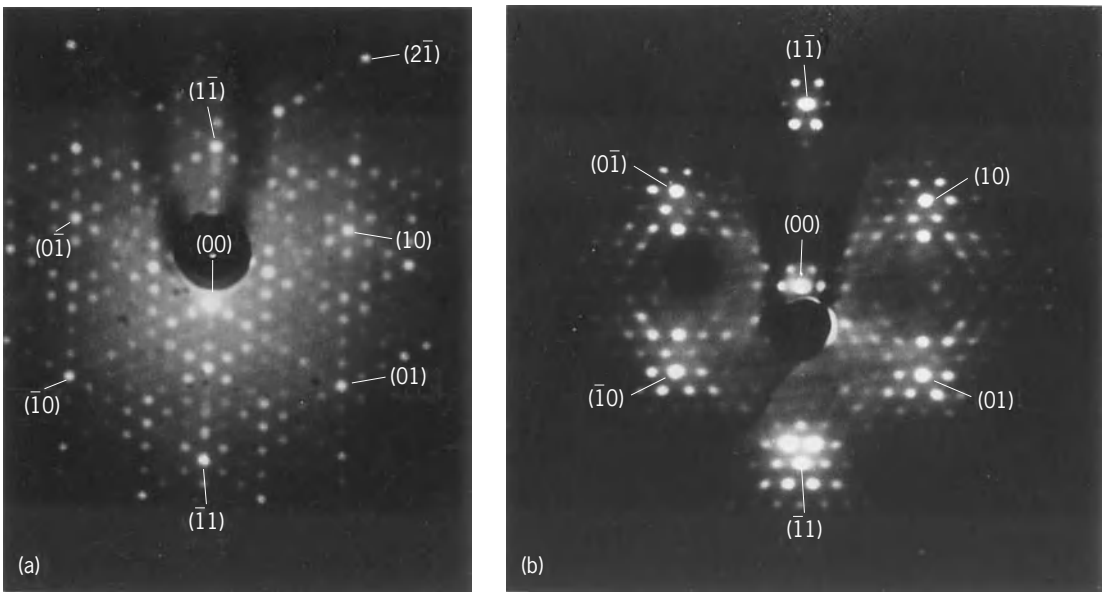

Fig. 1. LEED patterns from single-crystal surfaces. Numbers are indices (*h*, *k*) labeling diffraction spots of the surfaces. (*a*) Clean silicon (111) surface (7 × 7 structure). *E* = 85 eV. (*b*) Ordered carbon monoxide adsorption layer on a tungsten (110) surface. *E* = 120 eV.

is very large, and the main limitation of LEED is not caused by difficulties in obtaining LEED patterns but rather in evaluating them.

Interpretation of patterns. The difficulties in interpreting LEED patterns are caused by the strong elastic scattering of slow electrons by atoms. As a consequence of this, the electrons observed in the LEED pattern have been scattered several times on the average, in contrast to x-ray and neutron diffraction, in which single scattering is a good approximation. Thus evaluation of the spot intensities of LEED patterns, with the additional complications arising from the elastic scattering at the surface barrier and from inelastic scattering, requires large-scale computers. As a consequence, only the structures of surfaces whose unit meshes (two-dimensional unit cells) are small have been determined by LEED. For this reason, alternative evaluation methods have been developed in which multiple-scattering effects are reduced by proper averaging procedures.

The interpretation difficulties are considerably reduced if only the lateral periodicity of the surface, but not the location of the atoms in the unit mesh, is to be determined. The determination of the lateral periodicity of the surface structure and of the size and shape of domains with a given structure requires only the evaluation of the geometry of the pattern, that is, of the position, size, and shape of the diffraction spots but not of their intensity, and this evaluation is relatively easy. Thus the most important application of LEED is the study of surface phenomena associated with changes in lateral periodicity. A large amount of information on lateral disorder may be extracted from spot shapes and intensities. Special systems (spot profile analysis LEED, or SPALEED) have been developed for this purpose.

Surface structure. Clean surfaces of crystals frequently have the lateral periodicity that is expected from their bulk crystal structure. The analysis of *I*(*V*) curves of such surfaces (Fig. 2) in general reveals that the distribution of the atoms normal to the surface in the topmost layers differs somewhat from that in the bulk. These relaxations, that is, displacements of atoms from their positions in the bulk, are only a few percent on densely packed surfaces. On more open surfaces they can amount to 10% or more and are frequently oscillatory. There are, however, also surfaces which show complicated superstructures (known as reconstructed surfaces), that is, structures whose lateral periodicity is larger than that of the substrate (Fig. 1*a*). Usually, such surfaces are analyzed by LEED

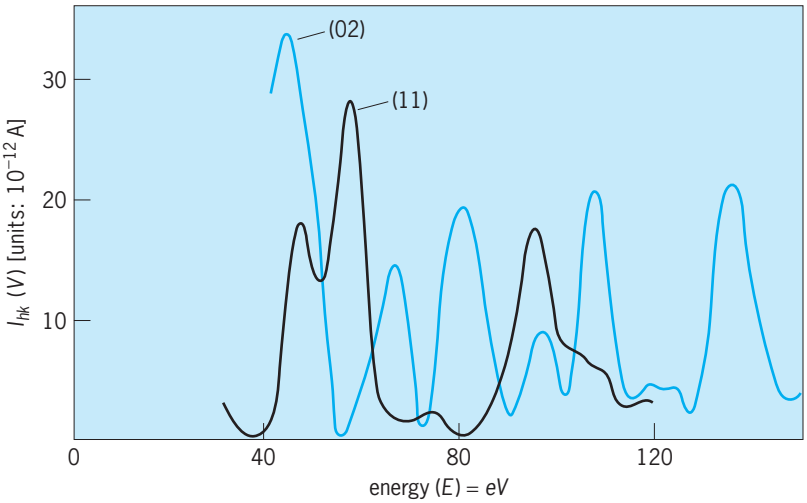

Fig. 2. *I*(*V*) curves of spots (*hk*) in the LEED pattern from a clean tungsten (110) surface. (*After G. A. Somorjai, The Structure and Chemistry of Solid Surfaces, Wiley, 1969*)

only after structural models have been proposed on the basis of other structure studies. On ideal surfaces, the decrease of the spot intensities with temperature has shown that the amplitudes of the thermal vibrations of surface atoms normal to the surface can be much larger than in the bulk. *See* CRYSTAL STRUCTURE; LATTICE VIBRATIONS.

Adsorption. Foreign atoms can interact with a clean surface in a variety of ways. The most important contribution of LEED is to the understanding of chemisorption, which precedes corrosion and, in many cases, epitaxy. Here, not only the structure of many adsorption systems, mainly of gases on metals, or metals on other metals and semiconductors, has been studied, but also the kinetics of the adsorption and desorption process as well as changes in the adsorption layer upon heating. The combination of LEED with Auger electron spectroscopy (AES) and with work-function measurements has proven particularly powerful in these studies, because such methods give the coverage and information on the location of the adsorbed atoms normal to the surface. Combining LEED with other complementary techniques such as ion scattering spectroscopy, electron energy loss spectroscopy, or photoelectron spectroscopy has become increasingly popular and can enable the elimination of ambiguities in the interpretation of many LEED results. *See* ADSORPTION; AUGER EFFECT; CORROSION; ELECTRON SPECTROSCOPY; SURFACE AND INTERFACIAL CHEMISTRY; SURFACE PHYSICS.

SPLEED. This is LEED with spin-polarized electrons. It gives information on the magnetic structure of surfaces and thin films due to the spin dependence of the diffraction of slow electrons in magnetic materials. This dependence is of particular importance in the spin-polarized version of low-energy electron microscopy (LEEM), called SPLEEM, which allows the magnetic microstructures of surfaces and thin films to be studied.

LEEM. Low-energy diffracted electrons can be used to image the surface of a specimen if it is made the cathode in an immersion objective lens of an ultrahigh-vacuum electron microscope. Because of the strong elastic backscattering of very slow electrons, the resulting LEEM images (**Fig. 3**) can be recorded at video rates. This allows the study of surface processes such as phase transitions or film growth in real time. *See* ELECTRON MICROSCOPE.

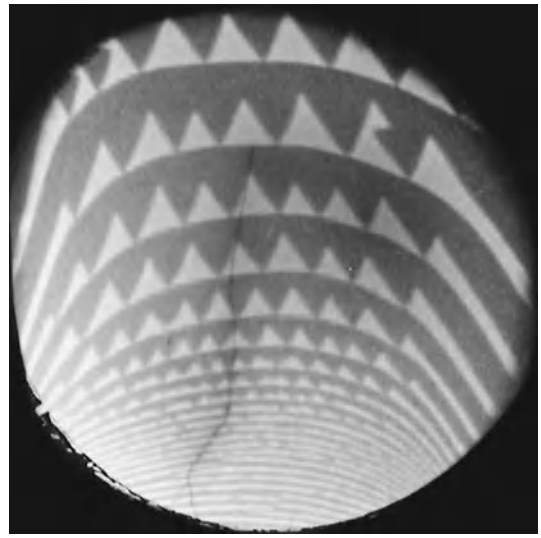

Fig. 3. LEEM image of a partially reconstructed silicon (111) surface. The bright triangles have (7 × 7) structure, the dark regions (1 × 1).

High-Energy Electron Diffraction

HEED is used mainly for the study of the structure of thin foils, films, and small particles (thickness or diameter of 10^{-9} to 10^{-6} m or 4×10^{-8} to 4×10^{-5} in.), of molecules, and also of the surfaces of crystalline materials. A monochromatic, usually nearly parallel, electron beam with a diameter of 10^{-3} to 10^{-8} m (4×10^{-2} to 4×10^{-7} in.) is incident on the target. The forward-scattered electrons (backscattering is negligible) are detected by means of a fluorescent screen, a photoplate, or some other current-sensitive detector, usually without the inelastically scattered electrons being eliminated.

Reflection HEED. The availability of UHV instruments has made it possible to study the structure of surfaces under clean conditions by reflection HEED (RHEED). Many experiments have confirmed the theoretical expectation that RHEED has a sensitivity comparable to that of LEED. Therefore, similar to LEED, RHEED can be used for the determination of the lateral arrangement of the atoms in the topmost layers of the surface (**Fig. 4**), including the structure of adsorbed layers. Although it is more convenient to deduce the periodicity of the atomic arrangement parallel to the surface from LEED patterns than from RHEED patterns, LEED frequently becomes inapplicable when the surface is rough. This usually occurs in the later stages of corrosion or in precipitation, for example, of silicon carbide on silicon, when small crystals grow on the surface. In such investigations RHEED is far superior to LEED because the fast electrons can penetrate the asperities and produce a transmission HEED (THEED) pattern. RHEED has become particularly important for thin-film growth monitoring via the specular beam intensity oscillations caused by monolayer-by-monolayer growth (**Fig. 5**). *See* CRYSTAL GROWTH.

Like LEED, RHEED gives little information on the chemical nature of the atoms that produce the diffraction pattern. Therefore, RHEED has been combined with analytical tools, such as Auger electron spectroscopy or x-ray emission spectroscopy. A second limitation inherent to both RHEED and LEED is the difficulty of determining the atomic positions normal to the surface. This is a consequence of dynamical effects and of absorption. Nevertheless, both techniques have given valuable information on corrosion, precipitation, adsorption, condensation, film growth, and other surface and interface phenomena that could not be obtained by other methods.

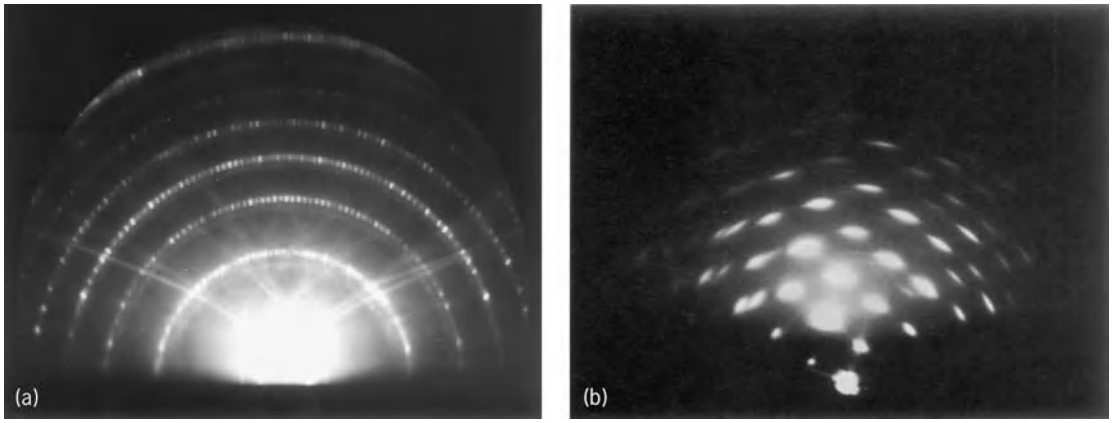

Fig. 4. RHEED pattern from (*a*) flat (111) surface of a silicon single crystal, and (*b*) rough surface of a polycrystalline evaporated film of calcium fluoride (CaF_2) with strong fiber texture, from which no LEED pattern can be obtained.

Scanning HEED. In scanning HEED (SHEED) the diffracted electrons are not recorded on photographic film but are directly measured electronically with sensitive detectors. By moving the detector across the diffraction pattern or by deflecting the diffracted electrons across a stationary detector (scanning), the intensity distribution in the diffraction pattern can be displayed quantitatively on an XY recorder. If an energy filter is put in front of the detector, the inelastically scattered electrons, which are usually not taken into account in the quantitative intensity evaluation, may be filtered out so that only the elastically scattered electrons are measured. This technique is particularly useful in transmission through polycrystalline samples, which produce a ring pattern (Debye-Scherrer diagram). It has also been used in single-crystal samples producing a spot pattern (Laue diagram) and even in reflection diffraction from surfaces. The main application of SHEED is in the study of processes which are accompanied by changes of the intensity distribution, such as the growth of thin films and annealing and corrosion processes.

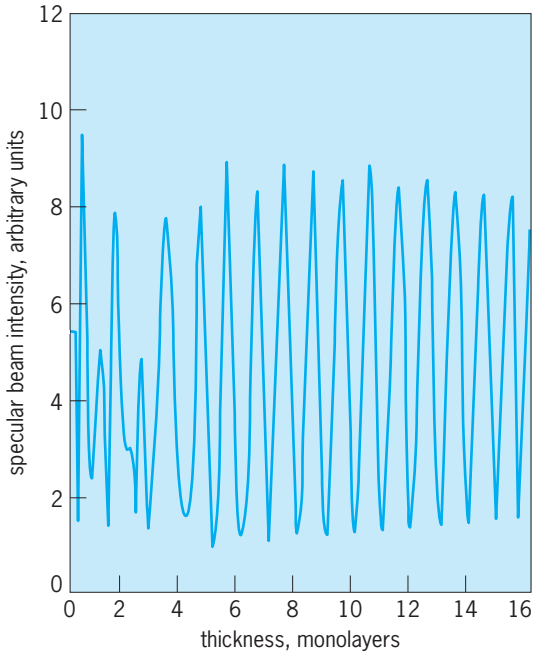

Fig. 5. RHEED specular beam intensity oscillations during molecular beam epitaxy of lead on a silicon (111) surface.

Transmission HEED. The technological importance of thin film and interface devices has led to an upsurge of thin film growth studies by conventional transmission HEED (THEED), usually combined with transmission microscopy. Information obtained this way has been mainly on the orientation of the crystallites composing the film (**Fig. 6**).

Diffraction electron microscopy. In transmission electron microscopy of crystalline specimens the main contrast mechanism is the diffraction contrast. To understand it, knowledge of the diffraction pattern and of diffraction theory is necessary. High-voltage electron microscopes (with electron energies of about 1 MeV) have become available. As a consequence, samples with thickness of the order 10^{-6} m (4×10^{-5} in.) can now be studied, as compared to the former upper thickness limit of about 10^{-7} m (4×10^{-6} in.). This has significance for metallurgy and the physics of metals and semiconductors. It is now possible to study the structure, distribution, and behavior of imperfections in metals, alloys, and semiconductors in bulk with little influence from the boundaries of the material. Other electron-microscope techniques that rely heavily on diffraction are the use of illumination systems producing tilted or conical illumination of the specimen and of objective apertures which transmit only diffracted electrons. With these techniques it has become possible to produce, by diffraction contrast, images of crystal planes 1.5×10^{-10} m (6×10^{-9} in.) apart. RHEED beams can be used for reflection electron microscopy (REM), though with considerable foreshortening due to the grazing incidence on the surface which is necessary for surface sensitivity and intensity.

Photoelectron diffraction. In photoelectron diffraction (PED), the electron wave emitted by one atom

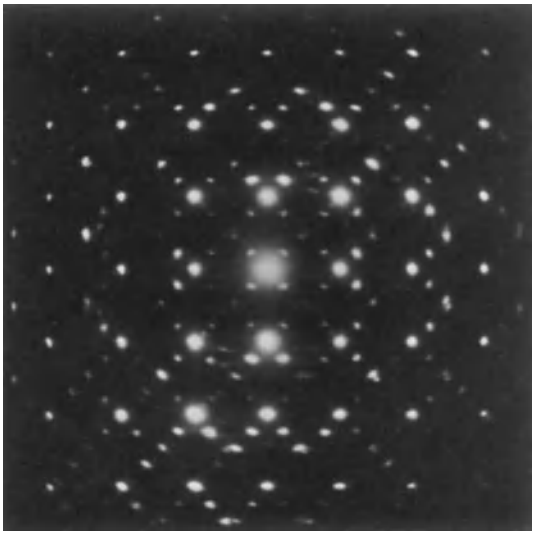

Fig. 6. THEED pattern from epitaxial gold particles on a lead sulfide (PbS) single-crystal film. (*From A. K. Green, J. Dancy, and E. Bauer, Growth of Au on PbS, PbSe, PbTe and SnTe thin film substrates, J. Vac. Sci. Technol., 10:494–502, 1973*)

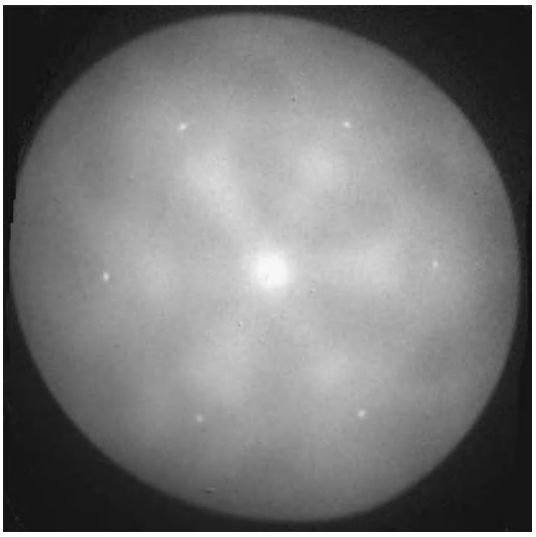

Fig. 7. PED pattern (diffuse features) with superimposed LEED pattern (sharp spots) of a lead film, five atomic layers thick, on a silicon (111) surface. Photon energy is 62 eV; electron energy, 39 eV. (*Courtesy of T. Schmidt*)

upon photoionization is diffracted by the surrounding atoms. The resulting diffraction pattern (**Fig. 7**) allows local structure analysis around the selected atom. At low energies (below several hundred electronvolts) the interpretation requires an effort comparable to that needed in LEED; at higher energies (above about 1000 eV) strong forward scattering frequently permits a simple qualitative analysis. Similar to LEEM, REM, and transmission electron microscopy (TEM), the photoemitted electron wave may also be used for imaging, a method known as x-ray photoemission electron microscopy (XPEEM) with synchrotron radiation. In this case the imaging is atom-specific, just as is PED. E. Bauer

Diffraction in Gases and Liquids

Electron diffraction in gases and liquids is similar in principle to that in solids; the differences arise from the lack in gases and liquids of any highly regular arrangement of the component atoms. In gases the low density makes it possible to study diffraction by individual atoms and molecules. The results obtained from monatomic gases represent the density of electronic charge in the atom as a function of the distance from the nucleus. The results from gaseous polyatomic molecules represent the equilibrium distances between the atomic nuclei and the average amplitudes of vibration associated with these distances. Liquids have been studied much less thoroughly, both in theory and in practice, than have gases.

It should be remembered that the structures of molecules in the gaseous state may be different from those in the liquid or solid states; for example, hydrofluoric acid in the gas forms an $(HF)_6$ ring, while in solution or liquid it forms an infinite chain (—FHFHFHFHFHFHF—); biphenyl in the gas is twisted with conformational equilibria, whereas in the solid state it is coplanar in one frozen conformer.

Applications. Typical questions of molecular structure studied by electron diffraction include those of configuration and size in many molecules, with special interest in the variation of chemical bond distances in related molecules (such as the 6.8-picometers or 2.7×10^{-10} in. decrease in C—F distance in the series CH_3F, CH_2F_2, CHF_3, CF_4), the variation in the angles between chemical bonds, the distinction between geometric isomers, the degree of restricted rotation around chemical bonds, and in general the relations between the geometry of molecules and their energy and chemical behavior. *See* CHEMICAL BONDING; MOLECULAR ISOMERISM; MOLECULAR STRUCTURE AND SPECTRA; STRUCTURAL CHEMISTRY.

The application of electron diffraction in liquids has been restricted because of the less precise information obtainable from molecules which lie in close contact but in irregular arrangements.

The liquids studied include certain metals (mercury, tin, tin-bismuth, and tin-aluminum alloys), paraffin oils, and thin layers of water. For monatomic molecules in the liquid state the distances between nearest neighbors are fairly uniform, whereas the separation between more distant neighbors varies considerably. Evaluation of the closest distance of approach can be made to about 10 pm (4×10^{-10} in.); major developments in the theory are still required.

The most interesting application of electron diffraction to a liquid or amorphous condensed phase has been in the examination of the surface layers on polished solids. The diffuse diffraction patterns often obtained may be the result of an amorphous arrangement in the surface or of the poor resolving power of very tiny crystalline particles. It is probable that both these states are produced in the polishing of different solids.

Theory and techniques. Structural information about gaseous molecules is obtained by having a fine beam of electrons pass through the gas and strike a photographic plate. The electrons in the beam interact with the charged particles in the atoms (electrons and nuclei) and are bent away from the original direction through varying angles, as registered in the pattern on the plate. The observed variation in the number of scattered electrons with increasing angle is interpreted as an interference effect; that is, electron scattering can be described in the language of diffracted waves for which the resultant is the sum of the component wavelets whose amplitudes and phases are influenced by the wavelength of the incident radiation and the relative positions of the scattering centers. The equivalent wavelength of the electrons, λ, is determined by their energy and is given by Eq. (1) or (2). The values of λ commonly used in diffraction by gases are in the range 7–5 $\times$ 10^{-12} m (3–2 $\times$ 10^{-10} in.).

The intensity I of electrons scattered by a spherically symmetrical atom is computed by the Schrödinger wave equation; the result is given by Eq. (3), where k is a constant and f is a complex func-

$$I = k\,|f|^2 \tag{3}$$

tion called the atomic scattering factor. If the atomic number Z is not too large, Eq. (4) holds, where a is the Bohr radius and F is given by Eq. (5). Here $\rho(r)$ is

$$|f|^2 = [2(Z - F)/as^2]^2 \tag{4}$$

$$F = 4\pi \int_0^\infty r^2 \rho(r)[(\sin sr)/sr]dr \tag{5}$$

density of electronic charge at distance r from the nucleus, $s = 4\pi(\sin\theta)/\lambda$ with 2θ equal to the angle between the scattered electron and the orignal beam, and λ is the electron wavelength. Experimental observations on I as a function of s in monatomic gases lead to the determination of $\rho(r)$.

The intensity of electrons scattered by a collection of independent molecules having all possible orientations is given by Eq. (6), where the summations are

$$\begin{aligned} I(s) = k \sum_i & [|f_i|^2 + (2/as^2)^2 S_i] \\ & + \sum_i \sum_j f_i^* f_j \int_0^\infty P_{ij}(r)[(\sin sr)/sr]\,dr \end{aligned} \tag{6}$$

taken over all the atoms in the molecule and $(2as^2)^2 S_i$ represents the intensity of inelastic scattering by the ith atom. The double summation has a term for each pair of atoms i and j. The relative probability of finding the distance between the atom i and j at various values is represented by $P_{ij}(r)$. In principle the use of the observed intensity to determine P_{ij} for each pair of atoms in the molecule constitutes a structure determination for the molecule. The expression is simpler when the atomic motions are nearly harmonic, as in the case of CCl_4. For this molecule the double summation has only two distinct terms. The first is given by expression (7); the second for Cl—Cl is similar.

$$8 f_C f_{Cl}\ \exp(-l_{CCl}^2 s^2)(\sin sr_{CCl})/sr_{CCl} \tag{7}$$

The four parameters which describe the structure are the equilibrium distances, r_{CCl} and r_{ClCl}, and the average amplitudes of vibration, l_{CCl} and l_{ClCl}. Molecular parameters such as these can be determined with high precision when a proper treatment of the scattering factors is included, and when accurate experimental intensities are obtained.

Special equipment is required to meet the conditions assumed in the scattering theory. The electron beam is accelerated with a steady voltage between 30,000 and 70,000 V, which should be constant within about 0.01%. The beam is focused by electrostatic and magnetic lenses so that it has a diameter of no more than 0.1 mm (0.004 in.) at the photographic plate. The whole path of the beam is enclosed in a high-vacuum chamber (pressure of about 10^{-5} mm Hg or 10^{-3} Pa or 10^{-7} lbf/in.) so that no appreciable electron scattering will occur in the residual air. The gas specimen is introduced in a fine jet so that the volume in which the electrons meet the gas is no more than 1.0 mm (0.04 in.) in diameter; high-speed pumping is required to remove the gas as rapidly as possible. In front of the photographic plate a rotating sector is mounted to modify the intensity of electrons reaching the plate so that the normally rapid decrease of intensity with increasing angle of scattering is leveled off in a known way and the emulsion is able to register the incident electrons over a wide range of angle.

Interpretation of results. The pattern observed is a set of light and dark circular bands whose spacings and relative intensities depend on the composition and structure of the specimen molecules. The pattern is scanned by a recording microphotometer which yields (after calibration of the photographic emulsion) a tracing or a digital output of the experimental data of I as a function of s. These data are interpreted with the aid of a Fourier transform of the intensity expression of Eq. (6) to give the P_{ij} functions for the pairs of atoms. From the positions and widths of the peaks for CCl_4 in **Fig. 8**, it is determined that Eqs. (8) hold (1 pm =

$$\begin{aligned} r_{CCl} &= 176.6 \pm 0.3 \text{ pm} \\ l_{CCl} &= 6.0 \pm 0.5 \text{ pm} \\ r_{ClCl} &= 288.7 \pm 0.4 \text{ pm} \\ l_{ClCl} &= 6.8 \pm 0.3 \text{ pm} \end{aligned} \tag{8}$$

10^{-12} m = 3.937×10^{-11} in.). *See* FOURIER SERIES AND TRANSFORMS.

This relatively high precision in determining the distances in gas molecules has been achieved since about 1950 with the development of new instrumentation and the application of high-speed computing methods. The relative intensities of scattered electrons registered on the photographic plate can be measured to better than 0.1%; the time saved by the

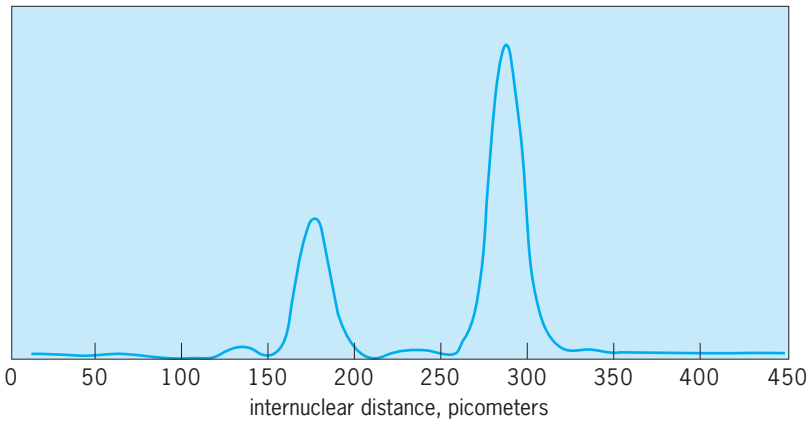

Fig. 8. Experimental distribution of internuclear distances in CCl_4. The two prominent peaks represent the C-Cl and Cl-Cl distances.

use of electronic computing methods in interpreting the data permits a more rigid application of the criteria for satisfactory agreement between experiment and theory.

Over a thousand gases have been studied by the methods giving internuclear distances within 0.4 pm (1.6×10^{-11} in.). In favorable cases precisions exceeding 0.1 pm (4×10^{-12} in.) have been obtained; when these are compared with spectroscopic results, good agreement is found if allowance is made for the difference in the nature of the distances measured. The molecular structures of some 500 gases have been reported by earlier electron diffraction procedures with uncertainties 10 times larger or more.

The applicability of the gas diffraction method alone is limited to simple molecules. Only when the three-dimensional structure can be derived uniquely from a one-dimensional spectrum of internuclear distances blurred by thermal motion is a complete structure determination possible from electron diffraction data alone; light atoms in the presence of heavy ones in the same molecule are less precisely located; distances as close as 3 pm (1.2×10^{-10} in.) can barely be resolved. The range of the method is greatly increased, however, when some structural features in the molecule can be assumed from the results of other methods of investigation. For example, if a six-membered ring of atoms is known to have trigonal symmetry, the number of structural parameters is decreased from 12 to 2. *See* NEUTRON DIFFRACTION; SCATTERING EXPERIMENTS (ATOMS AND MOLECULES); SCATTERING EXPERIMENTS (NUCLEI). Lawrence O. Brockway

Bibliography. W. Braun, *Applied RHEED: Reflection High-Energy Electron Diffraction During Crystal Growth*, Springer-Verlag, 1999; J. M. Cowley, *Diffraction Physics*, 3d ed., 1995; J. M. Cowley (ed.), *Electron Diffraction Techniques*, 2 vols. 1992, 1993; D. L. Dorset, *Structural Electron Crystallography*, 1995; C. S. Fadley, The study of structures by photoelectron diffraction and Auger electron diffraction, in R. Z. Bachrach (ed.), *Synchrotron Radiation Research: Advances in Surface and Interface Science*, 1992; A. Howie and U. Valdre (eds.), *Surface and Interface Characterization by Electron Optical Methods*, 1988; P. K. Larsen and P. J. Dobson (eds.), *Reflection High-Energy Electron Diffraction and Reflection Electron Imaging of Surfaces*, 1988; D. L. Misell and E. B. Brown, *Electron Diffraction: An Introduction for Biologists*, 1988; M. A. van Hove, W. H. Weinberg, and C. M. Chan, *Low Energy Electron Diffraction*, 1987.

Electron emission

The liberation of electrons from a substance into vacuum. Since all substances are built up of atoms and since all atoms contain electrons, any substance may emit electrons; usually, however, the term refers to emission of electrons from the surface of a solid.

The process of electron emission is analogous to that of ionization of a free atom, in which the latter parts with one or more electrons. The energy of the electrons in an atom is lower than that of an electron at rest in vacuum; consequently, in order to ionize an atom, energy must be supplied to the electrons in some way or other. By the same token, a substance does not emit electrons spontaneously, but only if some of the electrons have energies equal to, or larger than, that of an electron at rest in vacuum. This may be achieved by various means. If a substance is heated, the atoms begin to vibrate with larger amplitudes, and electrons may absorb sufficient energy from these vibrations to be emitted in the process known as thermionic emission. Electrons may also be liberated upon irradiation of the substance with light (photoemission). Electron emission from a substance may be induced by bombardment with charged particles such as electrons or ions in the phenomenon called secondary emission. Field emission, or cold emission, is the emission of electrons under influence of a strong electric field. Electrons may also be emitted from one solid into another, but this is usually referred to as electron injection. For example, a metal may inject electrons into an insulator under certain circumstances. *See* FIELD EMISSION; PHOTOEMISSION; SECONDARY EMISSION; THERMIONIC EMISSION. Adrianus J. Dekker

Bibliography. R. Gomer, *Field Emission and Field Ionization*, 1961, reprint 1993; G. Hohler, E. A. Nieksch, and J. Treusch (eds.), *Particle Induced Electron Emission*, 1991; A. Modinos, *Field, Thermionic and Secondary Electron Emission Spectroscopy*, 1984.

Electron-hole recombination

The process in which an electron, which has been excited from the valence band to the conduction band of a semiconductor, falls back into an empty state in the valence band, which is known as a hole. *See* HOLE STATES IN SOLIDS.

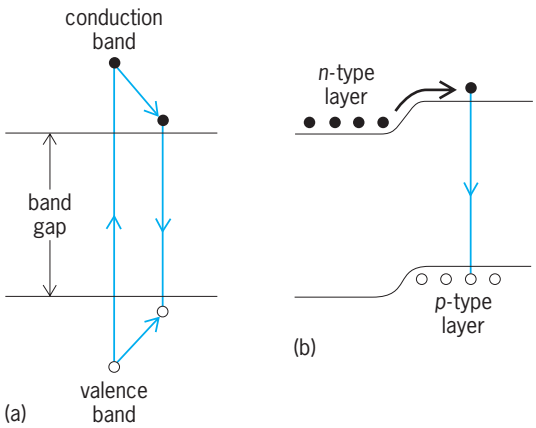

Fig. 1. Recombination of electrons and holes generated by (*a*) optical absorption and (*b*) a forward-biased *pn* junction.

When atoms come together to form a crystal, the discrete atomic energy levels broaden into a level continuum. The periodic potential of the crystal produces gaps in the continuum, separating it into a number of energy bands. In a pure semiconductor crystal, bonding electrons fill the valence band, and above the valence band in energy is an empty conduction band. The two bands are separated by the band gap, which has no levels, and the extent of this band gap is a characteristic of each semiconductor. The band gap is of order 1 eV for common semiconductor materials such as silicon (Si) or gallium arsenide (GaAs). *See* BAND THEORY OF SOLIDS.

Light with photon energies greater than the band gap can be absorbed by the crystal, exciting electrons from the filled valence band to the empty conduction band (**Fig. 1***a*). The state in which an electron is removed from the filled valence band is known as a hole. It is analogous to a bubble in a liquid. The hole can be thought of as being mobile and having positive charge. The excited electrons and holes rapidly lose energy (in about 10^{-12} s) by the excitation of lattice phonons (vibrational quanta). The excited electrons fall to near the bottom of the conduction band, and the holes rise to near the top of the valence band, and then on a much longer time scale (of 10^{-9} to 10^{-6} s) the electron drops across the energy gap into the empty state represented by the hole. This is known as electron-hole recombination. An energy approximately equal to the band gap is released in the process. Electron-hole recombination is radiative if the released energy is light and nonradiative if it is heat. *See* PHONON.

Recombination at pn junctions. Electron-hole recombination requires an excited semiconductor in which both electrons and holes occupy the same volume of the crystal. This state can be produced by purely electrical means by forward-biasing a *pn* junction. The *pn* junction is formed by adjacent *p*-type and *n*-type layers containing excess holes and excess electrons, respectively. In the *n*-type layer, donor impurities replace some of the host atoms. Each donor impurity has one more electron in its outer shell than the host atom, and these extra electrons go into the conduction band. Likewise, in the *p*-type layer, acceptor impurities replace some of the host atoms. The acceptor atoms contain one less electron in their outer shell than the host atoms. They complete their valence bonds by taking a valence electron from elsewhere in the crystal, thereby creating holes in the valence band.

The *pn* junctions forms a barrier that keeps the electrons and holes separated. By contacting these layers and applying a forward-bias voltage to the *pn* junction, the barrier height can be reduced, enabling electrons to flow into the *p*-region or holes into the *n*-region, thus allowing electron-hole recombination to take place. The current passing through a *pn* diode in electrons per second equals the rate of electron-hole recombination (Fig. 1*b*). A major application of this phenomenon is the light-emitting diode (LED), which generates light by means of radiative electron-hole recombination resulting from forward-biasing the *pn* junction. *See* LIGHT-EMITTING DIODE; LUMINESCENCE; SEMICONDUCTOR DIODE.

Radiative recombination. Semiconductor materials may be divided into two types: direct-band-gap semiconductors and indirect-band-gap semiconductors. Efficient radiative recombination between free electrons and holes takes place only in direct-band-gap semiconductors. The continuous energy levels of electrons in the conductor and valence bands are characterized by a definite value of momentum (or more precisely, wave vector). During an optical transition, momentum is conserved, and since the photon carries away negligible momentum, transitions take place only between conduction-band and valence-band states having the same momentum. This is easily satisfied in direct-band-gap semiconductors, because electrons and holes collect at the conduction band at minimum and the valence band at maximum, and both extrema have the same momentum (**Fig. 2***a*). However, for indirect-band-gap semiconductors (Fig. 2*b*), the conduction-band minimum and valence-band maximum have very

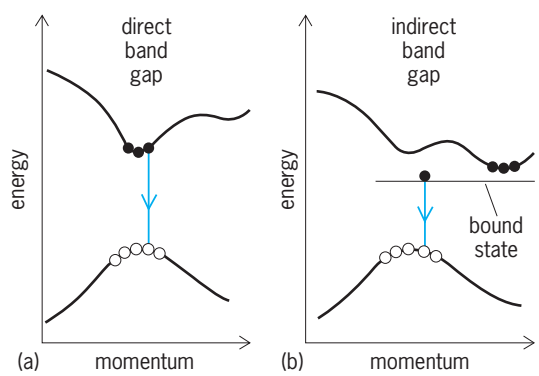

Fig. 2. Energy versus momentum for the conduction-band and valence-band states of (*a*) a direct-band-gap semiconductor and (*b*) an indirect-band-gap semiconductor. Optical transitions in indirect-band-gap semiconductors take place through bound states that have a momentum distribution as indicated.

different momenta, and consequently optical transitions between free electrons and holes are forbidden.

Radiative electron-hole recombination is possible in indirect-band-gap semiconductors when the transition is assisted by lattice phonons and impurities. Weak optical transitions can take place in which a momentum-conserving phonon is emitted along with the photon. Electrons can be trapped at impurities having energy levels within the energy gap. Unlike the free electron, the bound electron has a spread of momentum as a consequence of its localization (the uncertainty principle), with some components of momentum at the valence-band maxima (Fig. 2*b*). The isoelectronic impurity nitrogen substituted for phosphorus in gallium phosphide (GaP) and gallium arsenide phosphide (GaAsP) is used for this purpose. When heavily doped with nitrogen, these indirect semiconductors have moderate luminescence efficiency and are used as light-emitting diodes to generate green light. Bismuth impurities can be used in the same way to produce yellow light. It would be more desirable to use direct-band-gap semiconductors for this purpose, but currently there are no direct-band-gap semiconductors capable of forming *pn* junctions which emit these colors at room temperature. *See* CRYSTAL DEFECTS.

Apart from its application in light-emitting diodes and laser operation, radiative recombination, especially at low temperatures (approximately 2 K or −456°F), has been a very important tool for studying the interaction of electrons and holes in semiconductor crystals. At low temperatures, many weakly bound states such as excitons are able to form stably. These states interact only weakly with lattice vibrations and consequently emit much of their recombination radiation in the form of sharp (zero phonon) spectral lines. For example, at low temperatures, electrons become trapped at donor sites and holes at acceptor sites. The spectrum of donor-acceptor recombination contains hundreds of resolvable lines, one for each crystallographically inequivalent donor-acceptor pair. *See* EXCITON.

Nonradiative recombination. Competing with radiative recombination are the nonradiative recombination processes of multiphonon emission and Auger recombination. Nonradiative recombination of electrons and holes by multiphonon emission occurs in two steps. First, an electron is captured into a deep level near the middle of the energy gap, and subsequently the trapped electron state captures a hole from the valence band (**Fig. 3**). Multiphonon emission requires atoms very different from the host, or defects such as vacancy-impurity complexes, dislocations, or semiconductor-oxide interface states. Electron capture alters the positions of the surrounding lattice, in a process known as lattice relaxation that lowers the electronic energy of the bound electron (Fig. 3). The vibration of the lattice causes the electronic level to move within the energy gap and even cross into the bands. Capture takes place during the crossing.

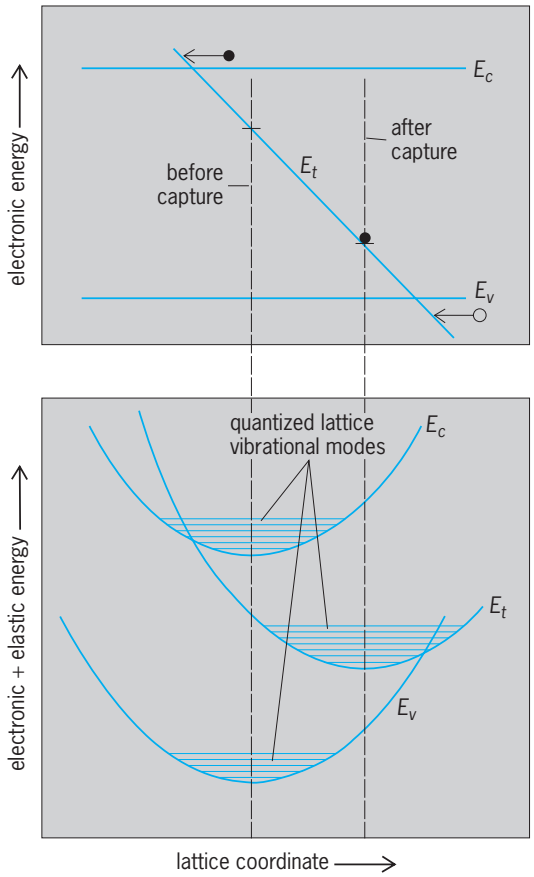

Fig. 3. Electron and hole capture into a deep level by multiphonon emission. After capture the lattice relaxes, lowering the electronic energy of the bound state E_t. E_c and E_v are the minima of the conduction band and maxima of the valence band, respectively.

In multiphonon emission, the energy lost during capture highly excites the vibrations of one or several atoms related to the defect site. This can greatly enhance the probability for an atom to jump to a different site, which also requires a highly excited atomic vibration. Such recombination-enhanced motion has been observed in detail for defects formed by ionizing radiation in GaAs and in Si. It is suspected that nonradiative recombination by multiphonon emission drives the movement of atoms at room temperature that are responsible for device degradation phenomena such as the climb of dislocations found in GaAs light-emitting diodes and lasers. *See* DIFFUSION; RADIATION DAMAGE TO MATERIALS.

Another source of nonradiative recombination is the Auger effect, in which an electron or hole takes up the energy released in the recombination of a pair of other charge carriers. Auger recombination has been observed for excitons bound to neutral donors in Si and GaP, a complex of two electrons and a hole. In these cases, the Auger recombination rate is several orders of magnitude greater than the radiative rate. Auger recombination has been shown to limit the performance of long-wavelength (1.3–1.6 micrometer) lasers and light-emitting diodes used in optical communication systems. The Auger effect reduces light-emitting diode

efficiency and greatly enhances the increase of laser threshold current with temperature. *See* AUGER EFFECT; LASER; OPTICAL COMMUNICATIONS; SEMICONDUCTOR. Charles H. Henry

Bibliography. V. N. Aba Kumov, V. I. Perel, and I. N. Yassievich, *Nonradiative Recombination in Semiconductors*, 1991; G. Gillessen and W. Schairer, *Light-Emitting Diodes: An Introduction*, 1987; P. T. Landsberg, *Recombination in Semiconductors*, 1992; J. Singh, *Semiconductor Devices An Introduction*, 1994; S. M. Sze, *Semiconductor Devices*, 1985.

Electron lens

An electric or magnetic field, or a combination thereof, which acts upon an electron beam in a manner analogous to that in which an optical lens acts upon a light beam. Electron lenses find application for the formation of sharply focused electron beams, as in cathode-ray tubes, and for the formation of electron images, as in infrared converter tubes, various types of television camera tubes, and electron microscopes.

Any electric or magnetic field which is symmetrical about an axis is capable of forming either a real or a virtual electron image of an object on the axis which either emits electrons or transmits electrons from another electron source. Hence, an axially symmetric electric or magnetic field is analogous to a spherical optical lens.

The lens action of an electric and magnetic field of appropriate symmetry can be derived from the fact that it is possible to define an index of refraction n for electron paths in such fields. This index depends on the field distribution and the velocity and direction of the electrons. It is given by the equivalencies shown below. Here e is the charge of the electron,

$$n = \sqrt{\phi + \frac{2e\phi^2}{mc^2}} - \sqrt{\frac{e}{2m}}\, A \cos\chi$$

$$= \sqrt{\phi + 0.978 \cdot 10^{-6}\phi^2} - 0.297A \cos\chi$$

m its rest mass, ϕ the potential of the point in space under consideration (so normalized that the kinetic energy of the electron vanishes for $\phi = 0$), c the velocity of light, A the magnetic vector potential, and χ the angle formed by the path of the electron with the direction of the magnetic vector potential. For an axially symmetric field the magnetic vector potential is perpendicular to the plane passing through the axis of symmetry and the reference point. Its magnitude is equal to the magnetic flux through the circle about the axis through the reference point divided by the circumference of that circle. The numerical coefficients in the final expression for n apply if ϕ is measured in volts and A in gauss-centimeters. *See* ELECTRON MICROSCOPE.

Electron lenses differ from optical lenses both in the fact that the index of refraction is continuously variable within them and that it covers an enormous range. Furthermore, in the presence of a magnetic field, n depends both on the position of the electron in space and on its direction of motion. It is not possible to shape electron lenses arbitrarily. *See* ELECTROSTATIC LENS; MAGNETIC LENS. Edward G. Ramberg

Bibliography. D. A. DeWolf, *Basics of Electron Optics*, 1990; E. M. Slayter and H. S. Slayter, *Light and Electron Microscopy*, 1993.

Electron microscope

A device for forming greatly magnified images of objects by means of electrons. Electron microscopes serve primarily two purposes: the visual examination of structures too fine to be resolved with ordinary, or light, microscopes, and the study of surfaces that emit electrons. The first function made transmission electron microscopes essential research tools in biology, chemistry, and metallurgy. Beginning in the 1960s the scanning electron microscope came to play an increasingly important role in the study of the surfaces of solid objects at more moderate magnifications. Various emission electron microscopes serve more specialized research purposes.

Instruments

A transmission electron microscope, shown in close analogy with a light microscope in **Fig. 1**, consists in its simplest form of a source supplying a beam of electrons of uniform velocity, a condenser lens for concentrating the electrons on the specimen, a specimen stage for displacing the specimen which transmits the electron beam, an objective lens, a projector lens, and a fluorescent screen on which the final image is observed. For permanent record of the

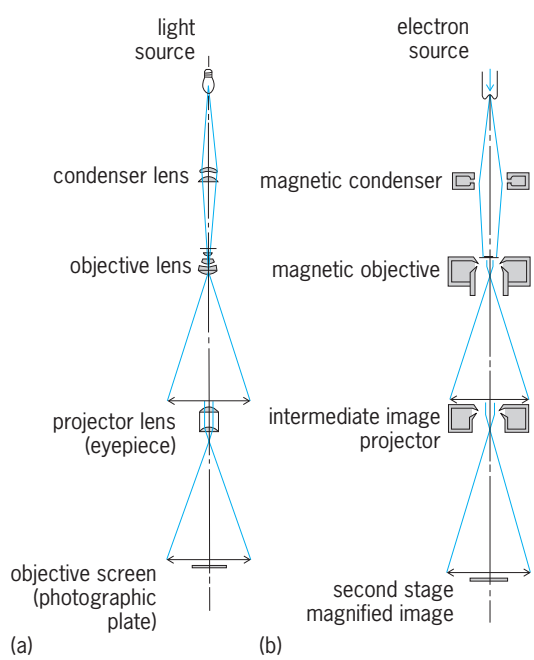

Fig. 1. Comparison between (*a*) light microscope and (*b*) magnetic electron microscope. (*After V. K. Zworykin et al., Electron Optics and the Electron Microscope, Wiley, New York, 1945*)

image, the fluorescent screen is replaced by a photographic plate or film.

Electrons are strongly scattered by all forms of matter including air. Hence the entire instrument must be evacuated to about 10^{-4} mmHg (10^{-7} atm or 10^{-2} pascal). Furthermore, the lenses cannot be material in nature. Instead, they are electric or magnetic fields, symmetrical about the axis of the instrument, that have the property of bending the electron paths toward the axis, just as converging glass lenses bend light rays toward their axis. The lenses in the instruments shown in Fig. 1 are magnetic fields formed at narrow gaps in the iron casings surrounding coils traversed by electric current. Lens strength is varied by varying the current. Most electron microscopes employ magnetic lenses of this type. These have yielded the highest resolution and magnification attained.

However, good results have also been obtained with electron microscopes employing unipotential electrostatic lenses and magnetic lenses excited by permanent magnets. *See* ELECTRON LENS; ELECTROSTATIC LENS; MAGNETIC LENS.

Resolution. A microscope can, at best, permit the discrimination of two point objects greater than $0.7\lambda/\sin\theta$ apart. Here λ is the wavelength of the illuminating radiation, and θ is the aperture angle of the cone of radiation that participates in forming the image (**Fig. 2**). For green light $\lambda = 500$ nanometers. Even for ultraviolet radiation in an immersion medium of refractive index 1.5, λ is no less than 170 nm.

Since light and ultraviolet microscope objectives can be designed to utilize practically all the radiation passing through the specimen, $\sin\theta \cong 1$ and the least resolvable distance for the ultraviolet microscope is about 100 nm. For 50- to 100-kV electrons, such as are commonly employed in electron microscopes, the wavelength range is 0.0053–0.0037 nm. Hence, even though a cone of radiation with an aperture angle less than 0.01 radian contributes to an image of optimum sharpness, object separations smaller than 0.3 nm have been resolved with the electron microscope.

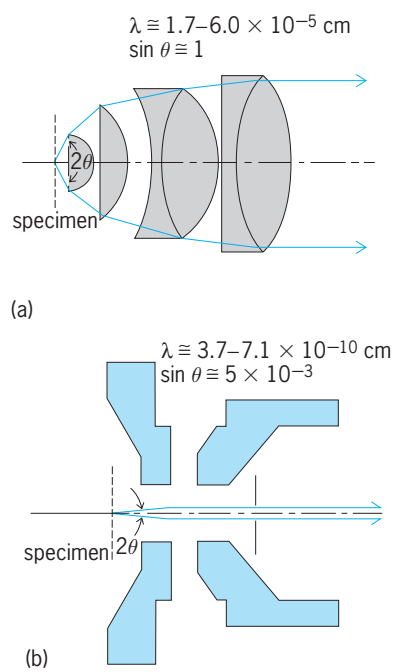

Fig. 2. Effective objective aperture in (*a*) light and (*b*) electron microscopes.

Thus the electron microscope has several hundred times the resolving power of the light microscope. Similarly, whereas the maximum useful magnification of the light microscope is about 2000, that of the electron microscope may approach 1,000,000. The maximum useful magnification is the least magnification of the image that reveals to the observer all the specimen detail that the microscope is capable of conveying.

Transmission electron microscope. Electrons are commonly emitted from the tip of a fine tungsten-wire hairpin filament or, to further reduce the size of the effective electron source, from a sharply pointed segment of wire welded to the filament tip (**Fig. 3**). The filament is maintained at a carefully stabilized negative potential of 50–100 kV with respect to the remainder of the instrument. Electrons enter the instrument through an anode aperture. The intensity and convergence of the electron beam that is falling on the specimen are adjusted by varying the coil current of the condenser lens.

Image contrasts are formed by the scattering of electrons out of the narrow cone that contributes to the formation of the image; denser or thicker portions of the specimen scatter more electrons and hence appear darker in the image. The sharpness of the image observed on the screen is adjusted by varying the objective coil current, and its magnification by varying the projector coil current. Both currents must be carefully stabilized to yield high resolution. In modern instruments the function of the projector is performed by two magnetic lenses in tandem. In the instrument shown in Fig. 3 the normal magnification range from 1400 to 200,000 is realized by varying the coil currents. The magnification range can be extended downward to 500 with the aid of a special specimen holder and projector polepiece. A binocular viewer provides an additional magnification by 10, needed for the fine adjustment of the objective current to yield images of maximum sharpness. Comparable magnifications of the final image are obtained by enlarging the fine-grain photographic plates on which the electron image is recorded.

The image intensity required for precise visual focusing and specimen exploration is considerably higher than that needed for photographic recording. This intensity, and the corresponding specimen damage by the electron beam, can be greatly reduced through the use of an image intensifier optically coupled to the image on the fluorescent screen of the electron microscope on the one hand and the target of a television camera tube on the other (**Fig. 4**). The picture signal generated by the camera tube can be employed to reproduce the microscope image instantaneously on the screen of a television receiver, can be recorded on tape by a video recorder for instant playback at a later time, or can be distributed

to a group of receivers for teaching purposes. Since the image can be photographed on the screen of the television receiver, this system makes it also possible to avoid contamination of the vacuum in the microscope column resulting from the introduction of photographic material. Furthermore, it permits arbitrary control of image contrasts.

In addition to the standard transmission microscopes operating at 50–100 kV, a number of very high-voltage instruments have been constructed. Instruments designed for operation up to 1500 kV (**Fig. 5**) are necessarily large, because clearances in the gun and the high-voltage supply are required to suppress field emission, and because the focal lengths of electron lenses increase as the accelerating voltage of the electrons is increased. The advantage of high-voltage electron microscopy does not lie in greater resolving power—the effect of a shorter electron wavelength (0.00087 nm at 1,000,000 eV) is more than outbalanced by the increased aberration of the lenses—but in increased penetration, which is particularly valuable in the direct study of metal sections prepared with a microtome.

Scanning electron microscope. A modern version of the scanning electron microscope is illustrated in **Fig. 6**. The microscope column, with the source at the top and the specimen chamber at the bottom, is at the left. The square cathode-ray tube display for visual observation with a long-persistence yellow phosphor is in the center, and the cathode-ray tube display for photographic recording with a short-persistence blue phosphor is at the right. **Figure 7** is a schematic diagram of the microscope column. The electron gun is similar to that in the transmission microscope, but operates at much lower voltages (1–20 kV). Three magnetic lenses image the first crossover, or narrowest cross section of the beam issuing from the filament tip, at a greatly reduced scale on the specimen. A pair of deflection coils actually mounted within the bore of the final lens serves to deflect the electron probe formed by the lenses over the surface of the specimen, the specimen and the cathode-ray displays being scanned in synchronism. The secondary electrons generated at the point of impact of the electron beam on the specimen are accelerated by 12 kV and impinge on a plastic scintillator optically coupled by a plastic light pipe to a multiplier phototube. The output current of this phototube controls the beam current, and hence the brightness in the display. *See* SCINTILLATION COUNTER.

The whole scanning electron microscope can be regarded as a slow-scan television system, with the microscope column constituting a very high-resolution camera tube, the specimen serving as camera tube target. Resolution is determined by the electron probe size at the specimen and magnification by the ratio of the scan amplitude in the display to that at the specimen. Thus the magnification can be varied, for example, in a range of 20–100,000, without any refocusing. The effective probe size is determined not only by the electron lens properties but also by the scattering of the electrons within the specimen. Since the range of electron diffusion decreases as the accelerating voltage of the electrons is reduced, relatively low operating voltages are desirable. The optimal resolution cited for the instrument in Fig. 6 is better than 20 nm, the normal operating resolution about 50 nm. At the same time, the great depth of focus (about 300 times that of a light microscope) endows images with considerable extent in depth with an extraordinarily lifelike, three-dimensional quality. This is lacking in light microscope images of the same magnification, since the image is sharp only at a rather sharply defined planar section through the object surface. *See* CATHODE-RAY TUBE; TELEVISION CAMERA TUBE.

Since the electron probe currents are very small (10^{-9} to 10^{-13}), the recording of noise-free high-resolution images demands relatively long scan periods, ranging from 2 to 1000 s; for visual observation

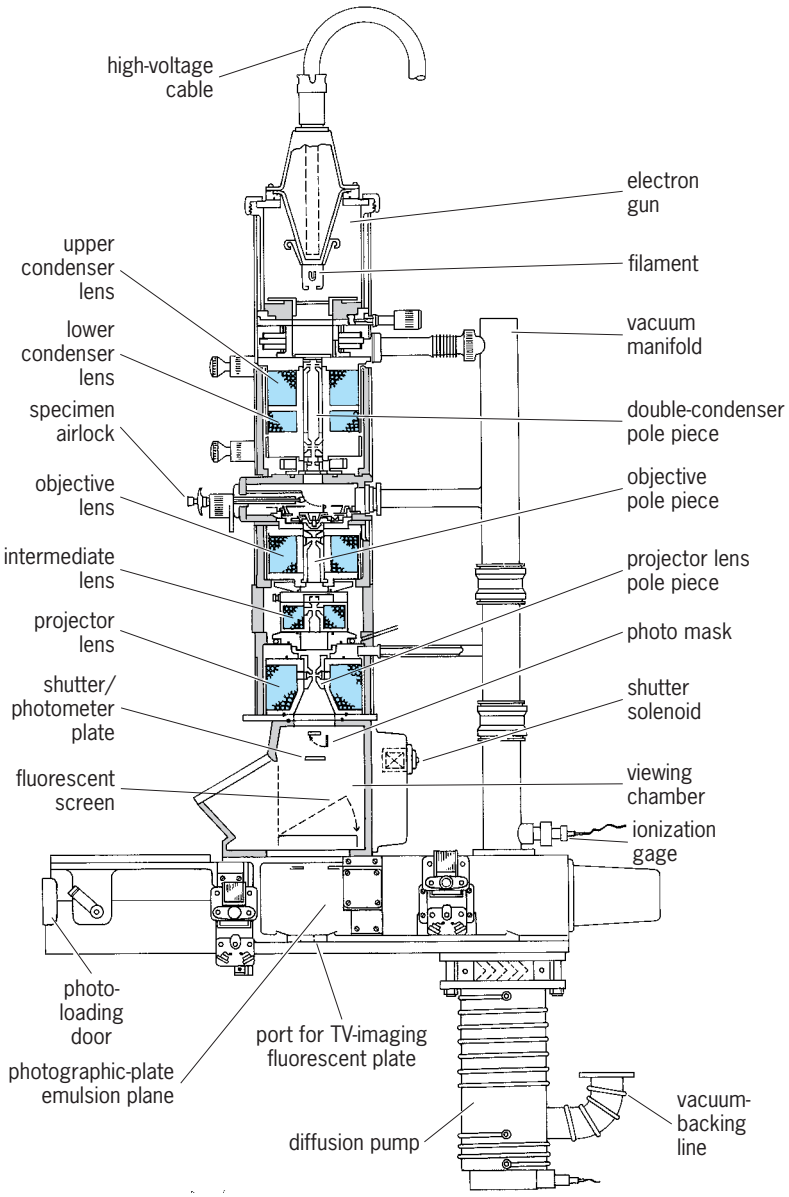

Fig. 3. Section through column of an electron microscope. This instrument has a normal magnification range from 1400 to 200,000. (*Radio Corporation of America*)

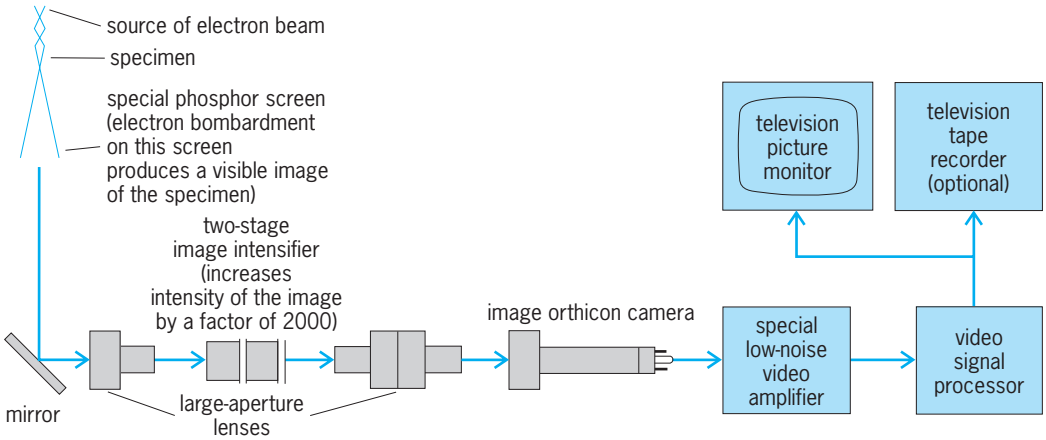

Fig. 4. Scanning electron microscope. (*Cambridge Instrument Co., Ltd.*)

on the long-persistence screen the frame period is reduced to 0.1 to 12 s.

The signal pickup system can be modified in various ways to show different properties of the object. In the examination of integrated circuit elements a comparison of the image obtained in the conductive mode (**Fig. 8**), the signal being derived from an appropriately chosen electrode on the specimen, with that obtained in the normal emissive mode is frequently useful in testing the operation of the system and locating defects. *See* INTEGRATED CIRCUITS; SCANNING ELECTRON MICROSCOPE.

Emission electron microscopes. These may take several forms (**Fig. 9**). In the immersion electron microscope, for which a rudimentary form is shown in Fig. 9*a*, the specimen is a flat conducting surface which may be heated, illuminated, or bombarded by high-velocity electrons or ions so as to emit low-velocity thermionic, photo, or secondary electrons. These are accelerated to a high velocity in an immersion objective or cathode lens and imaged as in a transmission electron microscope.

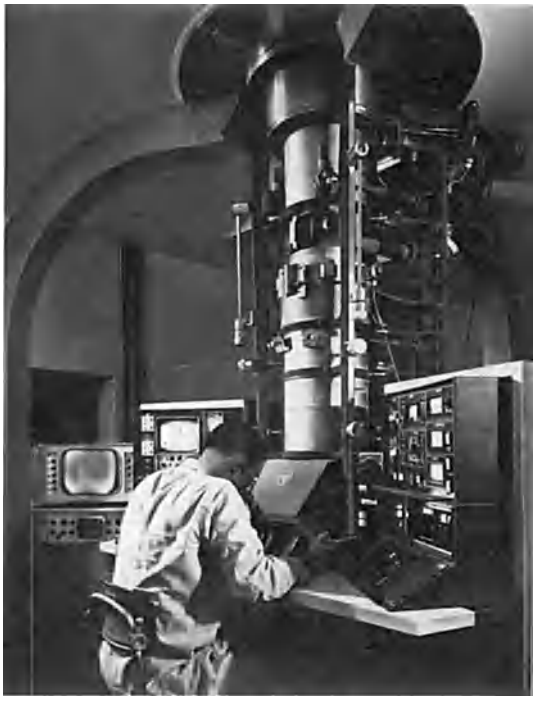

Fig. 5. TV image intensification system for the electron microscope. (*Radio Corporation of America*)

Other emission microscopes employ simple projection without lenses. One example (Fig. 9*b*) is the cylindrical microscope used for examining the thermionic emission from a thin wire stretched along the axis of a cylindrical tube whose walls at high positive potential are coated with a fluorescent substance. Another (Fig. 9*c*) is the field-emission microscope, in which field electron currents are drawn from an extremely fine, rounded point on the end of a wire of a metal such as tungsten. The electrons drawn out of the point follow nearly straight lines toward a positive-potential screen. The display on the screen shows variations in the emission characteristics (as determined by the work function and surface nonuniformities) over the surface of the point. *See* FIELD-EMISSION MICROSCOPY; THERMIONIC EMISSION.

Low-temperature techniques. Apart from freeze-etching, low-temperature techniques find several other applications in electron microscopy.

One of these is employment of a superconducting objective lens, as demonstrated by H. Fernandez-Moran. The lens field is produced by an air-core solenoid of niobium-zirconium alloy, cooled by a liquid-helium bath and short-circuited after excitation so as to remain in the persistent-current mode. While the immersion of the microscope column in liquid helium presents a very serious inconvenience, such superconducting solenoids may generate higher magnetic fields and can thus provide, in principle, lenses of shorter focal length and lower aberration than conventional iron pole-piece objectives. Furthermore, they are characterized by extraordinary stability. *See* SUPERCONDUCTIVITY.

Liquid air–cooling of the specimen chamber also serves to suppress contamination deposited on the specimen through the interaction of the electron beam with traces of organic vapor; without special precautions, contamination buildup can quickly

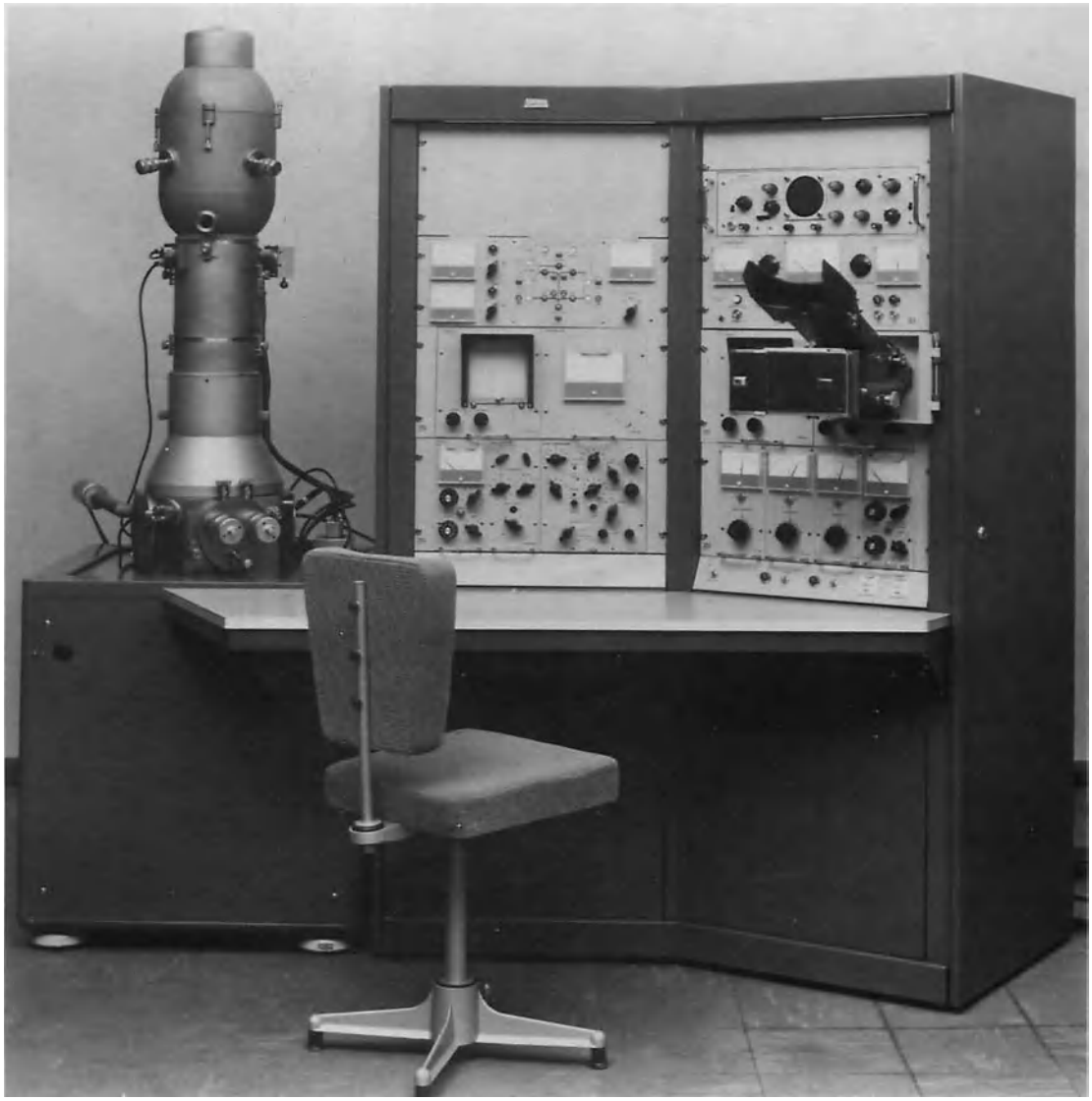

Fig. 6. Megavolt electron microscope at U. S. Steel research laboratory at Monroeville, Pennsylvania. (*Radio Corporation of America*)

obscure fine specimen detail at very high magnifications.

Finally, freeze-drying and freeze substitution provide ways of preparing organic specimens for observation with minimal distortion of their cell structure. In the former instance the specimen may be cooled extremely rapidly by immersion in liquid air–cooled propane and then be warmed slowly in vacuum, so that the ice sublimes rather than melts. The infusion of the embedding material then takes place either in vacuum or a dry-nitrogen atmosphere. In freeze substitution ice is removed by diffusion into a liquid solvent instead of sublimation in vacuum. Edward G. Ramberg

Electron Microscopy

The application of the electron miscroscope to examination and investigation of the ultramicrostructure of materials has become so extensive that there is hardly an area in biological and nonbiological research where electron microscopy does not play a role. In biological and medical research the development of sectioning techniques has extended electron microscopy down to observations at the macromolecular level for delineation of the complex organization of cell components such as membranes, mitochondria, endoplasmic reticulum, and ribosomes. Other techniques have been developed for studies on the structure of virus and even individual proteins and nucleic acids. Nonbiological solid materials have also become objects of extensive and fruitful investigation. Diffraction microscopy, together with the development of adequate techniques of sectioning and thinning crystalline materials such as metals, now makes the observation of defect structure in solids an important aspect of electron microscopy.

The very different character of the interaction of electrons and light with matter produces the greatest practical differences between electron microscopy and optical microscopy. Contrast in the optical microscope image is produced largely by

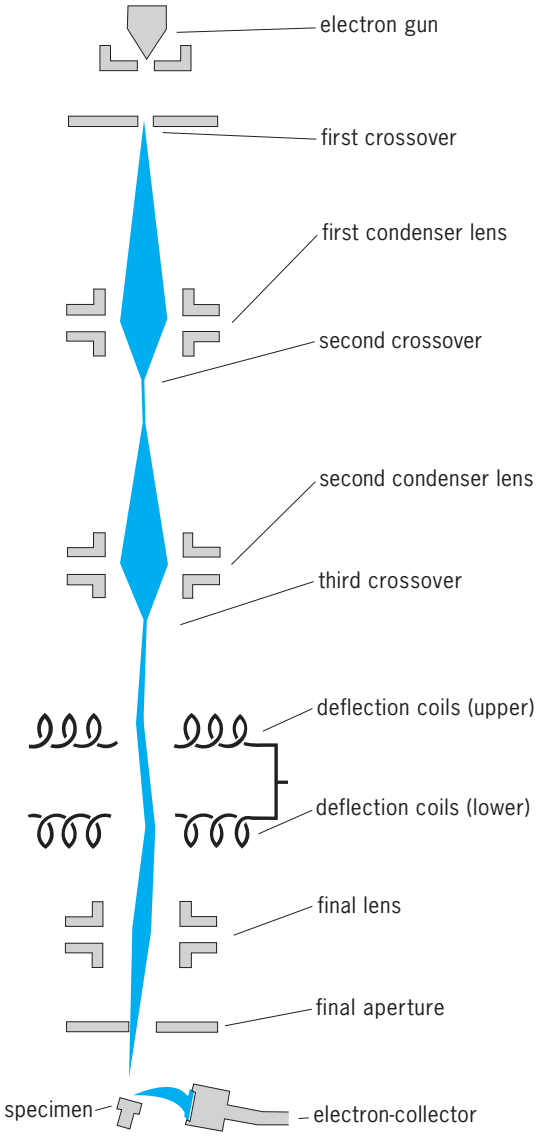

Fig. 7. Scanning electron microscope column. (*Cambridge Instrument Co., Ltd.*)

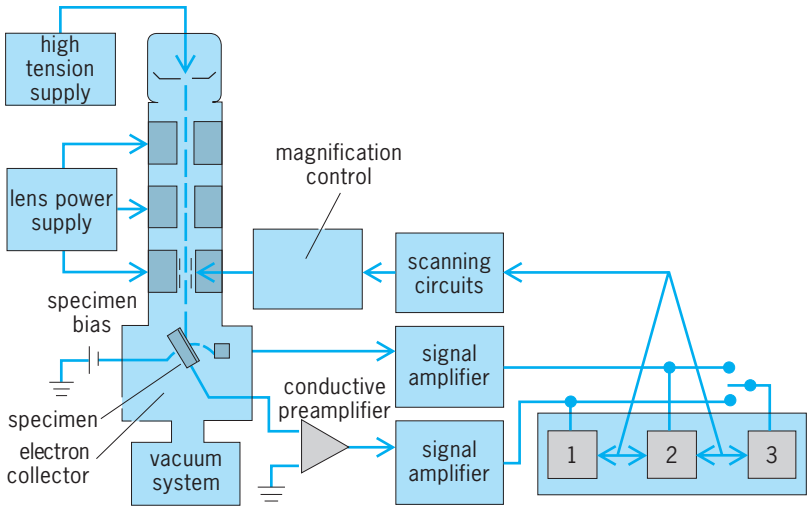

Fig. 8. Alternative operation of scanning electron microscope in conductive and emissive mode. (*Cambridge Instrument Co., Ltd.*)

spectral adsorption of light passing through the specimen. Contrast in the electron microscope image is produced primarily by the scattering of electrons from the specimen. Electrons in the 50–100-kV range, used in conventional commercial electron microscopes, are strongly scattered by matter. It is this strong interaction that determines the requirements for specimen preparation and places serious limitations on the examination of materials. As indicated previously, to avoid scattering of electrons by ambient gas molecules, an electron beam must be confined to a column evacuated to a pressure of the order of 10^{-4} to 10^{-5} mmHg (10^{-7} to 10^{-8} atm or 10^{-2} to 10^{-3} Pa). Thus, specimens must be so prepared that they can be placed in a high vaccum. Dehydration is required, and this requirement immediately places a severe restraint on all biological systems. Moreover, the total amount of material in a specimen must be very small to prevent excessive heating and the specimen must also be extremely thin. Substrates for mounting specimens are typically of the order of 10 nm thick. Sections of solid materials should be no thicker than the order of 0.1 micrometer. Biological tissues are usually sectioned to thicknesses of about 0.05 μm to obtain high-resolution detail. Metal foils are usually thicker, ranging from 0.1 to 0.2 μm to obtain more information on their defect characteristics. As thickness is increased, multiple scattering of electrons by matter destroys the contrast in the image; finally, energy transferred from the beam by inelastic collisions becomes so large that the specimen may be destroyed.

High-voltage electron microscopes have been developed in which the electrons are accelerated to energies of 500 keV to over 1 MeV, giving them greater penetrating power. Some of the severe limitations indicated above are relaxed to a degree in these high-voltage instruments. Thicker sections can be examined; less damage occurs in the specimen, and small specimen wet chambers with hydrated biological materials can be utilized. Development of high-voltage electron microscopy began in the mid-1960s. The problems and complexities of high-voltage electron microscopes limit their application, and most investigations will continue to be done within the limitations set by the 50–100-kV instruments.

Biological applications. Techniques and applications of electron microscopy are so specialized that it is convenient to discuss them in separate categories of biological and nonbiological, even though there are basic techniques that overlap both fields such as substrate preparation and shadow-casting.

Substrates. The material to be examined in the electron microscope must, in general, be mounted on an extremely thin supporting membrane which has a mass density per unit area less than that of the object being observed. In practice a membrane of the order of 10 nm thick is mounted on a fine electrodeposited metal mesh with openings 50–100 μm on a side, and the specimen material is deposited on these mounted substrates. Cellulose nitrate films prepared by casting a film on water from a solution of cellulose nitrate in amyl acetate produces a suitable substrate.

Another substrate material is Formvar, a polyvinyl formal plastic, cast on glass from solutions of the plastic in dichloroethane (ethylene dichloride). More durable substrates can be prepared by evaporating carbon on a glass slide, stripping off the carbon film by floating it onto a water surface, and mounting the film on the standard mesh specimen screens. Carbon-film substrates as thin as 5 nm can be produced and handled in this fashion.

Shadow-casting. Materials of low mass density will not produce appreciably greater scattering than the substrate film upon which they must be mounted. Contrast in electron images from these specimens will therefore be low, and since this low differential of mass density applies especially to organic materials, it presents a large obstacle in imaging very small biological objects. The technique of shadow-casting developed by R. C. Williams and R. W. G. Wyckoff is a very powerful means of enhancing contrast in electron microscopy of small objects of low mass density. It is now possible to resolve globular macromolecular material with diameters as small as 2–3 nm and linear structures such as strands of nucleic acid less than 1 nm thick. Resolution in shadow-cast specimens is set by the granularity of the shadow-casting material.

The technique involves vacuum deposition of an extremely thin metal layer on the specimen at an oblique angle. Any projection above the surface receives a heavier deposit of the metal on the side facing the metal evaporating source and casts a shadow in which no metal is deposited (**Fig. 10**). The metal shadowing layer produces a topographical representation of the surface of the specimen and, as discussed below, is an essential step in all replication of surfaces for study by electron microscopy.

After shadow-casting, the metal layer is the specimen imaged by the electron microscope, for the scattering from the shadow-cast specimen is predominantly that of the deposited metal layer. The areas in the shadow where no metal is deposited will scatter relatively few electrons, and the respective areas of the image will appear dark on the photographic plate that has been exposed to the high electron density incident in these portions of the image. Projecting areas facing the evaporation source will have heavy layers of the metal and will scatter most electrons, thus appearing light on the photographic record of the image. An example of a shadow-cast specimen is shown in **Fig. 11**, in which the image is reproduced in the customary manner with the same relative light and dark as on the original electron micrograph negative. The effect is as though strong illumination were falling obliquely on the specimen.

The metals most generally used are uranium, platinum, or chromium. The procedures developed for producing thin films by metal evaporation are applied. A tungsten filament charged with the metal to be evaporated is heated in a vacuum of about 10^{-5} mmHg (10^{-8} atm or 10^{-3} Pa) to a temperature at which the metal has a vapor pressure of the order 10^{-2} mmHg (10^{-5} atm or 1 Pa), and a layer of mean thickness of about 0.5 nm is deposited on the specimen at the desired angle.

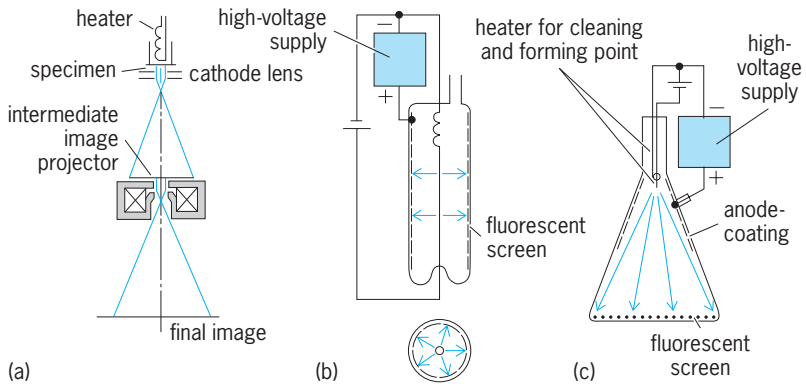

Fig. 9. Emission electron microscopes. (*a*) Immersion electron microscope. (*b*) Cylindrical thermionic-emission microscope. (*c*) Field-emission microscope.

Particulate materials. Examination of particulate biological materials such as bacteria, virus, or mitochondria of cells requires proper dispersion of a very small sample of the material on the substrate. Methods must be utilized that prevent aggregation and surface tension effects of drying.

Particles already dispersed and suspended in an aqueous medium can be deposited directly on plastic or carbon substrate films. The effects of surface tension on drying, the aggregation of the particles, and the effects of changing salt concentration and pH on drying must be considered. A method of minimizing these effects is to spray the particle suspension on the substrate in very small droplets of the order of 5–10 μm in diameter. If buffers are needed, volatile ones such as ammonium acetate, ammonium carbonate, or veronal acetate must be used. A further advantage of the spray technique is the rapid drying of the small droplets; this minimizes the effects of changing concentration on the biological materials. By depositing the droplets on substrates cooled to liquid nitrogen temperatures and subliming the ice

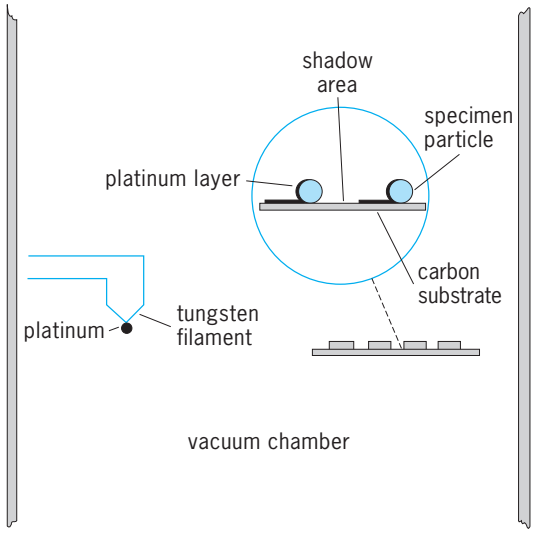

Fig. 10. Schematic of shadow-casting procedure.

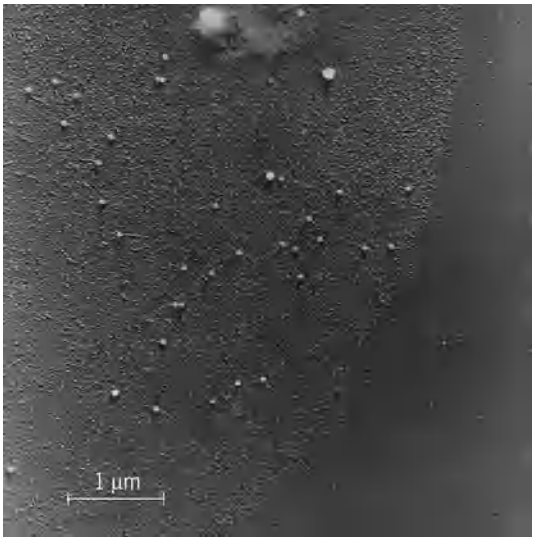

Fig. 11. Edge of spray droplet with bacteriophage.

under vacuum, the surface tension effects of drying can be further reduced.

The spray technique can be used for quantitative work, and assays can be made on suspensions of particles of unknown concentration in which a direct count is obtained on the number of particles per unit volume. Since each droplet is an aliquot sample of the total specimen, when the volume of the droplet is known, a count of the particles in a drop pattern gives the assay. By mixing the suspension to be assayed with a given volume of a known concentration of monodispersed polystyrene latex particles and counting the number of polystyrene in each drop, the volume of the drop is obtained. When a highly purified material is available for assay, it is possible to take an aliquot portion to dryness, to weigh it, and from the electron microscope particle count to calculate the molecular weight of the particle. **Figure 12** shows tobacco mosaic virus (rods) being assayed using the polystyrene spheres.

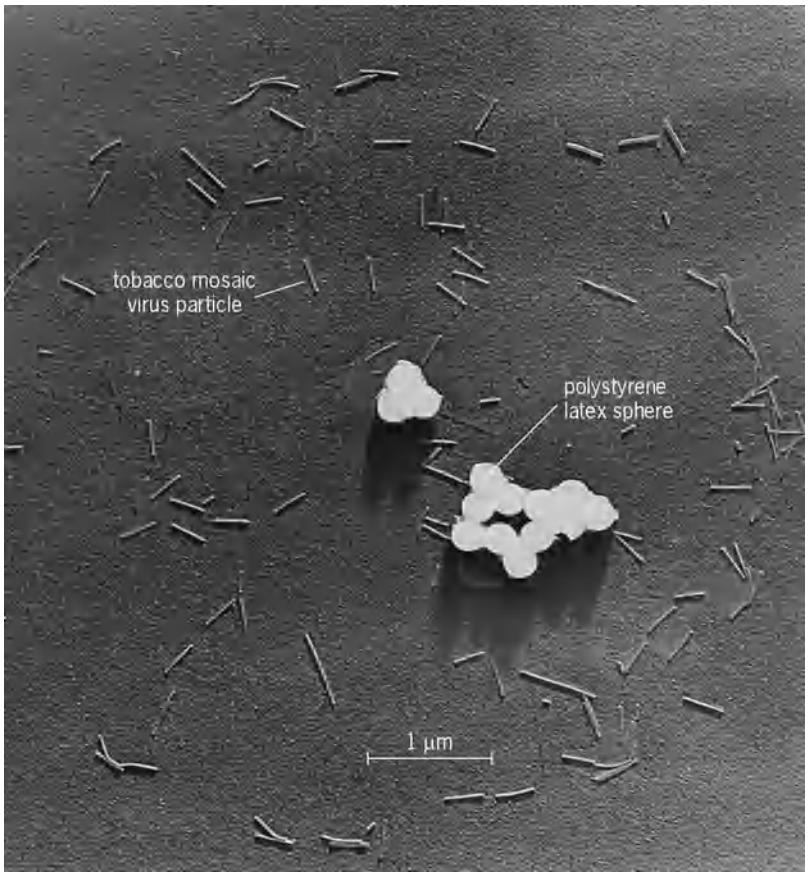

Fig. 12. Spray droplet pattern for assay of tobacco mosaic virus (rods) using polystyrene latex particles (spheres).

The electron microscope is capable of resolving macromolecules of biological interest. Spraying a purified sample of macromolecular material onto a very smooth surface such as freshly cleaved mica produces a suitably dispersed specimen. This specimen is shadowed directly on the mica with platinum, is backed with an evaporated substrate of carbon, and is stripped off the mica by using methods described below.

Negative staining is another important technique used to enhance the image contrast obtained at very high magnifications from biological specimens of a particulate nature such as subcellular components, bacterial surface structure, virus particles, and protein molecules. The particulate material is treated much as described above but with a soluble heavy metal salt added to the aqueous suspension medium. A typical negative staining procedure uses a 2% solution of potassium phosphotungstate buffered and adjusted to a neutral pH of 6.8 to 7.4. A droplet of the suspension either sprayed or deposited on the substrate film dries with an embedding matrix of the heavy metal salt surrounding it. Examination in the electron microscope produces an image of reverse contrast. The surrounding matrix strongly scatters the electron beam, producing a dark ground, while particles such as viruses which should not react with the stain appear light. If the virus or other particle under examination has substructure, the heavy metal matrix can penetrate the interstices and delineate the subunits of the particle (**Fig. 13**).

Tissues and cells. Successful development of techniques for sectioning biological tissues for examination at high resolution provided a means of directly observing the complex ultrastructure of organisms. The understanding obtained of the organization of the cell and the structure of its organelles profoundly changed the fields of histology and cytology and is the most significant contribution of electron microscopy (**Figs. 14** and **15**). *See* CELL (BIOLOGY).

The first successful thin sectioning using a modified conventional microtome was reported in 1948 by D. C. Pease and R. F. Baker. After that techniques developed rapidly and the methods for study of the ultrastructure of cells and tissues became highly developed. Microtomes, using fractured glass or polished diamonds for knives, cut serial sections for reproduction down to thicknesses as low as a few hundred angstroms, or in special cases as thin as 5 nm. Techniques for fixation, dehydration, embedding,

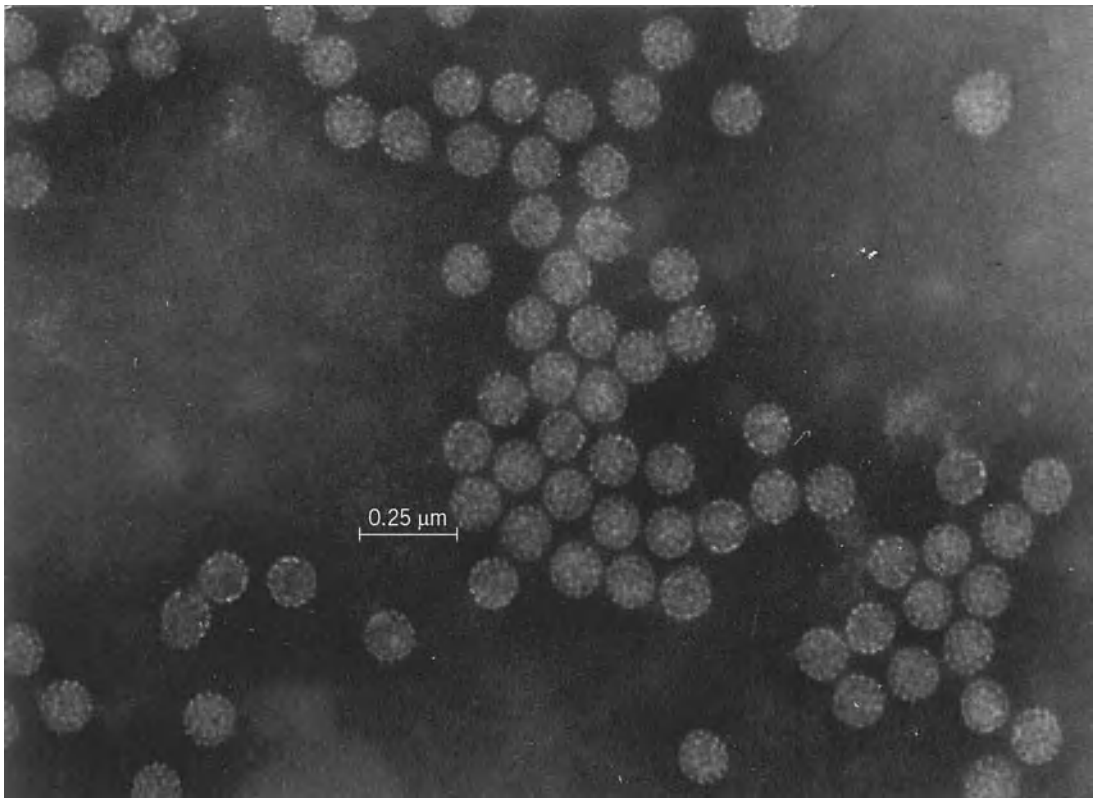

Fig. 13. Purified polyoma virus, negative stain 2% phosphotungstic acid buffered to pH 6. (*Courtesy of S. S. Breese, Jr.*)

and staining of tissues were developed which are adequate for a large range of observations on the complex detail of cell structure. The widely used fixatives are based on osmium tetroxide, potassium permanganate, or the aldehydes properly buffered to control pH. Embedding media were extended to a range of polymeric materials, the most successful being the epoxy resins and polyester resins. Staining of tissues either before or after sectioning can be very helpful to enhance contrast in very thin tissue sections. The general principle of staining for electron microscopy is to introduce heavy atoms such as lead or uranium into the tissue to make it more effective in scattering the electron illumination and thus to increase the visibility of those parts of the tissue to which the heavy atoms are attached. The specificity of stains available is limited, and the electron microscopist does not have available to him the range of specific staining materials with easily identifiable color contrasts that play such important roles in histological work with the light microscope. Among staining materials widely used are such heavy metal compounds as phosphotungstic acid, uranyl acetate, and lead hydroxide as well as the fixative osmium tetroxide.

There are special techniques for localization of specific entities in a cell. Immunochemical staining was developed by adapting the fluorescent antibody technique of light microscopy. Ferritin-labeled antibodies have been prepared that react with the homologous antigen and thus can be located in the cell by the increased contrast associated with this entity in the electron microscope image of the sectioned tissue. *See* FLUORESCENCE MICROSCOPE; IMMUNOCHEMISTRY.

Another technique adapted from light microscopy and applied with increased resolution is autoradiography. A tissue is labeled with a radioactive tracer material by injecting the animal or the nutrient medium before the tissue is taken. Tritium incorporated in thymidine or similar constituents is the most widely used label. The thin section containing the radioactive element is covered with a very thin layer of photographic emulsion and allowed to stand in the dark for weeks or months. Ionizing radiation exposes the photographic emulsion producing a latent image in the silver halide grains of the emulsion at those sites at which radioactive decay has occurred. After development and fixation of the photographic emulsion and staining of the tissue section, the composite section and developed photographic emulsion is examined in the electron microscope. The developed silver grains locate the sites at which the label was attached in the section with resolutions of better than 100 nm (Fig. 15). *See* AUTORADIOGRAPHY.

Great depth of focus, a characteristic of electron microscopes in general and of the scanning microscope in particular, gives the scanning electron microscope a decided advantage over the light microscope for topographical observations on biological structures (**Fig. 16**).

Nonbiological materials. The most severe limitation in observation of nonbiological materials is their bulk nature and the fact that only very small and extremely thin specimens can be examined. Since solid material often loses its bulk properties when reduced to

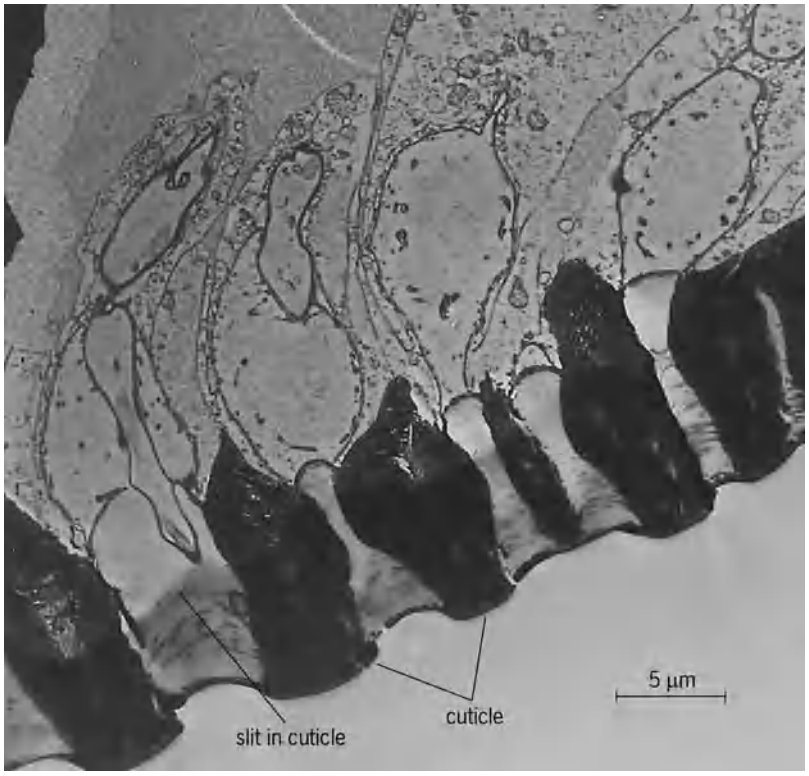

Fig. 14. Electron micrograph of longitudinal section of a spider's leg just above tarsal-metatarsal joint cut through the vibration receptor (ear). (*Courtesy of M. Salpeter*)

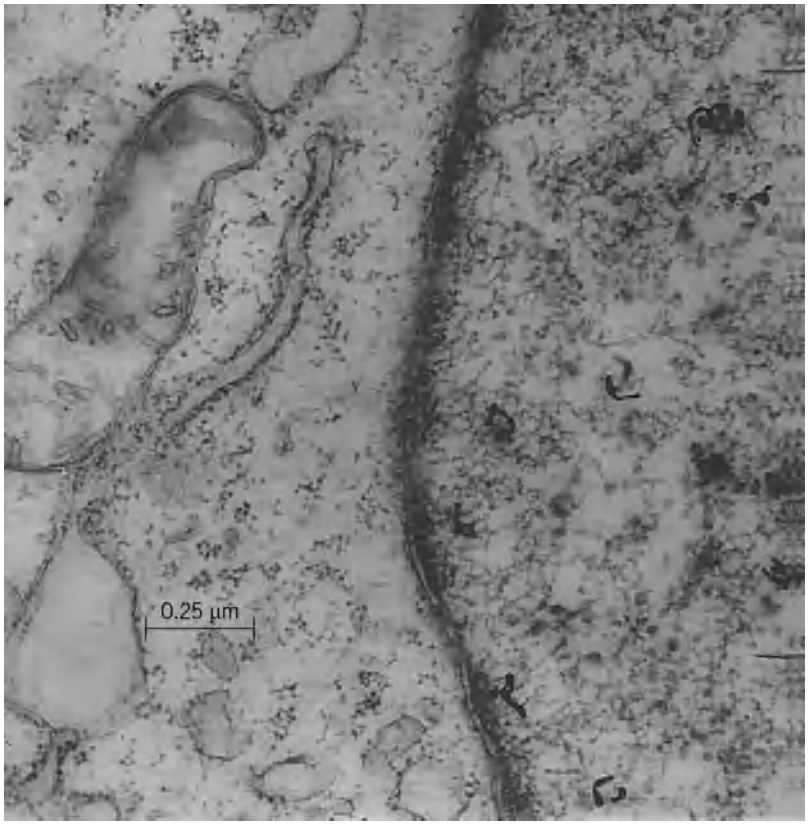

Fig. 15. Electron microscope autoradiogram of thin section of a mesenchymatous cell from the regenerating limb of adult newt (*Triturus*) labeled with tritiated thymidine. The dense black bodies over the nucleus on the right are developed silver grains, indicating radioactive decay in the tissue. (*Courtesy of M. Salpeter*)

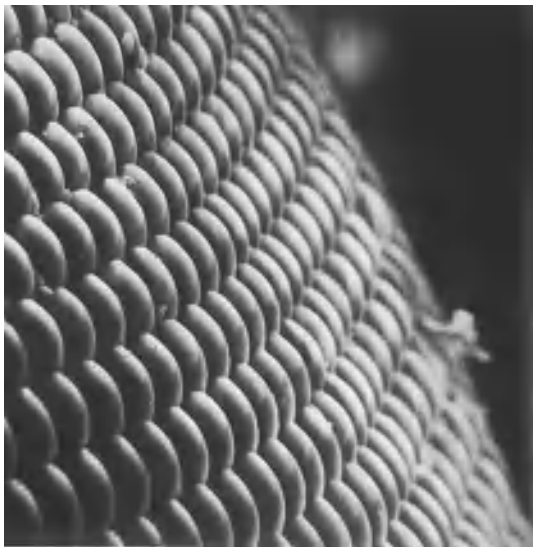

Fig. 16. Scanning electron micrograph of the surface of a fly's eye. (*RCA Research Labs, Princeton, New Jersey*)

very thin films or is too finely dispersed, high-voltage electron microscopes are used most actively for studies on nonbiological materials.

Particulate materials. Proper dispersion of the sample is the main consideration. If the material to be examined is a dry powder, a dispersion technique must be employed. One method used extensively is to mill the powder in a viscose solution of cellulose nitrate in amyl acetate until the aggregates are properly dispersed. The suspension is then diluted with amyl acetate to appropriate concentration for casting a suitably thin film with the particles suspended in it. Shadow-casting is usually employed to enhance the contrast or to give information on particle heights or shapes from the characteristics of their shadows.

Surfaces by replication. Surfaces of materials too thick to be placed directly in the electron microscope can be examined in the transmission electron microscope by preparing a thin film replica of the surface. The shadow-casting technique which produces contrast effects characteristic of the topography of the surface on which it is deposited greatly enhances the results obtainable by replication (**Fig. 17**).

A number of variations in method are employed in producing suitable replicas of surfaces. Each type of surface requires the proper choice from a variety of techniques and materials developed for application to different surfaces. The materials used for thin-film replicas include cellulose nitrate, Formvar, and evaporated carbon films. Carbon replicas are preferred for their greater strength and stability. A relatively smooth surface may be replicated directly by coating a film of Formvar on it by flooding the surface with a 0.5% solution of Formvar in ethylene dichloride. For greater ease of stripping, the replica may be backed by coating a cellulose nitrate film on the Formvar film. The backed replica can be wet-stripped, either by allowing the plastic film to float off on water if it will easily separate from

Fig. 17. Replica of a fractured surface of a magnesia ceramic. (*Courtesy of T. T. Sheng*)

the surface being replicated or the material being replicated can be dissolved away with a suitable solvent or acid to release replicating film. The replica film is mounted on specimen grids with the replication surface up, shadowed, and the backing film, if present, dissolved away in a solvent that does not attack the replica film. In the case of a Formvar replica backed with cellulose nitrate, amyl acetate is satisfactory. Dry stripping is accomplished by sandwiching the specimen grids between the film and Scotch tape and pulling off the film. Sometimes a stripping layer of a hydrophilic material is coated on the surface to be replicated. A detergent material such as Victawet can be evaporated onto the surface to be replicated, the metal shadowing deposited directly over the Victawet coat in the same vacuum chamber and, to complete the replica, a carbon substrate deposited by evaporation. This pre-shadowed replica is perhaps the most satisfactory and is used if the material to be replicated is suitable. On very rough surfaces, such as wood and fractured materials, it may be necessary to form a thick negative replica of a material such as polystyrene which can be molded to the surface by heat and pressure. This can then be replicated by shadowing, evaporating a carbon film on it, and releasing the replica by dissolving the polystyrene away with a suitable solvent. Materials containing water have been successfully replicated by using polyvinyl alcohol as the material for the negative replica. Purified crystalline virus in aqueous suspension and similar materials can be replicated by a rapid freezing technique. A frozen suspension is fractured to expose a fresh surface, placed in a vacuum chamber, and warmed to -110°F (-80°C) until the ice matrix is etched away slightly, leaving the material to be replicated in relief. This is cooled again to liquid nitrogen temperatures, the surface is replicated by shadowing, and a carbon substrate is evaporated on.

Crystalline films. The electron microscope finds wide application in investigations of crystalline materials. Two techniques are used to obtain specimens of bulk materials thin enough for direct transmission microscopy. Bulk samples can be reduced to a foil of 0.004–0.008 in. (0.1–0.2 mm) thick by processes such as jet machining, spark machining, or cutting and grinding. A final thinning is carried out by chemical or electrochemical methods which process the specimen to thicknesses of 0.1–0.2 μm for direct observation in the electron microscope. The other method utilizes various methods of deposition to grow films of suitable thicknesses and characteristics (**Fig. 18**).

Crystalline films obtained by these methods will strongly diffract the electron beam, and image contrast is determined by the nature of these diffraction interactions. When any part of the crystalline specimen is oriented so that a strong Bragg reflection occurs, the diffracted beam is scattered outside of the

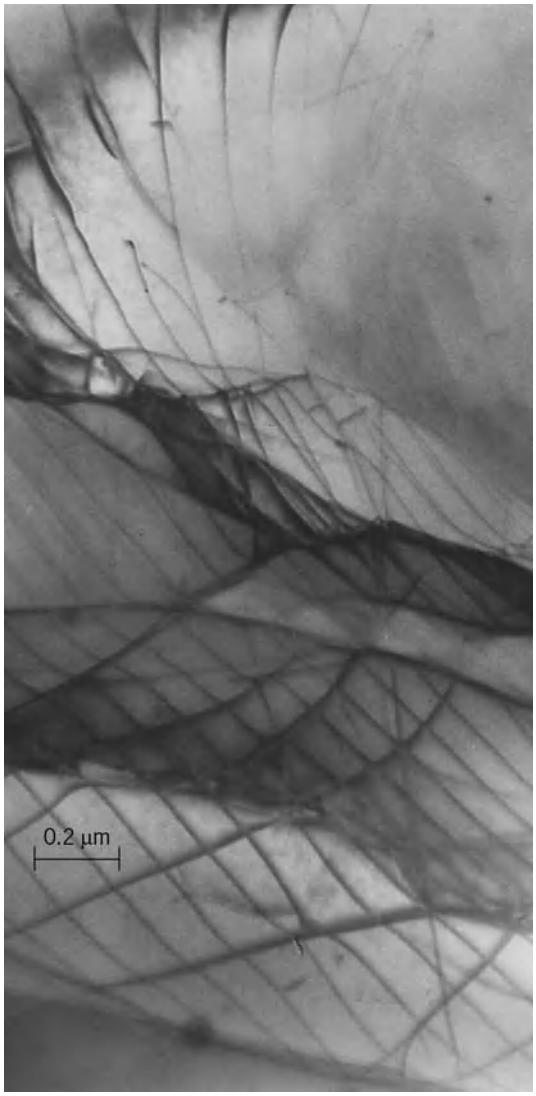

Fig. 18. Thin magnesia foil showing dislocation networks. (*Courtesy of T. T. Sheng*)

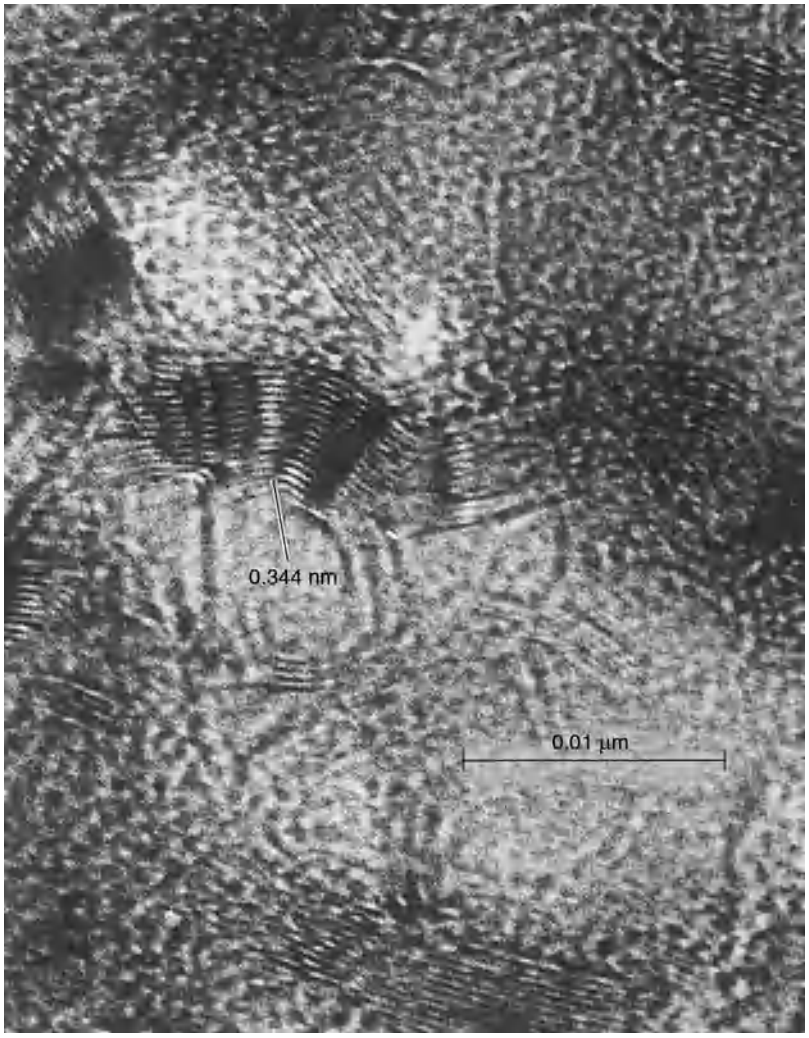

Fig. 19. Graphitized carbon showing lattice resolution of the 0.344-nm lattice spacing.

objective aperture and is no longer available in the imaging process. The image of this specimen then appears dark in areas corresponding to regions where Bragg reflection has occurred, so that high contrast is produced in the electron micrograph. Thus the crystal structure and its imperfections produce contrast effects in the imaging process that make this application of electron microscopy an important tool for study of defect structure in solids and the structure of thin films.

The electron optics of the electron microscope can be adjusted by a simple change of the focal length of one lens to produce the electron diffraction pattern from the same small area of specimen under observation with the normal electron optics for microscopy. *See* ELECTRON DIFFRACTION.

When the aperture stop of the objective lens is adjusted to allow one or two of the diffracted beams to contribute to the image formation in addition to the undeviated electron beam, a lattice resolution image is obtained that resolves lattice planes in crystalline material. Very thin deposited films or crystals which grow as thin platelets are required as specimens for these extremely high-resolution observations. By careful adjustment two or more lattice planes can be very highly resolved to spacings of less than 0.2 nm. If there are imperfections in the lattice, these defects will also be observed under certain orientations of the crystal. Graphitized carbon can also produce lattice resolution images as shown in **Fig. 19**.

Limitations on observations. Specimen artifacts, instrumental defects, and image characteristics are factors in the limitations on observations.

Specimen artifacts. Artifacts introduced in biological specimen preparation are primarily caused by the dehydration process necessary to bring the material under the vacuum of the electron microscope. Where effects such as distortion from surface tension and chemical changes from changing ionic concentration are severe, observations must be supplemented with other methods and careful evaluation made of the extent to which preparation artifacts may have changed a living, dynamic system.

The specimen will also be altered when exposed to the electron beam. Careful procedures will eliminate changes in the micromorphology of the material. Irradiation of a specimen with a low-intensity beam will cause chemical changes in the material, so that, for example, cellulose nitrate film becomes insoluble in all solvents that normally dissolve it. Also, the polymerized butyl methacrylate matrix used to embed tissue sublimes out when thin sections are exposed to low-intensity beams, leaving the cytological material and enhancing the contrast observed in the electron micrograph. On the other hand, exposure to high-intensity electron beams can produce a rise in temperature of hundreds of degrees Celsius in the specimen material. If a thin section with a polymeric embedding medium is subjected directly to a high-intensity beam, the polymer will melt, causing considerable distortion in the section. Thermal effects can cause considerable change in some materials, and care must be exercised in avoiding irradiation with too-high-intensity beams. Examination at high electron optical magnifications ($\times 1{,}000{,}000$) requires very intense beams, and special illuminating systems such as those employing a double condenser are utilized to minimize heating effects.

Contamination of the specimen also occurs under exposure to the electron beam. A layer of low-density, inert, carbonaceous material is laid down, sometimes at rates as high as several angstroms per second. The rate of deposition is a function of beam intensity, specimen temperature, and the amount of hydrocarbons present from pump oils and gasket grease. By surrounding the specimen with a chamber kept at -110°F (-80°C) the deposition can be reduced to a minimum. Special chambers designed to achieve this effect are essential for all high-resolution work where detail and contrast would soon be obliterated by deposition of a contamination layer.

Instrumental defects. High-resolution electron microscopes can be adjusted to attain an instrumental resolving power of better than 0.5 nm. Resolution obtained is usually limited by the resolution inherent in the specimen, especially for biological materials. Achievement of such a high level of instrumental performance requires a high degree of skill in the

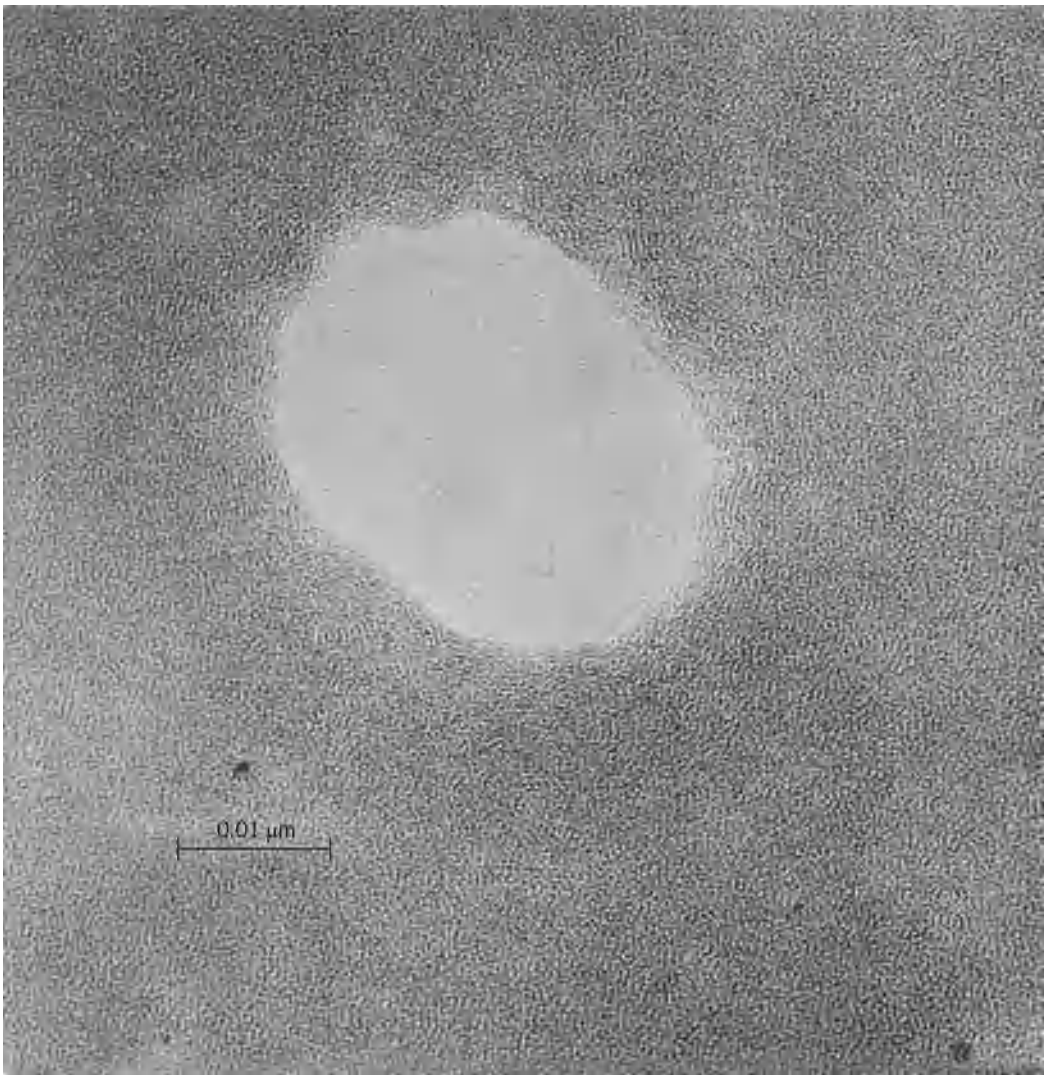

Fig. 20. Evaporated carbon film with small hole. Very slight underfocus shows Fresnel fringe inside hole. Phase contrast produces structural detail in film.

adjustment by the electron microscopist. The effect which first limits the resolving power even in a well-adjusted instrument is a deviation from axial symmetry in the objective lens producing an astigmatic image. Correction is made by introducing a compensating cylindrical field at right angles to the existing one, either by shims of ferromagnetic material oriented at the correct azimuth in the gap of the pole piece or by an electrostatic or magnetic field of correct orientation and magnitude. Externally adjustable stigmators, of both magnetic and electrostatic types, are available which enable the operator to compensate the electron microscope and obtain resolutions of 0.5 nm or better.

Contrast in the electron image can be enhanced by using a stop in the objective lens limiting the angular apertures that are necessary when relatively thick specimens such as sections and replicas are examined. However, an objective aperture can cause deterioration of resolving power, especially if asymmetric contamination forms on the aperture and causes increased astigmatism in the image. Good practice requires careful checking on instrumental performance using test specimens of a noncrystalline material such as carbon film with small holes to check asymmetry or astigmatism and the phase contrast in the out-of-focus image.

Image characteristics. One of the most prominent characteristics of the electron microscope is the contour effect. An instrument with a good, small effective source of illumination produces an image where Fresnel diffraction is clearly observable when the objective lens is out of true focus. In practice only the first-order fringe is strongly in evidence outlining any sharp boundaries in the image; hence the so-called contour effect. When the objective-lens current is too low, so that the focal length of the lens is too long, the image is said to be underfocused. At an edge where there is a sharp change in contrast, the outside of the edge is bordered by a dark contour. Where focal length is too short, the image is said to be overfocused and the contour now appears inside the edge. At true focus the Fresnel fringe vanishes and a soft image is obtained. Most electron

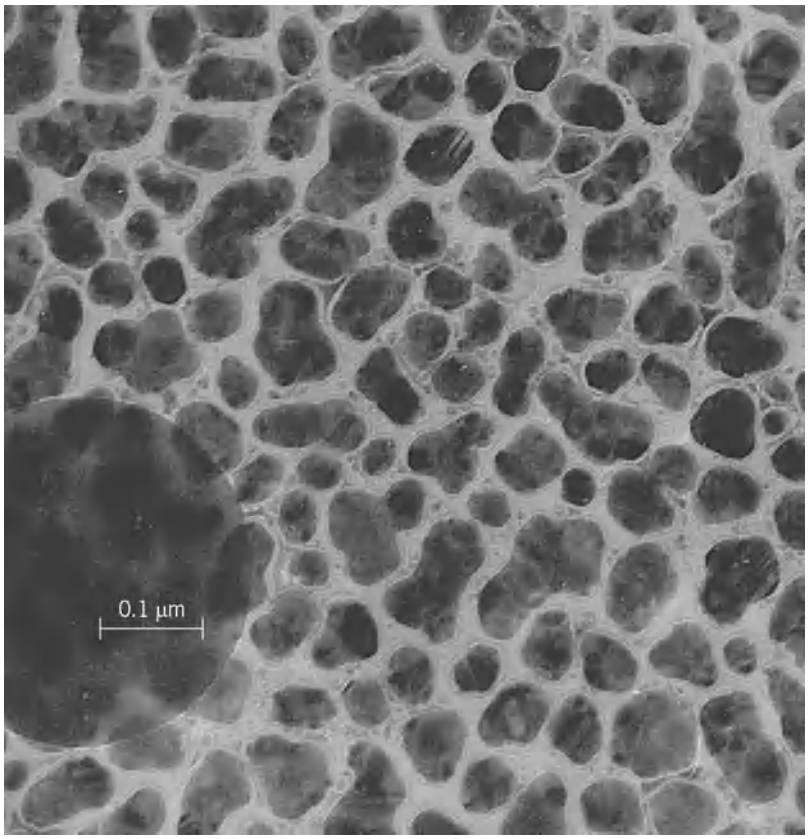

Fig. 21. Transmission electron micrograph of evaporated copper deposited on mica and transferred to a carbon substrate. Contrast effects are caused by different orientations and lattice imperfections.

micrographs are taken at a slightly underfocus setting of the objective lens, but for critical high-resolution work a focal series should be taken. Again, a carbon film with holes is a good test specimen (**Fig. 20**).

Crystalline material can produce anomalous contrast effects. If a crystal is oriented so that the incident electron beam is at an angle satisfying the Bragg condition for diffraction, the crystal can scatter the total incident beam into the Bragg angle while the same crystal when not oriented at the Bragg angle will scatter a fraction of the incident beam, depending on its thickness and its density.

The images obtained under the two different conditions of orientation can be of very different photographic density, and care must be taken in interpreting contrast effects of images of crystalline materials. **Figure 21** shows the strong contrast effects which can be produced by the electron interference interactions when scattered from a crystal lattice. The particles are of approximately uniform thickness, and the observed contrast effects are caused by different orientations and lattice imperfections.

The extension of the working resolving power of the electron microscope shows the significant role that diffraction phenomena play in the imaging characteristics of the instrument. Proper understanding of image contrast can be achieved only with a wave optical treatment of the imaging process. At levels of resolution of less than 0.5 nm, phase contrast plays a most important role. Phase contrast occurs with a defocusing of the image to adjust for the phase shift of the scattered electron wave by the spherical aberration of the objective lens. Phase contrast produces the contrast observed in the image of the carbon film shown in Fig. 19. Crystalline structures such as thin graphite layers show strong phase-contrast effects. The lattice resolution method is essentially a phase-contrast image. Biological macromolecular detail as small as 0.4–0.5 nm can be resolved by phase-contrast imaging when the specimen is suspended on the edge of a fiber support. Benjamin M. Siegel

Bibliography. P. B. Bell, Jr., *Scanning Electron Microscopy of Cells in Culture*, 1984; H. Bethge and J. Heydenreich (eds.), *Electron Microscopy in Solid State Physics*, 1987; R. Dickersin, *Diagnostic Electron Microscopy: A Text-Atlas*, 2d ed., 2000; V. A. Drits, *Electron Diffraction and High-Resolution Electron Microscopy of Mineral Structures*, 1987; H. Fujita (ed.), *History of Electron Microscopy*, 1987; P. Goodhew and F. Humphreys (eds.), *Electron Microscopy and Analysis*, 3d ed., 2000.

Electron motion in vacuum

Motion of electrons in a space freed sufficiently from matter so that collisions with other particles play a negligible role. The motion of electrons in vacuum is controlled by electric and magnetic fields whose force on the electrons is proportional to their magnitude. Electric and magnetic fields may arise from the presence of electrodes, currents, and magnets surrounding the evacuated space in which a particular electron moves, as well as from the presence of other charged particles within this space. This article deals with the nonrelativistic motion of electrons in static electric and magnetic fields, the effect of space charge on the electron paths, and motion in the time-varying fields that are encountered, for example, in the cathode-ray oscilloscope. For a discussion of the motion of electrons and other charged particles in cases where the velocities approach that of light *see* RELATIVISTIC ELECTRODYNAMICS

Static electric fields. An electron moving in a plane of symmetry of an electric field will remain indefinitely in that plane because the electrical forces acting on the electron lie within the plane. A plane of symmetry is here defined as one for which the mirror image of the potential distribution in front of the plane coincides with the potential distribution in back of the plane. Newton's second law of motion for the electron moving in the plane with rectangular coordinates x, y takes the form of Eqs. (1) and (2). Here $-e$ and m are the charge and mass of the

$$m\frac{d^2x}{dt^2} = -eE_x = e\frac{\partial\phi}{\partial x} \tag{1}$$

$$m\frac{d^2y}{dt^2} = -eE_y = e\frac{\partial\phi}{\partial y} \tag{2}$$

electron, t is time, E_x and E_y are the x and y components of the electric field, and ϕ is the electric

potential, normalized as stated by Eq. (3). Newton's

$$\frac{m}{2}\left[\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2\right]=e\phi \tag{3}$$

law as stated in Eqs. (1) and (2) implies that the speed of the electron is small enough so that its mass can be regarded as constant. For an electron having an energy of 10 kilovolts, the increase in mass is about 1%.

Path equation. Elimination of time from Eqs. (1) and (2) leads to the path equation (4). If potential distri-

$$\frac{d^2y}{dx^2}=\frac{1}{2\phi}\left[1+\left(\frac{dy}{dx}\right)^2\right]\left(\frac{\partial\phi}{\partial y}-\frac{dy}{dx}\frac{\partial\phi}{\partial x}\right) \tag{4}$$

bution $\phi(x,y)$ is known, and if position and velocity of the electron at one point within the field are also known, the electron's path can be determined by integrating Eq. (4). For simple electrode structures, the potential distribution can be determined analytically by solving Laplace's equation in the form of Eq. (5). More generally, it can be found by construct-

$$\nabla^2\phi=\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2}=0 \tag{5}$$

ing a large-scale model of the electrode structure and immersing it in an electrolytic tank so that the surface of the liquid (usually slightly acidified tap water) coincides with the plane of symmetry of interest. With potentials proportional to the actual potentials applied to the model electrodes, an equipotential line on the surface can be found by determining the points at which a probe at the potential in question draws no current.

The path equation (4) leads to Eq. (6) for the ra-

$$R=\frac{2\phi}{\partial\phi/\partial n} \tag{6}$$

dius of curvature R of the paths. Here $-\partial\phi/\partial n$ is the component of the electric field normal to the electron path. If an equipotential plot has been prepared, this relation permits graphical plotting of an electron path (**Fig. 1**). The path is approximated by a series of circular arcs, the radius of curvature between successive equipotential lines being computed from the preceding relation for R.

Paraxial-ray equation. Electrostatic fields having not only a plane of symmetry but an axis of symmetry, which represents the intersection of an infinite family of planes of symmetry, have particular practical importance. Equation (4) still applies here, provided that y is identified with r, the distance from the axis, and x with z, the distance measured along the axis. The Laplace equation in the new coordinates z, r takes the form of Eq. (7). Equation (7) is solved quite

$$\frac{\partial^2\phi}{\partial r^2}+\frac{1}{r}\frac{\partial\phi}{\partial r}+\frac{\partial^2\phi}{\partial z^2}=0 \tag{7}$$

generally by the series shown as Eq. (8), where $\Phi(z)$

$$\phi(z,r)=\Phi(z)-\frac{r^2}{4}\frac{d^2\Phi}{dz^2}+\frac{r^4}{64}\frac{d^4\Phi}{dz^4}\cdots \tag{8}$$

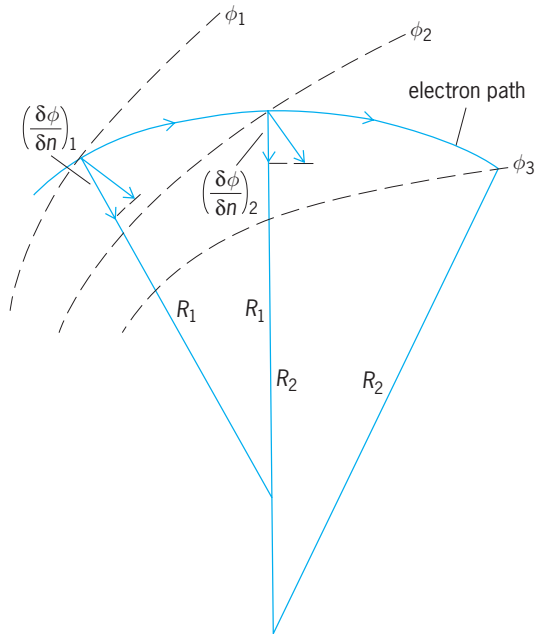

Fig. 1. Path plotting in an electrostatic field. Broken lines ϕ_1, ϕ_2, and ϕ_3 are equipotential lines. R_1 and R_2 are radii or curvature of the electron path.

is the potential on the axis of symmetry. Thus, the potential everywhere within the axially symmetric electrode structure is fully determined by the potential variation along the axis. Substitution of ϕ from Eq. (8) into Eq. (4) with retention of terms of the first order only in r and dr/dt leads to the paraxial ray equation (9). Equation (9) applies to electrons

$$\frac{d^2r}{dz^2}+\frac{1}{2}\frac{d\Phi}{dz}\frac{dr}{dz}+\frac{1}{4}\frac{d^2\Phi}{dz^2}r=0 \tag{9}$$

whose paths depart relatively little, both in slope and in distance, from the axis of the field.

Equation (9) is linear in r. Thus, if one electron path intersects the axis in two points, all electron paths passing through one of the points also pass through the other. In brief, the electric field images the one point into the other. It can be shown further that this imaging property is not limited to the axis but applies to extended areas about the axis, so that the axially symmetric electric field acts on the paths of electrons in the same manner as glass lenses act on light rays. Departures of the exact path from the paraxial equation result in image defects or

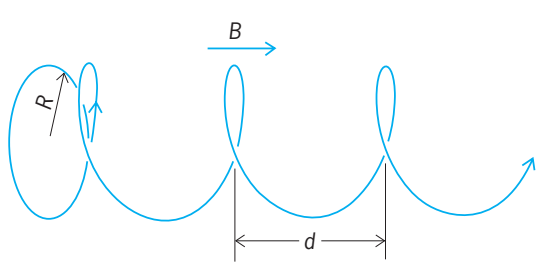

Fig. 2. Motion of an electron in a uniform magnetic field. Its path is in general a helix with pitch d, radius R, and axis parallel to the field.

aberrations similar in character to those observed for glass lenses. *See* ELECTROSTATIC LENS.

Magnetic fields. A magnetic field exerts on an electron of velocity **v** a force **F** perpendicular to both the direction of motion and the direction of the field. In vector notation this Lorentz force is given by Eq. (10). Here, **b** is the magnetic induction.

$$\mathbf{F} = -e(\mathbf{v} \times \mathbf{b}) \tag{10}$$

The components of the Lorentz force are given by Eqs. (11)–(13). Because this force is perpendicular

$$F_x = e\left(b_y \frac{dz}{dt} - b_z \frac{dy}{dt}\right) \tag{11}$$

$$F_y = e\left(b_z \frac{dx}{dt} - b_x \frac{dz}{dt}\right) \tag{12}$$

$$F_z = e\left(b_x \frac{dy}{dt} - b_y \frac{dx}{dt}\right) \tag{13}$$

to the direction of motion, it does no work on the electron, whose velocity consequently remains unchanged in magnitude. A uniform magnetic field parallel to the z axis is described by Eqs. (14). Newton's

$$b_z = B \qquad b_x = b_y = 0 \tag{14}$$

second law leads to a constant z component of the velocity. The magnitude of the velocity component v_{xy} in the xy plane is similarly constant, the square of the total velocity being equal to the sum of the squares of the components, For the motion projected on the xy plane, Newton's second law thus takes the form of Eq. (15). Here R is the radius of curvature of the

$$\frac{mv_{xy}^2}{R} = ev_{xy}B \tag{15}$$

projected path. R is seen to be a constant, so that the projected path is a circle with radius given by Eq. (16). Here α is the angle which the electron path

$$R = \frac{mv_{xy}}{eB} = \frac{\sin\alpha}{B}\left(\frac{2m\phi}{e}\right)^{1/2} = \frac{3.37\phi^{1/2}}{B}\sin\alpha \tag{16}$$

makes with the field direction, and ϕ is the accelerating potential of the electrons. If B is measured in gauss, and ϕ in volts, R is in centimeters. The frequency with which the circle is traversed by the electron is given by Eq. (17). This frequency, in s^{-1},

$$f = \frac{v_{xy}}{2\pi R} = \frac{eB}{2\pi m} = 2.8 \times 10^6 B \tag{17}$$

the cyclotron frequency, thus depends only on the magnetic field strength.

The complete motion of the electron (**Fig. 2**) is thus a helix about a magnetic line of force, with a pitch d in centimeters given by Eq. (18). All electrons passing through a point with equal axial velocity components pass through a series of points separated by d on the same magnetic field line. An initially divergent electron beam is held together by a uniform magnetic field, because an electron path which intersects a particular field line can never depart from it by more than twice the radius R of the helix. Uniform magnetic fields are widely used for keeping electron beams from spreading, for example, in traveling-wave tubes and klystrons. Electron beams will also follow magnetic field lines, if these are gently curved. This property is utilized in the magnetic deflection of beams in certain television camera tubes, such as the image orthicon and the vidicon. In these tubes, a weak transverse

$$d = \frac{v_z}{f} = \frac{2\pi\cos\alpha}{B}\left(\frac{2m\phi}{e}\right)^{1/2} = 21.08\,\frac{\phi^{1/2}}{B}\cos\alpha \tag{18}$$

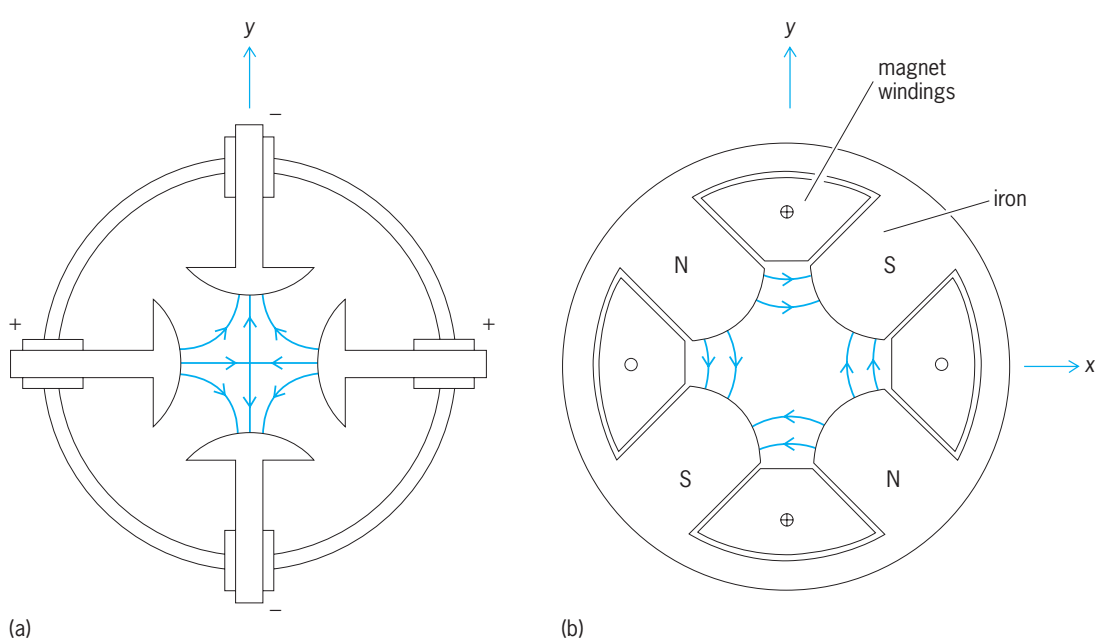

Fig. 3. Two types of quadrupole lens. (*a*) Electrostatic quadrupole. (*b*) Magnetic quadrupole.

magnetic deflection field is superposed on a strong longitudinal magnetic focusing field. *See* TELEVISION CAMERA TUBE; VAN ALLEN RADIATION.

Motion in nonuniform magnetic fields with axial symmetry is conveniently treated as a special case of motion in combined electric and magnetic fields.

Combined fields with axial symmetry. Motion is now expressed by Eqs. (19)–(21). Coordinates z, r, and θ

$$m\frac{d^2z}{dt^2} = e\left(\frac{\partial\phi}{\partial z} + b_r r\frac{d\theta}{dt}\right) \tag{19}$$

$$m\left[\frac{d^2r}{dt^2} - r\left(\frac{d\theta}{dt}\right)^2\right] = e\left(\frac{\partial\phi}{\partial r} - b_z r\frac{d\theta}{dt}\right) \tag{20}$$

$$m\frac{1}{r}\frac{d}{dt}\left(r^2\frac{d\theta}{dt}\right) = e\left(b_z\frac{dr}{dt} - b_r\frac{dz}{dt}\right) \tag{21}$$

represent distance along the axis, perpendicular distance from the axis, and azimuthal angle about the axis. Terms b_z and b_r are the axial and radial components of the magnetic induction. From Eqs. (19)–(21), Eq. (22*a*), a path equation expressing the variation of the radial distance r with the axial distance z, is derived, with ϕ^* defined by Eq. (22*b*). Here ϕD^2 is a shorthand symbol for the last term in Eq. (22*b*) and C is, except for a universal multiplying constant, the angular momentum of the electron about the axis at a point where the magnetic field vanishes, given by Eq. (22*c*). Here A is the magnetic vector potential, which is numerically equal to the magnetic flux through a circle about the axis through the reference point divided by the circumference of that circle.

$$\frac{d^2r}{dz^2} = \frac{1}{2\phi^*}\left[1 + \left(\frac{dr}{dz}\right)^2\right]\left(\frac{\partial\phi^*}{\partial r} - \frac{dr}{dz}\frac{\partial\phi^*}{\partial z}\right) \tag{22a}$$

$$\begin{aligned}\phi^* &= \phi\,(1 - D^2)\\ &= \phi - \left[\frac{C}{r} + \left(\frac{e}{2m}\right)^{1/2} A\right]^2\end{aligned} \tag{22b}$$

$$C = r^2\frac{d\theta}{dz}\phi^{1/2}\left[\left(\frac{dr}{dz}\right)^2 + r^2\left(\frac{d\theta}{dz}\right)^2 + 1\right]^{-1/2} - \left(\frac{e}{2M}\right)^{1/2} rA \tag{22c}$$

At the same time, the aximuth θ of the electron changes according to Eq. (23). Equations (22) can be solved by graphical and numerical methods useful for determining electron paths in electrostatic fields. The general paraxial equation is obtained by substituting the expansion of Eq. (24) into Eq. (22*b*).

$$\theta = \phi_0 + \int_{z0}^{z} \frac{D}{r(1 - D^2)^{1/2}} \times \left[1 + \left(\frac{dr}{dz}\right)^2\right]^{1/2} dz \tag{23}$$

$$A = \frac{r}{2}B(z) - \frac{r^3}{16}\frac{d^2B(z)}{dz^2}\cdots \tag{24}$$

Here $B(z)$ is the magnetic induction along the axis. Substitution of Eq. (24) and that for the electrostatic potential and retention of terms of the first order in r and $drPdz$ only lead to Eqs. (25). With B in gauss, ϕ in volts, and z in centimeters, $e/8m$ equals 0.022 volt/(gauss-cm)^2.

$$\frac{d^2r}{dz^2} + \frac{1}{2\Phi}\frac{d\Phi}{dz}\frac{dr}{dz} + \left(\frac{1}{4\Phi}\frac{d^2\Phi}{dz^2} + \frac{eB^2}{8m\Phi} - \frac{C^2}{\Phi r^4}\right)r = 0 \tag{25a}$$

$$\theta = \theta_0 + \int_{z0}^{z}\left[\frac{C}{r^2\Phi^{1/2}} + \left(\frac{e}{8m\Phi}\right)^{1/2} B\right]dz \tag{25b}$$

$$\frac{C}{r_0{}^2\Phi_0{}^{1/2}} = \left(\frac{d\theta}{dz}\right)_0 - \left(\frac{e}{8m\Phi_0}\right)^{1/2} B(z_0) \tag{25c}$$

Quadrupole fields. The principal use of axially symmetric electrostatic and magnetic fields is to converge electron pencils; thus, in the cathode-ray tube, electrons diverging from a point in front of the emitter, called the crossover, are converged to a small spot on the viewing screen. Similarly, in an electron-imaging device such as the electron microscope, electrons diverging from a point on an object are brought to focus at a corresponding point of the

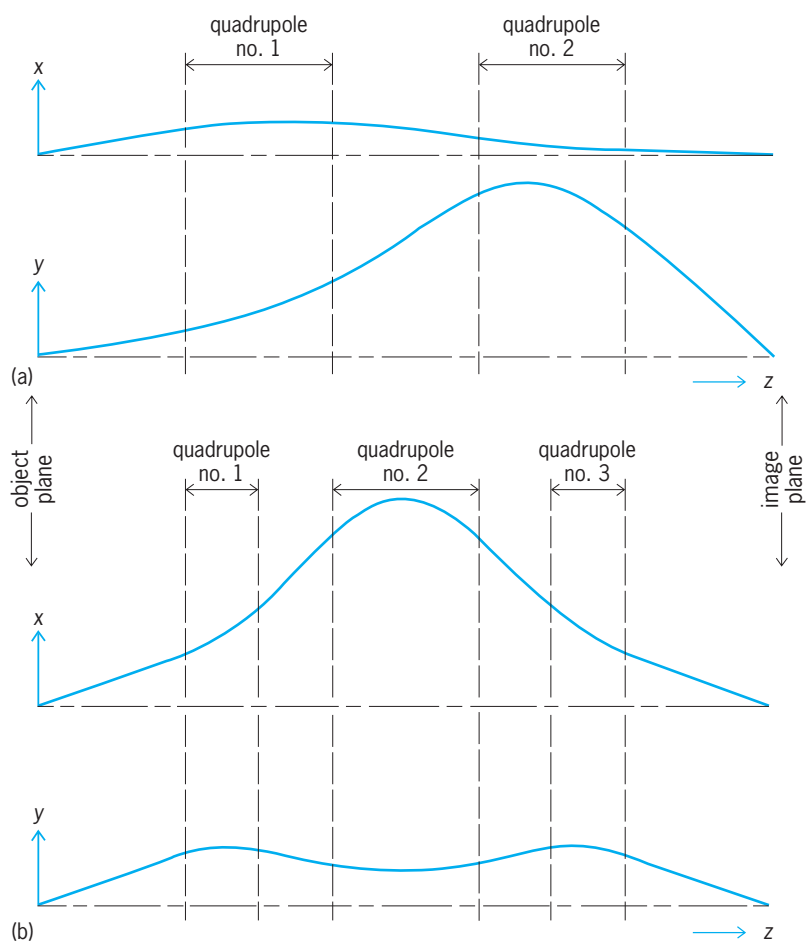

Fig. 4. Diagrammatic representation of electron paths in principal planes of image-forming symmetrical (*a*) doublet and (*b*) triplet.

image. However, this converging effect is only secondary; the primary effect of an axially symmetric electrostatic field is to accelerate or decelerate the electrons parallel to the axis, and the primary effect of an axially symmetric magnetic field is to give electrons diverging from a point on the axis a rotation about the axis. Correspondingly, the converging action or refractive power of a conventional electrostatic lens is proportional not to the first power but to the square of the ratio of the electrostatic field to the accelerating voltage, whereas the converging action of a conventional magnetic lens is proportional to the ratio of the square of the magnetic field to the accelerating potential at low electron energies (<0.5 MeV) and to the square of the ratio of the field to the accelerating potential at high electron energies (>0.5 MeV). Because breakdown phenomena place practical limits on the magnitude of electrostatic fields, and because pole-piece saturation limits (static) magnetic fields, conventional electron lenses are relatively ineffective in converging high-energy particles; they can be properly characterized as weak-focusing lenses.

Quadrupole fields, on the other hand, constitute strong-focusing lenses, insofar as their converging action is directly proportional to the electrostatic or magnetic field respectively. Consequently they (particularly magnetic quadrupole fields) have assumed great importance in the development of high-energy particle accelerators. In this application they permit the narrow confinement of the particle beam and consequently greatly reduce the cost of construction. Apart from the concentration and focusing of high-energy particle beams, strong-focusing lenses have found application as projector lenses (but not as objectives) of electron microscopes. Weak quadrupole fields (stigmators) also are used to compensate residual asymmetries in conventional electron lenses. *See* ELECTRON MICROSCOPE; PARTICLE ACCELERATOR.

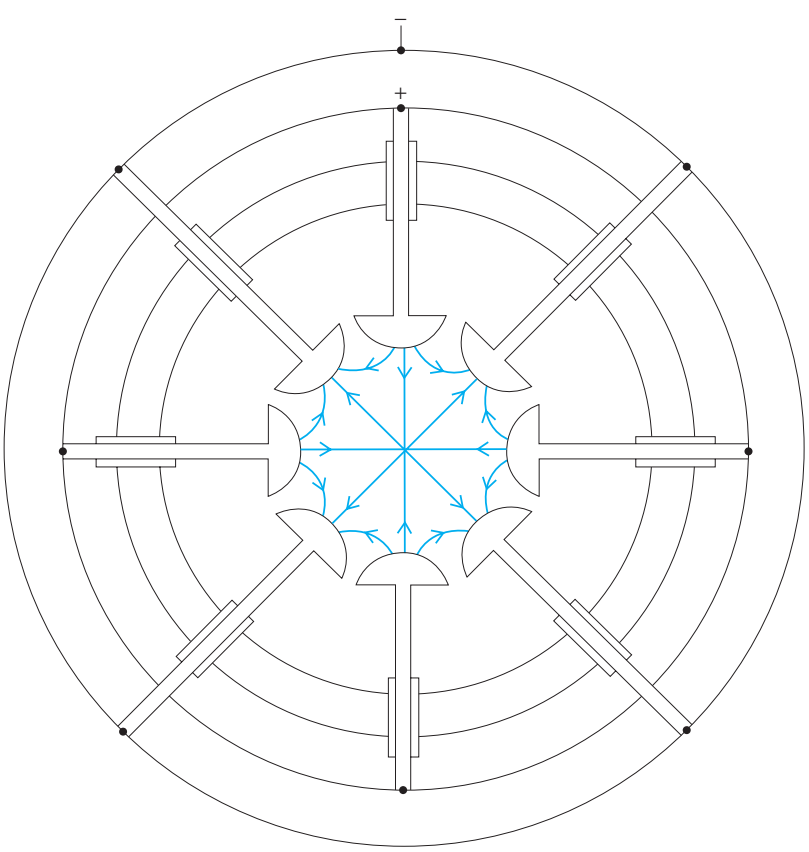

Fig. 5. Electrostatic octupole.

Figure 3 shows sections normal to the lens axis of an electrostatic and a magnetic quadrupole lens. In the first instance the electrostatic potential is symmetric with respect to the two principal (xz and yz) planes and antisymmetric with respect to the planes through the axis forming an angle of 45° with respect to the principal planes. In the second instance the magnetostatic potential is antisymmetric with respect to the principal planes and symmetric with respect to the two 45° planes. Fulfillment of Laplace's equation (5) now demands the form of Eq. (26) for

$$\phi(x, y, z) = \frac{a}{2}(x^2 - y^2) - \frac{d^2a}{dz^2}2(x^4 - y^4)\cdots \quad (26)$$

the electrostatic potential. Magnetic field B, on the other hand, may be written as Eqs. (27)–(29). In Eqs. (26)–(29), a and A are functions of z. Neglecting the

$$B_x = Ay - \frac{1}{6}\frac{d^2A}{dz^2}y(3x^2 + y^2)\cdots \quad (27)$$

$$B_y = Ax - \frac{1}{6}\frac{d^2A}{dz^2}x(x^2 + 3y^2)\cdots \quad (28)$$

$$B_z = 0 - \frac{1}{6}\frac{d^3A}{dz^3}xy(x^2 + y^2)\cdots \quad (29)$$

higher-order terms, which become negligible for a long lens, one finds that the force exerted by the field on the electron is given by Eqs. (30)–(32).

Electrostatic lens		*Magnetic lens*	
$F_x = -eax$	(30*a*)	$F_x = -eAv_2x$	(30*b*)
$F_y = eay$	(31*a*)	$F_y = eAv_zy$	(31*b*)
$F_z = 0$	(32*a*)	$F_z = 0$	(32*b*)

The force is thus, to a first approximation, entirely in a plane normal to the axis. In one principal plane it is such as to produce convergence (direction of force opposite to that of the displacement from the axis), and in the other, such as to produce divergence (direction of force equal to that of the displacement). Two successive quadrupole fields of opposite polarity (a doublet) are required to form a sharp, real image of an object on the axis. This image will have different magnifications in the two principal directions. If the image is to be undistorted, a triplet of three successive quadrupole fields is required. These electron paths are illustrated in **Fig. 4**.

The shaping of the poles influences only terms of the fifth and higher orders in the x and y coordinates in the field expressions. For long lenses all terms except the linear terms vanish if the poles are rectangular hyperbolic cylinders. In practice, circular-cylinder pole caps are more easily realized and give practically the same results. Like axially symmetric lenses, quadrupole lens systems exhibit aberrations, although these are more complex in character.

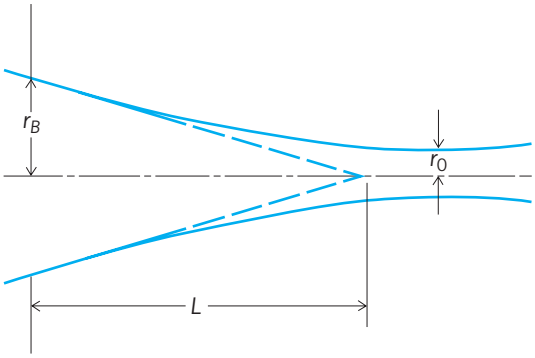

Fig. 6. Widening of electron beam by space-charge repulsion as it traverses from left to right.

Octupole fields (**Fig. 5**), which produce transverse forces proportional to the third power of the displacement from the axis, opposite in polarity in the principal planes and in the 45° planes, are commonly employed to correct aberrations of quadrupole lens systems.

Effect of space charge. Space charge of either positive or negative sign can influence the paths of electrons. Space charge of positive sign is formed by electron beams passing through an imperfectly evacuated space. The beam electrons collide with gas atoms and ionize them. The heavy ions remain in the path of the electron beams for some time and prevent it from spreading. The luminous nodular or thread beams which are produced in this way are favorite objects for demonstration. *See* SPACE CHARGE.

Electron beams in high vacuum, on the other hand, are subject only to the mutually repulsive forces between the electrons themselves. The repulsion is reduced, but never canceled, by the action of the magnetic fields that surround charges in motion; for two electrons moving with the same velocity v parallel to each other, the ratio of the magnetic attractive force to the electrostatic repulsive force is v^2/c^2, where c is the velocity of light. Hence, the magnetic force is significant only for electrons of very high energy.

The action of the remainder of the electrons in the beam upon any one electron can be approximated adequately by that of a continuous charge distribution equal to the average space-charge distribution. The behavior of the edge ray of a uniform circular beam of current I aimed at a point of convergence a distance L from the initial cross section of radius r_B may serve as an example (**Fig. 6**). If the variation of the potential along the axis of the beam is neglected (that is, if ϕ is assumed to be constant), and the charge density ρ is regarded as uniform within any beam cross section, ρ is given by Eq. (33) and the

$$\rho = \frac{I}{\pi r^2}\left(\frac{m}{2e\phi}\right)^{1/2} \tag{33}$$

path equation becomes Eq. (34). Here ϵ is the dielec-

$$\frac{d^2r}{dz^2} = \frac{\pi\rho}{\epsilon\phi}\,r = \left(\frac{m}{2e}\right)^{1/2}\frac{I}{\phi^{3/2}}\frac{1}{\epsilon r} \tag{34}$$

tric constant of vacuum. As the result of the repulsive force of space charge, the ray under consideration does not cross the axis, but reaches a minimum separation r_o from the axis and diverges from this point on. Integration of the differential equation, Eq. (34), gives the radius as Eq. (35). For example, if $r_B =$

$$\begin{aligned} r_0 &= r_B \exp\left[-\epsilon\left(\frac{e}{2m}\right)^{1/2}\frac{r_B{}^2\phi^{3/2}}{L^2 I}\right] \\ &= r_B \exp\left(-3.3\times 10^{-5}\frac{r_B{}^2\phi^{3/2}}{L^2 I}\right) \end{aligned} \tag{35}$$

1 mm, $\phi = 10{,}000$ volts, $L = 10$ cm, and $I = 0.001$ ampere, then $r_o = 0.037$ mm.

Time-varying fields. In the preceding discussion, it was assumed that the electric and magnetic fields traversed by the electrons were constant in time. The total energy of the electron, or the sum of the kinetic energy and the potential energy, is then a constant. Because the potential energy is a function of position only, so is the kinetic energy. This is no longer true if the fields change appreciably in a period corresponding to the transit time of the electrons. *See* KLYSTRON; MAGNETRON; PARTICLE ACCELERATOR; TRAVELING-WAVE TUBE.

Beam deflection in the cathode-ray oscilloscope ceases to be proportional to the potential difference V applied to the deflection plates if V changes appreciably during the passage of the electron beam through the deflection field. If $V = V_0 \cos 2\pi ft$, an integration of the transverse impulse impressed on the electron passing between two parallel plates of length l and separation d expresses the deflection angle α in the form of Eq. (36). Here ϕ is

$$\tan\alpha = \frac{V_0 l}{2\phi d}\frac{\sin u}{u}\cos(2\pi ft) \tag{36}$$

the accelerating potential of the beam, and t is the time the electron passes through the center of the

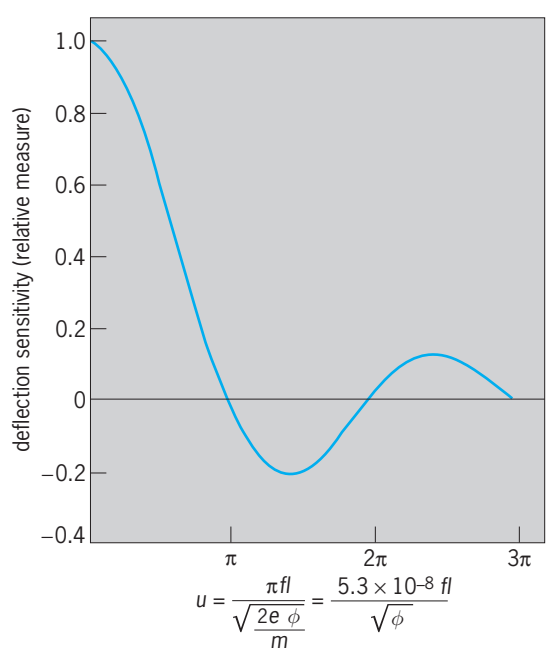

Fig. 7. Deflection sensitivity of cathode-ray oscilloscope as a function of frequency *f*.

deflection field. If f (the frequency of the applied voltage) is measured in s^{-1}, l in centimeters, and ϕ in volts, then Eq. (37) is applicable. The quantity

$$u = \pi f l/(2e\phi/m)^{1/2} = 5.3 \times 10^{-8} f l/\phi^{1/2} \quad (37)$$

$(\sin u)/u$ represents the ratio of the deflection sensitivity at frequency f to that at low frequencies ($f \rightarrow 0$) (**Fig. 7**). Thus, for a 10-kV beam and deflection plates 1 cm in length, the response is 95.4% at 1000 MHz (10^9 s^{-1}), 82.3% at 2000 MHz, 40.2% at 4000 MHz, and it drops to zero at about 5930 MHz. In this discussion, the deflection field is assumed to be a sharply cutoff uniform field, with the effects of fringe fields neglected. Edward G. Ramberg

Bibliography. D. A. DeWolf, *Basics of Electron Optics*, 1990; P. W. Hawkes and E. Kaspar (eds.), *Principles of Electron Optics*, 1996; M. Szilagy, *Electron and Ion Optics*, 1988.

Electron paramagnetic resonance (EPR) spectroscopy

The study of the resonant response to microwave- or radio-frequency radiation of paramagnetic materials placed in a magnetic field. It is sometimes referred to as electron spin resonance (ESR). Paramagnetic substances normally have an odd number of electrons or unpaired electrons, but sometimes electron paramagnetic resonance (EPR) is observed for ions or biradicals with an even number of electrons. EPR spectra are normally presented as plots of the first derivative of the energy absorbed from an oscillating magnetic field at a fixed microwave frequency versus the magnetic field strength. The dispersion may also be detected.

To overcome the intrinsic low sensitivity of the magnetic dipole transitions responsible for EPR, samples are placed in resonant cavities. Routine experiments are carried out in the steady state at a fixed microwave frequency of approximately 9 gigahertz by slowly sweeping the magnetic field through resonance. Free electrons resonate in a magnetic field of 3250 gauss (325 millitesla) at the microwave frequency of 9.1081 GHz, whereas organic free radicals resonate at slightly different magnetic fields characteristic of each particular molecule. *See* ELECTRON SPIN; MAGNETIC RESONANCE; PARAMAGNETISM.

Applications. EPR spectroscopy is used to determine the electronic structure of free radicals as well as transition-metal and rare-earth ions in a variety of substances, to study interactions between molecules, and to measure nuclear spins and magnetic moments. It is applied in the fields of physics, chemistry, biology, archeology, geology, and mineralogy. It is also used in the investigation of radiation-damaged materials and in radiation dosimetry.

The basic physics of transition-metal ions and rare-earth ions present in low concentrations in diamagnetic host crystals has provided a theoretical basis for how electronic structure is modified by the surrounding atoms. Particular applications include probing phase transitions in solids and studies of pairs and triads of magnetically interacting ions.

Applications of EPR in chemistry include characterization of free radicals, studies of organic reactions, and investigations of the electronic properties of paramagnetic inorganic molecules. Information obtained is used in the investigation of molecular structure. EPR is used widely in biology in the study of metal proteins, for nitroxide spin labeling, and in the investigation of radicals produced during reaction processes in proteins and other biomacromolecules. EPR has proved to be an important technique for interdisciplinary investigations of photosynthetic systems. By means of EPR, more than 20 proteins that function in the mitochondrial respiratory chains of mammals have been identified, and details regarding their electron transfer processes have been elucidated.

Sensitivity and resolution. EPR spectra due to pure transition-metal compounds or crystals show broad lines due to interactions between electron magnetic moments. In contrast, in nuclear magnetic resonance (NMR) the interactions between the much smaller nuclear magnetic moments are about 1 million times less intense. Small linewidths and usable resolution require use of dilute solutions in diamagnetic solvents or low concentrations of paramagnetic ions in isomorphous diamagnetic crystal lattices. Concentrations of typically less than 10^{16} molecules per cubic centimeter usually suffice for organic free radicals, but concentrations 10,000 times greater may be needed for inorganic molecules and for transition-metal and rare-earth ions in solids because of the greater linewidths, even for dilute systems. The number of spins required depends on the detectable limit of commercial spectrometers of 10 billion spins with a linewidth of 1 gauss (0.1 tesla). The larger the linewidths, the larger the number of spins required.

The observation of EPR spectra depends on spin-lattice relaxation, which is the exchange of magnetic energy with the thermal motion of the crystal or molecule. For transition-metal ions and rare-earth ions, experiments often require operation at or near liquid helium temperature (4 K; −269°C; −452°F). Organic free radicals can usually be studied successfully at room temperature. *See* MAGNETIC RELAXATION.

Solids. EPR spectra from single crystals clearly provide the greatest amount of information. These include crystals containing small concentrations of paramagnetic ions substituting for the regular ions in the crystal or, for organic molecules, small fractions of free radicals produced by ionizing radiation. Spectra which may contain up to several hundred lines are often highly anisotropic; that is, they change with the orientation of the magnetic field direction in the crystal. Transition-metal-ion and rare-earth-ion EPR spectra in crystals are generally much more anisotropic than free radicals due to the intrinsic anisotropy of the electron magnetic moments, and of other effects that are important when there is more than one unpaired electron.

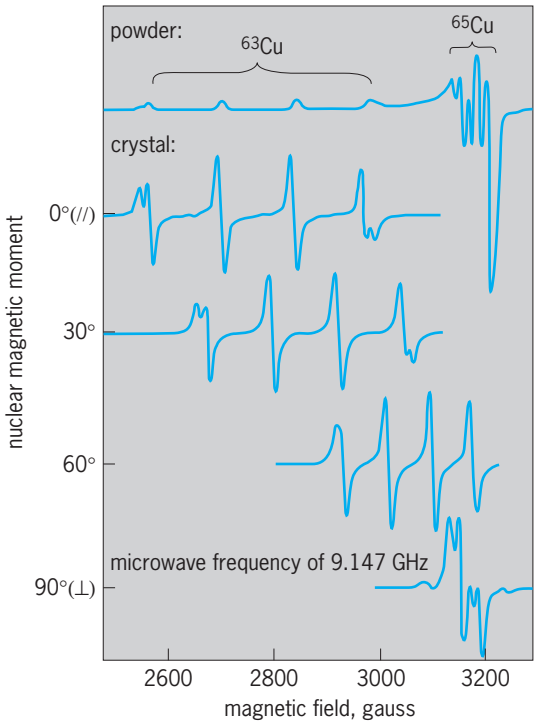

Fig. 1. Powder and single-crystal EPR spectra due to Cu^{2+} ions in calcium cadmium acetate hexahydrate at room temperature. The single-crystal spectra, with the magnetic field making angles of 0°, 30°, 60°, and 90° with the symmetry axis (//), illustrate anisotropy of the electron magnetic moment and the four-line copper hyperfine structure. The presence of the two copper isotopes, ^{63}Cu and ^{65}Cu, is revealed by the partly resolved features. The outer lines belong to ^{65}Cu, which has the larger nuclear magnetic moment. The lack of a characteristic four-line pattern at 90° is well understood. The powder spectra show absorptionlike features associated with 0° crystal spectra and evidence for the second copper isotope (^{65}Cu). (*After J. R. Pilbrow, Transition Ion Paramagnetic Resonance, Oxford University Press, 1990, © J. R. Pilbrow*)

The occurrence of many lines is due to interactions of the orbital motions of electrons with the electric potential of the local surrounding atoms, and to hyperfine interactions between the paramagnetic electrons and nuclear magnetic moments of the paramagnetic ion and surrounding atoms. In the case of free radicals, symmetric or nearly symmetric characteristic hyperfine patterns are observed. From knowledge of hyperfine interactions with nuclei whose spins and magnetic moments are known, the electron distribution throughout a molecule may be determined. Since hyperfine interactions vary as the reciprocal of the cube of the distance between the center of the free radical and the nucleus, structural information may be obtained in addition to electron densities.

The relative values of nuclear moments of different isotopes of a given element often can be determined from the ratio of the hyperfine splittings produced in the same chemical environment (**Fig. 1**).

Liquids and motional averaging. Spectra in solution due to free radicals are often quite simple as a result of motional averaging, and this clearly gives less information than would be obtained from a single-crystal investigation. Linewidths are very narrow (approximately 0.1 gauss or smaller). By varying the temperature above and below room temperature, EPR spectra range from the frozen solution at low temperatures, with a powderlike spectrum, to rapid motional averaging at room temperature where anisotropies are averaged out.

The intermediate region can provide information about slow molecular motions, which is especially important for nitroxide spin labels selectively attached to different parts of macromolecules such as the components of natural and synthetic phospholipid membranes, liquid crystals, and proteins. Such measurements have revealed important structural and functional information. Since the nitrogen hyperfine interaction in nitroxide free radicals has characteristic anisotropy, molecular reorientation times in the range 10^{-9} to 10^{-5} second can be determined. A typical free-radical spectrum in solution shows hyperfine splittings that split into two lines (**Fig. 2**). These splittings are due to two inequivalent nitrogen-14 (^{14}N) nuclei, a single phosphorus-31 (^{31}P) nucleus, and a weak coupling to a single hydrogen (^{1}H) nucleus (or proton). Other commonly occurring nuclear hyperfine splittings in free-radical EPR spectra may be due to deuterium (^{2}H), boron-11 (^{11}B), carbon-13 (^{13}C), or nitrogen-15 (^{15}N). When there are many equivalent protons, the splitting pattern can be calculated from binomial coefficients.

Electron transfer between molecules. Studies of chemical reaction processes involving electron transfer between molecules take advantage of stable free radicals with resolved hyperfine structure, where each line is associated with a particular precession frequency of the unpaired electron spin coupled to a particular configuration of nuclear spins. If an electron jumps to a different molecule nearby with a different arrangement of nuclear spins, its electron magnetic moment will then precess at a slightly different frequency. If the jump rate is small compared with the difference in precession frequencies, the observed linewidths will increase by the jump rate (in hertz). When the jump rates are fast compared with the difference between the precession frequencies, spectra characteristic of each molecule are observed. This latter situation also applies to inorganic transition-metal ion dimer complexes in crystals and frozen solutions where averaging is not observed and the anisotropic spectra provide useful structural clues. *See* ELECTRON-TRANSFER REACTION.

Electron nuclear double resonance (ENDOR). This multiple-resonance method involves simultaneous irradiation at both microwave frequencies and radio frequencies, the latter usually being swept over a range covering the nuclear frequencies. It can be understood as the detection of nuclear magnetic resonance (NMR) through a change in the EPR signal as the nuclear frequencies are swept through one or more resonances. ENDOR patterns for equivalent nuclei are much simpler than in a normal EPR spectrum. For example, for ^{1}H, the ENDOR pattern consists of two lines for any number of equivalent nuclei, whereas the normal EPR spectrum would

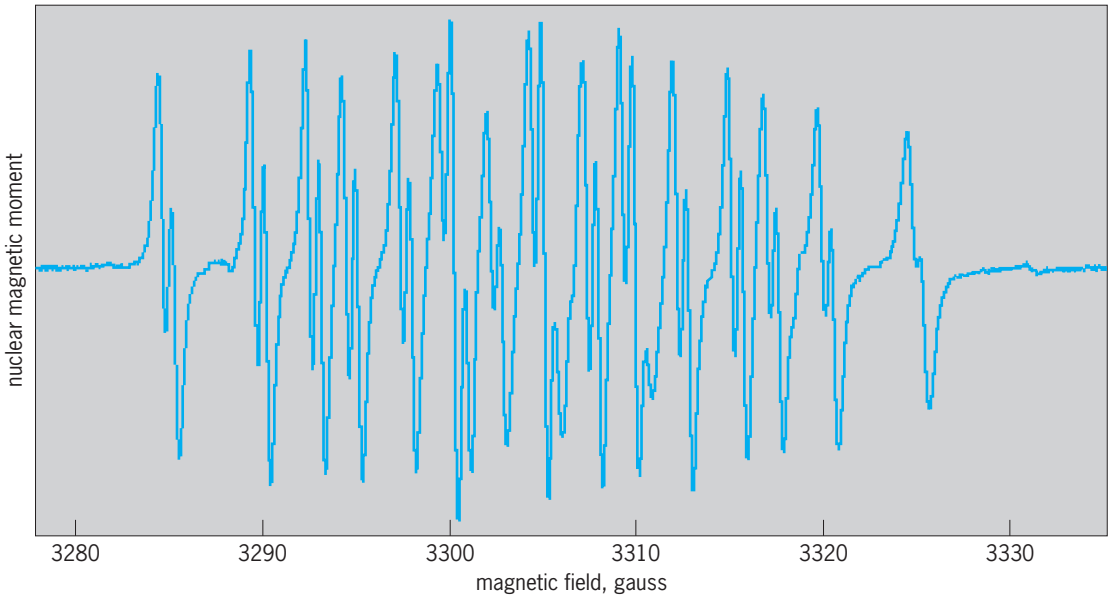

Fig. 2. Room-temperature solution spectrum of the betaine free-radical cation in toluene at a microwave frequency of 9.28877 GHz. The spectrum would be more symmetric at higher microwave frequencies. (*Courtesy of Dr. G. R. Hanson*)

yield multiline spectra with intenstites determined by the binomial coefficients. Linewidths are typically as low as 10 kHz compared with typical values in EPR of approximately 1 MHz for free radicals and 10 MHz for transition-metal ions. ENDOR has been successfully applied to determine properties of hydrogen atoms trapped in alkali halide and calcium fluoride crystals; to locate paramagnetic defects in diamonds, semiconductors, ferroelectrics; and to study organic free radicals.

Electron-electron double resonance (ELDOR). This method involves irradiation at two different microwave frequencies. The changes in the spectrum resulting from either sweeping the second frequency or the magnetic field are monitored through detection at the first frequency. This has been important for studies of slow motions using nitroxide spin labels.

Optical detection of magnetic resonance (ODMR). This method takes advantage of the much higher sensitivity of electric dipole transitions. In ODMR, paramagnetic states are optically excited, and the EPR signal is detected through changes in the optical absorption as the magnetic field is swept through one or more resonances. It has been particularly important for studying spin-triplet states that have nonmagnetic ground states and defects in solids, particularly in semiconductors. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

Pulsed methods. It is often of interest to study the EPR spectra of many chemical or biological systems as a function of time after the microwave power is suddenly switched off. In practice, this is usually accomplished with controlled short high-power pulses of microwave radiation. Pulses as short as 2 billionths of a second have been used. The free induction decay (FID) contains all of the spectral information, and this is mathematically transformed into the normal spectrum using the Fourier transform. The microwave circuit in EPR spectrometers places a limit on measurements, typically until at least 70 billionths of a second following the microwave pulse. While the free induction decay can be observed for many free radicals, it is normally difficult to observe for transition-metal and rare-earth ions. Use of the spin echo is resorted to in these cases.

Electron spin echoes occur whenever two or more short, intense pulses of microwave radiation are applied to a sample with inhomogeneously broadened lines due to many overlapping narrow lines resulting from resonances corresponding to different arrangements of nuclear spin splittings. In the simplest (two-pulse) echo, during the interval τ between the two pulses, the magnetic moments become dephased and lose their coherence. The second pulse is arranged so as to refocus the magnetic moments, thus producing an echo at a later time τ. This means that, provided relaxation times permit, the experimenter has considerable control as to when echoes will occur. If τ is increased in small steps in a two-pulse echo, or suitable time intervals are incremented for multiple-pulse echoes, the echo intensity as a function of time often shows one or more periodic modulations known as electron spin echo envelope modulation (ESEEM). These modulated echo patterns may be mathematically transformed to yield the spectrum of nuclear frequencies. Two-dimensional (2D) forms of spectroscopy are also possible by increasing two different pulse intervals. The resulting 2D plots enable precise correlation of spectra to structural and electronic properties of the active paramagnetic species.

ESEEM and ENDOR provide complementary information. In samples where the unpaired electron in a molecule, or from an ion in a crystal, has high

symmetry, ENDOR spectra are strongest near or along principal, or symmetry, directions; whereas ESEEM has its greatest intensities away from principal directions.

ESEEM has been used effectively to probe weak hyperfine interactions in solids, in assisting with the determination of molecular structure, and in illuminating the local environment of paramagnetic ions at active sites in proteins. J. R. Pilbrow

Bibliography. L. Kevan and M. K. Bowman (eds.), *Pulsed and Continuous Wave ESR*, 1990; J. R. Pilbrow, *Transition Ion Electron Paramagnetic Resonance*, 1990; C. P. Poole and H. A. Farach, *Handbook of Electron Spin Resonance*, 1994; J. A. Weil, J. R. Bolton, and J. E. Wertz, *Electron Paramagnetic Resonance: Elementary Theory and Practical Applications*, 1993.

Electron-positron pair production

A process in which an electron and a positron are simultaneously created in the vicinity of a nucleus or subatomic particle. Electron-positron pair production is an example of the materialization of energy predicted by special relativity and is accurately described by quantum electrodynamics. Pair production usually refers to external pair production, in which the positron (positively charged antielectron) and electron are created from a high-energy gamma ray as it passes through matter. Electron-positron pairs are also produced from internal pair conversions in nuclei, decays of unstable subatomic particles, and collisions between charged particles. *See* QUANTUM ELECTRODYNAMICS.

External pair production. In external conversion, the energy of an incoming gamma ray (a high-energy electromagnetic photon) is directly converted into the mass of the electron-positron pair (**Fig. 1**). The photon energy $h\nu$ (where h is Planck's constant and ν is the photon frequency) must therefore exceed twice the rest mass of the electron $2m_0c^2$, equal to 1.022 MeV (m_0 is the electron mass, c the velocity of light). In order to conserve both energy and momentum in this process, the pair must be created near a nucleus, which recoils to balance the momentum of the incoming photon with the momenta of the created electron and positron. Because the nucleus is so much heavier than the electron, it carries away almost no energy from the pair, and the energy of the photon in excess of $2m_0c^2$ is shared unequally as kinetic energy by the positron and electron. Individually, the electron and positron each exhibit a distribution of kinetic energies ranging from zero to the maximum available energy, $E_{max} = h\nu - 2m_0c^2$, correlated with one another so that their sum is equal to E_{max}. Similarly, the positron and electron are emitted over a broad range of angles, although they exhibit a tendency to move in the same direction, which reflects the momentum of the incoming photon. For incident photon energies above 5 MeV, external pair production is the dominant mechanism by which gamma rays are absorbed in matter. *See* GAMMA-RAY DETECTORS; GAMMA RAYS; PHOTON.

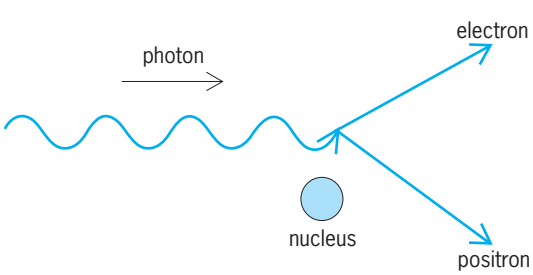

Fig. 1. External electron-positron pair production.

When a highly energetic charged particle collides with matter or with another particle, positron-electron pairs may be copiously produced by the external conversion of bremsstrahlung photons emitted as the incident particle is decelerated. This process is repeated as the created particles in turn produce pairs, until the average available energy is below the pair production threshold, giving rise to an electromagnetic cascade or shower. Showers are commonly observed in cosmic-ray interactions and are often used to detect high-energy electrons and photons produced by particle accelerators. *See* COSMIC RAYS.

Internal pair creation. Internal pair creation differs from external conversion in that the positron and electron are created directly from energy liberated by the deexcitation of an excited nucleus (produced, for example, in radioactive decay or nuclear collisions) to a state of lower energy, if the transition energy exceeds the pair mass threshold of $2m_0c^2$. Internal pair creation usually occurs only 10^{-3} times as often as deexcitation by gamma-ray emission, although the exact pair creation probability, as well as the angular correlation between the emitted positron and electron, depends on the nuclear charge and upon the energy and multipolarity of the nuclear transition. *See* RADIOACTIVITY.

Particle decay. Many unstable subatomic particles, such as the neutral Z boson and J/psi meson, decay into a positron-electron pair alone or with other particles. Since the decaying parent particle is massive, momentum is conserved without the presence of an additional nucleus as is required in external conversion. Decay into a single pair alone creates a positron and electron with equal and opposite momenta. They share the available kinetic energy equally, that is, the equation below is satisfied, where

$$E_{e^+} = E_{e^-} = \frac{Mc^2 - 2m_0c^2}{2}$$

M is the mass of the parent, and they are emitted back to back in the rest frame of the parent particle. *See* ELEMENTARY PARTICLE; MESON.

Superheavy nuclear collisions. In the 1980s, a new mode of positron-electron pair production was discovered in collisions of very heavy nuclei at high energies (**Fig. 2**). Narrow positron and electron peaks

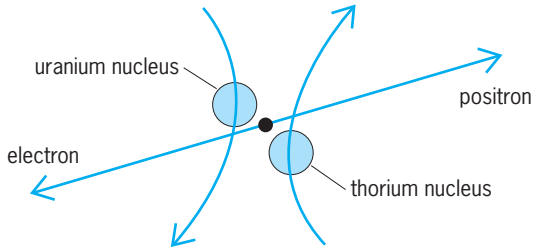

Fig. 2. Anomalous positron-electron pair creation in superheavy quasiatomic collisions.

were detected in coincidence from 1.4-GeV uranium-on-thorium quasiatomic collisions; three sets of correlated structures were found. In each case, the corresponding peak in the distribution of the sum of the positron and electron energies is narrower than the individual Doppler-broadened lines. This suggests the correlated cancellation of the positron and electron laboratory Doppler shifts, which is characteristic of an 180° back-to-back emission-angle correlation in the rest frame of their source. Explanations for these structures involving instrumental backgrounds, uncorrelated emission of separate positron and electron lines, external or internal pair production, and multibody final states are unable to describe the observed narrow width of the sum-energy peak. The data appear to reflect the production and decay of at least three previously undetected neutral objects with masses of 1.63, 1.77, and 1.84 MeV/c^2. Attempts to explain these data in terms of one or several new elementary particles have been unsuccessful so far. *See* POSITRON; QUASIATOM.

Thomas E. Cowan

Bibliography. M. L. Burns and A. K. Harding (eds.), *Positron-Electron Pairs in Astrophysics*, 1983; R. D. Evans, *The Atomic Nucleus*, 1955, reprint 1982; W. Greiner (ed.), *Physics of Strong Fields*, 1987; K. Siegbahn (ed.), *α-, β-, and γ-Ray Spectroscopy*, 1968; J. Thompson, *Soft Lepton Pair and Photon Production*, 1991.

Electron-probe microanalysis

A method used for determining the elemental composition of materials, based on the x-rays emitted by different elements when bombarded with high-energy electrons. It is a micro method that can detect x-ray photons emitted by the atoms within a small volume excited by an electron beam focused to 10 nanometers diameter or less. In biology electron-probe microanalysis can be used to determine the composition of cell organelles without isolating them and therefore altering the distribution of diffusible elements Studies with electron-probe analysis have thrown light on the major problems of excitation-contraction coupling in muscle, visual transduction in retinal rods, and the functions of the endoplasmic reticulum and mitochondria in the regulation of cytoplasmic calcium.

High-energy electrons can ionize atoms, ejecting an inner-shell electron. To fill the resultant vacancy, an outer-shell electron falls into the ionized shell (**Fig. 1**); the atom remains in the higher-energy excited state. The emission of a characteristic x-ray by the ionized atom is one of the mechanisms for releasing its excess energy. Another mechanism is the emission of Auger electrons; the probability of ejecting an x-ray instead of an Auger electron is the fluorescence yield. In the case of electron-probe microanalysis the source of the exciting electrons is the electron gun, which, in modern electron microscopes equipped with field-emission guns, can produce a focused beam narrower than 1 nm. The x-rays emitted as the result of atomic ionization are called characteristic x-rays, because their energy is characteristic of the core shell of the ionized element and of the shell from which the electron relaxed into the vacancy. X-rays excited by ionization of the innermost (K) shells are called K x-rays, those due to L-shell ionization, L x-rays. Typical energies of characteristic K x-rays of the biologically important elements are given in the **table**. *See* AUGER EFFECT; SURFACE PHYSICS; X-RAY FLUORESCENCE ANALYSIS.

X-ray detection and data processing. At the great sensitivity required for biological applications at high spatial resolution, x-rays are detected with energy-dispersive detectors. They are solid-state devices in which incident x-rays create electron-hole pairs. In those most commonly used, lithium-drifted silicon (Si-Li) detectors, each 3.8 eV of energy discharged by an x-ray photon creates one electron-hole pair. Hence, following the arrival of a high-energy x-ray, the charge deposited and then swept off the detector is a measure of the energy of the incident x-ray. The charge after pulse and digital processing, can be stored in one of the channels of a multichannel analyzer. This procedure builds up a spectrum of the number of emitted x-rays as a function of their

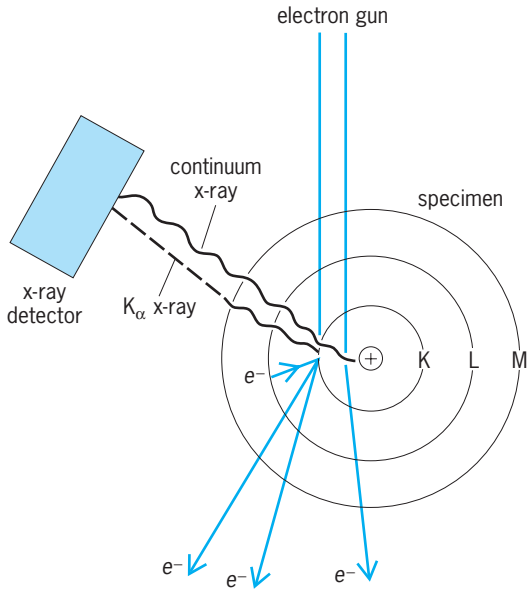

Fig. 1. Ionization of an atom in a specimen and production and detection of the characteristic and continuum x-rays generated by the incident electrons. Unscattered and elastically or inelastically scattered electrons are shown leaving the specimen plane.

Energies of K x-rays

Element	Energy, keV
Sodium	1.041
Magnesium	1.253
Phosphorus	2.013
Sulfur	2.306
Potassium	3.310
Calcium	3.691

energies (**Fig. 2**). The energy resolution of modern x-ray detectors, usually specified at the Kα manganese line, is approximately 150 eV. This represents significant broadening of the intrinsic (4 eV or less) width of the x-ray lines by electronic noise and other factors, and can cause overlap of adjacent x-ray peaks. The analysis of overlapping peaks, required for accurate determinations, is performed by computer.

Quantitative analysis. The concentrations of elements are determined from the relationship, given in the equation below, between the number of atoms

$$I_x = \omega_x \cdot q_x \cdot N_x \cdot T_x$$

(N_x) in an irradiated space and the number of characteristic x-rays (I_x) emitted. Here ω_x is the flourescence yield, q_x the ionization cross section (the probability of an atom being ionized), N_x the number of atoms of element X in the space irradiated, I the probe current, and T_x the detector transmission efficiency. Thus, the number of x-rays generated in a given space is directly proportional to the number of atoms and the number of exciting electrons. The two major influences on the performance of the detector are the geometric efficiency with which the emitted x-rays reach it and the extent of x-ray absorption by its window. The geometry of electron microscope lenses is such that less than 1% of emitted x-rays reaches an energy-dispersive detector. Low-energy x-rays are absorbed by the detector window: elements below sodium in atomic number are not detected, and the detection efficiency for sodium and magnesium is less than that for, say, potassium. Elements of low atomic number can be analyzed with windowless x-ray detectors or electron energy-loss spectroscopy. *See* ELECTRON MICROSCOPE; ELECTRON SPECTROSCOPY.

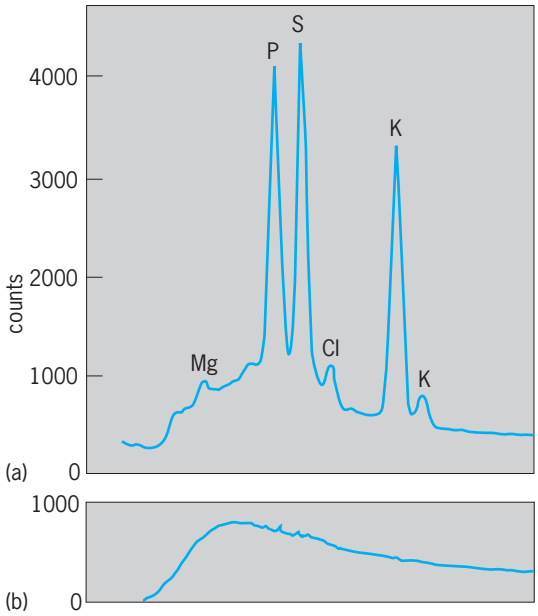

Fig. 2. Typical cellular x-ray spectrum, indicating high concentration of potassium: (*a*) a series of characteristic elemental peaks superimposed on (*b*) a background of x-ray continuum, shown with the characteristic peaks removed by computer processing.

X-ray photons are also absorbed by the material being analyzed, complicating the analysis of thick and bulky specimens. Spatial resolution is also reduced in thick specimens through broadening of the probe by electron scattering. Biological studies at the high spatial resolution required for analyzing cell organelles are performed on sections less than 200 nm thick.

The concentration of elements is also best determined in ultrathin specimens, because excitation of an atom with electrons also creates continuum x-rays, which result from inelastic scattering by atomic nuclei. Continuum x-rays are distributed over a broad energy band (Fig. 2*a*); their number is proportional to the total mass within the irradiated space. The same energy-dispersive x-ray detectors simultaneously detect characteristic and continuum x-rays, and the ratio of the two is a measure of elemental concentration, usually expressed as mmol-kg, dry weight. With modern instrumentation, concentrations of calcium as low as 0.3 mmol/kg can be detected within a 100-nm^3 volume or less.

X-ray mapping. X-ray mapping, a second mode of electron-probe microanalysis application, yields a two-dimensional map of elemental distribution. X-ray maps are obtained by scanning an area of a specimen with the electron beam and collecting the x-ray spectrum at each picture point (pixel) as the beam is rastered. In this manner, the concentration of different elements at each picture point can be determined, and the resultant values displayed as relative intensities on x-ray maps of, for example, phosphorus or calcium. Powerful electron sources are required with the small probes that give high spatial resolution, to obtain low-noise x-ray maps of biological specimens. Field-emission guns are the strongest available and provide the highest current density in a focused beam. The distribution of several elements in specific granules of the atrium of a rat heart, mapped with a field-emission gun, is shown in **Fig. 3**. *See* X-RAY MICROSCOPE.

Specimen preparation by rapid freezing. Extreme care must be taken to prevent the loss or translocation of diffusible elements during specimen preparation. Freezing at rates of almost 10,000°C/s (18,000°F/s) is used almost universally, followed by vacuum dehydration. Valuable information, albeit at lower resolution, has been obtained by analysis of frozen hydrated (ice) specimens of secretory epithelia.

Biological applications. Important biological structures have been analyzed with electron-probe

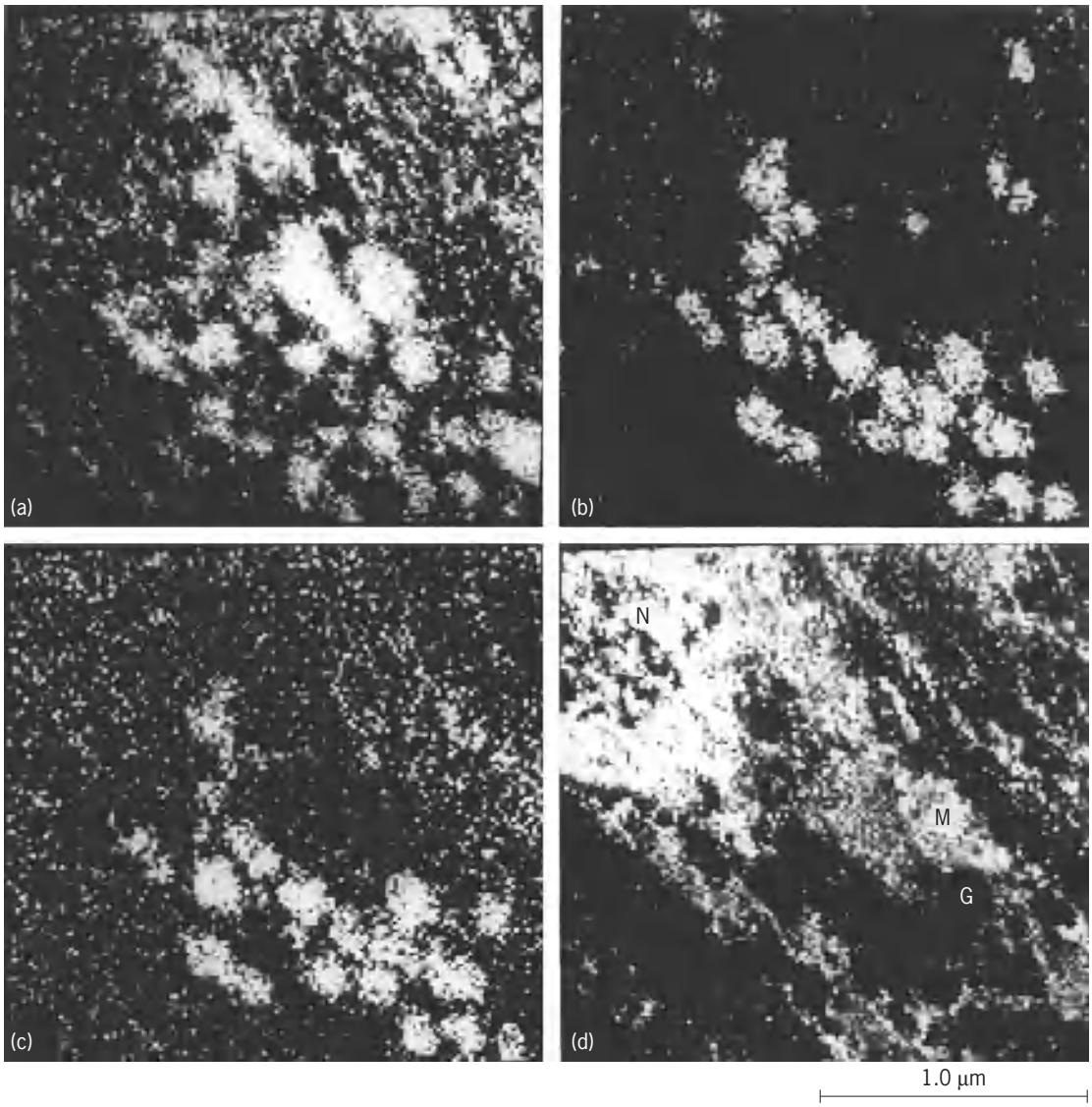

Fig. 3. X-ray maps of a rat heart cryosection, showing the distribution of (*a*) sulfur, (*b*) calcium, (*c*) chlorine, and (*d*) phosphorus in a 2-μm^2 region of the nuclear pole, containing mitochondria (M), atrial granules (G), and condensed chromatin inside the nuclear (N) envelope. The intensity of gray indicates distribution; black represents the fewest and white the most numerous atoms of each element.

microanalysis, including sarcoplasmic and endoplasmic reticulum, mitochondria, and nuclei. Largely as a result, mitochondria are now known to accumulate significant amounts of calcium only under pathological conditions when cytoplasmic calcium rises to abnormally high levels; in normal cells the endoplasmic (in muscle the sarcoplasmic) reticulum, in conjunction with plasma membrane calcium pumps, is the principal regulator of cytoplasmic calcium. Electron-probe microanalysis has revealed subcellular movements of magnesium in and out of mitochondria and sarcoplasmic reticulum that were not anticipated from studies of isolated organelles. Electron-probe microanalysis of bacteria and bacterial spores has revealed the distribution of magnesium, calcium, and other elements, as well as an increase in cell calcium during bacterial division. The application of this technique to abnormal red blood cells in sickle cell anemia has shown that their high calcium content is concentrated in small vacuoles. *See* MASS SPECTROMETRY; SECONDARY ION MASS SPECTROMETRY (SIMS); SPECTROSCOPY.

Andrew Paul Somlyo

Bibliography. A. Le Furgey, M. Bond, and P. Ingram, Frontiers in electron probe microanalysis: Application to cell physiology, *Ultramicroscopy*, 24:185–219, 1988; A. P. Somlyo, Cell calcium measurement with electron probe and electron energy loss analysis, *Cell Calcium*, 6:197–212, 1985; A. P. Somlyo, Compositional mapping in biology: X-rays and electrons, *J. Ultrastruc. Res.*, 83:135–142, 1985.

Electron spectroscopy

A form of spectroscopy which deals with the emission and recording of the electrons which constitute matter—solids, liquids, or gases. The usual form of spectroscopy concerns the emission or absorption

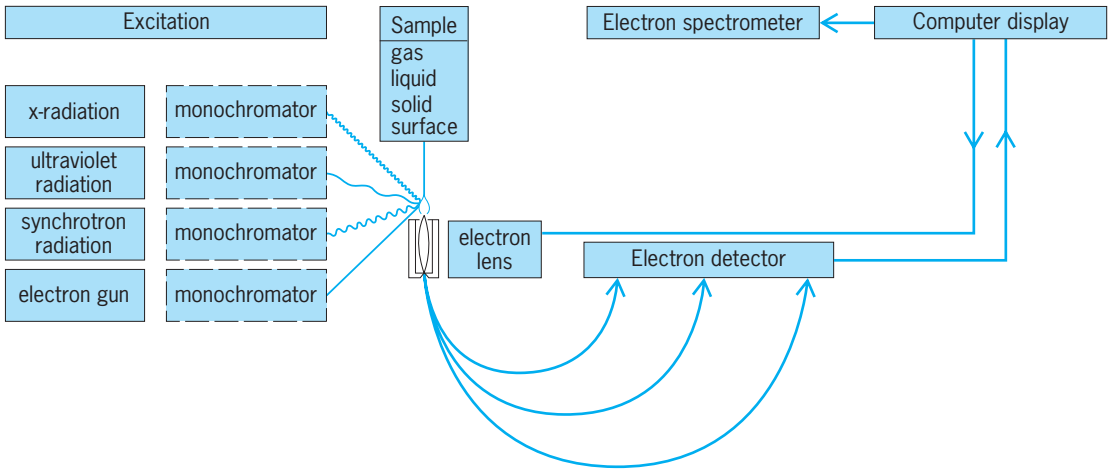

Fig. 1. Excitation of electron spectra recorded with high-resolution instruments. (*After H. Siegbahn and L. Karlsson, Photoelectron spectroscopy, Handbuch der Physik, vol. 31, 1982*)

of photons [x-rays, ultraviolet (uv) rays, visible or microwave wavelengths, and so on]. Electron spectra can be excited by x-rays, which is the basis for electron spectroscopy for chemical analysis (ESCA), or by uv photons, or by ions (electrons; **Fig. 1**). By means of x-ray or uv photons with energy E_{bv}, photoelectron spectra (PES) are produced when electrons with binding energies E_b are emitted with energy E_{kinetic} from bound molecular states, according to the equation below. E_r is a usually negligible small recoil

$$E_{\text{kinetic}} = E_{bv} - E_b - E_r - \varphi$$

energy, and φ a small work function correction which can be attributed to contact potentials. For insulating solid materials, precautions are taken for

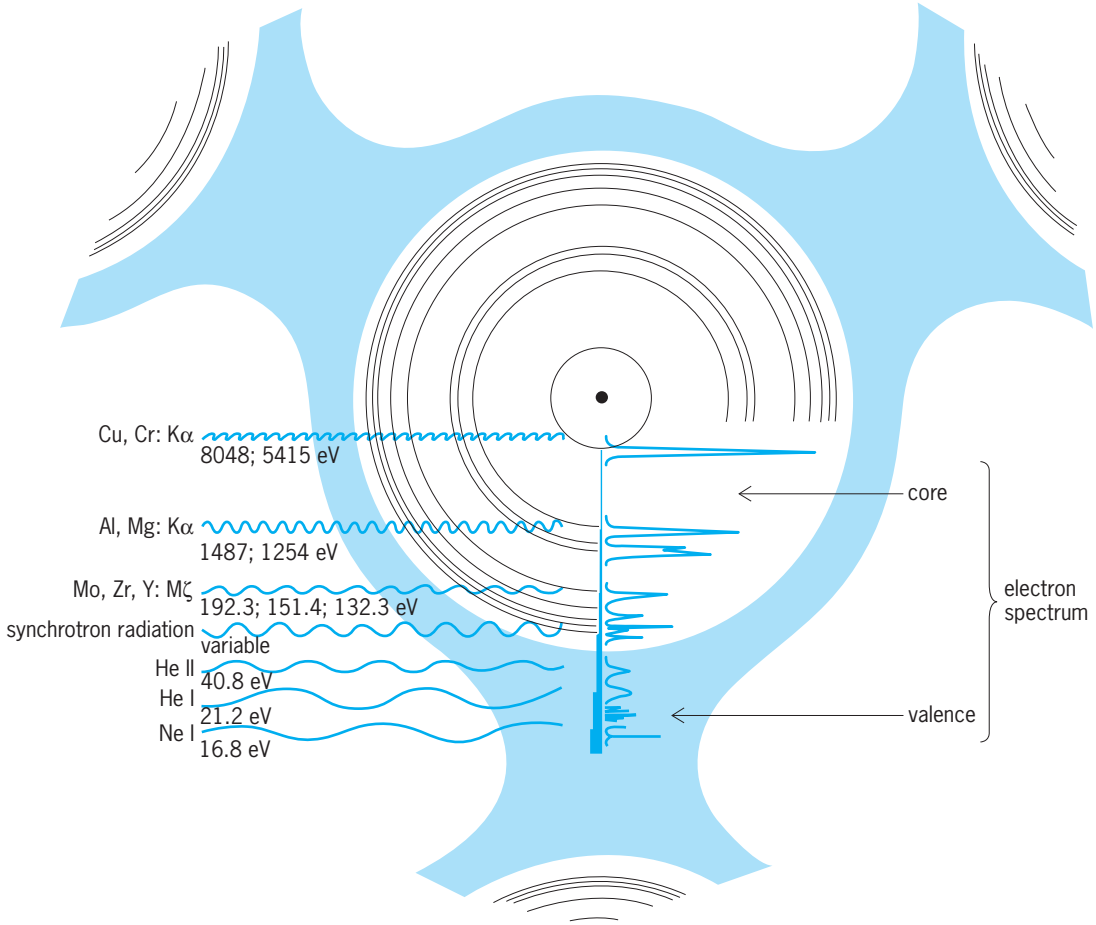

Fig. 2. The system of levels of an atom in a molecule can be divided into a valence electron region and an atomic-core region. Excitation of electron lines from the various regions can be made at different photon energies.

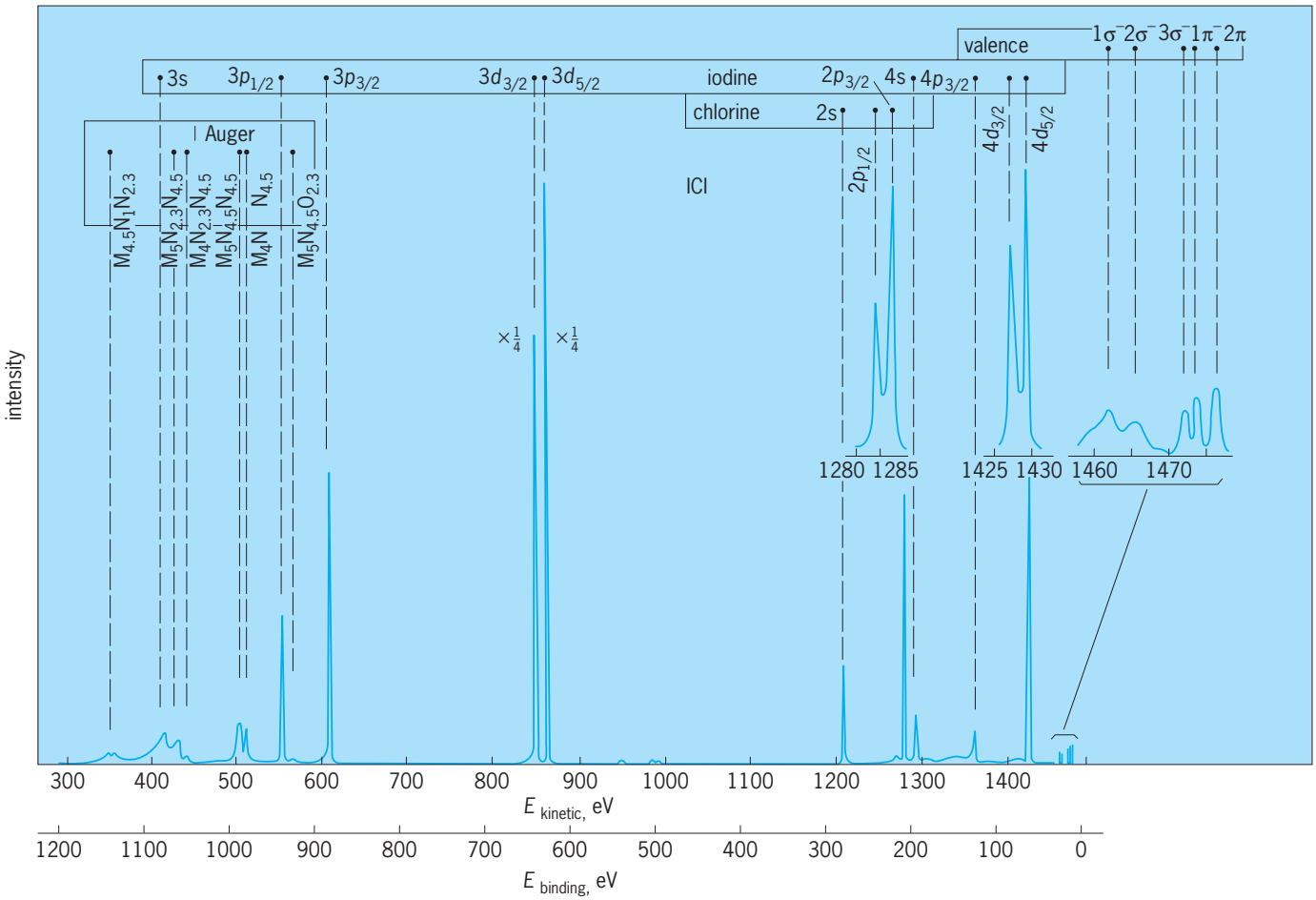

Fig. 3. Electron spectrum of the vapor of iodine chloride at a pressure of 0.04 torr (5 Pa), excited monochromatic Al Kα radiation with the ESCA instrument (scan spectrum). The M- and N-shell levels are excited in iodine, and the L levels in chlorine. At low binding energies (high kinetic energies), the valence spectrum of the molecule is recorded. The MNN Auger electron lines of iodine can also be seen.

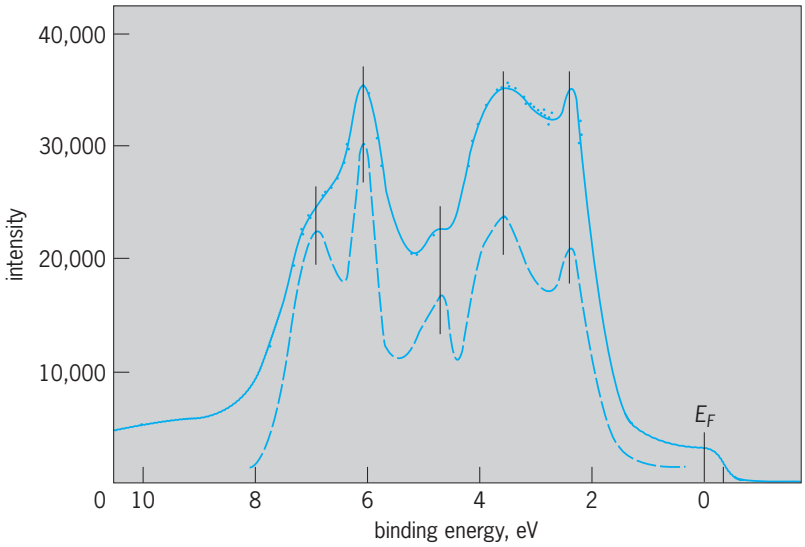

Fig. 4. Valence-band electron spectrum of gold excited by monochromatized Al Kα radiation (solid curve). Comparison is made with a theoretically calculated density of states (broken curve) which has been folded with a 0.25-eV gaussian corresponding to the experimental resolution. The energy scale of the theoretical density-of-states curve has also been expanded by 7.5%, which leads to a near-perfect correspondence between theoretical and experimental band profiles. E_F is the Fermi edge. $h\nu = 1486.7$ eV. Foil, $\theta = 45^\circ$. (*After U. Gelius et al., Experimental spectrum, Nucl. Instr. Meth., B1:85 1984; and N. E. Christensen and B. O. Seraphin, Theoretical density of states, Phys. Rev., B4:3321, 1971*)

stabilizing the surface potential. For solids, the binding energies are preferably referred to the Fermi level, and for gases, to the vacuum level.

Modes of excitation. By means of ESCA, complete sets of photoelectron lines can be excited from the internal (core) levels as well as from the external (valence) region (**Fig. 2**). Also, complete sequences of the Auger electron lines are automatically obtained in this mode. A convenient source of excitation is the x-radiation from Al at 1487 eV. With the use of spherically bent quartz crystals this radiation can be further monochromatized to an inherent width of 0.2 eV. The commonly used light source for uv excitation of photoelectron spectra within the valence electron region is a helium lamp that produces highly monochromatic radiation at 21.2 eV (internal width less than 10 mev) and (with less intensity) at 40.8 eV. Intermediate in energy for excitation are ultrasoft x-rays (Y Mζ at 132 eV). Another source of excitation in the intermediate region is provided by synchrotron radiation, which can be continually varied by means of a suitable monochromater. Excitation by means of an electron beam is an alternative mode to obtain Auger electron lines that has the advantage of ease of production. Ordinary photoelectron

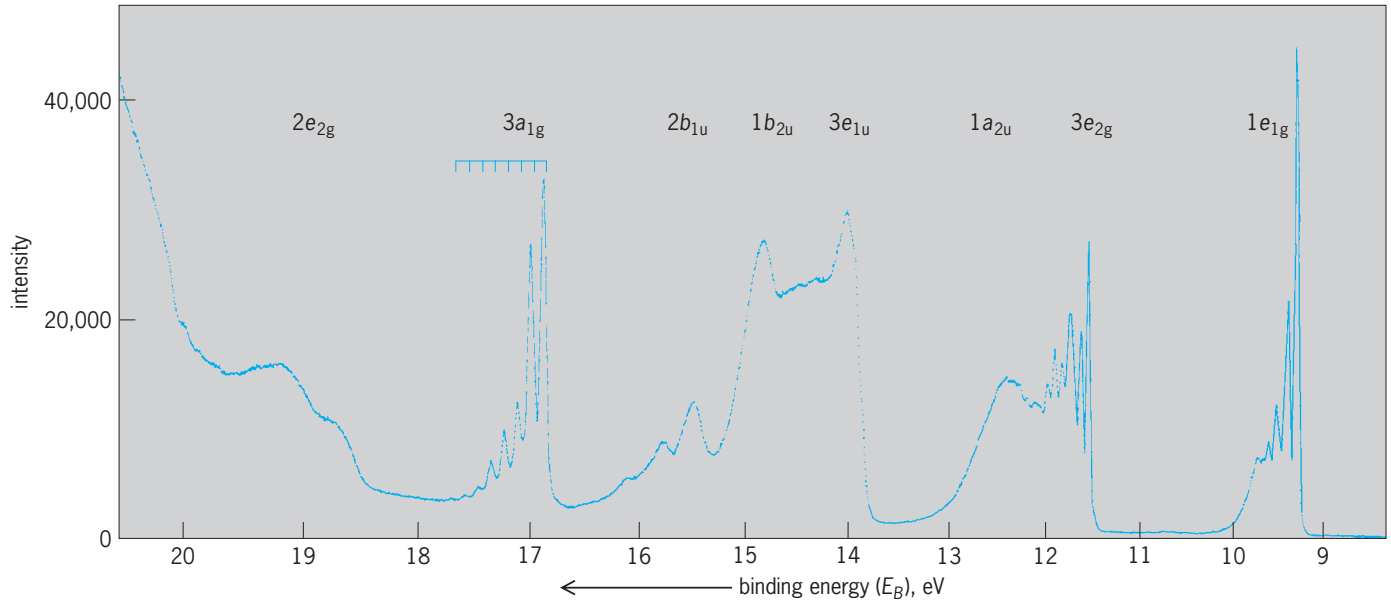

Fig. 5. Benzene valence electron spectrum excited by HeI radiation at 21.22 eV, showing vibrational structures.

lines are not produced by that radiation. In order to compensate for the low-signal-to-background ratio when Auger electron spectra are excited by an electron beam impinging on a solid surface, the spectral distribution is frequently differentiated. *See* AUGER EFFECT.

Applications. In ESCA, only electrons which are expelled from a surface layer of less than 50 nanometers of a solid material contribute to the electron line with the kinetic energy given above. Electrons from the interior of the material are scattered out from the line, and form a low background which does not interfere with the line character of the ESCA spectrum. The electron lines are extremely sharp and well suited for precision measurements. With a high- resolving ESCA spectrometer which has a magnetic or electrostatic focusing dispersive system, the electron lines have widths which are set by the limit caused by the uncertainty principle (the "inherent" widths of atomic levels). With a suitable choice of radiation, electron spectroscopy reproduces directly the electronic level structure from the innermost shells (core electrons) to the atomic surface (valence or conduction band; **Fig. 3**). Furthermore, all elements from hydrogen to the heaviest ones can be studied even if the element occurs together with several other

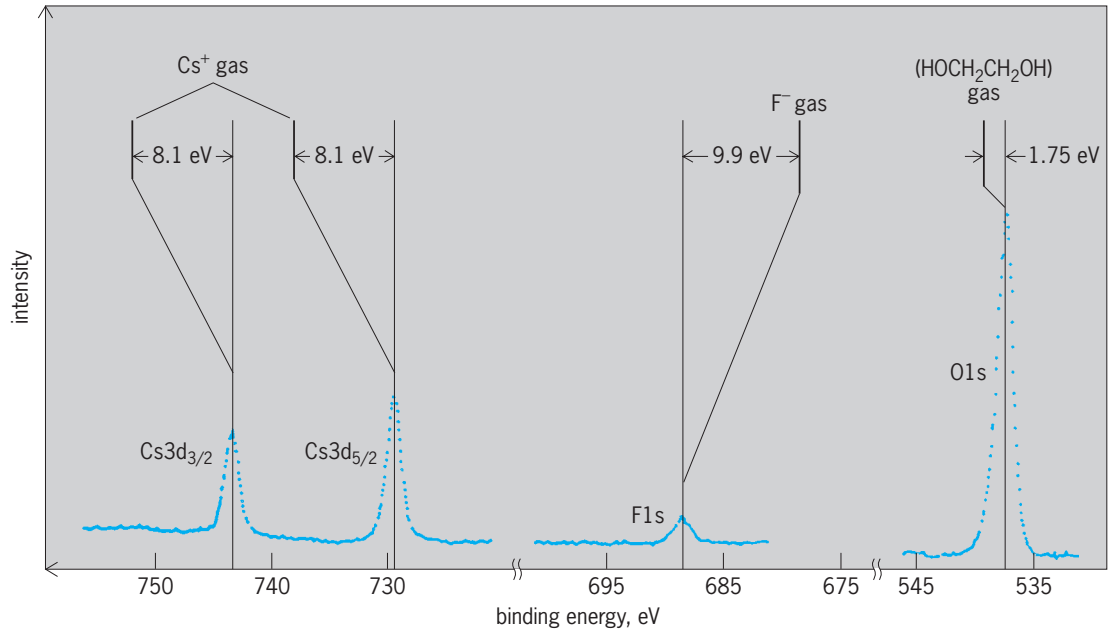

Fig. 6. ESCA spectrum of cesium fluoride (CsF) in glycol ($HOCH_2CH_2OH$) solution. Comparison is made with the corresponding electron-binding energies in the gas phase. (*After H. Siegbahn et al., Phys. Scripta, 27:431, 1983*)

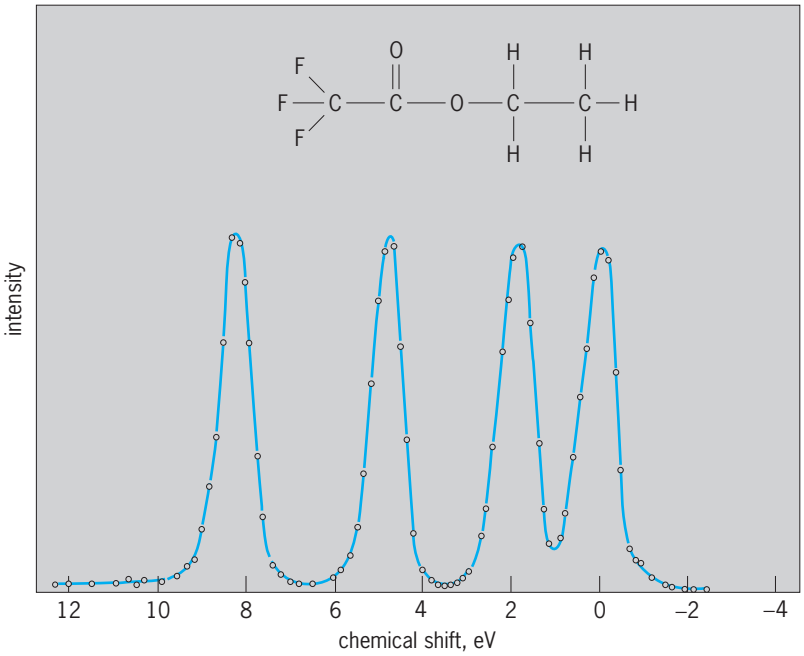

Fig. 7. Electron spectrum from the 1*s* level of carbon in ethyl trifluoroacetate; binding energy = 291.2 eV.

elements and even if the element represents only a small part of the chemical compound. *See* LINE SPECTRUM; UNCERTAINTY PRINCIPLE.

When applied to solid materials, ESCA is a typical surface spectroscopy with applications to problems such as chemical surface reactions, for example, corrosion or heterogeneous catalysis. In such cases, ultrahigh vacuum conditions are usually required, with a vacuum less than 10^{-9} torr (10^{-7} pascal). ESCA also reproduces bulk matter properties such as valence electron band structures (**Fig. 4**). With uv excitation, the resolution is sufficient to resolve vibrational structures in valence electron spectra for gases (**Fig. 5**). In fact, electron spectroscopy can supply a detailed knowledge of the valence orbital structure for all molecules which can be brought into gaseous form with pressures of 10^{-5} torr (10^{-3} Pa) or more. A convenient pressure region is 10^{-2} torr (10 Pa). Liquids and solutions of various compositions can also be studied by ESCA techniques (**Fig. 6**).

ESCA chemical shifts. When atoms are brought close together to form a molecule, the electronic orbitals of each atom are perturbed. Inner orbitals, that is, those with higher binding energies, may still be regarded as atomic and belonging to specified atoms within the molecule, whereas the outer orbitals combine to form the valence-level system of the molecule. These orbitals play a more or less active part in the chemical properties. The chemical bonds affect the charge distribution so that the original atoms can be regarded as charged to various degrees; a neutral molecule has a net charge of zero. The individual atoms in the molecule can be regarded as spheres with different charges set up by the transfer of certain small charges from one atomic sphere to the neighboring atoms taking part in the chemical bond. Inside each charged sphere the atomic potential is constant, in accordance with classical electrostatic theory. The result of this atomic potential is to shift the whole system of inner levels in an atom by a small amount, the same amount for each level. Levels belonging to different atoms in the molecule are shifted differently, however, and if the ESCA chemical shifts for individual atoms in the molecule are measured, a mapping can be made of the distribution of charge or potential in the molecule. This is a reflection of the chemical bondings between the atoms, which in turn can be described by the orbitals in the valence-level system.

A unique feature of ESCA is that, if the exact position of the electron lines characteristic of the various elements in the molecule is measured, the area of inspection can be moved from one atomic species to another in the molecular structure. **Figure 7** shows the electron spectrum from the 1*s* level of the carbon in ethyl trifluoroacetate. All four carbon atoms in this molecule are distinguished in the spectrum. The lines appear in the same order from left to right, as do the corresponding carbon atoms in the structure that is shown in the figure. If the structure of the molecule is known, the charge distribution can be estimated in a simple way by using, for example, the electronegativity concept and assuming certain resonance structures. More sophisticated quantum-chemical treatments can also be applied. Conversely, if, by means of ESCA, the approximate charge distribution is known, conclusions about the structure of the molecule can be drawn.

Experimental evidence obtained so far for various elements in a large number of molecules indicates strong correlations between chemical shifts and calculated atomic charges. A typical correlation curve for the 2*p* level in sulfur obtained from more than 100 compounds containing this element is shown in **Fig. 8**. Similar curves are obtained for other elements such as carbon, nitrogen, oxygen, and phosphorus. Chemical shifts are also observed in the electron lines

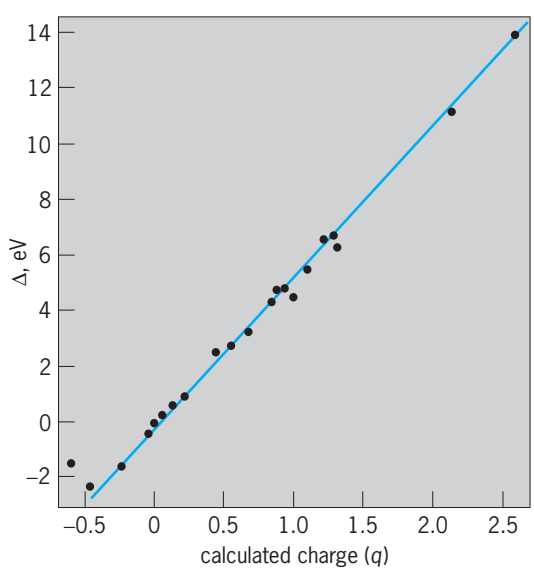

Fig. 8. Binding-energy shifts for the sulfur 2*p* electrons versus the calculated charge. The points indicate averages from more than 100 compounds.

due to the Auger effect. Second-order chemical shifts from groups situated farther away in the molecule (inductive effects) are also observed.

The theoretical importance of the chemical shift effect lies in the study of molecular electronic structure. In other contexts, its usefulness lies precisely in its ability to specify chemical composition, in particular, at surfaces. For example, a metal and its oxide give two distinctly different lines because of the chemical shift effect. It is often easy to follow the rate of oxidation at a surface, or the adsorption of a gas layer onto it. The chemical composition of a surface or the changes due to various chemical or physical treatments are amenable to detailed examination. *See* ATOMIC STRUCTURE AND SPECTRA; CHEMICAL BONDING; ELECTRON CONFIGURATION; ELECTRONEGATIVITY; MOLECULAR ORBITAL THEORY; SPECTROSCOPY.

Kai Siegbahn; Hans Siegbahn

Bibliography. T. L. Barr, *Electron Spectroscopy–Chemical Analysis (ESCA)*, 1994; C. R. Brundle and A. D. Baker (eds.), *Electron Spectroscopy, Theory, Techniques and Applications,* vols. 1–5, 1977–1984; M. Cardona and L. Ley (eds.), *Photoemission in Solids*, vols. 1 and 2, 1978, 1979; K. Siegbahn et al., *ESCA Applied to Free Molecules*, 1969; K. Siegbahn et al., *ESCA: Atomic, Molecular and Solid State Structure Studies by Means of Electron Spectroscopy*, 1967; G. C. Smith, *Surface Analysis by Electron Spectroscopy*, 1994; E. L. Wolf, *Principles of Electron Tunneling Spectroscopy*, 1985, reprint 1989.

Electron spin

That property of an electron which gives rise to its angular momentum about an axis within the electron. Spin is one of the permanent and basic properties of the electron. Both the spin and the associated magnetic dipole moment of the electron were postulated by G. E. Uhlenbeck and S. Goudsmit in 1925 as necessary to allow the interpretation of many observed effects, among them the so-called anomalous Zeeman effect, the existence of doublets (pairs of closely spaced lines) in the spectra of the alkali atoms, and certain features of x-ray spectra. *See* SPIN (QUANTUM MECHANICS).

All theory that concerns itself with electronic, nuclear, atomic, and molecular phenomena includes the electron spin in its formulation to obtain a theoretical structure consistent with experimental observation. The electron thus possesses the intrinsic property of spin angular momentum (rotational motion about an axis), in addition to the intrinsic properties of charge and mass. *See* ELECTRON.

The spin quantum number is s, which is always $^1/_2$. This means that the component of spin angular momentum along a preferred direction, such as the direction of a magnetic field, is $\pm{}^1/_2\hbar$, where $\hbar = h/2\pi$ and h is Planck's constant. The spin angular momentum of the electron is not to be confused with the orbital angular momentum of the electron associated with its motion about the nucleus. In the latter case the maximum component of angular momentum along a preferred direction is $l\hbar$, where l is the angular momentum quantum number and may be any positive integer or zero. The total orbital angular momentum is equal to $\sqrt{l(l+1)}\hbar$. In this discussion the terms angular momentum or magnetic dipole moment describe the maximum component of these quantities along a field direction. *See* ANGULAR MOMENTUM; QUANTUM NUMBERS.

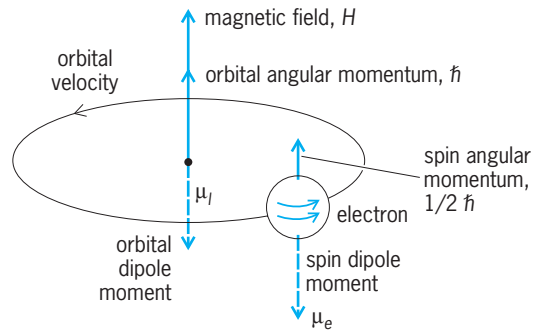

Fig. 1. Diagram of an electron in circular motion in a magnetic field, showing spin and orbital angular momenta and resulting magnetic moments.

Electron magnetic moment. The electron has a magnetic dipole moment by virtue of its spin. The approximate value of the dipole moment is the Bohr magneton μ_0 which is equal, in SI units, to $eh/4\pi m = 9.27 \times 10^{-24}$ joule/tesla, where e is the electron charge measured in coulombs, m is the mass of the electron, and c is the velocity of light. The orbital motion of the electron also gives rise to a magnetic dipole moment μ_l that is equal to μ_0 when $l = 1$. In **Fig. 1** is shown the simple case of an electron in circular motion in a magnetic field. The electron is shown with a spin angular momentum parallel to the orbital angular momentum. (Physically it could equally well be in the opposite direction.) Since the electron has a negative charge, the directions of the orbital dipole moment and spin dipole moment (both are vectors) are opposite to that of the orbital angular momentum. For a positron (a positively charged particle having the same mass and magnitude of charge as the negatively charged electron), the magnetic moments are positive, that is, in the same direction as the angular momentum. *See* MAGNETON; POSITRON.

The orbital magnetic moment of an electron can readily be deduced with the use of the classical statements of electromagnetic theory in quantum-mechanical theory; the simple classical analog of a current flowing in a loop of wire describes the magnetic effects of an electron moving in an orbit. The spin of an electron and the magnetic properties associated with it are, however, not possible to understand from a classical point of view. The classical radius of the electron is $e^2/(8\pi\epsilon_0 mc^2) = 1.141 \times 10^{-13}$ cm, where ϵ_0 is the permittivity of free space; a reasonable distribution of mass and electric charge within this radius which would lead to a magnetic moment μ_0 leads to the calculation of a peripheral velocity of the electron far greater than the velocity

of light. This is, of course, wholly precluded by the special theory of relativity. No theory of the structure of the electron has been formulated which makes the spin of the electron amenable to simple pictorial understanding. Nevertheless the interpretation of the spectra of atoms and molecules, the magnetic properties of materials, and other phenomena on an atomic scale unambiguously require that the electron have the property of spin. To the extent that physical theory assumes these properties and does not concern itself with questions about the structure of the electron, it is quite adequate to deal with a large range of physical phenomena in a highly quantitative way. *See* RELATIVITY.

In the Landé g factor, g is defined as the negative ratio of the magnetic moment, in units of μ_0, to the angular momentum, in units of $\hbar$. For the orbital motion of an electron, $g_l = 1$. For the spin of the electron the appropriate g value is $g_s \simeq 2$; that is, unit spin angular momentum produces twice the magnetic moment that unit orbital angular momentum produces.

The following discussion is limited to atoms which have a single electron outside of closed electron shells. Both the orbital and spin angular momenta of the electrons within closed shells add up in such a way that their net angular momentum is zero. The single electron outside closed shells may have $l = 0, 1, 2, \ldots$. By the usual rules developed from quantum mechanics the total angular momentum quantum number of the electron, which is called j, is $l \pm s$ (Fig. 1) except when $l = 0$, in which case the total angular momentum is s. For instance, when $l = 1$, $j = {}^1/_2$ or ${}^3/_2$. These relations may be represented by vector diagrams, as shown in **Fig. 2**. Since the revolution of the electron about the nucleus produces a magnetic field at the electron, and since the electron has a magnetic dipole moment, the energy of the atom is different when the vector s is parallel to $l(j = l + {}^1/_2)$ and when s is antiparallel (180°) to $l(j = l - {}^1/_2)$. Thus the spin of the electron causes a doubling of the energy levels in all single-electron atoms except when $l = 0$, in which case level is single. In the case of sodium, the familiar yellow lines (the D lines) comprise a closely spaced doublet that

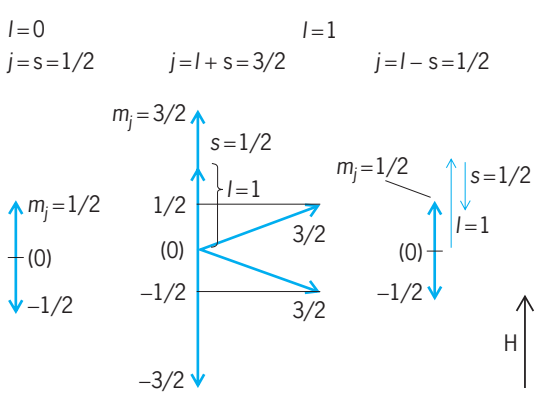

Fig. 2. Diagram illustrating addition of *l* and *s* vectors into a resultant *j* vector. The allowed values of m_j are also shown. In no case does the (0) give an allowed value. *H* = magnetic field.

arises from a transition from a state of $l = 1$ to one of $l = 0$. The doubling of the lines is a direct consequence of the existence of the electron spin.

Energy-level splitting. The total electronic magnetic moment of an atom depends on the state of coupling between the orbital and spin angular momenta of the electron (Fig. 2). In the single-electron case, an atom in the state for which $l = 0$ has only the magnetic moment associated with the spin, and this moment may be oriented in either of two directions with respect to an externally applied magnetic field. However, when the atom is in a state for which $l = 1$, j can be ${}^1/_2$ or ${}^3/_2$ and the Landé g factors for these two states are ${}^2/_3$ and ${}^4/_3$, respectively. That is, the magnetic moment per unit angular momentum is equal neither to that which characterizes the orbital motion nor to that which characterizes the spin motion. In a magnetic field, a single energy level characterized by l and j is split into several energy levels, each described by the component of j, m_j, along a magnetic field, where $m_j = j, j - 1, \ldots, -(j - 1), -j$. The energy of the level is the zero field energy plus the term $m_j g_j \mu_0 H$. In a transition between two energy levels that gives rise to a spectral line, the energy of the emitted or absorbed photon is the zero field energy difference between the two levels plus the difference between two magnetic energy terms as given previously. The resultant splitting of the line, which is single at zero magnetic field, into a line of two or more components is called the Zeeman effect. The Zeeman effect has been called anomalous when it is not explicable purely in terms of the orbital motion of the electron. The introduction of the electron spin allows the interpretation of all observed Zeeman effects. *See* ZEEMAN EFFECT.

Atomic beam measurements. Prior to 1940, spectroscopic measurements had been made only on optical spectra, and Zeeman effects were small effects superimposed on lines of considerable natural width. Exact measurements of the splitting of lines in a magnetic field, therefore, could not be made on such lines, and all data on the Zeeman effect were consistent with the statement that $g_s = 2g_l$. With the development of spectroscopy by the atomic beam method, a new order of precision in the measurement of the frequencies of spectral lines became possible. By using the atomic-beam techniques, it became possible to measure g_s/g_l directly, with the result $g_s/g_l = 2$ (1.001168 ± 0.000005). The magnetic moment of the electron therefore is not μ_0 but $1.001168\mu_0$, or equivalently the g factor of the electron departs from 2 by the so-called g factor anomaly defined as $a = (g_2 - 2)/2$ so that $\mu = (1 + a)_0$. Thus the first molecular beam work gave $a = 0.001168$. *See* MOLECULAR BEAMS.

Calculation of g-factor anomaly. It is not possible to give a qualitative description of the effects which give rise to the g-factor anomaly of the electron. The detailed theoretical calculation of the quantity is in the domain of quantum electrodynamics, and involves the interaction of the zero-point oscillation of the electromagnetic field with the electron.

Comparison of theoretical determination of a with its experimental measurement constitutes the most accurate and direct existing test of the theory of quantum electrodynamics.

Theoretical work on the g-factor anomaly, based on the principles of quantum electrodynamics, began simultaneously with its experimental discovery. The initial prediction of the value of the anomaly was made by J. Schwinger in 1948, who showed that the anomaly could be written as $a = 0.5(\alpha/\pi)$, where α, the fine-structure constant, is given by $\alpha = e^2/\hbar c \simeq 1/137$. Shortly thereafter, it was shown that a could be expressed as a power series in α (more customarily, α/π), that is, one can write $a = A(\alpha/\pi) + B(\alpha/\pi)^2 + C(\alpha/\pi)^3 + D(\alpha/\pi)^4 + \cdots$. Calculation of B, C, and D from quantum electrodynamics has proven to be very difficult, with errors occurring at various stages of the work. However, it is now generally agreed that $A = 0.500$, $B = -0.328478965\ldots$, $C = 1.181241456\ldots$, and $D = -1.7283 \pm 0.0035$. In order to find a theoretical value for a, an accurate value of α is required. There are several values of α, obtained from different sources, which may be used to determine a. The most accurate value of α is based upon results obtained by a measurement of the recoil velocity of rubidium atoms in an optical lattice and is given by $\alpha^{-1}{}_{\mathrm{Rb}} = 137.03599878 \pm 0.00000091$. The result of substituting this value of α into the expression for a gives the most recent theoretical value of a as a(theoretical) $= 0.0011596521887 \pm 0.0000000000077$. The uncertainty quoted for this number is due primarily to error in α. Finally, the weak and strong interactions have been calculated to add about 0.0000000000017 to a(theoretical). *See* ELECTRICAL UNITS AND STANDARDS; FUNDAMENTAL CONSTANTS; HALL EFFECT; QUANTUM ELECTRODYNAMICS; STRONG NUCLEAR INTERACTIONS; WEAK NUCLEAR INTERACTIONS.

Measurement of g-factor anomaly. The initial molecular-beam method for measuring a was based on measurement of g_s/g_l, which yields $1 + a$ rather than a itself. Such experiments have been superseded by two types of experiments which have the major advantages of measuring a directly, and of measuring a for electrons trapped in electric and magnetic fields but not bound to atoms. The second feature means that various corrections due to the presence of an atom are no longer necessary.

The new experiments may be readily understood, qualitatively, by noting that if an electron (spin angular momentum $\hbar/2$) moves at a low velocity $V(V/c \ll 1)$ perpendicular to a magnetic field B, the particle will rotate in the field at a frequency (the cyclotron frequency) given by $f_c = (^1/_2\pi)\,(eB/m)$, while its magnetic moment (and spin) will precess about the field at a frequency (the spin precession frequency) given by $f_s = (1 + a)f_c$. The difference between these frequencies, often called the $g - 2$ or difference frequency, is given by $f_D = f_s - f_c = af_c$. Thus f_D represents the rate at which the spin of the particle rotates relative to the particle's velocity.

One technique for measuring a consists of trapping very low-energy electrons in a magnetic field and measuring f_D and f_c simultaneously, with a being given by the relation $a = f_D/f_c$.

A slightly different version of the same technique consists of trapping electrons whose velocity is about half the speed of light (electrons accelerated to several hundred thousand volts) in a magnetic field and determining the same ratio. Relativistic effects manifest themselves in both the expressions for f_s and f_c, but cancel exactly when the ratio f_D/f_c is formed. The high-energy technique constitutes an excellent check of the separate measurements, as well as one of the most precise tests extant of the special theory of relativity, which predicts that the frequency ratio should be independent of velocity. *See* RELATIVISTIC ELECTRODYNAMICS.

The most accurate experimental result has been obtained by using the low-energy technique. This result is a(experimental) $= 0.00115965218085 \pm 0.00000000000076$, and it is in agreement, within the limits of error, with a(theoretical). The comparison of a(experimental) with a(theoretical) presented above constitutes the most precise confrontation of any experiment with a theoretical prediction in the history of science, and it provides the value of the fine-structure constant, $\alpha^{-1} = 137.035999710 \pm 0.000000096$, with a relative precision of 0.7 part per billion (10^9), which is an order of magnitude more precise than any other measurement available at present. *See* ATOMIC STRUCTURE AND SPECTRA; GYROMAGNETIC RATIO; QUANTUM MECHANICS.

Arthur Rich; Toichiro Kinoshita

Bibliography. R. Eisberg and R. Resnick, *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles*, 2d ed., 1985; G. Gabrielse et al., New determination of the fine structure constant from the electron g value and QED, *Phys. Rev. Lett.*, 97:030802, 2006; P. J. Mohr and B. N. Taylor, CODATA recommended values of the fundamental physical constants: 2002, *Rev. Mod. Phys.*, 77:1–107, 2005; B. Odom et al., New measurement of the electron magnetic moment using a one-electron quantum cyclotron, *Phys. Rev. Lett.*, 97:030801, 2006; H. C. Ohanian, *Modern Physics*, 2d ed., 1995; J. W. Rohlf, *Modern Physics from α to Z^0*, 1994.

Electron tomography

A powerful technique for determining three-dimensional (3D) structure from a series of two-dimensional (2D) images recorded in a transmission electron microscope. Transmission electron microscopy uses high-energy electrons to illuminate specimens at resolutions that in favorable cases surpass 1.0 nanometer (1/25,000,000 of an inch) and can resolve atoms. Resolution is not as great with electron tomography (4–8 nm) but is still much higher than can be achieved by light microscopy or medical imaging methods. However, the weak penetrating power of electrons requires operation under high vacuum and generally limits the technique to specimens less than about 1 micrometer in diameter. The most common application is

in biological research to study the 3D architecture of subcellular components such as mitochondria, nerve synapses, the Golgi apparatus, muscle cells during contraction, the arrangement of molecular components in bacterial cells, chromosome fibers, the kinetochore, and other structures. Such studies provide important data for understanding how cellular components carry out their vital functions. Electron tomography is also being used to study the structure of certain inorganic crystals used in the semiconductor industry.

Principles of operation. Transmission electron microscopy provides a translucent interior view of the specimen, analogous to what is seen in an x-ray image, rather than the surface view that we are accustomed to seeing in everyday life. Although this enables the microscopist to see inside the specimen, it has the disadvantage that the image formed is a 2D projection of the 3D object with features from different depths superimposed (**Fig. 1***a*). The solution is to reconstruct a 3D image of the original specimen from a series of 2D projections (Fig. 1*b*). This approach is also employed in medical imaging methods such as computerized axial tomography (CAT scanning) and magnetic resonance imaging (MRI).

The series of 2D images required for tomographic reconstructions is obtained by tilting the specimen in the electron beam. This is in contrast to medical imaging in which the patient is kept stationary while the beam source and detector are rotated to provide the necessary views. An important limitation of electron tomography is the inability to collect tilt data over the full 180° range required for complete angular coverage. (A 360° range is not necessary because, in projection imaging, views from the opposite side of the specimen provide the same information.) The difficulty in electron tomography is that typical electron microscope specimens are very thin (50–500 nm) and mounted on a support film that is more than 1000 times wider than the specimen. As a result, images can be recorded only over a limited tilt range of ±60 or 70° (a total range of 120–140°) before the specimen becomes too thick to view. This gives rise to a range of missing angular coverage that produces a well-characterized distortion that must be taken into account when interpreting electron tomographic reconstructions.

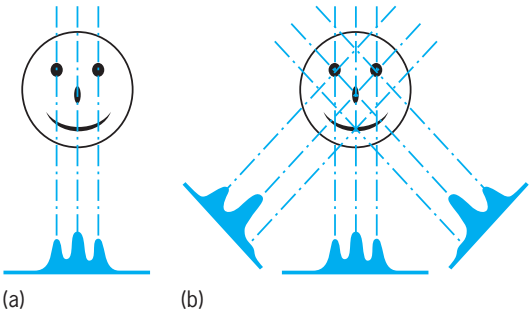

Fig. 1. Projection imaging and tomographic reconstruction. (*a*) Formation of one-dimensional (1D) projection of a two-dimensional (2D) smiley face. To form a true projection, the amount of penetrating radiation that strikes a unit length of the recording medium must be a measure of the amount of mass traversed by the beam in its path through the specimen. This is illustrated by the image density graph superimposed upon the 1D projection, and the broken lines that represent the radiation path through areas of maximum density in the specimen (that is, through the eyes or nose plus the mouth of the face). In transmission electron microscopy, true projections are formed as long as the specimen is not thicker than 0.2–0.5 micrometer. (*b*) Tomographic reconstruction of the 2D smiley face. The first step in tomography is to systematically record a series of 1D images from different viewing directions. The density in each 1D image is then projected back into a 2D area by simply distributing the density evenly along the original projection direction. When these "back-projection rays" from the different 1D images are summed, they intersect and reinforce one another only at the points where there was significant mass in the original object. In this way the 2D image is reconstructed from the 1D images. The same strategy is used to reconstruct a 3D image from 2D projection images. (*Adapted with permission from B. F. McEwen and M. Marko, Three-dimensional transmission electron microscopy and its application to mitosis research, Meth. Cell Biol., 61:82–111, 1999*)

Electron tomographic data sets are generally collected with 1 or 2° tilt angle intervals, because the resolution obtained in the resulting reconstruction is dependent upon how finely the angular range is sampled. Thus, a typical tilt series contains 60–140 images. Recording so many images was time-consuming and tedious with older microscopes, but recent improvements to mechanical design, along with incorporation of direct digital recording of the images and computer control of the microscope, have enabled newer microscopes to automatically record a tilt series of 140 images in less than 30 min. These technical developments in electron microscopy, coupled with the rapidly increasing computational power and the disk storage capacity of modern computers, are transforming electron tomography into a high-throughput tool that can be used ever more routinely in research and diagnostic settings.

Volume analysis and sample application. Perhaps the most surprising aspect of electron tomography is that it generally takes longer to analyze an electron tomographic reconstruction than it does to collect the data and compute the reconstruction. The difficulty is that the investigator is faced with displaying a three-dimensional tomographic reconstruction on a two-dimensional computer screen. However, in contrast to the original imaging situation, one now has a 3D image in the computer that can be taken apart and examined from any direction. Usually the quickest way to locate components of the structure is to view the reconstruction volume as a sequential stack of thin slices. However, understanding the 3D relationships of component parts and communicating the finding to a larger audience generally require constructing a 3D representation. For complex structures the most common approach is to segment select components away from the rest of the tomographic reconstruction by manual tracing and assignment of a color code.

Figure 2 illustrates the application of electron tomography to the study of the mammalian kinetochore. The kinetochore is a specialized structure that forms at a specific location on chromosomes during cell division. It functions to attach chromosomes to a bipolar spindle, generate chromosome motion, and exert control over the events of cell

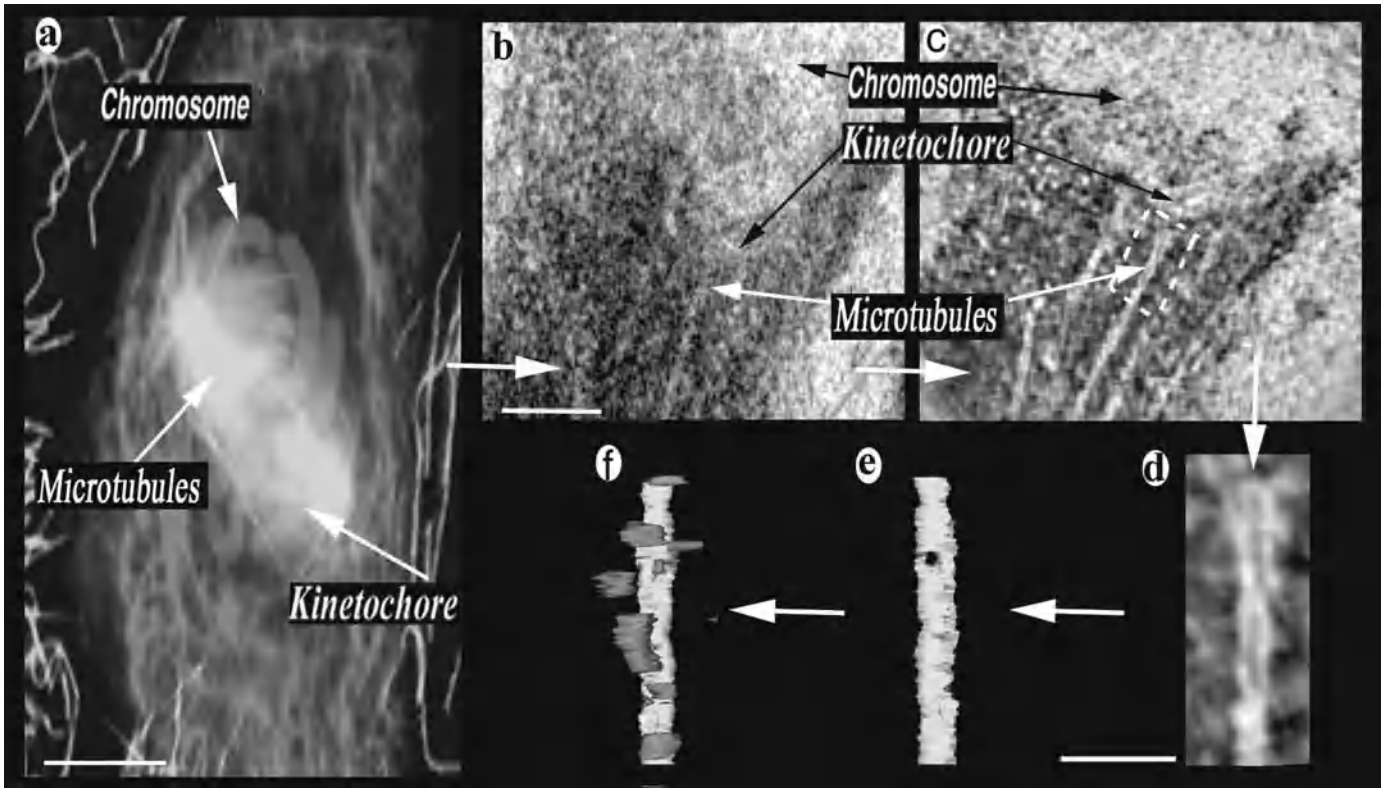

Fig. 2. Using electron tomography to analyze cell division. (*a*) An immunofluorescent light microscopy image of a mammalian cell during cell division. (*b*) Electron microscopy image of a kinetochore from the same type of cell. The kinetochore is the region of the chromosome that attaches to the mitotic spindle. (*c*) Single 2D slice from the 3D electron tomographic reconstruction of the kinetochore in *b*. Note the sharper microtubules in the tomographic reconstruction. (*d*) Single microtubule extracted from the tomographic reconstruction. (*e*) 3D model created by tracing and color-coding (light gray) the microtubule in *d*. (*f*) Same model as in *e* showing kinetochore components (dark gray) that are attached to the microtubule.

division. These functions are essential for proper cell division and the viability of the organism. Figure 2*a* is a gray-scale version of an immunofluorescence light microscopy image of a dividing cell, with arrows indicating a chromosome, location of a kinetochore, and microtubules. Microtubules are the chief structural components of the mitotic spindle. In the actual immunofluorescent image, microtubules are stained green, chromosomes blue, and keratin fibers, which form a cage around the spindle, red. One of the crucial yet poorly understood aspects of chromosome alignment is the interaction between the kinetochore and spindle microtubules. It is clear from Fig. 2*a* that the resolution of light microscopy is inadequate to study this interaction. Although single microtubules are visible in an electron microscopy projection image (Fig. 2*b*), details of the microtubule ends become visible only in single slices from the tomographic reconstruction (Fig. 2*c, d*). Manual tracings from trimmed slices such as in Fig. 2*d* were stacked into a 3D image, which are shown as a surface view in Fig. 2*e*. Kinetochore connections to the microtubules were also traced, and are shown in dark gray along with the microtubule in Fig. 2*f*. These and similar 3D models are being used to test current hypotheses on how the kinetochore interacts with microtubule ends.

Future directions and conclusions. Recent technical advances and the inherent versatility of electron tomography have accelerated its use and expanded the range of its applications. In particular, electron tomography is now being used for high-quality frozen-hydrated preparations of biological specimens. These preparations are made by rapidly freezing the specimen to prevent the formation of normal crystalline ice, which severely damages cellular structure. Such preparations are a snapshot of the specimen in its native hydrated environment and, thereby, superior to conventional preparations made by solvent extraction, heavy-metal staining, and plastic embedding. Previously the sensitivity of frozen-hydrated specimens to damage from electron irradiation had precluded such applications.

Electron tomography has been established as an important method for determining the structure of subcellular components and organelles, and the number and range of applications are likely to continue growing rapidly in the near future. *See* COMPUTERIZED TOMOGRAPHY; ELECTRON MICROSCOPE; FLUORESCENCE MICROSCOPE; MEDICAL IMAGING; OPTICAL MICROSCOPE.

Bruce F. McEwen

Bibliography. W. Baumeister, R. Grimm, and J. Walz, Electron tomography of molecules and cells, *Trends Cell Biol.*, 9:81–85, 1999; B. F. McEwen and M. Marko, The emergence of electron tomography as an important tool for investigating cellular ultrastructure, *J. Histochem. Cytochem.*, 49:553–563, 2001; B. F. McEwen and M. Marko, Three-dimensional transmission electron microscopy and its application to mitosis research, *Meth. Cell Biol.*, 61:82–111, 1999.

Electron-transfer reaction

A reaction in which one electron is transferred from one molecule or ion to another molecule or ion. Electron-transfer reactions are ubiquitous in nature. Some are deceptively simple [for example, reaction (1), where the asterisk is used to identify a

$$^{*}Fe(H_2O)_6^{3+} + Fe(H_2O)_6^{2+} \longrightarrow {}^{*}Fe(H_2O)_6^{2+} + Fe(H_2O)_6^{3+} \quad (1)$$

specific isotope]; others look very complicated (for example, the long-range electron transfers found in biology). The widespread occurrence of electron-transfer reactions has stimulated much theoretical and experimental work. *See* BIOLOGICAL OXIDATION; CHEMIOSMOSIS; PHOTOSYNTHESIS.

The development of electron-transfer theory, as well as the measurement of the rates of oxidation-reduction (redox) reactions, began in the late 1940s. A great deal of experimental work has been carried out by inorganic chemists, and during the 1960s and 1970s the reactivity patterns of a large number of complexes were uncovered. Using this body of knowledge as a reference point, bioinorganic chemists studying the rates and mechanisms of biological electron-transfer reactions are seeking parallels with the redox behavior of less complicated metal complexes. *See* BIOINORGANIC CHEMISTRY; COORDINATION COMPLEXES.

Metalloproteins are more than just metal ions in disguise. For example, the redox behavior as well as virtually every property of a protein (excluding its amino acid sequence) are dependent upon the solution pH; redox proteins are very large polyelectrolytes whose prosthetic groups are typically found buried in the protein interior. Thus pH-dependent behavior in a metalloprotein may be much more complex than in a simple metal complex. Another important distinction between the redox reactions of proteins and the electron transfers of small metal complexes is the magnitude of the electron donor-to-acceptor distance. The relevant distance for small molecules is generally taken to be van der Waals contact, which is roughly 0.3 nanometer. In contrast, several experiments have shown that electrons can jump at significant rates across distances of 1 nm or more in protein interiors. *See* INTERMOLECULAR FORCES; PH; PROTEIN.

Experimental investigation of the factors that control the rates of biological redox reactions has not come as far as the study of the electron transfers of metal complexes, because of the large number of variables that must be dealt with, and problems associated with protein purification and stability. Many experimental approaches have been pursued, including the covalent attachment of redox reagents to the surfaces of metalloproteins and the use of mutant proteins (obtained by gene cloning). *See* BIOELECTRONICS.

Self-exchange and cross reactions. The simplest reactions in solution chemistry are electron self-exchange reaction (2), in which the reactants and

$$^{*}A_{ox} + A_{red} \rightarrow {}^{*}A_{red} + A_{ox} \quad (2)$$

products are the same (the asterisk is used to identify a specific isotope). The only way to determine chemically that a reaction has taken place is to introduce an isotopic label. There is no change in the free energy ($\Delta G^{\circ} = 0$) for this type of reaction.

Much more common are cross reaction (3), where

$$A_{ox} + B_{red} \rightarrow A_{red} + B_{ox} \quad (3)$$

A_{ox} is the oxidized reactant, B_{red} is the reduced reactant, A_{red} is the reduced product, and B_{ox} is the oxidized product. For these reactions, $\Delta G^{\circ} \neq 0$. The experimental determination of cross-reaction rates is generally more straightforward than the determination of self-exchange rates. Either the reactants are simply mixed together or a thermodynamically unstable system is generated rapidly (via pulse radiolysis, flash photolysis, or temperature-jump relaxation) to initiate the redox reaction. In almost all cases, electronic absorption spectroscopy has been used to monitor the progress of protein cross reactions. *See* SPECTROSCOPY.

Inner-sphere and outer-sphere mechanisms. Both types of electron-transfer reactions (self-exchange and cross reactions) can be classified broadly as inner sphere or outer sphere. In an inner-sphere reaction, a ligand is shared between the oxidant and reductant in the transition state. An outer-sphere reaction, on the other hand, is one in which the inner coordination shells of both the oxidant and reductant remain intact in the transition state. There is no bond breaking or bond making, and no shared ligands between redox centers. Long-range electron transfers in biology are all of the outer-sphere type.

Electron-transfer theory. The simplest electron transfer occurs in an outer-sphere reaction. The changes in oxidation states of the donor and acceptor centers result in a change in their equilibrium nuclear configurations. This process involves geometrical changes, the magnitudes of which vary from system to system. In addition, changes in the interactions of the donor and acceptor with the surrounding solvent molecules will occur. The Franck-Condon principle governs the coupling of the electron transfer to these changes in nuclear geometry: during an electronic transition, the electronic motion is so rapid that the nuclei (including metal ligands and solvent molecules) do not have time to move. Hence, electron transfer occurs at a fixed nuclear configuration. In a self-exchange reaction, the energies of the donor and acceptor orbitals (hence, the bond lengths and bond angles of the donor and acceptor) must be the same before efficient electron transfer can take place. *See* FRANCK-CONDON PRINCIPLE.

The incorporation of the Franck-Condon restriction leads to the partitioning of an electron-transfer reaction into reactant (precursor complex) and product (successor complex) configurations.

Steps (4)–(6) go from reactants to products: K is the

$$A_{ox} + B_{red} \rightleftharpoons [A_{ox},\ B_{red}] \quad (4)$$

$$[A_{ox}, B_{red}] \xrightarrow{k_{et}} [A_{red}, B_{ox}] \quad (5)$$

$$[A_{red}, B_{ox}] \xrightarrow{\text{fast}} A_{red} + B_{ox} \quad (6)$$

equilibrium constant for the formation of the precursor complex [:A_{ox}, B_{red}], and k_{et} is the forward electron-transfer rate to produce the successor complex [A_{red}, B_{ox}].

Some degree of electronic interaction, or coupling between donor and acceptor, is required if the redox system is to pass from the precursor to the successor complex. (This passage involves an intermediate species known as a transition state or activated complex.) The magnitude of the electronic coupling determines the behavior of the reactants once the transition state is reached. Two cases can be distinguished. If the electronic coupling is very small, as in so-called nonadiabatic reactions, there is a very low probability that the reactants will undergo electron transfer, leading to very little product formation. If the electronic interaction is sufficiently large, as is the case for adiabatic reactions, the probability of electron transfer occurring when the reactants reach the transition state is unity. The degree of adiabaticity of a reaction is characterized by a transmission coefficient κ, whose value ranges from 0 to 1. For systems in which the electronic coupling is sufficiently large, κ equals 1. This situation occurs when the reacting centers are close together, the donor and acceptor orbitals overlap, and no substantial changes in geometry are involved. The transmission coefficient is much less than 1 for electron transfer over long donor-to-acceptor distances, as commonly found in protein electron-transfer reactions and in some electron transfers in large organic molecules.

Long-range electron transfer. The rate of long-range electron transfer between an electron donor (B_{red} and an electron acceptor (A_{ox}) depends on both the electronic coupling between A_{ox} and B_{red} (which is a function of the intersite $A_{ox}//B_{red}$ distance d, the nature of the intervening medium, and the relative $A_{ox}//B_{red}$ orientation, where // represents the protein medium that separates the donor and the acceptor) and an activation energy term. A standard theoretical rate equation (7) expresses k_{et} in terms of these

$$k_{et} = \nu\ [\exp(-\beta d)]\{\exp[-(\lambda + \Delta G^{\circ})^2/4\lambda RT]\} \quad (7)$$

factors; here ν is a frequency factor, β is a medium- and orientation-dependent quantity, d is the intersite λ is the reorganization energy, ΔG° is the reaction free energy of the electron-transfer process, R is the universal gas constant, and T is the absolute temperature of the system. Experiments in several laboratories have been designed to estimate the values of λ and β in modified metalloproteins, rigid organic molecules, and protein-protein complexes.

Modified metalloproteins. Long-range intramolecular electron transfer is being studied in synthetically modified metalloproteins. Ruthenium (Ru) ammines have been chosen as modification agents primarily because they can be attached specifically to surface histidines, and the resulting complexes are kinetically stable in solution in the relevant oxidation states. X-ray crystallographic structure determinations have been completed for the native proteins used in these studies, and so the distance and intervening medium between the donor and acceptor sites in the modified protein are accurately known. [X-ray crystallographic studies of a modified protein show that only minimal structural perturbations result from covalent attachment of ruthenium to a surface histidine (His).] The driving force ΔG° for the electron transfer between donor and acceptor of a modified protein can be determined accurately through standard electrochemical experiments. Thus electron transfer can be studied in a system where two of the important variables in Eq. (7), d and δG°, are reliably known. *See* X-RAY CRYSTALLOGRAPHY.

Sperm whale myoglobin (Mb) has four surface histidines that can be modified by using ruthenium ammine complexes. The pentaammine-ruthenium(III) [a_5Ru] complex has been covalently attached to these histidines. After separation and purification of the reaction products, four singly labeled myoglobin species are obtained, with the closest edge-to-edge distances from the heme ranging from 1.27 nm (His-48) to 2.2 nm (His-12). Electron transfer from the His-48 ruthenium complex to the iron (Fe)–containing heme has been measured by selectively reducing the Ru center on the exterior of the protein using flash photolysis techniques. This is followed by electron transfer from Ru^{2+} to the heme Fe^{3+} in the interior of the protein. The forward electron-transfer rate k_{et} (Ru^{2+} to Fe^{3+}) in a_5Ru-(histidine-48)-modified myoglobin [a_5Ru(48)MbFe] was found to be 0.02 s^{-1}; the reverse rate k_r (Fe^{2+} to Ru^{3+}) was uniquely determined to be 0.04 s^{-1} by employing a method developed to generate the $Fe^{2+}//Ru^{3+}$ mixed-valence species. These kinetic results clearly demonstrate that long-range electron transfer is reversible in this system.

Cytochrome *c* (from horse heart) also has been modified with the a_5Ru reagent. In contrast to the myoglobin system, cytochrome *c* has only one surface histidine (His-33) that readily binds ruthenium. The edge-to-edge distance from this histidine to the heme is 1.16 nm. By using flash photolysis, the Ru^{2+}-to-Fe^{3+} electron-transfer rate was found to be approximately 30 s^{-1}. Pulse radiolysis has also been employed on the same system, with a measured electron-transfer rate of 50 s^{-1}. *See* CYTOCHROME.

Driving-force dependences have been addressed in two ways with modified proteins. First, the ammine ligands on the ruthenium label have been varied to produce complexes with different reduction potentials. Second, the iron from the heme has been removed and replaced with zinc, magnesium, and palladium. After photoexcitation, the metalloporphyrin functions as an electron donor, leading to electron transfer to the surface-bonded ruthenium.

With these methods, a wide range of driving forces for intramolecular electron transfer in cytochrome *c* and myoglobin has been investigated. At the highest driving forces in cytochrome *c*, rates of roughly 3×10^6 s^{-1} were found. Values of the reorganization energy λ between 1.3 and 2 eV have been obtained from these experiments.

The distance dependence of k_{et} has been investigated in the a_5RuMbZn system. Values of β in the 8–10-nm^{-1} range have been extracted from these studies, thereby indicating that a protein is a good medium for electronic coupling of a donor and acceptor.

Rigid organic molecules. Long-range electron-transfer reactions have been investigated in molecules containing organic donors and acceptors separated by rigid hydrocarbon units. With hydrocarbon spacers of 1–1.2 nm, electron-transfer rates as high as 10^{10} s^{-1} have been observed. The rates in the organic molecules are 10,000 times faster than protein electron transfers at comparable donor-to-acceptor distances. In one study involving hydrocarbon spacers of lengths 0.7–1.4 nm, β was found to be 8.5 nm^{-1}.

Long-range electron-transfer reactions in rigid organic molecules have been studied only in organic solvents; electron-transfer theory suggests that the reorganization energy λ should be smaller in nonaqueous media than reorganization energies for comparable long-range electron transfers in aqueous solutions. Results obtained are consistent with theory; a reorganization energy of roughly 1 eV was found in a thorough investigation of driving-force effects on the rate of long-range electron transfer in a donor (hydrocarbon spacer)–acceptor system.

Protein-protein complexes. In biologically relevant partners, both redox centers are buried in a protein matrix; the entire electron transfer occurs through a protein medium and includes a protein–protein interface. The reorganization energy λ has been measured by determining both the temperature dependence and the free-energy (ΔG°) dependence of electron-transfer rates in these systems. These protein–protein complexes also are well suited for the study of the effects of the medium on biological electron-transfer rates. Two approaches have been used. One is to compare electron-transfer rates between homologous (same species) and heterologous (different species) protein redox pairs. The other is to use the technique of site-directed mutagenesis to change amino acids that lie between the two redox centers in the protein–protein complex.

The most commonly used experimental system involves metal substitution of zinc for iron in one of the porphyrin redox centers in the protein–protein complex. This system allows facile initiation of electron transfer through photoexcitation of the zinc porphyrin (ZnP). The excited zinc porphyrin (ZnP^*) may then decay to the ground state or transfer an electron: the electron-transfer reaction is $Fe^{3+}//ZnP^* \rightarrow Fe^{2+}//ZnP^+$. The Fe^{2+} in $Fe^{2+}//ZnP^+$ will transfer an electron back to ZnP^+, thereby regenerating the starting material. The forward rate ($ZnP^* \rightarrow Fe^{3+}$) is k_{et}, and the back rate ($Fe^{2+} \rightarrow ZnP^+$) is k_b. *See* PORPHYRIN.

Hybrid hemoglobins. Hybrid hemoglobin is the best characterized of the protein–protein complexes. In this system, the donor-acceptor separation distance and intervening medium are known from x-ray structure determinations. Both α(Zn)β(Fe) and α(Fe)β(Zn) hybrids have been studied. Electron transfer ($ZnP^* \rightarrow Fe^{3+}$) between the α_1 and β_2 subunits of hemoglobin has been observed: the metal-metal distance is 2.5 nm (edge-to-edge porphyrin distance is 2.0 nm). The ΔG° for the electron transfer is approximately 0.8 V, and at room temperature k_{et} is approximately 100 s^{-1}. The temperature dependence of k_{et} has been analyzed by both quantum-mechanical and semiclassical theories; values of λ in the 2–3-eV range have been obtained. *See* HEMOGLOBIN.

Effects of the medium. The effects of the medium on long-range electron transfer have been investigated in the cytochrome *c* peroxidase (CCP)//cytochrome *c* (cyt *c*) complex. By using cyt *c* Fe^{3+}//CCPZn, ΔG° for back electron transfer is approximately 0.85 V. When yeast cytochrome *c* peroxidase was used in this reaction, k_{et} for the complex with yeast cytochrome *c* was found to be 10-fold faster than that for complexes with tuna cytochrome *c* or horse cytochrome *c*. The rate k_b was found to be 1000-fold faster for yeast than for tuna cytochrome *c*. These data indicate that specific medium and interface contact effects may enhance the reaction in the homologous yeast pair. The fact that the homologous and heterologous pairs show such radically different rate enhancements for k_{et} and k_b suggests a conformational change after the forward electron transfer. *See* CONFORMATIONAL ANALYSIS.

By using site-directed mutagenesis, phenylalanine-87, which is located above the heme pocket, has been replaced by tyrosine and serine in yeast cytochrome *c*. Electron-transfer studies with CCPZn have demonstrated that k_{et} is not affected by these mutations. The back electron transfer, on the other hand, is unaffected for the tyrosine mutant but is decreased 10,000-fold for the serine mutant. These data suggest that aromatic amino acids act as mediators for electron transfer. Although reorganization energy effects caused by the mutation to serine cannot be ruled out, these experiments demonstrate that changes in the intervening protein medium can produce dramatic electron-transfer effects.

Driving-force dependence of ket. The ΔG° dependence of the electron-transfer rate has also been studied in cyt *c*//CCP by using a series of substituted horse cytochromes *c*. These experiments yielded an estimate of 1.5–2.0 eV for the reorganization energy λ. The λ values for this protein–protein complex and for hybrid hemoglobin are in reasonable agreement with those found for a_5Ru-modified proteins. The origin of the relatively large reorganization energies for protein electron transfers in aqueous solution is being actively investigated in several laboratories. *See* CHEMICAL THERMODYNAMICS; OXIDATION-REDUCTION.

Bruce E. Bowler; Walther R. Ellis, Jr.; Harry B. Gray;Thomas J. Meade

Bibliography. J. R. Bolton, N. Mataga, and G. M. McClendon (eds.), *Electron Transfer in Inorganic, Organic, and Biological Systems*, 1991; G. L. Closs and J. R. Miller, Intramolecular long-distance electron transfer in organic molecules, *Science*, 240:440–447, 1988; J. Mattay (ed.), *Electron Transfer I*, 1994; S. L. Mayo et al., Long-range electron transfer in heme proteins, *Science*, 233:948–952, 1986; R. A. Scott, A. G. Mauk, and H. B. Gray, Experimental approaches to studying biological electron transfer, *J. Chem. Educ.*, 62:932–938, 1985.

Electron tube

A device in which electrons can travel through a sealed chamber containing at least two electrodes and gas at a very low pressure. The gas pressure usually ranges from about 10^{-6} to 10^{-9} atm (10^{-1} to 10^{-4} pascal). At the low extreme of this pressure range, electron tubes are sometimes referred to as vacuum tubes, and at the high extreme as gas tubes. *See* GAS TUBE; VACUUM TUBE.

Electron emission. At least one of the electrodes must emit electrons, and at least one must collect electrons. The emitting electrode, the cathode, may emit electrons through one or more of four mechanisms: thermionic or primary emission, secondary emission, photoelectric emission, or field emission. Electrons must acquire more energy than they have in the conduction band of a metal in order to escape from the surface of a metal. They acquire this energy, respectively, in the four mechanisms listed above, from heat, electron or ion impact, a photon impact, or an external electric field. Photoelectric emission, first observed by H. Hertz (1887), is used in light-sensing devices, often in combination with secondary electron multiplication to amplify the current. Secondary emission, sometimes in combination with thermionic emission, plays an important role in magnetrons invented by A. W. Hull (1921) and in crossed-field amplifiers. Field emission is used in some experimental amplifiers, flat-panel display devices, and x-ray tubes, but by far the most common type of emitting electrode used in electron tubes, including cathode-ray tubes, is the thermionic cathode. *See* CATHODE-RAY TUBE; FIELD EMISSION; FLAT-PANEL DISPLAY DEVICE; MAGNETRON; PHOTOEMISSION; PHOTOMULTIPLIER; PHOTOTUBE; SECONDARY EMISSION; X-RAY TUBE.

Thermionic emission was first observed by T. A. Edison (1883) as "negative electricity" emitted from an incandescent filament (the Edison effect). J. J. Thompson (1897) identified the particle emitted from the heated filament as one of very small mass (an electron). O. W. Richardson (1902), M. von Laue (1918), and S. Dushman (1923) derived equations that describe the phenomenon. Pure metals such as tungsten and tantalum can reach temperatures sufficient to produce large currents (about 1 A/cm^2) through thermionic emission. Monolayers of elements such as thorium can reduce the filament temperature required for thermionic emission to about 1650°C (3000°F) by reducing the electron energy required of an electron to escape the surface (the work function). Tungsten or nickel coated with mixtures of barium, calcium, and strontium oxides discovered by A. Wehnelt (1903) or porous tungsten matrices impregnated with barium-calcium aluminate (the Philips cathode) can emit large currents at even lower temperatures (800–1150°C or 1500–2100°F). *See* THERMIONIC EMISSION.

Diodes. A diode is a two-electrode tube, with a cathode and a collecting electrode. Edison must have built a diode in order to discover thermionic emission. A. Fleming (1904) developed the first thermionic diode using an oxide cathode. Because the collecting electrode is usually operated at a positive potential with respect to the cathode in order to collect much of the available electron current from the cathode, it is called an anode. Even so, because of the thermal energy of thermionic electrons, the anode can collect some electrons when it has a slightly negative potential. In a diode, with a good vacuum, the electric field of the electrons in the space between the cathode and the anode just cancels the electric field produced by the anode at the cathode surface. Thus the number of electrons in this space must increase in proportion to anode voltage while their kinetic energy also increases in proportion to the voltage. The electron velocity is proportional to the square root of the kinetic energy, and because current is proportional to density times velocity, the current is proportional to the 3/2 power of the voltage. This is referred to as the space-charge-limited current. The proportionality constant is an invariant of the diode design and is known as the perveance. In gas diodes, ion neutralization of the electron space charge permits the maximum cathode current (determined by the supply voltage and the load resistance) to flow even at very low anode voltages. *See* CHILD-LANGMUIR LAW.

Diodes were used in great numbers as rectifiers until about 1950 or 1960. By that time, most had been replaced by silicon rectifiers, often used in series, to replace even very high voltage vacuum rectifiers. I. Langmuir (1923) and K. Blodgett (1924) calculated the space-charge-limited current and the potentials between electrodes in diodes with circular cylindrical and spherical electrodes. This theory remains important because it is used as a starting point in the design of highly convergent electron guns for very high power microwave tubes. *See* DIODE; SEMICONDUCTOR RECTIFIER.

Triodes. L. DeForest (1906) added a third electrode to a diode in order to control the current flow from cathode to anode. This third electrode, the grid, took the form of a fairly open array or mesh made of wires with a diameter small compared to their spacing. In this geometry, much of the electric field from the anode terminates on the grid, and the field from the grid that terminates on the cathode exerts a primary influence on the space-charge current that flows to or through the grid. When the grid is at a negative potential with respect to the cathode, current flows due to the anode field that leaks through the grid, but

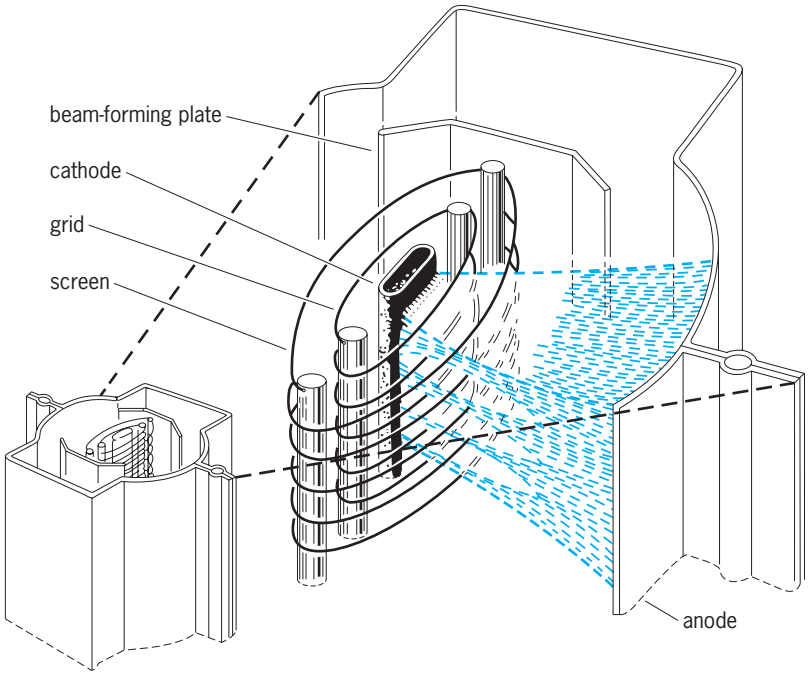

Fig. 1. Cutaway view of electrode arrangement in a beam-power tetrode. (*General Electric*)

the grid can collect no current. When the grid and anode are both positive, much more current flows and divides between the grid and anode.

By 1915, triode amplifiers and oscillators dominated the development of communications equipment. Gas-filled triodes, often called thyratrons and invented by Hull (1929), like gas-filled diodes, conduct much more current with much lower voltage drop than high-vacuum triodes because of space-charge neutralization by positive ions. The grid can start current flow and ionization, but cannot stop current flow until the ions are removed by reversing the polarity of the anode voltage.

Tetrodes and pentodes. Unfortunately, in triode amplifiers at high frequencies, the capacitance between the anode and grid electrodes, in combination with typical grid circuit reactances, can cause positive feedback, regeneration, or oscillation unless circuits that provide compensating negative feedback are used. For this reason W. Schottky (1919) invented the tetrode, which has a second or screen grid between the first or control grid and the anode. This grid was operated at a constant positive potential and effectively shielded the control grid from the anode. At large signal levels, it also created a problem by collecting secondary electrons emitted from the anode as a result of primary electron impacts when the instantaneous voltage on the anode was less than the screen grid voltage.

This problem was dealt with in two ways. G. Jobst and D. H. Tellegen (1926) introduced the pentode, which has a third very open suppressor grid between the screen grid and the anode. It was connected to the cathode. This created an electric field which returned secondary electrons to the anode. A more elegant solution to the secondary electron problem was provided in the beam-power tetrode (**Fig. 1**). In these tetrodes the anode was placed far enough from the screen grid that the charge of the electrons traveling between the screen grid and anode actually depressed the potential in the space between the screen and anode enough to return secondary electrons to the anode. This solution was analyzed by C. E. Fay, A. L. Samuel, and W. Schockley (1938). Beam-power tetrodes were popular in high-power transmitters and audio amplifiers because, in such applications, screen-grid-current interception and consequent power dissipation are problems. They are still popular in amateur radio equipment and high-power amplifiers used by musicians.

Replacement with solid-state devices. During 1950–1970, advances in solid-state diode and transistor technology, together with the development of small- and large-scale integrated circuits, resulted in the replacement of most low-power tubes with solid-state devices. There was really no contest because solid-state devices operate at much lower power input and at a much lower impedance level (lower voltage and high current) and hence provide higher bandwidth as well as efficiency. *See* INTEGRATED CIRCUITS; SEMICONDUCTOR DIODE; TRANSISTOR.

High-power tubes. However, for applications in which power serves a useful purpose (for example, transmitters), particularly at high frequencies, the outcome has been quite different. Because no electronic amplifier is 100% efficient, a high-power amplifier must dissipate power as heat. If the amplifier can run hot, it will be small and lightweight. If the temperature is limited to a low value, as it is when using most transistors, the amplifier will be large and heavy. Most of this size and weight will be in the metal heat sinks attached to the transistors.

The tubes and transistors discussed so far act as valves that control the flow of a current to a load. The potential energy of the current is derived from a direct-current power source. There is another, rather different class of electron tubes, most of which are referred to as microwave tubes, in which electrons are accelerated to a velocity at which they have a kinetic energy that is equivalent to the full voltage of the power supply that was used to accelerate them. If these electrons are bunched periodically in time, they can be made to give up their energy to the electric field in a gap or gaps in a very high frequency or microwave circuit. Microwave tubes include the inductive output tube, the klystron, traveling-wave tubes, crossed-field devices, and cyclotron-resonance devices.

There are no solid-state analogs to such tubes because the current flow in solid-state devices is too collisional. In electron tubes the electron density is very low, about 10^9–10^{11} electrons per cubic centimeter. At these densities, electrons do not collide with one another. Instead they move coherently and acquire a kinetic energy equal to the time integral of the scalar product of the electric field force and the velocity. As a result, in higher-voltage devices these electrons travel very rapidly (0.3 the velocity of light at 30 kV). Gap dimensions can be

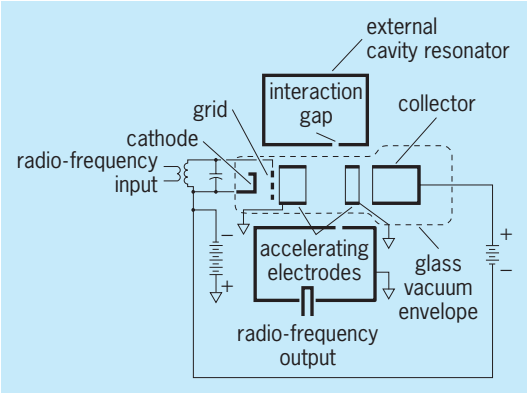

Fig. 2. Inductive output amplifier, showing separation of electron energy extraction from electron collection. (*After R. S. Symons, Tubes: Still vital after all these years, IEEE Spectrum, 35(4):52–63, April 1998*)

large for small transit-angle gaps. In a semiconductor, there are about 10^{19} carriers per cubic centimeter which have a great deal of thermal motion upon which is superimposed a low drift velocity proportional to the electric field. The carriers frequently collide with, and give up their energy to, the crystal lattice.

Inductive output tubes. In the inductive output tube (**Fig. 2**), invented by A. V. Haeff (1939), an electron beam is amplitude-modulated with a grid and then accelerated through a hole in the first accelerating electrode to form the high-velocity beam of electrons that passes through a gap in the center conductor of the coaxial external cavity resonator and into the collector. Inductive output tubes are used in many television transmitters operating between 470 and 900 MHz.

Klystrons. The klystron (**Fig. 3**), invented by R. Varian and S. Varian (1939), has a similar output cavity and collector, but has a beam which is first accelerated in a diode electron gun and then velocity-modulated in another reentrant cavity gap. Fast electrons overtake slowed electrons and yield an intensity-modulated beam by the time the electrons reach the output cavity. Additional cavities may be interposed between the input and output cavities (Fig. 3) to provide very high gain (often as high as 60 dB). *See* KLYSTRON.

Traveling-wave tubes. In traveling-wave tubes (**Fig. 4**), invented by R. Kompfner (1946), a high-velocity electron beam is velocity-modulated by, and gives up its energy to, periodically loaded or helical waveguides which slow the electromagnetic wave to a velocity nearly equal to that of the electron beam. Again very high gain is possible. *See* TRAVELING-WAVE TUBE.

Properties of linear-beam tubes. In all these linear-beam tubes, the electron beams are usually formed by highly convergent electron guns designed using techniques first devised by J. R. Pierce (1940). Computer simulation is now used to refine such designs. Steady magnetic fields provided by either electromagnets or permanent magnets are used to guide the electrons through the radio-frequency circuits. When the electrons leave these focusing fields, they repel each other due to their electric charge, and in addition, they cross radial magnetic field lines which also deflect them outward. The electron collectors on linear-beam tubes may have collection areas as much

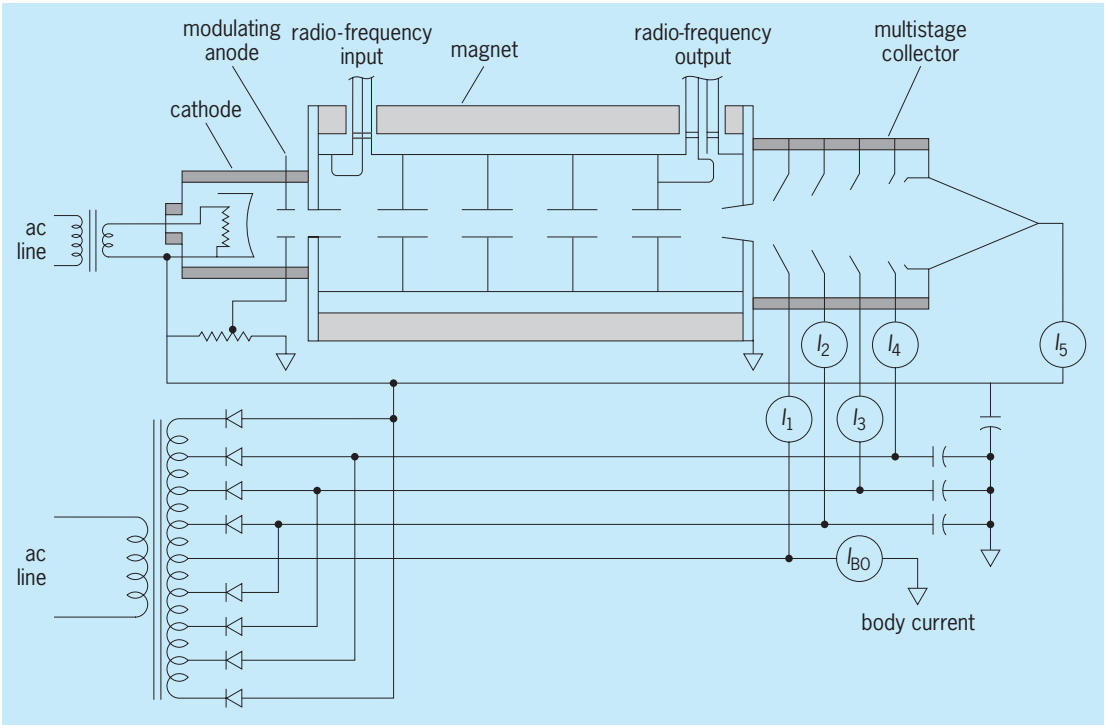

Fig. 3. High-efficiency, depressed-collector-klystron amplifier. (*After D. Christiansen, ed., Electronic Engineers' Handbook, 4th ed., McGraw-Hill, 1996*)

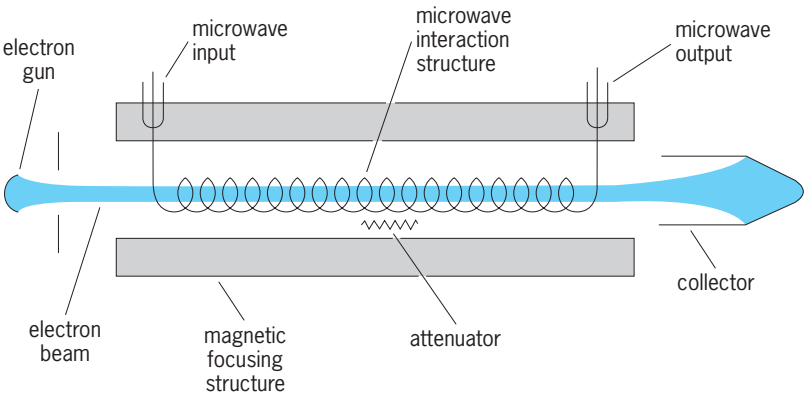

Fig. 4. Basic elements of a typical traveling-wave tube. (*After D. Christiansen, ed., Electronic Engineers' Handbook, 4th ed., McGraw-Hill, 1996*)

as 10,000 times the cross-sectional area of the electron beam, and continuous power as high as 1 MW has been achieved at frequencies as high as 10,000 MHz from such tubes. On klystrons and traveling-wave tubes, sometimes several collectors at different potentials are used (Fig. 3). These multistage depressed collectors, suggested by C. V. Litton (1940), can sort electrons by energy. Electrons that gave only small amounts of energy to the electric field in the radio-frequency circuit are collected on lower-potential electrodes. Tubes using these collectors are more efficient. R. Symons (1993) realized that an input-output tube with a multistage depressed collector could have the unique property of providing almost constant efficiency over a large range of signal amplitudes because both the current and the average collection voltage of the electrons could be made to go up and down with signal amplitude. Inductive-output tubes, klystrons, and traveling-wave tubes are used in television broadcasting, satellite communications systems, radar, scientific accelerators, medical accelerators used for cancer therapy, and military countermeasures equipment.

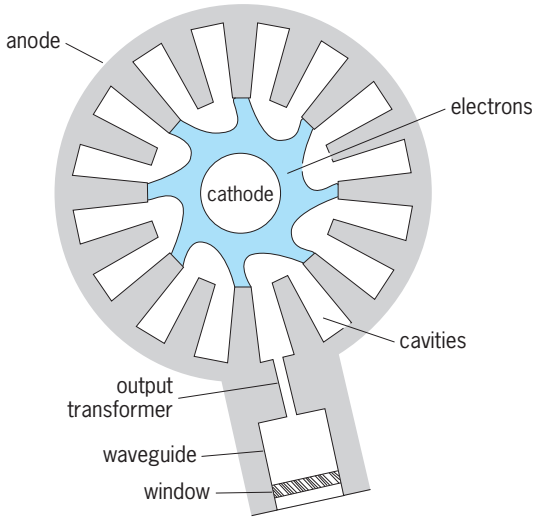

Fig. 5. Conventional microwave structure. (*After D. Christiansen, ed., Electronic Engineers' Handbook, 4th ed., McGraw-Hill, 1996*)

Magnetrons. In magnetrons (**Fig. 5**) and crossed-field amplifiers, electrons circulate about a cylindrical cathode in a radial static electric field and an axial magnetic field. Concentric with, and outside, the cathode is a periodically loaded transmission line that propagates a wave having circumferential electric field components that travel in synchronism with the rotating electron cloud. The electrons follow orbits that allow them to take energy from the radial static electric field and transfer it to the circumferential radio-frequency electric field of the wave on the circuit. Magnetrons are used in huge quantities in household microwave ovens. They and crossed-field amplifiers are also used in ground-based, shipboard, and airborne radars.

Cyclotron-resonance devices. Cyclotron-resonance devices including gyrotrons, gyroklystrons, and gyro-traveling-wave tubes again employ electrons that have been accelerated to the full energy provided by the electrical power supply. The beam is formed in a magnetic field so that it has a great deal of momentum perpendicular to the magnetic field, and the electrons follow helical paths. A radio-frequency electric field perpendicular to the axis of the electron trajectories will modulate the energy of the electrons and hence the relativistic mass and the cyclotron frequency. This azimuthal velocity modulation causes the electrons to draw into rodlike bunches that can give up their energy to a circuit supporting either the same alternating electric field that bunched them (in a gyrotron), or to an alternating electric field in another circuit (in a gyroklystron). Cyclotron-resonance devices are important because, up to a point, space-charge forces tend to create tighter bunches of electrons instead of driving electrons apart as they do in other linear beam tubes (inductive output tubes, klystrons, and traveling-wave tubes). For this reason, cyclotron-resonance devices can be built using very long circuits producing very weak electric fields, and as a result, having very low losses at very high frequencies. Efficient gyrotrons have been built at frequencies as high as several hundred gigahertz and have produced continuous power of hundreds of kilowatts. *See* GYROTRON; MICROWAVE TUBE.

Robert S. Symons

Bibliography. D. Christiansen (ed.), *Electronic Engineers' Handbook*, 4th ed., 1996; J. W. Gewartowski and H. A. Watson, *Principles of Electron Tubes*, 1965; A. S. Gilmour, Jr., *Microwave Tubes*, 1986; K. R. Spangenberg, *Vacuum Tubes*, 1948.

Electron wake

The pattern of electron density fluctuation and electromagnetic disturbance set up by the passage of a swift ion through condensed matter. In dense media that can sustain well-defined resonance oscillations at a frequency ω_0, wakes of periodic character will form behind swift charged particles having speed v. The periodicity in space, λ, the distance between

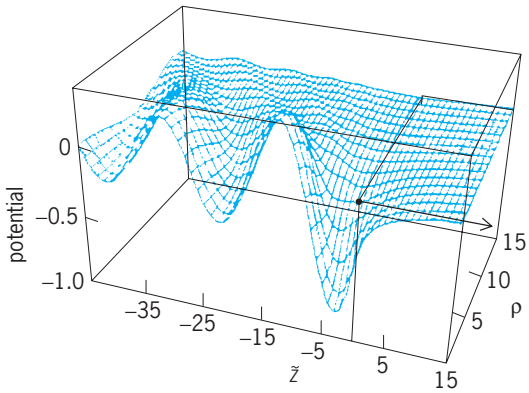

Fig. 1. Theoretical prediction of the scalar electric potential due to the electron wake of an ion penetrating a metallic medium having properties similar to those of aluminum. The ion has a speed three times the Bohr speed of 2.2 × 10^6 m/s. Coordinates $\tilde{z}$ and ρ are measured from the ion position in directions parallel and perpendicular to the ion velocity, respectively, in units of the Bohr radius, a_0 = 0.0529 nm. The potential, which does not include that of the bare ion charge, is scaled by the ion charge and is given in units of the potential at a distance of a_0 from a proton. (*After P. M. Echenique, R. H. Ritchie, and W. Brandt, Spatial excitation patterns induced by swift ions in condensed matter, Phys. Rev., B20:2567–2580, 1979*)

troughs of the wake, is given by Eq (1). The oscil-

$$\lambda = \frac{2\pi v}{\omega_0} \tag{1}$$

lations (**Fig. 1**) trail behind the ion, move with the ion velocity, and have the frequency ω_0. In addition, close collisions between the ion and electrons of the medium cause electrons to recoil to form the analog of a bow wave ahead of the ion. *See* RESONANCE (QUANTUM MECHANICS).

In metals such as aluminum, the resonant frequency ω_0 is just the plasma frequency, ω_p, of the conduction electrons. The plasma frequency is given by Eq. (2), in SI units, where n_0, e, and m^* are the

$$\omega_p = \left(\frac{n_0 e^2}{\epsilon_0 m^*}\right)^{1/2} \tag{2}$$

equilibrium electron density and the charge and effective mass of the electron, respectively, and ϵ_0 is the permittivity of free space. Frictional effects, due to collisions with ions of the medium, give plasma oscillations a finite lifetime. Consequently, the wake is damped at distances of the order of 10 λ, which is typically about 20 nanometers behind the ion. Dynamical screening effects limit the wake to lateral distances on the order of v/ω_0 from the track of the particle. The maximum amplitude of the electric potential is given approximately by $Ze\omega_0/(4\epsilon_0 v)$, in SI units, where Ze is the ion charge. Electron wakes may be generated in media such as amorphous carbon and in insulators as well. *See* ELECTRICAL UNITS AND STANDARDS; FREE-ELECTRON THEORY OF METALS; PLASMA (PHYSICS).

The wake at the position of the guiding ion is of special significance. The electric field there times the ion charge represents the reaction of the medium to the ion and yields the stopping power of the medium for the ion, that is, the energy loss per unit path length.

Coulomb explosion. When molecular ions are injected into a solid with speeds greater than v_0 = 2.2 × 10^6 m/s, the so-called Bohr speed (the speed of an electron in the ground state of hydrogen according to the Bohr model), the valence electrons are stripped, leaving atomic ions to propagate as clusters of correlated charged particles through the medium. A dicluster is composed of two atomic ions traveling close together at nearly the same velocity. A wake is formed given by a (generally nonlinear) superposition of wakes due to the individual ions of the cluster. The dynamically modified Coulomb repulsion between its constituents causes the cluster, in effect, to explode. A pair of ions traveling with the same initial velocity, and created exactly abreast of one another, will recede rapidly from one another because of the Coulomb force acting on them, until polarization of the medium screens out the repulsive force at separations on the order of v/ω_0. At the same time, the center of mass of the pair will continue to move in the original direction. In most cases of interest, the initial speed is much greater than the relative speed of recession acquired by the pair. In typical experiments, the cluster-particle interaction probes the slope of the cluster wake potential near the origin, because the foils used in most experiments are so thin that the separation of the ions after traveling through the foil does not greatly exceed their initial separation. An important effect of the wake interaction is to cause the cluster to lose energy at a faster rate than would its isolated constituents traveling at the same speed, because the wake field of a given ion in a cluster acts, in most experiments, to retard the other ions of the cluster.

Measurements of the angular deflections and energy losses of protons resulting from diclusters formed from swift $(HeH)^+$ or $(OH)^+$ ions bombarding thin foils yield angular distributions (**Fig. 2**) that have a circular character due to the action of a Coulomb explosion. Such distributions generally have a large peaked region on the perimeter due to the trailing protons that are focused by the wake of the other ion in the Coulomb explosion. There is also a much smaller peak due to protons that lead the ion. Such experiments have been important in establishing the structure of molecular ions that were formerly not well known.

In other work with diclusters, the oscillatory character of the wake is vividly displayed in a two-foil experiment. A cluster enters the first foil and, after passing through a vacuum separating the carbon foils, enters the second one. The trailing ion experiences the wake force of the leading one. The dependence of the yield of secondary electrons from a final target on the distance between carbon foils shows the characteristic oscillatory behavior. *See* COULOMB EXPLOSION.

Nonlinear effects. It has been proposed that a moving positron (and possibly an electron) may become self-bound in the wake that it creates in matter. This composite entity might be termed a wakeon.

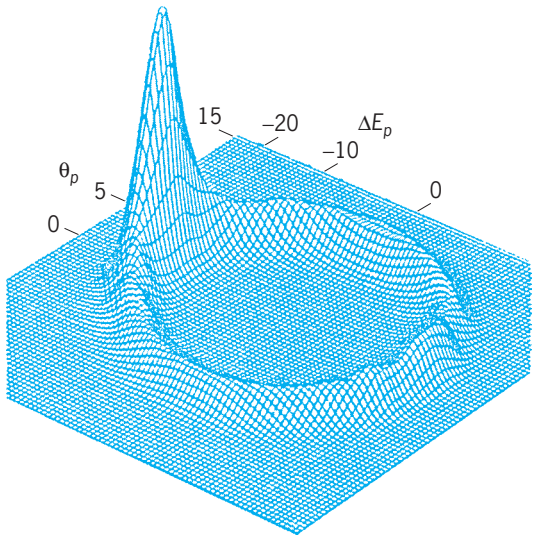

Fig. 2. Distribution of protons emerging from a carbon foil of 15.3 nm thickness that is bombarded by OH^+ molecular ions having an energy of 11.2 MeV. The distribution is given with respect to energy loss ΔE_p in kiloelectronvolts and angular deflection θ_p in milliradians measured with respect to the direction of the incident ions. The measurements were made in coincidence with the detection of the ion O^{6+}. (*After A. Breskin et al., Coulomb explosion of 11.2 OH^+ ion in carbon foils, Nucl. Instrum. Meth., 170:93–97, 1980*)

A theoretical formulation, allowing for the nonlinear response of the medium in the framework of soliton theory, has been invoked to determine the effective mass of a positron in different metals. It is found that the effective mass is appreciably larger than the mass in vacuum and appears to agree with experimentally determined values. *See* CHARGED PARTICLE BEAMS; POSITRON; SOLITON. Rufus H. Ritchie

Bibliography. P. M. Echenique, F. Flores, and R. H. Ritchie, Dynamic screening of ions in condensed matter, *Sol. State Phys.*, 43:229–308, 1990.

Electronegativity

Electronegativity, according to L. Pauling, is "the power of an atom in a molecule to attract electrons to itself." With the concept of electronegativity, a vast number of observations of chemical and physical properties have been either correlated or predicted. Quantitative definitions and scales of electronegativity have been based not on electron distribution itself but on properties which were assumed to reflect electronegativity.

The electronegativity of an element depends upon its valence state and thus is not an invariant atomic property. As an example, the electron-withdrawing ability of an sp_n hybrid orbital centered on carbon and directed toward hydrogen increases as the percentage of *s* character in the orbital increases in the series ethane < ethylene < acetylene. Thus, according to this concept of orbital electronegativity, each element exhibits a range of electronegativity values. In the following paragraphs, a few scales of electronegativity will be discussed.

The original scale, proposed by Pauling in 1932, is based upon the difference Δ between the energy of the A—B bond in the compound AB_n and the mean of the energies of the homopolar bonds A—A and B—B, as in Eq. (1). The A—B bond energy exceeds

$$\Delta = E(\text{A—B}) - \frac{E(\text{A—A}) + E(\text{B—B})}{2} \qquad (1)$$

the arithmetic means of the A—A and B—B bonds to an increasing extent as the elements A and B diverge in electron-attracting ability. The difference in electronegativities of bonded atoms is proportional to the square root of the energy difference Δ, as in Eq. (2). The proportionality factor 0.208 converts

$$\chi_A - \chi_B = 0.208\sqrt{\Delta} \qquad (2)$$

the units of energy from kilocalories to electronvolts. After the electronegativity of one element is arbitrarily assigned, other values of electronegativity can be calculated from thermochemical data. Values for selected elements in common oxidation states are presented in the **table**. Electronegativity increases with increasing oxidation state. For example, $\chi_{\text{Sn(II)}} = 1.80$ and $\chi_{\text{Sn(IV)}} = 1.96$.

R. S. Mulliken proposed that the electronegativity of an element is given by the average of the valence-state ionization potential and electron affinity: $\chi_M = (IP_V + EA_V)/2$. The quantities IP_V and EA_V are not observable properties of the ground state of an atom, but are energies for a hypothetical state of the isolated atom having the same electronic configuration (hybridization, electron-electron interaction, and so forth) as the atom in the molecule. The Mulliken approach has a sound theoretical basis, is consistent with Pauling's original definition, and gives orbital electronegativities, not invariant atomic electronegativities. Valence-state ionization potentials and electron affinities have been calculated from the equations $IP_V = IP_g + P^+ - P^0$ and $EA_V = EA_g + P^0 - P^-$, where IP_g and EA_g are ground-state potentials and

Average electronegativities from thermochemical data

Element	Value	Element	Value
H	2.20	Al	1.61
Li	0.98	Ga	1.81
Na	0.93	In	1.78
K	0.82	Tl	2.04
Rb	0.82	C	2.55
Cs	0.79	Si	1.90
Be	1.57	Ge	2.01
Mg	1.31	Sn	1.96
Ca	1.00	Pb	2.33
Sr	0.95	N	3.04
Ba	0.89	P	2.19
Sc	1.36	As	2.18
Ti	1.54	Sb	2.05
V	1.63	Bi	2.02
Cr	1.66	O	3.44
Mn	1.55	S	2.58
Fe	1.83	Se	2.55
Co	1.88	F	3.98
Ni	1.91	Cl	3.16
Cu	1.90	Br	2.96
Zn	1.65	I	2.66
B	2.04		

affinities, respectively, and P^+, P^0, and P^- are promotion energies of the positive ion, atom, and negative ion, respectively. The calculation of *d*-orbital electronegativity by the Mulliken method has not been accomplished for nontransition elements due to the lack of spectroscopic data. Since electronegativity is a sensitive function of *d*-orbital hybridization and since the extent of *d*-orbital participation generally cannot be ascertained quantitatively, the calculations of electronegativities for the heavier elements are limited.

The energy of an ion relative to the neutral atom can be expressed as a power series $E = aq + bq^2 + cq^3 + dq^4$, where q is the formal oxidation state or ionic charge for a particular state of ionization. Z. R. P. Iczkowaski and J. L. Margrave defined the electronegativity of a neutral atom as the derivative, $\chi_{IM} = (dE/dq)_{q=0}$. As a fairly good approximation, the last two terms in the above power series can be dropped, giving Eq. (3). The units of χ_{IM} are

$$\chi_{IM} = \frac{dE}{dq} = \frac{d(aq + bq^2)}{dq} = a + 2bq \qquad (3)$$

energy/electron and the magnitudes are the same, in accordance with theory, as those of Mulliken if E is evaluated only from the electron affinity and the first ionization potential. The quantity a is the Mulliken electronegativity for a neutral atom, and electronegativity is shown by Eq. (3) to increase linearly with increasing positive charge. By using IP_V and EA_V values, H. H. Jaffé and coworkers calculated the electronegativities of certain vacant, singly occupied, and doubly occupied orbitals.

Electronegativity was defined by A. L. Allred and E. G. Rochow as the force of attraction between a nucleus and an electron from a bonded atom. The electrostatic force was calculated simply from the effective nuclear charge and the atomic radius, as in Eq. (4).

$$\chi = \frac{0.359Z_{\text{eff}}}{r^2} + 0.744 \qquad (4)$$

A quantum-defect electronegativity scale has been developed from potentials based on atomic spectral data, and a nonempirical scale has been calculated by an ab initio method using floating gaussian orbitals.

Other methods for calculating electronegativities utilize such observables as bond-stretching force constants, electrostatic potentials, spectra, and covalent radii. The fact that the various scales of electronegativity have different dimensions (energy$^{1/2}$, energy/electron, force, potential, and so forth) or no dimension reflects the widespread results of differences in electron-attracting ability. The measurement of electronegativities involves observations of properties dependent upon electron distribution. Close agreement of electronegativity values obtained from measurements of several diverse properties lends confidence and utility to the concept. A. Louis Allred

Bibliography. F. A. Cotton and G. Wilkenson, *Advanced Inorganic Chemistry*, 6th ed., 1999; L. Pauling, *Nature of the Chemical Bond*, 3d ed., 1960; W. W. Porterfield, *Inorganic Chemistry: A Unified Approach*, 2d ed., 1993; A. G. Sharpe, *Inorganic Chemistry*, 3d ed., 1992.

Electronic display

An electronic component used to convert electrical signals into visual imagery in real time suitable for direct interpretation by a human operator. It serves as the visual interface between human and machine. The visual imagery is processed, composed, and optimized for easy interpretation and minimum reading error. The electronic display is dynamic in that it presents information within a fraction of a second from the time received and continuously holds that information, using refresh or memory techniques, until new information is received. The image is created by electronically making a pattern from a visual contrast in luminance between (1) individual electrically alterable picture elements (pixels) in a matrix array of pixels in flat-panel displays (FPDs) or (2) electrically excited and nonexcited areas in a phosphor film in cathode-ray tubes (CRTs). High-information-content (HIC) displays are those displays that have a sufficient number of pixels (75,000 to 2,000,000) to show standard or high-definition television images, or comparable computer images.

The use of electronic displays for presentation of graphs, symbols, alphanumerics, and still and video pictures has doubled every several years, in parallel with the rapid expansion of microelectronics. Electronic displays have largely replaced traditional mechanical devices, counters, galvanometers, and, to a degree, hardcopy (paper) means for presenting information. This change is due to the increased use of computers, microprocessors, inexpensive large-scale integration (LSI) electronics, and digital mass memories. The success of the digital watch, handheld calculator, and personal computer is directly attributable to the availability of inexpensive LSI electronics and electronic displays. *See* CALCULATORS; COMPUTER; COMPUTER GRAPHICS; COMPUTER STORAGE TECHNOLOGY; INTEGRATED CIRCUITS; MICROCOMPUTER; MICROPROCESSOR; WATCH.

Electronic transducers and four-digit (or more) FPDs have been used to replace the galvanometer movement, thermometer scale, barometer movement, and other forms of scientific instrumentation. Large signs, arrival and departure announcements, and scoreboards also use electronic means to portray changing messages and data. The major electronic display application is in home color television.

The computer terminal using a CRT or FPD is the most important industrial application of electronic displays. The personal computer terminal with a microprocessor and mass memory serve to replace the office paper, typewriter, and file cabinet. *See* COMPUTER PERIPHERAL DEVICES; WORD PROCESSING.

With advances in high-information-content FPD technology, the electronic display industry went through a dramatic change in the early 1990s. This

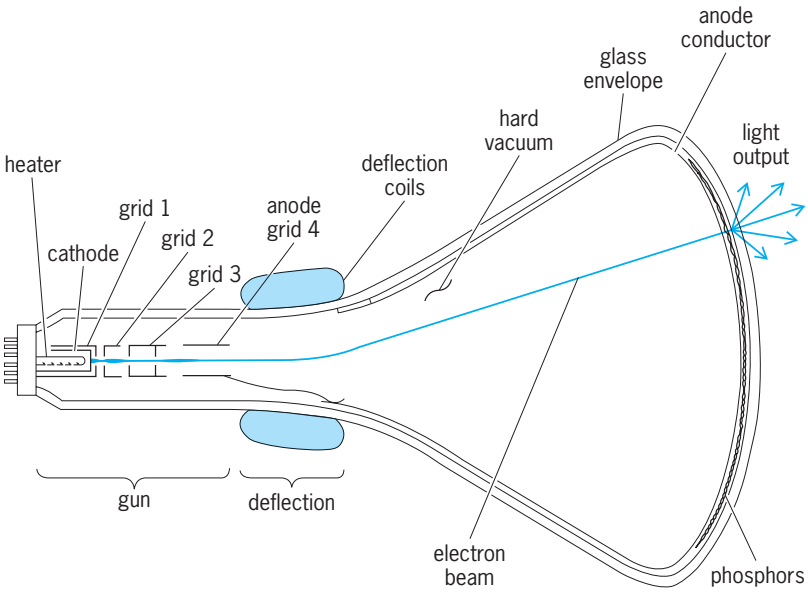

Fig. 1. Cathode-ray tube using electrostatic and magnetic deflection.

was primarily due to breakthroughs and manufacturing advances in liquid-crystal displays (LCDs). Color active-matrix LCDs (AMLCDs) have better performance than color CRTs for video and computer graphics and are thin and portable. However, the AMLCDs are several times more expensive than CRTs, so their use remains restricted to applications such as personal computers and television receivers where the CRT will not fit. *See* LIQUID CRYSTALS.

Cathode-ray tube. The primary applications of the CRT are in home entertainment television, computer monitors, scientific and electrical engineering oscilloscopes, radar display, and alphanumeric and graphic electronic displays. *See* OSCILLOSCOPE; RADAR; TELEVISION.

The CRT (**Fig. 1**) has a viewing screen coated with a phosphor which emits light when struck with a beam of high-energy electrons. The electrons are emitted from the cathode at the rear of the tube in a beam that is focused electrostatically (or magnetically) to a dot or spot on the phosphor screen and positioned in horizontal and vertical coordinates by magnetic (or electrostatic) forces. The cathode grids and electron-focusing lenses are incorporated into a subassembly called the gun. The beam is accelerated toward the phosphor by a high voltage of approximately 20 kV between the cathode and the anode.

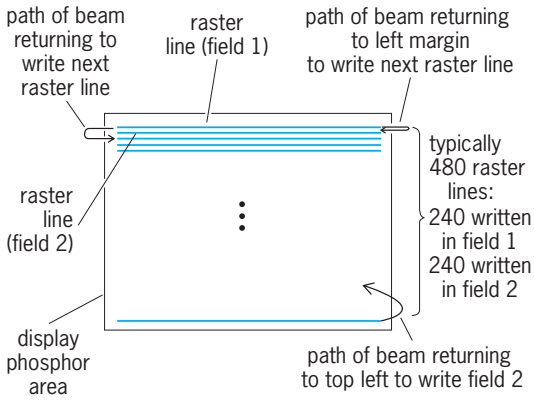

Fig. 2. Cathode-ray tube raster for television, using interlace.

Imagery is created on the screen as the CRT raster (**Fig. 2**) is traced out. The video signal is applied, after amplification, directly to grid 2 of the gun, and controls the amplitude of the electron beam and thus the luminous output at the display surface. The deflection coils steer the beam to trace out the raster. The horizontal-scan deflection signal causes the beam to trace out the horizontal lines and then fly back for the next line. The vertical-scan deflection signal causes the beam to be stepped down the raster and then retraced to the top left corner at the beginning of each frame. Two fields are interlaced, with the raster lines of one traced between those of the other. The purpose of the interlace is to minimize flicker in the picture and reduce video bandwidth. Computer CRTs do not use interlace, but continuous scan to prevent flicker. *See* CATHODE-RAY TUBE; PICTURE TUBE; TELEVISION RECEIVER.

Flat-panel displays. Because of the depth dimension of the CRT, there has been a concentrated effort to develop FPDs. A primary motivating factor has been to achieve a flat high-information-content display which could be hung on a wall or carried in a briefcase. Over the years, the electrical phenomena most extensively developed for FPDs have been gas discharge (plasma), electroluminescence, light-emitting diode, cathodoluminescence, and liquid crystallinity. The cost of FPDs is higher than CRTs on a character-per basis for HIC displays. *See* CATHODOLUMINESCENCE; ELECTROLUMINESCENCE; LIGHT-EMITTING DIODE.

Before the 1990s, cost and performance limitations restricted the use of FPDs to specialized applications. With the advent of two major breakthroughs in the 1980s, LCDs emerged as the leading FPD. These two breakthroughs were the invention of the supertwisted nematic (STN) LCD, often referred to as passive-matrix LCD (PMLCD), and the evolution of a manufacturable AMLCD. The STN inventions improve LCDs so that they can be made as a HIC FPD because of improvements in the matrix addressing capability. Both of these types are made in monochrome and full color. The AMLCD is fast enough for video and for multimedia displays, comparable to a CRT. The PMLCD is fast enough for computer use but not video; however, PMLCDs are about half the cost of AMLCDs.

Matrix addressing. Flat-panel displays are typically matrix-addressed. A row is enabled to accept display information in parallel via the column lines. The electronics commutate through the rows, serving the same purpose as the vertical deflection amplifier of the CRT. The column data are shifted into the column drivers, and at the proper time applied to the column lines.

The thickness of the flat-panel matrix addressing electrodes (**Fig. 3**) is approximately 0.1 in. (2.5 mm).

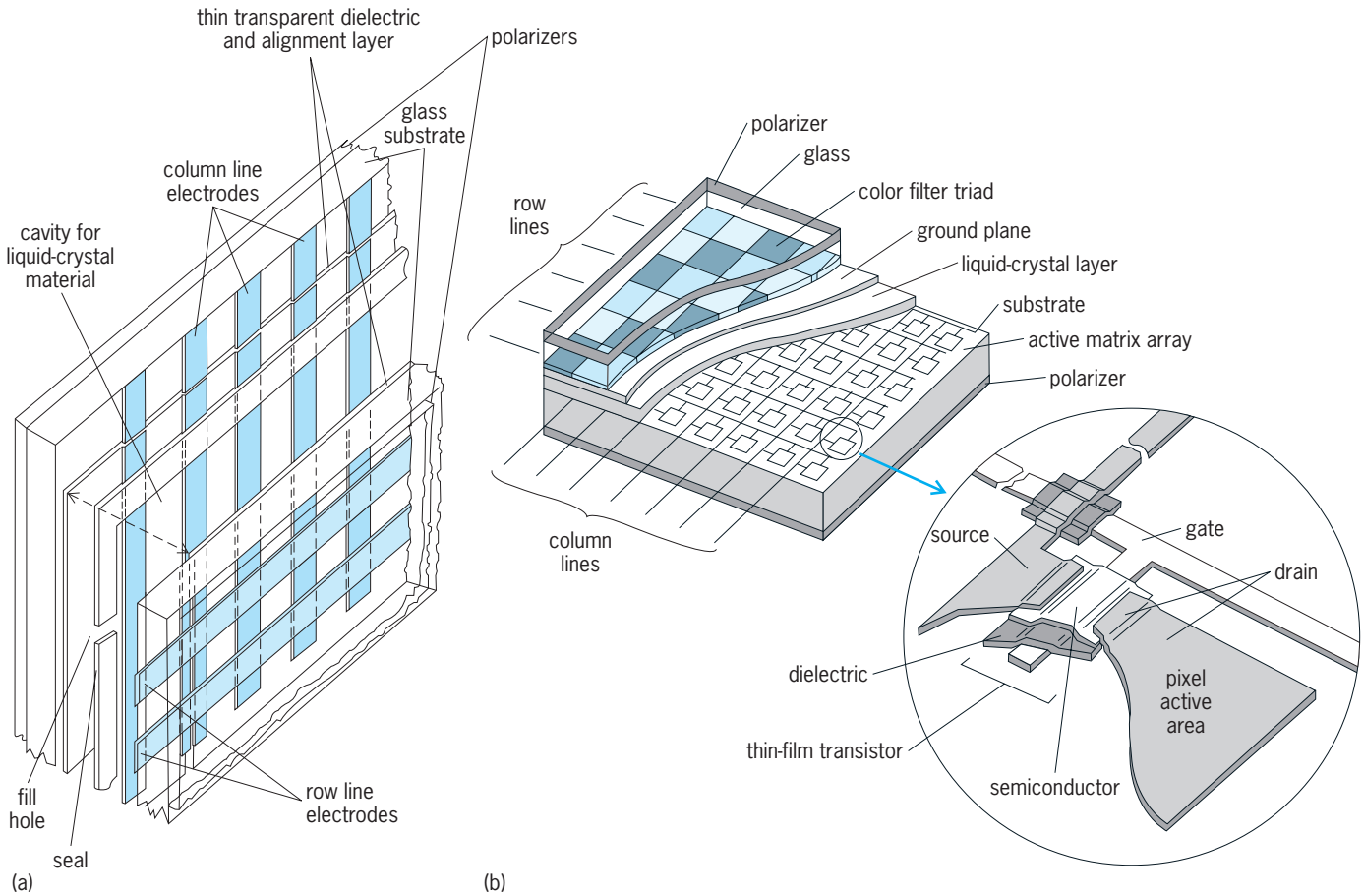

Fig. 3. Exploded sections of liquid-crystal displays (LCDs). (*a*) Passive-matrix LCD (PMLCD). Column-line and row-line electrodes are made of transparent indium–tin oxide. Liquid-crystal material aligns to the alignment layer (which has been mechanically rubbed with cotton cloth or similar means to import microgrooves on the surface) during filling steps. Alignment rub directions are rotated 90° for twisted nematic LCDs and 180–270° for supertwisted nematic LCDs. (*b*) Active-matrix LCD (AMLCD).

The row and column lines are spaced at 80 lines per inch (31.5 lines per centimeter). The intersection of each row line with each column line defines a pixel. A pixel is a picture element, and denotes the smallest addressable element for spatial information in an electronic display. To control the pixels, the row and column electronic drivers are attached at the edges of the glass, using matrix-addressing techniques.

Comparison of technologies. Each of the flat-panel technologies (**Table 1**) has unique advantages. The more important considerations include luminous efficiency, addressability, duty factor, gray scale, color, and cost. The capabilities in each area have evolved to the point of cost-effective applications to commercial and military products. For the vectorgraph and video display categories, liquid-crystal, electroluminescent, and gas-discharge or plasma displays have emerged as the most cost-effective approaches.

AMLCD technology. LCDs are manufactured in many configurations. The highest-performing and most widely used LCD is the AMLCD, where the active-matrix format is a thin-film (field-effect) transistor (TFT) [Fig. 3*b*]. The most common semiconductor used in the TFT is amorphous silicon; next is polysilicon; single-crystal silicon and cadmium selenide are also used. For medium-to-large displays, the semiconductor is almost universally amorphous silicon with a mobility typically of 0.5–1.0. This mobility, although low by semiconductor standards, is fast enough for electronic displays and, more importantly, the TFT can be made on an inexpensive borosilicate glass substrate. The polysilicon TFT is typically annealed amorphous silicon on quartz and is used for small (less than 2 in. or 5 cm in diagonal) displays. Single-crystal silicon, which has the highest mobility, is also used in small displays.

LCD applications. Applications of LCDs include portable computing devices and memory aids; electronic games; moving map devices in conjunction with global positioning satellites (GPS) for navigation of boats, aircraft, and motored vehicles; educational devices and video for portable handheld devices; and displays for television, entertainment, education, and information dissemination. LCDs are also used in helmet displays and virtual-reality displays. *See* SATELLITE NAVIGATION SYSTEMS; VIDEO GAMES.

The smaller AMLCDs are used in camcorder viewfinders, helmet displays, and projectors. The medium-size displays are used in all the applications that are viewed directly, such as portable computers or television receivers.

PMLCDs are used with the STN configuration of

TABLE 1. Flat-panel technologies

Technology	Phenomena
Emissive displays	
Gas discharge (GD)	Cathode glow from conducting gaseous discharge
Plasma panel (PDP)	Alternating-current capacitively coupled gas discharge
Light-emitting diode (LED)	Electron injection in a forward-biased *pn* semiconductor junction
Vacuum fluorescence (VFD)	Electron bombardment of phosphor in hard vacuum under control of a grid
Electroluminescence (ELD)	Electron conduction in polycrystalline phosphors due to high electric field
Flat cathode-ray tube (FCRT)	Electron bombardment of phosphor in hard vacuum under control of a grid or cathode
Field-emitter display (FED)	Electron bombardment of phosphor in hard vacuum under control of row and column electrodes, where the cathode is a field-emitter tip
Organic light-emitting diode (OLED)	Complex organic semiconducter compounds that emit photons when conducting electrons
Nonemissive displays	
Liquid crystallinity	Electrostatic rotation of organic compounds which exhibit liquid crystallinity
Electrochromism	Charging and discharging chemical systems (battery) which exhibit a color change in accordance with Faraday's law
Colloidal suspensions	Electrostatic transport or rotation of light-absorbing particles in a colloidal suspension
Micromechanical display (MMD)	Mirror displays on a silicon chip, deflectable with electrostatic force

LCD in devices where some reduction in performance is acceptable, such as portable computer displays that do not require video imagery. PMLCDs in the basic twisted nematic (TN) mode are used in inexpensive toys, automotive panels, meter displays, and watches.

Display categories. Electronic displays can be categorized into four classifications; pseudoanalog, alphanumeric, vectorgraphic, and video (**Table 2**). Each classification is defined by natural technical boundaries and cost considerations. The categorization is useful in visualizing the extent to which electronic displays are used.

Special-purpose displays. The above categorization emphasizes direct-view-type electronic displays, which are of primary interest to industry. There are other special-purpose displays used in very sophisticated applications.

Projection display. In the projection display an image is generated on a high-brightness CRT or similar electronic image generator, and then optically projected onto a larger screen. To illuminate screens larger than approximately 3×4 ft (1×1.3 m) and in color, multiple CRTs or light valves are used. The light valve is any direct-view display optimized for reflecting or transmitting the image, with an independent collimated light source for projection purposes. Light valves create images to control the reflection of light to be projected onto the screen. This permits powerful light sources such as xenon lamps to be used independent of the image-generating technique. Oil-film light valves, micromechanical displays, and liquid-crystal light valves are examples of devices used in large command and control and theater-size electronic display presentations.

Three-dimensional imagery. True three-dimensional imagery can be created electronically by several techniques. One technique requires goggles using PLZT (lead zirconate titanate modified with lanthanum) electrooptical ceramic eyepieces or liquid-crystal shutters over each eye. The eyepieces are electronically controllable shutters, with the ability to be reversibly switched from open to closed in microseconds. Two images from two television cameras placed to obtain the desired stereoscopic effects are electronically interlaced and displayed on one CRT monitor. Each image is sequentially displayed from each camera at television video rates. The goggles are synchronized to be opened and closed so that only the right eye sees the image from the right camera and the left eye sees the image from the left camera. The viewer sees true three-dimensional perspective while looking at the CRT monitor through the goggles. *See* ELECTROOPTICS.

Three-dimensional displays are made integral to and self-contained in a helmet and are called virtual-reality displays. A separate display is used for each eye to show the stereo pair of images. The display is typically either a miniature CRT or a miniature LCD, with appropriate optics to present the image at the proper focal length in front of the eye. In a virtual-reality display, the user cannot see beyond the displays as the intent is to completely control the images seen by the viewer. The viewer is visually engulfed in a synthetic visual environment. *See* VIRTUAL REALITY.

Helmet-mounted and heads-up displays. Helmet-mounted displays (sometimes called visually coupled displays) and heads-up displays are used in aircraft, automobiles, and toys. In both of these displays the image is projected, usually from a CRT onto a combining glass, and collimated to be in focus at infinity. The combining glass screen is designed to reflect the display imagery to the viewer, usually at selected wavelengths of light, while being sufficiently transmissive for the viewer to see the scene beyond. The primary application for the heads-up display is to present critical aircraft performance, such as speed and altitude, on a combining glass at the wind-screen for pilot monitoring while permitting the pilot to look out the window for other aircraft or the runway. The primary application for the helmet-mounted display is to present, on a combining glass within the visor of the helmet of a helicopter gunner, primary information for directing firepower. The angular direction of

TABLE 2. Electronic display spectrum of applications

Classification	Characteristics	Applications	Electronic technologies
Pseudoanalog	Dedicated arrangement of discrete pixels used to present analog or qualitative information	Meterlike presentations, go/no-go messages, legends and alerts, analog-like (watch) dial	Gas discharge, light-emitting diodes, liquid crystal, incandescent lamps
Alphanumeric	Dedicated alphanumeric pixel font of normally less than 480 characters; most common is 4- and 8-character numeric displays	Digital watches, calculators, digital multimeters, message terminals, games	Liquid crystal, light-emitting diodes, vacuum fluorescent, gas discharge, incandescent lamps
Vectorgraphic	Large orthogonal uniform array of pixels which are addressable at medium to high speeds; normally, monochromatic with no gray scale; may have memory; normally, over 480 characters and simple graphics	Computer terminals, arrivals and departures, scheduling terminals, weather radar, air-traffic control, games	Cathode-ray tube, plasma panels, gas discharge, vacuum fluorescent, electroluminescence, PMLCD, LED
Video	Large orthogonal array of pixels which are addressed at video rates (30 frames per second); monochromatic with gray scale or full color; standardized raster scan addressing interface, arrays of pixels approximately 240 rows by 320 columns and larger	Entertainment television, graphic arts, earth resources, video repeater, medical electronics, aircraft flight instruments, computer terminals, command and control, games	Cathode-ray tube, plasma panels, electroluminescence, AMLCD, LED

the helmet is sensed and used to control weapons or radar to point in the same direction in which the pilot gunner is looking. *See* AIRCRAFT INSTRUMENTATION.

Color. Color can be created on a CRT equipped with a shadow mask duplicating quite closely all colors that occur in nature. This is done in the CRT by using three different electron guns and three phosphors in a triad of red, green, and blue on the screen at each pixel. The shadow mask is a metal screen with one hole for each triad of phosphor dots (pixel). It is precisely located and aligned with the phosphor screen. Electron beams from each of three guns are constrained by each shadow mask hole to hit each respective phosphor dot. The gun, shadow-mask holes, and phosphor dots are aligned during manufacturing so that the three beams converge to pass through the single hole (or slit) in the shadow mask and then diverge as the beams emerge with sufficient separation to impact the three different phosphor dots. If all three guns are on simultaneously, the eye, upon close inspection, sees a red, a green, and a blue dot of light at each pixel. However, at a normal viewing distance, the three dots merge together when focused in the retina of the eye, and from the laws of additive color, the pixel appears white. The phosphor itself is whitish, but it emits different colors because of the dopants in the phosphor, and the image is created from the emitted light.

Penetration phosphors are also used to create color on CRT displays to eliminate the need for the shadow mask and extra guns. However, the color is limited and the brightness is low. Normally, two phosphors are placed on the screen in two layers or in microspheres of two layers. The gun and CRT anode are operated in two energy states to produce either a high-energy or a low-energy electron beam switchable in time. The high-energy beam penetrates the first phosphor layer and is stopped at the second layer. It then excites the second layer to produce its characteristic color. The low-energy beam is stopped by the first phosphor layer and excites it to produce its characteristic color. The two phosphor colors most often used are red and green. Intermediate-energy beams make it possible to fractionally excite both layers to get color combinations of red and green such as yellow and orange. Practical considerations limit this color approach to these four. Full color would require at least three primary colors such as red, green, and blue. Three layers of phosphors are limited in brightness due to practical considerations and have not been commercially available.

Monochromatic color is readily produced by FPD technologies. One color of a plasma panel is normally orange, owing to the neon gas. Full-color plasma panels are available in medium resolution. Other monochromatic colors are feasible. Light-emitting diode (LED) luminance is normally red, yellow, or green. Blue LEDs are available and have high luminous efficiency. Electroluminescence is normally yellow or green, owing to the manganese or copper activator, respectively. Full-color electroluminescent displays have been demonstrated but are not cost-effective in production. In flat-panel as in CRT displays, full color is produced by using a triad of red, green, and blue for each pixel, or sequential flashing at 180 Hz of red, green, or blue at each pixel area.

Full color is very important for entertainment television displays. Most industrial electronic displays do not necessarily need full color, but full color has been used more frequently as color display costs have decreased. In these applications, the color display instrument is usually a CRT using the shadow-mask color technique.

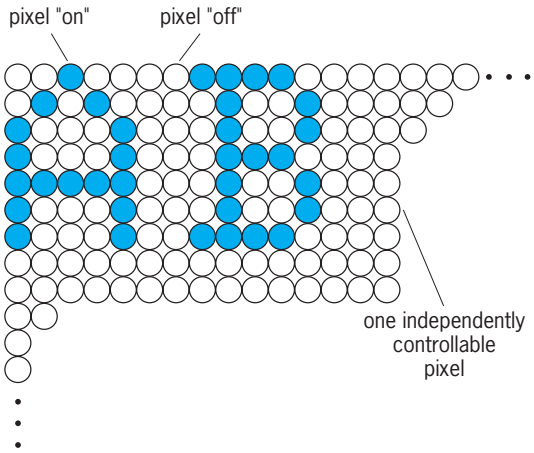

Fig. 4. Pixel array used for creating electronic display images.

Display technique. The essence of electronic displays is based upon the ability to turn on and off individual pixels (**Fig. 4**). A typical HIC display will have a quarter million pixels in an orthogonal array, each under individual control by the electronics. The pixel resolution is normally just at or below the resolving power of the eye at one minute of arc. Thus, a good-quality picture can be created from a pattern of activated pixels.

The pixel concept for electronic displays has evolved from the modern FPD technologies and digital electronics. It has been extended to the analog-raster-scan CRT in the following way: The electron beam from the gun is deflected magnetically (or electrostatically), so as to sweep across the phosphor and thereby cause a line to luminesce on the face of the CRT. In digitally modulated CRTs, the cathode is modulated by a sine wave as the beam is swept across the face of the CRT. Here, instead of a continuous line, a string of dots results. Each dot corresponds to a pixel. Each pixel is on when the beam density is high, and off when the beam is off. The beam is turned off between pixels. The pixels are refreshed from 30 to 60 times per second on a CRT.

Pixels are created in all the rows of the entire CRT raster in what is called a CRT-digital raster. This approach is commonly used in industrial applications and computer terminals, since it is easily interfaced with digital electronics. Home entertainment television uses an analog-raster-scan approach (Fig. 2).

There are some applications in which a nonraster approach is used to create alphanumeric characters and vectors on CRTs. The electron beam is deflected under control of the deflection amplifiers to stroke out each line of the image. When characters and vectors are generated this way, they are like Lissajous figures, as opposed to (digital or analog) raster characters and vectors. The Lissajous characters and vectors are best suited to large (25-in. or 63.5-cm diagonal) CRTs and where there are numerous vectors, straight lines, and curves. Vectors and curves drawn with the raster technique have stair steps. Lissajous vectors and curves are always smooth and continuous. Lissajous techniques have increasingly yielded to the digital raster as the cost of digital electronics has improved. *See* LISSAJOUS FIGURES.

Font. With flat-panel and CRT digital display techniques, alphanumeric character fonts are created by turning on the appropriate pixels in an array. One standard size is a 5×7 array with two pixels between characters and two pixels between rows (Fig. 4). All the letters and numbers can be created on this common array format. Several other combinations of pixels may be used to create the letter A. The viewer soon becomes accustomed to the minor variation. Readers do not read pixels but read letters and words, and therefore the exact detail of the character pixel pattern is of a secondary consideration. In general, the more pixels available in the basic array, the more esthetically pleasing is the character, at the cost of additional electronics to control the extra pixels. The viewing distance is normally far enough so that the pixels blur together.

A very efficient and elegant array has evolved for portraying numeric characters and many letters of the alphabet, called the seven-bar font (**Fig. 5**). Each bar is a pixel by definition. This font was initially considered crude when compared to the Leroy font and other more esthetic printer fonts. It is now universally accepted. A similar 14 bar font is sometimes used for alphanumeric characters.

Display electronic addressing. The numeric or alphanumeric display electronic drive may be performed in a single LSI metal oxide semiconductor (MOS) chip mounted in a single dual in-line package (DIP) suitable for direct assembly on a printed circuit board. All the timing, logic, memory, resistors, and drivers are contained in a single chip. A four- or seven-bit character code is serially fed into the chip for display (**Fig. 6**).

A computer terminal (**Fig. 7**) will incorporate a microprocessor unit and a CRT controller to perform master control and CRT housekeeping tasks. A line of video data is loaded into the shift register for serial drive of the video amplifier at the time a display line starts. The shift register is loaded during flyback

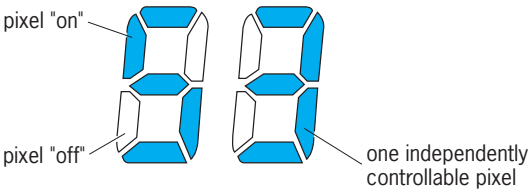

Fig. 5. Seven-bar numeric font.

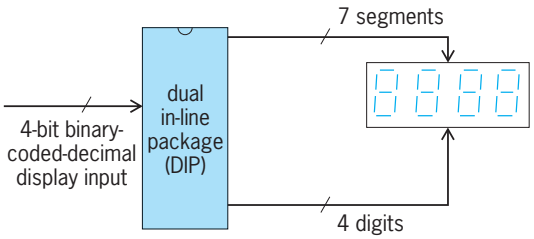

Fig. 6. Block diagram of LSI MOS electronic drive for numeric display.

time from the character generator. The character generator is a decoder transforming the alphanumeric information in coded form such as ASCII (American National Standard Code for Information Interchange) into pixel signals for each raster line. The ASCII code is stored in the display refresh memory. The entire cathode-ray-tube frame is stored in this memory, and is continuously displayed until changed under control of the microprocessor unit. New information can come from the mass memory, the keyboard, or other subsystems on the data bus and address bus, all under control of the microprocessor unit.

The CRT controller performs all the housekeeping tasks for proper CRT operation. These include the vertical- and horizontal-raster synchronization signals, the cursor control, blinking, blanking, interlace, paging, and scrolling. The CRT controller is a single LSI MOS chip, as are most of the other units in the terminal (Fig. 7).

Custom chip sets are available for controlling CRTs and flat-panel displays. The sets include processor, memory, and all programmable functions necessary for self-contained operation. The chip set operates off the main bus and formats the red, green, and blue signals and synchronization signals to the display. HIC FPDs typically come with their own electronics, with a red-green-blue plus sync interface similar to that used in CRTs.

Applications. Because of performance and cost issues, there now are only two major display technologies, CRTs and LCDs. Full-color displays are now expected by users, except for some niche areas such as medical imaging where resolution and gray shades are traditionally more important than color. Display technologies which cannot efficiently produce saturated red, green, and blue colors for electronically controlled additive mixing to produce full color can no longer find applications. Those technologies include electroluminescence, vacuum fluorescence, and field-emitter displays.

The CRT remains the most widely used display technology. It is made in the traditional 3:4 aspect ratio and a 9:16 aspect ratio for high-definition television (HDTV) with resolutions of over 2,000,000 pixels and sizes of up to 38 in. (1 m) in diagonal. For larger screen sizes, smaller projection monochrome red, green, and blue CRTs are used for consumer products up to 60 in. (1.5 m) and industrial products up to 100 in. (2.5 m) in diagonal screen size. The dominant position of the CRT is maintained because of its good performance and relatively low cost. The only limitation is the bulky size of the CRT which makes it impractical to ship and install when exceeding 40 in. (1 m) in diagonal size. CRT projectors using three or more smaller monochrome tubes are practical but more costly and have competition from other technologies in both price and performance.

If a CRT cannot be used due to its depth and volume, an LCD is probably the best alternative. LCDs can be made in almost any resolution and in sizes from 0.5 to 22 in. (1.2 to 55 cm) in diagonal, with possibilities of tiling four panels together. The LCD uses photolithography in manufacturing and is limited in size and resolution by the limitations in the photolithographic machinery and process. For larger sizes, LCDs are made as projectors.

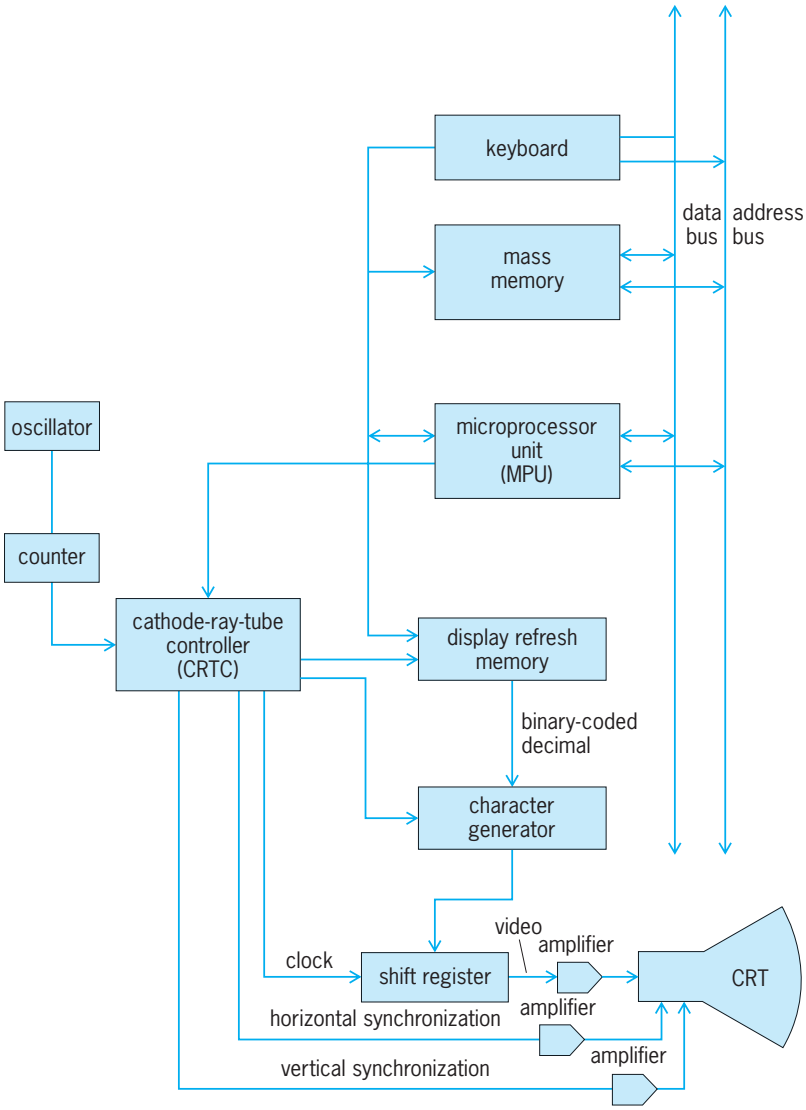

Fig. 7. Block diagram of LSI electronics for a smart computer terminal cathode-ray-tube display.

For direct-view sizes from 20 to 60 in. (50 to 150 cm) in diagonal, the plasma display panel is the only viable technology. PDPs use screen printing manufacturing technology, and therefore the larger sizes at HDTV resolutions are possible when 50 lines/in. (2 lines/mm) resolution is appropriate. PDPs are ideally suited to the television-on-the-wall picture applications, but they are too expensive and consume too much power for home television. They are used mainly in marketing booths, public exhibits, and industrial applications.

For low-resolution displays of 20 lines/in. (1 line/mm) or less and large-size direct-view displays, the LED has become the preferred technology. With the development of an efficient blue LED in combination with red and green LEDs, this technology is ideally suited to marquees and very large (billboard-size) displays in such applications

as sports arenas and roadside advertising. The LED sign is made from individual lamps assembled into printed circuit board modules tiled together into large arrays. The LED can be made bright enough to be read in full daylight but may consume kilowatts of power in large sign applications. Because of the lamp assembly technique, LEDs are relegated to low resolution and are not suitable for consumer television or computer monitors. Lawrence E. Tannas, Jr.

Bibliography. B. Bahadur, *Liquid Crystals, Applications and Uses*, 3 vols., 1991–1993; P. A. Keller, *The Cathode-Ray Tube: Technology, History, and Applications*, Palisades Press, 1991; L. W. MacDonald and A. C. Lowe (eds.), *Display Systems: Design and Applications*, Wiley, 1997; T. J. Nelson and J. R. Wullert II, *Electronic Information Display Technologies*, 1997; W. C. O'Mara, *Liquid Crystal Flat Panel Displays: Manufacturing Science and Technology*, 1995; S. Sherr, *Electronic Displays*, 2d ed., 1993; L. E. Tannas, Jr., Color in electronic displays, *Phys. Today*, 45(12):52–57, December 1992; L. E. Tannas, Jr., et al., *Flat-Panel Display Technologies: Japan, Russia, Ukraine, and Belarus*, Noyes Publications, 1995; L. E. Tannas, Jr., *Flat-Panel Displays and CRTs*, 1985; J. Whitaker, *Electronic Displays: Technology, Design, and Applications*, 1994.

Electronic equipment grounding

The connecting of electronic equipment to an electromagnetic reference common to itself, its power source, its environment, and the environment of its users. Electronic equipment is grounded to protect users from shock, to protect the equipment from spurious currents or voltages, and especially to isolate it from noise that contaminates its environment.

Need for thorough grounding. Usually electronic equipment is powered from the electric service that supplies the building where the equipment is used. To provide security against electric shock or fire ignition, equipment frames and power-conductor enclosures are connected and grounded. *See* GROUNDING.

In addition, grounding (together with the shielding, isolation, compensation, and equalization) is applied to minimize the entrance of extraneous signals (noise) into the equipment. Basically, electronic equipment can be treated as a sensor or signal source, a signal circuit, and central equipment (**Fig. 1**). The sensor may be any of a wide variety of input devices, such as a tape reader, card reader, memory readout, medical sensor, industrial transducer, or microphone. Information from the sensor travels along the signal circuit in digital or analog form at low level. At the central equipment the signal is processed to be used for such purposes as recording or control. Reliability of the output use depends directly on the integrity of the signal received at the central equipment.

The sophistication of electronic equipment has advanced tremendously. The necessary intensity of electrical signals has been diminished generally to below 1 V. The amount of time allocated for the transmission of one bit of information has been progressively diminished to less than 1 microsecond. These advances aggravate the likelihood that spurious error signals will be of sufficient magnitude to create an error response. The effect of an error response, if not identified and corrected for, can produce tragic results in critical areas such as manned space flight. The controlling parameter can be expressed as the signal-to-noise ratio. The problem is ultimately one of ensuring acceptably high levels of signal-to-noise ratio. *See* ELECTRICAL NOISE.

The most vulnerable circuits are typically those associated with the information-gathering function, because they generally operate at low signal level and are followed by high-gain amplifiers within the data-processing system. Such an input circuit is shown in Fig. 1. The signal-sensing element may be in the next room, on a different floor of the same building, in a different building, or even in a different city. Techniques that are effective in establishing a high signal-to-noise ratio in this critical circuit, and techniques needed to handle the less critical circuits, are reviewed here.

Noise coupling into signal circuit. Spurious unwanted electric impulses (noise) may be mixed with true information signals in numerous ways, the most common and important being electromagnetic induction, electrostatic induction, and conductive coupling (**Fig. 2**).

Magnetic induction. A changing magnetic field intensity in the space through which the signal circuits run (Fig. 2*a*) creates noise in the equipment. Such fields are present in the space surrounding any open conductor carrying substantial magnitudes of changing current. Intense fields exist around the anode leads of arc furnaces and the throat conductors of flash and spot welders. A helpful approach is to recognize that the inductive reactance of a current-carrying conductor, shown on the electrical engineer's one-line diagram as a property of the conductor, is in fact a property of the space-distributed magnetic field which surrounds the current-carrying conductor. Any electrical circuit which links a portion of such a space field will display a fractional part of the reactance voltage drop of the power conductor. *See* ELECTROMAGNETIC INDUCTION.

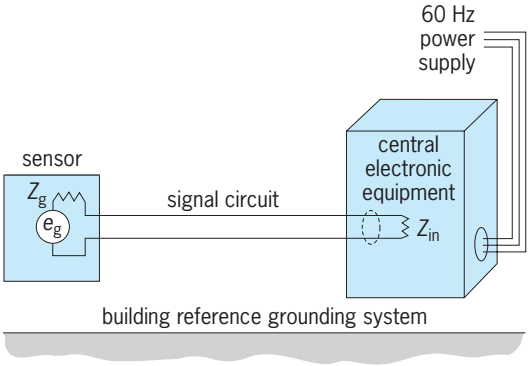

Fig. 1. The three essential components for communicating information electrically.

In Fig. 2*a* the fractional part of the changing space-distributed magnetic field ΔH_1 which threads between the two signal-circuit conductors will induce a normal-mode error signal in that circuit. The terminal sensor is often an electromagnetic device which, in the presence of a changing space-distributed magnetic field ΔH_3 in a particular direction, will display a normal-mode induced error voltage. A similar space field ΔH_2 threading between the signal circuit conductors and the nearby conducting structure which connects the central processing framework with the remote terminal structure will induce a common-mode potential which can appear on the signal circuit.

Electrostatic induction. A changing electrostatic space field (Fig. 2*b*) can induce a similar error voltage in the signal circuit by capacitive coupling. Electrostatic space-distributed fields will be created by unshielded, energized electrical conductors. Insulation over the conductor is not an attenuator. Intense fields can be expected in areas where ac high-potential testing is being done. Unshielded gaseous discharge lamps develop annoying space fields rich in harmonics and sharp step changes.

A changing-magnitude electrostatic field ΔE_1, if coupled to one of the signal-circuit conductors through a capacitance C_1, will impart to that conductor an error voltage. To the other signal conductor will be communicated a similar error voltage produced by space field ΔE_2 through coupling capacitance C_2. Even though ΔE_1 equals ΔE_2 and C_1 equals C_2, a normal-mode error signal voltage will be created if the signal circuit is not symmetrically balanced in its coupling to the central-equipment reference ground plane. Also, the terminal sensor may contain such parts as terminals, coils, and winding surfaces that are capacitively coupled to the ambient space field ΔE_3. Unless coupling capacitance C_3 displays a symmetrical balance to the two signal conductors, a portion of the induced error voltage will be of the normal-mode type. Even an intentional grounding connection on the signal circuit at the central equipment, if not precisely balanced, will cause a fractional part of the common-mode, electrostatically induced voltage to appear as a normal-mode type.

The presence of large-magnitude, common-mode voltages on signal circuits may directly damage or even destroy electronic components and may create troublesome normal-mode error signals if only slight deviations from perfect symmetry occur on signal circuits.

Conduction between units. Normal-mode error voltages are also created by conductive coupling (Fig. 2*c*). For the most part this origin of noise is identical with that commonly referred to as ground loop problems. The error voltage owes its origin to the fact that the reference grounding conductor system is not at a common potential throughout. Consider that one signal-circuit conductor is grounded at the central equipment in location G_1. This would be conventional practice when using coaxial signal circuits and is sometimes used on two-conductor wire circuits. The presence of a second ground connection on the same signal conductor at a different location G_2 will likely result in severe noise in low-level signal circuits. After making the first ground connection at G_1, it will be unlikely that another spot can be found along that conductor with a zero-voltage difference to the adjacent reference ground. When a second ground connection is made along a signal conductor, it forces the voltage difference, which had previously existed here, to vanish and appear as an impedance voltage drop (*IZ* drop) around the loop circuit formed by the second ground connection. *See* ELECTRICAL IMPEDANCE.

By reviewing the character of the impedance elements forming this loop, it becomes quite evident that the signal-circuit conductor between grounding points G_1 and G_2 will principally account for the ground loop impedance. The remainder of the loop will be made up of such low-resistance conductors as heavy building structural members and large pipes. Thus most of the voltage difference which, prior to the second ground connection, had existed between the signal conductor and the adjacent reference ground G_2 has, after the second ground connection, become an impedance voltage drop along only one signal conductor between locations G_1 and G_2. This voltage represents a normal-mode error voltage in the signal circuit.

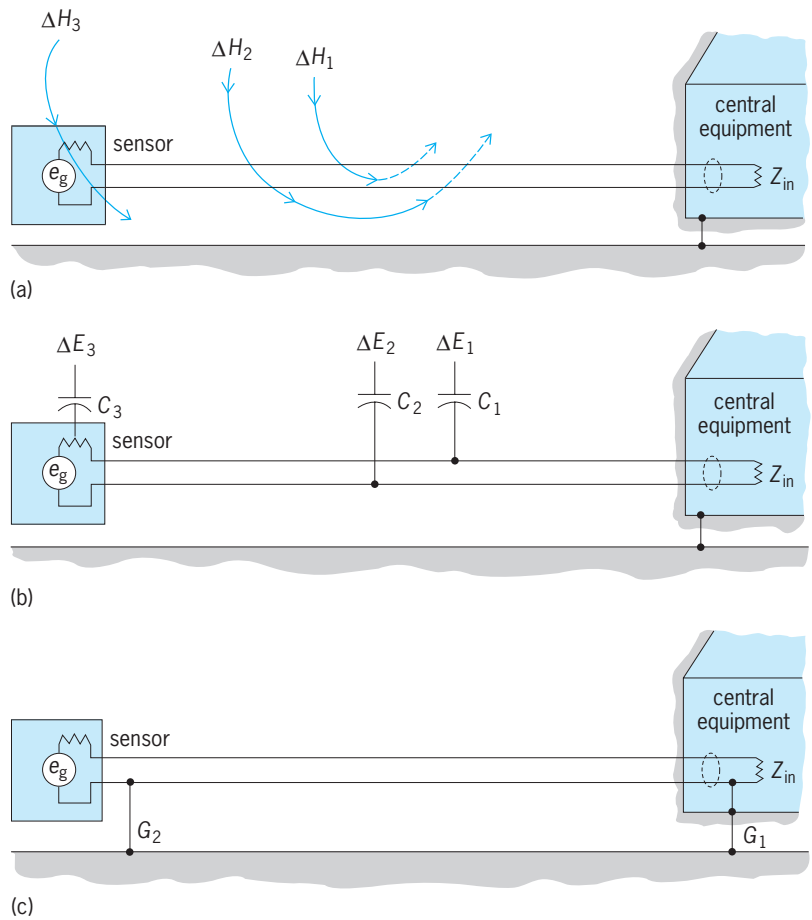

Fig. 2. Common sources of error signals. (*a*) Electromagnetic coupling. (*b*) Electrostatic coupling. (*c*) Conductive coupling.

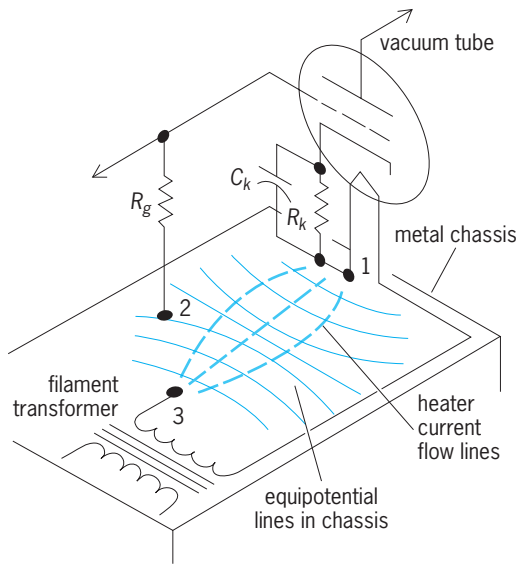

Fig. 3. Improper grounding of an individual amplifier stage to a metal chassis, so that the noise voltage is coupled into the signal circuit.

Conduction within one unit. At times the separation of only inches in the two ground connections can result in intolerable noise, as the following example illustrates. In a vacuum-tube amplifier (**Fig. 3**) the cathode and one heater lead are bonded to the chassis at point 1. One side of the filament transformer is bonded to the chassis at point 3; therefore, heater current flows through the chassis, as shown by the current flow lines. Normal to the current flow lines are equipotential lines, as shown. The grid resistor is bonded to the chassis at point 2. A noise signal e_{12} is thereby inserted into the grid circuit, resulting from the potential drop along the chassis between cathode and grid bonding points. This drop may amount to only a few microvolts, but if it enters an early stage of a high-gain amplifier, excessive noise may be noticed in the output.

To avoid such problems, all bonding connections for a given stage should be made to only one point on the chassis. Also, the flow of large currents through the chassis of any high-gain amplifier, whether vacuum-tube or solid-state, must be avoided. Where high-gain amplifiers are controlling high-power ac or dc circuits having grounded terminals, it is best to insulate the amplifier chassis from the frames of the rest of the machinery and to provide a single ground strap from chassis to frame. This will prevent large currents from passing through the amplifier chassis. *See* CIRCUIT (ELECTRONICS).

In general, each stage of a high-gain amplifier should have its own chassis ground. Each individual amplifier should have its own ground lead to a common amplifier ground point. All machinery and other power devices should be bonded together, with a single lead to the amplifier ground point. Finally, if additional shielding is provided, such as a metal cabinet, it should have a single ground lead to the common ground point for the amplifier and the power equipment. This common ground point would then be grounded to earth by bonding to a cold-water pipe or by other methods in conformance with the local electrical codes.

Ground gradients. The medium- to high-voltage substation may display large voltage differences between the potential on the local station grounding conductors and remote earth. Information-gathering (or sensing) circuits associated with electronic equipment extending into or close to such a substation area will generally require extreme design care to avoid dangerous common-mode potentials. It is not unreasonable to expect steady-state voltage levels of as much as 50 volts. During a line-to-ground fault on an outgoing overhead line, the potential of the substation ground mat might be elevated to several thousand volts, relative to remote mean earth potential by voltage gradients. Common-mode voltages from such causes can be compensated by insulating transformers or equalizing transformers or by other means of bridging large-magnitude, common-mode voltages.

The ordinary commercial building may present unexpected problems relative to common-mode voltage components due to voltage gradients in the building grounding conductor system. It is widely believed that the grounding requirements as spelled out in section NEC 250–23(a) of the National Electrical Code ensure that all electrical load current within the building will be returned to the service equipment (the point of electric service entrance) on power conductors, independent of the building structure.

An innocent-appearing exception can violate this concept. Section NEC 250–60 defines a grounding exception (rather generally used) for the frames of electric ranges, wall-mounted ovens, and counter-mounted cooking units. The code exception allows these appliance frames to be grounded by connection to the power system grounded conductor (white wire) if the electric service is a 240/120-V, single-phase, three-wire system or is taken from a 208Y/120-V, three-phase, four-wire system. It is unrealistic to assume that the appliance frames in question will not also be in contact with the building structure. The result is that the white load-current-carrying conductor becomes connected to the building frame through the heating appliance.

Thus a building devoted essentially to commercial activities, served with 208Y/120-V electric power, may have a snack bar installed on the sixth floor. Contained therein may be appliances grounded to the white wire as allowed by NEC 250–60. Instead of an electrically dry building frame, it may be found that between the sixth floor and the ground floor there is distributed the same electrical voltage drop as exists on the grounded power-system conductor between the snack bar connection and the service entrance.

When the service is three-phase and four-wire, all third-harmonic currents (and their multiples) in the entire array of line-to-neutral connected loads combine in an additive fashion in the neutral conductor (white wire) to aggravate the harmonic voltage drop

along the white wire. In both the three-phase and single-phase power-supply cases, the increasing use of time-modulated (SCR-controlled) current in line-to-neutral connected devices (such as fans and lighting units) makes for the presence of much step-front hash in the voltage drop along the neutral conductor, which becomes voltage gradient in the building structure if the NEC 250-60 exception is employed. *See* SEMICONDUCTOR RECTIFIER.

Efforts have been made to restrict the use of the NEC 250-60 exception to appliance circuits which originate as branch circuits at the service equipment. It must be recognized, however, that the ground reference potential on one floor of an office building may differ from that of another floor by as much as several volts (third harmonic) and contain substantial fast-front hash.

Section NEC 250-23 of the National Electrical Code clearly prescribes that there shall be no grounding connection to the grounded power conductor of the electrical system downstream of the service equipment for that establishment. When this rule is respected, all of the load-system power current is returned to the grounded conductor at the service via insulated conductors. There is no opportunity for electrical noise, as it exists on the grounded power conductor (the white wire) to be conductively transferred to the system grounding conductors. There are three exceptions in the NEC text which may permit that rule to be bypassed. The first one applies to an in-plant separately derived electrical system located remote from the service equipment of the establishment. It is common to apply a permanent grounding connection at the supply-machine neutral. This will create a cross bond between the grounded conductor (white wire) and the grounding conductor downstream of the main service equipment. The second one applies to the case of an electric-supply circuit run to a second independent building. The NEC requires a cross bond between the grounded and the grounding conductor at the point of entry to the second building, unless an independent grounding conductor has been included with the power conductors from the supply point in the main building. The third exception, pertaining to ranges, counter-mounted cooking units, wall-mounted ovens, and clothes dryers, can be troublesome. While the National Electrical Code says merely that the power-circuit grounded conductor (white wire) may be used as the grounding conductor, it is commonly found that the appliance frames so grounded by connection to the white wire are also in contact with building metal frame or metal piping systems. It is this "back door" connection which allows communication of the electrical noise to the establishment grounding conductor system.

Progress is being made in restricting the use of the white wire as a grounding conductor.

A few instances of a cross-bond between the grounded power conductor and the equipment grounding conductor downstream of the service equipment have been observed as a result of an inplant emergency or standby generator. These have been associated with 208Y/120-V and 480Y/277-V solidly grounded power systems serving hospitals, telephone exchanges, police centers, and other such important establishments. When operating on an emergency or standby basis, it is proper that a grounding connection exist at the local generator as prescribed for a separately derived electric power system in section NEC 250-26 of the National Electrical Code. It is common to find an intentional permanent grounding connection at the local-generator neutral junction. Unless the neutral line of the supplied load is switched, along with the phase conductors, the result is a permanent grounding connection on the "white wire" at the generator location, contrary to the planned design pattern. Location of the in-plant generator adjacent to the service equipment would eliminate the problem, as would also inclusion of the neutral conductor in the transfer switching operation.

There will occasionally occur brief intervals of much greater than normal voltage gradients along the building grounding conductor system. An insulation failure on a power conductor may permit a large-magnitude, ground-short-circuit current to flow through the building structure toward the service equipment location for the brief interval permitted by the overcurrent protector.

Importance of low-noise environment. The noise problems presented to the electronic equipment designer increase directly with increased ambient noise levels in the area. The quality level employed in grounding and shielding of power-system conductors has a marked effect on the ambient noise levels.

In an occupied building much can be learned of the general ambient noise levels by a study of the electric facility practices employed and by a general inspection of the building. Additional specific information may be obtained by test. It is important to look for unusual sources of high-energy radiation which may penetrate the building interior space. A nearby radio station antenna may create strong fields at one fixed frequency. High-power radar-scanning installations may be a source of annoying ambient fields of pulsed high frequency.

After an appraisal of the severity of the ambient electrical noise level, the problem of designing the grounding and shielding practices for the electronic equipment becomes a straightforward procedure.

Grounding and shielding practices. To establish an acceptably high signal-to-noise ratio demands that normal-mode noise be maintained at levels below a specified tolerable value.

Choice of conductor. The construction geometry of the signal conductor (**Fig. 4**) plays a prominent part. Of the many possible varieties, the four most commonly used are random-selected wires in a multiconductor cable, parallel-wire twin-lead, twisted-pair, and coaxial lines. Because of the superiority of the twisted-pair relative to random wires or twin-lead at modest cost increase, its use is universal when noise

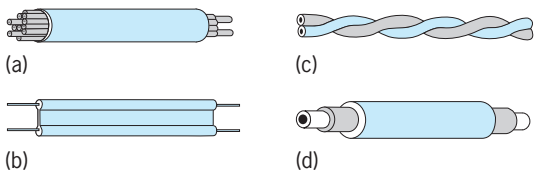

Fig. 4. Common signal-circuit geometries. (*a*) Two random wires. (*b*) Straight twin-lead. (*c*) Twisted pair. (*d*) Coaxial line.

reduction is a factor. A half twist every inch is both typical and effective.

An encircling metal raceway accomplishes substantial additional attenuation, particularly if it consists of steel conduit. In securing high-quality rejection of capacitively coupled noise voltages, it is common practice to avoid any grounding connection on the signal conductor (coaxial lines excepted) and to incorporate a balanced impedance connection to the central-equipment reference ground terminal in the input circuits of a differential amplifier.

The coaxial line (Fig. 4*d*) is a theoretically superior construction but may be faced with operating conditions which deteriorate these qualities. As it is inherently an unbalanced line, the exterior shell is commonly connected to the reference ground terminal at the central equipment. The existence of any connection to adjacent ground along the coaxial line (not easy to avoid) creates the ground loop situation described in Fig. 2*c* with a likely inevitable insertion of normal-mode error signal (noise). Even in the absence of a second ground connection, large-magnitude varying electric fields (Fig. 2*b*) can create measurable values of charging current flowing along the outer shell to the grounding connection at the central equipment, and in so doing can create an *IR* voltage drop along that conductor which will appear as a normal-mode noise voltage. An enclosing metal raceway grounded to the central-equipment ground reference terminal can be used to effectively suppress this noise source.

Conductor shielding. A multiplicity of signal circuits may be run within a common metal raceway, providing that the signal magnitude on these conductors is not great enough to create cross-coupled interference (cross talk). Of the various conductor geometries considered in Fig. 4, the random wire case in *a* would be most susceptible and would offer little hope for correction. With the conductor patterns in *b* and *c* the application of wrapped-on shielding tape (or equivalent concentric conducting shell) to each signal circuit, grounded at the central-equipment grounding terminal, would only accomplish high-quality cross-talk suppression. The coaxial line is inherently free of cross-talk interference if the outer conductor is grounded only at the central equipment.

It is possible (and not unlikely) that the outer terminal sensor may itself respond to ambient magnetic field noise (Fig. 2*a*) or electrostatic field noise (Fig. 2*b*) or both, and create normal-mode, signal-circuit noise. The magnetic field problem can be resolved by the application of appropriate magnetic shielding at the sensor, the enclosure of the complete sensor within a closed shell of conductive metal, or a combination of the two. The electrostatic problem can be resolved by enclosing the sensor winding and its circuit leads within a thin-wall shielding shell maintained at zero reference ground potential. The enclosing shell can be the same one used for magnetic induction suppression. The required connection to a zero potential reference point may require a unique grounding circuit extending from the central equipment (**Fig. 5**).

Control of grounds. The remaining perplexing noise abatement problem relates to the creation of reference grounding terminals at peripheral equipment locations which appear to be at the same potential as the central-equipment grounding terminal; an example is the electrostatic suppression problem discussed in the previous paragraph. An earlier discussion (Fig. 2*c*) develops the fact that between a building ground-reference point G_2 and a widely separated central-equipment grounding terminal G_1 there will, more often than not, appear a difference in electrical potential. The difference voltage can create normal-mode, error-signal voltages in the several ways described.

If the remote terminal G_2 is not in the same building as is terminal G_1, or if it is located in an area where large common-mode noise voltages are expected, the use of insulating or equalizing transformers in the signal lines or the adoption of other common-mode suppression techniques should be considered.

If the remote equipment is in the same metal-frame building or in a location where common-mode voltage difference between reference ground terminals is known not to exceed perhaps 10 V, acceptable attenuation should be possible without the introduction of isolation techniques on the signal wires themselves.

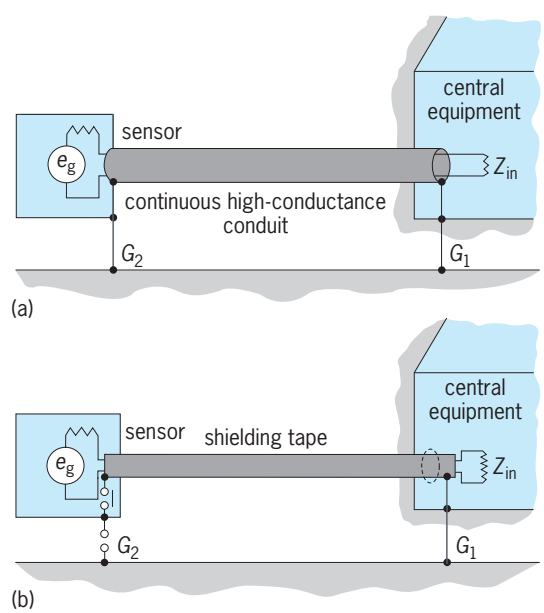

Fig. 5. Practical grounding techniques. (*a*) High-conductance raceway. (*b*) Electrostatic shielding.

The installation of a grounding conductor, run with the signal circuits, which interconnects grounding terminal G_2 with G_1 will be helpful but not as effective as might be expected. A substantial fraction of the original noise voltage will remain in the form of an *IZ* drop along the grounding conductor.

High-conductance raceway. By modifying the form of this grounding conductor, a tremendous gain in attenuation can be achieved. By forming the conductive metal of the grounding conductor into a hollow tube through which the signal conductors are passed, its effectiveness is increased enormously. The resulting tubular grounding conductor, called a high-conductance raceway (Fig. 5*a*), interconnects grounding terminals G_1 and G_2 as did the cable grounding conductor. Voltage difference $e_{G1} - e_{G2}$ appears as an *IZ* drop along the raceway. The conductance (tube diameter, wall thickness, and type of metal) is selected so that only a small portion of the total *IZ* drop appears as a resistance drop along the metal tube. Almost the entire voltage difference between grounding terminals G_1 and G_2 is accounted for by a reactive-voltage drop *IX* created by an induced magnetic field encircling the tubular raceway. This magnetic field also links each signal conductor contained within the raceway and thus induces in each one identically the same *IX* voltage drop as in the raceway. With the *IX* voltage component equalized, or canceled out, only the *IR* drop along the high-conductance raceway appears as a voltage difference between signal conductors and adjacent ground at G_2.

It is important to note that this cancellation of the *IX* drop occurs only on those signal conductors run within the high-conductance raceway. If some critical signal circuits must take a different route, an independent high-conductance raceway to contain them will be needed to accomplish the desired high-magnitude attenuation.

The high-conductance raceway technique is applied extensively to interconnect the several independent housings marking up the central electronics equipment to establish a nearly quiet ground reference potential throughout the central equipment. One prominent item of the central equipment is designated as the reference unit and is bonded to the building reference ground, preferably at only one point. All other equipment frames in the group would, by design intent, permit conductive connection only to the central reference structure in the form of high-conductance raceways. Power conductors and other high-noise-level conductors should be run in raceways independent of the sensitive signal circuits.

In extremely severe situations where the *IR* drop in the high-conductance raceway creates intolerable noise levels, a dual concentric system may be substituted. The exterior high-conductance raceway is identical in design and function with the one just described. Within this is a second metallic tubular raceway, insulated from the former except for an intentional cross bond at the central equipment grounding terminal G_1. At the remote equipment locations, the construction consists of an outer enclosure bonded to the building ground terminal G_2. The electronic equipment chassis containing the active electronics components is insulated from the outer housing and bonded only to the internal concentric raceway. By this modification even the resistance voltage drop along the outer tubular raceway is prevented from appearing as a voltage difference between signal conductors and their enclosing raceway.

Should the remote sensor terminal be small and of simple character, with possible noise being essentially of electrostatic coupling origin, the simple circuit pattern illustrated in Fig. 5*b* may prove adequate. A lightweight shielding tape enclosing the signal conductors preserves a zero-potential reference plane surrounding the signal conductors. The sensor frame (or enclosing electrostatic shield) is connected to the signal-circuit shielding tape and kept insulated from the local grounding terminal G_2. The presence of high-intensity electric field noise at the outer terminal may demand a higher-conductance signal-circuit shield to avoid objectionable noise voltages of *IR* drop origin.

Grounding, electromagnetic shielding (and cancellation), electrostatic shielding, and insulation are intimately intertwined in effecting an acceptable solution.

Data transmission. To an ever-increasing extent the high-density channels of information handling (data transmission) are being accommodated by means other than metallic conductors. The prominent transmission avenues are microwave links and optic fibers. The problems of circuit grounding along these links are simply nonexistent. The interconnecting wire circuits between the distributed sensors and the central processing center, and between the various functional systems at the center, and the transmission to the output terminal at that location will face the same signal-to-noise ratio problem as in the past. *See* MICROWAVE; OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

Optical isolator. In dealing with the aggravated problems associated with the large magnitudes of common-mode noise voltage encountered with circuits run to an outdoor electric substation, the optical isolator is a valuable tool. This device introduces a short optical transmission path into the electric signal circuit with a capability of withstanding several thousand volts of common-mode voltage without ill effects. *See* OPTICAL ISOLATOR. R. H. Kaufmann

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; Institute of Electrical and Electronics Engineers, *IEEE Recommended Practice for Powering and Grounding Sensitive Electronic Equipment*, IEEE Std. 1100-1992, 1992; M. Mardiguian, *Grounding and Bonding*, 1988; R. Morrison, *Grounding and Shielding Techniques in Instrumentation*, 1986; R. P. O'Riley, *Electricial Grounding: Bringing Grounding Back to Earth*, 5th ed., 1998.

Electronic listening devices

Devices which are used to capture the sound waves of conversation originating in an ostensibly private setting in a form, usually as a magnetic tape recording, which can be used against the target by adverse interests.

There are two kinds of electronic listening devices. One takes advantage of equipment already present on the target's premises, such as a telephone, radio, phonograph, television set, public-address loudspeaker, or tape recorder, to act as a microphone, transmitter, or power supply. The other does not. In the former case, the target's equipment is said to have been compromised.

These practices are unlawful in the United States and Canada except when carried out by law enforcement officers acting under authority of a warrant. In the United Kingdom, electronic eavesdropping by private parties may contravene the laws against trespass and unauthorized use of telephone equipment or radio frequencies, but it is not unlawful in and of itself. The same is generally true in western European countries.

Compromise. Compromise of the target's own equipment takes advantage of the fact that any loudspeaker is capable of functioning just as well as a microphone, that convenient sources of dc power are available within the equipment, or that the equipment is connected to power or signal lines that can transmit the intercepted conversation to some place where recording can conveniently be accomplished.

The equipment most frequently compromised is the telephone handset. The act of compromise may be as simple as bypassing the switch hook with a Zener diode so that the instrument can transmit conversations to a wiretap, or high-impedance parallel connection off the premises, while not signaling a busy tone to someone attempting to dial in. *See* TELEPHONE; ZENER DIODE.

It can be as complex as the infinity transmitter, which is activated by the eavesdropper's dialing in and transmitting, before the telephone rings, an audio tone to a tuned relay concealed in the instrument.

Bugs. Eavesdropping devices that can stand alone are known commonly as bugs. They take advantage of many developments of modern technology, such as microcircuits, miniature ceramic microphones, and miniature batteries. *See* INTEGRATED CIRCUITS; MICROPHONE.

The art of designing bugs has achieved its most advanced state of development in the national intelligence services of the Great Powers. One bug was discovered to have been inserted in the heel of a diplomat's shoe.

The smallest bug available to the private citizen is probably the Hong Kong "spider." It is no larger than a common postage stamp and less than a quarter-inch thick.

Electronically a bug is often just a two-stage frequency-modulated transmitter: an audio amplifier and a variable-frequency radio-frequency (rf) oscillator. *See* AMPLIFIER; OSCILLATOR; RADIO TRANSMITTER.

Bugs may operate on any frequency from 20 to 1000 MHz, but usually they are designed to operate at a frequency close to that of a powerful local frequency modulation (FM) or very high-frequency (VHF) television station.

A popular hybrid between a compromise device and a bug is the telephone drop-in. In this design, an FM transmitter is made in the form of a telephone microphone. The eavesdropper can casually unscrew the mouthpiece of the target's telephone handset and substitute the drop-in for the original microphone. The range of this device is about 250 ft (75 m). It has the added advantage of drawing its dc power from the telephone company central battery.

Defenses. The telephone line analyzer is used to defeat compromise devices. Under the control of a microcomputer, the line analyzer examines in turn each of the 50 or more telephone lines used by a typical commercial firm. It can detect the tiny amount of current flow into a Zener or other compromise device. A monotonically rising audio signal is then applied to trigger any infinity transmitter. If a compromise device is discovered, a pulse of 800 V is impressed upon the line to burn it out.

Technical surveillance sweeps, as they are known to the trade, are performed alternately with external signal lines connected and disconnected and with telephones on-hook and off-hook.

Often a sniffer is used to sense the presence of a bug. A sniffer may be a simple VHF diode detector with an output meter that reveals the presence of a local carrier. More sophisticated instruments sweep the frequency bands of interest and compare the audio output with locally generated signals, usually a tape recording of simulated business activity. Another type of sniffer affords discrimination between a genuine clandestine device and a strong local broadcasting station.

If the presence of a clandestine device is indicated, the sweep team may employ a panoramic intercept receiver. This instrument displays the amplitude of received signals on a cathode-ray-tube (CRT) trace calibrated in units of frequency. The frequency of the bug can be determined by observing which of the pips on the CRT trace appears to diminish and grow in direct response to the locally generated audio signal. Once its frequency is determined, the bug's physical location can be found by use of a tuned rf field-strength meter having a loop antenna. *See* ANTENNA (ELECTROMAGNETISM); CATHODE-RAY TUBE.

The objective of a sweep such as the one described is to locate the device and discover its technical characteristics with a view toward feeding the device false information that may discredit or compromise the individuals responsible for planting it.

National security forces encounter clandestine devices which are switched on and off by an agent using a control transmitter. Such a device cannot

be detected by the methods described above because they presuppose that the bug is continually on the air or can be activated by locally generated audio.

To detect silent bugs, advantage is taken of the fact that all subminiature transmitters contain one or more semiconductor junctions that are connected to an electromagnetic radiator. Inasmuch as these junctions are nonlinear impedances, a locally generated low-power ultra-high-frequency carrier will be reradiated by the bug, and the reradiated signal will contain strong harmonic components. *See* HARMONIC (PERIODIC PHENOMENA).

Natural devices such as rusty bedsprings will reradiate second-harmonic components, but only a real semiconductor junction will reradiate third (and higher) harmonics.

An ultra-high-frequency (UHF) sweeper is equipped with a meter that produces a negative deflection in the presence of reradiated second-harmonic energy and a positive deflection in the presence of reradiated third-harmonic energy. Such a deflection discloses the presence of a bug, albeit a silent one not detectable by sniffers or panoramic receivers. The UHF sweeper is portable; it bears a resemblance to a household appliance for vacuuming drapes. Increasing positive deflection indicates to the sweep team member that the bug is being approached. *See* HARMONIC ANALYZER.

Advanced devices. Human capacity for designing electronic instruments to spy on one's neighbors appears to be boundless. Some of the more exotic include laser radar which responds to infinitesimal motion of windowpanes caused by conversation in a room, bugs which include miniature television cameras, and minicomputers which are programmed to intercept messages to a resource-sharing computer and to return answers which are spurious.

The so-called pinhole camera, a television camera that uses charged-coupled devices, is less than 2 in. (5 cm) long. Its fast (*f*/1.8) lens provides the capability for excellent infrared vision and its wide-angle (11-mm focal length) characteristics permit viewing a whole room through a pinhole, or a hole no wider than 0.5 in. (13 mm) if a fiber-optic extender must be used to look through a thick wall. *See* CHARGE-COUPLED DEVICES; LENS (OPTICS); TELEVISION CAMERA.

Typical hiding places are in lighting fixtures, radios, television sets, and illuminated signs, and behind switch plates or electrical receptacles. Miniaturized video distribution systems are available that will transmit signals up to 4000 ft (1.22 km) over 24-gage twisted pair wires. In most jurisdictions these devices have not yet been declared illegal. John M. Carroll

Bibliography. Act IV Security Services Staff, *Surveillance Countermeasures*, 1994; B. Bruno, *Serious Surveillance for the Private Investigator*, 1992; S. French, *The Big Brother Game*, 1976; S. Gowrinathan and J. Shanley (ed.), *Surveillance Technologies II*, 1992; L. Lapin, *How to Get Anything on Anybody*, 1987, bk 2, 1991.

Electronic mail

The asynchronous transmission of messages by using computers and data-communication networks. Historically, electronic mail (or e-mail) referred to any of a number of technologies that allowed people to send documents to one another through electronic means. It was frequently used to describe both wirephoto [the precursor of the facsimile (fax) machine] and telegraphy. Subsequently, usage of the term focused upon the narrower sense given above. *See* FACSIMILE; TELEGRAPHY.

Early systems. The term e-mail began to disseminate in the early 1970s, when time-sharing on large mainframe computers became common. Many systems had the ability to send messages from one time-sharing user to another on the same system by directing it to the receiver's sign-on name. *See* MULTIACCESS COMPUTER.

With the creation of modems capable of being dialed under computer control, the relatively inexpensive creation of e-mail networks became possible. This development frequently was fairly haphazard, one computer dialing others a few times a day and transferring all the messages that had accumulated since the last phone call. As this process was repeated by all the cooperating computers, messages would get to their destinations. *See* MODEM.

Sending messages among a number of machines over dial-up connections required that the addressing scheme be expanded to give a computer identifier as well as the receiver's sign-on name. Since there was no central routing database, users were forced to route messages by hand within the address. For example, a!b!c!smith would send the message from system a to b, from b to c, and there deliver it to user smith on system c. In fact, there could be multiple computers named c, but only one to which b knew how to route mail.

Impact of personal computers. The availability of the personal computer in the 1980s further advanced the technology in three ways. First, since computers became more accessible, e-mail service bureaus became commercially viable. These e-mail systems used the mainframe messaging technology centrally and allowed subscribers to call in via the long-distance telephone network to send and receive their messages.

Second, personal computers coupled with local-area networks produced a number of proprietary e-mail networking systems for work groups. These systems offered better user interfaces and more functionality than the more general electronic-mail solutions at the expense of demanding uniform systems throughout the e-mail network. That is, work groups could communicate with each other only if they were all running identical software. *See* LOCAL-AREA NETWORKS.

Finally, there have always been two distinct actions in the delivery of e-mail: the delivery across the network to the recipient's mailbox, and the recipient's removing a mail item from the box to read it. Originally, both these actions would occur on the

same service bureau computer, where the recipient would connect online to the mail server and read the mail while connected. This forced the e-mail user to learn a completely different computing environment to read or send e-mail and to remain connected to the server while reading or writing messages. With the advent of personal computers with window-based point-and-click user interfaces, the operations of being a mail repository and a mail interface were split. New "mail clients" such as Eudora, Exchange, and Netscape Mail were created. These used special access methods such as POP (Post Office Protocol) or the newer IMAP (Internet Message Access Protocol) to manipulate mail remotely, including moving mail en masse to a recipient's workstation. Once there, the recipient could read it at leisure without being connected to the network. Outgoing messages would be kept until the next time the computer was connected to the Internet, when they would be sent on their way. This allowed the messages to be processed in a variety of new ways such as threading. Threading presents messages in chronological order within a topic rather than just chronologically. This allows a complete topic to be read (or discarded) as a unit. *See* HUMAN-COMPUTER INTERACTION; MICROCOMPUTER.

Message identification and routing. The use of electronic mail grew continuously until the late 1980s but never achieved widespread use outside of work groups or corporations. The limiting factor was the complicated addressing that had to be worked out before a message could be successfully transmitted.

There were two proposed methods to solve the problem of mail-system identification and routing. The Organization for International Standardization (ISO) formulated the X.400 standard, and the Internet community developed an extended use of the domain name system (DNS). Many impediments to the spread of X.400, such as high software costs and delays in standardization, caused the freely available DNS solution to become the de facto standard.

The DNS describes a worldwide distributed database in which each site maintains its own information about how to route messages to a computer within its administrative domain. A computer wishing to send a message to another asks the DNS for the routing information and uses the information returned to make the connection. This allows a person on virtually any online networking service to send mail to another person by giving only the personal identification and the e-mail system name of the recipient. *See* DISTRIBUTED SYSTEMS (COMPUTERS).

Multimedia transmission. From the time the usage of the term narrowed to exclude facsimile until the early 1990s, generally only coded textual information could be transferred via e-mail. The transmission of nontextual data required special preprocessing, postprocessing, and prior arrangements between the sending and receiving parties. It was very difficult to make these kinds of transfers if the sending and receiving computers were different types.

This restriction was lifted with the adoption of the MIME (Multimedia Internet Mail Enhancements) standard. It described a way of encoding an arbitrary list of media types within a normal textual message in an operating-system-independent manner. Finally, different types of systems could send executable, sound, picture, movie, and other kinds of files to each other via e-mail. *See* MULTIMEDIA TECHNOLOGY.

Attachments. The sender's mail package can attach a file to a standard text message. When the message is sent, the attached files are encoded into a single message. The receiver's e-mail package will decode the message into a base message plus an attachment for each file. The recipient's computer still must have the software required to view or process the file, but the transmission is transparent.

Security. The spread of electronic mail was also hampered by its lack of security. As mail was passed from one site to another closer to its destination, system administrators at each intermediate site could read messages. Also, the source of an e-mail message may be fairly easily forged to make it either untraceable or appear to come from another person. This limited the use of e-mail to so-called friendly applications. Public-key cryptography has been applied to e-mail messaging, notably in PEM (Privacy Enhanced Mail), in response to these security concerns. *See* COMPUTER SECURITY; CRYPTOGRAPHY.

Integration of messaging technologies. Many e-mail-to-fax gateways are in use. These allow messages that originate as e-mail to be converted to fax and then delivered via office fax machines to recipients that do not normally participate in the e-mail community.

Since the communications speeds required for e-mail are quite modest, messages are sometimes transmitted by radio. In the simplest applications of this technology, some pocket paging system providers allow short e-mail messages to be sent to their pagers. In more advanced applications, providers lease notebook-sized systems which can send and receive e-mail through Earth-satellite relay. *See* RADIO PAGING SYSTEMS.

Some voice-mail systems accept e-mail for their clients. The text of the e-mail is passed through a speech synthesizer to artificially read the e-mail text into voice storage. Clients can then hear their e-mail by accessing the voice mail system as they would for voice messages.

Infringement by other technologies. From the time that e-mail was developed until the late 1990s, it was the predominant way of staging a universally available discussion. This was done with a variety of mail forwarding utilities, such as LISTSERV, which would forward a message to a large group of discussion subscribers. Although this method is still in use and has some advantages in circumstances where the asynchrononous nature of e-mail is useful, some of these discussions have moved to various chat facilities on the World Wide Web and other technologies. These chat facilities provide a model of a room for a particular topic, and anything typed by someone in the room will be seen in close to real time by

others connected to the room. *See* COMPUTER; DATA COMMUNICATIONS; DIGITAL COMPUTER; ELECTRICAL COMMUNICATIONS; INTERNET; WORLD WIDE WEB.

Edward Krol

Bibliography. B. Shimmin, *Effective E-mail Clearly Explained: File Transfer, Security, and Interoperability*, AP Professional, 1997; D. Strom and M. T. Rose, *Internet Messaging: From the Desktop to Enterprise*, Harcourt Brace, 1998; D. Wood, *Programming Internet Email*, O'Reilly & Associates, 1999.

Electronic navigation systems

Systems that locate suitably equipped users by means of measurements made by electronic sensors. Typical measurements include distance or bearing from a vehicle to a known point, or direct determination of the present position of the vehicle in a particular coordinate system. From the knowledge of present position, the course and distance to a destination at a known location can be calculated. Most systems are based on the use of electromagnetic (radio) or acoustic waves or inertial sensing. The systems described here are typically used by aircraft, ships, and land and space vehicles. In addition, the use of satellites for land navigation by individuals has become common.

The use of radio waves has been found to be attractive because of their known and nearly constant velocity of propagation, namely, the speed of light (c), which is about 3×10^8 m/s (1.86×10^5 mi/s). Thus, if the time (t) of travel of the radio signal between two points is accurately measured, the distance or range (d) between the points can be accurately determined from $d = ct$. Ships also use systems based on underwater sound waves, with due allowance for the much slower speed of sound in water. Also, electromagnetic waves in the visible (optical) or near-visible spectrum can be used in a similar manner for distance measurement. Antennas designed for reception of radio signals of sufficiently high carrier frequency can also be used to produce tightly focused radiation patterns that can be used in measuring the bearing (angle) between two points. *See* ELECTROMAGNETIC RADIATION; LIGHT; SONAR; UNDERWATER NAVIGATION; UNDERWATER SOUND.

Parameters. The parameters that characterize the utility of an electronic radionavigation system include accuracy, coverage, fix dimensions, system capacity, signal characteristics, and spectrum.

Accuracy. In navigation, the accuracy of a measured position at a given time is the difference between the indicated position output of the system and the true position at that time. For horizontal accuracy, the circular error probable (CEP) is often used to statistically characterize the random system errors. The CEP represents the radius of a circle that contains 50% of the position fixes. For three-dimensional systems, the spherical error probable (SEP) is often used to statistically represent the random system errors. The SEP represents the radius of a sphere centered at the origin that contains 50% of errors.

Position accuracy can further be characterized as follows:

1. *Repeatable accuracy*: The accuracy with which a navigator can return to a location, the coordinates of which have been previously measured using the same onboard system.

2. *Absolute accuracy*: The accuracy with which a navigator can determine position in terms of an agreed-upon coordinate system distinct from the navigation system. A common coordinate system is the Earth Center Earth Fixed coordinate system, which is a Cartesian coordinate system whose center is nominally the Earth's center of mass, whose X axis goes through the intersection of the equatorial plane and the prime meridian (which passes through Greenwich, England), and whose Z axis goes through the North Pole. In conjunction with a model of the Earth's shape such as that defined by the WGS-84 ellipsoid (defined in the world Geodectic System 1984, in which positions of satellites in the Global Positioning System are specified), geographic latitude, longitude, and height with respect to the ellipsoidal surface may be determined. *See* GEODESY.

3. *Relational accuracy*: The accuracy with which a navigator can determine position with respect to another user of the same system. For example, if two ships both having the same navigation system were to rendezvous, the relational accuracy would determine their ability to meet at a common point.

Coverage. The coverage of a navigation system is that surface area or space volume in which the system is adequate to permit the navigator to determine position to a specified level of accuracy. Most navigation systems have been applicable to aircraft, ships, or land vehicles, and coverage has classically been defined in terms of these applications. More generally, coverage should include applicability to space vehicles, submersible craft, and indoor applications.

Fix dimensions. This characteristic defines whether the navigation system provides a linear, one-dimensional line-of-position, or a two- or three-dimensional position fix. The ability of the system to derive other quantities such as velocity, time, or attitude is also included.

System capacity. System capacity is the number of users that a system can accommodate simultaneously.

Signal characteristics. This parameter refers to signal power levels, frequencies, signal formats, coding sequences, data rates, and other characteristics necessary to define the means by which a user derives navigation information.

Spectrum. Spectrum refers to the frequency allocation of the signal. Spectrum allocation issues are addressed by the Federal Communications Commission (FCC) and National Telecommunications and Information Administration (NTIA) within the United States. International spectrum issues are addressed in conjunction with analogous agencies from other countries during the biannual World Radio Conferences. The spectrum allocation of current

Electronic navigation systems

System	Type[a]	Frequency, MHz	Accuracy (absolute)[b]	Number of stations[c]	Coverage, km (nmi)
Loran C	H	0.100	460 m (1500 ft), 2 drms	60	2200 (1200)
Beacons	R	0.200–1.6	±3° to ±10°, 2 sigma	5000	370 (200)
VOR	R	108–118	±1.4°, 2 sigma[d]	2000	185 (100)
DME	C	960–1215	185 m (600 ft), 2 sigma	1000	185 (100)
TACAN	R, C	960–1215	±1.0°, 180 m (600 ft)	800	185 (100)
ILS (azimuth)	R	108–112	±7.6 m (±25 ft), 2 sigma[e]	2000	33 (33)
ILS (elevation)	R	329–335	±2.1 m (±7.0 ft), 2 sigma[e]	2000	33 (18)
MLS (azimuth)	R	5000–5250	±4.0 m (±13 ft), 2 sigma[e]	44	37 (20)
MLS (elevation)	R	5000–5250	±0.6 m (±2 ft), 2 sigma[e]	44	37 (20)
MLS (DME)	C	960–1220	30 m (100 ft), 2 sigma	44	37 (20)
JTIDS-RelNav	C, PR	960–1215	Variable	Variable	Variable
PLRS	C	420–450	15–25 m (50–80 ft)	Variable	Variable
SSR	R, C	1030,1090	0.16°, 38 m (125 ft)	800	110–370 (60–200)
SSR/Mode S	R, C	1030,1090	0.04°, 7 m (24 ft)	150	110–370 (60–200)
TCAS	R, C	1030,1090	Variable	3500	Variable
GPS	PR	1575,1227	13 m (43 ft) horizontal; 22 m (72 ft) vertical	24[f]	Global
GLONASS	PR	1605–1250	30 m (100 ft)	14[f]	Global
Altimeter	S	4200–4400	2% of altitude		
Mapping radar	S	9375,9310[g]	Variable		450 (250)
Map matching	S	Various	Variable		
Doppler radar	S	13,325	0.2% or 1.0 m/s (2 knots)		Global
Inertial	S		0.25–0.5 m/s (0.5–1 knot)		Global

[a]H = hyperbolic; R = radial; C = circular, lines of position; PR = pseudoranging; S = self-contained; D/H = Doppler history.
[b]2 drms (twice distance root mean square) is the 95–98% probability, horizontal position error; 2 sigma is the 95% probability one-dimensional error.
[c]Worldwide implementation in 2005.
[d]Signal (flight check) accuracy.
[e]At Runway Threshold Approach Reference Datum.
[f]Satellites.
[g]Beacon frequency.

navigation systems spans the range from very low frequency to C band. *See* RADIO SPECTRUM ALLOCATIONS.

System classification. From an implementation viewpoint, all navigation systems can be classified as either cooperative or self-contained where cooperative systems exploit external radio or acoustic location signals. Examples of cooperative systems include the very high frequency (VHF) omnidirectional range (VOR), distance-measuring equipment (DME), TACAN, the instrument landing system (ILS), and the Global Positioning System (GPS). Examples of self-contained navigation systems include inertial navigation, Doppler speed-heading navigation, and data-base reference navigation systems (DBRNS). Alternatively, from a navigational viewpoint, systems are frequently classified as either positioning or dead-reckoning systems. Positioning systems determine an instantaneous location fix, whereas dead-reckoning systems accumulate position changes from a known initial location. Most positioning systems are cooperative systems, whereas most dead-reckoning systems are self-contained. The self-contained dead-reckoning systems may use electromagnetic or electromechanical technologies (Doppler radar, gyroscopes, accelerometers, or pressure transducers) to make observations from which the distance traveled and direction from the initial location are determined by continuous mathematical integration of velocity or acceleration. Many modern systems combine the data from cooperative and self-contained sensors to obtain a more accurate solution. Such systems have been called multisensor, integrated, or hybrid systems. *See* ACCELEROMETER; DEAD RECKONING; GYROSCOPE; PRESSURE TRANSDUCER.

The major electronic navigation systems (see **table**) will be discussed.

Self-Contained Systems

These systems can be classified as radiating or nonradiating. Radiating systems may be subject to jamming and to homing by radiation-seeking missiles, although the effects of jamming on the systems discussed below are typically small. *See* ELECTRONIC WARFARE; HOMING; JAMMING.

Radiating systems. These include the radar altimeter, airborne mapping radar, map matching, and Doppler radar.

Radar altimeter. This system (also known as radio altimeter) is a small radar with an antenna mounted on the bottom of an aircraft generating a beam toward the Earth's surface. The signal backscattered by the surface is received and processed by the radar, which generates a reading of altitude above the surface. *See* ALTIMETER; GROUND PROXIMITY WARNING SYSTEM.

Doppler radar. Radio waves that are transmitted from a moving aircraft toward the ground and backscattered to the aircraft experience a change of frequency, or Doppler shift, directly proportional to the ground speed of the aircraft. A Doppler radar consists of a transmitter and receiver and a frequency tracker that measures and tracks the Doppler shift of the signals in each of three or four antenna beams directed toward the Earth's surface. The time

integral of the Doppler frequency is proportional to the change in position during the time interval. The Doppler velocity measurement is input to a dead-reckoning navigation computation.

Doppler radar navigation systems are widely used on military helicopters. They have also been used in soft lunar and planetary landings. The Doppler velocity measurement concept has been incorporated into airborne search and mapping radars and is also used in sonar systems to measure the ship's velocity. *See* DOPPLER EFFECT; DOPPLER RADAR.

Airborne mapping radars. These radars scan the ground by using specially shaped beams that effectively map the terrain. They can recognize certain features, for example rivers and bridges, and use them for manually or semiautomatic position-fixing updates to a data-base reference navigation system. The output of an airborne mapping radar or a radar altimeter can be used to generate a map or terrain profile, which is then compared with onboard stored maps or terrain profiles to allow an aircraft or missile to automatically fly a prescribed track. *See* AIRBORNE RADAR; GUIDED MISSILE.

Data-base reference navigation systems (DBRNS). A DBRNS utilize a priori maps of a quantity such as bathymetry, topography, gravity, magnetic intensity, radar, or visual images to determine position by correlating or "map matching" to the corresponding sensor measurement aboard the vehicle. Most often they utilize a dead-reckoning navigation system such as an inertial navigator to provide relative navigation between the measurements and thereby correlate a profile of the measurements to the map. The position fix, used to reset an inertial navigation or Doppler dead-reckoning system, is often registered to the midpoint time of the profile. Examples of a DBRNS include bathymetric navigation, digital scene matching area correlation (DSMAC), terrain contour matching (TERCOM), and precision terrain aided navigation (PTAN).

In bathymetric navigation, a fathometer is used to determine a ship's depth below keel. A profile of these measurements is compared with the dead-reckoning system's indicated profile as referenced to the stored a priori map. Accuracy is better than 200 m (650 ft) when accurate maps based on surveys, for example using GPS position as a reference, are available. Coverage depends on the availability of accurate maps and the bathymetry having adequate variations and asymmetry. Flat ocean areas are not amenable to bathymetric navigation. Sonar frequencies of 10–20 kHz are employed.

DSMAC is a target-area missile guidance system. DSMAC utilizes maps made from photographs of the target area that are digitized. Once in the vicinity of the target, DSMAC will match the digitized photograph with the its real-time camera images and correct the missile guidance inertial navigation system. Accuracy is better than 100 m (330 ft).

TERCOM uses radar and barometric altimetry to determine a three-dimensional position by correlating measurement terrain profiles with prestored terrain map profiles. It is analogous to bathymetric navigation above, with the principle application being cruise missiles. Accuracy is better than 200 m (650 ft). Coverage depends on the availabillity of digitized land areas and the existence of adequate terrain signature.

PTAN is a planned system similar to TERCOM but with expanded coverage. Expanded coverage stems from terrain maps being available from the space shuttle or other space vehicles. PTAN is an alternative guidance technology to GPS for long-range cruise missiles.

Nonradiating systems. Nonradiating systems, such as inertial sensors, offer essentially complete protection against jamming. Inertial navigation equipment based on Newton's second law can determine aircraft acceleration and direction of travel. Acceleration may be determined by measuring the deflection of a spring attached to a known mass. The orientation of the accelerometers is determined by gyroscopes. The acceleration is then doubly integrated in the coordinate frame defined by the gyroscopes to determine the distance traveled since the last position fix. Thus, inertial navigation systems, consisting of accelerometers, gyroscopes, and computers, continuously determine position, velocity, acceleration, direction, and vehicle attitude (pitch and roll). *See* INERTIAL NAVIGATION SYSTEM.

Cooperative Systems

The two general categories of cooperative radio systems are point-source systems and multiple-source systems.

Point-source radio systems. Point-source systems typically determine position in a relative coordinate system by measuring the distance and bearing to a transmitting source at a known location. They may determine distance only or bearing only, if these are the only desired parameters. *See* RHO-THETA SYSTEM.

Bearing measurement. Perhaps the earliest form of a point-source radio system is the direction finder, in which the signal from a single transmitter source is received at two known points or elements. The direction from the vehicle to the source is determined by measuring the differential distance (as indicated by an electronic phase difference) of the signals received at the two points or elements. For operational convenience (size), it is desirable to have the two receiving points close together and to use common circuitry at both measuring points. A loop antenna fulfills both of these requirements. The loop is rotated until the currents in the two vertical arms of the loop are equal in amplitude and phase so that the output of the receiver is zero. The transmitter source is then located 90° from the plane of the loop. In simple loops, there is a 180° ambiguity, but this is typically resolved by temporarily connecting an omnidirectional antenna to the receiver input. As sources for shipboard and airborne direction finders, beacons are the oldest and most numerous navigation aids. Since the frequency bands of beacons are adjacent to the amplitude-modulation (AM) broadcasting band, receivers are easily designed to serve a dual purpose, and they are consequently widely used on

small boats and aircraft of all types. *See* DIRECTION-FINDING EQUIPMENT.

The angular measurement errors, particularly on aircraft and ships, can be quite large (3° or more). Furthermore, lateral navigational error decreases as the transmitter source is approached, a property common to all point-source angle-measuring systems. However, direction finders are still used, particularly for backup navigation and emergency homing, since only a single transmitter source (or beacon) on the ground and a simple rotating antenna and receiver on the vehicle are needed for operation.

Another class of point-source angle-measuring systems is based on the use of rotating or scanning antenna beams at the transmitter source and reception (or reflection) of the transmitted signal by the user vehicle receiver. For example, if a ground transmitter generates a rotating cardioid amplitude pattern at a fixed rate plus an omnidirectional reference signal, the user receiver can measure the relative phase difference between these two signals and thereby determine the bearing to the transmitter source. The very high frequency (VHF) omnidirectional range (VOR), used worldwide for short-distance aircraft navigation, is based on this principle. The bearing function of the military TACAN system is also based on this principle, except that in that system the source transmitter also generates a multilobe pattern for higher accuracy and the resulting ambiguity is resolved by an auxiliary technique.

VOR is the internationally standardized en route navigation system. It meets the Federal Aviation Administration (FAA) en route integrity requirements of a warning within 10 s if there is a failure of ground equipment. Many ground stations have converted to the Doppler mode, which reduces site error due to multipath. *See* DOPPLER VOR; VOR (VHF OMNIDIRECTIONAL RANGE).

Another point-source technique for bearing measurement is based on using a radiated dual-lobe structure, with the carrier in each lobe being modulated with different frequency tones and phases. The user receiver can then detect when the vehicle is operating at the intersection of the two beams. The instrument landing system (ILS), which is used worldwide for aircraft approach and landing, is based on that principle.

ILS is an internationally standardized system that provides a fixed electronic path at major runways throughout the world. It can be affected by close-by buildings and hills and does not allow curved approaches. *See* INSTRUMENT LANDING SYSTEM (ILS).

The known shortcomings of ILS with respect to obscuration and restriction to straight glide paths have led to the development of the microwave landing system (MLS). The use of high microwave frequencies, the nature of the ground-based antenna patterns, and the inclusion of a high-precision version of DME provide higher-accuracy performance and a capability for curved approaches. In view of the promising results with Local Area Differential GPS, also known as Local Area Augmentation System (LAAS), the development of MLS has been terminated in the United States. MLS is installed in about two dozen major airports in the United States and is scheduled to begin decommissioning in 2010. However, because of the unique requirements in the European airspace, acquisition and installation of MLS stations may proceed in European countries. Also, modular combined ILS/MLS/GPS receivers for approach and landing of aircraft are available. The U.S. military services have developed a mobile MLS (MMLS) that can be transported easily.

Secondary surveillance radar (SSR) is an outgrowth of the military Identification of Friend and Foe (IFF) system, using the same frequencies. It is the principal means by which air-traffic controllers identify and track civil and military aircraft worldwide. Beacon ground stations interrogate airborne transponders at 1030 MHz and receive replies at 1090 MHz in order to measure distance using two-way (round-trip) ranging and aircraft bearing using the SSR's narrow antenna beam. The replies are pulse-coded with identity (Mode A) and altitude (Mode C). Mode S, which has been added to the system, provides higher angular accuracy through monopulse techniques, discrete addressing of each aircraft, more efficient code modulation, and a much-higher-capacity data-link capability. *See* MONOPULSE RADAR; RADAR; SURVEILLANCE RADAR.

The Traffic Alert and Collision Avoidance System (TCAS) airborn equipment allows users to interrogate existing unmodified SSR 76 units and tracks the replies in range, bearing (in TCAS II), and altitude. It generates crew alerts if both range and relative altitude are predicted to be so small that a collision threat exists. Two types of alerts are displayed: traffic advisories (TAs) and resolution advisories (RAs). The resolution advisory shows a vertical escape maneuver (for example "climb") for avoiding a collision. TCAS II uses Mode S, air-to-air, data-link communication for exchanging escape maneuver intentions between aircraft. The simpler TCAS I version, intended for smaller aircraft, does not display resolution advisories. It tracks proximate aircraft in range and altitude only, and displays traffic advisories for intruders that represent a collision threat based on relative range and range rate. *See* AIRCRAFT COLLISION AVOIDANCE SYSTEM.

Distance measurement. Theoretically, there are two possible types of point-source systems for distance determination. One is based on the two-way (round-trip) ranging principle (**Fig. 1**). The interrogator, which may be located on the navigating vehicle, transmits a signal, typically a pulse or pair of pulses, at a known time, which is stored in the equipment. The signal is received at a transponder and, after a fixed known delay, is retransmitted toward the interrogator. When received by the interrogator's receiver, it becomes an input to the ranging circuit. This circuit measures the time difference between the original transmission time and the time of reception (less the known fixed delay), which (when multiplied by the speed of light) is a direct measure of the two-way distance.

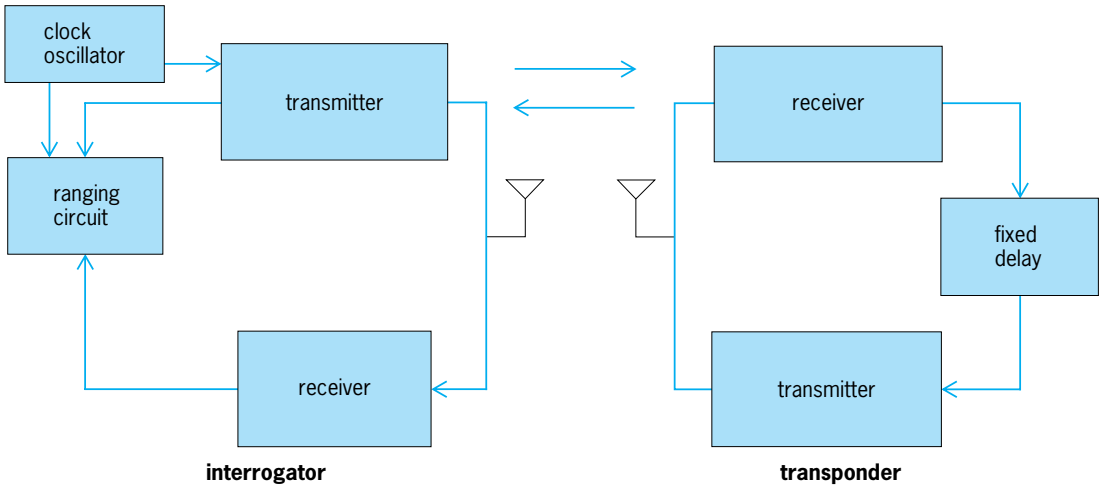

Fig. 1. Two-way (round-trip) ranging system.

An important advantage of this technique is that the signal returns to the point of initial generation for the time-difference measurement process. Therefore, the interrogator's clock oscillator need not be highly stable, since the error due to clock instability depends only on the round-trip time multiplied by the clock drift, and the round-trip time is very short since the signal travels at the speed of light.

Such a system is the basis of the distance-measuring equipment (DME), used for short-range navigation worldwide, and also of the Air-Traffic Control Radar Beacon System (ATCRBS), used for air-traffic control surveillance on a global basis. If the transponder is replaced by a passive reflector, for example an aircraft, the principle of operation is that of all primary surveillance and tracking radars, such as those used for air-traffic control, as well as those used for military applications in ground-based and airborne radars. *See* AIR-TRAFFIC CONTROL.

DME is an international standard which, together with VOR, has been the basis for the most widely used line-of-sight distance-bearing (rho-theta) aircraft navigation system. Airlines often use airborne equipment in DME/DME mode, obtaining distance to two or more DME transponders rather than in VOR/DME, thereby achieving improved accuracy. *See* DISTANCE-MEASURING EQUIPMENT.

TACAN is a military modification of DME, using the same channels and adding a bearing capability. The result is a ground antenna system which is more portable than that of the VOR, a property that is particularly useful for systems on aircraft carriers that guide aircraft home. The Vortac system provides DME service to civil aircraft and rho-theta service to military aircraft. *See* TACAN.

A second possible technique for point-source distance determination is one-way signal-delay (time-of-arrival) measurement between a transmitter source at a known location and a user receiver. In this case, a successful measurement is possible only if the transmitter oscillator (clock) and the user receiver oscillator (clock) are precisely time-synchronized (**Fig. 2**). Otherwise, true range (delay) cannot be determined, since the time of transmission is not known with respect to the user's clock. Since independent precise time synchronization between two equipments

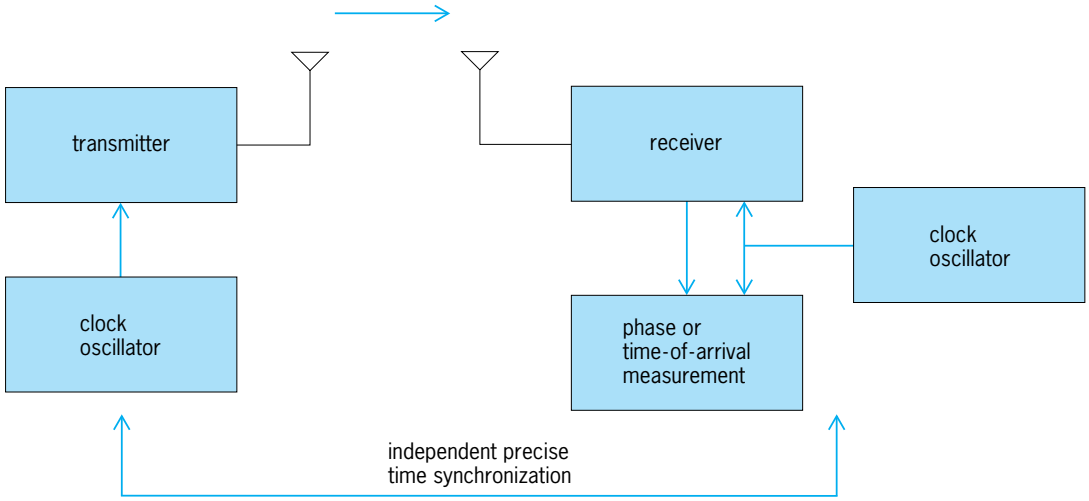

Fig. 2. One-way synchronous ranging system.

is frequently impossible at reasonable equipment cost, no practical point-source distance measurement system based on such synchronization has been developed, but several modern multisensor systems (discussed below) have modes that employ this principle.

Multiple-source radio systems. Multiple-source radio systems determine present position (and vehicle velocity) in an absolute or relative coordinate system by means of measurements to several signal sources at known locations. These systems therefore employ multiple transmitter sources, and user vehicle equipment typically consisting of a receiver or a receiver-transmitter. The four major categories of such systems (with some implementations using combinations of these) are hyperbolic systems, pseudoranging systems, one-way synchronous ranging (direct-ranging) systems, and two-way (round-trip) ranging systems.

Hyperbolic systems. These were the first to be developed, but they are still in widespread use. Three or more transmitter sources transmit time-synchronized continuous-wave or pulsed signals. The minimum-size chain (configuration of transmitters), a triad, usually consists of one so-called master and two secondary (slave) stations (**Fig. 3**). The user receiver measures time differences of arrival of signals from pairs of stations. Loci of constant time differences, or (equivalently) constant differences in distance, from two stations form hyperbolic lines of position (LOPs). The point where two lines of position cross is the position of the user vehicle.

One major advantage of this technique is that the user needs only a receiver and the receiver does not need an expensive high-quality clock oscillator, since only time differences are used. Theoretically, three pairs of sources are needed for a completely unique horizontal position fix, but in practice two pairs usually suffice because the user's altitude is known to acceptable accuracy. The differences in distance can be measured either in terms of differences in times of arrival (for example, of pulses) or in terms of differences in electrical carrier phase, or both. Achievable accuracy is very much a function of the relative geometric location of the sources and the user, and largely depends on the crossing angles of the lines of position (Fig. 3). For a given level of measurement noise, the smaller the crossing angles, the larger the position error. This accuracy degradation is called geometric dilution of precision (GDOP), and is essentially a multiplier on the basic time-difference or line-of-position root-mean-square measurement error. The GDOP is approximately $1/\sin\theta$, where θ is the crossing angle. The Loran C, Omega, and Decca systems are based on this hyperbolic principle, as were the now-terminated Omega and Decca systems. *See* HYPERBOLIC NAVIGATION SYSTEM.

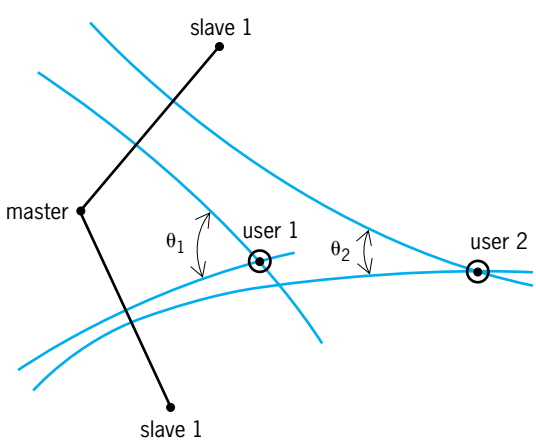

Fig. 3. Effect of relative geometry on accuracy of hyperbolic navigation systems. Accuracy largely depends on θ_1 and θ_2, the crossing angles at users 1 and 2.

The Loran C hyperbolic system was originally intended for shipping, both military and civil. The combination of low frequency and pulses provides long range and virtually eliminates multipath propagation interference from sky waves. The use of a form of phase measurement of the carrier cycles provides additional accuracy, and it inherently provides a fine-coarse readout of low inherent ambiguity. Since all Loran-C stations include an atomic clock for timing, precise time and time-interval information can be obtained from the system. For example, telephone systems use this capability for digital switching, signaling, and timing. Loran - C and GPS have been used in hybrid configurations to provide backup in case of outages, obtain precise time, increase overall system integrity, and provide a reference for differential GPS operation. These uses are likely to continue for some years, until GPS is widely accepted as the sole means of navigation. The use of digital computation in Loran C receivers has resulted in a tremendous growth in general aviation use. *See* LORAN.

Pseudoranging. This concept has some similarity in configuration to hyperbolic systems but does not use time differences. Several transmitter sources (for example, space vehicles or terrestrial stations), whose positions are made known to the user, transmit highly time-synchronized signals based on an established system time. With the transmission epochs of the signals from the sources provided or known to the user, the user measures the time of arrival (TOA) of each signal with respect to the user's own clock time, which normally has some offset from system time. The measured time of arrival minus the signal transmission time defines a range measurement that is called pseudorange, since it differs from the true range as a result of the offset of the user's clock relative to the system (transmitter) time. From successive or simultaneous time-of-arrival (pseudorange) measurements from four (or more) sources, the user receiver then calculates the three-dimensional position coordinates and its own clock offset (from system time). This is accomplished by solving four simultaneous quadratic equations (usually the computations are linearized), involving the three known position coordinates of the sources and the four unknowns, namely, the three user position coordinates and the user's clock offset.

Thus, such a system accurately determines not only the user's three-dimensional position but also the user clock offset relative to system time, which is easily related to Universal Time Coordinated (UTC). In addition, by properly combining the known

velocity of the sources and the measured Doppler shift of the signals received from the sources, the three user velocity coordinates and the user clock frequency offset can also be determined. In order to achieve high accuracy in this type of system, the transmitter sources must be highly synchronized through use of clock oscillators of very high stability, for example, atomic clocks. However, the user receiver clock oscillator does not need to be highly stable. *See* ATOMIC CLOCK; DOPPLER EFFECT; TIME.

The pseudoranging concept is the basis of operation of two major satellite navigation systems for worldwide use, the U.S. Global Positioning System (GPS) and the Russian Global Navigation Satellite System (GLONASS). In these systems, the satellite-borne clocks are rubidium or cesium atomic clocks, with long-term stabilities of the order of 1 part in 10^{13}. In order to provide high-accuracy time-of-arrival measurement capability, these satellite navigation systems use wide-bandwidth spread-spectrum modulation methods. The pseudoranging concept has also been applied in terrestrial navigation systems, notably in one mode of the Joint Tactical Information Distribution System–Relative Navigation (JTIDS-RelNav) system. *See* SATELLITE NAVIGATION SYSTEMS.

Global Positioning System (GPS). This satellite-based radio system was originally developed by the U.S. Air Force to provide worldwide coverage; high-accuracy, three-dimensional position, velocity, and time; and completely passive (receiver-only) operation by all types of dynamic users. It has found wide acceptance by both military and civil users. Two services are available, the Precise Position Service (PPS) for authorized users and the Standard Positioning Service (SPS) for all other users. The root-mean- square value of three-dimensional positioning errors for users of the Precise Positioning Service is specified to be less than 15 m (50 ft), although the achieved accuracy is of the order of 5 m (15 ft) or better in most circumstannces. The accuracy of the Standard Positioning Service (SPS) was improved on May 1, 2000, when the intentional degradation of service accuracy known as Selective Availability was discontinued. The specification value of horizontal positioning error that is obtained 95% of the time was reduced by this action from 100 m (330 ft) to 20–30 m (70–100 ft). The accuracy achieved since then is of the order of 5–10 m (15–30 ft), and further reduction is anticipated when additional civil signals become fully operational as soon as 2012.

The specified GPS consists of 24 satellites oriented in six orbital planes including four satellites each. The operational system has typically included 28 or more satellites because of the longer-than-anticipated lifetime of the satellites. Satellites are in 12-h orbits, with four satellites in six orbit planes inclined at 55°, all at an orbital altitude of 20,200 km (10,900 nmi). The satellites transmit highly synchronized, pseudonoise-coded (wide-bandwidth) signals, including data on their ephemerides and clock errors. The user receiver determines at least four pseudoranges by time-of-arrival measurements with respect to its own clock time, and four pseudorange rates (delta ranges) via Doppler measurements with respect to its own clock frequency. From these measurements, the user receiver computes its own three-dimensional position coordinates and its clock bias error, as well as (in some receivers) its three-dimensional velocity coordinates and clock frequency offset. The satellites transmit two codes, the 1.023-megabit-per-second (Mbps) C/A code, which can be tracked by civil users, and the 10.23-Mbps P (or Y) code, which is encrypted for exclusive use by military users. Each satellite has a unique code that distinguishes its signals from the other satellite transmissions that share the GPS frequency allocation by means of code-division multiplexing. System data are modulo-2 added to both codes at 50 bits per second. *See* MULTIPLEXING AND MULTIPLE ACCESS; SPREAD SPECTRUM COMMUNICATION.

Some applications, for example, aircraft landing and airport surveillance, require even higher accuracies than those available from the basic GPS service. For these, the differential GPS (DGPS) concept has been implemented, employing a reference station, whose position is precisely known. Relative positioning accuracy of 3 m (10 ft) is easily achievable with this technique. The Local Area Augmentation System (LAAS) is under development by the FAA to provide DGPS services at airports. Other augmentations to the civil use of GPS include the Wide Area Augmentation System (WAAS), in which two INMARSAT satellites transmit ionospheric delay corrections to users in the contiguous United States (CONUS) that have been determined by combining GPS measurements made at dozens of sites throughout the North America; Maritime and Nationwide DGPS (MDGPS and NDGPS), which transmit pseudorange and (sometimes carrier phase) corrections on the 285–325-kHz maritime radio beacon band; and Continuously Operating Reference Stations (CORS), which are operated by the National Geodetic Survey to support DGPS operation by surveyors, Geographic Information Systems providers, engineers, and scientists, by transmitting non-real-time GPS carrier phase and pseudorange measurements made at precisely surveyed locations.

Global Navigation Satellite System (GLONASS). This system, developed by the Russian government, is very similar in concept to the GPS. It uses 24 satellites in 25,500-km (13,800-nmi) orbits, with eight satellites in three orbital planes at an inclination of 64.8°. This inclination is somewhat higher than that of GPS, thus providing better polar coverage. GLONASS uses frequency-division multiplexing to distinguish signals broadcast from individual satellites. The future of GLONASS will likely be tied to international unification of Global Navigation Satellite System (GNSS) services as they relate to GPS and Galileo.

Galileo. Galileo is an initiative of the European Union, in collaboration with the European Space Agency and European Industry, to launch a European-financed global satellite navigation system under civilian control. It will have fundamental similarities to GPS with additional services for search and rescue and a guaranteed precise-accuracy service. Galileo will consist of a constellation of 30 satellites

in three orbital planes inclined at 54° and at an altitude of around 23,000 km (12,400 nmi). The compatibility, particularly with respect to military frequency allocations, between GPS and Galileo has been a major diplomatic issue. Although the two systems were distinct as of 2006, it is likely that future satellite navigation users will design receivers to utilize both simultaneously, which would further strengthen the dominance of satellite navigation for most terrestrial applications.

JTIDS. The Joint Tactical Information Distribution System (JTIDS) is a decentralized, military, spread-spectrum, data-communication and navigation system, using wide-bandwidth phase coding, frequency hopping, and time-division multiple access. The RelNav function permits all members of a JTIDS network, such as aircraft and ships, to accurately determine their position in both absolute and relative grid coordinates by means of highly precise time-of-arrival measurement of signals received by cooperating units when any member is located in an absolute reference system. Otherwise, all members of a net are positioned relative to one another. The system includes a means for independent precise time synchronization of each unit, and operates in both one-way synchronous ranging and pseudoranging modes.

JTIDS with the Relnav function is implemented in a large number of military aircraft and ships of the United States and a number of other countries. An international effort called the Multifunction Information Distribution System (MIDS) is under way toward further miniaturization and modernization of JTIDS terminals. In most aircraft applications, JTIDS-Relnav and an inertial navigation system (INS) are used in a fully integrated multisensor configuration. Combined JTIDS-RelNav/INS/GPS is planned in order to make use of the optimum combination of measurements from these three types of navigation sensors.

One-way synchronous ranging. The implementation of this concept between one of the transmitter sources and the user receiver was discussed above (Fig. 2). In order for a true one-way range measurement to be made, the clock of the user receiver must first be synchronized to that of the transmitter source. In some systems, this is accomplished by an independent means, for example, in the JTIDS-RelNav system and in the Position Location Reporting System (PLRS). The concept is also used in one mode of certain hyperbolic systems, notably Loran C and the decommissioned Omega. When applied to those systems, the concept has been called direct ranging. Two implementations are possible, the rho-rho method or the multiple-rho method.

The rho-rho method requires only two source transmitters, but also requires a highly stable user receiver oscillator (clock) and precise knowledge of the time of transmission from the source stations. A direct range is then developed from each station. The lines of position are circles rather than hyperbolas, with the intersection of the circles representing the user position, thereby leading to more favorable geometric conditions.

The multiple-rho method is an extension of the rho-rho scheme, requiring at least three stations. Using three lines of position permits clock oscillator self-calibration, much as the previously discussed pseudoranging concept, and therefore leads to a much less stringent clock oscillator requirement. Again, circular lines of position lead to higher position accuracy because of better geometric behavior.

Two-way (round-trip) ranging. This technique involves multiple two-way distance measurements of the type discussed above (Fig. 1). To obtain a completely unambiguous horizontal position fix, three such two-way ranges are required; however, two are usually sufficient. Since the lines of position are again circular, the geometric accuracy behavior is generally more favorable than for hyperbolic systems. The major disadvantage of this concept is that the user requires a transmitter as well as a receiver (Fig. 1). A typical application of this concept is DME-DME or multiple-DME operation for civil aviation navigation. Other applications are in the military PLRS and for the tracking of aircraft in military training ranges.

PLRS. This is a centralized military, spread-spectrum, position-location and navigation system for aircraft, land vehicles, and personnel. It uses wide-bandwidth phase coding, frequency hopping, and time-division multiple access, and also incorporates data communications capability. It provides military commanders with information on the location of all their elements, as well as their own positions and relative guidance information to each unit. Its operation is based on multilateration using highly precise time-of-arrival measurements of signals exchanged between units in both two-way (round- trip) ranging and one-way synchronous ranging modes.

PLRS and EPLRS, an enhanced version which provides more direct user-to-user communication, are in wide use by the U.S. Marine Corps and Army, respectively. A trend toward distributing the position- and range-bearing computations from the central processors to the individual user units is supported by the decreasing cost of digital processors and by the worldwide availability of GPS. A GPS receiver module can be embedded into a PLRS user unit, and GPS time-of-arrival-based positions can be combined with PLRS time-of-arrival-based measurements, resulting in the increased accuracy and availability of user positioning and navigation information. Walter R. Fried; Richard L. Greenspan; Marvin May

Bibliography. P. Enge and P. Misra, *Global Positioning System, Signals, Measurements and Performance*, 2d ed., 2006; J. Farrell, *Integrated Navigation Systems*, 5th ed., 2003; E. D. Kaplan and C. J. Hegarty (eds.), *Understanding GPS: Principles and Applications*, 2d ed., 2006; M. Kayton and W. R. Fried, *Avionics Navigation Systems*, 2d ed., 1997; B. W. Parkinson and J. J. Spilker (eds.), *Global Positioning System: Theory and Applications*, 2 vols., 1996; P. G. Savage, *Strapdown Analytics*, 2000; U.S. Department of Defense, *CJCS Master Positioning, Navigation, and Timing Plan*, biannually; U.S. Departments of Transportation and Defense, *Federal Radionavigation Plan*, biannually.

Electronic packaging

The technology relating to the establishment of electrical interconnections and appropriate housing for electrical circuitry. In particular, packaging for microelectronics focuses on interconnections between integrated circuits to achieve higher levels of assembly and electronic function, such as the storage and processing of information.

Electronic packages provide four major functions: interconnection of electrical signals, mechanical protection of circuits, distribution of electrical energy (that is, power) for circuit function, and dissipation of heat generated by circuit function. Thus, electronic packaging is a truly multidisciplinary science, involving such areas as physics, materials science, mechanical and electrical engineering, and computer-aided design.

Printed circuitry. As solid-state transistors started to replace vacuum-tube technology, it became possible for electronic components, such as resistors, capacitors, and diodes, to be mounted directly by their leads into printed circuit boards or cards, thus establishing a fundamental building block or level of packaging that is still in use.

Printed circuitry consists of a relatively stiff core of electrically insulating material in sheet form, with a patterned conducting layer on one or both sides. Holes drilled through the sheet at appropriate locations provide for mounting component leads that are subsequently soldered to ensure both mechanical and electrical connection to the package. Edges of printed circuit cards often provide metal lands (pads) for interconnection to other components. Consumer electronics, such as home video games, make extensive use of very basic printed circuit cards.

Although several different choices are available, the most common material set for printed circuit cards is a fiberglass core with copper conductor. Early patterns were formed by simple silk-screening and etching techniques, but these methods were later replaced by sophisticated photolithographic techniques to achieve very fine dimensions. For advanced applications, cards are now built in multiple layers that consist of conductor layers with intervening insulating layers, with interconnections known as through-vias (electrically conductive, orthogonal pathways between two parallel but isolated conducting layers). These vias are formed by drilling fine holes through the fabricated multilayer card and plating the inside surface of the holes, thus electrically connecting those conductor lines intersected by the drilled holes. The need for increasingly finer features (closer spacing of printed lines and smaller-diameter vias) has brought about significant advances in printed circuit technology, including use of laser-drilled vias and multilayered thin-film patterns on top and bottom of the printed circuit cards. The so-called microvia technology is the most advanced in the printed circuit industry. *See* PRINTED CIRCUIT.

Packaging hierarchy. Complex electronic functions often require more individual components than can be interconnected on a single printed circuit card. Multilayer card capability was accompanied by development of three-dimensional packaging of so-called daughter cards onto multilayer mother boards, providing interconnections for both power and card signals. Developed for the large mainframe computers during the period 1960–1985, this type of packaging hierarchy is now used for many personal computer applications.

Integrated circuitry allows many of the discrete circuit elements such as resistors and diodes to be embedded into individual, relatively small components known as integrated circuit chips or die. In spite of incredible circuit integration, however, more than one packaging level is typically required, driven in part because of the technology of integrated circuits itself.

Integrated circuit chips are quite fragile, with extremely small terminals. First-level packaging achieves the major functions of mechanically protecting, cooling, and providing capability for electrical connections to the delicate integrated circuit chip. At least one additional packaging level, such as a printed circuit card, is utilized, as some components (high-power resistors, mechanical switches, capacitors) are not readily integrated onto a chip. For complex applications, such as mainframe computers or network servers, a hierarchy of multiple packaging levels is required. A common packaging hierarchy consists of single-chip modules (first-level) soldered to multilayer printed circuit cards (second-level), which in turn are mechanically interconnected to system-level printed circuit boards (third-level) [**Fig. 1**]. *See* INTEGRATED CIRCUITS.

First-level packaging. In microelectronics, the first-level package is that which provides interconnection directly to the integrated circuit chip. The carrier upon which the chip rests (the substrate) must also provide for interconnection to the next (that is, second-level) package. Major emphasis has been placed on first-level packaging since the early 1970s.

Chip-to-package interconnection is typically achieved by one of three techniques: wire bond, tape-automated bond, and solder-ball flip chip (**Fig. 2**).

Wire bonding. The wire bond is the most widely used first-level interconnection; it employs ultrasonic energy to weld very fine wires mechanically from metallized terminal pads along the periphery of the integrated circuit chip to corresponding bonding pads on the surface of the substrate. Typical wires are made from aluminum or gold, with small alloying additions to achieve the desired strength required for handling. Wire diameters of only 25–50 micrometers (0.001–0.002 in.) are used because of the small size of the terminals of the integrated circuit chips. Metallization of the substrate and the pad of the integrated circuit chip is typically gold plating and evaporated aluminum, respectively; both are often modified with small alloying additions to customize properties. Bond pads are only a few micrometers thick and are on the order of 50–200 μm (0.002–0.008 in.) in diameter. Wire-bonded integrated circuit chips are

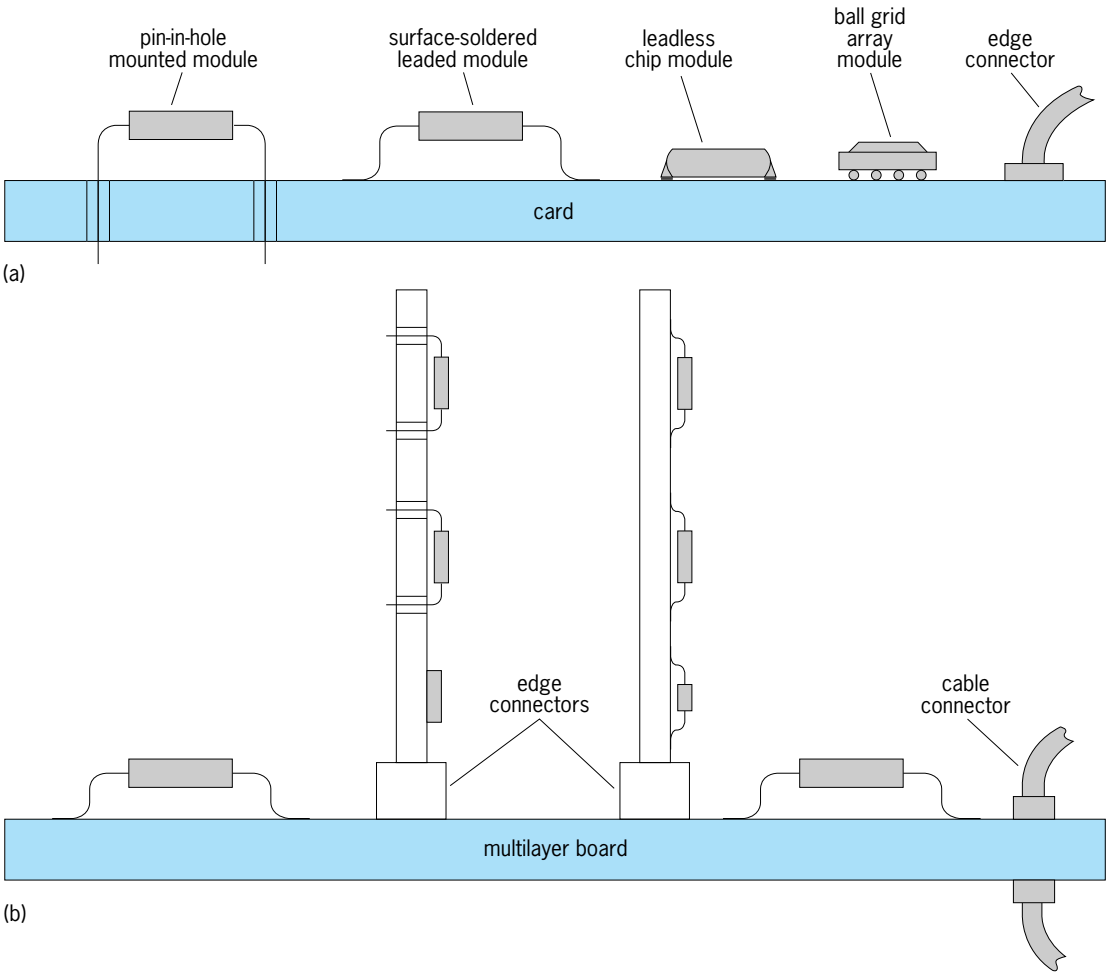

Fig. 1. Packaging hierarchy. (*a*) A card and single-chip modules. (*b*) Multilayer board.

attached to the substrate with active circuitry facing out, such that the back of each chip provides a surface for adhesive bonding. Heat is dissipated through the back-bonded interface.

Tape-automated bond. In the tape-automated bond, photolithographically defined gold-plated copper leads are formed on a polymide carrier that is usually handled like 35-mm photographic roll film, with perforated edges to reel the film or tape. The inner leads are connected to the integrated circuit chip through simultaneous thermocompression bonding of all leads to corresponding gold bump contacts on the terminals of the integrated circuit chips (gang bonding). Outer leads are soldered, thermocompression-bonded, or bonded ultrasonically to the substrate pads, which are similar to wire bond pads in structure.

Solder-ball flip chip. This technique involves the formation of solder bump contacts on the terminals of the integrated chips and reflowing the solder with the chip flipped in such a way that the bump contacts touch and wet to matching pads on the substrate. Lead-tin solder rich in lead is usually used, with flux to ensure adequate wetting to the surface of the substrate. In contrast to wire bonding, an area array configuration rather than peripheral leads is possible, leading to much greater density of interconnecting terminals per unit area of substrate. These chips are soldered with the circuitry facing toward the substrate (hence the name flip chip) in contrast to more common wire bonding.

Several alternate approaches to achieving flip chip, other than solder ball, are currently practiced. These typically include some form of metallic bump on the integrated circuit die connected to matching pads on the substrate with electrically conductive polymer adhesives.

Substrates. Substrates for first-level packages are quite varied; selection for a given application depends on several factors, including the required number of interconnections, electrical characteristics, reliability objectives, and cost. A major classification is whether the package supports a single integrated circuit chip (single-chip module) or more than one chip (multichip module). The former is by far the most common. Substrate insulator materials for multichip and single-chip modules are selected from one of two broad groups of materials, organics (including epoxy and polyimide) and ceramics (predominantly alumina, but also including silicon carbide and aluminum nitride). Conductors are chosen from a wide range of metals. Specific materials are selected on the

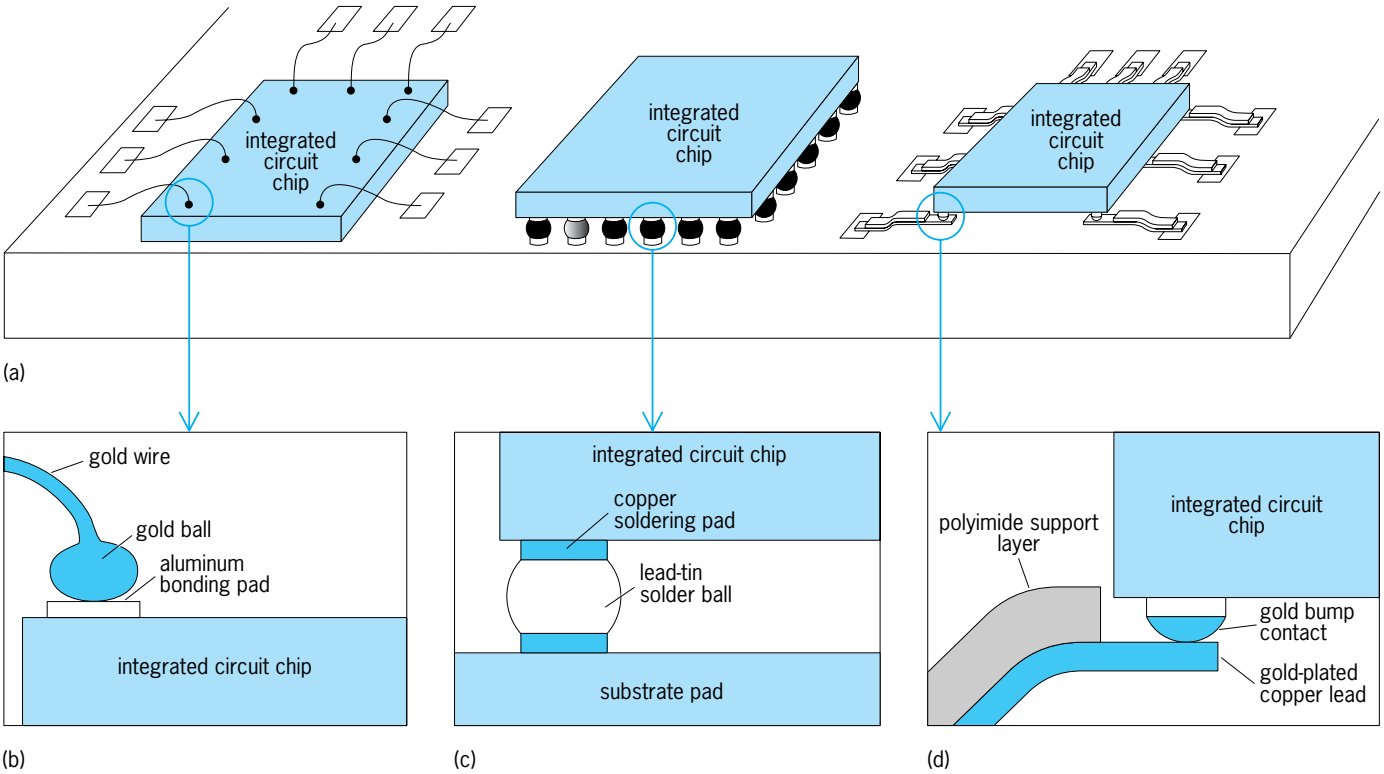

Fig. 2. First-level packaging. (*a*) Substrate. (*b*) Wire bond. (*c*) Solder-ball flip chip. (*d*) Tape-automated bond.

basis of electrical characteristics (for instance, low-dielectric-strength insulators and low-resistivity conductors) as well as properties required for processability and reliability, such as thermal coefficients of expansion, thermal conductivity, processing temperatures, and thermal stability. *See* CERAMICS; HETEROCYCLIC POLYMER; POLYETHER RESINS.

Analogous to the advances in printed circuitry, early first-level package substrates were fabricated with a single layer of conductor paste, screened, and then fired onto small pressed ceramic tiles prefabricated with holes through which pins were swaged to provide for second-level interconnections. Ceramic packages today can be made with upward of 60–70 alternating conductor and insulator layers, connected by conductor-filled vias. These complex multilayer ceramics, typically for applications involving multichip modules, are fabricated as green (that is, pressed but unfired particulate) laminates and then sintered to near-theoretical density. Surface metallization, required for first- and second-level interconnections, has advanced so that it can include photolithographically defined patterns.

Significant advances in organic-based substrate technology, notably those utilizing ball grid array interconnect, have led to multilayer plastic packaging for single-chip modules that is competitive in some aspects with traditional ceramic packaging. These substrates utilize many of the same fabrication techniques applied in advanced printed circuit technology discussed earlier.

Interconnection. Means for interconnecting to the second-level package often dictate the general form of the first-level package (**Fig. 3**). For instance, if the connection is to be pin-in-hole, the package must have pin-shaped leads that are either peripheral, such as a dual in-line package, or in an area array, such as the pin or ball grid array. If the package is to be soldered on the surface, it can be a leaded carrier, in which leads of various geometries are soldered to pads on the card, or a leadless carrier, in which solder from the card contacts metal pads on the edge of the package. Temperature hierarchy in processing must follow the packaging hierarchy, that is, assembly of first- to second-level packages should occur at

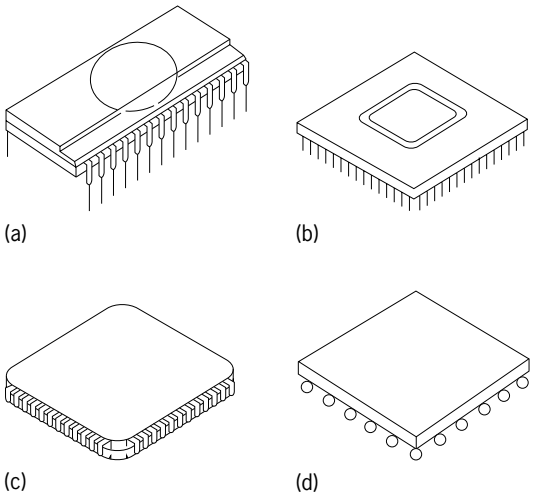

Fig. 3. Typical first-level packages. (*a*) Dual in-line package. (*b*) Pin grid array. (*c*) Leadless chip carrier. (*d*) Ball grid array.

process temperatures below that at which the chip-to first-level package interconnection was achieved. When more than two packaging levels are required, this guideline can become difficult to follow.

Field data have demonstrated that the most vulnerable portion of the package is the interconnection, either chip-to-substrate or module-to-board. Mismatch between the integrated circuit chip and the substrate materials and substrate-to-card materials caused by thermal expansion can generate cyclic fatigue stresses during machine cycles (that is, each time an electronic system is turned on and off, the integrated circuit chip and its package heat up and cool down). Reliability of both the integrated circuit chip and the interconnections can be enhanced by sealing out environmental moisture. Sealing (encapsulation) techniques range from inexpensive epoxies to high-performance hermetic sealing using soldered or glass-sealed caps to protect the integrated circuit chip and leads. Finally, heat sinks are often attached to the packages to assist in the dissipation of heat generated by operation of the integrated circuit chip. *See* ELECTRONICS. Peter J. Brofman

Bibliography. C. A. Harper, *Electronic Packaging and Interconnection Handbook*, 1991; J. Hayward, The problem of interconnect and the impact of printed circuit board technology on integrated circuit package development, *Future Circ. Int.* (London), no. 4, pp. 19–22, 1998; K. Matsuda et al., Simple method for flip-chip bonding on organic substrates, *Int. J. Microcirc. Electr. Packag.*, vol. 20, no. 3, 1997; D. P. Seraphim, R. C. Lasky and C. Li, *Principles of Electronic Packaging: Design and Materials Science*, 1989; R. Tummala and E. J. Rymaszewski, *Microelectronics Packaging Handbook*, 2d ed., 1995.

Electronic power supply

A source of electric power (voltage and current) to operate electronic circuits. Active electronic circuits contain such devices as transistors or vacuum tubes and require external power to amplify, filter, modify, or create electrical signals. The most common source of energy for electronic circuits is obtained by converting the electrical energy available in the conventional alternating-current (ac) electric power mains to an appropriate voltage or current. These converters, or electronic power supplies, can be implemented with a wide variety of circuits. Other power sources include batteries, mechanically driven generators, photovoltaic (solar) cells, and fuel cells. *See* ALTERNATING CURRENT; CONVERTER.

Most electronic circuits require a direct-current (dc) or constant voltage. If ac power is required, an oscillator or a simple transformer is used. Although some dc-to-dc converters are used, most dc power supplies convert the alternating power from the ac main to dc power. These ac-to-dc power supplies are classified according to the type of circuits used to realize the conversion: rectification, filtering, and regulation. Simple battery chargers are examples of power supplies that do not require filtering or regulation.

Rectification. An essential step in the conversion of ac to dc is a process called rectification. Rectification converts ac voltage to a waveform with average or dc value by passing only one polarity (half-wave) or by generating the magnitude or absolute value (full-wave). Three types of rectifier (diode) circuits are commonly used. Only one diode is required to obtain half-wave rectification. The diode passes current in one direction and blocks current flow in the other direction. Full-wave rectification can be obtained with four diodes connected in a bridge configuration or with two diodes and a center-tapped transformer (**Fig. 1**). The diodes permit current to flow from each half of the transformer secondary, but only in one direction through the load. The resulting waveform is filtered with a capacitor, and an optional regulator may be added for additional refinement of the dc output voltage. Transformers are normally used at the input of the rectifiers to increase or decrease the voltage and isolate the dc output from the ac input for safety purposes. *See* DIODE; RECTIFIER; TRANSFORMER.

Filtering. For most applications the ac or alternating portion of the rectified output is unwanted and may cause undesirable results, such as an annoying hum in audio systems. A capacitor can be used to reduce or filter the ac portion of the rectified waveform (Fig. 1). The capacitor is charged through the diodes to the peak ac voltage minus the diode forward voltage. Some of the charge stored on the capacitor is delivered to the load each cycle, but the next voltage peak recharges the capacitor. For a capacitance (C), the basic relationship between current (i) and voltage (v), Eq. (1), can be used to estimate the remaining peak-to-peak ripple, according to Eq. (2).

$$i = \frac{Cdv}{dt} \tag{1}$$

$$\text{Peak-to-peak ripple} = V_{pp} \cong \frac{I_{dc}}{fC} = \frac{E_{dc}}{fCR_L} \tag{2}$$

Here, the frequency, f, is the input frequency to the filter, I_{dc} and E_{dc} are the dc current and voltage, and R_L is the load resistance. The value for f can range from 50 Hz for a half-wave rectified European ac line, to several hundred kilohertz for an electronically switched power supply. For a full-wave rectified ac line in the United States, $f = 120$ Hz. Electrolytic capacitors are normally used for ac power-line applications but are not adequate at high frequencies. Other low-pass filter circuits using combinations of capacitors and inductors can be used to pass the dc and attenuate the ac part of the waveform. *See* ELECTRIC FILTER.

Regulation. Regulators are often used to make the power supply output insensitive to input voltage amplitude variations and further reduce the ripple voltage. The regulator may also be used to adjust or change the dc output voltage and limit the amount of current delivered by the power supply. Regulators are a form of dc-to-dc converter.

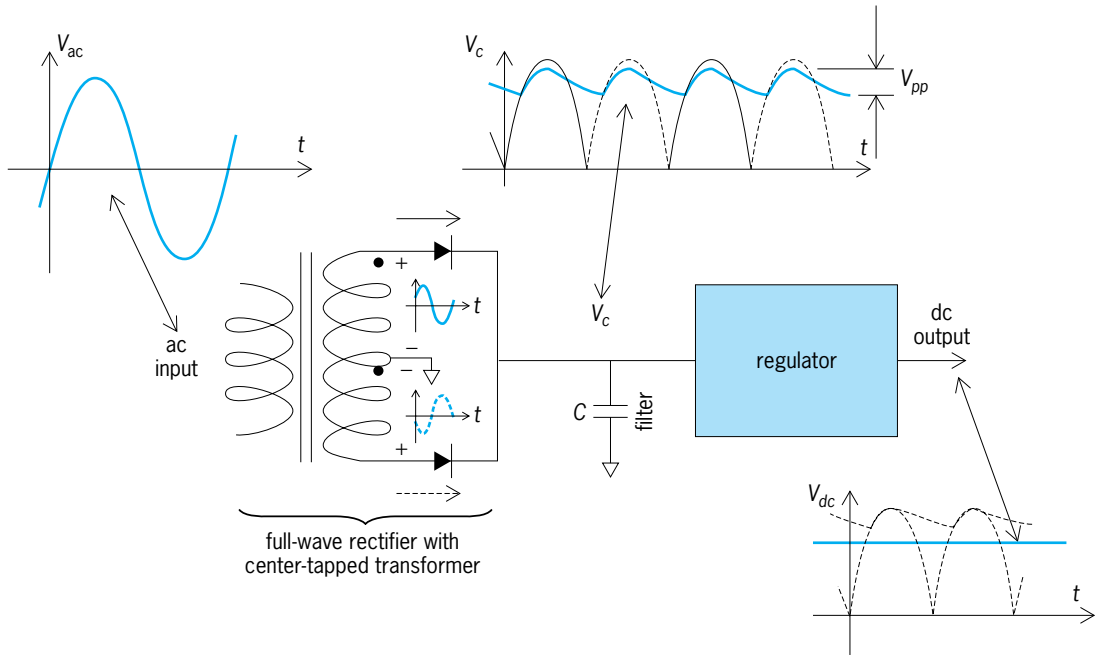

Fig. 1. Typical electronic power supply.

Linear regulators. The oldest and simplest type of regulator is the linear regulator. A simple linear regulator is the shunt type (**Fig. 2***a*). It consists of a Zener diode ($D1$) and a current-limiting resistor (R_s). The Zener diode establishes a fixed, or reference, voltage if it is properly reverse biased. Another common type of linear regulator is the series-pass regulator (Fig. 2*b*). A negative feedback control system automatically controls the input to the series-pass device (such as an *npn* bipolar transistor, $Q1$) to maintain the output equal to the reference or desired level, V_R. The power dissipation in the R_s, $D1$, or $Q1$ significantly reduces the efficiency of a power supply. *See* TRANSISTOR; ZENER DIODE.

Switching regulators. More modern power supplies have switching regulators, and are informally called switchers. There are more than a dozen different topologies (basic block diagrams) for switching regulators, but they all use one or more transistors acting as switches; either ON or OFF. In addition to the solid-state (transistor) switches, a typical switching regulator uses capacitors and inductors to store energy and diodes to direct the current. A negative-feedback control system is used to set and maintain the output level, similar to the linear series regulator. For a switching regulator, the feedback system controls the frequency, or duty cycle, of the switch control voltage. For a buck-or-forward switching regulator (**Fig. 3**), the output voltage must be smaller than the input dc voltage. When the switch ($Q1$) is on, the current increases through an inductor ($L1$), providing more current to the load, and charges the capacitor ($C1$). When the output voltage rises above the desired value, the control system turns the transistor ($Q1$) off. The current through the inductor continues to flow in the same direction but now through a diode ($D1$). The amplitude of the current decreases. The voltage drops slightly until the control system turns the switch on again. This cycle is repeated automatically to maintain a nearly constant dc output voltage. Other common types of switching regulators are called buck-boost (or flyback) and push-pull or (buck-derived). Switching regulators can be used to generate multiple outputs at different voltage levels. The improved efficiency of switching regulators is due to the fact that energy is stored very efficiently in inductors and capacitors. The remaining losses in the control circuitry, switches, and the diodes are small compared to linear regulators. *See* INDUCTOR; VOLTAGE REGULATOR.

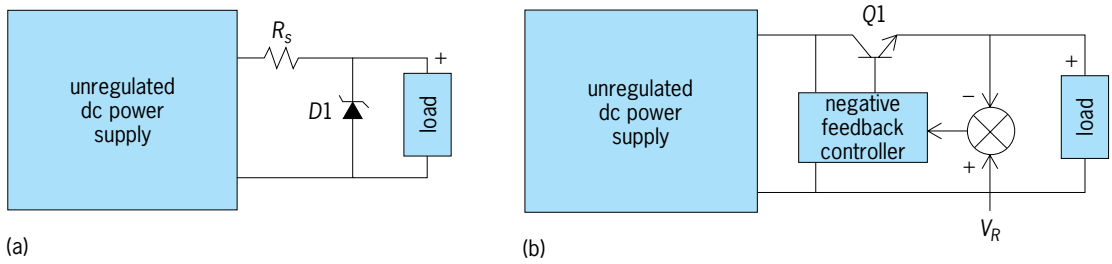

Fig. 2. Linear voltage regulators. (*a*) Zener-diode shunt type. (*b*) Transistor series-pass type.

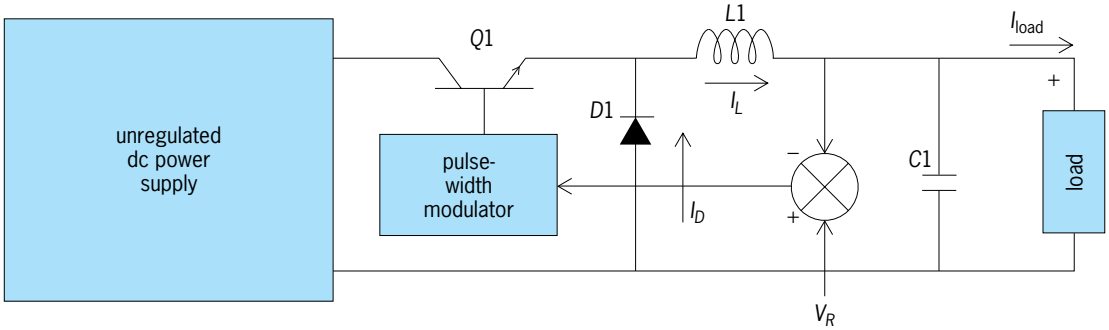

Fig. 3. Simplified circuit diagram of buck-or-forward switching regulator.

Linear versus switching supplies. The major drawback of linear power supplies is the power dissipation of the control device. For example, all of the load current must pass through the pass transistor in a series-pass regulator (Fig. 2*b*). The resulting power dissipation reduces the efficiency of the supply. Efficiencies of linear dc power supplies range 30–60% for typical output voltages and currents, while switching power supply efficiencies are typically 70–80%. Small linear regulators are relatively inexpensive. Completely integrated linear regulators that can carry up to 1.5 A are available in inexpensive plastic packages. Metal packages may be used for currents in the 5-A range, but power dissipation is a problem for higher-current applications. Size and weight is a problem with larger linear power supplies because large heat sinks and 60-Hz (or 50-Hz) input ac line transformers are required. The typical output power for a 10 cm × 10 cm × 10 cm (4 in. × 4 in. × 4 in.) linear supply is 15 W, while a switching power supply of this size can easily have an output power of more than 150 W. Most switching power supplies use small high-frequency transformers after the switch for isolation, so a large power-line-frequency transformer is not required. Switching regulators operate at a typical frequency of 40 kHz, so much smaller filter capacitors can be used. The increased efficiency of the switching regulator is due to the fact that inductors and capacitors store energy with very little power dissipation or energy loss. Also, the switching devices dissipate very little energy because they are either on or off. When they are off there is no current, and when they carry current in the on state the voltage is very small; thus the average power dissipation is small. Most of the power loss occurs in the transition between states. Therefore, device switching speed is very important. Using faster devices such as field-effect transistors (FETs), insulated-gate bipolar transistors (IGBTs), and metal-oxide-semiconductor (MOS) controlled thyristors results in improved performance over bipolar junction transistors (BJTs). Faster devices allow a higher switching frequency, which results in smaller filter capacitors and transformers.

Other power supplies. Ferroresonant transformer-based power supplies have some advantages for high-current applications such as battery chargers. Ferroresonant power supplies use nonlinear magnetic properties and a resonant circuit to regulate the output voltage, and they have efficiencies similar to switching power supplies. Power supplies driven by three-, six-, or twelve-phase ac are easier to filter, and they generate much lower harmonic distortion of the current in the ac power system.

Alternating-current power system quality is degraded by most electronic power supplies. This effect may become a serious issue as more electric vehicles are charged in the future or in buildings where electronic systems are the dominant ac load. Flyback power supplies and voltage-multiplier circuits are often used for high-voltage dc power supplies that require less current. National and international regulations for safety, power-line harmonics, and electromagnetic and radio-frequency interference (EMI-RFI) are a very important consideration for modern electronic power supplies of all types. *See* ELECTRICAL INTERFERENCE; VOLTAGE-MULTIPLIER CIRCUIT.

N. G. Dillman

Bibliography. S. G. Burns and P. R. Bond, *Principles of Electronic Circuits*, 2d ed., 1994; G. C. Chryssis, *High-Frequency Switching Power Supplies*, 2d ed., 1989; R. C. Dorf (ed.), *The Electrical Engineering Handbook*, 2d ed., 1997; M. J. Fisher, *Power Electronics*, 1991; D. A. Grant and J. Gowar, *Power MOSFETS, Theory and Applications*, 1989; A. I. Pressman, *Switching Power Supply Design*, 2d ed., 1997; R. S. Ramshaw, *Power Electronics Semiconductor Switches*, 2d ed., 1993.

Electronic switch

An electronic device in which one or more input signals can be routed to one or more outputs by the application of the appropriate electrical control signals. The term is most often applied when analog signals are involved, but the terminology is occasionally used when digital signals are involved.

Operation. Conceptually, an electronic switch can be visualized as a group of one or more mechanical electrical switches (such as light switches used in commercial wiring or toggle switches used in many electronic control panels) in which, instead of mechanically opening or closing the contacts, the physical opening and closing is achieved by applying appropriate electrical control signals to separate

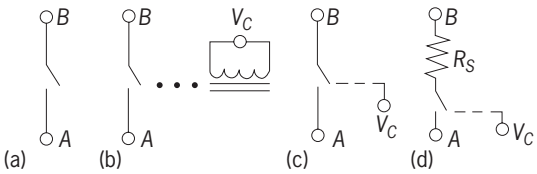

Switching devices. (*a*) Symbol for mechanical switch. (*b*) Symbol for electromechanical relay. (*c*) Symbol for electronic switch. (*d*) Simple model of electronic switch.

terminals on the switch in much the same way that a relay performs. Unlike the mechanical switch (**illus.** *a*) or the electromechanical relay (illus. *b*), however, the electronic switch does not contain mechanical contacts but semiconductor devices such as bipolar junction transistors or field-effect transistors. The basic electronic switch is depicted in illus. *c*. *A* and *B* are the terminals of the switch. When a control signal is applied to V_C, the switch closes. When the electronic switch is closed, a small residual resistance R_s remains between the terminals as depicted in the simple model of illus. *d*. The value of this resistance is termed the on resistance. This resistance varies considerably from one electronic switch to another and depends on the manufacturer of the switch. In most applications the nonzero on resistance does not prove problematic, but the user needs to be aware of this limitation. The electronic switch is typically bidirectional in the sense that the terminals *A* and *B* are interchangeable. *See* ELECTRIC SWITCH; RELAY; TRANSISTOR.

Advantages and disadvantages. Electronic switches offer attractive alternatives to mechanical switches and mechanical relays in several respects. They can be very small, allowing a large number of these devices to be placed in a small area; they are often significantly less expensive than their mechanical counterparts, particularly when small currents are being switched; they can be very fast, with on and off response times which are orders of magnitude faster than can be achieved with mechanical counterparts; they often require only minuscule amounts of energy at the input V_C to control the switch; the control inputs are often directly compatible with the voltage and current levels inherent in most electronic equipment; and they are generally considerably more reliable over a large number of cycles than their mechanical counterparts. On the other hand, they often have more stringent limitations on single voltage or current levels than do mechanical switches, and they do become more costly as the current- or voltage-handling capability increases.

Implementation. Electronic switches are available in stand-alone packages, often termed solid-state switches or analog switches. The term multiplexer, analog multiplexer, or data selector refers to a special class of multiple-input electronic switches in which either digital or analog signals are routed in a predetermined way to one or more outputs, often in a unilateral way (inputs and outputs not interchangeable). The electronic switching or data selection function can also be achieved with relatively simple circuits involving transistors or operational amplifiers along with a few passive components. *See* MULTIPLEXING AND MULTIPLE ACCESS; OPERATIONAL AMPLIFIER.

Randall L. Geiger

Electronic test equipment

Apparatus used for the evaluation of electronic components, subassemblies, and systems.

Early electrical instruments. The galvanometer was an early electromechanical instrument that converted current into angular rotation of a needle pointer across the face of a scale. The instrument consisted of a fine wire coil mechanically connected to a pivoting element. Also connected to the element were the needle pointer and a spring, which held the pointer to an initial position on the scale. The coil was located in the field of a permanent magnet so that when current flowed through the coil the induced magnetic field rotated the element, causing the needle to move across the scale. Greater current caused a larger deflection. This instrument could also be used as a voltmeter, since a higher voltage induces a larger current flow through the coil. *See* AMMETER; GALVANOMETER.

The advent of the cathode-ray tube (CRT) enabled the development of another widely used instrument, the oscilloscope. The CRT translates voltage into the deflection of an electron beam, which visibly activates a luminescent phosphor inside the face of the tube. Vertical and horizontal deflection plates correspond to x and y movement of the beam. Usually a sawtooth-shape waveform is connected to the horizontal plates, so that the x direction represents time. Therefore, when a voltage is connected to the vertical plates, the visual representation on the CRT face has the amplitude and polarity of the voltage displayed as a function of time. *See* CATHODE-RAY TUBE; OSCILLOSCOPE.

An important advance in electronic test equipment was the incorporation of circuits that directly converted the analog signal to be measured into a digital reading. Initially, these improvements provided easier and more accurate operator interpretation of the instrument's measurement. The most significant impact of these converters was to enable a computer interface to the equipment to be set up. Many new instruments were developed with computer bus capabilities, allowing direct computer monitoring and control of the instrument. Other equipments were designed with embedded computers, which provided very sophisticated analysis of the data within the instrument itself. *See* ANALOG-TO-DIGITAL CONVERTER; DIGITAL-TO-ANALOG CONVERTER; EMBEDDED SYSTEMS.

The best example of this enhancement is with spectrum analyzers. Digital signal processing (DSP) electronics have been included in these instruments which provide a frequency spectral analysis of the input signal. These instruments are widely used in the evaluation of radio communication signals, acoustic signal processing, and the analysis of

mechanical forces and vibration. *See* ACOUSTIC SIGNAL PROCESSING; SPECTRUM ANALYZER.

Automatic test equipment. The complexity and sophistication of electronic components also brought about the evolution of the test equipment. As components and subassemblies became more complicated, individual pieces of test equipment such as voltmeters, oscilloscopes, and signal generators could no longer efficiently evaluate the performance of the unit under test. The individual test instruments were integrated to form rack-and-stack automatic test equipment, where the individual instruments were stacked on top of each other and connected to a control computer by a common bus. After the initial test programming effort, these computer-controlled testers were very efficient at providing the input stimulus to the unit under test and monitoring the output response. Improved microelectronic components led to increasing capabilities, speed, and accuracy of stimulus and monitoring instruments, and component cost reductions eventually led to the complete integration of equipments into one complex automatic-test-equipment system (see **illus.**). *See* ELECTRONICS; INTEGRATED CIRCUITS; LOGIC CIRCUITS; MICROPROCESSOR; SIGNAL GENERATOR.

Characterization versus production testing. Two fundamentally different types of testing can require different types of electronic test equipment. Design verification or characterization tests are usually done on a small sample of units to validate the design or the process and to determine how much performance margin exists. Extensive data are taken, but the testing is a nonrecurring task, so test time and data acquisition efficiency are not too important. Product-specific test fixtures are designed, and general-purpose instruments are usually employed in an engineering laboratory environment, where the equipment can be shared or redeployed from one program to another.

Automatic test equipment for testing microelectronic devices. (*Rockwell International*)

Production testing, however, is done on all the units, so speed and efficiency are important. There is usually little data gathering, since the results of the tests are simply "pass" or "fail." Although there may be many different tests at various parametric conditions within one production test program, they are done in quick succession, and the tests typically require only a few seconds to carry out. Thus, only a few expensive production test systems can test thousands of units per day. The importance of short unit test times is reflected in the design of automatic test equipment. Stimulus- and response-measurement subsystems have been developed with ever briefer settling times (the time required to stabilize after a change in mode), and faster test-control computers have been included in automatic test equipment.

Test diagnosis. The ability of automatic test equipment to sort out good from bad units is only the first requirement of modern production test equipment. Since the repair of defective units can be costly, the design and program development of test equipment frequently must include special provisions to provide failure-mode analysis. Information such as the probable defective component on a printed circuit board or the probable defective board in a system is of great value in efficient repair. Some test systems include fault dictionaries, supplemental tests after first failure, or even artificial-intelligence features to assist in repair. *See* ARTIFICIAL INTELLIGENCE; FAULT ANALYSIS.

For microelectronic devices, where repair is usually impractical, failure-mode analysis is used to improve the manufacturing process. Bad devices can be tested under relaxed conditions to determine if they have a gross defect such as a short or an open, or if they have a parametric failure where they do not operate at rated speed or voltage. This information is used as feedback to the manufacturing process, where adjustments can be made to improve the yield, or number of good units, per manufacturing batch.

Quality level. The economic aspects of electronic system manufacturing have dictated the test thoroughness or quality level that must be achieved by the test equipment. In general, the cost to diagnose and repair a unit under test increases an order of magnitude for each successive level in manufacturing. If it costs $1 to detect a bad device, it would cost $10 to detect and replace that device in the next-higher assembly, the printed circuit board; it would cost $100 to detect and replace the bad board at the system level, and $1000 to repair that system once installed in the field. Therefore, the earliest detection of a defect is the most economical. Component customers who formerly asked for defect levels slightly less than a 1% now expect defect rates on the order of a few parts per million. The test equipment and the overall test strategy have been affected by this expectation of high quality level. The biggest impact is that testers with higher speed and improved accuracy are required. Worst-case conditions such as high temperature or low voltage must be determined for the units to be tested, and then all units are subject to

these conditions in a single test or even a multiple-pass test. Test equipment must be integrated with artificial environment chambers for the component, printed-circuit-board, and system tests. *See* ENVIRONMENTAL TEST; PRINTED CIRCUIT.

The importance of electronic test equipment far exceeds that of the simple measurements or even the complex components for which they have been designed. The common electronic automotive instrument cluster, for example, combines measurements of speed, angular speed, fuel level, and engine performance in one easily read panel, but behind that panel are the fundamental elements of test instruments. William R. Mann

Bibliography. K. Brindley, *Automatic Test Equipment*, 1991; M. Collons, *Computer Controlled Testing and Instrumentation: An Introduction to the IEC-625: IEEE 488 Bus*, 1983; Hewlett Packard, *Test and Measurement Catalog*, 1995; R. J. Traister, *Meters and Scopes: How to Use Test Equipment*, 1988; D. Wobschall, *Circuit Design for Electronic Instrumentation*, 2d ed., 1987.

Electronic warfare

The art and science of preserving the use of the electromagnetic spectrum for friendly forces while denying its use to the enemy. This is accomplished via use of techniques, devices, and equipment by an adversary to deny or counteract an enemy's use of radar, communications, guidance, or other radiowave devices. The principal techniques that are employed are called electronic countermeasures (ECM). Because of the growing use of optical and infrared techniques for communications, guidance, detection, and control, they are sometimes called electromagnetic, rather than electronic, countermeasures to convey more adequately the idea that countermeasures are not confined to the portion of the spectrum where electronic techniques alone are applicable but may be used throughout the electromagnetic spectrum. Other techniques include the use of antiradiation weapons designed specifically to home on and destroy radars, and the attempts to reduce the electromagnetic signatures of aircraft, missiles, and ground vehicles.

In a broader sense, electromagnetic warfare refers to the worldwide struggle by major military powers for supremacy in the electromagnetic spectrum. This largely silent, secret struggle occurs continually, independent of whether its participants are at war with one another, as was the case during the Cold War. Participants seek control of the spectrum because of the increasing dependence of military forces on its use for surveillance, communications, detection, measurement, guidance, and control. With mastery of the spectrum, one adversary could achieve an indispensable ingredient for conquering an enemy or discouraging a potential aggressor.

Since the late 1990s, there has been a close association of electronic warfare with information warfare (IW). Information warfare can be described as taking action to protect one's information systems from exploitation and corruption while simultaneously exploiting, corrupting, or destroying the adversary's information systems. Electronic warfare is considered the implementation tool or active component of information warfare.

Categories of ECM. Traditionally, ECM equipment and techniques are categorized as active and passive, depending on whether or not they radiate their own energy. The passive category includes reconnaissance or surveillance equipment that detects and analyzes electromagnetic radiation from radar and communications transmitters in a potential enemy's aircraft, missiles, ships, satellites, ground installations, and communications systems. The reconnaissance devices may be used to identify and map the location of emitters without in any way altering the nature of the signals they receive. Other types of passive ECM do seek to enhance or change the nature of the energy reflected back to enemy radars, but they do not generate their own energy.

Active ECM equipment generates energy, either in the form of noise to confuse an enemy's electromagnetic sensors or by radiating false or time-delayed signals to deceive radio or radar equipment and their operators. *See* ELECTRICAL NOISE.

Surveillance systems. Reconnaissance or surveillance ECM systems are carried by Earth-orbiting satellites, aircraft, ships, crewless (drone) aircraft (commonly known as unmanned aerial vehicles or UAVs), and automotive vehicles. Some are located on the ground (**Fig. 1**). A few systems are small enough to be carried by a person. Reconnaissance systems are

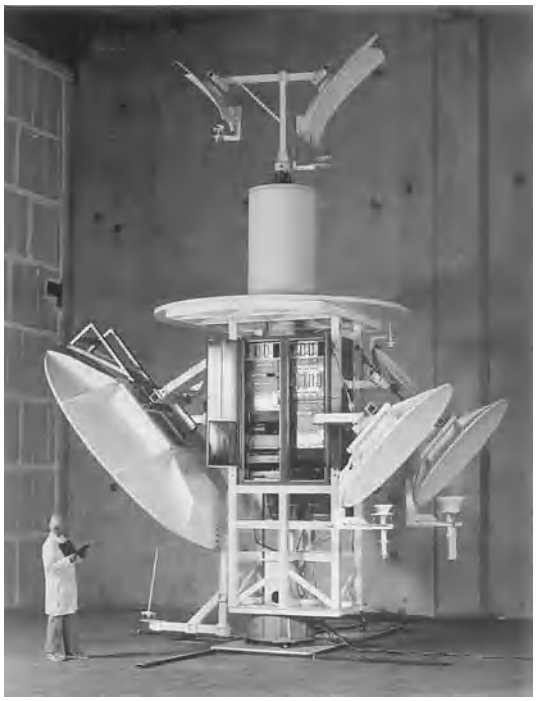

Fig. 1. Fully automatic land-based ELINT system, which monitors and documents the nature and location of hostile electromagnetic emitters and detects any changes in their location. (*Litton Applied Technology*)

interchangeably called ferret or electronic intelligence (ELINT) systems. *See* DRONE; MILITARY SATELLITES.

They consist of sensitive receivers electromechanically or electronically tuned over desired portions of the spectrum in search of transmissions of interest. Bearing to an intercepted signal can be determined by direction-finding techniques. Once secured, the signals can be displayed for analysis by an operator or stored on tape or other storage media for subsequent analysis or both. *See* DIRECTION-FINDING EQUIPMENT.

The type of information obtained by modern surveillance systems is illustrated by a standard system in use by the United States Air Force. It has a superheterodyne receiver and five separate radio-frequency tuning heads that span the spectrum from 1 to 18 GHz. (More advanced systems extend the range up to 40 GHz to cope with higher-frequency threats.) The receiver sequentially scans through those five heads at sweep rates up to 20 Hz, selectable by an operator. The operator has a panoramic display of the five separate amplitude-versus-frequency traces projected by a multigun cathode-ray tube. When a signal of interest is observed, a cursor can be superimposed over it, causing the receiver to discontinue the sweep mode and lock onto the intercepted signal. At this point, the display automatically switches into an analytical mode, offering an enlarged time display with repetitive pulsed signals on five separate logarithmically spaced time scales. Then the pulse shape, its repetition rate, and pulse width can be visually identified. Separately, an automatic digital indication of the signal's frequency in 100-kHz increments, pulse duration to 0.1 microsecond, and pulse repetition period to 1 microsecond can be obtained. *See* RADIO; RADIO RECEIVER.

Not all reconnaissance equipment is intended for spectral observation. Another form of reconnaissance equipment, carried by some U.S. Air Force aircraft, aids photo interpretation by detecting and partially identifying electromagnetic signals. Intelligence information gathered by the set is directly recorded on the film of a reconnaissance camera. The system, which covers six military frequency bands, locates an emitter in any one of eight sectors of the photograph. It helps in identifying the frequency, pulse-repetition frequency, and pulse width of the unknown intercepted radar signal.

Radar warning receivers. Radar warning-receiver (RWR) systems, which came into widespread use on United States tactical and transport aircraft during the Vietnamese war, are a more limited form of the ELINT system, intended as a means of self-protection. Unlike the latter, which is intended to search for signals over a broad range of the spectrum, the warning receiver is programmed to alert a pilot whose aircraft is being illuminated by a specific radar signal above predetermined power thresholds anticipated by ELINT systems. A typical airborne threat-warning system consists of a processor, receivers, power supply, antennas, display, and controls, and generates both visual and aural warnings of radar threats. When the pilot has been alerted, the aircraft can be maneuvered to evade the threat or initiate counteraction with onboard ECM capability. The principal radar threats, which appeared in the Vietnamese war and in the Middle East wars of 1971 and 1973, and again during Operation Desert Storm in the Persian Gulf in 1991, were Soviet-made surface-to-air missile (SAM) radar and anti-aircraft gunfire direction radars. Although both are ground-based threats, the warning-receiver concept is equally applicable to radar threats from opposing aircraft. Similarly, warning receivers are carried by ships or other vehicles.

Airborne warning systems usually are crystal video or superheterodyne receivers that indicate the presence of a radar threat and offer a coarse relative bearing to it. An associated receiver and extra antennas augment this capability, automatically determining direction to the threat and giving the pilot a means of homing on the target for weapons delivery by nulling an indicator that centers the aircraft on the signal.

The warning systems also exploit a particular weakness in SAM weapons, which requires that a command radar be turned on prior to missile launch so that the missile can be guided to its airborne target after launch. The receiver can detect characteristic changes in power level of the radar, thereby warning the pilot that a missile is about to be fired. SAM commanders often counter this technique by simply turning on the command radar frequently even when there is no impending target intercept. This can prompt aircraft to prematurely jettison fuel tanks or weapons in an effort to evade the supposed imminent launch of a SAM missile. *See* GUIDANCE SYSTEMS; GUIDED MISSILE; MISSILE.

These procedures illustrate how each advance in radar technology has a countermeasure and how each countermeasure is followed by a counter-countermeasure. Usually the countermeasure and the counter-countermeasure involve a technological advance, but frequently simple ingenuity or changes in tactics supply the new measure.

Intruding aircraft have a time advantage over defense radars because they can detect enemy radar at considerably greater distances than the hostile radar can spot them. The reason for this is that energy emitted by radar decays as a function of the square of the range from the radar to the target. *See* RADAR.

As technology progresses, the difference between the warning system and limited-capability ELINT receivers is narrowing with the advent of small digital processors, particularly for tactical military aircraft. The tactical aircraft needs to be able to react quickly to changes in power levels of specific radar threats, movement of transportable SAM radars, and frequency shifts by frequency agility or frequency hop radars.

Reflectors. One of the oldest passive ECM techniques is the use of chaff. These are metallic strips cut to lengths resonant at the defense radar frequency so that they return spurious radar echoes to enemy radar (**Fig. 2**). Chaff can confuse an enemy by generating false targets, or noise, forcing the enemy to take time to analyze the returns and sort real from false

Fig. 2. Metallized glass chaff, which returns spurious radar echoes. (*Lundy Electronics and Systems*)

targets. Chaff can screen or mask aircraft or higher-speed ships so that the enemy is unable to determine their presence, or can help an aircraft break track once it is alerted by its warning receiver that it is being tracked by radar.

Chaff has evolved from the aluminum strips hand thrown from bombers during World War II to 0.001-in.-diameter (25-micrometer) metallized glass fibers that are automatically ejected in great quantities from aircraft by electromechanical, pyrotechnic, or pneumatic dispensers. Typically, dispensing mechanisms aboard modern military aircraft forcibly eject cardboard packages, roughly $3 \times 5 \times {}^1/_2$ in. ($8 \times 13 \times 1.3$ cm) either manually as initiated by the pilot or automatically if the RWR is interfaced directly with the chaff/flare system. Each package, containing thousands of strands of aluminum-coated glass fibers, is torn open at release. The contents then blossom into clouds within the aircraft's slipstream.

In Vietnam, naval pathfinding aircraft would seed air corridors with chaff through which attacking aircraft would subsequently fly. The chaff masked their specific presence from enemy radars. In addition, strategic bombers are outfitted with chaff-dispensing rockets fired ahead of the bombers to assure that an adequate screen blooms before the aircraft arrives.

By the mid-1970s, U.S. Air Force and Navy tactical fighters were being outfitted to carry large chaff pods capable of dispensing 300 lb (136 kg) of lightweight chaff contained on six separate drums and blown out through separate exit pipes at the rear of the pod (**Fig. 3**). The dispensing rate and the duration and frequency of chaff bursts can be programmed into the pod on the flight line.

To assure availability aboard an aircraft of chaff cut to proper lengths, the military services have investigated dispensing systems that cut chaff to lengths comparable to the frequencies of radars that warning receivers indicate are illuminating the aircraft. A technique known as delayed-opening chaff (DOC), by which chaff packages are deployed from aircraft or ballistic missiles approaching their targets, also was developed. A DOC package is descended by parachute and then blown open by a time fuse. This might deceive an enemy into believing that there are targets far removed from the dispensing aircraft.

Other passive ECM techniques include the use of special radar-absorbing materials, such as ceramics or ferrites, which reduce reflection coefficients so that the amount of radar energy returned to the illuminating radar is reduced. A standard technique is the special shaping of bodies, specifically in missile reentry systems and airframes, that reduces the vehicle's radar cross section (the area of the vehicle that effectively returns echoes to the transmitter). The Air Force has successfully applied these and other signature-reducing techniques, called stealth, to diminish the radar observability of such aircraft as the F-117 fighter and the B-2 bomber. During the Gulf war, the F-117 achieved great success in Desert Storm, and was credited with covering 40% of the strategic targets in Iraq while flying only 2% of the combat sorties by coalition aircraft. The aircraft evaded radar and infrared tracking threats and was credited with being the only aircraft to strike in heavily defended downtown Baghdad. *See* MILITARY AIRCRAFT; RADAR-ABSORBING MATERIALS.

The use of corner reflectors, or Luneberg lenses, which concentrate the energy they reflect back to the radar, can provide a useful ECM aid by confusing an enemy. These reflectors, fixed to drone aircraft

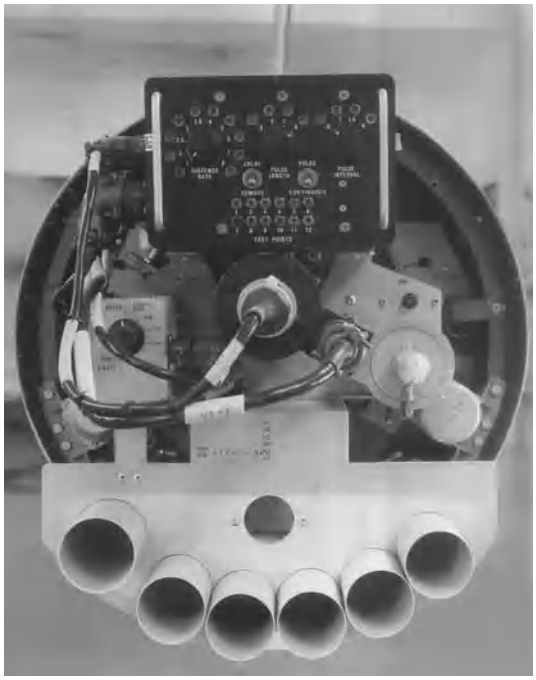

Fig. 3. Bulk-chaff dispensing pod which ejects chaff through six exit pipes.

Fig. 4. Countermeasures-dispensing system aboard an air transport used in humanitarian relief efforts ejecting flares to foil attacks by ground-launched infrared guided missiles. (*Lockheed Aircraft Service Co.*)

released from bombers, produce large returns on enemy radar screens, causing the drones to appear on radar like the larger bombers that released them. *See* ANTENNA (ELECTROMAGNETISM).

Infrared protection. United States aircraft carry sensitive infrared ECM receivers to protect them from heat-seeking infrared missiles. Because such missiles normally are launched toward the rear of the target so that they can home on the heat from the aircraft's jet exhaust, the receivers are located on the aircraft's tail. When the pilot is warned of impending danger from the infrared missile, infrared flares can be ejected through the chaff dispenser either automatically or manually in order to divert the heat-seeking missile.

Following the loss of an Italian transport to infrared guided missiles during a humanitarian relief mission into Sarajevo, Bosnia, in 1993, various United Nations countries began equipping their air-relief transports with missile-warning and flare-dispensing systems (**Fig. 4**). The flares are intended to lure away infrared guided missiles fired from shoulder-held launchers as the aircraft are descending on approach to their destinations.

U.S. Navy and Army aircraft are equipped with active infrared transmission devices designed to induce infrared missiles tracking an aircraft's engine exhaust to break target lock. These transmitters generate high-data-rate infrared pulses to jam signals derived within the missile seeker from the rotating action of the seeker's reticle in chopping incoming infrared energy.

U.S. Army and Marine rotary-wing aircraft have been outfitted with infrared countermeasures systems for protection from heat-seeking missiles, especially at low altitudes where they are vulnerable to attack by ground-fired, infrared-guided missiles (**Fig. 5**). The energy source in the transmitter of one such system radiates heat to foil heat-seeking missiles. The heat pattern generated by the 13-in.-high (33-cm), 28-lb (13-kg) transmitter confuses the missile's guidance and tracking system, throwing the threatening weapon off course.

Active ECM. The many active ECM techniques can be classified broadly either as noise or deception jamming. The former is the oldest, simplest, and most straightforward, but requires higher average power levels and is more expensive. Deception jamming is the more artful and sophisticated technique, operating on the characteristics of the pulse train generated by threat radars.

Noise jamming has a disadvantage in that it can alert enemies to the fact that they are being jammed. Yet in tactical situations where the enemy has little reaction time and noise contributes to confusion, this distinction has little importance.

Communications-jamming techniques include use of white-noise jamming, in which artificially produced white noise modulates the output of a transmitter. This is one of the simplest and easiest approaches, despite its excessive use of power. White-noise jamming is one of the more prevalent forms of noise jamming. In sweep-through jamming the transmitted carrier frequency is swept through a part of the spectrum at a high rate, producing pulses masking the incoming signal. This makes more economical use of transmitter power. *See* ELECTRICAL NOISE.

Deception-jamming techniques are predicated on the idea of operating on pulses received from the enemy so that the signal reradiated from the target deceives the enemy radar or its operators (**Fig. 6**). For instance, the ECM deception set may receive an enemy radar pulse, circulate it through a delay line, amplify it, and reradiate it back toward the enemy. Because the enemy determines the position of the target by the round-trip transit time of the radar energy, its radar decision circuitry will conclude that the target is at a greater distance than it actually is because of the deceptive pulse delay inserted in that round-trip period by the active ECM set. This technique is known as range gate walk-off (RGWO). Velocity gate walk-off (VGWO) is accomplished by applying some form of frequency translation technique such as serrodyning to the returned pulse. Serrodyning is an operation employed by the velocity deception electronic countermeasures (ECM) jammer to counter those radars utilizing a target's

Fig. 5. Transmitter of infrared countermeasures system, providing 360° protection, mounted atop fuselage of UH-IH utility helicopter. (*Sanders Associates*)

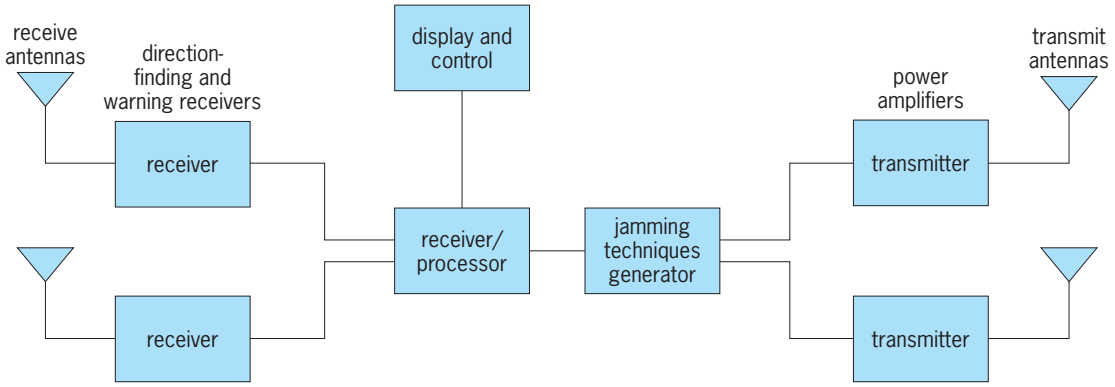

Fig. 6. Typical pulse repeater jamming system.

Doppler shift information. The jammer will frequently translate an incoming signal by a desired amount and at a specific rate during retransmission. If the radar follows the ECM signal, it is left without a target in its velocity gate when the ECM signal is turned off. The operation can be accomplished by either a traveling wave tube (TWT) or a digital phase shifter (DPS).

Similarly, the ECM deception set may operate on the radar pulse train, returning many pulses instead of one, in an effort to deceive the enemy into believing there are many targets spaced at different positions. The genuine return may be blanked out, or false signals may be generated at higher levels so that the enemy may think the genuine return is simply noise. By changing the character of the pulses in the train, the ECM set can deceive the antagonist in a number of ways. Newer ECM systems combine both noise and deception features.

The level of sophistication achieved in modern deceptive electronic countermeasures systems is indicated by a system deployed on some U.S. Navy fighter and attack aircraft. This microprocessor-controlled system employs multiple jamming techniques in countering air, ground, or sea-based pulse radar threats. When threat signals are received and identified by matching them against an emitter library in an electronic data file, the system chooses the optimum active technique to defeat the threat. In a high-threat-density environment, the 190-lb (86-kg) system automatically selects the appropriate countermeasures technique and power level for responding to the threat in accordance with threat priorities. Identification and prioritization of the threat and the choice of jamming response are all handled by the reprogrammable data file. The reprogramming capability enables the user to add or change threat characteristics as they become known.

ECM system. Because of the complex electromagnetic environment in warfare, no single technique is satisfactory for all occasions. A sophisticated warplane may have a penetration-aids subsystem with a number of separate systems operating interdependently to give aural and visual warning of radar threats. A cryogenically cooled infrared receiving set with its own decision circuitry located in the top of the aircraft's vertical stabilizer warns the pilot of threats from heat-seeking missiles. An ECM deception set automatically generates false target returns to tracking-radar threats in several frequency bands. A pneumatically driven chaff dispenser, working with other subsystems, selects and automatically ejects disposable packages including chaff and infrared flares (**Fig. 7**). A wideband radar homing and warning set detects enemy ground and airborne radar threats, identifies and determines the detected threat location, and calculates azimuth and elevation of the threat. It monitors the effectiveness of the ECM deception set and automatically supplies signals to deploy disposables. It also provides bearing

Fig. 7. Programmable countermeasures-dispensing system, which can dispense chaff, flares, and active expendables to protect aircraft from threats. ***(Tracor)***

inputs to the aircraft's inertial navigation computer and drives the lead computing gunsight in an attack mode so that the aircraft can attack a threatening emitter detected by the warning set.

ECM aboard missiles. Many ECM techniques developed for aircraft, ships, and ground operations were investigated and in some instances adapted by the United States to ensure penetration capability of its ballistic missiles. The techniques are similar to what was previously described, although active ECM techniques are less practical because of the severe environments and power limitations. Several ways of using active ECM have been investigated, however, including the possibility of installing jammers within small rocket-fired precursor decoys which would be fired ahead of a warhead. The jammers, each of which is tuned within different frequency bands, spread confusion among enemy radars.

The difficulties in securing extremely rugged microwave transmitting tubes to withstand missile environments and in generating a maximum amount of broadband power with a minimum of weight are formidable. A major hurdle was overcome during the mid-1960s by developing sources of kilowatt dc power having a high specific weight and high specific volume. *See* MICROWAVE TUBE.

To use active ECM properly, an adversary must have prior knowledge about the defensive radars so that ECM jammers can be tuned. In an aircraft or on a ship, there is more likely to be space and power for devices to sense hostile signals and tune transmitters. In a missile, where space and power are at a premium, this ability is far more costly.

In general, however, the military relies for missile penetration more heavily on passive ECM techniques, such as chaff and inflated metallized balloons, which reinforce enemy-radar returns. One concept is to use tethered chaff, which would be ejected from a missile system on atmospheric reentry and carried along near the vehicle to create large masking radar echoes. Regenerative chaff, a technique by which a missile or powered decoy would sequentially eject clouds of chaff, is another.

Trends in ECM. Newer ECM systems display greater integration among various subsystems to tighten up defense capability. They incorporate infrared warning receivers and in some cases new active infrared jamming systems, as well as attack warning from a data link.

The Navy is integrating its standard modular electronic warfare systems carried by a wide variety of ships. In one version, a passive ECM system would be coupled to rocket-propelled decoy and chaff packages to lure away radar-guided antiship missiles detected by the ECM system. A small hand-held active radar decoy can be used against monopulse radars that are particularly difficult to jam by conventional angle jamming.

As part of a continuing practice of updating the ECM capability of military aircraft to meet newly recognized threats, the U.S. Air Force has modified repeatedly its strategic bomber force. Its B-52 bombers have been retrofitted with improved jammers to counter higher-frequency radars carried by interceptors or located on the ground for control of surface-to-air missiles.

Fig. 8. Antenna for shipborne deception ECM system mounted on vessel's superstructure. (*Hughes Aircraft Co.*)

A number of allied aircraft were shot down or damaged by ground-launched infrared missiles during the Desert Storm operation. This experience reinforced efforts to find better sensors, either infrared or pulse-Doppler, to provide rear-sector warning of impending missile threats to aircraft, and to develop suitable missile alert warning systems.

The sinking of an Israeli destroyer by Soviet-made Egyptian cruise missiles in 1967 revealed a glaring inadequacy in the defensive systems of Western naval vessels that prompted a resurgence in shipboard ECM activities. This concern was heightened when the American destroyer *Stark* was taken unawares and was accidentally struck by two cruise missiles fired from an Iraqi aircraft during the Iran-Iraq war. A primary element of defense against a cruise missile is the brief interval of time, as little as 90 s, that a ship has from the instant a threatening missile rises above the horizon until it is upon the ship. This briefness requires that a ship quickly detect and identify the threat, determine its direction of flight, and take appropriate countermeasures. This may best be accomplished by integrating ship sensors, including ECM signal sensing, and identification and direction-finding capabilities, with antiaircraft weapons and improved command-and-control capability.

The Navy has developed systems to protect aircraft carriers and other large warships from radar-guided cruise missiles. One of these is a deception system that senses the threatening radar signal from the missile seeker and radiates a false electronic image of the

target ship, offset by a safe distance from the actual ship. Then the missile will home on this ghost image of the target and fall harmlessly into the sea (**Fig. 8**).

Military events have established and reaffirmed the crucial role played in modern warfare and in strategic and tactical defense by electronic warfare. This was apparent in the Persian Gulf conflict in 1991. Coalition electronic warfare aircraft were a significant factor in suppressing enemy air defenses with active jamming, passive location systems, and antiradiation missile delivery ability. *See* MILITARY AIRCRAFT.

Barry Miller; Paul J. De Lia

Bibliography. D. L. Adamy, *EW 101: A First Course in Electronic Warfare*, 2001; D. L. Adamy, *EW 102: A Second Course in Electronic Warfare*, 2004; J. P. R. Browne and M. T. Thurbon, *Electronic Air Warfare*, 1997; E. J. Chrzanowski, *Active Radar Electronic Countermeasures*, 1990; F. Neri, *Introduction to Electronic Defense Systems*, 2d ed., 2001; R. Poisel, *Introduction to Communication Electronic Warfare Systems*, 2002; A. Price, *The History of U.S. Electronic Warfare*, vol. 1, 1984, vol. 2, 1989; D. C. Schleher, *Introduction to Electronic Warfare*, 1986; R. J. Wiegand, *Radar Electronic Countermeasures System Design*, 1991; R. G. Wiley, *Electronic Intelligence: The Analysis of Radar Signals*, 2d ed., 1993; R. G. Wiley, *Electronic Intelligence: The Interception of Radar Signals*, 1985.

Electronics

The technological area involving the manipulation of voltages and electric currents through the use of various devices for the purpose of performing some useful action with the currents and voltages. This large field is generally divided into two primary areas, analog electronics and digital electronics.

Analog electronics. In analog electronics, the signals to be manipulated take the form of continuous currents or voltages. The information in the signal is carried by the value of the current or voltage at a particular time t. Some examples of analog electronic signals are amplitude-modulated (AM) and frequency-modulated (FM) radio broadcast signals, thermocouple temperature data signals, and standard audio cassette recording signals. In each of these cases, analog electronic devices and circuits can be used to render the signals intelligible. Processing signals in an analog fashion has advantages and disadvantages, discussed below.

Commonly required manipulations include amplification, rectification, and conversion to a nonelectronic signal. Amplification is required when the strength of a signal of interest is not sufficient to perform the task that the signal is required to do. For example, the signal obtained by the piezoelectric transducer on an old-fashioned record player is very weak and does not contain sufficient power to drive a set of speakers hard enough that humans can understand the information that is contained in the signal. The transducer signal must be amplified. The amplification process involves making a replica of the original signal containing sufficient power to perform the required task. However, the amplification process suffers from the two primary disadvantages of analog electronics: (1) susceptibility to replication errors due to nonlinearities in the amplification process and (2) susceptibility to signal degradation due to the addition, during the amplification process, of noise originating from the analog devices composing the amplifier. These two disadvantages compete with the primary advantage of analog electronics, the ease of implementing any desired electronic signal manipulation. The decision concerning whether to implement an electronic operation in analog form will depend upon the relative weights given to this advantage and the disadvantages cited above. *See* AMPLIFIER; DISTORTION (ELECTRONIC CIRCUITS); ELECTRICAL NOISE.

Historically, analog electronics was used in large part because of the ease with which circuits could be implemented with analog devices. This facility was primarily a result of the large variety of devices available, including vacuum tubes, transistors, and some special-purpose integrated circuits. However, as signals have become more complex, and the ability to fabricate extremely complex digital circuits has increased, the disadvantages of analog electronics have increased in importance, while the importance of simplicity has declined. Thus, modern-day circuit designers have opted more frequently to implement their designs digitally. This tendency is also the result of the error-correction possibilities inherent in digital electronics. Digital electronics is the most important class of electronics.

Digital electronics. The advent of the transistor in the 1940s made it possible to design simple, inexpensive digital electronic circuits and initiated the explosive growth of digital electronics. However, computers and other digital circuits had been implemented with electron tubes for at least two decades. Digital signals are represented by a finite set of states rather than a continuum, as is the case for the analog signal. Typically, a digital signal takes on the value 0 or 1; such a signal is called a binary signal. Because digital signals have only a finite set of states, they are amenable to error-correction techniques; this feature gives digital electronics its principal advantage over analog electronics. *See* ELECTRON TUBE; TRANSISTOR.

In common two-level digital electronics, signals are manipulated mathematically. These mathematical operations are known as boolean algebra. The operations permissible in boolean algebra are NOT, AND, OR, and XOR, plus various combinations of these elemental operations. The NOT operation performs an inversion of the digital signal; for example, a 1 signal is taken to 0, and a 0 signal is taken to a 1. The AND operation in boolean algebra is similar to the well-known multiplication operation. The digital signals are multiplied together, and the result is a 0 unless all of these signals are 1; in that case, the result is a 1 as well. The OR operation is closely related to addition. If any single input signal, or combination of input signals, is 1, then the result is a 1;

otherwise, if the inputs are all 0, the result is also 0. The XOR (exclusive or) operation resembles common addition, but the carry information is thrown away. In this case, if just one of the inputs is a 1, the output is a 1; otherwise the output is 0. Digital circuits exist to implement all of these operations, as well as more complex operations comprising a combination of these simple ones. As a result, signals can be manipulated in a large variety of ways; for example, signals can be added to each other or multiplied by each other, or transformations such as the Fourier transform can be made, allowing one type of signal to be converted into another. *See* BOOLEAN ALGEBRA; FOURIER SERIES AND TRANSFORMS; LOGIC CIRCUITS.

Electronic devices. Several common electronic devices serve useful purposes in circuit design. The devices most often used include resistors, capacitors, inductors, diodes, and transistors.

Resistor. A resistor is a two-terminal device that impedes the flow of electrons from one of its terminals to the other. The ratio of the voltage applied across the terminals of the resistor to the current which flows from one terminal to the other is the value of the resistance and is measured in units of ohms. One ohm is 1 volt per ampere. *See* ELECTRICAL RESISTANCE; RESISTOR.

Capacitor. A capacitor is a two-terminal device consisting of a set of parallel plates of conductive material on which electric charge can be stored. The ratio of the voltage applied across the terminals of the device to the steady-state charge stored on the plates is the capacitance of the device. Capacitance is measured in units of farads. One farad is 1 volt per coulomb. *See* CAPACITANCE.

Inductor. An inductor is a two-terminal device consisting of a series of conductor loops which generate a magnetic field. As current flows through the inductor, a magnetic field is created in which energy can be stored. The ratio of the voltage across the terminals of the device to the time-rate-of-change of the current flowing between the terminals is the inductance of the device. Inductance is measured in units of henrys. One henry is 1 volt-second per ampere. *See* INDUCTANCE; INDUCTOR.

Diode. A diode is a two-terminal device that permits electric current to flow in only one direction. The most common modern implementation of a diode takes the form of a semiconductor *pn* junction, that is, a junction between a semiconductor doped *p*-type and one doped *n*-type. *See* DIODE; JUNCTION DIODE; SEMICONDUCTOR DIODE.

Transistor. A transistor is a three-terminal device that uses the voltage on one terminal to control the current flowing between the two remaining terminals. There are several major classes of transistors, including field-effect transistors (FETs), bipolar junction transistors (BJTs), and minor devices such as point-contact transistors. Each type of device has characteristics which make it particularly suitable for certain applications. For example, the FET is relatively easy to fabricate in an integrated-circuit format, and so finds wide use in those types of applications. Generally, transistors are used to control one voltage or current in a circuit with another, in switching or amplifying circuits.

Logic families. A logic family consists of circuits fabricated to operate according to a particular specification. When integrated circuits of the same logic family are wired together on a printed circuit board, typically no buffering of the outputs or inputs is required for proper functionality. Thus, a multichip circuit can be built up almost exclusively from integrated circuits, circumventing the need for most discrete components. Some of the logic families are resistor-transistor logic (RTL), diode-transistor logic (DTL), transistor-transistor logic (TTL), emitter-coupled logic (ECL), and complementary metal-oxide-semiconductor (CMOS) logic.

The most successful logic family is the CMOS family. The widespread use of CMOS arises from its low power consumption and its acceptance of a wide range of voltages, giving it versatility that is not found in the other logic families. The low power consumption characteristic of CMOS is a result of the fact that it comprises FET devices which have extremely high input impedances and low operating bias currents. CMOS logic's ability to accept a wide range of supply and operating voltages is also due to the FET technology used, and serves to simplify interfacing the logic input and output points to the data collection and use points.

Electronic circuit design. Electronic circuits are composed of various electronic devices. In circuits built from discrete components, the components are typically soldered together on a fiberglass board known as a printed circuit board. On one or more surfaces of the printed circuit board are layers of conductive material which has been patterned to form the interconnections between the different components in the circuit. In some cases, the circuits necessary for a particular application are far too complex to build from individual discrete components, and integrated-circuit technology must be employed. Integrated circuits are fabricated entirely from a single piece of semiconductor substrate. It is possible in some cases to put several million electronic devices (transistors, resistors, capacitors) inside the same integrated circuit. Modern computer microprocessors are an example of this type of very large scale integration. Special techniques must be used to design such circuits because of the limited characteristics of some of the devices that can be realized. For example, in an integrated circuit it is very difficult to implement an inductor, and so priority is placed on circuit designs that function without inductors. Many integrated circuits can be fabricated on a single wafer of silicon at one time, and at the end of the fabrication process the wafer is sawed into individual integrated circuits. These small pieces, or chips as they are popularly known, are then packaged appropriately for their intended application. *See* INTEGRATED CIRCUITS; PRINTED CIRCUIT.

The microprocessor is the most important integrated circuit to arise from the field of electronics. This circuit consists of a set of subcircuits that can perform the tasks necessary for computation and are

the heart of modern computers. The microprocessors in current use can copy digital data elements from one transistor array (memory) location to another, load an instruction from memory, execute that instruction (typically to add, or otherwise manipulate, the data stored in memory), and copy the results of various instructions into memory, as well as carry out other processing necessary for computation. Microprocessors are programmed, or told what to do, via a series of instructions which together are called a program. Thus, the microprocessor is an example of a circuit that can be reconfigured on the basis of a set of instructions it receives from the programmer and executes. Microprocessors that understand large numbers of instructions are called complete instruction set computers (CISCs), and microprocessors that have only a very limited instruction set are called reduced instruction set computers (RISCs). Microprocessors have found wide acceptance in a large variety of applications and are used to control everything from the temperature of a private dwelling to the softness (calcium-ion concentration) of the water used in the fabrication foundries of the integrated-circuit manufacturers. *See* DIGITAL COMPUTER; MICROPROCESSOR.

Other circuit designs have been standardized and reduced to integrated-circuit form as well. An example of this process is seen in the telephone modem, used for performing digital communications over telephone or radio-frequency–optical links. Modulation techniques have been standardized to permit the largest possible data-transfer rates in a given amount of bandwidth, and standardized modem chips are available for use in circuit design. *See* MODEM.

The memory chip is another important integrated electronic circuit. This circuit consists of a large array of memory cells composed of a transistor and some other circuitry. As the storage capacity of the dynamic random-access memory (DRAM) chip has increased, significant miniaturization has taken place, to the point where modern DRAM integrated circuits have device feature sizes of approximately 0.2 micrometer. *See* CIRCUIT (ELECTRONICS); SEMICONDUCTOR MEMORIES. David R. Andersen

Bibliography. N. R. Malik, *Electronic Circuits: Analysis, Simulation, and Design*, 1995; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 3d ed., 1992.

Electronvolt

A practical unit of energy often used in atomic, nuclear, and particle physics. It is the energy acquired by an electron (or any singly charged particle) when it is accelerated by a potential difference of 1 V. The unit of energy in the International System (SI) is the joule: 1 eV is equal to 1.6022 10^{-19} J. Although the electronvolt is a non-SI unit, it is often expressed using SI prefixes, for example, meV (10^{-3} eV), keV (10^3 eV), MeV (10^6 eV), GeV (10^9 eV). *See* ELECTRICAL UNITS AND STANDARDS; ELECTRON; METRIC SYSTEM; PHYSICAL MEASUREMENT; UNITS OF MEASUREMENT.

The electronvolt is also commonly used for photon energies. Wavelength and photon energy E are related by equation below.

$$E(\text{eV}) = \frac{1239.84}{\text{wavelength (nm)}}$$

See PHOTON; QUANTUM MECHANICS. Fred Schlachter

Electrooptics

The branch of physics that deals with the influence of an electric field on the optical properties of matter, especially in crystalline form. These properties include transmission, emission, and absorption of light.

An electric field applied to a transparent crystal can change its refractive indexes and, therefore, alter the state of polarization of light propagating through it. When the refractive-index changes are directly proportional to the applied field, the phenomenon is termed the Pockels effect. When they are proportional to the square of the applied field, it is called the Kerr effect. *See* KERR EFFECT; POLARIZED LIGHT; REFRACTION OF WAVES.

Pockels cell. The Pockels effect is used in a light modulator called the Pockels cell. This device (**Fig. 1**) consists of a crystal C placed between two polarizers P_1 and P_2 whose transmission axes are crossed. A crystal often used for this application is potassium dihydrogen phosphate (KH_2PO_4). Ring electrodes bonded to two crystal faces allow an electric field to be applied parallel to the axis OZ, along which a light beam (for example, a laser beam) is made to propagate. The crystal has been cut and oriented in such a way that, in the absence of an electric field, the polarization of light propagating along OZ does not change, and therefore no light is transmitted past P_2. However, when an electronic driver V applies an electric field of the proper magnitude

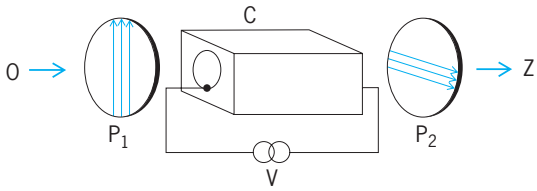

Fig. 1. Typical Pockels cell light modulator. The arrows (*OZ*) represent a light beam.

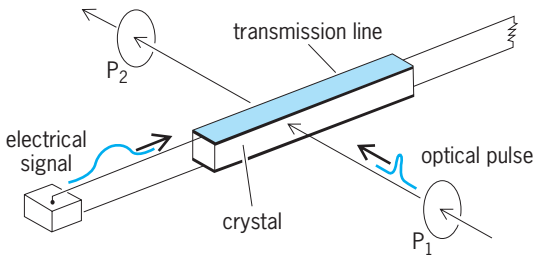

Fig. 2. Electrooptic sampling arrangement using a transmission-line Pockels cell. Examples of optical and electrical waveforms are shown.

(2000 V/cm or 5000 V/in. is a typical value) across the crystal, the crystal becomes birefringent, and light that has been vertically polarized by P_1 undergoes a 90° change in polarization in the crystal, and is now well transmitted past P_2. Pockels cells of this type can be switched on and off in well under 1 nanosecond. *See* BIREFRINGENCE; CRYSTAL OPTICS.

Electrooptic sampling. The linearity and high-speed response of the Pockels effect within an electrooptic crystal make possible a unique optical technique for measuring the amplitude of repetitive high-frequency (greater than 1 GHz) electric signals that cannot be measured by conventional means. The technique is known as electrooptic sampling and employs a special traveling-wave Pockels cell between crossed polarizers, P_1 and P_2 (**Fig. 2**). The electrooptic crystal (commonly lithium tantalate, $LiTaO_3$) is placed between electrodes that are part of a high-speed transmission line along which the unknown electric signal travels. Ultrashort optical pulses (on the order of 1 picosecond) from a mode-locked laser, synchronized with the electric signals and directed through the crystal, are used to repetitively sample a small portion of the electrical waveform. After passing through the second polarizer P_2, the optical pulses have an intensity proportional to the amplitude of that portion of the electrical waveform experienced as the pulse passed through the crystal. By varying the relative delay between the arrival of the optical pulse and electric signal in the crystal and graphing the transmitted intensity as a function of delay, an accurate replica of the electrical waveform is obtained. The availability of subpicosecond optical pulses makes it possible to measure electric signals containing frequencies up to 1 THz with similar temporal resolution. Electrooptic sampling is an important technique for the analysis of ultrafast electric signals such as those generated by high-speed transistors and optical detectors. *See* LASER; MICROWAVE SOLID-STATE DEVICES; OPTICAL DETECTORS; OPTICAL MODULATORS; OPTICAL PULSES. Michel A. Duguay; Janis A. Valdmanis

Self-electrooptic-effect device. A self-electrooptic-effect device (SEED) is a combination of a quantum-well electrooptic modulator with a photodetector which, when light shines on it, changes the voltage on the modulator. Although the device relies internally on an electrooptic effect, the output from the modulator is controlled by the light shining on the photodetector, giving an optically controlled device with an optical output. Most of these devices rely on the quantum-confined Stark effect in semiconductor quantum-well heterostructures as the electrooptic mechanism and utilize the changes in optical absorption resulting from this mechanism.

The quantum-confined Stark effect is conveniently observed by passing a light beam through a diode containing quantum wells in an undoped region (**Fig. 3**) between *p*- and *n*-doped regions. In the simplest self-electrooptic-effect device, this structure also can function simultaneously as the photodetector. The wavelength of the incident light can be so chosen that, as the voltage across the modulator decreases, the absorption of light by the modulator increases. When light is shone on the diode in the circuit of Fig. 3, photocurrent is generated, resulting in a voltage drop across resistor *R*. Consequently, the voltage across the modulator falls, increasing its optical absorption. This absorption increase itself results in a further increase in photocurrent, which in turn leads to a further reduction in voltage across the diode, and so on. This positive feedback can be so strong that it leads to a discontinuous switching into a highly absorbing state, giving the optically bistable input-output characteristic shown in **Fig. 4.** Thus this simple device can function as an optically controlled memory or switch with an optical output.

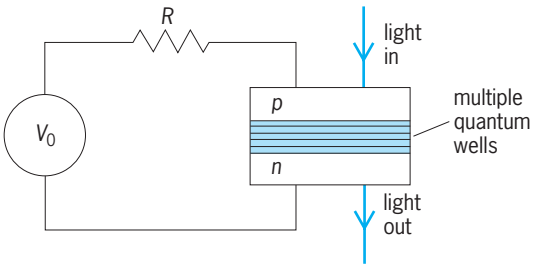

Fig. 3. Sample circuit for a self-electrooptic-effect devices.

Many different configurations of self-electrooptic-effect devices have been proposed for different functions. Substituting an inductor for the resistor gives an oscillator. Substituting a conventional photodiode for the resistor can give a bistable device whose switching thresholds are controlled by the light shining on the photodiode (a diode-based SEED or D-SEED). Substituting another quantum-well diode for the resistor can give a device that is bistable in the power ratio of the two light beams incident on the two diodes (a symmetric SEED or S-SEED). Transistors and more complex circuitry can also be incorporated to give yet more complex relations between optical inputs and outputs. Such devices can be used to perform all the elementary logical functions on light beams.

Self-electrooptic-effect devices can offer operating energy densities comparable to electronic devices when the modulators, photodetectors, and

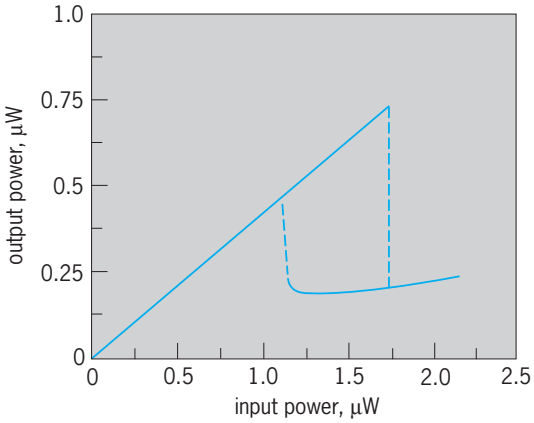

Fig. 4. Set of optical input-output characteristics for an integrated device.

any other circuitry are so integrated as to minimize any undesired capacitance. They can also be fabricated in two-dimensional arrays compatible with optical processing schemes. *See* ARTIFICIALLY LAYERED STRUCTURES; EXCITON; SEMICONDUCTOR HETEROSTRUCTURES; STARK EFFECT. David A. B. Miller

Bibliography. F. Argullo-Lopez, *Electro-Optics: Phenomena, Materials, and Applications*, 1994; H. Haug (ed.), *Optical Nonlinearities and Instabilities in Semiconductors*, 1988; M. A. Karim, *Electro-Optical Devices and Systems*, 1990; M. A. Karim (ed.), *Electro-Optical Displays*, 1990; R. Waynant and M. N. Ediger (eds.), *Electro-Optics Handbook*, 1994; A. Yariv, *Optical Electronics*, 5th ed., 1997.

Electrophilic and nucleophilic reagents

Electrophilic reagents are chemical species which, in the course of chemical reactions, acquire electrons, or a share in electrons, from other molecules or ions. Although this definition embraces all oxidizing agents and all Lewis acids, electrophilic reagents are ordinarily thought of as cationic species, such as H^+, NO_2^+, Br^+, or SO_3 (or carriers of these species such as HCl, CH_3 $COONO_2$, or Br_2), which can form stable covalent bonds with carbon atoms. Electrophilic reagents frequently are positively charged ions (cations). *See* ACID AND BASE.

Nucleophilic reagents are the opposite of electrophilic reagents. Nucleophilic reagents give up electrons, or a share in electrons, to other molecules or ions in the course of chemical reactions. Nucleophilic reagents frequently are negatively charged ions (anions). Typical nucleophilic reagents are hydroxide ion (OH^-), halide ions (F^-, Cl^-, Br^-, and I^-), cyanide ion (CN^-), ammonia (NH_3), amines, alkoxide ions (such as CH_3O^-), and mercaptide ions (such as $C_6H_5S^-$). *See* SUBSTITUTION REACTION.

Joseph F. Bunnett

Electrophoresis

The migration of electrically charged particles in solution or suspension in the presence of an applied electric field. Each particle moves toward the electrode of opposite electrical polarity. For a given set of solution conditions, the velocity with which a particle moves divided by the magnitude of the electric field is a characteristic number called the electrophoretic mobility. The electrophoretic mobility is directly proportional to the magnitude of the charge on the particle, and is inversely proportional to the size of the particle. An electrophoresis experiment may be either analytical, in which case the objective is to measure the magnitude of the electrophoretic mobility, or preparative, in which case the objective is to separate various species which differ in their electrophoretic mobilities under the experimental solution conditions.

Tiselius cell. The phenomenon of electrophoresis was first observed in 1807 by the Russian physicist F. F. Reuss, but electrophoresis was not employed as an experimental technique until the introduction of a new electrophoresis apparatus by Arne Tiselius in 1937. The apparatus of Tiselius detected electrophoretic motion by the moving-boundary method, in which a boundary is created between the solution of particles to be examined and a sample of pure solvent. As the particles migrate in an electric field, the boundary between solution and solvent can be observed to move, and if there are a number of species in the solution with different electrophoretic mobilities, a series of boundaries of various shapes and magnitudes can be detected. Using his apparatus, Tiselius demonstrated the heterogeneity of human blood plasma, and showed for the first time that the globulin molecules could be separated into different classes, which were designated alpha, beta, and gamma globulin. The moving-boundary method was used for three decades to separate complex mixtures of charged macromolecules in solution and to study the physical characteristics of solutions of proteins and other macromolecules of biological and industrial importance.

Gel techniques. The resolving power of electrophoresis was greatly improved by the introduction of the use of gel supporting media. The gel matrix prevents thermal convection caused by the heat which results from the passage of electric current through the sample. The absence of convection reduces greatly the mixing of the various parts of the sample, and therefore allows for more stable separation. The dimensions of the cross-links of the gel may also provide a molecular sieving effect, which increases the resolving power of the electrophoretic separation of molecules of different size. In addition, the gel media may support a gradient of a separate reagent, which assists in the separation of macromolecules. Gradients of pH and of reagents of various types may be combined in two-dimensional arrays for even greater resolving power. A very successful derivative of the gel technique is the determination of the molecular weights of protein molecules by electrophoresis of the molecules in a gel medium which contains substantial amounts of detergent. The detergent denatures the protein molecules, changing them from globular, compact structures to long, flexible polymers which are coated with detergent molecules. These polymers move in the electric field through the gel medium with a velocity which is determined by the length of the polymer, and therefore by the molecular weight of the protein unit. This method is the most common technique for the determination of molecular weights of proteins in biochemical studies. *See* GEL; PH; PROTEIN.

Isoelectric focusing. An important variation of the electrophoresis technique is isoelectric focusing. In this technique the medium supports a pH gradient which includes the isoelectric pH of the species being studied. Many charged macromolecules have both positive and negative charges on their surfaces, and the electrophoretic mobility is related to the net excess of charge of one type or the other. As

the pH becomes more acidic, the number of positive charges increases, and as the pH becomes more basic, the number of negative charges increases. For each molecule of this type, there is one pH at which the net charge on the surface is zero, so that the molecule does not move when an electric field is applied and thus has an electrophoretic mobility of zero. This pH is called the isoelectric pH. If the molecule is introduced into a pH gradient which includes its isoelectric pH, it will migrate to the position of the isoelectric pH and then become stationary. In this way, all molecules of a given isoelectric pH will migrate to the same region—hence the term isoelectric focusing. The method of isoelectric focusing is particularly good for the analysis of microheterogeneity of protein species and other species which may differ slightly in their chemical content. *See* ISOELECTRIC POINT.

Isotachophoresis. One important variant of electrophoresis is the phenomenon of isotachophoresis, in which ionic species move with equal velocity in the presence of an electric field. The isotachophoretic condition is maintained when particles of different mobilities form boundaries within the solution so that the most mobile ions form the leading edge of the moving sample, followed by the less mobile ions in the order of their mobility. The reason that the ions with different mobility can move with the same velocity is that the electric field in each of the regions is inversely proportional to the mobility of the ionic species in order to maintain a constant current throughout the sample. Isotachophoresis has a number of important analytical and preparative applications which are featured by its advantages of high resolving power, sensitivity, speed, and the ability to concentrate rather than to dilute the components which are being analyzed.

Particle electrophoresis. The methodology of electrophoresis may be modified considerably when the particles undergoing analysis are of sufficient size to be viewed either with the naked eye or with the assistance of an optical microscope. This general area of particle electrophoresis has its most important applications in the analysis of the surface charge of living cells and in the study of various types of particles used in industrial coating processes. The most straightforward technique, called optical cytopherometry or microelectrophoresis, is that in which a human experimenter views the particles in an electric field under an optical microscope and determines manually the amount of time necessary to traverse a given distance. Although time-consuming and tedious, this technique has had many important applications. Attempts to modernize this method have included the introduction of high-speed photography, television technology, and the study of particle electrophoresis by the laser Doppler effect. Particle electrophoresis can be conducted under many of the conditions which were originally developed for smaller macromolecules. The use of gradients of density and pH and the methods of isoelectric focusing and isotachophoresis are commonly applied to particles which are even large enough to be viewed with the naked eye (**Fig. 1**).

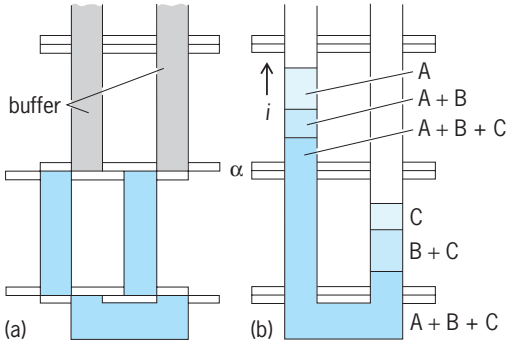

Fig. 1. Diagram of Tiselius electrophoresis cell. (*a*) Formation; buffer plus proteins A, B, and C. (*b*) Motion of electrophoretic boundaries.

A common interference effect in the performance of particle electrophoresis is the sedimentation of the particles in the field of the Earth's gravity. This effect can be minimized in importance by performing the experiments in a medium of equal density, by performing the electrophoresis in a vertical direction and accounting for the effect of sedimentation, or by performing the experiment in such a short period of time that the degree of sedimentation in the Earth's gravity is not a significant interfering effect.

Laser applications. Application of the optical laser to electrophoretic detection resulted in the development of a technique which can be used for analytical electrophoresis experiments on particles of all sizes. The basic principle is that the highly monochromatic (single-frequency) laser light impinges upon the particles, and is scattered from the particles in all directions. When observing the laser light which has been scattered from a moving particle, one can detect that there is a slight shift in the frequency of the light as a result of the motion of the particle. This is the Doppler effect, which causes the change in the apparent tone of passing trains or cars and which is the operating principle of other familiar techniques such as radar. The application of the laser Doppler principle to electrophoresis experiments, often called electrophoretic light scattering (ELS), is an important method for the rapid determination of electrophoretic velocities. The complete electrophoretic mobility distribution of a sample of many particles can be determined in a time as short as 1 s with a precision heretofore unobtainable by standard technology. *See* DOPPLER EFFECT; LASER.

In an electrophoretic light-scattering apparatus with a sample Doppler spectrum (**Fig. 2**), the very slight Doppler shifts caused by the electrophoretic motions are detected by the electronic "beats" which result when the scattered light is incident simultaneously with an unshifted reference beam, or "local oscillator," on a photodetector. An important application of electrophoretic light scattering is the characterization of leukemic cells (**Fig. 3**). The normal lymphocytes and leukemic cells isolated in the same way have distinctly different electrophoretic mobilities and mobility distributions, as shown by

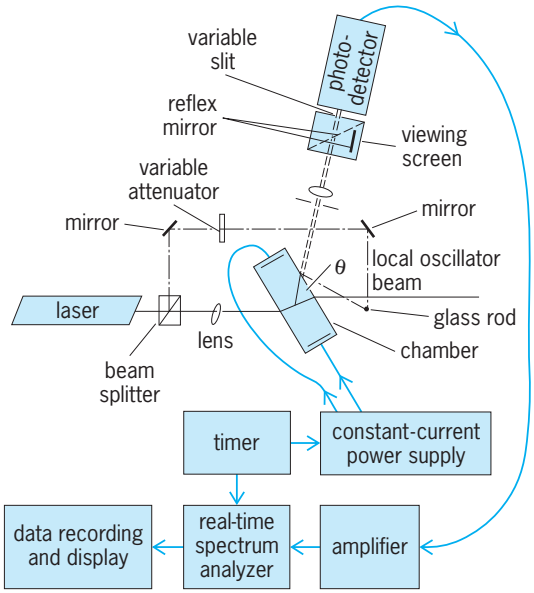

Fig. 2. Diagram of an electrophoretic light-scattering apparatus. Doppler-shifted light from the particles moving in the chamber is mixed with an unshifted reference beam (the local oscillator). The beat frequencies from the photodetector are spectrum-analyzed to produce the electrophorectic mobility histogram. (*After G. M. Hieftje, ed., New Applications of Lasers to Chemistry, ACS Symposium Ser. 85, 1978*)

the two spectra in Fig. 3*a*. The spectrum in Fig. 3*b* shows the simultaneous detection of leukemic and normal cells in a mixture, based solely on their Doppler-detected electrophoretic mobilities. Electrophoretic light scattering has been used for the study of many types of living cells, cell organelles, viruses, proteins, nucleic acids, and synthetic polymers. *See* ELECTROLYTIC CONDUCTANCE; SCATTERING OF ELECTROMAGNETIC RADIATION. B. R. Ware

Capillary electrophoresis. Electrophoresis can be performed in a capillary format. A typical system consists of two reservoirs and a capillary filled with a buffer solution (**Fig. 4***a*). A high voltage is applied across the capillary by using a high-voltage power supply. The very small diameter capillaries (typically 5–100 micrometers) employed in this technique allow for efficient heat dissipation. Therefore, much higher voltages can be employed than those used in slab gel electrophoresis, leading to faster, more efficient separations. Compounds are separated on the basis of their net electrophoretic mobilities.

Most often, the detector is placed on line and analytes are detected as they flow past the detector (Fig. 4*a*). Spectroscopic detection (ultraviolet and laser-based fluorescence) is usually performed in this manner by using the capillary itself as the optical cell. Alternatively, detectors can be placed off line (after the column; Fig. 4*b*). In this case, the detector is isolated from the applied electric field through the use of a grounding joint. Electrochemical detection and mass spectroscopic detection are generally accomplished in this manner, since the electric field can interfere with the performance of these detectors. The several different modes of capillary electrophoresis include capillary zone electrophoresis, micellar electrokinetic chromatography, capillary gel electrophoresis, isoelectric focusing, and isotachophoresis.

Capillary zone electrophoresis. This is the simplest and most widely used form of capillary electrophoresis. The capillary is filled with a homogeneous buffer, and compounds are separated on the basis of their relative charge and size. Most often, fused silica capillaries are employed. In this case, an electrical double layer is produced at the capillary surface due to the attraction of positively charged cations in the buffer to the ionized silanol groups on the capillary wall. In the presence of an electric field, the cations in the diffuse portion of this double layer move toward the cathode and drag the solvent with them, producing an electroosmotic flow. The resulting flow profile is flat rather than the parabolic shape characteristic of liquid chromatography. This flat flow profile causes analytes to migrate in very narrow bands and leads to highly efficient separations. The electroosmotic flow is also pH dependent, and it is highest at alkaline pH values.

In most cases, the electroosmotic flow is the strongest driving force in the separation, and all

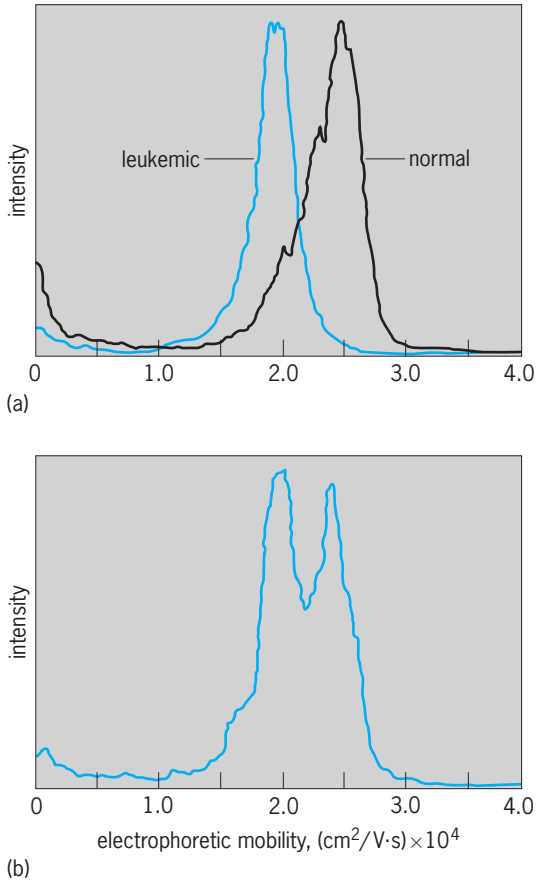

Fig. 3. Characterization of leukemic cells with electrophoretic light scattering. (*a*) Spectra of leukemic cells and normal cells. (*b*) Spectrum for mixture of the two cell types. (*After B. A. Smith, B. R. Ware, and R. S. Weiner, Electrophoretic distributions of human peripheral blood molecular white cells from normal subjects and from patients with acute lymphocytic leukemia, Proc. Nat. Acad. Sci. USA, 73:2388, 1976*)

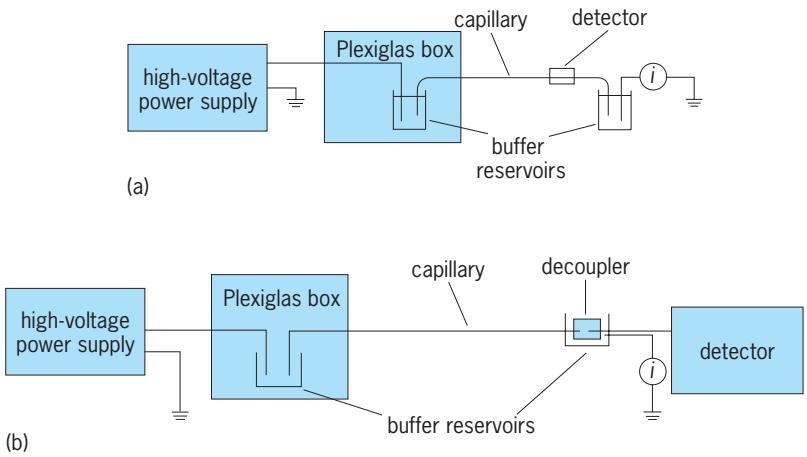

Fig. 4. Capillary electrophoresis. (*a*) Arrangement of a basic system with on-line detection. (*b*) System with off-line detection.

analytes, regardless of charge, migrate toward the cathode. Therefore, it is possible to separate and detect positive, negative, and neutral molecules in the same electrophoretic run, if the detector is placed at the cathodic end. Negatively charged compounds are attracted to the anode but are swept up by the electroosmotic flow and elute last. Neutral molecules, which are not separated from each other in capillary zone electrophoresis, elute as a single band with the same velocity as the electroosmotic flow. Positive compounds have positive electrophoretic mobilities in the same direction as the electroosmotic flow, and they elute first (**Fig. 5**). Capillary zone electrophoresis is generally employed for the separation of small molecules, including amino acids, peptides, and small ions, and for the separation of drugs, their metabolites, and degradation products.

Although the electroosmotic flow that is characteristic of the fused silica capillaries makes it possible to separate and detect both positive and negative ions in a single run, the presence of the negatively charged silanol groups can lead to irreversible adsorption to the capillary surface of proteins containing a large number of positive functional groups (high pI, the isoelectric point of the protein). Therefore, for many of the applications geared toward protein separations, the inside capillary surface is modified to mask the free silanol groups. This modification minimizes protein adsorption, but it also dramatically reduces the electroosmotic flow, leading to much longer separation times.

Additives to the buffer in the reservoirs (the run buffer) can also be used to change the mechanism of separation. For example, cyclodextrins have been employed for the separation of a number of optical isomers. Other additives, such as organic solvents and metal ions, have been utilized to improve selectivity or to change the electroosmotic flow. *See* ELECTROKINETIC PHENOMENA.

Micellar electrokinetic chromatography. In micellar electrokinetic chromatography, an ionic surfactant is added to the run buffer above the critical micelle concentration. Analytes partition in and out of the micelle based on their hydrophobicity (and charge). The migration time is determined not only by site and charge, as in capillary zone electrophoresis, but also by the amount of time that the analyte spends in the micelle relative to that spent in the run buffer. The most common surfactant employed is dodecyl sodium sulfate, which produces negatively charged micelles. In this case, hydrophobic compounds are retained longer on the column because of the high negative electrophoretic mobility of the anionic micelles. The resulting separation is very similar to that obtained by reversed-phase chromatography. This technique is especially useful for the separation of neutral compounds. However, charged compounds can exhibit additional ionic interactions with the micelles, which can affect their migration time. *See* MICELLE.

Capillary gel electrophoresis. This is the version of capillary electrophoresis most similar to conventional

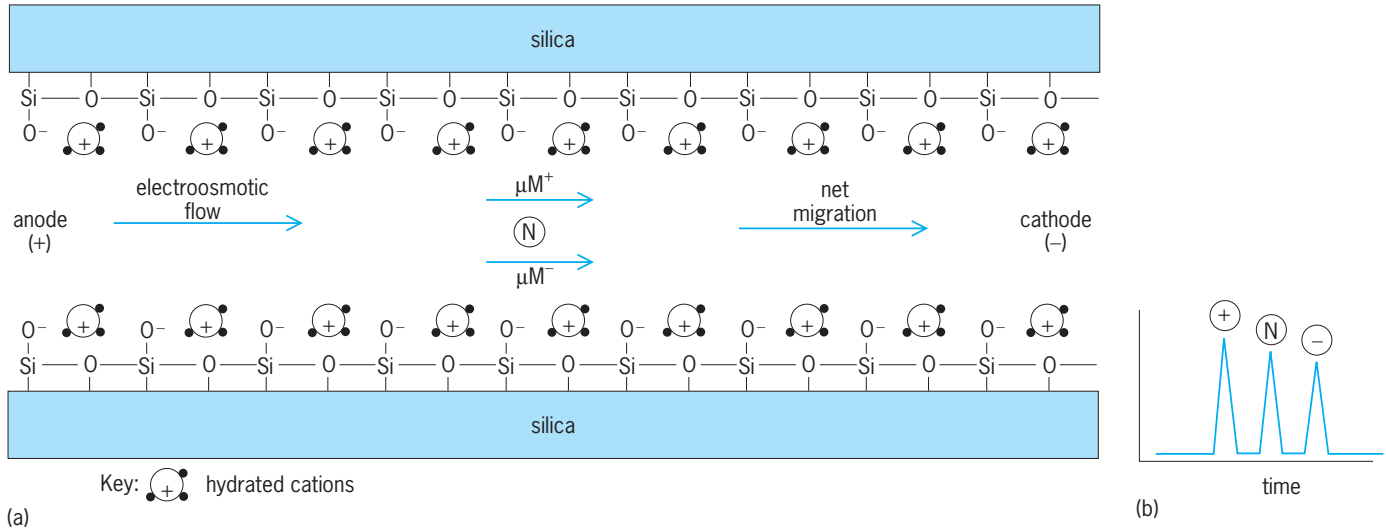

Fig. 5. Capillary zone electrophoresis. (*a*) Separation mechanism showing electrophoretic mobility of the positive ion (μM^+) and negative ion (μM^-); N is a neutral molecule. (*b*) Migration order of the ions.

slab gel electrophoresis. It has been applied extensively to the separation of oligonucleotides and proteins. In this form of capillary electrophoresis, a polyacrylamide gel (or other polymeric material) is placed inside the fused silica capillary, which has been modified so that the electroosmotic flow is essentially zero. The gel can be cross-linked, but separations with non-cross-linked gels have also been accomplished. Separation is based on size and charge. Since oligonucleotides exhibit high negative electrophoretic mobilities, their detection takes place at the anodic end of the capillary. Separations of oligonucleotides differing by only one base can easily be obtained with this technique. *See* OLIGONUCLEOTIDE.

Capillary isoelectric focusing. Isoelectric focusing can also be accomplished in the capillary electrophoresis format. As with capillary gel electrophoresis, a capillary with reduced electroosmotic flow must be employed. A pH gradient is created in the capillary by filling it with a mixture of the solute and a solution of an ampholyte, an electrolyte having both acidic and basic characteristics. For the focusing to take place, one buffer reservoir is filled with a basic solution and the other with an acidic solution. The high voltage is then applied. Under these conditions, both the solute and the ampholytes migrate until they reach a region where the pH is equal to their pI. When the isoelectric focusing process is completed, all charge is neutralized and the separation current goes to zero. The zones can then be mobilized for detection by pressure or through the addition of salt to one of the reservoirs. Proteins differing by less than 0.005 pI unit have been separated by using this technique.

Capillary isotachophoresis. This method, also known as displacement electrophoresis, is a completely different mode of capillary electrophoresis in which all the analytes move at the same velocity. Only one class of ions (anions or cations) can be separated in capillary isotachophoresis. In the case of an anion separation, the capillary is filled with an electrolyte containing an anion that has a higher mobility than any of the analytes. The sample solution is then loaded on the capillary, and the cathodic end of the capillary is submerged in an electrolyte solution containing anions of lower mobility than any of the analytes in the mixture. When the high voltage is applied, the anions in the lead electrolyte, which have the highest electrophoretic mobilities, move quickly toward the anode, producing a gap of low conductivity and high electric field strength. The electric field causes analyte ions to move toward the gap, with those exhibiting the largest electrophoretic mobilities moving first. Eventually, a steady state is reached in which each solute is isolated in its own band of differing field strength and all bands are moving at the same velocity. This method has also been employed as a preconcentration step for other modes of capillary electrophoresis.

All the capillary electrophoresis methods have the advantage of the ability to analyze small volumes (typical injection volumes are 1–50 nanoliters). This makes it possible to analyze very small samples or to use the same sample for several different analyses. One unique application of this technique is the determination of amino acids and neutrotransmitters in single cells. Susan Lunte

Alternating-field electrophoresis. Alternating-field agarose gel electrophoresis is a technique for separating very large molecules of deoxyribonucleic acid (DNA); fragments of DNA ranging in size from 30 to 10,000 kilobasepairs (kb) can be resolved. For the molecular biologist, this is a considerable advance over conventional agarose gel electrophoresis, which is limited to the resolution of less than 50 kb.

Conventional gel electrophoresis employs a single pair of electrodes to generate an electric field that is constant in both time and direction and that is uniform across the gel. DNA molecules are negatively charged with a uniform charge-to-mass ratio and thus migrate steadily toward the positive electrode. Although DNA is a linear molecule, in solution it tends to collapse into a random coil configuration. Agarose is a porous material that acts like a sieve, retarding the movement of the DNA; the larger the molecule (and therefore the larger the random coil), the more the retardation, and thus the molecules separate on the basis of size. However, above approximately 50 kb, the dimensions of the random coil are larger than the pore size of the agarose. The DNA can no longer be sieved through the gel, and resolution is lost.

The DNA can be visualized as a long, snakelike molecule. It tends to align itself parallel to an external electric field because it is uniformly charged along its length. This is the principle behind alternating-field electrophoresis. Contrary to conventional electrophoresis alternating-field electrophoresis does not use a constant electric field but one which regularly alternates in direction. With each change in field direction, the DNA molecules attempt to reorient themselves. When this happens an end or a small loop of the molecule, which has dimensions much smaller than those of the random coil of the entire molecule, may be positioned by a pore in the agarose. The electric field can then pull the DNA by the end through the hole. When the field regularly alternates from one direction to another, the DNA regularly reorients and is pulled, in a snakelike fashion, through an adjacent hole.

Not all molecules in the system will make equal progress under these conditions, because not all molecules can reorient themselves with equal speed. The larger the molecule, the more time it requires in a given field strength (determined by the applied voltage) to change directions and the less time it has left to move. This is the basis for the separation seen in alternating-field electrophoresis. Small molecules will reorient themselves quickly and spend most of the time migrating. Larger molecules will reorient slowly and be unable to move very far. Very large molecules may require such long times for reorientation that there is no time remaining for migration. Kathleen Gardiner

Bibliography. P. Bocek et al. (eds.), *Analytical Isotachophoresis*, 1988; P. Camilleri (ed.), *Capillary*

Electrophoresis: Theory and Practice, 2d ed., 1997; N. Guzman (ed.), *Capillary Electrophoresis Technology*, 1993; B. D. Hames (ed.), *Gel Electrophoresis of Proteins: A Practical Approach*, 3d ed., 1998; R. Weinberger, *Practical Capillary Electrophoresis*, 2d ed., 2000; R. Westermeier, *Electrophoresis in Practice: A Guide to Methods and Applications of DNA and Protein Separations*, 4th ed., 2005.

Electroplating of metals

A process for applying a coating to a metal surface by electrochemical means. Electroplating is extensively used to produce printed circuit boards. Its main advantage is that the circuit can be produced directly rather than having to be etched out of a piece of copper sheet. Electroplating is also widely used to impart corrosion resistance. Most parts of automobile bodies are zinc plated for corrosion resistance. Since zinc is more readily attacked by most corrosive agents that automobiles encounter than steel, it provides galvanic or sacrificial protection. An electrolytic cell is formed in which zinc, the less noble metal, is the anode and the steel, the more noble one, is the cathode. The anode corrodes and the cathode is protected. Zinc also provides a good base for paint. If a metal is more noble than the one on which it is electroplated, it provides protection against corrosion only if it is completely continuous. If a small area of the substrate is exposed such as under a pinhole, corrosion occurs there, rapidly forming a pit. *See* ELECTROCHEMISTRY.

An example of an electroplated coating applied primarily for wear resistance is hard chromium on a rotating shaft. Electroplating is also used to build up worn or undersized parts. Gold-plated jewelry is an example of a decorative application. Gold and also palladium are electrodeposited on electrical contacts. Here the absence of an oxide film avoids the rise in the electrical resistance of the contact. Nickel and chromium are plated for some decorative applications; however, their wide use in the automotive industry has diminished considerably, primarily because of the associated environmental problems. Magnetic components made of such alloys as permalloy can be manufactured by electroplating.

Process

The electroplating process consists essentially of connecting the parts to be plated to the negative terminal of a direct-current source and another piece of metal to the positive pole, and immersing both in a solution containing ions of the metal to be deposited (see **illus.**). The part connected to the negative terminal becomes the cathode, and the other piece the anode. In general, the anode is a piece of the same metal that is to be plated. Metal dissolves at the anode and is plated at the cathode. If the current is used only to dissolve and deposit the metal to be plated, the process is 100% efficient. Often, fractions of the applied current are diverted to other reactions

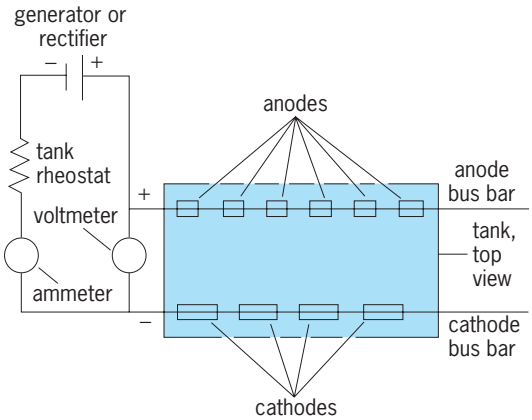

Typical connections for a simple electroplating process.

such as the evolution of hydrogen at the cathode; this usage results in lower efficiencies as well as changes in the acidity (pH) of the plating solution. In some processes, such as chromium plating, a piece of metal that is essentially insoluble in the plating solution is the anode. When such insoluble anodes are used, metal ions in the form of soluble compounds must also be added periodically to the plating solution. The anode area is generally about the same as that of the cathode; in some applications it is larger. *See* ELECTRODE; ELECTROLYSIS.

Most plating solutions are of the aqueous type. There is a limited use of fused salts or organic liquids as solvents. Nonaqueous solutions are employed for the deposition of metals such as aluminum that have overvoltages lower than hydrogen. Such metals cannot be plated in the presence of water, as hydrogen would be preferentially reduced. *See* OVERVOLTAGE.

In addition to metal ions, plating solutions contain relatively large quantities of various substances used to increase the electrical conductivity, to buffer, and in some instances to form complexes with the metal ions. Relatively small amounts of other substances, which are called addition agents, are also present in plating solutions to level and brighten the deposit, to reduce internal stress, to improve the mechanical properties, and to reduce the size of the metal crystals or grains or to change their orientation.

The quantity of metal deposited, that is, the thickness, depends on the current density (A/m^2), the plating time, and the cathode efficiency. The current is determined by the applied voltage, the electrical conductivity of the plating solution, the distance between anode and cathode, and polarization. Polarization potentials develop because of the various reactions and processes that occur at the anode and cathode, and depend on the rates of these reactions, that is, the current density. If the distance between anode and cathode varies because the part to be plated is irregular in shape, the thickness of the deposit may vary. A quantity called the throwing power represents the degree to which a uniform deposit thickness is attained on areas of the cathode

at varying distances from the anode. Good throwing power results if the plating efficiency is low because of polarization where the current density is high. A plating solution such as an alkaline cyanide bath, in which the polarization of the cathode increases strongly with increasing current density, has good throwing power. A plating solution such as an acid-sulfate copper bath, which is almost 100% efficient, has poor throwing power.

Two other deposition processes are closely related to electroplating: electroless plating and displacement plating. Both processes require no applied current. Electroless plating is a process which, once initiated, continues; that is, it is autocatalytic. Displacement plating occurs when the metal deposited is more noble than the substrate and the substrate dissolves. The reaction ceases when the substrate is completely covered by a pore-free deposit. *See* ELECTROLESS PLATING.

Preparation

In order for adherent coatings to be deposited, the surface to be plated must be clean, that is, free from all foreign substances such as oils and greases, as well as oxides or sulfides.

Cleaning. Three principal methods are employed to remove grease and attached solids. (1) In solvent cleaning, the articles undergo vapor degreasing, in which a solvent is boiled in a closed system and its vapors are condensed on the metal surfaces. Many solvents used in the past have been found to be toxic. (2) In emulsion cleaning, the metal parts are immersed in a warm mixture of kerosine, a wetting agent, and an alkaline solution. (3) In electrolytic cleaning, the articles are immersed in an alkaline solution, and a direct current is passed between them and the other electrode, which is usually steel. Heavily soiled articles are cleaned in solvent or alkaline spray machines. Cleaning solutions may contain sodium hydroxide, carbonate, phosphate, and metasilicate, plus wetting agents and chelating agents. More highly alkaline solutions are used for steel than for other metals. Most of the cleaning is accomplished by the scrubbing action of the evolved gases and the detergency of the components of the solution. The articles may be connected as anodes, as is usual for steel, or as cathodes, as is usual for other metals. Electrolytic cleaning is usually the last cleaning step.

Ultrasonic cleaning is also used extensively, especially for blind holes or gears packed with soils. Ultrasonic waves introduced into a cleaning solution facilitate and accelerate the detachment of solid particles embedded in crevices and small holes. Frequencies from about 18,000 to 24,000 Hz are usually employed. They produce cavitation, which causes rapid local circulation. *See* CAVITATION; ULTRASONICS.

Pickling. In this process, oxides are removed from the surface of the basis metal. For steel, warm, dilute sulfuric acid is used in large-scale operations because it is inexpensive; but room-temperature, dilute hydrochloric acid is also used for pickling because it is fast acting. In cathodic pickling of steel, attack of the metal is retarded while the oxide is being dissolved. In addition to rough pickling, acid treatments to activate the surface just prior to plating are often used.

Hydrogen embrittlement may be caused by the diffusion of hydrogen into steel during pickling and also in the certain plating operations. Especially with high-carbon steels, hydrogen causes cracking (a reduction in the fatigue strength and ductility). Hydrogen is gradually evolved on standing, and more rapidly evolved by heating to about 390°F (200°C). *See* DIFFUSION; EMBRITTLEMENT.

Electropolishing. Electropolishing is the preferential, electrochemical dissolution of asperities on surfaces. It results in a smoother and brighter surface. Thin electrodeposits tend to copy the morphology of the surface on which they are plated. Therefore, to attain a bright deposit, the surface on which it is plated must also be bright, a property which can be attained by electropolishing. The electropolished surface is smooth, clean, and free of deformed material, making it an excellent surface for plating. *See* ELECTROPOLISHING.

Equipment

The two classes of electroplating equipment are electrical and mechanical.

Electrical equipment. Most plating operations are conducted with direct current at potentials of 3–12 V. Since electricity is delivered as alternating current at 220 V, it is necessary to reduce the voltage and to rectify the current. Both motor generators and rectifiers are used. Selenium, germanium, and silicon rectifiers have been used extensively in plating. Special generators are used for pulse plating; the pulse variables are often computer controlled. *See* DIRECT-CURRENT GENERATOR; SEMICONDUCTOR RECTIFIER.

Mechanical equipment. The plating tanks are either poly(vinyl chloride) or steel, which requires no lining for alkaline solutions. For neutral and acid solutions, steel tanks are lined with rubber or plastic. For chromic acid baths, lead or plastic linings are used. Most large plating operations are conducted in conveyor tanks. In semiautomatic conveyors, the cathode racks are carried only through the plating tank. In fully automatic conveyors, the cathodes are carried successively through tanks that contain cleaning, pickling, and plating solutions, with intermediate rinses.

Plating of continuous strip is another important operation. Rolls of copper foil are produced by this method. A rotating hard-chrome-plated drum is partially immersed in the plating solution, where it is the cathode. Copper is deposited on the immersed part of the drum. As the copper foil does not adhere to chromium, it peels off as it emerges from the solution. After rinsing and drying, the foil can be rolled up. Tin is also plated on steel by the continuous-strip method.

Small objects are plated in "barrels," usually of hexagonal shape, with perforated plastic sides. The

barrels rotate on a horizontal axis in a tank containing the plating solution and anodes. The articles contact cathode connections as they tumble during the rotation of the barrels.

Plating of Specific Metals

Most metals can be electroplated from either aqueous or fused-salt solutions. The more important metals plated from aqueous baths are chromium, copper, gold, nickel, silver, tin, and zinc. Alloys can also be electroplated. Electrodeposited copper-zinc and lead-tin alloys are used extensively.

Chromium. Electroplated chromium is used primarily to produce wear- and corrosion-resistant coatings. Chromium is not deposited for decorative purposes as extensively as in the past because of the associated pollution problem from the discharge of hexavalent chromium. Chromium plating solutions consist primarily of chromic oxide, sulfuric acid, and water. The solutions in which a hard chrome finish is deposited contain a lower ratio of chromic oxide to sulfuric acid than the ones used for the decorative type, and are maintained at a somewhat higher temperature. Decorative chromium is plated in a high state of internal stress, generally leading to cracking except in very thin deposits. If there are only a few cracks, the undercoat, which is generally nickel, may be subjected to very high local corrosion currents. To avoid this problem, the chromium deposits are intentionally microcracked or made porous in order to increase the exposed area of the undercoat and thus reduce the current density. The throwing power and current efficiency of chromium plating solutions are rather poor. *See* CHROMIUM.

Copper. Electroplated copper is used extensively in the manufacture of printed circuit boards. As copper cannot be directly plated on the insulating-material substrates, they must be rendered electrically conductive first. The main advantage of using electrolytically deposited copper to produce printed circuits is that the areas of the board that are made conductive can be controlled. The actual circuit can be produced by selective etching or selective plating involving techniques such as photosensitizing, photoresist, and etch resist. Areas exposed through a mask can become conductive by coating them with electroless copper after activation with a solution of stannous and palladium chloride. Suitable organic materials are also used to render the board conductive. The areas made conductive serve as the substrate for copper electrodeposition. Some circuits are produced only with electroless plating of copper. Only the through holes are plated, generally with electroplated copper over electroless plated copper when the boards are made by laminating copper sheet to the plastic substrate. The copper segments of the printed circuit may be coated with an electroplated tin-lead alloy to facilitate subsequent soldering and also to protect them from oxidation. Gold deposits are also used for this purpose. There are other uses of electroplated copper in the electronic industry, such as in the production of microchips. Electroplated copper is also the undercoat for decorative nickel-chromium deposits. *See* COPPER; PRINTED CIRCUIT.

Gold. Gold plating is used for electrical contacts which must remain free of oxides, connections for microcircuits, certain information-storage devices, sold-state components, and jewelry. The use of gold for electrical and electronic application exceeds that for decorative purposes. Electroplated palladium is being substituted for gold in some applications. *See* GOLD; PALLADIUM.

Nickel. Nickel coatings covered with chromium provide corrosion-resistant and decorative finishes for steel, brass, and zinc-base die castings. The most widely used plating solution is the Watts bath, which contains nickel sulfate, nickel chloride, and boric acid. All-chloride, sulfamate, and fluoroborate plating solutions are also used. The sulfamate solution is used for low-stress applications. There are a number of compounds, mostly organic, which can be added to nickel-plating baths. When specularly bright nickel is plated, at least one from each of two groups of compounds called class I and class II brighteners must be added to the plating solution. Class I compounds, for example, benzene and naphthalene disulfonic acids, contain sulfur. Class II brighteners, for example, butyne diol, do not contain sulfur. Class II brighteners can also cause leveling, which is producing a deposit that is smoother than the surface on which it is plated. When class II brighteners are added to plating solutions without sulfur-containing agents, they result in semibright nickel deposits. For improved corrosion protection, a duplex nickel deposit, that is, a relatively thick, semibright, sulfur-free plate covered with a thinner bright coating, is frequently used. A thin chromium layer is plated over the bright nickel. If corrosion begins through a pore or crack in the chromium, it will penetrate to the more noble semibright nickel and then spread laterally. Thus, rust spots that result from penetration to the steel are less likely to develop. Exposure of the skin to nickel-plating solutions can cause a rash and severe itching. The sensitivity of persons to nickel salts varies considerably.

Electroless nickel deposits are used for decorative purposes on nonmetallic substrates. Adhesion occurs mainly mechanically by the penetration of the deposit into crevices caused by a prior selective-etch treatment of the substrate. The reducing agents in electroless nickel-plating solutions are either hypophosphites or borohydrides. The deposits are highly supersaturated solid solutions of phosphorus or boron in nickel. The grain size of electroless nickel containing less than 7% by weight of phosphorus is extremely small. Deposits of higher phosphorus content can be considered to be amorphous and are also nonmagnetic. Electroless nickel can also be precipitation-hardened by heat treatment. Electroless nickel plated on metallic substrates produces a uniform thickness over irregularly shaped parts. *See* HEAT TREATMENT (METALLURGY); NICKEL; SOLID SOLUTION.

Silver. The principal use of silver electrodeposits is for tableware because of their corrosion resistance

(except to sulfur-containing foods) and pleasing appearance. Other important uses are for bearings and electrical circuits, waveguides, and hot-gas seals. The plating solutions are of the cyanide type and generally contain additives that produce bright deposits. *See* SILVER.

Tin. Tin is used in electrodeposition as a component of solders. Solder is electroplated over copper to protect it from oxidation and to facilitate subsequent joining operations. The advantage of electroplating the solder is that it can be applied only where it is needed. Because of the desirability of eliminating lead in solders, those with higher tin contents are used.

Tin-plated steel for cans in which food is preserved has limited use, because lacquers are preferred to prevent contact between steel and food. Tin plating is employed for refrigerator coils and bearing surfaces. Bright pore-free surfaces can be produced by melting the electroplated tin and allowing it to "reflow."

Stannous chloride and stannous sulfate are the main components of acid tin-plating solutions. The alkaline solution contains sodium stannate or potassium stannate. Both types are used in the production of electrolytic tin plate. The main components of solutions for the electrodeposition of solders are lead fluoroborate and tin fluoroborate. Sulfonates are also used. *See* TIN.

Zinc. The sacrificial protection of steel against corrosion is the main reason for plating zinc. It is used extensively in the automobile industry for this purpose. Screws, bolts, and washers are barrel-plated with zinc also for corrosion protection. Continuous zinc electroplating of wire and strip is another application. The advantage of electrodepositing zinc over hot dipping is the ability to apply thinner coatings and higher purity. A conversion film formed by dipping in a chromate solution inhibits the formation of white corrosion products. Acid plating solutions are used preferentially over alkaline cyanide ones because of the pollution associated with the latter. *See* ZINC.

Properties of Electrodeposits

In the various applications of electrodeposits, certain properties must be controlled and, therefore, must be measured. Specifications have been supplied by the American Society for Testing and Materials, and by other technical societies and some governmental agencies, establishing the various properties and the methods of determining them. The properties of electroplated metal which should be considered, depending on the use of the deposit, are thickness, adhesion to the substrate, brightness, corrosion resistance, wear resistance, the mechanical properties of yield strength, tensile strength, ductility, and hardness, internal stress, solderability, density, electrical conductivity, and the magnetic characteristics.

Thickness. Several properties depend on the thickness of an electrodeposit and how it varies over a work piece. Several types of gages and also microscopic, gravimetric, and chemical methods are used to measure thickness.

X-ray fluorescence is most often used to measure deposit thicknesses. When the electrical conductivity of the substrate and the deposit differ considerably and neither is ferromagnetic, eddy-current gages can be used to measure deposit thickness. Beta backscatter gages can measure the thickness of a deposit if its atomic number is considerably different from that of the substrate, as is the case with gold and nickel. *See* NONDESTRUCTIVE EVALUATION; X-RAY FLUORESCENCE ANALYSIS.

Microscopic methods of thickness determination require that the cross section be metallographically prepared and be precisely perpendicular at the interface between substrate and deposit or at a known angle of inclination. The thickness of thin deposits should be determined by scanning electron microscopy, while optical microscopes have adequate resolution for thick ones.

Chemical thickness measurements involve a calibration of the time needed for a reagent to penetrate to a certain depth. Then the time for the same reagent to reach the substrate is measured and the deposit thickness calculated. If the substrate can be dissolved without attacking the deposit, the latter can be weighed. If the area and the density are accurately known, the average thickness can be calculated.

Adhesion. Good bonding between substrate and deposit is very important in almost all plating applications. Poor adhesion results in peeling or blistering of the deposit. Generally, the adhesion between two metals is strong. Poor bonding results if the substrate surface was not clean prior to plating, if foreign substances were absorbed, or if brittle phases form as a result of interdiffusion at elevated temperatures. The bond between a nonmetallic substrate and the electroless plate is primarily mechanical and can therefore vary considerably. Adhesive coatings on the substrate can improve the bonding to electroless deposits.

Adhesion tests are not well standardized. In the peel test, which is widely used, part of the deposit is loosened from the substrate, and the force required to pull off the rest is measured. A tensile test, in which the bonded interface is placed perpendicular to the applied force, is also used. Scraping tests and attempts to interpose a knife edge between substrate and deposit are examples of qualitative tests.

Brightness. For an electrodeposit to be specularly bright, the hills and valleys of the surface morphology should not vary in height by more than the wavelength of light. This requirement is generally fulfilled when the crystal size is very small. Addition agents in the plating solutions are also absorbed on faster-growing sites, permitting the microscopic recessed region to catch up. Quantitative measurements of brightness are rarely performed; comparisons by eye are generally made.

Corrosion. With sacrificial coatings, corrosion protection occurs as long as enough of the coating remains; thus protection depends mainly on the thickness of the plate. In order for a substrate to be

protected against corrosion by a deposit that is more noble, the coating must be completely free of pores, cracks, or discontinuities. The substrate must not be exposed at any point to the corrosive medium. When the less noble substrate is completely covered, the corrosion resistance of the coating itself is enhanced if the structure and chemical compositions are uniform.

Corrosion testing is performed either under actual or simulated service conditions or under accelerated exposure. The conditions under which accelerated exposure tests are performed are more severe than those encountered in service, so that the results can be seen more quickly. Salt-spray and acetic acid–modified salt-spray accelerated corrosion tests are most frequently used. The copper-accelerated acetic acid–salt-spray test and a test in which parts are exposed to a slurry of clay in a salt solution are examples of the second type. These tests are used to duplicate under accelerated conditions the corrosion of plated automotive parts in winter conditions. Electrochemical measurements of the corrosion current density and potential can also give an indication of corrosion resistance. *See* CORROSION.

Wear resistance. Although a high hardness value is an indication of good wear resistance, there are factors relating to wear resistance that are not well understood. There are also several types of wear, the two most common types being adhesive and abrasive wear. Corrosion and lubrication can greatly affect wear. Wear resistance is generally determined by moving two contacting surfaces relative to one another and noting the weight loss after some period of time. *See* WEAR.

Mechanical properties. In many respects, the mechanical properties of electrodeposits are similar to those of metals that have been heavily plastically deformed, that is, cold-worked, in that they exhibit higher strength and hardness and lower ductility than the same materials in the as-cast or annealed state. The number of crystal defects such as dislocations and lattice vacancies in electrodeposits is of the same order of magnitude as that found in severely cold-worked metals. Electrodeposited metals also recrystallize on heating. The small crystal size which electrodeposited metals frequently possess, and the inclusion in the crystal lattice of foreign materials from the plating-bath additives and side reactions, contributes to the strength. Many electrodeposited metals also have a large number of growth twins, which can increase strength. *See* CRYSTAL DEFECTS.

In general, the same factors that increase strength reduce ductility. Annealing electrodeposits can reduce strength and improve ductility. However, both strength and ductility are reduced upon annealing if a brittle grain-boundary phase is formed, as in the case of nickel deposits from plating solutions to which certain sulfur-containing agents have been added. Pores due to included gas also lower both strength and ductility. *See* GRAIN BOUNDARIES; METAL, MECHANICAL PROPERTIES OF.

The mechanical properties of electrodeposits can be determined by a tensile test. However, it is not advisable to use the tensile specimen used to test metal sheet, because the ratio of the width to the thickness is generally too large. It is preferable to scale the dimensions. Care must be taken so that the specimen is accurately aligned and the edges are not serrated from cutting. The bulge test is also used. Hardness is usually measured by an indentation test on a metallographically prepared cross section of the deposit. The Knoop indentor is particularly useful for evaluating relatively thin deposits. *See* HARDNESS SCALES.

Internal stress. Most electrodeposits are in a state of internal tensile stress. The stress can be reduced or made compressive by adding certain materials to the plating bath. The addition of saccharin to nickel-plating solutions has this effect. The causes of internal stress are not fully understood. There are indications that the coalescence of crystals or crystallites is one important cause of tensile stress. The diffusion of hydrogen so as to cause the surface layers to shrink can be another cause. Filling of voids with hydrogen gas so as to expand them can result in compressive stress.

Internal stress can be measured from the bending of a strip plated on only one side or from the change in length of a specimen plated on both sides. A long strip wound into a spiral, with one end fixed and the other free, is the most frequently used stress-measuring device. If this spiral is plated on one side, stress causes the free end to rotate. The rotation is calibrated to yield a stress value.

Solderability. Solderability is an important property, especially for plated segments of printed circuits. Solderability depends on the cleanliness of the surface, especially the degree of oxidation after prolonged storage. Coating copper deposits by plating them with solder, tin, or gold can prevent oxidation and thus improve solderability. Solderability is generally determined by how well molten solder pools spread. The spreading is also affected by the temperature and the flux. Because of the different coefficients of thermal expansion of the printed-circuit substrate and the copper deposit, high stresses can develop on heating for soldering. These stresses can result in the fracture of thin and brittle plated segments, especially at through holes. *See* SOLDERING.

Density, and electrical and magnetic properties. There are no special methods of measuring the density of electrical and magnetic properties of electrodeposits. The density of electrodeposits can be reduced because of the presence of gas or other foreign materials. Lattice vacancies can result in higher electrical resistivities of plated metals. The very fine grain size of electrodeposited or electroless plated ferromagnetic metals results in high coercive forces.

Special Applications

Special processes associated with electrodeposition include electroforming, anodizing, pulse plating, and laser plating.

Electroforming. Electroforming is a special type of plating in which thick deposits are subsequently removed from the substrate, which acts as a mold. The

process is particularly suitable for forming parts that require intricate designs on inside surfaces, for example, waveguides.

Intricate machining operations can be performed much more easily on outside surfaces. First, the outside surface of the substrate mold is machined; then, the contours of the design can be transferred to the inside surface of the deposit; and, finally, the deposit is separated from the mold. The mold or matrix can be either metallic or nonmetallic. Nonmetallic molds must be rendered electrically conductive by the application of a powder or by chemical reduction, electroless plating, or vapor deposition. For nonadherent deposits, substrate removal is easy; otherwise, the substrate must be dissolved or melted away. Important applications of electroforming are in the production of phonograph record masters, printing plates, and some musical instruments as well as in waveguides. *See* WAVEGUIDE.

Anodizing. In anodizing, a process related to plating, an oxide is deposited on a metal that is the anode in a suitable solution. The process is primarily used with aluminum, but it can be applied to beryllium, magnesium, tantalum, and titanium. Relatively thick oxide deposits can be produced, even though they are electrically insulating. The presence of small pores through which the solution can reach the metal surface permits the continuation of the reaction. The electrical current is carried by the electrolyte in the pores.

Solutions for anodizing aluminum generally contain either chromic or sulfuric acid. After anodizing, the pores should be sealed, to improve protection of the substrate, by a hot-water or steam treatment, which causes hydration and a resulting volume expansion of the oxide. Colored coatings can be produced by the incorporation of dyes.

Pulse plating. The maximum current density in electrode position is limited by the replenishment by diffusion or convection of the metal ions in the solution layer adjacent to the work piece. If the current is pulsed so that it is off for a certain time interval, higher current densities can be used because ions are replenished during the off time. The resulting higher deposition potentials can result in grain refinement, which, in turn, can lead to brighter and stronger deposits without the use of addition agents. Internal tensile stresses can also be reduced by pulse plating.

Laser plating. In many electronic applications it is necessary to produce very fine lines of deposits. When laser radiation is directed onto a substrate, the adjacent solution is heated and the plating rate so greatly increased that for all practical purposes deposition is limited to the illuminated area. Thus a laser can be used to draw the desired deposit design on the substrate. Directing a jet of plating solution onto the substrate can accomplish essentially the same result. Laser and jet plating can also be combined to give even better control over the dimensions and shape of the deposit. Rolf Weil

Bibliography. J. W. Dini, *Electrodeposition: The Materials Science of Coatings and Substrates,* 1993; L. J. Durney (ed.), *Graham's Electroplating Handbook,*4th ed., 1984; F. Lowenheim (ed.), *Guide to the Selection and Use of Electroplated and Related Finishes,* 1982; L. T. Romankiew and D. R. Turner (eds.), *Electrodeposition Technology: Theory and Practice,*1987; W. H. Safranek, *Properties of Electrodeposited Metals and Alloys,* 2d ed., 1986; P. W. Wild, *Modern Analysis for Electroplating,* 2d ed., 1991.

Electropolishing

A method of polishing metal surfaces by applying an electric current through an electrolytic bath in a process that is the reverse of plating. The metal to be polished is made the anode in an electric circuit. Anodic dissolution of protuberant burrs and sharp edges occurs at a faster rate than over the flat surfaces and crevices, possibly because of locally higher current densities. The result produces an exceedingly flat, smooth, brilliant surface.

Electropolishing is used for many purposes. The brilliance of the polished surface makes an attractive finish. Because the polished surface has the same structural properties as the base metal, it serves as an excellent surface for plating. Electropolishing avoids causing differential surface stresses, one of the requirements for the formation of galvanic cells which cause corrosion. Because no mechanical rubbing is involved, work hardening is avoided. Contaminants, which often are associated with the use of abrasives and polishing compounds, are also avoided. The surface is left clean and may require little or no preparation for subsequent treatment or use. Electropolishing also minimizes loss of high-temperature creep-rupture strength.

In electropolishing, the work is submerged in an electrolyte and connected to the positive terminal (anode) of a source of direct-current (dc) power. The negative terminal is connected to an electrode that will resist chemical interaction with the electrolyte. Carbon is often used. The electrolyte is usually a concentrated acid, although alkaline solutions, dilute acids, and salts have been successfully employed.

Current density on the work surface is crucial. The relationship of anode current to anode voltage

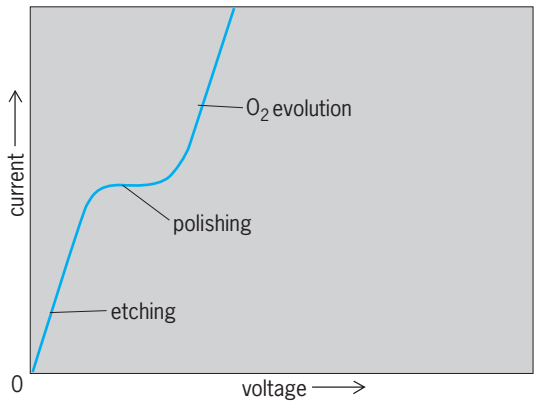

Anode current versus anode voltage in electropolishing.

Electropolishing processes

Metal to be polished	Electrolyte	Temperature, °F (°C)	Current density, A/ft^2 (A/m^2)	Polishing time, min
Aluminum	Phosphoric acid and ethylene glycol	180 (82)	140 (1500)	2–5
Brass and copper	Copper cyanide, zinc cyanide, and potassium cyanide	80 (27)	50 (500)	0.1–0.5
Molybdenum and niobium	Sulfuric acid and hydrofluoric acid in methyl alcohol	80–130 (27–54)	4500 (50,000)	0.05–0.4
Silver	Silver (powder), potassium cyanide, and potassium carbonate	80–125 (27–52)	15–20 (160–200)	0.1–0.2
Steel	Sulfuric acid and phosphoric acid, chromic or humic acids may be added	160 (71)	15–25 (160–270)	2–5
Tantalum and tungsten	Sulfuric acid and hydrofluoric acid in methyl alcohol	80–130 (27–54)	4500 (50,000)	0.05–0.4
Zinc*	Chromic oxide	65–90 (18–32)	1500 (16,000)	1–1.5

*After treatment, zinc surfaces must be successively dipped in dilute solutions of potassium chromate (2 min) and hydrochloric acid (5 s).

is shown in the **illustration**. Below a certain level of voltage, etching occurs. Above that level a constant-current region is reached in which polishing occurs. One mechanism proposed to explain this effect suggests that a film, or a systematic structuring of solution contents and ions, forms over the work surface and that this film, being thinner over burrs and sharp protuberances, allows selective dissolution of the surface to cause flatness. Whether or not this is correct, raising the voltage to a level sufficient to cause oxygen evolution results in solution agitation at the work surface and thereby interferes with polishing.

Acids commonly used as electrolytes, either alone or in combination, include acetic, chromic, citric, hydrofluoric, nitric, phosphoric, and sulfuric, with phosphoric playing a major role in many processes. In most processes the temperature of the electrolyte is usually maintained in the range 90–250°F (30– 120°C). Current densities vary widely, but 100–500 A/ft^2 (1100–5400 A/m^2) of surface will remove moderate amounts of metal. Current densities up to 5000 A/ft^2 (54,000 A/m^2) may be employed if large amounts of metal are to be removed. Polishing times vary from a few seconds to 15 min or longer (see **table**).

In general, higher electrolyte temperatures, higher current densities, and longer polishing times all tend to produce brighter finishes. The range of finishes includes all gradations from satin to mirror, control being obtained through selection of electrolyte, solution concentration, operating temperature, and polishing time.

Numerous metals and alloys electropolish well. Aluminum, both wrought and cast, can be electropolished and anodized to yield a finish that rivals chromium plate. Brass, copper, chromium plate, gold, nickel, nickel alloys, carbon steel, the 300 and 400 stainless steels, and zinc can all be electropolished. *See* ELECTROPLATING OF METALS.

William W. Snow

Bibliography. R. M. Burns and W. W. Bradley, *Protective Coatings for Metals*, 3d ed., 1975; D. R. Gabe, *Principles of Metal Surface Treatment and Protection*, 2d ed., 1978; V. Sedlacek, *Metallic Surface, Films, and Coatings*, 1993.

Electroscope

An instrument for detecting the presence and sign of an electric charge. It is the simplest type of ionization chamber. *See* IONIZATION CHAMBER.

The **illustration** shows a common type of simple gold-leaf electroscope. Gold leaf, shown as L, is used because it is an extremely thin conducting foil which has low mass per unit area and is very flexible. Hence, it responds quickly and vigorously to small electrostatic forces. Aluminum foil is almost as satisfactory as gold foil. In the illustration P is a metal support post for L; I is an insulator of high quality through which P passes and terminates in a metal knob K; and H is a cylindrical metal housing with flat ends and windows so located that the motion and final position of L are visible. H serves as an electrostatic shield that is connected to the environment, as well as a shield against air currents. The base B supports the electroscope.

The hard-rubber rod R (illus. *a*), which has been given a net charge by rubbing, has set up the charge distribution shown, by the process of electrostatic induction, since the isolated structure consisting of K, P, and L has to remain in total electrically neutral. P and L have the same sign of charge and so L is repelled from P. With R still present, K is connected to H, and then K, P, L, and H have a total induced positive charge. If K is disconnected from H and finally R is taken away, the remaining positive charge on K, P, and L induces a negative charge on H (illus. *b*). At this stage the electroscope has a positive charge on its leaf system. This charge and the leaf deflection decay in minutes through leakage currents across the surface of I.

If an electroscope has a charge of known sign, as in illus. *b*, it can be used to test the sign of an unknown charge, as in illus. *c*, where the metal test ball T, with its insulating handle J, has the unknown charge. In the situation pictured, L moves farther away from P

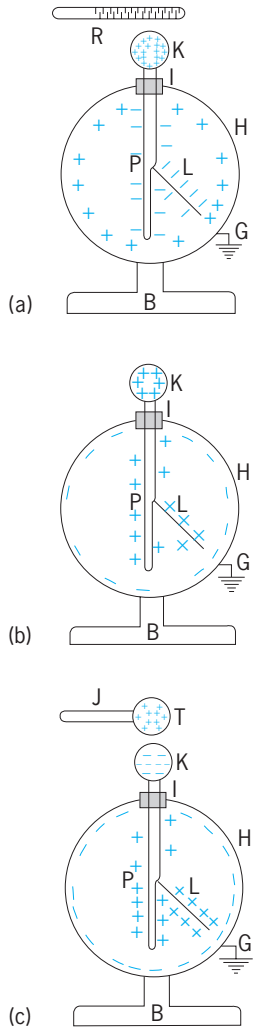

Electroscope. (*a*) Being charged by induction by negative charge on hard-rubber rod R. (*b*) Positive charge left on its leaf after induction process is complete. (*c*) Testing the sign of an unknown charge on test ball T.

as T is brought slowly up toward K, showing that T has a positive charge. If T had a negative charge, L would move toward P, as T slowly approaches K. The converse situation, if the leaf system in illus. *c* had a negative charge initially, can be readily visualized.

Although electroscopes have been built with a wide variety of geometries, the principle of operation is essentially the same for all. If an electroscope has a scale, permitting quantitative measurements, it is called an electrometer or electrostatic voltmeter. For information on electrometers *See* ELECTROMETER; VOLTMETER. Ralph P. Winch; Bryan P. Kibble

Bibliography. A. F. Kip, *Fundamentals of Electricity and Magnetism*, 2d ed., McGraw-Hill, 1969; Physical Science Study Committee, *Physics*, 7th ed., Kendall Hunt, 1991.

Electrostatic discharge testing

A stress testing procedure for determining the electrostatic discharge (ESD) susceptibility level of electrical and electronic devices and systems.

Beginning in the mid-1970s, active electronic devices such as semiconductor components and integrated circuits became susceptible to fast transient events called electrostatic discharge events. The cores of these more sensitive electronic devices began failing after human handling because there was no built-in protection against the fast transients from charged humans. Over the next 25 years, ESD design engineers were able to design-in ESD protection circuits to match the advanced technology changes. However, at the end of 1990s the dramatic increase in the combination of high speed, low power, high performance, and high pin count (over 1000 pins) resulted in extremely dense circuits in devices and system-in-package (SIP) assemblies. These, when combined in one device or assembly, led to increased sensitivity of the devices and therefore increased susceptibility to the fast ESD transients. *See* ELECTRIC TRANSIENT; ELECTROSTATICS; INTEGRATED CIRCUITS; STATIC ELECTRICITY.

There are four such ESD events: (1) the transient from a charged human, (2) the transient from a charged machine or piece of equipment, (3) the transient from the charged package of the device, and (4) the transient from a charged human holding a piece of conducting material. Each ESD event is unique, and therefore the associated pulses have very different parameters.

Since the mid-1970s, ESD stress testing methods have measured the differences in the failure voltage thresholds or the failure current thresholds of these ESD-sensitive devices to determine their susceptibility. ESD stress testing simulates (using a tester) what happens when a real-world ESD event "strikes" the pin of an ESD-sensitive device. The core of such a device is susceptible to the fast ESD transients, so ESD protective circuits are placed at every pin or pad of the device.

There are four ESD testing instruments in use to stress the electronic devices and assemblies, one for each of the different ESD events. The testing techniques are detailed in four corresponding ESD standards, which describe methods or procedures for stressing the devices and determining their susceptibility. With the advent of high-performance and high-pin-count devices came the need for modern ESD testing instruments and techniques, the major issue being one of added parasitics (added capacitance and inductance) in the ESD tester. These older tester parasitics modified the shape and other parameters (such as pulse rise time and pulse peak current) associated with the pulse from the ESD event.

Basic principles of stress testing. The basic principles associated with ESD stress testing are founded on the four different ESD events. The first three events each cause a single device or component to fail, and the testers that simulate these events are therefore used to qualify the devices as to their level of susceptibility before they are shipped. The fourth event (human + metal) has a different transient that causes assemblies to temporarily stop functioning (and requires rebooting) but does not destroy the devices, and the tester that simulates this event is

Data measured at 500 V for the four ESD model events

ESD transients	*RC* (Ω/pF)	Peak current, A	Rise time, ns	Decay time, ns	Scope bandwidth
Human body (fingers)	1500/100	0.33	2–10	150	500 MHz
Machine	0/200	8.75	8–15	66–90	350 MHz
Charged device	25/4	7.50	<0.2	<0.4	1–6 GHz
Human + metal	330/150	2.375	0.7–1.0	30	1.0 GHz

therefore used to qualify assemblies. Each device type must be stressed by each of the first three ESD transient types, and at several voltage levels, beginnning with the lowest (hundreds of volts) and ending with the highest (thousand of volts). At each voltage level, several devices are stressed to ensure that the results are repeatable and reproducible and that process variance is properly taken into account.

The **table** details the differences between the pulse parameters of the four transient ESD events. The values of the main circuit elements (resistance and capacitance) in the tester that simulate each ESD event are shown in the second column. The next three columns show the peak current, rise time, and decay time for each event. The scope displays the transient pulse from which the parameters can be recorded, and the different bandwidths confirm that each of the four ESD events is very different from the others.

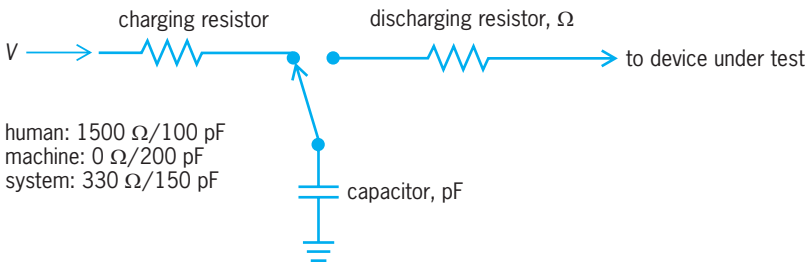

Fig. 1. *RC* (resistance–capacitance) network for three ESD events: human body, machine, and human + metal (system). Resistance and capacitance are shown for each network.

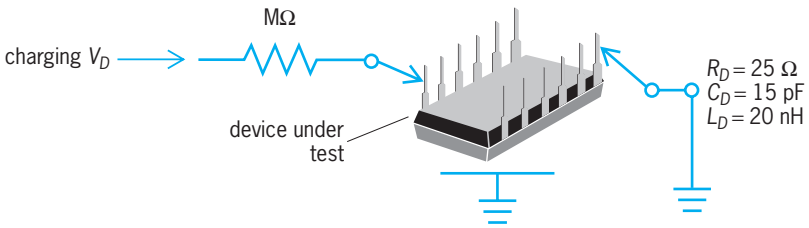

Fig. 2. *RCL* (resistance–capacitance–inductance) network for charged-device ESD event.

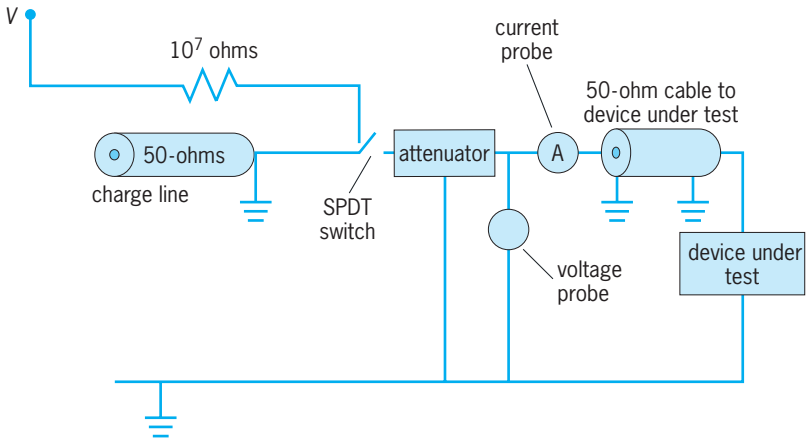

Fig. 3. TLP (transmission line pulse) tester in constant-impedance (50-ohm load) mode.

Traditional ESD testers. Figures 1 and **2** show the basic simulators with the resistance–capacitance (*RC*) network set up to test the devices. The actual testers have more parasitics (stray capacitances and inductances) than shown here, and although the ESD standards specify limits, the parasitic effects are large enough to modify the waveforms. Figure 1 can be used to represent three ESD simulators, depending on the *RC* network specified. For example, substituting the *RC* value of 1500 Ω/100 picofarads will produce the human-body circuit for simulation. Figure 2 represents the charged-device simulator-tester, where charging the device occurs through the V_{ss} (ground) bus pins or the V_{cc} (high-voltage) bus pins of the device. Here, R_D is the device resistance, C_D is the device capacitance, and L_D is the inductance. (The value $C_D = 15$ pF is an example from actual stressing of a real device, while the value of 4 pF in the table is used as a calibrator.)

Need for new stress testers. In addition to the inadequacies of additional parasitics in traditional ESD testers, the actual data provided only pass/fail results. Using this go/no-go procedure, the device fabrication processes are changed after a failure, and then the devices are redesigned and retested. This iteration is continued until the desired level for device operation is achieved. However, the increased number of I/O pins coupled with the need for faster and denser structures resulted in increased ESD sensitivity. To design a good ESD protection structure, the designer must understand the basic device failure mechanisms as well as the mechanisms of the various integrated circuit protection structures. By the mid-1980s, the sensitivity of the cores of electronic devices increased to such a high level that a new ESD test method, the transmission line pulse (TLP) ESD test method, had to be developed. This method used the cleaner and more reliable rectangular or square pulse to stress the device, and was employed in these early years to simulate and correlate with the human-body discharge event. The TLP method provided additional information beyond the human-body go/no-go test method.

The TLP ESD test method was used as an in-house laboratory diagnostic tool for engineering evaluation, characterization, and development of the

protection structures, with the goal of improving the robustness of the ESD protective structures in integrated circuits. It did so by allowing the engineer to determine precise current paths in the device and hence to develop improved and new circuits. The failure criteria in TLP testing were linked to the failure criteria of the traditional human-body ESD testing. The TLP pulse parameters and testing results were compared with those from the traditional human-body simulator in order to show that correlation exists between the two methods of ESD testing.

However, the need for higher pin count devices combined with the increased sensitivity of high-preformance semiconductor devices led to the need for more improvement in the TLP test method. The first commercially developed TLP tester-simulator was completed in 1999, and the first TLP standard test procedure was developed and published in 2003. It has since evolved from a laboratory diagnostic tool to a standard practice used for engineering evaluation and development, including failure analysis and subsequent redesign for increased robustness of the protective structures.

TLP tester. There are two basic TLP ESD testing configurations: the constant-current (500 → 1000-ohm load) mode and the constant-impedance (50-ohm load) mode (**Fig. 3**). The 50-ohm constant-impedance mode has inherent higher current capability and is charged by the 50-ohm charge cable line. The attenuator is used to reduce the unwanted reflections. The single-pole double-throw (SPDT) switch is used to go between charging and discharging from the high-voltage charge line. The TLP tester stresses the device under test in voltage steps starting from less than 1.0 V to as high as 500 V. Since a short circuit in this configuration produces 10 amperes, the impedance of the device will determine the device response current and corresponding response voltage.

The ESD protection structure of the device under test responds to the stresses by producing voltage and current values, so a plot of the device data results in a special current–voltage (*I-V*) curve, much like that seen on a curve tracer plot. Here, however, additional points and parameters [trigger voltage (V_{t1}) and trigger current (I_{t1}), holding voltage (V_{sb}), and device impedance slope] are obtained (**Fig. 4**). The TLP stress pulses are increased until the device produces a secondary breakdown point called V_{t2} and I_{t2}, where failure begins.

The current–voltage–leakage (*I-V-L*) plot tells about the complete performance of the ESD protection structure. The leakage current $I_{L\ device}$ is plotted against the device response current I_{device}. This dynamic *I-V*-L curve shows one "leakage current" point for each stress voltage. Figure 4 shows that there is no change in the leakage current (top *x*-axis scale) up to a specific current value I_{t2} (*y*-axis scale). Then there is a catastrophic change in the leakage current to coincide with the device second breakdown voltage, V_{t2}, also referred to as the total failure point. The gradual device leakage current evolution (from picoamperes

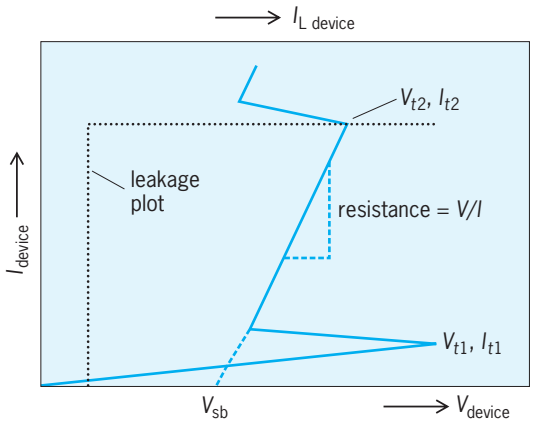

Fig. 4. Current–voltage–leakage (*I-V-L*) characteristic curve from a transistor device.

to nanoamperes) can change drastically well before the usually defined total failure point. The leakage is taken after each stress pulse because it can determine if a failure has occurred or if an early failure is to be expected. Leakage data are collected at the dc operaing voltage of the device. These data are critical because *I-V* plots alone cannot give any information about possibly early failures; only the leakage plot can.

General ESD stress testing will continue to be the simplest way to check and ensure the voltage or current level at which devices will continue to operate after being stressed by an ESD (transient) pulse. The human-body tester has been the main simulator for the human-body ESD event, but the new TLP test method provides information that is not available using the traditional methods. *See* ELECTRONIC TEST EQUIPMENT.

Leo G. Henry

Bibliography. J. E. Barth et al., TLP calibration, correlation, standards, and new techniques, *IEEE Trans. Electr. Pack. Manuf.*, 24(2):99–108, April 2001; ElectroStatic Discharge Association, *Electrostatic Discharge (ESD) Technology Roadmap*, 2005; H. Geske, DVI complaint ESD protection to IEC 61000-4-2 level standard, *Conformity*, pp. 12, 14–17, September 2004; L. G. Henry, All ESD testing standards are not created equal: A Rosetta Stone analysis, *Compliance Eng.*, July-August 2003; L. G. Henry et al., Charged device model metrology: Limitations and problems, *Microelectr. Reliab.*, 42:919–927, 2002; L. G. Henry and M. Chaine, Component level ESD testing, chap. 6 in ElectroStatic Discharge Association, *ESD Phenomena and the Reliability for Microelectronics* (white paper), 2002; T. J. Maloney and N. Khurana, Transmission line pulsing techniques for circuit modeling of ESD phenomena, *EOS-7: EOS/ESD Symposium*, p. 49, 1985.

Electrostatic lens

An electrostatic field with axial or plane symmetry which acts upon beams of charged particles of uniform velocity as glass lenses act on light beams. The action of electrostatic fields with axial symmetry is

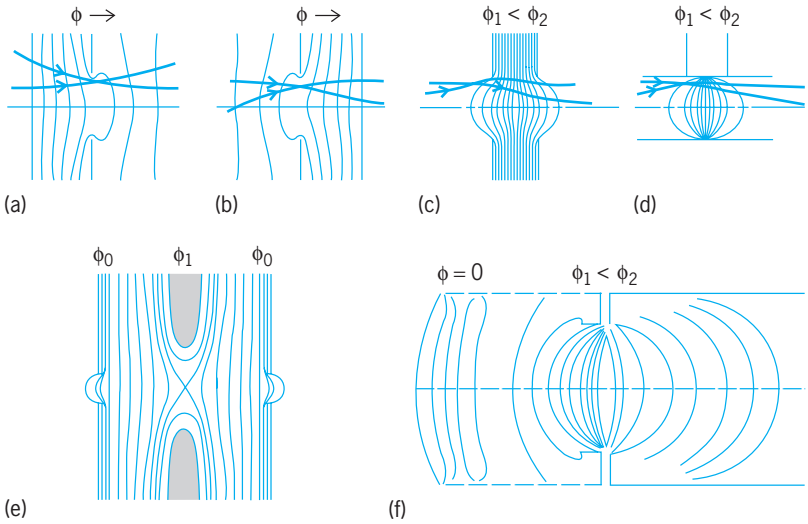

Fig. 1. Axially symmetric electrostatic lenses. (*a*) Single-aperture lens (decreasing field). (*b*) Single-aperture lens (increasing field). (*c*) Two-aperature lens. (*d*) Two-cylinder lens. (*e*) Unipotential lens. (*f*) Cathode lens (image tube). (*After E. G. Ramberg and G. A. Morton, J. Appl. Phys., vol. 10, 1939, and V. K. Zworykin et al., Electron Optics and the Electron Microscope, Wiley, 1945*)

analogous to that of spherical glass lenses, whereas the action of electrostatic fields with plane symmetry is analogous to that of cylindrical glass lenses. Plane symmetry as used here signifies that the electrostatic potential is constant along any normal to a family of parallel planes.

The action of an electrostatic lens on the paths of charged particles passing through it is most readily visualized with the aid of an equipotential plot of the fields in a plane of symmetry of the lens. The equipotential lines in the plot indicate the intersection with the plane of the drawing of surfaces on which the electrostatic potential is a constant. The paths of charged particles in the electrostatic field are bent toward the normals of the equipotentials as the particles are accelerated, and away from the normals as the particles are decelerated. *See* ELECTRON MOTION IN VACUUM.

Axially symmetric lenses. These lenses are generally formed at or between circular apertures and cylinders maintained at suitable potentials. A number of such lenses are shown with characteristic path plots in **Fig. 1**. For any of these it is possible to define focal points, principal planes, and focal lengths in the same manner as for light lenses and to determine with their aid image magnification for any object position (**Fig. 2**). For a thin electrostatic lens in particular, that is, a lens for which the extent of the variation in potential is small compared to its focal length, the object side focal length f_0 and the image side focal length f are given by Eq. (1). Here

$$\frac{\phi_o^{1/2}}{f_o} = \frac{\phi_i^{1/2}}{f_i} = \frac{3}{16}(\phi_o\phi_i)^{1/4}\int\left(\frac{\phi'}{\phi}\right)^2 dz \qquad (1)$$

$\phi(z)$ is the potential along the axis of the lens, ϕ its derivative with respect to z (that is, the electric field along the axis), and ϕ_o and ϕ_i are the potentials in object and image space, respectively. The integration is extended over the lens field. The quantity ϕ is here normalized so that it is equal to the accelerating potential of the particle in question.

Axially symmetric lenses are commonly divided into the four classes that follow.

Simple aperture lenses. These are the lens fields formed about circular apertures in a plane metallic electrode at potential ϕ with different electrostatic fields $-\phi'_o$ and $-\phi'_i$ on the two sides. In most cases the focal length f of such a lens is given to a sufficient degree of accuracy by the Davisson-Calbick formula for an aperture, Eq. (2). Simple aperture lenses are

$$\frac{1}{f} = \frac{\phi'_i - \phi'_o}{4\phi} \qquad (2)$$

encountered as parts of more complex electrostatic lens systems, as well as at the mesh openings of metal screens employed as electrostatic shields in vacuum tubes.

Bipotential, or immersion, lenses. In these lenses image space and object space are field-free, but at different potentials. Typical examples are the lenses formed between apertures or cylinders at different potentials (Fig. 1*c* and *d*). If the separation d of the two apertures is large compared to their diameters and if each component aperature lens satisfies the conditions for a validity of the Davisson-Calbick formula, the focal lengths of the bipotential aperture lens are given by Eq. (3). The distances of the principal planes

$$\frac{1}{f_o} = \left(\frac{\phi_i}{\phi_o}\right)^{1/2}\frac{1}{f_i} = \frac{3}{8d}\left[1-\left(\frac{\phi_o}{\phi_i}\right)^{1/2}\right]\left(\frac{\phi_i}{\phi_o}-1\right) \qquad (3)$$

from the plane of symmetry are given by Eqs. (4) and (5). Generally, the principal planes are displaced

$$h_o = \frac{-d}{2} - \frac{4d\phi_o}{3(\phi_i-\phi_o)} \qquad (4)$$

$$h_i = \frac{d}{2} - \frac{4d\phi_i}{3(\phi_i-\phi_o)} \qquad (5)$$

from the plane of symmetry toward the low-potential side, with the image-side principal plane closer to object space than the object-side principal plane.

For two cylinders of equal diameter D, whose difference in potential is small compared to their mean

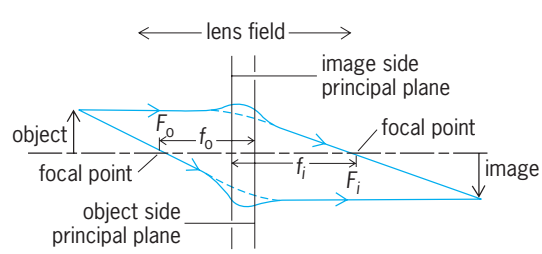

Fig. 2. Definition of principal planes, focal points, and focal lengths for axially symmetric electrostatic lens.

potential, Eq. (6) gives the focal lengths.

$$\frac{1}{f_o} = \left(\frac{\phi_i}{\phi_o}\right)^{1/2} \frac{1}{f_i}$$
$$= \left(\frac{\phi_i}{\phi_o}\right)^{1/4} 0.66 \left(\frac{\phi_i - \phi_o}{\phi_i + \phi_o}\right)^2 \frac{1}{D} \qquad (6)$$

Bipotential lenses, in particular lenses formed between two cylinders at different potentials, find wide application in beam-focusing devices such as electron guns. Like unipotential lenses, they invariably act as converging lenses.

Unipotential lenses. For this type the potentials are equal in object and image space. In their simplest form these lenses consist of three apertures of which the outer two are at a common potential ϕ_o and the central aperture is at a different, generally lower, potential ϕ_1. For such lenses with a central aperture of diameter D and the two outer apertures, of smaller diameter, separated a distance D from the plane of symmetry, the weak-lens focal length is given by Eq. (7).

$$\frac{1}{f} = \frac{0.2}{D}\left(\frac{\phi_o - \phi_1}{\phi_o}\right)^2 \qquad (7)$$

As ϕ_1 approaches zero, the quantity $1/f$ increases more rapidly than indicated by this formula; it attains a value of $0.7/D$ for $\phi_1 = 0$. Unipotential lenses operated at high potentials (for example, $\phi_o = 50$ kilovolts, $\phi_I = 0$) are employed as objectives and projection lenses in electrostatic electron microscopes. The electrodes are commonly made out of stainless steel and given a high polish. *See* ELECTRON MICROSCOPE.

Cathode lenses or immersion objectives. Here the lens field extends from the emitter surface up to field-free image space. Examples are the cathode region of an electron gun, the electron-optical system of an electrostatic image tube or image converter, and the objective of an emission electron microscope. In the electron gun the cathode lens converges the electrons emitted by the cathode to a small spot, the crossover, which is imaged by a second electron lens as the scanning spot on the cathode-ray tube screen. *See* CATHODE-RAY TUBE.

In the image tube the cathode itself—a transparent photoemissive surface on which a light picture is projected—is imaged on a fluorescent screen beyond the cathode lens. Frequently a cathode lens consists essentially of a uniform accelerating field followed by a short lens. The image magnification m is then given by Eq. (8). Here u is the distance between cathode

$$m = \frac{v}{2u} \qquad (8)$$

and short lens and v is the distance between short lens and image. The quantity v is given by Eq. (9),

$$1v = \frac{1}{f} - \frac{1}{2u} \qquad (9)$$

where f is the focal length of the short lens.

Lenses of plane symmetry. These lenses, analogous to cylindrical glass lenses, are formed between parallel planes and at slits, replacing the circular cylinders and apertures of lenses with axial symmetry. For the simple slit in an electrode at potential ϕ separating two regions of field $-\phi'_o$ and $-\phi'_i$, the focal length is given by the Davisson-Calbick formula for a slit as shown by Eq. (10).

$$\frac{1}{f} = \frac{\phi'_i - \phi'_o}{2\phi} \qquad (10)$$

Edward G. Ramberg

Electrostatic precipitator

A device used to remove liquid droplets or solid particles from a gas in which they are suspended. The process depends on two steps. In the first step the suspension passes through an electric discharge (corona discharge) area where ionization of the gas occurs. The ions produced collide with the suspended particles and confer on them an electric charge. The charged particles drift toward an electrode of opposite sign and are deposited on the electrode where their electric charge is neutralized. The phenomenon would be more correctly designated as electrodeposition from the gas phase. The practical aspects of the electrostatic precipitator were demonstrated in 1906 by F. G. Cottrell.

Construction. In its simplest form the experimental arrangement may consist of a vertical tube containing an insulated concentric wire (**Fig. 1**). A direct current potential of 10–100 kV may be applied to the center wire. A corona discharge occurs in a small area surrounding the wire. The suspended particles are ionized in the corona discharge and migrate to the wall of the vertical tube. If the suspended particles

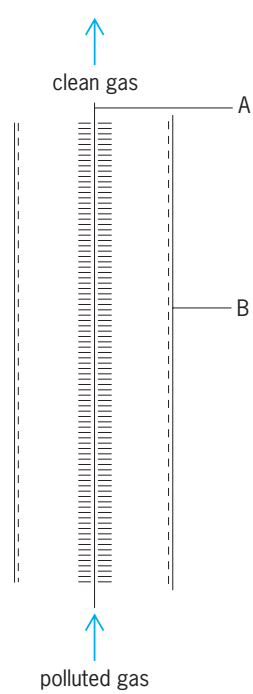

Fig. 1. Diagram of simple precipitator. A, corona wire; B, grounded tube.

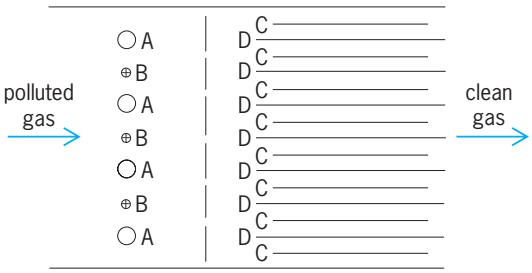

Fig. 2. Diagram of two-stage precipitator. A, grounded cylinders; B, corona wires; C, grounded collector plates; D, charged plates of polarity similar to B.

are liquid, they will accumulate on the wall and coalesce into droplets which can be drained away from the bottom of the tube. Solid particles may be displaced by mechanical vibration or scrapers and discharged into a conical collector at the bottom of the tube.

In more elaborate arrangements the ionization may occur in one vessel, whereas the deposition occurs in a second stage. A simplified two-stage apparatus is illustrated in **Fig. 2**. In the first chamber the particles become charged but are prevented from depositing on the grounded cylinders by suitable adjustment of the rate of flow of the gas. In the second chamber (consisting of alternately charged, loosely packed, parallel plates) precipitation can be achieved satisfactorily by applying a lower potential than is necessary in the charging chamber, since a corona discharge is not a requirement.

The corona discharge is usually produced by making the center wire negative, because precipitation efficiency is higher under these conditions. However, less ozone is produced by reversing the polarity, and a positive wire is commonly employed in the cleaning of air when the presence of ozone may be objectionable. The high-voltage direct current is commonly produced by mercury vapor or vacuum tube rectifiers. The power requirements vary from 2 to 5 kWh per 1,000,000 ft^3 (28,300 m^3) of gas being treated, depending upon the amount and nature of the particles being removed. Considerable difficulty is encountered in electrical leakage across the insulators because drops of liquid or solid particles are deposited on them.

H. J. White in 1951 stated that suspended particles become charged by two different mechanisms. The bombardment mechanism results from direct collisions between the suspended particles and ions produced in the corona discharge. The diffusion mechanism consists of the attachment of ions to the suspended particles by ion diffusion. Although both mechanisms operate simultaneously, the former is more important. However, the latter effect may become predominant for smaller particles, perhaps in the 0.1–0.2-micrometer range.

Application. The use of electrostatic precipitators has become common in numerous industrial applications. Each installation, however, is a separate design problem. There is no comprehensive theory applicable in every case for the separation of particles or droplets from a moving gas stream which is usually in turbulent flow. Consequently, empirical methods are generally used for designing precipitators. In fact, it is sometimes customary to make an accurate model of a proposed precipitator installation and to adjust and correct the model performance in the laboratory to meet specifications prior to the manufacture of a scaled-up unit.

In view of the lack of a firm theoretical basis the efficiency of a given unit must be stated carefully in reference to the particular unit, with respect to operating voltage, design geometry, flow rate, and particle removal. The efficiency is usually expressed as the weight percentage of the material removed from the input stream. Such an expression inadequately represents the effect of particle size. It is usually desirable to remove the smallest particles from effluent gases, since their higher degree of opacity makes their discharge more visible and therefore more objectionable to the public eye; units are designed accordingly. The large particles are therefore those which escape; however, the reverse condition may also be achieved by suitable design. The efficiency is thus greatly influenced by the particle or droplet size distribution in the stream, and an efficiency percentage based on weight must be referred to the size distribution, most correctly, before and after precipitation. The efficiency depends exponentially on the stream residence time in the precipitator, so that slower flow rates or longer vessels give better precipitation efficiencies.

Among the advantages of the electrostatic precipitator are its ability to handle large volumes of gas, at elevated temperatures if necessary, with a reasonably small pressure drop, and the removal of particles in the micrometer range. Some of the usual applications are (1) removal of dirt from flue gases in steam plants; (2) cleaning of air to remove fungi and bacteria in establishments producing antibiotics and other drugs, and in operating rooms; (3) cleaning of air in ventilation and air conditioning systems; (4) removal of oil mists in machine shops and acid mists in chemical process plants; (5) cleaning of blast furnace gases; (6) recovery of valuable materials such as oxides of copper, lead, and tin; and (7) separation of rutile from zirconium sand. *See* DUST AND MIST COLLECTION.

George S. Mill; W. O. Milligan

Electrostatics

The class of phenomena recognized by the presence of electrical charges, either stationary or moving, and the interaction of these charges, this interaction being solely by reason of the charges and their positions and not by reason of their motion. *See* ELECTRIC CHARGE.

At least 90% of the topics that are normally classified as electrostatics are concerned with the manipulation of charged particles by electric fields. When a particle becomes charged by rubbing or other means, it has either a surplus or a deficit of electrons. A body with a surplus of electrons is said to be negatively charged; a body with a deficiency,

positively charged. The amount or quantity of charge on a body is expressed in coulombs (positive or negative). A coulomb is an enormous amount of charge, and in most electrostatic situations charge levels of a small fraction of a coulomb give rise to significant effects. Electrostatic forces always exist between charged bodies. Bodies with like charge experience repulsive forces, while oppositely charged bodies experience attraction.

Coulomb's law. If two bodies are charged to Q_1 and Q_2 coulombs and are separated in vacuum by a distance of r meters, the force F in newtons between them is given by Eq. (1).

$$F = \frac{Q_1 Q_2}{4\pi\epsilon_0 r^2} \tag{1}$$

During the eighteenth century, Henry Cavendish carried out an ingenious experiment with two concentric metallized spheres which were electrically connected. He showed that after charging the outer sphere no charge was detectable on the inner sphere. He concluded that his observations proved that the force between charged bodies depended inversely on the square of the distance separating them. *See* ELECTRICAL SHIELDING.

Experimental work on the direct measurement of force between charged bodies was carried out by C. A. Coulomb several years later, and the law which expresses mathematically the force between charged bodies is known as Coulomb's law. In electrical science, ϵ_0 is an important constant known as the permittivity or dielectric constant of free space, and is also sometimes called the primary electric constant. It has the value $\epsilon_0 = 8.85416 \times 10^{-12}$ farad per meter. *See* COULOMB'S LAW; PERMITTIVITY.

Electric fields. Coulomb's law shows that a body charged to Q_1 experiences a force due to the presence of another body charged to Q_2. Q_2 may be considered to influence the whole of space surrounding it, because if Q_1 were to be positioned anywhere it would experience a force due to the presence of Q_2. The property of a charge to influence the whole of space can be modeled by a three-dimensional force field permeating the whole of the space surrounding the charge Q_2. This field is called the electric field, and is often symbolized E. It is a vector quantity having both magnitude and direction. From Coulomb's law, the force acting on Q_1 can be written as in Eq. (2).

$$F = Q_1 E \tag{2}$$

From Eq. (2) it follows that the electric field E may be defined as the coulombic force in newtons per coulomb of charge. The more familiar units of electric field are volts per meter. It again follows that the field E due to charge Q_2 is given by Eq. (3). When

$$E = \frac{Q_2}{4\pi\epsilon_0 r^2} \tag{3}$$

there are many charged bodies present in an environment, the force that would be exerted on a charged particle at any location can be found by calculating the field at the location due to the presence of each charged body separately, and the net field is obtained by adding up the individual components. An entire region may be considered in this way so that the electrostatic force acting on a particle at any position is readily determined. Hence the influence of charged bodies upon the trajectories of charged particles can be calculated. In a cathode-ray tube, for example, various fixed bodies in the form of deflection and focusing electrodes are appropriately charged to set up an electrostatic field which controls the trajectories of electrons. *See* CATHODE-RAY TUBE; ELECTRIC FIELD.

Potential energy. A system of charged particles or bodies is unstable unless the particles are prevented from moving, since the like-charged particles will repel each other until they are infinitely far apart, and oppositely charged bodies will attract one another and come together. The system has potential energy. The potential energy of two charged particles separated by a distance r can be shown to be given by Eq. (4). *See* ENERGY; POTENTIALS.

$$\text{Potential energy} = \frac{Q_1 Q_2}{4\pi\epsilon_0 r} \quad \text{joules} \tag{4}$$

Charging methods. The three principal methods of applying electric charge to objects are corona charging, induction charging, and tribocharging. *See* ENERGY.

Corona charging. The corona-charging method, which relies upon the impact of charged atoms or molecules (ions) on charged bodies, is particularly important in a number of technological applications. Copious quantities of ions may be generated by a corona discharge, which is a region in which an intense electric field acts upon air molecules and ionizes them so that free ions are produced. A sharply pointed electrode maintained at a high positive or negative potential induces a stream of positive or negative ions which may be used for charging surfaces. The stream of ions from a corona point is usually so intense that neutral air molecules become entrained in the flow to produce a corona wind which can be felt on the back of a hand or can deflect a candle flame.

Ions from a corona discharge may be used to charge isolated bodies, insulating surfaces or particles by simply directing a corona wind onto the surface to be charged. In the case of particles, it is normally sufficient for them to pass through a corona discharge region to receive a significant charge from ion-particle collisions. *See* CORONA DISCHARGE.

Induction charging. Surfaces may be charged by exposure to an electrostatic field. If the surface is a liquid and it is disrupted into droplets, they will be electrically charged. An example of induction charging of a liquid is illustrated in **Fig. 1**. With no voltage applied to the capillary tube, the liquid drips from the end of the tube. The drops are not significantly charged. When a voltage of a few hundred volts is applied to the capillary tube, an electric field is set up in the drop formation region. As each drop forms, the electric field induces surface charge, and an

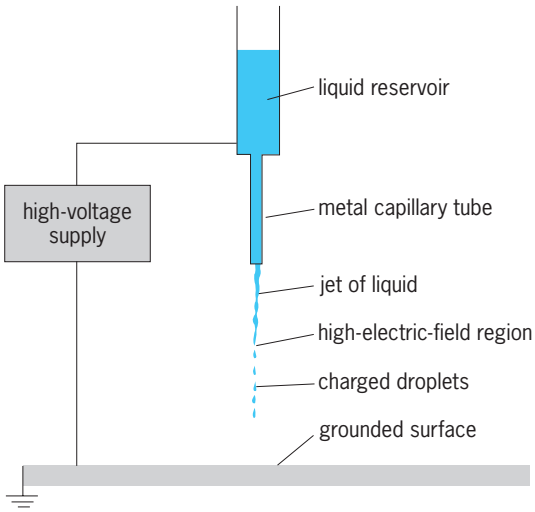

Fig. 1. Induction charging of a liquid.

electrostatic downward force acts on each drop in addition to the gravitational force. As a result, drops break off from the capillary tube before they are fully developed and are therefore smaller than normal. Thus the presence of the field increases the drip rate and, most importantly, causes the drops to be electrically charged. If the capillary-tube voltage is raised to a few thousand volts, the dripping changes to a jet or many jets of liquid which may subsequently break up into small charged droplets constituting a spray. *See* ATOMIZATION.

Induction charging of equipment and personnel may occur when they are exposed to an electric field. Personnel charged in this way may generate electrostatic discharges when approaching grounded surfaces. Sensitive microelectronic devices can be damaged and computer data can be corrupted by such discharges.

Applications. Electrostatics is put to good use in a wide variety of applications, many involving solid or liquid particles.

Air cleaners. The electrostatic precipitator enables smoke emissions from power-station chimneys, smelting plants, and other industrial plants to be reduced to relatively low, acceptable levels. On a smaller scale, efficient filters exist for removing dust from the air in offices, public places, and the home. In some filters, dust particles undergo corona charging as they are sucked by a fan through a duct, and are then collected on grounded electrodes; in others, permanently electrified filter material is used, made from thin plastic sheets which have been treated by surface bombardment from a corona ion source. Ions may attach so tenaciously to the surface of the sheets when conditions are carefully controlled that the treated material acquires a permanent electric charge. Such electrified material is known as an electret. An electret is the electrical equivalent of a permanent magnet, and will attract dust particles, fibers, and so forth, in the same way as a magnet attracts iron filings. *See* AIR FILTER; ELECTRET; ELECTROSTATIC PRECIPITATOR.

Spraying. In several applications which utilize electrostatics, solid or liquid particles are charged and sprayed onto grounded objects. Dry powder coating is used in many industries in preference to the wet-paint-spraying process. Crop spraying is another important application in which electrostatic forces help to efficiently apply herbicides or insecticides (**Fig. 2**). For the dry-powder-coating process, pigmented powder particles, which may incorporate resin, modifiers, and possibly a curing agent, are dispersed pneumatically and charged using a corona point at the end of a spray gun. Liquids are dispersed hydraulically, pneumatically, or by electrostatic forces alone, and droplets are charged by induction. As the sprayer nozzle is operated at several thousand volts, an intense electric field is set up between the nozzle and the grounded object that is to be sprayed. Charged particles tend to follow the field lines and are even guided to deposit on the far side of targets. This wrap-around effect minimizes

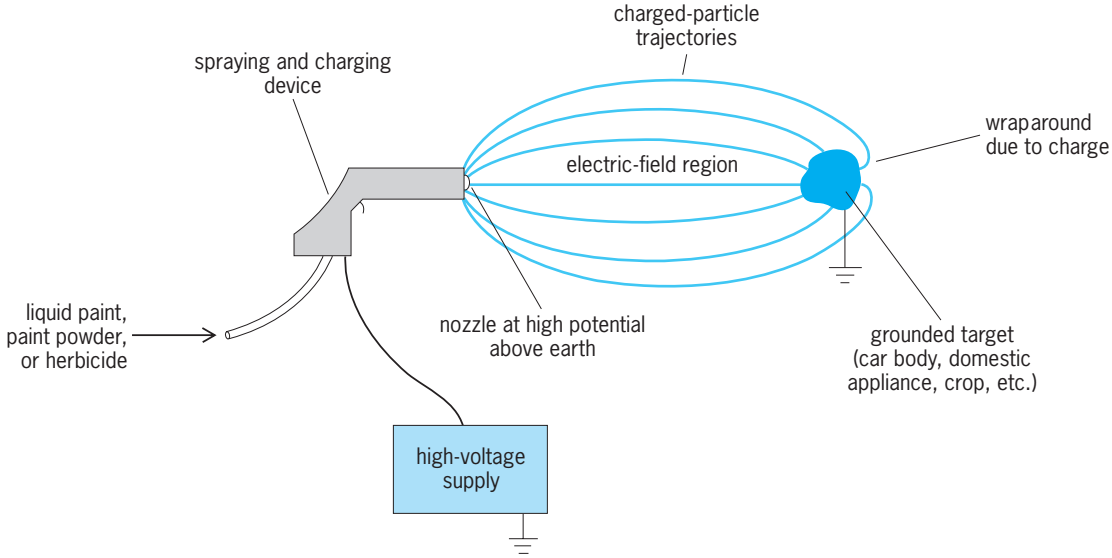

Fig. 2. Electrostatic spraying of powders or droplets.

target overspray, and is a particularly important feature of electrostatic spraying. In the powder-coating process, the powder adheres to the grounded workpiece by electrostatic forces known as image forces. A charged particle close to a conducting surface is attracted by an equal but opposite charge induced as a mirror image on the surface. Coated objects are baked or treated with infrared radiation, causing thermoplastics to melt and run to form a continuous film. In thermosetting powders, so-called cross-linking occurs, and once the coating is cured, it will not remelt. In the liquid-paint- and crop-spraying processes, centrifugal forces are sometimes used in conjunction with electrostatic fields, producing sprays consisting of droplets that are fairly uniform in size. This enables deposition to be more accurately controlled.

Research into electrospraying, sometimes called electrohydrodynamic atomization, is leading to new applications for the deposition of ceramic, glass, and polymer films and for powder particle production of special materials. The electrospraying of materials is also used for analysis by means of mass spectrometry, as the electrospray process is gentle and does not disrupt delicate complex molecules.

By using induction charging, it is possible to disperse fuel oils into charged sprays or mists. Charged electrodes allow some measure of control over the combustion of such mists; that is, it is possible to produce electrically controlled flames.

Electrophotography. In electrophotography an optical system is used to project the image to be copied onto a light-sensitive semiconducting surface precharged by a corona source. Exposure of the surface to light reduces the electrical conductivity of the material and allows surface charge to leak away to a back plate in proportion to the intensity of the light, so that bright parts of the image are regions that have lost most of the original charge while dark zones remain fully charged. Gray shades produce intermediate levels of charge. A mixture of very fine black toner particles (approximately 2 micrometers in diameter) and coarser carrier particles (approximately 100 μm in diameter) is then brought into contact with the charged surface. Transfer of only charged toner particles onto the latent charged surface occurs. The toner charges as it separates from the carrier particles. The amount of toner that attaches by electrostatic forces to each zone of the surface depends directly on the surface charge density. Dark regions of the image are highly charged and thus attract a relatively thick layer of toner. A sheet of paper is then laid over the toner-covered surface, and transfer of toner to paper occurs so that an image remains on the paper when it is peeled off the surface. The paper is then treated to fix the image. Electrophotography is technologically so advanced that color copying with excellent quality is possible. *See* PHOTOCOPYING PROCESSES.

Ink-jet printers. Modern ink-jet printers are of various types, not all of which utilize electrostatic principles. One type uses hydraulic pressure via arrays of acoustically excited nozzles which produce fine jets of ink, usually by hydraulic pressure. The jets travel a few millimeters and then break up into drops. By ensuring that drops are formed in the presence of an electrostatic field, they become charged and may be deflected electrostatically to a printing surface. Ink-jet printers are highly developed machines that print rapidly and silently. Colored inks may be used not only for alphanumeric printing but for high-resolution printing of pictures which approach photographic quality. *See* INK; PRINTING.

Metallic ion and droplet beams. Another development being commercially exploited is the production of metallic ion or droplet beams using electrostatic forces acting upon a liquid-metal surface. Considerable success has been achieved with many molten metals including gold and silver. Either ion or charged droplet beams may be formed depending on the operating conditions of the source. The beams so formed may be very well defined and directed with great accuracy onto targets where they can be used for ion implantation or for the formation of conducting tracks in the fabrication of microelectronic circuits. *See* INTEGRATED CIRCUITS; ION IMPLANTATION.

Electrostatic coalescers. Crude oil from wells, especially offshore, is invariably mixed with water or brine, with water content sometimes in excess of 80% of the volume. The extraction process subjects the mixture to considerable shearing, and stable oil-water emulsions form, which do not readily separate out. An oil continuous phase containing water drops will support an intense electric field which polarizes the individual drops. Each drop essentially becomes an electrostatic magnet with positively and negatively charged ends, aligned in the electric field. By virtue of this polarization, drops in proximity are attracted toward each other and coalesce to form larger drops. In this way a population of fine drops within an oil can be induced to coalesce rapidly to form much larger drops, which are then able to settle out under gravity from the oil. Electrostatic treators using electric fields not only to coalesce drops but also to move and deposit inorganic particles of sand, mud, and clay and organic particles have been developed.

Space exploration. Ion engines which produce thrust by electrostatically accelerating mercury or cesium ions have been successfully operated in space. Colloid thrusters, operating on exactly the same principles as electrostatic paint or crop sprayers, have also been developed. In these a propellant such as glycerol is atomized and accelerated from a nozzle by an electrostatic field. *See* ION PROPULSION. A. G. Bailey

Bibliography. A. G. Bailey, *Electrostatic Spraying of Liquids*, 1988; A. G. Bailey, The science and technology of electrostatic powder spraying, *J. Electrostat.*, 45:85–120, 1998; I. Berta, *Industrial Electrostatics*, 1994; J. S. Chang, A. J. Kelley, and J. M. Crowley (eds.), *Handbook of Electrostatic Processes*, 1995; J. A. Cross, *Electrostatics*: *Principles, Problems and Applications*, 1987; J. M. Crowley, *Fundamentals of Applied Electrostatics*, 1986, reprint 1991; Special issue on electrohydrodynamic atomization, *J. Aerosol Sci.*, 30(7):823–978, 1999.

Electrostriction

A form of elastic deformation of a dielectric induced by an electric field; specifically, the term applies to those components of strain which are independent of reversal of the field direction. Electrostriction is a property of all dielectrics and is thus distinguished from the converse piezoelectric effect, a field-induced strain which changes sign upon field reversal and which occurs only in piezoelectric materials. *See* DIELECTRIC MATERIALS; PIEZOELECTRICITY.

Electrostrictive strain is approximately proportional to the electric susceptibility, elastic compliance, and the square of the field strength, and is extremely small for most materials.

The electrostrictive effect in certain ceramics is employed for commercial purposes in electromechanical transducers for sonic and ultrasonic applications. *See* MICROPHONE. Robert D. Waldron

Electrotherapy

The use of electric current to treat disease. Electrotherapy is based on principles developed during the nineteenth and twentieth centuries following the first demonstration of "animal electricity" by Luigi Galvani in the eighteenth century. This article covers low-frequency electrotherapy, diathermy and hyperthermia, and electroconvulsive therapy.

Low-Frequency Electrotherapy

Low-frequency electrotherapy uses relatively weak alternating electric currents. They are delivered by electrodes that are placed under or on the surface of the body and are connected to pulse-shaping circuits that are located inside or outside the body.

Excitable tissues. Electrodes that stimulate electrically excitable cells, such as those in muscle and nerve tissues, are usually placed directly in or on the tissue by surgery or are inserted through a vein by catheterization. There are many applications for electrode stimulation: irregular heart rhythms can be controlled by pacemakers; muscles, such as those of the diaphragm and urinary bladder, that become paralyzed can be made to contract electrically; and long-term pain can be relieved by implanting electrodes in the spinal canal. When coupled to appropriate braces, sensors, and programmed computers, electrodes in contact with the muscle groups of a lower extremity have been shown to help persons with spinal cord injury become ambulatory. Surface electrodes are widely used for temporary relief of pain, a technique known as transcutaneous electrical nerve stimulation; for preventing muscle atrophy after injury or immobilization; and for treatment of curvature of the spine, or scoliosis. *See* CARDIAC ELECTROPHYSIOLOGY; MUSCLE; PAIN.

Nonexcitable tissues. A wide variety—if not all—of the body's nonexcitable cells alter their function in specific ways and at specific times in response to appropriate, very weak electrical stimuli. (That observation reemphasizes the central roles of physics in all living processes and the interaction of electric charges as the basis of the chemical aspects of those processes.) Much of the progress in electrotherapy has evolved from the observation that both hard and soft tissues, such as bone and arteries, become electrically charged when they are cyclically deformed by mechanical or hydrodynamic forces. Weak voltages, in the range of microvolts to millivolts per centimeter, occur because both cells and tissues contain piezoelectric molecules that respond to deforming forces by becoming electrically polarized. Furthermore, electrically charged entities such as ions, cells, and molecules are transported by hydrodynamic forces past sites of structurally fixed electric charge and, in the process, create a voltage. That phenomenon is an example of an electrokinetic event. By establishing the patterns and values of those stress-generated voltages in bone, researchers have been able to develop three methods of influencing the behavior of nonexcitable cells that are involved in the repair of nonunited fractures. The oldest method was an invasive procedure, and was followed by two noninvasive procedures. *See* BONE.

Electrode implantation. The first clinical method for treating nonunited fractures employed surgically implanted electrodes. Once placed at the fracture site, they delivered constant, direct currents similar in amplitude to those that occur naturally in bone after fracture, known as injury potentials, or as a response to mechanical deformation. Unfortunately, surgical methods carry a risk of infection, and direct currents can result in adverse electrochemical reactions around the electrodes.

Pulsed electromagnetic fields (PEMFs). Two noninvasive electrical methods have proved effective in altering cellular behavior. The first involves the placement of dynamically charged, insulated plates outside tissue-culture vessels or the fractured limbs of animals. Broad-scale application has not yet proven practical in humans, however. The second method uses one or more coils of wire coupled to a pulse generator to create a weak time-varying magnetic field that penetrates the body, much as radio waves enter a closed building. The field characteristics are designed to induce pulsing electric currents in the tissue, with waveforms, frequencies, and amplitudes similar to those produced normally by skeletal tissues during high-impact exercise. The waveforms of pulsed electromagnetic fields are quite asymmetrical and contain a broad range of frequencies, which are characteristics that distinguish them from power-line, radio-frequency, and microwave fields.

Depending upon the energy patterns generated by the magnetic fields in tissue, the function of cells involved in abnormal processes can be changed without producing heat. In nonunited fractures, the normal repair has been interrupted at an intermediate stage of healing, and soft, rather than hard, tissue forms a bridge between the bone fragments. The final stages of bony repair are not achieved until the soft tissues undergo calcification. Certain types of pulsed electromagnetic fields can initiate calcification, aid in the ingrowth of new blood vessels, and

increase bone formation, each of which is important in restarting the healing process and ultimately achieving successful union. Pulsed electromagnetic fields have been widely used to treat nonunited fractures, many of which had failed to heal after one or more operations and the person faced amputation of the affected part. In contrast to other electromagnetic fields, the pulse types do not carry any known risks, and hospitalization and a surgical procedure are usually unnecessary, since pulsed electromagnetic fields are applied in the doctor's office. The method, therefore, is substantially less costly than most other treatments available.

As the understanding of their mechanisms of action at the cellular and subcellular levels has increased, pulsed electromagnetic fields have been used successfully to treat other problems of bone and its surrounding soft tissue. For example, when avascular necrosis of the hip in young adults results in bone tissue death, revitalization has been achieved and hip function restored following pulsed electromagnetic field therapy. In older persons, shoulder pain from chronic inflammation of tendons that has proved resistant to classical forms of treatment has responded to pulsed electromagnetic fields. Therapeutic requirements for selective cellular effects appear to parallel those involved in the treatment of disease states with drugs. Future success will lie in encoding the appropriate physical (as opposed to chemical or drug) "information" in the pulsed electromagnetic field waveform to produce changes in cell function and thereby control or correct specific pathologic processes. Other medical applications for pulsed electromagnetic field, such as in cardiology, remain to be explored. C. Andrew L. Bassett

Diathermy and Hyperthermia

The therapeutic benefits of heat have been known for centuries, and modern medicine has used technology to provide controlled heat to diseased tissues.

Diathermy. Therapy for afflicted deep tissues that do not respond to conventional methods, such as infrared heating or heating pads, can often be achieved with diathermy. Heating results from the electrical resistance of tissue to high-frequency or microwave currents. Increasing the tissue temperature to a range of 106–113°F (41–45°C) increases the physiologic response and therapeutic benefit, which includes increased extensibility of collagen tissues in joint contractures, decreased joint stiffness in rheumatoid arthritis, and pain relief and reduction of muscle spasms through the local action of heat on nerve endings and receptors. Warming can also resolve inflammatory infiltrates, edema, and exudates and increase blood flow in diseased or damaged tissue. Heating has been used in cancer therapy under proper temperature control.

Various instruments have been developed to accommodate placement of the heating element on the body. They include (1) coupling of 13.56- or 27.12-MHz shortwave currents from electric fields by way of insulated or noncontacting capacitor plates or specially shaped electrodes for placement in natural body cavities; (2) induction of 13.56- or 27.12-MHz electric currents by magnetic fields from solenoidal coils that enclose the afflicted body member or from flat coils that are placed near the surface of the afflicted tissues; and (3) transmission of 433-, 915-, or 2450-MHz microwave energy into the tissues by radiation from dipole or cavity antennae. The shortwave magnetic and the 433- or 915-MHz microwave applicators are superior to the shortwave capacitive and 2450-MHz microwave applicators for treating tissues beneath subcutaneous fat layers, because they minimize the undesirable selective heating of the fat and the variation in heating levels between patients. Diathermy is contraindicated for areas that are anesthetized or noninnervated, for areas with inadequate vascular supply, in the presence of acute inflammation accompanied by the formation of pus and fever, whenever there is a possibility of hemorrhage, and with malignancies, when there is inadequate monitoring of tissue temperature.

Hyperthermia. Hyperthermia is an experimental method of treating malignant tumors that uses heat alone, heat in combination with ionizing (x-ray) radiation, or heat with chemotherapy. One form of heating involves the application of radio-frequency energy by using methods similar to—but more sophisticated and more carefully controlled than—those in diathermy treatment. The effective temperature range of hyperthermia is very narrow, 108–111°F (42.5–44°C); the benefits are minimal at lower temperatures, and damage to normal cells is probable at higher temperatures. Several mechanisms are thought to account for the selective destruction of tumor cells: (1) selective heating caused by the lower rate of blood flow in tumor tissues; (2) greater sensitivity of tumor cells to heat due to their hypoxic, acidic, and poor nutritional state; (3) synergism of ionizing radiation and hyperthermia due to thermal killing of cells that are hypoxic and are at those critical stages of growth when they are most resistant to radiation; and (4) increase in cell membrane permeability and sensitivity to chemotherapeutic drugs.

Exposures at frequencies other than those authorized for diathermy are performed in shielded rooms to prevent interference with radio communication facilities. The selective heating of tumor tissues uses the same basic techniques as those for diathermy: shortwave capacitive, shortwave inductive, and microwave applicators. However, because of the potentially life-threatening nature of malignant disease, more drastic measures can be justified, including the application of low-frequency currents by surface-contact and invasive electrodes and the implantation of magnetic seeds that can be heated by low-frequency magnetic fields. These approaches enhance the localization of tumor heating and minimize the destruction of normal tissues.

The frequencies used exceed 100 kHz to prevent excitation of nerves and muscles. Heating may be further localized through use of additional frequencies and more sophisticated applicators, such as phased-antenna arrays and cylindrical antennae designed for positioning in natural body cavities. Extensive

multipoint temperature-measuring equipment with automatic feedback power-control systems is necessary to maintain a sufficiently high temperature in the tumor without destructive temperature changes in normal tissues. *See* THERMOTHERAPY. Arthur W. Guy

Electroconvulsive Therapy

Electroconvulsive therapy is a procedure for treating severe psychiatric disorders. It is usually given as a series of treatments, typically numbering between 6 and 12, over a period of a few weeks. Each 10-min treatment is preceded by the administration of an anesthetic agent, to render the patient unconscious throughout the procedure, and a drug that blocks muscle movement in the body, to confine the provoked seizure to the brain and prevent convulsive movements in the body. Electrodes are then placed on the scalp and a small jolt of electricity is applied. The electric current provokes a generalized brain seizure, much like that experienced spontaneously by patients with grand mal epilepsy. The patient's brain waves are usually monitored with electroencephalography during the seizures, which typically last about 1 min.

In use since about 1940, electroconvulsive therapy has been shown to be effective in treating a specific set of psychiatric disorders, the most common of which is depressive illness. Depressed patients often undergo electroconvulsive therapy after failing to respond to other forms of therapy, particularly antidepressant medications. For others, electroconvulsive therapy is preferred because the medical risks associated with drug treatment are too great: in some elderly patients, the frequency of medical injury and death from antidepressants is believed to exceed that with electroconvulsive therapy. The fact that electroconvulsive therapy produces rapid improvement and has a high probability of success is another reason for its use. In patients at risk for death from suicide or medical conditions associated with depressive illness, electroconvulsive therapy may be preferred over other therapeutic methods. It has also been shown to be an effective treatment for patients with acute mania and schizophrenia. Less common uses are the treatment of a small set of medical and neurological disorders that are resistant to conventional therapies, such as Parkinson's disease, intractable epilepsy, and psychosis associated with toxic states. *See* AFFECTIVE DISORDERS; PARKINSON'S DISEASE; SCHIZOPHRENIA.

Electroconvulsive therapy has a characteristic set of side effects. Following each treatment, patients experience a period of confusion that may last from several minutes to several hours, and immediately following the complete course of treatments, they frequently have memory difficulties. Memory problems include difficulty in remembering newly learned information and in recalling events from the recent past. At the same time, other aspects of intellectual functioning, such as the ability to perform a task and solve a problem, are unchanged or show improvement over that displayed before electroconvulsive therapy. The improvements can be attributed to the beneficial effects of electroconvulsive therapy on the psychiatric conditions that caused the diminished intellectual performance. Objective evidence of persistent difficulties in learning or recalling new information is difficult to obtain several weeks after treatment, but some patients may experience permanent gaps in their memory for events that occurred in the weeks prior to, during, and following the treatment course. The beneficial and adverse effects of electroconvulsive therapy can likely be traced to changes in brain physiology and biochemistry. Studies in humans and animals indicate it is unlikely that electroconvulsive therapy causes structural damage to the brain, but the precise ways in which it produces beneficial and adverse effects are unknown. *See* ELECTROCONVULSIVE THERAPY. Harold A. Sackheim

Bibliography. R. Abrams, *Electroconvulsive Therapy*, 3d ed., 1997; A. Chiabrera, C. Nicolini, and H. P. Schwan (eds.), *Interaction Between Electromagnetic Fields and Cells*, 1985; T. Fried, R. Johnson, and W. McCracken, *Archives of Physical Medicine and Rehabilitation*, 65:228–231, 1984; P. L. Ludmer and N. Goldschlager, *N. Engl. J. Med.*, 311:1671, 1984; C. Polk and E. Postow, *Handbook of Biological Effects of Electromagnetic Fields*, 2d ed., 1996.

Electrothermal propulsion

Vehicular propulsion that involves electrical heating to raise the energy level of the propellant. In contrast, chemical rockets use the chemical energy of one or more propellants to heat and accelerate the decomposition products (monopropellants) or combustion products (bipropellants) for thrusting purposes. In both instances, the high-energy propellant gases are exhausted through a nozzle where they are accelerated to a high velocity (U_e), and thrust is produced by reaction. By decoupling the heating or energy addition process from the restraints of propellant chemistry considerations, electrothermal devices can be operated on a wide variety of materials, many of which would not otherwise be considered to be propellants. Water and space station liquid-waste streams are two examples of such propellants being considered for electrothermal propulsion purposes. *See* ROCKET PROPULSION.

Practical electrothermal thrusters come in two forms, resistojets and arcjets. In resistojets the electrical energy is first deposited in a heater or resistive element and then transferred to the propellant. The need to first heat a material limits the maximum operating temperature and the maximum enthalpy of the propellant. As a consequence, the essential simplicity of the device is balanced by well-defined limitations on exhaust velocity or specific impulse I_s. (Specific impulse is a measure of propellant utilization efficiency. It is defined as the thrust per unit weight of propellant expended to produce the thrust. Numerically, $I_s = U_e/g$, where g is the gravitational acceleration at the Earth's surface, 9.8 m/s^2 or 32.2 ft/s^2.) Arcjets circumvent this limitation by using the propellant as the heater element. High temperatures and

specific-impulse values can be achieved but only at the price of design complexity. *See* SPECIFIC IMPULSE.

One further and fundamental distinction is that electrical thrusters are power-limited whereas chemical thrusters are energy-limited. This distinction profoundly influences the form and applications of these two classes of propulsion systems. By definition, electrical propulsion systems must have an associated power supply for operation. Solar panels or nuclear power supplies can supply well-defined power to the thruster for essentially unlimited time, which is limited by the availability of propellant for propulsion. Consequently, although the power level is well constrained, the total energy available is virtually boundless. Chemical systems are exactly opposite. In this case, the total energy available for propulsion is well defined by the propellant volume. The propellant tanks (liquid) or casing (solid) can hold only a certain amount of propellant. However, the rate at which this propellant is used, the rate of energy usage per unit time, or the power can be exceedingly high. All of the space shuttle's propellant is exhausted in about 10 min, whereas electrical thrusters operate for months or years. In other words, electrical propulsion devices are characterized by low thrust (power), high exhaust velocity, and parsimonious use of propellant (their primary feature of interest), and chemical rockets are high-thrust systems that are profligate in the use of propellant. Chemical thrusters are ideal for escaping the Earth's gravity well; electrical thrusters are ideal for moving payloads in low-acceleration conditions removed from gravity wells, that is, for the more ambitious missions far removed from low Earth orbit.

Resistojet thrusters. These have been under development for many years and are now flying in a stationkeeping role on many communications satellites. An electric heating coil is used to add energy to the propellant (**Fig. 1**). The maximum temperature, and therefore the thruster performance limit, is set by the maximum temperature of the coil material and correspondings structural requirements. The heat of the coil is coupled either radiatively or directly to the propellant gas. In the radiatively coupled resistojet (Fig. 1*a*), the coil is placed in an evacuated cavity and

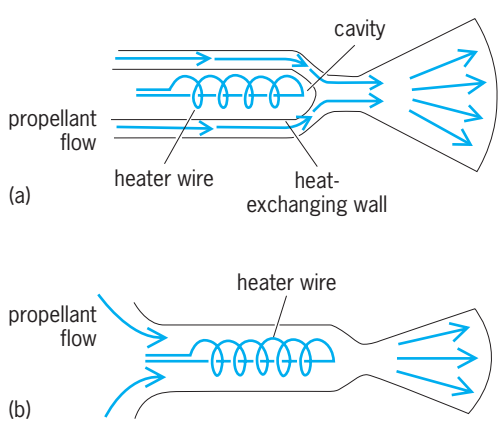

Fig. 1. Resistojet configurations. (*a*) Radiatively coupled. (*b*) Directly coupled.

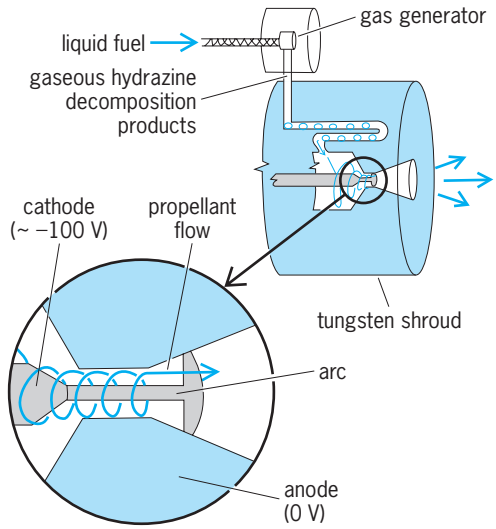

Fig. 2. Hydrazine arcjet thruster, showing enlargement of heat transfer region.

transfers energy by radiation to the inner cavity wall. The outer cavity wall acts as a heat exchanger which, in turn, is in direct contact with the propellant. Evaporation of the heater element is a life-limiting issue in this type of thruster. In a directly coupled resistojet (Fig. 1*b*) the heater element is fully immersed in the propellant and transfers energy principally by conduction and convection. Direct heating is more efficient but exposes the coil to the propellant, which can lead to erosion or corrosion. As in the radiatively coupled thruster, such processes limit the coil lifetime.

In principle, resistojets can be designed to operate at any power level. Existing devices typically require 300 to 800 W of electrical power and produce specific impulses in the range of 250 to 330 s (exhaust velocities from 2.5 to 3.3 km/s or 1.5 to 2.1 mi/s) with thrust levels that are typically on the order of 0.5 newton (0.1 lbf). Many different types of propellants, including hydrazine (N_2H_4), ammonia (NH_3), molecular hydrogen (H_2), and the organic waste gases such as carbon dioxide (CO_2) and methane (CH_4), have been successfully used. The omnivorous nature of the resistojet has led to its consideration as reboost propulsion for the NASA crewed space station using waste gases as propellants; and resistojets using water (H_2O) as a propellant have been selected for the reboost propulsion of the industrial space facility.

Arcjet thrusters. These are similar in concept to resistojets, but rather than transferring energy to the gas through a resistive coil, the arcjet utilizes an electric arc discharge which passes directly through the propellant. While this avoids the materials limitations imposed by the heater coils and the heat exchanger walls of the resistojet, a small portion of the gas must be ionized to carry the current. The interaction of this plasma with the electrodes is an important design consideration that has to be properly addressed to assure adequate longevity. *See* ARC DISCHARGE.

A schematic of an arcjet thruster is shown in **Fig. 2**. Low-speed propellant enters the arcjet plenum in a

tangential fashion to produce a vortical motion to assist plasma formation on the centerline. The gas accelerates around the cathode and through the constrictor and finally expands supersonically to ambient space. A direct-current arc is struck between the cathode and the anode (that is, the nozzle body). The high-temperature, high-velocity plasma which forms the laminar arc column is stabilized as it passes through the constrictor region by both energy (radiation and conduction) and mass transfer (diffusion and convection) to and from the propellant gas which surrounds it. The thruster walls are thus shielded from the hot arc plasma by the propellant. Electrons within the arc core are readily accelerated by the applied electric field and transfer energy via collision to their much more massive neutral and ionic neighbors. This process tends to cool the arc plasma and heat the surrounding propellant flow. Although the temperatures in the low-density central region may be in excess of 20,000 K, the bulk or mass-averaged temperature of the propellant is typically between 5000 and 8000 K, a level that substantially exceeds that of resistojets. As the arc plasma enters the diverging section of the nozzle, it is rapidly cooled and eventually expands to the anode wall. With proper design, high wall-current densities can be avoided so that the process of arc attachment remains is nondestructive.

Electrons produced at the cathode by a combination of thermionic and field emission provide most of the current in the arcjet. An equal number of ions, however, is generated in the anode fall region because of collisions of electrons with neutral gas molecules as they accelerate toward the wall. Thus, except in the cathode and anode fall regions (that is, within a few mean free paths of these surfaces), the arcjet plasma is spatially neutral. Langmuir probe measurements in arcjet plumes indicate that electron temperatures are on the order of 1 eV and number densities are on the order of 10^{13} cm^{-3}. For the typical arcjet flow rates, this represents an electron/ion mole fraction of less than 1%. *See* FIELD EMISSION; PLASMA DIAGNOSTICS; THERMIONIC EMISSION.

Arcjet thrusters can operate over a wide range of power by scaling the physical dimensions of the device. They may be broadly classified by their power requirements. If a nominal cathode voltage of −100 V with respect to the anode is assumed, the definitions given in the **table** apply. Typical thrust levels range from 0.25 N (0.05 lbf) for low-power arcjets to 2.5 N (0.5 lbf) for high-power arcjets. Arcjets are therefore considered as candidates for missions ranging from stationkeeping to orbit transfer. Like the resistojet, an arcjet can operate with many different types of propellant. Because the arcjet is an electrothermal device, operation is relatively insensitive to propellant selection; operation has been demonstrated by using argon (Ar), helium (He), molecular hydrogen (H_2), ammonia (NH_3), and hydrazine (N_2H_4). Depending on the choice of propellant, exhaust velocities can range from 4 to 9 km/s (2.5 to 5.5 mi/s) with corresponding specific impulses between 400 and 900 s. Development of prototype engines has been undertaken, with flights projected in the early 1990s. *See* SPACECRAFT PROPULSION. G. W. Butler

Classification of arcjet thrusters

Range	Power, kW	Current, A	Mass flow rate, g/s
Low	1–2	10–20	0.04
Medium	5–10	50–100	0.12
High	20–30	200–300	0.24

Bibliography. M. N. Hirsh and H. J. Oskam, *Gaseous Electronics*, 1978; R. G. Jahn, *Physics of Electric Propulsion*, 1968; *Proceedings of the AIAA/ASME/SAE/ASEE 24th Joint Propulsion Conference*, Boston, 1988; *Proceedings of the 1986 JANNAF Propulsion Meeting*, New Orleans, 1986; *Proceedings of the 19th International Electric Propulsion Conference*, Colorado Springs, 1987; J. A. Stone, D. C. Byers, and P. Q. King, *The NASA Electric Propulsion Program*, NASA Tech. Mem. 101324, 1988.

Electroweak interaction

One of the three basic forces of nature, along with the strong nuclear interaction and the gravitational interaction. The terms "force" and "interaction between particles" are used interchangeably in this context. All of the known forces, such as atomic, nuclear, chemical, or mechanical forces, are manifestations of one of the three basic interactions.

Until the early 1970s, it was believed that there were four fundamental forces: strong nuclear, electromagnetic, weak nuclear, and gravity. It was by the work of S. Glashow, S. Weinberg, and A. Salam that the electromagnetic and the weak nuclear forces were unified and understood as a single interaction, called the electroweak interaction. This unification was a major step in understanding nature, similar to the achievement of J. C. Maxwell and others a century earlier in unifying the electric forces and magnetic forces into the electromagnetic interactions. A goal of theoretical physics is to achieve a further simplification in understanding nature and describe the presently known three basic interactions in a unified way, usually referred to as the grand unified theory (GUT). Whether this is possible remains to be seen. *See* ELECTROMAGNETISM; FUNDAMENTAL INTERACTIONS; GRAND UNIFICATION THEORIES; GRAVITATION; MAXWELL'S EQUATIONS; STRONG NUCLEAR INTERACTIONS; WEAK NUCLEAR INTERACTIONS.

Properties of basic interactions. Some of the properties of the basic interactions are summarized in **Table 1**. The strong nuclear forces are the strongest, electroweak is intermediate, and gravity the most feeble by a huge factor. The ranges, that is, the distances over which the forces act, also differ greatly. The strong nuclear and the weak interactions have a very short range, while electromagnetism and gravity act over very large distances. Thus, at very short subatomic distances the strong nuclear force, which holds the atomic nucleus together and

TABLE 1. Basic forces in nature

Interaction		Relative strength	Property acted on	Force carrier	Range
Strong nuclear		1	Color charge (r, g, b)	Gluon (g)	10^{-13} cm
Electroweak	Electromagnetic	10^{-2}	Electric charge (q)	Photon (γ)	∞
	Weak nuclear	10^{-6}	Weak charges (t_3, y)	Bosons ($W^{\pm}$, Z^0)	10^{-16} cm
Gravity		10^{-40}	Mass (m)	Graviton (G)	∞

governs many interactions of the subnuclear particles, dominates. At larger distances the electromagnetic forces dominate, and hold the atom together and govern chemical and most mechanical forces in everyday life. At even larger scales, objects such as planets, stars, and galaxies are electrically neutral (have an exact balance of positive and negative electric charges) so that the electromagnetic forces between them are negligible, and thus the gravitational forces dominate in astronomical and cosmological situations.

Each of the basic forces acts on, or depends on, different properties of matter. Gravity acts on mass, and electromagnetic forces act on electric charges that come in two kinds, positive and negative. The strong nuclear forces act on a much less well-known property, called color charge, which come in three kinds, r, b, and g (often called red, blue, and green). The weak nuclear forces act on equally esoteric properties called weak isospin t and hypercharge y. While the mass and the electric charge are properties that are recognized in everyday situations, the color charge and the weak isospin and hypercharge have no correspondence in the large-scale everyday world. *See* COLOR (QUANTUM MECHANICS); ELECTRIC CHARGE; HYPERCHARGE; I-SPIN; MASS.

Fundamental constituents of matter. All known forms of matter are made of molecules and atoms, which are made up of the nucleus (protons and neutrons) and orbital electrons. These in turn can be understood to be made up of the fundamental constituents, the quarks and the leptons. Each of these comes in six kinds (**Table 2**). The six leptons are usually organized into three families:

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}$$

Similarly, the six quarks can be organized into three families:

$$\begin{pmatrix} d \\ u \end{pmatrix} \begin{pmatrix} s \\ c \end{pmatrix} \begin{pmatrix} b \\ t \end{pmatrix}$$

The masses of the leptons and quarks are frequently specified in units of electronvolts (1 MeV = 10^6 eV, and 1 GeV = 10^9 eV). In more usual units, the electron mass of 0.5 MeV is equal to 9×10^{-31} kg. For many years the masses of the three neutrinos (ν_e, ν_μ, and ν_τ) were believed to be zero. However, an effect called neutrino oscillations indicates that they have small but nonzero masses. The values of these masses are not yet known; only experimentally measured upper limits on them are available. *See* NEUTRINO.

All of the quarks and leptons have gravitational and weak interactions since they have nonzero values of mass and weak isospin and hypercharge. The particles with zero electric charge have no electromagnetic interactions, and the leptons have no strong nuclear interactions since they carry no color charge. *See* LEPTON; QUARKS.

Exchange forces and gauge bosons. The present understanding is that the basic forces are not contact forces but act over distances larger than the sizes of the particles (action at a distance). In this picture,

TABLE 2. Leptons and quarks, the fundamental constituents of matter

Particle	Symbol	Mass	Electric charge (q)	Color charge	Weak charges: Isospin (t_3)	Weak charges: Hyper-charge (y)
e neutrino	ν_e	≤12 eV	0	0	+1/2	−1
Electron	e^-	0.51 MeV	−1	0	−1/2	−1
Mu neutrino	ν_μ	≤0.2 MeV	0	0	+1/2	−1
Muon	μ^-	106 MeV	−1	0	−1/2	−1
Tau neutrino	ν_τ	≤18 MeV	0	0	+1/2	−1
Tau	τ^-	1777 MeV	−1	0	−1/2	−1
Down quark	d	3–7 MeV	−1/3	r, b, g	−1/2	1/3
Up quark	u	1.5–3 MeV	+2/3	r, b, g	+1/2	1/3
Strange quark	s	100 MeV	−1/3	r, b, g	−1/2	1/3
Charm quark	c	1.2 GeV	+2/3	r, b, g	+1/2	1/3
Bottom quark	b	4.3 GeV	−1/3	r, b, g	−1/2	1/3
Top quark	t	174 GeV	+2/3	r, b, g	+1/2	1/3

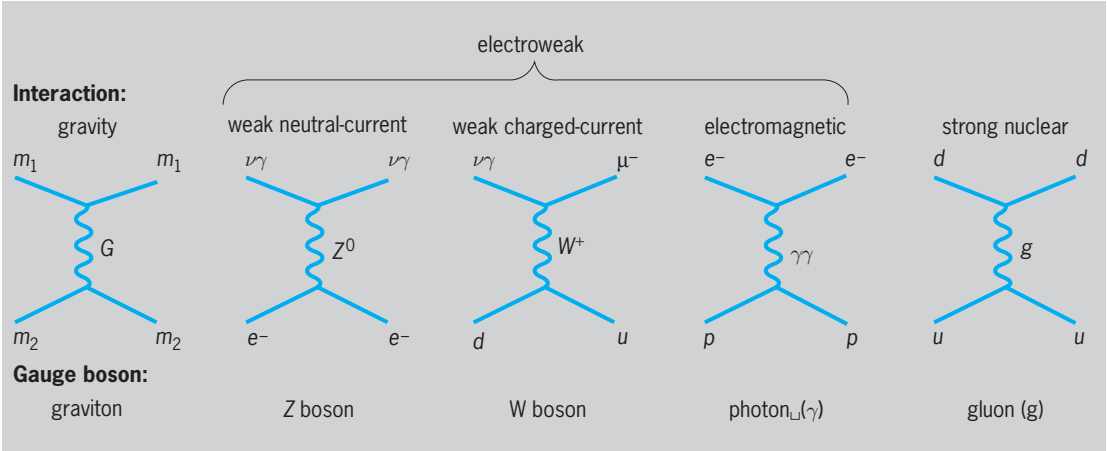

Fig. 1. Basic interactions mediated by gauge bosons.

based on field theory, the forces are carried or mediated by intermediate particles that are called gauge bosons. For example, the electromagnetic force between an electron and a proton is carried by the quantum of the electromagnetic field called the photon (γ). The strong nuclear force is carried by the gluon (g), and the gravitational force is carried by the graviton (G) [**Fig. 1**]. The weak nuclear force comes in two categories: the charge-changing (charged-current, for short) mediated by the $W^{\pm}$ bosons, and the neutral-current weak interactions mediated by the Z^0 boson (Fig. 1). *See* GLUONS; GRAVITON; INTERMEDIATE VECTOR BOSON; PHOTON.

Helicity and parity violation. All of the fundamental constituents, the quarks and the leptons, carry one-half unit of angular momentum (spin = 1/2) as if they were spinning around their own axis. (Such particles are called fermions.) By the rules of quantum mechanics, the direction of this spin is quantized to be either parallel or antiparallel to the direction of motion of the particle. Particles with spin direction parallel to their direction of motion have helicity +1 and are called right-handed, and particles with antiparallel spin have helicity −1 and are called left-handed. *See* ANGULAR MOMENTUM; HELICITY (QUANTUM MECHANICS); SPIN (QUANTUM MECHANICS).

One of the symmetries of nature is called parity, which is a symmetry between right-handed and left-handed coordinate systems. If parity symmetry holds, left-handed and right-handed particles must have the same interactions. In 1956 T. D. Lee and C. N. Yang proposed that parity symmetry is violated in the weak interactions, and this proposal was soon verified experimentally. It was found that the left-handed and the right-handed particles have different weak interactions. The weak isospin and hypercharge assignments given in Table 2 are for left-handed quarks and leptons. The right-handed particles have somewhat different assignments. In particular, the right-handed particles have no weak isospin, and thus only the left-handed particles participate in the charged-current weak interactions. *See* PARITY (QUANTUM MECHANICS).

Electroweak unification. Until the early 1970s, the electromagnetic and the weak interactions were believed to be separate basic interactions. At that time the Weinberg-Salam-Glashow model was proposed to understand these two interactions in a unified way. The model was based on an SU(2) × U(1) gauge symmetry in which the SU(2) part corresponds to a weak isospin triplet of gauge bosons, the W^+, W^0, and W^-, and the U(1) corresponds to a singlet, the B^0. The W's couple to the property called weak isospin, and the B^0 couples to weak hypercharge. *See* GAUGE THEORY.

In its original form, this model, based on an unbroken gauge symmetry, led to some physically unacceptable features such as zero masses for all the constituent particles and predictions of infinities for some measurable quantities. Through the pioneering work of G. 'tHooft, M. Veltman, and others, it was shown that the theory can be made renormalizable, removing the infinities and providing masses to the particles, by spontaneous breaking of the gauge symmetry and the introduction of one new particle, the Higgs boson. *See* HIGGS BOSON; RENORMALIZATION; SYMMETRY BREAKING.

The neutral gauge bosons, the W^0 and B^0, form a quantum-mechanical mixture, which produces the two physically observable gauge bosons, the γ and the Z^0, as given by Eqs. (1). The γ is the well-known

$$\begin{aligned} \gamma &= \sin\theta\, W^0 + \cos\theta\, B^0 \\ Z^0 &= \cos\theta\, W^0 - \sin\theta\, B^0 \end{aligned} \tag{1}$$

photon that mediates the electromagnetic interactions. The Z^0 mediates the neutral-current weak interactions, and the $W^{\pm}$ mediate the charged-current weak interactions (Fig. 1). In this way, all of these interactions are described by a common unified theory. The mixing angle θ in Eqs. (1), forming the γ and the Z^0, is called the weak mixing angle and is the fundamental parameter of the theory. The strength and nature of the interactions of the particles are determined by the vector and axial vector coupling constants g_v and g_A. In the electroweak model all of

these couplings are given in terms of the single parameter of the theory, the weak mixing angle, and the properties of the leptons and quarks given in Table 2. The model also gives a relationship between the electric charge q and the weak charges t_3 and y, Eq. (2).

$$q = t_3 + {}^1\!/_2\, y \tag{2}$$

See NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

The coupling constants that govern the electroweak interactions of all of the particles can be summarized as:

1. Electromagnetic interactions: $g_v = q$, $g_A = 0$
2. Charged current weak interactions: $g_v = -g_A = t$
3. Neutral current weak interactions: $g_v = t_3 - 2q \sin^2\theta$, $g_A = -t_3$

In the above expressions, t stands for the magnitude of the weak isospin, and t_3 is its projection along an axis of quantization.

The electroweak theory has great predictive power. Its first and most striking prediction was the existence of neutral-current weak interactions mediated by the Z^0 boson. Until the time of this prediction, the weak interactions were believed to be of charged-current nature only, with no neutral-current component. Some examples of neutral-current processes predicted by the theory were reactions (3).

$$\begin{aligned} \nu_\mu + e^- &\rightarrow \nu_\mu + e^- \\ \nu_\mu + p &\rightarrow \nu_\mu + p \\ \nu_\mu + p &\rightarrow \nu_\mu + p + \pi^0 \\ \nu_\mu + p &\rightarrow \nu_\mu + p + \text{other particles} \end{aligned} \tag{3}$$

See NEUTRAL CURRENTS.

The experimental discovery of such neutral-current weak interaction processes was considered very strong support for the validity of the theory. From the experimental measurements of the cross sections (that is, interaction probabilities) of such processes, an early estimate of the weak mixing angle was derived to be $\sin^2\theta \approx 0.23$. With this value of $\sin^2\theta$, the theory was able to predict the masses of the hypothetical $W^\pm$ and Z^0 gauge bosons to be approximately $M_W \approx 81$ GeV and $M_Z \approx 92$ GeV.

A second major triumph for the electroweak theory was the discovery of the W and Z bosons in 1983 at the proton-antiproton collider at the CERN Laboratory in Geneva, Switzerland, with masses very close to the values predicted by the theory. At this time the validity of the theory was considered to be firmly established. *See* PARTICLE ACCELERATOR.

Precision tests. A great deal of experimentation followed the discovery of the W and Z bosons to carry out detailed precision tests of the electroweak theory. These were done in a wide variety of contexts such as parity-violating effects in atomic physics, neutrino interactions, polarized electron scattering,

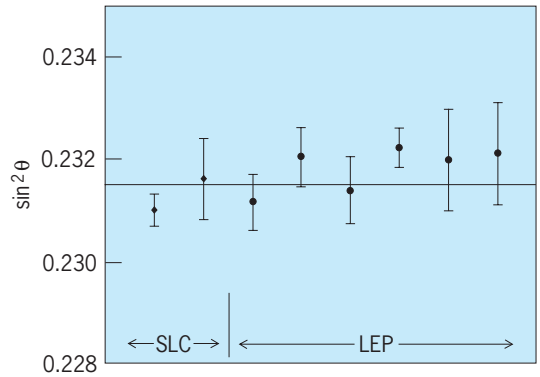

Fig. 2. Results of measurements of the weak mixing angle, θ, at the SLC (Stanford Linear Collider) and the LEP (Large Electron-Positron) Collider. The horizontal line indicates the average of all the measurements: $\sin^2\theta = 0.23155 \pm 0.0018$.

and precision measurements of the W and Z boson masses. All of the experimental results were in excellent agreement with the predictions of the theory. During the 1990s the most precise tests of the theory were carried out at two positron-electron (e^+e^-) colliders, the LEP (Large Electron-Positron) Collider at CERN and the SLC (Stanford Linear Collider) at Stanford Linear Accelerator Center (SLAC) in Stanford, CA, using reaction (4), where f can be any

$$e^+ + e^- \rightarrow Z^0 \rightarrow f + \bar{f} \tag{4}$$

one of the leptons or quarks, and $\bar{f}$ stands for their antiparticles. In these experiments, the value of the $\sin^2\theta$ parameter was measured in a large variety of ways. The agreement among the results is quite good (**Fig. 2**), providing the most sensitive test of the theory, and the most precise value of the weak mixing angle: $\sin^2\theta = 0.23155 \pm 0.00018$.

Prospects. The successful electroweak theory, combined with quantum chromodynamics (QCD), the theory describing the strong nuclear interactions, forms the so-called standard model of particle physics. This standard model has been brilliantly successful in accurately predicting and describing all experimental results over a huge energy range, varying from the electronvolt energies of atomic physics to the 100-GeV energy range of the largest existing particle colliders. As such, it represents a landmark achievement of both experimental and theoretical physics. *See* QUANTUM CHROMODYNAMICS; STANDARD MODEL.

However, in spite of these great successes, the story is not complete, and two major problems remain to be solved in this field. The first one has to do with the realization that the standard model cannot be complete in its present form since it cannot explain the masses of the fundamental constituents, the quarks and leptons. These masses vary over a large range (Table 2), from a few electronvolts to 174 GeV. The basic gauge symmetry on which the standard model is based would indicate that these masses should all be the same. There must therefore be an additional piece of the puzzle, which is usually referred to as the source of the electroweak

symmetry breaking, that remains to be found. Hypothetical ideas about this missing element of the model range from the prediction of a single additional particle, the Higgs boson, to complicated models such as supersymmetry that predict dozens of new elementary particles. The search for this new physics that will lead to a more complete theory motivates research in this field, including the construction of high-energy particle accelerators and colliders. *See* HIGGS BOSON; SUPERSYMMETRY; SYMMETRY BREAKING.

The second outstanding problem in this field is the search for a theory that not only describes the strong nuclear and the electroweak interactions but includes gravity as well. The standard model is based on the principles of quantum mechanics, while the current understanding of the gravitational forces is based on Einstein's theory of general relativity. No one so far has been able to combine these two theories; that is, a quantum theory of gravity does not, as yet, exist. The search for such a grand unified theory is a major focus of activity in theoretical physics. *See* ELEMENTARY PARTICLE; QUANTUM GRAVITATION; RELATIVITY. Charles Baltay

Bibliography. S. L. Glashow, Towards a unified theory: Threads in a tapestry, *Rev. Mod. Phys.*, 52:539–543, 1980; P. Renton, *Electroweak Interactions*, Cambridge University Press, 1990; A. Salam, Gauge unification of fundamental forces, *Rev. Mod. Phys.*, 52:525–538, 1980; S. Weinberg, Conceptual foundations of the unified theory of weak and electromagnetic interactions, *Rev. Mod. Phys.*, 52:515–523, 1980.

Element (chemistry)

An element is a substance made up of atoms with the same atomic number. Some common elements are oxygen, hydrogen, iron, copper, gold, silver, nitrogen, chlorine, and uranium. Approximately 75% of the elements are metals and the others are nonmetals. Most of the elements are solids at room temperature, two of them (mercury and bromine) are liquids, and the rest are gases.

Occurrence and classification. A few of the elements are found in nature in the free (uncombined) state. Some of these are oxygen, nitrogen, the noble gases (helium, neon, argon, krypton, xenon, and radon), sulfur, copper, silver, and gold. Most of the elements in nature are combined with other elements in the form of compounds. The most abundant element on the Earth is oxygen; the next most abundant is silicon. The most abundant element in the universe is hydrogen and the next most abundant is helium. *See* ELEMENTS, GEOCHEMICAL DISTRIBUTION OF.

The elements are classified in families or groups in the periodic table. Elements are also frequently classified as metals and nonmetals. A metallic element is one whose atoms form positive ions in solution, and a nonmetallic element is one whose atoms form negative ions in solution. *See* PERIODIC TABLE.

Atoms of a given element have the same atomic number, but may not all have the same atomic weight. Atoms with identical atomic numbers but different atomic weights are called isotopes. Oxygen, for example, is made up of atoms whose atomic weights are 16, 17, and 18. Hydrogen is made up of isotopes 1, 2, and 3; the isotopes of masses 2 and 3 are called deuterium and tritium, respectively. Carbon is made up of isotopes 11, 12, 13, and 14. Carbon-14 is radioactive and is used as a tracer in many chemical experiments.

All the elements have isotopes, although in certain cases only synthetic isotopes are known. Thus, fluorine exists in nature as ^{19}F, but the artificial radioactive isotope ^{18}F can be prepared. Many of the isotopes of the different elements are unstable, or radioactive, and hence they disintegrate to form stable atoms either of that element or of some other element. *See* ATOMIC MASS; RADIOACTIVITY.

Origin and uses. The origin of the chemical elements is believed to be the result of the synthesis by fusion processes at very high temperatures (in the order of 100,000,000°C or 180,000,000°F and higher) of the simple nuclear particles (protons and neutrons) first to heavier atomic nuclei such as those of helium and then on to the heavier and more complex nuclei of the light elements (lithium, boron, and so on). The helium atoms bombard the atoms of the light elements and produce neutrons. The neutrons are captured by the nuclei of elements and produce heavier elements. These two processes—fusion of protons and neutron capture—are the main ones forming the chemical elements. Furthermore, the energy of the Sun and the stars is derived primarily from the fusion of hydrogen nuclei and electrons to form helium nuclei. It is believed that this element-producing fusion process is occurring even today in hot stars. *See* NUCLEOSYNTHESIS.

The elements form the raw materials for the great chemicals industry today. Various metals are used for structural materials, protective coatings, ornamental devices, jewelry, and tableware. Such nonmetals as chlorine, bromine, hydrogen, sulfur, and nitrogen are important for the manufacture of many of the common chemicals of commerce. Neon is used to make neon light signs, and radon is used as a source of radioactive rays for therapy.

A number of elements, found in only very slight traces or not at all in nature, have been synthesized. Those are technetium, promethium, astatine, francium, and all the elements with atomic numbers above 92. These elements have been synthesized by a variety of nuclear reactions that involve transmuting atoms of one element into atoms of another by bombarding that element with neutrons or fast-moving particles (protons, deuterons, and alpha particles) which will change the atomic number to that of the new element. Not only have these elements been synthesized, but isotopes of all the other elements also have been synthesized. *See* ATOMIC STRUCTURE AND SPECTRA; ISOTOPE; TRANSURANIUM ELEMENTS. Alfred B. Garrett

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999;

N. N. Greenwood and A. Earnshaw, *Chemistry of the Elements*, 2d ed., 1997; S. Hofman, *On Beyond Uranium Journey to the End of the Periodic Table*, 2002.

Element 112

Element 112 was discovered in 1996. It should be a heavy homolog of the elements mercury, cadmium, and zinc, and is expected to be the last element in the *6d* shell. *See* CADMIUM; HALF-LIFE; MERCURY (ELEMENT); ZINC.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

A long-standing claim to having synthesized element 112 is made by a Jerusalem-European Organization for Nuclear Research (CERN) collaboration. A spontaneous fission emitter with a half-life of several weeks was found in 1971 in a tungsten-beam stop at CERN. It followed the chemistry of mercury, and was assigned to element 112. A two-step reaction mechanism was proposed. From the tungsten-beam stop of the 24-GeV particle accelerator strontium-88 is produced by high-energy fission of fissionable spallation products. In a second reaction the ^{88}Sr ions induce a fusion reaction with another tungsten nucleus. The nucleus ^{272}112 (the isotope of element 112) is synthesized at very low excitation energy by radiative capture and is detected after chemical separation by its fission. *See* NUCLEAR REACTION; SPALLATION REACTION; STRONTIUM; TUNGSTEN.

Element 112 was discovered on February 9, 1996, at GSI (Gesellschaft für Schwerionenforschung), Darmstadt, Germany, by detection of ^{277}112, which was produced by fusion of a zinc-70 projectile and a lead-208 target nucleus following the cooling down of the fused system by emission of a single neutron. The fused system was observed at an excitation energy of 12 MeV. Sequential alpha decays to darmstadtium-273, hassium-269, seaborgium-265, rutherfordium-261, and nobelium-257 allowed unambiguous identification by using the known decay properties of the last three members of the chain. In the decay chain (see **illustration**), the first three members are new isotopes. Isotope ^{277}112 has a half-life of 0.7 ms, and it is produced with a cross section of 0.5×10^{-37} cm^2. The new isotopes of darmstadtium and hassium are of special interest. Their half-lives differ by more than four orders of magnitude, and alpha energies are very different, which is characteristic of a closed-shell crossing. At the neutron number $N = 162$, a closed shell was theoretically predicted, and this is verified in the decay chain observed. Isotope ^{269}Hs has a half-life of 9 s, which is long enough to allow studies on the chemistry of this element. The crossing of the neutron shell at $N = 162$ is an important achievement in the field of research on superheavy elements. The stabilization of superheavy elements is based on high fission barriers, which are due to corrections in the binding energies found near closed shells. The shell at $N = 162$ is the first such shell predicted, and is now verified. *See* ALPHA PARTICLES; DARMSTADTIUM; HASSIUM; LEAD; NEUTRON; NOBELIUM; NUCLEAR STRUCTURE; RADIOISOTOPE; RUTHERFORDIUM; SEABORGIUM.

The methods used to produce element 112 were the same as those used for darmstadtium and roentgenium. The decay chain of the new element was observed in an irradiation time of about 3 weeks. The trend toward smaller cross sections has continued. *See* ROENTGENIUM.

Element 112 is unnamed, but there is no doubt that it was discovered. A further decay chain of ^{277}112 was detected in May 2000 by the GSI group, ending by spontaneous fission decay. In 2004, two more decay chains ending by spontaneous fission of ^{261}Rf were found at RIKEN, Japan. The decay chain of ^{277}112 feeds the isotope ^{269}Hs, which was produced in the reaction ^{248}Cm(^{26}Mg,5n) by the GSI nuclear

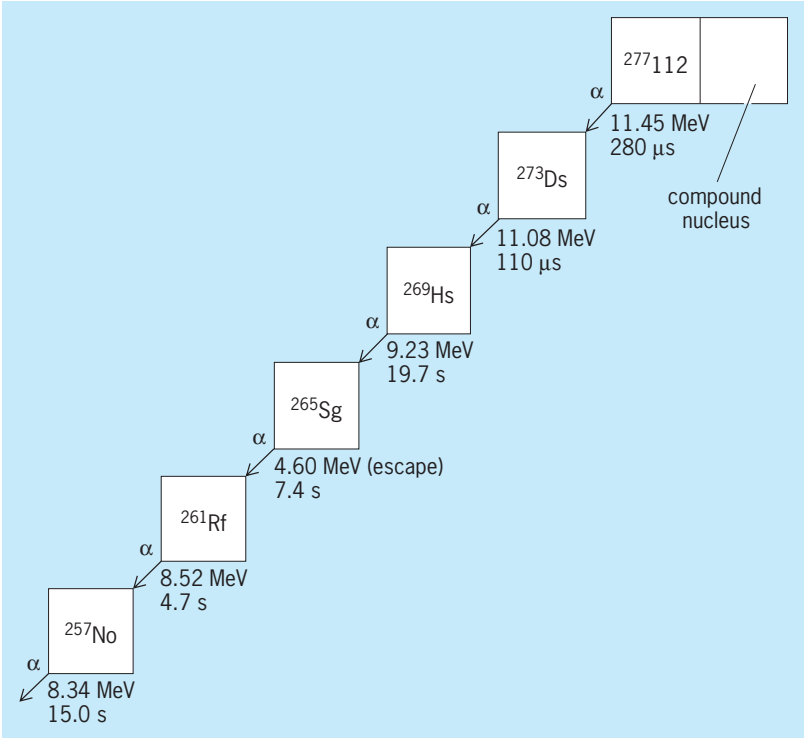

Sequence of decay chains that document the discovery of element 112. Numbers below boxes are alpha energies and correlation times. Element 112 is produced in the reaction ^{70}Zn + ^{208}Pb → ^{277}112 + 1*n*.

chemistry group in 2003. The decay chains observed confirm the pattern of isotopes from $^{277}112$.

Theoretical model calculations of the element 112 chemistry indicate that it may exhibit rather unusual properties. On the basis of its projected position in group 12 of the periodic table it should behave like a heavy metal similar to its homolog mercury (atomic number 80). However, it is predicted that an extra stabilization of its binding electrons may result in a relatively high volatility, possibly giving rise to some similarity with the noble gas radon (atomic number 86). *See* RADON. Peter Armbruster; M. Schädel

Bibliography. S. Hofmann et al., The new element 112, *Z. Phys. A*, 354:229–230, 1996.

Elementary particle

The idea that everything is made from a few basic elements originated in ancient Greece. In the nineteenth century the elementary pieces of matter were the atoms of the chemical elements, but in the first half of the twentieth century atoms were found to be compounds of electrons, protons, and neutrons. These became known as elementary particles, that is, particles that are not compounds of other particles. *See* ELECTRON; NEUTRON; PROTON.

Historical overview. The electrical attraction of opposite electrical charges is what grips negatively charged electrons around the positively charged atomic nucleus; it is the protons that give a nucleus its charge. The amount of positive charge on a proton is equal in magnitude but opposite in sign to that of an electron. This is crucial for the fact that matter in bulk is electrically neutral, whereby gravity controls the motion of planets and our attraction to the Earth's surface. However, why these two particles carry such precisely counterbalanced electrical charges is a mystery. *See* COULOMB'S LAW; ELECTRIC CHARGE; GRAVITATION.

A proton is some 1836 times as massive as an electron, their masses being respectively 939 and 0.511 MeV/c^2, where c is the speed of light A neutron has a mass of 940 MeV/c^2 and is very similar to a proton. In the immediate aftermath of its discovery in 1932, the neutron was thought to be a version of a proton with no electrical charge, though today their relationship is understood to be more profound.

An electron or a proton is stable, at least on time scales longer than the age of the universe. When neutrons are in the nuclei of atoms such as iron, they too may survive unchanged for billions of years. However, an isolated neutron is unstable, with a mean life of 886 s. Neutrons can also be unstable when large numbers of them are packed with protons in a nucleus, which leads to natural radioactivity of many elements. Such neutrons undergo beta decay, a result of the weak force, which converts a neutron into a proton and emits an electron and a neutrino. *See* BETA PARTICLES; RADIOACTIVITY.

Neutrinos have no electric charge and masses that are too small to measure with present techniques. The neutrino emitted in the neutron beta decay has a mass that is less than 3 eV/c^2, that is, less than 1/100,000 the mass of the electron. *See* NEUTRINO.

In the 1930s the known elementary particles were the above four together with the photon (γ), the quantum particle of electromagnetic radiation. The electron and neutrino are known as leptons and are still recognized as elementary in the sense that they are not composed of more fundamental constituents whereas the proton and neutron are built from quarks. Two varieties, or flavors, of quark are required to make protons and neutrons. They are the up quark, (denoted u), with electrical charge that is a fraction $+^2/_3$ of a proton's charge, and the down quark (d), whose fractional charge is $-^1/_3$. The combination uud is sufficient to make a proton and ddu makes a neutron. *See* LEPTON; PHOTON; QUARKS.

The up and down quarks, together with the electron and the above neutrino, form a basic family of what is known as the standard model of fundamental particles. These particles are fermions, defined by having an intrinsic angular momentum or spin of $^1/_2$ in units of Planck's quantum. The photon is a boson, with integer spin of 1. The exchange of photons between electrically charged particles transmits the electromagnetic force. *See* ANGULAR MOMENTUM; ELECTROMAGNETISM; QUANTUM STATISTICS; SPIN (QUANTUM MECHANICS); STANDARD MODEL.

This simple picture began to break down around 1950 with the discoveries of new forms of particle, first in cosmic rays and then in experiments at high-energy particle physics accelerators. With modern hindsight it is possible to classify the discoveries into two classes: leptons and hadrons. *See* COSMIC RAYS; PARTICLE ACCELERATOR.

The muon is a lepton with the same spin and electric charge as an electron but some 207 times more massive at 106.5 MeV/c^2. Today six members of the lepton family are known: three that are electrically charged, the electron (e), muon (μ) and tau (τ) [mass 1777 MeV/c^2]; and three varieties of neutrino known as the electron-neutrino (ν_e), muon neutrino (ν_μ), and tau neutrino (ν_τ) [**Table 1**]. The masses of the neutrinos have not yet been directly measured, but they are believed to have small but nonzero masses. All of these are fundamental particles. Leptons are unaffected by the strong (nuclear) force but feel the weak force and, if electrically charged (electron, muon and tau), the electromagnetic force. *See* STRONG NUCLEAR INTERACTIONS; WEAK NUCLEAR INTERACTIONS.

Particles that feel the strong force are known as hadrons. The cosmic rays revealed the existence of unstable hadrons, some of which became known as strange particles, such as the K meson and lambda hyperon (Λ). Experiments at particle accelerators enabled such particles to be produced when the kinetic energy of particles colliding with atomic nuclei was converted into ephemeral particles, which were revealed by their trails in bubble chambers (or nowadays by electronic detectors). Among these were short-lived heavier versions of the proton and

TABLE 1. The fundamental particles[a]

Gauge bosons $J_C^P = 1^-$ Self-conjugate except $\overline{W^+} = W^-$

Name	Symbol	Charge[b]	Mass and width, GeV	Couplings
Photon	γ	0	0	$A \Rightarrow \gamma A$
Gluon[c]	g	0	0	$A \Rightarrow gA'$
Weak bosons				
Charged[d]	$W^\pm$	± 1	80.4, 2.1	$d \Rightarrow W^+u$
Neutral[e]	Z^0	0	91.2, 2.5	$A \Rightarrow Z^0A$

Fermions[f] $J = {}^1/_2$ All have distinct antiparticles, except perhaps the neutrinos

Name	Charge[b]	Symbol and mass, GeV		Symbol and mass, GeV		Symbol and mass, GeV	
Leptons							
Neutrinos	0	ν_e	$<2\times10^{-9}$	ν_μ	$<.0002$	ν_τ	$<.02$
Charged leptons[g]	–1	e	.00051	μ	.106[h]	τ	1.777[h]
Quarks[c]							
Up type	$^2/_3$	u	.0015–.004	c	1.2	t	175[i]
Down type	$-^1/_3$	d	.004–.008	s	.10	b	4.5

[a] The graviton, with $J_C^P = 2^+_+$, has been omitted, since it plays no role in high-energy particle physics.
[b] In units of the proton charge.
[c] The gluon is a color SU_3 octet (8); each quark is a color triplet (3). These colored particles are confined constituents of hadrons; they do not appear as free particles, which is indicated by placing their entries in boxes.
[d] The branching ratios (%) of the decay modes of the W^+ are:

$u\bar{d}$, $c\bar{s}$	34 each
$\nu_e e^+$, $\nu_\mu \mu^+$, $\nu_\tau \tau^+$	11 each

[e] The branching ratios (%) of the decay modes of the Z^0 are:

$d\bar{d}$, $s\bar{s}$, $b\bar{b}$	15.6 each
$u\bar{u}$, $c\bar{c}$	11.6 each
$\nu_e\bar{\nu}_e$, $\nu_\mu\bar{\nu}_\mu$, $\nu_\tau\bar{\nu}_\tau$	6.7 each
e^+e^-, $\mu^+\mu^-$, $\tau^+\tau^-$	3.4 each

[f] The three known families (generations) of fermions are displayed in three columns.
[g] Any further charged leptons have mass greater than 100 GeV.
[h] The μ and τ leptons are unstable, with the following mean life and principal decay modes (branching ratios in %):

μ	$\tau_\mu = 2.2\times10^{-6}$ s	$e\bar{\nu}_e\nu_\mu$ 100
τ	$\tau_\tau = 2.9\times10^{-13}$ s	$\mu\bar{\nu}_\mu\nu_\tau$ 17, $e\bar{\nu}_e\nu_\tau$ 18, (hadrons) $\bar{\nu}_\tau$ 65

[i] The t quark has a width ≈ 2 GeV, with dominant decay to Wb.

neutron, known as resonances, whose lifetimes of the order of 10^{-23} s are similar to the time it takes a beam of light to traverse a proton. All of these hadrons are composed of quarks. *See* HADRON; PARTICLE DETECTOR.

There are six flavors of quark. The up (u), charm (c), and top (t) have electrical charge $+^2/_3$; the down (d), strange (s), and bottom (b) have charge $-^1/_3$ (Table 1). As individual quarks cannot appear in isolation, a direct measure of their mass is not possible, but approximate scales of mass have been determined. The up and down quarks, when trapped inside hadrons, have energies of around 300 MeV; most of this is due to their motion, their masses being at most 8 MeV/c^2. The strange quark is some 100 MeV/c^2 more massive, the charm quark mass is around 1.2 GeV/c^2 (1200 MeV/c^2), the bottom around 4.5 GeV/c^2, and the top around 175 GeV/c^2. The reason behind these values, or even their qualitative pattern, is not understood.

Strange hadrons contain one or more strange quarks. Hadrons containing charm or bottom quarks are known as charm and bottom hadrons (including a special category known as charmonium and bottomonium). Top quarks are so heavy that no hadrons containing them have been identified. It is even possible that top hadrons cannot form, as the top quark is so unstable that it decays (in a form of beta decay, $t \rightarrow be\nu_e$) before being able to grip to other quarks to make observable hadrons.

Fundamental particles and interactions. The fundamental forces that act on the particles are gravity, the electromagnetic force, and the strong and weak forces. These forces act by the exchange of particles, known as gauge bosons (Table 1). For the electromagnetic, strong, and weak forces, these have been identified, respectively, as the photon, gluon, and weak bosons; the graviton, the quantum of the gravitational force, is firmly predicted by theory, but the prospect of direct observation is exceedingly remote. The gravitational force is so feeble between individual particles that it can be ignored for all practical purposes.

The interactions of particles are responsible for their scattering and transformations (decays and reactions). Because of interactions, an isolated particle may decay into other particles. Two particles passing near each other may transform, perhaps into the same particles but with changed momenta (elastic scattering) or into other particles (inelastic scattering). The rates or cross sections of these transformations, and so also the interactions responsible

for them, fall into three groups: strong (typical decay rates of 10^{21}–10^{23} s^{-1}), electromagnetic (10^{16}–10^{19} s^{-1}), and weak ($<10^{15}$ s^{-1}). *See* FUNDAMENTAL INTERACTIONS; GLUONS; GRAVITON.

The photon, gluon, and weak bosons have spin 1. The photon and gluon have no mass, whereas the weak bosons are massive. The weak bosons may be electrically charged (the W^+ and W^-), with masses of 80.4 GeV/c^2, or neutral (the Z^0), with mass of 91.2 GeV/c^2. The fundamental couplings are known as gauge couplings (Table 1), in which a gauge boson is absorbed or emitted by a fundamental particle (fermion or gauge boson). *See* GAUGE THEORY; INTERMEDIATE VECTOR BOSON.

A massless photon with little or no energy can be emitted and absorbed by electrons in atoms or molecules within the normal laws of energy and momentum conservation. However, when a neutron undergoes beta decay, it converts to a proton by emitting a "virtual" W^-, violating energy conservation by some 80 GeV. The Heisenberg uncertainty principle of quantum mechanics states that such a violation can occur but only for a limited time, within which the W^- converts into an electron and neutrino, which are the real end products of beta decay. The energy violation of 80 GeV is so extreme that the ensuing time scale, and in consequence the distance through which the W^- can travel, are exceedingly short. This limits the effects of the forces associated with W exchange whereby it appears "weak." *See* ENERGY; UNCERTAINTY PRINCIPLE.

Data show that the fundamental couplings of a photon or of a W or Z^0 to a lepton or quark are similar in magnitude such that the electromagnetic and weak forces are now regarded as two manifestations of a single electroweak force. The apparent difference in their strengths is because the weak bosons are massive whereas the photon is massless. These differences in masses, and consequent differences in the effective forces, are examples of symmetry breaking. Theorists believe that the source of this broken symmetry is a spinless particle known as the Higgs boson. It is the interaction between the Higgs boson and the fundamental fermions and bosons that gives them their masses. The details of this interaction remain to be understood, and the first step is to confirm the existence of the Higgs boson. Theory predicts that its mass exceeds 100 GeV/c^2 and that it should be manifested at experiments at CERN in Geneva, Switzerland, commencing from 2007. *See* ELECTROWEAK INTERACTION; HIGGS BOSON; SYMMETRY BREAKING.

Stability. Most particles are unstable and decay into particles with lower masses. Massless particles, such as the photon, are stable. Neutrinos, long thought to be massless and stable, are now known to have the possibility of small masses; the lightest neutrino is the lightest fermion and is stable as its decay would be into bosons, which could not conserve angular momentum. The present view is that the only massive particles that are strictly stable are the electron and the lightest neutrino or neutrinos. The electron is the lightest charged particle; its decay

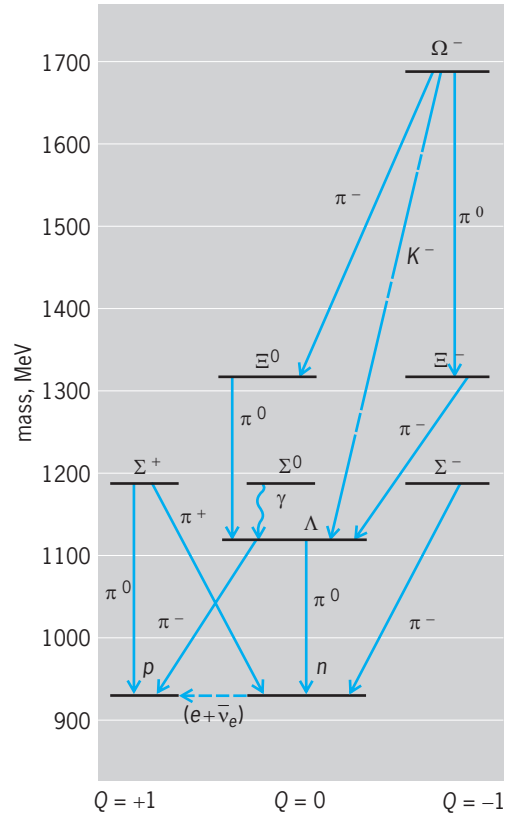

Fig. 1. Decay modes of quasistable baryons made from *u*, *d*, or *s* quarks.

would be into neutral particles and could not conserve charge. The proton may be stable, though there is no fundamental reason why it need be; it is to all practical purposes stable with a half-life exceeding 10^{30} years.

Unstable elementary particles must be studied within a short time of their creation, which occurs in the collision of a fast (high-energy) particle with another particle. Such high-energy particles occur naturally in the cosmic rays, but their flux is small; thus most elementary particle research is based on high-energy particle accelerators.

Hadrons can be divided into the quasistable and the unstable. The quasistable hadrons (**Fig. 1**) are those that are too light to decay into other hadrons by way of the strong interactions, such decays being restricted by the requirement that flavors be conserved. The quasistable hadrons that decay through weak interactions have long mean lives—more than 10^{10} times the characteristic time of strong interactions, $\hbar/(m_\pi c^2) = 0.5 \times 10^{-23}$ s, where m_π is the mass of the π meson or pion, and $\hbar$ is Planck's constant divided by 2π. Three hadrons, π^0, η, and Σ^0, can decay by way of the electromagnetic interaction. These three have mean lives of the order of $10^5 \times \hbar/(m_\pi c^2)$.

The important practical distinctions in the experimental study of interactions are among (1) the stable massive particles (electrons and nuclei), which can be used as target particles as well as in beams; (2) the particles with mean lives greater than 10^{-8} s (γ, ν, $\mu^\pm$, $\pi^\pm$, K_L, $K^\pm$), which can be used only

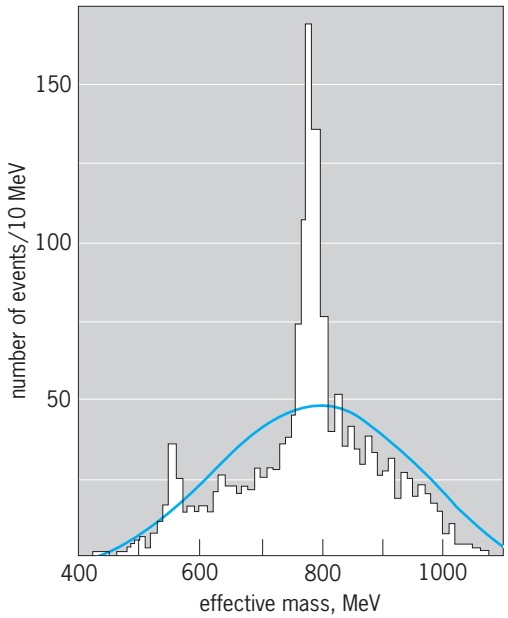

Fig. 2. Observation of the η meson ($m = 550$ MeV) and the ω meson ($m = 783$ MeV) as resonances in the reaction $\pi^+ p \rightarrow \pi^+ p\, \pi^+ \pi^- \pi^0$. Effective mass is the total relativistic energy of three of the emerging pions in their center-of-mass coordinate system. Peaks in the data indicate that collisions can create a short-lived particle of corresponding mass, which decays into three pions.

in beams; (3) the quasistable hadrons with mean lives of the order of 10^{-10} s ($\Xi^{0,-}$, Λ, $\Sigma^{\pm}$, K_S, Ω^-), which have only a small but usable chance of interacting in matter before decaying; and (4) the remaining hadrons, which have a vanishingly small chance of reinteracting except when produced within a nucleus.

The unstable hadrons are also called particle resonances. Their lifetimes, of the order of $\hbar/(m_\pi c^2)$, are much too short to be observed directly. Instead they appear, through the uncertainty principle, as spreads in the masses of the particles—that is, in their widths—just as in the case of nuclear resonances (**Fig. 2**).

The first unstable hadron observed was the $\Delta(1232)$ in pion-nucleon scattering. Flavorless vector mesons can be formed as resonances in collisions between electrons and positrons. This technique has led to the discovery of charmonium and bottomonium spectroscopies: quasiatomic systems made from a charm or bottom quark and its corresponding antiquark.

Antiparticles. To each kind of particle there corresponds an antiparticle, or conjugate particle, which has the same mass and spin but has opposite values of charge, strangeness, charm, or bottomness (quantum numbers that are conserved additively). Antiparticles are denoted by putting a bar over the symbol of the corresponding particle: antiproton $\bar{p}$, antiquark $\bar{q}$ and so forth. Thus a baryon, such as the proton, made of three quarks implies that an antibaryon is made of three antiquarks, $\bar{q}\bar{q}\bar{q}$. All baryons and antibaryons have half-integer values of spin and are fermions. Mesons are examples of hadrons that have integer spins and are made of an equal number of quarks and antiquarks. The simplest example of such a meson is thus $q\bar{q}$ and scores of examples are known. It is theoretically possible to form mesons made of $qq\bar{q}\bar{q}$, known as tetraquarks or molecules, or even of no quarks, made entirely from gluons, known as glueballs. There is some evidence that mesons with no spin, f_0 and a_0, having masses 980 MeV/c^2, are examples of $qq\bar{q}\bar{q}$, and that heavier examples of spinless mesons f_0 with masses of 1500 and 1710 MeV/c^2 may be quantum mixtures of $q\bar{q}$ and a glueball. *See* ANTIMATTER; BARYON; MESON.

Nomenclature. Hadrons of the same spin, of similar mass, and distinguished by their electrical charges are distinguished by writing charge as a superscript, for example, pi meson (π) and sigma hyperon Σ have π^+ and Σ^0. (But p and n are usually written instead of N^+ and N^0.)

With the discovery of large numbers of unstable hadrons, it has become impossible to name each with a different letter; instead letters are used to denote classes of hadrons. First, hadrons with a common value of spin (denoted J), whose quantum wave functions behave the same under space and charge parity, P and C, and whose masses are similar, form families with the same isospin, I. In recognition of these patterns, a naming scheme has developed (**Table 2**). For instance, π_2 is used for any meson with the properties $I = 1$ and $J^{PC} = 2^{-+}$, and a particular hadron is distinguished by giving its approximate mass, in MeV [for instance, π_2 (1680)]. If the hadron is the lightest one of its class, the mass can be omitted from its name; it is always omitted if the hadron is hadronically stable. The mass can be replaced by a specification of the quark state of the hadron if this is known, for instance $\psi(2S)$ instead of $\psi(3686)$. The scheme is described in full in the *Review of Particle Physics*, which is published biennially and updated more frequently on the World Wide Web. *See* I-SPIN; PARITY (QUANTUM MECHANICS); SYMMETRY LAWS (PHYSICS).

Since hadrons have an extended structure, they can have rotationally excited states, as do molecules and deformed nuclei. Such states form a sequence of hadrons with increasing spins ($J_0, J_0 + 1, \ldots$) and masses but with the same values of other quantum numbers (except for space and charge parity, P and C, which alternate in sign). For historical reasons such sequences of hadrons are called Regge trajectories.

TABLE 2. Quantum numbers and names of mesons

J^a	PC	$I = 1$	$I = 0$	$s\bar{s}$	$c\bar{c}$	$b\bar{b}$	$t\bar{t}^b$
0,2,4,...	$-+$	π	η	η'	η_c	η_b	η_t
1,3,5,...	$+-$	b	h	h'	h_c	h_b	h_t
1,2,3,...	$--$	ρ	ω	φ	ψ	Υ	θ
0,1,2,...	$++$	a	f	f'	χ	χ_b	χ_t

[a]Certain mesons, termed exotic states, may have other values of J. For example, the π_1 has $J^{PC} = 1^{-+}$.
[b]The t and $\bar{t}$ quarks may decay before the mesons in this column can form.

Quantum chromodynamics. In addition to their electrical charges and flavors, quarks carry a property called color, which is the source of the force between quarks. As photons couple to electric charge and mediate the electromagnetic force, so massless gluons transmit the color force. The quantum theory of the color force is called quantum chromodynamics (QCD). *See* COLOR (QUANTUM MECHANICS).

The gluon field is coupled to color. The coupling of the gluon to a particle is fixed by the color of the particle and just one universal coupling constant g, analogous to the electronic unit of charge e. There are three varieties of color for quarks and three of opposite sign for antiquarks. The rules of attraction and repulsion for electrostatics have an analogy in chromostatics: like colors repel and unlike colors attract. The latter property has two manifestations. One is directly analogous to electrostatics, in that a quark with a given color attracts an antiquark carrying the opposite sign of color charge. The other possibility is attraction between quarks, each carrying a different color, and in a quantum state that is antisymmetric under interchange of the two particles.

As electric charges of opposite sign lead to neutral atoms, so do the color charges of quarks and antiquarks combine to form overall colorless hadrons. Such states are known as color singlets in the language of the mathematical theory underpinning QCD. As hadrons are colorless, the long-range forces observed between them are no different than those between other particles. The two simplest combinations of quarks that can be colorless are $q_1\bar{q}_2$ and $q_1q_2q_3$; these are found in nature as the basic structure of mesons and baryons, respectively. The exchange of gluons between any of the quarks in these colorless combinations gives rise to an attractive force, which binds them together.

Not just quarks and antiquarks but also gluons carry color, and therefore gluons can mutually attract and repel one another. This situation is very different from electromagnetism, where the photon does not carry charge. The consequence of this self-coupling is that the interaction between two colored particles through the gluon field, which at short distances is an inverse-square Coulomb force, proportional to g^2/r^2 (where r is the distance between the particles), becomes stronger than this inverse-square force at larger r. This property is interpreted by saying that the coupling strength g is effectively larger at larger r; this coupling strength defines the so-called running coupling constant $g(r)$. Computer simulations of the QCD theory imply that this coupling becomes infinitely strong at distances of the order of 10^{-15} m, leading to confinement of the constituent quarks within colorless hadrons, in accord with experiment. The mutual interaction among gluons is predicted to lead to colorless hadrons made from gluons without the need for quarks. As mentioned above, such hadrons are known as glueballs.

This situation is very different from that in quantum electrodynamics (QED). There the long-range force is precisely of the inverse-square form at large distance, with coefficient equal to the product of the charges, e^2. In dimensionless form, this is $e^2/\hbar c \approx 1/137$, the dimensionless parameter α of QED. In QED, vacuum polarization shields a charge—just as in any polarizable medium—resulting in an increase in the apparent charge as one gets closer, penetrating the cloud of shielding charge; in QCD, vacuum polarization—the effect of the gluon self-coupling—is antishielding, resulting in an apparent decrease in the color strength as one gets closer. This short-distance property of QCD is known as asymptotic freedom. *See* FINE STRUCTURE (SPECTRAL LINES); FUNDAMENTAL CONSTANTS; QUANTUM ELECTRODYNAMICS.

At small r, or equivalently at large momentum transfer, the running coupling becomes small and consequently perturbation theory becomes reliable. Calculations in perturbative QCD, such as the gluonic radiative corrections to the cross section for $e^+e^- \rightarrow$ hadrons, are in good agreement with experiment. This is an important reason for the present view that QCD is the correct theory of the hadron glue. The ability to make precise computations in QCD for hadron interactions involving large transfers of momentum underpins much of the experimental planning at facilities such as the Large Hadron Collider (LHC). *See* QUANTUM CHROMODYNAMICS.

Quarkonium. According to the so-called naïve quark model, hadrons are bound states of nonrelativistic (slowly moving) quarks, analogous to nuclei as bound states of nucleons. The interactions between the quarks are taken qualitatively from QCD, namely a confining central potential and (exactly analogous to electromagnetic interactions) spin-spin (hyperfine) and spin-orbit potentials; quantitatively, these potentials are adjusted to make the energy levels of the model system fit the observed hadron masses. This model should be valid for hadrons composed of heavy quarks but not for hadrons containing light quarks (u, d, s), but in fact it succeeds in giving a good description of many properties of all hadrons. One reason is that many of these properties follow from so-called angular physics, that is, symmetry-based physical principles that transcend the specific model. *See* BARYON; MESON.

As mentioned above, mesons with the composition $c\bar{c}$ and $b\bar{b}$ are called charmonium and bottomonium. These names are based on the model of positronium, the atomic bound state of a positron and an electron; the generic name for flavorless mesons, $q\bar{q}$, is quarkonium (**Fig. 3**). The mesons are classified by their intrinsic angular momentum or spin, J, and the behavior of their quantum wavefunctions under parity P and charge conjugation C. *See* POSITRONIUM.

Some states of quarkonium are cleanly observable as resonances in e^-e^+ collisions: $e^-e^+ \rightarrow \gamma^* \rightarrow (q\bar{q})_1{}^{--}$, where γ^* is a virtual photon that materializes into a state of a quark and antiquark with $J^{PC} = 1^{--}$ (the quantum numbers of a photon). This vector quarkonium meson is made at rest in the center-of-mass frame of the colliding e^- and e^+ and is

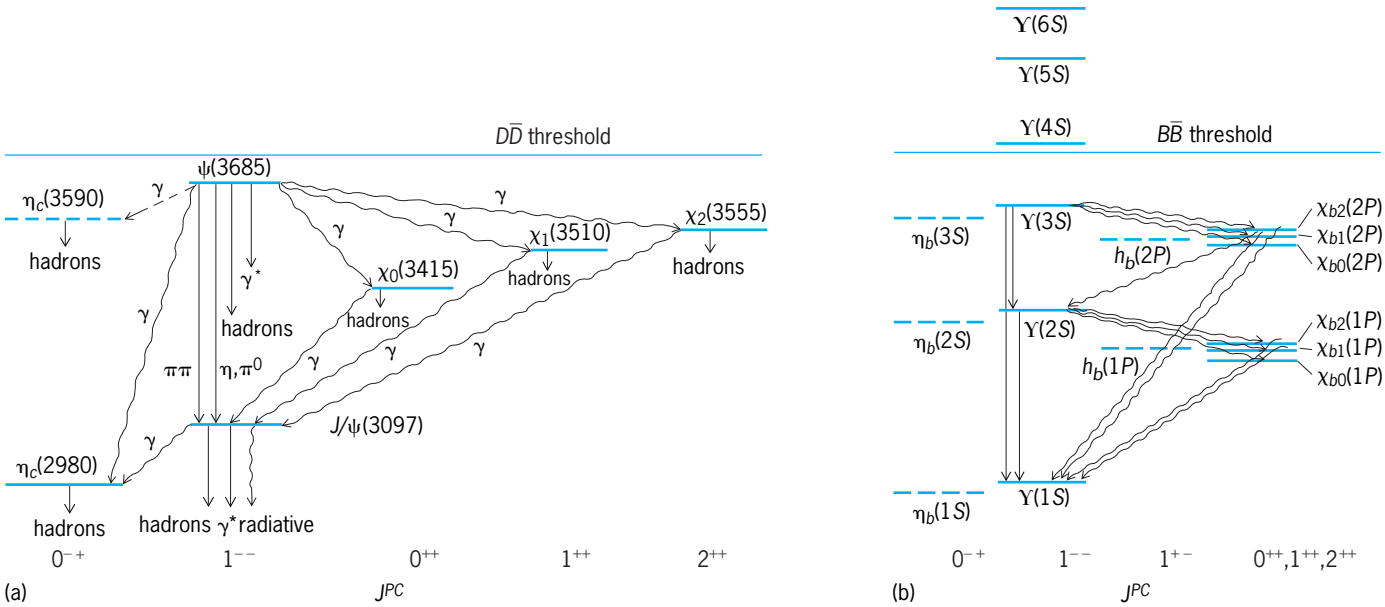

Fig. 3. Level schemes and transitions of (*a*) charmonium and (*b*) bottomonium. Observed radiative transitions are indicated by wavy lines, observed states and other observed transitions by solid lines, and uncertain states and transitions by broken lines. The notation γ^* refers to decay processes involving intermediate virtual photons, including decays to e^+e^- and $\mu^+\mu^-$. (*After Particle Data Group, Review of particle properties, Phys. Lett., 179B:1-350, 1986*)

unaccompanied by other hadrons. Quarkonium states of other quantum numbers can be observed as decay products of excited 1^{--} mesons. Charmonium has been produced also by the annihilation of antiprotons on protons, and observed in the decay products of bottom mesons (B), as in the process $B \rightarrow K + J/\psi(3097)$.

Since both heavy quarkonium and positronium are systems of a fermion bound to its antifermion by a central force, they are qualitatively very similar (quantitatively, their masses and excitation energies are in the ratios of $\sim 10^4$ and 10^7, respectively). However, the potential between the quarks is not $\sim 1/r$ (coulombic), which gives a different ordering in the energy levels for quarkonium and positronium. The pattern observed for quarkonium shows that the quark-antiquark potential is consistent with r (linear potential), which agrees with the prediction of QCD, as calculated by the lattice Monte Carlo method. Information on the spin-spin and spin-orbit parts of the $q\bar{q}$ interaction also comes from the quarkonium energy levels.

Just as positronium decays by annihilation of its constituents into photons, quarkonium can decay by annihilation of its q and $\bar{q}$ into gluons, which then fragment into hadrons. This mechanism results in a strong C dependence of decay rates; for example, the $\eta_c(2981)$ decays $\sim$200 times faster than the $J/\psi(3097)$. Vector quarkonium also has the decays $(q\bar{q})_1^{--} \rightarrow \Upsilon^* \rightarrow x\bar{x}$, where x is any charged fundamental fermion, lighter than the quark q; this decay mode accounts for about 28% of the decays of the J/ψ and 17% of the decays of the upsilon (Υ) meson. *See* J/PSI PARTICLE; UPSILON PARTICLES.

Weak decays of quarkonium, in which one of the heavy quarks is transformed to a lighter quark, are negligibly slow; for example, the chance that a J/ψ meson decays this way is $\sim 10^{-8}$.

Excited states of quarkonium have decay modes in which the heavy quarks neither annihilate nor transform. If the heavy quarks remain bound, this is a deexcitation, a decay into a lower state of the quarkonium with emission of one or more hadrons or a Υ (radiative decay). Radiative decays are the main source of the η_c, χ_0, χ_1, and χ_2 mesons.

Highly excited quarkonium states can decay into pairs of mesons. An example is the bottomonium decay $\Upsilon(10{,}575) \rightarrow B^-B^+$. The B meson is the lightest hadron which contains a b quark, and so $2m_B = 10{,}545$ MeV is the energy threshold for such decays; the $\Upsilon(10{,}575) = \Upsilon(4S)$ decays to $B\bar{B}$ nearly 100% of the time, with a decay rate 2000 times faster than the total decay rate of the next lower vector bottomonium state, $\Upsilon(10{,}355) = \Upsilon(3S)$. Generally, autoionization decay modes are the fastest decay modes of quarkonium whenever they are kinematically possible. Electron-positron annihilation at an energy of 10,355 MeV has been used as a source of large amounts of B mesons. Comparison of the properties of B and its antiparticle counterpart $\bar{B}$ is used to investigate possible asymmetries between matter and antimatter.

Heavy-quark hadrons. The known heavy quarks are the c (charm), b (bottom), and t (top) quarks, whose masses are larger than the mass of a proton, $\sim$1 GeV. A hadron which contains a single heavy quark resembles an atom where the heavy quark sits nearly at rest at the center, and is a static source of the color field, just as the atomic nucleus is a static source of the electric field. Just as an atom is changed very little (except in mass) if its nucleus is replaced by another of the same charge (an isotope), a

heavy-quark hadron is changed very little (except in mass) if its heavy quark is replaced by another of the same color. This is called heavy-quark symmetry. So, for example, the D and B mesons are similar, except in mass, which plays an important role in the quantitative analysis of their weak decays.

The only hadrons so far observed that contain more than one heavy quark are mesons containing a conjugate pair: $c\bar{c}$ and $b\bar{b}$. The creation of hadrons such as $c\bar{b}$ or ccu from ordinary matter, which contains only u and d quarks, requires the creation of two pairs of heavy quarks at nearly the same place and time; this is highly improbable, hence the difficulty in producing such states.

Bottom mesons B made from $b\bar{u}$ or $b\bar{d}$ and their antiparticle counterparts $\bar{B}$ made of $\bar{b}u$ or $\bar{b}d$ have special interest as their weak decays are not symmetric between particle and antiparticle. It is hoped that the study of this asymmetry can shed light on the nature of the large-scale asymmetry between matter and antimatter in the universe. To this end, over 10^9 B and $\bar{B}$ ("B-bar") pairs have been made at customized "B-factories": electron-positron colliders tuned to an energy of around 10.5 GeV where the production rate of these mesons is optimal. Such experiments have taken place at Stanford, California, and the Belle facility at KEK in Tsukuba, Japan, and definitive results have begun to emerge.

Weak interaction. The weak interaction changes the flavor of particles as in its most familiar manifestation, the beta decay of the neutron, which transmutes into a proton, electron, and neutrino. Neutrinos, which have no color or electric charge, and hence do ont respond to strong or electromagnetic forces, are ideal probes of the weak interaction.

All experimental results at low energies are consistent with the weak reactions being point four-Fermi interactions, that is, occurring when four fermions are at one point in space-time. This was the essence of Fermi's original theory of beta decay. As noted above, it is now known that such beta decays are the result of the exchange of a massive charged vector (spin-1) particle, the $W^{\pm}$ (W^+ and W^- are antiparticles of one another), which is emitted and absorbed by the so-called weak charged current. From the quark point of view, the so-called charged-current weak interactions are flavor-changing; here flavor is meant in a broad sense which includes leptons. This particle is part of a renormalized field theory, which is a spontaneously broken gauge theory (by the Higgs mechanism). This is the electroweak theory of S. Weinberg, S. Glashow, and A. Salam, which now is accepted as the standard model of electroweak interactions. It has four gauge fields, whose quanta are the W^+ and W^-, a massive neutral vector particle Z^0, and the photon.

The Z^0 is coupled to all particles with strengths controlled by a single parameter, the Weinberg (or weak) angle θ_W, whose value is determined by observation of the neutral-current weak interactions, which result from exchange of the Z^0. These were unknown before this theory suggested their existence. The particle that emits a Z^0 remains unchanged, as is the case with the emission of a photon.

Universality. An important property of the weak interaction is universality. This property was first observed in the fact that the interactions responsible for the beta decays $\mu \to e\nu_e\nu_\mu$ and $n \to pe\nu_e$ are nearly equal; that is, the couplings $W\nu_\mu\mu$ and Wud of the $W^{\pm}$ are nearly equal. (Likewise it is found that the couplings $W\mu\nu_\mu$ and $We\nu_e$ are equal.) This near-equality points to a deep similarity of leptons and quarks.

The universality is also seen in Wcs and Wtb couplings, which are also almost equal. In addition to these couplings within a generation, there are feebler couplings between the generations, at least for the quarks. This was originally found in the existence of two couplings involving the u quark, namely Wud (strangeness-conserving) and Wus (strangeness-changing), the latter about one-fourth as large as the former. This contrasts with the leptons, where the ν_e is involved in only one coupling, namely $W\nu_e e$ (about equal in strength to Wud), and $W\nu_e\mu$ vanishes. Nonetheless, an overall universality still occurs when all the possible transitions are accounted for in the mathematical matrix formalism of N. Cabibbo, M. Kobayashi, and T. Maskawa (involving the so-called CKM matrix). The numbers in the CKM matrix cannot, in general, be made real. As a result, the W coupling leads to an asymmetry between quarks and antiquarks, with potential implications for the large-scale asymmetry observed between matter and antimatter in the universe. However, while the values of the CKM matrix elements have been experimentally measured, so far there is no theory for them.

There is indirect evidence that a similar matrix formalism might apply to the leptons and that transitions such as $W\nu_e\mu$ occur. Such ideas also suggest that transitions between the electron and the muon might occur, albeit very rarely. This is an active area of research.

Weak bosons. The weak bosons W^+, W^- ($=\overline{W^+}$), and Z^0 were first observed in 1983 in proton-antiproton collisions. The $W^{\pm}$, and similarly the Z^0, are made by the annihilation of a quark pair in a $p\bar{p}$ collision, $q\bar{q} \to W \to$ decay products, where the q and $\bar{q}$ are contained in an incident proton p and antiproton $\bar{p}$, respectively. The branching ratios for the decays into quarks (Table 1) have been predicted from theory and confirmed in experiment. What the weak bosons do not decay into is important information about the nonexistence of additional fundamental particles (quarks, leptons, gauge, bosons, or Higgs particles) with masses less than 45 GeV. The decay modes $e^{\pm}\nu e$ and $\mu^{\pm}\nu_\mu$ of the W^+, and e^+e^- and $\mu^+\mu^-$ of the Z^0, are easy to observe—despite the rarity of the W or Z production (about 1 per 10^6 $p\bar{p}$ collisions)—because no other process makes many leptons with momenta as high as 40 GeV, at a large angle to the incident p and $\bar{p}$ beams. Because there are three times as many kinds of quarks as

leptons (each flavor of quark comes in three colors), most decays of the W and Z are into pairs of quarks, which fragment into 40-GeV jets; but these decays are hard to observe because of the background of similar jets made by the scattering process $q\bar{q} \rightarrow q\bar{q}$.

The Z is made cleanly by e^+e^- annihilation, enabling its decay modes to be easily seen. The decays $Z \rightarrow \nu_l \nu_l$, where ν_l is a neutrino, cannot be seen directly, but at least the total branching ratio into these modes can be deduced by observation of the total width of the Z. With the assumption of universality, this determines the number of light (mass < 45 GeV) neutrinos to be 3; that is, there are none beyond those presently known: ν_e, ν_μ, and ν_τ. By implication, if every neutrino is accompanied by a charged lepton, and this lepton pair is also accompanied by a pair of quarks (charges $^2/_3$ and $-^1/_3$), the three generations of leptons and quarks are the totality of such fermions.

Grand unified theory. The negative charge of the electron and the positive charge of a proton are equal to better than 1 part in 10^{19}. As the electron appears to be a fundamental fermion, as are the quarks that make up the proton, this equality is remarkable and hints at some deeper relation between leptons, such as the electron, and the quarks. The balance is achieved by a further act of balance: the electrical charges of quarks come in fractions that are $^1/_3$ of an integer (in units of a proton), and the strong color forces acting on the quarks cluster them in threes. The three-ness of the color charges conveniently conspires therefore with the third-integer nature of the electric charges, suggesting some relation between the electric and color charges.

These hints are further solidified when we look at the quantum field theories of the forces. The carriers of the electromagnetic and color forces are massless spin-1 particles, the photon and gluon. The weak force also is carried by spin-1 particles, the W and Z, though here they are massive. Apart from this mass effect that spoils the symmetry, and the fact that the weak interactions do not respect mirror symmetry (parity), the underlying quantum theories of the three forces show profound similarities.

Their strengths differ at room temperature, due to the different strengths of the interaction between the spin-1 particles and the fermions and also to the fact that the W and Z are massive. However, in experiments at higher energies these differences have been observed to die away. Experiments at the LEP accelerator at CERN have shown that when electrons and positrons annihilate at a total energy of 90 GeV, the Z boson is produced with a strength that is essentially the same as that of the electromagnetic interaction. Studies of the W boson at these extreme energies show that here too the weak interaction has an intrinsic strength that is united with the electromagnetic and has been obscured at low energies by the large masses of the W and Z, in contrast to the massless photon. Thus, as noted above, it is now accepted that the electromagnetic and weak forces are two manifestations of a single electroweak force.

The most familiar form of the weak interaction is the beta decay, such as that of a neutron into a proton, electron, and neutrino, which is ultimately driven by a down quark transmuting to an up with the emission of a virtual W^-. The coupling of the W to quarks appears to be identical to its interaction with leptons: were the W boson the only probe of leptons and quarks, they would be indistinguishable. This identity further strengthens the belief that leptons and quarks are profoundly related.

In theories that postulate such a relation, the leptons and quarks occur together in multiplets of the large symmetry group called families (or generations). The known fundamental fermions do seem to fall into three families (Table 1). Each family consists of a doublet of leptons (neutrino [charge 0] and charged lepton [charge $-e$]), and a color triplet of doublets of quarks (up-type [charge $^2/_3\ e$] and down-type [charge $-^1/_3\ e$]). (From the total width of the Z^0, it is deduced that there are no more families whose neutrino has mass <45 GeV.) However, there is no known connection between the leptons and the quarks of each family; for example, it is only because (ν_e,e) is the lightest lepton doublet and (u,d) is the lightest quark doublet that they are regarded as members of the same family.

The main difference between them is that quarks carry color and respond to the strong force whereas leptons do not. Here, though, there are further hints of unity at high energies. The strength of the color forces in QCD weakens with increasing energy; the strength of the analogous electroweak interaction, according to the quantum theory, increases. Extrapolating to higher energies, the theories imply that these forces all have a similar intrinsic strength at 10^{15} GeV, known as the grand unification energy. In grand unification theories, this energy is the mass m_{lq} of hypothetical superheavy "leptoquark gauge bosons," analogous to the $W^\pm$ and Z^0 bosons of the electroweak subtheory. At momentum transfers larger than this, the mass of the leptoquark bosons is irrelevant and the large symmetry is unbroken; for smaller momentum transfers, as at low energies, the symmetry is broken and the three forces appear different, the more so the lower the momentum transfer.

The couplings of the leptoquark gauge bosons turn leptons into quarks, or vice versa (this is the reason for the name leptoquark), or quarks into antiquarks. There are many theoretical ideas on the mathematical structure of the true unified theory. Different approaches that agree with known phenomena have widely different implications for the properties of new particles that are implied by the theory. Nonetheless, there are certain common features, such as that the exchange of a leptoquark boson can therefore result in the transformation $qqq \rightarrow lqq$, for example, $p \rightarrow e^+\ \pi^0$. This baryon- and lepton-number-violating interaction is a

much weaker interaction than the analogous ordinary weak interaction, because leptoquark bosons are much heavier than weak bosons. The implication that protons are unstable, albeit with a half-life that exceeds the age of the universe by many orders of magnitude, is a rather general consequence of attempts to build an empirically successful unified theory. Studies of large samples of matter in the hope of finding evidence for a single proton decay have been made, but no definitive evidence for proton decay has yet been found. *See* GRAND UNIFICATION THEORIES.

Accelerators. For a century, beams of particles have been used to reveal the inner structure of atoms. These beams have progressed from naturally occurring alpha and beta particles, courtesy of natural radioactivity, through cosmic rays, to intense beams of electrons, protons, and other particles at modern accelerators. By smashing the primary beams into a target, some of the energy can be converted into new particles, which can themselves be accumulated and made into secondary beams. Thus, beams of pions, kaons, muons, and neutrinos have been made, as well as antiparticles such as positrons and antiprotons. There are even beams of heavy ions—atoms stripped of their electrons—which enable violent collisions between heavy nuclei to be investigated.

There has also been a renewed interest in cosmic rays, where nature provides particles at energies far beyond anything that we can contemplate achieving on Earth. The problem is that such rays come at random and are much less intense than beams made at accelerators.

The basic principle is that electrically charged particles are accelerated by electric forces. If enough electric force is applied to an electron, say, it will be accelerated along a straight path, as in the linear accelerator at Stanford in California, which can accelerate electrons to energies of 50 GeV in a distance of 3 km (2 mi).

Under the influence of a magnetic field, the path of a charged particle will curve. By using electric fields to speed them, and magnetic fields to bend their trajectory, particles can be guided around circles over and over again. This is the basic idea behind huge rings, such as the 27-km (17-mi) LEP circular accelerator at CERN, which accelerated electrons and positrons to 100 GeV each.

Originally, the beams were directed at static targets. In a linear accelerator aimed at a static target, the debris of the collision is propelled forward, just as a stationary car is shunted forward when another car crashes into its rear. When a beam hits a stationary target, the hard-won energy of the beam particles is transferred largely into energy of motion—into moving particles in the target—and is effectively wasted. This problem is overcome if particles can be brought to collide head-on, so that their energy can be spent on the interaction between them. In such a collision the debris flies off in all directions, and the energy is redistributed with it—none is wasted in setting stationary lumps in motion. Thus, there has been an increasing strategy of making counterrotating beams of particles and antiparticles, such as electrons and positrons, or protons and antiprotons, and colliding them head-on.

The major application has been to enable collisions between particles and antiparticles, principally protons and antiprotons or electrons and positrons. A limit to the energy attainable at a circular accelerator of electrons or positrons is that these high-energy particles radiate away energy when they travel on a circular path. This "synchrotron radiation" is greater the tighter the radius of the orbit and the higher the energy of the particle. Protons and antiprotons also emit synchrotron radiation; but with a mass of nearly 2000 times that of an electron, they can reach much higher energies before the amount of energy lost becomes significant, and they also pack a bigger punch. Hence they are the prime choice when the aim is to reach out to previously unexplored higher energies, as at the Large Hadron Collider. *See* SYNCHROTRON RADIATION.

At CERN the 27-km ring of magnets that accelerated electrons and positrons to 100 GeV has been replaced by the superconducting magnets of the Large Hadron Collider, which can guide protons up to energies of 8000 GeV (8 TeV). Counterrotating beams will collide head-on at a total energy of 16 TeV. Experiments at this facility were scheduled to begin in late 2007.

In order to carry out fine-detail studies of new particles discovered at the Large Hadron Collider, there are plans to produce large numbers of them under more controlled conditions. To do so, the plan is to generate electron-positron collisions at the optimum energy. As this energy is expected to be several hundred GeV, two linear accelerators will be required—one for the electrons and the other for the positrons—which are aligned so as to produce head-on collisions of the beams.

Supersymmetry. The fundamental particles of matter, the leptons and quarks, are all fermions with spin $^1/_2$. The forces that act on them are mediated by the photon, the gluon, and the W and Z bosons, all with spin 1. Grand unification is based on this common feature of bosons as force carriers acting on the basic fermions. Supersymmetry (SUSY) theory implies that there is a further symmetry between the forces and the matter particles, such that the known fermions are partnered by new bosons, and the known bosons by new fermions, with novel forces transmitted by these fermions. It is hypothesized that this partnering may lead to a more complete unification between particles and forces.

In SUSY the families of bosons that twin the known quarks and leptons are known as superquarks and superleptons, more commonly referred to as squarks and sleptons. If SUSY were an exact symmetry, each variety of lepton or quark would have the same mass as its squark or slepton sibling. However, the selectron empirically must have a mass greater than 100 GeV, which implies that it must be hundreds of thousands of times more massive than the electron. Similar remarks follow for all the sleptons and

squarks, implying that SUSY is a very badly broken symmetry.

An analogous statement applies to the superpartners of the known bosons. The naming convention here is to add -ino (pronounced "eeno") to denote the superfermion partner of a standard boson. Thus there are predicted the photino, gluino, wino, and zino. The hypothetical graviton, carrier of gravity, is predicted to have a partner, the gravitino. Here too the SUSY is badly broken, and these "inos" must have masses far greater than their conventional counterparts.

Although SUSY must be badly broken, in the sense that the masses of the particles and sparticles differ greatly, the basic ideas are appealing mathematically and merge as a fundamental symmetry of space and time as encoded in Einstein's theory of relativity and the quantum theory. The resulting pattern of sparticles solves technical problems in the quantum theories of the forces at high energies and the response of the particles to those forces, in particular preventing nonsensical predictions such as that certain processes occur with infinite probability. Quantum effects, due to SUSY particles emerging fleetingly as virtual particles in accord with the uncertainty principle, can affect measurable quantities in present experiments. From such data it is predicted that the lightest sparticles may occur with masses of a few hundred GeV and as such be observable at experiments at the Large Hadron Collider. There is also the tantalizing possibility that the lightest sparticles are electrically neutral, such as the photino or gluino, and are metastable. As such they could form large-scale clusters under their mutual gravitational attraction and form a substantial part of the dark matter of the universe. *See* DARK MATTER; SUPERSYMMETRY; WEAKLY INTERACTING MASSIVE PARTICLE (WIMP).

Mass. The masses of the elementary particles have been discussed previously. Here, a description will be given of how these masses affect physical phenomena, and some of the enigmas associated with mass will be summarized.

The mass of matter in bulk is almost entirely due to the nucleons, protons, and neutrons, that form the nuclei of atoms. The mass of the proton is caused by the confining effects of QCD, whereby the quarks and gluons that make up the proton are entrapped within a sphere of radius approximately 10^{-15} m. Heisenberg's uncertainty principle implies that when spatially constrained to such a distance, the constituents gain momentum and energy of hundreds of MeV. The equation $E = mc^2$ then implies that this energy acts as an inertia or mass, leading to the 938-MeV/c^2 mass of the proton.

The individual up and down quarks within the nucleons carry masses of only a few MeV/c^2. Being permanently confined within hadrons, a direct measure of their masses is not possible, but they can be determined indirectly to be in the range 1.5–4 MeV for the up and 4–8 MeV for the down. It is the greater mass of the down quark that gives the neutron its slightly larger mass relative to the proton.

The strange quark is approximately 100 MeV/c^2 more massive than the up and down, which gives the extra mass to strange hadrons relative to their nonstrange counterparts. Charm, bottom, and top quarks have larger masses, approximately 1.2, 4.5, and 175 GeV/c^2, respectively. The fundamental leptons also have a spread of masses. The neutrinos have masses that are too small to determine, and they had long been thought to be massless; however, the phenomenon of neutrino oscillations shows that at least two of the three have nonzero masses, of the order of a few eV/c^2 at most, possibly only a fraction of this. The electron, muon, and tau leptons have masses of about 0.5, 106, and 1777 MeV/c^2, respectively. The reason for these patterns in the masses is not understood.

The symmetry properties of unified theories would be realized if all the fundamental particles were massless. In the case of the photon, gluon, and graviton, this is the case empirically, but the W and Z bosons have masses of 80.4 and 91.2 GeV/c^2, respectively. As described above, it is their large masses that enfeebles the weak forces transmitted by W and Z relative to the electromagnetic force transmitted by the massless photon.

Relative to the grand unification energy of 10^{15} GeV, all of these masses are negligible, and it is theorized that the true unified symmetries of nature would be exact but for the effect of masses. The nature and origin of these masses is thus a fundamental problem of great interest, and the phenomenon is known as electroweak symmetry breaking. The standard model explains mass by proposing that it is due to a new field called the Higgs field, after Peter Higgs who in 1964 recognized this theoretical possibility. Mass is then the effect of the interaction between the fundamental particles and the Higgs field. Photons do not interact with the Higgs field and thereby remain massless; the W and Z bosons do interact and thereby gain their large masses. The quarks and leptons are also presumed to gain their masses by such interactions.

The Higgs field is manifested in quantum theory by the appearance of particles known as Higgs bosons. In Higgs's original formulation, the boson has spin 0. However, if supersymmetry is realized in nature, there should be a family of such particles including the fermion super-Higgs or "Higgsino." The Higgs particles are predicted to occur with masses between 100 and 2000 GeV/c^2. The Large Hadron Collider at CERN has been constructed so as to be able to cover this entire range of possibilities and thereby to identify the origin of electroweak symmetry breaking.

Frank E. Close

Bibliography. F. Close, *The New Cosmic Onion*, Taylor and Francis, 2006; F. Close, *Particle Physics: A Very Short Introduction*, Oxford, 2004; B. Greene, *The Elegant Universe*, Jonathan Cape, 1999; S. Weinberg, *Dreams of a Final Theory*, Vintage, 1993; W. S. C. Williams, *Nuclear and Particle Physics*, Oxford, 1991; W.-M. Yao et al. (Particle Data Group), Review of particle physics, *J. Phys. G*, 33:1, 2006.

Elements, cosmic abundance of

The average chemical and isotopic composition of the solar system is appropriately referred to as cosmic, since this elemental abundance distribution is found to be nearly the same for interstellar gas and for young stars associated with gas and dust in the spiral arms of galaxies. The Sun makes up more than 99.9% of the mass of the solar system, so the bulk chemical composition of the solar system is essentially the same as that of the Sun. The cosmic abundances of the nonvolatile elements are determined from chemical analyses of a type of meteorite known as CI chondrites, whereas the relative abundances of the volatile elements are determined from quantitative measurements of the intensities of elemental emission lines from the Sun's photosphere. In most silicate-rich meteorites and the Earth, Moon, Venus, and Mars, the most abundant elements are oxygen, magnesium, silicon, iron, aluminum, and calcium. Average solar-system composition consists of 70.7 wt % hydrogen, 27.4 wt % helium, and only 1.9 wt % of all remaining elements, lithium to uranium. Cosmic abundances are now widely referred to as standard abundances in the astrophysical literature. *See* ASTRONOMICAL SPECTROSCOPY; ELEMENT (CHEMISTRY).

The cosmic abundances of the elements are given in numbers of atoms of one element relative to a certain number of atoms of a reference element (see **table**). Silicon is commonly taken as the reference element in the study of the composition of the Earth and meteorites, and the data are given in atoms per 10^6 atoms of silicon. Photospheric abundances in the Sun and other stars are often given in the logarithmic "dex" scale, relative to 10^{12} atoms of hydrogen [$A_{El} = \log(N_{El}/N_H) + 12.00$], where A_{El} is the photospheric abundanceof element El, N_{El} is the number

Cosmic abundances of the elements (expressed as number of atoms per 10^6 atoms of silicon)*.

Atomic number (Z)	Element symbol	CI chondrites†	Solar photosphere†	Cosmic abundance†	Atomic number (Z)	Element symbol	CI chondrites	Solar photosphere	Cosmic abundance
1	H	5.29×10^6	2.79×10^{10}	2.79×10^{10}	44	Ru	1.86	1.93	1.86
2	He	6.04×10^{-10}	2.72×10^9	2.72×10^9	45	Rh	0.344	0.367	0.344
3	Li	57.1	0.351	57.1	46	Pd	1.39	1.36	1.39
4	Be	0.73	0.70	0.73	47	Ag	0.486	0.243	0.486
5	B	16.9	9.9	16.9	48	Cd	1.61	1.64	1.61
6	C	7.58×10^5	9.23×10^6	9.23×10^6	49	In	0.184	1.27	0.184
7	N	5.99×10^4	2.32×10^6	2.32×10^6	50	Sn	3.82	2.79	3.82
8	O	7.66×10^6	1.88×10^7	1.88×10^7	51	Sb	0.309	0.279	0.309
9	F	843	1012	843	52	Te	4.81		4.81
10	Ne	2.36×10^{-12}	3.35×10^6	3.44×10^6	53	I	0.90		0.90
11	Na	5.74×10^4	5.96×10^4	5.74×10^4	54	Xe	3.50×10^{-13}		4.7
12	Mg	1.074×10^6	1.06×10^6	1.074×10^6	55	Cs	0.372		0.372
13	Al	8.49×10^4	8.22×10^4	8.49×10^4	56	Ba	4.49	3.76	4.49
14	Si	$\equiv 1.00 \times 10^6$	$\equiv 1.00 \times 10^6$	$\equiv 1.00 \times 10^6$	57	La	0.4460	0.412	0.4460
15	P	1.04×10^4	7.85×10^3	1.04×10^4	58	Ce	1.136	1.06	1.136
16	S	5.15×10^5	5.96×10^5	5.15×10^5	59	Pr	0.1669	0.143	0.1669
17	Cl	5.24×10^3	8.81×10^3	5.24×10^3	60	Nd	0.8279	0.881	0.8279
18	Ar	9.62×10^{-12}	7.00×10^4	1.01×10^5	62	Sm	0.2582	0.285	0.2582
19	K	3.77×10^3	3.67×10^3	3.77×10^3	63	Eu	0.0973	0.0902	0.0973
20	Ca	6.11×10^4	6.38×10^4	6.11×10^4	64	Gd	0.3300	0.367	0.3300
21	Sc	34.2	41.2	34.2	65	Tb	0.0603	0.0221	0.0603
22	Ti	2.40×10^3	2.92×10^3	2.40×10^3	66	Dy	0.3942	0.385	0.3942
23	V	293	279	293	67	Ho	0.0889	0.0507	0.0889
24	Cr	1.35×10^4	1.30×10^4	1.35×10^4	68	Er	0.2508	0.237	0.2508
25	Mn	9.55×10^3	6.84×10^3	9.55×10^3	69	Tm	0.0378	0.0279	0.0378
26	Fe	9.00×10^5	8.81×10^5	9.00×10^5	70	Yb	0.2479	0.335	0.2479
27	Co	2.25×10^3	2.32×10^3	2.25×10^3	71	Lu	0.0367	0.0320	0.0367
28	Ni	4.93×10^4	4.95×10^4	4.93×10^4	72	Hf	0.154	0.211	0.154
29	Cu	522	452	522	73	Ta	0.0207		0.0207
30	Zn	1.26×10^3	1.11×10^3	1.26×10^3	74	W	0.133	0.359	0.133
31	Ga	37.8	21.1	37.8	75	Re	0.0517		0.0517
32	Ge	119	71.6	119	76	Os	0.675	0.785	0.675
33	As	6.56		6.56	77	Ir	0.661	0.624	0.661
34	Se	62.1		62.1	78	Pt	1.34	1.76	1.34
35	Br	11.8		11.8	79	Au	0.187	0.285	0.187
36	Kr	1.64×10^{-13}		45	80	Hg	0.34		0.34
37	Rb	7.09	11.1	7.09	81	Tl	0.184	0.221	0.184
38	Sr	23.5	26.0	23.5	82	Pb	3.15	2.48	3.15
39	Y	4.64	4.84	4.64	83	Bi	0.144		0.144
40	Zr	11.4	11.1	11.4	90	Th	0.0335		0.0335
41	Nb	0.698	0.733	0.698	92	U	0.0090	<0.0094	0.0090
42	Mo	2.55	2.32	2.55					

*CI chondrite data are from Anders and Grevesse (1989), except for B for which data of Zhai and Shaw (1994) were used. Solar photosphere data are from Grevesse and Sauval (1998). Cosmic abundances are from Anders and Grevesse (1989), except for B, for which the newer data of Zhai and Shaw (1994) were used, and C, N, and O, for which the newer photosphere values of Grevesse and Sauval (1998) were used. For elements with long-lived radioactive isotopes, abundances are present-day values; abundances of these elements 4.5×10^9 years ago were higher.

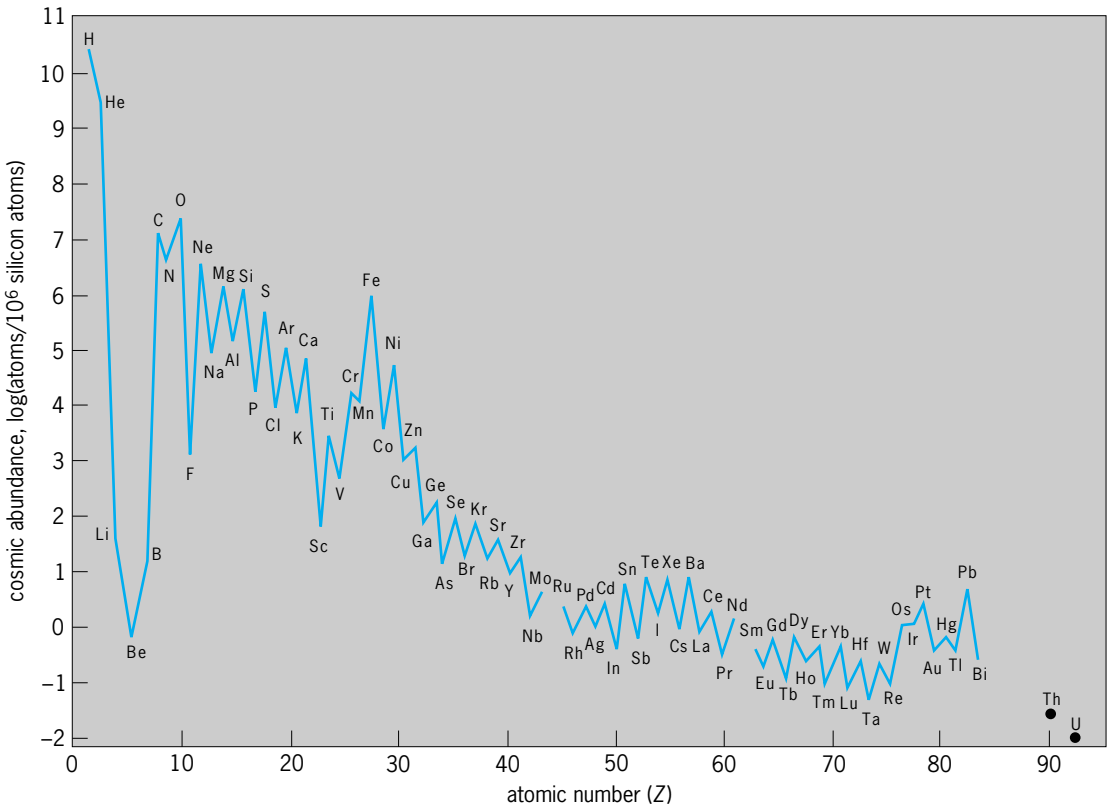

Fig. 1. Cosmic abundances of the elements, plotted as a function of atomic number (*Z*). Abundances are given in atoms per 10^6 atoms of silicon. For radioactive elements, abundances are those at the origin of the solar system 4.5×10^9 years ago. Elements with an even number of protons are more abundant than those with an odd number. (*After E. Anders and N. Grevesse, Abundances of the elements: Meteoric and solar, Geochim. Cosmochim. Acta, 53:197–214, 1989*)

of atoms of element El per 10^{12} atoms of hydrogen, and N_H is the number of atoms of hydrogen (10^{12}). *See* ELEMENTS, GEOCHEMICAL DISTRIBUTION OF.

Uses. Cosmic abundances of elements have several important uses. First, by comparing cosmic abundances to chemical analyses of various types of meteorites, inferences can be made about chemical fractionation processes that occurred in the primitive solar nebula, such as condensation and vaporization. Also, by comparing them to rock compositions, inferences can be made about processes that occurred early in the history of rocky planets, such as separation of a metallic core and differentiation of silicates into mantle and crust. Second, cosmic abundances serve as a standard of comparison for spectroscopic measurements of elemental abundances of the photospheres of other stars and for measurements of elemental and isotopic abundances in cosmic rays. Finally, nucleosynthesis occurs in many different stellar environments. Explanations of nucleosynthesis must account for how elements and isotopes from various astrophysical sources are made and then mixed to form the solar system's average chemical and isotopic composition. *See* NUCLEOSYNTHESIS; SOLAR SYSTEM.

CI chondrites. While there may be no obvious reason why a meteorite should have the average composition of the solar system, there are strong arguments that a particular group of meteorites, the CI chondrites, do in fact have the average nonvolatile elemental abundances of the solar system. The CI chondrites are primitive meteorites that do not actually contain chondrules; rather, they are composed primarily of clay minerals. Their textures indicate some hydrous alteration from an unknown precursor, but this alteration apparently did not affect bulk elemental composition. *See* CLAY MINERALS.

The processes of chemical fractionation in primitive meteorites are well known. The major sources of chemical fractionation are separation of elements by volatility, due to partial evaporation or incomplete condensation of the components within meteorites; separation of siderophile (metal-seeking) from lithophile (rock-seeking) elements, by magnetic or density separation; and separation by the passage of liquid water in parent bodies, whereby elements in water-soluble form, such as sodium (Na), potassium (K), calcium (Ca), and bromine (Br), can be fractionated from elements in water-insoluble forms. All of these processes could potentially have fractionated CI chondrites relative to the true average chemical composition of the solar system. However, when CI chondrite abundances are compared with those of the solar photosphere, the small deviations that are found are independent of volatility, siderophile character, or susceptibility to aqueous alteration. The only significant deviations found are that the nonvolatile elements lithium (Li), beryllium (Be), and boron (B) are depleted in the Sun relative to CI chondrites. Nuclear reactions in the Sun are known to

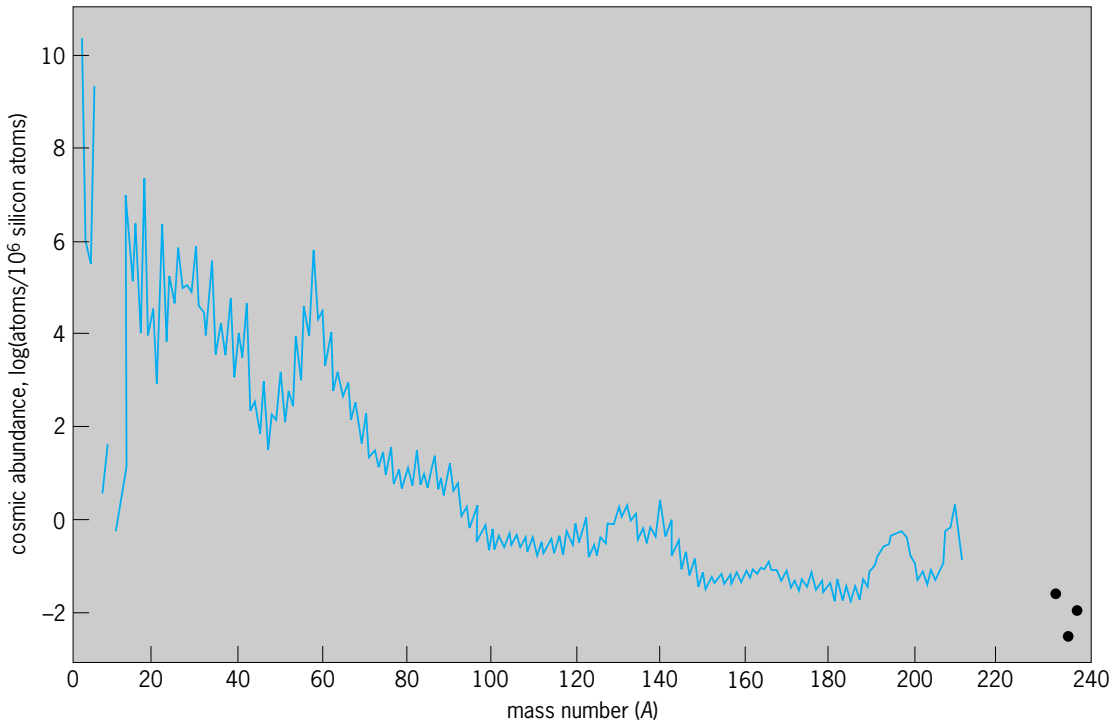

Fig. 2. Cosmic abundances of the isotopes, plotted as a function of mass number (*A*). Abundances are given in atoms per 10^6 atoms of silicon. For radioactive nuclides, abundances are those at the origin of the solar system 4.5×10^9 years ago. For some mass numbers, there is more than one element. A smooth curve can be drawn though the abundances of the odd-mass isotopes. (*After E. Anders and N. Grevesse, Abundances of the elements: Meteoric and solar, Geochim. Cosmochim. Acta, 53:197–214, 1989*)

consume these elements. Similarly, when abundances of odd-mass isotopes are plotted as a function of mass, there are no systematic deviations of groups of elements according to their chemical properties. *See* METEORITE.

Systematic patterns. There are a number of interesting features of cosmic abundances when they are plotted as a function of atomic number (Z) or mass number (A) (**Figs. 1** and **2**). When they are plotted as a function of Z, there are maxima at hydrogen (H), oxygen (O), iron (Fe), xenon (Xe), and lead (Pb); and there is a pronounced odd-even effect where elements with an even number of protons are more stable. There are a number of peaks and valleys that are seen more readily when cosmic abundances are plotted as a function of mass. Hydrogen-1 (^{1}H) and helium-4 (^{4}He) are the most abundant isotopes; there is a gap between ^{4}He and carbon-12 (^{12}C) where Li, Be, and B have very low abundances; there is a decrease in abundances from $A = 12$ (carbon) to $A = 45$ (scandium); a rise to a peak at $A = 56$ (iron) followed by a steep, then gradual decrease in abundances among the heavier elements, interrupted by several small peaks. After $A = 45$, it becomes readily apparent that even-mass isotopes are more abundant than odd-mass isotopes.

The idea that there might be some systematics in the abundances of the elements as a function of mass has a long history. It has been determined that elemental abundances in terrestrial rocks have been chemically fractionated by a variety of processes in the 4.5 billion years of Earth history. In 1947, it was found that the abundances of nuclides, especially of odd mass, are a smooth function of mass number. The most widely used compilation of cosmic abundances of elements indicates a level of precision where deviations from the smoothness of the odd-mass abundance versus mass-number curve (Fig. 2) have become apparent (see table). *See* ISOTOPE.

Comets and interplanetary dust. Other potentially useful samples of cosmic abundances are comets and interplanetary dust. Abundances in dust ejected from Halley's Comet were measured by mass spectrometers on the *Giotto* spacecraft that flew past Halley. Abundances match the solar photosphere for many elements, but there are two notable discrepancies: Halley is low in iron and high in silicon. Halley's Comet, at least the dust portion of it, does not appear to represent pristine solar nebular matter. The most pristine samples of interplanetary dust have been collected by aircraft flying at high altitude. Small grains (less than 15 micrometers in diameter) show no evidence of heating upon atmospheric entry. These grains have the mineralogy and chemistry of CI and CM chondrites. CM chondrites λ (type 2) have a bulk chemical composition indicating some nebular volatility fractionation when compared with CI chondrites, but are more primitive than other types of meteorites. Interplanetary dust particles are so small that analyses are not of high enough precision to judge whether they are a better sample of cosmic abundances than CI chondrites. *See* COMET; COSMOCHEMISTRY; HALLEY'S COMET; INTERPLANETARY MATTER; SPACE PROBE.

Andrew M. Davis

Bibliography. E. Anders and N. Grevesse, Abundances of the elements: Meteoritic and solar,

Geochim. Cosmochim. Acta, 53:197–214, 1989; N. Grevesse and A. J. Sauval, Standard solar composition, *Space Sci. Rev.*, 85:161–174, 1998; H. Suess and H. C. Urey, The abundance of the elements, *Rev. Mod. Phys.*, 28:53–74, 1956; M. Zhai and D. M. Shaw, Boron cosmochemistry, Part I: Boron in meteorites, *Meteoritics*, 29:607–615, 1994.

Elements, geochemical distribution of

The distribution of the chemical elements within the Earth in space and time. Knowledge of the geochemical distribution of the elements in the Earth, particularly in the Earth's crust, and of the processes that lead to the observed distributions make it possible to locate and use efficiently essential elements and minerals and to predict their dispersal patterns when they reenter the natural environment after use.

Formation of the Earth. To understand the present-day distribution of the elements in the Earth, it is necessary to go back to the time of Earth formation approximately 4.5 billion years ago. It is generally believed that the Earth and the other planets in the solar system formed by agglomeration of smaller fragments of solid material orbiting around the Sun. This material had precipitated from a cooling hot gas cloud (the solar nebula), with the most refractory materials condensing out first, the most volatile last. The distribution of elements in the solar system in this early phase thus had much to do with volatility, and the solid material that aggregated to form the planets was a mix of volatile and nonvolatile materials. In a general sense, this is reflected in the overall compositions of the planets: the volatile-rich material accumulated in the cooler regions away from the Sun and formed the gaseous planets such as Jupiter and Saturn; the less volatile material dominated in the so-called terrestrial planets of the inner solar system, for example, Earth and Mars. The material which accreted to form the Earth was probably similar to that found today in meteorites, a mixture of iron-nickel metal and silicate and oxide minerals. *See* COSMOCHEMISTRY; ELEMENTS, COSMIC ABUNDANCE OF; SOLAR SYSTEM.

Although the Earth may have been an approximately homogeneous mixture of accreted materials at the time of its formation, it is now made of many chemically distinct parts. At the fundamental level, these are the core, the mantle, and the crust. While chemical fractionation in the solar nebula depended upon volatility, chemical differentiation within the Earth took place by the separation of molten material from unmelted residue under the influence of gravity. Because large amounts of energy were released from accreting fragments, the early Earth was very hot, and during the accretion stage itself, temperatures in some parts exceeded the melting point of iron metal. Pools of dense molten iron, with dissolved nickel and other elements, aggregated and sank through the Earth under gravity to form the core, leaving behind a mantle of silicate and oxide minerals. The present core constitutes about 32.4% of the Earth's mass. The distinct parts of the Earth possess unique overall compositions (**Table 1**).

Dynamics of mantle. Many geological phenomena on the surface of the Earth are caused by the convection of rocks within the mantle. The dynamic activity of the mantle arises because temperature differences of only a few hundred degrees Celsius reduce the density of large volumes of rock in the mantle, causing them to rise toward the surface in the form of plumes. When the pressure on the heads of such rising plumes of hot rocks decreases, the melting temperature of the minerals is reduced, and melting occurs even though the temperature of the rocks has not increased.

The silicate and oxide minerals of rocks in the mantle melt over a considerable temperature range. The resulting liquid, called magma, is less dense than the unmelted residue, thereby allowing the magma to rise to the surface, where it is erupted to form volcanoes. The Hawaiian Islands and most of the island chains in the Pacific Ocean are examples of magma formation by decompression melting in mantle plumes. *See* SILICATE MINERALS.

This is the process by which the crust has formed over geologic time, and it still continues as is evident from the existence of active volcanoes. The crust constitutes only 0.4% of the Earth's mass; the mantle,

TABLE 1. Element distribution among some of the major subdivisions of the Earth*

Element	Continental crust	Oceanic crust	Upper mantle	Core†
Oxygen	45.3	43.6	44.2	—
Silicon	26.7	23.1	21.0	—
Aluminium	8.39	8.47	1.75	—
Iron	7.04	8.16	6.22	85.5
Calcium	5.27	8.08	1.86	—
Magnesium	3.19	4.64	24.0	—
Sodium	2.29	2.08	0.25	—
Potassium	0.91	0.13	0.02	—
Titanium	0.68	1.12	0.11	—
Nickel	0.011	0.014	0.20	5.5
Sulfur	NA	NA	NA	9.0

*Estimates of element abundances are in percent by weight and are arranged in order of decreasing abundance in the continental crust. Sulfur contents are not well known and are designated "not applicable."

†The estimate for the core is just one of several models. Others substitute light elements such as oxygen, carbon, or silicon for some or most of the sulfur shown here.

1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	57–71 La–Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84	85	86
87	88	89	90 Th	91	92 U												
rare earths		57 La	58 Ce	59 Pr	60 Nd	61	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu	

Key:
lithophile atmophile chalcophile siderophile

Geochemical classification of the elements in relation to the periodic table. The short-lived radioactive elements (43, 61, 84, 85, 86, 87, 88, 89, 91) have been omitted.

from which it and the core have separated, makes up 67.2%. *See* EARTH CRUST.

Large-scale distribution. The large-scale distribution of the elements in the Earth depends on the affinity of each element for specific compounds or phases. Those elements that alloy easily with iron, for example, are mostly sequestered in the Earth's core; those which form oxides and silicate minerals tend to be concentrated in the Earth's crust and mantle. Although many elements display multiple characteristics depending on the chemical environment, a classification according to geochemical affinity is nevertheless useful (see **illus.**). The categories in this classification include atmophile (elements that are gases and concentrate in the atmosphere), lithophile (elements that form silicates or oxides and are concentrated in the minerals of the Earth's crust), siderophile (elements that alloy easily with iron and are concentrated in the core), and chalcophile (elements such as copper which commonly form sulfide minerals if sufficient sulfur is available).

Although both the mantle and the crust are composed largely of lithophile elements, their specific compositions are quite different (Table 1). The differences result from the way in which the crust originates by melting of the Earth's interior. Partitioning between mantle minerals and the melts produced from them determines the crustal inventories of most elements. The major mineral phases in the portion of the mantle that melts to form crustal rocks are olivine [$(Mg,Fe)_2SiO_4$], pyroxene [$(Ca,Mg)SiO_3$], spinel ($MgAl_2O_4$), and garnet ($Mg_3Al_2Si_3O_{12}$). These are idealized chemical formulas; however, virtually every element in the periodic table occurs in the mantle, many dissolved in very small amounts in these mantle minerals, and some concentrated in other very rare minerals which may exist in some parts of the mantle. When melting occurs, those elements that do not fit easily into the structure of the major mantle minerals tend to be partitioned strongly into the melt. These are the elements that are enriched in the crust. In practice, they typically have very different ionic radii or valences than magnesium, iron, calcium, aluminium, or silicon, which are the major constituents of the mantle minerals. An example is the alkali element rubidium, which has a comparatively large ionic radius and therefore cannot fit into the structural sites normally occupied by magnesium, iron, or calcium in the mantle minerals. Estimates suggest that between one-third and two-thirds of the Earth's total inventory of rubidium is contained in the crust which, as mentioned, constitutes only 0.4% of the Earth. Elements such as rubidium that are strongly partitioned into melts are commonly termed incompatible elements by geochemists. The siderophile elements, as already noted, are concentrated in the core, and their abundances in the crust are very low regardless of their valence or ionic radius.

Earth crust. Although geochemists have a good general knowledge of the overall distribution of elements in the core and mantle, much more detailed information is available about the chemical composition of the crust, which is accessible. The crust is actually composed of two major parts with quite different compositions, thickness, and average age: the continental crust and the oceanic crust (Table 1).

The oceanic crust is continually renewed at the mid-ocean ridges and destroyed at subduction zones. It is a direct product of large-scale melting of the upper mantle, and its composition reflects this fact. Compared to the continental crust, it has higher concentrations of magnesium, iron, and calcium and lower concentrations of silicon, potassium, and the highly incompatible elements. Its overall composition is basaltic. *See* BASALT; MID-OCEANIC RIDGE; SUBDUCTION ZONES.

The continental crust is older than the oceanic crust and is more chemically fractionated relative to the mantle. There is good evidence that the continental crust itself is chemically zoned, with the upper parts being more highly enriched in the incompatible elements than the lower parts.

Although all elements are present, the crust is made almost entirely of just nine chemical elements: oxygen, silicon, aluminum, iron, magnesium, calcium, sodium, potassium, and titanium. Oxygen and silicon are by far the most abundant. The most common minerals in the crust are those of the silicate family, in which the basic building block is a silicon atom surrounded by four oxygen atoms in the form of a tetrahedron. The crust is essentially a framework of oxygen atoms bound together by the common cations. *See* CHEMICAL BONDING; SILICATE MINERALS.

A variety of processes act to make the crust chemically heterogeneous on many scales. Many of these processes involve liquid water. Running water physically sorts particles depending on size and density, which are ultimately related to chemical composition. It is also a superb solvent, carrying many elements in solution under different conditions of temperature and pressure, and depositing them when these conditions change. Processes involving water account for many ore deposits, in which extreme concentrations of some elements occur relative to their average abundance in the crust. One example is the circulating hydrothermal solutions in volcanically active parts of the crust, which can leach metals from their normally dispersed state in large volumes of volcanic rocks and deposit them in concentrated zones as the solutions cool and encounter different rock types. Another example is the action of weathering in tropical regions with high rainfall, which can leach away all but the least soluble components from large volumes of rock, leaving behind mineral deposits rich in aluminum or, depending on the original composition of the rocks being weathered, metals such as iron and nickel. *See* ELEMENT (CHEMISTRY); GEOCHEMISTRY; ORE AND MINERAL DEPOSITS; WATER; WEATHERING PROCESSES.

Rivers. The chemical composition of water in rivers depends on many factors, such as the geology of the drainage basin, the climate and its seasonal variation, the amount and composition of the suspended sediment, the acidity of the water, and the amount of biomass that the water contains. In addition, the chemical composition of rivers is affected by the discharge of industrial, municipal, and agricultural wastewater, all of which vary seasonally and enter rivers at different sites along the channel. Therefore, it is difficult to make an accurate determination of the chemical composition of the water of a particular river. However, when the available data are averaged by continent, eight different constituents are found to be dominant in all rivers: calcium, magnesium, sodium, potassium, chloride (Cl^-), sulfate (SO_4^{2-}), bicarbonate (HCO_3^-), and silica (SiO_2) [**Table 2**]. The concentrations of other elements are comparatively low and vary widely among the rivers of the world depending on the factors identified above. *See* RIVER; SILICATE MINERALS.

Seawater. The chemical composition of seawater depends on the amount of each element that enters the ocean each year and the amount that is removed from it. Chemical elements enter the oceans by the discharge of river water and by deposition from the atmosphere either dissolved in rain or as dry particulate fallout. Certain chemical elements (sodium and chloride ions) remain in the oceans for long periods of time and therefore have high concentrations in seawater. Other elements are removed from seawater because they form insoluble compounds (calcium and magnesium), because they are adsorbed on the surfaces of sediment particles (beryllium and potassium), or becausethey are nutrients that are

TABLE 2. Average chemical compositions of rivers in milligrams per liter*

	Ions in solution							
Continent	Ca^{2+}	Mg^{2+}	Na^+	K^+	Cl^-	SO_4^{2-}	HCO_3^-	SiO_2†
N. America	21.2	4.9	8.4	1.5	9.2	18.0	72.3	7.2
S. America	6.3	1.4	3.3	1.0	4.1	3.8	24.4	10.3
Europe	31.7	6.7	16.5	1.8	20.0	35.5	86.0	6.8
Africa	5.7	2.2	4.4	1.4	4.1	4.2	26.9	12.0
Asia	17.8	4.6	8.7	1.7	10.0	13.3	67.1	11.0
Oceania	15.2	3.8	7.6	1.1	6.8	7.7	65.6	16.3
World	14.7	3.7	7.2	1.4	8.3	11.5	53.0	10.4

* After E. K. Berner and R. A. Berner, *The Global Water Cycle*, Prentice-Hall, 1987.
† Silicic acid (H_4SiO_4) is reported as SiO_2.

TABLE 3. Concentrations of major ions in seawater of salinity 35‰ in milligrams per kilogram*

Ion	Concentrations
Ca^{2+}	412
Mg^{2+}	1,290
Na^{+}	10,770
K^{+}	399
Cl^{-}	19,354
SO_4^{2-}	2,712
HCO_3^{-}	120
$SiO_2^{\dagger}$	6

*After E. K. Berner and R. A. Berner, *The Global Water Cycle*, Prentice-Hall, 1987.
†Silicic acid (H_4SiO_4) is reported as SiO_2.

absorbed by marine plants (nitrogen and phosphorus).

The chemical composition of seawater is described in terms of the major ions whose concentrations depend only on the salt content of the water, and in terms of the other ions whose concentrations vary regionally within the oceans, or with increasing depth, or seasonally, or even on a daily basis (**Table 3**). The concentrations of the major ions in seawater are significantly higher than in average river water. Silica is the only exception because it is actively removed from seawater by certain sponges and diatoms which have siliceous skeletons. *See* SEAWATER.

Organic matter. The diversification of the chemical compositions of the rocks of the continental crust and of water that interacts with them is evident also in the composition of organic matter. The tissues of plants and animals are composed primarily of proteins, carbohydrates, lipids, and lignin, all of which contain carbon, oxygen, hydrogen, and minor amounts of other elements such as nitrogen, phosphorus, and sulfur. Carbohydrates and lignin are dominant in plant matter, whereas proteins, carbohydrates, and lipids are abundant in animal tissues. Plants derive certain nutrient elements from soil and from the atmosphere.The essential nutrient elements in plants (**Table 4**) are propagated to herbivorous and carnivorous animals with only minor shifts in composition. *See* PLANT MINERAL NUTRITION.

TABLE 4. Concentrations of essential nutrient elements in plants expressed in weight percent dry tissue

Element	Concentration
Carbon	45
Oxygen	45
Hydrogen	6
Nitrogen	1.5
Potassium	1.0
Calcium	0.5
Phosphorus	0.2
Magnesium	0.2
Sulfur	0.1
Chlorine	0.01
Iron	0.01
Manganese	0.005
Zinc	0.002
Boron	0.002
Copper	0.0006
Molybdenum	0.00001

Some chemical elements are toxic to plants and inhibit their growth even when present in soil or water at low concentrations. The toxic elements include beryllium, arsenic, antimony, selenium, mercury, nickel, copper, cobalt, lead, zinc, silver, cadmium, and tin. These elements may be present in the soils either because they occur naturally or because of environmental contamination by industrial or municipal waste products. Certain toxic elements, such as selenium, accumulate in plants tolerant of the element and enter the food chain when these plants are consumed by herbivorous animals. *See* SELENIUM; ELEMENTS, GEOCHEMICAL DISTRIBUTION OF.

G. Faure; J. D. MacDougall

Bibliography. E. K. Berner and R. A. Berner, *The Global Water Cycle*, Prentice-Hall, 1987; G. Faure, *Principles and Applications of Geochemistry*, 1998; A. E. Ringwood, *Composition and Petrology of the Earth's Mantle*, 1975; B. J. Skinner and P. B. Barton, Jr., Genesis of mineral deposits, *Annu. Rev. Earth Planet. Sci.*, 1:183–211, 1973; S. R. Taylor and S. M. McLennan, *The Continental Crust: Its Composition and Evolution*, 1985.

Elephant

The common name for three living species of mammals in the family Elephantidae, one of several families included in the order Proboscidea. The remaining families contain extinct animals, such as the mammoth. Two of the living species (*Loxodonta africana* and *L. cyclotis*) are indigenous to Africa, and the other (*Elephas maximus*) ranges throughout Southeast Asia.

These animals are terrestrial and entirely herbivorous. They have 26 teeth and the dental formula is I 1/0 M 0/0 Pm 3/3 M 3/3. The nostrils and upper lip are elongated into a proboscis, the trunk, which is a powerful and sensitive organ specialized to form a prehensile, food-gathering structure. The upper incisors protrude from the mouth on either side of the trunk as tusks, which continually grow throughout the life of the animal. The head is massive, and the leathery ears, which are large, especially in the African species, are used as cooling organs as well as for hearing. The hard, thick skin of the elephant is sparsely covered with hair and serves as an insulator, much as does hair or fat in other animals. The eyes are tiny but vision is keen. The tail is short, large columnar legs support the massive body, and the feet may measure 20 in. (50 cm) across.

Elephants live and travel in herds which were originally composed of several hundred individuals, but the usual number is now around 20 animals, with a mature bull, a number of cows and calves, and some younger bulls. Elephants are at ease in water, and they bathe and roll in the mud as a protective

***Loxodonta africana*, African bush elephant. (*Photograph by Robert Thomas and Margaret Orr; © 2001 California Academy of Sciences*)**

measure since, despite the thickness of the skin, they are sensitive to intense sun and insects. These animals are constantly on the move for pasturage, and when they travel they walk in single file with an experienced female in the lead and the males in the rear. During the heat of the day they take refuge in the cooler forested areas; however, there is insufficient herbage in these locations, and later in the day they move out to the tall-grass areas. Bulls reach sexual maturity at the age of 15, while cows mature earlier. The gestation period averages 20–22 months with a single calf being born. Twins are rare. The newborn, which may be 3 ft (1 m) tall and weigh 200 lb (90 kg), grows rapidly, and one elephant was reported to weigh 650 lb (293 kg) at 15 months. The female that is about to give birth is attended by other females, which in effect act as midwives and remain with her until the calf can follow its mother. Lactation continues for months, and the young uses its mouth and not the trunk for suckling.

The African elephant genus *Loxodonta* comprises two species, the larger African bush elephant (*L. africana*) (see **illustration**) and the smaller African forest elephant (*L. cyclotis*). The African bush elephant is the largest living land mammal. An adult bull may weigh as much as 6 tons (5.4 metric tons) and stand over 11 ft (3.3 m) high, while the cow usually weighs about 4 tons (3.6 metric tons). Both the bull and the cow have tusks, although they are small in the cow. The trunk appears annulated, the back is concave, and the tip of the trunk has two triangular lips. They have a life-span of 60–70 years and have not been domesticated.

The Indian elephant is smaller and rarely reaches 10 ft (3 m) in height. It weighs about 4 tons (3.6 metric tons), and the cow is tuskless. The trunk appears to be smooth, the ears are much smaller than those of the African elephant, the back is convex, and between the ear and the eye there is an opening to a gland, the musth gland, which may be associated with sexual activity. This animal has been domesticated and is a valuable beast of burden in much of the Far East. *See* MAMMALIA; PROBOSCIDEA.

Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999.

Eleutherozoa

One of the two subphyla into which the phylum Echinodermata had been customarily divided. The Eleutherozoa are now best considered as comprising at least two distinct subphyla: (1) the Echinozoa, spherical-bodied forms with meridional symmetry; and (2) the Asterozoa, star-shaped forms with radially divergent axes of symmetry. The Echinozoa would seem to have arisen from ancestors similar to cystoids, whereas the Asterozoa seem to be derived from crinoidlike forms. *See* CRINOIDEA; ECHINOZOA; PELMATOZOA.

At one time the Eleutherozoa were thought to comprise the free-living echinoderms, namely, the sea stars, the sea urchins, and sea cucumbers, in which the mouth is directed downward and the anus (if present) is usually placed on the upper surface. The Eleutherozoa were contrasted with the so-called Pelmatozoa, in which the body was attached to the substrate and the mouth and anus opened together on the upper surface, therefore making the gut U-shaped. However, further research showed that both groups are polyphyletic and that the eleutherozoan and pelmatozoan habit (with consequential morphological change) has been adopted independently more than once by various groups of echinoderms. Therefore the original definitions of the subphyla are abandoned because they described unnatural groupings of unrelated animals. *See* ECHINODERMATA.

Howard B. Fell

Bibliography. H. B. Fell, Phylogeny of sea stars, *Phil. Trans. Roy. Soc. London*, ser. B, 735, 246:381–435, 1963; H. B. Fell, The evolution of the echinoderms, *Smithson. Inst. Annu. Rep. 1962*, pp. 457–490, 1963; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Elevating machines

Materials-handling machines that lift and lower a load along a fixed vertical path of travel with intermittent motion. In contrast to hoisting machines, elevating machines support their loads instead of carrying them suspended, and the path they travel is both fixed and vertical. They differ from vertical conveyors in operating with intermittent rather than continuous motion. Industrial lifts, stackers, and freight elevators are the principal classes of elevating machines.

Industrial lifts. A wide range of mechanically, hydraulically, and electrically powered machines are classified simply as lifts (**Fig. 1**). They are adapted to such diverse operations as die handling and feeding sheets, bar stock, or lumber. In some locations with differences in floor level between adjacent buildings, lifts take the form of broad platforms to serve as floor levelers to obviate the need for ramps. They are also used to raise and lower loads between the ground and the beds of carriers when no loading platform exists. Lifting tail gates attached to the rear of trucks are similarly used for loading or

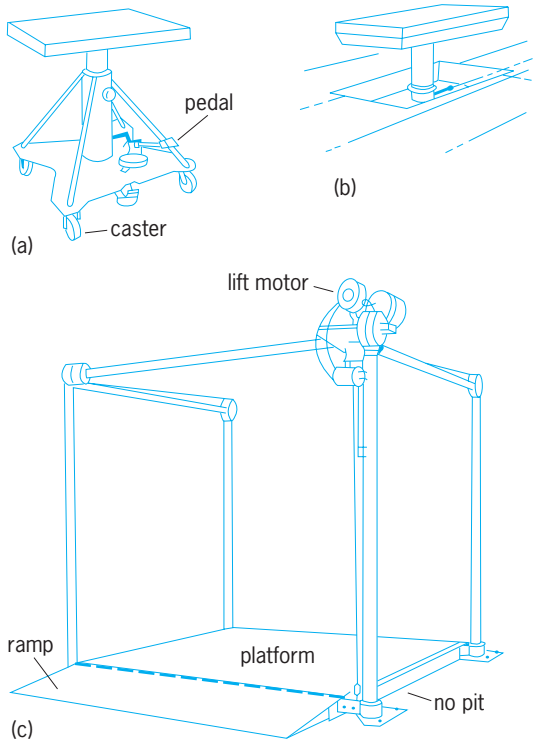

Fig. 1. Examples of industrial lifts. (*a*) Hydraulic elevating work table. (*b*) Hydraulic lift floor leveler. (*c*) Motor-driven floor leveler.

unloading merchandise on sidewalks or roads and at points where the lack of a raised dock would make loading or unloading difficult. These units are usually driven by battery-operated motors on a power takeoff from the drive transmission of the vehicle. Adjustable loading ramps are necessary because heights of truck and trailer beds vary. Advances in mechanized loading and unloading of vehicles have made necessary sturdier, more efficient dock boards or bridge plates, mechanically or hydraulically operated.

Stackers. Tiering machines and portable elevators used for stacking merchandise are basically portable vertical frames that support and guide the carriage, to which is attached a platform, pair of forks, or other suitable lifting device (**Fig. 2**). The operation of these units varies in the sense that the carriage can be raised and lowered by hand, by an electrically driven winch, or by a hydraulic cylinder, which actuates the system of chains or cables, operative by hand lever, pedal, or push button. Early electric motors used on stackers were plugged into adjacent power lines to receive current. This limited the flexibility of the stackers. Since the trend is to make the machines independent of this source, there are now models powered by either storage batteries or by small gasoline or gas engines. Horizontal movement is effected by casters on the bottom of the vertical frame, and can be accomplished manually, or mechanically, by using the same power source as the lifting mechanism. These casters are usually provided with floor locks bolted in position during the elevating or lowering operation.

The basic type of stackers can be varied in several particulars. Masts, which are part of the frame, can be hinged or telescopic, and the platforms can be plain, equipped with rollers, or constructed specially to handle a specific product. Some stackers have devices for tilting barrels and drums or for lifting and dumping free-flowing bulk materials. Used in conjunction with cranes, they are widely applied to the handling of materials on storage racks and die racks. Stackers have a significant place in the development of materials-handling equipment. They are the prototypes of completely powered noncounterbalanced platform and forklift trucks. *See* INDUSTRIAL TRUCKS.

Industrial elevators. Examples of industrial elevators range from those set up temporarily on construction jobs for moving materials and personnel between floors to permanent installations for mechanized handling in factories and warehouses. Dumbwaiters are a type of industrial elevator, having capacities up to approximately 500 lb (227 kg), with a maximum floor space of 9 ft^2 (0.8 m^2); they carry parts, small tools, samples, and similar small objects between buildings, but are not permitted to carry people.

Oil hydraulic plunger electric elevators are designed for low-rise, light- or heavy-duty freight handling. Although they can be installed without special building alterations, they are restricted to buildings with only a few floors because of the limitations of the plunger length and design.

The most common and economical elevator employs electric motors, cables, pulleys, and counterweights (**Fig. 3**). Powered machines impose a severe operating condition on elevators. Elevator platforms and structures are subjected to impact loading, off-balance loading, and extra static loading. To meet these forces, in addition to load forces, freight elevator design and construction provide greatly increased ruggedness over that of passenger units.

Special-purpose freight-handling elevators are equipped with platforms or arms for carrying specific articles such as rolls of paper, barrels, or drums. Some of these elevators load and discharge

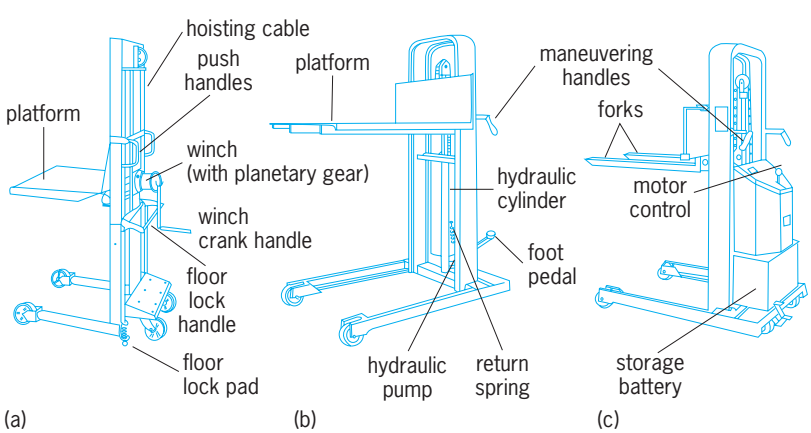

Fig. 2. Three types of electric and hydraulic stackers. (*a*) Hand type. (*b*) Hydraulic foot type. (*c*) Electric lift type.

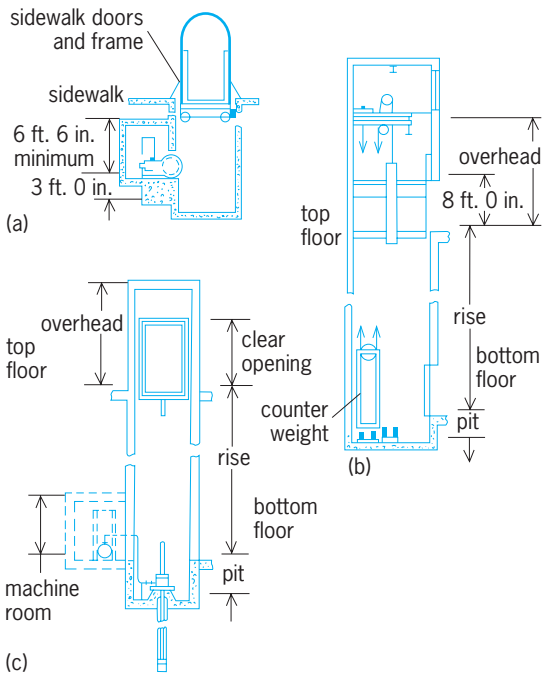

Fig. 3. Three types of industrial elevator. (*a*) Sidewalk elevator type. (*b*) Heavy-duty freight elevator type. (*c*) Hydraulic electric elevator type.

automatically and are arranged so that they can operate at any selected floor by means of remote control. *See* MATERIALS-HANDLING EQUIPMENT.

Arthur M. Perrin

Bibliography. American National Standards Institute, *Safety Code for Elevators and Escalators*, ANS A17.1–1981, ASME, 1981; American Society of Mechanical Engineers, *Safety Code for Elevators and Escalators: Handbook on A17.1*, 1997; E. A. Donoghe (ed.), *Safety Code for Elevators and Escalators*: *Handbook on A17.1*, 1993.

Elevator

A platform or enclosure that is raised and lowered in a vertical hoistway to transport freight or people. The term elevator can also encompass all of the hoisting equipment, motor, cables, and accessories (see **illus.**). *See* ELEVATING MACHINES.

Operation. The closed passenger car of a modern elevator rests inside a steel frame. The car and the car frame ride up and down on steel rails in an elevator shaft or hoistway. Guide shoes or rollers on the frame keep the car in place on the rails. Most elevators also have a heavy weight, called a counterweight, attached to the other end of the steel hoisting ropes that pass over the driving machine pulley. The counterweight offsets much of the weight of the car and passengers, thereby reducing power requirements.

Control devices. The typical elevator control system is made up of a speed-sensing device known as a governor, a clamping device (safety) mounted under each end of the car frame that grips the guide rail when tripped, a tension sheave (pulley) in the pit, and a steel rope. The governor rope makes a complete loop around the governor sheave and the tension sheave in the pit. Since the rope is fastened to and travels with the car, the governor sheave rotates at a speed directly proportional to the speed of the car.

The governor rope is connected to the safeties through mechanical linkages and lift rods. If the hoist ropes break or the car overspeeds, the governor

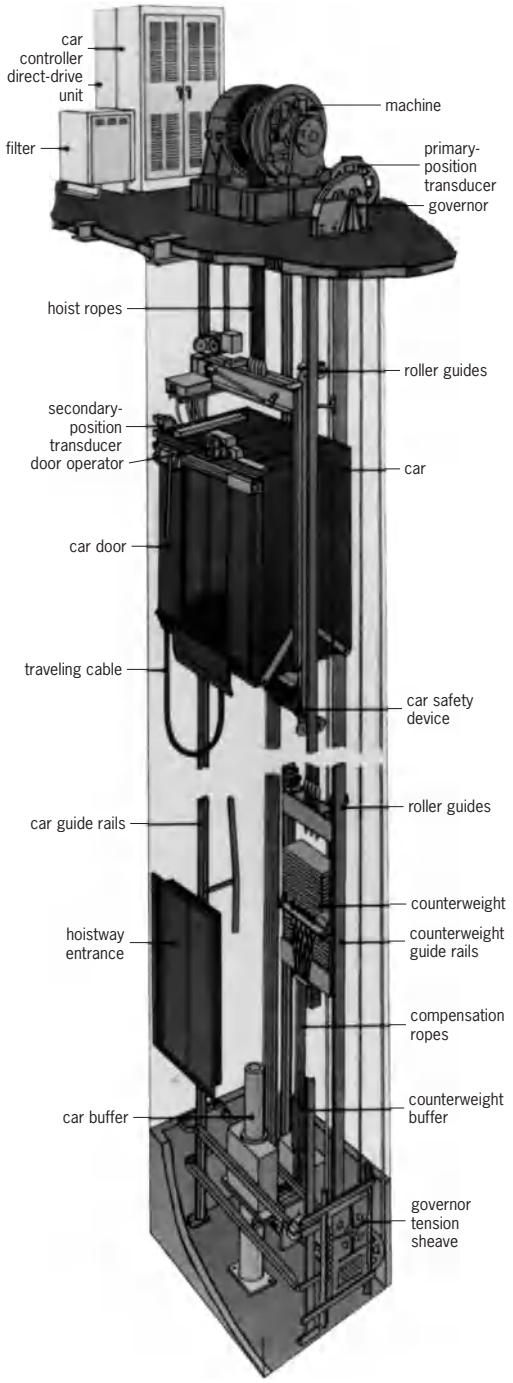

Typical installation of a traction elevator. (*Otis Elevator Co.*)

trips, releasing a clutching device that grips the governor rope. As the car continues its descent, the pull of the rope on the operating lever sets the safety. The safety, in turn, applies sufficient force against the guide rail to bring the car to a controlled stop with frictional force. *See* GOVERNOR.

Control devices are also built into the door and its control circuit. When the doors open, control circuits prevent the car from moving away from the landing, but permit releveling if the car moves as passengers enter or leave the elevator (load changes). Either a safety shoe or electronic detector, mounted on the doors, prevents the doors from continuing to close if a passenger or object is in the doorway. The safety shoe touches the person or object, retracts, and causes the door to reopen. Light-ray devices, often used with safety shoes, cause reversal of door movement whenever the light ray is broken by a passenger entering or leaving the car.

Elevator controller. Sophisticated elevator controllers began to replace human elevator operators in the 1950s. In modern elevators, microprocessor computer systems control elevator position, direction of travel, speed, door operation, passenger waiting time, flight time, energy consumption, and system diagnostics.

Systems are being developed that learn from past building traffic patterns and predict future patterns, assigning elevator cars to destinations in advance of actual demand and reducing passenger waiting times. In addition, in the future elevators will be equipped with laser devices that scan a floor for waiting passengers and signal the elevator accordingly to stop or continue. Modern elevators include Braille buttons and voice announcements of the floors to help the sight-impaired.

Types. There are three major types of elevators: gearless traction, geared traction, and hydraulic.

Gearless traction elevators are used in high-rise buildings over 10–12 stories, and travel at speeds from 400 to 2000 ft/min (120 to 610 m/min). They use large, slow-speed electric motors directly connected to a large, grooved drive sheave (pulley). The "hoist ropes" (steel cables) are attached at one end to the top of the elevator car, pass over the drive sheave and are attached at the other end to a counterweight that slides up and down the shaftway on its own guide rails. The full weight of the car and about half of its passenger load is balanced by the counterweight, which goes down as the car moves up. Thus, the electric motor does not have to lift the full weight of the car.

The principal difference between geared and gearless traction elevators is speed. A geared elevator usually travels at speeds from 25 to 450 ft/min (8 to 140 m/min), carries loads up to 30,000 lb (13,500 kg) or more, and uses a high-speed motor to drive the hoisting sheave through a gear reduction unit.

Hydraulic elevators are used extensively in low-rise buildings, usually up to five stories. With speeds rarely exceeding 150 ft/min (46 m/min), the hydraulic elevator does not need overhead hoisting machinery. The elevator is mounted on a piston inside a cylinder that extends into the ground to a depth equal to the height the elevator will rise. Relatively simple in design, the system uses an electric pump to force oil into the cylinder to give the elevator a controlled ascent. Electrically controlled valves release the oil for a controlled descent. Another form of hydraulic elevator is the so-called holeless model. A plunger that slides up and down on the side of the elevator is used, so that no hole is required beneath the standard shaftway space. Cynthia Di Tallo

Bibliography. A. Goetz (ed.), *Up, Down, Across: Elevators, Escalators, and Moving Sidewalks*, 2003; G. R. Strakosch (ed.), *The Vertical Transportation Handbook*, 3d ed., 1998.

Elevator (aircraft)

The hinged rear portion of the longitudinal stabilizing surface or tail plane of an aircraft used to obtain longitudinal- or pitch-control moments. The angular setting of the elevator is controlled by the human or automatic pilot through the flight-control system. A typical elevator control surface is shown in **Fig. 1**. Both leading-edge and horn-type aerodynamic balances and trailing-edge tabs are illustrated. These features reduce or eliminate the hinge moments required to deflect the elevator during flight. *See* FLIGHT CONTROLS.

General principles. The operating principles of elevator control surfaces are typical of all trailing-edge hinged control devices. Deflection of the elevator changes the camber of the entire surface. With the trailing edge down, high local flow velocities are obtained on the upper airfoil surface and low relative flow velocities are produced on the lower surface (**Fig. 2**). By Bernoulli's law, reduced pressure results on the upper surface, with increased pressures on the lower surface, creating lift. *See* BERNOULLI'S THEOREM.

The variation of elevator surface control effectiveness with percentage of total chord is shown in **Fig. 3**. Narrow-chord elevators are more effective than wide-chord elevator surfaces. Operating hinge moments vary as the square of the chord, and are thus greatly reduced on narrow-chord surfaces.

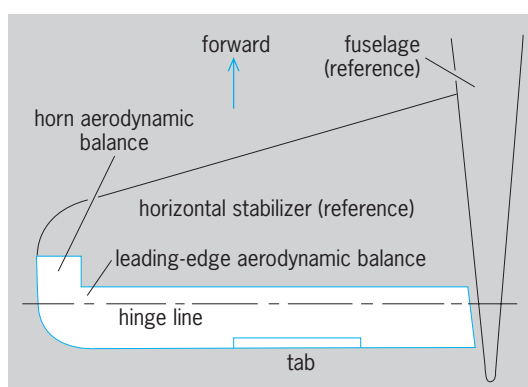

Fig. 1. Elevator control surface (left-hand side).

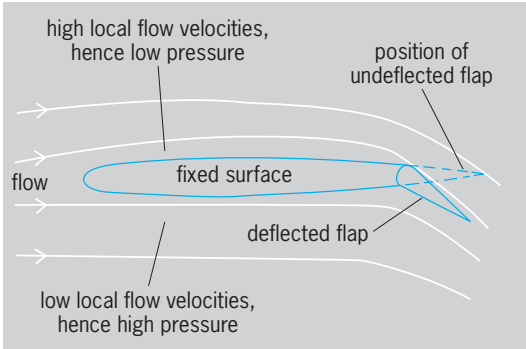

Fig. 2. Principle of operation of trailing-edge flap.

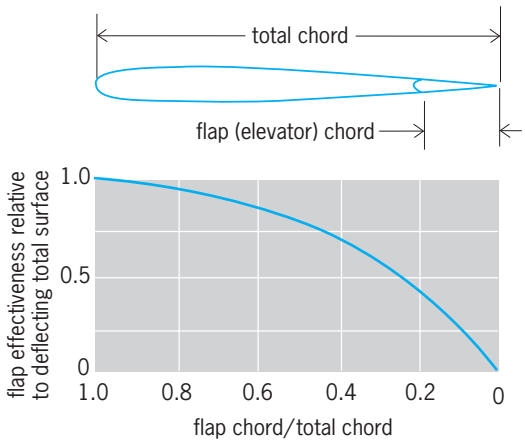

Fig. 3. Variation of trailing-edge flap effectiveness with flap chord to total chord ratio.

Flutter prevention. Elevator flutter is a divergent oscillation involving one or more degrees of freedom, such as rotation about the hinge line and flapping of the main surface. Hydraulic dampers, mass balancing, and structural stiffness are employed to prevent elevator flutter. *See* AEROELASTICITY; FLUTTER (AERONAUTICS).

Flight maneuvers. The elevator is used to perform pitching maneuvers, or maneuvers in which the aircraft's plane of symmetry is not disturbed. The maneuvers include airspeed adjustments and acceleration normal to the flight path (pull-ups or push-downs). The elevator also serves to adjust the aircraft's attitude with respect to the ground for takeoff and landing.

Special applications. All-moving horizontal stabilizers replace elevator control surfaces on supersonic aircraft and missiles to avoid large losses in effectiveness. The elevator may be geared to move at a fixed ratio to the deflections of an all-moving stabilizer. On aircraft without horizontal stabilizers (tailless aircraft), the elevator and aileron surfaces may be combined in surfaces that operate together as elevators and differentially as ailerons, for example, elevons. *See* ELEVON; STABILIZER (AIRCRAFT).

Malcolm J. Abzug

Bibliography. J. D. Anderson, *Introduction to Flight*, 4th ed., 1999; P. P. Wegener, *What Makes Airplanes Fly?*, 2d ed., 1996.

Elevon

A movable surface at the trailing edge of a tailless airplane that provides pitch and roll control. Elevons hinged on each side of the rear wing surface are shown in **Fig. 1**. They nose the airplane up or down, and roll one wing up and the other down. The term elevon is derived from elevator and aileron, and, in effect, elevons provide the same control as conventional elevators and ailerons. An example of an airplane employing elevons is the orbiter vehicle for the space shuttle. *See* AILERON; ELEVATOR (AIRCRAFT); SPACE SHUTTLE; TAIL ASSEMBLY.

The action of an elevon is illustrated in **Fig. 2**. A symmetrical airfoil is placed at an angle to the airflow, resulting in a lift force that acts approximately at a point located a quarter of the chord from the nose (Fig. 2*a*). The chord is defined as the distance from the nose, or leading edge, of the airfoil to the trailing edge. In Fig. 2*b* a trailing-edge flap on the same airfoil is deflected downward. This results not only in an increase in the lift on the airfoil but in a nose-down pitching moment about the quarter-chord point. Deflected upward, the flap produces a nose-upward pitching moment and a decrease in lift. *See* AIRFOIL.

When the elevons on both sides of the wing are deflected upward, they combine to produce a nose-up pitching moment on the wing and, hence, on the tailless airplane. If the elevon on the right wing is deflected upward and the one on the left wing downward, the differential movement does not result in any pitching moment, since the nose-up and nose-down pitching moments produced by the two elevons cancel each other. However, the lift on the

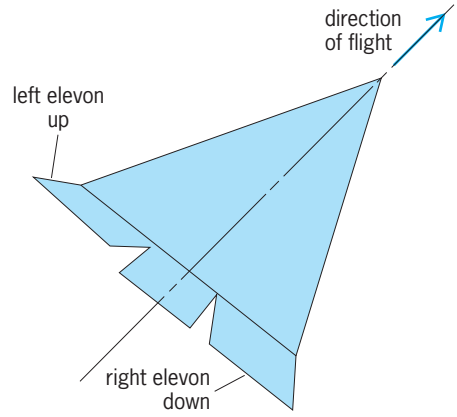

Fig. 1. Elevons on the trailing edge of a delta wing.

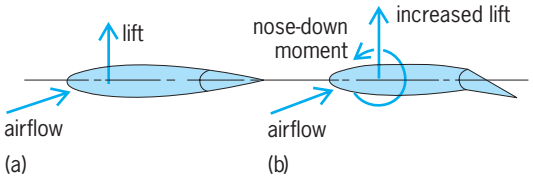

Fig. 2. Effect of flap on lift and moment of an airfoil at an angle of attack. (*a*) Symmetrical airfoil. (*b*) Symmetrical airfoil with flap deflected.

right wing decreases, and that on the left wing increases; the combined effect produces a moment on the airplane tending to roll it to the right. Deflecting the elevons in the opposite directions causes the wing to roll to the left, and deflecting both elevons downward produces a nose-down moment.

Pitching the nose up or down by elevon deflection changes the total lift on the wing, thereby accelerating the airplane vertically. Thus if the wing is in level flight, a pilot wishing to climb moves both elevons upward. Momentarily, the decreased lift causes the wing to drop, but the pitching moment quickly increases the angle of the wing relative to the airflow, the angle of attack; the result is a net increase of wing lift and an upward inclination of the flight path. *See* AIRPLANE; WING. Barnes W. McCormick, Jr.

Bibliography. J. D. Anderson, *Introduction to Flight*, 4th ed., 1999; B. Etkin and L. D. Reid, *Dynamics of Flight Stability and Control*, 3d ed., 1995.

Ellipse

A member of the class of curves that are intersections of a plane with a cone of revolution. The ellipse is obtained when the plane cuts all the elements of one nappe, and does not go through the apex. In the **illustration**, denote the distance between two points, F, F' of a plane by $2c$, $c > 0$, and let $2a$ be a constant, with $a > c$. The ellipse with foci F and F' and major axis $2a$ is the locus of points P of the plane such that $PF + PF' = 2a$, where PF denotes the distance of P and F. This suggests the following construction of an ellipse. Put pins at F and F', and slip over them a loop of thread of length $2a + 2c$, pulling the thread taut with a pencil. If the pencil is moved, keeping the thread taut, its point traces an ellipse. Another way to construct an ellipse is to drill a hole in a stick (at any point other than the midpoint) and move the stick so that its ends slide along two mutually perpendicular lines. The point of a pencil inserted in the hole will trace an ellipse. Limiting forms of the ellipse are (1) a circle, as the two foci approach coincidence, and (2) the segment FF', as c approaches a. If a circle is projected orthogonally on a plane not parallel to the plane of the circle, an ellipse is obtained, and every ellipse may be so obtained. Lines joining the foci to a point P of an ellipse make equal angles with the tangent to the ellipse at P, and consequently light or sound that emanates from one focus is reflected to the other focus. This property is used in construction

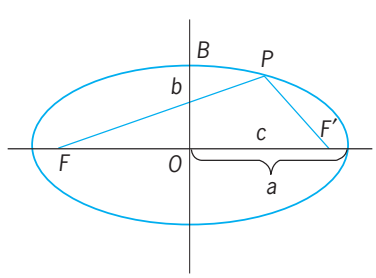

An ellipse; see text for explanation of symbols.

of "whispering galleries." *See* CONIC SECTION.

The midpoint of F, F' is the center O of the ellipse, and the chord through O perpendicular to the major axis is the minor axis, whose length is denoted by $2b$. If B is a point in which the minor axis intersects the ellipse, then $BF = BF' = a$, and so $c^2 = a^2 - b^2$. The ratio $c/a = \epsilon < 1$ is the eccentricity of the ellipse. If the half chords perpendicular to the major axis are multiplied by a/b, their extremities will lie on the circle whose diameter is the major axis. Hence, if Δ denotes the area of the ellipse, $\Delta \cdot a/b = \pi a^2$; that is, $\Delta = \pi ab$. The determination of the length L of an ellipse leads to the integral

$$\int_0^{\pi/2} [1 - \epsilon^2 \sin^2 \phi]^{1/2}\, d\phi$$

which is the complete elliptic integral of the second kind. It follows that

$$L = 2\pi a \left\{ 1 - (1/2)^2 \frac{\epsilon^2}{1} - \left(\frac{1 \cdot 3}{2 \cdot 4}\right)^2 \frac{\epsilon^2}{3} - \left(\frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\right)^2 \frac{\epsilon^2}{5} - \cdots \right\}$$

The volume bounded by the surface is obtained by revolving the ellipse with major axis $2a$ and minor axis $2b$ about its major axis $4\pi ab^2/3$. The area of the surface is

$$2\pi b^2 + 2\pi ab \frac{\sin^{-1} \epsilon}{\epsilon}$$

See ANALYTIC GEOMETRY; ELLIPTIC FUNCTION AND INTEGRAL. Leonard M. Blumenthal

Ellipsometry

A technique for determining the properties of a material from the characteristics of light reflected from its surface. The materials studied include thin films, semiconductors, metals, and liquids.

Principles. When an electromagnetic wave passes through a medium, it causes the electrons associated with the atoms of the medium to oscillate at the frequency of the wave. As a result, the wave is slowed so that its velocity v in the medium is less than its velocity c in empty space. Another result may be a transfer of energy from the wave to the electrons, thereby causing the amplitude of the wave to decrease as it penetrates into the material. These two processes are described phenomenologically by the complex refractive index, whose real part equals c/v and whose imaginary part equals $4\pi/\lambda$, where λ is the wavelength in vacuum. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES.

When an electromagnetic wave is incident on a medium, only part of it is transmitted into the medium (**Fig. 1**). The fraction reflected depends on the complex refractive index, the angle of incidence, and the polarization state of the wave. For multilayers with different complex refractive indices, the fraction also depends on the layer thicknesses. The two basic types of polarization are parallel, designated

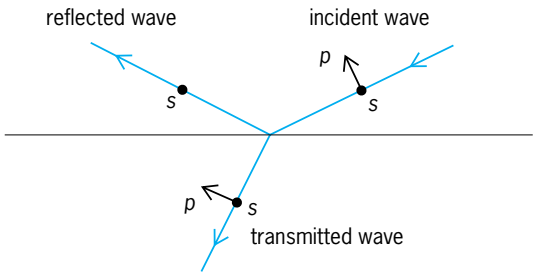

Fig. 1. Incident, transmitted, and reflected electromagnetic waves at the surface of a material. The angle of incidence shown is the Brewster angle, for which the *p*-polarized wave in the lower medium points along the reflection direction. The *s* wave is polarized perpendicular to the plane of the page.

p, and perpendicular, designated s. These terms refer to the orientation of the electric vector with respect to the plane of incidence, which is defined by the directions of the incident and reflected waves. The (intensity-independent) ratios of the amplitudes and phases of the reflected and incident p- and s-polarized electric fields are described by the complex reflectances r_p and r_s. *See* POLARIZATION OF WAVES; POLARIZED LIGHT; REFLECTION OF ELECTROMAGNETIC RADIATION.

The (also intensity-independent) ratio of the p- to the s-polarized component of such a wave is termed the polarization state. Simple examples include linear polarization, where the p- and s-polarized components are in phase, and circular polarization, where the p- and s-polarized amplitudes are equal but the phases differ by 90°. The geometric terms refer to the locus of the p and s (or y and x) components of the electric field when plotted in the complex plane (**Fig. 2**). The general polarization state is elliptical.

If an incident wave has electric-field components along both p and s, the components are independent during reflection and reassemble afterward with the changes in amplitude and phase described by r_p and r_s. That p- and s-polarized waves are affected differently upon reflection is readily apparent. The incident s wave causes the electrons in the medium to oscillate at right angles to the reflection direction for any angle of incidence (Fig. 1). These oscillating electrons generate a reflected s wave with its electric field transverse to the direction of propagation. However, an incident p wave causes the electrons in the medium to oscillate partly along the reflectance direction, a motion that does not contribute to the reflected wave. Thus the magnitude of the complex reflectance r_p is always less than that of r_s. In fact, for transparent materials there exists a particular angle of incidence, called Brewster's angle, where the oscillation is entirely parallel to the reflectance direction and the p-polarized component is suppressed completely.

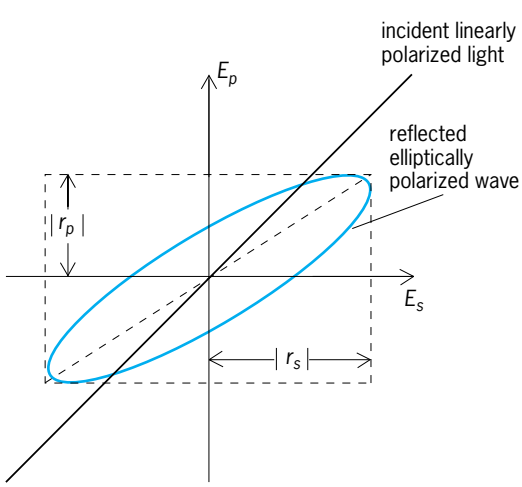

Fig. 2. Linearly polarized incident and elliptically polarized reflected waves. These represent the trajectories of the respective electric field vectors in the complex plane. E_p and E_s are the *p* and *s* components of the electric field; $|r_p|$ and $|r_s|$ are the magnitudes of the complex reflectances.

Since the complex reflectances r_p and r_s depend on the properties of the medium, the medium can be investigated by determining its reflectance for either p- or s-polarized light, that is, by determining the ratio of reflected and incident intensities $I_{\text{refl}}/I_{\text{inc}} = R = |r|^2$. This is the objective of reflectometry. Alternatively, because r_p and r_s are different, the complex reflectance ratio $\rho = r_p/r_s$, which is equal to the ratio of reflected and incident polarization states, can also be determined. This is the objective of ellipsometry. The ratio ρ is traditionally expressed in terms of angles ψ and Δ as in the equation below.

$$\rho = \frac{r_p}{r_s} = \tan\psi\, e^{i\Delta}$$

Because it deals with complex, intensity-independent quantities, an ellipsometric measurement is analogous to an impedance measurement. This gives ellipsometry certain advantages relative to reflectometry, such as higher accuracy and higher information content in a single measurement. A standard experimental approach, now used almost exclusively in spectroscopic applications, is to determine ρ by establishing a known state of polarization for the incident beam, for example, by passing it through a fixed linear polarizer, then determining the polarization state of the beam after reflection by passing it through a rotating linear polarizer, called an analyzer. The rotating analyzer essentially unrolls the polarization ellipse, allowing the azimuth angle of its major axis and its minor-major axis ratio to be determined by the phase and amplitude of the alternating-current component of the detected intensity. *See* ALTERNATING-CURRENT CIRCUIT THEORY; ELECTRICAL IMPEDANCE.

Applications. The primary application of ellipsometry is materials analysis, particularly the nondestructive analysis of thin films in semiconductor technology. Developing applications include the real-time monitoring and control of dynamic processes such as material deposition and etching. Spectroellipsometry, where the complex refractive index is measured and analyzed as a function of wavelength, has almost exclusively replaced reflectometry in materials analysis. Deposition and etching involve kinetic ellipsometry, where single-wavelength data are monitored as a function of time. *See* FILM (CHEMISTRY); INTEGRATED

CIRCUITS; NONDESTRUCTIVE EVALUATION; SEMICONDUCTOR HETEROSTRUCTURES.

Both applications rely on the fact that the complex refractive index depends directly on the ease by which electrons in a material can be set in motion by the incident electric field, which in turn depends directly on the properties of the material itself. Thus the optical properties of a metal such as chromium or aluminum, where the electrons may move freely, are quite different from those of an insulator such as plastic or glass, where the electrons are highly constrained. Further differences are observed if the material is crystalline or disordered or if two or more materials are mixed to form a composite. Other types of differences are observed in layered materials where the layers are thin enough to generate optical interference from back reflections within the sample. Such composites can be analyzed quantitatively for properties such as compositions and layer thicknesses, given an accurate database of optical responses for the pure materials. *See* AMORPHOUS SOLID; COMPOSITE MATERIAL. David E. Aspnes

Bibliography. F. Abeles (ed.), *Proceedings of the 1st International Conference on Spectroscopic Ellipsometry*, Paris, 1993: *Thin Solid Films*, vols. 233 and 234, 1993; R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light*, 1977, paper 1986; K. Riedling, *Ellipsometry for Industrial Applications*, 1988; H. Tompkins, *A User's Guide to Ellipsometry*, 1992.

Elliptic function and integral

In a certain sense, elliptic integrals are the simplest integrals not expressible in terms of elementary functions; elliptic functions arise as the inverse functions of certain elliptic integrals.

Let R be a rational function of x and y, and set $I = \int R(x,y)\,dx$. I can be expressed in terms of elementary functions if y^2 is a polynomial of degree 2 or less in x. If y^2 is a polynomial of degree 3 or 4 in x, I cannot in general be expressed in terms of elementary functions and is called an elliptic integral. (If y^2 is a polynomial of degree 5 or higher, I is called a hyperelliptic integral, and if y satisfies an algebraic equation whose coefficients are polynomials in x, I is called an abelian integral.) The standard elliptic functions are analogous to trigonometric functions. Trigonometric functions may be defined as the inverse functions of certain integrals of the form I; they satisfy differential equations, are periodic functions, and may alternatively be obtained as the "simplest" periodic functions. The standard elliptic functions are the inverse functions of certain elliptic integrals; they satisfy differential equations of order 1 and degree 2, are doubly periodic functions, and may alternatively be obtained as the "simplest" doubly periodic functions.

Applications. In geometry elliptic functions or integrals arise in determining the length of an arc of an ellipse, hyperbola, or lemniscate, the surface of an ellipsoid, geodesics on quadrics of revolution, parametric representations of plane cubic curves or, more generally, curves of genus 1, conformal representation problems, and other problems. In analysis there are applications to differential equations (Lamé's equation, diffusion equation, and others); in number theory, in a great variety of problems. In the physical sciences elliptic functions or integrals appear in potential theory both through conformal representations and in the potential of an ellipsoid, in the theory of elastica, the pendulum, in rigid body motion, in Green's functions in heat conduction and diffusion theory, and many other problems.

Reduction of elliptic integrals. By suitable homographic substitution, $x' = (ax + b)/(cx + d)$, $ad - bc \neq 0$, the elliptic integral I can be reduced to an elliptic integral in which the polynomial y^2 appears in normal form. The two customary normal forms are: Legendre's normal form, $y^2 = (1 - x^2)(1 - k^2x^2)$ where k, the modulus, is a real or complex number, $|k| \leqq 1$ and $k^2 \neq 1$, and it is usual to set $x = \sin\phi$; and the Weierstrass canonical form, $y^2 = 4x^3 - g_2x - g_3$ where g_2 and g_3, the invariants, are real or complex numbers. I can be expressed as a linear combination of an integral of rational functions and the elliptic integrals of the first, second, and third kinds defined in Legendre's normal form, respectively, as Eqs. (1)–(3), with $\Delta(t,k) = (1 - k^2\sin^2 t)^{1/2}$, and in the Weierstrass canonical form as expressions (4) with $y = (4x^3 - g_2x - g_3)^{1/2}$.

$$F(\phi, k) = \int_0^\phi dt/\Delta(t, k) \tag{1}$$

$$E(\phi, k) = \int_0^\phi \Delta(t, k)\,dt \tag{2}$$

$$\Pi(\phi, n, k) = \int_0^\phi \frac{dt}{(1 + n\sin^2 t)\Delta(t, k)} \tag{3}$$

$$\int \frac{dx}{y} \int \frac{x\,dx}{y} \int \frac{dx}{(x - c)y} \tag{4}$$

In Legendre's theory, $\mathbf{K} = \mathbf{K}(k) = f(\pi/2,k)$ and $\mathbf{E} = \mathbf{E}(k) = E(\pi/2,k)$ are called the complete elliptic integrals of the first and second kinds, respectively, $k' = (1 - k^2)^{1/2}$ is the complementary modulus, and $\mathbf{K}' = \mathbf{K}'(k) = f(\pi/2,k')$, $\mathbf{E}' = \mathbf{E}'(k) = \mathbf{E}(\pi/2,k')$. Complete elliptic integrals as functions of k satisfy linear differential equations of the second order and are hypergeometric functions of k^2. They also satisfy Legendre's relation, $\mathbf{KE}' + \mathbf{K}'\mathbf{E} - \mathbf{KK}' = \pi/2$ identically in k.

Periods and singularities. Elliptic integrals are many-valued functions. Any two determinations of I differ by a sum of integral multiples of certain real or complex numbers, the periods. E, F, and Π are many-valued functions of the complex variable $x = \sin\phi$. All three functions have branch points at $x = \pm 1$, $\pm k^{-1}$, and Π has branch points also at $x = \pm in^{-1/2}$. The periods of F are $4\mathbf{K}$ and $2i\mathbf{K}'$, those of E are $4\mathbf{E}$ and $2i(\mathbf{K}' - \mathbf{E}')$. Since the complete elliptic integrals are real when $0 \leqq k \leqq 1$, the first (second) of these periods is called the real (imaginary) period. Although E and F are many-valued functions of x, E is a

TABLE 1. Properties of jacobian elliptic functions

Function	Primitive periods	Zeros	Poles
sn u	$4\mathbf{K}$, $2i\mathbf{K}'$	$2m\mathbf{K} + 2ni\mathbf{K}'$	
cn u	$4\mathbf{K}$, $2\mathbf{K} + 2i\mathbf{K}'$	$(2m+1)\mathbf{K} + 2ni\mathbf{K}'$	$2m\mathbf{K} + (2n+1)i\mathbf{K}'$
dn u	$2\mathbf{K}$, $4i\mathbf{K}'$	$(2m+1)\mathbf{K} + (2n+1)i\mathbf{K}'$	

single-valued function of F provided that corresponding values of E and F are obtained by integration over the same path and using the same determination of $\Delta(t,k)$.

Inversion of elliptic integrals. Jacobian elliptic functions are determined by inversion of the functional relation $u = f(\theta,k)$. It is usual to write Eqs. (5).

$$\begin{aligned} \phi &= \text{am } u = \text{am } (u, k) \\ \sin\phi &= \text{sn } u = \text{sn } (u, k) \\ \cos\phi &= \text{cn } u = \text{cn } (u, k) \\ \Delta\ (\phi, k) &= \text{dn } u = \text{dn } (u, k) \end{aligned} \tag{5}$$

With the additional conditions sn $0 =$ cn $0 =$ dn $0 = 1$, it turns out that sn u, cn u, dn u are single-valued analytic functions of the complex variable u, and that they are doubly periodic and also regular except for simple poles. Nine additional functions are introduced by the notations $1/\text{pn } u = \text{np } u$, pn $u/\text{qn } u = \text{pq } u$ where p and q stand for any of the letters s, c, d. Similarly, the Weierstrass $\mathcal{P}$-function is introduced by writing relation, Eq. (6), in the form

$$z = \int_{\infty}^{w} (4t^3 - g_2 t - g_3)^{-1/2}\, dt \tag{6}$$

$w = \mathcal{P}(z) = \mathcal{P}(z; g_2, g_3)$, and $\mathcal{P}(z)$ turns out to be single-valued, doubly periodic, and regular except for double poles.

Doubly periodic functions. The term p is called a period of a single-valued analytic function $f(z)$, regular except for isolated singularities, if $f(z+p) = f(z)$. A nonconstant periodic function is either simply periodic, when all its periods are integral multiples of a single period, or else doubly periodic, when its periods are $2m\omega + 2n\omega'$ where m and n are integers, 2ω and $2\omega'$ are primitive periods, and $\text{Im } \omega'/\omega > 0$. Two points of the z plane are congruent if they differ by a period. A parallelogram in the z plane is a period parallelogram if every point in the plane is congruent to exactly one point of the parallelogram. If no singularity or zero of $f(z)$ lies on the boundary of the period parallelogram, the parallelogram is called a cell. An elliptic function is a doubly periodic function which is regular except for poles. An elliptic function has a finite number of poles in every cell; the sum of the residues at these poles is zero, and the sum of the orders of these poles is the order of the function.

Every elliptic function of order 0 is a constant. There is no elliptic function of order 1. An elliptic function of order $r > 1$ assumes in every cell each complex value r times (counting multiplicity). The difference of two elliptic functions with the same periods, same poles, and the same principal parts at each pole is a constant. The quotient of two elliptic functions with the same periods, poles, and zeros (including multiplicities) is a constant. All elliptic functions with a given pair of primitive periods form an algebraic field, and any two such functions are connected by an algebraic relation. Every elliptic function satisfies an algebraic differential equation of the first order. Every elliptic function possesses an algebraic addition theorem, that is, an algebraic relation connecting $f(u)$, $f(v)$, and $f(u+v)$. Conversely, any single-valued analytic function of z which is regular except for poles and possesses an algebraic addition theorem is either a rational function of z, or a rational function of e^{az} for some a, or else an elliptic function.

The simplest nontrivial elliptic functions are those of order 2. Choice of a basic function with two simple poles in a cell leads to Jacobi's functions, choice of a function with a double pole, to Weierstrass's functions.

Jacobian elliptic functions. Write $s = \text{sn } u$, $c = \text{cn } u$, $d = \text{dn } u$, $s' = ds/du$, and so on. the periods zeros, and poles are given in **Table 1** in which m and n are integers. These functions possess symmetry properties around the points $u = 0, \mathbf{K}, i\mathbf{K}'$, which are set out in **Table 2** and on account of which it is sufficient to study the functions in the parallelogram whose vertices are 0, $\mathbf{K}$, $\mathbf{K} + i\mathbf{K}'$, and $i\mathbf{K}'$. There is a very large number of identities for these functions. Some of the basic ones are shown in Eqs. (7). The addition theorem for s is given in Eq. (8). There are similar addition

$$\begin{aligned} &s^2 + c^2 = 1,\ k^2 s^2 + d^2 = 1,\ d^2 - k^2 c^2 = k'^2 \\ &s' = cd,\ c' = -sd,\ d' = -k^2 sc, \\ &\qquad s'^2 = (1 - s^2)(1 - k^2 s^2), \text{ etc.} \\ &\text{sn } (iu, k) = i \text{ sc } (u, k'),\ \text{cn } (iu, k) = \text{nc } (u, k'), \\ &\qquad \text{dn } (iu, k) = \text{dc } (u, k') \end{aligned} \tag{7}$$

$$\text{sn } (u + v) = \frac{\text{sn } u \text{ cn } v \text{ dn } v + \text{sn } v \text{ cn } u \text{ dn } u}{1 - k^2 \text{ sn}^2 u \text{ sn}^2 v} \tag{8}$$

theorems for c and d. These, in combination with the formulas for sn (iu), for example, serve to express sn $(u + iv)$ in terms of elliptic functions of u and v they provide formulas for sn $(2u)$, sn $(u/2)$, and so on. By means of these formulas the values of the jacobian functions at the points $m\mathbf{K}/2 + ni\mathbf{K}'/2$

TABLE 2. Symmetries of jacobian elliptic functions

v	$-u$	$2\mathbf{K} - u$	$2i\mathbf{K}' - u$
sn v	$-$sn u	sn u	$-$sn u
cn v	cn u	$-$cn u	$-$cn u
dn v	dn u	dn u	$-$dn u

TABLE 3. Degenerate elliptic functions

K	K′	k	k'	sn u	cn u	dn u
∞	$\pi/2$	1	0	tanh u	sech u	sech u
$\pi/2$	∞	0	1	sin u	cos u	1
∞	∞			0	1	1

(m, n integers) and at the points $m\mathbf{K} + in\mathbf{K}' + u$ can be obtained.

The elliptic functions reduce to elementary functions if one or both of the periods become infinite (degenerate elliptic functions, see **Table 3**).

Weierstrass's functions. With $w = 2m\omega + 2n\omega'$, sums and products running over all pairs of integers m, n except $m = n = 0$, there are these functions: Weierstrass sigma function, Eq. (9), which is

$$\sigma(z) = z\Pi\left\{\left(1 - \frac{z}{w}\right)\exp\left[\frac{z}{w} + \frac{1}{2}\left(\frac{z}{w}\right)^2\right]\right\} \tag{9}$$

an entire function; Weierstrass zeta function, $\zeta(z) = \sigma'(z)/\sigma(z)$, which is a meromorphic function; and Weierstrass $\mathcal{P}$-function, $\mathcal{P}(z) = \zeta'(z)$, which is an elliptic function of order 2 with double poles at $z = 0$ and congruent points. The invariants are $g_2 = 60\Sigma w^{-4}$ and $g_3 = 140\Sigma w^{-6}$. Legendre's relation becomes $\omega'\zeta(\omega) - \omega\zeta(\omega') = \pi i/2$. The $\mathcal{P}$-function satisfies the differential equation, Eq. (10), and possesses the addition theorem given as Eq. (11). It is a

$$\mathcal{P}'^2(z) = 4\mathcal{P}^3(z) - g_2\mathcal{P}(z) - g_3 \tag{10}$$

$$\mathcal{P}(u+v) = \frac{1}{4}\left[\frac{\mathcal{P}'(u) - \mathcal{P}'(v)}{\mathcal{P}(u) - \mathcal{P}(v)}\right]^2 - \mathcal{P}(u) - \mathcal{P}(v) \tag{11}$$

homogeneous function of degree -2 in z, ω, ω'. Every elliptic function with primitive periods 2ω, $2\omega'$ can be expressed in the form $R_1[\mathcal{P}(z)] + \mathcal{P}'(z)R_2[\mathcal{P}(z)]$, where $R_1(w)$ and $R_2(w)$ are rational functions of w; and there are also representations in terms of zeta and sigma functions. Degenerate cases lead to elementary functions.

Theta functions. The function in Eq. (12), with a

$$\theta(v|\tau) = \sum_{r=-\infty}^{\infty} e^{i\pi r^2\tau + i\pi r(2v+1)} \tag{12}$$

fixed τ and Im $\tau > 0$, is an even entire function of v. It has period 1, it is multiplied by $e^{-i\tau(2v+\pi)}$ when v is increased by τ, and it has simple zeros at the points $v = m + (n + 1/2)\tau$ (m, n integers). It is usual to consider four theta functions, Eqs. (13). $\theta(x/2, i\pi t)$

$$\begin{aligned}
\theta_1(v) &= -ie^{i\pi(v+\tau/4)}\theta\left(v + \frac{\tau}{2}\right)\\
\theta_2(v) &= e^{i\pi(v+\tau/4)}\theta\left(v + \frac{1+\tau}{2}\right)\\
\theta_3(v) &= \theta(v + 1/2)\\
\theta_4(v) &= \theta(v)
\end{aligned} \tag{13}$$

satisfies the partial differential equation $\partial^2 y/\partial x^2 = \partial y/\partial t$ and has a simple Laplace transform. Elliptic functions and elliptic integrals can be expressed in terms of theta functions: $\tau = \omega'/\omega$ in the case of Weierstrass functions and $\tau = i\mathbf{K}'/\mathbf{K}$ in the case of jacobian functions or Legendre's normal form of elliptic integrals.

Transformation theory. The set of periods of an elliptic function may be described by various pairs of primitive periods. The change from one pair of primitive periods to another pair is called a transformation of the elliptic function or integral. The quotient of the primitive periods, τ, undergoes a homographic substitution, $\tau = (a\tau + b)/(c\tau + d)$, where a, b, c, d are integers and $D = ad - bc$ is positive and is called the degree of the transformation. All transformations of degree 1 form the modular group. The study of these transformations is of great theoretical interest, has applications to number theory, and is also used for the numerical computation of elliptic functions. It is also connected with the study of elliptic modular functions, or analytic functions, $f(\tau)$, with the property that $f(\tau)$ and $f(\dot{\tau})$ are algebraically connected whenever τ and $\dot{\tau}$ are connected by a transformation of the modular group. *See* FOURIER SERIES AND TRANSFORMS. A. Erdelyi

Bibliography. M. Abramowitz and I. A. Stegun (eds.), *Handbook of Mathematical Functions*, 10th ed., 1972; N. Akhiezer, *Elements of the Theory of Elliptic Functions*, 1990; P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals for Engineers and Scientists*, 2d ed., 1971; K. Chandrasekharan, *Elliptic Functions*, 1985; A. Erdélyi et al., *Higher Transcendental Functions*, vol. 2, 1953, reprint 1981; S. Lang, *Elliptic Functions*, 2d ed., 1987; D. F. Lawden, *Elliptic Functions and Applications*, 1989.

Elm

Any species of *Ulmus*, a genus of hardwood trees in the Northern Hemisphere, with simple, serrate, deciduous leaves. The American or white elm (*U. americana*) is the most important species (**illus.** *a*). It is a large, typically vase-shaped tree which usually grows to about 80 ft (24 m) but may reach a height of 140 ft (42 m). The leaves are unequal at the base and doubly serrate; that is, the leaf teeth are themselves toothed. The winter buds are brown, scaly, and usually at one side of the leaf scar. The fruit, ripening in late spring, is a small, flat, winged, elliptical samara, fringed with hairs on the edge.

It ranges from the eastern half of the United States westward as far as the base of the northern Rockies and southward through central Texas to the Gulf of Mexico. The tree is also found in southern Canada.

The tree is common in moist soil, usually at low elevations. It is very popular as a shade tree and is perhaps better adapted as a street tree than any other species because the upper branches spread joining with elms across the street to form an arch.

The wood is heavy, hard, and difficult to split. It is used for wagon parts, barrel staves, shipbuilding, furniture, flooring, sporting goods, boxes, and baskets,

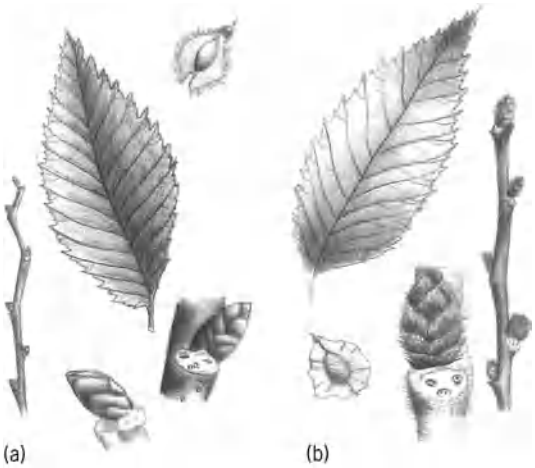

Diagnostic features of elm species. (*a*) Twig, leaf, fruit, and terminal and lateral buds of American elm (*Ulmus americana*). (*b*) Leaf, fruit, lateral bud, and twig of the slippery elm (*U. rubra*).

and to some extent, veneer and wall paneling.

Once abundant, the elm is being severely attacked by the lethal Dutch elm disease imported from Europe. Shade trees are particularly susceptible and have been virtually eliminated in many towns and cities. The future of the species is uncertain. *See* PLANT PATHOLOGY.

Slippery elm (*U. rubra*), a smaller tree with larger, rough leaves (illus. *b*) and mucilaginous inner bark, is commercially important in the east-central United States. The uses of the wood are similar to those of American elm.

Rock or cork elm (*U. thomasi*) of the northeastern United States is not so common. It can be recognized by its usually corky branches, small, spherical winter buds, and flowers and fruit in elongated clusters. The wood is similar in character and use to that of the American elm. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Elopiformes

An order of teleost fishes in the superorder Elopomorpha. All elopomorphs share the leptocephalus-type larvae. Elopiforms are further characterized by a slender body with abdominal pelvic fins and deeply forked caudal fin supported by seven hypurals (flattened bones along the ventral side of the urostyle); a large mouth, terminal or superior, bordered by premaxillae and toothed maxillae; teeth also present on parasphenoid and mesopterygoid bones, and tongue; mesocoracoid and postcleithra bones and a well-developed gular plate; cycloid scales; and small leptocephali, about 5 cm (2 in.) in length, with well-developed, forked caudal fin and strong teeth. Eel leptocephali lack a caudal fin (**Fig. 1**). There are two families of elopiforms, Elopidae and Megalopidae.

Elopidae (ladyfishes/tenpounders). Ladyfishes, also known as tenpounders, are elongate with a slightly compressed body, a large terminal mouth, large pseudobranchiae (small gill-like structures, developed on the underside of the gill cover), and no swim bladder (**Fig. 2**). The last ray of the dorsal fin is not elongate, the scales are small, and the lateral-line scales have unbranched tubes (pores). Members of the family are primarily coastal marine species in tropical and subtropical oceans, rarely entering brackish or freshwater. Maximum length is 1 m (3.3 ft). This family consists of one genus and about six species.

Megalopidae (tarpons). Tarpons have a rather strongly compressed body; a large terminal or superior mouth; no pseudobranchiae; and a highly vascular, lunglike swim bladder (**Fig. 3**). The last ray of the dorsal fin is elongate, the scales are very large, and the lateral-line scales have branched tubes. Tarpons occur in tropical and subtropical oceans, often entering freshwater, especially as juveniles. Maximum length is about 2.4 m (7.9 ft). There are two species: *Megalops cyprinoides* of the Indo-West Pacific and *M. atlanticus* (*Tarpon atlanticus*) of the western Atlantic and tropical seas off West Africa.

Classification. The leptocephalus larval stage is a commonality of the Elopiformes, Albuliformes, Anguilliformes, and Saccopharyngiformes, groups that comprise the Elopomorpha (Fig. 1). The elopomorphs as well as all the remaining extant teleosts are thought to have descended from amiiforms. In

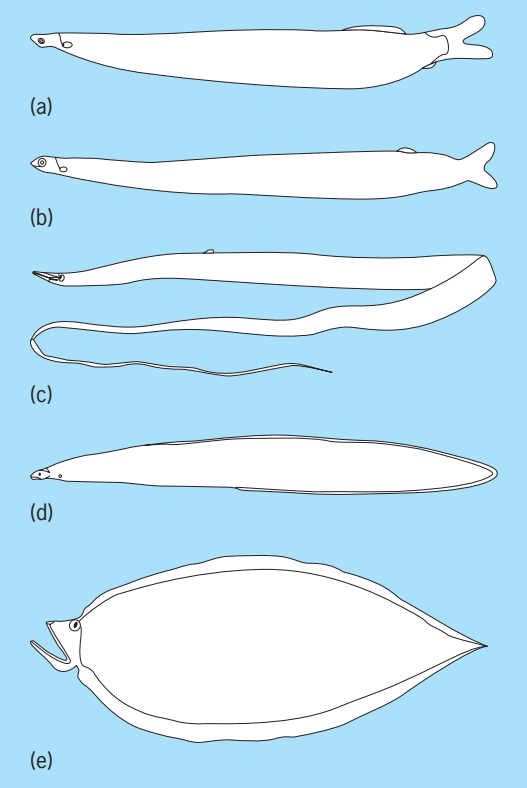

Fig. 1. Representative leptocephalus larvae of the major groups of Elopomorpha. Typical maximum length is in brackets. (*a*) Elopiformes [5 cm (2 in.)]; (*b*) Albuliformes [6 cm (2.4 in.)]; (*c*) Notacanthoidei [2 cm (0.8 in.)]; (*d*) Anguilliformes [12 cm (4.8 in.)]; and (*e*) Saccopharyngiformes [5 cm (2 in.)]. (*Courtesy of J. S. Nelson, Fishes of the World, Wiley, 2006*)

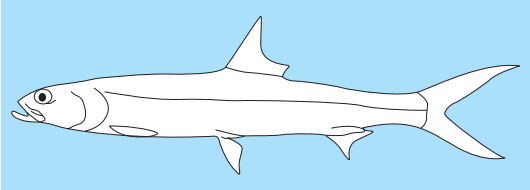

Fig. 2. Example from the family Elopidae (ladyfishes, or tenpounders). (*Courtesy of J. S. Nelson, Fishes of the World, Wiley, 2006*)

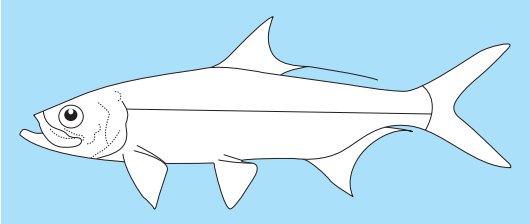

Fig. 3. Example from the family Megalopidae (tarpons). (*Courtesy of J. S. Nelson, Fishes of the World, Wiley, 2006*)

turn, albuliforms, anguilliforms, and saccopharyngiforms are believed to be descendants of elopiforms. The former order Notacanthiformes (spiny eels and halosaurs), which has leptocephalus larvae, is now classified in the order Albuliformes as the suborder Notacanthoidei. Some systemists reject the notion that monophyly (referring to a taxon that includes all the evolutionary descendants of the taxon's common ancestor and only those descendants) of the Elopomorpha is justified because the leptocephalus larval condition is believed to be a primitive rather than a derived condition. *See* ALBULIFORMES; ANGUILLIFORMES; SACCOPHARYNGIFORMES; TELEOSTEI.

Herbert Boschung

Bibliography. E. B. Böhlke (ed.), *Fishes of the Western North Atlantic*, Part 9, vol. 2: *Leptocephali*, pp. 657–1055, Sears Foundation for Marine Research, Memoir (Yale University), New Haven, 1989; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; C. H. Wang et al., Molecular phylogeny of elopomorph fishes inferred from mitochondrial 12S ribosomal RNA sequences, *Zool. Scr.*, 32:231–241, 2003.

Embedded systems

Computer systems that cannot be programmed by the user because they are preprogrammed for a specific task and are buried within the equipment they serve. The term derives from the military, where computer systems are generally activated by the flip of a switch or the push of a button. For example, a military aircraft pilot may wish to turn on the countermeasures equipment with a switch. There is no need for the pilot to be involved with the computer. The same holds true for a soldier who may direct a ground-to-ground missile against a target tank by the push of a button. In both cases, an embedded computer quickly goes to work.

Embedded systems are also used in industrial, automotive, consumer, and medical applications. Specific machine-tool routines can often be initiated by the machine operator by a switch or a button. The cruise-control feature in automobiles, which is activated by a front-panel or steering-column switch, is the result of an embedded microcomputer. Automatic scoring and unattended bowling lanes depend on the operation of embedded microcomputers. Many types of intensive-care and health-care equipment often operate unattended because of embedded computer systems.

The continual increase in the densities of ever-smaller microprocessors, the brains and core of microcomputers, on silicon chips that fit on a thumbnail, and the attendant decreases in their costs, has pushed the concept of embedded systems well beyond the original military applications. Embedded microprocessors provide value-added features to a wide range of products. They are used in consumer electronics, automotive vehicles, telecommunications, computer peripherals, process control, factory automation, robotics, and intelligent instruments. Microwave ovens, stereo equipment, washer-dryers, refrigerators, blenders, videocassette recorders, and even pressing irons use embedded microcomputers based on small microprocessors. Engine and drive-train controllers are two of the largest automotive applications. Embedded systems are also widely used in computer printers and modems. The first microprocessor chip, designed in 1971, can be considered an embedded computer, since it was originally created for an electronic calculator application.

CISC and RISC architectures. Most embedded microprocessors are of the CISC (complex-instruction-set computer) type, and most of these are used in applications where low cost is paramount and performance is secondary, such as consumer products. The later-generation microprocessors have wider bus widths, up to 64 bits, and thus can do more computations.

Since about 1990, microprocessors of the RISC (reduced-instruction-set computer) type have appeared, with much greater computational capability and at greater cost. RISC processors are used mostly in those embedded applications where performance is primary and low cost is secondary, such as in the military. They are also used in engineering workstations, where the computational burdens of high-resolution graphics require such processors. *See* COMPUTER-AIDED ENGINEERING; COMPUTER GRAPHICS; COMPUTER SYSTEMS ARCHITECTURE.

Software development. It is generally more difficult to develop software for RISC processors than for CISC processors. Indeed, modern high-performance processors, whether RISC or CISC, require considerably more software development than their predecessors, and placing such processors in embedded applications further complicates the task of software development. The chief issue in the design of embedded processors is now the nature of the end application rather than the architecture of the hardware, as

had been the case in earlier-generation devices. *See* INTEGRATED CIRCUITS; MICROCOMPUTER; MICROPROCESSOR; SOFTWARE ENGINEERING. Roger Allan

Bibliography. A. Berger, *Embedded Systems Design: An Introduction to Processes, Tools and Techniques*, 2001; Q. Li and C. Yao, *Real-Time Concepts for Embedded Systems*, 2003; T. Noergaard, *Embedded Systems Architecture: A Comprehensive Guide for Engineers and Programmers*, 2005; R. Zurawski (ed.), *Embedded Systems Handbook*, 2005.

Embioptera

A peculiar order of silk-spinning, orthopteroid insects related to termites, commonly called the embiids or web spinners. This order comprises about 1000 species which are chiefly tropical in distribution. The body is linear and supple. Adults vary in length from 0.12 to 0.8 in. (3 to 20 mm). The legs are short with three-segmented tarsi. The forelegs are adapted for spinning silk, and the hindlegs for reverse locomotion. Cerci are short, two-segmented, and tactile. The head is prognathus, with orthopteroid mouthparts. Metamorphosis is incomplete. The females are neoteinic and wingless (apterous) [**Fig. 1**]. Males are usually winged (alate), but in certain genera and species they are apterous. The wings are subequal and elongate, with the vannal fold obsolete. Venation is simple, with the veins centered in pigment bands and separated by hyaline stripes. The wings are flexible when in repose and folded over the back, but are stiffened when extended for flight by the blood pressure in the saclike R_1 (radial) veins. Flight is a poorly directed, whirling flutter.

Plantar surfaces of the mid and basal foretarsal segments bear many hollow, silk-spinning setae, each connected by a duct to a globular, syncytial silk gland; many such glands are massed in the greatly swollen basal segment. Hundreds of silk strands are emitted in unison as the tarsi brush a surface. A labyrinth of silk galleries is produced rapidly, which constitutes a safe shelter for all embiid activities except adult dispersal (**Fig. 2**). All embiids spin silk regardless of the species, developmental stage, or sex.

The galleries radiate on or in the food supply which also constitutes the habitat and consists of bark, lichens, moss, dead leaves, or grass. Many individuals, usually the brood of one female, may occupy one gallery system. *See* INSECTA; ORTHOPTERA. Edward S. Ross

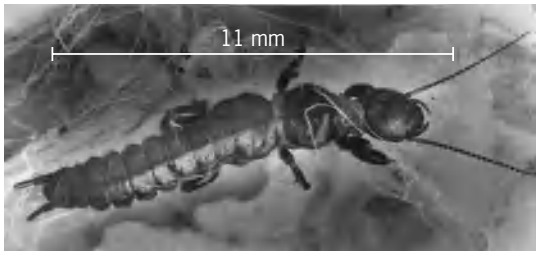

Fig. 1. Body form of a typical embiid (*Pararhagadochir trachelia*, female). (*From E. S. Ross, Insects Close Up, University of California Press, 1953*)

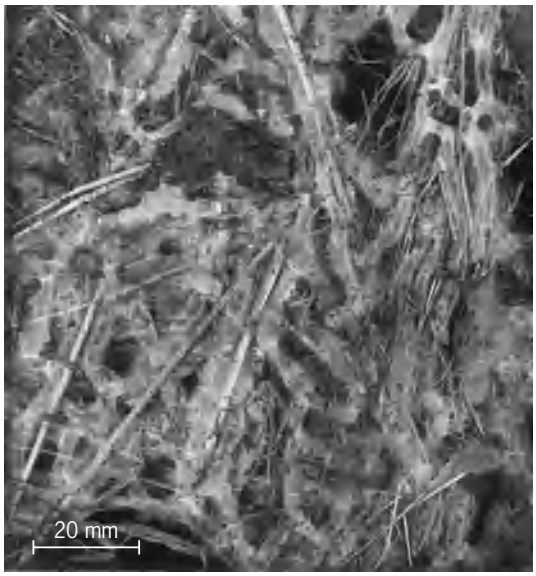

Fig. 2. Characteristic silk gallery system of the embiid *Haploembia solieri*.

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms,* 2 vols., 1982; E. S. Ross, The Embioptera of Europe and the Mediterranean region, *Bull. Brit. Mus. Entomol.*, 17:275–326, 1966; E. S. Ross, A revision of the Embioptera or web-spinners of the New World, *Proc. U.S. Nat. Mus.,* 94:401–504, 1940.

Embolism

The sudden blocking of an artery or vein by a clot or other substance which has been brought to its place by the blood current. The material carried in the circulation in this process is an embolus. Emboli may be composed of thrombi, fat, air, tumor cells, masses of bacteria or parasites, bone marrow, amniotic fluid, or atheromatous material from the vessel wall. A thrombus which has formed in the heart or one of the vessels is the usual form of embolus (see **illus.**). Emboli from the right side of the heart or the great venous system of the body come to lodge in the lungs; those from the portal system, in the liver; and those from the pulmonary veins or left side of the heart, in some segment of the peripheral arterial tree. Pulmonary emboli can result in infarction of the lung. However, if they are large enough to occlude the main pulmonary artery or one of the major branches, the individual may die of shock.

Embolization is a common method of spread of tumor cells, which is one reason for the frequency of metastatic tumors in the liver and lungs.

Injury to bone or adipose tissue may result in fat embolism. Following its escape from the injured fat cell, the fat gains access to the venous system. It may then come to rest in the capillaries of the lung, or it may pass through the lungs to find its way to some other tissue such as the brain or kidney.

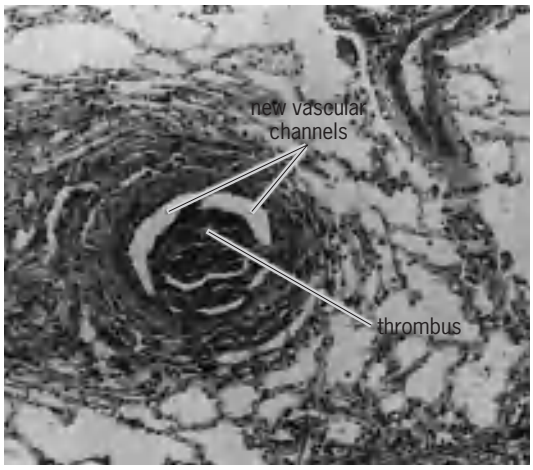

Thrombus, a type of embolus, appearing in the pulmonary artery of a dog's lung, with the subsequent formation of new vascular channels.

Air embolism may be a complication of a surgical procedure, particularly those about the neck. It can also result from rapid decompression, as in caisson disease, with the formation of bubbles of nitrogen in the blood. Infected emboli can form new foci of infection at their sites of lodgement. *See* DECOMPRESSION ILLNESS.

With complete obstruction of a vessel by an embolus an infarct may result. *See* INFARCTION.

Romeo A. Vidone

Bibliography. S. Z. Goldhaber, *Pulmonary Embolism and Deep Venous Thrombosis*, 1985; J. Hirsh et al., *Venous Thromboembolism*, 1981; G. Holm and M. Bjorkholm (eds.), *Biology of Vascular Disease*, 1987; O. Horwitz et al. (eds.), *Diseases of the Blood Vessels*, 1985; R. C. Little, *Physiology of the Heart and Circulation*, 4th ed., 1988.

Embrithopoda

An order established for the unique mammal *Arsinoitherium*, which has been found only in early Oligocene deposits in northern Egypt. This animal was of rhinoceros size with a large body and short pillarlike legs (see **illus.**). There were two huge, scimitarlike horns over the nose and two much smaller, peglike horns over the eyes. As in many herbivorous mammals, the cusps of the high-crowned teeth were connected to form two transverse lophs on each tooth. The exact relationship of this exotic order to other orders is not clear, but similarities to both the uintatheres, or Dinocerata, and Hyracoidea have been interpreted by some authorities as evidence of kinship. *See* DINOCERATA; MAMMALIA.

Malcolm C. McKenna

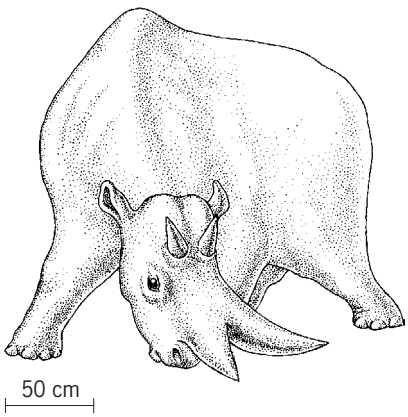

***Arsinoitherium*, the early Oligocene embrithopod from Egypt.**

Embrittlement

A general set of phenomena whereby materials suffer a marked decrease in their ability to deform (loss of ductility) or in their ability to absorb energy during fracture (loss of toughness), with little change in other mechanical properties, such as strength and hardness. Embrittlement can be induced by a variety of external or internal factors, for example, (1) a decreasing or an increasing temperature; (2) changes in the internal structure of the material, namely, changes in crystallite (grain) size, or in the presence and distribution of alloying elements and second-phase particles; (3) the introduction of an environment which is often, but not necessarily, corrosive in nature; (4) an increasing rate of application of load or extension; and (5) the presence of surface notches. This article is restricted to metals and alloys, and describes embrittlement in terms of the above factors. *See* BRITTLENESS.

Effect of test temperature. In body-centered cubic metals (for example, iron, tungsten) and hexagonal close-packed metals (for example, zinc, magnesium), a critical temperature exists below which the metal exhibits limited toughness. Fracture is usually brittle in nature, occurring either through the crystal lattice (cleavage) or along the boundaries separating the crystallites or grains (intergranular fracture). In simple terms, low-temperature embrittlement results from a competition between deformation and brittle fracture, with the latter becoming preferred at a critical temperature. For a material to be useful structurally, it is desirable that this critical temperature be below the minimum anticipated service temperature; in most cases, this is room temperature. At high temperatures, internal structural changes that lead to intergranular embrittlement can occur. Embrittlement usually occurs in the range of creep temperature at which deformation can occur under very low stresses; and the two processes are believed to be connected. *See* CRYSTAL STRUCTURE; PLASTIC DEFORMATION OF METAL.

Effect of heat treatment. In many metals, particularly structural steels, annealing or heat treating in certain temperature ranges sensitizes the grain boundaries in such a way that intergranular embrittlement subsequently occurs during service. In steels, for example, high strength is achieved by quenching from the stable high-temperature face-centered cubic phase (called austenite). This can

produce a room-temperature structure called martensite, a distorted, body-centered cubic phase which is extremely hard and strong, but brittle. The transformation of austenite to martensite is controlled by several metallurgical variables, including the alloy composition and the quench rate. To reduce the brittleness, the steel undergoes an annealing treatment called tempering, which, while decreasing the strength, usually increases the toughness. The exception to this trade-off occurs when the steel is tempered at 1000°F (538°C). This can lead to a mode of intergranular fracture called temper embrittlement; such a process has led to catastrophic failures in turbines, rotors, and other high-strength steel parts. It is associated with the preferential segregation of undesirable "tramp" elements such as tin and antimony to the grain boundaries; these elements should be kept to as low a level as possible to help alleviate the problem. Tempering at 550°F (287°C) can also lead to embrittlement in steels. In other metals, there are less specific but similar types of embrittlement resulting from critical heat treatments. *See* HEAT TREATMENT (METALLURGY); STEEL; TEMPERING.

Effect of environment. Metals can fracture catastrophically when exposed to a variety of environments. These environments can range from liquid metals to aqueous and nonaqueous solutions to gases such as hydrogen. The phenomenology of these processes and some of the corrective procedures used or envisaged are described below.

Liquid metal embrittlement. If a thin film of a liquid metal is placed on the oxide-free surface of a solid metal, the tensile properties of the solid metal will not be affected, but the fracture behavior can be markedly different from that observed in air. In some cases, specimens stressed above a critical value will fracture catastrophically. Although the embrittlement is usually intergranular, transgranular cleavage has also been observed. Such fractures can be induced in highly ductile metals such as aluminum and copper. Although many different liquid metals are capable of inducing embrittlement in a variety of solid metals, some of the more common couples, many of which have important engineering and design consequences, are mercury embrittlement of brass (an alloy of copper and zinc), lead embrittlement of steel, and gallium embrittlement of aluminum. The mechanism of liquid metal embrittlement is thought to result from the reduction by the liquid metal atoms of the force necessary to break apart two solid metal atoms at a crack tip.

Stress-corrosion cracking. If a metal is stressed and simultaneously exposed to an environment which may be, but is not necessarily, corrosive in nature, cracking and fracture can occur. Both stress and environment are required; if only one of these elements is present, the metal usually displays no embrittlement. *See* CORROSION.

The analogy of stress-corrosion cracking to liquid metal embrittlement is obvious, but in this case, the environment can be any aqueous or nonaqueous solution. In specific metals cracking has been observed in such diverse solutions as high-purity distilled water, salt water, caustics (for example, lye), ammonia, molten anhydrous salts, and organics. Stress-corrosion cracking can occur over a wide temperature range, and can be a very serious problem in many service applications. For example, brass cartridge cases which were internally stressed by a forming operation cracked in storage because of the presence of trace amounts of ammonia in the air; and titanium alloys used in aerospace applications cracked when cleaned with methyl alcohol. There are a host of situations in which moisture or salt water led to cracks or total fracture of structural materials.

A mechanism which has been proposed as an explanation for many of these cases has been associated with the stress-assisted preferential dissolution of metal along specific metallurgical features, leading to intergranular or transgranular cracks. This process can be likened to the action of a battery. In this case, the cracking region is the anode; when the anode is coupled to a cathodic region, the circuit is completed. This electrolytic behavior has been exploited as a protective device. Critical parts are made to serve as cathodes in local cells, allowing other noncritical parts to act as sacrificial anodes. This procedure, termed cathodic protection, is used to protect not only against stress-corrosion cracking but also against general corrosion; for example, zinc sacrificial anodes are used to minimize corrosion of ship propellers. Other mechanisms, such as corrosion, have also been put forward to explain the stress-corrosion cracking phenomena. *See* ELECTROCHEMICAL PROCESS.

Hydrogen embrittlement. This form of embrittlement is often considered to be a type of stress-corrosion cracking. Hydrogen atoms can enter a metal, causing severe embrittlement, again with little effect on other mechanical properties. This phenomenon was originally observed, and is most critical in, steels, but it is now documented to occur in titanium and nickel alloys, and may lead to cracking in other alloy systems as well.

The source of hydrogen can be extremely varied, making this type of embrittlement difficult to control. For example, hydrogen can be retained internally during the melting and casting of alloys; it can be discharged at cathodic areas in electrolytic cells (it is apparent that, at both the anode and the cathode, events which may lead to embrittlement occur); it can enter the metal during a plating operation; and it can come from an external molecular gas environment, even at very low partial pressures of hydrogen. When it is realized that coal gasification plants produce hydrogen or methane (a hydrogen-containing gas) as a fuel, that long-distance transport of hydrogen is envisaged as a means of alleviating the energy crisis, that liquid hydrogen is used as a rocket fuel, and that hydrogen is a by-product of many chemical reactions, it becomes apparent that structural materials must be resistant to hydrogen embrittlement. *See* COAL GASIFICATION; HYDROGEN.

In all forms of environmental embrittlement, the best remedial practice is to physically isolate the

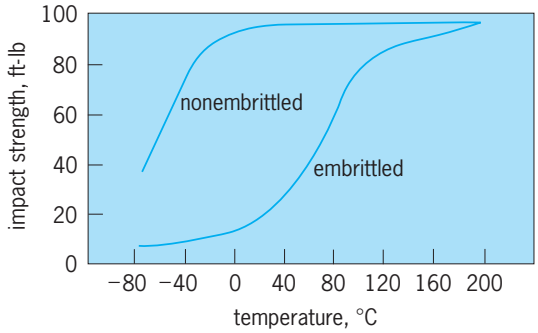

Effect of temper brittleness on the impact resistance of a low-alloy steel. 1 ft-lb = 1.36 J. °F = (°C × 1.8) + 32. (*After R. E. Reed-Hill, Physical Metallurgy Principles, Van Nostrand, 1973*)

metal from the dangerous environment. This is not always practical or possible. Studies have indicated that the susceptibility of a given material to environmental embrittlement can be reduced by controlling metallurgical factors, such as alloy composition and internal structure. This type of alloy design should help control environmental embrittlement, even where there is uncertainty as to the controlling mechanism involved in the process. *See* ALLOY.

Other factors. Factors such as notches (which both raise and modify the local stress state) and the rate of application of stress can modify the response of a material to a specific type of embrittlement. In general, notches or surface flaws always enhance embrittlement, both by acting as a stress raiser and by providing a preexisting crack. An increasing loading rate (impact) enhances low-temperature embrittlement, and possibly thermal embrittlement, but, interestingly, suppresses environmental embrittlement. For environmental embrittlement, species such as hydrogen must keep up with the moving crack, and they do so by moving through the crystal lattice at a rate that is largely determined by the temperature. If the crack is moving too quickly, as under impact conditions, hydrogen can not keep pace, and the severity of embrittlement decreases. For normal rates of loading, the maximum effect of hydrogen is unfortunately at, or near, ambient temperatures. The combined effects of various embrittlement phenomena are shown in the **illustration**. Temper embrittlement increases the impact ductile-to-brittle transition temperature from about −58 to 176°F (−50 to 80°C), making this steel unacceptable for room-temperature applications. *See* BRITTLENESS.

I. M. Bernstein

Bibliography. Z. Ahmad, *Principles of Corrosion Engineering and Corrosion*, 2006; N. E. Dowling, *Mechanical Behavior of Materials: Engineering Methods for Deformation, Fracture, and Fatigue*, 2d ed., 1998; M. G. Fontanta, *Corrosion Engineering*, 3d ed., 1986; G. Y. Lai, *High Temperature Corrosion of Engineering Alloys*, 1990; P. R. Roberge, *Handbook of Corrosion Engineering*, 1999; H. H. Uhlig and R. W. Revie, *Corrosion and Corrosion Control*, 3d ed., 1985.

Embryobionta

One of the two plant subkingdoms, the other being the Thallobionta. The Embryobionta (often also called Embryophyta) form a well-marked group with many morphological, anatomical, and physiological features in common, and taxonomists are agreed that the Embryobionta as a group are derived from the division Chlorophyta (green algae) in the subkingdom Thallobionta. The Embryobionta are here considered to include eight divisions, the Rhyniophyta, Bryophyta, Psilotophyta, Lycophyta, Spenoophyta, Polypodiophyta, Pinophyta, and Magnoliophyta. The Rhyniophyta are represented only by Paleozoic fossils, but the other seven divisions have both modern and fossil representatives. See the separate articles on each division. *See* CHLOROPHYCOTA.

Characteristics. The pigments of the chloroplasts of the Embryobionta are chlorophylls *a* and *b*, β- and usually α-carotene, and several xanthophylls, generally including lutein, cryptoxanthin, and zeaxanthin. Chlorophyll *a* is the principal chlorophyll, β-carotene is the principal carotene, and lutein is the principal xanthophyll. The carotenes and xanthophylls, collectively called carotenoid pigments, are masked by the more abundant chlorophylls. The most common carbohydrate reserve of the Embryobionta is starch. The primary cell wall consists largely of cellulose. Flagellated cells, when they are present (for example, the sperms of several divisions), have only the whiplash type of flagellum. In all these respects the Embryobionta resemble the green algae. *See* CAROTENOID; CHLOROPHYLL.

Life cycle. The Embryobionta differ from the green algae and from most Thallobionta in that the normal life cycle of the Embryobionta shows a well-marked alternation of generations in which the sporophyte (spore-producing, typically diploid) generation always begins as a parasite on the gametophyte (gamete-producing, typically haploid) generation. The young sporophyte is called an embryo. The only Thallobionta which have a definite sporophyte attached to the gametophyte are some of the red algae, and these are so different from the Embryobionta in other respects that no close relationship seems possible.

Reproductive structures. The more primitive divisions of Embryobionta have the gametes produced in multicellular sex organs (archegonia and antheridia), which is in contrast to the unicellular oogonia and antheridia of the Thallobionta in general. In the more advanced divisions (Pinophyta and Magnoliophyta) of Embryobionta, the antheridia and archegonia are highly modified or entirely suppressed, in conformity with the general reduction of the gametophyte generation.

Specializations. All divisions of Embryobionta except the Bryophyta have specialized conducting tissues (xylem and phloem) in the sporophyte. With the exception of most bryophytes, they also commonly have a characteristic stomatal apparatus which controls the opening and closing of numerous tiny pores (stomates) in the leaves and stems in response to

environmental conditions. These specializations, together with the progressive reduction of the gametophyte in the more advanced divisions, reflect the progressive evolutionary adaptation of the Embryobionta to life on dry land instead of in the ancestral water. The Embryobionta are therefore often called the land plants, in spite of the fact that many of them, such as the water lilies, have returned to an aquatic habitat. The seven divisions of Embryobionta which characteristically have xylem and phloem have sometimes been treated as a single comprehensive division under the name Tracheophyta. *See* PLANT KINGDOM; THALLOBIONTA.

Arthur Cronquist

Bibliography. A. Cronquist, *Introductory Botany*, 1961; A. Cronquist, A. Takhtajan, and W. Zimmerman, *Taxon*, 15:129–134, 1966.

Embryogenesis

The formation of an embryo from a fertilized ovum, or zygote. Development begins when the zygote, originating from the fusion of male and female gametes, enters a period of cellular proliferation, or cleavage. Cells of the embryo subsequently give rise to the tissues and organs of the body in a temporal and spatial pattern that creates a functional, multicellular organism. *See* CLEAVAGE (EMBRYOLOGY).

Animals

The animal zygote divides by mitosis to produce cells called blastomeres. Since there is little or no growth during cleavage, the blastomeres become smaller with each succeeding division, and the embryo maintains the relative size and shape of the zygote. In a cleaving cell, the spindle and asters of the mitotic apparatus are essential for division of the nucleus. The mitotic apparatus also determines the plane for division of the cytoplasm. A cleavage furrow arises on the cell surface along a line that is equidistant from the paired asters. As this furrow deepens, it gradually partitions the cytoplasm between the daughter blastomeres. The mitotic apparatus tends to situate in the central cytoplasm and orient parallel to the long dimension of a cleaving cell. The long dimension can, of course, change from one cleavage to the next, shifting the orientation of the mitotic apparatus and the plane of the cytoplasmic division. If cytoplasmic constituents, such as yolk granules, prevent the mitotic apparatus from assuming its customary position, the cleavage furrow will unequally distribute the cytoplasm, making the daughter blastomeres dissimilar in size and composition. *See* BLASTULATION; CELL DIVISION; MITOSIS.

Following cleavage, the cells of the animal embryo rearrange into three germ layers: an outer ectoderm, a middle mesoderm, and an inner endoderm. Cells, responding to intrinsic and extrinsic factors, eventually segregate from the germ layers and organize into the rudiments of the tissues (epithelia, connective tissues, muscles, and nerves) and the organs of the body. These rudiments alter both the size and the shape of the embryo. Cellular growth and differentiation are the principal processes that transform the rudiments into functional tissues and organs. Once the organs and organ systems are formed, further development consists primarily of growth. *See* ANIMAL MORPHOGENESIS; GERM LAYERS.

Michael J. Cavey

Plants. The zygote in higher plants develops into a multicellular embryo with apical meristems that ultimately produce the adult plant. Embryogenesis in both plants and animals involves a combination of growth, morphogenesis, and cellular differentiation. Significant differences exist between patterns of embryogenesis in higher plants and animals: (1) cell movement does not play a major role in plant embryogenesis because plant cells are surrounded by cell walls; (2) programmed cell death does not occur during plant embryogenesis; (3) plant embryos do not contain a germ line that functions only in the formation of gametes; and (4) the mature plant embryo does not resemble the adult because after germination, groups of undifferentiated cells known as meristems continue to produce in an accretionary manner the cells, tissues, and organs that make up the adult plant. *See* CELL WALLS (PLANT).

Major features of embryogenesis in flowering plants include the formation of root and shoot apical meristems; differentiation of primary vascular tissue; the transition from a heterotrophic zygote to an embryo capable of independent growth and development; and preparations for desiccation, dormancy, and germination, all of which appear to be associated with differential gene expression. Embryogenesis in flowering plants is accompanied by development of the endosperm tissue produced after double fertilization involving the fusion of a second male gamete with additional cells of the female gametophyte located within the immature ovule. This nutritive endosperm tissue may be consumed during embryogenesis or be included as part of the dormant seed. The endosperm does not appear to serve a regulatory function during the final stages of embryogenesis, because immature embryos from a wide range of plants have been removed from developing seeds and cultured to maturity on a simple nutrient medium. This process of embryo culture has been used to study the regulation of embryogenesis and to rescue aborted embryos produced by interspecific hybridization.

Under appropriate conditions somatic cells of higher plants can produce embryos in culture through a process known as somatic embryogenesis. Differentiated cells from a wide range of plants have been shown to produce somatic embryos that resemble zygotic embryos in both morphology and patterns of gene expression. Somatic embryogenesis has been used to demonstrate the totipotency of plant cells and to produce a large number of genetically similar plants, which have been used in fundamental research and hold promise for crop improvement. *See* APICAL MERISTEM; CELL DIFFERENTIATION; DEVELOPMENTAL BIOLOGY; EMBRYOLOGY; EMBRYONIC DIFFERENTIATION; EMBRYONIC INDUCTION.

David Meinke

Bibliography. L. W. Browder et al., *Developmental Biology*, 3d ed., 1997; S. F. Gilbert, *Developmental Biology*, 8th ed., 2006; R. Grossinger, *Embryogenesis*, 1986; V. Raghavan, *Embryogenesis in Angiosperms: A Developmental and Experimental Study*, 1986.

Embryology

The study of the development of an organism, commencing with the union of male and female gametes. Embryology literally means the study of embryos, but this definition is restrictive. An embryo is an immature organism contained within the coverings of an egg or within the body of the mother. Strictly speaking, the embryonic period ends at metamorphosis, hatching, or birth. Since developmental processes continue beyond these events, the scope of embryology is customarily broadened to encompass the entire life history of an organism. Embryology may, in this wider context, consider the mechanisms of both asexual reproduction and regeneration.

Animals

Development of a multicellular organism is a continuum of processes. It is useful, however, to discuss development from the perspective of several overlapping phases, such as gametogenesis, fertilization, cleavage, gastrulation, and histogenesis and organogenesis.

Gametogenesis. The production of male and female gametes occurs in the gonads. The differentiating gametes arise from diploid stem cells. Cell division by meiosis reduces the number of chromosomes carried by a mature gamete to one-half that present in the stem cell. Following meiosis, the male gamete achieves a significant reduction in weight by compacting its nucleus and shedding excess cytoplasm, and it also acquires the organelles needed to approach and fuse with the female gamete.

Specializations of the female gamete take place during, rather than after, meiosis. The female gamete often enters a protracted period of growth, accumulating the yolk that will meet the early nutritional needs of the embryo. A protective covering, or envelope, is deposited around the female gamete before it leaves the gonad, and other envelopes may be added along the female reproductive tract. The envelope(s) and the gamete constitute the egg. *See* GAMETOGENESIS; MEIOSIS; OOGENESIS; OVUM; SPERMATOGENESIS.

Fertilization. The union of spermatozoon and ovum creates a diploid zygote with the potential to form an entire organism. Two events must occur for successful fertilization: the ovum must respond to contact with the spermatozoon by making preparations for further development, a process called activation; and the haploid nucleus of the spermatozoon must combine with the haploid nucleus of the ovum, a process called amphimixis.

The zygotes of different species can be very dissimilar in size, shape, and composition. Because the spermatozoon contributes negligible cytoplasm to the zygote, these dissimilarities arise from the ovum. The cytoplasmic elements of an ovum are frequently localized in specific regions, and they are subsequently redistributed in consistent patterns during fertilization. Compositional differences, from one side of a zygote to the other and from the surface of a zygote to the interior, can be profound; disturbances of this organization can lead to serious developmental abnormalities.

Fertilization is the typical method to initiate development, but it is not the only way. In a few animals, the ovum develops independently by parthenogenesis, that is, without the participation of a spermatozoon. Parthenogenesis can sometimes be induced experimentally in organisms that would ordinarily opt for fertilization. Several physical and chemical agents are known to mimic the action of a spermatozoon in triggering the activation of an ovum. In such cases of artificial parthenogenesis, the ability of an ovum to develop rests largely on the response of its haploid nucleus. A normal sequence of developmental events is possible if the nucleus can, in some aberrant manner, become diploid. *See* FERTILIZATION.

Cleavage. A period of cell proliferation converts the unicellular zygote into a multicellular embryo. Cleavage is a modified form of cell division by mitosis, distinguished by little or no growth between the divisions. The cells of the embryo, or blastomeres, become smaller at the end of each division, so the embryo maintains the relative size and shape of the zygote. Small, fluid-filled spaces form between the cleaving blastomeres, and these spaces eventually coalesce to create an internal cavity, or blastocoele. Upon the appearance of a blastocoele, the cells of an embryo are referred to collectively as the blastoderm.

There is overwhelming evidence that the cytoplasm, not the nucleus, directs cleavage. The earliest demonstration of this control was achieved with subdivided ova, or merogones, produced by centrifugation. One merogone receives the haploid nucleus of the ovum, and its development can be initiated by fertilization. The development of the other merogone, which lacks a nucleus, must be initiated parthenogenetically. Both merogones show similar patterns of cleavage, implying cytoplasmic control over the process. The enucleated merogone cannot pass from cleavage to the next phase of development, suggesting that nuclear control is necessary in making the transition.

The isolation of blastomeres from cleaving embryos has also provided insights into early development. In amphibians, for example, the two blastomeres formed by the first cleavage can be separated and reared individually. Each blastomere continues to develop and produces a small but perfectly normal larva. The same experiment on the four blastomeres formed by the second cleavage gives different results. Two of the blastomeres develop into normal larvae, and two develop abnormally, forming larvae with conspicuous defects. Since the nuclei of the early blastomeres are identical, cytoplasmic

differences must specify whether a blastomere has the capacity to form an entire individual or only certain parts of that individual. Cleavage is, therefore, more than a period of rapid cell proliferation; it is also a time when cytoplasmic differences begin to guide the embryonic cells into specific pathways of development. *See* BLASTULATION; CLEAVAGE (EMBRYOLOGY); MEROGONY.

Gastrulation. Gastrulation is poorly delineated from cleavage because the cells of the embryo continue to divide. Gastrulation is distinguished from cleavage by extensive cell rearrangements that lead in most animals to the establishment of three germ layers: an outer ectoderm, a middle mesoderm, and an inner endoderm. Endodermal and mesodermal cells of the blastoderm migrate to the inside of the embryo, while ectodermal cells remain on the surface, where they spread to completely cover the body. Mapping experiments, using carbon particles, vital dyes, or radioisotopes to mark small areas of the blastoderm, illustrate that the cell movements during gastrulation are executed in a precise, orderly fashion. *See* FATE MAPS (EMBRYOLOGY).

Control of development passes from the cytoplasm to the nucleus immediately prior to gastrulation. Responding to cytoplasmic cues, the nuclei begin to specify the production of proteins that make the cells qualitatively different from one another. In a few invertebrates, the transfer of control from cytoplasm to nucleus actually fixes the developmental fate of a cell. In most other organisms and particularly in vertebrates, the determination of cell fate is not finalized until the blastoderm has rearranged into the three germ layers. *See* CELL LINEAGE; GASTRULATION; GERM LAYERS.

Histogenesis and organogenesis. The organization of cells into tissues (histogenesis) and tissues into organs (organogenesis) is closely allied with gastrulation. During creation of the germ layers, blastodermal rearrangements shift cells into new positions and bring about new intercellular relationships. To a considerable degree, the developmental fate of a cell can be the consequence of its new position. The influence exerted by one group of cells over the developmental fate of a neighboring group is called induction. Induction occurs by the transmission of chemical substances known as inducing agents.

An inducing agent presumably elicits a cytoplasmic change that, in turn, modifies the output of the nucleus. As cells specialize by accumulating different kinds of proteins, they begin to segregate from the germ layers and organize into four rudimentary tissues. These tissues, comprising epithelia, connective tissues, muscles, and nerves, subsequently combine to establish the rudiments of the organs. The spatial distribution of the organ rudiments alters both the size and shape of the embryo. Inductions between the rudimentary tissues are often required for conversion of the nascent organs into functional units. Growth and terminal differentiation are the principal processes in this conversion. Growth is possible in several ways: by augmenting the size of individual cells, by increasing the total number of cells, or by accumulating materials in the intercellular spaces. Terminal differentiation is dependent upon the acquisition of all necessary proteins by which a cell maintains its structure and function; this acquisition is manifest by changes in cell number, size, shape, motility, and adhesion. *See* EMBRYOGENESIS; EMBRYONIC INDUCTION.

Differentiation, determination, and morphogenesis. Differentiation, the process by which a cell becomes specialized, correlates to a reduction in the amount of genetic information that is expressed. Determination, the fixation of a developmental fate, occurs when a cell has such a limited amount of usable genetic information that it must commit to a terminal pathway of differentiation. *See* CELL DIFFERENTIATION.

Differentiation and determination are deliberate processes set in motion at the time of fertilization. The heterogeneous cytoplasm of a zygote, when distributed among the blastomeres by cleavage, is undoubtedly responsible for initial constraints placed upon the nuclei. If the nucleus of a blastodermal cell is transplanted to an enucleated zygote, development proceeds normally, indicating that constraints may limit the functions of a nucleus but do not alter its potential. The early restrictions on utilization of genetic information account for the establishment of the three germ layers and the origins of the lowly differentiated ectodermal, mesodermal, and endodermal cells. Further restrictions are imposed partly by the composition of the cytoplasm and partly by the influence of inducing agents. Inductions between the cells of the germ layers give rise to the moderately differentiated cells of the embryonic tissues, and subsequent inductions between the rudimentary tissues provide the highly (and terminally) differentiated cells of the functional organs. *See* GENE; GENE ACTION.

Cellular differentiation is just one aspect of morphogenesis, or the development of form. Morphogenesis must be considered at all levels of organization, ranging from the individual cell to the whole organism. Such a broad perspective complicates the formulation of general theories of development. Presently, no comprehensive theory exists, but there are some embryologists who anticipate that a theory is possible once activities of the DNA molecule have been fully integrated into the topic of development. *See* ANIMAL MORPHOGENESIS; REPRODUCTION (ANIMAL).

Michael J. Cavey

Plants

Reproductive development in multicellular plants is generally divided into three phases: gametogenesis, fertilization, and embryogenesis. The zygote produced by the fusion of male and female gametes divides to form a multicellular embryo with meristematic regions that ultimately produce the adult plant. Many different patterns of reproductive development have been described in plants; emphasis here will be placed on the most common pathways of gametogenesis and embryogenesis in angiosperms.

Gametogenesis. Development of the cell in flowering plants begins with a diploid megasporocyte located within the nucellar tissue of an immature ovule. This megasporocyte undergoes meiosis to form a tetrad of four haploid megaspores. In the most common pattern of development, three of these megaspores degenerate, leaving a single functional megaspore that undergoes several postmeiotic mitoses to form a mature megagametophyte (embryo sac) composed of seven cells and eight haploid nuclei. One of these haploid cells is the egg cell. Variations on this common pattern of development include differences in the number of functional megaspores and the number and arrangement of haploid nuclei present in the mature megagametophyte.

Development of the male gametes begins with numerous diploid cells (microsporocytes) located within the anthers of an immature flower. Each microsporocyte undergoes meiosis to form a tetrad of four haploid microspores, which then separate and enlarge to form mature pollen grains. Each microspore divides unequally to form a large vegetative cell, and a small generative cell located within the cytoplasm of the vegetative cell. The generative cell divides again, in either the maturing pollen grain or the elongating pollen tube, to form two genetically identical male gametes, the sperm cells.

Fertilization and embryogenesis. The zygote is produced as part of a unique process known as double fertilization. One of the male gametes fuses with the egg cell to form the diploid zygote, while the other male gamete fuses with two polar nuclei, located near the center of the embryo sac, to form a triploid endosperm nucleus. Following double fertilization, the zygote develops into an embryo composed of two parts, the embryo proper and the suspensor. The embryo proper ultimately differentiates into the mature embryo, whereas the suspensor degenerates during later stages of development and is not usually present at maturity. The suspensor appears to play both a passive role in attaching the embryo proper to surrounding maternal tissues, and an active role in regulating development of the embryo proper during early stages of embryogenesis. The triploid endosperm nucleus in most angiosperms undergoes a series of free nuclear divisions to form a nutritive endosperm tissue that usually becomes cellular during later stages of development.

Flowering plants can be divided into two groups, monocots and dicots. In most dicots, the endosperm tissue is gradually absorbed by the developing embryo and is not present in the mature seed. Nutrients required for the germination of dicot seeds are generally stored in the embryonic leaves known as cotyledons. In contrast, most mature monocot seeds contain a significant amount of starchy endosperm tissue that serves as a source of nutrients for the germinating seedling.

Two important regions of the mature embryo are the root and the shoot apical meristems. The entire shoot system (stems, leaves, and flowers) of the adult plant forms from cells that are located in the shoot apical meristem of the mature embryo. In some species of angiosperms, this embryonic meristem is composed of only a few cells, whereas in other species several true leaves are produced by the shoot apical meristem before the completion of embryogenesis. The root apical meristem that is formed during embryogenesis becomes active during the early stages of germination and ultimately produces the entire root system of the adult plant. *See* APICAL MERISTEM; ROOT (BOTANY).

The final stages of embryogenesis in angiosperms include maturation, desiccation, and preparation for seed dormancy. Lipids and storage proteins are produced at high levels during embryonic maturation, stored in membrane-bound vesicles, and utilized as a source of nutrients for the young seedling. Immature embryos are often capable of precocious germination in culture but become dehydrated within the fruit and usually germinate only when exposed to the appropriate environmental conditions. Release from seed dormancy may require abrasion of the seed coat, exposure to light, or cold treatment, in addition to sufficient moisture. *See* EMBRYOGENESIS.

Different patterns of embryo development are found in gymnosperms and in the more primitive vascular and nonvascular plants. Double fertilization and the development of a nutritive endosperm tissue are features unique to the angiosperms. The haploid microgametophyte (germinating pollen grain) in most gymnosperms contains two male gametes, but only one of these participates in fertilization. The nutritive function served by the endosperm tissue in angiosperms is served in gymnosperms by the large haploid megagametophyte. Early divisions of the zygote are also different in gymnosperms; the zygote typically undergoes a series of free nuclear divisions during the earliest stages of embryogenesis, and multiple embryos often arise from a single zygote through a process known as polyembryony. Even more striking differences in embryogenesis are found in ferns and mosses, where the haploid or gametophytic phase of the life cycle is much more extensive.

Several major differences also exist between embryogenesis in plants and animals. Plant cells are surrounded by a cell wall that limits the contact and movement between adjacent cells. Embryogenesis in plants therefore proceeds without the morphogenetic movements that are characteristic of animal development. Morphogenesis in plants is also not limited to embryo development, but occurs throughout the life cycle. The mature plant embryo is therefore not simply a miniature version of the adult plant. *See* PLANT MORPHOGENESIS.

Methods of study. Embryogenesis in plants has been approached through a combination of morphological, experimental, biochemical, and, more recently, genetic and molecular studies. Most of the basic anatomical features of reproductive development in angiosperms were elucidated through the observations of botanists in the nineteenth and early twentieth centuries. Some of the most detailed histological studies have dealt with *Capsella bursapastoris* (Cruciferae). Early stages of embryo development in *Capsella* are characterized by rapid growth of the endosperm tissue and the formation of a

proembryo which is composed of a filamentous suspensor and a small embryo proper. The suspensor is metabolically active during early stages of embryogenesis and reaches a mature size of 6–8 cells before degenerating during later stages of development. The embryo proper begins as a small sphere of cells that rapidly divides and becomes differentiated as it progresses through the globular, heart, torpedo, early cotyledon, and mature cotyledon stages of embryogenesis. The mature embryo contains two large cotyledons, a single hypocotyl, and both root and shoot apical meristems.

One of the most common experimental approaches to the analysis of embryo development in *Capsella* and other angiosperms has been the use of embryo culture. Immature embryos at the appropriate stage of development can be removed from the developing seed and cultured to maturity on a defined nutrient medium. This technique has been used not only to study the regulation of embryogenesis in plants but also to rescue aborted embryos produced by interspecific hybridization. It has also been possible in certain species of angiosperms to produce somatic embryos by growing cells derived from roots, leaves, or other sporophytic tissues on a defined nutrient medium with the appropriate growth regulators and then transferring the undifferentiated mass of cells onto a medium that supports the formation of nonzygotic embryos. This technique has demonstrated conclusively that the zygote is not the only plant cell capable of forming a normal embryo. Haploid plants have also been produced by culturing microspores on the appropriate nutrient medium. The immature microspores, in this instance, alter their pattern of development to form embryos instead of pollen grains.

The regulation of gene expression during embryogenesis in flowering plants has been approached through a variety of molecular and genetic techniques. Mutants defective in different aspects of embryo development have been analyzed, and many genes that are known to be expressed during embryogenesis are now being cloned and sequenced through the use of recombinant DNA technology. The genes that code for the major seed storage proteins have been studied in detail because they are transcribed into messenger ribonucleic acid at very high rates during later stages of embryogenesis. When more is learned about how the expression of these genes is regulated, it may be possible to alter the concentration and chemical composition of these storage proteins to improve the nutritional quality of seeds produced by important crop plants. It should also be possible to learn much more about how genes control the complex series of interactions that occur as the fertilized egg develops into a mature embryo. *See* DEVELOPMENTAL BIOLOGY; DEVELOPMENTAL GENETICS; REPRODUCTION (PLANT).

David W. Meinke

Bibliography. J. D. Bewley and M. Black (eds.), *Seeds: Physiology of Development and Germination*, 2d ed., 1994; L. W. Browder et al., *Developmental Biology*, 3d ed., 1997; B. M. Carlson, *Patten's Foundations of Embryology*, 6th ed., 2000; R. E. Coalson, *Embryology*, 2d ed., 1992; E. H. Davidson, *Gene Activity in Early Development*, 3d ed., 1986; G. M. Edelman, *Topobiology: An Introduction to Molecular Embryology*, 1993; V. Raghavan, *Embryogenesis in Angiosperms*, 1986; G. C. Schoenwolf and W. W. Mathews, *Atlas of Descriptive Embryology*, 6th ed., 2003; T. A. Steeves and I. M. Sussex, *Patterns in Plant Development*, 2d ed., 1989; F. H. Wilt and S. C. Hake, *Principles of Developmental Biology*, 2004; S. J. Wright, *A Photographic Atlas of Developmental Biology*, 2005.

Embryonated egg culture

Embryonated eggs are among the most useful and available forms of living animal tissue for the isolation and identification of animal viruses, for titrating viruses, and for quantity cultivation in the production of viral vaccines. The embryo proper, chorioallantoic membrane, yolk sac, allantoic sac, or amniotic sac may be inoculated in hens' eggs of various ages, so that a wide choice of types of tissue is available to fit the characteristics of the virus under study or for special studies. The chorioallantoic membrane is frequently used; in some infections, such as smallpox, vaccinia, and herpes simplex, characteristic lesions are produced which in some cases may resemble those in the natural host. For example, smallpox virus when cultured on the chorioallantoic membrane produces pocks and typical inclusions within the infected cells. When the embryo is inoculated, characteristic skin eruptions appear. Influenza virus, however, when inoculated into the amniotic cavity, does not give rise to pathology like that of the natural infection. *See* HERPES; INFLUENZA; SMALLPOX.

The growth of influenza virus in the allantoic cavity of embryonated eggs is probably the best-known use of eggs for animal virus culture. After inoculation into the allantoic cavity, the number of virus particles decreases exponentially, with almost 50% adsorbed to cells within 1 h. After adsorption the virus particles disappear (presumably by penetration of the cells), and an "eclipse period" of about 4 h follows. Evidence of viral multiplication first appears in the middle of the eclipse period, when soluble complement-fixing antigen (noninfective and nonhemagglutinating) is detectable. Hemagglutinating antigens can be found 3–4 h following inoculation, but in no greater amount than in the original inoculum; increases in hemagglutinating activity are seen only after 4 h or more. Infective virus is detectable 5–6 h after the allantoic cavity is inoculated; it is released almost at once into the allantoic fluid. Infective virus particles thus released then infect other cells of the allantoic sac, and virus continues to accumulate in the allantoic fluid for about 48 h, after which the system is exhausted. *See* ANIMAL VIRUS; ANTIGEN; COMPLEMENT-FIXATION TEST; FETAL MEMBRANE.

Joseph L. Melnick

Bibliography. D. O. White and F. Fenner, *Medical Virology,* 4th ed., 1994.

Embryonic differentiation

The process by which specialized and diversified structures arise during development of the embryo. The process involves (1) an increase in the number of cell types, and (2) an increase in morphological heterogeneity through the arrangement of cells into increasingly complex structural patterns in the form of tissues and organs. *See* HISTOGENESIS.

Differentiation begins in most organisms with fertilization of an egg with a sperm, after which the relatively large egg divides into many smaller cells called blastomeres. The blastomeres receive unequal portions of the cytoplasmic materials of the egg and are therefore initially somewhat different from each other. At the end of cleavage, the blastomeres are organized into a blastula, commonly either a hollow ball of cells or a flattened two-layered disk of cells. The cells of the blastula lie in different relative positions from those that will be occupied by their descendants in the adult organism. By a process known as gastrulation, they move to their approximate final positions and are arranged into three basic layers, called germ layers.

However, only two layers form in the simpler multicellular organisms. The outer layer is the ectoderm, from which arise the nervous system and the epidermal layer of the skin. The innermost germ layer, the entoderm, forms the epithelial lining of the digestive tract and contributes the essential tissue of associated organs. In all but the most primitive animals a third germ layer, the mesoderm, is formed by cells which come to lie in the area between the other two layers. In higher animals the mesoderm gives rise to most of the cells of the organism, such as those found in the muscles, skeleton, blood, connective tissue, kidneys, gonads, and certain other organs. The molding of groups of embryonic cells into such diverse tissues and organs proceeds through a variety of morphogenetic processes, such as migration, aggregation, dispersion, delamination, folding, and differential local growth of cells. *See* BLASTULATION; GASTRULATION; GERM LAYERS.

Cellular differentiation. Underlying the visible structural diversification of the embryo is the more fundamental and concomitant process of cellular differentiation (chemodifferentiation), by which embryonic cells are transformed into the highly specialized cells of the adult. A characteristic feature of this process is the production of manifold kinds of cells, all derived from a single precursor cell. The appearance of these new types or strains of cells is not abrupt but is the consequence of a long chain of progressive transformations in the molecular composition and organization of the cell. During these transformations each cell gradually loses its capacity to develop in alternative directions. At the same time it acquires characteristics essential for further differentiation along its prospective pathway.

The variety of adult cells produced by the divergent pathways of differentiation varies enormously in different kinds of animals, but in each individual the cells cooperatively discharge, in an ordered manner, the manifold functions of the organism. Simpler organisms are constructed of fewer kinds of cells, some of which seemingly perform a wider variety of functions than do the cells of more complex organisms. However, even the most specialized cells carry on a multitude of different chemical activities. Depending upon the criteria used, hundreds of different cell types may be recognized in the more complex animals. As methods of observation and analysis become more refined and penetrating, groups of apparently similar cells become resolvable into distinct types, in terms either of their functions or of their constituents.

Special synthetic products commonly distinguish many cell types, such as melanin in melanocytes, actomyosin in muscle cells, and hemoglobin in erythrocytes. But even these cell types may be further subdivided. Melanocytes of the skin are readily distinguished from those of the retina; and muscles of the heart, limbs, and digestive system each constitute distinctive cell strains. Moreover, the erythrocytes of the embryo have hemoglobin which is distinguishable from that found in adult erythrocytes. *See* CELL DIFFERENTIATION; CELL SENESCENCE; HEMATOPOIESIS.

Cell structure. A cell is an exceedingly complex structure, and presents numerous possibilities for alteration and diversification. Many of these possibilities have been realized in the specialized cells of adult organisms. In the center of the cell is the nucleus containing the chromosomes, which are responsible for determining the hereditary properties of the cell and for setting the limits of its developmental capacity. During cell division the chromosomes are duplicated and distributed in equal sets to the daughter cells; thus the cells of an organism are initially equipped with identical sets of chromosomes, one set derived from the egg, the other from the sperm. In each cell only a limited portion of this genetic endowment functions at any one time, and the most fundamental differences among cells stem from different sets of active genes. The genes are composed of deoxyribonucleic acid (DNA) that carries the code for synthesizing the proteins and enzymes that largely determine the properties of the cells of an organism. The more complicated the organism, the more DNA it requires: mammalian cells, for example, contain about 1000 times as much DNA as do bacterial cells. However, there are notable exceptions: among fish species the DNA content varies more than 150-fold and among amphibians more than 10-fold, yet these organisms are very similar in complexity. The significance of the vastly different amounts of DNA in the cells of these organisms is not understood at present. *See* DEOXYRIBONUCLEIC ACID (DNA).

During the differentiation of the cells of any organism the genetic code inscribed in the structure of active genes is transcribed into messenger ribonucleic acid (mRNA) molecules in the nucleus of the cell. These mRNA molecules then pass into the cytoplasm where they are translated into specific proteins or enzymes. The translation process requires two additional kinds of ribonucleic acid (RNA):

ribosomal RNA organized together with protein into small organelles called ribosomes, and transfer RNA responsible for carrying the amino acids from which proteins are made to the site of protein synthesis on the ribosomes. *See* GENETIC CODE; RIBOSOMES.

The activity of genes is dependent upon the chemical composition and physical structure of the chromosome. Morphological expressions of gene activity have been observed in some organisms: in certain dipteran insects, enlarged puffs appear at sites along the giant chromosomes during intense gene activity, while in amphibian oocytes, metabolically active loops extend laterally from the chromosomal axis. The molecular events underlying these phenomena are presently unknown, although at least two kinds of protein appear to be involved: the basic histones and the more acidic proteins. The histones are present in relatively constant amounts along the chromosomes and have been shown to inhibit gene-directed synthesis of ribonucleic acid. Chromosomes containing active genes also contain increased amounts of acidic or nonhistone proteins. These proteins may be involved in removing the histone inhibition of genes. *See* CHROMOSOME; MITOSIS; RIBONUCLEIC ACID (RNA).

In addition to the chromosomes, other constituents of the nucleus, the nucleoli, nucleoplasm, and the enclosing membrane also exhibit differences in various cell strains. However, the most conspicuous distinguishing characteristics of cells are found in the cytoplasm. In the cytoplasmic matrix are embedded various organelles, such as mitochondria, ribosomes, endoplasmic reticulum, centrosome, the Golgi complex, and a host of enzymes and other chemical substances, which by their proportions or type characterize the cell containing them. *See* CELL (BIOLOGY).

These components of the cell are organized into dynamic integrated patterns that confer on the cell the capacity to function and multiply. Moreover, once a terminal state of differentiation in the cell has been reached, it is conserved and perpetuated. Differentiated cells commonly do not multiply; but when division occurs under proper environmental conditions, the differentiated characteristics are preserved throughout descendant cell generations. Diverse cell strains are organized into still more complex patterns in the form of various tissues, which in turn are built up into organs. Tissues and organs vary greatly in the complexity of organization of their constituent cells but are generally composed of populations of heterogeneous cells held in rigid patterns by their mutual affinities, antagonisms, or both. The structure and function of tissues and organs undoubtedly are dependent upon the properties of the component cell; likewise, each cell is regulated in its function and course of differentiation by the cell population of which it is a part.

Mechanisms of differentiation. The mechanisms by which the course of cellular differentiation is realized are not precisely known. The factors involved may, however, be divided into two classes: (1) intrinsic, those operating within the cell, and (2) extrinsic, those brought to bear upon the cell from outside. Both classes of factors play a role in the differentiation of every cell. However, the relative importance of these factors varies considerably from one cell strain to another and also within the same cell at different stages in its development.

The fertilized egg begins development with a rich endowment, consisting of a nucleus with a set of paternal and maternal chromosomes together with a complexly organized cytoplasm. The activation of the egg by the sperm sets off a chain of actions and reactions that progressively transform the physical and chemical constitution of each descendant cell. The emergence of new cell characteristics may be attributed to an oscillating interaction between the intrinsic gene makeup of the cell and the surrounding cytoplasm. The dynamic imbalance existing between these interacting components drives the cell along its path of differentiation. In certain kinds of invertebrate embryos, interactions within each separate cell seem sufficient for guiding differentiation to its terminal state. Such embryos exhibit mosaic development. By contrast, in the embryos of vertebrates and certain invertebrates such as echinoderms, influences from adjacent cells are an essential part of the differentiation process. These embryos show regulative development.

Embryos showing mosaic development, in which cells differentiate autonomously, seem to be the rule among such groups as tunicates, mollusks, and annelids. In these organisms the destruction of a blastomere during cleavage results in a corresponding defect in later stages of development; that is, the injury is not repaired. Moreover, isolated blastomeres of such embryos tend to continue development as if they were still a part of an intact embryo. Throughout differentiation, the cells of these embryos are fixed in their developmental capacity and are able to differentiate in only one direction even when placed in abnormal surroundings. This type of differentiation is the product of an unfolding sequence of reactions and segregation of activities that follows a course predetermined in the structure of the egg; each step leads to the next without the intervention of influences from adjacent cells.

On the other hand, in regulative development the differentiation of cells and tissues is initiated and directed by inductive influences emanating from adjacent cells or tissues. In addition to the guiding influence of one tissue upon another, the general physicochemical environment established within a population of cells serves to regulate the further differentiation of the constituent cells. These population effects are referred to as embryonic fields or gradients. In contrast to the cells of mosaic embryos, regulative blastomeres exhibit great developmental plasticity. When transplanted to new locations in the embryo, the cells respond to their new environment by developing along pathways appropriate to their new location. Such developmental plasticity declines as differentiation proceeds, and the variety of pathways open to a cell is continuously restricted until the terminal state of differentiation is reached. It

should be emphasized that all gradations between the extremes of mosaic and regulative development exist and that even the tissues of highly regulative embryos in later stages of development exhibit increasingly autonomous differentiation. In all cases, however, differentiation is a gradual, progressive process resulting in the formation of specialized cells from generalized precursor cells. Cells, when fully differentiated, are very stable, and only under exceptional conditions of growth in tissue culture or during regeneration of injured tissues do cells lose their differentiated attributes and acquire the characteristics of other mature cells. *See* CELL LINEAGE; DEVELOPMENTAL BIOLOGY; EMBRYONIC INDUCTION.

Clement L. Markert

Embryonic induction

In the early development of many tissues and organs of complex, multicellular organisms, the action of one group of cells on another that leads to the establishment of the developmental pathway in the responding tissue. The groups of cells which influence the responding cells are termed the inducing tissue. Since specific inducing tissues cannot act on all types of cells, those cells which can respond are referred to as competent to react to the action of a specific inducer stimulus.

Embryonic induction is considered to play an important role in the development of tissues and organs in most animal embryos, from the lower chordates (for example, *Branchiostoma*) to the higher vertebrates (for example, mammals).

Primary induction. Perhaps the first major induction phenomenon occurs during the final stages of gastrulation of most animal embryos. Following fertilization, the egg divides to form a multicellular blastula-stage embryo. The cells of the blastula then undergo a series of movements which generate a more complex embryo, the gastrula, which contains three major groups of cells: ectoderm, mesoderm, and endoderm. The mesoderm actually arises as cells move from the surface of the embryo to the inside. Once inside, they induce the cells which reside over them, the surface ectoderm cells, to develop into the neural tube. The neural tube eventually forms the central nervous system. The first induction event of early embryogenesis is called primary embryonic induction (see **illus**.). The migratory cells which invaginate from the surface and induce the development of the neural tube are termed the embryonic organizer. The first step in the sequence of events termed primary embryonic induction is the acquisition by the mesoderm of neural inducing activity. Proteins such as fibroblast growth factor and activin, which belong to a category of so-called peptide growth factors, play key roles in programming the mesoderm cells to induce overlying ectoderm to differentiate into neural structures. *See* GASTRULATION.

The development of a large number of tissues and organs is influenced by embryonic inductions. Various eye structures (lens, optic cup, and so on), internal ear structures, as well as several tissues (for example, vertebral cartilage) emerge from cells which were acted upon by inducer tissues.

Sequential inductions. Following primary embryonic induction, the developing neural tube is influenced by the surrounding tissues, especially mesoderm, to undergo regional specialization. Noggin is a protein which is secreted by some cells, and can be purified from amphibian embryos. When added to organ cultures of ectoderm cells that have been separated from associated mesoderm, noggin induces the explant to develop neural characteristics.

Once brain tissue is induced, one region, the hindbrain, induces the overlying ectoderm tissue to develop into the ear placode. Later the ear placode induces surrounding tissue to develop into other structures of the ear. Likewise, one region of the forebrain makes contact with another area of the ectoderm and induces it to form a lens rudiment. The lens in turn exerts an inductive influence on the forebrain, causing it to change shape to form an optic cup. Finally, the optic cup and lens induce overlying tissue to

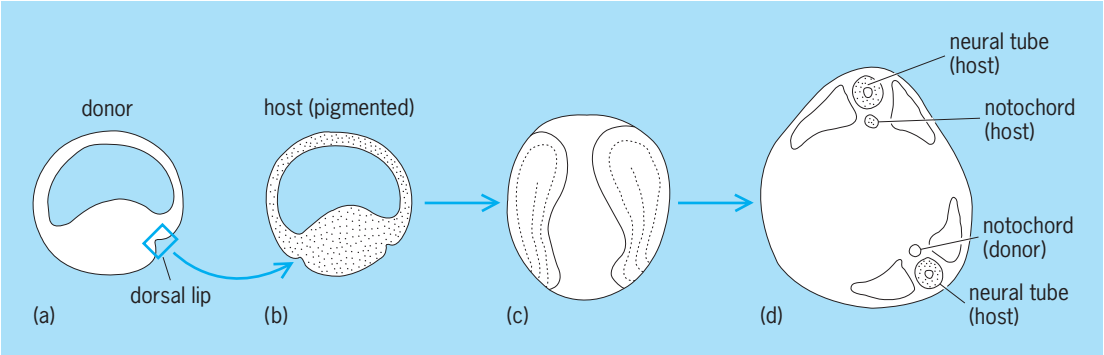

Experimental demonstration of primary embryonic induction in amphibian embryo. (*a*) Mesoderm tissue surrounding dorsal lip of donor embryo was grafted to (*b*) ventral side of pigmented host embryo. (*c*) External development of host reveals two neural plates. (*d*) Histological examination of later-stage embryo reveals two notochords as well—one developed from host cells, the other from donor cells. Both neural tubes were, however, of host origin. The donor tissue (for example, notochord) apparently induced the development of the second neural tube.

develop into the transparent cornea. Should any single step be omitted, or should actual cell-cell contact between inducing and responding tissue be prevented, subsequent events are adversely affected. Thus, the development of many organs depends upon a set of sequential, and in some cases (such as lens and optic cup) reciprocal, induction. *See* NERVOUS SYSTEM (VERTEBRATE).

Genetic influences. The inducing stimulus, whether from an intact tissue or a cell extract, can act only within the constraints of the responding tissue's genetic constitution. For example, the tissue which induces hair development in mammals can induce feather (but not hair) development in birds, as well as scale development in reptiles. *See* DEVELOPMENTAL GENETICS.

Underlying mechanism. Limbs, kidney, nasal structures, salivary glands, pancreas, teeth, feathers, and hair are organs which require inductive stimuli. It is not known whether a single common mechanism underlies each of those inductions. In some cases, such as lens induction, an instructional role for the inducing tissue (forebrain) is postulated. In other cases (such as the pancreas) a permissive role for the inducing tissue is imagined. In this role the inducer is required only to promote the successful self-differentiation of the responding tissue; that is, the inducer does not directly influence the choice of the developmental pathway.

Many scientists believe that inductive interactions are mediated by cell-cell contacts; that is, the developmental information which is transferred from the inducing tissue is thought to reside at the cell surface of that tissue. Perhaps the surface of the responding tissue recognizes the signal molecules present on the surface of the inducing tissue.

In other instances, a secreted protein such as noggin might move among various cells or tissues and exert its effects on competent cells.

Plant morphogenesis. The principles of animal development also apply to plants. A greater role is, however, usually played by the diffusion of small-molecular-weight signal molecules rather than cell-cell contacts or protein growth factors. The earliest stages of plant embryo development involve groups of cells acquiring the competence to respond to inductive signals. Later in development, inductive signaling also becomes important. For example, in flowering plants the distance between nodes along the stem elongates, and lateral buds form below the shoot apex. The buds are believed to develop in response to a concentration gradient of signal molecules which exists along the stem. Thus, a process which is analogous to embryonic limb bud formation in animals is played out, and both plant and animal inductions can be conceptualized in similar terms. *See* CELL DIFFERENTIATION; DEVELOPMENTAL BIOLOGY; EMBRYOLOGY; PLANT MORPHOGENESIS.

George M. Malacinski

Bibliography. S. F. Gilbert, *Developmental Biology*, 6th ed., 2000; J. P. Trinkaus, *Cells into Organs: The Forces That Shape the Embryo*, 2d ed., 1984.

Emerald

The medium- to dark-green gem variety of the mineral beryl, $Be_3Al_2Si_6O_{18}$, crystallizing in the hexagonal system. A flawless emerald with good color is one of the most sought after and highly prized of all precious gems. Emerald is restricted in its occurrence, and seldom are exceptional stones found; most emeralds are flawed and cloudy, and few stones command high prices. What constitutes a true emerald is well defined by the gem dealers, and the dividing line between emerald and mere green beryl is sharp.

Occurrence. In contradistinction to beryl and its other gem varieties, aquamarine, morganite, and goshenite, which almost exclusively occur in granite pegmatite druses, emeralds have only been found in mica schists or metasomatized limestones. The most outstanding occurrences include the Muzo and El Chivor mines in Colombia. Here emeralds occur in calcite veins associated with a dark limestone which has been metasomatized by magmatitic solutions. Noteworthy occurrences in mica schists include Tokovoja in the Ural Mountains, where emerald occurs with the beryllium minerals chrysoberyl (and its gem variety alexandrite) and phenakite; Habachtal, Austria; Transvaal, South Africa; and Kaliguman, India. The ultimate source of an emerald can often be assessed by a study of its inclusions.

Synthesis. Exceptional emeralds have been synthesized that rival the finest natural stones in quality and color. Broadly speaking, there are two main synthesis techniques. The molten-flux technique involves the compound oxides or gels of approximate beryl composition mixed with fluxes such as lithium molybdate and vanadium pentoxide. Fusion followed by slow cooling results in the growth of small prismatic crystals. The hydrothermal synthesis technique is capable of producing exceptional stones; emerald is hydrothermally grown on the plate of cut aquamarine which is used as a seed. The grown emerald is then sawed and used in turn as a seed for other emeralds. In all techniques 0.05–1.4% Cr_2O_3 is added. As in the natural emeralds, the Cr^{3+} chromophore is responsible for the green color. Detailed studies have been undertaken to define means of distinguishing natural emeralds from synthetics. Careful study (involving determination of birefringence and infrared absorption spectrum) will not only reveal whether a stone is natural or synthetic but will also distinguish the mode of synthesis if the emerald proved to be a synthetic stone. *See* BERYL; GEM.

Paul B. Moore

Bibliography. P. C. Keller, *Gemstones and Their Origins*, 1989; F. Ward, *Emeralds*, 2d ed., 2001.

Emergency medicine

The medical specialty that comprises the immediate decision making and action necessary to prevent death or further disability under emergency

conditions. It is based primarily in hospital emergency departments, but with extensive responsibilities for supervising emergency medical systems outside the hospital (paramedics).

Prehospital care. Prehospital care is professional emergency care delivered before the patient reaches the hospital. Before the 1960s there was almost no such care and ambulances merely provided transport. Care is provided by paramedics and emergency medical technicians. Paramedics are technicians trained in techniques such as electric shock for defibrillation of the heart, insertion of a tube into the trachea to assist respiration, and the use of intravenous drugs normally administered by physicians. Although possessing less training than registered nurses, they are authorized to carry out more procedures, and often function with less direct supervision. They operate under the supervision of an emergency physician, usually by radio communication. Emergency medical technicians (EMTs) are trained in first aid, give no medication, seldom communicate with a doctor, and often receive very little medical supervision. Paramedics are now found throughout the United States and are usually associated with ambulances and prehospital emergency care, but most such care is actually provided by emergency medical technicians.

The Federal Emergency Medical Services Act of 1974 established emergency medical service (EMS) districts throughout the United States. These districts set up regulations for paramedics and emergency medical technicians, administer the base hospitals that supervise the paramedics, and help establish rules for the management of trauma, burn, pediatric, and other emergencies. The control of emergency care is vested in different agencies in different regions; it may be the health department, the fire department, a private hospital, an ambulance company, or another body. Very few emergency medical service districts actually monitor the quality of care in a scientific manner, as hospitals are legally required to do. Nor does any impartial organization set standards for the districts' performance, audit the districts, or collect statistics about them. As a result, the quality of prehospital care varies widely and frequently is low.

There is little scientific knowledge about what constitutes effective prehospital care. Rapid transport of a victim of major trauma to a trauma center has been well studied and proven beneficial, as has prompt electrical defibrillation in cases of cardiac arrest. Whether these elaborate and expensive systems provide any benefit for the overwhelming majority of prehospital patients who suffer from neither of these conditions requires further research.

Trauma care. Trauma is physical injury, whether minor or major; it does not include medical emergencies due to other causes. In the United States, trauma constitutes only about 15% of all emergencies, and that includes minor trauma such as broken bones or lacerations. Life-threatening trauma needing emergency surgery constitutes only about 1–2% of all emergencies. In major trauma (usually caused by vehicular or industrial accidents, stabbing or gunshots wounds, or falls) the best care requires rapid transport to the hospital. However, such rapid transport is the wrong approach to the far more common medical emergencies, such as cardiac arrest, for which emergency treatment must begin at the site of occurrence.

Professional organizations, such as the American College of Emergency Physicians and the American College of Surgeons, have set standards for trauma centers, which are hospitals that treat a minimum number of serious trauma cases, and that meet the standards of these organizations in terms of staffing, experience, and medical outcome. A trauma center should belong to a systematic plan implemented by an emergency medical service region, which determines under what conditions an injured patient should be transported to a trauma center rather than the closest hospital emergency department. The purpose of such a plan is to ensure that only serious trauma is referred to trauma centers, which will provide the best care, but which should not be overloaded with the more common minor injury cases that can be well cared for in any properly staffed hospital emergency department.

Proper treatment of trauma requires prompt assessment by qualified paramedics, who may have to immobilize the neck to prevent spinal cord injury, protect the breathing passages, or administer intravenous fluid to replace blood lost by hemorrhage. Further evaluation and treatment by a qualified emergency physician must take place immediately upon arrival at the emergency department or trauma center. Head injuries account for a large percentage of trauma deaths in the United States; ensuring an adequate airway is crucial since the oxygen supply to the brain and other tissue must be maintained. Also, efforts must be made to control edema (swelling) of the injured brain, which leads to increases in intracranial pressure that can cut off blood supply and result in brain death. Fatalities in other types of trauma are due to conditions such as internal bleeding and punctured lungs (pneumothorax), which usually require surgery within hours. The administration of large quantities of intravenous fluid and blood has made it possible to save most patients who reach a trauma center. Before properly cross-matched blood is available, it has been traditional to maintain blood volume with 0.9% saline solution, but recent research has shown that 5% saline is much more effective; it is even helpful in the small volumes that can easily be administered before the patient reaches the hospital. Trauma care has also been revolutionized by the advent of computerized tomographic x-ray, which reveals details of intracranial and intraabdominal injury that previously could be determined only by exploratory surgery. Magnetic resonance imaging is an even more sophisticated noninvasive technique that can reveal subtle details of tissue density, blood flow, and anatomy that previously could not be discerned. *See* COMPUTERIZED TOMOGRAPHY.

Burn care. Burns are a very serious injury because the protective barrier of the skin is destroyed,

leading to massive loss of body fluids, development of shock, and invasion by bacteria. Patients with major burns are best treated in specialized burn centers and should be taken there promptly. Burn mortality has dropped significantly in recent decades because of prompt intravenous administration of very large volumes of fluid to replace body fluid losses, and the use of antibiotics. *See* BURN.

Many burns are caused by smoking in bed, and could be avoided by the marketing of available self-extinguishing cigarettes and by more widespread installation of smoke detectors.

Management of cardiac arrest. Cardiac arrest occurs when the heart suddenly goes into an abnormal rhythm so that it does not pump blood; it is the chief cause of death in heart attack, and causes 400,000 deaths a year in the United States. Cardiac arrest tends to occur in the first few hours of heart attack, often before the patient has even decided to go to the hospital, and sometimes is not preceded by any symptoms. If treated by defibrillation (electric shock to the heart) within the first 6 min, it is reversible. Otherwise, death is certain. Therefore, to save lives, emergency medical service programs must deliver paramedics to the patient's side within 6 min of the arrest, not an easy task. Nonetheless, in cities where such programs succeed, and particularly where many citizens are trained in basic cardiopulmonary resuscitation (CPR) to keep the heart and brain tissues alive until professional help comes, about 25% of patients are resuscitated and return to a normal life. This fact spurred the standardization and widespread teaching of cardiopulmonary resuscitation and certification of nurses and physicians in this skill by by the American Heart Association. Resuscitation of persons who have suffered cardiac arrest is a major advance that requires a public trained in cardiopulmonary resuscitation as well as public experienced and well-supervised paramedics. In most areas of the United States, too few of the key elements are in place, and survival rates from cardiac arrest are 10% or less.

Cardiopulmonary resuscitation includes the procedure of compressing the chest rhythmically 80 times a minute (in adults) to produce blood flow. Unfortunately, the amount of blood flow is a small fraction of normal and can keep heart and brain tissues alive only for a short time. Rescuers also must breathe for (ventilate) the patient; this can be done with a mask attached to a ventilating bag, or ideally, by placing an endotracheal tube in the patient's lungs and ventilating with 100% oxygen. The most successful cardiopulmonary resuscitation, and the usefulness of defibrillation, is limited to those patients who are in the abnormal rhythm known as ventricular fibrillation. However, about half of cardiac arrest victims are in other rhythms (asystole or electromechanical dissociation) for which no effective treatment is known.

Second to defibrillation, administration of epinephrine (to improve perfusion of the heart tissue during cardiopulmonary resuscitation) is the most useful treatment of cardiac arrest. This drug may return perfusion of the heart to adequate levels long enough to allow the heart to start beating spontaneously again.

A revolutionary development was the production of automatic defibrillators that interpret heart rhythms electronically and administer electric shocks when (and only when) needed. There are about 33,000 paramedics in the United States who are trained to use defibrillators. However, one of the 450,000 emergency medical technicians can usually reach a victim much sooner than a paramedic. Automatic defibrillators allow fire fighters, emergency medical technicians, and others with very little training to treat patients in cardiac arrest.

Most heart attack victims have a blockage of the arteries supplying blood to the heart, which until recently was not promptly treatable. However, two new drugs—streptokinase and tissue plasminogen activator (tPA)—can permit rapid treatment. Both drugs dissolve clots rapidly and can be administered as soon as the patient arrives in the emergency department, saving heart tissue that would otherwise die from lack of blood. The earlier these drugs are given, the better the results, and researchers are studying the possibility of administering them before the patient reaches the hospital.

Unfortunately, there is no way to resuscitate patients who have heart rhythms that do not respond to defibrillation. An even greater problem is how to prevent the brain damage that often occurs from lack of blood supply during cardiac arrest or after resuscitation. Brain resuscitation is a serious problem because in many instances the heart is resuscitated but within hours a relatively healthy brain will swell from widespread cellular damage. Then pressure increases inside the skull, the blood supply is further worsened, and irreversible brain damage eventually occurs.

Treatment of poisoning. Innumerable poisonings occur every year in the United States, involving thousands of products. Many are accidental (particularly in children and confused elderly), but many result from deliberate attempts at suicide or abuse. The physician cannot stay abreast of all toxic syndromes and their constantly evolving treatment. Therefore poison centers have been established in every state, where trained physicians and pharmacologists equipped with electronic and on-line references provide emergency telephone consultation. Poison centers have greatly improved the quality of care in poisoning cases.

It used to be traditional in poisoning to administer emetics such as syrup of ipecac. However, this does not empty the stomach completely. For most poisonings it is preferable simply to administer activated charcoal, which binds the poison in the gut and prevents its absorption. This has simplified and improved the treatment of poisoning. Michael Callaham

Bibliography. American College of Emergency Physicians, Guidelines for emergency department physician staffing, *Ann. Emerg. Med.*, 13:1165–1166, 1984; M. Callaham (ed.), *Current Therapy in Emergency Medicine*, 1987; P. Rosen et al. (eds.),

Emergency Medicine: *Concepts and Clinical Practice*, 2d ed., 1987.

Emery

A natural mixture of corundum with magnetite or with hematite and spinel. Emery has been used for centuries as an abrasive or polishing material. Because the mixture is very intimate and appears to be quite homogeneous, it was considered to be a single mineral species until the middle of the nineteenth century. The aggregate has a gray-to-black color and is extremely tough and difficult to break. The specific gravity varies from 3.7 to 4.3, depending upon the relative amounts of the constituent minerals. The hardness is about 8 (Mohs scale), less than that of pure corundum which is 9, and is more dependent upon the physical state of aggregation than on the percentage of corundum. *See* HEMATITE; MAGNETITE; SPINEL.

Since early times emery has been recovered from Cape Emeri on the island of Naxos and from other islands in the Grecian archipelago. Here it occurs as irregular beds and lenses and in loose blocks associated with crystalline limestone and schists. It is also found at several localities in Asia Minor under similar conditions, notably at Gumach Dagh, east of Ephesus, and at Kula, near Alashehr. Emery was worked during the latter part of the nineteenth century at Chester, Massachusetts, where it was associated with diaspore, margarite, and chloritoid. Because of its magnetic properties, resulting from the admixed magnetite, the Chester material was first worked unsuccessfully as an iron ore. Only after the similarity of the associated minerals with those of the Naxos emery was noted was its true nature determined.

Although synthetic abrasives have replaced emery in many of its earlier uses, it is still used as an abrasive and polishing material by lapidaries and in the manufacture of lenses, prisms, and other optical equipment. Emery wheels, emery paper, and emery cloth are used not only by lapidaries but also by machinists in the grinding and polishing of steel. *See* ABRASIVE.

Cornelius S. Hurlbut, Jr.

Emission spectrochemical analysis

An instrumental analytical technique used in qualitative or quantitative chemical analysis. It is conducted by monitoring and measuring the spectrum of light emitted by the material being analyzed. *See* ATOMIC SPECTROMETRY; QUALITATIVE CHEMICAL ANALYSIS; QUANTITATIVE CHEMICAL ANALYSIS.

In general, there are many ways in which to conduct an emission spectrometric measurement. The differences among approaches result from the choice of location within the electromagnetic spectrum at which to observe emitted radiation. Emission spectrochemical analysis, however, is the name given traditionally to those analytical determinations based on radiation in the visible through vacuum ultraviolet region of the electromagnetic spectrum. The included wavelengths are approximately 800–150 nanometers. The technique is used principally to detect (qualitative analysis) and determine (quantitative analysis) concentrations of metals. A few nonmetals can also be determined. Under optimum conditions, as little as 10^{-10} g of an element per gram of sample can be determined. Routine concentration ranges in which emission spectrometry is used are approximately 10^{-8}–10^{-2} g per gram of sample (1% by weight).

The steps in emission spectrochemical analysis are vaporization and atomization of the sample, excitation of the atomic vapor produced, dispersion of the emitted radiation, and observation and measurement of the dispersed radiation.

A number of approaches to emission spectrochemical analysis have been developed, all of which follow the four steps listed but which differ fundamentally in the first two. These steps are involved with the transfer of energy to the sample so that the constituents of the sample will emit characteristic radiation. The application, control, and transfer of different forms of energy pose significantly different equipment requirements. Therefore, different versions of emission spectrometry have been developed. The principal modern versions of vaporization, atomization, and excitation are alternating-current (ac) spark, direct-current (dc) arc, inert-gas electrical plasmas [dc plasma (DCP), inductively coupled plasma (ICP)], and chemical flame.

Vaporization and atomization. First it is necessary to produce a vapor (vaporization) in which all the compounds of the sample are broken down into their constituent atoms (atomization). It is the presence of the individual atoms that is detected and quantitated. If the sample is originally a solid, it can be made part of an anode-cathode pair. When an electric discharge is struck across these electrodes, producing an ac spark or dc arc, the sample is vaporized into the electrode gap. There are numerous ways to configure an electrode so that the sample is an integral part of it. The most desirable designs make it possible to reproduce the procedure simply and easily, so that the technique is routinely applicable.

A more recent development for the production of a cloud of atoms from a solid sample is to fire a high-power pulse of laser radiation at the sample. The material that is vaporized when the laser creates a small crater is in just the physical state required for an atomic emission experiment. This approach has become more extensively used as the sample vapor generated has become more reproducible. The method is favorable for qualitative or semiquantitative analyses or in situations where it is desired to do sampling and analyses in the field.

Dissolved solids or liquid samples can be broken down into microscopic droplets in a process called nebulization, and a stream of these droplets can be injected directly into a discharge of hot gas. The hot gas then evaporates the solvent in the droplet, vaporizes the resulting solid microparticle, and atomizes

the compounds of which the particle is made.

A variety of techniques have been developed in which sample nebulization systems and components for generation of a stable hot gas work together. The earliest and most successful of these techniques involves nebulizing the sample and mixing it with the fuel and oxidant in a burner manifold; as a flame is generated, the sample is coincidentally vaporized and atomized. Many approaches for nebulizing the sample, mixing the droplets and gases, and configuring the burner manifold and burner head have proven successful.

Excitation of atomic vapor. In the majority of spectrochemical instrument systems, the same process is used both to produce the atomic vapor and to excite the atoms to cause emission of characteristic radiation. Energy is provided to a hot gas, the hot gas and sample are mixed to cause vaporization, and more of the energy of the hot gas is used to excite the atoms of vapor.

The hot gas is produced in a number of different ways. A chemical flame produces a hot gas from the combustion of some sort of fuel, for example, acetylene or nitrous oxide, with the release of many thousands of calories of heat. Chemical flames normally have temperatures between 1000 and 4000 K (1300 and 7000°F).

An electrical discharge, which is caused by applied potentials or applied currents across an electrode gap (ac spark or dc arc), produces a hot gas of atmospheric species, electrode material, and electrons. Temperatures up to about 7500 K (13,000°F) result.

More advanced types of electrical discharges, known as inert-gas electrical plasmas, produce hot gases of much higher temperature. A plasma is a cloud of vapor in which more than 1% of the atoms are ionized; that is, they have their outer electrons stripped away. In general, the greater the extent of ionization, the greater the amount of energy resident in the plasma.

Modern electrical plasmas are produced in two principal ways. Direct-current plasmas are electrode-based plasmas in which a conventional dc arc is ignited in an inert gas (usually argon) in an electrode gap by applying a current. Unlike most dc arcs, the arc in a direct-current plasma has a highly controlled flow of cooling gas (also argon) surrounding it. The cooling flow decreases the electrical conductivity of the outer regions of the arc, causing the applied current to be carried in a path of reduced volume. This is called a thermal pinch of the arc. As a result, the current density of the arc rises significantly, and this causes the temperature of the gas through which the arc is carried to increase. Analytical direct-current plasmas generate plasmas of argon with temperatures between 6000 and 8000 K (11,000 and 14,000°F)

Inductively coupled plasmas (a generic term for non-electrode-based plasmas) also are generated using argon. The inert gas reaches a plasma state through the interaction of electrons, carrier gas atoms, and an applied radio-frequency (kHz to MHz) or microwave-frequency (MHz to GHz) field of energy. Electrons will gain energy from the electric field established at an induction coil surrounding a flow of carrier gas (see **illustration**) or within a microwave-frequency resonant cavity through which the carrier gas flows. The electrons eventually gain extremely high kinetic energies. Although the electrons that gain this high energy have low mass individually, they are able to impart this energy through numerous collisions with the more massive argon carrier gas atoms. The carrier gas atoms eventually obtain a sufficient amount of energy to reach excited states, which takes only microseconds. At this point the plasma has ignited, and the carrier gas can release its energy by emitting radiation, losing electrons (ionizing), or colliding with other atoms (thermal transfer of energy). Temperatures routinely attained in induced plasmas are between 6000 and 8000 K (11,000 and 14,000°F). *See* MICROWAVE; PLASMA (PHYSICS).

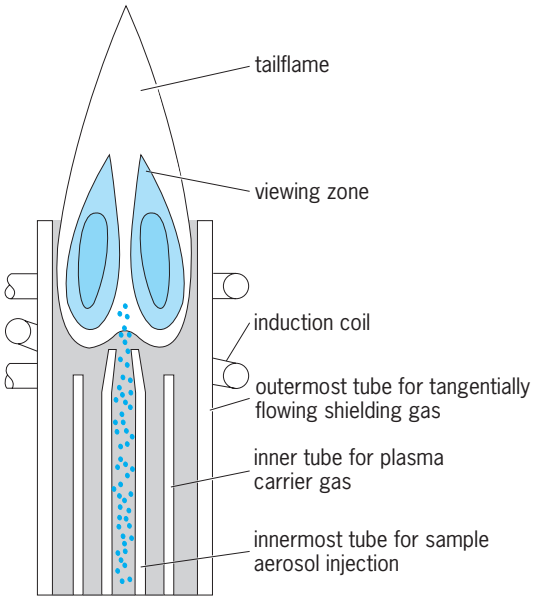

Inductivity coupled radio-frequency plasma.

Once the hot gas is produced, the vaporized and atomized sample is injected into it; or alternatively, the hot gas functions as both the vaporizer and atomizer, that is, cold sample is injected into it. In either case, the plasma is used to impart energy to the analyte atoms so that they may reach any number of excited states. Predominantly, collisional mechanisms of energy transfer contribute to the excitation of atoms. Other mechanistic pathways, such as resonance with applied fields and radiative transfer, probably occur. While more than one mechanism is followed as the analyte releases this energy, the most desirable route is that of radiational deactivation; that is, the analyte loses the energy by ejecting photons (emitting light).

Because the analyte atoms are allowed to attain a large number of excited states, a wide range of radiations (different wavelengths of light, photons of different energies) are emitted. The energy gained

and lost relates directly, through Planck's equation, $E = hc/\lambda$, to wavelengths between the infrared and vacuum ultraviolet regions of the spectrum. *See* PHOTOCHEMISTRY.

Dispersion and resolution. Identification of the element of interest and quantitative measurement of the radiation it emits requires that the light emitted be dispersed into its component wavelengths; that is, a spectrum of the light must be produced. This is done with a spectrograph using photographic detection or with a spectrometer using photoelectric detection. The instrument dispersion, its ability to separate wavelengths from one another, must be on the order of 0.01–0.1 nm/mm so that individual wavelengths are adequately separated at the detector. Prism spectrometers are no longer used since they do not have adequate dispersion over the entire wavelength range of interest. Diffraction grating instruments have mostly uniform dispersion and give quite high resolution. In general, spectrochemical analysis requires high spectral resolution—the ability to distinguish one dispersed wavelength from another. The reason is that the emission experiment can produce a larger number of spectral lines than other types of spectrometric approaches, such as fluorescence or absorption. *See* DIFFRACTION GRATING; SPECTROSCOPY; SPECTRUM.

Observation and measurement. Visual observation is not used in routine spectrochemical analysis, because the human eye does not possess any resolving power to sort out the complex spectra produced.

Photographic methods. Photographic observation has long been the method of choice for spectrochemical detection. It is used mainly for qualitative and quantitative analysis at the minor and trace constituent levels. Photographic glass plates are usually employed. Spectrograms are viewed with a projection comparator, which presents an image of both the sample and standard emulsion plates on a split-field screen, permitting any spectrum on one plate to be brought adjacent to and in register with any spectrum on the other. Any photographic emulsion is suitable for the 230–430-nm wavelength range. At higher wavelengths, special sensitization is needed. Emulsion sensitivity varies directly with the grain size in the emulsion. High sensitivity is obtained at the cost of resolving power, which varies inversely with grain size. Sensitivity and contrast vary with wavelength and age for all emulsions.

Measurement of line intensities requires calibration of the emulsion by constructing a relationship involving the blackening of the emulsion as a function of the intensity of the incident light. Lines whose intensities are to be compared photographically should be as close as possible in wavelength. Densitometers or microphotometers measure spectral-line blackening by scanning the illuminated spectrum with a fine slit or by projecting a fine illuminated line onto and through the spectrogram and onto a photocell.

The major advantage of photographic observation is that it is both the least expensive and the simplest method for obtaining the entire spectrum simultaneously and storing it nearly indefinitely. It is not the most sensitive method and is not useful for ultratrace-element analysis.

Photoelectric methods. Photoelectric observation affords extremely sensitive and rapid detection. The electric current output of this detecting device is directly proportional to the rate at which photons strike a radiation-sensitive anode. This current is collected and processed by various electronic means, and is amplified to give an electrical signal that is proportional to the amount of analyte present in the sample vapor. Photoelectric detection is best suited to trace and ultratrace determinations because of its high sensitivity. However, it presents significant problems for simultaneous multielement analyses, since it needs one detector for each wavelength to be measured. Arrays of photomultiplier tubes have been constructed for simultaneous observation of 60 or more wavelengths. *See* PHOTOELECTRIC DEVICES; PHOTOMULTIPLIER; PHOTOTUBE.

Another approach to multielement detection is to use only one photomultiplier tube and to adjust the spectrometer precisely so that successive wavelengths pass by the detector in time. This is known as spectrometer scanning. Its principal disadvantage is that each wavelength is viewed sequentially, and so it takes longer to monitor more wavelengths. Its principal advantage is that there is no limit to the wavelengths which can be viewed; all of them are accessible.

Solid-state arrays. Another approach to multielement, simultaneous detection is to mimic the photographic plate with an electronic imaging device, such as a photodiode array or charge-transfer device. The simultaneity and simplicity of photographic detection is obtained in conjunction with the high sensitivity and speed of photoelectric detection. *See* CHARGE-COUPLED DEVICES; PHOTODIODE ARRAYS.

Solid-state imaging devices have become useful for spectrochemical detection. These devices are semiconductor circuits that are responsive to radiation. Most of them are made from silicon. If the silicon is held at a discrete electrical potential (a voltage is impressed across it), the absorption of an incident photon by the silicon crystal will cause an electron to migrate out of position (this is called forming a hole). The hole-electron pair formation is a charge-separation process. The positively charged holes are collected at the negative end of the crystal of silicon or stored at a capacitor. More incident photons make more holes and lead to increased charge collection. The amount of external current needed to neutralize the holes produced is measured and is proportional to the number of photons absorbed, constituting a measure of the radiation emitted by the sample.

Physically discrete photon-detecting units are formed when the semiconductor devices are fabricated. These units are called picture elements, or pixels. The action of each pixel is similar to that of a single photomultiplier. The pixel responds to the radiation it receives when it is located at a specific wavelength position on the focal plane of a spectrometer. Combinations of pixels, called arrays, can be configured to match the shape of the focal plane.

The pixels, then, are equivalent to a photographic plate, only with very large grain size.

One advantage of using solid-state arrays for detection is that large numbers of pixels can be fabricated on a single array. Detector arrays with many millions of pixels can come close to mimicking the exact performance of a photoplate. Another advantage is that arrays can be produced by cost-effective and rapid semiconductor processes, since the same technology used to make computer chips can be used to make photodetectors. Yet another advantage of solid-state array detectors is that they are addressed electronically, and this can be done rapidly with a computer. While arrays of photomultiplier tubes can also be addressed electronically, there are rarely more than a few dozen detectors used at any one time. Photographic plates have many more equivalent pixels on them than solid-state arrays, but it is orders of magnitude slower in retrieving their recorded signals. Solid-state arrays also have an advantage of photoplates in the ability to integrate signal; that is, the more photons that strike the pixels, the more signal is collected, so that large signals can be accumulated over time. *See* INTEGRATED CIRCUITS.

A disadvantage of solid-state arrays is that their pixel density determines the resolving power of the spectrometer. Since each pixel has a finite size, on the order of tens of micrometers, greater resolving power has to come through the use of larger and more expensive spectrometers. Finally, solid-state arrays of silicon cannot compete completely with photomultipliers with respect to radiation sensitivity. The signal range over which silicon responds to radiation will always be smaller than that of a phototube, making the arrays less useful for ultratrace determinations.

Andrew T. Zander

Bibliography. J. A. C. Broekaert, *Analytical Atomic Spectrometry with Flames and Plasmas*, 2002; M. Cullen, *Atomic Spectroscopy in Elemental Analysis*, 2004; J. R. Dean, *Practical Inductively Coupled Plasma Spectoscopy*, 2005; J. D. Ingle and S. R. Crouch, *Spectrochemical Analysis*, 1988; R. Kellner et al., *Analytical Chemistry. A Modern Approach to Analytical Science*, 2d ed., 2004; A. Montaser and D. W. Golightly (eds.), *Inductively Coupled Plasmas in Analytical Atomic Spectrometry*, 2d ed., 1992; J. Nolte, *ICP Emission Spectrometry: A Practical Guide*, 2003; F. A. Settle (ed.), *Handbook of Instrumental Techniques for Analytical Chemistry*, 1997; R. Thomas, *Practical Guide to ICP-MS*, 2004; H. H. Willard et al., *Instrumental Methods of Analysis*, 7th ed., 1988.

Emissivity

The ratio of the radiation intensity of a nonblackbody to the radiation intensity of a blackbody. This ratio, which is usually designated by the Greek letter ϵ, is always less than or just equal to one. The emissivity characterizes the radiation or absorption quality of nonblack bodies. Published values are readily available for most substances. Emissivities vary with temperature and also vary throughout the spectrum. For an extended discussion of blackbody radiation and related information *see* HEAT RADIATION

There are several methods by means of which the emissivity can be determined. The one most commonly used is the cavity method. In this technique a fine hole is provided in a radiating surface, and the ratio of the radiation intensity from the surface to the radiation intensity from the hole yields the emissivity directly. This method is quite accurate. One can also use an optical pyrometer to determine the emissivity from the brightness temperatures of the hole and the surface in conjunction with Wien's law of radiation.

The total emissivity when introduced into the Stefan-Boltzmann law gives the total radiated energy W in joules per square centimeter of the real heat radiator as $W = \epsilon\sigma T^4$. Here T represents the absolute temperature and σ, the radiation constant, has the value 5.67×10^{-12} joule cm^{-2} K^{-4}. This energy is always smaller than the energy radiated by the blackbody, since ϵ is less than 1. For example, the total emissivity for tungsten is 0.32 at 2500°C (5032°F), which means that at the same temperature tungsten radiates approximately one-third the energy of a blackbody.

The spectral emissivity ϵ_λ (the subscript λ denotes the wavelength) provides information on the energy distribution. Any spectral emissivity value is valid only for a narrow wavelength interval. The wavelength at which ϵ_λ has been determined is indicated by a subscript, for instance, $\epsilon_{0.655}$. A spectral emissivity of zero means that the heat radiator emits no radiation at this wavelength. Strongly selective radiators, such as insulators or ceramics, have spectral emissivities close to 1 in some parts of the spectrum, and close to zero in other parts. Carbon has a high spectral emissivity throughout the visible and infrared spectrum, exceeding 0.90 in certain portions; thus carbon is a good blackbody radiator. Tantalum is the only metal with a spectral emissivity greater than 0.5 in the visible spectrum. All other metals have a lower spectral emissivity. Tungsten is a relatively good emitter, with a spectral emissivity of 0.43–0.47 within the visible region of the spectrum. *See* BLACKBODY.

Heinz G. Sell; Peter J. Walsh

Emitter follower

A circuit that uses a common-collector transistor amplifier stage with unity voltage gain, large input resistance R_i, and small output resistance R_o (see **illus.**). In its behavior, the emitter follower is analogous and very similar to the source follower in metal-oxide-semiconductor (MOS) circuits. Many electronic circuits have a relatively high output resistances and cannot deliver adequate power to a low-resistance load, or do suffer unacceptable voltage attenuation. In these cases, an emitter follower acts as a very simple buffer. Widely used, it is often found as the last stage of a multistage amplifier so that the circuit is better able to drive a low-resistance load. *See* AMPLIFIER.

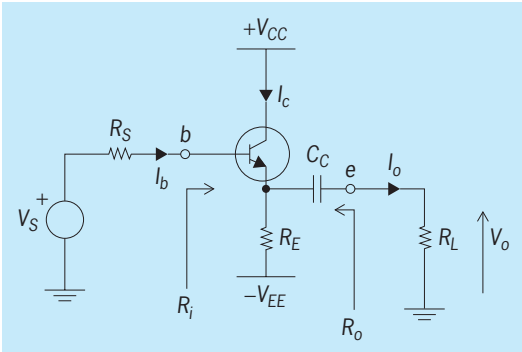

Schematic diagram of an emitter follower circuit.

Operation. In the emitter follower, the collector is connected to the positive power supply V_{CC}, that is, to signal ground, so that the collector is common to alternating-current input and output (see illus.). The emitter is connected through a (large) resistor R_E to the negative power supply, $-V_{EE}$. The resistor R_E helps set the direct-current bias of the transistor. In integrated circuits, the function of R_E is normally replaced by a current source. The implementation makes no difference in principle in the following discussion: R_E simply represents the effective resistance of the current source. The emitter follower circuit, between input terminal b at the base and output terminal e at the emitter, is inserted between a load R_L and a source V_S with its source resistor R_S. The coupling capacitor C_C, if used, prevents R_L from influencing the transistor bias. The capacitance C_C is chosen large enough to act as a short circuit at signal frequencies. *See* CURRENT SOURCES AND MIRRORS; INTEGRATED CIRCUITS.

Characteristics. Under some simple practical assumptions, the input resistance R_i of the emitter follower is approximately $\beta + 1$ times the load resistor, $R_i \approx (\beta + 1)R_L$, where the parameter β is the transistor's current gain I_c/I_b. A typical value for β is 100. The same assumptions lead to the output resistance R_o being approximately equal to the source resistor divided by $\beta + 1$, $R_o \approx R_S/(\beta + 1)$. The voltage gain V_o/V_S is approximately equal to, but always slightly less than, unity (in practice 0.9 to 0.99) whereas the current gain is large: the output current I_o delivered by the emitter follower is approximately $\beta + 1$ times the input current I_b, $I_o \approx (\beta + 1)I_b$. In summary, the emitter follower is an impedance transformer. It effectively divides the source resistor and multiplies the load resistor by $\beta + 1$. It has approximately unity voltage gain but a large current gain, $\beta + 1$. Consequently, the emitter follower provides power gain. The power V_oI_o delivered to the load is approximately equal to $\beta + 1$ times the power V_SI_b drawn from the source, $V_oI_o \approx (\beta + 1)\ V_SI_b$.

Performance limits. An emitter follower has a very large bandwidth so that the simple equations in the previous paragraph, especially the constant near-unity voltage gain with no phase shift, describe the circuit's functioning well until very high frequencies. The ultimate limits are given by the parameters of the chosen transistor. The negative feedback caused by R_E tends to reduce distortion in the emitter follower, but the signals must not get so large that the output current becomes equal to the bias current. Signal clipping will then result. When using an emitter follower, it must be observed that the forward-biased base-emitter diode imposes a 0.7-V direct-current voltage difference between the input (base) and output (emitter) terminals. Care must be taken to limit the current, should the output terminal be accidentally shorted to ground. *See* DISTORTION (ELECTRONIC CIRCUITS); TRANSISTOR. Rolf Schaumann

Bibliography. P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3d ed., Wiley, New York, 1993; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., Oxford University Press, New York, 1998.

Emotion

An umbrella concept in the common language, typically defined by instantiation by reference to a variety of mental and behavioral states. These range from lust to a sense of liking, from joy to hostile aggression, and from esthetic appreciation to disgust. Emotions are usually considered to be accompanied by some degree of internal, frequently visceral, excitement, as well as strong evaluative components. Emotions are also often described as irrational, that is, not subject to deliberative cogitation, and as interfering with normal thought processes.

These latter qualities are often exacerbated in the emotional behavior and expression seen in clinical cases. The expression of strong emotions is typically considered to be symptomatic of some underlying conflict, and even the positive emotions are used as indices of unusually strong attachments and atypical earlier experiences. Sigmund Freud introduced the concept of repression to describe a defense mechanism against the occurrence of strong emotional experiences. From the psychoanalytic point of view, what is repressed is not the emotion itself, since the very concept of emotion implies conscious experience, but rather the memory of an event which, if it became conscious, would lead to strong conflicts and emotional consequences. Many other defense mechanisms, such as rationalization and compulsive or obsessive neurotic symptoms, are also seen as serving the purpose of avoiding conscious conflict and emotional sequelae. *See* PSYCHOANALYSIS.

Classical views. The conventional view of emotion in classical psychological and biological theory, at least to the end of the nineteenth century, was embodied in Charles Darwin's classic study of emotional expression in animals and humans. Emotional expression was described as dependent on some prior emotional mental state on the one hand, and on the evolutionary history of both the expression and the prior mental state on the other. Some situation "elicited" the mental state of fear, for example, and this state was then manifested in transspecies patterns of facial (and other bodily) expression. The cumulative evidence speaks for the existence of

universal facial and muscular patterns which are in fact recognized, across societies and cultures, as communicating similar evaluative and "emotional" messages. However, whether these expressions are evolutionary products of prior emotional states or whether they are the remnants of primitive preverbal communicative systems is open to question. It certainly is the case that the facial and gestural acts can occur without an accompanying emotional state and can frequently serve purely communicative, evaluative purposes.

The first major break with the traditional position occurred in the late nineteenth century when William James and Carl Lange proceeded independently to turn the classical sequence upside down. As James proposed in 1884: "My thesis is . . . that the bodily changes follow directly the *perception* of the exciting fact, and that our feelings of the same changes as they occur *is* the emotion." This position required that for each discrete emotional experience there should be a prior discrete and different pattern of bodily (visceral and skeletal) changes. The lack of evidence for this and other aspects of the James-Lange position led the physiologist W. B. Cannon to a series of critical expositions which culminated in the thalamic theory of emotion, a neurophysiological, rather than psychological, position which has been influential in physiological approaches to emotion.

Neurophysiological theory. The thalamic theory became one of the bases for neurophysiological speculations and experiments about emotion. It suggested that emotions arise from the stimulation or excitation of areas in the thalamus but that these thalamic tendencies are normally inhibited by the activities of the cerebral cortex. Strong, innate stimuli bypass this cortical inhibition and act directly on the thalamus.

On the other hand, acquired emotions depend in the first instance on the recognition of, and memory for, certain classes of events that have "emotional" meaning. The memory for these events is activated in the cortex, and the inhibition by the cortex is released. When the thalamus is activated, signals from the thalamus serve two functions: they activate peripheral autonomic effects as well as cortical ones; the former produce the bodily expression of emotion, the latter its subjective experience. Removal of the cortex in animals does in fact produce hyperemotionality, but this persists even when the thalamus is removed. These studies and others have implicated the hypothalamus in the production of emotional experience and behavior. There is evidence of even wider areas of the brain in the emotional complex, with much emphasis placed on the role of the limbic system and the neocortex of the frontal lobes. *See* NERVOUS SYSTEM (VERTEBRATE).

Peripheral physiology. The essential aspect of subjective emotions appears to be the occurrence of visceral arousal. It is the peripheral autonomic nervous system that is primarily responsible for the reactions that are called visceral, with the sympathetic division focused on energy-expending function and the parasympathetic system devoted to energy-conserving functions. Cannon was principally responsible for noting the homeostatic function of the autonomic nervous system which "keeps the internal environment constant and fit for continued . . . action." However, the human homeostat was conceived as being a passive respondent, just like the physical thermostat. Subsequent evidence has shown that autonomic, particularly sympathetic, changes may have very active consequences on the organism. They mobilize not only internal balance but also the scanning of and attention to environmental changes. The autonomic system not only prepares the internal environment for action but also prepares the whole organism for action on the external environment. *See* AUTONOMIC NERVOUS SYSTEM.

Within the context of the homeostatic view of the autonomic nervous system, the events that produce autonomic arousal are usually listed as physical stressors, events that threaten the physical well-being of the organism, such as tissue injury, pain, or loss of support. With the additional emphasis on organismic action, it is possible to define a set of events that produce autonomic reaction, in addition to those that depend on physical insult. The interruption of ongoing action, the disconfirmation of current expectations, the violation of perceptual expectancies, all generally subsumed under the interruption of thought, perception, and action, produce visceral (autonomic) arousal. Such arousal, in turn, generates additional attention and action. This version of the function of the autonomic nervous system also permits a view of the emotions as being adaptive, rather than merely annoying or disruptive. Emotional (visceral) arousal will occur for events that are important for the organism, that is, events that diverge from its expectations, past experiences, and past habits. Thus, the emotional experience emphasizes important events in the organism's history, events that are different from the run-of-the-mill of everyday life. What is different is important, and what is important is underlined by the emotional experience. At the same time, the action of the autonomic nervous system makes it possible to cope better with these important events.

Cognitive/autonomic interactions. Modern views of emotion were significantly influenced by a series of experiments by S. Schachter and J. E. Singer which demonstrated the important coaction of cognitive (knowledge-oriented) and physiological (autonomic) events. They showed that in the presence of gross sympathetic arousal of unspecified origin, individuals will report subjective emotional experiences in keeping with their evaluations (cognitions) of their environment; that when the arousal can be ascribed to some specific cause, environmental evaluations are relatively ineffective; and also that emotional states are reported only to the extent that the individual experiences a state of physiological (visceral) arousal.

The experience of emotion is, therefore, the result of two different but complementary systems:

the arousal experience brought about by the autonomic nervous system, and a cognitive/evaluative state depending on the individual's current experiences, past evaluations of similar states, and the general accumulation of beliefs, values, and attitudes. Emotions arise when the current situation is different from that usually experienced, when it is experientially salient. The degree of divergence from the usual and the expected determines the intensity of the arousal and of the emotion. On the other hand, the current evaluation of the situation and the self determines the quality of the emotional experience. Thus, both positive and negative emotional states arise out of the combination of arousal and evaluation.

Within the complex interactions of cognitive and physiological factors, a variety of theoretical views can and have been accommodated. Thus, whether expressive or evaluative/communicative, the role of facial and gestural aspects of emotion are seen as part of the more general complex of the emotional experience. Biochemical and neurological factors that affect both cognitive and autonomic variables come into play, as well as pharmacological agents that exacerbate or suppress the state of arousal.

The perception of visceral arousal and the attention to the environmental origins and conditions of the individual's emotional state are factors that make for the apparent disorganizing effects of emotional experiences. The span of human attention or consciousness is known to be limited, and it is also known to be reduced by internal or external events that require some of that attentional capacity. Just as a loud noise or an extraneous thought reduces attentional capacity during a demanding task, so too does the internal "noise" of visceral activity. In addition, the external demands of an "emotional" situation intrude on a limited attentional capacity. As the attentional capacity is reduced during emotional states, the remaining attentional ability is focused on the aspects of the situation that the individual considers central. The judgment of centrality or importance is of course a subjective evaluation and may or may not be objectively correct. In moments of high emotional excitement or stress, it is frequently the case that attention is focused on events that may or may not be adaptive. In the latter case, emotion becomes disorganizing; in the former it is beneficial. Contemporary analyses of emotional states thus prevent any simple generalizations about emotions as being "good" or "bad" for the organism; they are properly seen as a component of the total adaptive complex of human thought and behavior rather than as a peculiar evolutionary hang-over. *See* CONSCIOUSNESS.

George Mandler

Bibliography. J. Callwood, *Emotions: What They Are and How They Affect Us*, 1986; M. Clynes and J. Panksepp (eds.), *Emotions and Psychopathology*, 1988; C. Darwin, *The Expression of Emotions in Man and Animals*, 1873; N. H. Frijda, *The Emotions*, 1987; C. E. Izard, *The Face of Emotion*, 1988; J. G. Thompson, *The Psychobiology of Emotions*, 1988.

Emphysema

The abnormal enlargement of air spaces distal to the terminal bronchioles in the lung. In the United States, the term emphysema is often used imprecisely to refer to any member of the group of illnesses known as chronic obstructive pulmonary disease, including emphysema, chronic bronchitis, and nonallergic asthma. The collective term obstructive pulmonary disease arises because many patients have elements of all three diseases. Although histological examination of lung tissue from surgical specimens or at autopsy is the only definitive diagnostic method, advances in physiological testing and radiologic techniques usually make an earlier diagnosis possible.

Types. There are two major anatomic patterns of emphysema, centriacinar or panacinar. In centriacinar emphysema, which is the form associated with heavy smoking, the grossly enlarged air spaces tend to appear in the respiratory bronchioles in the proximal part of the acinus, the unit of lung tissue; the upper portions of the lungs are most heavily involved. Histologically the lesion appears as a dilated or punched-out hole surrounded by relatively normal or only mildly thickened alveolar septa. In the panacinar type, the enlarged air spaces involve the entire acinus, from respiratory bronchioles to alveoli. Panacinar emphysema either favors the lower part of the lung or is diffuse, and is associated particularly with a deficiency of the glycoprotein α_1-antitrypsin.

Physiology. The destruction of alveolar walls results in a loss of the elastic recoil of lung tissue and an increase in lung compliance, that is, a greater than normal lung volume with any distending pressure developed by the respiratory muscles. With progressive loss of elasticity, the thorax enlarges and the distending or transpulmonary pressure falls. The drop in pressure across the lung results in a loss of support of the airways, narrowing of the airways, and progressive airflow limitation. With such limitation, the ability of the body to respond to the metabolic demands of exercise by increasing total ventilation is reduced, and the individual's tolerance for physical activity gradually declines. Enlargement of air spaces, with its concomitant loss of capillary bed, causes less efficient gas transfer. The "wasted" ventilation can be quantified by measuring the transfer of tiny amounts of carbon monoxide into the capillaries during breath holding. The destruction of blood vessels causes a rise in pulmonary vascular resistance and increased pumping demand from the right ventricle. Initially the right ventricle responds by increasing thickness and strength of its muscle wall, a condition known as cor pulmonale. Eventually it is unable to handle the load, and dilates and goes into right ventricular failure with systemic venous congestion. *See* HEART DISORDERS.

Pathogenesis. Emphysema is characterized by loss of elasticity of the lung due primarily to the destruction of elastin, an important structural protein in the lung's supporting connective tissue. The elastin is

detroyed when there is an imbalance between enzymes that can damage tissue and those that block this effect. In response to a threat to the airways, such as infection or cigarette smoke, alveolar macrophages and neutrophils move into the lungs to digest the foreign material and in so doing release some of their rich supply of proteases. Without the neutralizing effect of antiproteases, the proteases can damage alveolar tissue and create emphysematous lesions. *See* CONNECTIVE TISSUE.

Both animal and human research has demonstrated the protease-antiprotease action in emphysema. Human emphysema has been replicated in the lungs of laboratory animals by the instillation of papain. The active ingredient in papain is elastase, a proteolytic enzyme that destroys elastin. Human susceptibility to emphysema increases in individuals that are homozygously deficient in the serum antiprotease α_1-antitrypsin, which makes up about 90% of serum α_1-globulin and is the most important antiprotease in humans.

Tobacco smoking causes a low-grade pulmonary injury that attracts inflammatory cells into the lung, possibly setting off the chain of events that progresses to emphysema. Although virtually all smokers have increased inflammatory cells that can be recovered from their airways, only a minority develop emphysema. Since the level of α_1-antitrypsin does not distinguish those who will develop emphysema, other factors must be involved. In the absence of a smoking history, emphysema is distinctly unusual. Homozygous deficiency of α_1-antitrypsin—perhaps other, as-yet-unrecognized proteins—result in premature emphysema, especially in the presence of even modest smoking. Certain occupational exposures, such as contact with cadmium, can also cause emphysema in the absence of smoking.

Diagnosis and treatment. The diagnosis of emphysema is usually suspected when a smoker develops symptoms of chronic shortness of breath and cough. However, the disease may be present for many years before a person becomes symptomatic. Physical examination rarely leads to early diagnosis, although a decrease in the intensity of breath sounds and an increase in chest volume may be apparent. Pulmonary function tests may be normal in the very early stages but will eventually show airflow limitation, an increase in total volume of the lungs, and a loss of diffusing capacity for carbon monoxide. Measurement of lung compliance is technically feasible but rarely performed, because it requires passage of a balloon into the esophagus to measure the pressure that develops at different lung volumes. Chest x-rays may show flattening of the diaphragm and hyperinflation of the chest, but more significant is the appearance of regions of increased lucency, which reflects the enlargement of air spaces and loss of pulmonary blood vessels. Advances in computerized axial tomography of the chest allow radiologic diagnosis of emphysema at a stage of disease progression that would not be visible on ordinary chest x-rays. *See* MEDICAL IMAGING.

Because tissue destruction in emphysema is irreversible, the key to treatment is prevention, which usually involves cessation of smoking as early in the course of the disease as possible. Even when the disease is well established, however, smoking cessation may slow its progression, reduce cough and sputum production (from the accompanying chronic bronchitis), and prevent the recurrent respiratory infections that seem to hasten the negative course of the disease. Because muscle tone in the smooth muscle encircling the bronchi may be increased, drugs that relax this muscle tone (bronchodilators) may be helpful, especially when administered as an aerosol directly into the lungs. Drugs known as sympathomimetics may mimic the effects of stimulation of the sympathetic nervous system, and others referred to as anticholinergics can block the cholinergic nervous system. In addition, since some degree of inflammation often accompanies emphysema, corticosteroid medications may be of some use.

Even if the course of the disease cannot be altered, general measures such as avoidance of irritants, vaccinations against influenza and pneumococcus, prompt treatment of infections, adequate nutrition, family support, regular exercise, and participation in multidisciplinary rehabilitation programs may substantially improve well-being and the ability to function independently. *See* LUNG; RESPIRATORY SYSTEM DISORDERS. Michael Stulbarg

Bibliography. B. Burrows, An overview of obstructive lung disease, *Med. Clin. N. Amer.*, 65:455–471, 1981; N. J. Gross and M. S. Skorodin, Role of the parasympathetic system in airway obstruction due to emphysema, *N. Engl. J. Med.*, 311:421–425, 1984; M. Higgins, Epidemiology of COPD: State of the art, *Chest*, 85(suppl.):35–75, 1984; T. L. Petty (ed.), *Chronic Obstructive Pulmonary Disease*, 1985; T. L. Petty and L. M. Nett, *Enjoying Life with Emphysema*, 1983; G. L. Snider et al., The definition of emphysema (report of the National Heart, Lung and Blood Institute, Division of Lung Diseases Workshop), *Amer. Rev. Respir. Dis.*, 132:182–185, 1985; W. M. Thurlbeck, *Chronic Airflow Obstruction in Lung Disease*, 1976.

Empirical method

The empirical method is not sharply defined. It is generally characterized by the collection of a large amount of data before much speculation as to their significance, or without much idea of what to expect, and is to be contrasted with more theoretical methods in which the collection of empirical data is guided largely by preliminary theoretical exploration of what to expect. The empirical method is necessary in entering hitherto completely unexplored fields, and becomes less purely empirical as the acquired mastery of the field increases. Successful use of an exclusively empirical method demands a higher degree of intuitive ability in the practitioner. Percy W. Bridgman; Gerald Holton

Emulsion

A system of small liquid droplets dispersed in a second, immiscible liquid. Emulsions are a type of colloidal dispersion. By classical definition, emulsion droplets have diameters between 1 nanometer and 1 micrometer. However, for practical purposes, the principles of colloid science can be usefully applied to emulsions whose droplets are as large as tens or even hundreds of micrometers. The droplets in an emulsion are large enough that they do not behave like the atoms and molecules of classical chemistry. For example, emulsions do not generally behave like true solutions and may have undetectable freezing-point depressions. On the other hand, the droplets are small enough that they do not behave like the macroscopic particles of classical physics. For example, emulsion droplets may rise (cream) or fall (sediment) extremely slowly in apparent violation of Stokes' law. *See* COLLOID; SEDIMENTATION (INDUSTRY).

In most emulsions, one of the liquids is aqueous (that is, water containing dissolved substances), while the other is an oil (usually a hydrocarbon liquid). The two simplest kinds of emulsion are those in which oil droplets are dispersed in water (denoted oil-in-water, or O/W) and those in which waterdroplets are dispersed in oil (denoted water-in-oil, or W/O). In practice, emulsions can be much more complex and can consist of droplets within droplets. A double emulsion might contain aqueous droplets dispersed in oil droplets that are in turn dispersed in a continuous aqueous phase. Such an emulsion is termed water-in-oil-in-water, or W/O/W. The opposite kind of double emulsion is O/W/O. Double emulsion droplets can be quite large (tens of micrometers), and each can contain many droplets within it.

Preparation. If a small amount of vegetable oil is added to a cup of tap water and vigorously agitated, the oil will disperse into droplets: an emulsion has just been made. This emulsion will not be very stable, and in a short time the oil droplets will rise to the surface of the water and fuse (coalesce) together to produce a layer of oil floating on the water. Emulsions having a reasonable degree of stability contain an additional, emulsifying (stabilizing) agent. The emulsifying agents are usually surfactants or mixtures of surfactants and may include polymers or even fine solids. An ideal emulsifier has several properties. First, it makes the emulsion easy to create by reducing interfacial tension between the oil and aqueous phases. Second, it promotes creation of the desired kind of emulsion, whether O/W or W/O. Finally, it stabilizes the dispersed droplets against coalescence by providing an electrical and/or physical barrier at the droplets' surfaces. Beyond the oil and aqueous phases plus the emulsifier, chemical components may be added to an emulsion formulation to provide greater stability and to adjust other properties such as viscosity, color, and appearance. Properly formulated emulsions can remain stable for months or even years. *See* SURFACTANT.

Emulsions are prepared by the breaking-up of one liquid phase into droplets and then the further breaking-up, or comminution, of these droplets into smaller droplets. Typically, the emulsifying agent is dissolved into the phase where it is most soluble, after which the second phase is added and shear is applied to the mixture using either high-speed mixing or vigorous agitation. A variety of high-shear mixing/agitating devices are available for this purpose, including paddle, propeller, or turbine mixers, rotor-stator homogenizers, and ultrasound generators.

An important aspect of emulsion preparation is selecting an appropriate emulsifying agent. Although some experimentation will always be required, there are some rules of thumb that provide useful starting points for predicting the type of emulsion that will form under a given set of circumstances. For example, Bancroft's rule predicts that if an emulsifying agent is preferentially dissolved in, or wetted by, one of the liquid phases, that phase will likely form the external (continuous) phase. A related rule of thumb, the oriented wedge theory, predicts that soaps of monovalent metal cations, such as sodium, tend to produce O/W emulsions while those of polyvalent metal cations, such as calcium, will tend to produce W/O emulsions. The same basic concepts underlie the dimensionless hydrophile–lipophile balance (HLB) scale, in which the relative tendency of an emulsifying agent to prefer to dissolve in oil versus water is quantified. On the HLB scale, oil-soluble (lipophilic) emulsifiers have low values, usually less than 9, while water-soluble (hydrophilic) emulsifiers have high values, usually greater than 11. A very polar nonionic surfactant might have an HLB value of close to 20, while an ionic surfactant, such as sodium dodecyl sulfate, has an HLB of about 40. Good emulsifiers for W/O emulsions tend to have HLB values of about 3 to 8, while good emulsifiers for O/W emulsions tend to have HLB values of about 8 to 18. These rules of thumb provide indications of the kind of emulsion a potential emulsifying agent may promote, but not its efficiency. In addition, mixtures of emulsifiers tend to be more efficient at emulsifying than pure compounds having the same HLB. Laboratory emulsification tests are virtually always required in order to select an appropriate emulsifying agent (or mixture) and to determine the optimum amount of emulsifier to use.

Properties. Although emulsions can be prepared that will remain stable for weeks, months, or even years in some cases, most emulsions are not truly stable. Eventually, the dispersed droplets will fall or rise due to gravity (sedimentation/creaming). As they come into contact with each other, some will begin to stick together (aggregation), and some of those droplets that stick together will eventually fuse into larger droplets (coalescence). Over time, these processes will cause an emulsion to "break" or separate into two distinct liquid layers. A special kind of emulsion, called microemulsion, is actually thermodynamically stable; microemulsions do not break on standing or with centrifuging. Besides the emulsifying agent common to all conventional

emulsions, microemulsions contain an additional stabilizing agent (cosurfactant). The combined emulsifying agents promote the formation of very small size droplets, often 10 nm or less, and change the properties of the droplet surfaces so that the droplets behave almost as though they were truly dissolved. The droplets in a microemulsion can be so small that they scatter very little light and appear to be transparent.

Emulsions have some special properties. Having a system of small liquid droplets dispersed in a second, immiscible liquid means that emulsion droplets have high surface area–to–volume ratios. This means that, by being present in an emulsion, a very small volume of (dispersed) liquid can be made to present a very large amount of surface area, such as for reactions between species at the surface and species in solution.

Another consequence of having a dispersion of small droplets is an increased resistance to flow, or viscosity. An emulsion with a small volume of liquid (less than about 10% by volume) making up the dispersed droplets will have, essentially, a small number of small droplets so that individual droplets will not encounter each other very often and will have little effect on viscosity. However, as the volume of liquid making up the dispersed droplets increases above about 10% by volume, the number of droplets increases and so does the frequency of droplet-droplet encounters. As droplets encounter and interact with each other more often, there is a "crowding" effect, and it becomes more difficult to make the emulsion flow. As the total volume occupied by the droplets increases, so does the emulsion viscosity. In some emulsions the surfaces of the dispersed droplets carry electric charges, while in others the droplet surfaces have long surfactant or polymer molecules extending outward from them. In these cases, the electric fields and extended molecules cause additional forces between droplets when they encounter each other, which further acts to increase emulsion viscosity. *See* VISCOSITY.

In theory, the maximum volume possible for the liquid making up the dispersed droplets is 74%—the maximum volume for close-packed rigid spheres. In special cases, it is possible to make emulsions in which the droplets distort and occupy a greater volume. Generally when this happens, the emulsion inverts. For example, an O/W emulsion pushed to have a total oil droplet volume of 80% by volume will usually quickly invert to become a W/O emulsion having a 20% dispersed waterdroplet volume.

Characterization. Emulsions may be characterized in a number of ways.

1. The nature of the emulsion can be characterized as O/W, W/O, O/W/O, W/O/W, or O/W/O/W. The nature of the dispersed phase can often be determined using dyes and/or microscopy, whereas the nature of the external phase can be determined by electrical conductivity. The existence of multiple dispersed phases can usually be determined using optical or electron microscopy.

2. Droplet size and droplet size distribution can be determined by such techniques as optical or electron microscopy and light scattering.

3. Composition is usually determined by a titration method for simple emulsions. For more complex emulsions, composition is usually determined by separating the oil and water, and any gas or solids, and measuring their respective quantities.

4. Stability of the emulsion against breaking is usually determined using standardized bottle or jar tests, in which emulsion samples are allowed to stand for a specified period of time, with or without centrifuging, and then examined. *See* ELECTRON MICROSCOPE; ELECTROLYTIC CONDUCTANCE; LIGHT-SCATTERING TECHNIQUES.

Occurrence. Systems containing colloidal-sized particles, droplets, or bubbles are important because they feature prominently, in both desirable and undesirable contexts, in a wide variety of practical disciplines, products, and industrial processes (see **table**). The problems associated with colloids are usually interdisciplinary in nature, and a broad scientific base is required to understand them completely.

Emulsions are commonly used in many industries. Emulsion products include some foods, insecticides and herbicides, polishes, drugs, biological systems, personal-care creams, ointments, and lotions, greases, inks, paints, and varnishes. Double emulsions have applications in cosmetics, agriculture, food, photography, leather, and drug delivery. The use of emulsions to provide permanent and transient antifoams is an application that is important to an even broader variety of products and processes, including foods, cosmetics, pharmaceuticals, pulp and paper, water treatment, and minerals beneficiation. Emulsions may be encountered in virtually all of the process industries, including mineral processing, petroleum production and refining, asphalt paving, and paper de-inking. Emulsions also occur in the environment. Examples include the difficult-to-treat "chocolate mousse" type of emulsion that may be formed when crude oil is spilled on the ocean, and soil remediation, in which soil contaminants are removed in the form of an emulsion during a surfactant flushing process. A currently emerging area is that of nanoemulsions, having droplet sizes of about 1 nm or less, which are of interest in medical diagnostics and drug treatments because of their potential ability to cross the endothelial cells that protect the brain.

Demulsification. Demulsification refers to emulsion breaking. Some emulsions occur where they are not desired, such as in the process industries, and these emulsions have to be broken so the two liquid phases can be separated. The first step is to understand the type of emulsion being dealt with in terms of its nature (O/W, W/O, or multiple emulsion) and preferably in terms of the mechanism by which the emulsion was stabilized. Emulsion breaking involves two steps: causing the droplets to aggregate, or come into contact; and causing the droplets to coalesce, or fuse together. If these can be accomplished, the two liquid phases can be separated.

Some examples of emulsion occurrence

Application area	Emulsion examples	Emulsion types*
Environment	Insecticide and herbicide formulations	O/W
	Water and sewage treatment emulsions	O/W
	Oil spill emulsions	W/O, O/W
	Sulfide liquids in magma or lava	L/L
Mineral processing	Emulsion flotation media	O/W
	Oil flotation froth	O/W
Petroleum production	Oil well stimulation emulsion	O/W, W/O
	Emulsion drilling fluid	O/W, W/O
	Produced (well-head) emulsions	W/O
	Enhanced oil recovery in situ emulsions	O/W
Petroleum transportation	Transportation fuel emulsion (70% heavy oil)	O/W
	Heavy oil pipeline emulsions	O/W
	Oil spill mousse emulsions	W/O
	Tanker bilge emulsions	O/W
Manufacturing and materials	Asphalt (paving) emulsion	O/W
	Aqueous polymer-type paints	O/W
	Microemulsion metalworking oil	O/W
Food products	Milk, ice cream, creams, whippable toppings	O/W
	Mayonnaise, hollandaise, béarnaise sauces	O/W
	Butter, margarine, spreads, processed cheese	W/O
	Salad dressings	O/W, W/O
Agriculture	Pesticide, herbicide and fungicides	O/W, W/O/W
Biology and medicine	Blood	O/W
	Vesicles, vacuoles	W/W, W/W/W
	Emulsion encapsulated drugs	O/W, W/O, W/O/W
Personal care	Cosmetic and skin care creams	W/O, O/W, W/O/W
	Microemulsion hair dyes	W/O

*Liquid (L), water (W), oil (O).

Sometimes an emulsion can be broken by simply changing the temperature or by applying mechanical shear. A variety of emulsion-breaking devices and vessels are available under names such as separators, settlers, heaters, treaters, desalters, and free-water knockouts. All are designed to induce droplet aggregation and coalescence. In most cases, a chemical demulsifier is also needed.

Chemical demulsifiers generally act by neutralizing the stabilizing effect of the original emulsifier. For example, a W/O emulsion can sometimes be broken by adding just the right amount of a surfactant that would normally promote an O/W emulsion. Dosage is critically important. Adding too much of such a surfactant will usually create the reverse emulsion, replacing one problem emulsion with another. Similarly, if an emulsion is stabilized by electrical forces, demulsification could be brought about by adding a chemical agent that can overcome or reduce these forces. Following addition of an appropriate demulsifier, mechanical agitation is often all that is needed to cause the separation of the oil and water phases. A wide range of potential chemical demulsifiers is available, but selecting an appropriate demulsifier, and the optimum dosage, can involve considerable testing due to the large number of factors than can influence emulsion stability and demulsifier performance.

Laurier L. Schramm

Bibliography. P. Becher, *Emulsions: Theory and Practice*, 3d ed., American Chemical Society, 2001; P. Becher (ed.), *Encyclopedia of Emulsion Technology*, vols. 1–3, Marcel Dekker, 1983, 1985, 1988; B. P. Binks (ed.), *Modern Aspects of Emulsion Science*, Royal Society of Chemistry, Cambridge, 1998; L. L. Schramm, *Emulsions, Foams, and Suspensions, Fundamentals and Applications*, Wiley-VCH, Weinheim, 2005; A. L. Smith (ed.), *Theory and Practice of Emulsion Technology*, Academic Press, 1976.

Emulsion polymerization

A variety of free-radical polymerization—a method where individual molecules of high-molecular-weight polymer are rapidly generated by a free-radical chain reaction—in which a monomer with low water solubility is dispersed in small droplets (~10 micrometers in diameter) in water, and stabilized with surfactant to prevent separation of these two phases. A dilute "seed" latex of minute polymer particles in water may also be added. The reactants are agitated and heated, and a free-radical initiator is added, leading to the formation of a latex—a dispersion of polymer particles of colloidal dimensions in water (**Fig. 1**). *See* COLLOID; EMULSION; FREE RADICAL; POLYMER; POLYMERIZATION; SURFACTANT.

The term "emulsion polymerization" is a misnomer, as polymerization does not take place to any significant extent in the emulsified monomer droplets. The misnomer arose because the process was

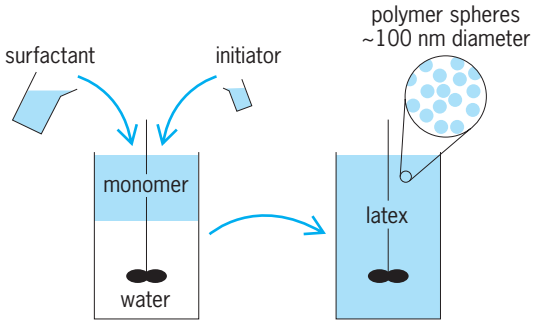

Fig. 1. Process of emulsion polymerization.

originally designed to polymerize in emulsion droplets, and the pioneers in the field did not realize that a much more complex process actually occurred. The first patents describe attempts to reproduce products similar to the natural polyisoprene latex produced by the rubber tree (*Hevea brasiliensis*), by dispersing monomers in water. Emulsion polymerization became an industrial process by the 1930s, with production of neoprene and of butadiene-acrylonitrile rubbers in Germany and the United States. Development and understanding of the process dramatically accelerated during World War II, when a research effort second only to the Manhattan Project was made in the United States for the industrial preparation of synthetic rubber.

Mechanism. A typical initiator of emulsion polymerization is an aqueous-phase thermal initiator (for example, ammonium persulfate, $NH_4S_2O_8$), although a redox couple (for example, potassium persulfate/sodium metabisulfite, $K_2S_2O_8/Na_2S_2O_5$) may also be used. Decomposition of the initiator in the aqueous phase generates hydrophilic radicals, which rapidly add to the small but not negligible amount of monomer present in the aqueous phase [reaction (1)]. The

$$SO_4^{\bullet -} + \text{methyl methacrylate (M)} \longrightarrow {}^{-}OSO_3M^{\bullet} \quad (1)$$

monomeric radical formed may then add a second monomer molecule [reaction (2)], and so on. Two

$$^{-}OSO_3M^{\bullet} + M \longrightarrow {}^{-}OSO_3MM^{\bullet} \quad (2)$$

of these radical species may combine to form an inert product (bimolecular termination) in the water phase. As more monomer units are added, the radical is increasingly hydrophobic. At a critical degree of polymerization z, it becomes surface-active and rapidly adsorbs to the nearest hydrophobic interface. In the early stages of a typical emulsion polymerization, this interface is a micelle—an aggregate of surfactant molecules with a typical diameter of a few nanometers—or a preformed polymer seed particle. Entry to emulsion droplets is highly unlikely, because although emulsion droplets are large they are much less numerous than micelles. In the environment of a micelle or small particle, the reactive center immediately encounters a concentration of monomer much greater than in the aqueous phase and propagates rapidly to form polymer [reactions (3) and (4)]

$$^{-}OSO_3(M)_{z-1}^{\bullet} + M \longrightarrow {}^{-}OSO_3(M)_z^{\bullet} \quad (3)$$

$$^{-}OSO_3(M)_z^{\bullet} + n(M) \longrightarrow \text{polymeric radical } (P^{\bullet}) \text{ in the particle phase} \quad (4)$$

(**Fig. 2**). The growing polymer chain will terminate either by bimolecular termination [reaction (5)] with

$$P^{\bullet} + P^{\bullet} \longrightarrow \text{polymer molecule} \quad (5)$$

an entering radical, or by abstraction of hydrogen from another species (typically monomer) to generate a radical capable of exiting to the aqueous phase [reaction (6)]. *See* INTERFACE OF PHASES; MICELLE.

$$P^{\bullet} + M \longrightarrow \text{polymer molecule} + M^{\bullet} \quad (6)$$

As polymerization proceeds, the polymer particles will grow until no free surfactant remains to stabilize additional polymer/water interfaces. The monomer droplets merely serve as reservoirs of monomer, which will diffuse from the droplets to the growing polymer particles. Once the micelles have disappeared, new z-meric radicals in the water phase will all enter latex particles. (While this micellar nucleation mechanism is the most common way that particles are formed, there are other mechanisms, for example, homogeneous nucleation, involving precipitation of polymer chains of degree of polymerization $j_{crit} > z$, which occurs in surfactant-free systems.) Eventually, all the monomer droplets will disappear, and the polymerization will consume the monomer present in the particles. It is common in industry to use a controlled feed, whereby ingredients, such as monomer, are added during the polymerization process. Thus, it is possible to maintain an approximately constant monomer concentration during much of the polymerization process, or to add a different monomer after the original monomer has been consumed.

Polymer solid contents of 50% by weight are routinely obtained in industry, with negligible amounts of coagulation and typical particle diameters of

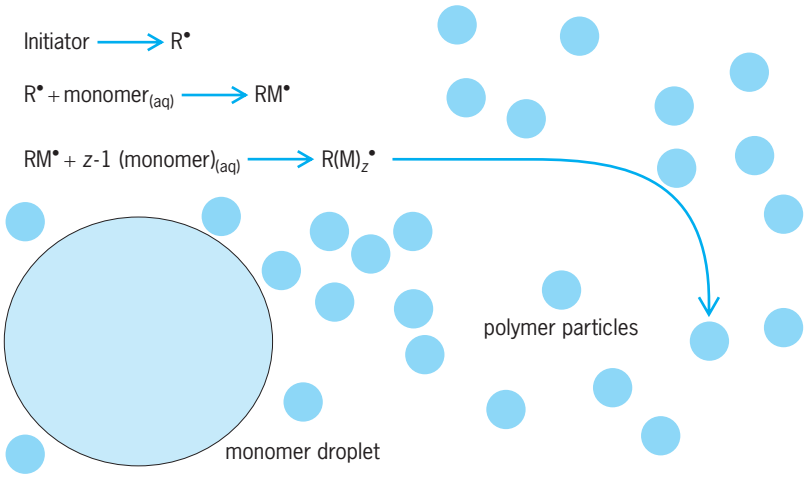

Fig. 2. Entry in emulsion polymerization.

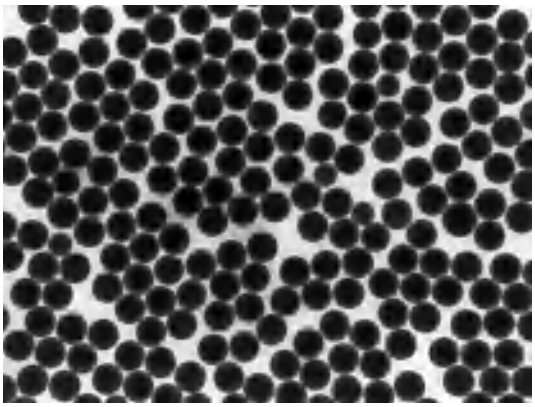

Fig. 3. Spontaneous packing of low-polydispersity poly(styrene) particles about 200 nm in diameter.

100 nanometers. Each particle contains a large number of polymer chains, but the size of a particle is not related to the molecular weight of the component polymer.

Benefits. Applications of emulsion polymerization include latexes for use as paints, adhesives, paper coatings, carpet backings, barrier products (for example, rubber gloves), and leather coatings. These latexes are frequently based on styrene, butadiene, acrylates, methacrylates, vinyl acetate, and acrylonitrile, with small amounts of water-soluble comonomers, such as acrylic acid, playing a crucial role in ensuring colloidal stability of the resulting latex. In addition to products used as a latex, there are many products where the polymer is isolated from the latex by coagulation: for example, styrene-butadiene rubber (SBR), used for tires, polytetrafluoroethylene (Teflon), and neoprene [poly(2-chlorobutadiene)].

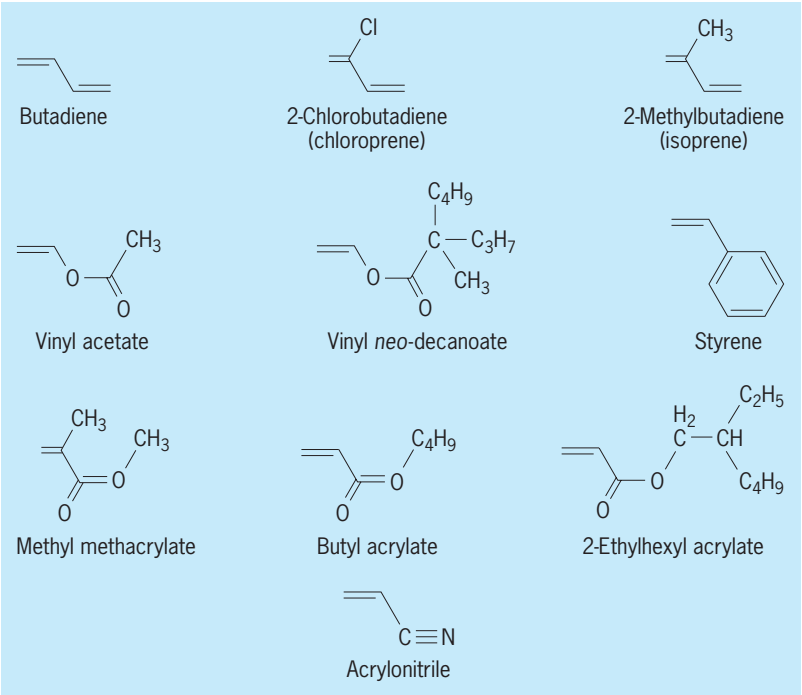

Fig. 4. Monomers used in emulsion polymerization.

Many advantages of emulsion polymerization arise from the compartmentalization of the growing chains, which reduces termination far below typical levels in bulk or solution polymerization to enable higher-molecular-weight products and faster reaction rates. Dispersion of the polymer in water keeps the viscosity low and makes the end product easy to handle, as opposed to the highly viscous and intractable materials produced by bulk or solution polymerization. As the process avoids use of flammable and toxic organic solvents, it has environmental and workplace safety benefits.

Recent developments. The mechanistic understanding of emulsion polymerization is now at a level where the size, composition, and topology of the colloidal particles can be controlled to achieve desired end-use properties.

The properties of a film cast from a latex are optimal when the particles can pack optimally, as is the case when they are all the same size (monodisperse) [**Fig. 3**]. While particle nucleation is now well understood, the process is sensitive to small disturbances (for example, changing amounts of oxygen present in the starting materials). Hence, the early stages of an emulsion polymerization in the absence of preformed latex may give an uncontrolled number of particles with a broad distribution of sizes. After nucleation, each particle grows at approximately the same rate in terms of volume, so the size distribution in terms of diameter narrows as the reaction proceeds. For this reason, industrial emulsion polymerizations are often seeded with a preformed latex of small diameter. Control of size and size distribution requires accurate measurement or prediction of particle-size distribution during polymerization, and this has been achieved by a number of groups in recent years by combining calorimetry or online spectroscopy with occasional sampling and novel process-control algorithms.

Most novel emulsion polymer products are likely to be made by combining, or more efficiently controlling, already common monomers, rather than by introducing new commodity monomers (**Fig. 4**). Possibilities to be considered are copolymers, polymer blends, and novel topologies. Indeed, the production of nanostructured particles has been routine in the industry for many years. *See* COPOLYMER.

In copolymers, different strategies must be used, depending on whether the two monomers should be distributed homogeneously within the particles or concentrated in separate domains. This can also be achieved by the online monitoring of the particle composition to optimize a mathematical model of the process that can be used for system control.

In some cases, it is desired to form two populations of particles with different copolymer compositions. For example, blocking of paints (the adhesion between two painted surfaces that can make a painted window frame difficult to open) can be minimized by generating a paint latex containing a population of small, relatively hard particles in addition to the larger, softer particles required for film formation. This may be done by varying the monomer feed

composition in a single, continuous feed process. Water-soluble monomers are frequently incorporated in latexes to generate surfactant, during the course of the reaction, which will be physically grafted to the surface of the polymer particles. Unlike conventional surfactants, these will not separate from the particle surfaces on freezing or shearing to give particle coagulation, nor will they diffuse out of the formed films into the surrounding environment.

Core-shell particles, in which a softer outer shell [for example, poly(butyl acrylate)] may give good film formation, while a harder core [for example, poly(methyl methacrylate)] provides strength, can be prepared by a two-stage controlled monomer feed.

The most rapid advances in emulsion polymerization in the near future are likely to arise from application of the methods of controlled radical polymerization (CRP). Reversible addition fragmentation chain transfer (RAFT), atom transfer radical polymerization (ATRP), and nitroxide-mediated radical polymerization have been used recently on the laboratory scale to give control of molecular weight and allow formation of novel block copolymers in emulsion polymerization systems. These methods use a reversible capping of the growing polymer radical to greatly reduce termination processes. Rather than the few growing chains of very short lifetime of conventional free-radical polymerization, CRP gives many (slowly) growing chains of very long lifetime. This leads to a much narrower distribution of molecular weights and the ability to create a range of polymer such as blocks and combs. Recent work suggests the practicality of eliminating all surfactant from emulsion polymerization through use of an aqueous-phase RAFT agent, and being able to make polymer chains of any desired architecture (**Fig. 5**). In the first stage, a "living" (termination-free) aqueous-phase polymer is generated, which reacts with the hydrophobic monomer added in the second phase to form polymeric surfactant molecules. These form rigid micelles which grow into polymer particles by continued feeding of the hydrophobic monomer. It has been suggested that because the individual polymer molecules in such a system are irreversibly attached to the micelles, unlike conventional surfactants which undergo rapid exchange, the number of micelles originally present can be correlated directly to the number of polymer particles. Thus any desired polymer microstructure can be built up. These methods also provide a pathway for incorporating specific functionalities on the surface of a polymer particle, for instance, to attach antibodies for use in latex agglutination tests in biomedicine. *See* CONTROLLED/LIVING RADICAL POLYMERIZATION; MACROMOLECULAR ENGINEERING.

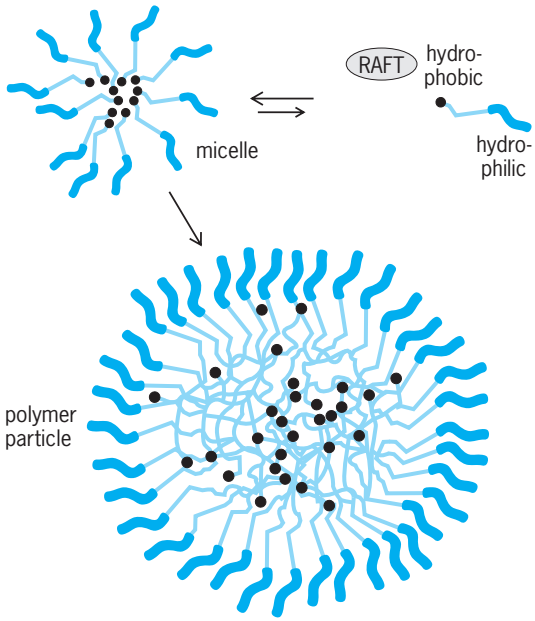

Fig. 5. Amphiphilic RAFT in emulsion polymerization.

Robert G. Gilbert; Christopher M. Fellows

Bibliography. C. Burguiere, C. Chassenieux, and B. Charleux, Characterization of aqueous micellar solutions of amphiphilic block copolymers of poly(acrylic acid) and polystyrene prepared via ATRP: Toward the control of the number of particles in emulsion polymerization, *Polymer*, 44:509–518, 2002; C. J. Ferguson et al., Effective ab initio emulsion polymerization under RAFT control, *Macromolecules*, 35:9243–9245, 2002; R. G. Gilbert, *Emulsion Polymerization: A Mechanistic Approach*, Academic Press, London, 1995; K. Matyjaszewski, New (co)polymers by atom transfer radical polymerization, *Macromol. Symp.*, 143:257–268, 1999; M. Vicente et al., Control of microstructural properties in emulsion polymerization systems, *Macromol. Symp.*, 182:291–303, 2002.

Enantiornithes

An extinct and specialized group of land fossil birds that lived during the Cretaceous Period. Enantiornithes represents the most diverse avian group in the Mesozoic. This group was not discovered until 1981, at which time some very unusual, isolated bones from the Late Cretaceous of Argentina were recognized as belonging to a new group of birds. Enantiornithine birds are now known to be widely distributed, with remains from Argentina, North America, Mexico, Mongolia, Australia, Spain, and China. Because their anatomy includes a mixture of specialized and primitive features, they are thought to represent an evolutionary side branch in early avian evolution that became extinct at the end of the Cretaceous and left no living descendants. *See* AVES; CRETACEOUS; FOSSIL BIRDS; MESOZOIC.

Characteristics. Enantiornithine birds are readily distinguishable from living birds. The most notable characteristic is the articulation between the scapula (shoulder blade) and coracoid (a bone connecting the scapula and the sternum). In enantiornithines, the scapula has a fossa for articulation with a process on the coracoid, whereas modern birds have the opposite arrangement in which the coracoid bears the fossa and the process is found on the scapula.

***Protopteryx*, the most primitive known enantiornithine bird, collected from the Early Cretaceous of Hebei Province, northern China.**

The group name refers to this difference: the Greek word "enanti" means directly opposite, and therefore Enantiornithes means "opposite birds." Enantiornithes is a natural (monophyletic) group, which is united by the shared possession of several unique features, including the characteristic contact between the scapula and coracoid, the distal (lower) end of the third metacarpal of the hand extending markedly past that of the second metacarpal, and a furcula (wishbone) with a long process (hypocleidium) that projects forward.

Enantiornithine birds were good fliers, possessing a more advanced flight apparatus than *Archaeopteryx*, including a short skeletal tail (comprising a few free vertebrae and a pygostyle or "parson's nose"), a sternum with a keel, and a wing that was generally similar to that of modern birds. Enantiornithines also possessed asymmetric flight feathers and the alula (bastard wing), indicating that they had acquired sophisticated flight capability. Conversely, some primitive avian characters, such as the presence of toothed jaws, were retained in many enantiornithes (although at least one form, *Gobipteryx* from Mongolia, completely lost the teeth). Primitive features were also retained in such structures as the pelvis and foot of enantiornithines.

Fossil record. Enantiornithine birds were initially described on the basis of often fragmentary Late Cretaceous forms such as *Enantiornis*, *Alexornis*, and *Gobipteryx*. However, the recognition of more complete Early Cretaceous enantiornithines has allowed significant progress to be made in our understanding of these birds during the last decade. Early Cretaceous enantiornithines have been discovered in Spain, China, Australia, and Russia. These early forms are generally small—much smaller than other contemporaneous birds such as *Confuciusornis* and *Jeholornis*. The proportions of the hindlimbs and toes, and the presence of large, curved toe claws, indicate that enantiornthines were mainly arboreal forms. Most enantiornithines, such as *Protopteryx* (see **illustration**), probably fed on insects. However, there are some exceptions: for example, *Longipteryx* (from the Early Cretaceous of China) had an elongated beak, rather like a living kingfisher, which suggests that it ate fishes (although in the case of *Longipteryx*, densely packed teeth were also present). Another Chinese Early Cretaceous enantiornithine, *Longirostravis*, probably had a probing feeding behavior. Discoveries of hundreds of specimens of Early Cretaceous enantiornithines from northeastern China (referable to 12 genera) have shown that by this time Enantiornithes had significantly differentiated not only in morphology, size, and flight capability but also in terms of dietary adaptation. *See* CONFUCIUSORNITHIDAE.

Zhonghe Zhou

Bibliography. L. M. Chiappe and C. A. Walker, Skeletal morphology and systematics of the Cretaceous euenantiornithes (Ornithothoraces: Enantiornithes), in L. M. Chiappe and L. M. Witmer (eds.), *Mesozoic Birds above the Heads of Dinosaurs*, pp. 240–267, University of California Press, Berkeley, 2002; A. Feduccia, *The Origin and Evolution of Birds*, 2d ed., Yale University Press, New Haven, 1999; L. D. Martin, The origin and early radiation of birds, in A. H. Bush and G. A. Clark, Jr. (eds.), *Perspectives in Ornithology*, pp. 291–338, Cambridge University Press, 1983; C. Walker, New subclass of birds from the Cretaceous of South America, *Nature*, 292:51–53, 1981; Z. Zhou, The origin and early evolution of birds: Discoveries, disputes, and perspectives from fossil evidence, *Naturwissenschaften*, 91:455–471, 2004.

Encalyptales

An order of the true mosses (subclass Bryidae) that grow in dull, dark tufts on soil or soil-covered rock, generally in calcareous areas. The Encalyptales consist of a single family and two genera, the better known being *Encalypta*, the extinguisher moss, so called because of its long calyptra of candle-snuffer form. The leaves resemble some of the Pottiales in shape and papillosity, but the ladderlike thickenings of the basal cells and the peristome characters indicate a fairly distant relationship.

Encalyptales are characterized by broad, papillose leaves and erect capsules covered by very long calyptrae. The stems are erect and simple or forked with folded, incurved leaves. Leaves are broadly acute to rounded and often abruptly short-pointed to

hair-pointed with a strong midrib ending at or beyond the leaf tip. The leaf margins are often revolute below. The upper cells are generally short and opaque because of numerous C-shaped papillae; the basal cells are smooth and rectangular with thin sidewalls and thickened, reddish crosswalls. The sporophytes are terminal, with elongate setae, and erect, cylindric, and ribbed capsules. The operculum is long-rostrate, and the peristome variable in structure. The spores are often large and quite various in sculpturing. The calyptrae are very long-mitrate and long-beaked, with the base often lobed or fringed. The chromosome number is usually 13 or multiples of 13, although 12 and 14 have also been reported. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA; POTTIALES. Howard Crum

Bibliography. D. G. Norton, A revision of the Encalyptaceae (Musci), with particular reference to the North American taxa, *J. Hattori Bot. Lab.*, 53:365–418, 1982, 54:357–532, 1983; D. H. Vitt and C. D. Hamilton, A scanning electron microscope study of the spores and selected peristomes of the North American Encalyptaceae (Musci), *Can. J. Bot.*, 52:1973–1981, 1974.

Endangered species

A species that is in danger of extinction throughout all or a significant portion of its range. "Threatened species" is a related term, referring to a species likely to become endangered within the foreseeable future.

Species diversity. The term "species" is defined in different ways. Most commonly, and for purposes of protecting from extinction, species are defined using the biological species concept: a species is a population or series of populations within which there is a significant amount of gene flow under natural conditions. Species extinction is a natural process. In fact, less than 1% of all species that ever existed are present today. Throughout the history of life, there have been certain periods of time with high extinction rates. For example, many dinosaur species became extinct within a short time, along with many other types of species, at the end of the Cretaceous Period. These periods of mass extinction have occurred five times. Currently, there is wide agreement among biologists that human activity has created a sixth mass extinction spasm, which continues today. According to some estimates, at least one in eight plant species and one in four mammal species are threatened with extinction.

The regions with the highest number of extinctions and endangered species are those with the greatest species diversity. Species diversity refers to the abundance of species in a given area. In general, larger areas contain greater species diversity. Species diversity tends to follow geographic patterns: there are more species at lower latitudes and altitudes. Tropical rainforests in countries such as Brazil and Indonesia harbor most of the world's species. In addition, many unique species live on islands, particularly tropical islands. Island species often exist under special conditions where they have comparatively few competitors and predators. They are restricted to a small range and commonly have small populations. These endemic species are particularly susceptible to extinction, because their ranges and population sizes are small.

Causes of species loss. The main factors that cause species to become endangered (and that have led to the current mass extinction spasm) are habitat destruction, invasive species, pollution, and overexploitation.

Habitat destruction. Habitat destruction is the single greatest threat to species around the globe. Natural habitat includes the breeding sites, nutrients, physical features, and processes such as periodic flooding or periodic fires that species need to survive. Humans have altered, degraded, and destroyed habitat in many different ways. Logging around the world has destroyed forests that are habitat to many species. This has a great impact in tropical areas, where species diversity is highest. Although cut forests often regrow, many species depend upon old-growth forests that are over 200 years old; these forests are destroyed much faster than they can regenerate. The threatened northern spotted owl, which lives in old-growth forests in the Pacific northwest of the United States, is a good example of a species that needs old-growth forests for the small mammals that they eat and for nest sites.

Agriculture has also resulted in habitat destruction. For example, in the United States, tallgrass prairies that once were home to a variety of unique species have been almost entirely converted to agriculture. Less than 2% of tallgrass prairies remain. Housing development and human settlement have cleared large areas of natural habitat. The construction of roads (which is associated with logging, mining, and housing developments) is a major threat to species, because animals are killed by vehicles or eventually avoid areas with too many roads and human activity. Mining has destroyed habitat because the landscape often must be altered in order to access the minerals (creating huge pits, removing mountaintops, or stripping off the top layer of large swaths of land). Mining also creates pollution from procedures for extracting minerals (for example, cyanide is used to extract gold). Finally, water development, especially in arid regions, has fundamentally altered habitat for many species. Dams change the flow and temperature of rivers and block the movements of species up and down the river. Also, the depletion of water for human use (usually agriculture) has dried up vegetation along rivers and left many aquatic species with insufficient water.

Invasive species. The invasion of nonnative species is another major threat to species worldwide. Invasive species establish themselves and take over space and nutrients from native species. Because they are transplanted, these invasive species often live in the absence of their natural predators and competitors. Invasive species are especially problematic for island species, which often do not have defensive mechanisms for the new predators or competitors.

For example, the brown tree snake was introduced to Guam and nearby islands. The snake multiplied vastly and ate so many birds and other animals that most of the bird species on those islands became extinct. Habitat destruction and invasion of nonnative species can be connected in a positive feedback loop: when habitat is degraded or changed, the altered conditions which are no longer suitable for native species can be advantageous for invasive species. In the United States, approximately half of all endangered species are adversely affected by invasive species.

Pollution. Pollution directly and indirectly causes species to become endangered. In some cases, pesticides and other harmful chemicals are ingested by animals low on the food chain. When these animals are eaten by others, the pollutants become more and more concentrated, until the concentration reaches dangerous levels in predators and omnivores. These high levels cause reproductive problems (birds with high levels of the pesticide DDT cannot reproduce) and sometimes death. In addition, direct harm often occurs when pollutants make water uninhabitable. Agriculture and industrial production cause chemicals such as fertilizers and pesticides to reach waterways. Lakes have become too acidic from acid rain, which contains previously airborne chemicals. Other human activities such as logging, grazing, agriculture, and housing development cause siltation (another form of pollution) in waterways. Largely because of this water pollution, two out of three freshwater mussel species in the United States are at risk of extinction. Pollution has also led to other phenomena which present risks to species. Most scientists agree that pollution is also causing climate change (often called global warming), as well as greater exposure to ultraviolet radiation from ozone-layer depletion.

Overexploitation. Many species have become endangered or extinct from killing by humans throughout their ranges. For example, the passenger pigeon, formerly one of the most abundant birds in the United States, became extinct largely because of overexploitation. This overexploitation is especially a threat for species that reproduce slowly, such as large mammals and some bird species. Overfishing by large commercial fisheries is a threat to numerous marine and fresh-water species.

Strategies for protection. Habitat destruction and other threats to species worldwide increased in the twentieth century; steps are under way to reduce those threats. Many people have realized that endangered species can signal current or eventual threats to human health and safety. Efforts to save species focus on ending exploitation, halting habitat destruction, restoring habitats, and breeding populations in captivity.

In the United States, the Endangered Species Act of 1973 protects endangered species and the ecosystems upon which they depend. Internationally, endangered species are protected from trade which depletes populations in the wild, through the Convention on International Trade in Endangered Species (CITES). Over 140 member countries act by banning commercial international trade of endangered species and by regulating and monitoring trade of other species that might become endangered. For example, the international ivory trade was halted in order to protect elephant populations from further depletion. In addition, the International Union for the Conservation of Nature identifies which species are in danger of extinction and initiates international programs to protect them.

Typically, the first step is identifying which species are in danger of extinction throughout all or part of their range and adding them to an endangered species list. In the United States, species are placed on the endangered species list if one or more factors puts it at risk, including habitat destruction or degradation, overutilization, disease, and predation. In addition, there is an assessment of whether the species is otherwise protected from those threats. Subspecies and distinct populations of species can be listed as endangered as well. For example, the Arizona population of the cactus ferruginous pygmy owl is endangered in the United States, even though this subspecies is more abundant across the border in Mexico. The International Union for the Conservation of Nature places species in categories of endangerment based on various factors, including population size, extent of population decline, and predicted population decline, either in population or area of occupancy.

From looking at patterns of where endangered species exist in the United States, it is clear that areas with high species diversity and high human population tend to have the most endangered species. Florida and California contain the most endangered species of all the contiguous 48 states. In addition, Hawaii has more endangered species than any other state. Hawaii, like other islands, has a diversity of unique species that occur nowhere else in the world. These species are also highly susceptible to endangerment because they tend to have small population sizes, and because they are particularly vulnerable to introduced competitors, predators, and disease.

Once a species is determined to be endangered, the U.S. Endangered Species Act protects and aims to restore endangered populations through two main strategies: halting exploitation, and habitat protection and restoration. It is prohibited to kill, harm, or harass an endangered species. In addition, habitat destruction is prohibited if it results in impairment of the animal's ability to forage, breed, or seek shelter. To complement prohibitions on harming endangered species, the act requires that critical habitat for a species be delineated and prioritized for protection. The act also requires a recovery plan for each species, which defines the status of the species, threats, measures to be taken to restore its numbers, and goals to be met so that delisting can occur. There are increasing numbers of cities and counties that develop conservation plans for one or more endangered species. These plans contain information on important habitat that should be preserved, as well as permits for housing development, agriculture, and

other activities to occur elsewhere. These protective efforts for endangered species also serve to protect other species in the same areas that might otherwise become endangered. All of these methods, from prohibitions to large-scale planning, are aimed at recovering species so that they are no longer in danger of extinction.

For many endangered species, a significant captive population exists in zoos and other facilities around the world. By breeding individuals in captivity, genetic variation of a species can be more easily sustained, even when the species' natural habitat is being destroyed. Some species exist only in captivity because the wild population became extinct. For a few species, captive individuals have been reintroduced into natural habitat in order to establish a population where it is missing or to augment a small population. Depending on the species, reintroduction can be very difficult and costly, because individual animals may not forage well or protect themselves from predators. For example, the golden lion tamarin has been nearly extirpated from Atlantic coastal rainforests in Brazil. These monkeys have been bred in captivity and trained for reintroduction at a nature reserve in Brazil. Although some individuals do not survive reintroduction, others are able to breed, and a viable wild population is being reestablished. *See* ECOLOGY; EXTINCTION (BIOLOGY).

Laura Hood Watchman

Bibliography. C. H. Flather, M. S. Knowles, and I. A. Kendall, Threatened and endangered species geography, *BioScience*, 48:365–375, 1998; G. K. Meffe and C. R. Carroll, *Principles of Conservation Biology*, 2d ed., Sinauer Associates, Sunderland, MA, 1997; E. O. Wilson, *The Diversity of Life*, Harvard University Press, Cambridge, MA, 1992.

Endocrine mechanisms

Those regulatory phenomena in animals or plants which involve, as intermediaries, one or more hormones.

Hormones

These are specific chemical entities, secreted by specialized cells, tissues, or organs and transported in solution in body fluids to other cells, tissues, or organs where they exert a specific physiological action at low concentrations (see **table**). The concept of a hormone was introduced in 1902 by W. Bayliss and E. Starling in a study of the coordination of digestive secretion; the term hormone was introduced 3 years later by Starling. The cells which form hormones are called endocrine because they pass their secretion into the blood or body fluid in contrast to exocrine cells, which secrete into the digestive tract or other regions considered outside the boundary of the body proper. *See* GLAND; HORMONE.

Regulation. Regulatory phenomena are those whereby the various activities of the component parts of the organism are modified so that they contribute to a coherent pattern of activity of the organism as a whole. The fact that hormones, like vitamins, exert their actions at very low concentrations has led to the suggestion that both kinds of substances act as catalysts. *See* VITAMIN.

Action mechanisms. Several mechanisms of action have been demonstrated for hormones, although there is often uncertainty as to whether a particular action is primary or is a consequence of some other more fundamental action of the hormone. Hormones may change the properties of the plasma membrane or of some internal membrane of the cell, induce a change in the conformation of a protein molecule and thus alter its activity as an enzyme, serve as a coenzyme in an enzymatic reaction, induce the formation of a coenzyme or other substance modifying enzyme action, or induce or suppress the synthesis of a protein.

Many hormones have been shown to act on target cells through the "second messenger" mechanism. When the hormone molecule is bound to the surface of the target cell by combination with a specific receptor molecule on the surface, this combination stimulates an increase in activity of the enzyme adenyl cyclase on the inner surface of the membrane. This enzyme catalyzes the conversion of adenylic acid (adenosine monophosphate, AMP) to cyclic 3′,5′-adenylic acid by attaching two free sites of the phosphate radical through esterification to form a ring structure. Cyclic AMP (cAMP) is formed by a great variety of cells in response to specific stimuli, and activates certain enzymes within a cell, thus increasing its normal secretory or other activity. The response of the cells of the thyroid gland to the thyroid-stimulating hormone of the pituitary is also modulated by a prostaglandin, a member of a ubiquitous class of lipoid substances, first isolated from the prostate gland but since shown to occur in several different forms and to act as part of the regulatory mechanisms of cells.

Any of these actions will alter the rate and direction of cellular activities and may thus result in widespread modification of physiological processes in the organism. Hormones appear to have more or less specific target organs or target cells; some hormones seem to act exclusively on one type of cell in a specific organ, while others act on many kinds of cells in many organs of the body. The specificity of target cells which react only to a certain hormone has been traced to the occurrence, at the cell surface or in the cytoplasm, of specific receptor molecules, each reacting only with a certain hormone. The hormone may be bound to the cell surface and exert its action there, as noted above, but other hormones, notably the steroids, enter the cell passively, and are then bound to specific receptors in the cytoplasm. In this latter case, they are carried, in combination with the receptor, to the cell nucleus, where they presumably effect their specific action by influence on the control function of the chromosomes.

There is a set of hormone-releasing factors formed by certain cells in the brain; each of these acts on a single type of cell in the anterior pituitary of a mammal, causing release from the pituitary cell of a specific hormone. On the other hand, hormones

Properties of hormones of vertebrate animals, unless noted otherwise

Names and abbreviations	Source; stimulus for secretion or release	Chemical nature	Target; receptor if known	Action or function
Acetylcholine (Ach)	Cholinergic neuron endings	Ester, $CH_3COOCH_2N(CH_3)_3$	Postsynaptic neuron or effector (muscle)	Transmitter; in mammals excites skeletal muscle, slows heartbeat, and stimulates intestinal contractions
Adrenocorticotropic hormone (ACTH)	Anterior pituitary; CRF	Polypeptide	Adrenal cortex	Stimulates secretion of corticosterone and other steriods
Aldosterone	Adrenal cortex; angiotensin	Steroid 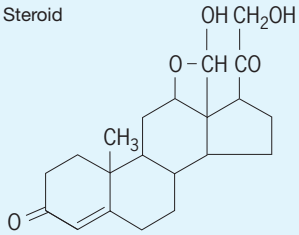	Cells of walls of kidney tubules; receptors are intracellular	Stimulates active reabsorption (retention) of salt (Na^+)
Androgen (testosterone)	Interstitial cells of testes; ICSH	Steroid 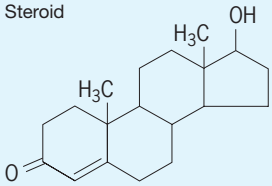	Male sex organs and ducts, muscle, and brain; receptors are intracellular	Stimulates development and maintains male secondary sex structures and behavior
Angiotensin	Blood, through enzymatic action of renin	Polypeptide	Muscle of arterioles, adrenal cortex	Stimulates contraction, increasing blood pressure; stimulates secretion of aldosterone
Auxins	Growing tips of plant shoots	Several known, including indoleacetic acid 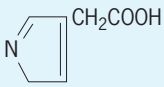	Growing plant cells and meristems	Stimulate growth of cells in roots and shoots, and formation of roots; inhibit development of buds in shoots
Calcitonin	Thyroid gland	Polypeptide	Bone	Inhibits reabsorption, promotes formation, decreasing Ca^{2+} in blood
Corticosterone, cortisone	Adrenal cortex	Steroids 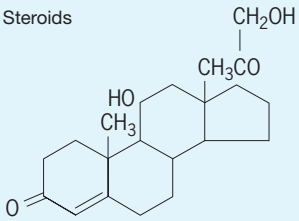	Many body cells; receptors are intracellular	Stimulate events of general adaptation syndrome in response to physical or emotional stress: breakdown of protein, synthesis of glycogen from amino acids, increased activity of immune mechanisms
Corticotropin-releasing factor (CRF)	Hypothalamus, posterior pituitary	Polypeptide	ACTH-secreting cells, anterior pituitary	Release of ACTH into blood
Corticotropin-release-inhibiting factor (CRIF)	Hypothalamus, posterior pituitary	Polypeptide	ACTH-secreting cells, anterior pituitary	Inhibits release of ACTH
Crustecdysone	Crustacean molt gland, Y organ	Steroid	Epidermis and other tissues concerned in periodic growth and development	Stimulates molting and associated processes
Ecdysone	Thoracic gland of larval and pupal stages of insects	Steroid 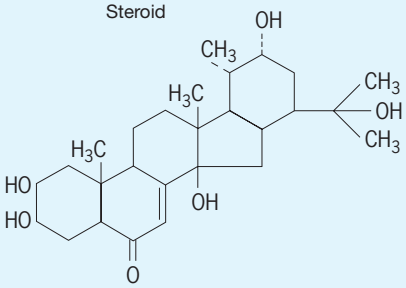	Epidermis and other tissues concerned in periodic growth and development	Stimulates molting and associated processes

Properties of hormones of vertebrate animals, unless noted otherwise (*cont.*)

Names and abbreviations	Source; stimulus for secretion or release	Chemical nature	Target; receptor if know	Action or function
Enterocrinin	Cells of small intestine, in response to high concentrations of sugar and fat	Polypeptide	Musculature of stomach	Inhibits gastric contractions, depresses appetite
Epinephrin (adrenalin)	Chromaffin cells of adrenal medulla, in response to nervous stimulation	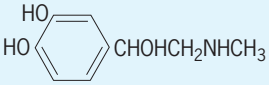	Liver, heart, blood vessels, other tissues	Increases conversion of glycogen to glucose, blood sugar, rate and force of heart; dilates blood vessels of skin; erects hairs
Estrogens: estradiol, estriol, estrone	Ovarian follicles, FSH	Steroids 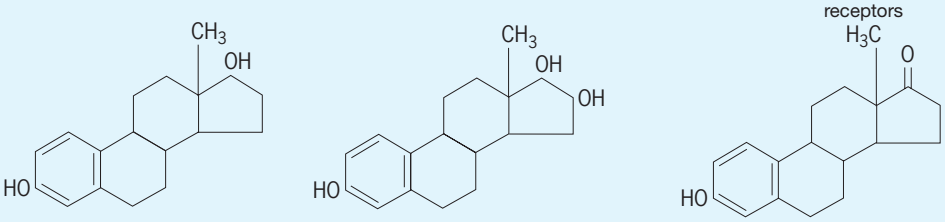	Endothelium (inner lining) of uterus and female sex organs and dusts; source of FSHRF; intracellular receptors	Stimulate proliferation of mature endothelium; promote development of secondary female sex characteristics; promote estrus in some mammals; inhibit secretion of FSHRF
FSH-releasing factor (FSHRF)	Hypothalamus, posterior pituitary	Polypeptide	FSH-secreting cells of anterior pituitary	Promotes release of FSH
Gamma-aminobutyric acid (GABA)	Some presynaptic neuron endings	Amico acid, $H_2NCH_2CH_2CH_2$ COOH	Postsynaptic neurons	Transmitter substance
Gastrin	Wall of stomach in response to filling	Polypeptide	Secretory cells of stomach	Stimulates production of gastric juice
Gibberellins	Plant tissues	Derivatives of gibberellic acid 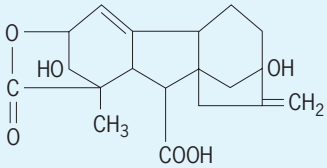	Growing cells of plants, leaves, fruits	Stimulate cell elongation, growth of leaves and fruits
Glucagon	Alpha cells of pancreatic islet tissue, in response to low blood sugar	Protein	Liver, kidney, intestine	Stimulates formation of glucose from glycogen; stimulates elimination of salt; inhibits contraction
Histamine	Wall of stomach; injured cells in general	Decarboxylated amino acid 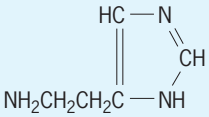	Stomach cells, arterioles, capillaries	Stimulates production of gastric juice; dilates, increasing blood flow; increases permeability and exudation of fluid
Hydrocortisone, cortisol	Adrenal cortex, ACTH	Steroid 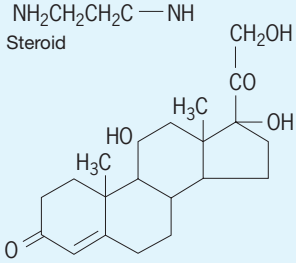	Tissues generally	Effects are intermediate between those of aldosterone and corticosterone
Insulin	Beta cells of pancreatic islet tissue	Protein	Muscle, adipose tissue	Immediate effects include increased uptake of glucose, decreasing blood sugar (surface receptors); long-term increase in utilization of glucose and synthesis of lipids (intracellular receptors)
Intermedin	Intermediate pituitary	Polypeptide	Pigment cells (chromatophores) in skin of lower vertebrates	Disperses melanin, darkening body surface
Interstitial-cell-stimulating hormone (ICSH) or luteinizing hormone (LH)	Anterior pituitary, ICSHRF	Protein	Interstitial cells of testes, ovarian follicle	Stimulates production of androgen; promotes transition to corpus luteum after ovulation

Properties of hormones of vertebrate animals, unless noted otherwise (*cont.*)

Names and abbreviations	Source; stimulus for secretion or release	Chemical nature	Target; receptor if known	Action or function
Isotocin	Posterior pituitary of fishes	Peptide: 4-serine, 8-isoleucine oxytocin	Testes, gills	Stimulates spawning; increases uptake of Na^+ in fresh water; turnover of Na^+ in sea water
Kinins (kinetin)	Plant tissues	Kinetin 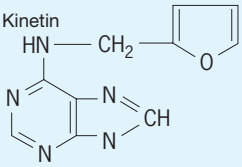	Meristematic tissues	Increase rate of cell division
Melatonin	Pineal body	Dimethyl serotonin 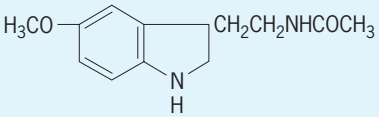	Pigment cells of skin of lower vertebrates	Concentrate melanin, lightening body surface
Mesotocin	Posterior pituitary of amphibians and reptiles	Peptide: 8-isoleucine oxytocin		Uncertain; may be similar to that of vasotocin
Neotenin (juvenile hormone)	Corpora allata of insects	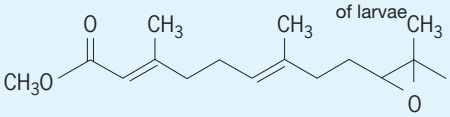	Epidermis and internal tissues of larvae	Prevents development of adult characteristics in larval molts
Noradrenalin (Norepinephrin)	Postganglionic neuron endings of sympathetic nervous system	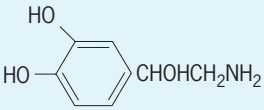	All structures with sympathetic innervation	Generally similar to those of epinephrin, antagonistic to those of acetylcholine
Oxytocin	Posterior pituitary	Peptide 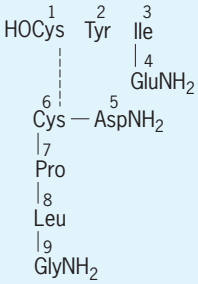	Kidney of lower vertebrates, uterus and mammary glands of mammals	Antidiuretic; stimulates contractions of uterus; stimulates flow of milk
Pancreozymin	Walls of duodenum, in response to HCl	Polypeptide	Exocrine cells of pancreas	Secretion and release of digestive enzymes
Parathormone	Parathyroid glands	Protein	Bone, kidney	Stimulates reabsorption of calcium; increases Ca^{2+} of blood; increase elimination of phosphates
Progesterone	Corpus luteum of ovary	Steroid 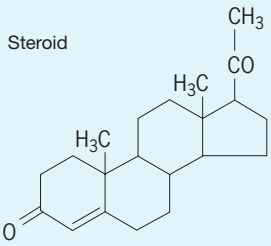	Endometrium of uterus	Maintains in secretory state, permitting implantation of embryo
Prolactin (luteotropic hormone)	Anterior pituitary	Protein	Corpus luteum	Maintains this structure and promotes secretion of progesterone
Renin	Juxtaglomerular cells of kidney	Protein	Precursor of angiotensin in blood	Liberates angiotensin by partial hydrolysis of precursor
Secretin	Walls of duodenum in response to HCl	Polypeptide	Exocrine cells of pancreas	Stimulates secretion of fluid containing salts, including $NaHCO_3$

Properties of hormones of vertebrate animals, unless noted otherwise (*cont.*)

Names and abbreviations	Source; stimulus for secretion or release	Chemical nature	Target; receptor if known	Action or function
Serotonin (SHT)	Central nervous system neurons	5-Hydroxy tryptamine 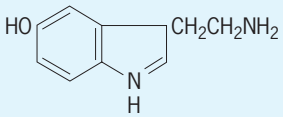	Postsynaptic neurons	Central transmitter substance; precursor of melatonin in pineal
Somatotropin (growth hormone, STH)	Anterior pituitary	Protein	Tissues in general, especially long bones and muscle	Promotes synthesis of protein and breakdown of fat, growth of young mammals; complemented by insulin
Somatotropin-releasing factor (STHRF)	Posterior pituitary	Polypeptide	Somatotropin-secreting cells of anterior pituitary	Promotes release of STH
Somatotropin-release-inhibiting factor (STHRIF)	Posterior pituitary	Polypeptide	Somatotropin-secreting cells of anterior pituitary	Inhibits release of STH
Thiamin	Leaves and seed germs of plants	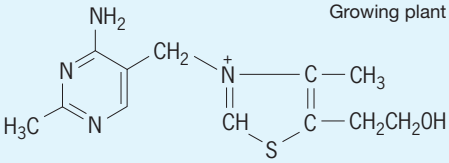	Growing plant tissues	Stimulates growth of roots and shoots; thiamin phosphate is coenzyme for carboxylation, must be supplied as a vitamin for animals
Thyrotropin (thyroid-stimulating hormone, TSH)	Anterior pituitary, TSHRF	Protein	Thyroid gland; intracellular receptors	Maintains gland and stimulates secretion
Thyrotropin-releasing factor (TSHRF)	Posterior pituitary; low levels of thyroid hormones	Polypeptide	Thyrotropin-secreting cells of anterior pituitary	Stimulates release of TSH
Thyrotropin-release-inhibiting factor (TSHRIF)	Posterior pituitary; high levels of thyroid hormones	Polypeptide	Thyrotropin-secreting cells of anterior pituitary	Inhibits release of TSH
Thyroxin (tetraiodo-thyronine, T4)	Thyroid gland	Iodinated amino acid 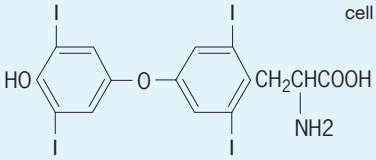	Cells generally (receptors in cell nuclei)	Stimulates metamorphosis in amphibians; stimulates metabolism in homeothermous animals (birds, mammals)
Triiodo thyronine (T3)	Thyroid gland	Iodinated amino acid 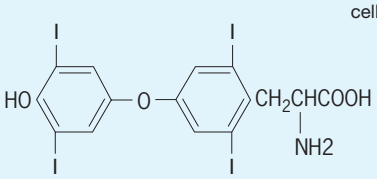	Cells generally (receptors in cell nuclei)	Stimulates metamorphosis in amphibians; stimulates metabolism in homeothermous animals (birds, mammals); the principal circulating thyroid hormone of mammals
Vasopressin (antidiuretic hormone, ADH)	Posterior pituitary, inhbited by increased internal fluid volume due to intake of fluid	Peptide, 3-phenylalanine, 8-arginine oxytocin or 3-phenylalanine, 8-lysine oxytocin	Kidney	Decreases volume of urine formed by increasing permeability of tubule walls and osmotic reabsorption of water
Vasotocin	Posterior pituitary of lower vertebrates	Peptide, 8-arginine oxytocin	Kidney, skin, and bladder wall of amphibians	Increases osmotic uptake of water from urine and medium by increasing permeability of membranes to water

like thyroxin or insulin act on many, perhaps all, cells in the body, and their actions on different cells may be distinct; thus insulin causes muscles to take up glucose from the blood and convert it to glycogen, and it causes cells of fatty tissues to liberate less fat into the blood.

Control demonstrations. Evidence for hormonal control of a process may take the form of demonstrations that (1) the process is influenced by events occurring elsewhere in the organism when there is no nervous connection involved; (2) surgical removal or pathological change in a particular organ or tissue is followed by changes in another organ or tissue elsewhere in the body; and (3) extracts of an organ or tissue, injected or otherwise administered, have effects opposite to those of surgical removal of the organ or tissue or similar to those of pathological hypertrophy of the tissue. The final proof of hormonal control, after points 1, 2, and 3 above have been established, comes with isolation, purification, and chemical identification of the hormone and the demonstration that the pure substance has the same effects as the extract of the tissue of origin. In some instances isolation and purification have preceded proof of hormonal function. Development of tracer techniques using radioisotopic or fluorescent labels and electron or fluorescence microscopy, of techniques of cell, tissue, and organ culture, and of extremely sensitive radioimmune assay techniques has greatly expanded the range of methods available for study of endocrine mechanisms, and has added much to knowledge of them.

The types of processes in which hormonal control is involved may be classified conveniently as coordinative, conservative, progressive, and cyclical. Any classification of hormones on the basis of mechanism of action is premature until more is known about such mechanisms.

Coordinative Control

This phenomenon was first demonstrated in the coordination of digestive secretion in the vertebrates, and has been much extended since the original discovery, by W. N. Bayliss and Ernest Starling, of the secretin mechanism. When the acidic contents of the stomach, during digestion, are ejected into the first part of the small intestine (duodenum), the hydrochloric acid of the gastric juice elicits liberation of the hormone secretin into the blood. When this reaches the exocrine cells of the pancreas, it stimulates them to form, and release into the ducts, a juice poor in enzymes but containing salts, including sodium bicarbonate, which neutralizes the gastric juice. The duodenal mucosa also forms pancreozymin, which stimulates liberation of digestive enzymes into the pancreatic secretion; cholecystokinin, which stimulates the gallbladder to contract and eject bile into the intestine; and enterocrinin, which inhibits gastric secretion. Contact of food with the stomach causes secretion. The effect of these hormonal secretions is to ensure a supply of digestive fluids at the time when food arrives in the appropriate region of the digestive tract and, in the case of enterocrinin, to terminate gastric secretion when intestinal digestion has begun. *See* DIGESTIVE SYSTEM.

Neurohumors. Hormonal substances formed by and liberated from special neurosecretory cells in nervous tissue (neurohumors) are often involved in coordinative control. In many and perhaps in all cases, the transmission of a state of excitation from nerve cell to nerve cell within the nervous system, or from nerve cell to muscle cell, involves a transmitter substance. The best known of these is acetylcholine, which has been identified as the transmitter between motor nerves and skeletal muscles, at parasympathetic nerve endings, between nerve cells of sympathetic ganglia in the vertebrates, and at some invertebrate nerve endings. At the vertebrate motor end plate in skeletal muscle, acetylcholine appears to act by increasing the permeability of the end plate membrane to salts; as a result, there is a flow of ions across the end plate which causes electrical depolarization sufficient to excite the muscle. The transmitter at postganglionic sympathetic fibers of vertebrates, such as those innervating the arterioles, is noradrenaline. Nerve endings at which adrenalinelike substances act as transmitters are called adrenergic. In some invertebrates, and possibly in vertebrates as well, certain nerve endings liberate gamma-aminobutyric acid (GABA) as a transmitter substance. There is growing evidence that GABA is a transmitter substance in the brains of vertebrates. Serotonin (5-hydroxytryptamine), first demonstrated as a transmitter substance in mollusks, is especially abundant in the pineal body in the roof of vertebrate brains, where it is the precursor of melatonin.

Color change. Control of color change in animals often involves neurohumors, and in crustaceans these substances act as true hormones, being liberated in one portion of the body and acting on structures in other parts of the body. Animals which can change color in response to visual and other stimuli, and thus match their body color or pattern to the background, do so by means of chromatophores. In crustaceans the movements of the pigments within the chromatophores are controlled by neurohumors formed in special groups of neurosecretory cells in the brain and eyestalk. Many of these cells have long extensions or axons terminating in the sinus gland, a small structure in the eyestalk. The secretions are apparently formed in the bodies of the neurosecretory cells and pass along the axon to be liberated from the ending in the sinus gland or elsewhere. None of the chromatophorotrophins, as these hormones are called, has been isolated, but partial purification has been accomplished, and it is clear that there are several substances, each active on one type of chromatophore. *See* CHROMATOPHORE.

Color change in the lower vertebrates is under a complex control which varies from species to species. In some, the only control is by means of intermedin, from the intermediate lobe of the pituitary gland. This substance has the effect of dispersing the black pigment melanin within the cells and thus darkening the body surface. Accumulating

evidence suggests that intermedin exerts this effect by changing the ionic balance of the cells, and the dispersion of pigment is a consequence of this change.

The pineal body, in the roof of a vertebrate brain, has a structure resembling the retina of the eye, and functions as an accessory photoreceptor, not quite a "third eye," in certain reptiles, where the skull over the pineal is relatively thin and translucent. A suspected endocrine function of the pineal—which Descartes considered, naively as it now seems, as the seat of the soul—has been verified by the demonstration that this body forms and releases into the blood the hormone melatonin. This substance, dimethyl serotonin, acts on melanophores (black pigment cells) in the skins of amphibians and reptiles to concentrate the melanin in the center of the cells, and thereby to lighten the body surface, acting thus in opposition to intermedin. These changes in color (or in shade) are adaptive in matching the shade of the animal to that of its background, and also in regulation of body temperature. A dark body surface absorbs sunlight, but also radiates heat, and a light surface absorbs and radiates less energy.

Intermedin and melatonin occur in birds and mammals, although these animals have no chromatophores and cannot change the color of their skin. Functions of these hormones in the higher vertebrates are not yet established. The secretory function of the pineal in mammals has been shown to be under the control of an adrenergic (with norepinephrine as transmitter) nerve tract originating in the hypothalamus, and the pattern of release of melatonin and of changes in the enzyme which forms it from serotonin suggests that melatonin may be concerned in regulation of the many circadian rhythms of activity and metabolism in mammals.

The effects of hormones are, in general, exerted slowly and in no pattern of local action other than that of the distribution of their targets. On the other hand, color change can be quite rapid in some lower vertebrates, and may occur in a variable pattern which matches the pattern of light and shade in the background. Professor Francis B. Sumner had a photograph of a flatfish which had matched the pattern of its skin to that of a checkerboard which had been placed underneath the fish. Rapid and patterned changes imply the action of the nervous system. *See* PROTECTIVE COLORATION.

In general, cholinergic nerves disperse pigment and adrenergic nerves concentrate it. Some vertebrates have only adrenergic nervous control; others have both.

Control of hormone secretion. Studies have demonstrated, in mammals, an elaborate system of controls centered in the hypothalamus, the lower portion of the midbrain, and acting primarily on the anterior pituitary.

Anterior pituitary. The anterior pituitary gland secretes six distinct hormones with varied actions. These are thyrotropin (TSH), which controls secretion of the thyroid hormones; adrenocorticotropin (ACTH), which controls secretion of steroids by the adrenal cortex; follicle-stimulating hormone (FSH); luteinizing hormone (LH), also known as the interstitial cell-stimulating hormone; prolactin, also known as the luteotrophic hormone (LTH)—which three hormones control secretion of sex steroid hormones by the gonads; and somatotropin (STH), which, besides its action on growth, controls secretion of insulin by the islet tissue of the pancreas. Each of these trophic (feeding, nourishing, or maintaining) or tropic (turning toward or acting upon) hormones acts to influence secretion of another hormone by its target cells. The secretion of the six hormones of the anterior pituitary has been traced to a specific cell type, one type for each hormone. *See* PITUITARY GLAND.

Neurosecretory activity. Control of these cells, in their function of secreting hormones, resides in a group of specialized neurosecretory cells located in the hypothalamus. These neurosecretory cells send fibers or axons into the stalk that attaches the pituitary gland to the base of the brain, along a ridge called the median eminence. In the stalk and the median eminence there is a specialized set of blood vessels arising from arteries supplying the brain. These arteries branch to form capillaries; as the blood flows through these capillaries, it picks up the products liberated by the endings of the fibers from the neurosecretory cells and carries them through a small "portal vein" directly to the anterior lobe of the pituitary gland.

Andrew Schally and Roger Guillemin, in work that earned them the 1977 Nobel prize, with their coworkers, demonstrated a set of releasing factors. These are specific peptide hormones formed in the hypothalamopituitary neurosecretory system; each of these hormones, on arrival in the anterior pituitary via the portal route, stimulates the release of a specific hormone from the cell type in which it is formed. In addition, release-inhibiting factors have been identified for most of the hormones. The entire set of systems and hormones assures the interaction between the nervous system and the endocrine system through the pituitary. This allows for rapid and integrated control of many aspects of body function, as well as for psychosomatic disorders when disturbances in brain function disrupt the integrated operation of the endocrine system and consequently normal body functions. The neurosecretory connection also takes part in the feedback control mechanisms by means of which the concentrations of various hormones in the blood are kept constant. *See* NEUROSECRETION.

Conservative Control

It is well known that hormones exert a conservative control of intermediary metabolism and of salt and water balance in mammals and other vertebrates.

Carbohydrate metabolism. The primary effect of insulin is to increase the utilization of glucose by the cells. This involves increased ability to transport glucose into the cells and increased synthesis of glycogen and oxidation of glucose. In many tissues, and especially in fatty tissues, the synthesis of fats from carbohydrate is increased, and the breakdown

of fats with the liberation of fatty acids into the blood is decreased. *See* INSULIN.

The pancreas also forms the hormone glucagon in the alpha cells of the islets. This hormone acts primarily on the liver, where it stimulates the breakdown of glycogen with consequent liberation of glucose into the blood; it also stimulates the synthesis of glycogen from amino acids. The action on glycogen breakdown involves the activation of the enzyme phosphorylase, which is the principal catalyst for the breakdown of glycogen. Glucagon and insulin are antagonists, with respect to the concentration of glucose in the blood (blood sugar); insulin decreases and glucagon increases the blood sugar. The normal level of sugar in the blood is controlled primarily by insulin, with glucagon serving to raise the level if it should fall below normal. *See* GLUCAGON; PANCREAS.

Experimental diabetes can be alleviated by surgical removal of the pituitary, and the injection of extracts of the anterior lobe of the pituitary into normal animals will temporarily increase the blood sugar (diabetogenic effect). The active principles of pituitary extracts act as antagonists to insulin, decreasing the utilization of glucose by the tissues. The growth hormone, somatotropin (STH), is the major insulin antagonist in pituitary extracts; this hormone also stimulates protein synthesis in the presence of insulin. Adrenocorticotropin (ACTH) also acts as an insulin antagonist in the intact animal, but this action is largely indirect and is exerted through stimulation of the cortical cells of the adrenal glands to form and liberate steroid hormones such as corticosterone and hydrocortisone. These steroids stimulate protein breakdown, the synthesis of liver glycogen from protein (gluconeogenesis), and the breakdown of liver glycogen to glucose (glycogenolysis) and hence cause an increase in the blood sugar. Removal of, or damage to, the adrenals is followed by disappearance of glycogen from the liver and inability of the animal to maintain blood sugar levels in fasting or in vigorous activity. *See* DIABETES.

In the normal animal these factors are in balance, and the blood sugar is held within well-defined limits. Increased levels of sugar in the blood act directly on the beta cells of the pancreas, stimulating secretion of insulin. The level at which this action takes place, however, is determined in part by STH; in the presence of STH, the beta cells secrete insulin at lower levels of blood sugar. Glucagon inhibits secretion of insulin. Secretion of glucagon occurs as a direct response of the alpha cells to decreased levels of sugar in the blood. The levels of adrenal steroids in the blood are maintained relatively constant, in normal health, by the ACTH feedback control, in which the rate of secretion of ACTH is increased when the levels of cortical steroids in the blood falls, and is decreased when the level increases. *See* CARBOHYDRATE METABOLISM.

Water balance. In mammals water balance is in part controlled by the posterior pituitary.

Pituitary gland. The injection of extracts of the posterior lobe of the pituitary, or of vasopressin (antidiuretic hormone, ADH), causes a distinct decrease in the volume of urine formed by a normally hydrated animal. Damage to certain neurosecretory cells in the hypothalamus, or section of the axons connecting these cells to the posterior pituitary, has the opposite effect, inducing the condition diabetes insipidus, in which large quantities of dilute urine are produced. This condition is entirely unrelated to the more common condition known as diabetes mellitus or sugar diabetes, which results from a deficiency of insulin secretion. ADH is normally released continually, and its release is inhibited through nervous action when the volume of the internal body fluid is increased, or the salt concentration of the fluid decreased, usually by intake of large quantities of liquids. The action of diuretic substances such as alcohol or caffeine or of some medications used to stimulate urine flow is exerted by inhibition of secretion of ADH. *See* EXCRETION.

Thyroid and parathyroid glands. The calcium and phosphate balance of blood, bone, and other tissues is controlled by the parathyroid and thyroid glands. Removal of the parathyroids is followed by a decrease in the urinary excretion of phosphate, increased calcium deposition in bone, and decreased levels of blood calcium. Nervous and muscular excitation are increased in consequence of the decreased calcium concentration, and tetany and death follow. Injection of parathormone decreases phosphate excretion and causes mobilization of calcium and phosphate from bone, increasing the blood calcium level. *See* PARATHYROID GLAND; PARATHYROID HORMONE.

The normal level of parathormone secretion is probably determined by the level of calcium in the blood, acting directly on the parathyroids. In addition to the well-known secretion of thyroxine, the thyroid gland of the vertebrates contains specialized cells that secrete a hormone known as calcitonin; this substance, when injected, causes a decrease in the calcium level of the blood and hence is antagonistic to parathormone. Calcitonin appears to act primarily by preventing the reabsorption of calcium and phosphate from bone into the blood and secondarily by increasing the loss of calcium in the urine. The stimulus for release of calcitonin seems to be an increase in levels of calcium in the blood. It acts rapidly, and for a short period of time, and consequently serves as a "fine control," with the long-term regulation being dependent upon the parathyroids. Other functions of the thyroid, in control of development and of metabolism, have been known much longer than that of control of calcium metabolism. Of two known products of secretion by this gland, thyroxine or tetraiodo thyronine (T4) and triiodo thyronine (T3), the latter appears to be the circulating form of thyroid hormone in mammals. Injected T4 has been shown to be converted to T3 in the body. *See* THYROID GLAND; THYROID HORMONES.

Total metabolic level. The primary effect of thyroid hormone in the warm-blooded or homeothermic animals—birds and mammals—is that of maintaining normal total metabolism, oxygen consumption, and heat production, and indirectly contributes to a constant body tempera-

ture. Thyroidectomy (removal of the thyroid gland) is followed by a decrease in metabolic rate, and injection of thyroid extract or of T4 or T3 increases this rate. Thyroidectomy has also been shown to decrease the resistance of animals to cold. *See* METABOLISM; THERMOREGULATION.

Integrity and normal function of the thyroid gland depends on the thyrotropic hormone (TSH), secreted in the anterior pituitary and released under the influence of the thyrotropin-releasing factor (TSHRF), from the hypothalamopituitary neurosecretory system. TSHRF is secreted when the level of circulating thyroid hormone falls below a set point. The response of the thyroid gland to TSH depends upon surface receptors and, through these, formation within the cells of cyclic AMP. The familiar analgesic aspirin has been shown to decrease blood levels of thyroid hormone, independent of TSH.

The receptors of target cells to T3 have been demonstrated in cell nuclei. This suggests that the thyroid hormones exert their actions on developmental processes and on metabolic rate through effects on cellular control processes, rather than by a direct effect on oxidation processes as such, as was formerly supposed.

Sodium ion concentration. The level of sodium ion concentration in the blood, and the ratio of sodium to potassium, is decreased after removal of the adrenal glands from mammals. This change is largely a consequence of an increase in sodium chloride excretion and a decrease in loss of potassium ion through the kidneys. These changes are reversed by aldosterone. When frogs are acclimated to salt solutions, the uptake of sodium ion across the skin and from the kidney tubules is decreased, and this change is accompanied by decreased levels of aldosterone in the blood.

Secretion of aldosterone by the adrenal cortex is in general not controlled by ACTH, which controls secretion of the other adrenal steroids. Renin is a protein secreted by specialized juxtaglomerular cells in the kidney in response to decreased blood flow to, or decreased concentration of sodium ion in the blood flowing to, the glomeruli. Renin functions as a hormone, but is also an enzyme, and its target in both functions is a protein component of blood. Renin hydrolyzes this component partially, to produce the peptide angiotensin, which then acts in two ways: At low concentration, it stimulates the adrenal cortex to secrete aldosterone. At higher concentrations, it acts on the circular muscle of the arterioles, causing them to contract and constrict these vessels, increasing the blood pressure. Aldosterone stimulates reabsorption of sodium ion from the kidney tubules thus retaining this ion and its salts in the body, and increased blood pressure increases blood flow to the glomeruli of the kidney.

Progressive Processes

The progressive processes under hormonal control are those of growth and differentiation in plants and animals. For example, the growth and development of higher plants is under complex control by at least three classes of hormones, known as auxins, kinins (kinetin), and gibberellins. Other examples are the periodic growth of arthropods (insects and crustaceans), the associated developmental processes, and the color changes of some crustaceans which are controlled by a neuroendocrine system having some interesting analogies with the vertebrate system.

Arthropods. The arthropod system is centered in the brain, with extensions into the eyestalks in some crustaceans. The thoracic gland of insects, and its homolog, the Y organ of crustaceans, is an ordinary endocrine gland (not neural in origin or connections) which secretes a steroid hormone (ecdysone in insects, and a modified version, crustecdysone, in crustaceans) that initiates and maintains the processes of molting, the shedding of the integument which is essential to growth and development in these animals. Secretion of ecdysone is initiated by a neurohumor formed in the brain. In larval insects the corpora allata, another pair of glands, secrete the juvenile hormone neotenin at each molt, and this hormone retards developmental changes leading to the adult condition until the final molt, when neotenin secretion ceases and the animal takes on adult characters.

Crustaceans have an androgenic gland associated with the testes which is responsible for the differentiation of male secondary sex characteristics, and the eyestalks form a hormone which restrains ovarian growth and delays the transformation of males into females, characteristic of the life history of certain crustaceans. In some crustaceans, also, the ovary secretes a hormone which induces development of a brood pouch and other temporary female characters. In insects the corpora allata appear to secrete one or more hormones concerned with development of female reproductive structures. In many other invertebrate animals there are correlations of neurosecretory activity with sexual maturation and the discharge of sexual products. In mussels and oysters, for example, neurosecretory cells show marked development before the maturation of the gametes and their discharge into the water; removal of the ganglia containing these cells initiates maturation and spawning. *See* ARTHROPODA; CRUSTACEA; INSECTA.

Plants. The growth and differentiation of plants is also under hormonal control. The growing tips of higher plants form auxins, which are then transported from the tips toward the base of the plant. The auxins have the effect of stimulating elongation of growing cells in proportion to auxin concentration, up to a limit; higher concentrations may inhibit growth. They also inhibit the growth of lateral buds and the abscission of fruits and leaves, and stimulate the enlargement of fruit cells and the initiation of new roots. The phenomenon of apical dominance, whereby the presence of a growing tip inhibits the growth of lateral buds in the basal regions of the stem, is attributed primarily to auxin production in the tip and polar transport toward the base. The phototropic responses of plants, whereby a growing stem tip turns toward the light and a growing root tip turns away, and the geotropic responses, in which

the stem grows upward and the root downward, are attributed to lateral polar transport of auxin in stem and root under the influence of light and gravity, respectively, and the inactivation of auxin by light.

The functions of two other plant-growth hormones remain uncertain. Kinetin is formed in many plants and has a stimulating effect on cell division. Its effects on intact plants and isolated plant parts depend partly on auxin concentration. Gibberellins, originally isolated from a fungus which attacks rice plants and other substances of similar nature, are also widely distributed. Like the auxins, the gibberellins stimulate cell elongation, and they are particularly effective in causing dwarf varieties of plants to grow to sizes characteristic of normal varieties. Unlike the auxins, gibberellins move in all directions in the plant and are not transported in polar fashion. *See* PLANT GROWTH; PLANT HORMONES; PLANT MORPHOGENESIS.

Cyclical processes. Cyclical processes of reproduction and molting have been noted above. The vertebrate female sexual cycle forms the best-known instance of hormonal control of such a cycle.

In mammals the anterior pituitary secretes a gonadotropin or follicle-stimulating hormone (FSH), which causes development of ovaries in the female and testes in the male. A second gonadotropin, the luteinizing (LH), or interstitial-cell-stimulating hormone (ICSH), stimulates growth of the interstitial cells of the testes and with FSH causes maturation of the ovarian follicle and ovulation. The interstitial cells of the testis secrete testosterone, which induces development of male sexual characteristics and male behavior. The ovarian follicle secretes estrogens, which induce development of female sexual characteristics. As the follicle grows, and estrogen secretion increases, the estrogen stimulates proliferation of the uterine lining and secretion of more LH by the pituitary. The LH, in turn, causes ripening of the follicle and ovulation. The ruptured follicle becomes a corpus luteum and begins to secret progesterone, which further stimulates growth of the uterine lining. If fertilization occurs, the fertilized egg is implanted in the uterus and the placenta secretes estrogen, progesterone, and LH; these maintain the corpus luteum and the uterine lining in active condition and inhibit the secretion of FSH by the pituitary. If fertilization does not occur, the pituitary begins to secrete FSH, the corpus luteum breaks down, and in the absence of progesterone the uterine lining breaks down, completing the cycle. *See* REPRODUCTIVE SYSTEM.

In most mammals the maximum period of estrogen secretion is during heat or estrus, the only time the female is receptive to the male. In some mammals, such as the rabbit, ovulation occurs only following copulation, as a result of neurosecretory stimulation to the anterior pituitary, leading to LH secretion. In other mammals, and many other vertebrates as well, the female cycle is initiated by an effect of changing lengths of daylight period, mediated through neurosecretory cells acting on the pituitary. In oviparous vertebrates the cycle is essentially the same up to the ovulation. Following deposition of eggs, or parturition in mammals, the anterior pituitary forms prolactin which induces maternal behavior and in mammals is essential to lactation. As the luteotropic hormone, it also maintains the ovarian corpus luteum in an active condition secreting progesterone. The flow of milk in lactating mammals is initiated by oxytocin. *See* ENDOCRINE SYSTEM (INVERTEBRATE); ENDOCRINE SYSTEM (VERTEBRATE); ESTRUS; LACTATION.

Bradley T. Scheer

Bibliography. K. L. Becker et al. (eds.), *Principles and Practice of Endocrinology and Metabolism*, 1990; I. Chester-Jones and I. W. Henderson (eds.), *Fundamentals of Comparative Vertebrate Endocrinology*, 1987; L. J. DeGroot, *Endocrinology*, 3 vols., 3d ed., 1995; J. E. Griffin and S. R. Ojeda (eds.), *Textbook of Endocrine Physiology*, 4th ed., 2000.

Endocrine system (invertebrate)

The chemical integrating system in animals that lack a vertebral (spinal) column. An endocrine system consists of those glandular cells, tissues, and organs whose products (hormones) supplement the rapid, short-term coordinating functions of the nervous system. Evidence has been presented for hormones in a wide variety of invertebrates. However, the large majority of the published reports pertains to the more highly evolved groups that will be discussed below, the annelids, mollusks, and particularly two classes of arthropods, the insects and crustaceans. Many of the hormones in invertebrates are neurohormones; that is, they are produced by nerve cells. Just as in the vertebrates, a wide variety of functions are regulated by hormones in the invertebrates. *See* NEUROSECRETION.

Insects. The endocrine system of insects has been studied more intensively than that of any other invertebrate group. Increase in linear dimensions of an insect, as in all arthropods, can occur only at periodic intervals when the restricting exoskeleton is shed during a process known as molting. The orderly sequence of postembryonic molts that leads from the immature insect, which has hatched from the egg, to the adult is controlled by three major classes of hormones. The brain is the source of prothoracicotropic hormone, which is the starting component of the endocrine sequence that leads to a molt. This hormone activates a pair of glands in the prothorax, the prothoracic glands, causing them to secrete ecdysone (**Fig. 1***a*). The ecydsone is then converted to 20-hydroxyecdysone in the Malpighian tubules and in the fat body, and released into the blood. 20-Hydroxyecdysone is the actual ecdysteroid that initiates the molt. Ecdysteroids are in fact the molt-inducing factors of all arthropods. The third class of hormones is the juvenile hormones (Fig. 1*b*), produced by the corpora allata, a pair of glands near the brain. Juvenile hormone functions during the juvenile molts to suppress the differentiation of adult characteristics, permitting growth but

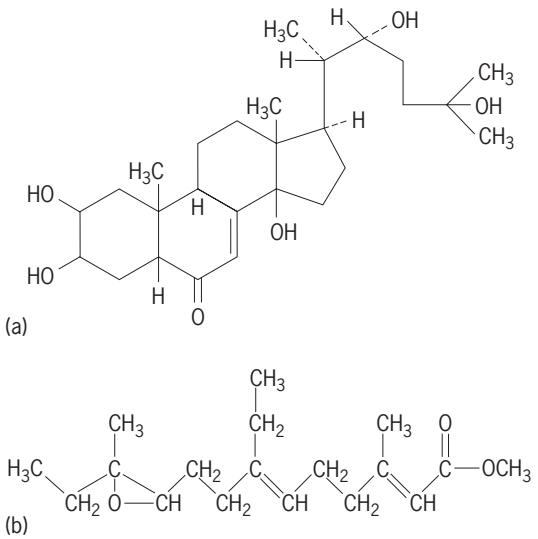

Fig. 1. Chemical structures of two insect hormones. (*a*) Ecdysone. (*b*) Juvenile hormone I.

preventing maturation. Surgical removal of the corpora allata early in development leads to the precocious development of an adult, which, however, is smaller than normal. Thus the character of the molt, whether juvenile characteristics are expressed or there is metamorphosis to the adult, is determined by the presence or absence of juvenile hormone in the blood. *See* MOLTING (ARTHROPODA).

Prothoracicotropic hormone appears to be a polypeptide or small protein. Ecdysone and 20-hydroxyecdysone are steroids. Juvenile hormone (JH I) was the first such isolated; at least three additional forms of juvenile hormone (JH 0, JH II, and JH III) that are structurally very closely related to JH I were isolated later. Prior to the isolation of JH I, a substance from the corpora allata was found to be necessary for the synthesis of yolk and its deposition in the developing eggs of an adult female. Now that pure juvenile hormone is available, it is clear that juvenile hormone, not another product of the corpora allata, has this role after the adult stage of development has been attained. *See* ECDYSONE; INSECT PHYSIOLOGY; INSECTA.

Water regulation. Two hormones with antagonistic actions are involved in regulating the water content of insects. One, the diuretic hormone, promotes water loss by increasing the volume of fluid secreted into the Malpighian tubules, the excretory structures, which are hollow threadlike structures that develop as outpocketings from the portion of the digestive tract anterior to the rectum, at the junction of the midgut and the hindgut. The second, the antidiuretic hormone, acts to conserve water by stimulating the wall of the rectum to increase the volume of water resorbed from its lumen.

Other hormones. Bursicon, a protein neurohormone, is responsible for the tanning and stiffening of a newly formed cuticle. The adipokinetic hormone is a peptide from the corpora cardiaca, which lie in the head near the corpora allata. This hormone produces an increase in the blood lipid concentration; the lipid is released from the fat body. This increase occurs normally during flight, the lipid presumably serving as a source of energy for the flight muscles. A few species of insects, notably the stick insect *Carausius morosus*, have the ability to change color. This insect has a circadian rhythm of color change, becoming darker at night and lighter by day as a result of the rearrangement of pigment granules within its epidermal cells. Darkening is due to a neurohormone. *See* PROTECTIVE COLORATION.

Crustaceans. In crustaceans, as in insects, hormones have vital roles in a wide variety of physiological processes. Higher crustaceans (**Fig. 2**) have a structure, the sinus gland, that in most stalk-eyed species lies in the eyestalk and is the storage and release site for several hormones, including a molt-inhibiting hormone. The Y-organs, a pair of glands found in the anterior portion of the body near the excretory organs, are the source of ecdysone. The ecdysone is then converted in other tissues such as the ovaries and testes to 20-hydroxyecdysone, which as in insects is the actual molting hormone. The sinus glands are neuroendocrine structures, but the Y-organs are nonneural. Molt-inhibiting hormone controls the time of onset of molting activity by preventing the output of ecdysone by the Y-organs. When molt-inhibiting hormone release is suppressed, the Y-organs become active. *See* CRUSTACEA.

Color changes. Crustaceans exhibit some of the most spectacular color changes seen in the animal kingdom. The pigment cells (chromatophores) are under hormonal control, as in insects. The chromatophorotropic hormones are released from the postcommissural organs, which lie near the esophagus, and from the sinus glands. Neuropeptides have been found to cause dispersion of the pigment within the chromatophores, as well as substances with the opposite action (those which cause the pigment to become punctate). A pigment-concentrating hormone of 8 amino acids and pigment-dispersing hormones of 18 amino acids have been purified, and their chemical structures have been determined. *See* CHROMATOPHORE.

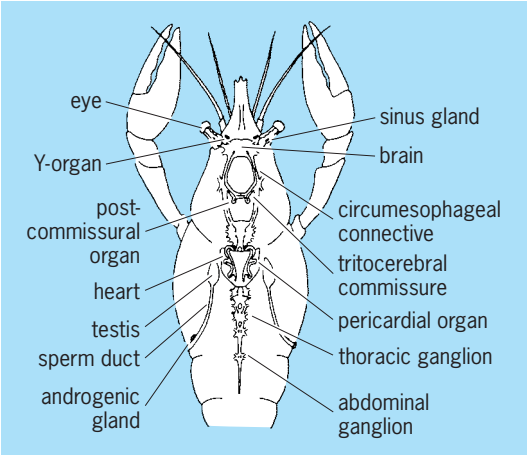

Fig. 2. Endocrine system of a male crustacean. (*After A. Gorbman and H. A. Bern, A Textbook of Comparative Endocrinology, Wiley, 1962*)

Reproduction. Although the sex of a crustacean is genetically determined, the structural and functional expression of the sex (male or female) is related to the presence or absence of functioning androgenic glands. These glands, which develop to a functional state only in males, are in all but a few species attached to the sperm ducts, and are required for the differentiation of a normal male. Not only are the androgenic glands essential for the normal development and functioning of the testes, but the androgenic glands (not the testes) are also responsible for the normal differentiation of a male's secondary sex characters, such as the copulatory appendages. In genetic females, however, in the absence of androgenic gland hormone, the gonadal rudiments autodifferentiate into ovaries; and the ovaries then produce a hormone that promotes the normal differentiation of the female secondary sex characters, such as those that are used to brood the developing young. A substance in the sinus gland inhibits the functioning of the gonads (the gonad-inhibiting hormone), whereas the brain and thoracic ganglia contain a gonad-stimulating hormone. The gonad inhibitor and stimulator act directly on the ovaries but only indirectly on the testes by acting directly on the androgenic glands. Through the actions of gonad-inhibiting and gonad-stimulating hormones, the reproductive cycles are properly timed.

Retinal pigments. The compound eye of crustaceans has three retinal pigments whose movements in accordance with changes in the illumination level control the amount of light impinging on the photosensitive cells. Of these pigments, the distal or outermost one has been studied most intensively. In the sinus gland is a neurohormone (the light-adapting hormone) that causes the migration of the distal pigment toward its light-adapted position, shielding the photosensitive cells, and another neurohormone (the dark-adapting hormone) that causes migration of this pigment toward its dark-adapted position, thereby uncovering the photosensitive cells. *See* EYE (INVERTEBRATE).

Other endocrine mechanisms. The pericardial organs, which are found near the heart, are neuroendocrine organs. The pericardial organ hormone produces an increase in the amplitude of the heartbeat. The sinus gland contains the hyperglycemic hormone, which causes a rise in the blood glucose concentration. Studies on the endocrine control of salt and water regulation revealed the existence of a neurohormone that prevents dilution of the blood of crustaceans placed into diluted seawater.

Annelids. Strong evidence exists for hormones in annelids, particularly from studies of nereids, which are polychaetes, and earthworms, which are oligochaetes. In nereids a hormone from the brain, that is, cerebral ganglia, is required for growth, which occurs normally by the formation of body segments posteriorly. This hormone is called the brain hormone. A substance from the nereid brain also inhibits precocious sexual maturation. These studies have been done mostly with nereid females, but the few studies done with nereid males suggest the control of sexual maturation is the same in both sexes. Removal of the brain from an immature worm stops growth but leads to precocious sexual maturation. It is possible that the brain produces two hormones, one that stimulates growth and another that inhibits reproduction, but the consensus is that the evidence favors the interpretation that both actions are due to a single substance. For example, there is a parallel decline in the rates of secretion by the brain of both the growth-stimulating hormone and the maturation-inhibiting hormone. The brain hormone in its combined role would stimulate growth in immature animals while exerting its inhibitory action upon sexual maturation processes. Later, secretion of this hormone would slow and eventually stop, and this leads to sexual maturation.

In contrast to the nereids, in earthworms a stimulatory hormone from the brain is required for sexual maturation. Removal of the brain from a sexually mature earthworm stops gametogenesis and leads to the regression of the secondary sexual characteristics, such as the clitellum that secretes the cocoon. The earthworm brain produces hormones that regulate other functions also. One such helps to control salt and water balance, possibly by lowering the permeability of the body wall to water influx or by acting as a diuretic hormone to stimulate excretion of excess water; still another one is a hyperglycemic hormone that, as in crustaceans, elevates the blood glucose concentration. However, the possible hormonal control of growth in earthworms is a subject that needs further study. There are apparently conflicting reports that normal growth is dependent upon the brain; brain removal does not prevent the formation of new segments; and brain implants inhibit growth. *See* ANNELIDA.

Mollusks. As in the annelids and arthropods, in mollusks hormones are known to control a wide variety of functions. In cephalopods, such as the octopus and squid, there is a pair of optic glands which lie on either side of the central part of the brain, on the stalks of the optic lobes. These glands secrete a gonadotropic hormone that stimulates maturation of the ovary and testis and also stimulates development of the male and female accessory sex organs. The optic glands are under nervous inhibitory control. Severing the nerve supply to these glands results in precocious sexual maturation. Similarly, pulmonate snails have paired hormone-producing organs, the dorsal bodies, attached to the dorsal surface of their cerebral ganglia. Pulmonate snails are hermaphroditic. The dorsal bodies secrete a female gonadotropic hormone that appears to be proteinaceous. Removal of the dorsal bodies inhibits maturation of egg cells, whereas reimplantation of the dorsal bodies restores the normal condition. Spermatogenesis in at least some pulmonates, such as the land slug *Limax*, is stimulated by a substance called the maturation hormone that appears to be released from two sites: the cerebral ganglia and a group of cells (the collar cells) in the optic tentacles.

The hormonal control of egg-laying behavior in snails has been studied extensively. The caudo-dorsal cells in the cerebral ganglia of the fresh-water snail *Lymnaea* and the bag cells which lie at the rostral margin of the abdominal ganglion in the marine snail *Aplysia* produce neuropeptides that evoke egg-laying behavior. This egg-laying behavior is a complex series of events, including ovulation of the oocytes, extrusion of the egg string, and its fixation to the substrate. The egg-laying hormones from the caudo-dorsal cells and bag cells consist of a chain of 36 amino acids, but are not identical. In fact, *A. californica* and *A. parvula* even show amino acid sequence differences in their bag-cell egg-laying hormone.

A growth hormone, which appears to be a peptide, has been found in the cerebral ganglia of several snails. It stimulates growth of both the shell and the soft body parts. A hyperglycemic factor has been found in the nervous systems of the mussel *Mytilus* and of *Lymnaea*. Also, hormones controlling salt and water balance have been found. For example, the pleural ganglia of *Lymnaea* produce a diuretic hormone. Cauterization of these ganglia reduces the rate of urine formation, resulting in increased body weight due to water retention. The visceral ganglia of the oyster *Crassostrea* appear to produce such a diuretic hormone. Also, the cerebral ganglia of *Lymnaea* produce a substance, which seems to be a peptide, that stimulates influx of sodium ions. Removal of the cerebral ganglia results in a decrease in the blood sodium concentration. *See* MOLLUSCA.

Milton Fingerman

Bibliography. D. E. Bliss and L. H. Mantel, *The Biology of Crustacea*, vol. 9, 1985; M. Fingerman, The endocrine mechanisms of crustaceans, *J. Crust. Biol.*, 7:1–24, 1987; P. M. Hopkins, Ecdysteroid titers and Y-organ activity during late anecdysis in the fiddler crab, *Uca pugilator, Gen. Comp. Endocrinol.*, 63:362–373, 1986; H. Laufer and R. G. H. Downer, *Endocrinology of Selected Invertebrate Types*, 1988; G. G. Lunt and R. W. Olsen, *Comparative Invertebrate Neurochemistry*, 1988; M. H. Schaffer, Functional and evolutionary relationships among the RPCH-AKH family of peptides, *Amer. Zool.*, 26: 997–1005, 1986.

Endocrine system (vertebrate)

A system of chemical communication among cells. The classical vertebrate endocrine system consists of a group of discrete glands that secrete unique products (hormones) into the bloodstream. These products travel in the blood to distant sites or targets where they cause specific physiological responses. Thus endocrine glands differ from exocrine glands, in that they lack ducts and deliver their secretions in the bloodstream. The classical definition of an endocrine system is harder to apply nowadays with the discovery of scattered cells rather than discrete glands that act as endocrine organs, of endocrine cells that affect themselves (autocrine effect) or nearby targets (paracrine effect) by diffusion through extracellular fluids rather than the bloodstream, and of neurons that secrete hormones (neurosecretion). All of these mechanisms, however, allow for chemical intercellular communication and can be considered part of the endocrine system. *See* NEUROSECRETION.

Homeostasis. One important function of the endocrine system, along with the nervous system, is to maintain homeostasis, that is, a constancy of the internal environment of an organism. Thus an organism reacts and adjusts physiologically to changes in its external environment. For example, certain species of fish migrate from fresh water to seawater, or vice versa, and adjust to changes in salt concentrations by endocrine and neuronal responses. The roles of the endocrine and nervous systems in maintaining homeostasis are many, complementary, and overlapping. *See* HOMEOSTASIS; NERVOUS SYSTEM (VERTEBRATE).

Nature of hormones. Hormones are the products of endocrine cells. They are either proteinlike (amino acids, peptides, proteins, glycoproteins) or steroidal. Peptide hormones are produced by protein synthetic mechanisms directed by the genes of the endocrine cells. Peptide hormones are usually produced as precursors (prohormones). In some cases, two or more different hormones are enclosed within the larger prohormone. These prohormones are cleaved to the smaller peptide hormone or hormones just before leaving the endocrine cell or when they reach their target. One reason that prohormones are produced may be that they are biologically inactive and thus do not cause physiological changes when stored. Protein hormones are stored in endocrine cells in secretory granules that bud off the endoplasmic reticulum and Golgi membranes, where protein synthesis occurs. The granules leave the cell by endocytosis and enter the bloodstream. *See* PROTEIN.

Steroid hormones, on the other hand, are produced from cholesterol by a number of well-characterized, enzyme-catalyzed steps. Cholesterol is thus converted stepwise to various steroid families: the hormones of the adrenal cortex (cortisol and aldosterone), the estrogens (estradiol) from the ovary, and the androgens (testosterone) from the testis. Steroid hormones diffuse across the endocrine cell plasma membrane to enter the circulation. *See* HORMONE; STEROID.

Receptors. Hormones cause physiological responses in their targets. Since in most cases the blood carries hormones throughout the body, there must be a system by which only certain tissues respond to each hormone. This is accomplished by receptors, which are binding sites either on the surface of the target cell or within its nucleus. Receptors are high-molecular-weight proteins; the structure of some, such as the insulin receptor, is known. In general, peptide hormones cannot cross the plasma membrane, so their receptors are located there, whereas steroid hormones do pass through the plasma membrane of their targets and

bind to nuclear receptors, probably located in the deoxyribonucleic acid (DNA).

Second messengers. In order for peptide hormones to stimulate physiological changes within the target cell, the "message" must be passed from the hormone-receptor complex of the plasma membrane to the interior of the cell. This process of signaling across the plasma membrane is accomplished by so-called second messengers of which the best known is adenosine 3′,5′-cyclic monophosphate (cyclic AMP). The breakdown of cell-membrane phospholipids in response to hormone–receptor interaction is another widespread mechanism of transmembrane signaling which mobilizes calcium ion within the target cells. Calcium ion, in turn, is an important second messenger. Steroid hormones, of course, have no need for second messengers since they are lipid-soluble and pass readily through the plasma membrane and into the cell. *See* CELL MEMBRANES.

Physiological responses. Once hormones are bound with their receptors and have stimulated their target cells, physiological responses occur. This may involve such biochemical processes as conversion of an inactive form of an enzyme into an active one, stimulation of critical enzymatic pathways, increase in transport of glucose or amino acids into cells, or synthesis of new proteins. These events may result in overall changes in cell or organ function, metabolism, growth, or even the behavior of the organism.

Feedback systems. The endocrine system is regulated by control mechanisms, the means by which homeostasis is achieved. The most common relationship between the hormone and its target is one of negative feedback, whereby the response to the hormonal stimulus turns off the original stimulus. For example, the endocrine pancreatic beta cells produce insulin in response to high blood sugar levels. Insulin is released into the blood, where it causes its target cells to take up glucose, thus reducing blood sugar. When blood glucose concentration falls, the secretion of insulin is turned off. The system is turned back on by the following process: when less insulin is secreted, the target cells take up less glucose, and thus the blood glucose content gradually increases again until insulin secretion is stimulated once more. Thus the blood level of hormone (insulin) and the substance responding to the hormone (glucose) are kept within normal ranges. *See* CARBOHYDRATE METABOLISM; INSULIN.

Most relationships between hormones and their targets involve negative feedback. However, positive feedback also occurs here and there, whereby the production of the hormone is stimulated rather than turned off by the product of the hormone–target interaction. *See* ENDOCRINE MECHANISMS.

Pituitary gland and hypothalamus. The pituitary gland, or hypophysis, is located near the base of the brain. It secretes many hormones and controls the function of other endocrine glands. The production and release of the various pituitary hormones are regulated in turn by small peptide-releasing hormones from the hypothalamus of the brain. These factors are produced by neurosecretory neurons and travel to the adenohypophysis (anterior lobe of the pituitary) by way of a portal blood system. The releasing hormones stimulate specific cells of the adenohypophysis to produce and release their hormones. Generally speaking, each of the adenohypophysial hormones is affected by a separate releasing hormone. Thus, the hypothalamic thyrotropin-releasing hormone stimulates the synthesis and release of thyroid-stimulating hormone (thyrotropin) by the adenohypophysis.

Other adenohypophysial hormones include adrenocorticotropic hormone, which stimulates the production of steroid hormones by the adrenal cortex; growth hormone, which stimulates protein synthesis and growth in many cells; prolactin, which stimulates the production of milk by the mammary glands and is important in salt and water balance and many other functions; follicle-stimulating hormone, which induces growth of the follicles of the ovary prior to ovulation; and luteinizing hormone, which induces ovulation in the ovary. The release of both follicle-stimulating and luteinizing hormones is governed by gonadotropin-releasing hormone.

Other hypothalamic hormones do not reach the pituitary by way of the bloodstream; instead they travel down the long axons of the neurosecretory cells into the neurohypophysis (posterior lobe of the pituitary). These hormones, oxytocin and vasopressin (or antidiuretic hormone), are released directly from the axon end bulbs into the blood. Oxytocin acts upon the mammary glands of female mammals to cause milk release in response to suckling by the young, and stimulates the uterus to contract at the end of pregnancy to aid in expulsion of the offspring. Vasopressin is important in water conservation by the kidney tubules (less urine excretion) and also produces an increase in blood pressure. Oxytocin and vasopressin are, therefore, hypothalamic hormones that are stored in the pituitary. *See* PITUITARY GLAND.

Thyroid gland. The thyroid gland lies in the neck region of mammals. It produces two closely related hormones, triiodothyronine and thyroxine. These both increase the metabolic rate of an organism, and increase enzyme activity and protein synthesis. The thyroid hormones act along with growth hormone to promote cell growth and development. These functions are well illustrated in cases of hypothyroid human infants, who are retarded both mentally and physically. Thyroid hormones are peptides but their three-dimensional structures may be similar to those of steroid hormones. Thus they are unusual in their ability to pass through the plasma membrane of their target cells and bind to nuclear receptors, directly affecting genes that control protein synthesis.

The control of hormone secretion by the thyroid (as well as by the adrenal cortex and gonads) involves more complex feedback relationships. These endocrine glands are affected by the levels of hormones from the adenohypophysis (which were affected by

releasing hormones from the hypothalamus). In the case of the thyroid gland, thyrotropin-releasing factor from the hypothalamus stimulates the release of thyroid-stimulating hormone by the adenohypophysis. In response, the thyroid secretes thyroxine and triiodothyronine. High blood levels of the thyroid hormones inhibit the secretion of both thyrotropin-releasing factor (long-loop feedback) and thyroid-stimulating hormone (short-loop feedback). *See* THYROID GLAND.

Calcium regulation. The parathyroid glands derive their name from the fact that in mammals they are embedded within the thyroids. These small glands are essential for life, as they regulate the concentration of calcium ion in blood and other extracellular fluids. If calcium is too low, the animal goes into tetanic convulsions and dies, whereas if calcium is too high, abnormal calcification and stone formation can occur. Thus the regulation of calcium ions is of utmost importance. Parathyroid hormone is a protein hormone that raises the blood calcium levels (hypercalcemia). The hormone acts upon bone to cause the release of calcium and phosphate, and upon the kidney to increase the reabsorption (conservation) of calcium and excretion of phosphate.

Vitamin D is now recognized as a steroidlike hormone, although it does not originate from an endocrine gland. It is synthesized from precursors present in the diet or produced after exposure of skin lipids to ultraviolet light. The two final steps in vitamin D biosynthesis involve the addition of hydroxyl groups to the molecule. The first hydroxylation occurs in the liver, the second in the kidney, thus forming the most active metabolite, 1,25-dihydroxyvitamin D_3. Vitamin D plays roles in calcium conservation by the kidney and in bone mineralization, but its most important function is to enhance calcium transport across intestinal cells and thus conserve dietary calcium. *See* VITAMIN D.

Calcitonin is a newly recognized peptide hormone produced by thyroid cells in mammals (different cells from those that produce thyroid hormones) and from the ultimobranchial glands of nonmammalian vertebrates. Calcitonin is hypocalcemic and acts by inhibiting calcium loss from bone. Of the three calcium-regulating hormones, it appears to be the least important. *See* PARATHYROID GLAND.

Carbohydrate regulation. Insulin and glucagon are peptide hormones produced by endocrine cells of the pancreatic islets. Insulin is a protein hormone produced by the pancreatic beta cells and is the only hormone that decreases blood sugar (glucose) levels. It acts on its target cells (skeletal muscle, fat cells) to increase the uptake of glucose, amino acids, and fatty acids. Once taken into cells, glucose is used in metabolic reactions or stored as glycogen, a large carbohydrate. Insulin also causes the conversion of amino acids to proteins and fatty acids to fats in the target cells. In the absence of insulin, as in diabetes mellitus, the target cells cannot take up glucose, and thus the body must utilize amino acids and fats as energy sources. These processes result in the accumulation of toxic metabolic products which eventually disrupt the acid–base balance, leading to coma and death. *See* DIABETES.

Glucagon, in contrast, is a hyperglycemic hormone. It is a small peptide from the pancreatic islet alpha cells that acts upon liver cells to cause the conversion of glycogen to glucose by activation of key enzymes in a complex metabolic pathway.

Many other hormones elevate blood sugar levels. For example, epinephrine (adrenalin), an amino acid derivative from the adrenal medulla, acts by the same pathway as glucagon to convert glycogen to glucose, except that the targets of epinephrine are skeletal and heart muscle. Epinephrine is secreted in times of stress and serves to prepare the body for an emergency by increasing the availability of energy in the form of glucose and by increasing the heart rate and blood pressure.

Growth hormone, a large protein hormone from the adenohypophysis, is secreted in response to low blood sugar levels. This hormone elevates blood sugar by blocking the uptake of glucose by cells and by favoring the utilization of fats rather than glucose as an energy source.

Many of the adrenal cortical hormones, such as cortisol, are known collectively as glucocorticoids, because they also elevate blood glucose levels. These steroid hormones favor the production of glucose from proteins and fats rather than glycogen (gluconeogenesis). Glucocorticoids also exert an anti-inflammatory action, which makes them useful for treatment of arthritis and other diseases. *See* ADRENAL GLAND.

Salt and water regulation. Several hormones affect the ability of the kidney to conserve or excrete salts and water. The antidiuretic hormone (vasopressin) promotes water reabsorption by the kidney tubules, so that the organism excretes less water. The secretion of vasopressin is regulated by hypothalamic neurosecretory neurons that are sensitive to the concentration of salts in the extracellular fluids. In the absence of vasopressin, an individual excretes great volumes of dilute urine, leading to severe dehydration (diabetes insipidus).

Salt excretion is regulated mainly by two hormones that act in opposition. Aldosterone is an adrenal cortical steroid that promotes the reabsorption of sodium by the kidney tubules and thus decreases its excretion in the urine. In contrast, atriopeptin (atrial natriuretic factor), a peptide that originates in heart muscle, acts upon the kidney to increase the excretion of sodium in the urine. *See* OSMOREGULATORY MECHANISMS.

Reproductive hormones. Probably the best-studied endocrine glands are the gonads, the testes of the male and the ovaries of the female. The gonads are regulated by the follicle-stimulating hormone and luteinizing hormone from the adenohypophysis. In the male, follicle-stimulating hormone stimulates the initiation of sperm formation by the testis tubules, and luteinizing hormone acts on the nearby Leydig cells of the testis to produce testosterone, the principal male sex hormone. Testosterone acts by a paracrine mechanism to cause the final maturation

of sperm, and by way of the blood to stimulate development of the male reproductive system and secondary sex characteristics (body shape, beard, muscle growth). *See* TESTIS.

In the female, follicle-stimulating hormone stimulates the growth of the ovarian follicles at the beginning of each reproductive cycle. As the follicles grow, they produce estradiol, an important female sex hormone. Increasing levels of estradiol cause feedback inhibition of the released gonadotropin-releasing hormone. High levels of estradiol also have an unusual positive feedback effect upon the hypothalamus and adenohypophysis to cause a surge in the secretion of luteinizing hormone, which results in ovulation (the release of the ovum). The corpus luteum, a remnant of the ovulated follicle, produces both estradiol and the second major female sex hormone, progesterone. Progesterone is necessary for the maintenance of a quiescent uterus during pregnancy, and both estrogen and progesterone are important in the regulation of the female reproductive cycle (the rhythmic timing of the menstrual cycle in primates, for example). Estradiol is also essential for the growth and maturation of the female reproductive system and secondary sex characteristics (breast growth, body shape, patterns of fat deposition). In both males and females, the sex hormones affect reproductive behavior. *See* OVARY; REPRODUCTIVE BEHAVIOR.

Other hormones. There are many other factors that act in various ways to achieve homeostasis or intercellular communication. Of these, the existence of a large number of peptides found in both gastrointestinal cells and the brain is of interest. These were recognized for many years as gastrointestinal hormones which aid in secretion of digestive juices and motility of the gastrointestinal tract. Their function in the brain appears to be different, and there is evidence that they act in pain reception or analgesia, as factors that stimulate or curb appetite, or in memory or other functions. This field of neuropeptide hormones is in its infancy and serves to emphasize the close relationship between the endocrine and nervous systems in intercellular communication.

Nancy B. Clark

Bibliography. A. Gorbman et al., *Comparative Endocrinology*, 1983; G. Litwak and A. W. Norman, *Hormones*, 2d ed., 1997; D. Randall et al., *Eckert's Animal Physiology*: *Mechanisms and Adaptations*, 4th ed., 1997.

Endocrinology

The study of the glands of internal secretion, the endocrine glands, and the hormones which they synthesize and secrete. These glands are ductless; the hormones are secreted directly into the blood to be carried to the target tissue or organ. The hormones, or chemical messengers, are highly specific and their action may be selective or generalized. *See* ENDOCRINE SYSTEM (VERTEBRATE).

Most of the physiological processes of the body are under hormonal and nervous system regulation. The quantities of circulating hormones are maintained in proper balance by the delicate feedback mechanisms involving endocrine gland interactions, blood levels of hormones, and physiological activities of the target organ. The pituitary is considered the master gland because so many of its hormones are trophic in nature; that is, they regulate the activities of other endocrines. *See* ENDOCRINE MECHANISMS; PITUITARY GLAND.

The hormones carry out their biological activities by regulating cellular metabolism; they act as inhibitors or accelerators of intracellular biochemical reactions to serve three general functions. These are control of growth and development, determination of secondary sex characteristics, and maintenance of homeostasis. *See* HOMEOSTASIS.

Chemically, there are three kinds of hormones: steroids, amines, and peptides or proteins. However, many subtypes and variations may appear in an individual. *See* HORMONE; STEROID.

Most hormones have been purified, analyzed, and standardized so that a whole gamut of synthetic analogs are available for the highly selective actions desirable in therapy. *See* ENDOCRINE SYSTEM (INVERTEBRATE).

Sybil P. Parker

Endocytosis

The process by which animal cells internalize particulate material (such as cellular debris and microorganisms), macromolecules (such as proteins and complex sugars), and low-molecular-weight molecules (such as vitamins and simple sugars). The size difference among these different nutrient sources requires cells to develop a hierarchical plan for endocytosis. Thus, cells engage in at least three different types of endocytosis: phagocytosis where cells engulf particulate material, receptor-mediated-endocytosis of macromolecules, and potocytosis of small molecules.

Some of the essential nutrients that a cell needs are scarce in the environment. The cells overcome this problem by expressing high-affinity receptors, or binding sites, on the membrane surface. Each type of receptor is specific for either macromolecules or molecules. These endocytic receptors are capable of concentrating their ligand at the cell surface before carrying it into the cell, thus increasing the efficiency of uptake so effectively that a malfunctioning endocytic receptor can cause a disease. For example, individuals who have a defective endocytic receptor for low-density lipoprotein (LDL), the major cholesterol-carrying molecule in the blood, develop a disease called familial hypercholesterolemia that causes premature heart attacks.

In all three endocytic pathways the internalization step begins with the invagination of plasma membrane and the conversion of this membrane into a closed vesicle called an endosome. Each of the pathways has its own set of molecules that control

internalization. These molecules assemble at the cell surface and physically deform the membrane into the shape of a vesicle. The vesicle, the endosome, then detaches and migrates to other locations within the cell. The same cell-surface assemblage of molecules also attracts endocytic receptors that are moving around on the cell surface, causing them to cluster over the site of internalization. Receptor clustering, which is essential for efficient uptake, is sometimes stimulated by ligand binding.

In general, molecules must be smaller than 1000 daltons to pass through a membrane barrier, because this size can easily fit within water-filled channels in the membrane. Therefore, particles and macromolecules can benefit the cell only after they are broken down into sugars, amino acids, and other molecules that can readily cross membranes. The breakdown of macromolecules is the major function of a membrane-bound digestive organelle called the lysosome. Therefore, endosomes that are generated by the phagocytic and receptor-mediated endocytic pathways often fuse with lysosomes that contain many different hydrolytic enzymes. Small molecules, by contrast, do not need further processing, so during potocytosis they are delivered directly to the cytoplasm. *See* CELL MEMBRANES; LYSOSOME.

Phagocytosis. Phagocytosis is a receptor-mediated process where the receptors function as adhesive elements that bond the plasma membrane to the particle. Inactivation of either the ligand on the particle or the receptor on the membrane inhibits phagocytosis.

The adhesive interaction of the phagocytic receptors with the membrane stimulates invagination. A critical molecule in this activity is actin, the same protein that provides power for muscle contraction. Surface membranes contain actin-binding proteins that link the phagocytic receptor to the actin cytoskeleton of the cell. Thus, when a particle binds to its endocytic receptor, a signal cascade is initiated that stimulates the recruitment of actin filaments to the site of phagocytosis. *See* CYTOSKELETON; MUSCLE PROTEINS.

Sometimes a cell initiates phagocytosis by extending lamellipodia that engulf the particle in a phagocytic cup. Lamellipodia are actin-rich ruffles of the plasma membrane that can assemble without a particle first binding to the membrane. They are particularly abundant in migrating cells, but in certain sessile cells, such as epithelial cells, they assemble in response to hormonal stimuli. The fusion of two lamellipodia that are next to each other can trap small amounts of extracellular fluid in an endocytic vesicle. This process is called pinocytosis, or cell drinking. All three endocytic pathways contribute to pinocytosis by nonspecifically trapping molecules and macromolecules present in the cell environment each time they generate an endocytic vesicle. *See* PHAGOCYTOSIS.

Receptor-mediated endocytosis. The clathrin-coated pit is the segment of cell membrane that is specialized for receptor-mediated endocytosis. Each pit can be recognized by the presence of a polygonal lattice on the cytoplasmic surface of the membrane. This lattice shapes the plasma membrane into a coated vesicle that immediately uncoats and fuses with endosomes. The endosome functions as a switching area that directs membrane and content molecules to specific locations within the cell. Each of the structural units of the coated pit, as well as any associated endocytic receptors, recycles many times during its life.

Endocytosis by coated pits takes place in several steps. There are four major groups of molecules to consider: receptors that use coated pits to internalize molecules; clathrin triskelions that form the distinctive polygonal lattice of the coat; assembly protein complexes (adaptors) that link the triskelion lattice to the membrane; and a membrane receptor for the adaptors that controls coated-pit assembly. Coated-pit formation starts when an adaptor binds to its receptor. The activated adaptors then aggregate and bind to triskelions that are free in the cytoplasm. The triskelions polymerize into a polygonal lattice that contains both hexagons and pentagons. Receptors migrate into fully formed coated pits and are retained there as the lattice transforms into a deeply invaginated pit. A molecular marker on the cytoplasmic tail of each receptor is the signal for retention. The coated pits then detach from the membrane.

Shortly after budding, the interior of the new vesicle becomes very acidic. A low pH is essential for the proper separation of a receptor from its ligand. Once the ligand is free in the lumen, the receptor is shunted into a vesicle that buds from one end of the endosome. These recycling vesicles return to the cell surface, while the ligand contained in the endosome proceeds to its next cellular destination. *See* CYTOPLASM.

Potocytosis. Potocytosis uses membrane proteins that are anchored by lipid rather than protein as endocytic receptors. The lipid anchor causes the attached proteins to migrate in the plane of the membrane and cluster in a membrane specialization called a caveola. Clustering ensures that any ligand bound to these receptors will be concentrated in this location. When caveolae close, they create a tiny compartment of uniform size that is sealed off from the extracellular space. When the ligand dissociates from its receptor, it reaches such a high concentration that it naturally flows through water-filled membrane channels into the cell.

Like coated pits, caveolae also have a distinctive membrane coat on their cytoplasmic surface. Presumably this coat plays an essential role in controlling the curvature of the membrane. A protein component of the coat, called caveolin, appears to be a regulatory molecule that helps to maintain a high concentration of cholesterol in the caveolar membrane. Unlike other endocytic organelles, the function of caveolae is strictly dependent on cholesterol. For example, both the shape of the caveolar membrane and the clustering of lipid-anchored endocytic receptors are abnormal in cells that have been treated to lower cholesterol levels.

Although caveolae are found on the surface of most cells, they are particularly abundant in endothelial cells. Caveolae in endothelial cells appear to deliver molecules present in the blood directly to tissue cells by a process called transcytosis. This involves the formation of a plasmalemmal vesicle at the apical surface of a special endosome, and this vesicle moves across the cell and fuses with the basal surface.

Caveolae work somewhat differently in nonendothelial cells. In epithelial cells, they appear to create a transient compartment that remains associated with the cell surface. It is not clear whether this involves the actual formation of a plasmalemmal vesicle or a simple opening and closing cycle without actual detachment from the plasma membrane. Nevertheless, there is no evidence that vesicles derived from caveolae ever merge with vesicles created by other endocytic pathways.

The closed caveolar compartment appears to be a unique space for the cell. It is transient, does not merge with other organelles, and can selectively concentrate extracellular molecules or ions and deliver them to the cytoplasm. In addition to importing molecules, cells can also use this space to store and process incoming or outgoing messengers that affect cell behavior. It has been determined that caveolae contain the molecular equipment for sending signals. For example, isolated caveolae are enriched in nonreceptor tyrosine kinases and heterotrimeric guanosine triphosphate (GTP) binding proteins, two sets of proteins that are known to be intermediates in a variety of signaling cascades. Furthermore, some of the lipid-anchored endocytic receptors are binding sites for signaling molecules like cyclic adenosine monophosphate (cAMP), while others are enzymes that can act on nearby substrates. Since the lipid anchor directs all of these molecules to the same caveola, a complex array of processing reactions can occur when it closes. *See* CELL (BIOLOGY).

Richard G. W. Anderson

Bibliography. R. G. W. Anderson, Dissecting clathrin-coated pits, *TIC,* 3:177–179, 1993; R. G. W. Anderson et al., Potocytosis: Sequestration and transport of small molecules by caveolae, *Science,* 255:410–411, 1992; J. L. Goldstein et al., Receptor-mediated endocytosis: Concepts emerging from the LDL receptor system, *Annu. Rev. Cell Biol.,* 1:1–39, 1985; J. H. Keen, Clathrin and associated assembly and disassembly proteins, *Annu. Rev. Biochem.,* 59:415–438, 1990; C. J. Steer and J. A. Hanover (eds.), *Intracellular Trafficking of Proteins,* 1991.

Endodermis

The single layer of plant cells that is located between the cortex and the vascular (xylem and phloem) tissues (**Fig. 1**). It has its most obvious development in roots and subaerial stems. In the aerial stem and in leaves it is not always detectable except by histochemical tests. However, aerial parts of many plants may develop a characteristic endodermis when subjected to belowground growing conditions.

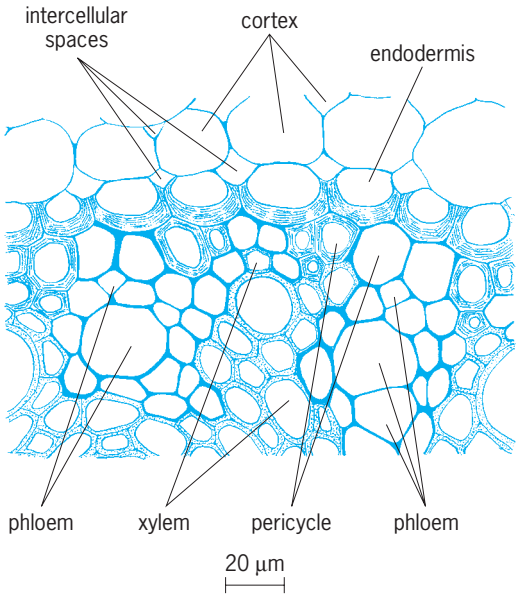

Fig. 1. Transection from part of *Zea* root illustrating endodermis in relation to contiguous tissues. (*After K. Esau, Plant Anatomy, 2d ed., John Wiley and Sons, 1967*)

Initial development. The primary phase in the development of the endodermis is identified by a thin band (Casparian strip) of suberin or ligninlike deposition around each cell in the anticlinal (perpendicular to the surface) walls, that is, the radial and transverse walls (**Fig. 2***a*). There are no intercellular spaces in the endodermis, and the anticlinal walls between the cells are blocked by the Casparian strip, giving a "watertight" appearance to this tissue (Fig. 2*b*). When plasmolized, the protoplasts of the endodermis appear to adhere to the Casparian strip. The contiguity of protoplast and the material of the Casparian strip (Fig. 2*c*) is supposed by some to prevent the movement of solute ions and water through the walls, or between the walls and the cytoplasm and thus restrict this movement to the cytoplasm. This relation, in turn, is thought to play a role in the selective absorption of ions and in maintaining

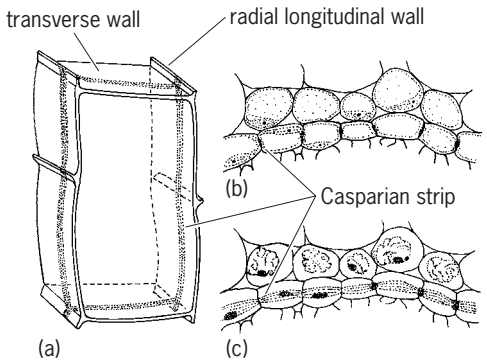

Fig. 2. Details of endodermal structure. (*a*) Diagram of cell showing location of Casparian strip. (*b*) Cells of endodermis and of ordinary parenchyma before treatment with alcohol. (*c*) Cells after treatment. Casparian strip is seen only in sectional views in *b* and *c*. (*After K. Esau, Plant Anatomy, 2d ed., John Wiley and Sons, 1967*)

hydrostatic pressure. Such a function of the endodermis, however, has not been proved conclusively. *See* CELL WALLS (PLANT).

Later development. The secondary phase in development of the endodermis consists of a complete suberization of all the walls. This event may be followed by a tertiary phase in which cellulose deposition takes place on the inner tangential and radial walls. An oxidation of phenols, naphthols, and anthrols to quinones occurs in the endodermis of roots and subaerial stems. These substances possibly serve as a barrier against the entry of pathogens such as bacteria, fungi, and nematodes. During the tertiary phase of wall development polymerized and oxidized inclusion products are deposited in the cellulose walls.

Chromosome complement. The endodermis has been found to have extra sets of chromosomes as compared with cortical and other cells in the plant. In some plants the chromosome numbers may be so high in the endodermis that four sets of chromosomes may occur in each endodermal cell. The larger amount of nuclear material and nucleic acid in the cells of the endodermis may in part account for the great capacity of endodermal cells to produce large amounts of chemical substances high in caloric energy.

Acetylenic oils. Unsaturated fatty acids develop in the endodermis, resulting in acetylenic oils in hundreds of species of plants. Acetylenic oils, called polyynes, may have the highest calorie content of any organic substance in the world. Large amounts of energy are required to produce these oils. Because of their rare chemical structure, the oils are often toxic to many organisms. The functional value of the endodermis includes the making of these and many other kinds of chemical molecules, which may serve as protective agents against invasion by bacteria and fungi and may render the plant unpalatable or poisonous to insects and herbivores. *See* SECRETORY STRUCTURES (PLANT).

Functions. The endodermis has many apparent functions: absorption of water, selection of solutes and ions, and production of oils, antibiotic phenols, and acetylenic acids. *See* CORTEX (PLANT); HYPODERMIS; PLANT TISSUE SYSTEMS.

Dick S. Van Fleet

Endophytic fungi

Fungi that live in the interior of plant host tissues without causing external symptoms. The term endophyte literally means "living within" and is contrasted with epiphyte, which means "living on the surface." These descriptive terms are used in several disciplines such as microbiology, mycology, and ecology, and at times their use is controversial. This article uses the term endophyte in a broad and topographical sense following the definition of O. Petrini (1991): "endophytes colonize asymptomatically the living, internal tissues of their hosts, even though the endophyte may after an incubation or latency period cause disease."

Endophytes range from biotrophic parasites to transient facultative saprotrophs (organisms that live on dead and decaying organic matter), and as a result their associations with hosts span the continuum from biotrophic mutualists (two species whose mutual interactions are beneficial to both) to necrotrophic, antagonistic pathogens. The endophytic habit apparently has evolved independently several times and is represented by fungi in various orders of Ascomycetes and Basidiomycetes. *See* ASCOMYCOTA; BASIDIOMYCOTA; FUNGAL ECOLOGY; FUNGI; MYCOLOGY.

Grass endophytes. The evolution and ecological relevance of endophytes are dramatically illustrated in the case of the clavicipitalean grass endophytes. These fungi are widespread in the grass family in many natural populations. Grass endophytes produce systemic infection in the aboveground parts of the grass host, developing hyphae [filaments composing the vegetative body (mycelium) of fungi] in the intercellular spaces of leaf sheaths, culms (jointed, usually hollow grass stems) [**Fig. 1**], developing flowers, and seeds. With most infections of grasses, the seed becomes the dissemination and survival structure for both the grass plant and the fungal endophyte. Another characteristic of the clavicipitalean grass endophytes is the high specificity of compatibility with the plant host. In general, a specific endophyte is adapted to a particular plant species; rarely do two strains coexist in the same plant host. This is in stark contrast to non-clavicipitalean endophytes, where many species may be present in a single host individual.

Taxonomically, grass endophytes are mainly anamorphic (a sexual state) forms of genus *Neotyphodium* in the Ascomycota family Clavicipitaceae. *Neotyphodium* spp. are frequently interspecific hybrid strains derived from the sexual forms in genus *Epichloë*. *Epichloë* causes choke disease (**Fig. 2**) in numerous grasses, producing total or partial abortion of inflorescences at the time of sporulation.

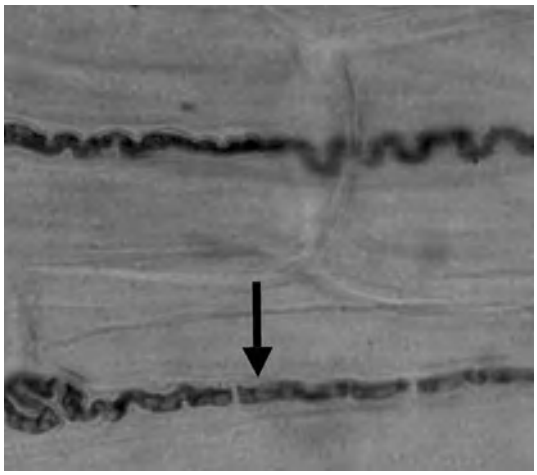

Fig. 1. ***Neotyphodium coenophialum*** **in intercellular spaces of tall fescue grass. Arrow shows hyphae: Magnification 1000×.**

Fig. 2. Choke disease caused by the clavicipitalean fungus *Atkinsonella texensis*.

The "sleepy grass" (*Achnatherum robustum*) common in western North America derives its name from the sleep-inducing alkaloid lysergic acid amide produced by an unidentified clavicipitalean endophyte inhabiting leaves, stems, and seeds of large natural populations of the grass. Horses that consume small quantities of this grass will sleep for up to 3 days, after which they recover gradually. *See* ALKALOID.

Similarly, in South America the clavicipitalean endophyte *Neotyphodium tembladerae* is widespread in several grasses—including *Poa huecu*, *Festuca hieronymi*, and *F. argentina*—where the fungus produces alkaloids that cause toxicity in grazing animals and result in muscular trembling and sometimes death. In northwestern China and Mongolia, an endophyte infects the grass commonly known as "drunken horse grass" (*Achnatherum inebrians*) and produces toxic alkaloids that cause a nervous disorder resulting in trembling and death. In South Africa, "drunk grass" (*Melica decumbens*) contains a clavicipitalean endophyte producing indole-terpenoid tremorgen (fungal mycotoxin that causes trembling) that produces nervous disorders in animals comparable to those of other endophytes.

Recognition that endophytes are responsible for many toxicoses of animal herbivores has developed only in the last three decades. C. W. Bacon et al. (1977) began the modern era of clavicipitalean endophyte investigation when they linked an endophyte (*Neotyphodium coenophialum*) in tall fescue (*Festuca arundinacea*) to widespread toxicities in cattle. The deluge of research that followed this historical epiphany included the pioneering ecological work of Keith Clay (1988) and the articulation of the "defensive mutualism hypothesis" that the clavicipitalean endophyte grass association is a mutualistic association where plants benefit from endophytes through production of fungal metabolites that are defensive against insect and mammalian herbivores.

It is now well supported that numerous clavicipitalean endophytes are defensive to their hosts. Recent research has extended endophyte defense of the hosts to include protection from fungal diseases of grasses. In certain grasses the endophytes exhibit an epiphytic stage where a network of mycelium extends over the epidermis of the grass leaf blades. These epiphyllous nets are believed to inhibit the colonization of leaf blades by potential pathogens of the host through a niche exclusion mechanism rather than through use of toxic metabolites. Endophyte-mediated resistance to drought in tall fescue grass has also been documented and may be interpreted as a defense against environmental challenges to plant survival. *See* FESCUE; GRASS CROPS.

Because of the many benefits of clavicipitalean endophytes in grasses, plant breeders have sought to take advantage of these endophytes to enhance performance of grasses for diverse applications. Seed containing endophytes can now be purchased from many seed suppliers. Turfgrass established with endophyte-infected seeds is hardier and more resistant to environmental extremes and pests requiring fewer chemical inputs to manage. For forage grasses, plant breeders have selected strains of endophytes that are low in production of alkaloids toxic to animals but high in alkaloids that impart insect resistance to grasses. *See* BREEDING (PLANT).

Although some fungi classified as endophytes may be early colonizing saprophytes, most endophytic fungi have arrived at endophytism generally from pathogenicity of the host through evolving reduced virulence or prolonged latency. In the Clavicipitaceae there is evidence that endophytes evolved initially from insect parasites. Clavicipitalean fungi are believed to have made a host shift to plants by infecting scale insects that are plant parasites themselves. This permitted clavicipitalean fungi to adapt to an epibiotic niche on plants prior to evolving endophytism. Once clavicipitalean fungi had adapted to plants as biotrophic pathogens (for example, sexually reproducing *Epichloë* spp.), evolutionary selection reduced virulence resulting in numerous asexual *Neotyphodium* endophytes. Because this process of ameliorating virulence occurred many times with different endophytes, asexual *Neotyphodium* endophytes are widespread in grasses in all habitats.

Curvularia endophytes. Non-clavicipitalean endophytes have shown similar enhancements in host resistance to environmental stresses. It has been demonstrated that a *Curvularia* species (fungal family Pleosporaceae) endophytic of rossette grass (*Dichanthellium lanuginosum*) substantially increases resistance of the infected plants to heat stress. Grass individuals containing the endophyte growing near hot springs were able to tolerate soil temperatures up to 65°C (149°F) for several days whereas endophyte-free grasses rapidly succumbed to heat stress. The mechanisms for endophyte-mediated tolerance to biological stresses such as drought and heat (thermo-tolerance) is currently unknown.

Fusarium endophytes. Another group of widespread endophytes and pathogens in many plant hosts from around the world are found in fungal genus *Fusarium* (Nectriaceae). One species, *Fusarium moniliforme*, is ubiquitous as an asymptomatic endophyte of corn wherever corn is grown. In some cases, *F. moniliforme* may exhibit a pathogenic phase (including ear, kernel, or stalk rots, and seedling blights), but generally it is entirely asymptomatic. It has been estimated that 90% of the corn grown worldwide may contain *F. moniliforme* endophytically and asymptomatically. Like the clavicipitalean endophytes, *Fusarium* endophytes produce chemical metabolites, including fumonisins and other mycotoxins. Some of these compounds are extremely toxic to animals that consume infected plants. The associations between plant and *Fusarium* endophytes have been proposed to be a defensive mutualism in that infected corn plants show enhanced resistance to several disease-causing fungi.

It is generally hypothesized that most fungal species encountered as asymptomatic endophytes may be in some way mutualistic with the host plant. However, it has been difficult in numerous cases to document the beneficial effects to hosts of some endophytes.

Woody plants: foliar endophytes. In contrast with grass endophytes, foliar endophytes of woody plants are highly diverse within the plant host. Foliar endophytes are transmitted horizontally via spore fall from plant to plant, leaves are free of the endophytes when emerging, and they become infected shortly thereafter. Fungal spores are disseminated from older leaves by water, wind, and probably insects that germinate on the leaf surface, and they infect the leaves via cuticular penetration or growth through stomata. Endophyte-enhanced defense against plant pathogens was demonstrated for the chocolate tree (*Theobroma cacao*). Taxonomically, foliar endophytes of woody plants represent a broad range of taxa in several orders and families in the Ascomycota and Basidiomycota.

Sources of bioactive compounds. Endophytic fungi have been shown to be a source of bioactive compounds that may have potential as new medicines. For example, the Pacific yew endophyte *Taxomyces andreaneae* was shown to produce the anticancer drug taxol. This chemical previously had been believed to be produced only by the host plant. *Serratia marcescens*, the bacterial endophyte of the aquatic plant *Rhyncholacis penicillata*, was found to produce the potent antifungal compound oocydin A. *Pestalotiopsis jesteri* has been shown to produce jesterone, another antifungal, apparently defensive compound. *Muscodor albus* from the cinnamon tree (*Cinnamomum zeylanicum*) was found to produce volatile antibiotics that protect the host from other microbes, a phenomenon termed mycofumigation.

Common methods to study endophytes. Endophyte detection in grass endophytes can be carried out through microscopic observation of different plant parts using fresh plants or dried herbarium material. Sheaths of internal leaves in the tiller are appropriate in the vegetative stage, and flowering culms or seeds are suitable for endophyte detection. Leaf sheaths can be processed using a clearing agent followed by staining of the endophyte. For examination of endophytes in culm tissues, culms should be cut longitudinally and the parenchyma tissue scraped out with a scalpel. Then the tissue is transferred to a drop of aniline blue or rose bengal stain on a slide and observed under the microscope. To determine the presence of endophytes in seeds, seeds are soaked overnight in a 5% solution of sodium hydroxide, after which they undergo several washes with tap water for removal of sodium hydroxide followed by microscopic observation using aniline blue or rose bengal stains. For clavicipitalean endophytes there are commercially available immunoblot assay kits for detection of the fungi in plant leaves and seeds of tall fescue and perennial ryegrass, and also for detection of ergot alkaloids produced by the endophytes in plants.

Fig. 3. Colony of *Neotyphodium lolii* isolated from *Lolium perenne*, perennial ryegrass, growing on potato dextrose agar.

For identification of endophytes in general, the common approach is direct isolation from the host plant. The procedure consists of surface sterilization of host tissues followed by culture in standard microbiological media (**Fig. 3**). Small pieces of host tissue are placed on the surface of agar medium (for example, potato dextrose agar). If the isolate sporulates, it is further characterized by its morphological features using a light microscope. If isolation results in sterile mycelia, the use of nucleic acid sequence approaches will add to strain characterization, allowing the researcher to determine the approximate phylogenetic position and assign a taxonomic category.

[Acknowledgments: The authors are grateful for support from the Fogarty International Center (NIH) under U01 TW006674 for International Cooperative Biodiversity Groups and the Rutgers Turfgrass Research Center.] Monica S. Torres; James F. White, Jr.

Bibliography. C. W. Bacon and J. F. White, Jr. (eds.), *Microbial Endophytes*, Marcel Dekker, 2000; K. Clay, Fungal endophytes of grasses: A defensive mutualism between plants and fungi, *Ecology*, 69:10–16, 1988; S. Faeth, Are endophytic fungi defensive plant mutualists?, *Oikos*, 98:25–36, 2002; O. Petrini, Fungal endophytes of tree leaves, in J. H. Andrews and S. S. Hirano (eds.), *Microbial Ecology of Leaves*, pp. 179–187, Springer-Verlag, New York, 1991; C. L. Schardl, A. Leuchtmann, and M. J. Spiering, Symbioses of grasses with seedborne fungal endophytes, *Annu. Rev. Plant Biol.*, 55:315–340, 2004; J. K. Stone, J. D. Polishook, and J. F. White, Jr., Endophytic fungi, pp. 242–270 in G. M. Mueller, G. F. Bills, and M. S. Foster (eds.), *Biodiversity of Fungi: Inventory and Monitoring Methods*, Elsevier/Academic Press, 2004.

Endoplasmic reticulum

An intracellular membrane system that is present in all eukaryotic cells. In most cells the endoplasmic reticulum is thought to consist of only one continuous membrane enclosing only a single space. However, in protozoa, some unicellular algae, and possibly some fungi, the endoplasmic reticulum occurs as separate, multiple vesicles. *See* CELL MEMBRANES.

Subdomains. Several morphologically and functionally distinct domains of this continuous membrane system can be distinguished. At the level of the nuclear pores, the inner nuclear membrane is continuous with the outer nuclear membrane; both membranes together are referred to as the nuclear envelope. The outer nuclear membrane in turn is continuous with the rough endoplasmic reticulum, which contains specialized regions, termed transitional elements, and is continuous with the smooth endoplasmic reticulum. The two membranes of the nuclear envelope enclose the perinuclear space. The rough and smooth endoplasmic reticula and the transitional element enclose a space called the intracisternal space, or lumen. Both intracisternal and perinuclear spaces form a single compartment. All nucleated cells contain at least a nuclear envelope, but the amount of smooth and rough endoplasmic reticula varies greatly among different cell types. *See* CELL NUCLEUS.

This single-membrane system undergoes an extensive and reversible fragmentation during mitosis. During the first stage of mitosis (prophase), extensive fission of the nuclear envelope and the endoplasmic reticulum takes place, yielding numerous closed vesicles. At the end of mitosis (telophase), the separated vesicles fuse and a single continuous membrane is restored. Fragmentation ensures that the two daughter cells each end up with approximately half of the mother's nuclear envelopes and endoplasmic reticulum. *See* MITOSIS.

When cells are broken by mechanical homogenization, the endoplasmic reticulum and often part of the outer nuclear membrane are fragmented. To a large extent, the fragments close at their ends to form vesicles that can be separated on the basis of their size and density, yielding fractions that represent rough microsomes, which are derived from rough endoplasmic reticulum, and smooth microsomes, which are derived from the smooth endoplasmic reticulum. Biochemical analyses of these fractions have revealed many of their functions.

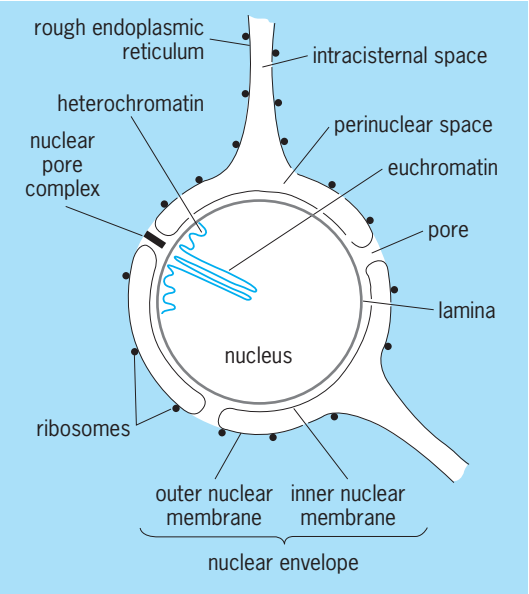

Fig. 1. Nuclear envelope and associated structures.

Nuclear envelopes. When cells are broken mechanically, most of the connections between nuclear envelope and rough endoplasmic reticulum (**Fig. 1**) are severed. The nuclei, still surrounded by their nuclear envelopes, can easily be separated from the homogenate because of their large size and high density. Following enzymatic digestion of the chromatin with nuclease, the nuclear envelope can readily be isolated from the nuclease digest with its principal associated organelles, the nuclear pore complexes, and the nuclear lamina still attached. Inner and outer nuclear envelope membranes have not yet been separated.

Associated with the nuclear envelope are the nuclear pore complexes (Fig. 1). An average mammalian nucleated cell contains approximately 5000 pore complexes, each with an estimated mass of 10^8 daltons. Pore complexes are thought to regulate macromolecular traffic into and out of the nucleus and to function in the three-dimensional organization of the genome.

Associated with the nuclear side of the inner nuclear membrane is the nuclear lamina, a fibrous meshwork 15–50 nanometers thick that is intercalated between inner nuclear membrane and chromatin. The nuclear lamina is composed of the lamins, which structurally belong to the family of intermediate filament proteins. The lamins (specifically lamin B) are physically connected to the cytoplasmic intermediate filament proteins at the level of the nuclear pore complex and are thought to function in the three-dimensional organization of the chromatin.

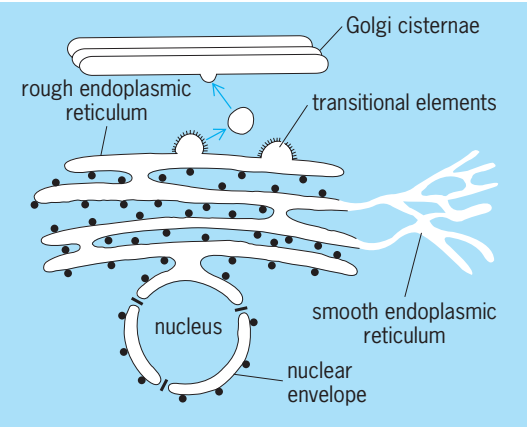

Fig. 2. Nuclear envelope, connected to the rough and smooth endoplasmic reticulum. The rough endoplasmic reticulum is linked to the cis cisterna of the Golgi complex by vesicles that shuttle between the two compartments.

Associated with the outer nuclear membrane are ribosomes. In fact, the outer nuclear membrane and the rough endoplasmic reticulum appear to be functionally equivalent.

The inner and outer nuclear membranes are likely to differ in their composition of integral membrane proteins. So far, only one such protein, the lamin B receptor, has been assigned exclusively to the inner nuclear membrane. This protein can therefore serve as a marker for the inner nuclear membrane. Another integral membrane glycoprotein of 190 kilodaltons has been assigned to the pore-wall region of the nuclear envelope; it is thought to function in anchoring nuclear pore complexes.

Rough endoplasmic reticulum. The term rough endoplasmic reticulum is based on the morphologic appearance of attached ribosomes, which are absent from smooth endoplasmic reticulum. Another morphologic distinction is the organization of the former in interconnected flattened sacs (called cisternae), whereas the latter forms a tubular network (**Fig. 2**). *See* RIBOSOMES.

The rough endoplasmic reticulum is the site of translocation of secretory and lysosomal proteins from the cytosol to the intracisternal space, and of integration into the membrane of integral membrane proteins. Except for integral membrane proteins of chloroplast, mitochondria, and peroxisomes, essentially all other integral membrane proteins are integrated into the endoplasmic reticulum and either remain there (resident endoplasmic reticulum membrane proteins) or are subsequently distributed to other cellular membranes.

The signal hypothesis was formulated to explain how these proteins are targeted to and then translocated across or integrated into the endoplasmic reticulum membrane. Its tenets are that all polypeptides targeted to this membrane contain a discrete sequence (termed the signal sequence), that a complex machinery recognizes this sequence, and that recognition triggers the opening of a proteinaceous channel (composed of several subunits) through which the polypeptide passes across the membrane. In the case of membrane proteins, the existence of an additional topogenic sequence, the so-called stop-transfer sequence, was postulated. This sequence is thought to trigger opening of the channel to the lipid bilayer to abort translocation and thus integrate the protein into the lipid bilayer.

Many predictions of the signal hypothesis have been verified. The presence of a signal sequence in these proteins has been demonstrated. The critical information for targeting is encoded in about 10–15 amino acid residues of the signal sequence, which in most cases is cleaved during or shortly after translocation by a special endoplastic reticulum–associated signal peptidase. The machinery for recognition of the signal sequence and targeting to the endoplasmic reticulum is complex. As soon as the signal sequence has emerged from within a channel of the large ribosomal subunit, it is recognized by the signal recognition particle, a ribonucleoprotein particle consisting of one 7S ribonucleic acid (RNA) molecule and six distinct polypeptide chains. The ensemble of ribosome, signal recognition particle, and nascent chain is then targeted to the signal recognition particle receptor (located exclusively in the endoplasmic reticulum). This receptor displaces the signal recognition particle and frees the signal sequence to interact with a second signal recognition system in the endoplasmic reticulum membrane. The second recognition step may be coupled to the opening of a proteinaceous channel; however, the existence of such a channel has not yet been demonstrated. The chain would then pass through that channel, or in the case of integral membrane proteins, the channel would be opened to the lipid bilayer by a stop-transfer sequence. After passage, the channel would close, terminating translocation.

The fact that the integration of integral membrane proteins requires proteins of the same type indicates that cells cannot assemble membranes de novo. Preexisting membranes with an integrated translocation machinery are required to make more such machinery and more membranes. Thus, if the daughter cells would not inherit endoplasmic reticulum from the mother cell, they would be able to synthesize but not integrate membrane proteins.

The rough endoplasmic reticulum also contains numerous enzymes, most of which are involved in the modification of the nascent chain on the cisternal side. For example, all the enzymes (and the required sugar nucleotide transporters) for the synthesis of a lipid-linked oligosaccharide and for the transfer of the oligosaccharide moiety to an asparagine residue of a polypeptide are located in the rough endoplasmic reticulum.

Thus the main function of the rough endoplasmic reticulum and the outer nuclear membrane is to serve as a port of entry of secretory, lysosomal, and integral membrane proteins and as the site of their initial modification. *See* CYTOCHEMISTRY.

Transitional elements. Secretory and lysosomal proteins as well as those integral membrane proteins that are not residents of the endoplasmic reticulum

are next transported to the cis Golgi cisternae. The transitional elements represent sites of transport from the rough endoplasmic reticulum. Coated vesicles carrying proteins to be transported form at these sites and, after uncoating, eventually fuse with the cis Golgi cisternae. *See* GOLGI APPARATUS.

Smooth endoplasmic reticulumm. Smooth endoplasmic reticulum contains enzymes for phospholipid biosynthesis, steroid biosynthesis, and drug detoxification. The principal phospholipid synthesized in the smooth endoplasmic reticulum is phosphatidyl choline. The enzymes for assembly from fatty acyl coenzyme A, glycerophosphate, and cytosine diphosphatecholine are integral membrane proteins with their active sites exposed to the cytosol. Phosphatidyl choline then becomes part of the cytoplasmic leaflet of the endoplasmic reticulum lipid bilayer. Transport to the luminal leaflet may be catalyzed by an enzyme called a flippase. Certain hydrophobic compounds that dissolve in the lipid bilayer can be hydroxylated and further modified by enzymes, thus becoming water-soluble and excretable. *See* CELL (BIOLOGY); CELL ORGANIZATION; ENZYME.

Günter B. Blobel

Bibliography. W. M. Becker, *The World of the Cell*, 4th ed., 2000; T. C. Cheng (ed.), *The Structure of Membranes and Receptors*, 1984; C. A. DeDuve, *A Guided Tour of the Living Cell*, 1985; E. Holtzman and A. B. Novikoff, *Cells and Organelles*, 3d ed., 1984; R. C. Warren, *Physics and the Architecture of Cell Membranes*, 1987; S. L. Wolfe, *Cell Ultrastructure*, 1985.

Endopterygota

A division (also known as Holometabola) of the subclass Pterygota, including those insects that undergo complete metamorphosis during their life cycle. That is, individual development goes through four distinct stages: egg, larva (trophic), pupa (reconstructive), and adult (reproductive). The larval and adult stages often live in very different adaptive zones and have very different life forms, necessitating a quiescent pupal stage in which extensive restructuring takes place. Typically growing through a sequence of immature forms (instars) punctuated by integumental molts, the larval instars correspond to those of nymphal Exopterygota in most ways, but are distinct in having wings, internalized as tiny wing buds. The buds are tucked out of the way of the larva, grow very rapidly in the pupal stage, and expand and quickly become functional in the newly hatched adult.

Internalization of the developing wings and the advent of the pupa combine to form a key general adaptation that breaks the constraining linkage between trophic and reproductive life forms, allowing each to evolve for its own special function. For example, most larvae have evolved life forms with legs that are reduced or absent, enabling them to live and feed cryptically within a substrate. This adaptation probably accounts for the rapid rise of the Endopterygota to a dominant position among the insects from the Permian to the present.

The division Endopterygota comprises a large majority of all insects, distributed in the orders Coleoptera, Neuroptera, Strepsiptera, Mecoptera, Siphonaptera, Trichoptera, Lepidoptera, Diptera, and Hymenoptera. It is thought to be monophyletic. *See* EXOPTERYGOTA; INSECTA; PTERYGOTA.

William L. Brown, Jr.

Bibliography. O. W. Richards and R. G. Davies, *Imm's General Textbook of Entomology*, 2 vols., 10th ed., 1994; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, vol. 2, 1982.

Endorphins

A family of endogenous, morphinelike peptides present within the central nervous system. The enkephalins, the first members of the endorphin family, were discovered by John Hughes and Hans Kosterlitz in 1975 and opened research into a wide number of centrally active neuropeptides. The term endorphin is generic, referring to all the opioid peptides, while specific peptides are given individual names, such as the enkephalins, dynorphins, and β-endorphin. Their isolation, shortly after the identification of the opioid receptors, brought together the concept of endogenous opioid peptide systems within the central nervous system which can modulate pain perception and through which opioids act. *See* OPIATES; PAIN.

Morphine, codeine, and their many synthetic and semisynthetic analogs are effective painkillers that act through specific recognition sites, or receptors, localized on the surface of neurons within selected brain regions. The existence of these highly specific receptors implied that, for example, morphine was mimicking endogenous compounds within the brain with morphinelike actions. This concept of an endogenous pain-modulating system provided an explanation for prior clinical and experimental observations. In a classic study, it was reported that soldiers wounded in World War II required fewer painkillers than civilians undergoing elective surgery, even though the soldiers had far more extensive injuries. Clearly, some aspect of the stress of combat affected the perception of pain. More than 20 years later, scientists reported that electrical stimulation of a specific brain region, the periaqueductal gray, produces pain relief, or analgesia, in animals; this was demonstrated later in humans. This pain relief was reversed by the drug naloxone, a potent and highly selective opiate antagonist that reverses morphine actions. The ability of naloxone to reverse this analgesia implied that the stimulation was releasing endogenous morphinelike, or opioid, materials, which have since been termed endorphins. Three families of opioid receptors involved with pain modulation have been identified that selectively interact with different classes of opioids and endorphins: mu (morphine and the endomorphins), kappa

TABLE 1. Structures of the opioid peptides

[Leu]enkephalin	**Tyr-Gly-Gly-Phe-Leu**
[Met]enkephalin	**Tyr-Gly-Gly-Phe-Met**
Dynorphin A	**Tyr-Gly-Gly-Phe-Leu**-Arg-Arg-Ile-Arg-Pro-Lys-Leu-Lys-Trp-Asp-Asn-Gln
Dynorphin B	**Tyr-Gly-Gly-Phe-Leu**-Arg-Arg-Gln-Phe-Lys-Val-Val-Thr
α-Neoendorphin	**Tyr-Gly-Gly-Phe-Leu**-Arg-Lys-Tyr-Pro
β-Endorphin	**Tyr-Gly-Gly-Phe-Met**-Thr-Ser-Glu-Lys-Ser-Gln-Thr-Pro-Leu-Val-Thr-Leu-Phe-Lys-Asn-Ala-Ile-Ile-Lys-Asn-Ala-Tyr-Lys-Lys-Gly-Glu
Endomorphin 1	Tyr-Pro-Trp-Phe-NH_2
Endomorphin 2	Tyr-Pro-Phe-Phe-NH_2

(dynorphins), and delta (enkephalins). *See* MORPHINE ALKALOIDS.

Structure and processing. The enkephalins were the first endogenous opioid peptides, or endorphins, to be isolated. These two pentapeptides have the same first four amino acids followed by either methionine or leucine. A number of opioid peptides have been isolated (**Table 1**). With the exception of the endomorphins, all contain the sequence of either methionine (Met) enkephalin or leucine (Leu) enkephalin at their NH_2 terminus. This common sequence among the various peptides raised questions as to whether the enkephalins might be breakdown products of larger peptides, such as β-endorphin and the dynorphins. However, three separate genes have been cloned that encode the precursor proteins for the enkephalins, the dynorphins, and β-endorphin, which clearly demonstrates the independence of each of these endorphin families of peptides.

The enkephalins are generated from a larger precursor peptide through the actions of enkephalin convertase, a carboxypeptidase. The pro-enkephalin gene codes for a peptide containing six copies of [Met^5]enkephalin and one copy of [Leu^5]enkephalin which is generated by enzymatic cleavage. The enkephalins are very sensitive to degradation by endopeptidases, particularly enkephalinase. Inhibitors of this enzyme have been synthesized and have modest analgesic actions, presumably by enhancing the enkephalin levels in the nervous system. The problem of metabolism has been greatly diminished by substituting an unnatural D-amino acid in the second position of a series of enkephalin analogs. A large variety of compounds have thus been synthesized, some of which are quite potent and active in animals and humans, even after systemic administration. Interestingly, alterations in the sequence of these peptides can change their selectivity toward the classes of opioid receptors.

The dynorphins and α-neoendorphin are derived from the pro-dynorphin gene, also termed the pro-enkephalin B gene, which is distinct from the one involved with enkephalin synthesis. Again, the presence of two basic amino acids directs processing to form dynorphin A, dynorphin B, and α-neoendorphin. These peptides have not been studied as extensively as the enkephalins, and few analogs are available. Like the enkephalins, all three are very sensitive to proteases and are broken down very rapidly in the brain.

β-Endorphin is perhaps the most interesting peptide. It is encoded by the pro-opiomelanocortin gene. This gene generates a large peptide, termed Big ACTH, which then is processed to form β-lipotropin and the stress hormone adrenocorticotropin (ACTH). β-Lipotropin is further processed to yield β-endorphin and β-MSH (melanocyte-stimulating hormone). Thus, β-endorphin is cogenerated with important, nonopioid hormones. *See* ADENOHYPOPHYSIS HORMONE.

Regional distributions and actions. The enkephalins are distributed unevenly throughout the brain with very high levels in the basal ganglia, the thalamus, and the periaqueductal gray (**Table 2**). In addition, there are high concentrations of enkephalins in the adrenal medulla, where they are coreleased with norepinephrine in response to stress, among other stimuli. The dynorphins and α-endorphin are located within the central nervous system with a distribution similar to that of the enkephalins. *See* STRESS (PSYCHOLOGY).

β-Endorphin has a unique distribution within the brain compared to the other opioid peptides. Unlike the enkephalins and dynorphins, which have high concentrations within a large number of brain

TABLE 2. Tissue distribution of opioid peptides

Location	Enkephalin	β-Endorphin	Dynorphin
Central nervous system			
Spinal cord	Dorsal horn (laminae I and II)	Absent	—
Brainstem	Cells and fibers	Fibers only	Fibers: medial forebrain bundle
Cerebellum	Absent	Absent	Absent
Diencephalon	High: thalamus, hypothalamus	Fibers: thalamus Cells: hypothalamus	Fibers: substantia nigra Cells: hypothalamus
Basal ganglia	High: globus pallidus	Low	Medium
Telencephalon	High: amygdala Low: hippocampus, cerebral cortex	Low	High: hippocampus Medium: cerebral cortex Low: amygdala
Peripheral tissue			
Pituitary gland	Low	Intermediate lobe	Posterior lobe
Gastrointestinal tract	Present	Present	Present
Adrenal medulla	High	—	—
Autonomic nervous system	Present	—	—
Placenta	—	High	—

regions, β-endorphin has been identified in only a single group, or nuclei, of cells within the hypothalamus that project to a number of regions. Its highest levels are in the pituitary gland. Within the pituitary, both ACTH and β-endorphin are derived from the same precursor protein and are located within the same cells. Stimuli that release ACTH, a stress hormone which in turn induces the adrenal gland to release steroids, also corelease β-endorphin at the same time. Thus, stressful stimuli that release ACTH and norepinephrine also release both β-endorphin from the pituitary and enkephalins from the adrenal into the blood. This is particularly intriguing in view of the decreased perception of pain reported under periods of stress. *See* ENDOCRINE SYSTEM (VERTEBRATE); PITUITARY GLAND.

All the endorphins are analgesic despite the fact that they act through different classes of opiate receptors. Unlike the stabilized synthetic enkephalin analogs, the natural endorphins are not active when given systemically due to their rapid degradation and their difficulty traversing the blood–brain barrier. The natural endorphins relieve pain only when given directly into the brain. However, the endorphins are likely to have a far more diverse range of actions beyond modulation of pain, as suggested by their presence in brain regions unrelated to pain perception, such as the striatum. Thus, many questions remain regarding the full physiological significance of these agents. *See* NERVOUS SYSTEM (VERTEBRATE).

Gavril Pasternak

Bibliography. B. M. Cox, A. Goldstein, and C. H. Li, Opioid activity of a peptide, β-lipotropin(61–91), derived from a β-lipotropin, *Proc. Nat. Acad. Sci. USA*, 73:1821, 1976; A. Goldstein et al., Porcine pituitary dynorphin: Complete amino acid sequence of the biologically active heptapeptide, *Proc. Nat. Acad. Sci. USA*, 78:7219–7223, 1981; J. G. Hardman and L. E. Limbird (eds.), *Goodman & Gilman's The Pharmacological Basis of Therapeutics*, 2001; J. Hughes et al., Identification of two related pentapeptides from the brain with potent opiate agonist activity, *Nature*, 258:577–579, 1975; D. J. Mayer and J. C. Liebeskind, Pain reduction by focal electrical stimulation of the brain: An anatomical and behavioral analysis, *Brain Res.*, 68:73–93, 1973.

Endotoxin

A biologically active substance produced by bacteria and consisting of lipopolysaccharide, a complex macromolecule containing a polysaccharide covalently linked to a unique lipid structure, termed lipid A. All gram-negative bacteria synthesize lipopolysaccharide, which is a major constituent of their outer cell membrane. One major function of lipopolysaccharide is to serve as a selectively permeable barrier for organic molecules in the external environment. Different types of gram-negative bacteria synthesize lipopolysaccharide with very different polysaccharide structures. *See* CELL MEMBRANES; LIPID; POLYSACCHARIDE.

Lipopolysaccharide. The polysaccharide component of lipopolysaccharide is often divided into two regions—the core polysaccharide and the O-antigen polysaccharide. The core region contains common hexose sugars, such as glucose and galactose (found in the outer core), and unique 7- and 8-carbon sugars, such as ketodeoxyoctulosonic acid (localized in the inner core). The O antigen consists of repeating oligosaccharide units, each containing four to six hexose sugars. The O antigen is different for each strain of microorganism, whereas the core polysaccharide is more consistent among species. Some pathogens, such as *Neisseria* and *Hemophilus*, synthesize a lipopolysaccharide which is devoid of O antigen. *See* GLUCOSE; GONORRHEA; MENINGITIS; OLIGOSACCHARIDE.

Covalently attached to the ketodeoxyoctulosonic acid of the inner-core polysaccharide is lipid A, with a structure that is very similar in almost all gram-negative microorganisms. This lipid consists of a backbone of two molecules of the sugar *N*-acetylglucosamine, to which are attached five to seven long-chain fatty acids. Sometimes attached to lipid A are additional phosphate or phosphorylethanolamine groups. Since lipid A is a chemically conserved structure among all species of gram-negative bacteria, the biological activity of endotoxic lipopolysaccharide resides almost entirely in the lipid A component

Host responses. When lipopolysaccharides are released from the outer membrane of the microorganism, either by natural cell death, by destruction of the bacterium following treatment with antibiotics, or by natural shedding, significant host responses are initiated in humans and other mammals. It is generally accepted that lipopolysaccharides are among the most potent microbial products, known for their ability to induce pathophysiological changes, in particular fever and changes in circulating white blood cells. For example, in humans as little as 4 nanograms of purified lipopolysaccharide per kilogram of body weight (0.00000008 ounce in a normal adult male) is sufficient to produce a rise in temperature of about 3.6°F (2°C) in several hours. This profound ability of the host to recognize endotoxin is thought to serve as an early warning system to signal the presence of gram-negative bacteria.

Unlike most microbial protein toxins (which have been termed bacterial exotoxins), endotoxin is unique in that its recognized mode of action does not result from direct damage to host cells and tissues. Rather, endotoxin stimulates cells of the immune system, particularly macrophages, and of the vascular system, primarily endothelial cells, to become activated and to synthesize and secrete a variety of effector molecules that cause an inflammatory response at the site of bacterial invasion. These effector molecules include a number of important cytokines, particularly tumor necrosis factor, interleukin-1 and -6, and interferon, as well as inflammatory mediators, including prostaglandins and leukotrienes. These mediator molecules promote the host response which results in elimination of the

invading microbe. One important consequence of endotoxin interaction with phagocytic cells is that these cells become more efficient at engulfing and killing bacteria. Endotoxin-stimulated macrophages also become more efficient at killing tumor cells. Thus, under these circumstances lipopolysaccharide is not a toxin at all, but serves an important function by helping to mobilize the host immune system to fight infection. *See* CYTOKINE; IMMUNOLOGY.

Even though endotoxin stimulation of host cells is important to host defense against infection, overstimulation due to excess production of endotoxin can lead to serious consequences. When excess mediator production occurs, the elicited pathophysiologic changes (termed the systemic inflammatory response syndrome) are expanded to include hypotension (abnormally low blood pressure), blood clotting, and changes in metabolism, heart rate, and respiration. These changes can severely damage many organs, including the lung, liver, and kidney. Endotoxin-induced multiple-organ failure continues to be a major health problem, particularly in intensive care; it has been estimated that as many as 50,000 deaths annually occur in the United States as the result of endotoxin-induced shock.

Immunization. Because of the extensive variability in chemical structure of the O antigen of various species, and the wide spectrum of gram-negative microorganisms that can cause disease, immunization of humans with endotoxin vaccines to protect against endotoxin shock has not been considered practical. In contrast, with bacterial protein exotoxins (such as diphtheria toxin, pertussis toxin, tetanus toxin) immunization leads to long-lasting immunologic protection against disease. Efforts to provide immunologic protection against endotoxin-related diseases have focused upon development of antibodies that recognize the conserved lipid A structure of endotoxin as a means of passive protection against the lethal effects of this microbial product. *See* BACTERIA; MEDICAL BACTERIOLOGY; VACCINATION.

David C. Morrison

Bibliography. D. C. Morrison and J. L. Ryan (eds.), *Bacterial Endotoxic Lipopolysaccharides*, vol. 1, 1992; E. T. Rietschel and H. Brade, Bacterial endotoxins, *Sci. Amer.,* 267:54–61, 1992; J. L. Ryan and D. C. Morrison (eds.), *Bacterial Endotoxic Lipopolysaccharides*, vol. 2, 1992.

Energy

One of the two fundamental constituents of the universe, interconvertible with the other constituent, matter. All physical processes involve exchange of energy, or conversion from one of its named forms (including relativistic mass-energy) into another. The total energy (including mass) of a closed system is a conserved quantity. *See* MATTER (PHYSICS).

Intuitive notions. On the human scale, energy is most familiar in the form of mechanical work and transfer of heat. The word "energy" is derived from Greek meaning "inner work." This is intuitively reasonable from the everyday notion of doing work because physical labor expends the body's energy.

The causal relationships among energy, force, work, power, kinetic energy, potential energy, heat, and temperature are also intuitive. A person pushing a rock up a slope applies the force of his or her muscles. While the rock is in motion, it has kinetic energy (relative to the ground) that could, for example, dent a tree trunk that might be in the rock's path. The amount of work done increases proportionally with the distance the rock is moved, and it results in a proportionate reduction of the person's internal energy resources; hence the notion that work and energy spent are equivalent. To push the rock faster, the person must apply more power, or energy per unit time. The work done on the rock is converted into an increase of gravitational potential energy, because the rock is now higher up on the slope, and into heat energy, resulting from friction, that raises the temperature of the rock and proximate ground. The potential energy of the rock is so-named because it has the potential to be converted into other forms of energy such as kinetic energy in the event that the rock should be allowed to roll freely downhill. *See* FORCE; HEAT; POWER; TEMPERATURE; WORK.

Formal and quantitative concepts. The intuitive notions of energy and its related parameters were first quantified with Newton's three laws of motion. Classical mechanics recognizes two types of energies:

1. Energy due to relative motion, or kinetic energy: A mass m moving with speed v relative to an observer has kinetic energy $T = {}^1/_2 mv^2$.

2. Energy that is stored in restraining motion against forces, or potential energy: For example, lifting a mass m from ground level to a height h increases its gravitational potential energy by $\Delta U = mgh$, where g is the downward acceleration due to Earth's gravitational field. *See* GRAVITATION; NEWTON'S LAWS OF MOTION.

With the development of atomic and subatomic physics came understanding of new types of forces with which potential energies could be associated. The electric and magnetic forces, with their associated potentials, act on electrical charges. In the nineteenth century, they were revealed to be two manifestations of the electromagnetic force. Twentieth-century research revealed that subatomic particles interact via the weak nuclear force and the strong nuclear force. *See* ELECTRIC FIELD; FUNDAMENTAL INTERACTIONS; MAGNETISM; MAXWELL'S EQUATIONS; STRONG NUCLEAR INTERACTIONS; WEAK NUCLEAR INTERACTIONS.

Most force fields of interest, such as gravity and electromagnetic force, are conservative. This means that taking a particle that "feels" that force and moving it through any round trip (starting from any point in the field and ending on the same point) will require no net energy input (the particle's potential energy will be unchanged).

Units of energy. The unit of energy in the International System (SI) is the joule (J), defined as the force of one newton (N) acting through a distance of one meter. Power, the time rate of flow of energy, is

measured in watts (W): one watt of power is defined as the rate of energy flow of one joule per second. Electrical utilities use the kilowatt-hour (kWh), the energy delivered by one kilowatt over a period of one hour; one kWh equals exactly 3.600×10^6 joules. Energies of chemical interactions are measured in kilocalories per mole, where one calorie equals 4.187 joules. A kilocalorie of heat will raise the temperature of a liter of water by about one degree Celsius. *See* METRIC SYSTEM; PHYSICAL MEASUREMENT; UNITS OF MEASUREMENT.

In high-energy physics, subatomic-particle interactions frequently involve mass changing to radiant energy, and vice versa. It is therefore more convenient to express both energy and mass by the same unit of measure, the electronvolt (eV): the change in potential energy of one electron when moved through an electrical potential difference of one volt; 1 eV equals 160.2×10^{-21} joule. *See* ELECTRONVOLT.

Law of conservation of energy. The law of conservation of energy, also known as the first law of thermodynamics, is one of the most important laws of physics. It states that the total energy of closed, isolated systems is always constant. (More precisely, it is the mass-energy that is conserved, of major concern only for relativistic velocities and nuclear reactions.) This law is a result (according to Nöther's theorem) of the homogeneity of the universe with respect to time; that is, with all else being equal, the outcome of any experiment is independent of when it is performed. *See* CONSERVATION LAWS (PHYSICS); CONSERVATION OF ENERGY; SYMMETRY LAWS (PHYSICS).

Many periodic phenomena are due to the cyclic interchange between potential energy and kinetic energy, with their sum always constant. The classical example is a swinging pendulum whose potential energy is maximum at the highest part of its swing, its kinetic energy at the lowest. *See* PENDULUM.

Energy and special relativity. Albert Einstein's special theory of relativity introduced the notion of mass-energy $E = mc^2$, where c is the speed of light. In the mass's own frame of reference, within which the mass is stationary, this formula gives the rest-mass energy. For small (nonrelativistic) velocities, the total energy approximately equals the rest-mass energy plus the term $^1/_2 mv^2$, exactly the classical formula for kinetic energy. *See* REST MASS.

When velocities are routinely relativistic, as in high-energy physics, it is convenient to express universally true physical relations as four-vectors. The momentum four-vector (also called the momenergy) embodies the conservation laws of both energy and momentum for all moving observers. *See* CONSERVATION OF MOMENTUM; MOMENTUM; RELATIVITY.

Energy in quantum mechanics. Quantum mechanics, and its offshoot quantum field theory, is founded on the principle of the quantum, the smallest amount of energy exchangeable in a given interaction. (Linear momentum, angular momentum, and mass are also quantized.) *See* QUANTUM FIELD THEORY; QUANTUM MECHANICS.

One interesting property of quanta of electromagnetic radiation is that they cannot have energies of zero even in empty space. This nonzero energy is called the zero-point energy and is experimentally confirmed by the Casimir effect: two parallel, closely spaced, uncharged conducting plates in empty space will attract each other.

Mechanical energy can be observed to be quantized: vibrational energy propagating through periodic media, such as silicon crystals, does so in quanta called phonons; essentially, these are quanta of sound energy. *See* LATTICE VIBRATIONS; PHONON.

A second principle of quantum mechanics, Heisenberg's uncertainty principle, limits the accuracy with which energy (as well as length, velocity, time, and position) may be measured. The error ΔE in measuring energy over time interval Δt is related by the inequality $\Delta E \cdot \Delta t \geq h/(2\pi)$, where h is Planck's constant. One consequence is that as masses (which, of course, are equivalent to energies) are examined at ever-shorter time scales, the uncertainties in the measured values increase. For short-enough particle interactions, particle physicists can invoke virtual particles that exist only briefly during the transformation of colliding particles into postcollision products. During their putative existence, they may violate the law of conservation of energy, a paradox resolved by the fact that they are never actually observed. *See* UNCERTAINTY PRINCIPLE.

System evolution, entropy, and the second law of thermodynamics. In a closed, isolated system with a certain distribution of energy within it there are many (usually practically uncountable) ways the energy could be redistributed. The second law of thermodynamics indicates what the ultimate distribution of energy within the system will be after a long time. It invokes the concept of entropy, a measure of the disorder within the system. Since the number of ways the water molecules can be arranged in a steam vessel is larger than the ways they can be arranged to form ice, a given quantity of water has a lower entropy as ice crystals than as steam. The second law states that a closed, isolated system will evolve over time so as to maximize the entropy. Thus, heat energy flows (if allowed) from hotter areas to colder ones, eventually equalizing the temperature. This maximizes the entropy; it also serves as an indicator of the flow of time (specifically, thermodynamic time). *See* ENTROPY; THERMODYNAMIC PRINCIPLES; TIME, ARROW OF.

All natural processes that locally decrease the entropy increase it elsewhere. When sodium and chlorine atoms migrate in aqueous solution to form a salt crystal, their entropy is decreased. However, the heat of crystallization that is released warms the solution, thereby increasing its entropy. Similarly, the biological processes of life create order but at the expense of the surroundings: consuming food creates highly ordered tissue but increases disorder via waste products and heat.

Types of energy. A number of named types of energy are found in scientific and technical literature. Most are macroscopic manifestations of

electromagnetic interactions at the atomic level.

Heat became understood, with the development of the atomic theory in the nineteenth century, as a manifestation of the collective kinetic energy of atoms in bulk materials. Temperature is a measure of their average kinetic energy. In solids, the motions are confined to vibrations of the atoms about equilibrium positions. In gases, the energy is equally partitioned into linear flight (between collisions), rotations, and vibration modes of the molecules. *See* KINETIC THEORY OF MATTER; STATISTICAL MECHANICS.

Elastic energy is potential energy stored when solids are deformed by various kinds of strains, such as compression or torsion. It is due to dislocation of the constituent atoms of the solid from their "relaxed" minimum-energy configurations. *See* ELASTICITY.

Surface-tension energy, which is responsible for the spherical shape of droplets, is the potential energy stored in the tangential cohesive forces between molecules at the surface of a liquid. *See* SURFACE TENSION.

The field of chemistry involves various kinds of forces, often expressed in terms of potential energies. Some are the ion-ion force, van der Waals force, and dipole-dipole force, of which the hydrogen bond is the strongest. The binding energy of a covalent or other bond between two atoms of a molecule is the energy required to pull them apart. Atomic cohesive energy is the energy required to separate the atoms of a solid far enough that they no longer feel cohesive forces. *See* CHEMICAL BONDING; COHESION (PHYSICS); DIPOLE-DIPOLE INTERACTION; INTERMOLECULAR FORCES.

Burning is a process of oxidation that converts fuel plus oxygen into combustion products. The total binding energy of fuel plus air is larger than of smoke plus ashes, and the difference is emitted as light and heat. Energy to power biological processes is stored in the binding energies of biomolecules such as adenosine triphosphate (ATP). This energy is also released through oxidation processes such as the Krebs cycle. *See* ADENOSINE TRIPHOSPHATE; BIOLOGICAL OXIDATION; CITRIC ACID CYCLE; COMBUSTION; METABOLISM.

Bending, stretching, and twisting of molecular bonds store and release energy. Analysis of molecules by infrared spectroscopy uses the fact that the various distortions of the bonds have telltale resonance frequencies. *See* INFRARED SPECTROSCOPY; MOLECULAR STRUCTURE AND SPECTRA.

The shapes that molecules assume are those that minimize the total potential energy due to the forces among their constituent atoms. This simple driving principle yields surprisingly complicated configurations in proteins. Deoxyribonucleic acid (DNA) codes for the one-dimensional sequence of amino acids, sometimes thousands of units long. These spontaneously crumple, by shedding potential energy, into precise and unique three-dimensional configurations that determine their highly specific biological functions. Predicting the ultimate shape of a given amino-acid chain is called the protein-folding problem, which has not yet been solved. *See* PROTEIN; PROTEIN FOLDING.

Nuclear energy results from release of binding energies of atomic nuclei, orders of magnitude larger than the binding energies of molecules in chemical reactions. For example, solar nuclear fusion combines pairs of deuterium nuclei into single nuclei of helium having a smaller binding energy; the difference is released as heat. The binding-energy differences are large enough to be evident as a conversion of mass to energy. Much of the mass of nucleons is due to the binding energies confining the quarks that comprise them. Chemical reactions, too, result in changes of mass, but generally too small in magnitude to be measurable. *See* ELEMENTARY PARTICLE; NUCLEAR BINDING ENERGY; NUCLEAR FUSION; PROTON-PROTON CHAIN; QUARKS.

Energy as a commodity. The production of electricity and its consumption illustrate how energy can go through many forms. The energy in the coal burned in a power plant originated during the Earth's Carboniferous Period, 300 million years ago, when the Sun's energy drove weather systems to carry ocean water as rain to land-based plants. The plants' chlorophyll also used the Sun's energy to convert rainwater and airborne carbon dioxide (CO_2) to produce plant tissue, mainly cellulose and other polysaccharides. Over eons, the energies of geological processes, such as the folding of tectonic plates, concentrated the carbon content through prolonged heat and pressure. This resulted in the fossil fuels, of which coal has the highest carbon content. *See* CARBONIFEROUS; COAL; FOSSIL FUEL; PHOTOSYNTHESIS.

In one type of power plant, coal is burned to heat water in boilers (sending the carbon back into the atmosphere as CO_2). The heat converts water to high-pressure steam that is passed to steam turbines, where the potential energy of steam pressure is converted to kinetic energy of the turbines' rotors. The rotors are coupled mechanically to the electric generators, whose rotors carry conductive metal windings that are rotated within a strong magnetic field. This results in electromotive force (emf) that manifests itself as electrical power flowing out of the generators. The electrical energy is passed to the power grid, which distributes it to the power plant's customers. *See* ELECTRIC POWER GENERATION; ELECTRIC POWER SYSTEMS; ELECTROMOTIVE FORCE (EMF); GENERATOR.

In the home, the electrical energy undergoes various conversions depending on its use. A few examples are heat energy in electric ranges, radiant energy (light) in lamps, kinetic energy in motorized appliances, and chemical energy in the recharging of batteries. After performing its desired functions, most of the energy will be dissipated as heat. *See* BATTERY; FLUORESCENT LAMP; INCANDESCENT LAMP; MOTOR.

Potential misconceptions. Some common ways of discussing matters related to energy could be misleading.

Energy utilities are often described as "producing" energy. In fact, they are only converting one form of energy into another.

We speak of a glass vase as "gaining potential energy" when raised from tile floor to tabletop. However, this energy does not reside in the vase as the words imply; rather, it resides in the new spatial relationship between the vase and the Earth's gravitational field. That the gained energy is real becomes evident if the vase should fall and shatter on the floor.

Similarly, an object's kinetic energy is an attribute not of the object alone but of its mass and the reference frame used. Thus, a speeding car has no kinetic energy with respect to the driver, but relative to a pedestrian it has a deadly amount.

Energy and cosmology. A few minutes after the big bang, the universe was uniformly filled with baryonic matter, mostly hydrogen and helium. Unlike a closed system of gas, however, this configuration did not represent maximum entropy and therefore stability—gravitational attraction among the atoms caused them to clump into stars. Those that exploded as supernovae generated the full range of the natural elements that made planets and life possible. The thermonuclear energy radiating from our Sun is therefore a direct result of gravity, and the Sun's radiated energy powers the entropy-lowering processes of life. *See* BIG BANG THEORY; SUPERNOVA.

Observations of star and galaxy dynamics imply that less than 5% of the mass of the universe is visible baryonic matter. The other 95% is hypothesized to be composed of invisible dark mass-energy, whose properties, other than its gravitational pull, are unknown. Since dark matter has mass, it represents a huge part of the universe's total energy budget. Suffusing the universe, too, is dark energy.

Dark energy may be at least partly due to the cosmological constant in Einstein's field equation for gravity that accounts for the expansion of the universe. Such energy could account for the recent evidence that the expansion of the universe is accelerating. Pressure itself, if large enough, can contribute to gravitational force, as occurs when stars collapse into black holes. The dark energy, however, can generate negative pressure, that is, tension, that would account for the acceleration. *See* ACCELERATING UNIVERSE; COSMOLOGICAL CONSTANT; COSMOLOGY; DARK ENERGY; DARK MATTER.

Search for the "true" nature of energy. Although energy flows and transformations can be accurately measured and calculated, the true nature of energy—its essence, or what it is in the ontological sense—is not known. To define energy in terms of work, as is often done, does not help since work itself is defined in terms of energy transferred.

Understanding of energy's true nature will likely come from a successor to the standard model that will successfully reconcile quantum mechanics with general relativity. In doing so, it will likely have to be valid for all stages of the evolution of the universe, including the very moment of the big bang itself at "time zero" where current theory breaks down.

Interestingly, the two main lines of research to solve this conundrum are founded on the conflicting world views of Isaac Newton and Gottfried Wilhelm Leibniz, archrival founders of classical mechanics. String theory posits that the superstrings, of which subatomic particles are composed, exist in a fixed frame of time and space (Newton's thesis); in contrast, loop quantum gravity is based on background independence, a concept that does away with absolute space and time (Leibniz's thesis); it also yields the radical implication that spacetime itself must be quantized. *See* SUPERSTRING THEORY.

Energy, in the form of fire and the natural forces of nature, baffled and amazed prehistoric peoples. It is remarkable that thousands of years later energy is still a key part of the mystery of existence.

Andrej Tenne-Sens

Bibliography. R. P. Feynman, R. B. Leighton, and M. L. Sands, *The Feynman Lectures on Physics: The Definitive and Extended Edition*, vol. 1, Addison-Wesley, 2005; B. Greene, *The Fabric of the Cosmos*, Vintage Books, 2005; J. A. Peacock, *Cosmological Physics*, Cambridge University Press, 1999; W. K. Purves et al., *Life: The Science of Biology*, 7th ed., Sinauer Associates and W. H. Freeman, 2003; L. Smolin, *Three Roads to Quantum Gravity*, Basic Books, 2001; T. W. G. Solomons and C. B. Fryhle, *Organic Chemistry*, 8th ed., Wiley, 2003; E. F. Taylor and J. A. Wheeler, *Spacetime Physics: Introduction to Special Relativity*, 2d ed., W. H. Freeman, 1992; J. R. Taylor, C. D. Zafiratos, and M. A. Dubson, *Modern Physics for Scientists and Engineers*, 2d ed., Prentice Hall, 2004; F. Wilczek, Whence the Force of F = ma? I: Culture Shock, *Phys. Today*, 57(10):11–12, October 2004; F. Wilczek, Whence the Force of F = ma? II: Rationalizations, *Phys. Today*, 57(12):10–11, December 2004 .

Energy conversion

The process of changing energy from one form to another. There are many conversion processes that appear as routine phenomena in nature, such as the evaporation of water by solar energy or the storage of solar energy in fossil fuels. In the world of technology the term is more generally applied to operations in which the energy is made more usable, for instance, the burning of coal in power plants to convert chemical energy into electricity, the burning of gasoline in automobile engines to convert chemical energy into propulsive energy of a moving vehicle, or the burning of a propellant for ion rockets and plasma jets to provide thrust. *See* ENERGY.

There are well-established principles in science which define the conditions and limits under which energy conversions can be effected, for example, the law of the conservation of energy, the second law of thermodynamics, the Bernoulli principle, and the Gibbs free-energy relation. Recognizable forms of energy which allow varying degrees of conversion include chemical, atomic, electrical, mechanical, light, potential, pressure, kinetic, and heat energy. In some conversion operations the transformation of energy from one form to another, more desirable form may approach 100% efficiency, whereas with others even a "perfect" device or system may have a theoretical

limiting efficiency far below 100%. *See* BERNOULLI'S THEOREM; CONSERVATION OF ENERGY; THERMODYNAMIC PRINCIPLES.

The conventional electric generator, where solid metallic conductors are rotated in a magnetic field, actually converts 95–99% of the mechanical energy input to the rotor shaft into electric energy at the generator terminals. On the other hand, an automobile engine might operate at its best point with only 20% efficiency, and even if it could be made perfect, might not exceed 60% for the ideal thermal cycle. Wherever there is a cycle which involves heat phases, the limitation of the Carnot criterion precludes 100% conversion efficiency, and for customary temperature conditions the ideal thermal efficiency frequently cannot exceed 50 or 60%. *See* CARNOT CYCLE.

In the prevalent method of producing electric energy in steam power plants, there are many energy-conversion steps between the raw energy of fuel and the electricity delivered from the plant, for example, chemical energy of fuel to heat energy of combustion; heat energy so released to heat energy of steam; heat energy of steam to kinetic energy of steam jets; jet energy to kinetic energy of rotor; and mechanical energy of rotor to electric energy at generator terminals. This is a typical elaborate and burdensome series of conversion processes. Many efforts have been made over the years to eliminate some or many of these steps for objectives such as improved efficiency, reduced weight, less bulk, lower maintenance, greater reliability, longer life, and lower costs. For a discussion of major technological energy converters *see* POWER PLANT; ELECTRIC POWER GENERATION.

Efforts to eliminate some of these steps have been stimulated by needs of astronautics and of satellite and missile technology and need for new and superseding devices for conventional stationary and transportation services. Space and missile systems require more compact, efficient, self-contained power systems which can utilize energy sources such as solar and nuclear. With conventional services the emphasis is on reducing weight, space, and atmospheric contamination, on improving efficiency, and on lowering costs. The predominant objective of energy conversion systems is to take raw energy from sources such as fossil fuels, nuclear fuels, solar energy, wind, waves, tides, and terrestrial heat and convert it into electric energy. The scientific categories which are recognized within this specification are electromagnetism, electrochemistry (fuel cells), thermoelectricity, thermionics, magnetohydrodynamics, electrostatics, piezoelectricity, photoelectricity, magnetostriction, ferroelectricity, atmospheric electricity, terrestrial currents, and contact potential. The electromagnetism principle today dominates the field. Electric batteries are an accepted form of electrochemical device of small capacity. *See* BATTERY; ELECTRIC ROTATING MACHINERY; ENERGY SOURCES; FUEL CELL; MAGNETOHYDRODYNAMIC POWER GENERATOR; SOLAR CELL; THERMIONIC POWER GENERATOR; THERMOELECTRIC POWER GENERATOR.

Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; H. A. Sorenson, *Energy Conversion Systems*, 1983; K. Weston, *Energy Conversion Engineering*, 1992.

Energy level (quantum mechanics)

One of the allowed values of the internal energy of an isolated physical system. This energy is not free to vary continuously above its minimum value, as predicted by classical mechanics, but is constrained to lie among a set or spectrum of particular values. This spectrum may consist of both an isolated discrete portion and a continuous component of restricted range. The term energy level usually refers to one of the allowed values in the discrete set.

Hydrogen spectrum. The primary indication for the existence of discrete energy levels came from the study of the spectrum of emissions of energetically excited atomic systems. Historically, the most important such spectrum is that of the simplest atom, hydrogen, a system of one proton and one electron bound together by their electromagnetic attraction. Within the framework of classical physics, the structure of the hydrogen atom poses fundamental problems. The first is the existence of a stable ground state: An electron in orbit around a proton is in constant acceleration, and therefore, according to Maxwell's classical electromagnetic theory, should continuously radiate away energy. Furthermore, the radiation emitted as the atom decays to a lower energy state should form a continuous spectrum of frequencies. However, the hydrogen atom both possesses a stable ground state and emits radiation at only a discrete set of frequencies. These frequencies, f, can be described, to a good approximation, by the formula (1) found by J. Balmer in 1885, where n_k and

$$f = \frac{R}{h}\left(\frac{1}{n_j^2} - \frac{1}{n_k^2}\right) \qquad (1)$$

n_j are both positive integers. Here, R is Rydberg's constant, and h is Planck's constant. *See* ELECTROMAGNETIC RADIATION.

Bohr model. In 1913 N. Bohr made a fundamental advance by postulating that the angular momentum of the electron-proton system could take on only a discrete set of values given by Eq. (2), where $\hbar$ is

$$L = (n-1)\hbar \qquad n = 1, 2, 3, \ldots \qquad (2)$$

Planck's constant divided by 2π. The angular momentum is said to be quantized. A consequence is that the excitation energies of the hydrogen atom also have a discrete spectrum, given by Eq. (3)

$$E_n = -\frac{R}{n^2} \qquad (3)$$

[see **illus.**]. In addition, Rydberg's constant could be calculated in terms of more fundamental

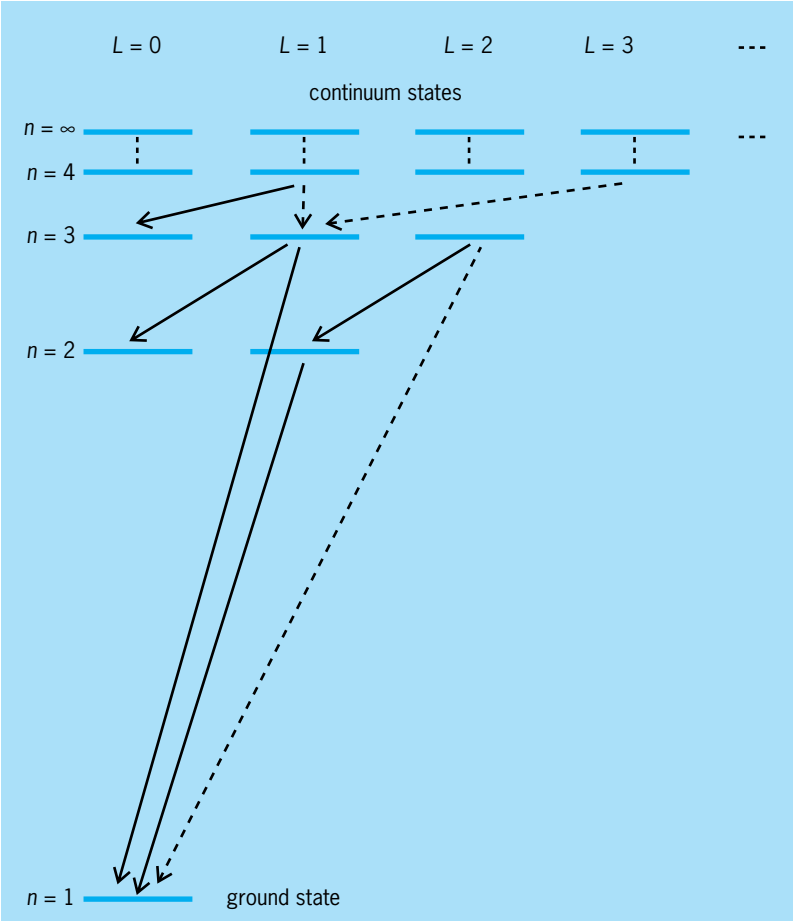

Spectrum of hydrogen energy levels in the leading order of approximation. L = angular momentum. Some typical allowed transitions are indicated by solid arrows. Transitions that are less probable because they violate approximate selection rules are indicated by broken arrows. Above the zero of energy, there exists a continuum of unbound states for all angular momenta.

constants, $R = me^4/(2\hbar^3)$, where e and m are the electron charge and mass respectively. The integer $n = 1, 2, \ldots$ now indexes the infinite number of discrete levels. They all have negative energy, relative to an infinitely separated electron-proton pair, corresponding to the bound nature of these states. The states with positive energy form a continuum and represent unbound states that can separate off to infinity. This is a generic feature of energy levels in quantum mechanics. *See* RYDBERG CONSTANT.

Bohr made the further postulate that the atom decays from an excited level, E_k, only by making a transition to a lower energy level, E_j, emitting a single light quantum (photon) in the process. The energy, E_γ, of this photon is given by the conservation of energy, $E_\gamma = E_k - E_j$. Then, by the Planck-Einstein relation, $E_\gamma = hf$, between the energy of the photon and the frequency, f, of the radiation, together with the expression (3) for the energies of the levels, the Rydberg formula (1) is derived. Although Bohr's postulates are in many ways without real foundation, they were later justified and extended by the development of quantum mechanics. *See* ATOMIC STRUCTURE AND SPECTRA.

Uncertainty relation. Crudely, the stability of the ground state of atomic systems in quantum mechanics is a consequence of the Heisenberg position-momentum uncertainty relation $\Delta x \Delta p \geq \hbar$, which implies that as the electron more closely approaches the proton (and thus is localized to within a smaller uncertainty Δx) the potential energy is reduced but its momentum, p, and thus its kinetic energy, increases. The competition between these two effects leads to a state of lowest energy. *See* UNCERTAINTY PRINCIPLE.

Complications. Other quantities label the excited states of hydrogen besides the integer n appearing in Eq. (3). At the simplest level of approximation, the n−th energy level comprises n independent but degenerate states of total angular momentum $L = 0, \hbar, \ldots, (n-1)\hbar$, each of which can have a component, m, of angular momentum along a given axis belonging to the set $m = L, L - \hbar, \ldots, 0, -\hbar, \ldots, -L$. Usually such a degeneracy is due to a symmetry of the system, possibly approximate. In the case of hydrogen, the exact rotational invariance of the system results in the exact degeneracy of the states with different values of m (for fixed n and L), while the special $1/r$ form of the Coulomb potential results in the degeneracy with respect to L. In fact, the potential is not exactly coulombic, and the levels with different L (for a fixed n) are found to be split by a small amount. The inclusion of the intrinsic angular momentum (spin) of the proton and electron further complicates the situation. *See* ANGULAR MOMENTUM; DEGENERACY (QUANTUM MECHANICS); SPIN (QUANTUM MECHANICS); SYMMETRY LAWS (PHYSICS).

Selection rules and broadening. An important modification of the simple picture presented above occurs when the decay of excited states is described in a more rigorous way considering the interactions of the hydrogen atom with the quantized electromagnetic field. Then, the excited levels are no longer infinitesimally narrow but are broadened by a calculable amount. This can be crudely understood by applying the Heisenberg energy-time uncertainty relation, $\Delta E \Delta t \geq \hbar$, which implies the shorter the lifetime of an excited state, Δt, the greater the intrinsic uncertainty in its energy, ΔE. Further, not all transitions are found to be equally probable. Since the intrinsic angular momentum of the photon is $\hbar$, conservation of angular momentum usually requires that the L values of the initial and final states differ by $\hbar$. Less frequently, the photon can also have some orbital angular momentum and mediate a transition with $\Delta L > \hbar$. Such differences in the relative probability of transitions are encoded in so-called selection rules. *See* LINEWIDTH; QUANTUM ELECTRODYNAMICS; SELECTION RULES (PHYSICS).

Quantization. The quantization of the allowed energy values that occurs in quantum mechanics has analogs for other physical quantities as well, such as angular momentum. The basic reason why such quantization occurs for bound systems of particles in quantum mechanics but not in classical mechanics is that in quantum mechanics particles have associated wavelike attributes, specifically a wave function which encodes the dynamical state of the

particle. (This is the content of wave-particle duality.) The wave function of a bound state satisfies an equation similar in many ways to the equation describing waves on a guitar string or drumhead of finite extent. Such musical instruments produce only certain specific notes, or frequencies, for a given length of string or size of drumhead. In other words, the frequencies are quantized. Similarly, the modes of oscillation of the wave function for a quantum system of finite extent are also quantized, leading to discrete energy levels, and so forth. An unbound quantum system, however, is analogous to a string of infinite length, which can play a continuous range of notes.

Eigenvalues. Formally, in quantum mechanics the energy spectrum of a physical system is given by the values of the energy, E_n, the energy eigenvalues, for which there exist nontrivial solutions of the time-independent Schrödinger equation (4). These

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}+V(x)\right)\psi_n(x)=E_n\psi_n(x) \qquad (4)$$

solutions, labeled $\psi_n(x)$, are the energy wave functions or eigenfunctions. For simplicity, the nonrelativistic Schrödinger equation has been chosen in one spatial dimension, x, for a single particle of mass m in a background potential $V(x)$. The energy eigenfunction, $\psi_n(x)$, encodes all possible information about the n−th quantum-mechanical state. In the general case, Eq. (4) becomes Eq. (5), where $\hat{H}$ is the hamil-

$$\hat{H}\psi_n = E_n\psi_n \qquad (5)$$

tonian operator appropriate for the system under study. *See* EIGENFUNCTION; EIGENVALUE (QUANTUM MECHANICS).

Complex systems. Energy levels are of great importance for many systems other than simple atoms such as hydrogen. For instance, they determine the interactions and binding of molecules in chemistry and biochemistry, the stability or decay of nuclei, and the macroscopic properties of solids, such as the optical properties of dyes or semiconductors. The observed spectroscopy of the energy levels of a system can also elucidate the properties of a new force, just as the study of hydrogen led to the development of quantum mechanics and quantum field theory. An outstanding example is the discovery, in high-energy experiments, of the J/ψ and Υ families of particles, comprising bound states of charmed–anticharmed quarks and bottom–antibottom quarks respectively. However, the binding force is no longer electromagnetism as in the case of hydrogen, but the strong force that binds nuclei, namely quantum chromodynamics (QCD). The study of the excited states of these systems has yielded greater understanding of the properties of the extremely complicated QCD force. *See* ELEMENTARY PARTICLE; J/PSI PARTICLE; MESON; NONRELATIVISTIC QUANTUM THEORY; QUANTUM CHROMODYNAMICS; QUANTUM MECHANICS; QUARKS; UPSILON PARTICLES. John March-Russell

Bibliography. A. Bohm, *Quantum Mechanics: Foundations and Applications*, 3d ed., 1994; A. Pais, *Inward Bound*, 1986; R. Shankar, *Principles of Quantum Mechanics*, 2d ed., 1994.

Energy metabolism

Energy metabolism, or bioenergetics, is the study of energy changes that accompany biochemical reactions. Animals, plants, and bacteria require energy to sustain life. Energy sustains the work of biosynthesis of cellular and extracellular components, the transport of ions and organic chemicals against concentration gradients (osmotic work), the conduction of electrical impulses in the nervous system, and the movement of cells and the whole organism. Sunlight is the ultimate source of energy for life. Photosynthetic cells use light energy to produce chemical energy and reducing compounds, used to convert carbon dioxide into organic chemicals such as glucose. The energy from the oxidation of carbohydrates, fats, and proteins sustains the biochemical reactions required for life.

Energy content of foodstuffs. The main sources of chemical energy for most organisms are carbohydrates, fats, and protein. Energy content is expressed in calories or joules. A calorie (cal) is equivalent to the amount of energy required to increase the temperature of 1 g (0.035 oz) of water by 1°C (1.8°F). One calorie is equivalent to 4.184 joules (J). The nutritional calorie, or kilocalorie (kcal), in foodstuffs is equivalent to 1000 calories. The energy content per gram of carbohydrate is 4 kcal (16 J); protein, 4 kcal (16 J); and fat, 9 kcal (36 J). The metabolism of foodstuffs yields chemical energy and heat.

Free energy changes. Energy is defined as the ability to do work, and metabolism represents the biochemical reactions that a cell can perform to produce energy. A chemical reaction is the process by which one or more substances are converted into other substances. For example, citrate (one biochemical compound) is converted into isocitrate (another compound) in the cell. The most important thermodynamic parameter in bioenergetics is the free energy change, ΔG, occurring at constant temperature and pressure (the usual conditions for chemical reactions inside the cell). The Gibbs free energy change is defined as the free energy content of the final state minus the free energy content of the initial state, as in Eq. (1). All feasible reactions occur with a nega-

$$\Delta G = G_{\text{final}} - G_{\text{initial}} \qquad (1)$$

tive free energy change; the final state has less free energy than the initial state; that is, $\Delta G < 0$ (process is exergonic). If the free energy of the final state is more than that of the initial state, ΔG is positive and the reaction is not feasible without the input of energy; $\Delta G > 0$ (process is endergonic). When the free energy change is zero, the reaction or process is at equilibrium; $\Delta G = 0$ (process is isoergonic). *See* FREE ENERGY; THERMODYNAMIC PRINCIPLES.

Standard free energy change. For general chemical reaction (2), the free energy change is expressed by

$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{C} + \mathrm{D} \qquad (2)$$

Eq. (3), where ΔG° = standard free energy change;

$$\Delta G = \Delta G^\circ + RT \ln [C][D]/[A][B] \quad (3)$$

R = gas constant, or temperature-energy coefficient (8.314 J K^{-1} $mole^{-1}$, or 1.98 cal K^{-1} $mole^{-1}$); K = absolute temperature in Kelvin (273 + degrees Celsius); ln = natural logarithm; [C][D]/[A][B] = the product of the molar concentrations, or activities, of C and D divided by the product of the molar concentrations, or activities, of A and B.

The standard free energy change for reaction (2) is the free energy change that accompanies the conversion of one mole each of reactants A and B into products C and D under standard conditions. Most biochemical reactions of interest occur in water, and the concentrations of reactants and products under standard conditions are taken to be 1 M (1 mole/liter).

The standard free energy change is a logarithmic function of the equilibrium constant,

$$K_{eq} = [C][D]/[A][B]$$

where the concentrations of A, B, C, and D are those at equilibrium. The relationship of the standard free energy change and the equilibrium constant is given by Eq. (4). When the equilibrium constant is

$$\Delta G^\circ = -RT \ln K_{eq} \quad (4)$$

greater than 1, the standard free energy change is negative. Under standard conditions (1 M in reactants and products), the reaction will proceed toward products (C + D). The free energy of the products is less than that of the reactants; that is, ΔG is negative. The actual free energy change (ΔG) can be larger or smaller than the standard free energy change (ΔG°). Increasing the concentrations of reactants or decreasing the concentration of products produces a larger decrease in free energy (the reaction is more exergonic). Decreasing the concentration of reactants or increasing the concentration of products produces a smaller decrease in free energy (the reaction is less exergonic).

Adenosine triphosphate. The complete oxidation of one mole of glucose ($C_6H_{12}O_6$) to carbon dioxide (CO_2) and water (H_2O) is associated with the liberation of free energy, as in reaction (5). Energy is

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O + 2870 \text{ kJ (686 kcal)} \quad (5)$$

released in a stepwise fashion and is coupled to the biosynthesis of adenosine triphosphate (ATP) from adenosine diphosphate (ADP) and inorganic phosphate (P_i). The reaction of ATP with water to produce ADP and P_i results in the liberation of a large amount of energy (30 kJ, or 7 kcal per mole). Such compounds are said to be energy-rich and to possess a high-energy bond. Lipmann's law is the cornerstone of energy metabolism: ATP serves as the common currency of energy exchange in living systems (animals, plants, and bacteria). The ATP-ADP couple receives and distributes chemical energy in all living systems. Creatine phosphate is an energy-rich compound found in vertebrate muscle and brain; it is a storage form of chemical energy and can energize the regeneration of ATP from ADP. Such a reaction occurs in vigorously exercising skeletal muscle when ATP is expended to produce contraction. *See* ADENOSINE TRIPHOSPHATE (ATP).

ATP synthesis. There are three general mechanisms for ATP generation in cells: substrate-level phosphorylation, oxidative phosphorylation, and photo-phosphorylation. Substrate-level phosphorylation is mediated by non-membrane-bound enzymes, and oxidative and photophosphorylation are mediated by an intricate membrane-associated ATP synthase. The energy for substrate phosphorylation is derived from exergonic chemical reactions. In the Embden-Myerhof glycolytic pathway for the conversion of glucose into two molecules of pyruvate, glyceraldehyde 3-phosphate is oxidized by nicotinamide adenine dinucleotide (NAD^+) to form 1,3-bisphosphoglycerate, an energy-rich compound. 1,3-Bisphosphoglycerate reacts with ADP to form ATP. This two-step process is one prominent example of substrate-level phosphorylation.

The energy for oxidative phosphorylation [oxidation of NADH (reduced NAD) or of succinate coupled to the phosphorylation of ADP to produce ATP] is derived from a proton gradient. Exergonic electron transport reactions generate a proton gradient (outside positive) across the inner mitochondrial membrane or the bacterial plasma membrane. Next, protons move exergonically down their electrochemical gradient from the outer to the inner aspect of the membrane. This process provides the energy to drive the endergonic portion of the phosphorylation reaction (6).

$$ADP + P_i \rightarrow ATP + H_2O \quad (6)$$

$$\Delta G^{\circ\prime} = +30 \text{ kJ/mole or } +7 \text{ kcal/mole}$$

Experiments with animal mitochondria show that the P:O ratio (number of ATPs formed from ADP + P_i per atom of oxygen consumed) using NADH as substrate is 2.5, and the P:O ratio with succinate as substrate is 1.5. There are three sites along the electron transport chain that are involved in proton translocation from the mitochondrial matrix to the outside of the inner mitochondrial membrane. Sites 1 and 3 are associated with the generation of a proton gradient sufficient for the formation of 1 ATP, and site 2 with 0.5 ATP. Transport of electrons from succinate dehydrogenase circumvents site 1 and results in the production of only 1.5 ATP molecules.

To recapitulate, there are four properties of oxidative phosphorylation. (1) Transport of electrons from reductant to oxygen along the respiratory chain is a very exergonic process. (2) Part of the chemical energy is conserved by the translocation of protons from the inside to the outside of the inner mitochondrial membrane or bacterial plasma membrane and the ensuing proton gradient. (3) The membrane is not freely permeable to protons. (4) Protons then move down their exergonic

electrochemical gradient and drive endergonic ATP formation in a process involving the ATP synthase of the inner mitochondrial membrane or plasma membrane (bacteria). Four protons need to be transported for the synthesis of each ATP in mitochondria: one for the translocation of phosphate and three for ATP synthase. *See* BACTERIAL PHYSIOLOGY AND METABOLISM.

In plant photophosphorylation, light energy fuels oxidation-reduction reactions that promote proton translocation in the chloroplast. Protons move down their electrochemical gradient to drive ATP biosynthesis by a chloroplast ATP synthase that is homologous to that in mitochondria and in bacteria. The generation of a proton gradient, or a proton-motive force, and its use to drive the generation of ATP biosynthesis is a cornerstone of bioenergetics, as first described by Peter Mitchell. The concept that osmotic energy derived from the proton gradient is used to drive the chemical reaction of ATP formation is called the chemiosmotic theory, or the Mitchell theory—a cornerstone of energy metabolism. *See* CYTOCHROME; ELECTRON-TRANSFER REACTION; MITOCHONDRIA; PHOTOSYNTHESIS.

Robert Roskoski, Jr.

Bibliography. D. G. Nicholls and S. J. Ferguson, *Bioenergetics 2*, 1992; R. Roskoski, Jr., *Biochemistry*, 1996; L. Stryer, *Biochemistry*, 4th ed., 1995.

Energy sources

Sources from which energy can be obtained to provide heat, light, and power. The term energy is used to describe an amount of work perfomed. There are two kinds of energy, kinetic energy, meaning work performed by the movement of matter, and potential energy, meaning work that is stored or at rest in matter. *See* ENERGY.

Energy Concepts and Needs

In the kinetic or potential state, energy takes on one of five forms: (1) Chemical energy results from changes in the chemical structure of substances, such as during combustion of fuel. (2) Electrical energy results from electrons and protons in motion in a stream called an electric current, or in temporary storage as in a battery or fuel cell. (3) Mechanical energy results from force applied or about to be applied to liquid, solid, or gaseous matter. (4) Thermal energy results from heat being applied to matter. (5) Nuclear fission is the splitting of the nucleus of an atom into two or more parts by collision with neutrons, with the consequent release of the force that binds protons and neutrons of the nucleus together. All living things on Earth depend on one or more of these forms of energy and must look to a wide variety of energy sources. *See* BATTERY; CHEMICAL ENERGY; FUEL CELL.

The Sun. The ultimate source of energy on Earth is the Sun, which produces energy sources in two ways on a daily basis and a stored-over-time basis (**Figs. 1** and **2**).

The Sun provides radiant heat on a daily basis to the Earth which drives many reactions. For example, solar heat evaporates water from the sea and the lakes, providing the moisture for cloud formations which break into rain in mountainous regions. The rain runs off into constructed reservoir lakes which feed water through hydroelectric dams, thus rotating electric generators that produce the kinetic electrical energy supplied through power lines to industrial plants, commercial buildings, and residential homes. Tidal power, wave power, solar power, and wind power are additional examples of daily energy sources.

On a stored-over-time basis, radiant heat from the Sun striking the Earth over millions of years provided the necessary energy input to convert vegetable matter into coal, petroleum, and natural gas. This phenomenon is still taking place, but the conversion process is so slow that it is meaningless in terms of its replacement capability, especially when compared to the enormous rate at which modern industrial society is consuming these resources.

Capital energy. Scientists refer to the stored-over-time energy sources collectively as capital energy, and this is subdivided into six categories: primary

Nonrenewable stored-over-time (capital)	Fossil:	coal peat crude petroleum natural gas	Combustion process
	Nuclear:	uranium thorium deuterium lithium beryllium	Noncombustion process
Renewable daily	Solar:	solar thermal conversion photoelectric energy conversion } direct photochemical conversion stored solar heat with heat pumps } indirect	
	Hydro:	river–reservoir energy conversion	
	Tidal:	tidal energy conversion	
	Wind:	windmill energy conversion	
	Oceans:	ocean heat conversion ocean current conversion wave energy conversion	
	Geothermal:	natural steam hot water hot dry rocks	
	Biomass:	wood and other vegetation	Combustion process

Fig. 1. Primary energy sources.

Nonrenewable	Electric:	electric power generation fuel cells	Noncombustion process
	Nuclear:	tritium plutonium	
	Fossil: (coal-derived)	coke char tar blast furnace gas water gas and carbureted water gas producer gas town gas briquetting coal slurries coal gasification coal methanol	Combustion process
	Fossil: (petroleum-derived)	gasoline kerosine petroleum coke oil shale petroleum fuel oils (No. 1, 2, 4, 5, and 6) liquefied natural gas (LNG) liquified petroleum gas (LPG) propane butane recycled lubricants and solvents	
Renewable	Biomass:	wood waste and bark bagasse hulls (grain, rice, cottonseed) peanut shells coffee grounds sugarbeets tobacco stems citrus rinds corncobs garbage and trash methane gas alcohol (ethanol, methanol)	

Fig. 2. Secondary energy sources.

energy, secondary energy, renewable energy, nonrenewable energy, combustion process, and noncombustion process.

Primary. This classification includes all forms of potential energy created mainly by the Sun in the Earth's crust that need no processing or treatment to transform them into usable energy (Fig. 1).

Secondary. This classification includes the forms of potential energy manufactured from primary energy forms generally by mechanical, chemical, thermal, or nuclear reaction means to transform them into usable energy.

Renewable. This term refers to forms of potential energy that constantly and rapidly renew themselves for steady, reliable use. This is a somewhat ambiguous definition, particularly when the word "rapidly" is added. Coal, crude petroleum, and natural gas are listed as nonrenewable in Fig. 1, but this is not totally accurate. These forms of potential energy are, in fact, constantly being created. However, modern industrialized societies are utilizing these energy sources so rapidly compared to the geological time period required for the formation of additional sources as to render the concept of renewable useless within any reasonable time frame. On the other hand, such energy forms as solar and wind power are clearly renewed on a timely basis and are labeled as renewable energy.

Nonrenewable. Any form of potential energy that does not fall within the definition accepted for renewable energy is considered nonrenewable. For example, the fossil fuels may be defined in specific considerations as nonrenewable energy sources.

Combustion process. Many of the forms of potential energy shown in Figs. 1 and 2 must be utilized in a combustion process before they will release their stored energy into a work process. This is part of the process of conversion that will be discussed below.

Noncombustion process. There are also ways to release energy without a combustion process, such as river water turning a waterwheel. This was an important source of energy for nineteenth-century factories.

Energy needs. Human use of energy, particularly capital energy, has accelerated over time due to an increase in the human population and the discovery of new technologies for utilizing energy. This acceleration took on exponential proportions during the Industrial Revolution of the nineteenth and twentieth centuries (see **table**). As the utilization of, and therefore the demand for, energy rose, scientists and engineers discovered ways to utilize new forms of energy. The availability and technical feasibility of many energy forms yielded a proliferation of energy choices, so that decisions had to be reached on the selection of energy sources. The need to make energy decisions has resulted in the development of an energy choice system (**Fig. 3**). This system involves an economic analysis which compares the cost of input versus the benefit of output for two specific energy sources. Among the factors included in this analysis are depletion cost, the cost of using nonrenewable energy; social cost, the health and environmental issues; availability cost, the cost of utilizing energy sources subject to interruptions of supply; and switching or conversion cost, converting plant and equipment to other energy uses.

Technical Aspects

Certain technical aspects must be considered in order that energy users can make intelligent decisions regarding energy sources.

Fossil fuels. Crude petroleum, natural gas, and coal were formed in the Earth's crust over the course of millions of years and exist today in subsurface locations. For instance, coal usually exists in seams ranging from 3 to 6 ft (0.9 to 1.8 m), although one seam in Wyoming averages 100 ft (30 m) thick with a maximum thickness of 220 ft (67 m). Crude petroleum is found trapped in the pores of rock (sandstone, limestone, or dolomite) or sand overlain with some kind of impervious cap rock that prevents the liquid from dispersing. Natural gas is normally found trapped in the Earth alongside or with crude petroleum.

The specific ways in which fossil fuels were formed is not known. The theory most dominant since the 1920s, and the one accepted by the United States oil, gas, and coal industry, is that they were

Energy consumption in the United States*

Item	1982	2000
Key energy use factors		
U.S. population, $\times 10^6$	232	268
Dwellings, $\times 10^6$	83	107
Passenger cars, $\times 10^6$	128	148
Other cars, $\times 10^6$	11	14
Buses and trucks, $\times 10^6$	20	25
Gross national product (1972$), $\times 10^9$	1476	2426
Fuel consumption		
Petroleum (total), 10^3 bbl (m^3)/day	15,254 (2425)	12,940 (2057)
Gasoline, 10^3 bbl (m^3)/day	6538 (1039)	4890 (778)
Jet fuel, 10^3 bbl (m^3)/day	1009 (160)	1400 (220)
Aviation gasoline, 10^3 bbl (m^3)/day	26 (4.1)	10 (2)
Diesel, 10^3 bbl (m^3)/day	1298 (206)	1800 (290)
Other distillates, 10^3 bbl (m^3)/day	1506 (239)	500 (80)
Petrochemical chemical feed, 10^3 bbl (m^3)/day	507 (80.6)	940 (150)
LNG/LPG, 10^3 bbl (m^3) /day	1537 (244)	1000 (200)
Residual fuels, 10^3 bbl (m^3)/day	1694 (269)	1640 (260)
Still gas, 10^3 bbl (m^3)/day	554 (88.1)	360 (57)
Asphalts/road oils, 10^3 bbl (m^3)/day	343 (54.5)	200 (30)
Lubricating waxes, 10^3 bbl (m^3)/day	154 (24.5)	200 (30)
Petroleum coke, 10^3 bbl (m^3)/day	247 (39.3)	100 (20)
Miscellaneous, 10^3 bbl (m^3)/day	91 (14)	—
Fuel alcohol, 10^3 bbl (m^3)/day	3 (0.5)	300 (50)
Natural gas, 10^3 ft^3 (m^3)/day	17.9 (0.5)	17.9 (0.5)
Coal (not including exports), 10^6 tons (metric tons)	678 (615)	1160 (1053)
Solar, $\times 10^6$ units	0.1	3.0
Hydroelectricity, 10^9 kWh (megajoules)	310 (1122)	340 (1224)
Nuclear, 10^9 kWh (megajoules)	283 (1019)	589 (2120)
Electricity from wind-waste, 10^9 kWh (megajoules)	2 (7)	18 (65)
Geothermal, 10^9 kWh (megajoules)	4 (14)	52 (187)
Electricity from geothermal, quad (joules)	—	1 (1.055×10^{18}
TOTALS, quad (joules)	70.9 (74.8×10^{18})	84.7 (89.4×10^{18})

*Based on U.S. Department of Commerce statistics, July 1983.

formed as a result of the fossilization and carbonization of trees, ferns, and other vegetable matter under intense pressure and temperature over exceedingly long periods of time in the Earth's crust.

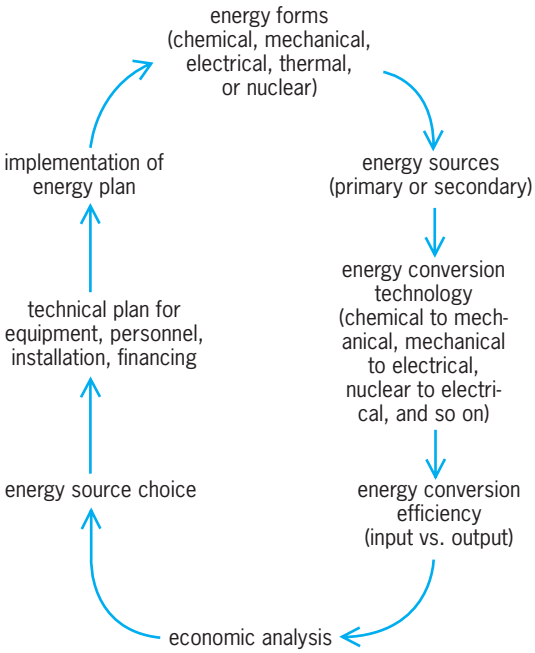

Fig. 3. Diagram of an energy choice system.

Another theory about the formation of fossil fuels is based on the deep-earth gas hypothesis formulated by T. Gold of Cornell University. According to Gold, incredibly huge quantities of methane gas (CH_4), the principal constituent of natural gas, may be trapped deep inside Earth. Tapping into this reserve would require wells drilled several miles below Earth's surface. According to this model, the known deposits of crude oil, natural gas, and coal were formed as a result of the mixing of the deep methane gas with deposits of vegetable and animal remains as the methane rose up through the earth.

Coal. This complex mineral substance is located principally in 29 of the 50 states of the United States and is actively and economically mined in 12 states in meaningful commercial quantities.

The principal chemical constituents of coal are water, carbon, hydrogen, sulfur, nitrogen, oxygen, and ash (noncombustible mineral residue). However, coal is not a uniform substance. It is almost infinitely variable in its composition from one location to the next, even within the same mine location. Nevertheless, 31 tests for defining coal have been established by the American Society for Testing and Materials. Different coals can be ranked according to the degree to which they have progressed from lignite through the bituminous stage to anthracite. This progression is roughly equivalent to the geologic time of development, although the other variables of depth-pressure,

heat, and vegetable matter constituents play an important part.

Coal is used in the United States for generating electric power, metallurgical production, general industrial processes, residential-commercial uses, and synthetic fuels. For the most part, coal is burned in fire-tube or water-tube boilers for raising steam that is used to generate electricity, provide heat for factories and buildings, and provide steam for production processes. *See* COAL; LIGNITE.

Peat. In the very early stages of coal formation, accumulations of decomposed and partially decomposed vegetation, trees, ferns, and mosses located in a wet, cold, and anaerobic (oxygen-deficient) environment are very likely to turn to peat at the rate of about 3 in. (7.5 cm) per 100 years. Only after hundreds or even millions of years would this material graduate to the status of coal.

Since the vegetable matter in peat, consisting of cellulose and other organic compounds, becomes only partially converted to carbon and hydrocarbons, peat has only one-third to one-half of the heating value of coals. It is used in very limited quantities for fuel after it is cut from the earth and formed into bricquettes. Peat is harvested and sold mainly as a soil conditioner. *See* PEAT.

Crude petroleum. Oil is also a complex substance derived from the carbonized remains of trees, ferns, mosses, and other types of vegetable matter. Like coal there is doubt about the exact nature of its origin. The principal chemical constituents of oil are carbon, hydrogen, and sulfur. Crude petroleum extracted from the earth will burn and produce thermal energy, but virtually all crude oil is processed in refineries, where it is converted into several useful fuels and special products (for example, feedstock for chemicals, plastics, food products, medicines, and tires, plus tar, asphalts, and lubricating oils). The various fuels made from crude oil are jet fuel, gasoline, kerosine, diesel fuel (or No. 2 fuel oil), and heavy fuel oils (or No. 4, 5, and 6 fuel oils).

Major oil comsumption in the United States is in the following areas: transportation, residential-commercial, industry boilers and other industrial uses, and generating electric power. *See* PETROLEUM; PETROLEUM PRODUCTS.

Natural gas. This energy source is 83–93% methane (CH_4), so that its principal chemical constituents are carbon and hydrogen. Natural gas is usually found in the immediate vicinity of crude petroleum, although some natural gas wells do not yield oil.

Of all the chemical or mineral sources of energy, natural gas may well be the most desirable because it can be pipelined directly to the customer, requires no storage vessels, is clean-burning, requires no air-pollution control equipment, produces no ash for disposal, and mixes with air easily to provide complete combustion at low excess air.

The principal uses for natural gas are residential, commercial buildings, industrial, transportation, and generating electric power. *See* NATURAL GAS.

Nuclear energy. The two ways to utilize radioactive fuel as an energy source are fission and fusion. In fission, heavy atoms are split into two principal elements that form the nucleus of two new, smaller atoms. In fusion, the nuclei of two small atoms fuse together into a single, larger nucleus. In both cases large quantities of energy are released.

Nuclear fission. The splitting of atoms cannot be effectively accomplished with electrically charged matter such as alpha particles, beta particles, or protons, which tend to be diverted or slowed down as they approach other charged matter. Neutrons, however, are not deflected by positive or negative electrical charges, and this fact makes them ideal candidates for atom smashing. The fission process, therefore, involves the bombardment of atoms with neutrons so that a sufficient number of collisions will take place on a statistically predictable basis and split those atoms into two or more separate nuclei while releasing massive quantities of thermal energy. *See* NUCLEAR FISSION.

Nuclear fusion. The fusion process is the opposite of fission. Instead of splitting atoms into two or more pieces, the fusion process causes two atoms to collide with such force that their natural electronic repulsion is overcome and their nuclei fuse into one.

The most suitable fuels for fusion are hydrogen isotopes (hydrogen atoms with the same number of protons but different numbers of neutrons). In particular, deuterium and tritium are thought to be ideal for fusion reactions. The former is relatively abundant and can be found in ordinary seawater. The latter is scarce. *See* DEUTERIUM; TRITIUM.

Although this energy source offers tremendous potential for the twenty-first century, there are formidable problems that must be overcome before it can be utilized. The temperature in a fusion reactor should be between 1.0 and 2.5×10^8 K, and no solid material has been developed for a container for such a reaction. A solution to this problem may be found by setting up an extremely powerful magnetic field around the reacting materials. However, conventional magnets may consume more energy than the fusion reaction puts out. This could be solved by using superconducting magnets that operate at very low temperatures to decrease their resistance to electrical flow, except that it will be difficult to operate such magnets adjacent to the heat of a fusion reactor.

At any rate, fusion power is still an important potential source of continuous low-cost energy, provided that the technical problems can be solved. *See* NUCLEAR FUSION; SUPERCONDUCTING DEVICES.

Solar energy. By far the most attractive energy source is the Sun itself because it is free, is clean and nonpolluting, and does not involve the use of dwindling, finite reserves of capital energy. The problem with solar energy is its cost, particularly for large industrial uses. Roughly 50% of the sunlight approaching Earth is reflected or absorbed by the atmosphere and the other 50% strikes Earth's surface. In the tropical and temperate zones, the Sun provides the equivalent of, on average, about 622 MW of energy per square mile (240 MW/km^2). Unfortunately, technology is able to convert only about 15% of the Sun's

radiant heat to usable work, so that a typical large electric generating station capable of producing 1000 MW would require an array of solar head collector-converters covering an area of at least 11 mi^2 (28 km^2). *See* SOLAR ENERGY.

The collector-converters are made of silicon, which produces electricity when it is exposed to light. The high cost of these devices derives from the cost of growing and cutting large crystals into thin slices of silicon. The more this process becomes automated, the lower the price and the more the process is likely to be employed. *See* PHOTOVOLTAIC CELL; SOLAR CELL.

Another idea is the solar power tower. This type of technology was demonstrated by an Italian scientist, G. Francia, in 1976. A 10-MW electric generating station is in operation in Daggett, California; it utilizes a 72-acre (29-hectare) field of mirrors to concentrate sunlight at the top of a central tower. The thermal energy from the sunlight superheats steam that is used to power electricity-generating turbines.

Finally, the concept of solar ponds has received attention in many parts of the world. These are artificial, salt-gradient ponds. They are derived from the work of a Russian scientist, A. Von Kaleczinsky, who in 1902 found that the temperature of the water in Lake Medve in Transylvania a few feet below the surface was 185°F (85°C) due to the variation in salinity of the water. Experiments have been conducted with ponds in which layers of different salinity allow the water to trap and hold solar heat. The hot brine can be used for space heating or for producing electricity with a thermoelectric device or organic Rankine cycle engine.

Hydro energy. One of the oldest energy-producing mechanisms uses water flowing in a river or falling from a height to rotate work devices, ranging from the waterwheels of the past to the massive modern hydroelectric dams that employ gigantic electricity-generating turbines. Internationally only 7% of the total potential hydroelectric power estimated at 2.9×10^6 MW is being used. Much of that potential, unfortunately, is located in remote places from which lengthy transmission lines would be prohibitively expensive. Furthermore, environmental protectionists are resisting new projects that might threaten natural resources. *See* ELECTRIC POWER GENERATION; WATERPOWER.

Tidal energy. The only major tidal energy project in operation is the Rance River project in Brittany, France; however, estimates indicate a worldwide potential for producing 3×10^6 MW of electric power from tidal movement. The French project employs a barrage type of dam across the estuary of a river. Turbines located in this barrage pump water into the estuary when the tide is rising. When a sufficient head of water is built up in the estuary, the water is permitted to flow back through the turbines to produce electricity. *See* TIDAL POWER.

Wind energy. This is also a form of energy that can be used to generate electrical or mechanical energy. Rotating devices known as windmills can convert the mechanical energy of wind to useful work.

The largest known wind generator has a 300-ft (90-m) span to produce 2.5 MW of electricity for a local electric power grid in Goldendale, Washington. This installation employs huge aircraft propeller-type blades which must always be facing into the wind; a more efficient design might be the Darrieus vertical-axis windmill that rotates regardless of the direction of the wind. Such a device was erected at Sandia Laboratories, Albuquerque, New Mexico, with funds from the U.S. Energy Research and Development Administration. It measures 61 ft (19 m) in height and produces 60 kW of electric power in a 28 mi/h (45 km/h) wind.

In other parts of the United States (such as California and Hawaii), the concept of wind farms is being developed that involves the erection and operation of perhaps dozens of windmills for the production of significant amounts of electric power. *See* WIND POWER.

Ocean energy. Ocean power may become an energy source in the future, either as wave power or as ocean temperature differential. Both forms are highly experimental and may require many years to become significant energy sources.

Wave power. Many different devices have been proposed and tested for exploiting the wave motion of the sea. In Scandinavia, one device is designed so that ocean waves cause massive quantities of water to flow into the confines of the device which then generates electricity as the water tries to escape back to the sea. In England, S. Salter demonstrated his so-called nodding duck device, which is a floating device that rotates due to the motion of waves rolling over it. The float is capable of driving a hyraulic pump that, in turn, can drive an electrical generating device.

Ocean temperature. An energy-producing system known as ocean thermal energy conversion (OTEC) is based on the temperature differential in the oceans near the Equator where the surface water is about 40°F (20°C) warmer than the water a few thousand feet down. This temperature difference can be utilized to vaporize a working fluid (such as ammonia) that can be run through an electricity-producing turbine.

Geothermal energy. As with ocean power, the geothermal energy source may be exploited in either or both of two ways: hot rocks and hydrothermal.

Hot rocks. It is known that hot granite rock (up to 400°F or 200°C) exists almost everywhere on Earth. The heat is generated, for the most part, by a slow radioactive decay process deep within the earth. According to one estimate, a 40-mi^3 (170-km^3) chunk of granite at 350°F (177°C) would yield the equivalent energy output of 1.2×10^{10} barrels (2×10^9 m^3) of oil, which is approximately the total yearly energy consumption in the United States. According to another estimate, the continental United States is underlain by hot rock with thermal energy amounting to 1.3×10^7 quad (1 quad = 10^{15} Btu) at a depth of just under 6 mi (10 km).

The technique (still experimental) for tapping this energy source involves conventional oil or gas

drilling expertise. The first hole drilled is used for injecting water pressurized to 5000 lb/in.^{2} (35 megapascals) that will cause the hot granite to fracture vertically. Once the rock is fractured, the amount of pressure is decreased to a normal pumping pressure. The next hole drilled and any subsequent holes must follow the fracture line so that the water heated by the hot granite can be brought back up to the surface and utilized for raising steam and generating electricity for space heating or for industrial process steam.

Hydrothermal. The earth emits steam in many locations, such as Iceland, the United States, the Philippines, and New Zealand, where it is captured and employed for space heating, generating electricity, or industrial process steam applications. Hydrothermal steam now provides the equivalent of about 1200 MW of electric power. It is believed that only about 1% of the total potential hydrothermal energy can be utilized in the world and converted to electricity at a 25% efficiency factor for a total contribution of only 3×10^6 MW. *See* GEOTHERMAL POWER.

Biomass energy. As applied to the field of energy, the term biomass energy encompasses a broad selection of energy sources: Any and all types of living matter that can be converted to a form of energy can be said to be biomass. Hence, scientists tend to think of such items as wood, wood waste, coffee grounds, corn husks, peanut shells, rice hulls, garbage, animal and human waste, sugarcane waste (bagasse), and organic effluent from streams and ponds as biomass.

The biomass considered to be most significant in terms of energy sources is wood and wood waste. There are approximately 9.6×10^9 acres (3.9×10^9 hectares) of forest land in the world, of which about 4×10^9 acres (1.6×10^9 hectares) are economically accessible. Up until the end of the nineteenth century, when coal took over as the leading energy source, wood was a preeminent provider of energy to the world. Wood is again being used on a limited scale at the residential level. The lumber, furniture, plywood, and pulp-and-paper industries all utilize wood waste items (for example, bark, shavings, sawdust, slabs, and end pieces) for raising steam that, in turn, is employed for space heating and industrial processes. *See* BIOMASS.

Conservation

An energy source that is being increasingly considered to be significant is energy conservation. Better insulation of buildings and homes could slash heating and air-conditioning requirements in half. More people could travel by public transportation than by using private vehicles. There could be greater use of smaller, fuel-efficient automobiles, bicycles, or low-gas-consumption "mopeds," motorcycles, and such. Greater use could be made of fluorescent lighting in homes and industry. Government agencies, businesses, and individuals could lower their thermostats in winter and raise them in summer. In fact, this movement was already quietly under way in the late 1970s until its effects were dramatically felt in the early 1980s in terms of mass decreases in total energy consumption. Conservation, coupled with the worldwide economic recession of 1982, resulted in the so-called oil glut in the mid 1980s. Therefore, conservation may acquire a major role in establishing energy source security. *See* CONSERVATION OF RESOURCES.

For further information on primary energy sources *see* BERYLLIUM; LITHIUM; NUCLEAR FUELS; OIL AND GAS, OFFSHORE; PETROLEUM RESERVES; THORIUM; URANIUM.

For further information on secondary energy sources *see* ALCOHOL FUEL; COAL GASIFICATION; COAL LIQUEFACTION; COKE; GASOLINE; KEROSINE; LIQUEFIED NATURAL GAS (LNG); LIQUEFIED PETROLEUM GAS (LPG); OIL SAND; OIL SHALE; PLUTONIUM.

William K. Fox

Bibliography. B. Aldridge, L. Crow, and R. Aiuto, *Energy Sources and Natural Fuels*, 1993; J. H. Harker and J. R. Backhurst, *Fuel and Energy*, 1981; J. T. McMullen, R. Morgan, and R. B. Murray, *Energy Resources and Supply*, 1976; D. Marier and L. Stoiaken (eds.), *Alternative Sources of Energy*, 1988; T. Ohta, *Energy Technology: Sources, Systems, and Frontier Conversion*, 1994; H. B. Smith, *Energy: Sources, Applications, Alternatives*, 1993.

Energy storage

The general method and specific techniques for storing energy derived from some primary source in a form convenient for use at a later time when a specific energy demand is to be met, often in a different location.

In the past, energy storage on a large scale had been limited to storage of fuels. For example, large amounts of natural gas may be stored under pressure in underground reservoirs during the summer and used to meet increased demands for heating fuel in the colder season. Petroleum and its products are stored at several points in the energy system, from the strategic petroleum reserve to the fuel tanks of automobiles. Since gasoline is a highly concentrated and readily portable form of energy, this method of energy storage makes the automobile independent of the supply system for appreciable distances and times. On a smaller scale, electric energy is stored in batteries that power automobile starters and a great variety of portable appliances. *See* OIL AND GAS STORAGE.

Energy storage in many forms is expected to play an increasingly important role in shifting patterns of energy consumption away from scarce to more abundant and renewable primary resources. For example, automobiles may be able to store transportation energy in the form of coal-derived synthetic fuels or as electricity in batteries charged with electric power from coal or nuclear power plants. Solar energy can already be made more usable by accumulating it during the day as warm water that can be stored for later use. Another example is the storage of electric energy generated at night by coal or nuclear power plants to meet peak electric loads during daytime periods. This storage is achieved by pumping water

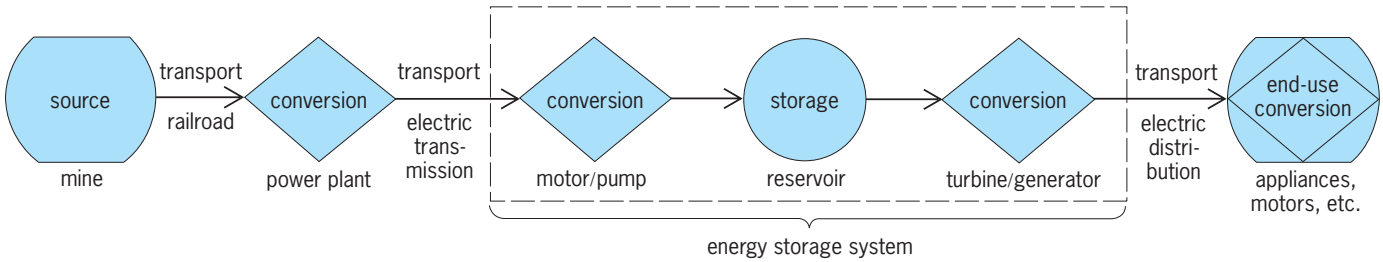

Fig. 1. Principal components of an energy storage system with specific reference to an electric power system.

from a lower to a higher reservoir at night and reversing this process during the day, the pump then being used as a turbine and the motor as a generator. As shown in **Fig. 1**, this example can be used to illustrate the conversion and storage functions of an energy storage system.

Broader application and use of new methods of energy storage could reduce oil consumption in the major energy-use sectors. Probably the largest practical potential for oil displacement is in the heating and cooling of buildings through storage of heat or coolness generated on site with solar energy or off-peak electricity (singly or in combination). In principle, the potential for oil displacement is largest in transportation, since highway vehicles consume large quantities of fuel. In practice, extensive displacement of conventional automobiles will be very difficult because of the very large energy storage capacity of fossil fuels compared to the most attractive alternative, electric storage batteries: a 20-gal (76-liter) tank of gasoline will give the average car a range of 300–600 mi (500–1000 km), while a lead-acid battery may provide a range of just 30–60 mi (50–100 km), weighs nearly a ton (0.9 metric ton), and costs between $1200 and $1500. The contrast between the two systems is illustrated in **Fig. 2**, which compares their specific ranges. Yet, because of their independence of oil and their more efficient as well as cleaner use of energy, electric vehicles powered by batteries are a possible alternative to conventional automobiles, especially for urban driving.

Energy storage in electric power systems has excellent potential for reducing the use of oil and gas for power generation, especially if pumped hydro storage can be supplemented by more broadly applicable techniques such as underground pumped hydro, compressed-air storage, and batteries. These techniques also could provide utilities with special operating advantages not found in conventional generating plants. Storage technologies that could find application in one or more of the major energy-use sectors are discussed below.

Pumped hydro storage. Until the late 1970s the only bulk energy storage method used by electric utilities, pumped hydro goes back to 1929 in the United States; the largest such facility in the world—with a capacity of 1.5×10^7 kWh—is the Ludington pumped storage plant on Lake Michigan. The construction of such facilities is limited by a variety of geographic, geologic, and environmental constraints, all of which may be considerably reduced if the power generation equipment and discharge reservoir are placed deep underground (**Fig. 3**). The size of the surface reservoir required for such an underground pumped hydroelectric system will be several times smaller than that used in the conventional system because of the greater pressure available to run the generators. The difference in height between reservoirs in the underground system may be several thousand feet (1 ft = 0.3 m), compared to several hundred feet common in the conventional pumped hydroelectric system.

Compressed-air storage. Off-peak electric energy can also be converted into mechanical energy by pumping air into a suitable cavern where it is stored at pressures up to 80 atm (8 MPa). Compared to pumped hydro storage, compressed air offers several advantages, including a wider choice of suitable geologic formations, lower volume requirements, and a smaller minimum capacity to be economically attractive. The world's first commercial compressed-air energy storage system (**Fig. 4**) is in Huntorf, Germany. In this installation, expanding air must be heated by the combustion of natural gas before it can be used to drive the generator and turbines. In a more efficient concept, the heat evolved during air compression would be stored in a bed of pebbles and reused to heat the expanding air.

Batteries. The development of advanced batteries with characteristics superior to those of the

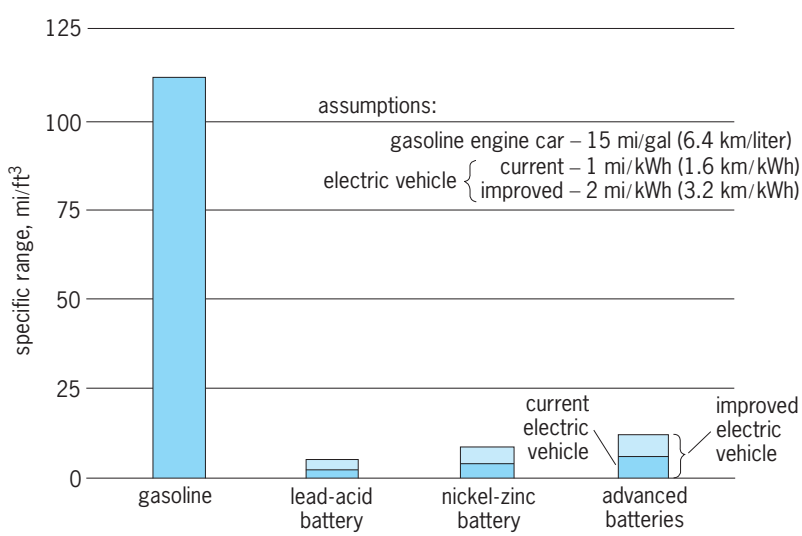

Fig. 2. Specific range of vehicles using energy stored in various forms. 1 mi/ft^3 = 57 km/m^3.

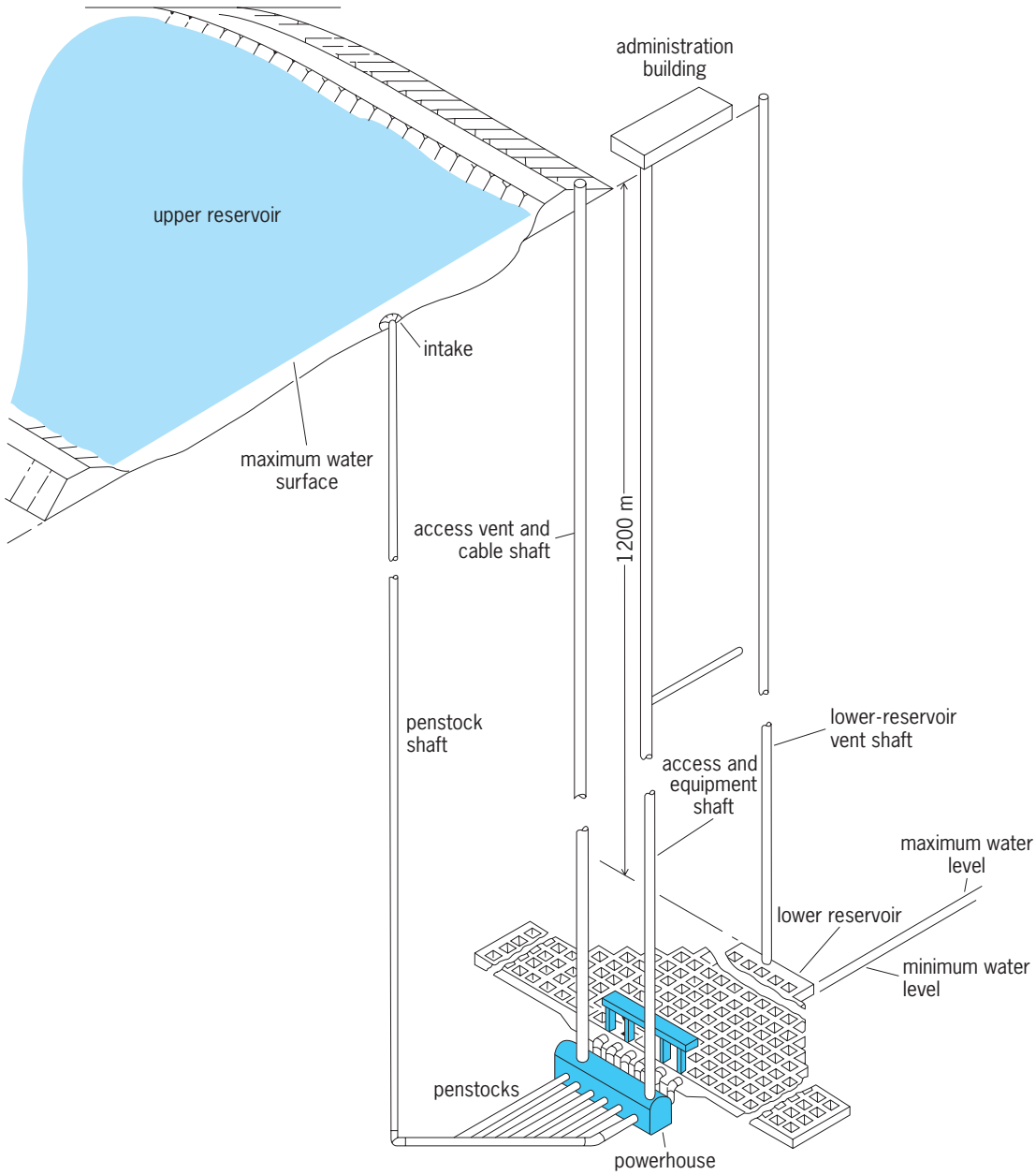

Fig. 3. Concept of an underground pumped hydroelectric storage system.

familiar lead-acid battery could result in use of battery energy storage on a large scale. For example, batteries lasting 2000 or more cycles could be used in installations with capacities of several hundred thousand kilowatt-hours in various locations on the electric power grid, as an almost universally applicable method of utility energy storage. Batteries combining these characteristics with energy densities (storage capacity per unit weight and volume) well above those of lead-acid batteries could provide electric vehicles with greater range, thus removing a major barrier to their broader use.

The development of several battery types has been undertaken to achieve systems with superior characteristics. Candidates include nickel-zinc and nickel-iron systems that could provide modest improvements over lead-acid. The zinc-chlorine, sodium-sulfur, and lithium-iron sulfide systems could yield substantial improvements over lead-acid.

The zinc-chlorine battery (**Fig. 5**) operates near ambient temperature with an aqueous solution of zinc chloride as the electrolyte. On charge, zinc is deposited on a graphite substrate and chlorine is released as a slightly soluble gas at the opposite electrode, composed of porous graphite. Chlorine is stored in a separate vessel by freezing it out of the aqueous solution as a solid compound of water (ice) and chlorine at about 48°F (9°C); in this form, chlorine is relatively safe to handle. On discharge, zinc redissolves and chlorine is reduced at the porous graphite cathode, reforming the original zinc chloride aqueous solution. The active materials and other components offer the promise of a low-cost battery; however, the complexity of the system (which uses

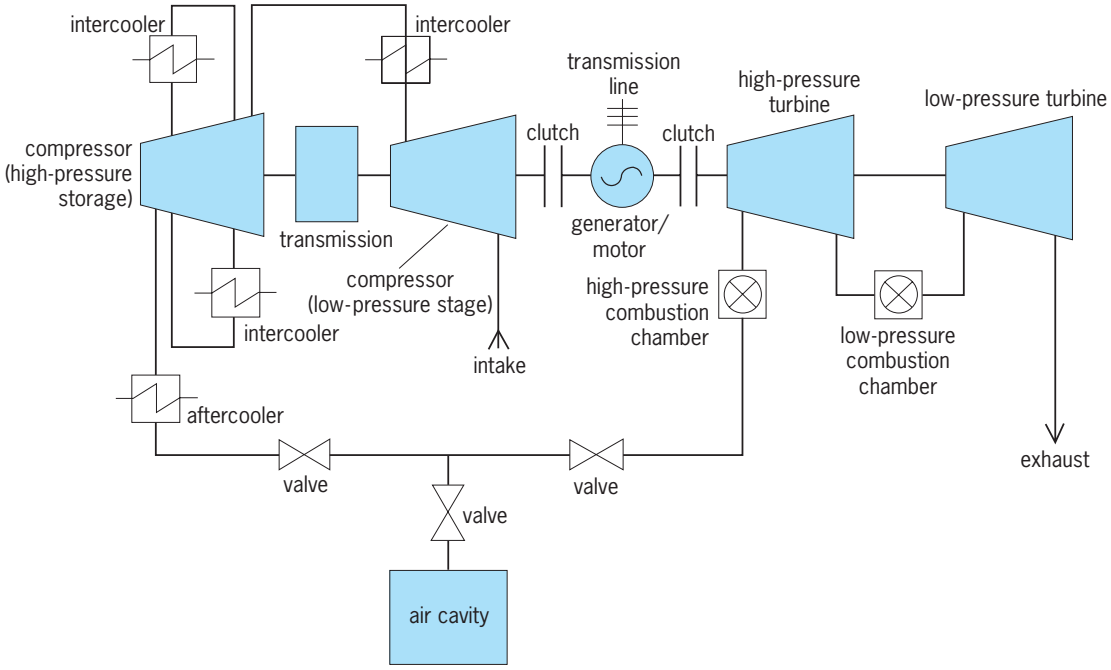

Fig. 4. Schematic diagram of compressed-air storage plant at Huntorf, Germany.

pumps, valves, and other auxiliaries) suggests that operating reliability could be a possible problem. *See* SOLID-STATE BATTERY.

The sodium-sulfur battery operates at temperatures above 570°F (300°C), where its electrodes exist as liquids separated by a unique solid ceramic electrolyte. Small experimental cells have been tested through 1000 charging cycles while showing no appreciable degradation. A disadvantage of the sodium-sulfur battery, however, is that it must operate at a temperature of 570–660°F (300–350°C). The operating costs and safety characteristics of large batteries consisting of thousands of single cells in electric series and parallel connection are also potential drawbacks. *See* BATTERY.

Thermal storage. Ceramic brick "storage heaters" that store off-peak electricity in the form of heat are used for heating buildings in Europe; the barriers to

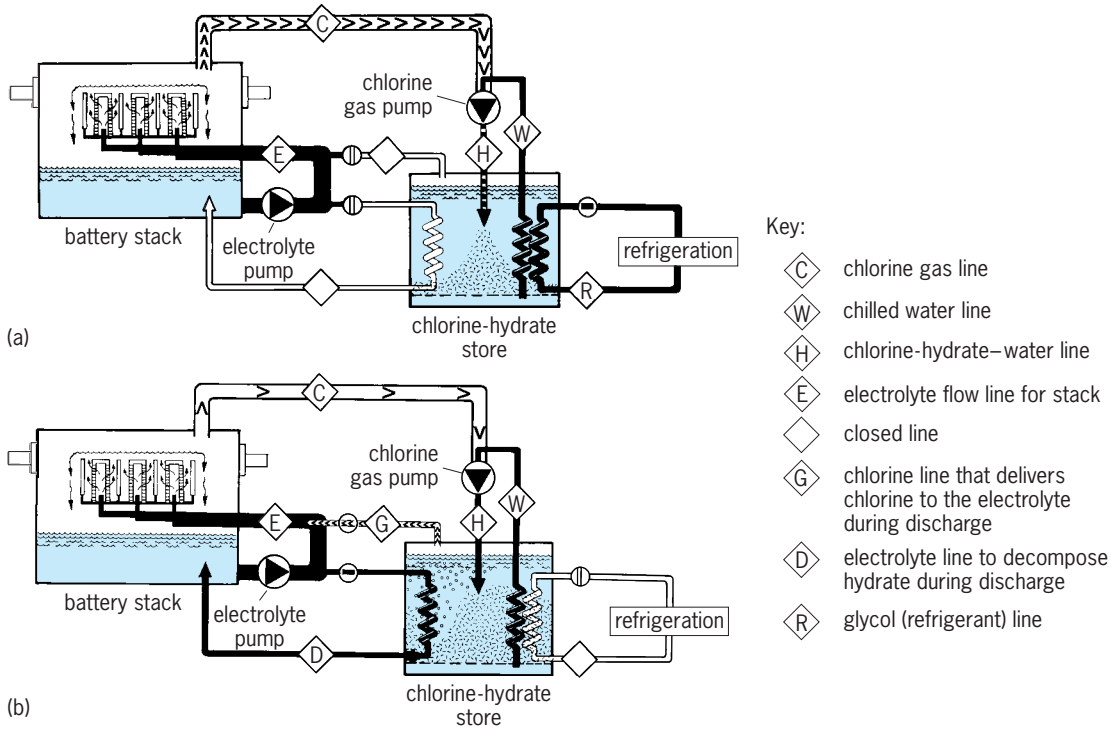

Fig. 5. Schematic diagram of zinc-chloride battery, showing two main compartments in the design, the battery stack and the chlorine hydrate store. (*a*) Charge cycle. (*b*) Discharge cycle. (*Gulf and Western Co.*)

their use in the United States are more institutional and economic than technological.

Testing has been done on prototype "coolness" storage systems, which use electric refrigeration to chill water or produce ice at night. Experimental installations indicate that daytime electric power demand for air conditioning could be reduced up to 75% by using such systems, but these systems are still relatively bulky and expensive.

Solar hot-water storage is technically simple and commercially available. However, the use of solar energy for space heating requires relatively large storage systems, with water or rock beds as storage media, and difficulties can arise in integrating this storage with existing buildings while keeping costs within acceptable limits. Innovative designs that may help overcome these problems include new types of heat transfer equipment, storage of heat in the walls and ceilings of buildings (passive solar heating), and the use of low-cost, roof-mounted water bags. *See* SOLAR ENERGY; SOLAR HEATING AND COOLING.

Chemical reaction systems. Heat or electricity may be stored by using these energy forms to force certain chemical reactions to occur. Such reactions are chosen so that they can be reversed readily with release of energy; in some cases the products can be transported from the point of generation to that of consumption, adding flexibility to the ways the stored energy can be used. For example, reactions which produce hydrogen could become attractive since hydrogen could be stored for extensive periods of time and then conveniently used in either combustion devices or in fuel cells.

Another possibility is to apply heat to a mixture of methane and water, converting them to hydrogen and carbon monoxide, which can be stored and transported to the end-use site, where a catalyst permits the reverse reaction to occur spontaneously with the release of heat. While not yet in practical use, chemical reaction systems are under extensive investigation as economic and flexible ways of storing energy. *See* FUEL CELL.

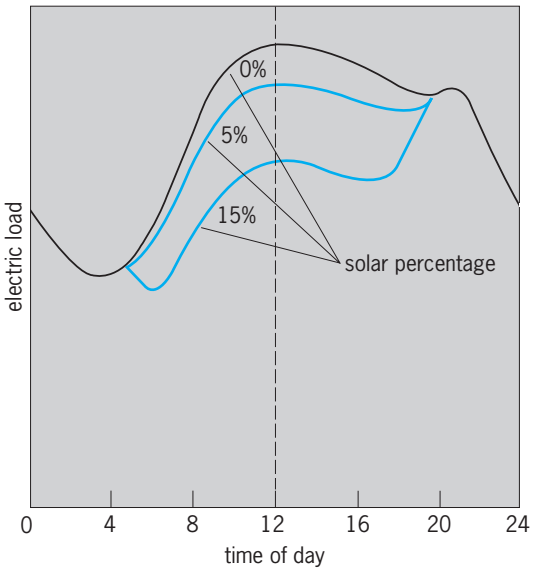

Fig. 6. Impact of solar-electric generation on utility load curve during a typical 24-h period.

Superconducting magnets. Electrical energy can be stored directly in the form of large direct currents used to create fields surrounding the superconducting windings of electromagnets. In principle such devices appear attractive because their storage efficiency is high, plant life could be long, and utilities would have few difficulties establishing the necessary conversion equipment. However, the need for maintaining the system at temperatures approaching absolute zero and, particularly, the need to physically restrain the coils of the magnet when energized require auxiliary equipment (insulation, vacuum vessels, and structural supports) which will represent a large expenditure. *See* SUPERCONDUCTING DEVICES.

Flywheels. Storage of kinetic energy in rotating mechanical systems such as flywheels is attractive where very rapid absorption and release of the stored energy is critical. However, even advanced designs and materials are likely to be too expensive for utility energy storage on a significant scale, and applications will probably remain limited to systems where high power capacity and short charging cycles are the prime consideration. Such applications do exist in pulse power supplies and in electric transportation for recovery of braking energy. *See* FLYWHEEL.

Combined systems. The rising importance of energy storage comes from its potential for shifting demand from scarce to plentiful primary energy sources. As such, the most successful storage devices are likely to be those that are adopted as components of larger systems designed specifically for resource conservation. In electric power generation, for example, system-wide storage may be combined with solar-electric systems to flatten load curves and reduce oil consumption (**Fig. 6**). If such complex systems are eventually to be developed and commercially introduced, financial incentives and more coherent national energy policy planning must be adopted. Thus, while advanced energy storage systems could have a major impact on United States energy consumption patterns with continued research, nontechnical factors such as regulatory strategies and pricing policies are likely to become decisive to their large-scale use. *See* ENERGY SOURCES.

Fritz Kalhammer; Thomas R. Schneider

Bibliography. G. Beckmann and P. V. Gilli, *Thermal Energy Storage*, 1984; F. Dinter, M. A. Geyer, and R. Tamme, *Thermal Energy Storage for Commercial Applications*, 1993; J. Jensen *Fundamentals of Energy Storage*, 1985; F. R. Kalhammer, Energy-storage systems, *Sci. Amer.*, 241(6):56–65, December 1979; S. Karkainnen and Y. El-Maghary (eds.), *Energy Storage Systems in Developing Countries*, 1988; D. Linden (ed.), *Handbook of Batteries,* 2d ed., 1995; B. Kilkis and S. Kakac (eds.), *Energy Storage Systems*, 1989; M. N. Wilson, *Superconducting Magnets*, 1987; F. De Winter, *Solar Collectors, Energy Storage, and Materials*, 1991.

Engine

A machine designed for the conversion of energy into useful mechanical motion. The principal characteristic of an engine is its capacity to deliver appreciable mechanical power, as contrasted to a mechanism such as a clock, whose significant output is motion. By usage an engine is usually a machine that burns or otherwise consumes a fuel, as differentiated from an electric machine that produces mechanical power without altering the composition of matter. Similarly, a spring-driven mechanism is said to be powered by a spring motor; a flywheel acts as an inertia motor. By definition a hydraulic turbine is not an engine, although it competes with the engine as a prime source of mechanical power. *See* ENERGY CONVERSION; HYDRAULIC TURBINE; MOTOR; PRIME MOVER; WATERPOWER.

Applications. A fuel-burning engine may be stationary, as a donkey engine used to lift cargo between wharf and ship, or it may be mobile, like the engine in an aircraft or automobile. Such an engine may be used for both fixed service and mobile operation, although accessory modifications that adapt the engine to its particular purpose are preferable. For example, the fan that draws air through the radiator of a water-cooled fixed engine is large and fitted in a baffle, whereas the fan of a similar but mobile engine can be small and unbaffled because considerable air is driven through the radiator by means of ram action as the engine propels itself along. *See* AIRCRAFT ENGINE; AUTOMOTIVE ENGINE; MARINE ENGINE.

Some types of engine can be designed for economic efficiencies in fixed service but not in mobile operation. Thus, the steam engine is widely used in central electric generator stations but is obsolete in mobile service. This is chiefly because, in a large ground installation, the furnace and boiler can be fitted with means for using most of the available heat. The engine proper can be a reciprocating (piston) or a rotating (turbine) type. Because shaft rotation is by far the most used form of mechanical motion, the turbine is the more common form of modern steam engine. For railroad service the steam engine has given place to diesel and gasoline internal combustion engines and to electric motors. *See* BOILER; LOCOMOTIVE; POWER PLANT; STEAM ENGINE; STEAM-GENERATING UNIT; STEAM TURBINE.

Types. Traditionally, engines are classed as external or internal combustion. External combustion engines consume their fuel or other energy source in a separate furnace or reactor. *See* NUCLEAR REACTOR; STEAM-GENERATING FURNACE.

Strictly, the furnace or reactor releases chemical or nuclear energy into thermal energy, and the engine proper converts the heat into mechanical work. The principal means for the conversion of heat to work is a gas or vapor, termed the working fluid. By extension, an engine which derives its heat energy from the Sun by solar radiation to working fluid in a boiler can be considered an external combustion type. To avoid loss of or contamination from nuclear fuel, the reactor and boiler are separated from (and may also be shielded from) the engine. *See* SHIP NUCLEAR PROPULSION.

The working fluid takes on energy in the form of heat in the boiler and gives up energy in the engine, the engine proper being a thermodynamic device. The device may be a turbine for stationary power generation or a nozzle for long-range vehicular propulsion. *See* NUCLEAR POWER.

In an engine used for propulsion, the rearward velocity with which the working fluid is ejected and, thus, the forward acceleration imparted to the vehicle depend on the temperature of the fluid. For practical purposes, temperature is limited by the engine materials that serve to contain the chemical combustion or nuclear reaction. To achieve higher exhaust velocity, the working fluid may be contained by nonmaterial means such as electric and magnetic fields, in which case the fluid must be electrically conductive. *See* ION PROPULSION.

The engine proper is then a magnetohydrodynamic device receiving electric energy from a separate fuel-consuming source such as a gas turbine and electric generator or a nuclear reactor and electric generator or possibly from direct electric conversion of nuclear or solar radiation. *See* INTERPLANETARY PROPULSION; SPACECRAFT PROPULSION; THERMOELECTRICITY.

A further basis of classification concerns the working fluid. If the working fluid is recirculated, the engine operates on a closed cycle. If the working fluid is discharged after one pass through boiler and engine, the engine operates on an open cycle. Closed-cycle operation assures the purity of the working fluid and avoids the discharge of harmful wastes. The open cycle is simpler. Thus the commonest types of engine use atmospheric air in open cycles both as the principal constituent of their working fluids and as oxidizer for their fuels.

If open-cycle operation is used, the next modification is to heat the working fluid directly by burning fuel in the fluid; the engine becomes its own furnace. Because this internal combustion type engine uses the products of combustion as part of the working fluid, the fuel must be capable of combustion under the operating conditions in the engine and must produce a noncorrosive and nonerosive working fluid. Such engines are the common reciprocating gasoline and diesel units. *See* DIESEL ENGINE; INTERNAL COMBUSTION ENGINE; ROTARY ENGINE; STIRLING ENGINE.

At low speeds the combustion process is carried out intermittently in a cylinder to drive a reciprocating piston. At high speed, however, friction between piston and cylinder walls and between other moving parts dissipates an appreciable portion of the developed power. Thus, where high power is developed at high speed, performance is improved by continuous combustion to drive a turbine wheel. *See* BRAYTON CYCLE; CARNOT CYCLE; DIESEL CYCLE; GAS TURBINE; OTTO CYCLE; TURBINE PROPULSION.

Engine shaft rotation may be used in the same way as in a reciprocating engine. However, for high-velocity vehicular propulsion, the energy of the

working fluid may be converted into thrust more directly by expulsion through a nozzle. Once the vehicle is in motion, the turbine can be omitted. Alternatively, instead of drawing atmospheric oxygen into the combustion chamber, the engine may draw both oxidizer and fuel from storage tanks within the vehicle, or the combustion chamber may contain the full supply of fuel and oxidizer. *See* JET PROPULSION; RAMJET; ROCKET PROPULSION; TURBOJET; TURBOPROP; TURBORAMJET.

Despite all the variation in structure, mode of operation, and working fluid—whether of moving parts, moving fields, or only moving working fluid—these machines are basically means for converting heat energy to mechanical energy. *See* THERMODYNAMIC PROCESSES. Frank H. Rockett; D. L. Anglin

Bibliography. W. H. Crouse and D. L. Anglin, *Automotive Engines*, 8th ed., 1994; L. Guzzella and C. H. Onder, *Introduction to Modeling and Control of Internal Combustion Engine Systems*, 2004; R. Stone, *Introduction to Internal Combustion Engines*, 3d ed., 1999.

Engine cooling

A cooling system in an internal combustion engine that is used to maintain the various engine components at temperatures conducive to long life and proper functioning. Gas temperatures in the cylinders may reach 4500°F (2500°C). This is well above the melting point of the engine parts in contact with the gases; therefore it is necessary to control the temperature of the parts, or they will become too weak to carry the stresses resulting from gas pressure. The lubricating oil film on the cylinder wall can fail because of chemical changes at wall temperatures above about 400°F (200°C). Complete loss of power may take place if some spot in the combustion space becomes sufficiently heated to ignite the charge prematurely on the compression stroke. *See* INTERNAL COMBUSTION ENGINE.

Fortunately, a thin protective boundary of relatively stagnant gas of poor heat conductivity exists on the inner surfaces of the combustion space. If the outer cylinder surface is placed in contact with a cool fluid such as air or water and there is sufficient contact area to cause a rapid heat flow, the resulting drop in temperature produced by the heat flow in the inside boundary layer keeps the temperature of the cylinder wall much closer to the temperature of the coolant than to the temperature of the combustion gas. The quantity of heat that crosses the stagnant boundary layer and must be carried away by the coolant is a function of the Reynolds number of the gas existing in the cylinder. In terms of practical engine quantities, the heat flow to the coolant varies approximately as in the relationship: (charge density × piston speed)$^{0.8}$. At full throttle and normal piston speed, this heat flow amounts to about 15% of the energy of the incoming fuel. *See* REYNOLDS NUMBER.

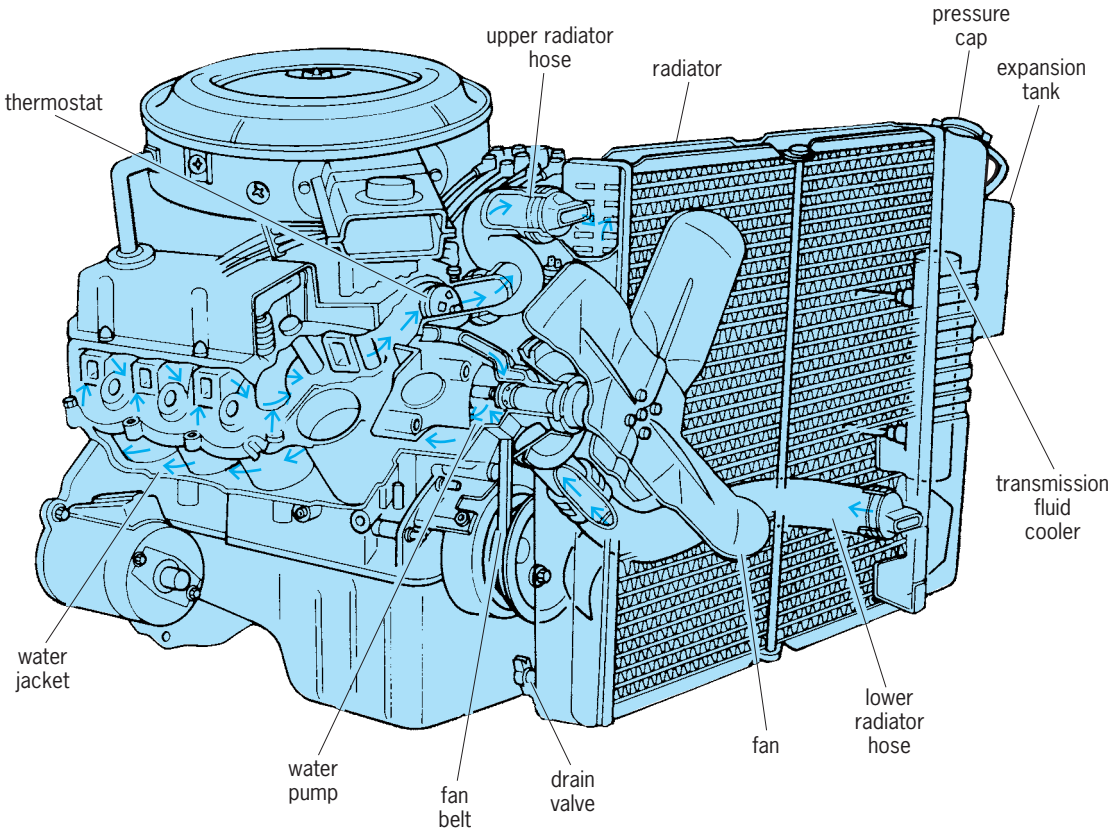

Fig. 1. Cooling system of a V-8 automotive spark-ignition engine. The arrows show the direction of coolant flow through the engine water jackets and cooling system. (*Ford Motor Co.*)

Liquid cooling. If the coolant is water, it is usually circulated by a pump through jackets surrounding the cylinders and cylinder heads. The water is circulated fast enough to remove steam bubbles that may form over local hot spots and to limit the water's temperature rise through the engine to about 15°F (8°C). In most engines in automotive and industrial service, the warmed coolant is piped to an air-cooled heat exchanger called a radiator (**Fig. 1**). The airflow required to remove the heat from the radiator is supplied by an electric or engine-driven fan; in automotive applications the airflow is also supplied by the forward motion of the vehicle. The engine and radiator may be separated and each placed in the optimum location, being connected through piping. To prevent freezing, the water coolant is usually mixed with ethylene glycol. *See* ANTIFREEZE MIXTURE; ETHYLENE GLYCOL; HEAT EXCHANGER.

Low water-jacket temperature contributes to corrosive wear of engine parts and increases piston friction losses. High water-jacket temperature increases the coolant loss by evaporation or by actual boiling. In automotive engines, temperature of the water jacket is often automatically maintained near 195°F (91°C) by a thermostat placed in the line from the engine to the radiator. When the engine outlet water is too cool, as when the engine is first started, the water is prevented from entering the radiator, and is usually recirculated in the engine block until warm enough to open the thermostat.

Many marine diesel engines have a closed cooling system in which fresh water is circulated through the engine. Instead of flowing through a radiator, the heated water flows through a heat exchanger, or cooler, immersed in seawater so that the heat is carried away by it instead of by air. In some engines, the heated water is used in a jacket to cool the exhaust manifold.

Air cooling. Engines are often cooled directly by a stream of air without the interposition of a liquid medium. The heat-transfer coefficient between the cylinder and airstream is much less than with a liquid coolant, so that the cylinder temperatures must be much greater than the air temperature to transfer to the cooling air the heat flowing from the cylinder gases. To remedy this situation and to reduce the cylinder wall temperature, the outside area of the cylinder, which is in contact with the cooling air, is increased by finning. The heat flows easily from the cylinder metal into the base of the fins, and the great area of finned surface permits heat to be transferred to the cooling air (**Fig. 2**). The ideal fin shape depends upon the conductivity of the fin material. In general, the fin is thickest at the base to permit heat flow from the cylinder. The fin should taper to a thin edge to give a good temperature gradient along its length. For reasons of mechanical strength, fins are usually made thicker than necessary for heat-transfer considerations. *See* HEAT TRANSFER.

High-output cylinders require many closely spaced fins. In these engines the area between adjacent cylinders is blocked off with sheet-metal baffles, which are also shaped to follow the fin tips part way around the cylinder. A pressure is built up in front of the baffles by means of a fan or because of the forward motion of the vehicle (ram effect). The pressure differential between front and rear of the engine forces the cooling air through the spaces between the fins.

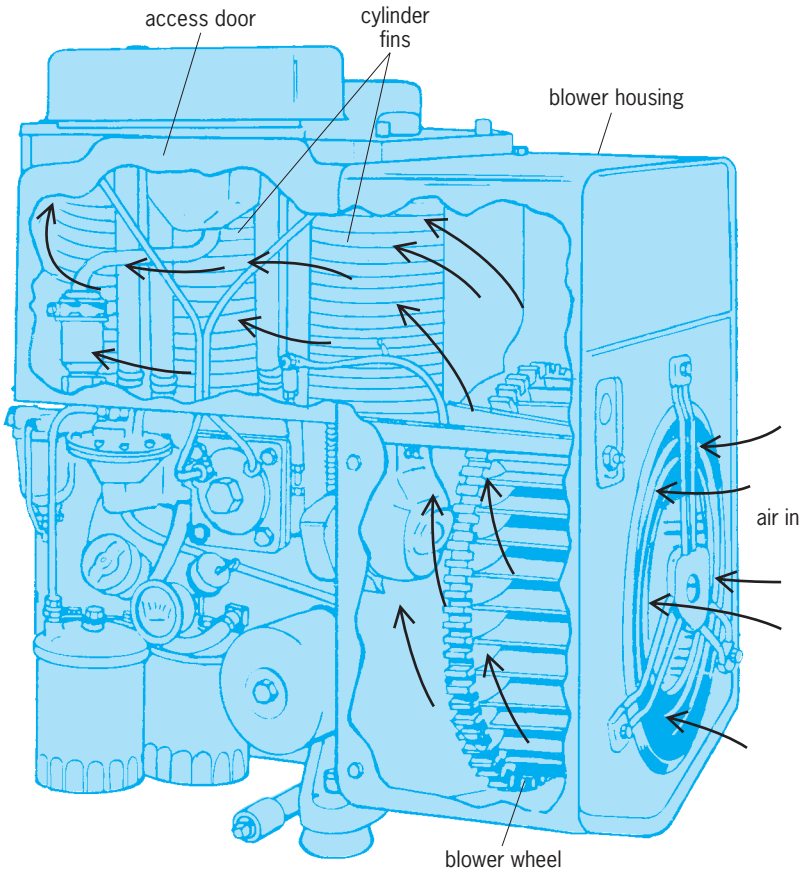

Fig. 2. Airflow around the cylinders and fins of an air-cooled engine. (*Onan Corp.*)

The power required to cool depends upon the quantity of cooling air used and the velocity at which it passes the fins. To keep the power required for cooling to a minimum, the fins should be long and close together so that a large heat-transfer area is served by a small coolant flow area. The temperature difference between fins and cooling air should be kept as high as possible so that less air velocity will be required. Cylinder temperatures of air-cooled engines are sometimes controlled by louvers or flaps, which may be set to restrict the cooling airflow until the engine becomes warm.

Adiabatic engine. This type of diesel engine has no cooling system through which heat is lost. Ceramic materials, which have higher melting points than metals, retain the heat generated in the combustion chamber and cylinder. This has the potential to make the adiabatic diesel engine a more efficient and economical power plant. *See* ADIABATIC PROCESS; CERAMICS.

Augustus R. Rogowski; Donald L. Anglin

Bibliography. W. H. Crouse and D. L. Anglin, *Automotive Fuel, Lubricating, and Cooling Systems*, 1981; Society of Automotive Engineers, *Engine Coolants, Cooling System Materials, and Components*, 1993.

Engine lubrication

In an internal combustion engine, the system for providing a continuous supply of oil between moving surfaces during engine operation. This viscous film, known as the lubricant, lubricates and cools the power transmission components while removing impurities, neutralizing chemically active products of combustion, transmitting forces, and damping vibrations. *See* INTERNAL COMBUSTION ENGINE; LUBRICANT.

Lubricant. Automotive and other four-stroke Otto-cycle engines are generally lubricated with petroleum-base base oils that contain chemical additives to improve their natural properties. Synthetic oils are used in gas turbines and may be used in other engines. *See* OTTO CYCLE.

Probably the most important property of oil is the absolute viscosity, which is a measure of the force required to move one layer of the oil film over the other. If the viscosity is too low, a protecting oil film is not formed between the parts. With high viscosity too much power is required to shear the oil film, and the flow of oil through the engine is retarded. Viscosity tends to decrease as temperature increases. Viscosity index (VI) is a number that indicates the resistance of an oil to changes in viscosity with temperature. The smaller the change in viscosity with temperature, the higher the viscosity index of the oil. *See* VISCOSITY.

Lubricating system. Small two-stroke cycle engines may require a premix of the lubricating oil with the fuel going into the engine, or the oil may be injected into the ingoing air–fuel mixture. This is known as a total-loss lubricating system because the oil is consumed during engine operation.

Most automotive engines have a pressurized or force-feed lubricating system in combination with splash and oil mist lubrication (see **illus.**). The lubricating system supplies clean oil cooled to the proper viscosity to the critical points in the engine, where the motion of the parts produces hydrodynamic oil films to separate and support the various rubbing surfaces. The oil is pumped under pressure to the bearing points, while sliding parts are lubricated by splash and oil mist. After flowing through the engine, the oil collects in the oil pan or sump, which cools the oil and acts as a reservoir while the foam settles out. Some engines have an oil cooler to remove additional heat from the oil. *See* WEAR.

Oil filter. This filter removes impurities from the oil passing through it and helps maintain the lubricating ability of the oil during the maintenance intervals. The impurities removed include solid particles such as combustion residue, metal particles, and dust.

Two types of filters have been used on internal combustion engines, full-flow and bypass. The full-flow filter protects the engine by filtering the entire oil-pump output and traps any particles during the first pass of the oil. However, a bypass valve is required to ensure continued supply of oil to the engine should the filter become clogged. Most automotive engines have pleated-paper filters of this type. Bypass filters remove about 10% or less of the oil passing through the lubricating system, and return this oil to the sump after cleaning. Some engines have both full-flow and bypass filters for the lubricating oil. Donald L. Anglin

Bibliography. *Bosch Automotive Handbook*, 1986; W. H. Crouse and D. L. Anglin, *Automotive Fuel, Lubricating, and Cooling Systems*, 1981; Deere and Co. Staff (eds.), *Engines*, 7th ed., 1991.

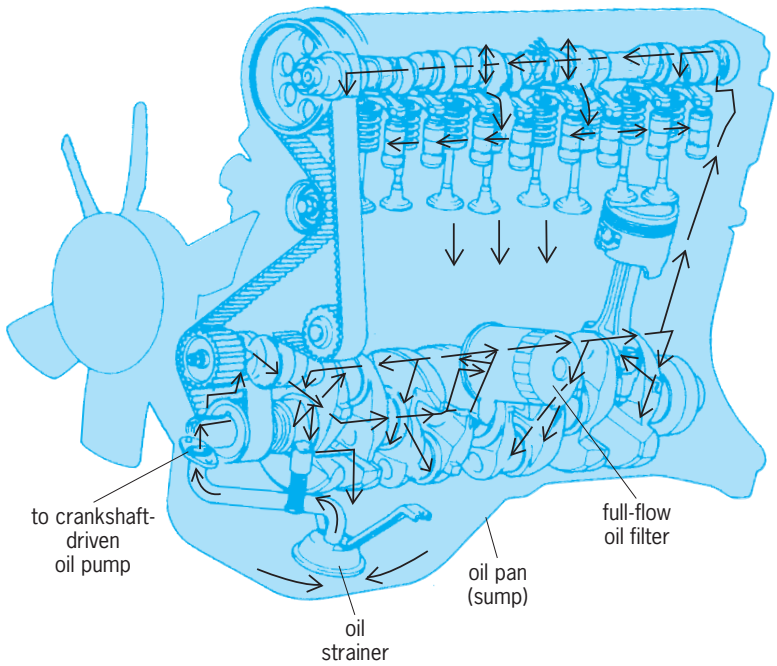

Pressurized lubricating system for an automobile engine. Arrows show the flow of oil through the engine. (*Toyota Motor Sales U.S.A. Inc.*)

Engine manifold

An arrangement or collection of pipes or tubing with several inlet or outlet passages through which a gas or liquid is gathered or distributed. The manifold may be a casting or fabricated of relatively light material. Manifolds are usually identified by the service provided, such as the intake manifold and exhaust manifold on the internal combustion engine (see **illus.**). Some types of manifolds for handling oil, water, and other fluids such as engine exhaust gas are often called headers. In the internal combustion engine, the intake and exhaust manifolds are an integral part of multicylinder engine construction and essential to its operation. *See* INTERNAL COMBUSTION ENGINE.

Intake manifold. The engine intake manifold is a casting or assembly of passages through which air or an air–fuel mixture flows from the air-intake or throttle valves to the intake valve ports in the cylinder head or cylinder block. In a spark-ignition engine with a carburetor or throttle-body fuel injection, the intake manifold carries an air–fuel mixture. In an engine with port fuel injection or in a diesel engine, the intake manifold carries only air. For the

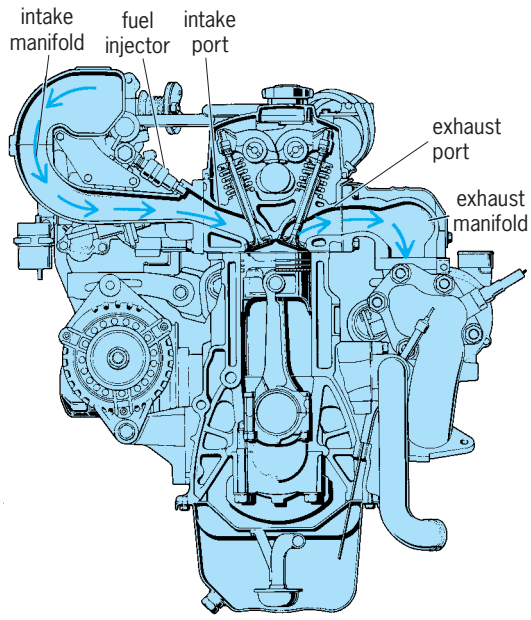

Cross-sectional view of a fuel-injected spark-ignition engine showing the intake and exhaust manifolds. Arrows show the flow of the air through the intake manifold and of the exhaust gases through the exhaust manifold. (*Ford Motor Co.*)

diesel engine, the air should be inducted with a minimum of pressure drop.

The purpose of the intake manifold is to distribute the air or air–fuel mixture uniformly to each of the cylinders and to assist in the vaporization of the fuel. The problems of manifold design are most severe when a large number of cylinders is charged through a single carburetor. Careful design is required to prevent unequal flow, localized condensation of fuel, and power losses due to overheating of the charge. The downdraft manifold is most frequently employed with automotive engines. To maximize volumetric efficiency, some engines have a tuned intake manifold, in which the port cross-sectional area and length are adjusted to a size that fills the cylinders most efficiently at a given engine speed, thereby causing engine torque to peak at that speed.

A leaking intake system may allow unfiltered air to reach the cylinders. In a spark-ignition engine, a leak may cause detonation, misfire, and driveability and exhaust-emission problems during engine operation, and evaporative hydrocarbon emissions when the engine is not operating. Leakage of air into the air–fuel mixture manifold unbalances the engine by producing lean mixtures at the leak and upsetting the calibration of the fuel-metering system. *See* AIR POLLUTION; FUEL INJECTION.

A leaking intake system on a naturally aspirated diesel engine may not be extremely critical. However, an explosive combustible vapor may be aspirated. Scavenging air manifolds or pressure-charged intakes lose valuable air, both quantitatively and in terms of pressure, if leaks occur. *See* SUPERCHARGER.

Exhaust manifold. The engine exhaust manifold is a casting or assembly of passages through which the products of combustion leave the exhaust-valve ports in the cylinder head or cylinder block and enter the exhaust piping system. The purpose of the exhaust manifold is to collect and carry these exhaust gases away from the cylinders with a minimum of back pressure. The pulsating flows in the exhaust manifold results in pressure variations that may improve scavenging or may result in increased back pressure for one or more cylinders. Since the frequency of the pressure variations for a given engine and exhaust system varies with engine speed, the exhaust manifold should be designed to operate efficiently over the expected speed range of the engine. The entire exhaust system, including the exhaust manifold, catalytic converter (in automotive engines), muffler, and piping, affects the efficiency of combustion-gas evacuation from the cylinders.

The volume of exhaust gas is greater than the volume of air handled by the intake system, and the exhaust gases are at much higher temperature than the intake. Some small pressure drop is tolerated in the exhaust system, but exhaust back pressure represents a direct loss of engine power. Stationary engines intended to operate indoors have water-jacketed exhaust manifolds which provide cooling while limiting the heat dissipated into the engine room.

Exhaust manifolds operate at high temperatures and may be subject to erosive or corrosive attack while pressures are low or moderate. Engines with turbochargers may have multiple exhaust manifolds to prevent pulsations or interferences in the gas flow to the power turbine. An engine with an exhaust-driven turbocharger will lose power and efficiency if leakage exists ahead of the power turbine. Exhaust manifolds may also be tuned to improve cylinder scavenging and reduce back pressure in the exhaust system.

Air manifold. Some automative spark-ignition engines have an air-distribution manifold as part of the exhaust-emission control system. This manifold distributes and proportions air to the individual exhaust ports through external tubing or integral passageways.

Donald L. Anglin

Bibliography. W. H. Crouse, *Automotive Engines*, 8th ed., 1994; W. H. Crouse and D. L. Anglin, *Automotive Fuel, Lubricating, and Cooling Systems* 1981; C. F. Taylor, *The Internal-Combustion Engine in Theory and Practice*, vol. 2, 1985.

Engineering

Most simply, the art of directing the great sources of power in nature for the use and the convenience of humans. In its modern form engineering involves people, money, materials, machines, and energy. It is differentiated from science because it is primarily concerned with how to direct to useful and economical ends the natural phenomena which scientists discover and formulate into acceptable theories. Engineering therefore requires above all the creative imagination to innovate useful applications of natural

phenomena. It is always dissatisfied with present methods and equipment. It seeks newer, cheaper, better means of using natural sources of energy and materials to improve the standard of living and to diminish toil.

Types of engineering. Traditionally there were two divisions or disciplines, military engineering and civil engineering. As knowledge of natural phenomena grew and the potential civil applications became more complex, the civil engineering discipline tended to become more and more specialized. The practicing engineer began to restrict operations to narrower channels. For instance, civil engineering came to be concerned primarily with static structures, such as dams, bridges, and buildings, whereas mechanical engineering split off to concentrate on dynamic structures, such as machinery and engines. Similarly, mining engineering became concerned with the discovery of, and removal from, geological structures of metalliferous ore bodies, whereas metallurgical engineering involved extraction and refinement of the metals from the ores. From the practical applications of electricity and chemistry, electrical and chemical engineering arose.

This splintering process continued as narrower specialization became more prevalent. Civil engineers had more specialized training as structural engineers, dam engineers, water-power engineers, bridge engineers; mechanical engineers as machine-design engineers, industrial engineers, motive-power engineers; electrical engineers as power and communication engineers (and the latter divided eventually into telegraph, telephone, radio, television, and radar engineers, whereas the power engineers divided into fossil-fuel and nuclear engineers); mining engineers as metallic-ore mining engineers and fossil-fuel mining engineers (the latter divided into coal and petroleum engineers).

As a result of this ever-increasing utilization of technology, people and their environments have been affected in various ways—some good, some bad. Sanitary engineering has been expanded from treating the waste products of humans to also treating the effluents from technological processes. The increasing complexity of specialized machines and their integrated utilization in automated processes has resulted in physical and mental problems for the operating personnel. This has led to the development of bioengineering, concerned with the physical effects upon humans, and management engineering, concerned with the mental effects.

Integrating influences. While the specialization was taking place, there were also integrating influences in the engineering field. The growing complexity of modern technology called for many specialists to cooperate in the design of industrial processes and even in the design of individual machines. Interdisciplinary activity then developed to coordinate the specialists. For instance, the design of a modern structure involves not only the static structural members but a vast complex including moving parts (elevators, for example); electrical machinery and power distribution; communication systems; heating, ventilating, and air conditioning; and fire protection. Even the structural members must be designed not only for static loading but for dynamic loading, such as for wind pressures and earthquakes. Because people and money are as much involved in engineering as materials, machines, and energy sources, the management engineer arose as another integrating factor.

Typical modern engineers go through several phases of activity during their careers. Formal education must be broad and deep in the sciences and humanities underlying the particular field. Then comes an increasing degree of specialization in the intricacies of the discipline, also involving continued postscholastic education. Normal promotion thus brings interdisciplinary activity as the engineer supervises various specialists. Finally, the engineer enters into the management function by interweaving workers, money, materials, machines, and energy sources into completed processes for the use of humankind.

For specific articles on various engineering disciplines *see* CHEMICAL ENGINEERING; CIVIL ENGINEERING; ELECTRICAL ENGINEERING; INDUSTRIAL ENGINEERING; MANUFACTURING ENGINEERING; MARINE ENGINEERING; MECHANICAL ENGINEERING; METHODS ENGINEERING; MINING; NUCLEAR ENGINEERING; SCIENCE; TECHNOLOGY.

Joseph W. Barker

Engineering, social implications of

The rapid development of human ability to bring about drastic alterations of the environment has added a new element to the responsibilities of the engineer. Traditionally, the ingredients for sound engineering have been sound science and sound economics. Today, sound sociology must be added if engineering is to meet the challenge of continued improvement in the standard of living without degradation of the quality of the environment.

Despite the fact that present and evolving engineering practices must meet the criteria of scientific and economic validity, these same practices generally cause societal problems of new dimensions. Consider, for example, exhaust gases emitted from tens of millions of internal combustion engines, both stationary and moving; stack gases from fossil-fuel-burning plants generating steam or electric power; gaseous and liquid effluents and solid waste from incinerators and waste-treatment systems; strip mining of coal and mineral ores; noise issuing from automotive vehicles, aircraft, and factory and field operations; toxic, nondegradable or long-lived chemical and particulate residues from ore reduction, chemical processing, and a broad spectrum of factory and mill operations; dust storms, soil erosion, and disruption of ground-water quality and quantity accompanying intensified mechanized farming in conjunction with massive irrigation and fresh-water diversion. *See* ECOLOGY, APPLIED.

Progress often results in the substitution of one set of problems for another. For example, in nuclear

electric power generating plants, replacement of fossil fuels by nuclear fuels relieves the burden of atmospheric pollution from stack gas emissions. Lower thermal efficiency of a nuclear plant, however, results in higher heat rejection rates and increased thermal pollution of sources of cooling water or air. The attendant consequences on atmospheric conditions or on the viability of aquatic life in the affected water are of great concern in the short and long terms. Ultimately, the cost and benefit considerations of nuclear power must be all-inclusive; in addition to usual considerations of economic length of plant life and so forth, one must account for all the economic and societal costs of the entire fuel cycle, from mining and refinement through use and ultimate recycling or safe disposal. The long-term effects of very low levels of radiation exposure (as such studies become available) will be an additional factor to consider.

The modern engineer must be increasingly conscious of the societal consequences of technological innovation. *See* ENGINEERING. Eugene A. Avallone

Engineering and architectural contracts

The legal documents pursuant to which most professional engineering and architectural services are rendered. A contract is the statement of promises agreed to by two or more parties in order to effectuate a specific goal. For the law to enforce these promises, certain elements must be present: mutuality of obligation, that is, each of the parties must have an obligation to the others; a meeting of the minds, that is, it must appear that the parties agree; and all material elements of the proposed transaction must be provided for.

Types. Engineers and architects, (that is, design professionals,) typically draw up three types of contract. (1) The design contract is an instrument under which the design professionals provide services to their client, who is usually the owner of the project. (2) The design subcontract is an instrument under which the original group of design professionals may retain other design professionals to assist them. (3) The construction contract is an agreement between the owner and construction contractors. It consists of the plans and specifications produced by the engineer, and such other provisions as may be needed to define the relative rights and responsibilities of owner and contractor in implementing the plans and specifications. Although usually not a party to the construction contract, the design professional normally prepares this document and is usually involved in carrying out the activities detailed in the design contract. Indeed, the purpose of most design contracts is the preparation of the construction contract.

Need for written contract. Except where the law requires certain types of agreements to be in writing, an oral contract is as enforceable as a written one. Within the United States, the law in most states requiring that certain types of contracts be in writing (generally called the Statute of Frauds) mandates written contracts in certain circumstances such as the purchase of real property. The purpose of requiring a written contract is to lessen disputes as to what was agreed upon.

Unlike the services of virtually all other learned professionals (for example, doctors, dentists, lawyers, and accountants), the services of design professionals are almost always rendered according to a written contract. The principal reason is that in the construction process many parties are involved (for example, the owner; architect; surveyor; soils, structural, electrical, heating, ventilating, and air-conditioning engineers; the various prime contractors and subcontractors; and manufacturers and suppliers). Without written agreements to define the rights and obligations of each of these parties, there would be hopeless confusion.

Contract enforcement. Contracts are enforced by the legal system. Contracts are interpreted and the rights of the parties are determined either by courts or by arbitration (most often under the Construction Industry Rules as promulgated and administered by the American Arbitration Association). Arbitrators have no power to enforce their decisions, but the laws in virtually all states provide that arbitrators' decisions will be enforced by the legal system.

Where a matter is submitted to the court, it attempts to apply established legal principles to enforce the terms of the contract and to give each party what was agreed to, even though the results might appear to be inequitable. Arbitrators, however, are free to ignore the law and make rulings based upon their personal views of what is equitable. Although it might appear that an equitable result having no legal foundation is better than an unfair legal result, disregard of legal principles by arbitrators destroys one of the functions of a legal system—to provide predictability so that contracts can be entered into with an awareness on the part of the signers of how their contracts will be interpreted.

Since both litigation in court and arbitration are time-consuming and costly processes, more and more construction disputes are resolved by alternate dispute resolution techniques such as mediation. Although the suggestions of a mediator are wholly unenforceable (unlike the rulings of an arbitrator), mediation has proved to be a very efficient approach to the resolution of construction industry disputes.

The most effective tool for avoiding disputes remains the carefully drawn contract clearly setting forth the rights and obligations of the parties and anticipating most of the situations that might arise during the course of a construction project.

Tort liability. While contract law enforces the rights of the parties to the contract, tort law enforces the rights and duties of those who are not parties to the contract (that is, in privity). A design professional has an obligation to persons not related by contract, such as the general public, construction workers, or construction contractors (third parties). In general, the design professional has the duty to be free from negligence in the performance of services. Where it is reasonably foreseeable that the negligent performance of services will cause harm to such third

parties, and such parties do suffer harm, the design professional will be liable for the damages suffered. This rule exists in the United States wherever the damage suffered by the third party is personal injury, loss of life, or physical damage to property. In most states the design professional is also liable to third parties if the professional's negligence causes only an economic loss; however, a significant minority of states do not impose liability for economic loss.

It is less clear to which third parties is owed the duty to exercise reasonable care. Courts usually look to the design professional's contracts to determine the class of persons to whom the design professional owes duties.

Standard forms of agreement. Design professionals frequently utilize standard contract forms that are intended to be used (with certain modifications) on a wide variety of projects. The American Institute of Architects (A.I.A.) publishes a wide range of such standard documents, as does the Engineering Joint Contract Documents Committee (EJCDC). The EJCDC, which is composed of representatives from the American Consulting Engineers Council, the American Society of Civil Engineers, the Construction Specifications Institute, and the National Society of Professional Engineers, publishes sample documents covering every phase of the design and construction process. The A.I.A. documents focus on building projects involving architects.

A standard contract is a widely applicable document that, while it covers the subject, does not fit the endeavor precisely, as a custom-tailored document would. Although the majority of the A.I.A. and EJCDC documents provide for additions of project-specific clauses and selecting alternative provisions to accommodate alternative methods of payment, what emerges is still a basic document that has been altered—it is still not custom tailored for its intended use. However, not all consumers can afford custom tailoring, and not all projects warrant such cost. Furthermore, many attorneys are not sufficiently versed in the construction process to produce a good custom-tailored contract, and most design professionals are not sufficiently versed in the law to produce a legally effective contract. Accordingly, the use of a standard document may be a better choice than an inadequate custom-tailored document. The greatest value of the standard documents may be to serve as a check list of what properly belongs in the contract, and to provide certain general clauses that serve to limit the design professionals' liability.

There are certain key provisions of the standard documents, provisions which in some form or other should appear in the design professionals' contracts. These include description and limitation of services, limitations of parties to whom duty is owed, and specific details of the construction contract.

Description and limitation of services. The key provisions in the contract are the establishment of the duties of each party. In the case of the owner, the obvious duties relate to the amount and timing of payment to the design professional (design contract) or to the construction contractor (construction contract). In the design contract, a detailed description of the design professional's duties is essential, especially where there may be others rendering services to the owner in connection with the design (for example, surveyors or soils engineers). During construction there are usually even more parties involved in providing goods and services, and there may be considerable overlap. A number of important concerns must be addressed, such as (1) whether the design professional who has performed the design is responsible for seeing that the construction contractor follows the plans, and if so, whether this duty applies only to the owner or also to the construction contractor; (2) whether the design professional is responsible if the contractor's operations result in injury to workers or others; and (3) whether the design professional is responsible if the method of construction chosen by the contractor is not feasible or unnecessarily costly.

A listing of specific services that the design professional is to perform is often insufficient to clearly define his or her responsibilities. The reason is that an obligation to perform certain other related services might be implicit in performing the listed services. For example, a requirement that the design professional design a building to be used as a diner might be interpreted to also require him or her to perform the property survey, collect and analyze soil samples, design the foundation, specify food-handling equipment, or perform environmental analyses. If it is not the intent of the parties that the design professional perform certain of these services, the contract should explicitly specify what is not to be done as well what is to be done.

Limiting parties to whom duty is owed. The law often looks to the design professional's contracts to determine whether a duty was owed to a third party. If the design contract requires the design professional to see that the construction contractor employs adequate safety precautions, the law will say that the design professional owes a duty to workers and the general public to see that the workplace is safe. If the design contract requires the design professional to inspect the work to guarantee to the owner that the construction conforms to the plans and specifications, the design professional has guaranteed to the owner that all construction will be in accordance with the contract and the design professional will be liable to the owner if the construction contractor has not followed the plans and specifications.

If it is not the intention that the design professional be responsible to third parties for the construction contractor's safety precautions, the design contract should explicitly so state and the design professional should not exercise any control over the contractor's operations. If it is not the intention that the design professional be the guarantor of the construction contractor's performance, the design contract should explicitly say this. Both the A.I.A. and the EJCDC standard documents provide language that clearly limits the engineer's responsibilities in these two areas.

Construction contract provisions. Although the design professional is not a party to the construction contract, he or she is nevertheless affected by that document. First, it is usually the design professional who prepares the construction contract documents, including the plans and specifications that define the construction contractor's duties. Second, the legal system looks to the construction contract should there be a question as to whether a duty is that of the design professional or the construction contractor, or both. For example, if the construction contract requires the contractor to prepare supplementary plans or details and submit them to the design professional for review and approval, it is reasonably clear the design professional has a duty to check the construction contractor's work. If the construction contract makes the construction contractor exclusively responsible for safety precautions and the means and methods of construction, in the absence of specific language also assigning this responsibility to the design professional the courts will not find that the design professional had any responsibilities in these areas. Again, the A.I.A. and EJCDC standard construction contract documents clearly define the relative responsibilities between the design professional and construction contractor.

Subcontracts for professional services. Many design professionals may be involved in a single project. While some owners contract separately with the various design professionals, most prefer to deal with a single professional who will be responsible for the performance of all design services. In the case of a building project the owner contracts with an architect, who in turn is expected to retain the services of structural and mechanical engineers and any other professional services required. The considerations necessary in preparing these subcontracts are essentially identical to the key provisions for the design contract. Both the EJCDC and the A.I.A. publish standard forms of subcontract agreements (but subject to the limitations of any standard contract as discussed earlier). *See* ARCHITECTURAL ENGINEERING; ENGINEERING.

Donald A. Ostrower

Engineering design

Engineering is concerned with the creation of systems, devices, and processes useful to, and sought by, society. The process by which these goals are achieved is engineering design.

The process can be characterized as a sequence of events as suggested in **Fig. 1**, with the recognition that no final, universally accepted description of so complex an intellectual and physical exercise, applicable to an enormously broad spectrum of products and processes, is possible. The process may be said to commence upon the recognition of, or the expression of, the need to satisfy some human want or desire, the "goal," which might range from the detection and destruction of incoming ballistic missiles to a minor kitchen appliance or fastener.

Concept formulation. Since the human aspiration is usually couched in nonspecific terms as a sought-for goal, the first obligation of the engineer is to develop more detailed quantitative information which defines the task to be accomplished in order to satisfy the goal, labeled on Fig. 1 as task specification. At this juncture the scope of the problem is defined, and the need for pertinent information is established. The source of the original request is questioned to establish the correspondence between the developing specifications and the initial definition of the goal.

But to know that a need exists and to have started on the task of qualitatively and quantitatively defining its substance and bounds should not be confused with the generation of ideas for possible solutions to the problem. This creative stage is called the concept formulation. When great strides in engineering are made, this represents ingenious, innovative, inventive activity; but even in more pedestrian situations where rational and orderly approaches are possible, the conceptual stage is always present.

The concept does not represent a solution, but only an idea for a solution. Initially evisceral and ephemeral, it can only be described in broad, qualitative, frequently graphical terms. Concepts for possible solutions to engineering challenges arise initially as mental images which are recorded first as sketches or notes and then successively tested, refined, organized, and ultimately documented by using standardized formats.

At this point it is important to note that the necessarily two-dimensional description of Fig. 1 and this sequential textural description should not be construed as an implication that the goal, task specification, and concept always appear in a simple, sequential temporal order. In fact, a central characteristic of the design process is the unpredictable emergence of, and iterations between, the various steps. This stochastic character can be suggested by drawing two-way connections between the various steps. For example, the definition of a task and emergence of a concept might precipitate restudy and possible alteration of the original goal. Consider the not infrequent experience of the serendipitous emergence of a new and interesting concept, and then subsequent

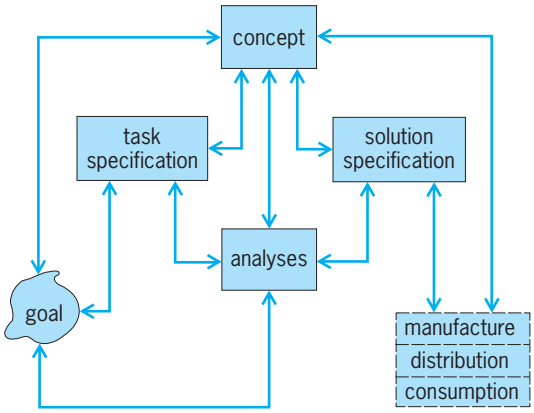

Fig. 1. Engineering design process.

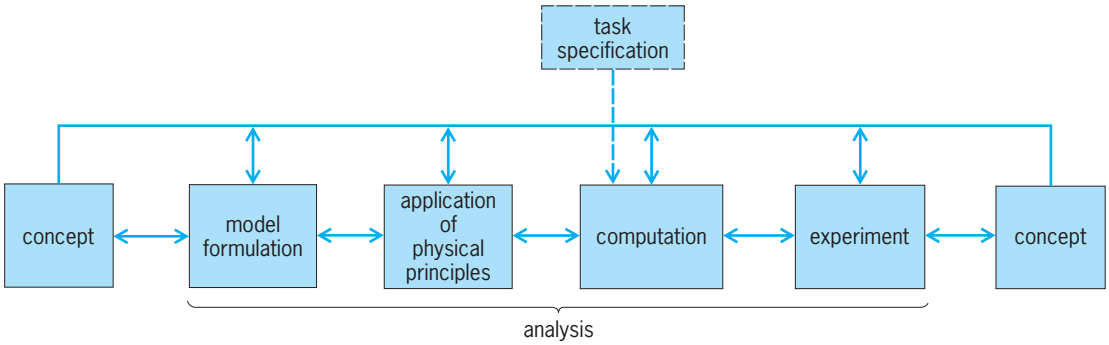

Fig. 2. Role of analysis in concept formulation.

search for a goal (application or market) to which to apply the idea.

Concepts are accompanied and followed by, sometimes preceded by, acts of evaluation, judgment, and decision. It is in fact this testing of ideas against physical, economic, and functional reality that epitomizes engineering's bridge between the art of innovation and science. The process of analysis is sometimes intuitive and qualitative, but it is often mathematical, quantitative, careful, and precise. The iteration of concept and analysis invariably gives rise to a focusing and sharpening, possibly a complete change, of the concept. Frequently, in light of the analyses conducted, what can or might be accomplished is illuminated and task specifications changed. Even goals may be altered. That which was sought may be beyond accomplishment, or perhaps solutions more sophisticated or more useful than initially undertaken may prove achievable.

Production considerations can have a profound influence on the design process, especially when high-volume manufacture is anticipated. Evolutionary products manufactured in large numbers, such as the automobile, are tailored to conform with existing production equipment and techniques such as assembly procedures, interchangeability, scheduling, and quality control. New techniques such as those associated with space exploration, where volume production is not a central concern, factor into the engineering design process in a very different fashion.

Similarly, the design process must anticipate and integrate provisions for distribution, maintenance, and ultimate replacement of products. Well-conceived and executed engineering design will encompass the entire product cycle from definition and conception through realization and demise and will give due consideration to all aspects.

The iterations of analysis and concept and considerations of manufacture develop information which defines a sequence of progressively specific solutions. In final form the solution specification consists of all drawings, materials and parts lists, manufacturing information, and so on necessary for construction of the device, system, or process.

It is important to note that, while the culmination of a particular engineering design process defines but one solution, inspection of the process during its evolution indicates a complex series of discrete partial, temporary, interim solutions which are compared one with another, and out of this comparison emerges the final compromise solution. The limitations on this scrutiny of alternative approaches are the practical contingencies of time and personnel, or other ways of measuring resources. When these assets reach their budgetary limits, the most satisfactory solution is accepted; however, in as dynamic a field as engineering the final design is not necessarily an ultimate or optimal best.

Figure 1 provides a broad overview of the engineering design process. Some expansion of the interaction between conceptualization and analysis is warranted, as shown in **Fig. 2**. On the left- and right-hand sides is concept; everything between is characterized as analysis.

Any physical entity, existing or hypothesized, of any degree of complexity cannot be analyzed in its entirety because of inadequate knowledge of the relevant physical laws, or inadequate time or facilities for the required computation, or a combination of these shortcomings. For these reasons, plus its inherent initial vagueness, the concept cannot be analyzed completely. Instead, simplified models are deliberately and precisely defined by applying established physical principles and laws to describe the model via mathematical equations. By using numerical values from the task specifications for parameters, the requisite computational tasks are performed.

In many engineering situations, particularly those where there is no body of experience with similar geometries, materials, and so on, the model from which the analysis is derived cannot be confidently assumed to characterize completely all significant attributes of the ultimate physical system. In some cases certain relevant processes are not completely understood; other times adequate resources to perform all pertinent analyses are not available. Thus, recourse to experiment is necessary. Since nature is the final arbiter of all physical proposals, however analyzed, experiment or test always precedes final acceptance of any proposal.

As previously suggested for Fig. 1, where all possible interconnections occurred between various stages and the process underwent a dynamic evolution, complex interactions occur also in the concept-analysis loop of Fig. 2. For example, the model is

defined, based on the concept recognizing existing physical knowledge about included processes and phenomena, and the effort that will be involved in the reduction of the resulting mathematical equations into quantitative results is anticipated. For example, an initial model might be a greatly simplified, or perhaps oversimplified, characterization of the concept in order that the computational results might quickly identify the utility of the idea or its ranges of applicability. Subsequently, more refined models of aspects of the concept of greater subtlety might be subject to scrutiny.

Where the function of the part is critical or the physical knowledge is inadequate, the engineer might, in fact, eliminate computation and move rapidly to experiment, perhaps on an analog or scaled-down version of important attributes of the concept.

Hierarchy of design. An adequate description of the engineering design process must have both general validity and applicability to a wide variety of engineering situations: tasks simple or complex, small- or large-scale, short-range or far-reaching. That is to say, there is a hierarchy of engineering design situations.

Systems engineering occupies one end of the spectrum. The typical goal is very broad, general, and ambitious, and concepts are concerned with the interrelationships of a variety of subsystems or components which, when taken together, make up the system to accomplish the desired goal. *See* SYSTEMS ENGINEERING.

At a subsidiary level of the design problem hierarchy, the same engineering design process applies to creation of a device which might be one component of the overall system. And at the most detailed end of this hierarchy the same process diagrams the engineering design of a single element of a component. Obviously, as the engineering design process is applied to create one of these several elements, components, or systems, different phases of the process come into play in different ways and to different degrees, depending upon the particular problem.

As an illustration of this hierarchy of design problems, all of which exhibit the schema of the engineering design process, consider a particular goal, the interplanetary inspection of the atmosphere of Venus by the *Mariner 2* space probe. At the system level this study encompasses considerations of possible Earth-Venus trajectories based on astronomical events, launching times, booster-spacecraft weight-thrust combinations, and so on. Projections of foreseeable booster capabilities balanced against estimates of the spacecraft weight, combined with the astronomical "windows" to Venus which result from Earth-Venus juxtaposition, give rise to detailed analyses of conceivable space paths which pose the need for midcourse maneuver capabilities.

The scientific measurements desired and the assumed Venus atmosphere postulate the functional characteristics of instruments: weight, volume, and power consumption. Control and telemetry requirements to and from the spaceship likewise augment

Fig. 3. Flight configuration of *Mariner 2* spacecraft.

projected weight, volume, and power consumption needs. Combinations of these requirements and others give rise to concepts for the spaceship design, the flight configuration of which is shown in **Fig. 3**. *Mariner 2* is an interesting example of new design, there being no established precedent as to what a Venus probe should look like. *See* SPACE PROBE.

Each major component of the spacecraft—structure, internal power supply, solar collecting vanes, radio telemetry antenna, and so on—undergoes a similar evolution from requirements and concepts through analysis. Each major subsystem is composed of subcomponents, each of which is composed of elements; **Fig. 4** shows the detailed configuration of the planetary boom. The same process by which the major spacecraft arrangement decisions were made applies again; individual concepts and analyses are permutated in order to arrive at the design configuration of elements, such as this planetary boom. Thus, the same general overall diagrams of Figs. 1 and 2 characterizing the engineering design process can be applied in an ever broadening and encompassing fashion from the most minor

Fig. 4. Planetary boom of the *Mariner 2* spacecraft.

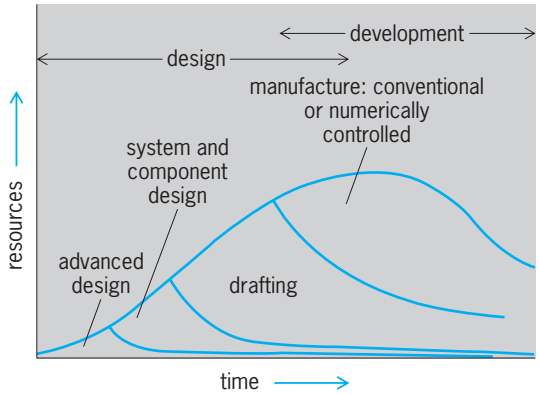

Fig. 5. Elapse of time and resources in an engineering design project showing various stages in sequence.

components to the most major system aspects of an overall design. In point of fact, in an actual situation the order is usually reversed, the process being applied first to the most sweeping question, and progressively to the more detailed aspects, although it is important to understand that the overall design is the sum of all its parts, that each part makes a necessary contribution to the whole, and that the performance of the whole is dependent on each part.

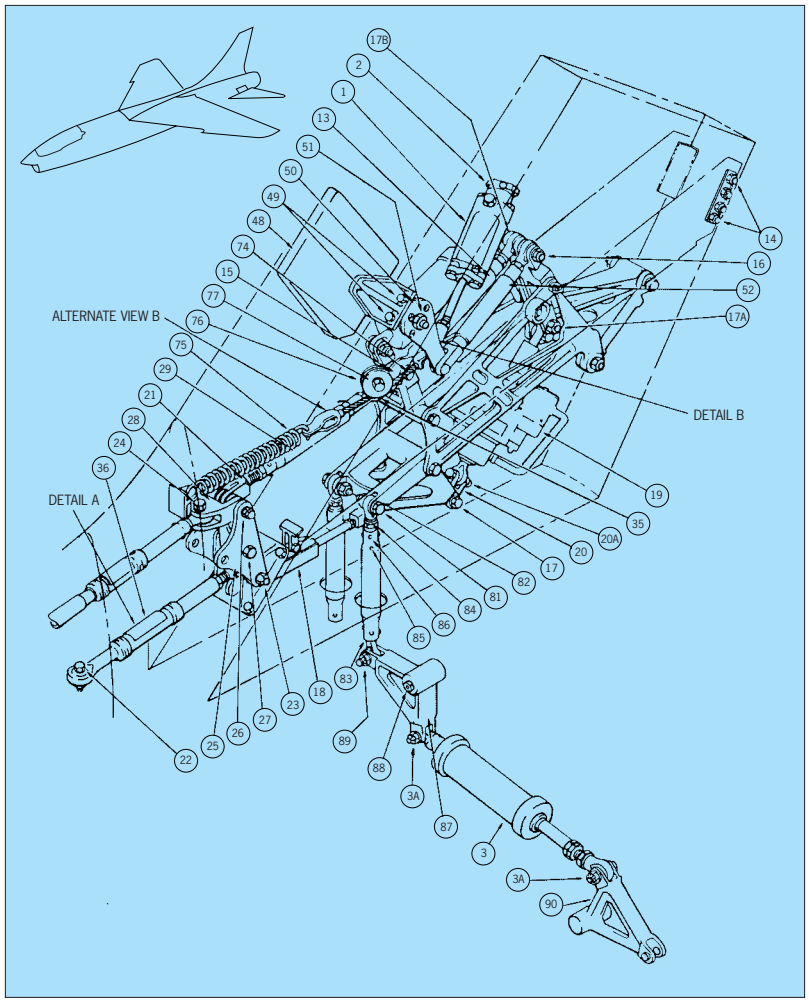

Fig. 6. One of the 10,000 sheets involved in the design engineering of a military aircraft. Numbers refer to the various components and are an indication of the tremendous bookkeeping involved in a project of this type.

The project is interesting as an example, not only because of its novel characteristics, adventuresome quest, and dramatic success, but because it also demonstrates the dynamics of the design process. The spaceship weight, of course, depended upon the particular booster rocket designated to bear it into interplanetary space. As the booster itself was under development during the spaceship development, the permissible weight of the Mariner vehicle varied in accordance with the predicted capability of the booster. Shortly before one of the scheduled departure times, major reductions in the vehicle were made mandatory by a change in the booster configuration. But this impermanence and the variability of the "ground rules" and "boundary conditions" of a design situation are always present, although perhaps not as dramatically as in this case.

Another aspect of the Mariner project particularly appropriate to this discussion is the importance of time and schedule to the success of the mission. Trajectory "windows" to Venus are only available at certain discrete, foreseeable times. Either the vehicle-booster combination would be ready to launch at the designated time, or else there would be no possibility of success. Again, most design situations are characterized by this urgency of time and schedule.

Finally, while the government-funded NASA program is perhaps not the best example of cost as the vital common denominator of all industrial and commercial engineering ventures, even in this case there was a budget, not a matter of profit or return on investment, but rather the limited skilled worker-hours available to meet the commitment.

Time–source dynamics. Another dimension of the dynamics of the engineering design process is the elapse of time and expenditure of worker effort in the evolution of an engineering design project. **Figure 5** plots time as the abscissa and resources (worker-hour or dollars) as the ordinate. The various stages of the engineering design process are set out in time sequence from left to right.

Goal refinement, task specification, and first-order concept and analyses iterations are conducted by one to a few engineers in the early stages to establish the feasibility of the idea and to block out possible approaches. This is called the advanced design stage.

As the design concept becomes more specific and substantive, more and more engineers, technicians, and draftsmen become involved in the project. For example, in the case of a modern aircraft probably fewer than a dozen extremely talented engineers carry through the early feasibility and configuration studies, while at a later stage hundreds of engineers, drafters, expediters, and coordinating personnel are involved. This deployment aspect of the design process cannot be overemphasized. In projects of significant size, the problem of coordinating and integrating the efforts of the many participants of different talents and skills becomes itself a major consideration. While some of the coordination involves judgment and decisions, much of this coordination is

purely clerical and some involves prosaic application of standard reference material. *See* PERT.

Ultimately the process culminates crescendolike in the solution specification—manufacturing drawings and specifications, parts lists, and so on. Where automated manufacturing techniques are warranted, the efforts of computer programmers are devoted to the transformation of graphic and numeric manufacturing information into a form decipherable by computers and machine-tool directors. Ultimately, the physical parts are realized through the manufacturing process.

The efforts of this large group of people demand a high level of coordination and integration if each of the thousands of separate parts of the aircraft are to satisfy its function, be compatible with one another, and be adequately strong and yet of minimum weight and volume. **Figure 6** suggests a few of the individual elements which must undergo design, one sheet of a 2-ft-high (0.67-m) pile of similar diagrams for a typical jet aircraft. The parts shown represent only a few of the better than 10,000 parts that make up the typical jet aircraft.

Use of the computer in design. Reference to the role of analysis in Fig. 1 and to computation in Fig. 2 highlights the ever increasing use of the computer, both analog and digital, in the engineering design process. As a high-powered successor to the slide rule and desk calculator, the computer is used routinely for much of the calculation, computation, and data reduction which constitute a major activity in design. Where repetitive series of calculations are carried out, programs are prepared to relegate to the computer more and more of the responsibility for analysis. Where economically justified, the overall engineering design process for a product is mechanized via computer programming. For example, **Fig. 7** describes the flow chart of a branching computer program which, given vital data on the requirements of an electrical power transformer, carries out the design automatically. In the process the program, having made assumptions on core size and copper windings, calculates temperature gradient, impedance, and so on; checks these against preprogrammed constraints; and optimizes its choices considering, among other things, the current cost of copper. **Figure 8** illustrates typical iterations of the computer design as it "homes" in on the lowest weight and cost design satisfying the initial data and the internal constraints. *See* COMPUTER; COMPUTER PROGRAMMING.

The speed, memory, and accuracy of the computer to iteratively calculate, store, sort, collate, and tabulate have greatly enhanced its use in design and encouraged the study, on their own merits, of the processes and subprocesses in the design process. These include optimization or sensitivity analysis, reliability analysis, and simulation as well as design theory. *See* DIGITAL COMPUTER.

Optimization analysis, given a model of the design and using linear and nonlinear programming, determines the best values of the parameters consistent with stated criteria and further studies the effects of

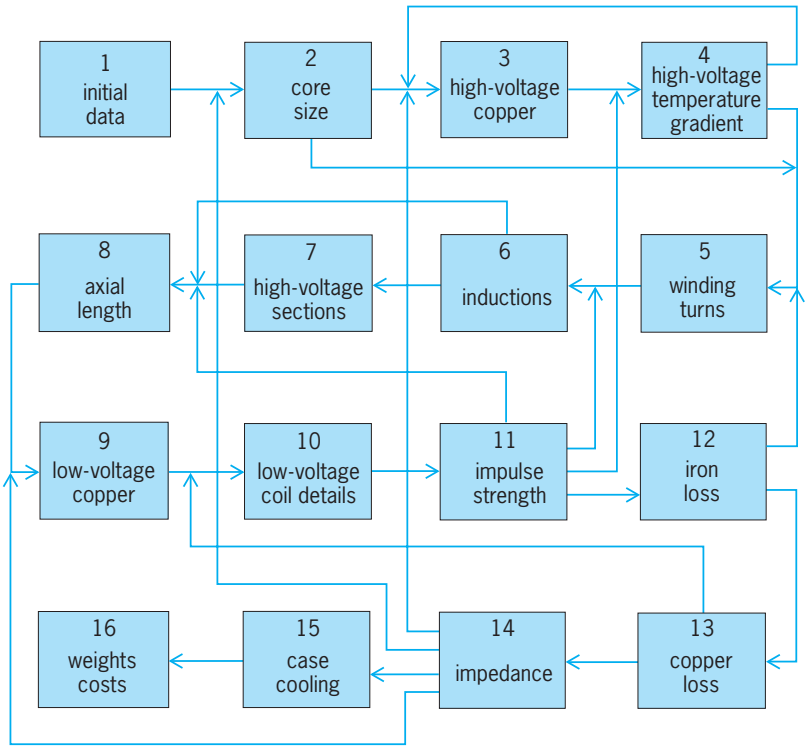

Fig. 7. Flow chart of a computer program.

variations in the values of the parameters. *See* OPTIMIZATION.

Reliability is a special case of optimization where the emphasis is to choose or evaluate a system so as to maximize its probability of successful operation, for example, the reliability of electronics. *See* CIRCUIT (ELECTRONICS); RELIABILITY, AVAILABILITY, AND MAINTAINABILITY.

Simulation as of dynamic systems, is mathematical modeling to study the response of a design to various inputs and disturbances. The analog computer was formerly widely used for simulation through its physical modeling of the mathematic analytical relationships of the proposed design. The digital computer's use of numerical data adapts it more readily to nonlinear or probabilistic situations by using random or Monte Carlo techniques, as well as to those situations requiring higher accuracy. *See* ANALOG COMPUTER; MONTE CARLO METHOD; SIMULATION.

Decision theory deals with the general question of how to choose between a great number of alternatives according to established criteria. It proposes models of the decision process as well as defining techniques, that is, programs or algorithms, of calculation by which to make choices. *See* DECISION THEORY.

Artificial intelligence or its less pretentious subdiscipline, expert systems, strives to identify and codify human thought processes and produce computer algorithms and programs for automatic problem solving, including facets of the design process. *See* ALGORITHM; ARTIFICIAL INTELLIGENCE; EXPERT SYSTEMS.

Although each of these techniques represents either a gross oversimplification or but a part of the

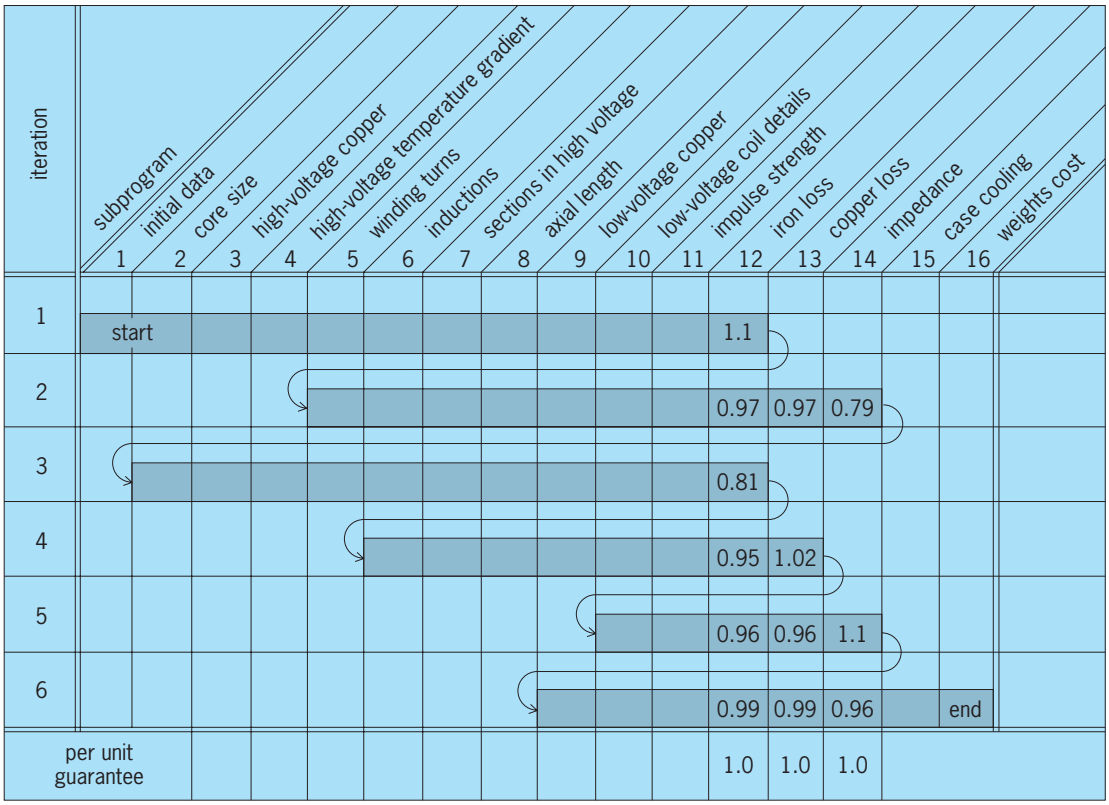

Fig. 8. Iterations of the computer program shown by the flow chart in Fig. 7.

overall design process of Fig. 1, their study and application are useful. Even where not strictly applicable, the awareness of a model of a process enhances its qualitative evaluation by providing a guide to action. The utility and comprehensiveness of these techniques are already advancing and will mature rapidly, largely because of the ability of the high-speed digital computer to carry out evaluations which heretofore were impractical. Finally, the record-keeping capacity of the computer makes possible the recording and reutilization of past action. This integration of prior results into new endeavors represents a "mechanization of experience."

Computer utilization. Notwithstanding the present contribution and unlimited promise of the computer to the design process, there are certain limitations. In reflecting on the model of the engineering design process and especially on the interaction between concept and analysis, it is useful to consider the nature of the activities involved. Concepts, or really new ideas, usually do not come about as a result of any ponderous, systematic, organized effort; rather they appear in an undisciplined, creative, spectacular fashion. The short-duration exultation of concept creation is then followed by usually very much longer, more regular, systematic periods of analysis. The typical use of a computer does not lend itself to the kind of person-analysis interaction implicit in the engineering design process and essential to its successful negotiation. Usually, the solution process for a piece of work must first be well thought out and known in the most intricate detail; only then can the user write a program to describe to the computer the sequence of steps necessary. But when the nature of the problem is only vaguely grasped and much learning has to be done, the computer may not be as helpful as pencil and paper.

The reason is that in the thinking process one needs to advance in steps and to test these steps frequently. While the steps may be large, as in the first gross examination of an idea, or small, as in an exercise of refinement, these tests need not be elaborate or even precise. What is wanted first is only a qualitative result or perhaps a quantitative result of only moderate precision, a confirmation or a denial of tentative guess work.

By conventional computer programming, it is virtually impossible to obtain quick answers to small discrete problems, even though the computer can work very rapidly. To solve a problem, the user must first prepare a program, a detailed ritual of calculations and comparisons of quantities for the computer to carry out. It is hoped that data so processed will yield the solution.

If there is a mistake in the program, the computer either will detect the mistake and refuse to waste time by trying to run the program, or if the mistake is more subtle, perhaps an error in principle, will go ahead and compute a mass of meaningless nonsense. This dependence on programming, waiting, and error finding and correcting frustrates the formulation and testing of concepts essential to a creative effort.

Another point is that the creative process is, virtually by definition, unpredictable. The sequence of the steps is never known at the beginning. If

it were, the whole process could be accomplished by the computer since the information prerequisite to the computer program would be available. Indeed, the creative process is the process of learning how to accomplish the desired result.

Clearly what is needed if the computer is to be of greater use in the creative process is a more intimate and continuous interchange between engineer and machine. This interchange must be of such a nature that all forms of thought congenial to humans, whether verbal, symbolic, numerical, or graphical, are also understood and acted upon by the machine in ways appropriate to the user's purpose. In reflecting on this human-machine symbiosis with the goal of designing a system with which to design, it is useful to identify the special attributes of each partner. *See* HUMAN-MACHINE SYSTEMS.

The speed of the computer is prodigious when measured against any term of reference and especially so when measured against the appropriate one, the human mind. A computer performs millions of arithmetic or logical operations per second. This rate outstrips by many orders of magnitude the human neurological responses to premeditated cerebrations.

The machine's memory is likewise extraordinary, although restricted to an extremely narrow but nonetheless useful class of memory impression. It can store binary numbers of 32 digits and recall them unerringly at microsecond speeds, while the human memory system is very much slower and not nearly so reliable. Of course, in contrast, the human memory encompasses an indescribable universe of types and kinds, far beyound the range of a computer. The reliability routinely exercised by the computer is unattainable by people. Once properly instructed, the computer executes its routines faithfully, untiringly, yielding to no boredom or carelessness, and introducing no humanlike errors. *See* COMPUTER STORAGE TECHNOLOGY.

Counterposed against the computer's assets stands an open-ended list of human attributes. In terms of the engineering design process, these human characteristics would include reflection on, and evaluation of, the social, esthetic, and economic aspects of the original goals; the formulation of previously unforeseen and unanticipated questions at many points in the investigation; an unflagging curiosity about the way things are done and about the way they are proposed to be done; the flexibility of mind to shift from one approach to another, to sort out the significant from a great mass of information and misinformation on the basis of very few discernible criteria; powers of mental association which can detect, correlate, and amplify useful relevancies between bodies of normally disparate information; the devising and structuring of new and unusual approaches; the willingness, ability, and intellectual integrity to make decisions, frequently on the basis of inadequate data; the ability to recognize that the job must get on and some decisions must be made rapidly, though the right to reverse decisions whenever facts so indicate is reserved; exercise of judgment based on prior experience in making decisions; and finally the desire and willingness to develop such adjudicative ability through experience. While the computer demonstrates the capability of encroaching upon some of these attributes, but only to the extent that humans thoroughly understand them and can program them, clearly others are forever denied it. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING; COMPUTER-AIDED ENGINEERING.

Time-sharing. Communicating directly with a computer by keyboard with the computer replying immediately on a cathode-ray-tube monitor removes the obstacle to thought caused by waiting hours for a reply. The user has no compunctions about posing smaller problem fragments to the computer, in a way more consonant with the small steps of the creative process. The results are immediately available, and the course of further action is guided more precisely.

Such intimate "conversations" with a computer are scarcely economically feasible when only one person uses the computer at a time because, measured against the speed of a computer, human beings are intolerably slow. A person likes to ponder a problem and often needs time to decide what to try next. Even when a decision is reached, the instruction to the computer takes several seconds to type on the keyboard, and during even this short time the computer could perform hundreds of thousands to millions of calculations were it free to do so. It is logical to arrange to have large main-frame computers in conversation with a large number of people, so that when one user is idle the computer can turn to another and answer that question. Should all users be idle at once, the computer can work on some large problem left as a backlog. This type of effective computer utilization is quite routine, and is called time-sharing. On a rough average, if a user is intensely busy for 1 h, the computer can discharge its responsibilities to this particular user in short bursts that total about 3 min. The rest of the time it serves the other users. *See* MULTIACCESS COMPUTER.

The dramatic reductions in cost and increases in power of contemporary minicomputers have expanded their use in time-share environments. The advent of the microprocessor, or computer on a chip, made economic the personal computer where only one keyboard operator can afford to monopolize a whole computer system. The power of such standalone systems is as yet still very limited for design application; therefore, microcomputers are currently used, together with their keyboards and graphic displays, as "intelligent" terminals on larger minicomputers or mainframe computers, providing the operator access to the computing power and vast memory of the larger machine. *See* MICROCOMPUTER; MICROPROCESSOR.

Graphical input/output. In many fields of design—notably architecture; design of airplanes, automobiles, and ships; consumer products; in almost all mechanical design and in electronic computer circuit design—the designer works largely in visual terms.

Initially limited to interpreting only binary digits, the digital computer was subsequently organized to understand ordinary numbers and words and combinations thereof via the various programming languages. Then, in the early 1960s Ivan Sutherland established bilateral graphical communication between the human and the digital computer. The essential link between the pictorial cognitive style of the design engineer and the speed, memory, and reliability of the digital computer had been forged, but there was a lag of more than a decade before the availability and economy of adequate main-frames and improved graphic interfaces made computer-aided graphics a practical reality. The computer has become an active partner in the act of drawing, so that it can provide a certain superskill in preparing the drawing once the intentions of the human operator have been made clear. *See* COMPUTER GRAPHICS.

Computer-aided design (CAD). The engineering design process, as aided by CAD capability, can be suggested via several examples.

Selecting from a task menu presented on a graphics console, a designer, using keyboard and cursor or light-pen, can then draw on the consol and into the computer the elemental parts of a machine; have the computer embody them as solids; assemble, section, and orient them under operator control; and finally document the parts following traditional drafting and specification standards, all electronically. The production engineer, with access to the common data base, can then invoke computer-aided manufacturing (CAM), optimize the machine tool cutter paths, with the coded information transferred to the actual numerically controlled (N/C) machine tool for part fabrication.

Comprehensive CAD systems provide for human-computer interaction in engineering analysis of the evolving design. A recurrent question in the concept-analysis iteration of Fig. 2 is: Will the strength be adequate? This is addressed in CAD, with a computer graphics portrayal of a hypothetical wrench turning a bolt. Using a stress-analysis technique known as the finite element method, the computer automatically superimposes a mesh into the region the engineer identifies as a potentially highly stressed region in the wrench. Then, having been provided the wrench material strength properties and the applied torque, the computer's finite element method program calculates, and the CAD graphics presents, coded in color, the resulting stress distribution. The engineer must then evaluate these data, accept the results, or intervene and alter the geometry, material, acceptable torque, and so on. The essence of the engineering design process has been retained in CAD, but the process has been vastly facilitated and accelerated. More iterations of more alternatives can be explored more rigorously and more rapidly when compared with traditional engineering design. The results are faster and better design and improved productivity. *See* FINITE ELEMENT METHOD.

Use of CAD (and CAM) will relegate increasingly to the computer tasks it can be programmed to perform automatically, leaving to the human the exercise of goal definition and setting and judgmental evaluation and decision. In one integrated system, progressive solutions of an automatic part design and analysis are produced by a computer program which, given functional geometrical constraints on a part and its structural strength requirements, automatically iterates geometrical designs and finite-element-method analyses to produce the minimum-mass part which satisfies the constraints.

Computer effectiveness. The effect of using computers in the design and development stages of a manufacturing process is twofold. First, the costs are greatly reduced because fewer people are involved in the stages, and less time is required to accomplish the tasks. For example, the translation of a designer's sketch of the shape of an airplane fuselage into a full-size, precise, geometrical shape description suitable for manufacturing purposes using traditional methods used to take about 6 months. Conventional, that is, non-CAD, computer methods reduce this time to a few weeks. But the emergent computer techniques will cause a further reduction of this time, perhaps to seconds. The results will be far more accurate than have previously been obtained and will be tailored much more closely to manufacturing requirements.

Second, an important hidden benefit is the greatly increased lead time which a CAD/CAM manufacturer gains over a noncomputerized (or less computerized) competitor; the former can be ready much sooner to begin production. The national military and the commercial aspects of this acceleration are obvious.

Robert W. Mann

Bibliography. J. L. Adams, *Conceptual Blockbusting*, 3d ed., 1990; G. L. Glegg, *The Design of Design*, 1969; M. P. Groover and E. W. Zimmers, Jr., *CAD/CAM: Computer-Aided Design and Manufacturing*, 1984; F. H. Jones, *Computer Aided Architecture and Design*, 1994; J. E. Shigley and L. D. Mitchell, *Mechanical Engineering Design*, 5th ed., 1989; M. F. Spotts, *Design of Machine Elements*, 7th ed., 1997.

Engineering drawing

A graphical language used by engineers and other technical personnel associated with the engineering profession. The purpose of engineering drawing is to convey graphically the ideas and information necessary for the construction or analysis of machines, structures, or systems.

In colleges and universities, engineering drawing is usually treated in courses with titles like Engineering Graphics. Sometimes these courses include other topics, such as computer graphics and nomography. *See* COMPUTER GRAPHICS.

The basis for much engineering drawing is orthographic representation (projection). Objects are depicted by front, top, side, auxiliary, or oblique views, or combinations of these. The complexity of an object determines the number of views shown. At times, pictorial views are shown. *See* DESCRIPTIVE GEOMETRY; PICTORIAL DRAWING.

Engineering drawings often include such features as various types of lines, dimensions, lettered notes, sectional views, and symbols. They may be in the form of carefully planned and checked mechanical drawings, or they may be freehand sketches. Usually a sketch precedes the mechanical drawing. Final drawings are usually made on tracing paper, cloth, or Mylar film, so that many copies can be made quickly and cheaply by such processes as blueprinting, ammonia-developed (diazo) printing, or lithography. *See* PHOTOCOPYING PROCESSES.

Section drawings. Many objects have complicated interior details which cannot be clearly shown by means of front, top, side, or pictorial views. Section views enable the engineer or detailer to show the interior detail in such cases. Features of section drawings are cutting-plane symbols, which show where imaginary cutting planes are passed to produce the sections, and section-lining (sometimes called cross-hatching), which appears in the section view on all portions that have been in contact with the cutting plane. When only a part of the object is to be shown in section, conventional representation such as revolved, rotated, or broken-out section is used

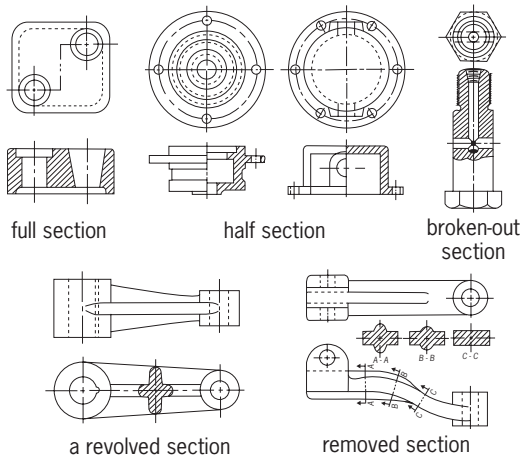

Fig. 1. Section views. (*After T. E. French and C. J. Vierck, Engineering Drawing, McGraw-Hill, 4th ed., 1978*)

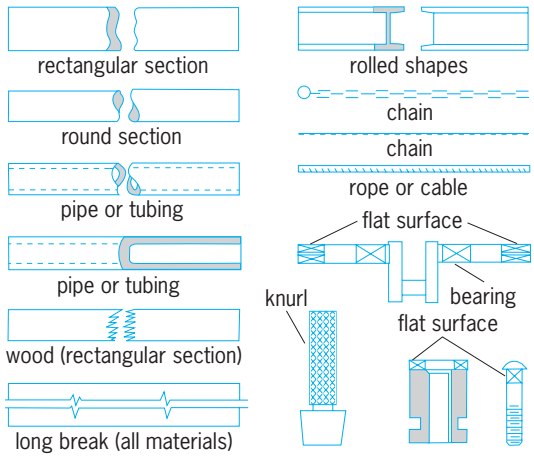

Fig. 2. Conventional breaks and other symbols to indicate details. (*After T. E. French and C. J. Vierck, Engineering Drawing, McGraw-Hill, 4th ed.,1978*)

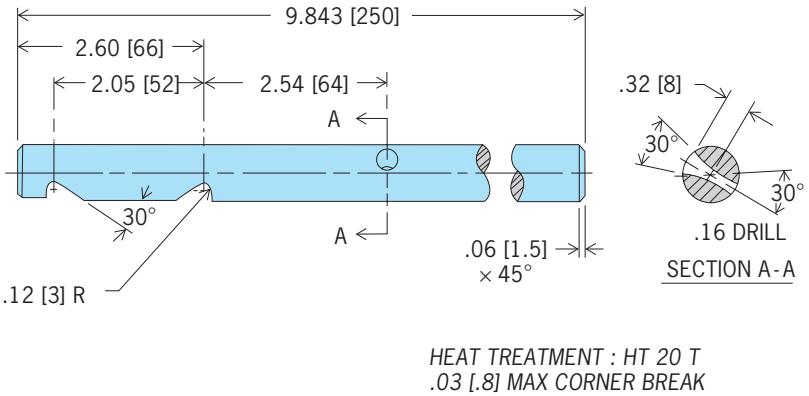

Fig. 3. Machine part which has been dimensioned in inches and millimeters (brackets).

(**Fig. 1**). Details such as flat surfaces, knurls, and threads are treated conventionally, which facilitates the making and reading of engineering drawings by experienced personnel (**Fig. 2**). Thus, certain engineering drawings will be combinations of top and front views, section and rotated views, and partial or pictorial views.

Dimensioning. In addition to describing the shape of objects, many drawings must show dimensions, so that workers can build the structure or fabricate parts that will fit together. This is accomplished by placing the required values (measurements) along dimension lines (usually outside the outlines of the object) and by giving additional information in the form of notes which are referenced to the parts in question by angled lines called leaders. In drawings of large structures the major unit is the foot, and in drawings of small objects the unit is the inch. A drawing containing dual dimensioning, inches and millimeters (in brackets), is shown in **Fig. 3**. In metric dimensioning, the basis unit may be the meter, the centimeter, or the millimeter, depending upon the size of the object or structure.

Working types of drawings may differ in styles of dimensioning, lettering (inclined lowercase, vertical uppercase, and so on), positioning of the numbers (aligned, or unidirectional—a style in which all numbers are lettered horizontally), and in the type of fraction used (common fractions or decimal fractions). If special precision is required, an upper and lower allowable limit are shown. Such tolerance, or limit, dimensioning is necessary for the manufacture of interchangeable mating parts, but unnecessarily close tolerances are very expensive.

Layout drawing. Layout drawings of different types are used in different manufacturing fields for various purposes. One is the plant layout drawing, in which the outline of the building, work areas, aisles, and individual items of equipment are all drawn to scale (**Fig. 4**). Another type is the aircraft, or master, layout, which is drawn on glass cloth or on steel or aluminum sheets. The object is drawn to full size with extreme accuracy. The completed drawing is photographed with great precision, and a glass negative made. From this negative, photo templates are made on photosensitized metal in various sizes and for

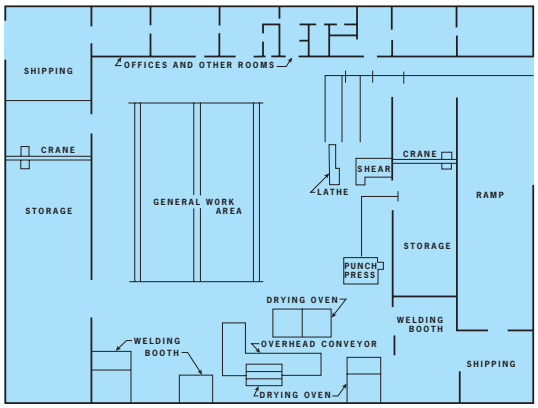

Fig. 4. A plant-layout drawing. (*After F. Zozzora, Engineering Drawing, McGraw-HIll, 2d ed., 1958*)

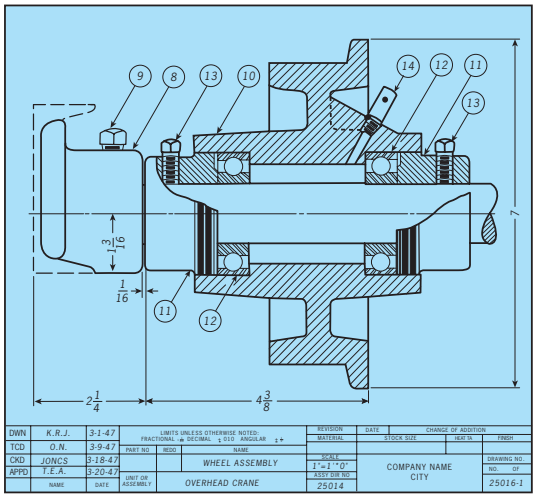

Fig. 5. A unit, or general, assembly. (*After T. E. French and C. J. Vierck, Engineering Drawing, McGraw-Hill, 4th ed.,1978*)

different purposes, thereby eliminating the need for many conventional detail drawings. Another type of layout, or preliminary assembly, drawing is the design layout, which establishes the position and clearance of parts of an assembly.

Assembly drawings. A set of working drawings usually includes detail drawings of all parts and an assembly drawing of the complete unit. Assembly drawings vary somewhat in character according to their use, as design assemblies or layouts; working drawing assemblies; general assemblies; installation assemblies; and check assemblies. A typical general assembly may include judicious use of sectioning and identification of each part with a numbered balloon (**Fig. 5**). Accompanying such a drawing is a parts list, in which each part is listed by number and briefly described; the number of pieces required is stated and other pertinent information given. Parts lists are best placed on separate sheets and typewritten to avoid time-consuming and costly hand lettering.

Schematic drawings. Schematic or diagrammatic drawings make use of standard symbols which indicate the direction of flow. In piping and electrical schematic diagrams, symbols recommended by the American National Standards Institute (ANSI), other agencies, or the Department of Defense (DOD) are used. In contrast to **Fig. 6**, the fixtures or components are not labeled in most schematics because the readers usually know what the symbols represent.

Additional information is often lettered on schematic drawings, for example, the identification of each replaceable electrical component. Etched-circuit drawing has revolutionized the wiring of electronic components. By means of such drawing, the wiring of an electronic circuit is photographed on a copper-clad board, and unwanted areas are etched away. On electrical and other types of flow diagrams,

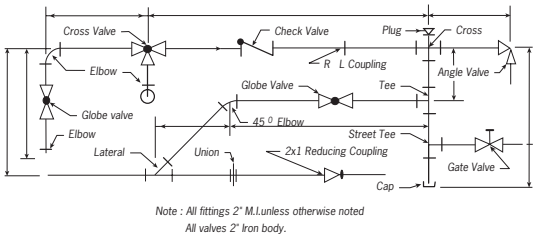

Fig. 6. Piping diagrammatic drawing. (*After T. E. French and C. J. Vierck, Engineering Drawing, McGraw-Hill, 1953*)

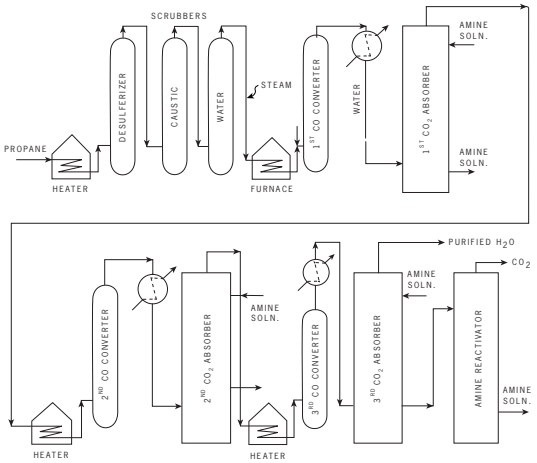

Fig. 7. Chemical engineering flow diagram. (*After P. H. Groggins, ed., Unit Processes in Organic Synthesis, 5th ed., McGraw-Hill, 1958*)

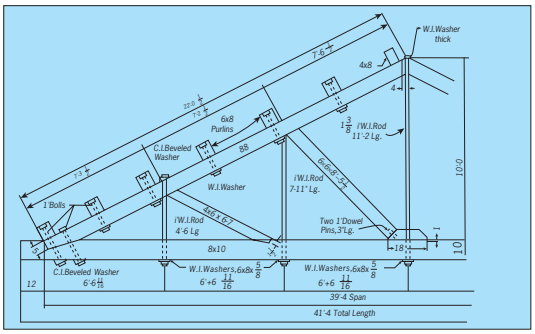

Fig. 8. Structural drawing of wood truss. A drawing of a steel truss would look similar, but would include symbols for such structural shapes as angles and channels. (*After T. E. French and C. J. Vierck, Engineering Drawing, McGraw-Hill, 4th ed., 1978*)

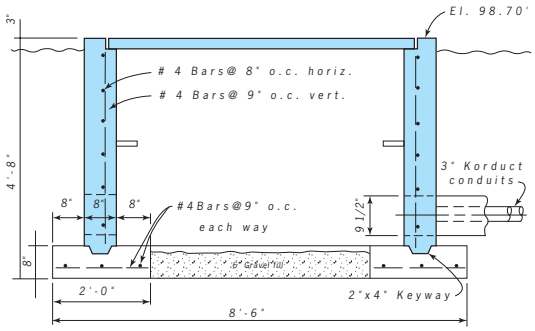

Fig. 9. Drawing of a reinforced concrete structure.

all single lines (often with arrows showing direction of flow) are drawn horizontally or vertically (**Fig. 7**); there are few exceptions. In some flow diagrams, rectangular enclosures are used for all items. Lettering is usually placed within the enclosures. *See* SCHEMATIC DRAWING; WIRING DIAGRAM.

Structural drawings. Structural drawings include design and working drawings for structures such as building, bridges, dams, tanks, and highways (**Figs. 8** and **9**). Such drawings form the basis of legal contracts. Structural drawings embody the same principles as do other engineering drawings, but use terminology and dimensioning techiques different from thoses shown in previous illustrations. *See* DRAFTING; GRAPHIC METHODS; NOMOGRAPH.

Charles J. Baer

Bibliography. C. J. Baer and J. R. Ottaway, *Electrical and Electronics Drawing*, 5th ed., 1986; T. E. French et al., *Mechanical Drawing*, 14th ed., 1993; T. E. French and C. J. Vierck, *Engineering Drawing and Graphic Technology*, 14th ed., 1993; C. Jensen and J. Helsel, *Engineering Drawing and Design*, 5th ed., 1997; J. R. Wirshing and R. H. Wirshing, *Civil Engineering Drafting*, 1983.

Engineering geology

The application of education and experience in geology and other geosciences to solve geological problems posed by civil engineering works. The branches of the geosciences most applicable are surficial geology, petrofabrics, rock and soil mechanics, geohydrology, and geophysics, particularly exploration geophysics and earthquake seismology. This article discusses some of the practical aspects of engineering geology.

Terminology. The terms engineering geology and environmental geology often seem to be used interchangeably. Specifically, environmental geology is the application of engineering geology in the solution of urban problems; in the prediction and mitigation of natural hazards such as earthquakes, landslides, and subsidence; and in solving problems inherent in disposal of dangerous wastes and in reclaiming mined lands.

Another relevant term is geotechnics, the combination of pertinent geoscience elements with civil engineering elements to formulate the civil engineering system that has the optimal interaction with the natural environment.

Engineering properties of rock. The civil engineer and the engineering geologist consider most hard and compact natural materials of the earth crust as rock, and their derivatives, formed mostly by weathering processes, as soil. A number of useful soil classification systems exist. Because of the lack of a rock classification system suitable for civil engineering purposes, most engineering geology reports use generic classification systems modified by appropriate rock-property adjectives. *See* ROCK; ROCK MECHANICS; SOIL MECHANICS.

Rock sampling. The properties of a rock element can be determined by tests on cores obtained from boreholes. These holes are made by one or a combination of the following basic types of drills: the rotary or core drill, the cable-tool or churn drill, and the auger. The rotary type generally is used to obtain rock cores. The rotary rig has a motor or engine (gasoline, diesel, electric, or compressed air) that drives a drill head that rotates a drill rod (a thick-walled hollow pipe) fastened to a core barrel with a bit at its end. Downward pressure on the bit is created by hydraulic pressure in the drill head. Water or air is used to remove the rock that is comminuted (chipped or ground) by the diamonds or hard-metal alloy used to face the bit. The core barrel may be in one piece or have one or two inner metal tubes to facilitate recovery of soft or badly broken rock (double-tube and triple-tube core barrels). The churn-type drill may be used to extend the hole through the soil overlying the rock, to chop through boulders, occasionally to deepen a hole in rock when core is not required or to obtain drive samples of the overburden soils. When the rock is too broken to support itself, casing (steel pipe) is driven or drilled through the broken zone. Drill rigs range in size from those mounted on the rear of large multiwheel trucks to small, portable ones that can be packed to the investigation site on a person's back or parachuted from a small plane. *See* DRILLING, GEOTECHNICAL.

The rock properties most useful to the engineering geologist are compressive and triaxial shear strengths, permeability, Young's modulus of elasticity, erodability under water action, and density (in pounds per cubic foot, or pcf).

Compressive strength. The compressive (crushing) strength of rock generally is measured in pounds per square inch or kilograms per square centimeter. It is the amount of stress required to fracture a sample unconfined on the sides and loaded on the ends (**Fig. 1**). If the load P of 40,000 lb is applied to a sample with a diameter of 2 in. (3.14 in.^{2}), the compressive stress is $40{,}000 \div 3.14 = 12{,}738$ lb/in.^{2} ($177{,}920$ N $\div$ 0.00203 m^2 = 87,645 kN/m^2). If this load breaks the sample, the ultimate compressive strength equals the compressive stress acting at the moment of failure, in this case 12,738 lb/in.^{2}. The test samples generally are cylindrical rock cores that have a length-to-diameter ratio (L/D) of about 2. The wide variety of classification systems used for rock results in a wide variation in compressive strengths for rocks

Compressive strength of rocks*

Type of rock	No. of tests	Weighted mean, kN/m²†	Minimum and maximum values, kN/m²†
Igneous			
Basalt	195	112,390	4,140–383,360
Diabase	6	150,655	4,140–275,100
Diorite	26	167,550	78,600–310,275
Gabbro	10	245,875	110,000–399,220
Granite	140	160,515	17,925–332,340
Rhyolite	44	162,100	11,720–453,690
Sedimentary			
Dolomite	62	87,840	6,205–358,540
Limestone	211	74,190	1,380–260,630
Sandstone	257	63,430	2,070–328,200
Shale	67	66,605	690–230,980
Siltstone	14	108,525	3,445–315,790
Metamorphic			
Amphibolite	14	152,240	24,820–280,625
Gneiss	103	133,760	35,850–292,350
Marble	31	100,600	30,340–273,730
Quartzite	25	292,350	25,510–628,825
Schist	16	50,265	6,895–162,030

*Based on data in W. R. Judd, *Statistical Relationships for Certain Rock Properties,* DDC AD735376, 1971, and *Statistical Methods to Compile and Correlate Rock Properties,* Purdue Res. Tech. Rep. 2, 1969.
†1 kN/m² = 0.1450 lb/in.²

having the same geologic name. The **table** gives a statistical evaluation of the compressive strengths of several rocks commonly encountered in engineering geology.

Most laboratory tests show that an increase in moisture in rock causes a decrease in its compressive strength and elastic modulus; what is not generally known, however, is that the reverse situation shown in **Fig. 2** has been encountered in certain types of volcanic rocks. In sedimentary rock the compressive strength is strongly dependent upon the quality of the cement that bonds the mineral grains together (for example, clay cement gives low strength) and upon the quantity of cement (a rock may have only a small amount of cement, and despite a strong bond between the grains, the strength is directly related to the inherent strength of the grains). Strength test results are adversely affected by microfractures that may be present in the sample prior to testing, particularly if the microfractures are oriented parallel to the potential failure planes.

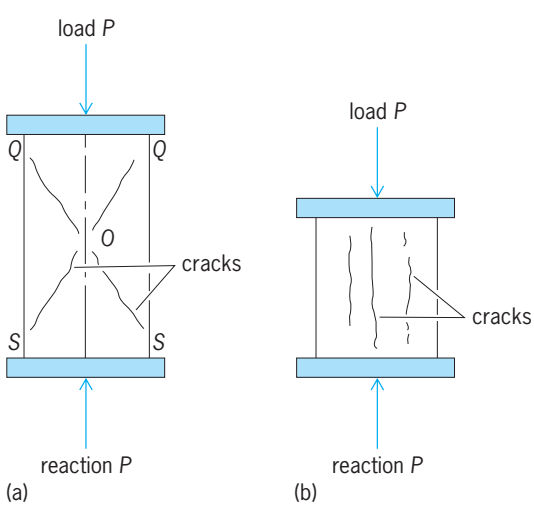

Fig. 1. Unconfined compression test. (*a*) Shear failure, showing failure planes *QS*. (*b*) Tension failure. (*After D. P. Krynine and W. R. Judd, Principles of Engineering Geology and Geotechnics, McGraw-Hill, 1957*)

The value of compressive strength to be used in an engineering design must be related to the direction of the structure's load and the orientation of the bedding, discontinuities, and structural weaknesses in the foundation rock. This relationship is important because the highest compressive strength usually is obtained when the compressive stress is normal to the bedding. Conversely, the highest Young's modulus of elasticity (E) usually results when the compressive stress parallels the bedding. When these strength and elastic properties apparently are not affected by the direction of applied load, the rock is described as isotropic; if load applied parallel to the bedding provides physical property data that are significantly different than those obtained when the load is applied normal (perpendicular) to the bedding, the rock is anisotropic or aeolotropic. If the physical components of the rock element or rock system have equal dimensions and equal fabric relationships, the rock is homogeneous; significant variance in these relationships results in a heterogeneous rock. Most rocks encountered in foundations and underground works are anisotropic and heterogeneous.

Shear in rocks. Shearing stresses tend to separate portions of the rock (or soil) mass. Faults and folds are examples of shear failures in nature. In engineering structures, every compression is accompanied by shear stresses. For example, an arch dam compresses the abutment rock and, if the latter is intersected by fissures or weak zones, it may fail in shear with a resulting tensile stress in the dam concrete that may rupture the concrete. The application of loads over long periods of time on most rocks will cause them to creep or even to flow like a dense fluid (plastic flow). *See* STRUCTURAL GEOLOGY.

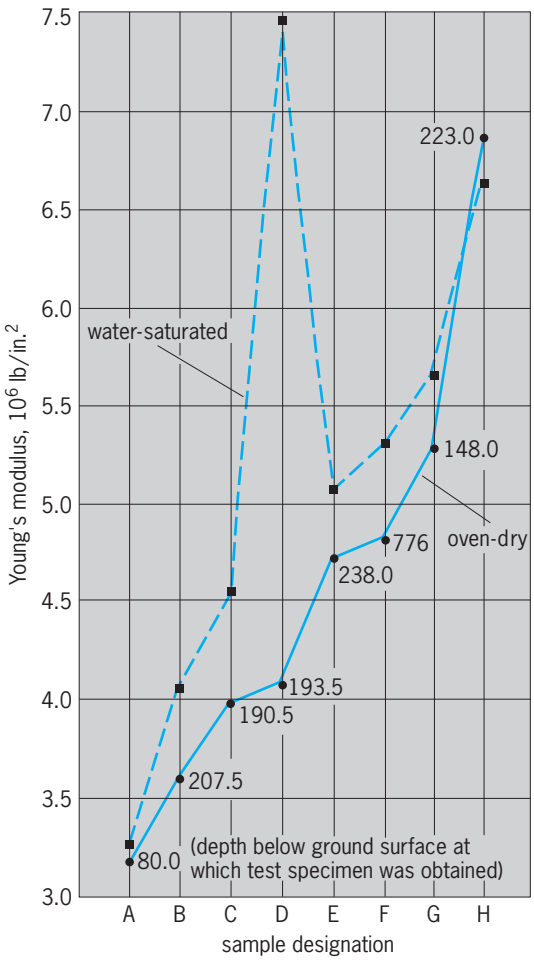

Fig. 2. Increase in Young's modulus caused by saturation of dacite porphyry. (*After J. R. Ege and R. B. Johnson, Consolidated tables of physical properties of rock samples from area 401, Nevada Test Site, U.S. Geol. Surv. Tech. Letter Pluto-21, 1962*)

Ambient stress. This type of stress in a rock system is actually potential energy, probably created by ancient natural forces, recent seismic activity, or nearby human-caused disturbances. Ambient (residual, stored, or primary) stress may remain in rock long after the disturbance is removed. An excavation, such as a tunnel or quarry, will relieve the ambient stress by providing room for displacement of the rock, and thus the potential energy is converted to kinetic energy. In tunnels and quarries, the release of this energy can cause spalling, the slow outward separation of rock slabs from the rock massif; when this movement is rapid or explosive, a rock burst occurs. The latter is a different phenomenon from a rock bump, which is a rapid upward movement of a large portion of a rock system and, in a tunnel, can have sufficient force to flatten a steel mine car against the roof or break the legs of a person standing on the floor when the bump occurs. *See* ROCK BURST; UNDERGROUND MINING.

One of two fundamental principles generally is used to predict the possible rock load on a tunnel roof, steel, or timber supports, or a concrete or steel lining: (1) The weight of the burden (the rock and soil mass between the roof and the ground surface) and its shear strength control the load, and therefore the resultant stresses are depth-dependent, or (2) the shear strength of the rock system and the ambient stresses control the stress distribution, so the resultant loading is only indirectly dependent upon depth. The excavation process can cause rapid redistribution of these stresses to produce high loads upon supports some distance from the newly excavated face in the tunnel. The geometry and span of the opening also influence the stress distribution. Lined tunnels can be designed so that the reinforced concrete or steel lining will have to carry only a portion of the ambient or burden stresses. *See* TUNNEL.

Construction material. Rock as a construction material is used in the form of dimension, crushed, or broken stone. Broken stone is placed as riprap on slopes of earth dams, canals, and river banks to protect them against water action. Also, it is used as the core and armor stone for breakwater structures. For all such uses, the stone should have high density ($\pm$165 lb/ft^3 or 2650 kg/m^3), be insoluble in water, and be relatively nonporous to resist cavitation. Dimension stone (granite, limestone, sandstone, and some basalts) is quarried and sawed into blocks of a shape and size suitable for facing buildings or for interior decorative panels. For exterior use, dimension stone preferably should be isotropic (in physical properties), have a low coefficient of expansion when subjected to temperature changes, and be resistant to deleterious chemicals in the atmosphere (such as sulfuric acid). Crushed stone (primarily limestone but also some basalt, granite, sandstone, and quartzite) is used as aggregate in concrete and in bituminous surfaces for highways, as a base course or embankment material for highways, and for railroad ballast (to support the ties). When used in highway construction, the crushed stone should be resistant to abrasion as fine stone dust reduces the permeability of the stone layer; the roadway then is more susceptible to settling and heaving caused by freezing and thawing of water in the embankment. Concrete aggregate must be free of deleterious material such as opal and chalcedony; volcanic rocks containing glass, devitrified glass, and tridymite; quartz with microfractures; phyllites containing hydromica (illites); and other rocks containing free silica (SiO_2). These materials will react chemically with the cement in concrete and release sodium and potassium oxides (alkalies) or silica gels. Preliminary petrographic analyses of the aggregate and chemical analysis of the cement can indicate the possibility of alkali reactions and thus prevent construction difficulties such as expansion, cracking, or a strength decrease of the concrete. *See* CAVITATION; CONCRETE; PETROGRAPHY; STONE AND STONE PRODUCTS.

Geotechnical significance of soils. Glacial and alluvial deposits contain heterogeneous mixtures of pervious (sand and gravel) and impervious (clay, silt, and rock flour) soil materials. The pervious materials can be used for highway subgrade, concrete

aggregate, and filters and pervious zones in earth embankments. Dam reservoirs may be endangered by the presence of stratified or lenticular bodies of pervious materials or ancient buried river channels filled with pervious material. Deep alluvial deposits in or close to river deltas may contain very soft materials such as organic silt or mud. An unsuitable soil that has been found in dam foundations is openwork gravel. This material may have a good bearing strength because of a natural cement bond between grains, but it is highly pervious because of the almost complete lack of fine soil to fill the voids between the gravel pebbles. *See* DELTA; FLOODPLAIN.

Concrete or earth dams can be built safely on sand foundations if the latter receive special treatment. One requirement is to minimize seepage losses by the construction of cutoff walls (of concrete, compacted clay, or interlocking-steel-sheet piling) or by use of mixed-in-place piles 3 ft (0.9 m) or more in diameter. The latter are constructed by augering to the required depth but not removing any of the sand. At the desired depth, cement grout is pumped through the hollow stem of the auger, which is slowly withdrawn while still rotating; this mixes the grout and the sand into a relatively impervious concrete pile. The cutoff is created by overlapping these augered holes. Some sand foundations may incur excessive consolidation when loaded and then saturated, particularly if there is a vibratory load from heavy machinery or high-velocity water in a spillway. This problem is minimized prior to loading by using a vibrating probe inserted into the sand or vibratory rollers on the sand surface or by removing the sand and then replacing it under vibratory compaction and water sluicing.

Aeolian (windblown) deposits. Loess is a relatively low-density ($\pm$0.044 ton/ft^3 or 1.4 metric tons/m^3) soil composed primarily of silt grains cemented by clay or calcium carbonate. It has a vertical permeability considerably greater than the horizontal. When a loaded loess deposit is wetted, it rapidly consolidates, and the overlaying structure settles. When permanent open excavations ("cuts") are required for highways or canals through loess, the sides of the cut should be as near vertical as possible: Sloping cuts in loesses will rapidly erode and slide because of the high vertical permeability. To avoid undesirable settlement of earth embankments, the loess is "prewetted" prior to construction by building ponds on the foundation surface. Permanently dry loess is a relatively strong bearing material. Aeolian sand deposits present the problem of stabilization for the continually moving sand. This can be done by planting such vegetation as heather or young pine or by treating it with crude oil. Cuts are traps for moving sand and should be avoided. The failure of Teton Dam in 1976 indicated, among other factors, that when loessial or silty soils are used for core materials in dam embankments, it is important to take special measures to prevent piping of the silts by carefully controlled compaction of the core, by using up- and downstream filters, and by extraordinary treatment of the foundation rock. *See* LOESS.

Organic deposits. Excessive settlement will occur in structures founded on muskeg terrain. Embankments can be stabilized by good drainage, the avoidance of cuts, and the removal of the organic soil and replacement by sand and gravel or, when removal is uneconomical, displacement of it by the continuous dumping of embankment material upon it. Structures imposing concentrated loads are supported by piling driven through the soft layers into layers with sufficient bearing power. *See* MUSKEG; TUNDRA.

Residual soils. These soils are derived from the in-place deterioration of the underlying bedrock. The granular material caused by the in-place disintegration of granite generally is sufficiently thin to cause only nominal problems. However, there are regions (such as California, Australia, and Brazil) where the disintegrated granite (locally termed DG) may be hundreds of feet thick; although it may be competent to support moderate loads, it is unstable in open excavations and is pervious. A thickness of about 200 ft (60 m) of DG and weathered gneiss on the sides of a narrow canyon was a major cause for construction of the Tumut-1 Power Plant (New South Wales) in hard rock some 1200 ft (365 m) underground. Laterite (a red clayey soil) derived from the in-place disintegration of limestone in tropical to semitropical climates is another critical residual soil. It is unstable in open cuts on moderately steep slopes, is compressible under load, and when wet produces a slick surface that is unsatisfactory for vehicular traffic. This soil frequently is encountered in the southeastern United States, southeastern Asia, and South America.

Clays supporting structures may consolidate slowly over a long period of time and cause structural damage. When clay containing montmorillonite is constantly dried and rewetted by climatic or drainage processes, it alternately contracts and expands. During the drying cycle, extensive networks of fissures are formed that facilitate the rapid introduction of water during a rainfall. This cyclic volume change of the clay can produce uplift forces on structures placed upon the clay or compressive and uplift forces on walls of structures placed within the clay. These forces have been known to rupture concrete walls containing 10.75-in.-diameter (19-mm) steel reinforcement bars. A thixotropic or "quick" clay has a unique lattice structure that causes the clay to become fluid when subjected to vibratory forces. Various techniques are used to improve the foundation characteristics of critical types of clay: (1) electroosmosis that uses electricity to force redistribution of water molecules and subsequent hardening of the clay around the anodes inserted in the foundation; (2) provision of adequate space beneath a foundation slab or beam so the clay can expand upward and not lift the structure; (3) belling, or increasing in size, of the diameter of the lower end of concrete piling so the pile will withstand uplift forces imposed by clay layers around the upper part of the pile; (4) treatment of the pile surface with a frictionless coating (such as poly(tetrafluoroethylene) or a loose wrapping of asphalt-impregnated paper) so the upward-moving

clay cannot adhere to the pile; (5) sufficient drainage around the structure to prevent moisture from contacting the clay; and (6) replacement of the clay by a satisfactory foundation material. Where none of these solutions is feasible, the structure then must be relocated to a satisfactory site or designed so it can withstand uplift or compressive forces without extensive damage. *See* CLAY.

Silt may settle rapidly under a load or offer a "quick" condition when saturated. For supporting some structures (such as residences), the bearing capacity of silts and fine sands can be improved by intermixing them with certain chemicals that will cause the mixture to "set" or harden when exposed to air or moisture; some of the chemicals used are sodium silicate with the later application of calcium chloride, bituminous compounds, phenolic resins, or special cements (to form "soil cement"). The last mixture has been used for surfacing secondary roads, for jungle runways in Vietnam, and as a substitute for riprap of earth dams. Some types of silt foundations can be improved by pumping into them soil-cement or clay mixtures under sufficient pressures to create large bulbs of compacted silt around the pumped area.

Geotechnical investigation. For engineering projects, these investigations may include preliminary studies, preconstruction or design investigations, consultation during construction, and the maintenance of the completed structure.

Preliminary studies. These are made to select the best location for a project and to aid in formulating the preliminary designs for the structures. The first step in the study is a search for pertinent published material in libraries, state and federal agencies, private companies, and university theses. Regional, and occasionally detailed reports on local geology, including geologic maps, are available in publications of the U.S. Geological Survey; topographic maps are available from that agency and from the U.S. Army Map Service. Oil companies occasionally will release the geologic logs of any drilling they may have done in a project area. Air photos and other remote sensing techniques such as pulsed or side-looking radar or false color can be used to supplement map information (or may be the only surficial information readily available). The U.S. Geological Survey maintains a current index map of the air-photo coverage of the United States. The photos are available from that agency, the U.S. Forest Service, the Natural Resources Conservation Service, and commercial air-photo companies; for some projects, the military agencies will provide air-photo coverage. The topographic maps and air photos can be used to study rock outcrop and drainage patterns, landforms, geologic structures, the nature of soil and vegetation, moisture conditions, and land use by humans (cultural features). Airborne geophysical techniques using magnetometers or gravimeters also may be useful to delineate surface and subsurface geologic conditions. *See* AERIAL PHOTOGRAPHY; LITERATURE OF SCIENCE AND TECHNOLOGY; REMOTE SENSING; TOPOGRAPHIC SURVEYING AND MAPPING.

Field reconnaissance may include the collection of rock and soil specimens; inspection of road cuts and other excavations; inspection of the condition of nearby engineering structures such as bridges, pavements, and buildings; and location of sources of construction material. Aerial reconnaissance is essential at this stage and can be performed best in helicopters and second-best in slow-flying small planes.

Preconstruction. Surface and subsurface investigations are required prior to design and construction. Surface studies include the preparation of a detailed map of surficial geology, hydrologic features, and well-defined landforms. For dam projects, a small-scale geologic map (for example, 1:5000) is made of the reservoir area and any adjacent areas that may be directly influenced by the project; in addition, a large-scale geologic map (for example, 1:500) is required of the specific sites of the main structures (the dam, spillway, power plant, tunnels, and so on). [This preferred means of designating map scales can be used for either customary or metric units. It means 1 unit of measurement on the map is equal to 5000 similar units on the ground; for example, 1 cm measured on the map is equal to 5000 cm measured on the ground.] These maps can be compiled by a combination of field survey methods and aerial mapping procedures. They should have a grid system (coordinates) and show the proposed locations for subsurface investigations.

Subsurface investigations are required to confirm and amplify the surficial geologic data. These may include test pits, trenches, short tunnels (drifts or adits), and the drilling of vertical, horizontal, or oblique (angle) boreholes. Geologic data obtained by these direct methods can be supplemented by indirect or interpreted data obtained by geophysical methods on the surface or in subsurface holes and by installation of special instruments to measure strain or deformation in a borehole or tunnel.

The geology disclosed by subsurface investigations is "logged" on appropriate forms. Tunnel logs display visual measurements of features and joint orientations (strike and dip); rock names and a description of their estimated engineering properties; alteration, layering, and other geologic defects; the location and amount of water or gas inflow; the size and shape of blocks caused by fracturing or jointing and the width of separation or the filling material between blocks; and the irregularities in the shape of the tunnel caused by the displacement of blocks during or after excavation (rock falls, rock bursts, chimneying, and overbreakage). Geophysical seismic methods may be used to define the thickness of loosened rock around the tunnel; geoacoustical techniques that detect increases in microseismic noise during tunneling may be used to determine if the excavation is causing excessive loosening in the tunnel rock. This detection of "subaudible rock noise" occasionally is used to detect the potential movement of rock slopes in open excavations.

The borehole data can be logged on a form such as shown in **Fig. 3**. These data can be obtained by direct examination of the core, by visual inspection of

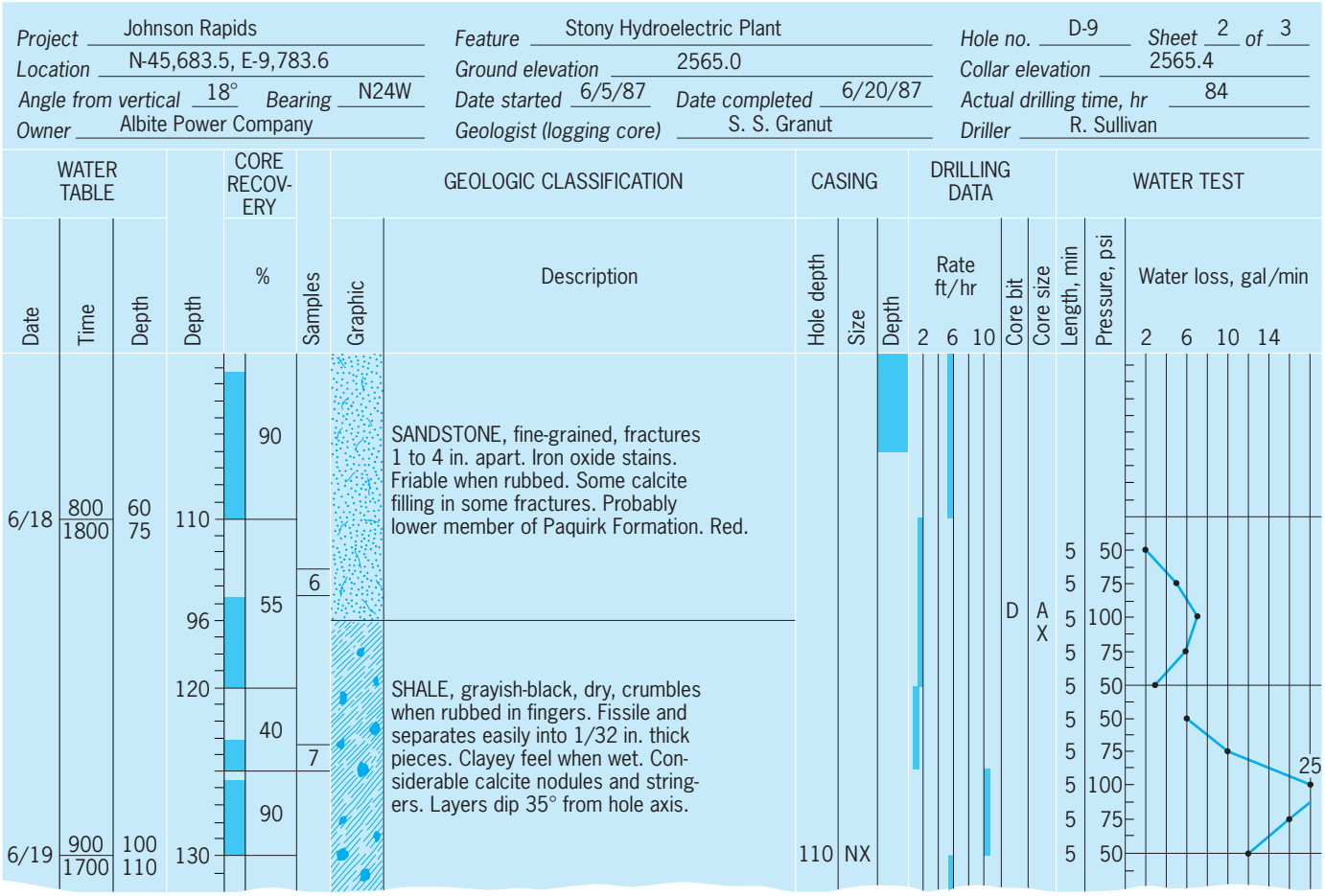

Fig. 3. Log recording of geological information from a borehole.

the interior of the borehole using a borehole camera (a specially made television camera) or a stratoscope (a periscopelike device), or by geophysical techniques. Direct viewing of the interior of the hole is the only positive method of determining the in-place orientation and characteristics of separations and of layering in the rock system. The geophysical techniques include use of gamma-gamma logging that evaluates the density of the rock surrounding the borehole or at depths as great as 150 ft (45 m) beneath the gamma probe; neutron logging to determine the moisture content of the rock system by measuring the depth of penetration of the neutrons; traversing the borehole with a sonic logger that, by calibration, measures differences in the velocity of wave propagation in different strata (and thus can determine in place Young's modulus of elasticity and the thickness of each stratum encountered by the borehole); and electric logging that uses differences in the electrical resistivity of different strata to define their porosity, moisture content, and thickness.

Occasionally a hole is drilled through a talus deposit containing the same type of rock as the underlying rock in place (bedrock). Because of the similarity in rock types, the talus-bedrock contact sometimes is best identified by determining the orientation of the remnant magnetism in the core: the magnetic lines in the core will have a regular orientation, but the talus magnetism will have random directions. This method is useful only in rocks that contain appreciable remnant magnetism such as some basalts.

Geophysical seismic or electrical resistivity methods also can be used on the ground surface to define the approximate depth of bedrock or various rock layers. The results require verification by occasional boreholes, but this is an inexpensive and satisfactory technique for planning and design investigations. The seismic methods are not useful when it is necessary to locate soft strata (wherein the seismic waves travel at relatively low velocity) that are overlain by hard strata (that have higher wave velocity); the latter conceal and block the signal from the soft strata. Also, difficulties may occur when the strata to be located are overlain by soil containing numerous large boulders composed of rock having higher velocities than the surrounding soil, or when the soil is very compact (such as glacial till) because its velocity characteristics may resemble those of the underlying bedrock. Another problem is that the seismic method seldom can identify narrow and steep declivities in the underlying hard rock (because of improper reflection of the waves).

Construction. Geotechnical supervision is desirable during construction in or on earth media. The engineering geologist must give advice and keep a record of all geotechnical difficulties encountered

during the construction and of all geological features disclosed by excavations. During the operation and maintenance of a completed project the services of the engineering geologist often are required to determine causes and assist in the preparation of corrective measures for cracks in linings of water tunnels, excessive settlement of structures, undesirable seepage in the foundations of dams, slides in canal and other open excavations, overturning of steel transmission-line towers owing to a foundation failure, and rock falls onto hydroelectric power plants at the base of steep canyon walls. The engineering geologist also is considered an important member of the team assigned to the task of assessing the safety of existing dams as now required by federal legislation.

Legal aspects. An important consideration for the engineering geologist is the possibility of a contractor making legal claims for damages, purportedly because of unforeseen geologic conditions (generally referred to legally as charged conditions) encountered during construction. Legal support for such claims can be diminished if the engineering geologist supplies accurate and detailed geologic information in the specification and drawings used for bidding purposes. These documents should not contain assumptions about the geological conditions (for a proposed structure), but they should show all tangible geologic data obtained during the investigation for the project: for example, an accurate log of all boreholes and drifts and a drawing showing the boundaries of the outcrops of all geological formations in the project area. The engineering geologist should have sufficient experience with design and construction procedures to formulate an investigation program that results in a minimum of subsequent uncertainties by a contractor. Numerous uncertainties about the geologic conditions not only can result in increased claims but also may cause a contractor to submit a higher bid (in order to minimize risks) than if detailed geologic information were available.

Special geotechnical problems. In arctic zones, structures built on permafrost may be heaved or may cause thawing and subsequent disastrous settlement. The growth of permafrost upward into earth dams seriously affects their stability and permeability characteristics. Obtaining natural construction materials in permafrosted areas requires thawing of the borrow area to permit efficient excavation; once excavated, the material must be protected against refreezing prior to placement in the structures. Permafrost in rock seldom will cause foundation difficulties. In planning reservoirs, it is essential to evaluate their watertightness, particularly in areas containing carbonate or sulfate rock formation or lava flows. These formations frequently contain extensive systems of caverns and channels that may or may not be filled with claylike material or water. Where extensive openings occur, grouting with cement slurry or chemicals can be used as a sealant; however, as demonstrated by the 1976 failure of the Teton Dam, such measures are not always successful. Sedimentation studies are required for the design of efficient harbors or reservoirs because soil carried by moving water will settle and block or fill these structures. In areas with known earthquake activity, structural design requires knowledge of the intensity and magnitude of earthquake forces. The prevention and rehabilitation of slides (landslides) in steep natural slopes and in excavations are important considerations in many construction projects and are particularly important in planning reservoirs, as was disastrously proved by the Vaiont Dam catastrophe in 1963. *See* EARTHQUAKE ENGINEERING; PERMAFROST.

Geohydrologic problems. In the foundation material under a structure, water can occur in the form of pore water locked into the interstices or pores of the soil or rock, as free water that is moving through openings in the earth media, or as included water that is a constituent or chemically bound part of the soil or rock. When the structure load compresses the foundation material, the resulting compressive forces on the pore water can produce undesirable uplift pressures on the base of the structure. Free water is indicative of the permeability of the foundation material and possible excessive water loss (from a reservoir, canal, or tunnel); uplift on the structure because of an increase in hydrostatic head (caused by a reservoir or the like); or piping, which is the removal of particles of the natural material by flowing water with a consequent unfilled opening that weakens the foundation and increases seepage losses.

The possibility of excessive seepage or piping can be learned by appropriate tests during the boring program. For example, water pressure can be placed on each 5-ft (2-m) section of a borehole, after the core is removed, and any resulting water loss can be measured. The water pressure is maintained within the 5-ft section by placing an expandable rubber ring (packer) around the drill pipe at the top of the test section and then sealing off the section by using mechanical or hydraulic pressure on the pipe to force expansion of the packer. When only one packer is used, because it is desired to test only the section of hole beneath it, it is a "single-packer" test. In a double-packer test, a segment of hole is isolated for pressure testing by placing packers at the top and the bottom of the test section. The best information on the permeability characteristics of the rock can be obtained by the use of three or more increments of increasing and then decreasing water pressure for each tested length of hole. If the water loss continues to increase when the pressure is decreased, piping of the rock or filling material in fractures may be occurring or fractures are widening or forming. The water-pressure test can be supplemented by a groutability test in the same borehole. This test is performed in the same way as the water test except, instead of water, a mixture of cement, sand, and water (cement grout) or a phenolic resin (chemical grout) is pumped under pressure into the test section. The resulting information is used to design cutoff walls and grout curtains for dams. The pressures used in water-pressure or grouting tests should not exceed the pressure exerted by the weight of the burden between the ground surface and the top of the test section. Excessive test pressure can cause

uplift in the rock, and the resulting test data will be misleading.

Included or pore water generally is determined by laboratory tests on cores; these are shipped from the borehole to the laboratory in relatively impervious containers that resist loss of moisture from the core. The cores with their natural moisture content are weighed when received and then dried in a vacuum oven at about 110°F (45°C) until their dry weight stabilizes. The percentage of pore water (by dry weight) is (wet weight − dry weight) × 100 ÷ dry weight.

Temperatures up to 200°F (about 90°C) can be used for more rapid drying, provided the dried specimens are not to be used for strength or elastic property determinations. (High temperatures can significantly affect the strength because the heat apparently causes internal stresses that disturb the rock fabric or change the chemical composition of the rock by evaporation of the included moisture.)

Protective construction. Civilian and military structures may be designed to minimize the effects of nuclear explosions. The most effective protection is to place the facility in a hardened underground excavation. A hardened facility, including the excavation and its contents, is able to withstand the effects generated by a specified size of nuclear weapon. These effects include the amount of displacement, acceleration, and particle velocity that occurs in the earth media and the adjacent structure. Desirable depths and configurations for hardened facilities are highly dependent upon the shock-wave characteristics of the surrounding earth media, for example, the type of rock, discontinuities in the rock system, free water, and geologic structure. Therefore, prior to the design and construction of such facilities, extensive geotechnical field and laboratory tests are performed, including an accurate geologic map of the surface and of the underground environment that will be affected by the explosion. The map should show the precise location and orientation of all geologic defects that would influence the wave path, such as joints, fractures, and layers of alternately hard and soft rock. *See* EXPLOSIVE; NUCLEAR EXPLOSION.

Application of nuclear energy. The use of nuclear energy for the efficient construction of civil engineering projects has been investigated in the Plowshare Program. Examples include rapid excavation, increasing production of natural gas by opening fractures in the reservoir rock, expediting production of low-grade copper ore by causing extensive fracturing and possible concentration of the ore, and by the underground "cracking" of oil shale. The production of electrical energy by nuclear fission requires engineering geology inputs during the planning and design of the power plant; for example, a major question to be answered is the presence or absence of faults and an estimate of when the last movement on the fault occurred. This question of "active" faults also is of increasing concern in the siting of dams.

Waste disposal. Another geotechnical problem occurs in the use of nuclear energy for generation of power or radioisotopes: safe disposal of the radioactive waste products. These products can be mixed with concrete and buried in the ground or ocean, but geohydrologic or oceanographic conditions must not be conducive to the deterioration of the concrete. One proposed solution is to excavate large caverns in rock or salt a thousand or more meters deep; however, such a solution must consider possible contamination of ground water in the event that the waste products' containers leak. The disposal of toxic chemical or biological products in deep wells no longer is considered safe. *See* RADIOACTIVE WASTE MANAGEMENT. William R. Judd

Bibliography. F. G. Bell, *Engineering Properties of Soil and Rocks*, 4th ed., 2000; B. M. Das, *Principles of Geotechnical Engineering*, 5th ed., 2001; R. E. Goodman, *Engineering Geology: Rock in Engineering Construction*, 1993; M. E. Harr, *Groundwater and Seepage*, 1992; R. B. Johnson and J. V. DeGraff, *Principles of Engineering Geology*, 1989; R. F. Leggett and A. W. Hatheway, *Geology and Engineering*, 3d ed., 1988; R. F. Legget and P. K. Karrow, *Handbook of Geology in Civil Engineering*, 3d ed., 1982; P. H. Rahn, *Engineering Geology: An Environmental Approach*, 2d ed., 1996; Q. Zaruba, *Landslides and Their Control*, 2d ed., 1982.

Enopla

A class of the phylum Rhynchocoela which is divided into the orders Hoplonemertini (free-living) and Bdellonemertini or Bdellomorpha (symbiotic). The proboscis in the Hoplonemertini is armed with stylets, and in both orders the mouth is anterior to the brain, the nervous system is inside the body wall musculature, and the vascular system is well developed.

The Bdellonemertini contain the single genus *Malacobdella* with four species, all of which are symbiotic with mollusks. *See* ANOPLA; BDELLONEMERTINI; HOPLONEMERTINI; NEMERTEA. J. B. Jennings

Enoplida

An order of nematodes characterized by having the cuticle of the cephalic region doubled or formed into a cap or helmet. The helmet results from the infolding of the cuticle over the extreme anterior end of the esophagus. The stoma may or may not possess armature in the form of teeth or movable jaws. Even though the stoma is surrounded by esophageal tissue, the anterior portion of the movable jaws is of cheilostomal origin. The well-developed esophagus is nearly cylindrical, and the esophagointestinal valve is prominent. The five esophageal glands are located in the posterior region of the esophagus; however, the orifices for these glands are located anterior to the nerve ring. The dorsal and two anterior subventral glands open at the base of the stoma or through the stomatal teeth. The amphid apertures may be transverse slits or elongate ovals. The cephalic sensilla are most often in two whorls: an anterior

circumoral whorl of 6 and a posterior whorl of 10 (6 + 4); in some taxa this whorl is clearly separated into the ancestral state of a circlet of 6 and one of 4. A single medioventral excretory cell is usually present; the orifice is generally at the level of or anterior to the nerve ring. Medioventral supplementary organs may or may not be present on males. The male spicules are paired and accompanied by a gubernaculum. Caudal glands and a cuticular spinneret are found in males and females.

There are two enoplid superfamilies, Enoploidea and Oxystominoidea. The Enoploidea are free-living marine nematodes comprising small to very large species in five well-defined families: Leptosomatidae, Lauratonematidae, Phanodermatidae, Thoracostomopsidae, and Enoplidae. While the feeding habits are largely unknown, it has been postulated that the group is predaceous or omnivorous. The group is almost exclusively marine in habitat, but a few species have been reported from brackish water or soils. Oxystominoidea contains species that maintain the most ancestral of characters known in the phylum. They are separated from other Enoplida by not having the lip region (=cephalic region) set off by a groove. *See* NEMATA (NEMATODA).

Armand R. Maggenti

Enstatite

A pyroxene close in composition to magnesium metasilicate ($MgSiO_3$). Like other pyroxenes, enstatite is typified by a pair of nearly perpendicular cleavages, both of which are parallel to the elongate dimension of the characteristically prismatic crystals. These cleavages reflect an atomic structure dominated by unbroken, parallel, linear chains of SiO_4 tetrahedra (silicon atoms surrounded by four oxygens). *See* PYROXENE.

The most common form of enstatite in nature is orthorhombic; it is part of a continuous solid-solution series with ferrosilite ($FeSiO_3$). By convention, the name enstatite is restricted to members of this solution series that are optically positive, which includes only compositions with less than about 12 mol % of the iron component. Pure magnesian enstatite is colorless, with refractive indices $n\alpha = 1.650$, $n\beta = 1.653$ $n\gamma = 1.658$; the substitution of iron increases the refractive indices and produces gray to pale-green coloration. Solid solution of calcic and aluminous components in orthoenstatite is common, but is limited to a few mole percent. *See* SOLID SOLUTION.

Compositions very close to $MgSiO_3$ may crystallize in a variety of different structures. Upon heating at atmospheric pressure, pure orthoenstatite transforms at 985°C (1805°F) to another orthorhombic mineral, protoenstatite, which is then stable up to its melting point near 1557°C (2835°F). A monoclinic form, clinoenstatite, is stable over a wide range of temperatures at pressures in excess of about 7.5–10 gigapascals (1.088–1.450 × 10^6 lb/in.^{2}). Clinoenstatite may also originate outside its field of stability, either by the inversion of protoenstatite on cooling or by intense shearing of orthoenstatite. Crystals of clinoenstatite produced by these mechanisms are repetitively twinned at very fine scale on the crystallographic plane (100). *See* CRYSTAL STRUCTURE.

Magnesium-rich members of the enstatite-ferrosilite series are commonly found in mafic and ultramafic rocks of the Earth's crust, such as basalts, gabbros, peridotites, dunites, serpentinites, and some mafic granulites. In these occurrences, associated minerals are typically olivine, calcic clinopyroxene, and calcic plagioclase. Alteration of enstatite to amphibole or to serpentine is common in such rocks. Enstatite is also an important constituent of the Earth's mantle, as indicated by its presence (along with olivine, diopside, spinel, or garnet) in ultramafic nodules included in lavas that originate at great depths, and in larger bodies of rock, with similar mineralogy, exposed in some deeply rooted mountain belts. Orthoenstatite and clinoenstatite are also found in both iron and stony meteorites.

Because the amounts of calcium and aluminum incorporated in orthopyroxene depend strongly on temperature and pressure, enstatite is an important geologic thermometer and barometer. Estimates of temperatures and pressures of mineral formation, based upon enstatite compositions in natural occurrences, permit geologists to determine environments of origin for many rocks of the Earth's lower crust and upper mantle. *See* GEOLOGIC THERMOMETRY; SILICATE MINERALS.

William D. Carlson

Bibliography. W. D. Carlson, Subsolidus phase equilibria on the forsterite-saturated join $Mg_2Si_2O_6$-$CaMgSi_2O_2$ at atmospheric pressure, *Amer. Mineralog.*, 73:232–241, 1988; D. H. Lindsley, Pyroxene thermometry, *Amer. Mineralog.*, 68:477–493, 1983; C. T. Prewitt, *Pyroxenes*, vol. 7 of *Reviews in Mineralogy*, 1980.

Enteropneusta

A class of Hemichordata with approximately 70 species commonly known as acorn worms. They are free-living, solitary animals with a very soft cylindrical body lacking external appendages. The size is highly variable, from 1 in. to 8 ft (2.5 cm to 2.5 m), and the color ranges from white to shades of violet. The body is covered with cilia and mucus and is always divided into proboscis, collar, and trunk (**Fig. 1**). Sometimes the animal smells like iodoform, and sometimes it shows luminescence.

The acorn worms usually live in shallow waters, buried in the sandy or muddy bottoms, but some species occur at depths of more than 9900 ft (3000 m). *Saxipendium coronatum* lives on hydrothermal vents. The species that live along the shore, sometimes in U-shaped or spiral burrows, can be detected at low tide by the presence of a spiral casting. The proboscis and collar (the acorn) are used in excavating burrows. During this activity, water, slime, sediment, and organic particles enter the mouth and pass into the gut. Seawater is filtered

and expelled through the gill slits, while the organic matter and sediment are retained to be digested.

The proboscis, or protosome, is generally conical and is continuous with the dorsal edge of the collar. The protocoel is more or less filled by muscle and connective tissue, confined to the posterior part, and lined by peritoneum. The protocoel opens to the dorsal surface by means of a proboscis pore.

The collar, or mesosome, is a short cylinder with a highly glandular epithelium usually differentiated into zones. There are two mesocoels, which extend anteriorly into the proboscis stalk. Each mesocoel opens to the exterior by way of a collar canal and pore ending into the first gill slit.

The trunk is differentiated into four regions: the branchial region, externally recognized by two longitudinal rows of dorsal gill pores; the genital region, characterized by the gonads, which occur dorsolaterally (in some genera, laterally from the genital wings); the hepatic region, distinguishable by a darker color and sometimes by the presence of external sacculations. The final region, the abdominal, is simply tubular with a terminal anus. The metacoels are closed.

The nervous system consists of an intraepidermal nervous plexus that becomes thickened to form the middorsal and midventral nerve cords. The dorsal trunk cord continues into the collar through the collar coelom to form the neurochord, or collar cord, which is isolated from the epidermis. A ciliated groove in the proboscis may be a chemoreceptor. Other sense organs are lacking.

The gut is a straight epithelial tube. The mouth opens at the anterior margin of the collar, and a buccal tube whose roof forms the buccal diverticulum by evagination occupies the center of the collar. The pharynx usually is differentiated into a ventral digestive part and a dorsal branchial part that is perforated by gill slits, which are primarily oval and elongated dorsoventrally. In all enteropneusts, however,

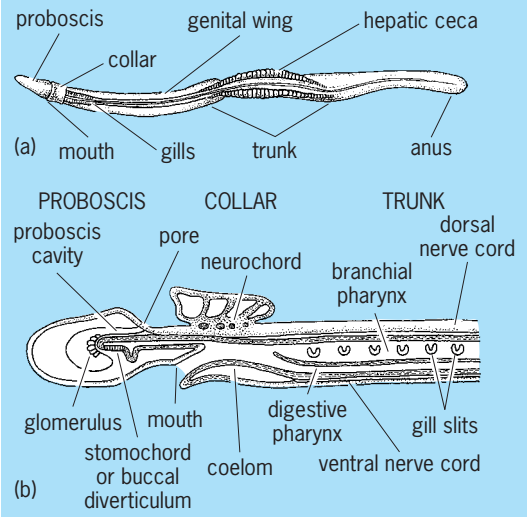

Fig. 1. ***Saccoglossus*, the tongue worm or acorn worm. (*a*) Dorsal view. (*b*) Median section of anterior portion. (*After T. I. Storer et al., General Zoology, 6th ed., McGraw-Hill, 1979*)**

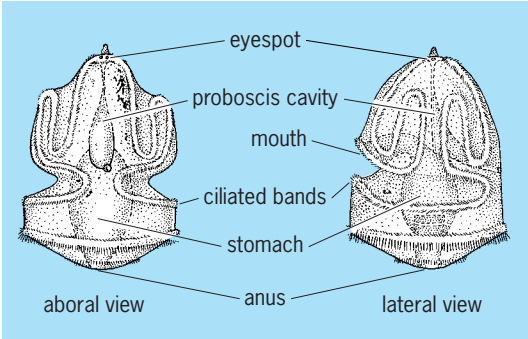

Fig. 2. Tornaria larva. (*After T. I. Storer et al., General Zoology, 6th ed., McGraw-Hill, 1979*)

the gill slits are transformed into U-shaped apertures because of a downgrowth from the dorsal end of the slit, the tongue bar. The gill slits lead to branchial sacs that open on the body surface by the gill pores. A muscular esophagus that molds the food particles follows the pharynx and leads to the hepatic region, where two rows of dorsal sacculations are distinguishable in some species. The intestine continues as a tube to the terminal anus.

The blood vascular system, which is open, consists of two main longitudinal vessels lined by peritoneal cells that lack an endothelium. Also present are a venous sinus, a central sinus, and a heart vesicle. The glomerulus is lined by podocytes.

The sexes are distinct. The sacciform gonads occur in from one to several longitudinal rows at the sides of the digestive tube in the branchial and genital regions. Each gonad opens externally in a pore through which the gametes are shed. Either the eggs are relatively yolky and develop directly, with a free-swimming larval stage that hatches into a miniature acorn worm, or they lack yolk and develop through a planktonic tornaria larval stage (**Fig. 2**). Only one species (*Xenopleura vivipara*) is viviparous, and only one (*Balanoglossus capensis*) reproduces asexually, but all enteropneusts have strong regeneration power. *See* HEMICHORDATA. Jesús Benito

Enterovirus

A genus of the family Picornaviridae. Enteroviruses include the polioviruses (three types), the coxsackieviruses (more than 29 types), and the echoviruses (more than 34 types). They are small (17–28 nanometers in diameter), contain ribonucleic acid (RNA), and are resistant to ether. In contrast to the rhinoviruses, the enteroviruses multiply chiefly in the alimentary tract and are stable under acid conditions (pH 3–5) for 1–3 h. They are protected by magnesium chloride against inactivation by heat. The polioviruses, most echoviruses, and a number of the coxsackieviruses can be grown in cell cultures of monkey origin, as well as in human cells. Certain of the coxsackieviruses will not grow in cultures; they are usually studied by infecting newborn mice. Those strains that can be grown in tissue cultures usually grow best if the cultures are kept

stationary and incubated at 98.5°F (37°C). *See* RIBONUCLEIC ACID (RNA).

Enteroviruses are widespread during summer and fall in temperate climates, but may circulate throughout the year in tropical areas. Different enteroviruses may produce the same symptoms; on the other hand, the same enterovirus may cause more than a single set of symptoms. The majority of enterovirus infections are benign and inapparent. However, when these viruses invade tissues other than the enteric tract, serious disease may result, as when poliovirus invades the spinal cord or when some of the coxsackievirus types invade the heart muscle. *See* ANIMAL VIRUS; COXSACKIEVIRUS; ECHOVIRUS; PICORNAVIRIDAE; POLIOMYELITIS; RHINOVIRUS; VIRUS CLASSIFICATION.

Joseph L. Melnick

Bibliography. H. Fraenkel-Conrat, *The Viruses: Catalogue, Characterization and Classification*, 1985; R. Rott and W. Goebel, *The Molecular Basis of Viral and Microbial Pathogenesis*, 1987; A. Scott, *Pirates of the Cell: The Story of Viruses from Molecule to Microbe*, 1987.

Enthalpy

For any system, that is, the volume of substance under discussion, enthalpy is the sum of the internal energy of the system plus the system's volume multiplied by the pressure exerted by the system on its surroundings. This may be expressed as $U + PV = H$, where U is the system's internal energy, P the pressure of the system, V the system's volume, and H the enthalpy of the system. The sum of $U + PV$ is given the special symbol H primarily as a matter of convenience because this sum appears repeatedly in thermodynamic discussion. Consistent units must, of course, be used in expressing the terms in the above equation. Previously, enthalpy was referred to as total heat or heat content, but these terms are misleading and should be avoided. Enthalpy is, from the viewpoint of mathematics, a point function, as contrasted with heat and work, which are path functions. Point functions depend only on the initial and final states of the system undergoing a change; they are independent of the paths or character of the change. Mathematically, the differential of a point function is a complete or perfect differential. *See* CALCULUS; MATHEMATICS; MAXWELL'S EQUATIONS.

Because the absolute value of internal energy of even a simple system is usually unknown, recorded values of enthalpy are relative values measured above some convenient but arbitrarily chosen datum. Thus in the steam tables of Keenan and Keyes, the datum is liquid water at 32°F (0°C) and under its own vapor pressure. At this state water is assumed to have an enthalpy equal to zero. Under this assumption the internal energy of water in this state is a negative quantity equal to PV. No complication is introduced by this fact, although visualization of negative energies of this kind may be disturbing to some. There is limited utility for absolute enthalpies because only the changes in enthalpy are measurable. It is instructive to examine the utility of the enthalpy function in terms of some simple but important thermodynamic processes.

The first law of thermodynamics is merely a statement of the law of conservation of energy. The first law alone indicates that:

1. For a chemical reaction carried out at constant pressure and temperature with no work performed except that resulting from keeping the internal and external pressure equal to each other as the volume changes, the change in enthalpy of the system (the material taking part in the chemical reaction) is numerically equal to the heat that must be transferred to maintain the above-mentioned conditions. This heat is often loosely referred to as the heat of reaction. More properly, it is the enthalpy change for the reaction.
2. So-called heat balances on heat exchangers, furnaces, and similar industrial equipment that operate under steady flow conditions are really enthalpy balances.
3. The work developed in a steadily running adiabatic engine or turbine is equivalent to the enthalpy change of the fluid passing through the engine.
4. The adiabatic, irreversible, steady flow of a stream of materials through a porous plug or a partially opened valve under circumstances where the change in kinetic energy of flow is negligible (a Joule-Thomson process) results in no change in enthalpy of the flowing stream. Although no change in enthalpy results from this process, there is a loss in the energy available for doing work as a result of the pressure drop across the plug or valve. *See* ENTROPY; THERMODYNAMIC PRINCIPLES; THERMODYNAMIC PROCESSES.

Harold C. Weber; William A. Steele

Entodiniomorphida

An order of the Spirotrichia. These are strikingly different-looking ciliates, covered with a smooth, firm pellicle. They are devoid of external ciliature except for the adoral zone of membranelles and, occasionally, one or two other tufts or zones of other specialized cilia. Internal organization of the body is very specialized and complex. These organisms are considered to be highly evolved. Entodiniomorphids occur exclusively as endocommensals of herbivorous mammals, either in the rumen and reticulum of ruminants or in the colon of certain higher mammals. The ophryoscolecids, which comprise the

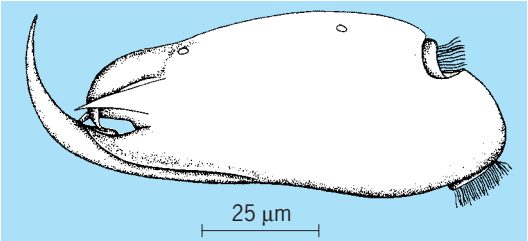

***Ophryoscolex*, an entodiniomorphid.**

majority of species in this order, are found in fantastic abundance in their hosts; for example, there have been estimations of as many as 10 billion per cow. Herbivores can survive free of their harmless protozoan guests, but the reverse is not true. *Epidinium*, *Ophryoscolex* (see **illus.**), and *Entodinium* occur in ruminants. *Troglodytella* is found in the colon of anthropoid apes. *See* CILIOPHORA; PROTOZOA; SPIROTRICHIA. John O. Corliss

Entomology, economic

The study of insects that have a direct influence on humanity. Though this includes beneficial as well as harmful species, most attention is devoted to the latter and how they become pests and are controlled. The emphasis on managing harmful insects reflects the immediacy and seriousness of pest problems, particularly the destruction of agricultural products and the transmission of disease. These are highly visible problems, whereas the benefits gained from useful insects are, in most cases, not so clearly understood or so well documented economically.

Economic thresholds. Central to the definition of a pest is determination of the economic threshold. Any insect population, when introduced into a favorable environment, increases numbers until reaching an environmental carrying capacity (see **illus.**). In pest insects, there exists a density above which the insect population interferes with human health, comfort, convenience, or profits. When this economic threshold is reached, a decision must be made to utilize some control measure to prevent further increase in numbers. Precise estimation of the economic threshold is difficult, since it depends on a myriad of factors such as weather, crop conditions, stage in the life cycle of the insect, market value of the crop, and cost of control, each of which varies. Often, the presence of even a single insect is sufficient to warrant control measures, for instance, when that insect is a flea harboring the plague bacillus, or a mosquito transmitting malaria. Also, consumer expectations in most markets of the United States, Canada, and western Europe are for insect-free produce, so the economic threshold is very low on items that people eat. However, some crops, such as field corn or forest trees, have a higher threshold since they can tolerate a low amount of insect damage before yield is reduced. Economic thresholds may also be higher for insects that damage only the inedible portions of crop plants such as the leaves of beans, tomatoes, and apple trees. In any case, knowledge of the amount of injury which is due to different densities of insects is an important prerequisite for efficient management.

Harmful insects. Insects cause damage in various ways. Direct damage occurs when insects eat foods destined for human consumption or otherwise decrease resource yields. Direct damage to crops by locusts has been a major factor limiting grain production in some semiarid countries of Asia and Africa. Damage to structures by termites and other wood-chewing insects runs to billions of dollars annually. Indirect damage may occur when an insect, feeding on nonedible portions of a crop plant, reduces yield of the edible portion. Red mites and aphids respectively feed on leaves and sap of apple trees, which results in lowered yield of fruit. Many of the most serious insect pests are those that affect humans indirectly by transmitting pathogens. Mosquitoes transmit malaria, yellow fever, viral encephalitis, and filarial roundworms. Such vectors (transmitters) are number one among insect pests. Fleas, ticks, lice, houseflies, cockroaches, and many others are capable of transmitting diseases to humans, while many aphids and leafhoppers are major vectors of plant virus diseases. Indirect damage may also occur from the mere inclusion of insects or their fragments in fresh or processed foods, especially when prohibited by law. Some persons exhibit extreme psychological reaction to a real or imagined insect infestation, and this, too, is damaging.

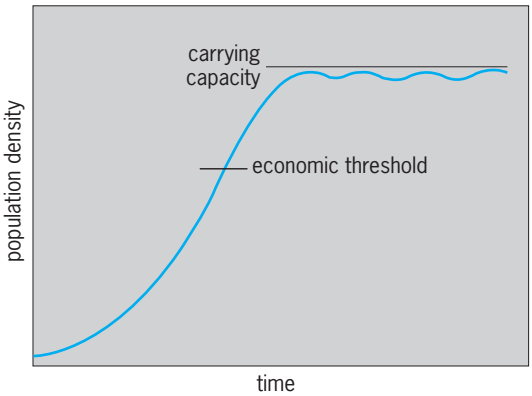

Growth of an insect population continues until reaching the environmental carrying capacity.

Pest management. Management of insect pests begins with prevention. Many of the United States' most noxious insects have been imported from overseas: most domestic cockroaches, the gypsy moth, Japanese beetle, corn borer, housefly, cabbageworm, and codling moth are just a few. Most major pests of foreign origin were introduced accidentally during the nineteenth century. Some North American insects have spread elsewhere, for example—the Colorado potato beetle to Europe and the fall webworm to Japan. To stem the flow of insect invasions, the federal government's Animal and Plant Health Inspection Service maintains inspection facilities for the examination of all incoming shipments of plant or other material that may harbor pests in order to intercept infested items.

Once a pest is established, its spread can sometimes be slowed by an efficient system of local quarantines, early detection, and local eradication. The gypsy moth, long an established defoliator of forests in New England, is slowly extending its range into the midwestern United States, with isolated individuals as distant as Oregon and California. States outside the infested region maintain a vigorous program of detection: sticky traps baited with sex attractant lure male moths, and when several are located at one site, a thorough inspection is followed by application of

insecticide. Gypsy moths lay their eggs on the undersides of vehicles, so that inspection and treatment of autos and trucks that have traveled in the infested area slows the westward spread of this pest.

Insecticides. Once noxious insects are firmly established in a region, there are many techniques to reduce their numbers below the economic threshold. The most widely used method is the application of synthetic chemical insecticides. Insecticides are probably the only efficient control technique for insects with an exceedingly low economic threshold such as occurs on commercial floral crops, fresh and processed fruits and vegetables, or nursery stock. There are scores of available insecticides which vary greatly in their characteristics; extension specialists at the county level or state university should be consulted for the latest information on available compounds. Most insecticides are poisonous to other animals, and handling requires appropriate caution. Insecticides were once regarded as a panacea for pest problems, but the development of resistant strains of major insect pests, together with the rising cost of materials and application and legal restrictions, has led to recognition that insecticides are more efficiently utilized in a program that integrates them with other techniques in a framework of total crop management. Research has been directed toward development of insecticides that minimize hazards to people and their environment. *See* INSECTICIDE.

Biological control. For insects whose economic threshold lies somewhat higher than the artificially low demands of "clean" produce, there is a wider choice of control techniques. Insects introduced into a new homeland where they become numerous sometimes have had their numbers checked in their old homeland by natural enemies. Economic entomologists have effectively reduced densities of several pests by releasing parasites or predators. Imported natural enemies have had major impact in controlling scale insects, aphids, whiteflies, alfalfa weevils, and many others. Natural enemies that are mobile and relatively restricted in diet are most effective in biological control. Diseases of insects often are important in halting outbreaks by killing large numbers in a short time. A few pathogens of caterpillars, mosquito larvae, Japanese beetle larvae, and others are formulated as biological rather than chemical insecticides and are applied commercially. *See* INSECT CONTROL, BIOLOGICAL.

Cultural practices. All populations are limited by their environment, and much can be done to augment the activity of natural control by manipulation of cultural practices. Crop rotation is a standard agronomic practice that often reduces damage due to insects. Rotation of alfalfa or soybeans with corn reduces populations of corn rootworms, wireworms, and white grubs. The physical disruption of autumnal plowing and disking destroys many insects, such as the corn borer, wheat stem sawfly, and cereal leaf beetle, that could overwinter in stubble or on the soil surface. Plant breeders and entomologists have cooperated to produce varieties of corn, wheat, soybeans, alfalfa, and other crops which retain high yields but which resist attacks from some of their more serious pests. The action of natural enemies can be enhanced by cultural practices such as leaving fencerows or other preserves for refuges at harvest. Strip cutting of alfalfa in the southwestern states has been particularly valuable in preserving predators of alfalfa aphids and caterpillars. *See* AGRICULTURAL SOIL AND CROP PRACTICES; BREEDING (PLANT).

Sanitation. The cleanup of breeding and gathering sites is useful, especially in management of medically important insects, many of which have evolved resistance to the commoner insecticides. Housefly larvae feed in decaying vegetable compost and dung; efficient disposal of garbage, manure, and sewage brings enormous relief from fly problems and associated diseases. Similarly, draining standing water with its crop of mosquito larvae reduces their numbers near human habitations. Personal hygiene and proper care of pets virtually eliminate problems with lice and fleas.

Special programs. Rarely, a unique program is developed to control a single pest insect. The screwworm fly, whose larvae infest open wounds of livestock, has been controlled in the United States and much of Mexico by weekly release of 180,000,000 sterile males. These mate with wild females (who only mate once), and the ensuing eggs do not hatch. Obviously, this technique is limited to species in which males are undamaging, though despite technical problems it is being developed for use against codling moths. The method is also being developed to control some species of mosquitoes, in which only females bite.

Integrated control. Most successful programs of insect pest management rely on integrated control or the use of several methods in concert to control a complex of pests. Apples, for example, are attacked by a host of insect pests, led by species such as the codling moth and apple maggot that attack the fruit directly and consequently have a low economic threshold (1/100 apples). Other pests, notably aphids and European red mites, feed on leaves and sap and are only pests when their numbers become enough to reduce yield of fruit.

Application of chemical insecticides to control codling moths killed predators of red mites and caused these creatures to become damaging secondary pests. Eventually, both the red mite and its major predators become partially resistant to most conventional insecticides. In the integrated control program, reduced dosages of insecticides conserve predators of the red mite while still yielding acceptable control of the fruit eaters.

The alfalfa weevil, whose larvae eat alfalfa tips in spring, is usually managed by use of resistant plant varieties and judiciously planned cutting times, with some additional control from six species of introduced parasites. Occasionally, the economic threshold is exceeded enough for an insecticide to be applied. Farmers must watch carefully as populations increase, or the crop may also be inspected by an agricultural consultant. The consultant's pest management scouts observe the fields regularly, and the farmers receive weekly (or more frequent) reports

on the status of their crops. Pest management programs involving crop scouting are now operational for corn, cotton, soybeans, fruits, tomatoes, tobacco, alfalfa, and other crops.

Beneficial insects. It has been estimated that the dollar value gained from a single insect, the honeybee, equals the loss from damage plus cost of control for all pests combined. Honeybees are managed for their honey and beeswax, but their most valued service is pollination of crop plants. Nearly all fruits and many vegetables, ornamental plants, and seed crops require pollination by honeybees or other insects. Bees of many species are the chief pollinators, though wasps, flies, moths, butterflies, and beetles pollinate as well.

Silk is produced by larvae of the silkworm, an insect so thoroughly domesticated that it cannot climb its food plant, mulberry, with its degenerated legs. The silkworm apparently no longer survives in the wild. In many uses, silk has been more recently replaced by less expensive synthetic materials.

The economic value of silkworms and honeybees is rather easily estimated from the cash value of their products. Other insects may be equally beneficial, but their value is not so easily calculated. Foremost among these are predatory insects of several orders. Some, chiefly parasitic wasps and predatory beetles, have been imported specifically for control of noxious pests. Others, such as dragonflies, mantises, lace wings, and such, are voracious predators and doubtless are a factor determining the environmental carrying capacity for many pests. These predators may prevent other insects from ever reaching an economic threshold and thus from becoming pests, as was demonstrated when predators of the European red mite were destroyed by chemical insecticides, allowing the mite population to reach damaging numbers on apple crops.

Innumerable insect species are scavengers, quietly but efficiently breaking down the remains of dead plants and animals. Their economic activity goes unnoticed, save when scavengers such as termites forsake logs on the forest floor for sills and siding on a summerhouse, or when flies invade a home. A lack of scavenging insects would, however, result in a great increase of decomposing organic material lying about.

Plant-eating insects have been set to beneficial use when their diets consist mainly of unwanted weeds. Alligatorweed in the southeastern United States and klamath weed in the northwestern states have been controlled by imported beetles, while prickly-pear cactus was similarly eliminated from Australian rangeland by larvae of moths from South America. Pasture thistles may soon be controlled by beetles in the eastern United States and adjacent Canada.

In many parts of the world, particularly the tropics and subtropics, termites, grubs, locusts, and other insects are routinely eaten by people, and for some they are a major source of dietary protein. In places like the United States, human consumption of insects is limited to expensive novelty items such as chocolate-covered ants, fried grasshoppers, and the like.

Certain rare and showy butterflies and beetles are sought after so that they have considerable economic worth. Conservation of rare and endangered insects incurs some expense as well. Habitat management to conserve rare insects is a valid and growing concern of economic entomologists.

Finally, insects have rendered invaluable service to science, and thus to humanity, as easily reared experimental animals for investigation of basic principles of genetics, biochemistry, development, and behavior. Pomace flies (genus *Drosophila*) have been extremely useful in this regard, and even cockroaches and houseflies are helpful in testing the effectiveness of new chemical insecticides and other insect control methods. *See* INSECTA. David J. Horn

Bibliography. R. H. Davidson and W. F. Lyon, *Insect Pests of Farm, Garden, and Orchard*, 8th ed., 1987; D. J. Horn, *Ecological Approach to Pest Management*, 1988; R. L. Metcalf and W. H. Luckmann, *Introduction to Insect Pest Management*, 3d ed., 1994; R. E. Pfadt, *Fundamentals of Applied Entomology*, 4th ed., 1985.

Entoprocta

A phylum of sessile, aquatic, often colonial invertebrates having a looped gut with both mouth and anus situated inside a circlet of tentacles, a pseudocoelomate body cavity, and no hardened skeleton.

Form and function. Entoprocts, for example, *Pedicellina*, are mainly marine, and may be encountered as small colonies on seaweeds and bryozoans near low tidemark. The colonies consist of threadlike, creeping stolons and erect zooids arising at intervals from them (see **illus.**). Noncolonial forms have similar zooids but lack the stolons. Each zooid is goblet-shaped, consisting of a basal stalk supporting a rounded or pyriform viscera-containing dilatation, the calyx. The zooids (or individuals) are usually in the range 0.016–0.08 in. (0.4–2.0 mm), but may reach 0.2 in. (5 mm) in species with long stalks. Usually undifferentiated, the stalk has alternating flexible and rigid sections in *Barentsia*.

The body wall comprises a cuticle-covered epidermis above a layer of muscle fibers. These fibers are particularly developed in the stalk. The body cavity, morphologically a pseudocoel, contains connective tissue and fluid. In the colonial forms, cells at the junction of stalk and calyx contract rhythmically, perhaps pumping or compressing fluid.

The U-shaped alimentary canal occupies most of the calyx, below a circlet of 6–36 ciliated tentacles. The entoprocts feed on particles filtered from the surrounding water. Long lateral cilia drive water inward between the tentacles, while shorter frontal cilia (that is, on the inner face) convey particles trapped in mucus down the tentacles. These frontal tracts join an annular ciliated groove, which dips into the mouth. The stomach is globular; the anus is raised on a papilla, leaving a space or atrium between

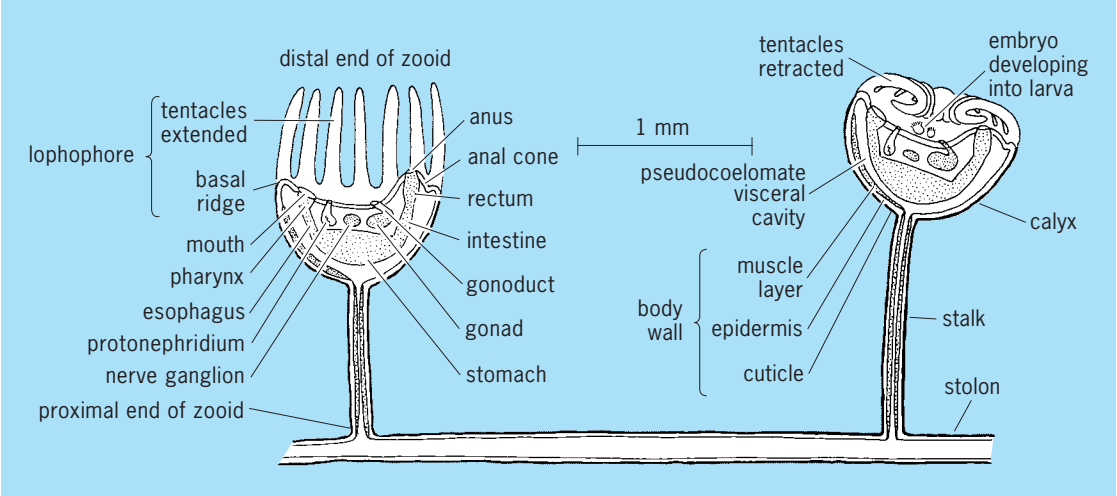

Morphological features of entoproct zooids, illustrated schematically in vertical cross section.

the rectum and the mouth. Two protonephridia discharge into the atrium, except in the fresh-water *Urnatella*, which has many nephridia of known osmoregulatory function. Adverse stimuli cause the tentacles to fold inward; the calyx is then closed by a muscular membrane.

Ecology and life history. In one order, Pedicellinida, zooids are colonial and free-living; in the other, Loxosomatida, they are noncolonial and live in association with various other invertebrates, for example, tubicolous polychaetes, sipunculids, sponges, and bryozoans. The loxosomatids are energy commensals, collecting food from the respiratory or feeding current of the host animal.

Entoprocts reproduce asexually and sexually. Loxosomatids bud off a succession of daughter zooids from the paired blastogenic zones on the calyx. These disperse only a short way from the parent, giving rise to characteristically dense aggregations on the host. Pedicellinids bud from the stolon: daughter zooids do not break free, so that a colony develops.

Dispersal is by means of sexually produced larvae. Zooids may be unisexual or hermaphroditic. Gonads are paired, with the ducts uniting before opening into the coelom. Fertilization is internal. Young embryos become attached to a part of the atrial wall, via which they obtain nourishment. The larva is possibly of the trochophore type, but the ciliary girdle is basal rather than equatorial. The *Pedicellina* larva has obvious similarity to an inverted and nontentaculate adult calyx. It attaches with mouth and anus toward the substrate, and metamorphosis involves reorientation of the gut to the adult position. Loxosomatid larvae may metamorphose or give rise directly to external or internal zooid buds, the latter type being released by rupture of the larva.

History and classification. The name Entoprocta was introduced in 1869 as a subdivision of the phylum Bryozoa, but was recognized as an independent phylum in 1888. Despite some possible similarities in larval morphology and metamorphosis, the differences between Entoprocta and Ectoprocta are numerous and fundamental. *See* BRYOZOA.

Alternative names are Calyssozoa and Kamptozoa. In view of the confusing similarity between the words Entoprocta and Ectoprocta, preference should be given to the distinctive and apposite name Calyssozoa. The phylum is divided into two orders, Loxosomatida (or Solitaria), with about 100 species, and Pedicellinida (or Coloniales), with about 30. No fossils are known. John S. Ryland

Entropy

A function first introduced in classical thermodynamics to provide a quantitative basis for the common observation that naturally occurring processes have a particular direction. Subsequently, in statistical thermodynamics, entropy was shown to be a measure of the number of microstates a system could assume. Finally, in communication theory, entropy is a measure of information. Each of these aspects will be considered in turn. Before the entropy function is introduced, it is necessary to discuss reversible processes.

Reversible processes. Any system under constant external conditions is observed to change in such a way as to approach a particularly simple final state called an equilibrium state. For example, two bodies initially at different temperatures are connected by a metal wire. Heat flows from the hot to the cold body until the temperatures of both bodies are the same. As another example, a vessel containing a gas is connected through a stopcock to an evacuated vessel. When the stopcock is opened, the gas expands to fill the whole of the available space uniformly. It is common experience that the reverse processes never occur if the systems are left to themselves; that is, heat is never observed to flow from the cold to the hot body, nor will the gas compress itself into one of the vessels. Max Planck classified all elementary processes into three categories: natural, unnatural, and reversible.

Natural processes do occur, and proceed in a direction toward equilibrium. Unnatural processes move

away with equilibrium and never occur. If A → B is a natural process between states A and B, then B → A is an unnatural process. A reversible process is an idealized natural process that passes through a continuous sequence of equilibrium states. Consider the evaporation of a liquid in the presence of its vapor at a pressure P. Let the equilibrium vapor pressure of the liquid be p. If $P < p$, liquid evaporates as a natural process. If $P > p$, evaporation is an unnatural process and will not occur; indeed, the opposite process—condensation—will take place. Finally, if $P = p$, both processes of condensation and evaporation are reversible and can be initiated by a very slight increase or decrease in the external pressure P.

A useful idea is that a reversible process may be exactly reversed by an infinitesimal change in the external conditions. If a hot object is placed adjacent to a much colder object, the heat-flow direction cannot be reversed by small changes in the temperature of either object. In reversible processes, work is accomplished through small pressure differences, and heat transfer occurs through small temperature differences.

Entropy function. The state function entropy S puts the foregoing discussion on a quantitative basis. The function is not derived in this article; but, rather, some of its properties are stated, and its implications are discussed mainly by example. Entropy is related to q, the heat flowing into the system from its surroundings and to T, the absolute temperature of the system. The important properties for this discussion are:

1. $dS > q/T$ for a natural change. $dS = q/T$ for a reversible change. It is necessary to introduce both S and T together. A formal derivation would show T^{-1} as an integrating factor leading to the complete differential dS.

2. The entropy of the system S is made up of the sum of all the parts of the system so that $S = S_1 + S_2 + S_3 \cdots$. *See* HEAT; TEMPERATURE; THERMODYNAMIC PRINCIPLES.

Heat flow. Consider two bodies, α and β, at different temperatures separated by an adiabatic (no heat transfer) wall. If the two bodies are connected by a fine wire that allows a small heat flow q from α to β, then $dS_\alpha = -q/T_\alpha$ and $dS_\beta = q/T_\beta$.

For the whole system, Eq. (1) holds. If $T_\alpha > T_\beta$,

$$dS = dS_\alpha + dS_\beta = q\left(\frac{1}{T_\beta} - \frac{1}{T_\alpha}\right) \qquad (1)$$

$dS > 0$, and heat flows from α to β as a natural process. The process could be continued until $T_\alpha = T_\beta$ and $dS = 0$.

Once the constraint of the adiabatic wall is abrogated, the entropy increases to a maximum value, and T_α becomes equal to T_β. This is a special case of the most important notion in thermodynamics; that is, the system will assume that equilibrium state which maximizes the entropy at constant energy, consistent with the constraints. *See* HEAT TRANSFER.

Nonconservation of entropy. In his study of the first law of thermodynamics, J. P. Joule caused work to be expended by rubbing metal blocks together in a large mass of water. By this and similar experiments, he established numerical relationships between heat and work. When the experiment was completed, the apparatus remained unchanged except for a slight increase in the water temperature. Work (W) had been converted into heat (Q) with 100% efficiency. Provided the process was carried out slowly, the temperature difference between the blocks and the water would be small, and heat transfer could be considered a reversible process. The entropy increase of the water at its temperature T is $\delta S = Q/T = W/T$.

Since everything but the water is unchanged, this equation also represents the total entropy increase. The entropy has been created from the work input, and this process could be continued indefinitely, creating more and more entropy. Unlike energy, entropy is not conserved. *See* CONSERVATION OF ENERGY.

Although the heat transfer is considered to be reversible in order to calculate the entropy increase, the overall process of converting work into heat is irreversible. The frictional process that converts kinetic energy into the heat of the metal blocks is a natural process. In fact, the impossibility of the reverse process is Lord Kelvin's statement of the second law of thermodynamics. Heat cannot be completely converted into work without other changes occurring in the surroundings. For example, a gas in a cylinder can be expanded reversibly by extracting heat from a large constant-temperature bath. All of the heat extracted from the bath is converted into work, but eventually the pressure of the gas system would be reduced to an unusable level. The system has changed, and the process cannot continue indefinitely. If one tries to convert heat into work through a system undergoing a cycle so that the system will return to its initial state, one finds that only a portion of the heat input does work and that the remainder must be rejected to a lower temperature; this is just the process which takes place in a heat engine. *See* THERMODYNAMIC CYCLE; THERMODYNAMIC PROCESSES.

Degradation of energy. Energy is never destroyed. But in the Joule friction experiment and in heat transfer between bodies, as in any natural process, something is lost. In the Joule experiment, the energy expended in work now resides in the water bath. But if this energy is reused, less useful work is obtained than was originally put in. The original energy input has been degraded to a less useful form. The energy transferred from a high-temperature body to a lower-temperature body is also in a less useful form. If another system is used to restore this degraded energy to its original form, it is found that the restoring system has degraded the energy even more than the original system had. Thus, every process occurring in the world results in an overall increase in entropy and a corresponding degradation in energy. R. Clausius stated the first two laws of thermodynamics as: "The energy of the world is constant. The entropy of the world tends toward a maximum."

Increasing entropy and mixing. Once the atomic theory of matter is accepted, the entropy concept

can be made much clearer. It is then found through statistical thermodynamics that the increase of entropy toward its maximum value at equilibrium corresponds to the change of the system toward its most probable state consistent with the constraints. The most probable state represents the most mixed or most random state. Mixing must be given a broad interpretation which includes particle or configurational mixing, and spreading of energy over the particles or thermal mixing. Diffusion of one gas into another represents obvious configurational mixing and increased entropy. Irreversible expansion of a gas represents configurational mixing of the molecules over the available space. Heat flow represents spreading of the kinetic energy between the particles. Friction spreads the kinetic energy of the body over the constituent particles. Sometimes the energy-spread entropy increase and the configurational entropy increase are not compatible, and a compromise is struck. A subcooled liquid adiabatically crystallizes to a lower configurational entropy but gains even more entropy through the additional energy levels made available. The same sort of behavior occurs in partially miscible liquids— some configurational entropy is sacrificed in order to gain a large amount of energy-spread entropy. *See* STATISTICAL MECHANICS.

Absolute entropy. The third law of thermodynamics (Nernst's heat theorem) refers to the vanishing of entropy at zero temperature. In 1912 Planck proposed that the theorem applied to pure crystalline solids. However, the theorem is now known to be applicable to gases and, by all reasonable expectation, is applicable to any system. Thus, any substance at finite temperatures has an absolute entropy, the value of which can be determined from either calorimetric or spectroscopic data. Absolute entropies, together with thermochemical data, are very useful in the calculation of equilibrium compositions of reaction systems.

The statistical viewpoint is that a thermodynamic state at finite temperatures corresponds to many microstates. During an observation the microstates of a system undergo continuous rapid transitions. Since entropy is proportional to the logarithm of the number of available microstates, the Nernst theorem implies that the thermodynamic state at zero temperature corresponds to a single microstate. At zero temperature, even a ferromagnetic material should exist in a single state, fully magnetized in a direction determined by its inevitable interactions with the environment. William F. Jaep

Measure of information. The probability characteristic of entropy leads to its use in communication theory as a measure of information. The absence of information about a situation is equivalent to an uncertainty associated with the nature of the situation. This uncertainty, designated H, is the entropy of the information about the particular situation, Eq. (2),

$$H(p_1, p_2, \ldots, p_n) = -\sum_{k=1}^{n} p_k \log p_k \qquad (2)$$

where $p_1, p_2, \ldots, p_n$ are the probabilities of mutually exclusive events, the logarithms are taken to an arbitrary but fixed base, and $p_k \log p_k$ always equals zero if $p_k = 0$. For example, if $p_1 = 1$ and all other p's are zero, the situation is completely predictable beforehand; there is no uncertainty and so the entropy is zero. In all other cases the entropy is positive. *See* INFORMATION THEORY.

In introducing entropy of an information space, C. E. Shannon described a source of information by its entropy H in bits per symbol. The ratio of the entropy of a source to the maximum rate of signaling that it could achieve with the same symbols is its relative entropy. One minus relative entropy is the redundancy of the source. Frank H. Rockett

Bibliography. H. B. Callen, *Thermodynamics and an Introduction to Thermostatics*, 2d ed., 1985; K. G. Denbigh, *Principles of Chemical Equilibrium*, 4th ed., 1981; R. M. Gray, *Entropy and Information Theory*, 1991; J. N. Kapur, *Measures of Information and Their Application in Science and Engineering*, 1994; K. S. Pitzer, *Thermodynamics*, 3d ed., 1995; R. C. Tolman, *The Principles of Statistical Mechanics*, 1980; K. Wark, *Thermodynamics*, 6th ed., 2000.

Environment

The sum of all external factors, both biotic (living) and abiotic (nonliving), to which an organism is exposed. Biotic factors include influences by members of the same and other species on the development and survival of the individual. Primary abiotic factors are light, temperature, water, atmospheric gases, and ionizing radiation, influencing the form and function of the individual.

For each environmental factor, an organism has a tolerance range, in which it is able to survive. The intercept of these ranges constitutes the ecological niche of the organism. Different individuals or species have different tolerance ranges for particular environmental factors—this variation represents the adaptation of the organism to its environment. The ability of an organism to modify its tolerance of certain environmental factors in response to a change in them represents the plasticity of that organism. Alterations in environmental tolerance are termed acclimation. Exposure to environmental conditions at the limit of an individual's tolerance range represents environmental stress. *See* ADAPTATION (BIOLOGY); ECOLOGY; PHYSIOLOGICAL ECOLOGY (ANIMAL); PHYSIOLOGICAL ECOLOGY (PLANT).

Abiotic factors. All the physical factors which affect an organism constitute the abiotic environment.

Light radiation. The spectrum of electromagnetic radiation reaching the Earth's surface is determined by the absorptive properties of the atmosphere (**Fig. 1**). Biologically, the most important spectral range is 300–800 nanometers, incorporating ultraviolet, visible, and infrared radiation. Visible light provides the energy source for most forms of life. Light absorbed by pigment molecules (chlorophylls, carotenoids, and phycobilins) is converted into chemical energy through photosynthesis. Photosynthetic bacteria

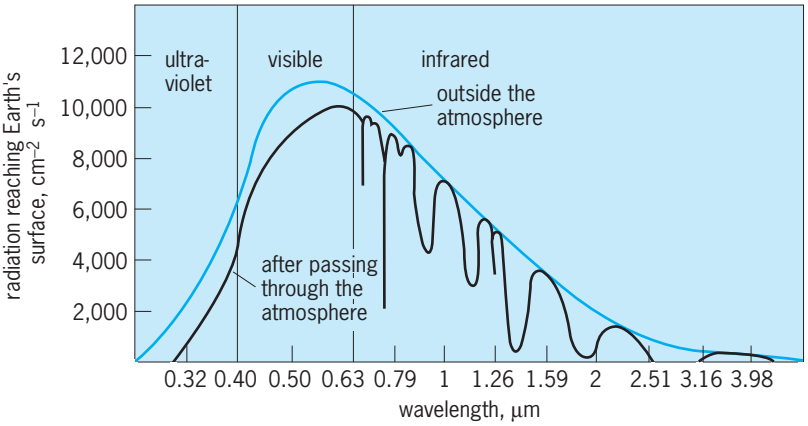

Fig. 1. Spectrum of electromagnetic radiation reaching the Earth's surface. Light quality is determined by the spectrum of light available from the Sun and by the absorptive properties of the atmosphere.

containing bacteriochlorophyll can absorb light into the far-red portion of the spectrum, up to 1000 nm, allowing them to survive in locations where the availability of visible light is low. Light availability is especially important in determining the distribution of plants. Plants receiving too little light (low light stress) may be unable to maintain an average photosynthetic rate high enough to support net growth. Excess light (high light stress) causes damage to leaf tissue, which in extreme cases may kill the plant. Photosynthetic organisms can exist within a wide range of light intensities. Full sunlight in the tropics is around 2000 μmol photons $\cdot$ $m^{-2} \cdot s^{-1}$. Photosynthetic organisms have survived in locations where the mean light is as low as 0.005% of this value. *See* INSOLATION; PHOTOSYNTHESIS; SOLAR RADIATION.

In addition to providing energy, light is important in providing an organism with information about its surroundings. The human eye, for example, is able to respond to wavelengths of light between 400 and 700 nm—the visible range. Within this range, sensitivity is greatest in the green part of the spectrum. This is the portion of the spectrum that plants absorb least, and so is the principal part of the spectrum to be reflected. The detection of light is not confined to animals with recognizable eyes. Many photosynthetic bacteria exhibit phototaxis, movement toward or away from light. Plants use light cues to determine the direction of growth and the extent of stem elongation. In such responses, light quality sometimes may be more important than the intensity of light. Plants possess photoreceptors that respond to ultraviolet, blue, red, and far-red light. Of particular importance in plant development is the ratio of red to far-red light. Red light is absorbed efficiently by leaves, whereas they are relatively transparent to far-red light. Hence, a low red to far-red ratio indicates that the leaf is being shaded by another leaf on the same or another plant. Such signals allow plants to optimize their growth form to maximize light capture and minimize shading from neighbors.

Temporal variation in light also provides an important stimulus. Life forms from bacteria upward are able to detect and respond to daily light fluctuations. Such a response may be directly controlled by the presence or absence of light (diurnal rhythms) or may persist when the variation in light is removed (circadian rhythms). In the latter case, regulation is through an internal molecular clock, which is able to predict the daily cycle. Such circadian clocks are normally reset by light on a daily basis. Processes controlled by circadian clocks range from the molecular (gene expression) to the behavioral (for example, sleep patterns in animals or leaf movements in plants). The ability to detect changes in day length is also essential to many biological processes. Day length changes are used by organisms to predict seasonal changes. For example, shortening of days, indicating the onset of winter, is used as a cue to trigger hibernation or migration in animals, and flowering, seed set, and leaf loss in plants. *See* PHOTOPERIODISM.

Ultraviolet radiation represents a form of stress. It has the ability to break chemical bonds and so may lead to damage to proteins, lipids, and nucleic acids. Damage to deoxyribonucleic acid (DNA) may result in genetic mutations. The amount of ultraviolet radiation reaching the Earth's surface is thought to be increasing due to depletion of the ozone layer in the stratosphere. Ozone is responsible for absorbing a large proportion of ultraviolet radiation reaching the outer atmosphere. As ozone is destroyed by the action of pollutants such as chloroflurocarbons, the proportion of ultraviolet radiation reaching the surface of the Earth rises.

Water. Water is ubiquitous in living systems, as the universal solvent for life. A sufficient supply of water is essential for biological activity. Many organisms have evolved the ability to survive prolonged periods in the total absence of water, but this is achieved only through the maintenance of an inactive state. Life first evolved in water and organisms later migrated onto land. Water availability remains a primary environmental factor limiting survival on land. Primitive land organisms possess little or no ability to conserve water within their cells and are termed poikilohydric. Examples include amphibians and primitive plants such as most mosses and liverworts. These are confined to places where water is in plentiful supply, for example close to ponds, or they must be able to tolerate periods of desiccation. Lichens can survive total water loss and rapidly regain activity upon rewetting. Such organisms must be able to minimize the damage caused to cellular structures when water is lost. Dehydration causes irreversible damage to membranes and proteins. This damage can be prevented by the accumulation of protective molecules termed compatible solutes.

Homeohydric organisms possess a waterproof layer that restricts the loss of water from the cells. Such waterproofing is never absolute, as there is still a requirement to exchange gas molecules and to absorb organic or mineral nutrients through a water phase. Hence, lungs and guts in mammals or stomata and roots in plants represent potential sites of water loss. Nevertheless, water conservation allows

organisms to live in environments in which the water supply is extremely low. In extremely arid environments, behavioral adaptations may allow the water loss to be minimized. Animals may be nocturnal, emerging when temperatures are lower and hence evaporation minimized. Cacti possess a form of photosynthesis, crassulacean acid metabolism (CAM), that allows them to separate gas exchange and light capture. Carbon dioxide is taken up and stored during the night, allowing the stomata to remain closed during the daytime. At the other extreme, inundation with water may be detrimental to organisms not adapted to wet environments. Asphyxiation through drowning in animals is paralleled in plants, where waterlogging of soils may lead to oxygen stress in the roots. Adaptations to survival in waterlogged soils includes air spaces in stems, aerenchyma, that allow stems to act like breathing tubes passing oxygen down to the roots and into the soil. *See* GROUND-WATER HYDROLOGY; OSMOREGULATORY MECHANISMS; PLANT-WATER RELATIONS.

Temperature. Temperature is a determinant of survival in two ways: (1) as temperatures decrease, the movement of molecules slows and the rate of chemical reactions declines; (2) temperature determines the physical state of water.

The slowing of metabolic activity at low temperatures is illustrated in reptiles. Such poikilothermic animals, unable to maintain their internal temperature, are typically inactive in the cold of morning. They bask in the sun to increase their body temperature and so become active. High temperatures will cause the three-dimensional structure of proteins to break down, preventing the organisms from functioning. Organisms adapted to extremely high temperatures need more rigid proteins that maintain their structure. Temperature also affects the behavior of cell membranes, made up of lipids and proteins in a liquid crystalline state. At low temperatures, the membrane structure becomes rigid and liable to break. At high temperatures, it becomes too fluid and again liable to disintegrate. In adapting to different temperatures, organisms alter the composition of the lipids in their membranes, whose melting temperature is thereby changed. This outcome also applies to storage lipids. Hence, cold-water fish are a useful source of oils, whereas mammals, with their higher body temperature, contain fats. The effect of temperature on membranes is thought to be a key factor determining the temperature range that an organism is able to survive.

The effect of temperature on the physical state of water is essential to determining the availability of that water to organisms. Poikilothermic organisms may find that the water in their cells begins to freeze at low temperatures. The temperature at which this occurs depends on a number of factors. For example, the presence of solutes in the cells depresses the freezing temperature. Water in cells can also remain liquid at temperatures below the freezing point (a phenomenon termed supercooling) depending on the absence of sites where ice crystals can form. Finally, certain species can survive total freezing, down to 77 K (-196°C or -321°F), through the prevention of ice crystal formation altogether. Water is frozen in an amorphous state, without forming crystals, which would otherwise damage cellular structures. To survive low temperatures, cells must be able to survive desiccation, and so low-temperature tolerance involves the formation of compatible solutes. High temperatures increase the rate of evaporation of water. Hence, where water supply is limiting, an organism's ability to survive high temperatures is impaired.

Mammals and birds, homeothermic organisms, are able to regulate their internal temperature, limiting the effects of external temperature variations. Temperature still acts as an environmental constraint in such organisms, however. Cooling is achieved through sweating and hence loss of water. Heat is produced through the metabolism of food, and hence survival in cold climates requires a high metabolic rate. *See* CRYPTOBIOSIS; HIBERNATION AND ESTIVATION; THERMOREGULATION.

Atmospheric gases. The atmosphere on Earth is thought to be determined to a large extent by the presence of life. At the same time, organisms have evolved to survive in the atmosphere as it is. The atmospheric constitutents with the most direct biological importance are oxygen (O_2) and carbon dioxide (CO_2). Oxygen makes up approximately 20% of the atmosphere and is due to the occurrence of oxygenic photosynthesis. This process involves the simultaneous uptake of CO_2 to make sugars. Aerobic respiration involves the reverse of this process, the release of CO_2 and the uptake of O_2 to form water. Hence, the current atmosphere represents the balance of previous biological activity. For most terrestrial organisms, neither CO_2 nor O_2 is limiting in the atmosphere; however, the need to get either or both of these gases to cells may represent a limitation on size or on the ability to tolerate water stress. Limitation of either gas may be important in aquatic environments, where the concentration of each is significantly lower. Conversely, oxygen may be toxic to certain anaerobic bacteria, which lack mechanisms to prevent oxidative damage. Such organisms are confined to anaerobic environments, such as waterlogged soils.

Nitrogen is also required by all organisms but cannot be used by most in the gaseous form. Nitrogen fixation, the conversion of N_2 gas into a biologically useful form, occurs in some species of bacteria and cyanobacteria or may be caused by lightning.

Atmospheric gases are important in determining the climate and the light environment. Absorption of electromagnetic radiation by the atmosphere determines the spectrum of light reaching the Earth's surface. Absorption and reflectance of infrared radiation by greenhouse gases such as CO_2 and water vapor regulate temperature. *See* PHOTORESPIRATION; RESPIRATION.

Ionizing radiation. Ionizing radiation entering the atmosphere as cosmic rays or generated through the decay of radioactive material may have a significant impact on the survival of organisms. High-level

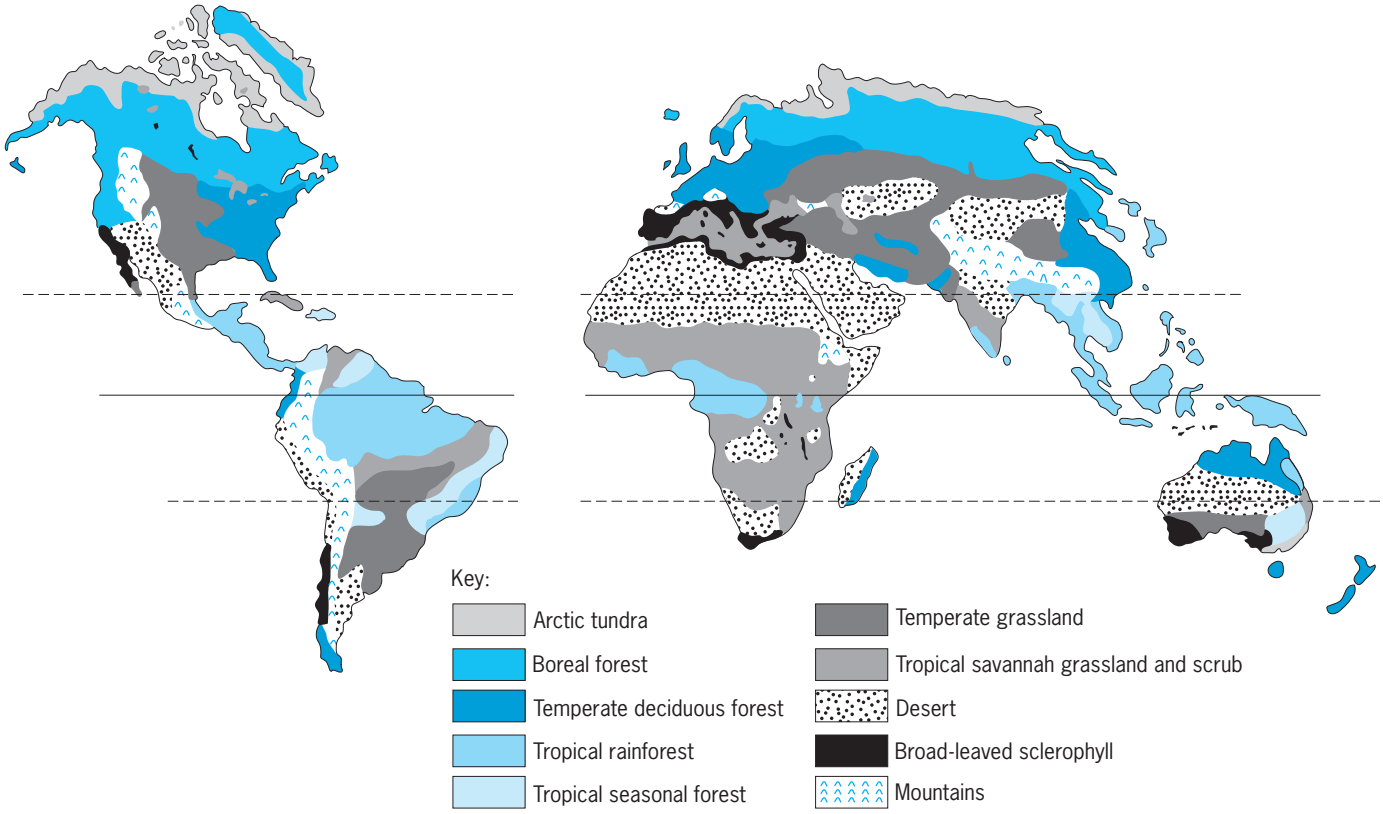

Fig. 2. Global distribution of major vegetation types (biomes).

exposure to radiation may directly damage tissues, ultimately killing the individual. Lower levels of radiation may cause damage to DNA, leading to the accumulation of genetic mutations. These mutations are frequently deleterious, but they also give rise to beneficial genetic variations. *See* PHOTORESPIRATION; RESPIRATION.

Other factors. Among other environmental factors determining the range and distribution and form of organisms are mechanical stimuli such as wind or water movement, and the presence of metals, inorganic nutrients, and toxins in the air, soil, or food. Wind, for example, may be a key factor in determining vegetation height, with plants tending to be shorter but having stronger stems when growing in exposed locations. Metals and other ions in the soil or food may be essential to the survival of an organism, as is true of magnesium, iron, and calcium. The concentration of these ions may represent a limit on the organisms in a particular location. Alternatively, ions may be nonessential or toxic (such as lead or aluminum). All inorganic ions may be toxic if present in sufficient quantities. Environmental factors particularly relevant in aquatic environments include water pressure and salinity.

Biotic factors. The biotic environment of an individual is made up of members of the same or other species. Intraspecific interactions involve the need to breed with other individuals, to gain protection through living in a group, and to compete for resources such as food, light, nutrients, and space. Low population densities exert a negative effect on an individual due to failure to breed or exposure to predation. High population densities are stressful as a result of increased competition. The optimal population density depends on the availability of resources and on the behavior, size, and structure of the organism. Interspecific interactions may also be positive or negative. For example, symbiotic relationships involve the mutual benefit of the individuals involved, whereas competition for resources is deleterious to both. Symbiotic relationships may be positive to one species and negative to the other. For example, a parasite gains benefit from its host but exerts a negative influence on that individual. Such relationships are not simple. Although predation exerts a negative influence on the population as a whole, the success of an individual may be enhanced if a predator removes one of its conspecific competitors.

The impact of humans on natural environments may be considered a special case among biotic factors. Humans alter their environment in ways that exceed the impact of all other organisms. For example, living in heated houses creates a hot, relatively dry environment that can be exploited by other organisms. Hence, cockroaches have exploited humans to extend their ecological range. Heating involves the release of greenhouse gases into the atmosphere, contributing to climate alterations over the entire planet. This in turn has impacts on the distribution of all other species. The release of pollutants into the environment brings organisms into contact with stresses to which they were not previously

exposed. This causes the evolution of new varieties, eventually perhaps new species, adapted to the polluted environments. *See* AIR POLLUTION; BIOSPHERE; HUMAN ECOLOGY; WATER POLLUTION.

Limiting factors and environmental stress. For any given organism, it is often possible to identify a factor in the environment that limits survival and growth. The limiting factor may change through time. Such a change may cause the organism to be at the limit of or outside its tolerance range for that or another environmental factor. In such cases, the organism is said to suffer stress. The relationship between the limitation of one factor and the stress observed may not be simple. For example, a plant suffering from water limitation closes its stomata, restricting the entry of CO_2 into the leaf and so reducing photosynthesis. The plant may then be vulnerable to light stress, due to light absorption exceeding the rate at which the energy is used.

If the stress to which an individual is exposed is extreme, it may result in irreversible damage and death. Exposure to moderate stress, however, results in a period of acclimation within the organism that allows it to adjust to the new conditions. Organisms exposed gradually to new conditions usually have a higher chance of survival than those exposed suddenly. This concept is familiar to gardeners who gradually expose plants to cold, hardening them off over several weeks before leaving them outdoors totally unprotected. *See* POPULATION ECOLOGY.

Environment and biomes. Where a particular environmental factor (or combination of factors) dominates the growth and development of organisms, it is often found that the adaptations and gross features of the landscape will be the same, even when the actual species are different. Thus, mediterranean vegetation is found not only around the Mediterranean Sea but also in California and South Africa, where the conditions of hot dry summers and warm wet winters occur. Regions with similar environmental conditions are classed as biomes (**Fig. 2**). The occurrence of such global vegetation types clearly illustrates the role played by the environment in determining the form and function of individual species.

Gites Johnson

Bibliography. R. Brewer, *The Science of Ecology*, Saunders College Publishing, 1994; P. Colinvaux, *Ecology 2*, Wiley, 1993; R. M. M. Crawford, *Studies in Plant Survival*, Blackwell Scientific, 1989; B. Freedman, *Environmental Ecology*, Academic Press, 1995; E. P. Odum, *Ecology: A Bridge Between Science and Society*, Sinauer Associates, 1997.

Environmental engineering

The division of engineering concerned with the environment and management of natural resources. The environmental engineer places special attention on the biological, chemical, and physical reactions in the air, land, and water environments and on improved technology for integrated management systems, including reuse, recycling, and recovery measures.

Scope. Environmental engineering began with the need for acceptable drinking water and for management of liquid and solid wastes. In fact, the terms public health and sanitary were used to identify the engineers seeking solutions for eliminating waterborne disease in the 1800s. Later, urbanization and industrialization contributed significantly to the formation of unsanitary conditions in many areas. Abatement of air and land contamination became new challenges for the environmental engineer, followed by toxic-waste and hazardous-waste concerns. The environmental engineer is also instrumental in the mitigation and protection of wildlife habitat, preservation of species, and the overall well-being of ecosystems.

During project development, environmental engineers conduct a life-cycle analysis (cost-effectiveness analysis) of project alternatives that consider the capital and the operation and maintenance costs over the life of the project. The National Environmental Policy Act and other legislation, regulations, and policies contain provisions to expand the alternative analysis to reduce ecological, health, and safety risks during project development. This process embraces the concepts often called green engineering.

Green engineering is the early application of environmentally favorable engineering project concepts in the design and development of processes, products, and facilities. Embracing the basic concepts of pollution prevention and waste minimization, green engineering considers innovative and alternative technologies. The project should be feasible and economical, while minimizing the generation of air, water, and solid-waste pollution at the source, as well as the risk to public health and the environment.

The U.S. Environmental Protection Agency has a Design for the Environment (DfE) program that includes a life-cycle assessment. This process provides a tool to examine the environmental impacts of products over their entire life cycle from materials acquisition to manufacturing, use, and disposition. Design for the Environment provides decision-makers with information, tools, and incentives to make informed decisions that integrate risk, performance, and cost issues. The life-cycle analysis evaluates impacts on human health, atmospheric resources/air quality, water quality, ecological health, and natural-resource consumption.

Education of professionals. Traditionally, environmental engineers drew their basic education and training from civil engineering programs. In order to broaden their perspective and capabilities, contemporary environmental engineers undertake course work and postgraduate training in a number of professional areas, including biology, epidemiology, chemical engineering, chemistry, and hydrology. Since environmental engineers deal with sensitive public issues, training in public education, risk assessment, public policy, and social sciences is desirable.

After satisfying experience and testing requirements, the environmental engineer obtains

professional engineering registration. Professional associations of interest to the environmental engineer include the American Academy of Environmental Engineers, the Air and Waste Management Association, the American Water Works Association, the Water Environment Federation, and the Solid Waste Association of North America.

Professional activities. Principal areas of employment for practicing environmental engineers include consulting, industry, and government. Other environmental engineers are academic or research faculty, or they direct the development and production of pollution-control equipment.

Consulting environmental engineers apply their knowledge and experience under contract to public and private clients. Their services may range from studies to preliminary design, final design, construction management, and operation and maintenance services. Specific areas of service include site evaluations, environmental impact studies and assessments, assistance in obtaining permits, and expert witness testimony.

Within the industrial sector, there is considerable demand for environmental professionals to achieve and maintain compliance with environmental statutes and regulations. Increasing public sensitivity to the environment and demands for a clean environment have expanded industry's need for environmental engineering technology.

Environmental engineers in public service provide technical expertise in all levels of government. Though most publicized for their role in the development of regulations and in enforcement, they also are involved in numerous research and development programs and technology transfer activities.

The principal environmental engineering specialties are air-quality control, water supply, wastewater disposal, stormwater management, solid-waste management, and hazardous-waste management. Other specialties include industrial hygiene, noise control, oceanography, and radiology.

Air-quality control. The term air pollution is used to describe the presence in the atmosphere of one or more contaminants in quantities, and with characteristics, that will be injurious to, or unreasonably interfere with, public health and welfare or other natural environmental processes. The extent of air-pollution problems ranges from relatively small areas, such as an industrial park impacted by one or more emission sources, to urban areas impacted by a number of contaminant sources.

Contaminants are categorized as particulate matter and gases and their associated forms, including dust, smoke, fumes, mist, and vapor. The primary gaseous air contaminants are carbon monoxide, hydrocarbons, nitrogen oxides, and sulfur oxides.

Meteorological and topographical factors contribute to the creation and continuation of air pollution under specific site conditions. Temperature inversions prevent upward diffusion, and very low wind speeds allow emissions to remain near their source. Some terrains cause emissions to follow specific patterns from one area to another.

Generally, sources of air contaminants may be classified as stationary, mobile, or fugitive. Respectively, they are attributed to point sources such as industrial stack emissions, transportation activities such as automobile emissions, and uncontrolled (fugitive) sources such as windblown dusts from stockpiles.

The environmental engineer is instrumental in developing particulate and gas controls for all sources of air contaminants. Source control is the first abatement method considered. Controls for particulates include settling chambers, inertial separators, wet scrubbers, and fabric filters. Gas controls include absorption, adsorption, condensation, flaring, and incineration. Other areas of practice address acid rain issues, fugitive emissions, odor control, indoor air quality, and noise abatement.

Indoor air quality is of increasing concern for health officials and the public. Industrial hygienists and various engineering disciplines are instrumental in identifying and controlling indoor air pollution. Environmental engineering principles can be used in technical assessments for identifying air pollutants and selecting source controls.

Indoor air quality control measures are very source-specific. The basic abatement strategies are source control, ventilation improvements, and air cleaners. Typically, source control is the most cost-effective abatement measure. *See* AIR POLLUTION; AIR POLLUTION, INDOOR.

Water supply. Historically, the environmental engineer has found methods to provide ample quantities of quality drinking water for domestic use as well as quality water for commercial and industrial uses. Water-supply issues include demand projections, quality requirements, surface-water and ground-water source evaluations, ground-water production, surface-water collection and storage, surface-water treatment, saline-water treatment, nonconventional water production, and treated-water distribution. *See* WATER RESOURCES; WATER SUPPLY ENGINEERING.

The typical surface-water treatment plant uses chemicals for enhancing removal of suspended solids and for disinfection. Physical treatment processes include simple settling and filtration. In sequence, the unit processes are rapid mix, coagulation, flocculation, sedimentation, filtration, and disinfection. Auxiliary systems are needed for chemical feed facilities and for sludge handling. Processes that are more specialized include carbon adsorption, ion exchange, and softening. *See* WATER SOFTENING; WATER TREATMENT.

Treatment of other sources generally requires site-specific determination of raw-water quality. Often, the quantity of demand will influence the cost-effective selection of treatment processes. Ground-water supplies require well development and treatment, such as aeration, softening, and disinfection. Likewise, brackish and saline waters require site-specific determination of treatment processes. Typical processes include membrane technology, such as reverse osmosis and electrodialysis. *See* RAW WATER; WATER DESALINATION.

The water distribution system includes service and distribution lines, transmission mains, and storage facilities. Elevated storage tanks with gravity distribution and ground storage tanks with distribution pumping are designed to provide the quantity and pressure required to satisfy system demands.

Wastewater disposal. Wastewater is the combination of liquid- and water-transported wastes from homes, commercial buildings, industrial facilities, and institutions, along with any ground-water infiltration and surface-water and stormwater flow that may enter the sewer system. As a minimum measure, treatment is required for suspended solids and dissolved organics. Special processes may be necessary to achieve removal of specific pollutants, such as phosphorus from a municipal source or heavy metals from an industrial plating facility. *See* SEWAGE COLLECTION SYSTEMS; SEWAGE DISPOSAL; WASTEWATER REUSE.

Minimum levels of treatment have been established by regulation. For example, in the United States 85% removal of oxygen-demanding organics and suspended solids followed by disinfection is the minimum level of treatment for domestic wastewaters. Additional treatment is dictated by the assimilative capacity of the receiving stream and by downstream water uses. *See* SEWAGE TREATMENT.

Physical processes are used to remove suspended solids that may damage or interfere with subsequent pumping and treatment units. Screens remove debris and other large solids, and gravity or aerated grit chambers capture sandy matter. Normally, gravity sedimentation is used to remove finer (organic) suspended solids. For special applications, centrifugation, dissolved air flotation, and filtration are used to remove suspended solids. *See* CENTRIFUGATION; FILTRATION; FLOTATION.

Dissolved organics generally are treated with biological processes. The more common systems are aerobic (with oxygen) and include aerobic-pond, trickling-filter, and activated-sludge processes. Concentrated wastes, such as primary sludges, or high-strength industrial wastewaters, such as meat-processing or brewery wastes, are considered for anaerobic (without oxygen) treatment processes. Sludges, principally from biological processes, require special handling. The sequence of processes includes stabilization, conditioning, dewatering, drying, and residual disposal. Land application and landfilling are the most practiced means of final disposal. Special concerns for land-applied and composted sludges arise because of the concentrating of contaminants, such as heavy metals, and presence of pathogens in these sludges. *See* SEWAGE SOLIDS.

Of the many other types of treatment process, one of the most important is natural systems, which historically included pond systems and, later, various modes of land application. Through technology, natural and constructed wetlands can provide high-quality effluents. *See* WETLANDS.

Stormwater management. While the study of stormwater management includes all elements of the hydrologic cycle, it focuses on how humans affect the production, movement, and control of surface runoff. In a natural system, the rate of surface runoff is controlled by the rainfall rate, soil conditions, vegetation, and subsurface geology. *See* HYDROLOGY.

Urbanization creates large impervious areas that increase the quantity and peak rate of runoff. Rainfall then washes deposited materials directly into surface waters, causing stream pollution. Organics create oxygen demands, nutrients accelerate lake eutrophication, and heavy metals accumulate in bottom sediments. *See* EUTROPHICATION; STREAM POLLUTION; WATER POLLUTION.

Environmental engineers apply modern stormwater management practices to use natural and engineered systems to minimize environmental damage. A complete stormwater management program contains many elements, including on-site infiltration and detention, collection and transport systems, regional flood control, and major stream-channel improvements.

In the past, sanitary wastes and stormwater were collected in the same (combined) sewer. During heavy rainfalls, the sewers would overflow, creating water pollution problems. Many of these combined sewers have been separated, while the remaining overflows are treated as point sources. Stormwater management practices are used to lessen the requirements and costs for combined sewer overflow treatment facilities.

Solid-waste management. The collection, transport, processing, and disposal of solid wastes is an important area of environmental engineering. Solid wastes are those materials that are deemed by their owner to possess no value and are discarded. They are generated by almost every activity, and the amount varies by source, season, geography, and time. As land becomes more limited and regulation increases, the solid-waste generator and handler must employ new and improved technology to reduce the quantities of materials requiring disposal. Furthermore, recovery and reuse are important elements of solid-waste management.

Historically, solid-waste disposal consisted of open dumping. However, the modern method of disposal uses double-lined landfills with collection and controls for gases and leachate. Other disposal means include composting and various incineration processes, which also may be used for co-disposal of wastewater treatment sludges. These disposal means typically require controls for created pollutants, such as leachate and odor from compost operations, and chemical and particulate emission from incinerator combustion.

Recovery and reuse are practiced widely. Source or central-facility separation is used for a variety of products, including paper, glass, plastics, ferrous metals, and nonferrous metals. Also, refuse-derived fuels may be used for energy production, and yard wastes may be composted to produce a humus soil conditioner.

Integrated solid-waste management plans include recovery and reuse as essential elements. Source reduction techniques and cost-effective recovery of solid wastes for reuse in industrial production

improve energy recovery and production, provide direct economic benefits, and lessen the overall requirements for solid-waste processing and disposal.

Source reduction is a priority pollution prevention measure and includes design, manufacture, or use of materials to reduce the amount and/or toxicity of materials before they enter the solid-waste management process. Examples include managing nonproduct organic wastes such as food scraps and yard trimmings, extending the useful life of products to postpone disposal, designing packaging to reduce the quantity for waste disposal, reusing existing products or packaging, and segregating and collecting waste materials onsite for recycle at residential, commercial, and manufacturing generator sites.

Principal energy recovery technologies are incineration furnace walls with integrated boiler tubes (waterwalls); modular incineration with heat recovery boilers or heat exchangers; refuse-derived fuels for co-fired or dedicated boilers; pyrolysis into gaseous, liquid, or solid fuels; and anaerobic digestion or landfill gas recovery for methane fuel development.

Hazardous-waste management. Whether they are pesticides from agricultural lands, gasoline leaks from service stations, heavy metals from plating solutions, medical wastes from hospitals, or radioactive wastes from nuclear power plants, hazardous wastes are present throughout the world as by-products of growth in developing nations. They pose unreasonable risks to human health and safety, property value and use, and all other components of the environment.

Hazardous-waste treatment involves liquid-waste treatment, solid-waste treatment, solidification and stabilization, thermal destruction, and land disposal. Remedial action is characterized as surface-water control, air-pollution control, or in-place treatment.

A significant portion of the technology for the treatment and disposal of hazardous wastes is refinement or adaptation of proven practices in air-quality control, wastewater treatment, and solid-waste management. Also, the environmental engineer must understand hydrogeology to assess the subsurface disposition of hazardous wastes. *See* AQUIFER.

At a hazardous-waste treatment or disposal site, the routes of the waste into the environment must be obstructed. Principal routes include surface-water contamination from runoff or overflows; groundwater contamination from leaks or leachate; air contamination from open burning, evaporation, or fugitive dusts; fire and explosion; and health risks from human contact.

U.S. Environmental Protection Agency regulations define hazardous wastes in specific lists based on characteristics of ignitability, corrosivity, reactivity, and an analytic procedure known as EP toxicity. Household wastes, domestic sewage, and certain other wastes are excluded from these regulations. In the United States, regulations for hazardous wastes include requirements of a manifest system for tracking wastes from generators through storage, transport, treatment, and disposal. In addition to treatment and disposal, it is necessary to provide the means to satisfy these regulatory controls as well as means of pollution prevention using waste reduction through process changes, including recovery and reuse. *See* CONSERVATION OF RESOURCES; HAZARDOUS WASTE; RADIOACTIVE WASTE MANAGEMENT.

Robert A. Corbitt

Bibliography. American Water Works Association et al., *Water Treatment Plant Design*, 4th ed., 2004; R. A. Corbitt (ed.), *Standard Handbook of Environmental Engineering*, 2d ed., 1999; H. M. Freeman (ed.), *Standard Handbook of Hazardous Waste Treatment and Disposal*, 2d ed., 1997; Metcalf and Eddy, Inc., *Wastewater Engineering: Treatment, Disposal, and Reuse*, 4th ed., 2002; J. Pichtel, *Waste Management Practices: Municipal, Hazardous, and Industrial*, 2005; M. P. Wanielista and Y. A. Yousef, *Stormwater Management*, 1993.

Environmental fluid mechanics

The study of the flows of air and water, of the species carried by them, and of their interactions with geological, biological, social, and engineering systems in the vicinity of a planet's surface. The environment on the Earth is intimately tied to the fluid motion of air (atmosphere), water (oceans), and species concentrations (air quality). In fact, the very existence of the human race depends upon its abilities to cope within the Earth's environmental fluid systems.

Meteorologists, oceanologists, geologists, and engineers study environmental fluid motion. Weather and ocean-current forecasts are of major concern, and fluid motion within the environment is the main carrier of pollutants. Biologists and engineers examine the effects of pollutants on humans and the environment, and the means for environmental restoration. Air quality in cities is directly related to the airborne spread of dust particles and of exhaust gases from automobiles. The impact of pollutants on drinking-water quality is especially important in the study of ground-water flow. Likewise, flows in porous media are important in oil recovery and cleanup. Lake levels are significantly influenced by climatic change, a relationship that has become of some concern in view of the global climatic changes that may result from the greenhouse effect (whereby the Earth's average temperature increases because of increasing concentrations of carbon dioxide in the atmosphere). *See* AIR POLLUTION; GREENHOUSE EFFECT; WEATHER FORECASTING AND PREDICTION.

Scales of motion. Environmental fluid mechanics deals with the study of the atmosphere, the oceans, lakes, streams, surface and subsurface water flows (hydrology), building exterior and interior airflows, and pollution transport within all these categories. Such motions occur over a wide range of scales, ranging from eddies on the order of centimeters to large recirculation zones the size of continents. This range accounts in large part for the difficulties associated with understanding fluid motion within the environment. In order to impart motion (or inertia)

to the atmosphere and oceans, internal and external forces must develop. Global external forces consist of gravity, Coriolis, and centrifugal forces, and electric and magnetic fields (to a lesser extent). The internal forces of pressure and friction are created at the local level, that is, on a much smaller spatial scale; likewise, these influences have different time scales. The winds and currents arise as a result of the sum of all these external and internal forces.

Global motion is the largest scale (greater than 5000 km or 3000 mi); synoptic scale motion is the next largest (100–1000 km or 60–600 mi). Mesoscale motion occurs over regional areas (10–100 km or 6–60 mi); local motion is commonly referred to as microscale motion (less than 100 m or 300 ft). Humans live in the microscale associated with atmospheric motion, that is, the boundary layer of air that extends about 1 km (0.6 mi) above the Earth's surface. The terrain of the Earth as well as ocean surface conditions significantly affects both the microscale and mesoscale motion of the atmosphere, and hence weather conditions. The scales of motion range from eddies on the order of centimeters to huge masses extending over thousands of kilometers. *See* MESOMETEOROLOGY; MICROMETEOROLOGY.

El Niño, the periodic flow of warm waters along the western coast of South America, disrupts the coastal ocean and the upwelling of cold waters, producing large amounts of precipitation, along with widespread destruction of plankton, fish, and sea birds (which prey on the fish). Major El Niño events occurred in 1925, 1941, 1957–1958, 1972–1973, 1982–1983, and 1992. It has been determined that the events are caused by changes in the surface winds over the western tropical Pacific, which periodically release and drive warm waters eastward to the South American continent. *See* EL NIÑO.

The study of air pollution falls within the category of environmental fluid mechanics, because the air within the lower atmosphere steers (or advects) and diffuses pollutants. Atmospheric winds near the Earth's surface are generally turbulent and gusty, which helps to clear polluted areas; the velocity varies with altitude, local stability (level of turbulence), and roughness of the terrain. However, when the winds are calm, stagnant conditions can occur which subsequently prevent pollutants from being cleared from a city, resulting in high levels of bad air quality and smog. Of particular importance on the mesoscale level is acid rain (whereby rainfall removes sulfates and nitrates within the atmosphere), which has resulted in serious environmental damage. Likewise, mixing of pollutants into the upper atmosphere can cause long-term changes in the ozone layer (even though the causes, such as propellants within spray cans, may have been generated within the microscale layer). The 1991 explosion of Mount Pinatubo in the Philippines resulted in the discharge of many tons of particulates into the Earth's atmosphere; these particulates in turn acted as seed nuclei for precipitation, and were the cause of much of the flooding and climatic changes over the following few years. *See* WEATHER MODIFICATION.

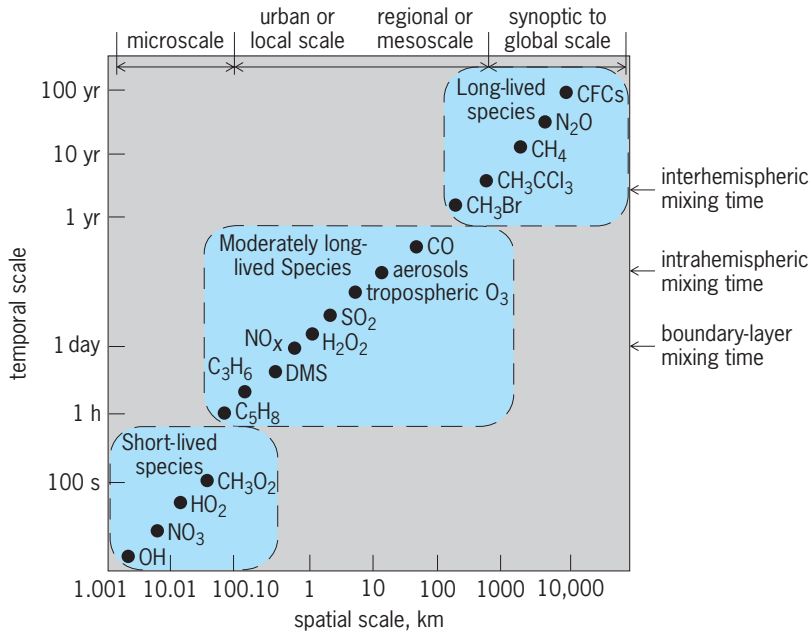

Fig. 1. Spatial and temporal scales of variability for atmospheric constituents. 1 m = 3.3 ft; 1 km = 0.6 mi. (*After J. H. Seinfeld and S. N. Pandis, Atmospheric Chemistry and Physics, Wiley, 1998*)

The various species found in the atmosphere have a wide range of lifetimes (residence time in the atmosphere). Species with short lifetimes have small spatial scales; those with lifetimes of years have spatial scales comparable to the entire atmosphere (**Fig. 1**). For example, the hydroxyl radical (OH) has a very short lifetime and small scale; methane (CH_4) has a lifetime of nearly 10 years and can become mixed over the entire Earth.

Governing equations. The foundations of environmental fluid mechanics lie in the same conservation principles as those for fluid mechanics, that is, the conservation of mass, momentum (velocity), energy (heat), and species concentration (for example, water, humidity, other gases, and aerosols). The differences lie principally in the formulations of the source and sink terms within the governing equations, and the scales of motion. These conservation principles form a coupled set of relations, or governing equations, which must be satisfied simultaneously. The governing equations consist of nonlinear, independent partial differential equations that describe the advection and diffusion of velocity, temperature, and species concentration, plus one scalar equation for the conservation of mass. In general, environmental fluids are approximately newtonian, and the momentum equation takes the form of the Navier-Stokes equation. An important added term, neglected in small-scale flow analysis, is the Coriolis acceleration, $2\Omega \times V$, where Ω is the angular velocity of the Earth and V is the flow velocity. *See* CONSERVATION LAWS (PHYSICS); CONSERVATION OF ENERGY; CONSERVATION OF MASS; CONSERVATION OF MOMENTUM; CORIOLIS ACCELERATION; DIFFERENTIAL EQUATION; DIFFUSION; FLUID-FLOW PRINCIPLES; NAVIER-STOKES EQUATION; NEWTONIAN FLUID.

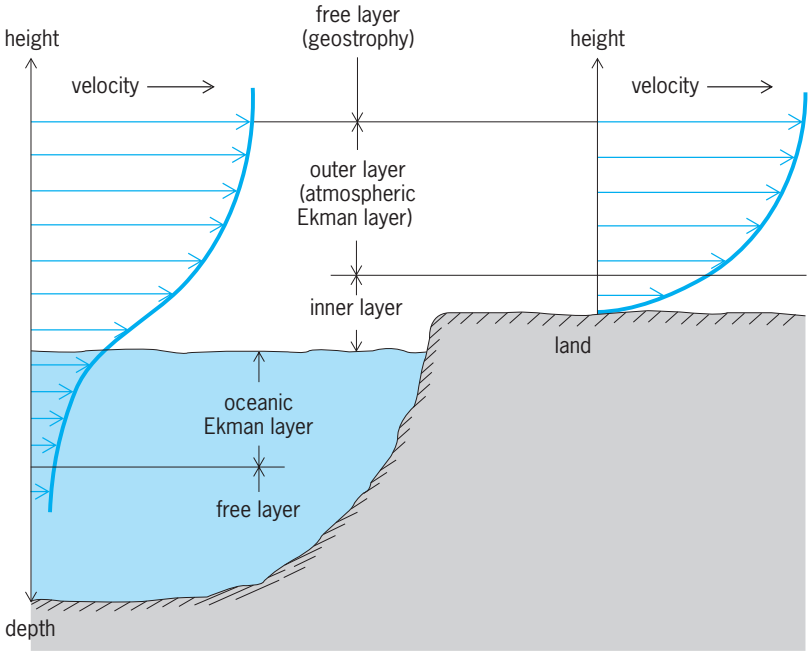

Fig. 2. Velocity distribution in the atmosphere and ocean. (*After S. Eskinazi, Fluid Mechanics and Thermodynamics of Our Environment, Academic Press, 1975*)

Driving mechanisms of flow. The mechanisms which drive the flow patterns in the atmosphere and oceans are vastly different. The atmosphere is thermodynamically driven, with the major source of energy coming from solar radiation. Short-wave radiation traverses the air and becomes partially absorbed by the land and oceans, which reemit the radiation at longer wavelengths. Long-wave radiation heats the atmosphere from below, creating convection currents in the atmosphere. *See* ATMOSPHERIC GENERAL CIRCULATION.

In the oceans, periodic gravitational forces of the Sun and Moon generate tides; in addition, the ocean surface is affected by wind stress that drives most of the ocean currents. Local differences between the air and sea temperatures generate heat fluxes, evaporation, and precipitation, which ultimately act as thermodynamical forces that create or modify wind-driven currents. *See* OCEAN CIRCULATION.

Environmental layers. Fortunately, not every term in the Navier-Stokes equation is important in all layers of the environment. The horizontal component of the motion is usually the most significant and is subjected to maximum frictional forces at atmosphere-ocean interfaces. This frictional force causes the formation of a boundary layer in which the velocity of air at the surface of the Earth is zero (relative to the Earth), and the velocity at the surface of the ocean is a minimum equal to the surface velocity of the water. The ocean current is primarily generated by the wind; hence, the water velocity at the surface is a maximum and decreases in depth, again as a result of frictional forces. In both instances, frictional forces cause strong velocity gradients and vorticity (rotation) within the boundary layer. **Figure 2** shows the velocity distribution in the atmosphere and ocean. *See* BOUNDARY-LAYER FLOW.

The thickness of the atmospheric boundary layer varies with the wind speed, degree of turbulence, and type of surface. For atmospheric flows, the layer is on the order of 1 km (0.6 mi) thick; within the ocean, it may be 30 m (100 ft) thick. Beyond this layer, the environmental flow is typically considered to be viscous-free (without turbulent shear), or inviscid. The rougher the terrain, or the larger the surface obstructions, the thicker the atmospheric boundary layer becomes, and the more gradual the increase of velocity with height (**Fig. 3**). The influence of the ground on the wind profile extends from a few hundred meters to over 500 m (1640 ft), depending on the roughness of the surface. Above this height, velocity is established from upper level meteorology. The wind speed is proportional to some power of height (empirically determined from experiments) above the surface.

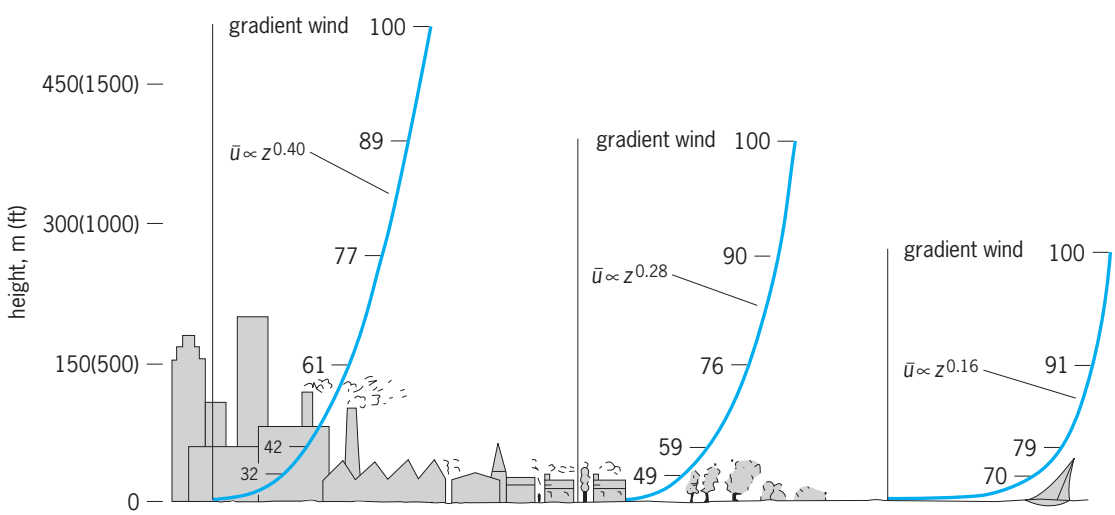

Fig. 3. Atmospheric boundary-layer profiles (plots of average wind speed $\bar{u}$ versus height z) over different terrains. Wind speeds are expressed as percentages of the upper level wind (referred to as the gradient wind) above the boundary (or surface) layer. (*After E. J. Plate, Aerodynamic Characteristics of Atmospheric Boundary Layers, AEC Critical Review Series, U.S. Department of Energy, 1971*)

Length, velocity, and time scales in the Earth's atmosphere and oceans*

Phenomenon	Length scale, km (mi)	Velocity scale, m/s (mi/h)	Time scale
Atmosphere			
Sea breeze	5–50 (3–30)	1–10 (2–20)	12 h
Mountain waves	10–100 (6–60)	1–20 (2–40)	Days
Weather patterns	100–500 (60–300)	1–50 (2–100)	Days to weeks
Prevailing winds	Global	5–50 (10–100)	Seasons to years
Climatic variations	Global	1–50 (2–100)	Decades
Ocean			
Internal waves	1–20 (0.6–12)	0.05–0.5 (0.1–1)	Minutes to hours
Coastal upwelling	1–10 (0.6–6)	0.1–1 (0.2–2)	Several days
Large eddies, fronts	10–200 (6–120)	0.1–1 (0.2–2)	Days to weeks
Major currents	50–500 (30–300)	0.5–2 (1–4)	Weeks to seasons
Large-scale gyres	Basin scale	0.01–0.1 (0.02–0.2)	Decades

*From B. Cushman-Roisin, *Introduction to Geophysical Fluid Dynamics*, Prentice Hall, 1994.

Because there are no shear stresses, the motion of the inviscid layer is governed only by the advection, pressure, and body-force terms. In atmospheric flows, the rotation of the Earth strongly influences this layer of flow, generally referred to as the geostrophic layer. Just above the surface, the mean velocities are small; the advection terms and the Coriolis force (which depends on the velocity) are negligible compared to the shear forces (viscous terms) which appear to be constant in this inner layer. However, within the outer, or Ekman, layer advection is still negligible and the viscous forces are small; this part of the boundary layer is in equilibrium with the Coriolis, pressure, and Reynolds stresses (turbulence). The **table** shows typical scales of length, velocity, and time for both atmospheric and oceanic motions. Oceanic motions are slower and more confined, and tend to evolve more slowly, than atmospheric motions.

Relative importance of terms. The key to being able to obtain solutions to the Navier-Stokes equation lies in determining which terms can be neglected in specific applications. For convenience, problems can be classified on the basis of the order of importance of the terms in the equations utilizing nondimensional numbers based on various ratios of values. The Rossby number (Ro) is the ratio of the advection (or inertia) forces to the Coriolis force, $\mathrm{Ro} = V/L\Omega$, where V is velocity, Ω is the Earth's angular velocity, and L is a specified reference length. When the Rossby number is much less than 1, the inertia forces become insignificant, implying that these types of flows are more geostrophic. The ratio of the viscous to Coriolis forces is defined by the Ekman number, $\mathrm{Ek} = \mu/\rho\Omega H^2$, where ρ is density, μ is viscosity, and H is a vertical reference height (or thickness). The ratio of inertia to viscous forces is referred to as the Reynolds number, $\mathrm{Re} = \rho VL/\mu$. The Rossby number divided by the Ekman number yields the Reynolds number, that is, $\mathrm{Re} = \rho VL/\mu = (V/\Omega L)(\Omega\rho H^2/\mu)(L^2/H^2) = (\mathrm{Ro}/\mathrm{Ek})(L/H)^2$. When the Rossby number is large and the Ekman number is small, the motion is geostrophic; when the Rossby number is small and the Ekman number large, an Ekman-type boundary layer develops. As the Reynolds number increases, the ratio of the flow velocity to viscosity increases (that is, the advection terms become more important than the viscous terms), with the flow eventually becoming turbulent. Since the Ekman number is generally small and the geometric ratio (L/H) is large (Ro is on the order of unity), the Reynolds number for geophysical flow is generally large and the flow turbulent. *See* DIMENSIONLESS GROUPS; GEOSTROPHIC WIND; REYNOLDS NUMBER; TURBULENT FLOW; VISCOSITY.

Measurements. Because of the scales of motion and time associated with the environment, and the somewhat random nature of the fluid motion, it is difficult to conduct full-scale, extensive experimentation. Likewise, some quantities (such as vorticity or vertical velocity) resist direct observations. It is necessary to rely on the availability of past measurements and reports (as sparse as they may be) to establish patterns, especially for climate studies. However, some properties can be measured with confidence.

Both pressure and temperature can be measured directly in the atmosphere and ocean with conventional instruments. In the ocean, depth is typically calculated from measured pressures obtained from instruments lowered into the sea. In the atmosphere, ground precipitation, radiative heat fluxes, and moisture content can be accurately measured. Likewise, the salinity of the ocean can be determined from electrical conductivity, and the levels monitored at shore stations. Concentration samples, collected at receptor sites over long periods of time, are examined to determine specific concentration levels and particulate sizes. These data are used to determine isopleth (concentration) levels and exposures over various atmospheric and oceanic conditions. Occasionally, inert tracer gases are released into the atmosphere to determine wind directions as well as atmospheric diffusion (turbulence levels) and plume trajectories.

Vector quantities such as horizontal winds and currents are typically measured by using anemometers and current meters. Anemometers atop buildings and towers, and current meters attached to mooring lines at fixed depths, offer fine temporal readings but are too expensive to adequately cover large areas. Instruments are routinely deployed on drifting platforms in the ocean, and balloons

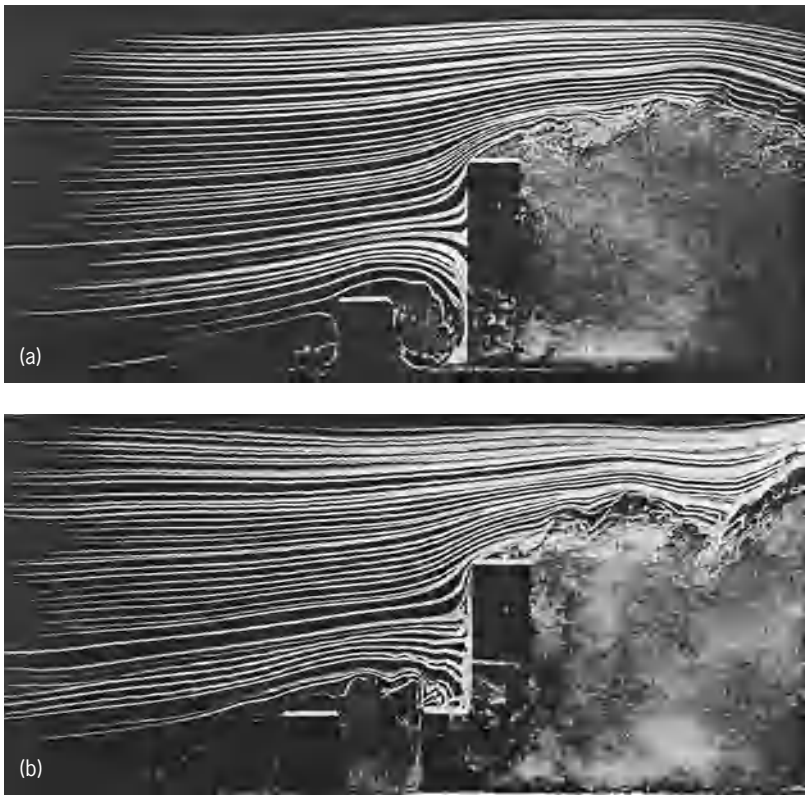

Fig. 4. Flow around two models of a tall building showing how minor design modifications can make a large difference in wind velocity at the pedestrian level. (*a*) Flow with vortex between the two buildings. (*b*) Flow with vortex removed by a slight change in the shape of the tall building. (*From H. Thomann, Wind effects on buildings and structures, Amer. Sci., 63:278–287, 1975*)

are released in the atmosphere. (However, such measurements are mixed in time and space.) Measuring the three-dimensional velocity components simultaneously and obtaining meaningful three-dimensional heat fluxes is difficult, and essentially relegated to small-scale laboratory experiments.

Advances utilizing satellite imagery, Doppler radar, acoustic sounding, and lidar (laser) have made it possible to obtain highly detailed data, including turbulence information, over much broader spatial distances. Doppler radar has yielded three-dimensional velocity data and rotational characteristics within thunderstorms that can be used reliably to predict the onset of tornadoes. *See* DOPPLER RADAR; LIDAR; METEOROLOGICAL INSTRUMENTATION; METEOROLOGICAL SATELLITES.

Modeling. There are two types of modeling strategies: physical models and mathematical models. Physical models are small-scale (laboratory) mock-ups that can be measured under variable conditions with precise instrumentation. Such modeling techniques are effective in examining wind effects on buildings and species concentrations within city canyons (flow over buildings; **Fig. 4**). Generally, a large wind tunnel is needed to produce correct atmospheric parameters (such as Reynolds number) and velocity profiles. Mathematical models (algebra- and calculus-based) can be broken down further into either analytical models, in which an exact solution exists, or numerical models, whereby approximate numerical solutions are obtained using computers. *See* WIND TUNNEL.

By far the most interesting and widely used models are the numerical models. The reason for their popularity is that it is possible to model more of the actual physics of the flow, that is, solve the Navier-Stokes equation, rather than make assumptions and eliminate key components of the physics just to obtain a solution. Although the Navier-Stokes equation is nonlinear, the partial differential equation can now be solved with some measure of confidence and reliability. In many instances dealing with environmental flows, the use of supercomputers is required. *See* SUPERCOMPUTER.

Numerical methods. Several broad classes of solution techniques are employed to solve the various derivatives and terms of the Navier-Stokes equation. The most common and widely used numerical methods are finite difference schemes (which are based on the use of truncated Taylor-series expansions); finite element schemes (which use an integral approach with local weighting and basis, or shape, functions); spectral methods (in which dependent variables are transformed to wave-number space by using a global basis function, such as the Fourier transform); pseudospectral methods (which use truncated spectral series to approximate derivatives); interpolation techniques (whereby polynomials are used to approximate the dependent variables in one or more spatial directions); and particle methods (which use lagrangian particles whose trajectories are calculated within a conventional eulerian grid). Such numerical models depend strongly on boundary and initial conditions; care must be exercised to correctly initialize and specify all variables at the boundaries of the computational model. All these schemes except the particle methods require knowledge of properties such as viscosity, dispersion coefficients, and thermal conductivity; particle methods require no constitutive models for particle viscosity or thermal conductivity, but do require a large number of particles for an accurate description of the flow field. The most popular modeling approaches are the finite difference, finite element, and interpolation schemes, especially for mesoscale and synoptic-scale simulations. *See* COMPUTATIONAL FLUID DYNAMICS; FINITE ELEMENT METHOD; INTERPOLATION; NUMERICAL ANALYSIS; SIMULATION.

Capabilities. The continuing rapid improvement in computational hardware has made it possible to model more complicated problems and include more physics (or mathematical terms) in the governing equations. Simulations of environmental fluid flow over microscale and mesoscale regions without simplifications of the equations of motion are now fairly common. Arrays consisting of millions of nodes can be calculated within a few hours on supercomputers, and three-dimensional graphical displays can be generated on work stations. By using satellite, radar, and conventional surface observations as input data to meteorological models, reasonably accurate local forecasts can be made for up to several days. Advances in numerical techniques as well as

computer hardware will continue, making it possible to perform more detailed calculations over broader expanses with improved accuracy over longer forecast periods.

Examples. An example of the simulation of fluid flow over a building complex is shown in **Fig. 5**. The three-dimensional equations of motion and species transport were solved using an adaptive finite element method. The flow field at the *x-y* midplane of the three-dimensional model domain (Fig. 5*a*) shows the development of a series of eddies as the flow moves downstream of the buildings. The flow field at the *x-z* vertical midplane (Fig. 5*b*) shows the formation of a large eddy between the two sets of buildings. Lagrangian particles, introduced upstream of the building array, clearly show the dispersion of concentration over and around the buildings (Fig. 5*c*).

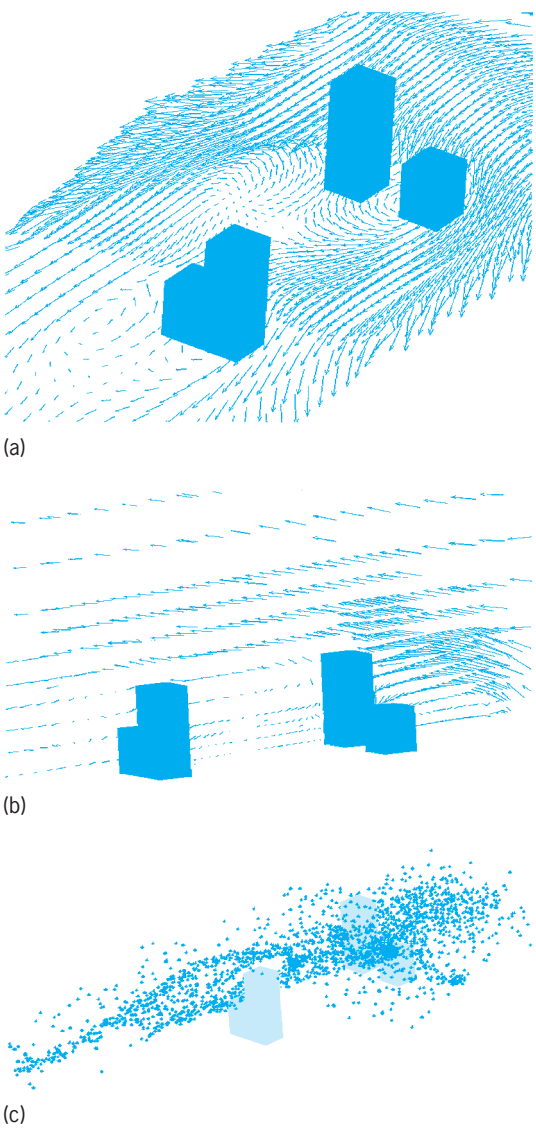

Fig. 5. Numerical simulation of flow over a building complex, carried out on a SGI-Cray Origin 2000 parallel computer. (*a*) Wind field in the horizontal (*x-y*) midplane. (*b*) Wind field in the vertical (*x-z*) midplane. (*c*) Lagrangian particles depicting species transport. (*After D. W. Pepper et al., eds., Development and Application of Computer Techniques to Environmental Studies, WITPress, 1998*)

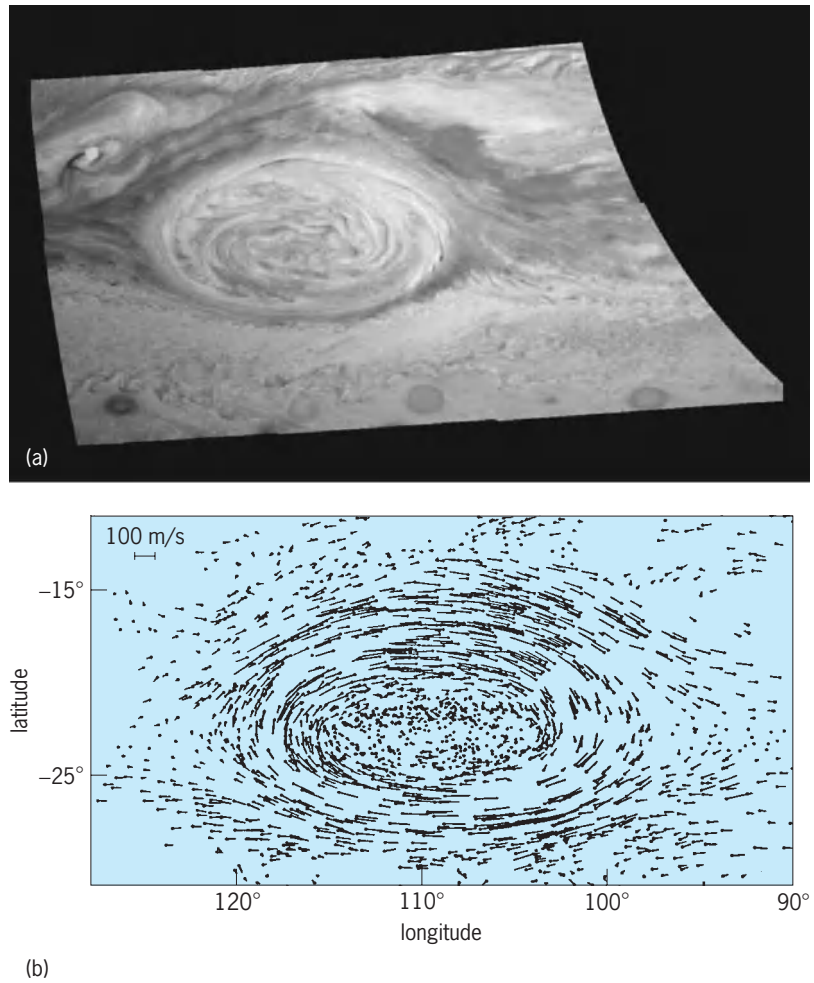

Fig. 6. Great Red Spot of Jupiter. (*a*) Image from *Galileo* spacecraft (*NASA*). (*b*) Velocity field in and around the Great Red Spot obtained by tracking small cloud formations in sequential *Voyager 1* images. (*From T. E. Dowling and A. P. Ingersoll, Potential vorticity and layer thickness variations in the flow around Jupiter's Great Red Spot and White Oval BC, J. Atmos. Sci., 45:1380–1396, 1988*)

Particles are being pulled into the large eddy between the two sets of buildings.

The flow circulation around the Great Red Spot of Jupiter (**Fig. 6***a*) is an example of the turbulent nature of fluid flow on a large scale. The velocity field in and around the Great Red Spot was obtained by tracking small cloud features over time (Fig. 6*b*). *See* FLUID MECHANICS; JUPITER.

Darrell W. Pepper

Bibliography. B. Cushman-Roisin, *Introduction to Geophysical Fluid Dynamics*, 1994; T. N. Krishnamurti and L. Bounoua, *An Introduction to Numerical Weather Prediction Techniques*, 1996; D. W. Pepper and J. C. Heinrich, *The Finite Element Method: Basic Concepts and Applications*, 1992; D. W. Pepper, P. Zannetti, and C. Brebbia (eds.), *Development and Application of Computer Techniques to Environmental Studies*, WITPress, 1998; M. L. Salby, *Fundamentals of Atmospheric Physics*, 1996; J. H. Seinfeld and S. N. Pandis, *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*, 1998; P. Zannetti, *Air Pollution Modeling*, 1990.

Environmental geology

The branch of geology that deals with the ways in which geology affects people. Examples of the effect of geology on human civilizations include (1) the ways that fertile soils develop from rocks and how these soils can become polluted by human activities; (2) how rocks and soils move down-slope to destroy roads, houses, and other human constructions; (3) sources of surface and subsurface water supplies and how they become polluted; (4) why floods occur where they do and how human activities affect floods; (5) locations of earthquakes and volcanic eruptions and the dangers they pose; (6) location of mineral resources such as copper, oil and gas, and uranium, and how mining these resources can pollute the environment; (7) how human activities can pollute the atmosphere and cause global warming, sea-level rise, and ozone depletion.

Soil. Most environmental concerns center on three materials essential to human life: soil, water, and air. Soils develop from chemical reactions between rocks, water, and atmospheric gases such as oxygen and carbon dioxide. The most abundant solid product of these reactions is a variety of clay minerals, which are major contributors to agricultural fertility. Ions on clay surfaces and decaying organic matter are the sources of nutrient elements for growing plants. Clay minerals are also responsible for destructive events, such as when they absorb water and cause soil to swell, resulting in foundation cracking and other structural problems in buildings. Changes in clay minerals as they lie in the soil are also the cause of many types of landslides. *See* CLAY MINERALS; SOIL.

The major environmental problems affecting soil are erosion and pollution. Almost half of American farmland is losing fertile topsoil faster than it can be formed. Most of the erosion results from poor farming practices and overzealous logging, which deplete the soil of plants and root systems that hold the soil in place. Most soil pollution is caused by synthetic chemical pesticides which not only kill the pests but also are absorbed by plants, which are then eaten by people and other animals. *See* EROSION; SOIL CONSERVATION.

Water. The amount of water on the Earth's surface is constant and more than enough to supply human needs. Water problems are generally concerned with local supply and its pollution. Serious water shortages in the United States occur in many western states because of increasing populations, as in Phoenix, Las Vegas, and Los Angeles.

In areas where surface water supplies are inadequate, fresh water can commonly be obtained from underground sources. The water occurs in the open spaces between the grains in sedimentary rocks and has accumulated in the rocks over many thousands of years. When this water is withdrawn faster than it is naturally replenished by rain- and snowfall, the underground supply can become seriously depleted, as is occurring today in many western cities. Such depletion is also a serious problem in the heavily irrigated agricultural areas of the central United States, because most irrigation water is drawn from subsurface supplies. Removal of subsurface water can also result in ground subsidence, a serious problem in central California, Houston, and New Orleans. *See* GROUND-WATER HYDROLOGY.

Pollution. Most water pollution problems in the United States are caused by synthetic organic compounds and heavy elements released in wastewater discharges, many by the military as well as by chemical plants. Oil refineries are also major polluters. Other sources of pollution include leaking sanitary landfills, pesticides in runoff from farms, and mixtures of sulfuric acid and cyanide from mining operations. *See* HAZARDOUS WASTE; WATER POLLUTION.

Hazardous waste water can be disposed of by underground injection, incineration, evaporation from open pits, or export to other areas. The most common domestic method is underground injection, but because the underground geology is inadequately known, liquid hazardous waste injected deep underground has seeped into underground water supplies in about half the states. About 20% of hazardous waste is incinerated, but the Environmental Protection Agency (EPA) estimates that 60% could be disposed of by this safer method. Incineration is more expensive than subsurface disposal, however. Disposal by evaporation from surface impoundments is very unsafe and has caused ground-water contamination in about three-quarters of the 50 states.

Floods. Floods are the most commonly experienced natural calamities and are the cause of about 60% of declared disasters. Flooding may be caused by hurricanes from the ocean or excessive water from landward areas. Hurricanes cannot be controlled, and the value of property destroyed by them has increased considerably over time because of people's desire to live near the seashore. Loss of life has decreased, however, because of improved forecasting and early warning. Flooding resulting from excessive rain on the land surface has been exacerbated by the human desire to live near river channels, to pave land surfaces and increase the rate of runoff, and to denude hill slopes by excessive logging. Without vegetation and porous soil, the rate of runoff of rainwater is increased and river channels cannot hold all the water they receive. The massive flooding in central China in 1998 was attributed to the stripping of vegetation over large areas.

Atmosphere. Concerns about air quality revolve around increases in the amount of carbon dioxide, decreases in the amount of ozone 10–20 mi above the surface, and pollution of the air that sustains life in the lower atmosphere. Increasing carbon dioxide levels result largely from the burning of coal and oil, the main sources of energy in industrial societies. The amount of carbon dioxide in the air has increased by nearly 30% during the past 150 years. Carbon dioxide absorbs radiation from the Earth's surface and has caused a measurable increase in global temperatures (greenhouse effect); this global warming may in turn result in rising sea levels (from melting glaciers), endangering coastal communities over the long term. *See* GREENHOUSE EFFECT.

The concentration of ozone in the stratosphere

absorbs most of the Sun's ultraviolet (UV) radiation, thus protecting people from excessive skin cancers, damaged immune systems, and other maladies. The thickness of the ozone layer has decreased significantly during the past 50 years because of the release to the atmosphere of chlorofluorocarbons (CFCs) which destroy ozone molecules. Production of CFCs has now nearly stopped worldwide, but the ozone layer will not restore itself completely for about another 50 years.

Noxious gases and particulates in the air result largely from combustion of coal and oil in automobiles and factories. It is the cause of many lung problems and premature deaths in industrial countries. *See* AIR POLLUTION.

Earthquakes and volcanoes. Earthquakes and volcanoes tend to be associated geographically and are prominent around the Pacific rim. Japan and the American west coast are particularly vulnerable, and both have suffered major quakes in recent years. In the United States, California is the most seriously affected by earthquakes, with major property damage and loss of life. Attempts to predict the timing and location of earthquakes have not yet been successful despite very large research efforts in many countries over several decades.

Dangers from volcanic eruptions include mud and rock flows, lavas, and acid gases. Most dangerous to people are the debris flows, which can move downslope at speeds as high as 50 mi/h for distances of tens of miles. Entire towns have been buried by such flow with much loss of life. Lava flows move slowly and hence cause few deaths, but their power is irresistible and property damage is extensive. Eruptions can sometimes be predicted, but surprises are common. *See* EARTHQUAKE; VOLCANO. Harvey Blatt

Bibliography. H. Blatt, *Our Geologic Environment*, Prentice Hall, Upper Saddle River, NJ, 1997; N. K. Coch, *Geohazards*, Prentice Hall, Upper Saddle River, NJ, 1995; T. E. Graedel and P. J. Crutzen, *Atmospheric Change*, W. H. Freeman, New York, 1993; P. Gralla, *How the Environment Works*, Ziff-David, Emeryville, CA, 1994.

Environmental management

The development of strategies to allocate and conserve resources, with the ultimate goal of regulating the impact of human activities on the surrounding environment. "Environment" here usually means the natural surroundings, both living and inanimate, of human lives and activities. However, it can also mean the artificial landscape of cities, or occasionally even the conceptual field of the noosphere, the realm of communicating human minds.

Approaches. Environmental management is a mixture of science, policy, and socioeconomic applications. It focuses on the solution of the practical problems that humans encounter in cohabitation with nature, exploitation of resources, and production of waste. In a purely anthropocentric sense, the central problem is how to permit technology to evolve continuously while limiting the degree to which this process alters natural ecosystems. Environmental management is thus intimately intertwined with questions regarding limiting economic growth, ensuring an equitable distribution of consumable goods, and conserving resources for future generations. Environmental management is a response to the increasing seriousness of the human impact on natural ecosystems. With a smaller global population base and a less pervasive use of technology, the environment might be able to recuperate on its own from human misuse, but it is now widely recognized that in many cases positive intervention is necessary if the environment is to recover.

There is, however, considerable disagreement about the course that such intervention should take, which has created a plurality of approaches to managing the environment. "Deep ecology" was born in the 1960s with the rise of movements that renounced technological development and decried the political basis of power and autocracy. However, "shallow" ecologists sought a compromise with those who argued that the solution to the world's environmental problems can come only through the generation of more technology. Environmental managers therefore fall within a broad spectrum that extends from conservationists to technocrats, from those who would limit human interference in nature to those who would increase it in order to guide natural processes along benign paths. Hence both conservationists and developers are represented. It is hoped that they will come together over the need to make economic development sustainable, without it being undermined by long-term damage to resources and habitats. This is the intention of the United Nations Convention on Environment and Development (the process that began at the Earth Summit in Rio de Janeiro in 1992), though underfunding and lack of commitment at the national level have severely limited the extent to which it has changed the global course of environmental management.

Participants. Participants in the process of environmental management fall into seven main groups: (1) governmental organizations at the local, regional, national, and international levels, including world bodies such as the United Nations Environment Programme and the U.N. Conference on Environment and Development; (2) research institutions, such as universities, academies, and national laboratories; (3) bodies charged with the enforcement of regulations, such as the U.S. Environmental Protection Agency; (4) businesses of all sizes and multinational corporations; (5) international financial institutions, such as the World Bank and International Monetary Fund; (6) environmental nongovernmental organizations, such as the World Wildlife Fund for Nature; and (7) representatives of the users of the environment, including tribes, fishermen, and hunters. The agents of environmental management include foresters, soil conservationists, policy-makers, engineers, and resource planners. The main link between these diverse groups of people is the need for accountability in the use of nature's riches. However, though there is much collaboration, relationships are often adversarial as objectives differ among the groups.

Intellectually, environmental management has assumed the status of a multidisciplinary academic field dedicated to furthering the human stewardship of natural resources. Biologists participate by virtue of their links with ecology and their interest in flora, fauna, and habitats. Geographers have a long-standing involvement in the ecology of people's relationships with their surroundings—in other words, in the two-way interaction between human communities and the landscapes that offer opportunities to prosper but that limit activities. Economists have become involved through the need to value environmental goods and services, assess the costs of pollution, and calculate materials balances. Engineers have created the field of ecotechnology, which is dedicated to the practical restoration of degraded or polluted environments. Finally, political scientists are the source of much environmental policy and regulatory know-how. Though these and other disciplines have a hand in environmental management, the links between them are complex, and as a result the field is remarkably diffuse. Indeed, it tends to lack a core of integrated concepts that might give it a clear popular identity. Some common themes of environmental management are as follows:

1. Bilateral and multilateral environmental treaties (transboundary ecological management)
2. Design and use of decision-support systems (practical utilization of environmental data; expert systems for environmental management)
3. Environmental policy formulation and enactment (participatory planning and public consultation regarding environmental programs)
4. Estimation, analysis, and management of environmental risk (risk perception and communication studies)
5. Formulation of environmental regulations (for dumping of wastes, emission of pollutants, and extraction of resources; monitoring and policing compliance)
6. Impacts and management of recreation and tourism (design and implementation of environmentally friendly "ecotourism" programs)
7. Natural resource conservation (designation and management of parks, preserves, and other protected areas; designation and protection of wilderness areas)
8. Positive environmental economics (economic justifications for investment in environmental protection)
9. Promotion of positive environmental values by education, debate, and information dissemination
10. Reduction of adverse environmental impacts
11. Resource evaluation and management
12. "Scoping" and investigation of environmental impacts (design of policies, norms, and procedures to limit impacts)
13. Strategies, methods, and programs for the rehabilitation of damaged environments (postpollution clean-up processes)

Improvements. The need to improve management of the environment has given rise to several new techniques. There is environmental impact analysis, which was first formulated in California and is codified in the U.S. National Environmental Policy Act (NEPA). Through the environmental impact statement, it prescribes the investigatory and remedial measures that must be taken in order to mitigate the adverse effects of new development. In this sense it is intended to act in favor of both prudent conservation and participatory democracy.

Another technique is environmental auditing, which uses the model of the financial audit to examine the processes and outcomes of environmental impacts. It requires value judgments, which are usually set by public preference, ideology, and policy, to define what are regarded as acceptable outcomes. Audits use techniques such as life-cycle analysis and environmental burden analysis to assess the impact of, for example, manufacturing processes that consume resources and create waste.

The atmosphere, surface and subsurface waters, growing plants, minerals, and so on, are sometimes considered to be beneficial resources. The process of exploitation often involves risks to the user, and if these are magnified the resource may turn into a hazard—for example, when excesses of water generate destructive floods. Similarly, pollution has been defined as "a resource in the wrong place at the wrong time." Thus the process of environmental management can be considered one of limiting resource usage to its more benign forms, thus reducing risks and hazards, and using human ingenuity to transform pollutants into recycled resources. This involves some thorny problems. For example, no human activity is completely devoid of risk, and few environmental pathologies have ever been eliminated. Moreover, even full-scale recycling is not without costs; for example, the energy required to reclaim waste materials will usually be generated at the expense of at least some pollution. In this sense, one of the most salutary lessons of recent decades has been that the so-called benign generators of electrical power—the renewable sources based on winds, tides, solar radiation, or waves—involve potentially large costs in terms of how they modify landscapes. This has limited their attractiveness in relation to nonrenewable sources, and has demonstrated the need to broaden the analyses that feed into decision-making about the environment so that hidden and unexpected costs are given their full weight. Thus, as the field has evolved, it has become correspondingly more sophisticated in its treatment of the variables that are considered when formulating policy.

New challenges. All of the main environmental problems of the late twentieth and early twenty-first centuries fall under the environmental management field. Most problems are controversial. Tropical deforestation, ozone depletion, and global warming have fueled debate over strategies for the management of the global environment. Transboundary pollution and the international exploitation of resources (for example, the appropriation of raw materials in one country and the patenting of their genetic derivatives in another) have underlined the need for bilateral, and often multilateral, agreements about sharing responsibilities. Radiation

emissions, toxic waste issues, and hazardous material spills have emphasized the need for secure and standardized methods of treating pollutants. The production of organic chemicals, for example, grew by more than two orders of magnitude during the second half of the twentieth century, and there was a corresponding growth in the number of catastrophic pollution episodes.

Environmental management has risen to meet many of these challenges. The field has expanded from a purely governmental preserve to one that encompasses the private sector as well. Indeed, the manufacture of pollution control equipment and the institutional management of environmental hazards have turned into growth areas. Yet the successes must be seen against a backdrop of deepening environmental crisis. Relentless population pressure, the unfettered nature of international capital, and the exposure of a record of significant environmental mismanagement in eastern Europe are examples of remaining problems. David Alexander

Bibliography. M. Atchia and S. Tropp (eds.), *Environmental Management: Issues and Solutions*, 1995; R. S. Dorney, *The Professional Practice of Environmental Management*, 1987; T. O'Riordan (ed.), *Environmental Science for Environmental Management*, 1995; R. Welford (ed.), *Corporate Environmental Management: Systems and Strategies*, 1996; R. R. White, *Urban Environmental Management*, 1994; G. A. Wilson and R. L. Bryant, *Environmental Management: New Directions for the 21st Century*, 1997.

Environmental radioactivity

Radioactivity that originates from natural and anthropogenic sources, including radioactive materials in food, housing, and air, radioactive materials used in medicine, nuclear weapon tests in the open atmosphere, and radioactive materials used in industry and power generation.

Natural radioactivity, which is by far the largest component to which humans are exposed, is of both terrestrial and extraterrestrial (cosmic) origin. About 340 nuclides are known in nature, of which 70 are radioactive and are found mainly among the heavy elements. Three nuclides which are responsible for most of the terrestrial component are potassium-40, uranium-238, and thorium-232.

Terrestrial nuclides. Potassium-40 has a half-life of 1.3×10^9 years and decays, through calcium-40, to stable argon-40. This nuclide is present to the extent of about 0.01% in natural potassium, thereby imparting a specific activity of approximately 800 picocuries per gram (3.0×10^4 becquerels per kilogram) of potassium. Potassium is of course an essential element of the human body, which contains about 0.1 microcurie (3.7×10^3 Bq) of the radioactive form of the element. *See* POTASSIUM.

Uranium-238, the most abundant of the uranium isotopes, decays through a chain of 18 nuclides to stable lead-206. This chain includes radionuclides of 10 heavy elements, the half-lives of which range from 4.5×10^9 years (uranium-238) to 1.64×10^{-4} s (polonium-214). *See* URANIUM.

An important characteristic of the uranium-238 chain is that it includes a radioactive noble gas, radon-222, with a half-life of 3.8 days. Radon emanates from the Earth's crust, diffuses into the atmosphere, and eventually decays to radionuclides of lead and polonium that attach themselves electrically to the dust normally present in the atmosphere. The dose of the lung from radon daughter products attached to atmospheric dust is much higher than the dose received from natural sources by any other part of the body. *See* RADON.

The radon daughters are also a source of natural "fallout" which, when released into the soil, results in broad-leaved plants having relatively high concentrations of both lead-210 and polonium-210. In the case of tobacco, the polonium is volatized in the course of smoking and delivers a higher-than-normal dose to the lung. It has been suggested that this phenomenon may play a role in the production of lung cancer from tobacco smoking, but this suggestion must be regarded as highly speculative.

Thorium-232 has a half-life of 1.4×10^9 years and decays through a series of 12 nuclides to a stable isotope of lead (lead-208). Human exposure from this series results mainly from the energetic gamma radiation from thallium-208. Internal irradiation results from absorption of radium-228 (6.7-year half-life) which finds its way into food. *See* THORIUM.

Dose rates. The average person in the United States receives 80–180 mrem/year (0.8–1.8 millisieverts/year) from natural sources of ionizing radiation, depending on the organ considered (see **table**). Most of this dose originates from radioactive materials in the Earth's crust. The external dose due to cosmic rays is an average of about 28 mrem/year (0.28 mSv/year), a value that increases with altitude due to reduced shielding of cosmic radiation by the atmosphere. At Denver, which is located 1 mi (1.6 km) above sea level, the annual dose from cosmic radiation is more than 50 mrem/year (0.5 mSv/year), or about twice the average dose. *See* COSMIC RAYS.

The human body is also exposed to radionuclides in food and water. Potassium-40 is the most important of these, with radium-226 and radium-228 of perhaps less importance from the point of view of the dose delivered.

A number of radionuclides are produced in nature by the interaction of cosmic rays with atmospheric gases, and these add slightly to the exposure received from natural sources. Carbon-14 is the most important of these nuclides which, like potassium-40, is present in the body to the extent of about 0.1 microcurie (3.7×10^3 Bq), but which delivers a dose of only about 1 mrem/year (10 μSv/year) because of the unusual softness, that is, low energy, of its beta emission. *See* COSMOGENIC NUCLIDE.

There are wide deviations from the average doses shown in the table. Thus, at one extreme, miners working underground in the presence of radioactive ore can be exposed to such high levels of atmospheric radon that they develop lung cancer. This

Average dose equivalent rates from various sources of natural background radiation in the United States

Source	Dose, mrem/year*				
	Gonads	Lung	Bone surfaces	Bone marrow	Gastrointestinal tract
Cosmic radiation	28	28	28	28	28
Cosmogenic radionuclides	0.7	0.7	0.8	0.7	0.7
External terrestrial	26	26	26	26	26
Inhaled radionuclides	—	100	—	—	—
Radionuclides in the body	27	24	60	24	24
Rounded totals	80	180	120	80	80

*1 mrem/year = 10^{-2} mSv/year.

was observed in mines in central Europe that were worked for precious metals long before radioactivity was discovered. Uranium miners in the southwestern United States have also been subject to a high incidence of lung cancer because of lack of control of radon.

Radioactive anomalies. There are also geographical areas where the levels of natural radioactivity are unusually high. Six types of anomalies that can be important from the point of view of population exposure are: monazite sands and other placers, alkaline intrusives and granites of the Conway type in New Hampshire, bauxites and intensely weathered soils, uraniferous phosphate rock (and soils), ground waters enriched in radium and radon, and black shales and related organic accumulations.

There are two locations, in India and in Brazil, where the external radiation levels are greatly elevated by the presence of thorium-bearing monazite sands on or near ocean beaches. The local inhabitants are thus exposed to doses several times those received under normal circumstances. In Kerala, India, more than 25% of the 70,000 people living in the monazite area receive a dose greater than 500 mrem/year (5 mSv/year), and a small percentage receive between 1000 and 2000 mrem/year (10 and 20 mSv/year). In Guarapari, Brazil, the 6000 residents receive a mean dose of about 600 mrem/year (6 mSv/year).

Another type of anomaly, of which the outstanding example is the Morro do Ferro near Pocos de Caldas, Brazil, is found in regions of alkaline intrusives. The Morro do Ferro is a hill located in the state of Minas Gerais. The plants absorb so much radioactivity that they autoradiograph themselves when placed on radiation-sensitive film (see **illus.**). It has been estimated that the lungs of rodents living underground on the Morro do Ferro receive as much as 3000 mrem/year (30 mSv/year) from inhalation of a short-lived (53 s) radon isotope found in the thorium chain. The source of the radioactivity on the Morro do Ferro is a deposit of about 13,000 tons (12,000 metric tons) of thorium located near the summit of the hill. The deposit has been in place for perhaps as long as 80,000,000 years, and is in an advanced state of weathering. Rainfall averages about 67 in./year (170 cm/year).

Because of marked chemical similarities between thorium and plutonium, a study of the history of the deposit and the manner in which it has been mobilized by hydraulic and other forces will provide useful information for predicting the behavior over geologic time of plutonium and other actinides in a radioactive waste repository that has been breached.

It is known that water from deep wells may have high concentrations of radium. About 1,000,000

Autoradiograph of plant from the Morro do Ferro, Brazil, made by the plant's natural radioactivity.

people in northern Illinois and southern Iowa drink water with abnormally high radium concentrations.

Technological modification. The natural radioactive environment can also be altered by human activities, as described below.

Building construction. The radon concentration with in buildings is influenced by the materials of construction. Construction materials that often have high concentrations of radium (the parent of radon) include pumice, concrete containing alum shale, and granite. Tailings from uranium and phosphate mining have been particularly troublesome in this regard, and it has been necessary to place restrictions on the use of these materials in construction of buildings. In Grand Junction, Colorado, the use of uranium mine tailings in home construction resulted in such high indoor radon concentrations that remedial action was necessary.

It is ironic that the need to conserve energy is resulting in higher exposures to radon. Weather stripping and other measures that reduce the rate of building ventilation cause less dilution of the radon that diffuses from the materials of construction, and hence tend to increase the radon concentration of the air within such buildings.

Combustion of fossil fuels. All fossil fuels contain naturally occurring radionuclides. Natural gas contains radon, which is sometimes present in detectable concentrations at the place of combustion.

Both oil and coal contain uranium and thorium and their daughter products. The radionuclides are discharged into the atmosphere in measurable amounts that are probably not significant to public health. It has been found that the radioactivity disseminated by burning coal is greater than that discharged into the atmosphere during normal operation of nuclear power reactors.

Aircraft travel. Flying at high altitudes increases the exposure from cosmic rays, and it is estimated that the dose received during a transcontinental trip across the United States is about 2 mrem (20 μSv).

Medical procedures. Procedures in which x-rays or radionuclides are used for diagnostic or therapeutic purposes are also a major source of radiation exposure, second in importance only to natural radioactivity. Doses of iodine-131 and iodine-125 are administered the most frequently. Other important nuclides include phosphorus-32, cobalt-60, gold-198, iron-59, and technetium-99. These nuclides are usually administered internally. *See* ISOTOPIC IRRADIATION; RADIOACTIVE TRACER.

The use of diagnostic x-rays subjects the general population to far more exposure than radiopharmaceuticals. Estimates of the per capita dose from diagnostic x-rays range from 75 mrem/year (0.75 mSv/year) in New Orleans and 50 mrem/year (0.5 mSv/year) in New York to as little as 1.25 mrem/year (12.5 μSv/year) in Thailand. The per capita dose tends to increase with the general level of economic development of the country and the availability of medical services. However, in all countries, developed and developing, the per capita dose can be minimized by improving the techniques used. *See* RADIOGRAPHY; RADIOLOGY.

Nuclear weapons testing. Environmental radioactivity from the testing of nuclear weapons diminished after the major nuclear powers declared a moratorium on atmospheric testing in 1963. Relatively small-scale tests have been conducted by France, China, and India, but these have not added significantly to the levels of radioactivity previously disseminated from tests by the United States, Soviet Union, and United Kingdom prior to the moratorium.

The residual radioactivity from weapons tests is largely from cesium-137 and strontium-90. Cesium-137 is chemically similar to potassium and tends to be distributed more or less uniformly throughout the soft tissues. The dose commitment (the mean lifetime dose from cesium-137 already deposited) was about 27 mrem (0.27 mSv) in the United States in 1976. Cesium-137 has a half-life of 30 years.

Strontium-90, which has a half-life of 38 years, is chemically similar to calcium and therefore tends to deposit in the skeleton. The dose from strontium-90 is highest in individuals whose skeletons were formed during the period of heavy weapons testing, from about 1954 to 1962. Among this segment of the population, the bone dose was annually about 2 mrem (20 μSv) and the 50-year dose in the temperate zone was estimated to have been about 87 mrem (0.87 mSv) in 1976. *See* RADIOACTIVE FALLOUT.

Nuclear power plants. Although various national and international regulatory organizations have proposed guidelines that limit the per capita dose received by individuals in the general population to 170 mrem/year (1.7 mSv/year), it has become evident that nuclear power plants can be routinely operated so that the general population will not be exposed to more than 1% of this limit. The U.S. Nuclear Regulatory Commission requires that the reactors be operated so that the dose to the maximum exposed individual, usually a hypothetical person located at the fenceline, is not more than 10 mrem/year (0.1 mSv/year).

The radioactive wastes discharged into the environment by a normally operated nuclear power reactor are minimal. As noted above, it has been shown that, at least so far as airborne radioactivity is concerned, coal- and oil-fired power plants discharge higher quantities of radioactivity into the environment.

Two of the oldest privately operated power reactors in the United States are Dresden I, a boiling water reactor operated by Commonwealth Edison Company near Chicago, and Yankee Rowe, a pressurized water reactor operated by the Yankee Atomic Electric Company in Rowe, Massachusetts. Surveys made by the U.S. Public Health Service after the reactors were in operation for about 10 years showed that there was no detectable residual radioactivity in the environs. Other reactors that have been studied have been found to expose people in the vicinity to no more than a few mrem per year, and these small levels should be reduced even further as the technology of reactor operation is improved.

A serious accident at the Three Mile Island Unit Two nuclear reactor early in 1979 resulted in worldwide concern about the danger from exposure of nearby Pennsylvania populations. Fortunately, despite extensive damage to the reactor core brought about by a series of mechanical and human errors, the radioactivity was well contained and exposures to the general population were minimal. The average dose, due to the accident, to the population in a 50-mi (80-km) radius was less than 2 mrem (20 μSv).

The principal source of this low-level radiation exposure in the past was gaseous plumes from boiling water reactors. It was at one time permissible to discharge the waste gases after only a 30-min holdup, and this permitted the discharge of short-lived noble gases in copious amounts, thereby accounting for the dose from the plume. However, in order to meet the more stringent standards set by the Nuclear Regulatory Commission in 1975, it is now necessary to provide longer holdup times to permit essentially complete decay of the shorter-lived noble gases.

The most significant nuclides in the liquid wastes are cesium-137 and cesium-134. The former is produced in both weapons tests and nuclear reactors. However, cesium-134 is produced only in nuclear reactors and provides a label by which it is possible to differentiate environmental cesium due to weapons testing from that due to reactors.

The standard methods of treating the low-level radioactive wastes are sufficient to limit the maximum dose to humans who consume fish or shellfish to a few mrem per year. The Environmental Protection Agency has forecast that by the year 2000 the per capita dose from the operation of the several hundred reactors (and associated fuel-reprocessing services) will be 0.4 mrem/year (4 μSv/year), which is less than 1% of the average dose people receive from nature and well within the range of variability in the natural radiation background. *See* NUCLEAR POWER; NUCLEAR REACTOR; RADIOACTIVE WASTE MANAGEMENT; RADIOACTIVITY.

Merril Eisenbud

Bibliography. M. Eisenbud et al., *Environmental Radioactivity*, 4th ed., 1997; R. L. Kathren, *Radioactivity in the Environment*, 1984; *Report of the President's Commission on the Accident at Three Mile Island*, 1979; V. Schultz and F. W. Whicker (eds.), *Radioecology: Nuclear Energy and the Environment*, 2 vols., 1982; L. Tommasino et al., *Radon Monitoring in Radioprotection, Environmental Radioactivity, and Earth Science*, 1990.

Environmental test

The evaluation of a physical system (engineering product) in conditions which simulate one or more of the environments that may harm the system or adversely affect its performance. Such testing is typically undertaken during the development of new products, with the aim of establishing or improving their reliability and durability.

Role of market forces. Almost every engineering product, from familiar household devices such as refrigerators to military hardware such as high-performance aircraft, undergoes some type of environmental testing before being offered to the public or government for sale. The increasing competitiveness of the world marketplace is a major impetus behind the drive for improved product reliability and durability, and has led to the expanded use of environmental tests by manufacturers worldwide.

In addition to the evaluation of a final, finished product, environmental testing can play an important role throughout a product's design/development cycle to ensure that the materials and manufacturing processes employed can meet the stresses imposed by the environment in which the product is likely to operate. By not waiting until a finished product is evaluated, manufacturers can use environmental testing to eliminate costly redesigns late in the design/development cycle.

Facilities. Because it is necessary to precisely control the environmental factors which define the test (for example, temperature, vibration level, or altitude), environmental testing is typically conducted in specially designed facilities, or environmental chambers. Some environmental chambers can generate extremely high and low temperatures and humidity levels. Others can simulate corrosive environments such as salt sprays. The expense of developing and maintaining facilities for conducting the wide variety of environmental tests that new products must undergo is often beyond the resources of the product manufacturers themselves. This has led to the formation of companies specializing in environmental tests. The increasing use of environmental test facilities is mirrored in the expenditures for the test equipment itself.

Standards and specifications. Products and equipment intended for military use are often subjected to the harshest and most variable of environmental conditions. The military has pioneered the creation of well-documented standards and specifications for the evaluation and testing of any products or equipment which it will purchase. In the United States, these standards and specifications are often cited in Department of Defense solicitations as requirements. Examples of standards documents (typically identified as *MIL STD-x* with *x* being a three- or four-digit number) published by the Department of Defense are *MIL STD 167-1 Mechanical Vibration of Shipboard Equipment; MIL STD-202 Test Methods for Electronic and Electrical Component Parts; MIL STD-750 Test Methods for Semiconductor Devices*; and *MIL STD-810 Environmental Test Methods.* Although a number of military standards and specification documents are being reviewed and canceled for the sake of efficiency, these documents still constitute an important source of information for environmental testing.

Civilian organizations, such as the Society of Automotive Engineers, publish standards for automotive and aerospace equipment. In addition, nearly 100 countries have adopted the International Organization for Standardization (ISO) 9000 series of standards for quality management and quality assurance. These standards are being implemented by thousands of manufacturing and service organizations,

both public and private. The ISO 9000 family of standards represents an international consensus of good management practice and is intended to give organizations guidelines as to what constitutes an effective quality management system. Of the ISO 9000 family standards, ISO 9003 covers quality assurance obligations of the manufacturer in the areas of final inspection and testing. Among the guidelines provided by ISO 9003 are those dedicated to (1) developing procedures to inspect, test, and verify that final products meet all specified requirements before they are sold; (2) developing procedures to control and calibrate the testing equipment; (3) ensuring that every product is identified as having passed or failed the required tests. ISO 9003 will have a significant impact upon the standardization of environmental testing in both civilian and military endeavors.

Types of tests. An incomplete but representative list of environmental tests requiring dedicated test chambers includes tests for altitude, dust, explosiveness, flammability, fungus, humidity, icing, acoustic vibration, overpressure, rain, salt, fog, sand, temperature, and wind. Tests typically not requiring chambers but still utilizing dedicated mechanical testing equipment include acceleration, fatigue cycling, transportation simulation, shock, and vibration. The size of environmental test chambers can range from less than a cubic meter for testing component parts (for example, exposing a sample of an automobile tire to controlled levels of ozone) to 1,000,000 m^3 (for example, testing entire aircraft in "cold-soak" conditions).

A representative example of environmental testing is provided by an ISO standard, also adopted by the Society for Automotive Engineers, that specifies environmental tests for aircraft equipment whose operating environment includes exposure to acoustic vibration (or noise). The standard specifies laboratory conditions for testing the ability of equipment in subsonic and supersonic civilian aircraft to withstand the stresses of acoustic vibration. Two types of tests are required to be performed: equipment functioning and acoustic endurance.

Equipment functioning refers to the ability of the equipment in question, for example, an avionics subsystem, to perform as desired in the presence of acoustic vibration representative of operational use. Tests for functioning are typically conducted both during equipment development and for certification of the equipment, once produced. In aircraft applications, acoustic vibration typically finds its source in external aerodynamics (that created by the aircraft passing through the atmosphere) or internal aerodynamics and combustion (that created by the jet engines). Noise or acoustic vibration in these environmental tests is quantified in decibels (dB). For example, sound pressure level (SPL) is defined as $\mathrm{SPL} = 10 \cdot \log_{10}(p^2_{\mathrm{rms}}/p^2_{\mathrm{ref}})$ decibels, where p_{rms} refers to the root-mean-square (rms) sound pressure (in pascals) in the acoustic source (for example, a jet engine) and p_{ref} is a reference sound pressure level, taken to be 0.0002 pascal (**Fig. 1**). Typically, sound pressure levels in excess of 125 dB are employed in the acoustic environmental tests for aircraft equipment. In addition to decibel rating, the frequency of acoustic vibration is used as an experimental variable.

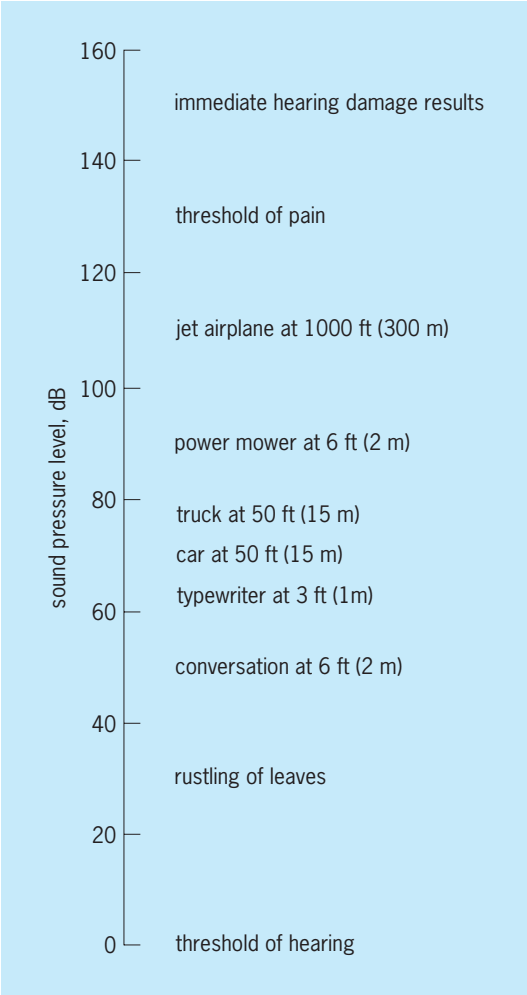

Fig. 1. Sound pressure levels of familiar sources.

Acoustic endurance refers to the length of time that equipment can perform to specification in the presence of noise of specified power. In environmental tests of acoustic endurance, so called wide-band noise is used, that is, noise with power distributed evenly across a broad range of frequencies. The level of power (test severity grade) is determined by the user according to either of two procedures: (1) If measurements of in-flight sound pressure levels at the location where the equipment is to be installed are available or can be calculated, they can be used to determine test sound pressure levels. (2) If such measurements are unavailable, the test levels may be obtained based on the type of aircraft in which the equipment is to be installed and the location of the equipment relative to the noise sources. *See* ACOUSTIC NOISE; AIRCRAFT TESTING; MECHANICAL VIBRATION; SOUND; SOUND PRESSURE; VIBRATION.

An example of an environmental test chamber is the large "sunspot chamber" for Earth orbital and deep-space simulations of spacecraft, operated by the National Aeronautics and Space Administration (**Fig. 2**). The chamber is equipped with a shroud that is liquid-nitrogen-cooled. Heat lamps are used to

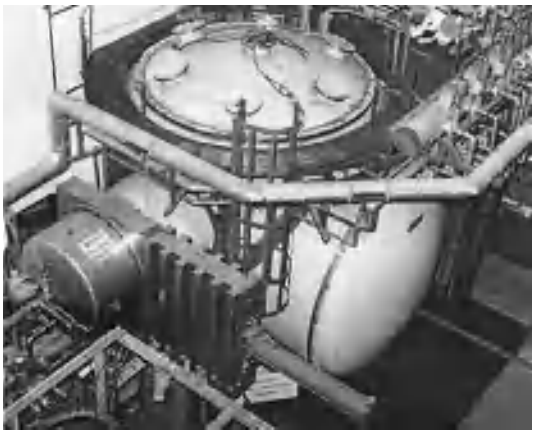

Fig. 2. Large "sunspot chamber" for Earth orbital and deep-space simulations of spacecraft, located at the Marshall Space Flight Center. (*NASA*)

simulate the radiance of the Sun. The chamber itself consists of a vertical cylinder 3.2 m (10.5 ft) in diameter and 3.7 m (12 ft) in height. To simulate the vacuum in deep space, the chamber pressure can be lowered to 1.3×10^{-4} pascal.

Environmental engineering specialists. The importance and ubiquity of environmental testing has led to the formation of an engineering subspecialty. An environmental engineering specialist is defined to be a person who is skilled in one or more environmental engineering areas, which include but are not limited to natural and induced environments and their effects on material; expertise in measuring and analyzing field environmental conditions; formulating environmental test criteria; determining when environmental laboratory tests are appropriate and valid substitutes for natural field/fleet environmental tests; and evaluating the effects of specific environments on material.

Challenger accident. The accident involving the NASA space shuttle *Challenger* illustrates the vital importance of environmental testing and the results of either inadequate testing or use of systems in environments which have not been adequately represented in tests. On January 28, 1986, the *Challenger* was launched from Cape Canaveral for a 7-day mission. Approximately 73 seconds into the flight, the vehicle exploded and the crew perished. After a lengthy investigation, the cause of the disaster was traced to the failure of an O-ring seal in one of the solid rocket boosters, which allowed hot exhaust gases to leak. The O-ring failure was attributed to the freezing temperatures of the launch environment, which severely degraded the flexibility of the O-ring, a flexibility that was required in the vibration and aerodynamic load environment of the ascent. *See* SPACE FLIGHT; SPACE SHUTTLE; SYSTEMS ENGINEERING.

Ronald A. Hess

Environmental toxicology

A broad field of study encompassing the production, fate, and effects of natural and synthetic pollutants in the environment. The breadth of this field depends on the definition of environment. It can be defined as narrowly as the home and workplace or as broadly as the entire Earth and its biosphere. Environmental toxicology is truly an interdisciplinary science. The effects of a pollutant on the environment depend on the amount released (the dose) and its chemical and physical properties. Pollutants can be grouped according to their origin and effects.

Nutrients. Pollution from nutrients is generally a problem of aquatic systems. Carbon, nitrogen, and phosphorus are essential nutrients and, when present in excess, can result in an overstimulation of microbial and plant growth. Nutrients enter the environment in runoff from fertilized agricultural areas, in effluents from municipal and industrial wastewater treatment facilities, and in dead and decaying plant material. Excess microbial growth can result in deoxygenation of a body of water. The overstimulation of algae and plant growth in aquatic systems due to excess nutrients of human origin is called cultural eutrophication. Decreased clarity of water, overgrowth of emergent vegetation, increased rate of sedimentation and filling-in of lake basins, and greatly shortened lifespan of an aquatic system are all characteristics of cultural eutrophication. *See* EUTROPHICATION.

Microbial pollution. Pathogenic bacteria and protozoa can be a major source of pollution in areas that receive untreated sewage, items from ocean dumping, and improperly discarded hospital waste. Toxic metabolites of fungal origin (mycotoxins) are also potential pollutants. Most important of the broad range of mycotoxins are those that contaminate human food and the feed of domestic animals. These contaminants include the ergot alkaloids, tricothecins, and aflatoxins. Ergot alkaloids are produced by *Claviceps* spp. and are strong neurotoxic and vasoconstrictive agents. Tricothecins are produced by a variety of species of the genus *Fusarium* and are potent inflammatory agents and tissue irritants. Aflatoxins are produced by species of the genus *Aspergillus* and contaminate improperly stored grains and nuts. Aflatoxins have been implicated in a variety of diseases in poultry and humans, have caused cancer in experimental animals, and have been linked epidemiologically to cancer in humans. Aflatoxins are the only mycotoxins to be regulated by the U.S. Food and Drug Administration. *See* AFLATOXIN; ERGOT AND ERGOTISM; TOXIN.

Suspended solids. Forest fires, volcanic eruptions, and dust storms can be major sources of suspended materials. These materials can also originate in runoff from agricultural areas, construction and mining sites, and roads and other paved areas. Truck and automobile exhaust and industrial discharge to the atmosphere are also sources of suspended solids. These materials reduce the amount of light in the atmosphere and bodies of water, interfering with photosynthesis. Physical irritation and abrasion, and obliteration of habitat by silting are the primary toxic effects of suspended solids.

Atmospheric pollutants. Metabolic processes and natural combustion and thermal activity (such as forest fires and volcanoes) can release large amounts of

gaseous by-products to the atmosphere. However, natural inputs are minor compared to atmospheric pollutants due to human activity. Although most anthropogenic air pollution is produced by the various forms of transportation, emission from stationary sources of fuel combustion (for example, factories and power plants) are responsible for the greatest amount of hazardous materials released.

Sulfur oxides, a primary combustion product of coal, are converted to acids in the atmosphere and precipitated in rain. In areas of limited buffering capacity such as northeastern North American and Scandinavia, pH levels are known to be greatly reduced, with loss of forests and aquatic life. *See* ACID RAIN.

Nitrogen oxides are formed during high-temperature combustion of petroleum products in refining operations and internal combustion engines. As with sulfur oxides, nitrogen oxides can form acid precipitation. They can also react with hydrocarbons and sunlight to form ozone and peroxyacetyl nitrates, the major constituents of photochemical smog.

Chlorofluorocarbons are important refrigerants and aerosol propellants. These gases accumulate in the upper atmosphere and have been implicated in a decrease in the protective ozone layer surrounding the Earth. Because the ozone layer absorbs harmful ultraviolet radiation from the Sun, it has been predicted that a decrease in this layer will result in an increase in skin cancer. In addition, the biological structures of aquatic and terrestrial ecosystems are strongly influenced by solar ultraviolet radiation, and most organisms exist very near their threshold of tolerance.

Carbon monoxide, formed by the incomplete combustion of fossil fuels, is primarily derived from transportation sources. It interferes with the oxygen-carrying function of hemoglobin in the bloodstream. The effects of carbon monoxide are acute but reversible, and thus pose a relatively small health hazard compared to other forms of atmospheric pollution. However, repeated anoxic episodes due to carbon monoxide can permanently damage the blood–brain barrier and the white matter of the brain.

Carbon dioxide is a by-product of respiration and fossil fuel combustion. Increasing amounts of carbon dioxide in the atmosphere have caused concern over a global warming trend, the so-called greenhouse effect. *See* AIR POLLUTION; GREENHOUSE EFFECT.

Metals. All living organisms require certain metals for physiological processes. These elements, when present at concentrations above the level of homeostatic regulation, can be toxic. In addition, there are metals that are chemically similar to, but higher in molecular weight than, the essential metals (heavy metals). Metals can exert toxic effects by direct irritant activity, blocking functional groups in enzymes, altering the conformation of biomolecules, or displacement of essential metals in metalloenzymes.

Solvents. Organic solvents are used widely and in large amounts in industries, laboratories, and homes (see **table**). They are released to the atmosphere as vapor and can pose a significant inhalation hazard. Improper storage, use, and disposal have resulted in the contamination of surface and ground waters and drinking water. *See* WATER POLLUTION.

Pesticides. The pesticides represent an important group of materials that can enter the environment as pollutants. They are highly toxic, and many nontarget organisms can suffer harmful effects if misuse or unintended release occurs. *See* PESTICIDE.

Insecticides. The most diverse group of pesticides is the insecticides and acaricides. Insecticides have been used for hundreds of years, but it has only been since World War II that synthetic insecticides have been widely used.

The chlorinated hydrocarbons include dichlorodiphenyl trichloroethane (DDT) and its analogs, the cyclodienes and related compounds, and the benzene hexachlorides (BHCs). Beginning in the 1940s, chlorinated hydrocarbons were used extensively in mosquito control and agriculture. These compounds exert their toxic effect by interfering with the transmission of nerve impulses, inhibiting the transport of ions across the axonal membrane. Many are very persistent in the environment and undergo biomagnification in food chains. Toxic effects on top predators such as birds and the contamination of human food supplies raised concern over their use. After the 1960s, most were replaced with other, less persistent compounds. *See* FOOD WEB.

Organophosphates inhibit the enzyme acetylcholinesterase, resulting in an overstimulation and excitation of nerves. Because of their mode of action, organophosphates are toxic not only to target insects but also to nontarget insects (bees and aquatic insects), birds, wildlife, and fish and other aquatic life. With few exceptions (malathion), organophosphates are toxic to mammals and humans. They are easily degraded and do not persist in the environment. However, because of their high nontarget organism toxicity, their use is limited to areas where undesired exposure can be minimized. *See* ACETYLCHOLINE.

Carbamate insecticides exhibit neurotoxic action similar to the organophosphates, by inhibiting acetylcholinesterase. However, this action is more readily reversible than that of the organophosphates.

Botanical insecticides are natural and synthetic derivatives of toxic plant materials and have been used for hundreds of years. Nicotine and its analogs are neuroactive agents. Rotenones are electron-transport inhibitors and are used as insecticides and pesticides. Pyrethroid insecticides were originally derived from chrysanthemum plants; however, most pyrethroids in use today are synthetic derivatives of natural plant toxins. Pyrethroids are nerve poisons, acting through interference of ion transport along the axonal membrane. The pyrethroids are the most widely used insecticides, primarily because of their low mammalian toxicity. Although not very toxic to mammals, pyrethroids are extremely toxic to nontarget arthropods, such as bees, aquatic insects, and crustaceans, and to nonmammalian vertebrates such as fish.

Uses and effects of selected organic solvents

Solvent	Uses	Effects
Aliphatic hydrocarbons		
Pentane, hexanes, heptanes, octanes	Commercial products and solvents	Depression of central nervous system and liver pathology
Halogenated aliphatic hydrocarbons		
Methylene chloride	Paint remover, aerosol solvent	Depression of central nervous system, respiratory poison
Chloroform	Chemical intermediate, solvent	Liver pathology, carcinogen
Carbon tetrachloride	Industrial and laboratory solvent	Lipid peroxidation, liver and renal pathology
Aromatic hydrocarbons		
Benzene	Organic synthesis, solvent, printing	Hematopoietic toxicity, immunosuppression, leukemia
Toluene	Solvent, chemical intermediate, paints, rubber	Narcotic, central nervous system pathology
Xylene	Solvent, chemical intermediate, pesticides, adhesives	Central nervous system pathology
Alcohols		
Methanol	Commercial and laboratory solvent, paints, fuel additive	Toxic metabolites (formaldehyde), optic nerve damage
Isopropyl alcohol	Cosmetics, glass cleaning solutions	Central nervous system depressant
Glycols		
Ethylene glycol	Heat exchangers, antifreeze, hydraulic fluids	Toxic metabolites

Fumigants are volatile substances used as soil pesticides and to control insects in stored products and scale insects on citrus. Common fumigants include ethylene dibromide, ethylene dichloride, ethylene oxide, carbon disulfide, methyl bromide, hydrogen cyanide, phosphine, and chloropicrin. Because they are nonselective and toxic to humans, many fumigants are now restricted or banned.

Organic thiocyanate insecticides are mild general poisons, and have been used as fly sprays and fumigants that are also toxic to nontarget animals and plants. Dinitrophenols were important as early insecticides and acaricides and still have limited use as dormant acaricides. They are toxic to mammals and plants. Fluoroacetate derivatives are general toxicants that form lethal metabolic products.

Inorganic insecticides are general toxicants and have been largely replaced with synthetic organic insecticides. However, two classes of inorganic insecticides are still being used, arsenicals and fluorides. Common arsenicals include lead and calcium arsenates and sodium arsenite. Sodium fluoride, sodium fluoraluminate, and sodium fluorosilicate are common fluorides. Both groups are stomach poisons, increasing in toxicity with increasing metal content. *See* INSECTICIDE.

Herbicides. These chemicals are selectively toxic to plants. Examples include the chlorophenoxy growth stimulators 2,4-dichlorophenoxyacetic acid (2,4-D) and 2,4,5-trichlorophenoxyacetic acid (2,4,5-T), the protein synthesis–inhibiting alachlor, the defoliant paraquat, and the chlorophenolic contact herbicides. Although herbicides have in large part not been an environmental problem because of their selectivity and low persistence, some can be very toxic to nontarget organisms. Paraquat, for example, can cause severe pulmonary symptoms. *See* HERBICIDE.

Fungicides. Fungicidal compounds are used widely to treat seed grains and wood. Common ones are pentachlorophenol and the organomercuricals. *See* FUNGISTAT AND FUNGICIDE.

Fossil fuels. Coal and petroleum-derived materials and by-products are major environmental pollutants. The world's economy is highly dependent on fossil fuels in energy production, industry, and transportation. Widespread use has led to enormous releases to the environment of distillate fuels, crude oils, runoff from coal piles, exhaust from internal combustion-fired power plants, industrial emissions, and emissions from municipal incinerators. Point-source leaks and spills, and non-point-source emissions have resulted in environmental contamination with millions of tons of petroleum hydrocarbons each year. Spills of crude oil and fuels have caused wide-ranging damage in the marine and fresh-water environments. Oil slicks and tar in shore areas and beaches can ruin the esthetic value of entire regions.

The toxicity of polycyclic aromatic hydrocarbons is perhaps one of the most serious long-term problems associated with the use of petroleum. They comprise a large class of petroleum compounds containing two or more benzene rings. They are present in fossil fuels and are formed in the incomplete combustion of organic materials. Polycyclic aromatic hydrocarbons are formed in nature by long-term, low-temperature chemical reactions in sedimentary deposits of organic materials and in high-temperature events such as volcanoes and forest fires. The major source of this pollution is, however, human activity. Polycyclic aromatic hydrocarbons accumulate in soil, sediment, and biota. At high concentrations, they can be acutely toxic by disrupting membrane function. Many cause sunlight-induced toxicity in humans and fish and other aquatic organisms. In addition, long-term, chronic toxicity has been demonstrated in a wide variety of organisms. Through metabolic activation, some polycyclic aromatic hydrocarbons form reactive intermediates that bind to deoxyribonucleic acid (DNA). For this reason many of these hydrocarbons are mutagenic, teratogenic, or carcinogenic. They are also suspected of interfering in the reproduction of aquatic life. Low

rates of reproduction and high rates of larval deformities and mortality have been observed in fish exposed to polycyclic aromatic hydrocarbons. *See* FOSSIL FUEL; MUTAGENS AND CARCINOGENS.

Other synthetic organic compounds. Polychlorinated biphenyls (PCBs) are produced by the chlorination of biphenyl, giving rise to mixtures of up to 210 possible products. They have been used worldwide in electrical equipment, vacuum pumps, hydraulic fluids, heat-transfer systems, lubricants, and inks. The related polybrominated biphenyls (PBBs) have been used as fire retardants. Major sources of polychlorinated biphenyls have included leaks from waste disposal facilities, vaporization during combustion, and disposal of industrial fluids. They have been identified in environments and organisms worldwide. Being lipid-soluble, they accumulate in the biota, and biomagnification in food chains can be demonstrated. They induce activity in hepatic enzymes and cause long-term liver damage, including cancer. Their use has been largely restricted or eliminated. Environmental concentrations are decreasing, but with their persistence they remain significant pollutants. *See* POLYCHLORINATED BIPHENYLS.

Chlorinated dibenzo-*p*-dioxins and dibenzofurans are formed during the heating of chlorophenols, and have been identified as potential contaminants in the herbicide 2,4,5-T. They can be formed during the incineration of municipal wastes, polychlorinated biphenyls, or plant materials treated with chlorophenols. In certain organisms (guinea pig) they are some of the most toxic known substances, exhibiting hepatotoxic, carcinogenic, and immunosuppressive activity. The dioxins and dibenzofurans have been implicated in several spectacular incidents, including the spraying of Agent Orange during the Vietnam War, the Yusho disease in Japan, the Seveso herbicide plant explosion in Italy, and the discovery of contamination at Love Canal, New York, at Times Beach, Missouri, and in the Great Lakes region of North America. *See* ENVIRONMENTAL ENGINEERING; TOXICOLOGY. James T. Oris

Bibliography. M. O. Amdur, *Casarett and Doull's Toxicology: The Basic Science of Poisons*, 5th ed., 1996; D. W. Connell and G. J. Miller, *Chemistry and Ecotoxicology of Pollution*, 1984; E. Hodgson and P. E. Levi, *A Textbook of Modern Toxicology*, 2d ed., 1997; T. N. Veziroglu (ed.), *The Biosphere: Problems and Solutions*, 1984.

Enzyme

A catalytic protein produced by living cells. The chemical reactions involved in the digestion of foods, the biosynthesis of macromolecules, the controlled release and utilization of chemical energy, and other processes characteristic of life are all catalyzed by enzymes. In the absence of enzymes, these reactions would not take place at a significant rate. Several hundred different reactions can proceed simultaneously within a living cell, and the cell contains a comparable number of individual enzymes, each of which controls the rate of one or more of these reactions. The potentiality of a cell for growing, dividing, and performing specialized functions, such as contraction or transmission of nerve impulses, is determined by the complement of enzymes it possesses. Some representative enzymes, their sources, and reaction specificities are shown in the **table**.

Characteristics

Enzymes can be isolated and are active outside of the living cell. They are such efficient catalysts that they accelerate chemical reactions measurably, even at concentrations so low that they cannot be detected by most chemical tests for protein. Like other chemical reactions, enzyme-catalyzed reactions proceed only when accompanied by a decrease in free energy; at equilibrium the concentrations of reactants and products are the same in the presence of an enzyme as in its absence. An enzyme can catalyze an indefinite amount of chemical change without itself being diminished or altered by the reaction. However, because most isolated enzymes are relatively unstable, they often gradually lose activity under the conditions employed for their study.

Chemical nature. All enzymes are proteins. Their molecular weights range from about 10,000 to more than 1,000,000. Like other proteins, enzymes consist of chains of amino acids linked together by peptide bonds. An enzyme molecule may contain one or more of these polypeptide chains. The sequence of amino acids within the polypeptide chains is characteristic for each enzyme and is believed to determine the unique three-dimensional conformation in which the chains are folded. This conformation, which is necessary for the activity of the enzyme, is stabilized by interactions of amino acids in different parts of the peptide chains with each other and with the surrounding medium. These interactions are relatively weak and may be disrupted readily by high temperatures, acid or alkaline conditions, or changes in the polarity of the medium. Such changes lead to an unfolding of the peptide chains (denaturation) and a concomitant loss of enzymatic activity, solubility, and other properties characteristic of the native enzyme. Enzyme denaturation is sometimes reversible. It is possible, in some cases, to obtain conditions under which the denatured enzyme refolds into its native conformation and regains its original catalytic activity. *See* AMINO ACIDS; PROTEIN.

Many enzymes contain an additional, nonprotein component, termed a coenzyme or prosthetic group. This may be an organic molecule, often a vitamin derivative, or a metal ion. The coenzyme, in most instances, participates directly in the catalytic reaction. For example, it may serve as an intermediate carrier of a group being transferred from one substrate to another. Some enzymes have coenzymes that are tightly bound to the protein and difficult to remove, while others have coenzymes that dissociate readily. When the protein moiety (the apoenzyme) and the coenzyme are separated from each other, neither possesses the catalytic properties of the original conjugated protein (the holoenzyme). By simply mixing the apoenzyme and the coenzyme together, the fully active holoenzyme can often be reconstituted. The

Some representative enzymes, their sources, and reaction specificities

Enzyme	Some sources	Reaction catalyzed
Pepsin	Gastric juice	Hydrolysis of proteins to peptides and amino acids
Urease	Jack bean, bacteria	Hydrolysis of urea to ammonia and carbon dioxide
Amylase	Saliva, pancreatic juice	Hydrolysis of starch to maltose
Phosphorylase	Muscle, liver, plants	Reversible phosphorolysis of starch or glycogen to glucose-1-phosphate
Transaminases	Many animal and plant tissues	Transfer of an amino group from an amino acid to a keto acid
Phosphohexose isomerase	Muscle, yeast	Interconversion of glucose-6-phosphate and fructose-6-phosphate
Pyruvic carboxylase	Yeast, bacteria, plants	Decarboxylation of pyruvate to acetaldehyde and carbon dioxide
Catalase	Erythrocytes, liver	Decomposition of hydrogen peroxide oxygen and water
Alcohol dehydrogenase	Liver	Oxidation of ethanol to acetaldehyde
Xanthine oxidase	Milk, liver	Oxidation of xanthine and hypoxanthine to uric acid

same coenzyme may be associated with many enzymes which catalyze different reactions. It is thus primarily the nature of the apoenzyme rather than that of the coenzyme which determines the specificity of the reaction. *See* BIOLOGICAL OXIDATION; COENZYME.

The complete amino acid sequence of several enzymes has been determined by chemical methods. By x-ray crystallographic methods even the exact three-dimensional molecular structure of a few enzymes has been deduced. *See* X-RAY CRYSTALLOGRAPHY.

Classification and nomenclature. Enzymes are usually classified and named according to the reaction they catalyze. The principal classes are as follows.

Oxidoreductases. These enzymes, which catalyze reactions involving electron transfer, play an important role in cellular respiration and energy production. Some of them participate in the process of oxidative phosphorylation, whereby the energy released by the oxidation of carbohydrates and fats is utilized for the synthesis of adenosine triphosphate (ATP) and thus made directly available for energy-requiring reactions. Most oxidoreductases require prosthetic groups or coenzymes, which participate in the catalytic reaction by serving as intermediate electron carriers. These coenzymes include the flavin nucleotides, the pyridine nucleotides, heme, and various metal ions, such as iron, copper, and molybdenum. Some oxidoreductases, the oxygenases, catalyze the direct incorporation of oxygen into their substrate. Oxidases are enzymes which utilize molecular oxygen as an electron acceptor, while dehydrogenases remove hydrogen atoms from their substrates and transfer them to an acceptor other than oxygen. For example, glucose oxidase catalyzes the reaction shown in reaction (1). This enzyme utilizes

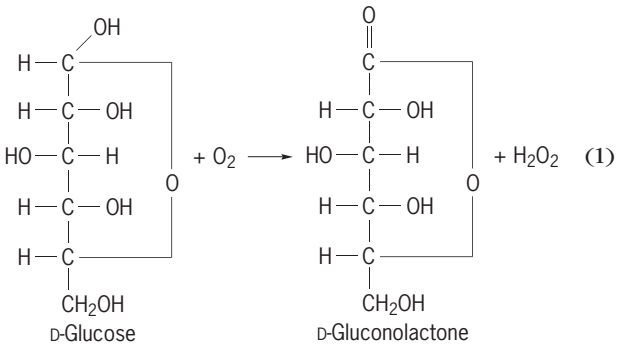

(1)

flavin adenine dinucleotide, a derivative of the vitamin riboflavin, as coenzyme. *See* RIBOFLAVIN.

Transferases. These enzymes catalyze the transfer of a particular chemical group from one substance to another. Thus, transaminases transfer amino groups, transmethylases transfer methyl groups, and so on. An important subclass of this group are the kinases, which catalyze the phosphorylation of their substrates by transferring a phosphate group, usually from ATP, thereby activating an otherwise metabolically inert compound for further transformations. The reaction catalyzed by transaminases may be represented by reaction (2). These enzymes require

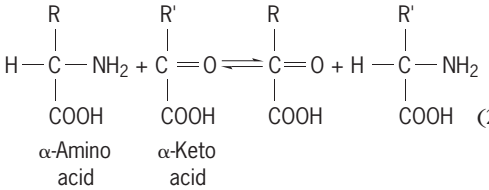

(2)

pyridoxal phosphate or pyridoxamine phosphate, derivatives of vitamin B_6, as coenzymes. *See* VITAMIN B_6.

Hydrolases. These are enzymes which catalyze the hydrolysis of proteins (proteinases and peptidases), of nucleic acids (nucleases), of starch (amylases), of fats (lipases), of phosphate esters (phosphatases), and of other substances. Many hydrolases are secreted by the stomach, pancreas, and intestine and are responsible for the digestion of foods. Others participate in more specialized cellular functions. For example, cholinesterase, which catalyzes the hydrolysis of acetylcholine, as shown in reaction (3), plays

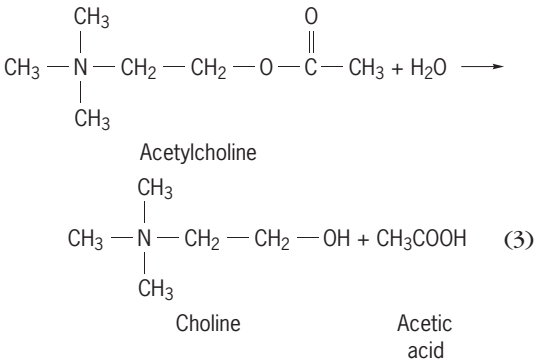

(3)

an important role in the transmission of nervous impulses. *See* ACETYLCHOLINE.

Lyases. These enzymes catalyze the nonhydrolytic cleavage of their substrate with the formation of a double bond. Examples are decarboxylases, which remove carboxyl groups as carbon dioxide, and dehydrases, which remove a molecule of water. The reverse reactions are catalyzed by the same enzymes, but are difficult to demonstrate in some cases because of an unfavorable equilibrium.

Thus pyruvic decarboxylase, which requires thiamine pyrophosphate and magnesium ions for activity, catalyzes the reaction that is shown in reaction (4).

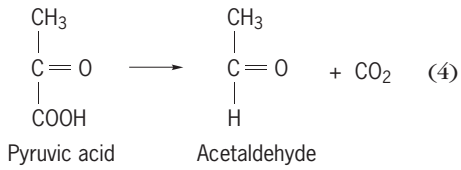

(4)

Isomerases. These are enzymes which catalyze the interconversion of isomeric compounds. For example, triose phosphate isomerase catalyzes the reaction shown in reaction (5).

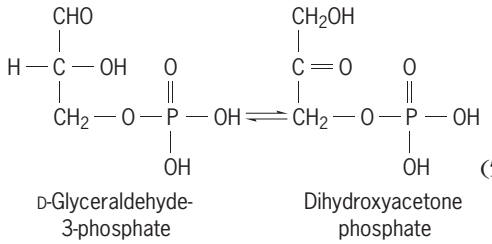

(5)

Ligases or synthetases. These enzymes catalyze endergonic syntheses coupled with the exergonic hydrolysis of ATP. They allow the chemical energy stored in ATP to be utilized for driving reactions uphill. An example is glutamine synthetase, which catalyzes the reaction shown in reaction (6). This enzyme, which

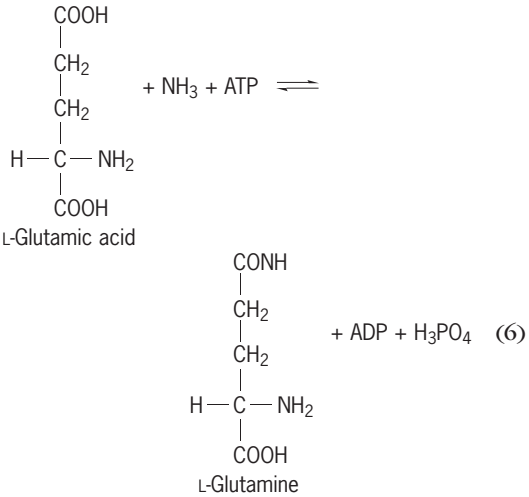

(6)

is abundant in brain tissue, requires magnesium or manganese ions for activity.

Active site. Since enzymes are large molecules and their substrates are often of low molecular weight, it appears probable that only a small portion of the enzyme protein, the active site, comes into contact with the substrate and is directly involved in catalyzing the reaction. The active site consists of a few amino acid residues, not necessarily adjacent to each other on the polypeptide chain but brought into proximity by the folding of the chain or chains. The active site also includes any coenzyme which may be required for activity. The function of the remainder of the protein molecule may be primarily to hold the components of the active site in the proper relative position and orientation. Part of the active site is involved in binding the substrate, while another part is responsible for the making or breaking of chemical bonds. There is much evidence to indicate that, for some enzymes, the binding of the substrates produces a conformational change which brings some reactive group of the enzyme in the proper position for the reaction to take place.

Because of their special environment at the active site of enzymes, coenzymes and certain amino acid residues acquire unique properties and reactivities which they do not possess in the free state. For example, when the proteinase chymotrypsin is treated with diisopropylfluorophosphate, one serine residue at the active site reacts rapidly with this reagent, resulting in complete inactivation of the enzyme. The other 27 serine residues of the protein do not react, and no reaction takes place when the protein molecule is first unfolded, that is, when it is denatured. On the other hand, amino acid residues remote from the active site of an enzyme may often be removed or chemically modified without affecting the catalytic activity.

Intracellular organization. Enzymes are not distributed uniformly throughout the cell but are localized within various subcellular structures or compartments. For example, the enzymes responsible for glycolysis, the degradation of glucose to pyruvate, are located in the soluble, nonparticulate portion of the cell, whereas the enzymes that catalyze the further oxidation of pyruvate to carbon dioxide and water and couple this oxidation to the production of ATP are localized in specialized particles, the mitochondria. Similarly, while the activation of amino acids for protein synthesis takes place in free solution, the assembly of the amino acids into proteins takes place on a particle, the ribosome. Other enzymes are localized on the cell membrane, in the nucleus, and in other subcellular particles. *See* CELL (BIOLOGY); MITOCHONDRIA; RIBOSOMES.

Control of enzyme activity. In order to preserve the stability of its internal chemical environment, a living organism must possess the means of regulating the rate of its metabolic reactions. This is accomplished by regulating both the amount and the activity of the enzymes catalyzing these reactions. Some of the control mechanisms involved are very complex and not yet completely understood, although much progress has been made in this field. A few examples are mentioned below.

Proenzyme activation. One way by which the activity of enzymes is controlled is as follows. The enzyme is produced in an inactive form, termed a zymogen or proenzyme, and converted into the active form only when needed. This is the case with thrombin, an enzyme which causes blood to clot by catalyzing the conversion of the soluble protein, fibrinogen, to

the insoluble fibrin. Blood does not normally contain thrombin but does contain its inactive precursor, prothrombin. It is only when prothrombin is converted to thrombin (by another enzyme) that clotting results. The mechanism is shown in reaction (7). An-

$$\text{Fibrinogen} \xrightarrow[\text{(active)}]{\text{Prothrombin (inactive)} \;\downarrow\; \text{Thrombin}} \text{Fibrin (clot)} \qquad (7)$$

other example of enzyme activation occurs with the proteinase trypsin, which is stored and secreted by the pancreas as the inactive trypsinogen. In the intestine, trypsinogen is converted to trypsin by the enzyme enterokinase. The conversion of trypsinogen to trypsin is also catalyzed by trypsin itself, the activation being therefore autocatalytic. The activation involves the splitting of a peptide from one end of the trypsinogen molecule.

Feedback inhibition. Other mechanisms of enzyme control are shown by reaction sequence (8), which

$$A \xrightarrow{E_1} B \xrightarrow{E_2} C \xrightarrow{E_3} D \qquad (8)$$

depicts a hypothetical metabolic pathway in which the starting material A is converted to the metabolite D required by the cell. One type of control, termed feedback inhibition, results from a specific, reversible, inhibition of enzyme E_1 by the end product D. This phenomenon is obviously of great advantage to the cell, since it allows the concentration of the product D to be accurately regulated: When D reaches a certain concentration, it decreases the rate of its own synthesis by preventing the conversion of A to B; the intermediates, B and C, do not accumulate; and the starting material, A, is spared for other reactions for which it may be needed. As soon as the concentration of D decreases, the inhibition of enzyme E_1 is relieved and the synthesis of D resumes. *See* ENZYME INHIBITION.

This type of control is also termed allosteric control, and the enzyme involved, an allosteric enzyme. The allosteric enzyme contains a specific binding site for the controlling metabolite, D, such that when D is bound to this site the enzyme is inhibited. Positive allosteric control, or allosteric activation, has also been observed. In this case, the binding metabolite, which may be the product of another pathway, causes activation of the enzyme. Some enzymes are susceptible to allosteric control, both positive and negative, by a number of different metabolities. *See* ALLOSTERIC ENZYME.

Induction and repression. Controls also exist at the level of enzyme synthesis. One of these controls is enzyme induction. For example, if the inducible enzymes E_1, E_2, and E_3 are required to utilize substance A, the enzymes are not produced by the cell when A is not available. Only in the presence of an inducer, which may be A or a related compound, are enzymes E_1, E_2, and E_3 synthesized. Another, related, type of control is enzyme repression. In this case the control is exerted by the end product of a metabolic pathway, for example, compound D. Thus, if enzymes E_1, E_2, and E_3 are repressible enzymes, the accumulation of D prevents their synthesis by the cell. Enzyme repression and induction involve complex regulatory mechanisms which are under direct genetic control.

Purification. One of the prerequisites for a study of the structure and mechanism of action of an enzyme is its availability in pure form. The enzyme must therefore first be extracted from the cells or tissue in which it is found and separated from many proteins and other substances present in the crude extract. Because of the instability of most enzymes, however, their isolation requires special techniques, some of which are described below.

Cell destruction. If the enzyme is present in cells, the cells must first be broken. This can be accomplished by grinding with abrasives, by freezing and thawing, or by exposure to ultrasonic vibrations. Although many enzymes are readily released in soluble form by these methods, some remain associated with membranes or subcellular particles and may be difficult to solubilize.

Heating with substrate. Some enzymes can be purified by taking advantage of the fact that they are stabilized in the presence of their substrate and can then be heated for a short time to a temperature at which less stable proteins are denatured and precipitated.

Selective precipitation. Enzymes contain many ionizable groups, such as carboxyl groups, which may be negatively charged, and amino groups, which may be positively charged. Since the degree of ionization of each group depends on the pH of the solution, the total charge on the protein varies greatly with pH. The pH at which the protein has no net charge, its isoelectric point, is usually the pH of minimal solubility. Thus an enzyme can often be selectively precipitated from a solution by adjusting the pH to its isoelectric point.

Electrophoresis. An enzyme can also be separated from proteins with different electric charges by electrophoresis. This method consists of applying an electric field to the solution and allowing each protein to migrate in a direction and at a rate which depend on its charge. *See* ELECTROPHORESIS.

Ammonium sulfate fractionation. Another useful method of separating proteins from each other is by ammonium sulfate fractionation. The solubility of proteins decreases as the concentration of ammonium sulfate is raised, and when the salt concentration is sufficiently high, the protein precipitates. Since different proteins precipitate at different ammonium sulfate concentrations, it is possible, by increasing the salt concentration in steps and collecting the precipitate that appears at each stage, to obtain considerable purification.

Dialysis. Salts and organic impurities of low molecular weight can be easily removed from enzymes by dialysis. This consists of placing the enzyme solution in a bag made from a semipermeable membrane, such as cellophane, and immersing it in water or in a buffer solution. The protein remains in the bag, while solutes of low molecular weight diffuse out. An enzyme may be separated from smaller or larger proteins by gel filtration chromatography or by

preparative ultracentrifugation. *See* CHROMATOGRAPHY; ULTRACENTRIFUGE.

Other techniques. Other methods frequently used for enzyme purification include precipitation at low temperatures with organic solvents, such as acetone or alcohol; selective adsorption on gels, such as calcium phosphate or alumina; and chromatography on ion-exchange resins or modified cellulose. For the complete purification of an enzyme, a combination of these methods must be employed. No generally applicable enzyme purification procedure exists. The most satisfactory method must be determined empirically for each enzyme.

The final step in the purification of many enzymes is crystallization. The first crystalline enzyme, urease, was prepared by J. B. Sumner in 1926, and since then many other enzymes have been crystallized.

Isozymes. It has been found that many enzyme preparations, even after crystallization, still consist of a number of distinct proteins, each capable of catalyzing the same reaction. These different enzymes, originating from the same organism or even from the same cells, have been named isozymes. They can usually be separated from each other by electrophoretic or chromatographic techniques. For example, in many tissues the oxidation of lactic acid to pyruvic acid is found to be catalyzed by five different isozymes. These isozymes differ not only in electrophoretic mobility but also in some catalytic properties. Each lactic dehydrogenase molecule is composed of four subunits, of which the two main types are the H and M subunits. The existence of five isozymes is accounted for by the fact that any combination of H and M subunits may assemble in a group of four to form an enzyme molecule. Thus there exist some lactic dehydrogenase molecules containing four H subunits (H_4), some with four M subunits (M_4), and some with intermediate properties having the compositions H_3M, H_2M_2. It may be that one of the physiological functions of isozymes is to give the cell more flexibility in metabolic regulation.

Specificity. The high degree of specificity exhibited by enzymes is one of their most characteristic properties.

Group specificity. The majority of enzymes catalyze only one type of reaction and act on only one compound or on a group of closely related compounds. There must exist between an enzyme and its substrate a close fit, or complementarity. This relationship has been compared to that between a lock and key. In many cases, a small structural change, even in a part of the molecule remote from that altered by the enzymatic reaction, abolishes the ability of a compound to serve as a substrate. An example of an enzyme highly specific for a single substrate is urease, which catalyzes the hydrolysis of urea to carbon dioxide and ammonia. A variety of urea derivatives which have been tested are unaffected by this enzyme. On the other hand, some enzymes exhibit a less restricted specificity and act on a number of different compounds that possess a particular chemical group. This is termed group specificity.

Stereospecificity. A remarkable property of many enzymes is their high degree of stereospecificity, that is, their ability to discriminate between asymmetric molecules of the right-handed and left-handed configurations. An example of a stereospecific, as well as a group-specific, enzyme is L-amino acid oxidase. This enzyme catalyzes the oxidation of a variety of amino acids of the type $R—CH(NH_2)COOH$. The rate of oxidation varies greatly, depending on the nature of the R group, but only amino acids of the L configuration react. Another enzyme, D-amino acid oxidase, specifically attacks amino acids of the D configuration. Conversely, when an enzyme catalyzes the formation of an asymmetric compound from symmetric substrates, the product is almost always optically active, that is, only one of the possible stereoisomers of the product is formed. For example, only the L isomer of malic acid is formed by the enzyme fumarase from fumaric acid and water, as shown in reaction (9).

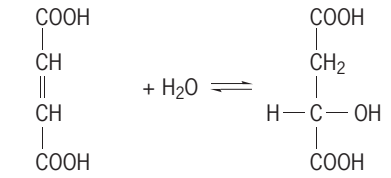

Fumaric acid L-Malic acid (9)

An enzyme may also act asymmetrically on a symmetrical substrate to yield a symmetrical product. Such specificity can be detected by the use of isotopes. Thus alcohol dehydrogenase, which catalyzes the conversion of ethanol to acetaldehyde, removes a specific hydrogen atom from carbon atom 1 of ethanol. This hydrogen atom is transferred to the nicotinamide ring nicotinamide adenine dinucleotide (NAD). If reduced NAD containing deuterium (D) is used in the reversal of this reaction, asymmetrically labeled ethanol is produced, as shown in reaction (10). No deuterium is found in

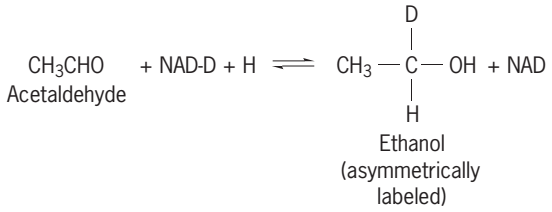

(10)

the acetaldehyde arising from enzymatic dehydrogenation of this ethanol. *See* NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD).

The enzyme, however, cannot distinguish between hydrogen and deuterium. This is shown by the fact that when acetaldehyde containing deuterium on carbon atom 1 is hydrogenated with unlabeled NAD, the other isomer of deuterium-labeled ethanol is formed, as shown in reaction (11). On subsequent

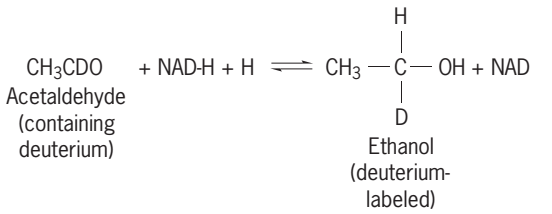

(11)

enzymatic dehydrogenation of this ethanol, all of the

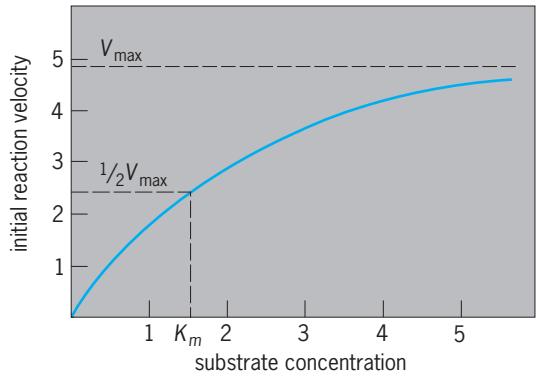

Fig. 1. Initial velocity of an enzyme-catalyzed reaction as a function of substrate concentration.

deuterium is found in the acetaldehyde. *See* PROCHIRALITY; STEREOCHEMISTRY.

Kinetics. Much information can be obtained concerning the mechanism of an enzyme-catalyzed reaction by studying the effect of pH, temperature, and the concentration of substrates, coenzymes, inhibitors, and activators on the velocity of the reaction. Such studies were already being carried out in the nineteenth century, long before anything was known about the chemical nature of enzymes.

Michaelis-Menten equation. It has been known for a long time that, unlike the velocity of many uncatalyzed chemical reactions, which increases linearly with the concentration of one of the reactants, the velocity of enzymatic reactions levels off and approaches a maximal value as the substrate concentration increases (**Fig. 1**). A mathematical theory accounting for this behavior was first proposed by V. Henri in 1903 and extended by L. Michaelis and L. M. Menten in 1913. The theory was based on the idea that the reaction velocity is proportional to the concentration of an enzyme-substrate complex *ES* formed by a specific combination of the enzyme *E* with its substrate *S*. This complex was assumed to react further, yielding a product *P* and regenerating a free enzyme. These reactions can be represented by reaction (12), where k_1, k_2, and k_3 are rate con-

$$\mathrm{E} + \mathrm{S} \underset{k_2}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{k_3} \mathrm{E} + \mathrm{P} \tag{12}$$

stants for the three reactions, whose velocities are v_1, v_2, and v_3, respectively.

The existence of an enzyme-substrate complex has been amply confirmed experimentally, and in some cases such a complex has been directly observed.

Under the conditions usually employed in kinetic studies, the concentration of enzyme is very small compared to that of substrate. When enzyme and substrate are mixed, the concentration of *ES* increases rapidly until a steady state is reached in which the rate of decomposition of the complex equals its rate of formation. Such a steady state is usually reached in less than a second and may last for several minutes. During this time, the velocity of the overall reaction is essentially constant. This velocity, which may be measured as the rate of product formation, is termed the initial velocity of the reaction. An expression for the initial velocity v_0 in terms of the substrate concentration $[S]$ can now be derived simply.

According to the law of mass action, $v_1 = k_1[S][E]$, $v_2 = k_2[ES]$, and $v_3 = k_3[ES]$; by definition, $v_0 = v_3$. The total enzyme concentration $[E_t]$ is $[E] + [ES]$, and the steady-state conditions state that $v_1 = v_2 + v_3$. These relations can be solved for v_0 to give Eq. (13).

$$v_0 = \frac{k_3[E_t]}{1 + \frac{k_2 + k_3}{k_1[S]}} \tag{13}$$

From Eq. (13), it is seen that, for any given substrate concentration, the initial velocity is directly proportional to the enzyme concentration. As the substrate concentration tends to infinity, the initial velocity approaches its maximal value (v_{max}), as expressed by notation (14).

$$v_{max} = k_3[E_t] \tag{14}$$

If K_m is defined as $k_2 + k_3/k_1$, the expression for v_0 becomes Eq. (15). This is the Michaelis-Menton equa-

$$v_0 = \frac{v_{max}}{1 + \frac{k_m}{[S]}} \tag{15}$$

tion, and the Michaelis constant K_m represents the substrate concentration required to obtain one-half the maximum velocity with a given concentration of enzyme. When k_3 is small compared to k_2, K_m approximates the dissociation constant of the enzyme-substrate complex and is therefore a measure of the affinity of an enzyme for its substrate. The equation predicts that, when v_0 is plotted as a function of $[S]$, a rectangular hyperbola should be obtained, as shown in Fig. 1. This closely approximates the results found with a number of enzymes. Most enzymatic reactions, however, involve more than one substrate, and a number of different enzyme-substrate complexes may be formed before the appearance of the product and the regeneration of free enzyme. It may also be necessary in some cases to take into account the effect of product, inhibitors, or activators on the velocity of the reaction. With new instrumentation, it has been possible to study the kinetics of the transient, pre-steady-state portion of the reaction and to determine the value of the individual rate constants for some enzymatic reactions.

Temperature optima. Chemical reactions are usually accelerated by an increase in temperature, because molecules require a minimum quantity of energy to react and at higher temperatures a greater proportion of the molecules possess this energy. When the rate of an enzymatic reaction is determined as a function of temperature, it is found that the rate increases with temperature up to a point and then abruptly decreases. This is because most enzymes are unstable and undergo rapid denaturation above a critical temperature. The temperature optimum for many enzymatic reactions is about 37°C (99°F).

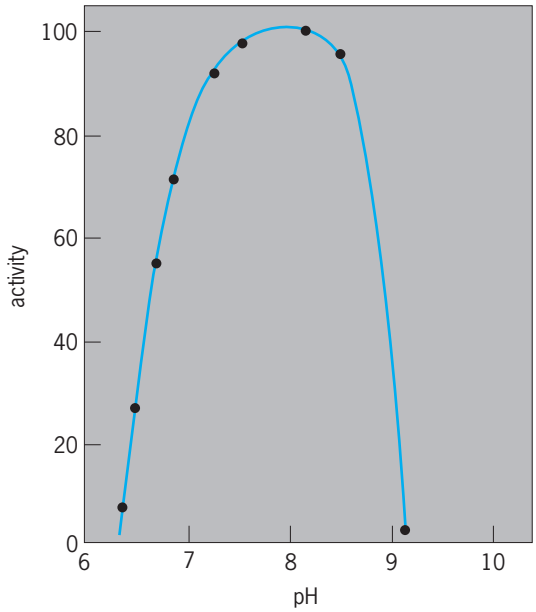

Fig. 2. Effect of pH on the velocity of the reaction catalyzed by the enzyme carbamyl phosphate synthetase.

pH optima. Enzyme activity is also influenced considerably by the hydrogen-ion concentration of the reaction medium (**Fig. 2**). An enzyme may exist in several different ionic forms, and the ionic species that predominates is a function of pH. Only one of these forms can be catalytically active, or one form may be more active than the others. The pH optimum for different enzymes varies over a broad range, from about pH 2 for pepsin to about pH 10 for arginase. Some enzymes exhibit a broad optimal range extending over several pH units, while others have a very sharp pH optimum. The pH optimum for an enzyme, like the temperature optimum, may vary, depending on the substrate being used and on other experimental conditions. Daniel Wellner

Immobilization. In many instances, the use of soluble enzymes is not feasible because of high cost. Immobilization of the enzymes offers the benefit of reuse of the catalyst, thus reducing expenses, and has a second possible advantage of enhanced reaction stability.

The principal techniques for enzyme immobilization are adsorption, covalent attachment, and entrapment within polymer gels. Many enzymes have a preponderance of either positive or negative amino acid side chains on the surface, and electrostatic attraction at multiple points between a solid support and an enzyme leads to a stable complex. Hydrophobic forces may also contribute to stability of noncovalent complexes of enzyme with water-insoluble materials. Covalent attachment requires reagents that create chemical linkages (between enzyme and support). For entrapment, the enzyme is mixed with polymerizable elements, such as vinyl monomers or "prepolymers." The polymerization yields a gel containing the enzyme entrapped within it.

Properties of immobilized enzymes. In most cases, enzymatic activity is reduced as a result of attachment of enzymes to solid supports. Loss of activity may occur either because of chemical reaction of a functional group at or near the enzyme active center or because of physical denaturation. When charged supports are used, microenvironmental effects result in the perturbation of the pH-rate profile and the apparent reactant-binding capability of the enzyme. Both effects are reduced by increased ionic strength of the bulk solution. The transport of reactant to the surface of an enzyme particle can be limited by the unstirred (Nernst) layer surrounding the particle. The supply of reactant to enzyme molecules in the interior of the supporting particle may also be rate-limiting.

Immobilized enzyme reactors. The attachment of enzymes to insoluble supports permits their use in reactors which resemble those employed in the chemical process industry. Tank reactors are often preferred for intermittent, small-scale production. Viscous reactant solutions and enzyme particles with low activity might also dictate the use of tank reactors. For large-scale processes, continuous packed-bed reactors are frequently selected. Upward-flow reactors are used when the enzyme particles are compressible and when gaseous products are produced. The fluidized-bed reactor, where the enzyme particles are kept in suspension by the upward flow of reactant, can be used with viscous reactant solutions or when gaseous products are produced.

The development of ultrafiltration membranes and hollow-fiber devices has led to the construction of reactors which depend on the fact that the enzyme (of high molecular weight) can be retained by the membrane, while the product (of low molecular weight) is allowed to pass. Such reactors are not appropriate for processes in which proteases are present by choice or as contaminants; proteases are often subject to autolysis, which means that they digest themselves as well as the reactant. Contaminating proteases can severely limit the lifetime of an enzyme in this type of reactor. In contrast, when the enzyme preparation is covalently immobilized on particles, the contaminants are also immobilized (or washed through), and proteolytic degradation is eliminated.

Application of immobilized enzymes. A large-scale process based on immobilized enzyme technology is the production of high-fructose corn syrup, much of which is used as a sweetener for soft drinks. In this process, glucose syrup (made enzymatically from cornstarch) is passed over a bed of immobilized glucose isomerase; about 50% of the glucose is isomerized to fructose. Because fructose is sweeter than glucose, high-fructose corn syrup is sweeter than the original glucose solution; its sweetness is equivalent to a solution of sucrose (table sugar) of comparable concentration.

In the production of 6-amino penicillanic acid from penicillin G, the phenylacetyl moiety is released from the amide linkage, but the lactam amide is left intact. The transformation occurs quantitatively under very mild conditions. This application is an excellent example of the exploitation of enzymatic specificity.

In the production of optically pure amino acids

from a racemic mixture with immobilized acylase, the enzyme exhibits preference for one optical isomer; only the L enantiomer is deacylated. The resulting L-amino acid crystallizes, and the acyl-D-amino acid remains. The process can be run continuously by reracemization of the D isomer to the D,L mixture.

Chemical analysis is a very important area of application for immobilized enzymes. Here again, enzyme specificity is exploited. Enzymes can selectively transform a single compound in a complex mixture. The immobilized enzyme may be coupled with a thermal detector or an ion-selective electrode. Coating the latter with an enzyme-polymer layer results in an "enzyme electrode." An even simpler approach is to cover the tip of a conventional electrode with an ultrafiltration membrane which contains the desired enzyme solution. A cation-selective electrode immersed in a solution of glutamate dehydrogenase has been successfully used for the detection of glutamic acid. *See* ION-SELECTIVE MEMBRANES AND ELECTRODES.

Medical applications also appear promising. For example, a deheparinization method for blood has been proposed. The use of extracorporeal devices for blood oxygenation during bypass surgery or kidney dialysis requires heparinization of the blood to prevent clotting. However, heparin treatment can lead to hemorrhagic complications. To prevent these problems, an immobilized heparinase reactor is placed in the external circuit, prior to the point of blood reentry. Heparinase in the column completely eliminates the heparin anticoagulant activity, and heparinization is restricted to the external circuit. Immobilized heparinase may find applications in other extracorporeal devices used in enzyme therapy.

Certain types of leukemia are asparagine-dependent; when the tumor cells are deprived of asparagine, proliferation slows dramatically. The enzyme L-asparaginase degrades asparagine. Depletion of asparagine by direct injection of asparaginase into a patient is accompanied by two problems which relate to the fact that the asparaginase most commonly used is of bacterial origin. The reticuloendothelial system recognizes the enzyme as alien and clears it from the bloodstream. Also, life-threatening allergic reactions may result following administration of the foreign protein asparaginase. To immobilize the enzyme within biocompatible capsules would appear to be a potential solution to the problems. Unfortunately, immobilized enzymes are not generally good catalysts at low concentration of reactants. Since the tumor cells can survive at low levels of asparagine, effective enzyme therapy requires reduction of the asparagine concentration to very low levels. The consequence is that leukemia therapy based on the use of immobilized asparaginase remains in the experimental stage.

G. P. Royer

Production

A unique property of enzymes, both proteinaceous and RNA-derived, is the ability to catalyze chemical reactions in a mild aqueous environment, that is, with relatively low temperatures (58–212°F or 15–100°C), atmospheric pressure, and moderate pH. Thus enzymes are indispensable where harsh chemical treatment or extreme conditions are unacceptable, such as in the baking, brewing, food processing, detergent, leather, and textile industries. Novel enzymes have been developed which catalyze chemical reactions in nonaqueous organic solvents, at elevated temperatures, under conditions of extreme pH. Moreover, the chemical specificity with which enzymes catalyze chemical reactions in mild or relatively harsh conditions has led to the use of enzymes for industrial-scale biochemical processes, including synthesis of peptides, chemical feedstock compounds, and specialty compounds. *See* FOOD MANUFACTURING; LEATHER AND FUR PROCESSING.

Sources. Enzymes for traditional industrial use (for example, food processing, laundry detergents, and leather processing) must be produced in large quantities and at modest cost. Consequently, they frequently are obtained as by-products from biological materials that are available in large quantities. Three common sources are plants (malt diastase, papain, bromelain, and ficin), animals (pancreatic enzymes, and digestive enzymes such as lipases, pepsin, and rennin), and microbes (amylases, proteases, cellulases, lipases, pectinases, invertase, and glucose oxidase). The number of plant- and animal-derived enzymes tends to be restricted by the availability of raw material and cost of recovery. For example, animal-derived enzymes are a low-cost by-product of the meat packing industry. However, the supply of animal-derived enzymes is restricted by the number of animals processed at meat packing plants.

To ensure a consistent supply of low-cost enzymes, industrial enzyme manufacturers have relied on microbes as a source. In fact, even specialty industrial enzymes used in the biomedical and biotechnology industries are obtained from microbes.

The use of bacteria, fungi, and yeast can be attributed directly to the fact that they can be grown rapidly in large quantities on inexpensive culture media. Moreover, microbes can be manipulated by traditional mutagenesis and modern genetic engineering techniques to increase the yield of specific enzymes. Thus, enzymes derived from microbes are available in virtually unlimited amounts at high yield and lower cost than those derived from plants and animals.

Fermentation. The process by which microbes are cultivated is called fermentation. When used by bacterial physiologists, this term refers to the anaerobic metabolism of growth substrates. However, in industry the term is applied to any process of cultivating microbial cells in a liquid medium. Microbial fermentation takes two forms, submerged fermentation and semisolid culture. In recent times, the former has predominated, and the latter has been used to only a very limited extent. Production can be divided into three stages: ferme recovery, by which a concentrated solution of enzyme is produced, enzyme recovery, and purification. *See* BACTERIAL PHYSIOLOGY AND METABOLISM.

A pure culture of the organism is inoculated into a sterilized liquid broth (250–1000 milliliters) contained in Erlenmeyer flasks. The growth medium

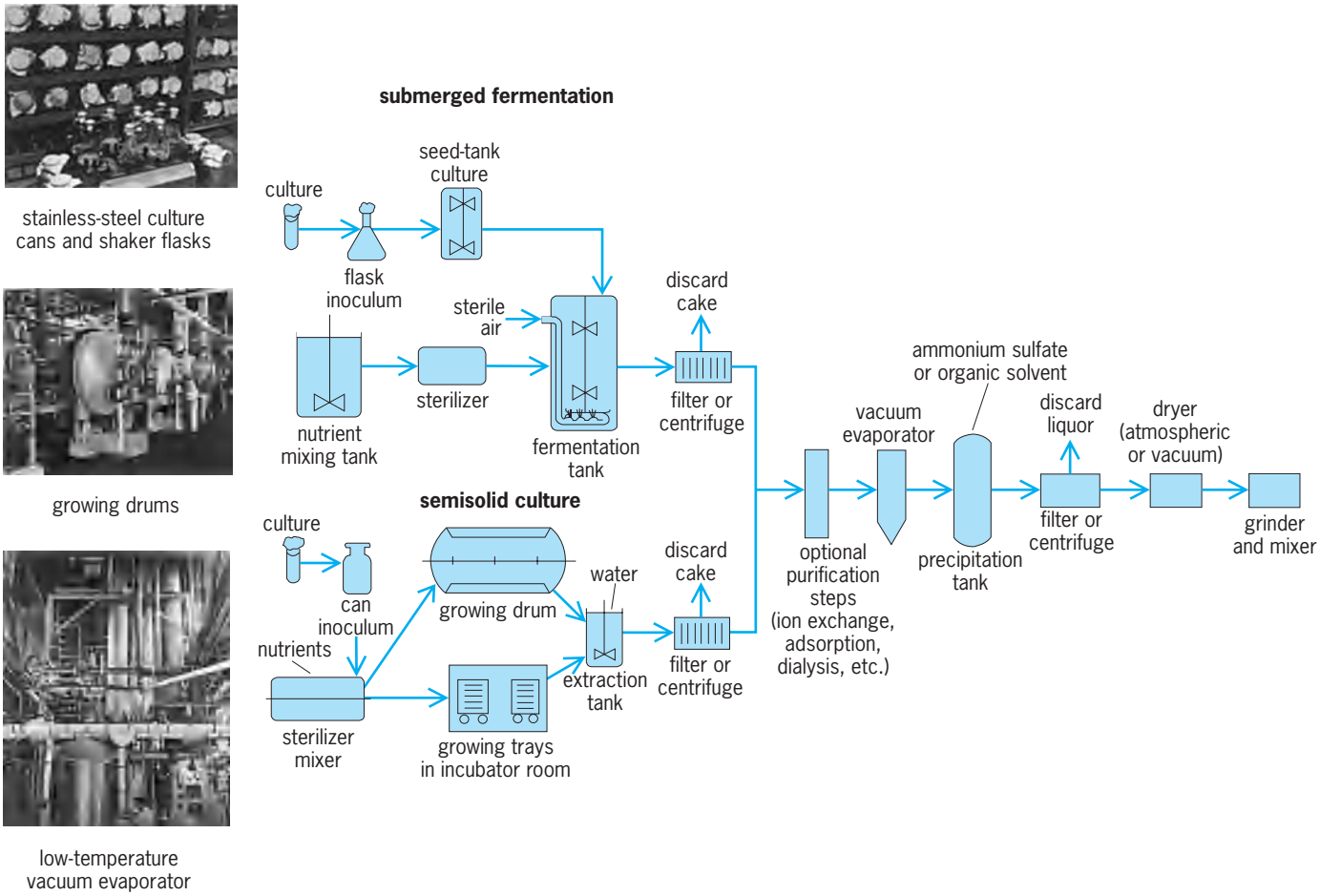

Fig. 3. Procedures for enzyme production by industrial fermentation. ***(Wallerstein Co., Inc.)***

typically employed in the initial and final stages of fermentation is composed of an inexpensive carbon and energy source (potato starch, corn syrup), a nitrogen and sulfur source (hydrolyzed casein or beef extract, ammonia and sulfur salts), minerals, and growth stimulants. The growth in flasks is used to inoculate seed tanks containing 10 to several hundred liters of sterilized culture broth. In the final stage of submerged fermentation, the contents of the seed tanks are transferred to a culture medium in production tanks (4000–100,000 liters). These are stainless steel tanks equipped for the passage of air sterilized by filtration, propeller agitation, temperature control by water jacket or thermal coils, pH control, and nutrient addition (**Fig. 3**).

Depending on the microorganism, enzyme production is maximal after 1 day to a week. The fermentation tank is then emptied for enzyme recovery. *See* CULTURE; FERMENTATION.

Enzyme recovery. The first step in recovery is to concentrate the enzyme preparation. The procedure depends on whether the enzyme is contained within the cell cytoplasm (glucose oxidase, invertase) or is secreted into the medium (protease, amylase, pectinase). Enzymes contained within cells can be concentrated by centrifugation. This is followed by disintegration of cells by autolysis or a mechanical method such as high-frequency sonic vibration. In the former procedure, enzymes produced by cells degrade the microbial cell wall, resulting in cell lysis and release of the contents. Enzymes are recovered by vacuum evaporation after intact cells and other insoluble constituents have been removed by centrifugation or filtration. *See* CENTRIFUGATION; FILTRATION.

Purification. Enzyme concentrates prepared as described above are frequently marketed without further purification after addition of suitable buffering and stabilizing agents. Depending on the application, further steps may be required before marketing. Additional purification may employ standard biochemical methods such as fractional precipitation by solvents (methanol, acetone, isopropanol) or salts (ammonium sulfate, calcium acetate). Fractional precipitation is rapid, inexpensive, and applicable to very crude enzyme suspensions. It is often the first step in purification of crude broths, primarily to remove low-molecular-weight compounds with objectionable odors or tastes.

Techniques such as adsorption, ion-exchange chromatography, and gel filtration chromatography, although uncommon for enzymes used for traditional industrial applications, are typical for the purification of enzymes used in the biomedical and biotechnology industries. The added purification steps are often included to eliminate cross-contamination by enzymes or other products that would render the enzymes unsuitable for use. This is particularly true for enzymes used in the biomedical field (such

as streptokinase, ceredase, and DNase) that require intravenous administration. Enzyme preparations contaminated with endotoxin, a natural byproduct of gram-negative bacteria, result in septic shock upon intravenous injection. This potentially life-threatening reaction can be eliminated only by rigorous purification of enzyme preparations prior to marketing. *See* BACTERIA; ENDOTOXIN.

Genetic engineering. The availability of industrial enzymes has been limited by identification of suitable, naturally occurring sources of the enzyme. Genetic engineering, including recombinant DNA cloning, site-directed mutagenesis, and the polymerase chain reaction, has made it possible to extend the range of naturally occurring sources and to create completely novel enzymes that are tailored to meet the specific requirements of industrial processes.

The recombinant DNA cloning and expression of foreign genes from animal, plant, and microbial cells in high-yielding bacterial cells has generated a virtually unlimited supply of enzymes. The cloning of enzyme-encoding genes on autonomously replicating extrachromosomal DNA elements (plasmids) frequently results in the high-level expression of the cloned gene product. High-level expression results because certain plasmids exist in hundreds of copies per bacterial cell. The presence of hundreds of copies of a gene that encodes a single enzyme enables the cell to synthesize an enormous quantity of the cloned gene product compared to the source from which the gene was obtained originally. Consequently, recombinant DNA cloning of an enzyme-encoding gene can provide, in certain instances, high yields of the enzyme irrespective of the gene's source.

Recombinant DNA manipulations have been used to further enhance enzyme production from cloned genes by genetically manipulating promoters. These are DNA sequences, located upstream from the region of the DNA encoding the gene product, that are required for transcription and translation of the gene product. Enzyme manufacturers are, therefore, no longer restricted to the use of enzymes from high-yield natural sources. It is possible to identify an enzymatic activity that suits the needs of an existing process, clone the gene encoding the desired enzyme, and manipulate the gene's expression in a bacterial or other suitable host to provide a continuous and relatively low cost supply of the enzyme.

Recombinant DNA methods also provide the means to create novel enzymes. The structure of individual genes can be manipulated by fusion or by the rearrangement of distinct genes from different organisms, thereby changing the characteristics of the product encoded by the gene. In combination, these methods are used to create and modify the activities of enzymes used in industrial processes.

Another extremely powerful technology is the use of random mutagenesis of cloned genes by the polymerase chain reaction technique combined with screening of the mutagenized gene product under selective conditions. Although the approach is referred to as random mutagenesis, the mutagenesis is restricted to a specific region of a cloned gene bounded by polymerase chain reaction primers. By restricting the location of mutations, it is possible to target mutations to a general region of cloned DNA, for instance the region corresponding to the active site of an enzyme. After generating DNA that has been randomly mutated within a defined region, the DNA is cloned into appropriate sites within the gene of interest. The mutated genes are introduced into a host organism that can express the altered enzymes derived from the mutated genes. Cells containing the mutated genes are then screened under selective conditions to reveal those cells that are expressing altered gene products with novel activities. This method has been used to obtain enzymes with altered activities compatible with industrial uses. The development of enzymes with catalytic activity in nonaqueous solvents has profound implications for industrial use, particularly since many industrial processes are performed in organic rather than aqueous solvents. Moreover, combining the use of random polymerase chain reaction mutagenesis with screening under selective conditions circumvents the need for time-consuming and expensive x-ray crystallographic analysis of enzyme three-dimensional structure. *See* GENETIC ENGINEERING; MOLECULAR BIOLOGY.

Edmund J. Stellwag

Bibliography. P. B. Chock et al. (eds.), *Enzyme Dynamics and Regulation*, 1987; P. A. Frey (ed.), *Mechanisms of Enzymatic Reactions: Stereochemistry*, 1986; P. Gacessa and J. Hubble, *Enzyme Technology*, 1987; C. P. Hollenberg and H. Sahm (eds.), *Microbial Genetical Engineering and Enzyme Technology*, 1987; T. Palmer, *Understanding Enzymes*, 1985.

Enzyme inhibition

The prevention of an enzymatic reaction due to interaction of an enzyme with a substance that decreases the rate of catalysis by the enzyme. The compound causing the effect is termed an inhibitor. Enzyme inhibitors are important regulators in the normal control of enzymatic reactions in living organisms, are useful in the study of enzyme mechanisms and cellular reactions, and are important chemotherapeutic agents. Inhibitors may have chemical structures that are similar to one of the substrates of the enzyme but may also be quite different.

Classification of inhibitors. Inhibitors bind reversibly to different forms of the enzyme during the catalytic reaction. They include dead-end inhibitors, which may resemble structurally one of the substrates but do not undergo the catalytic reaction; and reaction products, which can act to make the reaction proceed in the reverse direction. In reaction (1), substrate A is converted to product

$$A \rightleftharpoons P \tag{1}$$

P. The simplest enzymatic mechanism is shown in

reaction (2), where the enzyme E forms an enzyme–

$$E + A \rightleftharpoons EA \rightleftharpoons E + P \qquad (2)$$

substrate complex EA, which then converts A to P and releases the product. A dead-end inhibitor (I, where subscripts 1, 2, 3 refer to different chemical compounds) may bind to the free enzyme, reaction (3), or to the EA complex, reaction (4). In principle,

$$E + I_1 \rightleftharpoons EI_1 \qquad (3)$$

$$EA + I_2 \rightleftharpoons EAI_2 \qquad (4)$$

an inhibitor also could bind to both forms of enzyme. Most enzymatic reactions have more than one substrate and product, and the reaction of two substrates to form two products is shown in reaction (5). One possible bi-substrate enzyme mechanism, reaction (6) shows that the enzyme binds the substrates

$$A + B \rightleftharpoons P + Q \qquad (5)$$

$$E \overset{A}{\rightleftharpoons} EA \overset{B}{\rightleftharpoons} EAB \rightleftharpoons EPQ \underset{P}{\rightleftharpoons} EQ \underset{Q}{\rightleftharpoons} E \qquad (6)$$

and releases the products in an ordered sequence. For such a mechanism, an intermediate form of enzyme may bind inhibitors, as in reaction (7).

$$EQ + I_3 \rightleftharpoons EQI_3 \qquad (7)$$

Three kinds of inhibitors are classified and distinguished experimentally by a study of the effects on the rate of the enzymatic reaction at different concentrations of the inhibitor and varied concentrations of one of the substrates.

Competitive inhibition. Competitive inhibition results when the inhibitor binds reversibly to the same form of enzyme to which the substrate binds. For instance, the dead-end inhibitor I_1 in reaction (3) is a competitive inhibitor against varied concentrations of substrate A. The product P in reaction (2) or the product Q in reaction (6) are also competitive inhibitors against substrate A, as these compounds bind to enzyme form E.

Uncompetitive inhibition. Uncompetitive inhibition results when a dead-end inhibitor binds to a form of the enzyme after the substrate binds, for instance to EA, as shown in reaction (4). Uncompetitive inhibition can also result when the inhibitor binds to a subsequent intermediate form of enzyme, such as in reaction (7) after substrates A and B bind.

Noncompetitive inhibition. Noncompetitive inhibition can result from a variety of reactions, for instance, if a dead-end inhibitor binds to both free enzyme and to the enzyme–substrate complex, as in reactions (3) and (4), where I_1 and I_2 are the same compound. Inhibitor I_1 and product P are also noncompetitive inhibitors against substrate B in reaction (6). *See* CATALYSIS.

Kinetics of inhibition. The inhibition effects are described quantitatively by a general equation (8),

$$v = VS/[K_m(1 + I/K_{is}) + S(1 + I/K_{ii})] \qquad (8)$$

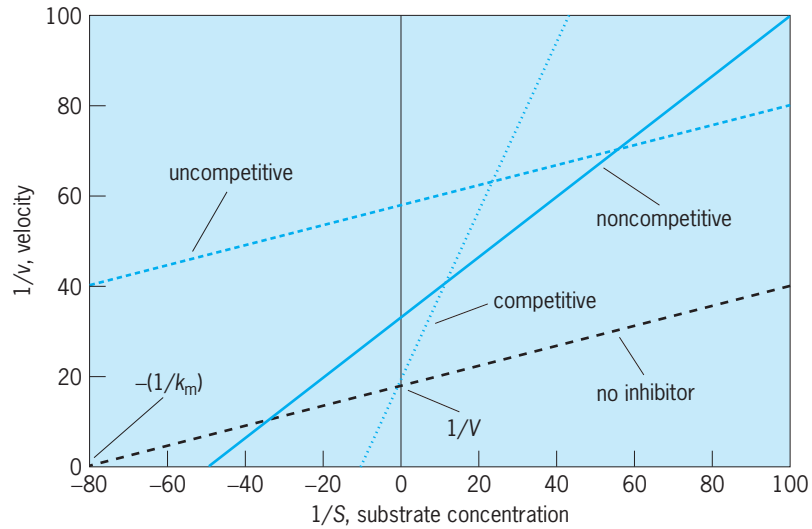

Fig. 1. Different kinds of inhibition.

which is a modification of the Michaelis-Menten equation and which can be written in its double reciprocal form (9). Here v is the initial velocity of

$$1/v = (K_m/V)(1 + I/K_{is})(1/S) + (1/V)(1 + I/K_{ii}) \qquad (9)$$

the reaction at various concentrations of a substrate (S), V is the maximum velocity of the reaction when the substrate concentration is much higher than the Michaelis constant (K_m), and K_{is} and K_{ii} are dissociation constants named for the effects of the inhibitors on slopes and intercepts, respectively, of a plot of $1/v$ against $1/S$. Thus, the different kinds of inhibition can be distinguished experimentally by studying the effects of one or more concentrations of inhibitor on the velocity as represented by a double reciprocal (**Fig. 1**).

A competitive inhibitor affects the slope but not the intercept ($1/V$) at infinite concentration of S of a plot of Eq. (9), because K_{is} has a finite value, whereas the term $(1 + K_{ii})$ is not present in the equation. Since the competitive inhibitor and the substrate bind to the same form of enzyme [such as reactions (2) and (3)], when the substrate concentration is sufficiently high to saturate the enzyme with substrate, the inhibitor has no effect, and the same maximum velocity is obtained as in the absence of inhibitor. In contrast, an uncompetitive inhibitor affects only the intercept of the plot, because K_{ii} has a finite value and the term $(1 + K_{is})$ is not present in the equation. Since the uncompetitive inhibitor and the substrate bind to different forms of enzyme [see reactions (2) and (4)], high substrate concentrations will not prevent the action of the inhibitor, the apparent maximum velocity will be lower, and the apparent K_m value will be higher than in the absence of the inhibitor. A noncompetitive inhibitor affects both the slopes and intercepts of the plot, because K_{is} and K_{ii} have finite, and generally different, values. In general, the apparent K_m value for the substrate, determined from the reciprocal of the intercept on the abscissa ($1/K_m$), will be altered by an inhibitor. If an enzyme is

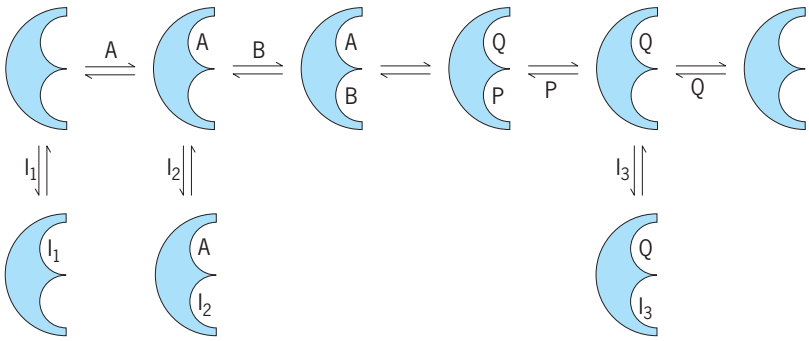

Fig. 2. A typical enzyme may have two binding sites and be subject to be different kinds of inhibition.

inactivated by an irreversible chemical reaction, the kinetics resemble noncompetitive inhibition with equal values for K_{is} and K_{ii}, simply because there is less enzyme present and the apparent K_m value is not changed by the inhibitor. *See* ENZYME.

Binding modes. Identifying where an inhibitor binds to an enzyme usually requires that the structure of a complex of the enzyme and the inhibitor be determined by a physical method such as x-ray crystallography or nuclear magnetic resonance (NMR). Defining the binding site by the type of inhibition is ambiguous. A competitive inhibitor may bind either to the same site as the substrate or to a different site. A noncompetitive inhibitor may also bind to the same or different sites.

With reference to reaction (6), a typical enzyme may have two binding sites (**Fig. 2**). Substrates A and B bind to adjacent sites and react to form products P and Q. As a product inhibitor, Q is competitive against varied concentrations of substrate A, since both compounds bind to free enzyme. In contrast, product P is a noncompetitive inhibitor against substrate B, since B binds to EA and P binds to EQ, and the presence of P reverses the reaction sequence. Thus, the inhibition is noncompetitive even though B and P bind to the same site. For the dead-end inhibitors, I_1 is competitive against A and noncompetitive against B, I_2 is competitive against B and uncompetitive against A, and I_3 is competitive against P and uncompetitive against either A or B.

Binding of an inhibitor or substrate may also change the conformation of the enzyme and prevent binding of the other, as shown in Fig. 2, where the inhibitor binds at an allosteric site, which is different from the site to which substrate binds. In this case, the inhibition is competitive, because the substrate and inhibitor bind to the same form of enzyme in a mutually exclusive manner. In general, an inhibitor binding at an allosteric site can be noncompetitive, because it can bind to both the free enzyme and the enzyme–substrate complex. *See* ALLOSTERIC ENZYME.

Inhibitors and metabolism. Enzymes catalyze the reactions in the metabolism of foodstuffs that yield energy and the materials of living cells. The reactions are controlled by the activities of the enzymes, which are inhibited by products of the reactions and inhibited or activated by various allosteric effectors. The controls are required to maintain the optimal steady-state concentrations of metabolites. For instance, the three dehydrogenases in the citric acid cycle produce the reduced form of nicotinamide adenine dinucleotide and are inhibited by this product when its concentration becomes high.

Another important control is feedback inhibition of an enzyme that catalyzes a rate-limiting step in a metabolic pathway. The end product of the pathway inhibits the enzyme and prevents production of compounds that are used only for the synthesis of the final product. A textbook example is a bacterial aspartate transcarbamoylase, which produces carbamoylaspartate from carbamoyl phosphate and aspartic acid, and is inhibited by cytidine triphosphate, the end product of the ten-step pathway. Interestingly, another nucleotide, adenosine triphosphate, activates the enzyme, and both nucleotides affect the Michaelis constant for aspartic acid and not the maximum velocity. Thus the nucleotides are "competitive," but they bind at an allosteric site, which is structurally distant from the active site where the substrate aspartic acid binds.

Antimetabolites and antibiotics. A major effort in the battle against disease is the synthesis and evaluation of new enzyme inhibitors. Many enzymes are targets of chemotherapy.

Inhibitors that are structurally similar to the substrate can bind to the same site on the enzyme as the substrate. For instance, malonic acid resembles succinic acid and competitively inhibits succinate dehydrogenase, an enzyme in the citric acid cycle (**Fig. 3**). As in Fig. 1, products bind to the same site as the corresponding substrate. However, substrate binding sites in enzymes are flexible and can accommodate structurally related molecules. This has led to the development of many kinds of antimetabolites that have been used to inhibit particular steps in metabolic pathways and to exploit the inhibition for therapeutic purposes. The antibiotic sulfanilamide competitively inhibits the incorporation of *p*-aminobenzoic acid into folic acid, an essential vitamin. The antibiotic is efficacious because bacteria cannot survive when they cannot synthesize folic acid, but humans do not make folic acid and normally consume sufficient amounts in the diet. *See* ANTIBIOTIC; CITRIC ACID CYCLE; SULFONAMIDE.

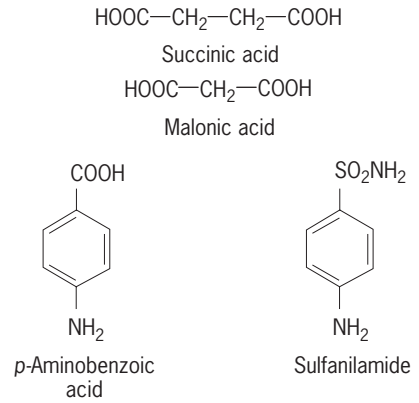

Fig. 3. Structural formulas for two essential metabolites (succinic acid and *p*-aminobenzoic acid) and their competitive inhibitors (malonic acid and sulfanilamide).

The anticancer drug methotrexate inhibits dihydrofolate reductase, which is in the pathway for the biosynthesis of thymidylate and is essential for synthesis of DNA. Methotrexate is a competitive inhibitor against dihydrofolate and binds in the same site of the enzyme where the folate binds. In contrast, nonnucleoside inhibitors of the reverse transcriptase encoded by human immunodeficiency virus bind at an allosteric site some distance from the active site. *See* CHEMOTHERAPY.

Uncompetitive and noncompetitive inhibitors offer a potential kinetic advantage for inhibiting a metabolic step, because high concentrations of the substrate cannot abolish the effect of the inhibitor. Examples are formamides and sulfoxides, which resemble chemically the product aldehyde that forms when alcohol dehydrogenase oxidizes an alcohol. The mechanism is given in reactions (6) and (7), where the alcohol is substrate B and the inhibitor I_3 binds to EQ after the aldehyde product P has dissociated. These inhibitors effectively inhibit alcohol metabolism.

Inhibitors that chemically react at a substrate binding site irreversibly inactivate the enzyme and are also very effective inhibitors. These may be called active-site-directed or suicide inactivators. Thymidylate synthetase is inactivated by a derivative of 5-fluorouracil, and synthesis of nucleotides required for DNA replication is inhibited, which is beneficial in treatment of certain cancers. Aspirin (acetylsalicylic acid) acetylates a serine residue in the active site of prostaglandin H synthetase, and inhibits the cyclooxygenase activity that initiates the conversion of arachidonic acid to physiologically active eicosanoids. Penicillin reacts with an enzyme that is required to make cross-links in bacterial cell walls and thus causes cell lysis in actively growing bacteria. These inactivators are also used to locate the binding site on the enzyme. *See* EICOSANOIDS. Bryce V. Plapp

Bibliography. P. F. Cook and W. W. Cleland, *Enzyme Kinetics and Mechanism*, Garland Science, 2007; S. J. Teague, Implications of protein flexibility for drug discovery, *Nat. Rev.*, 2:527–541, 2003; T. H. Venkataramaiah and B. V. Plapp, Formamides mimic aldehydes and inhibit liver alcohol dehydrogenases and ethanol metabolism, *J. Biol. Chem.*, 278:36699–36706, 2003; D. Voet and J. G. Voet, *Biochemistry*, Wiley, 2004.

Eoacanthocephala

An order (or a class according to some taxonomists) of the Acanthocephala characterized by the presence of a small number of giant subcuticular nuclei which are similar to the embryonic nuclei. Body spines may or may not be present and the chief lacunar vessels are dorsal and lateral. The cement gland of the male is a single syncytial organ with a specialized cement reservoir. Ligament sacs persist in the female. Proboscis hooks are few in number and arranged in circles. Eggs are ellipsoidal and thin shelled. These worms are parasitic in cold-blooded vertebrates (turtles, fish). The cystacanth occurs in crustaceans.

Neoechinorhynchus emydis is a parasitic species in North American turtles. The body is white and typically curved in extended specimens. The females are 0.4–0.9 in. (10–22 mm) long, whereas the males are 0.3–0.6 in. (6–14 mm). The body is 0.2–0.4 in. (0.5–1 mm) wide. There are typically six giant subcuticular nuclei of which five are middorsal and one midventral. The male reproductive organs are in the posterior half of the body, the testes are elongate, contiguous in tandem, and the large syncytial cement gland contains eight nuclei. The proboscis is globular with three circles of six hooks, each in quincuncial arrangement. The lemnisci are unequal in length; the longer contains two nuclei, the shorter only one nucleus. Eggs are oval with three membranes; the middle membrane does not completely encircle the acanthor. The first intermediate host is an ostracod (*Cypria maculata*); the second intermediate host is a snail (*Campeloma rufum*), with the cystacanth encysting in the foot of the snail. The definitive host is the map turtle (*Graptemys geographica*). In this life cycle the snail appears to be an essential second intermediate host and not simply a transport host. *See* ACANTHOCEPHALA. Donald V. Moore

Eocene

The second oldest of the five major worldwide divisions (epochs) of the Tertiary Period, that is, the epoch extending from the end of the Paleocene Epoch to the beginning of the Oligocene Epoch; also referred to as the middle epoch of the older Tertiary (Paleogene of some authors, Nummulitic of earlier French authors). *See* CENOZOIC; OLIGOCENE; PALEOCENE; TERTIARY.

CENOZOIC	QUATERNARY	Holocene
		Pleistocene
	TERTIARY	Pliocene
		Miocene
		Oligocene
		Eocene
		Paleocene
MESOZOIC	CRETACEOUS	

The Eocene/Oligocene boundary has been codified at the 63-ft (19-m) level in the Massignano section of the northern Apennines (Italy) and corresponds to the extinction level of certain planktonic foraminiferal groups. This boundary has an estimated age of 33.7 million years (Ma), based on radioisotopic ages in underlying and overlying ash beds. The Paleocene/Eocene boundary has been formally defined at the 5.2-ft (1.6-m) level in the Dababiya Quarry near the village of Dababya, approximately 22 mi (35 km) south of Luxor in the Upper Nile Valley, Egypt. This level coincides with a global carbon isotope excursion associated with significant climatic warming and biotic changes and is about 1 million years older than the base of the classic Ypresian Stage, normally considered the oldest stage of the

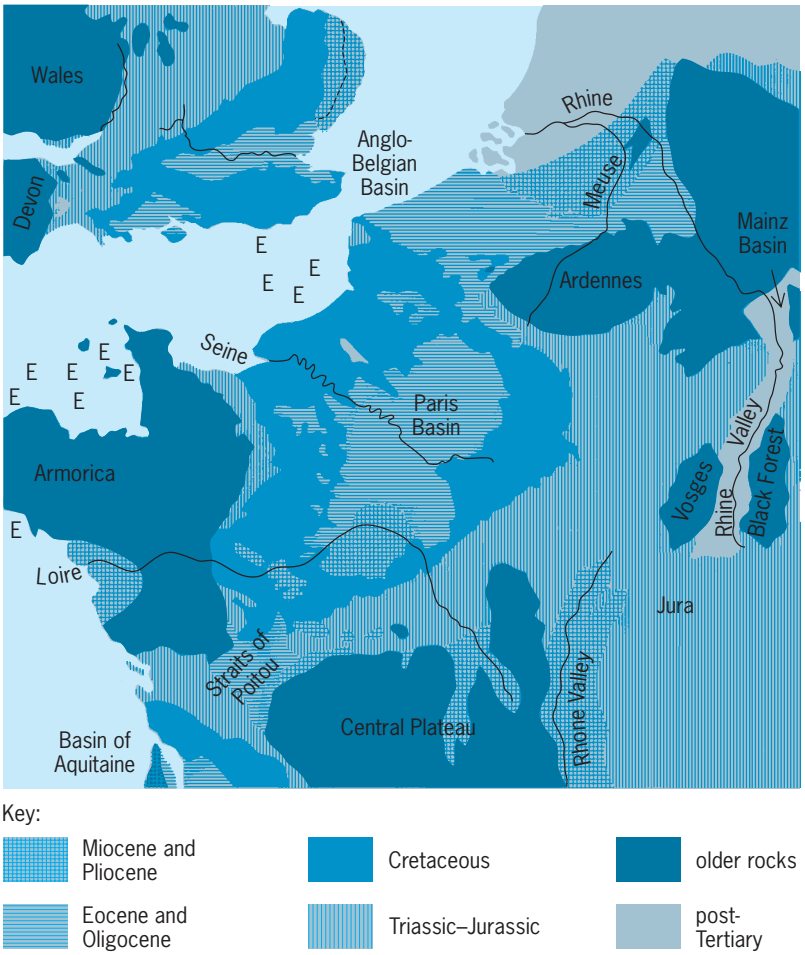

Fig. 1. The Paris and London (Anglo-Belgian) basins and their framework. E = Eocene outcrop. (*After A. M. Davies, Tertiary Faunas, vol. 2, George Allen and Unwin, 1934*)

Eocene. Currently there are two opinions regarding what to do with the associated stage boundaries. One proposes lowering the base of the Ypresian Stage to coincide with the newly relocated and defined Eocene base at the carbon isotope excursion with an estimated age of about 55 Ma. The other proposes retaining the Ypresian Stage in its present position, with an estimated age for its base of about 54 Ma based on the radiosiotopic ages of volcanic ash beds in northwestern Europe, and inserting the Sparnacian Stage as the lowest stage of the newly redefined Eocene. Lowering the Ypresian Stage by nearly 1 million years is counter to normal chronostratigraphic procedure. *See* PALEOMAGNETISM; ROCK AGE DETERMINATION.

Eocene strata are widespread throughout the world and on the deep ocean floor. They include the common sedimentary types and vary from terrestrial, to marginal (estuarine), to normal marine pelagic origin. Igneous activity, while not as extensive as in the later part of the Cenozoic, was notable in some areas such as East Greenland, Oregon, Washington, and British Columbia.

In 1833, Charles Lyell provided the first basic subdivision of the Tertiary Period into three basic units: Eocene, Miocene, and Pliocene. The basis was the percentage of living mollusk species present in each subdivision. The Eocene was said to have less than 5% of living forms in it. *See* MIOCENE; PLIOCENE.

The content (and extent) of the Eocene was subsequently modified with the creation of the Oligocene from the upper part of Lyell's original Eocene and the creation of the Paleocene from the lowermost part of Lyell's Eocene. Rocks of Eocene age are typically well developed in the Anglo-Belgian-Paris basins (**Fig. 1**). The Eocene is developed as a series of relatively shallow-water marls, limestones, and sands, with gypsum (typically developed in the Montmartre district of Paris) characterizing the uppermost part of the Eocene in the Paris Basin.

Subdivisions. The Anglo-Belgian-Paris basins have served as the type area for the establishment of a sequence of standard stage subdivisions of the Eocene. The lower Eocene corresponds to the Ypresian (Cuisian is synonymous), the Lutetian and the English Bartonian to the middle Eocene, and the Italian Priabonian to the upper Eocene. The Bartonian, previously thought to be the boreal, upper Eocene equivalent of the Mediterranean Priabonian, has since been shown actually to lie stratigraphically between the Lutetian and Priabonian and to be a part of the middle Eocene; while the French terms Auversian, Marinesian, and Ludian are seldom used anymore. In the American Gulf and Atlantic Coastal Plain, the terms Wilcoxian, Claibornian, and Jacksonian are used, although they do not correspond precisely to their European counterparts. A local stage classification is used in California based on benthic foraminiferans: Penutian, Bulitian, Ulatisian, and Narizian.

Terrestrial subdivisions include the Wasatchian, Bridgerian, Uintan, and Duchesnean land mammal ages. New Zealand marine stages include the Waipawan (upper part), Mangaorapan, Heretaungan (lower Eocene), Porangan, Bortonian, Kaiatan (middle Eocene), and Runangan (upper Eocene).

Climate and stable isotopes. Early Paleogene temperatures, including those of high latitudes, were the warmest of the Cenozoic; peak warming occurred in the early Eocene. The Earth was in a greenhouse state, with partial pressure of carbon dioxide (pCO_2) levels in the early Eocene estimated to have been six times higher than present-day values. During the late Paleocene to the early Eocene, deep-sea temperatures at high southern latitudes warmed by some 7–9°F (4–5°C), from about 50–52°F (10–11°C) to about 57–61°F (14–16°C), while surface temperatures increased by some 9–11°F (5–6°C), with maximum temperatures in excess of 68°F (20°C). At low latitudes, surface-water temperatures remained relatively constant and comparable to values of the present-day ocean. Superimposed on this long-term trend was a relatively abrupt (<10,000 years) 2.5–3% drop in $\delta^{13}C$ (the difference in isotopic ratios $^{12}C/^{13}C$ between a sample and a standard) and concomitant marine productivity that has been associated, in turn, with a major turnover (extinction of almost 50%) of the deep-sea benthic (bottom-dwelling) forminiferal fauna. This drop in $\delta^{13}C$ has been identified both in marine organisms and in mammalian

bone enamel and paleosol carbonates in terrestrial sections in the Big Horn Basin of the western interior of North America and in the type Sparnacian (that is, earliest Eocene) in the Paris Basin. *See* EXTINCTION (BIOLOGY); GEOLOGIC THERMOMETRY; PALEOSOL.

Theoretical calculations and empirical evidence (decrease in eolian grain size) suggest a decrease in effective atmospheric transport and concomitant upwelling and transport of nutrient flux to the oceanic photic zone during the early Eocene. With the advent of the middle Eocene (Lutetian; about 49 Ma), a long-term, stepwise 13–14°F (7–8°C) decline of low-latitude bottom-water temperatures set in; this change bears a strong resemblance to the fossil record and to stable isotope records of surface-water temperatures at high latitudes over the same interval, suggesting a strong linkage in the thermal evolution of these two areas. As the southern oceans cooled, a more vigorous thermohaline circulation developed and culminated in an accelerated cooling of deep waters in the late Eocene and relatively abrupt drop of 7–9°F (4–5°C) in earliest Oligocene time. This latter event was associated with the growth and expansion of global ice volume (most of which was probably present in the form of glaciers on Antarctica) estimated to have been in excess of 70% of present-day ice volume. The Earth had clearly passed into the icehouse state, in which it remains today.

Sea-level changes. Although the evidence for changes in sea level through time can be clearly read from the alternating history of transgressions and regressions in the stratigraphic record, the causal processes considered responsible for such changes are still unclear. An explanation may relate to global, or eustatic, changes of sea level; more local or regional tectonic events centered in the Earth's continental crust; or the major ocean ridges on the sea floor. A key method for distinguishing between these explanations is by precise stratigraphic correlation; synchrony between unconformities bounding sedimentary sequences in geographically and depositionally separated basins would strongly suggest that global, or eustatic, control of depositional sequences is responsible, whereas nonsynchronous or overlapping unconformities would suggest local or regional tectonic control. During the Neogene, glacioeustacy (globally synchronous and essentially quantitatively equivalent changes in sea level) is generally accepted as the driving force controlling sea-level oscillations; during ice-free periods of Earth's history, and during the early Paleogene (Paleocene-Eocene), this explanation does not account for observations in the rock stratigraphic record. Recent evidence suggests that glacioeustacy may have controlled sea-level fluctuations of 66–82 ft (20–25 m) during the Late Cretaceous. *See* CRETACEOUS; MID-OCEANIC RIDGE.

The history of sea level during the Cenozoic Era is one of gradual and inexorable withdrawal of the sea from the continents, punctuated by significant lowering of the sea level in the late Paleocene and the earliest Oligocene (**Fig. 2**). The latter event is generally considered to be linked with the first major buildup of ice on the Antarctic continent. Major sea-level

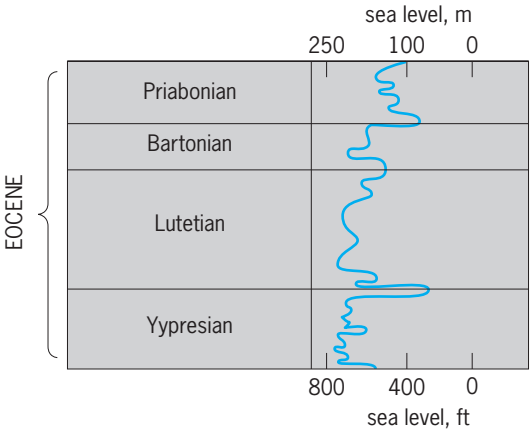

Fig. 2. Eustatic sea-level curve for the Eocene. (*After B. U. Haq, J. Hardenbol, and P. R. Vail, The chronology of fluctuating sea level since the Triassic, Science, 205:1156–1167, 1987*)

rises occurred during the early Eocene (Ypresian Age; around 55–54 Ma) and the early middle Eocene (Lutetian Age; around 49–47 Ma), whereas major falls in sea level took place at the early/middle Eocene (Ypresian/Lutetian) boundary (around 49 Ma) and at the middle/late Eocene (Bartonian/Priabonian) boundary (around 40 Ma).

A series of apparently correlative unconformities separating the synchronous alternation of shallow marine and brackish-water facies in the upper Paleocene–lower Eocene of the Paris-London-Hampshire and Belgian basins suggests a predominantly eustatic control on deposition in this passive margin setting. Comparable stratigraphic sequences have been delineated on the eastern and Gulf Coastal Plain of North America. *See* CONTINENTAL MARGIN; SEISMOLOGY; UNCONFORMITY.

Paleoslope modeling of the east coast of North America suggests that early and middle Eocene relative sea level stood 180–400 ft (55–120 m) higher than the present-day level and that the coastline would have shifted about 42 mi (70 km) during this change. Paleoslope modeling refers to the use of benthic forminiferans to reconstruct ancient environments on the sea floor along the slope of the continental shelf and slope. Different assemblages of benthic foraminiferans occupy different parts of the slope (with a certain degree of overlap), and each depth-related association is termed a biofacies. The faunal compositions of ancient biofacies are generally similar to those of the present day, so they can be used to approximate conditions in the past. On the assumption of a similar gradient of the slope in the past compared with today, the successional changes in benthic foraminiferal biofacies are used to reconstruct the ancient sea-floor environment.

The Eocene was a period of repeated transgressions and regressions of the sea; in particular during the middle Eocene, one of the most extensive global marine transgressions took place, leaving extensive shallow-water limestones with rich and diversified molluscan faunas in the Gulf Coast and the Paris Basin; the latter faunas became a focal point of

extensive studies in the nineteenth century and played a major role in the formulation of Lyell's concept of the Eocene.

Life. Early Paleogene temperatures, including those of high latitudes, were the warmest of the Cenozoic; peak warming occurred in the early Eocene. Equatorial surface waters may have been several degrees cooler than at present, and latitudinal thermal gradients were lower; this resulted in a poleward expansion of tropical life-forms. However, dramatic changes in climatic conditions appear to have taken place at the end of the Paleocene. A significant increase in the temperature of bottom waters of the ocean (about 5°F or 3°C) may have been coupled with a decrease in wind intensity, resulting in a decrease in surface productivity and lowered oxygen levels in bottom waters which may have triggered a major turnover in the deep-water benthic foraminiferans.

The diversification of life seen in the Paleocene continued in the Eocene, a reflection of the poleward expansion of the tropics, particularly during early Eocene time. In the oceanic realm, microplanktonic animals (foraminiferans) and plants (calcareous nannoplankton) flourished and diversified, as did true bony fishes and siphonate gastropods. In shallow, tropical waters, larger foraminiferans—*Nummulites* and discocyclinids—extended their geographic range to latitude 50° north, but the latter group disappeared at the end of the Eocene owing to cooling temperatures. Indeed, microplanktonic animals and plants experienced a gradual but inexorable decline in diversity starting in the late middle Eocene. Succeeding Oligocene faunas and floras were much reduced in diversity and much more uniformly distributed. *See* FORAMINIFERIDA; NUMMULITES.

On land, subtropical floras extended as far north as southern England and the North American Pacific coast of Puget Sound and southern Alaska. Indeed, the floras of southern England resembled those of modern-day China, Malaysia, and Australia. In the humid interior, thick and extensive mud deposits (the Green River Shale) in Colorado contain a beautifully preserved fresh-water fish fauna eagerly sought after by fossil collectors, as well as vast quantities of untapped oil comparable to that in the oil-rich tar sands of Alberta Province, Canada.

Europe was separated from the Eurasian landmass east of the Urals by a north-south seaway extending from the Arctic to the Tethys Sea—the Turgai Straits. Following the elimination of the elevated corridor that allowed transatlantic poleward migration between Europe and North America in late early Eocene time, middle and late Eocene time witnessed the development of extensive endemic animal evolution. Bats, flying lemurs, creodont carnivores, artiodactyls (cloven-hoof mammals, such as cattle, deer, and camels) and perissodactyls (odd-toed, hoofed mammals, such as rhinoceroses and horses), notoungulates (predominantly South American), and edentates reflect the diversification of primitive placental forms. The massive, rhinoceroslike herbivores called titanotheres and uintatheres appeared alongside the small early progenitors of the modern horse, *Hyracotherium* (known popularly as *Eohippus*). *See* MAMMALIA; PERISSODACTYLA.

In the Eocene, some mammals turned toward life in the sea; sea cows appeared in the middle Eocene, while the earliest whales (zeuglodonts) appeared in the North Atlantic–Gulf of Mexico region and the aquatic ancestors of the proboscideans appeared in the late Eocene.

The presence of warm southern ocean currents between Australia and Antarctica, which were in the process of separating, led to increased evaporation and precipitation on the great landmass of Antarctica and to gradual cooling, which, in turn, led inexorably to the initiation of continental ice on that continent in the succeeding epoch, the Oligocene. The Panamanian Isthmus was below sea level, and North and South America were separated from each other, resulting in the separate but parallel evolution of many mammals on the two continents. Marsupials, ancestral anteaters, armadillos, ground sloths, large rodents, and other now-extinct placentals evolved separately until the two landmasses were reunited by the elevation of the Panamanian Isthmus in the Pliocene some 3 Ma. *See* PALEOBOTANY; PALEONTOLOGY.

W. A. Berggren

Bibliography. M.-P. Aubry et al., Chronostratigraphic terminology at the Paleocene/Eocene boundary, in S. L. Wing (ed.) et al., *Causes and Consequences of Globally Warm Climates in the Early Paleogene* (Special Papers), Geological Society of America, vol. 369, pp. 551–566, 2003; M.-P. Aubry, S. G. Lucas, and W. A. Berggren (eds.), *Late Paleocene-Early Eocene Biotic and Climatic Events in the Marine and Terrestrial Records*, Columbia University Press, 1998; K. Ouda and M.-P. Aubry (eds.), *The Upper Paleocene-Lower Eocene of the Upper Nile Valley: Stratigraphy*, Micropaleontology Special Issue, vol. 49(1), 2003; D. R. Prothero, *The Eocene-Oligocene Transition*, Columbia University Press, 1993; D. R. Prothero, L. C. Ivany, and E. Nesbitt (eds.), *From Greenhouse to Icehouse*, Columbia University Press, 2002.

Eocrinoidea

A medium-sized class of primitive, brachiole-bearing, blastozoan echinoderms of the class Crinozoa that ranged from the Early Cambrian to the Middle Silurian, although few eocrinoids survived past the Middle Ordovician. About 32 eocrinoid genera have been described from North America, Europe, North Africa, and Australia; other occurrences of distinctive plates that may belong to eocrinoids have also been noted. Eocrinoids are the most diverse class of echinoderms known from the Cambrian with about 15 genera, and different members appear to have been ancestral to nearly all of the more advanced brachiole-bearing echinoderm classes that appeared in the Early or Middle Ordovician, such as rhombiferans, parablastoids, and coronoids. Eocrinoids have a globular, conical, or flattened theca or body, with

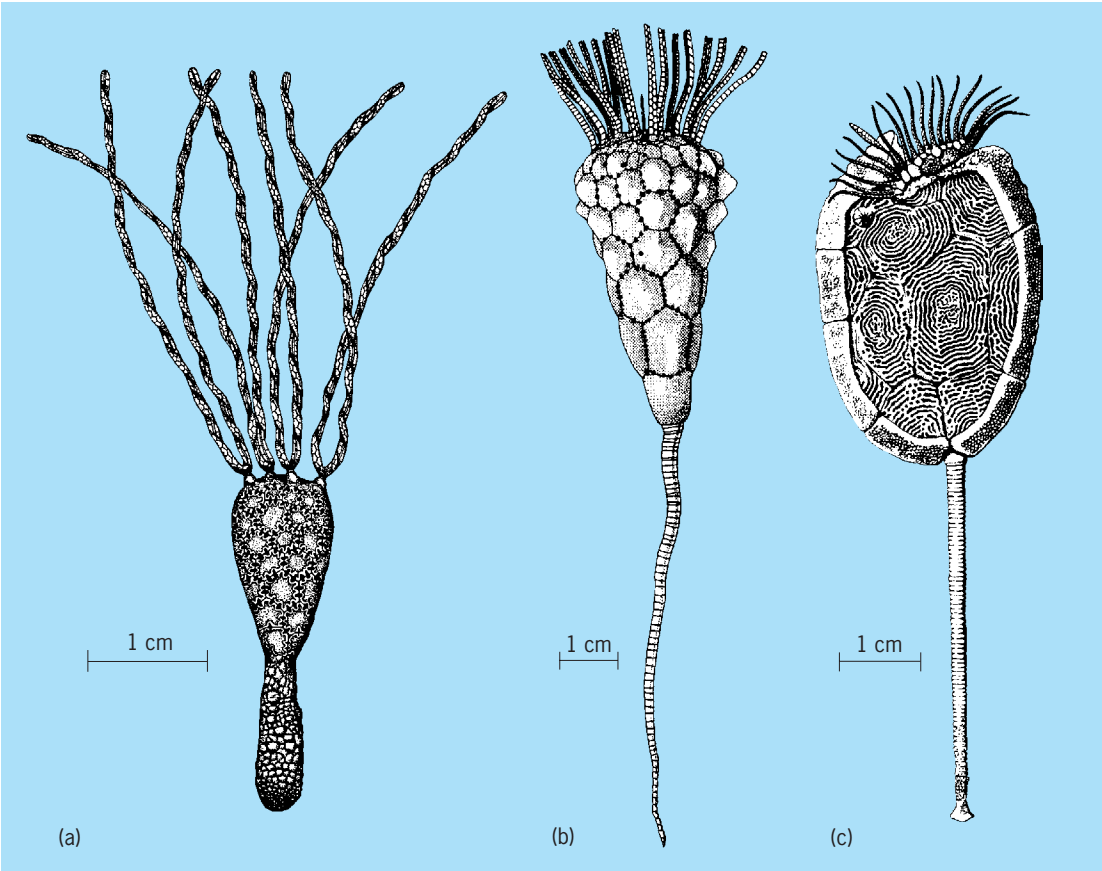

Reconstructions of Cambrian and Ordovician eocrinoids. (*a*) *Gogia spiralis*, from the Middle Cambrian of Utah, showing the primitive holdfast; sutural pores between the numerous thecal plates; and long, spiraled brachioles attached to the summit (*after R. A. Robison, Middle Cambrian eocrinoids from Western North America, J. Paleont., 39:360–363, 1965*). (*b*) *Rhopalocystis destombesi*, from the Early Ordovician of Morocco; note the more advanced columnal-bearing stem, a theca that still retains sutural pores, and incomplete brachioles on summit (*after G. Ubaghs, Eocrinoidea, in R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. S, 1968*). (*c*) *Mandalacystis dockeri*, from the Middle Ordovician of Oklahoma, showing the stem with its distal attachment, flattened theca with more massive plates around the margins, and short brachioles attached to the summit (*after R. D. Lewis et al., Mandalacystis, a new rhipidocystid eocrinoid from the Whiterockian Stage (Ordovician) in Oklahoma and Nevada, J. Paleontol., 61:1222–1235, 1987*).

many irregularly arranged to partly organized, imbricate or adjacent plates (see **illus.**). Most Cambrian genera have sutural pores on the plate margins, apparently for respiration. Many early eocrinoids have a multiplated, cylindrical to slightly inflated holdfast for attaching the theca to objects lying on the sea floor. Holdfasts apparently evolved into a true columnal-bearing stem in late Middle Cambrian eocrinoids, and most later genera have that advanced type of attachment structure. Eocrinoids typically have two to five short ambulacral grooves radiating from the mouth on the summit to many long, erect, biserial brachioles that were used for feeding. Most eocrinoids were attached, low- to medium-level suspension feeders that used the brachioles to collect small food particles drifting by the theca. Most researchers have argued that eocrinoids are a valid class containing genera that did not develop the fold-like respiratory structures found in more advanced blastozoan classes. However, others have recently proposed that eocrinoids are a paraphyletic stem group that should be discarded with the included genera reassigned to other classes. *See* ECHINODERMATA.

James Sprinkle

Bibliography. R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, pt. S, 1968; C. R. C. Paul and A. B. Smith, The early radiation and phylogeny of echinoderms, *Biol. Rev.*, 59:443–481, 1984; J. Sprinkle, *Morphology and Evolution of Blastozoan Echinoderms*, Mus. Compar. Zool., Harvard Univ., Spec. Publ., 1973.

Eolian landforms

Topographic features formed mainly by wind in regions with much loose sand, silt, or clay. An estimated 500 million metric tons (556 million tons) of dust per year are transported by the wind. Airflow in wind can be either laminar or turbulent, but is mostly turbulent. Wind blowing across the land surface contains many swirls and eddies, some of which produce upward components of movement. Generally, the higher the mean wind velocity, the greater the amount of upward turbulence. The ratio of the velocity of upward turbulent gusts near the ground to mean wind velocity varies, but it averages about 1:5. Thus, particles whose settling velocities are lower

than one-fifth of the mean wind velocity are likely to be conveyed upward by wind gusts and transported downwind. Larger particles remain mostly close to the ground until unusually high gusts lift them up. Small eddies occur under most atmospheric conditions, unless the air is unusually thermally stable. As wind velocity rises, corkscrew eddies develop with axes parallel to the mean wind direction; and larger, horizontal eddies, measured in hundreds of meters, are superimposed on the smaller, corkscrew eddies. *See* CLAY; DUST STORM; LAMINAR FLOW; SAND; VELOCITY; WIND.

Movement of particles by wind. Three modes of particle movement by the wind are suspension, saltation, and surface creep. Turbulent eddies are capable of transporting fine silt and clay particles upward and may suspend particles in the air. However, if the terminal velocity of sand grains is greater than the vertical component of the turbulent velocity, the particle trajectory path is a smooth curve with a trajectory height measured in centimeters, rather than in kilometers as is typical for fine particles in suspension. Sand-size particles traveling in low, smooth trajectories and bouncing along on the surface move by saltation. Particles that are too large to be lifted from the surface by the wind may be rolled or pushed along the surface by impact creep, produced by the impact of saltating particles. Of these, saltation is the most significant mode of transportation because fine sand grains are most easily moved by the wind. Smaller grains are more difficult to move because of particle cohesion and aerodynamic effects, and larger particles are more difficult to move because of their greater mass.

When a saltating grain returns to the surface, it may crash into a group of sand grains, driving them into the air. Saltation impact is a significant process in putting dust and other small grains into the air by winds that would otherwise not be strong enough to initiate grain movement. Saltating grains that strike surfaces composed of larger grains, too large to be moved by wind alone, push them downwind a short distance.

Sand grains saltating along a hard, rocky surface bound high into the air upon impact and are thus more easily kept in motion than particles hitting loose sand, where momentum is dissipated among the loose grains, often trapping the impacting grain. As golfers are aware, golf balls hit into a sand trap do not bounce nearly as much as balls landing on a green. Accumulation of sand grains in one place thus leads to the trapping of still more sand, rather than spreading out evenly over the ground surface. *See* MOMENTUM.

Coarse sand grains on the ground surface that are too large to be directly set in motion by the wind may be nudged forward by the impact of bombarding sand grains. The impact of a high-velocity, saltating grain can move a grain six times its diameter and more than 200 times its own weight. This type of downwind sand movement is known as impact creep or surface creep. Creep may account for about one-fourth of the total sand movement.

Wind is considerably limited in the size of particles it can transport. Not only does settling velocity determine the particle size of suspended grains, but also it is important in limiting the height to which saltating grains rise above the ground. In addition, the nature of the surface from which saltating grains rebound is important. The rebound height above a sand-mantled surface may be only 0.5 m (1.6 ft) or less, whereas above a rock surface it can be several meters. Abrasion by windblown sand rarely exceeds about 1 m (3.2 ft) above the ground. Wooden and steel telegraph posts in the Sahara severed about 1 m above the ground show the effective height of saltating grains there.

Wind is an unusually selective sorting agent. Eolian sediments are typically among the best-sorted deposits in nature. The most common grain sizes found by sieve analyses of dune sands are 0.30–0.15 mm in diameter. Diameters less than 0.08 mm are rare because most grains of this size are carried off in suspension.

Eolian abrasion and erosion. Wind erosion takes place by abrasion, the wearing away of material by frictional impact, and by deflation, the blowing away of loose material. Wind abrasion is caused by the collision of sand grains with surfaces. Spectacular examples occur when automobiles caught in sandstorms are stripped of paint and their windshields frosted in a matter of minutes. In nature, saltating sand grains, bouncing along the ground near the surface, collide with rocks or other sand grains that they abrade. Windblown sand blasts exposed rock faces and creates small-scale abrasion features whose details depend on the lithology, inclination, and orientation of the rock; the velocity and direction of the wind; and the size, hardness, and shape of the abrasive material. The results of wind abrasion include the polishing, pitting, grooving, fluting, and faceting of rocks and the rounding of sand grains in transport. *See* EROSION.

Among the most common features of wind abrasion are ventifacts, etched and pitted faces of exposed rock. If a rock face is normal to the prevailing wind direction, a faceted face is abraded at right angles to the wind, with sharp edges between the facet and the original surface. Rocks that periodically rotate so as to expose different sides to the direction of prevailing wind develop multiple-faceted sides. The abraded facets are typically pitted, fluted, grooved, and polished.

Larger-scale erosional effects of wind include yardangs, elongate ridges aligned parallel to the direction of prevailing wind with rounded windward faces that taper in the downwind direction. Yardangs vary in length from several meters to 1 km (0.6 mi), and from a few meters to 200 m (656 ft) high. Bedding and other structures in the material are truncated by the troughs, displaying their erosional origin. That they are wind-eroded is shown by their alignment parallel to the direction of prevailing wind and by their composition of easily eroded material. The elongate ridges are separated by intervening troughs, whose origin may also include erosion

by running water and weathering. *See* WEATHERING PROCESSES.

Where strong winds blow across surfaces lacking coarse material, removal of small debris by the wind forms deflation hollows. The size and shape of deflation hollows vary considerably, depending on the nature of the underlying material, strength of the wind, and local weathering conditions.

Sand dunes. A patch of sand makes an effective trap, which continues to grow in size until it becomes a mound of sand. Protrusion of the sand mound high enough into the air to disrupt laminar airflow over its surface significantly affects further accumulation and movement of sand, and a dune is formed. Once a sand grain becomes part of a dune, it advances downwind only at the rate of advance of the entire dune, rather than at its previous saltation rate. *See* DUNE.

When a critical height is reached, a slip face forms on the lee side of the dune. Sand moves up the windward side of the dune by saltation and is blown over the brink of the crest, increasing the slope until it exceeds the angle of repose for loose sand. An avalanche forms down the lee side, creating a slip face standing at the angle of repose of loose sand, about 30–34°. Successive slip faces become progressively buried until they are once again unearthed by erosion on the windward side of the dune as the dune gradually advances downwind. The most important factors are dune height and rate of sand flow. Measured rates of dune migration are 15–20 m/y (50–65 ft/y).

Barchans are crescent-shaped dunes with a steep slip face concave in the downwind direction. The average width of barchan dunes, measured from horn to horn, is about 37 m (121 ft), and the heights of the dunes are consistently about one-tenth the width. The horns of the crescent taper in the downwind direction because, even with constant lateral supply of sand, saltating grains will move more quickly over the hard ground adjacent to the dune patch and most slowly where the sand is thickest. Because the edges of the dune are lower and the rate of advance of the slip face is inversely proportional to the dune height, the edges move downwind at a faster rate than the thicker, central part of the dune, giving the dune its crescent shape. Barchans are typically best developed on barren desert floors where the supply of sand is meager, the prevailing wind is relatively uniform, and vegetation is scarce. They may occur as single dunes or in groups where individual barchans coalesce to form complex shapes.

Transverse dunes are elongate dunes that form perpendicular to the direction of prevailing winds. They originate in several ways. Some are clearly transitional forms with barchan dunes and consist of several barchans connected along single lines.

Parabolic dunes are crescent-shaped dunes in which the slip face of the dune is convex downwind and the horns of the crescent point upwind. Some are clearly transitional to barchan dunes and form as a result of anchoring of the horns by vegetation while the central portion continues to migrate downwind. Vegetation is an important factor in eolian processes because it is effective in breaking up airflow near the ground and serves to anchor thin sand deposits. The result is that the horns of a parabolic dune, anchored by vegetation, trail the thicker central part of the dune. In some cases, the horns of the parabolic dune may be left far behind, giving the dune a hairpinlike shape.

Longitudinal dunes are elongate ridges of sand parallel to the direction of prevailing winds. One of the places that they may form is where wind funnels sand through a gap in a ridge. As the wind blows through the gap, its velocity increases due to the Bernoulli effect; but when the wind passes through the gap, its velocity decreases to that of the general flow, and it deposits some of the sand carried by the enhanced velocity through the gap. *See* BERNOULLI'S THEOREM.

Seif dunes are longitudinal dunes that occur in long, straight, parallel chains over exceptionally long distances, sometimes 60–190 km (37–118 mi). Perhaps the most intriguing aspect of these unusual dunes is not only their length but also the regularity of their spacing. Most longitudinal dunes are separated by barren rock floors or desert pavements of rather uniform width. No definitive explanation has been demonstrated for these extraordinary dunes.

Star dunes typically consist of three or more sharp-edged ridges extending radially from a high, pointed, central peak. They seem to originate in areas where winds blow from more than one prevailing direction.

Loess. Loess consists of usually well sorted, wind-blown silt and clay deposited as a relatively homogeneous, unstratified blanket over the Earth's surface. Typical loess contains about 40–50% silt, up to 30% clay, and up to 10% fine sand. *See* LOESS.

An estimated 30% of the United States is mantled with loess, and as much as 10% of the Earth's land area is covered by loess. The sources of most loess deposits seem to be related to areas affected by Pleistocene glaciations and climatic changes, especially downwind from glacial outwash plains. Thick loess deposits occur in the Mississippi Valley region of the Central Plains; in Alaska; in the Columbia Plateau of eastern Washington; in portions of Europe south of the Scandinavian Ice Sheet; and in China, Mongolia, and the former Soviet Union. The largest known loess-covered region is 635,000 km^2 (245,110 mi^2) in north-central China. *See* CLIMATE HISTORY; GLACIAL EPOCH; PLEISTOCENE.

The loess deposits of China and the Central Plains of the United States are especially thick, reaching 300 m (984 ft). Elsewhere, loess thicknesses generally are several tens of meters. During the dust-bowl storms of the Central Plains in the 1930s, immense volumes of silt became airborne. In 1935 a dust storm near Wichita, Kansas, was estimated to hold about 5 million metric tons (5.6 million tons) of dust in suspension over a 78-km^2 (30-mi^2) area, and 300 metric tons/km^2 (333 tons/mi^2) was deposited in a single day near Lincoln, Nebraska. Deposition of loess in China presently occurs at a rate of several millimeters per year.

Don J. Easterbrook

Bibliography. R. Cooke, A. Warre, and A. S. Goudie, *Desert Geomorphology*, 1993; A. S. Goodie, S. Stokes, and I. Livingstone, *Aeolian Environments, Sediments and Landforms*, 2000; V. P. Tchakerian (ed.), *Desert Aeolian Processes*; 1995.

Ephedra

A genus of low, leafless, green-stemmed shrubs belonging to the plant class Ginkgoopsida of the division Pinophyta (see **illus.**). They grow in dry, alkaline soils around the world. In the southwestern United States these plants are called Mormon tea and jointfir. The drug ephedrine is extracted from the Asiatic species *Ephedra sinica* and *E. equisetina*. It is used medicinally in the treatment of colds, hay fever, and asthma. *See* EPHEDRALES; GINKGOOPSIDA; PINOPHYTA; PLANT KINGDOM.

Perry D. Strausbaugh; Earl L. Core

A species of *Ephedra*, growing in Utah. (*Courtesy of Tony Gauba, from National Audubon Society*)

Ephedrales

An order of the class Ginkgoopsida having about 35 species in the genus *Ephedra*. These plants are mostly of arid regions. They include freely branched, low shrubs with reduced, scalelike leaves and green photosynthetic twigs. The species are dioecious (seldom gynodioecious). The genus is known from Asia to Europe, in northern Africa, in the United States and Mexico, and from Bolivia to Patagonia. It is not known in the fossil record. *See* GINKGOOPSIDA; PINOPHYTA; PLANT KINGDOM.

Thomas A. Zanoni

Ephemeris

A source of data pertaining to features of various objects, usually astronomical, that depend on time. The ephemeris can be a printed or computer-readable table such as on a compact disk, or it can be a set of numerical data accessible by computer programs. Most ephemerides pertain to positions of celestial objects as seen from any place in the solar system. The term may also apply to other parameters, such as those describing the state of rotation or orientation of a body in space.

A basic ephemeris is usually derived from the known laws of physics that, in turn, are described by equations of motion. Before computers, these equations were solved and evaluated using elaborate analytical formulas requiring extensive manual effort. With computers, the process is done numerically. Initial ephemerides are compared with positional measurements (observations), the initial parameters are adjusted to produce better agreement (between ephemerides and observations), and the whole process is then iterated. The resultant ephemerides, stored in the computer, are accessible to produce tabular listings or to give other types of positional information at various requested times. *See* NUMERICAL ANALYSIS.

Traditional printed ephemerides are available in many national astronomical almanacs. On the Internet, there are interactive Web sites that enable customizing one's own ephemeris. Also available are computer subroutines for approximate accuracy that users can employ in their self-written programs. Finally, there are precision ephemerides for full accuracy, available in the form of computer data files and reading routines, to be used as a part of a user-written main computer program. Computer links to all of these sources are available on the Web site of the International Astronomical Union's Commission 4 (Ephemerides). *See* ALMANAC; CELESTIAL MECHANICS; INTERNET.

E. Myles Standish

Bibliography. G. Beutler, *Methods of Celestial Mechanics*, Springer, 2005; J. M. A. Danby, *Fundamentals of Celestial Mechanics*, Willmann-Bell, 1988; *Explanatory Supplement to the Astronomical Ephemeris and the American Ephemeris and Nautical Almanac*, H. M. Stationery Office, London, 1961; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, University Science Books, Mill Valley, CA, 1992.

Ephemeroptera

An order of insects commonly known as mayflies. They are aquatic and live in clean, fresh waters during their immature, or nymphal, lives. Nymphs are adapted to aquatic environments ranging from ponds to mountain streams. Most mayfly nymphs are vegetarians, and most species are present throughout the year as nymphs. The adult stage is very brief, lasting from a few hours to several days (see **illus.**). As

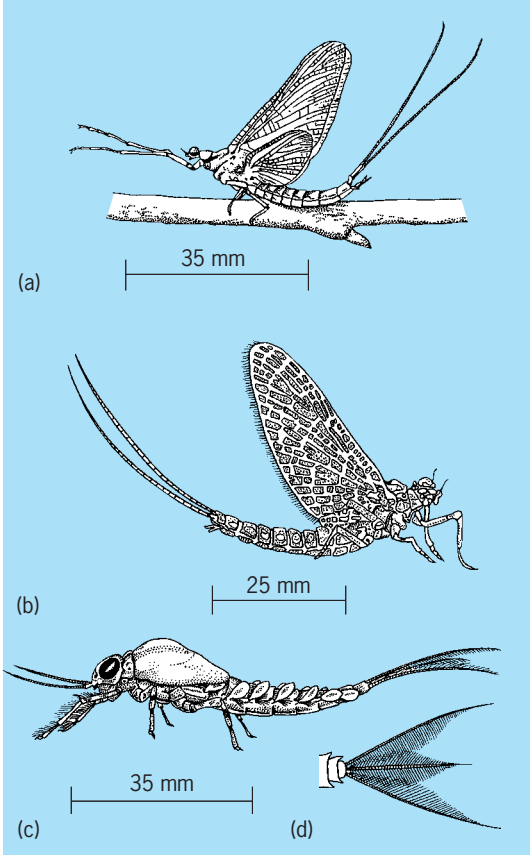

Ephemeroptera. (*a*) Mayfly, adult male, *Hexagenia* sp., in usual resting position (*after A. H. Morgan, Field Book of Ponds and Streams, G. Putnam's Sons, 1930*). (*b*) Mayfly, subadult male, *Callibaetis* sp., newly emerged from water, standing on surface film (*after A. H. Morgan, Kinships of Animals and Man, McGraw-Hill, 1955*). (*c*) Mayfly nymph, *Isonychia* sp., from rapid stream; (*d*) segmented tailpieces (*after A. H. Morgan, Field Book of Ponds and Streams, G. Putnam's Sons, 1930*).

immature insects, they constitute a year-round basic food supply for carnivores of their communities, especially for fishes, and trout in particular. Nymphs and adults are models for artificial bait flies.

The short-lived adult female mayflies drop their eggs in packets into the water or scatter them over its surface. Certain small species fix the eggs to surfaces of rocks partly submerged in rapid currents. The egg stage may last from a few days to several months and, after hatching, the nymph may mature rapidly (in a few weeks) or may grow slowly (requiring as much as 2 years to complete its development), according to the species. At the end of their growth period, most nymphs rise to the water surface, shed their skins, and become subadults, or subimagoes. Some nymphs may crawl from the water onto vegetation, rocks, or sticks prior to emergence. In a short time, varying from minutes to a day or two, they shed the subadult skin and become adults, or imagoes.

Mayfly nymphs are distinguished from all other aquatic insects by the paired tracheal gills on the back of each of the first seven abdominal segments. The body ends in two or three finely segmented tailpieces. The mouthparts are highly interesting examples of adaptations for scraping, cutting, and crushing the various plant cells, which are the chief food.

As subadults, mayflies have the general adult form and useless mouthparts. They are clothed by a thin, furry, grayish skin which accounts for the common name duns. The location of the subadult stage, however, between the nymphal and adult stages, suggests the sequence of stages in the metamorphosis of hemimetabolous insects such as larva, pupa, and adult of the butterfly. The subimago stage ends with the final molt of the mayfly's life. No other insects molt in their winged form.

The adult emerges with transparent wings, shining body, and useless mouthparts; the males generally have extremely large eyes, which may be divided, and long front legs. In a few mayflies, the legs of the female, and those of the male, except for the front ones, are nonfunctional and the brief adult life is spent entirely in flight. The quivering of the tailpieces during the mating flight gave them the name spinners. Thousands join in the twilight mating swarms, rhythmic dancing flights unequaled by other insects. Sometimes, the adults, which may be attracted by lights, accumulate in such masses that they become a hazard to vehicular traffic on highways and bridges over large rivers. *See* INSECTA.

Anne H. Morgan; Lewis Berner

Bibliography. L. Berner, *The Mayflies of Florida*, 1950; B. D. Burks, *The Mayflies, or Ephemeroptera, of Illinois*, 1953; J. W. Fannie and A. Leonard, *Mayflies of Michigan Trout Streams*, 1962; A. H. Morgan, *Field Book of Ponds and Streams*, 1930; J. G. Needham, J. R. Traver, and Y. C. Hsu, *Biology of Mayflies*, 1935; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Epicaridea

A suborder of the Isopoda which are parasitic on various crustaceans, mainly marine forms. The females are sometimes modified so strongly as to leave almost no indication of their isopod nature. The female pierces the host's skin with the aid of styliform mandibles to suck blood. It seizes the host's skin by means of its prehensile pereiopoda. The pleopoda are respiratory in function. The dwarf male attaches to the female, from which it takes nourishment, and retains its isopod structure. The first larva, the epicardium, is free-living, while the second, the microniscium, is temporarily ectoparasitic on Copepoda. The third larval stage, the cryptoniscium, is free-swimming and seeks a final host, on which it undergoes final development to the adult. The suborder is divided into two tribes, the Cryptoniscina and Bopyrina.

Cryptoniscina. These crustaceans live on Entomostraca and are protandrous hermaphrodites. The cryptoniscium larva attaches to a suitable female, develops testes without any change in its external appearance, and acts as a male. After the death of the host, it metamorphoses into a female, ultimately

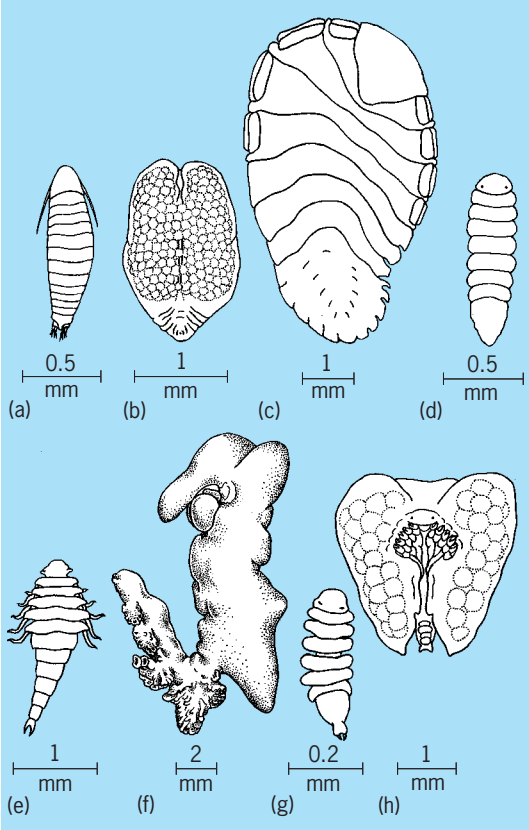

Epicaridea. (*a*) *Cyproniscus ovalis* male and (*b*) female. (*c*) *Bopyrus squillarum* female and (*d*) male. (*e*) *Portunion flavidus* male and (*f*) female. (*g*) *Prodajus bilobatus* male and (*h*) female.

undergoing anatomical degradation that leaves it a mere bulky sac distended with developing embryos. The brood chamber is internal and arises from paired invaginations of the sternum. *Hemioniscus* lives in the mantle cavity of Cirripedia, *Cyproniscus* (**illus.** *a* and *b*) lives in the shell of Ostracoda, and *Liriopsis* infests Rhizocephala, which is in turn a parasite of crabs. *Asconiscus*, *Podascon*, and *Cabirops* live in the brood pouch of Mysidacea, Amphipoda, and Isopoda, respectively.

Bopyrina. These isopods are dioecious, but it has been shown experimentally in some forms that the cryptoniscium is ambipotent, and a presumptive young male removed from the female develops into either a female or intersex. The tribe includes three families, rich in species. Bopyridae parasitize Decapoda. *Bopyrus* and its allies (illus. *c* and *d*) live in the branchial cavity of prawns, crabs, and hermit crabs; *Phryxus* lives under the abdomen of prawns; and *Athelges* lives above the abdomen of hermit crabs. The female shows an asymmetrical development of the segments, which are often partially fused, and it shows frequent disappearance of some of the pleopoda or one of their rami. Similar degeneration also occurs in the male, which is nevertheless symmetrical. The brood pouch is external, as in other Isopoda. Branchial parasites cause a swelling of the branchial cavity of the host. In some hosts the parasites produce an atrophy of the gonad and modification of secondary sexual characters. Entoniscidae, such as *Portunion* (illus. *e* and *f*), live in the visceral cavity of crabs and porcellanids. The female is vermiform, lacks metamerism, and is surrounded by a thin membrane derived from the host. Pereiopoda are atrophied, whereas pleopoda are excessively developed. Dajidae, for example, *Prodajus* (illus. *g* and *h*), live on Schizopoda, clinging to various parts of the host. The female is symmetrical, scarcely showing any metamerism. The fourth, fifth, and sixth pereiopods are reduced and pleophoda are abortive. *See* ISOPODA.

Sueo M. Shino

Bibliography. J. G. Baer, *Ecology of Animal Parasites*, 1951; M. Caullery, *Parasitism and Symbiosis*, 1952; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Epicycloid

A curve traced by a point on a circle that rolls on the convex side of a fixed circle. The term is also occasionally applied to the curve generated by a point on the prolongation of the radius of a circle as the circle rolls on a straight line. If the rolling circle has radius r and the fixed circle radius R, as in the **illustration**,

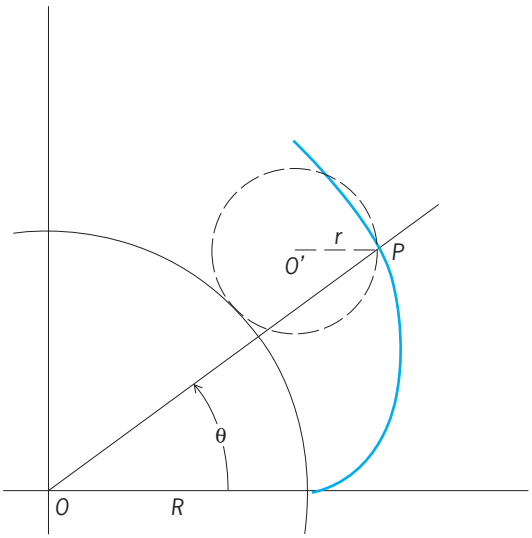

Diagram of an epicycloid.

parametric equations of the epicycloid are as shown below. The curve is closed provided R/r is a rational number. It is a cardioid when $r = R$. The curve was known to the Greek astronomer Hipparchus about 140 B.C., and was investigated by Gerard Desargues in 1639 and Leonhard Euler in 1781. Desargues made the first applications of epicycloids to the design of gear teeth. *See* CARDIOID.

$$x = (R + r)\cos\theta - r\cos(1 + R/r)\theta$$
$$y = (R + r)\sin\theta - r\sin(1 + R/r)\theta$$

Leonard M. Blumenthal

Epidemic

The occurrence of cases of disease in excess of what is usually expected for a given period of time. Epidemics are commonly thought to involve outbreaks of acute infectious disease, such as measles, polio, or streptococcal sore throat. More recently, other types of health-related events such as homicide, drownings, and even hysteria have been considered to occur as "epidemics."

Confusion sometimes arises because of overlap between the terms epidemic, outbreak, and cluster. Although they are closely related, epidemic may be used to suggest problems that are geographically widespread, while outbreak and cluster are reserved for problems that involve smaller numbers of people or are more sharply defined in terms of the area of occurrence. For example, an epidemic of influenza could involve an entire state or region, whereas an outbreak of gastroenteritis might be restricted to a nursing home, school, or day-care center. The term cluster may be used to refer to noncommunicable disease states.

In contrast to epidemics, endemic problems are distinguished by their consistently high levels over a long period of time. Lung cancer in males has been endemic in the United States, whereas the surge of lung cancer cases in women in the United States represents an epidemic problem that has resulted from increase in cigarette smoking among women in general. A pandemic is closely related to an epidemic, but it is a problem that has spread over a considerably larger geographic area; influenza pandemics are often global.

Origins of epidemics. Disease and epidemics occur as a result of the interaction of three factors, agent, host, and environment. Agents cause the disease, hosts are susceptible to it, and environmental conditions permit host exposure to the agent. An understanding of the interaction between agent, host, and environment is crucial for the selection of the best approach to prevent or control the continuing spread of an epidemic.

The importance of the agent-host-environment relationship is exemplified by the problem of epidemic influenza, which can result in high attack rates, or illness levels, and substantial mortality among certain population groups. The agent for epidemic influenza is the highly infectious influenza A or B virus. The host is someone who has never before been infected by the specific strain of influenza virus and lacks protective antibodies to it. If the environment is a closed one, such as a nursing home, the efficiency of virus transmission from a source, most likely another person with a case of influenza, to the susceptible host is heightened. *See* INFLUENZA.

For infectious diseases, epidemics can occur when large numbers of susceptible persons are exposed to infectious agents in settings or under circumstances that permit the spread of the agent. Spread of an infectious disease depends primarily on the chain of transmission of an agent: a source of the agent, a route of exit from the host, a suitable mode of transmission between the susceptible host and the source, and a route of entry into another susceptible host. Modes of spread may involve direct physical contact, such as touching or sexual intercourse, between the infected host and the new host, or airborne spread, such as coughing or sneezing. Indirect transmission takes place through vehicles such as contaminated water, food, or intravenous fluids; inanimate objects such as bedding, clothes, or surgical instruments; or a biological vector such as a mosquito or flea.

Responses to epidemics. The detection of epidemics in the United States is dependent on public health surveillance systems that monitor the incidence of different types of diseases. Surveillance for infectious disease epidemics or outbreaks relies heavily on reports of disease cases by physicians, hospitals, day-care centers, and other sources for health information. In the United States, state and local public health agencies bear primary responsibilty for investigating disease outbreaks, implementing control measures to halt their spread, and conducting long-term follow-up monitoring to detect recurrences. *See* PUBLIC HEALTH.

When potential outbreaks are detected, a public health agency may initiate an epidemic investigation. The investigation is a systematic procedure that combines aspects from the disciplines of epidemiology, clinical medicine, laboratory science, and communications, as well as common sense. The steps involved in an epidemic investigation help to determine who became ill, where the individuals were affected, and when they were affected. By examining those person, place, and time features, the investigator may be able to explain why the epidemic occurred and to identify appropriate control measures. However, because of the urgency associated with many epidemics, public health officials may need to take steps to control the problem while the investigation continues.

The primary purpose of an epidemic investigation is identification of the source or mode of disease transmission in order to recommend control measures. However, investigations of epidemics are conducted for other reasons. They include evaluation of intervention effectiveness, epidemiologic research in situ, and on-site training of public health workers and others engaged in epidemic control. In addition, such studies help fulfill the requirements of specific disease-control programs and provide more-immediate information in response to political or public concerns. *See* EPIDEMIOLOGY; INFECTIOUS DISEASE.

Richard A. Goodman

Bibliography. D. W. Fraser et al., Legionnaires' disease: Description of an epidemic of pneumonia, *N. Engl. J. Med.*, 297:1189–1197, 1977; W. W. Holland, R. Detels, and G. Knox (eds.), *Oxford Textbook of Public Health*, 3d ed., 1997; J. L. Kelsey et al., *Methods in Observational Epidemiology*, 2d ed., 1996; J. J. Sacks et al., A nurse-associated epidemic of cardiac arrests in an intensive care unit, *J. Amer. Med. Ass.*, 259:689–695, 1988; P. A. Schulte, R. L. Ehrenberg, and M. Singal, Investigation of

occupational cancer clusters: Theory and practice, *Amer. J. Public Health*, 77:52–56, 1987.

Epidemic viral gastroenteritis

A clinical syndrome characterized by acute infectious gastroenteritis with watery diarrhea, vomiting, malaise, and abdominal cramps with a relatively short incubation period (12–36 h) and duration (24–48 h). A viral etiology is suspected when bacterial and parasitic agents are not found. In the United States, no etiologic agent can be found in 70% of the outbreaks of gastroenteritis. Most of these may be due to viral agents, such as the Norwalk, Snow Mountain, and Hawaii agents, astroviruses, caliciviruses, adenoviruses, nongroup A rotaviruses, and paroviruses. Epidemics are common worldwide, and have occurred following the consumption of fecally contaminated raw shellfish, food, or water, although the virus may be spread by airborne droplets as well. Epidemics are most frequent in residential homes, camps, institutions, and cruise ships. Many individual cases of mild diarrhea may in fact occur in epidemics for which the source of the infection cannot be found. Epidemic viral gastroenteritis is distinct from rotavirus diarrhea, a seasonal disease in winter that is the most common cause of diarrhea in young children, and affects virtually all children in the first 4 years of life. *See* DIARRHEA; INFANT DIARRHEA.

Although many different agents may cause epidemic viral gastroenteritis, most of these viruses are difficult to detect even with the most sophisticated tests. Since the diarrhea is often mild and of short duration, attention should be given to rehydration therapy and prevention by identification of the source. Some epidemics have been noteworthy, such as those of group B rotavirus that occurred in China in 1982–1983, affecting more than 1 million people. Fatalities have been associated with severe dehydration and loss of fluids and electrolytes in the stool. Research may improve the ability to diagnose these agents, understand the epidemiology of the diseases they cause, and develop protective strategies. Some of these agents, such as astroviruses and caliciviruses, appear to affect young children, whereas other viruses such as Norwalk, Snow Mountain, Hawaii, and Taunton agents have been identified in older children and adults. The degree to which these agents induce an immune response is not known. *See* ANIMAL VIRUS. Roger Glass

Epidemiology

The study of the distribution of diseases in populations and of factors that influence the occurrence of disease. Epidemiology examines epidemic (excess) and endemic (always present) diseases; it is based on the observation that most disease does not occur randomly, but is related to environmental and personal characteristics that vary by place, time, and subgroup of the population. The epidemiologist attempts to determine who is prone to a particular disease; where risk of the disease is highest; when the disease is most likely to occur and its trends over time; what exposure its victims have in common; how much the risk is increased through exposure; and how many cases of the disease could be avoided by eliminating the exposure.

In the course of history, the epidemiologic approach has helped to explain the transmission of communicable diseases, such as cholera and measles, by discovering what exposures or host factors were shared by individuals who became sick. Modern epidemiologists have contributed to an understanding of factors that influence the risk of chronic diseases, particularly cardiovascular diseases and cancer, which account for most deaths in developed countries today. Epidemiology has established the causal association of cigarette smoking with heart disease; shown that acquired immune deficiency syndrome (AIDS) is associated with certain sexual practices; linked menopausal estrogen use to increased risk of endometrial cancer but to decreased risk of osteoporosis; and demonstrated the value of mammography in reducing breast cancer mortality. By identifying personal characteristics and environmental exposures that increase the risk of disease, epidemiologists provide crucial input to risk assessments and contribute to the formulation of public health policy.

Epidemiologic studies, based mainly on human subjects, have the advantage of producing results relevant to people, but the disadvantage of not always allowing perfect control of study conditions. For ethical and practical reasons, many questions cannot be addressed by experimental studies in humans and for which observational studies (or experimental studies using laboratory animals or biomedical models) must suffice. Still, there are circumstances in which experimental studies on human subjects are appropriate, for example, when a new drug or surgical procedure appears promising and the potential benefits outweigh known or suspected risks.

Descriptive studies. Descriptive epidemiologic studies provide information about the occurrence of disease in a population or its subgroups and trends in the frequency of disease over time. Data sources include death certificates, special disease registries, surveys, and population censuses; the most common measures of disease occurrence are (1) mortality (number of deaths yearly per 1000 of population at risk); (2) incidence (number of new cases yearly per 100,000 of population at risk); and (3) prevalence (number of existing cases at a given time per 100 of population at risk).

These measures can be calculated for all causes combined or for specific causes of disease and can pertain to an entire population or to specific subgroups (age or racial). When a rate for a specific subgroup is calculated, the number of people in the subgroup with the disease is the numerator and the population is the denominator. For example, to calculate the incidence rate for endometrial cancer in black women aged 50–54 in 1980, the numerator

should include the new cases of endometrial cancer diagnosed in black women in that age group in 1980 and the denominator should give the total number of black women in that age group in the population that year.

The risk of most diseases is related to age. For example, heart disease mortality rates increase markedly with age. Thus, a country with a young population may have low overall rates of heart disease compared to another country whose inhabitants are, on average, much older; still, the risk of heart disease in each specific age group could be similar in the two countries. To compare the risk of disease across populations, differences in the age structure of those populations must be taken into account; similarly, to compare risk in the same population over time, the changing age composition of that population must be taken into account.

The effects of age can be controlled in several ways. For example, with age adjustment by the direct method, the age-specific rates of the populations of interest are applied to the age distribution of a standard population, and the resulting numbers of expected deaths for each population are compared. By the indirect method, the age-specific rates of a standard population are applied to the number of people in each age group in the study population, and the resulting numbers of deaths are compared to the number actually observed. Similar procedures adjust rates for other factors that influence risk of disease, for example, sex or race.

Descriptive measures are useful for identifying populations and subgroups at high and low risk of disease and for monitoring time trends for specific diseases. They provide the leads for analytic studies designed to investigate factors responsible for such disease profiles.

Analytic studies. Analytic epidemiologic studies seek to identify specific factors that increase or decrease the risk of disease and to quantify the associated risk.

Observational studies. In observational studies, the researcher does not alter the behavior or exposure of the study subjects, but observes them to learn whether those exposed to different factors differ in disease rates. Alternatively, the researcher attempts to learn what factors distinguish people who have developed a particular disease from those who have not.

The observational analytic studies most often performed are the cohort study and the case-control study. In the former (also called prospective study), a disease-free group of people, often characterized by a common feature such as occupation, is identified; data on exposures of interest are collected; and the people are observed over time to see if they develop the disease. Occurrence can be compared between the entire cohort (factory workers exposed to certain chemicals) and an external group (the general population) or between subgroups in the study cohort who differ by exposure (workers in the same factory, some exposed and others not exposed to certain chemicals). The relative risk of disease associated with exposure is determined by dividing rates of disease in the exposed group by rates in the nonexposed comparison group. Cohort studies are ideal in the respect that data gathered on exposure predate development of disease. However, such studies require large numbers of people and long follow-up periods to ensure reliable estimates of disease occurrence. Cohort studies are typically expensive and administratively complex.

In case-control studies, people who have already developed the disease of interest (cases) and people who are free of the disease (controls) are studied. Data are collected from individuals in both groups on personal characteristics or previous exposure to factors that are suspected to be determinants of the disease, and the two groups are compared. Although the relative risk of disease associated with exposure cannot be calculated directly in case-control studies, it can be closely estimated by a quantity known as the odds ratio, that is, the odds of exposure in the cases divided by the odds of exposure in the controls. Because it is not necessary to observe study subjects over time for the occurrence of disease, case-control studies take less time than cohort studies and require fewer subjects to provide statistically meaningful results; they are therefore generally less expensive to conduct. However, it is not always possible to be certain that the factor suspected of causing the disease preceded the development of the disease; the time sequence is important for establishing causality. Furthermore, the accuracy of exposure information often depends on the recall of past events (dietary practices in early life, dosages of past medications, or occupational exposures to particular chemicals); and this recall may be poor.

Experimental studies. In experimental studies, the investigator alters the behavior, exposure, or treatment of people to determine the impact of the intervention on the disease. Usually two groups are studied, one that experiences the intervention (the experimental group) and one that does not (the control group). Ideally, subjects are randomly assigned to the groups and, where possible, those responsible for assessing outcomes are unaware of (blind to) the assignments. Intervention studies include risk-reduction trials of attempts to alter risk factors (such as cessation of smoking by middle-aged men at risk of heart disease); screening trials of the effectiveness of procedures that detect early subclinical cases (such as the Papanicolaou test to screen for cervical cancer); or clinical trials of the effectiveness of specific treatments (such as studies comparing survival after coronary bypass surgery to survival on medication alone).

Outcome measures include incidence, mortality, and survival rates in both the intervention and control groups. Some studies extend only to rates of compliance (the percentage of individuals assigned to the smoking cessation program who attend the program compared to those not assigned who may seek help on their own) or to risk-factor changes (the percentage of smokers assigned to the smoking cessation program who quit smoking compared to

those not assigned) rather than to measuring disease occurrence per se (heart disease rates in those assigned to the smoking cessation program compared to those not assigned).

Association versus causation. An association of a factor with a disease does not necessarily mean that the two are causally related. For example, people who carry matches may have higher rates of cancer, but matches do not cause the disease. Criteria for establishing causality between a factor and a disease include the strength of the observed relationship, supported by any biological evidence, consistency with results of other epidemiologic studies, and the correct time sequence between exposure to the factor and development of the disease.

Observed associations may be chance findings. Biostatistical methods, which are essential in modern epidemiology, are used to test the possibility that observed results are due to chance and consider the effects of other factors on the associations that are found. The association of a factor with a disease may arise because both are related to another factor (known as a confounding factor). For example, smoking is independently related to both alcohol consumption and oral cancer; thus, individuals who drink alcohol are at increased risk of oral cancer in part because of their smoking. The relative risk of oral cancer attributed to alcohol consumption would be artificially high if smoking (a confounding factor) were not taken into account. Design and analysis techniques deal with the effects of potential confounding factors. *See* DISEASE; EPIDEMIC.

Virginia L. Ernster

Bibliography. G. D. Friedman, *Primer of Epidemiology*, 4th ed., 1994; C. H. Hennekens and J. E. Buring, *Epidemiology in Medicine*, 1987; S. B. Hulley, *Designing Clinical Research*, 1988; J. L. Kelsey, W. D. Thompson, and A. S. Evans, *Methods in Observational Epidemiology*, 2d ed., 1996.

Epidermal ridges

Minute corrugations of the skin. They compose a sculpturing, termed dermatoglyphics, which characterizes the palmar and plantar surfaces of humans and the nonhuman primates. These areas lack hair and sebaceous (oil) glands, but sweat glands are numerous. In certain kinds of monkeys, a portion of the undersurface of the tail bears similarly specialized skin.

Epidermal ridges do not course in a uniform direction, like the ribs of corduroy, but are arranged in configurations of variable types, as exemplified in the familiar arches, loops, and whorls of the fingers. Ridges are dissimilar to the ribs of corduroy also in that they present frequent interruptions, branchings, and other irregularities. All these details are unchanged in lifetime except as they may be destroyed or damaged by accident or disease. Even a small area of skin has a unique individuality of these detailed features, which provide a basis for positive personal identification.

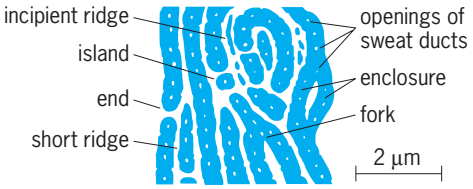

Characteristics of individual ridges and their configurations, seen in a print. (*After H. Cummins and C. Midlo, Fingerprints: Palms and Soles, McGraw-Hill, 1943*)

The ridges over the human hand as a whole, in young adult males, average 0.019 in. (0.48 mm) in breadth. They are slightly narrower in females, 0.017 in. (0.43 mm).

Ridges serve two functions: (1) They increase security of contact with objects, in the manner of the milling of a tool handle. Ducts of sweat glands open on the summits of ridges, and moistening of the skin augments the security of contact. (2) They enhance the sense of touch. In passing the fingers or palm over an object for judging its texture, the slight displacement of ridges heightens stimulation of the underlying nerve endings.

The characteristics of individual ridges and their collective configurations (see **illus.**) are not ordinarily studied by direct inspection of the skin. Instead, they are examined in prints, usually impressions of the inked surface, made for the purpose of record, or natural imprints made by chance contact. *See* FINGERPRINT; INTEGUMENTARY PATTERNS.

Harold Cummins

Epidermis (plant)

The outermost layer (occasionally several layers) of cells on the primary plant body. Its origin and structure are variable. In addition to a general description of the epidermis of plants, this article singles out five structural components of the tissue: (1) cuticle; (2) stomatal apparatus (including guard cells and subsidiary cells); (3) bulliform (motor) cells; (4) trichomes; and (5) root hairs.

In most roots the epidermis has a common origin with other tissues and therefore becomes established a short distance back of the apex. In shoots that have stratified meristems, the outermost layer (first layer of the tunica) usually gives rise to the epidermis. In many grasses, however, this layer divides and contributes to internal tissues of the leaf primordia before the epidermis is organized. In plants that have unstratified shoot meristems, the embryonic epidermis (protoderm) may originate on the flanks of the apical cone or on organ primordia. *See* APICAL MERISTEM.

Leaves, herbaceous stems, and floral organs usually retain the epidermis through life. Most woody stems retain it for one to many years, after which it is replaced. In roots it is usually short-lived. *See* PERIDERM.

Ordinarily the epidermis is compact and devoid of intercellular spaces, aside from the stomata. In surface view the individual cells commonly appear

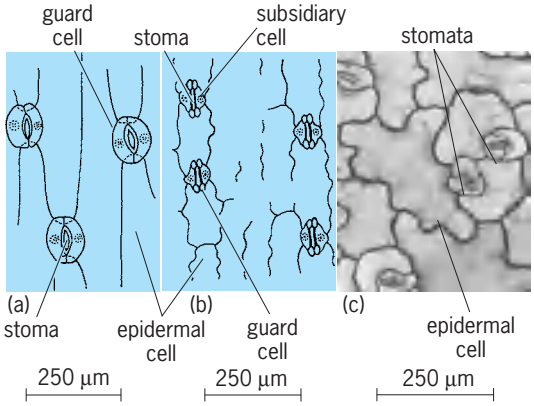

Fig. 1. Plant epidermis in surface view. (*a*) *Iris* leaf showing kidney-shaped guard cells. (*b*) *Zea* leaf showing dumbbell-shaped guard cells; two subsidiary cells are associated with each stoma. (*c*) *Sedum* leaf; each stoma is associated with three or four subsidiary cells.

polygonal or have wavy outlines (**Fig. 1**). Those on stems and linear leaves may be conspicuously elongated parallel to the axis of the organ. Specialized epidermal cells include the silica and cork cells in the leaves of some grasses, the lithocysts in *Ficus* (**Fig. 2***a*) and *Cannabis,* and fiberlike cells of *Raphia* and *Stylidium*. In the dry bulb scales of garlic the epidermis consists partly or entirely of heavily thickened sclereids. In the seed coats of many legumes it consists of a compact layer of macrosclereids arranged in columnar fashion. The epidermis of certain Polypodiaceae is photosynthetic. There are also plants in which the entire leaf, excluding the veins, consists of a double (*Elodea*) or even a single (*Hymenophyllum*) layer of epidermis. *See* SCLERENCHYMA.

Differential thickening of the walls characterizes most epidermal cells. In the more delicate types of epidermis, the outer tangential wall is commonly the thickest (Fig. 2*b*). In the leaves of many conifers, however, the walls of epidermal cells are so massively thickened that the cell cavities are almost obliterated (Fig. 2*c*).

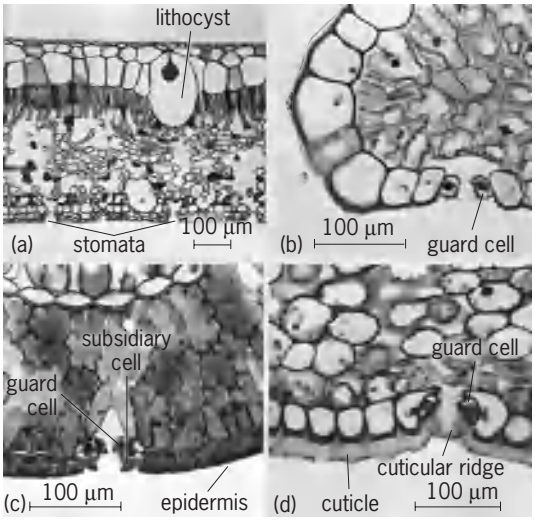

Fig. 2. Specialized epidermal cells in transections of leaves. (*a*) Leaf of *Ficus elastica*, each side having approximately three layers of cells in the multiple epidermis; lithocyst containing cystolith whose calcium deposits were removed during preservation. (*b*) Margin of leaf of *Lilium* with relatively thin-walled epidermis. (*c*) Leaf of *Pinus* with thick-walled epidermis and sunken guard cells. (*d*) Leaf of *Clivia* with heavy cuticle.

Multiple epidermis. This tissue is several layers of cells in depth. In its formation the protoderm undergoes periclinal (parallel with the circumference) divisions. A multiple epidermis occurs in many species of *Ficus, Begonia,* and *Peperomia*. The outermost layer of a multiple epidermis resembles a uniseriate (single series) epidermis, whereas the other layers constitute a water-storage tissue. In some peperomias, the water-storage tissue is thicker than the rest of the leaf. The velamen (parchmentlike covering) of orchid roots is a multiple epidermis.

Cutin, cuticle, and waxes. Cutin is a mixture of fatty substances characteristically found in epidermal cells. It impregnates the outer cell walls and occurs as a continuous layer (cuticle) on the outer surface (Fig. 2*d*). The cuticle covers the surfaces of young stems, leaves, floral organs, and even apical meristems. The cuticle can be stripped from some organs in relatively large sheets, a feature indicating its continuity. Cutin has been detected in the walls of cells adjacent to the intercellular spaces within leaves. There are some reports that it occurs in the epidermis of young roots.

Waxes appear as a deposit on the outside of the cuticle in many plants; the bloom on purple grapes and plums is an example. Most often the waxes are present in small quantity, but the leaves of some plants may be almost white with wax (*Echeveria subrigida*). The waxes of a few species are of great commercial value in the manufacture of polishes for floors, furniture, automobiles, and shoes. Carnauba wax is obtained from the leaves of a palm (*Copernicia cerifera*); candelilla wax, from the pencillike stems of *Euphorbia antisyphilitica*. Other substances, such as gums, resins, and salts, usually in crystalline form, may be deposited on the outside of the cuticle.

Stomata. The apertures in the epidermis which are surrounded by two specialized cells, the guard cells, are known as stomata. The singular form, stoma, is derived from the Greek word for mouth. However, some authorities prefer to include both aperture and guard cells within the concept of stoma.

Stomata occur most commonly on the leaves and young stems of vascular plants, with the exception of certain submerged water plants (*Elodea*). They are sometimes found on sporangia (*Botrychium*) and on the sporophytes (spore-producing generation) of certain nonvascular plants (mosses, *Anthoceros*). The apertures of stomata are contiguous with the intercellular space system of underlying tissues and thus permit gas exchange between internal cells and the external environment. *See* PLANT PHYSIOLOGY.

The number of stomata per square centimeter of epidermis varies from none to more than 100,000. In many plants stomata are restricted largely to the lower surface of the leaf. Occasionally they occur

only on the upper surface (*Castalia, Ammophila*). They may also occur in equal numbers on both sides of the leaf (*Abronia, Atriplex*). In oleander, stomata are restricted to special crypts on the lower leaf surface. In many dicotyledons stomata are arranged at random; in many monocotyledons and in the leaves of gymnosperms they are arranged with their long axes parallel to that of the leaf. The guard cells are frequently associated with subsidiary cells, which differ from other epidermal cells in size and form.

The guard cells of many plants are at about the same level as the other epidermal cells; in others they project above them. Xerophytes, plants which can subsist with a small amount of moisture, often have a thick-walled epidermis and the guard cells may be deeply sunken. In the majority of plants guard cells are kidney-shaped in surface view; in grasses, however, they resemble a pair of dumbbells. The guard cells undergo changes in form during the changes in turgor (water pressure), a phenomenon associated with the characteristic uneven thickening of the guard cell walls.

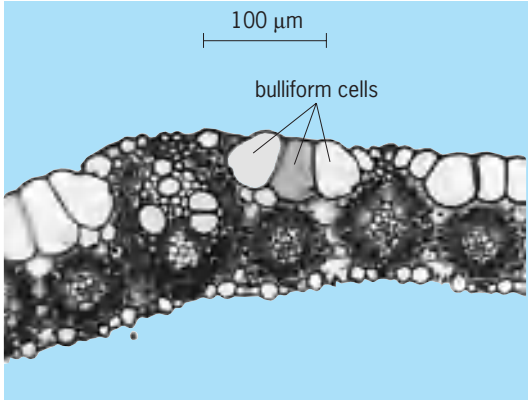

Fig. 3. Transection of leaf of *Andropogon* showing large bulliform cells in the upper epidermis.

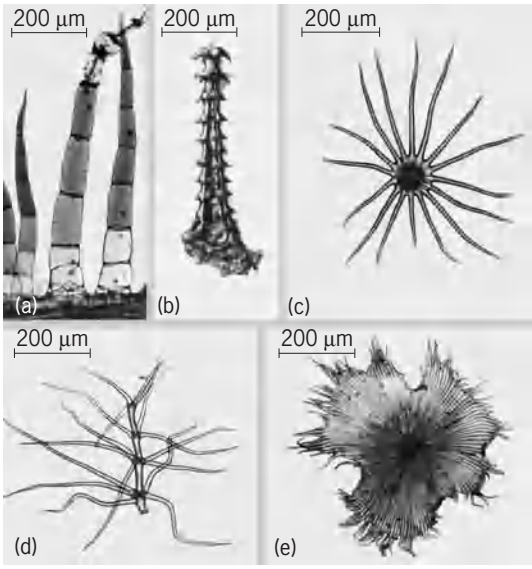

Fig. 4. Trichomes. (*a*) Uniseriate hairs of *Gynura*. (*b*) Unicellular anchor hair of *Mentzelia*. (*c*) Stellate hair of *Solanum*. (*d*) Branched (candelabra) hair of *Platanus*. (*e*) Peltate scale of *Shepherdia*.

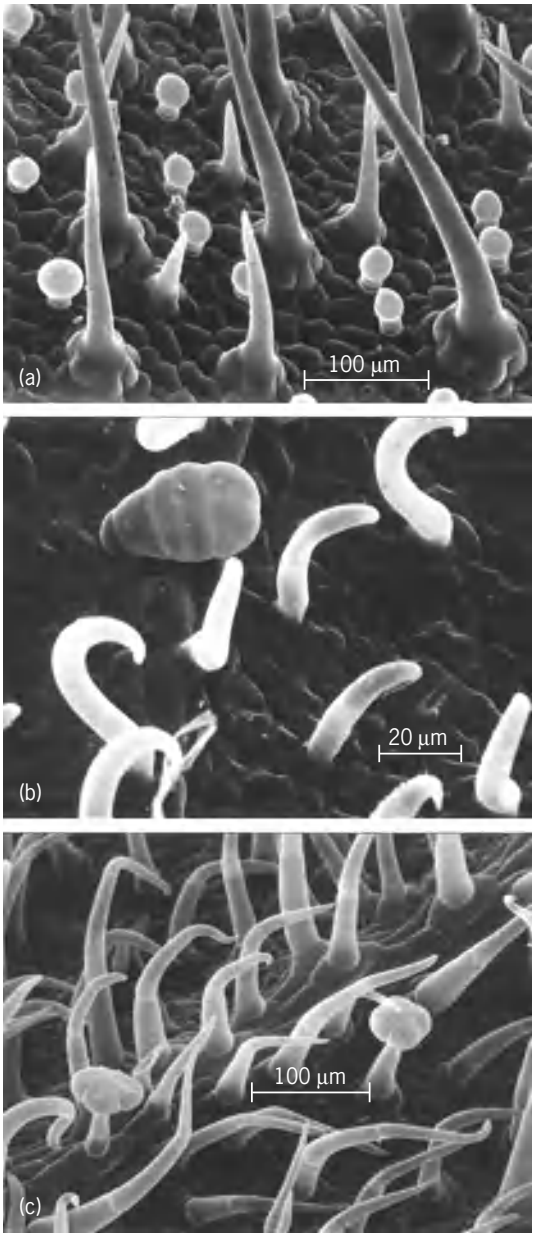

Fig. 5. Trichomes. (*a*) Unicellular and glandular (colleters) hairs of the geranium (*Pelargonium*). (*b*) Unicellular-hooked and uniseriate, club-shaped hairs of the bean (*Phaseolus*). (*c*) Uniseriate and glandular (colleters) hairs of the tomato (*Lycopersicon*).

The opening and closing of the stomatal aperture is caused by relative changes in turgor between the guard cells and surrounding epidermal cells. It is accompanied by hydrolysis, or condensation, of starch within the guard cells. *See* PLANT-WATER RELATIONS.

Bulliform (motor) cells. These large, highly vacuolated cells occur on the leaves of many monocotyledons but are probably best known in grasses (**Fig. 3**). They are thought to play a role in the unfolding of developing leaves and in the rolling and unrolling of mature leaves in response to alternating wet and dry periods. Not all investigators, however, subscribe to

this idea. The water vesicles in the epidermis of the ice plant (*Mesembryanthemum crystallinum*) are similar in appearance to bulliform cells.

Trichomes. Appendages derived from the protoderm are known as trichomes; the simplest are protrusions from single epidermal cells. Included in the concept, however, are such diverse structures as uniseriate hairs (**Fig. 4***a*), multiseriate hairs (*Begonia, Saxifraga*), anchor hairs (Fig. 4*b*), stellate hairs (Fig. 4*c*), branched (candelabra) hairs (Fig. 4*d*), peltate scales (Fig. 4*e*), stinging hairs, and glandular hairs (**Fig. 5**). Many of these structures arise on developing plant organs and are most conspicuous at this stage. They may be persistent or short-lived. Persistent trichomes are responsible for the silver-gray color of sagebrush (*Artemisia*), Spanish moss (*Tillandsia*), and Russian olive (*Elaeagnus*). Mature trichomes may or may not retain their protoplasts. The assignment of functions to trichomes is often uncertain, although they are useful in plant identification. Cotton and kapok fibers are unicellular epidermal hairs. Cotton fibers grow from the seed coat, those of kapok from the pod epidermis.

Root hairs. These are thin-walled extensions of certain root epidermal cells. They develop only on growing root tips and may arise from any epidermal cell, or from specialized cells known as trichoblasts. Root hairs increase the absorbing area of the root tip manyfold. The life of a given root hair is usually numbered in days. Persistent root hairs, reported for some plants, probably play no role in absorption. *See* PHLOEM; PLANT ORGANS; PLANT TISSUE SYSTEMS; ROOT (BOTANY); SECRETORY STRUCTURES (PLANT); XYLEM.

Norman H. Boke

Epidote

The group name for a family of minerals of general composition $Ca_2(Fe^{3+}, Al, Mn^{3+})Al_2O[SiO_4][Si_2O](OH)$ which occur widely in metamorphic and igneous rocks (**Fig. 1**). Epidote [octahedral ferric iron (Fe^{3+}) dominant] and clinozoisite [aluminum (Al) dominant] represent the most common compositions among the epidote group; a third composition, piemontite [manganese (Mn^{3+}) dominant], is less abundant. Allanite refers to compositions displaying significant rare-earth (such as lanthanum or cerium) substitution for calcium (Ca^{2+}), with corresponding replacement of Fe^{3+} by ferrous iron (Fe^{2+}). A fifth member, zoisite, is equivalent to clinozoisite, but it has a different crystalline system. Rare epidote-clinozoisites abundant in chromium (Cr), vanadium (V), and lead (Pb) and allanites rich in fluorine (F), beryllium (Be), and phosphorus (P) also exist. *See* CRYSTAL STRUCTURE; SOLID SOLUTION.

Structure. The epidote structure (**Fig. 2**) consists of chains of octahedra (M sites) elongated parallel to the *y* crystallographic axis and cross-linked by the silicate groups SiO_4 and Si_2O_7. There are three distinct kinds of octahedra (M1, M2, and M3) arrayed along two types of chains. One chain type consists entirely of edge-sharing octahedra (M2), whereas the other contains M1 octahedra alternating with M3 octahedra along its length. The resulting framework structure contains large cavities (A sites) where the Ca^{2+} cations are housed in 9–10-fold coordination. Like the chains of octahedra, epidote crystals themselves are also elongated parallel to the *y* axis. Such a relationship between internal crystal structure and external crystal form is a common theme among minerals in general. Epidotes exhibit cleavage on the *yz*(100) and *xy*(001) planes, and their hardnesses on the Mohs scale vary between 6 and 7.

In typical epidote-clinozoisites, Fe^{3+} and Mn^{3+} substitute for Al principally in M3 sites; M2 sites house only A1. M1 sites are usually dominated by A1 but

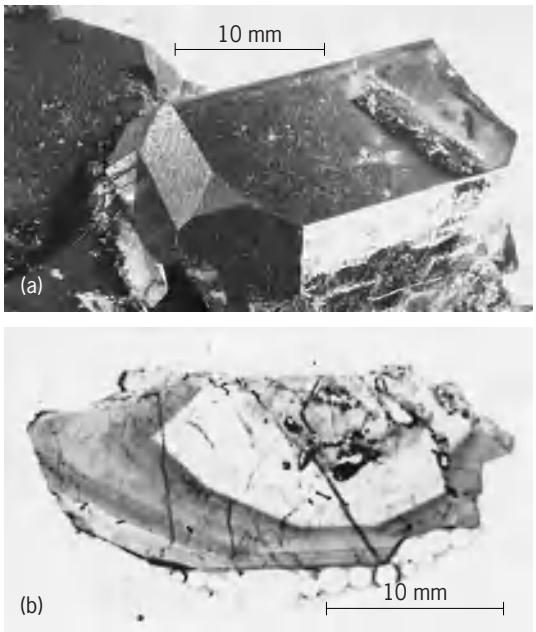

Fig. 1. Epidote. (*a*) Crystals from Prince of Wales Island, Alaska (*Yale University*). (*b*) Oscillatory zoning of these crystals.

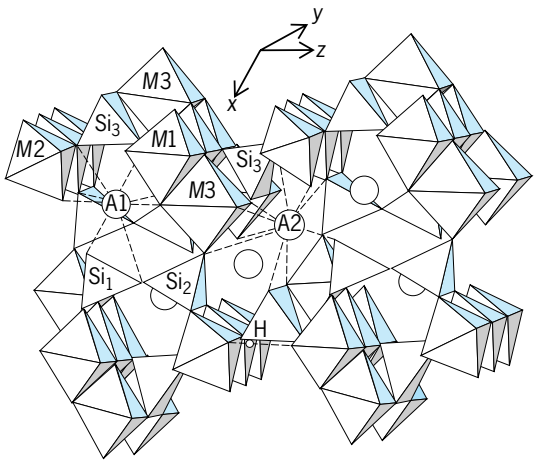

Fig. 2. Crystal structure of the epidote-group mineral clinozoisite. Si_3 represents the silicate group SiO_4; Si_1 and Si_2 represent the silicate group Si_2O_7. M1, M2, and M3 are distinct kinds of octahedra. A1 and A2 are cavities in the framework structure. (*After W. A. Deer, R. A. Howie, and J. Zussman, An Introduction to the Rock-Forming Minerals, Halsted Press, 1992*)

may contain Mn^{3+}. The explanation for the preference of these cations for the M3 site is twofold. First, both cations are larger than Al^{3+}, and the M3 octahedron is the largest in the structure. Second, the M3 sites are relatively distorted, and this characteristic may permit Mn^{3+} to gain additional stabilization energy here, over the other M sites.

Occurrence. Epidote group minerals, particularly epidote and clinozoisite, are common and widespread in regional- and contact-metamorphic rocks, both as primary and secondary (that is, alteration) minerals. They occur together as individual grains, as intergrowths, or as zoned crystals (Fig. 1*a*). Epidote and (clino)zoisite are found in aluminous limestones with grossularite, anorthite, microcline, quartz, and calcite; in mafic schists and gneisses with hornblende, albite, and chloritoid; in actinolite greenschists with chlorite, sphene, albite, quartz, calcite, and magnetite; in hornfels with diopside, actinolite, grossularite, and albite; in glaucophane schists; in quartzites; and in slates. Approximate depth-temperature conditions of their formation range from 5–25 km (3–15 mi) and 300–500°C (570–1020°F; low-grade) to 5–25 km (3–15 mi) and 450–650°C (840–1200°F; medium-grade). However, their stabilities are also sensitive to the pressure of oxygen in the rock during metamorphism. A zoned crystal will grow in a local environment of fluctuating f_{O_2} during metamorphism. Clinozoisite forms during periods of low f_{O_2}, whereas epidote grows during oscillations to high oxygen pressure. The cause of the oscillations in oxygen pressure is probably diffusion-controlled crystal growth, in which the growth rate exceeds that of oxygen diffusion to the crystallization site. *See* METAMORPHIC ROCKS.

In igneous rocks, epidote-(clino)zoisite occurs primarily in granitoids, as hydrothermal veins, as an alteration of plagioclase, or as a primary phase; in basalts, in cavities associated with augite, actinolite, grossularite, and sphene; and in mafic igneous rocks, as an alteration of plagioclase (that is, saussurite). *See* IGNEOUS ROCKS.

Piemontite and allanite are relatively restricted in occurrence. Piemontite is found chiefly in manganese ore deposits, altered volcanic rocks, and some schists. Allanite is an accessory mineral in granitoids, pegmatites, their volcanic equivalent rocks, and mafic schists and gneisses.

Epidote-group minerals occur as excellent crystals in several localities, notably near Salzburg, Austria; Isère, France; Piemonte, Italy; Riverside, California; and Prince of Wales Island, Alaska. *See* SILICATE MINERALS.

Peter S. Dahl

Bibliography. P. S. Dahl and L. M. Friberg, The occurrence and chemistry of epidote-clinozoisites in mafic gneisses from the Ruby Range, southwestern Montana, *Contrib. Geol.*, 18:77–82, 1980; W. A. Deer, R. A. Howie, and J. Zussman, *An Introduction to the Rock-Forming Minerals*, 2d ed., 1992; W. A. Dollase, Refinement of the crystal structures of epidote, allanite, and hancockite: *Amer. Mineral.*, 56:447–464, 1971; E. W. Heinrich, *Microscopic Identification of Minerals*, 1965; K. Langer and M. Raith, Infrared spectra of Al-Fe(III)-epidotes and zoisites, $Ca_2(Al_{1-p}Fe^{3+}{}_p)Al_2O(OH)[Si_2O_7][SiO_4]$, *Amer. Mineral.*, 59:1249–1258, 1974; C. Klein and C. S. Hurlbut, Jr., *Manual of Mineralogy*, 21st ed., 1993; G. H. Myer, New data on zoisite and epidote, *Amer. J. Sci.*, 264:364–385, 1966; E-an Zen and J. M. Hammarstrom, Magmatic epidote and its petrologic significance, *Geology*, 12:515–518, 1984.

Epinephrine

A hormone that is an important secretion from the adrenal medulla; also known as adrenaline. In 1901, J. Takamine isolated epinephrine in pure form from extracts of adrenal glands. Its structure (shown below) was finally proved by synthesis by F. Stolz in 1904 to be the catecholamine L-(3,4-dihydroxyphenyl)-2-methyl-aminoethanol. *See* ADRENAL MEDULLA; ADRENAL MEDULLA HORMONE; HORMONE.

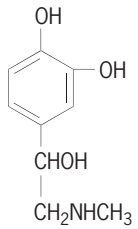

In mammals, the necessary components for the manufacture of the hormone by the body are the amino acids tyrosine or phenylalanine. Tyrosine is converted to dihydroxyphenylalanine (dopa), which is subsequently converted to dopamine, which in turn is modified to form norepinephrine. Norepinephrine is the immediate precursor for epinephrine. This final conversion requires the enzyme phenylethanolamine-*N*-methyltransferase (PNMT). PNMT is present in significant amounts only in the adrenal medulla and in some areas of the brain. Its expression is highly dependent on a supply of glucocorticoids, such as cortisol. Blood flowing from the adrenal cortex drains into the medulla, providing these cells with high glucocorticoid concentrations, sufficient to maintain PNMT production. *See* AMINO ACIDS; DOPAMINE; NORADRENERGIC SYSTEM.

Epinephrine and norepinephrine are stored in adrenal medullary cells (known as chromaffin cells) in granules bound to proteins called chromogranins. The granules also contain adenosine triphosphate (ATP). In response to release of acetylcholine from nerves innervating the medulla, the granules fuse with the chromaffin cell membrane, and their contents, including epinephrine, are released by exocytosis and quickly enter the blood.

Epinephrine is a sympathomimetic substance; that is, it acts on tissue supplied by sympathetic nerves, and generally the effects of its action are the same as those of other nerve stimuli. Epinephrine effects are mediated through alpha and beta adrenergic receptors, located in many tissues. There exist several

subtypes of alpha and beta receptors. Thus, epinephrine plays an important role in preparing the organism to meet conditions of physiologic emergency. Although medullary cells contain both epinephrine and norepinephrine, it is epinephrine that is released in sufficient quantities to have a greater physiologic impact on tissues.

When injected intravenously, epinephrine causes an immediate elevation in blood pressure and heart rate. Its effects enhance blood flow into muscle and additionally relax bronchial smooth muscle, allowing for greater oxygenation of blood. The chief metabolic changes following the injection of epinephrine are a rise in the basal metabolic rate and an increase of blood sugar. Epinephrine effects the latter action by causing an increase in the rate of glycogenolysis, or breaking down of stored sugar in the liver. These effects of epinephrine are transitory. *See* CARBOHYDRATE METABOLISM; GLYCOGEN.

Robert Kemppainen; Choh Hao Li

Bibliography. W. F. Boron and E. L. Boulpaep, *Medical Physiology*, Elsevier Science/Saunders, Philadelphia, 2005; W. F. Ganong, *Review of Medical Physiology*, 22d ed., McGraw-Hill, New York, 2005; P. R. Larsen et al., *Williams Textbook of Endocrinology*, 10th ed., Elsevier Science/Saunders, Philadelphia, 2003.

Epitaxial structures

Epitaxial interfaces in solids are a special class of crystalline interfaces where the molecular arrangement of one crystal on top of another is defined by the crystallographic and chemical features of the underlying crystal. Royer (1928) systematically showed how the geometry of the atomic arrangement on a substrate could affect the crystallographic orientation of the overgrowth. The term "epitaxy" was introduced to describe the importance of having parallelism between two lattice planes with similar networks of closely similar spacing. Epitaxial phenomena are important to study and understand, as they occur widely in nature (such as oxidation) and are the foundation by which modern semiconductor devices are grown and fabricated. *See* CRYSTAL GROWTH.

Epitaxial interfaces are a subset of a class of interfaces where lattice planes achieve a correspondence across an interface. If the matching is not perfect, such a correspondence can be achieved by a number of ways, including dilation and contraction of lattice planes; rotation of overgrowth (epilayer relative to the orientation of the substrate) until a set of closely matched lattice spacing can be found; and tilting of the epilayer with respect to the substrate (**Fig. 1**).

A combination of these effects can occur in some cases. When the epitaxial relationship is only approximate, a perfect matching of the two structures on an atomic scale would in general create prohibitively large elastic stresses in each crystal. Thus, small regions of perfect matching may describe the interface

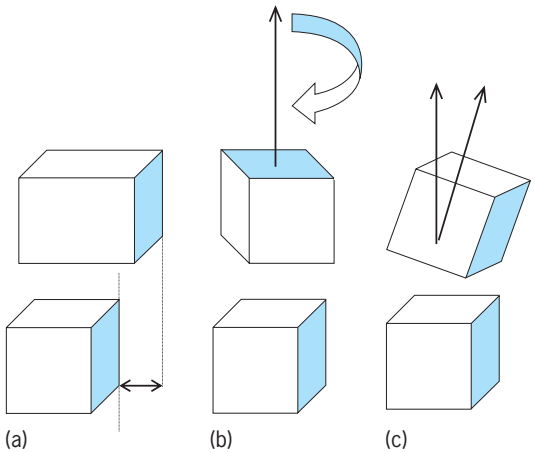

Fig. 1. Matching of lattice planes between the substrate and the epitaxial layer by (*a*) expansion or contraction of lattice plane, (*b*) in-plane rotation, and (*c*) tilting (out-of-plane rotation).

structure, separated by zones of mismatch. Depending upon the accommodation mechanisms between the crystals (that is, dilation, rotation, or tilting), the structural matching at the atomic or molecular level is often accompanied by the formation of interfacial defects which take the form of misfit dislocations. The crystallographic features of these misfit dislocations can be of varying complexity, ranging from one-dimensional features such as dislocations to two-dimensional defects such as atomic ledges and step (**Fig. 2**). *See* CRYSTAL DEFECTS; CRYSTAL STRUCTURE.

The extent to which these stresses can be accommodated before the breakdown of interfacial matching (or coherency) is very dependent upon the thickness of the film. Hence there exists a "critical thickness" before coherency at the atomic level breaks down. Numerous theoretical and experimental studies exist to understand the details of how these limits are reached. From a technological perspective, this is an important issue as one tries to grow epitaxial films which are coherent and relatively defect-free, and there is a need to know the mechanisms and means by which these limits can be extended. *See* SEMICONDUCTOR HETEROSTRUCTURES.

A simple but technologically important demonstration of defect accommodation at epitaxial interfaces is the epitaxial growth of cube-cube matching of unit cells where the spacing in the cube direction (or lattice parameter) is slightly different but the mismatch can still be accommodated by dilation processes (**Fig. 3**). In this example, the epilayer material has a slightly larger lattice parameter compared to the substrate. To achieve coherent matching of the planes that lie perpendicular to the interface, these planes in the epilayer must contract slightly to accommodate the equivalent lattice spacing of the substrate. As the substrate is of far greater thickness or volume than the epilayer, all the accommodation process occurs entirely in the thin film. This process of contraction leads to a distortion of the lattice symmetry of the epilayer, leading to a tetragonal structure

instead of maintaining the original cubic symmetry. The resulting epitaxial heterostructure now consists of a substrate and a strained layer. Multiple layers may be grown on top of the original epitaxial film, leading to a strained-layer superlattice. Such artificially layered architectures, especially in semiconductors, form an important class of materials which possess new and unique electronic properties different from the bulk forms of the same material. Introducing different levels of interfacial strain in these composite structures can alter the band gap. This unique blend of physics, mechanics, and crystal growth is the basis of developing new types of opto-electronic devices through band-gap engineering. *See* ARTIFICIALLY LAYERED STRUCTURES.

There are numerous modes of lattice accommodation apart from the example given above, including relaxation effects normal to the epitaxial interface; dislocation propagation and regeneration; formation and growth of planar defects; changes in surface morphology of the epitaxial film; and phase transitions such as ordering, and other forms of compositional modulations.

While this discussion has focused on issues governed by crystallography and structural issues, it should be noted that chemical effects play an equally important role. The concept of "matching" extends not only to geometrical matching but also to chemical compatibility. The extent to which the epitaxial films are mechanically stable due to coherency

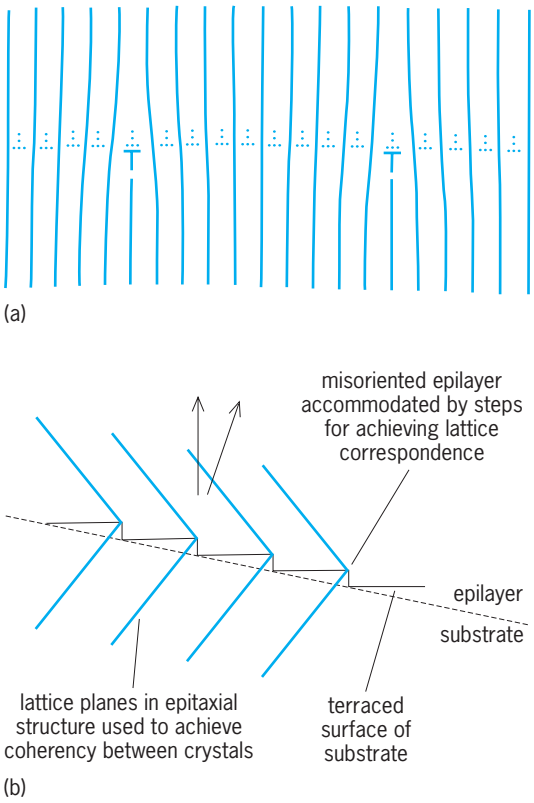

Fig. 2. Misfit accommodation occurs (*a*) between the epitaxial layer and the substrate through the presence of interfacial misfit dislocations and (*b*) by atomic steps at an interface where the epitaxial film is tilted with respect to the substrate.

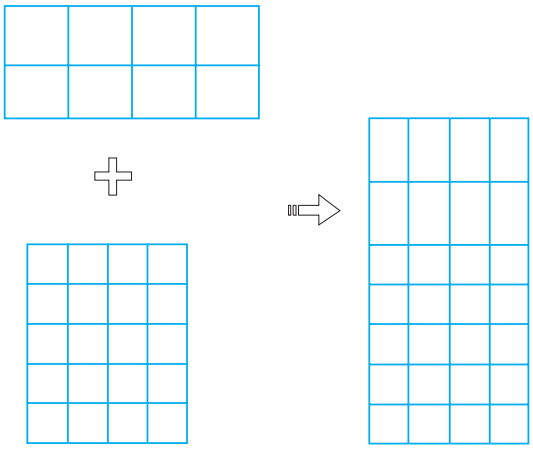

Fig. 3. Schematic illustration of the tetragonal distortion effect to accommodate purely dilational strained-layer epitaxial structure.

stresses is governed not only by the extent of lattice misfit but also by the strength of the chemical bond between the epilayer and the substrate. This property of adhesion is manifested by the extent to which the overlayer wets the substrate. There are two extreme cases by which one may describe this adhesion or wetting behavior at the molecular level. In one, proposed by F. C. Frank and J. H. van der Merwe, the adhesion between the epilayer and the substrate exceeds the binding energy between the crystal atoms, resulting in complete wetting (**Fig. 4***a*). The nucleation of an epitaxial layer occurs by the deposition of a monolayer which is easily deformed elastically and can adhere to the substrate crystal perfectly without the formation of dislocations up to a certain level of misfit (which in some cases can be as high as 10–15%). When the thickness of the layer increases, and hence also its rigidity, this state becomes unstable and epitaxial dislocations are introduced. This mode of growth is often sought in a variety of technological applications, which require the development of large-area, defect-free epitaxial layers. It should be noted that the earlier example for growing strained layer structures is based on the Frank–van der Merwe mechanism of wetting behavior. Extremely thin layers that are only a few atoms thick can be produced in this manner. Such thin layers form the microstructural foundation for the fabrication of quantum wells, which are extremely important in semiconductor device applications. *See* QUANTIZED ELECTRONIC STRUCTURE (QUEST).

The other extreme is known as the Volmer-Weber mode, which is typical of weak adhesion, resulting in poor wetting characteristics. The resulting film is not continuous but consists of small islands. This type of nanoscale architecture, however, is of value when trying to fabricate quantum dot structures, which involve having nanoscale epitaxial particles rather than continuous layers. Some systems exhibit an intermediate behavior known as the Stranski-Krastanov mode, where there is good adhesion for the first monolayer while the adhesion of the subsequent layers is weaker (Fig. 4*b* and *c*). These

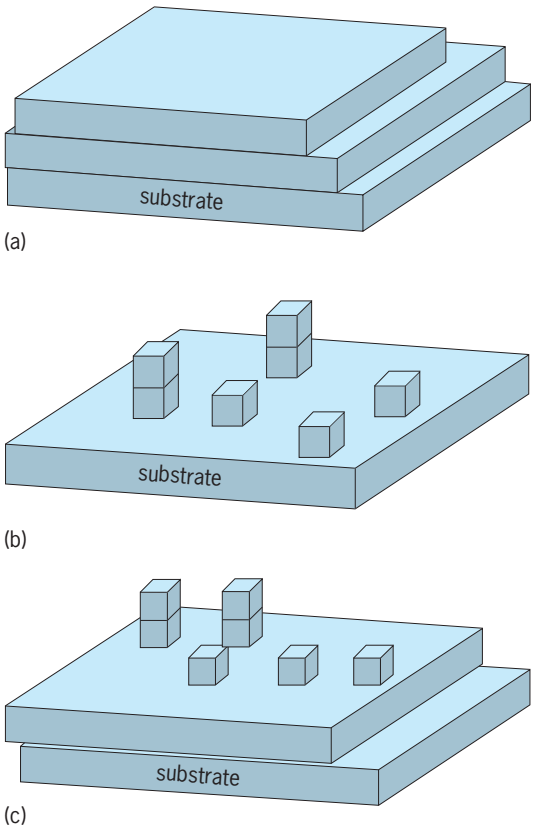

Fig. 4. Modes of epitaxial growth: (*a*) Frank–van der Merwe mode producing a complete wetting of epitaxial layers. (*b*) Volmer-Weber mode with totally incomplete wetting resulting in an island growth structure. (*c*) Stranski-Krastanov mode with initial layer growth accompanied by island growth.

latter modes of growth are important governing mechanisms in developing selective epitaxial growth on geometrically patterned substrates. *See* NANOSTRUCTURE.

Krishna Rajan

Bibliography. A. A. Chernov, *Modern Crystallography*, vol. III: *Crystal Growth*, Springer-Verlag, Berlin, 1984; P. Kordos and J. Novak (eds.), *Heterostructure Epitaxy and Devices*, NATO ASI Ser., vol. 48, Kluwer Academic, 1998; M. G. Lagally (ed.), *Kinetics of Ordering and Growth at Surfaces*, NATO ASI Ser., vol. 239, Plenum, 1990; F. R. N. Nabarro (ed.), *Dislocations in Solids*, Elsevier, Amsterdam, 1983; K. Rajan, J. Narayan, and D. Ast (eds.), *Dislocations and Interfaces in Semiconductors*, TMS, Warredale, PA, 1988; H. G. Schneider, V. Ruth, and T. Kormany (eds.), *Advances in Eptiaxy and Endotaxy*, Elsevier, Amsterdam, 1990.

Epithelium

One of the four primary tissues of the body, which constitutes the epidermis and the lining of respiratory, digestive, and genitourinary passages. The major characteristic of epithelium is that the cells are close together, separated by a very small amount of intercellular substance. The epithelium (endothelium) lining the inner cavities of the heart, blood vessels, and lymphatics differs from that of all the other groups because under abnormal conditions it may be transformed into a different type of cell characteristic of connective tissue, the fibroblast. Epithelium may be derived from any of the three primary germ layers of the very early embryo—ectoderm, endoderm, or mesoderm.

With very few exceptions (specifically, the stria vascularis of the cochlea, and the hypertrophied thyroid gland), epithelium is free of blood vessels. Nutrients reach the epithelium, and waste products leave it after passing through the ground substance of adjacent connective tissue.

Functions. The functions of epithelium are varied and include (1) protective function, by completely covering the external surface (including the gastrointestinal surface—and the surface of the whole pulmonary tree including the alveoli); (2) secretory function, by secreting fluids and chemical substances necessary for digestion, lubrication, protection, excretion of waste products, reproduction, and the regulation of metabolic processes of the body; (3) absorptive function, by absorbing nutritive substances and preserving water and salts of the body; (4) sensory function, by constituting important parts of sense organs, especially of smell and taste; and (5) lubricating function, by lining all the internal cavities of the body, including the peritoneum, pleura, pericardium, and the tunica vaginalis of the testis.

Structural integrity. The forces which hold the epithelial cells together are not satisfactorily understood. The intercellular substance between the cells, also called cement substance, is undoubtedly important. When the cement substance is weakened by removing calcium (with versene) or protein (with the enzymes trypsin or papain), the epithelial cells may be readily separated. The cement substance also contains a carbohydrate moiety. Interdigitation of adjacent cell surfaces and occurrence of intercellular bridges in certain cells may be important in holding the cells together. Finally, in certain cells local modifications of contiguous surfaces and the intervening intercellular substances, which together form the terminal bars, may be effective in the same way.

Terminal bars are visible with the light microscope after suitable staining and appear to close the spaces between the epithelial cells of the intestine at their free surface. The terminal bars are resolvable with the electron microscope into three components: (1) a very short outer tight junction, where the adjacent cells are so close that they obliterate the intercellular space; (2) an adherent zone immediately below this, where the cell membranes of adjacent cells are separated by a space of 15–20 nanometers; and (3) a desmosomal region. In the desmosomal region there is a short stretch of the cell membranes of adjacent cells where the cell membrane is slightly thickened and in direct contact with fine filaments which run into the adjacent cytoplasm. In the intercellular space between these thickened portions of the cell membranes is a thin layer of densely staining material. The intercellular bridges referred to above are essentially short processes between contiguous cells connected to each other by desmosomes. In

addition in certain sites (as in the skin), hemidesmosomes (thickened cell membrane plus cytoplasmic filaments) may be found in close contact with the subjacent basement membrane.

Structural polarity. A common feature of epithelial cells is their structural, as well as functional, polarity. In the intestinal epithelium, for example, the nucleus is somewhat toward the base of the cell. Between the nucleus and the base of the cell is a set of filamentous mitochondria, and between the nucleus and the surface is a set of granular mitochondria. Between the granular mitochondria and the nucleus is the Golgi apparatus. The striated border occurs only on the free surface of the epithelial cell. This border is visible with the light microscope and is resolvable into an enormous number of cell processes or cylinders 80–90 nm in diameter and up to about 1 micrometer long. Each process comprises a protoplasmic core enclosed in a cell membrane about 12.5 nm thick. After certain methods of preparation, the cell membrane appears as three sheets—two stainable layers separated by a nonstaining space. This trilaminar structure is now considered to be modified by the insertion of various sorts of protein molecules in the lipid coat. These proteins may be involved in antigenicity and cell recognition, permeability, cellular adhesion, and other cellular activities. Outside of this triple-layered cell membrane is another one which may contain protein or polysaccharide of some sort. It should be noted that, although the cell and the cell processes are polarized, the trilaminar structure of the cell membrane appears to be the same around the whole cell except for the specialized zones mentioned above (desmosomes, close junctions, and terminal bars). This cell polarity is equally marked in most epithelial cells and is even more marked in certain gland cells. For submicroscopic structure of cell organelles *see* CELL (BIOLOGY).

Cell and tissue affinities. It has long been known that many connective tissue cells which may be separate in their original site, except for small areas of contiguity, may in tissue cultures grow out as sheets of cells or epithelia. The factors that control this transformation are unknown but are believed to depend on some property of the cell surface not yet recognized. *See* CELL ADHESION; TISSUE CULTURE.

It is not altogether clear what forces prevent the sheet of epithelial cells from falling away from the underlying connective tissue. Undoubtedly the basement membrane of the connective tissue, when it occurs, has a cohesive property. When there is no basement membrane, the epithelium rests directly on the ground substance of the connective tissue, which has similar properties. The area of contact between epithelium and this ground substance is frequently increased by irregularities and processes of the basal surface of the epithelial cells.

Surface specialization. The outstanding property of the arrangement of most of the epithelium of the body is the economy of space achieved in the face of a broad exposure of the cell surfaces. The efficiency is achieved by the presence of numerous folds, which may be gross or microscopic and temporary or permanent. A part must also be attributed to

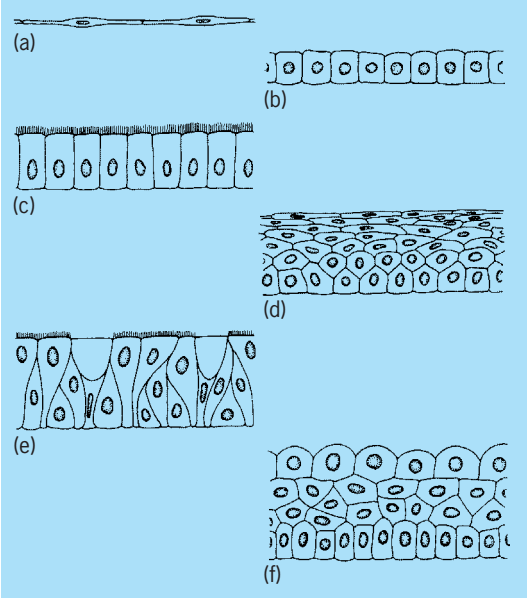

Cellular arrangements in epithelial tissues. (*a*) Squamous. (*b*) Cuboidal. (*c*) Columnar. (*d*) Stratified squamous. (*e*) Pseudostratified. (*f*) Transitional.

the surface specialization of the epithelial cells themselves, such as their minute, fingerlike processes. Another specialization of the surface or epithelial cell is the occurrence of motile cilia, each of which has a complicated structure. Especially prominent in the cilia are the nine pairs of filaments peripherally disposed in each cilium, with two central filaments. The peripheral, and frequently the central, filaments join at the base of the cilium to form a basal body, which is in turn connected with a rootlet that extends into the cytoplasm and may have a very complicated fine structure. *See* CILIA AND FLAGELLA.

Classification. Classification of epithelia is based on morphology, that is, on the shape of the cells and their arrangement (see **illus.**):

I. Single-layered.
  A. Squamous (mesothlium, descending loop of Henle in the kidney) flat.
  B. Cuboidal (duct, thyroid, choroid plexus)—cubelike.
  C. Columnar (intestine), sometimes ciliated (Fallopian tube, or oviduct)—tall.
II. Multiple-layered or stratified.
  A. Squamous (skin, esophagus, vagina)—superficial cells thin and flat, deeper cells cuboidal and columnar.
  B. Columnar (pharynx, large ducts of salivary glands), sometimes ciliated (larynx)—two or more layers of tall cells.
III. Pseudostratified (male urethra), sometimes ciliated (respiratory passages)—all cells reach to basement membrane but some extend toward the surface only part of the way, while others reach the surface.
IV. Transitional (urinary bladder)—like stratified squamous in the fully distended bladder, in the empty bladder, superfical cells rounded, almost spherical.

Special properties. An important property of epithelium is the ability of its cells to glide over surfaces. This allows replacement of dead cells to take place in the normal state, while presenting a closed surface to the external environment; replacement is especially important in wound repair, when it is necessary that a gap in the surface be filled quickly. Gliding ability is also manifested normally in the movement of cells which slide over each other in transitional epithelium, for example, when the urinary bladder is being distended or contracted. *See* GLAND. Isidore Gersh

Bibliography. P. S. Amenta, *Histology and Human Anatomy*, 6th ed., 1995; T. S. Leeson, C. R. Leeson, and A. A. Paparo, *Text-Atlas of Histology*, 1988.

Epoxide

A member of a class of three-membered ring cyclic ethers that are also known as oxiranes or alkylene oxides. The basic structure of an epoxide is analogous to that of the first member, ethylene oxide:

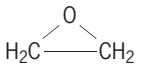

See ETHYLENE OXIDE.

Epoxides are made primarily by select oxidation reactions of alkenes; however, another classical preparation results from the ring closure of halohydrins by way of intramolecular nucleophilic displacement; that is, the reaction occurs within the same molecule, the alkoxide ion displacing the halogen to form a ring. The interest in epoxides results from their ease in preparation, and their usefulness as a reactive functional group that can give a variety of products after treatment with either electrophilic or nucleophilic reagents, or on occasion after treatment with oxidizing (for example, periodic acid) and reducing (for example, titanocene dichloride) agents. The ease of opening of the strained three-membered ring epoxides with attack of reagents in a stereospecific manner gives one or two stereochemical products (when applicable), usually in good yield. Epoxides are also used to prepare monomers, prepolymers, polymers, and copolymers, and to promote polymerization reactions, as shown in the scheme below, where the R terms are functional

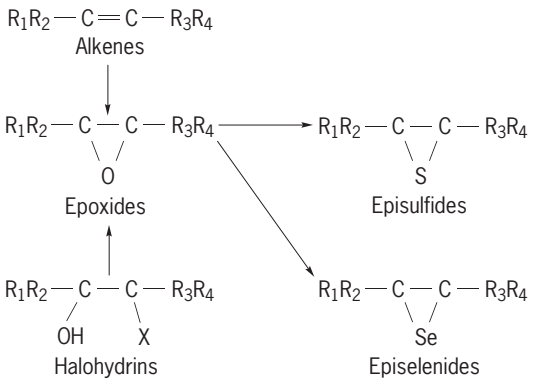

groups and X represents a halogen. *See* ALKENE; ELECTROPHILIC AND NUCLEOPHILIC REAGENTS; HETEROCYCLIC COMPOUNDS.

Preparation. Epoxides can be readily prepared by treatment of alkenes with peroxides (for example, hydrogen peroxide) or peroxy-acids (for example, *m*-chloroperoxybenzoic acid or magnesium monoperoxyphthalate). Several commercially available epoxides include ethylene oxide (oxirane, $R_1 = R_2 = R_3 = R_4 = H$), propylene oxide (methyloxirane, $R_1 = CH_3, R_2 = R_3 = R_4 = H$), styrene oxide (phenyloxirane, $R_1 = C_6H_5, R_2 = R_3 = R_4 = H$), cyclohexene oxide ($R_1/R_3 = -CH_2CH_2CH_2CH_2-$, $R_2 = R_4 = H$), and epi-chlorohydrin [1-(chloromethyl) oxirane, $R_1 = ClCH_2-, R_2 = R_3 = R_4 = H$].

Specialty epoxide preparations also include the treatment of compounds such as allyl alcohol with *tert*-butyl hydroperoxide (*t*-BuOOH), titanium tetraisopropoxide, and optically active diethyl tartrate; the treatment of two bromine atoms bonded to the same carbon atom (geminal dibromides; $R_1R_2CBr_2$) with lithium or butyllithium, followed by the condensation with carbonyl compounds (aldehydes or ketones; to variously substituted epoxides); the condensation of α-haloesters with carbonyl compounds (to epoxyesters); the addition of sulfur ylides (for example, dimethylsulfonium methylide) or diazomethane (CH_2N_2) to aldehydes or ketones (usually to geminally disubstituted epoxides); or the treatment of aldehydes with hexamethylphosphorous triamide (to symmetrically 2,3-substituted epoxides). *See* ALDEHYDE; ESTER; KETONE.

Reactions. Nucleophilic reagents used for the ring-opening of epoxides include hydroxide and water, alcohols or alkoxides, amines, bisulfite, hydrosulfide or mercaptide ion, acetylides, cyanide, dithiane anions, enamines, Grignard reagents, oxazine anions, and other reagents. Treatment of epoxides with select nucleophilic reagents can lead to other products: reactions of epoxides with triphenylphosphine form an olefin (epi-oxy-elimination); propylene oxide and certain other epoxides react with lithium diethylamide to form allyl alcohols; and reaction of sodium azide with an epoxide, such as styrene oxide, forms a β-azide alcohol, which can then be cyclized to an arizidine (three-membered heterocyclic nitrogen compound) by subsequent treatment with triphenylphosphine.

Common electrophilic reagents used for the ring-opening of epoxides include hydrochloric and hydrobromic acids (hydroiodic acid with difficulty), which results in the intermediate formation of oxonium intermediates from protonation of the epoxide oxygen atom. Straightforward acid-catalyzed hydrolysis of ethylene oxide can produce ethylene glycol (also, ethylene glycol, a 1–2 diol, can be cyclodehydrated, that is, lose a molecule of water to form the heterocyclic ethylene oxide). Similarly, treatment of epoxides with carboxylic acids yields β-hydroxyalkylcarboxylates ($RCOO—CH_2CH_2OH$). Also, select electrophilic reagents (for example, boron trifluoride etherate) rearrange epoxides to carbonyl compounds. Other reagents cause rearrangements of epoxy compounds. *See* REACTIVE INTERMEDIATES.

Uses. Poly(ethylene oxide) [—$(CH_2CH_2O)_n$] was one of the first polymers to be prepared in the laboratory. The use of numerous epoxides (as monomers, or for the synthesis of other epoxide monomers or prepolymers) in polymer synthesis has been quite extensive, and these polyethers have been prepared by both cationic and anionic polymerization techniques. The molecular weights of resulting polyethers generally range from 500 to 10,000, usually because of interfering chain-transfer reactions. An especially interesting synthesis is the opening of propylene oxide to give optically active poly(propylene oxides). Epoxides are used in the preparation of copolymers. *See* COPOLYMER; FURAN; POLYETHER RESINS; POLYMERIZATION.

Epoxides are promoters; that is, they are easy to polymerize, and their polymerization causes more difficult cyclic ethers such as tetrahydrofuran to polymerize. Epichlorohydrin has also been used for the preparation of prepolymers by condensations with the sodium salt of bisphenols. The epichlorohydrin/bisphenol A and other epoxide prepolymers have been used for the preparation of epoxy resins (thermosets), which are crossed-linked (3D) polymers, linear polymers that are connected by cross-linked molecular units. They are insoluble, infusible, and intractable, and sometimes are called epoxy resins. For example, a low-molecular-weight epoxy prepolymer can be treated and cross-linked with a polyamine or smaller molecules such as diethylenetriamine, or crossed-linked by the addition of carboxylic acid anhydrides, such as maleic anhydride. Other polymers can also be incorporated or participate in the cross-linking process, and include phenolics, ureas, and melamines. The products resulting from these prepolymer techniques are used for surface coating materials, molding, pipes, laminating, repair of damaged automobile bodies, manufacture of articles reinforced with glass fibers, durable and tough epoxy resin adhesives (glues), and many other applications.

Naturally occurring polymers such as rubber can be treated with hydrogen peroxide or peracetic acid to form an epoxidized polymer. Also, another naturally occurring polymer, cellulose, can be transformed into hydroxyethylcellulose by treating alkali cellulose with ethylene oxide. *See* CELLULOSE; RUBBER.

Related compounds. The three-membered sulfur analogs of epoxides are episulfides (also known as alkylene sulfides or thiiranes). Episulfides can be prepared from the treatment of diazoalkanes with sulfur, condensation of thioketones with sulfur ylides, and the reaction of epoxides with phosphene sulfide or thiourea/titanium tetraisopropoxide. These cyclic thioethers have been used as monomers, especially in copolymer preparations. There is some evidence for the preparation of several episelenides (replacement of the oxygen of epoxide with selenium; also known as ethylene selenides or seleniranes); but heating solutions of these less stable compounds results in the elimination of selenium and yields the alkene.

Charles F. Beam

Bibliography. T. L. Gilchrist, *Heterocyclic Chemistry*, 3d ed., 1997; R. B. Seymour and C. E. Carraher, Jr., *Seymour and Carraher's Polymer Chemistry*, 5th ed., 2000; M. Smith and J. March, *March's Advanced Organic Chemistry*, 5th ed., 2000.

Epsomite

A mineral with the chemical composition $MgSO_4 \cdot 7H_2O$. Epsomite, or epsom salt, occurs in clear, needlelike, orthorhombic crystals. More commonly it is massive or fibrous, although crystals from salt lakes on Kruger Mountain near Orville, Washington, are reported to be several feet long. Fracture is conchoidal. Luster varies from vitreous to silky. Hardness is 2–2.5 on Mohs scale and specific gravity is 1.68. The mineral has a salty bitter taste and is soluble in water.

Epsomite is found as a capillary coating in limestone caves and in coal or metal mine galleries. It loses readily some of its water of crystallization in dry air. It is also found associated with gypsum and in thin layers in salt deposits of oceanic origin or from salt lakes. *See* MAGNESIUM.

Edward C. T. Chao

Epstein-Barr virus

An antigenically distinct member of the herpesvirus group of viruses, whose genome is deoxyribonucleic acid (DNA). Under electron-microscopic observation, the mature virus is indistinguishable in size and structure from the other human herpesviruses; that is, it has a nucleocapsid about 110 nanometers in diameter, and when an envelope is present, the particle diameter is about 180 nm. The 78-nm central core, which is often observed in other herpesviruses, is not clearly visible in Epstein-Barr (EB) particles within infected cells, and few of the intracellular particles possess an envelope. Only about 10% of the extracellular virus particles are enveloped.

EB virus is the cause of one benign disease (infectious mononucleosis), and is associated with certain types of cancer; however, the great majority of EB virus infections are clinically inapparent.

The virus was detected initially by electron microscopy in a small proportion of cells in continuous lymphoblastoid cell lines derived from Burkitt's lymphoma (but particles have not been seen in cells of the tumor itself). This lymphoma is a tumor indigenous to children in central Africa. The virus also has been detected in cell lines derived from nasopharyngeal carcinomas, a type of cancer found with high frequency in persons from southern China. The virus is found in peripheral blood leukocytes from normal individuals and from patients with infectious mononucleosis. *See* INFECTIOUS MONONUCLEOSIS; LYMPHOMA.

With some strains of EB virus, attempts to infect human adult and cord blood lymphocytes and marmoset lymphocytes in cultures have resulted in the

establishment of continuous cell lines, suggesting that these cells have been transformed (that is, "immortalized") by the virus.

The high prevalence and high titers of EB antibody (detectable by immunofluorescence, complement-fixation, and gel-diffusion tests) in patients with Burkitt's lymphoma and those with postnasal carcinoma led to the assumption that the association might be etiologic. However, subsequent seroepidemiological surveys showed that infection with EB virus in normal populations is widespread not only in Africa and in Asia but also all over the world. In some areas, including urban areas of the United States, about 50% of children 1 year old, 80–90% of children over age 4, and 90% of adults have antibody to EB virus. Although the mechanism of virus transmission is unknown, it could take place through oropharyngeal excretion, which now has been found to be very common, especially in individuals with overt infectious mononucleosis but also in persons without symptoms. Furthermore, EB virus has an extreme predilection for cells of lymphoid origin, and hence could quite possibly be merely a passenger virus in the lymphoma cells. The question of whether EB virus is etiologically related to lymphoma or to postnasal carcinoma, or to both, remains to be answered. Indeed, other viruses (for example, herpes simplex virus and reovirus) have also been isolated from Burkitt's lymphomas. On the other hand, the possibility remains that EB virus is the agent responsible for the induction of these human cancers, or is one of the cofactors causing the malignancies.

The techniques for searching for viruses in human cancers have become quite sophisticated, but technical problems still make it difficult to determine whether a virus is merely a passenger inhabiting the tumor cells. Even more difficult is the identification of traces of a virus which may have caused the cancer but which is no longer present in its complete or infective form in the cancer it initiated. Such traces may include only selected portions of the viral genome residing in the cancer cell and molecular probes have been developed for this search. Since EB virus now has been shown capable of inducing lymphomas in the cottontop marmoset (a South American monkey), studies of this virus–primate host system should help to clarify the relationship between EB virus and human disease.

If EB virus is indeed confirmed as having a role in the development of human malignancies, then one major question to be resolved is how a virus so ubiquitous can be involved in so wide a variety of responses—ranging from asymptomatic infection through infectious mononucleosis to production of cancers. However, it should be recalled that many virus infections (for example, polio virus, hepatitis viruses, certain of the arboviral encephalitides) have a wide spectrum of outcomes, ranging from inapparent infection to severe syndromes. *See* ANIMAL VIRUS; ONCOLOGY.

Joseph L. Melnick

Bibliography. D. V. Ablashi et al. (eds.), *Epstein-Barr Virus and Human Disease*, 1990; *Peak Immunity: How to Fight Epstein-Barr Virus, Candida, Herpes Simplex, and Other Immuno-Depressive Disorders and Win*, Medical Sciences, General Medical Series, 1989.

Equalizer

An electronic filter that modifies the frequency response (amplitude and phase versus frequency) of a system for a specific purpose. While filters in some applications perform the conceptually simple operations of rejecting specific bands of frequencies and passing other bands of frequencies, equalizers typically realize a more complicated frequency response in which the amplitude response varies continuously with frequency, amplifying some frequencies and attenuating others. An equalizer may have a response fixed in time or may be automatically and continuously adjusted. However, its frequency response is usually matched to some external physical medium, such as an acoustic path or communication channel, and thus inherently needs to be adjustable.

Equalizers can be used in many applications. In music and sound reproduction, equalizers can compensate for artifacts of the electrical-to-sound conversion or for unwanted characteristics of the acoustic environment such as sound reflections or absorption. Sound-recording and sonar systems can use equalizers to reduce unwanted interference. Most analog recording and playback devices, such as audio and video tape recorders, incorporate equalizers to compensate for the undesirable aspects of the recording medium, such as high-frequency roll-off, as well as to reduce noise and maximize dynamic range. *See* ELECTRICAL INTERFERENCE; ELECTRICAL NOISE; SONAR; SOUND RECORDING.

An important application of equalization is to enhance the performance of systems that communicate or record digital signals (streams of bits). The transmission of digital (as opposed to analog) information is rapidly becoming dominant, and equalization is a critical component of almost all such systems. All communications and recording systems utilize a physical medium, such as wires; coaxial cables; radio, acoustic, or optical-fiber waveguides; or magnetic and optical recording media. These media cause distortion; that is, the output signal is different from the input signal. For example, on radio or acoustic channels there are often multiple paths from transmitter to receiver, each having a slightly different delay and superimposed at the receiver. An equalizer is an electrical device that compensates for this distortion, reversing the effect of the channel and returning a waveform approximating the input signal. The channel output signal in response to a particular input signal (. . . , 0, 0, 1, 0, 0, . . .) may differ from the input, but the equalizer output reproduces the channel input, at least to close approximation (**Fig. 1**). *See* DISTORTION (ELECTRONIC CIRCUITS).

If the characteristics of the channel are well known, the equalizer can be fixed or nonadaptive. More commonly, the detailed characteristics of a

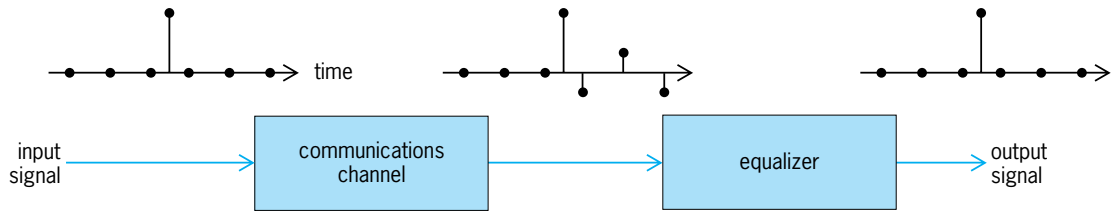

Fig. 1. Communications channel with an equalizer placed at the output. The equalizer recovers an accurate replica of the channel input signal.

channel are not known in advance. For example, an equalizer may be required to compensate for any length of wire, from very short up to a maximum (such as the connection from a telephone central office to the subscriber's telephone instrument). In other cases, the channel may be varying with time, as is characteristic of the radio channel from a fixed transmitter to a moving vehicle. (On a radio channel, the signal arriving at the vehicle may represent multiple superimposed paths, each path corresponding to a reflection from a different building or mountain and having a different delay. The net effect depends on the vehicle location, and hence changes as the vehicle moves.) An adaptive equalizer does not need to know the channel in advance, but is nevertheless able to adjust itself to compensate. Adaptive equalizers are important for achieving high bit rates in telephone computer modems, and also for digital communications over radio channels. *See* DATA COMMUNICATIONS; MODEM.

Prior to about 1970, equalizers were usually realized by continuous-time filters constructed from interconnected resistors, capacitors, and inductors. Using modern electronic technologies, equalizers are almost always realized in discrete time. That is, the continuous-time channel output, say $x(t)$ where t is time, is first sampled at uniformly spaced time intervals, yielding a sampled signal given by Eq. (1),

$$x_k = x(kT) \tag{1}$$

where T is the interval between samples. As long as T is sufficiently small, no essential information is lost. The discrete-time signal x_k, consisting of samples of the input signal, is input to the equalizer. Time is represented by the discrete variable k rather than the continuous variable t.

Structure. A common internal structure for a discrete-time equalizer is the transversal filter (**Fig. 2**). The output at any given time k is a linear combination of the current and a set of past input samples, given mathematically by Eq. (2) where the

$$y_k = \sum_{m=0}^{N-1} a_m x_{k-m} \tag{2}$$

current input and output samples are x_k and y_k respectively, and the set of N weighting factors, a_m, $1 \le m \le N-1$ are known as the filter coefficients. The coefficients are chosen or adjusted to control the equalization. A transversal filter can be thought of as introducing its own distortion controlled by the filter coefficients, and this distortion is chosen to reverse the distortion on the channel. This is usually possible as long as N is sufficiently large.

Modes of adaptation. In order to make a transversal filter adaptive and, in particular, to determine the channel distortion by looking at the output of the channel alone, the equalizer is normally operated in two modes: the training mode and the tracking mode. In the training mode, the channel characteristics are learned for the first time (for example, with a telephone computer modem at the beginning of a phone call), with very little prior knowledge of the channel necessary. In the tracking mode, the characteristics of the channel are followed, assuming those characteristics do not change too quickly. If the channel characteristics are fixed, the tracking mode is unnecessary.

In the training mode (**Fig. 3***a*), for purposes of learning the channel, a specially chosen training signal u_k is applied to the channel input, and is presumed to be known to the equalizer by prior agreement. (Normal communications over the channel is precluded during the training period.) The difference between the equalizer output and the training signal, known as the error signal e_k, is used to adapt the equalizer. The filter coefficients are adjusted so as to force this error signal, given by Eq. (3), to be

$$e_k = u_k - \sum_{m=0}^{N-1} a_m x_{k-m} \tag{3}$$

small. The e_k being consistently nearly zero is an indication that the equalizer is correctly compensating for the channel distortion. The channel characteristics are not learned directly; rather, the filter coefficients that cause the equalizer to reverse the channel distortion are determined.

The tracking mode (Fig. 3*b*) proceeds while actual communication is occurring, and hence it uses whatever channel input signal is to be communicated, that signal being unknown to the equalizer. A tracking mode is possible only for digital communications, where the transmitted signal u_k assumes only a

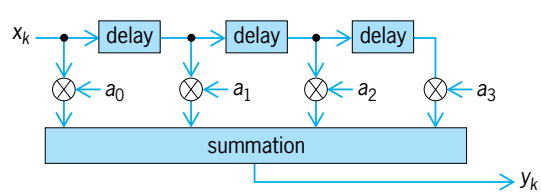

Fig. 2. Transversal filter, which forms a weighted summation of the current and past input signal samples (shown for $N = 4$ coefficients).

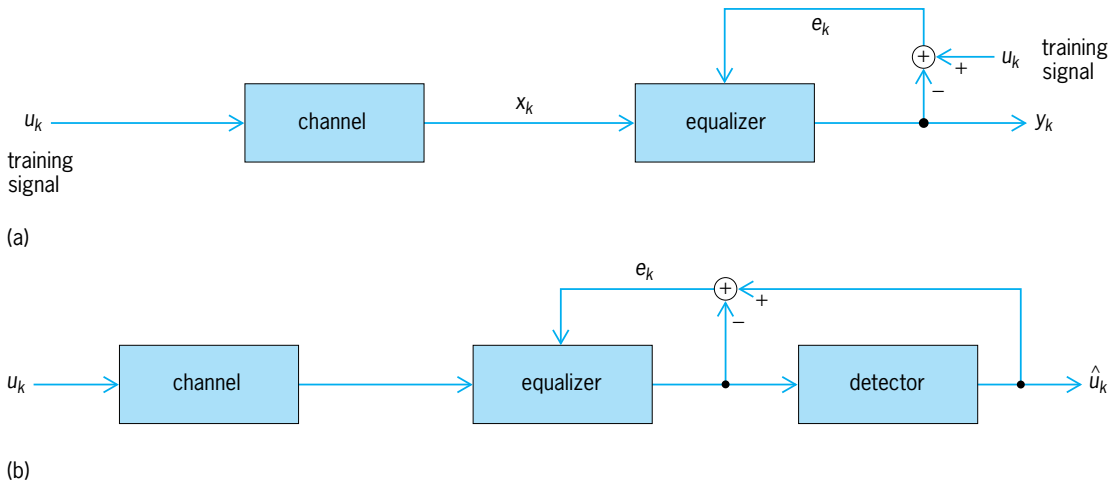

Fig. 3. Equalizer modes of operation. (*a*) Training mode. The channel input signal $\hat{u}_k$ is known to the equalizer, the adaptation of which has available the error signal e_k. (*b*) Tracking mode. The error signal is formed by using a detected version $\hat{u}_k$ of the channel input signal $\hat{u}_k$.

finite and discrete set of values at any given instant of time. (For example, input signal values may carry only one bit of information. That one bit can assume one of two values, 0 or 1, which can be represented by one of two signal values, say $u_k = +1$ or $u_k = -1$.) Tracking assumes that the equalization is always accurate enough that any residual distortion caused by incorrect coefficients does not cause errors in the detection of the input signal values. Thus, a detector placed at the output of the equalizer forms an accurate estimate $\hat{u}_k$ of the input signal u_k, giving a reference for equalizer adaptation. As in training, an error signal is formed, replacing the actual input signal u_k by the detected signal $\hat{u}_k$. Correct operation depends on $\hat{u}_k = u_k$, at least most of the time. It is quite possible for this scheme to fail if the channel changes too quickly, since the uncorrected distortion may become so large that the detector does not work properly. However, adaptation works properly in the presence of occasional detection errors (such as those caused by random noise) since the adaptation averages out these errors.

Adaptation algorithms. Adaptation chooses the filter coefficients a_m, $1 \le m \le N-1$, to minimize, in some sense, the squared error e_k^2. The goal is not to minimize this squared error at a particular time k, but to use a criterion that ensures that the squared error is consistently small. For tracking, the adaptation should have a finite memory; that is, it should not depend on the far past (when the channel was different) but the recent past. An example is given by the quantity in Eq. (4) for

$$\xi_k = \sum_{m=0}^{\infty} \beta^m e_{k-m}^2 \tag{4}$$

$0 < \beta < 1$. Since β^m is larger for small m and falls off rapidly for larger m, if ξ_k is minimized at time k, the recent channel characteristics are weighted much more heavily. It is a straightforward exercise in calculus to find the coefficients a_m, $1 \le m \le N-1$, that minimize ξ_k. A slowly changing channel can be tracked by re-solving for the coefficients at each time instant k.

There are a number of available algorithms for adapting the transversal filter coefficients, differing in the criterion that is minimized and the algorithmic formulation of the minimization. If greater computational complexity is acceptable, a greater rate of change in the channel characteristics can be tracked. Adaptive equalizers are usually implemented in a software program running on a special-purpose microprocessor known as a programmable digital signal processor. In an application such as digital communication with a moving vehicle by radio, where the channel characteristics can change rapidly, the adaptation algorithms are quite sophisticated and require a very high rate of computation. *See* ALGORITHM; DIGITAL FILTER; ELECTRIC FILTER; MICROPROCESSOR; OPTIMIZATION.

David G. Messerschmitt

Bibliography. C. Cowan and P. Grant, *Adaptive Filters*, 1985; K. Feher, (ed.) *Advanced Digital Communications Systems and Signal Processing Techniques*, 1987; E. Lee and D. Messerschmitt, *Digital Communication*, 1993; B. Mulgrew and C. Cowan, *Adaptive Filters and Equalizers*, 1988; B. Widrow and S. Stearns, *Adaptive Signal Processing*, 1985.

Equation of continuity

An equation of continuity appears in many branches of physics. In dynamic field theory, it is essentially a statement that charge is conserved or that the rate of increase of charge in any region equals the current $\mathbf{i}$ flowing into that region. If v is the volume enclosed by the surface S, this statement may be expressed in integral form, as given in Eq. (1), where ρ is the

$$\int_S \mathbf{i} \cdot \mathbf{n}\, dS = \int_v \nabla \cdot i\, dv = -\frac{\partial}{\partial t} \int_v \rho\, dv \tag{1}$$

charge density and $\mathbf{n}$ is a unit vector normal to S. The second integral comes from the first by Gauss' divergence theorem. When currents and charges are

confined to a surface S bounded by the curve s, this becomes Eq. (2), where θ is the angle between ds and

$$\oint i \sin\theta \, ds = -\frac{\partial}{\partial t}\int_S \rho \, dS \qquad (2)$$

$\mathbf{i}$ which is directed into the area S. This states that the current crossing the boundary of an area equals the rate of increase of charge in the area. The volume in Eq. (1) is arbitrary, so the integrands in the last two integrals are equal. This leads to the differential form in Eq. (3), where i_x, i_y, and i_z are the rectangu-

$$\nabla \cdot \mathbf{i} = \frac{\partial i_x}{\partial x} + \frac{\partial i_y}{\partial y} + \frac{\partial i_z}{\partial z} = -\frac{\partial \rho}{\partial t} \qquad (3)$$

lar components of $\mathbf{i}$. Maxwell's equations satisfy Eq. (3). For application to the motion of charged particles, Eq. (3) would be written as Eq. (4). When $\mathbf{i}$ and

$$\nabla \cdot \rho\mathbf{v} = \rho\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}\right) = -\frac{\partial \rho}{\partial t} \qquad (4)$$

ρ vary sinusoidally with time, they may be written as phasors, which are complex numbers such that when multiplied by $e^{j\omega t}$, the real part of the product gives the amplitude, phase, and time dependence. The only change in the preceding equations when $\mathbf{i}$ and ρ are phasor quantities is the replacement of $\partial/\partial t$ by $j\omega$. *See* FLUID-FLOW PRINCIPLES; GAUSS' THEOREM; MAXWELL'S EQUATIONS. William R. Smythe

Equation of time

The annual, cyclic variation between mean solar time shown on uniformly running clocks and apparent solar time displayed on sundials.

In the course of the Sun's daily east-to-west transit of the sky, the Sun crosses the meridian, an imaginary line running from north to south that passes overhead and divides the sky into equal halves. An observer in the middle of a time zone generally thinks of noon as being the moment that the Sun reaches the meridian. This event, however, corresponds to noon recorded by mechanical or electronic clocks on only four dates each year (approximately April 16, June 14, September 1, and December 25). On all other dates the Sun reaches the meridian either

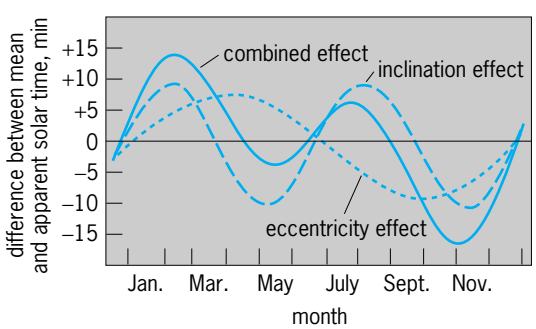

Fig. 1. Graph of the equation of time, showing how the equation results from combining the effects of the inclination of the Earth's axis and the eccentricity of its orbit. (*After B. M. Oliver, The shape of the analemma, Sky Telesc., 44:20–22, July 1972*)

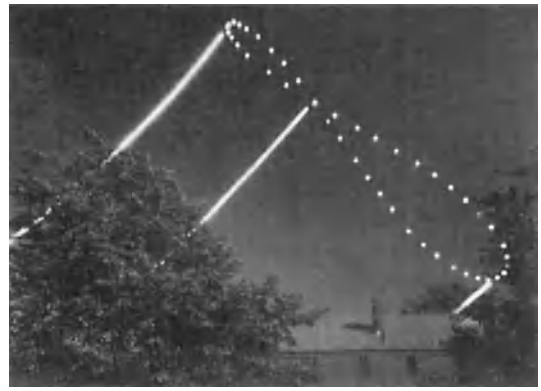

Fig. 2. Analemma recorded in a multiple-exposure, year-long photograph obtained with an east-facing, permanently mounted camera that made an exposure approximately once a week at exactly 8:30 a.m. Eastern Standard Time. On three occasions the shutter was left open from dawn until shortly before 8:30 to show the Sun's path across the sky. All images of the Sun were made through a special dense filter. The foreground was recorded without the filter one afternoon when the Sun was in the western sky and roughly behind the camera (as indicated by the shadow of the chimney). (*Photograph by Dennis di Cicco*)

early or late, with the extremes being 16.3 min early around November 3 and 14.3 min late around February 12. This difference is the equation of time, and results from the combined effects of Earth's axis of rotation being tipped 23° relative to Earth's orbital plane and the elliptical rather than circular shape of the orbit. *See* MERIDIAN.

The Earth's orbital motion causes the Sun to move eastward against the background stars along a path called the ecliptic. This motion is about 1° each day. Although the Earth turns once on its axis every $23^h\ 56^m4^s$, it must turn slightly more in order for the eastward-moving Sun to return to the meridian. Thus, the length of time from one meridian passage of the Sun to the next averages 24 h over the course of a year.

If the Earth had a perfectly circular orbit and its axis of rotation was perpendicular to the orbital plane, the Sun's motion would be along the celestial equator and uniform throughout the year. It is this idealized case of a mean Sun that shows the mean solar time kept by mechanical and electronic clocks and watches. however, the elliptical orbit brings the Earth closest to the Sun in January. This proximity, as J. Kepler discovered in the sixteenth century, causes the Earth to move more rapidly in its orbit than in July, when the Earth is farthest from the Sun. The changing orbital speed varies the Sun's apparent rate of motion along the ecliptic and is in part responsible for the equation of time. *See* CELESTIAL MECHANICS; KEPLER'S LAWS.

Slightly more influential is Earth's tipped axis, which varies the Sun's position north and south of the celestial equator according to the season. Around the time of the spring and autumn equinoxes the Sun moves at a steep angle relative to the celestial equator. Its daily motion projected onto the equator is less than at the solstices, when the Sun travels parallel to

the equator. This situation also creates a departure between the Sun's actual position and that of a mean Sun moving uniformly along the celestial equator. The effects of the inclination of the Earth's axis and the eccentricity of its orbit are combined (**Fig. 1**).

If the Sun's daily position is considered relative to the meridian at the moment a clock reads noon, from the equation of time it is seen that the Sun will sometimes be west of the meridian (early) and sometimes east of it (late). The coupling of this variation with the north-south movement of the Sun along the ecliptic causes the Sun to mark out a large figure-eight known as the analemma (**Fig. 2**). This pattern is sometimes shown on the tropical zone of Earth globes (usually in the Pacific Ocean). The analemma shows the latitude at which the Sun passes directly overhead on any given date and the equation of time. *See* EARTH ROTATION AND ORBITAL MOTION; TIME. Dennis Di Cicco

Bibliography. B. M. Oliver, The shape of the analemma, *Sky Telesc.*, 44:20-22, July 1972.

Equations, theory of

The branch of mathematics concerned with finding facts concerning the roots of algebraic equations and finding methods for obtaining them. The most important type of algebraic equation is the polynomial equation in one unknown which is an expression of the form $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$, where x is called the unknown, or variable; n is a positive whole number; and the a_i, with $i = 0, 1, \ldots, n$, are constants, or fixed numbers, called coefficients of the equation. The left member of the equation is called a polynomial in one variable of degree n. A root of such an equation is a number which, when substituted for the variable x, makes the left member zero. For example, 3 is a root of the equation $x^3 + 2x^2 - 13x - 6 = 0$. In addition, systems of equations in one or more variables are considered, and here the problem is to find values for the variables which simultaneously satisfy each equation of the system. *See* LINEAR SYSTEMS OF EQUATIONS; POLYNOMIAL SYSTEMS OF EQUATIONS.

The topics covered in a systematic study of the theory of equations can be placed in the following principal subdivisions: properties of a polynomial which do not depend on the particular number system containing the coefficients of the polynomial; factorization of polynomials; equations with coefficients which are rational, real, or complex numbers; determination of bounds for real roots, and systematic methods for approximating real roots of equations; the solution of quadratic, cubic, and quartic equations by radicals.

This article is limited to polynomials and to equations which have rational, real, or complex numbers as coefficients. Each of these number systems constitutes a number field.

Let $f(x) = a_n x^n + a_{n-1}x^{n+1} + \cdots + a_1 x + a_0$ and $g(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0 \neq 0$ be polynomials with coefficients in a number field F. Then the division algorithm, which is a formal statement of the ordinary division process for polynomials, states that there exist unique polynomials $q(x)$ and $r(x)$, with coefficients in F, such that $f(x) = q(x)g(x) + r(x)$, where either $r(x) = 0$ or $r(x)$ has degree less than $g(x)$. If $r(x) = 0$, then $g(x)$ divides $f(x)$. Immediate corollaries of the division algorithm are the remainder and factor theorems. If $g(x)$ is the linear polynomial $(x - a)$, then $f(x) = q(x)\,(x - a) + r$, where r is a constant. Then, substituting $x = a$, the result is $f(a) = r$. This is the remainder theorem. If $f(a) = r = 0$, then a is a root of the equation $f(x) = 0$, giving the factor theorem, which states that a is a root of $f(x) = 0$ if, and only if, $(x - a)$ is a factor of $f(x)$. It follows from the factor theorem that the equation $f(x) = 0$ has at most n roots in F, where $f(x)$ is of degree n.

The greatest common divisor (gcd) of two polynomials $f(x)$ and $g(x)$ is a polynomial $d(x)$ which divides both $f(x)$ and $g(x)$, and is divisible by every other polynomial which also divides $f(x)$ and $g(x)$. A process called the euclidean algorithm, based on successive application of the division algorithm, is used to find the gcd $d(x)$ having coefficients in the smallest number field that contains the coefficients of $f(x)$ and $g(x)$. The polynomials $f(x)$ and $g(x)$ are called relatively prime if their gcd is a constant.

The factorization of a polynomial $f(x)$ depends on the particular number field F under consideration. For example, the polynomial $x^5 - {}^1/_2x^4 - x^3 + {}^1/_2x^2 - 2x + 1$ having coefficients which are rational numbers, factors as $(x^2 + 1)\,(x^2 - 2)\,(x - {}^1/_2)$ over the rational numbers; as $(x^2 + 1)\,(x - \sqrt{2})\,(x + \sqrt{2})\,(x - {}^1/_2)$ over the real number; and as $(x + i)\,(x - i)\,(x - \sqrt{2})\,(x + \sqrt{2})\,(x - {}^1/_2)$ over the complex numbers. A polynomial $f(x)$ with coefficients in F is irreducible over F if it cannot be expressed as a product of polynomials of lower degree. Every polynomial can be expressed in essentially one way as a product of irreducible factors, although there is no general algorithm which enables one to obtain this expression. For polynomials with rational coefficients, there is such an algorithm, devised by L. Kronecker. There are methods for finding the repeated factors of a polynomial which are often useful as a first step in factoring a polynomial.

If a polynomial equation $f(x) = 0$ with rational coefficients is multiplied by a suitable whole number, an equation with whole number coefficients is obtained. If r/s, a rational number in lowest terms, is a root of such an equation, then r divides the constant term and s divides the leading coefficient. Hence, the rational roots of $f(x) = 0$ can be found by a finite number of trials.

The fundamental theorem of algebra states that a polynomial equation with complex coefficients has a complex root. From this it follows immediately that a polynomial of degree n with complex coefficients factors into n linear factors over the complex numbers.

If $f(x) = 0$ has real coefficients, then it can be shown that, if the complex number $a + bi$ is a root of $f(x)$, the conjugate complex number $a - bi$ is also

a root. Thus, the real, irreducible factors of a polynomial with real coefficients are linear or quadratic. A further consequence is that, if $f(x)$ with real coefficients has odd degree, then $f(x) = 0$ has at least one real root.

The property of a polynomial $f(x)$ with real coefficients, which is basic for the study of the equation $f(x) = 0$, is that $f(x)$ defines a continuous real function. The location principle which follows from this states that if there exist real numbers $a < b$ such that $f(a)$ and $f(b)$ have opposite signs, then $f(x) = 0$ has a real root r such that $a < r < b$. The location principle is used to isolate the real roots and is fundamental in systematic schemes such as Horner's method and Newton's method for approximating the real roots. A numerical method, known as Graeffe's method, can be used to approximate both the real and the complex roots.

Important results used in finding the real roots of $f(x) = 0$ with real coefficients include the following: Rolles' theorem from differential calculus, which states that, between two real roots of $f(x) = 0$, there is at least one real root of the derivative $f'(x) = 0$; Sturm's theorem, which gives an exact count of the number of real roots in an interval between two real numbers $a < b$; and Descartes' rule of signs, which states that the number of positive roots of $f(x) = 0$ equals the number of variations in sign of the coefficients of $f(x)$ minus a nonnegative even number. The negative roots of $f(x) = 0$ are positive roots of $f(-x) = 0$.

The results on the bounds for the real roots of $f(x) = 0$ are based on the fact that, for sufficiently large values of x, the sign of $f(x)$ is the same as the sign of the leading term $a_n x^n$ of $f(x)$. In particular, $f(x) = 0$ has no real roots for $|x| \geqq |a_k/a_0| + 1$, where a_k is the coefficient of $f(x)$ having the greatest numerical value.

Polynomial equations of degree 2, 3, and 4 are solvable by radicals. This means that there are formulas which give the roots in terms of the coefficients of the equation and that these formulas involve only the rational operations and the operation of extraction of roots. The principal methods for the solution of the cubic and quartic equations were devised by J. Cardan and L. Ferrari, respectively. By use of the Galois theory of equations, it can be proved that for $n > 4$ there cannot exist a formula involving only rational operations and root extractions for expressing the roots of every polynomial equation of degree n in terms of the coefficients. *See* DETERMINANT; MATRIX THEORY.

Ross A. Beaumont

Bibliography. R. A. Barnett and T. J. Kearns, *Elementary Algebra: Structure and Use*, 6th ed., 1994; D. Dobbs and R. A. Hanks, *Modern Course on the Theory of Equations*, 2d ed., 1992.

Equator

The great circle around the Earth, equally distant from the North and South poles, which divides the Earth into Northern and Southern hemispheres. It is the greatest circumference of the Earth because of centrifugal force from rotation, and resultant flattening of the polar areas.

The Earth's rotational axis is vertical to the plane of the Equator, and because the inclination of the axis is 66.5° (66.55°) from the plane of the ecliptic, the plane of the Equator is always inclined 23.5° from the ecliptic.

At noon on the days of the vernal and autumnal equinoxes (March 21 and September 23), the Sun on the Equator is 90° above the horizon. The lowest angle of the Sun is 66.5°. This occurs at noon, at the winter and summer solstices (June 21 and December 22), when the Sun is vertical at the Tropic of Cancer and Tropic of Capricorn, respectively. Days and nights are always equal because the plane of the ecliptic intersects the equatorial plane at the Earth's center. Consequently, the Sun at the Equator rises and sets at approximately 6 A.M. and 6 P.M. throughout the year.

The celestial equator in astronomy is equally distant from the celestial poles and is the great circle in which the plane of the terrestrial Equator intersects the celestial sphere. *See* ASTRONOMICAL COORDINATE SYSTEMS; MATHEMATICAL GEOGRAPHY.

Van H. English

Bibliography. A. H. Strahler and A. Strahler, *Introducing Physical Geography*, 4th ed., 2005.

Equatorial currents

Ocean currents near the Equator. The westward trade winds that prevail over the tropical Atlantic and Pacific oceans drive complex oceanic circulations characterized by alternating bands of eastward and westward currents (**Fig. 1**). The intense currents are confined to the surface layers of the ocean; below a depth of approximately 100 m (330 ft) the temperature is much lower, and the speed of ocean currents is much slower. The westward surface currents tend to be divergent—they are associated with a parting of the surface waters—and therefore entrain cold water from below. The water temperature rises as the currents flow westward, so that temperatures are low in the east and high in the west, except between 3 and 10°N where eastward surface currents create a band of warm water across the Pacific and Atlantic oceans. The distinctive sea surface temperature pattern in which surface waters are warm in the west and cold in the east, except for the warm band just north of the Equator, reflects the oceanic circulation (**Fig. 2**). A dramatic change in this pattern every few years during El Niño episodes, when the temperature of the eastern tropical Pacific Ocean rises, is associated with an intensification of the eastward currents and a weakening (sometimes reversal) of the westward currents. *See* EL NIÑO.

South Equatorial Current. This current flows westward in the upper ocean, has its northern boundary at approximately 3°N, and attains speeds in excess of 1 m/s (3.3 ft/s) near the Equator. It is directly driven by the westward trade winds and has its

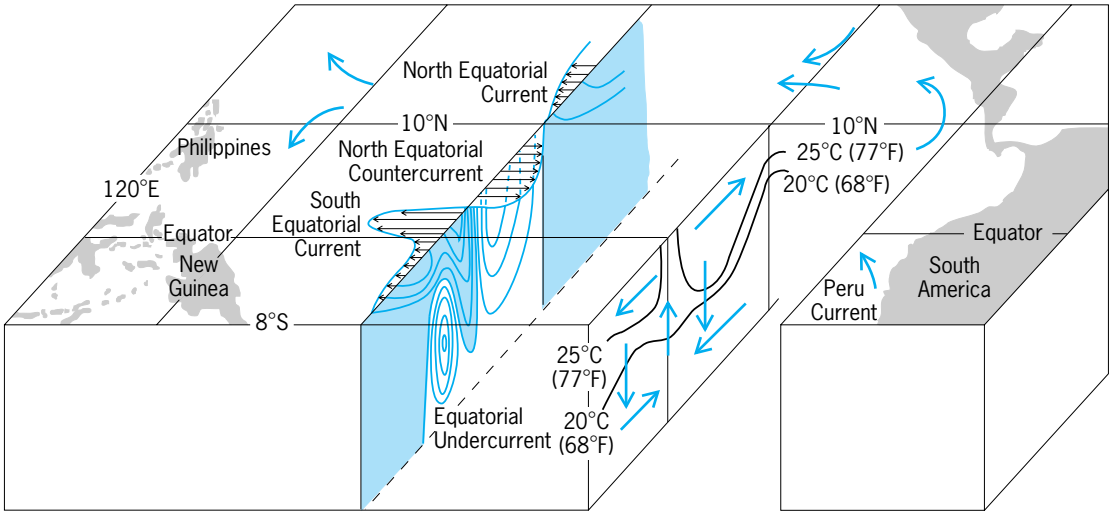

Fig. 1. Circulation of the tropical Pacific Ocean. The shaded areas indicate regions of westward flow. The principal eastward currents are the Equatorial Undercurrent centered on the Equator, and the North Equatorial Countercurrent to the north of the Equator. The heavy arrows in the vertical plane show intense upwelling at the Equator, poleward flow in the surface layers, and sinking motion off the Equator, sustained by equatorward flow at depth of 100 m (330 ft) approximately. This motion affects the thermal field as is evident from the two isotherms.

origins in the cold, northwestward-flowing Peruvian coastal current. Because the Coriolis force deflects water parcels to their right in the Northern Hemisphere and to their left in the Southern Hemisphere, this current is divergent at the Equator. As a consequence, cold water from below wells up along the Equator. *See* CORIOLIS ACCELERATION; UPWELLING.

North Equatorial Countercurrent. This current flows eastward immediately to the north of the South Equatorial Current. The boundary between these two currents is a sharp thermal front that is clearly evident in satellite photographs. The front can literally be a green line, hundreds of yards wide, because of the abundance of phytoplankton. (It frequently has westward-propagating undulations with a wavelength on the order of 1000 km or 600 mi. These are instability waves associated with the enormous shear of the two currents.) This current, which is counter to the wind, is driven by the torque (curl) that the wind exerts on the ocean. To its north is a colder westward current known as the North Equatorial Current. *See* OCEAN WAVES; PHYTOPLANKTON.

Equatorial Undercurrent. This current, which in the Pacific Ocean was originally known as the Cromwell Current, is an intense, narrow, eastward, subsurface jet that flows precisely along the Equator across the width of the Pacific. Its core, where speeds can be in excess of 1.5 m/s (5 ft/s), is at an approximate depth of 100 m (330 ft); its width is approximately 200 km (120 mi). A similar current exists in the Atlantic Ocean. In the Indian Ocean it is often present along the Equator, in the western part of the basin during March and April when westward winds prevail over that region. Such winds (including the trade winds over the Pacific and Atlantic oceans) pile up warm surface waters in the west while exposing cold waters to the surface in the east. Isotherms along the Equator therefore slope downward to the west (Fig. 1). Thus a column of water has a higher average temperature in the west than in the east. Furthermore, if horizontal pressure gradients near the ocean floor are negligible—the absence of strong currents near the ocean floor suggests that such is indeed the case—then the sea surface must have a higher elevation in the west than in the east. (The difference in height is on the order of 1 m or 3 ft.) Hence, westward winds, by driving the surface waters westward, also maintain this slope of the sea surface and thus create an eastward pressure. Off the Equator, the subsurface eastward flow driven by the pressure force is deflected equatorward, in both hemispheres, by the Coriolis force. At the Equator, where the

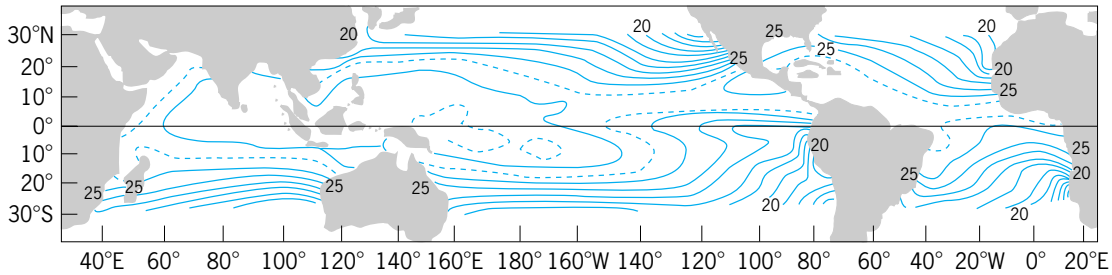

Fig. 2. Isotherms at the ocean surface. The contour interval is 1°C (1.8°F). Broken-line contours are 27°C (81°F) and 29°C (84°F).

Coriolis force vanishes, the convergent water flows in the direction of the pressure force and becomes the Equatorial Undercurrent. *See* ATLANTIC OCEAN; INDIAN OCEAN; PACIFIC OCEAN.

Numerical models that simulate the oceans are capable of reproducing the observed circulation when it is forced with the observed winds. From their results it is then possible to infer aspects of the circulation that are difficult to measure, for example, the extremely small vertical component of the currents. The vertical and meridional flows link the various eastward and westward equatorial currents (Fig. 1). A water parcel that flows eastward in the Equatorial Undercurrent gradually rises to the surface until it joins the westward South Equatorial Current, a divergent surface current in which the parcel flows poleward. If it is in the Northern Hemisphere, the parcel sinks in the neighborhood of 3°N, near the thermal front mentioned earlier. At depth, it flows back toward the Equatorial Undercurrent. *See* SIMULATION; WIND STRESS.

Time dependence of the currents. The winds in the tropics are subject to considerable seasonal and interannual fluctuations that force variations in the oceanic circulations and, in particular, in patterns of sea surface temperature. Changes in those patterns have a strong influence on the atmospheric circulation, and hence they can cause climate fluctuations such as El Niño. This circular argument—changes in sea surface temperature patterns both cause and are caused by changes in atmospheric conditions, especially the winds—implies that interactions between the ocean and atmosphere are the basis of climate variability. There is an asymmetry in the roles of the ocean and atmosphere because, whereas the atmosphere adjusts rapidly—that is, within a week or two—to a change in sea surface temperatures, the ocean has far more inertia and takes many months to adjust to a change in the winds. It is this memorylike character of the ocean that sets the time scale for the recurrence of El Niño phenomena.

The sudden onset of westerly winds over the equatorial Indian Ocean, occurring twice a year at the time of the equinoxes, provides valuable information about the memory of the oceans, and how equatorial oceans adjust to a sudden change in the winds. *See* EQUINOX.

The winds not only generate oceanic currents but also excite waves. Those waves propagate back and forth across the ocean basin and are instrumental in bringing the ocean into a state of equilibrium. The time they take to propagate across an ocean basin is a measure of the memory of the ocean—of the time it takes the ocean to be in equilibrium with the winds. These subsurface waves are most prominent along the layer that separates the warm surface waters from the cold deep water, and their speed increases with decreasing latitude. The memory of the ocean is therefore shortest near the Equator, and is shorter for a small ocean such as the Indian or Atlantic Ocean than for a larger one such as the Pacific. Thus, the seasonally varying winds force an equilibrium response in the equatorial Atlantic Ocean—the winds and currents fluctuate practically in phase—but not in the Pacific Ocean. The Pacific Ocean is so large that, even on interannual time scales, oceanic conditions at a given time are not simply determined by the forcing at that time but in part depend on the forcing at earlier times. Hence, at the height of El Niño, waves excited much earlier are still propagating around the basin. Those waves signal the termination of El Niño and set the stage for complementary conditions that prevail until waves excited during El Niño signal its return at a later date. *See* MARITIME METEOROLOGY; OCEAN CIRCULATION; TROPICAL METEOROLOGY. S. George Philander

Bibliography. Open University, *Ocean Circulation*, 2d ed., 2001; J. Pedlosky, *Ocean Circulation Theory*, 2004; S. G. Philander, *El Niño, La Niña and the Southern Oscillation*, 1990.

Equilibrium of forces

In a mechanical system the condition under which no acceleration takes place. Newtonian mechanics today is based upon two definitions which modify, but are essentially equivalent to, Newton's three fundamental laws. These definitions postulate the action of forces on particles. A particle is defined as a conceptual volume element that has mass and is sufficiently small to have point location. A body is defined as a system of particles. To develop the mechanics of a body, these definitions are applied to each of its particles and their influences summed. *See* ACCELERATION.

Newtonian particle laws. The law of motion is that, in a newtonian frame of reference, a particle of mass m acted on by resultant force $\mathbf{F}$ has acceleration $\mathbf{a}$ in accordance with the equation $\mathbf{F} = km\mathbf{a}$. Therein, k is a positive constant whose value depends upon the units in which $\mathbf{F}$, m, and $\mathbf{a}$ are measured.

The action-reaction law states that when one particle exerts force on another, the other particle exerts on the one a collinear force equal in magnitude but oppositely directed.

Equilibrium of a particle. A particle at rest or in uniform (unaccelerated) motion in a newtonian reference frame is in equilibrium. With few exceptions, a frame of reference fixed with respect to Earth is considered to be newtonian and the equation $\mathbf{F} = km\mathbf{a}$ applied. Accordingly, when a particle is in equilibrium on Earth, $\mathbf{a} = 0$ and thus $\mathbf{F} = 0$. Also, when $\mathbf{F} = 0$, $\mathbf{a} = 0$ because k and m are not zero. These considerations lead to a theorem and its corollary: The resultant of forces exerted on a particle in equilibrium is zero; and if the resultant of forces exerted on a particle is zero, the particle is in equilibrium.

Body equilibrium. A body acted upon by force is in equilibrium when its constituent particles are in equilibrium. The forces exerted on its particles (and therefore on the body) are either internal or external to the body. An internal force is one exerted by one particle on another in the same body.

An external force is one exerted on a particle or the body by a particle not of the body.

Being particle actions, internal forces obey the action-reaction law. They occur in pairs whose individual and combined resultants are zero. Hence a further theorem is that the resultant of all forces internal to a body is zero.

The resultant of forces on each particle and therefore on all particles of a body in equilibrium is zero. Because the forces are either internal or external and the resultant of internal forces is zero, it follows as a theorem that the resultant of all external forces acting upon a body in equilibrium is zero.

Equations. Forces whose resultant is zero have zero vectors of total force and total moment. Consequently, both the algebraic sum of their components in any direction and their sum of moments about any line are also zero. Specifically, relative to an *OXYZ* reference frame, the components and moments of forces externally applied to a body in equilibrium are related through the equations given below.

$$\Sigma X = 0 \qquad \Sigma M_x = 0$$
$$\Sigma Y = 0 \qquad \Sigma M_y = 0$$
$$\Sigma Z = 0 \qquad \Sigma M_z = 0$$

See RESULTANT OF FORCES.

Forces can be collinear, coplanar, parallel, concurrent, and so forth. According to the class of a particular external force system and its orientation in *OXYZ*, certain of these equilibrium equations may be either trivial (containing no terms on their left sides) or dependent (derivable by arithmetic operations on one or more previously written equations of the set). The number of equations that are independent (contain different information) is the number necessary to define the resultant of the particular class force system which acts.

Force interaction of bodies. Under certain conditions, particles of one body exert pertinent forces on particles of another. Such forces, being particle actions, obey the action-reaction law. Thus, the entire force system exerted by one body on another coexists with an equal but opposite system exerted by the other on the one. The resultants of the coexistent systems are also equal and opposite. These conditions lead to the theorem that the mutual force actions of bodies are equal and opposite.

Contact and body forces. External forces are either contact or body forces according to their manner of application.

Contact forces are those exerted between the contacting particles of bodies which touch physically. The resultant of such forces is a force, couple, or force and couple, depending partly upon the nature of the contacting surfaces. For example, rigid bodies whose contacting surfaces do not penetrate, interlock, or adhere exert resultant compressive (push) forces on one another which act at the point or within the area of contact. Additionally, if either surface is considered to be smooth (incapable of resisting or applying force tangent to the surface), the force is normal to the surfaces at the contact.

Body forces are those exerted upon certain or all particles of a body by matter which need not contact the body. Gravitational, magnetic, and electrostatic forces are examples. A body's weight is the force resultant of essentially parallel downward forces resulting from the gravitational pull of the Earth on each particle of the body. The weight force acts downward through the geometric center of a homogeneous body. *See* CENTER OF GRAVITY.

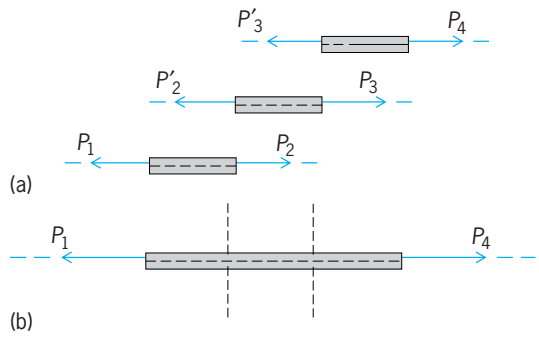

Examples of equilibrium forces. (*a*) Action forces and reaction forces acting on particles of a body. (*b*) Balanced external forces on a body.

Problem of equilibrium. To determine the forces which act upon matter at rest on Earth is the usual problem of equilibrium. The problem is statically determinate if solution is possible by employing the equations of equilibrium only. These equations relate all external forces; hence accurate solution requires that external forces be recognized and described according to their agencies of application. To summarize this information the free-body diagram is conveniently employed (see **illus.**). A free-body diagram is the sketch of a body of interest with all external forces acting thereon symbolized by arrows drawn on the sketch. *See* DYNAMICS; STATICS.

Nelson S. Fisk

Equine infectious anemia

A lentivirus-induced disease of the horse family with an almost worldwide distribution. It is characterized by recurring fever, platelet reduction, weight loss, edema, and anemia. Although death can occur, there is usually an eventual cessation of clinical signs. However, host defenses are unable to completely eliminate the virus, and the animal remains a persistently infected inapparent carrier.

Virus. The equine infectious anemia virus is in the *Lentivirus* genus of the family Retroviridae. Although it is closely related to the human immunodeficiency viruses (HIV-1, HIV-2), its genetic organization is the least complex of this group of viruses. The equine infectious anemia virus particle (approximately 115 nanometers in diameter) is composed of a lipid-bilayer membrane surrounding a cone-shaped core, which in turn encapsulates a diploid ribonucleic acid (RNA) genome. Transmembrane proteins protrude through the lipid bilayer to support surface unit proteins on the exterior of the particle. These surface unit proteins are believed to be responsible for attachment to host cells, which in the case of the equine infectious anemia virus are predominantly of

the monocyte-macrophage lineage (although some strains may also infect endothelial cell types). This function of the surface unit proteins makes them an important target for neutralizing antibodies. Following entry into the host cell, the single-stranded RNA genome is converted into a double-stranded, deoxyribonucleic acid (DNA) copy (termed the provirus) by the viral enzyme reverse transcriptase. Integrase, another viral enzyme, then inserts the proviral DNA into the DNA of the host, where it directs the synthesis of progeny viral particles.

Mode of transmission. Equine infectious anemia virus infects only members of the horse family. The mechanical transfer of blood between animals by large blood-feeding insects (mainly horse flies and deer flies) is probably the most important mechanism of natural transmission. However, the virus can be efficiently transmitted by humans if sterile techniques are not observed during veterinary procedures.

Symptoms and causes. Clinical responses to equine infectious anemia virus infection can range from an extremely severe disease resulting in death to an absence of obvious signs. This variation appears to be determined by both the virus strain and the genetic composition of the host. However, when disease occurs it is usually observed shortly after exposure to the virus (incubation periods of 10–45 days are common) and consists of fever, platelet reduction (thrombocytopenia), lethargy, loss of appetite, and petechial hemorrhages. Most horses survive this acute episode but in many cases progress to the chronic form of the disease, characterized by recurring fever, thrombocytopenia, anemia, weight loss, edema, and hemorrhage. Each fever episode is associated with massive amounts of viral replication that occurs predominantly in the mature tissue macrophages of the liver and spleen and results in the release of millions of viral particles into the bloodstream. This extensive replication triggers the release of powerful molecules called cytokines, such as tumor necrosis factor alpha (TNFα), that are normally associated with inflammatory reactions. As TNFα is both an inducer of fever and an inhibitor of platelet production, it appears that such molecules may account for many of the symptoms associated with equine infectious anemia. However, specific immune responses also contribute significantly to clinical signs such as anemia and thrombocytopenia. This is due to immune complexes that bind to receptors on erythrocytes and platelets, targeting these cells for destruction by complement or phagocytic cells.

Pathology. There are no pathological lesions unique to equine infectious anemia, although several abnormalities are typically associated with the chronic form of the disease. These abnormalities include enlargement of the spleen, liver, and lymph nodes along with hemorrhages on mucosal and serosal surfaces. At the microscopic level, immune complexes are frequently deposited in the glomeruli of the kidneys, Kupffer cells in the liver may contain stainable iron granules from the breakdown of erythrocytes, and the number of hematopoietic cells in the bone marrow may be reduced.

Immune responses and viral persistence. In most cases, immune responses eventually control the extensive viral replication associated with each fever episode and bring about a cessation of clinical signs. However, these responses require a lengthy period to fully develop and even then are incapable of eliminating the virus completely. Equine infectious anemia virus employs numerous mechanisms to avoid removal by host defenses. These mechanisms range from infecting and possibly compromising the efficiency of a key cell type in the immune system to possessing a surface unit protein with a fundamental structure that confers partial resistance to neutralizing antibodies. Another property of this protein is its ability to withstand considerable alterations in amino acid content without loss of function. This, in combination with a high mutation rate resulting from the possession of an error-prone reverse transcriptase, facilitates the emergence of variants that are completely resistant to the strain-specific neutralizing antibodies produced within the first months of infection. Although additional strain-specific neutralizing antibodies may be generated against these variants, the mutational capabilities of this virus permit still more resistant types to arise. This process explains why each clinical episode of equine infectious anemia is caused by a different antigenic variant and why termination of this cycle appears to coincide with the eventual production of neutralizing antibodies that are broadly cross-reactive instead of strain-specific. In addition to evading neutralizing antibodies, the ability to produce a constant stream of viral variants because of rapid mutation rates may also be valuable in combating cytotoxic lymphocyte responses. However, this has yet to be fully evaluated. *See* IMMUNITY.

Control. Several inactivated virus or genetically engineered viral protein vaccines have been tested, but these have been completely successful only against the strain of virus from which they were derived. In China, an attenuated live virus vaccine has been used which has reportedly induced a broader form of immunity, but only after multiple doses and a postvaccination incubation period of several months. The delay in the development of full immunity, along with the fact that this approach would confound the current protocols for the diagnosis and control of equine infectious anemia, has prevented the adoption of attenuated virus vaccines outside China.

A safe and effective vaccine that rapidly induces protection against all strains of the equine infectious anemia virus is not yet available; this virus is controlled in most areas of the world by serological testing. The most commonly used method is the agar immunodiffusion, or Coggins, test that detects the presence of antibodies against the major core protein of the virus. If a horse tests positive, further transmission of the equine infectious anemia virus can usually be prevented by segregation or quarantine. The application of these techniques resulted in more than a tenfold decrease in the reported rate of positive tests

in the United States between 1975 and 1998. *See* RETROVIRUS; VIRUS INFECTION, LATENT, PERSISTENT, SLOW.

R. Frank Cook; Charles J. Issel

Bibliography. R. F. Cook, C. J. Issel, and R. C. Montelaro, Equine infectious anaemia, in M. J. Studdert (ed.), *Virus Infections of Equines*, pp. 297–323, Elsevier Science, Amsterdam, 1996; W. Maury, Regulation of equine infectious anemia virus expression, *J. Biomed. Sci.*, 5:11–23, 1998; R. C. Montelaro, J. M. Ball, and K. E. Rushlow, Equine retroviruses, in J. A. Levy (ed.), *The Retroviridae*, vol. 2, pp. 257–360, Plenum Press, New York, 1993; D. C. Sellon, F. J. Fuller, and T. C. McGuire, The immunopathogenesis of equine infectious anaemia virus, *Virus Res.*, 32:111–138, 1994.

Equinox

One of the two places in the sky where the ecliptic crosses the celestial equator; or one of the two times of the year when the Sun crosses these points. The ecliptic is the great circle across the sky that marks the mean path of the Sun; the celestial equator is the great circle that is an extension into the sky of the Earth's mean Equator. These two great circles meet at two points, one of which is the vernal equinox and the other the autumnal equinox. The Sun passes the vernal equinox each year about March 20, and the autumnal equinox about September 22. The vernal equinox can occur as early as March 19 and as late as March 21; for most of the twenty-first century it will be on March 20. The autumnal equinox can occur as early as September 21 and as late as September 24; for most of the twenty-first century it will be on September 22. The dates and times drift with the difference between the actual solar years and 365 days, and are corrected by leap years. This results in a 4-year variation superimposed on a negative 11-min-per-year slope (**illus.** *a* and *b*). Since 2000 was an ordinary leap year in the Gregorian calendar, unlike 1900, the dates will continue to decline through 2100 (illus. *c*). *See* ASTRONOMICAL COORDINATE SYSTEMS; CALENDAR; TIME.

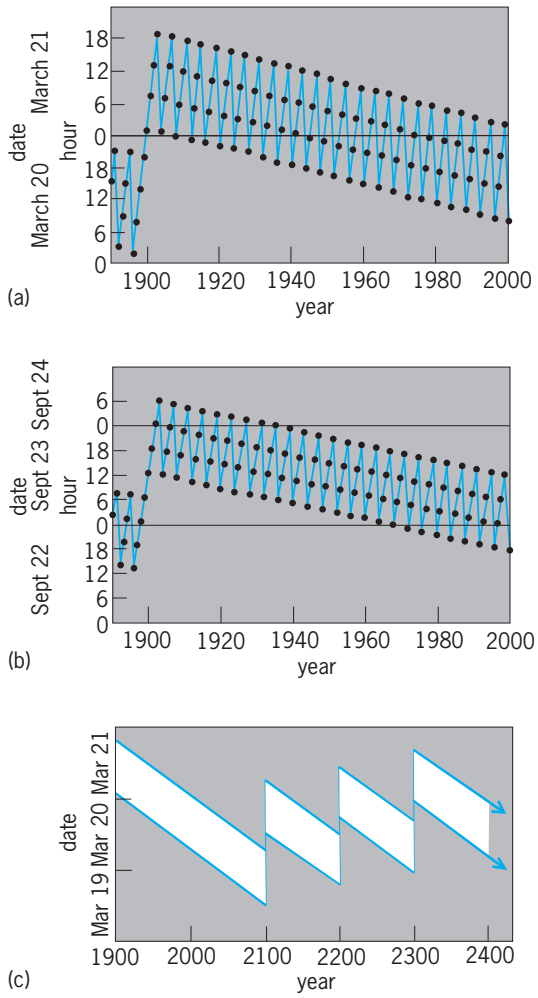

Date and time of the equinoxes (Universal Time or Greenwich mean time). (*a*) Vernal equinox, 1890–2000. (*b*) Autumnal equinox, 1890–2000. (*c*) Approximate 4-year range of times of the vernal equinox, 1900–2400. (*After R. L. Reese and G. V. Chang, The date and time of the vernal equinox: A graphical analysis of the Gregorian calendar, Amer. J. Phys., 55:848-849, 1987*)

At the vernal equinox, the Sun crosses from southern to northern declinations, marking the beginning of Northern Hemisphere spring and Southern Hemisphere autumn. At the autumnal equinox, the Sun crosses from southern to northern declinations, marking the beginning of Northern Hemisphere autumn and Southern Hemisphere spring. At the equinoxes, the Sun is directly above the Earth's Equator.

The term equinox is derived from the Latin for equal nights, indicating that the day and night are of equal duration. However, the actual duration of daylight is several minutes longer on the days of the equinoxes, and the actual dates of equal days and nights follow the autumnal equinox and precede the vernal equinox by a few days. The equinoctial dates are geometrical constructions in which the Sun is treated as a point; in actuality the top of its disk rises a few minutes ahead of its center. Furthermore, refraction in the Earth's atmosphere makes the Sun appear higher in the sky than it actually is, an effect that also lengthens daylight by several minutes. At sunrise, the top of the Sun is actually below the horizon, over which it is elevated by refraction. *See* METEOROLOGICAL OPTICS; REFRACTION OF WAVES.

Since precession of the equinoxes changes the equatorial plane over time, astrometrists distinguish between the dynamical equinox, defined by the intersection of the mean equatorial plane and the ecliptic, and the catalog equinox, which sets the zero value for the right ascensions in a star catalog for the particular epoch of that catalog. The precession, caused by the gravitational pull of the Sun and Moon on the Earth's equatorial bulge, moves the equinoxes along the ecliptic about 50 seconds of arc per year. Thus, the equinoxes occur 20 min earlier each year. *See* EARTH ROTATION AND ORBITAL MOTION; PRECESSION OF EQUINOXES.

Since the vernal equinox was in the constellation

Aries when Hipparchus studied it 2000 years ago, it is known as the first point in Aries. It is, however, in Pisces. Jay M. Pasachoff

Bibliography. R. L. Reese and G. Y. Chang, The date and time of the vernal equinox: A graphical analysis of the Gregorian calendar, *Amer. J. Phys.*, 55:848–849, 1987; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, rev. ed., 1992.

Equisetales

An order of the division Spenophyta, subkingdom Embryobionta. The Equisetales, commonly known as horsetails, is represented by a single living genus, *Equisetum*, with about 25 species found both in moist and dry habitats. These plants grow throughout the world, except in Australia and New Zealand. The plants range from herbaceous to shrubby and rarely exceed 3 or 4 ft (0.9 or 1.2 ft) in height, although some tropical species grow much taller.

The plant body is commonly composed of perennial underground stems (rhizomes) with various types of aerial stems (**Fig. 1**) in different species. Some of these are perennial, others annual; some are unbranched and reproductive, others much branched and vegetative. The bushy structure of the latter is suggestive of the common name, horsetail.

The leaves are reduced to scales which appear fused in whorls at the nodes. In the young leaf, the central region contains stomates and chloroplasts which function in photosynthesis for a short time. The burden of photosynthesis is carried on by cells in the cortex of the stem. There are no true leaf gaps in the stele. *See* LEAF.

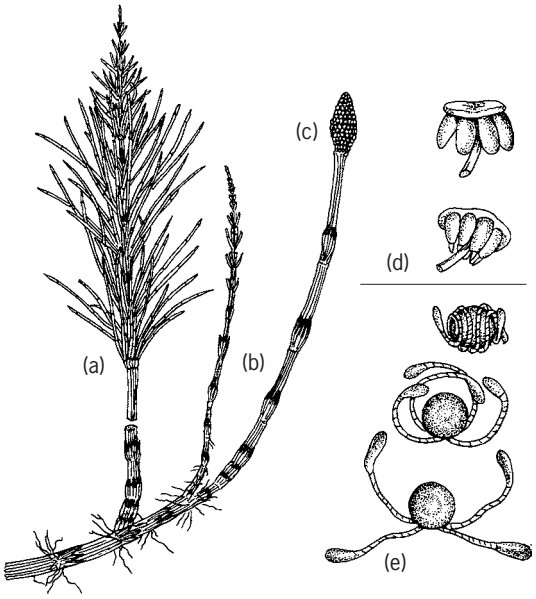

Fig. 1. ***Equisetum arvense*. (*a*) Sterile shoot. (*b*) Fertile shoot growing from an underground rootstock. (*c*) Cone. (*d*) Two views of shield-shaped sporangiophores. (*e*) Spores, greatly enlarged. As spore dries, elaters expand. (*After E. W. Sinnott and K. S. Wilson, Botany: Principles and Problems, 6th ed., McGraw-Hill, 1963*)**

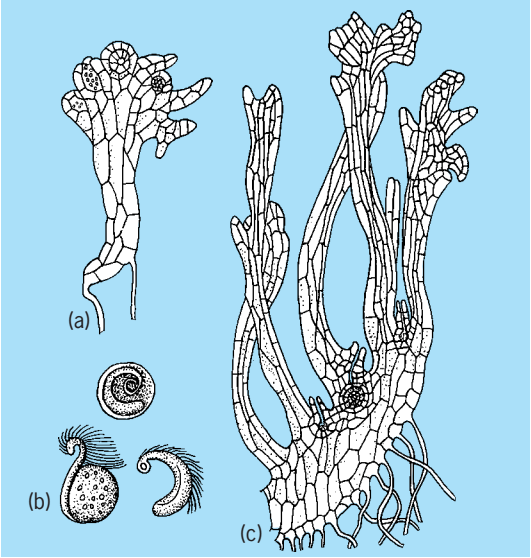

Fig. 2. ***Equisetum* gametophyte. (*a*) Small gametophyte, showing antheridia only. (*b*) Male gametes. (*c*) Larger gametophyte; branching lobes have archegonia at bases. (*After E. W. Sinnott and K. S. Wilson, Botany: Principles and Problems, 6th ed., McGraw-Hill, 1963*)**

The stem is composed of conspicuous nodes with long, strongly ribbed internodes. The cells of the stem originate in a single, pyramidal apical cell, giving rise eventually to a ring of vascular bundles with numerous openings or canals present except at the node. There is no cambium, but the basal part of each internode functions for some time as an intercalary meristem. The walls of sclerenchyma fibers in the cortex and of the epidermal cells are more or less impregnated with silica and aid in giving mechanical support to the stem. Early settlers, noting this, made use of the vegetative plants for scouring pots, pans, and floors, and as a result gave them the common name of "scouring rushes." *See* APICAL MERISTEM; CORTEX (PLANT); EPIDERMIS (PLANT); SCLERENCHYMA; STEM.

Except for the first-formed primary root, all roots are adventitious. Each adventitious root is derived from a single apical cell and develops an exarch protostele. There is no true cambium, but an intercalary meristem (region of cell division) has been observed.

The strobili (cones) are solitary and terminal, either on the main stem or on lateral branches. They are composed of a central axis with a whorled series of compact lateral branches known as sporangiophores. Sporophylls (spore-bearing leaves) are entirely suppressed. Because of the pressure resulting from their development, each sporangiophore appears compressed and hexagonal in surface view. At maturity the several sporangia ripen on the lower or abaxial portion of the sporangiophore. The spores produced are alike (homosporous). Each possesses a complex wall structure having four long, slender appendages known as elaters. These elaters are hygroscopic (coil and uncoil in response to changes in humidity) and thus aid in dissemination of spores.

Spores which reach a suitable substrate germinate immediately, producing a flat plate or mound

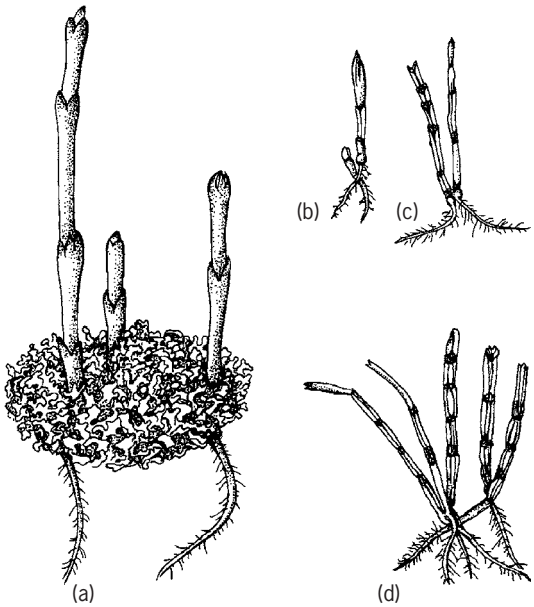

Fig. 3. Sporophytes. (*a*) *Equisetum* gametophyte with sporophytes. (*b*, *c*, *d*) The young sporophytes of *Equisetum hiemale*. (*After G. M. Smith*, *Cryptogamic Botany*, *vol. 2*, *2d ed.*, *McGraw-Hill*, *1955*)

of green tissue with numerous filamentous branches called the gametophyte (**Fig. 2**) which usually gives rise to the sex organs, antheridia (male) and archegonia (female). The motile sperms produced in the antheridia are relatively large, spirally coiled cells with numerous apical flagella. The archegonia are formed in the cushionlike portion of the gametophyte, with the chambers containing the eggs well buried in the thallus. Several eggs on the same gametophyte may be fertilized, following which each may develop an embryonic sporophyte (**Fig. 3**) without a suspensor (a chain of cells which serves to put the embryo in a favorable position in relation to its food supply). With the formation of the primary root and stem, the young sporophyte produces chlorophyll and becomes independent (self-sustaining). *See* EMBRYOBIONTA; REPRODUCTION (PLANT); ROOT (BOTANY).

Paul A. Vestal

Bibliography. L. Benson, *Plant Classification*, 2d ed., 1979; H. C. Bold, *Morphology of Plants and Fungi*, 5th ed., 1987; H. C. Bold and J. La Claire, *The Plant Kingdom*, 5th ed., 1987; G. M. Smith, *Cryptogamic Botany*, vol. 2, 2d ed., 1955.

Equivalent circuit

A representation of an actual electric circuit or electronic device by a simple circuit whose behavior is very near to that of the actual circuit over a specified range of conditions. When these conditions are satisfied, the equivalent circuit may be said to constitute a macromodel of the actual circuit. The use of equivalent circuits is important in many of the analysis and design operations associated with electronic circuits. For example, a frequently used electronic device is the operational amplifier. One of the better-known operational amplifiers, the μA 741, is realized as an integrated circuit containing over 20 transistors. Many electric circuits, for example, active-*RC* (resistance-capacitance) filters, require the use of large numbers of operational amplifiers. In such applications, the use of equivalent circuits greatly simplifies analysis and design operations. *See* ELECTRIC FILTER; INTEGRATED CIRCUITS; OPERATIONAL AMPLIFIER; TRANSISTOR.

Two of the best-known equivalent circuits are the Thévenin equivalent circuit and the Norton equivalent circuit. The Thévenin equivalent circuit consists of the series connection of a voltage source and a two-terminal circuit (**Fig. 1***a*); the Norton equivalent circuit consists of the parallel connection of a current source and a two-terminal circuit (Fig. 1*b*). In both circuits, the output of the voltage or current source is usually dependent on the electric signals applied to the input terminals. *See* NETWORK THEORY; THÉVENIN'S THEOREM (ELECTRIC NETWORKS).

Two common applications of equivalent circuits are the modeling of large circuits (such as the operational amplifier mentioned above), and the modeling of individual electronic solid-state devices such as transistors. An example of an equivalent circuit for an operational amplifier is shown in **Fig. 2***a*. The Thévenin equivalent circuit is used. The gain constant K is usually very large (of the order of 10^5). The output resistor R_{out} is usually very small and frequently is omitted. Solid-state devices such as transistors are inherently nonlinear in their behavior. For such nonlinear components, the use of equivalent circuits is usually restricted to modeling the actual circuit behavior under the conditions that the circuit variables have such small variations that the nonlinear relations may be approximated by idealized ones. In such a scenario, the equivalent circuits are usually referred to as small-signal equivalent circuits. An example of such a device is a first-order

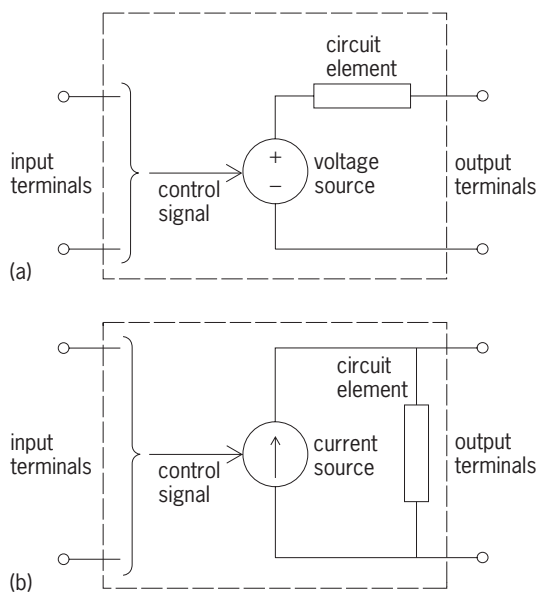

Fig. 1. Basic equivalent circuits. (*a*) Thévenin equivalent circuit. (*b*) Norton equivalent circuit.

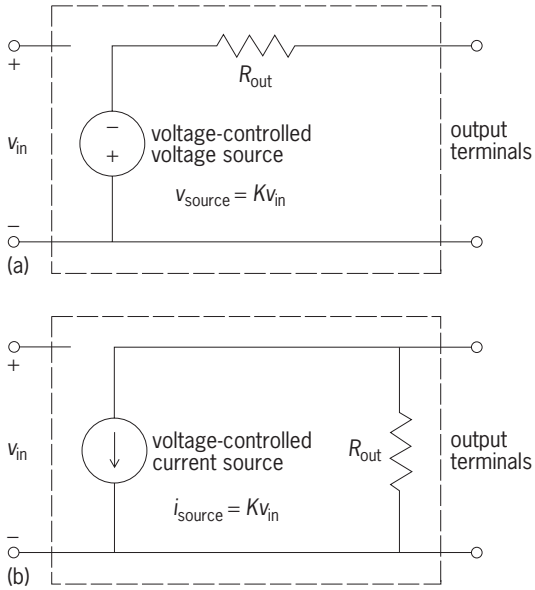

Fig. 2. Equivalent circuit applications. (*a*) Operational amplifier equivalent circuit applications. (*b*) MOSFET equivalent circuit.

small-signal model for a metal-oxide-semiconductor field-effect transistor (MOSFET; Fig. 2*b*). The Norton equivalent circuit is used. The output resistance R_{out} is usually very large, and frequently is omitted from the equivalent circuit. *See* CIRCUIT (ELECTRICITY); CIRCUIT (ELECTRONICS). Lawrence P. Huelsman

Bibliography. L. P. Huelsman, *Basic Circuit Theory*, 3d ed., 1991; D. E. Johnson, J. R. Johnson, and J. L. Hilburn, *Electric Circuit Analysis*, 1989; J. W. Nilsson and S. A. Riedel, *Electric Circuits*, 7th ed., Prentice Hall, 2004.

Equivalent weight

The number of parts by weight of an element or compound which will combine with or replace, directly or indirectly, 1.008 parts by weight of hydrogen, 8.00 parts of oxygen, or the equilvalent weight of any other element or compound. The term equivalent weight comes from the law of equivalent proportions, which states that the weights of two elements A and B which combine separately with identical weights of another element C are either the weights in which A and B combine together, or are related to them in the ratio of small whole numbers. A standard weight of 8.000 parts is chosen for oxygen. For all elements, the atomic weight is equal to the equivalent weight times a small whole number, called the valence of the element. *See* ATOMIC MASS; COMPOUND (CHEMISTRY); ELEMENT (CHEMISTRY); VALENCE.

An element can have more than one valence and therefore more than one equivalent weight. The use of the terms is explained below.

1. Ammonia, NH_3, contains 1 atom of nitrogen combined with 3 atoms of hydrogen. Since the equivalent weight of hydrogen is equal to its atomic weight, the equivalent weight of nitrogen is one-third its atomic weight and its valence is 3.

2. Magnesium oxide, MgO, contains 1 atom of magnesium combined with 1 atom of oxygen. Since the equivalent weight of oxygen is one-half its atomic weight, the equivalent weight of magnesium is also one-half its atomic weight and its valence is 2.

3. Phosphorus forms phosphorus trichloride, PCl_3, and phosphorus pentachloride, PCl_5. Since the equivalent weight of chlorine is equal to its atomic weight, in the trichloride the equivalent weight of phosphorus is one-third its atomic weight and its valence is 3, and in the pentachloride the equivalent weight is one-fifth its atomic weight and its valence is 5.

The equivalent weight of a compound depends on the reaction in which it takes part. Thus:

1. In the reaction between potassium iodate, KIO_3, and silver nitrate, $AgNO_3$, one molecule of silver iodate, $AgIO_3$, is precipitated for every molecule of silver nitrate. This is represented by reaction (1).

$$KIO_3 + AgNO_3 \rightarrow AgIO_3(\text{ppt}) + KNO_3 \qquad (1)$$

Since the equivalent weight of silver is its atomic weight, the equivalent weight of potassium iodate, in this reaction, is its molecular weight.

2. When potassium iodate, KIO_3, is reduced to iodine, I_2, by potassium iodide, KI, three molecules of iodine are produced per molecule of potassium iodate. This is represented by reaction (2). Since the

$$KIO_3 + 5KI + 6HCl \rightarrow 3I_2 + 3H_2O + 6KCl \qquad (2)$$

equivalent weight of iodine is one-half its molecular weight, the equivalent weight of potassium iodate, in this reaction, is one-sixth its molecular weight.

This concept, together with that of gram-equivalent weight, tends to have been abandoned, and relations are expressed in terms of balanced stoichiometric chemical equations and relative numbers of moles reacting. *See* ELECTROCHEMICAL EQUIVALENT; MOLE (CHEMISTRY); OXIDATION-REDUCTION; STOICHIOMETRY. Thomas C. Waddington

Erbium

A chemical element, Er, atomic number 68, atomic weight 167.26, belonging to the rare-earth group. The naturally occurring element is made up of the six

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

stable isotopes. The rose-pink oxide, Er_2O_3, dissolves in mineral acids to give rose-colored solutions. The salts are paramagnetic and the ions are trivalent. At low temperatures the metal is antiferromagnetic and at still lower temperatures becomes strongly ferromagnetic. For properties of the metal *see* RARE-EARTH ELEMENTS

Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

Ergosterol

A steroid, with the structure shown below, that be-

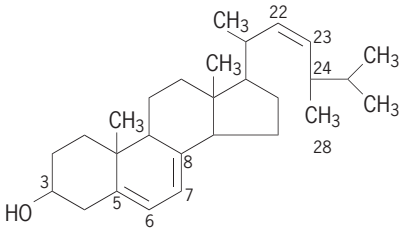

longs to the class of unsaponifiable lipids. It was first isolated from ergot bodies (black or dark-purple club-shaped bodies of hardened mycelium from the ergot fungus, which replace the seeds of various grasses, such as rye and wheat). Ergosterol is a white crystalline compound, insoluble in water and soluble in organic solvents. It is derived from the cyclopentanoperhydrophenanthrene skeleton:

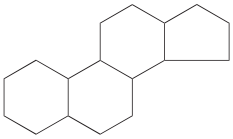

See ERGOT AND ERGOTISM.

The hydroxyl group at carbon atom 3 is of the β configuration. Ergosterol differs from cholesterol by having three positions of unsaturation, at carbon atoms 5-6, 7-8, and 22-23, and containing a methyl group (carbon atom 28) substituted for a hydrogen atom at carbon atom 24. *See* CHOLESTEROL.

Ergosterol makes up to 90–100% of the steroids in certain yeasts and can be extracted from these yeasts in commercial quantities. Ultraviolet irradiation of ergosterol leads to the formation of calciferol (vitamin D_2) by way of a series of irreversible reactions:

$$\text{Ergosterol} \rightarrow \text{lumisterol} \rightarrow \text{tachysterol} \rightarrow \text{calciferol}$$

Under controlled conditions, a more than 50% yield of vitamin D_2 can be obtained. The process is used commercially. *See* VITAMIN D.

The biosynthetic pathway of ergosterol in yeast is probably similar to that of cholesterol in animal tissues, since both acetate and squalene are converted to ergosterol by yeast:

$$\text{Acetate} \rightarrow \text{squalene} \xrightarrow{\text{methyl}} \text{ergosterol}$$

The amino acid methionine serves as the donor of the methyl group (carbon atom 28) of ergosterol. This reaction probably takes place after the steroid nucleus has been formed. *See* STEROID.

Willem J. Van Wagtendonk

Bibliography. H. Eriksson and J. A. Gustafsson, *Steroid Hormone Receptors: Structure and Function*, 1984; B. Green and R. E. Leake (eds.), *Steroid Hormones: A Practical Approach*, 1987; H. L. Makin, *Biochemistry of Steroid Hormones*, 2d ed., 1984.

Ergot and ergotism

Ergot is the seedlike body of fungi (molds) of the genus *Claviceps*; ergotism is a complex disease of humans and certain domestic animals caused by ingestion of grains and cereals infested with ergot. Ingestion of these long, hard, purplish-black structures called sclerotia (**illus.** *a*) may lead to convulsions, abortion, hallucinations, or death. During the Middle Ages, hundreds of thousands of people are believed to have died from this disease, often referred to as holy fire, St. Anthony's fire, or St. Vitus's dance. Epidemics in humans, although less prevalent in modern times, last occurred in 1951, and the potential danger is always present, as shown by annual livestock losses due to ergot poisoning. *See* POISON.

Fungus. Ergot of rye derives its name from *l'ergot*, the French word for cockspur, owing to the similarity in size and shape. Later, the word ergot was applied to all sclerotia of *Claviceps*, even though some were of globose or irregular shape. There are 32 recognized species, most of which infect members of the grass family. Only three species are parasitic on the rushes and sedges. *See* PLANT PATHOLOGY.

The other stages of the ergot fungus are less conspicuous and often go unnoticed. After the sclerotia overwinter or persist through drought, as is the case in some desert and tropical species, they germinate under favorable, moist conditions and give rise to small mushroomlike fruiting bodies (illus. *b*). The spiny spherical heads of the fruiting bodies contain

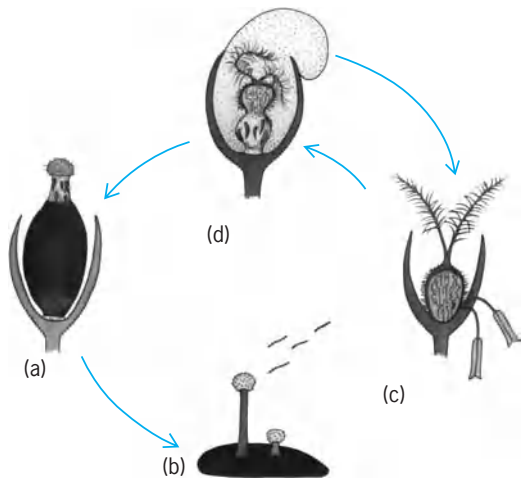

Ergot life cycle. Parts *a–d* explained in text.

long filamentous ascospores (sexual spores) which are ejected and blown about by the wind. Ascospores landing on a healthy stigma of a grass flower stick and enter, as does pollen, the ovary in the form of a fine mycelial network (illus. *c*). Soon after infection, a special convoluted structure called a sphacelium develops between the ovary and the base of the floral cavity, at which point further invasion of host tissue is restricted (illus. *d*). The sphacelium produces hundreds of thousands of conidia (asexual spores) embedded in a sugary matrix called honeydew. This sticky, mildly fetid substance attracts insects which feed on the sugars. When the insects move to other flowers, the conidia adhering to the legs and mouthparts infect healthy ovaries of other plants (illus. *c*). After several weeks the sclerotium develops between the sphacelium and the flower base, completing the cycle (illus. *a*). *See* ASCOMYCOTA.

Both honeydew and sphacelium are nontoxic and do not contain the poisonous alkaloids responsible for ergotism. The honeydew itself, composed almost entirely of sugars, is quite nutritious. The major sugars are glucose, fructose, sucrose, and fructose-containing oligosaccharides; in some species sugar alcohols are the major honeydew constituents. Honeydew sugars are known to be products of the parasite rather than products of the host, as was previously believed. By converting host sucrose into some of the above-mentioned sugars, the sphacelium creates a sucrose deficit in the floral cavity. The resulting flow of host sucrose is continuous, since fungal enzyme activity will not allow a sucrose equilibrium to exist.

Sclerotia have an unusual chemical makeup. They carry only 10% water by weight, and 50% of the dry weight is composed of fatty acids, sugars, and sugar alcohols, which make the ergot a storehouse of energy. Unfortunately, they also contain the poisonous alkaloids, ranging from 0 to 1.2% of the dry weight.

Disease. There are three types of ergotism (gangrenous, convulsive, and hallucinogenic). Their symptoms often overlap; the hallucinogenic form is usually observed in combination with one of the other two. The unusual combination of gangrenous and convulsive symptoms is sometimes observed in the Balkans and areas near the Rhine River.

The onset of gangrenous ergotism is generally characterized by lassitude, nausea, and pains in the limbs followed by alternating sensations of intense heat and cold. As bodily extremities become numb, livid watery vesicles may appear on the affected parts (usually arms and legs). Finally, the diseased area turns black, dry, and becomes mummified.

In convulsive ergotism, various parts of the body become grossly deformed as a result of clonic or tonic convulsions, or both. Generalized neurological stimulation causes epileptiform seizures, whereas specific stimulation might involve ravenous hunger, violent retching, tongue biting, or abnormal breathing patterns. This form involves a longer recovery period and often results in permanent nerve damage and subsequent sensitization.

The hallucinogenic form often includes symptoms of one of the other types. In its more pure form, it is referred to as choreomania, St. Vitus's dance, or St. John's dance. Vivid hallucinations are accompanied by psychic intoxication reminiscent of the effects of many of the modern psychedelic drugs (nervousness, physical and mental excitement, insomnia, and disorientation). Early reports state that the disease usually manifested itself in the form of strange public dances that might last for days or weeks on end. Dancers made stiff jerky movements accompanied by wild hopping, leaping, screaming, and shouting. They were often heard conversing with devils or gods, and danced compulsively, as if possessed, until exhaustion caused them to fall unconscious or to lie twitching on the ground. High mortality rates were associated with severe epidemics involving any of the three forms of ergotism.

History. The earliest undisputed occurrence of ergotism was documented in Germany in A.D. 857, about 700 years before ergot was described in any of the medieval herbals. However, the most serious epidemics were recorded between the late 900s and the 1800s. Although both gangrenous ergotism (mainly in western Europe) and hallucinogenic ergotism were reported throughout this period, the convulsive form was not described until the late 1400s. The epidemics came in times of famine, usually during the years of heavy rain following severe winters. Under these conditions rye became heavily infected with ergot.

Convulsive ergotism is little understood even today. It occurred mainly in northern Europe east of the Rhine, with the greatest number of fatalities among children and weaned babies. However, it did not affect nursing babies, even those with severely afflicted mothers. The disease was much less prevalent on farms and dairies or among individuals consuming bacon, eggs, or milk (or other vitamin A–rich food). These and clinical observations have led to the hypothesis that ergot intake combined with avitaminosis A was the cause of convulsive ergotism. This does not, however, explain the absence of the gangrenous form in eastern Europe.

Ergot was known to cause ergotism in France as early as the late 1500s. After this period, with the cleaning of grain and the increased use of wheat (wheat is a poor host for ergot), the incidence of ergotism fell rapidly in the West. In the East, ergot was not accepted as the cause of ergotism until the terrible 1771 epidemic; continued resistance to such ideas by many peasants caused widespread epidemics even into the twentieth century. Once ergot was known to cause the disease, many governments set safety limits on ergot ranging in grain from 0.1 to 0.3%. A level of 8 to 10% ergot in grain is considered to be fatal. In modern times ergotism has occurred seldom, and then usually from overdoses of the drug in medicine or in attempts to procure abortion.

The success with which the disease is controlled in humans has been brought about by (1) agricultural inspection, (2) use of wheat, potatoes, and maize instead of rye, (3) limited control of ergot, (4) reserves

of sound grain, and (5) forecasting severe ergot years. The most recent and best-recorded epidemic was in southern France in 1951 when an unscrupulous miller used moldy grain to make flour.

Ergot in medicine. Although the first description of the use of ergot in childbirth appeared in 1582, earlier reports indicate its use by midwives far back into European and Chinese history to stimulate uterine contractions and to control uterine bleeding. Its formal introduction into medicine did not come until the early 1800s. After commercial production began, powdered ergot preparations were found to be variable and unstable. Later in the century, ergot was successfully used to treat migraine, but this effect, too, was variable.

In the early twentieth century two ergot alkaloids (ergotoxine and ergotamine) were isolated. Unfortunately, they caused significant side effects and were not as specific or active as some of the crude aqueous preparations. Shortly after its discovery, ergotamine was found to be effective in the treatment of migraines. Both ergotoxine and ergotamine cause vasoconstriction that can lead to gangrene with chronic use. *See* HEADACHE.

In 1935 a new water-soluble ergot alkaloid, ergonovine, was synthesized. It is a small lysergic acid derivative, whereas ergotamine is a large, tetracyclic peptide linked to lysergic acid. Ergonovine is used to facilitate childbirth by stimulating uterine contractions. Ergotoxine was found to be a mixture of similar large alkaloids (ergocorine, ergokryptine, and ergocristine). Many other important lysergic acid derivatives have been produced by means of semisynthesis. Of these derivatives, LSD-25 (*d*-lysergic acid diethylamide) is the most famous. LSD has been used experimentally, mainly in psychiatry and neurophysiology. *See* PSYCHOTOMIMETIC DRUGS.

Through extensive research, many other uses for ergot alkaloids has been found. Ergotamine and dihydroergotamine are used to treat migraines, and methysergine is used in migraine prophylaxis. Dihydroergotoxine is prescribed for hypertension, cerebral diffuse sclerosis, and peripheral vascular disorders. Ergocorine and the less toxic agroclavine have been reported as unusual experimental birth-control agents. The drugs appear to inhibit implantation of the ovum. Several semisynthetic alkaloids are also active implantation inhibitors.

Alkaloid production. For more than a century, scientists have tried to produce sclerotia in pure culture, tissue culture, and sophisticated extractions of host tissues, but only the mycelial stage has been grown. Neither the sphacelia nor the sclerotia can be cultured away from the host. These failures have been due to a lack of understanding about the physiological requirements that are met during parasitism.

In 1951 the first ergot alkaloids were found in pure liquid cultures of mycelial mutants. The products were clavine alkaloids, which were interesting but of no medical value (except as possible birth-control agents). These compounds are ergoline derivatives (lysergic acid minus carbonyl group) and are not easily converted to lysergic acid derivatives. High yields of the more valuable alkaloids have been produced in liquid culture by new mycelial strains. Such production is not without its problems. Most important, the strains are unstable and may revert to nonalkaloid production with subsequent culture. However, bromocryptine and other ergot alkaloids have been used to alleviate some of the symptoms of Parkinson's disease. *See* ALKALOID; PARKINSON'S DISEASE.

R. L. Mower

Bibliography. E. L. Backman, *Religious Dances in the Christian Church and in Popular Medicine*, 1952; G. Barger, *Ergot and Ergotism*, 1931; F. J. Bové, *The Story of Ergot*, 1970; M. Goldstein et al. (eds.), *Ergot Compounds and Brain Function: Neuroendocrine and Neuropsychiatric Aspects*, 1980.

Ericales

An order of flowering plants, division Magnoliophyta (angiosperms), in the large asterid assemblage (often Asteridae in previous systems of classification). Nearly all of the 24 families assigned to the order have previously been considered members of several orders in the subclass Dilleniidae. Ericales are a diverse group that have general asterid characters: tenuinucellate ovules; flowers with fused sepals, and often, fused petals with the anthers fused at least basally to the petals, even when the petals are apparently free; cellular endosperm formation; and tegumentary tapeta.

The largest families are Ericaceae (1350 species), Primulaceae (1000 species), Myrsinaceae (1000 species), Sapotaceae (800 species), and Balsaminaceae (600 species). The two biggest families are largely temperate herbs (Primulaceae), and the next three are largely tropical trees or herbs (Balsaminaceae). A relationship between the first four families has been known for many years, but the last has nearly always been considered related to the rosid family Geraniaceae.

Familiar plants belonging to Ericales include rhododendrons (*Rhododendron*, Ericaceae), camellias (*Camellia*, also the genus of tea; Theaceae), primroses (*Primula*, Primulaceae), phlox (*Phlox*, Polemoniaceae), and impatiens (*Impatiens*, Balsaminaceae). *See* BLUEBERRY; DILLENIIDAE; MAGNOLIOPSIDA; PLANT KINGDOM.

Mark W. Chase

Eriocaulales

An order of flowering plants, division Magnoliophyta (Angiospermae), subclass Commelinidae of the class Liliopsida (monocotyledons). The order consists of the single family Eriocaulaceae, with about 1200 species. The Eriocaulales are Commelinidae with a reduced (or no) perianth and with unisexual flowers aggregated into a dense, involucrate head that is elevated above the clustered, basal leaves on a long peduncle. Although the individual

flowers are small and inconspicuous, the heads are more or less showy and pollination is usually by insects, in spite of the absence of nectar and nectaries. The perianth typically consists of two series of similar, white-hyaline tepals, three members in each series. Each locule of the ovary has a single, pendulous, orthotropous, tenuinucellate ovule. The order is of negligible economic importance. *See* COMMELINIDAE; LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM. Arthur Cronquist

Erosion

Erosion is the result of processes that entrain and transport earth materials along coastlines, in streams, and on hillslopes. Wind and water are common agents through which forces are applied to resistant rocks, soils, or other unconsolidated materials. Erosion types often are designated on the basis of the agent: wind erosion, fluvial erosion, and glacial erosion. Fluvial erosion usually has been regarded as the most effective type in shaping the land surface during recent geologic time. Under certain environmental conditions, however, wind erosion moves considerable quantities of earth materials, as demonstrated during the "dust bowl" years in the United States. Glacial erosion shaped much of the land surface during the Quaternary Period of geologic time. Each type of erosion produces distinctive landforms, contributing to the diversity of terrestrial landscapes. *See* DESERT EROSION FEATURES; EOLIAN LANDFORMS; GEOMORPHOLOGY; GLACIOLOGY; MASS WASTING; QUATERNARY; STREAM TRANSPORT AND DEPOSITION.

Many terms are used in association with erosion processes, although not with consistency. Erosion involves the entrainment and transportation of earth materials from source areas to deposition sites. Soil loss refers to material actually transported from a particular hillslope or field, and may be less in volume than erosion due to on-site deposition in microtopographic depressions. Sediment is the product of erosion, and sediment yield is the amount of sediment transported from a watershed. Other terms are found in the literature that differ somewhat in connotation within particular contexts and for particular scientific disciplines. *See* SEDIMENTOLOGY.

Erosion includes various processes, depending upon perspective, ranging from in-situ rock weathering, mass-movement, sheetwash on hillslopes, channel bed and bank scour, work of fauna, and deposition. Most scientists regard weathering, mass-movement, and deposition as separate processes and faunal erosion to be of minor importance except in special circumstances. *See* WEATHERING PROCESSES.

Forces exerted by erosion processes must exceed resistances of earth materials for entrainment and transportation to occur. Environmental conditions determine the magnitude of the forces, the resistances, and the relations among them. Erosion rates are highly variable in time and space due to changing relations between forces and resistances. The major factors governing wind-erosion rates are wind velocity, topography, surface roughness, soil properties and soil moisture, vegetation cover, and land use. The major factors governing fluvial-erosion rates on hillslopes are rainfall energy, topography, soil properties, vegetation cover, and land use. The major factors governing fluvial-erosion rates in stream channels are depth and velocity of water flow, together with the size and cohesiveness of the bed and bank materials. The major factors governing glacial-erosion rates are the depth and velocity of ice flow, together with the hardness of the bed and side-wall materials.

During the twentieth century, public and research interest focused on fluvial erosion. The erosion rate under natural, undisturbed conditions is known as geologic, normal, or natural erosion. The erosion rate following land disturbance, usually due to human activities such as agriculture, mining, or construction, is known as accelerated or anthropic erosion. Human activities are capable of changing erosion rates by two or three orders of magnitude. Conversely, soil-conservation practices can reduce erosion rates to approximate the natural rate if the properly designed, applied, and maintained. Effective erosion control is based upon either decreases in forces impinging on surfaces or increases in the resistances of surface materials. *See* SOIL CONSERVATION.

Accelerated erosion by fluvial processes may be the most important environmental problem worldwide because of its spatial and temporal ubiquity. Erosion rates commonly exceed soil-formation rates, causing depletion of soil resources. The effects of erosion are insidious due to the removal of the fertile topsoil horizon, compromising food production. Sediment frequently is transported well beyond the source area to degrade water quality in streams and lakes, harm aquatic life, reduce the water-storage capacity of reservoirs, and increase channel-maintenance costs. *See* SOIL. T. J. Toy

Bibliography. R. P. C. Morgan, *Soil Erosion and Conservation*, 3d ed., 2005; T. J. Toy, Accelerated erosion, process, problems, and prognosis, *Geology*, 10:524–529, 1982; T. J. Toy et al., *Soil Erosion: Processes, Prediction, Measurement, and Control*, 2002; F. R. Troeh, J. A. Hobbs, and R. L. Donahue, *Soil and Water Conservation*, 4th ed., 2003.

Errantia

A group of 34 families of Polychaeta in which the anterior, or cephalic, region is more or less fully exposed and the body is often long, linear to short, and depressed. The segments of these worms are similar or change gradually. The pharynx is often heavily muscularized and eversible, and its inner walls are fortified with calcified or chitinized plates or jaws. Some families are benthic throughout their lives, others are entirely pelagic, and some have pelagic larval and reproductive stages and benthic trophic development. Errantia occur in all seas, at all depths, and in inland seas or lakes. They range in length from a

few millimeters to 6 ft (2 m). *See* POLYCHAETA.

Scale bearers and allies. There are 6 scale-bearing and 3 allied families included in this large artificial category.

Aphroditidae. With 9 genera and about 85 species, they include the sea mouse (*Aphrodita*) (**Fig. 1**), the giant cold-water *Laetmonice*, and several other deep-water genera.

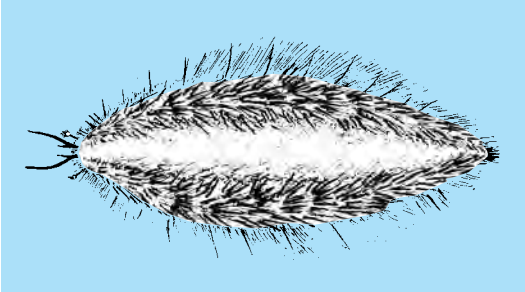

Fig. 1. ***Aphrodita*, of the Aphroditidae, dorsal view.**

Polynoidae. This is the largest of polychaete families, with about 75 genera and 600 species. The body may be short or long and large or minute; it is covered with elytra (**Fig. 2**). Representatives of this family are in all seas and at all depths. Most are benthic, some are pelagic (*Drieschia*, *Nectochaeta*), and others are commensal (*Arctonoe*, *Hesperonoe*, *Polynoe*); some are known only from abyssal depths (*Macellicephala*). Common genera are *Harmothoe*, *Lepidonotus*, and *Halosydna*.

Polyodontidae. The 8 or 9 genera and about 36 species are tubicolous and often large-bodied, with characteristic cephalic and parapodial structures. Some have spinning parapodial glands, which secrete the fibers to construct tubes. They are worldwide, occurring chiefly in tropical latitudes. Typical genera are *Euphanthalis*, *Panthalis*, and *Polyodontes*.

Sigalionidae. Known through 10 genera and about 130 species, they are sometimes large and long and seldom short and small (*Pholoe*). The best-known genera are *Leanira*, *Psammolyce*, *Sigalion*, *Sthenelais*, and *Thalenessa*. They are worldwide and occur in all latitudes, from littoral to abyssal depths.

Other families. The other 5 families are small, each with distinctive characteristics: The Peisidicidae have 1 genus and 3 species; the Pareulepidae, 1 genus and 8 species; the Pisionidae, 5 genera and 12 species; the Chrysopetalidae, 4 genera and 18 species; and the Palmyridae, 1 genus and 2 species.

Amphinomorpha families. There are 3 families included in this group. The Amphinomidae, with 18 genera and about 120 species, comprise the stinging or fire worms; a burning sensation results from handling them. Common genera are the large, brilliantly colored *Amphinome* and *Hermodice*—the "scorpions of the sea"—which may be 10–16 in. (25–40 cm) long and as thick as a finger. Most have a characteristic caruncle. They abound in circumtropical intertidal zones among corals and corallines. The chalky white, very numerous bristles can penetrate

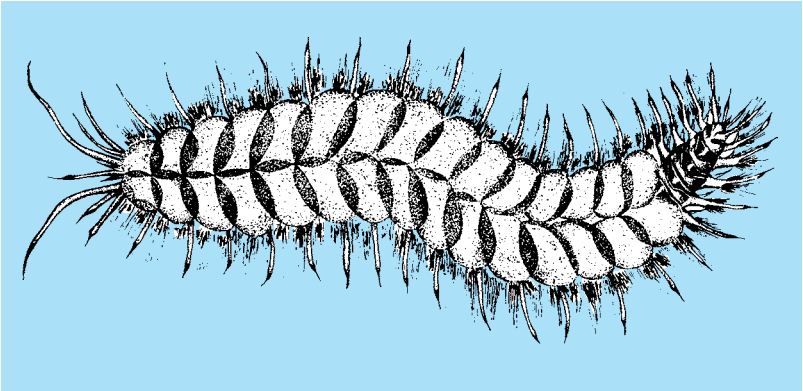

Fig. 2. ***Harmothoe*, of the Polynoidae, dorsal view.**

human skin and cause discomfort which may last for hours. The Euphrosinidae (**Fig. 3**) and the Spintheridae (**Fig. 4**) are smaller; the first family has 2 genera and 45 species and the second 1 genus and 8 species. They are short and depressed and are associated with sponges and other colonial animals.

Leaf-bearing and pelagic families. The leaf-bearing and pelagic groups are composed of 8 families. The Phyllodocidae, with about 30 genera and more than 240 species, are often brilliantly iridescent and are highly motile, with arresting colors and ornamentation. Most are long and linear; others are short

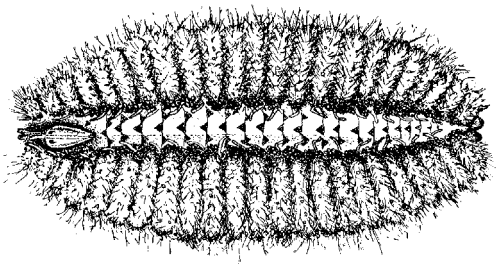

Fig. 3. ***Euphrosine*, of the Euphrosinidae, dorsal view.**

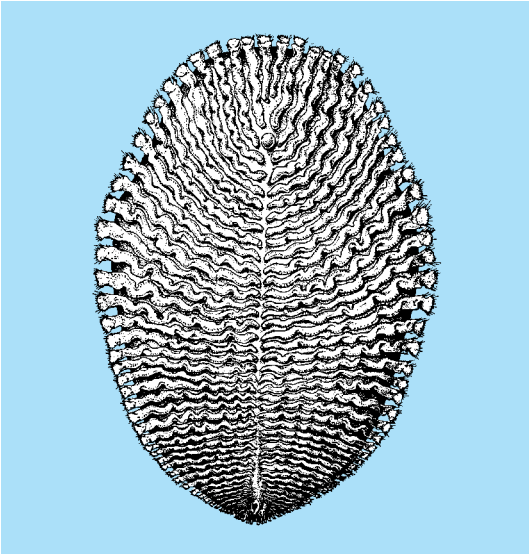

Fig. 4. ***Spinther*, of the Spintheridae, dorsal view.**

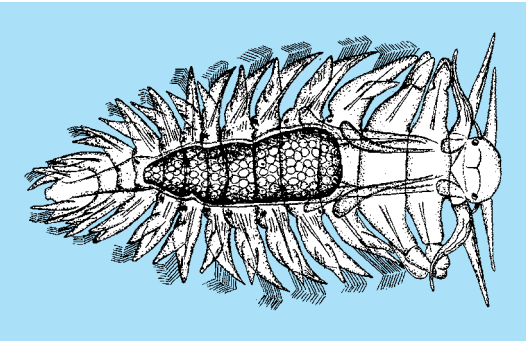

Fig. 5. *Lopadorrhynchus*, of the Lopadorrhynchidae, dorsal view.

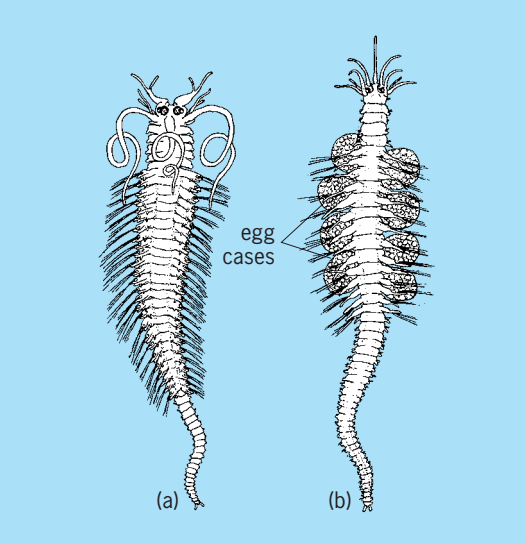

Fig. 6. *Procerastea* of the Syllidae (Autolytinae). (*a*) Epitokous male, dorsal view. (*b*) Epitokous female with attached egg cases, dorsal view.

and tumid. They occur in all seas and at all depths. Common genera are *Anaitides*, *Eulalia*, *Eteone*, and *Phyllodoce*.

The pelagic families allied to the Phyllodocidae are the Alciopidae, with 10 genera and 41 species; the Lopadorrhynchidae (**Fig. 5**), with 2 genera and 5 species; the Iospilidae, with 4 genera and 7 species; the Pontodoridae, with a single genus and species; and the Typhloscolecidae, with 3 genera and 14 species. The benthic Lacydonidae have 3 genera and 6 species, and the aberrant pelagic Tomopteridae, or glass worms, 2 genera and about 40 species. The development of *Tomopteris* demonstrates that its affinities with other polychaetes are doubtful, the resemblances being a result of evolutionary convergence.

Hesionidae and Pilargidae. The Hesionidae have 22 genera and more than 82 species, and the Pilargidae, 6 genera and about 25 species. They differ from other families chiefly in cephalic parts. Most are small, short, and depressed; a few are long and linear. Some are free-living; others commensal. Because of their small size, they escape collection unless special efforts are exerted. The best-known genera are *Hesionie*, *Leocrates*, *Oxydromus*, and *Podarke* in the Hesionidae and *Ancistrosyllis* and *Pilargis* in the Pilargidae.

Syllidae. These worms comprise the second largest of polychaete families; they have 55–60 genera and more than 400 species. Most occur along intertidal rocky shores among epiphytic plants or animals or in encrusting sponges. They are chiefly foraging predators or scavengers, identified by their long, linear, translucent bodies with articulated cirri. Some are very minute and are often found bearing eggs or juveniles on body segments. The largest syllid, *Trypanosyllis*, found in Antarctic seas, may be 4 in. (100 mm) long. Most syllids are a few millimeters long.

The Syllidae are easily divisible into four subfamilies: the Autolytinae, with sinuous pharnyx and without ventral cirri (*Autolytus* and *Myrianida*); the Eusyllinae, with a thick body and nonarticulated cirri (*Eusyllis*, *Pionosyllis*, and the luminescent *Odontosyllis*); the Syllinae, with many genera and species (*Typosyllis*, *Haplosyllis*, *Syllis*, and *Trypanosyllis*); and the *Exogoninae*, with a short, small body of few segments.

Reproduction is extremely diversified, both sexual and asexual, and sexes are sometimes dimorphic (**Fig. 6**). Schizogamy or reproduction by fission results in individuals with heads but no pharyngeal structures, as in *Chaetosyllis*, *Ioda*, *Tetraglene*, *Polybostrichus*, and *Sacconereis*.

Nereidae. A family known for about 27 genera and more than 360 species, they include some of the largest and best-known Errantia and are frequently used as laboratory animals. Most are marine, occurring in worldwide seas and at all depths; a few are euryhaline, entering tidal streams or inland seas, or occur in thermal springs and high mountain lakes. They have been studied for their neurosecretory cells in reproductive processes. Reproduction is usually sexual, with individuals similar or heteromorphid as epitokous adults. Development proceeds through pelagic, planktonic larvae; through lecithotropic (yolky) ova without pelagic

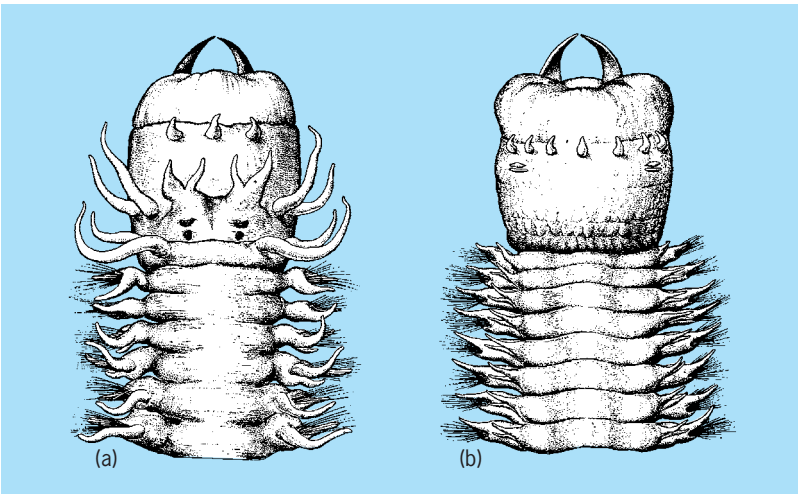

Fig. 7. *Ceratocephale*, of the Nereidae, with everted pharynx. (*a*) Dorsal. (*b*) Ventral.

larva; or by viviparity, in which juveniles are released from the female body. Common genera are *Nereis*, *Neanthes*, *Ceratocephale* (**Fig. 7**), *Platynereis*, and *Perinereis*. The subfamily Namanereinae occurs largely in fresh water. *Tylorrhynchus heterochaetus* is a pest in rice paddies in southeastern Asia.

Nephtyidae. With 4 genera and about 85 species, they are sometimes called sandworms for their occurrence in intertidal sands, where their highly opalescent colors and rapid movements render them almost invisible. The eversible pharynx is distinctive (**Fig. 8**). They occur in worldwide seas at all depths, and a few are euryhaline. Common genera are *Nephtys* and *Aglaophamus*.

Sphaerodoridae. This family has 2 genera and about 18 species; all are small (up to 1.4 in. or 35 mm long), short- to long-bodies, and usually papillated; usually their occurrence is rare.

Glyceridae. Known from 3 genera and about 66 species the Glyceridae are characterized by the enormous eversible proboscis used in food capture and forward progression. The largest genus, *Glycera*, has about 58 species and occurs in the shallowest to greatest ocean depths throughout the world. *Glycerella* is a polar genus, and *Hemipodus*, known for 16 species from shallow depths, occurs chiefly in the Western Hemisphere.

Goniadidae. Closely allied to the Glyceridae, the family Goniadidae has 7 genera and more than 60 species, from all seas and at all depths. *Glycinde*, *Goniada*, *Goniadides*, and *Ophioglycera* are worldwide in occurrence; *Goniadella* and *Progoniada* are limited to the North Atlantic; and *Goniadopsis* occurs in the Red Sea and Indian Ocean.

Eunicea. This is a superfamily which comprises 6 families. The Onuphidae are known for 10 genera and about 100 species; they are tubicolous, grazing, herbivorous, scavenging, are worldwide in occurrence, and exist in intertidal to abyssal depths. Conspicuous representatives are the quill worm (*Hyalinoecia*) and *Diopatra*, which is chiefly intertidal or in shelf depths, sometimes forming vast colonies. *Nothria* and *Paronuphis* are common in the shallow to greatest ocean depths. *Rhamphobrachium* is unique for its greatly prolonged setae in the first segment, which extend far back into the body and

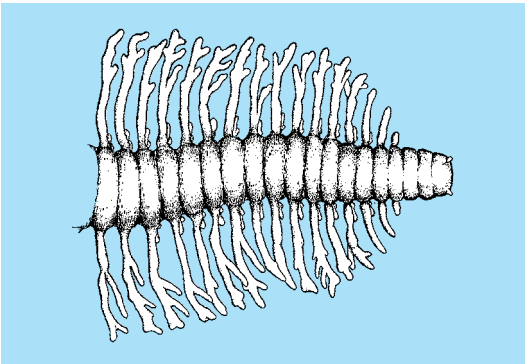

Fig. 8. Anterior end of *Aglaophamus*, of the Nephtyidae, with everted proboscis, dorsal view.

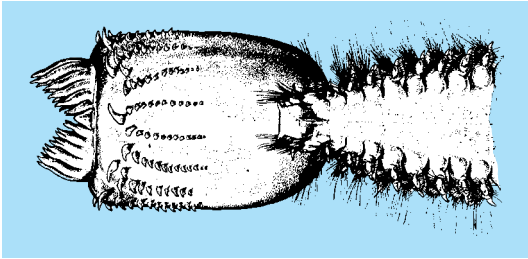

Fig. 9. *Iphitime*, of the Lysaretidae, dorsal view.

are capable of unusual forward extension, like the long shafts of knitting needles. Some onuphids display unusual care of the brood.

The Eunicidae have 7 genera and about 225 species, coming mainly from tropical seas. Included are the coral-eating *Lysidice*; the edible, swarming palolo of the South Pacific islands; and a bait worm, *Marphysa*. Some of its members are errantiate; others tubicolous. All have a highly characteristic pharyngeal armature consisting of maxillae and mandibles.

The Lumbrineridae have 3 genera and about 130 species; *Lumbrineris* is by far the largest genus and is known in all seas at all depths. *Ninoe* and *Cenogenus* are small, chiefly deep-water genera.

The Arabellidae are known for 9 genera and about 66 species. *Arabella*, *Drilonereis*, and some others are chiefly free-living, or partly parasitic, whereas *Haematocleptes*, *Oligognathus, Labidognathus*, and *Ophiuricola* are wholly parasitic. They retain the typical pharyngeal aramature of the superfamily Eunicea.

The Lysaretidae have 3 genera and 8 species; some are large fish-bait worms (*Lysarete* and *Halla*) in tropical regions where they occur. *Iphitime* (**Fig. 9**) is an ectoparasite on large crabs.

The Dorvilleidae have 3 genera and about 40 species. All are small to minute and are chiefly intertidal in epiphytic growth. The largest genus is *Dorvillea*, occurring in worldwide seas; *Protodorvillea* and *Ophryotrocha* are much smaller forms.

Two parasitic families allied to the Eunicea are the Histriobdellidae, ectoparasites of crayfishes, with 2 genera and 5 species, and the Ichthyotomidae, parasites of fishes, with a single genus and species. *See* ANNELIDA.

Olga Hartman

Bibliography. O. Hartman, *Catalogue and Index of the Polychaetous Annelids of the World*, University of Southern California, 1959–1965; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Erysipelothrix

A genus of gram-positive bacteria comprising at least two species, *E. rhusiopathiae* and *E. tonsillarum*. *Erysipelothrix rhusiopathiae* is the more pathogenic and causes disease in a variety of animal species, including humans. It is a rod-shaped

organism, 0.3 × 1–2 micrometers, that may form filaments in old cultures and in chronically infected tissues. It occurs in the tonsil of healthy swine and on the surfaces of fish and other aquatic species; it is shed in the urine, oral secretions, and feces of infected animals. Organisms deposited in the soil may survive for years and may be a source of infection for susceptible domestic and wild animals and birds.

Epidemiology and transmission. *Erysipelothrix rhusiopathiae* causes disease in swine (erysipelas), turkeys, lambs, and a variety of other domestic and wild mammals and birds. Infections are acquired via skin abrasions, or occur endogenously when organisms residing in the tonsil invade the bloodstream of swine stressed by excessive heat. The organism spreads via contaminated excretions and saliva. Infections in humans are associated with occupational exposure to the organism (veterinarians, butchers, pathologists, fish-handlers). They are usually acquired via minor cuts and abrasions on the hands or arms, leading to a painful local purplish erythema (reddening of the skin) termed erysipeloid. This infection may occasionally lead to invasion of the bloodstream or joints.

Clinical findings and pathology. In swine, the acute disease develops suddenly with high fever, prostration, and conjunctivitis, and death within a day or two of septicemia. A less severe form is associated with rhomboid, deep-red skin lesions and is called diamond skin disease. Chronic swine erysipelas takes the form of verrucous (wartlike) growths on the heart valves and a progressive immune-mediated arthritis resembling rheumatoid arthritis in humans. Polyarthritis is a common manifestation in lambs. The disease is often devastating in turkey flocks, causing mass mortality due to septicemia. A thin capsule on *E. rhusiopathiae* confers resistance to phagocytosis and intracellular killing by host phagocytes. Hyaluronidase and neuraminidase also contribute to virulence. Multiplication of the organism in, and subsequent killing of, phagocytes are critical in determining the outcome of infection. In the terminal stages, thrombi form in the small blood vessels, which may rupture and hemorrhage.

Diagnosis and treatment. *Erysipelothrix rhusiopathiae* is easily demonstrated microscopically and by culture of blood and tissues of affected animals. Penicillin is the antibiotic of choice for treatment of infected animals or humans.

Control and prevention. Protective immunity is stimulated by live attenuated strains or by inactivated virulent *E. rhusiopathiae*. A protein is involved in stimulation of protective antibodies. Live vaccines stimulate both enhanced killing activity of phagocytes and protective opsonizing antibodies. *See* MEDICAL BACTERIOLOGY. John F. Timoney

Bibliography. J. F. Timoney et al., *Hagan and Bruner's Microbiology and Infectious Diseases of Domestic Animals*, 8th ed., 1988; J. F. Timoney and M. H. Groschup, Properties of a protective protein antigen of *Erysipelothrix rhusiopathiae*, *Vet. Microbiol.*, 37:381, 1993; R. L. Wood and A. D. Leman et al. (eds.), *Diseases of Swine*, 1992.

Escalator

A continuous, moving stairway that transports passengers between two levels. Most escalators are installed to service floors less than 20 ft (6 m) apart, but some escalators have been constructed to rise 100 ft (30 m) or more between floors.

The modern escalator consists of a series of steps, fastened to an endless chain running in a track system that extends up an incline. The steps continue around a turnaround at the upper landing, pass back down the incline under the steps available to the passengers, and progress to another turnaround at the lower landing. This chain, known as a step chain, is powered by a machine that is usually connected by a gear, chain, or mechanical coupling to a shaft located at the upper landing turnaround. An alternating-current induction motor, selected for its simplicity and constant speed characteristics, is used to move the steps. Most modern escalators travel at 100 ft/min (0.5 m/s). *See* INDUCTION MOTOR.

The track system is designed to permit the steps to approach the upper and lower landings horizontally and to maintain a horizontal tread surface wherever the steps will be holding passengers. The treads of the steps are provided with cleats that match a pattern of combs at the landings, and each riser (the vertical front face of the step) is fitted with cleats that mesh with the adjacent step. The cleats are a safety feature, reducing the possibility of objects becoming caught at the landings or between steps.

The escalator has handrails on each side of the moving steps. They move at the same speed as the steps and extend a distance beyond the point where the steps pass under the combs at the landings. This provides support for passengers as they move onto or off the moving steps.

The escalator mechanism is covered with a balustrade that creates a safe environment for the passengers and enhances the product architecturally. There are two types of balustrades, glass and opaque. The glass type has glass panels below the handrails and gives the escalator an open appearance. The opaque type uses nontransparent panels instead of glass. Glass is usually preferred in installations in open atriums found in building lobbies or shopping malls.

The use of escalators is widespread because of their continuous accessibility and their ability to transport large numbers of people. The capacity, in passengers per hour, is determined by the step width and the speed at which the escalator travels. The most common step widths are 24 and 40 in. (60 and 100 cm). The capacity is based on one passenger per 40-in. (100-cm) step or one passenger on every other 24-in. (60-cm) step. These values are reliable for planning the number of escalators required in a building, even though theoretically a larger capacity can be achieved over a short time. Based on such step loading, an escalator running at 100 ft/min (0.5 m/s) can transport 4500 or 2250 passengers per hour with 40- or 24-in. (100- or 60-cm) step widths, respectively.

Cutaway view of an escalator. (*Otis Elevator Co.*)

Movement of people by escalators is very economical. The escalator is a balanced machine when operating empty. As passengers board in the down direction, they offset the friction in the system. When the down load exceeds approximately 10% capacity, power is generated by the escalator so that the total power consumption of the building containing the unit is reduced. Although power is required to carry passengers up, the cost to carry a passenger up one floor on a fully loaded escalator is extremely small, typically less than a fraction of a cent. D. L. Steel

Escape velocity

Minimum speed away from a parent body that a particle must acquire to escape permanently from the gravitational attraction of the parent. Escape velocity is also termed parabolic velocity. *See* ORBITAL MOTION.

Earth retains an atmosphere because the escape velocity is considerably higher than the mean velocity of the gas molecules in its atmosphere. For a space ship to escape from Earth and travel to another planet or orbit about the Sun, it must reach escape velocity. This velocity v can be calculated by equating the kinetic energy of the moving body m to the work necessary to overcome the gravity g_0 at the surface of the parent whose radius is r_0 in rising to a height h above the surface of the parent; thus Eq. (1) holds. From this Eq. (2) is derived, where for Earth $r_0 = 6.38 \times 10^6$ m and $g_0 = 9.8$ m/s^2 and for escape $h = \infty$ so that the term under the second radical is unity; thus $v_{\text{escape}} = 37.0 \times 10^3$ ft/s (11.2×10^3 m/s). For comparison, the mean velocity of a gas with molecules of mass m at absolute temperature T is $v = \sqrt{3kT/m}$, where k is Boltzmann's constant. Thus hydrogen has a mean velocity in the vicinity of 6.6×10^3 ft/s (2×10^3 m/s) and heavier gases in an atmosphere have lower velocities. Escape velocities for other bodies in the solar system are Mercury, 12.5×10^3 ft/s (3.8×10^3 m/s); Venus, 34.3×10^3 ft/s (10.4×10^3 m/s); Mars, 16.8×10^3 ft/s (5.1×10^3 m/s); Jupiter, 201.3×10^3 ft/s (61.0×10^3 m/s); Saturn, 121.1×10^3 ft/s (36.7×10^3 m/s); Uranus, 72.6×10^3 ft/s (22×10^3 m/s); and Neptune, 79.2×10^3 ft/s (24×10^3 m/s). *See* CELESTIAL MECHANICS; SATELLITE (SPACECRAFT); SPACE FLIGHT. R. L. Duncombe

$$\frac{mv^2}{2} = m\int_0^h g_0\left(\frac{r_0}{r_0+h}\right)^2 dh = \frac{mg_0r_0h}{r_0+h} \qquad (1)$$

$$v = \sqrt{2g_0r_0}\sqrt{\frac{h}{r_0+h}} \qquad (2)$$

Escapement

A mechanism in which a toothed wheel engages alternate pallets attached to an oscillating member. The escapement is found principally in timepieces but may be employed wherever oscillating motion is required. Its origins are ancient and obscure.

In the simple form of escapement (**Fig. 1**), oscillating member cc' is an open bar arranged to slide longitudinally in bearings CC, which are attached through a frame (not shown) to the bearing for toothed wheel a. Wheel a turns continuously in the direction of the arrow, and is provided with three teeth b, b', and b''. The oscillating member has two pallets c and c'. In the position shown, tooth b is just ceasing to drive pallet c to the right, and is escaping, while tooth b' is just coming in contact with pallet c', which it will drive to the left.

Although escapements are generally used to convert circular into reciprocating motion, as in the above example, the wheel being the driver, in many cases the action may be reversed. In Fig. 1, if the open slide bar were driven with reciprocating motion, the wheel would be made to turn in the opposite direction from its rotation as the input member. Also, there would be a short interval at the beginning of each stroke of the bar in which no motion would be given to the wheel. The wheel *a* must have one, three, five, or any other odd number of teeth upon its circumference.

Escapements are also adapted for use in automatic factory equipment to regulate the flow of parts from the magazine or feeder track. Such feeding and spacing devices are of various forms, many being adaptations of the ratchet and pawl. *See* PAWL; RATCHET.

In a mechanical clock or watch, the escapement intervenes between the energy source (spring or elevated weight) and the regulating device (pendulum or balance wheel). It is acted upon by both. The escape wheel is mounted on the same shaft as the last wheel of the gear train, and impulses are delivered from the escape wheel to operate the regulating device. The regulating device, which has a natural period of oscillation, determines the rate at which it will receive these impulses, and thus regulates the rate of going of the timepiece.

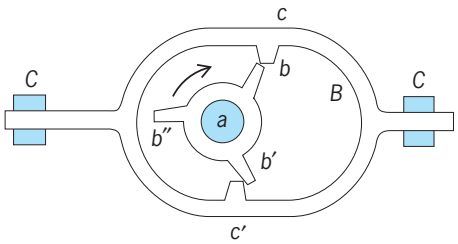

Fig. 1. Simple form of escapement.

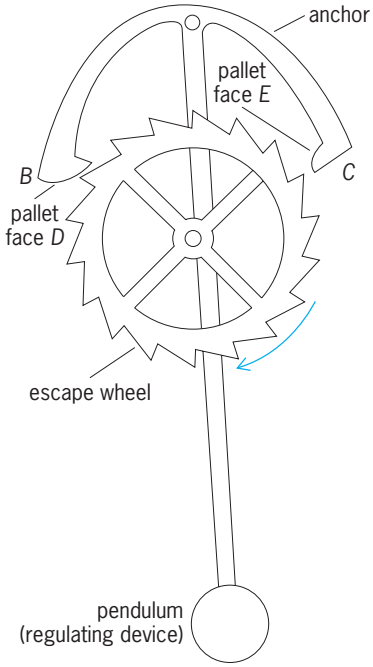

Fig. 2. Anchor recoil escapement.

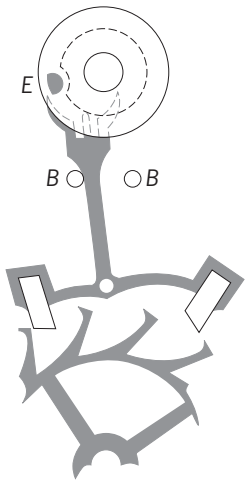

Fig. 3. Detached-lever escapement, a type used in watches.

The anchor recoil escapement (**Fig. 2**) is used with a pendulum and takes its name from the shape of its oscillating member and its action. This type of escapement appeared late in the seventeenth century and has survived, with modifications, to the present. To simplify the explanation, the pendulum is shown attached directly to the anchor, so that the pendulum and anchor swing as one.

In the position shown, pallet *B* has just received an impulse from the escape wheel, the impulse swinging the pendulum to the left. When pallet *B* has cleared the tooth, allowing the wheel to escape, pallet *C* will be in position to arrest the wheel. Recoil, or momentary reversal of the escape wheel, occurs just after it is arrested because the pendulum has not quite completed its swing. The wheel tooth in contact with pallet *C* will then give the oscillating parts an impulse in the opposite direction.

Deadbeat, an escapement without recoil, has arresting faces of the pallets described by a circular arc whose center is at the pivot point of the anchor. Escape-wheel teeth are contoured to give impulses to these pallet faces, over which they slide without recoil.

Modern watches generally employ a detached-lever escapement (**Fig. 3**), which has banking pins *B* to limit the oscillation of the anchor and its lever. An escapement is termed detached when the regulating device, in this case the balance wheel, is given an impulse during only a small part of its operating cycle. When the fork reaches the end of its swing, it is lightly locked by a wheel tooth and remains stationary until the returning impulse pin *E* causes sufficient recoil of the escape wheel to release the pallet. The chronometer escapement is a detached escapement that furnishes an impulse in only one direction of the swing of the balance wheel. *See* CHRONOMETER; CLOCK (MECHANICAL).

Some watches employ a light ratchet powered by a vibrating reed at 360 Hz in place of a balance wheel and escapement. *See* WATCH.

Douglas P. Adams

Bibliography. H. G. Harris, *Handbook of Watch and Clock Repairs*, rev. ed., 1972; H. H. Mabie and F. W. Ocvirk, *Mechanisms and Dynamics of Machinery*, 4th ed., 1987.

Escarpment

A long line of cliffs or steep slopes that break the general continuity of the land by separating it into two level or sloping surfaces. Some very high escarpments, or scarps, may form by vertical movement along faults. Often a whole block of land may be forced upward while the adjacent block is downfaulted. Such scarps are common in the tilted fault-block mountains of eastern California, Nevada and western Utah. In eastern Africa, prominent fault scarps mark the margins of the great rift valleys, whose floors are downfaulted by as much as 8200 ft (2500 m). *See* FAULT AND FAULT STRUCTURES.

Other types of escarpments form by differential weathering and erosion of contrasted rock types. Less resistant rocks, such as clay or shale, are often eroded from beneath resistant cap rocks, such as sandstone and limestone. With support removed from below, the cap rock fails and the escarpment retreats. Escarpments are often very prominent in arid regions, where hardened weathering products may form extensive cap rocks known as duricrusts. When inclined strata are eroded, they may produce cuesta escarpments, which have back slopes that approximate the dip of their sedimentary layers and steeper facing slopes (the scarp faces) that truncate the bedding.

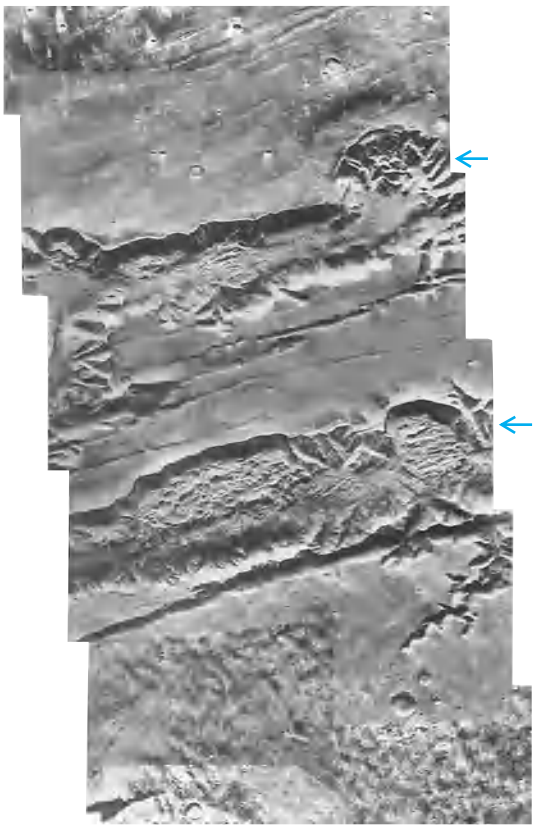

Tithonium Chasma (top) and Ius Chasma (bottom) in the great equatorial canyon system of Mars. The escarpments bounding them are 1.2–3.7 mi (2–6 km) high. The photographed scene is about 155 mi (250 km) across. Arrows indicate scars of the great landslides. (*NASA*)

Some of the largest known escarpments occur on the planet Mars, where erosion has presumably been much slower than on the Earth in reducing primary structural relief. The great shield volcano Olympus Mons is surrounded by a basal scarp that forms a circle approximately 430 mi (700 km) in diameter. Even more spectacular are the great escarpments that bound the Valles Marineris, a group of steep-walled equatorial canyons more than 2500 mi (4000 km) in length. The escarpments are as much as 4 mi (6 km) high and contain the scars of great landslides that contributed to scarp retreat (see **illus.**).

Victor R. Baker

Escherichia

A genus of bacteria named for Theodor Escherich, an Austrian pediatrician and bacteriologist, who first published on these bacteria in 1885. *Escherichia coli* is the most important of the six species which presently make up this genus, and it is among the most extensively scientifically characterized living organisms. *Escherichia coli* are gram-negative rod-shaped bacteria approximately 0.5 × 1–3 micrometers in size.

Escherichia coli can be differentiated from other closely related bacteria by a variety of assays for metabolic processes, including the fermentation of specific sugars and the enzymatic modification of amino acids. Modern molecular taxonomic analysis based on the nucleotide sequences of ribosomal ribonucleic acid (RNA) has revealed that *Shigella*, a bacterial genus of medical importance previously thought to be distinct from *E. coli*, is actually the same species.

The natural habitat of *E. coli* is the colon of mammals, reptiles, and birds. In humans, *E. coli* is the predominant bacterial species inhabiting the colon that is capable of growing in the presence of oxygen. The presence of *E. coli* in the environment is taken to be an indication of fecal contamination.

Most strains of *E. coli* are commensals which are harmless to the humans and other animals they colonize, but some strains can cause disease when given access to extraintestinal sites or the intestines of noncommensal hosts. *Escherichia coli* is the most important cause of urinary tract infections. Specialized protein structures on the surface of *E. coli*, known as adhesive fimbriae, permit some strains to colonize the epithelial cells lining the urinary tract, leading to inflammation and tissue damage. Women are more susceptible than men to *E. coli* urinary tract infections due to the proximity of the urethra to the anus. Four out of ten women experience at least one urinary tract infection in their lifetime. *Escherichia coli* urinary tract infections are also associated with catheterization and instrumentation of the urinary

tract. Men and women are equally susceptible in this setting. Urinary tract infections may extend into the bloodstream, especially in hospitalized patients whose defenses are compromised by the underlying illness. This may lead to a type of whole-body inflammatory response known as sepsis, which is frequently fatal. Certain *E. coli* strains with a specific polysaccharide capsule on their surface can invade the intestine of the newborn and cause sepsis and meningitis. These strains are acquired at birth from *E. coli* which have colonized the vagina of the mother.

Several different strains of *E. coli* cause intestinal infections. In the developing world, the most important of these are the enterotoxigenic *E. coli*. These strains produce specialized adhesive fimbriae which enable them to colonize the small intestine. Enterotoxigenic *E. coli* produce enterotoxins which act on the epithelial cells lining the small intestine, causing the small intestine to reverse its normal absorptive function and secrete fluid. This leads to a dehydrating diarrhea which can be fatal, especially in poorly nourished infants. Therapy consists of oral or, in serious cases, intravenous rehydration. Enterotoxigenic *E. coli* are transmitted by ingestion of fecally contaminated water and food, and are a common cause of diarrheal disease in travelers in developing countries. Enterotoxigenic *E. coli* are also an important cause of diarrheal disease in livestock animals, especially calves and piglets. These strains do not infect humans.

An important group of pathogenic *E. coli* in developed countries are the enterohemorrhagic strains, especially the serotype known as *E. coli* O157:H7. These strains are normal commensals in cattle but cause bloody diarrhea in humans. A complication of approximately 10% of cases is a potentially fatal disease known as hemolytic uremic syndrome. The virulence of these strains involves the close attachment of bacteria to epithelial cells lining the colon, resulting in alteration of the epithelial cell structure, and the production of Shiga toxin. The toxin enters the bloodstream after being absorbed in the colon and damages the endothelial cells lining the blood vessels of the colon, resulting in bloody diarrhea. In cases of hemolytic uremic syndrome, the toxin circulating in the blood damages blood vessels in the kidney, resulting in kidney failure and anemia. Antibiotics are not recommended at present due to concern that the action of antibiotics on the bacterial cells will increase the release of toxin. Enterohemorrhagic *E. coli* are acquired by the ingestion of undercooked beef, uncooked vegetables, or unpasteurized juices from fruits which have been contaminated with the feces of infected cattle. An infection can also be acquired from contact with a human infected with the organism and from contaminated water. Children and the elderly are at greatest risk of developing hemolytic uremic syndrome.

Other strains which are pathogenic in the human colon include the enteroinvasive *E. coli* (including *Shigella*) and the enteropathogenic *E. coli*. Enteroinvasive *E. coli* induce the epithelial cells lining the colon to engulf the bacteria. The bacteria then gain access to the interior of the epithelial cells, where they rapidly multiply and destroy the cell. An inflammatory response ensues, and the bacteria are ultimately destroyed by neutrophils. The disease is called bacillary dysentery and is characterized by bloody diarrhea. It can be fatal in poorly nourished children but is not usually associated with serious complications in developed countries, where it remains a common cause of diarrheal disease. Acquisition of enteroinvasive *E. coli* and *Shigella* is by ingestion of food or water contaminated with feces of a person infected with the organism, or by direct contact with infected feces. The disease is usually treated with antibiotics.

Enteropathogenic *E. coli* colonize the small intestine and interact closely with epithelial cells in its lining, altering the structure and function of the cells in a manner similar to enterohemorrhagic *E. coli*. Unlike enterohemorrhagic *E. coli*, enteropathogenic *E. coli* do not release toxins, and damage to the intestine appears to result from structural alterations to the cells. Enteropathogenic *E. coli* have been associated with protracted diarrhea in infants and can occasionally cause severe wasting. There are no routine methods available for the diagnosis of enteropathogenic *E. coli* infections, and the prevalence of disease caused by these organisms is unknown. Antibiotics are of uncertain efficacy. *See* DIARRHEA; TOXIN.

Steve L. Moseley

Bibliography. J. B. Kaper and A. D. O'Brien (eds.), *Escherichia coli O157:H7 and Other Shiga Toxin-Producing E. coli Strains*, ASM Press, 1998; H. L. T. Mobley and J. W. Warren (eds.), *Urinary Tract Infections: Molecular Pathogenesis and Clinical Management*, ASM Press, 1996; A. A. Salyers and D. D. Whitt, *Bacterial Pathogenesis: A Molecular Approach*, ASM Press, 1994.

Esociformes

An order of teleost fishes, also known as Haplomi and Esocae. Fishes of this small order, comprising two families and 10 species, can be identified by a combination of the following characteristics: dorsal and anal fins are posteriorly placed on a rather slender body; maxillae are toothless and in the gape of the mouth; small uniform teeth are on the tongue and basibranchial elements; and absent are adipose fin, breeding tubercles, pyloric caeca, and mesocoracoid bones. Both families occur in freshwater of the Northern Hemisphere. *See* OSTEICHTHYES; TELEOSTEI.

Esocidae (pikes). Esocids have a duckbill-like snout; a forked caudal fin; large canine teeth on the dentary, vomer, and palatine bones; small embedded cycloid scales; a complete lateral line; eight or more pores in the infraorbital canal; and abdominal pelvic fins (**Fig. 1**). The family consists of one extant genus, *Esox*, and five species: *E. lucius* (northern pike) is circumpolar in distribution; *E. reichteri* is endemic to the Amur River in eastern Asia; and three species—*E. americanus* (with two

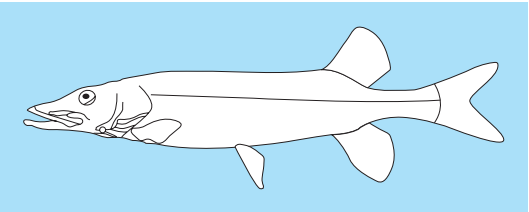

Fig. 1. Example of an esocid (pike). (*Courtesy of J. S. Nelson, 2006*)

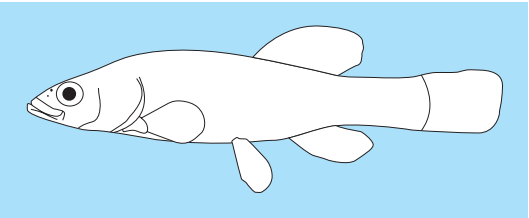

Fig. 2. Example of an umbrid (mudminnow). (*Courtesy of J. S. Nelson, 2006*)

subspecies, the redfin pickerel and the grass pickerel), *E. masquinongy* (muskellunge), and *E. niger* (chain pickerel)—are restricted to North America east of the Rocky Mountains.

Esocids are diurnal predators that feed mainly on other fishes, which are attacked from ambush. The muskellunge is the largest species, reaching a length of 1.4 m (4.6 ft).

Umbridae (mudminnows). Mudminnows are distinguished from esocids by the following characteristics: short snout, rounded caudal fin, faint or absent lateral line, and three or fewer pores in the infraorbital canal (**Fig. 2**). The family comprises three genera and five species: *Dallia pectoralis* (Alaska blackfish) occurs in Alaska and northeastern-most Siberia; *Novumbra hubbsi* (Olympic mudminnow) is limited to the Olympic peninsula in western Washington; *Umbra limi* (central mudminnow) occurs in east-central North America; *U. pygmaea* occurs in eastern United States; and *U. krameri* inhabits the Danube and Dniester river systems in southeastern Europe.

Umbrids inhabit densely vegetated swamps, bogs, ponds, sluggish creeks, as well as large rivers and lakes that support abundant vegetation. They are usually found on or near a substrate of mud or silt and feed on adult and larval aquatic insects, larval crustaceans, and snails. These hardy fishes are capable of surviving extreme cold and oxygen-depleted water. The maximum total length is 8–33 cm (3–13 in.), depending on the species. Herbert Boschung

Bibliography. T. Grande, H. Laten, and J. A. Lopez, Phylogenetic relationships of extant esocid species (Teleostei: Salmoniformes) based on morphological and molecular characters, *Copeia*, 2004(4):743–757, 2004; J. A. Lopez, P. Bentzen, and T. W. Pietsch, Phylogenetic relationships of esocoid fishes (Teleostei) based on partial cytochrome b and 16S mitochondrial DNA sequences, *Copeia*, 2000(2):420–431, 2000; J. A. Lopez, W. J. Chen, and G. Ortí, Esociform phylogeny, *Copeia*, 2004(3):449–464, 2004; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Esophagus

A section of the alimentary canal that is interposed between the pharynx and the stomach. Because of divergent specializations in the various vertebrates, the esophagus cannot be described in general terms and is not always distinguishable.

In humans it is a tube running the full length of the neck and the thorax, held in its position ventral to the vertebral centra by a tunica adventitia of loose connective tissue. It has an inner lining of folded mucous membrane with an exceptionally thick lamina propria, a submucosa of elastic and collagenous connective tissue, and two layers of muscle. The musculature is striated in the anterior third of its length, unstriated in the posterior third, and variably intermixed in the middle. It is supplied with autonomic nerve fibers.

Although normally collapsed, the human esophagus is capable of considerable distension during the rapid passage of swallowed material, under which condition the folds of mucous membrane and lamina propria are temporarily smoothed out. Numerous microscopic esophageal glands open into the lumen, extending their compound tubules out into the submucosa.

In humans the transition from the esophagus to the stomach occurs quite abruptly at the diaphragm. In many vertebrates the distinction between the two is not so clear. The pharynx narrows posteriorly like a funnel and the foregut may thereupon enlarge, but much of what appears to be stomach may have an esophageal character histologically. In a number of fish species there is no stomach enlargement at all. There is a metabranchial foregut anterior to the entrance of ducts from the liver and pancreas, but application to it of the names stomach and esophagus is of questionable value. At the other extreme, in some sharks and bony fishes there is a constricted esophagus with a sparsely glandular or cornified lining. The esophagus of some turtles is studded with harsh horn-tipped papillae. In some of the carnivorous birds the esophagus may be temporarily dilated for the storage of large masses of swallowed food before they can be digested. Many graminivorous (grain- and seed-eating) birds are provided with a permanent ventral esophageal pouch, the crop, for that purpose. Pigeons of both sexes feed their newly hatched young with a thin paste consisting of cells proliferated and sloughed from special areas of the crop under the stimulus of the hormone prolactin. This is pigeon's milk. *See* DIGESTIVE SYSTEM.

William W. Ballard

Essential oils

Volatile, fragrant oils obtained from plants. Essential oils are distinguished from those known as fixed oils, which are mainly triglycerides of fatty acids. *See* FAT AND OIL.

Essential oils have been obtained from over 3000 plants and are designated and defined by the plant species and sometimes the geographical location.

The sources of these oils are diverse, including flower petals (for example, rose and jasmine), spices (cinnamon and ginger), pine oil and turpentine, and citrus fruit peels. Compounds present in the juice that may contribute to the distinctive flavor of a fruit or berry are not, strictly speaking, components of the essential oil. Chemically, essential oils are extremely complex mixtures containing compounds of every major functional-group class. The oils are isolated by steam distillation, extraction, or mechanical expression of the plant material; often only certain parts, such as roots, buds, leaves, or flower petals, are used.

History. Essential oils have been produced and used for flavoring, incense, and medicinal purposes for many centuries. There are indications that crude preparations were known in ancient Egypt and Persia. Distillation of the essence of plants was described about the year 1500.

In the beginnings of modern organic chemistry, essential oils provided a fertile source of compounds for studying the structures and reactions of organic molecules. Early work revealed that several compounds from the oils of aromatic plants had in common a chemically inert structural unit, and these compounds were designated as aromatic. This unit was later determined to be a benzene ring (C_6H_6). From this association the modern term aromatic arose to denote the electronic structure of a large class of compounds of which benzene is the parent. Chemical investigations of essential oils also led directly to the development of terpene chemistry. *See* BENZENE; TERPENE.

Chemistry. Among the hundreds of compounds that have been separated from essential oils are many monoterpenes (10 carbons, giving 2 isoprene units) and a smaller number of sesquiterpenes (15 carbons giving 3 isoprene units). These terpenes may be open-chain, monocyclic, or polycyclic, and usually contain one or more double bonds and hydroxyl groups. Also present, generally in very small amounts, are numerous saturated aliphatic aldehydes and carboxylic esters. The major constituent of several important essential oils is a di- or trisubstituted benzene derivative. A few representative examples are shown in the **illustration**.

Monoterpene hydrocarbons make up the bulk of many essential oils that are used in flavoring foods and beverages. Some of the most valuable oils, used in perfumery, are obtained from tropical plants and contain a high proportion of bi- and tricyclic sesquiterpenes.

At one time essential oils were the principal or only source available for these compounds, but synthetic methods have advanced to the point that any compound that can be isolated from an essential oil can now be prepared by synthesis. Large-scale processes are used to manufacture a number of the major components present in commercially important oils. Linalool and geraniol can be prepared from acetone or isobutene, formaldehyde, and acetylene. Racemic menthol is available from thymol, which is in turn manufactured from toluene. A number of monoterpenes, in optically active form, are derived by simple reactions from the abundant pinenes in turpentine.

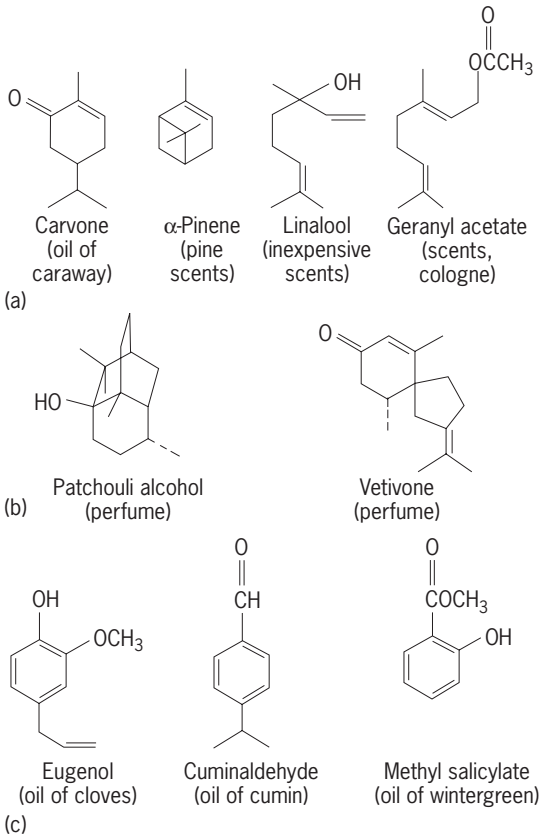

Examples of benzene derivatives in some essential oils: (*a*) monoterpenes, (*b*) sesquiterpenes, and (*c*) aromatics.

On the other hand, the most practical source of certain valuable perfume ingredients, such as patchouli alcohol and related complex sesquiterpenes, is the essential oil. *See* MENTHOL.

Characterization and analysis. The importance of essential oils for perfumes or flavors lies in the fact that they are extremely complex mixtures of as many as 200–300 compounds, usually a few major constituents and many minor ones. Distinctive odors and tastes may depend on the blend of all these components. In early work, only a few of the main constituents were isolated by fractional distillation and identified by various chemical methods. The total mixture composing the oil is characterized by such physical properties as color, boiling-point range, specific gravity, refractive index, and optical rotation. These properties and sometimes also information on chemical composition, such as total hydroxyl or aldehyde content, are used to specify essential oils of acceptable medicinal quality in the United States Pharmacopeia. For use as flavors or scents, organoleptic tests are an important part of the evaluation.

The development of gas chromatography in the 1960s and then the use of very long capillary columns, typically 0.001 in. × 165–330 ft (0.25 mm × 50–100 m), with a film of absorbent on the wall revolutionized the analysis of essential oils. With these wall-coated open-tube columns, coupled to the inlet of a mass spectrometer, it is possible to separate, identify, and quantitatively estimate over 100 compounds. Others can be separated and characterized

by general type, and still many more can be detected. *See* GAS CHROMATOGRAPHY.

Other analyses have been based on the infrared spectra of the components separated by capillary gas chromatography columns. An additional physical constant that is used in characterizing the individual terpenes and other components of essential oils is the Kovats index, which is a measure of the retention time of a given compound in a gas chromatogram relative to the retention times of *n*-alkane reference compounds. Carbon-13 nuclear magnetic resonance spectra of an essential oil can also be used for identification of the major components of the mixture. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

Variability from plant sources. The formation of essential oil in the plant, and consequently the yield and composition of the oil produced, depends on many factors. Genetic differences in plants of the same species that are otherwise indistinguishable (chemotypes) can result in widely different essential oil content. Geographic location and agricultural factors also influence oil production. These include soil, water, nutrients, and climatic variables, such as sunlight, temperature, and day length. The yield of oil from peppermint (*Mentha piperita*) grown in central Washington is threefold greater than that from the same strain grown in heavier soil in the midwestern United States.

Both the quantity and composition of the essential oil can change drastically as the plant matures. In the oil from coriander (*Coriandrum sativum*), the content of aliphatic aldehydes drops and that of the monoterpene alcohols increases tenfold from the stage of full flowering to green fruit.

Processing. The composition and quality of an essential oil is affected significantly by the method of isolation and subsequent processing steps. Enzymatic action (fermentation) in the crude plant material prior to distillation or extraction can bring about hydrolysis of glycosides and release of oil components. The composition of the oil may depend greatly on the isolation method. Steam distillation is the most common process, but sensitive compounds can undergo rearrangement or oxidation on heating. Oils from flower blossoms are extracted by pressing the petals with a purified fat (enfleurage process) or hydrocarbon solvent, and the essential oil is then extracted with alcohol. The alcohol solution is concentrated to give a liquid known as an absolute that is used in the manufacture of perfumes. Extraction of the plant with alcohols can lead to artifacts due to reaction with the solvent. *See* DISTILLATION; EXTRACTION.

Further processing steps depend on the ultimate use of the essential oil. Fractional distillation (rectification) at reduced pressure can be carried out to remove some of the terpene hydrocarbons and produce a concentrate (folded oil) in which the characteristic flavor or scent is enhanced. Distillation or chemical treatment may also be used to remove an undesired trace component. James Moore

Bibliography. E. Guenther, *The Essential Oils*, vols. 1–6, 1947–1952; B. D. Mokerjee and C. J. Mussinan (eds.), *Essential Oils,* 1981; *Perfumer and Flavorist*, vols. 1–10, 1976–1986; R. Teranishi, R. G. Buttery, and S. Sugisawa (eds.), *Bioactive Volatile Compounds from Plants*, ACS Symposium Series 525, 1993.

Ester

The product of a condensation reaction (esterification) in which a molecule of an acid unites with a molecule of alcohol with elimination of a molecule of water, reaction (1).

$$\underset{\text{Acid}}{RC(=O)OH} + \underset{\text{Alcohol}}{HOR'} \longrightarrow \underset{\text{Ester}}{RC(=O)OR'} + H_2O \qquad (1)$$

At one time it was thought that esterification was analogous to neutralization, and esters are still named as though they are "alkyl salts" of carboxylic acids, as shown in the following examples:

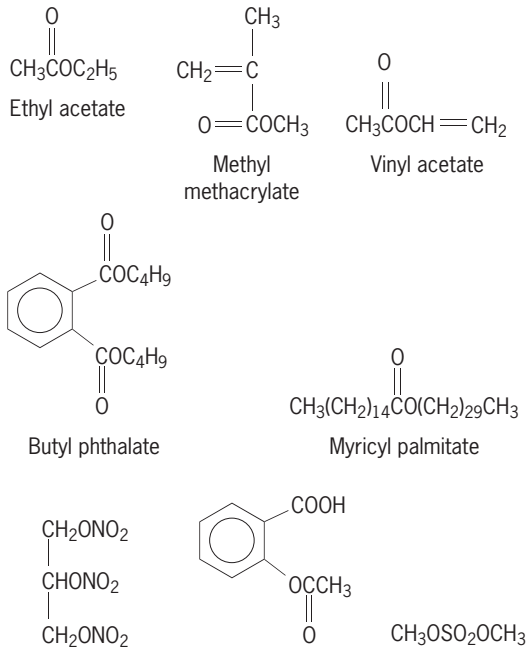

Properties and uses. Esters are generally insoluble in water and have boiling slightly higher than hydrocarbons of similar molecular weight. An ester may often be characterized by its infrared absorption spectrum. For example, saturated aliphatic esters have a distinctive, strong C═O absorption band at 1750–1735 cm^{-1}, and a second band assigned to C—O stretching at 1300–1000 cm^{-1}.

Ethyl and butyl acetates are volatile industrial solvents, used particularly in the formulation of lacquers. The ethyl acetate produced in the United States is primarily used as a solvent. Higher-boiling esters such as butyl phthalate are used as softening agents (plasticizers) in the compounding of plastics. The natural waxes of biological origin are largely simple esters. For example, a principal component of

beeswax is myricyl palmitate. *See* SOLVENT; WAX, ANIMAL AND VEGETABLE.

Esters of cellulose (cellulose triacetate) are used in photographic film and as a textile fiber (acetate rayon). Cellulose acetatepropionate and cellulose acetatebutyrate have become important as thermoplastic materials. Cellulose nitrate, containing 10.5–11% nitrogen, is called celluloid pyroxylin; with alcohol and camphor (a plasticizer), it forms celluloid. Dynamite cotton is cellulose nitrate of 11.5–12.3% nitrogen content, and gun cotton is cellulose nitrate of 12.5–13.5% nitrogen. Cordite and ballistite are made from gun cotton, which is plasticized with glyceryl trinitrate (nitroglycerin). Dimethyl and diethyl sulfates (esters of sulfuric acid) are excellent agents for alkylating organic molecules that contain labile hydrogen atoms, for example, starch and cellulose. *See* CELLULOSE; EXPLOSIVE.

Esters of unsaturated acids, for example, acrylic or methacrylic acid, are reactive and polymerize rapidly, yielding resins; thus, methyl methacrylate yields a poly(methyl methacrylate). Analogously, esters of unsaturated alcohols are reactive and readily react with themselves; thus, vinyl acetate polymerizes to poly(vinyl acetate). The polyester resins known as glyptals result from the polyesterification of glycerol with phthalic anhydride; the process can be controlled to yield either a fusible or an infusible resin. When the polyesterification is carried out in the presence of a long-chain, unsaturated acid of the drying oil type, the oxidative polymerization of the latter is superimposed upon the polyesterification, resulting in hard, synthetic, weather-resistant enamels, suitable for automobile finishes. Polyesterification of ethylene glycol with terephthalic acid results in a polyester fiber. If the material is formed in sheets, it is a useful photographic film. *See* POLYMER.

Many low-molecular-weight esters have characteristic, fruitlike odors: banana (isoamyl acetate), rum (isobutyl propionate), and pineapple (butyl butyrate). These esters are used to some extent in compounding synthetic flavors and perfumes. *See* ALCOHOL; CARBOXYLIC ACID; DRYING OIL; FAT AND OIL; POLYESTER RESINS; SOLVENT.

Esterification. In the broadest sense, esterification is any reaction in which at least one of the products is an ester. There are many routes to the formation of esters. Some of the more important reactions for preparing esters take place between the following pairs of compounds: (1) an acid and an alcohol, (2) an acid anhydride and an alcohol, (3) an acid chloride and an alcohol, (4) an acid and an unsaturated hydrocarbon such as an olefin or an acetylene, (5) an ester and an alcohol, (6) an ester and an acid, and (7) two different esters. This article treats esterfication in only a limited sense—reaction between a carboxylic acid (RCOOH) and an alcohol (R′OH) to give the ester and water. For discussions of reactions of an ester with an alcohol, an acid, or another ester *See* TRANSESTERIFICATION.

Esterification reactions are generally reversible and accompanied by relatively small heat effects of the order of a few kilocalories per mole of ester. Although the reactions generally take place in a single liquid phase in the presence of a catalyst, a limited number of esters have been prepared by passing the reactant vapors over a solid catalyst. In the presence of a catalyst, the reaction is commonly conducted at a temperature of about 100°C (212°F); in the absence of a catalyst, a temperature of about 250°C (480°F) is used to give a reasonable reaction rate. The pressure at which the reaction is conducted is determined only by the volatility of the components of the system. It is usually atmospheric pressure. In order to produce most esters economically, some means must be provided for completing the reaction by removing one or more of the products. *See* ACID ANHYDRIDE; ACID HALIDE.

In a typical industrial procedure for the preparation of ethyl acetate, a mixture of acetic acid, excess ethanol, and sulfuric acid is passed into an esterifying column heated to reflux. A ternary azeotrope containing 70% ethanol, 20% ester, and 10% water separates into layers, one of which contains 85% ethyl acetate. The ester may be purified by fractional distillation, and the recovered starting materials are recycled.

Other commercially important esters are prepared as follows. Dibutyl phthalate is prepared from phthalic anhydride and butanol in a stepwise reaction to form first the monoester and then the diester; cellulose acetate from purified α-cellulose and a mixture of acetic anhydride and acetic acid; alkyd resins from phthalic anhydride, unsaturated fatty acids, and glycerol; nitroglycerine (glycerol trinitrate) from glycerol and the proper mixture of nitric acid and sulfuric acid. Aspirin, the world's most used analgesic, is prepared by the reaction of salicylic acid with acetic anhydride below 90°C (190°F), and is purified by recrystallization. *See* ASPIRIN; AZEOTROPIC MIXTURE.

The fact that esterification involves an equilibrium was established in 1862 by M. P. E. Berthelot in his study of the ethyl alcohol–acetic acid system. If 1 mole each of acetic acid and ethyl alcohol react, it is found at equilibrium that $^2/_3$ mole each of ethyl acetate and water is present at room temperature, along with $^1/_3$ mole each of alcohol and acid. This can be applied to the equilibrium equation shown as Eq. (2), where K_E is the equilibrium constant and

$$K_E = \frac{[CH_3COOC_2H_5] \times [H_2O]}{[CH_3COOH] \times [C_2H_5OH]} \qquad (2)$$

the square brackets signify concentrations in moles per liter of the enclosed reagent. This gives for K_E the value 4. Unless the temperature is deliberately changed, this value is fixed for ethyl alcohol and acetic acid (different alcohol-acid systems have different though characteristic equilibrium constants); indeed, regardless of the starting concentrations of acid and alcohol, the value 4 is maintained.

The mechanism of direct esterification has been much studied. The use of isotopic oxygen (^{18}O) shows that in the reaction of an acid with an alcohol of primary or secondary type, the ester oxygen comes from the alcohol and the acid oxygen

goes to form water. Moreover, in an ordinary acid-catalyzed esterification, the rate of reaction is dependent upon the concentrations of both the carboxylic acid and the alcohol. These observations are accommodated by the mechanistic picture shown as reactions (3)–(6).

$$R-\overset{\overset{\displaystyle O}{\|}}{C}-OH + H^+ \text{ (from catalyst)} \rightleftharpoons RC(OH)_2^+ \quad (3)$$

$$RC(OH)_2^+ + R'OH \rightleftharpoons RC(OH)_2(OR')^+ + H^+ \quad (4)$$

$$RC(OH)_2(OR') + H^+ \rightleftharpoons RC(OH)(OR')^+ + H_2O \quad (5)$$

$$RC(OH)(OR')^+ \rightleftharpoons R\overset{\overset{\displaystyle O}{\|}}{C}-OR' + H^+ \quad (6)$$

In the case of tertiary alcohols, isotopic oxygen studies show that the ester oxygen comes from the carboxylic acid, and the hydroxyl from the alcohol goes to form water, implying the modified picture represented as reactions (7) and (8).

$$R-\underset{\underset{\displaystyle R}{|}}{\overset{\overset{\displaystyle R}{|}}{C}}-O-H + H^+ \text{ (from catalyst)} \rightleftharpoons R-\underset{\underset{\displaystyle R}{|}}{\overset{\overset{\displaystyle R}{|}}{C^+}} + H_2O \quad (7)$$

$$R-\underset{\underset{\displaystyle R}{|}}{\overset{\overset{\displaystyle R}{|}}{C^+}} + \underset{\underset{\displaystyle H}{|}}{O}-\overset{\overset{\displaystyle O}{\|}}{C}-R \rightleftharpoons R-\underset{\underset{\displaystyle R}{|}}{\overset{\overset{\displaystyle R}{|}}{C}}-O-\overset{\overset{\displaystyle O}{\|}}{C}-R + H^+ \quad (8)$$

Aromatic acids having substituents in both ortho positions are so hindered in their reaction with alcohols that direct esterification is impracticable. When such acids are dissolved in 100% sulfuric acid and the resulting solution is poured into an alcohol, a good yield of the ester is obtained in a few minutes. The mechanism of this reaction involves an intermediate acyl ion, as shown in reactions (9) and (10).

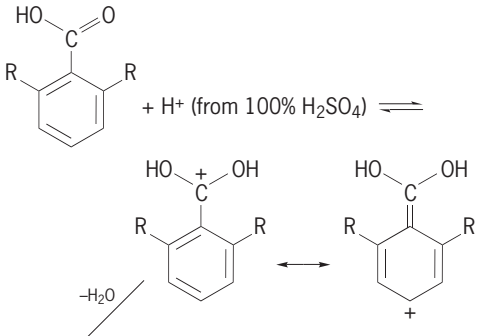

(9)

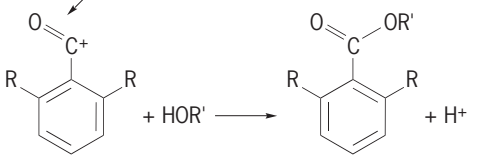

(10)

Hydrolysis. The splitting of esters in such a way as to regenerate the parent acid and alcohol is an example of hydrolysis. It is important, especially in dealing with naturally occurring esters such as those found in animal and vegetable fats, oils, and waxes. In the presence of dilute mineral acid, hydrolysis of an ester is the reverse of acid-catalyzed esterification; an excess of water is used to ensure complete splitting, and the reaction is carried out at elevated temperatures to speed up the process. Often alcohol is added to solubilize the reactants. Esters formed from glycerol and long-chain carboxylic acids (fats and oils), from long-chain acids and long-chain alcohols (waxes), and simple esters of mono-, di-, or polycarboxylic acids with primary, secondary, or tertiary alcohols, are hydrolyzable under acid conditions, using dilute hydrochloric or sulfuric acids or Twitchell's reagent (prepared from benzene or naphthalene, oleic acid, and concentrated sulfuric acid). However, esters of di-ortho-substituted aromatic carboxylic acids (for example, 2,6-dimethylbenzoic acid) are hindered with respect to hydrolysis, and must be treated according to the Newman technique, which involves first solution in 100% sulfuric acid, and then addition to excess cold water.

The reaction of an ester with a base to form an alcohol and salt of the acid is a type of hydrolysis historically called saponification. Ordinary household soaps are thus made from natural fats and oils of plant or animal origin. They are typically mixtures of the sodium salts of C_{12} and higher fatty acids, and a by-product is glycerol, reaction (11).

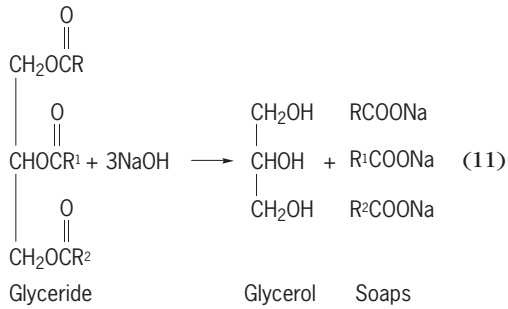

(11)

See DETERGENT; SOAP.

Catalytic reduction of esters can be effected at elevated temperature (250°C or 480°F) and pressure (15–20 atm or 1.5–2.0 megapascals), using molecular hydrogen and a copper chromite catalyst; this furnishes a convenient means for the preparation of long-chain mono- or dihydroxy alcohols from ester of the corresponding mono- or dicarboxyic acids. Thus, diethyl succinate ($C_2H_5OOC—CH_2CH_2—COOC_2H_5$) is reduced to form ethyl alcohol (C_2H_5OH) and butylene glycol ($HO—CH_2CH_2CH_2CH_2—OH$). For laboratory reductions, either sodium dissolving in alcohol (Bouveault-Blanc method), or lithium aluminum hydride is preferred.

Esters usually react well with Grignard reagents to yield tertiary alcohols. *See* GRIGNARD REACTION.

Acetoacetic ester synthesis. An ester of special importance in laboratory synthesis is ethyl acetoacetate, often called acetoacetic ester. Upon treatment with sodium ethoxide, acetoacetic ester forms the salt, ethyl sodioacetoacetate, which reacts with a primary alkyl bromide or iodide (RX) to form the alkyacetoacetate. Hydrolysis and decarboxylation of the latter product provides a general route to the

synthesis of methyl ketones, reaction (12).

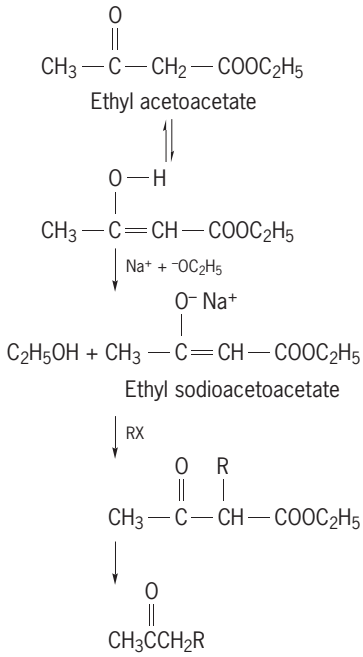

(12)

See KETONE.

Malonic ester synthesis. Ethyl malonate reacts with a strong organic base such as sodium ethoxide to form the sodiomalonate. Upon treatment with a primary alkyl halide (RX), a substituted malonic ester is

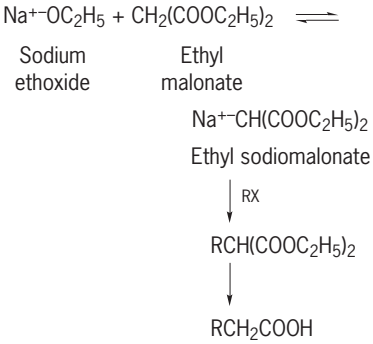

(13)

formed, which upon hydrolysis and decarboxylation yields a carboxylic acid. A variety of carboxylic acids have been prepared by reaction sequence (13).

Paul E. Fanta

Bibliography. R. T. Morrison and R. N. Boyd, *Organic Chemistry*, 6th ed., 1992; M. B. Smith and J. March, *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*, 5th ed., 2001.

Estimation theory

A branch of probability and statistics concerned with deriving information about properties of random variables, stochastic processes, and systems based on observed samples. Some of the important applications of estimation theory are found in control and communication systems, where it is used to estimate the unknown states and parameters of the system. For example, the position and velocity of a satellite is estimated from ground radar observations of its range, elevation, and azimuth. These observations are contaminated with random noise due to atmospheric propagation and radar circuitry. The statistical properties of random noise are assumed known except for some parameters which can be estimated from the data. Generally, the random noise is assumed to have a gaussian distribution, and its mean and covariance may be known or unknown. It is also assumed to be "white," that is, uncorrelated from one time instant to the next. The integral of white noise is a Wiener process or brownian motion process which plays a fundamental role in the theory of stochastic processes. *See* DISTRIBUTION (PROBABILITY); ELECTRICAL NOISE; STOCHASTIC PROCESS.

The estimation problem for dynamic systems may be divided into two parts: parameter estimation and state estimation. The basic difference between a parameter and the state is that the former either does not change at all or changes slowly in time, whereas the latter continuously evolves in time. For example, the state of a satellite is a six-dimensional vector consisting of three position variables and three velocity variables along the axes of an orthogonal coordinate system. The parameters of the satellite are its mass, inertia, and so on. In many control and communication problems, some of the system parameters are not known with desired accuracy. The problem of estimating these parameters from observed data is called parameter identification, though it is basically a problem of estimation. The more general problem of developing a mathematical model of the system from observed data is called system identification. On the other hand, the problem of state estimation is described by names such as signal processing, filtering, and smoothing. The problem belongs to the theory of stochastic processes and is also commonly known as time series analysis.

Approaches to estimation. Three basic approaches used for estimation are least-squares, maximum-likelihood, and bayesian. An estimator is defined as a function of the observations possessing certain desirable properties such as unbiasedness, consistency, and minimum variance. *See* STATISTICS.

Least-squares estimation. This was first used by the German mathematician K. F. Gauss in 1821. It is one of the most widely used techniques for estimation. In using the least-squares technique, it is postulated that the observations $\{y(t): t = 1, 2, \ldots, T\}$ are related to the unknown parameters, denoted by a vector θ, defined by Eq. (1), where $v(t)$ is a random error which

$$y(t) = f(x_t, \theta, t) + v(t) \qquad t = 1, 2, \ldots, T \tag{1}$$

is white and has zero mean, and x_t is a set of predetermined variables. The least-squares estimate of θ is one that minimizes a scalar function J defined by Eq. (2), where $\sigma_v^2(t)$ is the variance of $v(t)$.

$$J = \sum_{t=1}^{T} \left[y(t) - f(x_t, \theta, t)\right]^2 \frac{1}{\sigma_v^2(t)} \tag{2}$$

See LINEAR ALGEBRA.

The estimate is generally denoted as $\hat{\theta}_{LS}$ and is obtained by numerical minimization of J with respect to θ. The case in which the function $f(x_t,\theta,t)$ is linear

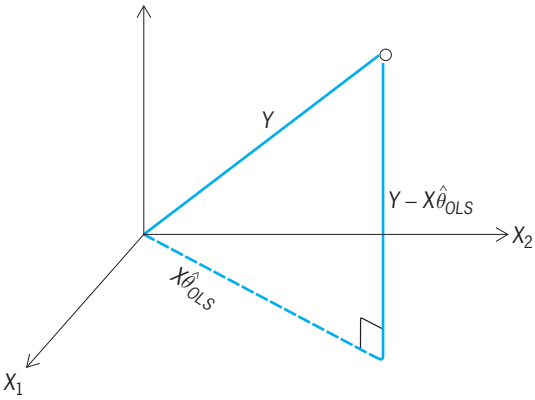

Fig. 1. Least-squares estimation for $T = 3$ and dimension of $\theta = 2$. Error vector is orthogonal to space spanned by columns of X.

in θ and σ_v^2 constant is of special importance and is known as ordinary least-squares. Equation (1) takes the form of Eq. (3).

$$y(t) = x_t^T\theta + v(t) \qquad t = 1, \ldots, T \tag{3}$$

Let Y denote a $T \times 1$ vector consisting of all observations $\{y(1), \ldots, y(T)\}$, where T (the number of observations) is greater than the dimension of θ (the number of unknown parameters). Similarly, defining X and V, Eq. (3) may be written as a single vector equation (4).

$$Y = X\theta + V \tag{4}$$

It is easily shown that for this case J is minimized by the estimator of Eq. (5).

$$\hat{\theta}_{\mathrm{OLS}} = (X^TX)^{-1}X^TY \tag{5}$$

See MATRIX THEORY.

The ordinary least-squares estimator has several interesting properties. It can be shown that in the class of all unbiased linear estimators, $\hat{\theta}_{\mathrm{OLS}}$ has the minimum variance. Furthermore, the error vector $(Y - X\hat{\theta}_{\mathrm{OLS}})$ is orthogonal to the space spanned by the columns of X (**Fig. 1**).

The best estimate of Y is $X\theta_{\mathrm{OLS}}$ which is a projection of Y on the space spanned by the columns of X. The concept of orthogonal projection is more general and carries over to infinite-dimensional spaces that are involved in state estimation. *See* LEAST-SQUARES METHOD.

Maximum-likelihood estimation. This was first used by R. A. Fisher in 1906 and is based on the concept of a likelihood function. Consider the conditional density function $p(Y|\theta)$, where Y represents the set of observations, and θ the set of unknown parameters. In general, $p(Y|\theta)$ is a function of both Y and θ, but if Y is set equal to its observed value, say Y^*, then $p(Y^*|\theta)$ regarded as a function of θ is known as the likelihood function. A justification for this name comes from the fact that if $p(Y^*|\theta)$ is twice as large for $\theta = \theta_1$ as it is for $\theta = \theta_2$, then the likelihood that θ_1 is the true value is twice as large as θ_2 being the true value. A maximum-likelihood estimate of θ is $\hat{\theta}_{\mathrm{ML}}$ if $p(Y^*|\hat{\theta}_{\mathrm{ML}})$ is the maximum of $p(Y^*|\theta)$ with respect to all possible values of θ. *See* PROBABILITY.

For example, if θ can take only two values, θ_1 or θ_2, and Y is a scalar, then $p(Y|\theta_1)$ and $p(Y|\theta_2)$ can be plotted as functions of Y (**Fig. 2**). If the observed value of Y is y, then θ_1 is the maximum-likelihood estimate (MLE) since $p(y|\theta_1) > p(y|\theta_2)$. Similarly, if y' is observed, then θ_2 is the maximum-likelihood estimate. But if y'' is observed, there is an ambiguity. If $p(Y|\theta_1) = p(Y|\theta_2)$ for all values of Y, then θ is said to be unidentifiable from Y.

When $p(Y|\theta)$ is gaussian, then the maximization of $p(Y|\theta)$ is equivalent to the minimization of a quadratic function in Y which is nothing but the least-squares criterion. Thus, for gaussian distributions, the maximum-likelihood estimate and least-squares estimate are equivalent.

Bayesian estimation. In least-squares and in maximum-likelihood estimation, θ is regarded as an unknown but constant parameter. In bayesian estimation, θ is regarded as a random variable and is assigned a prior probability $p(\theta)$. The probability $p(\theta)$ may be objective, that is, based on prior data, or subjective, based on a statistician's judgment. Once Y is observed according to a probability law $p(Y|\theta)$, Bayes' rule is used to obtain the posterior probability $p(\theta|Y)$ according to Eq. (6).

$$p(\theta|Y) = \frac{p(Y|\theta)p(\theta)}{\int p(Y|\theta)p(\theta)\,d\theta} \tag{6}$$

The observed value of Y is used in Eq. (6) so that $p(Y|\theta)$ is nothing but the likelihood function. The posterior probability $p(\theta|Y)$ summarizes all information regarding θ that can be obtained from Y. It is, therefore, much more informative than a point estimate of θ such as $\hat{\theta}_{\mathrm{LS}}$ or $\hat{\theta}_{\mathrm{ML}}$. On the other hand, computations of Eq. (6) are more tedious than those of point estimates. But there are cases in which these computations can be simplified.

In the gaussian case, if θ represents the conditional mean of Y and the prior distribution $p(\theta)$ is taken as gaussian, it is easily shown that $p(\theta|Y)$ is also gaussian and that Eq. (6) can be simplified to two equations, one for the mean of $p(\theta|Y)$ and the other for the covariance of $p(\theta|Y)$. If a sequence of observations are made, namely $\{y(1), \ldots, y(N)\}$, Eq. (6) can be used repeatedly to obtain $p(\theta|y(1), \ldots, y(N))$. The same procedure when used for state estimation in linear dynamic systems leads to a Kalman filter, discussed below. *See* BAYESIAN STATISTICS.

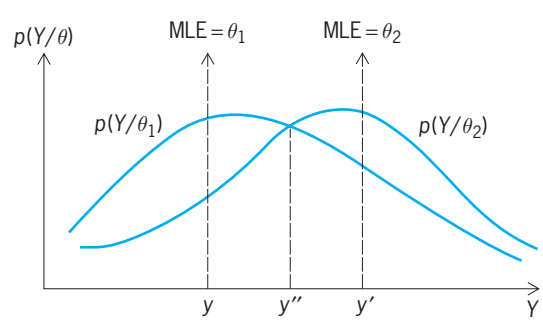

Fig. 2. Maximum-likelihood estimation (MLE).

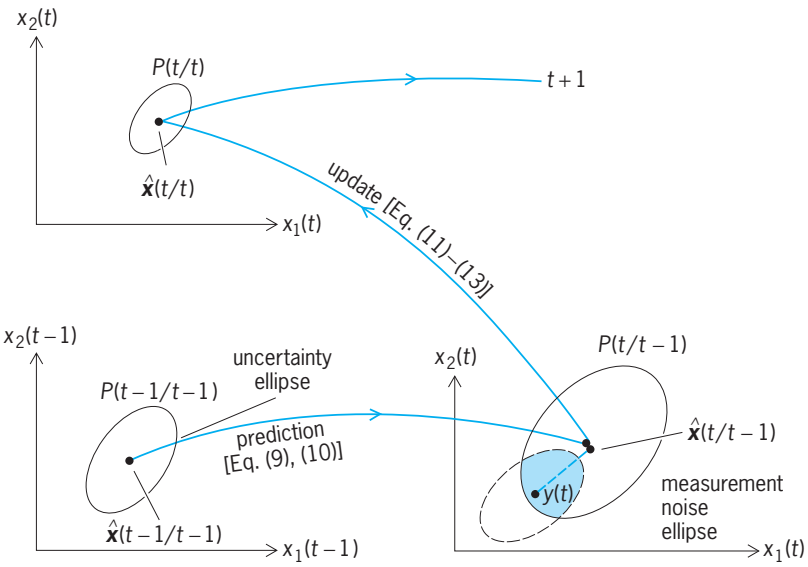

Fig. 3. Discrete-time Kalman filter in two dimensions. The steps involved in obtaining ($\hat{x}(t|t)$, $P(t|t)$) from ($\hat{x}(t-1|t-1)$, $P(t-1|t-1)$) are shown for a two-dimensional case.

Gauss-Markov processes. A stochastic process $x(t)$ is Markov if given its present state, its future is independent of its past. A fairly general model for Gauss-Markov processes is the Gauss-Markov sequence given by Eqs. (7) and (8), where $x(t)$ is an n-

$$x(t+1) = \Phi(t)x(t) + \Gamma(t)w(t) \tag{7}$$

$$y(t) = H(t)x(t) + v(t) \tag{8}$$

$$t = 0, 1, 2, \ldots, N$$

dimensional state vector and $w(t)$ is a q-dimensional zero-mean gaussian white noise sequence with the property that $w(t)$ and $w(\tau)$ are independent if $t \neq \tau$. The sequence $w(t)$ is called process noise, and $v(t)$ which has similar properties is called measurement noise; $y(t)$ is a p-dimensional vector of observations and $\Phi(t)$ is known as the transition matrix of the process. The initial state $x(0)$ also has a gaussian distribution.

The model of Eqs. (7) and (8) is also called a state vector model and arises naturally, in many cases, from a physical model. For example, if $x(t)$ consists of position and velocity components of a moving object subject to random shocks $w(t)$, then for small durations, Eqs. (7) and (8) would describe its motion. The state $x(t)$ takes on different meanings in different contexts, but it always represents the collection of all the information from the past and present behavior of the system that is sufficient to predict the future behavior of the system.

Three state estimation problems arise in connection with Eqs. (7) and (8). Let $\hat{x}(t|\tau)$ represent an estimate of $x(t)$ based on the observation set $\{y(1), \ldots, y(\tau)\}$. Based on the values of t and τ, the following names are given to the estimates: filtered estimate, $t = \tau$; predicted estimate, $t > \tau$; smoothed estimate, $t < \tau$.

Kalman filter. A Kalman filter provides estimates that are optimal in the least-squares, maximum-likelihood, and bayesian senses for the Gauss-Markov model.

Discrete-time case. A discrete-time Kalman filter is a set of recursive equations for the mean and covariance of the gaussian conditional density function $p(x(t)|y(t), \ldots, y(1))$ for $t = 1, 2, \ldots, N$. Let $\hat{x}(t|t)$ and $P(t|t)$ denote the mean and covariance at time t (**Fig. 3**). During the prediction phase, ($\hat{x}(t|t-1)$, $P(t|t-1)$) are obtained from Eqs. (9) and (10), where

$$\hat{x}(t|t-1) = \Phi\hat{x}(t-1|t-1) \tag{9}$$

$$P(t|t-1) = \Phi P(t-1|t-1)\Phi^T + \Gamma Q \Gamma^T \tag{10}$$

Q is the covariance of the process noise $w(t)$.

During the update phase, Bayes' rule, Eq. (6), is applied and gives Eqs. (11)–(13), where R is the co-

$$\hat{x}(t|t) = \hat{x}(t|t-1) + K(t)(y(t) - H\hat{x}(t|t-1)) \tag{11}$$

$$K(t) = P(t|t-1)H^T(HP(t|t-1)H^T + R)^{-1} \tag{12}$$

$$P(t|t) = (I - K(t)H)P(t|t-1) \tag{13}$$

variance of the measurement noise $v(t)$. Here $\hat{x}(0|0)$ and $P(0|0)$ are given as the initial distribution of $x(0)$.

Continuous-time case. The continuous-time analog of the model of Eqs. (7) and (8) is a linear differential equation forced by a white noise process, Eqs. (14) and (15).

$$\frac{dx(t)}{dt} = Fx(t) + Gu(t) \tag{14}$$

$$y(t) = Hx(t) + v(t) \qquad 0 \leq t \leq T \tag{15}$$

Here $x(0)$ is normally distributed with mean $\hat{x}_0$ and covariance P_0. The correlation functions of $u(t)$ and $v(t)$ are Dirac delta functions, $\delta(t - \tau)$, which are zero for $t \neq \tau$ and integrate out to one. These processes are also uncorrelated with $x(0)$ and with each other.

A simple way of deriving the continuous-time Kalman filter is to discretize Eqs. (14) and (15) with a small time-step Δt and express them in the form of Eqs. (7) and (8). Then the discrete-time Kalman filter equations (9)–(13) are used and the limit $\Delta t \to 0$ is taken. The filtering equations (16)–(19)are obtained.

$$\frac{d\hat{x}(t)}{dt} = F\hat{x}(t) + K(t)\,(y(t) - H\hat{x}(t)) \tag{16}$$

$$\frac{dP(t)}{dt} = FP(t) + P(t)F^T + GQG^T - P(t)H^TR^{-1}HP(t) \tag{17}$$

$$K(t) = P(t)H^TR^{-1} \tag{18}$$

$$\hat{x}(0) = x_0 \qquad P(0) = P_0 \tag{19}$$

A Kalman filter can be represented as a signal-flow block diagram (**Fig. 4**).

Both the discrete-time and continuous-time models can readily be generalized to the problems of prediction and smoothing.

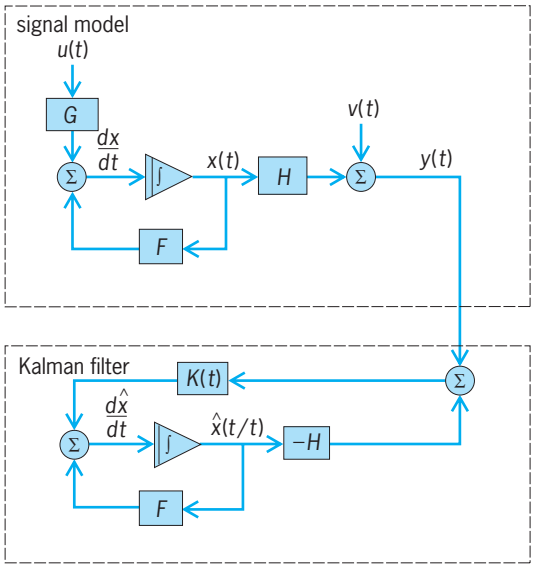

Fig. 4. Block diagram of a continuous-time Kalman filter and the signal model.

Other state estimation methods. The first treatment of the prediction problem was due to A. Kolmogorov (1941) for the discrete-time case. N. Wiener (1942) developed the continuous-time filter independently in the frequency domain. Both of these approaches considered stationary processes and made use only of the correlation or spectral properties of the stochastic processes. R. E. Kalman (1960), on the other hand, used the state vector model which can describe both stationary and nonstationary processes and can incorporate physical information regarding the system. The Kalman filter is equivalent to the Kolmogorov-Wiener filter for the stationary case.

Nonlinear filters. If the system equations (7) and (8) or (14) and (15) are nonlinear, the filtering density function is non-gaussian and it is necessary to evaluate higher-order moments or to propagate the complete density function. In the continuous-time case, this leads to partial differential equations which are very time-consuming to solve. A more practical, though suboptimal, approach is to linearize the nonlinear equations around the latest filtering and prediction estimates. This leads to an extended Kalman filter. More complicated versions are (1) the iterated extended Kalman filter in which the measurement nonlinearity effects are alleviated further by repeated linearization and update, (2) gaussian second-order filters in which certain bias terms due to the nonlinearities in the extended Kalman filter are eliminated, and (3) iterated smoothing filters in which one or more backward-step smoothing estimates are used for linearization.

Applications. Since their introduction in 1960, Kalman filters and their extensions have found numerous applications. Initially, these filters were developed for space applications such as satellite orbit determination, inertial navigation, Apollo lunar landing module guidance, and so on. The applications to power systems and industrial processes were developed shortly thereafter. Kalman filters have been used for forecasting, water quality prediction, hurricane tracking, and aircraft landing systems. Another area for application of Kalman filters is stochastic control, that is, control of systems with random inputs. It can be shown that an optimal controller for the so-called LQG systems (linear dynamics, quadratic criterion, and gaussian disturbances) consists of a Kalman filter cascaded with a linear-state feedback device. This concept has been used in the design of flight controllers, submarine depth-keeping control, process control, and pilot modeling, and in the study of optimal economic stabilization policies. *See* FLIGHT CONTROLS; GUIDANCE SYSTEMS; INERTIAL GUIDANCE SYSTEM; OPTIMAL CONTROL (LINEAR SYSTEMS); PROCESS CONTROL; SPACE NAVIGATION AND GUIDANCE; STOCHASTIC CONTROL THEORY.

Raman K. Mehra

System identification. This is the determination of a mathematical model of a system from the observed input and output data. System identification most often refers to dynamical system models, although this term can be extended to static input-output models. A dynamic system model is usually represented by differential equations (for a continuous-time model) or by difference equations (for a discrete-time model). Techniques to determine the model from measured data involve several key steps. The first step is the selection of the model type, including the selection of a continuous-time or discrete-time model; the selection of a linear or nonlinear model; the selection of the order of the model; and deciding whether there is prior knowledge of the system, or such knowledge is lacking, and parameters are to be determined from measurements. The other steps are the design of the experiment, including the choice of the experiment type as batch processing or recursive processing, and the choice of the input signal; the determination of parameter values in the model that provide the best fit to data (the parameter-estimation problem); and model validation. The selection of the model type is quite application dependent, while the remaining three steps involve one of numerous mathematical techniques for parameter estimation. *See* DIFFERENCE EQUATION; DIFFERENTIAL EQUATION; MODEL THEORY.

Model selection depends largely on the level of prior knowledge of the system. In some situations this knowledge may be sufficient to derive the dynamical model from known principles of physics, with uncertainty only about the values of some parameters in the model. The choice of the model in this case is then dictated by known laws of physics. This is the case of the identification of a partially known system. In other situations the knowledge of the system may be insufficient to derive a system of model equations. In this case the form of the model may be assumed, with all the parameters in it uncertain. The values of all the parameters are then estimated from input and output measurements. In contrast to the previous case, the parameters may have no physical meaning. An example of such a model is the linear discrete-time system given by Eq. (20).

$$y(t) + a_1 y(t-1) + \cdots + a_n y(t-n) = b_1 u(t-1) + \cdots + b_m u(t-m) + v(t) \quad (20)$$

This model has a single input $u(t)$, a single output $y(t)$, and noise $v(t)$, and there is no prior knowledge of the parameters $a_1, \ldots, a_n, b_1, \ldots, b_m$.

Batch versus recursive identification. If the input and output signals are prerecorded and the parameter-estimation procedure is carried out at some later time, the process is called a batch or off-line identification. The advantage of batch identification is that it is possible to try several models and find the one that produces the best fit to data. However, the identification results are not available in time for active control of the system while the signals are being recorded.

If the identification is carried out during the process operation and the parameter estimates are being modified as new measurements arrive, such a process is called a recursive or on-line identification. Recursive identification can be included in the feedback control algorithm, thus producing an adaptive control law, which can be used to control systems whose models are not completely known initially. *See* ADAPTIVE CONTROL.

Least-squares identification. As an example of system identification, the parameters of the model given by Eq. (20) can be determined by using the least-squares method in the following manner:

If two vectors are defined by Eqs. (21) and (22),

$$\phi(t) = [-y(t-1), \ldots, -y(t-n), u(t-1), \ldots, u(t-m)]^T \quad (21)$$

$$\theta = [a_1 \ldots a_n b_1 \ldots b_m]^T \quad (22)$$

then Eq. (20) can be rewritten as Eq. (23). The vector

$$y(t) = \theta^T \phi(t) + v(t) \quad (23)$$

of parameters θ can then be estimated by the least-squares method, as described above. This would result in a batch identification of the model. A recursive version of the same method is given by Eq. (24),

$$\hat{\theta}(t) = \hat{\theta}(t-1) + L(t)[y(t) - \hat{\theta}^T(t-1)\phi(t)] \quad (24)$$

where $\hat{\theta}(t)$ denotes the estimate of θ at time t.

The term $[y(t) - \hat{\theta}^T(t-1)\,\Phi(t)]$ represents the mismatch between the actual output $y(t)$ and the one predicted by using the previous parameter estimate $\hat{\theta}(t-1)$. The $L(t)$ is a column of $n + m$ time-dependent coefficients which dictate the relative size of the update made in each component of $\hat{\theta}$ from step $t-1$ to step t. The vector $L(t)$ is recursively computed from another formula (not shown here) in such a way that the sum of squared mismatch terms is minimal (the least-squares method).

The above is known as the recursive least-squares (RLS) algorithm. This is one of the most widely known, robust, and easy-to-implement algorithms.

The minimum number of data points required for the algorithm is equal to the dimension of the vector θ, that is, $n + m$. The algorithm requires initialization. The common choice of initial value is $\hat{\theta}(0) = 0$. If the sequences $v(t)$ and $\Phi(t)$ are uncorrelated, the estimate $\hat{\theta}(t)$ will be unbiased.

Choice of input signal. A recursive parameter estimation algorithm may not necessarily converge if the input signal is poorly chosen. A condition known as the persistent excitation condition is sufficient (but not necessary) for the convergence of several algorithms, including the RLS algorithm. In the case of model given by Eq. (20) with $m = n$, the persistent excitation condition requires that the power spectral density of the input signal $u(t)$ be nonzero at $2n$ or more frequencies. This condition is satisfied by a signal that contains at least $2n$ distinct frequencies, and also by white noise.

One particularly useful choice of the input signal is pseudorandom binary noise. This is a waveform switching between two values, for example, $+1$ and -1, according to a deterministic sequence designed in such a way that the autocorrelation function of the waveform approximates that of a white noise. Pseudorandom binary noise of small amplitude can be added to a constant input, thus providing a signal that does not disrupt normal system operation.

Applications of system identification. System identification is one of the most frequently applied parts of control systems theory. Both batch and recursive estimation algorithms have been used since the 1960s in many industries, in particular in the chemical and petroleum industries, paper-machine control, and ship steering. System identification subsequently became more closely integrated in the overall control loop in the form of adaptive control laws. Adaptive control has been applied to many problems in process control, motion control, and the control of aircraft and spacecraft.

Closely related to the problem of identification is the problem of failure detection in a dynamic system. It is a problem of determining a sudden change in a parameter value caused by a failure of one component in the system. This is usually accomplished by using algorithms which compare residuals in several redundant input-output relations. *See* CONTROL SYSTEMS; FAULT ANALYSIS; SYSTEMS ENGINEERING.

Andre Z. Manitius

Bibliography. S. M. Bozic, *Digital and Kalman Filtering*, 2d ed., 1995; R. G. Brown and P. Y. Hwang, *Introduction to Random Signal Analysis and Kalman Filtering*, 3d ed., 1996; H. F. Chen and L. Guo, *Identification and Stochastic Adaptive Control*, 1991; M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice*, 1993; L. Ljung, *System Identification: Theory for the User*, 2d ed., 1999; J. M. Mendel, *Lessons in Estimation Theory*, 1995; A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3d ed., 1991; H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, 2d ed., 1994.

Estrogen

A substance that maintains the secondary sex characters and organs, such as mammary glands, uterus, vagina, and fallopian tubes, of mammalian females. Naturally occurring substances with this activity

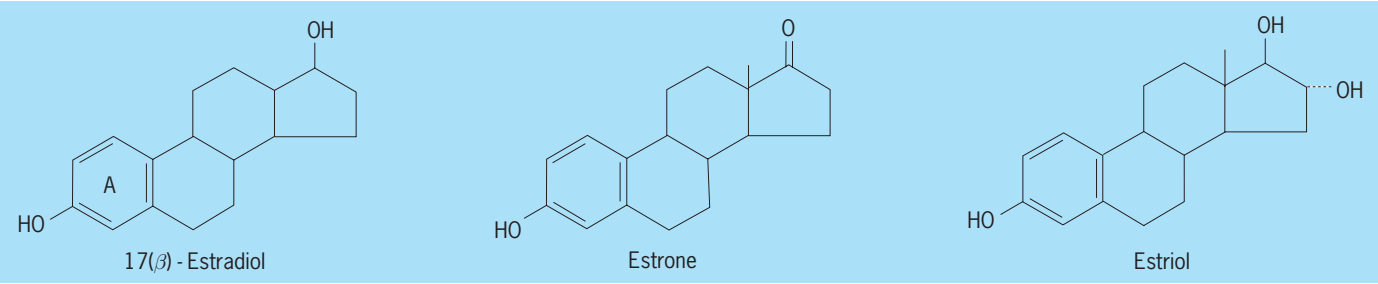

Fig. 1. Structural formulas of estrogenic hormones with ring A phenolic group.

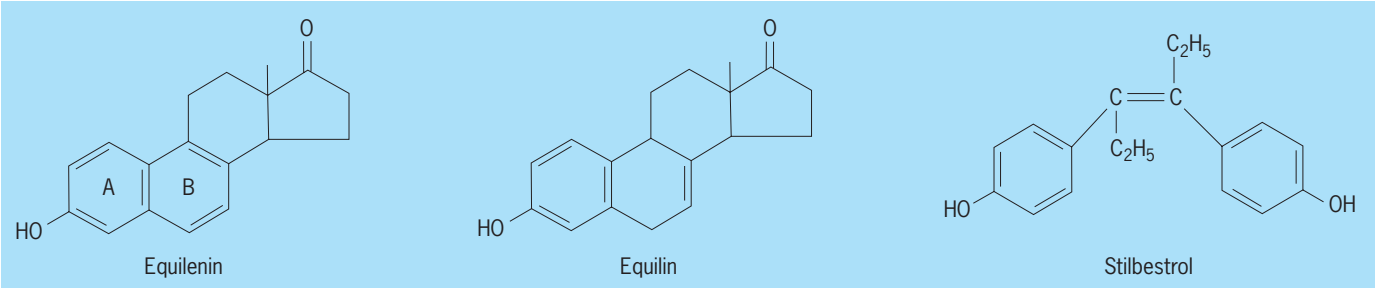

Fig. 2. Structural formulas of estrogenic hormones with ring B unsaturation.

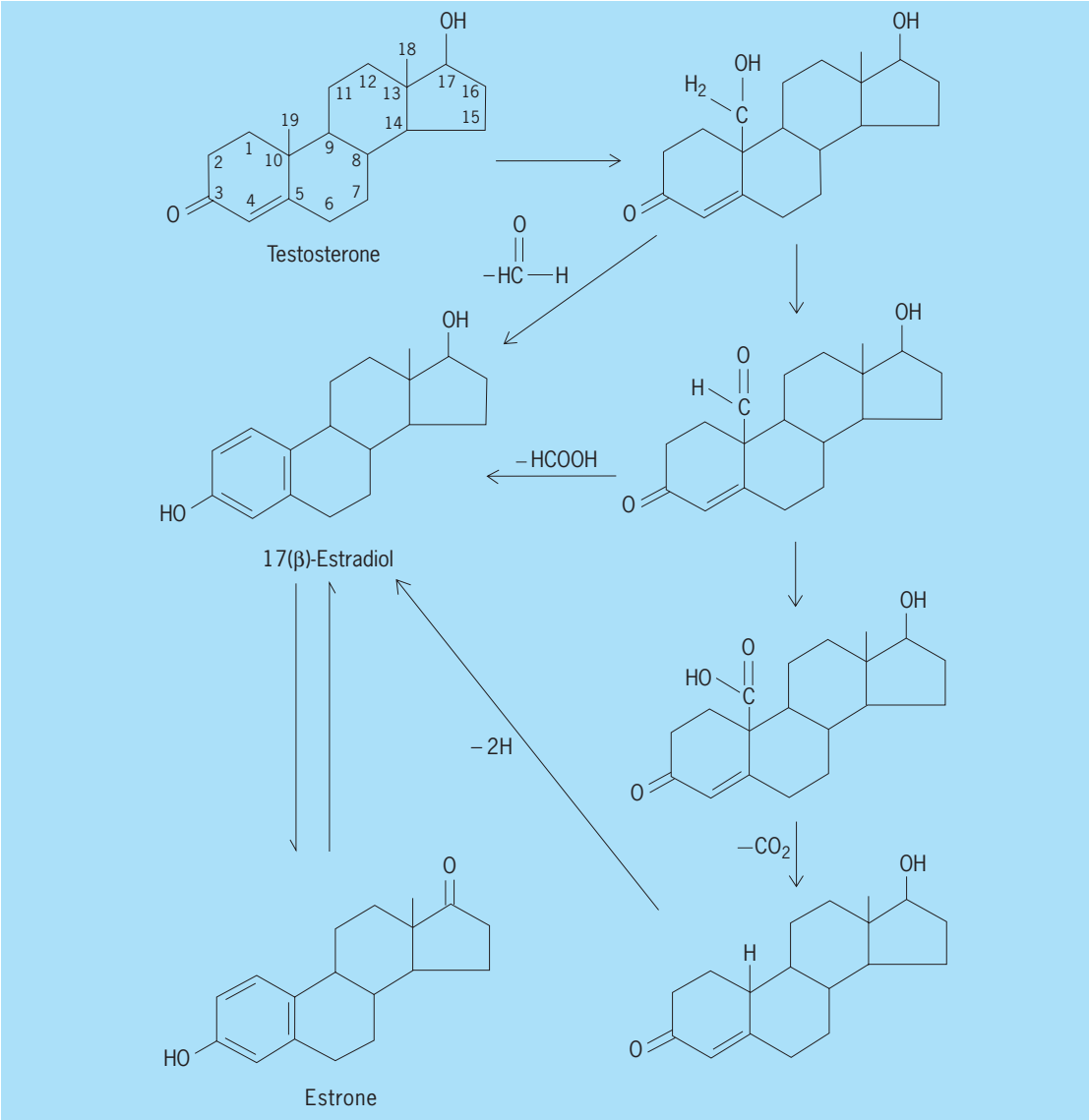

Fig. 3. Pathways for the biosynthesis of estrogens from testosterone.

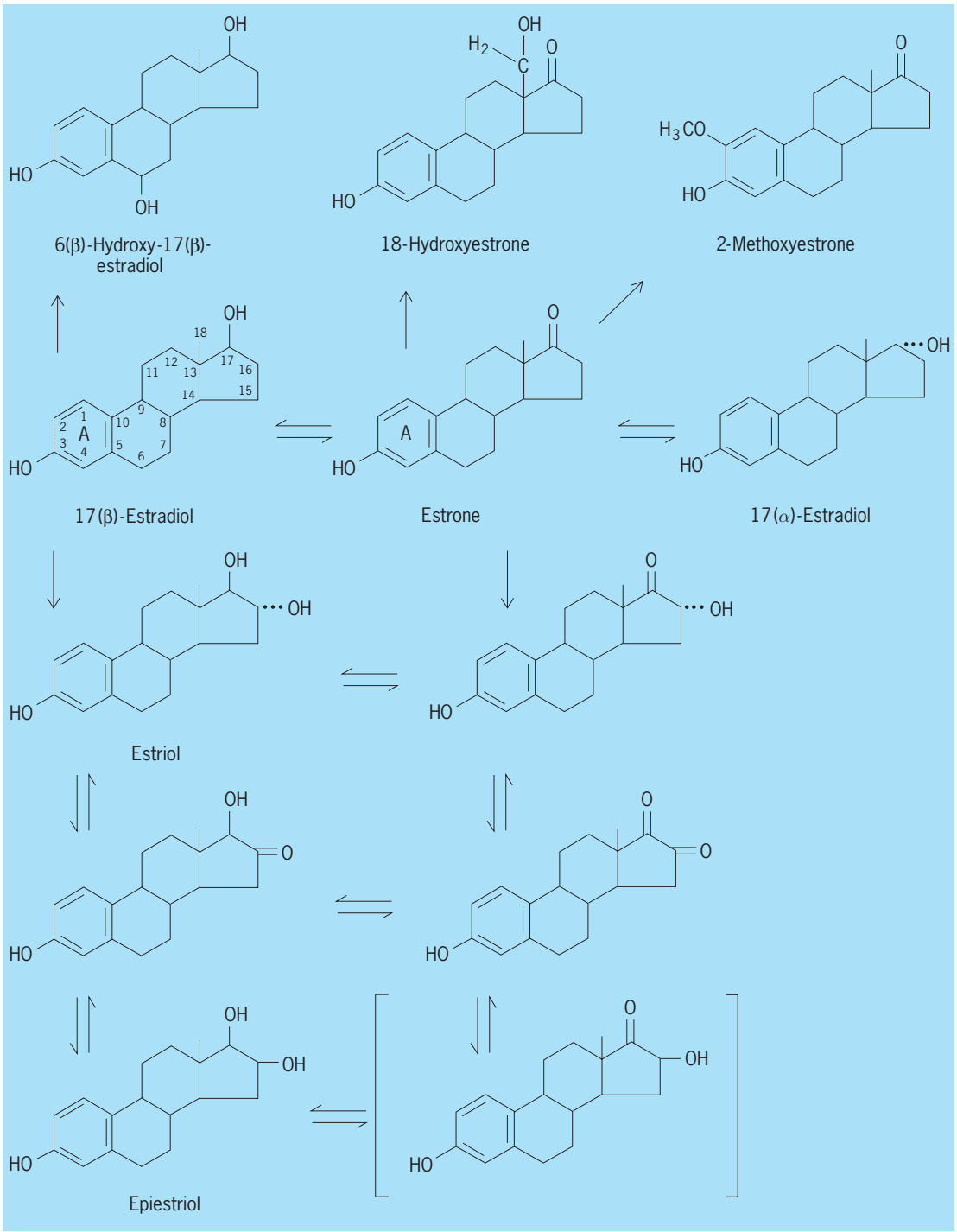

Fig. 4. Reactions involved in the metabolic breakdown of estrogens.

are steroid hormones. The principal estrogenic hormone substances are 17(β)-estradiol, estrone, and estriol. These compounds have a ring A phenolic group (**Fig. 1**). They are produced and secreted directly into the bloodstream by the ovary, testis, adrenal, and placenta of pregnancy. Two other naturally occurring estrogenic hormones, equilin and equilenin (**Fig. 2**), have been obtained only from the urine of pregnant mares and are apparently peculiar to that species. Stilbestrol, a synthetic compound with considerable estrogenic activity, has been used extensively in medical practice (Fig. 2).

Ring A phenolic group estrogens. The most active of the estrogens known to be produced by mammalian tissues is 17(β)-estradiol which is interconvertible with estrone. An enzyme present in liver, placenta, and other tissues catalyzes this oxidation-reduction. Estriol is considered to be derived from estrone.

Ring B unsaturated estrogens. Equilenin and equilin are naturally occurring representatives of ring B unsaturated estrogens. Equilenin also has been detected in human adrenal cancer tissue from a patient with a

feminizing syndrome. Stilbestrol is a nonsteroid synthetic estrogen. It has a biological activity equal to that of 17(β)-estradiol by injection and a higher activity by oral administration.

Biosynthesis of estrogens. This involves a series of reactions from androgens such as testosterone and Δ^4-androstene-3,17-dione. The transformation of androgens to estrogens involves the removal of carbon 19 and two hydrogen atoms, one each at carbons 1 and 2. The loss of the carbon atom most likely occurs by reactions indicated in **Fig. 3**, using testosterone as the example.

Catabolism of estrogens. The metabolic breakdown of estrogens involves various modifications at carbon atoms 2, 6, 16, and 18. These overall reactions are illustrated in the series of equations in **Fig. 4**. *See* HORMONE; STEROID. Ralph I. Dorfman

Bibliography. B. Green and R. E. Leake, *Steroid Hormones: A Practical Approach*, 1987; V. H. James and J. R. Pasqualini (eds.), *Hormonal Steroids*, 1984; H. L. Makin, *Biochemistry of Steroid Hormones*, 2d ed., 1984; S. Raam (ed.), *Immunology of Steroid Hormone Receptors*, 1988.

Estrus

The period in mammals during which the female ovulates and is receptive to mating. It is commonly referred to as rut or heat. From one estrus period to the next there occurs a series of changes, particularly in the ovary, uterus, and vagina, termed the estrous cycle. With reference to the ovary, the cycle can be divided into a follicular phase, during which the Graffian follicles are ripening, and a luteal phase, during which the corpora lutea develop in the ovulated follicles. During these two phases mainly estrogen and progesterone, respectively, are secreted, and these hormones control the uterine and vaginal changes. The beginning of the follicular phase is termed proestrus and the luteal phase metestrus. Following the latter, there is a period of relatively little change, termed diestrus. In species in which the latter is prolonged, it is termed anestrus.

During proestrus the endometrial lining of both the uterus and vagina thickens. The vagina regresses in metestrus, and vaginal washings exhibit mainly cornified epithelial cells. At the same time uterine glands continue to enlarge and become highly secretory. This is the stage during which the embryo would implant in the uterine lining if the previously ovulated eggs had been fertilized. Toward the end of this period, in the absence of fertilization, the uterine endometrium regresses, gradually in most mammals, but precipitously in primates. In the latter the regression involves considerable loss of tissue and bleeding. *See* MENSTRUATION.

Some mammals, such as the armadillo, bat, deer, wolf, shrew, and ferret, are monestrous, having a single cycle per year. Most are polyestrous, but considerable variations occur in different species. In dogs there are 2–3 cycles per year. Cats cycle about every 2 weeks during breeding seasons. Cows, horses, and sheep illustrate seasonal polyestrus with $2\frac{1}{2}$- to 3-week cycles, while laboratory rats and mice maintain a 5- to 6-day schedule throughout the year. In the rabbit there is persistent estrus, and ovulation occurs in response to the stimulus of coitus. This mating-induced ovulation is known to occur also in the cat and ferret during their more restricted period of estrus, and it probably also occurs in a number of other mammals.

The regulation of the estrous cycle, and of breeding seasons, is brought about by feedback-type hormonal interactions between the ovary and pituitary gland, influenced in many cases by environmental factors, such as length of day. Thus, for example, if additional illumination is provided in the fall, ferrets can start breeding during the winter instead of spring; and if the sequence of night and day is changed, mice, which ordinarily mate and ovulate during the night, will adapt to the new conditions. Such environmental stimuli (light, coition) can affect the reproductive cycle through releasing hormones produced by cells in the hypothalamus, which are connected to neural pathways in the brain. Specific releasing hormones control the amounts of gonadotrophic hormones discharged by the anterior lobe of the pituitary into the bloodstream. *See* ENDOCRINE MECHANISMS; OVUM; PITUITARY GLAND; REPRODUCTION (ANIMAL).

Albert Tyler; Howard L. Hamilton

Estuarine oceanography

The study of the physical, chemical, biological, and geological characteristics of estuaries. An estuary is a semienclosed coastal body of water which has a free connection with the sea and within which the seawater is measurably diluted by freshwater derived from land drainage. Many characteristic features of estuaries extend into the coastal areas beyond their mouths; and because the techniques of measurement and analysis are similar, the field of estuarine oceanography is often considered to include the study of some coastal waters which, by the above definition, are not strictly estuaries. Also, semienclosed bays and lagoons exist in which evaporation is equal to or exceeds freshwater inflow, so that the salt content either is equal to that of the sea or exceeds it. Hypersaline lagoons have been termed negative estuaries, whereas those with precipitation and river inflow equaling evaporation have been called neutral estuaries. Positive estuaries, in which river inflow and precipitation exceed evaporation, form the majority.

Topographic classification. Embayments are the result of fairly recent changes in sea level. During the Pleistocene ice age, much of the seawater was locked up in continental ice sheets, and the sea surface stood about 100 m (330 ft) below its present level. In areas not covered with ice, the rivers incised their valleys to this base level. During the ensuing Flandrian Transgression, when the sea level rose at about 1 m (3.3 ft) per century, these valleys became inundated. Much of the variation in form of the resulting estuaries

depends on the volumes of sediment that the river or the nearby coastal erosion has contributed to fill the valleys.

Where river flow and sediment discharge were high, the valleys have become completely filled and even built out into deltas. Generally, deltas are best developed in areas where the tidal range is small and where the currents cannot easily redistribute the sediment the rivers introduce. They occur mainly in tropical and subtropical areas where river discharge is seasonally very high. The distributaries, or passes, of the delta are generally shallow, and often the shallowest part is a sediment bar at the mouths of the distributaries. The Mississippi and the Niger are examples of this type of delta.

Where sediment discharge was less, the estuaries are unfilled, although possibly they are still being filled. These are drowned river valleys or coastal plain estuaries, and they still retain the topographic features of river valleys, having a branching, dendritic, though meandering, outline and a triangular cross section, and widening regularly toward the mouth, which is often restricted by spits. River discharge tends to be reasonably steady throughout the year, and sediment discharge is generally small. These estuaries occur in areas of high tidal range, where the currents have helped to keep the estuaries clear of sediment. They are typical of temperate regions such as the east coast of North America and northwestern Europe, examples being the Chesapeake Bay system, the Thames, and the Gironde. *See* COASTAL LANDFORMS.

In areas where glaciation was active, the river valleys were overdeepened by glaciers and fiords were created. A characteristic of these estuaries is the rock bar or sill at the mouth that can be as little as a few tens of meters deep. Inside the mouth, however, the estuaries can be at least 600 m (1800 ft) deep and can extend hundreds of kilometers inland. Fiords are typical of Norway and the Canadian Pacific coast. *See* FIORD.

Another estuarine type is called the bar-built estuary. These are formed on low coastlines where extensive lagoons have narrow connecting passages or inlets to the sea. Within the shallow lagoons, the tidal currents are small, but the deep inlets have higher currents. Again, a sediment bar is generally present across the entrance. In tropical areas, the lagoons can be hypersaline during the hot season. They are typical of the southern United States and of parts of Australia.

Estuaries are ephemeral features since great alterations can be wrought by small changes in sea level. If the present ice caps were to melt, the sea level would rise an estimated 30 m (99 ft), and the effect on the form and distribution of estuaries would be drastic.

Physical structure and circulation. Within estuaries, the river discharge interacts with the seawater, and river water and seawater are mixed by the action of tidal motion, by wind stress on the surface, and by the river discharge forcing its way toward the sea. The difference in salinity between river water and seawater—about 35 parts per thousand—creates a difference in density of about 2%. Even though this difference is small, it is sufficient to cause horizontal pressure gradients within the water which affect the way it flows. Density differences caused by temperature variations are comparatively smaller. Salinity is consequently a good indicator of estuarine mixing and the patterns of water circulation. Obviously, there are likely to be differences in the circulation within estuaries of the same topographic type which are caused by differences in river discharge and tidal range. The action of wind on the water surface is an important mixing mechanism in shallow estuaries, particularly in lagoons; but generally its effect on estuarine circulation is only temporary, although it can produce considerable variability and thus make interpretation of field observations difficult. *See* SEAWATER.

Salt-wedge estuaries. Freshwater, being less dense than seawater, tends to flow outward over the surface of seawater, which penetrates as a salt wedge along the bottom into the estuary (**Fig. 1**). This creates a vertical salinity stratification, with a narrow zone of sharp salinity change, called a halocline, between the two water masses, which can reach 30 parts per

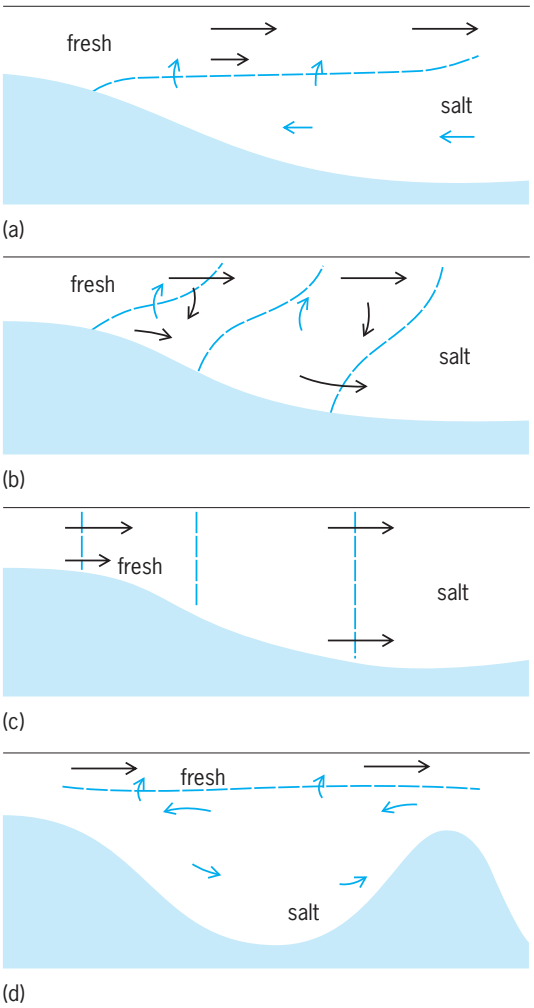

Fig. 1. Diagrams of mixing in estuaries. (*a*) Salt-wedge type. (*b*) Partially mixed type. (*c*) Well-mixed type. (*d*) Fiord.

thousand in 0.5 m (1.5 ft). If the sea is tideless, the water in the salt wedge is almost motionless. However, if the surface layer flowing toward the sea has a sufficiently high velocity, turbulent mixing can occur through a mechanism known as Kelvin-Helmholtz instability, which is a process where the denser salt water is drawn up into the overflowing freshwater in a coherent "rolled-up" pattern. Ultimately these instabilities break down completely, and the salt water is mixed entirely into the overlying fresh-water mass, increasing the salinity of the upper layer, before eventually being discharged to the ocean. This and similar processes are sometimes referred to as entrainment, which results in a net loss of fluid from the salt wedge. Consequently, for this loss to be replaced, there must be a compensatory flow of salt water toward the head, or landward portion, of the estuary within the salt wedge, but of a magnitude much less than that of the flow in the surface layer. There is a considerable velocity gradient near the halocline as a result of the friction between the two layers. Consequently, the position of the salt wedge will change according to the magnitude of the flow in the surface layer, that is, according to the river discharge. The Mississippi River is an example of a salt-wedge estuary. When the flow in the Mississippi is low, the salt wedge extends more than 160 km (100 mi) inland, but with high discharge the salt wedge extends only a mile or so above the river mouth. Some bar-built estuaries, in areas of restricted tidal range and at times of high river discharge, as well as deltas, are typical salt-wedge types.

Partially mixed estuaries. When tidal movements are appreciable, the whole mass of water in the estuary moves up and down with a tidal periodicity of about 12.5 h. Considerable friction occurs between the bed of the estuary and the tidal currents, and causes turbulence. The turbulence tends to mix the water column more thoroughly than occurs in salt-wedge estuaries, although little is known of the relationship of the exchanges to the salinity and velocity gradients. However, the turbulent mixing not only mixes the salt water into the fresher surface layer but also mixes the fresher water downward. This causes the salinity to decrease toward the head of the estuary in the lower layer and also to increase progressively toward the sea in the surface layer. As a consequence, the vertical salinity gradient is considerably less than that in salt-wedge estuaries. In the surface, seaward-flowing layer, the river discharge moves toward the sea; but because the salinity of the water has been increased by mixing during its passage down the estuary, the discharge at the mouth can be several times the river discharge. To provide this volume of additional water, the compensating inflow must also be much higher than that in the salt-wedge estuary. The velocities involved in these movements are only on the order of a few centimeters per second, but the tidal velocities can be on the order of 100 centimeters per second. Consequently, the only way to evaluate the effect of turbulent mixing on the circulation pattern is to average out the effect of the tidal oscillation, which requires considerable precision and care. The resulting residual or mean flow will be related to the river discharge, although the tidal response of the estuary can give additional contributions to the mean flow. The tidal excursion of a water particle at a point will be related to the tidal prism, the volume between high- and low-tide levels upstream of that point; and the instantaneous cross-sectional velocity at any time will be related to the rate of change of the tidal prism upstream of the section. In details, the velocities across the section can differ considerably. It has been found that in the Northern Hemisphere the seaward-flowing surface water keeps to the right bank of the estuary, looking downstream, and the landward-flowing salt intrusion is concentrated on the left-hand side (**Fig. 2**). This is caused by the Coriolis force, which deflects the moving water masses toward the right and is of increased importance for very wide estuaries. Of possibly greater importance, however, is the effect of topography, because the curves in the estuary outline tend to concentrate the flow toward the outside of the bends. Thus, in addition to a vertical circulation, there is a horizontal one, and the halocline slopes across the estuary. Because the estuary has a prismatic cross section, the saline water is concentrated in the deep channel and the fresher water is discharged in the shallower areas. Examples of partially mixed estuaries are the rivers of the Chesapeake Bay system.

Well-mixed estuaries. When the tidal range is very large, there is sufficient energy available in the turbulence to break down the vertical salinity stratification completely, so that the water column becomes

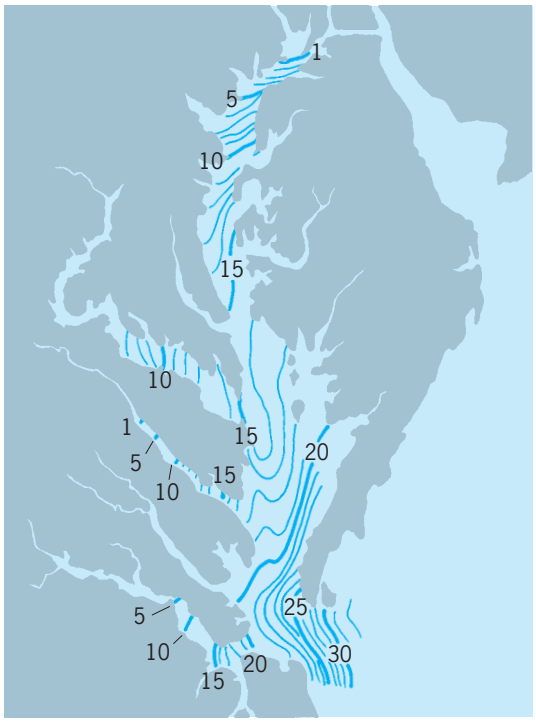

Fig. 2. Typical surface salinity distribution in Chesapeake Bay. Numbers indicate parts per thousand. (*After H. E. Landsberg, ed., Advances in Geophysics, vol. 1, Academic Press, 1952*)

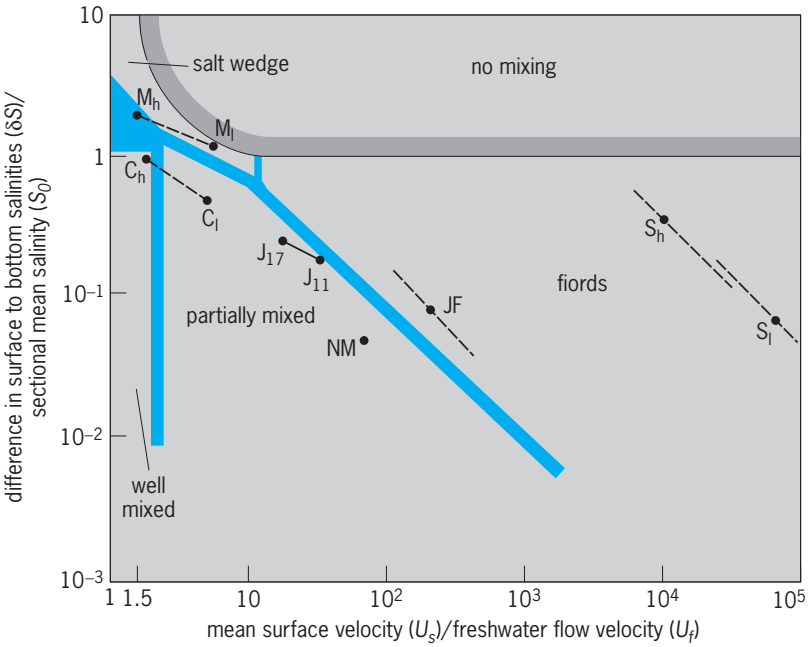

Fig. 3. Classification diagram for estuaries. An estuary appears as a line on the diagram; the upper reaches are less well mixed than the lower sections. Subscript letters refer to high (h) and low (l) river discharge; subscript numbers are distances from the mouth. J = James River; M = Mississippi; C = Columbia River; NM = Narrows of the Mersey; S = Silver Bay; JF = Strait of Juan de Fuca. (*After D. V. Hansen and M. Rattray, Jr., New dimensions in estuary classification, Limnol. Oceanogr., 11:319–326, 1966*)

vertically homogeneous. In this type of estuary there can be lateral variations in salinity and in velocity, with a well-developed horizontal circulation; or if the lateral mixing is also intense, the estuary can become sectionally homogeneous (also called a one-dimensional estuary). Because there is no landward residual flow in the sectionally homogeneous estuary, the upstream movement of salt is produced during the tidal cycle by salty water being trapped in bays and creeks and bleeding back into the main flow during the ebb. This mechanism spreads out the salt water, allowing salt to be maintained within the estuary, but it is probably an effective trapping mechanism for salt for only a small number of tidal excursions landward of the mouth.

Fiords. Because fiords are so deep and restricted at their mouths, tidal oscillation affects only their near-surface layer to any great extent. The amount of turbulence created by oscillation is small, and the mixing process is achieved by entrainment. Thus fiords can be considered as salt-wedge estuaries with an effectively infinitely deep lower layer. The salinity of the bottom layer will not vary significantly from mouth to head, and the surface fresh layer is typically no more than a few tens of meters deep. When the sill is deep enough not to restrict circulation, the inflow of water occurs just below the halocline, with an additional slow outflow near the bottom. When circulation is restricted, the replenishment of the deeper water occurs only occasionally, sometimes on an annual cycle and typically related to severe weather events. Between these replenishment episodes, the bottom layer can become anoxic, with very low dissolved oxygen (DO) levels.

The descriptive classification of estuaries outlined above depends on the relative intensities of the tidal and river flows and the effect that these flows have on stratification. A quantitative comparison between estuaries can be made using the diagram of **Fig. 3**, which is based on a stratification and a circulation parameter.

River plumes. The estuary represents the first stage of the blending of fresh river water into the sea. The water that is discharged from the estuary mouth typically has characteristics (such as salinity and temperature) that lie between those of the fresh river water and seawater, due to mixing and entrainment processes within the estuary. The process of blending with the coastal ocean continues in a region known as a river plume. With regard to the mixing and blending processes, a river plume can be divided into two regions. Certain estuaries may not exhibit a well-defined near-field region; and in some estuaries with particularly low freshwater flow, there may be no evidence of a plume at all.

A near-field plume is often considered an extension of the estuary out into the coastal ocean. It can be broadly defined as the region where the velocity of the outflowing estuarine water (related to river velocity) is sufficient to dominate the physical dynamics of the system. In this region, the less dense estuarine waters flow in a lens near the surface, spreading laterally, as well as mixing vertically with the ocean water underneath. The typical extent of a near-field region is approximately 1–3 km (0.6–2 mi).

The far-field plume begins at the outer boundary of the near-field region, or at the mouth if velocities are not sufficient to generate a near-field region, and often extends several hundreds of kilometers down the coast. In this region, the estuarine water continues to be mixed into the surrounding ocean by wind- and wave-generated processes. Because of effects caused by the Earth's rotation, a far-field plume will typically turn to the right outside of the river mouth and flow adjacent to the coast in the Northern Hemisphere or to the left in the Southern Hemisphere.

The mixing in the near-field region can be much more intense than the wind-driven mixing in the far-field plume, but acts across a very small region, compared to the size of the far field. Thus, several recent studies have suggested that the total amount of mixing occurring in each of the two regions may be roughly comparable.

Flushing and pollution-dispersal prediction. Much research into estuarine characteristics is aimed at predicting the distribution of effluents discharged into estuaries. Near the mouth of a partially mixed estuary, the salinity of the estuarine water is very near that of the adjacent ocean water, implying that the incoming fresh river water has been diluted significantly by mixing with a much larger volume of ocean water. Consequently, estuaries are more effective than rivers in diluting and removing pollutants. It has been observed that increased river flow causes both a downstream movement of the saline intrusion and a more rapid discharge of water to the

sea. The latter effect occurs because increased river discharge increases stratification; increased stratification diminishes vertical mixing and enhances the flow toward the sea in the surface layer. Thus, increased river discharge has the effect of increasing the volume of freshwater accumulated in the estuary, but to a lesser extent than the increase of the discharged volume. Obviously, it takes some time for the freshwater from the river to pass through the estuary. A rough estimate of the flushing time can be determined by dividing the total volume of freshwater accumulated in the estuary by the river flow. For most estuaries the flushing time is 5 to 10 days.

If a conservative, nondecaying pollutant is discharged at a constant rate into an estuary, the effluent concentration in the receiving water will vary with the tidal current velocity and will spread out by means of turbulent mixing. The concentrations will be increased during the next half cycle as the water passes the discharge point again. After several tidal cycles, a steady-state distribution will be achieved, with the highest concentration near the discharge point. Concentrations will decrease downstream but not as quickly as they do upstream. However, the details of the distribution will depend largely on whether the discharge is of dense or light fluid and whether the discharge is into the lower or upper layer. Since its movement will be modified by the estuarine circulation, the effluent will be more concentrated in the lower layer upstream of the discharge point, and it will be more concentrated in the upper layer downstream. To obtain maximum initial dilution, a light effluent would have to be discharged near the estuary bed so that it would mix rapidly as it rose.

For nonconservative pollutants, such as coliform sewage bacteria, prediction becomes more difficult. The population of bacteria dies progressively through the action of sunlight, and concentrations diminish with time as well as by dilution. The faster the mixing, the larger the populations at any distance from the point of introduction, since less decay occurs.

Because of the poor mixing of freshwater into a salt-wedge estuary, an effluent introduced in the surface layer will be flushed from the estuary before it contaminates the lower layer, provided that it is not too dense.

Mathematical modeling. Increasingly, mathematical modeling is being used, with reasonable success in many instances, to predict effluent dispersal with a minimum amount of field data. Although the governing mathematical equations can be stated, they cannot be solved in their full form because there are too many unknowns. To reduce the number of unknowns, various assumptions are made, including some form of spatial averaging to reduce a three-dimensional problem to two dimensions or even one dimension. Mixing parameters, about which little is known, are assumed constant or are considered as a simple variable in space, and are altered so that the model fits the available prototype data.

The first step is usually to model the flow and salinity distribution. Because the density field is important in determining the flow characteristics, density and flow are interlinked problems. Then, for pollutant studies the pollutant is assumed to act in the same manner as fresh or salt water, or the flow parameters are used with appropriate mixing coefficients to predict the distribution. Simple models consider the mean flow to be entirely the result of river discharge, and tidal flow to be given by the tidal prism. Segmentation is based on simple mixing concepts and crude mixing ratios. Salinity and pollutant concentrations can then be calculated for cross-sectionally averaged and vertically homogeneous conditions by using the absolute minimum of field data. These models are known as tidal prism models. One-dimensional models are very similar, but use a finer grid system and need better data for validation. Two-dimensional models assume vertical homogeneity and allow lateral variations, or vice versa. There are difficulties in including the effects of tidally drying areas and junctions; the models become more costly and require extensive prototype data, but they are more realistic. The ideal situation of modeling the flow and salinity distribution accurately simply on the basis of knowledge about the topography, the river discharge, and the tidal range at a number of points is still a long way off.

Ecological environments. Estuarine ecological environments are complex and highly variable when compared with other marine environments. They are richly productive, however. Because of the variability, fewer species can exist as permanent residents in this environment than in some other marine environments, and many of these species are shellfish that can easily tolerate short periods of extreme conditions. Motile species can escape the extremes. A number of commercially important marine forms are indigenous to the estuary, and the environment serves as a spawning or nursery ground for many other species.

River inflow provides a primary source of nutrients such as nitrates and phosphates which are more concentrated than in the sea. These nutrients are utilized by plankton through the photosynthetic action of sunlight. Because of the energetic mixing, production is maintained throughout, in spite of the high levels of suspended sediment which restrict light penetration to a relatively thin surface layer. Plankton concentrations can be extremely high, and then, higher levels of the food web—filter-feeding shellfish and young fish—have an ample food source. The rich concentrations provide large quantities of organic detritus in the sediments which can be utilized by bottom-feeding organisms and which can be stirred up into the main body of the water by tidal action. For a more complete treatment of the ecology of estuarine environments from the biological viewpoint *see* MARINE ECOLOGY.

The stratification present in estuaries tends to produce concentrated regions of detritus and microorganisms, which are attractive to other species as significant food sources. This can be particularly true where a region of strong stratification intersects the

bottom of the estuary, a region known as a front. Higher populations of organisms, from juvenile fish to seals, can often be seen congregating near fronts.

There is a close relationship between the circulation pattern in estuaries and the faunal distributions. Several species of plankton peculiar to estuaries appear to confine their distribution to the estuary by using the water-circulation pattern; pelagic larvae of oysters are transported in a similar manner. The fingerling fish (*Micropogon undulatus*), spawned in the coastal waters off the eastern coast of the United States, are carried into the estuarine nursery areas by the landward residual bottom flow.

Sediments. The patterns of sediment distribution and movement depend on the type of estuary and on the estuarine topography. The type of sediment brought into the estuary by the rivers, by erosion of the banks, and from the sea is also important; and the relative importance of each of these sources may change along the estuary. Fine-grained material will move in suspension and will follow the residual water flow, although there may be deposition and reerosion during times of locally low velocities. The coarser-grained material will travel along the bed and will be affected most by high velocities and consequently, in estuarine areas, will normally tend to move in the direction of the maximum current.

Fine-grained material. Fine-grained clay material, about 2 micrometers in size, brought down the rivers in suspension can undergo alterations in its properties in the sea. Base (cation) exchange with the seawater can alter the chemical composition of some clay minerals; also, because the particles have surface ionic charges, they are attracted to one another and can flocculate. Flocculation depends on the salinity of the water and on the concentration of particles. It is normally complete in salinities in excess of 4 parts per thousand, and with suspended sediment concentrations above about 300 parts per million (1 mg/liter), and has the effect of increasing the settling velocities of the particles. The flocs have diameters larger than 30 μm but effective densities of about 1.1 g/cm^3 because of the water closely held within. If the material is carried back into regions of low salinity, the flocculation is reversed and the flocs can be disrupted by turbulence. In sufficiently high concentrations, the suspended sediment can suppress turbulence. The sediment then settles as layers which can reach concentrations as high as 300,000 parts per million and which are visible as a distinctive layer of "fluid mud" on echo-sounder recordings. At low concentrations, aggregation of particles occurs mainly by biological action.

Turbidity maximum. A characteristic feature of partially mixed estuaries is the presence of a turbidity maximum. This is a zone in which the suspended sediment concentrations are higher than those in the river or farther down the estuary. This zone, positioned in the upper estuary around the head of the salt intrusion and associated with mud deposition in the so-called mud reaches, is often related to wide tidal mud flats and saltings. The position of the turbidity maximum changes according to changes in river discharge, and is explained in terms of estuarine circulation. Suspended sediment is introduced into the estuary by the residual downstream flow in the river. In the upper estuary, mixing causes an exchange of suspended sediment into the upper layer, where there is a seaward residual flow causing downstream transport. In the middle estuary, the sediment settles into the lower layer in areas of less vigorous mixing to join sediment entering from the sea on the landward residual flow. It then travels in the salt intrusion back to the head of the estuary. This recirculation is very effective for sorting the sediment, which is of exceedingly uniform mineralogy and settling velocity. Flocs with low settling velocities tend to be swept out into the coastal regions and onto the continental shelf. The heavier or larger flocs tend to be deposited.

The concentrations change with tidal range and during the tidal cycle, and fluid muds can occur within the area of the turbidity maximum if concentrations become sufficiently high. During the tidal cycle, as the current diminishes, individual flocs can settle and adhere to the bed, or fluid muds can form. The mud consolidates slightly during the slack water period, and as the current increases at the next stage of the tide, erosion may not be intense enough to remove all of the material deposited. A similar cycle of deposition and erosion occurs during the spring-neap tidal cycle. Generally, there is more sediment in suspension in the turbidity maximum than is required to complete a year's sedimentation on the estuary bed.

Mud flats and tidal marshes. The area of the turbidity maximum is generally well protected from waves, and there are often wide areas of mud flats and tidal marshes (**Fig. 4**). These areas also exchange considerable volumes of fine sediment with sediment in suspension in the estuary. At high water the flats are covered by shallow water, and there is often a long stand of water level which gives the sediment time to settle and reach the bottom, where it adheres or

Fig. 4. Aerial photograph of tidal flats showing the areas of pans, marshes, and vegetation between the channels, Scolt Head Island, England. (*Photograph by J. K. St. Joseph, Crown copyright reserved*)

is trapped by plants or by filter-feeding animals. The ebb flow is concentrated in the winding creeks and channels. At low water there is not enough time for the sediment to settle, and it is distributed over the tidal flats during the incoming tide. Thus, there is a progressive movement of the fine material onto the mud flats by a process that depends largely on the time delay between sediment that is beginning to settle and sediment that is actually reaching the bed. The tidal channels migrate widely, causing continual erosion. Thus, there is a constant exchange of material between one part of the marshes and another by means of the turbidity maximum. As the muds that are eroded are largely anaerobic, owing to their very low permeability, the turbidity maximum is an area with reduced amounts of dissolved oxygen in the water.

Coarse-grained material. Coarser materials such as quartz sand grains that do not flocculate travel along the bed. Those coming down the river will stop at the tip of the salt intrusion, where the oscillating tidal velocities are of equal magnitude at both flood and ebb. Ideally, coarser material entering from the sea on the landward bottom flow will also stop at the tip of the salt intrusion, which becomes an area of shoaling, with consequent decrease of grain size inland. Normally the distribution of the tidal currents is too complex for this pattern to be clear. Especially in the lower part of the estuary, lateral variations in velocity can be large. The flood and ebb currents preferentially take separate channels, forming a circulation pattern that the sediment also tends to follow. The channels shift positions in an apparently consistent way, as do the banks between them. This sorts the sediment and restricts the penetration of bed-load material into the estuary.

Salt-wedge patterns. In salt-wedge estuaries, the river discharge of sediment is much larger, though generally markedly seasonal. Both suspended and bed-load materials are important. The bed-load sediment is deposited at the tip of the salt wedge, but because the position of the salt wedge is so dependent on river discharge, the sediments are spread over a wide area. At times of flood, the whole mass of accumulated sediment can be moved outward and deposited seaward of the mouth. Because of the high sedimentation rates, the offshore slopes are very low, and the sediment has a very low bearing strength. Under normal circumstances, the suspended sediment settles through the salt wedge, and there is a zonation of decreasing grain size with distance down the salt wedge, but changes in river flow seldom allow this process to occur.

Fiord patterns. Sedimentation often occurs only at the heads of fiords, where river flow introduces coarse and badly sorted sediment. The sediment builds out into deltalike fans, and slumping on the slopes of the fan carries the sediment into deep water. Much of the rest of the fiord floor is bare rock or only thinly covered with fine sediment.

Bar-built estuary patterns. Bar-built estuaries are a very varied sedimentary environment. The high tidal currents in the inlets produce coarse lag deposits, and sandy tidal deltas are produced at either end of the inlets, where the currents rapidly diminish. In tropical areas, the muds that accumulate in the lagoons can be very rich in chemically precipitated calcium carbonate. Daniel G. MacDonald; K. R. Dyer

Bibliography. R. S. K. Barnes and J. Green (eds.), *The Estuarine Environment*, 1972; K. R. Dyer, *Coastal and Estuarine Sediment Dynamics*, 1988; K. R. Dyer (ed.), *Estuarine Hydrography and Sedimentation*, 1980; H. B. Fischer et al., *Mixing in Inland and Coastal Waters*, 1979; B. J. Neilson, J. Brubaker, and A. Kuo (eds.), *Estuarine Circulation*, 1989; B. J. Neilson and E. L. Cronin (eds.), *Estuaries and Nutrients*, 1981; C. B. Officer, *Physical Oceanography of Estuaries and Associated Coastal Waters*, 1976.

Ether

One of a class of organic compounds characterized by the structural feature of an oxygen atom linking two hydrocarbon groups, R—O—R′. Ethers are used widely as solvents, both in chemical manufacture and in the research laboratory. The most important ether is ethyl ether, $C_2H_5OC_2H_5$.

The hydrocarbon radicals R and R′ may be identical (simple ether) or different (mixed ether). They may be aromatic or aliphatic, and the names of the ethers correspond to the hydrocarbon groups present. Thus, $CH_3—O—CH_3$ is methyl ether, rarely dimethyl ether, and $C_6H_5—O—CH_3$ is phenyl methyl ether.

Properties. Ethers are less soluble in water than are the corresponding alcohols, but are miscible with most organic solvents. Low-molecular-weight ethers have a lower boiling point than the corresponding alcohols, but for those ethers containing radicals larger than butyl, the reverse is true. The boiling points approximate those of hydrocarbons of the same molecular weight and geometry, indicating that association of ether molecules in the liquid state is negligible. Inertness at moderate temperatures, an outstanding chemical characteristic of the saturated alkyl ethers, leads to their wide use as reaction media. The organic magnesium compounds known as the Grignard reagents, RMgX, perhaps the most used reagents in organic synthesis, are almost always prepared in ether solutions, and suspensions of alkali metals in ethers are often employed. At higher temperatures, however, ethers are split by the alkali metals and by the halide salts of metalloids.

Ethers may also be split by hydrogen halides. Hydrogen iodide, HI, for example, often reacts at room temperature to form an alcohol and an alkyl iodide. Ethers react with chlorine and bromine considerably more readily than do the corresponding hydrocarbons. The initial reaction involves the formation of hydrogen halide and the substitution of a halogen atom for one of the hydrogens on a carbon adjacent (alpha) to oxygen, as shown by reaction (1).

$$CH_3CH_2—O—CH_2CH_3 + Cl_2 \rightarrow CH_3CHCl—O—CH_2CH_3 + HCl \quad (1)$$

Because such α-halogens are reactive, the halogenated ethers are convenient intermediates for synthesis. Halogenated ethers are known in which the halogen is on a carbon other than that adjacent to the oxygen, but they are relatively inert.

On standing, ethers react with the oxygen of the air to form peroxides. Before distillation, it is essential that any considerable accumulation of peroxides be destroyed, by alkaline hydrolysis or by treatment with a reducing agent, such as ferrous hydroxide. On concentration and heating, ether peroxides detonate with dangerous violence. Some ethers form saltlike addition compounds with Lewis acids, the halogens, or picric acid. These addition compounds are theoretically related to, but usually much less stable than, the corresponding derivatives of amines. This property permits the separation of ethers from inert hydrocarbons by extraction with concentrated sulfuric acid.

Identification of ethers is difficult. Often the more reactive components of a mixture are removed by chemical reagents, and the residual ethers are identified by a combination of their failure to react and their specific physical properties. The inertness of ethers is utilized in the syntheses of complicated organic molecules, an objectionably reactive alcohol group being protected by converting it to an unreactive ether. Hydrogen iodide may be used to regenerate the alcohol from the ether when the need for protection has passed.

Unsaturated ethers undergo the reactions usually associated with the double bond. Vinyl ethers, in which the double bond is adjacent to the oxygen, are readily polymerized or copolymerized with such monomers as vinyl acetate to yield useful polymers. Vinyl ethers also react in the presence of acid catalysts with compounds that possess active hydrogens. Thus, with alcohols, they form acetals.

Manufacture and preparation. Simple ethers may be considered to be the anhydrides of alcohols and are manufactured from alcohols by catalytic dehydration, as in reaction (2), or from olefins by controlled catalytic hydration, as in reaction (3). Mixed

$$2ROH \rightarrow ROR + H_2O \qquad (2)$$

$$2CH_3CH{=}CH_2 + HOH \rightarrow (CH_3)_2CH{-}O{-}CH(CH_3)_2 \qquad (3)$$

ethers of definite structure may be prepared by the Williamson synthesis, shown by reaction (4). This

$$C_3H_7ONa + C_2H_5Br \rightarrow C_3H_7OC_2H_5 + NaBr \qquad (4)$$

synthesis was of considerable significance historically because a knowledge of the structure of ethers was important in developing the radical theory, a stepping-stone to the present extensive knowledge concerning the arrangements of atoms in the molecules of organic compounds. A closely related reaction is that which takes place between cellulose, alkali, and ethyl chloride to yield an important plastic, the polyethyl ether of cellulose known as ethyl cellulose.

Ethyl ether. The best hnown of the ethers is ethyl ether, sometimes called diethyl ether, $CH_3CH_2OCH_2CH_3$. It is used in industry as a solvent and in medicine as an anesthetic.

The older process of manufacture involved heating ethyl (grain) alcohol to moderate temperatures with catalytic quantities of sulfuric acid, H_2SO_4. Both ethyl ether and ethyl alcohol are manufactured by the controlled catalytic hydration of gasoline. Solid acid catalysts and a flow process are generally used instead of the sulfuric acid methods, and the proportion of alcohol to ether is controlled by variation of temperature and reactant concentrations.

The anesthetic properties were first noticed by Paracelsus (1490–1541) and were rediscovered by Michael Faraday in 1818, but it was not until 1846 that its potential as a surgical anesthetic was demonstrated, by W. T. G. Morton. The ethyl ether intended for anesthetic use differs from the ordinary variety in that possibly injurious impurities are removed. Peroxides are particularly harmful, and storage conditions must inhibit their formation. *See* ANESTHESIA.

When ether is used as a solvent, its high volatility can cause loss. However this volatility is advantageous in that ether can be readily removed from the concentrated or crystallized product. The toxicity to humans is low, and recovery from overexposure is rapid and complete. It readily forms explosive mixtures with air, and on standing in containers which have been opened, it forms dangerous peroxides. Its freezing point is $-117.4°C$ ($-179.3°F$); boiling point, 34.6°C (94.3°F); density, 0.7146; and refractive index, 1.35424. The solubility of ether in water is 6.18%, and of water in ether, 1.2%.

Cyclic ethers. Several cyclic ethers are of special importance and interest. The simplest of these is ethylene oxide or oxirane (**1**), made industrially by

(**1**)

the oxidation of ethylene with air over a silver catalyst. The major portion is used as an intermediate in the hydrolytic manufacture of ethylene glycol. Ethylene oxide is also used in the preparation of nonionic emulsifying agents, plastics, plasticizers, one type of synthetic rubber, and several important synthetic textiles. Another important use is as a gaseous sterilizing agent.

Dioxane or 1,4-dioxane (**2**) is prepared by the

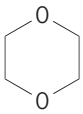

(**2**)

catalytic dimerization of ethylene oxide. It is unusual among substances of low dielectric constant (2.21) in that it is soluble in water in all proportions. Extensively used as a solvent industrially, it readily dissolves fats, waxes, natural and synthetic resins, cellulose

ethers, and lacquers, and it is employed by biologists to prepare paraffin-impregnated tissue sections.

Furan (**3**), made by the decarbonylation of furfural,

(**3**)

is the most important ether obtained from an agricultural source. Most of it is hydrogenated to form the useful solvent tetrahydrofuran (**4**).

(**4**)

Crown ethers. Certain large-ring polyethers, the crown ethers, are able to increase the solubility of alkali metal salts in nonpolar organic solvents. Specific metal complexes are formed by these crown ethers with alkali metal cations, the specificity for a given cation depending upon the hole in the middle of the crown ether structure. The 18-member cyclic polyether containing six oxygen atoms is known as 18-crown-6 (**5**). It readily forms a complex with

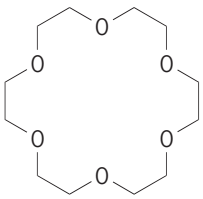

(**5**)

the potassium cation. Potassium permanganate dissolves in benzene in the presence of 18-crown-6, forming a purple solution that can oxidize alcohols, alkenes, and alkylbenzenes under neutral conditions. *See* FURAN; HETEROCYCLIC COMPOUNDS; MACROCYCLIC COMPOUND. Paul E. Fanta

Ethers for oxygenated fuel. In the 1990 the U.S. Clean Air Act was revised to promote the production and use of cleaner-burning fuels. As a result, methyl tertiary butyl ether (MTBE) became a very important petrochemical. Nine United States cities with the most severe ozone pollution have been mandated to have a minimum of 2% oxygen in a reformulated gasoline. Outside the United States, other cities and countries are requiring oxygenated gasoline. Other environmental driving forces such as the phasing out of lead compounds and the reduction of aromatics in gasoline have increased the demand for ethers.

Methyl tertiary butyl ether is the most widely produced ether for oxygenates. It is commonly produced by the dehydrogenation of isobutane and the subsequent reaction of isobutylene with methanol; it is also produced as a coproduct in the propylene oxide technology. Tertiary amyl methyl ether, also used as an oxygenate for fuel, is produced in fluid catalytic crackers and steam naphtha crackers. Because of their high octanes, these and other aliphatic ethers are good substitutes for aromatics in the gasolines. Also, their relatively low boiling temperature improves the distillation characteristics and drivability performance of the fuel. The oxygen in the ether also reduces the engine exhaust emissions. *See* GASOLINE; PETROCHEMICAL.

Tetrahydrofuran. Tetrahydrofuran is a solvent and monomer that is widely used. It has many applications in polymers, textiles, and composite materials. The most cost-effective commercial means of production is the oxidation of normal butane to maleic anhydride, which is subsequently reduced to tetrahydrofuran. Anne Gaffney

Bibliography. R. T. Morrison and R. N. Boyd, *Organic Chemistry*, 6th ed., 1992; M. B. Smith and J. March, March's *Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*, 5th ed., 2001.

Ether hypothesis

James Clerk Maxwell and his contemporaries in the nineteenth century found it inconceivable that a wave motion should propagate in empty space. They therefore postulated a medium, which they called the ether, that filled all space and transmitted electromagnetic vibrations.

During the last half of the nineteenth century, dozens of models were tried, but all broke down at some point. Direct experimental attempts to establish the existence of an absolute ether frame of reference, in which Maxwell's equations hold and light has the velocity c, have failed. The best known of these is the Michelson-Morley experiment, in which an attempt was made to measure the velocity of the Earth relative to the ether. *See* LIGHT.

Every hypothesis (ether drag, Lorentz contraction, and so on) invented to reconcile some experiment with the ether concept has been disproved by some other experiment. At present, there is no evidence whatever that the ether exists. *See* MAXWELL'S EQUATIONS. William R. Smythe

Bibliography. J. D. Jackson, *Classical Electrodynamics*, 3d ed., 1998; R. Resnick and D. Halliday, *Basic Concepts in Relativity and Early Quantum Theory*, 2d rev. ed., 1995; L. S. Swenson, Jr., *The Ethereal Aether: History of the Michaelson-Morley-Miller Aether-Drift Experiments, 1880–1930*, 1972; E. T. Whittaker, *A History of the Theories of Aether and Electricity*, 2 vols., 1954, reprint 1987.

Ethology

The study of animal behavior. Modern ethology includes many different approaches, but the original emphasis, as expounded by Konrad Lorenz and Niko Tinbergen, was placed on the natural behavior of animals. This contrasted with the focus of comparative psychologists on behavior in artificial laboratory situations such as mazes and puzzle boxes.

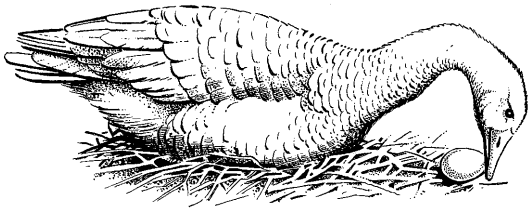

Fig. 1. Egg-rolling response of a greylag goose studied by Lorenz and Tinbergen is evoked by a great variety of shapes and objects near the nest.

Ethologists view the naturalistic approach as crucial because it reveals the environmental and social circumstances in which the behavior originally evolved, and prepares the way for more realistically designed laboratory experiments. The approach goes back to the stress that Charles Darwin placed on hereditary contributions to behavior in all species, including humans. Viewing behavior as a product of evolutionary history has helped to elucidate many otherwise puzzling aspects of its biology and has paved the way for the new science of neuroethology, concerned with how the structure and functioning of the brain controls behavior and makes learning possible.

Innate release mechanisms. A central concept in classical ethology still valuable today is that of the innate release mechanism. This has changed considerably since Lorenz's first characterization of the "innate schemata" of animals, but the major emphasis is similar. If a species has had a long history of experience with certain stimuli, especially those involving survival and reproduction, then to the extent that genes affect the ability to attend closely to such stimuli, natural selection leads to adaptations enhancing responsiveness to them. In the study of innate schemata a common first step was investigation of the development of responsiveness to such stimuli in infancy, focusing on situations that the ethologist as a naturalist knew to be especially relevant to survival. Study of infants, whether animal or human, with innate proclivities that have yet to be submerged by the accumulation of new experiences in growing to adulthood, is a rich source of insights into the biological foundations of behavioral development. Early investigations of the begging of nestling birds, serving to elicit feeding by their parents, indicated the inappropriateness of the term innate schema, which seems to imply that the bird inherits complete mental images. The evidence showed instead that birds are attuned to very specific, isolated "sign stimuli" or "releasers." Only after enrichment and transformation by experience can one confidently begin to speak in terms of mental imagery, which develops out of experiences garnered while responding to sign stimuli.

The later term "innate releasing mechanism," set forth by Tinbergen and Lorenz in an analysis of egg-rolling behavior in the greylag goose (**Fig. 1**), eschews notions of innate mental imagery. The concept has proved fertile in understanding how genes influence behavioral development, and in focusing attention of neuroethologists on inborn physiological mechanisms that permit learning while encouraging the infant to attend closely to very specific stimuli, the nature of which varies from species to species according to differences in ecology and social organization.

Such specific responsiveness can be imposed at many stages in the transformation of a sensation into a percept—sometimes centrally in the brain and sometimes in the very process of sensing a stimulus, especially in simple organisms. For example, the hearing organ on the antenna of a male mosquito is mechanically tuned to the wing tone of females of the species (**Fig. 2**). In a sense, she is the only thing he hears clearly. Similarly, the visual experience of a toad, although potentially rich and varied, with some capacity for change in the face of individual experiences, appears to be dominated by the "worm-likeness" or "snake-likeness" of shapes and images, the former associated with feeding and the latter with avoidance behavior (**Figs. 3** and **4**). When a human baby responds innately to simple components of a face, its brain is involved as well as its sense organs. It responds to faces as especially captivating and engrossing even to its untutored eyes, but unlike the toad, which learns only with deliberation, the baby quickly proceeds to learn more about detailed configurations of the face, especially its mother's. Soon it is able to distinguish men from women and eventually one stranger from another, displaying skills that can be traced back to its first innate responses to the simple stimuli conveyed even by a crude caricature of the human face (**Fig. 5**).

In birds, such as the herring gull, innate release mechanisms are also better thought of as having evolved to guide processes of perceptual learning, rather than to design animals as though they were automata (**Fig. 6**). Learning is as important in the development of the behavior of many animals as in the human species. Yet as the young organism interacts with social and physical environments with which it has evolved adaptive relationships, the

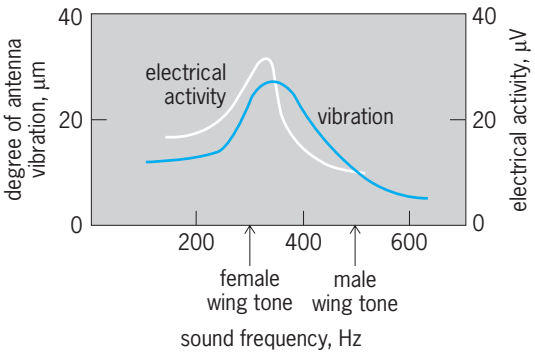

Fig. 2. Electrical activity and vibration of the male mosquito antenna. The hearing organs of the male mosquito, on the antennae, are tuned to resonate to the wing tone of females of the species, at about 300 Hz. This avoids jamming by the male's own wing tone, at about 500 Hz. (*After P. Marler and W. H. Hamilton, Mechanisms of Animal Behavior, John Wiley and Sons, 1966*)

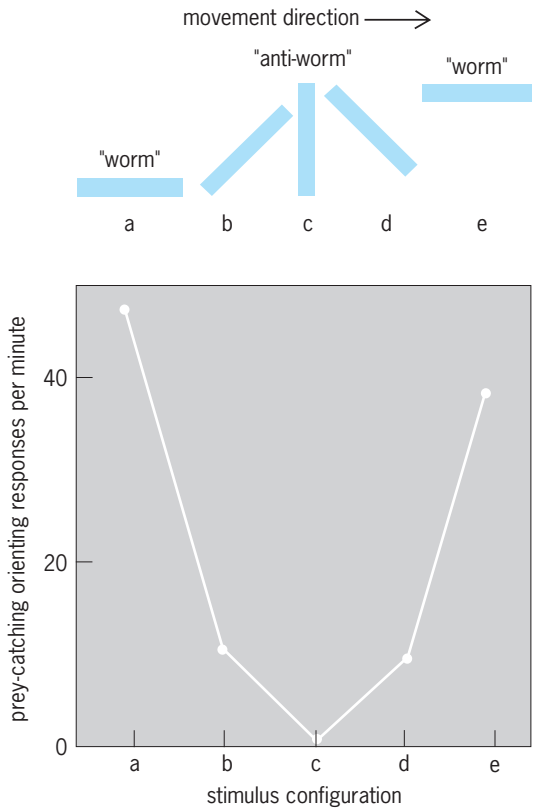

Fig. 3. A toad gives a feeding response to a "worm" stimulus but not to an "antiworm."

course that learning takes in nature is guided and profoundly influenced by the innate predispositions that the organism brings to bear on dealing with the situation.

It is important, however, to examine the innateness question closely, and to appreciate what the ethologist is *not* claiming by invoking innate responsiveness. Babies are clearly not innately equipped to respond to all characteristics of the human face. The capacity to modify and supplement innate release mechanisms accommodates for individual differences in experience. No two mothers are completely alike in voice, appearance, or in the ways in which they interact with their babies. Similarly in animals innate responsiveness focuses on a few key properties of a situation, with more complete and complex responsiveness learned by experience. Only in the earliest stages of avian and mammalian development is the world of the infant organism completely dominated by simple releasers. In adulthood, recognition may still depend heavily on sign stimuli, but never on them alone. They serve as introductions to the much more complex configurations of sights, sounds, smells, tastes, and textures that come to dominate behavior later in life. Yet the introductory role of innate release mechanisms is crucial in ensuring that some things are attended to closely and others are treated more casually, and in imposing some order on the process of building up impressions, progressing from the simple to the complex.

Fixed-action patterns. Perceptions of the external world provide a basis for both thought and action. It is a fundamental axiom of ethology that each organism's brain is armed with genetically determined programs of action which, in their own way, are as predictable and controlled as the genetic programs for developing anatomical structures such as a brain or a face. Ethologists have shown that it is possible to reconcile the need to modify patterns of action on the basis of experience with the possession of basic patterns of action that are coordinated by the brain, innately controlled, and often distinct from species to species.

Ethological pioneers such as Oskar Heinroth and Lorenz have provided the insight that what appears to the uninitiated to be a continuously varying series of unrepeatable actions typically proves to have at its core relatively stereotyped "fixed action patterns." In extreme cases, their structure can be used as a basis for historical reconstructions in much the same way that a paleontologist draws evolutionary inferences from variations in the structure of jaws, crania, and pelvic bones. Analyses of fixed action patterns have

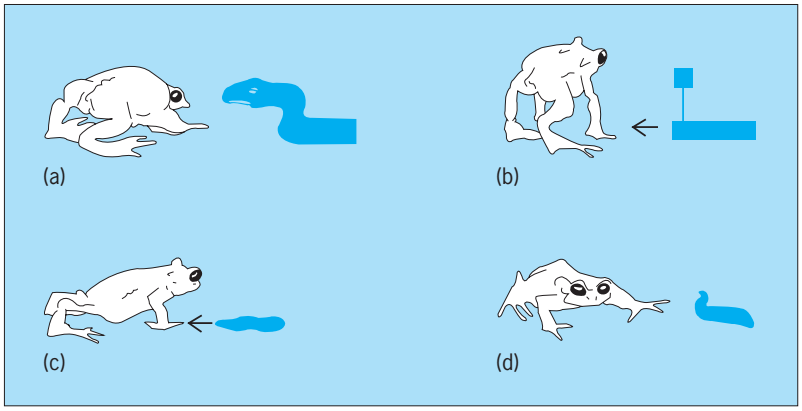

Fig. 4. Enemy images of the common toad. (*a*) A snake evokes avoidance. (*b*) A "head-rump" image also evokes avoidance. (*c*, *d*) A leech changes from a "worm" to an "enemy" when it rears its head. (*After J. P. Ewert, Neuroethology, Springer-Verlag, 1980*)

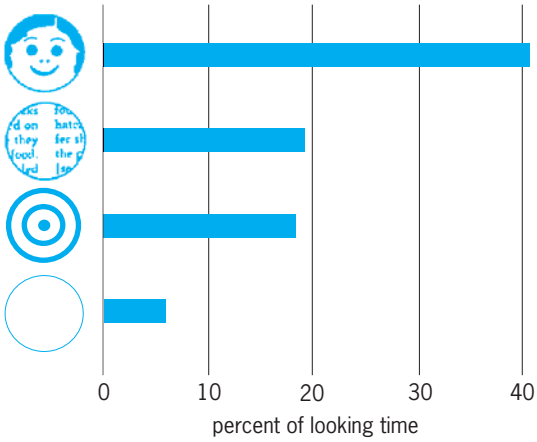

Fig. 5. Babies 2 to 3 months old spend more time looking at a facelike drawing than at newsprint or a bull's-eye. (*After R. L. Fantz, The origins of form perception, Sci. Amer., 204:66–72, 1961*)

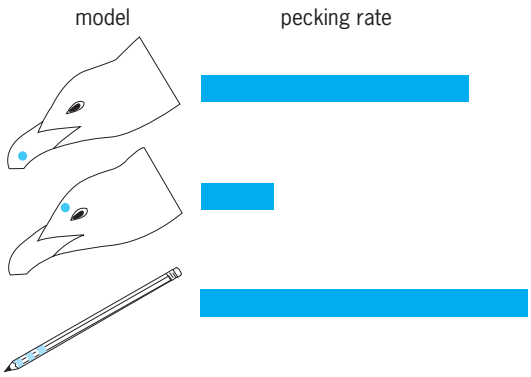

Fig. 6. Herring gull chick pecks at the spot, red in true life, on its mother's bill to get the food that is presented. The spot is critical for getting the newly hatched chick to peck. At this early age it pecks as much or more at a red-banded stick.

been used for reconstructing phylogenetic histories of organisms as diverse as fruit flies, fiddler crabs, fishes, ducks, gulls, and monkeys.

The concept of the fixed action pattern, like that of the innate release mechanism, has gone through many refinements and metamorphoses, but the underlying concept remains fundamental to understanding the development of the ability to act. Innate "motor programs," generated by the brain, are the natural units out of which behavior emerges during development. Each of these programs designates not a single completely stereotyped action, but a range of options which are limited but sufficient that selection among them allows for adjustments through experience. Close study reveals that sometimes the modifiability of actions lies in the potential flexibility of orientation, timing, and sequencing of actions rather than in the basic patterns or coordinations from which complex actions are built up. Thus nervous systems make some behavioral adjustments promptly and easily and others only with much greater difficulty, in harmony with species differences in the requirements for patterns of action as dictated by the species' structure and mode of life. Ethologists have found repeatedly that while social experience is vital in many animals for normal development of actions and responses, animals reared in restricted environments may still develop many units of action that are normal. The animal has to learn, however, how to put them together in an adaptive sequence.

Even when behavior improves as young animals mature, great care must be taken not to confuse effects of experience with consequences of maturation of the nervous system. A classic experiment demonstrated the emergence of normal movement in salamander embryos that were anesthetized through entire phases of development; thus the behavior is not due to practice or learning. In many cases, however, feedback from the performance of actions may be critically important for normal development, as when birds learn to sing.

Imprinting. Modern research on the ethology of learning began when Lorenz discovered imprinting in geese. He found that if he led a flock of newly hatched goslings himself they became imprinted on him (**Fig. 7**). When mature, they would court people as though confused about their own species identity. Learning occurred very rapidly and tended to be restricted to a short sensitive phase early in life. A newly hatched gosling would normally follow its parents as they led the young away from the nest, and it was only by interposing himself as a parental surrogate that Lorenz discovered that a learning process is involved. The learning is highly focused by genetically determined preferences both to follow a parent-object with particular appearance and emitting species-specific calls, and also to learn most quickly and accurately at a particular stage of development. The interplay between nature (genetic predisposition) and nurture (environmental influence) in learning is displayed especially clearly in imprinting, hence its special interest to biologists and psychiatrists. Indications are that it is not concerned so much with learning about species as with learning to recognize individual parents and kin, both to ensure mating with one's own kind and to avoid incestuous inbreeding.

There are many forms of imprinting. So-called filial imprinting, ensuring that ducklings and goslings follow only their parent, is distinct from sexual imprinting, affecting mate choice in adulthood; the sensitive phases for learning are different in each case. Imprinting-like processes also shape the development of food preferences and abilities to use the sun and stars in navigation.

Song learning has many of the same features as imprinting. There are local dialects in many birdsongs so that, for example, in the San Francisco Bay area one could get a good idea of location just by listening

Fig. 7. Goslings following Lorenz, on whom they are imprinted, adopting him as their mother.

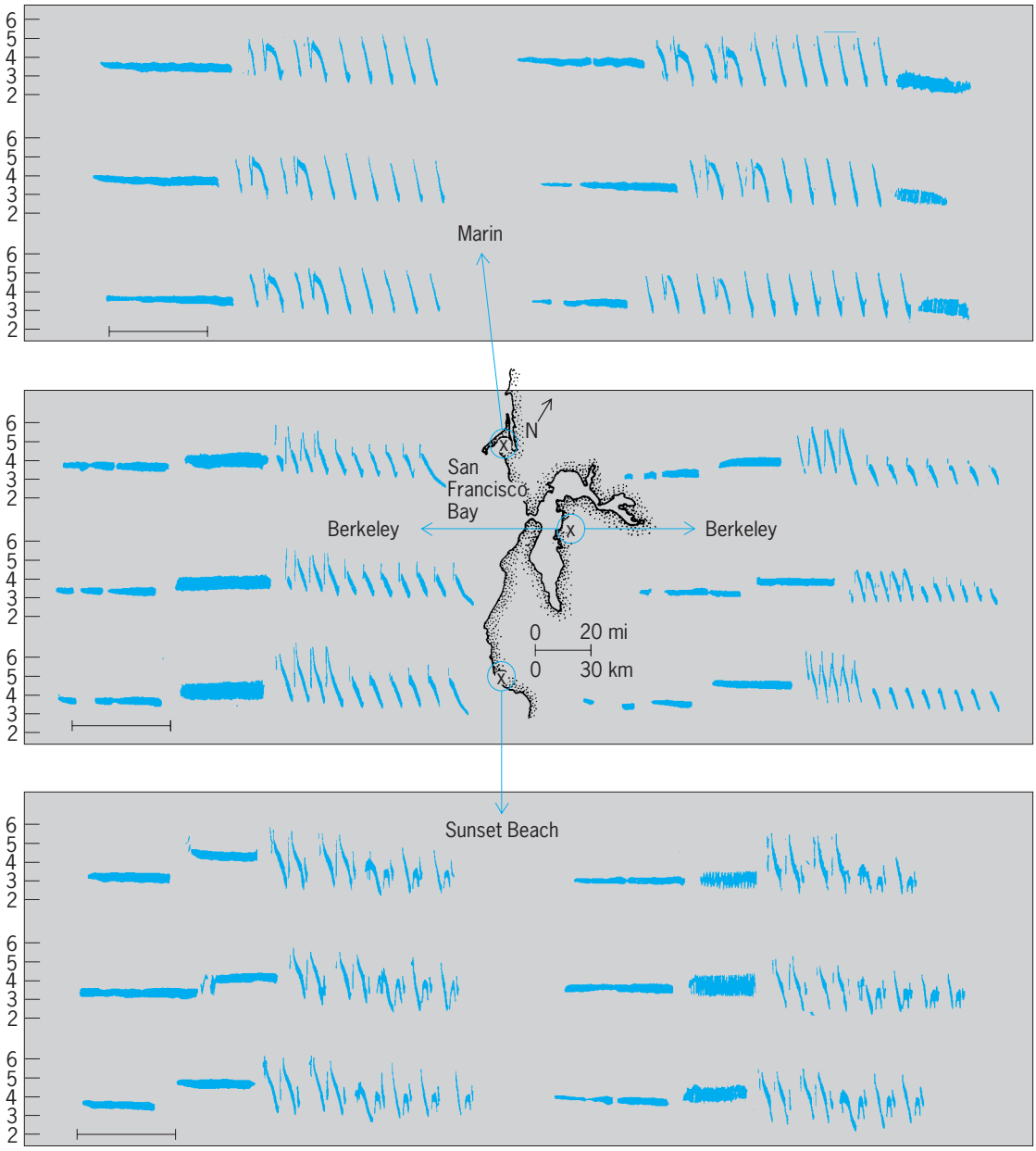

Fig. 8. Sound spectrograms of songs of 18 white-crowned sparrows in the San Francisco Bay area. Marin, Berkeley, and Sunset Beach sparrows each have a different dialect. Sound spectrograms can be read from left to right with sound frequency (in hertz) as a measure of pitch on the vertical axis. (*After P. Marler, A comparative approach to vocal learning: Song development in white-crowned sparrows, J. Comp. Physiol. Psychol., 71(no. 2, pt. 2):1–25, 1970*)

to the white-crowned sparrow (**Fig. 8**). A sparrow reared away from members of its own kind will sing an abnormal song, lacking any trace of the local dialect. If taped songs are played to a young sparrow, roughly between 2 and 7 weeks of age, it will then memorize the song patterns (**Fig. 9**). Later it will not only sing normally, but will reproduce that particular dialect, even if the dialect is from a different area than the sparrow's birthplace.

This experiment, which has been conducted with many songbirds, shows clearly that songs are nurtured by experience. But nature also intrudes at many points in the learning process. If two species are reared in isolation, both will sing abnormally (**Fig. 10**), but their songs are quite distinct from one another, preserving enough normal features so that their species identity can easily be distinguished. The deficiencies are clear if the songs are played over a loudspeaker on the territories of wild males. Normal songs are attacked more than those of males reared in isolation, but the latter still get more responses than the songs of another species. The size of individual male song repertoires varies from species to species, ranging from one into the hundreds. Although innate repertoires are always smaller than learned ones, species differences nevertheless persist.

Another way to demonstrate the interaction of nature and nurture is to give the young sparrow a choice of songs of more than one species. Some

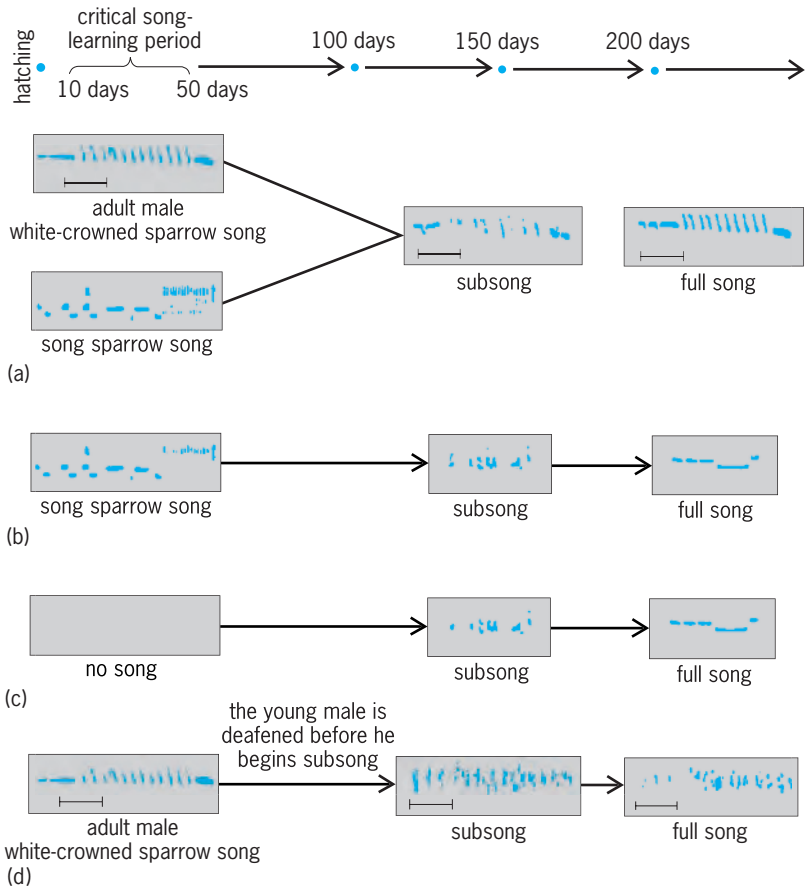

Fig. 9. Diagrams of song development in white-crowned sparrows: (*a*) when played recordings of normal song; (*b*) when given songs of a different species, the song sparrow; and (*c*) when reared in isolation. (*d*) Song development in a deaf male. Songs of other species may be rejected for learning, as shown with song sparrow vocalization in *a* and *b*.

birds, such as starlings and mockingbirds, learn the songs of other species. African wydah birds lay their eggs in other birds' nests, and learn songs from their hosts. But these birds are exceptions. If choices are offered, birds are likely to favor songs of their own kind. Ethologists theorize that songbirds inherit an auditory template in the brain that both helps them in selecting songs of their own kind for learning, and also helps the young male to develop his own song at a later stage. However, he must be able to hear his own voice to achieve this. This is evident from the buzzy, tuneless song of a bird that is deaf, a song so anonymous that it lacks even those characteristics that the bird develops normally in isolation from its own kind. Thus innate auditory templates are responsible for some of the normal features of a birdsong, exerting their influence by way of auditory feedback.

Research on song learning is one of the clearest cases in which neuroethologists have been led to fundamentally new discoveries about how brains support learning. The new insights stem directly from research in the ethological tradition, demonstrating genetic contributions to the song-learning process. Thus unlike psychological studies of animal learning in the laboratory, which have tended to favor the "blank-slate" view of the brain's contribution to learning, ethology emphasizes the need to understand all aspects of the biology of a species under study before one can hope to understand how the animal learns to cope with the many complexities of individual existence and social living. Thus ethology may lead not only to an understanding of how natural behavior evolves, but also to new insights into how

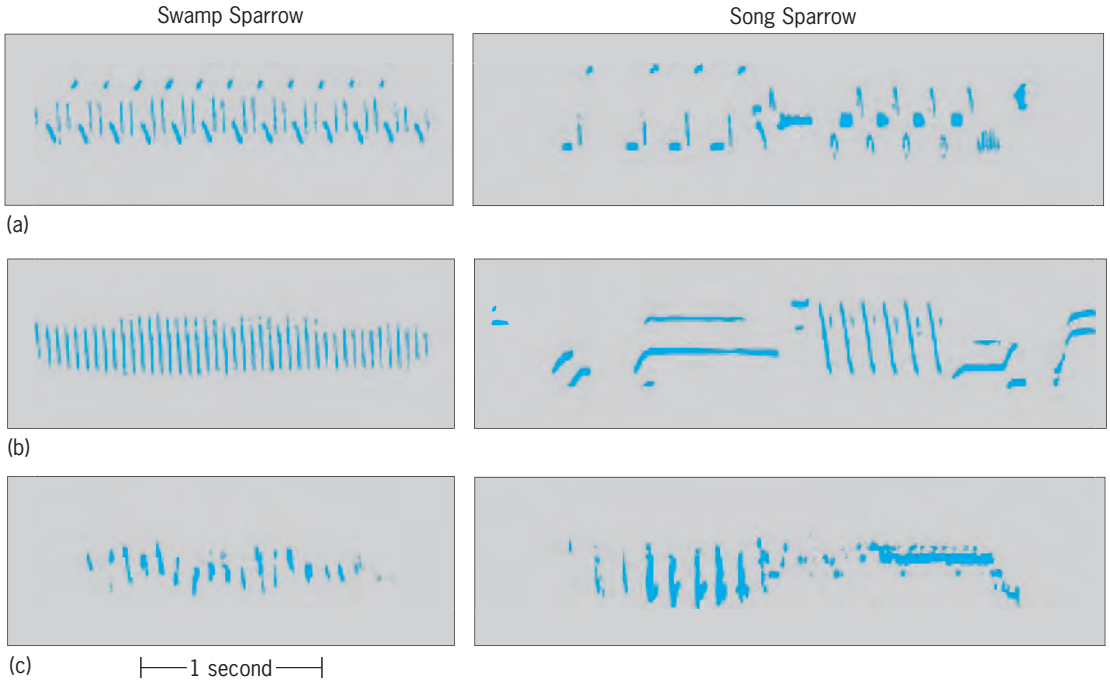

Fig. 10. Comparison of songs of the swamp sparrow and the song sparrow: (*a*) as heard in nature (normal); (*b*) as heard from males reared from the egg without any opportunity to hear the normal song of their species (isolated); and (*c*) as produced by deaf birds. (*After P. Marler, Song learning: innate species differences in the learning process, in P. Marler and H. Terrace, eds., The Biology of Learning, Dahlem Conference, Springer-Verlag, 1984*)

brains help organisms learn to cope with social and environmental problems confronting them as individuals. *See* ANIMAL COMMUNICATION; BEHAVIOR GENETICS; INSTINCTIVE BEHAVIOR. Peter R. Marler

Bibliography. J. L. Gould, *Ethology: The Mechanisms and Evolution of Behavior*, 1982; R. A. Hinde, *Ethology: Its Nature and Relations with Other Sciences*, 1982; J. R. Krebs and N. B. Davies, *An Introduction to Behavioural Ecology*, 3d ed., 1993.

Ethyl alcohol

Probably the best known of the alcohols, ethyl alcohol, formula CH_3CH_2OH, is also called alcohol, ethanol, grain alcohol, industrial alcohol, fermentation alcohol, cologne spirits, ethyl hydroxide, and methylcarbinol. Pure ethyl alcohol is a colorless, limpid, volatile liquid which is flammable and toxic and has a pungent taste. It boils at 78.4°C (173°F) and melts at −112.3°C (−170.1°F), has a specific gravity of 0.7851 at 20°C (68°F), and is soluble in water and most organic liquids. It is one of the most important industrial organic chemicals. Ethyl alcohol is produced by chemical synthesis and by fermentation or biosynthetic processes. *See* BIOSYNTHESIS; ORGANIC SYNTHESIS.

Ethyl alcohol is used as a solvent, extractant, antifreeze, fuel and fuel additive, and intermediate in the synthesis of innumerable organic chemicals. It is also an essential ingredient of alcoholic beverages.

Various grades of ethyl alcohol are produced, depending on their intended use. U.S. Pharmaceutical (USP XV) grade is the water azeotrope of ethyl alcohol and is 95% ethyl alcohol by volume. National Formulary (NFX) grade is 99+% ethyl alcohol by weight; it is also called absolute, or anhydrous, alcohol. This grade is generally prepared by azeotropic dehydration with benzene and therefore usually contains about 0.5% benzene. Denatured alcohol contains a small amount of a malodorous or obnoxious material to prevent the use of this grade of ethyl alcohol for beverage purposes. *See* AZEOTROPIC MIXTURE.

The major use of ethyl alcohol is as a starting material for various organic syntheses. Bimolecular dehydration of ethyl alcohol gives diethyl ether, which is employed as a solvent, extractant, and anesthetic. Dehydrogenation of ethyl alcohol yields acetaldehyde, which is the precursor of a vast number of organic chemicals, such as acetic acid, acetic anhydride, chloral, butanol, crotonaldehyde, and ethylhexanol. Reaction with carboxylic acids or anhydrides yields esters which are useful in many applications. The hydroxyl group of ethyl alcohol may be replaced by halogen to give the ethyl halides. Treatment with sulfuric acid gives ethyl hydrogen sulfate and diethyl sulfate, a useful ethylating agent. Reaction of ethyl alcohol with aldehydes gives the respective diethyl acetals, and reaction with acetylene produces the acetals, as well as ethyl vinyl ether. Treatment of ethyl alcohol with ammonia produces acetonitrile, which may be reduced to ethylamine. These and other ethyl alcohol-derived chemicals are used in dyes, drugs, synthetic rubber, solvents, extractants, detergents, plasticizers, lubricants, surface coatings, adhesives, moldings, cosmetics, explosives, pesticides, and synthetic fiber resins. Herman J. Phaff; Emil M. Mrak

Bibliography. J. R. Critser, Jr., *Biotechnical Engineering: Equipment and Processes*, 1994; M. Roehr (ed.), *The Biotechnology of Ethanol: Classical and Future Applications*, 2001; H. C. Vogel and C. C. Todaro (eds.), *Fermentation and Biotechnical Engineering Handbook: Principles, Process Design, and Equipment*, 2d ed., 1996.

Ethylene

A colorless gas, formula $CH_2{=}CH_2$, with a boiling point of −103.8°C (−155°F) and a melting point of −169.4°C (−273°F). Ethylene is the most important synthetic organic chemical in terms of volume, sales value, and number of derivatives. About half of the ethylene produced is used in the manufacture of polyethylene; ethylene dichloride and vinyl chloride production uses about 20%, synthesis of ethylene oxide and derivatives account for about 12%, and styrene production consumes about 8% of the ethylene. Other important derivatives are ethanol, vinyl acetate, and acetaldehyde.

Production. Thermal cracking of hydrocarbons in the presence of steam is the most widely used process for producing ethylene. Cracking is done at about 1600°C (2912°F) and 30 lb/in.^{2} absolute (21 kilopascals absolute) pressure, followed by rapid quenching to below 1000°C (1760°F). Ethylene is recovered by low-temperature fractionation at 500–550 lb/in.^{2} absolute (340–380 kPa absolute) and purified by very low-temperature (approximately −65°C or −85°F) gas-separation procedures to remove hydrogen, methane, and ethane.

In the United States, hydrocarbon gases—ethane through the butanes—account for about 65% of the feedstocks used to produce ethylene. With these clean raw materials, yields are good (about 50%), and efficiencies and conversion are high (about 80%). Liquid refinery products such as naphthas, kerosines, and gas oils constitute about 25% of process raw materials in the United States. While with these more complex feedstocks the yields and efficiencies are about 30%, the trend has been toward the use of the more available, heavier feedstocks. In some countries where petroleum raw materials are expensive, ethylene is produced by the dehydration of fermentation ethanol.

Derivatives. Polyethylene, the most important derivative of ethylene, is produced by both high- and low-pressure processes to make high- and low-density, high-molecular-weight thermoplastic polymers. Aluminum alkyl catalysts (Ziegler polymerization) are used to polymerize ethylene to relatively low-molecular-weight, straight-chain hydrocarbon

derivatives which are convertible to even-numbered-carbon linear olefins, alcohols, and acids. Commercial processes use palladium-catalyzed oxidation of ethylene to produce acetaldehyde, or if acetic acid is used as the solvent, vinyl acetate. Chlorination and oxychlorination processes are used to make vinyl chloride. Ethylene oxide is produced by silver-catalyzed oxidation of ethylene. Acid-catalyzed hydration of ethylene produces ethanol competitively with fermentation processes. *See* ACETYLENE; ALKENE; ETHYL ALCOHOL; ETHYLENE OXIDE; POLYMER.

Robert K. Barnes

Function as plant hormone. Ethylene exerts a major influence on many aspects of plant growth, development, and senescence. It is produced by plants and acts in trace amounts; the effects on plants are spectacular and commercially important.

The ability of certain gaseous agents existing in smokes, fumes, or illuminating gas to stimulate fruit ripening, to induce pineapple flowering, to hasten the coloring ("degreening") of harvested lemons, or to shed tree leaves prematurely, was recognized long before ethylene was identified as a plant growth regulator. In 1901, D. Neljubow identified ethylene as the agent in illuminating gas which most effectively caused the abnormal growth of dark-grown pea seedlings. Since then, those growth-regulatory effects of illuminating gas and various fumes and smokes have been attributed to the presence of ethylene in the air.

Although fruit shippers were aware that ripe or rotting fruits could hasten the ripening of other fruits stored with them, it was not until 1934 that chemical proof of production of ethylene by ripe apples was provided. Fruit physiologists have since established that ethylene functions as a ripening hormone. Subsequently, ethylene has been identified as a natural product not only of fruits but of all plant parts, including leaves, stems, flowers, and roots. The rate of ethylene production varies with both the type of tissue and its stage of development. In germinating seeds, high rates of ethylene production occur when the radicle (primary root of the seedling) starts to penetrate the seed coat and during the period when the seedling forces its way through soil. In fruits and flowers, ethylene production is induced at the initial stage of fruit ripening and flower fading. The pathway of ethylene biosynthesis in plants has been recently elucidated: ethylene is derived from C-3,4 of methionine, an amino acid, via *S*-adenosylmethionine and 1-aminocyclopropane-1-carboxylic acid as intermediates.

Demonstration of the natural function of ethylene in fruit ripening has stimulated the search for other functions. Among the most diverse physiological effects which have been recognized are: breaking of seed and bud dormancy, stimulation of cortical air space (aerenchyma) development and adventitious root formation, inhibition of root growth, stem strengthening and internode shortening, stimulation of flowering in pineapple, induction of female flower in cucumber, promotion of fruit and leaf degreening, stimulation of fruit ripening, promotion of fruit and leaf abscission, promotion of fading in various types of flowers, promotion of stem growth in some aquatic and semiaquatic plants, modification of gravitropism and stimulation of latex flow in rubber trees. *See* ABSCISSION; DORMANCY.

The use of ethylene gas to regulate plant responses of commercial value is restricted to harvested crops (such as green banana, tomato, or citrus fruits) in special closed chambers. This problem was largely overcome with the introduction and commercialization of ethylene-releasing compounds, such as 2-chloroethylphosphonic acid, which is spontaneously degraded into chloride, ethylene, and orthophosphate upon dilution with water. This "liquid ethylene" can be applied to preharvest or postharvest crops, and is now registered for more than 20 crops. It is used for promoting ripening of fruits such as tomatoes, apples, coffee berries, and grapes; facilitating harvesting of cherries, walnuts, and cotton by accelerating abscission or fruit dehiscence; increasing rubber production by prolonging latex flow in rubber trees; increasing sugar production in sugarcane; synchronizing flowering in pineapple; and accelerating senescence of tobacco leaves. *See* PLANT GROWTH; PLANT HORMONES.

S. F. Yang

Bibliography. F. B. Abeles, *Ethylene in Plant Biology*, 2d ed., 1997; L. Albright, *Novel Production Methods for Ethylene, Light Hydrocarbons, and Aromatics*, 1991; M. Arshad and W. T. Frankenberger, Jr., *Ethylene: Agricultural Sources and Applications*, 2001; J. E. Brady and F. Senese, *Chemistry: Matter and Its Changes*, 4th ed., 2004.

Ethylene glycol

A colorless, nearly odorless, sweet-tasting, hygroscopic liquid, formula $HOCH_2CH_2OH$. It is relatively nonvolatile and viscous and is the simplest member of the glycol family. Ethylene glycol freezes at −3°C (8.6°F), boils at 197.6°C (387.7°F), and is completely soluble in water, common alcohols, and phenol. Low molecular weight, low volatility, water solubility, and low solvent action on automobile finishes make ethylene glycol ideal as a radiator antifreeze and coolant; water mixtures (58–80% glycol) freeze below −46°C (−50°F). *See* ANTIFREEZE MIXTURE; GLYCOL.

Worldwide, about half of all ethylene glycol is used in polyester resins, and about a third goes into antifreeze. Other uses for this commodity chemical are in explosives, brake and shock-absorber fluids, and alkyl-type resins. *See* POLYESTER RESINS.

The original commercial process for ethylene glycol involved hydrolysis of ethylene chlorohydrin derived from chlorine and ethylene. All commercial ethylene glycol is now produced by the vapor-phase oxidation of ethylene. A commerical process was developed to produce ethylene glycol directly from ethylene. This process uses a tellurium oxide/bromide ion catalyst to oxidize ethylene in acetic acid solution. Another patented process describes

production of ethylene glycol from synthesis gas (carbon monoxide and hydrogen).

Ethylene glycol undergoes the simple reactions of alcohols such as etherification, condensation, oxidation, and esterification. Mono- and dialkyl ethers of ethylene glycol are formed by its reaction with sodium hydroxide and dialkyl sulfates or alkyl halides. Dioxane, a cyclic diether, as well as diethylene glycol, can be prepared by acid-catalyzed dehydration of ethylene glycol. Di-, tri-, and polyethylene glycols are best produced by the acid- or base-catalyzed reaction of ethylene oxide with ethylene glycol. Commercially di-, tri-, and some tetraethylene glycols are produced as by-products of industrial ethylene glycol processes. Acid-catalyzed reactions of aldehydes and ketones with ethylene glycol produce five-membered cyclic acetals and ketals called 1,3-dioxolanes. Vapor-phase catalytic dehydrogenation (oxidation) of ethylene glycol can produce either 2-hydroxymethyl-1,3-dioxolane or glyoxal. Liquid-phase nitric acid or air oxidation produces oxalic, glycolic, and formic acids, formaldehyde, and glycol aldehyde and glyoxal. With nitric acid under dehydrating conditions, ethylene glycol forms the dinitrate ester, which is employed in conjunction with nitroglycerin to produce explosives that have low freezing points.

Dibasic acids or anhydrides react with ethylene glycol to form polyester condensation polymers. Reaction with terephthalic acid or its esters produces polyester resins which can be spun to fibers that find wide use in clothing and general fabrics applications. The polymer also has important film applications. Condensation of glycol with unsaturated diacids, for example maleic acid, followed by free-radical cross-linking reactions with polymerizable olefins such as styrene produces another commercially important polymer class identified as unsaturated polyester resins. Saturated and unsaturated polyester resins constitute the major use of ethylene glycol in the world today. *See* ETHER; ETHYLENE OXIDE; POLYETHYLENE GLYCOL. Robert K. Barnes

Bibliography. J. E. Brady and F. Senese, *Chemistry: Matter and Its Changes*, 4th ed., 2004.

Ethylene oxide

The simplest cyclic ether or epoxide, with the formula C_2H_4O, and the structure

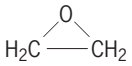

It is also called epoxyethane and oxirane. Ethylene oxide was discovered in 1859 by C. A. Wurtz. His preparation from ethylene chlorohydrin and aqueous base became the first commercial process, practiced during the early 1920s and until the mid-1940s. Thereafter the direct, silver-catalyzed vapor-phase oxidation of ethylene was the process of choice. Commercial processes use either air or oxygen to oxidize ethylene at low conversion and high selectivity to ethylene oxide. *See* ETHYLENE; HETEROCYCLIC COMPOUNDS.

Ethylene oxide is a colorless gas boiling at 10.4°C (50.7°F) and melting at −112°C (−170°F), with refractive index 1.3597 at 7°C (45°F), and density 0.8969 kg/liter at 0°C (32°F). Its vapors are flammable and explosive, and it is considered a relatively toxic liquid and gas. It is miscible in all proportions with water, alcohols, ethers, and other organic solvents. *See* ETHER.

Ethylene oxide reacts, usually by acid or base catalysis, with most active hydrogen compounds such as water, alcohols, acids, phenols, amides, hydrogen sulfide, and hydrogen cyanide; it also reacts with other organic compounds such as carbon dioxide, amines, ammonia, and Grignard reagents and with itself. Ethylene oxide's reactivity can be ascribed to the strained ring structure.

About 50% of the ethylene oxide produced is converted to ethylene glycol by reaction with water. Primary uses for ethylene glycol are in the manufacture of polyester resins and as an automobile antifreeze. About 30% of the ethylene oxide produced is used in the manufacture of nonionic surfactants, glycol ethers, and ethanolamines. Gaseous ethylene oxide in CO_2 or difluoromethane is used as a fumigant and a sterilizing agent for medical equipment. *See* ETHYLENE GLYCOL; POLYETHYLENE GLYCOL. Robert K. Barnes

Bibliography. J. E. Brady and F. Senese, *Chemistry: Matter and Its Changes*, 4th ed., 2004; T. W. G. Solomons, *Organic Chemistry*, 7th ed., 1999.

Ethylenediaminetetraacetic acid

A chelating agent for metallic ions, abbreviated EDTA. The structural formula is

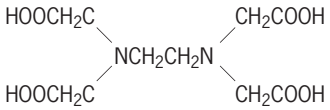

Tetrasodium EDTA is the most common form in commerce, but other metallic chelates are marketed, for example, iron, zinc, and calcium. Tetrasodium EDTA is a white solid, very soluble in water and forming a basic solution. Prepared from ethylenediamine, formaldehyde, and sodium cyanide in basic solution, or from ethylenediamine and sodium chloroacetate, EDTA is a strong complexing and chelating agent. It reacts with many metallic ions to form soluble chelates. As such, it is widely used in analysis to retain alkaline earths and heavy metals in solution. The iron chelate is useful in lawn management and gardening as a replacement for ferrous sulfate (copperas). Calcium EDTA is used to control the deterioration of natural seawater in aquariums. Calcium disodium salt of EDTA is used in pharmaceuticals to prevent calcium depletion of the body during therapy. *See* CHELATION; COORDINATION COMPLEXES. Frank Wagner

Bibliography. J. E. Brady and F. Senese, *Chemistry: Matter and Its Changes*, 4th ed., 2004.

Eucalyptus

A large and important genus of Australian forest trees; includes about 500 species in the family Myrtaceae. Only two species occur naturally outside Australia in the adjacent islands. Eucalypts occur throughout Australia except in coastal tropical and subtropical rainforests in Queensland and New South Wales and in temperate rainforests in Victoria and Tasmania. They are confined to water courses in the extensive arid zones of central and northwest Australia. Eucalypts grow from sea level to tree line (6600 ft or 2000 m).

Because of its large geographic range the genus exhibits many habits, from tall trees to multistemmed, shrubby species called mallees. The mountain ash (*Eucalyptus regnans*) of Victoria and Tasmania is the tallest hardwood in the world, reaching heights over 330 ft (100 m); only the coast redwoods of California are taller. Having epicormic buds in branches and stems, and lignotubers in the roots, many species are well adapted for surviving fire and drought. Some species have smooth bark, in which case they are referred to as gum-barked, while others have rough bark. Half-barked species have rough bark on the lower trunk and smooth bark on the upper stem and branch. Many species shed their bark in long ribbons.

Eucalyptus is an evergreen genus with four different leaf types—seedling, juvenile, intermediate, and adult—depending on plant maturity. Juvenile leaves of some species, particularly those that are silvery blue and oval, are extensively used for floral decorations. Most species have white or cream flowers; however, some species, particularly those from Western Australia, produce spectacular red (*E. ficifolia*) or yellow (*E. erythrocorys*) flowers and are planted widely as ornamentals. These ornamental species are small in habit and drought-resistant. Eucalypt seed cases are woody, and many species produce so-called gumnuts.

The major commercial species include mountain ash (*E. regnans*), alpine ash (*E. delegatensis*), messmate stringybark (*E. obliqua*), flooded gum (*E. grandis*), blackbutt (*E. pilularis*), spotted gum (*E. maculata*), jarrah (*E. marginata*), karri (*E. diversicolor*), tallowwood (*E. microcorys*), blue gum (*E. globulus*), and river red gum (*E. camaldulensis*). Heartwood is generally dense and varies in color from light brown, such as blackbutt, to dark red, such as jarrah.

Eucalyptus globulus (see **illus.**), with gray or brownish bark at the base and peeling bark above, is the common species throughout California. The red flowering eucalypt (*E. ficifolia*) is a popular ornamental in southern California. *Eucalyptus camaldulensis* has also been planted in Florida, and Hawaii has plantations of many species. Eucalypts have been planted widely for commercial use in Brazil and other South American countries, Africa, the Indian subcontinent, and the Middle East. They are used extensively for fuel and construction and are an important component of third world economies. Foliage of some species yields essential oils for medicines and perfumes. Tannins are extracted from the bark of certain species. *See* ESSENTIAL OILS; MYRTALES; TREE.

Robert L. Edmonds

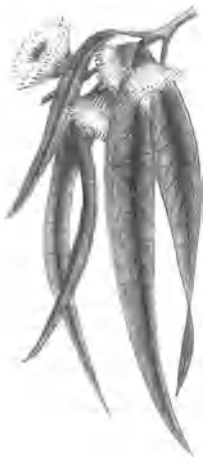

Eucalyptus globulus, **showing spray of mature foliage.**

Bibliography. D. J. Boland et al., *Forest Trees of Australia*, 1984; A. Keast (ed.), *Ecological Biogeography of Australia,* 1981; C. D. Pryor, *The Biology of Eucalypts*, 1976.

Eucarida

The largest and most highly evolved superorder of the crustacean class Malacostraca. It contains the orders Euphausiacea, Amphionidacea, and Decapoda. Despite the great morphological and ecological diversity of eucaridans, all share (1) a well-developed carapace that is fused to all the thoracic somites; (2) stalked and movable eyes, although in a few these have been secondarily reduced; (3) mandibles that lack a mobile element between the incisor and molar processes; (4) a telson without a caudal furca; (5) heart and gills that are thoracic in position; and (6) typically metamorphic larval development.

The Euphausiacea are pelagic, open-water crustaceans. The Amphionidacea, although also pelagic and broadly distributed, are represented by a single rather bizarre species. The Decapoda, including the shrimps, lobsters, crabs, and crablike crustaceans, abound from the intertidal to the abyssal seas but are found as well in terrestrial habitats and fresh-water rivers, streams, and lakes. Their morphological adaptations, although based on a common plan, are as ubiquitous as their habitats. *See* AMPHIONIDACEA; CRUSTACEA; DECAPODA (CRUSTACEA); EUPHAUSIACEA; MALACOSTRACA.

Patsy A. McLaughlin

Bibliography. R. C. Brusca and G. J. Brusca, *Invertebrates*, 2d ed., 2002; P. A. McLaughlin, *Comparative Morphology of Recent Crustacea*, 1980; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; F. R. Schram, *Crustacea*, 1986; F. R. Schram (ed.), *Crustacean Phylogeny, Crustacean Issues*, 1983.

Eucestoda

A subclass of tapeworms including most of the members of the class Cestoda, which is a taxon of the platyhelminthes. All species are endoparasites of vertebrates, living in the intestine or related ducts.

Morphology. Like other members of the class, the eucestodes have no digestive tract or mouth. Nutrition presumably occurs by absorption of food through the body surface. The body, or strobila, is usually very elongated and tapelike and frequently divided into segments, or proglottids, with replication of the hermaphroditic reproductive systems. In a few species there is duplication of both male and female organs within a single segment. The anterior end is usually modified into a holdfast organ, the scolex. Since a digestive tract is completely absent, the scolex is of solid construction, typically highly muscular with sucking depressions and hooks (**Fig. 1**). Behind the scolex there is usually a region of undifferentiated tissue, termed the neck, from which proliferation of new segments occurs. In some species a neck is not apparent, and segments begin immediately behind the scolex. The strobila, or chain of segments, typically arises by growth and differentiation, the smallest and least-differentiated segments being nearest to the neck. Young segments are always broader than long, but, as the segments grow in volume, they may become square or longer than broad. The body is clothed with what has been called a tegument, but it has been shown that this noncellular layer is relatively complex (**Fig. 2**). This might be expected since these animals actively absorb such nutrients as carbohydrates and amino acids.

Nervous system. The nervous system consists basically of a pair of longitudinal lateral nerves running through the length of the strobila. Additional nerve trunks are frequently present. The trunks are connected by a ring-shaped commissure in each segment and by thickened commissures, a "brain," in the scolex. Although the tapeworms lack special sensory structures, the body surface has many sensory nerve endings.

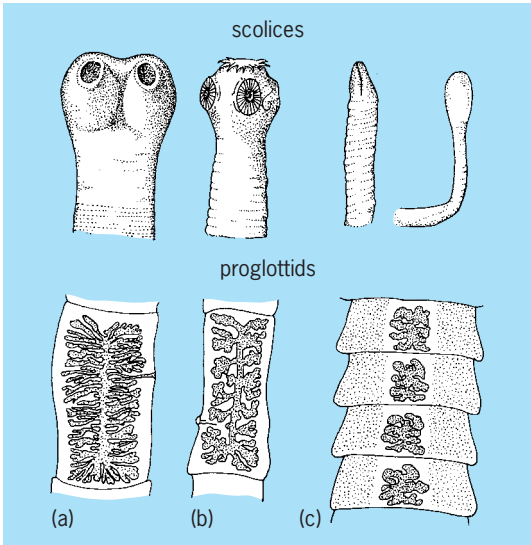

Fig. 1. Scolices and mature proglottids of (*a*) *Taenia saginata*, (*b*) *T. solium*, and (*c*) *Dibothriocephalus latus*. (*After T. I. Storer and R. L. Usinger, General Zoology, 3d ed., McGraw-Hill, 1957*)

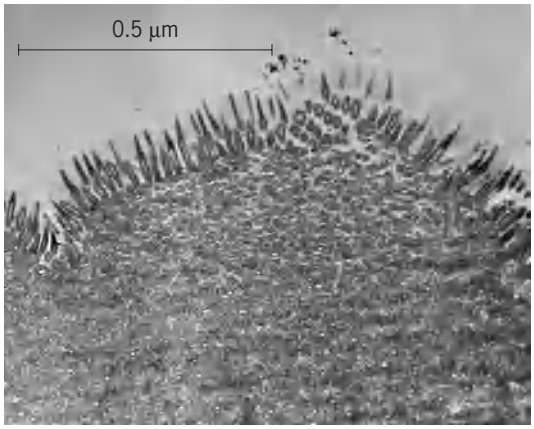

Fig. 2. The tegument of *Hymenolepis diminuta*. (*Electron photomicrograph by A. H. Rothman*)

Excretory system. The excretory or osmoregulatory system consists of protonephridia, or flame bulbs, which connect ultimately through a system of fine ducts to a pair of longitudinal lateral canals on either side. These canals traverse the length of the body.

Reproductive system. The reproductive systems develop progressively along the strobila, the male system usually becoming mature before the female. In some forms there are two complete hermaphroditic systems. The male and female genital ducts commonly open into a common antrum, usually situated on the lateral margin of the segment. The testes are usually small rounded bodies, ranging in number from 1 to 1000 but most often about 100. Sperm are carried in a system of ducts to an eversible copulatory organ, the cirrus, which is frequently armed with spines or hooks. The ovary is single, commonly having two lobes. In the orders other than Cyclophyllidea, yolk glands occur as numerous follicles. Yolk cells from these glands are gathered in a duct system which enters the oviduct from the ovary. The oviduct also receives the vagina and eventually enlarges to form the uterus. The uterus begins developing after the gonads and their associated ducts have matured and varies in shape from a simple sac to a ramified network.

With growth, the reproductive systems differentiate, and fertilization occurs in a modified region of the oviduct where the ovum is enclosed with yolk cells in a shell. These eggs develop into embryos, or oncospheres, within the uterus. Fertilization may occur within a single proglottid, between segments in the same strobila, or between proglottids of separate worms. Hypodermic impregnation occurs in a few cases. The posterior proglottids become distended with embryos. In some orders, such as Tetraphyllidae, the sexually mature proglottids are shed from the body before significant development of embryos, and the free proglottids undergo "ripening" in the host intestine. In most other orders the gravid,

embryo-filled proglottids are shed from the strobila and pass out of the host, with or without degeneration of the segment tissue. In the Pseudophyllidea, gravid proglottids are retained and the shelled embryos are shed through a uterine pore. The growth rates of eucestodes show great variation. Different species of the same genus may show marked differences in growth rate and in the time during which a high growth rate, without senility, is maintained. The life of the sexually reproducing tapeworm varies from 4 or 5 days in some pseudophyllideans to 20 or more years in some cyclophyllideans. *See* CESTODA; TETRAPHYLLIDEA.

Larval development. The details of larval development are known for only a few tapeworms. Cleavage leads to the formation of a six-hooked embryo, the oncosphere, which is enclosed in a ciliated embryophore in the Pseudophyllidea. Further embryonic development occurs in some other host species. Except for Cyclophyllidea, two intermediate hosts are usually required for embryonic development. Except for some taeniaoid tapeworms, the first intermediate host is always an arthropod, while the second is often a fish, but may be an invertebrate. Invariably, the embryo attains entrance to the intermediate host by being eaten, and the vertebrate host is infected by eating the intermediate host.

The postoncosphere development of all eucestodes includes a procercoidlike stage. In most orders this is followed by the development of a plerocercoid larva in a second intermediate host. However, considerable modification of this developmental pattern occurs in the Cyclophyllidea, in which the postproceroid development continues in a single intermediate host and the plerocercoid is modified in one of several different ways. Other modifications of development in Cyclophyllidea include asexual reproduction of new individuals by budding. After differentiation and maturation of the scolex in the larval form, growth ceases until such time as the intermediate host is eaten by a vertebrate in which development of the strobila can occur. In general, the sexual strobilate phase of tapeworms shows considerable specificity with regard to the kind of vertebrate in which development occurs. It is rarely known whether this is due, in a given case, to a failure of contact of worm and host or to a physiological incompatibility. In many instances, such as the eucestodes of birds, tapeworms occur in only a single vertebrate host species. *See* PSEUDOPHYLLIDEA.

Nutrition. The nutritional relationships of adult tapeworms have been extensively studied. The worms require carbohydrate for growth and reproduction and this requirement is satisfied only from the host ingesta. On the other hand, nitrogenous nutrients and many micronutrients may be obtained from the body stores of the host; deleting such materials from the host's diet has no appreciable effect on the worms. In at least one case, *Dibothriocephalus*, it has been shown that the anemia appearing in some humans harboring the worm is attributable to the absorption of vitamin B_{12} by the eucestode. The energy metabolism of tapeworms is primarily fermentative. In air the rate of oxygen uptake is low and accounts for the oxidation of a small fraction of the carbohydrate metabolized.

Phylogeny. Most authorities agree that the tapeworms are ancient parasites, probably evolving as parasites of the earliest fishes. It seems probable that the tapeworms did not evolve from trematodes or other present-day groups of parasitic flatworms. The ancestry of the tapeworm may be directly derived from the acoele or rhabdocoele turbellarians and represents a line of evolution which is completely independent of other parasitic flatworms. These relationships remain obscure in the absence of any fossil records. *See* TURBELLARIA. Clark P. Read

Euclidean geometry

The word geometry is derived from two Greek words meaning earth measurement. It seems probable that many of the early discoveries in geometry were motivated by the need to make measurements of distances and areas on the Earth. However, euclidean geometry has a broader meaning. It is the chief subject matter of the monumental 13-volume work called *The Elements*, written about 300 B.C. by the Greek mathematician Euclid, who taught and founded a school of geometry at Alexandria. One of the milestones in the history of scientific thought, these books of Euclid still occupy an important position in mathematical instruction today.

Geometry, as developed by Euclid, was a systematic body of mathematical knowledge, built by deductive reasoning upon a foundation of three main pillars: (1) definitions of such things as points, lines, planes, angles, circles, and triangles; (2) the assumption of certain geometrical postulates regarded as true but perhaps not self-evident; and (3) the assumption of certain axioms or common notions which were taken to be self-evident truths. The body of Euclid's great work consists of a set of propositions or theorems, each derived systematically and logically from the definitions, axioms, and postulates of his foundation and from theorems already proved.

Definitions. Several of the more important definitions are given in this article. Other definitions and explanations are given in separate articles. *See* CIRCLE; POLYGON.

Point. A point, in the euclidean definition, is that which has no part. A point, according to most modern authors, is an undefined element in a geometry, such that to each pair of points corresponds a unique set of points called a line. In posteuclidean metric geometry, each pair of points A and B determines a (real) number called the distance $\overline{AB}$, which is positive or zero according to whether A and B are distinct points or the same point.

Line. A line, according to the euclidean definition, is breadthless length, the extremities of a line are points, and a straight line is a line that lies evenly with the points of itself. Three points A, B, and C are said to be collinear if, and only if, they lie in

the same straight line. In modern usage the finite straight line of Euclid is called a line segment, and the single word line in geometry commonly refers to a euclidean straight line produced indefinitely in both directions. In metric geometry the point P is said to lie between A and B if $\overline{AP} + \overline{PB} = \overline{AB}$. Three points are collinear if one lies between the other two. A line segment $[AB]$ consists of two distinct points A and B and all the points between them. A line AB consists of two distinct points A and B and all the points collinear with them. A half line, or ray, $A(B)$ consists of the segment $[AB]$ and all points P such that B lies between A and P.

Any two lines (indefinitely extended) are congruent. Two lines having two distinct points in common are coincident; two lines having just one point in common are called intersecting lines. Any two rays (or half lines) are congruent. They are coincident if they have in common the vertex and one other point. Two line segments are congruent if and only if they have the same length.

Skew lines. Two lines that are not coplanar are called skew. Given any two skew lines in space, and a point P not on either, there is a unique line through P that intersects both skew lines. There is a unique plane that contains the first of two given skew lines and that is parallel to the second.

Parallels. Two lines (each extended indefinitely in both its directions) are called parallel if, and only if, they lie in the same plane but do not intersect. Two planes, or a line and a plane, are called parallel if they do not intersect. There is one and only one line parallel to a given line through a given point not on the line (Euclid's parallel postulate). Two distinct lines parallel to the same line (in space or in a plane) are parallel to each other. Two distinct planes parallel to the same plane are parallel to each other. If each of two parallel planes is intersected by a third plane, the lines of intersection are parallel (**Fig. 1**).

Perpendicular. When two straight lines intersect so that adjacent angles at their point of intersection are equal, these angles are called right angles, and the lines are said to be perpendicular, or normal, or orthogonal to each other (**Fig. 2**). Two skew lines are called perpendicular if one of them intersects at right angles a line parallel to the other. In general, a line perpendicular to one of two parallel lines is perpendicular to the other also. In a plane, but not in general in space, two lines each perpendicular to the same line are parallel. One, and only one, perpendicular can be drawn to a given line through a given point not on the line. The same is true in the plane, but not in space, if the given point is on the line. A line that intersects a plane in a point P is called perpendicular

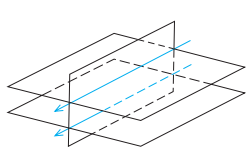

Fig. 1. Parallel lines and planes.

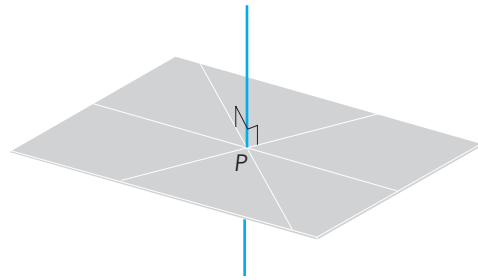

Fig. 2. Normal to a plane.

or normal to the plane if it is perpendicular to every line through P that lies in the plane.

The plane is then called perpendicular to the line. A line perpendicular to each of two intersecting lines is perpendicular to their plane and to every line in the plane. The distance from a point to a line or from a point to a plane is measured along the unique perpendicular from the point to the line or plane. A line perpendicular to a segment at its midpoint is called a perpendicular bisector of the segment. A perpendicular drawn from a vertex of a triangle to the opposite side is called an altitude of a triangle. The three altitudes of a triangle meet in a point H which is called its orthocenter. The three perpendicular bisectors of the sides of a triangle meet in a point P, called its circumcenter, equally distant from all three vertices.

Congruence. Two geometric figures are congruent to each other if the points in the one figure can be made to correspond in a one-to-one manner with the points of the other figure so that the distance between any two points in the one figure is equal to the distance between corresponding points in the other figure. (This is a posteuclidean version of the definition.)

Postulates. Five postulates followed Euclid's list of definitions and may be paraphrased as:

1. A unique straight line can be drawn from any point to any other point.
2. A finite straight line can be extended continuously in either direction in a straight line.
3. A circle can be described with any given center and radius.
4. All right angles are equal.
5. If a straight line falling on two straight lines makes interior angles on the same side with a sum less than two right angles, the two straight lines, if produced indefinitely, meet on that side on which the angle sum is less than two right angles.

This fifth postulate is known as the parallel postulate and is essentially equivalent to the statement that "a unique line parallel to a given line can be constructed through any point not on the line." For 2000 years many unsuccessful attempts were made to prove that this postulate was a consequence of Euclid's other postulates, definitions, and axioms. Only as recently as the nineteenth century did N. I. Lobachevski (1793–1856), J. Bolyai (1802–1860), and G. F. B. Riemann (1826–1866) show that the parallel postulate was independent of the others,

by constructing so-called noneuclidean geometries in which the fifth postulate is not valid. *See* NONEUCLIDEAN GEOMETRY.

Axioms. Certain axioms were stated by Euclid and treated as self-evident truths.

1. Things which are equal to the same thing are equal to each other.
2. If equals are added to equals, the wholes are equal.
3. If equals are subtracted from equals, the remainders are equal.
4. Things which coincide with one another are equal to one another.

A fifth axiom, ascribed by Proclus to Euclid, is believed by Paul Tannery and T. L. Heath to have been added by later writers.

5. The whole is greater than the part.

This has also been replaced by the statement: The whole is equal to the sum of its parts.

Euclid's books of elements. Euclid's great work of systematic reasoning was divided into 13 books. Two other books, at one time ascribed to Euclid, are now believed to have been written by others.

The first six books of Euclid include most of the subject matter commonly taught in high school geometry; books 7, 8, 9, and 10 are concerned with number theory, square and cube roots, and incommensurable magnitudes; books 11 and 12 are concerned with solid geometry and mensuration; and book 13 is concerned with extreme and mean ratio and the five regular solids.

Book 1, after listing a large number of definitions and the postulates and axioms described above, contains 48 propositions, including many properties relating to perpendicular and parallel lines, angles, and congruent triangles, and the book culminates with the important pythagorean theorem.

Book 2 contains 14 propositions on geometrical algebra, establishing relationships between the areas of certain related squares and rectangles. For example, proposition 7 is equivalent to the identity

$$(a + b)^2 = a^2 + b^2 + 2ab$$

Book 3 deals with circles, chords, inscribed angles, and other figures related to the circle; book 4 with inscribed and circumscribed circles and polygons; book 5 with ratios; and book 6 with areas and similar triangles.

Book 7 is number-theoretical, not geometrical, and deals with units, primes, and squares and with the euclidean algorithm for least common divisor. Book 8 includes continued proportions and cube roots. Book 9 includes the factoring of numbers in continued proportion, a theorem on the infinitude of prime numbers, and a formula for even perfect numbers.

Book 10 includes a thorough treatment of quadratic irrationalities and related incommensurable magnitudes.

Book 11 introduces the fundamentals of three-dimensional geometry, including theorems on planes and lines, perpendiculars and parallels, solid angles,

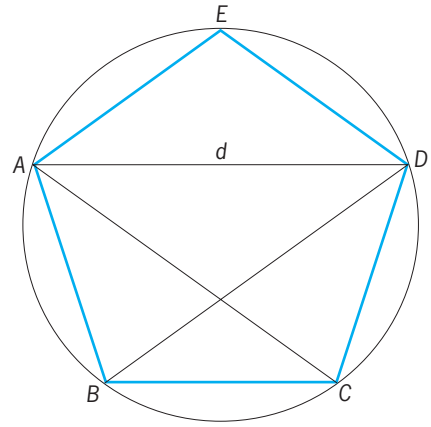

Fig. 3. Regular pentagon *ABCDE* showing the diagonal *d*.

and the volumes of parallelepipedal solids. Book 12 deals with areas and volumes of pyramids and cones and spheres, deriving formulas for the latter by the method of exhaustion due to Eudoxus. Book 13 starts with the division of a segment in extreme and mean ratio, in such a manner than one part of the segment is to the second part as the second part is to the whole. This ratio is the so-called golden ratio of the sides of a regular pentagon to its diagonal d (**Fig. 3**). The ratio is equal to $(\sqrt{5} - 1)/2$, its reciprocal is $(\sqrt{5} + 1)/2$, and it plays a fundamental role in the study of two of the regular solids. The latter half of book 13 is concerned with a study of relations among the volumes and edges and diameters of the five regular solids, often called platonic solids. *See* DIFFERENTIAL GEOMETRY; POLYHEDRON; PROJECTIVE GEOMETRY; SOLID (GEOMETRY). J. Sutherland Frame

Bibliography. H. L. Baldwin, *Essentials of Geometry*, 1993; Euclid, *Elements*, ed. by T. L. Heath, 1926; R. D. Gustafson and P. D. Frisk, *Elementary Geometry*, 3d ed., 1991; P. G. O'Daffer and S. R. Clemens, *Geometry: An Investigative Approach*, 1992.

Eucoccida

The most important order of the protozoan subclass Coccidia, in the class Telosporea, subphylum Sporozoa. In this group there is alternation of asexual and sexual phases of the life cycle, and the parasitic stages occur within the cells of vertebrates or invertebrates. There are three suborders.

In the suborder Adeleina, the macrogamete and microgametocyte are joined during development (syzygy) and the microgametocytes produce only a few microgametes. Members of this suborder include parasites of invertebrates and lower vertebrates; a few even occur in higher vertebrates. In most of them, asexual development takes place in one host and sexual development in another, but the patterns of development and the hosts are too numerous to be described here.

In the suborder Eimeriina, there is no syzygy and the microgametocytes produce a large number of microgametes. This group contains several hundred species of coccidia, most of which occur in the

intestinal cells of vertebrates. They multiply asexually in their host cells, and some species cause coccidiosis, a disease that produces diarrhea, dysentery, and even death. They are especially important in domestic animals.

In the suborder Haemosporina, there is no syzygy and the microgametocytes produce a moderate number of microgametes. Asexual development takes place in the blood cells of vertebrates, and sexual development in blood-sucking insects or mites. This suborder contains *Plasmodium*, species of which cause malaria in humans, lower primates, birds, and other vertebrates. Other genera in this suborder are *Hepatocystis* (which is found especially in African monkeys), *Haemoproteus* (which occurs in birds; wild birds are often infected, but the parasites do not appear to be very pathogenic), and *Leucocytozoon* (which occurs in both wild and domestic birds and may cause a fatal disease in them). *See* COCCIDIA; MALARIA; PROTOZOA; SPOROZOA; TELOSPOREA. Norman D. Levine

Eucommiales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Hamamelidae of the class Magnoliopsida (dicotyledons). The order consists of a single family, genus, and species, *Eucommia ulmoides*, of China. It is a simple-leaved tree with reduced, solitary flowers that lack a perianth, and leaves without stipules, the paired basal appendages found on many leaves. The ovary has two unitegmic (with a single integument) ovules and ripens into a one-seeded, winged fruit. It is widely cultivated as an ornamental. *See* HAMAMELIDAE; MAGNOLIOPSIDA; PLANT KINGDOM. Arthur Cronquist

Eudicotyledons

One of the two major types of flowering plants (angiosperms), characterized by possession of three apertures in their pollen; the other major type is magnoliids. Although this difference in pollen development and form has been known for a long time, it has become clear, as a result of several studies of deoxyribonucleic acid (DNA) sequences, that this difference is very significant. The high degree of coincidence of the genetic data with this pollen distinction means that it is more important to recognize this distinction than the number of seed leaves, as previously thought. *See* DICOTYLEDONS; FLOWER; MAGNOLIALES; MONOCOTYLEDONS.

Mark W. Chase; Michael F. Fay

Euechinoidea

A subclass of Echinoidea (sea urchins, sand dollars, and heart urchins) in which the test plates are consistently arranged in double rows or columns. Five double rows bear the tube feet (the ambulacral rows), and these alternate with five which do not (the interambulacral rows). Eighteen orders are included within the Euechinoidea. *See* ECHINACEA; ECHINODERMATA; ECHINOIDEA; NEOGNATHOSTOMATA. Andrew C. Campbell

Euglenida

An order of the protozoan class Phytamastigophorea. This order of protozoa is also known as Euglenoidina, and botanists classify the group in the class Euglenophyceae. The euglenids include the largest green noncolonial flagellates, *Euglena ehrenbergii*, which are 400 micrometers long. Many of the colorless members are also large. They have one or two equal or subequal flagella. There are relatively few genera of Euglenida, even if doubtful ones such as *Clorachne* and *Ottonia* are included. The freshwater genera *Euglena*, *Lepocinclis*, *Phacus*, and *Trachelomonas* have many species, as does the colorless genus *Petalomonas*. Others have few species, including the marine genus *Eutreptia*. *Calkinsia* is yellow, some euglenas are red at times, and all others are green or colorless, although *Trachelomonas* tests are frequently colored by iron. *See* EUGLENOPHYCEAE; PHYTAMASTIGOPHOREA.

Morphology. Euglenids generally show an anteroposterior elongation; however, *Trachelomonas volvocina* is practically spherical, as shown by young individuals, and *Euglena acus* is extremely needle-like. Most free-swimming members are round in cross section, but species of *Phacus* are flattened, as are *Petalomanas*, *Anisonema*, *Triangulomonas*, and others (see **illus.**). Many are enclosed in a thick pellicle sculptured into keels (*Phacus, Pleotia*, and *Tropidoscyphus*); others, such as *Euglena oxyuris* and *Phacus pyrum*, show strong torsion of the pellicle. This pellicle is sometimes ornamented with striae or small wartlike protuberances. It prevents distortion of the body according to its degree of thickening, but species which have a very thin pellicle (*Astasia*, *Peranema*, *Heteronema*, and some euglenas) shorten on their axis, then elongate. Such contortions are termed euglenoid or metabolic movements. There are sometimes fine openings in the pellicle through which mucous threads are extruded under certain conditions.

Colony formation is restricted to *Colacium*, in which a flagellated euglenoid cell usually attaches to a copepod by its anterior end and develops into an arboroid structure of several nonflagellated cells. Tests or shells are common; *Klebsiella* is a free-swimming marine form in an urn, *Ascoglena* a freshwater epiphyte in an urn. In *Trachelomonas* and *Strombomonas* the tests vary widely in shape and ornamentation and have a narrow pore through which the flagellum emerges.

Distinguishing features. Euglenids have two distinguishing features. They synthesize a starchlike polysaccharide, paramylum, of constant shape for, and often characteristic of, a given species. This

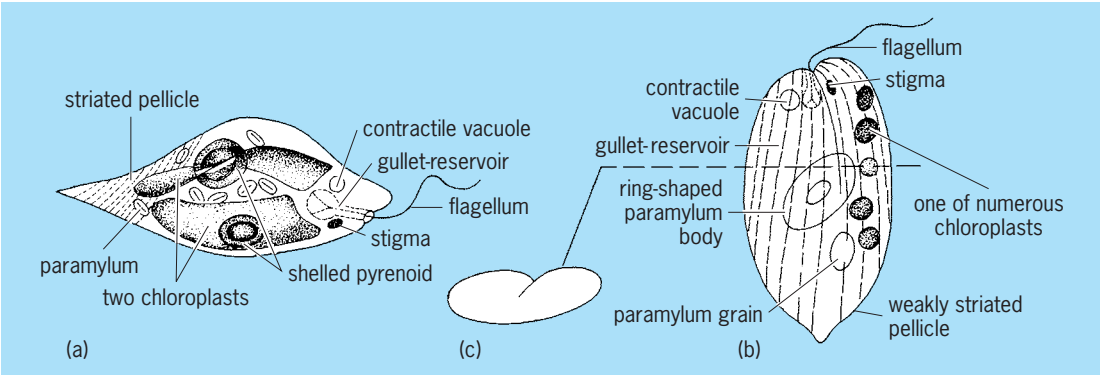

Euglenids. (*a*) *Euglena pisciformis*. (*b*) *Phacus strokesii*. (*c*) Cross section of cell.

assimilation product of both green and colorless forms is often large and lamellated.

The second characteristic is the gullet-reservoir system. An anterior mouth opens into a short gullet which widens into a reservoir. The flagella emerge from its floor, and at one side a contractile vacuole pulses, except in marine species. This system is present in some Chloromonadida (*Gonyostomum semen*) and is approximated by certain Cryptomonadida, but all euglenoids possess it. The mouth may serve to ingest food (*Peranema trichophorum*), but in most cases does not. Green forms possess a stigma near the reservoir, and a thickening on the longer of the two flagella is found in most of them, as well as in some colorless species.

Flagella and chromatophores. In most green species the flagellum is long, uniform, and movable throughout. *Euglena mutabilis* has an extremely short one, as does *E. spirogyra.* According to G. F. Leedale, there is no bifurcation of the locomotor flagellum; instead, a second short flagellum is found separate, but it does not extend beyond the mouth. *Calkinsia* has a long, uniform, anterior one, movable throughout, and a shorter, fine, trailing one. *Peranema* has a long tapering flagellum, vibratile at its tip during steady progression, and a very short, fine, posterior one. *Anisonema acinus* has a rather short, fine, anterior flagellum and a tapering, long, and very heavy posterior one. Chromatophores in the green members may be circular, stellate, or irregular disks, or variously shaped plates and short bands. The chromatophores are reticulate plastids containing bright green pigments closely related to, if not identical with, the chlorophylls of higher plants. The chromatophores of Euglenida never contain other masking pigments, though they are often associated with pyrenoids. In *Euglena pisciformis* both chromatophores have a central pyrenoid which has a shell of paramylum. Pyrenoids stain deeply with hematoxylin and may be proteinaceous. Besides paramylum, fats occur, and glycogen has been reported.

Nuclei. Nuclei are large, vesicular, and contain one or more endosomes which are deeply staining chromatoid bodies. These are surrounded by dispersed chromatin granules. In division, a specific number of chromosomes is formed, but a spindle is lacking, and chromosomes are oriented around and parallel to the drawn-out endosomal mass. Cell division usually occurs in the active state, although *Euglena gracilis* divides while encysted. Palmella stages have been reported but are not common. Sexual processes are unknown. Daily isolations of several clones of *Entosiphon sulcatum* for many months showed no depression of the division rate, and the usual type of mitotic division was enough to maintain species vitality.

Nutrition. Nutrition varies greatly in Euglenida. Green members are holophytic or autotrophic, but there is evidence that some of them absorb soluble organic matter with no evidence of ingestion of solid matter. It seems questionable that more than a few green members are strict autotrophs which do not require organic carbon and nitrogen but form carbohydrates and proteins from carbon dioxide and inorganic salts. Certainly the colorless forms are heterotrophs (dependent on organic food), and some of the green ones readily assume a heterotrophic existence, if they ever were autotrophic. Thus *Euglena gracilis* and species of *Phacus* become colorless in an organic medium in the dark, whereas *Trachelomonas reticulata* is devoid of chlorophyll. *Peranema* ingests large food bodies; however, most of the colorless forms are saprozoic rather than holozoic and are readily cultivated in soluble organic media. They also vary widely in oxygen requirements; the same species occur in well-oxygenated or anaerobic situations. In a rich organic medium and under a film of oil, *Entosiphon* thrives, but often forms somatellalike masses, in which, however, cell division does not become complete. Parasitism is unknown in the class, although *Euglenomorpha* occurs in the intestine of tadpoles, where it is probably commensal.

Ecology. In habitat the Euglenida are widespread. Green members occur mostly in fresh water and frequently in such numbers as to form blooms. This is especially true for pools contaminated by cattle droppings. *Strombomonas* in the Ohio Valley of the United States is a potamophile form, *Trachelomonas* a lake or pond form. *Eutreptia* blooms densely in marine bays and estuaries, and the mud-water

interface in marine waters is a rich source of the colorless creeping members. Iron seepages and coal-mine seepages containing sulfuric acid are often bright green with *Euglena mutabilis* even at pH values of 1.5, whereas cedar swamps have blooms of *Euglena polymorpha*. Unlike the Phytomonadida, the Euglenida appear in warmer waters but do not seem to occur in hot springs. They abound in citrus-waste lagoons and in later stages of sewage-treatment waters, an indication of their partially saprozoic nutrition.

Relationships within the class are not well defined at present; each classification within recent years has departed from previous ones. Organisms are readily assigned to the Euglenida on a structural basis, because the class is quite specialized. However, intraclass relationships are difficult to assign. *See* CILIA AND FLAGELLA; EUGLENOPHYCEAE; MASTIGOPHORA; PROTOZOA.

James B. Lackey

Euglenophyceae

A class coextensive with the division Euglenophycota, comprising unicellular colorless or photosynthetic flagellates with very distinctive cytological characters. In protozoological classification, these organisms constitute an order, Euglenida, of the class Phytomastigophora. Although photosynthetic euglenoids are like Chlorophycota in containing chlorophyll *a* and *b*, in most other respects they are so different from those algae as to suggest an independent phylogenetic origin of their pigments. About 1000 species have been described and classified into about 40 genera and 6 orders. *See* EUGLENIDA.

Most euglenoids are free-swimming and have two flagella, one of which may be nonemergent, arising from an anterior invagination known as a reservoir (see **illus.**). *Euglenamorpha* and *Hegneria*, both inhabitants of the rectum of tadpoles, have three to seven flagella. A few rhizopodial forms lack flagella. In *Colacium*, mucilaginous stalks join *Euglena*-like cells in dendroid colonies attached to small aquatic animals. In some species of *Euglena* and *Eutreptia*, cells cohere in palmelloid sheets. *Ascoglena*, although flagellate, is sessile.

Structure. Euglenoid cells have helical symmetry, usually with superimposed bilateral symmetry, and are 15–500 micrometers long. *Euglena gracilis*, the best-known species, is spindle-shaped, while other species are spherical, flattened, or twisted. All euglenoids have a characteristic proteinaceous pellicle composed of flexible or rigid, interlocking, spirally disposed strips lying beneath the cell membrane. In *Trachelomonas*, *Strombomonas*, and *Ascoglena*, an inorganic envelope (lorica) is deposited outside the cell membrane. In species with flexible pellicles, flagellar locomotion may be supplemented or supplanted by so-called euglenoid movement, a peristalsis involving the entire cell. The surface of all euglenoids is coated to a varying extent with mucilage secreted by muciferous bodies (mucocysts) that traverse the pellicle.

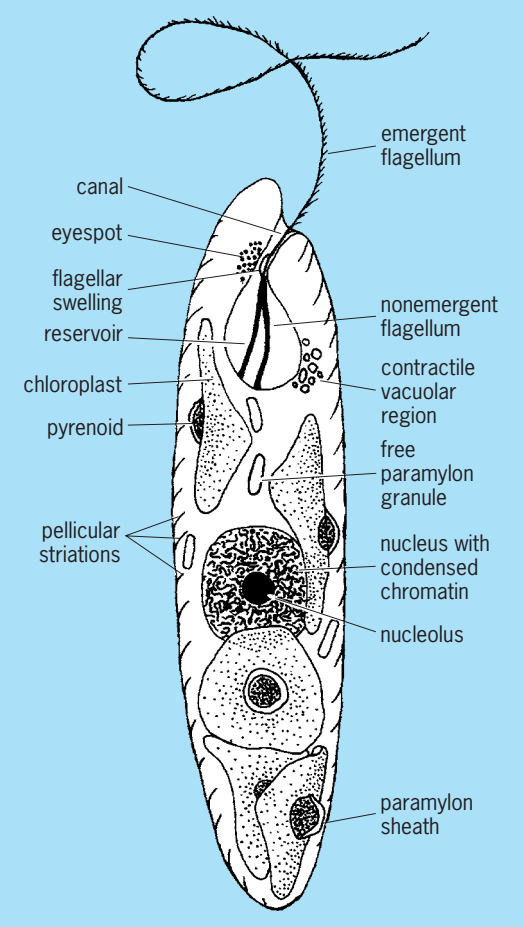

***Euglena* sp., showing overall morphology and intracellular structures.**

Emergent flagella, which are covered with fine hairs, contain a paraflagellar rod in addition to the usual 9 + 2 array of axoneme microtubules. At the base of one flagellum is a swelling believed to be a photoreceptor that acts in concert with the eyespot in phototaxis. The eyespot, which is situated in the cytoplasm at one side of the reservoir, consists of lipid globules not surrounded by membranes. With the possible exception of a few marine and parasitic species, all euglenoids have a contractile vacuolar region adjoining the reservoir. A single nucleus is present in each cell. As in dinoflagellates, the chromatin is permanently condensed so that chromosomes are visible throughout the cell cycle. The nuclear membrane remains intact during mitosis. *See* DINOPHYCEAE.

Chloroplasts. Photosynthetic euglenoids contain one to many grass-green chloroplasts, which vary from minute disks to expanded plates or ribbons. Each chloroplast is surrounded by three membranes, the outer of which is discrete and not associated with endoplasmic reticulum. Photosynthetic lamellae are composed of three appressed thylakoids. They contain β-carotene and various xanthophylls in addition to chlorophyll *a* and *b*. An extraplastidial carotenoid causes a few species of *Euglena* to be red. Pyrenoids are present in many species. The reserve product of

photosynthesis is paramylon, a β-1,3 glucan stored outside the chloroplast in the form of rods or disks. Paramylon is also stored in colorless species. The chloroplast of some species can be experimentally eliminated or rendered nonfunctional by heat, ultraviolet light, or antibiotics.

Nutrition. Colorless euglenoids depend on osmotrophy or phagotrophy for nutrient assimilation. Some phagotrophs, such as *Peranema*, have a cytostome, an opening in the anterior of the cell through which particulate food is ingested. The cytostome, which is sometimes accompanied by other specialized ingestion organelles, is independent of the reservoir. Some photosynthetic euglenoids are facultative osmotrophs, but none are facultive phagotrophs.

Reproduction, use, and habitat. Euglenoids reproduce solely by longitudinal binary fission, sexuality being unknown. Thick-walled resting stages (cysts) are common in many genera. *Euglena gracilis*, which can be grown easily in bacteria-free culture, has been used extensively in the study of photosynthesis, nucleic acid synthesis, membrane generation, and other aspects of physiology and molecular biology. Its requirement for vitamin B_{12}, the antipernicious anemia factor in animals, makes it useful in medical bioassays.

Euglenoids are found most commonly in fresh water rich in organic matter, but they also occur in marine or brackish habitats, on mud or sand, and in ice or snow. A few species prefer very acidic water. There are three genera of endoparasites, inhabiting fresh-water flatworms, annelids, copepods, and tadpoles.

Paul C. Silva; Richard L. Moe

Bibliography. D. E. Buetow (ed.), *The Biology of Euglena*, vols. 1 and 2, 1968, vol. 3, 1982; G. F. Feedale, *Euglenoid Flagellates*, 1967; M. Gojdics, *The Genus Euglena*, 1953; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Eukaryotae

The vast array of living and fossil organisms with complex cells in which the genetic material is organized into chromosomes (linear pieces of DNA bound with histone proteins and visible as compact structures during mitosis and meiosis) and contained within a membrane-bound nucleus. The Eukaryotae have a common origin from organisms with prokaryotic cells (bacteria and archaea) that do not have their DNA in chromosomes and lack membrane-bound organelles. The Eukaryotae thus include all multicellular plants, fungi, and animals, as well as a collection of the unicellular protists. *See* CELL NUCLEUS; CHROMOSOME.

Characteristics. All organisms in this group possess a eukaryote cell that is structurally more complex than the prokaryote cell of archaea and bacteria. In addition to a nucleus, the cell contains a number of other membrane-bound cytoplasmic organelles, such as mitochondria, Golgi bodies, lysosomes, peroxisomes, endoplasmic reticulum, and, in photosynthetic forms, plastids or chloroplasts. Rather than being locked within a rigid cell wall like prokaryotes, eukaryotes have evolved a complex cytoskeleton of microfilaments and microtubules that give structural support to the cell and allow the cell membrane to grow and change shape. Characteristically, centrioles, cilia, or flagella are also present, the latter locomotory organelles composed mainly of tubulin with microtubules generally in a 9 + 2 pattern. Therefore, there is a major discontinuity, at the cellular level alone, between the Prokaryotae and the Eukaryotae. *See* CELL (BIOLOGY); PROKARYOTIC CELL.

Biology. Organization at the organismal level ranges from solitary unicellular (as is common in many of the algal and protist groups) to colonies of cells, mycelial, syncytial (coenocytic), and truly multicellular with extensive tissue differentiation. Modes of nutrition run the gamut: absorptive, ingestive, photoautotrophic, plus combinations of these three. Life cycles vary tremendously and may include both asexual and sexual methods of reproduction. Asexual reproduction can be by cell division for unicellular forms, or by vegetative growth or budding in multicellular forms. Exchange of genetic material (sexual reproduction) takes place in the majority of eukaryotes by fusing haploid gametes such as egg and sperm or pollen to form a diploid embryo, but variations such as conjugation (where nuclei are exchanged) may also occur.

Aerobic metabolism is commonly exhibited, especially by aquatic and terrestrial forms; anaerobic mechanisms exist, however, for numerous species found in poorly oxygenated habitats, including various sites within bodies of host organisms.

Diversity. Size of species ranges considerably (from unicellular protists to large multicellular plants and animals). The size of the eukaryotic cell is usually 20–50 micrometers in diameter, but mammalian red blood cells are just 8 μm in diameter while some single-celled protists, such as the ciliate *Spirostomum*, are over 1 mm (1000 μm) in length. Habitats cover all possible ecological niches: aquatic, terrestrial, and aerial, for free-living forms; and in or on all kinds of hosts, for symbiotic or parasitic forms (internal habitats, including cells, tissues, organs, or various body cavities). Dormant stages include cysts, spores, and seeds; these are often involved in dispersion or propagation of the species.

Some species that cause fatal diseases in other eukaryotic organisms (including humans) are as important economically and medically as some of the parasitic members of the superkingdom Prokaryotae. Parasitic life cycles may involve multiple hosts, specialized vectors, and so on. *See* ANIMAL KINGDOM; FUNGI; PLANT KINGDOM; PROTOZOA.

Diana L. Lipscomb; John O. Corliss

Bibliography. L. Margulis and K. Schwartz, *Five Kingdoms: An Illustrated Guide to the Phyla of Life on Earth*, 3d ed., 1998; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Eulamellibranchia

The largest subclass of the class Bivalvia in the phylum Mollusca. Eulamellibranchs are ciliary filter feeders with greatly enlarged ctenidia, in each of which the elongated gill filaments are held together in parallel series to form folded lamellae. The 30,000 living species constitute the most successful of molluscan groups, in terms both of individual numbers with their biomass and of ecological bioenergetics (calorie transfer in food chains).

Classification. There are several proposed systems of ordinal classification for the subclass Eulamellibranchia, but little agreement on the best structural bases. One system frequently employed comprises six orders: Anisomyaria, Taxodonta, Heterodonta, Schizodonta, Anomalodesmata, and Adapedonta. There is more general acceptance of the (possibly 18) superfamilies of eulamellibranchs, each with characteristic shell structure, hinge dentition, ligament, and degree of fusion of the mantle lobes involving the origin of the siphons. The more important superfamilies include Mytilacea (mussels), Ostreacea (true oysters), Cardiacea (true cockles), Unionacea and Corbiculacea (fresh-water clams), Myacea (including softshell clams), Mactracea (surf clams), Tellinacea, and Pholadacea (=Adesmacea). Within each larger superfamily there is a wide range of adaptive radiation, with different genera living as shallow or deep burrowers or as sessile attached forms. There can be considerable evolutionary convergence shown by genera from different superfamilies, and deep burrowers with massive fused siphons are found in five superfamilies. *See* BORING BIVALVES.

Filter feeding and digestion. The four folded lamellae of the paired gills are morphologically homologous (in blood vessel and ciliary arrangements) with the archetypic molluscan ctenidia and their functional orientation is also similar. The lateral cilia on the filaments produce a nearly continuous water current between adjacent filaments which flows from the outer (ventral) inhalant part of the mantle cavity to the inner (dorsal) exhalant part. All suspended material, including food organisms brought into the outer mantle cavity by the water current, is filtered and accumulates on the inhalant faces of the gill lamellae. Some of the filtration is accomplished by the frontal cilia, but most is due to the laterofrontals which are compound cilia with a finely pinnate structure. Accumulated material and food is then moved by the frontal cilia toward the ventral edges of the gills, and the fine food particles become concentrated in the food grooves. Coarser and heavier materials (such as sand grains) do not enter the food grooves but are rubbed or twitched off onto the mantle wall. Sorting is carried out in food grooves and on the labial palps on a size basis; palp cilia carry fine material into the mouth, from which ciliary tracks carry it slowly and continuously into the esophagus and to the stomach, where it undergoes further sorting.

In the stomach, extracellular digestion begins. In all eulamellibranchs, the peculiar secreted structure termed the crystalline style is used as a cilia-driven gastric stirring rod and enzyme store, allowing a slow continuous release of amylase and glycogen-breaking enzymes. Food particles then pass into the tubules, the cells of which are phagocytic so that final intracellular digestion takes place. Indigestible material rejected from the stomach complex passes to the posterior intestine and rectum, where water reabsorption and a neutral pH consolidate the true feces into very hard pellets. This process avoids self-contamination of the feeding mechanism.

The coarser particulate material filtered off by the feeding structures but not entering the gut is termed pseudofeces. It is collected by the cleansing cilia of the inside of the mantle wall into ciliary vortices. The pseudofeces are expelled from the bivalve by spasmodic contractions of the adductor muscles which force water and accumulated pseudofeces out through the normally inhalant openings to the mantle cavity. The anus and the renal and genital openings are in the exhalant part of the mantle cavity (as in all mollusks); therefore, expulsion of wastes or reproductive products is not accomplished by these spasmodic cleansing movements but by the normal and continuous water flow of the feeding and respiratory current.

Feeding life-styles. A wide variety of eulamellibranch life-styles and corresponding shell forms can be based on this successful (and relatively uniform) filter-feeding mechanism. The majority of eulamellibranchs are suspension feeders filtering living phytoplankton and bacteria from the overlying water. A substantial minority are detritus feeders that siphon in and filter organic materials (living and dead) from the surface of the sea-bottom deposits. Different styles of suspension feeders include relatively active shallow burrowers with globular shells and short siphons such as cockles and quahogs, sedentary deep burrowers (and some borers) with long fused siphons such as softshell clams, attached forms such as mussels, and permanently cemented forms such as oysters. Detritus feeders (principally superfamily Tellinacea) are all active deep burrowers with long extensible and separate siphons, the inhalant siphon actively sweeping the sea bottom, and with particularly intensive sorting on gill lamellae and palps.

Ecology. The marine eulamellibranchs as suspension feeders consume a major part of the largest plant crop in the world, the marine phytoplankton. Certain eulamellibranch species are the most abundant marine benthic animals; beds of mussels and of oysters are found in satellite images of coasts and shallow seas, and less obvious burrowing clams can be as abundant. Computations have suggested that, in size of standing crop biomass for individual species (annual means), certain eulamellibranchs outrank all other animal species on the planet. In turn, eulamellibranchs are a major food of all bottom-living fishes (including cod, halibut, plaice, and haddock), and are second only to planktonic copepods (such as *Calanus*) in annual turnover of calories through animal tissues for all marine food webs. Birds, marine

mammals, and humans also feed directly on eulamellibranchs. *See* FOOD WEB; MARINE ECOLOGY.

Economic significance. A few eulamellibranch species have negative economic significance for humans. Teredinids or shipworms (of superfamily Pholadacea) bore destructively in marine wood structures such as dock pilings and boats. Certain small eulamellibranchs (*Corbicula* spp. and the zebra mussel, *Dreissena*) are recent introductions to fresh waters in North America and have quickly achieved pest status, fouling water intakes and clogging filters at power plants and municipal water supplies. *See* BIVALVIA; MOLLUSCA; SHIPWORM. W. D. Russell-Hunter

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; W. D. Russell-Hunter, *A Life of Invertebrates*, 1979; K. M. Wilbur (ed.), *The Mollusca*, 11 vols., 1982–1989; C. M. Yonge and T. E. Thompson, *Living Marine Molluscs*, 1976.

Euler angles

Three angular parameters that specify the orientation of a body with respect to reference axes. They are used for describing rotating systems such as gyroscopes, tops, molecules, and nonspherical nuclei. These parameters are not symmetrical in the three angles but are simpler to use than other rotational parameters.

Unfortunately, different definitions of Euler's angles are used, and therefore it is confusing to compare equations in different references. The definition given here is the majority convention according to H. Margenau and G. Murphy.

Let *OXYZ* be a right-handed cartesian (right-angled) set of fixed coordinate axes and *Oxyz* a set attached to the rotating body.

The orientation of *Oxyz* can be produced by three successive rotations about the fixed axes starting with *Oxyz* parallel to *OXYZ*. Rotate through (1) the angle ψ counterclockwise about *OZ*, (2) the angle θ counterclockwise about *OX*, and (3) the angle ϕ counterclockwise about *OZ* again. The line of intersection *OK* of the *xy* and *XY* planes is called the line of nodes.

Denote a rotation about *OZ*, for example, by *Z* (angle). Then the complete rotation is, symbolically, given by Eq. (1) where the rightmost operation is done first.

$$R(\psi, \theta, \phi) = Z(\phi)X(\theta)Z(\psi) \qquad (1)$$

A point *P* will have coordinates (x, y, z) with respect to the body axes and (X, Y, Z) with respect

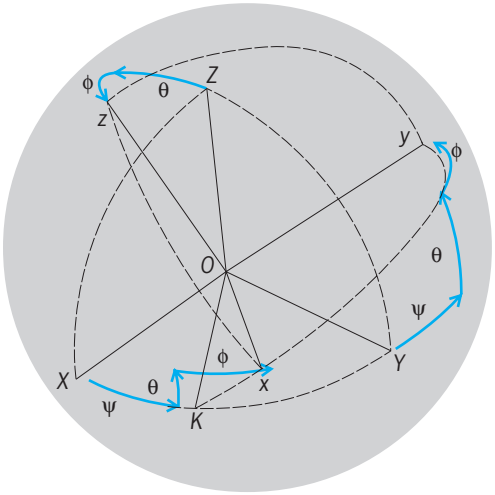

Euler angles. The successive movements of the axes on a unit sphere described in the text are shown by arrows. The complete rotation may also be obtained by a different sequence of rotations, namely, first through ϕ about *OZ*, then through θ about the displaced *x* axis (which is *OK*), then through ψ about *OZ*.

to the fixed axes. These are related by linear Eqs. (2)–(4), where (x, X) is the angle between the

$$x = X\cos(x, X) + Y\cos(x, Y) + Z\cos(x, Z) \qquad (2)$$

$$y = X\cos(y, X) + Y\cos(y, Y) + Z\cos(y, Z) \qquad (3)$$

$$z = X\cos(z, X) + Y\cos(z, Y) + Z\cos(z, Z) \qquad (4)$$

axes *Ox* and *OX*, and so forth. The nine direction cosines are expressed in terms of the three Euler angles in the **table**.

Inspection of the **illustration** makes it apparent that no operation in $R(\psi, \theta, \phi)$ can be replaced by a combination of the other two. Therefore, three parameters are needed to specify the orientation, and the amounts of the angles are unique (barring additional 360° rotations). In dynamical problems of rotating bodies, ψ, θ, and ϕ can be used as independent angular coordinates.

Molecules and nuclei undergo oscillatory changes in shape while rotating. The body axes apply to the average shape. For another set of rotational parameters *see* CAYLEY-KLEIN PARAMETERS.

Bernard Goodman

Bibliography. A. P. Arya, *Introduction to Classical Mechanics*, 2d ed., 1997; R. Baierlein, *Newtonian Dynamics*, 1983; A. L. Fetter and J. D. Walecka, *Theoretical Mechanics of Particles and Continua*, 1980, reprint, 2003; H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., 2002.

Direction cosines stated in terms of Euler angles

	X	*Y*	*Z*
x	$\cos\psi\cos\phi - \sin\psi\sin\phi\cos\theta$	$\cos\psi\sin\phi + \sin\psi\cos\phi\cos\theta$	$\sin\psi\sin\theta$
y	$-\sin\psi\cos\phi - \cos\psi\sin\phi\cos\theta$	$-\sin\psi\sin\phi + \cos\psi\cos\phi\cos\theta$	$\cos\psi\sin\theta$
z	$\sin\psi\sin\theta$	$-\cos\phi\sin\theta$	$\cos\theta$

Euler's equations of motion

A set of three differential equations expressing relations between the force moments, angular velocities, and angular accelerations of a rotating rigid body. They are equations of motion in the usual dynamical sense, having the form of Eqs. (1)–(3). The formula-

$$I_1(d\omega_1/dt) + (I_3 - I_2)\omega_2\omega_3 = M_1 \quad (1)$$

$$I_2(d\omega_2/dt) + (I_1 - I_2)\omega_3\omega_1 = M_2 \quad (2)$$

$$I_3(d\omega_3/dt) + (I_2 - I_1)\omega_1\omega_2 = M_3 \quad (3)$$

tion employs as coordinate axes the three principal axes of rotational inertia of the body that can rotate about a body-fixed point, which is the center of mass if constraints are absent. These reference axes, which form a right-hand set, are indicated by subscripts 1, 2, and 3 in the equations, where I_1, I_2, and I_3 represent the principal moments of inertia; ω_1, ω_2, and ω_3 are the angular velocities about the axes; M_1, M_2, and M_3 are the corresponding force moments; and t is the time.

In the general case, these equations cannot be integrated, but solution is possible in special cases. Soluble problems of interest include that in which force moments are absent, the resulting complex behavior being called Poinsot motion, and that in which two of the principal moments of inertia are identical and only one force moment is present. The latter case includes spinning tops and gyroscopes. *See* RIGID-BODY DYNAMICS.

Russell A. Fisher

Bibliography. R. Baierlein, *Newtonian Dynamics*, 1983; H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; T. W. B. Kibble and F. H. Berkshire, *Classical Mechanics*, 5th ed., 2004.

Eumagnoliids

An informal set of flowering plant orders that contains Laurales, Magnoliales, Piperales, and Winterales, plus the monocotyledons, and in total contains roughly one-third of angiosperm species. Formerly, many have equated this group (minus the monocotyledons) to the subclass Magnoliidae, but studies of sequences of deoxyribonucleic acid (DNA) have demonstrated that orders such as Nymphaeales and Ranunculales are distantly related to the others and that the monocotyledons are much closer. A formal taxonomy for this group is not yet feasible because until more data are collected the exact relationships of the eumagnoliids to the eudicots are not clear, and before formal names are proposed it is preferable to understand these taxa better. The eumagnoliids, including the monocotyledons, exhibit a set of characteristics that reflect their relatively close relationships; these include a tendency to trimerous flowers and a suite of chemical compounds restricted to the group. They also have a set of traits that are "primitive" by most estimations; these include vesselless wood, laminar stamens, and laminar placentation. It now seems quite clear that although the eumagnoliids exhibit some primitive traits they are neither the most archaic nor the earliest of angiosperms, and they contain some of the most advanced taxa, such as many of the monocotyledons (grasses, palms, and orchids). *See* LAURALES; LILIOPSIDA; MAGNOLIALES; PIPERALES; PLANT TAXONOMY.

Mark Chase

Eumalacostraca

The major subdivision (subclass) of the crustacean class Malacostraca. It comprises three, four, or five Recent superorders: Syncarida, Peracarida, and Eucarida are most commonly included and all that are recognized here. However, some carcinologists still insist on ranking the Hoplocarida among eumalacostracans. Acceptance of a separate superorder, Pancarida for the Thermobaenacea, has not received general support, and thermobaenaceans are commonly now included with peracaridan taxa. *See* EUCARIDA; PERACARIDA; SYNCARIDA.

Formerly, the concept of the Eumalacostraca as a homogeneous evolutionary unit generated virtually no controversy. The subclass was defined by a series of characters referred to as the caridoid facies: carapace enclosing the thorax; stalked, movable eyes; biramous antennules; scalelike antennal exopod; thoracopods with natatory exopods; abdomen with well-developed, complex musculature designed for flexing it; telson and uropods forming a tailfan; biramous pleopods 1–5; and internal organs primarily excluded from the abdomen. Opponents of the caridoid theory point to the fact that several of these attributes are not restricted to eumalacostracans, and conversely, not all eumalacostracans possess all of them. The carapace, for example, is absent in syncarids and the peracaridan isopods and amphipods, but present in all other eumalacostracans. Eyes are sessile or lacking in bathynellaceans, thermobaenaceans, cumaceans, isopods, and amphipods, and either stalked or sessile in anaspidaceans. Similarly, the antennal scale is lacking in most Peracarida (except the Mysidacea).

Basic structural plan. The basic body plan common to all eumalacostracans includes a head of five somites, a thorax of eight somites, and an abdomen of six somites, plus a terminal telson. Zero to three pairs of thoracic somites may be fused to the head and their thoracopods and modified to form feeding appendages (maxillipeds). The thorax usually has five to eight pairs of ambulatory or specialized appendages (pereopods); the abdomen may or may not have paired appendages (pleopods) on the first five somites; and the sixth somite usually has a pair of well-developed appendages (uropods). Female gonopores (coxal or sternal) are present at the level of the sixth thoracic somite, while male gonopores are at the level of the eighth.

Carapace. The carapace, its origin and evolution, has provided the greatest controversy. Classically, the carapace, or cephalic shield, was believed to have had its origin as a fold of integument arising

from the posterior border of the maxillary somite of the cephalon. With expansion, anteriorly and posteriorly, cephalothoracic segments became dorsally fused to it. In one alternative hypothesis, the carapace is thought to originate as branchiostegal folds that grow outward and upward from one or more thoracic somites. A second suggests that the carapace is a dorsal expansion of the protocephalic area (antennal somite when differentiated), and that the cephalic and thoracic tergites do not enter into the process at all.

Optics. Three distinct types of optical mechanisms are present in eumalacostracan eyes: apposition optics, often with hexagonal facets; refracting superposition eyes with hexagonal facets; and superposition eyes with square facets. Apposition optics are common to decapod larvae and most adult brachyuran crabs. Refracting, superposition eyes with hexagonal facets occur in mysids and euphausiids, whereas similar eyes, but with square facets using mirror optics, are found in adult shrimps, lobsters, and dromiid crabs. Apposition optics are thought to reflect the primitive condition, with a trend to sessile eyes and eventually total eye loss associated with adaptation to an interstitial life. *See* EYE (INVERTEBRATE).

Locomotion. Two caridoid attributes shared by the majority of eumalacostracans are the strong abdominal musculature and well-developed tailfan. Together with accompanying neural elements, these appear to be designed as a highly efficient propulsion mechanism, that is, an escape reaction. Rapid flexion of the abdomen brings the tailfan forward, providing for the sudden propulsion of the animal backward. Studies suggest that the antennal scale also may play a role in this escape reaction.

Despite major differences in habitat and behavior, other aspects of ambulatory locomotion show considerable intraordinal resemblance. These similarities are attributed to the internal continuity in pereopodal skeletomusculature that exists among the various eumalacostracan orders. In thoracic appendages not modified as maxillipeds, the entire limb lies in a plane that remains unchanged with normal locomotion, but can bend to accommodate actions such as feeding, swimming, and burrowing. Walking is accomplished by extension and flexion within the limb plane, whereas "rowing" motions occur from the limb base. The presumably primitive pattern, found in syncarids and eucarids, involves the development of the pereopodal coxa as a gimbal. Dicondylic body-coxa articulation allows promotion-remotion, and dicondylic coxa-basis articulation allows abduction-adduction. In contrast, the body-coxa articulation in peracaridans, other than mysids, is either immobilized or capable of only limited abduction-adduction, whereas the coxa-basis articulation is monocondylic and capable of performing a complete suite of motions.

In swimming, which has been hypothesized to have evolutionarily followed ambulation, both the natatory exopods of the thoracic appendages and the pleopods work in concert. However, in interstitial eumalacostracans, some peracaridans, and many higher decapods, there is a tendency toward reduction of the pleopods. Swimmers among these groups rely almost exclusively on the thoracic exopods which are provided with long, marginal setae.

Respiration. Respiration in its simplest form is provided by the thoracic epipods. These vascularized, saclike outgrowths of the coxae form primitive gills; the blood vessels are marginal with transverse connectors. Epipodal gills are found in the Syncarida and several orders of the Peracarida. However, in isopods the pleopods have taken on the respiratory function. These gills may vary from simple branchial epithelium to tufts of branchial filaments, and in terrestrial species specialized air sacs (pseudotrachiae) may develop from evaginations on the inner sides of the exopods. Gills of the Eucarida are more complex, each consisting of a central axis with numerous branches, and may also involve countercurrent circulation.

Circulation. The eumalacostracan circulatory system typically consists of an elongate dorsal blood vessel, part of which is expanded in the thoracic region to form a heart with one or more paired ostia, often with a descending (sternal) artery, and with anteriorly, posteriorly, and ventrally directed arteries. Exceptions are found in most peracaridans where a sternal artery is absent. In isopods where the heart is located in the abdomen, lateral arteries provide blood to the pereopods, pleopods, and pleotelson, and the anterior aorta supplies the head. The length of the heart is substantially reduced in euphausiids and decapods but remains positioned in the thorax.

Reproduction. Most eumalacostracans are dioecious; however, some isopods, tanaids, and a few decapods are hermaphroditic. Syncarids lay their eggs on the substrate after copulation. Peracarids brood their embryos in ventral (or in thermobaenaceans, dorsal) brood pouches. Female eucarids, for the most part, carry their eggs in egg sacs (euphausiids) or attached directly to the pleopods; however, penaeid shrimp and some euphausiids shed their eggs directly into the sea. Fertilization is by means of spermatophores and may be external or internal.

Syncarid sperm ultrastructure (based on *Anaspides*) is described as having a very elongate subacrosomal filament which is filiform and coiled and bypasses the nucleus. A nuclear membrane persists. In the Peracarida, spermatozoa consisting of two convergent linear components, the principal sperm body and a transversely striated, taillike, nonflagellate structure, are found in mysidaceans, amphipods, isopods, and cumaceans. Only in the Tanaidacea are the sperm rounded and lacking appendages, a condition similar to that of syncarids. Little is known of euphausiid spermatozoa except that they are ovoid and lack appendages. Both euphausiids and stenopodidean decapods appear to lack an acrosome. Sperm in dendrobranchiate and procarid shrimp have a single acrosomal spine, but rarely with arms corresponding to those characteristic of other decapods.

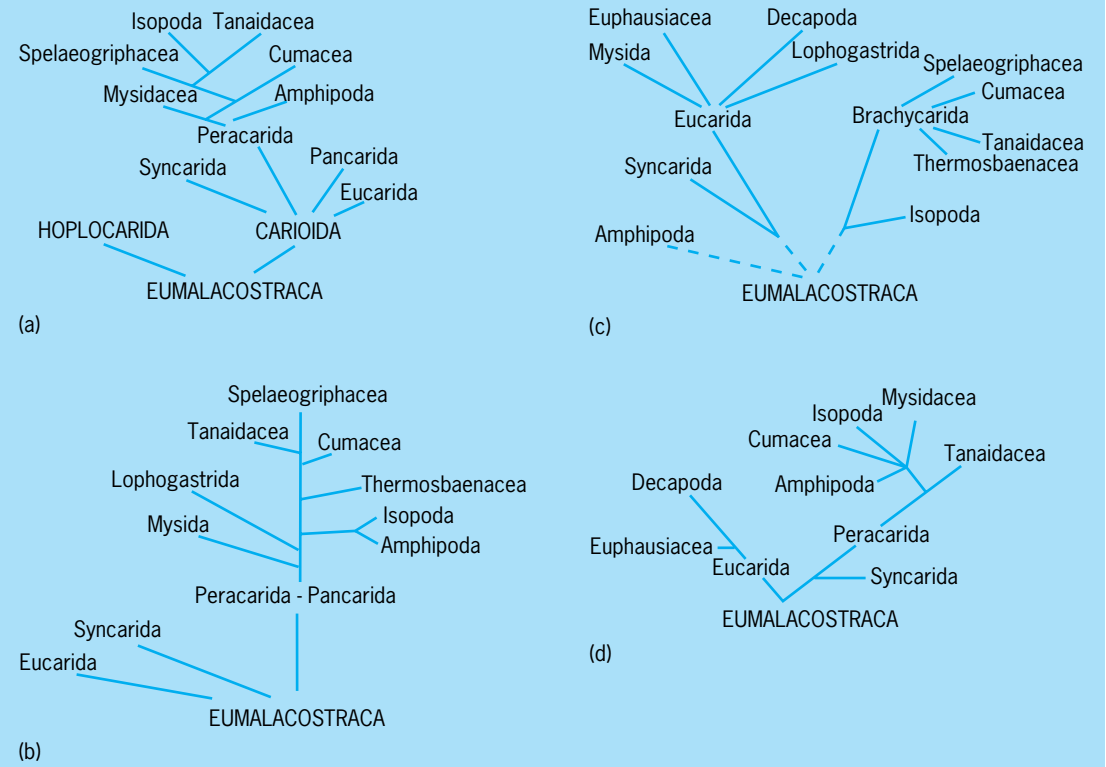

Phylogeny of the Eumalacostraca. (*a*) Based on the caridoid facies concept (*after F. Schram, ed., Crustacean Phylogeny, Crustacean Issues, A. A. Balkema NE, vol. 1, 1983*). (*b*) Based on general morphological characters (*after F. R. Schram, Crustacea, Oxford University Press 1986*). (*c*) Based on selected morphological characters (*after F. Schram, ed., Crustacean Phylogeny, Crustacean Issues, A. A. Balkema NE, vol. 1, 1983*). (*d*) Based on spermatozoa ultrastructure (*after B. G. M. Jamieson, Ultrastructure and phylogeny of crustacean speratozoa, Mem. Queensland Mus., vol. 31, 1991*).

Several spikes are characteristic of lobsters and crabs.

Systematics and phylogeny. A cursory review of eumalacostracan classifications reflects the current lack of consensus. Central to the differing views is disagreement on the origin and evolution of the carapace, and the identity of a eumalacostracan precursor. Proponents of retention of the Hoplocarida (Stomatopoda) in the Eumalacostraca have argued not only that the carapace has a maxillary origin but that the presence of the carapace in phyllocarids establishes its genesis well in advance of the advent of eumalacostracans. Additionally, certain caridoid attributes are primitive (plesiomorphic) or probably characterized the eumalacostracan ancestor. Advanced caridoid attributes (apomorphies) are, for the most part, related initially to adaptations to pelagic life and improved benthic ambulation.

Opponents point to the fact that if these advanced caridoid attributes were derived from an ancestral eumalacostracan, those eumalacostracans that lack such attributes, particularly the carapace, would have had to have arisen through degeneration of these attributes. Rather than the carapace being a primitive caridoid attribute, it is viewed then as an advanced character, evolved initially to serve a hydrodynamic function. Unity among the three eucaridan superorders relies on the presumed apomorphies of the escape reflex.

The possibility that the peracarids branched off from a presyncarid ancestor before the eucarids also has been proposed. Attainment of the advanced caridoid attributes then arose independently within the Peracarida and Eucarida (see **illus.**). *See* CRUSTACEA; MALACOSTRACA.

Patsy A. McLaughlin

Bibliography. D. E. Bliss et al. (eds.), *The Biology of Crustacea,* vols. 1, 2, and 5, 1982–1983; R. C. Brusca and G. J. Brusca, *Invertebrates,* 1990; A. A. Fincham, Eyes and classification of malacostracan crustaceans, *Nature*, 287:729–731, 1980; R. R. Hessler, The structural morphology of walking mechanisms in eumalacostracan crustaceans, *Phil. Trans. Roy. Soc. Lond.,* B296:245–298, 1982; P. A. McLaughlin, *Comparative Morphology of Recent Crustacea,* 1980; F. Schram (ed.), *Crustacean Phylogeny, Crustacean Issues,* 1983.

Eumycetozoida

An order of Mycetozoia. These are slime molds which form a plasmodium, a multinucleate stage that may in some species measure a foot or more across. The plasmodium, typically found on decaying plant material, is a migratory phagotroph preying on microorganisms. A mature plasmodium is a sheet of protoplasm containing a network of channels through which rapid endoplasmic streaming occurs. Direction of flow is reversible at intervals (shuttle-flow movement), but with a net unidirectional movement

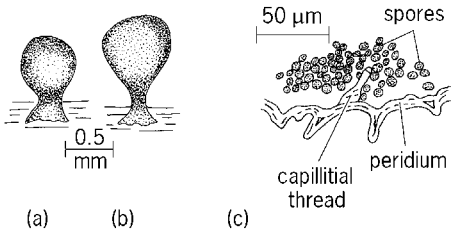

Fig. 1. ***Arcyria cinerea*. (*a*, *b*) Development of sporangium. (*c*) Mature sporangium. (*After R. P. Hall*, *Protozoology*, *Prentice-Hall*, *1953*)**

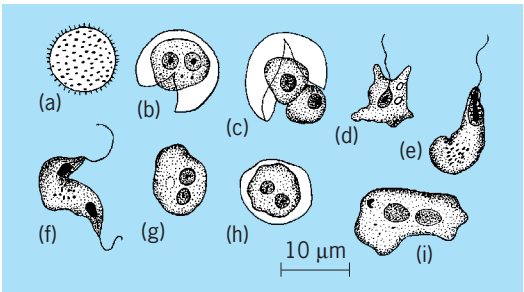

Fig. 2. Stages in the cycle of *Physarum polycephalum*. (*a*) Spore; (*b*, *c*) hatching of spore; (*d*, *e*) myxoflagellates; (*f–h*) zygote before nuclear fusion; (*i*) excysted zygote after first nuclear division. (*After R. P. Hall*, *Protozoology*, *Prentice-Hall*, *1953*)

of the plasmodium. The cytoplasm often contains pigment (yellow, orange, green, or other colors).

Resistant cysts (spores), possibly haploid, are produced in sporangia (**Fig. 1**). Liberated spores (**Fig. 2**) hatch into myxoflagellates or myxamebae (which may change into flagellates). Either type may represent gametes, and mating types have been reported in certain species. In some species a large resistant stage may be formed by dehydration of the plasmodium and secretion of a membrane (sclerotium) around part or all of the plasmodium. Within the membrane, multinucleate cysts are produced. Such a mature sclerotium may remain viable for several years, and can be reactivated by rehydration. *See* PROTOZOA; RHIZOPODEA; SARCODINA; SARCOMASTIGOPHORA. Richard P. Hall

Eumycota

True fungi, a group of heterotrophic organisms with absorptive nutrition, capable of utilizing insoluble food from outside the cell by secretion of digestive enzymes and absorption. Glycogen is the primary storage product of fungi. Most fungi have a well-defined cell wall that is composed of chitin and glucans. Spindle pole bodies, rather than centrioles, are associated with the nuclear envelope during cell division in most species. Typically, the fungal body (thallus) is haploid and consists of microscopic, branched, threadlike hyphae (collectively called the mycelium), which develop into radiating, macroscopic colonies within a substrate or host. The filamentous hypha may be divided by cross walls (septa) into compartments. Hyphal growth is apical. Some species are coenocytic (without cross walls); others, including yeasts, are unicellular. Reproductive bodies are highly variable in morphology and size, and may be asexual or sexual.

Great changes in understanding the phylogenetic relationships of fungi have been brought about by the use of characters derived from deoxyribonucleic acid (DNA) sequences and the use of computer assisted phylogenetic analysis; these changes are reflected in current classification schemes. The Oomycota and Hyphochytriomycota, characterized by heterokont flagella and cellulose-containing cell walls, have been shown to occur in a clade (lineage) that includes organisms with heterokont flagella and chlorophylls *a* and *c*, including diatoms and brown algae. Myxomycota and other groups of slime molds also are excluded from the Eumycota. The Chytridiomycota has been recognized as the earliest diverging group of fungi, and many members are distinguished from all other fungi by their possession of a posterior whiplash flagellum, a character that signals the close relationship of fungi and animals. About 70,000 species of fungi have been described; however, some estimates suggest that there may be 1.5 million species. A modern classification of Eumycota follows:

Eumycota
- Phylum: Chytridiomycota
- Phylum: Zygomycota
  - Class: Zygomycetes
    - Trichomycetes
- Phylum: Ascomycota
  - Class: Archiascomycetes
    - Hemiascomycetes
    - Euascomycetes
- Phylum: Basidiomycota
  - Class: Hymenomycetes
    - Urediniomycetes
    - Ustilaginomycetes

Of cosmopolitan distribution, fungi are found in practically every type of habitat (soil, air, fresh or brackish water, ocean, and so forth). Most are strictly aerobic, although a few are anaerobes that live in the gut of herbivores. Some species are thermophilic. Many fungi form saprobic (including parasitic) relationships with animals and plants; the majority are saprobes. As now recognized, the Eumycota are a monophyletic group of the crown eukaryotes, presumed to have been present in the fossil record 900–570 million years ago. *See* FUNGI; MYXOMYCOTA; OOMYCOTA. Meredith Blackwell

Bibliography. D. J. S. Barr, Evolution and kingdoms of organisms from the perspective of a mycologist, *Mycologia*, 84:1–11, 1992; D. Bhattacharya et al., Algae containing chlorophylls $a + c$ are paraphyletic: Molecular evolutionary analysis of the Chromophyta, *Evolution*, 46:801–1817, 1992; N. J. Dix and J. W. Webster, *Fungal Ecology*, Chapman and Hall, London, 1995; D. Griffin, *Fungal Physiology*, 2d ed.,

Wiley-Liss, New York, 1993; D. D. Leipe et al., The straminopiles from a molecular perspective: 16S-like rRNA sequences from *Labyrinthula minuta* and *Cafeteria roenbergensis*, *Phycologia*, 33:369–377, 1994.

Euphausiacea

A well-defined marine order of the class Crustacea. These planktonic malacostracans are closely allied with, and sometimes included in, the Decapoda. The order contains two families, one with a single species, the other with about 84 species divided among 10 genera. Euphausiids are predominantly pelagic organisms of the open ocean, but a few species are neritic. All are shrimplike in appearance (see **illus.**).

Euphausiids form a significant proportion of the total planktonic biomass. This is especially true in higher latitudes and most pronounced in the Antarctic Ocean, where *Euphausia superba*, which attains a body length of some 2.4 in. (60 mm), occurs in vast numbers.

Distribution and ecology. All species are strictly marine and do not occur in fresh-water or brackish environments. Species are found in all oceans and seas, except the brackish Baltic and Black seas.

Known as krill to whalers, they constitute the diet of many whales, particularly the baleen whales. The main whale feeding grounds coincide with the areas of greatest concentration of euphausiids: areas of convergence, backwaters, vortices of mixed layers, centers of gyres, and fronts. They contribute to the diet of many other animals such as seals, herring, sardines, many birds, and even humans. Their concentration in areas of high productivity is correlated with their own diet, which consists mainly of phytoplankton, small crustaceans, and detrital matter. Their feeding appendages, which consist of not only the mouthparts but also the anterior pairs of thoracic legs, are elaborate and can be used in several ways; this enables the animal to filter food particles from the water, collect organic material from the surface of the sediment, or catch live prey.

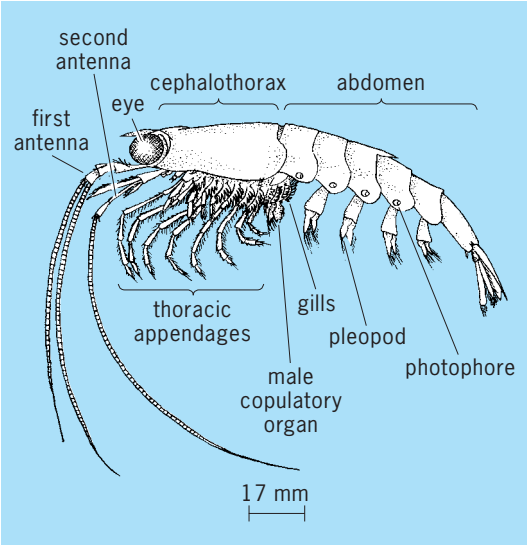

Diagram of a euphausiid crustacean.

Most euphausiids live at considerable depth during daylight hours, and many undertake extensive diurnal vertical migrations to the surface layers at night. Their vertical position in the water column is thought to be photoregulated; during the day most species live at depth under conditions of blue-green light of low intensity. A few species live in the bathypelagic environment, having been captured at depths greater than 6000 ft (2000 m); these species have small eyes and attain considerable size, the maximum recorded body length of one species being 6 in. (150 mm).

Morphology. Euphausiids, with the exception of *Bentheuphausia amblyops*, possess photophores that emit a brilliant blue-green light.

The eyes are compound, bilobed in some genera, and contain three pigments: a carotenoid (astaxanthin), a melanoid, and a photolabile substance that is probably a visual pigment. The spectral sensitivity of the eye is greatest to blue-green light. *See* PHOTORECEPTION.

Respiration is by means of foliose, digitiform gills located at the bases of the second to eighth thoracic appendages.

The blood is a pale, leukocyte-bearing fluid with hemocyanin as the respiratory pigment. The heart is compact and has two pairs of ostia. *See* RESPIRATORY PIGMENTS (INVERTEBRATE).

Reproduction. The male copulatory organs are extremely complex and form the main criterion for specific identification. The phenomenon of swarming is common in many genera and is usually associated with reproduction. The male transfers a spermatophore to the spermatheca of the female at copulation. The early embryos usually live freely in the sea, but those of some 25 species are carried attached to the ventral regions of the thorax.

Life cycle. The larval stages are numerous. A nauplius hatches from the egg membranes and develops to a metanauplius. Three calyptopis stages follow during which the abdomen and stalked eyes develop. All species pass through these stages. The following developmental sequences, some 6–12 molts as furciliae, vary between species and within species in different localities and at different times. Sexual maturity is normally achieved at an age of 1 year, although some species such as *E. superba* and those living in the meso- and bathypelagic environments require 2 or more years. One species, *Thysanoessa longicaudata*, produces two generations each year in warmer parts of the North Atlantic.

Commercial exploitation. The possibilities of exploiting the Antarctic *E. superba* and other species in other regions have been investigated in some detail. Estimates of the possible yield of a commercial fishery for *E. superba* range 27–180 $\times$ 10^6 ton (30–200 $\times$ 10^6 metric tons) per year. This species forms

large surface and subsurface aggregations that can be located visually or by echosounder. Ring and purse nets operated from small catchers or stern trawls operated from larger ships can be used to capture the swarms.

The euphausiids can be processed to yield large quantities of protein. Krill pastes and meal can be manufactured to feed domestic animals or used in therapeutic diets. Krill sausages, krill-stuffed eggs, and shrimp butter are already marketable products in Japan, Russia, and Germany.

In addition to *Euphausia superba*, some 30 species of euphausiids form surface or subsurface swarms or aggregations and are also therefore potentially exploitable resources. *Euphausia pacifica* is marketed for human consumption and also used, as is *Meganyctiphanes norvegica* in the North Atlantic, as food for fish in mariculture systems. *See* CRUSTACEA; DECAPODA (CRUSTACEA). John Mauchline

Bibliography. G. C. Eddie, *The Harvesting of Krill*, Food and Agricultural Organization, UN Development Program, Rome, Rep. GLO/SO/77/2, 1977; G. J. Grantham, *The Utilization of Krill*, FAO, UN Development Program, Rome, Rep. GLO/SO/77/3, 1977; J. W. Martin and G. E. Davis, *An Updated Classification of the Recent Crustacea*, Natural History Museum of Los Angeles, Science Ser. 39, 2001; J. Mauchline and L. R. Fisher, The biology of euphausiids, in *Advances in Marine Biology*, vol. 7, 1969; P. M. Mikkelsen, The Euphausiacea of eastern Florida (Crustacea: Malacostraca), *Proceedings of the Biological Society of Washington*, 100:275–295, 1987.

Euphorbiales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Rosidae of the class Magnoliopsida (dicotyledons). The order consists of the large family Euphorbiaceae (about 7500 species) and 3 small satellite families which have fewer than 100 species among them. The Euphorbiales are a group of few-ovulate, mostly simple-leaved Rosidae in which the flowers have become unisexual and then undergone further reduction accompanied by aggregation. The trend toward reduction and aggregation culminates in the very large genus *Euphorbia*, with perhaps 1500 species, in which the pseudanthium has a pistillate flower consisting of a naked, tricarpellate pistil surrounded by several staminate flowers, each of which consists of a single stamen. The illusion that this cluster constitutes a single flower is fostered also by the cup-shaped involucre which often bears showy, petallike appendages. The Christmas poinsettia (*E. pulcherrima*) and the para rubber tree (*Hevea brasiliensis*) are well-known members of the Euphorbiaceae. Many African euphorbiads are spiny stem-succulents which resemble cacti in habit. Aside from the pronounced floral differences, the euphorbiads have a milky juice, which the cacti do not. *See* MAGNOLIOPSIDA; PLANT KINGDOM; ROSIDAE. Arthur Cronquist; T. M. Barkley

Europe

Although long called a continent, in many physical ways Europe is but a great western peninsula of the Eurasian landmass. Its eastern limits are arbitrary and are conventionally drawn along the water divide of the Ural Mountains, the Ural River, the Caspian Sea, and the Caucasus watershed to the Black Sea. On all other sides Europe is surrounded by salt water. The stretches of water between the Mediterranean and Black seas, however, are narrow, and it is possible to swim across the Bosporus to Asia Minor. Not even the Mediterranean Sea is a real physiographic boundary inasmuch as the Atlas Mountains of northwestern Africa are structurally part of the great Tethys geosyncline. Of the oceanic islands of Franz Josef Land, Spitsbergen (Svalbard), Iceland, and the Azores, only Iceland is regarded, mainly on historical grounds, as an integral part of Europe; thus the northwestern boundary is drawn along the Danish Strait.

Maritime patterns. Europe has a large ratio of shoreline to land area reflecting a notable interfingering of land and sea. Excluding Iceland, the maximum north-south distance is 3529 mi (5680 km); and the greatest east-west extent is 3860 km. Of Europe's area of 3,880,000 mi^2 or 10,050,000 km^2 (that of the United States with Alaska and Hawaii is 3,620,000 mi^2 or 9,375,000 km^2), 73% is mainland, 19% peninsulas, and 8% islands. Altogether, Europe has a shoreline length of 23,600 mi (37,900 km), that is, 0.205 mi/1000 ft^2 (3.55 km/1000 m^2). Also, 51% of the land is less than 150 mi (250 km) from shores and another 23% lies closer than 300 mi (500 km). This situation is caused by the inland seas that enter, like arms of the ocean, deep into the northern and southern regions of Europe, which thus becomes a peninsula of peninsulas. The most notable of these branching arms of salt water are the White Sea, the North Sea, the Baltic Sea with the Gulf of Bothnia, the English Channel (La Manche), the Mediterranean Sea with its secondary branches—the Tyrrhenian, Adriatic, and Aegean Seas—and finally, the Black Sea. Even the Caspian Sea, presently the largest salt-water lake of the world, formed part of the southern seas before the folding of the Caucasus. Of these inland seas, only the Mediterranean and Black seas have depths greater than 400 fathoms (730 m). The penetration of the landmass by these seas brings marine influences deep into the continent and provides Europe with a balanced climate favorable for the evolution and settlement of humans. The favorable conditions are enhanced by the North Atlantic Drift (of Gulf Stream waters), which causes the northwestern coastal areas of Europe to be much warmer than the latitudes would suggest; for example, Labrador is much colder than the British Isles.

Land patterns. Europe has a unique diversity of landforms and natural resources. The relief, as varied as that of other continents, has an average elevation of 900 ft (300 m) as compared with North America's 1320 (440). However, 51% of the land lies below the 600-ft (200-m) contour lines, 27% is situated between 600 and 1500 ft (200 and 500 m), only

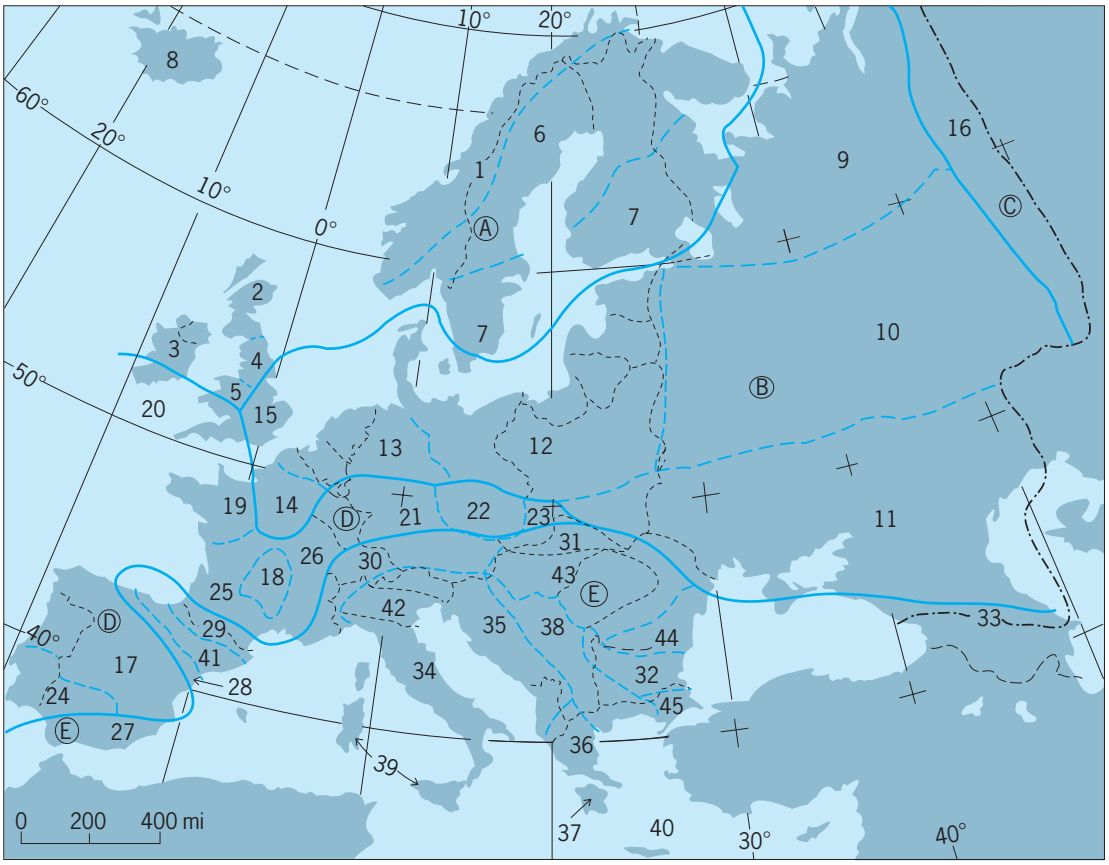

Fig. 1. Physiographic regions of Europe showing major units and subdivisions. A, Northern Uplands: 1, Kjöllen Mountains; 2, Scottish Uplands; 3, Irish Uplands; 4, Pennines; 5, Wales; 6, Scandinavian Hill Land; 7, Southern Lake Region; 8, Iceland and Arctic Islands. B, Central Lowlands: 9, Arctic Lowlands; 10, Great Russian Plain; 11, Ukrainian Steppes; 12, Baltic Lowlands; 13, North Sea Plains; 14, Paris Basin; 15, Anglican Plains. C, Eastern Uplands: 16, Ural Mountains. D, Western Plateaus, Mountains, and Basins: 17, Iberian Meseta; 18, Massif Central; 19, Armorican Uplands; 20, Cornwall and South Ireland; 21, Ardennes, Shale Mountains, Vosges, and Black Forest; 22, Bohemian Massif; 23, Lysa-Gora; 24, Portuguese Lowlands and Andalusian Basin; 25, Aquitanian Basin; 26, Rhone Depression. E, Southern Highlands and Basins; 27, Baetic Cordillera; 28, Cantabrian Cordillera; 29, Pyrenees; 30, Alps; 31, Carpathians; 32, Balkan Ranges; 33, Caucasus and South Crimea; 34, Apennines; 35, Dinarides; 36, Pindos Mountains; 37, Peloponnesus; 38, Rhodope; 39, West Mediterranean Islands; 40, Aegean Islands; 41, Aragonian Basin; 42, Po Basin; 43, Carpathian Basin; 44, Wallachian Plain; 45, Thracian Basin. (1 mi = 1.6 km.)

5% is higher than 3000 (1000), and only 1% exceeds 6000-ft (2000-m) above mean sea level.

The shape and the overall physiographic aspect of the great peninsula are controlled by geologic structure that delimits the major regional units (**Fig. 1**). The triangular Central Lowlands (B on the map) with its acute apex in the Anglican Plains (B-15) and with younger strata covering the Precambrian Russian Table in the eastern parts (as in B-9, 10, 11) are bounded on each of the three sides by systems of mountain ranges with adjacent upland areas. On the northwestern side the erosional remnants of the once higher Caledonian ranges, folded during the early Paleozoic, are broken up into the Irish Uplands, Scottish Highlands and Uplands, and the Kjöllen Mountains (A-3, 2, 1) bordering the predominantly granitic Baltic Shield of Precambrian age. In the east, the late Paleozoic Variscan mountain revolution folded up the Ural Mountains (C-16) and welded them to the Russian Table. The southern side of the triangle is more complex. Here the Paleo-Mesozoic Tethys geosyncline was folded twice into mountain chains; the Variscan revolution produced the Hercynian Mountains and set the stage for the great Tertiary Alpine mountain revolution which elevated the chains of the Sierra Nevada, Pyrenees, Alps, Carpathians, Dinarides, Pindos, Balkan Ranges, and the Caucasus (E-27, 29, 30, 31, 35, 36, 32, and 33). These mountains form the Southern Highlands, whereas the eroded and block-faulted remnants of the Hercynian Mountains, from Iberia to the Lysa-Gora, complete the highly complex regional unit of Western Plateaus, Mountains, and Basins (D).

Climate. This aspect of Europe is determined by a number of factors. Probably the most important are a favorable location between 35° and 71°N latitudes on the western or more maritime side of the world's largest continental mass; the west-to-east trend (rather than north-south) of the lofty southern ranges and the Central Lowlands, as well as of the inland seas, which permit the prevailing westerly winds of these latitudes to carry marine influences deep into the continent; the beneficial influence of the North Atlantic Drift, which makes possible ice-free coasts far within the Arctic Circle; and the low elevation of the northwestern mountain ranges and the Urals,

which allows the free shifting of air masses over their crests.

Pressure, air mass, and moisture trends. The high atmospheric pressures built up in winter over central Asia obstruct the eastward passage of cyclonic storms, but the summer low of Asia, occurring simultaneously with the weakening of the Icelandic low, permits the prevailing westerlies to carry precipitation as far as the Urals and even beyond. Southeastward escape of these rain-bearing summer storms is barred by the Alps and Carpathians. Also, the seasonal shifts of air masses from and to the Icelandic and the subtropical Azores pressure centers produce precipitation maxima north of a line along the Pyrenees-Alps-Carpathian-Balkan Ranges and Caucasus Mountains during the warm months of the year, whereas south of this line the summers are more or less rainless. These main climatic trends are broken up into a mosaic of local climates, depending on the closeness to shore lines, local elevation, and slope direction, and influenced by human activities such as urbanization, industrialization, and agriculture.

Maritime climates. Among the marine climatic types, the Atlantic, the Mediterranean, and even the boreal polar climates are characteristic. True oceanic Atlantic climate occurs along the western fringes of Europe, from northern Portugal to central Norway and in the British Isles and Iceland. This west-coast marine climate is characterized by mild winters, with rarely a freezing period, and by somewhat cloudy, cool summers. The mean monthly range of temperature seldom exceeds 20°F (11°C). Precipitation, mostly in the form of drizzly rain, is regularly distributed over the year. Hence, runoff is small and soil erosion is at a minimum. Heather, wood anemone, small-leafed linden, English oak, and European beech are the representative plants of this climate. They are also prevalent in central Europe and along some of the shores of the western Mediterranean. In the Iberian and the Italian peninsulas, as well as in southern France, this climate is tempered by warm Atlantic influences reaching as far north as southern England and Denmark and eastward into Greece and the Aegean archipelago, with a plant community characterized by the ivy, English holly, box tree, and fig.

Mitigated by oceanic climate along northwest Norway's irregular shores, the tundra climate of the Barents Sea coast occupies but a narrow coastal strip. Average temperature of the warmest month is below 50°F (10°C) but above 32°F (0°C). Precipitation is scarce, soil is permanently moist, and true summers are absent. Tundra vegetation is only herbs, dwarf plants, and the characteristic tree of boreal regions, the dwarf birch.

On a relatively narrow strip along Europe's southern coasts the Mediterranean climate prevails. It influences to a certain extent some inland areas as far north as a line through the Pyrenees, Alps, Carpathian Basin, Crimea, and Caucasus. Mild winters with sunny weather and occasional rains change, through a delightful temperate spring, into dry and rainless summers with high temperatures and bright blue skies. Some of the most characteristic plants of this climate are associated in the brushy macchia formation; others are mostly cultivated plants, such as the vine, maize (corn), chestnut, olive, and orange.

Continental climates. The counterparts to the marine climates are continental climates extending into Europe, across the Urals, from the great Asiatic landmass with its climatic belts running parallel to northern latitudes. Easterly air masses can move easily across the low, broadbacked Ural Mountains, and their impact reaches westward over Europe's central lowlands, where it mingles, especially in the summer, with the oceanic climate pushing eastward on the wings of the westerly winds. The northernmost zone of this continental influence has a subpolar continental climate, the southern boundary of which runs close to 60°N latitude. Brief summers are marked by temperatures above 50°F (10°C). The relatively meager precipitation covers the ground for almost one-half of the year in the form of snow, and late spring floods often occur as a result of the blocking effect of frozen estuaries of rivers flowing northward into Arctic waters. The vegetation consists mainly of coniferous forests known as taiga. Norway spruce and Scotch pine are the main species, intermixed northward with the characteristic boreal tree, the dwarf birch.

South of this zone, in the Great Russian Plains, the long cold-winter climate predominates. During 3 months of the winter period, temperatures drop below the freezing point and the snow is blown along land surfaces by icy continental winds. Temperatures averaging over 50°F (10°C) and the maximum rainfalls occur in the June-to-August growing season, but precipitation rarely exceeds 20 in. (50 cm) per year. Scotch pine on sandy soil and deciduous forests prevail in the northern parts of this zone, whereas oak forests alternate with the fertile cropland on chernozem soils to the south.

The southernmost continental climatic zone, covering roughly the Ukrainian Steppes, which are a western extension of the mid-latitude steppe and desert regions of Asia, is characterized by the semiarid steppe climate. With less than 5 in. (12.5 cm) of rainfall per year in the area of the lower reach of the Volga River, these eastern parts are a desert. Westward the 15–20 in. (37.5–50 cm) of yearly precipitation, with the maxima during the growing season, and mean annual temperatures just below 50°F (10°C) predominate in the natural grasslands. The most characteristic plant of these regions is the Stipa grass, which also appears to the westward in the Wallachian Plain and in the central portion of the Carpathian Basin.

Patterns of local climates. Among the more pronounced local climates of Europe, several are of importance (**Fig. 2**). An outstanding example is the semiarid climate of the Iberian Peninsula, partly due to the plateau character of the meseta upland and partly to warm Atlantic climatic influences. The peninsula has much scrubby vegetation and rocky soil covered with low grass.

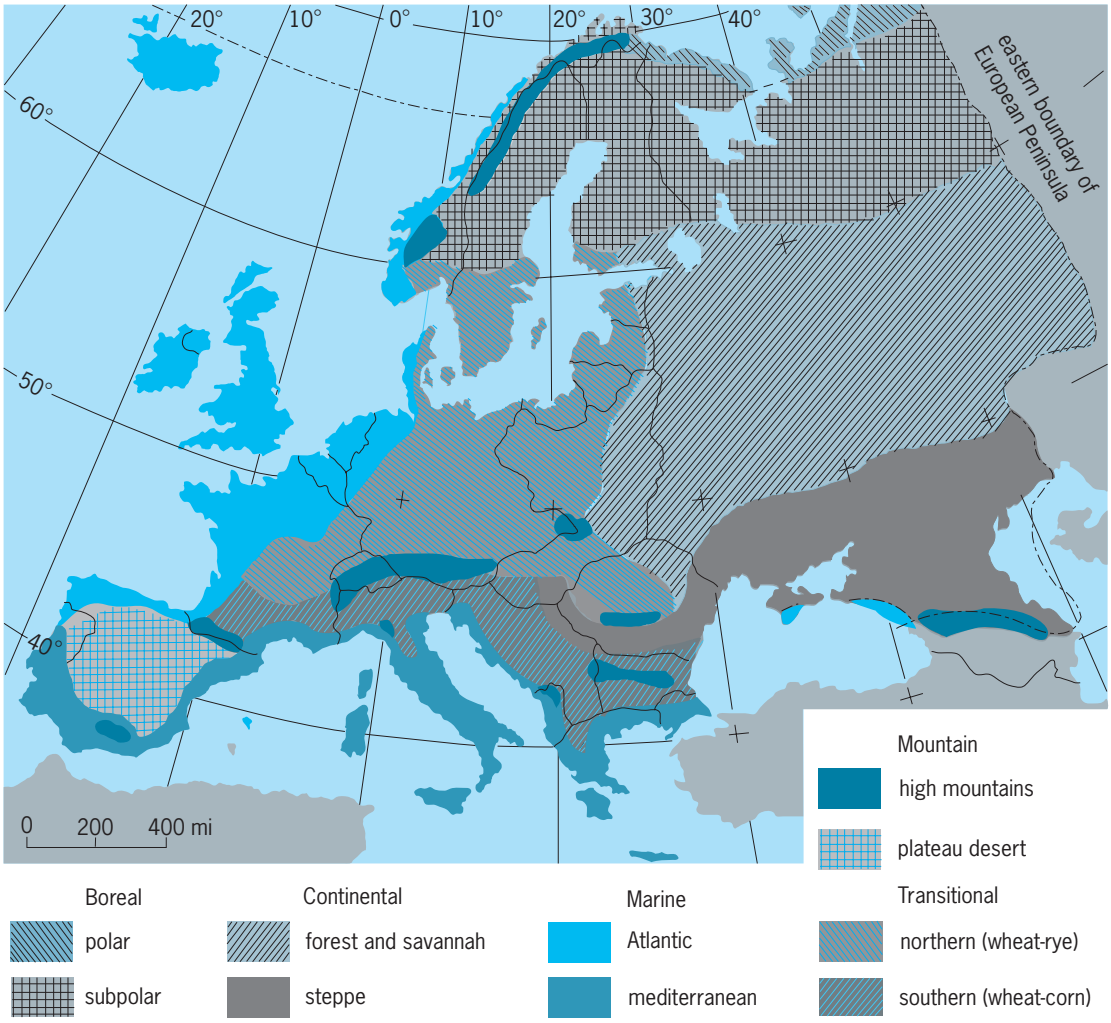

Fig. 2. Map of climatic regions of Europe. Regions and subdivisions are based on weather and vegetation data. 1 mi = 1.6 km.

Mountain climate is more a group of local climates than a single climatic type. Depending upon the elevation above sea level, the latitude in which they occur, the proximity to seas or landmasses, and the general slope directions, these climates show extreme variability. With glacial influences in the Alps, cordilleran character in the Pyrenees and in the southern Carpathians, subtropical influences in the western Caucasus, and boreal imprint in the Kjöllen Mountains, they do not have much in common. Only the occurrence of a fir line and a snow line, due to temperature decrease in the lofty sections of these mountain ranges, gives them a certain uniformity. Local winds in such areas have special names, such as the föhn of the Alps and the nemere of the eastern Carpathians. These climatic types intermix and compete in a nearly triangular area. This is comparable to Europe's structurally defined shape, with the apexes in the southern part of the Gulf of Bothnia, the western Pyrenees, and the western portion of the Black Sea. The center of this area is approximately near Wien, Austria. The northern transition zone is more tempered, the southern zone shows extremes of temperature, and toward the east both transitional climatic regions are apt to have variable extremes of precipitation which quite often cause flood and drought years.

Drainage patterns. The intricate relief and the climates of Europe are well reflected in the drainage system. Extensive drainage basins with large slow-flowing rivers are developed only in the Central Lowlands, especially in the eastern part. Streams with the greatest discharge empty into the Black Sea and the North Sea, although Europe's longest river, the Volga, feeds the Caspian Sea. Second in dimension is the Danube, which crosses the Carpathian Basin and cuts its way twice through mountain ranges at the Gate of Bratislava and at the Iron Gate. The Rhine and Rhone are the two major Alpine rivers with headwater sources close to each other but feeding the North Sea and the Western Mediterranean Basin, respectively. Abundant precipitation throughout the year, as well as the permeable soils and the dense vegetation which temporarily store the water, provides the streams of Europe north of the Southern Highlands with ample water throughout the seasons. The combined effects of poor vegetation, rocky and desolate limestone karstlands, and slight annual precipitation result in intermittent flow of the rivers along the Mediterranean coast, especially on the eastern side

of peninsulas. Only the Alpine rivers carry enough water, and if it were not for the Danube and Rhone, both originating in regions north of the Alps, the only major river of the Mediterranean basin would be the Po. The main water divide between northern oceanic and southern Mediterranean drainage runs in a general east-northeast direction from Gibraltar along the Sierra Nevada, Cantabrian Mountains, Pyrenees, Massif Central, Western Alps, Black Forest, Czech Forest, Northern Carpathians, Pripyat Marshes, and Valdar Hills to the northern central Urals. The Urals form the water divide between Europe and Asia.

Summary of physical patterns. Geologic structure, relief and drainage, climate, soils and vegetation, and human activities have shaped the physiographic aspects of Europe (**Tables 1–4**). Every regional

TABLE 1. Lakes

Name	Area, mi^2 (km^2)	Countries, and area, mi^2 (km^2)	Elevation, ft (m)	Approx. depth, ft (m)	Important tributaries	Important cities
Balaton	260 (660)	Hungary	348 (106)	36 (11)	Zala (outlet, Sio)	Siofok
Berre, Étangde	60 (156)	France	43 (13)	36 (11)		
Biesbosch, Lake	80 (200)	Netherlands	0	?		
Bodensee (Lake of Constance)	208 (539)	Germany, 127(328)	1,296 (395)	827 (252)	Rhine*	Konstanz, Bregenz
		Switzerland, 82 (211)				
Byelosero, Lake	434 (1,125)	CIS	400 (122)	33 (10)		
Caspian Sea	168,000 (436,000)	CIS, Iran	-85 (-26)	3,100 (946)	Volga, Ural	Astrakhan, Baku
Chiemsee	30 (80)	Germany	1,700 (519)	240 (74)	Alz*	
Como, Lake	56 (145)	Italy	650 (198)	1,350 (410)	Adda*	Como
Corrib, Lake	73 (190)	Rep. of Ireland	26 (8)	144 (44)		
Femunnen, Lake	79 (205)	Norway	2,990 (910)	?		
Fertö, Neusiedler, Lake	120 (320)	Austria, 90 (232)	371 (113)	13 (4)		
		Hungary, 34 (88)				
Garda, Lake of	140 (370)	Italy	210 (65)	1,140 (346)	(outlet, Mincio)	
Geneva (Leman) Lake	225 (582)	France, 90 (234)	1,230 (375)	1,020 (310)	Rhone*	Geneva, Lausanne
		Switzerland, 134 (348)				
Ilmen Lake	354 (918)	CIS	59 (18)	10 (3)		
Inari, Lake	535 (1,385)	Finland	387 (118)	260 (80)	Ivalo	
Jalpug, Lake	85 (220)	CIS		?	Jalpug*	
Jannitza, Lake	40 (100)	Greece	1,345 (410)	72 (22)		
Ladoga, Lake	7,019 (18,180)	CIS 3,925 (10,166)	16 (5)	820 (250)	Svirj* (from Lake Onega)	
		Finland 3,094 (8,014)				
Lugano, Lake of	19 (49)	Italy, 7 (18)	899 (274)	915 (279)		Lugano
		Switzerland, 12 (31)				
Luzern, Lake of		Switzerland	1,433 (437)			Luzerne
Lomond, Loch	27 (71)	Scotland		?		
Maggiore, Lake	82 (212)	Italy, 73 (190)	636 (194)	367 (372)	Ticino*	Locarno
		Switzerland, 9 (22)				
Malaren, Lake	440 (1,140)	Sweden	3 (1)	210 (64)		
Mjösa, Lake	139 (359)	Norway	397 (121)	1,450 (443)		
Neagh, Lough	153 (396)	North Ireland	49 (15)	100 (31)		
Neuchâtel Lake of	83 (216)	Switzerland	1,400 (427)	510 (154)		Neuchatel
Ohrida, Lake	142 (367)	Albania, 46 (119) Macedonia	2,290 (698)	938 (286)	Drin*	
Onega, Lake	3,687 (9,550)	CIS	128 (39)	407 (124)		Petrozavodsk
Oulöjarvi, Lake	387 (1,002)	Finland	404 (123)	?		
Päijänne, Lake	503 (1,304)	Finland	26 (78)	305 (93)		
Peipus, Lake	1,383 (3,583)	CIS	28 (31)	59 (18)	Narva*	
Prespa, Lake	110 (286)	Greece, Albania, Macedonia	2,970 (906)	180 (54)		
Ree, Lake	64 (165)	Rep. of Ireland	130 (39)	120 (36)		
Saimaa, Lake	460 (1,200)	Finland	250 (76)	190 (58)		
Sasik, Lake	75 (195)	CIS	0	?		
Scutari (Shkoder), Lake	140 (370)	Albania, 57 (148) Montenegro	18 (6)	145 (44)	Zeta (outlet, Bojana)	Shkoder (Scutari)
Segosevo, Lake	460 (1,200)	CIS	358 (109)	325 (99)		
Thingullatn, Lake	33 (85)	Iceland	2,050 (625)	?		
Trasimene, Lake	50 (129)	Italy	846 (258)	23 (7)		
Vänern, Lake	2,140 (5,550)	Sweden	145 (44)	320 (98)	Klar (outlet, Gotäelv	Karlstad
Vättern, Lake	730 (1,900)	Sweden	290 (88)	390 (119)		Jönköping
Vierwaldstätter, Lake	44 (114)	Switzerland	1,430 (437)	702 (214)		
Zuider Zee (artificial)	?	Netherlands	0	?		Amsterdam
Zürich, Lake of	34 (89)	Switzerland	1,340 (409)	469 (143)		Zürich

*Tributaries which cross the lake.

TABLE 2. Selected peaks higher than 2800 m (9240 ft)

Name	Elevation, ft (m)	Location by mountain ranges, divisions and subdivisions, or island
Adamello	11,660 (3,554)	Prealps, Lombardian Alps, Alps of Chamonix
Adula	11,180 (3,406)	Western Alps, Lepontine Alps, Adula group
Aneto, Pico d'	11,170 (3,404)	Central Pyrenees, Maledetta group
Antelao	10,710 (3,263)	Eastern Alps, Dolomites, Antelao group
Argentea	10,820 (3,299)	Western Alps, Maritime Alps, Ligurian Alps
Bernina, Piz	16,290 (4,052)	Eastern Alps, Rhaetian Alps, Bernina group
Birkkar	9,700 (2,956)	Prealps, Bavarian Alps, Karwendel Mts.
Blanc, Mont	15,770 (4,807)	Western Alps, Pennine Alps, Mont Blanc group
Casse, Grand	10,110 (3,081)	Western Alps, Graian Alps, Massif de Vanoise
Coll' Alto	11,270 (3,435)	Eastern Alps, Noric Alps, Alps of Pusterthal
Corno, Monte	9,583 (2,921)	Appennines, Abruzzian Apennines, Gran Sasso
Cristallo, Monte	10,550 (3,216)	Eastern Alps, Dolomites, Tofano group
Dachstein	9,829 (2,996)	Eastern Alps, Salzkammergut
Dufour	15,300 (4,663)	Western Alps, Pennine Alps, Monte Rosa group
Écrins, Barre des	13,460 (4,103)	Prealps, Dauphinean Alps Pelvoux group
Elbrus	18,570 (5,660)	Western Caucasus
Etna*	10,760 (3,279)	Island of Sicily
Finsteraarhorn	13,970 (4,257)	Western Alps, Bernese Alps
Fluchthorn	11,165 (3,403)	Eastern Alps, Rhaetian Alps, Silvretta group
Gran Paradiso	13,320 (4,061)	Western Alps, Graian Alps
Grand Casse	10,110 (3,081)	Western Alps, Graian Alps, Massif de Vanoise
Grossglockner	12,460 (3,798)	Eastern Alps, Noric Alps Hohe (High) Tauern
Grossvenediger	12,010 (3,660)	Eastern Alps, Hohe Tauern
Hochfeiler	11,520 (3,510)	Eastern Alps, Noric Alps, Alps of Zillerthal
Hochgolling	9,393 (2,863)	Eastern Alps, Noric Alps Niedere (Low) Tauern
Hochkönig	9,639 (2,938)	Prealps, Alps of Salzburg
Jel Tepe	9,580 (2,920)	Rhodope, Pirin Planina
Jungfrau	13,670 (4,166)	Western Alps, eastern Bernese Alps
Kazbek	16,550 (5,045)	Eastern Caucasus
Leone, Monte	11,650 (3,552)	Western Alps, Lepontine Alps
Marmolata	10,970 (3,342)	Eastern Alps, Dolomites, Marmolata group
Matterhorn	14,780 (4,505)	Western Alps, Pennine Alps
Midi, Pic du	9,465 (2,885)	Central Pyrenees
Monte Rosa	15,190 (4,630)	Western Alps, Pennine Alps, Monte Rosa group
Mulhacen	11,420 (0.481)	Sierra Nevada
Mus Alla	9,596 (2,925)	Rhodope, Rila Planina
Olympos	9,573 (2,918)	Eastern Greek Mts., Chazia
Ortler	12,790 (3,899)	Eastern Alps, Rhaetian Alps, Ortier group
Paradiso, Gran	13,323 (4,061)	Western Alps, Graian Alps
Parseier	9,967 (3,038)	Prealps, Bavarian Alps, Alps of Lechtal
Rosa, Monte	15,190 (4,630)	Western Alps, Pennine Alps, Monte Rosa group
Sandspitze	9,393 (2,863)	Eastern Alps, Carnic Alps, Gailthal group
Scesaplana	9,734 (2,967)	Eastern Alps, Rhaetian Alps, Rhaetikon
Tabor, Mount	10,420 (3,177)	Western Alps, Cottian Alps (northern)
Todi	10,710 (3,623)	Western Alps, Glarnian Alps
Tofana	10,640 (3,243)	Eastern Alps, Dolomites
Triglav	9,393 (2,863)	Eastern Alps, Julian Alps
Vezzana, Cimi di	10,470 (3,191)	Eastern Alps, Dolomites, Pale group
Viso, Monte	12,600 (3,841)	Western Alps, Cottian Alps (southern)
Wildhorn	10,710 (3,264)	Western Alps, Bernese Alps (western)
Wildspitze	12,380 (3,774)	Eastern Alps, Rhaetian Alps, Alps of Oetzthal
Zuckerhütl	10,030 (3,057)	Eastern Alps, Rhaetian Alps, Stubaier Alps
Zugspitze	9,724 (2,864)	Prealps, Bavarian Alps, Wetterstein Mts.

unit comprises regions with traits of similarity; but similarity does not mean uniformity. For example, the Alps have been sculptured by glaciers of the great Ice Age and their valleys are broad and long, whereas the Pyrenees and Caucasus, which bore few glaciers, have narrow, steep-sloped, V-shaped valleys, which give them a rougher aspect than that of the much higher Alps. Each region has its own characteristics of local relief, vegetation, animal life, and mineral resources. Some of the minerals occur in Precambrian and Paleozoic rocks of the ancient shields, blocks, and mountain chains, for example, the iron ores of Scandinavia, England, Alsace-Lorraine, Spain, and Bohemia, or the coalfields of England, the Ruhr, Silesia, and the Donets Basin. Oilfields are mainly concentrated in the Carpathian region in Romania, Hungary, Poland, and Austria; a great variety of ores is bound to the metamorphic rocks and Tertiary volcanic areas affected by the Variscan and Alpine mountain revolutions; bauxite and manganese occur in pockets of the deeply weathered Mesozoic limestone areas, mainly in France, Hungary, and former Yugoslavia; and precious metals are most abundant in the Urals

TABLE 3. Major rivers

Name	Length, mi (km)	Drainage area, mi^2 (km^2)	Discharges into	Main tributaries*	Source region and regions crossed
Bug (Polish)	484 (779)	28,370 (73,470)	Vistula	(R) Narew	Podolian Upland, Ukrainian and Baltic Lowlands
Bug (Ukrainian)	466 (750)	28,290 (73,280)	Black Sea		Podolian Upland, Ukrainian Steppes
Danube	1,800 (2,890)	312,192 (808,578)	Black Sea	(L) Vah, Tisza, Olt, Siret, Prut; (R) Inn, Drava, Sava, Morava	Black Forest, Bavarian Lowland, East Alpine Foreland, Carpathian Basin, Iron Gate, Wallachian Plain, Dobrudja, Moldavia boundary
Dnieper	1,340 (2,150)	204,000 (527,000)	Black Sea	(R) Pripyat, Beresina; (L) Seim	Great Russian Plain, Ukrainian Steppes
Dniester	850 (1,370)	29,700 (76,900)	Black Sea		Eastern Carpathians, Bessarabia (Ukrainian Steppes)
Don	1,160 (1,860)	166,000 (429,800)	Sea of Azov	(R) Donets	Great Russian Plain, Ukrainian Steppes
Donets	670 (1,078)	?	Don		Ukrainian Steppes
Duero	485 (780)	37,983 (98,375)	Atlantic Ocean	(R) Esla	Iberic Mts. (Cantabrian Cordillera), Iberian Plateau
Dvina	1,110 (1,780)	139,900 (362,300)	White Sea	Upper reach called Vicegda	Timan Mts., Arctic Lowlands
Ebro	580 (930)	33,000 (86,000)	Mediterranean Sea	(L) Segre	Cantabrian Cordillera, Aragonian Basin
Elbe	717 (1,154)	57,044 (147,744)	North Sea	(L) Saale	Sudeten Mts., Bohemian Basin, North Sea Plains
Garonne	400 (650)	32,741 (84,800)	Atlantic Ocean	(R) Lot, Dordogne;	Pyrenees, Aquitanian Basin
Guadalquivir	360 (579)	22,050 (57,120)	Atlantic Ocean	(L) Genil	Baetic Cordillera, Andalusian Basin
Guadiana	510 (820)	26,193 (67,840)	Atlantic Ocean		Sierra Morena, Iberian Plateau, Portuguese Lowland
Inn	320 (510)	10,000 (25,700)	Danube	(R) Salzach	Bernese Alps, Engadin, Bavaria Alps and Lowland
Kama	1,260 (2,030)	?	Volga	(L) Bjelaga: (R) Vjatka	Permian Hill Land, Western Ural foothills, Russian Plain
Kemijoki, Kemi	795 (494)	20,000 (50,000)	Gulf of Bothnia	(R) Ounas	Maan Hills, Northern Hill Land
Kuban	512 (824)	?	Black Sea		Caucasus, North Caucasian foothills
Loire	623 (1,002)	46,520 (120,500)	Atlantic Ocean	(L) Vienne	Plateau Central, South Paris Basin, Armorican Upland
Mernel	582 (936)	37,900 (98,100)	Baltic Sea		Pripyat Marshes, Baltic Lowlands
Meuse	575 (925)	13,000 (33,000)	North Sea		Plateau of Langres, North Sea Plains
Mosel	339 (545)	10,900 (28,200)	Rhine		Vosges, Hunsrück
Oder	561 (903)	48,136 (124,671)	Baltic Sea	(L) Neisse; (R) Warta	Sudeten Mts., Baltic Lowland
Pechora	932 (1,500)	123,700 (320,300)	Barents	(L) Ishma	Northern Ural, Arctic Lowlands
Po	418 (672)	28,950 (74,970)	Adriatic Sea	(L) Adda, Ticino	Cottian Alps, Po Basin
Rhine	824 (1,326)	86,640 (224,400)	North Sea	(L) Mosel; (R) Neckar, Main	Lepontine Alps, Vosges and Black Forest boundary, North Sea Plains
Rhone	505 (812)	38,000 (99,000)	Gulf of Lions	(R) Saone; (L) Durance	Lepontine Alps, Lake Geneva, Rhone Depression
Sava	564 (907)	38,800 (100,519)	Danube	(L) Drina	Julian Alps, Carpathian Basin
Seine	483 (776)	30,000 (77,800)	English Channel (La Manche)	(R) Marne, Oise	Plateau of Lengres, Paris Basin
Severn	210 (338)	8,100 (21,000)	Bristol Channel		Cambrian Mts., Wales
Tejo (Tagus)	628 (1,010)	31,250 (80,950)	Atlantic Ocean		Iberian Meseta, Portuguese Lowland
Thames	209 (336)	5,920 (15,340)	North Sea	Cotswold Hills, Anglican Plains	
Tiber (Tevere)	252 (405)	6,629 (17,169)	Tyrrhenian Sea		Tuscan Apennines, Apennine foothills (East)
Tisza	597 (961)	60,670 (157,135)	Danube	(L) Szamos, Maros, Koros (two sources, Black and White Tisza)	North-Eastern Carpathians, Carpathian Basin
Ural	1,480 (2,380)	84,900 (219,000)	Caspian Sea	(L) Ilek	Southern Urai, Kirghiz Steppes, Caspian Depression
Varda	200 (322)	?	Aegean Sea	(R) Crna	Albanian Alps, Thessalonian Plain
Vistula	862 (1,387)	76,560 (198,290)	Gulf of Gdynia	(R) Bug	Beskides (Western Carpathians), Baltic Lowlands
Volga	2,200 (3,570)	548,000 (1,420,000)	Caspian Sea	(R) Oxa; (L) Kama	Valdai Plateau, Great Russian Plain, Ukrainian Steppes, Caspian Depression
Warta	470 (760)	21,000 (54,000)	Oder		Lysa Gora foothills, Baltic Lowlands

*(L) left bank; (R) right bank.

TABLE 4. Volcanoes

Name (locate)	Last erupted	Elev. ft (m)	Rank*
Askja (Iceland)	1926	2,290 (698)	6
Etna (Sicily)	1979	10,760 (3,279)	1
Hekla (Iceland)	1948	5,108 (1,557)	3
Katla (Iceland)	1918	5,300 (1,600)	2
Santorin (Cyclades, Greek Aegean)	1928	1,920 (584)	7
Stromboli (Lipari Islands)	1934	3,038 (926)	5
Vesuvius (Apennine Peninsula)	1944	3,904 (1,190)	4
Vulcano (Lipari Islands)	1926	1,600 (500)	8

*By altitude.

and the Bihar Mountains. *See* AFRICA; ASIA; ATLANTIC OCEAN; BALTIC SEA; BLACK SEA; CONTINENT; EAST INDIES; MEDITERRANEAN SEA. Geza Teleki

Bibliography. E. A. Koster (ed.), *The Physical Geography of Western Europe*, 2005; M. Shahgedanova (ed.), *The Physical Geography of Northern Eurasia*, 2003; A. N. Strahler and A. H. Strahler, *Modern Physical Geography*, 4th ed., 1991; T. Unwin (ed.), *A European Geography*, 1998.

Europium

A chemical element, Eu, atomic number 63, atomic weight 151.96, a member of the rare-earth group. The stable isotopes, ^{151}Eu and ^{153}Eu, make up the naturally occurring element. The metal is the second most volatile of the rare earths and has a considerable vapor pressure at its melting point. It is very soft, is rapidly attacked by air, and really belongs more to the calcium-strontium-barium series than to the rare-earth series. *See* PERIODIC TABLE.

1												13	14	15	16	17	18
1 H	2																2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

The element is attractive to the atomic industry, since the elements can be used in control rods and as nuclear poisons. These poisons are materials added to a nuclear reactor to balance the excess reactivity at start-up, and are so chosen that the poisons burn out at the same rate as the excess activity decreases. The television industry uses considerable quantities of phosphors, such as europium-activated yttrium orthovana-dates, and other europium-activated yttrium phosphors have been patented. These phosphors give a brilliant red color and are used in the manufacture of television screens. *See* RARE-EARTH ELEMENTS. Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner, Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

Eurypterida

An extinct Paleozoic group of aquatic arthropods, belonging to the subphylum Chelicerata and class Merostomata and thus related to the living marine xiphosurans (horseshoe crabs) and terrestrial arachnids (spiders, scorpions). Although most eurypterids were less than 10 in. (25 cm) in length, some members of the group were the largest arthropods of all time, reaching sizes up to 6 ft (2 m). *See* MEROSTOMATA.

The anterior portion of the eurypterid body is the prosoma. The prosoma is covered dorsally by the carapace, which bears a large pair of lateral compound eyes and a small pair of median simple eyes or ocelli (see **illus.**). Visible on the ventral surface of the prosoma are six pairs of appendages. The first pair, adjacent to the mouth, comprises small pincers known as the chelicerae. This feature unites the eurypterids with the xiphosurans and the arachnids. The chelicerae in some forms are enormous. The next four pairs of appendages were probably used in walking and food gathering. The distal portions of the sixth pair of appendages in many eurypterids are flattened and laterally expanded. They closely resemble the "swim paddles" of living blue crabs and were probably used to generate thrust and lift for

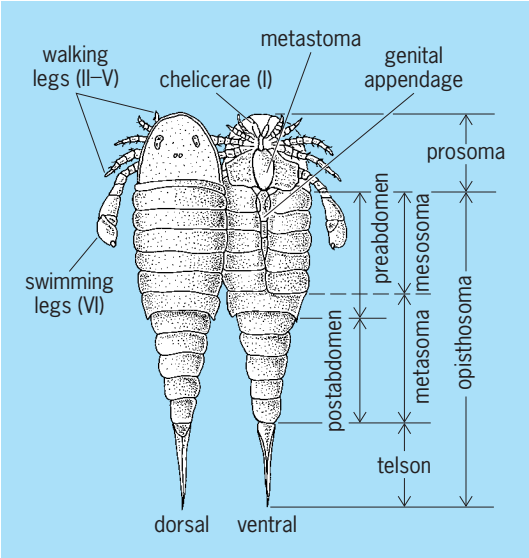

General morphology of eurypterids, in dorsal and ventral views.

swimming. The metastoma is a large plate which lies ventral to the bases of the sixth pair of appendages. It may represent a fused seventh pair of appendages and is unique to the eurypterids.

The posterior opisthosoma consists of 12 unfused segments and the terminal telson or tail spine. The opisthosoma can be divided into the anterior six-segmented mesosoma, which bears appendages, and the posterior six-segmented metasoma, which lacks appendages. The five pairs of platelike mesosomal appendages cover gills located on the ventral body wall. The first two mesosomal appendages carry the sexually dimorphic external genitalia. The opisthosoma can also be divided into a broad seven-segmented preabdomen and a narrower five-segmented postabdomen.

The telson of most eurypterids is styliform and closely resembles that of horseshoe crabs. In some forms the telson is distinctly curved, whereas in others it forms a broad flat plate, which may have functioned as a rudder during swimming.

Eurypterid fossils are found worldwide. The chitinous body of eurypterids is not mineralized, and so fossil remains are rare and generally restricted to a limited number of ancient aquatic environments. Almost all of the approximately 300 described species, belonging to about 65 genera, are from single localities or regions. The earliest definite eurypterids are found in the Ordovician, in both shallow and deep marine settings. The group diversified during the Silurian, peaked during the Late Silurian, and then declined during the Devonian. Silurian and Devonian eurypterids are found in both marine and nonmarine environments. Carboniferous (Mississippian Pennsylvanian), and Permian eurypterids are rare and are exclusively nonmarine (for example, coal swamps). The group became extinct sometime during the Permian. *See* DEVONIAN; MISSISSIPPIAN; ORDOVICIAN; PENNSYLVANIAN; PERMIAN; SILURIAN.

Eurypterids were almost exclusively aquatic, although some forms may have been amphibious. A number of forms were dominantly benthonic, but most were active and agile swimmers. Most eurypterids were probably carnivores, although they lacked the ability to crush heavily armored prey.

Although the eurypterids have no living descendants, they are very similar to (and are often found with) the earliest scorpions. The two groups probably share a common ancestor. *See* ARTHROPODA.

Roy Plotnick

Bibliography. R. Plotnick and T. Baumiller, The pterygotid telson as a biological rudder, *Lethaia*, 21:13–27, 1988; P. A. Selden, Autecology of Silurian eurypterids, *Spec. Pap. Palaeontol.*, 32:39–54, 1984; R. E. Snodgrass, *A Textbook of Arthropod Anatomy*, 1965.

Eustigmatophyceae

A small class of nonmotile, photosynthetic, unicellular algae, segregated from the class Xanthophyceae (Tribophyceae) on the basis of cytological, ultrastructural, and biochemical features of the vegetative cell and zoospore. Because of their lack of chlorophyll *b*, these organisms may be placed in the division Chromophycota, but their lack of chlorophyll *c*, in contrast to other chromophytes, supports recognition at a higher level (such as the division Eustigmatophyta). Only a dozen species in three genera are known. These live chiefly in fresh water, but also in marine habitats and in soil.

Vegetative cells usually are free-floating (attached in one species) and spherical, polyhedral, or elongate, with a maximum dimension of 5–60 micrometers. The cell wall is fibrillar-mucilaginous, of undetermined composition. Each cell contains one or more nuclei and usually a single chloroplast. The chloroplast is surrounded by three membranes, the outermost of which is continuous with endoplasmic reticulum that apparently is not confluent with the nuclear envelope. A distinctive pyrenoid—stalked and spherical or polyhedral—projects from the inner face of the chloroplast. Photosynthetic lamellae consist of three appressed thylakoids containing chlorophyll α, β-carotene, and violaxanthin. The presumed food storage product, of unknown composition, is deposited around the pyrenoid outside the chloroplast and appears lamellate in electron micrographs.

Sexual reproduction is unknown. Asexual reproduction is by the formation of two or four autospores or by zoospores. Zoospores are especially diagnostic of the class, although not known for all species. They are elongate, naked, and with two unequal flagella, only one of which is emergent in most species, the other being reduced to a basal body. The longer flagellum has two opposite rows of stiff hairs (tubular mastigonemes). Each zoospore has a single chloroplast without a pyrenoid. A large, reddish-orange eyespot not surrounded by membranes is situated in the cytoplasm near the base of the flagella. *See* ALGAE; XANTHOPHYCEAE.

Paul C. Silva; Richard L. Moe

Bibliography. D. J. Hibberd, Notes on the taxonomy and nomenclature of the algal classes Eustigmatophyceae and Tribophyceae (synonym Xanthophyceae), *Bot. J. Linn. Soc.*, 82:93–119, 1981; D. J. Hibberd and G. F. Leedale, Observations on the cytology and ultrastructure of the new algal class, Eustigmatophyceae, *Ann. Bot.*, 36:49–71, 1972.

Eutardigrada

An order of tardigrades, lacking a cirrus lateralis, a sensory cephalic appendage, and a clava, or club-shaped appendage. Pharyngeal pockets are strengthened by separated rods or macroplacoids or are without thickenings. Claws are of different size, arranged in two pairs in which a larger and smaller claw are united. In *Milnesium* the claws are separated. *Haplomacrobiotus* has two simple claws. Eight longitudinal muscles are associated with the midgut. At the beginning of the hindgut, there are three excretory glands, called the vasa malpighii, which are specifically stretched or trilobed. Each gland is composed of three cells. The gonoducts open into the rectum,

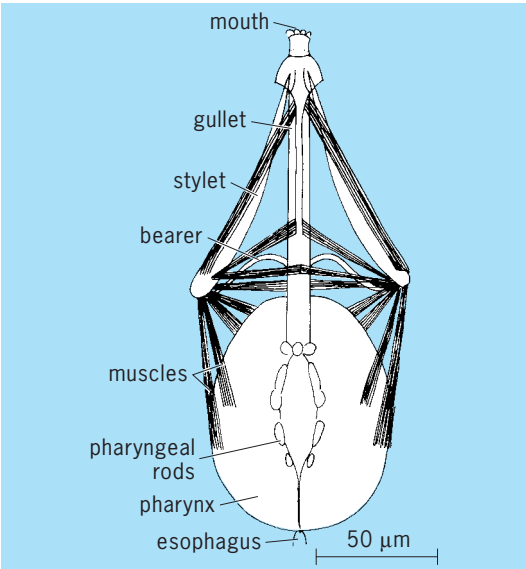

Buccal apparatus of Eutardigrada

resulting in a single opening known as the anogenital pore or cloaca.

Thermozodium esakii, from a warm spring in Japan, evidently represents a third order, the Mesotardigrada. This species combines such echiniscoidean features as the cirrus, clava, claws, and allusive plates with eutardigradan characters: pharyngeal rods (see **illus**.), vasa malpighii, and cloaca. *See* HETEROTARDIGRADA; TARDIGRADA.

Ernesto Marcus; Eveline Marcus

Eutectics

The microstructures that result when a solution of metal of eutectic composition solidifies. The eutectic reaction must be distinguished from eutectic microstructures.

Eutectic reaction. The eutectic reaction is a reversible transformation of a liquid solution to two or more solids, under constant pressure conditions, at a constant temperature denoted as the eutectic temperature T_E. The eutectic phase diagram is notable for displaying negative slopes of both pairs of solidus and liquidus lines, relative to the pure metal or compound terminal components A and B (**Fig. 1**). *See* PHASE EQUILIBRIUM.

Eutectic reactions often constitute only a portion of a complex phase diagram, as in the Cu-Si system; other alloy systems, for example, the Cu-Mg system, may exhibit multiple eutectic reactions between intermediate phases. The transformation may occur between metals only, between metals and nonmetals (as in Ni-TaC), or between nonmetals only (KCl-NaCl).

In all cases the central liquid phase L touches the eutectic temperature line from above, as shown in Fig. 1; the eutectic alloy, at composition c_e, has the lowest melting point of any mixture of components A and B. Solidification of the eutectic composition occurs with the simultaneous crystallization of the two solid phases α and β.

In terms of Gibbs' phase rule, at constant pressure, $F = C + 1 - P$, where F = number of degrees of freedom, C = number of components, and P = number of phases. Binary equilibrium involving three phases allows no degree of freedom: $F = 2 + 1 - 3 = 0$. Consequently, the eutectic reaction is known as an invariant reaction; the three phases must be at the equilibrium temperature and have fixed compositions.

A very similar invariant reaction involving liquid and solid phases is the peritectic reaction, in which two solids and one liquid are at equilibrium; in this case the central solid phase meets the invariant line from below (1). In the solid state the equivalent reactions with decreasing temperature are the eutectoid (2) and peritectoid (3), where S_i denotes solid phases.

$$S_1 + L \xrightarrow{\text{cooling}} S_2 \qquad (1)$$

$$S_1 \rightarrow S_2 + S_3 \qquad (2)$$

$$S_4 + S_5 \xrightarrow{\text{cooling}} S_6 \qquad (3)$$

Eutectic microstructures. Microstructures which are wholly eutectic in nature can occur only for a single, fixed composition c_e in each alloy system demonstrating the reaction (Fig. 1). However, any alloy whose composition passes through the eutectic invariant line (composition c_1 to c_2) undergoes the eutectic reaction as part of its solidification process. Initially, however, the liquid forms some primary crystals of the nearest end-point phase. Whatever liquid remains in the two-phase mixture at the equilibrium temperature then transforms by the eutectic reaction, and the primary crystals are left unchanged.

Microstructure classification. The great variety of patterns observed in eutectic microstructures, in some cases even in a single alloy system, has led to several attempts to develop classification systems. The most successful of these is due to E. Scheil, who classified eutectics on the basis of their mode of crystallization rather than on morphological characteristics, and this classification will be followed here.

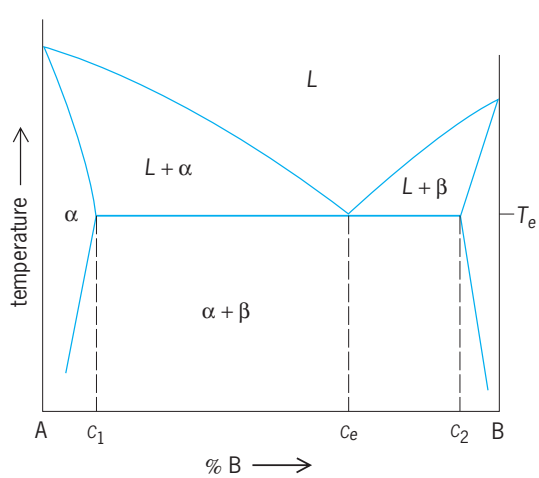

Fig. 1. Schematic eutectic phase diagram between components A and B.

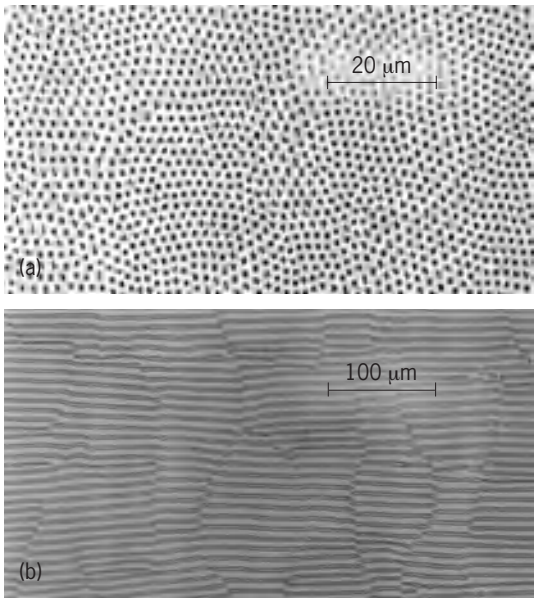

Fig. 2. Cross sections of aligned eutectic composites. (*a*) Fibers of Al_3Ni in Al matrix. (*b*) Alternating lamellae of Mg and $Mg_{17}Al_{12}$. Lighter etching phase is $Mg_{17}Al_{12}$.

Binary eutectic microstructures can be divided into normal and abnormal classes. Normal microstructures comprise primarily lamellar or fibrous types that are formed by simultaneous growth of the two solid phases in the form of parallel fibers in a continuous matrix (**Fig. 2***a*) or as parallel lamellae (Fig. 2*b*). Normal growth occurs when the two solid phases grow locally at the same velocity, resulting in a planar solid-liquid interface. The fiber or lamellar axis tends to be perpendicular to the interface. When freezing begins at more than one center, each center gives rise to a eutectic grain, within which all lamellae or fibers have the same crystallographic orientation. Simple crystallographic relationships between coexisting eutectic phases have been noted; for example, in the lamellar Cd-Zn eutectic the following planes are parallel: $(0001)_{Cd}$ and $(0001)_{Zn}$; and the following directions are parallel: $(0110)_{Cd}$ and $(0110)_{Zn}$. Preferred growth directions have been reported for many systems, but there is no general agreement concerning their existence. *See* CRYSTAL STRUCTURE.

Abnormal eutectics are those in which the two solid phases do not grow at equal velocity. The faster-growing phase (usually that present in smaller volume) grows in a branching or dendriticlike pattern, and the other phase forms from the melt remaining between the branches.

Eutectic fusible alloys. Although technologically important alloy systems, particularly the Fe-C system (all cast irons used commercially pass through a eutectic reaction during solidification), exhibit at least one eutectic reaction, there has been little exploitation of wholly eutectic microstructures for structural purposes. A number of relatively low-melting eutectics are found in binary and higher-order mixtures of bismuth, lead, cadmium, tin, and indium. **Table 1** shows the compositions and melting temperatures of eutectic fusible alloys which are used as solders, as heat-transfer media, for punch and die mold and pattern applications, and as safety plugs. A silver-copper eutectic alloy is also used for high-temperature soldering applications. *See* SOLDERING.

Eutectic composites. During the early 1960s it was recognized that directional solidification of eutectic alloys so as to create a microstructure well aligned parallel to the growth direction could produce high-strength, multiphase composite materials. High-temperature gradients at the melt-solid interface, necessary to produce aligned alloys, are achieved by cooling the base of the crucible containing the liquid alloy as the crucible is withdrawn from the furnace. Excellent mechanical properties have been reported for metal-metal as well as nonmetal-reinforced metallic systems; examples of such alloys, all of which melt above 1832°F (1000°C), are given in **Table 2**. Note that lamellar eutectics are generally characterized by much higher-volume fractions V_f of reinforcing phase than fibrous alloys are. Many of the alloys listed demonstrate elevated temperature creep, stress-rupture, and fatigue properties which are superior to those of the best current conventional alloys. *See* COMPOSITE MATERIAL.

Several factors have been shown to exert major influences on the mechanical behavior of eutectic composites. These include: (1) Degree of alignment; the more perfect the alignment the stronger the alloy.

TABLE 1. Compositions and melting temperatures of eutectic fusible alloys

	Composition, percentage by weight				
T_E, °F (°C)	Bi	Pb	Sn	Cd	Other
116.2 (46.8)	44.70	22.60	8.30	5.30	19.10 In
136 (58)	49.00	18.00	12.00	—	21.00 In
158 (70)	50.00	26.70	13.30	10.00	—
196.7 (91.5)	51.60	40.20	—	8.20	—
203 (95)	52.50	32.00	15.50	—	—
216.5 (102.5)	54.00	—	26.00	20.00	—
255 (124)	55.50	44.50	—	—	—
281.3 (138.5)	58.00	—	42.00	—	—
288 (142)	—	30.60	51.20	18.20	—
291 (144)	60.00	—	—	40.00	—
351 (177)	—	—	67.75	32.25	—
361 (183)	—	38.14	61.86	—	—
390 (199)	—	—	91.00	—	9.00 Zn
430.3 (221.3)	—	—	96.50	—	3.50 Ag
457 (236)	—	79.7	—	17.7	2.60 Sb
477 (247)	—	87.0	—	—	13.00 Sb

TABLE 2. High-strength-high-temperature eutectic alloys

System	T_E, °F (°C)	Form	V_f
Ni-W	2732 (1500)	Fibrous	0.06
Ni,Cr,Al-TaC	2458 (1348)	Fibrous	0.05
Co-TaC	2556 (1402)	Fibrous	0.10
Ni-Ni_3Nb	2318 (1270)	Lamellar	0.26
Ni,Al,Cr-Ni_3Nb	2282 (1250)	Lamellar	0.33
Ni_3Al-Ni_3Ta	2480 (1360)	Fibrous	0.35
Ni_3Al-Ni_3Nb	2336 (1280)	Lamellar	0.44
Co-Cr_7C_3	2377 (1303)	Fibrous	0.30
Co-NbC	2489 (1365)	Fibrous	0.12

(2) Growth speed; $\lambda^2 R$ = constant, where λ = interphase separation and R = growth rate; strength is inversely proportional to $\lambda^{-1/2}$. (3) Perfection of reinforcement; faults in the lamellae or fibers tend to reduce strength. (4) Ductility of coexisting phases; cracks nucleate readily in brittle phases, thereby limiting ductility of the composite. (5) Presence of intentional alloying additions; eutectics with up to nine components have been successfully aligned and reveal superior mechanical properties.

Among the major advantages of these alloys are extraordinary thermal stability of unstressed microstructures, retention of high strength to very close to the eutectic temperature of the respective alloys, and the ability to optimize strength by appropriate alloying additions to induce either solid-solution strengthening or intraphase precipitation of additional phases.

Future applications. The most likely applications for aligned eutectics are as gas turbine engine materials (turbine blades or stator vanes) or in nonstructural applications such as superconducting devices in which directionality of physical properties is an important requirement. *See* ALLOY; METAL, MECHANICAL PROPERTIES OF.

Norman S. Stoloff

Bibliography. K. Budinski and M. K. Budinski, *Engineering Materials*: *Properties and Selection*, 8th ed., 2004; J. C. M. Li (ed.), *Microstructure and Properties of Materials*, 2000.

Eutheria

A higher-level taxon that includes all mammals except monotremes and marsupials. Eutheria (Placentalia) is variously ranked as an infraclass or cohort within Mammalia. Eutheria includes over 4000 living species arranged in 18 orders; another 12 orders are known only from fossils. An ecologically diverse group, Eutheria includes primates, insectivores, bats, rodents, carnivores, elephants, ungulates, and whales. Like other mammals, eutherians are generally fur-covered and produce milk to nourish their young. In part because they can make their own body heat and regulate their body temperature, eutherians are widely distributed over most continents and occur in all oceans.

Reproduction. Eutherians, often called placental mammals, have a unique reproductive system in which unborn young are nourished for an extended period via a placenta. The placenta forms inside the uterus of the mother, and the developing fetus is attached to it by an umbilical cord. The placenta provides an interface between the bloodstream of the fetus and that of the mother, allowing the transport of oxygen and nutrients from the mother to the fetus and the transport of carbon dioxide and wastes from the fetus to the mother. This system permits retention of the young in the protective environment of the uterus during most of early development. Fetal survival rates are high under most conditions. Young are born in a relatively advanced state of development and are never sheltered in a pouch after birth. Gestation time ranges from 20 days (for example, shrews and hamsters) to 22 months (elephants). Many eutherians have only one or two young per pregnancy, but as many as 20 offspring may be produced at a single birth in some species (such as some tenrecs).

Morphology and ecology. Eutherians range in size from insectivores and bats that weigh only a few grams to blue whales that can weigh over 190,000 kg (420,000 lb). All have a relatively large brain in which the two hemispheres are connected by a corpus callosum, which facilitates information transfer between the hemispheres. Eutherians exhibit more variation in ecology than any other group of vertebrates, and these differences are reflected in their morphological specializations. Most eutherians have a complex dentition, and structure of the teeth is indicative of diet. Plant-eating eutherians generally have teeth with grinding ridges, while meat-eating taxa have teeth for puncturing and shearing. Some groups (for example, rodents) have enlarged, ever-growing incisors that are used for gnawing. The sharp, bladelike incisor teeth of vampire bats are used to inflict tiny cuts in the skin of their prey, and the bats obtain a blood meal by lapping at the wound. Many ant- and termite-eating eutherians have a reduced dentition and use a long, sticky tongue to obtain their food. Long tongues are also used by other eutherians to feed on nectar in flowers (such as glossophagine bats). Some whales lack teeth and use flexible sheets of baleen to sieve small organisms from seawater. Baleen is made of keratin, the same material used to make hair, fingernails, claws, hooves, and the outer covering of horns.

Body form and limb structure reflect ecological habits. Eutheria includes the only mammals to evolve a fully aquatic life-style (for example, whales and sirenians), and these taxa have a fusiform body and paddlelike limbs for underwater locomotion. Whales have lost the hindlimbs and use a new structure, the tail fluke, for propulsion. Eutheria also includes bats, the only flying mammals, which have wings that consist of skin membranes supported by elongated forelimbs and fingers. Subterranean moles have forelimbs modified into shovellike structures for digging, arboreal sloths have hooklike hands for climbing and hanging under tree branches, ungulates have elongated limbs and hooves for running, and elephants have columnar limbs and modified foot pads to support great body weight. All of these modifications evolved from a primitive mammalian body plan that probably included small body size (less than 1 kg or 2 lb), an insectivorous diet, quadrupedal locomotion, and scansorial (scrambling) habits.

Sensory systems. Most eutherians have good vision, hearing, and a keen sense of smell. The relative importance of these sensory systems varies among taxa. Most diurnal eutherians (for example, ungulates, squirrels, and humans) have large eyes and well-developed vision. The eyes are reduced in many fossorial eutherians (such as moles). All bats retain functional eyes. Old World fruit bats (Megachiroptera) have excellent vision; echolocating bats

(Microchiroptera) rely on a sophisticated sonar system for navigation. High-frequency sounds produced in the larynx are emitted from the nose or mouth; the bats listen for and analyze returning echoes to obtain information about their environment. Some whales (odontocetes) also use echolocation.

Social behavior. All eutherians exhibit complex behaviors, and many live in social groups. Vocalizations, facial expressions, and scent marking play an important part in social interactions. Diurnal eutherians (for example, many primates and ungulates) often have colorful coat markings that may be used in social displays. Nocturnal forms often rely on olfactory rather than visual cues. The basic social bond is between mother and young. Apart from mother-offspring interactions, some species are solitary and interact with others only to mate. Even in social species, the sexes may remain separate except at mating season. However, many eutherians live in complex, relatively stable social groups (such as primates, elephants, and wolves). Monogamy is uncommon, occurring in fewer than 3% of eutherian species. Harem groups, in which large groups of females are guarded by a single male, are common, but many eutherians live in groups with multiple males and multiple females. One unusual social system occurs in naked mole rats (*Heterocephalus glaber*), which live in underground colonies organized like those of social insects, such as honeybees and ants.

Fossil record. The fossil record of Eutheria extends back at least into the Cretaceous Period. Several differences in the skull and dentition distinguish fossil eutherians from early members of other mammal lineages (for example, marsupials). The earliest eutherians were apparently small, nocturnal mammals that may have resembled some modern insectivores. Although Cretaceous eutherians are known from most continents, diversification of the modern orders apparently did not occur until the Paleocene and Eocene. Many species are known only from fossil teeth and skull fragments, but exceptionally well-preserved fossils of many groups have been found at unique sites such as Messel, Germany, and La Brea, California. *See* ARTIODACTYLA; CETACEA; CHIROPTERA; DENTITION; MAMMALIA; METATHERIA; THERIA.

N. B. Simmons

Bibliography. B. Grizmek, *Grizmek's Animal Life Encyclopedia: Mammals*, vols. 10–13, Van Nostrand, 1972; L. B. Halstead, *The Evolution of Mammals*, Eurobook, 1978; D. W. MacDonald, *The Encyclopedia of Mammals*, Facts on File, 1984; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; W. Voelker, *The Natural History of Living Mammals*, Plexus Publishing, 1986.

Eutriconodonta (Triconodonta)

One of the taxonomically most diverse mammalian groups currently known from the Jurassic and Cretaceous (approximately 200 to 65 million years ago) when they were represented by 24 genera and 37 named species (all extinct). The name Eutriconodonta refers to the characteristic shape of their posterior cheek teeth, which consist of three or sometimes four cusps aligned anteroposteriorly (**Fig. 1**). The lower molariform teeth in many eutriconodontans are linked by a tongue-in-groove articulation formed by a projection from the posterior end of one tooth that fits into a groove in the anterior end of the following tooth. *See* DENTITION; MAMMALIA.

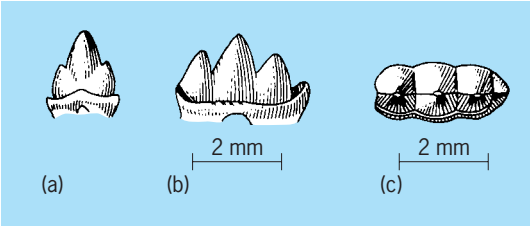

Fig. 1. Eutriconodontan molars had three main cusps. (*a*) External view of a lower molar of *Amphilestes*, an amphilestid eutriconodontan. (*b*) Internal view and (*c*) occlusal view of a lower molar of *Priacodon*, a triconodontid eutriconodontan.

Classification. In recent years, discovery and analysis of new fossils—some nearly complete skeletons—and revised interpretations of evolutionary relationships have greatly improved our understanding of the pattern of mammalian evolution during the Jurassic and Cretaceous. Previous interpretations grouped the mammals discussed here with morganucodontans and docodontans in the group Triconodonta. These three groups are now recognized as not being particularly closely related. For the sake of clarity, the core members of the Triconodonta have been united in a group designated Eutriconodonta, which includes three families. The Amphilestidae is made up of forms representing the primitive structural grade of the eutriconodontans. Although there is still uncertainty about their interrelationships, the other two families, Gobiconodontidae and Triconodontidae, appear to have evolved from an amphilestid ancestry. *See* ANIMAL EVOLUTION; DOCODONTA.

Fossil record. The oldest, but questionable, record of a eutriconodontan is from Early Jurassic deposits in India. Unquestionable records of the group are known from Middle Jurassic sites in England, China, and Mexico. The Late Jurassic and Early Cretaceous were the times of the eutriconodontans' greatest abundance, diversity, and biogeographic range, which included North America, western Europe, Asia, and Africa. Eutriconodontans survived into the latest Cretaceous of North America and possibly South America but disappeared from the fossil record before the mass extinctions marking the end of the Cretaceous. *See* EXTINCTION (BIOLOGY).

Characteristics. In comparison to modern mammals, the skulls of eutriconodontans were large relative to their bodies. Overall their skeletons were robustly built. Eutriconodontans exhibited a broad range of body sizes. Some were small, about the size

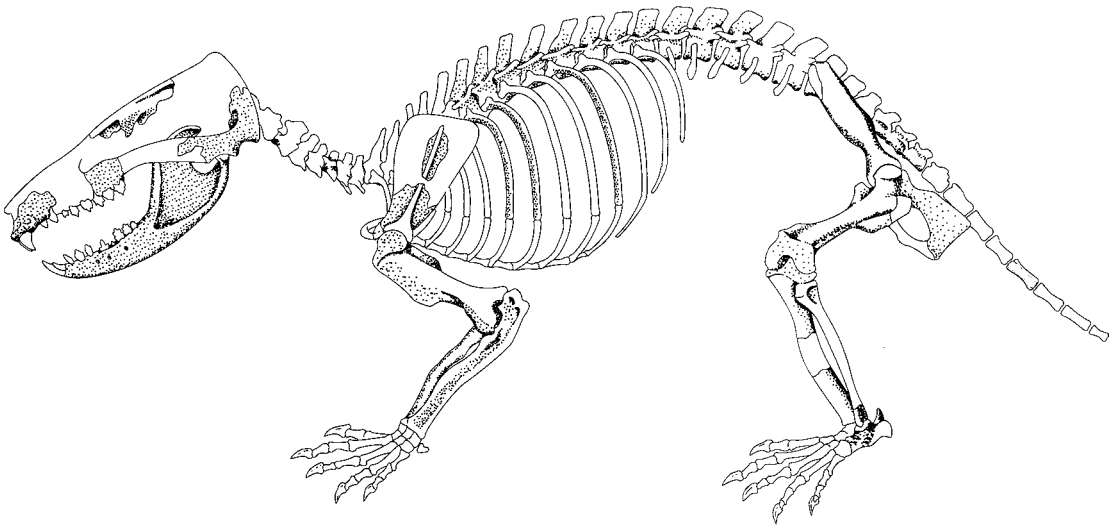

Fig. 2. Partial restoration of the skeleton of *Gobiconodon*, a gobiconodontid eutriconodontan. The length of the tail of *Gobiconodon* is unknown. In other gobiconodontids the tail contributed approximately 50% of the total length of the body. (*After F. A. Jenkins, Jr., and C. Schaff, The Early Cretaceous Mammal Gobiconodon (Mammalia, Triconodonta) from the Cloverly Formation of Montana, J. Vert. Paleontol., 8:1–24, 1988*)

of modern mice. The largest eutriconodontans are found among the gobiconodontids. *Gobiconodon*, for example, grew to about the size of an adult Virginia opossum, *Didelphis virginiana* (approximately 5.5 kg or 12 lb) [**Fig. 2**]. The largest currently known Mesozoic mammals are species of the gobiconodontid *Repenomamus*, which are particularly well known from skeletons found in Early Cretaceous deposits in Liaoning Province, China. The largest species, *R. giganticus*, is estimated to have weighed 12–14 kg (26–30 lb). In any Late Jurassic or Early Cretaceous terrestrial fauna, particularly on continents in what is now the Northern Hemisphere, eutriconodontans usually were the largest mammals.

Evolution of an articulation between the dentary (the tooth-bearing bone of the lower jaw) and squamosal of the skull is a derived character of modern mammals. This joint replaced the primitive articular-quadrate articulation of their ancestors, the nonmammalian synapsids. A dentary-squamosal articulation was fully functional in all eutriconodontans. In contrast, in at least some eutriconodontans the primitive articular-quadrate articulation was retained. The articular remained connected to the dentary, while the quadrate contacted the squamosal. The articular and quadrate were not completely freed from their function of providing part of the articulation of the lower jaw to the skull. Possibly, in other eutriconodontans these bones were excluded from the jaw articulation and formed the malleus (articular) and incus (quadrate), which function as part of the sound-transmitting apparatus characteristic of the mammalian middle ear. *See* EAR (VERTEBRATE).

In other skeletal characteristics the eutriconodonts exhibit a mosaic of primitive and advanced conditions. For example, in the few eutriconodontan skeletons discovered so far, the posture of the forelimb with the elbow held relatively close to the body resembles the posture of most modern quadrupedal mammals. In contrast, the hindlimb preserved a primitive, sprawling posture. This and many other lines of evidence support the hypothesis that during the Jurassic and Cretaceous mammalian evolution was not characterized by simple, straight-line acquisition of new characters but involved a mosaic of changes in various lineages.

Probably many if not most eutriconodontans were carnivorous—predators, scavengers, or both. The morphology of their lower jaws, lacking an angular process but with a high coronoid process, indicates evolution of strong mandibular adductor musculature giving them a powerful bite. Evolution of a tongue-in-groove articulation linked the blades of the lower molars into an effective chopping or shearing mechanism. The canines and, in some species, the anterior incisors were large, stabbing teeth. Direct evidence of the diet of *R. giganticus* was provided by discovery of a skeleton with the remains of a juvenile dinosaur, *Psittacosaurus*, preserved in its stomach region.

Recent discoveries of new species of eutriconodonts have dispelled the view that during the Jurassic and Cretaceous mammals were uniformly small, insectivorous forms that scurried through the trees or hid on the ground to avoid the dominant dinosaurs. In addition to documenting the diversity of eutriconodonts, discoveries of aquatic otter- and beaverlike docodonts add to the evidence that mammals occupied a wide variety of niches in the ecosystems of the time. William A. Clemens

Bibliography. T. S. Kemp, *The Origin and Evolution of Mammals*, Oxford University Press, 2005; Z. Kielan-Jaworowska, R. Cifelli, and Z.-X. Luo, *Mammals from the Age of Dinosaurs: Origins, Evolution, and Structure*, Columbia University Press, New York, 2004.

Eutrophication

The deterioration of the esthetic and life-supporting qualities of lakes and estuaries, caused by excessive fertilization from effluents high in phosphorus, nitrogen, and organic growth substances. Algae and aquatic plants become excessive, and when they decompose, a sequence of objectional features arises. Water for consumption from such lakes must be filtered and treated. Diversion of sewage, better utilization of manure, erosion control, improved sewage treatment and harvesting of the surplus aquatic crops alleviate the symptoms. Prompt public action is essential. *See* WATER CONSERVATION.

Extent of problem. In inland lakes this problem is due in large part to excessive but inadvertent introduction of domestic and industrial wastes, runoff from fertilized agricultural and urban areas, precipitation, and groundwaters. The interaction of the natural process with the artificial disturbance caused by human activities complicates the overall problem and leads to an accelerated rate of deterioration in lakes. Since population increase necessitates an expanded utilization of lakes and streams, cultural eutrophication has become one of the major water resource problems in the United States and throughout the world. A more thorough understanding must be obtained of the processes involved. Without this understanding and the subsequent development of methods of control, the possibility of losing many of the desirable qualities and beneficial properties of lakes and streams is great. *See* WATER POLLUTION.

Cultural eutrophication is reflected in changes in species composition, population sizes, and productivity in groups of organisms throughout the aquatic ecosystem. Thus the biological changes which are caused by excessive fertilization are of considerable interest from both the practical and academic viewpoints. *See* FRESH-WATER ECOSYSTEM; LAKE; LIMNOLOGY.

Phytoplankton. One of the primary responses to eutrophication is apparent in the phytoplankton, or suspended algae, in lakes. The nature of this response can be examined by comparing communities in disturbed and undisturbed lakes or by following changes in the community over a period of years during which nutrient input is increased. *See* ECOLOGICAL COMMUNITIES; PHYTOPLANKTON.

The former approach was utilized in studies at the University of Wisconsin, in which the overall structure of the phytoplankton communities of the eutrophic Lake Mendota, at Madison, and the oligotrophic Trout Lake, in northern Wisconsin, was analyzed. These investigations showed that in the eutrophic lake the population of species is slightly lower, although the average size of organisms is considerably larger, indicating higher levels of production than in the oligotrophic lake. When compared in terms of an index of species diversity, the community of the eutrophic lake displayed values lower than those observed in the oligotrophic lake. Seasonal changes and bathymetric differences in the index of diversity were also more apparent in the eutrophic lake.

Often the low species diversity of the phytoplankton in eutrophic lakes is a result of high populations of blue-green algae, such as *Aphanizomenon flosaquae* and *Anabaena spiroides* in Lake Mendota. Frequently, however, species of diatoms such as *Fragillaria crotonensis* and *Stephanodiscus astrae* also attain high degrees of dominance in the community. Dense populations of the blue-green algae *Oscillatoria rubescens* were indicative of the deteriorating conditions in Lake Zurich (Switzerland) and Lake Washington (Seattle), but since diversion of sewage from these lakes, they are no longer a problem. The same species has been observed in several other lakes that have undergone varying degrees of cultural eutrophication. The relatively high nutrient concentrations in eutrophic waters appear to be capitalized on by one or two species that outcompete other species and periodically develop extremely high population levels. Because of the formation of gas vacuoles during metabolism, senescent forms of the blue-green algae rise to the surface of the lake, causing nuisance blooms.

In addition to nuisance scums in the pelagial, or open-water regions, the rooted aquatic plants and the attached algae of the littoral, or shoreward, region often prove to be equally troublesome in eutrophic lakes. Species of macrophytes, such as *Myriophyllum* and *Ceratophyllum*, and algal forms, such as *Cladophora*, frequently form dense mats of vegetation, making such areas unsuited for both practical and recreational uses.

Bottom fauna. Often in eutrophic lakes the bottom fauna display characteristics similar to those observed in the algal community. Changing environmental conditions appear to allow one or two species to attain high degrees of dominance in the community. Generally higher levels of production are associated with the change in structure of the community—the result being nuisance populations of organisms. In Lake Winnebago, in Wisconsin, for example, the lake fly or midge *Chironomus plumosus* develops extremely high populations, which as adults create an esthetic as well as an economic problem in nearby cities.

Great Lakes. It was generally thought that eutrophication would not be a major problem in large lakes because of the vast diluting effect of their size. However, there has been evidence that eutrophication is occurring in the lower Great Lakes. Furthermore, the undesirable changes in the biota appear to have been initiated in relatively recent years. Utilizing long-term records from Lake Erie, qualitative and quantitative changes have been observed in phytoplankton of that large body of water owing to cultural eutrophication. Total numbers of phytoplankton have increased more than threefold since 1920, while the dominant genera have changed from *Asterionella* and *Synedra* to *Melosira*, *Fragillaria*, and *Stephanodiscus*.

Other biological changes usually associated with the eutrophication process in small lakes have also

been observed in the Great Lakes. Of the five lakes, Lake Erie has undergone the most noticeable changes due to eutrophication. In terms of annual harvests, commercially valuable species of fish, such as the lake herring or cisco, sauger, walleye, and blue pike, were replaced by less desirable species, such as the freshwater drum or sheepshead, carp, and smelt. Similarly, in the organisms living in the bottom sediments of Lake Erie, drastic changes in species composition have been observed.

Oxygen demand. It is apparent that the increase in organic matter production by the algae and plants in a lake undergoing eutrophication has ramifications throughout the aquatic ecosystem. Greater demand is placed on the dissolved oxygen in the water as the organic matter decomposes at the termination of life cycles. Because of this process, the deeper waters in the lake may become entirely depleted of oxygen, thereby destroying fish habitats and leading to the elimination of desirable species. The settling of particulate organic matter from the upper, productive layers changes the character of the bottom muds, also leading to the replacement of certain species by less desirable organisms. Of great importance is the fact that nutrients inadvertently introduced to a lake are for the most part trapped there and recycled in accelerated biological processes. Consequently, the damage done to a lake in a relatively short time requires a many-fold increase in time for recovery of the lake.

Action programs. Lake eutrophication represents a complex interaction of biological, physical, and chemical processes. The problem, therefore, necessitates basic research in a wide variety of scientific disciplines. Moreover, to be profitable, such research must be coordinated into well-integrated team-research efforts requiring extensive monetary support.

Studies are needed in such areas as monitoring the amount of those nutrients that reach critical levels in lake waters and lead to the development of nuisance growths of plants and algae. Nitrogen and phosphorus are undoubtedly important.

Sewage effluent is the major contributor of nitrogen and phosphorus to lakes, followed by runoff from manured and fertilized land. Considering nitrogen alone, rain adds more than any single source. Its nitrogen content comes from combustion engines and smokestacks. Where sewage effluent and agricultural drainage have been diverted from lakes, an improvement in nuisance conditions occurs. Dead lakes become alive again; hence this treatment is the first step in alleviation.

The conditions observed in lakes and streams reflect not only the processes operating within the body of water but also the metabolism and dynamics of the entire watershed or drainage basin. After precise identification of the critical nutrient compounds, it is necessary to determine the nutrient budget of the whole drainage basin before acting to alleviate the undesirable fertilization of a lake or stream.

Methods of treating sewage plant effluent are being explored, and further support for these efforts is justified. Agricultural practices such as low tillage reduces erosion, hence the amount of fertilizer in the runoff. *See* AGRICULTURAL SOIL AND CROP PRACTICES; EROSION; SEWAGE TREATMENT.

It is known that aquatic plants concentrate nutrients from the lake waters in their tissues. The removal or harvesting of aquatic plants in eutrophic lakes, consequently, is a good potential method for reducing nutrient levels in these lakes. Similarly, significant amounts of nitrogen and phosophorus are concentrated in fish flesh. Efficient methods of harvesting these organisms are important in impoverishing a well-fertilized lake.

Utilization or land disposal of farm manure is a major problem. Animal manures are largely unsewered, yet in the Midcuest it is equivalent to the sewage of 350,000,000 people.

In addition to improvements in waste disposal, more research and development are needed on the profitable utilization of surplus algae, aquatic plants, fish, manure, and sewage. The overfertilized lake needs to be impoverished of its nutrients as well as protected from inflowing sources. Chemicals have been used to poison the plants and algae, but this is not a good conservation practice because the plants and algae rot and provide more nutrients. Moreover, eventual harm to other species has not been assessed.

It would seem desirable to set aside certain lake areas for research purposes. More information is needed to decide upon the best plans for allowing the domestic development of these areas with the least disturbance to the water resources. Steps will be necessary to devise optimum zoning laws and multiple-use programs in light of the intense economic and recreational uses made of water resources. Legislation and law enforcement in relation to public interactions undoubtedly will be a complex problem to overcome in this respect.

Cultural eutrophication is a paradoxical condition, since it is in large part due to human economic, agricultural, and recreational activities and at the same time eventually conflicts with these same activities of society in general. Arthur D. Hasler

Bibliography. G. D. Cooke, *Restoration and Management of Lakes and Reservoirs*, 3d ed., 2005; F. de Jong, *Marine Eutrophication in Perspective: On the Relevance of Ecology for Environmental Policy*, 2006; A. J. Horne and C. R. Goldman, *Limnology*, 2d ed., 1994; M. C. T. Scholten et al., *Eutrophication Management and Ecotoxicology*, 2005.

Evaporation

The process by which a liquid is converted into a vapor. In the liquid phase, the substance is held together by intermolecular forces. As the temperature is raised, the molecules move more vigorously, and in increasingly high proportion have sufficient energy to escape from their neighbors. Evaporation is therefore slow at low temperatures but faster at higher

temperatures. In an open vessel, the molecules escape from the vicinity of the liquid, and there is a net migration from the liquid to the atmosphere. In a closed vessel, net evaporation continues until the number of molecules in the vapor has risen to the stage at which the rate of return from the vapor to the liquid is equal to the rate of evaporation. At this stage there is a dynamic equilibrium between the liquid and its vapor, with evaporation and its reverse, condensation, occurring at the same rate. The pressure of the vapor in the closed vessel is called the vapor pressure of the substance; its value depends on the temperature. Boiling occurs in an open vessel (but not in a closed vessel) when the vapor pressure is equal to the ambient pressure. *See* BOILING POINT; KINETIC THEORY OF MATTER; VAPOR PRESSURE.

Evaporation is an endothermic (heat-absorbing) process because molecules must be supplied with energy to overcome the intermolecular forces. The enthalpy of vaporization, $\Delta_{vap}H$ (formerly, the latent heat of vaporization), is the heat required at constant pressure per mole of substance for vaporization. The entropy of vaporization, $\Delta_{vap}S$, at the boiling point, T_b, is equal to $\Delta_{vap}H/T_b$. According to Trouton's rule, for many liquids the entropy of vaporization is close to 85 J/K · mol. This value reflects the similar change in disorder that occurs when a liquid is converted into a gas. However, certain liquids (water and mercury among them) are more structured than others, and have a bigger entropy of vaporization than Trouton's rule suggests. *See* ENTHALPY; ENTROPY; THERMODYNAMIC PRINCIPLES.

Volatile liquids evaporate more rapidly than others at the same temperature. Such liquids have relatively weak intermolecular forces. In general, the rate of evaporation depends on the strengths of the intermolecular forces and the rate at which heat is supplied to the liquid. *See* INTERMOLECULAR FORCES; LIQUID. P. W. Atkins

Bibliography. P. W. Atkins and J. de Paula, *Physical Chemistry*, 8th ed., 2006.

Evaporator

A device used to vaporize part or all of the solvent from a solution. The valuable product is usually either a solid or a concentrated solution of the solute. If a solid, the heat required for evaporation of the solvent must have been supplied to a suspension of the solid in the solution; otherwise the device would be classed as a drier. The vaporized solvent may be made up of several volatile components, but if any separation of these components is effected, the device is properly classed as a still or distillation column. When the valuable product is the vaporized solvent, an evaporator is sometimes mislabeled a still, such as water still, and sometimes is properly labeled, such as boiler-feedwater evaporator. In the great majority of evaporator installations, water is the solvent that is removed. *See* DRYING.

Uses. Evaporators are used primarily in the chemical industry. Common salt is made by boiling a saturated brine in an evaporator. The salt precipitates as a solid in suspension in the brine. This slurry is pumped continuously to a filter, from which the solids are recovered and the liquid portion returned for further evaporation. In the manufacture of pulp and paper, the waste liquor from cooking the wood is a dilute solution of inorganic cooking chemicals and soluble, organic wood constituents. The liquor is disposed of by concentrating it in evaporators to a strength at which the liquor will burn in a boiler. This avoids a water-pollution problem, recovers the inorganic chemicals for reuse, and provides sufficient heat energy not only to operate the evaporator but also to supply other needs of the mill as well. In the alkali industry, salt brine passed through a diaphragm-type electrolytic cell yields a dilute solution of salt and caustic soda. Evaporation to a strength of about 50% NaOH purifies the caustic by causing precipitation of practically all of the salt. It also simplifies shipment by bringing about a more than fivefold reduction in volume. Further evaporation of the caustic solution to a final temperature of about 700°F (371°C) produces a practically anhydrous product that freezes at about 600°F (316°C) and is shipped as a solid. Evaporators are widely used in the food industry, usually as a means of reducing volume to permit easier storage and shipment. Evaporators are also the most commonly used means of producing potable water from seawater or other contaminated sources.

Classification. The vaporization of solvent requires large amounts of heat. Provisions for transferring this heat to the solution constitute the largest element of evaporator cost and the principal means of distinguishing between types of evaporators. Practically all evaporators fall into one of the following categories:

1. Those heated by a flame that burns below the liquid surface, and in which the hot combustion gases are bubbled through the liquid.
2. Those in which the flame and combustion gases are separated from the boiling liquid by a metal wall, or heating surface.
3. Those in which steam or other condensable vapor is the source of heat, and in which the steam condenses on one side of the heating surface and the heat is transmitted through the wall to the boiling liquid.

Submerged-combustion evaporators (type 1) are used primarily to concentrate solutions that would deposit a heat-insulating blanket of scale on the solid heating surfaces of other types of evaporators. Since the evolved vapor is mixed with the combustion gases, neither the solvent vapor nor its heat content can be recovered easily.

Direct-fired evaporators (type 2) are typified by the steam boiler and the old maple-syrup kettle. They are not commonly used for the concentration of solutions, primarily because local overheating can cause formation of insulating deposits on the heating surface, which then becomes overheated and may melt or burn through. Also, a large heating surface is needed to recover the heat in the combustion gases, and there is usually no cheap metal that will resist

attack by both the combustion gases and the boiling liquid.

Steam-heated evaporators (type 3) are by far the most common, primarily because condensing steam gives up its heat so readily. Condensing steam film coefficients are usually in excess of 1000 Btu/(h)(ft^2)(°F) [5.7 kW/(h)(m^2)(°C)]. Thus the design of the evaporator and materials of construction can be suited to the solution being concentrated instead of being dictated by the problem of getting heat to the heating surface. The heating surface is usually in the form of metal tubes, since this represents the most economical method of putting the largest heating surface in the smallest volume. The tubes may be vertical or horizontal, and the boiling liquid may be either inside or outside, depending on the characteristics of the solution, such as viscosity, ratio of feed to evaporation, and whether or not a salt is deposited. *See* DISTILLATION.

Operation. The vaporization capacity of an evaporator is determined by the usual rules of heat transfer, and is directly proportional to the area of heating surface, to the difference between condensing steam and boiling liquid temperatures, and to the coefficient of heat transfer. The heat-transfer coefficient is usually limited by conditions on the boiling liquid side, although the condensing steam film and the resistance of the metal wall have some influence. Various means are employed to increase the boiling-film coefficient and all involve movement of the liquid relative to the heating surface. A great many evaporators use only natural convection to accomplish this circulation. Typical is the long-tube vertical type shown in the **illustration**. Feed liquid enters the bottom of a nest of vertical tubes and begins to boil as it passes up the tubes. The boiling causes a large increase in volume, which accelerates the liquid to high velocities and gives good heat-transfer performance. The vapor-liquid mixture is separated in the chamber at the top of the tubes. The liquor may all be discharged as product, or part may be recirculated to the feed inlet.

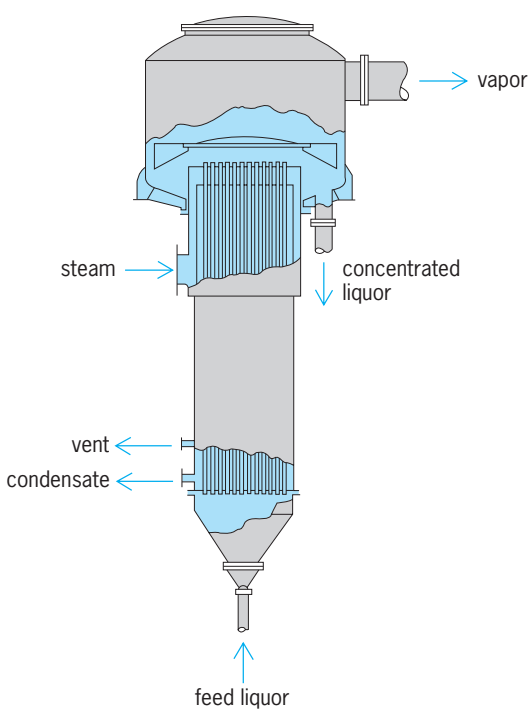

Long-tube vertical evaporator.

A pump or agitator may be used to force the liquid past the heating surface if even higher heat-transfer coefficients are needed, for example, when corrosive conditions dictate use of the smallest possible area of an expensive alloy. Forced circulation evaporators are also used for scaling liquids or those from which a salt is to be crystallized, since there is less tendency for solids to form on the heating surfaces. Such evaporators usually consist of a flash chamber or crystallizing chamber, a conventional shell-and-tube heat exchanger, and a pump to circulate fluid from the chamber to the exchanger and back to the flash chamber. Another type of forced-circulation evaporator for extremely viscous, heat-sensitive, or foamy materials either rotates the heating surfaces or employs wipers that sweep the material around the walls of a stationary surface.

The water vapor evolved in an evaporator is usually about the same in quality and quantity as the steam used to heat the evaporator. The only difference is that the vapor has a lower pressure and hence lower condensing temperature. It is possible to compress the vapor so that it can be used as the heating steam in the same evaporator. Such thermocompression evaporators require far less energy for the compressor than would be needed to generate fresh steam. However, to keep the compressor cost and power consumption within reason, it is necessary to use a narrow compression ratio and this requires a large and expensive evaporator.

Assuming a perfect compressor, the power requirement is given by the Carnot equation: $W = Q\Delta T/T$, where W and Q are work required and amount of heat pumped in the same units, ΔT is the difference between saturated steam temperatures at compressor discharge and suction pressures, and T is the absolute suction temperature. For pure water boiling at atmospheric pressure (212°F or 100°C), the latent heat is 970.3 Btu/lb (2257 kilojoules/kg). At a temperature difference of 10°F (5.6°C), the ideal work required is only (970.3)(10)/(672) = 14.4 Btu/lb (33.6 kilojoules/kg), or 35.3 kWh per 1000 gal of water evaporated (33.6 MJ per 1000 liters). This is only 1.5% of the heat energy required for a simple evaporator, but the energy is expensive mechanical energy rather than low-grade heat energy.

If pure water is being boiled, this 10°F (5.6°C) temperature difference established by the compressor is available as the driving force to transfer heat from the compressed steam through the heating surface to the boiling liquid. If the boiling liquid is an aqueous solution, it has a boiling temperature higher than that of pure water at the same pressure and the difference in these temperatures is called the boiling point elevation (bpe). The bpe cannot be utilized as a part of the

driving force. Thus, if the solution being evaporated had a bpe of 30°F (16.7°C), the compressor would have to work across a 40°F (22.2°C) compression range and would require four times the above power to establish the same 10°F (5.6°C) driving force in the evaporator.

An alternative method of reducing the energy requirement of an evaporator is to use the water vapor evolved in one part to heat another part in which liquid is boiled at a lower temperature. The water vapor evolved in this part, termed an effect, can then be used to heat another effect boiling at still lower temperature (and pressure), and so on. The ultimate limit is determined by the need to discharge to a heat sink the heat contained in the vapor from the last effect, which amounts to most of the heat supplied by steam to the first effect (the balance of the entering heat leaves with the condensate and concentrate, which are generally hotter than the feed solution). The heat sink is usually water from a river or other source, which limits the boiling temperature in the last effect to about 100°F (38°C).

If the condensate of last-effect vapors is valuable, the vapors are condensed in a shell-and-tube heat exchanger by the cooling water. If the solution is the desired product, the vapors are brought into direct contact with a shower of cooling water. This is usually done in a barometric condenser, which is elevated to such a height (about 34 ft or 10.3 m) that the high vacuum cannot prevent the water from draining out by gravity. To maintain the vacuum in the condenser, it also is necessary to remove noncondensible gases with a vacuum pump. These gases originate in the feed, as air dissolved in the condenser water and as air leakage into the evaporator. Only in the rarest of circumstances, as when a very low boiling temperature is needed to avoid degradation of the solution, is it feasible to compress the entire last-effect vapor for discharge at a higher temperature to the heat sink.

Such multiple-effect evaporators are more expensive than single-effect units because each effect can operate at only a fraction of the total difference, ΔT, between the temperature at which heat is accepted from the prime steam and the temperature at which it is rejected to the heat sink. For a single effect, the amount of heat transfer surface A required is given by the equation: $A = Q/U\Delta T$, where Q is the rate at which heat must be transferred to achieve the desired evaporation rate and U is the coefficient of heat transfer.

In a triple effect, each effect evaporates only about one-third of the water and must transfer only about one-third of the total heat, but the total temperature difference must also be divided between the three effects. Thus the heating surface in each effect may be written: $A_n = (Q/3)/U(\Delta T/3) = Q/U\Delta T$. Each effect must have substantially the same amount of surface as a single-effect evaporator. The steam required, however, goes down in inverse proportion to the number of effects. The choice of the proper number of effects involves an economic balance between the first cost of equipment and the continuing costs of steam and cooling water. The great majority of evaporator installations employ the multiple-effect principle and as many as a dozen effects have been used. Another means of achieving multiple-effect steam economy is the multistage flash cycle, which has been developed primarily for producing fresh water from seawater. *See* HEAT EXCHANGER; HEAT TRANSFER; SALINE WATER DESALINATION. Ferris C. Standiford

Bibliography. J. A. Kent (ed.), *Kent and Riegel's Handbook of Industrial Chemistry and Biotechnology*, 11th ed., 2007; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineer's Handbook*, 7th ed., 1997; R. Smith, *Chemical Process: Design and Integration*, 2d ed., 2005.

Evergreen plants

Plants that retain their green foliage throughout the year. Popularly, needle-leaved trees (pine, fir, juniper, spruce) and certain broad-leaved shrubs (rhododendron, laurel) are called evergreens. In warm regions many broad-leaved trees (magnolia, live oak) are evergreen, and in the tropics most trees are evergreen and nearly all have broad leaves. Many herbaceous biennials and perennials have basal rosettes with leaves close to the ground that remain green throughout the winter. *See* FOREST AND FORESTRY; LEAF; PLANT TAXONOMY. Nelle Ammons

Evolutionary computation

Evolutionary computation is a rapidly growing interdisciplinary science area that is concerned with modeling aspects of natural evolution in order to solve real-world problems. Living organisms, as well as those long extinct, demonstrate optimized complex behavior at all levels: cells, organs, individuals, and populations. Charles Darwin wrote of "organs of extreme perfection" when describing the ability of evolution to craft ingenious solutions to complex problems such as vision. Evolution is the great unifying principle of biology, but it extends beyond biology and can be used as an engineering principle where individuals in a population of candidate solutions to some particular problem undergo random variation (mutation and recombination) and face competition and selection based on their appropriateness for the task at hand.

Process. The common use of an evolutionary algorithm requires four elements: (1) an evaluation (fitness) function that describes the quality of any candidate solution in quantitative terms, (2) a representation or data structure that the computer uses to store solutions, (3) a random variation operator (or operators) that transform "parents" into "offspring," and (4) a means for selecting which solutions will survive to the next generation and which will be eliminated. In addition, the process must be initialized with a population of candidate solutions to the task at hand. This is often accomplished by seeding the first population with completely random solutions; however, if domain-specific knowledge is available

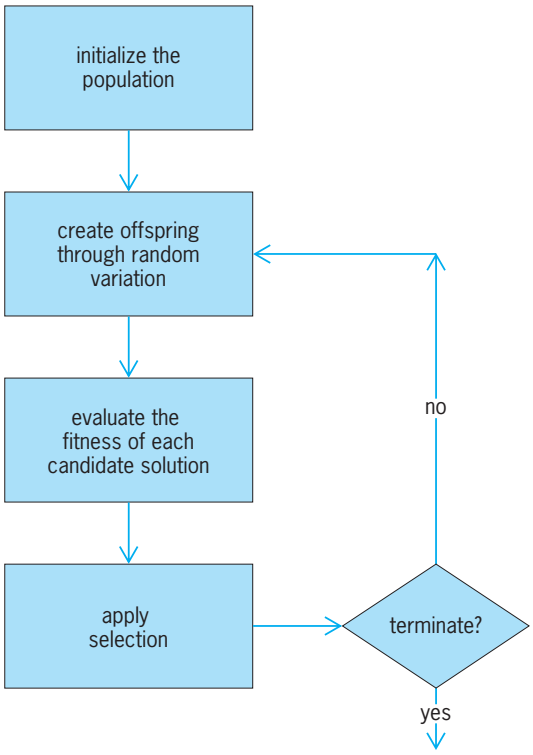

Fig. 1. Typical flowchart for an evolutionary algorithm.

regarding which solutions may be better than others, these hints can be used to bias the initial population and may accelerate the evolutionary optimization procedure.

An evolutionary algorithm (**Fig. 1**) is executed over a series of generations of random variation and selection. The variation to the existing solutions can come in the form of single-parent or multiparent operators. Alternative choices offer different sampling distributions from the space of all possible solutions. Each of the individuals in the population is scored with respect to how well the individual accomplishes the task at hand (fitness), and selection is used to eliminate some subset of the population or to amplify the percentage of above-average solutions. The algorithm terminates when some extrinsic criterion has been satisfied, such as prescribed maximum number of generations, or a suitable error tolerance. Over time, by eliminating poor solutions and extending the evolutionary search around those that appear better, the population can discover superior solutions to complex problems.

Origins. Although this idea of using evolutionary computation to solve problems is currently undergoing a renaissance that promises to have a significant impact on computing, its origins date back to the 1950s and 1960s. Some of the first experiments performed on John von Neumann's computer involved what would now be called artificial life. A program written for this machine by Nils Barricelli in 1953 simulated an environment, separated by cells in a grid. Numbers resided in each cell and migrated to neighboring cells based on a set of rules. When two numbers collided in the same cell, they competed for survival. Even with very simple rules for propagating throughout the environment, certain numeric patterns would evolve and could persist only when other patterns were also present. These first experiments already demonstrated something akin to symbiosis.

The idea of simulating evolution on a computer arose as many as eight to ten times independently from 1953 to 1969 in diverse fields of study. Evolutionary algorithms were proposed in 1957 and again in the mid-1960s to simulate biological processes with the aim of understanding adaptation in nature. Another proposal (1960) was to use Darwin's ideas to generate artificial intelligence by evolving models that would predict future events in an environment. The use of evolutionary algorithms was also proposed (1962) for function optimization, and this led to work on the prospects of evolving neural networks. And in 1964, principles of evolution were used to search for improved designs for physical devices in fluid-mechanics problems (such as a supersonic flashing nozzle). The multiple simultaneous beginnings of evolutionary computation attest to the compelling nature of the idea of simulating the fundamental processes of evolution in fast time.

Putting evolution to work. As a stochastic search technique, an evolutionary algorithm is a means for generating useful solutions rather than perfect solutions. For example, a traveling salesperson problem comprising 1000 cities involved finding the minimum tour length required to cover each city once and only once and return to the starting position. The initial best solution required over 50,000 units to accomplish this tour, while the best found solution, after examination of 4×10^7 possible solutions, had a distance of just over 2000 units (**Fig. 2***a*). This solution was not the best possible solution for this problem, but it was within about 5–7% of the expected minimum length for all 1000-city traveling salesperson problems. Moreover, the improvement in the quality of the solution was quite rapid in the initial phases of the evolutionary search (Fig. 2*b*). The size of the search space was over 10^{1200} (1 followed by 1200 zeros), indicating that the evolutionary algorithm had searched only a small fraction of all possible solutions while finding one of very high quality. Real-world applications of the traveling-salesperson problem include scheduling, the optimization of distribution networks, and computer chip layout and design.

Adaptability of technique. In comparison to traditional problem-solving techniques, evolutionary algorithms are often much faster and more adaptable to changes in the environment because whatever knowledge has been learned about how to solve a problem is contained in the collection of individual solutions that has survived up to that point. In contrast with, say, dynamic programming, an evolutionary algorithm need not be restarted when facing new data. This feature promises to have a significant potential in future problem solving. As the spread of data and processing speeds continue to increase,

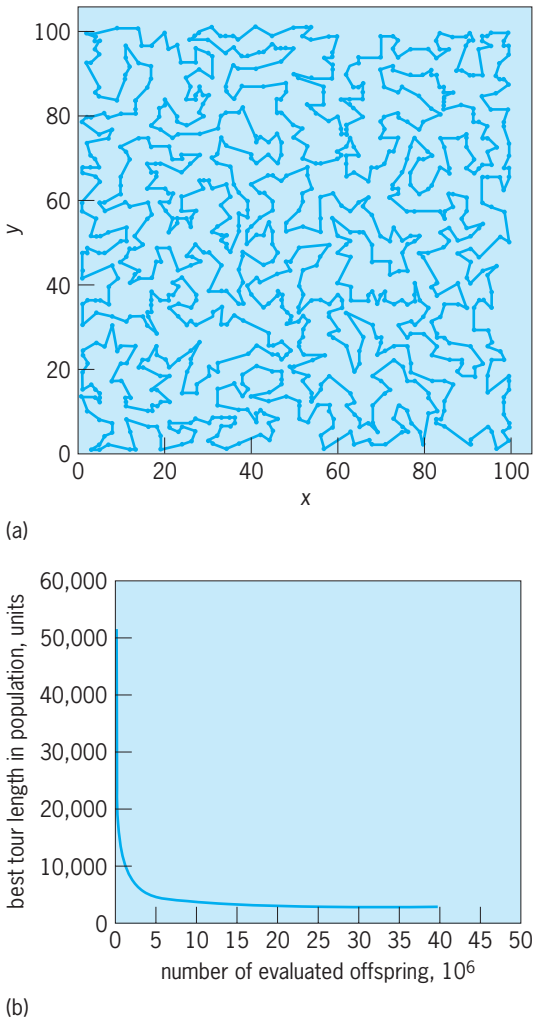

Fig. 2. Traveling salesperson problem comprising 1000 cities. (*a*) Best-evolved solution after examination of 4×10^7 possible solutions. (*b*) Rate of evolutionary optimization. (*After D. B. Fogel, Applying evolutionary programming to selected traveling salesman problems, Cybernet. Sys., 24:27–36, 1993*)

it will be necessary to make decisions ever more quickly. The evolutionary approach of bootstrapping off the current basis set of knowledge may eventually become the standard approach to real-world, real-time problem solving.

Contrast with conventional approaches. The utility and applicability of evolutionary algorithms requires a shift in thinking that may require some time to overtake conventional engineering. The typical approach to problem solving is to think in terms of reducing complex situations into simpler elements (building blocks) and then working to optimize each of these components. The hope is that when brought back together, these optimized components will serve to make the best design. But it is well known that outstanding individuals do not necessarily make a great team. Evolution works in an entirely different manner. Only the entire cohesive functional unit is measured by selection. Owing to the complex effects of genetic interactions, where a single genetic change can affect myriad behaviorial traits (a relationship known as pleiotropy), evolution tinkers with the underlying genetics and simultaneously generates a diverse array of new behaviors. Selection then culls out the least appropriate designs.

Choice of evolutionary algorithms. Even though there have been some attempts to analyze evolutionary algorithms as a parallel to human engineering, focusing mainly on recombination as a mechanism for bringing together "good ideas," there is no reason to expect that the evolutionary model of learning is the model that is used by engineers. Moreover, there is little reason to expect that it will be possible to improve evolutionary algorithms by capturing idealized models of genetic operators that occur in nature. The effects of those operators may be wholly different in the natural setting. A flapping wing will have different effects on a seagull in contrast to a jet aircraft. Airplanes can be designed to have increased biological fidelity by gluing feathers to their wings, but in the end all that will be accomplished is to increase their drag. What is termed the phenotypic effect, the manner of behavioral response to the environmental demands, is all that selection can assess. By consequence, in evolutionary algorithms, variation operators must be tailored to the task at hand in order to gain the most benefit and achieve the greatest efficiency.

This notion has been written down in formal mathematics in the "no free lunch theorem." Within some overarching assumptions, all algorithms perform exactly the same on average when applied across all possible functions. The consequence of this result is that if an evolutionary algorithm is good for a certain class of problem, then it is not good (worse than a random search) on some other class. There is no single best construction for an evolutionary search (or any other search algorithm). Much of the early theory that supported using binary representations, one-point crossover, and proportional selection has now been shown to be either incorrect or of no practical value. What remains is a significant challenge for the future: Analyzing classes of optimization problems and determining useful means for identifying which types of evolutionary algorithms are appropriate in each case. This is a wide-open area of investigation and it promises to be quite difficult.

Prospects. Computing power has now caught up with the concept of evolving solutions to complex problems. Current real-world applications include supply chain optimization, screening for new drug leads, electronic circuit design, scheduling logistic and transportation services, and gaming.

The processing speed of modern desktop machines is rated at about 10^9 floating-point operations per second (flops). But evolution is an inherently parallel process in which potential solutions can often be evaluated simultaneously rather than sequentially. In these cases, a parallel distributed computing architecture can offer a tremendous speed-up, since many more trials can be conducted per unit of time. An evolutionary algorithm can be implemented with multiple populations on independent processors,

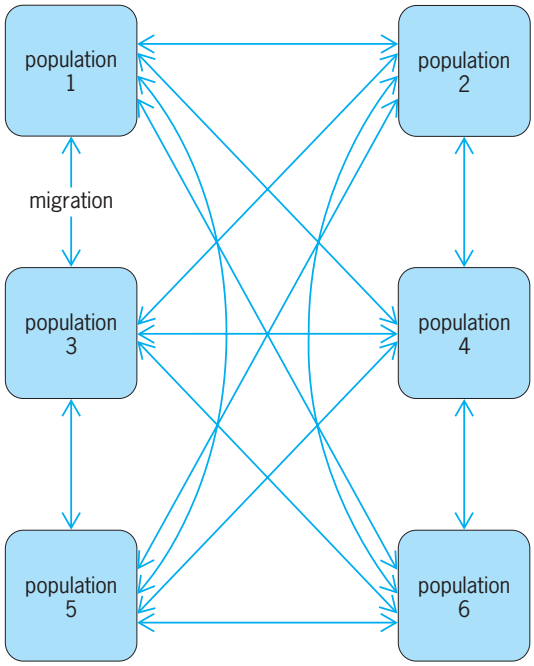

Fig. 3. Arrangement for distributing an evolutionary algorithm across multiple processors. Each processor holds its own population that undergoes random variation and selection. Occasionally, solutions will migrate from one population to another and begin competing for survival in their new environment.

with occasional migration being applied to introduce solutions from different populations (**Fig. 3**). In this manner, the parallel nature of evolution can be captured and many more trials can be conducted per unit of time. The best manner in which to set up such a distributed network of populations is problem-dependent and is an open question for research.

Experiments are being conducted with distributed architectures that comprise 1000 350-MHz computers. This computing power is roughly what can be expected for desktop machines around 2010. The products that will emerge from these massively parallel designs remain a matter of speculation; some foresee the use of evolutionary algorithms to design surrogate artificial brains that will supplant human cognition. Whether or not this conjecture becomes reality, with the price of desktop computers continuing to fall, the parallel approach to supercomputing is increasingly attractive and appears to be perfectly suited to evolutionary problem solving. *See* ALGORITHM; CONCURRENT PROCESSING; DISTRIBUTED SYSTEMS (COMPUTERS); GENETIC ALGORITHMS; MULTIPROCESSING; OPTIMIZATION; ORGANIC EVOLUTION; SUPERCOMPUTER. David B. Fogel

Bibliography. T. Bäck, *Evolutionary Algorithms in Theory and Practice*, Oxford University Press, New York, 1997; D. B. Fogel (ed.), *Evolutionary Computation: The Fossil Record*, IEEE Press, Piscataway, NJ, 1998; Z. Michalewicz and D. B. Fogel, *How To Solve It: Modern Heuristics*, Springer-Verlag, Berlin, 2000.

Exchange interaction

A consequence of the quantum mechanics of interacting identical particles that is responsible for magnetic interactions between electrons and atoms in solids, and magnetic order such as ferromagnetism and antiferromagnetism. The exchange interaction is actually not one of the fundamental interactions such as electromagnetism, but an effect that arises from the interplay of electromagnetism with quantum mechanics. *See* ELECTROMAGNETISM.

Quantum mechanics describes the state of a system of particles by a wave function that depends both on the positions of the particles and on any additional internal degrees of freedom needed to describe them (such as the spin of an electron). A fundamental axiom of quantum mechanics states that elementary particles fall into two classes, fermions and bosons. If two identical particles are interchanged, the wave function describing the state of the system is unchanged if they are bosons, and merely changes sign if (like electrons) they are fermions. *See* ELECTRON; ELEMENTARY PARTICLE; QUANTUM STATISTICS; SPIN (QUANTUM MECHANICS).

This antisymmetry of the wave function describing a many-electron state means that if the spins of two electrons are parallel the wave function must vanish when their positions coincide, making such electrons less likely to be found close together than a pair with antiparallel spins. This effect lowers the mutual electrostatic energy of electrons with parallel spins. Since electrons have a magnetic moment parallel to their spin, this effect helps explain the occurrence of magnetic moments in transition-metal and rare-earth atoms with a partially filled electronic shell (intraatomic exchange). This mechanism is used to explain the origin of ferromagnetism in metals such as iron. *See* ATOMIC STRUCTURE AND SPECTRA; ELECTRON CONFIGURATION; ELECTRON SPIN.

Similar considerations hold in the case of interatomic exchange of electrons between two such magnetic atoms in close contact, as in a magnetic solid. However, whether ferromagnetic (parallel) or antiferromagnetic (antiparallel) alignment of the two magnetic moments is favored now depends on the details of the most likely path by which exchange of electrons by quantum-mechanical tunneling between the atoms occurs. This path depends on the structure of the atoms and the solid, and may involve processes on nonmagnetic atoms in between the magnetic atoms. Both ferromagnetic and antiferromagnetic exchange interactions occur, though the latter are more common. Such interatomic exchange processes are used to explain the origin of magnetic order in nonmetallic magnetic materials. *See* TUNNELING IN SOLIDS.

Magnetic atoms also have a direct magnetic dipole interaction (like that between bar magnets) which can be understood purely in terms of classical electromagnetism, without invoking quantum mechanics. However, these interactions between neighboring atoms are much weaker than the exchange interaction and are not responsible for the

occurrence of magnetic order. Since the exchange interaction involves a quantum-mechanical tunneling process, its strength falls off rapidly (exponentially) with distance between atoms, and the long-range dipole interactions dominate at greater distances, controlling the large-scale magnetic-domain structure of ferromagnets. *See* ANTIFERROMAGNETISM; DIPOLE-DIPOLE INTERACTION; FERROMAGNETISM; QUANTUM MECHANICS.

F. Duncan; M. Haldane

Bibliography. D. C. Mattis, *Theory of Magnetism* vol. 1, 1981, reprint 1988; G. T. Rado and H. Suhl (eds.), *Magnetism*, vol. 1, 1963, vol. 4, 1966; R. M. White, *Quantum Theory of Magnetism*, 2d ed., 1983.

Excitation

Any of a number of different phenomena which (a) alter a system in some way or (b) result in some type of response. In electrical network theory, an excitation initiates a response or a response sequence. In other areas of technology, an excitation establishes an altered condition which causes an apparatus or system to exhibit useful response capabilities.

In electrical network theory, the term excitation designates a time-varying independent voltage or current in an n-port system or network. The independent voltage or current—which is also referred to as the excitation or the excitation signal—causes a network response which affects the voltage or current at the dependent port. The nature of the network response, and therefore of the dependent voltage or current, derives from the characteristics of the network. Thus an excitation signal at the input of a bistable multivibrator causes the output to shift from one of the circuit's stable states to the other; the next excitation signal causes the output to return to its former state. *See* MULTIVIBRATOR; NETWORK THEORY.

In atomic physics, excitation means the addition of energy to an atom at ground state to produce an excited state. *See* ATOMIC STRUCTURE AND SPECTRA; EXCITATION POTENTIAL.

In other contexts, excitation means the application of energy to one portion of a system or apparatus in a manner that enables another portion to perform a specialized function. Excitation energy may differ from the output energy in source, form, level, or location. That is, an excitation produces a primary effect that is linked, through an intermediate physical phenomenon, to a dependent secondary effect (see **illus.**).

Some examples will make this relationship apparent: A dynamic loudspeaker uses an excitation current in a field coil to generate a magnetic field; only then can a second magnetic field, generated by an audio signal, actuate the voice cone and produce sound waves. The sound-producing portion of a motion picture projector uses a lamp to provide excitation illumination; as the light passes through the sound-track portion of the moving film, its intensity is

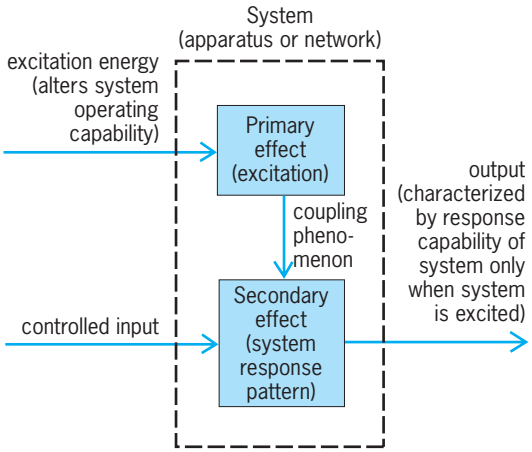

Diagram showing function of excitation.

modulated, producing audio-frequency fluctuations, which the photosensitive pickup converts into an electronic audio signal. *See* LOUDSPEAKER; OPTICAL RECORDING.

In some cases, the application of excitation energy is all that is necessary to generate the secondary response. A fluorescent lamp, for example, uses an excitatory current passing through an enclosed gas to generate ultraviolet radiation; when the phosphor coating inside the tube absorbs ultraviolet energy, it emits visible light. The human heart, to cite an example from biology, uses for excitation the decreasing electrical transmembrane potential of the pacemaker to generate an electrical impulse; when the surrounding striated muscle tissue receives this impulse, it transmits the electrical stimulus to other adjacent tissue and utilizes metabolic energy to contract, a response which produces the heartbeat. *See* CARDIOVASCULAR SYSTEM; FLUORESCENT LAMP.

William W. Snow

Excitation potential

The difference in potential between an excited atomic or molecular state and the ground state. The term is most generally used in connection with electron excitation, but it can be applied to excited molecular vibrational and rotational states.

A closely related term is excitation energy. If the unit of potential is taken as the volt and the unit of energy as the electronvolt, then the two are numerically equal. According to the Bohr theory, there is a relationship between the wavelength of the photon associated with the transition and the excitation energies of the two states. Thus the basic equation for the emission or absorption of energy is as shown below, where h is Planck's constant, c the

$$\frac{hc}{\lambda} = E_i - E_f$$

velocity of light, λ the wavelength of the photon, and E_f and E_i the energies of the final and initial states, respectively. If the final state is the ground state, then the difference in energies is just the excitation energy of the initial state. If neither of the two states

is the ground state, then the difference is numerically equal to the difference in excitation energies of the two states. This suggests that the excitation energy for many states may be determined spectroscopically. In fact, the careful measurement of the wavelengths associated with the transitions, along with the identification of the levels involved, permits the assignment of an excitation potential to each available energy state. This may be accomplished by either emission or absorption spectra. *See* ATOMIC STRUCTURE AND SPECTRA; ELECTRONVOLT.

Measurements have also been made by using electron impact means. Here electrons from a hot filament are accelerated by a grid and collected on an outer electrode after passing through an intervening space where they can interact with a gas of molecular or atomic species. The potentials are so adjusted that the electrons can reach the collector only if they lose no energy after passing the grid. As the accelerating grid potential is increased, a series of sharp drops is noted in the collector current. Each of these is interpreted as representing a case where the electrons have obtained just enough energy to produce excitation, and hence no longer have the energy needed to climb the potential barrier to the collector. The evidence thus obtained is very direct, but the accuracy is not as good as that for spectroscopic data. However, this is the method that was used in the historic Franck-Hertz experiment, which demonstrates the quantization of atomic energy levels. *See* GROUND STATE; IONIZATION POTENTIAL. Glenn H. Miller

Bibliography. S. Gasiorowicz, *Structure of Matter: A Survey of Modern Physics*, 1979; I. I. Sobel'man, L. A. Vainshtein, and E. A. Yukov, *Excitation of Atoms and Broadening of Spectral Lines*, 2d ed., 1995.

Excited state

In quantum mechanics, a stationary state of higher energy than the lowest stationary state or ground state of a particle or a system of particles. Customarily, only bound stationary states, which generally are at most denumerably infinite in number, are spoken of as excited, although the formal quantum theory often treats the noncountable unbound stationary states on an equal footing with the bound states. Conventionally, the excited states are ranked in order of increasing energy; that is, the second excited state has higher energy than the first, which lies higher than the zeroth or ground state.

For actual physical systems it is usually, but by no means necessarily, the case that energy levels of high degeneracy (that is, energy levels associated with a comparatively large number of stationary states of the same energy) are encountered in excited states rather than in ground states. Illustrative systems possessing highly degenerate ground states include the 10-fold degenerate germanium-72 nucleus (32 protons, 41 neutrons) and the 13-fold degenerate uranium atom (92 electrons in the field of a massive positive point charge equal to 92 e, where e is Sobel'man charge). *See* DEGENERACY (QUANTUM MECHANICS); GROUND STATE; METASTABLE STATE; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS. Edward Gerjuoy

Bibliography. S. Gasiorowicz, *Quantum Physics*, 2d ed., 1995; R. B. Leighton, *Principles of Modern Physics*, 1959.

Exciton

An excited electronic state in a large molecule, or a semiconducting or insulating solid composed of a negatively charged electron and a positively charged hole bound together by their electrical attraction. Its creation through internal charge separation is most frequently caused by the absorption of light, and its demise is occasioned by the emission of light. The exciton is thus an important portal for the study of the properties of an insulator by optical means and for applications such as optoelectronics. Its external neutrality enables it to move through the system as an energy carrier, an important agent in energy transfer processes of special importance in biological functions such as photosynthesis. When the excited electron and hole are restricted to a plane, a wire, or a microscopic box in an artificially fabricated structure known, respectively, as a quantum well, quantum wire, or quantum dot, a new dimension is added to the ability to explore fundamental nonlinear and nonequilibrium physics and coherent quantum optics as well as quantum control of microscopic objects. *See* ELECTRON; HOLE STATES IN SOLIDS; PHOTOCHEMISTRY; PHOTOSYNTHESIS; QUANTIZED ELECTRONIC STRUCTURE (QUEST).

Creation and recombination. The many-electron ground state of the insulator is immune to excitation until the excitation energy reaches a threshold known as the energy gap. When external influences such as the electromagnetic field and lattice vibrations are ignored, the exciton may be viewed as a robust state of an excited electron and the hole which it has left behind in the sea of electrons. The hole acquires its positive charge from the loss of a charge from the sea whose total charge is neutralized by that of the ions in the molecule or solid. *See* BAND THEORY OF SOLIDS; ELECTRIC INSULATOR.

The photon–exciton interaction is responsible for the optical excitation (though not necessarily in the visible frequency range) of the exciton and for its spontaneous recombination emitting a photon (a quantum unit of light). The dipole matrix element responsible for the transition between the energy states is strong when the electron and hole wave functions overlap in space or match in wave vector. From Planck's law in quantum theory, the frequency of the emitting light is proportional to the energy loss in returning the exciton state back to the ground state. *See* ELECTRON-HOLE RECOMBINATION; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Excitons in semiconductors. A semiconductor is distinguished from an insulator qualitatively by a smaller

energy gap, with the frequency of the emitting light from the exciton spanning the range from visible light to very far infrared. Because of the higher mobility of their excited electrons and holes, semiconductors are of supreme technological importance. For example, the electrons and holes can be electrically injected to form excitons whose recombination leading to light emission enables the conversion of electrical energy to light in such devices as light-emitting diodes and lasers. *See* LASER; LIGHT-EMITTING DIODE; SEMICONDUCTOR.

In addition to the direct "resonant" formation by the absorption of a photon (in which the photon energy equals the exciton formation energy), the exciton is commonly formed by the photoluminescence process during which the optically excited electron and hole with energy much higher than the exciton energy relax to the lowest states to form an exciton. The recombination of the exciton emitting a photon then completes the luminescence process. *See* LUMINESCENCE.

Indirect excitons. In an indirect-gap semiconductor such as silicon, where the momentum of the exciton does not match that of the photon, the excitons are generally formed after relaxation of optical excitations with initial energy much higher than the gap. The indirect exciton has a long lifetime because its recombination with the emission of a photon requires the conservation of momentum to be satisfied by the assistance of a lattice vibration or trapping by a defect. Consequently, the excitons have time to form a large pool known as an electron-hole drop.

In a direct-gap material, the spatial separation of the electron and hole can be enforced by housing them in two layers sufficiently close to maintain their electric attraction. The recombination of such indirect excitons may then be controlled by changing the separation with an electric field. An interesting phenomenon is the laser spot excitation of these indirect excitons, leading to their migration away from the laser spot as optically inactive excitons for long distances, some radiating in a bright spot when trapped by a defect, and others surviving the traps to form two concentric luminous circles centered at the laser spot. The outer one is a necklace of evenly spaced bright spots, thought to be a lattice of vortices. Since L. V. Keldysh predicted the condensation of excitons as bosons, the observation of exciton condensation has become a highly prized objective of research in the field. This system of excitons in a double quantum well is considered a good candidate for condensation at low temperature even though the necklace has a noncondensation origin. *See* BOSE-EINSTEIN CONDENSATION; TRAPS IN SOLIDS.

Frenkel and Wannier excitons. If the constituent electron and hole of the exciton are mostly localized at an ion, the exciton is localized, but with some probability to hop from site to site. Such a Frenkel exciton is common in molecules and molecular solids. At the other extreme, if the electron and hole wave functions are widespread as extended orbitals in a molecule or Bloch waves in a crystal, their bound state as the exciton can have their center of mass moving through the system with ease. Such Wannier (or Wannier-Mott) excitons are most common in broad-band and small-gap semiconductors. They resemble the hydrogen atom or, more closely, the positronium system composed of an electron and a positron. Because of the dielectric screening of the electrical force in small-gap materials and sometimes the small effective mass of the electron, the Wannier exciton radius is from 10 to 100 times larger than the positronium radius of about 0.1 nm. The absorption spectrum of excitons also resembles that of a hydrogen atom, with a series of lines with rapidly decreasing spacing finally blurring into a continuum whose energy edge is theoretically the energy gap but practically difficult to discern. Such optical spectra are usually obtained at the temperature of liquid helium (around 2 to 4 K or −456 to −452°F) to avoid the broadening of spectral lines by lattice vibrations. *See* ATOMIC STRUCTURE AND SPECTRA; BLOCH THEOREM; POSITRONIUM.

Exciton polaritons. Because of the interaction between electrons and light, not only can an exciton decay irreversibly into a photon or vice versa, but it can also exchange roles with the photon in a quantum-mechanically coherent fashion. Thus, the exciton may exist in the solid as a quantum superposition of an exciton and a photon, known as a polarition. Since the photon energy varies linearly with its momentum at the speed of light, the exciton at most values of its center-of-mass momentum has much too low an energy to mix with the photon. However, at low momentum the exciton and the photon can approximately match both their momentum and energy values, the coupling mixes the two states into two superpositions of photon and exciton with an energy splitting. Thus, the massless photon is slowed down by the massive exciton by the quantum-mechanical superposition—a version of slow light. *See* PHOTON.

Microcavities. Confinement of light in a microcavity enhances the coupling of the exciton with the confined light mode. It was discovered that two-dimensional exciton polaritons, when pumped resonantly and coherently, can be scattered into a low-energy mode where they accumulate to a macroscopic (Avogadro) number, causing a high-intensity luminescence. This phenomenon has been interpreted as an optically maintained boson condensation state. Experiments have shown that when quantum dots are planted in a nanocavity, their excitons couple strongly with cavity photons. This strong coupling is a promising basis for cavity quantum electrodynamics in a semiconductor for the proposed quantum control of excitons via photons. *See* NANOSTRUCTURE; QUANTUM ELECTRODYNAMICS.

Biexcitons. For the same reason that two hydrogen atoms can bind into a molecule, two excitons can bind into a biexciton. A biexciton takes two photons to create and can spontaneously emit two photons, usually one at a time. Nonlinear response of a semiconductor to three electric fields may be viewed as numerous replicas of the event with three photons,

some incoming and some outgoing, each supplied by an electric field. The biexciton created in the process plays a central role. Resonance of field frequencies with exciton or biexciton energy enhances the response. Not only has nonlinear optics flourished with the time-resolved measurement of dynamics, first at the picosecond and then at the femtosecond scale, but also the coherence of the laser light has been matched with the quantum-mechanical coherence of the exciton, giving rise to the study of the coherent optics of quantum properties in the solid-state system. The coherence of the laser light derives from the phase relation between different parts of the electromagnetic wave and the coherence of the exciton from the superposition of its constituent states. The delicate four-wave mixing has yielded information that extends beyond the structure of the excitations of the excitons to their dynamics far from thermal equilibrium. The exploration of the exciton dynamics of the first few femtoseconds after ultrafast intense laser excitation has contributed profoundly to our knowledge of nonequilibrium physics. *See* COHERENCE; NONLINEAR OPTICS.

Trions. Excitons can also form complexes. When an exciton is tethered to a positively charged impurity ion, the electron system in a donor is a bound state of two electrons and a hole, known as a negatively charged trion. Since the electron-hole pair in the trion may be coupled to light, light may be used to control the dynamics of an electron through the trion in an electronic Raman process in which no phonons need be involved. *See* DONOR ATOM; PHONON; RAMAN EFFECT.

Exciton spin and light polarization. The exciton possesses the quintessential quantum-mechanical property of an intrinsic angular momentum, known as spin, of one fundamental quantum unit $\hbar$ (Planck's constant divided by 2π), which matches the intrinsic angular momentum of the photon perfectly. Now the plus-one- and minus-one-quantum-of-angular-momentum states of the photon are the building blocks, respectively, of right- and left-handed circularly polarized light. Angular momentum conservation in the photon–exciton conversion process means that it is possible to "polarize" excitons, all with their spins pointing along (or opposite) to the light propagation direction by excitation with right- (or left-) circularly polarized light. The electron constituent of the exciton has spin one-half quantum unit, which points in the opposite direction to the optically generated exciton. (Angular momentum conservation is maintained by the hole contribution of 3/2 units from its orbital and spin angular momenta and no orbital angular momentum from the electron.) Thus, the optical excitation of excitons by circularly polarized light provides a good source of spin-polarized electrons (that is, with spin pointing mostly in one direction) when the electron and hole constituents of an exciton are separated by an electric field, for example, in a *pn* diode. Time-resolved photoluminescence by polarized light is an excellent probe of the spin dynamics of the excitons. *See* ANGULAR MOMENTUM; ELECTRON SPIN; POLARIZATION OF WAVES; POLARIZED LIGHT; SPIN (QUANTUM MECHANICS).

Quantum wells. Forcing the electron and hole together by confinement increases their attraction and makes the exciton binding stronger. The exciton binding energy is the separation of its energy from the energy continuum of the unbound electron and hole pair. With the advent of molecular-beam epitaxy, it is possible to build a sandwich of semiconductors of different band gaps so that the electron and hole can be effectively confined in a quantum well. The exciton whose constituents are confined to move effectively in a plane can have as much as four times the binding energy of its counterpart in three dimensions. This increase is sufficient to make the exciton line visible in the optical spectrum at room temperature, whereas the exciton line in three dimensions would be indistinguishable from the continuum due to thermal broadening. For the same reason, the biexciton also has a larger binding energy in quantum wells. The study of exciton dynamics and spin dynamics flourishes in quantum wells and other confined structures. The exciton in the quantum well is also more susceptible to the influence of an external electric field. The modification of the exciton spectrum by an electric field has been applied to self-electrooptic-effect devices. *See* ARTIFICIALLY LAYERED STRUCTURES; ELECTROOPTICS; SEMICONDUCTOR HETEROSTRUCTURES.

Quantum dots. A semiconductor quantum dot of submicrometer dimensions serves as a trap of a single electron or an exciton. It behaves like a giant atom with the advantages of easier electrical and magnetic control because of the lower energy spacings and larger dipole transition matrix elements than real atoms. However, it also has the concomitant disadvantage of shorter radiative recombination time and decoherence time. It is possible to study the dots in ensemble or singly. In the method of microluminescence, by suitable masking to expose only a small area to light, it is possible to obtain the spectra of excitons, biexcitons, and trions in a single quantum dot from its luminescence. (The trion is generated by exciting an exciton in a dot doped with a single electron.) Similarly, coherent nonlinear spectroscopy can be performed by optically pumping and probing a single dot. This spectroscopy opens up the possibility of quantum control of a microscopic state such as the exciton. Just like coherent quantum optics in atoms, the coherent driving of the Rabi rotation between the ground state and the exciton has been demonstrated by several groups. The Rabi rotation is a change of the relative proportion of two states in a quantum superposition by a coherent electromagnetic pulse. This rotation stops when the pulse stops, a deterministic event to be distinguished from the probabilistic excitation from one state to another that is governed by the Fermi golden rule when the different frequency components of the electromagnetic field are not phase-coherent.

Quantum control. The demonstration of Rabi rotation is the first step toward operations on a quantum unit of information (qubit). The biexciton in a dot is

bound in the sense that it has a lower energy than the sum of energies of its constituent excitons when they have antiparallel spins. The biexciton can be coherently formed by driving the dot with two laser pulses with opposite circular polarizations. One might then imagine the absence and presence of each exciton as a qubit, thus forming a two-qubit system in the dot. The difference in energy between adding an exciton of spin up, say, to the ground state or to the single-exciton state of the opposite spin direction is key to the ability to drive the Rabi rotation between two double-digit states (0,0) and (1,0) or between (0,1) and (1,1) with light of the same polarization but a choice of different frequencies. Either rotation is a conditional dynamics (control of one exciton depending on the second exciton being absent or present). That is sufficient for the entanglement of the two excitons, which is the fundamental resource for quantum information processing. Such processing has indeed been demonstrated experimentally in a quantum dot at 4 K (−452°F). It is remarkable that it could be done in a dot that was housed in a solid-state system full of potentially disruptive decoherence sources, quite unlike the trapped-ion system in which such a feat had been demonstrated earlier. Housing two exciton qubits in a dot is not a model on which one can build a similarly controllable system of an arbitrary number of qubits. This two-qubit model is said to be "unscalable," and thus not directly useful for the purpose of building a practical quantum computer. However, the demonstration of the optical control of an exciton in a single dot makes it possible to imagine the optical control (via Raman processes) of an electron in a dot whose two spin states constitute the qubit and of two electrons in separate dots, scalable to a large system. Thus, the exploration of quantum control of microscopic objects, now being electron spins residing in separate dots, can continue. *See* QUANTUM COMPUTATION.

Lu Jeu Sham

Bibliography. R. S. Knox, *Theory of Excitons* (*Solid State Physics, Advances in Research and Applications, Supplement* 5), Academic Press, New York, 1963; E. I. Rashba and M. D. Sturge (eds.), *Excitons*, North-Holland, Amsterdam, 1982; W. Schäfer and M. Wegener, *Semiconductor Optics and Transport Phenomena*, Springer, Berlin, 2002; J. Shah, *Ultrafast Spectroscopy of Semiconductors and Semiconductor Nanostructures*, 2d ed., Springer, Berlin, 1999; P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors*, 3d ed., Springer, Berlin, 2004.

Exclusion principle

No two electrons may simultaneously occupy the same quantum state. This principle, often called the Pauli principle, was first formulated by Wolfgang Pauli in 1925 and, for time-independent quantum states, it means that no two electrons may be described by state functions which are characterized by exactly the same quantum numbers. In addition to electrons, all known particles having half-integer intrinsic angular momentum, or spin, obey the exclusion principle. It plays a central role in the understanding of many diverse phenomena, including the periodic table of the elements and their chemical activities, the electron contribution to the specific heat of metals, the shell structure in the atomic nucleus analogous to that of electrons in atoms, and certain symmetries in the scattering of identical particles. *See* ANGULAR MOMENTUM; NONRELATIVISTIC QUANTUM THEORY; QUANTUM NUMBERS; SPIN (QUANTUM MECHANICS).

Symmetry of wave functions. When a system of identical particles is described by a wave function, the indistinguishability of the particles implies that the wave function must have certain symmetry properties when the coordinates of any two of the particles are interchanged. Specifically, the wave function either remains unchanged or is changed only in sign when such a coordinate interchange is made. For the two-particle system, Eq. (1) holds, where $\mathbf{x}_i$ denotes

$$\psi(\mathbf{x}_1\mathbf{x}_2) = \pm\psi(\mathbf{x}_2\mathbf{x}_1) \tag{1}$$

all the coordinates of the particle i, such as space coordinates, spin, and any others necessary, and ψ is the wave function.

When the plus sign applies, the wave function is said to be symmetric; the particles then are termed bosons and obey Bose-Einstein statistics. When the minus sign applies, the wave function is antisymmetric under particle exchange; the particles then are termed fermions and obey Fermi-Dirac statistics. Bosons are particles with integer spin, while fermions have half-integer spin, measured in units of Planck's constant. This connection between spin and statistical laws obeyed by the particles may be proved mathematically under appropriate assumptions. *See* BOSE-EINSTEIN STATISTICS; FERMI-DIRAC STATISTICS.

The Pauli principle follows from the symmetry of the many-fermion wave function in the special case when this wave function may be written as a product of single-particle wave functions. For example, consider two identical fermions and two single-particle states. Let $\psi_\alpha(1)$ denote particle 1 in state α, $\psi_\beta(2)$ particle 2 in state β, and so on, where α and β represent all quantum numbers labeling the states. Then a possible state of the two-particle system is shown by Eq. (2), which obviously changes sign if 1 and

$$\psi(1, 2) = \psi_\alpha(1)\psi_\beta(2) - \psi_\alpha(2)\psi_\beta(1) \tag{2}$$

2 are interchanged. If $\alpha = \beta$, then ψ is identically zero; that is, no such state exists which is exactly the statement of the Pauli principle. Note that ψ may be expressed as a determinant, as in Eq. (3).

$$\psi = \begin{vmatrix} \psi_\alpha(1) & \psi_\beta(1) \\ \psi_\alpha(2) & \psi_\beta(2) \end{vmatrix} \tag{3}$$

For more than two particles, one may still form product wave functions which are also determinants. The exchange of a pair of particles is equivalent to the interchange of two rows of the determinant which is known to change its sign. Furthermore,

if any two of the column labels, the α and β above, are the same, the determinant vanishes identically, which again is the Pauli principle. The most general wave function is a linear combination of all possible determinant wave functions, and the antisymmetry still holds, but the Pauli principle does not, since it applies only if there is a one-to-one correspondence between the number of single-particle states and the number of particles.

Spin. In the example above a single function was used to denote a single particle state. In many physical applications it is more convenient to express the wave function of a single particle as a product of a function which describes its spatial properties and one which describes its spin orientation. For spin s there are $2s + 1$ possible spin orientations ranging from $-s$ to s in integer steps. Electrons, protons, and neutrons all have $s = {}^1/_2$. The wave function for a collection of identical particles may then also be expressed as the product of a spatial wave function for the collection times a spin wave function for the collection. For the case of two particles, the Pauli principle dictates that, if the spatial wave function is symmetric, the spin function must be antisymmetric and vice versa. For more than two identical fermions, mixed symmetries in each part may occur, but they must be complementary in the sense that the overall wave function obeys the Pauli principle.

Isotopic spin. It is often convenient to regard protons and neutrons as simply two different states of the same fermion, termed a nucleon. This approach takes into account that the proton is charged and the neutron is not, but ignores the small mass excess of the neutron. The additional quantum number needed to describe the two possible charge states is called isobaric spin. In this case the wave function of a single nucleon may be taken to be the product of a spatial part, an ordinary spin-orientation part, and a part to describe its charge state, that is, isotopic spin. The ideas about the symmetry apply similarly for a collection of nucleons. In this case one speaks of the generalized Pauli principle, in that the overall wave function is antisymmetric. For two nucleons this is attained by choosing the symmetries of the parts of the wave function as indicated in the **table**, where + denotes symmetric and − antisymmetric. Again, for more than two nucleons mixed symmetries occur, but the overall wave function obeys the generalized Pauli principle. *See* I-SPIN; SUPERMULTIPLET.

Atomic structure. One important consequence of the exclusion principle is that, if there are N different single-particle states corresponding to a given energy level, then, at most, N identical fermions may have that same energy. For example, in the atomic case the single-electron states have allowed energies labeled by the quantum number $n = 1, 2, 3, \ldots$ corresponding to the lowest energy, next lowest, and so on. For a given value of n, the orbital angular momentum l may have any one of the values $l = 0, 1, 2, \ldots, n - 1$. For each l value there are $2l + 1$ states labeled by m_l, which range in integer steps from $-l$ to l, and specify the orbit orientation. In addition, there are the two possible spin orientations, $\pm{}^1/_2$. For $n = 1$, then, there are two different single-particle states, while for $n = 2$, there are eight. In general, $2n^2$ is the maximum number of electrons that may have an energy specified by n.

Using the fact that a system will try to occupy the state of lowest possible energy, the electron configuration of atoms may be understood by simply filling the single-particle energy levels according to the Pauli principle. This is the basis of Niels Bohr's explanation of the periodic table. For example, the one electron of hydrogen goes into the $n = 1$ level, as do both of the electrons of helium. The $n = 1$ level is said to be full. The next element in the periodic table is lithium, and its third electron must go into the $n = 2$ level. For neon, of atomic number 10, two of its 10 electrons go into the $n = 1$ level and eight into the $n = 2$ level, so that both are full. Now, the chemical activity of an element is determined mainly by its outermost (largest n) electrons, and a filled level is essentially inert. One sees therefore why helium and neon are noble gases and have very low chemical activity, whereas lithium, which has one more electron than the very stable helium configuration, is very active chemically. By extending these ideas, such phenomena as binding energies and valences may be understood. Even if the Pauli principle applied only to this one area, it would still be a very important contribution to physics. *See* ATOMIC STRUCTURE AND SPECTRA; VALENCE.

S. A. Williams

Bibliography. M. Chester, *Primer of Quantum Mechanics*, 1987, reprint 1992; R. Eisberg and R. Resnick, *Quantum Physics of Atoms, Molecules, Solids, Nuclei and Particles*, 2d ed., 1985; M. Jammer, *The Conceptual Development of Quantum Mechanics*, 1966, reprint 1989; D. A. Park, *Introduction to Quantum Theory*, 3d ed., 1992; R. L. Sproull and W. A. Phillips, *Modern Physics*, 3d ed., 1980, reprint 1989; R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics and All That*, 1964, reprint 1980.

Choice of symmetries for two nucleons

Space	Spin	Isotopic spin
+	+	−
+	−	+
−	+	+
−	−	−

Excretion

The removal of excess material from a cell or a living animal. Such a definition is so broad that it is necessary to set some limits to the discussion that follows. The first limitation will be to the animal kingdom since excretion appears not to be a major problem in the plant kingdom—partly because of the difference in physiology.

Carbon dioxide is one of the chief products of the metabolism of cells. In the dark, plant cells

consume oxygen and produce carbon dioxide, but in the light, the actively synthesizing plant cells require a source of carbon dioxide and produce oxygen in excess. Gaseous diffusion over relatively short distances meets these respiratory problems in plants and very small animals, but in larger animals special mechanisms have evolved for the transport of carbon dioxide in blood. The excretion of carbon dioxide might properly be discussed here, but it has been found more satisfactory in the science of physiology to treat carbon dioxide and oxygen together under the heading of blood-gas transport. This case is cited to exemplify the assignment of excretion of heat to body-temperature regulation and of salts to osmoregulation. *See* PHOTOSYNTHESIS; RESPIRATION; THERMOREGULATION.

Physiological Problems

Though animals exhibit a wide variety of excretory structures, depending upon their individual needs, the excretory mechanisms are employed to overcome certain basic physiological problems.

Water and salts. One of the central problems of excretory physiology is the handling of water. To understand the role of water, it is important to speculate briefly about the very origins of life. Current views assume life to have begun in the form of self-reproducing molecules which grew by incorporating the already formed smaller molecules from the surrounding medium into their ordered structure. But, as essential ingredients in this medium diminished in quantity, self-reproduction became dependent upon a supply of the necessary material, obtained perhaps by modification of a closely related molecule. A primitive metabolism can thus be considered to have evolved. It would become essential to slow the diffusion processes which would inevitably carry away so valuable a material. This could be accomplished by the development of a relatively impermeable membrane from the many molecules which would be absorbed onto the surface of a complex molecule of this sort. But when a semipermeable membrane is placed around a quantity of organic matter osmotic pressure develops, and the primitive cell faces its first problem of excretion, namely, getting rid of the water which osmosis continuously brings into it. One answer to this problem was that developed by plant cells. By surrounding the protoplasm with an extremely strong wall of cellulose, osmotic pressures of many atmospheres could be allowed to develop. The cell walls prevent the cell rupture that would otherwise follow. Whole plants take advantage of differences in osmotic pressure between external cells and internal cells to move water along the roots and stems. The turgor of cells, petioles, and leaves which develops as a result of osmotic pressure is also used to close stomata or to alter the position of a leaf. Development of cellulose cell walls has therefore been made useful in manifold ways. *See* OSMOREGULATORY MECHANISMS; PREBIOTIC ORGANIC SYNTHESIS.

The planet provides an array of habitats—major ones being marine, brackish water, fresh water, and terrestrial—and animals have established themselves successfully in each one. Particular osmotic problems are offered by each habitat, and the evolution of means of meeting these problems will be given brief consideration in phyletic sequence.

Nitrogen. Carbohydrate and fat are metabolized to carbon dioxide and water, and no other major excretory problems result from their use. Protein is in a different category. The amino acids used as an energy source must first be freed of nitrogen (deaminized), and the ammonia which is the first result of this process immediately constitutes a formidable excretory problem. Carnivorous animals of necessity have a diet high in protein, and omnivorous animals may prefer a high protein intake. One additional source of nitrogenous wastes, the nucleic acids, should be put in perspective. Each cell contains a small portion of this material in its nucleus so that purines may be present in the body in excess. An entire section below is devoted to the problem of nitrogen excretion. *See* NUCLEIC ACID; PROTEIN METABOLISM.

Other compounds. The kidneys of various animals have had to deal with compounds of considerable variety originating in plants taken as food. A surprising number of animal species employ all three of the mechanisms basic to urine production: filtration, reabsorption, and secretion.

Filtration of blood usually initiates the process and conserves blood cells, blood proteins, and fat droplets for the body. In the passage of the filtrate through tortuous channels, the substances most needed are recaptured by a secretory process termed reabsorption. Salts, amino acids, and carbohydrates fall into this category, and water may also be reabsorbed in quantity, thus concentrating the waste products left in the urine. The special compounds ingested but not used in the body may thus be excreted, but at this point the third mechanism, that of secretion of material from the blood into the urine, becomes pertinent. Some compounds reach a concentration in the urine far beyond that which would be possible by the combination of the first two processes.

Of major interest here is the ubiquity with which certain categories of substances are actively secreted. Phyla of animals separated by hundreds of millions of years of independent evolution have nevertheless retained the capacity for secreting the same compounds. The renal physiologist uses the phenolphthalein dye phenol red as a tool in exploring this field. This compound is actively secreted by almost every animal in which it has been tested. But it seems obvious that phenol red is not a compound found in nature. Its structure simply fits the transport units of an active pump retained to excrete other members of a large family of compounds. Hippuric acid, formed by the conjugation of the benzene rings of benzoic acid with glycine, might serve better as a model. This compound also fits the transport units which carry phenol red and is highly concentrated in the urine of many animals. The criticism of this compound, directed at some suggested

detoxification mechanisms, is that the products of the supposed detoxification are still somewhat toxic. The advantage may be simply the use of a compound whose concentration in the body is quickly lowered by the excretory pump. Benzoic acid, for example, is not as rapidly excreted as is hippuric acid.

Some of these compounds are excreted so efficiently that blood flowing through the kidneys is effectively cleared of the material. It appears that one mechanism may transport such diverse materials as phenol red, hippuric acid, penicillin, and diodrast. These compounds are but representatives of whole classes of compounds possessing the necessary structure to some degree.

Excretory Structures

Having listed the major excretory problems faced by animals, it will be profitable to proceed systematically through the animal kingdom to examine the structures that have developed to meet the problems and to follow their evolutionary sequence wherever possible.

Contractile vacuoles. In some plant gametes and in the protozoa and sponges, animals commonly placed at the bottom of the evolutionary array, there is a conspicuous mechanism for ridding the body of excess water and probably of wastes as well. These structures are called contractile vacuoles. Measurements show that some animals may have to put out a volume of water equal to the total body volume every 4 min. One view is that the water is secreted into channels in the protoplasm, through which it is gathered into the vacuole for extrusion. *See* PORIFERA; PROTOZOA.

It is possible that the contractile vacuole represents simply the interiorization of a phenomenon common to most cells. The reverse process, the taking up of water from a medium, has been watched and termed pinocytosis or cell-drinking. In such a case the surface membrane is very active. Drops of external fluid are trapped in infoldings of the membrane, which then leave the boundary of the cell to become vacuoles in the protoplasm. If the membrane is digested away, the droplet becomes a part of the protoplasm. It has been suggested that the secretion of water may be accomplished by a reverse of this process. The many internal membranes that are being revealed by the electron microscope surround droplets of water separated from the protoplasm. These vacuoles approach the surface membrane, incorporate into it, and discharge the droplet to the exterior. In the case of the contractile vacuole, the droplets discharge, instead, into the internal channels leading to the contractile vacuole because in these cellular forms it is important for the external pellicle to have considerable mechanical strength, whereas in the internal cells of multicellular animals the cell membrane can be a delicate and mobile one. *See* ENDOCYTOSIS.

Samples of fluid taken from the contractile vacuoles with micropipets show that, in the process of excreting water, salts are conserved by the cell; the excreted fluid is hypoosmotic to the protoplasm. *See* VACUOLE.

Coelenterata. This important group of metazoan animals is usually placed just above the sponges and is known as the phylum Coelenterata or Cnidaria. This phylum warrants mention here because, while the cells do not possess contractile vacuoles, no specific mechanism for excretion or osmotic regulation has been found. Members of the group live in marine, brackish, and fresh-water environments, and it may safely be forecast that the cellular adaptations that make this possible will someday be identified and studied. *See* CNIDARIA.

Nephridia. A most significant development in the evolution of excretory processes occurs in the phyla beyond the Protozoa, Porifera, and Coelenterata. Structures termed nephridia are present in the Platyhelminthes, Rotifera, Nemertea, Acanthocephala, Priapuloidea, Entoprocta, Gastrotricha, Kinorhyncha, Cephalochorda, and some Archiannelida and Polychaeta, as well as in larval stages of the Polychaeta, Archiannelida, Echiuroidea, Mollusca, Phoronidea, and Cephalochorda. An eminent authority, E. Goodrich (1945), said that these organs (protonephridia) are in fact widely distributed and may be inferred to have been present in the common ancestor of all the metazoan Triploblastica.

Aside from the fact of its presence in all the animals at this intermediate level of evolution, the nephridium represents for the process of excretion the specialization of tissues that enable animals to become larger and more independent of their immediate environment. The cells of the body no longer live in antagonism to seawater or fresh water, but are bathed by an internal medium which is regulated to represent an optimal environment. Numerous tissues have as their special function the regulation of the composition of this medium. It is the part of the nephridium to clear the body fluids of wastes, without at the same time losing from the body the substances needed by the cells. As in the case of the contractile vacuole, it is not understood clearly how the nephridium accomplishes these functions, but assumptions can be made about several of the processes involved. *See* URINARY SYSTEM.

Two particular sorts of nephridia are worth describing briefly, the protonephridium and the metanephridium. The protonephridium is the primitive structure consisting of a single cell, shaped like a flask, within the neck of which there is found a cylindrical cavity, or lumen. In this lumen are found one or more threadlike structures, or cilia, which beat continuously, the appearance thus giving rise to the old term flame cell. The flame cell appears to filter through itself a part of the animal's circulating fluid, but its polarity is different from that of the other cells of the body, and a "urine" passes from the lumen to the exterior by way of a nephridiopore.

The metanephridium is a structure for which the functions can be more nearly assigned. The tubular structure opens into a body or coelomic cavity. There is always fluid in this cavity, resulting from filtration through the cells making up its lining. This fluid

contains not only wastes but also materials of value to the animal, and as it is propelled down the nephridium by a lining of cilia, it is subjected to changes as it passes the cells making up the tubule. Some salts and organic substances are reabsorbed before they leave the body. In the few cases which have been studied, it has been noted that phenol red is secreted into the urine by nephridial cells.

Higher invertebrate kidneys. The phylum Annelida has produced the peak development of small independent nephridia. Some individual animals may possess thousands, of which some are directed to the exterior and others to the digestive lumen. An evolutionary line from the group leads to the Mollusca and Arthropoda, and along this line it is found that the nephridial system is replaced by quite different structures. *See* ANNELIDA.

Mollusca. The trochophore larvae of the mollusks possess nephridia. As the larvae settle down and undergo metamorphosis, the nephridia are lost and replaced by kidneys proper. Filtration through the coelomic lining begins the process of urine formation in the animals possessing metanephridia. Now the kidney itself originates from the remnants of the coelom, which is very much reduced in modern mollusks. It is to be expected that filtration would be retained as a first step in urine formation, and this is found in all the mollusks so far studied. In the bivalve, or pelecypod, mollusks the filtration is actually through the walls of the heart; in the gastropods the filtration is probably from capillaries in the walls of the kidneys; and in cephalopods the filtrate appears to be formed by special organs, the branchial heart appendages. Filtration pressures are not high, but the mollusks generally have a low protein content in their blood, and so the reabsorptive forces which oppose filtration are not high either. As a result, surprisingly large volumes of filtrate have been obtained in the few studies made on the subject. From the pericardial space, or branchial heart appendage, the filtrate passes next though the renopericardial canal into the kidney proper, where two processes are simultaneously at work. Because the process of filtration cannot be finely discriminating, it merely separates water and solutes from blood proteins and cells. Many useful solutes remain in the filtrate and would be lost from the body, but they are absorbed by the cells of the kidney from the urine as it passes over the many folds into which the lining of the kidney is thrown. At the same time other cells of the kidney actively secrete materials into the urine. These three processes, filtration, reabsorption, and secretion, occur repeatedly in phylum after phylum, regardless of the origin of the kidney tissues.

Mollusks may be exposed to desiccation during tidal excursions or at certain seasons of the year. A low metabolic rate and the usual presence of an impermeable shell make them well suited for surviving prolonged drought. The kidneys may release no urine whatever during such a period, the chief nitrogenous excretion, uric acid, simply forming crystals within the kidney or in the kidney cells. At the next wet period normal function is restored. *See* MOLLUSCA.

Arthropoda. Of the very large phylum Arthropoda, only two major subgroups are treated here, the crustaceans and the insects.

1. *Crustacea*. The crustaceans which have been most studied originate their urine at a thin-walled sac held open by strands of connective tissue passing from the external surface of the sac or in a network of fine capillaries in its walls results in formation of a filtrate upon which other changes may later be made. The filtrate formed in the end sac runs next through a long canal lined with cuboidal or columnar epithelium before it is voided to the exterior. While the filtrate passes through the canal, salts and water may be reabsorbed to conserve them for the body, while other materials are secreted into the urine and thus excreted at a higher rate than they would have been by the process of filtration alone.

2. *Insecta*. Excretion in the insects follows an entirely different pattern from any described to this point. In his early work on the histology of the silkworm, M. Malpighi described the tubes which were later named the Malpighian tubes when their excretory function was learned. These tubes, 2–150 in number, lie in the abdominal cavity; when they are few in number they are long, and when they are more numerous they are short. The Malpighian tubes begin development as blind pouches, lie free in the body cavity, are convoluted or not depending upon their length, and finally open most frequently into the midgut but sometimes into the hindgut. There is no obvious mechanism for filtration of blood, and it is thought that the urine in all the members of this large group of animals is formed entirely by secretion of the cells lining the tubes. In fact the tubes sometimes secrete other than excretory substances; for example, silk for the cocoon is secreted by the Malpighian tubes of some of the Collembola and Neuroptera. In the absence of filtration, it might be surmised that there would be no need for reabsorption, but it has been shown that uric acid is secreted in the form of the relatively soluble sodium and potassium salts in the upper segment, and that the free uric acid is precipitated in the lower segment where the water and the liberated bases are reabsorbed and used again. In insects much reabsorption of water occurs in the hindgut. *See* INSECT PHYSIOLOGY; INSECTA.

Nephrons and the Vertebrata. The large number of phyla possessing nephridia has been emphasized above. One major evolutionary line whose forebears, though uncertain, were undoubtedly animals possessing nephridia, has progressed along a path that produced the echinoderms and led to the chordate group of animals. *See* CHORDATA; ECHINODERMATA.

The phylum Echinodermata, composed of relatively inactive marine animals, does not shed much light on the process since these animals do not have specialized excretory organs. Experimentally injected foreign substances are picked up by wandering ameboid cells, the coelomocytes, and delivered to the digestive tract, from which they are excreted with the feces.

Among the chordates, the cephalochordates have nephridia comparable to those of the annelids. The vertebrates have a common origin with the other chordates, but the nephridia are called nephrons and are regarded as the unit structures of the kidney. With minor modifications, each class of vertebrates possesses kidneys made up of nephrons like those of its relatives, all the way from the cyclostomatous fishes up to the mammals. The small and primitive animal is characterized by few nephrons, and the large and highly evolved animal by many nephrons, so that a single kidney of the human may contain 1,250,000 of these units. *See* KIDNEY.

A nephron may be said to be composed first of a Malpighian body consisting of a tuft of capillaries, the glomerulus, surrounded by a close-fitting sac, Bowman's capsule. Blood pressure in the tuft of capillaries filters a large amount of fluid into the capsule, from which it passes into a much-convoluted tubule draining the fluid away through a narrow passage between closely packed kidney cells. These cells carry out the complicated processes of reabsorbing for the body the materials of usefulness and of adding to the urine, from the blood, materials of such a nature that special mechanisms could usefully be developed for their transportation across the kidney cells. Nevertheless, the utilization of the three processes of filtration, reabsorption, and secretion used throughout the rest of the animal kingdom can be noted. A few special features of excretion in the different vertebrate classes are worthy of comment.

Fishes. In the Insecta there is a specialization of structure such that secretion is the primary mechanism used by the Malpighian tube in fulfilling the function of excretion for the animal. A similar device has been developed by some marine fishes; the Malpighian body of the vertebrate nephron is gone, leaving only the tubule, closed at the end where the glomerulus should be found. Since the means of accomplishing filtration is gone, it is clear that secretion is the only remaining mechanism for accomplishing the excretory functions of the kidneys in these animals. These nephrons make up what is called an aglomerular kidney, and the aglomerular fishes represent a successful group of animals, being numerous and widely distributed. It is thought that the reduction or loss of the glomerulus represents a means of water conservation. Fish blood is osmotically inferior to the salt-water environment. The fishes consequently continuously lose water osmotically through the gills and pharyngeal membranes, and any means of reducing excretory water loss has survival value.

The general osmotic problem faced by marine fishes is of such overriding nature that an entirely extrarenal excretory mechanism has developed to meet it in the gills of fishes. The characteristic cation of living cells is potassium, whereas the abundant cation of body fluids and saline waters is sodium. Cell membranes are permeable to these small ions, so that a majority of all animal cells continuously pump sodium out of the cell and potassium into the cell and must do so to continue normal life. Certain cells of the fish gill direct the pumping action not generally outward but polarize it, so that sodium entering from the blood and tissue fluid is pumped against a steep gradient into the seawater. In this way the kidney of the fish is spared the task of excreting the sodium ion, a task which the kidney of most terrestrial animals must perform. *See* RESPIRATORY SYSTEM.

Amphibia. Amphibians are characteristically bound to bodies of fresh water. Their osmotic problems are therefore quite different from those of the marine fishes just discussed. Their nephrons begin the process of urine formation with vigorous filtration. As the filtrate passes down the tubules, more salt than water is reabsorbed, water ordinarily being present in excess, and a hypotonic urine results. Even so, salt is continuously being lost in the urine and so is at a premium in these animals. The deficit is made up by the skin, which has the power of taking sodium and chloride ions from the environment into the blood against a steep gradient.

Reptiles. The discussion of excretory specialization in reptiles will be deferred until the section dealing with the excretion of nitrogenous wastes.

Birds. Marine birds face a situation similar to that of the marine fishes. Although they may not lose water by osmosis directly to the seawater, some of them remain in environments where fresh water is available. At the same time they may feed on marine invertebrates whose body cells are in osmotic equilibrium with the seawater, and so through the digestive tract they are exposed to salt loads that constitute a serious problem to the body. In these cases research shows that it is not the kidney which has assumed the excretion of salt as a higher rate but special glands of the head region, the nasal glands. These glands produce a salt solution very hypertonic to the blood, which is blown, drips, or is shaked from the beak. They are thus able to meet the imposed load easily. The function of these glands has been well integrated into the physiology of the animal and is controlled from the osmoreceptors of the body by way of the central nervous system. *See* SALT GLAND.

Mammals. Desert mammals meet somewhat the same excretory problem that marine fishes and some marine birds must meet, namely a limitation on the water available for excretion. It has been shown that their kidneys have developed the capacity to excrete salt far beyond that of the ordinary mammal. These animals can comfortably eat dry food and drink seawater and, because they excrete a urine hypertonic to seawater, maintain themselves in satisfactory water balance. Some marine mammals are not only limited in their fresh-water intake but eat marine invertebrates containing large amounts of salt. These animals show the same adaptation as the desert mammals: great ability to concentrate salt in the kidney tubules. In fact only in the mammals has the kidney been the organ of choice for the production of a hypertonic excretory fluid. The mechanism developed makes use of the ordinary structures of the kidney, but in an entirely new way.

Excretion of Nitrogenous Wastes

Nitrogenous wastes, the group of compounds which, it may be assumed, necessitated the development of excretory systems, are ammonia, urea, and the purine derivatives. From the standpoint of the enzymologist, these compounds make up a single family, all the members of which can be synthesized by well-known enzymes starting either from the ammonia generated from amino acids or with purines generated from nucleic acids. The intermediates are illustrated in **Fig. 1**. It has been estimated that 95% of the nitrogenous wastes of the average animal comes from proteins and only 5% from nucleic acids, but because of the metabolic interconnection, one cannot predict from these percentages what the form of excretion will be. Instead it turns out that the habitat and the need for conservation of water determine the form of excretion. In animals which spend their lives in an abundance of water—for example, amphibian larvae and many marine invertebrates—the ammonia split from protein is simply liberated to the water and carried away. The animals are termed ammonotelic. But these animals also produce nitrogenous wastes from nucleic acids. Instead of liberating adenine and guanine directly to the environment, the complement of enzymes shown in Fig. 1 is invoked, the material is split all the way to ammonia, and the ammonia released therefore originates from protein and nucleic acid. This general principle of a common denominator of nitrogen excretion was enunciated by M. Florkin. The results of the principle, and some exceptions, are shown in **Fig. 2**.

Elaboration of urea. Animals of larger size, in particular those which might not have a steady supply of water, cannot tolerate the accumulation of ammonia which would result from the use of ammonia as the single nitrogenous excretory product, and it was necessary that a less toxic material be substituted. In retrospect it is clear that the next evolutionary step, the elaboration of urea, was both most important and very difficult. Early biochemists speculated that urea was formed from two molecules of ammonium carbonate in what appeared to be a relatively simple synthesis. Progress in working out the actual pathways was slow in spite of the fact that the action of the enzyme arginase in splitting urea from the amino acid arginine had long been known. There was simply not enough arginine available to account for the amounts of urea formed. It was through the genius of H. A. Krebs that the solution was found. Only catalytic amounts of arginine need to be present. Arginase splits off urea, leaving a molecule of ornithine. This is condensed with an ammonia to form arginine again. Ornithine and citrulline had long been known in nature as constituents of a few plants, but their possible roles in such a cycle had not been suggested until Krebs demonstrated that catalytic amounts of these compounds greatly accelerated the test-tube synthesis of urea by liver tissues. Since that time, much detailed information has been accumulated about this reaction, including the energy sources needed for the synthesis. Unlike the simple liberation of ammonia, the synthesis of urea demands the provision of a special complement of enzymes and the utilization of some of the body's supply of energy. The result for the whole animal, however, is the substitution of urea, a relatively soluble and nontoxic material, for the highly toxic ammonia molecule. Animals which synthesize and excrete urea are spoken of as being ureotelic. Most fishes and many amphibians fall into this category. There is a decided tendency for the

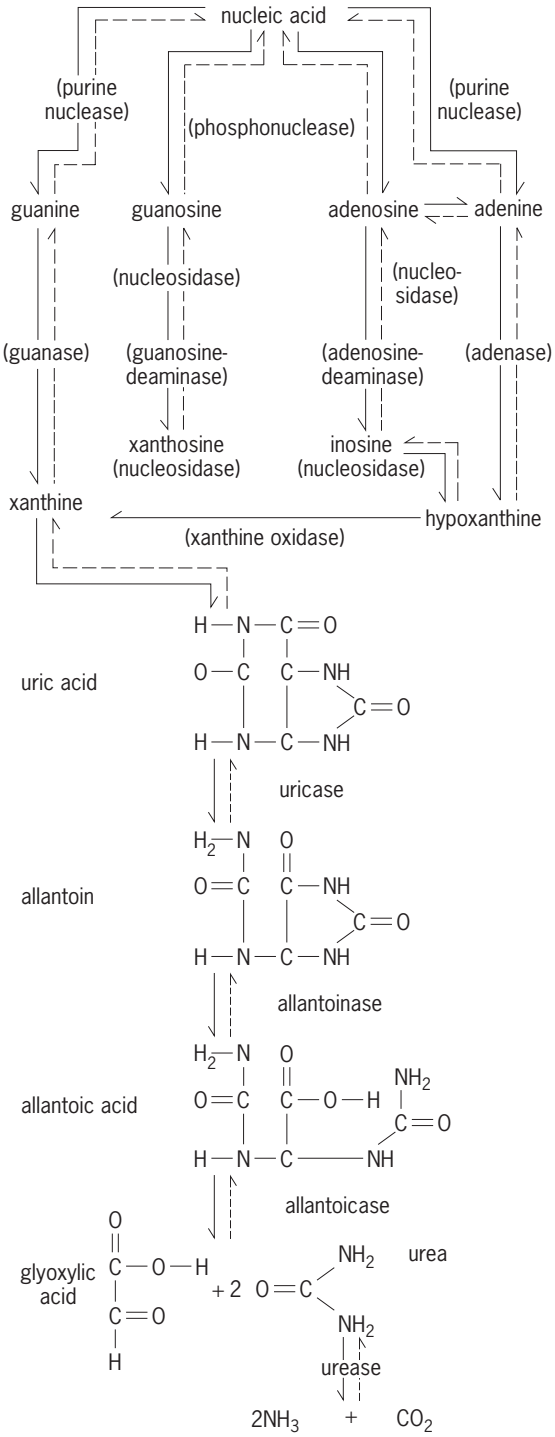

Fig. 1. Sequence of breakdown from nucleic acid to ammonia. Reverse pathways broken indicate either lack of knowledge of the mechanism or that other enzymes than those shown are involved. (*After M. Florkin, Biochemical Evolution, Academic Press, 1949*)

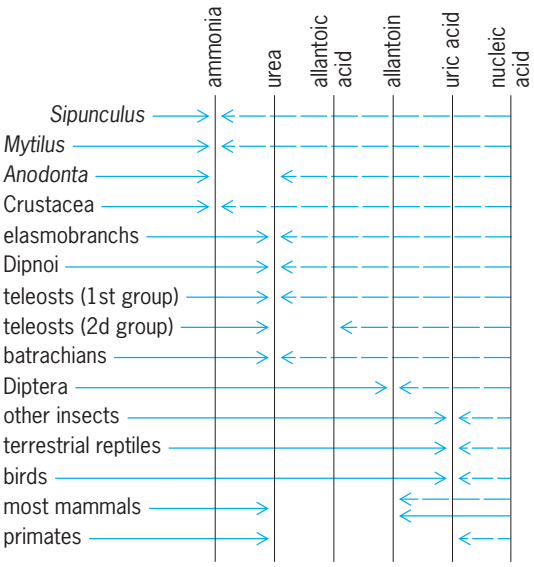

Fig. 2. Chart of the principle of excretion of a common nitrogenous waste, whether derived from protein or nucleic acid. (*After M. Florkin and G. Duchateau, Arch. Int. Physiol., 53:267–307, 1943*)

purine bodies to be degraded to the level of urea and excreted in this form (Fig. 2). *See* NITROGEN EXCRETION.

Urea synthesis and excretion. The synthesis and excretion of urea made animals independent of a direct and abundant source of water. The solubility of urea, which appears actually to have been one of its virtues for the development of the primitive kidney, turns out to be a disadvantage to animals such as the birds and desert lizards which, for conservation of weight and water, need an excretory material which is both nontoxic and insoluble. At the same time the material must be capable of being secreted by kidney cells in solution in water and precipitated only in regions from which it can be expelled safely from the body. From the evolutionary point of view, such a material had been available all the time to the more primitive animals. It will be recalled that in the fish and amphibians (Fig. 1) the nitrogen from the purines was degraded by way of uric acid to urea. The birds and most reptiles arrest this degradation at the uric acid stage. By enzymatic processes not yet fully understood, they carry the synthesis of the protein ammonia past the stage of urea synthesis and all the way up to uric acid. The uric acid is precipitated in the ureters and cloaca, and the pasty material is eliminated through the rectum. In this way another great step is made in the conservation of water. These animals are called uricotelic. It may be noted from Fig. 2 that such a mechanism was independently developed by the insects, and for the same reasons. It should be pointed out that when such mechanisms develop there is not necessarily a clear-cut departure from the previous condition. Careful study shows therefore that various animals excrete varying mixtures of ammonia, urea, and uric acid, instead of only a single material. Arthur W. Martin

Bibliography. P. Dejours et al. (eds.), *Comparative Physiology: Life in Water and on Land,* 1987; N. Heisler (ed.), *Acid-Base Regulation in Animals,* 1986; B. L. Muir, *Pathophysiology: An Introduction to the Mechanisms of Disease,* 2d ed., 1988.

Exon

In split genes, a portion that is included in the ribonucleic acid (RNA) transcript of a gene and survives processing of the RNA in the cell nucleus to become part of a spliced messenger RNA (mRNA) or structural RNA in the cell cytoplasm. Split genes are those in which regions that are represented in mature mRNAs or structural RNAs (exons) are separated by regions that are transcribed along with exons in the primary RNA products of genes, but are removed from within the primary RNA molecule during RNA processing steps (introns). *See* INTRON; RIBONUCLEIC ACID (RNA).

Exons comprise three distinct regions of a protein-coding gene. The first is a portion that is not translated into protein, but contains the signal for the beginning of RNA synthesis, and sequences that direct the mRNA to ribosomes for protein synthesis. The second is a set of exons containing information that is translated into the amino acid sequence of a protein. The third region of a gene that becomes part of an mRNA is an untranslated end portion that contains signals for transcription termination and for the addition of a polyadenylate tract at the end of a transcript. *See* GENETIC CODE; PROTEIN; RIBOSOMES.

The evolutionary origin of the exon-intron organization of eukaryotic genes has not been clarified. However, it has been determined from comparisons of structural-functional domains in proteins with the exon domains in corresponding genes that there is often a correlation of the two. For example, the gene for one of the human cell-surface proteins that are important in the recognition of self and nonself by the immune system, the HLA-B7 gene, has nine exons. Exon 1 encodes the 5′ (beginning) untranslated portion of the mRNA that is presumably responsible for steps in the initiation of protein synthesis from the mRNA. Exon 2 encodes the site of translation initiation and the sequence of amino acids (the signal peptide) that designates the protein being made as one that will either be a component of membrane or pass through a membrane to be secreted. Exons 3, 4, and 5 encode the three domains of the HLA-B7 protein that are modified by the addition of carbohydrates and are exposed to the cell exterior. Exon 6 encodes the transmembrane portion of the protein, being composed of codons for a series of amino acids that are able to interact with the lipids of the cell membrane. Exons 7 and 8 encode the cytoplasmic ("anchor") portion of the HLA-B7 protein, and exon 9 contains polyadenylation signals.

The mechanism by which the exons are joined in RNA copies of genes is called RNA splicing, and it is part of the maturation of mRNAs and some transfer and ribosomal RNAs (tRNAs and rRNAs) from primary transcripts of genes (see **illus.**). Three different RNA splicing processes have been identified. One involves mRNA precursors in nuclei, and

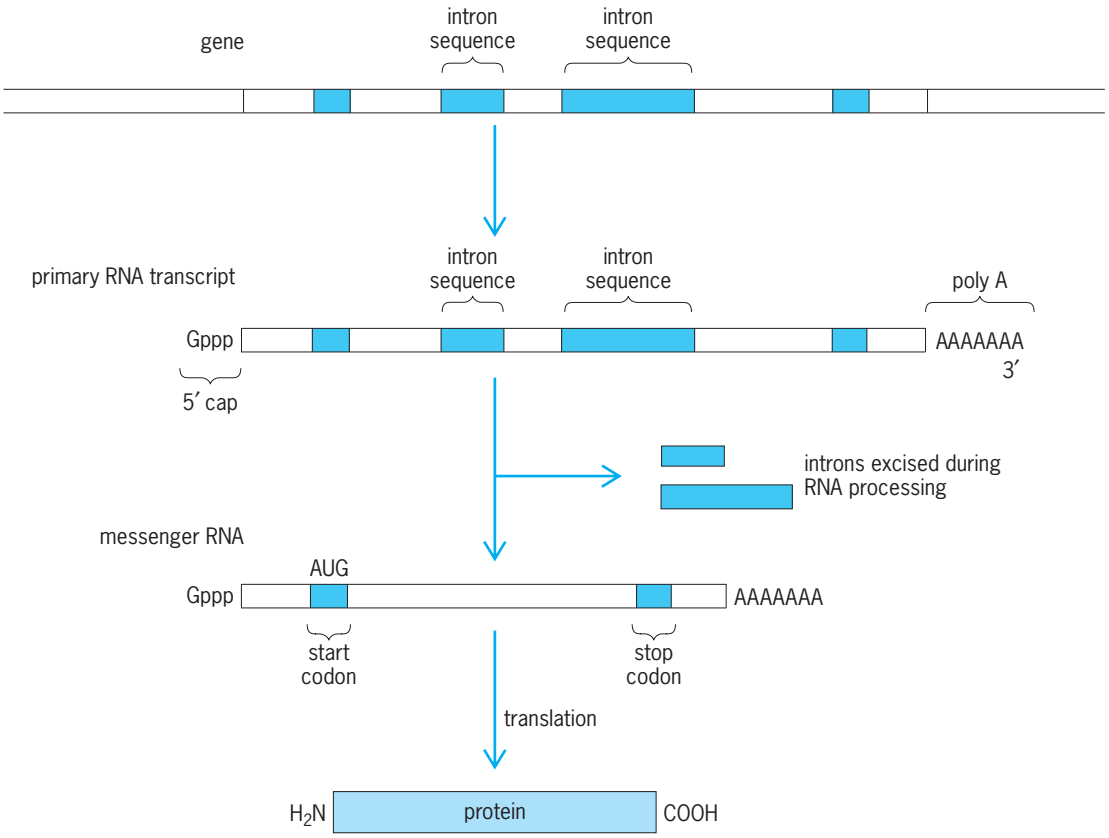

Map of a eukaryotic gene, primary RNA transcript, and messenger RNA for a protein. Exons, the unmarked segments of the map, are separated by intron sequences, which are spliced out during synthesis of messenger RNA. The appearance of the 5′ cap, a modified guanine nucleotide, precedes or accompanies any splicing activity. The 3′ end, always a series of adenine nucleotides, is added at the same time. The start codon, AUG, specifies methionine. The stop codon, which signals the termination of protein-chain translation, is a ribonucleotide triplet.

specific sequences at exon-intron junctions that are recognized by certain nuclear ribonucleoprotein particles that facilitate the cleavage and ligation of RNA. Another applies to nuclear precursors of tRNA, where splice sites are determined by structural features of the folded RNA molecules. The third form of splicing was discovered in studies of protozoan rRNA synthesis, and has also been shown to be a part of the maturation of both rRNA and mRNA in yeast mitochondria; it is an autocatalytic process that requires neither an enzyme nor added energy such as from adenosine triphosphate. In bacteria, it has been observed that certain mRNAs in bacteriophage-infected cells result from the same sort of self-splicing that has been described for protozoan rRNA. *See* GENE.

Peter M. M. Rae

Bibliography. J. Darnell et al., *Molecular Cell Biology*, 1986; J. D. Watson et al., *Molecular Biology of the Gene*, 5th ed., 2000.

Exopterygota

A division (also known as Hemimetabola) of the subclass Pterygota, including those insects that show relatively slight change in body form with growth. They develop through a series of immature forms (nymphs) from hatchling to adult, so that wings grow as external pads and enlarge with each molt. The nymphs are often scaled-down copies of the adults, but in some cases, particularly in those orders with aquatic nymphs and aerial-terrestrial adults (such as in Ephemeroptera and Odonata), a considerable difference in body form exists between adults and their immature forms. This development, called incomplete metamorphosis, is characteristic of all Paleoptera and some members of the infraclass Neoptera (all ancestrally winged insects except Endopterygota). Most exopterygotan taxa have freely mobile nymphs with well-developed legs, although conspicuous exceptions occur in the form of the scale insects (Coccoidea) and some of their highly sedentary sap-feeding allies.

Exopterygota are a very diverse group, encompassing plant feeders, predators, and animal parasites, and living in nearly all habitats and areas where insects are found. Common examples of Exopterygota are mayflies, dragonflies, grasshoppers, termites, cockroaches, aphids, plant bugs, biting and sucking lice, and thrips. *See* ENDOPTERYGOTA; INSECTA; PTERYGOTA.

William L. Brown, Jr.

Bibliography. Commonwealth Scientific and Industrial Research Organization, *Insects of Australia: A Textbook for Students and Researchers*, 2d ed., 1991; O. W. Richards and R. G. Davies, *Imm's General Textbook of Entomology*, 2 vols., 10th ed., 1994.

Exotic nuclei

Nuclei with ratios of neutron number N to proton number Z much larger or much smaller than those of nuclei found in nature. Studies of nuclear matter under extreme conditions, in which the nuclei are quite different in some way from those found in nature, are at the forefront of nuclear research. Such extreme conditions include nuclei at high temperature and at high density (several times normal nuclear density), as well as those with larger or smaller N/Z ratios. The N/Z ratio depends on the nature of the attractive nuclear force that binds the protons and neutrons in the nucleus and its competition and complex interplay with the disruptive Coulomb or electrical force that pushes the positively charged protons apart.

A chart of the nuclides is shown in **Fig. 1**. The squares are the stable nuclei ($Z \leq 83$) and the very long-lived nuclei (with half-lives of the order of 10^9 years or more) found in nature. The first jagged lines to either side are the limits of the presently observed nuclei; very little is known about those at the edges. The light, stable nuclei (such as ^{4}He) have $Z = N$, reflecting the preference of the nuclear forces for $N = Z$ symmetry, but as Z increases, the strength of the Coulomb force demands more neutrons than protons to make a particular element stable, and $N \approx 1.6\,Z$ for the heaviest long-lived nuclei (such as ^{238}U. For $Z > 92$, the Coulomb force causes most nuclei to spontaneously fission. *See* NUCLEAR FISSION.

Stable nuclei lie in the so-called valley of beta stability. As the N/Z ratio decreases (proton-rich nuclei) or increases (neutron-rich nuclei) compared to that of the stable isotopes, there is, respectively, energy for a proton or neutron in the nucleus to undergo beta (β^+, β^-) decay to move the nucleus back toward stability. Most knowledge of nuclear structure and decay has been gained from nuclei in or near the valley of beta stability. *See* RADIOACTIVITY.

Structure and decay. The spherical shell model was developed to explain the so-called spherical magic numbers for protons and neutrons, which give nuclei with these numbers a very stable structure and spherical shape. Spherical magic Z and N of 2, 8, 20, 28, 50, 82, and 126, and the weaker magic Z and N at 40, are shown in Fig. 1 as horizontal and vertical lines. The nuclear shell model resembles the atomic shell model, where the noble gases (helium, neon, argon, and so forth) have filled shells and then there is a gap in the binding energy to the next electron shell (or orbit). A major question concerns what happens to these spherical magic proton and neutron gaps (orbits) in exotic nuclei. *See* NUCLEAR STRUCTURE.

Another major question concerns the decay modes whereby a nucleus rids itself of excess energy and returns to the stable forms of nuclear matter. As N/Z decreases or increases from the stable values, a point is reached for a given Z where, if one more neutron is pulled out, one proton becomes unbound (an isotope of that element cannot exist with that number of neutrons), or, if one more neutron is added, that neutron is not bound to that nucleus. These limits define, respectively, the proton and neutron drip-lines (the jagged lines furthest from the valley of stability in Fig. 1).

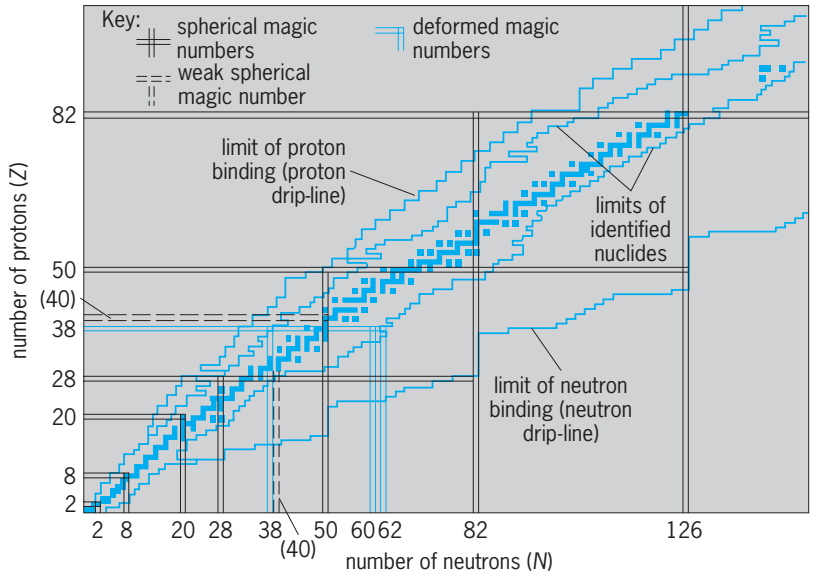

Fig. 1. Chart of the nuclides. The squares are the stable and very long-lived nuclei found in nature.

Answers to the above questions give significant insights into the structure and decay modes of nuclear matter that make it possible to test and extend theoretical models of the nucleus and the understanding of the nature of both the strong and weak nuclear forces. These insights could not be gained by studies of nuclei in and near the valley of stability.

Shape coexistence. These insights include the discovery of nuclear shape coexistence, which was first observed in the very light mercury and selenium isotopes. Competing bands of levels occur in one nucleus, which overlap in energy but are quite separate in their decays because they are built on quite different coexisting nuclear shapes. Shape coexistence is now known to be important in many nuclei throughout the periodic table. Even multiple (four, and possibly more) different-shape coexisting structures are observed, for example, in ^{185}Au and ^{187}Au.

Deformed magic numbers. A significant advance was made in the nuclear shell model with the discovery of new magic numbers associated with shell gaps in the energies of the proton and neutron orbitals. These new numbers may be called deformed magic numbers because they stabilize a nucleus in a deformed shape, just as the spherical magic numbers give stability to a spherical shape. The deformed magic numbers (shell gaps) identified so far include N and Z of 38 and N of 60 and 62. The deformed shell gaps at 38, 60, and 62 appear in **Fig. 2**, along with spherical shell gaps at 28 and 50, and other expected shell gaps are shown.

Shell gap reinforcement. These gaps led to a further advance in the shell model: the importance of proton and neutron shell gap reinforcement. Two new islands of the strongest ground-state deformations

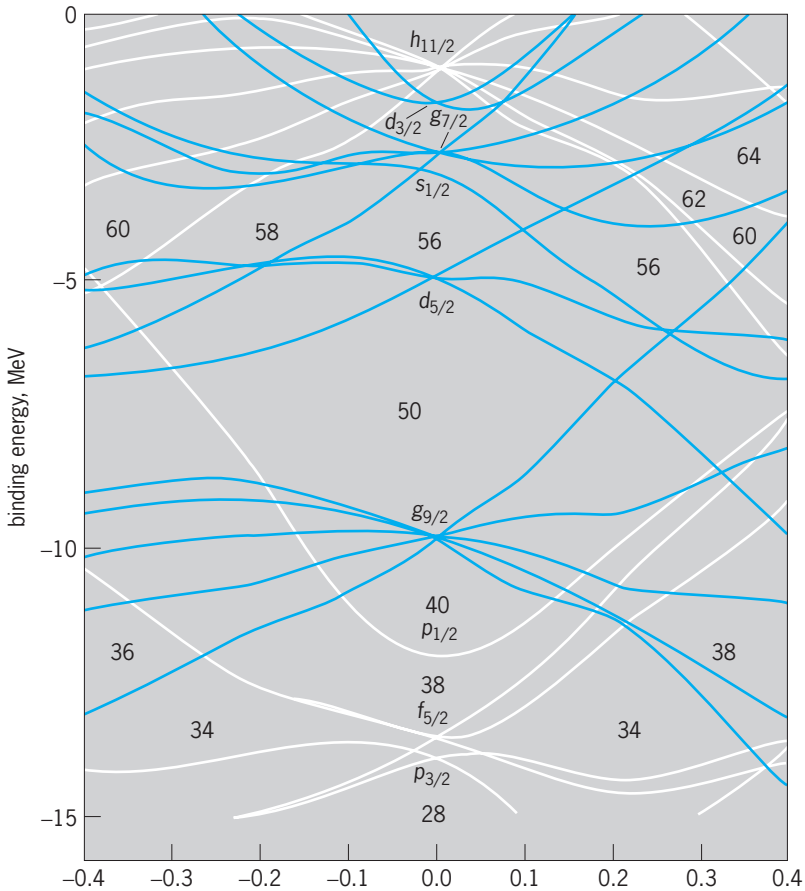

Fig. 2. Binding energies for neutrons and protons in nuclei in the mass region $A = 60$–120. These energies are plotted against the nuclear deformation parameter ϵ. The energies for a spherical shape occur in the center ($\epsilon = 0$), for prolate shapes to the right, and for oblate shapes to the left. A gap between levels occurs when a particular shell or orbital is filled. Numbers in gaps give total numbers of neutrons or protons in closed-shell or closed-subshell configurations. Subscripted letters give orbital and total angular momenta of single-particle states. *(After R. Bengtsson et al., Nuclear shapes and shape transitions, Phys. Scripta, 29:402–430, 1984)*

known above $A = 30$ (where $A = Z + N$ is the mass number) have been discovered centered on $N = Z = 38$ and $Z = 38$, $N = 60, 62$. In these cases, the protons and neutrons reinforce each other for the same strong deformed shape and override the weaker spherical shell gaps found at 38 and 40 in Fig. 2. On the other hand, the discovery of spherical ${}^{68}_{28}\mathrm{Ni}_{40}$ shows that 40 is a good number for a spherical shell when it is reinforced by a proton partner (or neutron as in ${}^{90}_{40}\mathrm{Zr}_{50}$) with a strong spherical shell gap like 28 (or 50). Shell gap reinforcement is important for both spherical and deformed shapes.

With all the shell gaps for different deformations in Fig. 2, the small number of known deformed magic numbers requires explanation. The 38, 60, and 62 deformed magic numbers are important only when the protons and neutrons reinforce each other for the same deformation. This reinforcement is likely to be necessary to observe the other gaps in Fig. 2. The competition between the different shell gaps also leads to shape coexistence as observed. It is very difficult to make exotic nuclei with Z and N shell gaps at the same deformation, but the possibility of discovering additional deformed shell gaps is of much current interest.

Modes of decay. Exotic nuclei exhibit decay modes not seen near stability, such as proton radioactivity and beta delayed-particle emission. (After beta decay, the highly excited nucleus can emit one or more particles, such as one or two protons, an alpha particle, or one or two neutrons.) Joseph H. Hamilton

Neutron-rich nuclei. With calculations based on relatively simple models of the nucleus (for example, the liquid-drop model), it can be shown that nuclei with much greater neutron excess than those in, or near, the stability valley should remain stable with respect to one- and two-neutron emission, the limit that defines the neutron drip-line. This limit is approximately given by $N \approx 2.4Z$ for $Z > 10$. Between the stability valley, the neutron drip-line, and the limit imposed by spontaneous fission lie approximately 4000 neutron-rich nuclei, which should be stable with respect to the strong nuclear forces. They can, of course, decay by alpha and beta decay with lifetimes of $\gtrsim 1$ microsecond, which are very long times compared to the typical strong-force decay times of 10^{-23} s. Out of the 4000 neutron-rich nuclei, less than 1000, which lie close to the stability valley and can be reached by simple nuclear reactions, have been studied in great detail. Together with another ~700 nuclei on the proton-rich side of the valley, these constitute the sample on which almost all present knowledge of nuclear forces and nuclear structure is based. The remaining proton-rich (~1500) and neutron-rich (~3000) nuclei, which lie far from the stability valley and are difficult to reach, have earned the appellative exotic.

The exotic nuclei provide unusual opportunities for testing models of nuclear structure that have been proposed and optimized for nuclei in, or near, the stability valley. The neutron-rich exotics are especially important because of their astrophysical connection. The neutron-rich isotopes of the lightest elements, hydrogen and helium, provide models for neutron matter, and those of medium-weight elements play important roles in nucleosynthesis via the *r*-process. The heaviest neutron-rich nuclei, the superheavies, not only test the limits of nuclear stability but are of practical interest because of their expected unusual fission properties. *See* NEUTRON STAR; NUCLEOSYNTHESIS.

Techniques. A great variety of experimental techniques have been developed to create and study neutron-rich exotic nuclei. These include multinucleon transfer and charge-exchange reactions, fission, fusion, and projectile fragmentation and target fragmentation (spallation) reactions. The capabilities of these various techniques have been greatly enhanced by the use of medium- to high-energy beams of heavy ions, pions, and radioactive nuclei, together with stable and radioactive targets. The reaction products of interest are mechanically as well as electromagnetically transported and analyzed, and the stability, masses, half-lives, radii, spins, electromagnetic moments, and decay schemes have been

studied with ingenious techniques. *See* NUCLEAR REACTION; PARTICLE ACCELERATOR.

Limits of identified nuclei. The neutron drip-line appears to have been reached for elements with $Z \leq 7$; the heaviest particle-stable isotopes identified are ${}^{3}_{1}H_2$, ${}^{8}_{2}He_6$, ${}^{11}_{3}Li_8$, ${}^{14}_{4}Be_{10}$, ${}^{19}_{5}B_{14}$, ${}^{22}_{6}C_{16}$, and ${}^{23}_{7}N_{16}$. For oxygen the most neutron-rich isotope reached so far is ${}^{24}_{8}O_{16}$, although particle stability is expected to extend to ${}^{28}_{8}O_{20}$. For the alkali-metal sodium it has been possible to go as far as ${}^{34}_{11}Na_{23}$. For heavier elements the most neutron-rich isotopes reached so far seem to be at $N/Z \simeq 1.60 \rightarrow 1.70$. Examples are ${}^{50}_{19}K_{31}$, ${}^{99}_{37}Rb_{62}$, ${}^{148}_{55}Cs_{93}$, and ${}^{228}_{87}Fr_{141}$ from among the alkali metals, and ${}^{78}_{29}Cu_{49}$, ${}^{102}_{38}Sr_{64}$, ${}^{133}_{50}Sn_{83}$, ${}^{146}_{54}Xe_{92}$, ${}^{164}_{64}Gd_{100}$, ${}^{214}_{82}Pb_{132}$, ${}^{264}_{100}Fm_{164}$ from among the other elements. These N/Z ratios are still far from the neutron drip-line value of $N/Z \approx 2.4$; thus, much work remains to be done.

Nuclear structure. Several interesting nuclear structure results have emerged from the measurements of neutron-rich exotic nuclei.

1. Several nuclei that at one time or another were predicted to be unbound (mainly on the basis of assumed spherical shapes) have been found to be bound (for example, ${}^{8}_{2}He_6$, ${}^{11}_{3}Li_8$, and ${}^{14}_{9}Be_{10}$).

2. Mass extrapolations using relations based on systematics for the nuclei in, or near, the valley of stability, such as the Garvey-Kelson relations, have been found not to work well for the extremely neutron-rich nuclei.

3. An indirect measurement of the matter radius of ${}^{11}Li$ shows it to be very large, approximately the same as that of ${}^{28}Si$. It has been suggested that this reflects a unique phenomenon, a loosely bound neutron halo around a ${}^{8}Li$ or ${}^{9}Li$ core.

4. Particle transfer experiments with heavy ions have been used to infer that the nucleus ${}^{6}H$ is unbound with respect to one-neutron emission by only 2.7 MeV. If true, this would be of great interest. Unfortunately, a high-sensitivity pion double charge exchange experiment has failed to verify this claim.

5. Despite the fact that ${}^{9}He$ is found to be neutron-unstable by 1.1 MeV, the ground state and at least two excited states of ${}^{9}He$ are found to be ≤ 0.5 MeV wide. Surprisingly, the binding energies of all three states are well predicted by shell-model calculations whose parameters are optimized for the nuclei in the valley of stability.

6. Two-neutron separation energies, differences in charge radii, and in many cases location of the first 2^+ states have been studied for a large number of neutron-rich nuclei across the major shell-closing (magic) neutron numbers. Rapid unexpected onsets of large deformation have been found, for example, at $N \simeq 60$ (${}^{98}_{38}Sr_{60}$, ${}^{100}_{40}Zr_{60}$).

7. Elements up through $Z = 112$ have been successfully synthesized, as well as $Z = 114$, $Z = 116$, and $Z = 118$. *See* DARMSTADTIUM; ELEMENT 112; MEITNERIUM; ROENTGENIUM; TRANSURANIUM ELEMENTS.

Proton-rich nuclei. Highly unstable, proton-rich nuclei that define the limits of nuclear stability offer unique opportunities to study new or unusual radioactive decay phenomena. This boundary region, in which the proton binding energy goes to zero, has been delineated up to nuclides with mass $A \approx 60$ and has been established to a significant degree in the region $A = 100$–185. It is relatively easier to reach this proton drip-line than the equivalent neutron drip-line, since, as successive neutrons are removed from a given stable isotope, the effects of the strong coulombic repulsion between the protons in the nucleus cause a rapid increase in the decay energy.

Apart from the observation of new radioactive decays (for example, proton radioactivity), studies of exotic proton-rich nuclei determine the nuclear mass surface under extreme conditions, test the effects of charge independence of nuclear forces, study the weak interaction in nuclei, and evaluate complex shell-model calculations of nuclear structure.

Experimental techniques. These must be capable of observing radioactive nuclei produced in low yield and with half-lives of 5–100 milliseconds to observe beta decays of proton-rich nuclides; further, at the drip-line, other decay modes such as proton decay may occur with considerably shorter half-lives than are predicted for beta decay.

Proton-rich nuclides can be produced by bombarding appropriate targets with either light-ion or heavy-ion beams. In principle, the observation techniques involve either prompt detection of the reaction products or on-line radioactivity measurements. An example of a prompt-detection technique is the use of one or more magnetic spectrometers followed by sophisticated detectors to study the fragmentation of heavy-ion projectiles with in-flight separation of the exotic species of interest. Examples of techniques involving radioactivity measurements are the helium-jet technique (in which nuclear reaction products are thermalized in helium, attach to an aerosol, and are transported rapidly to a detection area) and the use of on-line isotope separators.

Decay modes. The large beta-decay energies of these exotic nuclei permit the population of highly excited states in their daughters, so that beta-delayed proton and alpha-particle emission can be observed. More exotic decays are also possible.

Beta-delayed particle emission. Experimental information on how the many open beta-decay channels of a quite proton-rich nuclide manifest themselves can frequently be derived by measurements of delayed radiations (protons or alpha particles) from excited states of the daughter nuclide that are populated by beta (positron) decay. These decay modes are characteristic of nuclides with $N - Z \leq -3$. **Figure 3** illustrates this decay for the lightest aluminum isotope that exists, ${}^{22}Al$ ($N - Z = -4$), with a half-life of approximately 70 ms. This abbreviated decay scheme shows a favored positron decay of ${}^{22}Al$ (with 18.5-MeV available decay energy) to the analog state in its ${}^{22}Mg$ daughter, with subsequent emission of a proton (p) to either the ground state or first excited state of ${}^{21}Na$. (The analog state arises if nuclear forces are charge-independent, and it has the same nuclear structure as the parent state in ${}^{22}Al$.)

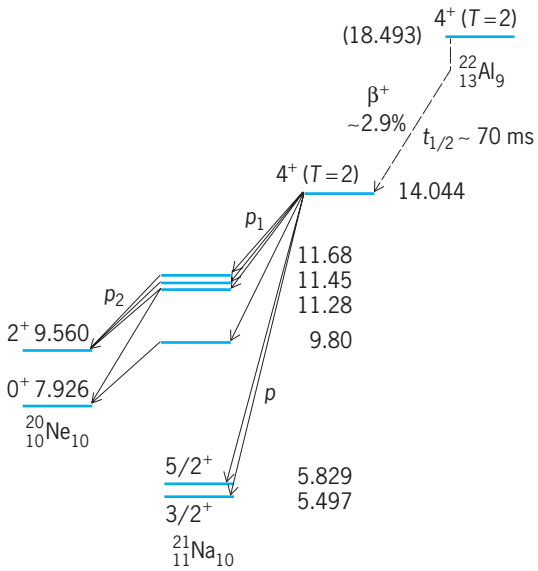

Fig. 3. A partial decay scheme for aluminum-22 (^{22}Al). Numbers above the bars at left give spins and parities of levels; numbers above the bars at right or to right of the bars give energies relative to the ground state of ^{22}Mg in MeV. *T* is the total isobaric spin. (*After M. D. Cable et al., Beta-delayed two-proton decays of ^{22}Al and ^{26}P, Phys. Rev., C30:1276–1285, 1984*)

The high-energy proton decay from the analog state is detected (but not the preceding positron decay). *See* ANALOG STATES.

Beta-delayed two-proton emission. The large decay energy of ^{22}Al shown in Fig. 3 led to the nuclide being the first discovered (1983) in which the phenomenon of beta-delayed two-proton radioactivity could be observed. Interest centers on whether this additional branch in the analog-state decay is that of a proton pair being simultaneously emitted or of the sequential emission of the two protons. The mechanism is now known to be sequential emission of a proton (p_1) to a highly excited state in ^{21}Na followed by a second proton (p_2) to the final nuclide ^{20}Ne. This decay mode has also been established in the decays of six other light nuclei up to ^{39}Ti.

Proton radioactivity. This decay provides direct nuclear structure information and is the single-step emission of a proton from either the ground state or a long-lived isomeric state. As in alpha decay, the decay rate is determined by quantum-mechanical tunneling through the coulombic and angular-momentum barriers.

The first proton radioactive nucleus discovered (1970) was a 247-ms isomer of ^{53}Co, a nuclide that lies fairly close to the valley of stability. Ground state proton emission was first observed in 1981 in the decay of ^{151}Lu. Subsequent studies have observed this decay mode in 13 elements from antimony to bismuth.

Two-proton radioactivity. Direct two-proton radioactivity has been predicted since 1960 but has yet to be experimentally observed. This simultaneous emission of two protons from the nuclear ground state, or perhaps from a many-particle isomeric state, is expected to occur in nuclei near the proton dripline. Should this decay be observed, its mechanism would be of considerable interest, since this phenomenon could be compared with the tunneling of electrons between metals in superconducting and normal states, and thus could be said to be a nuclear Josephson effect. *See* JOSEPHSON EFFECT; SUPERCONDUCTIVITY.

Joseph Cerny

Bibliography. D. A. Bromley (ed.), *Treatise on Heavy Ion Science*, vol. 8, 1989; J. Cerny and A. M. Poskanzer, Exotic light nuclei, *Sci. Amer.*, 238(6):60–72, June 1978; J. H. Hamilton and J. Maruhn, Exotic atomic nuclei, *Sci. Amer.*, 255(1):80–89, July 1986; D. N. Poenaru (ed.), *Nuclear Decay Modes*, 1996; M. de Saint Simon and O. Sorlin (eds.), *International Conference on Exotic Nuclei and Atomic Masses*, 1995; B. M. Sherrill and D. J. Morrissey (eds.), *Enam 98: Exotic Nuclei and Atomic Masses*, AIP Conf. Proc. 455, 1999.

Exotic viral diseases

Viral diseases that occur only rarely in human populations of developed countries. However, many of these diseases cause significant human morbidity and mortality in underdeveloped areas, and have the proven capacity to be transported to population centers in developed countries and to cause explosive outbreaks or epidemics. Most of the exotic viruses are zoonotic, that is, they are transmitted to humans from an ongoing life cycle in animals or arthropods; the exception is smallpox.

Smallpox may be the most exotic pathogen on Earth, since it was successfully eradicated by a worldwide program launched in 1967, and now exists only in a few laboratories. Eradication is not possible for most other exotic viruses, however, since their life cycles include animal hosts. Control programs that may include vaccinations, insect control, or quarantine must be used in areas affected by outbreaks.

There are a number of important diseases caused by exotic viruses. (1) Yellow fever, caused by a mosquito-borne virus, particularly threatens population centers in tropical areas. (2) Venezuelan equine encephalitis is caused by a mosquito-borne virus that often infects humans. (3) Rift valley fever had a treacherous outbreak in Egypt in 1977. (4) Tick-borne encephalitis occurs in limited geographic areas where human habitats overlap with those of small mammals that normally serve as hosts to the causative virus. (5) Crimean hemorrhagic fever is caused by a tick-borne virus but is capable of transfer from human to human contact. (6) Rabies is found principally in dogs and many wild mammals. (7) Lassa fever is caused by a virus endemic to a single rodent species of west Africa. (8) Hemorrhagic fever with renal syndrome is caused by a rodent-borne virus and found in Russia and elsewhere. (9) Marburg and Ebola hemorrhagic fevers, usually lethal, are caused by two viruses of unknown natural history.

Control of these diseases is often very difficult because of the lack of detailed knowledge about the natural history of the viruses in their natural animal hosts, and because of the difficulty of controlling natural populations of alternative hosts such as insects or rodents. *See* ANIMAL VIRUS; ARBOVIRAL ENCEPHALITIDES; RABIES; SMALLPOX; YELLOW FEVER.

Frederick A. Murphy

Experiment

The test of a hypothesis under controlled conditions. The experiment is one of the distinctive tools of the scientist. It enables the scientist to put questions to nature and receive answers. These answers lead to new problems whose solutions require more complex experiments (seeking smaller differences), improved techniques, detailed plans, and better analysis of results.

Results of experiments are usually recorded numerically. Statistical methods are used to reduce the data to summary forms with estimates of the magnitude of such effects as averages, variances (dispersal of the data around the average), and the relationships between the variables being measured. Associated with these estimates are tests of significance that permit the researcher to go beyond the samples and make inferences about the characteristics (parameters) of the populations from which the samples are drawn. Statistics helps the researcher understand the magnitude of the imperfections of the experimental data. *See* BIOMETRICS; STATISTICS.

Experimental error. Results of experiments are affected not only by the treatments (experimental procedures whose effects are being measured and compared), but also by certain extraneous variations that tend to mask the effect of these treatments. Two main sources of experimental error which must be distinguished are (1) an inherent quality of the experimental material (the material on which the treatments are acting) and (2) lack of uniformity in the conduct of the experiment or failure to standardize techniques.

The presence and cause of experimental error need not concern the investigator, provided the results are sufficiently accurate to permit definite conclusions. Often, however, the results of experiments are greatly influenced by experimental errors, making decisions difficult.

Planning experiments. Inferences that can be made from results depend upon the way the experiment is conducted; thus planning of an experiment must include a detailed description of the proposal, together with such considerations as the following.

Objectives of the experiment. The purpose of the experiment must be clearly defined in terms of questions to be answered, hypotheses to be tested, specifications to be met, or effects to be estimated. The statement should indicate the extent to which inferences from the data will be applied.

Usually a sample is used. The size (that is, the number of units analyzed) of the sample needed depends upon the accuracy desired for the results, the variability of the experimental material, the errors of measurements, the magnitude of the differences to be measured, and the time and money available for the study. The results of any piece of research are no better than the sample used. A sample is interesting only if it furnishes information about the population to which the conclusions are to be applied.

Description of the experiment. After the objectives have been determined, the researcher should describe the five following aspects of the experiment.

1. *Treatments*. The treatments whose effects are to be tested and measured are selected. It is important to understand how these treatments will help reach the objectives of the experiment. The choice of treatments may also have a substantial effect upon the precision of the experiment. Often it is desired to test many varieties, levels of chemicals, sources of vitamins, or kinds of spray material. These treatment combinations make up single-factor experiments. Striking gains in precision may be achieved in factorial experiments, where several factors are investigated simultaneously. Such experiments are useful in exploratory work to determine whether factors have any effect, to find interactions among factors, and to permit recommendations that apply to a wide variety of conditions.

2. *Experimental material*. The experimental material is chosen and its inherent variability is evaluated. It is useless to try to get the chemical determination to check within 2% when there is a 20% variation among samples of the source material. At the other extreme, it may be unwise to select uniform material because the responses obtained may not apply to the regular, unselected material.

The term experimental unit denotes the group of material to which a treatment is applied in a single trial of the experiment. The unit may be a patient in the hospital, an electronic tube, a group of pigs, a tree, or a missile. It is the characteristic of such units that they produce different results even when subjected to the same treatment.

3. *Size of the experiment*. This factor helps determine the sensitivity of the experiment, that is, the size of the detectable differences between the effects of treatments. One method of improving sensitivity is to increase the size of the experiment. This can be done by using more replications, that is, by increasing the number of times a complete set of treatments is run, or by using larger experimental units. Whatever the source of variation, replication of the experiment decreases the error associated with the difference between the average effects of two treatments, provided the precaution of randomization within replications has been used and inclusion of more experimental units or larger ones has not introduced additional variation. Statistics and the scientist's experience can assist in determining the efficient size for the experiment.

4. *Experimental techniques*. The principal objectives of good techniques are to secure uniformity in applying treatments, to impose sufficient control over external influences so that every treatment

produces its effect under comparable and desirable conditions, to provide unbiased measures of the effects of treatments, and to prevent gross errors. Refining the techniques helps to reduce the experimental variation.

5. *Related variables*. These variables, which predict to some extent the performance of the experimental units, if measured and used, will increase the efficiency of the experiments. Analysis of covariance enables the experimenter to estimate from the data the extent to which the results were influenced by variation in these supplementary variables.

Experimental designs. Experimental error often can be minimized by choosing an efficient design, a set of rules for allocating treatments to experimental units. Each restriction imposed by the experimental design has a definite purpose. Of the many types of designs available, only the basic ones and a few of the variations are to be mentioned.

Completely randomized. In this type the treatments are allocated to experimental units entirely by chance. These units should be handled in random order at all stages of the experiment at which order is likely to affect the results. This design is desirable for laboratory research, especially in physics, chemistry, bacteriology, or experimental cookery, in which mixing produces quantities of homogeneous material that can be tested under uniform conditions.

Randomized complete block. In this design the experimental units are in compact or homogeneous groups, each group containing enough units for all treatments. Usually several groups (replications) are needed to give an estimate of the effects being tested. During experimentation a uniform technique should be employed for all units in a group. The experimental error arises from variation within groups.

Latin square. The treatments are grouped into replications in two ways, thus providing an opportunity for two reductions in error by skillful planning. The location of leaves on a plant and the difference between plants may determine the basis for grouping.

Incomplete block. This design is used for experiments in which a large number of treatments is to be investigated and the number of homogeneous experimental units which can be grouped is small. These designs are composed of blocks or groups of experimental units smaller than a complete replication.

An example of how designs are selected to fit into experimental situations is given. In tests of mosquito repellents which involve exposure of treated arms to mosquitoes, the block consists of two arms of a subject at one time. To test six repellents the design in the **table** provides for each of five subjects to submit his two arms to treatments three times.

Incomplete block design

	Individuals				
Day	A	B	C	D	E
1	1 2	1 3	6 2	3 6	4 5
2	6 5	4 6	4 1	5 1	2 3
3	3 4	5 2	3 5	2 4	6 1

Every possible pair of treatments occurs once on some individual on one day. Repellents 1 and 2 are assigned at random to the right and left arms of the individual A during the first test day. On the second day the arms are submitted to repellents 6 and 5. After 3 test days, the arms of each individual have been submitted to all six repellents.

Incomplete block plans, such as the one illustrated by the table, often are 20–100% more sensitive than randomized complete block designs; that is, smaller differences can be determined in an equivalent number of trials.

Split-plot. This design is used for experiments in which an extra factor is introduced by dividing each experimental unit into two or more parts. Any of the plans presented above could have each unit divided into parts permitting the use of an extra factor. In industrial experimentation frequently one series of treatments requires rather large bulks of experimental material, such as types of furnaces for the preparation of alloys. Another series, such as the molds into which the alloy is poured, can be compared through use of smaller units.

Factorial experiment. In factorial experiments more than one factor is included in the treatment combinations. The quantitative or qualitative factors may be either independent or not independent.

If factors are independent, these factorial experiments may be arranged in completely randomized, complete or incomplete block designs. As the number of factors or levels of each factor increase, the number of treatment combinations increases rapidly. When the number of homogeneous units in groups is limited, incomplete block designs are desired. For factorial experiments, a special series of incomplete block designs has been developed where information on certain treatments is sacrificed in order to secure better information on other treatments or where higher-order interactions are completely or partially confounded with blocks.

If quantitative factors are used regardless of independence, one thinks of the yield or response as a function of the level of the variables. Sequential experimentation is used to estimate the optimal point on a response surface and to explore the nature of this surface near this optimum.

The limitations of experimental materials, laboratory facilities, or time and convenience of investigators often dictate the specifications of the design, which is a logical plan. An adequate experimental design involves not only a satisfactory plan for conducting the experiment but also includes appropriate methods for evaluating results.

Analysis. The analysis of data calls for careful selection from the available efficient and flexible set of statistical methods. The use of these methods requires judgment and an understanding of the basic assumptions involved. It should be emphasized that statistical analysis cannot increase the validity of data. The accuracy of mathematical inferences from data must be limited by the precision, accuracy, and adequacy of the measurements of the observations. If

the experiment has been conducted so that treatments are confounded with extraneous effects, statistics cannot give reliable estimates of treatment effects. For well-planned and well-executed experiments, the sample of observations will provide an unbiased estimate of treatment effects, with measures of the uncertainty of these estimates.

Interpretation. The statistician should help the experimenter plan his experiment so that the results will answer definite questions. The experimental results should be presented in concise and organized form so that others can draw accurate conclusions. Drawing conclusions without the help of statistics is similar to building a suspension bridge without knowledge of the tensile strength of cables. A perfectly safe bridge can be built without such knowledge, but it will be unnecessarily expensive. Gertrude M. Cox

Bibliography. M. N. Das and N. C. Giri, *Design and Analysis of Experiments,* 2d ed., 1986; C. R. Hicks, *Fundamental Concepts in the Design of Experiments*, 5th ed., 1999; H. E. Klugh (ed.), *Statistics: Essentials of Research,* 3d ed., 1986; M. Lentner and T. Bishop, *Experimental Design and Analysis,* 1986.

Expert control system

A type of intelligent control system which can emulate the reasoning procedures of a human expert in order to generate the necessary control action. Expert control systems seek to incorporate knowledge about control system design, practical operations, and abnormal system recovery plans to automate tasks normally performed by experienced control engineers (the experts). Techniques relating to the field of artificial intelligence are usually used for the purpose of acquiring and representing knowledge and for generating control decisions through an appropriate reasoning mechanism. As it operates essentially on a knowledge base, an expert control system is often referred to as a knowledge-based control system. One of the most important benefits associated with the use of an expert control system is the inherent capability of the system to deal with uncertainty in information. Information provided to these systems need not be complete or precise. It can be general, qualitative, or vague. These properties do not cause difficulty because, like humans, expert control systems possess functionalities to perceive, reason, infer, and deduce new information. They can learn, gain new knowledge, and improve their performance through experience.

Motivation. Conventional control systems operate, explicitly or implicitly, on the basis of a mathematical model of the system to be controlled. However, for a complex system, a mathematical model representative of all important aspects of the system either does not exist or provides at most an imperfect local approximation to the actual system. When the model is inaccurate, the performance of the control system may not be acceptable. Furthermore, conventional control systems often need a complete set of data to be available, including sensory information and parameter values, in order to produce control actions. When these data are unavailable or unreliable, possibly due to failure of system components, the control system cannot function adequately. Therefore, in complex engineering systems, where great quantities of interacting numeric, symbolic, and quantitative information must be handled, and abnormal or emergency situations frequently arise, "hard" or "crisp" control with conventional control systems becomes inadequate. Human experts may be more suited to operate these complex systems, but unfortunately this is true only over a limited time span. People are quickly affected by fatigue, stress, emotion, and other environmental factors, which influence their performance to varying degrees. Therefore, for complex systems, intelligent control based on expert systems is useful to provide an efficient mimic of human expert judgment and control, yet with a robust level of proficiency and consistency unattainable with human experts.

Components. The structure of expert control systems may differ slightly due to the different application domains and the correspondingly different requirements. However, the principal components almost certain to be present in all expert control systems are the knowledge source, the database, the inference engine, the control algorithms, and the interface between the expert control system and humans (see **illus.**).

Central to an expert system is the knowledge source (or knowledge base), which contains knowledge in a specific domain (control in this context). This knowledge consists of domain-specific facts and heuristics, usually in the form of rules or frames, useful for solving problems in the domain. The knowledge source is implemented as an identifiable and separate part of the program. The integrity and correctness of the information within the knowledge source is the key to obtaining correct results or

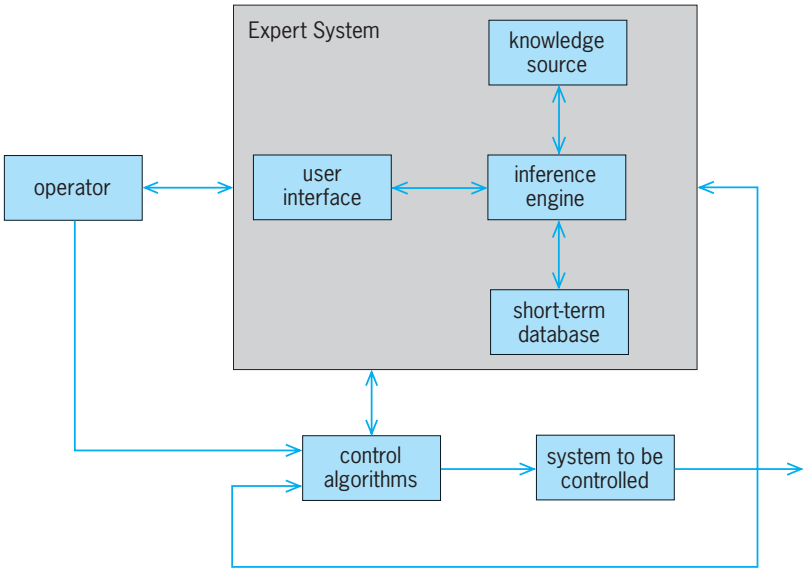

Basic architecture of an expert control system.

solutions from the expert systems.

The database is a short-term memory component which contains the current problem status, inference states, and the history of solutions to date for reference purposes.

The inference engine operates on the information from the knowledge source, from the associated database, or from the user; guides the search process according to the programmed strategy and search algorithms; and uses the inferencing mechanisms, usually in the forms of rules of logic, to solve problems, arrive at conclusions, and activate the final control actions.

The control algorithms are the tools for the expert systems to perform the final control action. A rich library of control algorithms, with advanced features, is usually made available for the expert system to configure and select the appropriate one for commissioning. The algorithms may reside physically in the field controllers outside the expert system, and the control parameters may be downloaded dynamically from the expert system.

The remaining basic component is the user interface, which is provided for the user to interact with the overall system, browse through the knowledge source, edit rules, and perform many other interactive tasks.

Although not strictly a principal component, an expert system shell is an important tool in the development of expert systems. The shell is basically a computer program used to develop expert systems. Often a shell also has provisions for changing the reasoning processes of the inference engine to adapt to the specific problem. Expert system shells differ significantly from each other and offer the user a wide variety of capabilities.

A well-developed expert system will also have a degree of self-awareness or self-knowledge which allows reasoning and explanation of the decisions to the user. This ability is one of the most valuable features of expert systems. The user can request a complete trace for a consultation, and an explanation as to how the conclusion was inferred and why any particular information was needed. This feature allows the verification and validation of expert systems and improves their overall reliability and efficiency.

Applications. Expert systems have found application in control systems, design and engineering, computerized tasks, manufacturing, finance, science and medicine, geological work, and training. One prominent successful application of expert control systems is in process control. Most process control systems are typical large-scale complex engineering systems, where multiple control loops (there can be several thousand in an enterprise), with interacting and time-varying dynamics, are run concurrently. The processes are often subject to unpredictable material variations, extraneous disturbances, and interprocess loading effects. The control system must maintain an efficient level of performance in the face of these difficulties. This is one application domain naturally akin to the functions of expert systems. In process control, the expert control system typically performs the following functions:

1. *Process operations and monitoring*: Supervise the normal operation of the system (process) and field controller, monitor the relevant signals, and take corrective actions when necessary.
2. *Diagnostics*: Examine possible failure or fault of the system components, replace these faulty components, or revise control algorithms to maintain the necessary performance of the system. *See* FAULT ANALYSIS.
3. *Control configuration*: Select suitable control strategies for specific situations to adapt to the variation of the system parameters and the environment.
4. *Scheduling and planning*: Coordinate the capabilities or individual controllers within the system to optimize production or increase efficiency. *See* ARTIFICIAL INTELLIGENCE; CONTROL SYSTEMS; EXPERT SYSTEMS; PROCESS CONTROL.

T. H. Lee; K. K. Tan; C. C. Hang; K. C. Tan; L. C. Woon

Bibliography. K. J. Åström, J. J. Anton, and K.-E. Årzèn, Expert control, *Automatica*, 22:276–286, 1986; C. W. de Silva and T. H. Lee, Knowledge-based intelligent control, *Meas. Control*, 164:102–113, 1994; T. H. Lee, Q. G. Wang, and K. K. Tan, A knowledge-based approach to dead-time estimation for process control, *Int. J. Control*, 61:1045–1072, 1995.

Expert systems

Methods and techniques for constructing human-machine systems with specialized problem-solving expertise. The pursuit of this area of artificial intelligence research has emphasized the knowledge that underlies human expertise and has simultaneously decreased the apparent significance of domain-independent problem-solving theory. In fact, a new set of principles, tools, and techniques have emerged that form the basis of knowledge engineering.

Expertise consists of knowledge about a particular domain, understanding of domain problems, and skill at solving some of these problems. Knowledge in any specialty is of two types, public and private. Public knowledge includes the published definitions, facts, and theories that are contained in textbooks and references in the domain of study. But expertise usually requires more than just public knowledge. Human experts generally possess private knowledge that has not found its way into the published literature. This private knowledge consists largely of rules of thumb or heuristics. Heuristics enable the human expert to make educated guesses when necessary, to recognize promising approaches to problems, and to deal effectively with erroneous or incomplete data. The elucidation and reproduction of such knowledge are the central problems of expert systems.

Importance of expert knowledge. Researchers in this field suggest several reasons for their emphasis on knowledge-based methods rather than formal representations and associated analytic methods. First, most of the difficult and interesting problems do not have tractable algorithmic solutions. This is

reflected in the fact that many important tasks, such as planning, legal reasoning, medical diagnosis, geological exploration, and military situation analysis, originate in complex social or physical contexts, and generally resist precise description and rigorous analysis. Also, contemporary methods of symbolic and mathematical reasoning have limited applicability to the expert system area; that is, they do not provide the means for representing knowledge, describing problems at multiple levels of abstraction, allocating problem-solving resources, controlling cooperative processes, and integrating diverse sources of knowledge in inference. These functions depend primarily on the capacity to manipulate problem descriptions and apply relevant pieces of knowledge selectively. Current mathematics offers little help in these tasks.

The second reason for emphasizing knowledge is pragmatic: human experts achieve outstanding performance because they are knowledgeable. If computer programs embody and use their knowledge, they too attain high levels of performance. This has been proved repeatedly in the short history of expert systems. Systems have attained expert levels in several tasks: mineral prospecting, product configuration, chemical structure elucidation, symbolic mathematics, chess, medical diagnosis and therapy, electronics analysis, manufacturing control, and credit analysis, among others.

The third motivation for focusing on knowledge is the recognition of its intrinsic value. Knowledge is a scarce resource whose refinement and reproduction creates wealth. Traditionally, the transmission of knowledge from human expert to trainee has required education and internship periods ranging from 3 to as much as 20 years. By extracting knowledge from human beings and transferring it to computable forms, the costs of knowledge reproduction and exploitation can be greatly reduced. At the same time, the process of knowledge refinement can be accelerated by making the previously private knowledge available for public test and evaluation.

In short, expert performance depends critically on expert knowledge. Because knowledge provides the key ingredient for solving important tasks, it reflects many features characteristic of a rare element: it justifies possibly expensive mining operations; it requires efficient and effective technologies for fashioning it into products; and a means of reproducing it synthetically would be "a dream come true." *See* DATA MINING.

Distinguishing characteristics. Expert systems differ in important ways from both conventional data-processing systems and systems developed by workers in other branches of artificial intelligence. In contrast to traditional data-processing systems, artificial intelligence applications exhibit several distinguishing features, including symbolic representations, symbolic inference, and heuristic search. In fact, each of these characteristics corresponds to a well-studied core topic within artificial intelligence. Often a simple artificial intelligence task may yield to one of the formal approaches developed for these core problems. Expert systems differ from the broad class of artificial intelligence tasks in several regards. First, they perform difficult tasks at expert levels of performance. Second, they emphasize domain-specific problem-solving strategies over the more general, "weak" methods of artificial intelligence. Third, they employ self-knowledge to reason about their own inference processes and to provide explanations of justifications for the conclusions they reach. As a result of these distinctions, expert systems represent an area of artificial intelligence research with specialized paradigms, tools, and system-development strategies.

Accomplishments. The first laboratory examples of expert systems appeared during the latter half of the 1970s, the first commercial applications appeared in 1981–1983, and by 1988 the number of applications in daily use worldwide was estimated to exceed a thousand. Since 1988, the number of applications has grown considerably. During this period, expert systems demonstrated relevance to a broad class of industrial and commercial problems, particularly order entry and configuration, field service, eligibility analysis and candidate screening, and process control. In addition, many companies claimed to have developed both cost-saving and strategic types of applications.

The following were among the first applications to demonstrate the practical utility of expert systems: PROSPECTOR discovered a molybdenum deposit whose ultimate value will probably exceed $100 million; XCON (originally R1) configured customer requests for computer systems; DENDRAL supported hundreds of international users daily in chemical structure elucidation; CADUCEUS embodied more knowledge of internal medicine than any human (approximately 80% more) and could correctly diagnose complex test cases that baffled experts; and PUFF integrated knowledge of pulmonary function disease with a previously developed domain-independent expert system for diagnostic consultation and routinely provided expert analyses at a California medical center.

Among later applications, the following illustrate the diversity and importance of commercial expert systems: ESPM remotely monitored the error logs of thousands of computers for signs of impending disk failures and scheduled appropriate maintenance preventively; ACE performed a function similar to ESPM for scheduling maintenance of telephone cables; BRUSH accepted product performance requirements from an engineer and automatically designed, drew, and documented the appropriate electric motor parts; MIX used expert rules of thumb to blend ingredients into finished products that satisfied customer requirements and production constraints; and CHARLIE collected physical vibration data from rotating equipment and diagnosed all types of mechanical problems. In each of these applications, the expert systems met the high performance requirement of equaling or surpassing the capabilities of highly trained human experts.

Undoubtedly, most applications that developers label expert systems would not generally meet this

Generic categories of knowledge engineering applications

Category	Problem addressed
Interpretation	Inferring situation descriptions from sensor data
Prediction	Inferring likely consequences of given situations
Diagnosis	Inferring system malfunctions from observables
Design	Configuring objects under constraints
Planning	Designing actions
Monitoring	Comparing observations to plan vulnerabilities
Debugging	Prescribing remedies for malfunctions
Repair	Executing a plan to administer a prescribed remedy
Instruction	Diagnosing, debugging, and repairing students' knowledge weaknesses
Control	Interpreting, predicting, repairing, and monitoring system behaviors

same high performance standard. Instead, companies have developed many so-called competence systems that automate more-routine decision-making. The appeal of the expert systems approach for these applications derives from three factors. First, the rule-based formalism of expert system shells makes it easy to express simple judgmental systems. Second, many knowledgeable people want to automate and distribute their particular know-how. Third, the computing equipment required to run the development tools and applications has become nearly ubiquitous.

Types of systems. Most of the knowledge-engineering applications fall into a few distinct types (see **table**).

Interpretation systems infer situation descriptions from observables. This category includes surveillance, speech understanding, image analysis, chemical structure elucidation, signal interpretation, and many kinds of intelligence analysis. An interpretation system explains observed data by assigning symbolic meanings to them that describe the situation or system state accounting for the data. *See* CHARACTER RECOGNITION; COMPUTER VISION.

Prediction systems present the likely consequences of a given situation. This category includes weather forecasting, demographic predictions, traffic predictions, crop estimations, and military forecasting. A prediction system typically employs a parametrized dynamic model with parameter values fitted to the given situation. Consequences which can be inferred from the model form the basis of the predictions. *See* WEATHER FORECASTING AND PREDICTION.

Diagnosis systems predict system malfunctions from observables. This category includes medical, electronic, mechanical, and software diagnosis. Diagnosis systems typically relate observed behavioral irregularities with underlying causes by using one of two techniques. One method essentially uses a table of associations between behaviors and diagnoses, and the other method combines knowledge of system design with knowledge of potential flaws in design, implementation, or components to generate candidate malfunctions consistent with observations.

Design systems develop configurations of objects that satisfy the constraints of the design problem. Such problems include circuit layout, building design, and budgeting. Design systems construct descriptions of objects in various relationships with one another and verify that these configurations conform to stated constraints. In addition, many design systems attempt to minimize an objective function that measures costs and other undesirable properties of potential designs. This view of the design problem can subsume goal-seeking behavior as well, with the objective function incorporating measures of goal attainment. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Planning systems design actions. These systems specialize in problems with the design of objects that perform functions. They include automatic programming, robot, project, route, communication, experiment, and military planning problems. Planning systems employ models of agent behavior to infer the effects of the planned agent activities. *See* ROBOTICS.

Monitoring systems compare observations of system behavior to features that seem crucial to successful plan outcomes. These crucial features, or vulnerabilities, correspond to potential flaws in the plan. Generally, monitoring systems identify vulnerabilities in two ways. One type of vulnerability corresponds to an assumed condition whose violation would nullify the plan's rationale. Another kind of vulnerability arises when some potential effect of the plan violates a planning constraint. These correspond to malfunctions in predicted states. Many computer-aided monitoring systems exist for nuclear power plant, air traffic, disease, regulatory, and fiscal management tasks, but few autonomous expert systems have been reported publicly. Many companies are using applications of this sort either in pilot applications or in advisory capacities.

Debugging systems prescribe remedies for malfunctions. These systems rely on planning, design, and prediction capabilities to create specifications or recommendations for correcting a diagnosed problem. Computer-aided debugging systems exist for computer programming in the form of intelligent knowledge base and text editors, but none of them qualify as expert systems.

Repair systems develop and execute plans to administer a remedy for some diagnosed problem. Such systems incorporate debugging, planning, and execution capabilities. Expert systems participate with human technicians and more traditional computing applications in the domains of automotive, network, avionic, and computer maintenance, as well as others.

Instructional systems incorporate diagnosis and debugging subsystems that specifically address student behavior. These systems construct a hypothetical description of the student's knowledge. From this model they diagnose weaknesses in the student's knowledge and identify an appropriate remedy. Finally, they select or plan a tutorial interaction

intended to convey the remedial knowledge to the student. Such systems, called intelligent tutoring systems or knowledge-based tutors, typically include additional expert systems that allow the instructional system to solve problems assigned to the student. This domain expertise capability allows the tutor to demonstrate problem-solving skills, answer questions, and compare expert knowledge and problem solving to the inferred model of the student's knowledge and problem solving.

An expert control system adaptively governs the overall behavior of a system. To do this, the control system must repeatedly interpret the current situation, predict the future, diagnose the causes of anticipated problems, formulate a remedial plan, and monitor its execution to ensure success. Problems addressed by control systems include air traffic control, business management, battle management, mission control, and others. Prototype systems developed as part of the U.S. Strategic Computing Initiative have illustrated these capabilities in two different kinds of applications. In the autonomous land vehicle, vision, planning, and real-time control functions were combined to drive a test vehicle at slow speed on both paved streets and open terrain. Although this system achieved unprecedented results in the field of mobile robotics using artificial intelligence techniques, none of the individual components achieved a very high level of expertise, measured by human standards. In the second program, a collection of independent but cooperating expert systems combined to perform the function of a pilot's associate, an artificial "backseater" who would replace the second human in future one-seater aircraft. This kind of application combines independent expert systems addressing separate areas of expertise, including situation assessment, mission planning, tactical planning, system status, pilot–vehicle interface, and overall mission management. Collectively, the pilot's associate subsystems incorporate most of the generic categories of knowledge engineering applications. *See* CONTROL SYSTEMS; INTELLIGENT MACHINE.

Systems components. The ideal expert system (see **illus.**) contains a language processor for problem-oriented communications between the user and the expert system; a blackboard for recording intermediate results; a knowledge base comprising facts plus heuristic planning and problem-solving rules; an interpreter that applies these rules; a scheduler to control the order of rule processing; a consistency enforcer that adjusts previous conclusions when new data or knowledge alters their bases of support; and a justifier that rationalizes and explains the system's behavior.

The user interacts with the expert system in problem-oriented languages, usually some restricted variant of English, and in some cases through means of a graphics or structure editor. The language processor mediates information exchanges between the expert system and the human user. Typically, the language processor dissects, or parses, and interprets user questions, commands, and volunteered information. Conversely, the language processor formats information generated by the system, including answers to questions, explanations and justifications for its behavior, and requests for data.

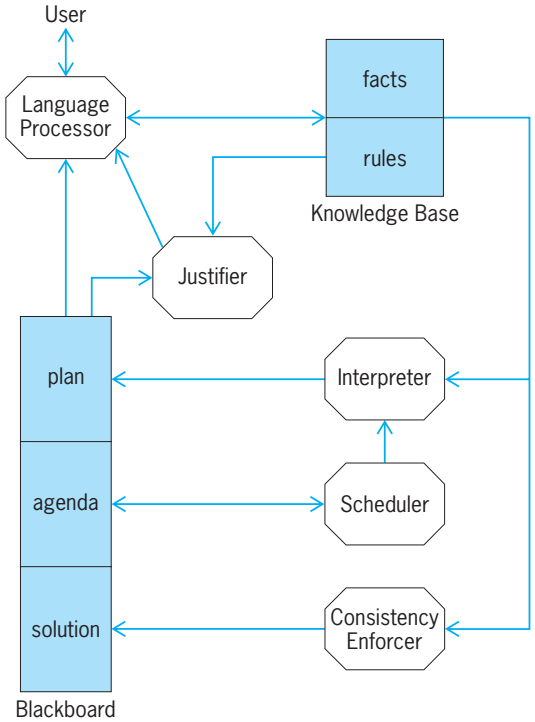

Anatomy of an ideal expert system. No existing system contains all the components shown, but one or more components occur in every system.

Existing expert systems use a wide variety of language processors, depending on the type of user. In many commercial applications, the ideal format for dialog is the business form, depicted graphically on a cathode-ray tube. In military applications, both natural language and annotated maps provide more appropriate languages for system–user dialog. In other applications, the most common languages of dialog are graphic illustrations, such as computer-aided design (CAD) drawings, or semiformal languages, such as part-assembly specifications. These language processors increasingly employ stand-alone interface packages developed specifically for reuse by multiple applications with similar dialog requirements. *See* COMPUTER GRAPHICS; PROGRAMMING LANGUAGES.

The blackboard is a global database that records decisions, which the expert system manipulates. Every expert system uses some type of intermediate decision representation, but only a few explicitly employ a blackboard for the various types of ideal expert system decisions. Three types of decisions that are recorded on the blackboard are plan, agenda, and solution elements. Plan elements describe the overall or general attack the system will use to solve the current problem, including current plans, goals, problem states, and contexts. For example, a plan may recommend processing all low-level sensor data first; formulating a small number of most-promising hypotheses; refining and elaborating each of these hypotheses until the best one emerges; and, finally,

focusing exclusively on that candidate until the complete solution is found. This kind of plan has been incorporated in several expert systems. The agenda elements record the potential actions awaiting execution. These generally correspond to knowledge-base rules that seem relevant to some decision placed on the blackboard previously. The solution elements represent the candidate hypotheses and decisions that the system has generated thus far, along with the dependencies, called links, that relate decisions to one another.

The scheduler maintains control of the agenda and determines which pending action should be executed next. Schedulers may utilize considerable abstract knowledge, such as "do the most profitable thing next" and "avoid redundant effort." To apply such knowledge, the scheduler needs to prioritize each agenda item according to its relationship to the plan and to other solution elements. To do this, the scheduler generally needs to estimate the effects of applying the potential rule.

The interpreter executes the chosen agenda item by applying the corresponding knowledge-base rule. Generally, the interpreter validates the relevance conditions of the rule, binds variables in these conditions to particular solution blackboard elements, and then makes rule-prescribed changes to the blackboard. Interpreters of this sort are generally written in LISP because of its facilities for manipulating and evaluating programs, but other languages are also suitable. *See* PROGRAMMING LANGUAGES.

The consistency enforcer attempts to maintain a consistent representation of the emerging solution. This may take the form of likelihood revisions when the solution elements represent changing hypothetical diagnoses and when some new data are introduced. Alternatively, the enforcer might implement truth maintenance procedures when the solution elements represent changing logical deductions and their truth-value relationships. Most expert systems use some kind of numerical adjustment scheme to determine the degree of belief in each potential decision. This scheme attempts to ensure that plausible conclusions are reached and inconsistent ones avoided.

The justifier explains the actions of the system to the user. In general, it answers questions about why some conclusion was reached or why some alternative was rejected. To do this, the justifier uses a few general types of question-answering plans. These typically require the justifier to trace backward along blackboard solution elements from the questioned conclusion to the intermediate hypotheses or data that support it. Each step backward corresponds to the inference of one knowledge-base rule. The justifier collects these intermediate inferences and translates them to English for presentation to the user. To answer "why not" questions, the system uses a heuristic variant of this technique. It can identify a possible chain of rules that would reach the questioned conclusion but that did not apply because the relevance condition of some rule failed. The justifier explains the system's decision to reject a possible conclusion by claiming that such failed conditions impede all reasoning chains that can support the conclusion.

Finally, the knowledge base records rules, facts, and information about the current problem that may be useful in formulating a solution. While the rules of the knowledge base have procedural interpretations, the facts play only a passive role. The language PROLOG is the most widely available general-purpose programming language that explicitly supports this fact-plus-rule view of programs. However, relatively few expert systems are developed directly in PROLOG. Instead, most developers adopt a higher-level knowledge engineering tool or expert system shell as their programming language. These shells offer additional features for enhanced representations, inference, and control. Some of the most popular enhancements include expressive notations and built-in mechanisms for uncertainty; procedural blocks for more explicit control; classes and type hierarchies for organizing semantic relationships among categories, specializations, and instances; and change-driven control mechanisms to support modeling, reactive systems, and message-passing styles of programming. The portrait of the ideal expert system ignores these embellishments which, although occasionally important, often obscure the common essence of the expert systems technology.

Agents. Agents can vary from simple programs, lacking knowledge, to sophisticated autonomous real-time control systems that provide both deliberative reasoning (for example, including sophisticated planning and scheduling) and fast real-time response. The vast majority of simple utilitarian agents, such as off-line browsers that automatically download web sites, or agents that monitor web pages for changes, are not knowledge-based. Most knowledge-based agents are still experimental.

Agents are variously defined as autonomous programs; programs that communicate with an agent communication language (for example, KQML); mobile programs that can travel from one computer to another; or programs that collaborate with humans in solving tasks, offloading tedious tasks, and acting as monitors and liaisons with other agents. Agents can also be viewed as higher-level architectural components, providing distributed open architectures in which the agents can be registered, requests can be brokered, and different organizational structures can be used to solve higher-level tasks. In this architectural view, they are distinguished from objects by their persistence, autonomy, communications capabilities, and behavior. Another key aspect of most agent definitions is the agent's dynamic and uncertain environment. Agents also may or may not have graphical representations, commonly called avatars or personas.

Internet agents may make use of information retrieval techniques such as term frequency and inverse document frequency analyses. In these approaches, the frequency of terms in documents

relative to their overall frequency in a document corpus is compared to the same measure for the terms in a query (for retrieval) or in another document (to gauge similarity). Most Internet agents rely more on information retrieval techniques such as these or machine learning techniques (such as collaborative filtering, a variant of the nearest-neighbor algorithm) than the reasoning and representation techniques used in most knowledge-based applications. These Internet agents can provide recommendations, provide personalized searches and web page recommendations, monitor multiple sources of information, negotiate, and so forth, using the Web as the medium they are situated in for perception and action. Stand-alone robotic or software agents designed for entertainment may have a different architecture comprising reactive structures or simple learning architectures (for example, neural networks). *See* INTERNET; NEURAL NETWORK; WORLD WIDE WEB.

Knowledge-based agents overlap with expert systems: When a knowledge-based program is a real-time autonomous program or if its software architecture comprises a multiple-agent system, these knowledge-based systems are agents or, conversely, the agents are also knowledge-based systems.

An example of a knowledge-based agent is Guardian, an on-line agent designed to operate in an intensive-care unit to monitor and control a ventilator unit. It has sensors and effectors, and operates in a dynamic, uncertain environment. It also has models of human biological responses, disease processes, and intervention strategies to handle crises that arise. It can be viewed as an agent that incorporates multiple expert systems to handle different tasks (diagnosis, monitoring, and repair). Guardian has been tested only in a simulated environment.

The New Millennium Remote Agent (NMRA) is an autonomous expert system and agent that operates in deep space. Developed by the National Aeronautics and Space Administration, it controls the functions of an entire spacecraft. It includes planning and monitoring functions and operates entirely autonomously. Like Guardian, it has both reactive and deliberative capabilities. It has a procedural executive for fast real-time response coupled with a model-based diagnosis system and planner that can be invoked by the executive for recovery from plan execution errors. The model-based diagnosis system is used to infer the state of the satellite components and to provide input to the procedural executive.

Falling between heavy-weight knowledge-based agents such as Guardian and NMRA and simple utilitarian agents that do not rely on knowledge-based techniques, there are agents that rely on a changing knowledge base to provide useful assistance, but do not constitute expert systems in the traditional sense. Rather than expert knowledge, they provide informed guidance. Examples are guides to web sites and virtual guides to software and other repositories.

Another area of overlap between expert systems and agents is the use of pedagogical agents in intelligent tutoring systems. Such systems operate in a dynamic, uncertain environment, and are, by definition, knowledge-based, having knowledge of the subject matter and the student, and some qualitative model of the teaching process. Many intelligent tutoring systems use graphical personas as pedagogical agents. For example, one program uses a talking character to teach botany to middle-school students, and another uses a virtual-reality instructor to teach troubleshooting and maintenance of Navy propulsion systems. In both cases the tutor agents are knowledge-based: in the first case, the tutor knows about botany, and in the second case, the tutor knows about equipment operation and the current state of its controls and internal parts.

Expert systems and agents can also overlap architecturally: agents can be used as higher-level building blocks to build knowledge-based systems. ARCHON is one example of a real-time expert system that is architecturally composed of agents. It provides a methodology for building a multiagent community that incorporates existing control systems. ARCHON has been operationally deployed in an electrical transmission system in northern Spain. It wraps existing on-line control and expert systems as agents so they can communicate, share knowledge, and act in concert where this was not previously possible.

Similarly, expert systems can be used as building blocks of agents. The NMRA agent can be viewed as a collection of expert systems to handle different spacecraft systems. Intelligent tutoring systems and pedagogical agents typically have one expert system to provide domain expertise, and may incorporate additional expert systems for student modeling and pedagogical control.

Frederick Hayes-Roth; William R. Murray

Evaluation. Expert systems can be evaluated from three perspectives: subjectively by evaluating the system's stated requirements and level of usability, technically by evaluating how well the knowledge base was built, and empirically by evaluating the system's effect on user performance.

Subjective evaluations. The goal of subjective evaluations is to assess the system from the perspective of potential users and sponsors. To accomplish this, users should be involved in specifying the system's performance requirements. Users are increasingly involved in requirements validation, that is, evaluating whether the system is being built to achieve the right performance criteria.

Users are also involved in system evaluation through judging the adequacy of system performance and usability. User involvement may take the form of seeing a general system demonstration, systematically walking through each system feature, or actually using the expert system to solve decision problems representative of their actual job.

Technical evaluations. Static testing focuses on detecting logical problems in the knowledge base without actually executing it by using test cases. Static testing may be performed manually or with automated static testors. The more prevalent logic problems that can be detected include redundant rules,

subsumed rules, conflicting rules, circular rules, unnecessary "if" conditions, and unreferenced or illegal attribute values.

Dynamic testing uses experts who have not participated in development and test cases to evaluate the knowledge base's functional completeness and predictive accuracy. Ideally, there will be correct answers for the test cases; otherwise, it is necessary to rely on the judgment of the evaluation expert or, preferably, the consensus judgments of a group of experts. If the knowledge base's predictive accuracy is acceptable, the evaluation experts will be unable to tell whether they or the system generated the best conclusions. This approach is referred to as a Turing test.

Empirical evaluations. Experiments help to generalize from a test sample to the larger population. There are two kinds of experiments. The first tests the system against objective benchmarks representing performance constraints. If the system passes, it proceeds further; if it fails, it undergoes further development or is set aside.

The second kind of experiment is a factorial design where one or more factors are systematically varied as the independent variables, and system performance is measured by one or more dependent variables. Dependent variables include objective measures (such as performance and speed), observational measures (such as regarding how the system is used), and subjective measures (such as user confidence in the solution). The goal is to show that people using the expert system perform better than those who have access only to their typical tools.

Ideally, field experimentation would be used to assess if the expert system significantly improves performance in the targeted setting. When the sample size and randomization requirements of true experiments are not possible, quasi-experimental designs can be used. These include time-series designs, where the performance of an organizational unit is measured before and after receiving the system; multiple time series designs, which use a control group that does not receive the system; and nonequivalent (and nonrandomized) control-group designs, which rely on statistical techniques to compare performance improvement in expert-system and control groups. *See* ARTIFICIAL INTELLIGENCE.

Leonard Adelman

Bibliography. L. Adelman, *Evaluating Decision Support and Expert Systems*, 1992; J. C. Giarratano and G. Riley, *Expert Systems: Principles and Programming*, 3d ed., 1998; U. Gupta (ed.), *Validating and Verifying Knowledge-Based Systems*, 1991; B. Hayes-Roth, An architecture for adaptive intelligent systems, *Artif. Intell.*, 75:195–240, 1995; F. Hayes-Roth and N. Jacobstein, The state of knowledge-based systems, *Commun. ACM*, 37(3):27–39, 1994; W. L. Johnson (ed.), *Proceedings of the 1st International Conference on Autonomous Agents*, ACM Press, 1997; S. C. Shapiro (ed.), *Encyclopedia of Artificial Intelligence*, 2 vols., 2d ed., 1992; D. A. Waterman, *A Guide to Expert Systems*, 1986.

Explosive

A substance containing a large amount of stored energy that can be released suddenly, thereby converting the substance into compressed gases or numerous fragments that expand with great force or velocity. An explosion is a sudden expansion of matter into a much larger volume than it formerly occupied, or a sudden increase in the pressure exerted by confined matter.

Explosions

Examples of explosions are expansion of highly compressed gas and throwing of fragments following brittle rupture of a pressurized container; rapid combustion of a fuel-air mixture in an internal combustion engine; deflagration of smokeless powder in a gun followed by sudden expulsion of the projectile and gases from the muzzle; rapid deflagration of a gas-producing pyrotechnic mixture such as black powder; detonation of a charge of chemical explosive such as TNT; release of nuclear energy in a mass of nuclear explosive such as uranium-235 with sufficient rapidity to suddenly vaporize it and any adjacent material; explosive eruption of a volcano; and explosion of a star to produce a supernova. *See* INTERNAL COMBUSTION ENGINE; NUCLEAR EXPLOSION; SUPERNOVA; VOLCANO.

Blast and shock. Divergent motion, in the form of a pressure wave of high amplitude above ambient pressure (that is, high overpressure), propagates outward from the source of an explosion into the surrounding medium, typically air, water, or rock. In those parts of the wave where the overpressure is low, the velocity of wave propagation is not much higher than the velocity of sound in the unperturbed medium. Here, the velocity of the medium itself (the particle velocity) is relatively low. In those parts of the wave where the overpressure is high, the sudden compression has increased the temperature and the sound velocity in the medium, and here the particle velocity is also relatively high. The outward velocity of a local perturbation in the wave is the sum of the sound velocity and the particle velocity. Therefore, perturbations and the energy they carry move outward faster in the high-pressure parts of the wave, thereby changing the shape of the wave. Consequently, if the amplitude of the wave is sufficiently high, it finally evolves into one or more shock waves of characteristic shape (**Fig. 1***a* and *b*) regardless of the initial wave form.

A shock wave has a shock front with a typical thickness of about 3×10^{-7} ft (1×10^{-7} m) which travels at supersonic velocity and at which the pressure jumps almost instantaneously from ambient pressure to a value high above ambient. Immediately ahead of the shock front, the medium is usually at rest, while immediately behind it the medium is moving forward with a high particle velocity. The kinetic energy of this motion, per unit volume of the medium, is called the dynamic pressure. The overpressure of a shock wave decays with the distance traveled, not only due to divergence, which causes its energy to be spread

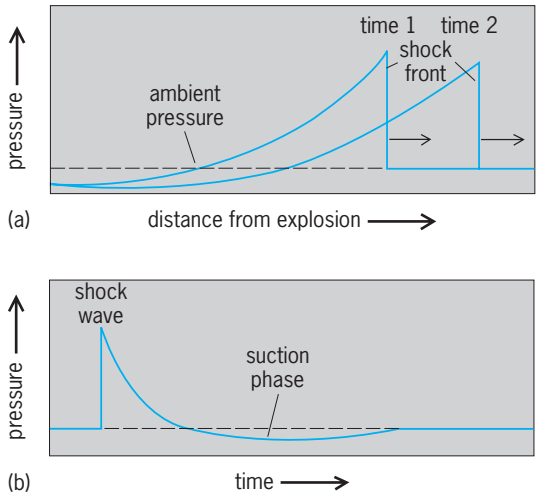

Fig. 1. Pressure waves from explosions. (*a*) Pressure versus distance near the shock front at two times. (*b*) Pressure versus time in air at one distance.

over a surface of increasing size, but also due to an irreversible degradation of its energy to heat resulting from the discontinuous jump in pressure at the shock front. Its velocity of propagation approaches that of sound (about 1050 ft/s or 320 m/s in air, and about 5000 ft/s or 1520 m/s in water) as the overpressure approaches zero. *See* SHOCK WAVE.

In air, the particle velocity of the pressure wave is sensed as a sudden blast of wind, so the wave is sometimes called a blast wave whether or not it is also a shock wave.

Effects of scale. The smaller the amount of energy released, the shorter must be the time and the smaller the volume in which it is released in order to obtain the transient effects of an explosion. Conversely, as the amount of energy increases, explosion effects can be produced with larger durations and larger volumes for the energy release. The release of energy may produce a pressure wave that never develops into a shock wave because the wave has too slow a rise time, too low an overpressure, or is in a medium that rapidly degrades much of the energy of deformation into heat. Such an energy release may nevertheless produce motion that is sufficiently sudden to produce damage or other mechanical effects associated with explosions, and may therefore be classified as an explosion.

Under certain conditions, it is possible to use relatively simple relationships to describe the behavior of explosions over a very large range of energies. Necessary conditions are that the effects of radiation have negligible effect on the explosion phenomenon of interest and that the material in which the explosion occurs not be changed. Then the desired simple relationships are obtained when the independent variable is the scaled time $tW^{1/3}$ or the scaled distance $DW^{1/3}$, and the dependent variable is a gravity-independent variable, where t is the time after explosion, D the distance from the explosion, and W the explosion energy. W is usually expressed as the mass of TNT having the same energy yield as the explosion.

When gravity effects cannot be neglected, as for explosion phenomena resulting from heaving rock, soil, or water upward against gravity or against centrifugal force, similar results will also be obtained for explosions at equal values of the scaled time $tW^{-1/3}$ and scaled distance $DW^{-1/3}$, provided that equal values of tg and Dg are also maintained, g being the gravitational or centrifugal acceleration of the explosion environment.

Such scaling laws permit the effects of very large explosions to be predicted from the behavior of very small explosions. In some cases, these may be made in a centrifuge to obtain a high value of g commensurate with the small values of t and D.

Damage to structures. The damage done to a structure by an explosion or a rapid series of them can be the result of pressure exerted against the outside or inside of the structure, impulse delivered to it, or vibrational energy imparted to it.

When a shock wave in air strikes a solid surface head-on, part of the wave is reflected. The change in momentum associated with the resulting impact of the moving air on the surface can generate an overpressure that is more than twice that of the incident wave, and for very strong shocks and very rigid surfaces can be almost eight times as great. If the natural period of vibrations of a structural member is very short compared to the duration of the positive phase, the damage-causing potential of the explosion depends on the peak pressure or the reflected peak pressure appropriate for the angle of incidence of the wave on the member. This is commonly the case for very large explosions, such as nuclear explosions, or for very thin, light members, such as windowpanes. If the natural period of vibration of a member is long compared to the duration of the positive phase, the damage-causing potential of the explosion tends to depend on the positive-phase overpressure impulse per unit area.

A pressure wave can be described as a sum of sinusoidal waves of various frequencies and amplitudes. If the arriving wave is not a single pulse but a wave train with multiple peaks resulting from a series of explosions or the effects of the intervening media that transmitted it, then energy carried by the wave train when plotted as a function of frequency will have many high peaks and deep valleys. If one or more peaks of sufficient height in this power spectrum contain frequencies that match natural frequencies of vibration of the structure, the wave train may excite vibrations in the structure of sufficient amplitude to cause damage. The approximate vibration levels that buildings of various types can tolerate without damage are known, and industrial blasts near buildings are designed so as to stay well below those levels. In rock blasting, which is commonly done by detonating a number of charges with short delays between detonations, the delays are chosen to maximize rock breakage and minimize air blast and ground vibration.

Even a mild explosion that produces no easily perceptible wave in the open air can be very damaging if it occurs inside an enclosed structure, such as a

building, where rapid venting cannot occur. Most structures of this type will fail at internal pressures easily generated by the burning of a combustible dust or vapor in air.

Combustion and detonation. Chemical explosives are metastable compositions that can react rapidly within themselves to produce large quantities of hot gas. The reaction is generally initiated by heat or by mechanical deformation of the explosive that produces heat. Once initiated, the reaction may proceed through the mass of explosive at a relatively slow (subsonic) rate, called combustion, burning, or deflagration, or at a relatively rapid (supersonic) rate, called detonation. In combustion, the rate at which the reaction front advances into the unreacted explosive is governed by the conduction and convection of heat from the hot products of reaction to the unreacted explosive.

In detonation, the rate of advance of the reaction front is the supersonic velocity of a shock wave, called the detonation wave, driven by the high-pressure products of reaction. The chemical reaction is initiated in fresh material by the mechanical deformation produced by this shock wave. The velocity of a detonation wave is determined by the energy released in the chemical reaction, the equation of state of the reaction products, and the constraints of conservation of mass, momentum, and energy. A detonating explosive is classified as a primary explosive if heat alone will cause a small unconfined volume of it to detonate, and as a secondary explosive if not. Primary explosives can be detonated by an influence that will produce even a miniscule hot spot, such as contact with a small flame or hot particle, a static electrical discharge, or a relatively low level of friction or impact.

Combustion of an explosive can accelerate and transform into a detonation, and vice versa. Many of the damage effects of explosions can be produced by sufficiently rapid combustion as well as by detonation, but detonations can produce much higher peak pressures with a more sudden onset than combustion, and damage effects characteristic of such suddenly applied high pressure. *See* COMBUSTION; FLAME.

The propellants used in guns and rockets are explosives whose combustion under the conditions of use occurs in a reproducible and controlled manner without converting to a detonation, which would burst the chamber. Explosives such as propellants that combust in normal use are called low explosives; those that detonate are called high explosives. But given suitable conditions, low explosives will detonate, and high explosives will combust, even though these conditions may not occur in normal practice.

Double-base powder, the propellant most commonly used in artillery and firearms, is made from nitrocellulose with various amounts of glycerine trinitrate (nitroglycerine) and the nitrates of mono-, di-, and triethyleneglycol, plus other ingredients that modify the burning rate, prolong shelf life, reduce barrel erosion, and reduce muzzle flash and smoke. When a substantial amount of nitroguanidine is added to such propellants in order to reduce muzzle flash and barrel temperatures, they are called triple-base powders. The term powder is often used for explosives and propellants, whether or not their physical form is powder.

Modern solid rocket propellants are also often of the double-base type, but those with the highest performance may be made from high-explosive compounds such as RDX, sometimes with an added high-energy fuel (usually powdered aluminum), an oxidizer (usually ammonium perchlorate), an uncross-linked monomer, a cross-linker, and other minor additional additives. After the viscous mixture is poured into a mold or rocket casing, the monomer cross-links to form a solid elastomer, resulting in a body of high-performance solid propellant that will yield rather than fracture under the transient stresses imposed during initiation and combustion. A fracture could produce a disastrous increase in the area of the burning surface.

Although solid propellants are inherently detonable, the shock or chamber pressures required to initiate detonation in them are so high that in practice they deflagrate rather than detonate. *See* PROPELLANT.

No sharp distinction exists between energy-releasing compositions that are explosives or can produce an explosion, and those that are not and cannot. The term pyrotechnics covers a continuum of such compositions, ranging from high explosives and low explosives, through fuel/oxidant gaseous mixtures, incendiaries that burn with a flame but do not explode, to mixtures of solids that react to produce high temperatures but little or no gas or flame and no explosive effects. *See* PYROTECHNICS.

Combustion and detonation fronts move through a mass of explosive at rates typically within the ranges 3×10^{-2} to 3×10^{2} ft/s (10^{-2} to 10^{2} m/s) and 3×10^{3} to 3×10^{4} ft/s (10^{3} to 10^{4} m/s), respectively. For both modes of reaction, the velocity of propagation tends to increase with the energy of the explosive, the amount of confinement provided by the surroundings, and the diameter of the explosive mass normal to the direction of propagation. For combustion, the velocity also tends to increase with the surface area of the mass of explosive and its permeability to gas. For detonation, it increases with the density of the mass of explosive if the charge has a sufficiently large diameter.

The velocity s with which the combustion front moves into a single grain of a typical propellant follows a relationship of the form $s = AP^n$, where P is the gas pressure and A and n are constants characteristic of the propellant. The propellant is generally formulated to have a pressure index n less than 1 to prevent the burning rate from accelerating to a level that could burst the chamber. To control the burning rate, propellant grains are also shaped to maintain a nearly constant surface area as they burn (**Fig. 2**).

Gas and dust explosions. Explosions that result from the ignition of a mixture of air with fuel in a confined space are a frequent source of industrial accidents. It is possible for such explosions to be

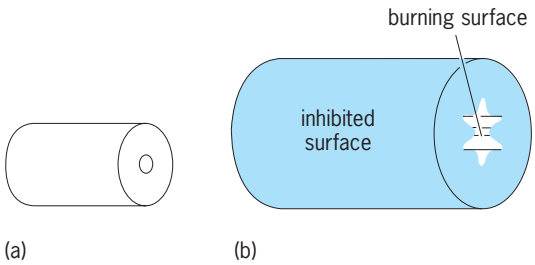

Fig. 2. Propellant grains. (*a*) A single powder grain in the form of a perforated cylinder, which may have a wall thickness (or web) from less than 0.04 in. (1 mm) to 0.25 in. (6 mm). (*b*) A rocket grain, which may be several feet or more in diameter, depending on the size of the rocket.

detonations (the knocking of an internal combustion engine being an example), but they are usually much slower, burning reactions. The damaging explosion pressure results from the confinement of the gases as they are heated by the reaction.

Environments that are potentially explosive include storage bins, tanks, exhaust ducts, and underground spaces containing combustible vapors or dusts. Such environments are often encountered in grain elevators, ships, manufacturing spaces, and coal mines, where flammable gases or powders are handled or created. Examples of gases that frequently create such a potential hazard are methane and hydrogen; examples of dusts are starch, synthetic-resin dust, coal dust, and powdered aluminum, magnesium, and zirconium. Dust and dusty air can generate static voltages and electrical discharge when they flow, particularly at low humidity. Such discharges have been sources of ignition for many dust explosions.

The maximum pressure that can be produced is limited by the amount of heat and the resulting pressure that can be produced by the combustion of all the available oxygen in the air. This is about 10 bars (1 megapascal) for conventional fuels in air initially at 1 bar (10^5 pascals). The volume in which the explosion occurs and the surfaces against which the pressure acts are usually very large, so the total energy and force mobilized in such an explosion can be very damaging. Ease of ignition and burning rate of dusts increase with decreasing particle size and vary greatly with their composition. Similarly, the ease of ignition of gas mixtures varies greatly with the nature of the gases and their proportions, ignition being possible at all only within certain limits of composition. The most sensitive methane-air mixture requires about 20 times the energy for ignition as the most sensitive hydrogen-air mixture. An air-methane explosion in a gassy part of a coal mine can produce a pressure wave and flame front that first disperses coal dust settled on beams, ledges, and floors and then ignites it, producing an explosion that can propagate for long distances through the mine. The ignitability of coal dust is reduced by diluting it with powdered limestone spread where the dust collects. The ignitability of methane is reduced by maintaining good ventilation. Other practices that reduce the possibility of a mine explosion include use of explosives formulated to be able to detonate in an air-methane mixture without igniting it, use of water-filled bags rather than coal dust to stem the charges, and use of electrical devices designed to function without producing sparks or flame. *See* MINING.

Types of Explosives

For reasons of safety as well as economics, large explosive charges are usually made of explosives that are relatively difficult to initiate. The amount of very sensitive explosive that is easy to initiate is kept to a minimum by using only enough of it to reliably initiate a larger charge of a less sensitive explosive, and using that to initiate a still larger charge of a still less sensitive explosive.

Either failure to detonate or premature explosion can be disastrous. Therefore, it is important that explosives be formulated, manufactured, and used so as never to explode accidentally and to detonate with high reliability when desired. Achieving this balance is an overriding consideration in all work with explosives, both industrial and military.

Industrial explosives. Black powder (a mixture of charcoal, sulfur, and potassium nitrate) was the only industrial explosive until Alfred Nobel's introduction of liquid nitroglycerine for rock blasting in 1863. Black powder is still used extensively in pyrotechnic devices. The modern era of industrial explosives began with Nobel's invention of the detonator in 1865 and of the first dynamites and blasting gelatin in the period 1867–1875. These inventions were a great and lasting contribution to the world's industrial development because dynamite had much greater safety and shattering power than black powder, and much lower cost.

Uses of industrial explosives include blasting ore, coal, and rock in mining and construction, generating vibrations in seismic prospecting for oil and gas, stimulating and perforating gas and oil wells, bonding sheets of dissimilar metals to each other, and synthesizing industrial diamonds.

Rock, ore, and coal are blasted by drilling them with a pattern of holes, filling most of each hole with explosive containing a detonator, plugging the exit of the hole with a material such as drill cuttings or bags of water (stemming), and detonating the charges in a rapid and carefully timed sequence that results in efficient rock breakage and low vibration of nearby structures.

In contrast to military explosives, industrial explosives tend to have larger critical diameters, lower densities, lower detonation velocities, and lower explosion pressures; and to have more complex compositions and lower cost. Most military explosives are rigid solids, whereas most industrial explosives are formulated to be plastic, pumpable, or free-flowing to permit filling the cross section of deep, rough, and irregular holes in rock.

Industrial explosives usually contain separate fuel and oxidizer ingredients in intimate combination. Usually, they also contain a sensitizer to aid in the initiation and propagation of detonation. Other ingredients may be used to increase or decrease density, to

increase explosion energy, to prevent the detonation from igniting methane in coal mines, to provide plasticity, pumpability, or flowability, to prevent setting or stiffening during storage or at low temperature, to prevent separation of ingredients and chemical instability during storage, and to provide resistance to desensitization by water, low temperature, hydrostatic pressure, or transient pressure from explosions in nearby holes. Industrial explosives are usually packaged in bags or cartridges of polymer film or paper, or are carried in bulk form to the blasting site, where they are blown or pumped into the holes through hoses. Sometimes they are mixed at the blasting site.

To avoid toxic fumes underground or fumes that can give secondary explosions when mixed with air, industrial explosives are usually oxygen-balanced to minimize carbon monoxide (CO), nitrogen oxides (NO_x), and hydrogen (H_2) in the reaction products. In a typical modern dynamite, nitroglycerine containing a nitrocellulose thickener is adsorbed on a mixture of ammonium nitrate and cellulosic fuel to form a cohesive, plastic mass that is packaged in paper cartridges. The composition, including the cartridge, is approximately oxygen balanced; that is, the oxygen balance is about zero.

In an oxygen-balanced composition, the amount of oxygen present is just sufficient to oxidize all the carbon to CO_2 and all the hydrogen to water (H_2O) and any metals present to their oxides. If the oxygen is insufficient to do this, the oxygen balance is said to be negative, and if it is more than sufficient, it is said to be positive. Explosives containing only carbon (C), H, nitrogen (N), and oxygen (O) exhibit their maximum energy at a slightly negative oxygen balance and a minimum production of toxic NO_x and CO at about zero oxygen balance.

Most industrial blasting is done with ANFO, a free-flowing mixture of ammonium nitrate (AN) in the form of small spheres (prills) and 6% No. 2 diesel fuel oil, sometimes with aluminum powder to increase the energy. Ammonium nitrate prills are made by the solidification of droplets of molten ammonium nitrate during free fall in a shot tower. ANFO is a very low-cost, oxygen-balanced explosive with adequate sensitivity and energy for most industrial blasting in dry holes of relatively large diameter. Prilled ammonium nitrate alone is a relatively weak explosive having a positive oxygen balance. It is too difficult to initiate and has too large a critical diameter to be a useful explosive in itself. *See* PRILLING.

Aqueous explosives are used for much blasting where the holes have small diameter or contain water, or have too thick a burden of rock for ANFO to break. These compositions contain an aqueous phase saturated with inorganic nitrates or perchlorates, one or more fuels such as ground coal, starch, oil, wax, aluminum, or sugar, and a sensitizer such as trinitrotoluene (TNT), monomethylammonium nitrate, pigmentary aluminum, microscopic hollow glass spheres, or bubbles. This fluid mixture is thickened, gelled, or emulsified through the use of other ingredients. The resulting semisolid consistency prevents separation of the ingredients, provides water resistance, and prevents loss of explosive through fissures in the rock. Aluminum (Al) fuel, when used in sufficient quantities to leave little or no water in the explosion products, can be particularly effective in giving aqueous compositions a very high energy

Combustion and detonation parameters for various explosives

Composition	Charge density, Mg/m^3	Detonation velocity, ft/s*	Chapman-Jouguet pressure, kPa	Approximate critical diameter, in. (mm)
$2H_2 + O_2$ (moles of gas at 1 bar and 20°C)		9.28×10^3	1.80×10^3	
$2H_2 + O_2 + 5N_2$ (moles of gas at 1 bar and 20°C)		5.97×10^3	1.44×10^3	
Nitroglycerine (liquid at 20°C)	1.60	25.3×10^3	$2.53 \times 10^{7\dagger}$, $2.51 \times 10^{7\ddagger}$	0.08 (2)
TNT (liquid at 100°C)	1.44	21.4×10^3		1.6–2.4 (40–60)
TNT (particulate)	0.95	15.9×10^3		0.08–0.9 (2–23)
TNT (cast)	1.62	22.1×10^3	$2.10 \times 10^{7\dagger}$, $2.23 \times 10^{7\ddagger}$	0.47–1.4 (12–35)
RDX	1.77	28.8×10^3	$3.38 \times 10^{7\dagger}$, $3.48 \times 10^{7\ddagger}$	0.2 (5)
PETN	1.76	27.1×10^3	$3.35 \times 10^{7\dagger}$, 3.32×10^7	
Tetryl	1.71	25.8×10^3	$2.60 \times 10^{7\ddagger}$	0.43 (11)
Lead azide	3.8	18×10^3		0.02 (0.5)
Black powder	1.0	4.43×10^3		
40% special gelatin dynamite	1.59	16.4×10^3	8.10×10^6	0.55 (14)
ANFO made with prilled AN	0.85	15.6×10^3	4.82×10^6	4 (100)
Amatol 80/20	1.6	17×10^3		3 (80)
Cyclotol 75/25	1.76	27.2×10^3		0.16 (4)
Pentolite 50/50	1.70	24.7×10^3	$2.55 \times 10^{7\ddagger}$	0.3 (7)
Composition B-3	1.72	25.9×10^3	$2.87 \times 10^{7\dagger}$	0.16 (4)
HBX-3	1.84	23.4×10^3		
PBX-9407	1.60	26.0×10^3	$2.87 \times 10^{7\dagger}$, $3.00 \times 10^{7\ddagger}$	

*1 ft/s = 0.3048 m/s.
†Measured.
‡Calculated.

from the reaction below, with all carbon appearing

$$2Al + 3H_2O \rightarrow Al_2O_3 + 3H_2 + 8.86\ MJ/kg$$

in the products as CO. *See* EMULSION; GEL.

Military explosives. Most military explosives are simple compositions formulated for high energy density, loading in munitions plant, and long storage life. Most of them are based on explosive chemical compounds that incorporate both oxidizer and fuel components in the same molecule. Many of these compounds are used for special purposes in industrial explosives also (see **table**).

TNT, which was once used extensively as a military explosive, either alone or in Amatol (80% AN, 20% TNT), is a relatively insensitive and inexpensive explosive of moderate energy. TNT melts at 177°F (81°C), which allows it and compositions containing it to be easily melt-cast. It is mainly used in admixture with more powerful explosives such as RDX and PETN. Examples are Cyclotol 75/25 (75% RDX, 25% TNT), composition B-3 (60% RDX, 40% TNT), HBX-3 (31% RDX, 29% TNT, 35% Al, 5% wax), and Pentolite 50/50 (50% PETN, 50% TNT; **Fig. 3**).

RDX (hexagen) is a fairly insensitive explosive of high density and melting point with the highest detonation pressure of any compound in common use. It is often combined with an elastomeric or polymeric binder (Fig. 3). An example is PBX-9407 (94% RDX, 6% vinyl chloride–chlorotrifluoroethane copolymer). HMX is an analog of RDX, having an eight-membered rather than a six-membered ring, and has a higher density and therefore a higher detonation pressure.

PETN has moderate sensitivity, a small critical diameter, and a high detonation pressure, which makes it useful in detonating cord (Fig. 3). Detonating cord is flexible line having an explosive core covered sometimes with lead but usually with a layer of fibers and plastic or wax. It is used to transfer a detonation from one point to another, both in military devices and in commercial blasting.

Tetryl, like PETN, is a relatively sensitive explosive, but it is easier to press into pellets for use as booster charges. Booster charges are detonated with a still smaller charge of an even more sensitive explosive, such as lead azide, and in turn they detonate the main charge consisting of a relatively insensitive explosive.

HNS (Fig. 3) is a secondary explosive, having unusually high stability at elevated temperatures and the unusually small critical diameter of 0.020 in. (0.5 mm). TATB (1,3,5-triamino-2,4,6-trinitrobenzene) is another secondary explosive. It is used in munitions because it has an unusually high resistance to transition from deflagration to detonation in a fire.

Mercury fulminate (Fig. 3) is the primary explosive that was used in the first dynamite detonators (blasting caps). It has since been replaced by explosives such as lead azide, lead styphnate (2,4,6-trinitroresorcinate), and lead mononitroresorcinate, which have greater stability at high temperatures or are less subject to failure when pressed to high density.

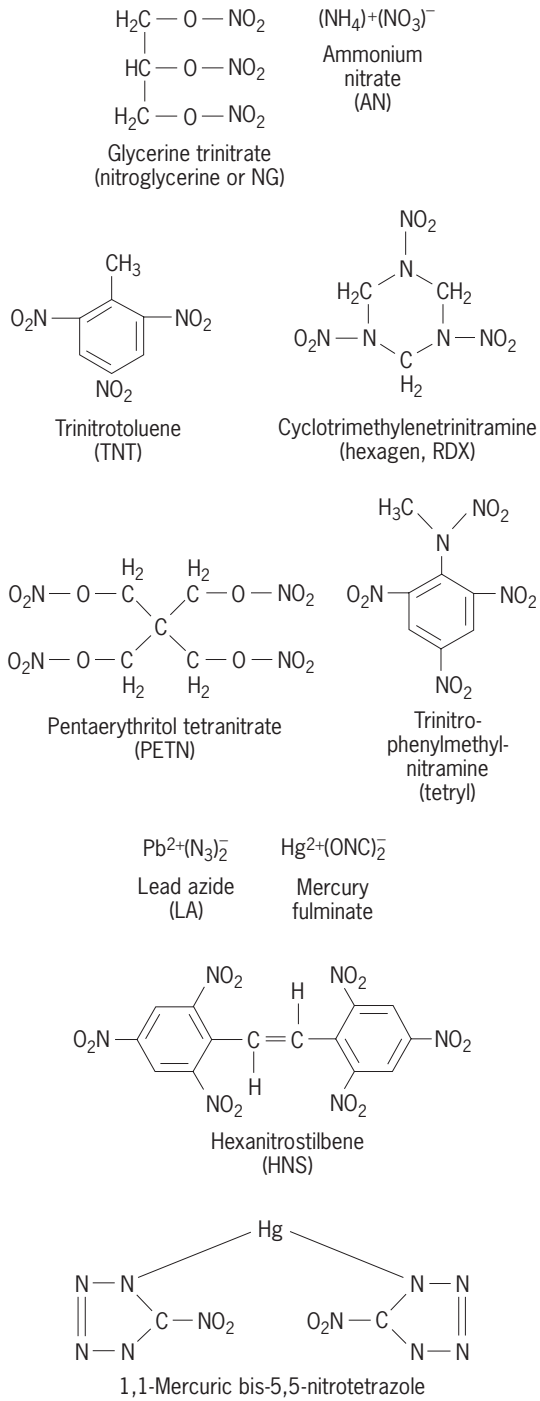

Fig. 3. Structural formulas of some common explosives.

Lead azide (Fig. 3) is a primary explosive much used in detonators. It has good thermal stability and detonates promptly without undergoing a perceptible stage of deflagration, but it is subject to hydrolytic instability. When stored in contact with copper or brass, it can react to form dangerously unstable copper azide.

1,1′-Mercuric bis-5,5′-nitrotetrazole (DXN-1; Fig. 3) has high thermal stability and detonates promptly upon application of heat without exhibiting a well-defined burning zone as do mercury fulminate and lead styphnate.

Detonators. Large charges are usually made of explosives that are relatively insensitive. The amount of very sensitive, more dangerous, and more costly primary explosive required for initiating detonation is kept to a minimum. Only enough is used in an assembly known as a detonator to reliably initiate a larger charge of a less sensitive explosive. This in turn initiates a still larger charge of a still less sensitive explosive.

Most detonators employ increasing violence to produce an explosion accompanied by a shock wave that initiates explosion of the small primer charge of a pyrotechnic composition. The latter's flame contacts a charge of a very sensitive primary explosive, such as lead azide, in contact with another small base charge of a powerful secondary explosive, such as PETN or RDX.

The base charge of the detonator is placed in or against the receptor charge. Upon being heated or shocked, for example, by a fuse or an electrically ignited match-head, the charge of primary explosive detonates. This in turn detonates the base charge. The resulting shock wave or impact of missiles detonates the receptor charge.

In other detonators, the charge of primary explosive is detonated by impact, friction, an intense flash of light, or an intense surge of electric current through a bridgewire or across a spark gap. In still others, the primary explosive is omitted and the secondary explosive is detonated by igniting it while it is strongly confined in a thick-walled metal tube, or by the impact of a thin metal plate driven by an electrically ignited charge of secondary explosive.

Detonators are frequently provided with one or more safety devices, such as an electrical shunt across the lead wires or a temporary barrier between the detonator and the receptor charge. These safety devices are removed only when preparing to fire the charge.

In mining and civil engineering, blasts usually use many separate charges, each in a separate borehole and each with a separate detonator cap. These charges are detonated in a timed sequence chosen to reduce the vibration of nearby buildings, increase rock breakage, and control the throw of the broken rock. The timing of the detonations may be controlled by timing the pulse of electricity used to initiate each detonator, or by providing a pyrotechnic delay train or an internal electronic time delay circuit for each detonator. The delay train may be in the detonator or in a line of detonating cord or shock tube, if that is used to initiate the detonator.

Detonators can be accidentally detonated by rough handling, electric current, or heat, and, though deceptively small, have sufficient energy to maim. Therefore, extensive precautions are required in their transportation, storage, and use.

Primers, boosters, igniters, and squibs. The charge that is initiated by a detonator (if used to detonate a main charge that is still larger and more insensitive) is called a booster or primer. Bags of a water gel explosive, cartridges of dynamite, or cast cylinders of pentolite are used as boosters in industrial blasting. Pressed pellets of tetryl are often used as boosters in military ordnance.

Sometimes the term primer is also used loosely to mean detonator. The term so used is not to be confused with the same term more correctly used to mean a flame-producing pyrotechnic charge or an assembly used to ignite a fuse or a larger mass of gun propellant, or to provide the heat to detonate a small charge of primary explosive in a detonator. A primer may be initiated by percussion, stab action, friction, or heat.

The terms igniter and squib are variously used to mean heat or flame-producing devices for igniting rocket propellants and other deflagrating pyrotechnics.

The amount of primer used is usually about 0.0035 oz (0.1 g) or less. A common example is the tiny globule that may be seen on the ends of the fuse wire post in a photographic flashbulb; this material serves to ignite the aluminum foil.

Although primer compositions deflagrate rapidly, they do not detonate. In fact, a detonating action would be undesirable because it would tend to blow away rather than to ignite the charge. The total burning time of a primer is about 500 microseconds.

Primer mixtures sensitive primarily to friction may consist simply of potassium chlorate as the oxidizer and antimony sulfide as the fuel. Such a mixture produces hot particles as well as gases. Ground glass or silicon carbide may be added to increase the frictional effect, and sulfur may be added to make the mixture more sensitive. Also, very fine grained black powder (meal powder) is sometimes added. The ingredients are made into a paste with a small amount of gum arabic that binds them together when the paste dries.

Percussion primer for small-arms ammunition usually contain a primary explosive such as mercury fulminate in addition to the materials mentioned above. However, primary explosives are not essential; they may be replaced by other materials such as lead thiocyanate. Lead compounds seem to be preferred, probably because of the hot particles of lead oxide that they produce. Lead styphnate (2,4,6-trinitroresorcinate) is used as a percussion-cap component.

In detonators, the primer is often of a match-head composition. In fact, the formulation of primers is quite similar to that of matches and pyrotechnics; and ingredients such as potassium chlorate, sulfur, and antimony sulfide are common in all applications.

Fuse. This is conduit that leads a hot or high-pressure chemical reaction front from one place to another. The word is not to be confused with "fuze" (sometimes spelled fuse), which is a mechanism used to fire an item of explosive or pyrotechnic ordnance after it has been launched. A fuze is usually a mechanical or electrical device but may contain a fuse.

A typical deflagrating fuse is miner's safety fuse, used in commercial blasting and comprising a train of black powder sheathed in woven and waxed fabric and made in types burning at 120 and 90 s/yd

(131 and 98 s/m). Firecracker fuse is a train of black powder wrapped in tissue paper. Gasless delay fuse comprises a core of reactants [for example, lead oxide (Pb_3O_4) plus boron (B)] that react exothermally to produce mainly solid products, encased in a short tube (often of lead), made so as to give relatively accurate overall burning times ranging between 10^{-3} and 10^{+2}; it is used in delay detonators for commercial blasting with multiple charges, and also is used in fuzes.

A detonating fuse contains a train of detonating explosive. The most common type, ordinary detonating cord, comprises a train of PETN or RDX sheathed with wax- or asphalt-impregnated fiber and detonating at 17,000–26,000 ft/s (5200–7900 m/s), and is used for initiating multiple charges in commercial blasting. Shock tube is a flexible, plastic tube on the inner wall of which is an adherent coating of very fine HMX and aluminum powders. Initiation with a shotgun shell primer results in a shock wave supported by a dust explosion, which travels through the tube at 1950 m/s (6400 ft/s), leaving the tube unbroken and producing no noise. It is used in commercial blasting for detonating multiple charges, particularly when airborne noise is to be minimized. D. Linn Coursen

Bibliography. A. Bailey and S. G. Murray, *Explosives, Propellants, and Pyrotechnics*, 2d ed., 2000; International Society of Explosives Engineers, *The Blasters Library*, 1971–1993; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., vol. 9, 1998; B. Lewis and G. von Elbe, *Combustion, Flames, and Explosions of Gases*, 3d ed., 1987; P. A. Persson, R. Holmberg, and J. Lee, *Rock Blasting and Explosive Engineering*, 1993; R. A. Sickler, *Explosive Principles: An Essential Guide to Understanding Explosives and Detonations*, 1992; U.S. Department of Defense and Department of Energy, *The Effects of Nuclear Weapons*, 3d ed., 1977, reprint 1983.

Explosive forming

The shaping or modifying of metals by means of explosions. The explosives may be of either the detonating or deflagrating type. Explosive gas mixtures or stored gas at high pressure may also provide the motive power (**Fig. 1**). *See* EXPLOSIVE.

Most types of explosive forming involve a mold into which a flat sheet of metal is pressed by the explosion. The metal is stretched uniformly by the explosive impulse. Even welds in the original blank survive the deformation without damage. The advantages of explosive forming over conventional forging methods accrue especially in the case of intricate shapes of which only a few items are required. Tooling costs are low.

Cold welds can be made between dissimilar metals by driving the two parts together under explosive impact. In other applications of explosive-forming methods, powders are pressed into solid billets.

In a quite different application, high explosives are used to cut large blocks of metal and even to split thin sheets into two layers of exactly one-half the original thickness. When an explosive charge is detonated in contact with the metal, it produces a compression wave. On reflection from a free surface, the compression wave turns upside down to produce a tension wave. Along the planes in the metal where the tension in the wave front exceeds the strength of the metal, rupture occurs.

Explosives can also be employed to extrude metal shapes and to punch hard metals with the aid of dies. Shapes produced explosively are very exact and free from the fine cracks that sometimes result when pressure is slowly applied (**Fig. 2**). Forces far exceeding those of the largest hydraulic presses can be applied by explosives.

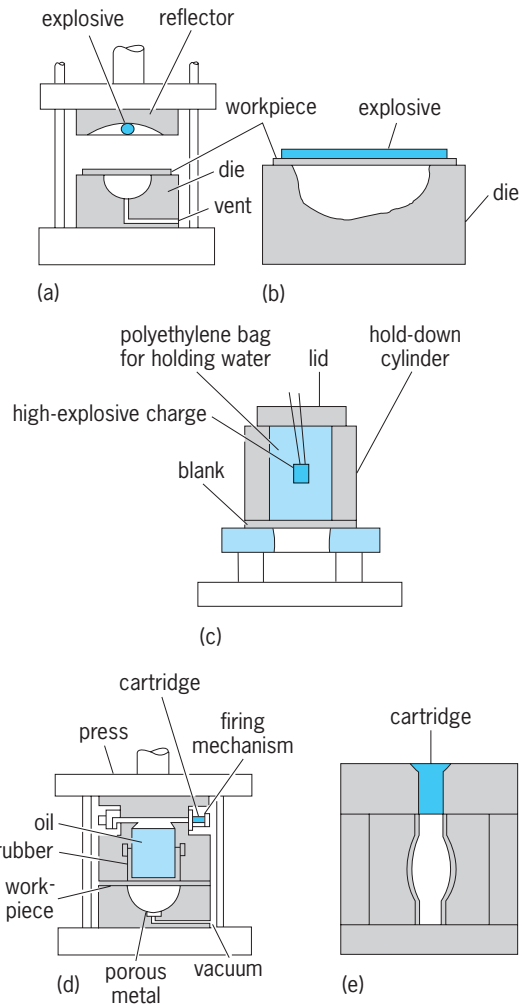

Fig. 1. Five methods of explosive forming. (*a*) Shaped charge and parabolic reflector. Shock and pressure are motive power. Reflector and charge must be shaped to direct explosive force toward workpiece. (*b*) Flat high explosive is simply placed on top of sheet which lies on female die. Forming over a punch produces wrinkles. (*c*) The hold-down cylinder is often a carton of water or a bolted ring on lower die. (*d*) Gunpowder cartridges are usually applied in a press or enclosed die. This one uses oil and rubber to distribute force. Vacuum prevents air pocket which would retard action. (*e*) Bulge forming with cartridge power is comparatively simple and safe. Such devices properly vented are less noisy than many standard forming operations. (*After Explosives form space age shapes, Steel, August 25, 1958*)

Fig. 2. Typical metal shapes produced by explosive forming. Note weld in bowl. (*National Northern Corp.*)

Metals can also be hardened under explosive impact with results that compare favorably with those from slow cold-working methods. *See* METAL FORMING.

William E. Gordon

Exponent

In mathematics, a symbol or number written to the right of and above another symbol or number to denote how many times the latter is to be multiplied by itself. For example, $7^3 = 7 \cdot 7 \cdot 7$, and $a^3 = a \cdot a \cdot a$. By use of this convenient mathematical device, the number 420,000,000,000,000 can be expressed unmistakably and more compactly as 4.2×10^{14}, which is read as, "four and two-tenths times ten to the fourteenth power." This abbreviated notation is particularly valuable in expressing the extremely large or extremely small numbers encountered in modern scientific work. For example, 0.000000143 may be expressed as

$$\frac{143}{1,000,000,000} \quad \text{or} \quad \frac{143}{10^9} \quad \text{or} \quad 1.43 \times 10^{-7}$$

Operations with exponents are governed by the following rules:

1. $x^a x^b = x^{a+b}$.
   Example, $x^2x^3 = xx \cdot xxx = x^5$.

2. $(x^a)^b = x^{ab}$.
   Example, $(x^2)^3 = (x^2)\,(x^2)\,(x^2) = x^6$.

3. $(xy)^a = x^a y^a$.
   Example, $(xy)^3 = xyxyxy = xxxyyy = x^3y^3$.

4. If $y \neq 0$, then $\left(\dfrac{x}{y}\right)^a = \dfrac{x^a}{y^a}$.

   Example, $\left(\dfrac{x}{y}\right)^3 = \dfrac{x}{y} \cdot \dfrac{x}{y} \cdot \dfrac{x}{y} = \dfrac{x^3}{y^3}$.

5. If $x \neq 0$, then $\dfrac{x^a}{y^a} = x^{a-b}$.

   Example, $\dfrac{x^5}{x^2} = \dfrac{xxxxx}{xx} = xxx = x^3$.

6. If $x \neq 0$, and $-n$ is a negative integer, then $x^{-n} = \dfrac{1}{x^n}$.

   Example, $\dfrac{x^2}{x^5} = \dfrac{1}{x^3} = x^{-3}$.

7. If $x \neq 0$, then $x^0 = 1$.

8. If p/q is any rational number expressed with positive denominator q, then $x^{p/q} = (\sqrt[q]{x})^p$.

   Example, $x^{-2/3} = \sqrt[3]{x^{-2}} = (\sqrt[3]{x})^{-2} = \dfrac{1}{(\sqrt[3]{x})^2}$.

9. If a is any irrational number, then x^a is equal to the limit of the sequence, x^{a1}, x^{a2}, $x^{a3}, \ldots$, x^{an}, where $a_1, a_2, a_3, \ldots, a_n$, etc., are the one-place, two-place, ..., n-place approximations of a.

   Example, x^π is equal to the limit of the sequence, $x^3, x^{3.1}, x^{3.14}, x^{3.141}, x^{3.1416}, \ldots$.

Any number which can be written in the form $a + ib$, where a and b are real numbers and $i = \sqrt[q]{-1}$, is called a complex number. Exponents which are complex numbers are frequently encountered. For example, any number may be expressed as a power of $e = 2.718\ldots$, the base of the system of natural logarithms. Many quantities in natural phenomena are so expressed. Where e has a complex exponent, e^{a+ib}, by rule 1, $e^{a+ib} = e^a \cdot e^{ib}$. The value of e^{ib} may be found by substituting ib for x in the expansion

$$e^x = 1 + \frac{x}{1} + \frac{x^2}{1 \cdot 2} + \cdots + \frac{x^n}{1 \cdot 2 \cdot 3 \cdots n} + \cdots$$

By separating the real and imaginary parts, the Euler equation is obtained, $e^{ib} = \cos b + i \sin b$. Thus, $e^{a+ib} = e^a(\cos b + i \sin b)$. *See* ALGEBRA.

Arnold N. Lowan; Salomon Bochner

Exposure meter

An indicating instrument used in photography to determine lens aperture and shutter speed. An exposure meter may be used either in the darkroom to determine approximate printing time for a contact print or an enlargement or, more usually, with a camera to determine exposure of film.

Exposure meters (see **illus.**) are of two basic types, photovoltaic and photoconductive. In photovoltaic meters, a selenium cell converts photons to electrons, producing a current directly proportional to received light. A sensitive microammeter indicates this current to the user. The photocell has rapid response, spectral sensitivity closely approximating that of panchromatic film, and indefinitely long life. Because the power to deflect the microammeter comes entirely from the received light, design of the exposure meter is a compromise between a

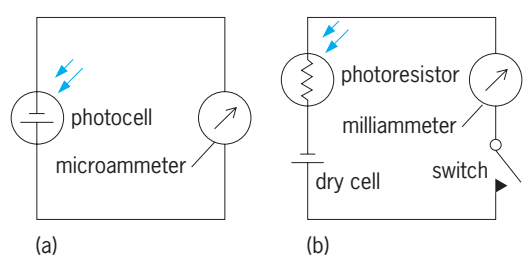

Basic exposure meter circuits. (*a*) Photovoltaic type circuit. (*b*) Photoconductive type circuit.

cell of large area and a meter movement of delicate construction. *See* PHOTOVOLTAIC CELL.

In the photoconductive meter, a cadmium sulfide cell changes conductivity in proportion to the received light. A battery (usually a mercury cell) supplies power through the cadmium sulfide cell to the meter movement. Because the received light serves to control the power from the battery, this type of instrument is about three orders of magnitude more power-sensitive than the photovoltaic type. As a consequence, cell area can be small and meter movement can be somewhat more sturdy. Spectral response is greatest in the green and the yellow portion of the spectrum and least in the blue, so a filter is built into the meter. The photoconductive cell recovers slowly from bright light; therefore, it is preferably covered when not in use. Also, a switch opens the electric circuit when the meter is not in use to conserve battery life, nominally about a year. *See* PHOTOCONDUCTIVE CELL.

For the photovoltaic meter, a honeycomb (or fly-eye) lens with acceptance angle comparable to that of a camera lens of normal focal length gathers light for the cell. A multi-iris cover may serve to reduce sensitivity for use in bright light, with a mechanically linked change in meter scale. For the photoconductive meter, a small lens controls acceptance angle; or the cell may be mounted inside the camera and thus can use the camera lens, which determines acceptance angle. As an accessory meter separate from the camera, the exposure meter may be fitted with a lens having an acceptance angle of about 3° and a corresponding viewfinder, so that the meter can be used to measure from a distance the brightness of small portions of the subject.

Used at the camera, either type of meter serves as a brightness meter to measure reflected light from the subject. Used at the subject, with a diffuser head over the exposure meter lens, the instrument serves as an illumination meter to indicate incident light. *See* PHOTOMETRY.

To aid in interpreting the indicated light as exposure conditions, the meter may carry a circular slide rule. If built into the camera, the meter movement may either operate an indicator in the field of the viewfinder, or it may control the camera iris to produce an average exposure. In any case, the user presets the rated speed of the emulsion in use into the meter. *See* CAMERA; PHOTOGRAPHY. Frank H. Rockett

Bibliography. J. Neubart, *The Photographer's Guide to Exposure*, 1988; L. Stroebel and R. D. Zakia (eds.), *The Focal Encyclopedia of Photography*, 3d ed., 1996.

Extended x-ray absorption fine structure (EXAFS)

The structured absorption on the high-energy side of an x-ray absorption edge. The absorption edges for an element are abrupt increases in x-ray absorption that occur when the energy of the incident x-ray matches the binding energy of a core electron (typically a 1*s* or a 2*p* electron). For x-ray energies above the edge energy, a core electron is ejected from the atom. The ejected core electron can be thought of as a spherical wave propagating outward from the absorbing atom. The photoelectron wavelength is determined by its kinetic energy, which is in turn determined by the difference between the incident x-ray energy and the core-electron binding energy. As the x-ray energy increases, the kinetic energy of the photoelectron increases, and thus its wavelength decreases. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; LIGHT; PHOTOEMISSION; QUANTUM MECHANICS.

The x-ray-excited photoelectron will be scattered by the neighboring atoms surrounding the absorbing atom. The portion of the photoelectron wave that is scattered back in the direction of the absorbing atom is responsible for the EXAFS oscillations. If the outgoing and backscattered photoelectron waves are out of phase and thus interfere destructively, there is a local minimum in the x-ray absorption cross section. At a higher x-ray energy (shorter photoelectron wavelength), constructive interference leads to a local maximum in x-ray absorption (see **illus.**). EXAFS thus arises from photoelectron scattering, making it a spectroscopically detected scattering method. *See* INTERFERENCE OF WAVES; SCATTERING OF ELECTROMAGNETIC RADIATION.

EXAFS typically refers to structured absorption from approximately 50 to 1000 eV or more above the absorption edge. X-ray absorption near edge structure (XANES) is often used to refer to the structure in the near (around 50 eV) region of the edge. X-ray absorption fine structure (XAFS) has gained some currency as a reference to the entire structured absorption region (XANES+EXAFS).

Form of oscillations. The energy can be given in units proportional to the inverse photoelectron wavelength [the photoelectron wavevector, or k, defined as in Eq. (1), where m_e is the electron mass, E is

$$k = \sqrt{2m_e(E - E_0)/\hbar^2} \qquad (1)$$

the incident electron energy, E_0 is the core-electron binding energy, and $\hbar$ is Planck's constant divided by 2π]. If this is done, the EXAFS oscillations take

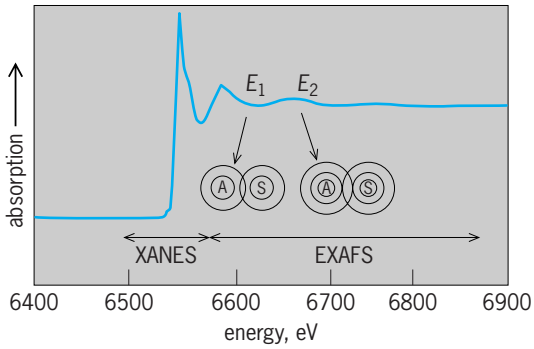

X-ray absorption spectrum for manganese, showing XANES and EXAFS regions. As the x-ray energy increases from E_1 to E_2, the interference of the outgoing and backscattered photoelectron wave (shown schematically by concentric circles around the absorbing, *A*, and scattering, *S*, atoms) changes from destructive to constructive.

on a particularly simple form. For a single absorber-scatterer pair (for example, a diatomic gas) the alternating constructive and destructive interference gives rise to regular, sinusoidal oscillations in the absorption coefficient. The fractional modulation in the x-ray absorption coefficient, $\chi(k)$, can then be described by an equation similar to (2), where the

$$\chi(k) = \sum_s \frac{N_s A_s(k)}{kR_{as}^2} \exp\left(-2k^2\sigma_{as}^2\right) \cdot \sin[2kR_{as} + \phi_{as}(k)] \quad (2)$$

right-hand side is a sum over the various types of scattering atoms.

In Eq. (2), the parameters of greatest interest are the number of scattering atoms of type s, N_s, the absorber-scatterer distance, R_{as}, and the mean-square deviation in R_{as}, σ_{as}^2. To a first approximation, these can be related to the depth of modulation (that is, the EXAFS amplitude), the frequency of the modulations, and the shape of the amplitude envelope describing the oscillations, respectively. In order to determine these structural parameters, the intrinsic scattering amplitude, $A_s(k)$, and the scattering phase shift, $\phi_{as}(k)$, which also appear in Eq. (2), need to be known. These can be obtained by fitting the EXAFS data for structurally characterized models (models where N_s, R_{as}, and σ_{as}^2 are known), or they can be calculated from first principles, typically with empirical calibration. Both $A_s(k)$ and $\phi_{as}(k)$ depend on the chemical identity of the scatterer and thus provide the information necessary to identify the scattering atom. However, both $A_s(k)$ and $\phi_{as}(k)$ depend only weakly on scatterer identity, and thus it is difficult to identify the scatterer with precision. This typically limits scatterer identity to the nearest row of the periodic table (that is, sulfur and oxygen can be distinguished, while nitrogen and oxygen cannot).

All absorber-scatterer pairs contribute to the observed oscillations, as reflected by the sum in Eq. (2). In practice, however, it is not usually realistic to refine all of the different absorber-scatterer interactions. Consequently, it is necessary to group absorber-scatterer interactions into shells. A shell is a group of similar scatterers at approximately the same distance from the absorber.

Fourier transform. Most commonly, EXAFS data are not presented in the form shown in the illustration, but as the Fourier transform of the data. The Fourier transformation separates the frequency-space signal $\chi(k)$ into its different constituent frequencies, with the peaks in the Fourier transform representing the different shells of scatterers. The Fourier transform thus resembles a radial distribution function of electron density around the absorbing atom. However, the Fourier transform peaks are shifted to lower R by about 0.5 Å (50 picometers) due to the phase shift $\phi_{as}(k)$, and the amplitude of the peaks cannot be related directly to electron density due to the effect of the amplitude terms $A_{as}(k)$. *See* FOURIER SERIES AND TRANSFORMS.

Structural information. EXAFS spectra contain structural information comparable to that obtained from single-crystal x-ray diffraction. The principal advantage of EXAFS in comparison with crystallography is that EXAFS is a local structure probe and does not require the presence of long-range order. This means that EXAFS can be used to determine the local structure in noncrystalline samples. In ideal circumstances, EXAFS data can be analyzed to determine the absorber-scatterer distance with an accuracy of around 0.02 Å (2 pm) and a precision that is substantially better (0.004 Å or 0.4 pm, and perhaps even better under carefully controlled conditions). Coordination numbers can be determined with an accuracy of around 25%. *See* X-RAY CRYSTALLOGRAPHY.

The information available from EXAFS is relatively limited. Scattering atoms can be resolved into different shells only if they are separated by around 0.15 Å (15 pm) or more and only scatterers within about 4–5 Å (0.4–0.5 nanometer) of the absorbing atom contribute to the signal. Equation (2), which describes single scattering, does not contain any angular information. In some cases, multiple scattering, in which the photoelectron is scattered off more than one atom, can give EXAFS sensitivity to geometry. However, multiple scattering is a large effect only if the structure contains a nearly linear arrangement of atoms. Consequently, EXAFS does not generally provide significant angular detail. Despite these limitations, EXAFS often provides the only available structural information for noncrystalline systems. This has made EXAFS one of the preferred structural probes in many areas, particularly for studies in materials science, catalysis, and biological systems.

James Penner-Hahn

Bibliography. J. Goulon, C. Goulon-Ginet, and N. B. Brookes (eds.), X-ray Absorption Fine Structure, *J. Phys. IV*, 1997; D. C. Koningsberger and R. Prins (eds.), *X-ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS, and XANES*, 1988; P. A. Lee et al., Extended x-ray absorption fine structure: Its strengths and weaknesses as a structural tool, *Rev. Mod. Phys.*, 53:769–806, 1981; J. Penner-Hahn, X-ray absorption spectroscopy in coordination chemistry, *Coord. Chem. Rev.*, 190–192:1101–1123, 1999; B. K. Teo, *EXAFS: Basic Principles and DATA Analysis*, 1986.

Extinction (biology)

The death and disappearance of a species. The fossil record shows that extinctions have been frequent in the history of life. Mass extinctions refer to the loss of a large number of species in a relatively short period of time. Episodes of mass extinction occur at times of rapid global environmental change; five such events are known from the fossil record of the past 600 million years. Human activity is causing extinctions on a scale comparable to the mass extinctions in the fossil record.

Nature and record. An extinction may be of two types; phyletic or terminal. Phyletic extinction occurs when one species evolves into another with time; in this case, the ancestral species can be called

extinct. However, because the evolutionary lineage has continued, such extinctions are really pseudoextinctions. In contrast, terminal extinction marks the end of an evolutionary lineage, termination of a species without any descendants. Most extinctions recorded in the fossil record and those occurring today are terminal. *See* ORGANIC EVOLUTION.

It has been estimated that 99% of all species that have ever lived are now extinct. Because of problems of fossil preservation and sampling, the measurement of extinction in the fossil record is not simple. Except in a few circumstances, the fossil record of individual species is not used. Rather, a higher taxonomic level (typically the family) is used as a crude measure of species-level extinction. This is done because of the potential of sampling error in the geologic record. The absence of a species from some geologic stratum may be the result of poor preservation of fossils, inappropriate rock types, or inadequate sampling rather than true extinction. The absence of all species within a family may be a more reliable indicator of actual extinction. Because the number of species varies from family to family, the number of family extinctions may be an inaccurate measure of the number of species extinctions. Yet, where extinctions have been studied in detail, at species, genus, and family levels, high rates of extinction at the higher taxonomic level result from high rates at the lower level. For example, the record of family extinctions since the Permian reflects the level of genus extinctions.

An additional difficulty in measuring extinctions in the fossil record is estimating their time of occurrence. Although radiocarbon dating gives precise estimates for extinctions of the past 30,000 years, the method cannot be used for older rocks and fossils. The extinction of a species (or genus or family) is marked by the horizon where it last appeared in the rock record. Such horizons can be dated with a precision of a few million years at best. Generally speaking, the farther back in the rock record, the more difficult it is to pinpoint the time of extinction.

The fossil record is best known for marine organisms. The extinctions of marine organisms over the past 570 million years were relatively minor, ranging from one to seven families per million years. The extinction rates never dropped to zero, indicating that extinction is a normal part of the history of life. Such low levels are sometimes referred to as background extinctions, those that are always occurring as a consequence of normal environmental changes or local catastrophes. Background rates of family extinction appear to have declined through geologic time. Although this may reflect an increase in average fitness through evolution, the decline in family extinction rate may also be an artifact; there has been an increase in the number of species per family since the Cambrian. Because a family becomes extinct only when all its component species become extinct, families with many species are less likely to become extinct than families with few species.

The mass extinctions of the marine fossil record

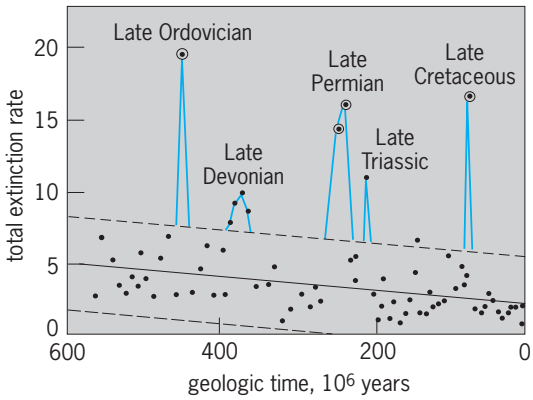

Fig. 1. Record of extinction of families of marine organisms. Spikes show times of mass extinctions. (*After D. M. Raup and J. J. Sepkoski, Jr., Mass extinction in the marine fossil record, Science, 215:1501–1503, 1982*)

are from two to five times as severe as background extinctions. The mass extinctions are indicated by the five spikes in **Fig. 1**: the Late Ordovician, Late Devonian, Late Permian, Late Triassic, and Late Cretaceous. These mass extinctions affected a variety of organisms in many different ecological settings.

Figure 2 shows the effect of these mass extinctions on the diversity of marine life since the latest Precambrian. Each sharp drop in diversity is associated with a mass extinction. The most severe drop is in the Late Permian, when 52% of marine families became extinct. As many as 96% of all species may have become extinct at that time.

A 26-million-year periodicity has been detected in the marine extinctions of the past 250 million years; that is, rates of family and genus extinction seem to increase, sometimes to mass extinction levels, every 26 million years. The most recent episode was about 13 million years ago in the Middle Miocene. Although there is debate regarding the quality of both the data and the analytical methods, the implications are profound. Periodic extinctions suggest the periodic causes.

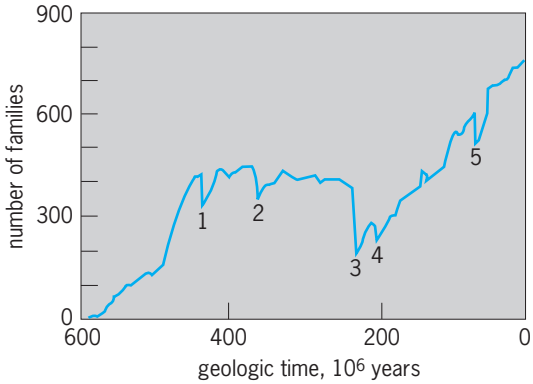

Fig. 2. Diversity of marine families in the fossil record. Sharp drops (1–5) indicate times of mass extinctions. 1 = Late Ordovician (−12%). 2 = Late Devonian (−14%). 3 = Late Permian (−52%). 4 = Late Triassic (−12%). 5 = Late Cretaceous (−11%). (*After D. M. Raup and J. J. Sepkoski, Jr., Mass extinction in the marine fossil record, Science, 215:1501–1503, 1982*)

The fossil record of terrestrial animals and plants is less well known than for marine organisms. Terrestrial animals and plants are not as likely to be preserved as fossils. As a result, it is more difficult to estimate the time and magnitude of the extinction of terrestrial organisms.

Terrestrial and marine mass extinctions seem to occur at about the same time. The Late Permian, Late Triassic, and Late Cretaceous are also times of extinction for terrestrial vertebrates; the most dramatic extinction of terrestrial vertebrates took place at the end of the Cretaceous, when the last dinosaurs died off. Correlation between the terrestrial and marine record is difficult, but it appears that the terrestrial and marine extinctions at the end of the Cretaceous occurred at the same time. *See* DINOSAURIA.

The best record of terrestrial vertebrate extinction is that of the Pleistocene. Late Pleistocene extinctions in North America are especially well known—33 genera of mammals vanished during the last 100,000 years. These extinctions were concentrated among the large mammals—those over 100 lb (44 kg) in weight—and most occurred during a short time interval approximately 11,000 years ago.

Causes. Ever since the work of Georges Cuvier, the French naturalist who demonstrated the reality of extinction, explanations have fascinated both scientists and the general public. Cuvier invoked sudden catastrophic events, whereas his contemporaries favored more gradual processes. These two themes, catastrophism and gradualism, are still debated.

In 1980 high concentrations of iridium were reported precisely at the Cretaceous-Tertiary boundary. Iridium is rare in most rocks but more abundant in meteorites. It was proposed, therefore, that an asteroid struck the Earth 65 million years ago. The impact darkened the atmosphere with dust, caused a catastrophic short-term cooling of the climate, and thus led to the extinction of dinosaurs and many other Cretaceous species. The iridium-rich layer at the boundary marks this terminal Cretaceous event.

Subsequent research has strengthened the case. Mineral grains with textures characteristic of high shock pressures have been found in the boundary layer. Soot, possible evidence of fires ignited by the impact, has also been recovered at the boundary.

Astronomical theories have been put forward to explain the Late Cretacous extinctions as well as the 26-million-year periodicity. In one theory, the Sun has a distant companion star that would pass in orbit near the solar system's cloud of comets every 26 million years. This might perturb many comets, sending a few into the Earth. A comet would produce the same effects as an asteroid. Although this companion star has yet to be discovered, it has been named Nemesis, and a search for it has been undertaken.

Opponents of the impact theory contend that many features of the Late Cretaceous rock and fossil record are better explained by terrestrial causes such as volcanism, lowered sea level, and climatic change. They argue that the fossil record indicates that Late Cretaceous extinctions occurred over several millions of years and were not catastrophic.

Although the impact theory has been offered as an explanation for almost all the extinctions in the fossil record, well-confirmed discoveries of iridium-rich sediments in association with extinctions have been made only at the Cretaceous-Tertiary boundary and in the late Eocene.

Other explanations for mass extinctions include lowered sea level, climatic cooling, and changes in oceanic circulation. It has been pointed out that

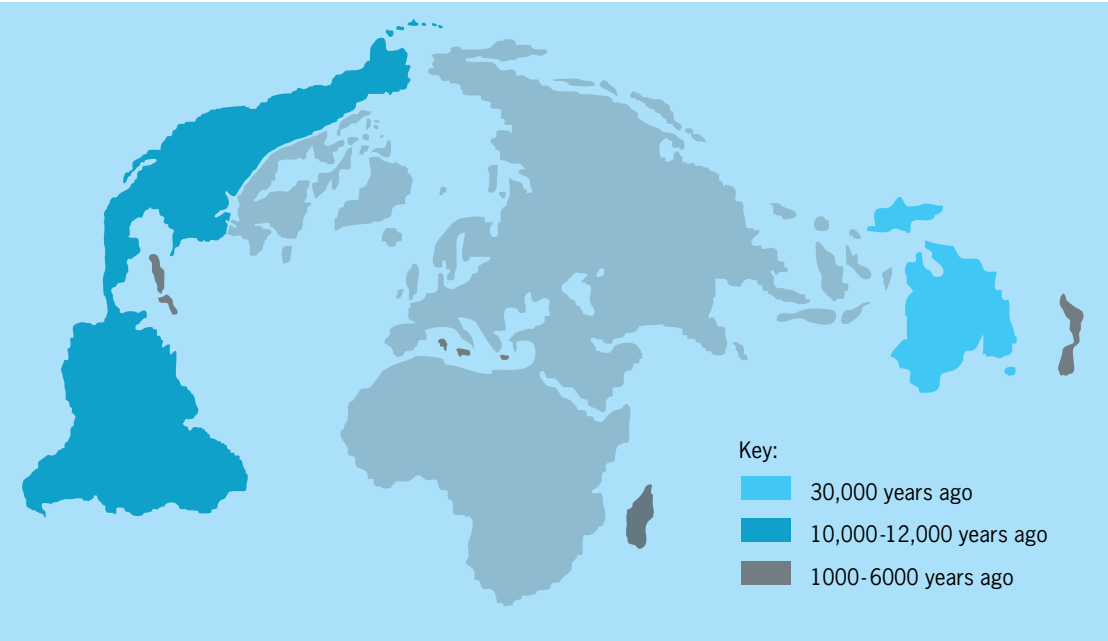

Fig. 3. Times and places of extinctions of large mammals during the late Pleistocene. (*After P. S. Martin and R. G. Klein, eds., Quaternary Extinctions, 1984*)

the Late Ordovician extinctions coincided with an episode of global cooling and sea-level lowering prompted by the growth of the polar ice caps. Explanations for the Late Devonian, Late Permian, and Late Triassic mass extinctions are more elusive. Some evidence for climatic cooling is present in the Late Devonian, though other evidence suggests more sudden causes. The Late Permian marks a time in the Earth's history when sea level was exceptionally low. Habitat for marine organisms may have been severely restricted, and global climates may have been strongly seasonal. The Late Triassic extinctions are also associated with a lowered sea level.

Biotic processes such as disease, predation, and competition may also cause the extinction of species but are difficult to prove from the fossil record because they leave little evidence. Thus there is no evidence that mammals preyed upon dinosaur eggs. In any event, biotic factors usually affect only one or a few interdependent species. Disease, predation, and competition probably played only minor roles in mass extinctions.

Predation has been implicated in the extinction of Pleistocene mammals, which may have resulted from human hunting. Their large size and their disappearance at a time of human immigration support this possibility. **Figure 3** shows the times of large-mammal extinctions during the past 30,000 years. Pleistocene extinctions have also been attributed to rapid climatic change.

Predation and competition are important causes of more recent extinctions, which continue today. Human activities such as hunting and fishing (predation), habitat alteration (competition for space), and pollution have probably destroyed thousands of species. These activities, together with continued tropical deforestation and resulting changes in climate, are likely to cause extinctions that will be comparable to the mass extinctions seen in the fossil record. *See* FOSSIL. Karl W. Flessa

Bibliography. J. C. Balovet, *Extinct Species of the World: 40,000 Years of Conflict*, 1990; P. Ehrlich and A. Ehrlich, *Extinction*, 1981; D. K. Elliott (ed.), *Dynamics of Extinction*, 1986; P. S. Martin and R. G. Klein (eds.), *Quaternary Extinctions*, 1984; M. H. Nitecki (ed.), *Extinctions*, 1984; M. J. Novacek and Q. Wheeler (eds.), *Extinction and Phylogeny*, 1992; D. M. Raup, *The Nemesis Affair*, 1986, revised 1999; S. M. Stanley, *Extinction*, 1987; E. O. Wilson (ed.), *Biodiversity*, 1988.

Extraction

A method of separating the constituents of a mixture utilizing preferential solubility of one or more components in a second phase. Commonly, this added phase is a liquid, while the mixture to be separated may be either solid or liquid. As a mundane example, the preparation of tea or coffee is a process of liquid/solid extraction whereby water selectively dissolves certain components of the mixture, leaving behind the insoluble residue (as tea leaves or coffee grounds). If the starting mixture is a liquid, then the added solvent must be immiscible or only partially miscible with the original and of such a nature that the components to be separated have different relative solubilities in the two liquid phases.

Principles. The ratio of the concentrations of a particular dissolved substance (solute) in two coexisting liquid phases at equilibrium is shown in Eq. (1),

$$D_A = \frac{x_A}{y_A} \tag{1}$$

where D is the distribution coefficient, and x_A and y_A are the concentrations of A in the two phases. The ease of separation of two components in a mixture is conveniently measured by the separation factor α, which is the ratio of the distribution coefficients of the two components between the two solvents (the equivalent of relative volatility in distillation), as in Eq. (2).

$$\alpha_{AB} = \frac{D_A}{D_B} \tag{2}$$

Although the basic concepts of equilibrium outlined above define the potential for separation, they give no indication of the rate of the process. When two phases that are not at equilibrium are contacted together, the rate of transfer of solute between them depends on the extent to which the concentrations of the solute in the two phases differ from the equilibrium value as fixed by the distribution coefficient. In the classical two-film theory, it is assumed that the two phases are actually in equilibrium at the interface and that the resistance to mass transfer is concentrated in thin films on either side of the interface. Mass transfer through these films takes place by molecular diffusion. Thus, Eq. (3) applies, where

$$\text{Rate} = k_x I(x_A - x'_A) = k_y I(y'_A - y_A) \tag{3}$$

k is the film mass-transfer coefficient for a particular phase; I is the interfacial area over which mass transfer takes place; x'_A and y'_A are the solute concentrations in the bulk of a phase; and x_A and y_A are the concentrations in a phase adjacent to the interface.

Important variables affecting rate of mass transfer are temperature and agitation. When two immiscible liquids are mixed together, one will break into droplets, forming the dispersed phase while the other remains coherent as the continuous phase. Increasing the degree of agitation gives smaller droplets, hence greater total surface area, enhancing the mass-transfer rate.

Having achieved dispersion and mass transfer, it is necessary to segregate the phases for separate removal. Coalescence of droplets then becomes important, and various chemical and mechanical aids are available for speeding this step.

Chemistry of extraction. Many applications of extraction, particularly those concerned with the separation of metals, depend upon chemical interaction between solute and solvent. In nonreacting systems, phase behavior is determined by physical differences

such as polarity and can be analyzed by using thermodynamics. *See* CHEMICAL THERMODYNAMICS.

Ion association systems. Ion association systems involve the pairing of oppositely charged ions under the influence of electrostatic attraction. Metals may be incorporated as either the cationic or anionic species in an aqueous phase; by association with an oppositely charged ion, an uncharged molecule is formed which can be extracted into the organic phase. Examples of metals in the anionic form include halides and thiocyanates, while the organic solvent is typically an oxygenated compound (for example, ethers, alcohols, ketones). An example is the extraction of uranium by tributyl phosphate (TBP), shown in reaction (4). The position of equi-

$$\begin{array}{l} \text{Aqueous } UO_2^{2+} + 2NO_3^- \rightleftharpoons UO_2(NO_3)_2 + 2TBP \\ \text{Organic} \qquad\qquad\qquad\qquad \Downarrow \\ \qquad\qquad\qquad\qquad\qquad UO_2(NO_3)_2\ 2TBP \end{array} \tag{4}$$

librium in metals extraction reactions of this type depends upon the pH of the system. Control of the pH therefore allows sharp separations between different metals.

Chelate systems. Metal chelates are cyclic coordination compounds containing a metal atom in the ring. A reagent that has found widespread application is 8-quinolinol (**I**) which contains two atoms (N and O) that will coordinate with metal ions to form a five-membered ring as shown in reaction (5). Thus

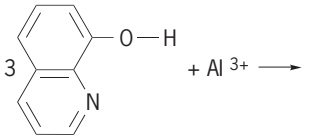

(1)

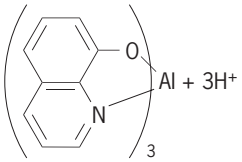

(5)

(2)

aluminum 8-quinolinate (**II**) is an uncharged chelate that resembles many other unionized molecules in its extremely low water solubility and relatively high solubility in organic solvents. Metal separations may be achieved by the control of pH and oxidation state. Other chelating agents which have proven to be useful in extraction and dithiozone (diphenylthiocarbazone) and acetylacetone. *See* CHELATION.

Nonreacting systems. Separation of a liquid mixture into two liquid phases in thermodynamic equilibrium can be understood via energy considerations. The total energy of the system is described by the Gibbs free energy G, defined in Eq. (6), where H

$$G = H - TS \tag{6}$$

is the enthalpy, T is temperature, and S is entropy. When a mixture is prepared from pure components, the change in Gibbs free energy, or energy of mixing,

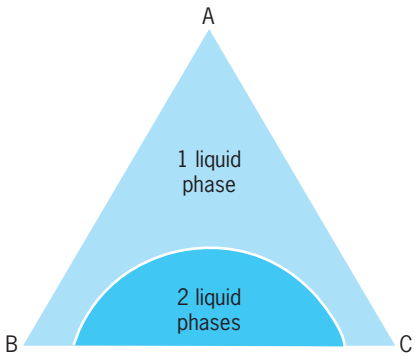

Fig. 1. Phase diagram of a three-component system.

g_m, can be expressed as in Eq. (7), where g_i is the

$$g_m = g_i + g_e \tag{7}$$

energy of mixing for an ideal component and g_e is the "excess" Gibbs free energy. When g_e is zero, an ideal system exists and no phase separation occurs. Highly nonideal systems with significant g_e contributions exhibit liquid-phase separation over certain ranges in composition.

The phase behavior of three-component systems is commonly represented by using a ternary diagram (**Fig. 1**) which forms a useful tool for quantitative analysis and prediction. A, B, and C define the three pure components; binary mixtures correspond to the sides of the figure, while any point within the triangle is a ternary composition, which may be miscible (one liquid phase) or form two liquid phases.

Contacting. The three basic methods for contacting solvent (S) with the feed mixture (F) are shown in **Fig. 2**.

Countercurrent extraction is the most efficient and is the choice for commercial operation whenever possible. In this latter approach, raffinate and extract phases flow countercurrently and emerge at opposite ends of the contactor. Many stages can be used to give a high degree of separation while maintaining a modest solvent requirement. *See* COUNTERCURRENT TRANSFER OPERATIONS.

Liquid/solid extraction. Liquid/solid extraction may be considered as the dissolving of one or more

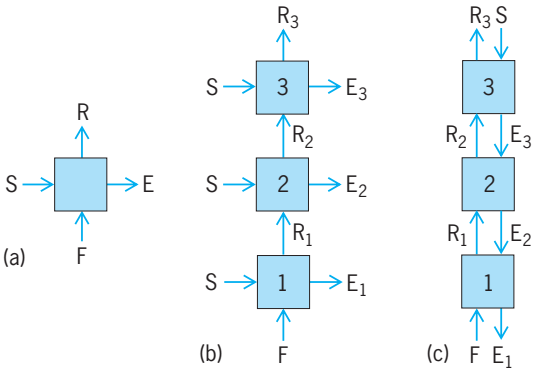

Fig. 2. Contacting methods used in extraction: (*a*) single stage; (*b*) crosscurrent; and (*c*) countercurrent.

components in a solid matrix by simple solution, or by the formation of a soluble form by chemical reaction. The largest use of liquid/solid extraction is in the extractive metallurgical, vegetable oil, and sugar industries. The field may be subdivided into the following categories: leaching, washing extraction, and diffusional extraction. Leaching involves the contacting of a liquid and a solid (usually an ore) and the imposing of a chemical reaction upon one or more substances in the solid matrix so as to render them soluble. In washing extraction the solid is crushed, permitting the valuable soluble product to be washed from the matrix. Sugar recovery from cane is a prime example of this method. In diffusional extraction the soluble product diffuses across the denatured cell walls (no crushing involved) and is washed out of the solid. The recovery of beet sugar is an excellent case in point.

Particle size is significant in all cases since it is a direct function of the total surface area that will be available for either reactions or diffusion. In ores, porosity and pore-size distribution greatly affect the rate of extraction, because the solvent must flow or diffuse in and out of the pores; and, in many cases, the solute moves through the pores to the particle surface by diffusion.

Liquid/liquid extraction. In crosscurrent extraction, fresh solvent is added at each stage; enhanced separation is obtained, but solvent requirement is high.

Liquid/liquid extraction has found application for many years in the coal tar industry. The recovery of tar acids from crude tar oil by washing with an aqueous solution of alkali is an example where chemical interaction between solute and solvent determines differential solubility.

On a smaller scale, extraction is a key process in the pharmaceutical industry for recovery of antibiotics from fermentation broths. Penicillin is obtained by extraction into solvents such as amyl acetate at relatively low pH values (2 to 2.5) and is then stripped from the organic phase by treatment with a buffered aqueous solution at about pH 7 to 7.5. Other examples in this field are the recovery and separation of vitamins and the production of alkaloids from natural products.

Liquid/liquid extraction separates the components of a homogeneous liquid mixture on the basis of differing solubility in another liquid phase. Because it depends on differences in chemical potential, liquid/liquid extraction is more sensitive to chemical type than to molecular size. This makes it complementary to distillation as a separation technique. One of the first large-scale uses was in the petroleum industry for the separation of aromatic from aliphatic compounds. The original process employed liquid sulfur dioxide as solvent. More recently, sulfolane (thio-cyclopentane-1,1-dioxide) has replaced sulfur dioxide for extraction of lighter aromatics due to its greater selectivity and ease of recovery. For the selective separation of higher-molecular-weight aromatics and aliphatics as in lubricating oil manufacture, phenol and furfural are the most widely used solvents.

Supercritical fluid extraction. With a conventional liquid solvent and a given system, temperature is the only variable which can be manipulated to influence solvency. However, if the solvent is a fluid near its critical temperature, changes in pressure significantly affect the fluid's density, hence solvency for a given solute. One advantage of this approach is that the solvent can be separated again merely by reducing the pressure, offering energy savings. A commercial application of the technique is the use of supercritical carbon dioxide for selective removal of caffeine from coffee extract.

Equipment. On a laboratory scale, the separatory funnel is the simplest device for achieving mixing and subsequent phase separation. To obtain continuous extraction of a small feed sample, the Soxhlet extractor is employed.

If the starting mixture of the sample is a solid, it must first be ground down to the required particle size. On a commercial scale, the equipment which is used for achieving efficient contacting with the liquid solvent tends to be specific to the type of industry and may employ screw-type conveyors, successive mixing/settling tanks, or a variety of proprietary designs.

Liquid/liquid extraction using countercurrent processing is preferably carried out in a vertical column. The internals are designed to disperse one phase into droplets, and to accomplish mass transfer between this phase and the continuous phase in a series of stages. These stages may be physical partitions comprising plates or trays, or theoretical concepts in the case of differential contactors such as a packed column.

A novel means of supplying mechanical energy to enhance mixing in an extractor is by pulsing the inlet flow. This concept of pulsed extraction has been applied to both packed and perforated-plate columns, but the mechanical problems associated with pulsing in large-capacity units have limited the commercial adoption of this technique. *See* CHEMICAL EQUILIBRIUM; CHEMICAL SEPARATION TECHNIQUES; COUNTERCURRENT TRANSFER OPERATIONS; SOLVENT EXTRACTION; TRANSPORT PROCESSES.

Bruce M. Sankey

Bibliography. J. Rydberg, *Solvent Extraction Principles and Practice*, 2d ed., 2004; L. T. Taylor, *Supercritical Fluid Extraction*, 1996; J. Wisniak and A. Tamir, *Liquid-Liquid Equilibrium and Extraction: A Literature Source Book*, 1987.

Extrapolation

A process in mathematics used to find the value of a function outside its tabulated values. This is done as in interpolation by assuming that over a small range of x the function may be closely approximated by a polynomial or some other readily computed function. *See* INTERPOLATION.

Formulas. Any of the interpolation formulas can be used, therefore, and the desired value of x substituted in them. Thus, for example, if $y = f(x)$ has been

Difference table

x	$y = \log x$	δy	$\delta^2 y$	$\delta^3 y$
1.00	0.0000 000			
		43 214		
1.01	0.0043 214		−426	
		42 788		8
1.02	0.0086 002		−418	
		42 370		9
1.03	0.0128 372		−409	
		41 961		8
1.04	0.0170 333		−401	
		41 560		7
1.05	0.0211 893		−394	
		41 166		$7 + \epsilon$
1.06	0.0253 059		$-387 + \epsilon$	
		$40\,779 + \epsilon$		$7 + \epsilon$
1.07	$0.0293\,838 + \epsilon \times 10^{-7}$		$-380 + 2\epsilon$	
		$40\,399 + 3\epsilon$		$7 + \epsilon$
1.08	$0.0334\,237 + 4\epsilon \times 10^{-7}$		$-373 + 3\epsilon$	
		$40\,026 + 6\epsilon$		
1.09	$0.0374\,263 + 10\epsilon \times 10^{-7}$			

tabulated at $x = x_{-N}, x_{-N+1}, \ldots, x_{-1}, x_0$, the Gregory-Newton interpolation formula (1) may be used to

$$y = y_0 + u\,\delta y_{-1/2} + \frac{1}{2!}u(u+1)\,\delta^2 y_{-1} + \frac{1}{3!}u(u+1)(u+2)\,\delta^3 y_{-3/2} + \cdots + \frac{1}{m!}u(u+1)\cdots(u+m-1)\,\delta^m y_{-m/2} \quad (1)$$

determine a polynomial equation passing through the $m + 1$ ordinates $y_{-m}, y_{-m+1}, \ldots, y_{-1}, y_0$. These differences give expression (2), and the differences

$$u + \frac{x - x_0}{h} \quad (2)$$

$\delta^k y_{-k/2}$, $k = 1, 2, 3, \ldots$, are the same as those used in interpolation.

If $-1 < u < 0$, then $x_{-1} < x < x_0$ and the formula is used to interpolate. On the other hand, substitution of positive values of u permits its use for extrapolation for y beyond y_0, the last value tabulated.

If the function $y = f(x)$ is known, the error introduced by using a polynomial to extrapolate for the value of the function can be expressed by adding to Eq. (1) a remainder term, shown as notation (3), where now ξ is any value of x lying be-

$$\frac{1}{(m+1)!}u(u+1)\cdots(u+m)f^{(m+1)}(\xi) \quad (3)$$

tween the smallest and the largest of the numbers $x_{-m}, x_{-m+1}, \ldots, x_{-1}, x_0$, and x. This term will be larger for extrapolation, $u < 0$, than for interpolation, $u < 0$, for two reasons. First, since u is positive for extrapolation, the coefficient of $f^{(m+1)}(\xi)$ will be larger. Second, since the range of values permitted ξ is larger, $|f^{(m+1)}(\xi)|$ must be assumed to be larger. It is necessary, in calculating the error, to take the largest absolute value of this $(m + 1)$th derivative of $f(x)$ in the above range of ξ. If there is a singularity of $f(x)$ or of its derivatives near the value of x required in the extrapolation, the remainder term in notation (3) would indicate that the extrapolation could involve a large error.

If it is necessary to extrapolate a distance greater than h, the interval of the **table**, beyond the limits of a table, it may be helpful to proceed by first extending the entries in the table by extrapolation. This may be done by assuming, for instance, that some order of difference remains constant, or by letting u take on positive integral values in Eq. (1). An estimate of the errors introduced by extending the table in this way can be made by attempting to extrapolate for the last few entries in the table from the earlier entries. For values not near a singularity of the function tabulated, these errors should be of about the same size as those introduced in extending the table. Of course, the same number of entries should be added in the two cases.

Having extended the entries in the table by extrapolation, one can look upon the problem of finding y for an intermediate value of x as just the problem of interpolation. If the degree of the interpolating polynomial is large enough, the error of this interpolation can be ignored in comparison with the error in extrapolating for the additional entries.

Example. From the portion of the difference table lying above the line, find by extrapolation the logarithms of 1.07, 1.08, and 1.09 and determine a probable limit for the error.

If it is assumed the third difference stays constant at seven units in the seventh decimal place, the difference table can be extended, as shown below the line, by working from right to left. For this purpose the ϵ's added to the number are ignored.

To determine the maximum error in the extrapolated values it must be recognized that the true third differences would not necessarily be equal to 7×10^{-7}. The maximum error would occur if all these third differences assumed were too high or too low. Therefore the true values of logarithms, rounded off to seven decimal places, will differ from those computed above by less than the ϵs attached to the numbers in the table.

Since a reasonable value for ϵ is $\pm$ 1 reasonable upper limits for the errors in the extrapolation for the logarithms of 1.07, 1.08, and 1.09 are, respectively, 1, 4, and 10 units in the last decimal place. Comparison of these values with the true values reveals that the errors are actually 0, -1, and -2 units in the last decimal place. It should be clear from this example, however, that the error in extrapolating beyond the limts of a table can be expected to grow rapidly for each new entry added. *See* GRAPHIC METHODS.

Kaiser S. Kunz

Extrasolar planets

Planets that orbit stars other than the Sun. Since 1995, the number of such planets known has increased from zero to more than 100. Prior to the discovery of extrasolar planets, theoretical studies suggested that most planetary systems would look like the solar system, with small rocky planets orbiting near the star and the larger gas giants farther out, all in nested concentric quasicircular orbits. These theories have been disproven, and in their place a much richer understanding of planet formation and evolution has emerged.

More than 200 years ago, the observation that the orbits of the major solar system planets are quasicircular, coplanar, and aligned with the Sun's equator led Pierre-Simon Laplace and Immanuel Kant to the conjecture that the planets formed from a common disk of material that must have circled the early Sun. This theory has been greatly strengthened by observations over the past 20 years that more than half of all nearby young stars are swaddled in disks of gas and dust. *See* PROTOSTAR; SOLAR SYSTEM.

It is extremely challenging to detect planets orbiting Sun-like stars. Planets shine only by the reflected light of their host stars. The largest (and therefore easiest to detect) planet in the solar system, Jupiter, has 300 times the mass and 11 times the diameter of the Earth. It is nonetheless 10^9 times fainter than the Sun. Taking a picture of a Jupiter-like planet orbiting another star would be similar to snapping a photo from New York of a firefly buzzing around a megawatt searchlight in Los Angeles. Though this is beyond current technology, a number of strategies are being studied that might be capable of taking direct images of Jupiter-like, or even Earth-like, planets within 20 to 50 years.

In the intervening time we are left to deduce the presence of extrasolar planets by the small effects a planet has on the host star. A planet will alter the position and velocity of its host star by gravitationally tugging on it. For example, the Sun and Jupiter jointly orbit a common center of mass, which lies on the line connecting the Sun and Jupiter, just outside the surface of the Sun. A hypothetical alien astronomer could detect the presence of Jupiter by noting that the position of the Sun is periodically wobbling against the background stars (astrometry) or by measuring the periodic velocity variation of the Sun (Doppler spectroscopy; **Fig. 1**). Such measurements would reveal the orbital period and the magnitude of the wobble. From these data the hypothetical alien astronomer would be able to calculate the orbital radius and mass of the unseen Jupiter from Kepler's third law of planetary motion (1619) and Newton's principle of momentum conservation. *See* ASTROMETRY; CELESTIAL MECHANICS; DOPPLER EFFECT; KEPLER'S LAWS.

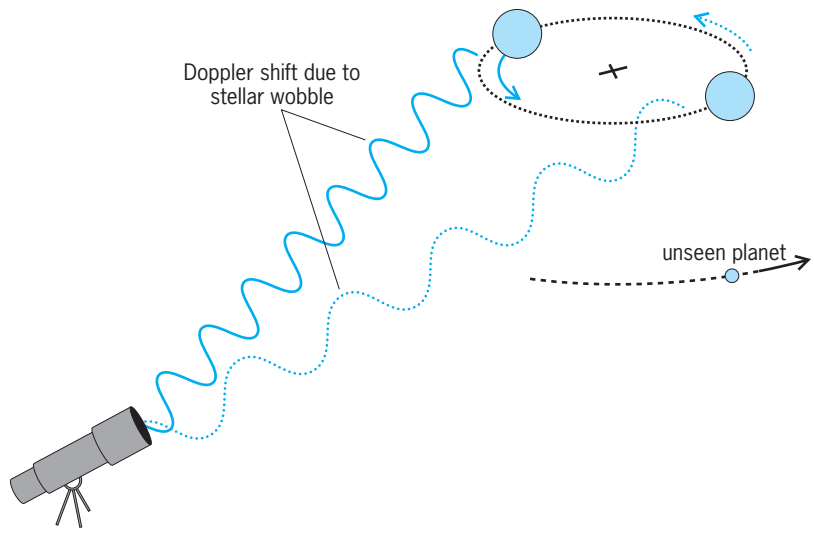

Fig. 1. Stellar wobble that results when an unseen planet orbiting a star gravitationally tugs the star in a small counterorbit. The planet can be detected by observing the star's periodic wobble relative to background stars (astrometry) or by observing the periodic change in the star's line-of-sight velocity (Doppler velocity). (*Courtesy of California & Carnegie Planet Search Program*)

If the Sun-Jupiter orbital plane were aligned edge-on to the alien's line-of-sight, Jupiter would block 1% of the Sun's light as it passed directly in front of (transited) the Sun once per orbit. From this fact, the alien could deduce that Jupiter has 1% the surface area, and hence 10% the diameter, of the Sun.

The planetary zoo. The Doppler velocity method (Fig. 1) is responsible for the discovery of essentially all of the known extrasolar planets, although new space-based telescopes should begin making astrometric detections by about 2015. The Doppler velocity technique can detect only massive planets, and is biased toward detecting planets in small orbits. This bias is due to two effects. More than half an orbit (and preferably more than two orbits) needs be observed to detect a planet. This duration favors the detection of planets in short-period orbits. In addition, the closer a planet is to its star, the stronger the mutual gravitational attraction and hence the larger the velocity that the planet imposes on the star.

In place of the outdated notion that all planetary systems would be similar to the solar system, recent discoveries reveal three primary types of planets. Among the systems of multiple planets that have been discovered to date, some have more than one type of planet. These varying types of planets and systems provide valuable clues about the formation and evolution of planetary systems.

Solar system analogs. Of the 135 known extrasolar planets (as of 2004), only a couple bear even passing resemblance to the giant planets of the solar

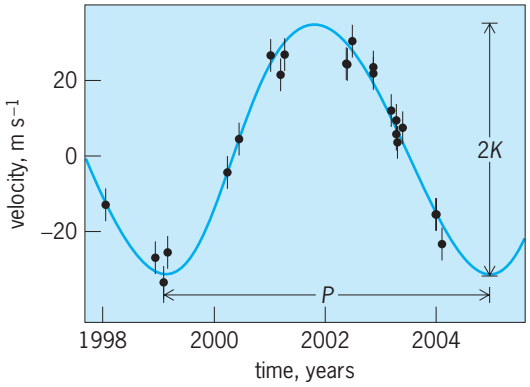

Fig. 2. Doppler velocity curve of the solar system analog HD 70642. The orbital period *P* of the planet can be directly read off the figure, while the mass of the planet can be calculated from velocity amplitude *K*. Here the period is 6 years. The semiamplitude (*K*) is 30 m/s (100 ft/s) and the full amplitude (2*K*) is 60 m/s (20 ft/s). The planet thus orbits at 3.3 AU (compared to Jupiter at 5.2 AU), with a mass about twice that of Jupiter. The shape of the velocity curve reveals that the orbit is approximately circular. (*The Anglo-Australian Planet Search Program*)

system with quasicircular orbits at several astronomical units (1 AU equals the Earth-Sun distance, about 93 million miles or 150 million kilometers). The planet orbiting HD 70642 is the system that most closely resembles Jupiter in the solar system, with an orbital distance of 3.3 AU, twice the mass of Jupiter, and a circular orbit. **Figure 2** shows the discovery data for this planet. The orbital period, 6 years, is immediately obvious from the Doppler velocity curve, while the velocity amplitude yields the planet mass.

Detection of true solar system analogs, Jupiter- and Saturn-like planets, requires 10 to 30 years of monitoring. Doppler velocity programs with sufficient sensitivity to detect these planets have been in existence for only a few years. The first crop of long-period planets should emerge by about 2010.

Eccentric planets. The vast majority of the planets found to date have markedly elliptical (noncircular) orbits. The planet-star distances for these systems vary by 10 to 70% over the course of the orbit. The surface temperatures of these planets have huge seasonal variations as the heating from the host star changes by hundreds of degrees between "winter" and "summer."

The chance of a life-supporting planet in a system with a giant eccentric planet is very small. Many of the eccentric planets have "Earth-crossing" orbits that would spell immediate doom for small Earth-like planets. Even for the eccentric planets that are not "Earth-crossing," the interaction with a small Earth-like planet would gravitationally destabilize the orbit of the small planet.

The unexpected discovery of eccentric planets has driven new theoretical understanding of the evolution of planetary systems. Planets forming in protoplanetary disks are initially in circular orbits. The gravitational interaction between a planet and the disk can cause the planet's orbit to migrate, moving toward or away from the central star, and driving orbital eccentricity.

For planets that survive the disk phase of formation and evolution, another danger awaits: planet-planet interactions. A system with multiple giant planets might stably orbit for millions or even billions of years, then interact chaotically with some planets thrown into smaller orbits and others thrown farther out or even ejected from their system. Such a fate would have befallen the solar system if Saturn had ended up about twice as massive as it did. The Jupiter-Saturn interaction, like billiard balls on a gravitationally warped pool table, would have been devastating to the orbital stability of the "little" planets of the inner solar system, including Earth. *See* CHAOS.

Hot Jupiters. About 20 planets have been found orbiting their host stars with periods of 3 to 5 days. Due to the biases of the Doppler velocity method, these are the easiest type of planets to detect. The orbital distance of these planets is typically 0.04 AU, only about 10 stellar radii from the central star itself. At this distance, these planets are beyond blazing hot, about 2500°F (1400°C). Even at these blast furnace temperatures, hydrogen gas is gravitationally bound to Jupiter- and Saturn-mass planets. The proximity of the star raises enormous tides on the planets, and theorists agree that such tides will circularize a planet's orbits typically in a million years or less. As a result, all the "hot Jupiter" planets are in circular orbits.

The chance that a given planet will transit (pass directly in front of) its host star is inversely proportional to the orbital distance. A hot Jupiter has about a 10% chance of transiting, while the odds for a planet at the Earth's orbital distance are only one in 200. Of the 20 known hot Jupiters found from precision Doppler surveys, three are known to transit their host stars. Another eight transiting hot Jupiters have been found from photometric surveys. The prototype, HD 209458, passes directly in front of its star for about 2 hours every orbital period. During the transit the planet blocks about 1.4% of the host star's light. The combined information from Doppler velocities and the transit yield the exact mass (63% of Jupiter) and size (40% larger than Jupiter) of the planet. The density of the planet, about half that of Saturn, implies that the planet is composed primarily of hydrogen and helium, similar to Jupiter and Saturn. This planet is "puffed up" relative to Jupiter and Saturn due to the intense heat of the nearby host star, like a balloon that expands when heated. *See* TRANSIT (ASTRONOMY).

The hot Jupiter planets could not have initially formed so close to their host star. The enormous heat and tidal forces of the star would prevent the planet from gathering the necessary amount of hydrogen and helium. The hot Jupiters must have formed farther out and then migrated into their roasting hot orbits.

Multiple-planet systems. The solar system consists of multiple planets (eight or nine depending on the current status of Pluto). While multiple-planet systems are probably common, about 80% of the known planets are single. This is almost certainly due to the biases of the Doppler detection technique, which cannot detect small earth-mass planets, and are most sensitive to giant planets in small orbits. Even with

these limitations, several multiple-planet systems have already emerged. These include systems with just eccentric planets (HD 38529, HD 217107) and systems with both a hot Jupiter and eccentric planets (upsilon Andromedae, 55 Cancri). Systems with multiple circular planets, such as the solar system, have yet to emerge.

It is expected that, as the time baseline of precision Doppler surveys increases, many of the single-planet systems will be revealed as multiple systems. When next-generation techniques capable of detecting small Earth-like planets come on line, virtually all planetary systems may be shown to be multiple.

Mass function. Since the Doppler technique is sensitive only to motion perpendicular to the "plane-of-the-sky," Doppler velocities produce only a lower limit on a planet's mass. For the case of an orbit that is edge-on as seen from Earth, the Doppler technique yields the full mass of the planet, while it is insensitive to orbits that are face-on or in the "plane-of-the-sky." Without knowledge of the orbital inclination (the angle of the orbit with respect to the "plane-of-the-sky"), Doppler velocities give the product of the planet's mass M times the sine of the inclination angle i (so-called M sine i).

For studies of planet mass, the uncertainty of the inclination angle is not a major hurdle. Assuming random orbital inclinations as seen from Earth (it would be beyond extraordinary if the orbital axes of planetary systems were systematically pointed either toward or perpendicular to the Earth), two-thirds of all planets will have a true mass within a factor of 2 of the observed M sine i, and only one in a thousand planets would have a true mass ten times larger than the derived M sine i. The typical planet will have a true mass about 1.5 times larger than the minimum M sine i mass.

Figure 3 shows the observed planet mass function, the number of known planets as a function of the minimum M sine i mass of the planet. As this diagram shows, there are very few planets with a minimum mass greater than 5 Jupiter masses. There are many more low-mass planets relative to high-mass planets. The real planet mass function must increase

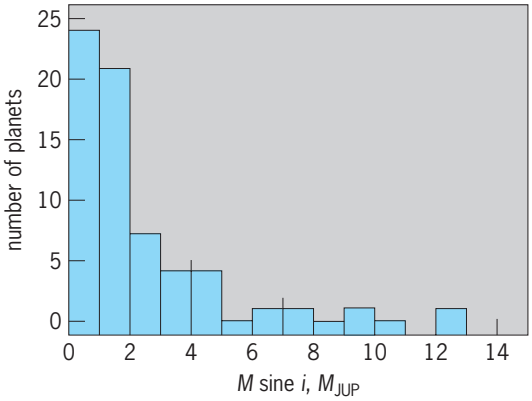

Fig. 3. Planet mass histogram, showing the number of known planets as a function of the minimum *M* sine *i* planet mass, measured in units of Jupiter's mass, M_{JUP}. (*California & Carnegie Planet Search Program*)

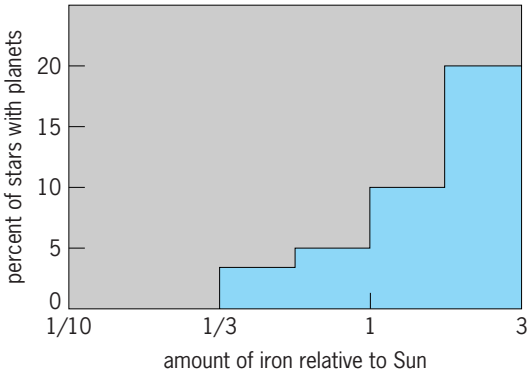

Fig. 4. Percent of nearby stars with known planets as a function of stellar metallicity. The plot shows that stars rich in elements heavier than hydrogen and helium (called "metals" by astronomers) are much more likely to have detectable planets, while stars with less than one-third the Sun's heavy-element abundance have yet to yield any planets. (*Courtesy of Debra Fischer and Jeff Valenti*)

even more rapidly, since massive planets are the easiest to detect with the Doppler velocity technique, while the less massive planets are likely to be overlooked. Planets less than a Jupiter mass are more likely to be missed by the Doppler velocity technique, and below a Neptune mass (15 Earth masses) the Doppler technique is currently insensitive.

While the observed planet mass function reveals that there are roughly 20 planets of 1 Jupiter mass for every planet of 5 Jupiter masses, it would be wrong to conclude that there should be many tens or hundreds of Earth-mass planets for every Jupiter-mass planet. The solar system provides the only example of a system where both large and small planets can be detected. The inventory of the solar system consists of two massive planets (Jupiter and Saturn), two intermediate-mass planets (Uranus and Neptune), and four small planets (Mercury, Venus, Earth, Mars).

Metallicity. The two most common elements in the universe, accounting for about 99% of all matter, are the two simplest, hydrogen and helium. Not surprisingly, stars and giant planets also consist primarily of hydrogen and helium. The abundance of heavier elements in a star can be calculated from the spectrum of the star. Stars with greater than average abundances of heavier elements are much more likely to yield detectable planets (**Fig. 4**). This suggests that planets initially form from an accumulation of dust, rocks, and ices in the protoplanetary disk. Theoretical models show that protoplanets that reach 10 Earth masses can then gravitationally grab hydrogen and helium gas from the protoplanetary disk and rapidly grow to 100 Earth masses or more. *See* STELLAR POPULATION.

Prospects. Precision Doppler programs will continue to be the principal means of detecting planets orbiting nearby stars, at least until about 2015. All 2,000 Sun-like stars (stars with masses between 0.5 and 1.1 solar masses) within 160 light-years (50 parsecs) have been placed under survey, primarily by the two groups that have discovered about 90% of the known planets. By about 2015, these surveys will be sensitive to Jupiter- and Saturn-mass planets

orbiting beyond 4 AU. The central looming question is whether these planets will commonly be found in circular orbits, or alternatively if the architecture of the solar system is rare.

Of the greatest anthropocentric interest are Earth-like planets in quasicircular orbits capable of supporting life. The National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA) have plans for new space-based telescopes to detect possible terrestrial analogs, including the transit missions *COROT* and *Kepler*, which are scheduled to launch in 2006 and 2007, respectively. These may be sensitive to Earth-size planets at 1 AU, providing valuable information on the incidence of such planets. As Earth-size planets block only 1 part in 10,000 of the light from the host star, these missions may be subject to astrophysical and instrumental effects that make interpretation of potential transit signals ambiguous. Transit missions cannot determine orbital eccentricity, and thus cannot determine if planets are in circular, solar system–like, orbits. These space-based transit missions are targeting stars at several hundred light-years, making follow-up by other techniques difficult.

The NASA Space Interferometry Mission (SIM) is extended to beyond 2015. A primary goal of SIM is the astrometric detection of planets as small as 3 Earth masses in Earth-like orbits around the nearest stars. SIM has the promise of measuring both the masses and orbital eccentricities of these planets, thus determining if they are solar system analogs.

NASA and ESA have long-range plans to build giant space-based telescopes capable of directly detecting (imaging) extrasolar planets around the nearest stars. Two techniques currently being researched, coronagraphs and nulling interferometers, block or "null" the light from the central star before it reaches the detector, allowing the faint light from the planets to be seen. The resulting images would be unresolved (a planet would appear as a single dot), but it would be possible to take spectra of the planets for the first time. Spectra could reveal water, oxygen, and other possible signposts of life. Though many technological hurdles remain to be overcome, these systems may optimistically be in place in 20 to 50 years. *See* ASTRONOMICAL SPECTROSCOPY; CORONAGRAPH; INTERFEROMETRY; PLANET.

R. Paul Butler

Bibliography. D. Goldsmith, *Worlds Unnumbered*, University Science Books, Sausalito, CA, 1997; M. D. Lemonick, Searching for other worlds, *Time*, 147(6): 52–57, Feb. 5, 1996; G. Marcy and P. Butler, Hunting planets beyond, *Astronomy*, 28(3):42–47, March 2000; D. Shiga, Imaging exoplanets, *Sky Telesc.*, 97(4):44–52, April 2004.

Extraterrestrial intelligence

Postulated entities beyond Earth with a level of intelligence and comprehension at least equal to that of humans at the present time. While extraterrestrial intelligence is usually envisioned as an advanced civilization, populated by creatures that have evolved via Darwinian evolution on a planet vaguely similar to Earth, it could conceivably be artificial intelligence initially created by biological beings. Extraterrestrial intelligence is a subset of astrobiology, which encompasses all aspects of the existence of, and search for, extraterrestrial life. Astrobiology is sometimes referred to as exobiology or bioastronomy. *See* ASTROBIOLOGY.

Scientific rationale. The expectation that extraterrestrial intelligence exists derives from two facts and one assumption: (1) The universe is vast, with approximately 10^{11} galaxies (a total of about 10^{22} stars) within the reach of telescopes. This number is so large that even if the emergence of intelligence is improbable, such intelligence could still have arisen frequently. (2) The physics and chemistry of the universe are everywhere the same. This is known from astronomical observation. (3) Habitable, Earth-like planets of the type that might spawn intelligence, with thick atmospheres and liquid water on their surface, are not extraordinarily rare. This is a hypothesis, sometimes called the principle of mediocrity. According to this principle, the Earth is not extraordinary in any of its important properties. The principle dates, in its modern form, to Nicolaus Copernicus (1473–1543), who dethroned the Aristotelian idea of an Earth-centered cosmos. *See* UNIVERSE.

In addition to these general arguments, research gives support to the idea that extraterrestrial biology (not necessarily intelligent) might be plentiful. Since 1995, astronomers have detected planets around other, Sun-like stars. At minimum, 10–20% of such stars are now thought to have a solar system consisting of at least one orbiting body. The discovery technique is to measure the small motions of the star induced by the planet, a scheme that is most sensitive to massive worlds in tight orbits. It is still unknown what fraction of stars have small, rocky planets similar to Earth, but the *Kepler* and *Darwin* space-based telescopes, now being planned, should decide this question. *See* EXOPLANETS.

The possibility that some other worlds in the solar system could have spawned life has also increased. There is good photographic evidence that Mars once had lakes and possibly oceans, and may still harbor liquid water hundreds of meters below its surface. Direct evidence for ancient Martian life, claimed to have been found in a meteorite that is known to have come from the Red Planet, is highly controversial. A surprising discovery has been the growing indication for enormous oceans beneath the surface crust of several of Jupiter's satellites (Europa in particular, but also Callisto and Ganymede). Even Titan, a large satellite of Saturn swathed in a thick, hydrocarbon-laced atmosphere, is considered a possible (albeit unlikely) habitat for simple life. Whereas Earth was once thought to be the only solar system body that could support life, there are now several other candidates. If any of these has spawned indigenous life, it would demonstrate that biology is a commonplace occurrence. *See* JUPITER; MARS; SATURN.

But even if life is widespread, can intelligent life be expected to frequently evolve? This question will

probably remain unanswered until we either detect extraterrestrial intelligence or learn what drove the emergence of intelligence on Earth.

Search schemes. Given the enormous distances between the stars (the nearest is 4.2 light-years distant, or 4.0×10^{13} km or 2.5×10^{13} mi), it is beyond human capability to send robotic probes to directly search for intelligence elsewhere. A more promising approach is to look for signals that are either deliberately or inadvertently transmitted from their world to Earth. While there are many possible ways to signal, the most promising is electromagnetic radiation, and more specifically, light and radio waves. In the twentieth century, radio was recognized as an effective way to send information across space; and as early as 1900, Nikola Tesla mistakenly thought he had picked up transmissions from Mars. *See* ELECTROMAGNETIC RADIATION.

Radio waves travel at the speed of light which, according to current understanding of physics, is the fastest possible. Stars produce relatively little radio emission, and the universe is very "quiet" at radio frequencies, particularly in what is called the microwave part of the spectrum (approximately 1000–100,000 MHz), thereby making communication easier. Microwave signals also pass unperturbed through the gas and dust clouds that float between the stars. In 1959, Philip Morrison and Giuseppe Cocconi made calculations regarding the equipment and power necessary to signal over interstellar distances. It turned out that the requirements were not much beyond the type of installation that humans could build now. Consequently, the two physicists urged that a search be made for signals broadcast by other societies that had reached or surpassed the human level of technology. They also noted the advantages of the microwave band, and suggested that 1420 MHz (21 cm wavelength), which happens to be the frequency at which interstellar hydrogen naturally radiates, was the best part of this band to monitor. Since hydrogen is the most abundant element in the universe, this frequency will be known to all technologically sophisticated civilizations. *See* MICROWAVE; RADIO ASTRONOMY.

Frank Drake had independently reached the same conclusions, and in April 1960 he searched for artificial radio emissions from two nearby Sun-like stars, Tau Ceti and Epsilon Eridani. Drake used an antenna at the National Radio Astronomy Observatory in Green Bank, West Virginia, that was 26 m (85 ft) in diameter, and tuned his receiver near the hydrogen frequency. His search, whimsically named Project Ozma, became the prototype for today's more comprehensive experiments, known as SETI (Search for Extraterrestrial Intelligence).

In the 1970s, the National Aeronautics and Space Administration (NASA) began a modest SETI program to build equipment and develop search strategies. In late 1992, the NASA program initiated experiments using the 305-m (1000-ft) Arecibo Radio Telescope in Puerto Rico and the 34-m (111-ft) Goldstone antenna in California. However, action by the U.S. Congress stopped the program within a year. Today, radio SETI bears a strong likeness to the halted NASA effort, although advances in digital technology have improved the equipment. The two largest contemporary experiments are Project Phoenix, which used the Arecibo antenna and other large radio telescopes in a sensitive hunt for signals from nearby Sun-like stars, and SERENDIP, which sweeps large sections of the sky, using the same telescope, in a survey for powerful transmissions. Some of the data from SERENDIP are freely distributed for analysis by home computers, a project known as SETI@home. There are also SETI experiments in Australia, Italy, and Argentina. *See* RADIO TELESCOPE.

So far, no confirmed extraterrestrial signals have been found. However, as technology improves, radio SETI will speed and extend its search for artificially produced transmissions. A major development is the construction of the Allen Telescope Array, which can be used nearly full-time for SETI. It will eventually consist of hundreds of 6-m-diameter (20-ft) antennas located at the Hat Creek Observatory in northern California.

A second search method is to monitor star systems for brief flashes of light, signals that might be deliberately beamed in the direction of Earth with powerful lasers. By using mirrors to focus the lasers, it is straightforward to produce flashes lasting a nanosecond (10^{-9} s) or less that can greatly outshine the light from the transmitting civilization's home star. Optical SETI, as searches for such light pulses are called, has already examined several thousand star systems. The research is being conducted at observatories in California and at Harvard and Princeton universities. An optical SETI telescope that uses a sky-scan strategy, rather than concentrating on individual star systems, has been built at Harvard.

While searching for signals is generally regarded as the most promising scheme for proving the existence of extraterrestrial intelligence, other approaches have been suggested. One might look for examples of astroengineering by very advanced societies, or possibly the radiation produced by high-powered, interstellar spacecraft. In much of the public's mind, the many thousand sightings of unidentified flying objects (UFOs) each year are proof that some extraterrestrials are actually visiting Earth. Nearly all scientists are skeptical of such claims, and one of their principal objections is the lack of convincing physical evidence, despite more than a half-century of publicized sightings.

In addition to the expected improvements in search technology, advances in astronomy will aid SETI efforts. In particular, it will soon be possible to construct space-based telescopes that can directly image planets around other stars. By analyzing the light reflected from these planets' atmospheres, it should be possible to find biomarkers such as oxygen and methane that would indicate the presence of biology. Future SETI experiments will benefit from being able to concentrate their efforts on star systems where worlds with life are known to exist.

Drake equation. In 1961, Drake devised a simple formula to calculate N, the number of broadcasting

societies in the Milky Way Galaxy. The computation multiplies the rate at which such societies arise, R, by the average length of time they survive, L. This is akin to computing the number of students at a university as the product of the number admitted each year times the number of years (4) before the average student graduates. The rate R is composed of terms that estimate the birth rate of stars, the fraction that have habitable planets, the chances that biology will appear, and the fraction of the worlds with life that develop intelligent, technically competent beings. While many of the terms in the Drake equation remain very speculative (particularly L, which depends on sociological factors), the formula remains a highly useful framework for discussing the subject of extraterrestrial life.

Fermi paradox. Enrico Fermi is said to have posed the question "Where is everybody?" in 1950. His remark was intended to point out that, while humans are in no position to colonize other star systems, advanced extraterrestrials—if they exist—might do so. Even if they require thousands of years to travel from one star to the next, an ambitious society could spread itself throughout the entire Milky Way Galaxy in only a few tens of millions of years. Since this is far less than the age of the Galaxy (about 12 billion years), it suggests that, if sophisticated and ambitious societies arose in the past, evidence of their presence should now be everywhere. Since that evidence is lacking, the implication is that the Galaxy is inhabited solely by humans.

There are many suggestions of how this paradox might be resolved; in other words, how the failure to see local evidence of alien activity could be reconciled with a Galaxy that we think might house many sophisticated societies. For example, it could be that interstellar colonization is so daunting that no one ever undertakes it. Perhaps humans are incapable of recognizing the widespread presence of intelligence. Or the extraterrestrials could know about humans, but have arranged for a "one-way mirror," whereby they can watch, but humans cannot detect them (the "zoo hypothesis"). The Fermi paradox, while intriguing, remains a point of discussion rather than a key to new knowledge.

Consequences of discovery. If a signal or other verifiable proof of extraterrestrial intelligence is discovered, what would be the effects on human society? Clearly, this depends on how distant the intelligence is, and whether (in the case of a signal) an embedded message can be deciphered. It is possible that any intelligence that is found will be hundreds or even thousands of light-years distant, in which case two-way conversation or direct, physical contact would be exceedingly difficult.

In the case of a SETI detection, it is quite likely that the transmitting intelligence will be far more advanced than our own. This is because there is a good chance of detection only if L is large (at least thousands of years), and it is unlikely that the first signal received would be from a society that is only a few hundred years beyond us. (Of course, signals will not be found from societies that are not at least as accomplished as ours.) If such advanced beings wish to transmit useful information, human society would be greatly altered. On the other hand, it may be that humans would never be able to decode the transmissions from such a sophisticated civilization. Even in that case, merely learning that intelligence has developed elsewhere in the cosmos would be a profound event.

Seth Shostak

Bibliography. I. Shklovskii and C. Sagan, *Intelligent Life in the Universe*, Holden-Day, New York, 1966, reissue edition, Emerson-Adams Press, 1998; S. Shostak, *Sharing the Universe*, Berkeley Hills Books, Berkeley, CA, 1998; S. Shostak and A. Barnett, *Cosmic Company*, Cambridge University Press, 2003; S. Webb, *If the Universe Is Teeming with Aliens, Where Is Everybody? Fifty Solutions to Fermi's Paradox and the Problem of Extraterrestrial Life*, Copernicus Books, New York, 2002.

Extrusion

The forcing of solid metal through a suitably shaped orifice under compressive forces. Extrusion is somewhat analogous to squeezing toothpaste through a tube, although some cold extrusion processes more nearly resemble forging, which also deforms metals by application of compressive forces. Most metals can be extruded, although the process may not be economically feasible for high-strength alloys.

Hot extrusion. The most widely used method for producing extruded shapes is the direct, hot extrusion process. In this process, a heated billet of metal is placed in a cylindrical chamber and then compressed by a hydraulically operated ram (**Fig. 1**). The opposite end of the cylinder contains a die having an orifice of the desired shape; as this die opening is the path of least resistance for the billet under pressure, the metal "squirts" out of the opening as a continuous bar having the same cross-sectional shape as the die opening. By using two sets of dies, stepped extrusions can be made. The small section is extruded to the desired length, the small split die is replaced by the large die, and the large section is then extruded.

Extrusion pressures and speed vary considerably, depending upon the size and shape of the section

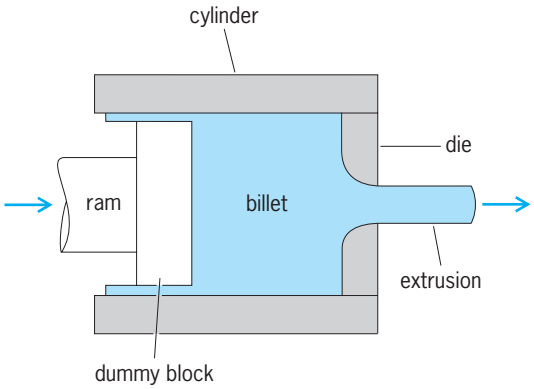

Fig. 1. Schematic representation of the direct extrusion process (hot).

and the mechanical properties of the metal. Some metals, such as magnesium and some aluminum alloys, require slow speeds of a few feet per minute, while others, such as some copper alloys, lead, and steel, are extruded at speeds of over 1000 ft/min (300 m/min). The extrusion speed is also somewhat dependent upon the temperature of the alloy. Considerable heat is generated by the process; if the extrusion speed is high, this heat cannot be dissipated, resulting in a rise in temperature. In some instances, the rise in temperature may be sufficient to melt or at least weaken the metal to the point at which the frictional stresses at the surface cause cracking.

The flow of metal is not uniform during extrusion of the billet. Because of the restraining effect of the die face and the frictional effects at the cylinder walls, the outer zones of the billet resist deformation and the flow occurs most rapidly at the center of the billet. Eventually, as the billet shortens, the different rates of flow at the center and surface result in the billet's becoming hollow. If the extrusion ram travels further, defects appear at the center of the extruded section. Therefore, a portion of the billet may be left in the cylinder and discarded. To prevent the oxidized surface layers of the billet from getting into the extruded product, the dummy block (or follower plate) in front of the ram is of slightly smaller diameter than that of the cylinder; thus, a thin sleeve of the billet is extruded out of the ram end of the cylinder and is subsequently discarded.

Lubricants are used to minimize friction and protect the die surfaces. Graphite is a common lubricant for nonferrous alloys, whereas for hot extrusion of steel, glass is an excellent lubricant.

Indirect, or inverted, extrusion was developed to overcome such difficulties as surface friction and entrainment of surface oxide of direct extrusion. In the indirect process, the ram is hollow, the die opening is in the dummy block, and the opposite end of the cylinder is closed. As the ram advances, the billet does not move as in the case of direct extrusion, and the metal is extruded backward through the die and the hollow ram. However, the process is not very popular because the hollow ram is weaker, resulting in lower machine capacity; trouble-free operation requires that the extruded product be straight and not hit the inside of the ram.

Tubular shapes are produced by a mandrel of the desired inside shape of the product. The mandrel may be fastened to the dummy block in a direct extrusion press if a pierced billet is used; or, in the case of a solid billet, a mandrel must first pierce the billet, after which the main ram advances.

In the production of lead cable sheathing or cored solder wire, the core material (cable or rosin flux) passes through a core tube (or die block) around which is heated lead under pressure. The core material passes from the core tube into a die cavity and the lead is extruded out with it, forming a casing around the cable or rosin core. The process is semicontinuous, molten lead being added to a vertical cylinder and pressure applied periodically. The lead is solid when it reaches the core material.

Cold extrusion. The extrusion of cold metal is variously termed cold pressing, cold forging, cold extrusion forging, extrusion pressing, and impact extrusion. The term cold extrusion has become popular in the steel fabrication industry, while impact extrusion is more widely used in the nonferrous field.

The original process (identified as impact extrusion) consists of a punch (generally moving at high velocity) striking a blank (or slug) of the metal to be extruded, which has been placed in the cavity of a die. Clearance is left between the punch and die walls; as the punch comes in contact with the blank, the metal has nowhere to go except through the annular opening between punch and die. The punch moves a distance that is controlled by a press setting. This distance determines the base thickness of the finished part. The process is particularly adaptable to the production of thin-walled, tubular-shaped parts having thick bottoms, such as toothpaste tubes.

A process requiring less pressure than backward extrusion is the forward-extrusion process, originally called the Hooker process (**Fig. 2**). A formed blank (usually a thick-walled cup) is placed in a die cavity and struck by a punch having a shoulder or enlarged section a short distance from the end. Upon contact with the blank, the nose or end of the punch starts to push the center of the blank through the die cavity, in a manner similar to the action occurring in deep drawing of sheet metal. After the punch has advanced a short distance, the shoulder comes in contact with the top of the thick wall of the blank. The punch shoulder then extrudes the metal through the annular space between the die and the end of the

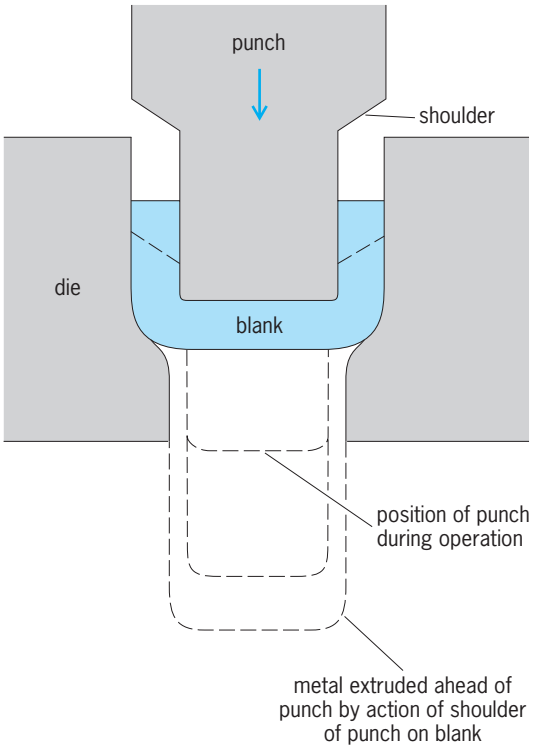

Fig. 2. Schematic of the forward cold-extrusion process, originally called the Hooker process.

punch. Thus, in forward extrusion, the metal moves in the same direction as the punch, whereas in backward extrusion the metal moves in the opposite direction. *See* SHEET-METAL FORMING.

The application of the cold-extrusion process to steel was developed in Germany about 1930 but was not released until after World War II. The success of cold extrusion of steel hinged on the discovery of a suitable lubricant and surface treatment. The steel is given a phosphate surface coating, which absorbs and holds the lubricant (soap emulsion, vegetable oils, or dry metal stearates) and prevents seizure between the metal and tools.

Advantages of cold extrusion are higher strength because of severe strain-hardening, good finish and dimensional accuracy, and economy due to fewer operations and minimum of machining required. *See* METAL FORMING. Ralph L. Freeman

Bibliography. M. Bauser, G. Sauer, and K. Siegert (eds.), *Extrusion*, 2d ed., 2006; P. Saha, *Aluminum Extrusion Technology*, 2000; S. L. Semiatin (ed.), *ASM Handbook Volume 14A: Metalworking: Bulk Forming*, 10th ed., 2005.

Eye (invertebrate)

An aggregation of photoreceptor cells together with any associated optical structures. Eyes occur almost universally among animals, and are possessed by some species of virtually every major animal phylum. However, the complexity of eyes varies greatly, and this sense organ undoubtedly evolved independently a number of times within the animal kingdom.

Structure and occurrence. The simplest invertebrate organs that might be considered to be eyes are clusters of photoreceptor cells located on the surface of the body. Pigment cells are usually interspersed among the photoreceptors, giving the eye a red or black color. Accessory structures, such as the lens and cornea, are usually absent. Simple eyes of this type, called pigment spot ocelli, are found in such invertebrates as jellyfish, flatworms, and sea stars (**Fig. 1**).

The most basic image-forming type of invertebrate eye probably arose from such patches of photoreceptor cells by an in-sinking of the sensory epithelium to form a cup. The cup, which represents the retina, may have remained as a pit open to the surface, may have sunk beneath the surface, or may have become closed in conjunction with the evolution of a cornea and lens. Such an evolutionary history is clearly suggested by the embryology and comparative anatomy of many invertebrates. In gastropod mollusks (snails), for example, the eyes vary from simple pigment cup ocelli to large complex eyes (**Figs. 2** and **3**). In embryonic development, the cups form as an in-sinking of the surface ectoderm. The eyes of free-living flatworms and marine annelids indicate a similar history (**Figs. 4–7**).

The photoreceptor cells in the wall of the cup are generally cylindrical or rounded with an axon fiber directed toward the brain. The photoreceptor cells may be directed toward, or away from, the light source, and the eye is said to be direct or indirect respectively. The distribution of direct and indirect eyes is variable. The eyes of mollusks are direct; the

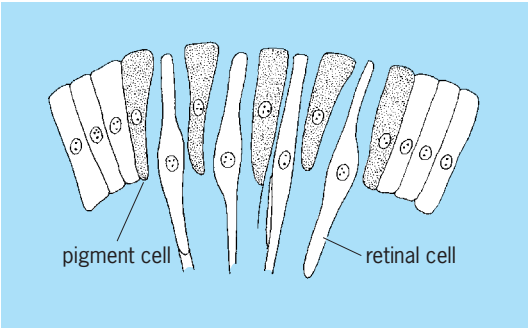

Fig. 1. Pigment spot ocellus, the simpest invertebrate eye. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., W. B. Saunders, 1968*)

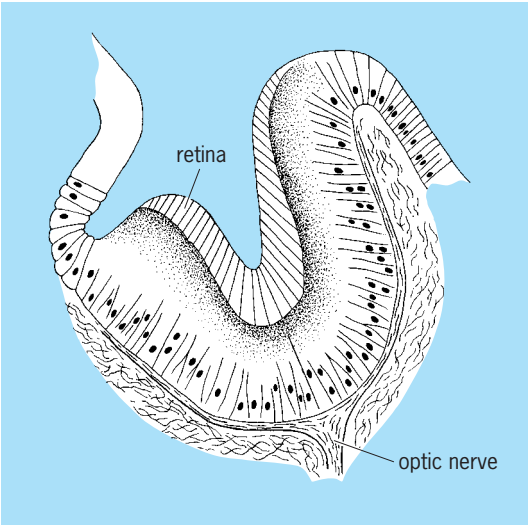

Fig. 2. Simple open eye of the marine limpet *Patella*. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., W. B. Saunders, 1968*)

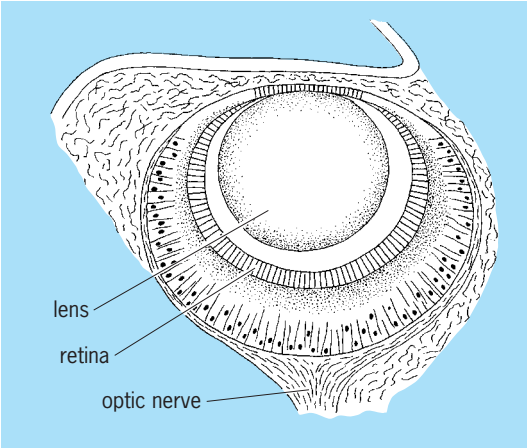

Fig. 3. Large, complex eye of the marine snail *Murex*. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., W. B. Saunders, 1968*)

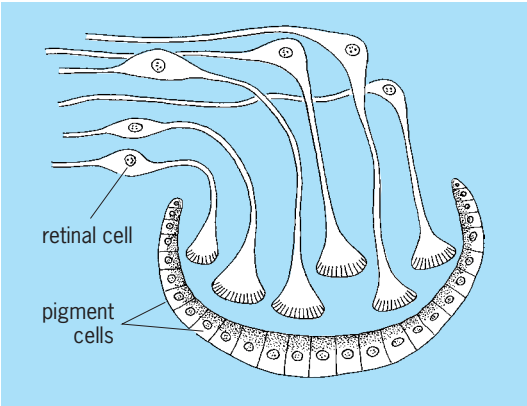

Fig. 4. Simple pigment cup ocellus of flatworm. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., W. B. Saunders, 1968*)

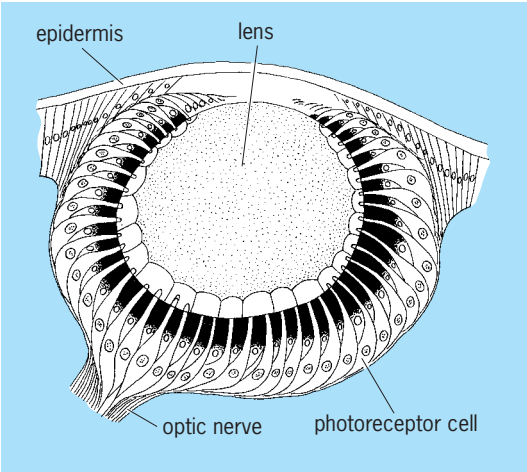

Fig. 5. Large, complex eye of the clam worm *Nereis*. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., W. B. Saunders, 1968*)

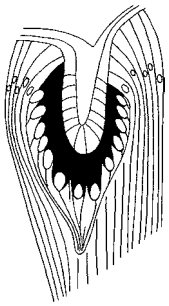

Fig. 6. Simple eye of *Ranzania*, a marine annelid. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., W. B. Saunders, 1968*)

eyes of many flatworms, like those of vertebrates, are indirect. Most spiders possess 8 eyes, of which 2 are direct and 6 are indirect.

In bilateral cephalic invertebrates, the eyes are typically paired and located at the anterior end of the body. Although one pair is usual, as in mollusks and many arthropods, multiple pairs are not uncommon. Some polychaete annelids have 4 eyes, and scorpions may have as many as 12. The greatest number of eyes is found in marine flatworms, where there may be over 100 ocelli scattered over the dorsal anterior surface and along the sides of the body.

The occurrence of eyes on parts of the body other than the head is usually correlated with radial symmetry or unusual modes of existence. Thus eyes are found around the bell margin of jellyfish and at the tips of the arms of some sea stars. Scallops possess many eyes along the margin of the mantle (**Fig. 8**). Some tube-dwelling fanworms have eyes on the plumelike head appendages, which project from the opening of the tube and are used in filter feeding. *See* FEEDING MECHANISMS (INVERTEBRATE).

Function. The primitive function of animal eyes was merely to provide information regarding the intensity, direction, and duration of environmental light. Many invertebrate eyes, even those with lens and cornea, retain those primitive functions. Cornea and lens merely provide the eye with a protective covering and concentrate the light on the retinal surface. The perception of objects is dependent upon several other factors, namely, the number of photoreceptors in the retina, the quality of the optics, and central processing of visual information.

Image formation. This has evolved as an additional capacity of the eyes of some invertebrates. The number of photoreceptor cells composing the retinal surface is of primary importance, since each photoreceptor cell or group of cells acts as the detector for one point of light. An image is formed by the retina through the association of points of light of varying intensity, much as an image is produced by an array of pixels on a computer monitor. The ability of an eye to form an image and the coarseness or fineness of the image are, therefore, dependent upon the number of points of light that are distinguished which, in turn, is dependent upon the number of photoreceptor cells composing the retina. A large number of photoreceptor cells must be present to produce even a coarse image. The great majority of invertebrate eyes cannot form a detailed image because they do not possess a sufficient number of photoreceptor cells. The number of photoreceptor cells might be sufficient to detect movement of an object, but is inadequate to provide much information about the object's form. Clearly, the question as to whether an animal can "see" depends on the meaning of the word. *See* PHOTORECEPTION; TELEVISION.

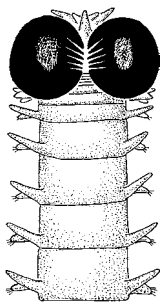

Fig. 7. Highly developed eyes at anterior end of pelagic marine annelid, Alciopidae. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., W. B. Saunders, 1968*)

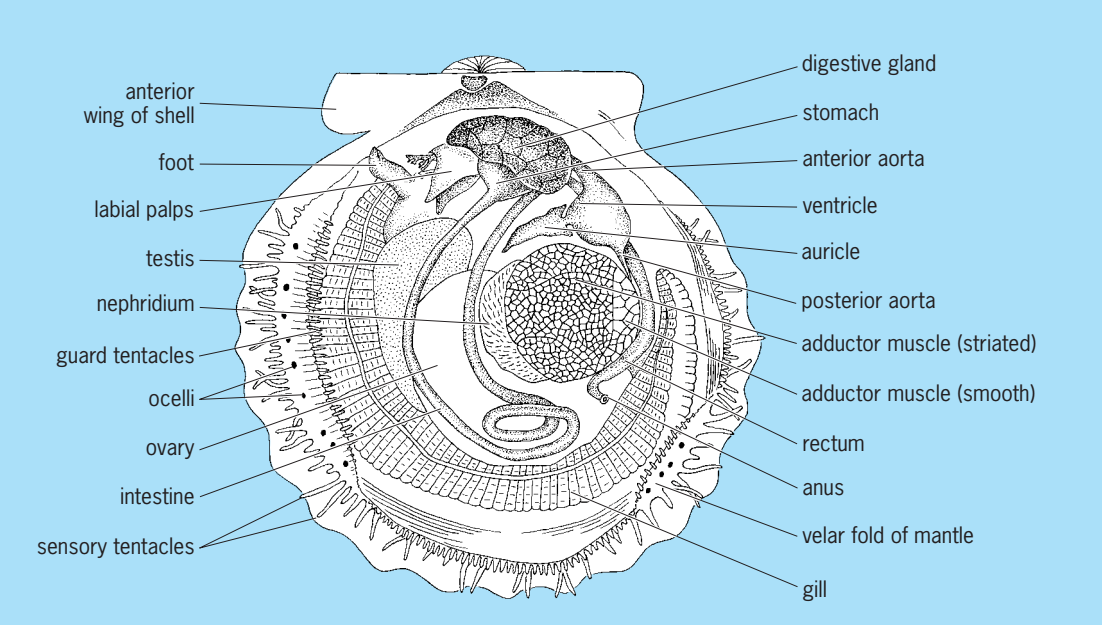

Fig. 8. Scallop *Argopecten* with one valve removed showing ocelli along the margin of the mantle. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., W. B. Saunders, 1968*)

Focus. The focusing mechanisms of invertebrate eyes vary considerably. The focus of arthropod eyes tends to be fixed, that is, the distance between the optical apparatus and the retina cannot be changed. Thus objects are in focus only at a certain distance from the eye, determined by the distance between the lens and the retina. For example, jumping spiders, which are able to detect an object 30 cm away (a considerable distance for so small an animal), have tubelike anterior median eyes (**Fig. 9**). The tubular form increases the distance between the lens and the retina, enlarging the image for distant vision.

The oceanic family of swimming polychaete worms, Alciopidae, have the most highly developed eyes of any of the segmented worms (Fig. 7). The eyes of these worms are focused hydrostatically. A bulb to one side of the eye injects fluid into the space between the retina and the lens, forcing the lens outward. Another mechanism is employed in octopods whereby lens movement is brought about by a ciliary muscle attached to the lens (as in aquatic vertebrates, like fish) [**Fig. 10**].

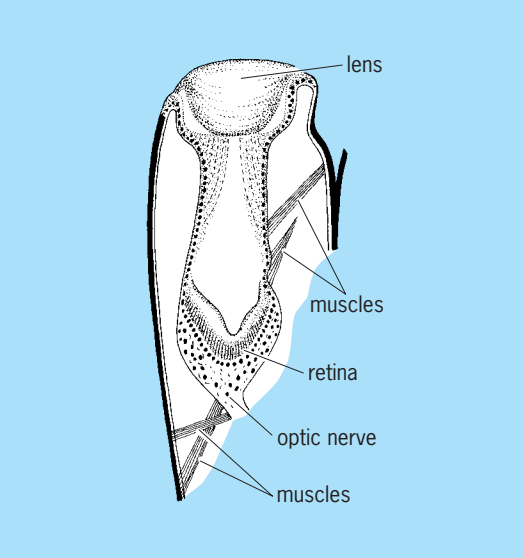

Fig. 9. Longitudinal section through the tubelike anterior median eye of jumping spider *Salticus*. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., W. B. Saunders, 1968*)

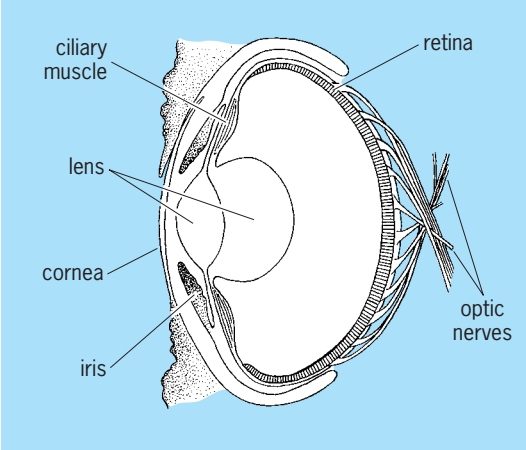

Fig. 10. *Octopus* eye. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., W. B. Saunders, 1968*)

Interpretation. Despite the emphasis placed upon eye structure, an image is ultimately an interpretation by the central nervous system of signals received from the photoreceptors. Visual interpretation is still very poorly understood among invertebrates. The little information available is based upon a few experimental animals, notably the octopus and the honeybee.

Color discrimination has been demonstrated in crustaceans and also in insects, such as bees, flies, and butterflies. Invertebrate color vision can be

extremely complex, involving up to eight primary color classes in some crustaceans.

Visual object perception in invertebrates is correlated with a variety of behavioral adaptations. In fact, behavior frequently indicates that the eye is capable of some degree of image formation. Invertebrates which commonly possess complex eyes, which are probably capable of detecting and recognizing objects, include predators such as cursorial hunting spiders, squids and octopods, certain insects, some crabs and other crustaceans; those with highly controlled flight, such as bees, wasps, flies, dragonflies, and butterflies; those displaying courtship behavior, such as jumping spiders and fiddler crabs; or those whose orientation is by celestial clues, such as bees and some crustaceans.

Compound eyes. The compound eye of crustaceans, insects, centipedes, and horseshoe crabs has a sufficiently different construction from that of other invertebrates to warrant separate discussion. The structural unit of the compound eye is called an ommatidium (**Fig. 11**). The outer end of the ommatidium is composed of a cornea, which appears on the surface of the eye as a facet. Beneath the cornea is an elongated, tapered crystalline cone; in many compound eyes the cornea and cone together function as a lens. The receptor element at the inner end of the ommatidium is composed of one or more central translucent cylinders (rhabdome), around which are located several photoreceptor cells (typically 7 or 8).

The rhabdome is the initial photoreceptive element, and it in turn stimulates the adjacent photoreceptor cells to depolarize. Note that the photoreceptor element of each ommatidium functions as a unit and can respond only to one point of light. Thus

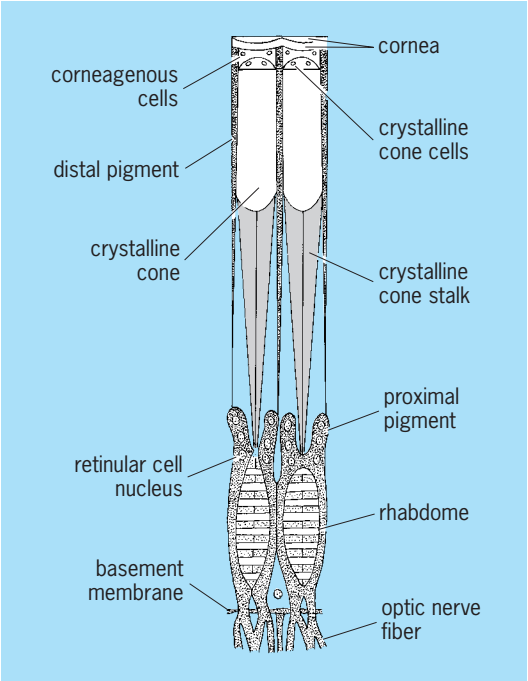

Fig. 11. Two ommatidia from compound eye of crayfish *Astacus*. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., W. B. Saunders, 1968*)

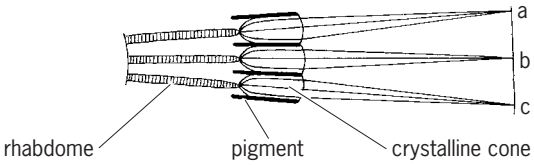

Fig. 12. Compound eye adapted for opposition image formation. Light rays from points a, b, and c. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., W. B. Saunders, 1968*)

image formation is dependent upon the number of photoreceptor units present. The number of ommatidia composing a compound eye varies greatly. Some wood lice (crustaceans) possess no more than 25 ommatidia per eye, whereas the lobster possesses 14,000 and the dragonfly perhaps 30,000.

Pigment granules surround the ommatidium proximally and distally, forming a light screen that separates one ommatidium from another. The pigment granules migrate, depending upon the amount of light. In bright light the ommatidium is adapted by funneling light directly down to the rhabdome (**Fig. 12**), by extending the pigment screen, so that light received by one ommatidium is prevented from stimulating the rhabdome of another. Under these conditions the image produced is said to be appositional, or mosaic. The term mosaic has been misinterpreted to mean that a given ommatidium forms a separate image, even if only a part of the image. In general, however, the compound eyes function like any other eye—each photoreceptor unit represents one point in visual space. It is not obvious whether or not compound eyes have any special advantages over other eye designs, despite their universal occurrence in crustaceans and insects. However, in many arthropods the total corneal surface is greatly convex, resulting in a wide visual field. For example, the compound eye at the end of the eyestalk of a crab may cover an arc of 180°.

Nocturnal vision. The eyes of many nocturnal invertebrates are adapted for functioning in reduced light. Wolf spiders, hunters that run over the ground at night, illustrate eyes of this type. Of the eight eyes, six are indirect and are supplied with a layer of reflecting pigment, called a tapetum, which reflects light onto the photoreceptors. Like those of a cat, the eyes of these spiders appear to glow in the dark when facing a beam of light; the eyes of specimens the size of a quarter can be seen from a distance of 15 m (50 ft) or more!

Reflecting pigment is also found in the compound eyes of certain nocturnal insects and crustaceans. One type of compound eye forms a superpositional image by withdrawal of the screening pigment, so that light received by one ommatidium can strike the rhabdome of an adjacent ommatidium (**Fig. 13**). Under these conditions, a given rhabdome will receive light entering a number of facets so the image will appear much brighter. However, visual acuity is much lower in a superposition eye as compared to an apposition eye.

Polarization vision. Many invertebrate eyes are capable of seeing and analyzing patterns of polarized

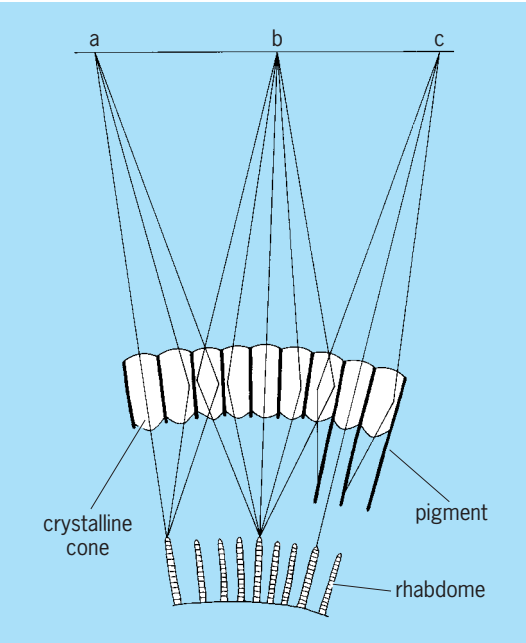

Fig. 13. Compound eye which is adapted for superposition image formation. The light from points a, b, and c is concentrated on an ommatidium. (*After R. D. Barnes, Invertebrate Zoology, 2d ed., W. B. Saunders, 1968*)

light in nature. This capacity reaches its apex in compound eyes, as well as in the simple eyes of cephalopods (octopuses, squids, and cuttlefish). Cuttlefish are known to communicate with each other with displays produced on their body surfaces that are visible only to animals that have polarization vision. Most invertebrates with polarization vision, however, use this ability to navigate with the assistance of patterns of polarization in the sky that occur naturally due to scattering of sunlight by the atmosphere. Bees and ants can find their way back to their nests or hives using only these celestial polarization cues. *See* EYE (VERTEBRATE).

Thomas W. Cronin; Robert D. Barnes

Bibliography. M. A. Ali (ed.), *Photoreception and Vision in Invertebrates*, 1984; H. Autrum (ed.), *Handbook of Sensory Physiology*, vol. VII, pt. 6A–C: *Vision in Invertebrates*, 1981; P. A. Meglitsch and F. R. Schram, *Invertebrate Zoology*, 3d ed., 1991; D. G. Stavenga and R. C. Hardie (eds.), *Facets of Vision*, 1989.

Eye (vertebrate)

A sense organ that acts as a photoreceptor capable of image formation. The eye of vertebrates is constructed along a basic anatomical pattern (**Fig. 1**) which, in the diversification of animals, has undergone a variety of structural and functional modifications associated with different ecologies and modes of living.

Often compared with a camera (**Fig. 2**), the vertebrate eye is conveniently described in terms of its wall, cavities, and lens. *See* CAMERA.

The Wall

The wall of the eye consists of three distinct layers or tunics which, from outward to inward, are termed the fibrous, vascular, and sensory tunics.

Fibrous tunic. This continuous, outermost fibrous tunic comprises a transparent anterior portion, the cornea, and a tough posterior portion, the sclera. In the human, the cornea represents about one-sixth of the fibrous tunic, the sclera five-sixths; the transitional area where they meet is of particular functional importance and is termed the sclerocorneal junction or limbus (Fig. 1 and **Fig. 3**).

Cornea. The vertebrate cornea exhibits very few modifications in structure regardless of environmental influences. Its major constituent is connective tissue (both cells and fibers), regularly arranged and bordered on both anterior and posterior surfaces by an epithelium. The anterior epithelium is stratified, ectodermal in origin, and continuous with the (conjunctival) epithelium lining the eyelids. The posterior epithelium (termed the mesothelium or endothelium) is only one cell thick, perhaps mesodermal in origin, and ends at the limbus. The transparency of the cornea is attributed to the geometric organization of its connective tissue elements, its

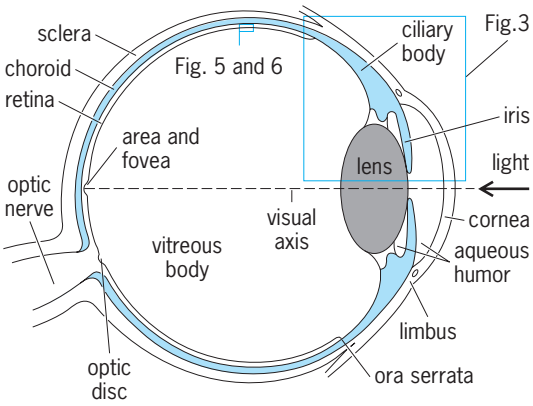

Fig. 1. Horizontal section through human eye.

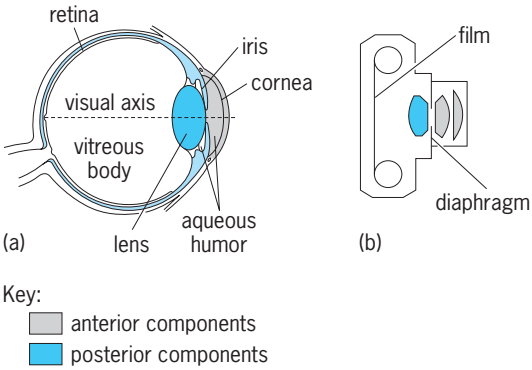

Fig. 2. Comparison between (*a*) eye and (*b*) miniature camera. In both, the refractive part consists of anterior and posterior components, separated partially from one another by an iris diaphragm. The sensory retina corresponds to the light-sensitive film. (*After E. Gardner et al., Anatomy: A Regional Study of Human Structure, 4th ed., Saunders, 1975*)

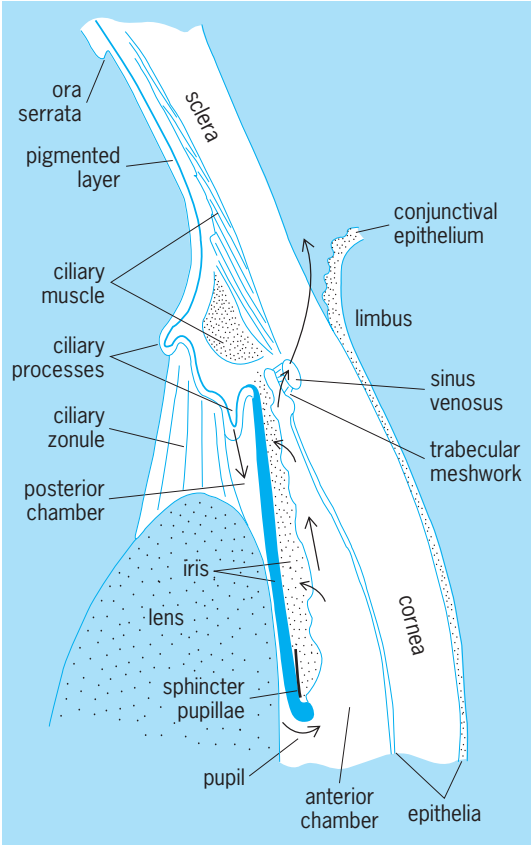

Fig. 3. Meridional section through anterior portion of human eye showing ciliary region and iridocorneal angle. The aqueous circulation (source, route, discharge) is indicated by arrows. (*After E. Gardner* et al., *Anatomy: A Regional Study of Human Structure, 4th ed., Saunders, 1975*)

constant state of deturgescence, and its chemical composition. It is the first ocular component traversed by the incoming light. For animals living in air, the cornea bends (refracts) this light toward the optic axis so that the visual image becomes placed on the sensory portion (retina) of the eye. The cornea provides the greatest refractive medium in these animals, whereas it appears to be optically functionless in animals living in water. Refractive errors such as myopia and astigmatism are common. *See* EYE DISORDERS; REFRACTION OF WAVES.

Sclera. This connective tissue tunic provides support for the eye and serves for the attachment (insertions) of the muscles that move it. It is cartilaginous in all elasmobranchs, anurans, reptiles (except snakes), and birds, as well as in some teleosts, urodeles, and mammals (monotremes).

Limbus. Located at the angle of the anterior chamber (Figs. 1 and 3), this small, circular transitional zone between the cornea and the sclera houses the major route for the discharge of aqueous humor from the anterior chamber: a circular venous sinus (canal of Schlemm) for transport of this fluid to the venous system, and a very porous mass of tissue (trabecular meshwork) which links this sinus with the anterior chamber. In birds and reptiles (except crocodiles and snakes), the limbus is further characterized by the presence of a circle of flat (sometimes concave), overlapping bones, the scleral ossicles, forming a sclerotic ring. Usually 14 in number, these ossicles support a broad and strong limbus with an annular depression which is believed to aid in accommodation (changing the contour or the position of the lens for both near and far vision) in these animals.

Vascular tunic. The vascular tunic or uvea makes up the middle layer of the wall of the eye. It does not form a continuous layer around the eye but is deficient anteriorly, where the opening is termed the pupil. Beginning at the pupil, three continuous components of the uvea can easily be recognized: the iris, ciliary body, and choroid (Fig. 1).

Iris. The iris is a spongy, circular diaphragm of loose, pigmented connective tissue separating the anterior and posterior chambers and housing a hole, the pupil, in its center (Fig. 3). When heavily pigmented, the human iris appears brown; when lightly pigmented, blue. The anterior surface is lined by connective tissue cells with extensive interstices which permit the passage of aqueous humor. The posterior surface is covered by a double layer of pigmented epithelial cells, the iridial part of the retina. Situated in front of this epithelium are two muscle masses concerned with the opening (dilator pupillae) and closing (sphincter pupillae) of the pupil and, therefore, the regulation of light impinging upon the retina.

Pupil. The shape of the pupil varies considerably among vertebrates (**Fig. 4**). Both circular and slitlike pupils are found in all vertebrate classes. Circular forms are the most common; slitlike pupils are adaptations for nocturnality. Most slitlike pupils are oriented vertically, although horizontal slits are present in some species of all classes. In some fish, the iris possesses a single pupillary process or operculum which is capable of expanding into the pupil, relegating it to a circular slit. The greatest variety of pupillary shapes is exhibited by the amphibians.

Ciliary body. The ciliary body, triangular in section, is continuous with the root of the iris. The posterior epithelium of the iris continues along the internal surface of the ciliary body as a double layer of cells (ciliary processes) for the attachment of the ciliary zonule (suspensory ligament) of the lens. This ligament holds the lens in position and shape, and

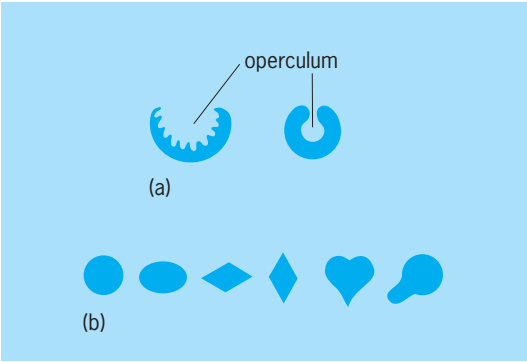

Fig. 4. Examples of various shapes of pupils in vertebrates: (*a*) pupillary opercula in fishes; (*b*) partially contracted pupil in different amphibians.

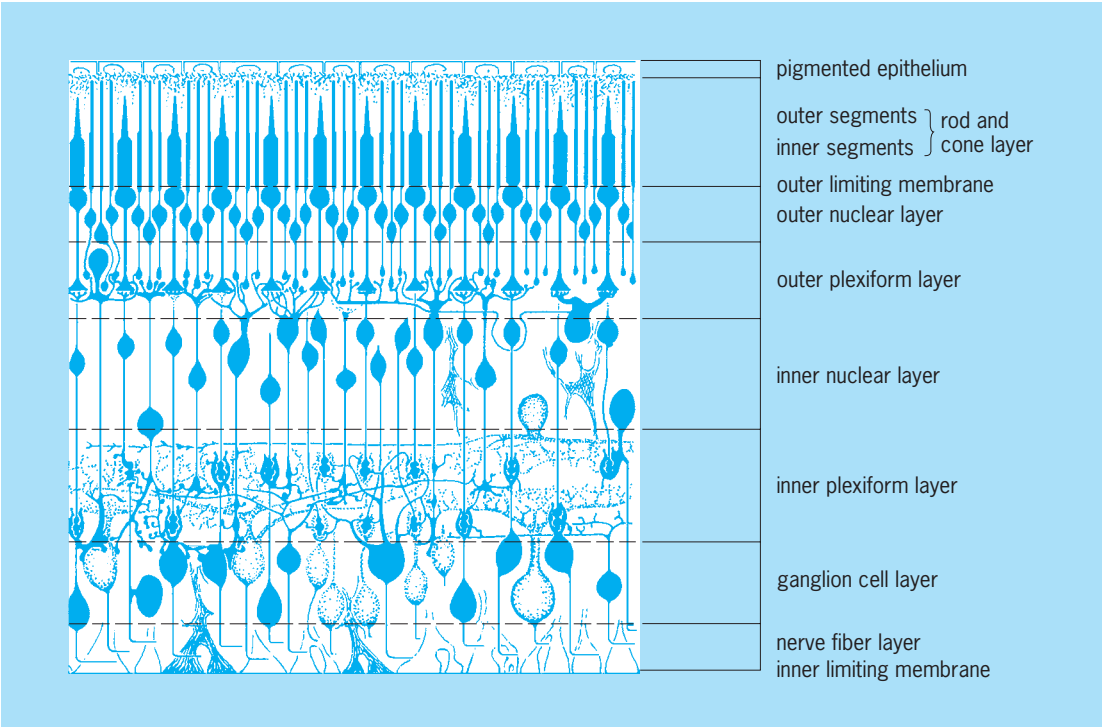

Fig. 5. Cross section of the vertebrate retina.

marks the posterior boundary of the posterior chamber. The inner layer of the ciliary epithelium contains no pigment. It produces aqueous humor which flows into the posterior chamber and thence into the anterior chamber (via the pupil) for final discharge through the sinus venosus (canal of Schlemm) via the trabecular meshwork. The continual production and removal of this fluid maintain the intraocular pressure of the eye (which is increased in glaucoma). The mass of the ciliary body, on the other hand, consists of muscular tissue (smooth in most vertebrates, skeletal in birds and reptiles) concerned with accommodation.

Choroid. The choroid is the most posterior portion of the uvea. It is directly continuous with the subepithelial portion of the ciliary body and consists primarily of blood vessels embedded within deeply pigmented connective tissue. The blood vessels are an important nutritional source for the retina (particularly the visual cells), which lies immediately internal to the choroid; the pigment serves to absorb stray light, thereby preventing internal reflections. In teleosts the choroid in the vicinity of the optic nerve develops a large capillary mass termed the choroidal gland. Although the function of this "gland" is not understood, it has been shown to be well suited for supplying the oxygen requirements of the teleost retina.

Sensory tunic. The retina is the sensory tunic of the eye. It has the form of a cup closely applied to the inner portion of the choroid, and, internally, it is slightly adherent to the semisolid vitreous body. Anteriorly, the sensory layers of the retina terminate at the liplike ora serrata, where they become continuous with the nonsensory epithelium lining the ciliary body (ciliary part of the retina). Derived from the brain, the vertebrate retina contains the light-sensitive receptors (visual cells) and a complex of well-organized impulse-carrying nerve cells (neurons), all arranged into discrete layers in an orderly and remarkably similar fashion (**Fig. 5**). Basically, this stratification is produced by the spatial alignment of a three-neuron linear chain oriented perpendicularly to the surface of the retina, and by the orderly synaptic connections that these neurons make with each other and with other nerve cells (**Fig. 6**).

The visual cells or photoreceptors represent the first neurons in this series. They are nerve cells specialized for the reception of light stimuli and the subsequent conversion of this energy into electrical impulses. Two physiological types of visual cells occur in vertebrates: rods, which operate in dim light (scotopic vision) and are involved with visual sensitivity; and cones, which function in bright light (photopic vision) and are responsible for visual acuity and color perception. Both rods and cones transmit their responses to a second group of neurons, the bipolar cells, by means of a specialized contact or junction called a synapse. The bipolar cells, in turn, carry the signals by a similar means to a third group of neurons, the ganglion cells. Single nerve processes (axons) from each of the ganglion cells then leave the retina together as the optic nerve, and convey their information to higher centers in the brain. Further complicating this circuit is the presence of two other neurons (horizontal and amacrine cells) which maintain similar positions in all vertebrate retinas and which, by their synaptic contacts, must greatly influence the responses carried along the three-neuron

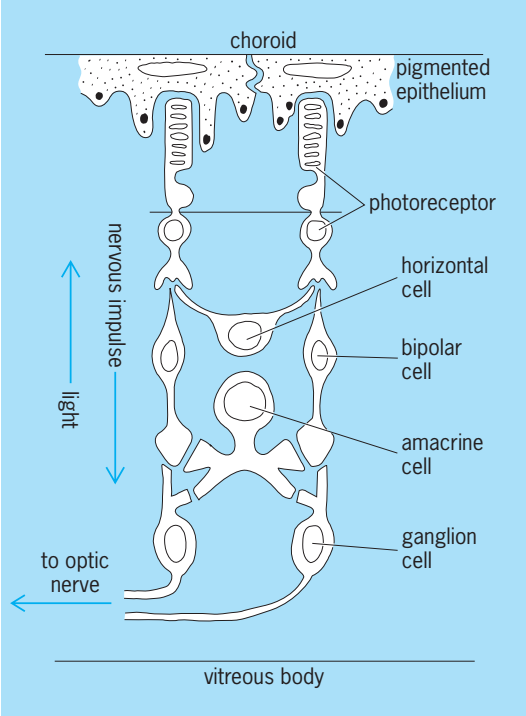

Fig. 6. Basic organization of the vertebrate retina. (*After J. E. Dowling, Synaptic organization of the frog retina: An electron microscopic analysis comparing the retinas of frogs and primates, Proc. Roy. Soc. London, 170B:205–228, 1968*)

chain. Horizontal cells exhibit their influence in the region of the receptor-bipolar junction, amacrines by the bipolar- ganglion cell contacts. Discrete, well-demarcated layers are formed as a result of the spatial alignment of these different cells and their synapses.

Aggregations of nuclei form three layers: outer nuclear, by nuclei from the photoreceptors; inner nuclear, by nuclei of bipolar, horizontal, and amacrine cells; and ganglion cell layer, by ganglion cell nuclei. The synaptic contacts between these nuclei form the plexiform layers. It should be emphasized that the vertebrate retina is an inverted one; that is, its receptors are situated in the deepest (most external) portion of the retina and point away from the incoming light. Light, therefore, must traverse almost the entire retina before striking the photosensitive elements. So that these relationships can be more easily understood, the retina is conveniently divided into 10 layers, beginning with its most external layer, the pigmented epithelium (Fig. 6). Finally, it should also be emphasized that the visual image is not imprinted on the retina, as on a film, but is coded and transmitted to the brain by way of the optic nerve, somewhat as in television. *See* TELEVISION.

Pigmented epithelium. Cuboidal in shape and heavily pigmented, this simple layer of epithelial cells forms an important barrier between the light-sensitive receptors (visual cells) and their blood supply, the choroid. As in the choroid, the pigmentation serves to absorb light and prevent its reflection. In fish, frogs, birds, and some reptiles the pigment granules effectively screen scattered light by migrating among the visual cells in the light and by receding into the cell body in the dark. In higher vertebrates, including the human, this close association between pigmented epithelium and visual cells involves the participation of the epithelium in the turnover of the rod photoreceptors, continuously phagocytizing the growing tips of the outer segments of the rods. Remnants of extruded rod tips, termed phagosomes, have been revealed in the electron microscope within the cytoplasm of the pigment epithelium in various stages of degeneration. A similar process has not been confirmed in cones. Aggregations of tightly packed disks or saccules, termed myeloid bodies, are also found in the cytoplasm of amphibians, birds, and reptiles, but their significance is not known. At the ora serrata the pigmented epithelium continues as the pigmented layer of the ciliary part of the retina.

Tapeta lucida. Many nocturnal or crepuscular vertebrates possess a reflecting layer, or tapetum lucidum, situated within the pigment epithelium (retinal tapetum) or choroid (choroidal tapetum). Densely packed crystals (guanine, riboflavin), rodlets (zinc complex), and lipid droplets or fibrils (collagenous) are responsible for the eyeshine produced by these reflectors, an adaptation for increasing light sensitivity at the expense of visual acuity. Retinal tapeta are found in bony fish, alligators, and the opossum; choroidal tapeta are characteristic of cats and ungulates.

Visual cells. The rods and cones of vertebrates generally occur as single units, but combinations of each type are frequently encountered in several vertebrate classes (**Fig. 7**). Double rods are scarce and are confined to some nocturnal lizards and snakes. Double cones, on the other hand, are distributed widely among vertebrates, for example, amphibians, birds, reptiles, and nonplacental mammals. Physiologically, they are regarded as one of the most obscure elements in vertebrate retinas. Whereas both double rods and double cones involve unequal, unfused members, the twin cones, which are another visual cell type unique to bony fish, exhibit identical components which are fused along the entire extent of their inner segments.

Cones appear to be adapted for photopic, or daylight, vision, based on correlations with the visual habits of the animals involved. Rods, which predominate in nocturnal vertebrates, are adapted for scotopic, or night, vision. Except for their external process, the structure of these cells does not reflect these differences. Indeed, both are organized in a very similar fashion and exhibit almost identical compartmentalization of organelles (**Fig. 8**) in all vertebrates. Shaped like a bipolar cell, each visual element contains a cell body with a nucleus, and two processes, one directed externally (toward the choroid) and functioning as the receptor end organ, the other directed internally (toward the vitreous body) for transmission of the visual image to the bipolar cells. Since the external processes are either rod- or cone-shaped and are easily recognized as such

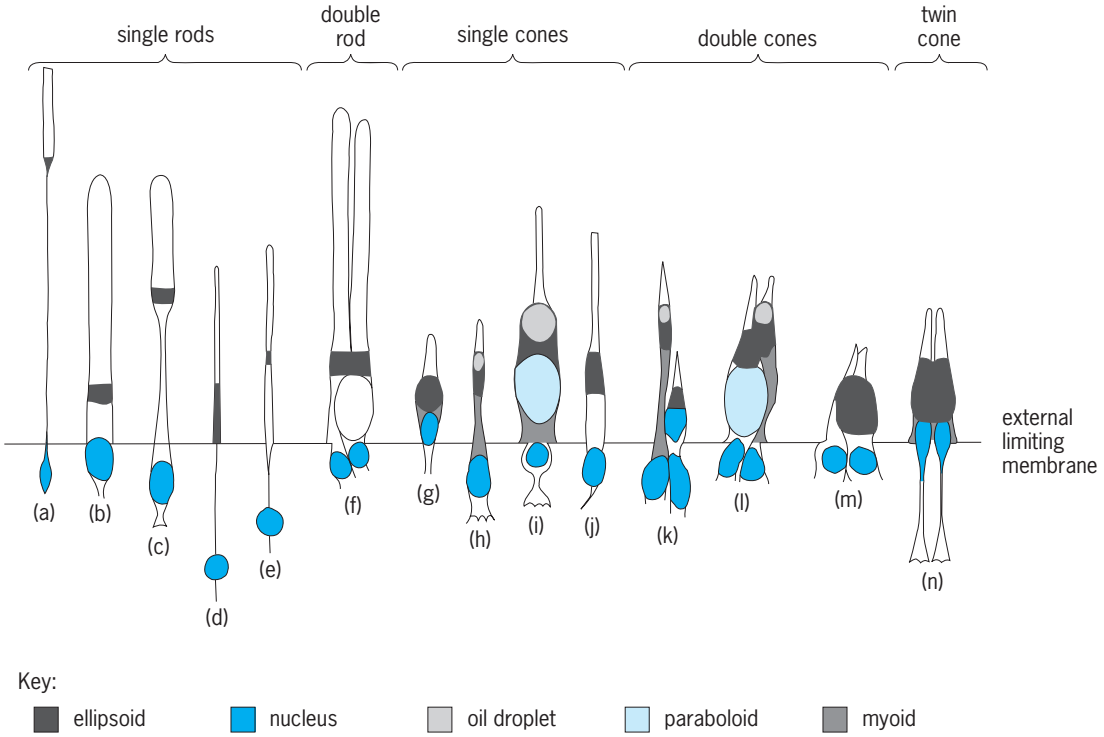

Fig. 7. Types of rods and cones found in different vertebrates: (*a*) goldfish, (*b*, *c*) frogs, (*d*) squirrel, (*e*) human, (*f*) gecko, (*g*) goldfish, (*h*) frog, (*i*) turtle, (*j*) human, (*k*) frog, (*l*) turtle, (*m*) snake, and (*n*) fish.

microscopically, the term "rod" or "cone" has been used to designate the entire cell. The external processes occupy, therefore, the rod and cone layer (Fig. 5). Each such process is divided rather evenly into a proximal inner segment and distal outer segment joined by a thin connecting stalk (Fig. 8). The outer segment represents the photosensitive receptor and characteristically consists of a stack of membranous disks containing a light-sensitive visual pigment. A multiplicity of visual pigments occurs among vertebrates, but all contain vitamin A conjugated to a protein. The response of a vertebrate eye to light is determined by the type of visual pigment, the most well known of which are rhodopsin (rod) and iodopsin (cone). *See* PHOTORECEPTION; VITAMIN A.

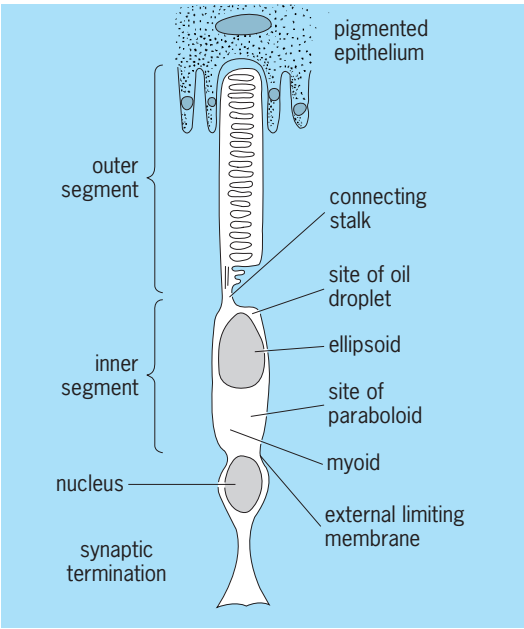

Fig. 8. Typical vertebrate visual cell showing the compartmentalization produced by the spatial arrangement of its organelles and inclusions. An oil droplet and a paraboloid are present only in some species.

The inner segments house most of the cellular organelles and inclusions in a very characteristic compartmental pattern. The mitochondria (respiratory organelles), for example, are always concentrated at the distal ends of the inner segments in the form of an ellipsoid, the name adopted for this entire mass, whereas the rest of the inner segment, termed the myoid, contains the Golgi apparatus and the bulk of the endoplasmic reticulum. In most vertebrates, except mammals, "myoid" is a very appropriate name for this region, because in these animals it has the ability to contract, like muscle. Unlike muscular contractions, however, these contractions are under the influence of light and in coordination with the migration of the pigment of the pigmented epithelium. In light, the pigment migrates toward the visual cells, while the rods elongate and the cones contract. These photomechanical responses shield the outer segments of the rods from the light and make those of the cones more accessible to it. The opposite reactions occur in the dark, when the rods become the dominant visual element.

In birds, reptiles (except snakes), and nonplacental mammals, the distal ends of the cone ellipsoids are further characterized by the presence of brightly colored (red, orange, yellow, or green) oil droplets

which, by their strategic position (to the incoming light) and their coloration, undoubtedly filter and modify the light destined for the outer segment. Rod myoids and some cone myoids in amphibians, birds, and reptiles likewise become compartmentalized by the presence of discrete deposits of glycogen which have a hyperboloid and paraboloid shape, respectively. The function of the rod hyperboloids and cone paraboloids, as they are called, is not known, although they have been thought to improve visual acuity or to provide an energy source for visual cell metabolism, or to perform both functions. The vertebrate retina displays a variety of other structural adaptations, some of which are believed to improve visual abilities under various environmental stresses; others are little understood and appreciated. Among the adaptations that have been shown to be characteristic features of some vertebrate groups are specialized retinal areas for acute vision, with or without a foveal depression and supplementary nutritive devices for the retina, for example, the pecten oculi and the conus papillaris.

Areas and foveae. An important adaptation for improving visual detail in vertebrates is the formation of circumscribed thickenings (circular, oval, or linear) of the retina resulting from localized increases in the number of visual cells and the other retinal neurons associated synaptically with them. Such thickenings, termed areas of acute vision, appear in some members of all vertebrate classes and reach their greatest development in birds, in which one to three distinct areas may be found in the same retina. Only a single area occurs in humans; it is colored yellow and is called the macula.

These areas most commonly develop a depression or pit of various depth and size called a fovea. Formed by a lateral displacement of the cellular components of the retina, foveae most often lie in the central or lateral region, or both, of the fundus (fovea centralis or fovea lateralis), and in birds are positioned so as to lie directly in the visual axis for either monocular (fovea centralis) or binocular (fovea lateralis) vision (**Fig. 9**). In the human only a single fovea is present; it is situated in the center of the fundus and contains only cones.

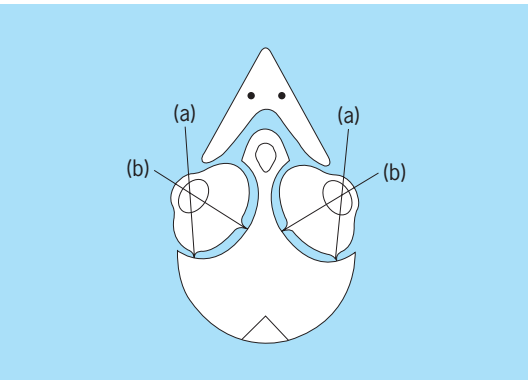

Fig. 9. Visual axes of bifoveate retinas: (*a*) for the lateral fovea and binocular vision; and (*b*) for the central fovea and monocular vision.

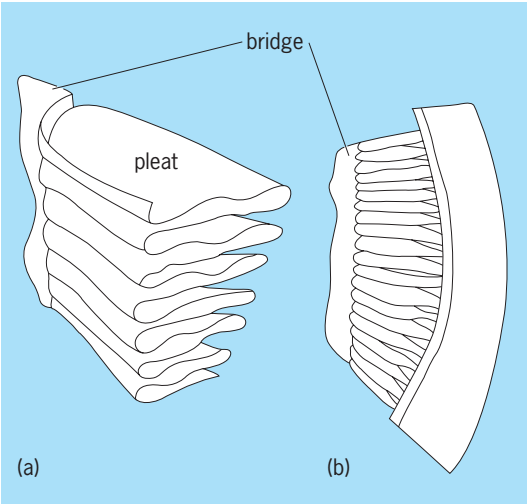

Fig. 10. Pleated pecten of the domestic chicken: (*a*) pleats and "bridge" in three-dimensional view; (*b*) longitudinal view. (*After G. K. Smelser, ed., The Structure of the Eye, Academic Press, 1961*)

Pecten and conus papillaris. In addition to the choroid, the nutritional supply to the vertebrate retina draws on several different sources, the most bizarre of which are the heavily pigmented, vascular bodies arising from the optic nerve head and projecting into the vitreous body of birds (pecten oculi) and lizards (conus papillaris). In lieu of retinal or vitreal vessels, these homologous organs are believed to provide nourishment for the inner layers of the vertebrate retina. The avian pecten has been the most extensively studied and has been found to assume a conical, vaned, or pleated appearance. The pleated type (**Fig. 10**) is by far the most common and varies considerably in size, shape, and number of pleats. The largest and most pleated pectens are found in active diurnal birds with high visual acuity, for example, hawks and eagles. Nocturnal birds with poor vision tend to have smaller and simpler ones. Because of its unique intravitreal location, the pecten has also been implicated in many other functions, particularly those concerned with vision. For example, the pecten is believed to cast shadows upon the retina that aid in the detection of movement and perhaps also, by acting as a sextant, enable the bird to estimate the position of the Sun and thereby aid in navigation. *See* VISION.

The Cavities

Three cavities or chambers are present within the vertebrate eye: anterior, posterior, and vitreous (Fig. 1). The anterior and posterior chambers are continuous with one another at the pupil and are filled with a fluid, the aqueous humor, which is continuously secreted by the ciliary body and removed by the sinus venosus (canal of Schlemm). The eye is normally maintained in a distended state by the (intraocular) pressure created by this fluid. The vitreous cavity, on the other hand, is filled with a semisolid material, the vitreous body, which is fixed in amount and relatively permanent. Its consistency is not uniform

in all vertebrates, however. In mammals and owls, for example, the vitreous body is completely gelatinous, whereas in many birds (pigeon, chicken, turkey) and fish (tuna), it is partly fluid and partly gelatinous.

The Lens

The lens is a transparent body, supported by thin suspensory fibers (ciliary zonule) and by the vitreous body behind and by the iris in front (Figs. 1 and 3). It is completely cellular, the anterior cells forming a thin epithelium, and the posterior cells, much elongated, forming the so-called lens fibers. The entire lens is surrounded by an elastic capsule which serves for the attachment of the ciliary zonule.

In all vertebrates the lens functions in accommodation, either by moving backward and forward or by changing its shape. Some vertebrate lenses are colored yellow, for example, in the lamprey, some snakes and diurnal geckos, most squirrels, and tree shrews and humans. In ground squirrels and prairie dogs they are almost orange in color. Colored lenses, like colored oil droplets, are considered as intraocular filters; they represent adaptations for increasing visual acuity by enhancing color contrasts. In humans lenticular color increases with age. An opacity of the lens is a cataract.

David B. Meyer; Ronan O'Rahilly

Electrophysiology of Rods and Cones

Visual information perceived by the vertebrate eye is fed to the brain in the form of coded electrical impulses that are initiated by the light-sensitive, visual-pigment-containing outer segments of the rods and cones. Light striking the outer segments is absorbed by these pigments, which form an integral component of their membranous discs resulting—in the case of rhodopsin, for example—in the isomerization of the 11-*cis*-retinal chromophore to all *trans*-retinal. The outcome of this photolytic process is a change in electrical activity at the plasma membrane enclosing the outer segments, and a sudden and drastic decrease in its permeability (particularly to Na^+). The net result is a hyperpolarization response, or increased negativity of membrane potential, with the rods exhibiting slightly slower but somewhat longer-lasting responses than cones. Hyperpolarization generates a membrane current that spreads to the inner segment and finally to the synaptic terminal, where it regulates the release of neurotransmitter and thus controls the flow of information from the visual cells to other retinal cells (bipolars, horizontals, other photoreceptors).

Whereas the outer-segment plasma membrane is depolarized in the dark and is permeable to Na^+ and other ions (K^+, Ca^{2+}, Cl^-), the light-adapted hyperpolarized membrane exhibits closed ionic channels (with resultant decrease in intracellular Na^+) in response to a decrease in the level of cyclic guanosine 3′,5′-monophosphate (cyclic GMP) owing to its rapid hydrolysis. Cyclic GMP is directly responsible for regulating the permeability of the plasma membrane by opening these channels (in the light). Its

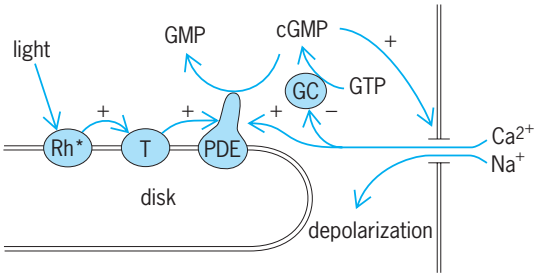

Fig. 11. Summary of the interactions occurring in the rod outer segments during light activation. Within the disc membrane, photolytic rhodopsin (Rh*) activates transducin (T), which in turn activates phosphodiesterase (PDE), leading to the breakdown of cyclic GMP (cGMP) to inactive GMP. In the dark, cyclic GMP maintains open ionic channels in the outer-segment plasma membrane, permitting Na^+ to enter the cell. In the light, decreased cyclic GMP results in a closure of these channels, a decrease in intracellular Na^+ levels, and a hyperpolarization of the cell. A decrease in Ca^{2+} concentration acts primarily as a negative feedback, leading to an increase in cyclic GMP levels. (*After J. E. Dowling, The Retina: An Approachable Part of the Brain, Harvard University Press, 1987*)

concentration is controlled by a peripheral membrane enzyme, phosphodiesterase, which in turn is activated by transducin, an intracellular messenger protein generated by a photolytic intermediate of rhodopsin (**Fig. 11**). Since one molecule of photoactivated rhodopsin can react with many molecules of transducin, an amplification of the visual cells' response is produced, the final amplitude being enhanced by degradation of cyclic GMP by phosphodiesterase and subsequent inhibitory feedback (cones) from the horizontal cells.

Since photoreceptors are depolarized in the dark, their axon terminals continually release a transmitter (for example, acidic amino acids such as L-glutamate) that hyperpolarizes (inhibits) the bipolar cell, and since this cell is hyperpolarized in the dark, it is prevented from releasing its excitatory transmitter at the ganglion cell synapse so that the synapse is not excited. In the light, on the other hand, hyperpolarization of the visual cells causes a decrease in the amount of inhibitory transmitter released at the bipolar synapse, leading to a depolarization of the latter, which in turn increases the amount of excitatory transmitter released at the bipolar-ganglion synapses and affecting the ganglion cells.

A change in the light energy taking place across the retina also initiates a transient action potential, the electroretinogram, which is recorded as a difference in potential between the cornea and the back of the eye and exhibits three principal deflections: A, B, and C waves. The C wave has been attributed to the pigment epithelium but may be triggered by the photoreceptors, which have also been implicated as the site of origin of the A wave.

David B. Meyer

Accommodation

Ocular adjustments for the sharp focusing of objects viewed at different distances is termed accommodation. Each vertebrate class has its own mechanism of accomplishing accommodation, and most of

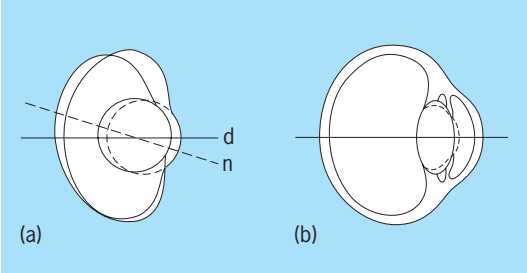

Fig. 12. Methods of accommodation in vertebrates. (*a*) In the teleost (bony fish) eye, on the approach of an object the lens returns to its normal relaxed (forward) position for near vision; it accommodates for distant vision. (*b*) In the mammalian eye, on accommodation for near vision, the lens becomes more curved. The visual axes for near (n) and distant (d) vision are indicated.

these mechanisms involve specific muscular actions. Two basic methods of accommodation are used by vertebrates: movement of the lens, either forward (elasmobranchs, amphibians, snakes) or backward (teleosts, lampreys), and changes in the shape of the lens (reptiles except snakes, birds, and mammals) [**Fig. 12**].

Cyclostomes, fish, and amphibians have developed specialized muscles for lenticular movement—the cornealis, retractor lentis, protractor lentis, respectively—whereas higher vertebrates, including humans, utilize the musculature of the ciliary body to change the shape of the lens. In birds and reptiles (except snakes) lens shape change is accomplished by a squeezing of the lens at the equator. In mammals, contraction of the ciliary muscles relaxes the tension that normally exists in the ciliary zonule, thereby permitting the lens to relax and change shape (become more spherical for nearer objects).

Adaptations

The preference of an animal for a certain type of illumination has produced a variety of ocular adaptations which have served to improve its chances for survival. On the basis of their visual habits, for example, vertebrates are classified into three major groups: diurnal, nocturnal, and arrhythmic.

Diurnal vertebrates. Members of this group are active primarily during daylight, and thus their eyes are suited to sharp vision at the expense of light sensitivity. Lizards, birds, and primates, which feed on small objects (seeds and insects), require such vision, as do animals that depend upon great speed and dexterity to capture prey or elude enemies. Such vertebrates possess large eyes with a large retinal surface housing chiefly cones, as well as rather flattened lenses and a long "throw" of the image from the lens to the retina (**Fig. 13***a* and *b*).

Nocturnal vertebrates. These vertebrates are active chiefly at night. Their eyes, therefore, are modified for greater sensitivity to light at the expense of sharp vision. Vision in these animals, for example, bats, is not as important as sound and touch. Nocturnal eyes are characterized by a rod-rich retina, a larger and broader lens with a larger pupil, and a noticeable decrease in the distance, or throw, from the lens to the retina (Fig. 13*c*). Some nocturnal eyes (for example, in owls, Fig. 13*d*) are tubular in shape, which results in a small retinal image but an increased throw from the lens to the retina.

Arrhythmic vertebrates. Arrhythmic, or 24-h, vertebrates are equally active by day and by night. Among them are the bony fish, frogs, crocodiles, and larger land mammals (elephants, bears, lions, wolves). Because the eyes of these vertebrates must function equally well under both extreme conditions of illumination, they have combined those ocular adaptations which are not mutually exclusive, for example, duplex retinas (rods and cones), visual cell and pigment migration, and mobile pupils. Their lenses are generally intermediate in size (between nocturnal and diurnal), and the throw to the retina is also intermediate in length (Fig. 13*e*).

Development

The development of the human eye is summarized in **Fig. 14**. At 3 weeks, a groove (the optic sulcus) appears on the internal surface of the brain wall (neural plate) (Fig. 14*a*). The sulcus rapidly becomes deeper (Fig. 14*b*), and the communication (rostral neuropore) between the cavity of the brain and the outside of the embryo becomes narrower and soon closes. The optic sulcus has now been transformed into a cavity, which, together with its wall, is known as the optic vesicle (Fig. 14*c* and *d*). The right and left optic vesicles communicate widely with that part of the forebrain termed the diencephalon (Fig. 14*e*). The optic vesicle (**Fig. 15**) induces the overlying surface ectoderm to form a lens disk (Fig. 14*f*), and the subjacent wall of the vesicle constitutes a retinal disk. During the next few days the lens and retinal disks become invaginated to form a lens pit and the optic cup, respectively (Fig. 14*g*). Subsequently, the cavity of the optic cup becomes obliterated, and the two retinal layers maintain contact with each other. The outer layer of the cup becomes the pigmented epithelium of the retina, and the inner, invaginated layer develops into the sensory retina (Fig. 14*h*). By this time (5 weeks) the lens has developed into the lens vesicle and has lost contact with

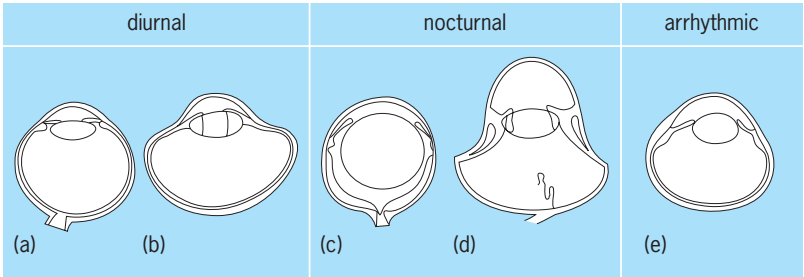

Fig. 13. Intraocular proportions of various species in relation to habit; (*a*) human, (*b*) pigeon, (*c*) mouse, (*d*) owl, and (*e*) dog. A tubular eye, such as that shown in *e*, is also found in birds of prey.

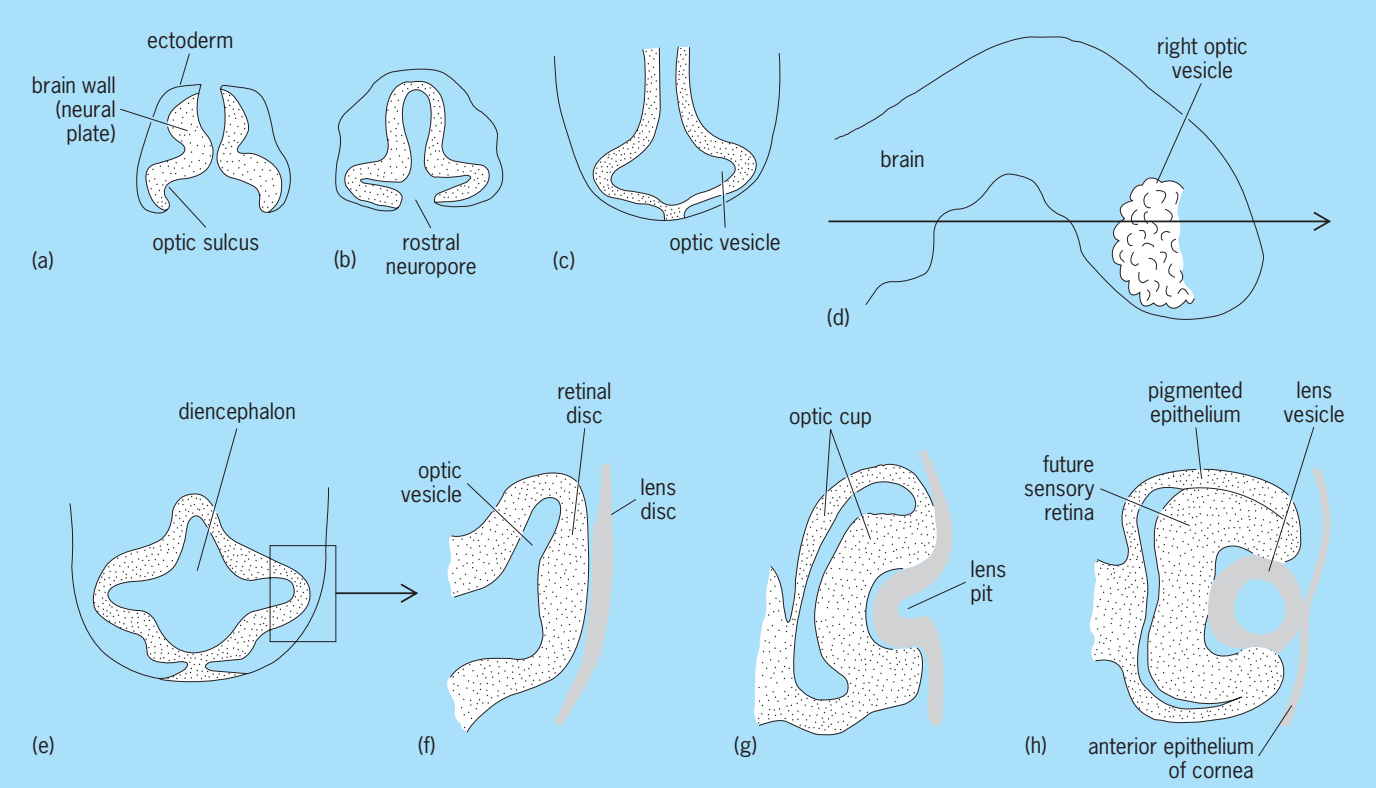

Fig. 14. Development of the eye in the human embryo at (*a*, *b*, *c*) $3^1/_2$ weeks; (*d*, *e*, *f*) 4 weeks; (*g*) $4^1/_2$ weeks; and (*h*) 5 weeks. (*After R. O'Rahilly, Contr. Embryol. Carnegie Inst., 38:1–42, 1966*)

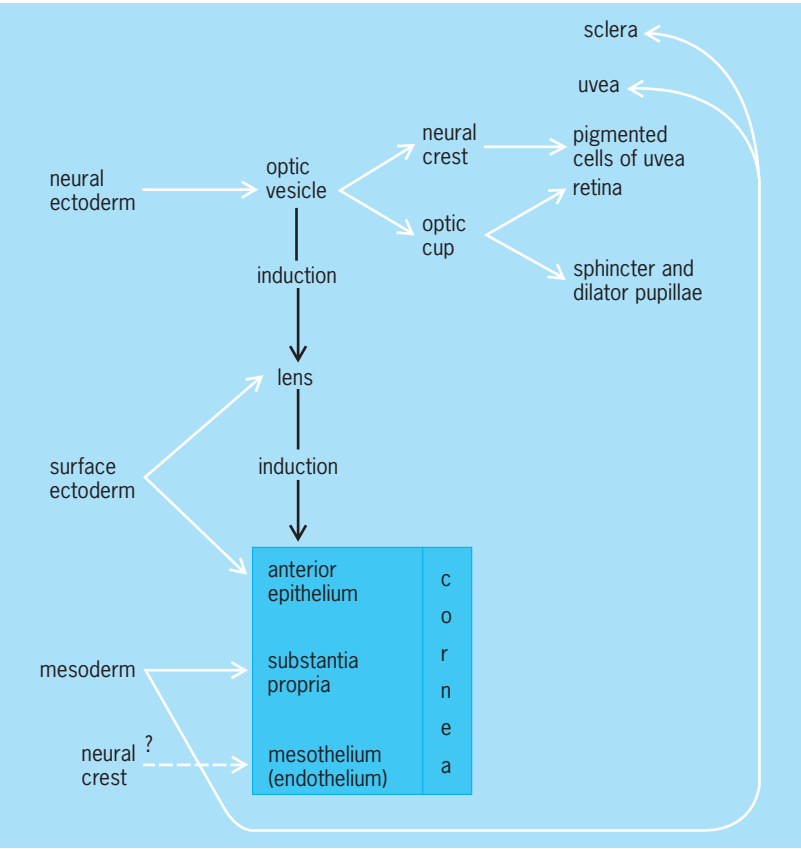

Fig. 15. Simplified scheme showing the chief embryological sources of the ocular tissues.

the surface ectoderm, which becomes the anterior epithelium of the cornea (Fig. 14*b*). Mesodermal invasion contributes to the rest of the cornea, the uvea, and the sclera. The posterior cells of the lens vesicle grow forward and obliterate the lens cavity. A simplified scheme of the derivation of the ocular tissues is provided in Fig. 15. *See* NERVOUS SYSTEM (VERTEBRATE); NEURULATION.

David B. Meyer; Ronan O'Rahilly

Bibliography. M. A. Ali and M. A. Klyne, *Vision in Vertebrates*, 1985; H. Davson, *Physiology of the Eye*, 5th ed., 1991; J. E. Dowling, *The Retina: An Approachable Part of the Brain*, 1987; J. Heggie, *The Use of the Eyes in Movement,* 1985; M. H. Wake (ed.), *Hyman's Comparative Vertebrate Anatomy*, 3d ed., 1992.

Eye disorders

Disorders of form and function that affect the eye. Disorders of form affect the way the eye looks or feels; disorders of function affect vision. Normal vision requires proper performance from the entire visual system, from the precorneal tear film to the occipital cortex. *See* EYE (VERTEBRATE); VISION.

External eye. The eyelids are susceptible to a variety of disorders, many of which are common to the facial skin in general. The eyelids contain many glands (Meibomian, Zeis, Moll, Wolfring, Krause) that are susceptible to acute or chronic infection. Acute infection produces a hordeolum, or sty, which is a

localized nodule up to several millimeters in diameter. Moderate discomfort and lid swelling may develop rapidly. The lids are easily distensible, and a soft, puffy swelling may involve the entire upper or lower lid. Chronic infection results in a granulomatous nodule, or chalazion, which involutes spontaneously; however, this regression may take several months to a year or more. In many cases, surgical removal is indicated.

Conjunctival and corneal disorders. The conjunctiva is a mucous membrane that lines the inner surface of the eyelids and covers the outer surface of the anterior third of the globe except for the corneal epithelium, with which it is contiguous. Because the conjunctiva is connected to the inner surface of the eyelids and the anterior surface of the globe, it forms a barrier between the outside world and the retrobulbar space. Conjunctivitis is an inflammation of the conjunctiva, regardless of cause, which may include viral, bacterial, or chlamydial infection; mechanical or chemical trauma; allergy; and so on. The symptoms of conjunctivitis depend on the cause and generally include conjunctival redness (hyperemia), swelling (chemosis), mild to moderate discomfort, and tearing. If the conjunctiva is the only structure involved, vision is usually affected minimally or not at all.

Infectious conjunctivitis is often termed pink eye because of the bright red appearance of the conjunctiva. Bacterial infections are usually responsible for pink eye in children. Infectious conjunctivitis in adults is commonly caused by viral agents. Chlamydial infections are responsible for inclusion conjunctivitis and ophthalmia neonatorum in the United States and trachoma in many arid regions of the third world. Repeated chronic chlamydial infections characteristic of trachoma cause progressive conjunctival scarring and shortening, resulting in an inturning of the eyelids, and trauma to the cornea by eyelashes (trichiasis), exposure to the elements, a lack of lubrication, and abrasion by the scarred inner surface of the eyelids. In many villages in northern Africa, for example, trachoma is the leading cause of blindness in adults. Like endemic infections in any part of the world, trachoma is treated by behavior modification and specific therapy.

The cornea and conjunctiva are adjacent. The cornea is the window through which light passes to the retina; harm to its form or function is likely to result in visual impairment ranging from mild image distortion to blindness. Trachoma is a case in point. The initial infection involves the conjunctiva, but ultimately the disease results in blindness due to a scarred, vascularized, opaque cornea. *See* VISUAL IMPAIRMENT.

If the cornea and conjunctiva are both involved, the condition is referred to as keratoconjunctivitis. Adenovirus infection causes keratoconjunctivitis. If lubrication by the lacrimal glands or accessory lacrimal glands is interrupted by infection, trauma, or autoimmune disease, one of the dry-eye syndromes may ensue. The most common of these syndromes frequently affects the aging population, causing a nonspecific loss in the quantity and quality of the tears. The symptoms include a mild foreign-body sensation, light sensitivity (photophobia), and gritty feeling in the eyes that progresses throughout the day. Paradoxically, increased tearing (epiphora) can be a symptom of dry eyes. In such cases, tear function is poor because the normal mucoid and oily components of the precorneal tear film are insufficient. Localized irritation results and leads to reflex increase in the secretion of the lacrimal or watery component of the tear film. Keratoconjunctivitis sicca is an autoimmune disease in which the lacrimal and salivary glands are diffusely infiltrated with an inflammatory reaction. Therapy may be difficult in severe cases.

Glaucoma. Glaucoma is a serious eye disorder with many subclassifications. Combined, they affect 1–2% of the population of the United States and Europe, and as much as 10–12% of the population in certain isolated regions of the world. Glaucoma is characterized by intraocular pressure sufficiently high to cause characteristic damage to the optic nerve. Acute glaucoma is associated with pain, redness of the eye, and sudden, often reversible, visual loss. Chronic glaucoma generally causes no symptoms other than progressive peripheral visual loss. Chronic glaucoma is often called the silent thief of vision because the person may not notice a problem until most of the vision has been destroyed. If untreated, the glaucomatous eye may lose all visual functions. *See* GLAUCOMA.

Cataract. The leading cause of blindness in the world is cataract, an opacity within the crystalline lens. Cataracts can be caused by trauma, chronic drug use, or endocrine imbalance, but the overwhelming majority have no specific cause and are associated with aging. Except for rare cataracts caused by medication or associated with systemic disorders, there is no non-surgical treatment. Effective therapy almost always involves surgical removal of the lens, during which a prosthetic intraocular lens is usually implanted. The typical symptom of cataract formation is a gradual painless loss of vision in one or both eyes. Because the process is largely limited to a physical clouding of a single refractive structure, surgery frequently restores excellent vision. *See* CATARACT.

Retinal disorders. The retina is the sense organ of the eye. Its rods and cones, or photoreceptors, and their supporting cellular elements receive light energy and convert it to electrical impulses which travel along the optic nerve to the brain. Degeneration of the retina produces a painless distortion or loss of vision. The function may be disrupted in many ways. The retina itself may be affected by by infectious processes. Cytomegalovirus retinitis, for instance, is a frequent cause of visual loss in individuals with acquired immune deficiency syndrome (AIDS). Primary retinal degeneration may occur in many forms. Retinitis pigmentosa tends to manifest itself early in life and progresses from peripheral to central visual loss; prognosis depends in part on the hereditary pattern of the disease. Involutional macular degeneration is fairly common in the elderly. It affects the central vision first and rarely spreads

to the peripheral vision. Between these two extremes is a very large number of progressive retinal dystrophies and degenerations, involving central or peripheral vision or both. Although the ultimate visual loss is variable, it is often severe.

Retinal detachment occurs when the sensory layers of the retina are separated from their supporting foundations. It is classified as rhegmatogenous (caused by a retinal hole) or nonrhegmatogenous. Rhegmatogenous retinal detachment occurs spontaneously or following trauma. Nonrhegmatogenous retinal detachment occurs as a final stage of such pathologic conditions as retinopathy of prematurity or diabetic retinopathy. The symptoms of retinal detachment are a painless and sudden segmental or total visual loss in one eye. Treatment is aimed at reestablishing the connection between the sensory-neural retina and its supporting structures. In some complicated cases, or when there has been pathologic involvement of the vitreous body, the vitreous may need to be removed surgically. This is accomplished with vitrectomy instruments that amputate and aspirate small amounts of the vitreous gel.

Diabetic retinopathy is a common cause of severe retinal disease. Research has developed many methods of interrupting the progression of diabetic retinopathy. The most effective include laser therapy for localized retinal edema, hemorrhage, and neovascularization, and vitrectomy for late-stage diabetic retinopathy with vitreous hemorrhage and retinal traction. Michael Drake

Bibliography. M. L. Callaham (ed.), *Current Therapy in Emergency Medicine*, 1987; F. W. Newell, *Ophthalmology: Principles and Concepts*, 8th ed., 1996; D. Vaughan and T. Asbury, *General Ophthalmology*, 15th ed., 1998.

Eyeglasses and contact lenses

Lenses worn to improve vision. The primary function of the eye is to focus light energy on the retina. In the optics of a normal eye, parallel rays of light are refracted into sharply focused points of light on the retina. Refraction is necessary because all natural light rays are divergent. *See* EYE (VERTEBRATE); REFRACTION OF WAVES.

The degree of divergence (or minus power) varies with distance: the closer the observer to the light source, the greater the divergence. The unit of measure of divergence is the diopter; the divergence of light rays in diopters is the reciprocal of the distance of the observer from the light source. For example, if the observer is 0.1 m (4 in.) from the light source, the light rays reaching that observer's eye have a divergence of 10 diopters; 6 m (20 ft) from the source the light rays have a divergence of one-sixth of one diopter. For practical purposes, the human eye cannot distinguish rays with one-sixth of one diopter from parallel rays, and therefore 6 m (20 ft) is referred to as optical infinity. The eye is approximately $1/\sqrt[3]{60}$ m (1.67 cm or 0.65 in.) and therefore requires about 60 diopters of converging (plus) power to bring light rays parallel at the cornea into focus on the retina. *See* GEOMETRICAL OPTICS; VISION.

Myopia. If the eye is longer than 1/60 m (0.65 in.), or if its converging media (cornea and lens) impart more than 60 diopters to parallel light rays, they come into focus in the vitreous cavity in front of the retina, and the image on the retina is made up of blurred circles rather than focused points of light. In either case the eye possesses too much converging or plus power to form a good image from parallel light rays.

This may be remedied in two ways. First, the observer may move closer to the object. At the correct distance from a myopic eye, the divergence due to proximity to the object (the minus power of the rays reaching the eye) neutralizes the excess plus power of the eye's refracting system, thereby giving a sharp image on the retina. This explains the term nearsighted. A myopic person clearly sees things close to the eye.

A second method of correcting nearsightedness is to introduce a minus lens into the path of light striking the eye. This can be done with concave lenses, thus adding divergence to the light entering the eye. This divergence combines with the excess convergence of the eye to give clear images.

Hyperopia. The converse is true for farsightedness. The farsighted or hyperopic eye is too short or too weak in refracting power. The convergence applied to incoming light is insufficient to focus it at the retina, and blurred circles form rather than focused points of light. To correct this, plus lenses that add convergence to incoming light are used.

Presbyopia. Accommodation is the ability of the eye to change its refractive power for sharp viewing across different distances. When objects are held close for reading or detailed examination, the refracting (plus) power of the normal eye should increase to focus the more divergent light on the retina. Ability to accommodate decreases with age, and by the mid-forties or early fifties it is often difficult to read print at a normal distance. This condition is known as presbyopia and is corrected with reading glasses, or bifocals in which the distance prescription is in the upper segment and plus power in a lower segment.

Contact lenses. In addition to eyeglasses worn in frames, contact lenses can be used to assist the eye in forming a sharp image. The contact lens rests directly on the precorneal tear film; its front surface becomes the initial refracting surface of the eye. The space between the lens and the cornea is filled with the precorneal tear film. Contact lenses can be either divergent or convergent.

Contact lenses are classified as hard, soft, or gas-permeable hard according to the materials that they are made of, their rigidity or flexibility, and their permeability to oxygen and other gases. Traditional hard contact lenses are the most rigid. They float on the precorneal tear film, through which the cornea must eliminate most of its waste. Hard contact lenses tend to be smaller in diameter than the cornea and to move freely over it if fitting properly. These lenses

provide excellent vision correction where suitable, but because they require an active precorneal tear film, they tend to be tolerated better at a younger age.

Soft lenses are usually larger in diameter than hard contact lenses and are very thin and flexible. They are made of different materials than the hard lenses and are permeable to oxygen and other gases. Soft contact lenses tend to take the shape of the cornea, and if it has an irregular curvature, may not correct the vision as effectively as hard contact lenses. Soft contact lenses tend to be more comfortable and better tolerated, particularly by older people whose precorneal tear film would not support a hard contact lens.

A gas-permeable contact lens is an improvement of the hard contact lens. It is more rigid than a soft lens but more permeable to gases than the older hard contact lenses. Such lenses may give satisfactory correction where the corneas are irregular or for other reasons are not corrected satisfactorily with soft contact lenses. They tend to be better tolerated than ordinary hard contact lenses.

Contact lenses are not entirely harmless, and have been associated with a specific immune response known as giant papillary conjunctivitis and with dangerous infections capable of destroying an eye. In some users, extended-wear contact lenses may be left in place for several weeks to months at a time. Disposable contact lenses are also available, which can be left in the eye for 1–2 weeks. For many individuals, contact lenses are an effective alternative to spectacles. *See* LENS (OPTICS). Michael Drake

Bibliography. D. D. Michaels, *Visual Optics and Refraction: A Clinical Approach*, 3d ed., 1985; M. L. Rubin, *Optics for Clinicians*, 2d ed., 1974, revised 1993.

Eyepiece

A lens or optical system which offers to the eye the image originating from another system (the objective), at a suitable viewing distance. The image can be virtual. *See* OPTICAL IMAGE.

In modern instruments, most eyepieces (also called oculars) are not independently corrected for all errors. They are designed to balance out certain residual aberrations of the objective or (as in the microscope) of a group of objectives, for instance, chromatic difference of magnification. *See* ABERRATION (OPTICS).

The Ramsden eyepiece consists of two planoconvex lenses, the field lens and the eye lens, with their plane sides out. Both of these lenses have the same power and focal length; their separation is equal to their common focal length. The field lens is near the image formed by the objective. The Ramsden acts as a field stop. The Kellner eyepiece is a Ramsden eyepiece with an achromatic eye lens. *See* GEOMETRICAL OPTICS; LENS (OPTICS).

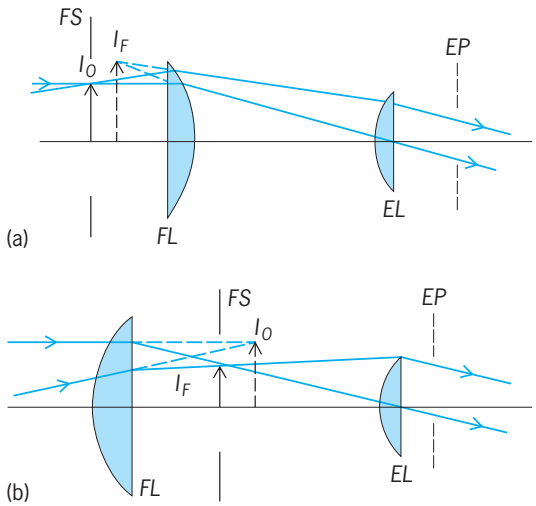

Eyepieces: (*a*) Ramsden and (*b*) Huygens. *FL* = field lens; *EL* = eye lens; *FS* = field stop; *EP* = exit pupil or eye point; I_O = image (real or virtual) formed by the preceding system; I_F = image (virtual or real) formed by the preceding system and the field lens.

The Huygens eyepiece also consists of two planoconvex lenses, but the plane sides of both lenses face the eye. The focal length of the field lens is in general three times that of the eye lens, and the separation is twice the focal length of the eye lens. The Huygens eyepiece is usually corrected for astigmatism and distortion. A type of Huygens eyepiece in which the eye lens is achromatized to compensate for the color errors of the objective is called a compensating eyepiece. The field lens may also be achromatized.

Huygens and Ramsden eyepieces (see **illus.**) are in general used for low magnifying powers. For higher powers, eyepieces assume more complex forms to correct astigmatism, field curvature, and distortion for larger field angles.

The orthoscopic eyepiece consists of a single lens, made up of three cemented elements, to which a planoconvex lens is added.

Complex types of oculars enable corrections to be made over a field angle of 70° or more on the image side. They are used especially in military instruments. Max Herzberger

Bibliography. M. Laikin, *Lens Design*, 1990; D. Malacara and Z. Malacara, *Handbook of Lens Design*, 1994; Optical Society of America, *Handbook of Optics*, 2 vols., 2d ed., 1994; W. J. Smith, *Modern Optical Engineering*, 3d ed., 2000; W. J. Smith and Genesee Optics Software, *Modern Lens Design: A Resource Manual*, 1992.

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

7 F - GEN

www.MHEST.com

F

Fabales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Rosidae of the class Magnoliopsida (dicotyledons). The order consists of three closely related families (Mimosaceae, Caesalpiniaceae, Fabaceae) collectively called the legumes. Members of the order typically have stipulate, compound leaves, 10–many stamens which are often united by the filaments, and a single carpel which gives rise to a dry fruit (legume) that opens at maturity by splitting along two sutures, releasing the non-endospermous seeds. Many, or perhaps most, members of the order harbor symbiotic nitrogen-fixing bacteria in the roots.

Mimosaceae includes some 40 genera and 2000 species, most of which are woody and occur in the warmer regions of the world. The flowers are regular, and the stamens at least twice as many as the petals, and often very numerous. Common members of the family are *Acacia, Albizia, Mimosa,* and *Prosopis* (mesquite).

Caesalpiniaceae includes about 150 genera and 2200 species, most of which are woody and of tropical or subtropical regions. The flowers are somewhat irregular, typically with five (or fewer) unequal petals and 10 or fewer stamens. The family includes the redbud (*Cercis*), Kentucky coffee tree (*Gymnocladus*), honey locust (*Gleditsia*), and numerous tropical ornamentals.

Fabaceae (sometimes termed the Papilionoideae) has about 400 genera and at least 10,000 species of trees, shrubs, vines, and herbs, which are widespread in the tropical and temperate regions. The flowers are characteristically irregular, with a large upper petal (the banner or standard), two wing petals, and the lower two petals fused to form a boat-shaped keel. Usually there are 10 stamens with 9 of them united by their filaments, and the tenth one more or less separate. Well-known members of the family include beans (*Phaseolus*), peas (*Pisum*), sweet peas (*Lathyrus*), clover (*Trifolium*), vetch (*Vicia*), alfalfa (*Medicago*), and lupine (*Lupinus*). *See* ALFALFA; BEAN; CLOVER; COWPEA; KENTUCKY COFFEE TREE; KUDZU; LESPEDEZA; LOCUST (FORESTRY); LUPINE; PEA; PEANUT; SOYBEAN; VETCH. T. M. Barkley

Facies (geology)

Any observable attribute of rocks, such as overall appearance, composition, or conditions of formation, and changes that may occur in these attributes over a geographic area. The term "facies" is widely used in connection with sedimentary rock bodies, but is not restricted to them. In general, facies are not defined for sedimentary rocks by features produced during weathering, metamorphism, or structural disturbance. In metamorphic rocks specifically, however, facies may be identified by the presence of minerals that denote degrees of metamorphic change.

Sedimentary Facies

The term "sedimentary facies" is applied to bodies of sedimentary rock on the basis of descriptive or interpretive characteristics. Descriptive facies are based on lithologic features such as composition, grain size, bedding characteristics, and sedimentary structures (lithofacies); on biological (fossil) components (biofacies); or on both. Individual lithofacies or biofacies may be single beds a few millimeters thick or a succession of beds tens to hundreds of meters thick. For example, a river deposit may consist of decimeter-thick beds of a conglomerate lithofacies interbedded with a cross-bedded sandstone lithofacies. The fill of certain major Paleozoic basins may be divided into units hundreds of meters thick comprising a shelly facies, containing such fossils as brachiopods and trilobites, and graptolitic facies. The scale of an individual lithofacies or biofacies unit depends on the level of detail incorporated in its definition, and is determined only by criteria of convenience, or

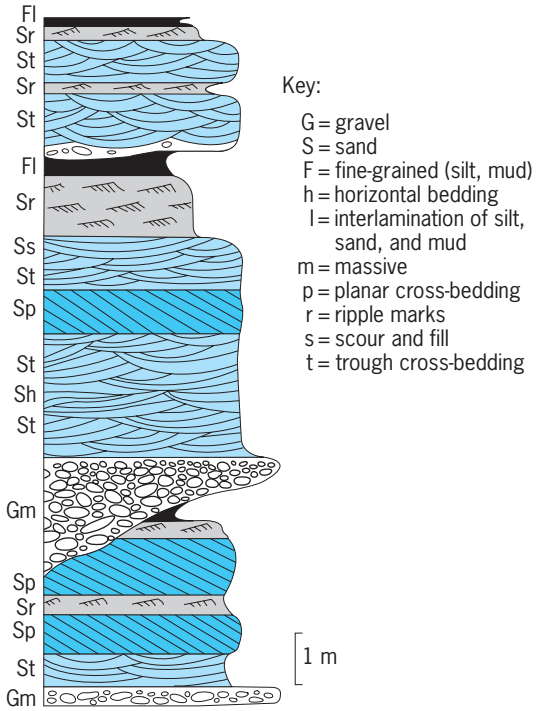

Fig. 1. Stratigraphic section through a fluvial deposit illustrating its subdivision into lithofacies and the use of a lithofacies code scheme. 1 m = 3.3 ft. (*After A. D. Miall, A review of the braided river depositional environment, Earth Sci. Rev., 13:1–62, 1977*)

availability of outcrop and subsurface data, depending on whether a geologist wishes to describe a single outcrop or an entire sedimentary basin.

Facies analysis. The term "facies" can be used in an interpretive sense for groups of rocks that are thought to have been formed under similar conditions. This usage may emphasize specific depositional processes such as a turbidite facies, or a particular depositional environment such as a shelf carbonate facies, encompassing a range of depositional processes. *See* TURBIDITE.

Facies analysis is an important component of stratigraphic and sedimentary geology because it provides a methodology of systematic description that is fundamental to orderly stratigraphic documentation and to interpretations of depositional environment and paleogeography. Most sedimentological basin studies should now begin with such an analysis, in which the succession is divided into its component lithofacies and their characteristics described carefully. Lithofacies may be identified by name or by a code system devised by the researcher, the latter method being convenient for abbreviated description and notetaking and for computer processing (**Fig. 1**).

Increasing use is being made of three-dimensional geometric (architectural) information regarding the extent of facies units, because such information commonly is environmentally diagnostic.

Mappable stratigraphic units commonly consist of a single lithofacies or a small assemblage of lithofacies. This provides the basis for a hierarchical subdivision of stratigraphic sequences ranging from the individual bed or lithofacies unit through member or lithosome and the higher-ranking subdivisions such as stratigraphic sequence, formation, and group. *See* FACIES (GEOLOGY); STRATIGRAPHY.

Facies assemblages. Groups of facies (usually lithofacies) that are commonly found together in the sedimentary record are known as facies assemblages or facies associations. These groupings provide the basis for defining broader, interpretive facies for the purpose of paleogeographic reconstruction. For example, a delta plain facies may be defined, consisting of distributary channel sandstones, levee sandstones and siltstones, and backswamp mudstones and coal. This may pass seaward into a delta front facies consisting of distributary mouth bar sandstones and interdistributary bay mudstones and shell beds. Still further seaward this may pass into a fine-grained prodelta facies, and landward there may be a fluvial facies and a lacustrine facies. An example of a carbonate facies assemblage is illustrated in **Fig. 2**.

The shelly facies and graptolitic facies described earlier are biofacies terms, but they correspond to distinct lithofacies assemblages dominated by shelf carbonate sediments and by deep-water mudstones and siltstones, respectively. Use of the terms provides a convenient way to describe the broad paleogeography of some of the major Paleozoic sedimentary basins around the margins of North America and elsewhere. *See* PALEOGEOGRAPHY.

Many facies assemblages are cyclic in the sense that the individual lithofacies components tend to follow each other in a similar order in a stratigraphic succession, reflecting the repeated occurrence of the same depositional processes. For example, coal beds commonly (but not invariably) follow a seat

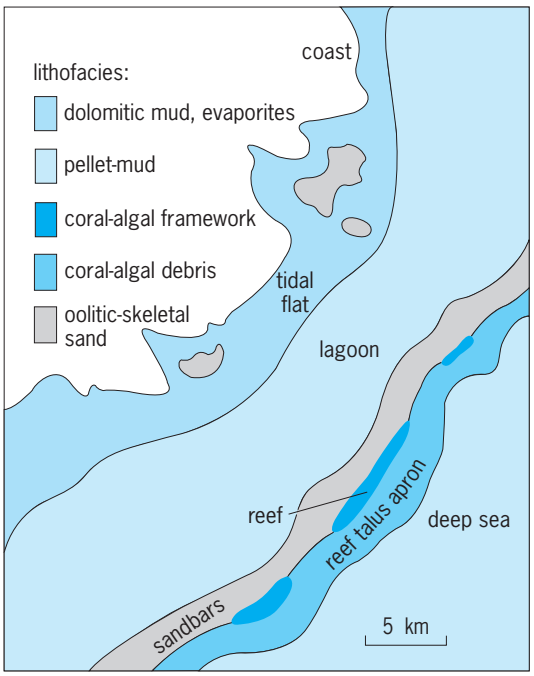

Fig. 2. Hypothetical carbonate shelf showing relationship of lithofacies to environmental facies; based on the modern shelves of Florida and the Bahamas. 1 km = 0.6 mi.

earth or root bed. Such assemblages are termed cyclothems or, simply, cycles. Much sedimentological research has been devoted to analysis and interpretation of cycles, which occur in most depositional settings. For example, the Carboniferous coal-bearing deposits of Europe and of central and eastern North America consitute the original cyclothems. Problems have arisen because these and other cycles can originate from several causes, the effects of which may be very similar and may be superimposed. Thus, cyclothems are caused partly by the advance and abandonment of delta lobes and partly by oscillatory changes of sea level. Second, progressive changes in a sedimentary basin, such as the slow seaward progradation of a delta or tidal flat, may be interrupted by rare random events, for example, a hurricane, which can erode much of the previous cyclic sedimentary record and leave its own unique lithofacies instead. Analysis of cyclicity is based on application of Walther's law. This law, first defined in the nineteenth century, stemmed from the recognition that vertical successions of strata are formed largely by lateral migration of sedimentary environments. The law states that only those facies found side by side in nature can occur superimposed on each other in vertical succession. Recognition of widespread cyclic units is the basis of sequence stratigraphy. Sequences record regional or global changes in sea level brought about by tectonics or eustacy. *See* CYCLOTHEM; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS.

Facies models. These have been defined for most sedimentary environments. Each model is intended as a summary of the depositional processes and products occurring in a specific setting, such as a tidal flat or a submarine fan. Such models are useful guides when carrying out preliminary interpretations of an unknown sedimentary succession, because they provide a range of criteria relating to lithofacies, sedimentary structures, biofacies, and cyclic characteristics that the geologist can use to confirm or negate a particular interpretation.

Facies maps. These are constructed to illustrate lithologic variations across a sedimentary basin, and they usually reveal patterns or trends that can readily be interpreted in terms of depositional environments and paleogeography. Various quantitative techniques have been devised for emphasizing relevant features. Ratio maps are particularly useful, for example, ones showing the ratio of total thickness of sandstone beds to total shale or total carbonate thickness. In elastic coastal environments such a map may reveal the location and shape of major sandstone bodies, which can then, from their size, shape, and orientation, be interpreted in terms of an appropriate sedimentary model (delta, barrier island system, submarine fan, and so on) in conjunction with diagnostic sedimentological criteria. Location and trend may be of considerable economic importance if, for example, porous sandstone bodies are known to contain pools of oil or gas. *See* BARRIER ISLANDS; PETROLEUM GEOLOGY.

Another technique is to plot the percentage of sandstone or aggregate sandstone thickness in a section. **Figure 3** illustrates a map of this type, and reveals a hitherto unsuspected pattern of lobate sandstone bodies containing a readily identifiable deltaic shape. Thicker sandstone beds lie closer to the sediment source, and so in this example transport was clearly from northwest to southeast.

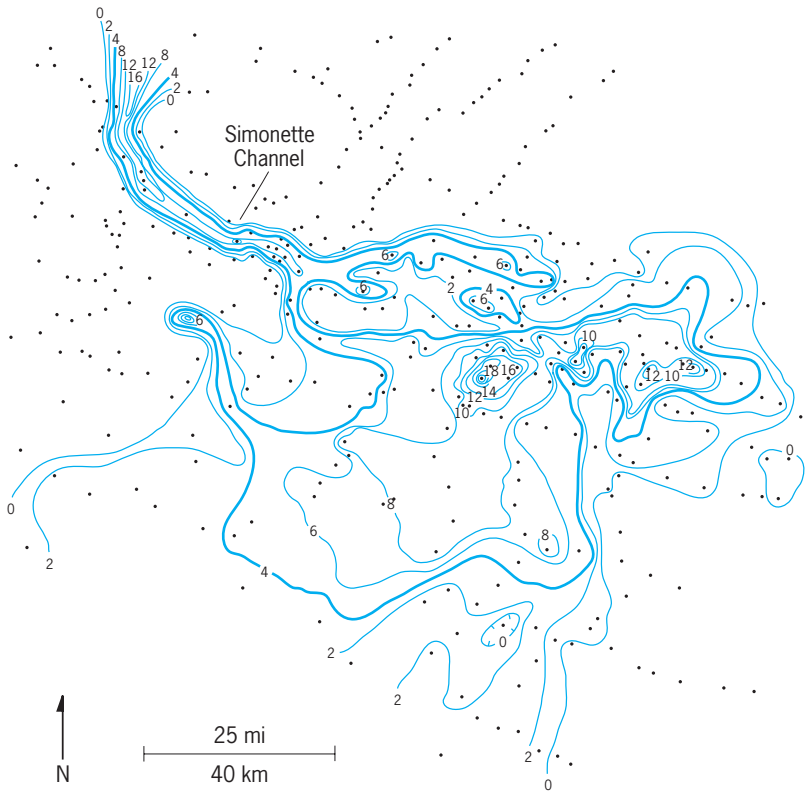

Fig. 3. Plot of sandstone thickness, in meters, in a Cretaceous sandstone-shale unit, derived from subsurface log analysis. Note the narrow sandstone-filled channel representing a delta distributary channel entering the area from the northwest, and the broad lobe of sandstone formed at the mouth of this channel where the delta splayed out into the sea. (*After J. Bhattacharya, Regional to subregional facies architecture of river-dominated deltas in the Alberta subsurface, Upper Cretaceous Dunvegan Formation, in Society for Sedimentary Geology, Concepts in Sedimentology and Paleontology, vol. 3, 1991*)

Similar techniques can be applied to many successions of carbonate rocks. Thus carbonates are typically classified in terms of the ratio of biogenic or detrital fragments to carbonate and matrix. This classification reflects depositional energy (wave and current strength) because the mud matrix tends to be winnowed out in high-energy environments. Therefore, a plot of the ratio can be a useful paleogeographic indicator. However, carbonates are particularly susceptible to diagenetic alteration, which can obscure primary depositional characteristics and hinder detailed paleogeographic reconstruction. *See* DIAGENESIS; SEDIMENTARY ROCKS; SEDIMENTOLOGY; STREAM TRANSPORT AND DEPOSITION.

Andrew D. Miall

Metamorphic Facies

A metamorphic facies is a collection of rocks containing characteristic mineral assemblages developed in response to burial and heating to similar depths and temperatures. It can represent either the diagnostic

mineral assemblages that indicate the physical conditions of metamorphism or the pressure–temperature conditions that produce a particular assemblage in a rock of a specific composition.

Changes in pressure and temperature, due to tectonic activity or intrusion of magma, cause minerals in rocks to react with one another to produce new mineral assemblages and textures. This is the process that transforms igneous and sedimentary rocks into metamorphic rocks. The mineral assemblage that is preserved in a metamorphic rock reflects the composition of the original igneous or sedimentary parent and the pressure–temperature history experienced by the rock during metamorphism. Hence, rocks of the same initial composition that are subjected to the same pressure–temperature conditions will contain identical mineral assemblages.

Facies names and boundaries. The metamorphic facies to which a rock belongs can be identified from the mineral assemblage present in the rock; the pressure and temperature conditions represented by each facies are broadly known from experimental laboratory work on mineral stabilities. Several metamorphic facies names are commonly accepted, and these are shown, along with their approximate pressure–temperature fields, in **Fig. 4*a***. These facies names are based on the mineral assemblages that develop during metamorphism of a rock with the composition of a basalt, which is a volcanic rock rich in iron and magnesium and with relatively little silica. For example, the dominant mineral of the blueschist facies (in a rock of basaltic composition) is a sodium- and magnesium-bearing silicate called glaucophane, which is dark blue in outcrop and blue or violet when viewed under the microscope. Characteristic minerals of the greenschist facies include chlorite and actinolite, both of which are green in outcrop and under the microscope. Basaltic rocks metamorphosed in the amphibolite facies are largely composed of an amphibole called hornblende. The granulite facies takes its name from a texture rather than a specific mineral: the pyroxenes and plagioclase that are common minerals in granulite facies rocks typically form rounded crystals of similar size that give the rock a granular fabric. *See* AMPHIBOLE; AMPHIBOLITE; BASALT; BLUESCHIST; GLAUCOPHANE; GRANULITE; PYROXENE.

Rocks of different composition show different mineral assemblages in each of the facies. A shale metamorphosed to greenschist facies conditions would contain chlorite, biotite, and perhaps garnet as its key indicator minerals, in contrast to the chlorite + actinolite assemblage that would be found in a rock of basaltic composition. In the amphibolite facies, the indicators in a metamorphosed shale would be garnet, staurolite, or kyanite rather than amphibole. *See* SHALE.

The boundaries separating individual facies from one another are chemical reactions called isograds that transform one key mineral assemblage into another. For example, a typical greenschist is transformed into an amphibolite by the hornblende isograd reaction chlorite + actinolite + plagioclase = hornblende, and an amphibolite can be transformed into a granulite by the reaction hornblende = clinopyroxene + orthopyroxene + plagioclase. In general, these reactions occur over a range of pressure–temperature conditions that depend upon the specific compositions of the minerals involved. In most cases, facies are separated by several reactions that occur within a restricted interval of pressure–temperature space. Thus the boundaries between facies are diffuse and are represented by tinted areas rather than individual lines in Fig. 4. *See* HORNBLENDE; PHASE EQUILIBRIUM.

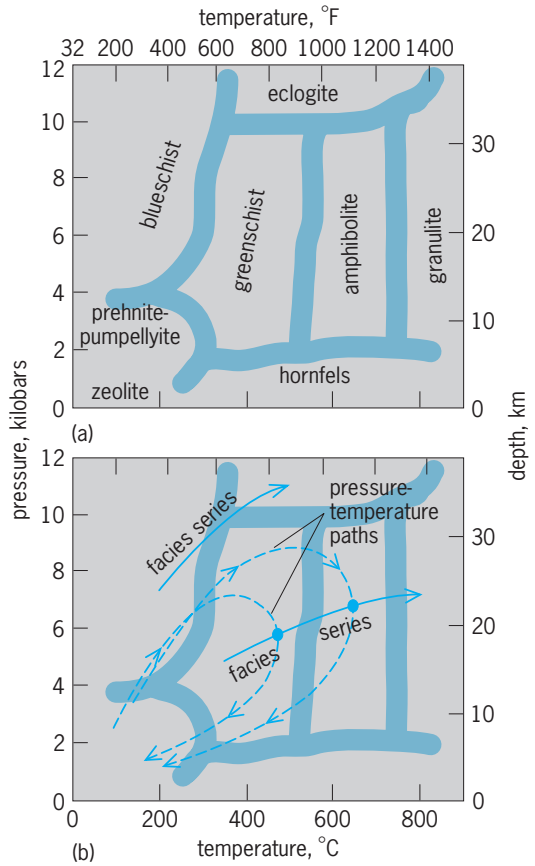

Fig. 4. Pressure–temperature fields. (*a*) Diagram showing the relationship between the different metamorphic facies and the physical conditions of metamorphism. Boundaries (tinted areas) between facies are regions where isograd reactions occur. (*b*) Diagram showing the relationships between metamorphic facies, individual rock pressure–temperature paths, and two typical facies series. Note that the facies series connects maximum temperature points on pressure–temperature paths followed by different rocks. 1 kilobar = 10^8 pascals. 1 km = 0.6 mi.

Facies series. Different geological terranes show different combinations of metamorphic facies. For example, one area may show a progression from blueschist to eclogite to greenschist facies over a distance of a few kilometers (1 km = 0.6 mi), whereas another area may be characterized by a sequence from greenschist to amphibolite to granulite facies. Because the specific facies are related to the pressure and temperature conditions of metamorphism, these different facies series indicate differences in the pressure–temperature histories of the two areas.

The blueschist-eclogite-greenschist facies series develops in areas that have experienced high-pressure metamorphism at relatively low temperatures. This type of pressure–temperature history is produced in subduction zones where two lithospheric plates collide with one another. In contrast, the greenschist-amphibolite-granulite series forms in regions that have undergone heating at middle to lower crustal depths. This facies series is exposed in many of the mountain belts of the world (Appalachians, Alps, Himalayas) and appears to result from tectonic thrusting of one continent over another. Lowpressure hornfels or greenschist-amphibolite facies series develop in response to intrusion of hot magma at shallow levels of the Earth's crust; this is a process referred to as contact metamorphism. As illustrated by these examples, metamorphic facies series can be used as an indicator of the plate tectonic history of a geologic region. *See* ECLOGITE; MAGMA; PLATE TECTONICS.

Pressure–temperature paths. Because burial and uplift rates may differ from rates of heat transfer, individual rocks may follow paths through pressure–temperature space that are quite distinct from the facies series recorded by the host terrane. Mineral reaction rates are strongly dependent upon temperature and proceed faster during heating than during cooling; pressure has little effect on these rates. Hence, metamorphic facies and facies series generally reflect equilibration at the thermal maximum of metamorphism, regardless of the complexity of the actual pressure–temperature path followed by the rock in response to tectonic movements. As shown in Fig. 4*b*, individual pressure–temperature paths in many cases intersect the facies series path at a high angle. Disequilibrium mineral assemblages and chemical zoning in some minerals can sometimes be used to reconstruct portions of these pressure–temperature paths and hence gain information on how the depth of burial and the temperature of an individual rock varied through time. This information can be used to refine tectonic interpretations of an area based on the metamorphic facies series. *See* METAMORPHIC ROCKS; METAMORPHISM; PETROLOGY.

Jane Selverstone

Bibliography. A. D. Miall, *Principles of Sedimentary Basin Analysis*, 3d ed., Springer-Verlag, 1999; A. Miyashiro, *Metamorphic Petrology*, 1994; G. Nichols, *Sedimentology and Stratigraphy*, Blackwell, 1999; F. J. Turner, *Meatamorphic Petrology: Mineralogical, Field, and Tectonic Aspects*, 2d ed., 1981; R. G. Walker and N. P. James (eds.), *Facies Models: Response to Sea Level Change*, Geological Association of Canada, 1992.

Facilities planning and design

The design and planning of the physical environment of an activity to best support the execution of this activity. The activity may be one of many different types, such as manufacturing (plant layout), health care (hospital layout), logistics (warehouse design), transportation (airport layout), and services (bank layout, stadium layout).

Facilities design is a typical engineering design project and as such exhibits the main characteristics of any (large) long-term engineering design project. It is a complex problem that is ill defined. Many of the objectives and constraints are implicitly assumed or can only be stated qualitatively. Different shareholders in the design and the use of the facility have a variety of definitions and objectives of the project. Hence, an analytical problem statement or a mathematical solution method, such as linear programming, to find the "best" design does not exist. Many variant and competent designs may be generated, and the selection of the final design will involve both quantitative and qualitative measures. *See* ENGINEERING DESIGN.

Facilities design is an interdisciplinary team effort; no single engineer or engineering discipline can complete the full design or has all the required expertise. Industrial engineers typically are responsible for the logical design of the facility and focus on space allocation and functional flow in the facility, while civil engineers focus on the physical construction. Communication between all the user groups and the design team is essential for the final acceptance of the facility. *See* CIVIL ENGINEERING; INDUSTRIAL ENGINEERING.

Most facilities exist for extended periods of time, which may span many decades. Activities and the requirements for their supporting facilities will change constantly over time, and the function and the detailed layout of the facility will be redesigned many times during its life cycle. A primary characteristic of a good facility design is that it is flexible and easily adaptable. Integration in a spatial and temporal master plan is essential for the future efficient use of the facility.

Projects. Facilities design projects must identify the following two basic elements during the early phases: activity centers and the affinities between these centers in the facility.

Activity center. An activity center, commonly called a department, is a compact area that has a homogeneous material flow with the rest of the facility so that it can be treated as a single unit (with respect to the material flows in the facility). Its area should not be too small, since it is not important enough to be considered separately. A rule of thumb is to have an area larger than 3% of the total facility area. This means that the number of departments should be limited to approximately 35. Its area should also not be too large, since the material flow might no longer be homogeneous over the area of this department.

Affinity. The pairwise relationship between two departments expresses the affinity between these two departments, based on material flow and environmental considerations. Relationships can be numerical (quantitative) if accurate information is available or symbolic (qualitative). The closer one would like these departments to be, the more positive their relationship. A negative relationship means that it is desirable to keep the two departments separate, such

TABLE 1. Qualitative relationships representation

Closeness	Letter	Lines	Color
Absolute necessary	A	4	Red
Especially important	E	3	Orange
Important	I	2	Blue
Ordinary closeness	O	1	Green
Unimportant	U	No	None
Undesirable	X	Wave	Brown

TABLE 2. Qualitative relationship chart

	A	B	C	D	E	F	G
A		E	O	I	O	U	U
B			U	E	I	I	U
C				U	U	O	U
D					I	U	U
E						A	I
F							E
G							

as for noise or vibration pollution. The table of relationships is called a relationship chart or relationship matrix. Possible representations of qualitative relationships are given in **Table 1**. An example of a qualitative relationship chart, as originally proposed by R. Muther, is shown in **Table 2**.

Design sequence. The following sequence of steps can be used to identify departments and their affinities in the project.

1. Identify the major materials-handling flows in the project. This material flow can be of many different types depending on the facilities design application. In a hospital, patients, medical personnel, drugs, and medical supplies are all important material flows. In an airport, passengers (arriving and departing), crew, planes, and luggage are all important material flows. For a bank customer, personnel, documents, and money are examples of important material flows. In manufacturing, parts, personnel, and tools can be examples of the major materials-handling flows. *See* MATERIALS HANDLING.

2. For each of the materials-handling flows identified in step 1, identify the major processing steps. Again, this includes a wide variety of tasks depending on the application. Examples are a customer dining in a restaurant, luggage dispersal to travelers in an airport, drug storage in the pharmacy of a hospital, order picking in a warehouse, and so on. The processing steps are executed in activity centers or departments. At this particular stage, a functional description of the steps is required, such as store, move, or transform.

3. Construct a graph with a column for each of the materials-handling flows and a row for each of the activity centers. For each of the materials-handling flow, a line is drawn between the activity centers to indicate the order in which the centers are visited. An example of such a graph is a multiproduct process chart. *See* METHODS ENGINEERING.

4. Given the materials-handling flows and the activity centers, determine the affinities between the activity centers in a common unit. If only materials-handling flows are present, then this is a relatively easy step. The materials-handling load, such as a pallet of forklift truck trip, might be the common materials-handling unit. If other affinities are also present, then their size must be carefully determined to be consistent with the materials-handling flow units. Examples of positive affinities are the desirability of having windows in offices and cafeterias, which implies a location on the perimeter of the layout. Examples of negative affinities are welding (which produces sparks) and painting (which produces combustible vapors) departments, vibration pollution between a stamping and a measuring department, and noise pollution between a heavy machining and an office department. Once a common handling unit has been defined, the flows in the above graph can be quantified. All materials-handling flows and other affinities are then summarized in a two-dimensional relationship matrix, also called a from–to matrix, which has as elements the sum of the affinities between two departments for all major materials-handling flows.

5. Given the estimated materials-handling flows that need to be processed in each of the activity centers, compute the number of individual servers required in each activity center. Servers are the generic term for the people or machines that execute the processing step in the activity center. In a manufacturing context, a server is typically a machine and all its surrounding support equipment. In a hospital emergency room, a server may be an examination area surrounded by a privacy curtain. In a grocery store checkout area, a server may be the cashier and the station. Deterministic capacity calculation, queuing theory, and digital simulation are three methods used to compute the number of servers, which are ranked by increasing level of accuracy and required effort. Finally, based on the number of servers and the required area for each server and possible space for waiting area for the server, compute the required area for each activity center.

Space-relationship diagrams. Based on the data for the activity centers, such as required area and other constraints, and on the affinities between the centers, a logical diagram of the centers and the material flows can now be constructed. R. Muther defined such a diagram and called it a space-relationship diagram. The objective of the logical design of the facility is to place activity centers with large positive affinities close to each other and to separate centers with significant negative affinity, subject to space and site constraints. Three major types of diagrams can be identified. For historical reasons their names are based on their implementation in manufacturing activities, but the prototypes apply to all types of facilities.

Product layout. A product layout is recommended when most of the materials or products follow the same material flow path. This corresponds to a flow shop in manufacturing. Most warehouses also correspond to a flow shop, since the product inventory

has a limited number of paths from receiving to shipping. The objective is to have a simple configuration for the major material flow path, such as straight line or L- and U-shaped. The activity centers are located according to the sequence of processing steps required by the product.

Process layout. A process layout is recommended when products have different paths through the facility. This corresponds to a job shop in manufacturing. The typical example is a repair shop with different types of machining centers. The objective is to minimize the total amount of material movement. Because of the complexity of the combined flow paths, optimization models can provide some support in the early design phases. Since accurately predicting the future material flows over the lifetime of the facility is nearly impossible, the facility should be designed to accommodate reconfiguration.

Project layout. A project layout is recommended when materials are transported to a site to be integrated into a final product. A typical example is ship construction.

Conceptual layout. In the next phase, the space-relationship diagram is converted in a conceptual layout, where departments are shown without internal details and without material flow corridors. This conversion is most often executed by the design engineer with drawing computer-aided design (CAD) tools but without formal problem definition or algorithmic support. In recent years, significant progress has been made by the academic community in formalizing and automating this conversion. But these advances have not been transferred to the practice of facilities design. An example of mapping from the functional flow diagram to the conceptual layout is shown in the **illustration**. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

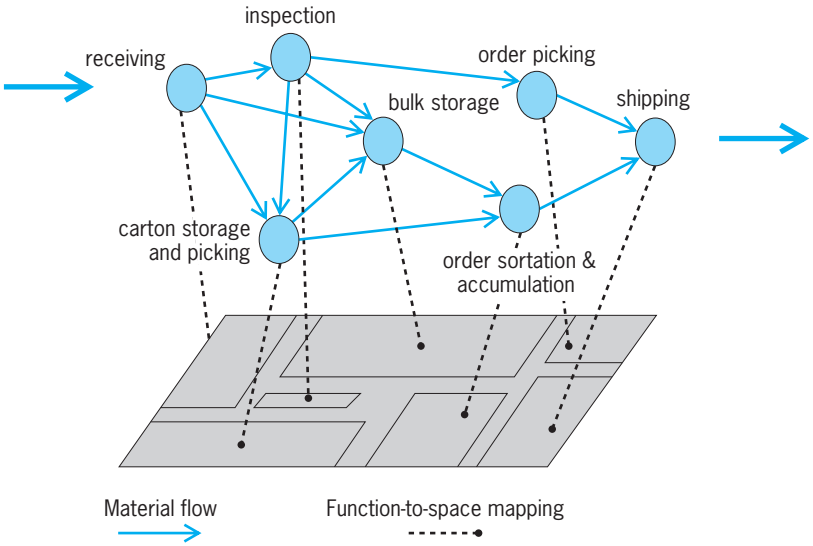

Mapping from the functional flow diagram to the conceptual layout. (*After L. F. McGinnis et al., 2000*)

Detailed layout. The conceptual layout is then converted into a final detailed layout. Space is added for materials-handling corridors and for support functions, such as restrooms, utility closets, and so on. The layout elements corresponding to the requirements of federal regulations such as the Occupational Safety and Health Administration (OSHA), Environmental Protection Administration (EPA), Americans with Disabilities Act (ADA), and of local fire and safety regulations are integrated into the design. If structural elements such as support pillars are present, the design is adapted to incorporate these constraints.

Evaluation. During the design process, numerous engineering design decisions are made that may result in many alternative layouts. The comparative evaluation of facility designs is based on a mixture of engineering economy and qualitative factors. All quantifiable costs for a particular design over the expected life span of the facility are computed and plotted along a time line. This includes one-time construction and material transportation equipment purchasing costs, as well as recurring operating costs. Typical examples of recurring costs are those for energy and material transportation. Costs occurring in the future are reduced to their net present value (NPV) based on a monetary discount factor. Since a facility supports a particular activity, the net present value of the initial investments and recurring costs typically are negative. The facility with the least negative NPV, subject to budget constraints, is preferred on pure economic grounds. However, a number of qualitative evaluation criteria are typically present. A weight is assigned to the economic evaluations and the qualitative evaluations and the facility design with the highest weighted score is finally selected. The weights are based on a polling of the different stakeholders in the facility and are not uniquely defined or constant over time. To provide support to the decision makers, sensitivity analysis on the weight has to be performed.

Simulation tools. It is clear that the above facilities design method incorporates many ill-defined objectives and constraints. Consequently, there has been a very limited development of computer-based facilities design tools beyond drawing tools and facilities management tools. The most often used software during the design of facilities is digital simulation coupled with animation. Simulation is capable of very detailed analysis of a proposed facility design and of the evaluation of some minor design variations, but it cannot generate significantly different designs. As a result, the design of facilities remains largely a person-based activity supported by computer-based drawing and simulation tools. *See* COMPUTER-AIDED ENGINEERING; SIMULATION.

Marc Goetschalckx

Bibliography. R. L. Francis, L. F. McGinnis, and J. A. White, *Facility Layout and Location: An Analytical Approach*, 2d ed., Prentice Hall, 1992; S. Heragu, *Facilities Design*, PWS Publishing, 1996; J.-G. Kim and M. Goetschalckx, An integrated approach for the concurrent determination of the block layout and i/o point locations based on the contour distance, *Int. J. Prod. Res.*, 43(10):2027–2047, 2003; S. Konz, *Facilities Design*, Wiley, 1985; A. Kusiak

and S. S. Heragu, The facility layout problem, *Eur. J. Oper. Res.*, 29:229–251, 1987; L. F. McGinnis et al., Rethinking warehouse design research, *Proceedings of the Material Handling Research Colloquium*, York, June 10–14, York, PA (published in R. Graves et al., eds., *Progress in Material Handling Research: 2000*, pp. 206–216, 2000; R. Meller and K.-Y. Gau, The facility layout problem: recent and emerging trends and perspectives, *J. Manuf. Sys.*, 15(5):351–366, 1996; H. Sherali, B. M. P. Fraticelli, and R. D. Meller, Enhanced model formulations for optimal facility layout, *Oper. Res.*, 51(4):629–644, 2003; R. Muther, *Systematic Layout Planning*, 2d ed., Cahners Books, 1973; D. R. Sule, *Manufacturing Facilities: Location, Planning and Design*, 2d ed., PWS Publishing, 1988; J. A. Tompkins et al., *Facilities Planning*, 3d ed., Wiley, 2003.

Facsimile

The process by which a document is scanned and converted into electrical signals which are transmitted over a communications channel and recorded on a printed page or displayed on a computer screen. The scanner may be compared with a camcorder, and the recorder is similar to an office copier or a computer printer. As an alternative to scanning, a document stored in computer memory can be transmitted. As an alternative to recording, a text facsimile (fax) image can be captured in computer memory and converted into computer-processable text by optical character recognition (OCR) software. Telephone lines or satellites provide the communication channel. More than 99% of facsimile units are International Telegraph and Telephone Consultative Committee Group 3, used for sending business documents. The remainder have specialized designs tailored for a number of important applications that cannot be fulfilled by Group 3.

Group 3 facsimile. Group 3 digital facsimile handles far more messages than any other communications system except voice. Most facsimile units communicate over the Public Switched Telephone Network, alternatively called the General Switched Telephone Network. A built-in high-speed digital modem automatically selects the highest modem speed (28,800–2400 bits/s) common to both facsimile units. If the telephone-line quality is not good enough for this transmission speed, a lower speed is negotiated during initialization. A very small percentage of the fax units send or receive via a cellular phone connection. Group 3 standards support alternative use of the Integrated Services Digital Network, including the 64-kilobit/s channel. *See* INTEGRATED SERVICES DIGITAL NETWORK (ISDN); MODEM; TELEPHONE SERVICE.

Unattended-reception capability allows transmission with automated recording. Some fax units can place calls and send documents without an operator. The cost of a facsimile call usually is far less than a voice call since facsimile sends information much faster, at a rate of 3–10 or more pages per minute.

Scanning. An image of the original page is formed by a lens in a way similar to that of an ordinary camera (see **illus.**). A charge-coupled device linear array of small photodiodes is substituted in the facsimile scanner for the camera film. The portion of the image falling on the linear diode array is a thin line, 0.005 in. (0.13 mm) high, across the top of the page being transmitted. Typically, 1728 diodes are used to view this line for a page $8^1/_2$ in. (216 mm) wide. The photodiode corresponding to the left edge of the page is first checked to determine whether the very small portion of the image it detects is white (the paper background) or black (a mark). The spot detected by a single photodiode is called a picture element (a pel for short if it is recorded as either black or white, or a pixel if a gray scale is used). Each of the 1728 diodes is checked in sequence, to read across the page. Then the original page is stepped the height of this thin line, and the next line is read. The step-and-read process repeats until the whole page has been scanned. *See* CAMERA; CHARGE-COUPLED DEVICES; PHOTODIODE; TELEVISION CAMERA.

Another class of flatbed scanner uses a contact image sensor linear array of photodiodes whose width is the same as the scanned width. One version has a linear array of fiber-optics rod lenses between the page being scanned and the sensor array. Light from a fluorescent lamp or a linear light-emitting-diode array illuminates the document beneath the rod lenses. The reflected light picked up by the sensor generates a signal that is proportional to the brightness of the spot being scanned. A second

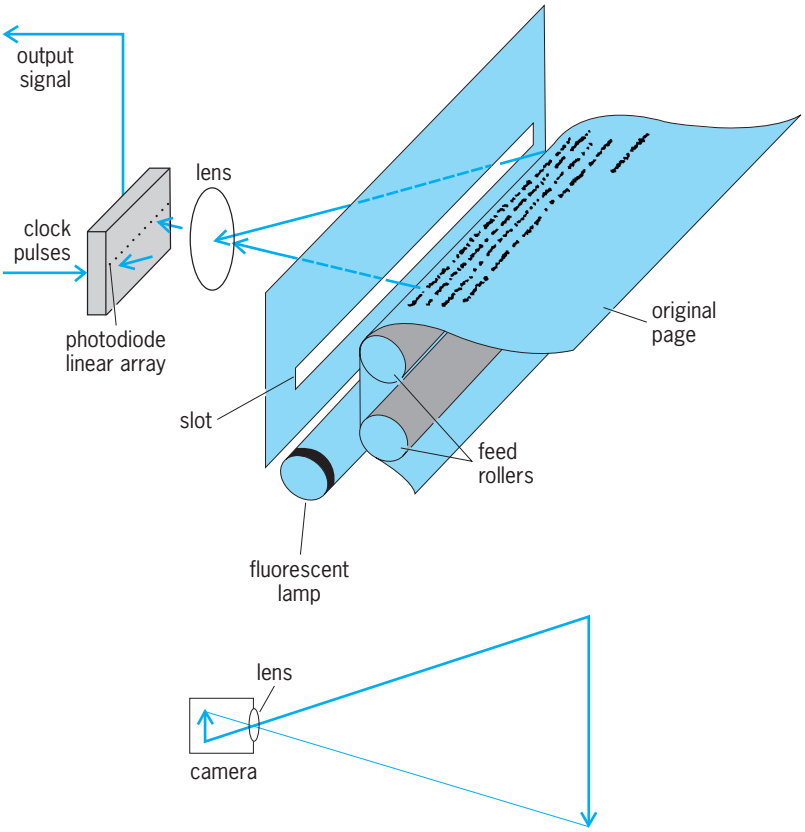

Photodiode facsimile scanner.

version has a hole in the center of each square pixel sensor element. Light from a light-emitting diode passes through this hole to illuminate the area of the document page at this pixel. No lenses or other optical parts are used. *See* FIBER-OPTICS IMAGING; LIGHT CURVES.

In drum-type scanning, the original sheet of paper is mounted on a drum that rotates while the scan head with a photosensor moves sideways the width of one scanning line for each turn of the drum. Drum-type scanners are used mainly for remote publishing facsimiles and for color scanning in graphic arts systems.

Recording. In the recording process, facsimile signals are converted into a copy of the original. Facsimile receivers commonly print pages as they are received, but in an alternative arrangement pages may be stored and viewed on a computer screen. Facsimile capability is possible with computers, including portable varieties.

One of the most common methods is thermal recording. There are 1728 very fine wires positioned in a row across the recording paper. These wires touch the paper and produce very small hot spots as current passes through them. These hot-spot sections of the wires form a straight line across the page at a resolution of 200 dots per inch (8 dots per millimeter). A wire takes only a few milliseconds to go from a cool, nonmarking condition to a hot, marking condition. Each recording wire corresponds to a photodiode in the linear array of the transmitting scanner. Direct-writing, thermally sensitive coated paper is used to produce the image, but a more expensive system of thermal transfer that uses ordinary white paper is offered in some units.

Xerographic recording with a laser may be used where a higher resolution of 400 dots per inch and higher writing speeds are needed. The laser beam, modulated by the picture signal, is swept in a line across the photoreceptor drum (similar to the operation of laser printers used with computers). The drum steps ahead one recording line height of 0.0025 in. (0.064 mm) before writing the next line. The toning, fixing, and paper-feed systems resemble those of computer printers. A linear array of light-emitting diodes or liquid-crystal shutters may be used instead of the laser beam. *See* COMPUTER PERIPHERAL DEVICES; LASER; PHOTOCOPYING PROCESSES.

Ink-jet recording is one of the least expensive methods of plain-paper recording. An ink-jet facsimile recorder sprays ink droplets on plain paper to reproduce black markings scanned at the facsimile transmitter. At resolutions of up to 300 dots per inch, the copy quality compares favorably with that produced on a laser printer. The ability to print either inch-based resolution or metric resolution on the same printer gives the ink-jet printer an advantage.

Advanced Group 3 capabilities. Group 3 standards can support computer functions with fax, adding many optional features without impairing compatibility with Group 3 fax machines that lack them.

Group 3C. This 64 kb/s option for Group 3 facsimile on public digital networks uses full-duplex or half-duplex communication with the circuit-switched mode and error-correction mode. This method is considerably simplified compared with the protocols used by Group 4 facsimile. Sending or receiving documents from another G3-64 fax machine is at 64 kb/s using an Integrated Services Digital Network B channel. Communication with a standard Group 3 machine on the Public Switched Telephone Network is possible via the Integrated Services Digital Network.

Cellular radio. Automobiles equipped with cellular radio enable communication by facsimile as well as by telephone. The brief loss of signal that occurs when switching between cells is heard as a click and the voice signals continue normally, but with facsimile each click causes a streak across the page and may obliterate information. The frequency of cell changing is unpredictable, and facsimile pages may have no streaks or up to 10 streaks per page. Switching may also result in a failure message at the facsimile transmitter, caused by too many received errors, or may cause a disconnect. Cellular networks also are subject to loss of signal, multipath, and fading due to physical obstructions, reflections from buildings, and distance. Use of the error-correction mode is desirable. *See* MOBILE RADIO.

Computer facsimile. Personal computers can emulate Group 3 facsimile for transmission of text and graphic files. A so-called fax board is plugged into a vacant slot in the computer, and a software program converts computer language files into Group 3 facsimile files. A standard facsimile modem on the fax board makes the signals and the communication protocol identical with Group 3. Documents sent to or received from Group 3 facsimile or other personal computers can be viewed and edited on the display screen, or a hard copy can be made on the computer's printer. By adding a scanner, documents can also be sent (for complete emulation of Group 3 facsimile). *See* MICROCOMPUTER.

Computer file transfer. Mainframes, file servers on local area networks, workstations, business computers, personal computers, and even palm-size personal digital assistants often have the ability to send and receive Group 3 facsimile documents. Several options additionally allow automatic transmission of data by the Group 3 facsimile high-speed modem and worldwide standardized facsimile protocols. High-level data-link control (HDLC) frames are sent as the logical equivalent of an error-corrected facsimile message, with the protocols and transmission speeds being automatically established. This file transfer method sends any data file with or without additional information concerning the file.

Other facsimile systems. A number of facsimile systems besides Group 3 are in operation.

Group 4 facsimile. This system sends error-free copies over public data networks at 200 and 300 lines per inch (7.9 and 11.8 lines per millimeter), with 240 and 400 (9.4 and 15.7) optional. Group 4 operates over high-speed error-free digital networks using the Open Systems Interface (OSI) model. Communication is also possible with Group 3 units by a

so-called dual-mode Group 4 Class 1 fax. Dual-mode Group 4 facsimile may have standard Group 3 with a fax modem, Group 3c (64-kb/s capability), or both. There are slight page size distortions caused when interworking between Group 3 metric-based resolutions and similar Group 4 inch-based resolutions. Group 3 now provides most of the same features as Group 4b, using simplified protocols.

Newsphoto facsimile. Most pictures printed in newspapers have been sent by facsimile. News services have nationwide networks for distribution of newsphotos. Thousands of newspapers receive the latest news pictures at the same time by a broadcast operation, coming from almost anywhere in the world. These systems reproduce the gray scale very well, with each pixel representing one of 64 shades. Color separations are also sent as a set of four gray-scale originals representing the primary colors plus black. An electronic picture desk may store the pictures in memory as they are received.

Remote publishing. Newspapers and magazines are published from full-page facsimile masters received in plants thousands of miles away from the place where the pages are composed. The facsimile recording is made on a film negative or directly on a printing plate. High-resolution, high-speed facsimile units send full-size newspaper pages over wideband satellite or land lines. Resolutions of 800–1000 lines/inch may be used. Major newspapers transmit to many printing plants via direct satellite broadcast to receiving stations located at each plant. Color pages are sent by making three or four transmissions of the original color page by using color separation filters.

Weather facsimile. Weather networks cover the United States, constantly making copies of weather charts. Cloud-cover photographs and multispectral band images are recorded by facsimile from signals sent by orbiting and geostationary satellites. *See* METEOROLOGICAL SATELLITES. Kenneth R. McConnell

Bibliography. K. R. McConnell, *Facsimile Technology and Applications Handbook*, 2d ed., 1992.

Factor analysis

A method of quantitative multivariate analysis with the goal of representing the interrelationships among a set of continuously measured variables (usually represented by their intercorrelations) by a number of underlying, linearly independent reference variables called factors. Although the term factor analysis has come to represent a family of analysis methods, the two most commonly used approaches are the full component model, in which the entire variance of the variables (represented by unities inserted in the principal diagonal of the correlation matrix) is analyzed, and the common factor model, in which the proportion of the variance that is accounted for by the common factors (represented by communality estimates inserted in the principal diagonal) is analyzed.

The method was developed in England around the turn of the century and was first applied to the study of the structure of intellectual abilities. Since then it has been used in many disciplines, from agriculture to zoology, in which the underlying structure of multiple variables and their representation in that structure are of interest. Another application of factor analysis is to represent parsimoniously the variables in the set on which the observations are made by a smaller number of underlying reference variables or factors.

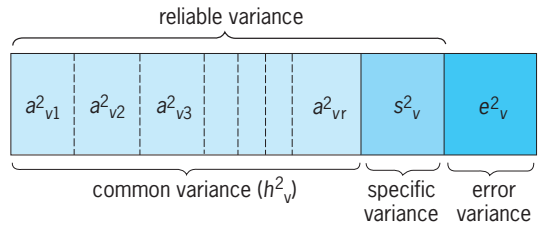

Fig. 1. Diagram of the total variance of a variable in standard score units subdivided into types of variances.

Equations and matrices. The defining equation for the full component model is shown in Eq. (1), where

$$z_{iv} = a_{v1}F_{1i} + a_{v2}F_{2i} + \cdots + a_{vr}F_{ri} \qquad (1)$$

z_{iv} represents the standard score for case i on variable v, a_{vr} is the weight for variable v on factor r, and F_{ri} is the amount (factor score), in standard scores, on factor r possessed by case i.

The defining equation of the common factor model is shown in Eq. (2), where the first elements

$$z_{vi} = a_{v1}F_{1i} + \cdots + a_{vr}F_{ri} + a_{vu}U_{iv} \qquad (2)$$

are the same as in the full component model, a_{vu} is the weight given variable v's unique factor, and U_{iv} is case i's unique factor score for variable v. Each variable's unique factor includes random errors, influences affecting that variable only, and all other sources of error and bias that do not contribute to the correlations of that variable with other variables in the set, and therefore do not enter into determining common factors. Usually, only the common factors are evaluated from the data.

The total variance of each variable, σ^2_v, may be separated into its constituent variances, as represented by Eq. (3).

$$\sigma_v^2 = \sigma_{v1}^2 + \sigma_{v2}^2 + \cdots + \sigma_{vr}^2 + \sigma_{vs}^2 + \sigma_{ve}^2 \qquad (3)$$

Dividing both sides of the equation by σ^2_v to convert to standard scores yields Eq. (4) in terms of

$$1.0 = a_{v1}^2 + a_{v2}^2 + \cdots + a_{vr}^2 + s_v^2 + e_v^2 \qquad (4)$$

proportions of variance. This set of proportions of variances may also be represented diagrammatically (**Fig. 1**).

In applying the common factor model, usually only the common variances are evaluated, and their sum for a variable is represented by h^2_v and is termed the communality of the variable. The full component

model accounts for the components making up the total variance. In practice, however, a "truncated" analysis is usually performed, so that only the common variances (or factors) are extracted from the correlation matrix, even though the components model is used.

The correlation r_{jk} between two variables can be accounted for in terms of their common factor loadings. For orthogonal factors (axes) this can be represented by Eq. (5).

$$r_{jk} = a_{j1}a_{k1} + a_{k1} + \cdots + a_{jr}a_{kr} \tag{5}$$

Thus the correlation between two variables can be accounted for in terms of their common factors. The corresponding equation in matrix notation for a set of variables is $R = FF'$.

The correlation between two variables may also be represented geometrically by the scalar product of the length of two vectors and the cosine of the angle (ϕ) between them. This relationship is represented by Eq. (6). If each variable is indicated by a

$$r_{jk} = V_j V_k \cos \phi_{jk} \tag{6}$$

unit length vector (representing its total variance), the correlation is equal to the cosine of the angle between them. Representations of several r's in terms of unit length vectors and the angle between them are shown in **Fig. 2**. Since the cosine of an angle of 90° is 0.0, an r of zero may be represented by two unit length vectors separated by an angle of 90°, and so on. Positive correlations have angles between 0 and 90°, negative correlations between 90 and 180°.

Figure 3 depicts an attempt to represent the following correlation matrix geometrically in two dimensions:

Variable	1	2	3
1	1.000	.000	.707
2	.000	1.000	−.505
3	.707	−.500	1.000

Since the sum of 45° plus 90° is not equal to 120°, the representation is erroneous; it requires three dimensions to represent the vectors and the angles between them correctly. The minimum number of dimensions required to represent a matrix of intercorrelations may be taken to indicate the rank of the matrix. One of the primary uses of factor analysis is to determine the number of independent reference variables or dimensions needed to account for the intercorrelations of a set of continuous variables. With empirical data that contain measurement and other sources of error, the rank can only be approximated.

Figure 4 is a geometric representation of the following matrix of intercorrelations among four variables, which can be represented correctly in two dimensions:

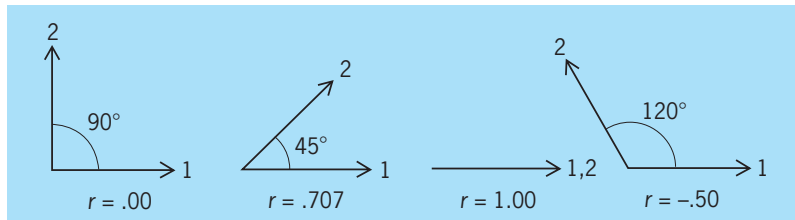

Fig. 2. Vectorial representation of correlation coefficients.

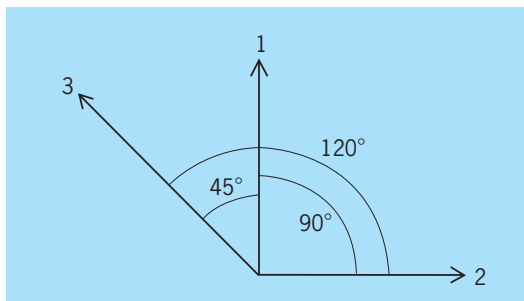

Fig. 3. Representation demonstrating that the three-variable correlation matrix cannot be represented correctly in two dimensions.

Variable	1	2	3	4
1	(1.00)	.80	.96	.60
2	(.80)	(1.00)	.60	.00
3	(.96)	.60	(1.00)	.80
4	(.60)	.00	(.80)	(1.00)

A pair of orthogonal cartesian axes has been inserted into the vector space, with axis I coinciding with vector 2, and axis II coinciding with variable 4.

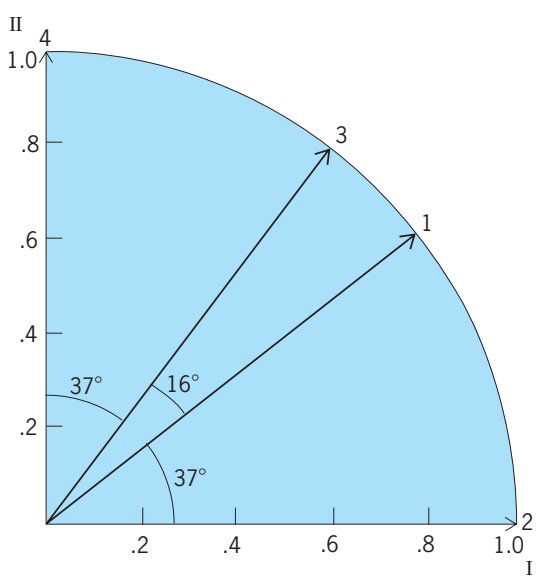

Fig. 4. Vectorial representation of the intercorrelations among four variables in a two-dimensional space.

If it is preferred to have the reference axes (which actually extend from +1.0 to −1.0) in another position relative to the variable vectors (on either orthogonal or oblique axes), they may be transformed to that position by "rotating" them about the origin. When an analysis has been performed, the resulting values are usually written in an abbreviated factor matrix form, with the entries in the body of the table (that is, the factor coefficients) being the coordinates of the variables on the axes, and h^2 (the sum of the squared factor coefficients on orthogonal axes) representing the communality of the variable:

	Factor		
Variable	I	II	h^2
1	.8	.6	1.0
2	1.0	.0	1.0
3	.6	.8	1.0
4	.0	1.0	1.0

For fallible data the values of the communalities would be less than 1.0. The nature of the underlying characteristic represented by a factor is frequently inferred from the common element or characteristic of the variables having high loadings on the factor.

Procedure. The most commonly performed steps in factor analysis are as follows:

1. *Factoring the correlation matrix*: The most commonly used procedure is to obtain the eigenvectors and eigenvalues of the matrix by least-squares procedures with unities inserted in the principal diagonal (principal components), or with communality estimates in the principal diagonal (principal factors). In the latter case the factoring process is frequently iterated one or more times (using the communalities computed from the previous factoring) to stabilize the communality estimates. A number of other factoring procedures have been devised, including diagonal, centroid, multiple-group, image, alpha, nonlinear, nonmetric, and maximum-likelihood solutions.

2. *Determining the number of factors*: Since, with falliable data, there is not an exact number of factors, a large number of approximate criteria have been developed. Some commonly used ones with principal-components analysis are retaining those factors with eigenvalues ≥ 1.0, and Cattell's screen test.

3. *Rotation and interpretation of factors*: The position of the axes determined by the mathematical factoring methods are usually not optimal for inferring the nature of the factors based on the content of the variables and knowledge of the field of investigation. Axes may be rotated about the origin, yielding orthogonal or oblique axes solutions according to several models that are more appropriate for interpretation.

Historically, one of the first models was C. Spearman's two-factor theory. He observed that measures of mental abilities intercorrelated positively, and developed a theory of a general intellectual ability, which he called the two-factor theory. All the mental ability measures loaded on a common factor (labeled "g," for general intellectual ability), and each test also had its own specific factor. K. J. Holzinger modified the model to the bifactor theory. In addition to the general or g-factor, groups of the mental ability tests have factors in common; these group factors represented cognitive abilities such as verbal, numerical, and reasoning.

A more general model is L. L. Thurstone's multiple-factor theory, which is based on the principle of "simple structure." For the case where each variable is factorially less complex than the set of variables taken together, the following are some considerations for evaluating a simple structure: Each variable should have at least one factor loading close to zero. Each factor should have at least as many near-zero loadings as there are factors. For every pair of factors there should be several variables with near-zero loadings on one factor, but not on the other. For analyses with four or more factors, a large proportion of the variables should have negligible loadings on any pair of factors. Only a small number of variables should have appreciably large loadings on any pair of factors.

A secondary consideration in guiding rotations of factors based on sets of variables with positive intercorrelations is "positive manifold"; that is, after rotation of axes all factor loadings are zero or positive. **Figure 5** shows a possible simple structure for 10 variables with a positive manifold configuration in three factors on the surface of a sphere, and **Fig. 6** shows one for 12 variables, with bipolar factors having negative as well as positive loadings. Several computer programs have been devised for axis rotation to simple structure. The most widely used one for orthogonal solutions is H. F. Kaiser's normalized-varimax algorithm, and for oblique rotations the oblimin, biquartimin, promax, and maxplane methods. In an oblique solution, the degree of obliqueness of the axes, as well as the matrix of factor coefficients, must be taken into consideration. The matrix representing the obliquities of the axes

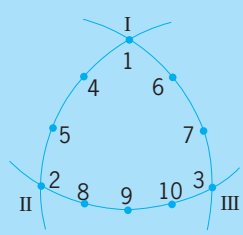

	factor pattern		
variable	I	II	III
1	+		
2		+	
3			+
4	+	+	
5	+	+	
6	+		+
7	+		+
8		+	+
9		+	+
10		+	+

Fig. 5. A simple structure configuration for 10 variables in three dimensions, shown on the surface of a sphere, with reference axes I, II, and III at the intersections, and the corresponding factor-pattern matrix.

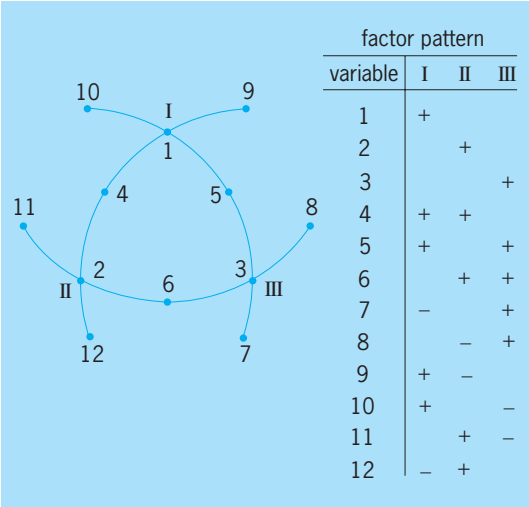

factor pattern

variable	I	II	III
1	+		
2		+	
3			+
4	+	+	
5	+		+
6		+	+
7	−		+
8		−	+
9	+	−	
10	+		−
11		+	−
12	−	+	

Fig. 6. A simple structure for 12 variables in three dimensions with bipolar factors, and the corresponding factor-pattern matrix.

in an oblique solution may be factored to obtain a second-order factor solution, and if several oblique factors are obtained, higher-order analyses may be performed.

"Procrustes" is another approach to rotation of axes. It involves setting up a hypothesized factor pattern for a set of variables based on theory or previous results, and rotating the factor matrix as closely to the hypothesized matrix as possible while still reproducing the correlation matrix as accurately as can be done from the unrotated factors. This procedure is also used to make factor solutions based on common sets of variables as congruent as possible before relating factors across studies.

Factor scores. It will be recalled that the defining equations given at the beginning of this article included F_{ri}'s, or factor scores in standard score units. Since the standard scores on the variables and the factor coefficients are available after a factor analysis, the weights for factor scores can be solved for, or estimated, by linear regression. Theoretically the factor scores can be solved for exactly in the full components model, and only estimated in the common factor model, since estimates of the unique factor scores and weights are usually not available in that model. A more direct, less formal approach to estimating factor scores is to use one or more of the variables that have high coefficients on a factor, but with only negligible coefficients on other factors, to represent the factor.

Statistical inference. The usual statistical approach to factor analysis is descriptive, although some tests of significance and other inferential procedures have been developed. Factor analytical studies may be classified as exploratory or confirmatory. In connection with the latter, analysis of covariance structures is a procedure that is used to test the fit of the data to a model of specified factor coefficients. The number of cases used to obtain the sample data is usually several times the number of variables in order to minimize problems of statistical significance and help ensure that the correlation matrix will be gramian.

Designs. While the intercorrelation of variables over cases on one occasion (R-technique) is the most common design, others such as Q-technique (the correlation of cases over variables on one occasion), and P-technique (the correlation of variables over occasions for one case) have been used. Also, three-mode factor analysis permits the simultaneous analysis of variables, cases, and occasions. Factor analysis has been used in many fields of science to infer the underlying structure of overlapping measured variables, and especially in psychology to analyze measures of individual differences and their underlying structure. *See* STATISTICS. Benjamin Fruchter

Bibliography. D. Child, *The Essentials of Factor Analysis*, 2d ed., 1991; A. L. Comrey and H. B. Lee, *A First Course in Factor Analysis*, 2d ed., 1992; R. P. McDonald, *Factor Analysis and Related Methods*, 1985; R. A. Reyment and K. G. Joreskog, *Applied Factor Analysis in the Natural Sciences*, 2d ed., 1993.

Fagales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the superorder Rosidae of Eudicotyledon. The order consists of 8 families (Betulaceae, Casuarinaceae, Fagaceae, Juglandaceae, Myricaceae, Nothofagaceae, Rhoipteleaceae, Ticodendraceae) and approximately 30 genera and nearly 1000 species. The Fagales are either simple or compound-leaved, woody plants. Flowers are mostly

Butternut or white walnut (*Juglans cinerea*), a North American species showing compound leaves and slender drooping male catkins. (*Ken Sytsma, University of Wisconsin*)

unisexual and much reduced for wind pollination. The female flowers produce one- or two-seeded nut fruits (for example, acorn, chestnut, walnut); the male flowers are grouped into pendant catkins (see **illus.**). Birch (*Betula*), beech (*Fagus*), walnut (*Juglans*), and oak (*Quercus*) are members of the Fagales. *See* BEECH; BIRCH; MAGNOLIOPHYTA; OAK; PLANT KINGDOM; ROSIDAE. K. J. Sytsma

Falconiformes

A worldwide order of diurnal predacious birds without obvious affinities to other orders of birds. They are not closely related to the owls; any similarities between these two orders are the result of convergence. *See* AVES.

Classification. The Falconiformes are divided into two suborders. The first, Accipitres, has three families: Pandionidae (osprey; one species; worldwide, including Australia; *see* **illustration**); Accipitridae (hawks, eagles, kites, harriers, and Old World vultures; 217 species; worldwide); and Sagittariidae (secretary bird; one species; Africa). The second suborder, Falcone, contains a single family, Falconidae (falcons, caracaras; 62 species; worldwide). The New World vultures, which had been included in the Falconiformes, are now placed with the Ciconiiformes. *See* CICONIIFORMES.

The inclusion of all these families in one order has been questioned. Some scientists conclude that the falcons are not related to the Accipitres but are allied to the owls. The position of the largely ground-dwelling and hunting secretary bird has been questioned, with some researchers suggesting that it may be a predacious member of the Gruiformes. *See* GRUIFORMES; STRIGIFORMES.

Osprey (*Pandion haliaetus*). (*Photo by Glenn and Martha Vargas, © 2002 California Academy of Sciences*)

Characteristics. The falconiforms (or diurnal raptors) range in size from sparrow-sized falconets (*Microhierax*) of southeastern Asia to the gypaetine vultures with wingspans of 8 ft (2.5 m); most are medium-sized to larger birds. They have strong feet, usually with three toes pointing forward and one pointing backward, and ending with sharp claws. Their beaks are hooked and powerful. Their well-developed wings vary in shape according to the type of flight; they are long and pointed in falcons, short and broad in accipitrine hawks, and long and broad in vultures, eagles, and buzzards. Falconiforms generally hunt from the air, with their wing shape depending on the mode of hunting flight, and they feed on animal prey from insects to sizable vertebrates, including fish, and on carrion. However, the palm-nut vulture (*Gypohierax angolensis*) of Africa is partly vegetarian, eating the fruits of oil palms.

Hawks are flying specialists, although most species can walk well. They are monogamous, with a strong pair bond, and they may mate for life. The young are downy, stay in the nest, and are cared for by both parents. Many northern species are migratory, with spectacular concentrations observed along mountain ridges and the coast during the fall migration.

The osprey (see illustration) is specialized for catching fish with a reversible fourth toe, well-developed claws, and spiny scales along its toes. The one species is found almost worldwide.

The secretary bird of Africa is long-legged and specialized for a terrestrial way of life. It runs rapidly, flies only infrequently, and catches poisonous snakes as well as many other vertebrates.

Falcons are fast-flying birds with long, pointed wings. They hunt from the air, diving and catching their prey in the air or on the ground. They possess an extra toothlike projection on the lateral edges of the upper jaw just behind the terminal hook.

Fossil record. Hawks, falcons, and ospreys are known from Eocene-Oligocene times and are well represented in the fossil record. Secretary birds, now restricted to Africa, are known from the Oligocene and Miocene of France. Most interesting is that the Old World vultures (Gypaetinae) have an extensive fossil record in North America.

Economic significance. Hawks had once been regarded as vermin, but now that their value in controlling rodent pests is understood, they are widely protected. The presence of concentrations of dichlorodiphenyltrichloroethane (DDT) and other pesticides in their food has caused drastic decreases in the number of many species and the near elimination of the osprey, peregrine falcon, and bald eagle over large parts of their range. With decrease in the use of DDT and other pesticides, they are regaining their former numbers. Still, many species remain endangered because of habitat loss, decrease in prey, and continued use of toxic pesticides. More recently, several species of Old World vultures in India have decreased greatly in number with the apparent cause being susceptibility to antibiotics present in the dead

cattle eaten by these scavengers; the drugs have been heavily overused and had built up in the body of the cattle. *See* PESTICIDE. Walter J. Bock

Bibliography. L. Brown and D. Amadon, *Eagles, Hawks and Falcons of the World*, Country Life Books, Feltham, 1968; T. J. Cade, *The Falcons of the World*, Cornell University Press, 1982; J. Ferguson-Lees et al., *Raptors: An Identification Guide to the Birds of Prey of the World*, Houghton-Mifflin, 2001.

Fallopian tube

The upper part of the female oviduct present in humans and other higher vertebrates. The fallopian tube extends from the ovary to the uterus and transports ova from the ovary to the cavity of the uterus. Each tube is about 5 in. (12.5 cm) long; one lies on either side of the uterus and is attached at the upper portion. Each curves outward to end in a hoodlike opening, the infundibulum, with many fingerlike projections, the fimbriae; the cavity of the fallopian tube is continuous with the cavity of the coelom. The ovaries lie below and inside the tubal curve. When ovulation occurs, about the middle of each menstrual cycle in humans, the ovum is picked up by the fimbriae and drawn into the infundibulum and hence the oviduct. The lining of this tube is an epithelium with many hairlike processes, cilia, on its surface. These cilia, and probably muscular action, move the ovum along the tube. The ovum remains viable for about 1–3 days only. If fertilization occurs, the ovum moves into the cavity of the uterus and then implants on its wall. If fertilization fails to occur, the ovum degenerates in the uterus. Occasionally, a fertilized ovum fails to enter the uterus, or may be freed into the abdominal cavity, so that an ectopic pregnancy results if the ovum finds a site for implantation. *See* PREGNANCY DISORDERS; REPRODUCTIVE SYSTEM. Walter Bock

Fan

A fan moves gases by producing a low compression ratio, as in ventilation and pneumatic conveying of materials. The increase in density of the gas in passing through a fan is often negligible; the pressure increase or head is usually measured in inches of water.

Blowers are fans that operate where the resistance to gas flow is predominantly downstream of the fan. Exhausters are fans that operate where the flow resistance is mainly upstream of the fan.

Fans are further classified as centrifugal or axial (**Fig. 1**). The housing provides an inlet and an outlet and confines the flow to the region swept out by the rotating impeller. The impeller imparts velocity to the gas, and this velocity changes to a pressure differential under the influence of the housing and ducts connected to inlet and outlet.

Performance. In selecting a fan for a particular application, requirements of primary interest are the quantity Q of gas to be delivered by the fan and the head H which must be developed to overcome the resistance to flow of the quantity Q in the connected system. These operating conditions establish the fan dimensions of diameter D and rotational speed N. Performance of a fan of diameter D is rigorously described by its characteristic curves (**Fig. 2**). Fans of different types and sizes are conveniently compared by converting their characteristics to dimensionless form. **Figures 3** and **4** show selected fan performance on the percentage basis, where 100% rating is defined as the peak of the efficiency curve.

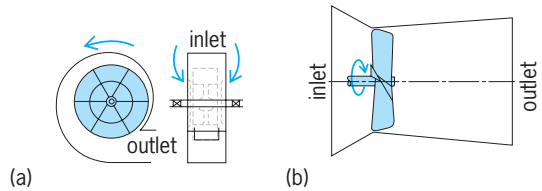

Fig. 1. Fan types. (*a*) Centrifugal. (*b*) Axial.

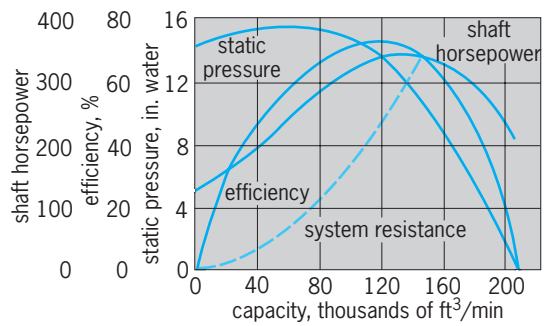

Fig. 2. Static pressure characteristics of backwardly curved blade centrifugal fan at constant speed. 1 horsepower = 746 W. l in. water = 249 Pa. 1 ft^3/min = 0.00047 m^3/s.

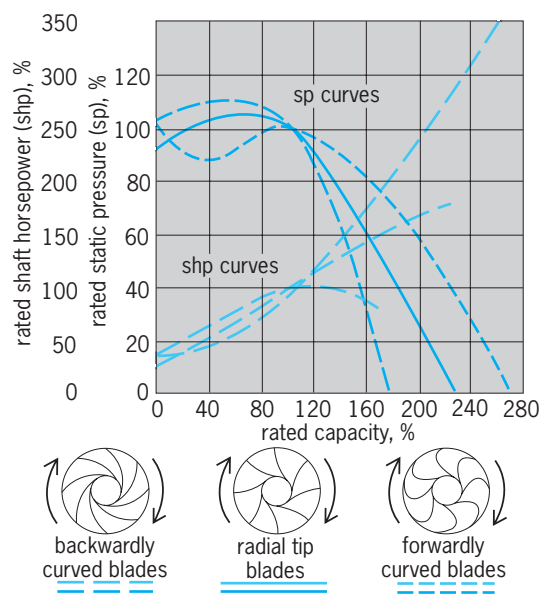

Fig. 3. Percentage characteristics compare performance of three forms of centrifugal fans. 1 horsepower = 746 W.

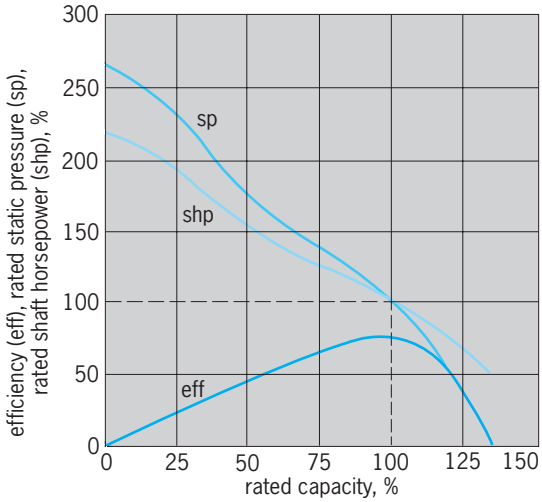

Fig. 4. Percentage characteristics of one axial fan. 1 horsepower = 746 W.

The load placed on the fan must correspond to a condition on its operating characteristic. If the system in which the fan operates presents less resistance than is overcome by the head developed by the fan at the required capacity, more resistance must be introduced by a damper, or the fan speed must be changed, or the excess capacity must be diverted elsewhere.

Performance of a given fan, operating at a given point on its efficiency curve, varies with speed in accordance with the following rules:

Capacity Q proportional to speed N
Head H proportional to speed squared N^2
Horsepower P proportional to speed cubed N^3

For a series of similar fans, operating at a given point on their efficiency curves and at the same speed, performance varie according to the following rules:

Capacity Q proportional to diameter cubed D^3
Head H proportional to diameter squared D^2
Horsepower P proportional to diameter to fifth power D^5

These relations can be combined to express the performance of a family of similar fans in terms of dimensionless coefficients, as in Eqs. (1)–(3).

$$\text{Capacity coefficient } C_Q = \frac{Q}{ND^3} \tag{1}$$

$$\text{Head coefficient } C_H = \frac{gH}{N^2D^2} \tag{2}$$

$$\text{Power coefficient } C_P = \frac{gP}{wN^3D^5} \tag{3}$$

where
Q = capacity, ft^3/s
H = head, ft of fluid
P = shaft horsepower, ft-lb/s
g = gravitional acceleration, ft/s^2
N = revolutions/s
w = weight density, lb/ft^3
D = wheel diameter, ft

Values of these coefficients for a selected group of fans are listed in the **table**.

Comparison of different types of fans is further facilitated by the elimination of the diameter term. Equations for capacity coefficient and head coefficient are solved simultaneously to give specific speed N_S, as in Eq. (4). Specific speed is an inherent performance criterion; it is usually employed as a dimensionally impure coefficient. In the selection of a fan for a given application, a fan is chosen that has the desired value of N_S in its region of peak efficiency. Low specific speed corresponds to a fan of low rotative speed and large impeller diameter.

$$N_S = \frac{NQ^{1/2}}{(gH)^{3/4}} \tag{4}$$

Alternatively, speed can be eliminated, giving

Performance and dimensions of a group of fans

	Centrifugal fans			
	Backwardly curved blades	Steel plate	Forwardly curved blades	Axial fan (pressure blower)
Performance				
Specific speed (N_S)	0.25	0.11	0.21	0.75
Peak efficiency, %	75	65	65	72
Tip speed, fpm × 10^3	18	11	3.5	16
C_H at shutoff	3.8	6.3	11.2	3.2
C_Q at peak efficiency	0.42	0.18	1.2	0.4
C_P at peak efficiency	2.1	1.7	2.1	0.55
Proportions as functions of wheel diameter D				
Inlet area ($\times D^2$)	0.7	0.4	0.8	0.8
Outlet area ($\times D^2$)	0.6	0.3	0.7	
Axial blade length ($\times D$)	0.3	0.4	0.6	
Number of blades	18	8	64	6

specific diameter D_S, as in Eq. (5).

$$D_S = \frac{(gH)^{0.25}D}{Q^{0.5}} \tag{5}$$

Types. The shape of a fan characteristic as a function of capacity depends on the fan type. For example, backwardly curved blade centrifugal fans have a steeply falling head characteristic and a self-limiting horsepower characteristic as shown in Fig. 3.

Fans with forwardly curved blades have rising head and horsepower characteristics. The axial fan has a falling horsepower characteristic so that at shutoff the fan may require more power than at high flow rate (Fig. 4). *See* DUCTED FAN. Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; A. H. Church, *Centrifugal Pumps and Blowers*, 1944, reprint 1972; V. L. Streeter, B. Wylie, and K. W. Bedford, *Fluid Mechanics*, 9th ed., 1998.

Faraday effect

Rotation of the plane of polarization of a beam of linearly polarized light when the light passes through matter in the direction of the lines of force of an applied magnetic field. Discovered by M. Faraday in 1846, the effect is often called magnetic rotation. The magnitude α of the rotation depends on the strength of the magnetic field H, the nature of the transmitting substance, the frequency v of the light, the temperature, and other parameters. In general, $\alpha = VxH$, where x is the length of the light path in the magnetized substance and V the so-called Verdet constant. The constant V is a property of the transmitting substance, its temperature, and the frequency of the light.

The Faraday effect is particularly simple in substances having sharp absorption lines, that is, in gases and in certain crystals, particularly at low temperatures. Here the effect can be fully explained from the fundamental properties of the atoms and molecules involved. In other substances the situation may be more complex, but the same principles apply.

Rotation of the plane of polarization occurs when there is a difference between the indices of refraction n^+ for right-handed polarized light and n^- for left-handed polarized light. Most substances do not show such a difference without a magnetic field, except optically active substances such as crystalline quartz or a sugar solution. It should be noted that the index of refraction in the vicinity of an absorption line changes with the frequency (**Fig. 1***a*). *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; POLARIZED LIGHT.

When the light travels parallel to the lines of force in a magnetic field, an absorption line splits up into two components which are circularly polarized in opposite directions; that is the normal Zeeman effect. This means that, for one line, only right-handed circularly polarized light is absorbed, and for the other one, only left-handed light. The indices of refraction n^- and n^+ bear to their respective absorption frequencies the same relation as indicated in Fig. 1*a*; that is, they are identical in shape but displaced by the frequency difference between the two Zeeman components. It is evident that $n^+ - n^-$ is different from zero (Fig. 1*b*), and the magnetic rotation is proportional to this difference. The magnitude of the rotation is largest in the immediate vicinity of the absorption line and falls off rapidly as the frequency of the light increases or decreases. *See* ZEEMAN EFFECT.

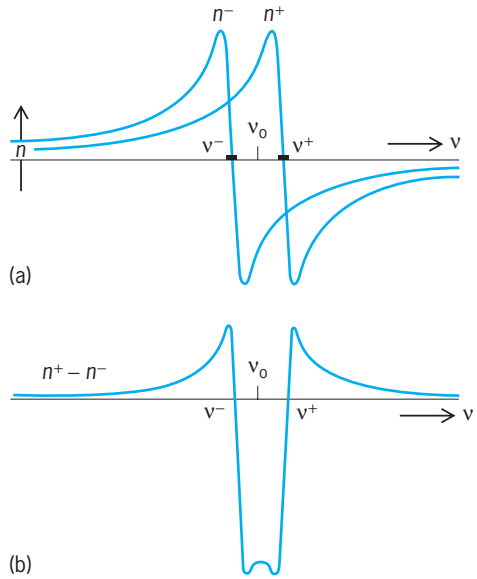

Fig. 1. Curves used in explaining the Faraday effect. (*a*) Index of refraction for left-handed circularly polarized light (n^-) and right-handed light (n^+) in the vicinity of an absorption line split into a doublet (ν^-, ν^+) in a magnetic field. (*b*) Difference between two curves, $n^+ - n^-$. Magnetic rotation is proportional to this difference.

The Faraday effect may be complicated by the fact that a particular absorption line splits into more than two components or that there are several original absorption lines in a particular region of the spectrum.

The case represented in Fig. 1 is independent of the temperature, and the rotation is symmetric on both sides of an absorption line. This case is called, not quite correctly, the diamagnetic Faraday effect. It occurs when the substance is diamagnetic, which means the splitting of the absorption line is due to the splitting of the upper level only (**Fig. 2***a*), and

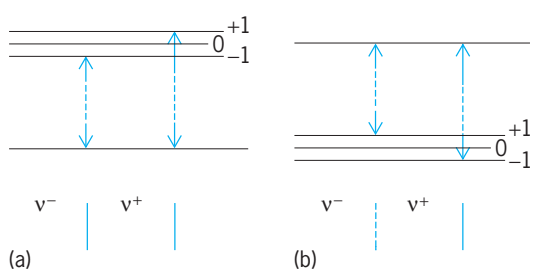

Fig. 2. Two cases of Faraday effect. (*a*) Diamagnetic case, which is temperature-independent, and (*b*) paramagnetic case, which is temperature-dependent.

the lower level of the line is not split. The same situation prevails in general when the intensity of the two Zeeman components is equal. This holds for all substances except paramagnetic salts at very low temperatures (Fig. 2*b*), in which case the v^- component is absent.

In the latter case at high temperatures there are an equal number of ions in the $+1$ and -1 states, and the two Zeeman components have equal intensities, as in the previously discussed situation. When the temperature is lowered, however, the ions concentrate more and more in the lower level (-1), and therefore absorption by ions in the upper level ($+1$) disappears. At very low temperatures only the high-frequency component in the Zeeman pattern is left. In this case the n^- refraction coefficient is not affected by the presence of the absorption line and is a constant. The difference $n^+ - n^-$ will therefore have the shape of the n^+ curve in Fig. 1*a*. Here the rotation is not symmetric with respect to the absorption line. If it is right-handed on one side of the line, it will be left-handed on the other. As the temperature is raised, the other line comes in with increasing strength until the two are nearly equal. In the transition region the Faraday effect depends strongly on the temperature. This is called the paramagnetic Faraday effect.

It is possible to modulate laser light by use of the Faraday effect with a cylinder of flint glass wrapped with an exciting coil. But since this coil must produce a high magnetic field of about 19,000 gauss, this method of modulation has been little used. For a discussion of other phenomena related to the Faraday effect *see* MAGNETOOPTICS. G. H. Dieke; W. W. Watson

Faraday's law of induction

A statement relating an induced electromotive force (emf) to the change in magnetic flux that produces it. For any flux change that takes place in a circuit, Faraday's law states that the magnitude of the emf ϵ induced in the circuit is proportional to the rate of change of flux as in expression (1).

$$\epsilon \propto -\frac{d\Phi}{dt} \tag{1}$$

The time rate of change of flux in this expression may refer to any kind of flux change that takes place. If the change is motion of a conductor through a field, $d\Phi/dt$ refers to the rate of cutting flux. If the change is an increase or decrease in flux linking a coil, $d\Phi/dt$ refers to the rate of such change. It may refer to a motion or to a change that involves no motion.

Faraday's law of induction may be expressed in terms of the flux density over the area of a coil. The flux Φ linking the coil is given by Eq. (2), where α is

$$\Phi = \int B \cos\alpha \, dA \tag{2}$$

the angle between the normal to the plane of the coil and the magnetic induction B. The integral is taken over the area A enclosed by the coil. Then, for a coil of N turns, Eq. (3) holds.

$$\epsilon = -N\frac{d\Phi}{dt} = -N\int \frac{d(B\cos\alpha)}{dt}A \tag{3}$$

See ELECTROMAGNETIC INDUCTION; ELECTROMOTIVE FORCE (EMF). Kenneth V. Manning

Farm crops

The farm crops may be roughly classed as follows: (1) food crops—the bread grains (wheat and rye), rice, sugar crops (sugarbeets and sugarcane), potatoes, and dry legume seeds (peanuts, beans, and peas); (2) feed crops—corn, sorghum grain, oats, barley, and all hay and silage; and (3) industrial crops—cotton (lint and seed), soybeans, flax, and tobacco. *See* BARLEY; BEAN; CORN; COTTON; FLAX; GRAIN CROPS; OATS; PEA; PEANUT; POTATO, IRISH; POTATO, SWEET; SORGHUM; SOYBEAN; SUGAR CROPS; TOBACCO.

Regional cultivation. Crop production is regionalized in the United States in response to the combination of soil and climatic conditions and to the land topography, which favors certain kinds of crop management. In general, commercial farm crops are confined to land in humid and subhumid climates that can be managed to minimize soil and water erosion damage, where soil productivity can be kept at a relatively high level, and where lands are smooth enough to permit large-scale mechanized farm operations. In less well-watered regions, cropping is practiced efficiently on fairly level, permeable soils, where irrigation water can be supplied. The tilled crops, such as corn, sorghums, cotton, potatoes, and sugar crops, are more exacting in soil requirements than the close-seeded crops, such as wheat, oats, barley, rye, and flax. The crops planted in soil stands, mostly hay crops (as well as pastures), are efficient crops for lands that are susceptible to soil and water erosion.

United States productivity. The United States has the highest production of any area of its size in the world with regard to farm crops, together with pasture and rangelands that support livestock. Productivity per acre has increased tremendously over the years because of a combination of factors that have resulted in more efficient systems of farming. These factors include greatly improved use of mechanical and electric power and a high degree of mechanization applied to land and water management, methods of land preparation, planting, crop protection (against insects, diseases, and weeds), harvesting, curing, and storage. There have been marked advances in improving varieties of all crops and in seed technology. The development and application of effective pesticide chemicals and the greatly increased use of commercial fertilizers have been included in the newer farming systems. Soil and water conservation programs are widespread, resulting in better choice of many kinds of crops and of farming practices that are most suitable to the local climatic conditions and

to the capabilities of the local classes of soils. *See* AGRICULTURAL SCIENCE (PLANT); AGRICULTURAL SOIL AND CROP PRACTICES; FERTILIZING; LAND DRAINAGE (AGRICULTURE).

Farming regions. The major farming regions of the United States are named from the predominant kinds of crops grown, even though there is tremendous diversity within each region. The Corn Belt includes a great central area extending from Nebraska and South Dakota east across much of Iowa, Missouri, Illinois, and Indiana to central Ohio. To the north and east of this region is the Hay and Dairy Region, which actually grows large quantities of feed grains. To the south of the Corn Belt is the Corn and Winter-Wheat Belt, but here also extensive acreages of other crops are grown. The southern states, once the Cotton Belt, now concentrate on hay, pasture, and livestock, with considerable acreages of soybeans and peanuts. The cultivated portions of the Great Plains, extending from Canada to Mexcio, with annual rainfall of 15–25 in. (37.5–52.5 cm) are divided into a spring-wheat region in the Dakotas and a winter-wheat region from Texas to Nebraska, with grain sorghum a major crop in all portions of the Great Plains where soil conditions and topography favor tillage. The Intermountain Region, between the Rocky Mountains and the Cascade–Sierra Nevada mountain ranges, is cropped only where irrigation is feasible, and a wide range of farm crops is grown. In the three states of the Pacific Region, a great diversity of crops is grown. Cotton is now concentrated in the irrigated regions from Texas to California.

Marketing. From a world viewpoint, the United States is known most widely for its capacity to produce and export wheat. However, this nation has become a major producer of soybeans for export, and rice exports have become important. The ability to produce feed grains in abundance and at relatively low cost has created a large world market for United States corn, sorghum grains, oats, and barley. Although United States cotton once dominated the world market, its total production is now only about one-quarter of the world cotton supply. Most of the other farm crops are consumed within the United States. Tobacco and sugar crops are high acre-value crops, as are potatoes, peanuts, and dry beans. The production of these high acre-value crops is concentrated in localized areas where soils and climate are particularly favorable, rather than in broad acreages.

Howard B. Sprague

Fat and oil

Naturally occurring esters of glycerol and fatty acids that have commercial uses. Since fats and oils are triesters, they are commonly called triglycerides or simply glycerides. A glyceride may be designated a fat or oil, depending on its melting point. A fat is solid and an oil is liquid at room temperature. Some liquid waxes are incorrectly referred to as oils. *See* ESTER; TRIGLYCERIDE; WAX, ANIMAL AND VEGETABLE.

Structure and properties. The structure of triglycerides is shown below, where R_1, R_2, and R_3 represent the alkyl chain of the fatty acid.

$$\begin{array}{c} H_2COC(O)R_1 \\ | \\ HCOC(O)R_2 \\ | \\ H_2COC(O)R_3 \end{array}$$

The physical and chemical properties of fats and oils are determined to a large extent by the types of fatty acids in the glyceride. It is possible for all the acids to be identical, but this is rare. Usually there are two or even three different acids esterified to each glycerol molecule.

In all commercially important glycerides, the fatty acids are straight-chain, and nearly all contain an even number of carbon atoms. There are often traces of acids with an odd number of carbon atoms as well, but these are normally ignored. With the exception of castor oil, which contains the hydroxy substituted acid, ricinoleic acid, the alkyl chain is unsubstituted.

Most fats and oils are based on C_{16} and C_{18} acids with zero to three ethylenic bonds. There are exceptions, such as coconut oil, which is rich in shorter-chain acids, and some marine oils, which contain acids with as many as 22 or more carbons and six or more ethylenic linkages. The fatty acid composition of a number of commercial glycerides is shown in the **table**. The effect of chain length and degree of unsaturation on the chemical and physical properties is evident.

The majority of fats and oils are not important commercially. However, the noncommercial group contains the most unusual fatty acids, such as those with acetylenic bonds, epoxy rings, and cyclic structures.

Sources and isolation. The majority of fats and oils come from only a few sources. Plant sources are nuts or seeds, and nearly all terrestrial animal fats are from adipose tissue. Marine oils come principally from the whole body, although a small amount comes from trimmings.

Plant fats and oils are obtained by crushing and solvent extraction. Crushing is the older process, and while it is simpler to use than solvent extraction, it is less efficient. Sometimes high-fat-content sources are crushed or prepressed prior to solvent extraction to facilitate the extraction operation. *See* SOLVENT EXTRACTION.

Animal and marine oils are nearly all recovered by rendering. This is a process of heating fatty tissue with steam or hot water to melt and free the glyceride, followed by separating the oil or fat from the aqueous layer. The glyceride is processed for direct sale; the aqueous layer is first dried and then sold as a feed supplement. Tallow is obtained almost exclusively from beef adipose tissue; grease comes from swine. *See* ADIPOSE TISSUE.

Processing. There are some uses for fats and oils in their native state, but ordinarily they are converted to more valuable products. The most important changes are hydrolysis and hydrogenation.

Percentage fatty acid composition and some empirically determined properties of selected commercial glycerides[a]

	Saturated[b]							Unsaturated[c]										Average[d]			
Source	08	10	12	14	16	18	20	16:1	18:1	18:2	18:3	18:4	20:1	20:2	20:4	20:5	22:6	IV	SV	mp, °C	Titer, °C
Castor[e]					1	1			3	4								85	181	−11	
Coconut	8	7	48	17	8	4			5	3								8	255	24	22
Corn					11	2			27	59								124	190	−11	
Cottonseed				1	25	3			17	53								103	193	0	34
Crambe					2		2		17	9	6		61	1				91	170		
Linseed					6	4			19	15	57							175	192	−20	20
Palm				1	47	4			38									52	200	37	44
Palm kernel	1	3	51	18	9	2			15	1								19	250	25	24
Rapeseed[f]					4	1	1		17	13	5		56					81	175	−9	14
Soybean					10	3			22	55	8		1					130	193	−22	20
Tallow				3	26	22		3	43	1								50	200		42
Herring				1	14	1		10	24			11		3	9	13	6	148	190		24
Menhaden				7	24	3		10	17	1		4			3	12	9	178	191		28

[a]Values are approximate, do not include trace amounts of acids, and do not show the variation due to variety and growing conditions.
[b]08, 10, and so on, indicate total number of carbons in the fatty acid.
[c]18:1, 18:2, and so on, indicate total number of carbons in the fatty acid and the number of ethylenic bonds.
[d]IV = iodine value; SV = saponification value.
[e]Castor oil contains about 90% 12-hydroxy-9-octadecenoic acid.
[f]Newer varieties have less than 1% C 20:1 unsaturated acid.

Hydrolysis is commonly called splitting. The purpose is to hydrolyze the ester into its constituent glycerol and fatty acids, which are valuable intermediates for many compounds. Splitting is normally a continuous process in which the glyceride and water are fed into a tall cylindrical reactor at a temperature of about 260°C (500°F) and pressure of approximately 5200 kilopascals (750 lb/in.^{2}). Fatty acids come out the top, and sweetwater, aqueous glycerol, comes out the bottom; both are recovered and purified. *See* GLYCEROL; HYDROLYTIC PROCESSES.

Hydrogenation is the catalytic addition of hydrogen to ethylenic bonds, and is applicable to both acids and glycerides. Nickel is the customary catalyst, and the reaction is normally run at pressures to 3100 kPa (450 lb/in.^{2}) and temperatures to about 200°C (390°F). The purpose of hydrogenation is usually to raise the melting point or to increase the resistance to oxidation. *See* HYDROGENATION.

Analysis. At one time the characterization of fats and oils was exceedingly difficult due to their molecular weight and similar structure. Consequently, various empirical tests were devised to ensure quality and identity. Gas-liquid chromatography has largely supplanted these old tests, but some still remain and form the basis for commercial standards and trading rules.

Iodine value. The iodine value of a glyceride is related to its unsaturation: the higher the iodine value, the greater the unsaturation and the greater the liquidity. It is expressed in grams of iodine absorbed by 100 g of oil.

Saponification value. In this indirect procedure the ester is saponified and then back-titrated to determine the mean molecular weight of the glyceride molecule. The lower the molecular weight, the greater the saponification value. *See* DETERGENT; SOAP.

Acid value. The acid value of an oil or fat indicates the amount of nonesterified fatty acid present. The determination is made by alkaline titration.

Solidification point. This is the temperature at which a liquefied sample solidifies under certain standardized conditions. The solidification point is not necessarily the same as the melting point.

Titer. Titer is the temperature at which the first permanent cloud appears when a molten sample is cooled in a specified manner. It is an important test for detecting small amounts of contaminants.

Color. Color is almost invariably related to quality: the darker the color, the poorer the quality. There are many methods for determining color, most of which were developed for a specific product or group of products.

Gas-liquid chromatography. Gas-liquid chromatography is the most valuable analytical procedure available to the chemist of fats and oils. By selecting the proper column and operating temperatures, it is possible to obtain a quantitative estimate of the chain length, the amount of unsaturation, and the types of substitution associated with the fatty-acid chain.

Since gas-liquid chromatography requires so little sample and can be quantified, it has been used to determine, indirectly, the iodine value and the acid value of fats and oils when there has been insufficient sample for conventional analysis. Because gas-liquid chromatography is so sensitive to small amounts of trace substances, it is especially useful for distinguishing between two sources of glycerides or their derivatives. *See* CHROMATOGRAPHY; DRYING OIL; ESSENTIAL OILS; FAT AND OIL (FOOD).

Howard M. Hickman

Bibliography. R. J. Hamilton and A. Bhati (eds.), *Fats and Oils*: *Chemistry and Technology*, 1981; R. J. Hamilton and A. Bhati (eds.), *Recent Advances in Fats and Oils*, 1987; K. S. Markley, *Fatty Acids*:

Their Chemistry, Properties, Production, and Uses, 5 pts., 2d ed., 1983; D. Swern, *Bailey's Industrial Oil and Fat Products*, 4th ed., 2 vols., 1979–1982.

Fat and oil (food)

One of the three major classes of basic food substances, the others being protein and carbohydrate. Fats and oils are a source of energy. They also aid in making both natural and prepared foods more palatable by improving the texture and providing a more desirable flavor. *See* CARBOHYDRATE; FOOD; PROTEIN.

Fats and oils are esters of glycerol and fatty acids. They contain small amounts of other fat-soluble compounds—some common to all fats and others depending on the natural source. Fats and oils are soluble in organic solvents such as petroleum hydrocarbons, ether, and chloroform, but are insoluble in water. By definition, fats are more or less solid at room temperature, whereas oils are liquid. Some tropical oils, such as coconut oil, are solid at normal room temperature, but are liquid in their natural habitat.

Fats are grouped according to source. Animal fats are rendered from the fatty tissues of hogs, cattle, sheep, and poultry. Butter is obtained from milk. Vegetable oils are pressed or extracted from various plant seeds, primarily from soybean, cottonseed, corn (germ), peanut, sunflower, safflower, olive, rapeseed, sesame, coconut, oil palm (pulp and kernel separately), and cocoa beans. Marine oils are not consumed in the United States, but commonly are elsewhere. They are obtained mostly from herring, sardine, and pilchard.

Nutritive value. Fats and oils are important in the diet. They are the most concentrated form of food energy, contributing about 9 cal/g (38 joules/g), as compared to about 4 cal/g (17 joules/g) for carbohydrates and proteins. Fats make a meal more satisfying by creating a feeling of fullness, and also delay the onset of hunger. Contrary to popular belief, fats are highly digestible, with 94–98% of the ingested fat being absorbed from the intestinal tract.

The polyunsaturated fatty acids, primarily linoleic and arachidonic, are essential nutrients; that is, they are not synthesized by the body but are required for tissue development. Absence of these fatty acids from the diet results in an essential fatty acid syndrome and in a specific form of eczema in infants. Vegetable oils are an excellent source of linoleic acid, while meat fats provide arachidonic acid in small but significant amounts. Fats and oils are carriers of the oil-soluble vitamins A and D, and are the main source of vitamin E. They also have a sparing action on some of the B complex vitamins. *See* NUTRITION; VITAMIN.

Chemical constitution. Fats and oils, also called triglycerides, are esters of the trihydric alcohol glycerol, $C_3H_5(OH)_3$, and various fatty acids (see **table**). Most naturally occurring fatty acids are straight carbon chains with a terminal carboxyl group and 4–24 carbon atoms in even numbers. A few acids have an odd number of carbons and some have a cyclic group or branched chain, but all of these are relatively rare. Triglyceride structure is usually represented as shown in structure (**1**), where R, R′, and

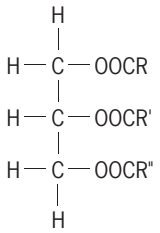

(**1**)

R″ represent carbon chains of the fatty acids. *See* TRIGLYCERIDE.

The fatty acid may be saturated or unsaturated. In a saturated acid, such as palmitic acid (**2**), all bonds between carbon atoms are single bonds, with hydrogen atoms attached to all the carbon atoms, except that of the carboxyl radical. In an unsaturated acid, each of two adjacent carbons lacks one hydrogen, so that the carbon bonds link together to form a double bond. Oleic acid (**3**), with one double bond, is a monounsaturated acid. Linoleic, linolenic, and arachidonic acids are polyunsaturated, having two, three, and four double bonds, respectively.

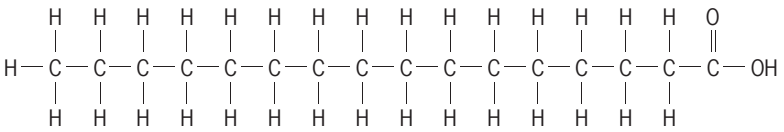

(**2**)

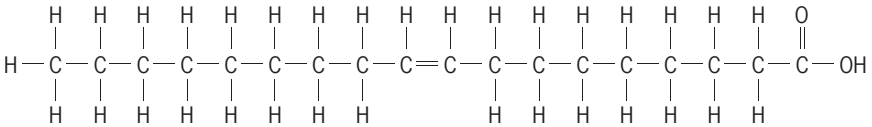

(**3**)

Typical fatty acid composition of common fats and oils[a]

Type	Saturated			Unsaturated		
	Palmitic	Stearic	Other[b]	Oleic	Linoleic	Other
Lard	24.8	12.3	2.6	45.1	9.9	5.3[c]
Beef fat	25.5	21.6	6.5	38.7	2.2	5.5[c]
Butterfat	26.2	12.5	26.5	28.2	2.9	3.7[c]
Palm oil	45.1	4.7	1.6	38.8	9.4	0.4
Cocoa butter	25.8	34.5	1.2	35.3	2.9	0.3
Soybean	11.0	4.0	0.5	23.4	53.2	7.9[d]
Cottonseed	24.7	2.3	1.1	17.6	53.3	1.0[e]
Corn	12.2	2.2	0.1	27.5	57.0	1.0
Peanut	11.6	3.1	5.7	46.5	31.4	1.7
Sunflower	6.8	4.7	1.1	18.6	68.2	0.6
Olive	13.7	2.5	0.9	71.1	10.0	1.8
Coconut	8.4	2.5	81.1	6.5	1.5	0.0

[a]As weight percentages of component fatty acids.
[b]Butterfat and coconut oil contain saturated fatty acids having 4–14 and 6–14 carbon atoms, respectively. Peanut oil contains fatty acids with 20–24 carbon atoms.
[c]Mainly palmitoleic acid and about 0.4% arachidonic acid.
[d]Mostly linolenic acid.
[e]Includes 0.5–1% malvalic and sterculic acids with a terminal cyclopropene ring structure.

Unsaturated acids can be converted to saturated acids by the addition of hydrogen.

Chain length is also important. The most common saturated fatty acids in edible fats and oils are palmitic (16 carbons), stearic (18), and lauric (12). The important unsaturated acids have 18 carbons, except for arachidonic with 20.

The physical characteristics of fat depend on the distribution of fatty acids on the various triglyceride molecules. The distribution is complex and nonrandom. Lard, for example, has palmitic acid almost exclusively on the second or middle carbon atom of glycerine. Vegetable oils have unsaturated fatty acids in this position, with saturates exclusively on the end carbons of glycerine. Cocoa butter consist almost exclusively of 2-oleo-palmitostearin and 2-oleodistearin.

Most fats contain 0.5–2% nonglyceride components which are classified as unsaponifiables, including squalene, carotenoids, tocopherols, and sterols. Tocopherols, also known as vitamin E, have antioxidant activity. Cholesterol is a sterol found only in animal fats. Plant sterols are chemically related to cholesterol, but are physiologically inactive in humans, except that they may competively inhibit absorption of cholesterol. *See* CHOLESTEROL; STEROL; VITAMIN E.

Production methods. Processing of fats and oils is carried out in a series of individual steps: extraction, refining, bleaching, and deodorization. Additional processes for specific products include winterization, hydrogenation, and texturizing. Typically there are many variations in both equipment and technique for each process. The selection of equipment also depends on the oilseed or fatty tissue to be processed.

Fats and oils are contained in seed or animal tissues within proteinaceous cell walls. The proteins are coagulated, usually through the use of heat, to release the oils. Animal tissues need no prior preparation other than cutting or comminuting them. Oilseeds are prepared in a variety of ways before oil can be removed: cottonseed must be dehulled; soybeans may or may not be dehulled; sunflower seeds are rarely dehulled; coconuts, palm kernels, and peanuts must have their shells removed; corn germ must be separated from the kernel before its oil can be removed. *See* CORN; COTTON; PEANUT; SOYBEAN; SUNFLOWER.

Fat rendering. Beef fat or tallow is usually dry-rendered under vacuum. The cracklings are then strained from the bulk of the fat and pressed to remove more tallow. Prime steam lard is wet-rendered by heating hog fat under 40–60 lb/in.^{2} (276–414 kilopascals) pressure by using steam injection into the vessel. After cooking for 4–6 h, the mixture is allowed to settle. The separated lard is then drawn off. In low-temperature rendering, comminuted fatty tissues are heated to 115–120°F (46–49°C) to melt the fat, which is separated by centrifugation. The tissue residues may be used in edible meat products, such as sausage.

Oil extraction. Oil is removed from oilseeds by pressing, solvent extraction, or a combination of both. First, however, the seeds must be crushed or flaked

and cooked to denature the proteins. The moisture content must also be adjusted to an optimum level for the particular seed and extraction method. Some mills use high pressure presses which also develop high temperatures in the oil and cause some degradation. Extraction with hexane, the most commonly used solvent, gives the best-quality oil, but is not practical for all oilseeds. Prepressing seeds at low pressure (and consequently low temperature) and removing the balance of the oil with solvent is frequently the best compromise.

Oil refining. Crude oils received from the mill must be refined for use. The first step in the process is called refining; the term "fully refined" refers to an oil which has been refined, bleached, and deodorized. Refining as such is carried out by mixing the crude oil with a water solution of sodium hydroxide at about 150°F (66°C). The alkali reacts with the free fatty acids to form soap, which is insoluble in oil. Mucilaginous gums, oil soluble in dry form, are hydrated and become insoluble. Other minor components, such as part of the sterols and tocopherols, are also incorporated in the insoluble "soap stock." The entire mass is pumped through a continuous centrifuge to remove the oil-insoluble matter. The oil is then water-washed, recentrifuged, and dried under vacuum. *See* CENTRIFUGATION.

Some vegetable oils, especially soybean oil, may first be degummed by using water to hydrate the gums. The crude gums which separate from the oil are dried to yield crude lecithin, which is useful as an emulsifier. The degummed oil is then alkali-refined as before.

Animal fats and water- or phosphoric acid–degummed vegetable oils may be steam-refined, stripping free fatty acids, monoglycerides, unsaponifiable waxes, and some pigments. During this process, the oil is held under high vacuum (4–6 torr or 0.5–0.8 kPa) and high temperature (about 450°F or 230°C).

Bleaching. Bleaching, the process for removing pigments from fats and oils, is usually carried out by adding about 1% bleaching clay to oil under vacuum at about 225°F (107°C), agitating, and then filtering to remove the clay. High temperature drives moisture from the clay so that it will adsorb the pigments. Neutral clay removes most pigments, but acid-activated earth also removes chlorophyll. Carotene, found in palm oil and tallow, is not removed by adsorption but can be decolorized by heat bleaching during deodorization. Some pigments are fixed in the oil, and cannot be removed.

Deodorization. Deodorization is used to remove volatile materials from the oil product. Along with removal of compounds which contribute flavor and odor, free fatty acids, monoglycerides, and some color bodies are distilled off. Peroxides and carotenes are decomposed. The resulting product is bland in flavor and odor. The process is carried out by blowing steam through the oil held at 400–500°F (200–260°C) and a vacuum of 5–6 torr (0.7–0.8 kPa).

Fractional crystallization. Fractional crystallization of fats is the separation of the solidified fat into hard fractions (stearine) and soft fractions (oil). It is carried out by allowing the fat to crystallize at a preselected temperature and filtering or pressing the mass to separate the oil from the stearine. The process is used for preparing high-stability liquid frying oils from partially hydrogenated soybean oil and from palm oil.

Hard butters are made by fractionating hydrogenated oils or naturally hard fats from solvent to isolate the disaturated, monounsaturated triglycerides, which resemble cocoa butter in melting characteristics.

Winterization. Winterization is a specialized form of fractional crystallization of fats. Natural cottonseed oil and partially hydrogenated soybean oil contain solid fats at low to moderate room temperature. This makes these oils undesirable in appearance, especially when refrigerated, and useless for mayonnaise, where the fat crystals would break the emulsion. Winterization is achieved by holding the oil at 40°F (4°C) for 2–4 days to crystallize the solid fats, which are removed by filtration. Crystal inhibitors, primarily oxystearin and polyglycerol esters of fatty acids, may be added to the clear oil to retard further crystallization of fats, which are rarely completely removed by winterization. Some oils, particularly from sunflower and corn, have a high level of wax, which causes a cloudy appearance. These oils are dewaxed in a process resembling winterization.

Hydrogenation. Hydrogenation is the chemical binding of hydrogen to the double bonds of unsaturated fatty acids. Glyceride molecules with two or more unsaturated acids are liquids at room temperature. Saturating one or more of these acids makes the glyceride more solid and also more stable against oxidation. The hydrogenation process calls for vigorous agitation of a mixture of oil, finely divided nickel metal catalyst (0.025–0.3%), and hydrogen gas at high temperature (usually 250–400°F or 120–200°C) and gas pressure ranging from atmospheric (14.7 lb/in.^{2} or 101 kPa) to 60 lb/in.^{2} (414 kPa). The end point is measured by the refractive index, which correlates with the iodine value, and is predetermined by experience to give the desired hardness of product. Characteristics of the hardened oil depend on process temperature, type and amount of catalyst, purity and pressure of hydrogen, and other similar variables. The fat is rebleached and filtered to remove the catalyst as a final step. *See* HYDROGENATION.

Interesterification. Interesterification is used to rearrange a nonrandom mixture of fatty acids on various triglyceride molecules into a randomized distribution. It is used mainly to change lard from a coarse crystalline fat to a smooth, textured shortening which will cream in bakery products, and to convert a mixture of fully hydrogenated coconut or palm oils and fully hydrogenated cottonseed or other nonlauric oils to hard butters for confectionery use. Interesterification involves heating the fatty starting material with a catalyst, usually sodium metal or sodium methoxide, at 95–135°F (35–57°C) for

0.5–1 h, killing the catalyst with water, and removing the resulting soaps.

Texturizing. Texturizing of shortening compounds, usually a blend of partially hydrogenated oils or interesterified lard and 8–12% fully hydrogenated cottonseed or palm oils, is carried out by rapid cooling in a scraped-surface internal-chilling machine and whipping the cold, partially crystallized fat with air or nitrogen. The shortening is filled into containers to harden on standing. The process is analogous to the freezing of ice cream.

Chemical adjuncts. A number of additives are used to modify the performance characteristics of edible oils and shortenings. Emulsifiers are the largest and most varied group of additives. Monoglycerides are the basic emulsifiers for incorporation of air into cakes and icings and for retarding staling in yeast-raised doughs. These attributes are enhanced by the addition of lactylated or ethoxylated monoglycerides, polyglycerol esters of fatty acids, propylene glycol monostearate, and polysorbates. Lecithin, a very potent emulsifier, is useful as an antisticking and antispattering agent in grill frying oils and margarine and as a texture modifier in chocolate and confectionery coatings.

Acetylated monoglycerides have some emulsifying properties, but are more useful as stable oils or waxes for coating meat products, cheese, nuts, raisins, cereals, and so forth to retard moisture transfer during storage.

Antioxidants are used to retard oxidative rancidity in fats and oils. Tertiary butyl hydroquinone (TBHQ), butylated hydroxyanisole (BHA), butylated hydroxytoluene (BHT), and propyl gallate are the most widely used antioxidants. Citric acid is used as a trace-metal scavenger to inactivate the prooxidant effect of heavy metals, primarily iron and copper. Citric acid is added without exception to all oils in the final stages of deodorization. *See* CITRIC ACID.

Methyl silicones are added to some fats used in deep fat frying as an antifoam agent.

Margarine, popcorn oil, pan frying oils, and some shortenings may be colored yellow through the addition of carotene or annatto pigments. Buttery flavors are also added to margarine and popcorn oil to enhance their appeal. *See* MARGARINE.

Testing procedures. Crude oils are purchased according to rigid specifications which include free fatty acid, refining loss, bleached color, and peroxide value. Free fatty acid is found by titrating the oil with standard alkali. Refining loss is done either by cup refining, a laboratory simulation of plant refining, or by adsorption of impurities from the oil by using powdered aluminum oxide. Refined oil is bleached with clay in the laboratory for determination of color by visual comparison with a series of color standards. The most common standards are Lovibond glasses in graduated steps of yellow and red. Peroxide value, a measure of the extent of oxidative deterioration of the oil, is determined by iodine titration, in which the peroxide equivalent is found in terms of elemental iodine liberated from potassium iodide. Other evaluations of crude oil include moisture, insoluble impurities (meal or crackling residues), and volatiles such as residual solvents.

Finished salad oils must pass a cold test: the length of time the oil takes to show the first trace of fat crystals in an ice bath. The standard test requires the oil to be clear after 5.5 h. Iodine value, the amount of iodine reacting with the oil, determines the total unsaturation. Iodine values are calculated from the fatty acid composition of the oil as determined by gas-liquid chromatography. *See* GAS CHROMATOGRAPHY.

Naturally hard fats and hydrogenated vegetable oils are actually mixtures of solid and liquid triglycerides. The solid fat index (SFI) is a measure of the solid glyceride levels at various temperatures below the melting point of the fat. It is a valuable tool for determining the potential firmness of a plasticized shortening or margarine. The consistency of these products is evaluated after plasticizing by observing package penetration. A standardized needle or other pointed object is allowed to penetrate the product for several seconds. The depth of penetration must fall within predetermined limits for the material involved. Hardness is also given in terms of melting point. Fats do not have a sharp melting point, so that several empirical methods have been developed, each giving slightly different values from the others.

Potential stability of fats toward oxidation is usually determined by an accelerated test known as the active oxygen method (AOM), in which stability is measured by bubbling air through the sample at a specified temperature, usually 208°F (98°C), and determining the length of time the fat takes to reach a peroxide value of 100. Smoke point, the temperature at which a fat gives off a steady stream of smoke, is determined for frying fats. Monoglyceride content is an important evaluation for bakery shortenings. Finished oils are also evaluated for peroxide value, free fatty acid content, moisture, color, and flavor.

Many other tests exist to give special information where needed, but are not in general use.

Deterioration factors. Several forms of deterioration may occur in fats and oils.

Flavor reversion. Flavor and odor may develop after deodorization of a product to complete blandness. The flavor is generally characteristic of the oil source and is therefore usually acceptable. However, soybean oil can develop disagreeable flavors described as beany, grassy, painty, fishy, or like watermelon rind. Beef fat can become tallowy, which is also objectionable. Reversion is apparently caused by changes in substances which have been oxidized prior to, but not removed by, deodorization. Ordinary chemical tests do not indicate any change in the oil. It is strictly an organoleptic observation, although work using gas chromatography has demonstrated that a large number of volatile compounds form during development of reverted flavor. Those compounds which have been identified indicate that

reversion is related to oxidation of unsaturated fatty acids.

Oxidative rancidity. This is a serious flavor defect and highly objectionable. It starts with the formation of hydroperoxides at the double bond of fatty acids, primarily linolenic and linoleic, which then decompose to form aldehydes which have a pungent, disagreeable flavor and odor. Oxidative rancidity is detectable chemically by peroxide value and by one of several tests for aldehydes. Peroxides and aldehydes in crude oils are removed by deodorization. They are reformed in finished oils by further oxidation, which is catalyzed by exposure of the oil to light, to some metals, especially copper and iron, and to excessive heat. Retardation of oxidation is brought about by using opaque, airtight containers, or nitrogen blanketing if clear glass bottles are used. Antioxidants, which retard oxidation by interrupting the reaction of oxygen radicals, are required in meat fats, since lard, tallow, and so on contain no natural antioxidant material. Vegetable oils contain tocopherols. Additional antioxidant, with the exception of TBHQ, has little benefit for these oils. Copper metal and copper-bearing alloys are extremely active prooxidants, and must never be used in oil-handling or oil-processing equipment. Iron rust is a prooxidant, but recently cleaned iron, coated with oil to prevent rusting, is inert, and is commonly used for oil-processing and oil-storage equipment.

Hydrolytic rancidity. This type of rancidity results from the liberation of free fatty acids by the reaction of fats and oils with water. While most fats show no detectable off flavors, coconut and other lauric acid oils develop a soapy flavor, and butter develops the strong characteristic odor of butyric acid. In some foods, active lipase from materials such as raw nuts or low-temperature–pasteurized milk or from contaminating microorganisms catalyzes the hydrolysis of fats and oils. Packaged coconut-oil products and lauric-type hard butters sometimes contain added lecithin, which acts as a moisture scavenger, thereby retarding hydrolytic rancidity development.

Frying fat breakdown. Fats and oils used in deep fat frying can break down under adverse conditions, especially where frying is intermittent or the fryer capacity is not fully used. This results in a low fat turnover rate, that is, an insufficient percentage of fresh fat being added to the fry kettle at regular intervals. The breakdown is detected first by an increase in acidity, measured by free fatty acid determination but actually caused by development of acidic breakdown products. Further deterioration results in the oil becoming very dark in color, viscous, foul-odored, and foaming badly during frying. It becomes oxidized and then polymerized, requiring that it be discarded, since it imparts strong off flavors to the fried food.

Crystal transformation. Crystal structure transformation of packaged shortening results in formation of a grainy, soft product, which may also lose incorporated gas and take on the appearance of petroleum jelly. In extreme cases, liquid oil pockets form in the shortening mass. In addition, there is a loss of creaming ability in baking of cakes and icing preparation. Crystal transformation is caused by improper hydrogenation, poor formulation, unstable hard fat used for texturizing, or the wrong chilling conditions. Unfortunately, such transformation takes several days after filling the shortening before it becomes obvious. Similar changes in crystal structure can cause bloom in chocolate coatings, a defect which gives a white haze or even open grain on an originally smooth, glossy surface. Chocolate bloom can be inhibited by the addition of lecithin or polysorbates or both. *See* COCOA POWDER AND CHOCOLATE; FOOD MANUFACTURING; LIPID. Theodore J. Weiss

Bibliography. H. J. Deuel, Jr., *The Lipids: Their Chemistry and Biochemistry*, 3 vols., 1951–1957; E. W. Eckey, *Vegetable Fats and Oils*, 1954; A. Kramer and B. A. Twigg (eds.), *Quality Control for the Food Industry*, vol. 2, 3d ed., 1972; V. C. Mehlenbacher, *The Analysis of Fats and Oils*, 1960; *Official and Tentative Methods of the American Oil Chemists' Society*, 1946–1983; D. Swern, *Bailey's Industrial Oil and Fat Products*, 4th ed., vol. 1, 1979, vol. 2, 1982; T. J. Weiss, *Food Oils and Their Uses*, 2d ed., 1983.

Fate maps (embryology)

Diagrams of embryos showing what will happen to each small region in the course of development. These diagrams show where the cells move and what structures they develop into. *See* EMBRYOGENESIS.

Animals. In animals, fate maps are usually constructed by applying small patches of label to the embryo and locating the position of the labeled cells at subsequent times during development (**Fig. 1**). In the early twentieth century, vital dyes or fine carbon particles were used for labeling. Vital dyes are stains that can color living tissues without harming them, although they do tend to diffuse and fade. In the late twentieth century, the carbocyanine dyes diI and diO were introduced; they are intensely fluorescent and are retained within the membranes of cells.

Fate maps may also be constructed by grafting small explants of tissue from a labeled embryo to the same position in an unlabeled host and observing the movements and later development of the cells of the graft. The donor for such a graft would usually be an embryo carrying a marker gene such as the gene for green fluorescent protein. There are also lineage labels, such as the enzyme horseradish peroxidase or fluorescent derivatives of the polysaccharide dextran, that can be injected into cells and do not diffuse out again. These may be used either for labeling the donor embryos for grafts or to label individual cells by injection.

A combination of these methods has enabled accurate fate maps to be produced for the types of embryos most often used in biological research: those of the amphibian (**Fig. 2**), chick, mouse, fruit fly,

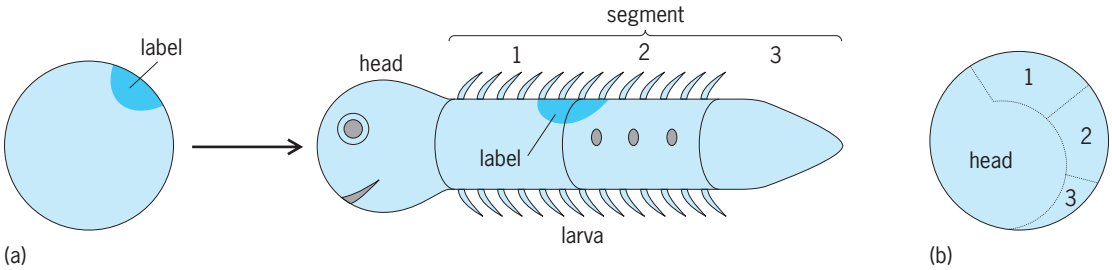

Fig. 1. Principle of fate mapping. (*a*) Label placed in a certain position on the egg ends up in a reproducible position on the later animal. (*b*) Possible fate map for this type of egg. (*Reprinted with permission from J. M. W. Slack, Essential Developmental Biology, 2d ed., Blackwell Science, 2005*)

and zebrafish. For embryos in which there is no increase in size during development and no random mixing of cells, the fate map can be precise from the fertilized egg onward. This is the case for some invertebrates such as gastropods or ascidians. For embryos in which there is some local cell mixing, as in vertebrates, the fate map cannot be so precise, and represents the average behavior of a population of embryos.

Although it is a common misconception, a fate map of multicellular embryo regions does not, in itself, give any information about the intrinsic developmental commitment of the cells of the embryo to form particular structures or cell types. For example, in cases in which the commitment of a region is acquired in response to an inductive signal from another region, the commitment may differ from the fate shown on the fate map at the time before the inductive signal operates. States of commitment must be deduced by other types of experiment such as culturing embryo regions in isolation or grafting them to different positions in a host embryo.

An associated concept to the fate map is the cell lineage, which shows the family tree of cells descended from a single precursor cell. Unlike a fate map, a cell lineage does not show the spatial relationships between cells. The lineage of every cell has been established from the fertilized egg to the adult in the nematode *Caenorhabditis elegans*, in which every individual embryo undergoes exactly the same sequence of cell divisions.

The term clonal analysis is used to refer to the deduction of developmental mechanisms from the study of cell lineage. For example, if one labeled cell predictably forms a whole structure in the later organism, then it must be the sole precursor to that structure. Conversely, if one labeled cell gives rise to progeny that span two structures, then it could not have been committed to form either structure at the time of labeling. *See* CELL LINEAGE; DEVELOPMENTAL BIOLOGY; EMBRYOLOGY. J. M. W. Slack

Plants. In clonal analysis of plants, cells are labeled by exposing the seed or developing plant to ionizing irradiation or a chemical mutagen to produce chromosome mutations that result in distinct phenotypic alterations, usually deficiencies in the pigments chlorophyll or anthocyanin. The low frequency of mutations produced indicates that these are single-cell events. Because all of the progeny of a labeled cell will carry the same chromosome mutation, a shoot meristem cell labeled with a chlorophyll deficiency mutation, for example, will produce a sector, or clone, of white tissue in the developing plant (**Fig. 3**). It is possible to deduce the number and fate of meristem cells labeled at one stage of development by examining the size and position of sectors present in the plant at a later stage of development. For example, a sector extending the entire length of the shoot would be generated by a permanent initial cell at the center of the meristem; if the width of

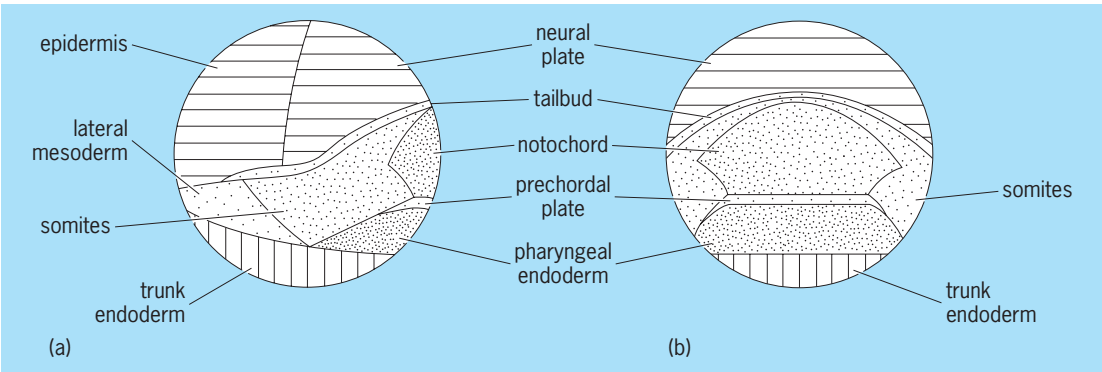

Fig. 2. Fate maps of urodele blastula from (*a*) lateral and (*b*) dorsal aspects, deduced from vital staining. (*After J. M. W. Slack, From Egg to Embryo: Determinative Events in Early Development, Cambridge University Press, 1983*)

that sector occupied one-third of the circumference of the shoot, then the shoot must have been formed by three such initial cells. *See* MUTATION; PLANT GROWTH.

Fate maps have been constructed for the shoot apical meristems of maize (*Zea mays*) and sunflower (*Helianthus annuus*) embryos based on pigment-deficient sectors generated in plants grown from irradiated seed (**Fig. 4**). The pattern of sectors in both plants indicates that several domains of cells from the periphery to the center of the meristem give rise to successively more apical portions of the mature plant. It also shows that the contribution of cells occupying similar positions in the meristem is variable. Thus the fate of cells in the meristem of these two plants is not fixed in the embryo: no cells are committed to a particular developmental fate, such as the production of the reproductive parts of the plant. The number of cells and cell domains predicted by the fate maps of maize and sunflower is consistent with the actual number and arrangement of cells observed in histological sections of the meristems.

Similarly, the fate of cells in developing leaf primordia of tobacco (*Nicotiana tabacum*) has been established by irradiating primordia at different developmental ages. A different type of pigmentation chimera that was produced by certain interspecific crosses has been used to construct a fate map of cells contributing to the cotyledons, to the first true leaves, and to the shoot meristem in the developing cotton (*Gossypium*) embryo. Earlier, the fate of cells in the different cell layers of the shoot meristem of several plants was deduced from studies of chimeras in which different cell layers of the meristem are genetically distinct for pigmentation or chromosome number. That research showed that whereas cells in a particular layer of the meristem generally contribute to a particular part of the shoot, such as the outermost layer to the epidermis or the third layer to the vascular tissue and pith, the cell layers do not have fixed developmental fates. *See* CHIMERA; DEVELOPMENTAL BIOLOGY; EMBRYOLOGY. Dorothy E. Jegla

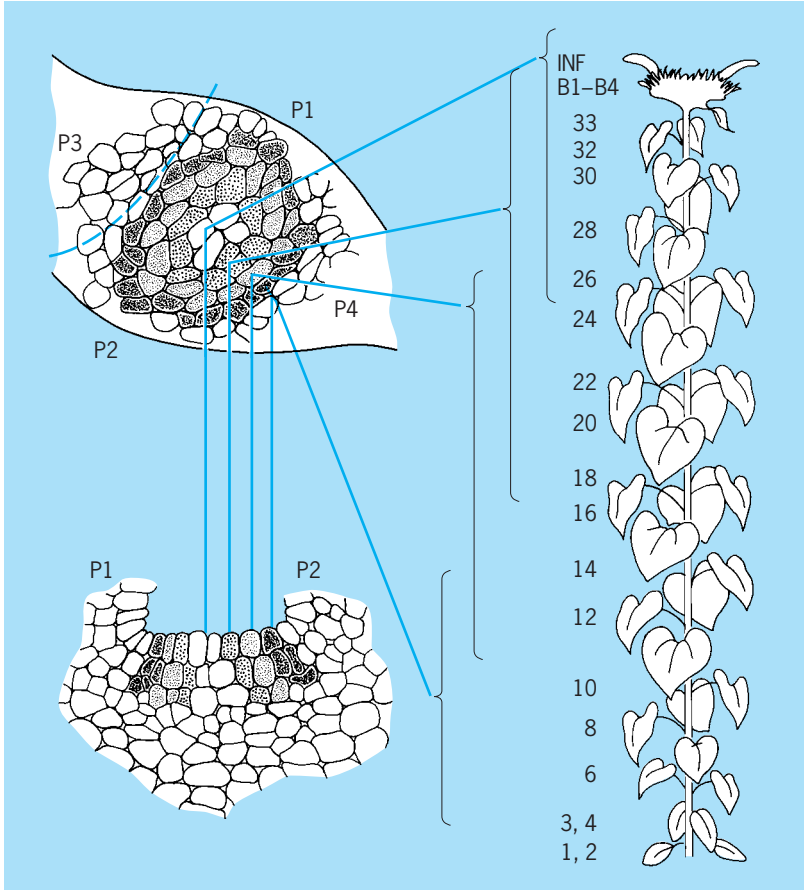

Fig. 4. Fate map of the shoot apical meristem of the sunflower embryo. Cell drawings on the left show the actual number and arrangement of cells in the surface layer (top) and median longitudinal section (bottom) of a typical meristem. Positions of the first four leaves that are present in the embryo are labeled P1–P4. Cells are differentially shaded to show four domains of cells that give rise to the successively more apical portions (indicated by brackets) of the mature plant diagrammed on the right. Numbers refer to node number: B1–B4 to the involucral bracts, and INF to the inflorescence. (*After D. E. Jegla and I. M. Sussex, Cell lineage patterns in the shoot meristem of the sunflower embryo in the dry seed, Dev. Biol., 131:215–225, 1989*)

Fig. 3. Chlorophyll-deficient sector in a sunflower plant grown from irradiated seed. The sector represents a clone of tissue derived from a single cell present in the meristem of the embryo at the time of irradiation.

Bibliography. D. E. Jegla and I. M. Sussex, Cell lineage patterns in the shoot meristem of the sunflower embryo in the dry seed, *Dev. Biol.*, 131:215–225, 1989; J. M. W. Slack, *From Egg to Embryo: Regional Specification in Early Development*, 2d ed., Cambridge University Press, 1991; J. M. W. Slack, *Essential Developmental Biology*, 2d ed., Blackwell Science, 2005.

Fault analysis

The detection and diagnosis of malfunctions in technical systems. Such systems include production equipment (chemical plants, steel mills, paper mills, and power stations), transportation vehicles (ships, airplanes, automobiles), and household appliances (washing machines, air conditioners). In any of these systems, malfunctions of components may lead to damage of the equipment itself, degradation of its

function or product, jeopardy of its mission, and hazard to human life. While the need to detect and diagnose malfunctions is not new, advanced fault detection has been made possible only by the proliferation of the computer. Fault detection and diagnosis actually means a scheme in which a computer monitors the technical equipment to signal any malfunction and determines the components responsible. The detection and diagnosis of the fault may be followed by automatic actions, enabling the fault to be corrected such that the system may operate successfully even under the particular faulty condition.

Diagnostic concepts. Fault detection and diagnosis applies to both the basic technical equipment and the actuators and sensors attached to it. In the case of a chemical plant, the former includes the reactors, distillation columns, heat exchangers, compressors, storage tanks, and piping. Typical faults are leaks, plugs, surface fouling, and broken moving parts. The actuators are mostly valves, together with their driving devices (electric motors and hydraulic or pneumatic drives). The sensors are devices measuring the different physical variables in the plant, such as thermocouples, pressure diaphragms, and flow meters. Actuator and sensor fault detection is very important because these devices are prone to faults.

The on-line or real-time detection and diagnosis of faults means that the equipment is constantly monitored during its regular operation by a permanently connected computer, and any discrepancy is signaled almost immediately. On-line monitoring is very important for the early detection of any component malfunction before it can lead to more substantial equipment failure. In contrast, off-line diagnosis involves monitoring the system by a special, temporarily attached device, under special conditions (for example, car diagnostics at a service station).

The diagnostic activity may be broken down into several logical stages. Fault detection is the indication of something going wrong in the system. Fault isolation is the determination of the fault location (the component which malfunctions), while fault identification is the estimation of its size. On-line systems usually contain the detection and isolation stage; in off-line systems, detection may be superfluous. Fault identification is usually less important than the two other stages.

Fault detection and isolation can never be performed with absolute certainty because of circumstances such as noise, disturbances, and model errors. There is always a trade-off between false alarms and missed detections, with the proper balance depending on the particular application. In professionally supervised large plants, false alarms are better tolerated and missed detections may be more critical, while in consumer equipment (including cars) the situation may be the opposite.

Approaches. A number of different approaches to fault detection and diagnosis may be used individually or in combination.

Limit checking. In this approach, which is the most widely used, system variables are monitored and compared to preset limits. This technique is simple and appealing, but it has several drawbacks. The monitored variables are system outputs that depend on the inputs. To make allowance for the variations of the inputs, the limits often need to be chosen conservatively. Furthermore, a single component fault may cause many variables to exceed their limits, so it may be extremely difficult to determine the source. Monitoring the trends of system variables may be more informative, but it also suffers from the same drawbacks as limit checking.

Special and multiple sensors. Special sensors may be applied to perform the limit-checking function (such as temperature or pressure limit sensors) or to monitor some fault-sensitive variable (such as vibration or sound). Such sensors are used mostly in noncomputerized systems. Multiple sensors may be applied to measure the same system variable, providing physical redundancy. If two sensors disagree, at least one of them is faulty. A third sensor is needed to isolate the faulty component (and select the accepted measurement value) by "majority vote." Multiple sensors may be expensive, and they provide no information about actuator and plant faults.

Frequency analysis. This procedure, in which the Fourier transforms of system variables are determined, may supply useful information about fault conditions. The healthy plant usually has a characteristic spectrum, which will change when faults are present. Particular faults may have their own typical signature (peaks at specific frequencies) in the spectrum. *See* FOURIER SERIES AND TRANSFORMS.

Fault-tree analysis. Fault trees are the graphic representations of the cause–effect relations in the system. On the top of the tree, there is an undesirable or catastrophic system event (top event), with the possible causes underneath (intermediate events), down to component failures or other elementary events (basic events) that are the possible root causes of the top event. The logic relationships from bottom up are represented by AND and OR (or more complex) logic gates. Fault trees can be used in system design to evaluate the potential risks associated with various component failures under different design variants (bottom-up analysis). In a fault diagnosis framework, the tree is used top down; once the top event is observed, the potential causes are analyzed by following the logic paths backward.

Parameter estimation. This procedure uses a mathematical model of the monitored system. The parameters of the model are estimated from input and output measurements in a fault-free reference situation. Repeated new estimates are then obtained on-line in the normal course of system operation. Deviations from the reference parameters signify changes in the plant and a potential fault. The faulty component location may be isolated by computing the new physical plant parameters and comparing them with those from the model. *See* ESTIMATION THEORY; MODEL THEORY.

Consistency checking. This is another way of using the mathematical-system model. The idea is to check if the observed plant outputs are consistent with the

outputs predicted by the model (**Fig. 1**). Discrepancies indicate a deviation between the model and the plant (parametric faults) or the presence of unobserved variables (additive faults). This testing concept is also called analytical redundancy since the model equations are used in a similar way as multiple sensors.

In preparation for fault monitoring by analytical redundancy methods, a mathematical model of the plant needs to be established. This may be done from "first principles," relying on the theoretical understanding of the plant's operation, or by systems identification using experimental data from a fault-free plant.

The actual implementation of fault monitoring usually consists of two stages (Fig. 1). The first is residual generation, where residuals are mathematical quantities expressing the discrepancy between the actual plant behavior and the one expected based on the model. Residuals are nominally zero and become nonzero by the occurrence of faults. The second stage is residual evaluation and decision making, where the residuals are subjected to threshold tests and logic analysis. Disturbances and model errors may also cause the residuals to become nonzero, leading to false alarms.

Fault isolation requires specially manipulated sets of residuals. In the most frequently used approach, residuals are arranged so that each one is sensitive to a specific subset of faults (structured residuals). Then in response to a particular fault, only a fault-specific subset of residuals triggers its test, leading to binary fault codes.

Principal component analysis (PCA). In this approach, empirical data (input and output measurements) are collected from the plant. The eigenstructure analysis of the data covariance matrix yields a statistical model of the system in which the eigenvectors point at the "principal directions" of the relationships in the data, while the eigenvalues indicate the data variance in the principal directions. This method is successfully used in the monitoring of large systems. By revealing linear relations among the variables, the dimensionality of the model is significantly reduced. Faults may be detected by relating plant observations to the normal spread of the data, and outliers indicate abnormal system situations. Residuals may also be generated from the principal component model, allowing the use of analytical redundancy methods in this framework. *See* EIGENFUNCTION.

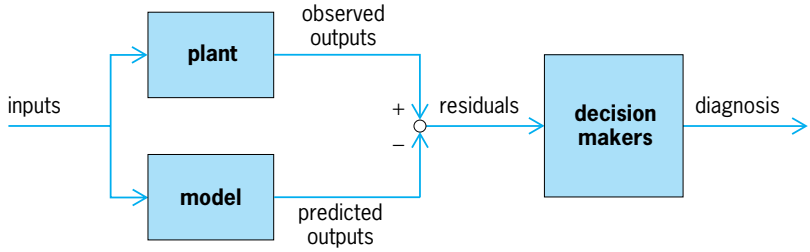

Fig. 1. Two stages of model-based fault detection and isolation. (***After R. Isermann and B. Freyermuth, eds., Proceedings of the IFAC SAFEPROCESS Symposium, Baden-Baden, Germany, September 10–13, 1991, Pergamon, 1992***)

Example of fault-tree analysis. The schematic of a simple electrical circuit in which a light is operated by a pair of three-way switches is shown in **Fig. 2**. (Such circuits are used in long hallways.) **Figure 3** shows the detailed fault tree of the circuit. The tree goes down to subcomponents (contacts of the switches) in order to illustrate more complex logic relations on this simple system. Note that nonfailure events (operating conditions) are also among the basic events because such conditions (the position of each switch) determine whether a particular failure event triggers the top event.

Example of consistency checking. Traditionally, a few fundamental variables, such as coolant temperature, oil pressure, and battery voltage, have been monitored in automobile engines by using limit sensors. With the introduction of onboard microcomputers, the scope and number of variables that can be considered have been extended. Active functional testing may be applied to at least one actuator, typically the exhaust-gas recirculation valve. Model-based schemes to cover the components affecting the vehicle's emission control system are gradually introduced by manufacturers. One approach (**Fig. 4**) uses analytical redundancy to monitor two groups of actuators (fuel injectors and exhaust gas recirculation) and four sensors (throttle position, manifold pressure, engine speed, and exhaust oxygen). By the appropriate selection of the model relations, the residuals are insensitive to the load torque and the vehicle's mass. The structured residual technique is used to support fault isolation. The critical issue

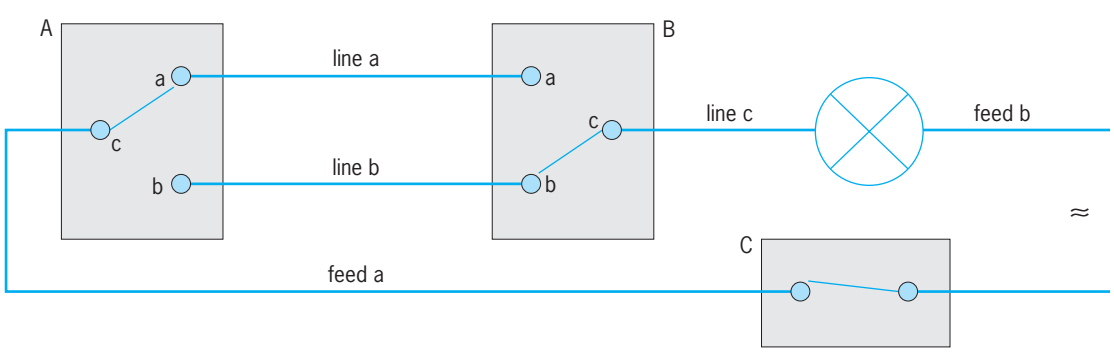

Fig. 2. Simple electrical circuit: a lamp operated by a pair of three-way switches.

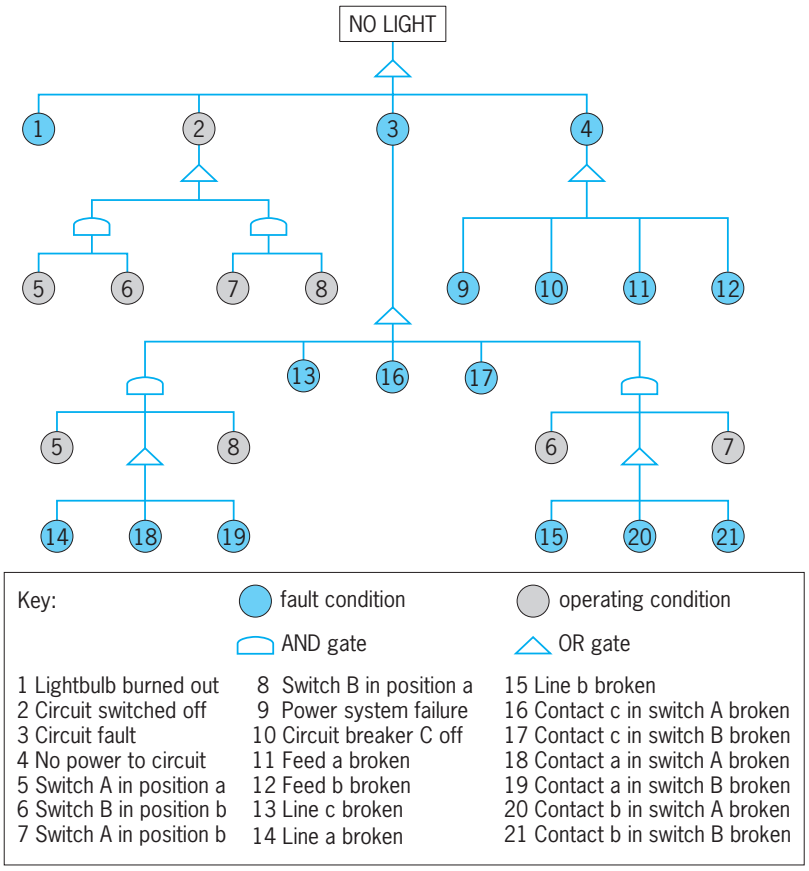

Fig. 3. Detailed fault tree of the circuit shown in Fig. 2.

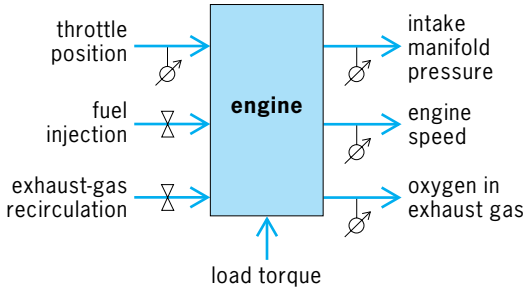

Fig. 4. Car engine system with onboard fault detection and diagnosis. (*After R. Isermann and B. Freyermuth, eds., Proceedings of the IFAC SAFEPROCESS Symposium, Baden-Baden, Germany, September 10–13, 1991, Pergamon, 1992*)

is to find sufficiently general models so that a single scheme may function well across an entire automobile product line and under widely varying operating conditions. *See* AUTOMOTIVE ENGINE; MICROCOMPUTER; MICROPROCESSOR. Janos J. Gertler

Bibliography. L. H. Chiang, R. D. Braatz, and E. Russel, *Fault Detection and Diagnosis in Industrial Systems*, Springer, 2001; P. L. Clemens, *Fault Tree Analysis*, 4th ed., Jacobs Sverdrup, 2002; J. Gertler, *Fault Detection and Diagnosis in Engineering Systems*, 1998; J. Gertler, Survey of model-based failure detection and isolation in complex plants, *IEEE Control Sys. Mag.*, 8(7):3–11, 1988; R. Patton, P. Frank, and R. Clark (eds.), *Fault Diagnosis in Dynamic Systems*, 1989.

Fault and fault structures

Products of fracturing and differential movements along fractures in continental and oceanic crustal rocks. Faults range in length and magnitude of displacement from small structures visible in hand specimens, displaying offsets of a centimeter (1 cm = 0.4 in.) or less, to long, continuous crustal breaks, extending hundreds of kilometers (1 km = 0.6 mi) in length and accommodating displacements of tens or hundreds of kilometers. Faults exist in deformed rocks at the microscopic scale, but these are generally ignored or go unrecognized in most geological studies. Alternatively, where microfaults systematically pervade rock bodies as sets of very closely spaced subparallel, planar fractures, they are recognized and interpreted as a type of cleavage which permitted flow of the rock body. Fractures along which there is no visible displacement are known as joints. These include shear joints, formed by fracturing and imperceptible movement of the walls of the fractures parallel to fracture surfaces, and tension joints, formed by negligible or barely visible displacement of the walls of the fractures perpendicular to the fracture surfaces. Large fractures which have accommodated major dilational openings (a meter or more) perpendicular to the fracture surfaces are known as fissures. Formation of fissures is restricted to near-surface conditions, for example, in areas of crustal stretching of subsidence. When faulting takes place under conditions of high temperature or pressure, zones of penetrative shear flow may develop which are best described as ductile fault zones.

Locating faults. The recognition of faults in continental regions of moderate to excellent rock exposure is generally straightforward. Classically, the process of systematic geological mapping has proved to be a powerful method for locating faults. In essence, faults can be identified and tracked by recognizing in mapped patterns the truncation and offset of one or several bedrock units. Depending on the nature of the exposure and movement on the fault, truncation and offset might produce a simple horizontal shifting of the dominant mapped pattern of bedrock units; alternatively, the faulting might lead to a pattern of repetition or omission of specific rock formations within the geologic column.

Valuable physical signatures which reveal the presence of faults include abnormally straight topographic lineaments or fault-line scarps; aligned springs issuing from fractured and favorably displaced bedrock; intensely fractured rocks, perhaps with zones of angular chunks of brecciated, rotated materials; fracture surfaces naturally polished through the movement process and etched with striations or grooves; dragging (folding) of rock layers out of their normal orientation; loosely consolidated,

ground-up rock flour or paste, commonly referred to as gouge; radically crushed, cataclastically deformed rocks known as mylonites; and alteration, silicification, or mineralization brought about by circulation of hot fluids through shattered bedrock. Faults are simple to locate in areas of active or very recent mountain building, especially where faulting has broken the ground surface and produced scarps. *See* MYLONITE.

In continental areas of very poor rock exposure, in the subsurface, and in ocean basins, faults are much more difficult and costly to locate. However, major faults are routinely discovered through application of geophysical methods, especially seismic, gravity, and magnetic surveying. Abrupt contrasts in the geophysical signatures of rocks at depth signal the sharp truncation of bedrock by faults and allow the pattern of faults to be mapped. The geophysical exploration of fault and fault structures in the ocean floor has completely changed the way in which geologists view the Earth and earth dynamics. It is known that major fracture zones, unlike any recognized in continental regions, exist in the ocean floor. These fractures, hundreds to thousands of kilometers in length and spaced at tens of kilometers, pervade the Mid-Oceanic Ridge system and are usually oriented perpendicular to the crest of the ridge segment which they occupy. Interpreting these to be enormous faults which accommodate the movement of newborn oceanic crust as it spreads bilaterally from the Mid-Oceanic Ridge system, J. T. Wilson named them transform faults. *See* GEOPHYSICAL EXPLORATION; MID-OCEANIC RIDGE; MARINE GEOLOGY.

Transform faults. Three fundamental types of transform faults exist: the first connects one ridge segment to another; the second connects a ridge segment (where new oceanic crust forms) to a trench (where oceanic crust is consumed through subduction); the third connects two trenches. Ridge-ridge transform faults, perhaps the easiest to visualize (**Fig. 1**), link parallel but offset ridge segments, and are characterized by shallow earthquake activity restricted to the part of the transform between the two ocean-ridge segments. The sense of movement along each ridge-ridge transform fault, as deduced from fault-plane solutions, is opposite to that which would explain the offset of the oceanic ridge. In essence, the movement along the transform fault does not produce the offset, but rather records the differential sliding-past (shearing) of the new ocean floor moving in opposite directions from the two ocean-ridge segments which are connected by the transform. Thus the transform faults serve, in this example, as part of the divergent plate boundary which accommodates sea-floor spreading and the movement of the two plates can be described by an imaginary pole of rotation located on the Earth's surface at a position defined by the intersection of great circles drawn perpendicular to points on the array of transform faults. The rate of relative movement along the ridge-ridge transform is a function of the rate of generation of new oceanic crust along the ridge and is generally on the order of several centimeters per year. *See* MOUNTAIN SYSTEMS; PLATE TECTONICS; TRANSFORM FAULT.

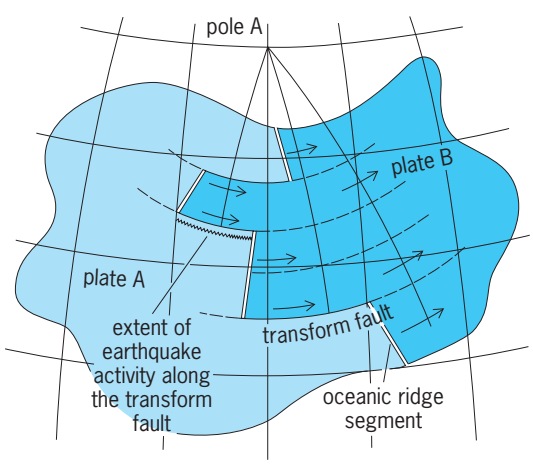

Fig. 1. Transform faults linking oceanic ridge segments. (*After W. J. Morgan, Rises, trenches, great faults, and crustal blocks, J. Geophys. Res., 73:1959–1982, 1968*)

Transform fault patterns are quite complicated, as is their evolution through time. Boundary zones between three adjacent tectonic plates (or triple junctions) are particularly difficult to evaluate with respect to geometry and kinematics (motions). Nonetheless, plate-tectonic analysis, including detailed assessment of transform faults, has had a revolutionary impact on understanding of the Cenozoic tectonic evolution of the Earth's crust. For example, scientists generally accept Tanya Atwater's hypothesis that the infamous San Andreas fault system of California is a transform fault boundary separating two enormous crustal plates, the North American and the Pacific. Relative movement between these plates is horizontal and right-handed, and in magnitude amounts to hundreds of kilometers.

Movements. This interpretation of the San Andreas Fault clearly demonstrates that some major fault systems in continental regions can be better understood in the context of the "new rules" afforded by study of transform faulting and plate tectonics. However, most fault systems in continental regions cannot be clearly and quantitatively linked to specific plate-tectonic movements or configurations, present or past, and they are analyzed and understood in an entirely different way; the guidelines for analysis are well established. In addition to describing the physical and geometric nature of faults and interpreting time of formation, it has been found to be especially important to determine the orientations of minor fault structures (such as striae and drag folds) which record the sense of relative movement.

Evaluating the movement of faulting can be difficult, for the apparent relative movement (separation) of fault blocks as seen in map or outcrop may bear little or no relation to the actual relative movement (slip). The slip of the fault is the actual relative movement between two points or two markers in the rock that were coincident before faulting

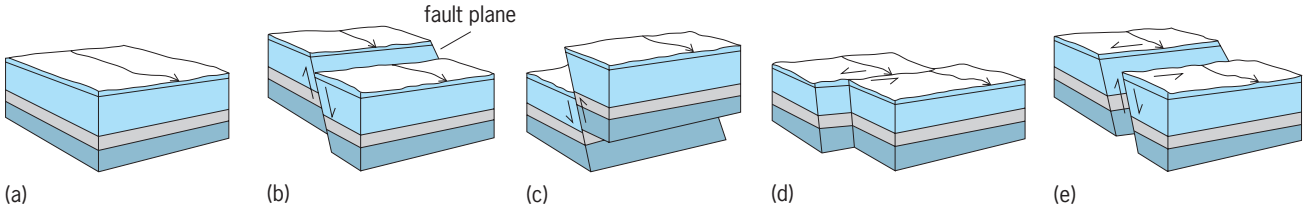

Fig. 2. Slip on faults. (*a*) Block before faulting; (*b*) normal-slip; (*c*) reverse-slip; (*d*) strike-slip; (*e*) oblique-slip. (*After F. Press and R. Siever, Earth, 2d ed., W. H. Freeman, 1978*)

(**Fig. 2**). Strike-slip faults have resulted in horizontal movements between adjacent blocks; dip-slip faults are marked by translations directly up or down the dip of the fault surface; in oblique-slip faults the path of actual relative movement is inclined somewhere between horizontal and dip slip. Strike-slip faults are described as left- or right-handed, depending on the sense of actual relative movement; dip-slip faults are described as normal-slip, thrust-slip, or reverse-slip, depending on the sense of actual relative movement and on the dip and dip direction of the fault surface. Listric normal-slip faults are a type of normal-slip fault in which the inclination of the fault decreases. Movement on such a curved fault surface produces a profound backward rotation of the upper block.

Recognizing even the simplest translational fault movements in nature is often enormously difficult because of complicated and deceptive patterns created by the interference of structure and topography, and by the absence of specific fault structures which define the slip path (**Fig. 3**). While mapping, the geologist mainly documents apparent relative movement (separation) along a fault, based on what is observed in plan-view or cross-sectional exposures. In the separation sense, left- and right-lateral faults are those displaying apparent horizontal shifts of rock in map view. Such shifts are said to be apparent because the left- or right-lateral offset might actually have been produced by dip-slip, not strike-slip, movements. Normal, reverse, and thrust are separation terms for faults, and again the usage of each is based on apparent offset in cross-sectional view and on the dip and dip direction of the fault.

Movement and offset on large, regional fault systems must be evaluated on the basis of displaced geologic terrains and abnormal stratigraphic relations. For example, 435 to 500 mi (700 to 800 km) of left-slip fault movement has been postulated in northern Mexico during the time period 150,000,000 to 170,000,000 years ago. The basis of the interpretation is truncation and offset of terrains of Precambrian and Paleozoic rocks in California and Arizona. Low-angle thrust movements in western Utah produced in Cretaceous time a 44-mi (70-km) west-to-east transport of thick Precambrian through Paleozoic miogeoclinal strata onto thin shelf and platform strata of the same age.

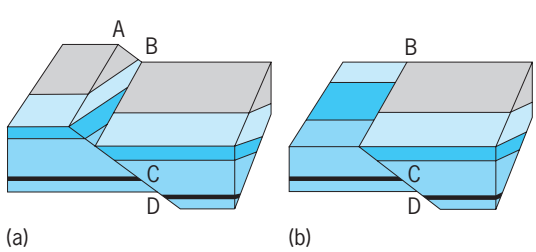

Fig. 3. Slip versus separation. (*a*) AB is the dip-slip. (*b*) After erosion of top of footwall block. The block has undergone a right separation. (*After M. P. Billings, Structural Geology, 3d ed., Prentice-Hall, 1972*)

Stress conditions. Theory on faulting is based on applied physics and engineering, and focuses on the stress conditions under which rocks break. The theory is almost exclusively concerned with the brittle behavior of crustal rocks; as such, it is most applicable to faulting at upper crustal layers in the Earth. Results of deformational experiments under controlled temperature, pressure, and strain-rate conditions bear importantly on modern understanding of the dynamics of faulting. The results of theoretical and experimental work provide insight into why faults can be conveniently separated into categories of normal-slip, thrust-slip, and strike-slip.

Forces which act on a rock body may be resolved by vector analysis into components of force acting in specific directions. These, in turn, can be converted to magnitudes and directions of stresses which tend to deform the body. This is done by dividing the force component by the surface area (of the body) on which it acts. Two types of stresses are distinguished, normal stress (σ), which acts perpendicular to a given surface, and shear stress (τ), which acts parallel to the given surface. Stress analysis, in effect, evaluates the magnitude of shear and normal stresses acting in all directions throughout a body and predicts the orientations of the surfaces along which faulting should occur.

Evaluating the distribution of stresses in a body that is acted upon by forces discloses that there are three unique directions within the stress field, called principal stress directions. These stress directions are mutually perpendicular, and the value of shear stress equals zero only along these three directions. Furthermore, one of the principal stress directions is characterized by the maximum value of normal stress (σ_1) within the system, and another (σ_3) is characterized by the minimum normal stress. Maximum shear stress values (τ_{max}) occur along lines oriented 45° to the principal stress directions.

Two-dimensional mathematical analysis demonstrates that the magnitude of normal stress (σ) on any plane (that is, on any potential surface of faulting)

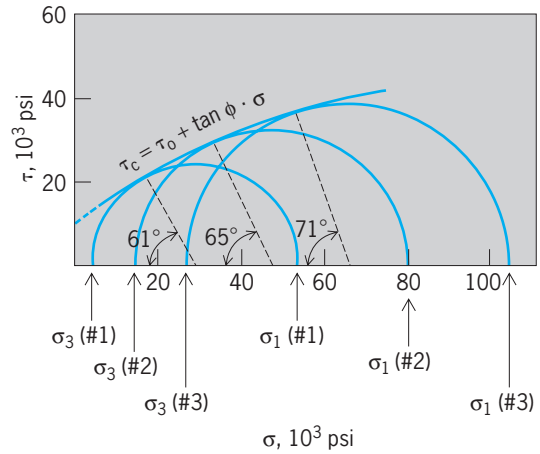

Fig. 4. Mohr circle diagram constructed from three sets of compressional tests of limestone. Angles (61°, 65°, and 71°) are 2θ values. 1 psi = 6.89 kPa.

in the body is given by Eq. (1), where σ = greatest

$$\sigma = \frac{\sigma_1 + \sigma_3}{2} - \frac{\sigma_1 - \sigma_3}{2} \cos 2\theta \qquad (1)$$

principal normal stress, σ_3 = least principal normal stress and θ = angle between the greatest principal stress axis and the direction of shearing stress (in the plane for which σ is being evaluated). The magnitude of shear stress (τ) on that same plane is given by Eq. (2). The distribution of paired values of normal

$$\tau = \frac{\sigma_1 - \sigma_3}{2} \sin 2\theta \qquad (2)$$

stress and shear stress as a function of θ is such that shear stress (τ) is zero at values of 0°, 90°, 180°, and 270°, but attains maximum values at 45°, 135°, 225°, and 310°. Normal stress reaches a maximum at 90° and 270°, but is minimal at 0° and 180°. These variations in normal and shear stress values may be portrayed on a Mohr circle diagram, a graphical representation of the above equations (**Fig. 4**). Points on the periphery of the circle have coordinates (σ, τ) which correspond in value to normal stress and shear stress on a plane which makes an angle of θ with σ_1, the greatest principal stress direction.

Given such a stress distribution, and assuming that the differential stress conditions ($\sigma_1 - \sigma_3$) exceed the strength of the rock body, the orientation of faulting can be determined. The Mohr-Coulomb law of failure [Eq. (3)] predicts that faulting should occur

$$\tau_c = \tau_0 + \sigma \tan \phi \qquad (3)$$

at a critical shear stress level (τ_c) where τ_0 = cohesive strength of the rock, σ = normal stress on the fault plane, and ϕ = angle of internal friction. The coefficient of internal friction, tan ϕ, equals σ/τ at failure. For most rocks, the coefficient of internal friction has a value between 0.4 and 0.7; thus the angle of internal friction, ϕ, commonly varies from 20 to 35°. The value of θ for such internal friction typically ranges from 27 to 35°, and is often 30°. In practice, the failure points on Mohr circle diagrams, as generated through experimental deformation, do not conform to the ideal straight-line failure envelope predicted on the basis of the Mohr-Coulomb law. Rather, the failure envelopes are smoothly curved in a way that describes an increase in 2θ with an increase in confining pressure (σ_3). Predictions as to when and at what angle faulting should occur are further complicated by variables of temperature, strain rate, pore-fluid pressure, and presence of fractures. In fact, fundamental questions have been raised regarding the extent to which theory and short-term experimental work can be applied to some natural geological systems.

What arises from theory and experiments is that fractures form in an orientation such that they contain the axis of intermediate stress (σ_2) and make an angle of θ (commonly around 30°) with σ_1 and 90 − θ with σ_3. Since the Earth's surface has no shear stress, principal stress directions near the Earth's surface tend to be vertical and horizontal, and depending on the relative configuration of the principal stresses, will give rise to thrust-slip, normal-slip, or strike-slip faults. The direction of movement on these faults is such that the wedge receiving the greatest compressive stress moves inward, whereas the wedge receiving the least compressive stress moves outward.

Examples. There are many excellent natural examples of normal-slip, thrust-slip, and strike-slip faults. The Basin and Range Province of the western United States displays a unique physiographic basin/range style because of pervasive large-scale normal-slip faulting (**Fig. 5**), which resulted from regional crustal extension within the last 15,000,000 years. Normal-slip faults accommodate extension.

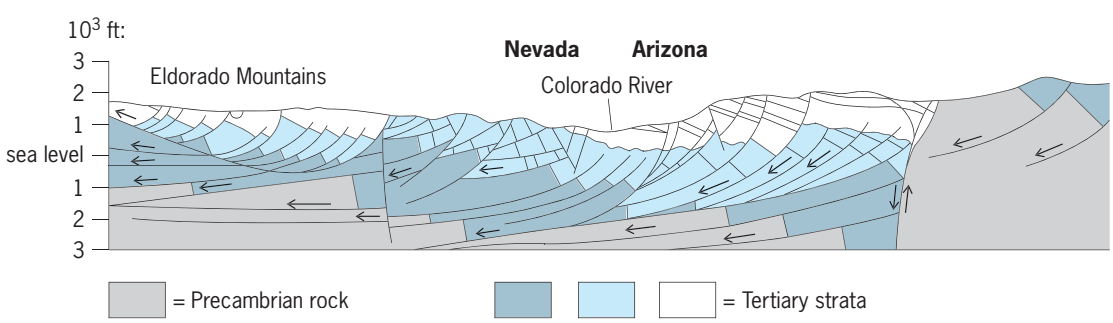

Fig. 5. Cross section of listric normal faults in the Basin and Range Province. (*After R. E. Anderson, Geologic Map of the Black Canyon 15-Minute Quadrangle, Mohave County, Arizona, and Clark County, Nevada, USGS Map GQ-1394, 1978*)

Thrust-slip faults are an integral part of the tectonic framework of the southern Appalachian Mountains, the Sevier orogenic belt of western Utah and western Wyoming/eastern Idaho, and the Canadian Rockies. In mountain belts throughout the world, thrusting has played a major role in accommodating crustal shortening during mountain building. Major mechanical questions have arisen as to how enormous masses of crustal rocks can be thrust tens or hundreds of kilometers, since the apparent force required should have crushed the rock mass before it moved. The paradox of regional overthrusting has led to theories of faulting that emphasize the importance of factors such as high pore-fluid pressure, gravitational sliding, viscous creep of ductile materials, and underthrusting.

The best-documented strike-slip faulting is concentrated near margins of lithospheric plates, but such faulting has occurred in foreland tectonic regions as well, including the Rocky Mountains of the American West and mainland China. A major theme that has emerged from the study of high-angle faults, especially strike-slip faults, is that ancient faults in basement rocks are commonly reactivated in post-Precambrian time, producing zones of concentrated, superposed strain. This interpretation has been used to explain the classic monoclinal uplifts of the Colorado Plateau. *See* GRABEN; HORST; STRUCTURAL GEOLOGY.

George H. Davis

Bibliography. M. P. Billings, *Structural Geology*, 4th ed., 1986; J. G. Dennis, *Structural Geology: An Introduction*, 1987; R. E. Powell and J. C. Matti (eds.), *San Andreas Fault System: Displacement, Palinspastic Reconstruction, and Geologic Evolution*, 1993; F. Press and R. Siever, *Understanding Earth*, 3d ed., 2000; R. J. Twiss and E. M. Moores, *Structural Geology*, 1992; J. T. Wilson, A new class of faults and their bearing on continental drift, *Nature*, 207:343–347, 1965.

Fault-tolerant systems

Systems, predominantly computing systems, telecommunication systems, and other computer-based systems, which tolerate undesired changes in their internal structure or external environment. Such changes, generally referred to as faults, may occur at various times during the evolution of a system, beginning with its specification and proceeding through its utilization. Faults that occur during specification, design, implementation, or modification are called design faults; faults that occur during utilization are referred to as operational faults. Design faults are due to mistakes made by humans or by automated design tools in the process of specifying, designing, implementing, or modifying a system. Operational faults can be physical or human and can occur internally or externally. Examples include component failures (physical-internal), temperature and radiation stress (physical-external), mistakes by human operators who are integral parts of the system (human-internal), and inadvertent or deliberate interference by humans who interact with the system (human-external).

Dependability and performability. Both design faults and operational faults can affect a system's ability to serve its user(s). With respect to a specification of expected service, such ability is referred to as dependability (the ability of a system to deliver the specified service). System service is classified as proper if it is delivered as specified; otherwise it is improper. System failure is identified with a transition from proper to improper service and is a consequence of errors caused by faults. Dependability is quantified by reliability, which is a measure of continuous delivery of proper service, and availability, which quantifies the alternation between deliveries of proper and improper service. More refined assessments of a system's ability to serve are provided by measures of performability (the probability distribution of how well a system performs over a specified period of time). In particular, such measures can account for degraded levels of service that, according to failure criteria, remains proper.

Specific requirements for dependability and performability can vary dramatically according to the nature of the service provided. At one extreme, degraded performance or failure may simply inconvenience the user; in this case, the requirements are relatively modest. At the other extreme, a system failure may result in the loss of human lives; here, continuous proper service is essential and, hence, ultrahigh reliability becomes the dominating requirement. *See* RELIABILITY, AVAILABILITY, AND MAINTAINABILITY.

Fault prevention and fault tolerance. To satisfy a specified set of dependability and performability requirements, faults can be dealt with by prevention or tolerance. Fault prevention is the more traditional method. It attempts to ensure a fault-free system through methods of fault avoidance, which eliminate faults—for example, design methodologies, quality control methods, or radiation shielding; or methods of fault removal, which find and remove faults prior to system utilization—for example, testing and verification. Fault tolerance techniques are based on the premise that a complex system, no matter how carefully designed and validated, is likely to contain residual design faults and is likely to encounter unpreventable operational faults.

To describe how systems tolerate faults, it is important to distinguish the concepts of fault (defined above) and error. Relative to a description of desired system behavior at some specified level of abstraction, an error is a deviation from desired behavior caused by a fault. Since desired behavior may be described at different levels of abstraction—for example, the behavior of computer hardware may be described at the circuit level, logic level, or register-transfer level—a variety of errors can be associated with a given fault. Moreover, a fault can cause an error at one level without causing an error at another level. A fault is latent until it causes an error,

and it is possible for a fault to remain latent throughout the lifetime of a system, thus never causing an error.

The distinction between fault and error is particularly important in the case of design faults. For example, a software design fault (commonly referred to as a bug) may remain latent for years before the software is finally executed in a manner that permits the bug to cause an execution error. The sense in which a system tolerates a fault is typically defined as some form of restriction on the errors caused by that fault. Often, though not always, this term refers to the prevention of errors at a level that represents the user's view of proper system behavior. Errors at this level coincide with system failures and, accordingly, a fault is tolerated if it does not cause a system failure. A system is fault-tolerant if all faults in a specified class of faults are tolerated.

Fault tolerance techniques. A variety of fault tolerance techniques are employed in the design of fault-tolerant systems. Generally, such techniques attempt to prevent lower-level errors (caused by faults) from causing system failures. Using various types of temporal, structural, and informational redundancy, such techniques either mask a fault (no errors are propagated to the faulty subsystem's output) or detect a fault (via an error) and then effect a recovery process which, if successful, prevents system failure. In the case of permanent internal faults, the recovery process usually includes some form of reconfiguration (replacement of a faulty subsystem with a spare; use of an alternate program) which prevents the fault from causing further errors. Many of these methods invoke multiple computations whose results are used to mask faults, via a decision mechanism that produces a single result, or to detect faults through discrimination of differences among their values. This assumes that, with high probability, faults do not cause similar errors among multiple results. Hence, in the case of design faults, simple replication computation is precluded; instead, multiple computations need to be realized by diverse designs which minimize the probability of similar errors.

Typically, a fault-tolerant system design incorporates a combination of fault tolerance techniques which complement the techniques used for fault prevention. The choice of these techniques and, more generally, the specific nature of the system's hardware and software architecture is highly dependent on both types of faults anticipated and the dependability and performability requirements. *See* COMPUTER ARCHITECTURE; INFORMATION THEORY; SOFTWARE ENGINEERING.

John F. Meyer

Bibliography. S.-T. Levi and A. K. Agrawala, *Fault Tolerant System Design*, 1994; J. F. Meyer et al. (eds.), *Dependable Computing for Critical Applications, Two*, 1992; D. K. Pradhar, *Fault-Tolerant Computing: Theory and Techniques*, vol. 1, 2d ed., 1993; D. P. Siewiorek and R. S. Swarz, *Reliable Computer Systems: Design and Evaluation*, 3d ed., 1998.

Feather

A specialized keratinous outgrowth of the skin, which is a unique characteristic of birds. Feathers are highly complex structures that provide insulation, protection against mechanical damage, protective coloration, and also function significantly in behavior. One special functional role is in flight, where feathers provide propulsive surfaces and a body surface aerodynamically suitable for flight. Feathers are used in maintenance of balance and occasionally in the capture of prey and various specialized displays.

Distribution. The number of feathers on birds is determined by a variety of parameters but, in general, large birds have more feathers than small birds. Feather counts vary from 940 on the ruby-throated hummingbird to about 4000 on sparrow-sized birds, 5000 on blackbirds, and more than 25,000 on a whistling swan. There is a significant correlation between body weight and the relative weight of the feathers, so that large birds have proportionately more plumage per surface area than small birds. No significant taxonomic correlation is found in feather number or weight; presumably the relative feather mass is involved in the conservation of metabolic heat. Thus the feather mass of various species is more likely to be influenced by ecological and physiological factors than by taxonomic position.

The uniform distribution of feathers, which appear to cover a bird almost completely, is only superficial. Definitive (as opposed to down) feathers grow from restricted tracts (pterylae) and are absent in the intervening areas (apteria). Number of feathers often varies according to body part, and the head and neck often carry 40–50% of the total number. The pterylae have well-defined patterns which are at least partially specific for different taxonomic groups. The study of these patterns is termed pterylography.

Structure. A representative definitive feather (**Fig. 1**) contains a single long central axis which supports a row of small branchlike structures along each side (barbs). The central structure consists of a hollow, vaneless, cylindrical basal section, called the calamus, and a solid, angular portion, the rachis. The calamus has openings at both ends (the upper and lower umbilicus), the function of which is only incompletely known. The umbilici and calamus are especially well developed in powder down, and the umbilici may be related to the growth of these feathers. Barbs form the vane, or web, of the feather. Individual barbs branch off the rachis at variable angles and point toward the outer tip of the feather.

The barbules are small branches from the barbs. They lie in the same plane as the barbs and arise in rows from their anterior and posterior surfaces. The anterior barbules have a flattened base and a series of small hooklike projections called hamuli; the hamuli attach to the proximal ridge of the posterior barbules of the next barb, forming an interlocking structure characterized by its great strength and light weight (**Fig. 2**). All feather types consist basically of these structural elements. The distribution

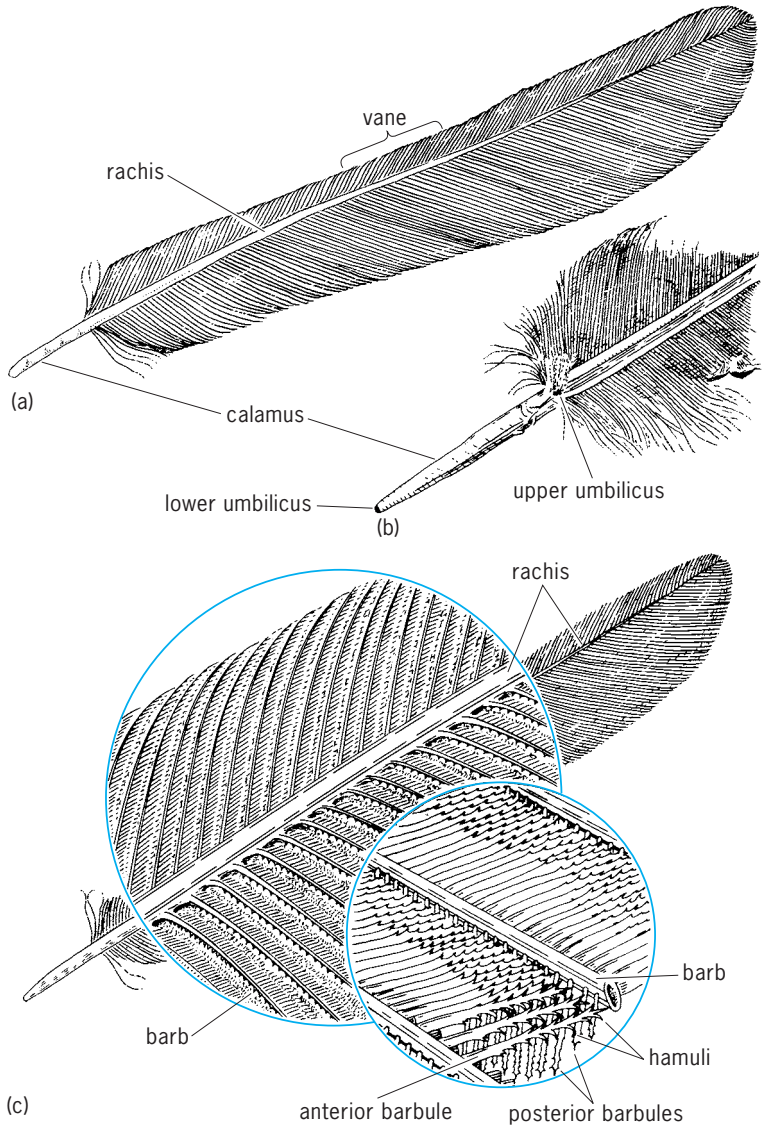

Fig. 1. Typical flight feather. (*a*) Full view. (*b*) Detail of base. (*c*) Detail of vane. (*After J. Van Tyne and A. J. Berger, Fundamentals of Ornithology, 2d ed., John Wiley and Sons, 1976*)

and interrelationships of these elements produce the three-dimensional, macroscopic appearance of the feather.

Specializations. Most of the superficial feathers are contour feathers (pennae). These include the large flight feathers (remiges) of the wing and the long tail feathers (rectrices). Other common feather types include the down feathers (plumulae), intermediate types (semiplumes), and filoplumes (**Fig. 3**). Pennae grow only in pterylae and form the major contours of the body. Semiplumes lack hamuli and thus lack a firm vane. They occur at the periphery of pterylae, integrate structurally with pennae, and serve to provide insulation, bouyancy, and increased flexibility at the base of moving parts. Filoplumes are specialized structures and may be decorative or sensory or both. They are always associated with the papillae of contour feathers. Filoplumes have only a reduced vane at their tip or lack a vane completely. The reduced vane (vibrissa) is hairlike in appearance and presumably has a tactile function. Vibrissae are well developed in the eyelashes of several species and in the stiff (rictal) bristles found at the base of the bill. Down feathers on adults are usually concealed by the contour feathers. On hatchlings of most species the entire feather complement consists of downy feathers. The degree of nestling feathers (neossoptiles) ranges from sparse to dense and, in many cases, is related to chick behavior and the ecology of the nest. Most natal downs lack a rachis and are not confined to pterylae. Their primary function is insulative. Powder down, found only in certain taxonomic groups, are continually growing feathers. They produce a soft, friable material used in the cleaning and maintenance of other feathers.

Naked, or featherless, areas of adult birds serve several functions. For example, combs, wattles, bare eye rings, casques, and other structures on or near the head are used in sexual and mating displays; parts of the legs and feet may be used in the regulation of body heat, or for the protection, by shading, of the eggs and young; the naked head and neck in carrion feeders may serve a sanitary function.

Cytology. Adult feathers are essentially dead structures. However, developing feathers consist of three rather distinct cell types. These form the cuticle, cortex, and medulla, the morphologically distinct elements of the feather. The proportions of these elements vary within any particular type of feather. In the medulla, especially, the relationship between inter- and intracellular spaces may vary and influence the appearance of the feather, particularly its color. Cells of the developing cortex show high cytoplasmic ribonucleic acid (RNA) activity, as occurs in cells undergoing rapid proliferation and protein synthesis. *See* RIBONUCLEIC ACID (RNA).

Chemistry and molecular structure. Feathers are ectodermally derived proteinaceous material. The main protein, β-keratin, has a molecular weight of approximately 11,000 and is characterized by its general insolubility, a result of the large numbers of intramolecular disulfide bonds. Although the amino acid composition of feather parts may differ, serine, glycine, cystine, and proline are the most frequent residues. *See* PROTEIN.

The solubilized proteins from feathers are highly heterogeneous upon separation by electrophoresis.

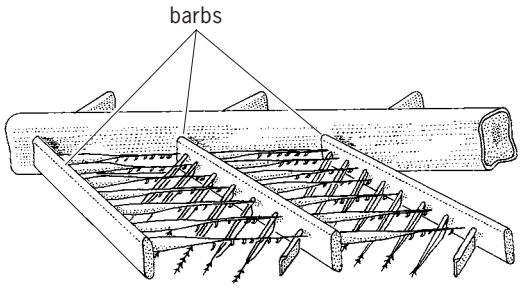

Fig. 2. Parts of three barbs from a vaned feather, showing the hooking mechanism of the overlapping barbules. (*After J. C. Welty, The Life of Birds, Saunders, 1962*)

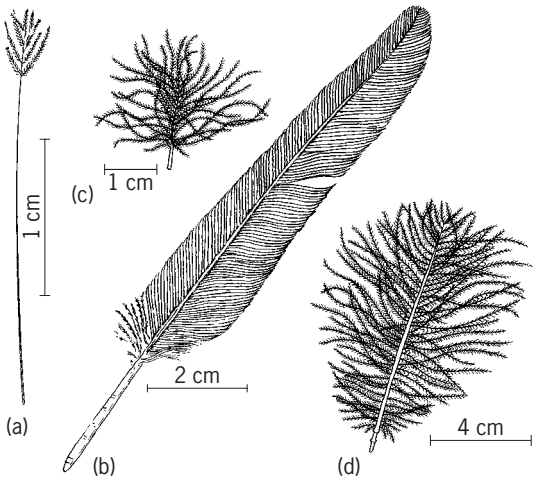

Fig. 3. Types of feathers, (*a*) Filoplume. (*b*) Vane or contour. (*c*) Down. (*d*) Semiplume. (*After J. C. Welty, The Life of Birds, Saunders, 1962*)

Combined with differences in physiochemical and immunological properties, the complexity indicates the presence of a large number of monomeric, individual gene products. This interpretation is supported by the existence of species-specific keratin polypeptides and experiments with messenger RNA (mRNA) and complementary deoxyribonucleic acid (DNA) which indicate the presence of over 200 keratin genes in the total genome. Differences in molecular protein structure and distribution reflect the interaction of structural and regulatory gene products. In certain genetic mutants the abnormal gross feather structure is the result of the same interactions.

The complete amino acid sequence is known for keratin monomers from two species. Physical analysis shows each monomer, or gene product, to consist of a folded and an amorphous portion. Sequence differences appear limited to the latter. The folded portions interact between units to form larger, highly organized, molecular assemblies. The folded portion may also contain the immunogenetic site of the protein. Thus, structural conformation is related to an evolutionary conservativeness. The individual polypeptides form small microfibrils. These, in turn, are organized into filaments which are cemented together in an amorphous matrix.

Development and molt. Feather development is a complex process and has been studied in detail only in the contour and flight feathers of domestic birds. The penna develops initially as a domelike elevation of the epidermis that has a dermal core. Each developing feather grows outward by a rapid proliferation of cells at its base (**Fig. 4**). A tapered epidermal cylinder is produced with a dermal core that contains vessels which provide nutrients to the developing structure. Subsequent feather development depends on accurate alignment, in both time and space, of rows of epidermal cells to form the rachis, barbs, and barbules. This is especially true in the highly structured adult contour and flight feathers, in which growth is adjusted for the subunits of the vane to hook together properly. In addition, specialized feather structures, such as the tips on the feathers of cedar waxwings and similar structures in other species, are produced by modification of the same basic mechanism.

Histogenesis. As development proceeds, three epidermal areas can be distinguished. The outer layer, which forms a sheath; a thick intermediate layer, representing the primordia of the rachis and barb, which later becomes keratinized; and an inner layer, which forms an envelope about the pulp. All three layers are derived from a collarlike ring of embryonic epidermal cells. In all cases a remnant of the original dermal cells, surrounded by epidermal cells, remains at the papilla. This is necessary for formation and production of subsequent generations of feathers. In newly hatched chicks the epidermal sheath ruptures as it dries, allowing feather parts to unfurl. As outward growth proceeds, the pulp is simultaneously withdrawn toward the feather base, producing a completely keratinized, nonliving structure.

Endocrine regulation. Many aspects of feather growth are dependent on hormones. Evidence exists which suggests that thyroid hormones are immediately responsible for molt by acting directly on the feather papilla, but there is a great deal of interspecific difference in the quality and timing of response to thyroxine administration. Pigmentation is, in varying degrees, determined by thyroid, sex, and gonadotrophic hormones. Hormone imbalances, especially in reference to thyroid secretion, can cause changes in the pattern and melanin content of feathers, as well as affect the timing of the molt.

Periodicity of molt. Feathers normally undergo attrition because of the physical abuse attendant to the normal activity of birds. In most species, feathers are replaced completely at least annually, and many of the feathers are replaced more frequently. The actual molt involves an exquisite integration of the sequence and relative timing of a series of events and

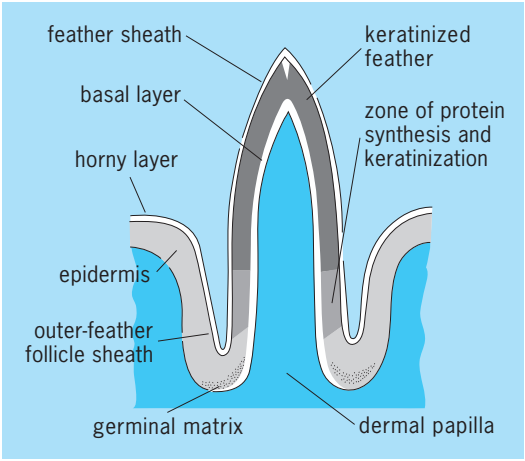

Fig. 4. Section through growing feather follicle. (*After R. I. C. Spearman, The keratinization of epidermal scales, feathers, hairs, Biol. Rev., 41:59–69, 1966*)

is clearly of adaptive value. Considerable energy is expended in molt, in the biosynthesis of new keratin, in the associated changes in thermal conductance, and in associated behavioral changes. For this reason molting is usually scheduled so as not to interfere with other energy-demanding phenomena of the annual cycle, most notably reproduction and migration.

Molt patterns. The sequence of feather molt is surprisingly orderly. Penguins, which shed large patches of feathers in an irregular pattern, are an exception. In most species the power of flight is retained during molt. The molt, that is, the normal shedding of feathers and their replacement by a new generation of feathers, is a single growth process which is actively concerned only with the production of the new generation of feathers. The old feathers are pushed out of the follicles passively.

Coloration. The patterns and coloration in feathers are controlled by mechanisms active during their formation. Obviously, hormonal and nutritional conditions are important during this period. Less well understood are the systems which produce and control the repetition and register so apparent in the plumage pattern. Such patterning involves the control of follicle spacing, feather angle, and the initiation and termination of the pigmentation process in individual feathers. Superficial patterns of this type provide the basis for species recognization.

There are basically two types of feather colors: schemochromes, which are colors that originate within the feather structures; and biochromes, which are colors that are caused by the presence of true pigmentation.

Schemochromes. The schemochromes are derived from morphological variations in the rachis, barbs, and barbules. Pigments are not necessary for the production of structural colors, but they are generally present and modify the colors considerably. The simplest case of structural coloration is whiteness, which is the result of almost complete reflection and refraction of incident light by the minute, colorless, transparent feather surfaces, interstices, and air-filled spaces. Noniridescent blues, such as in the blue jay, are another common structural color produced when a fine, colorless framework of barbs overlies a dark layer which contains melanin granules. The colorless layer reflects the blue wavelength, while other wavelengths pass through this layer and are absorbed by the pigment below. Other noniridescent colors, such as various greens or olive, are also associated with the presence of pigments in the absorbing layer. These include both melanins and carotenoids. *See* TYNDALL EFFECT.

Iridescent feathers, such as are found on hummingbirds, sunbirds, and the rock dove, are produced independently of pigmentation by the interference of light by a thinly laminated structure in the barbules. Barbs of iridescent feathers are invariably broad and flat with blunted ends, distinctly pigmented, and generally lacking barbules. These broad surfaces often overlap, and the feather surface then acts as a diffraction grating. Colors are caused by the interference of light waves reflected from the surfaces of extremely thin, colorless laminae or film-like plates which compose the outer surfaces of the barbules. The various colors are produced by differences in the angle of the surfaces and changes in the angle of light incident to the surface.

Biochromes. The most common biochromes are melanins. Melanins are synthesized from the amino acid tyrosine in specialized cells (melanophores) and are deposited as discrete particles or granules. There are probably no consistent chemical or structural differences between the eumelanins and phaemelanins even though they produce slightly different colors. Melanins are responsible for a wide range of colors, from dull yellow through red-brown to dark brown and black. *See* PIGMENTATION.

Carotenoids. The most spectacular colors in feathers are produced by the carotenoid pigments. Carotenoids are found mainly in the apical portion of the barbs, but in intensely pigmented feathers the entire barb may be pigmented. In certain cases intense pigmentation is accompanied by morphological changes in the feather, such as flattening of the barb in one plane and the absence of barbules. This provides a better exposure of the pigment. Early workers reported the absence of carotenoids in the feathers of young birds, but more evidence for such statements is necessary. *See* CAROTENOID.

As birds cannot synthesize carotenoids, they depend on their food for these pigments. Feather carotenoids are generally highly oxidized forms (commonly keto or hydroxy derivatives) of ingested pigments. Thus, birds obviously possess the metabolic capability of modifying these pigments. Accumulating evidence indicates that these pathways may be common to many groups of birds. Elucidation of the metabolic pathways for the production of feather carotenoids may be fruitful in understanding both the evolution of such pathways and the birds which possess them.

Color variability regulation. Many of the seasonal displays of feather pigments, especially in the males of migratory songbirds, depend on carotenoids. These deposits, as well as the pigments responsible for sexual dichromatism, presumably are controlled by hormonal mechanisms. In the brightly colored scarlet tanager, for example, the biochemical difference between the feathers of the males and the females, and between the colors of breeding and nonbreeding males, represents only a relatively small difference in the chemistry of the pigment molecule. This difference produces an enormous visual effect, however, and is of great significance to the biology of this species. The chemical differences in the feather pigments of the sexes and the seasonal differences in the males may be mediated by a single enzyme, determined by a single gene. Conceivably, the activity of this gene could be controlled by a single hormone. Additional studies which link pigment chemistry, diet, genetics, and natural history will further understanding in this area. Other genetic effects on pigments include the absence of pigments in normally pigmented individuals

(albinism) or an overabundance of melanins (melanism) or carotenoids (xanthophyllism). *See* SEXUAL DIMORPHISM.

Porphyrin compounds. An unusual pigment has been isolated from members of the touraco family (Musophagidae). This pigment occurs in two forms in feathers, the red turacin and a green, oxidized derivative, turacoverdin. Chemically, this pigment is a copper-containing uroporphyrin. A related pigment, but one lacking copper, has been found in the bustards (Otididae). Free porphyrins (coproporphyrin) have also been reported in the down feathers of owls.

Adherent colors. Another type of coloration, produced by neither internal pigmentation nor structure, is the adherent colors. They are chemicals, commonly iron oxides, that are applied, presumably unintentionally, to feathers during preening.

Function. Many functional aspects of feathers are dependent on their posture and position. Feather posture is adjusted chiefly by bands of nonstriated muscle fibers in the dermis (m. arrectores plumarum). Typically each pennaceous feather follicle is connected by pairs of muscles to four adjacent follicles. The depressor and erector muscles in each pair provide fine control over the position of individual feathers. Receptors in the skin near the feather may sense pressure or vibration of the feather and thereby provide information for further adjustment of position.

Thermoregulation. A major physiological role of feathers is to provide insulation. This is accomplished by regulating the configuration of feather and skin in such a way that differing amounts of air are trapped in the dead space so formed. A second mechanism for control of heat dissipation is the balance of the exposure of feathered and unfeathered body parts. *See* THERMOREGULATION.

Waterproofing. Feathers act as a protective boundary in their role of providing waterproofing. Water repellency is a structural feature of feathers and is the result of precise geometric relationships between the diameter and spacing of barbs and barbules. Preening appears to be more important in the maintenance of this structure than it is for the application of oils or any other natural product, as was once thought.

Behavior. A third function of the surface configuration and overall pattern of feathers is in the area of behavioral adaptations. These may be of two types. First is concealment, when the bird is cryptically marked to match its background and escape detection. Other types of concealment include disruptive patterns, which break the familiar outline of the bird, and mimicry, a rare situation when one species resembles another. The second type consists of various types of advertisement. For example, unique combinations of feather color and pattern may label the species as inedible, serve to maintain contact among members of a flock in a gregarious species, or provide a means of attracting and stimulating members of the opposite sex. *See* PROTECTIVE COLORATION.

Evolution. Feathers presumably derived from scales or scalelike structures present on the reptilian ancestor of birds. Chemical and structural evidence indicates the presence of similar proteins in these structures. The form and process of derivation of the primitive feather are unknown. However, there is some agreement that primitive feathers were rather generalized in structure and arose in conjunction with the metabolic and thermoregulatory demands of avian precursors. This generalized feather, in turn, gave rise to the more specialized downy and contour plumages. *See* SCALE (ZOOLOGY). Alan H. Brush

Bibliography. A. H. Brush (ed.), *Chemical Zoology*, vol. 10, 1976; A. H. Brush (ed.), *Symposia on Feather Structure: Proceedings of the 16th International Ornithological Congress*, *Canberra*, *1974*, 1976; A. M. Lucas and P. B. Stettenheim, *Avian Anatomy*: *Integument*, pts. 1 and 2, USDA Handb. 362, 1972.

Feedback circuit

A circuit that returns a portion of the output signal of an electronic circuit or control system to the input of the circuit or system. When the signal returned (the feedback signal) is at the same phase as the input signal, the feedback is called positive or regenerative. When the feedback signal is of opposite phase to that of the input signal, the feedback is negative or degenerative.

The use of negative feedback in electronic circuits and automatic control systems produces changes in the characteristics of the system that improve the performance of the system. In electronic circuits, feedback is employed either to alter the shape of the frequency-response characteristics of an amplifier circuit and thereby produce more uniform amplification over a range of frequencies, or to produce conditions for oscillation in an oscillator circuit. It is also used because it stabilizes the gain of the system against changes in temperature, component replacement, and so on. Negative feedback also reduces nonlinear distortion. In automatic control systems, feedback is used to compare the actual output of a system with a desired output, the difference being used as the input signal to a controller. These two points of view will be considered separately in the following discussion. However, the analysis of both feedback amplifiers and electromechanical control systems can be made on a common basis; from the point of view of analysis the two have much in common.

Amplifier feedback. Feedback can be introduced into an amplifier by connecting to the input a fraction of the output signal. An amplifier will, in general, have better frequency-response characteristics when the system has feedback than when there is no feedback. The system can be designed to have a wider bandwidth and more nearly ideal frequency-response characteristics. Further, harmonic distortion due to nonlinear transistor

or tube characteristics can be reduced by using feedback. *See* DISTORTION (ELECTRONIC CIRCUITS).

The use of feedback in an improperly designed system, however, can produce a system with worse characteristics. Amplifiers can become oscillators when feedback is used in an improperly designed system. Furthermore, feedback is not always intentional, and amplifiers, particularly radio-frequency amplifiers, are often unstable because of unintentional (parasitic) feedback paths in the circuit. For a discussion of amplifier frequency response and bandwidth *see* AMPLIFIER

A system with feedback can be shown in block diagram form (see **illus.**). The sinusoidal input signal is $\mathbf{E}_s$ and the amplifier gain is $\mathbf{A}$, which is a function of frequency. When there is no feedback, $\mathbf{E}_o = \mathbf{A}\mathbf{E}_s$ because $\mathbf{E}_i = \mathbf{E}_s$. When there is feedback, $\mathbf{E}_o = \mathbf{E}_s + \mathbf{E}_f$.

Since $\mathbf{E}_o = \mathbf{A}\mathbf{E}_i$, and $\mathbf{E}_f = \beta\mathbf{E}_o$, the overall gain of the system with feedback is then shown by the equation below.

$$\frac{\mathbf{E}_o}{\mathbf{E}_s} = \frac{\mathbf{A}}{1 - \mathbf{A}\beta}$$

This formula for the gain of an amplifier with feedback indicates the effect of feedback upon the frequency response of the amplifier without feedback. The $(1 - \mathbf{A}\beta)$ term in the denominator is a complex number. Therefore, the magnitude and phase angle of the gain of the amplifier with feedback will differ from the gain of the amplifier without feedback. The amount of the difference depends upon the value of the $\mathbf{A}$ and the β terms, and no general statements can be made.

If there is a frequency for which $\mathbf{A}\beta = 1$, the denominator of the expression for the gain will be zero while the numerator will not be zero. When this occurs, the amplifier will oscillate at approximately the frequency for which $\mathbf{A}\beta = 1$. Furthermore, if there is a frequency for which the magnitude of $\mathbf{A}\beta$ is greater than unity and the phase angle of $\mathbf{A}\beta$ is 0°, 360°, or any integral multiple of 360°, the amplifier will oscillate.

Positive and negative feedback. The terms positive feedback and negative feedback are used to denote the type of feedback found in certain electronic circuits. *See* NEGATIVE-RESISTANCE CIRCUITS.

When the magnitude of the denominator of the feedback equation is greater than unity, the overall gain with this (negative) feedback will be less than the gain with no feedback, and stable operation will result. If the magnitude of $\mathbf{A}\beta$ is much greater than one, $\mathbf{A}/(1 - \mathbf{A}\beta)$ is approximately $-1/\beta$ and does not depend on $\mathbf{A}$, thus minimizing the effect of unstable, inaccurate elements on the circuit. This result is the principal benefit of negative feedback.

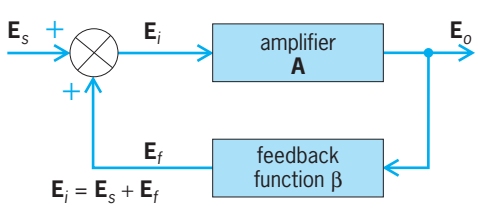

Block diagram of feedback circuit.

If the magnitude of the denominator in the displayed equation is less than unity, then the overall gain with this positive feedback will be greater than the gain without feedback. Under these circumstances the circuit may be unstable and either oscillate or drive its output to the limit of the power-supply voltage level and remain there.

It follows that both the magnitude and the phase angle θ of the product $\mathbf{A}\beta$ are important in determining whether the feedback is positive.

The effect of feedback upon the frequency response of an amplifier can be determined from the expression $\mathbf{A}/(1 - \mathbf{A}\beta)$. The performance in a particular case depends upon the behavior of $\mathbf{A}$ and β as functions of frequency.

Analysis and design. There are several methods for analyzing and designing feedback amplifiers. In those cases where the $\mathbf{A}$ and β terms in the previous equation can be identified from the circuit, the possibilities of oscillation due to the feedback can be determined by plotting the amplitude and phase of $\mathbf{A}\beta$ over the frequency range from zero to a sufficiently high frequency. The examination is often made by plotting the magnitude of $\mathbf{A}\beta$ against the phase angle in polar coordinates. It may be done also by plotting the magnitude of $\mathbf{A}\beta$ in decibels and the phase of $\mathbf{A}\beta$ against the logarithm of the frequency for a sufficiently wide frequency range. The second method also indicates whether the amplifier will be satisfactory for audio applications.

There are many cases where the circuit is sufficiently complex that more powerful tools are required. The simple block diagram of the illustration is not applicable. Instead, the Laplace transform method is used to obtain the transfer function (the output transform divided by the input transform) of the system. The poles and zeros of the transfer function define the system's behavior. Typically, the system is unstable if there are poles in the right half of the s plane (where s is the Laplace transform variable). The use of the computer greatly reduces the tedium of the design process by displaying on a monitor pole-zero plots on the s plane as parameters are varied, the frequency response, the transient response, and other items of interest. The possibility of poles in the right half of the s plane can also be determined by using the Routh-Hurwitz criteria. *See* CONTROL SYSTEM STABILITY; LAPLACE TRANSFORM.

Computer-aided design is particularly helpful in the design process because circuit components such as transistors, resistors, and capacitors can have a significant variation around their nominal values. For example, resistors with a 20% tolerance might be used because they are inexpensive. The amplifier can be described by using the state-variable method in which a system of first-order differential equations for the circuit is determined. The aforementioned properties of stability, response, and so forth, can be determined. *See* LINEAR SYSTEM ANALYSIS.

The computer also provides great help in determining the effects of adding compensating networks to improve the frequency response, the effects of different feedback configurations, and so forth. Thus, the time needed for breadboarding and experimenting is reduced. However, the ultimate test is the performance of the final circuit since the computer prediction is only as good as the equations that model the circuit. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Input and output impedances. The above example (see illus.) shows voltage feedback where a function of the output voltage is fed back to the input. Current feedback, where a voltage proportional to the output current or some function of the output current is fed back, may also be used. Furthermore, an amplifier may have both voltage and current feedback.

The input and output impedances of a feedback amplifier depend upon whether the output current or voltage is sampled and upon whether this signal is fed back in series or in parallel with the input excitation. It is possible for the impedance with feedback to be greater than or less than what it was before the feedback was added. An emitter-follower is an example of a feedback amplifier where the output impedance has been reduced to a very low value and the input impedance has been increased considerably.

Oscillator feedback. An oscillator can be viewed as a feedback amplifier with an input voltage supplied from its output. Referring to the illustration, this would mean that $\mathbf{E}_s$ equals zero and $\mathbf{E}_i$ equals $\mathbf{E}_f$.

From this viewpoint the condition for oscillation at a frequency f_o is (to a first and usually very good approximation) that $\mathbf{A}\beta = 1$ at $f = f_o$. This means that the feedback must be positive.

Servomechanism feedback. The purpose of feedback in a servomechanism is to obtain better control of the system. As an example, consider a position control system which is used to position an object at a point determined by a reference signal.

The command to move the object is derived by comparing the reference signal with a signal indicating the instantaneous location of the object. The command signal is an error signal derived from the comparison of the actual and desired signals; any error signal drives the system in such a direction as to reduce the error, that is, to make the actual position agree with the desired reference position.

If feedback were not used in the position control system, a precisely calibrated control device would be needed to position the object in each position dictated by the reference. In general the required control could not be built with sufficient accuracy. *See* CONTROL SYSTEMS; SERVOMECHANISM.

Harold F. Klock

Bibliography. P. M. Chirlian, *Analysis and Design of Integrated Electronic Circuits*, 2d ed., 1987; J. D'Azzo and C. Houpis, *Linear Control System Analysis and Design*, 3d ed., 1988; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Feeding mechanisms (invertebrate)

The feeding methods of invertebrates are as diverse as the invertebrates themselves, which are adapted to all kinds of habitats, in freshwater, in the sea, and on land. The feeding mechanisms are best classified by the method used: browsing, suspension feeding, deposit feeding, carnivorous, and phytophagous (plant-eating). An alternative classification often adopted, but perhaps less satisfactory, may be based on the size of particles ingested. Thus, the same invertebrate may be described either as microphagous (feeding on minute organisms) or as dependent on substances in solution. Both classification schemes may be further subdivided. Carnivorous feeders, for example, include predators and animal parasites; both share dependence on other (living) animals as a food source. Some methods will be restricted to particular habitats. Suspension feeders, for example, can only be aquatic, while the phytophagous habit may be found wherever edible plants occur. The feeding mechanisms employed will depend, at least partly, on the habit and habitat of the plants utilized.

The insects are by far the most abundant in species and variety. Their feeding mechanisms depend on modification of the mouthparts to suit different foods and on adaptation of the gut to cope with different diets. The mouthparts comprise three pairs of appendages borne on the head—the mandibles, the maxillae, and the labium, each capable of adaptation for biting, piercing, lapping, or sucking. Different adaptations are shown by the mouthparts of, say, a locust for biting off pieces of living plants, a mosquito for piercing mammalian skin and sucking the blood from dermal capillaries, an aphid greenfly for tapping plant sap, and a housefly for lapping fluids, but all may be referred to the same basic plan. Both aphids and locusts are phytophagous in that they depend on plants for food, but their feeding mechanisms are entirely different. *See* INSECTA.

Browsers. The gastropod mollusks afford good examples of browsing invertebrates. The buccal mass houses a rasping radula that acts as a food grater, removing the layer of encrusting organisms (mainly algae) from the rocks. The rasp lines of limpets can often be seen on algae-encrusting rocks in the intertidal zone. The chitons also have a radula and feed by browsing. The radula is a narrow band of tissue bearing rows of minute chitinous teeth; it is produced continuously to replace the anterior part as it wears away. The different dental arrangements are characteristic of each group. The radula is manipulated by a complex system of muscles, for the radula serves not only for rasping the food off the rocks but for conveying it into the buccal cavity. *See* MOLLUSCA.

Carnivores. It is not only the more active invertebrates such as squids and octopus or the planktonic arrowworms (Chaetognatha) that pursue prey; jellyfish and sea anemones catch the animals when they brush against them, and eat them. Nor is it only the larger invertebrates that are predatory. The ciliate

Didinium is voracious in attacking other ciliate protozoa.

The cnidarians are characteristically armed with stinging cells (nematocytes) containing nematocysts (capsules with eversible threads) that may both paralyze and catch fishes and other animals that touch them. The same basic mechanism is adapted by jellyfishes, by the pelagic Portuguese man-of-war (*Physalia*), anemones, and, on a smaller scale, the individual hydranths of colonial hydroids and the solitary freshwater *Hydra*. There is great variety and specialization in the different types of nematocysts for stinging and capture. The comb jellies (Ctenophora) "fish" in the same manner, though their cells (collocytes) are different. After the food is secured (by adhesion), muscular contraction of the tentacles serves to bring the prey to the mouth. *See* CNIDARIA; CTENOPHORA.

Few annelid worms are predators, apart from the leeches, though some polychaetes do prey on other invertebrates, which they catch by means of their jaws and eversible proboscis. Nemertine ribbon worms feed on other animals, mostly when dead or moribund, by protrusion and withdrawal of their long proboscis. Advanced leeches (such as the medicinal leech) pierce the skin of the host by means of three bladelike jaws edged with minute teeth, sucking the blood which is prevented from clotting by anticoagulants in the saliva. Such predators (or parasites) take large meals at infrequent intervals and have large crops for storage. Some leeches such as *Trochaeta* ingest whole earthworms, and some freshwater leeches ingest other small invertebrates. Mollusks such as *Urosalpinx* prey on barnacles, gastropods, or bivalves, boring through the shells by a combination of chemical and mechanical means. The process is often lengthy and laborious but is rewarded by a substantial supply of food. Preliminary digestion may be external, the partially digested tissue being sucked up by means of a tubular proboscis. *See* ANNELIDA; MOLLUSCA.

Some sea stars use their tube-feet to pull apart the valves of mussels, the star everting part of its stomach between the valves to digest and absorb the meat. *See* ECHINODERMATA.

Scavengers. Many invertebrates are scavengers, depending on animal or plant remains, or on animals in a moribund condition. Many crustaceans are scavengers, including isopods and amphipods; and many of the larger crabs also fall into this category. The multiple mouthparts of such animals are adapted to carry out different parts of the feeding process. Large pieces of food may be grasped by pincerlike (chelate) appendages and passed via other pairs of appendages that comminute the food into fragments of a size suitable for ingestion. *See* CRUSTACEA.

Suspension feeders. Feeding on particles in suspension is particularly widespread among intertidal invertebrates, for the action of the waves and the high productivity of coastal waters ensure a relatively abundant supply of detritus and living organisms. Some filter feeders have remarkably efficient filtering mechanisms (**Fig. 1**), and the adoption of suspension feeding is often combined with a sedentary life. The mechanisms employed may be subdivided into setose, ciliary, and mucous-bag feeding.

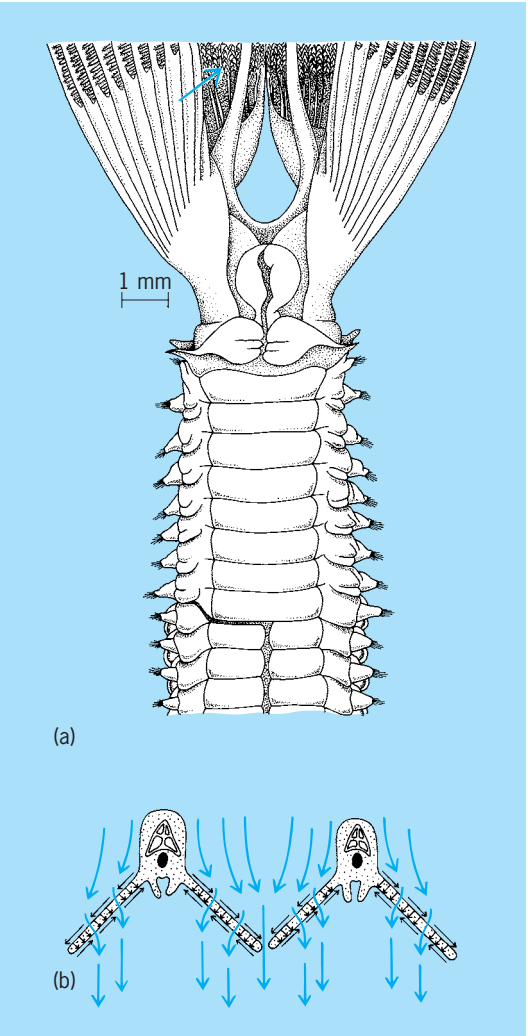

Fig. 1. ***Sabella pavonina*. (*a*) Ventral view of base of tentacular crown and first 12 body segments. (*b*) Diagrammatic transverse section of two tentacles, as indicated by arrow on *a*, to show the direction of the flow of water entering the crown (colored arrows) and the direction of the ciliary beat on the pinnules (black arrows). The longitudinal food groove of each tentacle is overarched by expanded bases of pinnules; particles from pinnules enter it between these bases. (*After V. Fretter and A. Graham, A Functional Anatomy of Invertebrates, Academic Press, 1976*)**

Setose appendages. Suspension feeding by means of setose appendages is characteristic of many crustaceans such as acorn barnacles, which secrete a limy carapace and use six pairs of appendages (the cirri) to strain particles from the water and direct them to the mouth. In strong currents the cirri need only be held out into the current, the particles are caught on the setae, and the cirri are contracted from time to time to convey the catch to the mouth. Some crabs, such as the porcelain crabs, also depend on particles in suspension. Various appendages participate in the filtering process, but in general setose extensions of mouthparts or maxillipeds are used. Some sand crabs

and amphipods have some appendages adapted for filtering and others adapted for dealing with larger particles so that they can utilize a wide range of food as opportunity permits.

Ciliary mechanisms. The fan worms such as *Sabella* depend on currents drawn through a crown of pinnate filaments by long cilia along the edges of the filaments, particles being trapped by mucus produced by glandular cells and conveyed toward the mouth along defined tracts by motion of shorter cilia. The cilia beat in particular directions so that a stream of trapped particles converge from each pinnule to the base of the filament, where they are sorted according to size. Unsuitable particles are rejected by other ciliary currents sweeping them away from the crown (Fig. 1).

Some intertidal gastropods have adapted their respiratory currents into feeding currents and their gills into ciliary filters, a mechanism that may have evolved from the necessity not only of keeping the gills irrigated but of keeping them clean. Gastropods such as the slipper limpet, *Crepidula fornicata*, have therefore become sedentary, for by transferring the gill cleanings to the mouth there is no necessity for locomotion. The Chinaman's-hat limpet (*Calyptraea*) feeds in a similar manner. So does the tall *Turritella* to a certain extent, but since much of its food is sucked up from the surface of the substratum, it is principally a deposit feeder.

Clams and oysters are the most efficient filterers in that the water is passed through the gills, the particles being sorted mainly by size on special flap-shaped palps by the mouth. The enormous mussel (*Mytilus*) banks on rocky coasts bear testimony to the amount of food in suspension in such places. Mussels avoid being swept away by secreting threads (byssus), but rock oysters and many other filterers cement themselves to rocks and pilings in order to take advantage of these nutritive currents.

The protochordate sea squirts have achieved the same type of feeding by other means. The filter is a branchial basketlike elaboration of pharyngeal gill slits. The feeding current is taken in through the mouth and passed through the slits and out through the atrium and exhalant aperture (**Fig. 2**). The food is trapped in a mucous string produced from a ventral gutter (endostyle), the mucus flowing across the basket to the esophagus. Like many other filterers, the sea squirts have become sedentary.

The bryozoans, common below tide on many shores, have developed a filtering crown of tentacles, the lophophore, superficially like that of hydroid cnidarians; and they are also sedentary and colonial. The bryozoan inhalant current is directed by cilia down into the center of the crown and out between the bases of the tentacles, the food particles being directed along prescribed paths to the mouth. The solitary phoronids are similar; and the brachiopod lamp shells, so numerous in earlier times, have a lophophore shaped like a double horseshoe of numerous slender filaments within the lower valve of the Roman lamp–shaped shell. The mechanism is functionally like that of a mussel, but the construction is entirely different. *See* BRACHIOPODA; BRYOZOA; PHORONIDA.

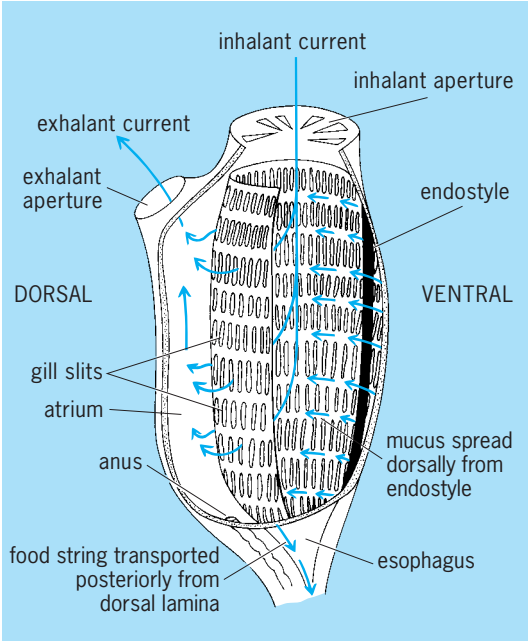

Fig. 2. Diagrammatic sagittal section of *Clavelina* showing the branchial basket and feeding currents. The side of the tunic and part of the branchial basket are cut away. (*After R. C. Newell, Biology of Intertidal Animals, Marine Ecological Surveys Ltd., Faversham, Kent, England, 1979*)

Mucous bag. Secretion of mucus, denied to the crustaceans but an indulgence in soft-bodied invertebrates, is a common means of trapping food. In a few animals the mechanism is more complex. In general, nets or bags of mucus are spun from the mouth; when the device is full, both the net and the catch are eaten. The net is filled by a water current that may be actively provided by muscular or other means, or the feeder may depend on the natural current. Some insect larvae use this method in freshwaters, while the annelid *Chaetopterus* and the echiuroid *Urechis* do so in the sea. Mucous-bag filters trap minute particles. *Chaetopterus* (**Fig. 3**) secretes a parchmentlike tube whereas *Urechis* lies in a burrow, but both make a current through a secreted mucous bag. *Chaetopterus* produces its current with muscular balers or "fans," *Urechis* by peristalsis of the body wall. In *Chaetopterus* the bag is subtended by two armlike extensions that secrete it, the bag rolled up in a cup-shaped structure on the segment behind, and the bolus passed forward along a groove to the mouth.

Integumentary absorption. Some invertebrates absorb all their nutrients through the integument. This is true not only of many parasites bathed in the nutritious media of their hosts' body fluids or gut contents but also of the pogonophoran annelids, which live in cold deep-water marine deposits and have neither mouth nor gut. Slow and lethargic, they live in permeable horny tubes through which their nutrients are actively absorbed from solution. Some other invertebrates are capable of utilizing

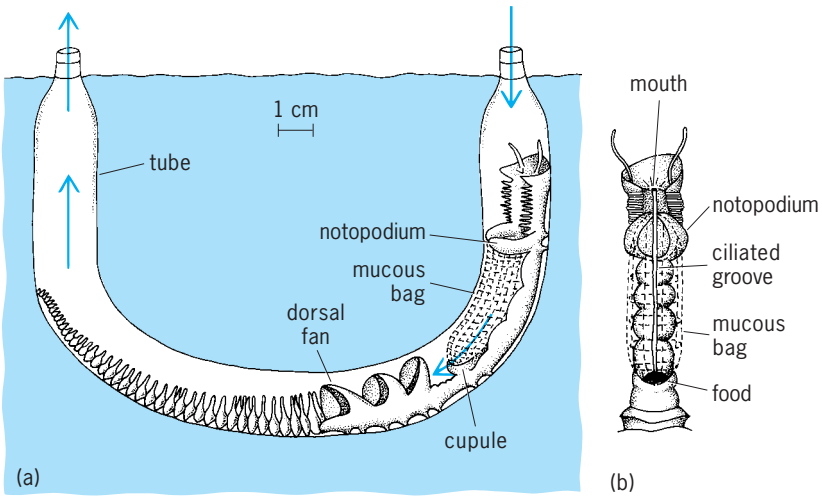

Fig. 3. ***Chaetopterus*. (*a*) Feeding with a mucous bag, the arrows indicating water current. (*b*) Dorsal view of the anterior end. (*After V. Fretter and A. Graham, A Functional Anatomy of Invertebrates, Academic Press, 1976*)**

dissolved substances, amino acids or complex carbohydrates, by active uptake. This form of nutrition has often been overlooked, but it may form a small but significant proportion of the energy budget. *See* POGONOPHORA.

Deposit feeders. Many invertebrates feed on deposits of mud and sand. The nutritional content may be relatively low, but it is usually highest at the surface where there is a microflora of photosynthetic organisms and where detritus is also deposited. There are various mechanisms for collecting the most nutritious surface layer, and guts of such feeders are specialized for coping with large volumes of deposit continuously delivered.

In the intertidal zone the polychaete annelids afford a variety of examples of deposit-feeding mechanisms. The lugworm (*Arenicola*) lies at a depth of 8–16 in. (20–40 cm) in an L-shaped burrow, with the head in the toe of the L, and feeds on sand. As it eats, sand caves in from above, forming hollows on the surface. From time to time, the worm backs up the vertical shaft to form the castings of sand that has passed through the gut. The cycles of eating and defecation continue with clockwork regularity, for there is little to disturb the even tenor of its ways. Other polychaetes living beneath the surface, or in crevices for protection, have developed tentacles that can be extended onto the surface. *Terebella* is one such, with numerous extensible tentacles arising, medusalike from the head, each provided with muscles and a ciliated groove to assist in conveying morsels to the mouth. Other invertebrates, such as ampharetid worms and sea cucumbers (*Holothuria*), have shorter sticky tentacles that are licked clean of adherent particles by the mouth. *See* POLYCHAETA.

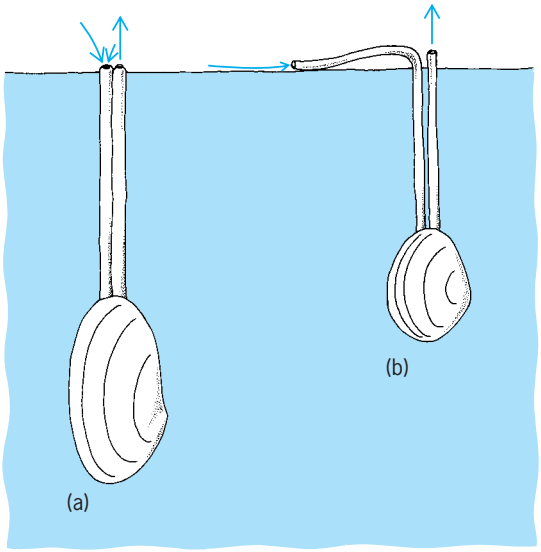

Fig. 4. Use of siphons in (*a*) a suspension-feeding bivalve (*Mya arenaria*) and (*b*) a deposit-feeding bivalve (*Scrobicularia plana*). Arrows indicate the inhalant and exhalant streams. (*After R. C. Newell, Biology of Intertidal Animals, Marine Ecological Surveys Ltd., Faversham, Kent, England, 1979*)

Most bivalves which are not filterers ingest the surface deposits by means of the inhalant siphon, which acts rather like a domestic vacuum cleaner. The suspension-feeding *Mya* lies well buried in the mud, with inhalant and exhalant siphons externally joined together and extending to the surface (**Fig. 4***a*). The deposit-feeding *Scrobicularia* has the inhalant and exhalant siphons separated, enabling the inhalant tube to search a wide area of surface while the rest of the animal remains at a safer depth (Fig. 4*b*). Many gastropods, like the tiny *Hydrobia*, feed on surface mud. The larger pelican's-foot snail (*Aporrhais*) has an elaborate mechanism for rasping off organisms from the deposits which are taken in from the surface by means of an inhalant current. Both inhalant and exhalant currents are produced by the proboscis and consolidated by the mucus it secretes. As in *Arenicola*, the surface is eaten while the animal remains concealed. *Corophium*, an amphipod that burrows in fine marine mud, takes the rich surface into the opening of the burrow by means of its long second antennae. The beating of the more posterior appendages (pleopods) creates a current that carries the raked-up material into a setose sorting area near the mouth.

In terrestrial environments there are, of course, no filter feeders, and deposit feeders live mainly in the soil to avoid light and water loss as well as predators. Terrestrial vegetation provides food for browsers such as slugs and snails and for numerous insect larvae with mouthparts adapted to deal with particular plants. Plant litter in various stages of decay provides food for other invertebrates both on and beneath the surface. Such invertebrates have not only mouths and teeth to deal with the food but also gut floras to assist digestion. Cellulose and wood are often broken down with the aid of gut flagellates. *See* DIGESTION (INVERTEBRATE). R. Phillips Dales

Bibliography. R. S. K. Barnes, P. Calow, and P. J. W. Olive, *The Invertebrates: A Synthesis*, 3d ed., 2001; R. C. Brusca and G. J. Brusca, *The Invertebrates*, 2d ed., 2003; P. A. Meglitsch, *Invertebrate Zoology*,

3d ed., 1991; V. Pearse et al., *Living Invertebrates*, 1989; C. L. Prosser, *Comparative Animal Physiology*, 4th ed., 1991; C. M. Yonge, Feeding mechanisms in the invertebrates, *Biol. Rev.*, 3:21–70, 1928.

Feldspar

Any of a group of aluminosilicate minerals whose crystal structures are composed of corner-sharing $[AlO_4]$ and $[SiO_4]$ tetrahedra linked in an infinite three-dimensional array, with charge-balancing cations [primarily sodium (Na), potassium (K), and calcium (Ca)] occupying large, irregular cavities in the framework of the tetrahedra. Collectively, the feldspars constitute about 60% of the outer 8–10 mi (13–17 km) of the Earth's crust. They are nearly ubiquitous igneous and metamorphic rocks, and are a primary constituent of arkosic sediments derived from them. The importance of the many feldspars that occur so widely in igneous, metamorphic, and some sedimentary rocks cannot be underestimated, especially from the viewpoint of a petrologist attempting to unravel earth history. *See* ARKOSE; MINERALOGY; PETROLOGY; SILICATE MINERALS.

With weathering, feldspars form commercially important clay materials. Economically, feldspars are valued as raw material for the ceramic and glass industries, as fluxes in iron smelting, and as constituents of scouring powders. Occasionally their luster or colors qualify them as semiprecious gemstones. Some decorative building and monument stones are predominantly composed of weather-resistant feldspars. *See* CLAY MINERALS; IGNEOUS ROCKS; METAMORPHIC ROCKS.

Chemical Composition

The general formula AT_4O_8 characterizes the chemistry of feldspars, where T (for tetrahedrally coordinated atom) represents aluminum (Al) or silicon (Si). The A atom is Ca^{2+} or barium (Ba^{2+}) for the $[Al_2Si_2O_8]^{2-}$ alkaline-earth feldspars and Na^+ or K^+ for the $[AlSi_3O_8]^-$ alkali feldspar series of solid solutions and mixed crystals. A complete range of compositions is observed in the plagioclase feldspar series, $Na_yCa_{1-y}Al_{2-y}Si_{2+y}O_8$ ($0 \leq y \leq 1$), and in a somewhat analogous, though rare, $K_xBa_{1-x}Al_{2-x}Si_{2+x}O_8$ series called hyalophanes. The $BaAl_2Si_2O_8$ end member is called celsian (symbolized Cn).

The majority of natural feldspars have the approximate formula

$$K_xNa_yCa_{1-(x+y)}Al_{2-(x+y)}Si_{2+(x+y)}O_8$$

where $0 \leq (x+y) \leq 1$, and x is equal to the mole fraction of the potassium feldspar end member (Or, for orthoclase), y is equal to the mole fraction of the sodium feldspar (Ab, for albite), and $1-(x+y)$ is equal to the mole fraction of the calcium feldspar (An, for anorthite).

Knowledge of a feldspar's composition (**Fig. 1**) and its crystal structure is indispensable to an understanding of its properties. However, it is the distribution of the Al and Si atoms among the available tetrahedral sites in each chemical species that is essential to a complete classification scheme, and is of great importance in unraveling clues to the crystallization and thermal history of many igneous and metamorphic rocks.

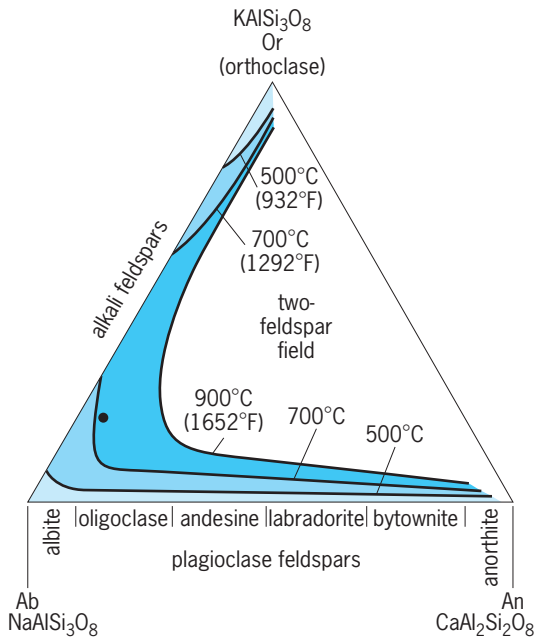

Fig. 1. The feldspar Or-Ab-An ternary as determined at a water pressure of 1000 atm (6.9 MPa). The solid circle represents a feldspar quenched above 700°C (1290°F) whose composition is $Or_{20}Ab_{75}An_5$ or $K_{0.20}Na_{0.75}Ca_{0.05}Al_{1.05}Si_{2.95}O_8$. The range of feldspar compositions is shown as a function of increasing temperature in the successively less shaded areas.

Crystal Structures

The structures of all feldspars share similar characteristics. The primary structural units are four-membered rings of AlO_4 and SiO_4 tetrahedra linked to similar rings by sharing certain oxygen atoms (**Fig. 2**). In an idealized concept (Fig. 2*b*), crankshaft-like chains are shown as extending through the crystal parallel to the *a* unit cell dimension, and they occur in pairs that are mirror images (or approximate mirror images) of one another, across the (010) crystallographic plane. Parallel to the *b* unit cell dimension, the chains are joined together by the T-O-T, bonds, and in the *c* direction slabs of chain pairs are joined by T-O-T bonds through oxygen.

Large A cations are located between the (001) slabs; Na^+ and Ca^{2+} are bonded to six or seven oxygen atoms and K^+ to nine oxygen atoms. The prominent (010) and (001) cleavages of feldspars are expressions of the weakness of the A-O bonds.

Complete disorder of Al and Si is virtually unknown in natural potassium feldspars. At lower temperatures, Al increasingly concentrates in T_1 sites and Si in T_2 sites, but this ordering process is limited by the number of Al atoms per formula unit. If all the Al is to end up in a single T_1 site (which in nature is the

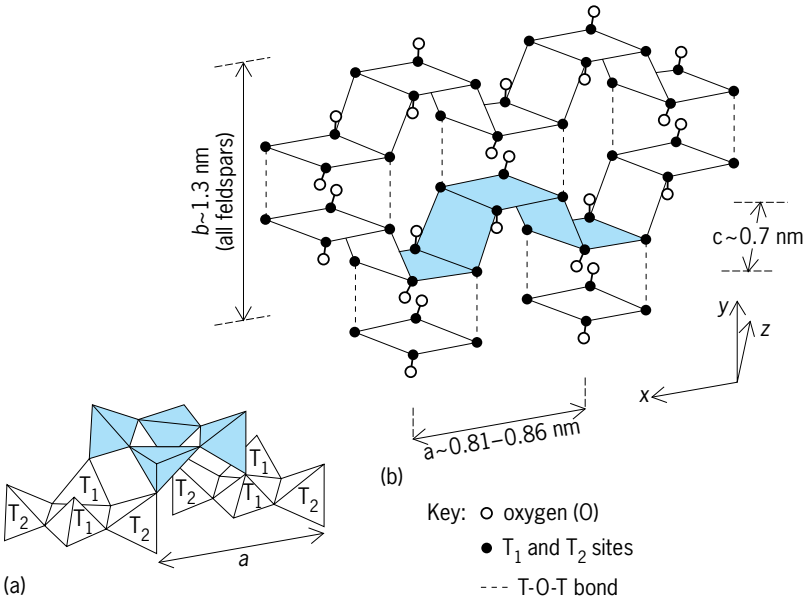

Fig. 2. Feldspar structure. (*a*) Perspective view showing the linkage of $Al0_4$, and SiO_4 tetrahedral units, T_1 and T_2, into fourfold rings; oxygen atoms are at the corner of the tetrahedra. The shaded tetrahedral ring is related to the unshaded ones by an *a*-glide plane parallel to (010). (*b*) Expanded and idealized representation, showing both T_1 and T_2 sites as black dots. These are a pair of mirror- or pseudomirror-related crankshaftlike chains. The shaded area is that portion of this drawing that is projected onto the (010) plane in Fig. 3*a*.

configuration with the lowest free energy for alkali feldspars), the T_1 sites must become symmetrically nonequivalent (**Fig. 3*b***). An O suffix is arbitrarily assigned to the T_1 site into which A1 is concentrated; the other site is called T_1m because in the feldspar structure it is related to T_1O by a pseudomirror (m) operation. The T_2 sites are likewise differentiated.

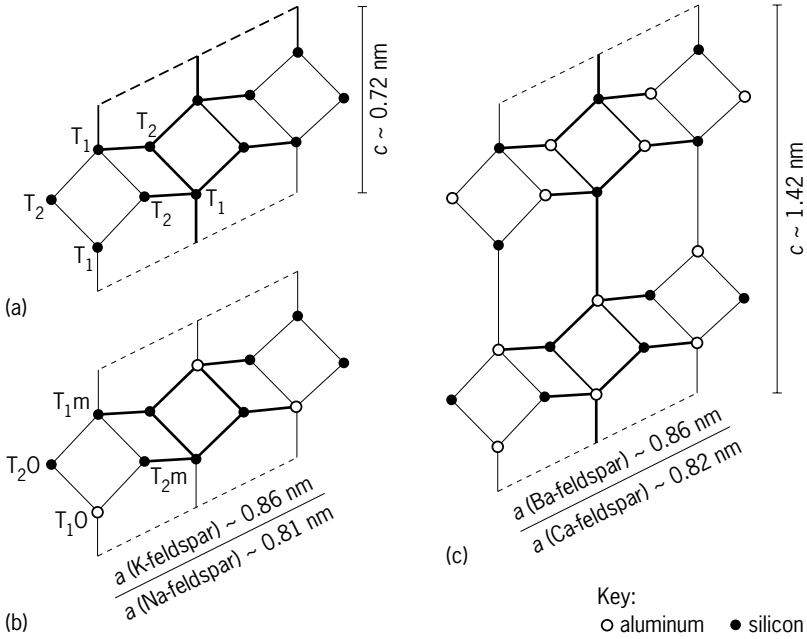

Fig. 3. Feldspar structures showing the shaded portion of Fig. 2*b* projected onto (010). (*a*) Monoclinic sanidine and monalbite. (*b*) Triclinic low albite and low microcline [Al ordered into the T_1O sites]. (*c*) Monoclinic celsian or triclinic anorthite [note the doubled *c* cell dimension]. The angle between the *a* and *c* unit cell edges is called β; $\beta \sim 116^\circ$ in all feldspars.

There is distortion of the fourfold ring, and the monoclinic space group (C2/m) feldspar becomes triclinic ($C\bar{1}$).

Feldspars containing more than a few percent of divalent cations may have very complex crystal structures or exsolve into two phases, or both, primarily because A1:Si ratios depart from the 1:3 ratio of the relatively simple alkali feldspars discussed above. In nature, the Al and Si atoms of the celsian and anorthite end members alternate regularly in the tetrahedral framework. Such an ordered Al,Si distribution (Fig. 3*c*) requires that the *c* cell dimension be doubled to about 1.4 nm from the approximately 0.7-nm characteristic of the C2/m and $C\bar{1}$ feldspars (Fig. 3*a*,*b*). Celsian is monoclinic (space group, I2/c). Anorthite, whose A site is occupied by a much smaller atom than that in celsian [radius of Ca $\simeq$ 0.10 nm, that of Ba $\simeq$ 0.14 nm], has a collapsed triclinic framework (space group $I\bar{1}$ or $P\bar{1}$). *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Nomenclature

The feldspar minerals are broadly divided into two classes, alkali feldspars and plagioclase feldspars (**Tables 1** and **2**).

Alkali feldspars. Names are assigned to the polymorphs of $KAlSi_3O_8$ and $NaAlSi_3O_8$ in accordance with their symmetry and the Al content of their tetrahedral sites. *See* ALBITE; ORTHOCLASE; MICROCLINE.

The names pericline and cleavelandite apply to particular morphological varieties of albite, and adularia is a variety of K-rich feldspar. Anorthoclase is

TABLE 1. Alkali feldspars, in order from high- to low-temperature forms

Potassium-rich feldspars	
Monoclinic (space group C2/m)	
High sanidine	$0.5 \leq 2t_1 \leq 0.67$
Low sanidine	$0.67 < 2t_1 < 0.75$
Orthoclase	$0.75 \leq 2t_1 \leq 1.00$
Triclinic (space group C1)	
Intermediate microcline	$1.0 < t_1o > t_1m > t_2o = t_2m$
Low microcline or maximum microline	$t_1 \simeq 1.0$; $t_1m = t_2o = t_2m \simeq 0.0$ (fully ordered)
Sodium-rich feldspars	
Monoclinic (space group C2/m)	
Monalbite	$0.5 \leq 2t_1 \leq \sim 0.6$ (exists only at temperatures above 980°C or 1800°F)
Triclinic (space group C1)	
Analbite	$0.5 \leq 2t_1 \leq \sim 0.6$ (the lower-temperature form of monalbite)
High albite	$t_1o > t_1m \geq t_2o = t_2m$ (highly disordered; all *t* values near 0.25)
Intermediate albite	$t_1o > t_1m \geq t_20 = t_2m$ (probably not stable in nature)
Low albite	$t_1o \simeq 1.0$; $t_1m = t_2o = t_2m \simeq 0.0$ (fully ordered)

TABLE 2. Plagioclase feldspars

Name	Composition
Sodic (acid) plagioclases	
Albite	An_0–An_{10}
Oligoclase	An_{10}–An_{30}
Andesine	An_{30}–An_{50}
Calcic (basic) plagioclases	
Labradorite	An_{50}–An_{70}
Bytownite	An_{70}–An_{90}
Anorthite	An_{90}–An_{100}

a triclinic solid solution of composition $Or_{37}Ab_{63}$–Or_0Ab_{100} containing up to 10 mol % anorthite, or more. *See* ANORTHOCLASE.

Ternary feldspars are those containing at least 10% of all three end members; in nature most anorthoclases and ternary feldspars will have exsolved into two discrete phases.

Plagioclase feldspars. Nomenclature of the plagioclases is based on their compositions (Fig. 1). Plagioclases containing significant amounts of exsolved K-rich feldspar are called antiperthites. It is only in once-molten rocks quenched at very high temperatures that the full range of so-called high plagioclases exist as simple solid solutions. With very slow cooling over millions of years, complex textures develop in most feldspar crystals as a coupled NaSi, CaAl ordering. *See* ANDESINE; BYTOWNITE; LABRADORITE; OLIGOCLASE; SOLID SOLUTION.

Properties

The variable properties of feldspars are determined by their structure, symmetry, chemical composition, and crystallization and subsequent history of phase transformation, exsolution, and alternation or deformation.

Color. Very few feldspars are transparent and colorless; many are white or milky due to internal reflections of light from inclusions, exsolution interfaces, and fracture or cleavage surfaces. In fact, the name for the sodium feldspar is albite (from the Latin *albus*, white). Fine-scale alteration to sericite or clay produces a cloudy appearance, and if tiny dispersed crystals of hermatite (Fe_2O_3) are present, a cloudy pink to orange-brown to brick-red color results. It is thought that black-clouded feldspars (most often calcic plagioclases) are caused by minute, dispersed iron oxides or iron-titanium oxides. Yellow feldspars contain minor Fe^{3+}. Blue to green feldspars (including amazonite) contain Pb^+ and structurally combined H_2O. Smoky feldspar, like smoky quartz, results from ionizing radiation.

Oriented intergrowths resulting from exsolution produce a variety of effects due to scattering, reflection, and interference of incident light. Aventurine and sunstone are feldspars in which thin flakes of hematite have exsolved in a particular crystallographic orientation and reflect light as glittering rosy or gold schiller (play of color); copper platelets produce pink schiller. Sometimes, oriented mica flakes or metal oxide platelets and needles cause a silvery sheen or chatoyancy. Exsolution textures in cryptoperthitic alkali feldspars may result in an ethereal blue to whitish appearance. These feldspars are known as moonstones. In the three regions of exsolution in the plagioclase subsolidus, the resulting two-phase intergrowths are essentially planar. Light interference, like that from an oil slick on water, produces brilliant red, yellow, green, and blue colors.

Mechanical properties. Plagioclases are slightly harder (6–6.5) on Mohs scale than K-rich feldspars (6). Feldspars are brittle and, when broken, cleave along the (001) and (010) crystallographic planes. Both cleavages are parallel to the crankshaftlike chains (Fig. 2), and both represent planes of weak A–O bonding. Quartzlike conchoidal fracture is occasionally seen in crystals that have a glassy appearance.

Morphology. Feldspars that have grown relatively unimpeded in fluids or cavities usually have the following primary crystal faces: $\{110\}$, $\{1\bar{1}0\}$, $\{010\}$, $\{001\}$, $\{20\bar{1}\}$, and $\{10\bar{1}\}$; faces of lesser importance are $\{130\}$, $\{1\bar{3}0\}$, $\{1\bar{1}\bar{1}\}$, and $\{1\overline{11}\}$ (**Fig. 4**). The principles controlling crystal growth and favoring the development of one or more crystal faces over others are complex.

Twinning. As a group, no minerals have more complex and ubiquitous twins than feldspars. If only two (or a few) individuals are observed, the twins are called simple twins (**Fig. 5***a*, *b*). Complex twins result when two twin laws with the same composition plane are simultaneously operative in a single crystal; Albite-Carlsbad is a relatively common example in plagioclases. Albite and Pericline twins are nearly always multiple or polysynthetic twins (Fig. 5*c*, *d*), and they are observed only in triclinic feldspar. The (010) composition planes of albite twins often show up as straight lines (actually shallow grooves) on (001) faces (Fig. 5*c*). Traces of Pericline twins are seen on (010) cleavage planes as parallel straight lines (Fig. 5*d*), inclined to the trace of the (001) cleavage.

In practice, all types of twins are most easily observed and identified with a petrographic microscope, which is widely used in the study of rocks and minerals. Genetically there are three main categories of twins: (1) growth twins, (2) inversion or transformation twins, and (3) deformation or mechanical twins. *See* PETROGRAPHY.

Density. Densities are measured in terms of mass per unit volume. Densities are near 2.56 g/cm^3 for

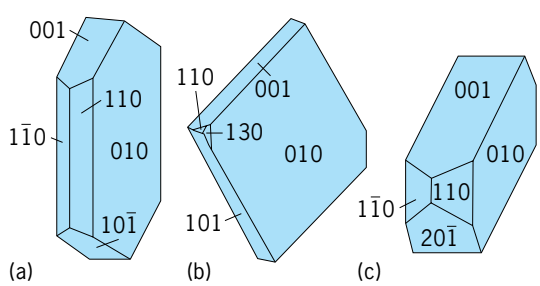

Fig. 4. Various feldspar morphologies: (*a*) blocky, typical of microcline; (*b*) platy, typical of quenched feldspars called microlites; (*c*) prismatic, typical of orthoclase.

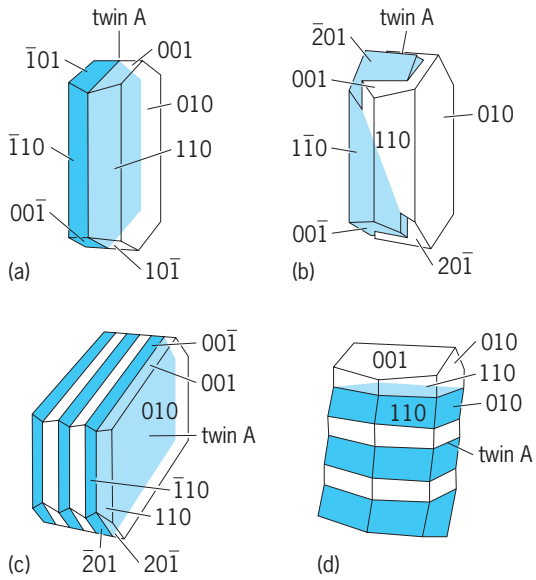

Fig. 5. Simple and multiple twins in feldspars. (*a*) Carlsbad (contact), (*b*) Carlsbad (interpenetrant), (*c*) albite, (*d*) pericline. (*After T. Zoltai and J. H. Stout, Mineralogy: Concepts and Principles, Burgess, 1984*)

K-feldspars and vary regularly with composition between 2.62 g/cm^3 for Na-feldspar and 2.78 g/cm^3 for Ca-feldspar. Ba-feldspar has a density of 3.37 g/cm^3.

Optical properties. Feldspars are optically biaxial, exhibiting three refractive indices whose values and vibration directions within the crystal can be diagnostic of composition and degree of Al,Si order-disorder.

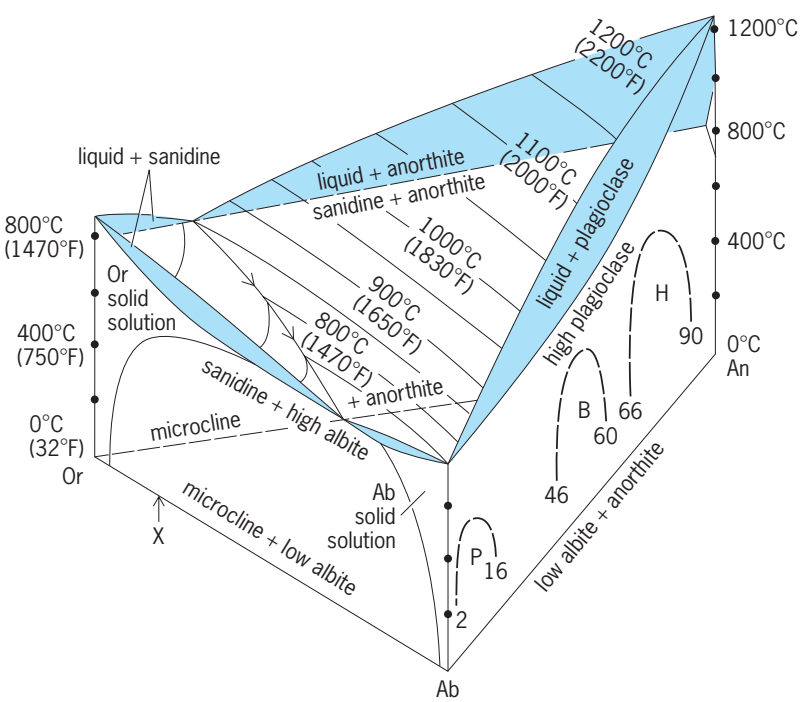

Fig. 6. Temperature-composition diagram for the Or-Ab-An system at a water pressure near 5000 kg · cm^{-2} (490 MPa). Details of the interior of this pseudo phase diagram are complex and have been omitted. P = peristerite intergrowth. B = Bøggild intergrowth. H = Huttenlocher intergrowth. Numbers at intergrowths represent approximate ranges (in mol % An) over which exsolution occurs. Temperatures are given in degrees Celsius (Fahrenheit).

Although these optical properties are relatively complicated, they are widely and routinely used, together with extinction angles of twins, by geologists in their study of the history of crystallization, metamorphism, or mechanical deformation of rocks in which these ubiquitous feldspars occur.

Melting properties. In a dry (water-free) system K-rich feldspar melts incongruently, forming a mixture of leucite ($KAlSi_2O_6$) and silica-rich liquid above about 1075°C (1967°F); albite and anorthite melt congruently at 1118°C (2044°F) and about 1545°C (2813°F), respectively, forming glasses when rapidly quenched. When water is present, the melting temperatures of feldspars are lowered substantially.

Phase Equilibria of the Feldspar Binary Systems

Several isotherms on the Or-Ab-An composition triangle at about 1000 atm (6.9 MPa; Fig. 1) have been determined experimentally. At or above 500°C (930°F), feldspar compositions exist as single phases, that is, as crystalline solid solutions. At or above 700°C (1300°F), all compositions from the 700°C (1300°F) isotherm out to the limiting triangle are single phases, and likewise for the 900°C (1650°F) isotherm. At 900°C (1650°F) or below, bulk compositions in the two-feldspar field are immiscible and form two chemically distinct phases. Phase relations of the K-Na-Ca feldspars are better understood with reference to a diagram (**Fig. 6**) of the three binary systems, Or-An, Or-Ab (alkali feldspars), and Ab-An (plagioclases), versus temperature for an approximate water pressure of 5000 kg cm^{-2} (490 MPa). In such a diagram the top surface (the liquidus) is contoured for temperatures at which a particular composition becomes completely molten. Individual binary diagrams give information on the range of temperatures and compositions at which a crystalline phase is in equilibrium with a liquid. At temperatures below the solidus, only crystalline (solid) phases exist; depending on temperature and composition, there may be a single homogeneous phase or two crystalline solid-solution phases in equilibrium. *See* PHASE EQUILIBRIUM.

System $KAlSi_3O_8$-$CaAl_2Si_2O_8$. A phase diagram, known as a binary eutectic, has been determined experimentally in the laboratory, but it is of little practical importance in nature because bulk compositions in this range do not occur in igneous rocks.

System $KAlSi_3O_8$-$NaAlSi_3O_8$. This system is of considerable importance geologically, and alkali feldspars have been studied in great detail. It has been impossible to duplicate experimentally the textures and structures of many of these feldspars which, in nature, may have had many millions of years to attain their present state. Because of the extremely sluggish kinetics of Al,Si migration in the solid state, true equilibrium assemblages of feldspars in rocks are by no means the rule. However, using combinations of laboratory synthesis and detailed analysis of natural feldspars, petrologists have reached a fairly sophisticated level of understanding of the K-Na system (Fig. 6).

In the pure Or-Ab binary, the structural state and composition of the phase or phases observed will depend on three parameters: (1) the initial bulk composition, (2) the temperature at which the material is observed (or from which it is rapidly quenched to room temperature to facilitate analytical studies), and (3) the annealing conditions, including in particular the rate at which the material has been cooled. Below the solvus the initially homogeneous single crystal begins to exsolve, separating in the solid state into two feldspars. The resulting composite feldspar is called perthite. *See* PERTHITE.

Both the Or and Ab solid-solution phases change not only in composition but also in structural state with slow cooling. Increasingly ordered Al,Si distributions are to be expected (Table 1).

System $NaAlSi_3O_8$-$CaAl_2Si_2O_8$. Crystallization from a liquid in the pure binary system is relatively simple, but in nature, spurts of cooling and reheating, as well as local fluctuations of magma composition (including H_2O and other volatiles), will more than likely produce zoned plagioclase grains. Petrologists often analyze the compositional variation of such feldspars—together with other mineralogical and textural variables—in attempts to decipher the crystallization histories of the rocks in which they occur. *See* MAGMA.

Once all the fluid phases in a rock system are exhausted, zoning patterns may persist throughout geologic time, as long as the rock is not thermally metamorphosed. But whether chemically zoned or homogeneous, the relatively disordered high-plagioclase crystals that form just below the solidus are subject to further structural and textural modification in the solid state.

An Ab-An phase diagram (Fig. 6) can be simplified and yet is adequate to serve as a guide to a discussion of the nature of most plagioclases that are found in igneous and metamorphic rocks.

In nature, most plagioclases in the range An_2 to An_{16} have exsolved into two phases: one is nearly pure low albite, the other is an oligoclase of composition $Ab_{80\pm5}An_{20\mp5}$. The resultant lamellar intergrowths are called peristerites whether or not they display a white, pale blue, or occasionally brilliant range of spectral colors. The colors are due to interference of light reflected from multiple thin films of albite and oligoclase. Paul H. Ribbe

Bibliography. W. L. Brown (ed.), *Feldspars and Feldspathoids: Structures, Properties and Occurrences*, 1984; I. Parsons (ed.), *Feldspars and Their Reactions*, 1994; P. H. Ribbe (ed.), *Feldspar Mineralogy*, 2d ed., vol. 2 of *Reviews in Mineralogy*, 1993; J. V. Smith, *Feldspar Minerals*, vols. 1 and 2, 1974, rev. 1986.

Feldspathoid

A member of the feldspathoid group of minerals. Members of this group are characterized by the following related features: (1) All are aluminosilicates with one or more of the large alkali ions (for example, sodium, potassium) or alkaline-earth ions (for example, calcium, barium). (2) The proportion of aluminum relative to silicon, both of which are tetrahedrally coordinated by oxygen, is high. (3) Although the crystal structures of many members are different, they are all classed as tektosilicates; that is, they consist of a three-dimensional framework of aluminum and silicon tetrahedrons, each unit of which shares all four vertices with other tetrahedrons, giving rise to a ratio of oxygen to aluminum plus silicon of 2:1. Ions of other elements occupy interframework sites. (4) They occur principally in igneous rocks, but only in silica-poor rocks, and do not coexist with quartz (SiO_2). Feldspathoids react with silica to yield feldspars, which also are alkali–alkaline-earth aluminosilicates. Feldspathoids commonly occur with feldspars.

The principal species of this group are the following:

Nepheline	$KNa_3[AlSiO_4]_4$
Leucite	$K[AlSi_2O_6]$
Cancrinite	$Na_6Ca[CO_3\|(AlSiO_4)_6] \cdot 2H_2O$
Sodalite	$Na_8[Cl_2\|(AlSiO_4)_6]$
Nosean	$Na_4[SO_4\|(AlSiO_4)_6]$
Haüyne	$(Na,Ca)_{8-4}[(SO_4)_{2-1}\|(AlSiO_{46}]$
Lazurite	$(Na,Ca)_8[(SO_4,S,Cl)_2\|(AlSiO_4)_6]$

The last four species (sodalite group) are isostructural, and extensive solid solution occurs between end members; but members of the sodalite group, cancrinite, leucite, and nepheline have different crystal structures.

Nepheline is an essential constitutent of intrusive nepheline syenites and nepheline syenite pegmatites (as in the extensive deposits of the Kola Peninsula, Russia) and of their extrusive equivalents, phonolites. It also occurs in rocks of metamorphic character, having originated by metamorphism of silica-poor source rocks or, more commonly, by metasomatic replacement of rocks such as marbles and gneisses (as near Bancroft, Ontario, Canada). Cancrinite occurs principally with, and as a replacement of, nepheline in plutonic rocks, having been derived by the reaction of preexisting nepheline with $CaCO_3$-containing solutions. Sodalite also occurs principally in association with nepheline in syenites and related silica-poor intrusive igneous rocks, but it may also occur in equivalent extrusive rocks. Both haüyne and nosean are found almost exclusively in extrusive, silica-deficient rocks such as phonolites. Leucite occurs in potassium-rich basic lavas, frequently in large phenocrysts having the form of trapezohedrons. Pseudoleucite is a mixture of feldspar and nepheline which has the crystal form of preexisting leucite; it occurs in both extrusive and intrusive rocks. Unlike most other members of the feldspathoid group, lazurite occurs in metamorphic rocks, specifically, metamorphosed impure limestones. Lapis lazuli is the name of a blue lazurite-rich rock which is used as a gem material. *See* CANCRINITE; LAZURITE; NEPHELINITE; SILICATE MINERALS.

Donald R. Peacor

Feline infectious peritonitis

A fatal disease of both domestic and exotic cats (particularly cheetahs) caused by feline infectious peritonitis virus, a member of the Coronaviridae family. There are multiple strains of the virus which vary in virulence. Feline infectious peritonitis virus is closely related morphologically, genetically, and antigenically to other members of the Coronaviridae and arises as mutants from feline enteric coronaviruses, which infect cats but generally induce very mild or inapparent gastroenteritis. All coronaviruses are single-stranded ribonucleic acid (RNA) viruses with poor or no error correction during replication, resulting in relatively high mutation rates.

Epidemiology and transmission. Feline coronaviruses are contracted primarily during exposure to infectious cat feces in the environment, but also via ingestion or inhalation during cat-to-cat contact. Feline infectious peritonitis virus is relatively labile once outside the cat's body but may be able to survive for as long as 7 weeks if protected from heat, light, and desiccation. It is readily inactivated by most disinfectants.

Cases of feline infectious peritonitis occur either sporadically or in outbreaks with 50–70% mortality in homes with more than one cat. Most sporadic cases occur when a cat is infected with a feline enteric coronavirus which then mutates within the individual cat into the pathogenic feline infectious peritonitis phenotype. Cats with feline infectious peritonitis may shed either benign enteric coronavirus, potentially fatal feline infectious peritonitis virus, or neither. Many cats do not succumb to feline infectious peritonitis even after prolonged extensive exposure to infected cats.

In contrast to the sporadic transmission of feline infectious peritonitis, the transmission of the parental feline enteric coronavirus is frequent and is difficult to detect clinically. Prevalence of infections with enteric coronavirus is related to cattery size and density: in homes with five or more cats, approximately 100% of the cats have been exposed. Following exposure (often as kittens), cats may periodically shed enteric coronavirus in feces for weeks or a few months. After recovery, cats are not immune to the virus. Thus, they are likely to become rapidly reinfected and to shed again. Approximately 5% of kittens become chronic shedders after infection. Very few of the infected cats will show any clinical signs of enteritis; however, 2–5% of the cats may experience mutation to the feline infectious peritonitis phenotype and develop clinical signs of the virus. The likelihood of the mutant form developing and not being rapidly eliminated by the cat's immune system depends upon whether the cat is immunosuppressed (as by concurrent feline leukemia or feline immunodeficiency virus infection) and overall viral load. It is not known exactly at what point following infection with enteric coronavirus that the mutation occurs. The peak age distribution for development of feline infectious peritonitis is 6 months to 3 years of age.

Pathogenesis. In contrast to feline enteric coronavirus, whose replication is largely restricted to intestinal epithelial cells, feline infectious peritonitis virus targets feline macrophages (white blood cells). The virus can persist and replicate within the macrophages and spreads systemically through the bloodstream and into tissues throughout the body. The basis for development of clinical disease associated with feline infectious peritonitis depends upon a balance between antibodies and cell-mediated immunity (comprising T lymphocytes, natural killer cells, macrophages, and other innate immunity). The presence of antibodies to the virus enhances the access of the virus into the macrophage and may produce accelerated disease, although this phenomenon appears to be restricted to laboratory strains and infection protocols. Virus-antibody complexes are also deposited in the walls of small blood vessels, where they activate the complement system, attracting other blood vessels and initiating the profound inflammatory lesions that characterize the disease.

Clinical signs. Signs of infection with feline infectious peritonitis virus depend upon the severity of infection, the relative ability of the immune system to minimize some of the lesions, and the organ systems affected. If antibody-virus complexes are diffusely distributed along small blood vessels in the abdomen, the resulting immune-mediated sequelae may produce vasculitis which results in protein and fluid leakage out of the blood vessels and into the abdominal cavity (or less frequently thoracic and scrotal spaces). This form of disease is called wet feline infectious peritonitis and occurs in cats with overwhelming infection, with poor immunity, or during late stages of other forms of the virus. In contrast, if a cat has a moderately competent (but ultimately ineffectual) immune response to the virus infection, granulomas may arise in infected tissues which attempt to localize virus with an influx of uninfected macrophages, neutrophils, and lymphocytes. This dry form commonly affects kidneys, liver, lymph nodes, mesentery, diaphragm, the outer surface of the intestine, and the neurological system.

Cats with either wet or dry feline infectious peritonitis may exhibit weight loss, fever (often cyclical), and lethargy. Disease associated with a particular organ system reflects impairment or failure of that system, such as vomiting associated with liver disease, and seizures and ocular opacity associated with neurological or ocular feline infectious peritonitis. Wet or effusive feline infectious peritonitis is characterized by abdominal swelling, jaundice, and typically difficulty in breathing. Signs of upper respiratory disease or diarrhea are not directly associated with the virus. Suspicion of feline infectious peritonitis is based upon consistent history and signalment (typically cats originally from multiple-cat households which are 1–3 years old) and presence of consistent clinical signs. Suggestive laboratory testing includes elevated serum globulins, any positive coronavirus titer (which will not, however, discriminate

between feline infectious peritonitis virus and feline enteric coronavirus), and mild elevations of neutrophils in blood and mild reductions in lymphocytes. The only accurate testing is biopsy of lesions or necropsy.

Treatment and control. As with most viral infections, there is no specific antiviral drug of proven efficacy in the treatment of feline infectious peritonitis. Clinical management rests upon palliative treatment of the specific signs, exhibited by each cat, and upon antibiotics, when indicated, to reduce secondary bacterial infections. The most important therapeutic approach involves the administration of immunosuppressive doses of corticosteroids to reduce the cat's immune response to the virus.

Attempts at prevention are frustratingly difficult, but include vaccination, testing and removing infected cats, and early weaning of kittens from queens. No one method is completely effective at eliminating enteric coronavirus or preventing feline infectious peritonitis. The best protection is by reducing group size to five or less. *See* ANIMAL VIRUS; VIRUS. Janet E. Foley

Bibliography. C. E. Greene, *Infectious Disease of the Dog and Cat*, Saunders, 1998; N. C. Pedersen, *Feline Infectious Diseases*, American Veterinary Publications, 1988.

Feline leukemia

A type of cancer caused by the feline leukemia virus, a retrovirus which affects only a small percentage of freely roaming or domestic cats. The feline leukemia virus is genetically and morphologically similar to murine leukemia virus, from which it presumably evolved several million years ago.

Epidemiology and transmission. About 1–5% of healthy-appearing wild or freely roaming domestic cats have lifelong (persistent) infections. These carrier cats shed the virus in urine, feces, and saliva. The principal route of infection is oral. Transmission in the uterus, transmission from mother to kitten through nursing, and intraspecies biting are of secondary importance.

The infection rate among freely roaming cats rises progressively from 6 months onward, and by 3–5 years of age a majority have been exposed. However, infections occurring in nature are usually inapparent or mild, and 95% of such cats recover without any signs of illness. The remainder become infected for life and shed the virus for months or years before they finally succumb. Mortality due to a feline leukemia virus infection occurs mainly among persistently infected cats and at a rate of around 50% per year. Almost 90% of these infected cats die within 3 years.

Although feline leukemia virus is of minor importance among freely roaming cats, the mortality can be 30–50% among certain high-risk indoor cat populations. Populations at risk include breeding catteries, shelters, managed wild cat colonies, pet stores, and homes with a large number of pet cats. The infection is more severe in these instances because indoor cats often share the same litter and food and water dishes, and frequently groom each other. Emotional stresses among large groups of confined cats further lower their resistance. Kittens are much less resistant than adults to persistent infection, making breeding groups of indoor animals particularly susceptible.

The typical cat infected with feline leukemia virus is an unvaccinated, nonpurebred cat 1–7 years of age that has been allowed to roam freely or has been inadvertently housed with an infected animal.

About one-half of cats infected with feline leukemia virus develop cancers (leukemias) arising from cells within the bone marrow, thymus gland, lymph node, or spleen. Another one-fourth die from anemia, and an equal number are affected by a type of acquired immunodeficiency syndrome with complicating opportunistic infections. A smaller proportion of cats infected with feline leukemia virus develop neurologic or ocular disorders, or other rare types of cancers.

Treatment. There is no treatment that eliminates the virus. Once the illness appears, a medical examination will determine precisely which feline leukemia virus-related disorder has occurred. Supportive or symptomatic treatment may prolong life for weeks or months, depending on the particular disease manifestation.

Prevention. All cats should be tested for the presence of the virus prior to putting them in contact with feline leukemia virus–free animals. Healthy-appearing or ill infected cats should not be in intimate contact with noninfected cats, even if the latter have been vaccinated. If they have to be kept in the same premises, they should be isolated from other cats and their litter and food and water dishes kept separate. The virus is very unstable once it is shed from the body; therefore there is very little transmission by contaminated clothing or hands.

Inactivated and genetically engineered subunit feline leukemia virus vaccines are available and should be administered annually. Because they are not totally effective in preventing infection, feline leukemia virus vaccines should not be considered a substitute for testing, elimination, and quarantine procedures. Vaccination is not essential for virus-free cats that are kept strictly confined indoors, providing precautions are taken to prevent contact with infected animals. *See* LEUKEMIA; RETROVIRUS. Niels C. Pedersen

Bibliography. J. A. Levy (ed.), *The Retroviridae*, Vol. 2, 1993; N. C. Pedersen, *Feline Infectious Diseases*, 1988.

Feline panleukopenia

An acute virus infection of cats, also called feline viral enteritis and (erroneously) feline distemper. The virus infects all members of the cat family (Felidae) as well as some mink, ferrets, and skunks (Mustelidae); raccoons and coatimundi (Procyonidae); and the

binturong (Viverridae). Panleukopenia is the most important infectious disease of cats.

Infectious agent. Feline panleukopenia virus is classified as a parvovirus, and is one of the smallest known viruses. It is antigenically identical to the mink enteritis virus, and only minor antigenic differences exist between feline panleukopenia virus and canine parvovirus. It is believed that canine parvovirus originated as a mutation from feline panleukopenia virus. *See* ANIMAL VIRUS.

Epidemiology. This disease occurs worldwide, and nearly all cats are exposed by their first year because the virus is stable and ubiquitous; the disease is therefore rarely seen in older cats. Infected cats shed large amounts of virus in their secretions, vomit, and feces. The virus is highly resistant to inactivation and can remain infectious for months or years. It can easily by carried on shoes, clothes, feed bags, and toys. In addition, recovered cats may shed the virus for some time.

Pathogenesis. In order for parvoviruses to replicate, the cells they infect must be in a replicative stage. Consequently, only tissues with high cell proliferation rates are significantly affected by this virus group. These include leukocytes, bone marrow, intestinal crypt cells (cells that produce absorptive cells of intestinal villi) and many additional cell types in the developing fetus or neonate. Thus, the nature of the disease caused by feline panleukopenia virus depends upon the cat's age or pregnancy status. The typical form of the disease occurs in cats 2 weeks of age and older, which are exposed by nose or mouth. The virus attaches to and replicates in lymphoid cells lining the pharynx and adjacent tissues, and these cells carry the virus internally. A generalized infection occurs when new lymphoid cells are infected and the virus is released into the blood, infecting leukocytes, bone marrow, lymphoid cells of the thymus, lymph nodes, spleen, and follicles in the intestine. The virus then infects the proliferative crypt cells of the intestinal mucosa. Cells infected by feline panleukopenia virus die when the virus replicates, accounting for the clinical signs observed with panleukopenia. First, there is a precipitous drop in leucocyte counts 2–4 days after exposure. (The name panleukopenia, meaning absence of leukocytes in the blood, relates to this clinical feature.) Twenty-four to forty-eight hours later, destruction of the intestinal crypt cells results in the shortening and disruption of villi, causing diarrhea, vomiting, and loss of body fluids. Injury to the intestinal mucosa may allow bacteria or their toxins to enter the body. This, along with the destruction of the disease-fighting white blood cells, predisposes the cat to blood poisoning. Without treatment, this disease is often fatal.

The rapidly developing cat fetus may become infected and die, resulting in complete fetal resorption or delivery of a mummified or stillborn fetus. A unique syndrome, characterized by underdevelopment of the cerebellum, can occur in kittens infected by feline panleukopenia virus during the last 2 weeks of gestation or the first 2 weeks of life. The kitten may appear completely healthy, but it lacks muscle coordination because of destruction of parts of the cerebellum.

Diagnosis. Perhaps the majority of cats infected with feline panleukopenia virus do not become ill. However, the disease is severe and life threatening in 20–50% of cases. The cat is depressed and may refuse food or water; vomiting and diarrhea are common, resulting in severe dehydration. The cat may have a fever or a subnormal temperature. A low white blood cell count confirms the diagnosis as panleukopenia. Diagnosis can be made by autopsy and evidence of the destruction of the intestinal crypts and villus shortening.

Immunity and prevention. Both modified-live-virus and inactivated vaccines that are highly effective and safe are available for the prevention of panleukopenia. Only inactivated vaccines should be given to pregnant or immunosuppressed cats. The major obstacle to successful immunization of kittens is the presence of antibodies in the mother's milk. These antibodies block vaccination, and may last for the first 6–14 weeks of life, depending of the immune levels of the mother. For this reason, a series of two or three vaccinations, starting at 8 weeks of age and spaced 2–3 weeks apart, is given to kittens to ensure successful vaccination; a booster is often given once a year. If the immune status of a cat is unknown, and the cat will be in an environment with a high probability of exposure to feline panleukopenia virus (such as a cat show or a kennel), immune serum can be administered. This will protect the cat for 2–4 weeks, but the cat will have to be vaccinated later for permanent immunity.

Premises contaminated by feline panleukopenia virus are extremely difficult to disinfect because it is resistant to most of the common disinfectants; chlorine bleach, formaldehyde, or a certain quaternary ammonium disinfectant will destroy the virus. Because carpets, furniture, or other easily damaged objects in a home cannot be satisfactorily disinfected, a cat should be successfully immunized before being introduced to premises where a panleukopenia-infected cat previously lived. Jack H. Carlson

Bibliography. F. Fenner et al. (eds.), *Veterinary Virology*, 1993; B. N. Fields et al. (eds.), *Virology*, 3d ed., 1996; C. E. Greene (ed.), *Infectious Diseases of the Dog and Cat*, 2d ed., 1998.

Fennel

Forniculum vulgare, a culinary herb of the parsley family (Apiaceae). It is grown for the dried, ripe fruits or seeds which are used in bread, pickles, liqueurs, and meat sauces and dishes. Although similar in odor to anise, fennel seed can be distinguished by its warm, sweet character. *See* APIALES.

Fennel is one of approximately five species in the genus *Forniculum*. The plant is an erect, branching, short-lived perennial growing 3–5 ft (1–1.8 m) high. The leaves are three- to four-pinnately compound, threadlike, and 1.5 in. (3.7 cm) long. Common giant

fennel (*Ferula communis*), a close relative, grows twice as large (6–12 ft or 2–4 m), but has no culinary use.

The three fennel varieties of commercial importance are *vulgare*, *dulce*, and *piperitum*. The variety *vulgare* is grown for its seed and the essential oil, obtained by steam distillation and known as bitter fennel oil. The variety *dulce* (finochio or Florence fennel) is grown for four products; seed and leaf for culinary use, the enlarged leaf base for a vegetable, and for the essential oil from the seeds (sweet fennel oil). The young stems of Italian fennel, variety *piperitum*, are used for flavoring in salads.

Fennel is native to southern Europe and the Mediterranean region and is presently cultivated as an annual or biennial plant in southern and eastern Europe, India, Argentina, China, and Pakistan. The plants do best in a mild climate and on limey, well-drained soils with plenty of sun. Sown in early spring at 8–9 lb/acre (9–10 kg/ha) with rows spaced 2–3 ft (0.6–0.9 m) apart, the fennel plants are thinned to 8–12 in. (20–30 cm) apart when established. The ripe seeds are greenish gray and fairly hard. When the plants are harvested, yields are low the first season but increase to 700–1000 lb/acre (800–1100 kg/ha) by the second year. Fennel should be located away from dill to avoid the resulting flavor effects from cross-pollination. Aphids and mildew can cause damage to fennel.

Though it is primarily a culinary herb, commercial interest in fennel also lies in its essential oil, bitter or sweet, for cooking, medicine, perfumery, and as a masking agent for insecticides and room sprays. *See* SPICE AND FLAVORING. Seth Kirby

Fermentation

Decomposition of foodstuffs generally accompanied by the evolution of gas. The best-known example is alcoholic fermentation, in which sugar is converted into alcohol and carbon dioxide. This conversion, described by the equation below, was established

$$\underset{\text{Sugar}}{C_6H_{12}O_6} \rightarrow \underset{\text{Carbon dioxide}}{2CO_2} + \underset{\text{Alcohol}}{2C_2H_5OH}$$

by J. L. Gay-Lussac in 1815. *See* DISTILLED SPIRITS.

Before 1800 the association of yeast or leaven with fermentation had been noted, but the nature of these agents was not understood. Experiments of C. Cagniard-Latour, of F. T. Kützing, and of T Schwann in 1837 indicated that yeast is a living organism and is the cause of fermentation. This view was opposed by such leading chemists as J. von Liebig and F. Wöhler, who sought a chemical rather than a biological explanation of the process. The biological concept became generally accepted following the work of Louis Pasteur, who concluded that fermentation is a physiological counterpart of oxidation, and permits organisms to live and grow in the absence of air (anaerobically). This linked fermentation and putrefaction as comparable processes; both represent decompositions of organic matter brought about by microorganisms in the absence of air. The difference is determined by the nature of the decomposable material; sugary substances generally yield products with pleasant odor and taste (fermentation), whereas proteins give rise to evil-smelling products (putrefaction). *See* YEAST.

Pasteur also discovered the lactic acid and butyric acid fermentations, and from his experiments concluded that each kind of fermentation was caused by a specific microbe. Later work supported this idea to a large extent, and considerably increased the number of specific fermentations.

During fermentation organic matter is decomposed in the absence of air (oxygen); hence, there is always an accumulation of reduction products, or incomplete oxidation products. Some of these products (for example, alcohol and lactic acid) are of importance to society, and fermentation has therefore been used for their manufacture on an industrial scale. There are also many microbiological processes that go on in the presence of air while yielding incomplete oxidation products. Good examples are the formation of acetic acid (vinegar) from alcohol by vinegar bacteria, and of citric acid from sugar by certain molds (for example, *Aspergillus niger*). These microbial processes, too, have gained industrial importance, and are often referred to as fermentations, even though they do not conform to Pasteur's concept of fermentation as a decomposition in the absence of air. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; CITRIC ACID; INDUSTRIAL MICROBIOLOGY; VINEGAR. Cornelis B. Van Niel

Bibliography. D. R. Berry (ed.), *Physiology of Industrial Fungi*, 1988; R. Dubos, *Louis Pasteur, Free Lance of Science*, 1960, reprint 1986; B. M. McNeil and L. M. Harvey (eds.), *Fermentation: A Practical Approach*, 1990; O. P. Ward, *Fermentation Biotechnology*, 1989.

Fermi-Dirac statistics

The statistical description of particles or systems of particles that satisfy the Pauli exclusion principle. This description was first given by E. Fermi, who applied the Pauli exclusion principle to the translational energy levels of a system of electrons. It was later shown by P. A. M. Dirac that this form of statistics is also obtained when the total wave function of the system is antisymmetrical. *See* EXCLUSION PRINCIPLE; NONRELATIVISTIC QUANTUM THEORY.

Distribution function. Such a system is described by a set of occupation numbers $\{n_i\}$ which specify the number of particles in energy levels ϵ_i. It is important to keep in mind that ϵ_i represents a finite range of energies, which in general contains a number, say g_i, of nondegenerate quantum states. In the Fermi statistics, at most one particle is allowed in a nondegenerate state. (If spin is taken into account, two particles may be contained in such a state.) This is simply a restatement of the Pauli exclusion principle, and means that $n_i \leqq g_i$. The probability of having a

set $\{n_i\}$ distributed over the levels ϵ_i, which contain g_i nondegenerate levels, is described by Eq. (1),

$$W = \prod_i \frac{g_i!}{(g_i - n_i)!n_i!} \tag{1}$$

which gives just the number of ways that n_i can be picked out of g_i, which is intuitively what one expects for such a probability. In Boltzmann statistics this same probability is given by Eq. (2). The equi-

$$W' = \prod_i \frac{g_i^{ni}}{n_i!} \tag{2}$$

librium state which actually exists is the set of n's that makes W a maximum, under the auxiliary conditions given in Eqs. (3). These conditions express

$$\Sigma n_i = N \tag{3a}$$

$$\Sigma n_i \epsilon_i = E \tag{3b}$$

the fact that the total energy E and the total number of particles N are given. *See* BOLTZMANN STATISTICS.

Equation (4) holds for this most probable distribu-

$$n_i = \frac{g_i}{\frac{1}{A}\epsilon^{\beta\epsilon_i} + 1} \tag{4}$$

tion. Here A and β are parameters, to be determined from Eq. (3); in fact, $\beta = 1/kT$, where k is Boltzmann's constant and T is the absolute temperature. When the 1 in the denominator may be neglected, Eq. (4) goes over into the Boltzmann distribution; this provides a procedure for identifying β. It is known that in classical statistics the Boltzmann distribution may be obtained if specific assumptions are made as to the number of collisions taking place. It is there assumed that the number of collisions per second in which molecules with velocities in cells i and j in phase space produce molecules with velocities in cells k and l is given by Eq. (5). Here $a^{kl}{}_{ij}$ is a geomet-

$$A_{ij}^{kl} = n_i n_j a_{ij}^{kl} \tag{5}$$

rical factor. This leads to the Boltzmann distribution. The Fermi distribution, Eq. (4), may be obtained if instead of Eq. (5) one assumes Eq. (6) for the number

$$A_{ij}^{kl} = a_{ij}^{kl} n_i n_j \left(\frac{g_k - n_k}{g_k}\right)\left(\frac{g_l - n_l}{g_l}\right) \tag{6}$$

of collisions. One observes from Eq. (6) the interesting quantum theoretical feature that the probability for a collision depends on the occupation numbers of the states into which the colliding particles will go. In particular, if these final states are filled up ($n_k = g_k$), no collision with that state as a final state can occur.

The distribution f_i is often used; it is defined by Eq. (7).

$$n_i = f_i g_i \tag{7}$$

Applications. For a system of N electrons, each of mass m in a volume V, Eq. (4) may be written as Eq. (8), where h is Planck's constant. Equations (3a)

$$f\left(v_x v_y v_z\right)\, dv_x\, dv_y\, dv_z = 2\left(\frac{m}{h}\right)^3 V \frac{dv_x\, dv_y\, dv_z}{\frac{1}{A}e^{mv^2/2kT} + 1} \tag{8}$$

and (3b) may now be transformed into integrals, yielding Eqs. (9). Here $\lambda = (h^2/2\pi mkT)^{1/2}$ is the ther-

$$\frac{N}{V}\frac{\lambda^3}{2} = U_{1/2}(A) \tag{9a}$$

$$\frac{E}{3kT}\frac{\lambda^3}{V} = U_{3/2}(A) \tag{9b}$$

mal de Broglie wavelength, and $U_\rho(A)$ is the Sommerfeld integral defined by Eqs. (10), where $u = mv^2/2kt$

$$U_\rho(A) = \frac{1}{\Gamma(\rho+1)}\int_0^\infty \frac{u^\rho\, du}{\frac{1}{A}e^u + 1} \tag{10a}$$

$$U_\rho(A) \cong A \quad (A \ll 1) \tag{10b}$$

$$U_\rho(A) \cong \frac{(\ln A)^{\rho+1}}{\Gamma(\rho+1)} \quad (A \gg 1) \tag{10c}$$

and Γ is the usual Γ-function. Very often one writes instead of A in the Fermi distribution a quantity μ defined by Eq. (11). It may be shown (for instance,

$$\mu = kT \ln A \qquad A = e^{\mu/kT} \tag{11}$$

by going to the classical limit) that μ is the chemical potential. It may be seen that if ϵ is large $(1/kT)(\epsilon - \mu) \gg 1$; hence the Fermi distribution goes over into a Maxwell-Boltzmann distribution. It is easy to verify that this inequality may be transcribed so as to state expression (12). Physically this is reasonable because

$$\frac{V}{N} \gg \lambda^3 \tag{12}$$

expression (12) says that classical conditions pertain when the volume per particle is much larger than the volume associated with the de Broglie wavelength of a particle. For example, $V/N\lambda^3$ is about 7.5 for helium gas (He^3) at 4 K (−452°F) and 1 atm (10^2 kilopascals). This indicates that classical statistics may perhaps be applied, although quantum effects surely play a role. For electrons in a metal at 300 K (80°F), $V/N\lambda^3$ has the value 10^{-4}, showing that classical statistics fail altogether for electrons in metals. When the classical distribution fails, a degenerate Fermi distribution results. Numerically, if $A \gg 1$, a degenerate Fermi distribution results; if $A \ll 1$, the classical results are again obtained. For example, Eqs. (9a), (9b), and (10b) show that $E = {}^3/_2NkT$ and the specific heat should be ${}^3/_2R = {}^3/_2Nk$. However, for an electron gas this does not apply since such a system is degenerate for normal temperatures and $A \gg 1$. A somewhat lengthy calculation yields the result that, in the case A $\gg$ 1, the contribution to the specific heat

is negligible. This resolves an old paradox, for, according to the classical equipartition law, the electronic specific heat C should be $^3/_2Nk$, whereas in reality it is very small; in fact, $C = \gamma T$, where γ is a very small constant. This is a consequence of the fact that an electron gas at normal temperature is a degenerate Fermi gas. The electrical resistance of a metal can be understood on a classical picture, but the lack of a specific heat is a pure quantum effect. *See* BOSE-EINSTEIN STATISTICS; FREE-ELECTRON THEORY OF METALS; GAMMA FUNCTION; KINETIC THEORY OF MATTER; QUANTUM STATISTICS; SPECIFIC HEAT OF SOLIDS; STATISTICAL MECHANICS. Max Dresden

Fermi surface

The surface in the electronic wavenumber space of a metal that separates occupied from empty states. Every possible state of an electron in a metal can be specified by the three components of its momentum, or wavenumber. The name derives from the fact that half-integral spin particles, such as electrons, obey Fermi-Dirac statistics and at zero temperature fill all levels up to a maximum energy called the Fermi energy, with the remaining states empty. *See* FERMI-DIRAC STATISTICS.

The fact that such a surface exists for any metal, and the first direct experimental determination of a Fermi surface (for copper) in 1957, were central to the development of the theory of metals. The Fermi surface is a geometrical shape which could be determined by experiment and calculated theoretically, providing the needed test of the quantum theory of metals. Neither step could have been taken much earlier. High-purity materials were needed for the experimental studies, and high-speed digital computers for the theoretical determinations. Before 1957, many scientists believed that the one-electron approximation was hopelessly naive. Soon afterward the Fermi Surface Conference (Cooperstown, New York, 1960) considered the many Fermi surfaces that had by that time been determined experimentally.

Nearly free electrons. An early surprise arising from the known Fermi surfaces was that many of the shapes were close to what would be expected if the electrons interacted only weakly with the crystalline lattice. The long-standing free-electron theory of metals was based upon this assumption, but most physicists regarded it as a serious oversimplification. *See* FREE-ELECTRON THEORY OF METALS.

The momentum p of a free electron is related to the wavelength λ of the electronic wave by Eq. (1),

$$p = \frac{2\pi\hbar}{\lambda} \tag{1}$$

where $\hbar$ is Planck's constant divided by 2π. The ratio $2\pi/\lambda$, taken as a vector in the direction of the momentum, is called the wavenumber $\boldsymbol{k}$. If the electron did not interact with the metallic lattice, the energy would not depend upon the direction of $\boldsymbol{k}$, and all constant-energy surfaces, including the Fermi surface, would be spherical. This is illustrated in

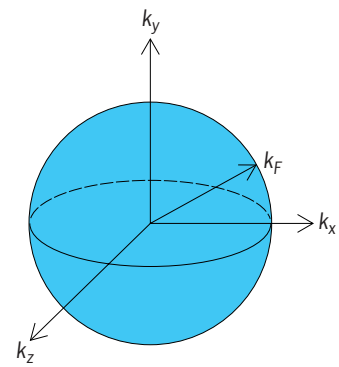

Fig. 1. Free-electron Fermi surface in the space defined by the components of wavenumbers of the electrons.

Fig. 1, with the radius k_F related to the number N of electrons per unit volume by Eq. (2).

$$k_F{}^3 = 3\pi^2 N \tag{2}$$

The Fermi surface of copper was found to be distorted (**Fig. 2**) but was still a recognizable deformation of a sphere. The polyhedron surrounding the Fermi surface in the figure is called the Brillouin zone. It consists of Bragg-reflection planes, the planes made up of the wavenumbers for which an electron can be diffracted by the periodic crystalline lattice. The square faces, for example, correspond to components of the wavenumber along one coordinate axis equal to $2\pi/a$, where a is the cube edge for the copper lattice. For copper the electrons interact with the lattice so strongly that when the electron has a wavenumber near to the diffraction condition, its motion and energy are affected and the Fermi surface is correspondingly distorted. The Fermi surfaces of sodium and potassium, which also have one conducting electron per atom, are very close to the sphere shown in Fig. 1. These alkali

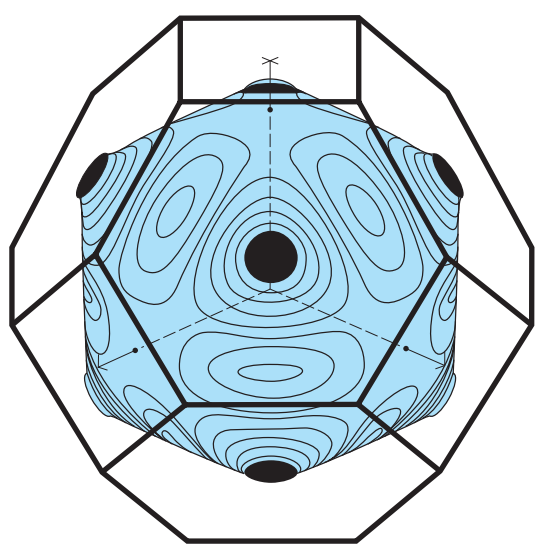

Fig. 2. Fermi surface of copper, as determined in 1957; two shapes were found to be consistent with the original data, and the other, slightly more deformed version turned out to be correct.

metals are therefore more nearly free-electron-like. *See* BRILLOUIN ZONE.

Metals with more than one electron per atom, such as calcium and aluminum, have higher electron densities and correspondingly larger Fermi wavenumbers k_F; the Fermi sphere overlaps the Bragg reflection planes. Again the interaction with the lattice is weak and deforms the Fermi surface only slightly. However, electrons accelerated by a magnetic field and with wavenumbers in a Bragg plane are diffracted to other parts of the Fermi sphere. Thus their motion in this wavenumber space is discontinuous, and the corresponding motion of the electron in real space shows sharp turns between smooth circular sections of orbit. It is possible to describe these complex motions in a more continuous fashion by translating various sections of the Fermi sphere back into the original Brillouin zone (Fig. 2), so that the motion of the wavenumber is continuous across these more complicated Fermi surfaces. Such constructions are shown in **Fig. 3** for calcium (valence 2) and aluminum (valence 3). The motion of any electron in a uniform magnetic field is along the intersection of one of these surfaces and a plane perpendicular to the magnetic field. The electron orbit in the real crystal is easily seen to be of exactly the same shape, though rotated by 90°. The size of the real orbit, for attainable magnetic fields, is very much larger than the lattice distance as the electron successively accelerates and diffracts. In the real metal, the Fermi surface and the corresponding orbits are deformed and rounded at the edges but are still close to the ideal surfaces shown.

The weakness of the interaction between electrons and lattice was a surprise, since the potential that describes this interaction is large. It turns out, however, that the electrons move so rapidly past the atoms that the effective potential, called a pseudopotential, is weak. It was the understanding of the nearly free-electron Fermi surfaces which suggested the weak pseudopotential and the possibility of simple theories of metals treating the pseudopotential as a small perturbation in the electron states.

In transition metals there are electrons arising from atomic *d* levels, in addition to the free electrons, and the corresponding Fermi surfaces are more complex than those of the nearly free-electron metals. However, the Fermi surfaces exist and have been determined experimentally for essentially all elemental metals.

Determination from experiments. The motion of the electrons in a magnetic field discussed above provides the key to experimentally determining the Fermi surface shapes. The simplest method conceptually derives from ultrasonic attenuation. Sound waves of known wavelength pass through the metal and a magnetic field is adjusted, yielding fluctuations in the attenuation as the orbit sizes match the sound wavelength. This measures the diameter of the orbit and Fermi surface. The most precise method uses the de Haas–van Alphen effect, based upon the quantization of the electronic orbits in a magnetic field. Fluctuations in the magnetic susceptibility give a direct measure of the cross-sectional areas of the Fermi surface. The method used in 1957 to determine the Fermi surface of copper did not use a magnetic field, but was based upon the anomalous skin effect, which determines the number of electrons moving parallel to the metal surface and from that deduces the curvature of the Fermi surface. *See* DE HAAS-VAN ALPHEN EFFECT; SKIN EFFECT (ELECTRICITY); ULTRASONICS.

Fermi surface determination has become an end in itself, with findings tabulated in the literature. There is intense interest in determining the Fermi

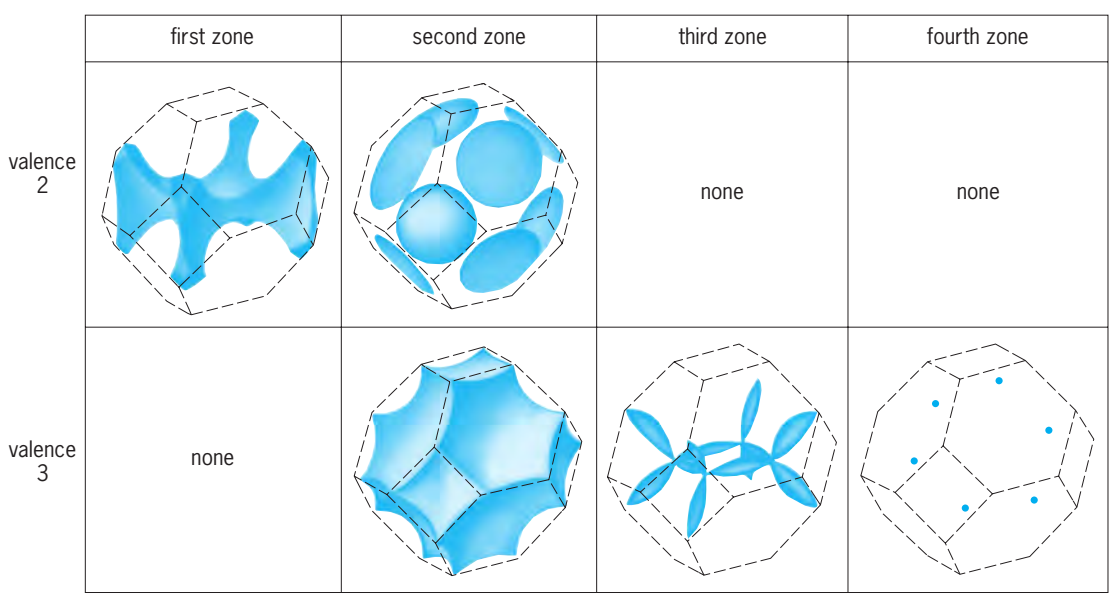

Fig. 3. Free-electron Fermi surfaces for face-centered cubic metals of valence 2 (calcium) and valence 3 (aluminum). (*After W. A. Harrison, Electronic structure of polyvalent metals, Phys. Rev., 118:1190–1208, 1960*)

surface, if one exists, in the high-temperature superconductors, but no experimental determination has yet been convincing. *See* BAND THEORY OF SOLIDS; SOLID-STATE PHYSICS; SUPERCONDUCTIVITY.

Walter A. Harrison

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; A. P. Cracknell (ed.), *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology*, *New Series*, Group 3: *Crystal and Solid State Physics*, vol. 13c: *Electron States and Fermi Surfaces of Elements*, 1984; A. P. Cracknell and K. C. Wong, *The Fermi Surface*: *Its Concept*, *Determination*, *and Use in the Physics of Metals*, 1973; W. A. Harrison and M. B. Webb (eds.), *The Fermi Surface*, 1960; J. Kondo and A. Yoshimori, *Fermi Surface Effects*, 1988.

Fermium

A chemical element, Fm, atomic number 100, the eleventh element in the actinide series. Fermium does not occur in nature; its discovery and production have been accomplished by artificial nuclear transmutation of lighter elements. Radioactive isotopes of mass number 244–259 have been discovered. The total weight of fermium which has been synthesized is much less than one-millionth of a gram. *See* ACTINIDE ELEMENTS; PERIODIC TABLE; RADIOACTIVITY.

1												13	14	15	16	17	18
1 **H**	2																2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Spontaneous fission is the major mode of decay for ^{244}Fm, ^{256}Fm, and ^{258}Fm. The longest-lived isotope is ^{257}Fm, which has a half-life of about 100 days. Fermium-258 decays by spontaneous fission with a half-life of 0.38 millisecond. This suggests the existence of an abnormality at this point in the nuclear periodic table. *See* NUCLEAR CHEMISTRY; NUCLEAR REACTION; TRANSURANIUM ELEMENTS.

Glenn T. Seaborg

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; J. J. Katz et al. (eds.), *The Chemistry of the Actinide and Transactinide Elements*, 3d ed., 2006; G. T. Seaborg and W. D. Loveland, *The Elements Beyond Uranium*, 1990.

Ferret

One of three carnivorous mammals, also known as polecats, in the family Mustelidae. The European polecat (*Mustela putorius*) inhabits Europe west of the Ural Mountains; the steppe polecat (*M. eversmanni*) is found in the steppe zone from Austria to Manchuria and Tibet; and the black-footed ferret (*M. nigripes*) inhabits the plains region of North America from Alberta and Saskatchewan to northeastern Arizona and Texas. The domestic ferret (*M. putorius furo*) is generally thought to be a descendant of one or both of the Old World species.

General morphology. Ferrets and polecats are weasellike and have long slender bodies, small rounded ears, and short legs. The limbs bear five digits with nonretractile curved claws. The European polecat is the darkest of the three species with a general coloration of dark brown to black with pale yellow underfur. The steppe polecat is generally straw yellow or pale brown with the chest, limbs, and terminal third of the tail being dark brown to black. There is a dark mask across the face. The black-footed ferret is generally yellow buff with the forehead, muzzle, and throat being nearly white. The top of the head and middle of the back are brown, while the face mask, feet, and tip of the tail are black. The domestic ferret is generally white or pale yellow in color. The dental formula is I 3/3, C 1/1, Pm 3/3, M 1/2 × 2 for a total of 34 teeth. The senses of hearing, smell, and sight are well developed. The head and body length of these three species ranges 205–500 mm (8–20 in.), and the tail length is 70–190 mm (2.7–7.5 in.). Weight ranges 205–2050 g (0.75–4.5 lb). Males are larger than females. *See* DENTITION.

Range and occurrence. The original range of the black-footed ferret coincided with the range of prairie dogs (*Cynomys*), its primary prey. The widespread use of sodium fluoroacetate ("1080") to poison prairie dogs resulted in secondary poisoning of black-footed ferrets that fed on them. In addition, the large-scale eradication of prairie dogs from the range of the black-footed ferret greatly reduced their

Black-footed ferret, *Mustela nigripes*. (© *Joseph Dougherty/www.ecology.org*)

food supply. Today, the ferret is considered the most endangered mammal in North America. One of the last known populations of black-footed ferrets was extirpated in South Dakota during the early 1970s. The species was feared to be extinct until a small population was found in northwestern Wyoming in 1981. Several animals from the Wyoming population were caught for a captive breeding program before the wild population disappeared. Successes in the captive breeding program allowed reintroductions at a southcentral Wyoming site from 1991 until 1995, but disease and other factors caused that program to be suspended. Beginning in 1994, reintroduction efforts were begun in Montana and South Dakota. In 1996, Arizona became the recovery program's fourth reintroduction site. At present, there are no known nonintroduced wild populations. *See* PRAIRIE DOG.

Behavioral ecology and reproduction. Relatively little is known about the ecology of black-footed ferrets. They are primarily nocturnal and spend most of their time in underground prairie dog burrows which they use for shelter and travel. Adult ferrets are solitary except during the breeding season. Adults of the same sex are not known to inhabit the same prairie dog town, whereas adults of the opposite sex are found in the same town but not in the same burrow system. While prairie dogs appear to be the primary food, black-footed ferrets also prey on a variety of other small mammals. They kill their prey by attacking the neck and base of the skull. Potential predators of ferrets are badgers, coyotes, bobcats, rattlesnakes, eagles, hawks, and owls.

Ferrets breed in the spring with litters born after a gestation period of 42–45 days. Litter size ranges 1–6 with an average of 3.5 kits. Unlike many weasels, there does not seem to be delayed uterine implantation. Newborn young are altricial (immature and helpless) and have their eyes closed. Sexual maturity is attained by one year of age. Captives have lived to about 12 years of age.

The domestic ferret is usually tame and playful, and is found in captivity throughout the world. In some areas it is used to control rodents and to drive rabbits from their burrows. *See* CARNIVORA; WEASEL.

Donald W. Linzey

Bibliography. G. A. Feldhamer, B. C. Thompson, and J. A. Chapman, *Wild Mammals of North America: Biology, Management, and Conservation*, 2d ed., Johns Hopkins University Press, 2003; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Ferricyanide and ferrocyanide

The common names for hexacyanoferrate(III) and hexacyanoferrate(II), respectively.

Ferricyanide. Hexacyanoferrate(III) is a compound containing the $[Fe(CN)_6]^{3-}$ complex ion. The oxidation state of iron is a tripositive ion [Fe^{3+} or Fe^{III}] and is low spin (spin paired), consistent with the strong field nature of the cyanide (CN^-) ligand. The complex $[Fe(CN)_6]^{3-}$ adopts an octahedral geometry (see **illus.**).

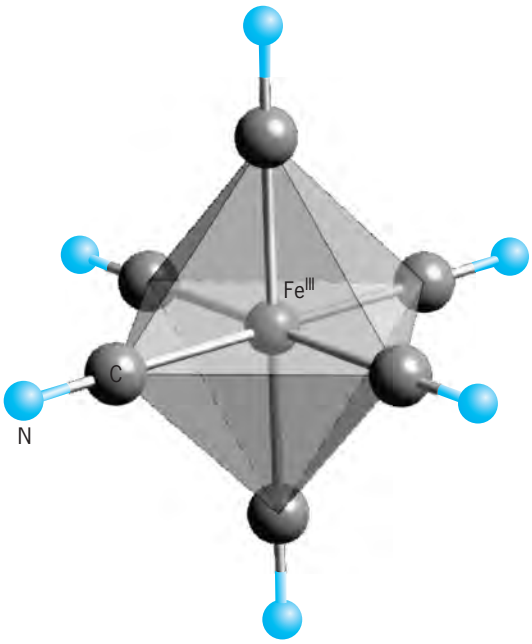

Diagram showing octahedral shape of ferricyanide.

The $[Fe(CN)_6]^{3-}$ ion is kinetically unstable, as it dissociates to give the free cyanide anion, CN^-. It is therefore toxic (LD_{50} = 2690 mg/kg versus 10 mg/kg for KCN). In contrast, the ferrocyanide ion $[Fe(CN)_6]^{4-}$ is stable and less toxic (LD_{50} = 6400 mg/kg). Substituted derivatives, such as the nitroprusside ion $[Fe(CN)_5NO]^{2-}$, are known. The redox potential for the aqueous reduction of ferricyanide to ferrocyanide is ($[Fe(CN)_6]^{3-} + e^- = [Fe(CN)_6]^{4-}$), $E^\circ = 0.36$ V (versus standard hydrogen electrode), but is quite solvent dependent.

The sodium ($Na_3[Fe(CN)_6]$) and potassium ($K_3[Fe(CN)_6]$) salts have been isolated as ruby-red crystals and are photosensitive. The potassium salt reacts with metallic silver to produce silver ferrocyanide, which is used in photographic processes. Addition of Fe^{II} ions to ferricyanide forms Prussian blue $Fe^{III}_4\,[Fe^{II}(CN)_6]_3 \cdot xH_2O$, $x \sim 15$. This pigment is used in blueprint materials and as a mild oxidizing agent in organic synthesis. *See* PRUSSIAN BLUE.

Ferrocyanide. Hexacyanoferrate(II) is a compound containing the complex ion $[Fe(CN)_6]^{4-}$. The oxidation state of iron is dipositive (Fe^{2+} or Fe^{II}) and is low spin (spin paired), consistent with the strong field nature of the cyanide (CN^-) ligand.

The $[Fe(CN)_6]^{4-}$ ion is very stable, with an octahedral geometry similar to that of $[Fe(CN)_6]^{3-}$. The free acid, hydrogen hexacyanoferrate(II) ($H_4[Fe(CN)_6]$), is soluble in water and is isolated as a white powder by precipitation of the ion in strongly acidic solutions. Salts of sodium ($Na_4[Fe(CN)_6]$) are prepared by heating an aqueous mixture of sodium cyanide (NaCN), iron(II) sulfate ($FeSO_4$), and other salts. In contrast to ferricyanide $[Fe(CN)_6]^{3-}$, ferrocyanide is substitutionally inert—that is, it is slow to liberate

cyanide. Salts of ferrocyanide are yellow hydrates, which are insoluble in most organic solvents. Ferrocyanide is also used as an antidote for thallium poisoning.

Hexacyanoferrate(II) may be converted to hexacyanoferrate(III) (ferricyanide) by strong oxidizing agents such as peroxides and permanganate ion.

Ferrocyanide is commonly used as a reducing agent. Its also is used as an additive to make free-flowing table salt and as a U.S. FDA-approved colorant, as well as in the preparation of dyes and photographic fixatives and for stabilizing synthetic and natural latex foams.

Replacement of Fe^{II} or Fe^{III} ions with other metal ions leads to other Prussian blue–structured materials, which are magnets. Use of Cr^{III} in place of Fe^{II}, as well as V^{II} in place of Fe^{II}, leads to a family of magnets with ordering temperatures above room temperature. *See* COORDINATION CHEMISTRY; COORDINATION COMPLEXES; CURIE TEMPERATURE; CYANIDE; IRON.

Joel S. Miller

Bibliography. K. R. Dunbar and R. A. Heintz, Chemistry of transition metal cyanide compounds: Modern perspectives, *Progr. Inorg. Chem.*, 45:283–392, 1997; J. S. Miller, 3-D-network structured cyanide-based magnets, *MRS Bull.*, 25(11):60–64, 2000; A. G. Sharp, *The Chemistry of Cyano Complexes of the Transition Metals*, Academic Press, New York, 1976.

Ferrimagnetic garnets

A class of ferrimagnetic oxide materials that have the garnet crystal structure. The crystal structure contains three types of nonequivalent sites in the oxygen lattice which are occupied by metal ions. The sites are referred to as 16(*a*), 24(*c*), and 24(*d*). A metal ion in a 16(*a*) site is surrounded by six oxygen ions which are at the corners of the octahedron. The metal ions in the 24(*d*) sites are found in nonregular tetrahedrons formed by oxygen ions, while the metal ions in the 24(*c*) sites have eight oxygen-ion nearest neighbors to form a triangular dodecahedron at the corners of a distorted cube. *See* CRYSTALLOGRAPHY; FERRIMAGNETISM; FERRITE.

The typical formula of a ferrimagnetic garnet is $X_3Fe_5O_{12}$, where the trivalent X ion is yttrium, or any of the rare-earth ions with an atomic number greater than 61. (The rare-earth ions below samarium apparently have radii too large to fit into the garnet structure.) However, the larger rare earths and lanthanum can substitute partially for some of the yttrium or smaller rare earths. The iron occupies both the 16(*a*) and 24(*d*) sites, while the yttrium, lanthanum, or rare-earth ion, or a combination of two or more of these ions, occupies the 24(*c*) sites. The 24 iron ions (per unit cell) on the (*d*) sites are strongly coupled, with an antiferromagnetic coupling, to the 16 iron ions in the (*a*) sites. If the 24 ions on the (*c*) sites have a magnetic moment, it is weakly aligned antiparallel to the moment of the ions on the (*d*) sites. A significant amount of divalent calcium ion can be substituted for rare-earth ions, if an equal amount of a quadravalent ion such as germanium is substituted for iron to maintain charge balance. *See* RARE-EARTH ELEMENTS.

Ferrimagnetic garnets are of great theoretical interest because they have a highly ordered structure and because they accommodate rare-earth ions, some of which have a small contribution to their magnetism due to the orbital motion of electrons, in addition to the magnetism due to electron spin. The first practical engineering interest in the ferrimagnetic garnets was due to the yttrium iron garnet, which is used in certain microwave ferrite devices because of its very narrow ferromagnetic resonance absorption line (less than 1 oersted or 80 A/m at 10,000 MHz). Development of magnetic bubbles for use in solid-state nonvolatile memory devices resulted in a large research effort on complex rare-earth iron garnets. In these garnets several different rare-earth ions are simultaneously incorporated in the structure in concentrations designed to optimize a number of physical and magnetic properties which influence the memory parameters. An example is the garnet $(Y_{1.73}Lu_{.22}Sm_{.20}Ca_{.85})(F_{4.15}Ge_{.85})O_{12}$ used in practical bubble memories with a storage density of 3×10^5 bits/cm^2. *See* COMPUTER STORAGE TECHNOLOGY; FERRITE DEVICES; GYRATOR.

Single crystals of ferrimagnetic garnets can be grown directly from a melt of the appropriate oxides, from a solution of the oxides dissolved in a PbO/B_2O_3 flux, or from a gas of the chlorides, which are volatile at high temperatures. Rare-earth garnets for magnetic bubble devices are grown from the flux onto the surface of crystallographically compatible nonmagnetic gadolinium gallium garnet in the form of single-crystal epitaxial films a few micrometers thick. Under the growth conditions used, these films have anisotropic magnetic properties, even though the crystal structure is cubic, as a result of growth- and stress-induced effects. *See* CRYSTAL GROWTH.

Morton E. Jones

Bibliography. J. Goldman, *Modern Ferrite Technology*, 1990; A. Paoletti (ed.), *Physics of Magnetic Garnets*, 1979; C. Srivastava and M. J. Patni, *Ferrites*, 1990; H. Watanabe (ed.), *Ferrites*, 1982.

Ferrimagnetism

A specific type of ordering in a system of magnetic moments or the magnetic behavior resulting from such order. In some magnetic materials the magnetic ions in a crystal unit cell may differ in their magnetic properties. This is clearly so when some of the ions are of different species. It is also true for similar ions occupying crystallographically inequivalent sites. Such ions differ in their interactions with other ions, because the dominant exchange interaction is mediated by the neighboring nonmagnetic ions. They also experience different crystal electric fields, and these affect the magnetic anisotropy of the ion. A collection of all the magnetic sites in a crystal with identical behavior is referred to as a

magnetic sublattice. A material is said to exhibit ferrimagnetic order when, first, all moments on a given sublattice point in a single direction and, second, the resultant moments of the sublattices lie parallel or antiparallel to one another. The notion of such an order is due to L. Néel, who showed in 1948 that its existence would explain many of the properties of the magnetic ferrites. *See* FERRITE; FERROMAGNETISM.

At high temperatures all magnetic systems are disordered. As the temperature of a potentially ferrimagnetic system is lowered, there comes a point, the Curie temperature, at which all sublattices simultaneously acquire a moment and arrange themselves in a definite set of orientations. The Curie temperature and the temperature dependence of the sublattice moments depend in a complicated way on the magnetic properties of the individual ions and upon the interactions between them. In general, there is a net moment, the algebraic sum of the sublattice moments, just as for a normal ferromagnet. However, its variation with temperature rarely exhibits the very simple behavior of the normal ferromagnet. For example, in some materials, as the temperature is raised over a certain range, the magnetization may first decrease to zero and then increase again. Ferrimagnets can be expected, in their bulk properties, measured statically or at low frequencies, to resemble ferromagnets with unusual temperature characteristics. *See* CURIE TEMPERATURE.

The most versatile of ferrimagnetic systems are the rare-earth iron garnets. The garnet unit cell has three sets of inequivalent magnetic sites, differing in their coordination to neighboring oxygen ions. Two of the sets, with 24 and 16 sites respectively per unit cell, are each occupied by Fe^{3+} ions, and the corresponding sublattices orient antiparallel. The remaining set of 24 sites may be occupied by nonmagnetic ions, such as Y, or by a magnetic rare-earth ion. This sublattice, when magnetic, usually lies parallel to the 24-site Fe^{3+} lattice. Rare-earth ions of various species have widely different magnetic moments, exchange interactions, and crystalline anisotropies. By a suitable choice of rare-earth ions, it is possible to design ferrimagnetic systems with prescribed magnetizations and temperature behavior. *See* FERRIMAGNETIC GARNETS.

Laurence R. Walker

Bibliography. D. J. Craik (ed.), *Magnetic Oxides*, 1975; D. J. Craik, *Magnetism*, 1994; C. Srivastava and M. J. Patni, *Ferrites*, 1990; H. Watanabe (ed.), *Ferrites,* 1982.

Ferrite

Any of the class of magnetic oxides. Typically the ferrites have a crystal structure that has more than one type of site for the cations. Usually the magnetic moments of the metal ions on sites of one type are parallel to each other, and antiparallel to the moments on at least one site of another type. Thus ferrites exhibit ferrimagnetism, a term coined by L. Néel. *See* FERRIMAGNETISM.

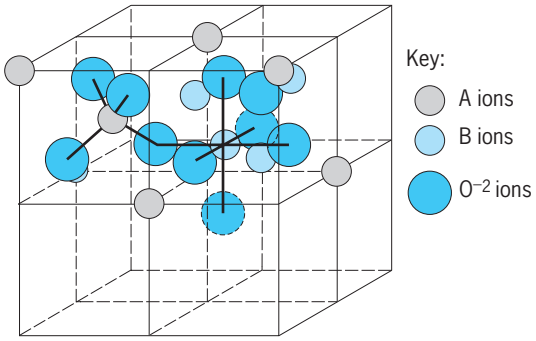

Fig. 1. Two octants of the spinel unit cell showing A ions on tetrahedral sites and B ions on octahedral sites. (*After E. P. Wohlfarth, ed., Ferromagnetic Materials, vol. 2, North-Holland, 1980*)

Commercial types. There are three important classes of commercial ferrites. One class has the spinel structure, with the general formula $M^{2+}Fe_2^{3+}O_4$, where M^{2+} is a divalent metal ion. In the unit cell of this structure, there are 8 tetrahedral and 16 octahedral sites for the locations (**Fig. 1**). The locations in the tetrahedral and octahedral sites are surrounded by four and six oxygen ions, respectively. The sites of each type constitute a sublattice. The moments of the two sublattices in the spinels are antiparallel, and the net moment is given by the difference in the moments of the sublattices. The magnetic moment of a spinel is then determined by the magnetic moment of M^{2+} and Fe^{3+} together with their distribution in the two sublattices. So-called linear ferrites used in inductors and transformers are made of Mn and Zn (for frequencies up to 1 MHz) and Ni and Zn (for frequencies greater than 1 MHz). MgMn ferrites are used in microwave devices such as isolators and circulators. Until the late 1970s, ferrites with square loop shapes held a dominant position as computer memory-core elements, but these gave way to semiconductors. *See* COMPUTER STORAGE TECHNOLOGY; GYRATOR; SEMICONDUCTOR.

The second class of commercially important ferrites have the garnet structure, with the formula $M_3^{3+}Fe_5^{3+}O_{12}$, where M^{3+} is a rare-earth or yttrium ion. The Fe^{3+} ions form two antiparallel sublattices composed of tetrahedral and octahedral sites. The M^{3+} ions form a third sublattice of dodecahedral sites in which each M^{3+} ion is surrounded by eight oxygen ions (**Fig. 2**). Yttrium-based garnets are used in microwave devices. Thin monocrystalline films of complex garnets have been developed for bubble domain memory devices. *See* FERRIMAGNETIC GARNETS.

The third class of ferrites has a hexagonal structure, of the $M^{2+}Fe_{12}^{3+}O_{19}$ magnetoplumbite type, where M^{2+} is usually Ba, Sr, or Pb. The Fe^{3+} ions occupy both tetrahedral and octahedral sites, in addition to a site surrounded by five oxygen ions. Because of their large magnetocrystalline anisotropy, the hexagonal ferrites develop high coercivity and are an important member of the permanent magnet family.

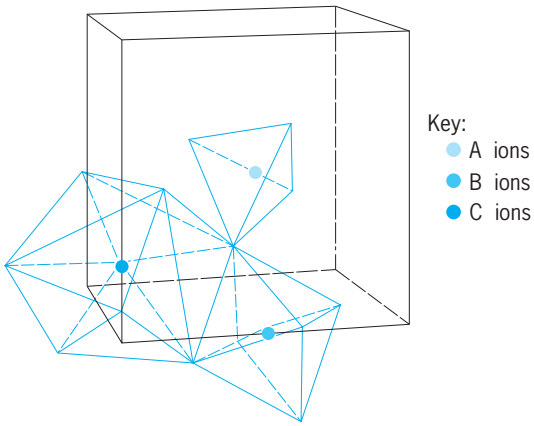

Fig. 2. One octant of the $M_3Fe_5O_{12}$ garnet unit cell showing A ions on tetrahedral sites, B ions on octahedral sites, and C ions on dodecahedral sites. (*After M. A. Gilleo and S. Geller, Magnetic and crystallographic properties of substituted yttrium-iron garnet, $3Y_2O_3 \cdot xM_2O_3 \cdot [5-x]\ Fe_2O_3$, Phys. Rev., 110:73–78, 1958*)

Another magnetic oxide, γ-Fe_2O_3, also has the spinel structure, but has no divalent cations. It is the most commonly used material in the preparation of magnetic recording tapes. *See* CRYSTAL STRUCTURE; MAGNETIC RECORDING.

Properties. The important intrinsic parameters of a ferrite are the saturation magnetization, Curie temperature, and magnetocrystalline (K_1) and magnetostrictive (λ_s) anisotropies. These properties are determined by the choice of the cations and their distribution in the various sites. Commercially important ferrites invariably contain two or more M elements; and the iron content often deviates from stoichiometry to optimize the magnetic properties. High-permeability MnZn ferrites, for example, contain about 26% MnO, 22% ZnO, and 52% Fe_2O_3 ($Mn_{0.51}Zn_{0.44}Fe_{2.05}O_4$), at which composition both K_1 and λ_s are nearly zero.

In addition to the intrinsic magnetic parameters, microstructure plays an equally important role in determining device properties. Thus, grain size, porosity, chemical homogeneity, and foreign inclusions dictate in part such technical properties as permeability, line width, remanence, and coercivity in polycrystalline ceramics. In garnet films for bubble domain device applications, the film must essentially be free of all defects such as inclusions, growth pits, and dislocations.

Preparation. There are a number of methods by which ferrites may be prepared; the choice is dictated by the final form desired. Thin monocrystalline garnet films for bubble domain devices are grown by a liquid-phase epitaxy technique on top of a lattice- matched nonmagnetic gadolinium gallium garnet (GGG) crystal substrate. The GGG single crystals are usually grown by the Czochralski pulling technique. Spinel ferrite single crystals for use in recording heads are often grown by the Bridgman technique or the Czochralski technique. Oxide powders prepared for magnetic recording are obtained by precipitation from aqueous salt solutions. *See* SINGLE CRYSTAL.

Polycrystalline ferrites are most economically prepared by ceramic techniques. Component oxides or carbonates are mixed, calcined at elevated temperatures for partial compound formation, and then granulated by ball milling. Dispersants, plasticizers, and lubricants are added, and the resultant slurry is spray-dried, followed by pressing to desired shape and sintering. The last step completes the chemical reaction to the desired magnetic structure and effects homogenization, densification, and grain growth of the compact. It is perhaps the most critical step in optimizing the magnetic properties of commercial ferrites. *See* CERAMICS; FERRITE DEVICES.

Gilbert Y. Chin

Bibliography. A. J. Baden-Fuller, *Ferrites at Microwave Frequencies*, 1987; D. J. Craik (ed.), *Magnetic Oxides*, 1975; J. Goldman, *Modern Ferrite Technology*, 1990; E. C. Snelling, *Soft Ferrites: Properties and Applications*, 2d ed., 1988; E. C. Snelling and A. D. Giles, *Ferrites for Inductors and Transformers*, 1984; C. Srivastava and M. J. Patni, *Ferrites*, 1990; R. Valenzuela, *Magnetic Ceramics*, 1994; H. Watanabe (ed.), *Ferrites*, 1982.

Ferrite devices

Electrical devices whose principle of operation is based upon the use and properties of ferrites, which are magnetic oxides. Ferrite devices are divided into two categories, depending on whether the ferrite is magnetically soft (low coercivity) or hard (high coercivity). Soft ferrites are used primarily as transformers, inductors, and recording heads, and in microwave devices. Since the electrical resistivity of soft ferrites is typically 10^6–10^{11} times that of metals, ferrite components have much lower eddy current losses and hence are used at frequencies generally above about 10 kHz. Hard ferrites are used in permanent-magnet motors, loudspeakers, and holding devices, and as storage media in magnetic recording devices. In this article, discussion of hard ferrites will be restricted to recording only. One type of soft ferrite, referred to as square-loop ferrites, was once deployed in huge quantities as cores for digital computer memories but has now been replaced by semiconductor integrated circuits. *See* FERRIMAGNETISM; FERRITE.

Chemistry and crystal structure. Modern soft ferrite devices stemmed from the contributions on spinel ferrites made by the Japanese and Dutch scientists during World War II. The general formula for the spinel is MFe_2O_4, in which M is a divalent metal ion. In special cases, the divalent ion M can be replaced by an equal molar mixture of univalent and trivalent ions. Thus lithium ferrite can be thought of as having the formula $Li_{0.5}Fe_{0.5}Fe_2O_4$. However, the commercially practical ferrites are those in which the divalent ion represents one or more magnesium (Mg), manganese (Mn), iron (Fe), cobalt (Co), nickel (Ni), copper (Cu), zinc (Zn), and cadmium (Cd) ions. The trivalent Fe ion may also be substituted by other trivalent ions such as aluminum (Al). The compositions are carefully adjusted to optimize the device

requirements, such as permeability, loss, ferromagnetic resonance line width, and so forth.

The ferrimagnetic garnets were discovered in France and the United States in 1956. The general formula is $M_3Fe_5O_{12}$, in which M is a rare-earth or yttrium ion. Single-crystal garnet films form the basis of bubble domain device technology. Bulk garnets have applications in microwave devices. *See* FERRIMAGNETIC GARNETS.

Hard ferrites for permanent-magnetic device applications have the hexagonal magnetoplumbite structure, with the general formula $MFe_{12}O_{19}$, where M is usually barium (Ba) or strontium (Sr). Hexagonal ferrites exhibit large coercivity ($H_c \simeq 200$ kA/m or 2.5 kilooersteds) owing to a large magnetocrystalline anisotropy. Particulates of this material are also under study as potential storage media in magnetic recording.

The material of choice for magnetic recording is γ-Fe_2O_3, which has a spinel structure. The particles are usually modified with a thin cobalt layer on the surface to increase the coercivity for enhanced performance. *See* MAGNETIC RECORDING.

With the exception of some single crystals used in recording heads and special microwave applications, and particulates used as storage media in magnetic recording, all ferrites are prepared in polycrystalline form by ceramic techniques.

Applications. A summary of the applications of ferrites is given in the **table**. These may be divided into nonmicrowave, microwave, and magnetic recording applications. Further, the nonmicrowave applications may be divided into categories determined by the magnetic properties based on the *B-H* behavior, that is, the variation of the magnetic induction or flux density *B* with magnetic field strength *H*, as shown in the **illustration**. The plot in the illustration is termed a hysteresis loop, with the area encompassed by the loop being proportional to the power loss per unit volume within the ferrite. The categories are linear *B-H*, with low flux density, and nonlinear *B-H*, with medium to high flux density. The highly nonlinear *B-H*, with a square or rectangular hysteresis loop, was once exploited in computer memory cores.

Linear B-H devices. In the linear region, the most important devices are high-quality inductors, particularly those used in filters in frequency-division multiplex telecommunications systems and low-power wide-band and pulse transformers. Virtually all such devices are made of either MnZn ferrite or NiZn ferrite, though predominantly the former.

In the design of inductors, the so-called μQ product of a material has been found to be a useful index of the quality of the material. In this product, μ is the initial permeability and Q is equal to $\omega L/R$,

Summary of ferrite applications

Ferrite chemistry	Device	Device function	Frequencies	Desired ferrite properties
		*Linear B-H, low flux density**		
MnZn, NiZn	Inductor	Frequency selection network	<1 MHz (MnZn)	High μ, high μQ, high stability of μ with temperature and time
		Filtering and resonant circuits	~1–100 MHz (NiZn)	
MnZn, NiZn	Transformer (pulse and wideband)	Voltage and current transformation Impedance matching	Up to 500 MHz	High μ low hysteresis losses
NiZn	Antenna rod	Electromagnetic wave receival	Up to 15 MHz	High μQ, high resistivity
MnZn	Loading coil	Impedance loading	Audio	High μ high B_s, high stability of μ with temperature, time, and dc bias
		*Nonlinear B-H, medium-to-high flux density**		
MnZn, NiZn	Flyback transformer	Power converter	<100 kHz	High μ, high B_s, low hysteresis losses
MnZn	Deflection yoke	Electron-beam deflection	<100 kHz	High μ, high B_s
MnZn, NiZn	Suppression bead	Block unwanted ac signals	Up to 250 MHz	Moderately high μ, high B_s, high hysteresis losses
MnZn, NiZn	Choke coil	Separate ac from dc signals	Up to 250 MHz	Moderately high μ, high B_s, high hysteresis losses
MnZn, NiZn	Recording head	Information recording	Up to 10 MHz	High μ, high density, high μQ, high wear resistance
MnZn	Power transformer	Power converter	up to 200 kHz	High B_s, low hysteresis losses
		Microwave properties		
YIG†, MgMn‡, Li§, NiZn	Isolators, attenuators, circulators, switches, modulators	Impedance matching, power level control, power splitting	1–5 GHz (YIG), 2–30 GHz (YIG, MgMn, Li), 30–100 GHz (NiZn)	Controlled B_s, high resistivity, high Curie temperature, narrow resonance line–width
		Magnetic recording properties		
γ-Fe_2O_3, Co-γ-Fe_2O_3	Audio and video tapes, computer disks	Information storage		High B_s, high H_c

*After P. I. Slick, Ferrites for nonmicrowave applications, in E. P. Wohlfarth (ed.), *Ferromagnetic Materials*, vol. 2, North-Holland, 1980.
†May contain Al, Gd.
‡May contain Al, Zn.
§May contain Ti, Zn.

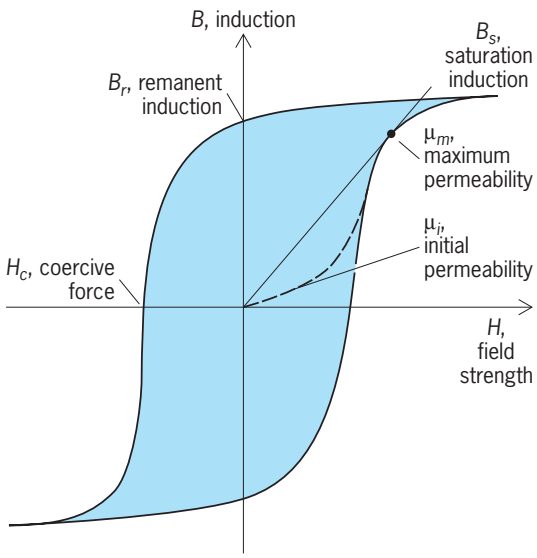

Plot of magnetic induction *B* as a function of magnetic field strength *H*.

where ω is the angular frequency, L the inductance, and R the effective series resistance arising from core loss. The higher the value of ωQ, the better is the material, and ferrites have the highest ωQ product of any commercially available magnetic material. Typical ωQ values measured at 20 kHz for various materials are iron dust, 2000; Permalloy powder, 10,000; 12.5-micrometer-thick molybdenum Permalloy tape, 100,000; and MnZn ferrite, 250,000. Values of ωQ greater than 10^6 at 20 kHz have been achieved commercially in specially prepared MnZn ferrites.

As compared with NiZn ferrites, MnZn ferrites have lower residual and hysteresis losses, higher permeabilities, and lower resistivities. A high permeability lowers the frequency of ferromagnetic resonance, which is accompanied by a large rise in losses. This factor, along with increased eddy current losses coming from lower resistivities, restricts the useful upper frequency range for MnZn ferrites in high-Q inductors to about 1 MHz. Above that, NiZn ferrites are preferred. *See* ELECTRIC FILTER; INDUCTOR; Q (ELECTRICITY); TRANSFORMER.

For transformer applications the highest value of ω over the operating frequency range is desired. Values of ω in the 18,000 range (10 kHz) for MnZn ferrite are commercially available, although values up to 40,000 have been obtained in the laboratory.

Nonlinear B-H devices. The largest usage of ferrite measured in terms of material weight is in the nonlinear *B-H* range, and is found in the form of deflecting yokes and flyback transformers for television receivers. The cores for these devices must have high saturation induction B_s along with high maximum permeability ω_m at the knee of the *B-H* curve to frequencies as high as 100 kHz, the effective flyback frequency used in scanning a television tube. Again, MnZn and NiZn ferrites dominate the use in these devices. *See* TELEVISION RECEIVER.

A rapidly growing use of ferrites is in the power area, where ferrite transformers are extensively used in switched mode (alternating current to direct current) and converter mode (direct current to direct current) power supplies. Such power supplies are widely used in various computer peripheral equipment and private exchange telephone systems. Here a large value of B_s together with low hysteresis losses is important. MnZn ferrites are superior to NiZn ferrites in both aspects. *See* ELECTRONIC POWER SUPPLY.

Although growth area for ferrites in the nonlinear *B-H* range is in recording heads. As compared with metals, such as Permalloy and Sendust, ferrites have higher electrical resistivity and wear resistance and are hence preferred in video and high-frequency recording applications.

Microwave devices. Microwave devices make use of the nonreciprocal propagation characteristics of ferrites close to or at a gyromagnetic resonance frequency in the range of 1–100 GHz. The most important of such devices are isolators and circulators. Materials having a range of B_s are needed for operation at various frequencies since for resonance, B_s is less than ω/γ where γ is the gyromagnetic ratio. In the 1–5-GHz range, yttrium iron garnet of $B_s = 0.02$–0.18 tesla is used. In the 2–30-GHz range, MgMn, MgMnZn, MgMnAl, and Li ferrites of $B_s = 0.06$–0.25 T are used along with the garnets. At 30–100 GHz, NiZn ferrites with B_s up to 0.50 T are used. The garnets have highly desirable, small, ferromagnetic-resonance linewidths, particularly in single-crystal form. In device development there is a strong trend toward realizing the conventional waveguide components in microstrip form whereby both the transmission and the gyromagnetic function are provided by a ferrimagnetic substrate, or the gyromagnetic function is provided by a ferrimagnetic insert on a nonmagnetic ferrite substrate. *See* GYRATOR; MICROWAVE.

Magnetic recording devices. Vast amounts of audio and video information and digital data from computers are stored in magnetic tapes and disks. Here magnetic recording materials function as hard magnetic materials with coercivity in the 25–100 kA/m (310–1250 oersteds) range. The materials are generally prepared in particulate form by chemical precipitation techniques and then dispersed in an organic slurry, either on a plastic substrate (tape or flexible disk) or on a rigid disk (usually an Al-Mg alloy). In a more advanced technique, continuous films are deposited by using sputtering or evaporation, particularly metal alloys rather than ferrites, on the smaller rigid disks. The most widely used particles in magnetic recording are γ-Fe_2O_3 and co-modified γ-Fe_2O_3. The basic γ-Fe_2O_3 is generally used in audio recording, while the higher-coercivity Co–γ-Fe_2O_3 [$H_c \simeq$ 50 kA/m (620 Oe) versus 25 kA/m (310 Oe) for γ-Fe_2O_3] dominates video recording. The hexagonal ferrites, Ba(or Sr)$Fe_{12}O_{19}$, are under study for use in magnetic recording. Since the easy axis of magnetization is normal to the plane of the particle platelets, such particles may be used as perpendicular recording media, in contrast to the normal

longitudinal mode. *See* COMPUTER STORAGE TECHNOLOGY.

Gilbert Y. Chin

Bibliography. J. Goldman, *Modern Ferrite Technology*, 1990; National Advisory Board, *Report of the Committee on Magnetic Materials*, Publ. NMAB-426, National Research Council, 1985; E. C. Snelling, *Soft Ferrites: Properties and Applications*, 2d ed., 1988; E. C. Snelling and A. D. Giles, *Ferrites for Inductors and Transformers,* 1983; G. Winkler, *Magnetic Garnets*, 1981; E. P. Wohlfarth (ed.), *Ferromagnetic Materials*, vol. 2, 1980.

Ferroalloy

An important group of metallic raw materials required for the steel industry. Ferroalloys are the principal source of such additions as silicon, Si, and manganese, Mn, which are required for even the simplest plain-carbon steels; and chromium, Cr, vanadium, V, tungsten, W, titanium, Ti, and molybdenum, Mo, which are used in both low- and high-alloy steels. Also included are many other more complex alloys. Ferroalloys are unique in that they are brittle and otherwise unsuited for any service application, but they are important as the most economical source of these elements for use in the manufacture of the engineering alloys. These same elements can also be obtained, at much greater cost in most cases, as essentially pure metals. The ferroalloys contain significant amounts of iron and usually have a lower melting range than the pure metals and are therefore dissolved by the molten steel more readily than the pure metal. In other cases, the other elements in the ferroalloy serve to protect the critical element against oxidation during solution and thereby give higher recoveries. Ferroalloys are used both as deoxidizers and as a specified addition to give particular properties to the steel. *See* STEEL MANUFACTURE.

Many ferroalloys contain combinations of two or more desirable alloy additions, and well over 100 commercial grades and combinations are available. Although of less general importance, other sources of these elements for steelmaking are metallic nickel, Ni, silicon carbide, molybdic oxide, and even misch metal (a mixture of rare earths). Analyses of a few typical ferroalloys are given in the **table**.

The three ferroalloys which account for the major tonnage in this class are the various grades of silicon, manganese, and chromium. For example, 13 lb (5.9 kg) of manganese is used on the average in the United States for every ton of open-hearth steel produced. Elements supplied as ferroalloys are among the most difficult metals to reduce from ore.

The most common grade of ferromanganese is a blast-furnace product (standard ferromanganese), the major variation from pig iron production being the use of manganese-rich ore. Other grades of ferromanganese with low C are made by the processes described in the next sections. A low-Si ferroalloy can be made in the same way, but the tonnage grades are made by other methods.

The most general method of ferroalloy manufacture is the submerged-arc furnace (see **illus.**). Its use will be described in connection with the production of 50% ferrosilicon. A modern furnace of this type is 26 ft (7.9 m) in diameter and 10 ft (3 m) deep with the three carbon electrodes supplying an average of 5 kWh per pound (2.4×10^8 joules per kilogram) of silicon produced. The furnace produces about 2 tons (1.8 metric tons) of alloy per hour. The furnace is charged intermittently from the top with a mixture of the required amounts of SiO_2 (quartzite rock), C (coke), and Fe (as turnings), and the ferrosilicon is periodically tapped from the bottom. The overall operation is essentially continuous, like that of a blast furnace, but the heat is supplied by the resistance of the charge and some arcing, and the coke serves primarily as a reducing agent. After solidification in flat molds, the ferroalloy is crushed to specified sizes for delivery. The overall reaction for the process can be

Analysis of typical ferroalloys, % weight

Type of ferroalloy*	Mn	Si	C	Cr	Mo	Al	Ti	V
Ferromanganese								
Standard	78–82	1.25	7.5					
Medium carbon	80–85	1.25–2.5	1–3†					
Low carbon	80–85	1.25–7.0	0.75					
Ferrosilicon								
50% regular	—	47–52	0.15					
75% regular	—	73–78	0.15					
Ferrochromium								
High carbon	—	1–2	4.5–6.0†	67–70				
Low carbon	—	0.3–1.0	0.03–2.0†	68–71				
SM low carbon	4–6	4–6	1.25	62–65				
Ferromolybdenum								
High carbon	—	1.5	2.5	—	55–70			
Ferrovanadium								
High carbon	—	13.0	3.5	—	—	1.5	—	30–40
Ferrotitanium								
Low carbon	—	3–5	0.1	—	—	6–10	38–43	

* In all cases the balance is Fe, with the exception of minor impurities. The latter are usually specified, such as 0.10% max P.
† In several specified grades within this range.

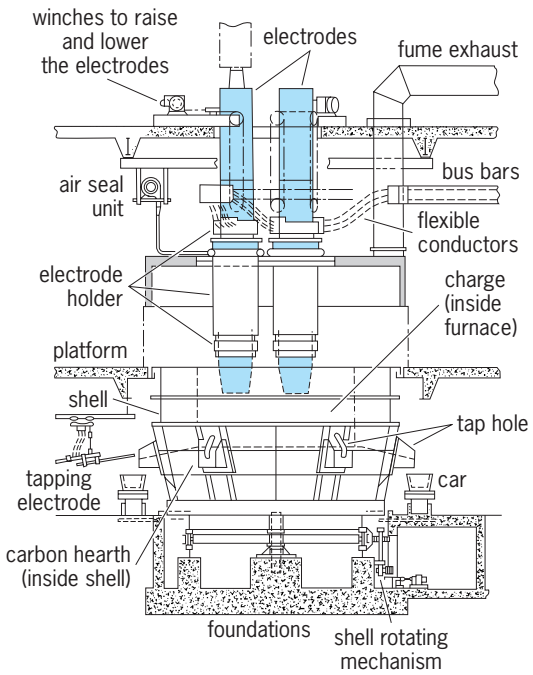

Submerged-arc furnace used in ferroalloy manufacture. (*Lectromelt Furnace Division*, *McGraw-Edison Co.*)

written as (1). The Si reduced in this way dissolves in

$$SiO_2 + C \rightarrow Si + 2CO \qquad (1)$$

the molten iron which is also present and simplifies the operation of the process.

Because SiO_2 is present as gangue in many ores, and because silicate slags have sufficiently low melting points to be controlled conveniently, many other ferroalloys contain appreciable amounts of Si when they are manufactured by this same process. The inherent low solubility of C in Si alloys makes this combination desirable when Si is not objectionable. Low-silicon grades can be made, however, by further addition of metal oxide as illustrated by the general reaction (2). Thus ferrosilicon can be used as an

$$2MO + Si \rightarrow SiO_2 + 2M \qquad (2)$$

intermediate to produce other ferroalloys, and this scheme is common.

Likewise, aluminum can be used as the reducing agent by a process called the thermit reaction, as indicated in (3), in which all or part of the heat required

$$2Al + Fe_2O_3 \rightarrow 2Fe + Al_2O_3 + heat \qquad (3)$$

results from the reduction reaction. The oxide of various desirable alloys may be substituted for Fe_2O_3, and Ca or Mg may be used for Al, a variety of such combinations being possible. The product is low in both Si and C.

Thus, depending upon the reducibility of the ore, and the amounts of C, Si, and Fe that can be tolerated in the product, as well as economic considerations, ferroalloys are produced in the blast furnace, or in the submerged-arc or similar variations of the electric furnace, or by aluminothermic or silicothermic reactions. These ferroalloys are then supplied to the steel industry for use as deoxidizers or alloy-addition agents. *See* ARC HEATING; IRON ALLOYS; MANGANESE; MOLYBDENUM; SILICON; STEEL; VANADIUM.

Gerhard Derge

Bibliography. P. D. Deeley et al., *Ferroalloys and Alloying Handbook*, 1981; R. Serjeantson, *Ferro-Alloy Directory and Databook*, 3d ed., 1992; Y. S. Touloukian and C. Y. Ho, *Properties of Selected Ferrous Alloying Elements*, vols. 1–3, 1989.

Ferroelectrics

Crystalline substances which have a permanent spontaneous electric polarization (electric dipole moment per cubic centimeter) that can be reversed by an electric field. In a sense, ferroelectrics are the electrical analog of the ferromagnets, hence the name. The spontaneous polarization is the so-called order parameter of the ferroelectric state, just as the spontaneous magnetization is the order parameter of the ferromagnetic state. The names Seignette-electrics or Rochelle-electrics, which are also widely used, are derived from the name of the first substance found to have this property, Seignette salt or Rochelle salt. *See* FERROMAGNETISM.

The reversibility of the spontaneous polarization is due to the fact that the structure of a ferroelectric crystal can be derived from a nonpolarized structure by small displacements of ions. In most ferroelectric crystals, this nonpolarized structure becomes stable if the crystal is heated above a critical temperature, the ferroelectric Curie temperature; that is, the crystal undergoes a phase transition from the polarized phase (ferroelectric phase) into an unpolarized phase (paraelectric phase). The change of the spontaneous polarization at the Curie temperature can be continuous or discontinuous. The Curie temperature of different types of ferroelectric crystals range from a few degrees absolute to a few hundred degrees absolute. As a rule, the ferroelectric phase is the low-temperature phase; however, there are crystals which are ferroelectric in a relatively narrow temperature range only, and others stay polarized up to the temperature of decomposition or melting.

Classification. From a practical standpoint ferroelectrics can be divided into two classes. In ferroelectrics of the first class, spontaneous polarization can occur only along one crystal axis; that is, the ferroelectric axis is already a unique axis when the material is in the paraelectric phase. Typical representatives of this class are Rochelle salt, KH_2PO_4, $(NH_4)_2SO_4$, guanidine aluminum sulfate hexahydrate, glycine sulfate, colemanite, and thiourea.

In ferroelectrics of the second class, spontaneous polarization can occur along several axes that are equivalent in the paraelectric phase. The following substances, which are all cubic above the Curie point, belong to this class: $BaTiO_3$-type (or perovskite-type) ferroelectrics; $Cd_2Nb_2O_7$; $PbNb_2O_6$;

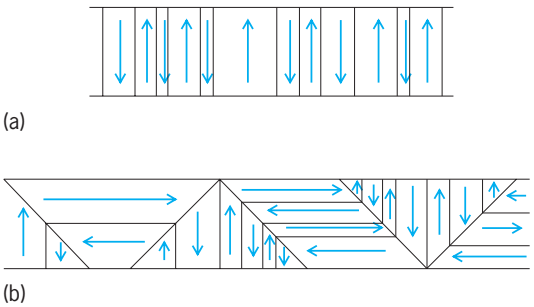

Fig. 1. Domain configurations (simplified) encountered in ferroelectric crystals. (*a*) First class. (*b*) Second class.

certain alums, such as methyl ammonium alum; and $(NH_4)_2Cd_3(SO_4)_3$. Some of the $BaTiO_3$-type ferroelectrics have, below the Curie temperature, additional transition temperatures at which the spontaneous polarization switches from one crystal axis to another crystal axis. For example, $BaTiO_3$ and $KNbO_3$ polarize with decreasing temperature first along a [100] axis, then the polarization switches into a [110] axis, and finally into a [111] axis.

From a scientific standpoint, one can distinguish proper ferroelectrics and improper ferroelectrics. In proper ferroelectrics, for example, $BaTiO_3$, KH_2PO_4, and Rochelle salt, the spontaneous polarization is the order parameter. The structure change at the Curie temperature can be considered a consequence of the spontaneous polarization. The unit cell of the crystal in the ferroelectric phase contains the same number of chemical formula units as the unit cell in the paraelectric phase. In improper ferroelectrics, the spontaneous polarization can be considered a by-product of another structural phase transition. The unit cell in the ferroelectric phase is an integer multiple of the unit cell in the paraelectric phase. Examples of such systems are $Gd(MoO_4)$ and boracites. The dielectric elastic and electromechanical behavior of the two types of ferroelectrics differ significantly.

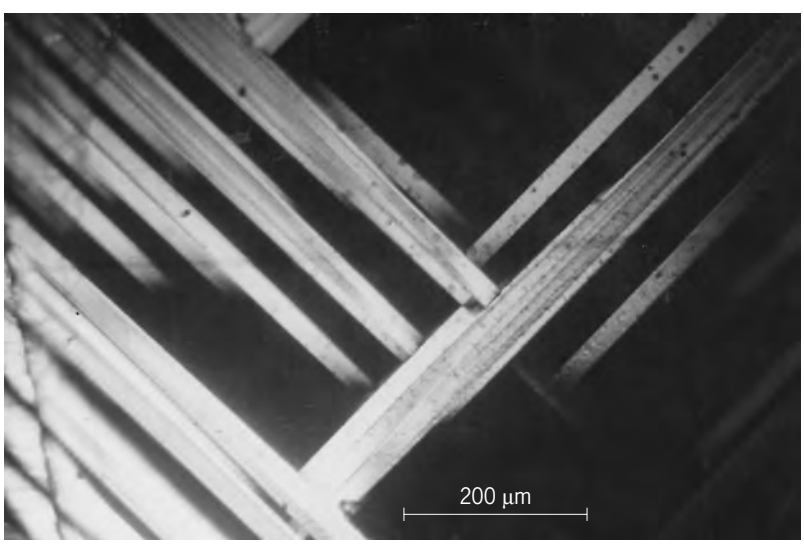

Fig. 2. Ferroelectric domains in $BaTiO_3$ photographed through a polarizing microscope. Ferroelectric domains range from macroscopic to submicroscopic size.

Ferroelectric domains. The spontaneous polarization can occur in at least two equivalent crystal directions; thus, a ferroelectric crystal consists in general of regions of homogeneous polarization that differ only in the direction of polarization. These regions are called ferroelectric domains. Ferroelectrics of the first class consist of domains with parallel and antiparallel polarization (**Fig. 1***a*), whereas ferroelectrics of the second class can assume much more complicated domain configurations (Fig. 1*b*). The region between two adjacent domains is called a domain wall. Within this wall, the spontaneous polarization changes its direction. The wall between antiparallel domains is probably only a few lattice spacings thick, whereas the wall between domains polarized at a right angle to each other is probably thicker. Ferroelectric domains can be observed in a number of substances by means of the polarizing microscope (**Fig. 2**) because of their birefringence, or double refraction. The ferroelectric domains range in size from macroscopic (millimeters) to submicroscopic. *See* BIREFRINGENCE.

Ferroelectric hysteresis. When an electric field is applied to ferroelectric crystal, the domains that are favorably oriented with respect to this field grow at the expense of the others, for example, by sidewise motion of domain walls. In addition, favorably oriented domains can nucleate and grow until the whole crystal becomes one single domain. When the field is reversed, the polarization reverses through the same processes. The relation between the resulting polarization P of the whole crystal and the externally applied electric field E is given by a hysteresis loop (**Fig. 3**). The shape of the hysteresis loop depends strongly upon the perfection of the crystal as well as upon the rate of change of the externally applied field E. A simple circuit that permits observation of ferroelectric hysteresis loops with an oscilloscope is shown in **Fig. 4**. In some ferroelectrics, polarization can be reversed within a fraction of 1 microsecond.

Spontaneous polarization. The magnitude of the permanent or spontaneous polarization P_s of a domain can be obtained from the hysteresis loop by extrapolating the saturation branch to zero external field (Fig. 3). For most ferroelectrics, the values of P_s are between 10^{-7} and 10^{-4} coulomb/cm^2 (**Fig. 5**). In nonferroelectric dielectrics, electric fields between 10^5 and 10^8 V/cm would be necessary in order to achieve such large polarizations.

Dielectric properties. As a rule, the dielectric constant ϵ measured along a ferroelectric axis increases in the paraelectric phase when the Curie temperature is approached. In many ferroelectrics, this increase can be approximated by the Curie-Weiss law, shown in the equation below. Here T designates the

$$\epsilon = \frac{C}{T - T_0}$$

temperature of the crystal, and T_0 is equal to or somewhat smaller than the transition temperature. C is the so-called Curie constant. For $BaTiO_3$, this law holds

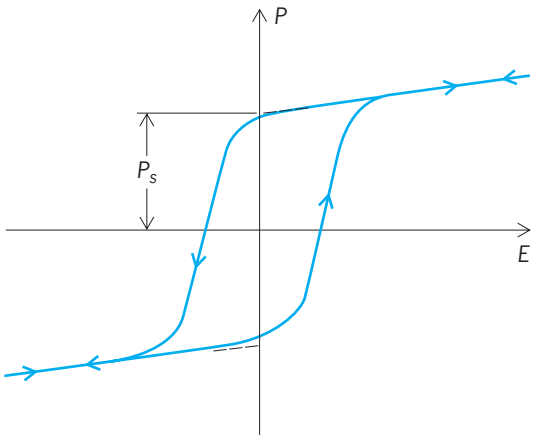

Fig. 3. Net polarization *P* of a ferroelectric crystal versus externally applied electric field *E*.

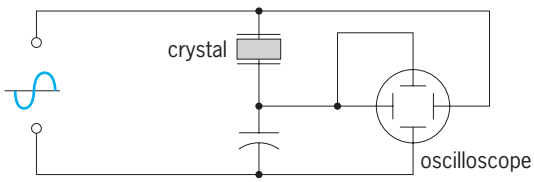

Fig. 4. Circuit for the display of ferroelectric hysteresis loops on an oscilloscope.

unaltered up to frequencies of 2.4×10^{10} Hz. Dispersion sets in the far-infrared. The dielectric constant drops when the crystal becomes spontaneously polarized (**Fig. 6**). In the ferroelectric phase, the dielectric constant has two components. The first component is the dielectric constant of the individual domains. It is independent of the frequency and of the electric field generally up to far-infrared frequencies. The second component is due to domain wall motions, that is, to partial reversal of the spontaneous polarization. This process can give rise to large dielectric losses, and it depends strongly upon the frequency, the electric field strength, the domain structure, and the temperature. In uniaxial ferroelectrics, the dielectric constant measured perpendicular to the ferroelectric axis generally does not show a very pronounced anomaly near the Curie temperature, and in some cases it has even the same order of magnitude and temperature dependence as for any normally behaving dielectric crystal. *See* CURIE-WEISS LAW; DIELECTRIC MATERIALS; PERMITTIVITY.

Piezoelectric properties. Ferroelectrics can be divided into two groups according to their piezoelectric behavior.

The ferroelectrics in the first group are already piezoelectric in the unpolarized phase. Those piezoelectric moduli which relate stresses to polarization along the ferroelectric axis have essentially the same temperature dependence as the dielectric constant along this axis, and hence become very large near the Curie point. The spontaneous polarization gives rise to a large spontaneous piezoelectric strain which is proportional to the spontaneous polarization. In KH_2PO_4-type ferroelectrics and in Rochelle salt, for example, this strain is a shear in the plane perpendicular to the axis of polarization. It reaches 27 min of arc in KH_2PO_4 and about 1.8 min of arc in Rochelle salt. The piezoelectric modulus decreases as the spontaneous polarization increases. But with sufficiently large stresses, it is possible to align the domains and reverse the spontaneous polarization (**Fig. 7**). The relation between the resulting polarization of the whole crystal and the mechanical stress is given by a hysteresis loop analogous to the loop of Fig. 3 (piezoelectric hysteresis). This effect can simulate a very large piezoelectric modulus.

The ferroelectrics in the second group are not piezoelectric when they are in the paraelectric phase. However, the spontaneous polarization lowers the symmetry so that they become piezoelectric in the polarized phase. This piezoelectric activity is often hidden because the piezoelectric effects of the

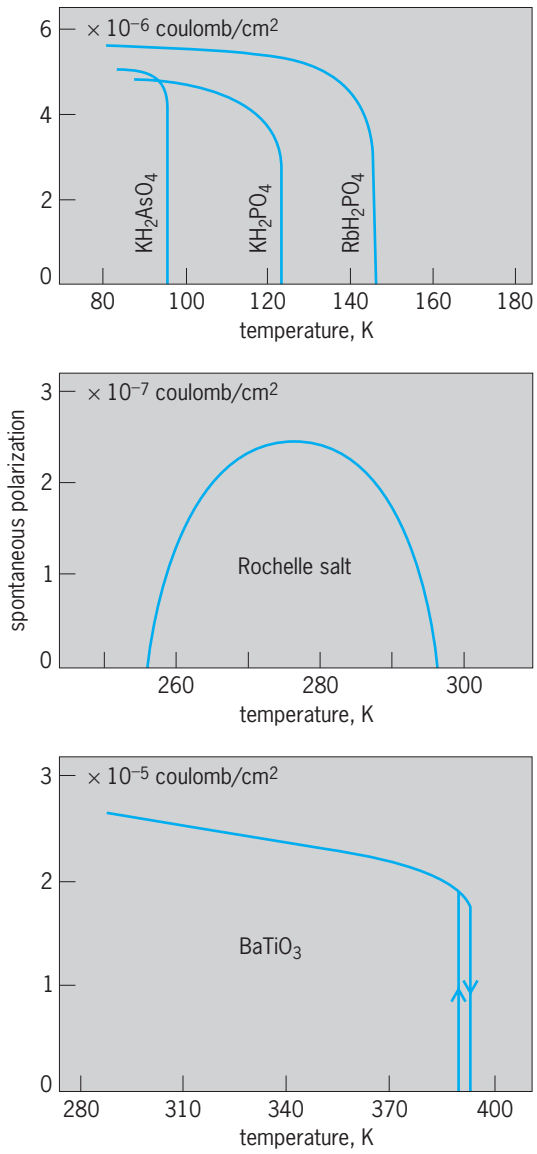

Fig. 5. Dependence upon temperature of the spontaneous polarization of some ferroelectrics. °F = (K × 1.8) − 459.67. °C = K − 273.15.

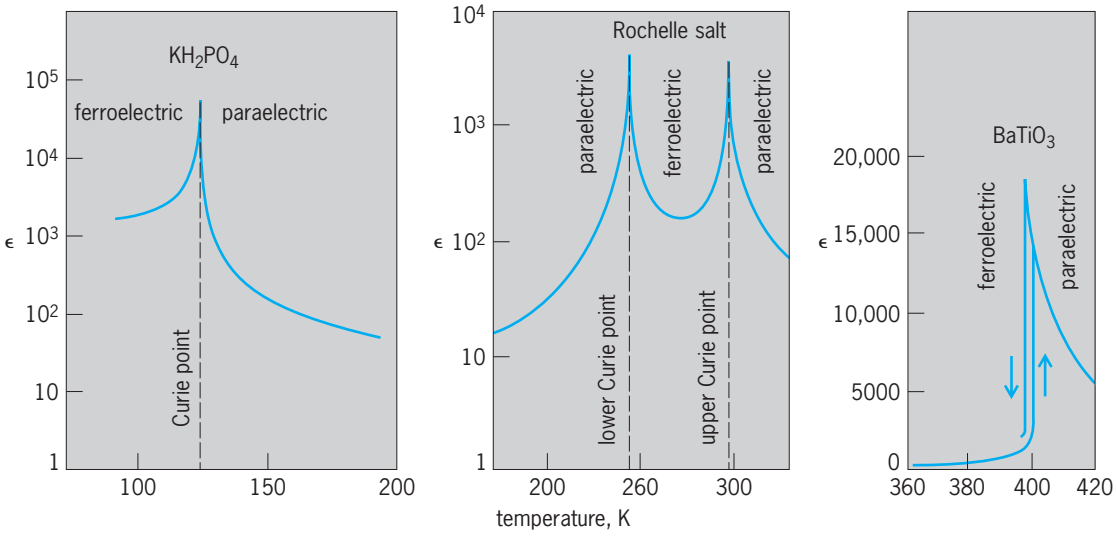

Fig. 6. Anomalous temperature dependence of relative dielectric constant of ferroelectrics at transition temperature. °F = (K × 1.8) − 459.67. °C = K − 273.15.

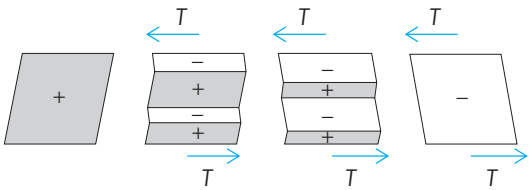

Fig. 7. Schematic representation of the reversal of the spontaneous polarization by a mechanical shear stress *T* in KH_2PO_4 and Rochelle salt.

various domains can cancel. However, strong piezoelectric activity of a macroscopic crystal or even of a polycrystalline sample occurs when the domains have been aligned by an electric field. The spontaneous strain is proportional to the square of the spontaneous polarization. In $BaTiO_3$, for example, the crystal (which has cubic symmetry in the unpolarized phase) expands along the axis of polarization and contracts at right angles to it. The strain is of the order of magnitude of 1%. The spontaneous polarization cannot be reversed by a mechanical stress in ferroelectrics of this group. *See* PIEZOELECTRICITY.

Crystal structure. The structures of different types of ferroelectrics are entirely different, and it is not possible to establish a general rule for the occurrence of ferroelectricity. The structures of a number of ferroelectrics and the minute changes that they undergo when spontaneous polarization occurs are known in great detail from x-ray diffraction and neutron diffraction studies. In a qualitative way, the process of polarization is best understood for ferroelectrics of the $BaTiO_3$ type. **Figure 8** shows schematically the structure of the unit cell of a $BaTiO_3$ crystal in the unpolarized state, and the arrows indicate the direction in which the ions are slightly displaced when the lattice becomes spontaneously polarized along the axis z. The order of magnitude of the displacement is 1% of the unit cell dimension. However, these displacements do not account quantitatively for the observed polarization, because other changes of the electronic structure occur as well. *See* CRYSTAL STRUCTURE.

In KH_2PO_4-type ferroelectrics, hydrogen bonds O—H · · · O play an important part in the ferroelectric effect. Above the Curie temperature, the hydrogen ions are statistically distributed over the two possibilities O—H · · · O and O · · · H—O, whereas below the Curie point, one or the other of these two possibilities is strongly favored, depending upon the sign of the spontaneous polarization.

Antiferroelectric crystals. These materials are characterized by a phase transition from a state of lower symmetry (generally low-temperature phase) to a state of higher symmetry (generally high-temperature phase). The low-symmetry state can be regarded as a slightly distorted high-symmetry state. It has no permanent electric polarization, in contrast to ferroelectric crystals. The crystal lattice can be

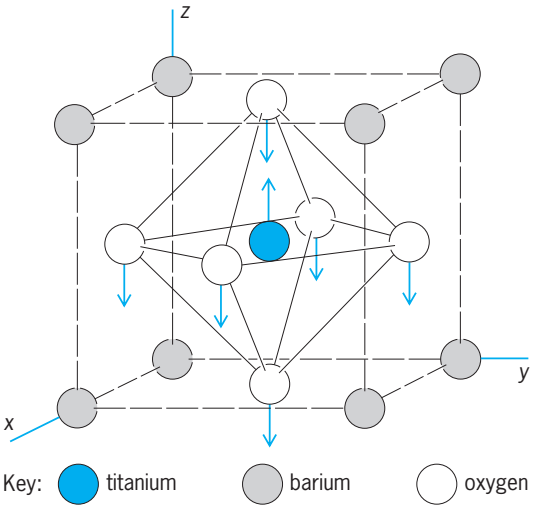

Fig. 8. Crystal lattice of $BaTiO_3$. Arrows indicate displacements of the ions when crystal becomes polarized.

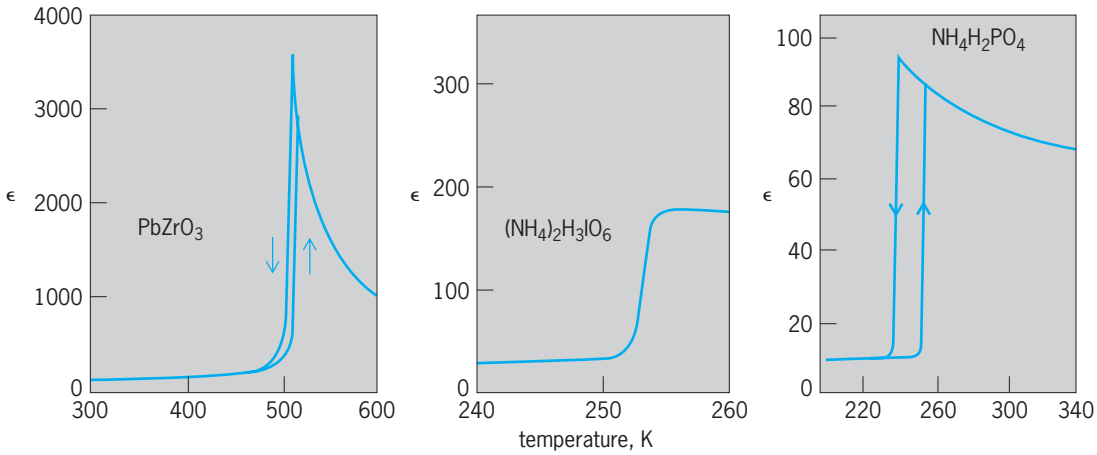

Fig. 9. Anomalous temperature dependence of relative dielectric constant of antiferroelectric crystals. Note increase as transition temperature is approached and drop when antipolarization occurs. °F = (K × 1.8) − 459.67. °C = K − 273.15.

regarded as consisting of two interpenetrating sublattices with equal but opposite electric polarization. This state is referred to as the antipolarized state.

In a certain sense, an antiferroelectric crystal is the electrical analog of an antiferromagnetic crystal. In the high-symmetry phase, the sublattices are unpolarized and indistinguishable. In general, antiferroelectric crystals have more than one axis along which the sublattices can polarize. Therefore, the low-symmetry phase consists of regions of homogeneous antipolarization which differ only in the orientation of the axis along which antipolarization has occurred. These regions are called antiferroelectric domains and can be observed by the polarizing microscope. Because these domains have no permanent electric dipole moment, an electric field generally has little influence on domain structure. *See* ANTIFERROMAGNETISM.

The dielectric constant of antiferroelectric crystals is generally larger than it is for nonferroelectric crystals and has an anomalous temperature dependence. It increases as the transition temperature is approached and drops when antipolarization occurs (**Fig. 9**). In some antiferroelectrics the phase transition is discontinuous; in others it is continuous.

The structure of antiferroelectric crystals is generally closely related to the structure of ferroelectric crystals. Some antiferroelectrics even undergo phase transitions from an antipolarized state into a spontaneously polarized, ferroelectric state; in others a sufficiently strong electric field applied along an antiferroelectric axis reverses the polarity of one of the sublattices so that a ferroelectric state results. The crystal reverts, however, to the antiferroelectric state when the electric field is removed. **Figure 10** shows net polarization versus externally applied field for such a case.

Compounds with antiferroelectric properties are $PbZrO_3$, $PbHfO_3$, $NaNbO_3$ (isomorphous with ferroelectric $BaTiO_3$), WO_3 (structure related to $BaTiO_3$), $NH_4H_2PO_4$ and isomorphous NH_4 salts (isomorphous with ferroelectric KH_2PO_4), $(NH_4)_2H_3IO_6$, $Ag_2H_3IO_6$, and certain alums.

Origin of phase transition. The ferroelectric phase transition results from an instability of one of the normal lattice vibration modes. On approaching the transition temperature, the frequency of the relevant normal mode decreases (soft mode). The restoring force of the mode displacements tends to zero. When the stability limit is reached, the displacements corresponding to the soft mode freeze in, and the ferroelectric phase results. The ferroelectric soft mode is polar (infrared-active) and of infinite wavelength. The antiferroelectric phase transition, on the other hand, emerges from a soft lattice mode with a finite wavelength equal to an integer multiple of a lattice period.

Applications. The piezoelectric effect of ferroelectrics (and certain antiferroelectrics) finds numerous applications in electromechanical transducers. The large electrooptical effect (birefringence

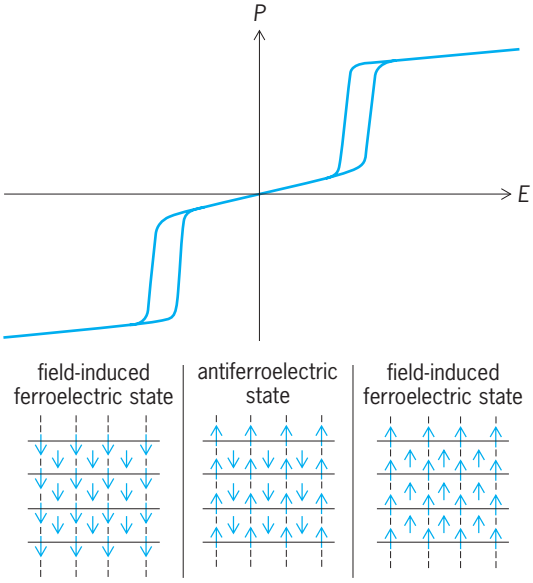

Fig. 10. Polarization *P* of antiferroelectric $PbZrO_3$ versus externally applied electric field *E*. Strong fields "switch" the antiferroelectric crystal into a ferroelectric state, as shown here schematically.

induced by an electric field) is used in light modulators. In some ferroelectrics (such as $BaTiO_3$, $LiNbO_3$, $KTaNbO_3$, and $LiTaO_3$), light can induce changes of the refractive indices. These substances are used for optical information storage and in real-time optical processors. The temperature dependence of the spontaneous polarization corresponds to a strong pyroelectric effect that can be exploited in thermal and infrared sensors. Werner Kanzig

Bibliography. J. M. Herbert, *Ferroelectric Transducers and Sensors*, 1982; F. Jona and G. Shirane, *Ferroelectric Crystals*, 1962, reprint 1993; M. E. Lines and A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials*, 1977; T. Mitsui, *An Introduction to the Physics of Ferroelectrics*, 1976; A. Moulson and J. M. Herbert, *Electroceramics: Materials, Properties, Applications*, 1992; N. Setter and E. L. Colla (eds.), *Ferroelectric Ceramics*, 1992; G. A. Smolenskii, *Ferroelectrics and Related Materials*, 1984; X. Yuhuan, *Ferroelectric Materials and Their Applications*, 1991.

Ferrofluid

A type of magnetic fluid that consists of a colloidal suspension of nanoscopic magnetic particles in a carrier liquid. A ferrofluid behaves as a fluid that is attracted to an external magnetic field without a large change in its viscosity.

Magnetic fluids are liquids that respond in many interesting ways to external magnetic fields. Most magnetic fluids combine conventional liquids (oil or water) with very small magnetic particles. Magnetorheological fluids are oils loaded with micrometer-sized magnetic particles. These materials experience a large change in viscosity, going from a thick fluid to a peanut butter consistency, when exposed to a magnetic field. Applications include computer-controlled vehicle suspension systems, where the damping in shock absorbers is controlled by magnetically regulating the magnetorheological fluid viscosity inside the shock absorber. Ferrofluids are similar to magnetorheological fluids except the particle sizes are 2–3 orders of magnitude smaller. These magnetic particles are so small (~10 nanometers in diameter) that thermal agitation in the fluid is great enough to keep the particles in suspension. One significant difference between magnetorheological fluids and ferrofluids is that ferrofluids behave as a colloidal suspension. The particles are so small and finely divided that they will not settle under the action of gravity or magnetic fields. Unlike magnetorheological fluids, ferrofluids experience a much smaller change in viscosity when exposed to magnetic fields. *See* COLLOID; MAGNETIC FIELD.

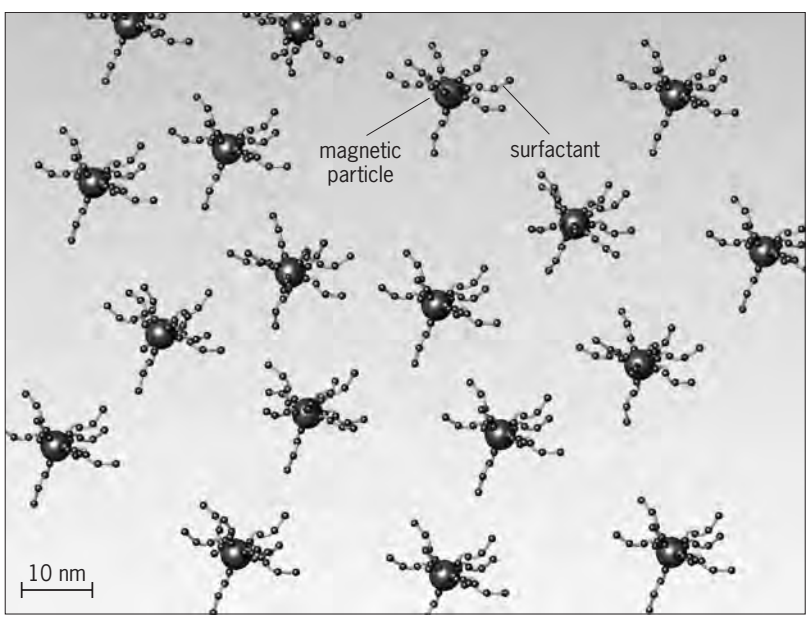

Fig. 1. Basic composition of a ferrofluid—magnetic nanoparticles coated with surfactant in a carrier fluid.

Composition. Ferrofluids consist of three basic components: a carrier fluid, magnetic nanoparticles, and a surfactant covering the nanoparticle (**Fig. 1**). Carrier fluids are generally organic solvents or water. Oil-based ferrofluids are used in applications such as magnetic seals, heat conduction for loudspeakers, and damping in stepper motors. Water-based ferrofluids are finding many useful biomedical applications such as magnetic cell sorting and biochemical sensing. The most common type of magnetic nanoparticle is magnetite (Fe_3O_4). The particle size is so small that the particle is considered superparamagnetic, meaning that there is only one magnetic domain in the particle. When an external field is applied to the fluid, the particles line up along the same field lines (**Fig. 2**). When the field is removed, thermal agitation in the fluid is great enough to cause the particles to resume their random orientation. The primary role of the surfactant is to aid in dispersing the particles within the carrier fluid. One end of the surfactant is absorbed on the particle surface, while the other interacts with the fluid and other particles. *See* MAGNETITE; NANOPARTICLES; SOLVENT; SURFACTANT.

There are a number of different methods for creating magnetite nanoparticles. The most common methods are ball grinding and chemical coprecipitation. In ball grinding, small particles are made from larger ones; whereas in chemical coprecipitation, nanoparticles are made in solution. Ball grinding, or size reduction, is a wet grinding process in which larger magnetite particles are mixed with a surfactant and rubbed together for long periods (~1000 hours). Chemical coprecipitation is based on mixing solutions of divalent and trivalent iron salts. As sodium hydroxide is dripped into the solution, magnetite nanoparticles precipitate out. This process requires careful control of the solution concentration, temperature, and pH. *See* PRECIPITATION (CHEMISTRY).

Applications. Most applications of ferrofluids exploit the combined fluid and magnetic behavior of the fluid. Most electromagnetic devices (motors, speakers, generators) are thermally limited. Some high-quality speakers use ferrofluids to improve

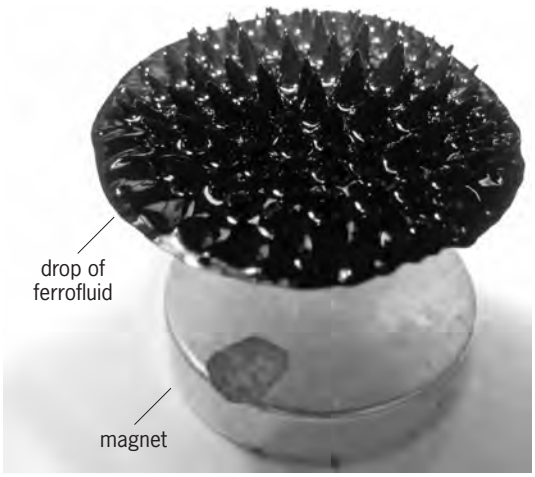

Fig. 2. Ferrofluid drop exposed to a magnetic field showing spikes due to magnetic field lines.

performance. The ferrofluid is captured magnetically in the air gap of the voice coil. The fluid serves two purposes: increased thermal conductivity to provide greater power capacity (reduce magnet and coil size) and mechanical damping to reduce harmonic distortion. Many stepper motors and optical disc actuators for CD and DVD players experience high-frequency vibration or ringing. A small volume of ferrofluid in the air gap of these actuators provides a very simple means of damping high-frequency vibration without additional design features. Ferrofluid seals, based on magnetically containing a small volume of ferrofluid between two concentric magnetic components, provides a liquid O-ring that serves as a hermetic seal with no friction and long life. *See* LOUDSPEAKER; STEPPING MOTOR. Lonnie J. Love

Bibliography. B. Berkovskii and V. Bashtovoi, *Magnetic Fluids and Applications Handbook*, Begell House, 1996; E. Blums, A. Cerbers, and M. Maiorov, *Magnetic Fluids*, Walter de Gruyter, 1997; R. E. Rosensweig, *Ferrohydrodynamics*, Dover, 1985.

Ferromagnetism

A property exhibited by certain metals, alloys, and compounds of the transition (iron group), rare-earth, and actinide elements in which, below a certain temperature called the Curie temperature, the atomic magnetic moments tend to line up in a common direction. Ferromagnetism is characterized by the strong attraction of one magnetized body for another, a phenomenon known before 600 B.C.

Atomic magnetic moments arise when the electrons of an atom possess a net magnetic moment as a result of their angular momentum. The combined effect of the atomic magnetic moments can give rise to a relatively large magnetization, or magnetic moment per unit volume, for a given applied field. Above the Curie temperature, a ferromagnetic substance behaves as if it were paramagnetic: Its susceptibility approaches the Curie-Weiss law. The Curie temperature marks a transition between order and disorder of the alignment of the atomic magnetic moments. Some materials having atoms with unequal moments exhibit a special form of ferromagnetism below the Curie temperature called ferrimagnetism. *See* CURIE TEMPERATURE; CURIE-WEISS LAW; ELECTRON SPIN; FERRIMAGNETISM; MAGNETIC SUSCEPTIBILITY; PARAMAGNETISM.

The characteristic property of a ferromagnet is that, below the Curie temperature, it can possess a spontaneous magnetization in the absence of an applied magnetic field. Upon application of a weak magnetic field, the magnetization increases rapidly to a high value called the saturation magnetization, which is in general a function of temperature. For typical ferromagnetic materials, their saturation magnetizations, and Curie temperatures, *See* MAGNETIZATION.

The tasks of a theory of ferromagnetism are to account for the spontaneous magnetization below the Curie point, the temperature dependence of the saturation magnetization, and the nature of the magnetization process, or magnetization curve.

Weiss theory. The Weiss molecular field theory of ferromagnetism (P. Weiss, 1907) represents the first realistic attempt to account for the properties of a ferromagnet. This theory rests on two hypotheses.

1. Below the Curie point, a ferromagnetic substance is composed of small, spontaneously magnetized regions called domains. The total magnetic moment of the material is the vector sum of the magnetic moments of the individual domains. It is now known that these assumed domains really exist and are usually between 0.01 and 0.1 cm wide.

2. Each domain is spontaneously magnetized because a strong internal magnetic interaction tends to align the individual atomic magnetic moments within the domain.

The consequence of these assumptions is that, while each domain is spontaneously magnetized, the directions of magnetization of the domains do not coincide; therefore the overall magnetization of the sample may be much smaller than if it were composed of a single domain. Application of a relatively weak field of the order of 10–100 kiloamperes/m or 100–1000 oersteds (and often very much less) is sufficient to align the directions of magnetization of the domains, thereby achieving a large magnetization.

The second hypothesis of the Weiss theory leads to the existence of a Curie temperature below which a domain may be spontaneously magnetized in the absence of an applied magnetic field. This comes about in the following way. If the domain is spontaneously magnetized, there must be some sort of interaction between the atomic magnetic moments which tends to align them. Otherwise the domain would behave paramagnetically. The average strength of this interaction may be represented by an internal magnetic field, the Weiss molecular field, which is proportional to the magnetization of the domain. Thus the effective field acting on any atomic magnetic moment within the domain may be written as Eq. (1),

where H_0 is an externally applied magnetic field and

$$H = H_0 + \lambda M \tag{1}$$

λM (M = magnetization) is the Weiss molecular field. For iron $M \simeq 135$ kA/m (1700 Oe) and $\lambda \simeq 5000$, so that $\lambda M \simeq 10^9$ A/m (10^7 Oe), a huge value. It is relatively easy to deduce the magnetic susceptibility above the Curie point since it is known that the ferromagnetic substance behaves like a paramagnet above the Curie temperature. If the Curie law holds, Eq. (2) is valid, where C is the Curie constant. In the

$$\frac{M}{H} = \frac{C}{T} \tag{2}$$

Curie law, H is taken to be the effective magnetic field acting on an atomic magnetic moment, Eq. (1). From Eqs. (1) and (2), Eq. (3) is obtained. The susceptibil-

$$\frac{M}{H + \lambda M} = \frac{C}{T} \tag{3}$$

ity χ is the magnetization per unit applied field H_0, so that from Eq. (3), in the electromagnetic system of units, Eq. (4) is derived. The Curie temperature

$$\chi = \frac{M}{H_0} = \frac{C}{T - C\lambda} = \frac{C}{T - T_c} \tag{4}$$

is defined by $T_c = C\lambda$, and the susceptibility follows the Curie-Weiss law above the Curie temperature. The form of the Curie-Weiss law leads to a nonzero magnetization when $H_0 = 0$ at T_c.

Below the Curie point, where the Curie-Weiss law breaks down, the Weiss theory leads to a spontaneous magnetization and also predicts the temperature dependence of the saturation magnetization. This comes about if it is taken into account that the Curie law is only an approximation valid for weak fields or high temperatures; it does not allow for saturation effects. The correct quantum-mechanical expression to replace Eq. (2) is given by Eqs. (5) and (6), where N is the number of atoms per unit volume,

$$M = NJg\mu_B B_J(a^*) \tag{5}$$

$$a^* = \frac{g\mu_B(H_0 + \lambda M)J}{kT} \tag{6}$$

each with angular momentum quantum number J, g is the spectroscopic splitting factor (the measure of the energy level splittings), μ_B is the Bohr magneton, k is Boltzmann's constant, T is the absolute temperature, and $B_J(a^*)$ is the Brillouin function. For $T < T_c$, the value of the magnetization for an applied field of H_0 is obtained by solving Eqs. (5) and (6) simultaneously. In particular, the spontaneous magnetization is obtained by setting $H_0 = 0$ in Eq. (6). *See* GYROMAGNETIC EFFECT; GYROMAGNETIC RATIO; MAGNETON.

The solution for the spontaneous magnetization ($H_0 = 0$) may be obtained graphically for any temperature T, as shown in **Fig. 1**. The broken line of Fig. 1 represents the line $M = a^*kT_c/g\mu_B J\lambda$ and is the line for the largest value of T for which there is a solution. This is the Curie point, and for $T > T_c$ there is no spontaneous magnetization. The value

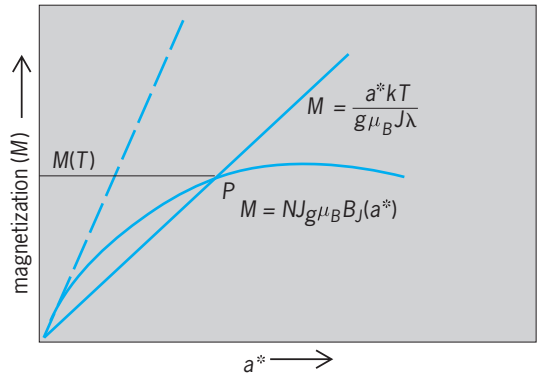

Fig. 1. Spontaneous magnetization below the Curie point. Intersection *P* determines *M*(*T*).

of T_c obtained in this way [the slope of Eq. (5) at $a^* = 0$] is $T_c = Ng^2\mu_B{}^2J(J + 1)\lambda/3k$, a result which is consistent with the earlier definition $T_c = C\lambda$. The results obtained from Fig. 1 are plotted as a function of temperature in **Fig. 2**. The value of J chosen is $^1/_2$.

The plot of Fig. 2 is in overall general agreement with the experimental results for iron, nickel, and cobalt. The low-temperature behavior, however, is better described by magnon theory. *See* MAGNON.

Heisenberg theory. The Heisenberg theory of ferromagnetism (W. Heisenberg, 1928) treats the origin of the Weiss molecular field on an atomic basis. It may be remarked at the outset that ordinary dipole-dipole interactions among atomic magnetic moments are much too small to account for the Weiss field. The foundation of the Heisenberg theory is the Pauli exclusion principle. The electrons providing the atomic magnetic moments are described by wave functions containing both space coordinates and spin angular momentum coordinates. The Pauli principle demands that the electron wave functions change sign (that is, be antisymmetric) when the space and spin coordinates of a pair of electrons are switched, or exchanged. As a consequence, the energy of the electrons depends on their relative spin orientation. Heisenberg realized that the essential

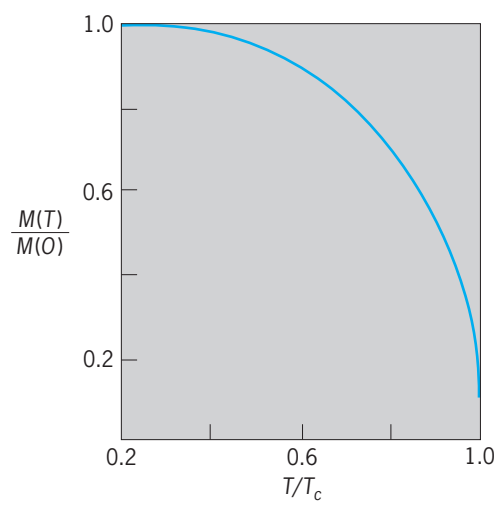

Fig. 2. Temperature dependence of spontaneous magnetization for $J = {}^1/_2$.

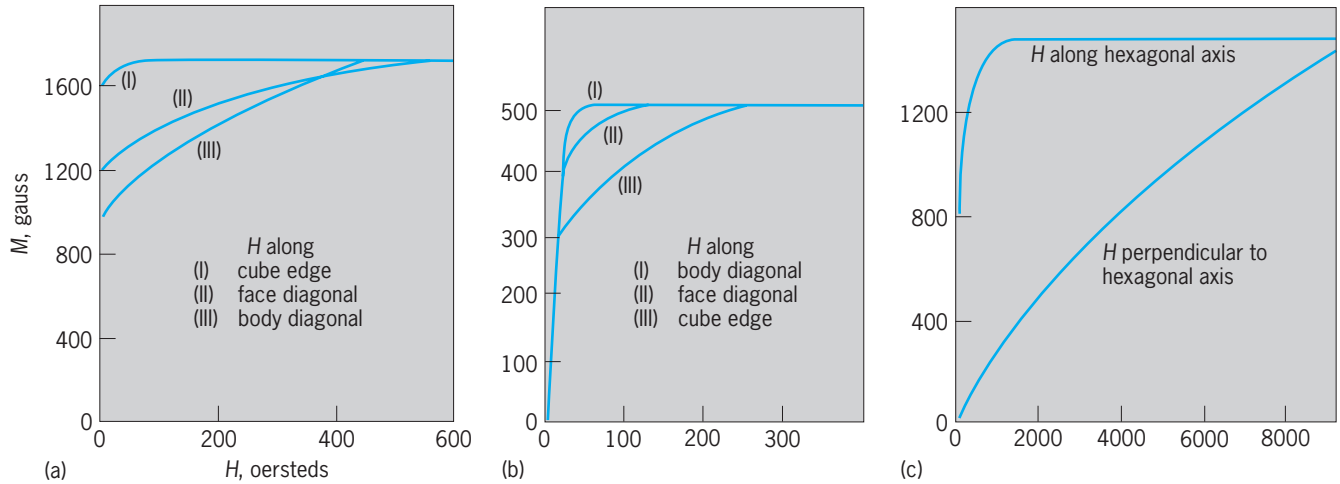

Fig. 3. Magnetization curves, for various directions of applied field *H*, indicated for single crystals (*a*) iron, (*b*) nickel, and (*c*) cobalt. 1 gauss = 10^{-4} tesla; 1 oersted = 79.6 A/m.

physics involves treating the electrostatic Coulomb repulsion between electrons within the constraint of the Pauli exclusion principle. He showed that the situation can be viewed in terms of an effective interaction between the electron spins, the exchange interaction energy, which had first been derived by P. A. M. Dirac. In Eq. (7), $\mathbf{S}_i$ and $\mathbf{S}_j$ are the spin angu-

$$\text{Exchange energy} = -2J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j \qquad (7)$$

lar momentum vectors of the two electrons i and j, and J_{ij} is the so-called exchange integral between the electrons. The exchange integral decreases rapidly with distance between the electrons and depends in a complicated way upon the spatial distribution (wave function) of the electrons. It is extremely difficult to compute. If the exchange integral is positive, the parallel arrangement is favored, and if J is large enough, ferromagnetism should result. If J is large and negative, antiferromagnetism or ferrimagnetism supposedly arises. The order of magnitude of J given by $J \simeq kT_c \simeq 10^{-13}$ erg. *See* EXCLUSION PRINCIPLE.

There seems to be no question that the Heisenberg theory correctly accounts for the tendency of electrons in the same ferromagnetic atom to exhibit parallel spins. However, whether or not it leads to the correct explanation of the interatomic alignment of spins is still a subject of controversy. In insulators it usually proves possible to express the coupling between atomic spins by Eq. (7), J_{ij} is interpreted as an effective exchange integral. In metals the problem is much more complicated. There is little doubt, however, that the basic Heisenberg idea is correct and the molecular field arises from the interplay between electrical forces and the effects of Pauli exclusion.

Magnetocrystalline anisotropy energy. This accounts for the experimental fact that ferromagnets tend to magnetize along certain crystallographic axes, called directions of easy magnetization. For example, a single crystal of iron, which is made up of a cubic array of iron atoms, tends to magnetize in the directions of the cube edges. At room temperature the magnetization of iron can be moved into a hard direction along a body diagonal only if an external magnetic field greater than about 30 kA/m (400 Oe) is applied in that direction. This can be seen in **Fig. 3**, which shows single-crystal magnetization curves for various directions of applied field H in iron, nickel, and cobalt.

N. S. Akulov showed (1929–1931) that the anisotropy energy U_A could be expressed conveniently in an ascending power series of the direction cosines between the magnetization and the crystal axes. For cubic crystals the lowest-order terms take the form of Eq. (8), where α_1, α_2, and α_3 are direction

$$U_A = K_1(\alpha_1{}^2\alpha_2{}^2 + \alpha_2{}^2\alpha_3{}^2 + \alpha_3{}^2\alpha_1{}^2) + K_2(\alpha_1{}^2\alpha_2{}^2\alpha_3{}^2) \qquad (8)$$

cosines with respect to the three cube edges, and K_1 and K_2 are temperature-dependent parameters characteristic of the material, called anisotropy constants. In general, $|K_1| > |K_2|$, and further terms are unnecessary. In iron, K_1 is positive, and therefore U_A is a minimum when any single direction cosine $\alpha_i = 1$ and the other $\alpha_j = 0$. That is, the cube edges are easy directions. In nickel, K_1 is negative and hence the body diagonals are easy directions (Fig. 3).

For crystals of other than cubic symmetry the energy U_A must be expressed in a form which is different from Eq. (8) and appropriate to the particular symmetry.

Anisotropy constants can be determined from (1) analysis of magnetization curves, (2) the torque on single crystals in a large applied field, and (3) single-crystal ferromagnetic resonance. *See* MAGNETIC RESONANCE.

The Heisenberg exchange energy, Eq. (7), is isotropic and cannot account for the observed anisotropy, which probably has its origin in a complicated interplay of spin-orbit coupling, crystalline electric fields, and overlap of orbital wave functions. The anisotropy energy also depends on the state of strain of the crystal, giving rise to magnetostriction,

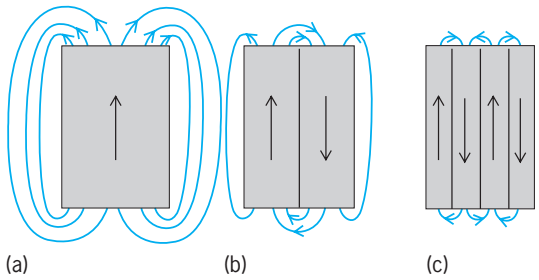

Fig. 4. Lowering of magnetic field energy by domains. (*a*) Lines of force for a single domain. (*b*) Shortening of lines of force by division into two domains. (*c*) Reduction of field energy by further subdivision.

that is, changes in length of a substance when it is magnetized. *See* MAGNETOSTRICTION.

Ferromagnetic domains. Small regions of spontaneous magnetization, formed at temperatures below the Curie point, are known as domains. As shown in **Fig. 4**, domains originate in order to lower the magnetic energy. In Fig. 4*b* it is shown that two domains will reduce the extent of the external magnetic field, since the magnetic lines of force are shortened. On further subdivision, as in Fig. 4*c*, this field is still further reduced.

Another way to describe the energy reduction is to note that the interior demagnetizing fields, coming from surface poles, are much smaller in the long, thin domains of Fig. 4*c* than in the "fat" domain of 4*a*.

The question arises as to how long this subdivision process continues. With each subdivision there is a decrease in field energy, but there is also an increase in Heisenberg exchange energy, since more and more magnetic moments are aligning antiparallel. Finally a state is reached in which further subdivision would cause a greater increase in exchange energy than decrease in field energy, and the ferromagnet will assume this state of minimum total energy.

Bloch wall. Also because of exchange energy, the reversal of magnetization between domains does not occur abruptly but takes place gradually through a transition zone called the Bloch wall (**Fig. 5**). To understand the reason for this wall, consider the exchange energy involved. Let the angle between the magnetization of neighbor planes be ϕ and assume that, as is usual, the exchange energy is appreciable only between neighbor atoms. According to Eq. (7), the exchange energy between atoms on neighbor planes varies as $-\cos\phi$, or as a constant $+\phi^2/2$ if ϕ is small. If total reversal occurs in N planes, $\phi = \pi/N$, and the total exchange energy of the wall is given by expression (9). This equation shows that the total ex-

$$U_E\ (\text{wall}) \propto \text{constant} + N\left(\frac{\pi}{N}\right)^2 \qquad (9)$$

change energy decreases as the number of transition planes increases; that is, the exchange energy tends to make the domain wall thicker.

The question now arises as to why N should not continue to increase until the entire crystal becomes a single Bloch wall. This does not happen because the decrease in exchange energy is accompanied by an increase in anisotropy energy. Many of the intermediary planes of Fig. 5*b* must of necessity have their magnetization along hard directions. The larger the value of N, the greater must be the number of such planes. The thickness of the wall can be determined by finding that value of N which makes the sum of exchange and anisotropy energies a minimum. In iron the wall is ~50 nanometers thick and has the total energy ~2 ergs/cm^2.

Domain arrangement. The orientation of domains in a crystal is determined primarily by the need to minimize the magnetic energy (Fig. 4). It is possible to eliminate all surface magnetic poles by forming flux-closure domains (**Fig. 6**). Here the normal component of magnetization is continuous across all domain boundaries. The demagnetizing fields are zero everywhere, except for the trivial effect of surface poles in the Bloch walls. In a uniaxial crystal, that is, a crystal with a single easy direction, an arrangement as shown in Fig. 6*b* will be preferred since it has a lower density of magnetization normal to the easy direction, or in the hard direction. Even in cubic crystals, in which all directions of magnetization in Fig. 6 may be easy, Fig. 6*b* will be preferred because of magnetostriction. In iron, for example, each domain increases in length along the direction of magnetization by a fraction $\sim 2 \times 10^{-5}$. Thus the domains of Fig. 6 can be fitted smoothly together only by straining them elastically against this magnetostriction, and the required elastic strain energy will be smaller in Fig. 6*b* than in Fig. 6*a*.

In polycrystals the domain structure is much more complicated, depending upon such variables as grain

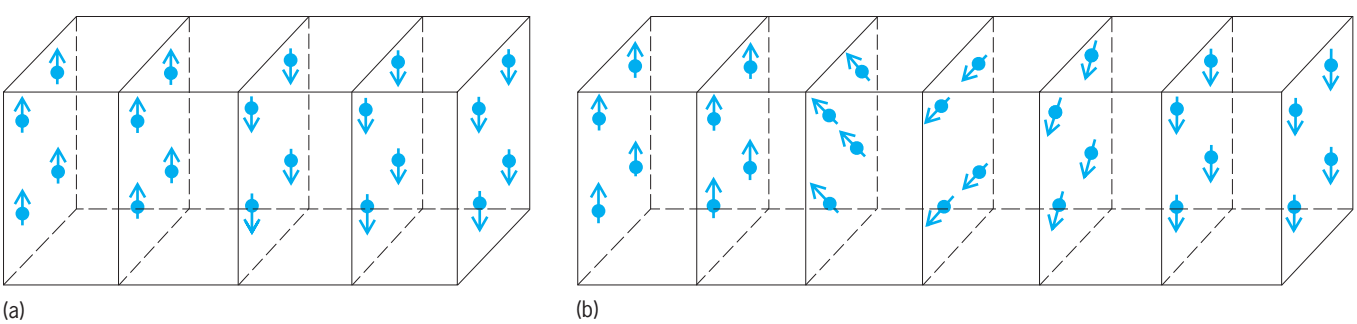

Fig. 5. Lowering of exchange energy by the transition zone known as Bloch wall. The reversal of magnetization between domains (*a*) does not take place abruptly but (*b*) by degrees.

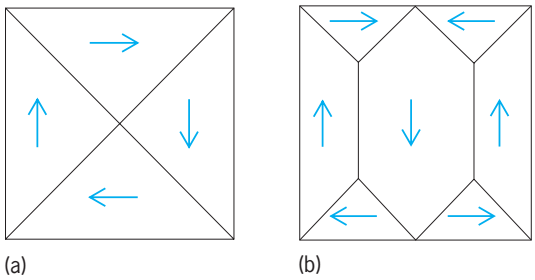

Fig. 6. Flux-closure domains. (*a*) Large domains at right angles. (*b*) Reduction in their size, causing reduction of anisotropy energy of uniaxial crystals or of strain energy of cubic crystals.

orientation and grain boundaries. It is possible, however, for domains to cross grain boundaries.

On minimizing the total contributions from (1) magnetic, or demagnetizing, energy, (2) anisotropy, (3) magnetostriction, (4) elastic strain, and (5) Bloch wall energy, it is found that, depending upon the composition and shape of the crystal, the theoretical domain thickness should vary from about 0.001 to 0.1 cm.

Bitter powder patterns. Direct experimental evidence of the existence of domains is furnished by Bitter powder patterns. First a surface is prepared by careful polishing with fine abrasive, followed (for metallic samples) by electropolish. Then a drop of a colloidal suspension of magnetic Fe_2O_3 particles is placed on the surface and covered by a microscope cover glass. These particles gather at surface regions where the magnetic field is largest and can be observed easily with a microscope. It is seen from Fig. 5 that the surface field is largest at the center of the Bloch wall; that is, lines of surface poles will demarcate the wall. Thus the colloidal particles concentrate along domain boundaries, as shown in **Fig. 7**. If the surface deviates by even a few degrees of angle from a simple lattice plane, the pattern becomes very complicated. The beginnings of this may be seen in the "fir tree" patterns at the lower right of Fig. 7.

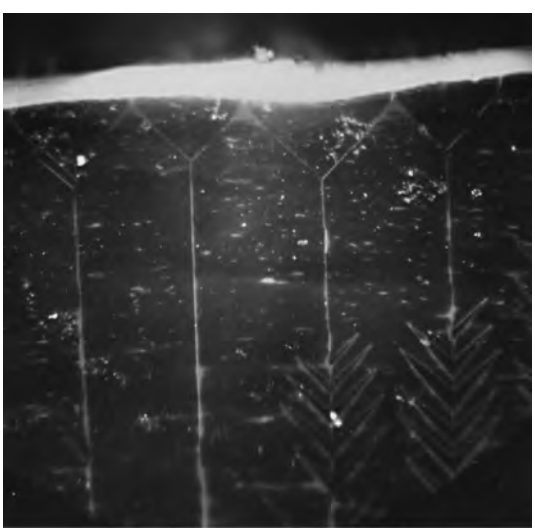

Fig. 7. Bitter powder patterns on a (100) surface of silicon-iron. (*Photograph by H. J. Williams*)

The powder patterns only delineate the boundaries and do not themselves show the direction of magnetization. This direction may, however, be inferred from a study of how particles gather around a scratch made on the surface, coupled with observation of domain growth when the sample is placed in a magnetic field.

The Kerr magnetooptic effect (rotation of the plane of polarization of light reflected from a magnetic surface) has also been used to study domains. J. F. Dillon demonstrated that single-crystal slabs of ferromagnetic yttrium iron garnet (YIG) are transparent to visible light and that the Faraday effect (rotation of the plane of polarization of light transmitted along a magnetic field) can be used to observe domains. The effect is very striking, and the growth and diminution of domains are readily visible. *See* FARADAY EFFECT; KERR EFFECT.

It is also possible to observe domains by means of transmission or reflection of electrons in an electron microscope. This technique, and also the optical techniques mentioned above, are particularly useful if the Bloch walls are extremely thick (because of small magnetocrystalline anisotropy) and the resultant Bitter powder patterns very blurred.

Bloch wall motion. This accounts for the initial portion of the magnetization curve. As the wall passes through regions of crystal strain, or over a number of foreign atoms, it may suddenly go over a potential energy hump and into a minimum. There will be a small, almost discontinuous jump in the magnetization, easily detected by a search coil (a device for measuring change of flux density) and amplifier. This is called the Barkhausen effect. It was once thought erroneously that this effect comes from a sudden, complete reversal of magnetization within a domain. However, H. J. Williams and W. Shockley showed that many Barkhausen jumps accompany the motion of a single domain wall. *See* BARKHAUSEN EFFECT; DOMAIN (ELECTRICITY AND MAGNETISM).

Soft magnetic materials. Materials easily magnetized and demagnetized are called soft; these are used in alternating-current machinery. The problem of making cheap soft materials is complicated by the fact that readily fabricated metals usually have many crystalline boundaries and crystal grains oriented in many directions. In such metals the magnetization process is accompanied by much irreversible Bloch wall motion and by much rotation against anisotropy, which is usually irreversible.

The ideal cheap soft material would be an iron alloy fabricated by some inexpensive technique which results in all crystal grains being oriented in the same or nearly the same direction. Various complicated rolling and annealing methods have been discovered in the continued search for better grain-oriented or "cube-textured" steels.

In some situations the cost of fabrication is secondary, and alloys with the smallest possible crystalline anisotropy and magnetostriction are demanded. The first such need appears to have been for the inductive loading of submarine telegraph cables. This gave rise to the perfection of permalloy, an

alloy of 21.5% iron and 78.5% nickel. The constant K_1 of Eq. (8) takes opposite signs in pure iron and pure nickel, and at the permalloy composition the resultant anisotropy goes to zero. More precisely, the composition for zero anisotropy is not quite the same as for zero magnetostriction, and permalloy represents a compromise. Addition of a third element, generally molybdenum, can simultaneously drive anisotropy and magnetostriction to zero and, as a bonus, decrease resistivity.

Hard magnetic materials. Materials which neither magnetize nor demagnetize easily are called hard; these are used in permanent magnets. The properties of a hard magnet material are described by its B versus H hysteresis curve, where B is the magnetic induction and H the applied magnetic field. Important material parameters specified by this curve include the remanence B_r (the induction for zero applied field), the coercivity or coercive force H_c (the reverse applied field required to drive B to zero), and the maximum energy product $(BH)_{max}$, the maximum product of B and H in the second quadrant of the hysteresis curve. The parameter $(BH)_{max}$ is the most widely used single figure of merit for a permanent magnet material; it provides a measure of the magnetic field that a unit volume of the material can produce.

Lodestone was the first magnetically hard material known. Principally composed of the ore magnetite (Fe_3O_4) it was long applied in navigational compasses. In the twentieth century, five classes of permanent-magnet materials have enjoyed technological importance. The magnet steels contain carbon, chromium, tungsten, or cobalt additives, serving to impede domain wall motion and thus to generate coercivity. These materials feature high remanence ($B_r \simeq 1$ tesla or 10 kilogauss) but low H_c and low $(BH)_{max}$ (less than 8 kT/m^3 or 1 megagauss-oersted). Alnicos are aluminum-nickel-iron alloys containing finely dispersed, oriented, elongated particles precipitated by thermal treatment in a field. The coercivity arises from the shape of the particles, and $(BH)_{max}$ can be as large as 80 kT/m^3 (10 MG · Oe). Hard ferrite magnets are based on the oxides $BaFe_{12}O_{19}$ and $SrFe_{12}O_{19}$. Their coercivity derives from the magnetocrystalline anisotropy, rather than the shape, of small, single-domain particles, and can exceed that of alnicos. Although $(BH)_{max}$ is no larger than 40 kT/m^3 (5 MG · Oe), hard ferrite magnets are relatively inexpensive and are used in a great variety of commercial applications.

Substantial increases in $(BH)_{max}$ have been afforded by rare earth–transition metal materials whose rare earth component provides huge magnetocrystalline anisotropy which can be translated into large coercivity in a practical magnet, while the magnetization arises chiefly from the transition metal component. Samarium-cobalt magnets based on the $SmCo_5$ or Sm_2Co_{17} intermetallic compounds can have coercivities of ~1.6 MA/m (20 kOe) and maximum energy products in the range of 160–240 kT/m^3 (20–30 MG · Oe). Magnets containing the more abundant elements neodymium and iron rather than samarium and cobalt have been developed from neodymium-iron-boron materials based on the ternary compound $Nd_2Fe_{14}B$. Such magnets have been prepared both by the powder metallurgy method used for ferrite and samarium-cobalt materials and by the rapid solidification technique of melt spinning in which molten alloy is directed onto a rapidly rotating substrate wheel. Neodymium-iron-boron magnets with $(BH)_{max}$ approaching the enormous value of 400 kT/m^3 (50 MG · Oe) have been produced. Elihu Abrahams; Frederic Keffer; J. F. Herbst

Bibliography. B. Barbara, D. Grignoux, and C. Vettier, *Lectures on Modern Magnetism*, 1988; R. M. Bozorth, *Ferromagnetism*, 1994; S. Chikazumi and S. H. Charap, *Physics of Magnetism*, 1964, reprint 1978; B. D. Cullity, *Introduction to Magnetic Materials*, 1972; R. A. McCurrie (ed.), *The Structure and Properties of Ferromagnetic Materials*, 1994; D. C. Mattis, *Theory of Magnetism*, vol. 1, 2d ed., 1988, vol. 2, 1985; E. P. Wohlfarth (ed.), *Ferromagnetic Materials*, 3 vols., 1980–1982.

Ferry

A ship specifically configured for carrying passengers between two points. It permits persons to make their way from one place to another across a body of water, and it may carry vehicles, including commercial vehicles. A ferry is distinct from a cruise ship or a cargo ship. For a cruise ship the voyage itself is the destination, whereas with a ferry the journey's end point is the destination. While a ferry may carry cargo, in contrast to a cargo ship, this cargo is contained in a commercial vehicle accompanied by a driver.

Technology. Ferries may range in appearance from ships that are outwardly indistinguishable from cruise ships to small craft indistinguishable from pleasure boats. Some ferries with overnight runs on the Baltic Sea carry well over a thousand passengers, housing them in luxurious cabins and offering casinos and theaters for entertainment. At the other end of the spectrum, many of the world's harbors are served by water taxis—small ferries which may carry as few as six persons for a 5–15-min run.

Ferries may range in speed from slow vessels of 10 knots (18.5 km/h) or less to ultrafast ferries with speeds exceeding 60 knots (110 km/h). They may carry passengers only, or they may carry a mixture of passengers and private automobiles, or passengers, cars, and trucks. Ferries may be characterized by their hull form and their propulsion plant.

Hull forms. Ship hull forms are characterized by the way they support the weight of the ship. This is depicted in the sustention triangle (**Fig. 1**). The vertices of this triangle depict the three modes employed to support a ship. The buoyant, or archimedean, support mode is the fundamental means of support of all ships and boats.

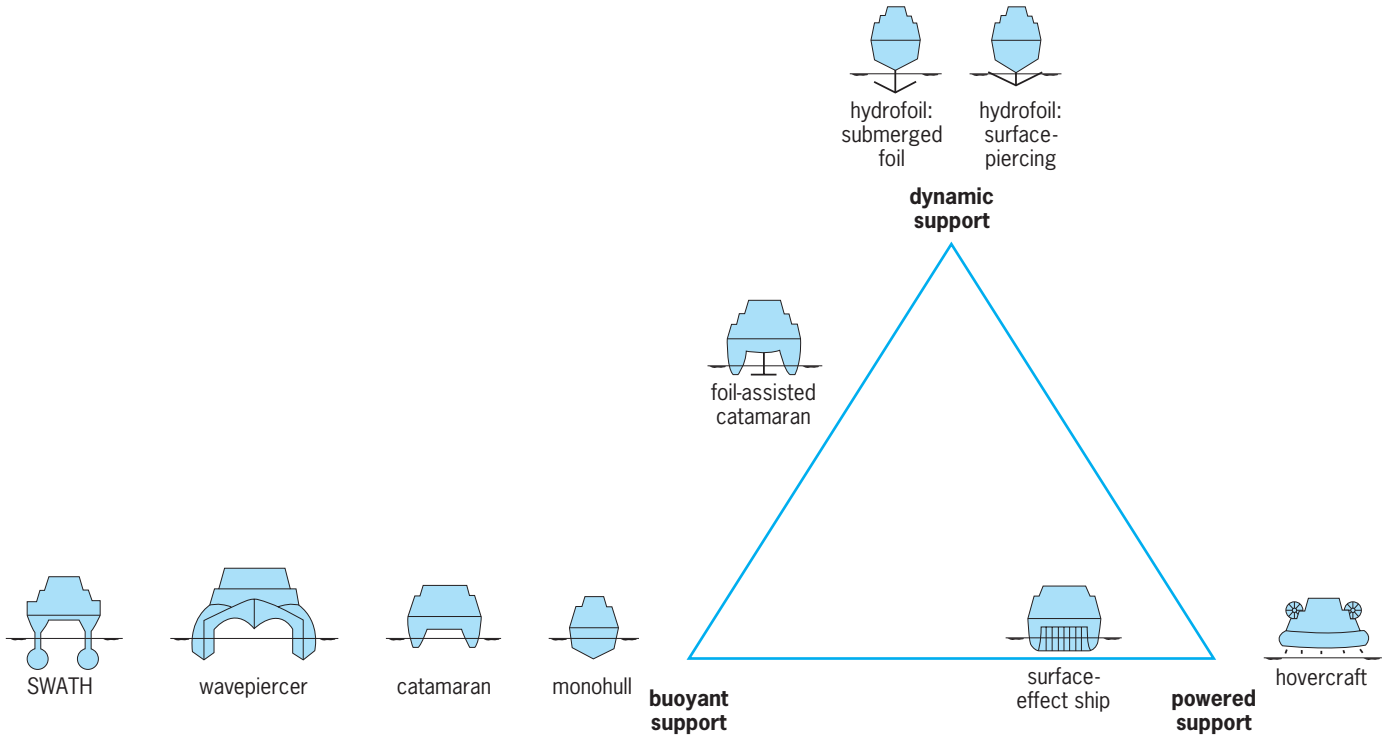

Fig. 1. Sustention triangle depicting the three principal modes of supporting the weight of a ferry and the types of ships that employ them.

Buoyant support. In buoyant support mode, a vessel's weight is borne by the upward force produced by the static displacement of the water in which it floats. The weight of the volume of water displaced is exactly equal to the weight of the ship. Any increase in the weight of the ship (for example, due to the loading of passengers or cargo) requires the ship to displace an increased volume of water, which it does by sinking to a slightly deeper draft. *See* ARCHIMEDES' PRINCIPLE; BUOYANCY; HYDROSTATICS.

The most common hull design for ferries is the displacement monohull ship form, which has one hull. Monohull ferries have been built in all sizes and speeds. Among monohulls a unique form is the double-ended ferry (**Fig. 2**). On short runs, it makes operational sense to build a ship which is identical fore and aft. For example, in use for river crossing, this ship simply shuttles from one dock to the other, putting one "bow" into the slip at one port, the other "bow" at the other port.

The catamaran hull is a very popular form, particularly for high-speed ferries (**Fig. 3**). A catamaran is a twin-hulled ship, having two hulls side by side. Because the two hulls are slender, a catamaran may have less wavemaking drag than a monohull of the same displacement. The two hulls are generally separated by a distance about equal to the beam of one hull. This results in a broad main deck area, which is more nearly square than the slender deck found on a monohull ship. Because of the broad expanse of the deck area, and because its low-aspect-ratio shape is easier to load with vehicles, the catamaran hull form has been very successfully incorporated into ferry service.

The drag (resistance to forward motion) of a displacement ship has three main components: air drag associated with moving the above-water portion of the ship, and two hydrodynamic drag components, skin friction and wavemaking drag. Skin friction is the drag of water sliding along the wetted surface of the hull.

Fig. 2. Double-ended ferry. (*Washington State Ferries*)

Fig. 3. High-speed catamaran ferry. (*Hitachi Zosen*)

When a ship moves through the water, it produces waves known as the ship's wake. Forming waves in the water surface requires a tremendous amount of energy. In order to move forward, a displacement ship must force aside a mass of water equal to the mass of the ship. Fast forward movement requires this substantial mass to be displaced rapidly, and the energy required grows. To a first order of magnitude, the power required to move a displacement-type hull will increase as the cube of the speed. *See* SHIP POWERING, MANEUVERING, AND SEAKEEPING.

The steep slope of the cubic speed-power relationship has driven engineers to seek ship hull forms with less drag. The quest for reduced drag has focused on finding means of support other than displacement, so that the ship no longer relies on thrusting aside a large mass of water.

Dynamic support. Shown in the sustention triangle (Fig. 1), dynamic support is produced by the movement of the ship and disappears when the vessel is at rest. The most common example of dynamic support is planing, the skimming of a craft across the surface of the water. In this case, the dynamic support is produced by the shape of the hull itself. Most small fast motor boats are planing craft.

A hydrofoil is a form of dynamic support craft which uses submerged winglike appendages to produce dynamic lift. These foils lift the main hull of the craft clear of the sea surface. The wavemaking portion of craft drag is thereby reduced; lifting the craft above the surface reduces the energy dissipated in wavemaking. The frictional component of drag, however, is undiminished, or may even be increased due to the increased surface area of the hydrofoils. Since the advantages of this type of craft come from the lifting of the ship above the water, it becomes important to reduce the basic weight of the craft. Lightweight materials and components are frequently employed on dynamically supported hulls. *See* HYDROFOIL CRAFT.

Powered support. The third vertex of the sustention triangle denotes powered support. In the marine world the only example is the air-cushion vehicle (ACV), including the related surface-effect ship (SES). With the air-cushion vehicle, an engine-driven fan produces aerostatic pressure beneath the bottom of the craft, lifting the rigid hull above the water surface. In this case there may be no contact between the ship's structure and the water, and thus there may be no frictional drag. The air pressure does, however, displace the water surface, and thus there is still wavemaking drag. Again, since the powered lift craft expends energy to raise the ship, it is in the designer's interest to reduce the weight of the ship.

A pure air-cushion vehicle has an air cushion underneath all parts of the ship—there is no contact between hard structure and the water. This produces some engineering challenges, the most obvious of which is that it is difficult to use underwater propulsors such as marine screw propellers. To overcome this, a hybrid type of vehicle was invented which is called a surface-effect ship in the United States. In the United Kingdom this concept is called a sidewall hovercraft. The surface-effect ship is a marriage of the catamaran and the air-cushion vehicle. A pair of slender catamaran-like hulls have an air cushion between them. Surface-effect ships are well suited to very high speeds (as much as 80 knots or 150 km/h). A few surface-effect ship ferries exist. *See* AIR-CUSHION VEHICLE.

Propulsion systems. Ferry propulsion systems depend upon the speed and hull form of the ship. The most common propulsion system is the conventional screw propeller. This is a submerged helical screw at the stern of the ship, driven by the ship's engines. Double-ended ferries require some form of reversible or double-ended propulsion system. Large double-ended ferries, such as those operated by the State of Washington and City of New York, have rudders and propellers at each end, sometimes having separate engines for each end. Other, much smaller double-ended ferries use a form of reversible propulsion, such as an azimuthing drive located amidships. This drive is steered 180° when the vessel reverses course. *See* PROPELLER (MARINE CRAFT).

High speed ferries—none of which are double-ended—face special propulsion issues. Whereas the diesel engine is the prime mover of choice for low-speed ferries, fast ferries demand much higher power. The higher weight sensitivity of a fast ferry frequently leads to use of gas turbine engines because of their light weight. *See* MARINE ENGINE.

Virtually all low-speed ferries and many high-speed ferries are propeller-driven. However, at speeds above about 40 knots (75 km/h) the hydrodynamic efficiency of the waterjet propulsor becomes superior to that of a propeller. As a consequence, almost all the fastest high-speed ferries use waterjet propulsion (**Fig. 4**).

Operation and economics. The condition which makes ferry service attractive is simple: a population divided by water. This might be, for example, a large city which is penetrated by a harbor or lake. In such a case, persons wishing to go to a point directly opposite them across the lake or harbor have the choice of taking a long land route or a considerably shorter water route. This same principle may

Fig. 4. Large waterjet propulsors used in high-speed catamaran ferry. (*Hitachi Zosen*)

apply on a much large scale to, for example, the Baltic or Mediterranean seas. It is, in fact, possible to drive from Italy to Africa, but the route is ridiculously long, and the maritime route is greatly preferred. Island communities are other obvious users of ferries.

It is important that ferries be able to load and unload rapidly. This is best accomplished by maximizing the number of lanes of traffic that can move in parallel, and eliminating obstacles in those lanes. Thus a fast-loading ferry may have all vehicle lanes continuous across the shore ramps, and may be capable of loading all vehicles straight-ahead, without need for any turning, backing, or other maneuvering. Similar requirements apply to pedestrian loading, wherein it is important to have wide lanes and eliminate bottlenecks created by narrow doorways, gate agents, and so forth. The rapidity of load and unload is a major factor in the economic success of a ferry, since a ferry which loads inefficiently may not offer any time savings compared to nonferry alternatives.

Laws and regulations. Most nations have federal maritime safety agencies regulating the operation of ferries within their national waters. In the United States, this role is performed by the Coast Guard. Ferries which run on international routes generally also fall under the authority of the International Maritime Organisation.

A significant secondary role is played by a group of private commercial entities known as classification societies. These societies were originally developed by insurers to assess the safety and insurability of a ship. In the centuries of their existence, they have published a large body of guidance on design, construction, operation, and maintenance of ferries. These guidelines have been adopted by some flag states and given the stature of de facto national law.

The largest international classification societies are the American Bureau of Shipping, Lloyds Register, and Det Norske Veritas. There are many others, and the selection of a classification society is a commercial decision made by the ship owner. *See* SHIP DESIGN.

Chris B. McKesson

Bibliography. E. V. Lewis (ed.), *Principles of Naval Architecture*, 2d ed., 3 vols., Society of Naval Architects and Marine Engineers, Jersey City, 1988; *Rules for Building and Classing Steel Vessels*, American Bureau of Shipping, Houston, 1998; R. Taggurt (ed.), *Ship Design and Construction*, 3d ed., Society of Naval Architects and Marine Engineers, Jersey City, 1980; Title 46, *US Code of Federal Regulations* (46 CFR).

Fertilization (animal)

The fusion of a female gamete (egg) with a male gamete (sperm) to form a single cell, the zygote. Gametes are haploid because their nuclei contain only one set of chromosomes. In contrast, most animal cells are diploid since their nuclei contain two sets of chromosomes, one set derived from the egg and the other from the sperm. The reduction from the diploid to the haploid number occurs in a specialized set of two cell divisions, known as meiosis, during which gametes are formed in the gonads of an adult. Fertilization restores the diploid chromosome number. In the human, normal eggs and sperm contain 23 chromosomes each, whereas the zygote and most of its daughter cells contain 23 pairs, or 46, chromosomes per cell nucleus. *See* GAMETOGENESIS; MEIOSIS; ZYGOTE.

Structure of eggs and sperm. Eggs and sperm contain the standard components of cells, including a nucleus and an outer membrane, the plasma membrane, which is a double layer of phospholipids with embedded proteins. However, each gamete also has specialized organelles for its particular function. The egg is the larger gamete. It sustains the embryo until the embryo can obtain nutrients from outside. This stage is reached quickly by the mammalian embryo when it implants in the uterus, so that the mammalian egg can be very small, typically about 0.1 mm in diameter. Other animal eggs are much bigger, as they contain storage materials that support embryonic development to more advanced stages. The mammalian egg is surrounded by the zona pellucida (or zona), a transparent envelope consisting of glycoproteins (**Fig. 1**). Other animal eggs, which are often released into the external environment, have additional protective layers. *See* CELL (BIOLOGY); OVUM.

The sperm's function is to find an egg of the same species and fuse with it. A mammalian sperm is sleek, with a head containing the highly condensed nucleus, a midpiece with mitochondria serving as the sperm's power plants, and a whiplike tail. Capping the nucleus is the acrosome, a bag of enzymes that will be released when the sperm penetrates the zona (**Fig. 2**). *See* SPERM CELL.

Attraction of sperm to eggs. In many animal species, sperm are attracted by chemicals released from the egg or—in mammals—from the ovarian follicle at ovulation. Evidently, the sperm can sense the attractant and swim to where the attractant is most concentrated. The sperm propels itself by the beat of its tail. In mammals, sperm need to move from

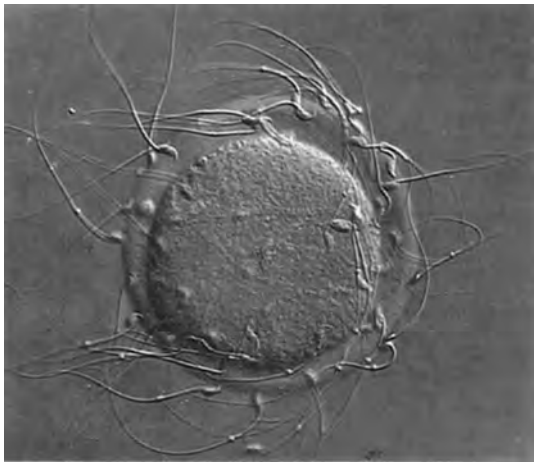

Fig. 1. Photomicrograph of mouse sperm bound to zona pellucida of an unfertilized egg under laboratory conditions (in vitro).

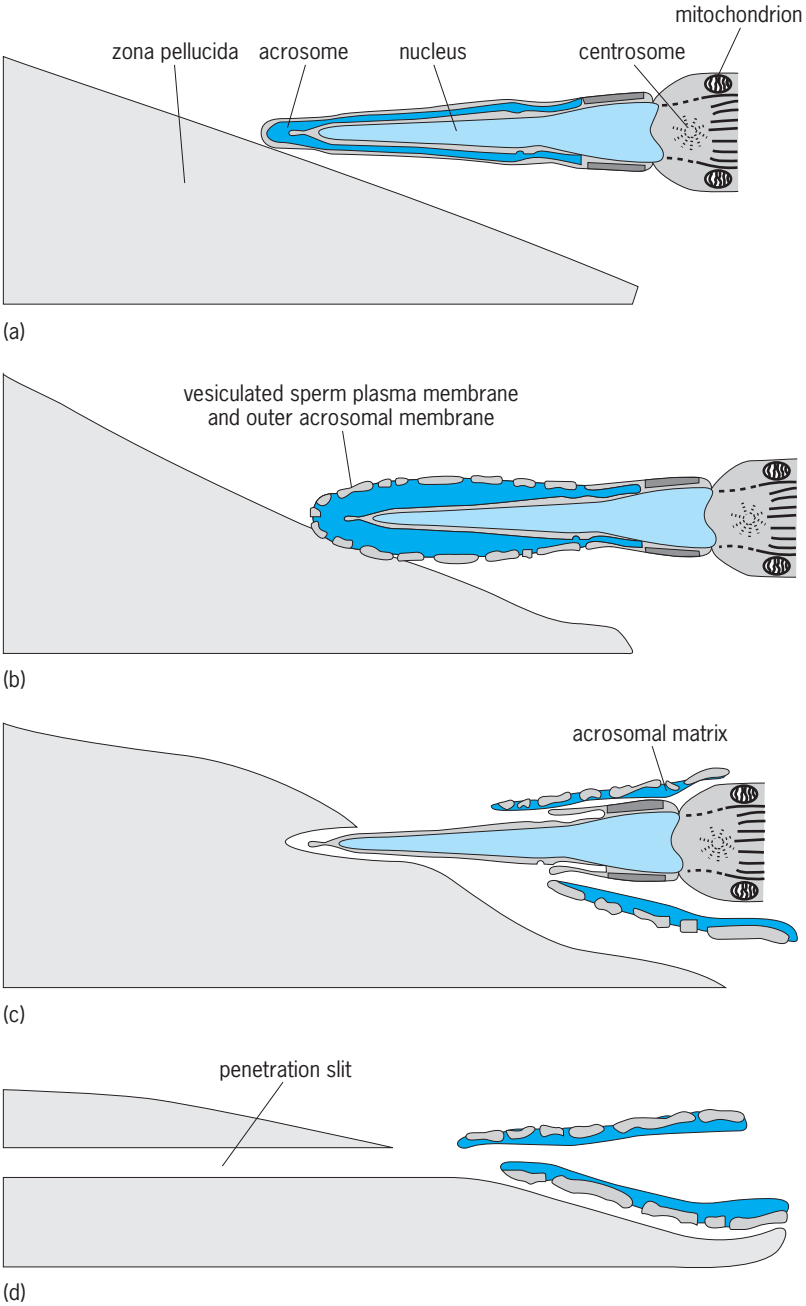

Fig. 2. Mammalian sperm (*a*) adhering to zona pellucida, (*b*, *c*) undergoing the acrosome reaction, and (*c*, *d*) penetrating the zona pellucida on its way to the egg cell proper. (*Modified from K. Kalthoff, Analysis of Biological Development, 2d ed., McGraw-Hill, 2001*)

the vagina to the site of fertilization in the upper third of the oviduct. Although this travel is facilitated by movements of the female genital tract, only a few per million ejaculated sperm reach their destination. In humans, both egg and sperm remain viable in the oviduct and capable of fertilization for about 3 days. *See* REPRODUCTION (ANIMAL); REPRODUCTIVE SYSTEM.

Sperm–egg adhesion. An important step in the fertilization process is the adhesion of the sperm to the zona (or equivalent egg envelope in other animals) [Fig. 1 and Fig. 2*a*]. This step guarantees much of the species specificity of fertilization because sperm adhere selectively to eggs of the same species. This selectivity is mediated by matching molecules on the head of the sperm and the zona. Sperm–egg adhesion is associated with an event in the sperm known as the acrosome reaction (Fig. 2*b* and *c*). The sperm plasma membrane and the outer acrosomal membrane fuse at many points, breaking up the two membranes into many vesicles, so that the sperm tip is now covered by the inner acrosomal membrane. Lytic enzymes released from the acrosome, along with the sperm's propulsive action, help it to penetrate the zona (Fig. 2*c* and *d*).

Gamete fusion. In the defining event of fertilization, the plasma membranes of egg and sperm fuse to generate one membrane covering the contents of both gametes. This membrane fusion depends on a sperm membrane protein, Izumo, named after a Japanese marriage shrine. The egg now actively engulfs the contents of the sperm and draws them inside. The zygote then contains two haploid nuclei, called pronuclei. The female pronucleus originates when the egg completes its meiotic divisions. The male pronucleus is formed by decondensing the sperm nucleus. The sperm also contributes the centrosome, a pair of cylindrical bodies that had been located at the base of its tail. The centrosome duplicates, and each daughter centrosome forms one pole of a mitotic spindle. The spindle incorporates the chromosomes from both pronuclei and organizes the division of the zygote in two embryonic cells, or blastomeres. Both blastomeres will be diploid, combining the genetic information encoded in the chromosomes of egg and sperm. *See* CLEAVAGE (DEVELOPMENTAL BIOLOGY).

Egg activation. Gamete fusion triggers a rapid series of events collectively called egg activation. Within seconds, the electric potential across the zygote's plasma membrane changes from negative on the cytoplasmic side (resting potential) to positive (fertilization potential). Most of this change occurs by admitting sodium ions from outside the egg to the inside. A minute later, calcium ions are discharged from stores in the egg cytoplasm in a wave spreading from the sperm entry point over the entire egg. The released calcium activates regulatory proteins, which in turn trigger an acceleration of cell metabolism and the resumption of the cell division cycle, which had been arrested at the end of meiosis. *See* BIOPOTENTIALS AND IONIC CURRENTS; CELL CYCLE; CELL DIVISION.

Blocks to polyspermy. It is critical to allow only one sperm to fertilize an egg. Entry of more than one sperm into the egg cytoplasm, an event called polyspermy, would be disastrous because multiple introduced centrosomes would set up competing mitotic spindles, which would generate blastomeres with incomplete sets of chromosomes. Two complementary mechanisms to prevent polyspermy are known as the fast block and the slow block. The fast block is triggered by the fertilization potential, which involves a change in sperm receptors on the egg cell membrane and the release of calcium from the egg. The fertilization reaction does

not allow any additional sperm to fuse with the egg. The fast block is instantaneous, but it lasts only for about a minute, after which the fertilization potential disappears. By that time, a slow but permanent block has been triggered by the calcium wave that passes over the egg. It triggers a dramatic change known as the cortical reaction because it occurs at the egg surface (cortex). Thousands of membrane-bound vesicles (cortical granules) fuse with the zygote's plasma membrane, thus shedding their contents into the space under the zona. These contents include enzymes that modify the zona glycoproteins so that already adhering sperm are released and newly arriving sperm do not adhere.

Klaus Kalthoff

Bibliography. K. Kalthoff, *Analysis of Biological Development*, 2d ed., McGraw-Hill, Boston, 2001; P. M. Wassarman (ed.), *Elements of Mammalian Fertilization*, vols. 1 and 2, CRC Press, 1991.

Fertilizer

Materials added to the soil, or applied directly to crop foliage, to supply elements needed for plant nutrition. These materials may be in the form of solids, semisolids, slurry suspensions, pure liquids, aqueous solutions, or gases.

The chemical elements nitrogren, phosphorus, and potassium are the macronutrients, or primary fertilizer elements, which are required in greatest quantity. Sulfur, calcium, and magnesium, called secondary elements, are also necessary to the health and growth of vegetation, but they are required in lesser amounts compared to the macronutrients. The other elements of agronomic importance, provided for plant ingestion in small (or trace) amounts, include boron, cobalt, copper, iron, manganese, molybdenum, and zinc. These minor elements are called micronutrients. All these fertilizer elements, along with other chemical elements, occur naturally in agricultural soils in varying concentrations and mineral compositions which may or may not be in forms readily accessible to root systems of plants. The addition of fertilizer to soils used for the production of commercial crops is necessary to correct natural deficiencies and to replace the components absorbed by the crops in their growth.

Plants also show toxic responses to certain elements (including some of the micronutrients) naturally present in some soils. Much laboratory and field testing has been done to develop an understanding of crop sensitivity and response to chemical elements in their various natural combined forms. Further research has led to methods of soil adjustment to enhance positive response and to suppress toxicity. Advances in sensitive analytical chemical techniques have done much to extend this knowledge. Plant species vary in sensitivity, and in some cases boundaries of the desirable operating concentration region between deficiency and toxicity are difficult to define. It is also possible that alternation of other soil components may shift availability of an in-place micronutrient to create a more favorable, or less favorable, condition without changing its concentration in the soil.

Commercial Fertilizers

The major effort of the chemical process industry, in regard to fertilizer, is directed to the production of compounds of the primary fertilizer elements. Crop requirements of fertilizer components could be satisfied by the spreading of individual materials for each element deficient in the soil. However, economy favors the single application of a balanced mixture that satisfies all nutritional needs of a crop. Since each primary element can be supplied in any number of different chemical compounds, the practice has given rise to a great diversity of formulations. Liquids, slurry suspensions, solids, and solid blends are marketed in a wide array of compositions to suit the needs of specific crops and the specific soils.

Formulas. The compositions of fertilizer mixtures, in terms of the primary fertilizer elements, are identified by an N-P-K code: N denotes elemental nitrogen; P denotes the anhydride of phosphoric acid (P_2O_5); K denotes the oxide of potassium (K_2O). All are expressed numerically in percentage composition, or units of 20 lb each per short ton (10 kg per metric ton) of finished fertilizer as packaged. Formula 8-32-16 thus contains a mixture aggregating 8 wt % N in some form of nitrogen compounds, 32 wt % P_2O_5 in some form of phosphates, and 16 wt % K_2O in some form of potassium compounds, to give a product with a total of 56 fertilizer units. The commercial N-P-K formulas are generally in whole numbers. The chemical nature of the primary element sources are identified elsewhere on the label by the manufacturer. None of the N-P-K formulas totals 100% plant nutrients because the formulas indicate only the nutrient portions of the primary-element compounds and do not account for any other materials present.

Other unidentified materials present in the fertilizer products include free water and water content of the hydrated primary components, uncredited cation and anion portions of salts (sodium, silica, chloride, sulfate, and so on), impurities remaining in the chemical ingredients as processed, and deliberately added inert materials (called ballast), to make whole-number formula compositions. Low-analysis formulas contain substantial proportions of ballast. Formulations prepared to correct specific soil deficiencies may include secondary or micronutrient fertilizer elements. The other than primary elements are identified in the fourth and successive positions of the formula. Formula 16-20-0-14S contains the equivalent of 16 wt % N, 20 wt % P_2O_5, 0 wt % K_2O, and 14 wt % S (elemental sulfur).

Form of fertilizers. The fertilizer elements rarely occur in nature in the elemental chemical form, and most would not be useful for plant nutrition if supplied in that form. Growing plants can assimilate only fertilizer elements in the combined state of inorganic compounds that are amenable to osmotic absorption. Many modern fertilizer materials consist of compounds that are immediately usable by the

crops to which they are applied. Others are quickly converted within the soil to forms that can be assimilated. Some fertilizer chemicals are specifically designed to dissolve slowly or to delay reaction within the soil and therefore prolong the release of easily absorbed compounds to provide sustained feeding over the growth cycle of the plants. Fertilizers are derived from many different types of raw materials. The earlier forms of fertilizers deliberately applied to soils were readily available waste products from the activities of other enterprises. Primary, secondary, and trace elements are now compounded into finished fertilizer materials as major products of the chemical process industry. Large amounts of fertilizer chemicals are also generated as by-products in the production of other chemicals.

Production, marketing, and use. The fertilizer industry of the United States is organized into separate segments for producing and marketing nitrogen, phosphate, or potash intermediates and products. The relatively few plants for production of large quantities of these basic materials are usually located near the mining operations. In the case of nitrogen fertilizer products, the large production complexes are located near natural gas sources, which may be wells or major natural gas pipelines. There are overlaps in the business, and some producers make a full line of products, including basic grades of bulk granular N-P and N-P-K materials. High-analysis products are made to minimize freight cost per unit of plant nutrient. Bulk products are transported to market by ship, barge, rail, and truck. Substantial quantities of anhydrous liquid ammonia also move by pipeline from plants in natural gas–producing areas to major fertilizer markets in the north-central and midwestern states. One innovation has been the construction of large refrigerated atmospheric-pressure storage tanks, barges, and oceangoing ships for transport of cold liquid ammonia. Nitrogen solutions are also produced in bulk and shipped by pipeline as well as by conventional carriers. The fertilizer needs of farmers are serviced by blending plants located near their markets. Services by some distributors include soil analysis and rental of application equipment.

UAN solutions. Aqueous solutions of urea, ammonia, and ammonium nitrate (UAN solutions) are used directly by the farmers as well as in the preparation of granular N-P-K products by mixing with other materials, such as normal superphosphate and triple superphosphate. UAN solutions are also spread directly by field application or used to prepare complete N-P-K fertilizer solutions or suspensions. Solution injection into irrigation water is gaining in popularity. Solutions completely free of solids and materials that tend to precipitate in hard water are important to prevent equipment plugging in trickle irrigation systems. Clear solutions are also important in pivotal irrigation systems for circular, quarter-section plots. Low volatility of the fertilizer components of the solutions assures minimum evaporation and most economic use.

Suspension fertilizers. These fertilizers consist of aqueous slurries of fine crystals in saturated solutions that are stabilized by small amounts of gelling materials, such as attapulgite clay. Many grades of high-analysis N-P-K suspensions, including secondary nutrients and micronutrients, are produced by blending plants. Suspension fertilizers can be produced in a simple agitated tank with very low expenditure of energy. The suspensions can be maintained in uniform composition during spreading on the fields, and give better dispersion than granular material.

Nitrogen. Only a few species of bacteria and algae are alone able to directly utilize nitrogen from its stable molecular form in the atmosphere. Many members of the Leguminosae family of plants maintain a symbiotic relationship with specific members of the bacterial genus *Rhizobium*; this allows fixation of atmospheric nitrogen in nodules on the roots of the plants. The rhizobia are capable of prolonged independent existence in the soil, but fixation takes place only when symbiosis is established. Under favorable conditions this fixation can annually provide up to 100 lb of nitrogen per acre (112 kg/ha). Though virgin soils accumulate substantial amounts of nitrogen from the symbiosis process and return of decaying vegetation over long periods of time, commercial cropping rapidly depletes available nitrogen in the soil.

It is probable that on average no more than 50% of the applied chemical fertilizer nitrogen contributes to plant growth. Several factors are responsible for this inefficiency. Nitrates are the most readily absorbed and usable form of nitrogen, but the nitrates are all very water-soluble compounds and the nitrate ions are not retained by reaction in the soils. The portions of nitrate not immediately absorbed by plant roots can also be lost by bacterial conversion to diatomic nitrogen gas (denitrification) under anaerobic conditions, with loss accelerated in warm environment and pH greater than 7. Liquid ammonia immediately vaporizes in the soil and may be lost to the atmosphere in significant proportion by expansion of the gas if it is not applied deep enough to be captured by soil colloids and water. All alkaline soils contain mineral carbonates and bicarbonates that react to form unstable ammonium carbonates. Ammonium ions held on colloids are converted to nitrate by bacterial nitrification and then become subject to the same losses as nitrate applied directly. The presence of urease in practically all soils assures rapid hydrolysis of uncoated urea, and similar organic compounds, with the development of local alkalinity, favoring ammonium carbonate formation. Excessive water from rain or irrigation after fertilization can physically remove soluble nitrogen compounds in the runoff. Significant deficiency in any of the growth requirements, such as sunlight, soil moisture, and other vital fertilizer elements, will reduce the effectiveness of applied nitrogen fertilizer and thus result in loss of nitrogen benefit. Proper attention to the balance of all soil requirements gives the best assurance of optimum response from any fertilizer component. *See* NITROGEN FIXATION.

Ammonia. Anhydrous ammonia is the richest source of nitrogen (82 wt % N) of any of the fertilizer

materials available. Ammonia is the basic component of aqueous solution fertilizers because of the increased solubility that the presence of ammonia imparts to other important solution chemicals, such as ammonium nitrate and urea. Ammonia is the most water-soluble of the common gases. A saturated aqueous solution at room temperature contains about 30 wt % ammonia in the form of ammonium hydroxide, NH_4OH, which has a strongly alkaline nature and is similar in chemical behavior to the hydroxides of alkali metals.

Practically all modern nitrogen fertilizer is derived from ammonia synthesized from hydrogen and nitrogen. The high-pressure catalytic process for ammonia synthesis is known as the Haber-Bosch process. The process employs an iron catalyst promoted with small amounts of alkali, magnesium, aluminum, and calcium. The effect of the ammonia synthesis process upon the chemical and fertilizer industries has been profound.

The preparation and purification of synthesis gas for ammonia production is of comparable importance to the synthesis reaction system itself. A major requirement is the economic production of large quantities of hydrogen free of impurities that would cause undesirable side reactions or would be damaging to catalysts. Electrolysis of water gives a high-purity hydrogen product, but it is expensive and delivered at near-atmospheric pressure. The only practical raw materials available for quantity production of hydrogen for the initial operation of ammonia synthesis were coal and coke. Coke was selected because its reaction with steam produced the less complicated raw hydrogen gas stream. Hydrogen production for the process in the United States quickly shifted to sweet natural gas, which was abundant and consisted mainly of methane with few troublesome contaminants. Methane, CH_4, is the richest hydrocarbon feedstock and the simplest to convert to hydrogen by catalytic steam reforming [reaction (1)]. Sulfur is a poison to most of

$$CH_4 + H_2O \rightarrow CO + 3H_2 \qquad (1)$$

the catalysts in the system and must be removed. In the case of natural gas or other hydrocarbons, the sulfur is removed by pretreatment of the feedstock. The preparation of ammonia synthesis gas requires the addition of nitrogen to the net hydrogen stream.

An alternative scheme, used to a lesser extent, is the injection of pure nitrogen obtained from cryogenic separation of air apart from the ammonia synthesis system. The residual carbon monoxide in the crude synthesis gas is then converted to carbon dioxide to the maximum extent possible with two stages of water-gas shift conversion. Carbon dioxide is selectively absorbed from the cooled gas by any one of a number of commercially available acid-gas absorption solutions. The rich solution is stripped of its CO_2 content and recycled to the absorber. Residual traces of carbon oxides in the gas from the absorber are then catalytically converted to methane before compression and delivery of the synthesis gas to the ammonia synthesis system.

The need for fertilizer production in densely populated areas, such as the Netherlands and Japan and in developing countries with rapidly growing populations, but without natural gas supplies, stimulated research for reforming catalysts that could accommodate heavier hydrocarbons. Naphtha, a liquid petroleum product that could be easily transported to the plant sites, was plentiful because of lesser demand for gasoline in those parts of the world and thus more economically available than in the United States. Several ammonia plants were built in the 1960s to use naphtha as feedstock for hydrogen production.

Increasing dependence of the industrial nations upon the oil-producing nations was forcefully brought to attention during the oil embargo of 1973. The sharply increasing price of petroleum and natural gas after that time dictated an intensive search for ways to reduce all energy use, including requirements for the production of ammonia and other fertilizer materials. Special attention has been given to heat recovery and reactor design. Renewed interest in catalyst development brought forth new manufacturing techniques promising significant improvements in properties that allow reduction in ammonia plant operating cost. Several techniques were devised to recover and recycle hydrogen and ammonia from the purge gas that was previously lost or burned as fuel. Pressure swing absorption, permeable-membrane diffusion, and cryogenic separation are examples of unit operations that are being used to advantage to recover products and reactants from purge streams. *See* AMMONIA.

Nitric acid. Nitric acid, HNO_3, is the second-most important industrial mineral acid. Nitric acid is not used by itself as a fertilizer, though the pure liquid contains 22 wt % N. The greatest portion of commercial nitric acid in the United States is used to produce fertilizer-grade ammonium nitrate, which has the N-P-K formula 35-0-0.

Many large-scale attempts have been made to achieve direct oxidation of nitrogen in air. The fast high-temperature oxidation of nitrogen, whether it be the diatomic gas in air or in the ammonia molecule, results in almost exclusive formation of nitrogen oxide, NO. Further reactions to higher oxidation states are favored by much lower temperature. The air oxidation of NO to NO_2 is carried out simultaneously with the disproportionation reactions (2).

$$\begin{aligned} 2NO + O_2 &\rightarrow 2NO_2 \\ 3NO_2 + H_2O &\rightarrow 2HNO_3 + NO \end{aligned} \qquad (2)$$

This complex pair of reactions is highly exothermic and requires a multitray absorption column, with each tray cooled separately for countercurrent absorption of the nitric acid, HNO_3, in water. Nitric acid forms an azeotrope with water at about 68 wt % HNO_3. Several dehydrating techniques have been developed to concentrate nitric acid above 68%. Most fertilizer processes do not require higher than 60% concentration, and only a small portion of the

product is concentrated for other markets. *See* NITRIC ACID.

Ammonium nitrate. Ammonium nitrate, AN, is a white, water-soluble, inorganic salt that contains 35 wt % N. As a fertilizer, ammonium nitrate is dissolved readily by moisture in the soil and diffuses outward from its placement. The nitrate portion of this salt is immediately available to plants through the root system, and the ammonium portion is retained by the soil colloids, where it is oxidized gradually to the nitrate form with the aid of bacterial organisms in the soil.

Ammonium nitrate is made by the simple acid-base reaction of nitric acid and ammonia in aqueous solution, with the heat release of 26,000 calories (109 kilojoules), per gram mole. The heat of reaction is usually used to vaporize the ammonia fed to the process and to evaporate water to concentrate the salt solution produced. After the conclusion of World War II, intensive research brought forth several process variations for the production of ammonium nitrate. The Stengel process produces a low-moisture-content melt in a packed reactor, and the melt which solidifies in a thin ribbon on a water-cooled stainless steel belt was flaked and screened to uniform particle size. Crystalline product is made to a controlled crystal size in a recirculation neutralizer-crystallizer when virgin acid and anhydrous ammonia are used. The crystals are centrifuged from the mother liquor and dried. The neutralization reaction has been carried out under various pressure levels from vacuum to as much as 5 atm (32 kilopascals). Graining is the process of kneading a congealing melt to form round granules which are then cooled, screened, coated, and packaged. Dilute ammonia gas streams are scrubbed in direct contact with acid in towers, and the resultant salt solution is concentrated in an evaporator to the desired melt composition. Solutions concentrated to greater than 99.5 wt % form a melt from which dense prills can be made. Prills produced for fertilizer are made as dense as possible to provide strength in handling and resistance to moisture intrusion. Prilling is the process of forming uniform droplets of a molten material at the top of a tower and cooling the droplets to solid spherical beads as they fall through a rising column of air. A common means of forming the droplets is to pass the molten material through a shower-head type of distributor at the top of the tower. This process makes a very uniform product of size set by the design of the system. Ammonium nitrate can also be granulated in any of several types of granulation devices to particle size as desired to match other materials for bulk blending. Another desirable feature of granules is the greater crushing strength compared to that of prills.

Ammonium nitrate is also a component part of other nitrogen fertilizers, such as ammonium sulfate-nitrate. ASN (26% N), calcium-ammonium nitrate, CAN (28% N), and ammoniated nitrophosphate. A troublesome characteristic of ammonium nitrate is the several crystal phases it exhibits with change in temperature at 257, 183, 90, and −0.4°F (125, 84, 32, and −18°C). These different crystal structures have significantly different densities, and the transition from one form to another tends to fracture the solids. Temperature cycling (particularly above and below 90°F or 32°C) can reduce well-shaped flakes, prills, or granules to powder within storage piles or containers. Much research has been done to alleviate this problem.

Between 1950 and 1980 ammonium nitrate was the most important nitrogen fertilizer from the standpoint of world consumption. However, in spite of its continued application in large quantity, the percentage use has been steadily declining relative to the use of urea as a source of nitrogen. A large amount of ammonium nitrate is used to make nitrogen solutions for direct application, for the preparation of slurries, and for increasing the nitrogen content of solid fertilizers in granulation operations. The presence of ammonia increases the solubilities of ammonium nitrate and urea in aqueous solution, and a number of high-analysis nitrogen solutions are made by mixing ammonium nitrate, ammonia, and urea. The solutions are simple to prepare and easy to apply to farmlands. *See* PRILLING.

Sodium nitrate. Sodium nitrate is a fertilizer that provides 16% N, and has been particularly favored for cotton and beets. While natural deposits (caliche) of sodium and potassium nitrates have been reported in widely scattered parts of the world, the only deposits of commercial significance are in the northern desert of Chile. Caliche is a complex mixture of natural double salts of sodium, potassium, calcium, and magnesium. Nitrates and sulfates predominate, but recoverable portions of chlorides, iodides, and borates are also present.

An early high-temperature recovery method, known as the Shanks process, was introduced in Chile late in the nineteenth century. Another process came into operation in 1922. This was the Guggenheim process, which was designed to treat low-grade ores with efficient nitrate extraction at moderate temperature and crystallization with mechanical refrigeration. This process soon demonstrated the ability to economically treat ore with nitrate concentration as low as 7%. Large-scale operations were initiated in the late 1920s and early 1930s which greatly extended the life of Chile's deposits. The recovery of iodine and other mineral values in the caliche also enhanced the economics and helped to extend the operational life.

Following World War I, considerable effort was exerted to devise methods for producing synthetic sodium nitrate to guard against industry crisis by interruption of supply of the natural product from Chile. The five different process schemes that were commercialized used either nitric acid directly or oxides of nitrogen produced by the catalytic oxidation of ammonia. Sodium was supplied in the form of soda ash, common salt, or caustic soda. In one of the processes the reaction of nitric acid with salt (sodium chloride) also produced chlorine without generating by-product sodium hydroxide, as in the case with electrolytic production of chlorine. All the synthetic processes yielded aqueous solutions from which the sodium nitrate was recovered by

crystallization. Production peaked in the 1930s and continued brisk through the 1940s. The fertilizer use of sodium nitrate was then gradually replaced by ammonium nitrate and other nitrogen materials. Some sodium nitrate is being used, but it is unlikely that it will again enjoy its previous favored position in the fertilizer market.

Other nitrogen salts. Ammonium sulfate, $(NH_4)_2SO_4$, is a water-soluble, crystalline material containing 21% N and 24% S. Ammonium sulfate was first made in substantial amounts by the absorption of the ammonia content of coke-oven gas in scrubbers with aqueous solution of sulfuric acid and subsequent crystallization of the salt from the concentrated scrubber liquor. It is also made by reacting calcium sulfate with ammonium carbonate, filtering out the precipitated calcium carbonate, and crystallizing ammonium sulfate from the mother liquor. Large quantities of ammonium sulfate have been produced by the reaction of fresh ammonia with spent sulfuric acid from chemical operations and oil refinery catalytic alkylation processes. A substantial amount of ammonium sulfate has been made by the reaction of virgin sulfuric acid and ammonia. Ammonium sulfate produced as by-product in the manufacture of caprolactam and other large-tonnage chemicals has supplemented the world supply of fertilizer nitrogen. The intensive interest during the 1970s in environmental protection from industrial sulfur dioxide emissions led to development of several processes for recovery of the sulfur values, which include production of ammonium sulfate. Most of this pollution control by-product also finds its outlet in the fertilizer market.

Ammonium chloride or sal ammoniac, NH_4Cl, is a water-soluble, crystalline salt with 25 wt% N. Ammonium chloride is more corrosive than other ammonium fertilizers, and it also tends to increase the residual chloride content of soils. In spite of these characteristics, it has been a popular fertilizer material for rice and some field crops. A rice fungus that reduces sulfates to toxic sulfides under anaerobic conditions does not affect chlorides, and thus ammonium chloride is preferred over ammonium sulfate in such applications.

The salt can be easily made by reacting ammonia with hydrochloric acid and crystallizing the product. However, the major manufacturing method for ammonium chloride production is the double-salt process, which is a modification of the Solvay ammonia-soda process for making sodium carbonate. A sodium chloride solution is ammoniated and blown with carbon dioxide in a carbonating tower until the ammonia is carbonated to the bicarbonate, whereupon it reacts as formed with sodium chloride to give sodium bicarbonate and ammonium chloride. The Solvay process reaches equilibrium at about 75% conversion in the salt exchange, sodium bicarbonate is recovered by filtration, and the ammonia is recovered by the addition of calcium hydroxide, leaving a solution of calcium chloride. The ammonium chloride in the double-salt process regenerated after separation of the sodium bicarbonate is salted out by ammoniating the mother liquor, cooling, and adding fresh sodium chloride. The ammonium chloride is removed by centrifuge, and the cycle is repeated. The sodium bicarbonate is then calcined to sodium carbonate, and the carbon dioxide from calcination is used to carbonate the ammoniated salt solution in the tower.

Calcium cyanamide. Calcium cyanamide, $CaCN_2$, is an alkaline fertilizer material with fixed nitrogen. It is black and dark gray because of the residual unreacted carbon remaining from the reduction step in its production. The fertilizer nutrient content is 35% N and 70% CaO. Decomposition after application to the soil converts calcium cyanamide to calcium carbonate and ammonia according to reaction (3). It has

$$CaCN_2 + 3H_2O \rightarrow CaCO_3 + 2NH_3 \qquad (3)$$

additional agricultural value as a herbicide and as a defoliant for mature cotton plants to aid machine harvesting of the fiber. *See* DEFOLIANT AND DESICCANT; HERBICIDE.

Three manufacturing steps are involved in the production of calcium cyanamide. The basic raw materials are limestone and coke. Limestone, $CaCO_3$, is first calcined to quicklime, CaO. Calcium carbide, CaC_2, is then produced in an electric arc furnace (3600–4000°F or 2000–2200°C) by the reduction of quicklime with coke. This reaction is highly endothermic. The calcium cyanamide is then made by passing nitrogen gas, N_2, directly over heated calcium carbide in a furnace. This high-temperature reaction (1650–1830°F or 900–1000°C) is sufficiently exothermic to be self-sustaining, once it is iniated. Oxygen-free nitrogen is supplied from an air separation plant. The major use of calcium cyanamide since its discovery in 1898 has been agricultural. Its cost as a fertilizer is high in today's market relative to other fixed-nitrogen materials. Further agricultural use of calcium cyanamide is therefore likely to be limited to special services that can justify its high cost.

Urea. Fertilizer-grade urea is a water-soluble, synthetic, organic solid that contains 46% N. It has the highest nitrogen nutrient value of any of the solid fertilizers. Urea is less deliquescent than ammonium nitrate, but the critical relative humidity is 72% at 86°F (30°C), and so urea must be stored in a controlled low-relative-humidity environment to remain free-flowing. The many good characteristics of urea as a nitrogen fertilizer are responsible for its rapid worldwide acceptance. Urea is a stable material with a convenient melting point (270.9°F or 132.7°C) that allows it to be easily prilled or granulated for bulk blending with other solid fertilizer materials. Urea has high solubility in water and can be used to prepare high-analysis solutions. It forms an adduct with phosphates which allows the production of homogeneous granular solids. Low-cost by-product carbon dioxide from ammonia production related to making urea helps keep the cost competitive with other nitrogen materials. These two production units are commonly associated in a single fertilizer complex facility.

The process for urea synthesis in commercial use today is the pressure synthesis from ammonium carbamate. Ammonia and carbon dioxide are reacted under pressure of 170–400 atm (1.2–2.8 megapascals), with melt temperature maintained in the range of 347–410°F or 175–210°C. The selection of operating conditions is a compromise among several interactive factors. Yield is favored by high temperature, high pressure, and minimum excess ammonia. Unfortunately, these same conditions are also the worst conditions from the standpoint of equipment corrosion, and much attention has been given to this problem. The partial dehydration of ammonium carbamate formed in the reactor produces urea and water. The mixture let down in pressure from the reactor is stripped of ammonia and carbon dioxide, and the aqueous solution of urea is concentrated for direct granulation or prilling. The simplest and least expensive design for a plant is once-through operation. Then the off-gas containing excess ammonia from the reactor and the carbamate decomposer are used to make other products in associated process units. Plants with various degrees of gas and liquid recycle have also been constructed. Most plants today are designed for total recycle with maximum possible heat recovery. Some licensors offer plant designs having several levels of pressure letdown with vapor absorption and solution recycle arranged to minimize process energy consumption.

The hot urea melt is time-temperature-sensitive with respect to decomposition by condensation of 2 moles of urea to biuret with the loss of 1 mole of ammonia. Many crops show a toxic response to biuret, and fertilizer-grade urea has strict limits on the allowable biuret content. Most urea is concentrated by evaporation of water with the minimum possible melt residence time before entering the prill tower distributors to keep the biuret content below 1 wt %. Especially low-biuret urea can be made by the more expensive concentration with low-temperature crystallization from the aqueous solution, and prilling or granulating the melted crystals. Granulation avoids the atmospheric emission problem of prilling and has the advantage of adjusting particle size to match other solids for bulk blending. The matching of particle size is important to prevent resegregation of the different materials in the blend after packaging or in the bulk blend transported to the field. The granulated product also has much higher crushing strength than the prills. Sulfur-coated urea is made by spray-coating granular urea with molten sulfur in a rotating drum and seal-coating the product with a wax or low-molecular-weight polyethylene. The smaller-diameter urea prills are unsuitable for coating because the high surface-to-volume ratio would require too much sulfur and result in product too low in nitrogen content. Treating urea in this manner has several advantages. In addition to providing controlled release of nitrogen and sulfur, the coated material has excellent storage and handling characteristics that allow its use in bulk blending in humid climates. It may also be blended with single or triple superphosphates, which are incompatible with uncoated urea. *See* UREA.

Ureaform. Ureaform contains at least 35% N in a largely insoluble form that slowly and continuously becomes available in the soil. It is one example of the several slow-release fertilizer materials on the market. In the late 1930s an aqueous ammoniating solution of urea and formaldehyde was used to ammoniate superphosphate in the manufacture of mixed fertilizer. This operation produced a water-insoluble reaction product of urea and formaldehyde called ureaform. The advantage of ureaform as a fertilizer is its relative insolubility in cold water. This slows the nitrification process in the soil and thus provides an extended release of usable nitrogen to the plants. The condensation of urea and formaldehyde with the loss of water is catalyzed by acid in aqueous solution. Technology developed rapidly, and it is now made commercially by either the dilute solution process or the concentrated solution process. In either process the granulated and dried product is an odorless, white solid containing about 38% N.

Ureaform is not a single compound, but a mixture of methylene urea polymers that are capable of continued polymerization at reduced pH and elevated temperature. Ureaform is not hygroscopic. The conditions of manufacture strongly influence the quality of the product as determined by the activity index, AI, which is the percentage ratio of polymer soluble in hot water (212°F or 100°C) to the polymer insoluble in cold water (77°F or 25°C). As the degree of polymerization increases beyond the desired level, the hot-water solubility of the ureaform decreases. The best-quality ureaform has a high percentage of cold-water-insoluble material and a high AI value. This property makes ureaform an ideal direct-application nitrogen fertilizer for turf grass, ornamental shrubs, and greenhouse plants. A single annual application can satisfy the nitrogen requirements without the danger of burning the crop, as can be the case when more active nitrogen fertilizers are applied in excess or with too little water. Ureaform is also employed in batch and continuous processes to make dry-mix products and homogeneous granular N-P-K fertilizers. Care must be taken during these operations to avoid prolonged exposure to high temperature or to pH outside the range of 5.4–6.2. Urea is the preferred nitrogen supplement. Pulverized dolomitic limestone and hydrated lime are commonly used to compensate for the acidity of granulation ingredients such as superphosphate. Storage temperature of these N-P-K products must be maintained below 100°F (38°C) to avoid degradation of the ureaform.

Phosphorus. Phosphorus is a constituent of lecithin and other lipoids in plants. The average phosphorus, P_2O_5, content of whole plants is about one-third the nitrogen content. Highest concentrations are found in the germ and bran of grain crops. Phosphorus is the eleventh most abundant element in the Earth's crust. The elemental form of phosphorus does not exist in nature. Its minerals are widely distributed in the Earth's crust in moderate

concentrations. Many deposits in various parts of the world are used in commercial production of phosphate rock. Numerous forms of phosphate materials, including untreated phosphate rock, have been used for fertilizer. Large amounts of phosphates are needed just to replace the quantity carried to the sea by rivers in the form of land erosion, sewage discharge, and other chemical losses. It is estimated that combined annual phosphate losses to the sea from the United States alone are equivalent to about 60×10^6 tons (54×10^6 metric tons) of phosphate rock.

Phosphate rock. The principal source of phosphate for the fertilizer industry is phosphate rock, which is a generic name for ores of phosphorus-containing minerals. The chief phosphatic component of these ores is fluorapatite, with the general formula $3Ca_3(PO)_2 \cdot CaF_2$. The name phosphorite is given to sedimentary apatites. The phosphate content of the amorphous phosphorites is more easily released by the action of mineral acids than is that of the hard crystalline apatite. The phosphorites are therefore the favored commercial ores. The phosphate content of phosphate rock is frequently given in bone phosphate of lime (BPL) or % $Ca_3(PO_4)_2$, as well as in terms of P_2O_5. A high-grade rock of 75 BPL is equivalent to 34.4 wt % P_2O_5. Large tonnages of ground phosphate rock have been used directly for soil fertilization without further treatment. The availability of the P_2O_5 content of this phosphorite application is dependent upon the pH and mineral content of the soil. Ground phosphate rock is the least costly phosphate fertilizer, but release to the crop is slow under the best conditions. *See* APATITE.

Phosphate rock as mined is usually beneficiated (concentrated) to remove gangue and increase the P_2O_5 content for feed to other processes. Beneficiation includes washing, screening, crushing oversize, treating fines by flotation, and agglomerating the phosphate concentrage.

Superphosphate. To produce superphosphate, concentrated sulfuric acid, water, and ground rock are continuously proportioned into a mixer which intimately blends the reactants. The hot acidulated slurry from the mixer, such as the Kuhlman turbine which emulsifies the ground-rock acid, discharges into a continuous den. The den encloses a slow-moving conveyor, or a rotating plate where the slurry sets up to a spongy mass (consisting mainly of monocalcium phosphate hemihydrate, $Ca(H_2PO_4)_2 \cdot H_2O$, and calcium sulfate hemihydrate, $CaSO_4 \cdot {}^1/_2H_2O$) with the evolution of gas. At the end of travel the plastic superphosphate mass is broken up by a cutter and discharged onto a belt conveyor which delivers the product into storage to cure. The curing process allows complete reaction of the sulfuric acid by penetration of the phosphate rock, which maximizes the availability of the P_2O_5 content. Superphosphate recovered from curing storage is crushed and screened for sale as run-of-pile.

The superphosphate is nearly always dark in color because of residual carbon left in the product from charring of organic material in the rock and the spent acid during the acidulation reaction. Ammonia will react with monocalcium phosphate according to reaction (4). The heat of this reaction and the cement-

$$Ca(H_2PO_4)_2 \cdot H_2O + NH_3 \rightarrow CaHPO_4 + NH_4H_2PO_4 + H_2O \quad (4)$$

ing action of ammonium phosphate formed allow granulation of the mixture with little additional expenditure of energy. Overammoniation will cause a reversion to $Ca_3(PO_4)_2$ and a loss of P_2O_5 availability. The fertilizer value of the ammoniated superphosphate enhanced with the nitrogen content caused a strong demand for this product. The simplicity of the reaction and the effectiveness of the product allowed superphosphate to remain unchallenged in the top position of phosphate fertilizers for more than a hundred years from the time of the first patent, issued in 1842. Superphosphate, sometimes called normal or single superphosphate, continued to represent half of the world production of phosphatic fertilizers as recently as 1962. However, superphosphate fell in popularity because of its low P_2O_5 content and the rising cost of freight. To a large extent, it has been replaced by high-analysis fertilizers such as a triple superphosphate (TSP) and the ammonium phosphates (MAP and DAP).

Phosphoric acid. Phosphoric acid is a major commodity of the fertilizer industry and the chief source of phosphorus in modern commercial fertilizers. It is a strong mineral acid and is not applied directly to the soil as a fertilizer. The merchant grade of orthophosphoric acid, H_3PO_4, contains 54 wt % P_2O_5. *See* SOLVENT EXTRACTION.

1. *Production*. Phosphoric acid is made by either of two general routes: the electric furnace process or acidulation of phosphate rock.

Furnace acid, or white acid, is made by the electric furnace process. Coke and silica are mixed with coarse phosphate rock to provide a porous bed that is melted by heating to a high temperature with electric current passed through the electrodes and carbon blocks forming the bottom of the slag pool in a closed furnace that also serves as a retort. The phosphorus is reduced to the element, which vaporizes in the furnace. Dust is removed from the vapor with electrostatic precipitation. Air is excluded from the system, and the vapor is condensed to liquid white phosphorus by direct contact with water sprays. The liquid phosphorus is collected under a layer of water and pumped to a burner where it is oxidized with air to phosphorus pentoxide, P_2O_5, in a water-cooled, graphite-block, combustion chamber. The white fumes of P_2O_5 are absorbed in water or dilute acid in a spray tower to produce water-white, high-purity orthophosphoric acid. The combustion and hydration can also be done in a single tower. Because other minerals remain with the slag in the furnace, the furnace process can accept lower-grade phosphate rock and can tolerate large amounts of silica and more trivalent metal contaminant than the wet process. Ferrophosphorus is a byproduct of this process that can be used for steelmaking. The furnace reduction process is important for

production of elemental phosphorus for munitions and some important chemicals. The white acid was used for the production of high-analysis fertilizers when electric power was cheap and plentiful.

Wet process, or green, phosphoric acid is made by the acidulation of phosphate rock with dilute solutions of strong mineral acids, such as sulfuric, hydrochloric, or nitric. Practically all the phosphoric acid used in modern fertilizer production is made by the wet process. Sulfuric acid is most commonly used in this process. There are several variations in design among the proprietary processes, but in general, finely ground phosphate rock is digested with diluted acid in an agitated reactor, or in a series of them, where phosphoric acid is formed along with calcium sulfate, which is precipitated in a manner to yield a filterable dihydrate crystal slurry, according to reaction (5). A slight excess of sulfuric acid

$$3Ca_3(PO_4)_2 \cdot CaF_2 + 10H_2SO_4 + 20H_2O \rightarrow 6H_3PO_4 + 10CaSO_4 \cdot 2H_2O + 2HF \quad (5)$$

is used to maximize phosphorus recovery. The relatively insoluble calcium sulfate dihydrate (gypsum) crystals are filtered and washed with two or three progressive washes of filtrate followed by a water wash. The strongest wash is returned to the digestor with fresh sulfuric acid feed. The washed gypsum is usually discharged to a storage pond in a water slurry, but may be used for ballast in granulation of N-P-K fertilizers or reprocessed for manufacture of cement and wallboard. The filtered phosphoric acid is recovered at 28–30% P_2O_5, which is then usually concentrated in rubber-lined vacuum evaporators to merchant-grade acid. When the phosphate rock supplied to this process is first calcined to burn off the organic matter in it, the acid from the filter often has a rich green color due to the metallic elements in the ore. This coloration gave rise to the name green acid for wet-process phosphoric acid. The organic matter in uncalcined phosphate rock used in the digestion becomes charred by the reaction, and the colloidal carbon particles give the phosphoric acid a dense black color. In either case, the acid produced by this process is supersaturated with respect to a number of complex salts of iron, aluminum, magnesium, potassium, fluorine, and silica. The continuous precipitation of these complex salts has caused problems in storage and shipment of the acid. Much effort has been spent by industry to find a means of preventing the formation of the hard deposits of these materials.

2. *Purification*. Among the many purification processes for wet-process phosphoric acid produced with sulfuric acid, urea phosphate crystallization is the simplest and has the lowest energy demand. The calcination of phosphate rock to remove organic matter before acidulation is an effective means of achieving clear wet-process phosphoric acid. However, this energy-intensive process is expensive and is practiced only with phosphate rock containing high concentration of organics. Several proprietary extraction and ion exchange processes have been developed for the purification of phosphoric acid produced by rock acidulation with hydrochloric acid and with nitric acid. These processes are all expensive with respect to plant investment and operating costs. The limited cases of their application have resulted from unusual circumstances. It is unlikely that any of these processes will be in widespread use.

3. *Uranium recovery*. Phosphate rock contains a small percentage of uranium that can be recovered from wet-process phosphoric acid by organic solvent extraction. Much of the phosphoric acid produced in the United States is now being profitably extracted for uranium recovery. The dramatic increase in energy cost has prompted renewed efforts to improve the economics of phosphoric acid production. One variation in the design of the wet process is crystallization of calcium sulfate hemihydrate, $CaSO_4 \cdot {}^1/_2H_2O$, in the acid digestion, which occurs at higher temperature and greater P_2O_5 concentration than is the case with gypsum formation. Filtration of the hemihydrate requires greater filter surface because the crystals do not drain as rapidly as the gypsum crystals. Other variations include the foam process and the two-step recrystallization of hemihydrate to gypsum. The goal of these variations is the production of phosphoric acid of sufficient strength (40% P_2O_5) for granulation without further concentration of the filtered acid. Most facilities that manufacture phosphoric acid use all or part of the product captively to make granulated N-P-K fertilizers.

4. *Phosphate rock supply*. The highest grades of phosphate rock are preferred for long-distance shipments because freight has become an ever larger fraction of the delivered cost. The large demand for phosphate fertilizers in the past few decades has led to depletion of the highest grades of phosphate rock in currently active mines. Very large reserves of phosphate rock remain in the United States and other producing parts of the world, but the once readily available 75 BPL rock is in short supply and 68 BPL is now commonly shipped long distances. Even more important than P_2O_5 content of the rock to the production of wet-process acid is the contamination of specific minerals, of elements such as iron and aluminum that cause loss in production by troublesome precipitations.

5. *Superphosphoric acid*. Superphosphoric acid is the viscous product formed by further concentration of merchant grade to increase the P_2O_5 content to 70% or more. Wet-process phosphoric acid at 54% P_2O_5 concentration is equivalent to about 75 wt % H_3PO_4, with a viscosity of about 12 centipoise (0.012 pascal-seconds) at 100°F (38°C), and practically all phosphate is in the form of orthophosphoric acid. As the acid solution is further concentrated by the loss of free water, the temperature rises and the phosphoric acid gradually begins to polymerize to the dimer, pyrophosphoric acid ($H_4P_2O_7$), and higher condensed forms by the progressive loss of combined water. The viscosity increases rapidly with the increase of mineral contaminants. The viscosity of furnace-grade superphosphoric acid containing 40% polyphosphate at 100°F (38°C) is about

250 centipoise. The viscosity of wet-process superphosphoric acid also containing 40% polyphosphate at 100°F (38°C) will vary between 2000 and 20,000 (2 and 20 Pa · s) centipoise, depending upon the amount of impurities present. The superphosphoric acid must therefore be stored and handled in heated systems to retain fluidity. The furnace-grade superphosphoric acid is made by using less water in the absorption tower receiving the P_2O_5 from the burner. The temperature range of superphosphoric acid concentration (446°F or 230°C) is well above the acceptable limit for rubber-lined equipment, and so the concentration must be done in expensive alloys or carbon brick equipment. One superphosphoric acid concentrator design holds the acid in a conical pool of a chamber lined with carbon brick while it is heated by submerged combustion or the direct contact with hot combustion gases blown through the pool of acid. The large amount of acid mist generated by this operation is recovered by means of a high-energy venturi scrubber followed by a stationary entrainment separator. In addition to the high P_2O_5 content of superphosphoric acid made from wet-process acid, it has the ability to sequester impurities in the wet-process acid and prevent their troublesome precipitation.

Triple superphosphate. TSP, also called concentrated superphosphate, is made by the action of phosphoric acid on ground phosphate rock to give a product that is predominantly water-soluble monocalcium phosphate monohydrate. Unlike single superphosphate, TSP is not diluted with calcium sulfate, so that the P_2O_5 concentration is generally 40–49%, depending upon the quality of the rock used. The presence of other minerals in the rock leads to the formation of some citrate-soluble phosphates, such as iron and aluminum phosphates. These and other trace elements are of benefit to granulation of the TSP and contribute to micronutrient fertilizer value. Citrate solubility of phosphates is acceptable as P_2O_5 availability in the United States, but the legal requirement of water solubility is much more important in Europe and the United Kingdom.

Acidulation of phosphate rock with phosphoric acid releases silicon tetrafluoride and hydrofluoric acid, as does the acidulation with sulfuric acid, but the amount of fluoride released is dependent upon the phosphoric acid concentration, which also influences the temperature of the reaction with the phosphate rock. The speed of the reaction and the setting time of the reaction mass are also determined by the temperature and acid concentration.

Several processes have been developed which include liquid and solid reaction types with batch or continuous operation to make powdered or granular product. The equipment and reactions are similar to those for single superphosphate, but the mixing periods are shorter and the reaction mass solidifies more quickly in the production of triple superphosphate. The initial moisture content is greater and the curing time is longer. Most TSP is manufactured by continuous granulation of the reaction products, without the intermediate curing in a pile. The heat from the dryer in the granulation circuit speeds up the curing process, and the dry granular material is conveyed to storage. Shipment of TSP is seasonal in many parts of the world, and large bulk storage must be provided.

Ammonium phosphate. Wet-process monoammonium phosphate, $NH_4H_2PO_4$ (11-48-0), and diammonium phosphate, $(NH_4)_2HPO_4$ (18-46-0), have become popular fertilizers because of their high fertilizer analyses and good storage and handling characteristics. Both materials are commonly made in granulation circuits by reacting anhydrous ammonia with wet-process phosphoric acid. The high degree of ammoniation requires carefully designed scrubbing systems to recover and recycle excess ammonia carried in the air streams from the granulator and the dryer along with dusts generated in the solids-handling equipment. The fresh incoming acid is usually used to scrub the ammonia vapors and collect the dust. This partially ammoniated acid also containing dissolved ammonium phosphate dusts is then further ammoniated in a preneutralizer. The preneutralized hot slurry is distributed onto a bed of semiwet granular solids where more ammonia is introduced within the bed to react with phosphoric acid on the surface of the granules. The moist, blended granules are dried in a cocurrent rotary dryer and screened on double deck screens. The oversize particles are crushed and returned to the granulator along with the fines. The intermediate-size particles are withdrawn as required for production, cooled, and conveyed to storage, while the excess is returned to the granulator.

Many different types of reactor circuits and scrubbing systems have been designed to achieve these purposes. Two granulation devices most used are the ammoniating drum granulator and the pug mill granulator, or blunger. The drum granulator produces the harder and more uniformly round granules. Either type of circuit can be used to make a full range of N-P-K fertilizers by adding potassium compounds to the granulator feed. Consumption of the ammonium phosphates has steadily increased since the mid-1950s till these materials now constitute over 60% of the world usage of P_2O_5.

The SAI process makes a stable powder form of monoammonium phosphate (MAP) containing 6–12% moisture that has good storage and handling properties similar to superphosphate. The MAP is made with a slurry of wet-process phosphoric acid ammoniated to 1.35 N:P ratio, which is then reduced to 1.0 N:P ratio by the addition of more acid. The great difference in solubility of saturated ammonium phosphates causes the mixture to congeal with minimum solubility at the composition of MAP by the addition of phosphoric acid to the very soluble mixture of DAP and MAP. Mineral impurities in the wet-process acid influence the crystal growth of MAP in the process. The congealed MAP is broken up and dried at low temperature (257°F or 125°C) to no less than 6% moisture, which retains the good granulating characteristics of the material. Development of the pipe-cross reactor made possible the direct contact of ammonia vapor with phosphoric acid or

sulfuric acid at high temperature. This high reaction temperature allows ammonium phosphate production from orthophosphoric acid without the need for a preneutralizer. Much of the water in the phosphoric acid is evaporated directly in this reactor with reaction heat, which reduces the recycle rate and the energy requirement in the granulation circuit. The low-moisture-content melt (less than 2%) fed to the granulator from the pipe-cross reactor also eliminates the need of a dryer for some N-P-K grades. The reduction of materials handling within the process unit also reduces air-pollution problems. The favorable economics of this operation has caused many producers to retrofit granulation plants with the pipe-cross reactor to manufacture high-analysis granular N-P and N-P-K products.

Potassium. Potassium has several important functions in plant metabolism. It appears to play an essential role in the water economy of plants through the action of its soluble compounds on osmotic pressure. Potassium is involved in oxidative reactions as a carrier for iron required by important enzymes that convert sugar to starch and amino acids to protein. Sugarbeets, potatoes, and other plants that produce large quantities of carbohydrates use large amounts of potassium.

Complex potassium minerals are widely distributed, and most are insoluble. However, some extensive mineral deposits of single and double salts of chlorides and sulfates are soluble and provide the major sources of potassium supply. Most of the potassium supplied to the fertilizer industry is in the form of the commercial-grade crystalline chloride salt (60% K_2O). Potassium sulfate (50% K_2O) is supplied for some crops that require potash fertilization, such as tobacco, and are sensitive to chloride. The potassium salts are usually mixed with other fertilizer materials in dry-blended formulas or in homogeneous granules. Solution and suspension fertilizers containing potassium compounds gained in popularity because of lower cost resulting from processes that require less energy than those involving granulation or drying and blending of solids.

Potassium salts are also present in minor concentration in the caliche deposits of Chile. Development of the Guggenheim process made possible the economic recovery of potassium nitrate, KNO_3, along with other minor minerals in conjunction with production of sodium nitrate. Potassium nitrate is an expensive, but totally soluble high-analysis fertilizer that combines both N and K_2O in a single material containing 60 units of plant nutrients. It is favored for special formulations that are applied in solution form to some chloride-sensitive crops. *See* POTASSIUM.

Sulfur. Sulfur is widely distributed in plant tissue. It is a component of the amino acids methionine and cysteine, and is also found in some hormones and in oil glycosides that impart odor and flavor to mustard, garlic, and onions.

Large deposits of elemental sulfur trapped geologically in deep salt domes along the Gulf coast of the United States and Mexico are economically mined with superheated steam by the Frasch process. Sulfur is recovered in variable quantities in the form of hydrogen sulfide from natural gas, once burned in large quantity in flares, that is now recovered as elemental sulfur by the Claus process. Environmental regulations have brought about the recovery of sulfur dioxide from the metallurgical processing of sulfide ores of iron, copper, lead, zinc, mercury, and antimony. It is usually most economic to convert the recovered sulfur dioxide directly to sulfuric acid, provided that the acid can be used close to the site of production. The recovery of sulfur dioxide from the combustion of oil and coal is an environmental necessity that is a source of sulfur and sulfuric acid for the fertilizer industry. Enormous deposits of high-quality gypsum (calcium sulfate dihydrate) are located close to the surface and available for recovery of sulfur values in the event that the more easily converted forms of sulfur should become insufficient for world needs. *See* GYPSUM.

Most soils contain substantial amounts of sulfates; and fertilizers such as superphosphate, ammonium sulfate, and ammonium sulfophosphate, commonly used in the past separately or in mixtures, had ample sulfur replacement value to maintain growth requirements. These and other fertilizer materials were generally credited for their primary fertilizer elements, and the presence of other nutrient elements, such as sulfur and calcium, were ignored. The trend to high N-P-K analysis fertilizers to offset the rising cost of freight led to evidence of sulfur deficiency in some heavy-sulfur-consuming crops. Sulfur deficiency manifests itself by marked chlorosis and retardation of plant growth. The extensive switch to triple superphosphate and ammonium phosphates from normal superphosphate has practically eliminated the generous amount of sulfate that was formerly provided without charge or recognition and applied to crops. Continued and expanded evidence of sulfur deficiency has brought about the focusing of greater attention to the secondary elements in commercial fertilization programs. From a tonnage standpoint, sulfur is even more important to the fertilizer industry in the form of sulfuric acid as a means of providing phosphorus availability from phosphate rock. This is by far the biggest worldwide use of sulfur. *See* SULFUR; SULFURIC ACID.

Calcium. It is believed that calcium is utilized in continuous cell division and is involved in nitrogen metabolism because the absorption and assimilation of nitrates in plants suffer retardation when calcium is deficient. Calcium is not a mobile element in plants in the manner of phosphorus. Besides the fertilizer function of calcium, it is applied to the soil to correct excessive acidity and improve the physical structure of the soil. The application of lime to an acid soil in an amount to change the pH from 6.5 to 7.5 creates a condition favorable to the normal activities of beneficial soil microbes. Calcium is leached out of soils in the form of calcium biocarbonate. The trend to produce very high-analysis fertilizers has resulted in severe reductions in calcium that was previously applied along with N-P-K formulas prepared from normal superphosphate. Though calcium is a widely

distributed element in nature, the practice of eliminating its inclusion in fertilizers normally applied to the fields may in the future result in deficiency that is not evident at present. Calcium deficiency is less likely to appear in Europe where the acidulation of phosphate rock with nitric acid is still practiced and nitrogen is frequently applied in the form of calcium nitrate and calcium ammonium nitrate and some calcium cyanamide. Phosphate fertilizer in the form of basic slag from steel production contains about 17 wt % P_2O_5 as tetracalcium phosphate and other compounds of calcium. Basic slag is a common fertilizer item in Europe but of little importance in the United States. Phosphate rock is a rich source of calcium, and it is also available in huge natural deposits of limestone and gypsum. *See* CALCIUM.

Magnesium. Magnesium is the central atom in the structure of chlorophyll, which is vital to photosynthesis; most of the magnesium found in plant tissues is concentrated in the green parts, which contain maximum chlorophyll. Magnesium is credited with translocation of phosphorus in the plant and participation in the activity of enzymes. The disruption of plant processes due to a deficiency of magnesium is manifested by chlorosis, or yellowing of the green tissues. Magnesium is the eighth most abundant element in the earth's crust and is widely distributed in the soil in the form of carbonates and silicates. Magnesium minerals are more soluble than their calcium counterparts, and leaching losses occasionally cause a soil deficiency. Magnesium supplements are often supplied by liming with dolomite, $CaMg(CO_3)_2$. Other minerals commonly used as magnesium fertilizer are langbeinite, $K_2SO_4 \cdot 2MgSO_4$, and magnesium sulfate. *See* CHLOROPHYLL; MAGNESIUM; PHOTOSYNTHESIS.

Micronutrients. The micronutrient elements essential to plant growth are usually present in the soil in concentrations adequate for the small quantitative demands of the crops. Natural distribution of the elements is, however, nonuniform, and some shortages do occur. Also, long-term usage of the land can deplete the natural supply of these elements. Moreover, it is possible that applications of pesticides and fertilizer materials to supply primary and secondary elements may react in a way to reduce the availability of one or more of the micronutrients. Replacement of these fertilizer elements required in trace amounts poses a problem because of the narrow concentration margin between deficiency and toxicity. It is necessary that the desirable compound of the micronutrient be diluted in the applied material to avoid local high concentrations that would occur in the soil in the immediate vicinity of a concentrated granule of the micronutrient. This is achieved by adding a small percentage of the compound in the manufacture of granulated N-P-K formulas, liquids, or slurry mixtures. These dilute forms can then be distributed over the soil to assure uniformity of application. The distribution of a dry blended mixture of solids of the same overall concentrations cannot achieve the necessary degree of dispersion that is possible with the diluted materials.

Boron. Boron is an essential micronutrient that is required in very low concentration. The role played by boron in the life process of plants is unknown, but specific deficiency symptoms traced to shortages of boron include leaf-roll of potato, browning of cauliflower, and brown heart of turnip and sugarbeets. The desirable soil concentration ranges very considerably from crop to crop. Tomatoes do best with boron concentration below 1.0 ppm, while asparagus development is favored with boron concentration as high as 15 ppm. Boron-deficient soils can be adjusted by applications of borax (sodium tetraborate), boric acid, or boron frits in the order of 2.7–16 lb/acre (0.5–3.0 kg/hectare). Faster corrections can be made with foliar applications, but the range of concentration between optimal feed and toxicant is narrow. *See* BORON.

Cobalt. Cobalt concentrations in the range of 0.1–0.3 ppm are important to healthy plant growth, but the function of cobalt in plant metabolism is obscure. Fertilizer forms of cobalt are the soluble salts of sulfate, chloride, or nitrate, which may be applied admixed in aqueous solutions slurries, or homogeneous granules of N-P-K products. *See* COBALT.

Copper. Copper is an important nutrient that must be present in very low concentration. Copper has an essential, though not fully understood, role of participation in oxidizing-reducing enzymes, such as tyrosinase and ascorbic acid oxidase. Excess amounts can be very toxic. Copper sulfate (blue vitriol) is used as an agricultural poison and as an algicide in water purification. *See* COPPER.

Iron. Iron as a micronutrient is most effective in the ferrous state, $(Fe)^{2+}$. Though it can be absorbed in the ferric state, $(Fe)^{3+}$, iron is reduced to ferrous form in plant tissues. Iron is intimately associated with several enzymes which take part in the oxidative reactions of living cells. It is required as a catalyst in the elaboration of chlorophyll. Iron is not mobile within the tissues, and there is little redistribution of the plant's reserve under conditions of supply deficiency. The availability of iron in the soil is strongly influenced by the pH (or degree of acidity) and is favored by soil conditions that are slightly acidic. The amount of iron found in green plants varies within the range of 10–1500 ppm with deficiency evidenced by chlorosis. Where soil conditions tend to lock iron in less available form, soil or foliant applications of organic chelated iron compounds can relieve the plant deficiency. *See* IRON.

Manganese. Manganese is widely distributed in nature, and only small soil concentrations are required for healthy plant growth. In the manganous form this element is an activator of some enzyme systems, such as carboxylases and dehydrogenases. Plants suffering from manganese deficiency exhibit a chlorosis distinctly different from the symptoms of magnesium or iron deficiency. Soil availability of manganese is favored by slight acidity and much restricted by alkaline conditions. Excess manganese can be very toxic to plants. An excess of this element acts as an oxidizing agent to convert ferrous iron to the ferric state, which is physiologically inactive in plants and

thus also contributes to iron deficiency. Manganese is believed to be essential in animals for utilization of vitamin B_1. *See* MANGANESE.

Molybdenum. Molybdenum is the plant nutritional element required in the smallest concentration. The requirement varies among crops, but deficiency in tomato plants is prevented by concentration in the order of 1 part in 100,000,000 in the cultural solution. One of the important functions of molybdenum in plant metabolism is the reduction of nitrates. It is also associated with tannin synthesis and legume nodule formation in nitrogen fixation. *See* MOLYBDENUM.

Zinc. Zinc is an essential growth element for both plants and animals. The amounts of zinc found in whole green plants vary between 3 and 150 ppm. Zinc has a function in relation to the enzyme carbonic anhydrase and is known to be required for the synthesis of indoleacetic acid. Excess soil concentrations can be harmful. Zinc is less available in alkaline soils than those on the acid side, but the effect is less pronounced than it is in the case of manganese or iron. The sulfate is commonly used for fertilizer in zinc-deficient soils. Organic chelates of zinc are sometimes used in foliant applications to ornamental plants, such as gardenias. *See* CHELATION; PLANT MINERAL NUTRITION; ZINC.

Alan Longacre

Organic Fertilizers

Organic fertilizers are organic materials of vegetable and animal origin which contain certain macro, secondary, or micro nutrients that can be utilized by plants after application to agricultural soils. Some materials can be applied beneficially to crops in an unprocessed state, or after processing by such technologies as composting, drying, fermentation, anaerobic digestion, co-composting, and chemical stabilization. The primary nutrient sources of vegetable origin are crop residues, green manures, oilseed cakes, seaweeds, and miscellaneous food processing and distillery wastes. Also included in this category is biologically fixed nitrogen from legumes in association with root-nodulating bacteria of the genus *Rhizobium*. Animal sources include animal manures and urine, sewage sludge, septage, latrine wastes, and to a lesser extent materials such as blood meal, bone meal, and fish scraps. Often organic fertilizers are of mixed animal and vegetable origin, such as most farmyard manures, rural and urban composts, and sewage effluents and sludges. *See* NITROGEN FIXATION; SOIL MICROBIOLOGY.

The **table** lists the nitrogen, phosphorus, and potassium content, or fertilizer value, and the carbon:nitrogen ratio of some of the most widely used organic fertilizers. The plant nutrient content for most of these materials is quite low compared with most modern synthetic chemical fertilizers. There is also considerable variation in the nutrient content among these organic materials. Those richest in nitrogen include wastes of animal origin, oilseed processing, and legume green manures. Cereal straws, corn stover (the stalks or nongrain part of the crop), seaweeds, farmyard manures, and composts are among the sources that are generally lowest in nitrogen. Likewise, fertilizers of animal origin are richer in phosphorus than crop residues, seaweeds, composts, and farmyard manures, which have the lowest phosphorus content of the materials listed in the table. Sources highest in potassium include plant residues and animal manures, while sewage sludge and some composts are lowest. The nutrient percentages listed in the table are mean values, which can vary considerably depending on the age and diet of animals, processing and storage conditions, and the soil and climatic conditions under which plants are grown.

Use. Organic fertilizers are often used to supplement chemical fertilizers (or vice versa) or as a soil amendment to improve the physical and chemical properties of soil. In some farming systems, such as organic or biological agriculture, they are the primary nutrient sources applied to field and horticultural crops. When organic fertilizers are used primarily as nutrient sources, their chemical composition is important. However, the availability of their contained nutrients for plants, especially nitrogen, may depend more on the C:N ratio than on the nutrient content itself. Materials with a high C:N ratio (for example, wheat straw and corn stover) generally have a low potential for release, or mineralization, of nitrogen as well as other nutrients to a plant-available form. When applied to soil, such materials may lead to a temporary decrease in plant-available nitrogen because of the concomitant demand for nitrogen by soil microorganisms engaged in residue decomposition. Materials with a lower C:N ratio, such as animal or green manures, would release their nitrogen and other nutrients to crops much more readily with little danger of microbial immobilization.

Types. Thus, organic fertilizers can be divided into general types. First, those which have a high nutrient availability index (NAI) are rapidly decomposed by soil microorganisms, and are essentially unstable organic materials. Examples include fresh animal

Nutrient composition of selected organic fertilizers of vegetable and animal origin*

Fertilizer	C:N	N, %	P, %	K, %
Vegetable				
Wheat	100	0.5	0.1	1.1
Corn stover	55	0.6	0.3	1.3
Green manure (alfalfa)	19	2.5	0.3	2.0
Cottonseed cake	10	6.4	1.3	1.8
Seaweed	17	0.7	0.3	1.2
Animal				
Cattle manure	19	1.9	0.6	1.4
Poultry manure	12	3.8	1.9	1.8
Latrine wastes	8	4.0	1.4	1.3
Bone meal	—	3.4	10.9	0
Blood meal	3.5	11.1	0.7	0
Sewage sludge	10	4.0	1.4	0.3
Fish scraps	4.5	7.5	2.8	0.8
Mixed				
Farmyard manure	15	0.8	0.2	0.7
Composted sewage sludge	19	1.3	0.9	0.2
Composted cattle mature	17	0.4	0.2	0.2

*Most data compiled by the Food and Agriculture Organization (1983).

manures, sewage sludge, and latrine wastes. A second group of materials are characterized by a high organic stability index (OSI), are decomposed relatively slowly in soil, and have a relatively low NAI. An example is composed organic wastes. Materials having a high NAI generally have a low OSI. Farmers in both developed and developing countries utilize both types of materials depending on their need to release nutrients rapidly to a vegetable crop, or to improve the physical properties and productivity of a marginal soil.

Production methods. The value of some organic wastes and residues as organic fertilizers and soil conditioners can often be increased through certain biological or physical processes which enhance consumer demand, storage, handling, and field application. Some of the more common methods include composting, drying, and shredding or grinding.

Compost is produced by the controlled decomposition of organic materials such as manure, sewage sludge, and plant residues by microorganisms under specific conditions of time, aeration, temperature, and moisture. Composting provides organically stable, humuslike products which no longer resemble the input raw materials, and which can be more easily stored, handled, and applied. The process is usually carried out in a compost pit or by heaping the material in stacks aboveground. Composting also reduces bulk and destroys pathogenic microorganisms which may be present in some raw materials.

Drying, usually by exposure to wind and sunlight, is used to remove the large amounts of water present in some materials such as sewage sludges and animal manures. Grinding or shredding can reduce the bulk of coarse materials such as bones and some woody plant residues which are difficult to use in their original form.

Advantages and disadvantages. A distinct advantage of organic fertilizers over concentrated chemical fertilizers is that they release their plant nutrients more slowly over a greater period of time. This slow-release characteristic can greatly minimize the potential for loss through leaching and runoff. Moreover, the organic component is highly beneficial for improving soil physical properties, resulting in increased infiltration and retention of water, soil aeration, and plant growth. Organic fertilizer may also serve as a buffering agent for stabilizing soil pH, and as a complexing agent for reducing phytotoxic effects of aluminum, iron, and manganese under acid soil conditions. Disadvantages of organic fertilizers include their relatively low nutrient analysis, which involves greater bulk and weight than high-analysis chemical fertilizers, and their considerable cost in transportation and labor for application. Continued use of some organic materials as sole sources of fertilizer can lead to an undesirable imbalance of nutrients in the soil. However, where both vegetable and animals sources are applied, there is little likelihood of this occurring. Other potential disadvantages of organic fertilizers are related to public acceptance of some materials (for example, latrine wastes and sewage sludge), health hazards from pathogens and fly breeding, and malodors. However, most of these problems can be overcome.

Quantities available. Organic fertilizers are a major source of plant nutrients for crop production in the United States and in many developing countries. For example, of the approximately 23×10^6 tons (21×10^6 metric tons) of nitrogen returned annually to cropland by United States farmers, 55% originates from organic sources, including biologically fixed nitrogen. This includes 3×10^6 tons (2.7×10^6 metric tons) from crop residues, 1.4×10^6 tons (1.3×10^6 metric tons) from animal manures, and 7.2×10^6 tons (6.5×10^6 metric tons) from biologically fixed nitrogen in legume and green manure crops. Crop residues and animal manures account for about 75% of the organic materials produced annually in the United States, while municipal refuse (garbage or solid waste) accounts for about 18% of the total, and sewage sludge less than 1%. More than 70% of the crop residues and animal manures (up to 35% of the sewage sludge, but only a small percentage of municipal refuse) are utilized as organic fertilizers. It is estimated that the amount of nitrogen, phosphorus, and potassium, respectively, returned to the soil annually in the form of organic fertilizers is equivalent to 85, 55, and 150%, respectively, of the total amount of these nutrients applied as commercial fertilizers. *See* AGRICULTURAL SOIL AND CROP PRACTICES; FERTILIZING; SOIL CHEMISTRY.

R. I. Papendick; J. F. Parr

Bibliography. O. P. Englestad (ed.), *Fertilizer Technology and Use*, 3d ed., 1986; A. Finck, *Fertilizers and Fertilization: Introduction and Practical Guide to Crop Fertilization*, 1982; Food and Agriculture Organization, *Comparative Effects of Mineral Fertilizers and Organic Manures*, Project Field Doc. 14, Regional Project RAS/75/004, FAO/UNDP, New Delhi, 1983; T. P. Hignett (ed.), *Fertilizer Manual*, 1985; J. A. Kent (ed.), *Riegel's Handbook of Industrial Chemistry*, 9th ed., 1992; W. Lockeretz (ed.), *Environmentally Sound Agriculture*, 1983; V. Sauchelli, *Phosphates in Agriculture*, 1965; A. V. Slack and G. R. James, *Ammonia*, pt. 1, Fertilizer Science and Technology Series, vol. 2, 1973; U.S. Department of Agriculture, *Improving Soils with Organic Wastes*, 1978; U.S. Department of Agriculture, *Report and Recommendations on Organic Farming*, 1980; W. H. Waggaman, *o-Phosphoric Acid, Phosphates, and Phosphatic Fertilizers*, 2d ed., ACS Monogr. 34, 1952; R. D. Young and F. P. Acorn, *Trends in U.S. Fertilizer Technology*, Nat. Fert. Dev. Center, TVA, Bull. Y-133, 1978.

Fertilizing

Addition of elements or other materials to the soil to increase or maintain plant yields. Fertilizers may be organic or inorganic. Organic fertilizers are usually manures and waste materials which in addition to providing small amounts of growth elements also serve as conditioners for the soil. Commercial fertilizers are most often inorganic. Fertilizer analysis and

systematic application began about 1850 and marked the beginning of scientific crop production.

Grades. Commercial fertilizers are mainly designed to supply one or more of the three major elements nitrogen, N, phosphorus, P, and potassium, K, in suitable chemical form. The fertilizers are graded in the order N-P-K, the numbers indicating the percentage of the total weight of each of the three components. Hence, the numbers 5-10-10 represent a mixture containing 5% nitrogen, N, 10% phosphorus pentoxide, P_2O_5, and 10% potassium oxide, K_2O.

Physical forms. Chemical fertilizers are marketed in both dry and liquid forms. Dry forms include powdered, granulated, and pelleted fertilizers. Liquid fertilizers are obtainable in high-pressure and low- or nonpressure forms. Mixed fertilizer solutions are usually nonpressure liquids.

Application methods. Methods of applying fertilizers vary widely and depend on such factors as kind of crop and stage of growth, application rates, physical and chemical properties of the fertilizer, and soil type. Two basic application methods are used, bulk spreading and precision placement. Time and labor are saved by the practice of bulk spreading, in which the fertilizer is broadcast over the entire area by using large machines which cover many acres in a short time. Precision placement, in which the fertilizer is applied in one or more bands in a definite relationship to the seed or plants, requires more equipment and time, but usually smaller amounts of fertilizer are needed to produce a given yield increase.

When applied at planting, the fertilizer is usually placed in a single band 1–3 in. (2.5–7.5 cm) to the side of and 2–6 in. (5–15 cm) below the seed (**Fig. 1**). Care must be taken to have enough soil between the seed and fertilizer to prevent damage to the seed or developing seedlings. For crops highly susceptible to fertilizer injury, only a small amount of starter fertilizer may be applied at planting, with the remainder being applied as a side-dressing after the plants have emerged. Several side-dressings may be applied during the growing season to those crops with a high requirement for nitrogen or other elements which leach easily. Gaseous or volatile liquid fertilizers are applied 5–6 in. (12.5–15 cm) deep by using a narrow chisel tool and sealing the opening with a press wheel to prevent losses.

Fig. 1. Efficient placement of fertilizer for potatoes. Attachments to make fertilizer placement at the time of planting are available on most field row planters. (*USDA, Agriculture Research Service*)

For some deep-rooted plants, subsoil fertilization to depths of 12–20 in. (30–50 cm) is advantageous. This is usually a separate operation from planting, and uses a modified subsoil plow followed by equipment to bed soil over the plow furrow, thereby eliminating rough soil conditions unfavorable for good seed germination. Top-dressings are usually applied by broadcasting over the soil surface for closely spaced crops such as small grains. Side-dressings for row crops may be placed at depths from surface to 5 in. (12.5 cm), but care must be taken to avoid damage to plant roots from the plow foot.

Equipment. Since solid fertilizers range from dense heavy materials to light powders and liquid fertilizers range from high pressure to zero pressure, a variety of equipment is required for accurate metering and placement. In addition, application rates may be as low as 50 lb/acre (56 kg/hectare) or as high as 6 tons/acre (13.3 metric tons/hectare). Large bulk spreaders usually use drag chains or augers to force the material through a gate or opening whose size is varied to regulate the amount passing through and falling on the spreader. One type of spreader consists of rapidly rotating horizontal disks with vanes mounted on their tops. The fertilizer is spread by centrifugal force over widths varying from 10 to 50 ft (3 to 15 m). Although this method is simple, the accuracy of distribution is affected by particle size, wind velocity, and slope of land on which the application is made. Another spreader consists of long perforated tubes through which the fertilizer is moved by augers.

Often the main supply bin contains simple agitators to prevent heavy or damp materials from bridging. The star-wheel metering unit is commonly used on drills and some row-crop fertilizer hoppers. This dispenser operates horizontally, with fingers carrying increments of fertilizer under a shield which is adjustable in height. Some drills have hoppers with rounded perforated bottoms. The hole size is adjustable by a slide, and the fertilizer is forced through by agitators mounted above the bottom of the hopper. Other mechanisms have been used, but basically all regulate the flow by varying opening size or speed of star wheel or agitator.

Liquid fertilizer of the high-pressure type (for example, anhydrous ammonia) is usually regulated by valves or positive displacement pumps. The size of the orifice may be controlled manually or automatically by pressure-regulating valves. Low-pressure solutions may be metered by gravity flow through orifices, but greater accuracy is obtained by using compressed air or other gases to maintain a constant pressure in the tank. This method eliminates the effect of temperature and volume changes.

Nonpressure solutions may be metered by gravity or by gear, roller, piston, centrifugal, or hose pumps. The accuracy of the gravity type can be improved by the use of a constant head device by which all air is introduced into the tank at the bottom.

All of the pumps except the piston and hose types must be calibrated at a fixed pump speed since their output is not linear. Therefore, the overall accuracy of placement is dependent upon maintaining a constant driving speed, which is very difficult on rolling terrain. Errors of 15% or higher may be introduced by variations in driving speed. The piston and hose pumps can be driven by a wheel rolling over the ground, making their ouput proportional to driving speed. The hose pump consists of any number of hoses which are compressed by rollers in a sequential action which squeezes the liquid along the hoses. Since the output is determined only by the speed at which the pump is rotated, no manifolds, valves, or orifices are needed. Nonvolatile fertilizer solutions are often pumped into the supply lines of irrigation systems to allow simultaneous fertilization and irrigation. With the exception of bulk spreaders and other broadcasters, most fertilizer application devices are built as attachments which can be mounted in conjunction with planters, cultivators, and herbicide applicators. Often the tanks, pumps, and controls used for liquid fertilizers are also used for applying other chemicals such as insecticides.

J. G. Futral

Effect on growth and yield. For plants to develop, all the nitrogen and mineral elements essential for growth must be supplied by way of the root system or, to a limited extent, through the leaves. The total supplies of mineral elements in certain soils are often adequate for many years of crop production, but the rate at which those elements become available for plant use may be too slow. Each nutrient must be in adequate supply in intensity and capacity, and in a reasonably favorable balance with all the others. Hence potassium, phosphorus, and any of several micronutrients may be deficient in rate of supply because of fixation or adsorption by the soil; lack of nitrogen may become a limiting factor through adverse microbial conditions or excessive loss by leaching. Moreoever, the rate of supply of nutrients must be favorable throughout the various ontogenetic stages of growth.

Desirable characteristics. Fruit or storage organs, rather than merely vegetative yield, are often desired. Further, it is frequently the total content of some substance within the harvested organ that is important rather than the weight or size of the organ alone. Thus the quality of tomatoes or grain or the percentage of sugar in cane or beets is economically important. The mineral supply at later stages of growth may determine the economic value of the crop. For example, high nitrogen acquisition at later development stages may lead to increased protein but decreased carbohydrate content. Thus mineral nutrition must be in favorable accord with development of desirable characteristics of vegetative growth and the harvested product.

Nutritional requirements. Plant growth and the quality of the harvested product are dependent on proper nutrition. Mineral materials are usually derived from the soil and, while total supply is important, the rate of effective availability is paramount. Plant and soil form an integrated system, in which the root is the nutrient-absorbing organ. While the type of root development is a species characteristic, its distribution in space may depend on environmental conditions in the medium. Thus it may vary in lateral and vertical disposition and in production of root hairs. Modifications also depend on physical and chemical conditions in the soil. Cultural management, such as cultivation, water, fertilization, spacing (competition), and cropping regime, may be determinative. If rooting into areas of nutrient adequacy is extensive (other factors being satisfactory), additional fertilizer may not be essential. Soil depletion, however, requires resupply to meet plant demand; deficiencies must be corrected for optimum growth. Growth is dependent on nutrient supply. Conversely, since plant need varies with rate of growth and the stage of ontogenetic development, the time, placement, and rate of fertilization are important. Some typical fertilizer placements are illustrated in **Fig. 2**. Plant growth, nitrogen distribution and total sorption in corn are shown in **Fig. 3** and in the **table**.

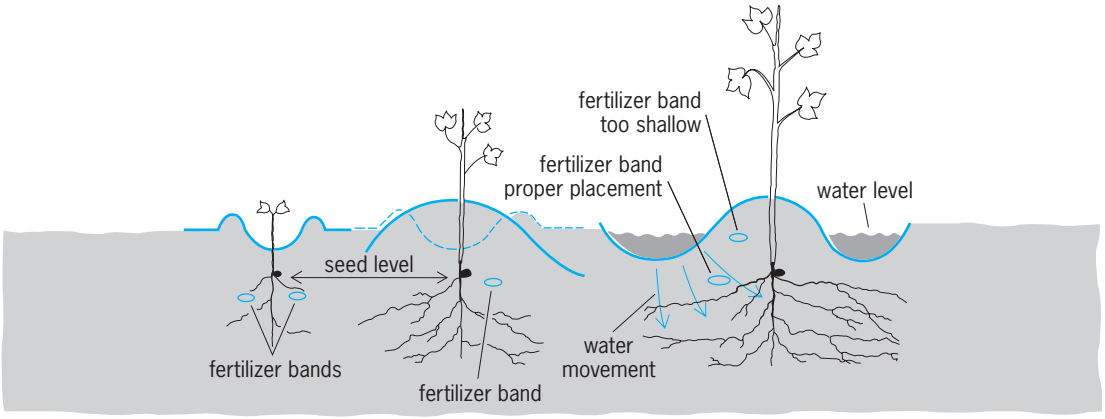

Fig. 2. Relationship of the placement levels of plant food materials to the movement of water. (*After Western Fertilizer Handbook, California Fertilizer Ass., 1965*)

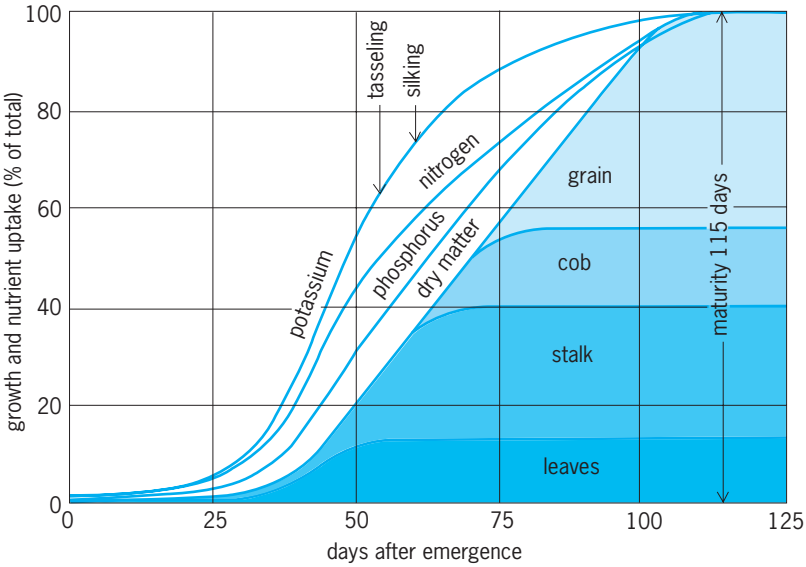

Fig. 3. Graph showing the growth and the uptake of nutrients in relation to the accumulation of dry matter by corn plants. (*After C. E. Millar, L. M. Turk, and H. D. Foth, Fundamentals of Soil Science, Wiley, 1965*)

Nonlegumes may require greater nitrogen fertilization than legumes. In this respect, the microbiological population and other soil factors, such as temperature, aeration, pH, and organic matter, are important. Bacterial species and their numbers are an important facet of fertility relations in general, especially insofar as nitrogen fixation and the presence of organic matter are concerned. Pathological disease incidence varies with fertilization. Fertility relations affect evapotranspiration by modifying the ratio of canopy cover to soil surface area, thus determining efficiency of water use. Rate of replenishment of nutrients from the soil or other medium may limit growth, as root sorption may create depletion zones in the rhizosphere of roots. *See* NITROGEN FIXATION; SOIL MICROBIOLOGY.

Fertilizers. For improving the mineral nutrition of plants, single materials or favorable combinations of chemicals are compounded and used. Chemical compounds may include all of the mineral elements that are essential to growth. Micronutrients, occurring as impurities in fertilizer-grade chemicals, in some cases may be inadequate for maximum plant development and extra supplies may have to be added. Some latitude is permissible in the balance of elements supplied, but both marked deficiencies or marked excesses of any one element must be avoided.

Balanced fertilizers. Commercial fertilizers, manures, lime, or other materials may be added to provide essential mineral elements at required levels of availability or as amendments to make environmental conditions more favorable, improve structure, or enhance microbial activity within the soil so that such factors do not become limiting to growth and development. Sometimes reference is made to the use of balanced fertilizers. The balance that is important is not in the fertilizer but in the soil after the fertilizer has been added and has reacted with the soil. The same fertilizer may produce different effects in different soils. Addition to the soil does not necessarily ensure availability; certain elements, such as potassium, phosphorus, and at least some of the micronutrients, are strongly fixed in some soils. This is especially true where the colloidal clay content of the soil is relatively high. In any case, what is important is the ability of the plant to acquire essential nutrients in proper quantities favorable for the species at any particular time during growth and development, assuming other factors to be optimal.

Soil type and composition. When considering application of fertilizer to soil, general information should be sought regarding the types of soil in the area and their physical and chemical compositions. Crops favorable to the representative soil and climatic conditions should be considered. Local agricultural information on these matters is usually available. "Feeding" zones vary with species and soil management. Shallow-rooted species absorb from and deplete zones different from those of deep-rooted crops. As roots grow, absorb, and die and after root decomposition, minerals may be redistributed thereby in the soil and thus provide a more available source of nutrients for succeeding plantings. Fallowing allows replenishment from parent and residual organic materials. Total salt and pH tolerance of species, as well as the total specific salt content and pH of the soil, must be considered. Protracted applications of single or even some mixed fertilizers may correct unfavorable conditions or cause adverse changes in total salt level, exchangeable ion ratios, or pH of the soil. The

Approximate plant growth, sorption of primary nutrients, and water use for a 100-bu corn crop per acre or hectare

Plant part	Basis	Dry weight	Nutrient sorption* N	Nutrient sorption* P	Nutrient sorption* K	Water use†
Shoots	Acre	13,000 lb	155 lb	26.3 lb	104 lb	6.5×10^4 lb
	Hectare	14,600 kg	173.7 kg	29.5 kg	116.5 kg	7.3×10^4 kg
Roots	Acre	5,000 lb	50 lb	8.7 lb	33 lb	2.5×10^4 lb
	Hectare	5,600 kg	56.0 kg	9.7 kg	37.0 kg	2.8×10^4 kg
Shoots and roots	Acre	18,000 lb	205 lb	35 lb	137 lb	5.5×10^6 lb
	Hectare	20,200 kg	229.8 kg	39.2 kg	153.4 kg	6.2×10^6 kg

*Apart from certain possible luxury accretion, the total sorption of nutrients represents that necessary during growth. That part in roots is restored to the soil. Therefore the net removal of nutrients is that contained in the shoots, unless stover is returned to the soil.

†Total water use during growth (5.5×10^6 lb or 6.2×10^6 kg) includes that for growth and that expended in evapotranspiration. The water in shoots is in large part lost to the air on crop maturity; that in roots is largely returned to the soil.

nature of the materials and of all interrelated facets of cropping and soil dynamics determine the result. For example, simple leaching may correct total salinity; sulfur, with microorganism activities, may correct alkalinity; lime may correct acidity; and gypsum may correct an unfavorable ion-exchange balance and the physical structure of soils.

Types of fertilizers. Fertilizer materials may be organic (urea) or inorganic (ammonium phosphate, potassium nitrate); natural (Chilean sodium nitrate, calcium sulfate) or synthetic (ammonium sulfate); simple, mixed (two or more), or complete (containing N, P, and K). They may be applied to soil in gaseous (ammonia), liquid (ammonia, or solutions), or solid (rock phosphate) states. Filler or make-weight material may be added to solid concentrates. Numerous factors must be considered in deciding the most productive mode, place, amount, balance, and time of application to fulfill known or anticipated plant needs. Knowledge of specific crop requirements at various stages of development (seedling, vegetative, reproductive), soil and irrigation water analysis, periodic plant tissue tests, and previous growth and harvest results will assist in suggesting management practice.

Maintaining adequate soil conditions. Natural and artificial materials can replenish nutrients as well as improve structure and water-holding capacity of many soils. Temperature, pH, aeration, and activity of microorganisms determine the rate of decomposition and supply or loss of nutrients. Natural manures are good sources of N, P, and K; stubble may be low in N, P, and Ca, but may be supplemented with ammonium sulfate, superphosphate, and limestone. Cover crops, especially legumes, are a means of effectively supplying natural organic matter to soils and favorably altering nutrient availability. Similarly, crop rotation can usefully provide protracted land use while meeting economic supply and demand. It is often necessary on a long-term basis to obviate or correct development of certain soil pathogens. Nurse crops under some conditions (for example, in orchards) can be beneficial. Manures, though beneficial on many soils, are not necessarily indispensable. Fully satisfactory plants or their products can be attained with soilless culture. Although soilless culture cannot supplant usual agricultural methods generally, it can supplement them where mechanics and economics justify such methods.

Choice of fertilizer. The best fertilizers to use depend on the anticipated plant need and interactions with the soil, prior to or during cropping, which may determine availability for plant growth. Organics and many inorganics are useful where suitable. Need must be known, and the appropriate combinations of elements must be applied to favor soil conditions and plant growth. A single inorganic salt or combinations of such salts may accomplish this. The ratios of essential elements in a fertilizer should be considered on their atomic equivalent basis per unit weight, not as the oxides or otherwise, which is the way the plant utilizes them. Further, size of particle and nature of the chemicals should be considered;

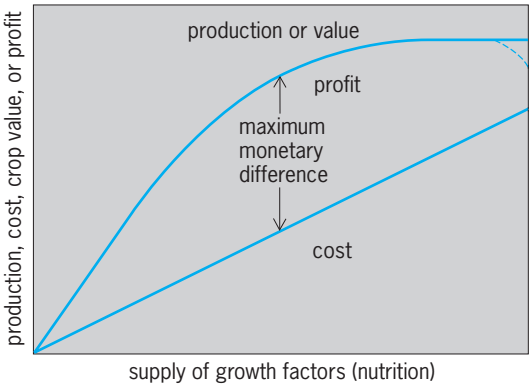

Fig. 4. Representation of the economics of crop production per unit area. Schematic shapes of curves (increases) may vary with conditions.

granules or larger particles and relatively insoluble compounds may dissolve slowly, which may be desirable or disadvantageous. Some applied ions may be fixed on soil colloids (such as K or P) or precipitated (such as Ca, Fe, or P) through chemical double decomposition. Some solubles (such as NO_3) may be leached by excessive amounts of water to subsoil or drainage systems and thus become ineffective or lost.

So-called sustained-release fertilizers are marketed; they are coated particles that release the elements slowly and protractedly. To lessen fixation of micronutrient cations by soil colloids, these metals are effectively supplied in chelated form, which involves weak chemical binding. In this form they may be sprayed upon foliage and thus effectively fertilize the plant. Natural materials may supply micronutrients as an appreciable impurity, though often inadequately.

Soil conditioning. Soil amendments or conditioners are often applied to improve structure for aeration, water penetration, and other soil conditions. Natural organic materials may be helpful, if not supplied excessively. Gypsum can favorably increase granulation and flocculation of soil colloids, being especially useful in nonsaline alkali soils. Some synthetics have been proposed to improve soil structure but thus far are not economically practical. Some have been suggested to decrease excessive evapotranspiration but, here too, further experimental trial is necessary for practical and economic feasibility.

To properly fertilize a crop, the agriculturist should know fairly well the nature and properties of soils, fertilizers, and plants, their many interactions, and the economics of fertilization (**Fig. 4**). *See* AGRICULTURAL MACHINERY; AGRICULTURAL SOIL AND CROP PRACTICES; FERTILIZER; PLANT GROWTH.

Theodore C. Broyer

Bibliography. J. Archer, *Crop Nutrition and Fertilizer Use*, 2d ed., 1988; California Fertilizer Association, *Western Fertilizer Handbook*, 8th ed., 1994; G. W. Cooke, *Fertilizing for Maximum Yield*, 3d ed., 1988; O. P. Engelstad (ed.), *Fertilizer Technology and Use*, 3d ed., 1985; S. L. Tisdale and W. L. Nelson, *Soil Fertility and Fertilizers*, 6th ed., 1998; K. K. Vyas and S. W. Saxena, *Biofertilizers*, 1990.

Fescue

A group of approximately 100 species of grass; more than 30 are represented in the United States. Tall fescue (*Festuca arundinacea*), a perennial cool-season plant introduced from Europe, occupies about 35×10^6 acres (15×10^6 hectares), primarily in the humid south-central region of the United States. It is popular because of its ease of establishment, vigor, wide range of adaptation, long grazing season, tolerance to abuse, sufferance of drought and poor soils, pest resistance, good seed production, and esthetic value when used for turf, ground cover, and conservation purposes. Its dominance of pastures and roadsides indicates that it fills an important ecological niche. It is used primarily as pasture and hay for beef cattle, with lesser use for dairy cows or replacement heifers, sheep, and horses. The leafy and vigorous plants can grow to 3–4 ft (0.9–1.2 m) if undisturbed; under grazing or clipping, they can form a dense sod when sufficient water and fertility are available.

Tall fescue toxicosis. Cattle and sheep that graze tall fescue can suffer from a condition known as tall fescue toxicosis. During warm weather, signs of this toxicosis are poor animal gains [less than 450 g/day (16 oz/day) for steers], intolerance to heat, excessive salivation, rough haircoat, slightly elevated body temperature, nervousness, lowered milk production, and abnormally low conception rate. In cool weather, the toxicosis can be expressed as fescue foot, a gangrenous condition of cattle feet. Bovine fat necrosis, characterized by fatty deposits in cattle abdomens, is associated with applications of several tons of chicken manure per hectare to tall fescue pastures. When pregnant mares graze tall fescue, particularly during the last trimester of gestation, their placentas become thickened, which sometimes causes delivery of dead foals, little or no milk production even when the foal survives birth, and sometimes death of the mare.

An endophytic symbiotic fungus, *Acremonium coenophialum*, growing between the cells of tall fescue tissues, is responsible for the toxicosis. Over 80% of all tall fescue plants contain the endophyte; that is, they are E^+.

Cultivars. There are over two dozen cultivars of tall fescue, the most widespread being Kentucky 31 in variants that either contain the fungus endophyte or are endophyte free. The use of endophyte-free cultivars to improve animal performance is important. The turf industry uses cultivars that contain the endophyte for their greater tolerance to environmental stress; special storage is used to guarantee viability of both endophyte and seed.

Other important fescues include meadow fescue (*F. elatior*), used to a limited extent in pasture; red fescue (*F. rubra*) and Chewings fescue (*F. rubra* var. *commutata*), a fine-leaved, shade-tolerant species with several cultivars used for turf; and Idaho fescue (*F. idahoensis*) and sheep fescue (*F. ovina*), fine- or bristle-leaved species adapted to drier regions and poorer soils. *See* CYPERALES; GRASS CROPS.

Henry A. Fribourg

Bibliography. C. W. Bacon and M. R. Siegel, Endophyte parasitism of tall fescue, *J. Prod. Agr.*, 1:45–55, 1988; D. M. Ball, J. F. Pedersen, and G. D. Lacefield, The tall-fescue endophyte, *Amer. Sci.*, 81(4):370–379, 1993; H. A. Fribourg, C. S. Hoveland, and K. D. Gwinn, Tall fescue and the fungal endophyte: A review of current knowledge, *Tenn. Farm Home Sci.*, 160:30–37, 1991.

Fetal alcohol spectrum disorder

A range of effects, including physical, mental, behavioral, and/or learning disabilities, that can occur in the offspring of women who consume alcoholic beverages during pregnancy. The term fetal alcohol spectrum disorder encompasses several conditions. Fetal alcohol syndrome is the most severe disorder, and is characterized by abnormal facial features and growth and central nervous system problems.

The frequency of adverse pregnancy outcome of chronic alcoholic women is around 40%. There is an eightfold increase in perinatal mortality, and a smaller percentage has the dysmorphic features noted above.

Characteristics. Children who are born to women who consume alcohol during pregnancy may exhibit pre- and postnatal growth deficiency. At age 1, they often continue to be small and fail to thrive. Similarly, some children score below average on standard intelligence and motor function tests. Physical abnormalities include typical facial changes, such as small eye openings and a lack of distinction (or a groove) in the area from below the nose to the upper lip. The presence of these characteristic abnormal facial features is sufficient to alert physicians to possible unrecognized alcoholism in the mother and to enable the diagnosis of fetal alcohol syndrome during the neonatal period, although the typical facial changes are more fully recognizable after age 3.

The severity of the dysmorphic features appears to correlate with the extent of cognitive dysfunction. With age, more manifestations of complex and pervasive neurobehavioral deficits become apparent. These include cognitive dysfunction, attention deficit and hyperactivity, learning disability, poor adaptive behavior, and conduct disorder. The spectrum of deficits can vary widely. Fetal alcohol spectrum disorder encompasses all these manifestations. *See* ATTENTION DEFICIT HYPERACTIVITY DISORDER.

Level of alcohol consumption. Although full-blown fetal alcohol syndrome is seen in some children born to severely alcoholic mothers, lesser degrees of change may be found in some offspring of women who consume smaller amounts of alcoholic beverages. It has yet to be determined whether there is a threshold dose of alcohol consumption during pregnancy below which no fetal injury results. However, fetal alcohol spectrum disorder has been described only among offspring of problem-drinking mothers.

It is not known whether there is a safe level of alcohol consumption during pregnancy. Even at a level of two drinks per day, infants average about 2–5.5 oz

(60–160 g) below the expected birth weight for gestational age. Whether small amounts of alcoholic beverages produce injury or death of the embryo during these early stages of development remains to be established. Women should be advised to refrain from any alcoholic beverages during gestation.

Pathogenesis. The mechanisms by which ethanol produces the changes associated with fetal alcohol spectrum disorder in the developing fetus are not known. One potential mechanism implicates the major ethanol metabolite, acetaldehyde, which may directly affect the proliferation of fetal cells. Alcoholic women tend to have higher levels of circulating acetaldehyde, probably due to their development of alcohol-related liver disease. However, there is other evidence to suggest that ethanol itself can produce the changes noted. New evidence points to oxidative damage. Antioxidants can ameliorate experimentally induced brain injuries caused by ethanol administration. *See* ANTIOXIDANT.

Whether the concurrent use of alcohol and tobacco or the common nutritional deficiencies associated with alcoholism have a potentiating role remains to be determined. The impairment of growth and development and cognitive functions are possible following the maternal use of a number of central nervous system–active agents. Thus, the syndrome described may be augmented by a variety of agents or may be the result of a common mechanism induced by a variety of agents. *See* ALCOHOLISM; BEHAVIOR TOXICOLOGY.

Prevention. Medical advice regarding the consumption of alcoholic beverages during pregnancy involves several principles: heavy drinking is clearly deleterious to the fetus; abstinence from alcoholic beverages immediately preceding and during pregnancy is the only assured safe course of action presently known; and therapeutic attention should also be directed to the problems leading to the alcoholism so that the psychological environment of the offspring may be improved.

Testing for the presence of fatty acid ethyl esters (ethanol metabolites) in meconium (first neonatal stool) is becoming a clinical tool to identify babies exposed in utero to extensive maternal drinking and thus identify infants at risk of developing fetal alcohol spectrum disorder.

Gideon Koren; Irena Nulman; N. Karle Mottet

Bibliography. E. L. Abel, *Fetal Alcohol Syndrome*, 1990; D. Chan et al., Fetal exposure to alcohol as evidenced by fatty acid ethyl esters in meconium in the absence of maternal drinking history in pregnancy, *Ther. Drug Monit.*, 26:474–481, 2004; S. Clarren et al., Brain malformations related to prenatal exposure to alcohol, *Pediatrics*, 92:64–67, 1978; K. Jones and D. W. Smith, Recognition of the fetal alcohol syndrome in early infancy, *Lancet*, 2:999–1001, 1973; G. Koren and I. Nulman, *Motherisk Handbook for the Diagnosis of Fetal Alcohol Syndrome*, Motherisk Program, Toronto, 2002; G. Koren et al., Fetal alcohol spectrum disorder, *CMAJ*, 169:1181–1118, 2003; A. P. Streissguth et al., Teratogenic effects of alcohol in humans and laboratory animals, *Science*, 209:353–361, 1980.

Fetal membrane

Any of four membranous structures surrounding the embryo during its developmental period. Because these membranes are external to the embryo proper, they are called extraembryonic membranes. Throughout embryonic development, every embryo must contend with a few basic needs. All developing animals need some source of nutrition, a way to remove wastes, some form of respiration, and protection from the outside environment. Fetal membranes collectively serve these functions.

Amniote egg. Invertebrates and lower vertebrates usually undergo rapid development in an aqueous environment until the free-swimming, feeding larval stage. Their eggs are characterized by membranes that are produced by the maternal follicle cells or by secretions from specialized cells of the maternal reproductive tract. As animals made the transition to life completely on land, they adapted a mechanism to enclose the developing embryo in a self-maintained aqueous state. This led to the formation of the amniote egg. Reptiles, birds, and mammals (collectively known as amniotes) undergo a longer developmental period than fish and amphibian embryos and thus require greater protection from the external environment. Whether they are encased in a shell or the uterus, amniotic embryos form four extraembryonic membranes—yolk sac, allantois, chorion, and amnion—to assist their development in the terrestrial environment. These fetal membranes are formed by the tissues of the developing embryo itself and, therefore, differ from membranes or shells that are of maternal origin. *See* AMNIOTA.

Development and structure. In early development, there is no initial delineation between embryonic and extraembryonic structures. The fetal membranes are made up of the same sheets of cells as the embryo: the ectoderm, the endoderm, and the lateral plate mesoderm. As development progresses and the body folds form by differential proliferation rates, the body becomes separate from the yolk and thus embryonic and extraembryonic structures become defined. The extraembryonic ectoderm and mesoderm combined (called the somatopleure) form the amnion and the chorion. The extraembryonic endoderm and mesoderm combined (called the splanchnopleure) form the yolk sac and allantois. The space between these two mesodermal layers is continuous with the body cavity (coelom) of the embryo and is called the extraembryonic coelom. *See* COELOM; GERM LAYERS.

Yolk sac. The yolk sac is the first of the four extraembryonic membranes to form. It is of endodermal origin and is connected to the midgut by the yolk duct. As the embryo develops, the nutrients within the yolk sac are used up and it decreases in size, ultimately becoming incorporated into the gut of the embryo.

Chorion and amnion. The amnion and chorion initially form over the dorsal region of the embryo, but then extend ventrally and fuse to become the external layers. To do this, the ectoderm plus the somatic mesoderm form a double-layered fold (the anterior amniotic fold) anterior to the embryonic axis

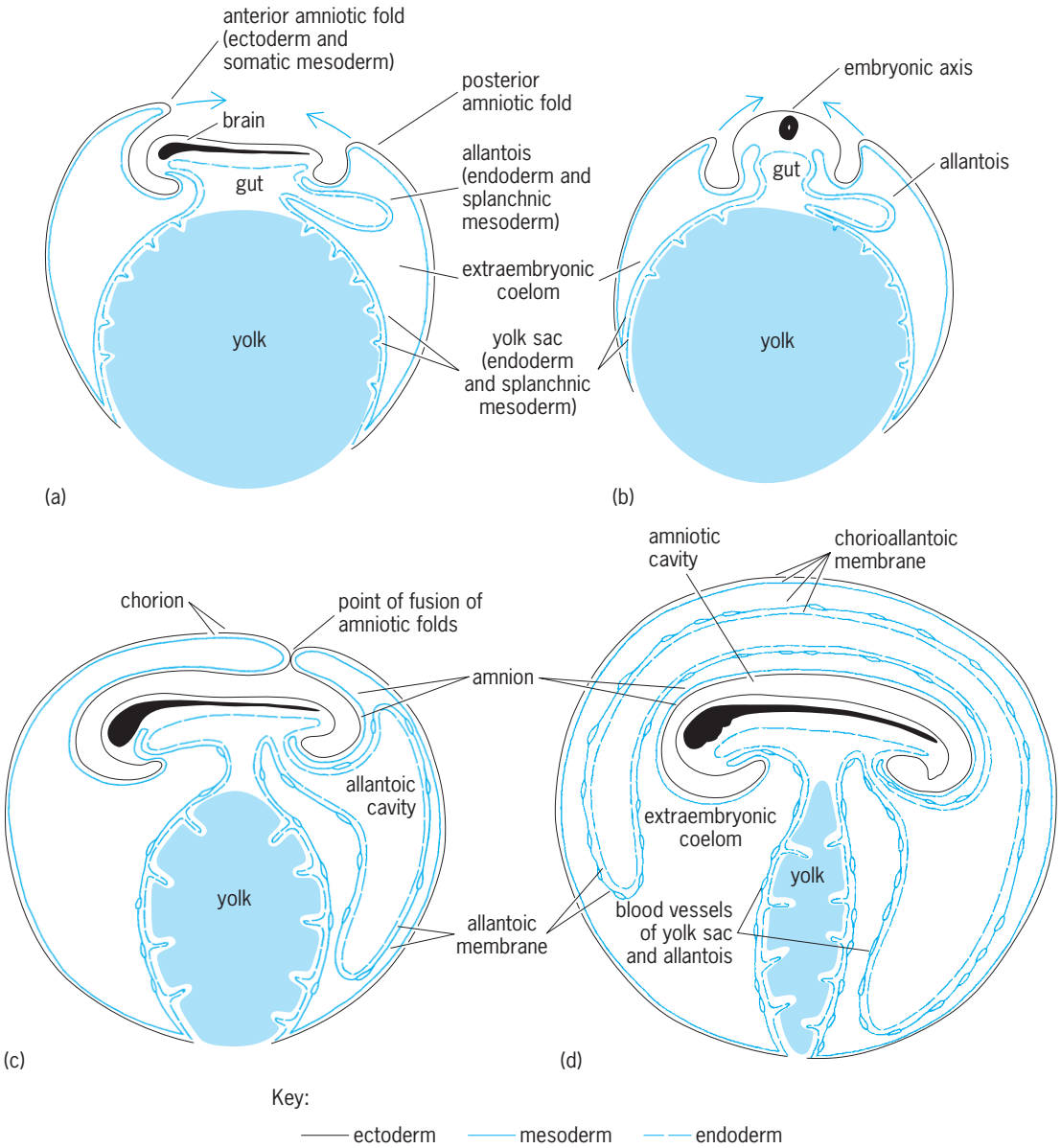

Relationships of fetal membranes to each other and to the developing embryonic body. (*a*) Midsagittal section during amniotic folding. (*b*) Cross section showing lateral amniotic folds. (*c*) Midsagittal section just before fusion of amniotic folds. (*d*) Midsagittal section showing fully developed fetal membranes.

(**illus.** *a*). A similar fold (the posterior amniotic fold) appears posterior to the axis a short time later. As these folds grow toward each other over the dorsal side of the embryo, lateral folds also appear and approach each other (illus. *b*). The result is that the exposed surface of the embryo gradually disappears beneath the advancing margins of the folds (illus. *c*), which soon meet and fuse over the posterodorsal region of the embryonic axis. Following fusion of the two sides, the upper double-layered membrane (now the chorion) separates from the lower (now the amnion), and the extraembryonic coelom becomes continuous over the dorsal surface of the embryo. The chorion becomes the outermost, and the amnion the innermost, membrane surrounding the developing embryo (illus. *d*).

Allantois. In a similar way, the allantois begins as a blind sac growing out from the ventral wall of the hindgut (and is, therefore, lined with endoderm), pushing ahead a covering of splanchnic mesoderm (illus. *a*, *b*, and *c*). As the allantois increases in size, it expands into the extraembryonic coelom, and its splanchnic mesoderm becomes closely associated with, if not fused with, the somatic mesoderm of the chorion. The two membranes together are called the chorioallantoic membrane (illus. *d*).

Function. Each of the four fetal membranes serves a different primary purpose: the yolk sac contains the source of nutrients, the allantois is responsible for waste management, the chorion allows for gas exchange, and the amnion produces and maintains the amnionic fluid, which protects the embryo from drying out.

Yolk sac. The primary function of the yolk sac is for the external storage and supply of nutrients. Instead of the nutrients being acquired directly through the yolk duct, nutrients are obtained through the blood vessels found in the mesoderm of

the splanchnopleure around the yolk sac. In placental mammals, the yolk sac is not necessary for nutrient storage, as the mother supplies what the embryo needs through the placenta; however, a vestigial yolk sac remains. This is possibly due to the secondary functions it performs: the yolk sac endoderm is the source of primordial germ cells, as well as the source of original red and white blood cells, which originate from mesodermal cells lining the yolk sac endoderm. This endoderm also contributes in development as an inducer of specific organs and organ systems.

Allantois. The primary function of the endodermally lined allantois is the storage of urinary waste. This is most important in reptiles, birds, and nonplacental mammals undergoing external development, as it limits the embryo's exposure to toxins and osmotic stress. In birds and reptiles, this waste is primarily in the form of uric acid, which is relatively insoluble in water and must be stored away from the embryo in the allantois. The close association of this membrane (in conjuction with the chorion) with the shell also allows for some level of gas exchange in these organisms. Among mammals, there is great species variation, and in each species the level of allantois involvement is inversely related to the amount of vascular exchange occurring in the placenta. Thus, the allantois also has some function in gas exchange.

Chorion. The chorion, the outermost of the four membranes, is the site of exchange with the environment. In bird and reptile eggs the chorion's principal function is gas exchange, whereas in mammals the chorion develops into the placenta, and its functions include respiration, nutrition, excretion, filtration, and endocrine and immune responses. Depending on the species of placental mammal, the placenta may be easily separated from the uterus (as in pigs) or may be deeply integrated with the maternal components of this tissue (as in humans). *See* PLACENTATION.

Amnion. The amnion is responsible for the production and maintenance of the amniotic fluids. This aqueous environment is essential to prevent the embryo from drying out, to allow for proper folding of the developing embryo and membranes, and to allow for exchange of gases and waste. Smooth muscle fibers in the amnion spontaneously contract and gently rock the embryo before it develops the capacity for movement.

Fate. Although these four fetal membranes are essential to development, very little of them remains once the animal is born or hatched. The yolk sac is incorporated into the midgut, and the proximal portion of the allantois becomes the urinary bladder. The amnion and chorion separate entirely from the animal, and the placenta (including part of the maternal uterine mucosa, in some cases) is delivered separately as afterbirth. *See* ALLANTOIS; AMNION; CHORION; YOLK SAC. Cheryl Schaeberle

Bibliography. B. I. Balinsky, *An Introduction to Embryology*, 5th ed., 1981; F. C. Battaglia and G. Meschia, *An Introduction to Fetal Physiology*, 1986; B. M. Carlson, Extraembryonic membranes and placenta, in *Patten's Foundations in Embryology*, 5th ed., Chap. 7, pp. 250–289, McGraw-Hill, New York, 1988; J. C. Cross, Formation of the placenta and extraembryonic membranes, *Ann. N.Y. Acad. Sci.*, 857:23–31, 1998; S. F. Gilbert, *Developmental Biology*, 7th ed., Sinauer Associates, Sunderland, MA, 2003; N. K. Wessels, *Tissue Interactions and Development*, 1977.

Fever

An elevation in the central body temperature of warm-blooded animals caused by abnormal functioning of the thermoregulatory mechanisms. Fever accompanies a wide variety of disease states, both infectious and noninfectious, and in the great majority of instances is due to an abnormality in the regulation of body temperature by the central nervous system. *See* THERMOREGULATION.

Endogenous pyrogens. Experimental studies on the cause of fever suggest that leukocytes (circulating granulocytes and monocytes) of the host, as well as the fixed macrophages of the reticuloendothelial system in the liver and spleen, can be activated by various stimuli to produce a fever-inducing substance—endogenous pyrogen. Activators of microbial origin include the lipopolysaccharide endotoxins of gram-negative bacteria, cell walls and soluble antigens of other bacteria, viruses, and fungi. Endogenous pyrogen has been characterized as a protein of relatively low molecular weight (13,000) with an essential lipid moiety. Pyrogens from different tissues and from rabbits and humans appear to have similar characteristics. Circulating endogenous pyrogen, liberated by activated cells, is believed to cross the blood-brain barrier and act upon specialized neurons, presumably located in the anterior hypothalamus, that initiate heat conservation and increase heat production.

Mechanisms. The mechanisms responsible for elevating body temperature include: reduction in heat loss by constriction of peripheral vessels whose tone is under control of the sympathetic nervous system; inhibition of panting and sweating, the latter by way of the cholinergic nerves; and increased heat production by means of shivering in voluntary muscles innervated by somatic motor nerves.

Pharmacologic pyrogens. Various naturally occurring pharmacologic agents, including serotonin (5-hydroxytryptamine) and catecholamines, are normally present in high concentration in areas of the brain that control body temperature and can cause fever when injected under certain experimental circumstances. The role of these agents in most experimentally produced or clinical fevers is obscure. Certain naturally occurring steroids are also pyrogenic and are known to activate human leukocytes in the test tube.

Effects. There is no clear evidence that elevated body temperature evoked by most infections is directly injurious to microbial invaders. As fever regularly accompanies inflammation, however, increased body temperature may well accelerate certain biochemical reactions of use to the host in

combating infection. *See* HOMEOSTASIS; INFLAMMATION. Elisha Atkins

Bibliography. H. Laudien, *Temperature Regulations in Animals and Man*, 1986; J. B. Walter, *Pathology of Human Disease*, 1989; B. W. Zweifach et al. (eds.), *The Inflammatory Process*, 2d ed., 1974.

Feynman diagram

A pictorial representation of elementary particles and their interactions. The diagrams show the paths of particles in space and time as lines, and the interactions between the particles as points where the lines meet.

These diagrams were introduced by R. P. Feynman in 1949 in the context of quantum electrodynamics, the quantum field theory of electromagnetic interactions of charged particles. The **illustration** shows Feynman diagrams for electron-electron scattering. In each diagram, the straight lines represent space-time trajectories of noninteracting electrons, and the wavy lines represent photons, particles that transmit the electromagnetic interaction. External lines at the bottom of each diagram represent incoming particles (before the interactions), and lines at the top, outgoing particles (after the interactions). Interactions between photons and electrons occur at the vertices where photon lines meet electron lines. *See* ELECTRON; PHOTON.

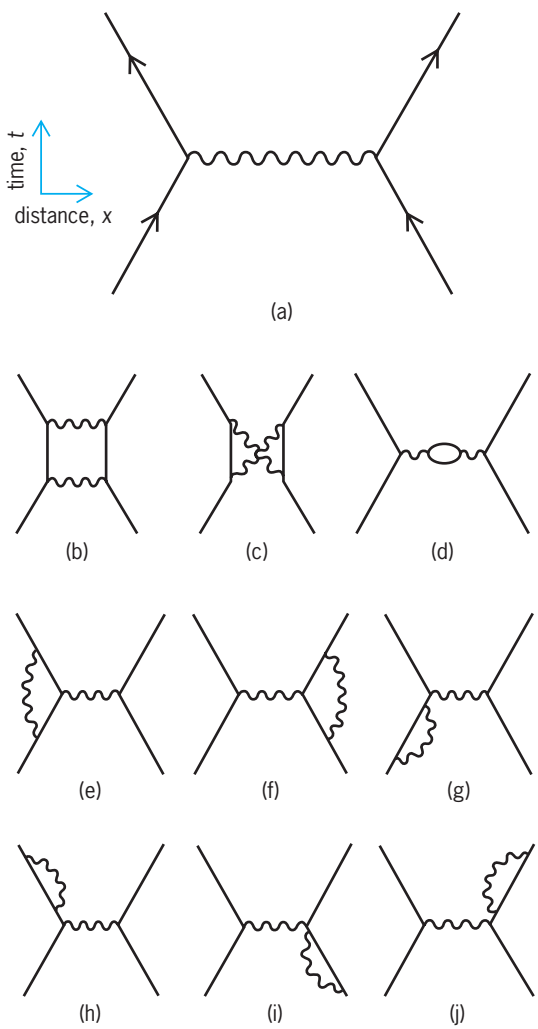

Feynman diagrams for electron-electron (Møller) scattering: (*a*) second-order diagram (two-vertices); (*b–j*) fourth-order diagrams.

Each Feynman diagram corresponds to the probability amplitude for the process depicted in the diagram. This amplitude can be calculated from the structure of the diagram according to specific rules. It is a product of wave functions for external particles, a propagation function for each internal electron or photon line, and a factor for each vertex that is proportional to the strength of the interaction, summed over all possible space-time locations of the vertices. The propagation function, corresponding to a line joining two vertices, is the amplitude for the probability that a particle starting at one vertex will arrive at the other. *See* PROPAGATOR (FIELD THEORY).

The set of all distinct Feynman diagrams with the same incoming and outgoing lines corresponds to the perturbation expansion of a matrix element of the scattering matrix in field theory. This correspondence can be used to formulate the rules for writing the amplitude associated with a particular diagram. The perturbation expansion and the associated Feynman diagrams are useful to the extent that the strength of the interaction is small, so that the lowest-order terms, or diagrams with the fewest vertices, give the main contribution to the matrix element. *See* PERTURBATION (QUANTUM MECHANICS); SCATTERING MATRIX.

Since their introduction in quantum electrodynamics, Feynman diagrams have been widely applied in other field theories. They are employed in studies of electroweak interactions, certain situations in quantum chromodynamics, acoustooptics, and in many-body theory in atomic, nuclear, plasma, and condensed matter physics. *See* ACOUSTOOPTICS; ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS; QUANTUM CHROMODYNAMICS; QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; WEAK NUCLEAR INTERACTIONS. Peter Mohr

Bibliography. J. Bjorken and S. Drell, *Relativistic Quantum Fields*, 1965; J. Bjorken and S. Drell, *Relativistic Quantum Mechanics*, 1964; C. Itzykson and J. Zuber, *Quantum Field Theory*, 1980; M. Kaku, *Quantum Field Theory: A Modern Introduction*, 1993; F. Mandl and G. Shaw, *Quantum Field Theory*, paper 1994; M. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, 1995; P. Ramond, *Field Theory: A Modern Primer*, 2d rev. ed., 1990.

Feynman integral

A technique, also called the sum over histories, which is basic to understanding and analyzing the dynamics of quantum systems. It is named after fundamental work of R. Feynman. The crucial formula

gives the quantum probability density for transition from a point q_0 to a point q_1 in time t as expression (1), where S(path) is the classical mechanical action

$$\int \exp\left[iS(\text{path})/\hbar\right]\, d\left(\text{path}\right) \qquad (1)$$

of a trial path, and $\hbar$ is the rationalized Planck's constant. The integral is a formal one over the infinite-dimensional space of all paths which go from q_0 to q_1 in time t. Feynman defines it by a limiting procedure using approximation by piecewise linear paths. *See* ACTION.

Feynman integral ideas are especially important in quantum field theory, where they not only are a useful device in analyzing perturbation series but are also one of the few nonperturbative tools available.

An especially attractive element of the Feynman integral formulation of quantum dynamics is the classical limit, $\hbar \to 0$. Formal application of the method of stationary phase to expression (1) says that the significant paths for small $\hbar$ will be the paths of stationary action. One thereby recovers classical mechanics in the hamiltonian stationary action formulation. *See* LEAST-ACTION PRINCIPLE.

Feynman-Kac formula. Mark Kac realized that if one replaces the time by a formal purely imaginary time (that is, replaces the Schrödinger equation by a heat equation), then one can make precise mathematical sense out of Feynman's integral in terms of the stochastic process introduced by Norbert Wiener in his theory of brownian motion. The resulting formula for the semigroup e^{-tH}, with $H = -{}^1\!/_2\Delta + V$ the quantum hamiltonian (Δ is the laplacian operator and V is the potential energy), is given by Eq. (2).

$$(e^{-tH}f)(x) = E_x\left[\exp\left(-\int_0^t V[b(s)]\,ds\right)f[b(t)]\right] \qquad (2)$$

Here E_x is the expectation with respect to Wiener's process, $b(s)$, starting at the point x, and f is an arbitrary bounded measurable function to which e^{-tH} is applied. Equation (2) is the Feynman-Kac formula. It makes a variety of probabilistic ideas applicable to a rigorous mathematical analysis of quantum theory. *See* BROWNIAN MOVEMENT; STOCHASTIC PROCESS.

Euclidean quantum field theories. The passage to imaginary time initiated by Kac is especially useful in relativistic quantum field theory. Since the Lorentz covariance of relativity is replaced by euclidean group covariance under this replacement, the new theory is often called euclidean field theory. The corresponding Feynman integral is often called a euclidean path integral and has become the standard formulation of Feynman integrals in most modern discussions of quantum field theory. The analogs of stationary action solutions are called instantons. *See* INSTANTON.

Statistical mechanics analogy. The appearance of the exponential in Eq. (2) suggests an analogy between quantum dynamics and classical statistical mechanics. The analog of the total space-time dimension in a euclidean quantum field theory is then the dimension of space in statistical mechanics. Just as the phenomenon of phase transitions occurs in statistical mechanics, the analogous phenomena of spontaneously broken symmetries and dynamical instability occur in quantum field theory. These ideas are basic to all modern theories of elementary particle physics. They are usually analyzed within a Feynman integral framework. *See* ELEMENTARY PARTICLE; NONRELATIVISTIC QUANTUM THEORY; PHASE TRANSITIONS; QUANTUM FIELD THEORY; STATISTICAL MECHANICS; SYMMETRY LAWS (PHYSICS). Barry Simon

Bibliography. R. Cenni et al., *Feynman Integrals in Theoretical, Nuclear, and Statistical Physics*, 1989; R. P. Feynman, Space-time approach to nonrelativistic quantum mechanics, *Rev. Mod. Phys.*, 20:367–387, 1948; R. P. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals*, 1965; M. Kac, On some connections between probability theory and differential equations, *Proceedings of the 2d Berkeley Symposium on Mathematical Statistics Problems*, pp. 189–215, 1950; B. Simon, *Functional Integration and Quantum Physics*, 1979.

Fiber bundle

A decomposition of a space E into a family of identical subspaces (fibers). The space of subspaces is called the base space. Fiber bundles arise naturally in many physical situations. The theory of fiber bundles has been applied to gauge theory in physics, and there is a lively interaction centered on ideas related to fiber bundles and gauge theories in which significant results have been contributed in both areas. The development of the mathematical theory of fiber bundles was begun in the 1930s and has numerous applications within mathematics.

Definition. A fiber bundle consists of three spaces, a fiber F, a total space E, and a base B, together with a map $\pi: E \to B$ from the total space to the base. These sets and the map π are supposed to satisfy the following requirements:

1. The spaces F, E, and B are topological spaces: A topology is specified by giving meaning to the phrase: two points are close to one another. It is required that each of the spaces F, E, B has such a notion of closeness. *See* TOPOLOGY.

2. The map π is required to preserve the notion of closeness; if points x and y are sufficiently close to one another in E, then the points $\pi(x)$ and $\pi(y)$ are close in B; that is, π is a continuous map.

3. Each fiber can be identified with F: The map $\pi: E \to B$ decomposes E into subsets; two points are in the same subset (or fiber) if they have the same image in B. The set of points in E that map to a given point b in B is called the fiber over b. It is required that each fiber be topologically equivalent (that is, homeomorphic) to F.

4. The total space E has a local product structure: For a set U of points in B that are sufficiently close together, the points $\pi^{-1}(U)$ in E that map to U can be identified with the set of pairs (b, x) with b an

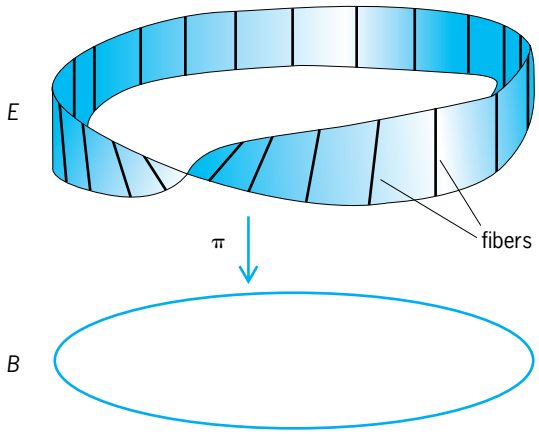

Möbius strip, the simplest nontrivial example of a fiber bundle.

element of U and x an element of F; that is, $\pi^{-1}(U)$ is homeomorphic to $U \times F$.

Examples. The trivial examples of fiber bundles are those where E is a product space $E = B \times F$, the set of pairs (b, x), where b is a member of B, and x is a member of F.

Möbius strip. The simplest nontrivial example of a fiber bundle is the Möbius strip (see **illus.**). It is constructed by twisting one end of a piece of paper and then gluing the two ends together. The fibers are line segments, and the space of line segments is a circle. The map π takes all the points in a fiber into the same point on the circle.

Tangent bundles. Tangent bundles are among the earliest examples of fiber bundles. For spaces M (called manifolds) which are locally like a euclidean space, the space $T(M)$ consisting of pairs (x, v) with x a point in M and v a vector tangent at x to a path in M is called the tangent bundle to M; the fibers form a vector space, and the base space is M. Fiber bundles in which the fibers are vector spaces are called vector bundles. *See* MANIFOLD (MATHEMATICS).

Applications to physics. Laws of physics are often invariant under the action of an appropriate group. Fiber bundles arise in physics in situations where such a group appears as an internal or local symmetry group; the local symmetry group is the fiber. *See* GROUP THEORY; SYMMETRY LAWS (PHYSICS).

For example, by Maxwell's equations, a vector potential A for a magnetic monopole of charge g located at the origin in three-dimensional euclidean space, $\boldsymbol{R}^3$, would have a curl whose magnitude is equal to g/r^2, where r is the distance from the origin to the point p at which the curl of A is evaluated, and whose direction is that of a vector directed from the origin to p. There is no vector potential that both satisfies this condition and is defined on all of $\boldsymbol{R}^3 - (0, 0, 0)$ [three-dimensional space with the origin removed]. However, potentials that satisfy this condition can be defined on sufficiently small open sets in $\boldsymbol{R}^3 - (0, 0, 0)$. For two such open sets, a gauge transformation, Ψ is a mapping from the intersection of U and $\tilde{U}$ to $U(1)$, the group of rotations of a circle, here represented by the complex numbers of modulus one. The vector potentials A, defined on U, and $\tilde{A}$, defined on $\tilde{U}$, are said to be related by the gauge transformation Ψ if, at each point in the intersection of U and $\tilde{U}$, A is equal to the sum of $\tilde{A}$ and a term that is equal to $i\hbar c/e$ times the product of the gradient of Ψ with Ψ^{-1}. Here, $\hbar$ is Planck's constant divided by 2π, c is the speed of light, and e is the electric charge of the monopole. Potentials that are related by gauge transformations and defined on sets whose union covers all of $\boldsymbol{R}^3 - (0, 0, 0)$ determine a fiber bundle with fiber $U(1)$ and base $\boldsymbol{R}^3 - (0, 0, 0)$. The condition that the potentials are related by gauge transformations is equivalent to Dirac's quantization condition relating the electric and magnetic charge; namely, that $2ge/(\hbar c)$ is an integer. *See* CALCULUS OF VECTORS; COMPLEX NUMBERS AND COMPLEX VARIABLES; GAUGE THEORY; MAGNETIC MONOPOLES; MAXWELL'S EQUATIONS.

Applications within mathematics. There are numerous, fundamental applications of fiber bundles within mathematics. Many of these applications are obtained through associating numerical invariants to vector bundles. The geometry of manifolds is then studied in terms of invariants of the tangent bundle of the manifold.

One such application involves the concept of cobordism. Two manifolds that are disjoint (have no elements in common) and whose union can be realized as the boundary of another manifold are called cobordant. A manifold that can be realized as the boundary of manifold is called cobordant to zero. For example, a circle is the boundary of disk and hence is cobordant to zero; likewise, a two-dimensional sphere is the boundary of a solid ball. However, the two-dimensional space whose points are lines through the origin in euclidean three-space is not cobordant to zero.

A set of invariants, called Stiefel-Whitney numbers, of vector bundles detect whether a manifold is cobordant to zero, and whether two manifolds are cobordant: Specifically, the condition that a manifold be cobordant to zero is equivalent to the condition that all Stiefel-Whitney numbers of the tangent bundle of the manifold are zero; and the condition that two manifolds are cobordant is equivalent to the condition that the two manifolds have the same Stiefel-Whitney numbers.

Richard Porter

Bibliography. D. Husemoller, *Fiber Bundles*, 3d ed., 1993; K. B. Marathe and G. Martucci, *The Mathematical Foundations of Gauge Theories*, 1992; R. Porter, *Introduction to Fibre Bundles*, 1977; N. E. Steenrod, *The Topology of Fibre Bundles*, 1951.

Fiber crops

Many plants are grown as fiber crops because of their content or yield of fibrous material which is used for many commercial purposes and for home industry. The production and use of plant fibers date back many centuries and form the basis for one of the earliest industrial processes. Fibers may be extracted from various parts of different plants. *See* NATURAL FIBER; PLANT ANATOMY.

Long, multiple-celled fibers can be subdivided into hard, or leaf, fibers that traditionally are used for cordage, such as sisal for binder and baler twine and abaca or manila hemp for ropes; soft, or bast (stem), fibers that are used for textiles, for example, flax for linen, hemp for small twines and canvases, and jute and kenaf for industrial textiles such as burlap; and miscellaneous fibers that may come from the roots, such as "broom" root for brushes, or stems, as Spanish moss for upholstery, or fruits, as coir from coconut husks for cordage and floor coverings. *See* FRUIT; LEAF; ROOT (BOTANY); STEM.

There are many exceptions to these uses. For example, some sisal near the sources of production is used for bags, and jute is used for cordage. However, the large quantities that enter international trade are used as indicated above.

Short, one-celled fibers come from the seeds or seed pods of plants such as cotton and kapok. Cotton is the world's most widely grown and used textile fiber. Kapok, because of its low tensile strength, goes primarily into upholstery, pillows, and life preservers. Many plant fibers are being replaced wholly or in part by synthetic or manufactured fibers. Mixtures of natural and synthetic fibers often exhibit qualities superior to either used alone. *See* ABACA; CELLULOSE; COIR; COTTON; FLAX; HEMP; JUTE; KAPOK TREE; LINEN; MANUFACTURED FIBER; RAMIE; SISAL.

Elton G. Nelson

Fiber-optic circuit

The path of information travel, usually from one electrical system to another, in which light acts as the information carrier and is propagated by total internal reflection through a transparent optical waveguide. An electrooptic modulator and an optoelectric demodulator are required to convert the electrical signals into light and back again at the transmit and receive ends of the link, respectively.

Application. A fiber-optic circuit, or link, is used for data transmission when a shielded twisted pair of wires or a coaxial cable fails to meet the performance criteria of the system designer. Depending upon fiber type, the product of distance and bandwidth of a fiber can be tens to thousands of times greater than that for electrical transmission through wire. An optical communication fiber is a nearly perfect waveguide for light, meaning that little or no energy escapes through radiation. Thus, the data traveling in the fiber are secure from eavesdropping, as well as being harmless in or around equipment sensitive to electromagnetic interference. Telecommunication optical fibers have a very small diameter, 5–10 micrometers (0.0002–0.0005 in.), which allows telecommunication cables to be fabricated with a much higher packing density. Furthermore, the most common materials used to make the fibers, silica and plastic, are less dense than copper, making the cable lighter. Lastly, since the fiber is a dielectric, it can be used in volatile or sensitive environments that require electrical isolation. *See* COMMUNICATIONS CABLE; ELECTRICAL INTERFERENCE; OPTICAL COMMUNICATIONS; OPTICAL FIBERS; WAVEGUIDE.

Optical transmitters. In order to pass electrical information through a fiber, an electrooptic modulator, or transmitter, is used to convert electrical impulses into modulated light, which is then launched, or focused, by the transmitter into the fiber (**Fig. 1**). The transmitter generally consists of a silicon integrated circuit that converts input voltage levels from a personal computer, workstation, or telecommunications company mainframe computer into current pulses. These current pulses, in turn, drive a light-emitting diode (LED). The output light from the diode is then focused by one or more lenses into the fiber. Faster and more expensive circuits use laser diodes, which are semiconductor chips that emit coherent laser light at a very low cost. *See* INTEGRATED CIRCUITS; LASER; LIGHT-EMITTING DIODE.

Optical receivers. An optoelectric demodulator, or receiver, is used to convert the detected light back into electrical impulses at the other end of the fiber (Fig. 1). Light from the fiber is focused onto a reverse-biased *pn*-junction photodiode that generates an electron-hole pair for each photon impinging on or near its active area. The applied reverse bias creates an electric field across the *pn* junction, which causes the generated electron-hole pairs to be swept apart and give rise to photocurrent before they can recombine. Systems requiring higher speed or sensitivity use more expensive PIN (positive-intrinsic-negative) diodes. Another circuit, usually a silicon integrated circuit, amplifies this electron-hole current and converts it into voltage levels suitable for interfacing with the computer at the receiving end. *See* PHOTODIODE; PHOTOVOLTAIC CELL.

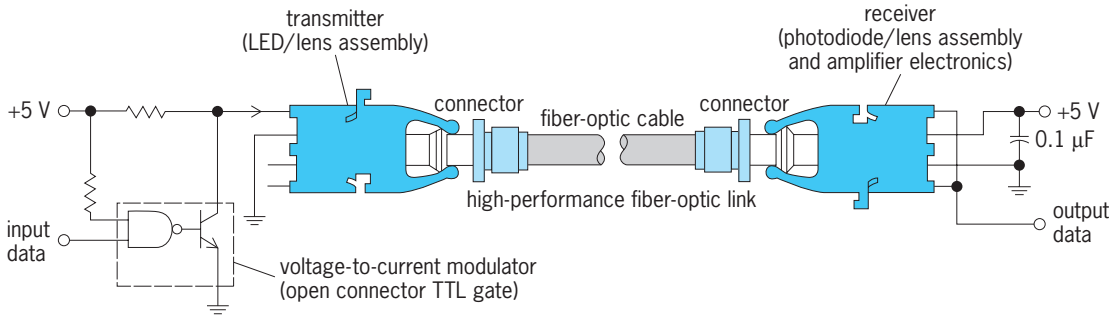

Fig. 1. Typical interface circuit for fiber-optic link that transmits data at 5 megabits per second. (*Hewlett-Packard*)

Link budget. The parameters involved in the design of a fiber-optic link include the optical output power, spectrum, fiber loss, dispersion, extinction ratio, bandwidth, rise and fall times, and near- and far-field radiation patterns of the transmitter; and the sensitivity, bandwidth, signal-to-noise ratio, bit error rate, eye pattern, and intersymbol interference of the receiver. Link design involves trade-offs with these parameters and the costs to obtain the desired link budget. The link budget is the most important parameter in the design of the system. It is defined as the power detected at the receiver in excess of that required to maintain a given signal-to-noise ratio while the link is operating at the maximum specified frequency; it is expressed in decibels (dB). The link budget makes it possible to calculate how far the link will carry signals without a repeater in attenuation-limited systems, or how many connectors and splices can be used at a given distance in dispersion-limited links. *See* SIGNAL-TO-NOISE RATIO.

Links. The simplest and least expensive link uses plastic fiber and operates with visible light having a typical peak wavelength of about 665 nanometers. The transmitter is usually a transistor-transistor logic (TTL) line driver which shunts current from a resistor and an inexpensive aluminum gallium arsenide (AlGaAs) light-emitting diode, while the receiver uses a silicon photodiode integrated with the receiver electronics. Links of this type generally transmit data at 5 megabits per second and are limited to several tens of meters because of the high optical attenuation of the plastic cable, typically 0.25 dB/m or 0.075 dB/ft (**Fig. 2**). *See* LOGIC CIRCUITS.

To transmit data over longer distances or at higher rates, silica fiber, 820-nm light-emitting diodes, and silicon PIN detectors are used. Silica fiber has a typical attenuation of 2.8 dB/km (4.5 dB/mi) at 820 nm. Silica fiber also has a much higher bandwidth than plastic fiber, allowing data transmission at hundreds of megabits per second. Optical dispersion limits the 820-nm light-emitting diode links to about 100 megabits per second at 1 km (0.6 mi; Fig. 2).

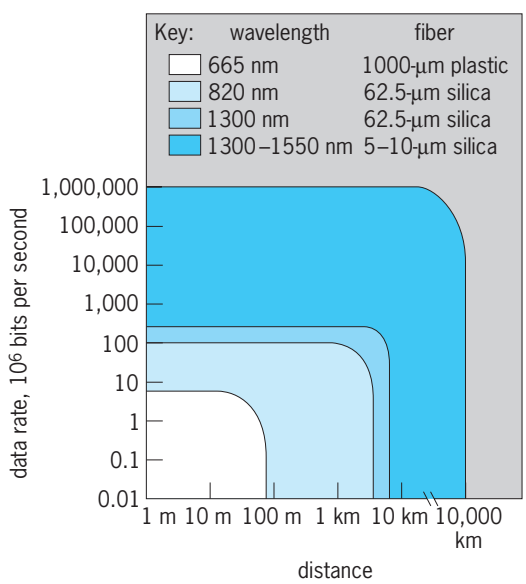

Fig. 2. Performance of four types of fiber-optic links. The limits shown for wavelengths of 1300–1550 nm over 5–10-μm silica fiber are the expected limits for 2010 with wavelength division multiplexing (WDM).

Optical dispersion is near zero with silica fibers at 1300 nm. Light-emitting diodes and PIN photodiodes made of indium phosphide and indium gallium arsenide phosphide (InP/InGaAsP) operate at 1300 nm, have longer operating lifetimes than the AlGaAs light-emitting diodes, and are capable of data rates of over 200 megabits per second at 2 km (Fig. 2). The transmitter and receiver electronics as well as packaging parasitics are much more complicated and interrelated at these data rates. The 820-nm laser diodes can be used for longer distances at higher data rates, but their reliability is much less than that of both 820- and 1300-nm light-emitting diodes. Typical systems such as those described above operate over several kilometers at data rates of several hundred megabits per second. By replacing light-emitting diodes with vertical cavity surface emitting lasers (VCSELs), these local area network links are capable of operating in the gigabit-per-second region. *See* LOCAL-AREA NETWORKS.

Commercial telecommunication links operate at the rate of low tens of gigabits per second. They use single-mode fiber with laser transmitters operating between 1300 and 1580 nm. Trade-offs between modal and chromatic dispersion versus cost, wavelength, fiber diameter, data rate, and distances from tens to hundreds of kilometers without repeaters are made, depending on what is most important for the desired link (Fig. 2). Most telecommunication transmitters use 1530–1580-nm edge emitting semiconductor lasers to be compatible with optical amplifiers.

Optical amplifiers. Despite the tremendous information-carrying capacity of optical fibers, the performance of most communication systems is limited by electronics speed in signal regeneration. Erbium-doped fiber amplifiers (EDFAs) overcome the speed and cost limitations of electronically "repeatered" communication systems and have the advantages of temperature stability, quantum-limited noise figure, polarization insensitivity, and immunity to interchannel crosstalk. Erbium-doped fiber amplifiers expand on the physics of laser diodes allowing amplification of light. For each photon that goes into an erbium-doped fiber amplifier, many coherent photons come out, thus providing amplification with minimal signal degradation. Erbium-doped fiber amplifiers provide a more efficient and cost-effective technology for both transoceanic and terrestrial communication links where the demands for bandwidth continue to grow at an exponential rate as telecommunication markets grow and the Internet gains popularity. *See* INTERNET.

Optical routers. With the surge of networking in the 1990s, optical routers are an obvious extension. Networks allow multiple computers to connect all at once. A router takes data from one source computer link and directs it to the proper data link going

toward the destination computer on a network. Typical routers are packed with fiber-optic transceivers where input optical data are converted to electricity, routed to the proper transmitter, and then converted back to light and sent to the next router or destination computer. This optical-to-electrical-to-optical conversion is both relatively expensive and slow. Advanced routers now are doing this without the conversion and are replacing the electronics with micromachined mirrors that steer the light. These routers direct data traffic from one segment of a network to another at up to 10 terabits, or 10^{13} bits, of data per second. These advanced routers can direct the information equivalent of whole libraries in seconds.

Wavelength-division multiplexing. Wavelength-division multiplexing (WDM) is another means of getting more data though a fiber, thus delaying or eliminating the need to install more strands of fiber. Wavelength-division multiplexing is accomplished by encoding data in different wavelengths of light and multiplexing these in a single-mode fiber. These separate single-mode fibers are then fused together and spliced to a single fiber that is usually already installed. At the other end of the fiber, the light is similarly demultiplexed and converted to electrical signals with wavelength-tuned receivers. Usually, the detectors are identical, and narrow Bragg grating filters are used to select which wavelength will hit each detector. *See* MULTIPLEXING AND MULTIPLE ACCESS.

Kirk W. Lindstrom

Bibliography. P. C. Becker et al., *Erbium-Doped Fiber Amplifiers: Fundamentals and Technology*, 1999; D. Derickson (ed.), *Fiber Optic Test and Measurement*, 1997; I. P. Kaminow and T. L. Koch (eds.), *Optical Fiber Telecommunications III and IIIb*, 1997; H. Kressel (ed.), *Semiconductor Devices for Optical Communication*, 2d ed., updated 1990; J. C. Palais, *Fiber Optic Communication*, 4th ed., 1998.

Fiber-optic sensor

A sensor that uses thin optical fibers to carry light to and from a location to be probed. In performing the sensing, light can be lost from the fibers or modified in velocity by the action of the phenomena on the fiber. Fiber-optic sensors are ideal for probing in remote or hostile locations, where miniature sensors are required such as in the body, or where extreme sensitivity is required. Two classes of fiber sensors have evolved: intensity sensors, in which the amplitude of light in the fiber is changed during sensing, and interferometric sensors, in which the velocity of light or its phase is modified during sensing. The latter class has proved to be extremely sensitive; intensity sensors are used where moderate performance is acceptable and lower cost is important.

Intensity sensors. The intensity sensor configurations shown in **Fig. 1** illustrate several techniques used in fiber sensors to probe the environment; over 100 sensor designs are currently available. Depending on their design, fiber sensors can respond to pressure, temperature, liquid level, position, flow, smoke, displacement, electric and magnetic fields, chemical composition, and numerous other conditions.

With the pressure sensor (Fig. 1*a*), the displacement of a diaphragm which responds to pressure is detected by light probes. Light is carried to the diaphragm by the central fiber and is reflected. The amount of light captured by the fibers surrounding the central fiber is proportional to the displacement of the diaphragm and thus to the pressure. *See* PRESSURE MEASUREMENT.

In the liquid-level sensor (Fig. 1*b*), light carried by the fiber is reflected up when the fiber is in air; however, in a liquid the reflection is dispersed and light lost into the liquid. By monitoring of the reflected light at the input end, the presence or absence of the liquid can be determined. *See* LEVEL MEASUREMENT.

The remote-detection sensor (Fig. 1*c*) uses the fiber to carry light to the location to be probed. Light is scattered by the gas or liquid at the end of the fiber and partially recaptured by the fiber. Analysis of the scattered light at the input end can yield information on chemical composition or the presence of species such as pollutants in water. A sensor of this type is used on over 1000 ships to monitor oil contamination in discharge water. These sensors have also been tested clinically to determine pH and oxygen and carbon dioxide pressure in blood. The small size of the fiber probe (100 micrometers) makes these sensors ideal for invasive testing.

Encoder sensors (Fig. 1*d*) utilize a bundle of fibers to read bar codes printed on the sample. A wide variety of sensitized-tip sensors (Fig. 1*e*) for temperature, pressure, and other parameters utilize a material attached to the fiber with optical properties that change with the value of the measured quantity. Temperature sensors using this principle typically have sensitivities of 0.1°C (0.2°F) with ranges over 200°C (360°F). In the microbend sensor (Fig. 1*f*), light can be lost from the fiber by squeezing it at sharp,

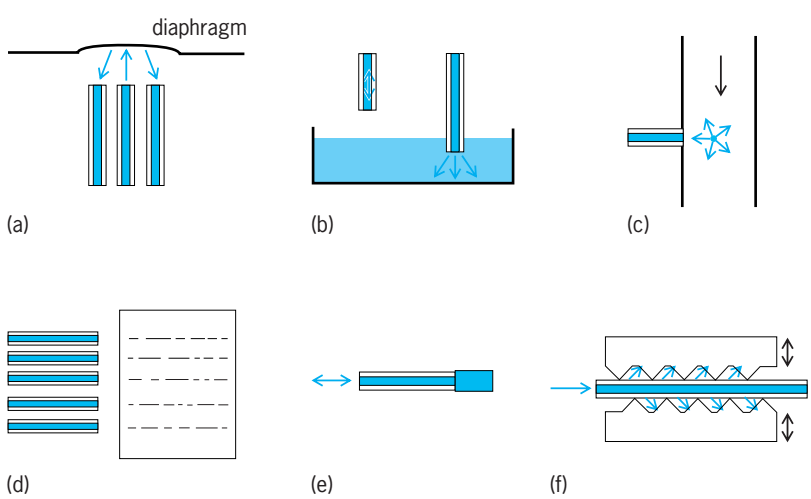

Fig. 1. Intensity fiber-optic sensors. (*a*) Pressure. (*b*) Liquid-level. (*c*) Remote detection. (*d*) Encoder. (*e*) Sensitized-tip. (*f*) Microbend.

distinct points. Pressure, temperature, liquid-level, and intruder sensors have all been built by using this principle.

Interferometric sensors. Optical interferometry is one of the most sensitive means of detecting displacements as small as 10^{-13} m. Interferometric fiber sensors apply this technology to sense many physical phenomena.

Fiber gyro. A fiber gyro based on the Sagnac effect (**Fig. 2***a*) is formed by making a fiber loop which, when rotated, causes the light traveling in both directions in the loop to experience different velocities with or against the rotation. Light is coupled into the fiber from a light source, and the fiber coupler divides the input light equally in both directions. After traversing the loop, the light is again combined in the coupler and detected photoelectrically. The phase shift of the light, which is a measure of the rotation rate, is given by Eq. (1), where λ is the wavelength of

$$\phi_s = \frac{\pi L D \Omega}{\lambda c} \tag{1}$$

the light, c is the speed of light, L the length of fiber, D the diameter of the loop, and Ω the rotation rate to be measured. By increasing the length of fiber L or the loop diameter D, the gyro can be made more sensitive. Minimum rotation rates of 10^{-4}°/h have been demonstrated in the laboratory. The type of fiber is known as polarization-preserving (Fig. 2*b*); that is, the polarization of the light in the fiber is controlled.

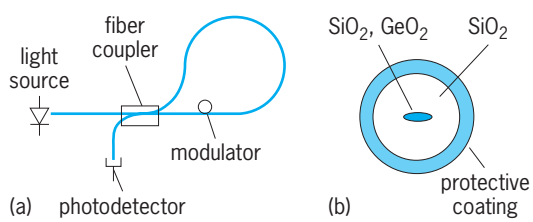

Fig. 2. Fiber gyro based on the Sagnac effect. (*a*) Configuration. (*b*) Cross section of polarization-preserving fiber.

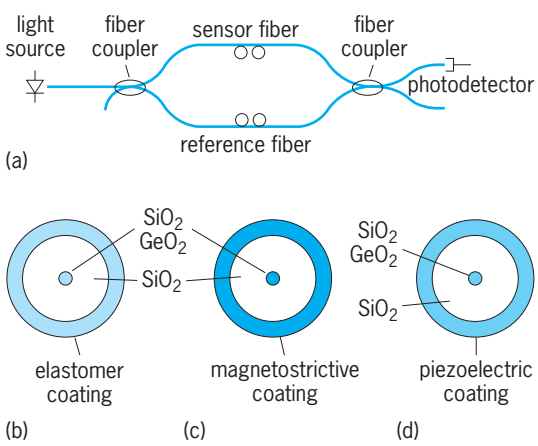

Fig. 3. Interferometric sensor based on Mach-Zehnder interferometer. (*a*) Configuration. (*b*) Cross section of fiber used in acoustic sensors, (*c*) magnetic sensors, and (*d*) electric sensors.

Polarization control is required in order to obtain the maximum sensitivity from Sagnac interferometers. *See* GYROSCOPE; POLARIZED LIGHT.

Mach-Zehnder sensor. A second type of interferometric sensor is constructed by using the Mach-Zehnder interferometer (**Fig. 3***a*). Light from a semiconductor laser is divided equally by a fiber coupler into the sensor and reference arms of the interferometer. The sensor fiber is treated with material which responds selectively to the phenomena to be sensed, whereas the reference fiber is selected not to respond. Coatings are applied to the fibers to achieve the desired sensitivity: elastomers for acoustic sensors (Fig. 2*b*), magnetostrictive material for magnetic sensors (Fig. 2*c*), piezoelectric material for electric sensors (Fig. 2*d*), and so on. These coatings stretch the fiber in response to the phenomena being measured; the induced phase shift caused by the stretching is given by Eq. (2), where λ and L are the wave-

$$\phi = \frac{2\pi}{\lambda}\left(\frac{1}{L}\frac{dL}{dP}\right)PL \tag{2}$$

length of the light and length of fiber respectively and P is the field strength being measured. This phase shift is converted into an amplitude modulation by the passage of the light through the second coupler. The modulated light or signal is detected photoelectrically. Sensitivities equal to or surpassing the best conventional technologies have been achieved in these sensors. For example, fiber hydrophones have demonstrated sensitivities of 30 dB re 1 micropascal (that is, the ability to detect signals whose intensity is 30 dB with respect to a sound wave whose acoustic pressure is 1 μPa) whereas the best piezoelectric acoustic sensors have shown sensitivities of 40 dB re 1 μPa. at 100 Hz; thermal fiber sensors have shown sensitivities of 10^{-7}°C with 5-kHz bandwidth, whereas the best conventional thermometer sensitivities are about 10^{-5}°C. *See* DECIBEL; INTERFEROMETRY; MAGNETOSTRICTION; PIEZOELECTRICITY.

Thomas G. Giallorenzi

Bibliography. K. Grattan and B. Meggitt, *Optical Fibre Sensor Technology*, 1992; B. P. Pal, *Fundamentals of Fibre Optics in Telecommunication and Sensor Systems*, 1993.

Fiber-optics imaging

The use of fiber optics in image transmission, based on the ability of a precisely aligned bundle of optical fibers to transmit an image from one end of the bundle to the other. Each fiber transmits one element of the image so that the complete image is made up of a matrix of dots (pixels) which blend into a recognizable image like a halftone picture on a printed page.

Light is transmitted through individual fibers by means of total internal reflection from the fiber walls. For efficient transmission, each fiber has a highly transparent core (usually glass) coated with a layer (cladding) of lower refractive index (also usually

glass). This cladding prevents light from leaking into neighboring fibers (crosstalk) and protects the core from contamination and wear. Fiber bundles may be either flexible or rigid. *See* CROSSTALK; OPTICAL MATERIALS; REFRACTION OF WAVES.

Flexible fiber bundles. Most flexible bundles are made by either winding or leaching. In the former a single or a multiple fiber is drawn from a preform, which is slowly lowered through an electric furnace, and is precisely wound onto a rotating drum or slotted frame. The helical ribbons thus formed are cemented along a short portion of their circumference, stactional area, and permanently bonded at the cemented portion. The loop formed is cut open and the end faces ground and polished to form a coherent image-transmitting fiber bundle.

In the leaching process the preform consists of a core rod with a double-glass cladding, the outer cladding being acid-soluble. This is drawn into small rods 0.04–0.08 in. (1–2 mm) in diameter, which in turn are assembled into a multifiber preform and drawn to a multifiber bundle. In some cases, this process is repeated. The final rigid rod is cut to the desired length and leached (except at the ends) to form a flexible bundle. In both cases, the fibers must be precisely aligned to avoid image distortion and loss of resolution.

A third (hybrid) form of the flexible image bundle consists of a fused bundle of clad silica-based fibers drawn in a silica tube to a diameter that is small enough to be flexible. In this all-fused type, the degree of flexibility strongly depends on the bundle diameter, which is not the case in the leached or wound types of bundles.

The effective image resolution is determined by the total number of fibers (pixels) which the bundle contains. The number is limited by the size of the bundle, since the minimum size of an individual fiber is about 8 micrometers (5 μm in all-fused types). In general, wound bundles are used for larger sizes from about 0.12 in. (3 mm) to as much as 1 in. (25 mm) in diameter. The leached type ranges from 0.12 in. (3 mm) to about 0.02 in. (0.5 mm). The total number of pixels can be as low as 2000 to more than a million in the largest wound bundles.

Flexible image bundles are used in a wide variety of industrial and medical fiberscopes or endoscopes to transmit the image from an objective lens at the distal end to an eyepiece in the control handle. In general, several components compete for space within the flexible shaft of the instrument, including the control wires, the fiber light guides (to illuminate the object), and channels to admit instruments.

Fiberscopes range in diameter from 0.04 in. (1 mm) to 0.6 in. (15 mm) or more. The smaller versions are of necessity simpler with little or no control of distal articulation and at most a single operating channel. Larger-diameter instruments have remote-tip articulation either in a single plane or, with dual control knobs, in all planes. All are equipped with fiber-optic illumination.

Industrial fiberscopes ranging in length from 3 to 8 ft (1 to 2.5 m) are used in aircraft engine inspection, for example, examination of turbine blades. Longer industrial fiberscopes are used for pipe and weld inspection, including those in nuclear power plants. Here radiation-resistant silica fiber bundles are employed for image transmission and illumination.

Medical endoscopes range in size from less than 0.08 in. (2 mm) for viewing inside the arteries (cardioscope) to 0.6 in. (15 mm) for examining the colon (colonoscope). The latter has an operating channel large enough to perform surgery via remotely controlled forceps or electrosurgical snares. Intermediate-sized endoscopes are used to view the bronchi (bronchoscope), the kidneys (nephroscope), the bladder (cystoscope), the throat (laryngoscope), and the stomach (gastroscope).

Rigid or fused fiber bundles. These are made from glass preforms by a series of drawing steps leading to multifibers and multi-multifibers, which are fused under controlled heat and pressure into a block or boule of fibers. Precise alignment is achieved by assembling multi-multifibers that are large enough to be accurately fitted together. The resulting boule contains millions of individual fibers with sizes ranging from 10 to 4 μm.

Since many applications of fused-fiber components involve the transmission of diffused light or light from large-aperture lenses, the components are designed to provide a wide acceptance angle (high numerical aperture). For this, glasses for the core and cladding are selected with a large refractive index difference. *See* GEOMETRICAL OPTICS.

Fused-fiber bundles often contain means for eliminating stray light caused by light entering the cladding and scattering throughout the component. This is done by adding a pattern of minute light-absorbing glass fibers among the transmitting fibers. These intercept the stray light but do not absorb the signal light.

Fused-fiber boules are the starting point for various image-transmitting components, including faceplates (windows) for image-intensifier tubes and cathode-ray tubes, image inverters (twisters), magnifiers, and rigid conduits. Faceplates of various shapes and sizes are cut and machined directly from the boule. The entrance and exit faces are ground and polished by normal lens-making processes. *See* CATHODE-RAY TUBE; IMAGE TUBE (ASTRONOMY); LIGHT AMPLIFIER.

One faceplate of an image intensifier transfers light from the image formed on its entrance face to the photocathode layer deposited on its exit face. This inner face may be concave to correct the aberrations of the electron lens in the intensifier. A second faceplate transmits the image formed on a phosphorescent screen deposited on the inner face of this plate to the exit face, where it is coupled to another fiber-optic component or an eyepiece.

Image intensifier tubes are used in military night-vision devices, low-light television cameras, and astronomical telescopes. In a cathode-ray tube the fiber-optic faceplate is used to transmit the image on the phosphor screen to the exit face for printing directly onto photographic material.

Image inverters and magnifiers require further hot-forming operations in their manufacture. In the former the boule is clamped at both ends, heated in the middle, and one end twisted through 180° (or any desired angle). Fiber-optic magnifiers are similarly made by stretching instead of twisting the boule. The resulting hourglass bundle is cut into two tapers and the ends are polished. Since the diameters of the two ends of the taper are different, they can be used to magnify or reduce an image. Image conduits are made by stretching (or drawing) a boule into a long rod of uniform diameter. Although rigid, these can be bent by heating into various shapes to pass around corners. *See* OPTICAL FIBERS.

Walter P. Siegmund

Fibrinogen

The major clot-forming substrate in the blood plasma of vertebrates. Though fibrinogen represents a small fraction of plasma proteins (normal human plasma has a fibrinogen content of 2–4 mg/ml of a total of 70 mg protein/ml), its conversion to fibrin causes a dramatic gelation which blocks the flow of blood. Upon injury, sufficient amounts of the clotting enzyme, thrombin, are generated in about 5 min clotting time to produce a gel.

Although clotting in the circulation (thrombosis) can be extremely dangerous, clotting is an essential and normal response (hemostasis) for preventing the loss of blood. Individuals born with the hereditary absence of fibrinogen (afibrinogenemia) suffer from severe bleeding, which can be counteracted by transfusing normal plasma or purified fibrinogen. There are also rare cases of abnormal fibrinogens (dysfibrinogenemia) which may be associated with a tendency for bleeding. Depletion of fibrinogen might occur by disseminated intravascular coagulation when substances which promote clotting enter the circulation (as in placenta abruptio). Premature degradation of fibrinogen by the proteolytic enzyme plasmin could also lower the fibrinogen concentration of plasma to levels where life-threatening hemorrhage might ensue. Malignant cells produce large amounts of activators of this fibrinolytic system, and the problem is often compounded by the fact that the fibrinogen-degradation products themselves interfere with clotting. Drugs such as ϵ-aminocaproic acid, which inhibit the conversion of plasminogen to plasmin, might be useful in such a situation. *See* HEMORRHAGE; THROMBOSIS.

Fibrinogen is synthesized by the hepatocytes in the liver, and the synthetic rate can be stimulated by hormones. Significant amounts of carbohydrates become attached to the protein before it is secreted into the circulation; fibrinogen is a plasma glycoprotein. Alterations in its carbohydrate composition as found in some liver diseases can give rise to abnormal fibrinogens with defective clotting properties.

Structure. In electron micrographs, fibrinogen has a length of about 45 nanometers, and the molecule appears to be symmetrical. The structure can be represented best by a model comprising globular as well as rodlike regions (**Fig. 1***a*). The diameters of the nodular domains (starting from either end) are 4.8, 4.8, 3.4, and 4.5 nm. The rods are 2.2 nm thick. Though not shown in Fig. 1, there are also two flexible appendages (a short and a long one) protruding from both ends of the molecule. The longer appendages are highly exposed, and they can be easily removed by proteolytic enzymes.

During clotting, an end-to-end association ensues (Fig. 1*b*), followed by a side-to-side aggregation of the filaments (Fig. 1*c*), producing a half-staggered overlap of individual molecules with a characteristic cross-striation of about 22.5 nm. In this array, the central nodule of one molecule interacts with the two end-to-end associated, terminal globules in the neighboring filament.

Composition. As far as chemical composition is concerned, human fibrinogen, with a molecular weight of 337,897, comprises two sets of three open polypeptide chains (Aα, Bβ, and γ) held together by disulfide bonds; thus, the composition can be expressed as $A\alpha_2B\beta_2\gamma_2$. Only the B$\beta$ and γ chains carry carbohydrate residues. Each of the three polypeptides is encoded by a distinct messenger ribonucleic acid, which indicates that the synthesis of fibrinogen is the result of the concerted expression of three genes. The Aα chain is the largest (mol wt 66,026), followed by Bβ (mol wt 54,358) and γ chain (mol wt 48,529). Though the exact mode of assembly of these constituent chains in the model given in Fig 1 is not yet known, they are arranged such that all N termini are located in the central globular region and all C termini are found on the outermost structures. The long, flexible appendages protruding from each end of the molecule comprise the C-terminal regions of the Aα chains and the short appendages the C termini of γ chains. The complete amino acid sequences of all chains of human fibrinogen are known. Considerable evolutionary variations exist in vertebrates

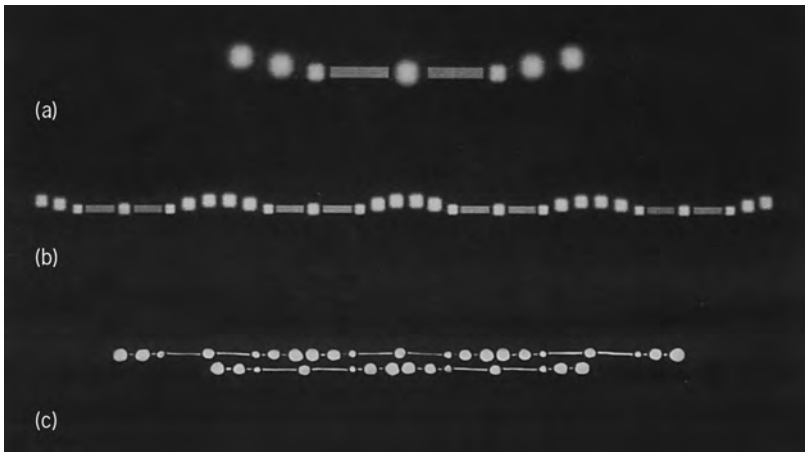

Fig. 1. Fibrinogen models. (*a*) Computer image of the symmetrical molecule. (*b*) Model for end-to-end aggregation. (*c*) Half-staggered molecular overlap between side-to-side aggregated filaments in a clot network. (*Parts a and b from C. Cohen et al., The structure of fibrinogen and fibrin, pt. I, Ann. N. Y. Acad. Sci., 408:194–213, 1983; c* from *J. S. Weisel, G. N. Phillips, Jr., and C. Cohen, pt. II, Ann. N. Y. Acad. Sci., 408:367–378, 1983*)

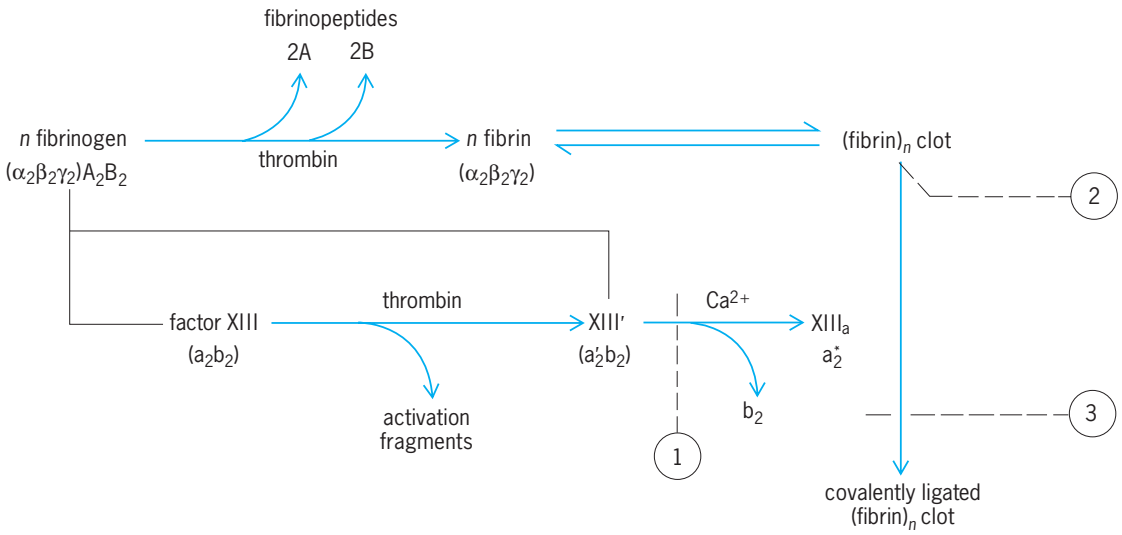

Fig. 2. Physiological pathway for the clotting of fibrinogen in normal plasma. Clotting time is measured by forming the (fibrin)$_n$ network; this, however, is not the end product. Unless this clot undergoes ligation by γ-glutamyl-ϵ-lysine cross-links, bleeding may ensue. Such disorders of fibrin stabilization occur in the hereditary absence of factor XIII or when inhibitors to one of the reaction steps (marked 1–3) appear in the blood circulation. (*After T. H. Spaet, ed., Progress in Hemostasis and Thrombosis, vol. 5, pp. 245–290, Grune and Stratton, 1980*)

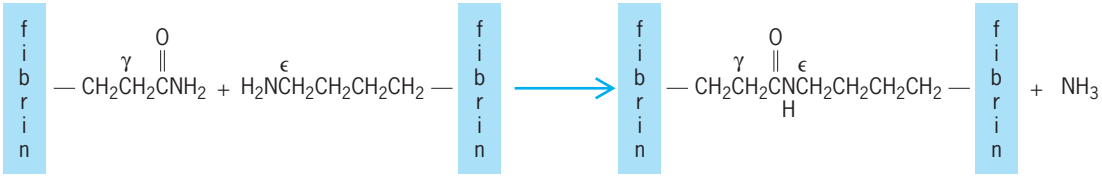

Fig. 3. Formation of γ-glutamyl-ϵ-lysine peptide bridges between fibrin molecules is the last enzyme-catalyzed reaction in blood clotting.

with regard to the sequences of N-terminal peptides in the Aα and Bβ chains (called fibrinopeptides A and B) and in the exposed C-terminal half of the Aα chain.

Clotting mechanism. It is now possible to reconstruct with reasonable certainty the sequence of events surrounding the clotting of fibrinogen in normal plasma, and each of the reaction steps in the scheme in **Fig. 2** can be studied separately with the use of purified components. Clotting is regulated by two enzymes, thrombin and factor XIII$_a$ (fibrinoligase, activated fibrin-stabilizing factor, transglutaminase). Thrombin is a protease similar to trypsin but of much greater specificity. It exerts a dual control by regulating the rate of fibrin formation as well as producing factor XIII$_a$ by processes of limited proteolysis. In the fibrinogen-fibrin conversion, about 97% of the parent protein becomes fibrin and 3% appears as fibrinopeptides A and B; in the reaction of thrombin with factor XIII, activation fragments are released from the N termini of the *a* subunits of factor XIII. The presence of fibrinogen greatly accelerates the latter reaction so that the production of fibrin and the activation of factor XIII are harmonized. In the plasma milieu, the fibrin molecules readily aggregate into a clot. In order to obtain a clot structure of a strength sufficient to stem bleeding, it is necessary for the thrombin-modified factor XIII (indicated as XIII′ in Fig. 2) to be activated to XIII$_a$. This is promoted by fibrinogen and is accomplished by the action of Ca^{2+} ions which dissociate the catalytic subunits (a^*_2) from the inert, carrier subunits (b_2). Factor XIII$_a$ acts as a transamidating enzyme which strengthens the fibrin clot by creating a few γ-glutamyl-ϵ-lysine cross-links (**Fig. 3**). Without such cross-links, a clot structure would be like a brick wall without mortar.

Individuals with the hereditary absence of factor XIII often suffer from severe bleeding, even though their clotting times are in the normal range. In addition, various molecular disorders of fibrin stabilization are known in which some acquired inhibitor suddenly makes its appearance in the circulation, directed either against the activation of factor XIII or against the cross-linking of fibrin by XIII$_a$. Several such inhibitors were found to be autoimmune IgG antibodies. Primary amines also specifically inhibit cross-linking without interfering with fibrin formation. *See* BLOOD; HEMOPHILIA; IMMUNOGLOBULIN.

L. Lorand

Bibliography. D. W. Chung et al., Cloning of fibrinogen genes and their cDNA, *Ann. N.Y. Acad. Sci.*, 408:449–456, 1983; R. F. Ebert, *Index of Variant Human Fibrinogens*, 1991; C. Y. Liu and S. Chen (eds.), *Fibrinogen, Thrombosis, Coagulation, and Fibrinolysis*, 1990; L. Lorand, Fibrinopeptide: New aspects of the fibrinogen-fibrin transformation, *Nature*, 167:992, 1951.

Fibromyalgia syndrome

A chronic condition characterized by widespread pain in muscles, tendons, and ligaments, and the presence of multiple tender points on the body. The illness occurs in up to 5% of the general population, and approximately 90% of affected individuals are women.

Symptoms and diagnosis. Like many other syndromes, fibromyalgia has no single specific feature but is a symptom complex of self-reported or elicited findings. In 1990, the American College of Rheumatology published diagnostic criteria, which include widespread pain for more than 3 months and muscle tenderness in at least 11 out of 18 so-called tender points, places on the body that are sensitive to the application of pressure. Most tender points are located at tendon insertion areas. Besides musculoskeletal pain and mechanical tenderness, most patients with fibromyalgia also complain of insomnia, fatigue, and distress. In addition, fibromyalgia appears to coaggregate with major mood disorders. *See* AFFECTIVE DISORDERS; MUSCLE; MUSCULAR SYSTEM.

Pathogenesis. The hypersensitivity of patients with fibromyalgia is not limited to pain but also includes light, sound, and smell. The cause for this heightened sensitivity is unknown, but central nervous system sensory processing abnormalities have been reported in several studies. Most of these studies found evidence of central sensitization of neurons of the spinal cord and the brain, including areas that are related to pain (such as the dorsal horn of the spinal cord and the thalamus). The pathogenesis of central sensitization in fibromyalgia is unclear but may be related to prior stressors, including infections and traumas. Because most patients report feeling pain in deep tissues, particularly muscles, these structures may play an important role in the initiation and maintenance of central sensitization and pain. No consistent tissue abnormalities have been reported in patients with fibromyalgia that would explain persistent pain. However, focal muscle areas, like trigger points, may be relevant for the pain of fibromyalgia because little nociceptive input from peripheral tissues is required for the maintenance of central sensitization. *See* NERVOUS SYSTEM (VERTEBRATE); PAIN.

Treatment. Treatment of fibromyalgia needs to be individually tailored to each patient's needs. The identification of pain generators is essential for an effective treatment plan. It is also important to assess biopsychosocial abnormalities (particularly anxiety and depression), which can be detected in most patients with fibromyalgia. Pharmacotherapy has been most successful with the use of antidepressant, muscle relaxant, and anticonvulsant drugs. These drugs affect the release of various neurochemicals (such as serotonin, norepinephrine, and substance P) that have a broad range of activities in the brain and spinal cord, including modulation of pain sensation and tolerance. None of these drugs, however, is currently approved by the U.S. Food and Drug Administration for the treatment of fibromyalgia. *See* NORADRENERGIC SYSTEM; PSYCHOPHARMACOLOGY; SEROTONIN.

The efficacy of patients to manage their pain seems to correlate with their functional status (change to physical function). Brain imaging and psychological profiles have identified at least three fibromyalgia subgroups, that is, patients who display high dysfunction, interpersonal distress, or effective coping mechanisms. Such studies provide an explanation of why some treatments seem to be differentially effective in individual patients.

Most patients' symptoms respond to low-dose tricyclic medications, such as amitriptyline and cyclobenzaprine, as well as cardiovascular exercise, cognitive–behavioral therapy, patient education, or a combination of these approaches. Thus, optimal management of fibromyalgia requires a combination of pharmacological and nonpharmacological therapies, with patients and health care professionals using a team approach. *See* PSYCHOTHERAPY.

Roland Staud

Bibliography. D. D. Price et al., Enhanced temporal summation of second pain and its central modulation in fibromyalgia patients, *Pain*, 99:49–59, 2002; R. Staud et al., Abnormal sensitization and temporal summation of second pain (wind-up) in patients with fibromyalgia syndrome, *Pain*, 91(1–2):165–175, March 2001; R. Staud et al., Body pain area and pain-related negative affect predict clinical pain intensity in patients with fibromyalgia, *J. Pain*, 5(6):338–343, August 2004; R. Staud et al., Ratings of experimental pain and pain-related negative affect predict clinical pain in patients with fibromyalgia syndrome, *Pain*, 105(1–2):215–222, September 2003; R. Staud and M. L. Smitherman, Peripheral and central sensitization in fibromyalgia: Pathogenetic role, *Curr. Pain Headache Rep.*, 6:259–266, 2002; C. J. Vierck et al., The effect of maximal exercise on temporal summation of second pain (wind-up) in patients with fibromyalgia syndrome, *J. Pain*, 2(6):334–344, 2001.

Fidelity

The overall input-output accuracy of a signal storage or transmission system. Many attributes contribute to fidelity. In an audio recording-playback system, some of the individual attributes are distortion, bandwidth, signal-to-noise ratio, group delay, and loudspeaker dispersion. Most system designers would agree that noise and distortion are of chief importance. The fidelity of video systems depends on many attributes, including optical quality, bandwidth, dynamic range, and phosphor brilliance, whereas that of a motion picture system depends on optical quality, mechanical stability in camera and projector, and the exposure-density relationships of the film. *See* DISTORTION (ELECTRONIC CIRCUITS); ELECTRICAL NOISE; SIGNAL-TO-NOISE RATIO.

In a given system, fidelity is improved when the linearity of each constituent element in the system has been improved. However, fidelity may be

increased by redefining the system. For example, stereophonic sound transmission is inherently of superior fidelity compared with monophonic transmission in that it provides auditory perspective. It can be argued that two stereophonic channels of moderate quality offer greater fidelity than a single monophonic channel of the highest quality, thus underscoring the subjective basis for assessing fidelity. Similarly, most viewers would agree that color video transmission presents greater fidelity than monochrome transmission. *See* CINEMATOGRAPHY; LINEARITY; SOUND-REPRODUCING SYSTEMS; STEREOPHONIC SOUND; TELEVISION. John Eargle

Field emission

The emission of electrons from a metal or semiconductor into vacuum (or a dielectric) under the influence of a strong electric field. In field emission, electrons tunnel through a potential barrier, rather than escaping over it as in thermionic emission or photoemission. The effect is purely quantum-mechanical, with no classical analog. It occurs because the wave function of an electron does not vanish at the classical turning point, but decays exponentially into the barrier (where the electron's total energy is less than the potential energy). Thus there is a finite probability that the electron will be found on the outside of the barrier. This probability varies as $e^{-cA^{1/2}}$, where c is a constant and A the area under the barrier. *See* NONRELATIVISTIC QUANTUM THEORY; PHOTOEMISSION; THERMIONIC EMISSION.

For a metal at low temperature, the process can be understood in terms of the **illustration**. The metal can be considered a potential box, filled with electrons to the Fermi level, which lies below the vacuum level by several electronvolts. The distance from Fermi to vacuum level is called the work function, ϕ. The vacuum level represents the potential energy of an electron at rest outside the metal, in the absence of an external field. In the presence of a strong field F, the potential outside the metal will be deformed along the line AB, so that a triangular barrier is formed, through which electrons can tunnel. Most of the emission will occur from the vicinity of the Fermi level where the barrier is thinnest. Since the electron distribution in the metal is not strongly temperature-dependent, field emission is only weakly temperature-dependent and would occur even at the absolute zero of temperature. The current density J is given by the Fowler-Nordheim equation below, where B is a field-

$$J = BF^2 e^{-6.8\times10^7 \phi^{3/2}/F}$$

independent constant of dimensions A/V^2 (A = amperes or state), ϕ is work function in electronvolts, and F is applied field in V/cm. The factor $\phi^{3/2}/F$ is proportional to the square root of the area under the barrier at the Fermi level. Appreciable emission requires fields of 4–7 × 10^7 V/cm, depending on ϕ. *See* FREE-ELECTRON THEORY OF METALS; WORK FUNCTION (ELECTRONICS).

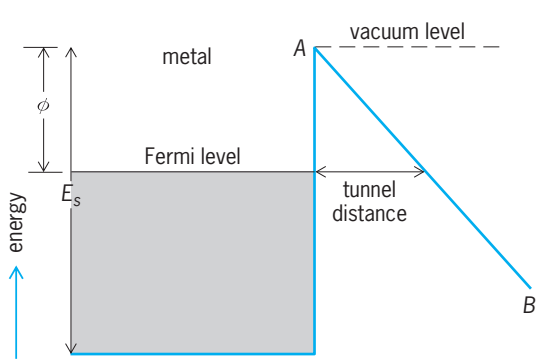

Energy-level scheme for field emission from a metal at absolute zero temperature.

Field emission is most easily obtained from sharply pointed metal needles whose ends have been smoothed into nearly hemispherical shape by heating. Tip radii r_t equal to or less than 100 nanometers can be obtained in this way; because of its small size, the emitter is generally a single crystal. If an emitter is surrounded by a hemispherical anode raised to a voltage V, then the field F is approximately $V/5r_t$ at the emitter. Thus, modest voltages suffice for emission. The electric lines of force diverge radially from the tip; since the electron trajectories initially follow the lines of force, they also diverge, and a highly magnified emission map of the emitter surface can be obtained, for instance, by making the anode a fluorescent screen. This constitutes a field emission microscope, invented by E. W. Müller in 1936. Since work function and hence emission are affected by adsorbed layers, the field emission microscope is very useful for studying adsorption, particularly surface diffusion. Field emitters are widely used in "ordinary" and scanning electron microscopes as high-brightness quasi-point sources of electrons, since emission occurs as if it originated from the center of the emitter cap. *See* ELECTRON MICROSCOPE.

Field emission can also occur from electrode asperities into insulating liquids, and thus isolated atoms or molecules in high fields, and is then called field ionization. Field ionization forms the basis of the field ion microscope, and is a useful method of generating ions in analytical mass spectrometry. Internal field emission can occur from the valence to the conduction band of a semiconductor in a high field, and is then known as Zener breakdown. *See* ELECTRON EMISSION; FIELD-EMISSION MICROSCOPY. Robert Gomer

Field-emission microscopy

A technique that uses field emission of electrons or positive ions from a needle-shaped emitter to produce a magnified image of the emitter surface on a fluorescent screen. In the field electron microscope, the image reveals the variation in work function of the emitter surface. Due to the large lateral

velocity of the emitted electrons, a resolution of only about 2.5 nanometers can be achieved. The large lateral velocity arises from a diffraction effect of the de Broglie wave and the large kinetic energy of electrons inside the metal; these effects are intrinsic quantum properties of particles and cannot be eliminated. The field electron microscope has been used to study adsorption and desorption of gases and vapor-deposited materials, surface migration of adsorption layers and absorbed atoms on single crystal faces, and surface reactions in catalysis. Medium-sized individual molecules such as phthalocyanin have been made visible also. *See* DE BROGLIE WAVELENGTH; ELECTRON MICROSCOPE; FIELD EMISSION; WORK FUNCTION (ELECTRONICS).

Field ion microscope. In the field ion microscope, the emitter is kept at a high positive potential while the microscope chamber is filled with helium, neon, or argon at a pressure of 10^{-5} to 10^{-4} torr (10^{-3} to 10^{-2} pascal) [**Fig. 1**]. Under a field of several tens of volts per nanometer, an image gas molecule can be ionized above a protruding surface atom when an atomic electron tunnels into the metal. The ion is then accelerated to the screen by the applied field. Every second about 1000 ions are formed above the same surface atom; thus the atom is continuously imaged. When the tip is cooled down to the temperature of liquid hydrogen or nitrogen, the thermal energy of the image gas molecules is greatly reduced. In combination with their very short de Broglie wavelength, a resolution of about 0.25 nm can be achieved. This is sufficient to resolve the atomic structure of most surfaces. In addition, surface atoms can be evaporated by the applied field if it is gradually increased. Therefore the atomic structure of lattice defects inside the bulk can be revealed also. By using a channel plate for image intensification and image gases of high or low ionization energy, a wide range of metals from beryllium to uranium, semiconductors, and some compounds such as high-temperature oxide superconductors can be imaged. The field ion microscope (Fig. 1) can also be used as a field electron microscope by simply reversing the polarity of the tip voltage and operating it in high vacuum. *See* TUNNELING IN SOLIDS.

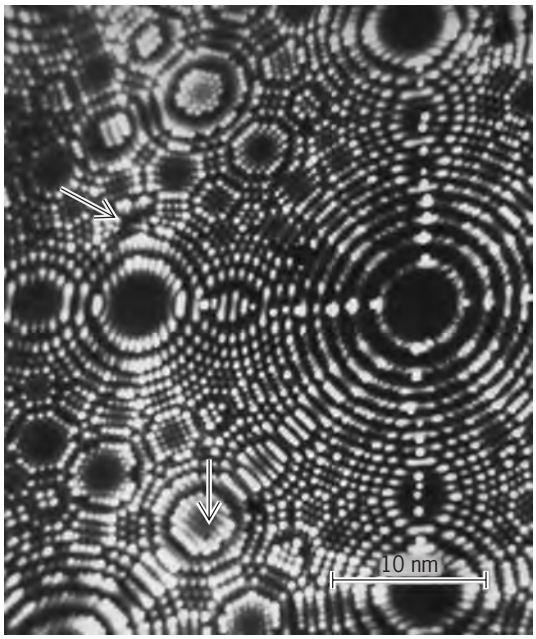

Fig. 2. Field ion micrograph of the atomic structure of an iridium crystal surface, with a vacancy and a dislocation shown by arrows, and many fully resolved net planes of atoms.

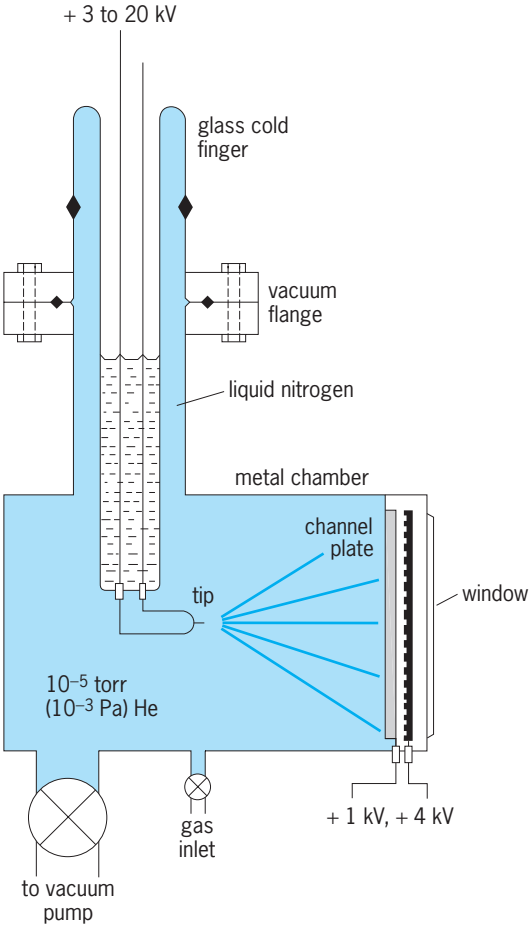

Fig. 1. Field ion microscope.

The field ion microscope has been used as a research tool for studying lattice defects such as vacancies (**Fig. 2**), interstitials, dislocations, grain boundaries, and radiation damages in metals and alloys. It has also been used to observe directly the behavior of single adsorbed atoms on metal surfaces. Random-walk diffusion of single atoms, interactions between single atoms, formation of atomic clusters from single atoms and their dissociation, and adsorption-layer superstructure formation have been observed and detailed atomic steps revealed. The field ion microscope has also been employed to study atomic reconstruction of solid surfaces and the formation of precipitates, oxide layers, thin silicide layers, and amorphous thin films. *See* CRYSTAL DEFECTS.

A slightly modified version of the field ion microscope is the field desorption microscope, which uses a double-channel plate. No image gas is needed; field desorbed ions form their own images. The desorption image reveals directly the spatial distribution of the adsorbed species on the surface. A shadowing technique has also been developed to image biological molecules. The field desorption images of adsorbed methane can reveal the shape of the biological molecules embedded beneath the methane layers.

Atom-probe field ion microscope. The instrument combines the field ion microscope with a single ion detection sensitivity mass spectrometer, usually a time-of-flight spectrometer. The tip is mounted on a gimbal system, and a channel plate screen assembly with a small probe hole is used. Behind the probe hole is a flight tube 3–24 ft (1–8 m) long (**Fig. 3**). The probe hole usually covers several atomic image diameters. The chemical identity of atoms chosen by the investigator from the field ion image can be analyzed one by one by adjusting the gimbal until the image of the atoms falls into the probe hole. Nanosecond high-voltage pulses or subnanosecond laser pulses are then applied to field-evaporate surface atoms one by one. Because of the ion optics, only those atoms that have their image covered by the probe hole can go through the probe hole and be detected. From the flight times of these ions, their mass-to-charge ratios are calculated and chemical species identified.

The atom-probe field ion microscope represents the ultimate in chemical analysis, that is, an atom selected by an investigator can be analyzed. As field evaporation starts from the edges of a surface layer, composition of a solid surface can also be analyzed atomic layer by atomic layer. The spatial resolution in atom-probe chemical analysis is approximately 5 to 10 angstroms in the lateral direction and much better than 1 angstrom in the vertical direction.

The three-dimensional atom probe uses a wide-angle position-sensitive ion detector for measuring the flight times of field-evaporated ions from a large area of the emitter surface. With this instrument, it is possible to analyze the three-dimensional elemental distribution of a material with a spatial resolution of approximately 0.5 nm.

The atom-probe field ion microscope has been successfully employed in surface science and materials science to study problems such as compositional variations in alloys due to segregation of alloy species and impurities to the surface, grain boundaries and phase boundaries, atomic exchange in surface diffusion, composition of small alloy phases such as precipitates and clusters, range profiles in particle irradiations, gas-surface interactions, and surface-catalyzed chemical reactions. The advantages of the atom-probe field ion microscope over other microanalytical instruments are the good spatial resolution and the high degree of sensitivity. Thus, very small precipitates and compositional modulations of subnanometer sizes in materials, where other instruments fail, can be analyzed with the atom-probe. *See* ALLOY STRUCTURES; GRAIN BOUNDARIES; SURFACE PHYSICS; TIME-OF-FLIGHT SPECTROMETERS.

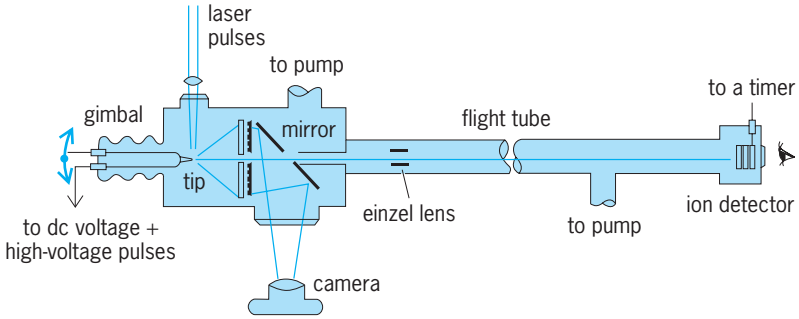

Fig. 3. Time-of-flight atom-probe field ion microscope. The einzel lens is an electrostatic one-piece lens for focusing the ion beam.

Tien Tzou Tsong

Bibliography. E. W. Müller and T. T. Tsong, *Field Ion Microscopy: Principles and Applications*, 1969; J. A. Panitz, Surface characterization by single-atom mass spectroscopy, *Crit. Rev. Solid State Sci.*, 5:153–178, 1975; T. T. Tsong, *Atom-Probe Field Ion Microscopy*, 1990; T. T. Tsong, Atom-probe field ion microscopy, *Phys. Today*, 46:24, 1993.

Field-programmable gate arrays

Very large scale integrated (VLSI) circuits that can be electrically programmed to implement complex digital logic. By combining many of the advantages of custom chips with the programmability of microprocessors, they are useful in a wide range of electronics applications. The name derives from their structure—a grid of programmable logic elements (gate array) and interconnection wires that can be configured at an engineers workbench, or even inside a consumer electronic device (programmed in the field), as opposed to other technologies that must be preconfigured during chip fabrication. *See* INTEGRATED CIRCUITS.

Design. Field-programmable gate arrays generally consist of a large number of programmable logic blocks embedded in a flexible routing structure (**Fig. 1**). The logic blocks implement the logic functions and flip-flops (simple memory elements that allow the system to remember past inputs and behaviors), while the routing structure wires these elements together into a complete system (**Figs. 2** and **3**). Typically, the logic blocks contain look-up tables, small 1-bit-wide memories that store the desired output for any possible combination of their inputs. To configure a look-up table to compute the function "return true if both A and B are true, or C is true," the table would have false for the three cases where C is false and at least one of A and B are false (ABC = {000, 010, 100}), and true for the other five cases. In a similar manner, any 3-input combinational Boolean function can be implemented in a 3-input look-up

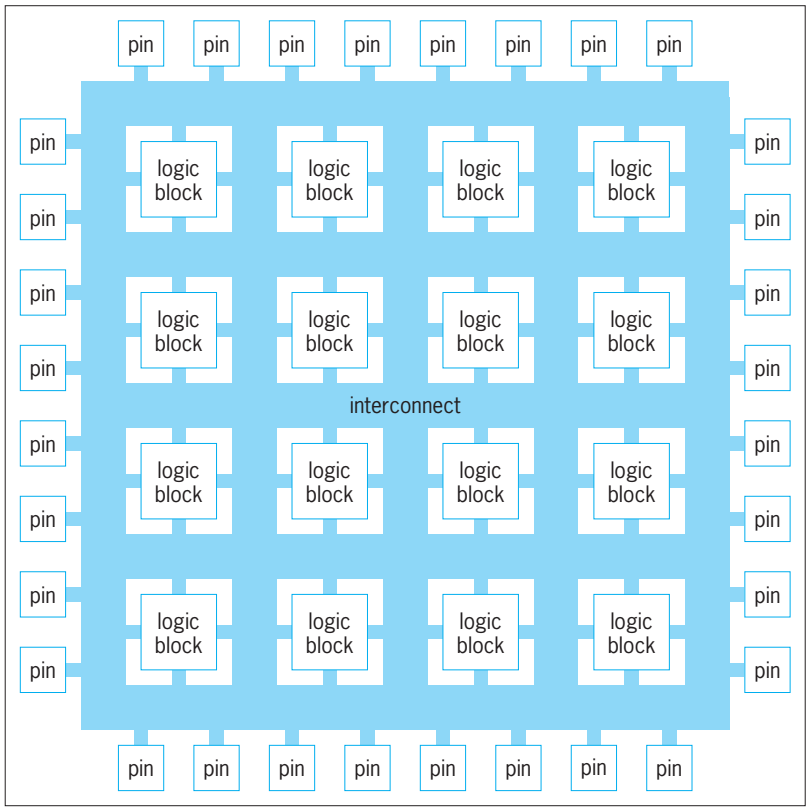

Fig. 1. Basic structure of a typical field-programmable gate array.

table. *See* LOGIC CIRCUITS; SEMICONDUCTOR MEMORIES.

Look-up tables are usually coupled with an optional D flip-flop (Fig. 2). All D flip-flops in the field-programmable gate array receive a periodic signal called a "clock," which transitions millions to hundreds of millions of cycles a second. Every time the clock changes from true to false, the D flip-flop's input D is captured and sent on to the output Q. In this way, the global clock signal synchronizes the operation of all memory elements in the system, and allows the design to remember and react to past inputs and behaviors.

The muxiplexor (Fig. 2) picks between the outputs of the look-up table and the flip-flop. If the programming bit P stores a 0 (false) value, the look-up table's output is sent to the output of the logic block; while if the programming bit is 1 (true), the D flip-flop's output is sent. In this way, the logic block can be programmed to compute a clocked or unclocked function.

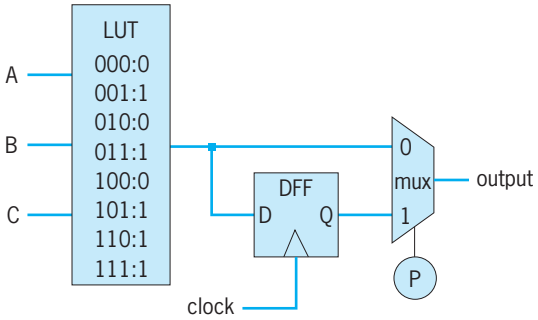

Fig. 2. Abstract representation of a logic block in a typical field-programmable gate array. LUT = Look-up table. DFF = D flip-flop (input D; output Q). Mux = muxiplexor (programming bit P).

The routing structure provides fairly flexible wiring connections between the outputs and inputs of the logic blocks on the chip. This allows the relatively simple logic blocks to be combined to support a complex computation, perhaps computing a function equivalent to a million or more basic 2-bit Boolean functions. The routing structure also connects to the chip input and output pins. The routing wiring pattern is generally fixed once the field-programmable gate array is programmed, instead of the dynamically changing communication pattern found in a typical computer network.

Modern field-programmable gate arrays often augment these basic elements with more complex units. Logic blocks have elements to speed up additions and subtractions of large values, while other parts of the array replace the logic blocks with more special-purpose elements. These might include multipliers, memories, and even complete microprocessors. Also, the input/output (I/O) pins often include the special hardware needed to support high-speed communication protocols to other chips or communication networks.

Programming. All of these elements within the field-programmable gate array are electrically programmable. The chip contains huge numbers of programming bits, individual elements that optionally turn on or off different features in the array. For example, each entry in a look-up table will be a bit programmable to represent either a true or a false. These programming bits might be implemented using antifuse, SRAM (static random-access memory), or EEPROM/Flash (electrically erasable programmable read-only memory/flash memory) technology. Antifuse programming points are normally disconnected wires in the array that can be "blown" by a high voltage to make a permanent electrical connection. SRAM are memory bits like the cells in a computer memory; they can hold a value for as long as power is maintained to the chip, but can also be changed at will. EEPROM/Flash is a hybrid of the two, based on isolated capacitors that can hold a value without power being applied, yet can be reprogrammed via high voltages.

The basic programming technology of a field-programmable gate array affects the way that it can be used. Chips based upon SRAM technology generally must have some external memory to hold the configurations, since they lose their values when the power is turned off; antifuse and EEPROM/Flash do not. Antifuse devices can be programmed only once; SRAM or EEPROM/Flash devices can be reconfigured to fix problems or change configurations. Because of this reconfigurability, SRAM-based and EEPROM/Flash-based FPGAs will be referred to as reconfigurable field-programmable gate array here. Overall, SRAM-based devices are generally the most

successful commercially, since they do not need special fabrication techniques and thus can use the latest chip technologies.

To configure a field-programmable gate array for a specific circuit, a designer typically starts with code written in a hardware description language such as Verilog or VHDL. These are languages similar to C or Ada but are optimized for specifying hardware designs. This code is then converted to basic logic gates, such as ANDs and ORs, by a logic synthesis tool. A "technology-mapper" takes these logic gates and groups them into chunks that fit in the look-up tables in the field-programmable gate array. Finally, these look-up-table size groupings are assigned into specific logic blocks and routed together by the "placement" and "routing" tools respectively. At the end of this process, a file with the assignments to the field-programmable gate array's configuration bits (known as a bitstream) is generated. When this file is downloaded into a field-programmable gate array, it is configured to implement the user's design. Each of these steps is highly automated by software tools. *See* PROGRAMMING LANGUAGES.

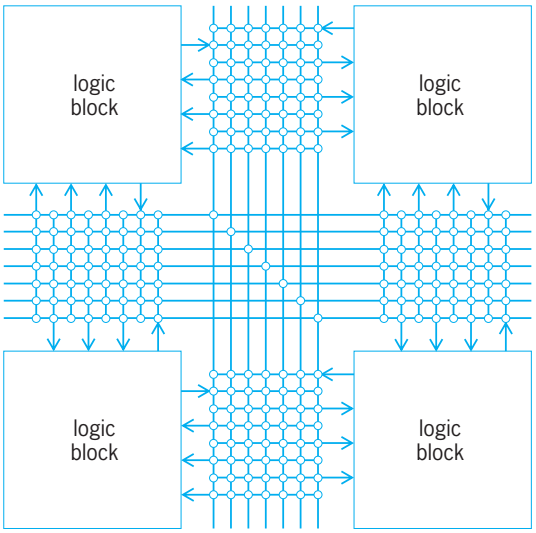

Fig. 3. Abstract representation of the interconnect structure for a typical field-programmable gate array. The circles are programmable connections between crossing wires.

Applications. Field-programmable gate arrays can be viewed as occupying a middle ground between programmable processors and custom chips. Compared to custom chips developed for a specific application (application-specific integrated circuits, or ASICs), field-programmable gate arrays are generally slower, are less power-efficient, and have a smaller logic capacity. However, an ASIC requires a much greater design time and a large up-front cost to set up fabrication for the first chip. Furthermore, unlike reconfigurable field-programmable gate arrays, ASICs cannot be changed after fabrication to handle bug fixes or functionality upgrades.

Compared to programmable processors, such as microprocessors and digital signal processors, field-programmable gate arrays can be faster, denser, and more power-efficient. However, field-programmable gate arrays are only more efficient than processors on tasks involving long repetitive tasks, such as multimedia processing and networking. Also, designing for field-programmable gate arrays is more complex than for microprocessors, and field-programmable gate arrays usually cannot support running multiple applications simultaneously (multitasking). *See* MICROPROCESSOR.

Although a microprocessor-based solution is usually preferred over an implementation with either an ASIC or a field-programmable gate array, some applications have speed or power requirements that make programmable processors unsuitable. Because of their relative strengths versus ASICs, field-programmable gate arrays are typically used in the following applications.

Low- to medium-volume hardware. Because ASICs have very high initial costs that must be amortized over each system sold, custom hardware may be affordable only for products that will sell a large number of units. And while in the past field-programmable gate arrays were able only to implement small parts of an application, their significant increase in capacity have allowed field-programmable gate arrays to now support complete systems in a single chip.

Fast time-to-market. Even for applications with enough sales volume to justify the initial costs of an ASIC, the design time may be too long. Field-programmable gate arrays, which require less design time and have no extra delays for custom fabrication, can help get a product to market much faster than an ASIC.

Rapid prototyping. When a designer is developing a new circuit, it is often crucial to create a working prototype of the system for testing. This prototype will need to be changed multiple times as bugs and limitations are found. While ASICs cannot be changed, the reprogrammability of reconfigurable field-programmable gate arrays makes them ideal for prototyping.

Upgradability. For some applications, it is clear the circuitry will change after the system is sold. This includes devices supporting evolving protocols, complex designs susceptible to bugs, and applications where new features might be added over time. Reconfigurable field-programmable gate arrays can easily handle such systems, with the logic changed simply by providing a new bitstream.

Multimode systems. Some systems will need to support many different applications. For example, a single cellphone design might need to handle multiple protocols, depending on the country and the cellphone network provider. A field-programmable gate array can implement each protocol simply by providing it with the appropriate configuration file.

Run-time reconfiguration. A more aggressive use of a field-programmable gate arrays reconfigurability is to change the loaded configuration rapidly during the operation of the system. A field-programmable gate array otherwise is too small to implement a large computation but can do so if that computation is broken down into multiple, smaller pieces. This

significantly increases the virtual capacity of the field-programmable gate array. ASICs would have to implement all of the configurations simultaneously.

Field-programmable gate arrays were introduced commercially in 1985, then became a multibillion-dollar industry. Due to their reprogrammability, ease of use, and low cost of hardware and support software, SRAM-based field-programmable gate arrays are one of the best hardware technologies for the novice or hobbyist. Scott A. Hauck

Bibliography. S. Brown and Z. Vranesic, *Fundamentals of Digital Logic with Verilog Design*, McGraw-Hill, 2002; K. Compton and S. Hauck, Reconfigurable computing: A survey of systems and software, *ACM Comput. Surveys*, 34(2):171–210, June 2002; S. Hauck, The roles of FPGAs in reprogrammable systems, *Proc. IEEE*, 86(4):615–639, April 1998; J. Villasenor and W. H. Mangione-Smith, Configurable computing, *Sci. Amer.*, pp. 66–71, June 1997.

Field theory (mathematics)

In algebra, the term field is used to designate an algebraic system or structure containing at least two elements and having two binary rules of composition: addition and multiplication (that is, if a and b are any two elements of the field, then $a + b$ and ab are defined and are elements of the field). The structure rules are as follows: The elements form an abelian (commutative) group under addition with the additive identity denoted by 0; that is, $a + 0 = a$ for all elements a. The set of nonzero elements (and there are some since the field has at least two elements) form an abelian group under multiplication with the multiplicative identity denoted by 1. It follows that all nonzero elements have a multiplicative inverse or reciprocal. The two rules of composition are related by the distributive law: $(a + b)c = ac + bc$ for all elements a, b, c. It follows from the distributive law that $a \cdot 0 = 0$ for all elements a, since $1 \cdot a = (1 + 0)a = 1 \cdot a + 0 \cdot a$, whence $0 = 0 \cdot a$. *See* GROUP THEORY.

The most familiar example of a field is the set $\mathbf{Q}$ of all rational numbers ($0, 1, 2, {}^1/_2, -1, -2, -{}^1/_2, 3, {}^1/_3, {}^2/_3, \ldots$) with the usual rules for equality, addition, and multiplication. Other examples are the set $\mathbf{R}$ of all real numbers and the set $\mathbf{C}$ of all complex numbers (both with the usual rules of addition and multiplication). The set of rational functions (formal quotients of two polynomials with rational coefficients) also constitute a field under the usual rules for equality, addition, and multiplication. *See* ALGEBRA; COMPLEX NUMBERS AND COMPLEX VARIABLES; REAL VARIABLE.

A field satisfies all the axioms for a commutative ring with unity and in addition has the property that it contains the reciprocal of each nonzero element. A field is also an integral domain; in particular, if $ab = 0$ and $a \neq 0$, then $b = (a^{-1}a)b = a^{-1}(ab) = a^{-1} \cdot 0 = 0$. *See* RING THEORY.

There are fields that contain only finitely many elements. The simplest example is obtained by considering the residue class ring of the integers modulo a prime p. This field, which may be denoted by $\mathbf{F}_p$, has as its elements the p distinct arithmetic progressions or residue classes modulo the prime p, namely $I_0, I_1, \ldots, I_{p-1}$, where I_j consists of the integers $j, j \pm p, j \pm 2p, \ldots$. The rules of composition are given by Eqs. (1) and (2), and can be used to construct addition and multiplication tables (**Table 1**). The progression I_0 is the additive identity since $I_0 + I_k = I_k$ for all k, and I_1 is the multiplicative identity since $I_1 I_k = I_k$ for all k. In $\mathbf{F}_p$, if any I_k is added to itself p times, I_0 is obtained. The residue class rings modulo a composite number do not form a field, since in those rings some nonzero elements do not have multiplicative inverses.

$$I_j + I_k = I_s \text{ if } j + k \text{ lies in } I_s \tag{1}$$

$$I_j I_k = I_t \text{ if } jk \text{ lies in } I_t \tag{2}$$

Subfields and extensions. If F is a field and a subset A of F is a field under the rules of composition of F, then A is said to be a subfield of F and F is said to be an extension of A. Thus the field of rational numbers $\mathbf{Q}$ is a subfield of the field of real numbers $\mathbf{R}$ and $\mathbf{R}$ is an extension of $\mathbf{Q}$. If a field F is an extension of the field A, then F is a vector space over A. If, as a vector space, F is of finite dimension n over A, then F is said to be an extension of A of degree n. In particular, $\mathbf{C}$ is an extension of $\mathbf{R}$ of degree 2. *See* LINEAR ALGEBRA.

Two fields F and F' are said to be isomorphic if there is a one-to-one mapping f of the elements of F onto the elements of F' such that $f(a + b) = f(a) + f(b)$ and $f(ab) = f(a)f(b)$ for all pairs of elements a, b in the field F. Every field F contains a unique subfield which is isomorphic either to the field $\mathbf{Q}$ of rational numbers or to $\mathbf{F}_p$ for some prime p. The finite field $\mathbf{F}_p$ occurs when repeated addition of 1 to itself gives 0. It follows that if F is a finite field, then it is an extension of $\mathbf{F}_p$, for some prime p, and hence is a vector space of finite dimension over $\mathbf{F}_p$. Thus the cardinality, or number of elements, of a finite field must always be a power of a prime. All

TABLE 1. Addition and multiplication tables for the residue class ring of the integers modulo $p = 5$

+	I_0	I_1	I_2	I_3	I_4
I_0	I_0	I_1	I_2	I_3	I_4
I_1	I_1	I_2	I_3	I_4	I_0
I_2	I_2	I_3	I_4	I_0	I_1
I_3	I_3	I_4	I_0	I_1	I_2
I_4	I_4	I_0	I_1	I_2	I_3

×	I_0	I_1	I_2	I_3	I_4
I_0	I_0	I_0	I_0	I_0	I_0
I_1	I_0	I_1	I_2	I_3	I_4
I_2	I_0	I_2	I_4	I_1	I_3
I_3	I_0	I_3	I_1	I_4	I_2
I_4	I_0	I_4	I_3	I_2	I_1

TABLE 2. Addition and multiplication tables for the field J of four elements

$+$	0	1	a	$a+1$
0	0	1	a	$a+1$
1	1	0	$a+1$	a
a	a	$a+1$	0	1
$a+1$	$a+1$	a	1	0

$\times$	0	1	a	$a+1$
0	0	0	0	0
1	0	1	a	$a+1$
a	0	a	$a+1$	1
$a+1$	0	$a+1$	1	a

finite fields of the same cardinality are isomorphic to each other. For example, a field of four elements is isomorphic to the field J (**Table 2**), which contains $\mathbf{F}_2$ and is an extension of $\mathbf{F}_2$. In a finite field there exists an element whose powers give all nonzero elements. In the case of $\mathbf{F}_5$, both I_2 and I_3 are such elements. In the field J (Table 2), there are two such elements, which may be designated a and $a+1$.

Given an integral domain D, there exists (to within isomorphism) a unique smallest field $Q(D)$ containing D. The field $Q(D)$ is called the quotient field of D, and it stands in relation to D as the field $\mathbf{Q}$ of rational numbers stands in relation to the ring of integers; indeed $\mathbf{Q}$ is the quotient field of the ring of integers.

Solutions of polynomial equations. Every linear equation $ax + b = 0$ with $a \neq 0$, and a, b in a field F has a solution in F, namely $-(b/a)$. But this need not be the case for equations of higher degree; for example; neither $x^2 - 2 = 0$ nor $x^2 + 1 = 0$ has a solution in the field $\mathbf{Q}$ of rational numbers. Given a field F and a polynomial $p(x) = a_0x^n + \cdots + a_n$, $n > 0$, with $a_0 \neq 0$, $a_1, \ldots, a_n$ in F, there exist extensions of F containing a solution for $p(x) = 0$. For example, the set B, consisting of all elements of the form $a + b\sqrt{2}$, where a and b are rational numbers, with the composition rules $(a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2}$ and $(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2}$ is a field which extends the field of rational numbers and contains a solution of the polynomial equation $x^2 - 2 = 0$. The field J contains a root of the polynomial $x^2 + x + 1$, which is inducible over $\mathbf{F}_2$. If $p(x)$ is a polynomial with coefficients in a field F and if $p(x)$ is irreducible over F, that is, $p(x)$ does not factor into a product of polynomials with coefficients in F of lower degree, then there is a unique (to within isomorphism) extension K of F containing a solution for the polynomial equation $p(x) = 0$. If the degree of a field A over F is n, then every element of A satisfies an irreducible polynomial over F of degree at most n.

If a field F is such that every polynomial equation over F of positive degree has a solution in F, then F is said to be algebraically closed. The field of all complex numbers is algebraically closed; but this is a theorem in analysis. The field of rational numbers, the fields of residue classes modulo a prime p, and the field of real numbers are not algebraically closed. Given any field F, there exists a unique (to within isomorphism) minimal extension $\tilde{F}$ of F such that $\tilde{F}$ is algebraically closed. If the field F is countable (that is, can be put into a one-to-one correspondence with the positive integers), then $\tilde{F}$ is countable. Since the field $\mathbf{Q}$ of rational numbers and the fields $\mathbf{F}_p$ are countable, it follows that $\tilde{\mathbf{Q}}$ and $\tilde{\mathbf{F}}_p$ are countable. In particular, it follows that $\tilde{\mathbf{Q}} \neq \mathbf{C}$, since $\mathbf{C}$ is not countable. It is easily seen that $\mathbf{C}$, the field of complex numbers, is isomorphic to $\tilde{\mathbf{R}}$, the algebraic closure of the field of real numbers.

If K is an extension of a field F, an element of K is said to be algebraic over F if that element is a solution of a polynomial equation with coefficients in the field F. If an element of the larger field K is not algebraic over F, the element is said to be transcendental over F. The real numbers $\sqrt{2}$ and $\sqrt[3]{1 + 3\sqrt{5}}$ are algebraic over the field of rational numbers, while the numbers e, π, and

$$\theta = \sum_{n=1}^{\infty} 10^{-n!}$$

are transcendental over the field of rational numbers. The proofs are not obvious and use analysis. It is an important and difficult problem in number theory to determine if a given real number is algebraic or transcendental over the field of rational numbers. If an element η of K is transcendental over the field F, then the smallest field containing F and η is isomorphic to the field of rational functions in one variable over F; that is, the quotient field of the ring of polynomials $F[t]$ with coefficients in F. If all the elements of some extension K containing the field $\mathbf{Q}$ of rational numbers has all its elements algebraic over $\mathbf{Q}$, then one seeks to study the arithmetic structure of K; for example, to answer whether K has elements which uniquely generate its multiplicative group as the primes generate the multiplicative group of nonzero rationals. *See* CIRCLE; E (MATHEMATICS); NUMBER THEORY; ROOT (MATHEMATICS).

Valuations. An important class of fields are those which possess a valuation. The simplest valuations are functions v from the field to the nonnegative real numbers with the properties: [i] $v(a) = 0$ if and only if $a = 0$, [ii] $v(a + b) \leq v(a) + v(b)$, [iii] $v(ab) = v(a)v(b)$. The ordinary absolute value is a valuation on the fields $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{C}$. The field $\mathbf{Q}$ has other valuations. For example, if p is a prime, write $a/b = p^{\alpha}(c/d)$, where α is an integer and p does not divide cd; then $v_p(a/b) = p^{-\alpha}$ is a valuation on $\mathbf{Q}$ known as the p-adic valuation. When a field has a valuation, one can study its completion in the sense of a corresponding topology. The fields $\mathbf{R}$ and $\mathbf{C}$ are complete under the ordinary absolute value and $\mathbf{R}$ is the completion of $\mathbf{Q}$ under this valuation. The completion of $\mathbf{Q}$ under v_p creates the fields of p-adic numbers, $\mathbf{Q}_p$. The fields $\mathbf{Q}_p$ and their extensions occur in the study of arithmetic properties of fields algebraic over $\mathbf{Q}$. *See* TOPOLOGY.

Automorphisms of fields. A very useful tool in the study of algebraic structures is the group of automorphisms of the structure. An automorphism of a field is an isomorphism of a field onto itself. The study of relations between extensions of a field by means of automorphism groups is known as Galois theory. Important problems can often be formulated as questions about solutions of the polynomial equation, and these in turn can be resolved by considering the automorphism group of the minimal extension containing all solutions of the equation. Among such problems are the question of squaring the circle and ruler and compass constructions. D. J. Lewis

Bibliography. M. Artin, *Algebra*, 1995; L. N. Childs, *A Concrete Introduction to Higher Algebra*, 2d ed., 2000; J. B. Fraleigh, *A First Course in Abstract Algebra*, 6th ed., 1998; I. N. Herstein, *Abstract Algebra*, 3d ed., 1996; N. Jacobson, *Lectures in Abstract Algebra*, vol. 3, 1964, reprint 1981; S. Lang, *Algebra*, 3d ed., 1993; O. Zariski and P. Samuel, *Commutative Algebra*, 1975, reprint 1986, 1991.

Fig

A species of deciduous tree, *Ficus carica*, of the mulberry family (Moracaeae). Leaves are characteristically palmate, with 3 to 7 lobes. A milky latex containing the protein-digesting enyzme ficin is produced. It is of southwest Asian origin and is now cultivated in the subtropical regions of the world for its edible fruit, especially in Asia, Africa, Europe, and the United States.

Flowers and fruit. The fig is a multiple or collective fruit (see **illus.**). It develops from a fleshy compound inflorescence in the form of a hollow structure (syconium) with an apical canal or pore (ostiole) and numerous tiny unisexual flowers lining the inner wall. The flowers are long-style pistillate (female); short-style pistillate (female); and staminate (male). They are significant in the life history of a fig wasp species, *Blastophaga psenes*. The true fruits are tiny, juicy stone fruits (drupelets) that develop from the ovaries of the pistillate flowers. *See* FLOWER; FRUIT.

Tree forms. Two forms of trees characterize *F. carica*: fig and caprifig. The fig bears long-style pistillate flowers exclusively and produces the edible figs of commerce. Depending on cultivar, it may produce two crops annually: a spring crop, commonly called the first or breba crop, and a summer crop, commonly called the main crop or the second crop, even if there is no first crop.

***Ficus carica*, details of branch and fruit.**

The caprifig bears short-style pistillate flowers and staminate flowers which are clustered around the inner opening of the ostiole. It produces three crops annually: an overwinter crop, which is initiated in fall and matures in early spring (the mamme crop); a crop that is initiated in early spring and matures in late spring or early summer (the profichi crop); and a crop that is initiated in summer and matures in fall (the mammoni crop). Caprifigs are the source of pollen for certain cultivars of figs whose flowers must be pollinated to set a crop of fruit.

Cultivars. More than 600 cultivars of figs have been named. Depending on the kinds of crops they bear, they are classified into three horticultural types: common, San Pedro, and Smyrna. Trees of the common type produce good summer crops. Depending on cultivar and climatic conditions, they may or may not produce a spring crop. Caprification is not necessary for setting a crop. Important cultivars, along with names by which they are better known in the trade, are Brunswick (Magnolia), Dottato (Kadota), Franciscana (Mission), Genoa, Malta (Celeste), San Piero (Brown Turkey), and Verdone (Adriatic).

Trees of the San Pedro type generally produce good spring crops and poor summer crops, if any. The latter may or may not need to be caprified to set a crop. Important cultivars of this type are San Pedro, Blanquette, Dauphine, King, and Lamperia. Trees of the Smyrna type seldom bear a spring crop. They set good summer crops, but only if the syconia are caprified, that is, if the pistillate flowers are pollinated with subsequent development of the ovaries into drupelets. Important cultivars are Sari Lop (Lob Injir or Calimyrna) and Kalamata. Fig cultivars are propagated almost entirely by the rooting of cuttings.

Caprification. Caprification is effected by the females of the fig wasp, which colonizes caprifig trees. A wasp develops in the ovule of each pistillate flower in which an egg has been laid. The females become covered with pollen when they crawl through the staminate flowers before leaving the syconium through the ostiole. They enter the fig syconia through the ostiole and pollinate the flowers as they crawl over them while attempting to lay eggs in the ovules. However, they fail in the attempts to oviposit because the styles of the flowers are longer than the ovipositor. This procedure takes place only in late spring and early summer, when the profichi crop of the caprifig trees ripens and the flowers of the fig trees are receptive to pollination.

Harvesting and utilization. Figs that are to be eaten or marketed fresh, canned, candied, pickled, or made into jam or preserve usually are picked from the trees by hand. The cultivars most widely used are Brunswick, Franciscana, King, Malta, San Piero, and Verdone.

The bulk of the world's fig crop either is eaten dry or is marketed as dried figs or as fig paste, which is used for making fig bars and other bakery products. Figs to be used in these ways usually are allowed to dry on the trees and drop to the ground

for harvesting either by hand or by machinery. The cultivars grown most widely for dried figs are the following: in California, Sari Lop, Dottato, Franciscana, and Verdone; in Greece, Kalamata; in Italy, Dottato; in Turkey, Sari Lop. *See* FRUIT, TREE. William B. Storey

Bibliography. I. J. Condit, *The Fig*, 1947; A. J. Downing, *Fruit and Fruit Trees of America*, 1993; J. Galil, Fig biology, *Endeavour*, 1(2):52–56, 1977.

Filbert

A nut from any plant in the genus *Corylus*; also called hazelnut. About 15 species are recognized, distributed widely over the North Temperate Zone and ranging in size from shrubs to tall trees. Filberts in commerce are derived mostly from the European species *C. avellana* and *C. maxima*. Hybrids of these with the American species, *C. americana*, are grown sparingly in the eastern United States.

Filberts belong in the birch family (Betulaceae), which produces the staminate (male) flowers in catkins on the previous season's growth and the pistillate (female) flowers, which form the nuts, at the end of the current season's shoots. The windblown pollen is shed from the catkins in late winter or early spring, when the pistillate flowers are receptive to fertilization. Cross pollination is necessary to produce a crop of nuts. The round or oval nuts, $^1/_2$ to $^3/_4$ in. (1.3 to 1.9 cm) in diameter, are developed in a leaflike husk, from which they fall at maturity.

Filberts are grown widely in Europe and the Mediterranean basin, where much of the crop is consumed locally. Imports to the United States come mostly from Turkey. In the United States filberts are grown in Oregon and Washington.

American filberts are grown as small trees with single trunks. This requires frequent removal of the sucker shoots, which normally grow from the base and produce a bush. In Oregon harvesting is done with machines which rake the nuts into windrows and pick them up from the smooth-rolled ground.

Filberts are sold in shell, or the kernels are roasted and used in confectionery and baked goods. *See* NUT CROP CULTURE. Laurence H. MacDaniels

Film (chemistry)

A material in which one spatial dimension, thickness, is much smaller than the other two. Films can be conveniently classified as those that support themselves and those that exist only as layers on top of a supporting substrate. The latter are known as thin films and have their own specialized science and technology.

Thin films. Thin films, from one to several hundred molecular layers, are generally defined as those that lie on a substrate, either liquid or solid. Thin films may form involuntarily at a surface or they may be created by some external agency.

Involuntary formation. The molecules at the surface of a liquid or solid are subjected to an imbalance of intermolecular attractions. At a liquid surface, this imbalance gives rise to the phenomenon of surface tension and all its related effects known as capillarity. At a solid surface, the lattice is distorted to reduce its free energy, to a degree determined by the polarizability of the atoms or ions and the nature of the contiguous phase. Compounds such as mercury(II) iodide (HgI_2), lead iodide (PbI_2), and molybdenum sulfide (MoS_2), or elemental sulfur develop such low surface energies by this process that they become naturally hydrophobic. Even water-soluble salts such as chromium chloride ($CrCl_3$) may develop a hydrophobic surface. Other substances, composed of less polarizable atoms, continue to reduce their surface energies spontaneously by adsorbing molecules from the atmosphere; this is an example of a thin film formed involuntarily. Another example of the rearrangement of a surface layer is the presence of a thermodynamically stable liquid film on the surface of ice, even at temperatures 5–10°C (9–18°F) below the melting point, forming a transitional zone between the bulk crystalline phase of the ice and the surrounding air. Clearly, this layer, which must be a few hundred molecular diameters in thickness, cannot be the same as ordinary water. *See* ADSORPTION; INTERFACE OF PHASES; SURFACE TENSION.

Solutes that lower the interfacial free energy are said to be surface-active and are known as surfactants. Molecules of such solutes are more concentrated in the surface layer than in the bulk phase. Well-known examples in water are the soaps and detergents. These are examples of spontaneously formed films. The phenomenon is more familiar in liquid solutions but it is also known to occur in solid solutions. *See* DETERGENT; SOAP; SOLID SOLUTION; SURFACTANT.

Formation by external action. Monomolecular films on the surface of water can be made by adding appropriate materials in extremely small quantity to the surface. These materials are inappreciably soluble in water, rendered so by a large hydrocarbon or fluorocarbon functional group on the molecule or by other means; they are, however, able to dissolve in the surface of water by virtue of a hydrogen-bonding group or groups on the molecule. These monolayers can be lifted off the surface of water and deposited onto a solid substrate, such as glass or platinum. By repeated deposition on top of underlying layers, builtup films that are many monolayers thick can be obtained. These are known as Langmuir-Blodgett films; they occur in biological membranes and have applications in microlithography, integrated optics, electrooptics, electronic displays, photovoltaic cells, two-dimensional magnetic arrays, and field-effect devices. *See* MICROLITHOGRAPHY; MONOMOLECULAR FILM.

Films can also be deposited directly onto solid substrates by evaporation of material in a vacuum, by sputtering, by chemical reaction (chemical vapor deposition), by ion plating, or by electroplating. Suitable materials for vacuum deposition are metal alloys or binary inorganic compounds such as oxides, halides, sulfides, selenides, and tellurides. Some polymers are also suitable, but compounds that

decompose on heating are not. This process is used extensively for thin-film metallization of glass and ceramics and in microelectronic circuitry. A modification of the method is atomic layer epitaxy, in which a chemisorbed monolayer is left after less strongly held molecules have been thermally desorbed. The process is repeated to build up successive chemisorbed layers, ultimately producing a monocrystalline deposit. This technique has been used to create a monocrystalline film of diamond on a silicon substrate, a development that may lead to a powerful new generation of computer chips. In sputtering, the material is removed from a source (the cathode), is carried by a plasma, and is deposited on a substrate (the anode). In ion plating, the material is evaporated in the positive glow region of a gas discharge at a pressure of a few pascals. A negative bias of several kilovolts with respect to the plasma is applied to the substrate. *See* CHEMICAL VAPOR DEPOSITION; DIAMOND; ELECTROPLATING OF METALS; PHYSICAL VAPOR DEPOSITION; SPUTTERING; VACUUM METALLURGY.

Thin films play an extensive role in both traditional and emerging technologies. Examples are foams and emulsions, the active layers in semiconductors, the luminescent and protective layers in electroluminescent thin-film displays, and imaging and photoelectric devices. *See* ELECTROLUMINESCENCE; ELECTRONIC DISPLAY; EMULSION; FOAM; SOLAR CELL.

Self-supported films. Films in this category have a nominal thickness not larger than 250 micrometers. Films of greater thickness are classified as sheets or foils.

Films are commonly composed of organic polymers, either thermoplastic resins or cellulose-based materials. The three produced in largest quantity are polyethylene, poly(vinyl chloride), and polystyrene. These are used chiefly in packaging applications. The chief requirements of such films are cost economy, transparency, uniform thickness, tear resistance, absence of pinholes, and inertness to moisture and chemicals. *See* POLYMER.

The lower limit of thickness of a self-supported film is about 25 μm. Gold leaf, which historically was the first such film to be produced, can be brought to a thickness of 0.01 μm, but at this degree of attenuation it is too fragile to exist without a supporting substrate. Sydney Ross

Bibliography. M. Ohring, *The Materials Science of Thin Films*, 2d ed., 2001; K. R. Osborn and W. A. Jenkins, *Plastic Films: Technology and Packaging Applications*, 1992; J. A. Venables, *Introduction to Surface and Thin Film Processes*, 2000.

Filtration

The separation of solid particles from a fluid-solids suspension of which they are a part by passage of most of the fluid through a septum or membrane that retains most of the solids on or within itself. The septum is called a filter medium, and the equipment assembly that holds the medium and provides space for the accumulated solids is called a filter. The fluid may be a gas or a liquid. The solid particles may be coarse or very fine, and their concentration in the suspension may be extremely low (a few parts per million) or quite high (>50%).

The object of filtration may be to purify the fluid by clarification or to recover clean, fluid-free particles, or both. In most filtrations, the solids-fluid separation is not perfect. In general, the closer the approach to perfection, the more costly the filtration; thus the operator of the process cannot justify a more thorough separation than is required.

Gas filtration. It is frequently necessary to remove solids (called dust) from a gas-solids mixture because (1) the dust is a contaminant rendering the gas unsafe or unfit for its intended use; (2) the dust particles will ultimately separate themselves from the suspension and create a nuisance; or (3) the solids are themselves a valuable product that in the course of its manufacture has been mixed with the gas. One method of separating solids from a gas is by gas filtration. Among other methods are centrifugal separation, infringement, scrubbing, and electrostatic precipitation.

Three kinds of gas filters are in common use. Granular-bed separators consist of beds of sand, carbon, or other particles which will trap the solids in a gas suspension that is passed through the bed. The bed depth can range from less than an inch to several feet. Bag filters are bags of woven fabric, felt, or paper through which the gas is forced; the solids are deposited on the wall of the bag. A simple example is the dust collector of a household vacuum cleaner. Air filters are light webs of fibers, often coated with a viscous liquid, through which air containing a low concentration of dust, normally less than 5 grains/1000 ft^3 (approximately 28 m^3), can be passed to cause entrapment of the dust particles. The filter installed in the air-intake lines of most household furnaces is an air filter. Gas filters can be used to remove particles in the size range from 1 nanometer to 100,000 nanometers. *See* DUST AND MIST COLLECTION.

Liquid filtration. Liquid-solids separations are important in the manufacture of chemicals, polymer products, medicines, beverages, and foods; in mineral processing; in water purification; in sewage disposal; in the chemistry laboratory; and in the operation of machines such as internal combustion engines. Filtration is the most versatile and widely used operation for such applications. In some cases, however, gravity or centrifugal sedimentation is preferred to filtration. *See* CENTRIFUGATION; MECHANICAL SEPARATION TECHNIQUES; THICKENING.

Liquid filters are of two major classes: cake filters and clarifying filters. The former are so called because they separate slurries carrying relatively large amounts of solids. They build up on the filter medium as a visible, removable cake which normally is discharged "dry" (that is, as a moist mass), frequently after being washed in the filter. It is on the surface of this cake that filtration takes place after the first layer is formed on the medium. The feed to cake

filters normally contains at least 1% solids. Clarifying filters, on the other hand, normally receive suspensions containing less than 0.1% solids, which they remove by entrapment on or within the filter medium without any visible formation of cake. The solids are normally discharged by backwash or by being discarded with the medium when it is replaced.

Cake filters. There are three classes of cake filters, depending on the driving force producing the separation. In pressure filters, superatmospheric pressure is applied by pump or compressed gas to the feed slurry, and filtrate is discharged at atmospheric pressure, or higher. In vacuum filters, the pressure of the atmosphere is the driving force, a vacuum being applied to the filtrate collection chamber. In centrifugal filters, centrifugal force causes the filtrate to flow through the cake and filter medium.

Cake filters are also classified as batch-operating or continuous. In general, pressure filters are batch or intermittent filters, vacuum filters are continuous machines, and centrifugal filters may be either.

Pressure filters. There are three principal industrial classes of pressure filters: filter presses, horizontal plate filters, and leaf filters. They may be operated at 250 lb/in.^{2} (1724 kilopascals) or higher feed pressure.

Filter presses are the simplest and most versatile pressure filters. One form, the plate-and-frame press, is shown in **Fig. 1**. This filter consists of a number of rectangular or circular vertical plates, alternating with empty frames, assembled on horizontal rails. One fixed head and one movable head hold the plates and frames together. Both sides of each plate and the inner surfaces of the heads constitute the filtration area. The medium consists of filter cloths hung over the plates and covering all filter areas. The plates and frames are compressed by a powerful screw, so that the cloths also act as gaskets to prevent the filtrate from leaking between the plates. Filtrate collects on the plate surfaces and leaves the press through drainage channels; cake collects in the frames and is dumped out after the filtration cycle is finished.

The horizontal plate filter consists of a number of horizontal circular perforated drainage plates operating in parallel and placed one above another within

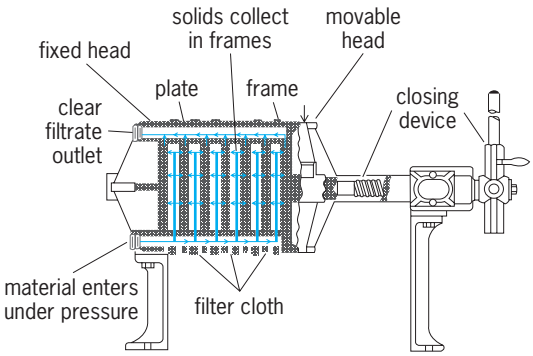

Fig. 1. Corner-feed, closed-discharge filter press, a type of plate-and-frame press. (*T. Shriver and Co.*)

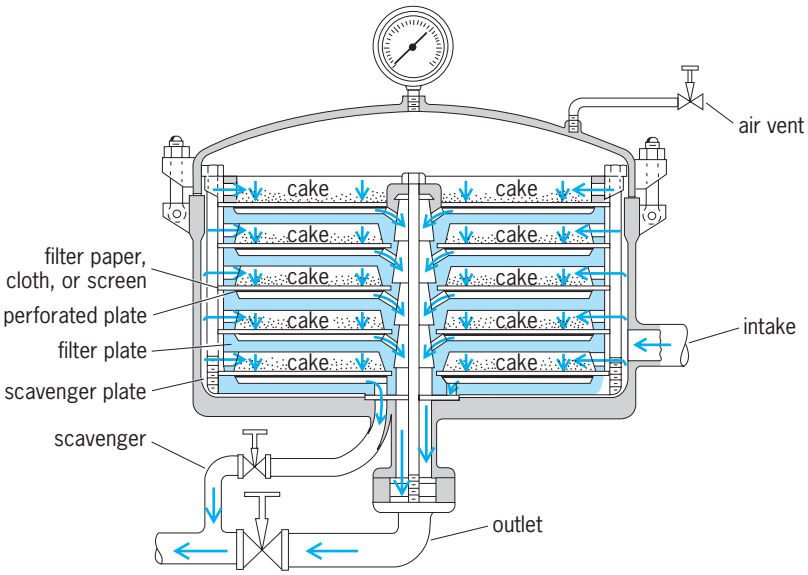

Fig. 2. Sparkler horizontal plate filter. (*After R. H. Perry and D. Green, eds., Perry's Chemical Engineers' Handbook, 6th ed., McGraw-Hill, 1984*)

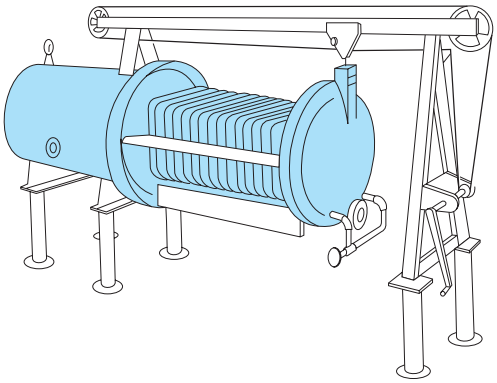

Fig. 3. Horizontal tank pressure leaf filter. (*After R. H. Perry and D. Green, eds., Perry's Chemical Engineers' Handbook, 6th ed., McGraw-Hill, 1984*)

a cylindrical shell (**Fig. 2**). A filter cloth or paper is placed on each plate, which has a drainage chamber beneath. During filtration, the whole shell is filled with slurry under filtering pressure. At the end of the cycle, the shell is drained and filter cake removed.

Pressure leaf filters consist of pressure tanks filled with slurry into which a number of filter leaves connected to drainage pipes are immersed. The leaves may be grids covered with fitted cloth bags, shallow boxes of screen cloth, or porous cylindrical tubes of metal or porcelain. They may be horizontal or vertical. A horizontal tank filter with vertical leaves is shown open for cleaning in **Fig. 3**. When the filter cake can be flushed away instead of recovered dry, a pressure leaf filter can be operated for many cycles over a period of months without being opened. Such filters are adaptable to automatic programmed operation.

Continuous-vacuum filters. These filters are used principally for free-filtering solids. Their most important forms are rotary-drum, rotary-disk, horizontal belt, and horizontal pan or table filters.

The horizontal drum filter is a cloth-covered horizontal cylinder partially immersed in the slurry (**Fig. 4**). The periphery of the drum consists of a number of individual filter elements, each connected to one of the trunnions (drum supports) by pipes. The trunnion contains a special valve which applies suction to each element from the time it enters the slurry until just before it reaches the doctor blade. It then directs compressed air to the element until the element again submerges. Suction is then reestablished. Filtrate flows to the trunnion under the action of the vacuum and then out of the filter, as does wash water sprayed on the cake. The cake is sucked dry of wash and doctored off the filter.

In a disk filter, the rotating drum is replaced by an array of parallel disks, the faces of which are the cloth-covered filtering surface (**Fig. 5**). Each disk is divided into trapezoidal filter elements controlled by a valve, as in the drum filter.

Horizontal filters are useful when heavy solids are to be filtered or when thorough washing is required. They consist either of an endless belt or a rotating annular table, the top surface of each being the filtering surface. The belt passes over a suction box; the table is divided into filter elements connected to a valve, as in the drum filter. Instead of a table being used, the filter elements can be separated into individual trays, annularly arrayed, that can be tipped to assist in the cake dumping.

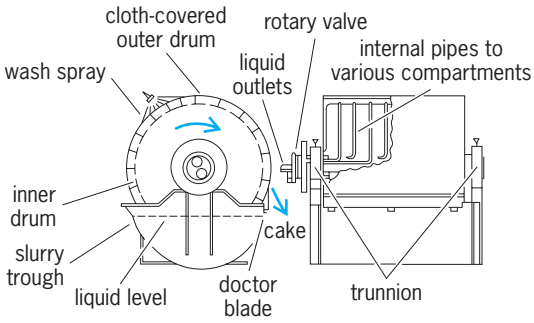

Fig. 4. Continuous vacuum rotary filter. (*After W. L. McCabe and J. C. Smith, Unit Operations of Chemical Engineering, McGraw-Hill, 3d ed., 1976*)

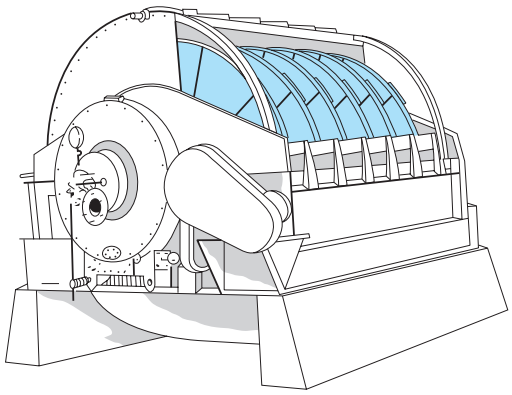

Fig. 5. American disk filter. (*After R. H. Perry and D. Green, eds., Perry's Chemical Engineers' Handbook, 6th ed., McGraw-Hill, 1984*)

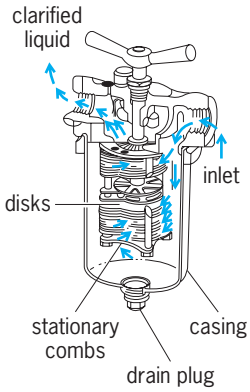

Fig. 6. Cartridge filter. (*After W. L. McCabe and J. C. Smith, Unit Operations of Chemical Engineering, McGraw-Hill, 3d ed., 1976*)

Clarifying filters. When the object of filtration is to produce a solids-free liquid from a dilute suspension, a clarifying filter may be used. Alternatively, sedimentary clarifiers or centrifuges may be chosen. *See* CLARIFICATION; SEDIMENTATION (INDUSTRY).

Clarifying filters are pressure filters that employ a very fine filter medium, the pores of which are sufficiently small to prevent passage of particles of the size that must be removed. In cake filtration, cake particles bridge relatively large openings in the medium and create a new filtering surface; in clarifying filtration, each particle is collected on or within the filter medium.

Clarifying filters may be in the form of plate filters (analogous to shallow-frame filter presses), disk filters (analogous to horizontal plate filters), cartridges (analogous to leaf filters), and granular beds. The last-mentioned type is a special case; the principal example of the use of granular beds is in water filtration, commonly carried out by sand or coal filters operated by gravity instead of under pressure. A cartridge filter is shown in **Fig. 6**.

Although the filter medium is important in every filtration, it assumes paramount importance in clarifications. Aside from granular beds, there are four general types of clarifying mediums: spaced wires that strain out solids (lower limit, 25 micrometers); depth mediums, masses of fibers or interconnected pores that trap particles by interception (lower limit, 5 μm); micronic sheets or pads of carefully sized fibers (lower limit, 0.01 μm); and hyperfiltration membranes (lower limit, 1 nm). A special case of granular-bed clarification occurs when a cake of diatomaceous earth is laid down as a precoat on a coarse filter, thereby converting it to a clarifying filter.

S. A. Miller

Bibliography. N. P. Cheremisinoff and D. S. Azbel, *Liquid Filtration for Process and Pollution Control*, 1989; W. L. McCabe, J. C. Smith, and P. Harriott, *Unit Operations of Chemical Engineering*, 7th ed., 2004; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; A. Rushton, A. S. Ward, and R. G. Holdich,

Solid-Liquid Filtration and Separation Technology, 2d ed., 2000; R. L. Wakeman (ed.), *Progress in Filtration and Separation*, vols. 1–4, 1986–1991.

Fine structure (spectral lines)

A term referring to the closely spaced groups of lines observed in the spectra of the lightest elements, notably hydrogen and helium. The components of any one such group are characterized by identical values of the principal quantum number n, but different values of the azimuthal quantum number l and the angular momentum quantum number j.

According to P. A. M. Dirac's relativistic quantum mechanics, those energy levels of a one-electron atom that have the same n and j coincide exactly, but are displaced from the values predicted by the simple Bohr theory by an amount proportional to the square of the fine-structure constant α. The constant α is dimensionless, and nearly equal to $^1/_{137}$. Its value is actually $0.007297352533 \pm 0.000000000027$. In 1947 deviations from Dirac's theory were found, indicating that the level having $l = 0$ does not coincide with that having $l = 1$, but is shifted appreciably upward. This is the celebrated Lamb shift named for its discoverer, Willis Lamb, Jr. Modern quantum electrodynamics accounts for this shift as being due to the interaction of the electron with the zero-point fluctuations of the electromagnetic field. *See* ATOMIC STRUCTURE AND SPECTRA; QUANTUM ELECTRODYNAMICS; QUANTUM NUMBERS; RELATIVISTIC QUANTUM THEORY.

In atoms having several electrons, this fine structure becomes the multiplet structure resulting from spin-orbit coupling. This gives splittings of the terms and the spectral lines that are "fine" for the lightest elements but that are very large, of the order of an electronvolt, for the heavy elements.

F. A. Jenkins; W. W. Watson

Fingerprint

Distinctive ridges that appear on the bulb of the inside of the end joint of each finger and thumb. These ridges have definite contours and appear in three general pattern types, the arch, the loop, and the whorl, each with general and specific variations of the pattern, dependent upon the shape and relationship of the ridges.

Patterns and Classification

Fingerprints are an infallible means of identification. In addition to their value in the apprehension of criminals, fingerprints can ensure personal identification for humanitarian reasons, such as in cases of amnesia, missing persons, or unknown deceased. Fingerprints are invaluable in effecting identifications in tragedies such as fire, flood, and vehicle crashes. In criminal matters, besides establishing the identity of the arrested person, fingerprint records provide a history of known offenders, or indicate when a person is a first offender. The vast majority of fingerprints maintained in the Criminal Justice Information Services Division of the Federal Bureau of Investigation (FBI) of the United States, the largest repository of fingerprints in the world, are civil records from military and other federal employees.

Patterns. Fingerprints fall into three general types of patterns. Each group bears the same general characteristics or family resemblance. The three general pattern types may be further divided into subgroups by means of smaller differences existing between the patterns in the same general group. The arch group includes the plain arch and the tented arch. The loop group includes the radial and ulnar loops. The whorl group includes four types of whorls, the plain whorl, central pocket loop, double loop, and accidental whorl.

Pattern area. The pattern area is that part of a fingerprint in which appear the cores, deltas, and ridges that are used for classification. The pattern areas of loops and whorls are enclosed by type lines, which may be defined as the two innermost ridges which start parallel, diverge, and surround or tend to surround the pattern area.

Focal points. Within the pattern areas of loops and whorls are the focal points which are used in the detailed classification. These points are called the delta and the core. The delta is that point on a ridge at or in front of and nearest the center of the divergence of the type lines. A core may be defined as that point on a ridge which is located in the approximate center of the finger impression.

Loop. The loop (**Fig. 1**) is that type of pattern in which one or more of the ridges enter on either side of the impression, recurve, touch or pass an imaginary line drawn from the delta to the core, and terminate or tend to terminate on or toward the same side of the impression from which such ridge or ridges entered. In the loop-type pattern the ridges intervening between the delta and core are counted.

The terms "radial" and "ulnar" are derived from the radius and ulna bones of the forearm. The classification of loops is determined by the way the loops flow on the hand. Loops which flow in the direction of the ulna bone, toward the little finger, are called ulnar loops, and those which flow in the direction of the radius bone are called radial loops.

Arches. In plain arches (**Fig. 2***a*) the ridges enter on one side of the impression and flow or tend to

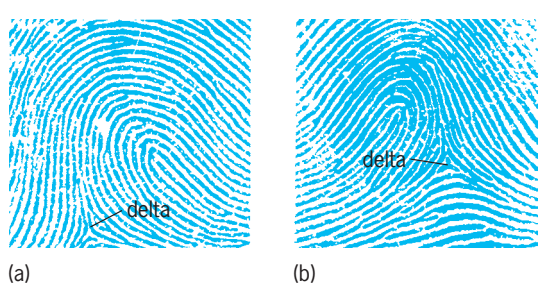

Fig. 1. Loop fingerprint patterns. (*a*) Twelve-count ulnar loop, right hand. (*b*) Radial loop, right hand. (*Federal Bureau of Investigation*)

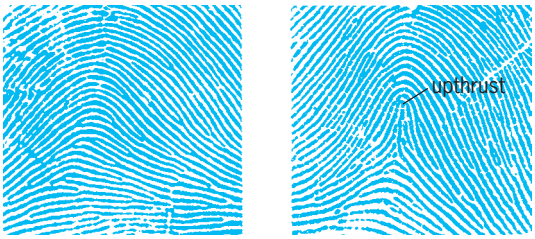

(a) (b)

Fig. 2. Arch fingerprint patterns. (*a*) Plain arch. (*b*) Tented arch. (*Federal Bureau of Investigation*)

flow out the other side with a rise or wave in the center.

A tented arch (Fig. 2*b*) is that type of pattern which possesses either an angle, an upthrust, or two of the three basic characteristics of a loop. That is, the tented arch may lack a delta, ridge count, or recurve, which are characteristics of the loop.

Whorl. The whorl is that type of pattern in which at least two deltas are present with a recurve in front of each. This definition is very general. The pattern may be subdivided into four subgroups for extension purposes in large groups where whorls predominate.

Plain whorl. This pattern is the simplest and commonest of the whorl subdivisions. It has two deltas and at least one ridge making a complete circuit, which may be spiral, oval, circular, or any variant of a circle. An imaginary line drawn between the two deltas must touch or cross at least one of the recurving ridges within the inner pattern area.

Central pocket loop. This pattern is that type of whorl which has two deltas and at least one ridge making a complete circuit. An imaginary line drawn between the two deltas must not touch or cross any of the recurving ridges within the inner pattern area. In lieu of a recurve in front of the delta in the inner pattern area, an obstruction at right angles to the line of flow will suffice. The inner line of flow is determined by drawing an imaginary line between the inner delta and the center of the innermost recurve or looping ridge.

Double loop. This pattern consists of two separate loop formations and two deltas.

Accidental whorl. This is a pattern consisting of a combination of two "different" types of pattern, with the exception of the plain arch, with two or more deltas; or a pattern which possesses some of the requirements for two or more different types; or a pattern which conforms to none of the definitions (**Fig. 3**).

The technique of whorl tracing depends upon the establishment of the deltas. Every whorl has two or more deltas. When the deltas have been located, the ridge emanating from the lower side or point of the extreme left delta is traced until the point nearest or opposite the extreme right delta is reached. The number of ridges intervening between the tracing ridge and the right delta is then counted. If the ridge traced passes inside of (above) the right delta, and three or more ridges intervene between the tracing ridge and the delta, the tracing is designated as "inner" or I. If the ridge traced passes outside (below) the right delta, and three or more ridges intervene between the tracing ridge and the right delta, the tracing is designated as "outer" or O. All other tracings are designated as "meeting" or M. When the ridge traced ends abruptly, the tracing drops to the ridge immediately beneath and continues from there.

Classification. Although the system used in manual (noncomputerized) fingerprint files by the Federal Bureau of Investigation is based on the system devised by Edward Henry, it has been necessary to amplify and extend the Henry system with numerous additional subdivisions. The 10 fingers are considered as a unit in obtaining the complete classification. This classification is expressed in a formula consisting of a combination of letters and numerals. It reflects the general characteristics of the patterns in all fingers and permits filing in proper sequence for reading reference.

The fingerprint pattern types are indicated at the bottom of each finger block. Under the index fingers the appropriate capital letters are placed for every pattern except the ulnar loop, which is designated by a diagonal line slanting in the direction of the loop. Whorls in any finger are designated by the letter W. Under all other fingers, the appropriate small letter or diagonal line is placed for every pattern.

Formula. The classification formula is generally composed of the following divisions: primary, secondary, small-letter group, subsecondary, major, final, and key.

For the purpose of obtaining the primary classification, numerical values are arbitrarily assigned to each of the 10 finger blocks. Fingers numbers 1 and 2 have a value of 16, fingers numbers 3 and 4 have a value of 8, fingers numbers 5 and 6 have a value of 2, and fingers numbers 9 and 10 have a value of 1. Wherever a whorl appears, it assumes the value

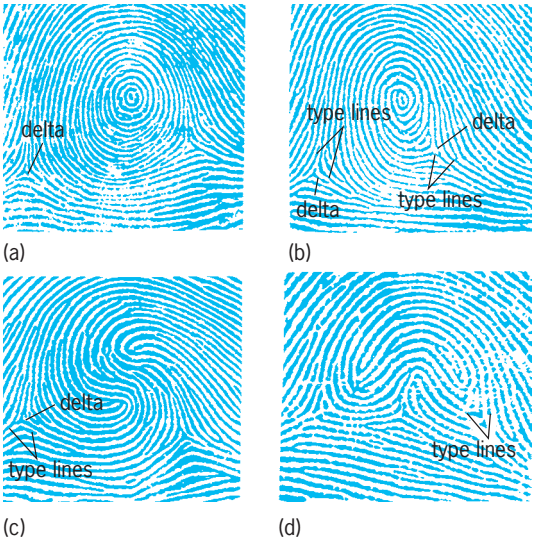

(a) (b) (c) (d)

Fig. 3. Whorl fingerprint patterns. (*a*) Plain whorl, O tracing. (*b*) Central pocket loop, O tracing. (*c*) Double loop, I tracing. (*d*) Accidental whorl, I tracing; this is a whorl over a loop. (*Federal Bureau of Investigation*)

of the space in which it is found. The summation of the numerical values of the whorl type pattern, if any, appearing in fingers 1, 3, 5, 7, 9, plus an arbitrary 1, furnishes the denominator of the primary. The summation of the values of the whorls, if any, in fingers 2, 4, 6, 8, 10, plus an arbitrary 1, furnishes the numerator of the primary. For example, the primary classification on the set of fingerprints (**Fig. 4**) is derived from the location of the whorls in fingers 7, 9, and 10. The numerator is based on the whorl in the tenth finger, which has a value of 1. This plus the 1 arbitrarily added gives the 2 in the formula. The denominator depends upon the values of 2 and 1 given the whorls in fingers 7 and 9. This total of 3 plus the arbitrary 1 yields the 4 that is shown in the denominator.

It should be noted that, after the primary is obtained, the entire remaining portion of the classification formula is based upon the arrangements of the impressions appearing in the right hand as the numerator over the impressions appearing in the left hand as the denominator.

Secondary. The secondary classification appears to the right of the fractional numerals which represent the primary. It is shown in the formula by capital letters representing the types of patterns appearing in the index fingers of each hand.

Small letters. The small-letter group includes the prints with an arch or tented arch in any finger or a radial loop in any except the index fingers. Such small letters are brought up into the classification formula in their proper relative positions, either to the left or right of the secondary. A dash is used to indicate the absence of each small letter between the index fingers and another small letter, or between two small letters.

Subsecondary. The subsecondary classification is obtained by grouping the ridge counts of loops and the ridge tracings of whorls. The ridge counts are translated into small and large, represented by symbols I and O. The whorl tracings are brought up as 1, M, or O, denoting inner, meeting, or outer ridge tracings of the whorl types. Only six fingers are used to obtain the subsecondary, numbers 2, 3, 4, in the

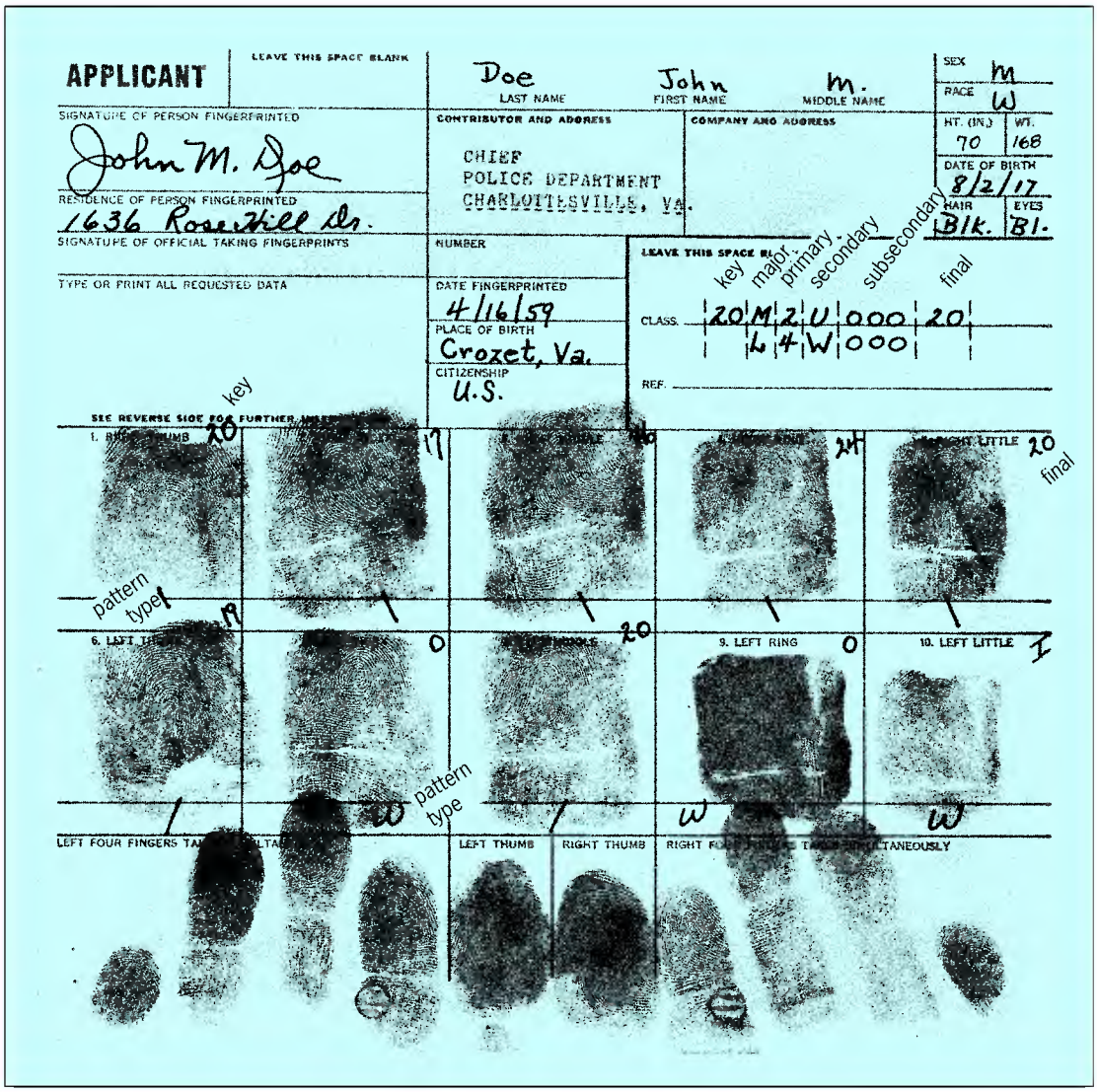

Fig. 4. Set of fingerprints showing classification and pattern types. (*Police Department, Charlottesville, Virginia*)

numerator, and 7, 8, and 9, in the denominator. The subsecondary is placed to the right of the secondary in the classification formula.

Major. The major classification is derived from the thumbs. When whorls appear, the major reflects the whorl tracings. When loops appear, however, a table of ridge count values is used to translate the ridge counts into small, medium, or large groups, designated by the letters S, M, L. The major is placed to the left of the primary in the classification formula.

Final. The final provides a definite order of filing the prints within the aforementioned subdivided groups. It is the ridge count of the loop in the right little finger. If the right little finger is not a loop, the ridge count of the left little finger is used. It is indicated at the extreme right of the classification formula.

Key. The key is obtained by counting the ridges of the first loop appearing on the fingerprint card, beginning with the right thumb, exclusive of the little fingers which are never considered for the key because they are reserved for the final. The key is always placed on the extreme left of the numerator of the classification formula.

File system. The use of the classification formula results in a file system susceptible to multiple expansion and segregates prints into groups on the basis of similar pattern characteristics in corresponding fingers. This enables the location of a record within minutes by examining a limited number of the millions of individual cards on file.

Latent prints. The Latent Fingerprint Section of the Federal Bureau of Investigation deals with the identification of single or latent (hidden) fingerprints developed at the scene of a crime or upon articles of evidence. This generally involves the examination of fragmentary latent finger, palm, or even foot impressions developed by appropriate processes on objects associated with various crimes.

The distinctive ridges of the palms of the hands and the soles of the feet closely resemble those of the fingerprint and can be identified in the same way. Footprints of persons having no fingers are taken for record purposes. Fingerprint, palm print, and sole print identifications have all been accepted as evidence by the courts. *See* EPIDERMAL RIDGES.

J. Edgar Hoover

Automated Fingerprint Identification Systems

An automated fingerprint identification system (AFIS) uses computer technology to collect, store, and identify friction skin ridge impressions from human fingers or palms.

Continuously evolving since the 1880s, fingerprint identification filing systems are used to identify one person's prints among a collection of record (known) fingerprints. Limited success in managing large hand-filed collections was achieved with the constant refinement and increase of classification subgroupings. For example, cards from persons with plain arch patterns on all ten fingers might be filed in a drawer, with further subgroupings according to five-year age groups and whether any of their fingerprint pattern areas include permanent scars.

With the introduction of computers into most large government record systems in the 1960s and 1970s, automated filing began evolving for the task of fingerprint identification in large collections. During the 1980s and 1990s, large fingerprint card repositories throughout most of the world began utilizing AFIS to greatly improve throughput for comparison and identification of fingerprint records. There was also an increase in accuracy for detecting duplicate records (previously undetected aliases). For latent prints collected from crime scenes or other evidence, AFIS enables identifications impossible in large manual (noncomputerized) finger and palm print databases. AFIS also links together unsolved crimes when the same (as yet unknown) person's prints are found. This creates additional leads for investigators. *See* DATABASE MANAGEMENT SYSTEM.

Automated palmprint identification system (APIS). These were introduced in the 1990s and operate similar to AFIS. The individualized nature of friction skin means that any ridge detail on the hands and feet can be encoded, searched, and identified in an automated system. *See* AUTOMATION.

Card reader. Scanners of various configurations are used to input record fingerprint cards into AFIS. Some automatic-feed-type card readers can input over 100 fingerprint cards per hour. Others are simple desktop models similar to those used with personal computers. *See* COMPUTER.

Consolidation. The discovery of two fingerprint records in a repository for the same person using different aliases is called a consolidation.

Encoding. To encode a latent print typically means to electronically mark (**Fig. 5**) the ridge details in preparation for launching a search against the fingerprint repository. Encoding record fingerprint cards often includes entering personal data about the person, verifying computer evaluation of legible friction ridge detail areas of prints, and verifying computerized classification of the pattern types for each finger.

Integrated Automated Fingerprint Identification System (IAFIS). This is the name of the FBI's nationwide AFIS repository located in Clarksburg, West Virginia. IAFIS links all 50 states and all federal agencies.

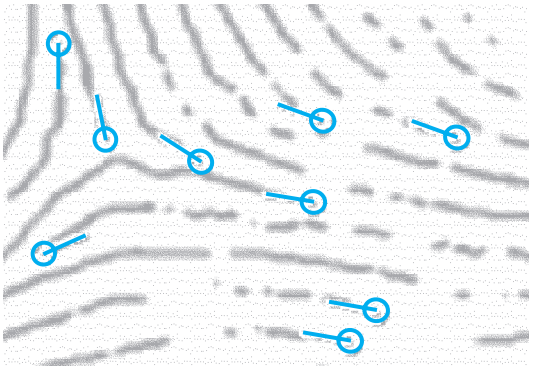

Fig. 5. An AFIS is electronically marked for plotting location (x/y axis) and orientation (rotation).

Latent print searches. All chance or accidental evidentiary finger or palm prints left on an item of evidence or at the scene of a crime are referred to as latent prints. This includes impressions in paint, ink, or blood, as well as those truly invisible before chemical development. About one-third of the latent fingerprints found at most crime scenes are suitable for AFIS searching. The remainder are often from lower finger joints, finger tips, sides of fingers, or other areas not normally found in an AFIS.

Latent workstation. This is usually a personal computer networked with an AFIS, and having latent print capture, encoding, searching, and comparison capabilities. It often includes a digital camera or a flatbed scanner for image capture.

Live scan. These tabletop or stand-alone devices for scanning friction skin ridges capture electronic finger or palm prints by using light or ultrasound.

Record print. All types of known, purposely recorded finger or palm prints are referred to as record prints. Live scan, inkless fingerprint cards, and powder-developed impressions lifted from skin are all record prints.

Matching score. AFIS compares fingerprints using mathematical matching algorithms, resulting in a matching score. For record print comparisons, an identification threshold may be set for the matching score in some systems. All matching scores above the threshold will be positive identifications with zero errors. Matching scores below the threshold may also be identifications as well, but require human analysis, comparison, evaluation, and verification to complete the identification process. Latent prints (**Fig. 6**) normally have insufficient area or clarity to enable automatic identification using a matching score.

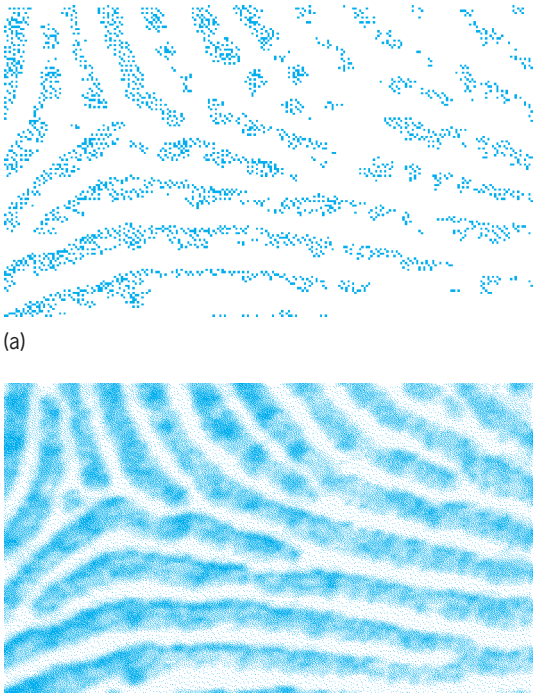

(a)

(b)

Fig. 6. For comparison purposes, (*a*) evidence prints are seldom as legible as (*b*) prints on file.

Point. Ridge endings, bifurcations, and dots are all routinely called points, characteristics, Galton details, minutiae, or features.

Skeleton. Fingerprint images are sometimes thinned or skeletonized during automated feature extraction, enabling easier automated computer detection and encoding.

Smart card. Personal identification cards containing automated fingerprint-feature encoding of the authorized user are one version of a biometric smart card widely used. The user inserts the card into a card reader and places the appropriate finger on a live scan device. A fast and simple one-to-one verification is then completed, using the stored minutiae and the person's finger live scan image. Smart cards are used for automatic teller machine (ATM) cards for banks, access to military restricted areas or restricted computers, and a variety of private and government access control programs. Use of both a fingerprint and a personal identification number (PIN) is required with some smart cards.

Tracing. Tracing ridge detail on translucent plastic over latent print photographic enlargements was required for many early AFIS systems. The resulting ridge detail was similar to the skeleton sometimes produced during AFIS ridge point encoding, and lacked the typical background interference present on evidence. The hand-produced tracings were then reduced to life size and "read" into the AFIS through a camera or scanner. Modern AFIS procedures do not require tracing. However, it is still practiced by many agencies which believe it is faster and more accurate insofar as isolating valid ridge details versus electronically marking points on a computer monitor.

Finger, palm, and sole impression comparison and identification is based on three levels of friction ridge details: Level I—ridge flow or general pattern configuration (**Fig. 7**); Level II—ridge endings, bifurcations, and dots (**Fig. 8**); and Level III—ridge shape, including the shape of ridge edges and pores (**Fig. 9**). Level I information can be captured by sampling images at 150 pixels per inch (ppi). Level II detail needs 1000 ppi or higher resolution for best observation.

Automated fingerprint filing attempts in the early 1960s were computerized search programs which reported the groups of fingerprint cards to be searched based on Level I pattern and other class or group characteristics such as age, sex, race, and geographic location. Some systems, such as Miracode, linked together microfilm rolls of fingerprint images and computer records. These were the first fingerprint comparison workstations. None were efficient, and intensive human evaluation, classification, comparison, and verification were constantly involved.

Through collaborative efforts between the FBI, Rockwell International, and other laboratories in the United States, France, and the United Kingdom, the first AFIS using Level I and II details for searching and comparison was developed between 1967 and 1972. The Level II information recorded included the x and

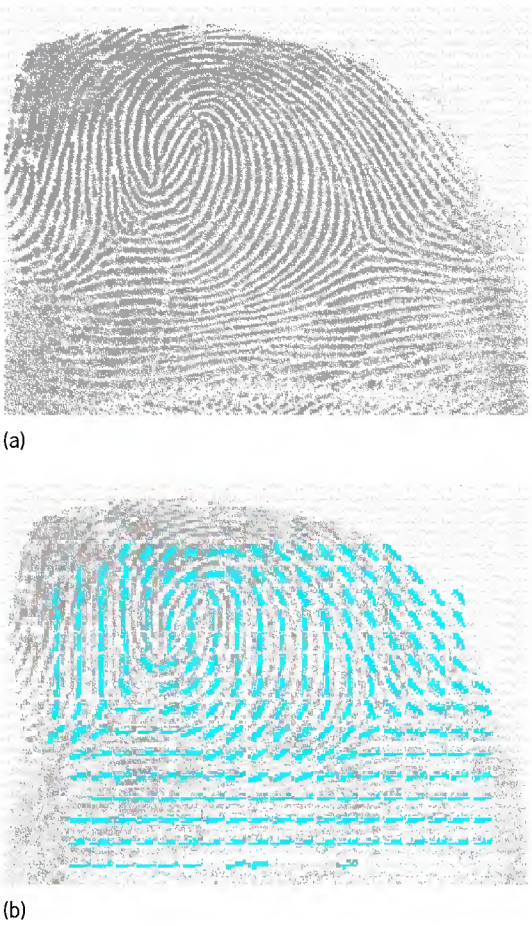

Fig. 7. Typical (*a*) inked fingerprint with (*b*) colored lines on the fingerprint representing Level I detail.

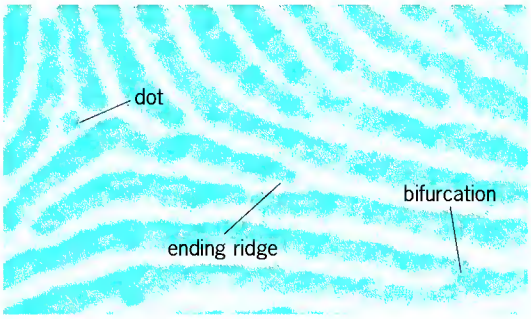

Fig. 8. Enlarged fingerprint showing Level II detail.

y coordinate axis, as well as the rotation or orientation of the ridge ending and bifurcation. Most AFIS ignore dots.

Smaller and more advanced AFIS have been developed in the decades since. Most use not only the x/y axis and rotation, but also intervening ridge counts between individual points and the nearest 3–8 neighboring points. Accuracy for automated comparison between inked prints have been enhanced to the point that in some systems a matching score threshold can be set to enable completely automated comparison and "full" identification with no human intervention and verification and zero error rate.

AFIS operations typically are divided between fingerprint repository (known print) and crime laboratory (questioned print) activities. All record or known prints are first checked by name and date of birth in an attempt to verify identification without wasting AFIS feature computer matching resources. If there is no personal-data-based identification, a technical search using fingerprint minutiae is launched in an attempt to detect existing fingerprint records under an alias name. In a background operating mode, newly arriving fingerprint cards each day are also compared against all unsolved-case latent prints in the AFIS. Crime laboratory experts proactively launch latent print searches against the record fingerprints, and also against other unsolved-case latent prints in an attempt to link related cases.

The National Crime Information Center Year 2000 (NCIC2000) is an FBI service that enables near-instant wanted-person checks using a single fingerprint scanned on a portable live scan unit carried in a patrol car. For example, a person's right index fingerprint can be transmitted electronically to the FBI, searched against a database of 250,000 wanted persons, and an arrest or not-wanted response returned within 2 min. Miniaturization even enables most cities' entire fingerprint repositories to be stored and managed in a personal computer with full AFIS crime-solving capabilities.

Cooperation between police agencies in the 1980s and 1990s resulted in the establishment of limited AFIS definition standards through the American National Standards Institute/National Institute of Standards and Technology (ANSI/NIST). As a result, most record or known fingerprints are captured, stored, and transmitted as 500 or 1000 pixels per inch (ppi) grayscale images using wavelet scalar quantization (WSQ) compression. Most evidence or latent prints are captured as 1000 ppi (or higher) grayscale or color images with no compression; that is, bitmap or tagged image format files are used. Sampling rates of less than 1000 ppi do not capture sufficient Level

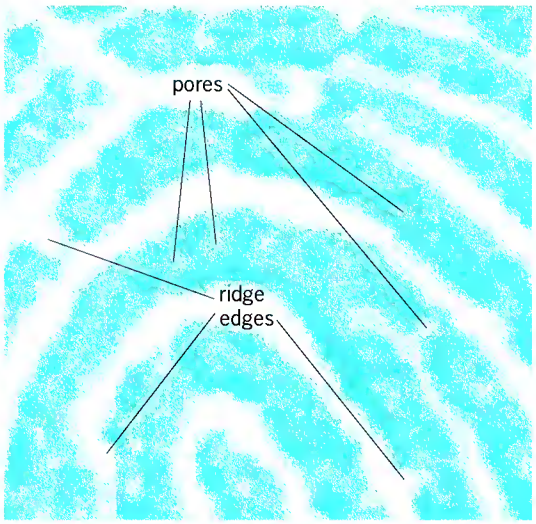

Fig. 9. Enlarged fingerprint showing Level III detail.

III detail to enable meaningful comparison with difficult evidence prints. Some agencies with 500 ppi card scan systems retain the paper fingerprint cards solely to enable latent print comparisons. *See* IMAGE PROCESSING. Edward R. German

Bibliography. D. R. Ashbaugh, *Quantitative-Qualitative Friciton Ridge Analysis*, CRC Press, Boca Raton, FL, 1999; D. Caudle, Personal communication, The Phoenix Group, Pittsburg, KS, 1999; H. Lee and R. Gaensslen, *Advances in Fingerprint Technology*, CRC Press, Boca Raton, FL, 1994; R. Olsen, *Scott's Fingerprint Mechanics*, Thomas Publishing, Springfield, IL, 1978.

Finite element method

A numerical analysis technique for obtaining approximate solutions to many types of engineering problems. The need for numerical methods arises from the fact that for most engineering problems analytical solutions do not exist. While the governing equations and boundary conditions can usually be written for these problems, difficulties introduced by either irregular geometry or other discontinuities render the problems intractable analytically. To obtain a solution, the engineer must make simplifying assumptions reducing the problem to one that can be solved, or a numerical procedure must be used. In an analytic solution, the unknown quantity is given by a mathematical function valid at an infinite number of locations in the region under study, while numerical methods provide approximate values of the unknown quantity only at discrete points in the region. In the finite element method, the region of interest is divided into numerous connected subregions or elements within which approximating functions (usually polynomials) are used to represent the unknown quantity.

Development. The physical concept on which the finite element method is based has its origins in the theory of structures. The idea of building up a structure by fitting together a number of structural elements is quite a natural one (**Fig. 1**) and was used in the truss and framework analysis approaches employed in the design of buildings and bridges in the early 1900s. By knowing the characteristics of individual structural elements and combining them, the governing equations for the entire structure could be obtained. This process produces a set of simultaneous algebraic equations. The limitation on the number of equations that could be solved imposed a severe restriction on the analysis. *See* STRUCTURAL ANALYSIS.

The finite element method evolved into a practical analysis tool as a result of development work in aircraft structural analysis. This was due to demanding analysis requirements coupled with several technological advances in the 1950s and 1960s. The first of these was the formulation of structural problems in terms of matrix algebra, thus forming the governing algebraic equations in a routine manner. Next, the introduction of the digital computer made possible the solution of large-order systems of equations. Finally, the theoretical concepts of finite element analysis crystallized during the same period. *See* DIGITAL COMPUTER; LINEAR ALGEBRA; MATRIX THEORY.

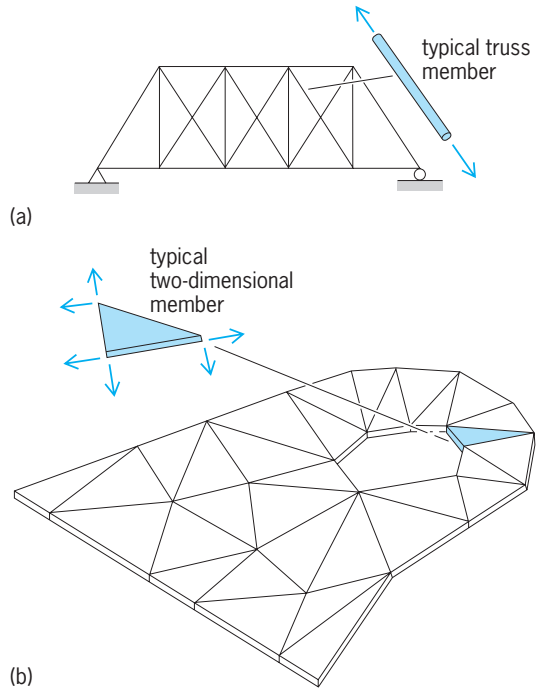

Fig. 1. Structures modeled by fitting together structural elements: (*a*) truss structure; (*b*) two-dimensional planar structure.

Development work since then has caused the finite element method to emerge as one of the most powerful approaches for obtaining approximate solutions to a wide range of problems in mathematical physics. The method has achieved acceptance in nearly every branch of engineering and is the preferred approach in structural mechanics and heat transfer. Its application has extended to soil mechanics, fluid flow, magnetic field calculations, and other areas. *See* HEAT TRANSFER; SOIL MECHANICS.

Mathematical methods. Due to the broadened scope of the finite element approach and the recognition that it is one of the most powerful methods for the approximate solution of boundary value problems, it has attracted the attention of mathematicians who have endeavored to establish a solid mathematical basis for the method. It is, in essence, a variational method of approximation, employing the Rayleigh-Ritz-Galerkin approach. A given region is represented as a collection of a number of geometrically simple regions (finite elements) connected at nodal points. Simple mathematical functions, generally polynomials, are chosen for each element. *See* CALCULUS OF VARIATIONS.

Conventional finite element approaches use Lagrange polynomials as displacement shape functions. An alternate approach, the P-finite element method, uses Legendre polynomials. In the P method, more sophisticated elements are generated

by using high-order Legendre polynomials to produce "elasticity"-type solutions for a wide range of problems. *See* HYPERGEOMETRIC FUNCTIONS; LEGENDRE FUNCTIONS; ORTHOGONAL POLYNOMIALS.

Another advance in finite element technology is the resurgence of integral methods, which has led to the development of the boundary element method into an accepted engineering analysis tool. Only the boundary of the region is discretized into individual elements. For two-dimensional problems the elements are one-dimensional lines on the boundary, while for three-dimensional problems the elements are two-dimensional surface elements. In contrast to conventional finite element approaches, there are no nodes within a region. This has two advantages: (1) Only the boundaries need to be defined, so that less input is required, and (2) a smaller set of equations needs to be solved. Moreover, the boundary element method, like the P method, produces high-accuracy solutions. *See* INTEGRAL EQUATION.

Element formulation and types. The formulation of the individual finite element matrices is central to the method, since the total solution to a problem is constructed of linear combinations of the individual element behavior. In structural mechanics, the unknown field variables of displacements or stresses are defined in terms of values at the node points, which are the unknowns of the problem.

The accuracy of the solution depends not only on the number and size of the finite elements used but also on the functions selected for the elements. These functions must be carefully chosen and must satisfy specific criteria to ensure convergence to the "correct" solution.

Using these functions, a number of approaches can be used to formulate the element matrices (stiffness, stress, induced strain, mass, and so on). The first of these is called the direct approach because physical reasoning is used to establish the element matrices. The other approaches are more theoretical and versatile, allowing the method to be extended to other fields. Here the calculus of variations is used to derive energy principles, which are then employed in deriving the element formulations.

In solid mechanics problems, the element shape functions represent a displacement field within each element (commonly called the displacement method). They could also be chosen to represent stresses (the stress or equilibrium method) or a combination of displacements and stresses (the hybrid method). For most problems in solid mechanics, the displacement method is the easiest to apply, and consequently it is the most widely used approach.

Some of the more commonly used finite elements are depicted in **Fig. 2**. The bar (Fig. 2*a*) is the simplest type of element and can resist only axial loads. These elements can be used to model truss-type structures. The beam element (Fig. 2*b*) is similar to the bar element, but can also take moments and shear loads necessary for representing frame structures. The membrane elements (Fig. 2*c*), grouped together, are basic in finite element analysis of two-dimensional problems. They are thin-membrane elements, triangular and quadrilateral in shape, that are loaded only in their own plane.

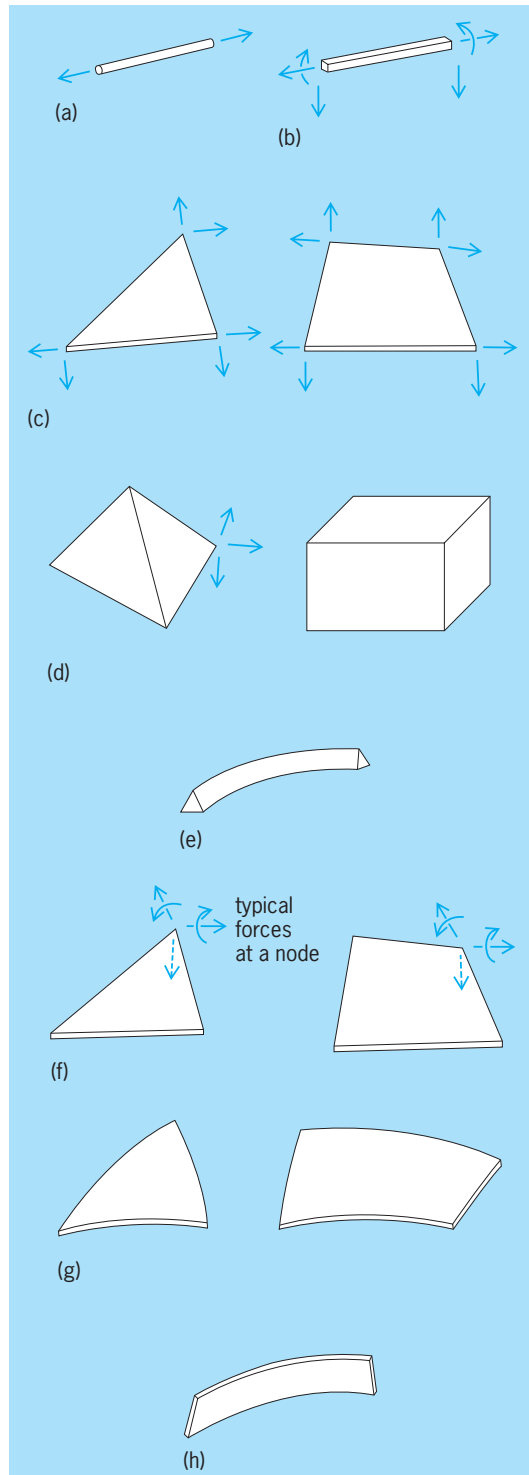

Fig. 2. Element types: (*a*) bar; (*b*) beam; (*c*) membrane elements; (*d*) solid elements; (*e*) axisymmetric solid; (*f*) plate-bending elements; (*g*) curved-shell elements; (*h*) axisymmetric thin-shell element.

Extension of the concept to three dimensions results in a family of solid elements (Fig. 2*d*), of which the tetrahedron and hexahedron are the most

common shapes. These elements are used for the solution of three-dimensional stress-analysis problems. Another type of solid element is the solid of revolution (Fig. 2*e*). The remaining elements in Fig. 2 are classified as bending elements in that they can also take loads normal to their surfaces. The plate-bending elements (Fig. 2*f*) are used to model flat plates and shells or thin-walled members. A more accurate representation of a shell structure can be obtained by using curved shell elements (Fig. 2*g*). Finally, the shell-of-revolution element (Fig. 2*b*) is applicable to a restricted class of thin-shell problems. Elements with additional nodes placed between the corner nodes, using higher-order functions and providing more accurate results, are becoming more accepted.

Application and implementation procedure. The complexity and size of problems which can be addressed by the finite element method has increased. This has been due to advances in element formulation and in solution algorithms, as well as in computer software and hardware technology. The use of the method is illustrated by a representative application taken from the aerospace industry.

Figure 3 shows a finite element model of a business jet which is typical of modeling in aircraft structures. Both the fuselage and wing construction consist of thin metal sheet wrapped around a framework. Bending effects are small enough to be neglected in the skin, which is modeled with membrane elements (Fig. 2*c*). The framework structure (spars and ribs in the wing, frames and longerons in the fuselage) is modeled with bars and beams (Fig. 2*a* and *b*) as well as membrane and bending plate elements (Fig. 2*c* and *f*). The model shown is for a medium-complexity business jet; larger aircraft, such as the Boeing 747, require more complex models. The finite element analysis of the 747 wing-body region alone required the solution of over 7000 equations. Problems requiring the solution of hundreds of thousands of equations are now common. *See* AIRFRAME; FUSELAGE; WING STRUCTURE.

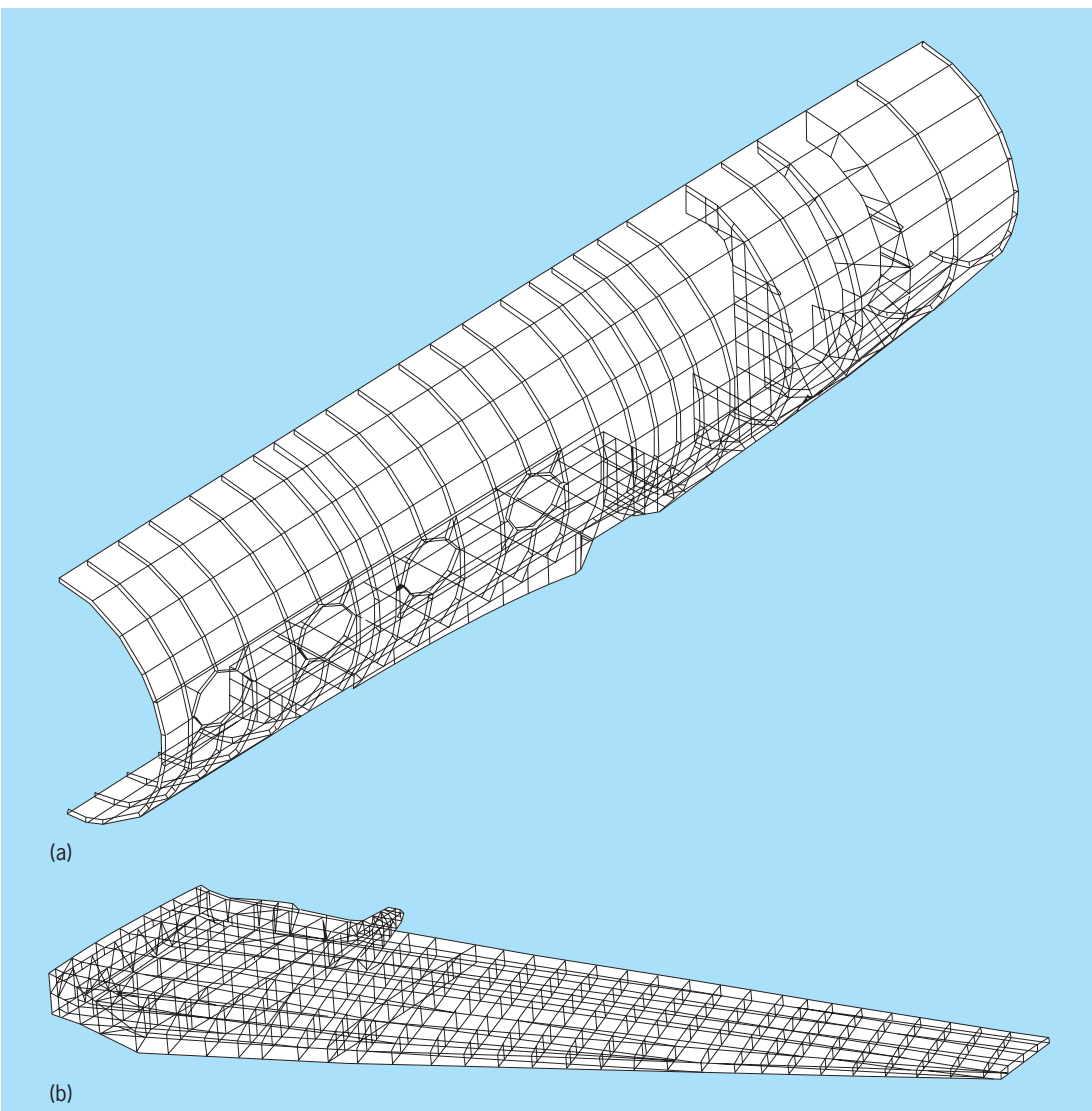

Fig. 3. Finite element method of (*a*) fuselage section and (*b*) wing of a business jet.

While aircraft companies were the first to use the finite element method extensively, today all major designers and manufacturers of automobiles, ships, power plants, bridges, and computers use this analytical tool. The finite element method enables the automobile industry to produce vehicles that are lighter and thus more efficient (using structural optimization techniques) as well as safer (using crash worthiness analysis with the method). In addition to applications in static structural analysis, the method is used to compute vibration modes and perform dynamic analysis as well as to calculate buckling loads for failure analysis.

Finite element analysis is most often implemented with large general-purpose programs capable of handling many types of problems. These programs are very costly to develop and are often key components of computer-based design systems which incorporate interactive graphics and data management systems. The development and maintenance of these programs has shifted from companies that are primarily users today, to companies that specialize in the development of these enhanced products. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

The tasks necessary to perform a finite element analysis are the following:

1. *Model generation*. The first step requires a great deal of modeling expertise. The level of complexity required to meet the intent of the analysis must be established, as well as the types of elements to be employed. Input data describing the coordinates of the grid points and the element topology and properties, as well as boundary conditions and load data, must then be generated. Today these tasks are done automatically in preprocessing modules that are part of the commercial products. In addition, improved interfaces with computer-aided design systems allowing automatic access to wire-frame and solid geometry are moving toward an integrated computer-aided engineering environment. *See* COMPUTER-AIDED ENGINEERING.

2. *Data verification*. The large amounts of data must be checked, since the consequences of error are costly. This is accomplished with graphics modules and special data checking routines as well as error guidelines which are contained in the programs.

3. *Analysis*. The governing matrix equations for the finite element model are generated and solved. For static structural analysis problems, a set of linear algebraic equations are solved, yielding the deflections at the grid points. Substitution into an additional set of equations gives the stresses or forces in the individual elements.

4. *Postprocessing*. The output of the analysis is organized and edited to produce efficient-to-use formats. This includes conversion to report formats, and extensive use of computer-generated graphics showing internal loads and deflections.

The finite element method has found a tremendous number of applications. This is due to its generality in analyzing almost any type of structure without any restrictions on material properties or boundary conditions. In addition, it is readily adaptable to the design process, wherein crude models are employed in the conceptual stages, and more detailed models as the design evolves. Moreover, it is used to reduce the cost of test programs by using analytical models to minimize the number of tests required. Significant advances, including crack propagation capabilities, shape optimization as an outgrowth of structural optimization, the boundary element method, and the P method incorporating automatic error estimation, have further enhanced the scope and the use of the method. Theodore Balderes

Bibliography. K.-J. Bathe, *Finite Element Procedures*, 2d ed., 1995; C. A. Brebbia and J. Dominguez, *Boundary Elements: An Introductory Course*, 2d ed., 1992; R. D. Cook, D. S. Malkus, and M. E. Plesha, *Concepts and Applications of Finite Element Analysis*, 3d ed., 1989; R. H. MacNeal, *Finite Elements: Their Design and Performance*, 1994; Ch. Schwab, *p- and hp-Finite Element Methods: Theory and Applications in Solid and Fluid Mechanics*, 1999; O. C. Zienkiewicz and R. L. Taylor, *The Finite Element Method*, 2 vols., 4th ed., 1994.

Finite mathematics

These parts of mathematics which deal with finite sets. More generally, finite mathematics is taken to include those fields which make no use of the concept of limit; thus linear programming, which finds maxima and minima through the four arithmetic operations, is included, whereas classical optimization, which uses differential calculus, is excluded. *See* LINEAR PROGRAMMING; OPTIMIZATION.

In general, finite mathematics has grown in popularity because it represents, in a sense, the maximum that can be taught without introducing the difficult concepts of limits; thus it can serve as a valuable college course for nonscience majors. There is, however, only the above criterion to unify the field. Brief descriptions will be given of several of the subjects generally included in finite mathematics.

Set theory. An algebra of finite sets can be developed by introducing the operations of union, intersection, and difference. This is necessarily a very elementary set theory, and the usual paradoxes of infinite sets are avoided. *See* BOOLEAN ALGEBRA; SET THEORY.

Logic. As with sets, an algebra of propositions (the propositional calculus) can be developed with the operations of conjunction, disjunction, negation, and implication. There is a close parallelism between this algebra and the set algebra discussed above. Thus if P and Q are the sets of those x's for which the statements $p(x)$ and $q(x)$ respectively are true, then the intersection $P \cap Q$ is the set of x's for which the conjunction $p(x) \wedge q(x)$ is true. *See* LOGIC.

Algebraic structures. Finite groups, rings, and fields are interesting both in themselves and because of their applications. Thus, chemists are interested in finite groups as representing the possible symmetries of a molecular structure, and statisticians might

study finite fields for experimental design purposes. *See* COMBINATORIAL THEORY; CRYSTAL STRUCTURE; GROUP THEORY; RING THEORY.

Linear algebra. Matrices and finite-dimensional vectors represent linear systems; they are thus of great interest, for example, in economics, in which linearity assumptions are at least approximately valid. Thus, if additivity and linear returns to scale are assumed, some economic systems can be represented by a technological matrix $A = (a_{ij})$, where a_{ij} is the amount of good j needed in the production of one unit of good i. The input and output of the system, vectors x and y respectively, are then related by the equation $x = Ay$. *See* LINEAR ALGEBRA; MATRIX THEORY.

Probability. Finite probability spaces and the related random variables are of interest. Most concepts in probability theory—except the limit theorems—can be explained within the context of finite mathematics. In particular, finite Markov chains can be studied by the introduction of matrices. *See* PROBABILITY; STOCHASTIC PROCESS.

Linear programming. An important application of linear algebra arises in economics if it is assumed, as above, that input and output are related by a linear equation, and that profits are also a linear function of output, of the form $P = c^t y$ (where c is also a vector, and the superscript t indicates that it is transposed, to form a row vector). Then, if a maximum input vector b is permitted, the economic problem is to find an output (or activity) vector y which will maximize $c^t y$, subject to the constraints $Ay \leq b$, $y \geq 0$ (because all outputs or activities must be nonnegative). The generic problem of maximizing (or minimizing) a linear function subject to linear (inequality or equation) constraints is known as a linear program. It can be proved, first, that the set of feasible points (those satisfying the constraints) is a polyhedral subset of space, and has therefore a finite number of extreme points (vertices); and second, that the maximum or minimum (if it exists) can always be found at one of these extreme points. Methods then exist that permit a systematic examination of these extreme points to determine the optimum.

Game theory. In the theory of finite two-person games, an $m \times n$ matrix $A = (a_{ij})$ is given. Players one and two choose respectively a row i and a column j of the matrix; player two then pays player one the amount (payoff) a_{ij} which appears in the ith row and jth column of the matrix. An entry a_{kl} is known as a saddle point if it is simultaneously the largest element in its column and the smallest in its row, and it is generally recommended that the players choose the row and column which give the saddle point, if one exists. Unfortunately, games do not generally have saddle points; this has led to the concept of a mixed strategy, which is best defined as a probability distribution over the rows (for player one) or columns (for player two) of the matrix. Mixed strategies take the form of vectors $x = (x_1, \ldots, x_n$ and $y = (y_1, \ldots, y_n)$ respectively, and give rise to an expected payoff equal to the matrix product xAy^t. It can then be shown that, with this payoff, a game always has a saddle point in mixed strategies. Linear programming can be used to compute the strategies (known as optimal strategies) which give rise to the saddle point. *See* GAME THEORY.

Dynamic programming. Certain physical systems can be represented by multistage processes for which, at each stage, the output is a function of that at the preceding stage (the state), an action taken by a decision maker (the control), and possibly a random event. An example occurs in inventory control, in which the inventory of a given item depends on the previous day's inventory (the state), the amount produced (the control), and the external demand (random event). This gives rise to a constraint space of very high dimension. The technique of dynamic programming reduces this sequence to one of one-dimensional problems, which are generally much simpler to solve.

Networks. A network is a collection of points (nodes), together with arcs joining certain pairs of nodes which may represent communications links between various locations, roads joining certain cities, and so on. Numbers representing, for example, the capacity or length of the arcs are frequently assigned. Problems considered may be of a combinatorial nature (finding the number of possible connections between two nodes) or of an optimization nature (finding the shortest path, or maximizing the amounts shipped between two nodes).

Guillermo Owen

Bibliography. S. J. Farlow, *Finite Mathematics and Its Applications*, 2d ed., 1994; J. G. Kemeny et al., *Introduction to Finite Mathematics*, 3d ed., 1974; D. P. Maki and M. Thompson, *Finite Mathematics*, 4th ed., 1995; A. Mizrahi and M. Sullivan, *Finite Mathematics: An Applied Approach*, 7th ed., 1995.

Fiord

A segment of a troughlike glaciated valley partly filled by an arm of the sea. It differs from other glaciated valleys only in the fact of submergence. The floors of many fiords are elongate basins excavated in bedrock, and in consequence are shallower at the fiord mouths than in the inland direction. The seaward rims of such basins represent lessening of glacial erosion at the coastline, where the former glacier ceased to be confined by valley walls and could spread laterally. Some rims are heightened by glacial drift deposited upon them in the form of an end moraine.

Fiords occur conspicuously in British Columbia and southern Alaska (**Figs. 1** and **2**), Greenland, Arctic islands, Norway, Chile, New Zealand, and Antarctica—all of which are areas of rock resistant to erosion, with deep valleys, and with strong former glaciation. Depths of the floors of many fiords exceed 2600 ft (800 m); the deepest on record is 6379 ft (1933 m), in Skelton Inlet, Antarctica. Depths are the result of enhanced glacial erosion caused by the rapid flow due to confinement of the glacier between valley sides.

Fig. 1. Fiord in Alaska. (*W. W. Atwood*, *USGS*)

Fig. 2. Typical fiord-side scenery of southeastern Alaska in June. Steep forested slopes give way upward to rocky ice-scoured upland. (*R. M. Chapman*, *USGS*)

Some shallow fiords were glaciated while their floors stood above sea level, and were submerged later. Others were deepened below the level of the sea at the time of glaciation. A valley glacier flowing into seawater continues to exert stress on its floor until flotation occurs. A vigorous glacier of density 0.9 and thickness of 3300 ft (1000 m) would continue to erode its floor even when submerged to a depth of nearly 3000 ft (900 m). As net submergence of coasts through postglacial rise of sea level is unlikely to have much exceeded 300 ft (100 m), excavation below the sea level of the time probably occurred in most valleys that are now deep fiords. *See* GLACIATED TERRAIN; OCEAN CIRCULATION. R. F. Flint

Bibliography. R. F. Flint, *Glacial and Quaternary Geology*, 1971; H. J. Freedland et al. (eds.), *Fjord Oceanography*, 1980; J. W. Gregory, *The Nature and Origin of Fiords*, 1913; J. P. Syvitski et al., *Fjords*, 1986.

Fir

Any tree of the genus *Abies*, of the pine family, characterized by erect cones, by the absence of resin canals in the wood but with many in the bark, and by flattened needlelike leaves which lack definite stalks. The leaves usually have two white lines on the underside and leave a circular scar when they fall. The native fir of the northeastern United States and adjacent Canada is *A. balsamea*, which attains a height of 75 ft (23 m) and has resinous buds. Its principal uses are for paper pulp, lumber, boxes, and crates, and as a source of the liquid resin called Canada balsam. In the eastern United States the fir is commonly used as a Christmas tree. It does not do well as an ornamental tree in areas where the summers are hot. *See* PINALES.

The Fraser fir (*A. fraseri*) is a similar species found in the southern Appalachians. In this species the bracts grow out beyond the cone scales (see **illus.**).

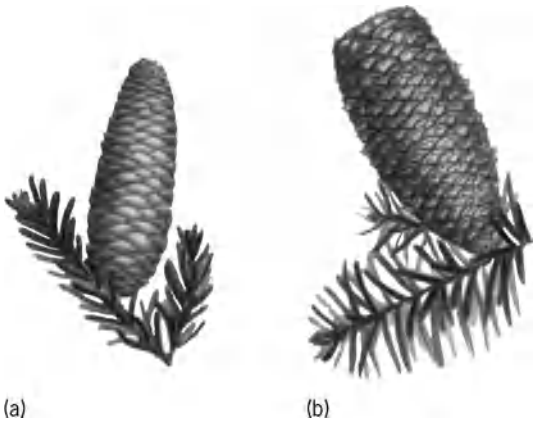

Cones and branches of (*a*) balsam fir (*Abies balsamea*) and (*b*) Fraser fir (*A. fraseri*).

Several species of *Abies* grow in the Rocky Mountains region and westward to the Pacific Coast. The most important is the white fir (*A. concolor*), also known as silver fir, a large tree which may reach a height of 120 ft (36 m) and is characterized by blue-green needles 2–3 in. (5–7.5 cm) long. It is often cultivated as an ornamental in the eastern United States, and also furnishes lumber for construction timber, boxes, crates and general millwork.

Other western species of commercial importance are the subalpine fir (*A. lasiocarpa*), grand fir (*A. grandis*), Pacific silver fir (*A. amabilis*), California red fir (*A. magnifica*), and noble fir (*A. procera*). *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Fire

A rapid but persistent chemical reaction accompanied by the emission of light and heat. The reaction is self-sustaining, unless extinguished, to the extent that it continues until the fuel concentration falls below a minimum value. Most commonly, it results from a rapid exothermic combination with oxygen by some combustible material. Flame and heat also may result from a reaction involving an agent other than oxygen. Thus, certain reactive metals such as zinc will burn in an atmosphere of chlorine.

Flame, the visible manifestation of fire, results from a heating to incandescence of minute particulate matter composed principally of incompletely burned fuel. The color of the flame depends upon the material undergoing reaction and the temperature. *See* COMBUSTION; FLAME; FOSSIL FUEL.

Francis J. Johnston

Fire detector

A device designed to initiate a desired action in response to a change in its surroundings caused by an unwanted fire. Heat detectors respond when the surrounding temperature exceeds a specific minimum value or when the rate of rise exceeds a specified rate. Heat-detector sensing elements are normally mechanical, employing components which melt or expand when heated, but can also be electronic, using elements which produce thermoelectric voltages or change in conductivity.

Smoke detectors are also commonly used fire detectors. They respond when the concentration of solid or liquid particles in the air exceeds a selected value. In ionization-type smoke detectors, a small radioactive source is employed to make the air within the sensing chamber slightly electrically conductive. The presence of particles reduces this conductivity in proportion to the number of particles present. In photoelectric-type smoke detectors, a light source and light-sensitive receiver are arranged so that light from the source does not normally strike the receiver. Smoke particles reflect light from the source onto the receiver in proportion to the number of particles present.

Flame detectors contain receiving elements sensitive to light in certain specific infrared or ultraviolet wavelengths characteristic of flames. Flame detectors can detect a fire within a few milliseconds and are therefore most often used to activate explosion-suppression systems.

The actions initiated by fire detectors include sounding of audible and visual alarms, signal transmission to remote monitoring stations, activation of extinguishing systems, and emergency shutdown of equipment or processes which might increase the severity of the fire hazard. *See* AUTOMATIC SPRINKLER SYSTEM; FIRE EXTINGUISHER.

Richard W. Bukowski

Bibliography. J. L. Bryan, *Fire Suppression and Detection Systems*, 3d ed., 1993; National Fire Protection Association, *Fire Protection Handbook*, 18th ed., 1997.

Fire extinguisher

Fire may be extinguished by the following methods: (1) cooling the burning materials; (2) blanketing the fire with inert gas that chokes it for lack of oxygen; (3) introducing materials that inhibit combustion; and (4) covering the burning matter with a blanket or a layer of solid particles that prevent access of air. Fire extinguishers operate on one or a combination of these principles.

The method or methods that are used to put out a fire depend upon its nature. Burning magnesium reacts with water according to the equation: $Mg + H_2O \rightarrow Mg + H_2$. This reaction is itself a form of combustion which is supported by water instead of oxygen. In fact, the hydrogen evolved will burn in the air and possibly create a worse situation. However, to pursue this example still further, the action of water on a magnesium fire will depend not only on these chemical factors but also on the relation between the amount of water and the amount of burning magnesium (the surface area of magnesium covered by flame). A small amount of water on a large magnesium fire would make matters worse; but a very large amount of water on a small magnesium fire would quench it because of the cooling effect. Therefore, both the chemical and the physical nature of the fuel and the size of the fire have an important bearing on the effectiveness of a given extinguishing technique.

Water. Water is the most effective cooling agent used in fire extinguishing. The generation of steam also drives away the air and forms a blanket, but being less dense than air, it is rapidly displaced.

Wetting agents and foaming agents increase the effectiveness of water. In the case of burning cotton bales, for example, the addition of a wetting agent will cause the water to penetrate the burning material and extinguish the fire more efficiently. Foaming agents make it possible to produce a heavy blanket of foam over the fire. This excludes air and cools the burning material at the same time. Attachments that feed wetting and foaming agents into the fire hose stream are now common.

The use of water in the form of fog rather than as a spray or stream is very effective under some conditions. In fighting an oil fire, for example, fog is effective in cooling the fire and providing a blanket, whereas a stream of water would serve only to disperse the oil and spread the fire. The greatest benefit of fog is that it provides a protecting screen and allows the fire fighter to work in close to the fire.

In the small, first-aid, water fire extinguishers, a propellant must be provided. Usually this is carbon dioxide, which is either generated when needed (see **illus.**) or stored in a cartridge. When a foaming agent is used in such an extinguisher, the foam is rendered still more effective by the inert gas that fills the bubbles.

Antifreeze agents must be added if the water extinguisher is placed outdoors in cold climates. Commonly, calcium chloride is used, but in certain cases (the so-called loaded stream extinguishers), the antifreeze salts may be such as to have a definite fire retarding effect.

There are many fires, as pointed out above, on which water should not be used. On oil fires, for example, water is effective only as a foam or in some cases as a fog; a water stream should never be used. On electrical fires, the use of water can prove extremely dangerous, and in the case of electric generators and motors, it is completely ruinous. Fog nozzles

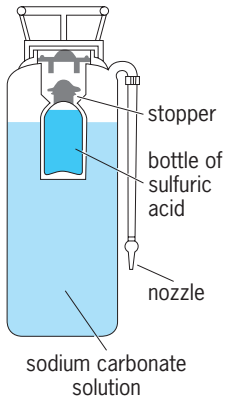

Soda-acid fire extinguisher. When it is inverted, the stopper falls out of place, allowing sulfuric acid to react with sodium carbonate to produce carbon dioxide.

can sometimes be used effectively, but there is always an element of added danger to the fireman whenever water is used near an electrical fire. Carbon dioxide or dry extingishers are decidedly preferred.

Automatic water sprinkler systems are now the most common form of fire protection in industrial plants and large buildings. The installation of these systems is the greatest single factor to be credited for the sharply reduced incidence of disastrous fires. *See* AUTOMATIC SPRINKLER SYSTEM.

Carbon dioxide. A safe and effective extinguisher for all confined fires, carbon dioxide, CO_2, acts as an inert blanket, and because it is heavier than air, it will exclude oxygen very efficiently from a fire on the floor of a building or in a vat or similar vessel. It is not effective in an elevated location or outdoors where the wind can blow the gas away.

Carbon dioxide is supplied as liquid in a tank. The pressure in the tank is about 800–900 lb/in.^{2} (5500–6200 kilopascals); small leaks are therefore serious, and the tank should be weighed frequently to ensure that the contents are still intact. Because of the high pressure, large carbon dioxide containers are not practical. For small, hand-carried extinguishers, the carbon dioxide tank has a short hose with a horn at the end through which the gas may be directed onto the fire. Large installations involve a battery of carbon dioxide cylinders that are attached to a manifold line.

In certain industrial applications, distribution systems are used for carbon dioxide analogous to the sprinkler systems for water. For example, when a fire starts in electric power generating stations, jets are opened to blow CO_2 into the generator housings. This quickly blankets the critical area with the gas. Similar equipment is provided in large spray-painting operations.

In some industrial plants, provision is made to use flue gases to extinguish fires. The flue gas must be cooled and washed before it is put on the fire. Steam from a boiler is similarly used in certain special instances, but its application is definitely limited. *See* CARBON DIOXIDE.

Dry chemical extinguishers. A dry powder, consisting principally of sodium bicarbonate, is also used as a fire extinguisher. The powder must have the correct particle size and contain materials that prevent it from caking. The action of the powder is threefold. Probably its most important action is to generate carbon dioxide by the reaction $2NaHCO_3 \rightarrow Na_2CO_3 + H_2O + CO_2$. Because this reaction is caused by the heat of the fire, the carbon dioxide is generated precisely where it is needed. Secondly, the powder serves to cool the burning material. Thirdly, it provides shielding to prevent access of air. Dry chemical is useful for small fires, and especially electrical fires. As added advantages, sodium bicarbonate is nontoxic, does little or no secondary damage, and is easily cleaned up after the fire.

Halogenated hydrocarbons. Carbon tetrachloride, CCl_4, has had a long history as a fire-extinguishing agent. As it is customarily used, in small quantities, the principal action is to supply a heavy blanket of vapor over the fire. Because it is 3.5 times as dense as carbon dioxide, it is very effective for this purpose.

In addition, carbon tetrachloride, in common with all the halogenated compounds, has a definite chemical inhibiting effect on combustion. Even vapor percentages in the air as low as 4.5% will sometimes put out a fire. Larger percentages are usually needed; but the halogenated materials are more effective by far in suppressing the inflammability of gasoline vapor, for example, than inert gases such as carbon dioxide and nitrogen.

Other halogenated hydrocarbons that have been used for fire extinguishing besides CCl_4 are chlorobromomethane, CH_2ClBr, and several of the fluorinated hydrocarbons known as freons. An effective agent that has been of interest is Freon 13B-1, CF_3Br.

The principal difficulty, however, with all the halogenated hydrocarbons, and with CCl_4 in particular, is toxicity. The vapors of carbon tetrachloride are themselves quite dangerous. They may render a person unconscious and prevent escape from the fire. More serious still are the gases that are produced from CCl_4 by reactions in the fire. In addition to other noxious fumes, small percentages of phosgene, $COCl_2$, may be generated, and this gas is most deadly. Certain of the freons are much less toxic than CCl_4 and less apt to give dangerous products. *See* HALOGENATED HYDROCARBON.

Other extinguishing methods. Small fires are best treated by choking off the supply of air. In addition to the several methods for doing this already described, others should be mentioned that require no special equipment. For a household fire, especially involving clothing on a person, a blanket or a rug provides an effective means to smother the fire. Small fires around a kitchen stove may be snuffed out with salt or, better still, with bicarbonate of soda. A bucket of sand, strategically located, is also useful against domestic fire hazards.

A very novel technique for putting out a fire burning at the liquid surface in an oil tank or vat is to supply agitation by blowing air in at the bottom of the vessel. (A built-in means must be provided for doing this.) As a result of the agitation, the cold oil

from the bottom is brought up to displace the hot burning layer of oil at the top. This method, where applicable, is amazingly effective.

Still other firefighting methods are called for in various special cases. One of these is the explosion of firedamp (methane) in mines. The explosion generates a wind in the mine shaft because the air is pushed ahead of the flame. This accelerates the progress of the explosion and increases the devastation. It has been found that stone dust will stop such fires when it is spread along the mine passages so as to be stirred up readily by the wind. The stone dust probably cools the hot gases to the point where the flame will not propagate.

In the case of forest fires, different techniques, such as back-firing, are used. These depend on wind conditions, moisture, and fuel type. Water suspensions of sodium calcium borate have proved effective in preparing a fire break and even in direct attack on the fire. Borates are well known as fire retardants for paper and fabrics. *See* COMBUSTION; EXPLOSIVE; FIRE DETECTOR; FLAME; FLAMEPROOFING; FOREST FIRE.

William E. Gordon

Bibliography. National Fire Protection Association, *Fire Protection Handbook*, 18th ed., 1997; C. V. Walsh and L. Marks, *Firefighting Strategy and Leadership*, 2d ed., 1976.

Fire technology

The application of results of basic research and of engineering principles to the solution of practical fire protection problems, but entailing, in its own right, research into fire phenomena and fire experience to provide an improved level of understanding of the nature of unwanted fire and its threat to life and property. The defense against destructive fire has been largely empirical and improvisational; although this approach has been reasonably effective in the past, it has not kept pace with urbanized and industrialized life. However, an impressive array of scientific and engineering disciplines has been brought to bear on the complex practical problems of fire defense. This article reviews the current state of fire technology, principally as it applies to urban fires.

Fire Fighting and Protection

The fundamental techniques used by fire fighters consist primarily of putting water on a fire. Water serves to cool a burning material down to a point where it does not produce gases that burn. While water is the most practical and inexpensive extinguishing agent, modern technology has provided not only additives to water to render desirable properties such as easy flow or enhanced sticking, but also chemical and physical extinguishants such as fluorocarbons, surfactant film-forming proteins, and foams. *See* FIRE EXTINGUISHER.

Fog or spray nozzles. Water cools more efficiently if it is broken up into tiny droplets which expose more water surface to heat, hence the increasing popularity of the so-called fog or water-spray nozzles. Wetting agents and other surfactant materials are also employed to make water more efficient.

With the introduction of fog and spray nozzles, there is a tendency to use smaller hose lines, principally the $1^1/_2$-in. (38-mm) booster lines, and smaller quantities of water to fight fires in dwellings, apartments, and other smaller occupancies. Fog nozzles have been used successfully on nearly all types of fires, even oil, tar, grease, and electrical fires where water had previously been considered ineffectual or dangerous. It also has been found that water spray or fog absorbs some of the toxic gases of the fire and tends to clear away the smoke. As a result, the sprays are widely used in fighting smoky interior fires.

Fire fighters generally agree that the way to fight a fire is to get inside the building, find the source, and get water directly on it. However, when alarms are delayed and the fire is fully developed before the arrival of the fire fighters, hose streams may be required to reduce the fire to the point where it can be approached. Among various modern developments in water sprinklers, the most noteworthy is the "on-off" system which turns itself off when the fire is under control and comes on again if the fire rekindles.

Ventilation. An important fire-fighting tactic is that of building ventilation, accomplished by the fire fighters' practice of chopping holes in roofs and breaking windows in portions of the building which do not appear to be involved in the fire. This tactic is used to ventilate the building and clear it of heavy smoke and heated gases so that the building can be entered and the fire reached. Ventilation reduces hazards to life and damage due to smoke and heat. Ventilation should be attempted only by experienced fire officers since improper ventilation can spread the fire and increase the damage and can even introduce the risk of explosion. Fire departments may also use electrically driven fans to help clear buildings of smoke.

Fire Prevention and Loss Reduction

The contribution of the practices of fire prevention is potentially much greater than that of the actual fire-fighting activities. Fire-prevention and loss-reduction measures take many forms, including fire-safe building codes, periodic inspection of premises, fire-detection and automatic fire-suppression systems in industrial and public buildings, the substitution of flame-retardant materials for their more flammable counterparts, insurance to offset the financial loss of fire, and the investigation of fires of suspicious origin, serving to deter the fraudulent and illegal use of fire. *See* AUTOMATIC SPRINKLER SYSTEM; FIRE DETECTOR.

The revision of building codes, the establishment of meaningful fire-resistive tests, and many other activities of fire-protection engineering constitute a field in themselves. Similarly, the subject of fire insurance is highly specialized and complex. There is much interest in improving methods for investigating fires of unknown origin. Fire insurance companies employ or retain investigators to determine the

cause of a large number of fires of suspicious origin each year. In addition, many fire fighters are trained to recognize the clues pointing to the origin of a fire.

Fire Problems of the Future

Fire involves the chemical reaction between oxygen (air) and a fuel which is raised to its ignition temperature by heat. This definition may be depicted by the fire triangle (**Fig. 1**), which shows the three ingredients required for a fire. The reduction or removal of a sufficient amount of any one of these will stop the fire. This concept formed the rudimentary beginnings of the technology of fire control.

Classification of fires. Fires are divided into four classes according to the type of combustible involved.

Class A. Fires involving solids such as wood, paper, and most plastics are termed class A fires and were almost the only type that used to concern people. Water is the most effective, the cheapest, and the most easily applied extinguishing agent for class A fires.

Class B. These are fires of organic liquids such as gasoline, fuel oil, benzene, and acetone. Such fires became important with the advent of the petroleum industry and other modern chemical processing industries. These fires can be controlled by removing the air with a blanketing agent, such as carbon dioxide or a water-based foam. Water in the form of high-pressure fog is an effective extinguisher, but conventional water streams are unsuitable because they spread the fire.

Class C. A class C fire is any fire in which energized electrical equipment is involved. Because of the hazard of electric shock, any extinguishing agent may be used that does not form electrically conductive paths, including high-pressure water fogs. Aqueous foams are not suitable.

Class D. Fires fueled by sodium, magnesium, titanium, and other reactive metals or their reactive compounds, such as sodium hydride, compose this class and are latecomers to the fire scene. These fires are best controlled by removal of air by a blanket of unreactive powder, such as sodium chloride or graphite. Carbon dioxide, sand, and water cannot be used either to blanket or to cool reactive-metal fires because these substances are a source of oxygen for these fuels.

Future hazards. Parodoxically, some of the remedies devised to alleviate one aspect of the fire problem may be adding worse problems in another area. A case in point is the use of fire-retardant formulations in synthetic and natural polymers. Some of these alter the pyrolysis chemistry, producing more toxic effluents. Increased use of these flame retardants may actually increase the life hazards in a fire.

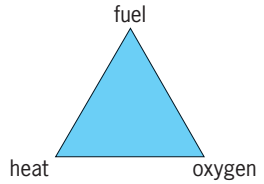

Fig. 1. Triangle of essential ingredients of a fire.

So far, fire has been discussed in terms of the fuel involved without regard to the oxidant because it has been assumed that the oxidant is supplied by air. Fires involve oxidizing agents other than the oxygen in air. Oxidizers such as oxygen, fluorine, and chlorine as liquified gases, and concentrated solutions of hydrogen peroxide find wide industrial use and are transported by highway trucks and railroad cars in growing quantities, entering into the urban scene with increasing regularity. These materials are hazardous because they oxidize, or burn, many common materials of construction, such as wood, paints, plastics, and rubber, and because the temperature required to ignite the combination may be considerably lower than that required for a fire in air alone. Finally, increased use of concentrated oxygen atmospheres and high-pressure, or hyperbaric, atmospheres constitutes a new danger.

Control tactics. The various tactics used in fire prevention and control essentially involve modification of the fire characteristics to alter the time factor to save lives and minimize property damage. Such modifications include preventing fire by reducing possible ignition sources and increasing the ignition time to infinity; early detection of the fire and prompt sounding of the alarm; reducing the fire department arrival time; facilitating quick attack, easy egress, and rescue; decelerating the fire growth; delaying the flashover; and enhancing structural fire endurance. As the National Fire Protection Association (NFPA) points out, fire-safe construction starts on the drafting board, where careful attention can be given to such complexly interrelated factors as the selection of materials, details of architecture, proposed use of the structure, building codes, and legal requirements. All too often, fulfillment of the minimum safety requirements of the codes is no guarantee against a disaster, for what goes into the building changes with time and people.

Fire safety strategies. Of the many specific technological strategies to improve fire safety in an urban structure, the 10 most important are discussed below.

Location and exposure control. This requires placement (zoning) of structures in a community as well as placement of hazardous processes on the periphery of a structure or area. Factors to be taken into account include not only the combustibility of construction materials and the hazards of occupancy, but also the dangers of possible ignition of the building or its contents from exposure to a fire in an adjoining building or other exterior source of ignition.

Occupancy control. This strategy advocates either prohibition or restricted entry of certain materials into a structure with regard to their life hazard and combustibility.

Early detection and alarm systems. The sooner a fire is detected, the easier it is to put out and the less the threat to the occupants.

The five principal types of fire detectors in common use are:

1. *Electronic combustion-gas detector*. This device detects the combustion gases at an early stage of a fire, often long before the smoke would be noticed. Such a detector can usually monitor in excess of 500 ft^2 (50 m^2 or more) of floor area and is preferred in occupancies where life safety is critical.

2. *Radiant-emission detector*. Such a detector responds to infrared or ultraviolet radiation. In some cases, this detector is designed to respond to rapid fluctuations in the radiation, the rate of fluctuations being chosen to correspond approximately to the flicker rate of fires but to exclude the background radiation from other sources. Some versions have a response delay to further screen out background signals. Such a detector can monitor as much as 10,000 ft^2 (1000 m^2) of floor area if its field of view is unobstructed.

3. *Light-scattering (or refracting) detector*. This device appears in various versions, each employing a different principle. One commonly used detector of this type is based on the scattering of light by smoke particles. A new device uses the refraction of a multiply reflected laser beam to detect convection cells produced by a fire. The smoke detectors are relatively slow, and they sample only small volumes of room air, but the laser-beam approach appears to suffer neither disadvantage. *See* LASER.

4. *Fixed-temperature detector*. This detector triggers an alarm upon reaching a preset temperature. Typical versions respond at a temperature level of 158°F (70°C). They are also available in temperatures ranging from as low as 122°F (50°C) to as high as 266°F (130°C). The fixed-temperature detector ordinarily will protect an area of about 100 ft^2 (10 m^2). Relatively sluggish in response, this type of detector is unsuited to life safety and is usually employed in protection of unoccupied property, such as in automatic sprinkler systems for warehouses and the like.

5. *Rate-of-temperature-rise detector*. Such a type responds to a quick rise in room temperature. A typical version responds when the rate of rise exceeds about 0.4°F/s (0.2°C/s); first, however, the temperature must exceed a certain fixed value. A rate-of-rise detector usually monitors about 200 ft^2 (20 m^2) of floor area.

Fire loading control. This strategy could be a part of occupancy control. It involves limiting the maximum amount of combustible material allowed per unit floor area of a structure. Fire loading does not tell much about the ease of ignition and speed of fire growth, but in conjunction with the geometry of the structure, it tells something about the severity of the fully involved fire. The quantity of fuel and its rate of burning determine the duration of a fire.

Regulation of materials. A structure built of steel, concrete, and brick is, of course, safer than one built of wood. When wood is used, application of gypsum or asbestos board improves the safety. The fire properties of building materials fall broadly into two categories: combustibility and fire resistance. Combustibility is further divided into ease of ignition, fire spread, fuel contributed, and smoke developed. The fire resistance of materials and assemblies is expressed in hours of endurance to the exposure from a standard time-temperature furnace.

Compartmentation, subdivision, and ventilation. An effective way of restraining spread of fire through a structure is to suitably seal off each unit of the structure with walls, ceilings, floors, windows, and doors which are impervious to fire. Such a practice is frequently followed in large modern buildings subdivided by fire walls and fire curtains.

In subdivision of buildings, factors must be considered such as the occupant population and the available modes of egress. Pressurized stairwells are under consideration to facilitate safe egress atmospheres. The concept of ventilation is related peripherally to compartmentation and subdivision in that the sealed atmosphere of a modern high-rise building requires, upon detection of a fire, automatic opening of certain vents and closing of others to keep smoke and toxic gases from entering the entire air circulation system.

Structural stability and fire integrity. Since a modern structure is a complex system of many interconnected components, and since a single weak link can defeat the purpose of the entire chain, it is necessary to examine the complex structure for its overall stability and integrity in a potential fire. An otherwise well-designed building and its occupants could, for instance, be placed in jeopardy by improper and inadequate installation of building service equipment.

Accessibility. The need for all-around accessibility of a structure often contradicts the need for security in nonfire situations. Basements, especially, offer some difficulty in accessibility. Exit facilities in accord with the Life Safety Code serve the dual function of providing means of exit from and access to the upper stories of buildings, especially those high-rises that are too tall for fire truck ladders.

Suppression. For most of the problems, advocates in the insurance industry give an all-embracing answer—automatic sprinkler protection. Many commercial and industrial properties are deemed uninsurable unless they are provided with automatically (thermally) actuated water sprinklers. Rigid inspection and maintenance of sprinkler systems are recommended, for insurance company records show that a building with no sprinklers and one with inoperative sprinklers are equally destructible by fire.

Codes, standards, and laws. Codes are, of necessity, generalities, and since fire situations vary over a broad range, literal application of the codes in the design and maintenance of a structure seldom leads to guaranteed safety. Codes and laws are usually based on standards which themselves are deduced from the best available scientific and technological knowledge on fires. Unfortunately, the little available scientific and technological knowledge becomes obscured when filtered through standards to codes. Codes consequently vary widely from place to place in the United States and Canada, but there is a strong movement in the United States to bring

uniformity. True uniformity is possible only when enough factual data are developed from science and technology.

The existing codes are in three categories. Life safety codes primarily concern exits—their adequate number and size and unrestricted approach. Building codes provide rules for safe construction. Most municipalities have their own codes, which are based largely on five model codes: National Building Code (NBC), Uniform Building Code (UBC), Southern Standard Building Code (SSBC), Basic Building Code (BBC), and National Building Code of Canada (NBCC). Finally, fire prevention codes exert an important influence on the occupancy and operation of a building by further refining the building codes.

Chemically active fire-suppression agents. A general approach to fire control has been developed involving use of flame inhibitors. Unlike older fire-extinguishing materials such as water and carbon dioxide, these agents operate indirectly in that they interfere with those reactions within a flame that lead to sustained release of heat. As a result, temperature of the system falls below ignition temperature.

Chemical flame inhibitors offer one of the greatest hopes for better control of fires of all sizes. These agents have achieved such importance that they have, in effect, added a new dimension to the fire triangle, making it a tetrahedron, with one additional mechanism for interfering with the combustion process (**Fig. 2**).

Chemical extinguishers are of two types: liquid (or liquefied gas) and dry powder.

Liquid chemical extinguishers. The most effective liquids are the halogenated hydrocarbons such as chlorobromomethane (CB) and bromotrifluoromethane (better know as Halon 1301) which are colorless, odorless, and electrically nonconductive. *See* HALOGENATED HYDROCARBON.

Automatic Halon 1301 extinguishing systems are ideally suited to situations in which damage of equipment or furnishings by water or dry chemicals would cause irreparable harm. It is used in automatic fire protection systems by insurance companies, banks, universities, and other organizations with computer facilities and valuable, irreplaceable records on data-processing cards or tapes to protect equipment, subfloor electrical wiring, and record libraries.

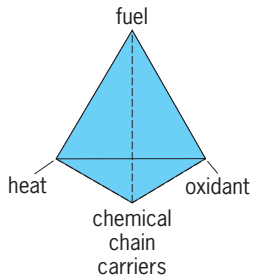

Fig. 2. Chemical fire suppressants add another dimension to the fire triangle, forming a fire tetrahedron.

Tests with Halon 1301 have been successful in hyperbaric (pressurized) atmospheres as well. The atmospheres in hyperbaric chambers are oxygen-rich if not pure oxygen. Materials that are fire-resistant in normal air burn enthusiastically in such oxygen-rich environments, but can be extinguished by Halon 1301 if they can be detected in time.

Dry-powder chemical extinguishers. Ammonium dihydrogen phosphate is the most useful of this type and is rated for class A, B, and C fires. Other dry-powder inhibitors are salts of alkali metals (which include lithium, sodium, potassium, rubidium, and cesium). For example, the dry powders based on sodium bicarbonate and the more effective potassium bicarbonate are rated as extinguishers for class B and C fires. The rubidium and cesium salts have not been studied in detail, and lithium salts have no appreciable activity. Dry chemicals can be effective even when they are applied as a fog in water solutions.

Water. Notwithstanding the success of chemical inhibitors, water is still the prime suppressant because of its great cooling power, general availability and low cost. However, application in a solid stream is regarded as wasteful of water and damaging to the components and contents of structures. Water spray and fog have been highly effective, as mentioned above, even in quenching gasoline fires, for which water is normally unsuitable. Moreover, spray and fog minimize water damage.

Foams. Foams are also widely used. Protein-type, low-expansion foams, particularly useful in quenching burning volatile petroleum products, are used in crash-rescue operations. High-expansion foams are available for fire suppression in enclosed areas. Some foams of this type are generated at a rate of 15,000 ft^3/min (424.8 m^3/min). They can contain sufficient air to allow a human to breathe inside of them.

Light water. A film-forming solution of a specific fluorocarbon surfactant in water, known as light water, was developed by the U.S. Navy for use with dry chemicals to fight aircraft crash fires. It may be used either as a liquid or as a low-expansion foam to interfere with the release of flammable vapors from the burning fuel. Light water is also useful in extinguishing petroleum storage tank fires and may find application to urban fires once the cost is no longer prohibitive. Staney B. Martin; A. Murty Kanury

Bibliography. R. E. Colburn, *Fire Protection and Suppression*, 1975; R. L. P. Custer and R. G. Bright, *Fire Detection: The State of the Art*, NBS Tech. Note 839, 1974; W. M. Haessler, *The Extinguishment of Fire*, NFPA, rev. ed., 1974; H. E. Hickey, *Hydraulics for Fire Protection*, 1980; R. F. Mendes, *Fighting High-Rise Building Fires*, National Fire Protection Association, 1975; National Commission on Fire Prevention and Control, *America Burning*, 1973; National Fire Protection Association, *Fire Protection by Halons*, collection of articles from *Fire Journal* and *Fire Technology*, NFPA SPP-26, 1975; National Fire Protection Association, *Inspection, Testing and Maintenance of Sprinkler Systems*, 1987.

Fire-tube boiler

A steam boiler in which hot gaseous products of combustion pass through tubes surrounded by boiler water. The water and steam in fire-tube boilers are contained within a large-diameter drum or shell, and such units often are referred to as shell-type boilers. Heat from the products of combustion is transferred to the boiler water by tubes or flues of relatively small diameter (about 3–4 in. or 7.5–10 cm) through which the hot gases flow. The tubes are connected to tube sheets at each end of the cylindrical shell and serve as structural reinforcements to support the flat tube sheets against the force of the internal water and steam pressure. Braces or tension rods also are used in those areas of the tube sheets not penetrated by the tubes.

Fire-tube boilers may be designed for vertical, inclined, or horizontal positions. One of the most generally used types is the horizontal-return-tube (HRT) boiler (**Fig. 1**). In the HRT boiler, part of the heat from the combustion gases is transferred directly to the lower portion of the shell. The gases then make a return pass through the horizontal tubes or flues before being passed into the stack.

In the Scotch marine boiler, one or more large flues (approximately 18–24 in. or 45–60 cm in diameter) are used for furnaces or combustion chambers within the shell, and the hot gases are returned through groups of small-diameter tubes or flues. The flues that form the combustion chamber are corrugated to prevent collapse when subjected to the water and steam pressure. These boilers may be oil-fired, or solid fuel can be burned on grates set within the furnace chambers. Scotch marine boilers have, with few exceptions, been superseded by water-tube marine boilers.

Gas-tube boilers are sometimes used for the absorption of waste heat from process gases or the exhaust from internal combustion engines, particularly when their installation provides a simple and economical means for the recovery of low-grade heat. The boiler shell may be augmented by an external steam-separating drum and downcomer and riser connections to provide for proper circulation (**Fig. 2**).

Shell-type boilers are restricted in size to 14 ft (4.3 m) in diameter and an operating pressure of 300 lb/in.^{2} (2.1 megapascals) by the stresses in the large-diameter shells and the necessity to design the flat-tube sheets for practicable thicknesses. They are best suited for low-pressure heating service and for portable use because the integrated structure can be transported easily and the generous reserve in water capacity requires minimum attendance. However, there is the risk of catastrophic damage if overheating and rupture of the shell or tube sheet occur as a result of low water levels or the formation of an insulating layer of internal scale produced by water impurities. *See* BOILER; STEAM-GENERATING UNIT.

George W. Kessler

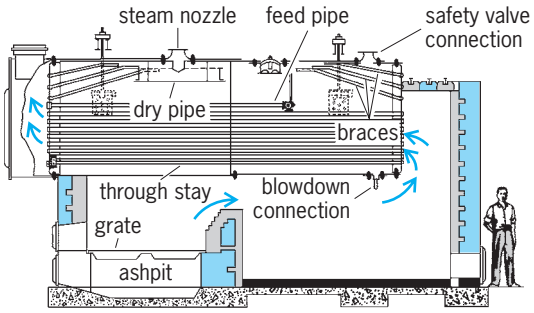

Fig. 1. Horizontal-return-tube boiler.

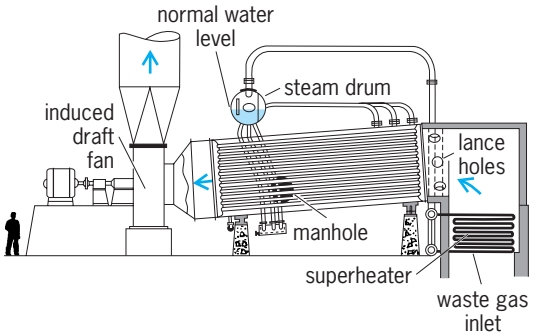

Fig. 2. Fire-tube waste-heat boiler.

Fired heater

A direct-fired heater used in processing petroleum. The basic feature of these heaters is that the fluid (liquid, vapor, or a mixture of both phases) is heated in carbon- or alloy-steel tubes, which in turn receive their heat from hot products of combustion. The fuel may be liquid, such as heavy fuel oil, or gas, such as natural gas or cracked refinery gas. Burners of the steam, gas, or mechanically atomizing type may be used for liquid fuels, and burners of the simpler ring-type, spider, or center-nozzle type for gas fuels. Combination burners designed to handle both fuels are in wide use.

The fired heater is also known as a tube-still heater. Use of "still" in the name comes from the fact that the first continuous heaters were those for heating the topping still, which separated the various fractions of the virgin crude by distillation at atmospheric pressure. Later, to permit distillation of higher-boiling fractions, the distillation was carried out under subatmospheric pressure, whence the name vacuum-pipe still. With the advent of thermal cracking, the tubular heater served the dual purpose of heating the oil to be cracked to the reaction temperature and of maintaining it under the proper temperature, pressure, and time conditions while providing the required endothermic heat of reaction. Today it is the general-purpose tool for supplying heat at the higher temperature levels, that is, over 650–700°F (343–371°C), wherever required in the refinery for distillation, for preheating, for reactions such as pyrolysis, or for reboiling.

Petroleum heaters may heat a single fluid stream, such as the crude charge to a unit. In this case, the

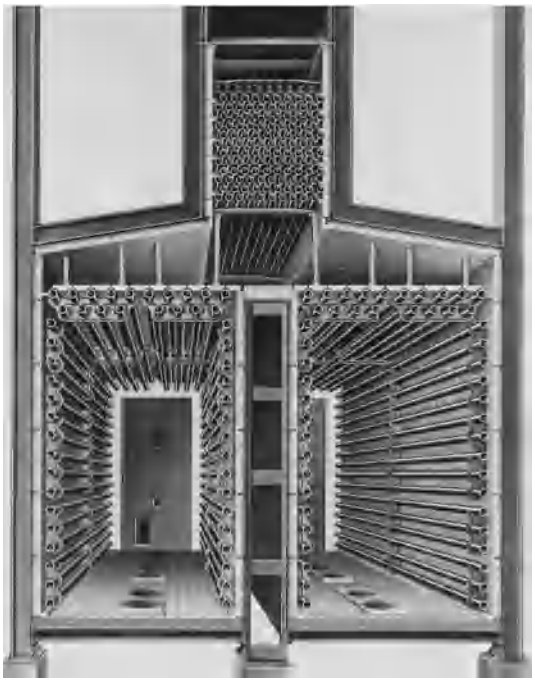

Fig. 1. Double radiant box furnace. (*Lummus Co.*)

tubes are connected in series flow by suitable return bends, which may be provided with cleanout openings. Two or more parallel fluid streams may be heated within the same box in order to keep the friction drop through the heater down to desirable limits. Sometimes streams for different services are heated in the same furnace. However, flexibility today usually dictates that the streams be fired and controlled separately. In certain cases, for example, in steam-methane reformers, where large volumes of gas are passed through catalyst-filled tubes and pressure drop must be minimized and temperatures carefully controlled, all furnace tubes are manifolded in parallel, with all tubes given essentially the same treatment.

Most modern heaters are of one of two types, the box type, usually with horizontal tubes (**Fig. 1**), or the cylindrical type with vertical tubes arranged in a circle (**Fig. 2**). Within the radiant section, the combustion zone in which the burners are located, heat is given up to the tubes from four sources: radiation from the flame, radiation from the hot molecules (carbon dioxide and water vapor), radiation from the gases within the chamber. The convection section is basically a recuperative zone in which the flue gases can be reduced in temperature while heating the coldest fluid to obtain economic heat recovery. Higher overall efficiency may be obtained with air preheaters. In this latter zone, heat is transferred to the air used in combustion by radiation from the hot gases passing through the staggered bank of tubes, by convection from such gases and by radiation from the hot refractory confining the section. The first two rows to receive the hot flue gases from the radiant section often are called the shield or shock tubes, because, as a result of their position, they also receive heat either directly from the combustion box, as do the radiant tubes, or indirectly from large areas of hot radiating refractory.

The design of fired heaters is based as much on experience as on engineering science. The designer must consider not only structural arrangement and heat-transfer rates, but also the effects of fly ash on refractories and metals, strains from heating and cooling, corrosion and erosion, inlet and outlet pressures, methods of operating, and methods of control.

The calculation of pressure drop on the fluid side of the tubes is often complicated because the conditions are nonisothermal, and the relative volumes and compositions of the liquid and gas phases may be changing rapidly.

Tube stills have been designed for operating temperatures up to 1900°F (1037°C), for pressures as high as 4500 lb/in.^{2} (30 megapascals), and as low as 60 mmHg (8 kilopascals). Radiant-section heat-transfer flux may be as high as 30,000–40,000 Btu/(h)(ft^2) [95,000–126,000 W/m^2],

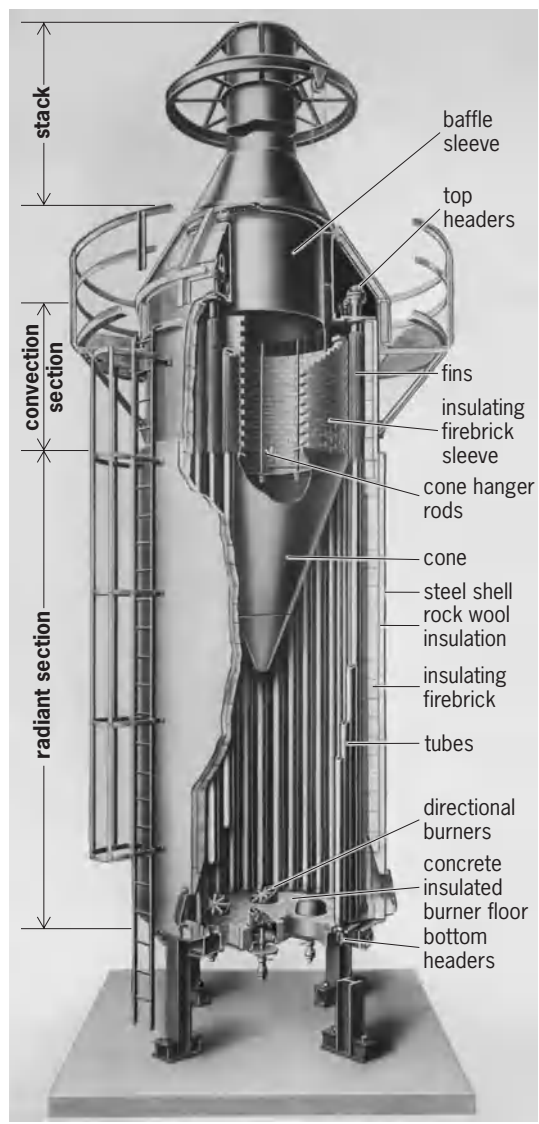

Fig. 2. Cylindrical-type furnace. (*Petro-Chem Development Co.*)

though the more common figure is 12,000–14,000 Btu/(h)(ft^2) [38,000–44,000 W/m^2], for the cleaner, lighter oils and 8000–10,000 Btu/(h)(ft^2) [25,000–32,000 W/m^2] for the heavier, dirtier ones. The severe conditions often encountered in furnace tubes were a prime incentive in the earlier development of strong, high-temperature alloys. *See* FURNACE CONSTRUCTION; HEAT TRANSFER; PETROLEUM PROCESSING AND REFINING. Walter E. Lobo

Fischer-Tropsch process

The synthesis of hydrocarbons and, to a lesser extent, of aliphatic oxygenated compounds by the catalytic hydrogenation of carbon monoxide. The synthesis was discovered in 1923 by F. Fischer and H. Tropsch at the Kaiser Wilhelm Institute for Coal Research in Mülheim, Germany. The reaction is highly exothermic, and the reactor must be designed for adequate heat removal to control the temperature and avoid catalyst deterioration and carbon formation. The sulfur content of the synthesis gas must be extremely low to avoid poisoning the catalyst. The first commercial plant was built in Germany in 1935 with cobalt catalyst, and at the start of World War II there were six plants in Germany producing more than 4,000,000 bbl (6,000,000 m^3) per year of primary products. Iron catalysts later replaced the cobalt.

Following World War II, considerable research was conducted in the United States on the iron catalysts. One commercial plant was erected at Brownsville, Texas, in 1948, which used a fluidized bed of mill scale promoted by potash. Because synthetic oil was not competitive with petroleum, the plant was shut down within a few years. Fischer-Tropsch plants constructed in South Africa (SASOL) have produced gasoline, waxes, and oxygenated aliphatics. The SASOL plants gasify inexpensive coal in Lurgi generators at elevated pressure. After purification, the gas is sent to entrained and fixed-bed synthesis reactors containing iron catalysts. *See* COAL GASIFICATION; COAL LIQUEFACTION.

Methanation. In the years since 1960, interest has grown in the United States in catalytic methanation to produce high-Btu gas from coal. Synthesis gas containing three volumes of hydrogen to one volume of carbon monoxide is reacted principally to methane and water, as shown in reactions (1). Nickel catalysts,

$$\begin{aligned} (2n+1)H_2 + nCO &\xrightarrow{\text{Co catalysts}} C_nH_{2n+2} + nH_2O \\ (n+1)H_2 + 2nCO &\xrightarrow{\text{Fe catalysts}} C_nH_{2n+2} + nCO_2 \end{aligned} \quad (1)$$

first discovered by P. Sabatier and J. B. Senderens in 1902, are important catalysts for methanation. Active precipitated catalysts and Raney nickel catalysts were developed, and the Bureau of Mines developed a technique for applying a thin layer of Raney nickel on plates and tubes by flame-spraying the powder. This led to development of efficient gas recycle and tube-wall reactors.

Reactions. Typical Fischer-Tropsch reactions for the synthesis of paraffins, olefins, and alcohols are reactions (1), (2), and (3), respectively.

$$\begin{aligned} 2nH_2 + nCO &\xrightarrow{\text{Co catalysts}} C_nH_{2n} + nH_2O \\ nH_2 + 2nCO &\xrightarrow{\text{Fe catalysts}} C_nH_{2n} + nCO_2 \end{aligned} \quad (2)$$

$$\begin{aligned} 2nH_2 + nCO &\xrightarrow{\text{Co catalysts}} C_nH_{2n+1}OH + (n-1)H_2O \\ (n+1)H_2 + (2n-1)CO &\xrightarrow{\text{Fe catalysts}} C_nH_{2n+1}OH + (n-1)CO_2 \end{aligned} \quad (3)$$

The primary reaction on both the cobalt and the iron catalysts yields steam, which reacts further on iron catalysts with carbon monoxide to give hydrogen and carbon dioxide. On cobalt catalysts, at synthesis temperatures (about 392°F or 200°C), the reaction $H_2O + CO = CO_2 + H_2$ is much slower than on iron catalysts at synthesis temperatures (482–617°F, or 250–320°C). All these synthesis reactions are exothermic, yielding 37–51 kilocalories/mol (160–210 kilojoules/mol) of carbon in the products or 4700–6100 Btu/lb of product (11,000–14,000 kilojoules/ kg).

The hydrocarbons formed in the presence of iron catalysts contain more olefins than those formed in cobalt catalyst systems. The products from both catalysts are largely straight-chain aliphatics; branching is about 10% for C_4, 19% for C_5, 21% for C_6, and 34% for C_7. Aromatics appear in small amounts in the C_7 and in larger amounts in the higher boiling fractions. Operating conditions and special catalysts required, such as nitrides and carbonitrides of iron, have been ascertained for the production of higher proportions of alcohols. Joseph H. Field

Bibliography. R. B. Anderson, *The Fischer-Tropsch Synthesis*, 1984; *Kirk-Othmer Encyclopedia of Chemical Technology*, 3d ed., vol. 11, 1998; H. H. Lowry (ed.), *Chemistry of Coal Utilization*, vols. 1 and 2, 1945, suppl. vol., 1963.

Fisher

A medium-sized, weasellike, carnivorous mammal classified in the family Mustelidae along with weasels, mink, martens, wolverines, otters, and badgers. Fishers (*Martes pennanti*) occur only in North America. They are found in forested areas of northern North America and southward in the Sierra Nevada Mountains to central California; and southward in the Appalachian Mountains to West Virginia. Their original range, which has been significantly reduced, particularly in the United States, nearly coincided with the combined distributions of northern hemlock, northern hardwood, western mountain, and boreal forests. The fisher is known sometimes as the Pekan, Pennant's marten, black cat, or American sable, although fisher appears to be the most popular name. Oddly enough, the animal does not fish.

Fisher, *Martes pennanti.* (*From Living with Fisher in Massachusetts, © Mass Wildlife, www.mass.gov./dfwele/dfw/dfw_fisher.htm*)

General morphology. The fisher has a long body with short legs, small rounded ears, and a long, bushy tail. The head is wedge-shaped and tapers to a narrow black nose. The body, which generally resembles that of a stocky weasel, is set low to the ground. The pelage (coat), which is dark brown to nearly black, is long except on the face. It differs between the sexes and changes with the seasons. The pelage of females is darker and silkier than that of males, whose fur is coarser and more grizzled. During winter, the tail, rump, and legs are glossy black. White-tipped guard hairs give the face, neck, and shoulders a frosted appearance. Molting occurs in late summer and autumn. The feet are large with each foot consisting of five toes with sharp, curved, unsheathed claws. The large feet, which are heavily furred in winter, presumably assist in walking on snow. Locomotion is digitigrade (that is, they walk on the digits). They can swim if necessary. Both sexes possess a pair of lateral anal sacs that contain a brown, musky-smelling, viscous liquid. The dental formula is I 3/3, C 1/1, Pm 4/4, M 1/2 × 2 for a total of 38 teeth. Males are larger, heavier, and more muscular than females. Adult head and body length is 490–639 mm (19 – 25 in.), and tail length is 253–425 mm (10 – 16 in.). Males weigh 2.6–5.5 kg (5.6 – 12 lb), and females weigh 1.3 – 3.2 kg (2.8 – 7 lb). *See* DENTITION.

Habitat. The fisher inhabits large, heavily wooded areas consisting of spruce, fir, or mixed hardwood trees. It also inhabits timbered northern bogs and swamps. It avoids open places with no overhead cover such as fields, roads, burns, and open bogs. Temporary dens for cover and protection may be located under logs, brush piles, or tree roots, in hollow trees or ground burrows, or under the snow. Temporary dens are seldom used for more than 2 or 3 days. Nesting dens occur most often high in hollow trees.

Fishers are shy and elusive animals that are rarely seen, even in areas where they are abundant. They are most active at night, but daytime activity is not uncommon. They are active year-round. They climb well and, although they spend a considerable amount of time foraging and resting in trees, their activity is predominantly terrestrial. They descend trees head first like a squirrel with their hindlimbs rotated outward and caudally. They are largely solitary, except during the breeding season, with members of each sex defending territories only against members of their own sex. Male territories may range from 20 to 80 km^2 (7.7 to 31 mi^2), whereas female territories range from 4 to 35 km^2 (1.5 to 13.5 mi^2). Territory size depends on habitat quality and prey availability.

The fisher is a generalized, opportunistic predator. Mice, squirrels, hares, rabbits, porcupines, birds, and even carrion are the predominant foods, although birds and their eggs, insects, amphibians, reptiles, and various fruits and nuts are consumed when available. The fisher is one of the few predators to successfully prey on porcupines. When a porcupine is found on the ground, the fisher attacks its face, which is not protected by quills. After 30 min or more, it may have bitten the porcupine's face several times, weakening the porcupine sufficiently to turn it over and begin eating from the ventral surface where there are few quills.

Reproduction and development. Although the fisher breeds in March or April, implantation of the zygote is delayed for about 11 months until the following February or March, after which an approximately 30-day active gestation ensues. Thus, total gestation ranges from 327 to 358 days. The average litter of three young is born from February to May, with the female breeding again within 10 days. Thus, except for a few days following parturition, adult females are pregnant almost all the time. Maternal dens are located in a tree cavity high in a large tree and are used for the first 8–10 weeks. Once the kits become mobile, they are moved from the maternal den to one on or below the ground. It is believed that the maternal den functions to protect the helpless young from aggressive male fishers and ground predators. Newborn young are helpless, blind, and sparsely furred (that is, they are altricial). The eyes open at about 53 days, and they venture outside the den for the first time at about 10–12 weeks of age. Weaning occurs at about 15 weeks. The young travel with the female until late summer or early fall, at which time they disperse to begin their solitary lives. Females are sexually mature when a year old, produce their first litters when they are 2 years old, and breed annually thereafter. Fishers have lived 10 years in the wild and over 10 years in captivity. Humans, by hunting and trapping, are the only predator of adult fishers, which are highly valued for their fur. Kits may be preyed upon by hawks, owls, foxes, bobcats, lynxes, and black bears. Overharvesting and habitat destruction have resulted in the fisher's extirpation from much of the southern portion of its original range. Restocking has been attempted in a number of areas with varying amounts of success. *See* CARNIVORA; DENTITION.

Donald W. Linzey

Bibliography. D. W. Linzey, *The Mammals of Virginia*, McDonald & Woodward, Blacksburg, Virginia, 1998; R. A. Powell, *The Fisher: Life History, Ecology, and Behavior*, 2d ed., University of Minnesota Press, Minneapolis, 1982; R. A. Powell, *Martes pennannti*, *Mam. Species*, no. 156, American Society of

Mammalogists, 1981; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Fisheries ecology

The study of ecological processes that affect exploited aquatic organisms, in both marine and freshwater environments. Because this field is primarily motivated by an attempt to harvest populations, special attention has been given to understanding the regulation of aquatic populations by nature and humans. The foundations of fisheries ecology lie in population and community ecology. *See* ECOLOGY.

A growing emphasis on the conservation of aquatic resources and on the development of aquatic culture in captivity is leading to a diversification of the issues examined by fisheries ecologists. Ideas, concepts, and methods from physiology, genetics, molecular biology, and epidemiology are increasingly relevant in fisheries ecology. Also, because of the multiplicity of issues associated with harvesting biological resources, fisheries ecology must go beyond biology and ecology into sociology and economics.

Harvesting. The problem of regulating the exploitation of aquatic organisms in order to ensure sustained harvest lies at the core of fisheries ecology. Commercial fishing in the twentieth century showed that unregulated or inadequately regulated harvest frequently results in overexploitation and collapse of fishery systems. During the two decades following World War II, the harvest of marine resources increased threefold. The increase continued as new resources were found and technology magnified the efficiency of commercial fishery. Examples of overexploitation abound, including the collapse of the Peruvian anchoveta in the early 1970s and the collapse of Atlantic cod stocks in the maritime provinces of Canada in the 1990s. Most experts agree that the harvest of marine resources has peaked and increased yields are likely to come only from fine-tuning of regulations on stocks that are fully exploited. Consequently, regulation is becoming more critical if sustained harvests from world fisheries are to be ensured.

The typical unit at which management efforts are directed is the exploited population, customarily termed a stock in fisheries science. A central aspect of assessing fisheries resources is to identify these stocks and determine how isolated they are from other stocks of the same species. For example, there are 20 major recognized stocks of cod in the North Atlantic Ocean, each on the whole isolated from every other and distinct with regard to several biological characteristics that determine the potential for harvest. The three main population processes that govern the size and productivity of given stock are somatic growth, mortality, and recruitment. (Recruitment is a term that was coined in fisheries science to designate the incorporation of new individuals into the population through birth.) A typical assessment of an exploited stock includes the study of these three processes, as well as some protocol to estimate abundance. From this information, and aided by mathematical models and statistical tools, fisheries biologists produce recommendations on how many individuals to harvest, of what size or age, when, and where.

Most of the natural variability in numbers experienced by exploited marine fish populations comes from interannual variability in egg production and survival at the juvenile stages, the two processes that determine recruitment. The study of environmental effects on recruitment is the subject of the field known as recruitment fisheries oceanography. Through the study of the recruitment process, fisheries oceanographers attempt to understand those factors that control abundance and to develop techniques that can help predict future population sizes. Another purpose of fisheries oceanography is the development of tools and procedures to detect, monitor, and analyze changes in biological oceanic processes as a consequence of large-scale environmental changes.

Whereas many fishery studies are performed at the single-species level, exploited populations interact with other species as prey or predators and compete for food or space. For example, krill in the Antarctic oceans is the main food source for numerous species of fish, seabirds, and marine mammals. Its exploitation can thus have far-reaching effects at the community level. Many fisheries, such as those exploiting groundfish in the North Pacific and North Atlantic, have as a target an assemblage of species, each of which may have a different productivity. A key task in these fisheries is to administer the harvest of the conglomerate of species by relatively nonselective fleets.

The seemingly foolish tendency to manage most fisheries from around the world on the basis of single-species approaches is due more to the technical difficulties related to addressing the complexities of species interactions than to a lack of awareness about their significance. In recent years, greater attention has been given to the development of fisheries assessment methods that consider the multispecies nature of exploited systems.

Unlike terrestrial populations, aquatic populations cannot be assessed by direct counts. Most aquatic resources are surveyed indirectly through expensive sampling protocols or from catch data from fisheries. Knowledge about the state of these populations and about their response to exploitation is hence tainted by significant errors. Moreover, a fishery involves not only the biological system but also a social and economic system, with its own rules and behavior. Economic, political, or social realities often overruled wise plans for biological management, leading to prominent failures in fisheries management. Substantial efforts in modern fisheries studies are devoted to finding management strategies and protocols that take into account the uncertainties and complexities associated with exploited populations. *See* ADAPTIVE MANAGEMENT.

Aquaculture. The total world supply of fish has increased steadily in recent years. Much of this positive trend is attributable to aquaculture, which contributed 18.5% of the total world production of finfish and shellfish in 1995 (or 112.9 million metric tons). Aquaculture is the farming of aquatic organisms in fresh, salt, and brackish waters. It is an activity close to agriculture and animal husbandry in its fundamental objective: the intensive production of organisms in controlled and stable environments. That objective is reached by providing adequate quality and quantity of food, shelter from predators, and protection against disease. Fishes, crustaceans, mollusks, and algae are the favorite cultured aquatic species.

Aquaculture activities can be classified by the extent of human intervention in the process. One extreme of that spectrum is supplementation or stocking, by which large numbers of artificially produced individuals are added to natural systems to increase natural production or to mitigate the effects of harvest and other human-induced impacts. This is a widespread practice with salmonids in fresh water and has been the technique used to introduce them around the world, such as in New Zealand, Australia, and South America.

Ranching consists of producing juvenile stages of aquatic organisms in captivity and releasing them into the wild. They are harvested at an older age after reaching commercial sizes as free-ranging organisms. Salmon ranching is a widespread practice in Europe and North America.

Intensive aquaculture is developed through an end-to-end intervention and accounts for most of the world aquaculture production. Organisms are reared from egg to adult in controlled environments. Prominent examples are carp production in Asia and salmon production in marine net pens in Norway and Chile.

While ecological studies to guide harvesting plans emphasize population and community ecology, support for aquaculture activities comes more from those areas of biology and ecology related to physiology, nutrition, genetics and epidemiology. *See* AQUACULTURE.

Conservation. The early history of modern fishing was characterized by a rather utilitarian urge to ensure an increasing supply of aquatic organisms at all costs. In the last two decades, that view has yielded to a more realistic perspective, increasingly heedful to the natural limits of aquatic resources and to many environmental aspects of harvested systems. These include habitat degradation, overfishing, incidental mortality of nontarget species, and the indirect effects of species removal at the ecosystem level.

Habitat degradation is a major threat for aquatic species, especially in fresh-water and coastal systems where human activities are most intensive. Dramatic examples of habitat alterations are found in rivers fragmented by dams. For instance, the construction of hydroelectric projects is one of the prime factors affecting hundreds of salmon stocks at risk in the Pacific Northwest of North America by impeding migration, by flooding prime fresh-water habitat, and by killing fish passing through turbines. Deforestation and erosion of river banks also have significant effects on the survival and reproduction of fresh-water fish, as does the withdrawal of water from rivers for agriculture and pollution by urban and industrial waste. Some fisheries operating on the continental shelf, such as commercial bottom trawling and dredging, produce a profound physical disturbance on sea-floor communities.

The way in which fisheries are conducted or regulated often has important implications for the conservation of species other than those targeted. Incidental capture of accompanying nontarget species (known as "bycatch") is a significant problem in many fisheries, such as in most shrimp fisheries. In the Gulf of Mexico, the ratio of fish discards to shrimp caught ranges from 3:1 to 20:1. In multispecies fisheries, the imposition of catch limits for some species can promote the discarding of otherwise perfectly good fish. In the United States west coast groundfish fishery, discards in excess of 80% of the catch were observed.

Another serious environmental problem associated with some fishing operations is the incidental take of seabirds and marine mammals. Albatrosses and other seabirds in the North Pacific and in the South Atlantic oceans are attracted to baited hooks in longline fisheries for tuna, swordfish and squid as the gear is deployed. When caught on a hook, the birds drown. The impact of incidental mortality on some of these species is yet unknown, but available data indicate that it might be significant. In other cases, the effects of given techniques were unequivocal and led to sweeping regulations, such as the United States ban on driftnet fishing established in 1992. In a few cases, the problem of incidental catch was corrected by relatively simple technological adjustments in the fishing gear or in the fishing operation. The incidental catch of loggerhead turtles in the Gulf of Mexico shrimp fishery, for example, was significantly curtailed by the development of a device that lets turtles escape from fishing nets while retaining shrimp. Similarly, the incidental mortality of dolphins in tuna purse seiners operating in the Pacific Ocean was greatly reduced by a combination of gear modification and changes in fishing protocols.

The indirect effects of removing key species on aquatic communities are extremely difficult to study and detect. Nonetheless, commercial fisheries are suspected to be major degrading factors in some aquatic ecosystems. For example, the decline experienced by populations of marine mammals and seabirds in the Gulf of Alaska and the Bering Sea coincided with the development of major fishing operations in the region. The limitation of forage fish resulting from removals by fisheries has been proposed as a main cause for this decline.

Aquaculture activities can affect the environment and neighboring communities in different ways. Fish waste and excess food from net pens can significantly affect habitat and bottom communities of lakes and

coastal areas. Another form of pollution from aquaculture is through the supply of antibiotics to hatchery fish in net pens, which can kill naturally occurring bacteria in bottom communities and have a cascading effect on the rest of the community. In some other cases, the infrastructure required for aquaculture has a substantial impact on the environment. Such is the case of shrimp aquaculture in Latin America, where the construction of rearing ponds has led to the loss of thousands of acres of coastal mangroves.

Another undesirable effect of aquaculture is the crossing of escaped hatchery with wild organisms, which can diminish the ability of future hybrid generations to adapt to local wild conditions. Such a situation is found in Norway, the leading producer of cultured salmon in the world. Small escapes from this enormous hatchery production into the wild can dwarf the relatively small wild populations. Not only are the ecological impacts of hatchery fish of concern, but also the danger of introducing diseases that could be deadly for wild fish.

A major problem associated with the development of aquaculture is that it promotes a major trade in exotic species around the world. Species for which culture techniques have been mastered in one corner of the world are being introduced to other regions with the promise to jump-start aquaculture. Often, once they are introduced for aquaculture purposes, individuals of these alien species escape from the confines of culture facilities and establish populations in the wild, with unexpected effects on native communities. A contemporary example is provided by the Atlantic salmon introduced on the Pacific coast of North America, where it has become a favorite species for net pen culture. Atlantic salmon escaping from net pens are regularly found in rivers of the Pacific Northwest, and evidence of successful spawning has been recently gathered. *See* MARINE CONSERVATION; MARINE FISHERIES.

Miguel Pascual; Patricia Dell'Arciprete

Bibliography. L. R. Brown et al., *State of the World 1995, Worldwatch Institute*, Norton, New York, 1995; FAO Inland Water Resources and Aquaculture Service, Fishery Resources Division, Review of the State of the World of Aquaculture, *FAO Fisheries Circ.*, no. 886 FIRI/C886 (Rev. 1), Rome, 1997; FAO Marine Resources Service, Fishery Resources Division, Review of the State of the World Fishery Resources: Marine Fisheries, *FAO Fisheries Circ.*, no. 920 FIRM/C920 (En), Rome, 1997; R. Hilborn and C. Walters, *Quantitative Fisheries Stock Assessment*, Chapman & Hall, London, 1992.

Fissidentales

An order of the true mosses (subclass Bryidae) consisting of a single family, the Fissidentaceae, and about 800 species in the genus *Fissidens* plus a few others distributed in four or five segregate genera. *Fissidens* is the largest genus of mosses. The Fissidentales are unique in leaf form and arrangement.

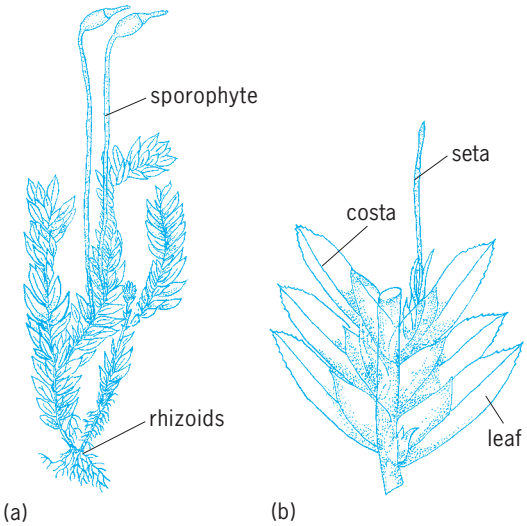

***Fissidens adiantoides*. (*a*) Entire plant. (*b*) Enlarged portion. (*After W. H. Welch, Mosses of Indiana, Indiana Department of Conservation, 1957*)**

The plants, erect and simple or forked, grow (except at the earliest stages) from a two-sided apical cell. The leaves, arranged edgewise to the stem in two rows, have two sheathing blades which clasp the stem and the base of the next leaf above (see **illus.**) Beyond the sheathing blade and toward the stem is a ventral blade; on the side away from the stem is a dorsal blade extending to the leaf base or nearly so. The costa is single and well developed or rarely lacking. The cells are isodiametric and often papillose, though leaf margins are often bordered by linear cells. The archegonia are usually terminal. The capsules are generally exserted, erect, and symmetric, with a peristome that consists of 16 teeth that are characteristically forked and spirally ridged at the tips. The calyptra is cucullate and the chromosome numbers are 5, 6, 8, and 10. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA.

Howard Crum

Fission track dating

A method of dating geological and archeological specimens by counting the radiation-damage tracks produced by spontaneous fission of uranium impurities in minerals and glasses. During fission two fragments of the uranium nucleus fly apart with high energy, traveling a total distance of about 25 micrometers (0.001 in.) and creating a single, narrow but continuous, submicroscopic trail of altered material, where atoms have been ejected from their normal positions. Such a trail, or track, can be revealed by using a chemical reagent to dissolve the altered material, and the trail can then be seen in an ordinary microscope. The holes produced in this way can be enlarged by continued chemical attack until they are visible to the unaided eye. *See* URANIUM.

Theory. Track dating is possible because most natural materials contain some uranium in trace amounts and because the most abundant isotope of uranium, ^{238}U, fissions spontaneously. Over the lifetime of

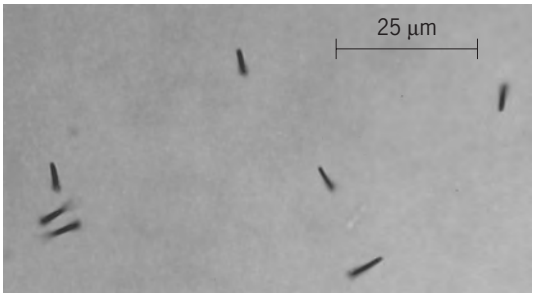

Etched natural fission tracks in muscovite mica crystal. The sample was etched in concentrated hydrofluoric acid.

a rock substantial numbers of fissions occur; their tracks are stored and thus leave a record of the time elapsed since track preservation began. The **illustration** shows etched fission tracks in a mica crystal. The number of tracks produced in a given volume of material depends on the uranium content as well as the age, so that it is necessary to measure the uranium content before an age can be determined. This measurement is most conveniently made by exposing the sample to a known dose of thermal (slow) neutrons, which induce new fissions whose number then gives the uranium content.

The age A of a sample is given by the relation shown in the equation below, where ρ_s is the number of spontaneous fission tracks per unit area, ρ_i is the number induced per unit area, f is the fraction of the uranium caused to fission by the neutron irradiation, and λ_D and λ_F are the total decay rate and the spontaneous-fission decay rate (decays/atom per unit time) for ^{238}U.

$$\frac{\rho_s}{\rho_i} = \frac{[\exp(\lambda_D A) - 1]\lambda_F}{\lambda_D f}$$

Many minerals and glasses are suitable for fission track dating. To be usable, a sample both must retain tracks over geological time and must contain sufficient uranium so that the track density is high enough to be countable in a reasonable period of time. Thus quartz and olivine are highly track retentive but too low in uranium to be usable, whereas calcite may have enough uranium but appears to be inadequately track retentive. Minerals that have been successfully used for fission track dating include the following: zircon, apatite, allanite, epidote, sphene, hornblende, albite, diopside, enstatite, muscovite, phlogopite, and assorted natural glasses.

Applications. One feature unique to this dating technique is the time span to which it is applicable. It ranges from less than 100 years for certain synthetic, decorative glasses to approximately 4,500,000,000 years, the age of the solar system. This great flexibility in age span allows fission track dating to fill the gap that previously existed between the upper limit of 50,000 years, in which carbon dating is applicable, and the lower limit of roughly 1,000,000 years, below which potassium-argon dating becomes excessively laborious.

A second useful feature is that measurements can sometimes be made on extremely minute specimens, such as chips of meteoritic minerals or fragments of glass from the ocean bottom.

A third useful feature is that each mineral dates the last cooling through the temperature below which tracks are retained permanently. Since this temperature is different for each mineral, it is possible to measure the cooling rate of a rock by dating several minerals—each with a different track-retention temperature. It is similarly possible to detect and date the reheating of an old rock (caused, for example, by a local igneous intrusion or regional metamorphism) by the finding of a low age for the less retentive minerals and a greater age for the more retentive ones. *See* AMINO ACID DATING; EARTH, AGE OF; GEOLOGIC TIME SCALE; MINERAL; ROCK AGE DETERMINATION. R. L. Fleischer; P. B. Price; R. M. Walker

Bibliography. G. Faure, *Principle of Isotope Geology,* 2d ed., 1986; G. A. Wagner and P. Van Den Huate, *Fission Track Dating*, 1992.

FitzGerald-Lorentz contraction

The contraction of a moving body in the direction of its motion. In 1892 G. F. FitzGerald and H. A. Lorentz proposed independently that the failure of the Michelson-Morley experiment to detect an absolute motion of the Earth in space arose from a physical contraction of the interferometer in the direction of the Earth's motion. According to this hypothesis, as formulated more exactly by Albert Einstein in the special theory of relativity, a body in motion with speed v is contracted by a factor $\sqrt{1 - v^2/c^2}$ in the direction of motion, where c is the speed of light. *See* LIGHT; RELATIVITY. E. L. Hill

Flame

An exothermic reaction front or wave in a gaseous medium. Consider a uniform body of gas in which an exothermic chemical reaction (that liberates heat) is initiated by raising the temperature to a sufficiently high level; the reaction is started by a localized release of heat, as by a sufficiently energetic spark, and then spreads from the point of initiation. If the reaction is relatively slow, the whole gas will be involved before the initial region has finished reacting. If, however, the reaction is relatively fast, the reaction zone will develop as a thin front or wave propagating into the unreacted gas, leaving fully reacted gas behind. If the front, in addition, shows luminosity (emission of light), the flame may be considered a classical example. However, perceptible emission of visible radiation is not essential to the definition, and some flames are indeed vanishingly faint under ordinary viewing conditions.

Sufficient reaction rates may also be attained under special conditions (when the gas is very slowly heated inside a closed vessel) without very high temperatures if free radicals are generated in good concentration; this gives so-called cool flames.

Flame-producing reactions. The most common flame-producing reaction is combustion, which is broadly defined as a reaction between fuel and an oxidizer. The oxidizer is typically oxygen (usually in air), but a variety of other substances (for example, bromine with hydrogen) can play the same role in combination with the right fuel. While the overall theoretical reaction in a combustion flame—namely, fuel and oxidizer making fully oxidized products such as carbon dioxide and water vapor—is invariably simple, the actual reaction mechanism is typically very complex, involving many intermediate steps and compounds. Free radicals are generally present and figure prominently in the mechanism. *See* FREE RADICAL.

An overall reaction involving just one reactant is chemical decomposition, for example, ozone decomposing into oxygen. Decomposition flames are usually simpler chemically than combustion flames.

Combustion flames. Combustion flames are broadly divided into premixed flames and diffusion flames. Premixed flames occur when fuel and oxidizer are mixed before they burn. Diffusion flames occur when fuel and oxidizer mix and burn simultaneously. The intermediate case, with partial premixing, has been of relatively low theoretical and practical interest. Flames are further categorized on the basis of shape, time behavior (stationary or moving), flow regime (laminar or turbulent), buoyancy regime (forced convection or natural convection), presence or absence of confinement (as by combustion chamber walls), and flow complications (such as swirling flow and crosswind).

Premixed flames. A flame can be produced in a premixed gas of fuel and oxidizer only if the composition falls between the flammability limits. Once initiated, the flame continually advances into the unburned gas and leaves burned gas behind. The situation is simplest when the unburned gas is still or in laminar motion. Like a sound wave, the flame may be considered to be moving everywhere in the direction normal to its plane. The speed of its advance relative to the unburned gas is the laminar flame speed υ_L (the name and symbol most frequently seen in the older literature are burning velocity and S_u). Away from walls and wherever the flame is thin relative to its curvature, υ_L depends only on the gas properties, and it is this value of υ_L that is usually used. The range of υ_L for near-stoichiometric mixtures of organic fuels with air at 77°F (25°C) and atmospheric pressure is about 0.8–80 in./s (2–200 cm/s), and 16 in./s (40 cm/s) is average. Simplest theory suggests that υ_L is independent of pressure, but the actual behavior is more complex. The effect of increasing the absolute unburned gas temperature T_u is generally to increase υ_L, roughly in proportion to the $^3/_2$ power of T_u.

Figure 1 illustrates some typical features of a premixed laminar flame front. An observer at rest with respect to the flame sees unburned gas (reactants) flowing into the reaction zone from one side and burned gas (products) emerging from the other. The concentrations of fuel and oxidizer fall across the reaction zone, those of intermediate products rise and then fall, and those of final products continually rise. The flame is divided into several zones by the phenomena. The luminous zone is embedded in the reaction zone. The temperature rise accompanying the reaction produces a strong temperature gradient across the flame. Heat is conducted down this gradient, from the departing hot gas into the entering cold gas. The reaction zone is thus preceded by a preheat zone in which the unburned gas is warming but not yet reacting. From this description, it might be assumed that heat conduction into the cold gas, bringing it up to reaction temperature levels, is the controlling mechanism in laminar flame propagation in premixed gases. However, diffusion of free radicals toward the incoming unburned gas also plays a part in kindling reaction and may sometimes, as in cool flames, be dominant.

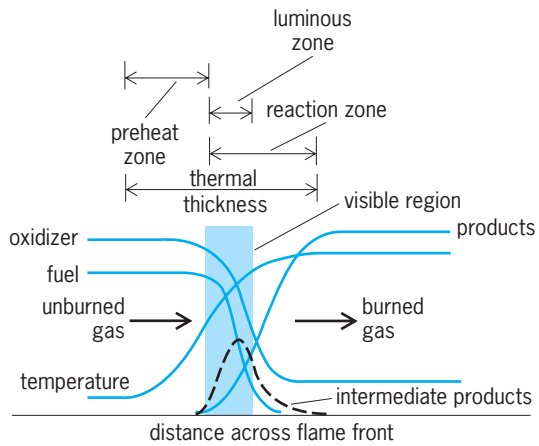

Fig. 1. Structure of a premixed laminar flame front.

The thickness of a laminar premixed flame is commonly measured by the thermal thickness—the thickness of the zone of appreciable temperature gradient. In combustion flames of organic fuels with air at atmospheric pressure, cold reactant temperature [77°F (25°C)], and near-stoichiometric composition, the thermal thickness λ is about 0.04–0.2 in. (1–5 mm). It is very nearly inversely proportional to the pressure P, and the product $\upsilon_L \lambda P$ is roughly constant.

The effect of turbulence on flame propagation in premixed gases is complex. A rough approximation is that the flame is locally laminar, but is everywhere more or less strongly convoluted or "wrinkled" by the turbulence. The volume of the reaction zone is thereby increased and, with it, the rate of burning. A common way of characterizing turbulent flame propagation is to define a surface following the average position of the front (smoothing out the wrinkles). The speed of advance of this mean front relative to the mean motion of the unburned gas is called the turbulent flame speed υ_T. The turbulent flame speed is generally greater than the laminar flame speed υ_L, and is a complicated function of υ_L and of the properties of the turbulence in the unburned gas (particularly the spatial scale and the level of velocity fluctuation).

Premixed flames occur in a great variety of situations, both contrived and accidental. In technical applications with continuous combustion, burner flames are most commonly used. The bunsen burner is a familiar example. Typically, premixed gas issues from an orifice or pipe over which the flame hovers in the form of a hollow cone, stabilized at its base on the burner rim. The stabilization mechanism is too complex for simple explanation. The flow of the unburned gas and the propagation of the flame into that gas are exactly balanced, so the flame stands still. When the gas flow is increased, the flame cone lengthens, and when it is decreased, it shortens. When the flow is made too high, stabilization fails and the flame is swept away downstream (blowoff). Stabilization also fails when the flow is made too low, in which case the flame propagates upstream into the burner (flashback). A flame in a stream of premixed gas can also be stabilized in the wake of a nonstreamlined solid body placed in the stream, a so-called bluff-body stabilizer. In applications where flame is initiated in a finite body of gas rather than a continuing stream, as in spark-ignited internal combustion engines, the flame sweeps through the gas from ignition point to boundaries, leaving it more or less completely burned. In engines, high unburned-gas temperatures produced by precompression and high turbulence levels promoted by intake-valve and combustion-chamber design give high effective flame speeds needed for efficient high-speed operation. *See* COMBUSTION CHAMBER; INTERNAL COMBUSTION ENGINE.

Diffusion flames. In technical applications (for example, furnaces, boilers, gas heating devices, and jet and gas-turbine engines), these are typically burner flames designed for continuous combustion, and are much more widely used than premixed flames for such applications. Industrial diffusion flames of any size are also generally turbulent. However, laminar flames do occur in a number of important situations: small gas-jet or vapor-flow flames, as on gas-range burners and candles; the burning of liquid-fuel droplets and metal particles; and the burning of particles or small pieces of solid organic fuels such as wood and coal.

All burning in flames associated with solid and liquid fuels remains in the gas phase. Heat from the flame melts fuels such as candle wax and magnesium metal (at least at the surface) and vaporizes the liquid. The vapor diffuses away and, at the flame front, meets oxygen diffusing toward it from the surrounding gas and burns. Liquid droplets similarly generate vapor, which then burns. In the case of wood and coal, heating at first drives off combustible gases (volatile combustible matter), which burn in a diffusion flame. The subsequent burning of the residual char, however, is exceptional and may involve no flame. The melting and boiling points of carbon are so high that vaporization is normally impossible. Instead, oxygen is usually considered to diffuse to the char surface and there react directly, a mode of reaction called heterogeneous combustion. For some circumstances, though, a mechanism is favored in which carbon dioxide diffuses to the char surface and reacts to produce carbon monoxide. The carbon monoxide then diffuses away, meets incoming oxygen at a flame front, and burns there, forming carbon dioxide, of which half diffuses back to the char to react with it and the other half diffuses away as final product.

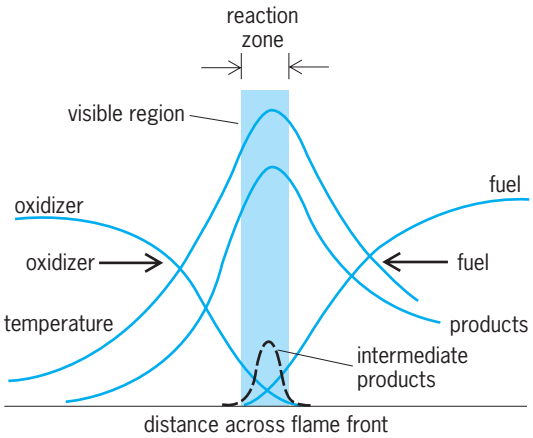

Fig. 2. Structure of a laminar diffusion flame front.

Figure 2 illustrates the structure of a laminar diffusion flame front. Fuel and oxidizer diffuse toward each other and meet and react at the flame front. The reaction zone is the region of overlap between the profiles of fuel and oxidizer, and the temperature peaks in this zone. The reaction zone is usually very thin, which means that the reaction rate of fuel and oxygen is chemical-kinetically very fast, so the fuel and the oxidizer react virtually as fast as they can get together by diffusion. This is why nonpremixed flames are commonly called diffusion flames. Of course, intermediate products of combustion appear in, and tend to be confined to, the reaction zone. Final products are found on both sides of the reaction zone. Any inerts accompanying fuel or oxidizer (such as nitrogen in air) are also found on both sides of the reaction zone. A diffusion flame cannot propagate through the gas, neither into the oxidizer on one side nor into the fuel on the other, but is fixed at the front where fuel and oxidizer meet. The concept of flame speed here is meaningless, and the situation is altogether different from a premixed flame.

The effect of turbulence bears a vague similarity to that in the premixed flame in that here, too, turbulence produces a distorted and greatly convoluted reaction zone. However, turbulence affects the meeting of fuel and oxidizer quite profoundly. A crude way of explaining this is to say that turbulent diffusion is added to, and usually completely overshadows, molecular diffusion. *See* CHEMICAL DYNAMICS; EXPLOSIVE; OXYGEN.

Henry A. Becker

Bibliography. N. A. Chigier (ed.), *Progress in Energy and Combustion Science*, vol. 10, 1986; Combustion Institute, *Combustion Flame*, vol. 55, 1984; *Combustion Science and Technology* (journal), vol. 35, 1984; F. M. Fristrom, *Flame Structure*

and Processes, 1994; I. Glassman (ed.), *Combustion*, 3d ed., 1996; K. Kuo, *Fundamentals of Combustion*.

Flameproofing

The process of treating materials so that they will not support combustion. Although cellulosic materials such as paper, fiberboard, textiles, and wood products cannot be treated so that they will not be destroyed by long exposure to fire, they can be treated to retard the spreading of fire and to be self-extinguishing after the igniting condition has been removed.

Treatments. Numerous methods have been proposed for flameproofing cellulosic products. One of the simplest and most commonly used for paper and wood products is impregnation with various soluble salts, such as ammonium sulfate, ammonium phosphate, ammonium sulfamate, borax, and boric acid. Special formulations are often used to minimize the effects of these treatments on the color, softness, strength, permanence, or other qualities of the paper.

For some applications, these treatments are not suitable because the salts remain soluble and leach out easily on exposure to water. A limited degree of resistance to leaching can be achieved by the addition of latex, lacquers, or waterproofing agents. In some cases the flameproofing agent can be given some resistance to leaching by causing it to react with the cellulose fiber (for example, urea and ammonium phosphate).

Leach-resistant flameproofing may also be obtained by incorporating insoluble retardants in the paper during manufacture, by application of insoluble materials in the form of emulsions, dispersions, or organic solutions, or by precipitation on, or reaction with, the fibers in multiple-bath treatments. The materials involved are of the same general types as those used for flameproofing textiles and include metallic oxides or sulfides and halogenated organic compounds. Because of the higher material cost and more elaborate equipment required, such treatments are used on paper only when unusual conditions justify their cost.

Action of retardants. Various theories have been proposed to explain flameproofing action. In general, these are similar to those considered in the flameproofing of textiles. Two types of combustion may be involved. In one, the volatile decomposition products burn with a flame; in the other, the solid material undergoes flameless combustion, or afterglow. In general, the alkaline types of retardants are effective in preventing afterflame, and the acid types control afterglow. A few, such as ammonium phosphates and halogenated products, reduce both flame and glow.

The theory for prevention of afterflame having the most support is a chemical one, which holds that effective flame retardants direct the decomposition of the cellulose when heated to minimize the formation of volatile flammable products and increase the amount of water and solid char formed. The prevention of afterglow is usually attributed to a modification of the flameless combustion to make it less exothermic and therefore incapable of maintaining itself, for example, by formation of carbon monoxide instead of carbon dioxide. Other theories include the formation of a coating or froth which excludes oxygen and smothers the combustion, the formation of nonflammable gases which dilute the flammable products and exclude oxygen, and the thermal theories which hold that heat is dissipated by endothermic reactions or by conduction. It is probably that these actions do contribute in some cases.

The susceptibility of paper or fiberboard to combustion may also be modified by the introduction of noncombustible but otherwise inert materials, such as mineral fillers, asbestos, and glass or ceramic fibers, into the paper or board during manufacture. The high proportions required result in products with properties determined largely by noncombustible components rather than by cellulosic fibers.

The behavior of paper, fiberboard, or wood products exposed to fire depends also on structural factors. The density of the material and the ratio of mass to surface exposed affect the ease of ignition and the rate of spreading or penetration by the fire. Fire-resistant coatings minimize the spread of flame. Heat-reflective coatings or insulating coatings prevent or retard the material beneath them from reaching the ignition temperature. Intumescent coatings are formulated to form a thick insulating foam when exposed to fire.

A wide variation exists in methods for evaluating the efficiency of a flameproofing treatment or of suitability for a given use. Paper is usually tested by igniting with a flame of specified character for a predetermined period of time and then noting duration of afterflame and of afterglow and the amount of char length or char area. Structural materials are tested by other methods which involve features of construction in addition to those of composition. Methods have been standardized by the Technical Association of the Pulp and Paper Industry (TAPPI), American Society for Testing and Materials (ASTM), and various governmental agencies.

For a discussion of textile flameproofing *see* COMBUSTION; FIRE EXTINGUISHER; FLAME; TEXTILE CHEMISTRY.

T. A. Howells

Bibliography. Fire Retardant Chemicals Association, *Flame Retardants 101: Basic Dynamics*, 1998; A. R. Horrocks and D. Price, *Fire Retardant Materials*, 2001; National Research Council, *Toxicological Risks of Selected Flame-Retardant Chemicals*, 2000.

Flash welding

A form of resistance welding that is used for mass production. The welding circuit consists of a low-voltage, high-current energy source (usually a

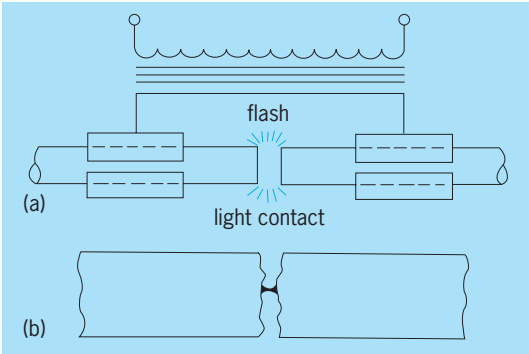

Fig. 1. Elements of flash-welding operation. (*a*) Two pieces of horizontally clamped metal are brought together to make a light contact. At that time the welding circuit is closed to the transformer. (*b*) Current starts flowing when small areas of the workpiece make initial contact. (*After A. L. Phillips, ed., Welding Handbook, 4th ed, American Welding Society, 1957*)

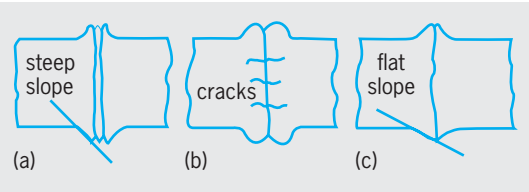

Fig. 2. Flash-weld quality indicated on the surface. (*a*) Satisfactory heat and upset. (*b*) Cracks caused by insufficient heat. (*c*) Insufficient heat or upset force, or both. (*After A. L. Phillips, ed., Welding Handbook, 4th ed., American Welding Society, 1957*)

welding transformer) and two clamping electrodes, one stationary and one movable.

The two pieces of metal to be welded are clamped tightly in the electrodes, and one is moved toward the other until they meet, making light contact (**Fig. 1***a*). Energizing the transformer causes a high-density current to flow through small areas that are in contact with each other (Fig. 1*b*). Flashing starts, and the movable workpiece must accelerate at the proper rate to maintain an increasing flashing action. After a proper heat gradient has been established on the two edges to be welded, an upset force is suddenly applied to complete the weld. This upset force extrudes slag, oxides, molten metal, and some of the softer plastic metal, making a weld in the colder zone of the heated metal (**Fig. 2**). Flash and upset material are frequently removed, although when the design permits, a portion of the upset material may be retained as a reinforcement. *See* ARC WELDING; RESISTANCE WELDING. Eugene J. Limpel

Bibliography. R. Carr and R. O'Con, *Welding Practices and Procedures*, 1983; H. B. Cary, *Modern Welding Technology*, 4th ed., 1997; J. W. Giachino and W. R. Weeks, *Welding Skills and Practices*, 1985.

Flat-panel display device

An electronic display in which a large orthogonal array of display elements, such as liquid-crystal or electroluminescent elements, form a flat screen. Since the days of the first television tube, display engineers have dreamed of producing flat-panel displays (FPDs) with the visual attributes of a cathode-ray tube (CRT). The term "flat-panel display" is actually a misnomer, since thinness is the distinguishing characteristic. Most television sets and computer monitors currently employ cathode-ray tubes. Cathode-ray tubes cannot be thin because the light is generated by the process of cathodoluminescence whereby a high-energy electron beam is scanned across a screen covered with an inorganic phosphor, one point, or pixel, at a time, activating light-emitting centers in the material. The cathode-ray tube must have moderate depth to allow the electron beam to be magnetically or electrostatically scanned across the entire screen. *See* CATHODE-RAY TUBE; CATHODOLUMINESCENCE; ELECTRONIC DISPLAY.

For a flat-panel display technology to be successful, it must at least match the basic performance of a cathode-ray tube by having (1) full color, (2) full gray scale, (3) high efficiency and brightness, (4) the ability to display full-motion video, (5) wide viewing angle, and (6) wide range of operating conditions. Flat-panel displays should also provide the following benefits: (1) thinness and light weight, (2) good linearity, (3) insensitivity to magnetic fields, and (4) no x-ray generation. These four attributes are not possible in a cathode-ray tube.

Flat-panel displays can be divided into three types: transmissive, emissive, and reflective. A transmissive display has a backlight, with the image being formed by a spatial light modulator. A transmissive display is typically low in power efficiency; the user sees only a small fraction of the light from the backlight. An emissive display generates light only at pixels that are turned on. Emissive displays should be more efficient than transmissive displays, but due to low efficiency in the light generation process most emissive and transmissive flat panel displays have comparable efficiency. Reflective displays, which reflect ambient light, are most efficient. They are particularly good where ambient light is very bright, such as direct sunlight. They do not work well in low-light environments.

Most flat-panel displays are addressed as an X-Y matrix, the intersection of the row and column defining an individual pixel (see **illus.**). Matrix addressing provides the potential for an all-digital display. Currently available flat-panel display devices range from 1.25-cm (0.5-in.) diagonal displays used in head-mounted systems to 125-cm (50-in.) diagonal plasma displays.

Currently, most commercially manufactured flat-panel display devices are liquid-crystal displays (LCDs). The benchmark for flat-panel display performance is the active matrix liquid-crystal display (AMLCD). Most portable computers use AMLCDs. Competing flat-panel display technologies include electroluminescent displays, plasma display panels, vacuum fluorescent displays, and field-emission displays. Electroluminescent displays are often used in industrial and medical applications because of their ruggedness and wide range of operating

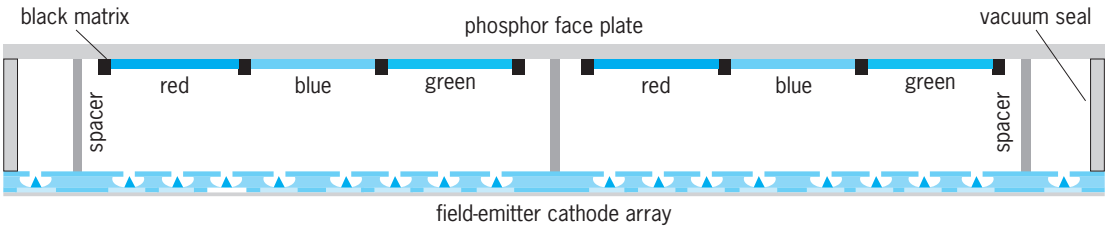

Field-emission display, an example of a matrix-addressed display. Each pixel is addressed in an X-Y matrix. In the case of a color display, each pixel is subdivided into a red, a blue, and a green subpixel.

temperatures. Plasma display panels are most often seen as large flat televisions, while vacuum fluorescent displays are used in applications where the information content is fairly low, such as the displays on appliances or in automobiles. Field-emission displays are the most recent of these flat-panel technologies. There are also emerging flat-panel display technologies based on laser-pumped-polymer–switched matrices and electrophoretically switched reflective displays. They offer the potential to provide low-cost, very large displays. Another area of active research in flat-panel displays is the development of flexible displays produced on plastic substrates, which have the potential for a tremendous number of applications.

Liquid-crystal displays. The liquid-crystal display is a spatially addressed light valve that transmits or reflects light of a given polarization, depending on the orientation of the liquid crystal molecule. The orientation of the molecule is determined by how the material is affixed to the substrate (the pre-tilt angle) and the applied electric field placed across the molecule. There are several variations of the liquid-crystal display, including the twisted nematic, the ferroelectric, the polymer dispersed, and the cholesteric liquid-crystal display. Each of these technologies has specific strengths and weaknesses, ranging from high-speed switching for ferroelectric liquid crystals to bistable operation for cholesteric liquid crystals. Advanced liquid-crystal displays sometimes include active matrix addressing. In an active matrix addressing scheme, one or more thin-film transistors are placed at each pixel to more accurately control the applied voltage on the pixel, thus increasing both the refresh rate and the dynamic range of gray scale. Amorphous silicon is used in the manufacture of thin-film transistors; for high-performance applications, polycrystalline silicon is used.

Liquid-crystal displays can operate in either a transmissive or a reflective mode. In the transmissive mode, a backlight and a polarizer are required. To produce a color display, each pixel is usually divided into at least three subpixels, with a red, blue, or green filter placed above each subpixel. In the reflective mode, ambient light is reflected from the liquid crystal by using a reflective material on the backside of the liquid-crystal display, or the liquid-crystal material itself can be reflective. The deficiencies of current liquid-crystal displays include (1) limited viewing angle for both intensity and color purity, (2) limited operating temperature range, (3) low efficiency for light utilization (50% polarization loss, 66% loss for color filters), and (4) slow refresh rates for reflective displays. These deficiencies are being reduced with continuing research. Advanced AMLCDs exhibit resolutions greater than 140 lines per inch (color; 55 lines per centimeter) and diagonal dimensions as large as 30 in. (76 cm). *See* LIQUID CRYSTALS.

Electroluminescent displays. Electroluminescent (EL) displays are emissive displays in which light is generated by the passage of current through an electroluminescent layer. Traditionally they have been based on inorganic materials, but recently there has been significant interest in organic-based light-emitting materials. In inorganic electroluminescent displays the light is generated through impact ionization of the active sites in the phosphor layer by energetic electrons. In organic-based light-emitting diodes, light is generated through electron-hole pair recombination. Inorganic displays are high-voltage devices (more than 100 V), while organic displays are generally low voltage (less than 10 V). The electroluminescent process is typically not as efficient in generating light as the cathodoluminescent process used in cathode-ray tubes. A main advantage of the electroluminescent display is that it is completely solid-state. Large-area [greater than 3-in. (7.6-cm) diagonal] electroluminescent displays are normally matrix-addressed devices, similar to a passive matrix liquid-crystal display. Recently, active matrix addressing has been introduced to small diagonal [less than 5 cm (2 in.)] electroluminescent displays, where the display is produced directly on a silicon substrate. Active matrix addressing provides faster displays with larger dynamic range, as well as higher brightness due to higher-frequency operation. Also, active matrix addressing reduces the interconnect density for high-resolution displays. Active matrix electroluminescent displays with 12-micrometer-size pixels [2000 lines per inch (800 lines per centimeter)] are currently under development for small-format displays. *See* ELECTROLUMINESCENCE; LIGHT-EMITTING DIODE.

Plasma display panels. The plasma display panel (PDP) is an emissive display in which light is generated via an excited plasma in a low-pressure gas. In a color plasma display panel, the plasma generates deep ultraviolet photons, which in turn excite photoluminescent phosphors to produce the colors for the display. Resolution is limited by the physical barriers required to separate the plasma in the individual subpixels. Plasma display panels are typically

low- to moderate-resolution devices that have relatively poor power efficiency. They are best suited for large, low- or moderate-resolution display applications such as wall television. Plasma display panels generally are not suitable for portable applications. There have been significant improvements in the power efficiency and the lifetime of color plasma panels. Plasma panels as large as 125-cm (50-in.) diagonal are in production, and prototype panels with a resolution of 80 lines per inch (32 lines per centimeter) are being produced.

Vacuum fluorescent displays. The vacuum fluorescent display is based on the thermionic emission of electrons from a small wire, and the cathodoluminescent production of light. Most vacuum fluorescent displays operate at low anode voltage, typically less than 50 V, and use low-voltage phosphors such as zinc oxide. The display is formed by segmenting the phosphor and applying the anode bias to the segments to be illuminated. The resolution is moderate to low, such as for small alphanumeric displays. The main advantages of the vacuum fluorescent display are high brightness, low operating voltage, and a wide operating temperature range, making it attractive for automotive applications.

Field-emission displays. A field-emission display (FED) is similar to a cathode-ray tube in many ways. The major difference is the type and the number of electron emitters. The electron guns in a cathode-ray tube produce electrons by thermionic emission from a cathode. CRTs typically have one or three electron guns, depending on the configuration of the electron scanning system. The extracted electrons are focused by the electron gun, and while the electrons are accelerated toward the anode plate, electromagnetic deflection coils are used to scan the electron beam across a phosphor-coated faceplate. This requires a large distance between the deflection coils and faceplate. The larger the cathode-ray tube viewing area, the greater the depth (anode to cathode distance) required to scan the beam.

By contrast, in a field-emission display, electrons are produced by cold cathode emission. Several electron emitters or cathodes are associated with each pixel (see illus.). This eliminates the need for beam steering. Field-emission of electrons is achieved by applying a bias of less than 100 V between the cathode and a control electrode. These extracted electrons are then accelerated without deflection toward a phosphor-coated flat faceplate, biased between 300 and 10,000 V. Unlike the cathode-ray tube, electrons in a field-emission triode (cathode-gate-anode) device receive little focus. Thus, in a field-emission display, in order to achieve a small spot size, a small gap (less than 1 mm) is maintained between the cathode array and the anode faceplate. As a result, a field-emission display is much thinner than a cathode-ray tube. Also, since the emitters are placed in an addressable X-Y matrix, a field-emission display does not suffer from the traditional nonlinearity and pin cushion effects associated with a cathode-ray tube. A field-emission display, therefore, combines the best features of a cathode-ray tube with the thinness and linearity of a flat-panel display. *See* FIELD EMISSION.

Bruce E. Gnade

Bibliography. B. R. Chalamala, Y. Wei, and B. Gnade, FED up with flat tubes, *IEEE Spectrum*, 35:42, 1998; G. Gu and S. Forrest, Design of flat-panel displays based on organic light emitting devices, *IEEE J. Selected Top. Quant. Electr.*, 4:83, 1998; Liquid crystal devices display their empire, *Electr. Des.*, 45:40, 1997; L. W. MacDonald and A. C. Lowe (eds.), *Display Systems: Design and Applications*, John Wiley, New York, 1997; L. E. Tannas, Jr. (ed.), *Flat-Panel Displays and CRTs*, Van Nostrand Reinhold, New York, 1985; K. Werner, The flowering of flat displays, *IEEE Spectrum*, 34(5):40–49, May 1997.

Flavonoid

A large category of natural plant products that derive from γ-pyrone (**1**). All flavonoid compounds, which

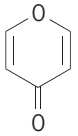

(**1**)

are derived from either 2-phenylbenzopyrone (**2**) or 3-phenylbenzopyrone (**3**), can be classified into 10 groups (**Fig. 1**): chalcones, flavanones, flavones,

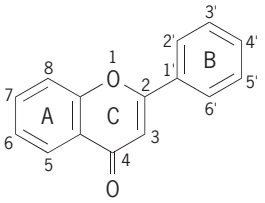

(**2**)

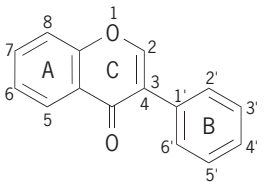

(**3**)

flavonols, anthocyanidins (flavylium cations), flavan 3-ols (catechins), flavan 3,4-diols (proanthocyanidins), biflavonoids and oligomeric flavonoids, isoflavonoids, and the aurones. They differ in the oxidation level or substitution pattern of their heterocyclic ring (ring C).

Physical and chemical properties. Most flavonoids show characteristic absorbance in the visible and ultraviolet region of the spectrum. Depending on the basic flavonoid skeleton, the visible absorption bands may be located between 330 and 370 nanometers. The absorption band in the ultraviolet region is less affected by structural variations and is usually between 260 and 280 nm. Because of the close

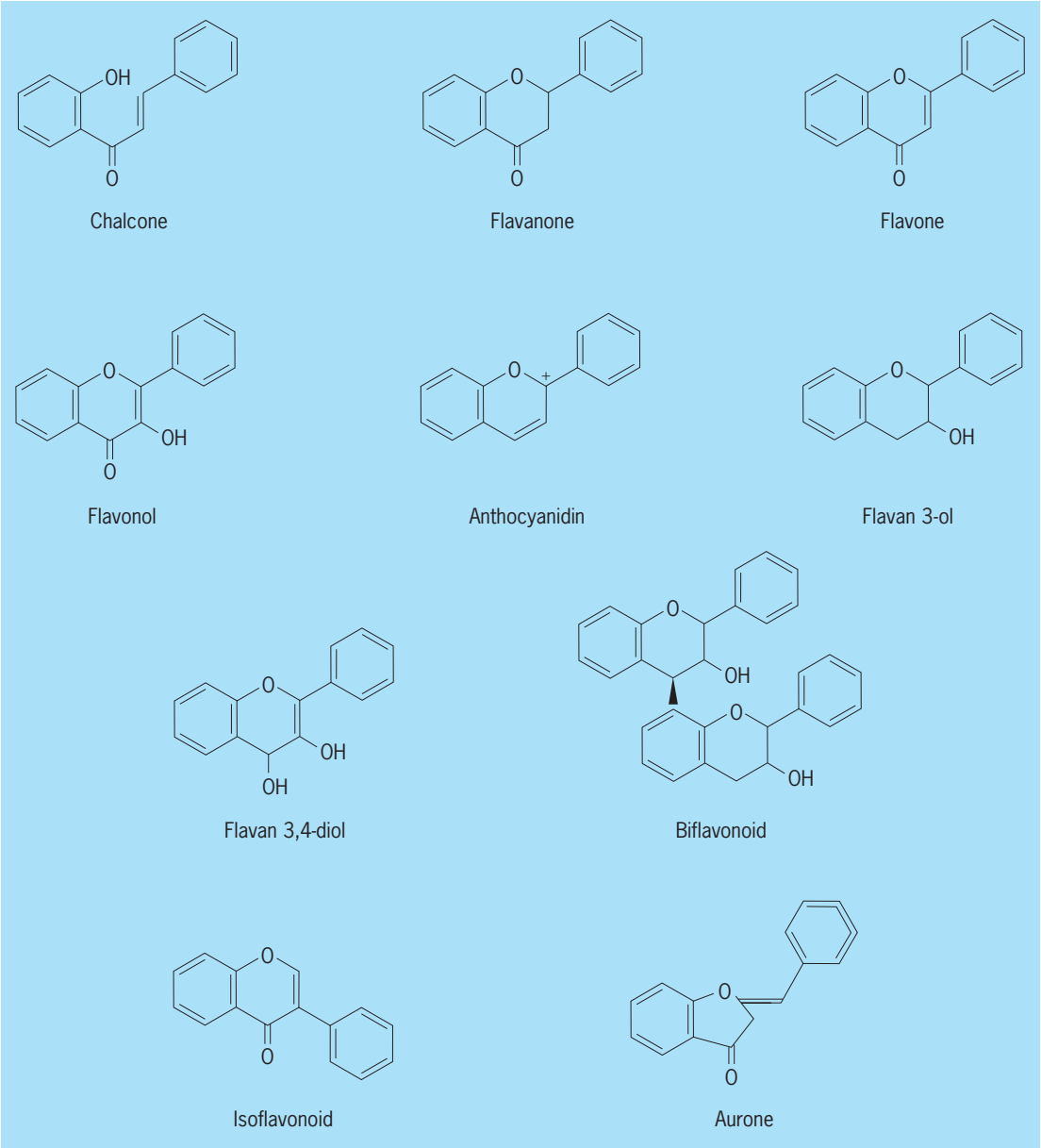

Fig. 1. The 10 classes of flavonoid compounds and their structural relationship.

relationship between the positions of these absorption bands and the basic structure of the flavonoid skeleton, they have proven essential in structural determinations.

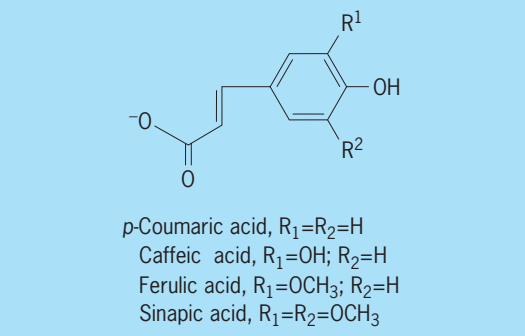

Fig. 2. Aromatic acids that may be attached to the sugar moiety of the flavonoid compounds.

More than 1300 different flavonoid compounds have been isolated from plants. Whereas the basic differences between the various classes of flavonoids lie with the oxidation stage of the heterocyclic ring and the site of attachment of the phenyl ring (ring B) to the heterocycle, individual flavonoids in a group differ from each other by the number and position of the hydroxy, methoxy, and sugar substituents. Thus positions 3′, 4′, 5, and 7 (structures 2 and 3) are the most commonly substituted among the flavonones and flavones, and 3, 3′, 4′, 5, and 7 among the flavonols, anthocyanidins, and flavan 3-ols. The number of the substituents may vary from zero to eight, but a combination of four to six hydroxyl and methoxyl groups is found most frequently. As a rule, flavonoid compounds occur in plants as glycosides, with hexoses such as glucose, galactose, and rhamnose, and pentoses such as arabinose and xylose as the most commonly found sugars. The sugars can be

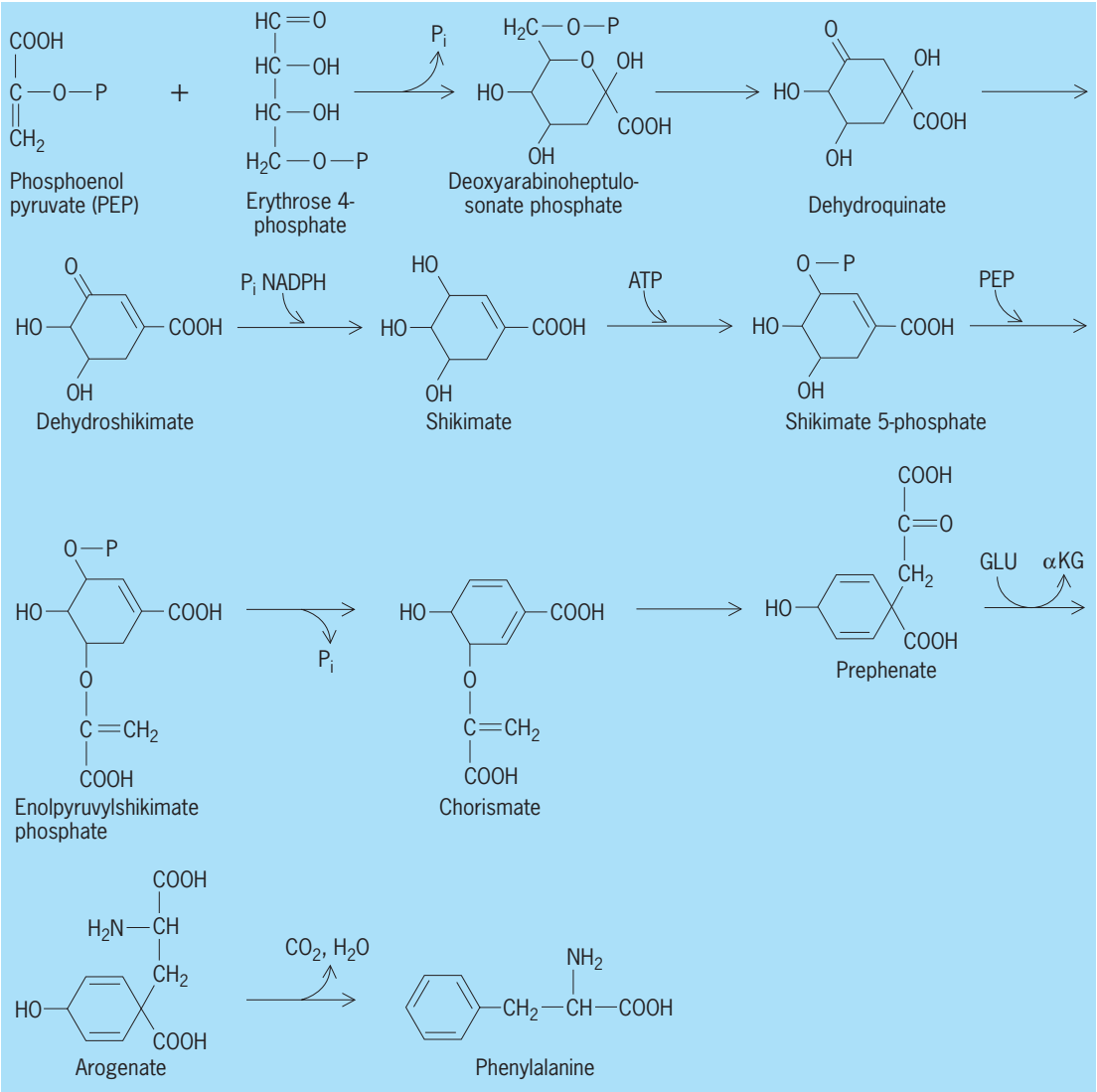

Fig. 3. Shikimate section of the plant aromatic pathway that produces phenylalanine (and the other aromatic amino acids tyrosine and tryptophan). P_i = inorganic phosphate; GLU = glutamate; αKG = α-ketoglutarate.

attached singly or in combination with each other. The most common position of glycosylation of the flavanones and flavones is the 7-hydroxyl, in flavonols the 3- and the 7-hydroxyls, and in anthocyanins the 3- and 5-hydroxyls. Glycosylation renders these compounds water-soluble and permits their accumulation in the vacuoles of cells. *See* GLYCOSIDE.

The sugars in the molecules often have aliphatic (acetic, malonic) or aromatic (*p*-coumaric, caffeic, ferulic, and sinapic) acids attached (**Fig. 2**). The aliphatic acids are difficult to detect by spectral investigation, but the aromatic acids are easily found by the appearance of a new band between 307 and 340 nm in the spectra. Sugars may also be attached by carbon-carbon (C-C) bonds to flavonoids (C-glycosides); the most commonly found attachment sites are positions 6 and 8. Flavonoids that carry terpenoid or alkaloid side chains are known to exist, but their occurrence is rare. Chalcones and aurones are intensely yellow and are the cause of

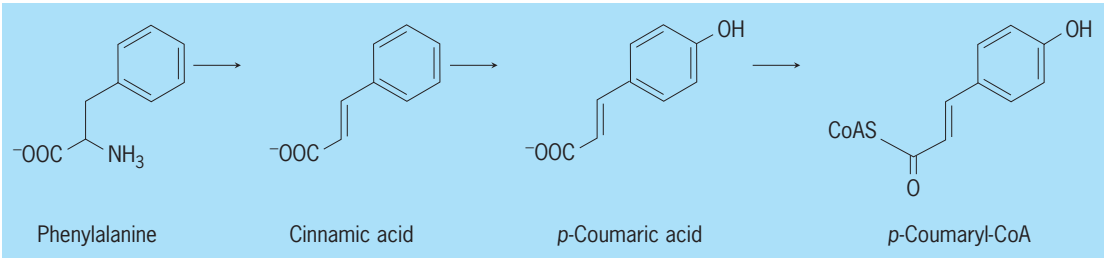

Fig. 4. Phenylpropanoid section of the plant aromatic pathway.

flower color in some plants (chalcones in carnations and Dahlia species and aurones in the daisy family).

Biosynthesis. As complicated as the structures of the diverse flavonoids may appear, their biosynthesis starts from simple components. The pathway that leads to their formation, referred to as plant aromatic metabolism, is one of the best understood in plants. It comprises the shikimate pathway, the phenylproponoid segment, and the flavonoid segment. The shikimate pathway produces the aromatic amino acids phenylalanine, tyrosine, and tryptophan; the phenylpropanoid segment transforms phenylalanine to cinnamic acid derivatives; and the flavonoid segment condenses the activated cinnamic acid derivative, *p*-coumaryl-CoA, with three molecules of malonyl-CoA to form the first flavonoid compound, naringenin chalcone, from which all flavonoids derive. *See* AMINO ACIDS.

The shikimate section of the pathway starts with a simple sugar derivative, erythrose 4-phosphate (**Fig. 3**); its condensation with another simple compound, phosphoenol pyruvate, yields a seven-membered sugar, deoxyarabinoheptulosonate phosphate. That sugar cyclizes and undergoes a series of transformation reactions to form shikimate phosphate, for which the pathway segment is named. Shikimate phosphate then undergoes a further condensation reaction with phosphoenol pyruvate, rearrangement of this newly added side chain, amination, and in the case of phenylalanine, a loss of water.

One part of the phenylalanine so formed is deaminated by the first enzyme of the phenylpropanoid pathway segment, phenylalanine ammonia lyase, to form cinnamic acid (**Fig. 4**). Introduction of a hydroxyl group in the para position and condensation with coenzyme A produces *p*-coumaryl-CoA, the activated form of the phenylpropanoids. That may be further hydroxylated and methoxylated to the aromatic acid derivatives shown in Fig. 2 and may participate in the production of the plant structural component lignin. *See* LIGNIN.

Condensation of the activated *p*-coumaric acid with three similarly activated malonic acids by the enzyme chalcone synthase produces the first flavonoid compound, naringenin chalcone, from which all flavonoids derive. The flavonoid segment of the plant aromatic metabolism is shown in **Fig. 5**, using the example of kaempferol 3-glucoside, the simplest flavonol derivative in plants. A series of modifications of naringenin chalcone leads to the formation of all flavonoid classes and their glycosylated forms that occur naturally in plants. From an evolutionary point of view, the first enzymatic step of the flavonoid pathway section shows the condensation of three malonyl CoA's with the starter compound *p*-coumaryl-CoA to resemble in its mechanism that of the fatty acid synthesis that uses acetyl-CoA as a starter. This similarity with the otherwise different fatty acid metabolism suggests that flavonoid and fatty acid metabolism may have had common ancestral origins.

The biosynthetic mechanism of the isoflavonoids has two unique reactions not found in the pathway segments of other flavonoids. One is the formation of 5-deoxy isoflavonoids in legumes, wherein a specific reductase coacts with chalcone synthase to remove the hydroxyl prior to cyclization of the A ring. The other reaction is the migration of the aromatic B ring from the 2 to the 3 position. This reaction is catalyzed by the enzyme isoflavone synthase, a monooxygenase that requires reduced nicotinamide adenine dinucleotide phosphate, and is believed to take place via an enol-epoxide intermediate.

Controls over biosynthesis. In order to maintain a cell's metabolism and prevent overproduction of one or the other metabolite, biochemical events have to be strictly regulated. This is generally accomplished in cells and organisms by a mechanism called feedback control. In this mechanism, overproduction of

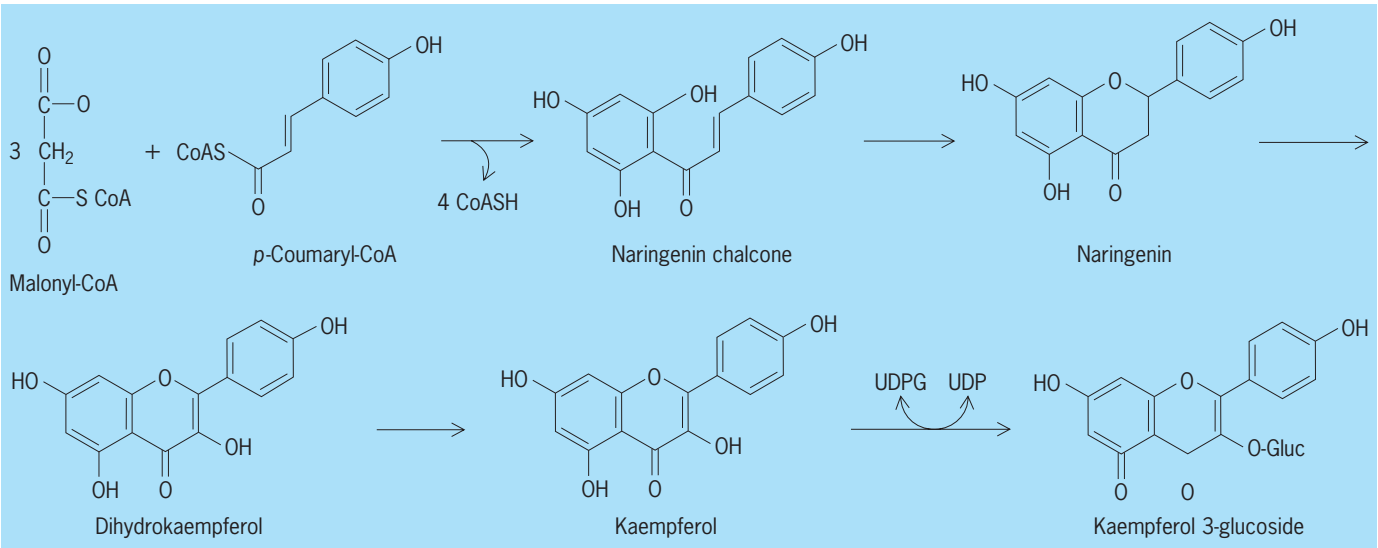

Fig. 5. Flavonoid section of the plant aromatic pathway. UDPG = uridine-5′-diphosphoglucose; UDP = uridine-5′-diphosphate.

one or more compounds has an inhibitory effect on enzymes located further upstream in the pathway, shutting them down until the concentration of the inhibitory product diminishes. Such a feedback control has been shown to exist in the shikimate segment of the flavonoid pathway, but no such control has been found with the other two segments. They are controlled at the genetic level by de novo synthesis of the enzymes from their corresponding messenger ribonucleic acids (**Fig. 6**). The pathway segments are thought to function as enzyme complexes in which the product of the one enzyme is used as substrate by the next until the final product is formed.

Sites of accumulation and synthesis. There are few data on cell or tissue-specific accumulation of flavonoids, primarily because of the difficulties encountered in isolating individual tissues that are free of cross contamination. The few reports available indicate that flavonoids accumulate in epidermal tissues, with approximately 70% in the upper and 30% in the lower epidermis. The development of methods to isolate intact, mature plant vacuoles on a large scale led to investigations that showed the subcellular accumulation site of flavonoids to be vacuolar.

Although the vacuoles are probably the only site of flavonoid accumulation in the cells, they are clearly not the site of synthesis. Instead, synthesis of flavonoids takes place in the cytoplasm, and evidence indicates that the different pathway segments may function as membrane-associated enzyme complexes.

Physiological role in plants. Flavonoid compounds can no longer be regarded as stray end products of metabolism; some are now known to be physiologically active. The symbiotic interaction of *Rhizobium meliloti* and some members of the family Leguminosae results in the formation of nitrogen-fixing root nodules. Specific genes are required for the early host responses of cortical cell division and root hair curling. Induction of this gene expression was found to be caused by host-symbiont signaling, and a number of flavonoid compounds were discovered to be the host-specific signal molecules. In alfalfa, the specific host signal is the 3′,4′,5,7-tetrahydroxyflavone, luteolin (**4**). *Trifolium* species produces 4′,7-dihydroxyflavone (**5**),

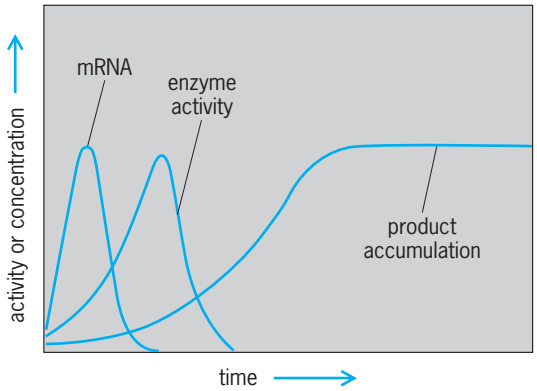

Fig. 6. Control of flavonoid synthesis in plants. A signal (either light or stress) induces messenger ribonucleic acid (mRNA) production, which synthesizes the corresponding enzyme. The enzyme in turn synthesizes the product, which accumulates when the pathway is blocked, or acts as the substrate for the next enzyme, which is concomitantly induced.

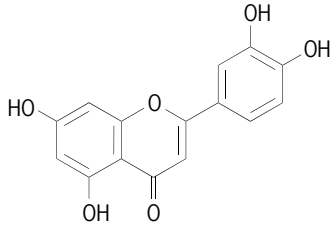

(4)

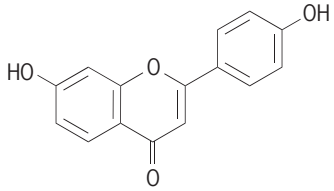

(5)

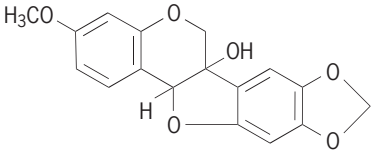

(6)

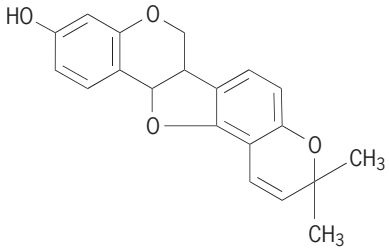

(7)

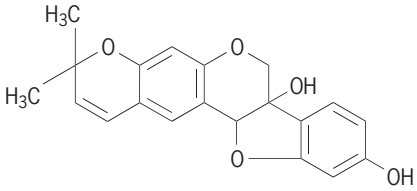

(8)

4′,7-dihydroxy-3′-methoxyflavone, and 4′-hydroxy-7-methoxyflavone as the specific host signals for nodulation. *Vicia sativa* synthesizes a number of nodulation-inducing compounds; at least one was reported to be a flavonoid. *See* NITROGEN FIXATION.

Light is an essential energy source for plants, and it provides external signals that regulate the developmental processes and adjustments to changes in the environment. In tissues exposed to potentially

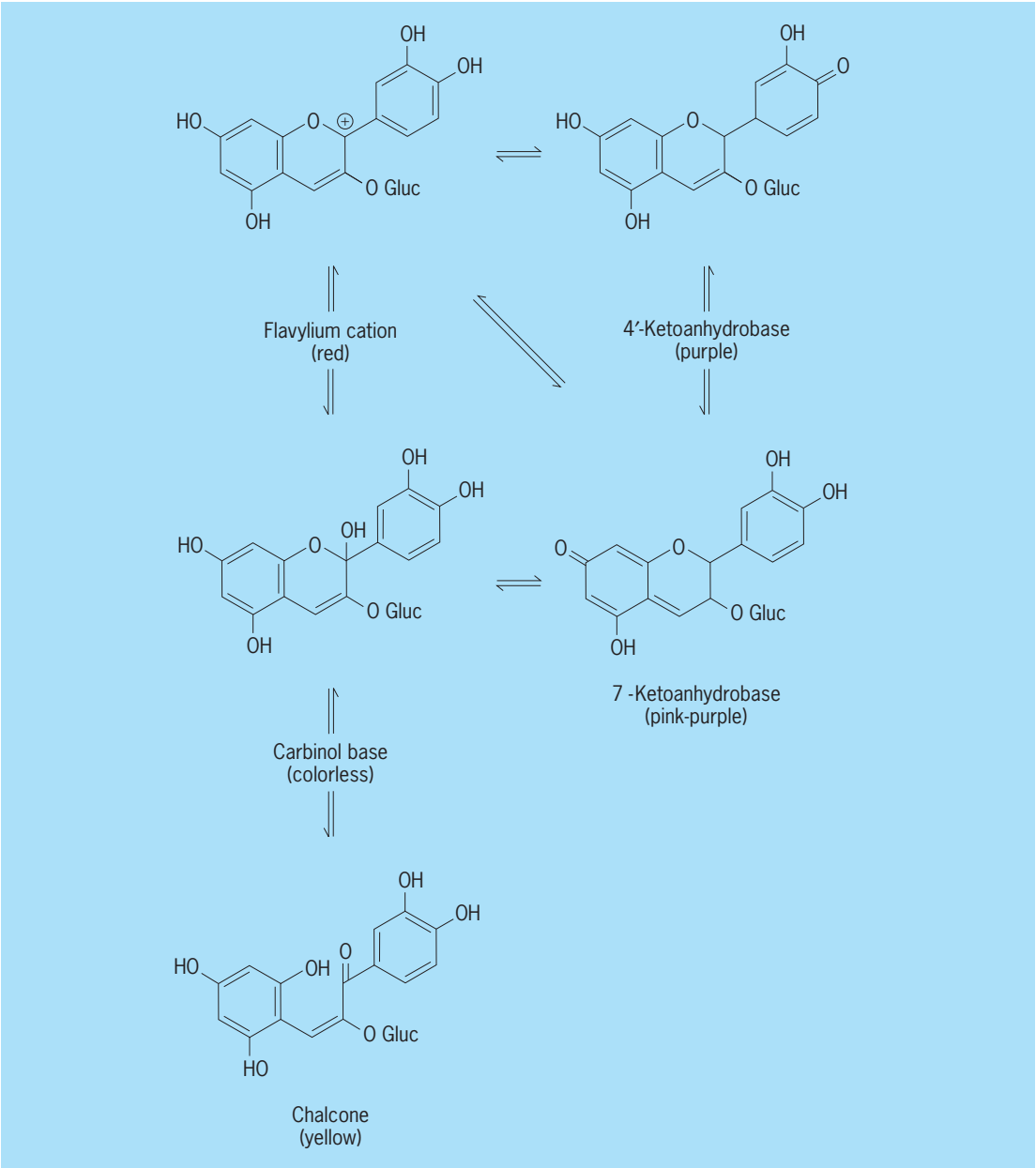

Fig. 7. Structural transformation reactions of anthocyanins as shown by the example of cyanidin 3-glucoside, in the pH range 1–7, which contributes to color variation in plants. O Gluc = O glucose.

damaging ultraviolet radiation, a major form of adjustment is the rapid accumulation of ultraviolet-blocking compounds. The enzymatic machinery for flavonoid production is induced by ultraviolet irradiation. Flavonoids accumulate in the vacuoles of epidermal cells and absorb strongly in the critical range of 280–380 nm, where damage caused by ultraviolet radiation occurs. Flavonoids have therefore been linked to the mechanism of plants for protection from ultraviolet radiation.

Many plant species synthesize phytoalexins upon invasion by microorganisms. The production of these antimicrobial compounds represents an active means of defense against potential pathogens. The majority of phytoalexins produced by legumes are isoflavonoids, and each plant species seems to produce a specific compound. Thus peas (*Pisum sativum*) produce pisatin (**6**), beans (*Phaseolus vulgaris*) produce phaseollin (**7**), and soybeans (*Glycine max*) produce glyceollins (**8**) as part of their defense mechanism.

Anthocyanins. Because of their strikingly vivid color, ranging from deep red through purple to deep blue, anthocyanins represent the most visible class of flavonoid compounds. Cyanidin is probably the most widely distributed anthocyanin. Other common anthocyanidins are pelargonidin, peonidin, delphinidin, petunidin, and malvidin, all of which occur in plants as glycosides. Anthocyanins, the glycosides of anthocyanidins, are most obvious in flowers and fruits, but they are also present in roots, stems, leaves, seeds, and other parts of the plant. In autumn,

when the cool nights slow down respiration, photosynthate accumulates, chlorophyll breaks down, and anthocyanins accumulate in the leaves. The accumulated anthocyanins, together with carotenes, provide the varied colors characteristic of that season. Anthocyanins are also produced when plants are subjected to other stress, such as ultraviolet radiation, injury by insects, malnutrition, or unusual concentrations of metal ions.

Color variation in plants results from many factors. Variation in color intensity may be caused by differences in concentration of the same pigment or by variations in the acidity of the cell sap. Anthocyanins undergo dramatic color changes when the acidity of their solution is varied, a response caused by reversible structural transformation reactions (**Fig. 7**).

In the pH range of 1 and below, all anthocyanin molecules of a solution are present in the flavylium cation form, which is red. The color of the flavylium cation is caused by a positive charge in the 2 position of the molecule and the double bond between carbons 3 and 4, which permits electron flow between the two aromatic rings A and B. As the pH of an anthocyanin solution is raised to between 4 and 4.5, the flavylium cation may undergo two different reactions: the nucleophilic attack of water in the 2 position, which yields the colorless carbinol base; and the relocation of electrons from ring B or A to produce the corresponding anhydro bases with the 4′-keto or 7-keto structures. Both have a purplish color that increases in stability as the pH rises toward neutrality. The anhydrobases also undergo hydration, which results in the formation of the colorless carbinol base. Around neutrality, the carbinol base undergoes ring fission to produce the yellow chalcone. All such structural transformation reactions are freely reversible, and the red color of an anthocyanin solution can be reestablished fully upon readjusting the pH to 1 or below.

Flower petals exhibit great color variation that ranges from pink to blue. Whereas in old, senescing petals the blue color may be due to a rise of pH in the vacuole, the color of most flowers is caused by complex formation of the anthocyanins with other phenolic, or flavonoid, compounds.

The varied colors of fruits, on the other hand, are not caused by complexes with other compounds but by variation of the pigment concentration in the tissues and cells, and the pH of the cell sap. Thus the blue color of plums or some grapes is not caused by a blue pigment in the cells but by a combination of a very high concentration of anthocyanins that reflects light and a waxy layer that helps to diffuse it. *See* PHYTOCHROME; PLANT METABOLISM. Geza Hrazdina

Bibliography. M. Dean, *Naturally Occurring Oxygen Ring Compounds,* 1963; J. B. Harborne, *The Flavonoids: Advances in Research since 1980*, 1988; J. B. Harborne and T. J. Mabry (eds.), *The Flavonoids*: *Recent Advances in Research*, 1982; C. F. Van Sumere and P. J. Lea (eds.), *The Biochemistry of Plant Phenolics*, 1985.

Flavor

Any of the six different varieties of quarks. All hadronic matter is composed of quarks, the most elementary constituents of matter. The six different flavors are labeled u, d, s, c, b, and t, corresponding to up, down, strange, charmed, bottom, and top. Quarks are all spin-$^1/_2$ fermions. The u, c, and t flavors carry a positive electric charge equal in magnitude to two-thirds that of the electron; the d, s, and b flavors have a negative charge one-third that of the electron. Different flavored quarks have vastly different masses ranging from the lightest, the u quark, with a mass around 5 MeV/c^2 (where c is the speed of light), equal to the mass of about 10 electrons, to the top quark, with a mass 35,000 times greater, or 175 GeV/c^2, about the mass of a gold atom. Because of its large mass, only in 1994 was evidence presented for existence of the top quark. Quarks of any flavor are further characterized by three additional quantum numbers called color: red, green, and blue. Each quark has an antiquark counterpart carrying the corresponding anticolor. All of these properties have been well established by experiment. *See* COLOR (QUANTUM MECHANICS).

Physically realizable particles are colorless. For example, the proton, composed of two u- and one d-flavored quark, has one red, one green, and one blue quark to make the ensemble colorless. The fact that quarks carry color and physical entities must be colorless accounts for their never appearing as isolated objects. Mesons are composed of quark-antiquark (q,$\mathbf{q}^-$) combinations. A positively charged pion consists of a u and anti-d quark; the anticolor of the anti-d quark cancels the color of the u quark. *See* BARYON; MESON; PROTON.

The strong interquark forces are produced by the exchange of massless, spin-1 bosons called gluons. All evidence suggests that these forces are flavor independent despite the vastly different masses of quarks of different flavors. Furthermore, electromagnetic forces are also flavor independent. In contrast, the weak interactions can induce a change in flavor. For example, at the quark level the beta decay of the neutron is viewed as the decay of the d quark to a u quark plus an electron and a neutrino. *See* GLUONS.

The six different quark flavors are frequently grouped with the leptons into three so-called generations, as shown below, where their electrical

$$\begin{bmatrix} u \\ d \\ e \\ \nu_e \end{bmatrix} \begin{bmatrix} c \\ s \\ \mu \\ \nu_\mu \end{bmatrix} \begin{bmatrix} t \\ b \\ \tau \\ \nu_\tau \end{bmatrix} \begin{matrix} {}^2/_3 \\ -{}^1/_3 \\ -1 \\ 0 \end{matrix}$$

charges, row by row, are given on the right. There is a corresponding set of antiparticles, quarks and leptons carrying the opposite charge. The weak interactions between quarks of the same generation are highly favored. For example, a t quark preferentially decays to a b quark, a key to the ultimate identification of the t quark. In the standard model of the electroweak interactions, neutral

currents cannot cause a change in the quark flavor. *See* ANTIMATTER; ELECTROWEAK INTERACTION; ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS; LEPTON; QUARKS; STANDARD MODEL; WEAK NUCLEAR INTERACTIONS. Val L. Fitch

Flax

The flax plant (*Linum usitatissimum*) is the source of two products: flaxseed for linseed oil and fiber for linen products. Plants with two distinct types of growth are used for seed and fiber production (**Fig. 1**, but fiber is sometimes a by-product of linseed production, and seed is an important by-product of fiber flax. Flax is an annual plant with a stem size that varies between about $^1/_{16}$ and $^1/_4$ in. (1.6 and 6.4 mm) in diameter. The flax flower has five petals, which are blue, white, or pale pink (**Fig. 2**). The color of the flower depends on the variety. Seeds of most varieties are varying shades of brown and yellow. Flax must be grown in rotation with other crops and not planted on the same land more frequently than once every 4 to 6 years; otherwise disease-producing organisms build up in the soil and yields may be greatly reduced unless resistant varieties are grown. *See* SEED; STEM.

Flax for fiber. Flax for fiber requires fertile, well-drained, and well-prepared soil and a cool, humid climate. It is planted in the early spring at 80 lb/acre (70 kg/hectare) or more. The high rate of seeding provides a dense stand of slender stems that branch only at the very top. Some chemical weed control may be necessary. A few weeds such as wild oats, cereal crops, perennial thistles, or morning glories will destroy the value of the crop for fiber, and all fields badly infested with weed seeds of any kind must be avoided for planting fiber flax. *See* AGRICULTURAL SOIL AND CROP PRACTICES; HERBICIDE.

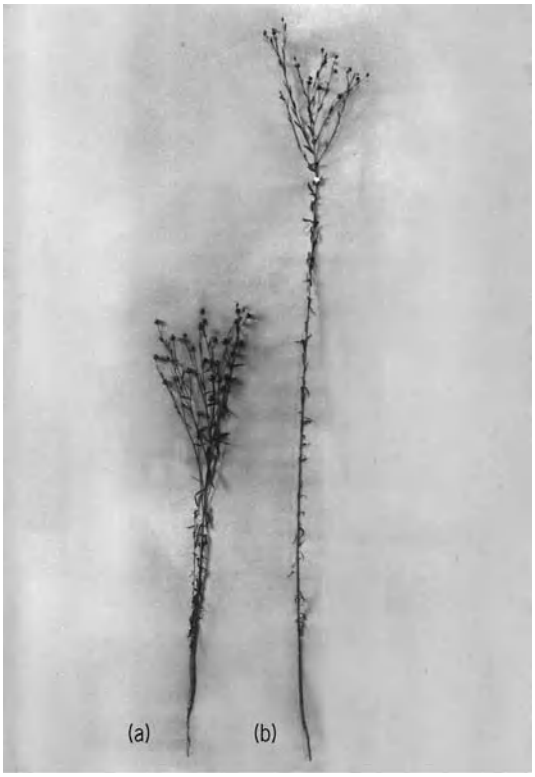

Fig. 1. Types of flax plant. (*a*) Seedflax. (*b*) Fiber flax. (*USDA*)

Fig. 2. Flowering top of flax plant showing buds and seed capsules. (*USDA*)

Harvesting. When about half the seed pods (bolls) have turned yellow, harvesting must be started. If the bolls are allowed to reach full maturity before harvesting, quality is greatly reduced.

Fiber flax is harvested by a machine that pulls it out of the ground and ties it in bundles. Some flax is still pulled by hand. Stems are kept parallel while the seed is removed; bundles of the straw are then packed vertically in concrete tanks and covered with warm water to ret, a process in which bacterial action loosens the fiber in the stem. The bundles of wet, retted straw are set up with the tops held together and the root ends spread out so the bundles form "gaits" or "wigwams" to dry in the sun. The dried flax is retied in bundles and stored under cover until it is scutched, a process in which the straw is first crushed to break the woody portion of the stem into small pieces (called shives) and is then subjected to a beating operation that removes the outer bark and the woody parts from the line (long parallel) fiber. During the beating operation, short tangled fiber called tow also comes out of the line fiber with the shives and must be further cleaned. In Western Europe the scutching operation is mechanized. There, the straw is spread in a thin layer, held near the center by a gripping device on a narrow belt, and carried between fluted rollers to break the woody parts into shives. The shives are then beat out of one end, the device grips the cleaned portion, and the other end is cleaned, leaving long (16–40 in. or 40–100 cm) strands. *See* AGRICULTURAL MACHINERY.

Uses. The best-known use for flax fiber is in the manufacture of fine linen fabrics; other uses are

linen thread, linen twine, toweling, and canvas. Seed from fiber flax is used for replanting and for oil. The major producer of flax fiber is Russia, but the world's best fiber comes from Belgium and adjoining countries. Most Irish linen is manufactured from fiber produced in Belgium. *See* LINEN; NATURAL FIBER; TEXTILE.

Flax for seed. The sowing rate for seedflax is lower than for fiber flax, about 40–50 lb/acre (45–55 kg/ha). The varieties used produce short straw and a relatively high yield of seed with 35–44% oil. Machines used for planting and harvesting cereals are also used for flax for seed. Planting is usually done in the early spring on a well-prepared seedbed on land relatively free from weed seeds. A small acreage of flax is sown in the fall in Texas. Herbicides may be used to control weeds. *See* CEREAL; FAT AND OIL.

The world's major producers of flaxseed are Argentina, Canada, India, Russia, and the United States. Most of the flaxseed acreage in the United States is in Minnesota, North Dakota, and South Dakota.

The principal uses of linseed oil are in the manufacture of protective coatings—paint, varnish, and lacquer. It is also used in linoleum, printer's ink, and patent and imitation leathers, and as core oil for making sand forms in metal casting. Uses that promise to become of great importance are as antispalling compounds for concrete highway construction and as curing agents for concrete. The linseed cake remaining after the oil is extracted is used as a high-protein livestock feed. Large quantities of clean (weed-free) straw from seedflax are used in the production of a high-grade paper. This product is used to make cigarette and other fine papers. Elton G. Nelson; Joseph O. Colbertson

Diseases. Of nearly two dozen diseases of oil and fiber flaxes throughout the world, fungi cause about 16 diseases, a virus causes crinkle, a mycoplasma-like organism causes aster yellows, a chlorophylless seed plant known as dodder (*Cuscuta epilinum*) is parasitic on flax stems, and several nonparasitic agents account for other diseases such as heat canker and iron chlorosis.

Diseases of flax can result either in complete losses of stands or in reduced yields at harvest, depending upon the disease, cultural practice, seed quality, environment, and resistance in the flax cultivar. Diseases destructive in one locality may be of negligible importance in another. For example, pasmo is important in Minnesota and the Dakotas, but it is of little importance in California and is unknown in Mexico; similarly, stem break (caused by *Polyspora lini*) is troublesome in Europe but the disease is rarely found in the United States.

Rust (*Melampsora lini*) is potentially the most destructive disease of flax wherever flax is grown in the world, primarily because its airborne spores can generate an epidemic rapidly within a season in one area. Severe rust occurred in North America in 1939–1941, 1944–1945, 1949–1951, and 1963, following the introduction of cultivars that proved to be susceptible to rust. Because no one cultivar is resistant to all of the more than 300 races of rust, cultivars must be grown that are resistant to the races prevalent in a given area, and this is the only practical method of control.

Wilt caused by *Fusarium oxysporum* was the limiting factor in flax production in the United States early in this century. Complete losses in seedling stands occurred in commercial cultivars (**Fig. 3**). As wilt resistance was bred into cultivars, some with moderate seedling resistance wilted after flowering, resulting in 8–20% losses in yield. The fungus is seed- and soil-borne, and once established in soil it survives there for decades, even as long as 50 years in the absence of flax. This makes crop rotation and

Fig. 3. Wilt resistance plot in Minnesota. Three rows of susceptible cultivar in the center are compared with rows of resistant cultivars on each side, in a plot where flax has been grown each year for 50 years.

Fig. 4. Pasmo disease of flax. Dark-colored bands indicate places of infection with the pasmo fungus which girdles stems and kills plants.

Fig. 5. Blight of flaxseed caused by several fungi. (*Minnesota Agricultural Experiment Station*)

fungicides ineffective; breeding for resistance is the only feasible method of control, and this method is eminently successful today.

Pasmo, caused by *Septoria linicola*, occurs in Europe, North and South America, and New Zealand, and was introduced into the United States from Argentina about 1915. It became destructive within 30 years. The stems turn brown, often in bands (**Fig. 4**), and leaves fall off, and at harvest the seeds are shriveled. Because no resistant cultivars are known, crop rotation and sanitation are the only possible control measures.

Seedling blight caused by several fungi (*Alternaria* spp. *Colletotrichum lini, Pythium* spp., and *Rhizoctonia solani*) and root rot caused by *R. solani* can be important at certain times and places. Sometimes seed condition determines whether seedling blight and root rot occur. Damaged flaxseed, known also as cracked, chipped, split, fractured, weathered, or blighted seed, is common in all flax-growing regions of North America (**Fig. 5**). Sowing damaged seed can result in 10–50% lower stands and sometimes 20% lower yields because organisms in the soil can enter openings in damaged seeds or infect crippled seedlings growing from damaged seeds. The best control for seedling diseases is to select sound seeds and treat them with an appropriate fungicide. Some seed-borne, seedling blight fungi are also controlled by fungicides.

Only one virus is important on flax, the oat blue dwarf virus, which causes crinkle. A mycoplasma-like organism (*Spiroplasma kunkelii*) causes aster yellows in flax. Both of these pathogenic agents reduce seed yields and are transmitted by the same vector, the aster leaf-hopper. Control is by growing resistant cultivars. *See* PLANT VIRUSES AND VIROIDS.

There are no important bacterial or nematode diseases of flax. Flax is susceptible to both hail and insect damage which predispose plants to disease. *See* PLANT PATHOLOGY. Thor Kommedahl

Bibliography. H. H. Flor and V. E. Comstock, Flax cultivars with multiple rust-conditioning genes, *Crop Sci.,* 11:64–66, 1971; T. Kommedahl, J. J. Christensen, and R. A. Frederiksen, *A Half Century of Research in Minnesota on Flax Wilt Caused by Fusarium oxysporum,* Minn. Agr. Exp. Sta. Tech. Bull. 273, 1970; P. Neergaard, *Seed Pathology*, vols. 1 and 2, 1977.

Flexibilia

An extinct subclass of stalked or creeping Crinoidea which includes some 50 Paleozoic genera, ranging from Ordovician to Permian times. A flexible tegmen was present with open ambulacral grooves. One of the five oral plates served as a madreporite. The anus was at the tip of a short siphon. About three of the lower brachial ossicles of each arm were loosely incorporated into the dicyclic calyx. The three infrabasals were sometimes obscured by the stem, but the five basals and five radials remained conspicuous and movably articulated. The uniserial arms lacked pinnules but branched freely, arching inward to form a globular crown (see **illus.**). The cylindrical stem

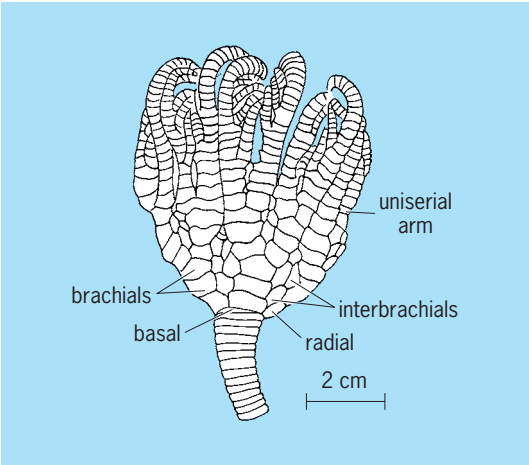

***Talanterocrinus* sp., a member of the Flexibilia.**

lacked cirri and was sometimes discarded, as in *Edriocrinus*, to produce a creeping adult. *See* CRINOIDEA; ECHINODERMATA. Howard B. Fell

Flexible manufacturing system

A factory or part of a factory made up of programmable machines and devices that can communicate with one another. Materials such as parts, pallets, and tools are transported automatically within the flexible manufacturing system and sometimes to and from it. Some form of computer-based, unified control allows complete or partial automatic operation.

Programming. The programmable machines and devices are numerically controlled machine tools, robots, measuring machines, and materials-handling equipment. Each programmable machine or device typically has its own controller which is, in effect, a dedicated digital computer; programs must be written for these controllers. For example, a numerically controlled machine needs a program which supplies the sequence of machining operations to be carried out on a given part. A robot needs a program which specifies where and how it must move and employ its gripper. These programs are usually written in special-purpose languages designed to handle the geometry and machining. Increasingly, numerically controlled machines are being programmed by graphical presentations on computer screens, that is, graphical computer interfaces. This allows the programmer to follow the machining operation and specify desired operations without the need for statements in a programming language. Robots have usually been programmed by so-called teaching, where the robot is physically led through a sequence of movements and operations; the robot remembers them and carries them out when requested. Robots are also programmed with robot-programming languages and, increasingly, with graphical computer interfaces. Similar choices in programming method are available for other devices. Programming a flexible manufacturing system may involve several different languages and methods, resulting in complicated and expensive software development and maintenance. *See* COMPUTER GRAPHICS; COMPUTER NUMERICAL CONTROL; COMPUTER PROGRAMMING; INTELLIGENT MACHINE; MATERIALS-HANDLING EQUIPMENT; PROGRAMMING LANGUAGES; ROBOTICS.

Communication. The programmable machines and devices communicate with one another via an electronic connection between their controllers. Increasingly, this connection is by means of local-area networks, that is, communication networks that facilitate high-speed, reliable communication throughout the entire factory. These connections allow the programs in the various programmable devices to interact. For example, when a numerically controlled machine has completed a part, it can ask the robot to put it on a vehicle. *See* LOCAL-AREA NETWORKS.

Material transport system. The automatic material transport system is usually a guided, computer-controlled vehicle system. The vehicles are usually confined to a fixed network of paths, but typically any vehicle can be made to go from any point in the network to any other point. The network is different from a classical assembly line in that it is more complex and the flow through it is not in one direction. Paths in the system can be tracks to which the vehicles are confined or a marker, such as wire buried in the floor, that a vehicle senses and follows.

Since the vehicles must transport various kinds, shapes, and sizes of objects to various places, a way to accommodate this automatically is needed. One widely used method is to employ pallets standardized in the sense that common fittings allow each pallet to mount onto each vehicle and each location. For example, all pallets might be designed to mount on four vertical conical pins, or all might have the same slots machined into their sides. There are other possibilities, but in all cases a standardized mounting is the key. Beyond this, it must be possible to attach each object to be transported to a pallet, sometimes automatically and sometimes manually. Typically a part is attached to a pallet that is loaded into the system, picked up by a vehicle, and loaded into the numerically controlled machine.

Unified control. Commands and orders to the flexible manufacturing system are sent to its computer-based, unified control. For example, an order can be given to make 25 parts of type 76-998. The control, in turn, issues orders for the transport of various kinds of material, the transfer of needed programs, the starting and stopping of programs, the scheduling of these activities, and other activities. Typically, the control is in a central computer dedicated mainly to it; this computer is usually not a device controller.

Flexibility. Flexible manufacturing systems are flexible in the sense that their device controllers and central control computer can be reprogrammed to make new parts or old parts in new ways. They can also often make a number of different types of parts at the same time. However, this flexibility is limited. Each system is designed with a certain family of parts in mind, for example, axles. It can often be reprogrammed to a make a new kind of axle, and it can usually make a number of different kinds of axles at the same time. But its flexibility would almost certainly not extend to making engine blocks.

Advantages. Various rationales for flexible manufacturing systems have been given. The reduction of the work force involved is rarely the most important one. In fact, direct labor costs are a relatively small part of total costs in many manufacturing systems, particularly those that are candidates for conversion to flexible system. Thus, direct labor reduction usually has a minor impact on total costs.

Because the flexible manufacturing system is under continuous computer control, the movement of parts, material, and tools is efficient in the sense that neither time nor space is wasted. If a part needs to be moved it is done immediately. Such a performance is difficult to obtain in nonautomated systems other than highly structured, inflexible ones. Therefore, the number of partially completed parts that have to be in the system (referred to as work in progress) for a given level of production is significantly reduced, resulting in major cost reductions. For example, if producing 1000 axles a month

requires a work in progress of 10,000 partially completed axles and if introduction of a flexible manufacturing system reduces work in progress to 2000 axles, the cost of carrying an inventory of 8000 partially completed axles will be saved. *See* INVENTORY CONTROL.

Similarly, the processing time from raw material to finished part is reduced, and this allows for a reduced inventory of finished products and quicker response to market demands. The first reduces inventory costs, and the second, though more difficult to quantify, improves competitiveness.

Computer control also reduces the idle time for the machines. Moreover, flexible manufacturing systems are increasingly designed to continue unattended during nights, weekends, and holidays. This is an important source of additional production. Fewer machines and less floor space are needed.

Flexible manufacturing systems typically produce a better product for several reasons. The controlling computers execute the required programs day and night with few mistakes. Since the mechanical handling of parts is automatic, it must be designed to be done repeatedly with very few mistakes. Typically, quality control measurements are built into the system. This is important because it allows deviations to be detected, diagnosed, and corrected at once, often with little interruption in production. Moreover, computer control allows problems to be anticipated and preventive actions to be carried out. For example, many systems keep a history for each tool and can replace it in anticipation of breakage. In addition to a better-quality product, this reduces the number of parts that have to be discarded or reworked, and translates into higher production or a smaller system. *See* QUALITY CONTROL.

In addition, flexible manufacturing systems allow production levels to be easily and quickly changed; the production of sample new designs can be intermingled with ongoing production, and product composition can be readily changed. Product changeovers are also easier. Furthermore, they allow a reduction in indirect labor because complete information about manufacturing operations is available immediately from the computer.

Disadvantages. Even though flexible manufacturing systems have existed for over 40 years, they still are complex systems that are difficult to create, especially if flexibility is emphasized. For example, they require high skills and training, particularly with respect to computers. Development, maintenance, and modification of all needed computer programs often present major difficulties, which is the main reason that flexibility of programmable devices is often not fully exploited.

Flexible manufacturing systems are best suited for specialized manufacturing. On the one hand, they are not viewed as a replacement for so-called hard automation as exemplified by the assembly line. On the other hand, they are not intended to compete with the small-job shop. In the first case, flexible manufacturing systems have unneeded and expensive flexibility, and in the other they are not flexible enough.

It has also been widely noted that introduction of flexible automation usually requires a major change in the way that a firm organizes and manages its staff; the difficulty of defining and making these changes is comparable with that presented by programming.

Renewed attention in the late 1980s to the management of manufacturing systems with automation levels significantly lower than those of flexible manufacturing systems, with particular attention to personnel management, resulted in important improvements in productivity and quality.

Computer-integrated manufacturing. Flexible manufacturing systems are part of a larger computer-based technology, termed computer-integrated manufacturing (CIM), which encompasses more than the movement and processing of parts on the factory floor. For example, computer-aided design (CAD) of products is a major part of CIM, as are production and capacity planning. Indeed, the scope of CIM is wide, including the computerization of a manufacturing company and its suppliers and customers. *See* AUTOMATION; COMPUTER-AIDED DESIGN AND MANUFACTURING; COMPUTER-INTEGRATED MANUFACTURING.

Arch W. Naylor

Bibliography. D. M. Considine and G. D. Considine (eds.), *Standard Handbook of Industrial Automation*, 1986; B. Luggen, *Flexible Manufacturing Cells and Systems*, 1990; R. Maleki, *Flexible Manufacturing Systems: Technology and Management*, 1990; A. Ross, *Dynamic Factory Automation*, 1992; H. Tempelmeier and H. Kuhn, *Flexible Manufacturing Systems: Decision Support for Design and Operation*, 1993; D. Tijunelis and K. E. McKee, *Manufacturing High Technology Handbook*, 1987; H. J. Warnecke and R. Steinhilper (eds.), *Flexible Manufacturing Systems*, 1989.

Flight

The movement through the atmosphere or space, sustained by aerodynamic reaction or other forces. Animal flight includes gliding and flapping flight. Four animal groups evolved flapping flight, namely insects, pterosaurs, birds, and bats. Flapping flight in vertebrates was probably preceded by gliding; in insects it may have originated by leaping and gliding, by surface skimming on water, or (if small enough) by passive floating in the air. Flying insects show greater variation than flying vertebrates, and their flight spans a wider range of Reynolds numbers, which is the ratio of inertial forces to viscous forces in the flow. Flight of tiny insects is in the lower range of Reynolds numbers, where viscous forces are dominant, whereas large insects and vertebrates operate in the higher range, where inertial forces are important.

Flight is very expensive in terms of energy cost per unit time. However, flying animals can travel several times faster than nonflying ones, and the cost of carrying a unit of weight of an animal through a unit

distance (cost of transport) is lower for flight than for running; but higher than for swimming.

Characteristics. The flight characteristics of large insects and vertebrates can be understood by aircraft aerodynamics. In steady level flight, an animal and an aircraft must generate forces that support weight against gravity and provide propulsive thrust against drag forces. The forces acting on an airfoil (the shape of the cross section of the wings) moving through the air depend upon the flow pattern around it. Because of the asymmetric profile on an airfoil, the air flowing over the upper surface travels farther and flows faster than air passing underneath. According to Bernoulli's principle, pressure falls when speed rises in a moving fluid, resulting in a pressure difference between the upper and lower sides of the airfoil **(Fig. 1*a*)**. This pressure difference has to disappear gradually toward the wingtips, and some air will flow upward around the wingtips. The air moves downward behind the wing as trailing vortices (Fig. 1*b*), and the reaction of this momentum flow is experienced by the wing as lift. The stronger the vortex, the greater the lift generated, but with some energy loss to drag. The lift force is responsible for weight support (its vertical component) and thrust or drag (its horizontal component). *See* AERODYNAMIC FORCE; AIRFOIL; BERNOULLI'S THEOREM.

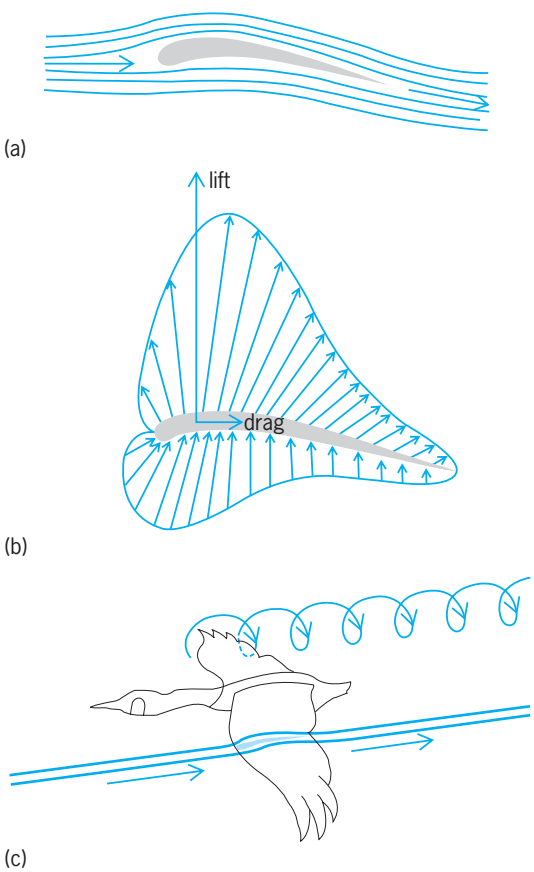

Fig. 1. Aerodynamics of flight. (*a*) Airflow around a typical wing profile. (*b*) Pressure distribution around a typical wing profile. (*c*) The difference in pressure disappears toward the wingtips as trailing vortices in the wake. (*After M. Brooke and T. Birkhead, eds., Cambridge Encyclopedia of Ornithology, Cambridge University Press, 1991*)

Energy expenditure. Energy must be expended to generate the trailing vortices in the wake and to overcome friction on the wing and body surfaces. These energy losses are experienced as drag forces, which act parallel to the direction of movement of the airfoil. The drag caused by trailing wingtip vortices is conventionally termed induced drag, the frictional drag of the wings is termed profile drag, and the frictional drag of the body is the parasite drag. The magnitude of lift and drag varies with the velocity of the flow, the size and shape of the wing, and the angle at which the flow meets the airfoil (angle of incidence). When the stalling angle is exceeded, the airstream separates from the upper surface of the wing with a sudden fall of lift and increase of drag. A good airfoil maximizes the pressure difference between its upper and lower sides and minimizes the drag: the lift:drag ratio of avian wings averages between 5:1 and 10:1.

Since drag is a retarding force, the animal must either descend (glide) through the air at such an angle that a component of its weight balances the drag, or do mechanical work with its flight muscles (flap its wings) to overcome it. The rate at which this work is done is the mechanical power required to fly, and it equals speed times drag [measured in watts (W)]. The flight muscles also produce heat when they contract, so the total metabolic power expenditure is the sum of this heat loss and the mechanical power. The metabolic power is estimated to be four to five times the mechanical power, and is dependent on the size of the animal. *See* WORK.

Gliding flight. Compared with active flight, gliding flight is very inexpensive, and is found in a wide range of animals, such as squirrels, marsupials, lizards, frogs, fishes, and even one snake. It is the main component in soaring flight used by many birds and some bats, when the animals use thermals or updrafts. Gliding in birds costs only two to three times the basal metabolic rate, because the flight muscles do not perform any mechanical work other than for stability and control of movements, and produce mostly static forces to keep the wings down on the horizontal plane, opposing the aerodynamic force.

When gliding, the wings of the animal leave behind a continuous vortex sheet that rolls up into a pair of vortex tubes (wingtip vortices), as in fixed-wing airplanes (Fig. 1*b*). The lift produced balances the animal's weight, but potential energy must be used to overcome the drag and the animal loses height. An animal gliding at steady speed descends at an angle to the horizontal (glide angle), established by the ratio of lift to drag (glide ratio). The best glide ratios in birds range from 10:1 to 15:1 for vultures and birds of prey and reach 23:1 in the wandering albatross, whereas modern gliders can achieve 45:1. An animal cannot glide more slowly than its stalling speed, which in birds can be reduced by splaying the wingtip primaries, or by spreading the alula (a digit of the wing) at the wrist, or both. An animal can increase its gliding speed by flexing the wings and reducing the wing area.

Flapping flight. In horizontal forward flight, the velocity of the air that meets a wing section depends on the forward speed of the animal and the flapping speed of the wing section. The local forces on the wings are therefore different along the wing and vary throughout the wing-beat cycle. In the downstroke the inner parts of the wing produce small lift and drag forces, and the outer parts mainly lift and thrust. The structure of the wake varies with flight speed. As the wings travel in flapping flight, they leave behind them a vortex sheet that is formed during the first half of the lift-producing downstroke. During the upstroke the wings are more or less flexed to reduce drag, have no bound vorticity, and contribute little or no lift. The rapid acceleration and deceleration of the wings at the beginning and end of the downstroke alters the circulation on the wing, and the shed trailing vortices then close into vortex loops, one for each downstroke (**Fig. 2***a*). During faster flight at low wing-beat amplitude and frequency, the vortex wake consists of a pair of continuous undulating vortex tubes or line vortices behind the wingtips (Fig. 2*b*), as in kestrels in cruising flight, indicating that also the upstroke provides useful aerodynamic forces.

Hovering flight represents the most expensive form of animal flapping flight. The essence of hovering is to produce a vertical force to balance body weight. The wake consists of a chain of vortex rings continuously shed during the wing strokes. In hummingbirds and insects, lift is produced during both the downstroke and upstroke of the wings (symmetrical hovering), and two vortex rings are produced during each wing stroke (Fig. 2*c*). In other hovering animals the wings are flexed during the upstroke (asymmetrical hovering), and the rings are produced during the downstrokes only (Fig. 2*d*).

Wing shape and flight behavior. Different foraging behavior and associated flight modes put different demands on body size and wing form. The relative size and shape of the wings can be quantified by the wing loading and the aspect ratio. Wing loading is estimated as the ratio of body weight to wing area [in units of newtons (N) per square meter]. It increases with body size and ranges from around 10 N/m^2 in small flycatchers and owls to 230 N/m^2 in the whooper swan. Flight speed is proportional to the square root of wing loading, and the radius of a banked turn is proportional to wing loading. Slow and maneuverable flight is thus obtained by a low wing loading.

Aspect ratio is calculated as the wing span squared divided by wing area. It is dimensionless and ranges from around 4.5 in the great tit and some galliforms and tinamous to around 18 in albatrosses. Aspect ratio is a measure of the aerodynamic efficiency, and lower energy cost of flight is obtained in animals with wings of high aspect ratio, and particularly in combination with a low wing loading (long, narrow wings).

Many seabirds, swifts, and swallows have long and narrow wings, and they can thus afford to fly continuously during foraging. Animals with short and narrow wings, that is, with higher wing loadings (including loons, mergansers, ducks, auks, and free-tailed bats) are capable of fast and, for their speed, rather economical flight, used for commuting and migration.

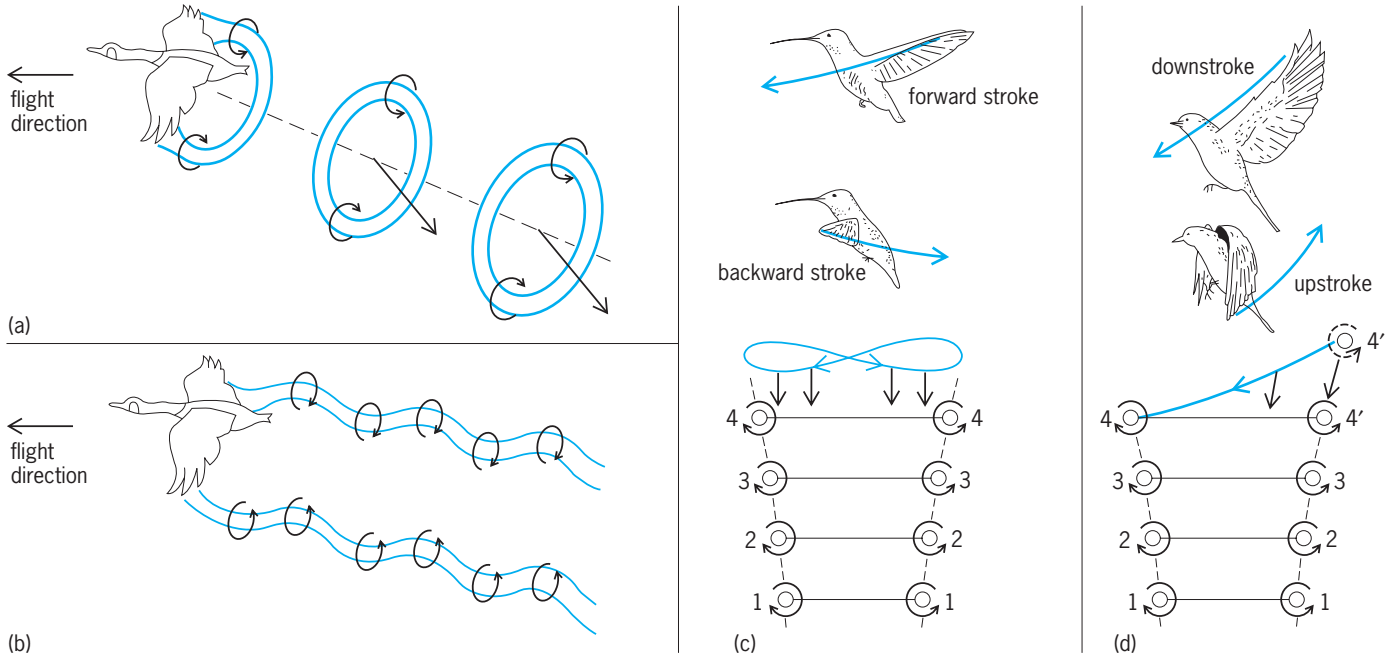

Fig. 2. Vortex wakes behind a bird in flapping flight, and below a hovering bird. (*a*) In slow flight, vorticity is shed only during the downstroke, resulting in a stack of vortex rings. (*b*) In faster flight, lift can be generated also during the upstroke, and the wake consists of an undulating pair of line vortices (*after M. Brooke and T. Birkhead, eds., Cambridge Encyclopedia of Ornithology, Cambridge University Press, 1991*). (c) Formation of vortex rings in four wingstrokes in symmetrical hovering and (*d*) in asymmetrical hovering. (*After U. Norberg, Vertebrate Flight, Springer, 1990*).

Soaring birds and animals flying among vegetation benefit from low weight and broad wings for slow flight and high maneuverability. Birds in this group often have a large alula and the ability to separate the wingtip primaries, so as to increase lift at large angles of incidence, thus permitting slow speeds. The very low aspect ratios in animals flying among vegetation are connected with high flight costs, and the animals are, therefore, usually perchers or spend much of their foraging time walking, climbing, and hanging. *See* AERODYNAMICS; AVES. Ulla M. Norberg

Bibliography. M. Brooke and T. Birkhead (eds.), *Cambridge Encyclopedia of Ornithology*, 1991; R. F. Johnson (ed.), *Current Ornithology*, vol. 5, 1988; U. M. Norberg, *Vertebrate Flight*, 1990; C. J. Pennycuick, *Bird Flight Performance*, 1989.

Flight characteristics

The behavior of an aircraft while airborne. An aircraft, while in the air, is free to move in six degrees of freedom (**Fig. 1**). The aircraft can translate along any of its three axes, X, Y, and Z. The corresponding translational velocities are identified as U (forward velocity), V (side velocity), and W (up or down velocity), respectively. In addition, the aircraft can rotate about any of its three axes, X, Y, and Z. The corresponding rotational velocities are labeled as P (roll rate), Q (pitch rate), and R (yaw rate) respectively. *See* DEGREE OF FREEDOM (MECHANICS).

The mission of most aircraft is to carry a payload from one point on Earth to another. To carry out that mission the aircraft must satisfy several requirements: (1) It should be able to develop enough lift to overcome its weight. (2) It should be able to develop enough thrust to overcome its drag. (3) It should have sufficient fuel on board to cover a useful range with a useful payload. (4) It should be able to take off and land on reasonably sized airfields. (5) It should be easily and safely maneuverable from one flight condition to another.

For any given airplane, these five requirements are given specific numerical values which depend on the mission of the aircraft. To predict whether or not a given aircraft design can meet these requirements, its flight characteristics are separated into two categories: performance and flying-quality (also called handling-quality) characteristics. Both are subject to airworthiness regulations by civilian and military authorities because of safety concerns. These regulations form important constraints on the design of aircraft and aircraft systems. *See* AIRCRAFT DESIGN.

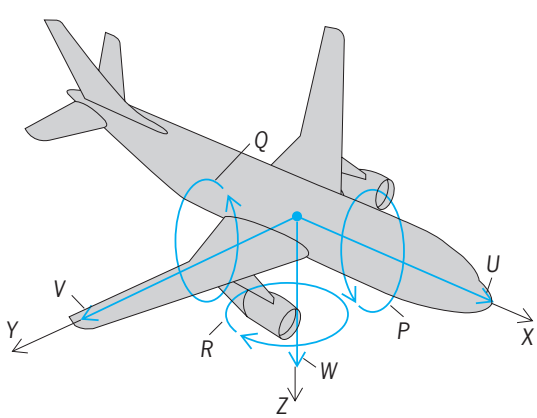

Fig. 1. Aircraft axes *X*, *Y*, *Z* and degrees of freedom *U*, *V*, *W*, *P*, *Q*, *R*.

Performance characteristics. To predict and describe aircraft performance characteristics, it is usually sufficient to address the following performance categories: takeoff distance, climb rate or climb gradient, time to climb to altitude, maneuverability, cruise speed and altitude, payload-range and fuel used, loiter, descent rate, descent time, and landing distance. These performance categories are normally summarized with the help of a flight envelope (**Fig. 2**) and a mission profile (**Fig. 3**). Cruise speed and payload-range will be discussed in some detail to highlight important aircraft design principles and design parameters.

Cruise speed. The cruise speed of an airplane depends on weight, drag, thrust, and altitude. The cruise drag, which must be overcome by thrust, is itself the sum of two contributions: drag at zero lift (mostly caused by friction of the air over the surfaces of the airplane) and drag due to lift (also called induced drag). The thrust required to overcome the cruise drag (in pounds or newtons) is referred to as the required thrust (T_{required}). The combined engine thrust should be able to overcome the cruise drag. The combined engine thrust at some thrust setting and altitude is referred to as the available thrust ($T_{\text{available}}$).

At the cruise speed, the thrust required and the thrust available are equal to each other. The cruise speed U can be determined from the airplane cruise weight; the zero-lift drag coefficient of the airplane; the air density at cruise altitude; the dimensions of the wing; the wing aspect ratio; and Oswald's efficiency factor, which accounts for the fact that lift distributions on practical wings are not perfectly elliptical. The cruise speed U can be determined graphically from a plot of required and available thrust (**Fig. 4**). *See* AERODYNAMIC FORCE; AIRFOIL; PROPULSION; THRUST.

Payload-range. Most aircraft are designed to carry a specified payload over a given distance, at a given speed and altitude, while using a minimum amount of fuel. It is important to understand the relationship between various weights [in pounds-force or newtons]. In the equation below W_{TO} is the airplane

$$W_{\text{TO}} = W_{\text{OE}} + W_{\text{PL}} + W_{\text{F}}$$

takeoff weight, W_{OE} is the airplane operating empty weight, W_{PL} is the payload weight, and W_{F} is the mission fuel weight. All airplanes are designed to some maximum allowable takeoff weight mostly to prevent overloading. The operating empty weight normally consists of the actual empty weight plus the weight of the crew needed to operate the airplane. The fuel weight consists of the fuel used during the

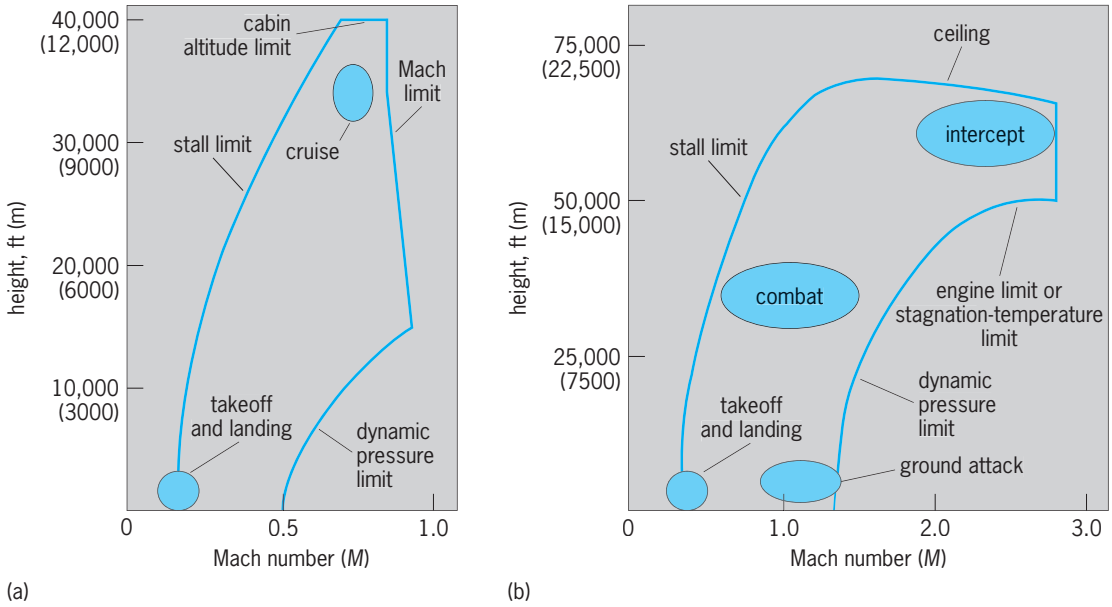

Fig. 2. Examples of operational flight envelopes. (*a*) Civilian airplane. (*b*) Military airplane.

mission plus the reserve fuel required by civil or military regulations.

The maximum allowable takeoff weight and the empty weight are defined in the flight manual for any specific airplane. The difference between the takeoff weight and the operating empty weight is referred to as the useful weight. From Eq. (1) it follows that the useful weight is the sum of payload weight and mission fuel weight. Operators of airplanes must decide which combination of payload and fuel to carry for a given mission. It is usually not acceptable to exceed the maximum allowable takeoff weight. Most airplanes also have a limited amount of fuel volume available.

The range (in nautical miles or kilometers) which can be covered by a jet airplane while cruising at constant speed and angle of attack may be computed from the airplane cruise speed, the specific fuel consumption, the lift-to-drag ratio in cruise, and the ratio of airplane weight at the beginning of cruise to that at the end of cruise. Roughly, the weight at the beginning of cruise may be thought of as equal to the takeoff weight. The weight at the end of cruise is then lower by the amount of fuel actually used. *See* SPECIFIC FUEL CONSUMPTION.

High cruise speed and low fuel consumption, coupled with a high lift-to-drag ratio and a high weight ratio, are conducive to long range. Operators of commercial and military transports are interested in the relationship between range and payload. This relationship is normally represented in the form of a payload-range diagram. The maximum range with full payload is referred to as the harmonic range. The range with maximum fuel (from a volumetric viewpoint) can be attained at a reduced payload only. The maximum range at zero payload is called the ferry range.

Flying-quality characteristics. In the case of airplanes that are flown by human pilots, the interaction between pilot cockpit control inputs and airplane response to the cockpit control inputs must be such that the pilot can achieve the mission objectives with reasonable physical and mental efforts. In other words, the airplane is required to have acceptable flying qualities (sometimes also called handling qualities) anywhere inside the operational flight envelope.

The flying qualities of an airplane must be such that

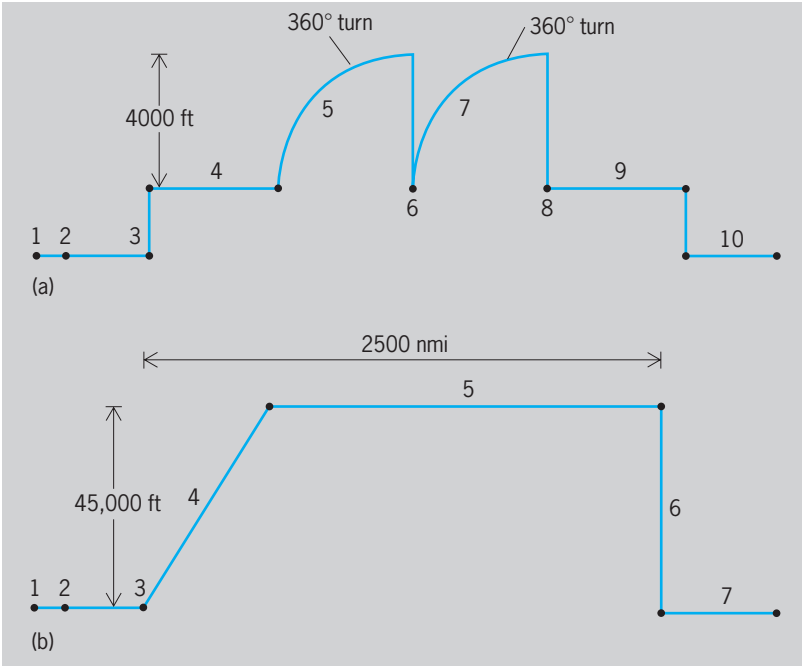

Fig. 3. Examples of typical mission profiles and flight phases. (*a*) Military attack airplane. Numbers indicate flight phases: (1) engine start and warmup; (2) taxi; (3) take off and accelerate to 350 knots (650 km/h) at sea level; (4) dash 200 nm (370 km) at 350 knots; (5) 360°, sustained, 4.5-*g* turn, including a 4000-ft (1200-m) altitude gain; (6) release two bombs and fire 50% ammunition; (7) 360°, sustained, 4.5-*g* turn, including a 4000-ft altitude gain; (8) release two bombs and fire 50% ammunition; (9) dash 200 nm at 350 knots; (10) land, taxi, shut down. (*b*) Civilian, passenger-transport airplane. Numbers indicate flight phases: (1) engine start and warmup; (2) taxi; (3) take off; (4) climb to 45,000 ft (13,500 m); (5) cruise; (6) descend; (7) land, taxi, shut down. 2500 nm = 4630 km.

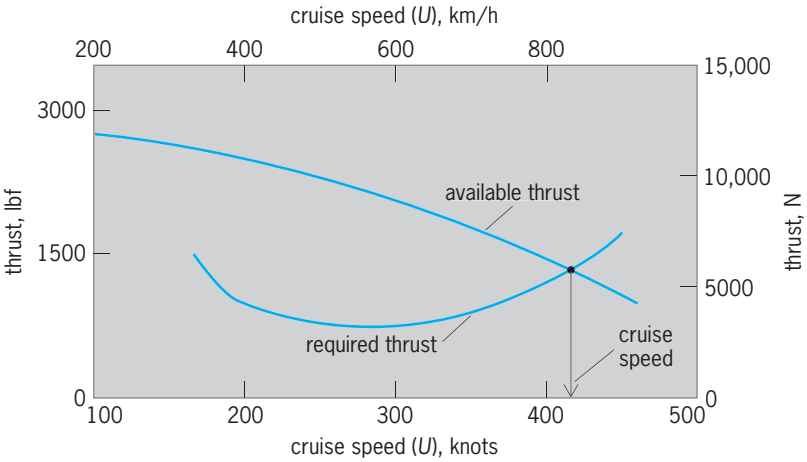

Fig. 4. Determination of cruise speed for a typical jet airplane. Thrust setting is for maximum continuous thrust in cruise. Data apply to one weight only.

the following characteristics are present anywhere inside the operational flight envelope:

1. The airplane must have sufficient control power to maintain steady-state, straight-line flight (level or not) as well as steady-state maneuvering flight consistent with mission objectives.
2. The airplane must be maneuverable from one (steady-state) flight condition to another.
3. The airplane must have sufficient control power to accomplish the transition from ground operations to airborne operations (takeoff run, lift-off, and initial steady-state climb), and the transition from airborne operations to ground operations (steady-state approach, touch-down, and landing run).

These three characteristics must be present whether the engine or engines are operating or not. Included are certain conditions of asymmetrical power which can be expected to occur. The whole is referred to as the minimum control speed condition.

The physical pilot efforts required to fly an airplane while meeting its mission objectives are expressed in terms of cockpit control activity and cockpit control forces. The cockpit control force magnitudes and their variation with certain motion parameters (examples are speed, load factor, and sideslip) are important to pilots. Examples of flying-quality parameters that deal with physical pilot efforts are maximum required stick force, stick force per g, stick-force–speed gradient, and aileron wheel force. None of these parameters should be too large or too small.

The mental efforts required to fly an airplane while meeting its mission objectives are usually referred to as pilot compensation. If an airplane reacts either too slowly or too quickly to a pilot command, the pilot must compensate for this behavior by adjusting his or her own gain or by leading the airplane.

An example of pilot gain is the amount of lateral cockpit control activity generated by a pilot to keep the wings level while flying in turbulence. A high gain in this case would mean a large wheel deflection to correct for a small bank-angle deviation. A low gain would mean the opposite. *See* GAIN.

The need for a pilot to lead an airplane may be illustrated by the example of a pilot flying an airplane with a high rolling moment of inertia. When the pilot moves the lateral cockpit controls to start a turning maneuver, the airplane will be slow to react to that input. However, once this airplane begins to roll, it will be equally reluctant to stop rolling once the bank angle required for a given turning maneuver is attained. The pilot must therefore neutralize the lateral cockpit controls well before the desired bank angle is reached. In other words, the pilot must lead the airplane.

A pilot should not have to excessively lead an airplane or have to provide either too much or too little gain during any given mission segment.

Equations of motion. To predict the dynamic response characteristics of an airplane in flight and thus assess the flying qualities requires a mathematical model that describes the motions of the airplane. Since the airplane in flight has three translational and three rotational degrees of freedom (Fig. 1), it takes six equations to describe the airplane motions. In nearly all cases it is possible to describe the dynamic response behavior of an airplane by assuming that the deviations (perturbations) of the airplane motion variables from a known steady-state (constant forward speed, U), straight-line flight condition are small. This is referred to as the small-perturbation assumption. As a result of this assumption, the equations of motion can be cast in the form of six linear differential equations with constant coefficients. The six equations in turn can be split into two sets of three equations each: the longitudinal equations (drag, lift, and pitching-moment equations) and the lateral-directional equations (side-force, rolling-moment, and yawing-moment equations) respectively. *See* DIFFERENTIAL EQUATION; PERTURBATION (MATHEMATICS).

Longitudinal open-loop response. One example of the motion characteristics which result from this mathematical model concerns the longitudinal motions of the airplane. It will be assumed that the pilot moves the longitudinal cockpit control (which, in turn, moves either the elevator or the stabilizer) in a step fashion. The airplane responds to this control input with small oscillations in speed, angle of attack, and pitch attitude angle. [The speed is the true airspeed of the airplane, that is, the velocity of the airplane relative to the air. The angle of attack is the angle between the speed or velocity vector and the X axis. The pitch attitude angle is the angle between the X axis and the horizon.] *See* ELEVATOR (AIRCRAFT); STABILIZER (AIRCRAFT).

Each oscillation itself is broken down into two suboscillations: one is characterized by a low frequency, and the other by a relatively high frequency. The low-frequency motion is referred to as the phugoid mode of motion of the airplane, while the high-frequency motion is referred to as the short-period motion. In most cases, from a flying-quality viewpoint, the short-period motion is the more important. If the short-period frequency is too low, pilots call the airplane sluggish. If the short-period frequency is too high, the airplane reacts too quickly.

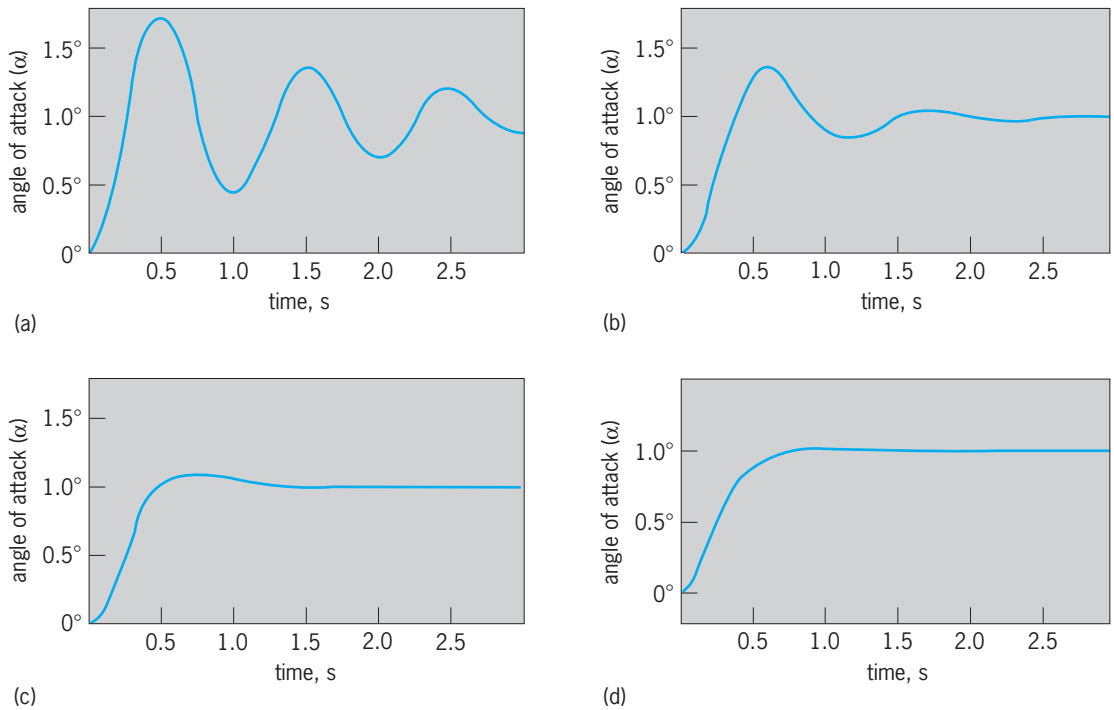

Fig. 5. Effect of short-period damping ratio. ($\zeta_{\text{short period}}$) on angle-of-attack oscillations. Typical oscillations are shown for $\zeta_{\text{short period}}$ equal to (*a*) 0.1, (*b*) 0.3, (*c*) 0.6, (*d*) 0.9.

For good control during various tracking tasks, the short-period frequency must lie between some maximum and minimum allowable values. The damping of the short period must also be acceptable (**Fig. 5**). For a damping ratio of 0.9 the airplane does not exhibit oscillatory behavior, but for a damping ratio of 0.1 there is much oscillatory activity. *See* DAMPING; MODE OF VIBRATION; VIBRATION.

Certain factors determine the short-period frequency and damping of an airplane. The short-period frequency may be estimated from factors that include the airplane lift-curve slope (the tangent of the angle between the lift and the angle of attack), the airplane center-of-gravity location, the airplane aerodynamic center location, the atmospheric density, the speed of the airplane, the wing area, and the pitching moment of inertia of the airplane. The frequency of the short period depends on the distance between the airplane aerodynamic center and the center of gravity, on the atmospheric density, and on the pitching moment of inertia. *See* MOMENT OF INERTIA.

At a forward center-of-gravity position, the airplane tends to have a higher frequency than at an aft center-of-gravity position. Also, at high altitude the frequency is significantly lower than at sea level, and airplanes tend to become sluggish. Finally, if the airplane configuration is such that large concentrated masses exist at the nose or tail (such as engines, propellers, radars, and jamming pods), the airplane would have a large pitching moment of inertia, and that also would tend to make the airplane sluggish.

The damping of the short-period oscillation is much affected by the size and the moment arm of the horizontal tail. Because the undamped natural frequency is itself proportional to the square root of atmospheric density, the damping ratio is also. Therefore, the damping ratio will deteriorate with altitude. For this reason, nearly all high-altitude airplanes have a pitch damper in their autopilot.

Closed-loop response. A typical closed-loop response scenario concerns the airplane yaw damper (**Fig. 6**). It is assumed that the airplane experiences

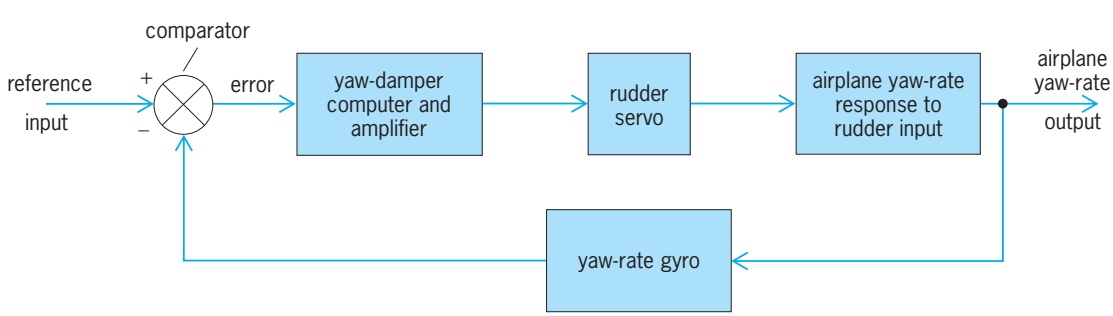

Fig. 6. Block diagram for a yaw damper.

a disturbance in the yaw rate, the angular rotation rate about the Z axis (Fig. 1). Such a disturbance could be induced by atmospheric turbulence. The yaw rate is sensed by the yaw-rate gyro and sent to a device known as the comparator. The comparator subtracts the sensed yaw-rate signal from the reference input signal of the yaw damper. In steady-state straight-line flight, this reference input signal would be zero. The output of the comparator is called the error signal, and it is sent to the yaw-damper computer. The computer decides how much rudder deflection is required and in which direction the rudder is to be moved. That signal is amplified and sent to the rudder actuator. The airplane responds to the rudder movement with a yaw rate which will tend to cancel the original yaw-rate disturbance. The yaw-damper functions to drive the error signal to zero and thereby assures that the airplane yaw rate output is equal to the reference input, which is the desired result.

This type of response scenario is called closed-loop because a loop is closed in the block diagram of the yaw damper (Fig. 6) around the airplane open-loop yaw rate response to a rudder input. The yaw damper is also called a stability-augmentation system. Its primary function is to enhance the damping of the dutch roll, which in many modern airplanes is inherently poor. *See* STABILITY AUGMENTATION.

Closed-loop systems are used to control airplanes automatically in many cases. Automatic navigation, automatic approach, and landing are typical applications. Many automatic systems can control an airplane more precisely and safely than pilots can. To ensure adequate safety in the case of system failures, many closed-loop systems are made redundant so that the failure of one or more systems is of little or no consequence. *See* AIRPLANE; AUTOMATIC LANDING SYSTEM; AUTOPILOT; CONTROL SYSTEMS; FLIGHT CONTROLS.

Jan Roskam

Bibliography. B. Etkin and L. D. Reid, *Dynamics of Flight-Stability and Control*, 3d ed., 1996; J. Roskam, *Airplane Design*, pt. 1–8, 2d ed., 1984, reprint 2000–2003; J. Roskam, *Airplane Flight Dynamics and Automatic Flight Controls*, pt. 1 and 2, 1979, reprint 2001, 2003; D. Stinton, *The Design of the Aeroplane*, 2d ed., 2001.

Flight controls

The devices and systems which govern the attitude of an aircraft and, as a result, the flight path followed by the aircraft. Flight controls are classified as primary flight controls, auxiliary flight controls, and automatic controls. In the case of many conventional airplanes, the primary flight controls utilize hinged, trailing-edge surfaces called elevators for pitch, ailerons for roll, and the rudder for yaw (**Fig. 1**). These surfaces are operated by the human pilot in the cockpit or by an automatic pilot. In the case of vertically rising aircraft, a lift control is provided. *See* AILERON; AIRCRAFT RUDDER; ELEVATOR (AIRCRAFT); ELEVON.

Controls to govern the engine power and speed, while not usually classified as flight controls, are equally important in the overall control of the aircraft. This is especially true if the engine exhaust can be directed to produce pitch or yaw motions.

Auxiliary flight controls may include trimming devices for the primary flight controls, as well as landing flaps, leading-edge flaps or slats, an adjustable stabilizer, a wing with adjustable sweep, dive brakes or speed brakes, and a steerable nose wheel.

Automatic controls include systems which supplement or replace the human pilot as a means of controlling the attitude or path of the aircraft. Such systems include automatic pilots, stability augmentation systems, automatic landing systems, and active controls. Active controls encompass automatic systems which result in performance improvement of the aircraft by allowing reductions in structural weight or aerodynamic drag, while maintaining the desired integrity of the structure and stability of flight.

Cockpit controls. The control system incorporates a set of cockpit controls which enable the pilot to operate the control surfaces. Because of the approximately fixed size and strength of the human pilot and the need to standardize the control procedures for airplanes, the primary controls are similar in most types of airplanes. The cockpit controls incorporate a control stick which operates the elevators and ailerons, and pedals which operate the rudder. Sometimes a column/wheel arrangement is used to operate the elevators and ailerons, respectively. The control stick or column/wheel is mounted on the floor between the pilot's legs. Some fighter and transport airplanes have the control stick mounted on the side of the cockpit, just beyond the end of the pilot's armrest. This space-saving arrangement, called a side-stick controller, can be selected for a variety of reasons, including improved handling qualities during maneuvering air combat and prevention of display obscuration by the control stick or column/wheel.

The longitudinal or pitching motion of the airplane is produced by fore-and-aft motion of the control stick. Forward motion of the control produces a nose-down moment of the airplane, and rearward motion a nose-up moment. Right or clockwise motion of the stick or wheel produces rolling moments to the right, and left or counterclockwise motion produces moments to the left. Moving the right rudder pedal forward produces a yawing moment to the right, and vice versa. A forward angular rotation of the rudder pedals may be used to control wheel brakes for ground control. In cases in which a lift control is provided, upward motion of the lift control lever causes increased lift.

The cockpit controls for auxiliary control devices are not as completely standardized as those for the primary controls. A nose-wheel steering control is provided on large airplanes (such as commercial

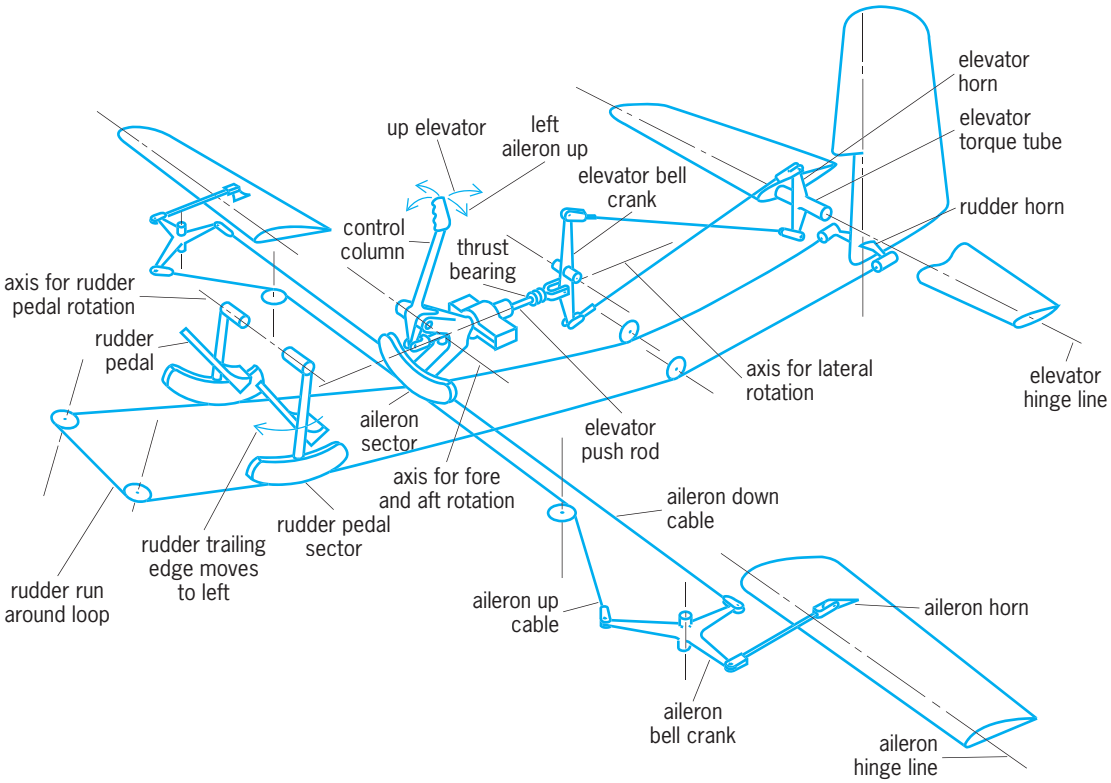

Fig. 1. Simplified diagram of a typical primary flight-control system, governing attitude and direction. (*After C. H. Chatfield, C. F. Taylor, and S. Ober, The Airplane and Its Engine, 5th ed., McGraw-Hill, 1949*)

airliners) or on airplanes that must taxi in a limited ground area (such as an aircraft carrier). The nosewheel steering may be commanded by a dedicated control wheel located on the side of the cockpit or by using the rudder pedals in combination with a switch. Trim controls, which are used to reduce the control forces to zero in a steady flight condition, may be operated by a set of control wheels or by buttons or switches which operate electric motors. In general, the directions of actuation are made consistent with those of the primary controls.

Types of control systems. Control systems with varying degrees of complexity are required, depending on the size, speed, and mission of an airplane (**Fig. 2**). In relatively small or low-speed airplanes, the cockpit controls may be connected directly to the control surfaces by cables or pushrods, so that the forces exerted by the pilot are transmitted directly to the control surfaces (Fig. 2*a*). In large or high-speed airplanes, the forces exerted by the pilot may be inadequate to move the controls. In these cases, either an aerodynamic activator called a servotab or spring tab (not illustrated) may be employed, or a hydraulic activator may be used (Fig. 2*b*). Hydraulic actuators are universally used on the controls of supersonic airplanes, because of the ineffectiveness of tabs at supersonic speeds and because of the large and sometimes irregular variations of the hinge moments on the control surfaces. In some airplanes, particularly those with swept wings and those which fly at high altitudes, the provision of adequate static stability and damping of oscillations by means of the inherent aerodynamic design of the airplane becomes difficult. In these cases, stability augmentation systems are used. These systems (Fig. 2*c*) utilize sensors such as accelerometers and gyroscopes to sense the motion of the airplane. These sensors generate electrical signals which are amplified and used to operate the hydraulic actuators of the primary control surfaces to provide the desired stability or damping. *See* STABILITY AUGMENTATION.

Auxiliary servomechanisms. Small auxiliary hydraulic or electric servomechanisms are often used to feed the inputs from the sensors to the primary control system. These devices, called secondary actuators, may be linked to the control system in such a way that no effects are felt in the cockpit, for example, by means of an extensible link or a summing linkage in the control system (a so-called series servo), or they may exert a force on the control system that causes movement of the cockpit controls (a so-called parallel servo). In some cases, the provision for both electrical and mechanical inputs is made within the package of the primary actuator. *See* SERVOMECHANISM.

Stability augmentation system failure. In the case of failure of the stability augmentation system, the human pilot should be able to override the system and maintain control of the airplane. In many airplanes, this condition is provided by limiting the travel of a series servo, or the force output of a parallel servo,

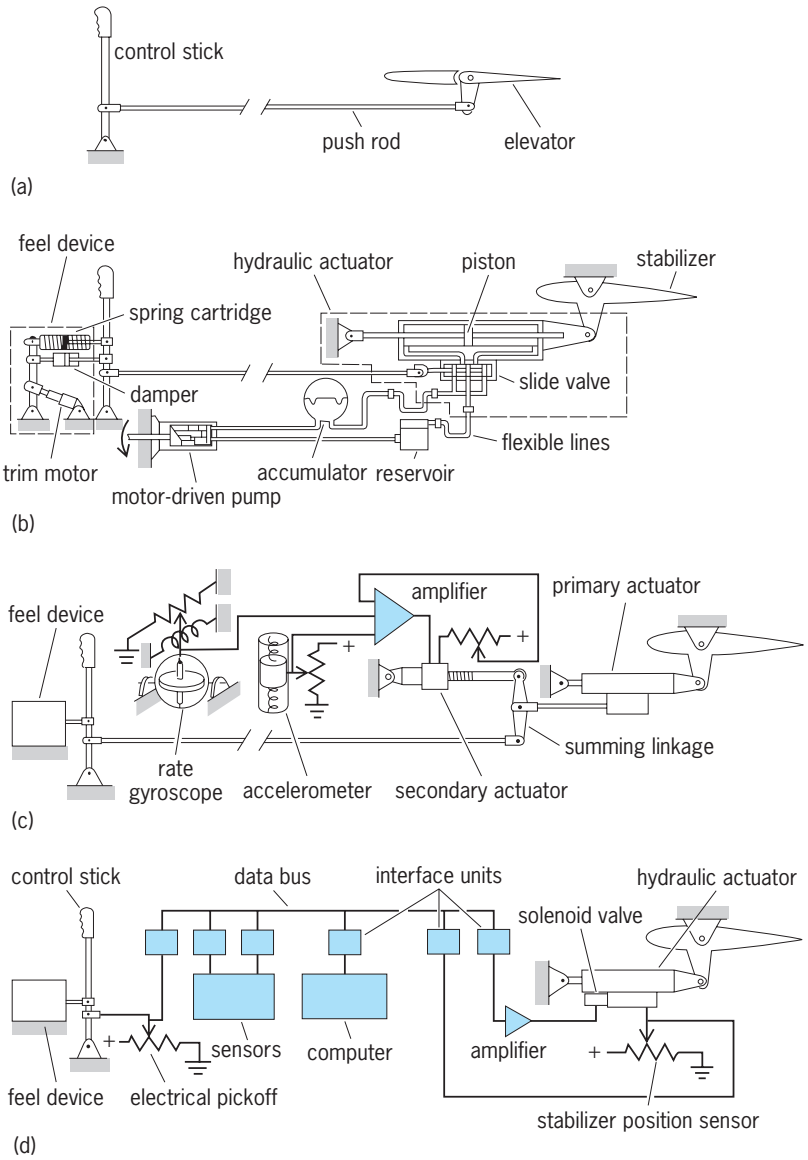

Fig. 2. Typical airplane control systems of varying degrees of complexity, illustrated for the case of longitudinal (elevator or stabilizer) control. Similar systems are used for the other control surfaces. (*a*) Mechanical control system. (*b*) Mechanical control system with control feel device and hydraulic actuator. (*c*) Control system with stability augmentation. (*d*) Digital fly-by-wire control system.

to a small fraction of that provided by the human pilot. This technique, called limiting the authority of the control, may not be acceptable if the automatic control system performs functions which require large control displacements. However, if the automatic system has large authority, its failure might result in an abrupt, large control deflection which would impose loads sufficient to cause structural failure of the airplane. The only way to avoid this problem is to make the automatic system as reliable as the airframe itself.

Fly-by-wire systems. The weight and complication of mechanical control linkages and the extensive reliance on electrical signals in automatic controls led to the development of control systems in which the control inputs from the pilot, as well as those from the stability augmentation sensors, are transmitted to the primary control actuators by electrical signals. Systems of this type are called fly-by-wire systems (Fig. 2*d*). The electrical signals are readily compatible with computers, typically digital, which can perform the functions of combining the signals from the pilot and the sensors. *See* DIGITAL COMPUTER.

Fly-by-light systems. Fly-by-wire systems can malfunction when exposed to high-intensity electromagnetic fields. The malfunctions occur when the electromagnetic fields interfere with the flight control signals being transmitted via wires between the digital computers, sensors, and actuators. The solution to this problem has been to shield the transmission media and extensively test the system before certifying it, adding cost and weight to the system. These difficulties, the intrinsic immunity of optical technology to electromagnetic interference, and the availability of optical-fiber-based transmission media from the communications industry led to the development of fly-by-light systems. These systems use optical fibers to transmit signals instead of wires; the interface units (Fig. 2*d*) are replaced with optical-electrical converters. Ultimately, the entire system, including sensors, actuators, and computers, is likely to be based on optical technology. *See* OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

Digital flight control computers. There are three major reasons why digital flight control computers are used in modern airplanes. First, digital flight control computers can enhance the pilot's control of the airplane by optimizing the movement of the control surfaces in response to pilot commands, over the operating flight conditions of the airplane. Second, as a result of their ability to rapidly monitor and interpret multiple sensor inputs, digital flight control computers can often exceed the performance of an unassisted pilot in compensating for critical situations which might otherwise result in loss of airplane control. Third, digital flight control computers permit input directly from remote control and navigation devices such as digital automatic landing systems, assisting the pilot in zero-visibility conditions or freeing the pilot for other airplane management tasks.

Since digital flight control computers have a critical role in flight safety, they have stringent performance and reliability requirements. Typical requirements include a probability of failure of 10^{-9} over 10 h (for commercial airplanes); an instruction cycle time on the order of 100 Hz; and a transport lag of about 5 milliseconds. Methods of achieving the performance requirements include the use of multiple, redundant processors, fault-tolerant architectures, and parallel multitask processing techniques in both hardware and software. Control systems that make use of digital flight control computers often consist of multiple, physically separated computers with different power sources to reduce the impact of single-point failures. For fully fly-by-wire systems, triple redundancy with a fail-safe backup is considered adequate. This level of redundancy permits two

failures while remaining in operation and a third simultaneous failure while maintaining the safety of flight. This design philosophy is carried through from the input sensors to the control surface actuators and their positional feedback sensors. *See* FAULT-TOLERANT SYSTEMS; MULTIPROCESSING.

Since the inputs to a digital flight control computer are usually analog at the source, they must be converted to digital signals for use by the computers. Three types of signal input are used by the flight control computers: information about the flight status of the airplane and its operating environment; information about the control surface positions; and attitude commands from the pilot or from other airplane systems, such as navigation, weapons control, terrain following/avoidance, autopilot, or automatic landing systems. *See* ANALOG-TO-DIGITAL CONVERTER.

The prime outputs of the digital flight control computer are command signals to the control surface actuators. These actuators are typically constructed so that digital signals from the computer are converted to analog signals. In addition to commanding the control surface actuators, digital flight control computers provide feedback to the pilot. This may be in terms of feel-force actuators on the cockpit controls or updated visual displays of the airplane's attitude and flight status. This feedback serves two purposes: it gives the pilot a sense of identity with the airplane, and it warns when the pilot is asking something beyond the capability of the airplane in terms of maintaining its structural integrity or dynamic stability. *See* DIGITAL-TO-ANALOG CONVERTER.

A digital flight control computer evaluates its inputs based on precomputed models of the airplane's expected behavior under various conditions of flight in order to produce command signals for the control surface actuators. Typically, wind-tunnel data are used to derive and verify the precomputed models used in the computer. Test flights and computer simulations are also used extensively to verify computer operation. *See* AIRCRAFT TESTING; WIND TUNNEL.

The primary calculation tasks of the digital flight control computer involve evaluation of analog-source input signals to produce the necessary motion of the control surfaces. Since this should result in smooth changes in attitude which are proportional to the input command signals, the computer must incorporate a considerable degree of selective signal filtering in its operation. An important function of these filters is to prevent oscillation in the control system which might otherwise result from time or value sampling differences between input, output, and feedback signals. Potential control system oscillations are typically caused by delays and mechanical resonance in the airplane's structure as it interacts with the atmosphere, as well as by the digital nature of the computers themselves. *See* ESTIMATION THEORY.

Handling qualities. Because the control system is the link through which the pilot flies the airplane, its characteristics have an important influence on the handling qualities, those characteristics which influence the ease and safety of flight. In the case of a simple mechanical control system (Fig. 2*a*), a desirable condition exists in that the control forces required of the pilot are directly related to the moments applied to the airplane, because both arise from aerodynamic effects. In such airplanes, much attention is directed to the detailed design of the control surface shapes to produce the desired hinge moment characteristics. In the case of hydraulically actuated controls (Fig. 2*b*), a control feel system must be provided to generate the control forces which are no longer available from the control surface hinge moments. Likewise, the stability augmentation system should be designed to provide response to the pilot's control inputs which have the desired magnitude and speed of response and which avoid excessive overshoot or oscillations. The techniques of servomechanism analysis are employed to design both the mechanisms themselves and the overall system, including the airplane, control system, and pilot modeling. *See* CONTROL SYSTEMS; FEEDBACK CIRCUIT.

Active controls. The availability of reliable, full-authority electronic control systems makes possible the use of closed-loop control systems to improve the performance of the airplane. The applications most frequently considered in order to take advantage of this capability are reduced static stability, maneuver load control, gust load alleviation, and structural mode damping (**Fig. 3**).

Reduced static stability. The operation of the elevators or rudder in such a way as to augment the stabilizing moment normally provided by a fixed surface (Fig. 3*a*) may allow the areas of the horizontal or vertical stabilizing surfaces to be reduced. The reduced tail area results in decreased aerodynamic drag on the airplane. In addition, in the case of the horizontal tail, a more rearward center of gravity position may be allowable, thereby reducing the down load on the tail required for trim in steady flight or in turns. The extent to which the area of stabilizing surfaces may be reduced is determined by conditions requiring full deflection of the controls, such as raising the nose for takeoff, recovery from stall, or flight with asymmetric power. Careful attention to design features which reduce these requirements allows greater benefits from active controls.

Maneuver load control. The ailerons, in addition to being deflected asymmetrically to produce roll, can be moved symmetrically upward in a turn or pull-up, to reduce the load on the wing tips (Fig. 3*b*). As a result, the bending moment on the wing is reduced, allowing a lighter structure. Alternatively, the wingspan may be increased to reduce the drag. *See* WING STRUCTURE.

Gust load alleviation. Gust load alleviation (Fig. 3*c*) may likewise be provided by moving the ailerons, or possibly the wing flaps, symmetrically to offset the load on the wing caused by gusts. The surfaces may be moved in response to the signals from an

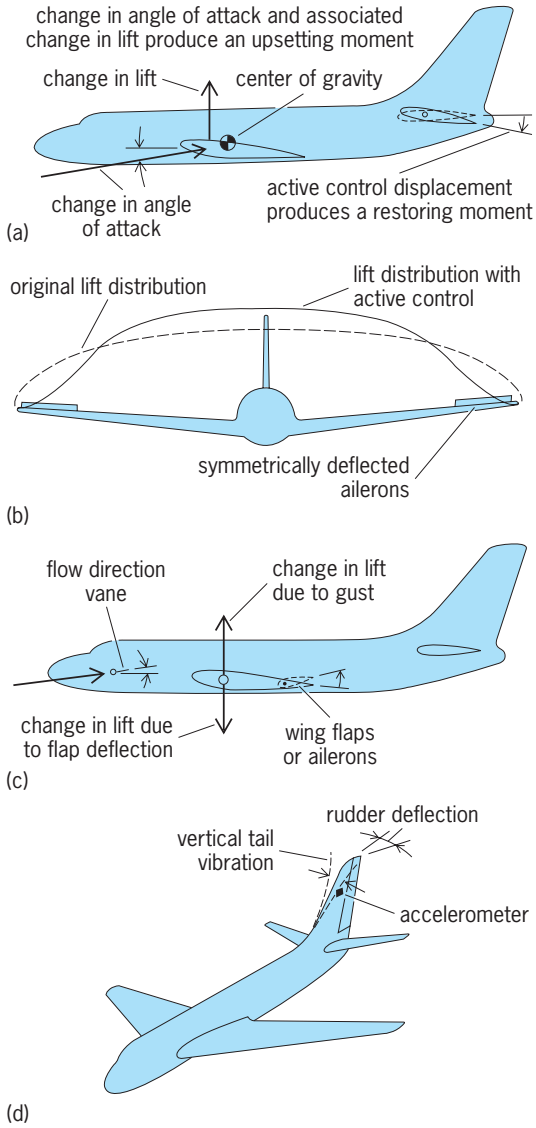

Fig. 3. Examples of active control. (*a*) Active control to allow reduced static stability. Change in angle of attack and associated change in lift produce an upsetting moment. (*b*) Active maneuver load control. (*c*) Active control for gust load alleviation. (*d*) Active control for structural mode damping.

angle-of-attack or flow-direction vane (Fig. 3*c*) or an accelerometer. The reduced gust loads may allow lighter wing structure, as well as improve passenger comfort.

Structural mode damping. Structural mode damping refers to the damping of structural vibrations caused by turbulence (Fig. 3*d*). These repeated loads are largely responsible for reducing the fatigue life of an airplane. The structural weight may therefore be reduced if these loads are decreased. In practice, the oscillations are sensed by accelerometers or rate gyroscopes at various locations in the airplane, and the control forces to damp the oscillations are applied by the primary control surfaces or by small auxiliary surfaces.

For example (Fig. 3*d*), the rudder can be moved in response to signals from an accelerometer in the vertical tail to damp lateral vibrations in the tail assembly. In more extreme applications of this technique, the flutter speed of the airplane may be increased—giving the designer greater freedom to use a lighter structure or a thinner surface, or to redistribute items of mass in a way to improve performance. *See* FLUTTER (AERONAUTICS).

Automatic flight controls. The flight of an aircraft may be controlled automatically by providing the necessary signals for navigation as inputs to the control system. In practice, automatic pilots are used to relieve the human pilot of routine flying for long periods, and automatic control systems are used to make precision landings and takeoffs under conditions of reduced visibility. The reference signals for such navigation functions come from radio signals generated by ground stations or by satellites, and from inertial navigation systems. Inertial navigation systems determine the position of the airplane over the surface of the Earth by use of precise on-board gyroscopes and accelerometers. *See* AUTOPILOT.

The requirements for reliability of the navigation systems for cruise flight are less critical than those for the primary control system, because of the time available to recover from a failure. However, the requirements for reliability of automatic landing systems are similar to those for the primary control system. *See* AUTOMATIC LANDING SYSTEM; ELECTRONIC NAVIGATION SYSTEMS; INERTIAL GUIDANCE SYSTEM; SATELLITE NAVIGATION SYSTEMS.

Flight management. The use of a flight-control system to provide the desired responses to control inputs, to limit structural loads, and to fly on precise paths defined by inertial or external navigation references relieves the human pilot of many tasks requiring skill and dexterity. The pilot is thereby given the opportunity to concentrate on management of the systems which control the airplane. However, even this task can overwhelm the pilot. Therefore, designers are constantly searching for ways to automate airplane control functions. Applications of artificial intelligence concepts, including the use of symbolic manipulation programming languages, are resulting in the development of embedded expert systems to perform a variety of automatic control functions, including automatic selection of navigation aids and other sensors, mode control selection based on flight condition, and automatic fault detection and control system reconfiguration. Although more and more control functions are being automated, designers are careful to provide advisory displays which keep the pilot aware of the status of various sytems and to provide mechanisms which permit the pilot to override an automatic control function if necessary. *See* AIRCRAFT INSTRUMENTATION; ARTIFICIAL INTELLIGENCE; EMBEDDED SYSTEMS.

William H. Phillips; P. Douglas Arbuckle

Bibliography. J. D. Anderson, *Fundamentals of Aerodynamics*, 3d ed., 2001; J. Blakelock, *Automatic Control of Aircraft and Missiles*, 2d ed., 1991; J. Curran, *Trends in Advanced Avionics*, 1992; B. Etkin and L. D. Reid, *Dynamics of Flight Stability and Control*, 3d ed., 1996; G. F. Franklin,

J. D. Powell, and A. Emami-Naeini, *Feedback Control of Dynamic Systems*, 5th ed., 2006; J. Roskam, *Airplane Flight Dynamics and Automatic Flight Controls*, pts. 1 and 2, 1979, reprint 2001, 2003; C. Spitzer, *Digital Avionics Systems: Principles and Practice*, 2d ed., 1993, reprint 2000.

Floodplain

The relatively broad and smooth valley floor that is constructed by an active river and periodically covered with water during intervals of flooding. Water chemistry and suspended sediment patterns indicate that floodwaters can come from combined main channel overflow, local ground-water rise, tributary inputs, and precipitation directly onto a saturated floodplain surface. Engineers consider the floodplain to be any part of the valley floor subject to occasional floods that threaten life and property, and commonly designate floodplain zones based on the recurrence interval with which they are flooded (for example, 100-year floodplain). Various channel modifications, from levees and channelization to flow regulation, can be used to restrict the natural flood process. These modifications have altered most of the world's large river-floodplain systems. For example, an estimated 96% of the floodplain ecosystem historically present along the lower Mississippi River has been lost because of levees and channelization. *See* RIVER; RIVER ENGINEERING.

Geomorphologists consider the floodplain to be a surface that develops by the active erosional and depositional processes of a river. Floodplains are underlain by a variety of sediments, reflecting the fluvial history of the valley, and floodplain stratigraphy is commonly used to infer paleoenvironmental changes. The chronology of these changes is determined using radiocarbon, thermoluminescence, and other dating techniques. Large rivers can have extensive floodplain deposits; the Mississippi River has up to 300 ft (100 m) of fluvial sediment beneath its floodplain, and the floodplain of the Amazon River reaches 30 mi (50 km) across. *See* DATING METHODS; FLUVIAL SEDIMENTS; GEOMORPHOLOGY.

Deposits and topography. Most floodplains consist of the following types of deposits: colluvium—slope wash and mass-wasting products from the valley sides, as is common along steep channels with narrow floodplains; channel lag—coarse sediment marking the bottoms of former channels; lateral accretion deposits—sand and gravel deposited as a meandering river migrates laterally; vertical accretion deposits—clay and silt deposited by overbank flooding of the river; crevasse-splay deposits—relatively coarse sediment carried through breaks in the natural river levees and deposited in areas that usually receive overbank deposition; and channel-fill deposits—sediment filling abandoned river channels. Channel fills can be coarse for sandy rivers. Muddy rivers such as the Mississippi, however, fill their abandoned channels with cohesive clay. The clay "plugs" in these old channels then resist further migration of the river. Changes in the river through time create a complex floodplain topography (**Fig. 1**). *See* STREAM TRANSPORT AND DEPOSITION.

Fig. 1. Aerial view of the floodplain of the Amazon River showing complexity of floodplain topography. 20 km = 12.4 mi. (*After L. A. K. Mertes et al., Channel-floodplain geomorphology along the Solimões-Amazon River, Brazil, Geol. Soci. Amer. Bull., 108:1089–1107, 1996*)

On a section across a typical broad floodplain, such as that of the Mississippi, the highest elevations occur at the crests of natural levees flanking the river channels. The levee back slopes flatten away from the channels toward backswamp basins. These basins receive fine overbank deposits during occasional floods. Their high water tables promote irregular lakes and swamp or marsh vegetation. Levee crests can rise 33 ft (10 m) or more above adjacent backswamps along certain portions of the Mississippi. Crevasse splays through the levees can locally extend sandy deposits into backswamp areas.

The floodplain of the world's largest river, the Amazon, is annually inundated for 4 to 5 months. Water level rises 30 ft (9 m) vertically along some parts of the river, and spreads up to 30 mi (50 km) across the floodplain. Vertical change in water level as a function of discharge is related to floodplain width, which varies substantially along the length of most rivers. At the downstream end of a river that enters a body of standing water, the floodplain grades into a delta. *See* DELTA.

Floodplain functions and dynamics. Floodplains influence physical processes along rivers because the floodplain surface often has greater hydraulic roughness, associated with vegetation or structures,

Fig. 2. Flooded forest along the Amazon River. Water depths can reach several meters in areas such as this during the peak of the annual flood.

and shallower flows than the main channels. Thus floodwater moves across floodplains more slowly, dissipating flow energy and attenuating the flood hydrograph. Floodplains also strongly influence river ecological processes by providing spawning and nursery habitat for many species during floods, and nutrients to the main channel as floodwaters recede. Habitat complexity and abundance increases substantially during floods as wetlands, backwaters, flooded forests (**Fig. 2**), and floodplain lakes all become accessible to organisms migrating out from the main channel. These ecological processes are described by the flood-pulse model, which highlights the ecological necessity of regular seasonal floods along many alluvial rivers. River restoration efforts around the world now attempt to restore channel–floodplain connections through levee setbacks, changes in flow regulation, and removal of channelization measures.

Sediment can remain on a floodplain for hundreds or thousands of years, or the sediment can have a very short residence time. Residence time depends on the recurrence interval of high-energy floods capable of depositing or eroding floodplain sediments (**Fig. 3**). Laterally stable rivers can alternate between gradual floodplain construction through overbank deposition over hundreds of years, and episodic erosion that strips the floodplain to a basal lag deposit from which it slowly reforms. Laterally mobile rivers are more likely to have regular erosion and redeposition of new floodplain sediments as the river migrates. These different patterns of floodplain storage can be particularly important where contaminants such as heavy metals or organochlorine compounds are attached to silt and clay particles. In this case, floodplain dynamics determine how these contaminants are stored and mobilized along a river network. *See* SEDIMENTOLOGY; STREAM TRANSPORT AND DEPOSITION.

River terraces. Many rivers display abandoned floodplains that formed when the river flowed at higher levels than at present. These abandoned floodplains are known as terraces. Terraces reflect the numerous complex controls on river gradients and patterns, including climatic changes, uplift or subsidence of the drainage basin, sea-level changes, and land-use changes. Distinguishing between these various controls for a given field setting is often very

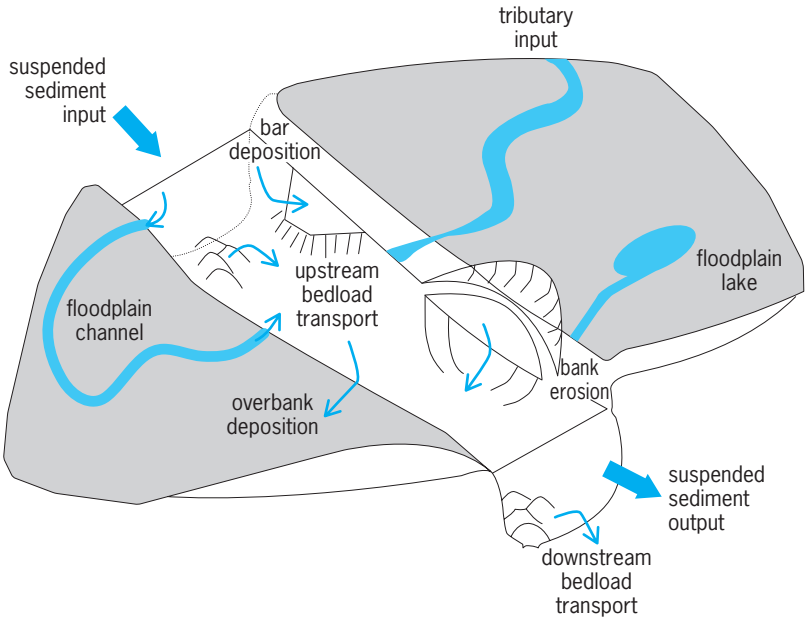

Fig. 3. Processes governing sediment storage along a channel–floodplain reach. Sediment enters a reach of channel from upstream, tributaries within the reach, and bank erosion. Sediment leaves the channel reach by channel transport, deposition on bars, diffuse overbank flow, and through floodplain channels. (*After T. Dunne et al., Exchanges of sediment between the flood plain and channel of the Amazon River in Brazil, Geol. Soc. Amer. Bull., 110:450–467, 1998*)

difficult, but has the potential to provide important insights into environmental changes and river response through time. *See* FLUVIAL EROSION LANDFORMS. Ellen E. Wohl; Victor R. Baker

Bibliography. P. B. Bayley, The flood pulse advantage and the restoration of river-floodplain systems, *Regul. Rivers Res. Manage.*, 6:75–86, 1991; P. B. Bayley, Understanding large river-floodplain ecosystems, *BioScience*, 45:153–158, 1995; J. S. Bridge, *Rivers and Floodplains: Forms, Processes, and Sedimentary Record*, Blackwell, Oxford, 2003; T. Dunne et al., Exchanges of sediment between the flood plain and channel of the Amazon River in Brazil, *Geol. Soc. Amer. Bull.*, 110:450–467, 1998; C. W. Martin, Heavy metal storage in near channel sediments of the Lahn River, Germany, *Geomorphology*, 61:275–285, 2004; L. A. K. Mertes, Inundation hydrology, pp. 145–166, in E. Wohl (ed.), *Inland Flood Hazards*, Cambridge University Press, 2000; R. E. Sparks, Need for ecosystem management of large rivers and their floodplains, *BioScience*, 45:168–182, 1995.

Floor construction

A floor of a building generally provides a wearing surface on top of a flat support structure. Its form and materials are chosen for architectural, structural, and cost reasons.

A ground-supported floor may be of almost any firm material, ranging from compacted soil to reinforced concrete. It is supported directly by the subsoil below.

An elevated floor spans between, and is supported by, beams, columns, or bearing walls. It is designed to be strong and stiff enough to support its design loading without excessive deflection; to provide for an appropriate degree of fire resistance; and to supply diaphragm strength to maintain the shape of the building as a whole, if necessary. A ceiling may be hung from the underside of the floor assembly as a finish surface for the room below. The optimum floor design meets these criteria while being as thin as possible and economical to construct. *See* BEAM; COLUMN; LOADS, TRANSVERSE.

Wood. Wooden floors are generally used in light residential construction. Such flooring generally consists of a finish floor installed on a subfloor of tongue-and-groove planking or plywood, spanning between wooden beams that are commonly called joists (see **illus.**). The flooring is usually about 1 in. (2.5 cm) thick; the joists are, for example, 2 in. by 8 in. (5 cm by 20 cm), spaced at 16 in. (40 cm) or 24 in. (60 cm) on center. Such distances are widely recognized in modular-type construction. The joists usually span about 10–20 ft (3–6 m) between their supports. The dimensions are determined by architectural requirements, code-stipulated design loads, and allowable design stresses for the floor components. The wearing surface of such wooden floors is typically of vinyl tiles or hardwood flooring.

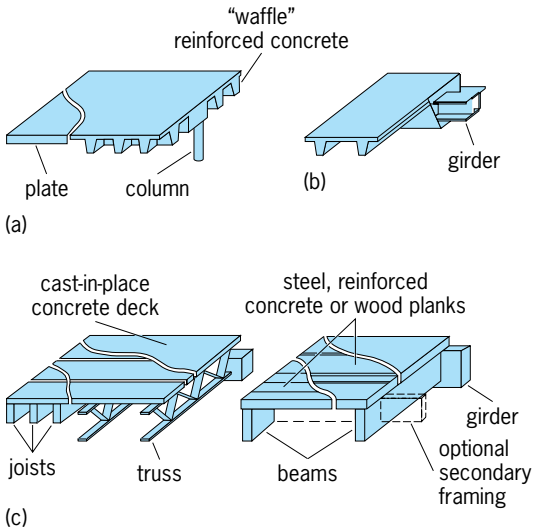

Some types of framing and decking floor construction. (*a*) Cast-in place. (*b*) Precast. (*c*) Site-assembled. (*After Sweets Catalog File, General Building and Renovation, 1993*)

Reinforced concrete. Slabs fabricated of reinforced concrete are a common type of floor for heavier loading. The concrete is cast on forms, and reinforced with properly placed and shaped steel bars (rebars), so as to span between steel or reinforced concrete beams or between bearing walls. Flat slab floors are supported primarily by columns alone. Sometimes reinforced concrete planks precast in a factory are used for the slab structure, covered by a thin finish cover of a strong concrete mix that is troweled smooth. *See* REINFORCED CONCRETE.

Composite. This type of floor is commonly used in modern office building construction. Concrete is cast on, and made structurally integral with, corrugated metal deck, which spans between steel joists of either solid-beam or open-web types, generally spaced between about 16–48 in. (40–120 cm) on center. The corrugated metal deck may include closed tubular regions to provide space for electrical conduit. Sometimes such composite deck is also structurally connected to the supporting steel beams so as to make them act as composite beams and thus be stronger, stiffer, and more economical for their function. *See* COMPOSITE BEAM.

Prestressed concrete. This is used for long span slabs. Highly prestressed high-tension steel wires within the high-strength concrete slab produce a thin, stiff, and strong floor deck. *See* PRESTRESSED CONCRETE.

Lift slab. This type of construction has evolved for economy and efficiency. A concrete slab is first formed at ground level, reinforced and cured to adequate strength, and then carefully jacked up into its final position on supporting columns. *See* STRUCTURAL DESIGN. Milton Alpern

Bibliography. American Society of Civil Engineers, *Minimum Design Loads for Buildings and Other Structures*, ASCE 7–88, 2000; D. K. Ballast, *Architect's Handbook of Construction Detailing*, 1990;

H. J. Cowan, *Design of Reinforced Concrete Structures*, 2d ed., 1988; K. F. Faherty and T. G. Williamson (eds.), *Wood Engineering and Construction Handbook*, 2d ed., 1995.

Floriculture

The segment of horticulture concerned with commercial production, marketing and sale of bedding plants, cut flowers, potted flowering plants, foliage plants, flower arrangements, and noncommercial home gardening.

Florist crop production. Commercial crops are grown either in the field or under protected cultivation, such as in glass or plastic structures. Field production is confined to warm climates or to summer months in colder areas. Typical field crops are gladiolus, peonies, stock, gypsophila, asters, and chrysanthemums. Greenhouse production is not as confined by climate or season, but most greenhouses are located in areas that have advantages such as high light intensity, cool night temperatures, or ready access to market. Jet air transportation resulted in major changes in international crop production. For example, fast transport of inexpensive cut flowers from Bogotá, Colombia, to Europe and the United States greatly affected production on all three continents. Other newer crop production centers, such as Israel, could cause additional changes in the world production and distribution of cut flowers. The cost of shipping potted plants has generally confined large-scale potted plant production to population centers.

Costs. Greenhouse production and profits are calculated on a square-foot basis rather than on an acre basis as with most agricultural crops. Costs are high because of labor, expensive greenhouse structures and equipment, fuel required for heating, and handling and shipping of the perishable products. Commercial flower growers are quick to utilize research results to increase efficiency, quality, production and profit.

Soilless media. A striking change that has occurred in commercial greenhouse production is the almost universal acceptance of soilless mixes as growing media, particularly for potted plants. In some areas the most popular medium is peat moss and vermiculite, or peat and perlite, while in other areas such as the southeastern United States it is pine bark humus and sand. Many of the media are formulated by processors who sell the potting mixture to growers. Advantages of the prepared mixes can be uniformity of product from one shipment to the next, freedom from pests, and lightness. The media are also well drained with adequate pore space. Soilless media gained in popularity when field soils became less available or of questionable quality at greenhouse ranges located in or near metropolitan areas.

Controls in the greenhouse. The greenhouse environment can be controlled to such an extent that flowers can be produced on very precise schedules year-round. The phenomenon of photoperiodism is used by growers to produce chrysanthemum flowers when the natural day length is too long for flower bud initiation and development. Growers simulate short-day conditions by pulling black cloth over the plants each evening. The short-day plants can be kept in a vegetative state by interrupting the dark period with lights, when the natural day length would be conducive for flowering. Flower bud dormancy of azaleas and hydrangeas can be overcome by exposing plants to cool temperatures. Proper nutrient levels are maintained with sophisticated fertilization programs. Pests are controlled with systemic fungicides and insecticides, and biological control is often used. *See* PHOTOPERIODISM.

Growth and flowering also are regulated with chemicals. Some chemicals enhance rooting of cuttings, while other stimulate lateral shoot development, inhibit growth, stimulate flower bud initiation or development, break dormancy, or affect flower color. Perhaps the most commonly used chemicals are the growth retardants which prevent potted plants such as poinsettias and chrysanthemums from becoming too tall.

Breeding. Pronounced improvements in cultivars have been realized because of excellent breeding programs conducted by commercial propagators and by some horticulture departments. Modern cultivars have traits such as more attractive flower colors and forms, longer-lasting flowers, better growth habit, increased resistance to insects and disease organisms, or ability to grow and flower at cooler night temperatures. *See* BREEDING (PLANT).

Field production. Most environmental factors cannot be as well controlled in field production as in the commercial greenhouse, but growers follow some procedures to regulate growth and flowering. Fields can be lighted at night to keep chrysanthemums vegetative until growers wish to provide short days for flowering. Black cloth or plastic can be pulled over plants at night in the summer to promote flowering, though the expense of this practice has reduced its popularity and feasibility. Plants often are covered with Saran cloth to reduce light intensity to prevent injurious leaf and flower temperatures. Irrigation systems are installed to ensure adequate water, and soluble fertilizers can be injected through the irrigation system. Fields are fumigated prior to planting to control pests, and crops can be grown at closer spacings since cultivation is not required to control weeds. Grading and packing operations have been improved to get the flowers to market in good condition.

Wholesale marketing. The wholesaler serves as a valuable link between the grower and florist. Cut flowers and florists' greens primarily are marketed by the wholesaler, though some potted plants pass through this channel. Wholesalers usually are located in metropolitan areas, near retail florist shops, but some wholesalers use trucks to service shops quite far from the wholesale house. The wholesale florist also has a large inventory of supplies, such as ribbon,

foil, containers, and other items essential in floral design.

Retail floriculture. The retail florist often is the final link in the movement of flowers and potted plants from the field to greenhouse to the customer. Florists are especially busy at holidays, but many floral arrangements and corsages are purchased for special occasions throughout the year. The florist offers various customer services, such as delivery, artistic design, and credit. An organized network of florist shops makes its possible for flowers to be ordered from one shop and delivered by another in a distant city.

Cut flowers and potted plants also are sold by mass market outlets such as supermarkets, department stores, and discount centers. Flowers and potted plants are purchased in large quantities by these outlets. Most sales are on a cash-and-carry basis, eliminating costs of delivery and billing. A large percentage of bedding plants are sold in supermarkets.

The purchase of flowers on a weekly basis is not yet as common in the United States as in Europe, but the increased interest in the environment has had pronounced effects on the floriculture industry.

Home gardening. Expanded interest in the environment has increased the number of individuals who now enjoy gardening as an avocation. Bedding plants, such as petunias, marigolds, geraniums, and impatiens, can produce pleasing effects throughout the summer at little cost. *See* PLANT GROWTH.

Roy A. Larson

Bibliography. M. N. King (ed.), *Introduction to Greenhouse Production*, 1992; R. Larson (ed.), *Introduction to Floriculture*, 2d ed., 1992; G. MacDaniel, *Ornamental Horticulture*, 1991.

Flotation

A process used to separate particulate solids, which have been suspended in a fluid, by selectively attaching the particles to be removed to a light fluid and allowing this mineralized fluid aggregation to rise to where it can be removed. The principal use of the process is to separate valuable minerals from waste rock, or gangue, in which case the ground ore is suspended in water and, after chemical treatment, subjected to bubbles of air. The minerals which are to be floated attach to the air bubbles, rise through the suspension, and are removed with the froth which forms on top of the pulp. Although most materials subjected to flotation are minerals, applications to chemical and biological materials have been reported.

Fields of application. Froth flotation was first used to recover sulfide minerals which were too fine to be recovered by gravity concentrators such as jigs, tables, and sluices. These early operations removed all of the sulfide minerals into one concentrate. Although good recoveries were achieved by this bulk flotation, the concentrate grades were poor.

As the understanding of the chemistry of flotation improved, better reagents became available and differential flotation became a reality. It became possible, in many cases, to float only the valuable sulfide minerals and leave the unwanted pyrite (fool's gold) with the gangue. The advent of differential flotation revolutionized the mineral-processing industry. Most plants discarded their gravity-concentrating equipment and turned to flotation exclusively.

Froth flotation is also used to separate the silicate minerals. It is quite common to produce a concentrate of albite ($NaAlSi_3O_8$), muscovite [$KAl_2(AlSi_3O_{10})(HO_2)_1$], quartz ($SiO_2$), spodumene ($LiAlSi_2O_6$), and zircon ($ZrSiO_4$) by flotation.

Flotation is also used to concentrate oxide minerals such as hematite (Fe_2O_3) and pyrolusite (MnO_2), and native elements such as sulfur, silver, gold, copper, and carbon (both graphite and diamond), as well as many salts. Slightly soluble salts such as fluorite (NaF), barite ($BaSO_4$), and calcite ($CaCO_3$) can be floated with conventional flotation techniques, but soluble salts such as halite (NaCl) and sylvite (KCl) are floated not from water but from saturated brines.

Diamonds and phosphates are floated by immersing the ore in water on which there is a thin layer of oil. The gangue sinks, and the diamonds (or phosphates) float on the water with the oil. This type of flotation is known as skin flotation.

Flotation has been widely used in the coal-cleaning industry; the coal is floated, and the pyrite and shales, which form ash, are depressed into the pulp.

Principles. The reagents used in flotation can be classified as collectors, frothers, and modifiers.

Collectors. Collectors are organic molecules which have two distinct parts: polar and nonpolar. The polar end of the collector molecule attaches to the mineral surface; the nonpolar portion of the molecule is exposed to the water. The thin coating of collector transforms the minerals from hydrophylic to hydrophobic. When the coated mineral particle comes into contact with an air bubble, the particle adheres to the bubble and rises through the pulp, as shown in **Fig. 1**. Collectors are classified as either cationic or anionic. Examples of anionic collectors are xanthates, sulfonates, thiophosphates, and organic acids. The principal cationic collectors are amines.

The mechanism of adsorption of collectors onto mineral surfaces and the resulting bubble attachment has been debated. Adsorption is controlled by a combination of driving forces, any one of which may dominate in any given case. Among the recognized driving forces are: chemical reaction between the mineral and the collector; attraction of opposite charges between the mineral surface and the polar portion of the collector; and association between the nonpolar collector tails. *See* ADSORPTION.

Frothers. Frothers are added to the mineral pulps to help form a stable froth which can be removed. Frothers perform this function by becoming more concentrated at air-water interfaces than in the body

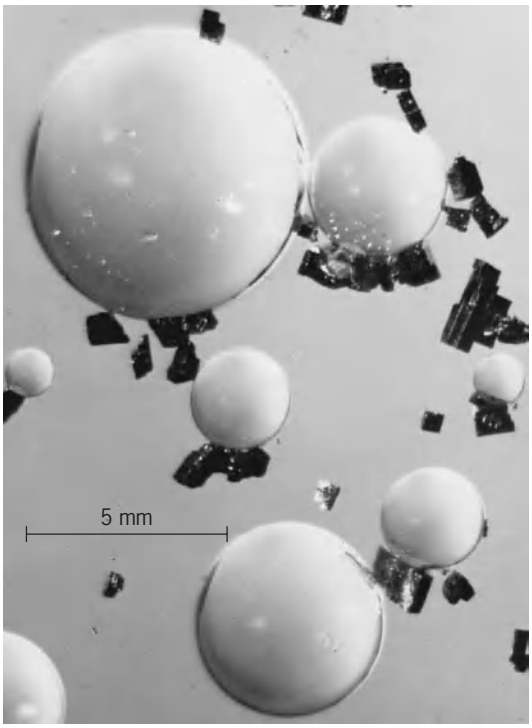

Fig. 1. Photomicrograph of particle-bubble attachment in flotation. Particles of xanthate-conditioned 65/100-mesh galena are clinging to air bubbles and are being rafted to the surface. View is horizontal through the side of a glass flotation cell. (*Courtesy of W. R. Speden*)

of the water. This changes the surface free energy of the air-water interface. When a bubble strikes the pulp-air interface, it does not break, because the kinetic energy associated with its movement can be adsorbed by diffusion of more frother to the interface. Because the surface must absorb the energy, it is important that the concentration of frother be such that a small change in frother concentration produces a large change in surface energy. For this reason, an excess of frother can "kill" froth.

Frothers, somewhat like collectors, tend to be heteropolar molecules, with a polar and nonpolar portion. Some collectors, such as oleic acid and the amines, also produce a froth. It is important that the frother not act as a collector and float an unwanted mineral.

The chemicals used as frothers are primary alcohols, secondary alcohols, and ethers.

Modifiers. Modifiers are added to flotation systems to alter the chemical behavior of the collector or the mineral, or both. If cyanide (CN^-) is added to a pulp which contains pyrite (FeS_2), an iron-cyanide complex is formed on the surfaces of the pyrite which prevents the attachment of the collector and, therefore, prevents the flotation of the pyrite.

Another example is the addition of copper ion (Cu^{2+}) to activate sphalerite (ZnS). Xanthates make excellent collectors for copper sulfide minerals, but poor collectors for zinc minerals. When copper ions are added to the pulp, there is an exchange of zinc for copper at the surface of the sphalerite which causes the surface of the zinc mineral to respond, chemically, as if it were a copper mineral, and hence it can be easily floated with the xanthate collectors.

The interaction of the air bubble and particle in the pulp, and the subsequent attachment, appears to be a dual process. The pulp is supersaturated with gas (air), and a bubble nucleates on the collector-coated mineral. The collision of this particle with a free-moving air bubble results in an easy attachment of the two.

Plant practice. Commercial-scale flotation is carried out in a continuous operation with equipment similar to that shown in **Fig. 2**. These cells can be as small as $^1/_3$ ft^3 to as large as 500 ft^3 (14 m^3).

Among the operational variables are the pulp density (ratio of solids to liquid), the amount of reagent added, the time that the pulp spends in the cells, the height of the froth, the speed of the agitator, the temperature of the pulp, and the size of the solids in the pulp. There are also several variables associated with each ore which cannot be controlled. In light of these variables, it is safe to say that no two flotation plants have the same flow sheet or reagent schedule.

There are many approaches which can be taken to separate a mixture of valuable minerals. One is to float all of the valuable minerals together and then separate them. This approach is used to recover chalcopyrite ($CuFeS_2$) and molybdenite (MoS_2). Both minerals are floated; xanthates are used to collect the chalcopyrite, and fuel oil is used to activate the molybdenite. It should also be noted that cyanide and lime [$Ca(OH)_2$] are added to the pulp to prevent the

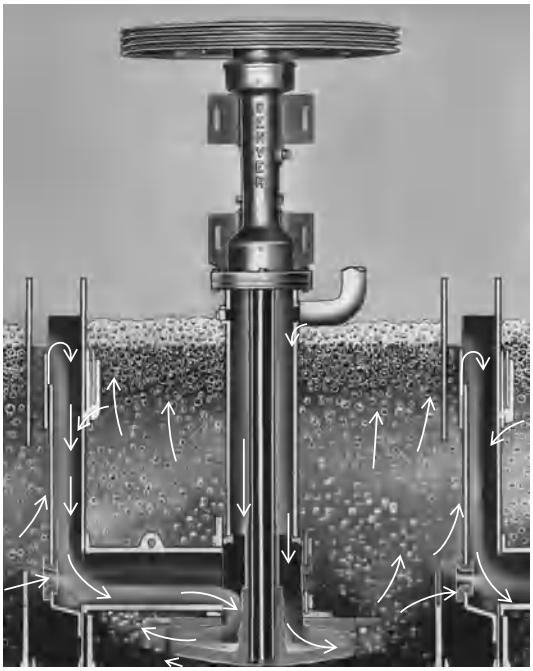

Fig. 2. Cutaway view of section of a flotation cell. (*Denver Equipment Co.*)

pyrite from floating. The chalcopyrite-molybdenite concentrates are then treated with a polysulfide, known as Noke's reagent, which prevents the chalcopyrite from floating. The next stage of flotation produces a molybdenite concentrate in the froth and a chalcopyrite concentrate in the tails.

This separation can also be approached by floating each mineral separately. The molybdenite is floated, the xanthate can be added to the pulp, and the chalcopyrite recovered. This approach works well, but a plant almost twice as large as that used for the bulk float approach described above is needed.

Separating some ores requires both approaches. An example is the separation of a mixture of chalcopyrite ($CuFeS_2$), galena (PbS), sphalerite (ZnS), and pyrite (FeS_2) in a calcite gangue. It is easy to float the chalcopyrite and the galena together and then separate them. After the galena and chalcopyrite are removed, the sphalerite can be floated.

Trends in flotation have been toward using larger cells with automatic equipment to control such variables as pH, pulp density, and froth height. The amount of material treated by flotation is enormous. Plants which can handle in excess of 100,000 tons (90,000 metric tons) per day are common. *See* INTERFACE OF PHASES; ORE DRESSING; SURFACTANT; UNIT OPERATIONS. Ronald L. Atwood

Bibliography. R. D. Crozier, *Flotation*, 1992; M. C. Furstenau and J. D. Miller, *Chemistry of Flotation*, 1985; M. W. Ranney (ed.), *Flotation Agents and Processes: Technology and Applications*, 1981.

Flow-induced vibration

Structural vibrations caused by a flowing fluid. Whenever a structure is exposed to a flowing fluid, the fluid forces may cause the structure to vibrate. In many cases, such as a gentle wind rustling the leaves of a tree, the structural response does not produce any damage. However, in other cases, the structural motion can couple with the fluid excitation in such a way that the energy from the flowing fluid is transferred to the structure and damaging vibrations can occur. One of most spectacular examples of this is the 1940 collapse of the Tacoma Narrows Suspension Bridge. In this case, the 2800-ft (850-m) central span oscillated wildly for several hours before finally failing at a relatively modest wind speed just over 40 mi/h (65 km/h). In fact, a similar catastrophe had occurred nearly 100 years earlier with the Wheeling Suspension Bridge in Ohio. This bridge had a span of 1000 ft (305 m), the longest in the world at the time (1854). Any lessons learned from this experience had apparently long been forgotten when the Tacoma Narrows Bridge was designed. Fortunately, not all excessive flow-induced vibrations lead to failure if they are recognized and appropriate action is taken. The Golden Gate Suspension Bridge was completed in 1937 and has a central span of 4200 ft (1280 m). During a wind storm in 1951, the bridge deck suffered peak-to-peak vibration amplitudes of 12 ft (3.5 m) and peak-to-peak rotational oscillations of 22°. Subsequent torsional stiffening of the deck stabilized the bridge response to acceptable levels. *See* BRIDGE; FLUID MECHANICS; MECHANICAL VIBRATION; VIBRATION.

Impact of changing design approaches. Traditionally, structures were designed with large factors of safety, which were required to compensate for the effects of limited knowledge of loading and strength as well as for the relatively simple analysis tools available at the time. The resulting structures were massive and stiff, and they did not respond significantly to fluid loading. Even when light weight was a design criterion, the structures were stiff, such as is seen in the early biplane and triplane designs. However, the introduction of the monoplane for increased performance also substantially reduced wing stiffness. Thus, the wings of the Fokker D8 were found to oscillate abruptly at a certain wind speed and to break off after a few oscillations. The flexibility of the wing permitted bending and twisting vibrations that coupled with the wind to produce a hitherto unknown flow-induced vibration phenomenon. This phenomenon has become known as bending-torsion flutter, and it is a critical design criterion for modern fixed-wing aircraft. *See* AEROELASTICITY; FLUTTER (AERONAUTICS); STRUCTURAL ANALYSIS.

The technology explosion since the mid-twentieth century has revolutionized our understanding of materials and structural loading as well as our analysis tools. This has permitted a much improved approach to design and the establishment of rational factors of safety. These advances, on top of the demand for higher efficiency and performance, have resulted in lighter, more flexible structures exposed to higher-velocity flows. All of these factors have increased the likelihood of flow-induced vibrations. The problems are ubiquitous, from the fluttering of artificial heart valves and the galloping of bridge decks to the oscillations of marine structures in waves and tidal currents and damaging vibrations of tubes in nuclear steam generators. Unfortunately, the fields of application are so wide ranging that the literature is quite scattered, tending to be focused in specialized areas of application such as aeronautics, wind engineering, ocean engineering, hydraulics engineering, nuclear engineering, and biomechanics. Because of the particular difficulties in analyzing problems of dynamic instability and unsteady fluid flows with moving boundaries, substantial research has been carried out in the fields of theoretical mechanics and computational fluid dynamics. The latter is especially challenging and a field of much current research activity. *See* COMPUTATIONAL FLUID DYNAMICS.

Basic mechanisms. Despite the great diversity of the physical applications of flow-induced vibrations, they can generally be classified into three basic mechanisms: forced vibrations, self-excited vibrations, and self-controlled or resonant vibrations. These distinguish the phenomena both physically and mathematically and can be illustrated generically by plotting structural response amplitude against flow velocity

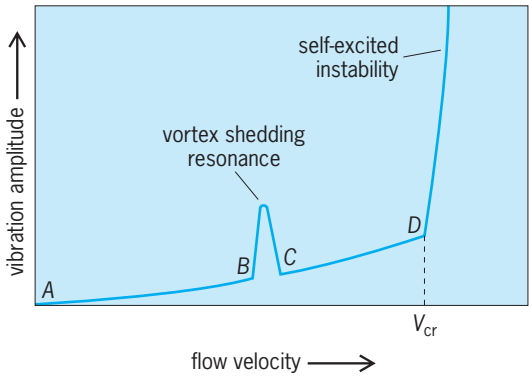

Fig. 1. Generic vibration amplitude response as a function of flow velocity.

(**Fig. 1**). The details of the response characteristics are discussed separately with each excitation mechanism below.

Forced vibration. Forced vibration phenomena are characterized by fluid excitation forces which do not interact strongly with the structural motions; that is, the fluid forces are not affected by the structural response and contain no significant periodicities that could excite structural resonance. This is the simplest form of flow-induced vibration.

Mathematically, such problems involve excitation forces that are functions of time alone, and the solution represents a steady-state equilibrium between the energy input to the structure from the flow and the energy dissipated by structural damping. As the flow velocity increases, the energy in the flow increases and so too will the structural response (Fig. 1, regions *A–B* and *C–D*).

Typically, modeling the structural dynamics is fairly straightforward and uses empirical damping factors. Unfortunately, modeling the fluid mechanics in many cases is much more challenging and measured data are usually required. However, since the structural motion does not affect the fluid forces, the experimental determination of the forces is relatively simple and can be obtained from rigid scaled wind-tunnel models or from measurements on full-scale structures. *See* WIND TUNNEL.

The most common example of this type of flow-excitation mechanism is the turbulent buffeting of structures by the wind. The wind near the Earth's surface behaves like a turbulent boundary layer, the details of which depend on climatic conditions, local topology, and the height above the Earth's surface. The forces are random in nature and, therefore, the structural response can only be determined statistically. Thus, mean and extreme wind loading statistics are used for structural design and form the basis of the wind engineering aspects of building codes. *See* BOUNDARY-LAYER FLOW; ENVIRONMENTAL FLUID MECHANICS; TURBULENT FLOW; WIND.

Of course, there are innumerable other practical examples of this type of flow-induced vibration from the turbulent buffeting of heat exchanger tubes to the gust loading of airplane wings. However, the approach to engineering analysis is basically the same: Compute the structural response statistics using measured flow excitation data. There will always be some turbulence response, but proper design can keep this within acceptable limits. Usually, structural failure due to turbulent buffeting takes a long time to develop and appears in the form of high-cycle fatigue or wear in the case of the rubbing or impacting of loose components. *See* PLASTIC DEFORMATION OF METAL.

Structural failures due to extreme loading caused by such events as hurricane force winds are due to the extremely high average value of the fluid forces and are typically not due to flow-induced vibrations.

Self-excited vibrations. Self-excited vibrations are characterized by periodic fluid excitation forces which depend totally on the structural response; that is, in the absence of structural response, the periodic forces associated with self-excitation mechanisms do not exist. This makes these phenomena much more difficult to study because not only do the structural geometry and fluid mechanics need to be properly scaled but so do the structural dynamics and the fluid-structure interaction. Thus, a model must deform and vibrate just like the full-scale structure (structural dynamic similarity) with the further requirement that the model response in the flow scale to the prototype (fluid-structure similitude). This is often difficult or impossible to achieve fully, and significant judgment may be required for modeling such problems. *See* DIMENSIONAL ANALYSIS.

Mathematically, self-excited vibrations are dynamic stability problems in which the fluid forces are dependent on time and structural displacement. When the flow velocity reaches the critical velocity for instability, V_{cr}, a dramatic increase in oscillation amplitude occurs (Fig. 1, point *D*). The velocity V_{cr} usually represents the practical upper limit for the flow velocity to prevent structural failure.

Bending-torsion flutter of aircraft wings is an excellent example of self-excited vibrations. As the flow velocity increases, the flow over the wings couples with the structural natural frequencies and modifies them. At the critical velocity for flutter, the aerodynamic coupling causes coalescence of the bending and torsional wing vibration modes and the oscillation amplitudes grow extremely rapidly until catastrophic failure occurs. This happens without warning and so quickly that there is no time for pilot reaction. Thus, bending-torsion flutter is a critical design criterion for fixed-wing aircraft. Interestingly, modern feedback control technologies are used on tactical aircraft to sense impending flutter, and automatically adjust controls to effectively delay the critical velocity. *See* CONTROL SYSTEMS.

There are numerous other examples self-excited vibrations that can cause structural failures in service. Tube-and-shell heat exchangers are used widely in chemical, fossil, and nuclear power plants. Since the 1960s, there have been many failures of tubes in such heat exchangers due to a self-excited phenomenon called fluid-elastic instability. Above the critical velocity, the tubes vibrate with such violence that they can fail rapidly from fatigue or wear due to

Fig. 2. Heat exchanger tubes cut through at supports and damaged by impacting from broken tubes due to flow-induced vibration.

tube-to-tube clashing. This causes plant shutdowns, loss of productivity, and very costly repairs. In the case of nuclear steam generators it causes leakage of irradiated primary side coolants. Clearly, fluid-elastic instability must be avoided in such equipment to prevent serious damage (**Fig. 2**). *See* HEAT EXCHANGER.

Other examples of self-excited vibrations are seen in the so-called galloping of bluff bodies, such as bridge decks, certain structural shapes, and iced power transmission lines, and the destructive vibrations of hydraulic gates, such as are used for controlling flows in waterways, hydroelectric power plant, and so forth. While the structural geometries and fluid flow details in all these cases are very different, the underlying excitation mechanism is basically the same. The flexibility of the structure permits displacements due to the flowing fluid which fundamentally alters the fluid forces. If these structural-displacement-dependent fluid forces act in such a way as to increase the structural displacement, then the possibility of instability exists. Dynamic instability occurs when the energy transferred from the flow to the structure overcomes the structural damping. Increasing structural stiffness or modifying the structural shape are usually effective means of increasing the maximum permissible flow velocity for self-excited vibrations. Both methods have been used extensively to stabilize the galloping of bridge decks.

Self-controlled or resonant vibrations. Self-controlled vibrations occur when the fluid forces contain strong periodicities. As long as these periodicities are well removed from the structural natural frequencies, the structural response is small, there is no effect of the structural motion on the fluid forces, and the problem can be treated as a forced vibration. However, if the flow periodicity coincides with a structural natural frequency, a large-amplitude resonant response can occur. In this case, the structural response interacts with the flow excitation, increasing its strength and controlling its frequency; thus the appellation "self-controlled." This response is seen as a sharp peak in the amplitude response (Fig. 1, *B–C*). Resonance can be distinguished from a self-excited instability by the fact that the resonance response disappears at higher flow velocities. Instability responses continue to develop more violently with velocity increases above the stability threshold. *See* RESONANCE (ACOUSTICS AND MECHANICS).

The most common example of the self-controlled excitation mechanism is vortex-induced vibrations of cylinders in a cross flow. As the flow passes the cylinder, vortices are shed alternately into the wake. The flow pattern in the wake is called a von Kármán vortex street and the vortex shedding frequency, f_v, is linearly proportional to the flow velocity, V (**Fig. 3**, *A–E*). This process is scaled using the dimensionless parameter called the Strouhal number, St, where the characteristic length used is the cylinder diameter D. Thus, $\mathrm{St} = f_v D/V$, and has a value of about 0.2 for isolated circular cylinders. *See* DIMENSIONLESS GROUPS; KÁRMÁN VORTEX STREET; VORTEX; WAKE FLOW.

The antisymmetry of the wake vortices produces fluctuating lift forces on the cylinder at the frequency of the shedding of each pair of vortices. Thus, the cylinder vibrations occur primarily in the direction transverse to the flow direction. As the flow velocity is increased, the vortex shedding frequency, f_v, approaches the structural natural frequency, f_n. At frequency coincidence, $f_v = f_n$, resonance occurs (Fig. 3 point *B*). As the vibration amplitudes build up, the fluid lift forces increase in strength and the structural motion dominates the fluid mechanics, locking the vortex shedding frequency onto the structural natural frequency (Fig. 3, *B–C*). Ultimately, the lock-in is lost, the vibration amplitudes drop precipitously, and the vortex shedding frequency jumps back to the natural shedding frequency (Fig. 3, *C–D*). As the flow velocity is further increased, resonance in higher modes may occur.

Probably the most common example of the self-controlled mechanism is the vortex-induced vibrations of smoke stacks in a cross wind. However, other examples abound from submarine periscopes and marine structures in tidal currents to thermocouple probes in pipelines and bluff structures in a wind such as bridge decks. While vortex-induced vibrations in air flows rarely produce structural failures, it is common to use devices such as spiral strakes on smokestacks to minimize the vibrations. On the other hand, vortex shedding in liquid flows can cause serious structural (fatigue) failures so that design of such structures must consider these effects.

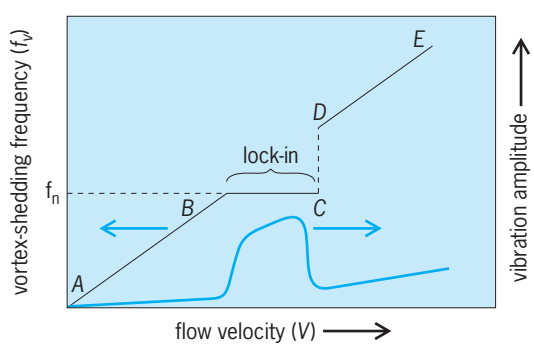

Fig. 3. Vortex-shedding frequency and vibration-response amplitude plotted against flow velocity, showing resonance and frequency control (lock-in).

Some structures are subject to both vortex shedding and self-excited vibrations. For example, bluff bodies such as bridge decks may be excited by both vortex shedding and galloping, and heat exchanger tube arrays experience both vortex shedding and fluid-elastic instability. In some cases, a large-amplitude vortex-shedding resonance can actually precipitate the onset of self-excited vibrations. In other cases, these phenomena occur at quite different flow velocities and behave independently.

David S. Weaver

Bibliography. M. K. Au-Yang, *Flow-Induced Vibration of Power and Process Plant Components*, ASME Press, 2001; R. D. Blevins, *Flow-Induced Vibration*, 2d ed., Van-Nostrand Reinhold, 1990, reprint, Krieger, 2001; E. Naudascher and D. Rockwell, *Flow-Induced Vibrations: An Engineering Guide*, Dover, reprint, 2005; M. P. Païdoussis, *Fluid-Structure Interaction: Slender Structures and Axial Flow*, vol. 1, Academic Press, London, 1998, vol. 2, Elsevier Academic Press, London, 2004; E. Simiu and R. Scanlan, *Wind Effects on Structures: Fundamentals and Applications to Design*, 3d ed., Wiley, 1996.

Flow measurement

The determination of the quantity of a fluid, either a liquid, vapor, or gas, that passes through a pipe, duct, or open channel. Flow may be expressed as a rate of volumetric flow (such as liters per second, gallons per minute, cubic meters per second, cubic feet per minute), mass rate of flow (such as kilograms per second, pounds per hour), or in terms of a total volume or mass flow (integrated rate of flow for a given period of time).

Flow measurement, though centuries old, has become a science in the industrial age. This is because of the need for controlled process flows, stricter accounting methods, and more efficient operations, and because of the realization that most heating, cooling, and materials transport in the process industries is in the form of fluids, the flow rates of which are simple and convenient to control with valves or variable speed pumps. *See* PROCESS CONTROL.

Measurement is accomplished by a variety of means, depending upon the quantities, flow rates, and types of fluids involved. Many industrial process flow measurements consist of a combination of two devices: a primary device that is placed in intimate contact with the fluid and generates a signal, and a secondary device that translates this signal into a motion or a secondary signal for indicating, recording, controlling, or totalizing the flow. Other devices indicate or totalize the flow directly through the interaction of the flowing fluid and the measuring device that is placed directly or indirectly in contact with the fluid stream. **Table 1** shows various metering devices, their operating principles, and flow ranges.

Volume Flow Rate Meters

Flow rate meters may modify the flow pattern by insertion of some obstruction or by modification of the shape of the walls of the conduit or channel carrying the fluid or they may measure the velocity without affecting the flow pattern.

Differential-producing primary devices. Differential-producing primary devices (sometimes called head meters) produce a difference in pressure caused by a modification of the flow pattern. The pressure difference is based on the laws of conservation of energy and conservation of mass. The law of conservation of energy states that the total energy at any given point in the stream is equal to the total energy at a second point in the stream, neglecting

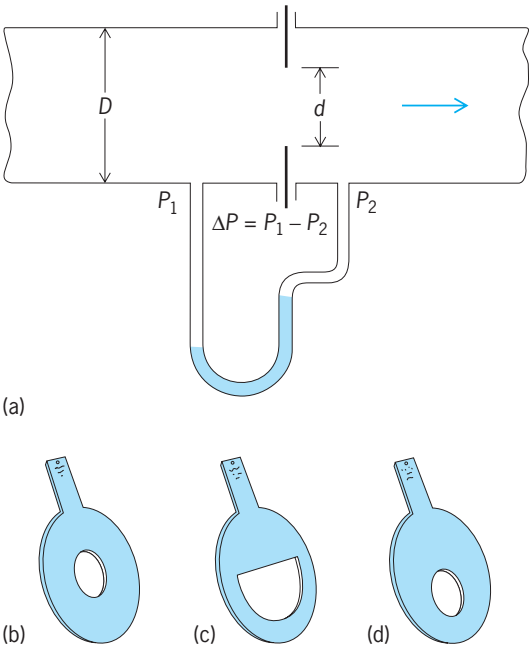

Fig. 1. Orifice plate primary device. (*a*) Schematic cross section. (*b*) Concentric plate. (*c*) Segmental plate. (*d*) Eccentric plate.

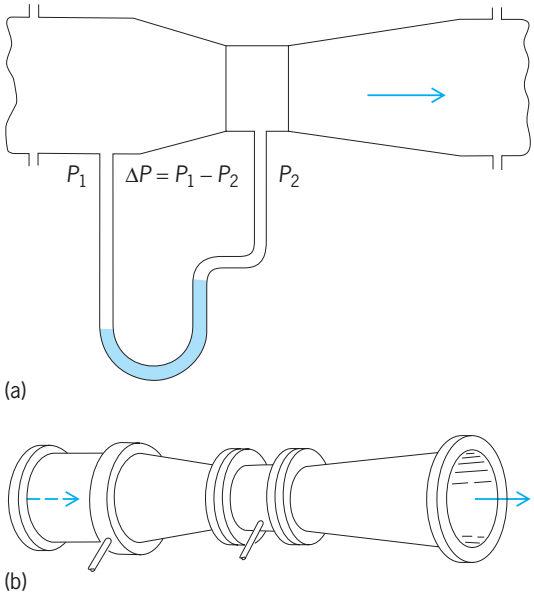

Fig. 2. Venturi tube. (*a*) Schematic cross section. (*b*) Perspective view.

TABLE 1. Examples of fluid meters

Name or type	Operating principle	Useful flow range
Orifice	Differential pressure (ΔP) over a restriction in the flow	3.5:1
Venturi		4:1
Nozzle		4:1
Flow tubes		4:1
Pitot-static	Compares impact pressure with the static pressure	3.5:1
Averaging pitot		3.5:1
Pipe elbow	Pressure from change in it	3.5:1
Target	Force on circular obstruction	3.5:1
Weir	Variable head, variable area	75:1
Flume	Variable head, variable area	up to 100:1
Rotameter (free float)	Automatic variable area	10:1
Piston		10:1
Plug		10:1
Magnetic	Electromagnetic induction	20:1
Tracere	Transit time	10:1
Nutating disk	Positive displacement; separation into and counting of discrete fluid quantities (most of these are available with high accuracy for custody transfer, with a less expensive, less accurate one for monitoring uses)	20:1
Rotary vane		25:1
Oval-shaped gear		50:1
Lobed impeller		10:1
Abutment rotor		25:1
Drum gas meter		40:1
		10:1 to 25:1
Gas turbine meter	Converts velocity to proportional rotational speed	10:1 to 25:1
Liquid turbine		10:1 to 50:1
Bearingless		10:1
Insertion turbine		25:1
Propeller		15:1
Helical impeller		15:1
Vortex cage		15:1
Axial flow mass	Momentum effect due to change in direction of fluid path	10:1
Gyroscopic/Coriolis		>10:1
Gyroscopic		10:1
Thermal loss	Heat extracted by fluid	10:1
Vortex flowmeter	Counting vortices shedding from an obstruction	8:1 to 100:1
Ultrasonic detection of vortex		8:1 to 100:1
Swirl	Precessing rotating vortex	10:1 to 100:1
Fluidic oscillator	Coanda wall attachment	30:1
Ultrasonic contrapropagating: clamp-on built-in spool piece	Time difference upstream/downstream	25:1
Ultrasonic contrapropagating: multiple path in spool piece		
Ultrasonic contrapropagating: axial in spool piece		
Ultrasonic correlation	Transit time of turbulence	10:1
Ultrasonic correlation: deflection (drift)	Deflection of beam across pipe	5:1
Ultrasonic contrapropagating: Doppler	Doppler shift-reflection from particles	10:1
Noise measurement	Spectral distribution of noise	5:1
Nuclear magnetic resonance	Transit time of nuclear disturbance	10:1
Laser Doppler velocimeter	Frequency shift of light reflected from particles	>1000:1

frictional and turbulent losses between the points. It is possible, however, to convert pressure (potential energy) to a velocity (kinetic energy), and vice versa. By use of a restriction in the pipe, such as an orifice plate (**Fig. 1**), venturi tube (**Fig. 2**), flow nozzle (**Fig. 3**), or flow tube (**Fig. 4**), a portion of the potential energy of the stream is temporarily converted to kinetic energy as the fluid speeds up to pass through these primary devices. *See* BERNOULLI'S THEOREM.

The secondary device measures this change in energy as a differential pressure ΔP, which is related to the flow by Eq. (1), where V is the velocity of flow

$$V = K_1\sqrt{\Delta P/\rho_f} \tag{1}$$

through the restriction and ρ_f is the mass density of the flowing fluid. The constant K_1 includes the orifice or throat diameter and the necessary dimension units, sizes, local gravity, and so forth. It also contains the discharge coefficient which has been experimentally determined and reduced to equation form for certain standard differential producers [namely, the square-edged orifice with specified pressure taps, ASME (American Society of Mechanical Engineers) nozzles, and classical venturis].

Mass flow for liquids, q_m, is given by Eq. (2), where K_2 is given by Eq. (3), and the quantities in Eqs. (2)

$$q_m = K_2\sqrt{\Delta P\rho_f} \tag{2}$$

$$K_2 = NSD^2 \tag{3}$$

and (3) are given in **Table 2**. The approximate value

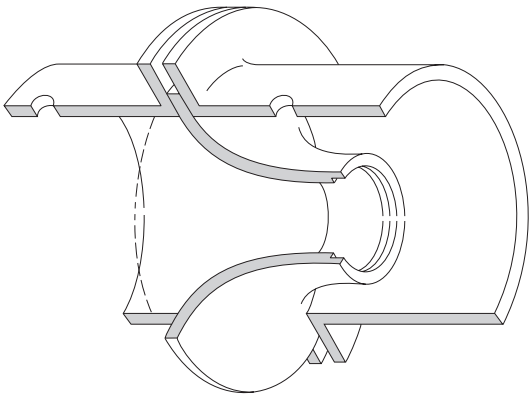

Fig. 3. Cutaway view of a flow nozzle.

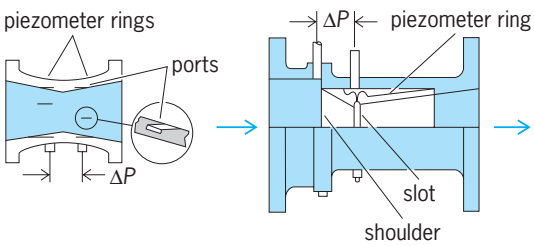

Fig. 4. Two types of flow tube. (*After 1969 Guide to process instrument elements, Chem. Eng., 76(12):137–164, 1969*)

of the sizing factor S, upon which K_2 depends, can be found from **Fig. 5**.

Beyond the restriction, the fluid particles slow down until the velocity is once more the same as that in the upstream conduit section. However, the static pressure is less due to the effect of turbulent and frictional losses associated with the acceleration and deceleration at the restriction. The magnitude of the loss depends upon the configuration of the primary device. *See* METERING ORIFICE.

Orifice plate. The common orifice plate (Fig. 1) is a thin plate inserted between pipe flanges, usually having a round, concentric hole with a sharp, square upstream edge. The extensive empirical data available and its simplicity make it the most common of the primary devices. Segmental and eccentric orifices are useful for measurement of liquids with solids, and of vapors or gases with entrained liquids.

Venturi tube. This device (Fig. 2) has the advantage of introducing less permanent loss of pressure than the orifice plate. The converging inlet cone permits solids and dirt to be flushed through it, and the outlet cone reduces the turbulent losses.

Flow nozzle. The flow nozzle (Fig. 3) also has a converging inlet, but it has no diverging outlet section. The pressure loss is similar to that with an orifice plate for the same differential pressure. It is used where solids are entrained in the liquid and also where fluid velocities are particularly high (it will pass more fluid than an orifice of the same diameter).

Flow tubes. Several designs of primary devices have been developed to give lower permanent pressure loss than orifices and shorter overall length than standard venturis, and they sometimes incorporate piezometer chamber averaging of the pressure tap. Figure 4 shows two such designs.

Pitot tubes. In industrial flow measurement, pitot tubes are used in spot checks and for comparative measurements to determine trends. A pitot tube measures the difference between total impact pressure and static pressure. In the device shown in **Fig. 6** the central tube receives the impact pressure, while the holes in the outer tube are subjected to the static pressure. A secondary device senses the difference in these pressures and converts it into an indication of fluid velocity at the location of the tip. There are a wide variety of other tube configurations available for sensing both impact and static pressure or combinations thereof. Several readings should be taken across the pipe along a diameter (preferably two diameters) of the pipe and the average velocity determined. The pitot tube is used extensively as a laboratory device for the point velocity measurement and also for speed measurement of vehicles and in wind tunnel work.

The averaging pitot tube stimulates an array of pitot tubes across the pipe. It consists of a rod extending across the pipe with a number of interconnected upstream holes and a downstream hole for the static pressure reference. In one design, the diamond shape of the rod is claimed to present a very

TABLE 2. Quantities used in liquid-flow calculation

Symbol	Quantity	Unit or value: SI units	U.S. Customary units	
N	Constant that depends upon the system of units used	1.11	0.0999	0.52
D	Inside diameter of pipe	meter	inch	inch
ΔP	Pressure difference	pascal	in. H_2O*	lbf/in.^{2} (psi)
ρ_f	Density of the flowing fluid	kg/m^3	lb$_m$/ft^3	lb$_m$/ft^3
β	Ratio, (d/D)	dimensionless		
C	Discharge coefficient	dimensionless		
S	Sizing factor $= C\beta^2/\sqrt{1-\beta^4}$	dimensionless		
d	Orifice or throat diameter	meter	inch	inch
q_m	Mass flow rate	kg/s	lb$_m$/s	lb$_m$/s

*The pressure of a column of water 1 in. (25.4 mm) high at 68°F (20°C) and acceleration of gravity of 32.174 ft/s^2 (9.80665 m/s^2).

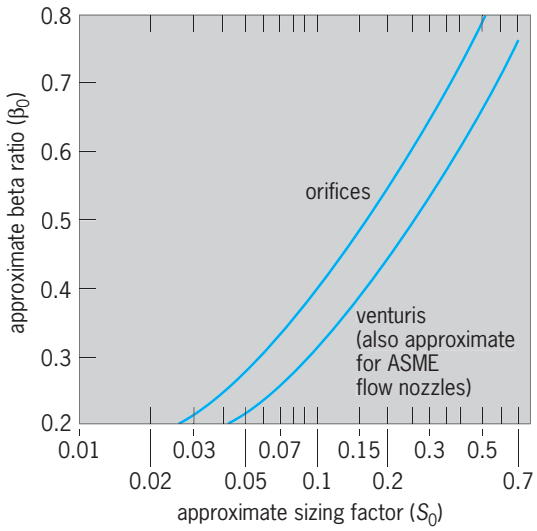

Fig. 5. Sizing factors for differential-primary-producing elements, for liquid flows, based on $C = 0.6 + 0.06\,S$ for orifices, $C = 0.995$ for venturis. The factors do not include the Reynolds number correction. (*Foxboro Co.*)

low flow obstruction and still, with the interconnected ports, to provide a differential pressure that is representative of the flow relatively unaffected by profile differences. *See* PITOT TUBE.

Pipe elbow. Fluid flow around the bend in a pipe elbow (**Fig. 7**) produces a centrifugal force greater on the outside than on the inside, and thereby produces a difference in pressure between pressure connections on the inside and outside of the bend. This differential pressure is used as an indication of the volume flow rate. This device is often used for rough flow measurement because it is not necessary to disturb existing piping. Carefully installed and calibrated, the device can provide an accurate measurement.

Differential-pressure secondary devices. All the primary devices discussed above create a difference in pressure between two points. This differential pressure must be measured and converted by the equations given to obtain flow. Differential pressure may be measured accurately by simple liquid-filled (usually mercury) U-tube manometers or by more refined types of meters, such as the float-operated mechanism or the weight-balanced, ring-type meter, to provide controlling, recording, and totalizing functions. These devices (**Fig. 8**) are ordinarily connected directly to the pressure connections on the primary device and are therefore exposed to the process fluids. Sometimes an inert-liquid-seal fluid is used to isolate the secondary device from corrosive process fluids. In either case, the manometers are best located adjacent to the primary device. These manometers are accurate if the mercury is clean and the fluid flow rate is not changing too rapidly. *See* MANOMETER.

Dry-type sensors. Diaphragm- and bellows-type differential-pressure measuring devices were developed to eliminate manometer fluids and to provide faster response and easier installations. **Figure 9** shows a representative bellows-type differential-pressure sensor. The range of these instruments is changed by removing one cover and adjusting the range spring or replacing it with one of a different spring rate. Thermal expansion of the fill fluid contained in the two bellows is compensated for by the bimetal-supported additional convolutions shown at the left of the figure. These are very stable and reliable instruments.

Many direct-connected meters have been modified to provide a pneumatic or electric signal to remote instrumentation. Serviceable transmitters have been developed. Two types of meters, the force-balance and deflection types, are suited to either pneumatic or electric transmission of the flow signal.

The force-balance type (**Fig. 10**) causes the variation of an air pressure in a small bellows to oppose the net force created by the differential pressure on opposite sides of a diaphragm or diaphragm capsule. This air pressure is used as a signal representing the flow rate. An electronic force-balance design utilizes a restoring force due to a force motor and uses an inductive detector for unbalance; the current in the motor is the transmitted (4–20 mA) signal. One electronic version provides an output signal proportional to flow by making the feedback force the interaction between two motors, with the current flowing through them (the force is squared).

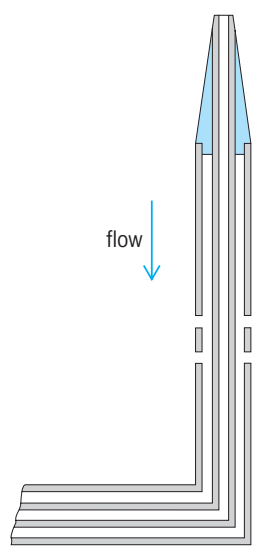

Fig. 6. Laboratory-type pitot tube.

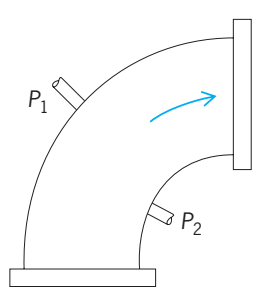

Fig. 7. Pipe elbow. Difference between outside pressure P_1 and inside pressure P_2 indicates volume flow rate.

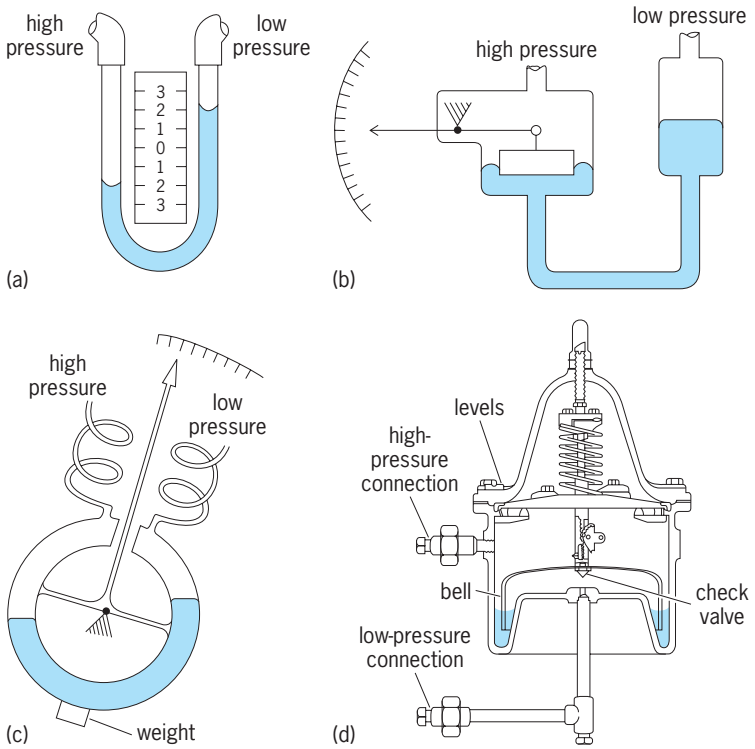

Fig. 8. Differential-pressure secondary devices. (*a*) U-tube manometer. (*b*) Mercury float mechanism. (*c*) Weight-balanced ring-type meter. (*d*) Bell type (*Foxboro Co.*).

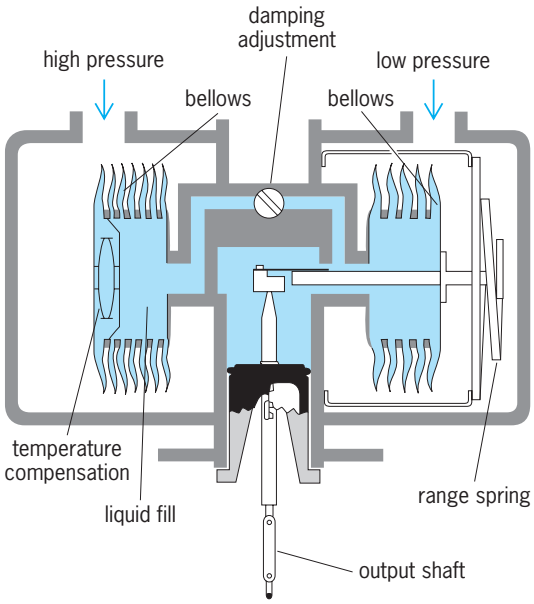

Fig. 9. Bellows-type differential-pressure-measuring element. (*Foxboro Co.*)

Many types of deflection differential-pressure cells are offered commercially. Several apply the differential pressure across a diaphragm containing or connected to a strain gage (bonded-semiconductor, diffused strain gage in a silicon diaphragm, and so forth; **Fig. 11***a*). A representative two-wire circuit is shown in Fig. 11*b*. The signal and power are carried by the same two wires.

Some differential-pressure cells use variable capacity or differential capacity between a diaphragm and an adjacent wall (Fig. 11*c*). Some use the moving diaphragm to cause a variable reluctance or variable inductance. In most cases the measuring elements are protected from the fluid being measured by a pair of seal-out diaphragms or bellows, and the elements are submerged in silicon oil. In one type the measuring diaphragm provides a variable tension on a wire that is placed in a magnetic field and connected to a self-oscillating circuit that provides a frequency that varies with the differential pressure (Fig. 11*d*).

Readout scales. Since these secondary devices measure the differential pressure developed, their outputs are proportional to the square of the flow rate. In the case of direct-reading devices, they must be fitted with square root scales and charts for reading flow rate. These scales are particularly difficult to read at low values, thereby limiting the useful rangeability. Uniform flow-rate scales are obtained with manometers with specially shaped chambers or with shaped plugs (Ledoux bell) or with special cams in the linkage between the manometer and the indicating pointers or pens. Some differential-pressure transmitters provide linear-with-flow transmitted signals by incorporating squaring functions in the feedback, or mechanical or electronic square rooting can be done either at the transmitter or at the receiver location.

Target meter. The target-type meter (**Fig. 12**), besides being particularly suitable for dirty fluids, has the important advantage of a built-in secondary device and is without the disadvantage of fluid-filled connection pipes to the secondary device.

Open-channel meters. Flow in open channels is determined by measuring the height of liquid passing through a weir or flume. Consistent results require that the measuring device be protected from direct impact of the flowing stream. Because the height-flow velocity relationship is exponential, the secondary device frequently uses cams to obtain a uniform flow scale.

Variable-area meters. These meters work on the principle of a variable restrictor in the flowing stream

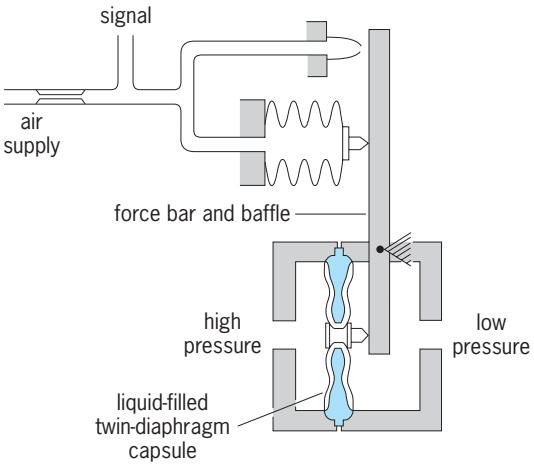

Fig. 10. Diaphragm meter, force balance type.

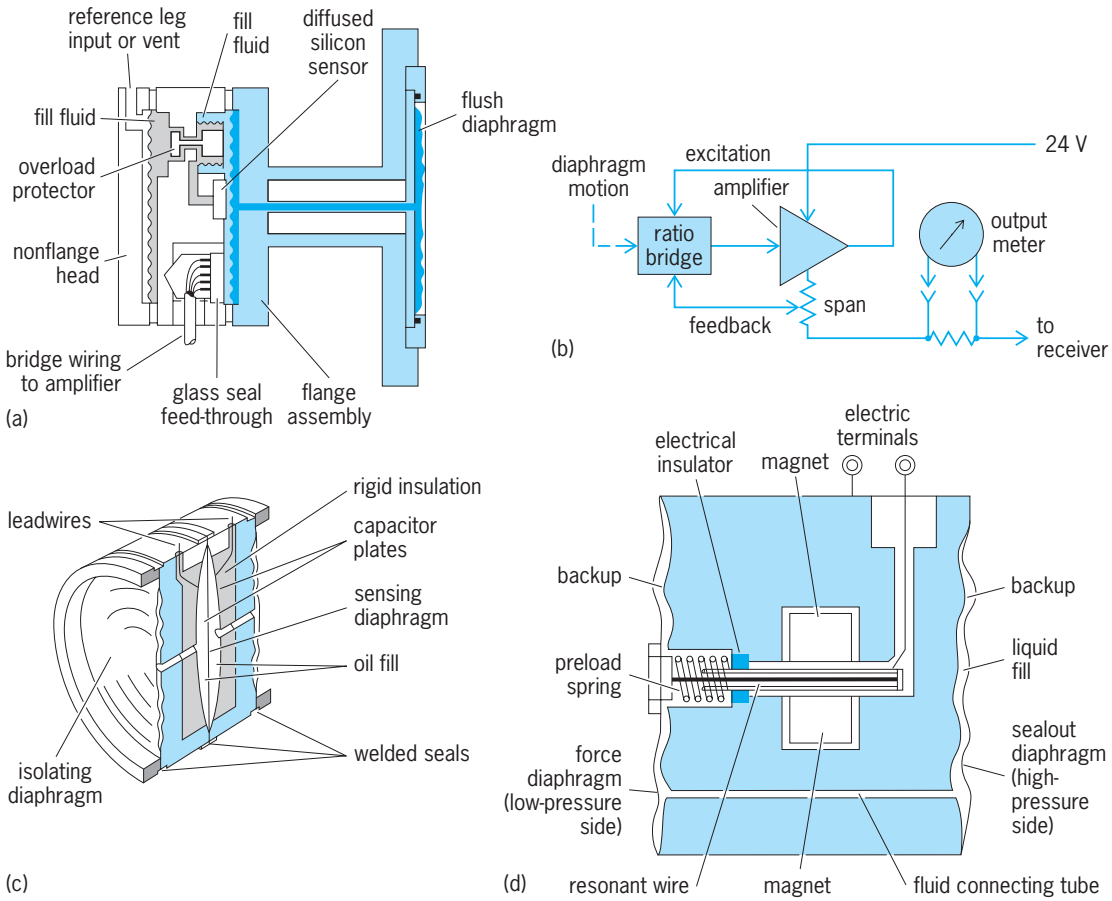

Fig. 11. Differential-pressure sensors. (*a*) Silicon-chip type (*Honeywell Industrial Division*) with (*b*) 4- to 20-mA two-wire transmitter circuit (*after 1969 Guide to process instrument elements, Chem. Eng., 76(12):137–164, 1969*). (*c*) Capacitance type (*Rosemount Engineering Corp.*). (*d*) Resonant-wire type (*Foxboro Co.*).

being forced by the fluid to a position to allow the required flow-through. Those that depend upon gravity provide essentially a constant differential pressure, while those that supplement the gravity with a spring will have a variable differential pressure.

In the tapered-tube rotameter (**Fig. 13***a*), the fluid flows upward through a tapered tube, lifting a shaped weight (possibly misnamed a float) to a position where the upward fluid force just balances the float weight, thereby giving a float position that indicates the flow rate. For remote indication, and also where the pressure requires a metal tube, a center rod attached to the float carries the core of an inductance pickup for transmission of the float position through a nonmagnetic tube seal. In other designs the rod carries one side of a magnetic coupling to a follower mechanism outside of the nonmagnetic tube.

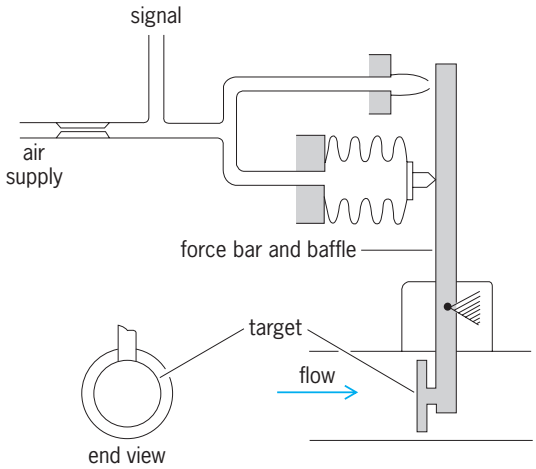

Fig. 12. Target flowmeter. (*Foxboro Co.*)

In a position type (Fig. 13*b*) the bouyant force of the liquid carries the piston upward until a sufficient area has been uncovered in a slot in the side of the vertical tube to allow the liquid to flow through the slot. The position of the piston indicates the flow rate and, by properly shaping the slot, this type of meter may be made with a uniform flow scale. A spring is used in some designs to supplement the gravity of the plug and further characterize the indication.

In the tapered-plug design (Fig. 13*c*), the plug, located in an orifice, is raised until the opening is sufficient to handle the fluid flow. Readouts can be the same as in the tapered-tube design.

A spring-loaded variable-area meter (Fig. 13*d*) provides a differential pressure for a secondary

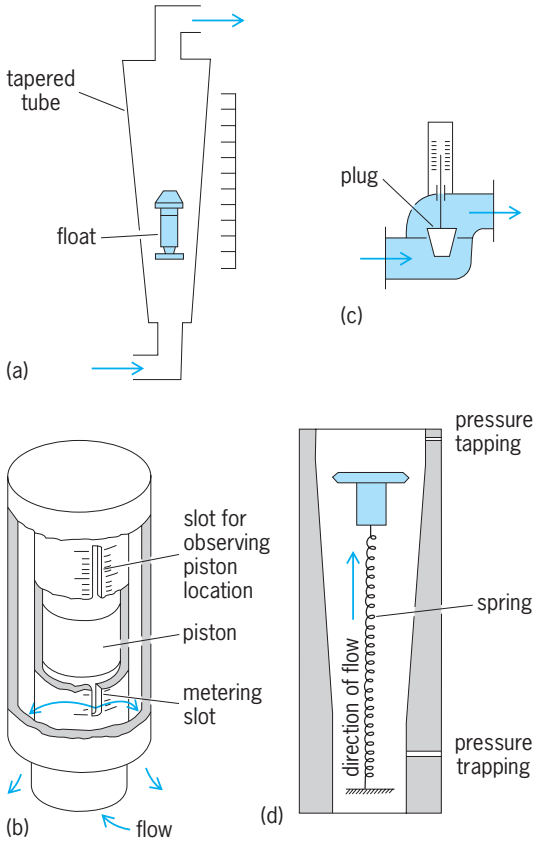

Fig. 13. Variable-area flowmeters. (*a*) Rotameter type. (*b*) Piston type. (*c*) Plug type (*after 1969 Guide to process instrument elements, Chem. Eng., 76(12):137–164, 1969*). (*d*) Spring-loaded differential-pressure producer (*after Institute of Mechanical Engineers, How to Choose a Flowmeter, 1975*).

device. By combining the characteristics of a spring-loaded variable-area meter and a standard differential-pressure measurement across the float, it is claimed that the rangeability of the measurement can be raised to as high as 50:1.

Electromagnetic flowmeter. A magnetic field is applied across a metering tube carrying a flowing liquid, and the voltage generated between two perpendicularly located electrodes contacting the liquid is measured by a secondary or transmitter instrument. Since this voltage is strictly proportional to the average velocity, a linear volume flow signal results. Practical flow measurements can be made on any liquid that is at all conductive. The linear scale allows accurate flow measurements over a greater range than is possible in the differential-pressure meters. The lack of obstruction to fluid flow makes it possible to measure thick slurries and gummy liquids. In order to avoid errors due to electrodes becoming coated with an insulating material, some electromagnetic flowmeters are equipped with ultrasonic electrode cleaners. *See* ELECTROMAGNETIC INDUCTION.

Classical electromagnetic flow tubes power the field coils with a continuous ac mains voltage and measure the induced ac voltage on the electrodes. After an approximately 90° phase shift, a ratio is made of the two voltages to correct for mains voltage change. Some electromagnetic flowmeters apply a periodic dc voltage pulse to the coils and measure the dc voltage across the electrodes at a carefully selected time relative to the field excitation. These meters are available with nonmagnetic stainless steel pipe and various types of insulating lining, and with all-plastic construction. When the field coils are mounted inside the pipe, a magnetic iron may be used for the pipe.

Tracer method. This classic method consists of injecting a small quantity of salt into the flowing stream and measuring the transit time of passage between two fixed points by measuring electrical conductivity at those points. Inaccuracies are introduced by variations in flow cross section in the pipe or channel being measured. Because of the convenience for a quick check of flow under comparatively adverse conditions, standards have been written around the use of salt, radioactive material, and even pulses of heat. Appropriate sensors are offered with correlation techniques for improving the accuracy.

Quantity Meters

Quantity meters are those that measure or pass fixed amounts of fluid. By introducing a timing element, all quantity meters can be converted to volume flowrate meters. Quantity meters may be classified as positive-displacement meters or rotating impeller-type meters. These are illustrated in **Figs. 14** and **15**.

Positive-displacement meters. The flowing stream is separated into discrete quantities by these meters, which capture definite volumes one after another and pass them downstream. Some representative examples will be discussed.

Nutating-disk meter. A circular disk attached to a spherical center is restrained from rotation by a vertical partition through a slot in the disk (Fig. 14*a*). It is allowed to nutate so that its shaft is in contact with a conical guide. The liquid entering the inlet port moves the disk in a nutating motion until the liquid discharges from the outlet port. The double conical shape of the measuring chamber makes a seal with the disk as it goes through the nutating motion. Normally a mechanical counter indicates the number of cycles as the fluid flows through smoothly with no pulsations.

Rotary-vane meter. An eccentrically mounted drum carries radial spring-loaded vanes (Fig. 14*b*). These vanes slide in and out, and seal against the meter casing to form pockets that carry a measured amount of fluid on each cycle.

Oval-shaped gear meter. Two carefully formed gears provide their own synchronization, and with close-fitting teeth to the outer circular-shaped chambers they deliver a fixed quantity of liquid for each rotation (Fig. 14*c*). Rotation of the output shaft can be recorded smoothly by a contact on each revolution. If a higher-frequency signal is desired for flow control

or closer monitoring, a compensating gear or optical disk can be used to make the output uniform for a constant liquid flow.

Lobed impeller. The lobed impeller (Fig. 14*d*) is similar to the oval gear except that the synchronization is done by external circular gears and therefore the rotation is smooth.

Rotary-abutment meter. Two displacement rotating vanes interleave with cavities on an abutment rotor (Fig. 14*e*). The three elements are geared together, and the meter depends upon close clearances rather than rubbing surfaces to provide the seals.

Liquid-sealed drum-type gas meter. A cylindrical chamber is filled more than half full with water and divided into four rotating compartments formed by trailing vanes (Fig. 14*f*). Gas entering through the center shaft from one compartment to another forces rotation that allows the gas then to exhaust out the top as it is displaced by the water.

Remotely indicating meters may incorporate a simple electronic contact, or a magnetic or optical sensor which can transmit one or more pulses per revolution.

Rotating impeller meters. These convert velocity to a proportional rotational speed. They include various types of turbine meters, propellers, helical impellers, and vortex cage meters.

Turbine meters. The gas turbine meter (Fig. 15*a*) is characterized by a large central hub that increases the velocity of the gas and also puts it through the tips of the rotor to increase the torque. Readout can be by a directly connected mechanical register or by a proximity sensor counting the blade tips as they pass by. Low-friction ball bearings require special lubricants and procedures.

Most of the available types of liquid-turbine meters (Fig. 15*b*) are designed to cause a fluid flow or hydraulic lift to reduce end thrust, and several depend upon the flowing fluid to lubricate the bearings.

One bearingless-turbine meter is designed to float a double-ended turbine on the flowing fluid and thereby avoid the problem of wearing bearings. The turbine rotor can be made of various materials to suit the fluid, making it an attractive flow measurement device for nonlubricating and corrosive fluids.

In one insertion meter a small propeller or turbine rotor is mounted at right angles to the end of a support rod. Readout is normally through an inductive pickup and appropriate electronics for obtaining rate of flow. The propeller can be inserted into a flowing stream or into a closed pipe with a direct readout of the flow rate at the location of the propeller. Proper location can give an average flow rate either locally or remotely.

Propeller, helical impeller, and vortex cage meters. Since the elements in these meters (**Fig. 16**) rotate at a speed that is linear with fluid velocity, the useful rangeability is large and limited mostly by the bearings. For local indication, revolutions can be counted by coupling to a local-mounted counter. Proximity detectors, giving pulses from 1 to as many as 100 pulses per revolution, can be used for transmitting to remote integrators or flow controllers.

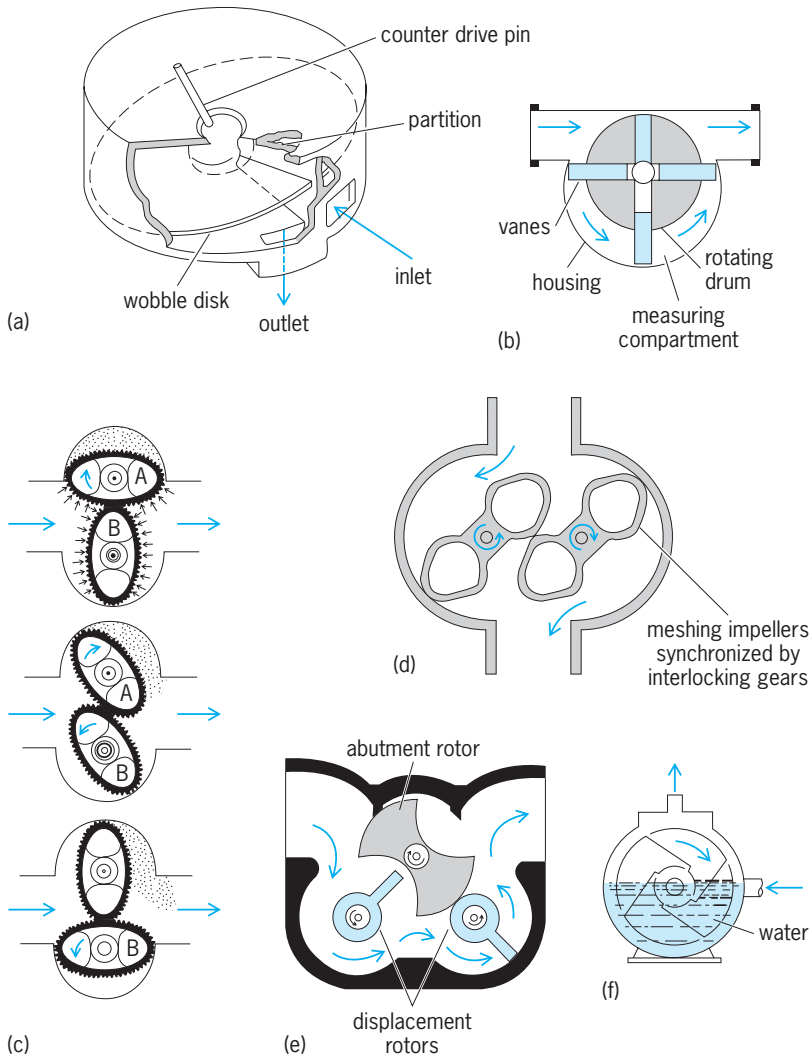

Fig. 14. Positive-displacement flowmeters. (*a*) Nutating disk. (*b*) Rotary vane. (*c*) Oval-shaped gear (*Brooks Instrument Division*). (*d*) Lobed impeller. (*e*) Rotary abutment. (*f*) Liquid-sealed drum-type gas flowmeter (*after 1969 Guide to process instrument elements, Chem. Eng., 76(12):137–164, 1969*).

Flow correction factors. The accuracy of quantity meters is dependent to varying degrees on the physical properties of the fluid being measured. Changes from the reference (calibration) conditions of viscosity and density must be considered if accurate measurement is required. It is usually the practice to take pressure and temperature measurements of vapors and gases, or temperature measurement of liquids, at the flow-metering device so that corrections either may be applied by the mechanism of a recording or totalizing device or introduced manually. Some meters are available in which these corrections are introduced automatically with great accuracy, while others are made in which only a first-order approximation is introduced. Proper selection is determined by the requirements of the installation. A meter compensated for fluid conditions can become a mass flowmeter.

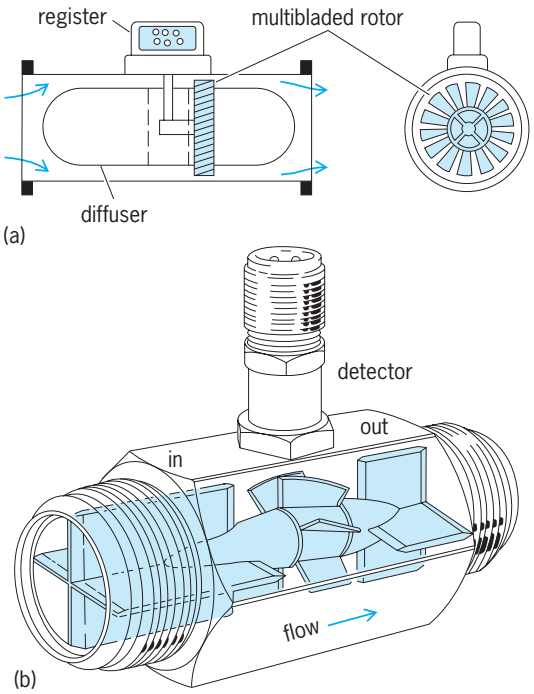

Fig. 15. Turbine flowmeters. (*a*) Gas turbine flowmeter (*after Gas meters, Factory Mag., 9(1):39–45, January 1976*). (*b*) Liquid turbine flowmeter (*Foxboro Co.*).

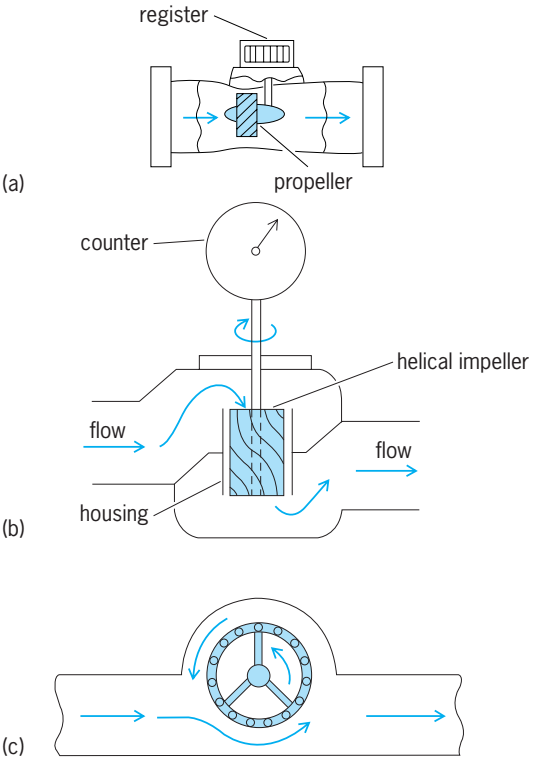

Fig. 16. Types of quantity flowmeters. (*a*) Propeller. (*b*) Helical impeller. (*c*) Vortex cage.

Mass Flow Rate Meters

There have been many attempts to develop practical industrial mass flowmeters. The most consistently satisfactory method of determining mass flow has been through calculation using volume flow and density determined either by measurement or inference from pressure and temperatures (**Fig. 17**). Only a few of the mechanical type will be mentioned here.

In one type, the fluid in the pipe is made to rotate at a constant speed by a motor-driven impeller. The torque required by a second, stationary impeller to straighten the flow again is a direct measurement of the mass flow (**Fig. 18*a***).

A second type, described as gyroscopic/Coriolis mass flowmeter, employes a C-shaped pipe and a T-shaped leaf spring as opposite legs of a tuning fork (Fig. 18*b*). An electromagnetic forcer excites the tuning fork, thereby subjecting each moving particle within the pipe to a Coriolis-type acceleration. The resulting forces angularly deflect the C-shaped pipe an amount that is inversely proportional to the stiffness of the pipe and proportional to the mass flow rate within the pipe. The angular deflection (twisting) of the C-shaped pipe is optically measured twice during each cycle of the tuning-fork oscillation (oscillating at natural frequency). A digital logic circuit converts the timing of the signals into a mass-flow-rate signal.

In another configuration (using the Coriolis effect) a ribbed disk is fastened by a torque-sensing member within a housing, and both are rotated at constant speed. The fluid enters at the center of the disk and is accelerated radially; the torque on the disk is a direct measure of the mass flow rate. Unfortunately, the practical problem of maintaining the rotating seals in the device has hindered its commercial development. *See* CORIOLIS ACCELERATION.

Another type is based on the principle of the gyroscope (Fig. 18*c*). The flowing stream is sent through a pipe of suitable shape so that the mass of material flowing corresponds to a gyro wheel rotating about axis C. The entire gyrolike assembly is rotating about axis A, and a torque is produced about axis B proportional to the angular momentum of the simulated gyro wheel. This torque is therefore directly proportional to the mass rate of flow. The pretzel-like configuration of pipe is introduced to eliminate centrifugal and other extraneous effects. *See* GYROSCOPE.

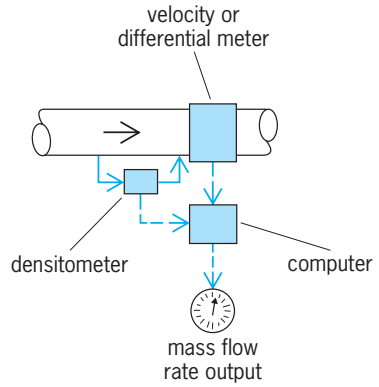

Fig. 17. Inferential mass-flowmeter which computes real-time flow rate in terms of volume rate times gas density. (*After Gas meters, Factory Mag., 9(1):39–45, January 1976*)

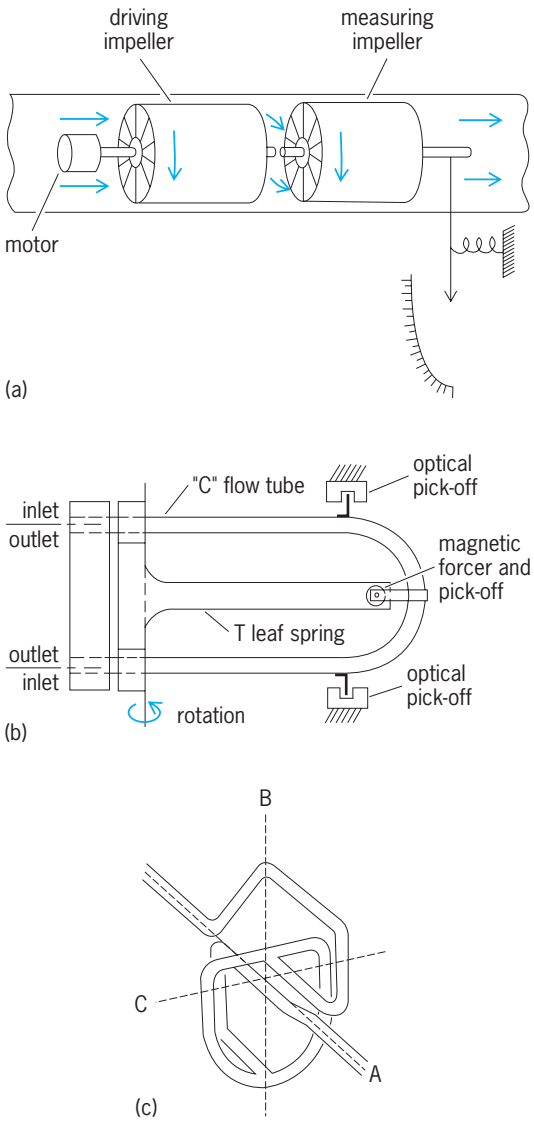

Fig. 18. Mass flowmeters. (*a*) Axial-flow type (*General Electric*). (*b*) Gyroscopic/Coriolis type (*after K. O. Plache, Coriolis/Gyroscopic Flow Meter, ASME Pap. 77-WA/FM/4, American Society of Mechanical Engineers, 1977*). (*c*) Gyroscopic type.

Thermal Flow Measurement Devices

Several heated or self-heating types of sensor are used for gas-velocity or gas-mass flowmeters.

Hot-wire anemometer. A thermopile heated by a constant ac voltage is cooled by the gas flow. The degree of cooling, proportional to local flow velocity, is indicated by the thermopile millivoltage output. *See* THERMOCOUPLE.

Heat-loss flowmeter. A constant current is supplied to a thermistor bead. The temperature of the bead depends upon the cooling effect of the mass flow of the fluid, and the measure of the voltage drop across the bead is therefore an indication of the flow. A second thermistor bead is used to compensate for the temperature of the fluid. *See* THERMISTOR.

Another heat-loss flowmeter measures the amount of electrical heating that is required to keep the temperature of a platinum resistor constant since the amount of heat carried away changes with the velocity of the fluid. A second sensor is used to compensate for the temperature of the fluid.

Vortex Flowmeters

Flowmeters have been developed based on the phenomena of vortex shedding and vortex precession. *See* VORTEX.

Vortex-shedding flowmeters. One of the newest successful methods of measuring flow is based on one of the oldest observed phenomena relating to flow— the eddies (vortices) forming behind an obstruction in a flowing stream. Close observation has shown that each vortex builds up and breaks loose, followed by a repeat in the opposite rotation on the other side of the obstruction, and so alternately the vortices shed and go downstream in what is called a vortex street. Further observation shows that the distance between successive vortices is directly proportional to the blocking width of the obstruction. The distance is also dependent of the fluid velocity over a large range, and thus the frequency of the shedding is directly proportional to the velocity of the fluid. *See* KÁRMÁN VORTEX STREET.

Application of vortex-shedding flowmeters is normally limited to the turbulent-flow regime, and various shedding-element shapes and internal constructions have been developed in order to improve the shedding characteristics and the detection of the shedding. The common characteristic is sharp edges around which the vortices can form.

Sensing of the shedding vortices is done by (1) measuring the cyclical change in flow across the front surface of the shedder (self-heated resistance elements); (2) sensing the change in pressure on the sides of the "tail" of the shedder (inductive sensing of motion of a "shuttlecock" in a cross passage, or piezoelectric or capacitive sensing of the pressure on diaphragms sealing a cross passage); (3) sensing the forces on the shedder or an extended tail (piezoelectric sensing of the bending of the shedder, or strain-gage sensing of tail motion); (4) sensing of the vortices after they have left the shedder (by an ultrasonic beam across the wake); (5) sensors mounted outside of the pipe at the junction between two tubes connected to openings in the sides of the tail of the shedder.

One design, which uses a rod or a taut wire as the obstruction and as the ultrasonic detector, is offered for open-channel measurement as well as ship velocity.

Vortex flowmeters are available for measuring the flow of liquids, vapors, and gases. The pulse rate output is particularly convenient for input to integrators for recording total flow and for digital computers. The linear-with-flow output makes the measurement particularly attractive for widely varying flows. However, the relatively low frequency (as low as 1 Hz) of the large blockage designs must be considered when using this measurement in a flow-control loop.

Vortex-precession flowmeter. In this instrument, used to measure gas flows, a swirl is imparted to

the flowing fluid by a fixed set of radial vanes. As this swirl goes into an expanding tube, it forms a precessing vortex that is detected by a heated thermistor sensor. The frequency of the precession is claimed to be proportional to the volumetric flow rate within an accuracy of 0.75%.

Fluidic-Flow Measurement

Two types of flowmeters based on fluidic phenomena have been developed: the fluidic oscillator meter and the fluidic flow sensor.

Fluidic-oscillator meter. This type of meter works on the principle of the Coanda effect, the tendency of fluid coming out of a jet to follow a wall contour. The fluid entering the meter will attach to one of two opposing diverging side walls (**Fig. 19***a*). A small portion of the stream is split off and channeled back through the feedback passage on that side to force the incoming stream to attach to the other side wall (Fig. 19*b*). The frequency of the oscillation back and forth is directly proportional to the volume flow through the meter. A sensor detects the oscillations and transmits the signal. This principle is used in both liquid and gas flowmeters.

Fluidic-flow sensor. Figure 20 shows this device (also known as a deflected-jet fluidic flowmeter) as used in gas flows, particularly in dirty environments. It consists of a jet of air or other selected gas directed back from the outer nozzle onto two adjacent small openings. The flow of the gas being measured will deflect the jet to change the relative pressure on the two ports and thereby give a signal corresponding to the gas velocity.

Ultrasonic Flow Measurement

Ultrasonic flowmeters are available in many physical arrangements and are based on a great variety of fundamental principles. Most are found in one of the following categories.

Contrapropagating meter. Contrapropagating diagonal-beam designs (**Fig. 21***a*) have transducers mounted in sockets machined in the pipe or welded on wedges, and also as clamp-on elements. The difference between upstream and downstream "flight" times combined with the speed of sound in the fluid will give fluid velocity. Since the flow near the middle of the pipe is faster, a multiple parallel-beam design (Fig. 21*b*) will give a better average. In a symmetrical profile a beam offset from the center of the pipe may also give a better average.

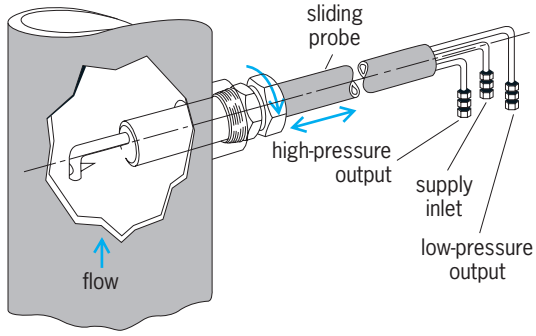

Fig. 20. Fluidic flow sensor. (*Fluid Dynamics Devices*)

Axial meter. For small pipes, the short path length across the fluid stream makes measurement difficult. Figure 21*c* shows a contrapropagating arrangement for measuring lengthwise over a relatively longer path where timing measurements are more practical.

Correlation meter. Two pairs of elements with beams across the pipe can be used with a correlation circuit to detect the time it requires for discontinuities in the fluid stream to pass between the two detectors (Fig. 21*d*).

Deflection (drift) meter. An ultrasonic beam directed across the pipe will be deflected by an amount depending on the relative velocity of sound in the fluid and the flowing rate of the fluid (Fig. 21*e*). One arrangement detects this deflection with two receivers mounted on the other side of the pipe. The relative intensity of the two receivers is an indication of the flowing-fluid velocity. A change in the velocity of sound in the fluid will introduce an error.

Doppler meter. Flow rate can be detected by the Doppler shift of a reflection from particles or discontinuities in the flowing fluid. A transceiver on one side of the pipe may have an error due to the reflections from the slower particles near the wall. Figure 21*f* shows the measurement being made by opposing elements. *See* DOPPLER EFFECT.

Noise-type meter. One flow measurement device is available where the noise that is generated by the flowing fluid is measured in a selected frequency band, with a useful but relatively inaccurate readout (Fig. 21*g*).

Open channels. A different application of ultrasonics to flow measurement can be found in flumes and weirs where the level of the flowing fluid, measured by ultrasonic reflection, is used to calculate the flow rate. In unrestricted open channels the level can be measured with one reflecting transducer (either above or below the surface) and the flowing rate with a pair of submerged detectors. An electronic

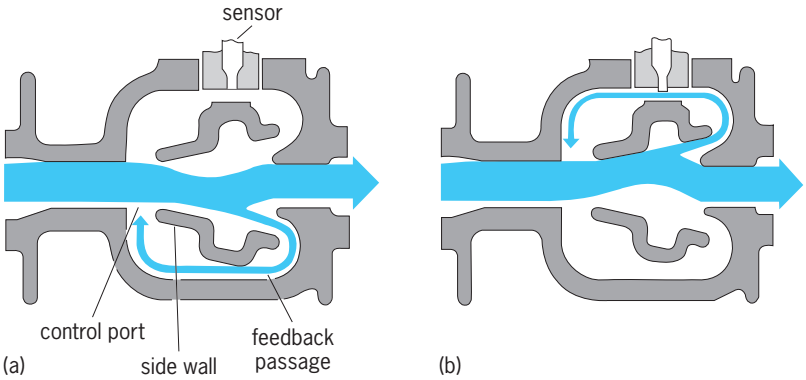

Fig. 19. Fluidic oscillator meter in which feedback stream attaches to (*a*) first one wall, and (*b*) then the other. (*After S. J. Bailey, Tradeoffs complicate decisions in selecting flowmeters, Control E Eng., 27(4):75–79. 1980*)

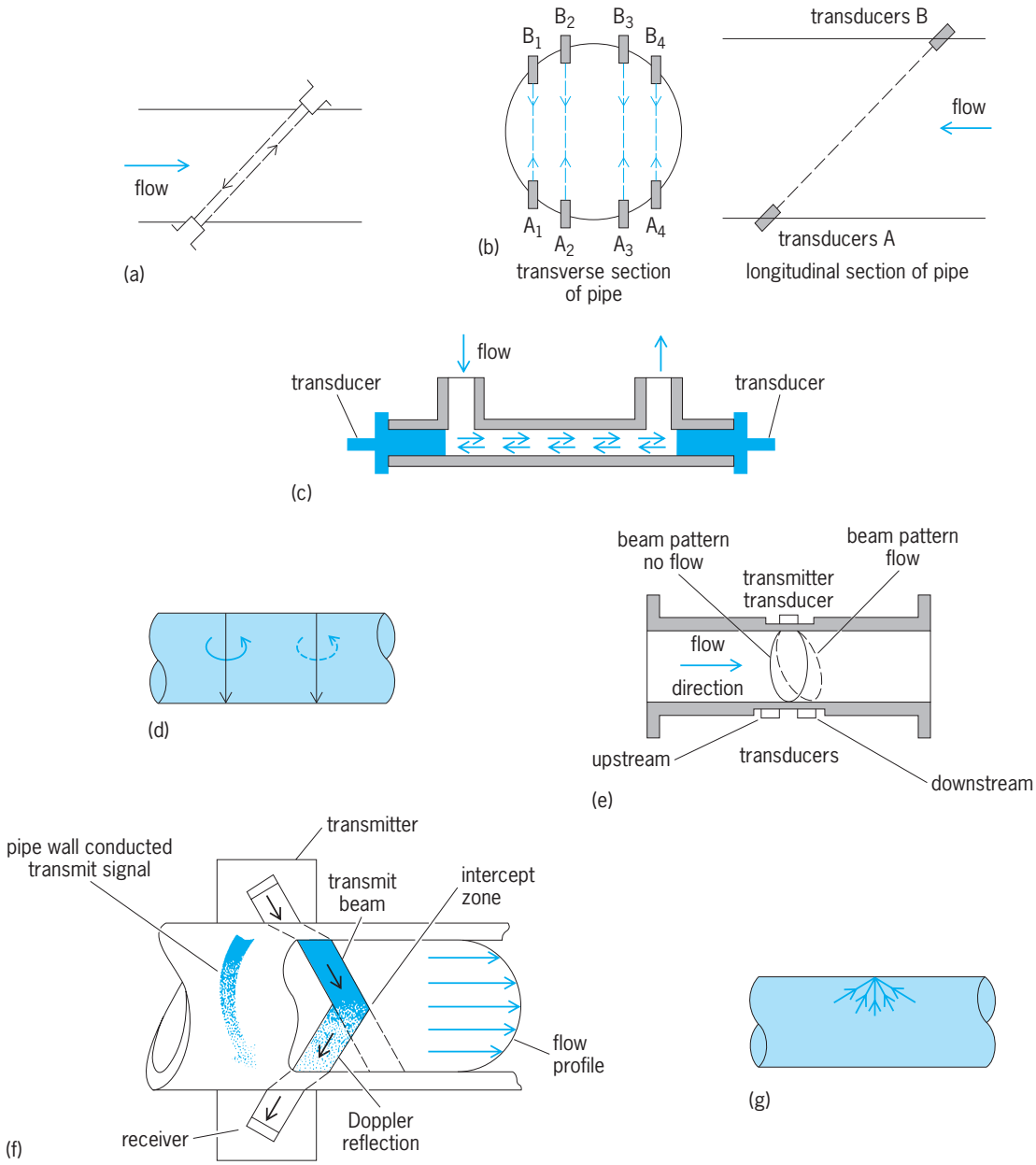

Fig. 21. Ultrasonic flowmeters. (*a*) Contrapropagating type (*after Fischer and Porter Co., Publ. 12257, 1957*). (*b*) Multiple-beam contrapropagating type (*after Institute of Mechanical Engineers, How To Choose a Flowmeter, 1975*). (*c*) Axial-transmission contrapropagating type (*after Dupont Co., Industrial Products*). (*d*) Correlation type (*after W. P. Mason and R. Thurston, eds., Physical Acoustics, vol. 14, Academic Press*). (*e*) Deflection (drift) type (*after H. E. Dalke and W. Walkowitz, A new ultrasonic flowmeter for industry, ISA J., 7(10):60–63, 1960*). (*f*) Doppler type (*after H. M. Morris, Ultrasonic flowmeter uses wide beam technique to measure flow, Control Eng., 27(7): 99–101, 1980*). (*g*) Noise type (*after W. P. Mason and R. Thurston, eds., Physical Acoustics, vol. 14, Academic Press*)

computation incorporating the cross-section shape of the channel will give the flow rate.

Nuclear Magnetic Resonance Flowmeter

In this device (**Fig. 22**), nuclei of the flowing fluid are resonated by a radio-frequency field superimposed on an intense permanent magnetic field. A detector downstream measures the amount of decay of the resonance and thereby senses the velocity of the fluid. The most effective fluids are hydrocarbons, fluorocarbons, and water-bearing liquids because this resonance is most pronounced in fluorine and hydrogen nuclei. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

Laser Doppler Velocimeter

When light is scattered from a moving object, a stationary observer will see a change in the frequency of the scattered light (Doppler shift) proportional to the velocity of the object. In the laser Doppler velocimeter (**Fig. 23**) this Doppler shift measures the velocity of particles in a fluid. From the particle velocity the fluid's velocity is inferred. A laser is used as the light source because it is easily focused and is

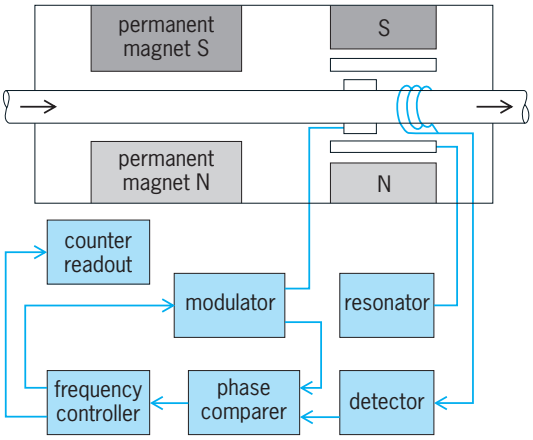

Fig. 22. Nuclear magnetic resonance type. (*After 1969 Guide to process instrument elements, Chem. Eng., 76(12):137–164, 1969*)

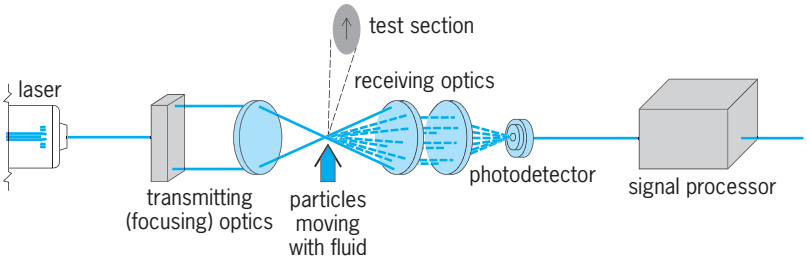

Fig. 23. Laser Doppler velocimeter. (*Thermo-Systems, Inc.*)

coherent and thereby enables the measurement of a frequency shift that is due to a Doppler effect. *See* LASER.

Accuracy of Measurement

The practical accuracy statement must be for the total flow-measurement system (combined primary and secondary devices). This has been customarily expressed as a percentage of the upper-range value ("full scale") in the case of rate meters used for industrial process control, and as a percentage of total flow rate for quantity meters. However, linear meters have often been specified in percentage of flow regardless of the use. Well-established standards for terminology, test procedures, and statements of operating environments have made it relatively easy to know what to expect for performance of secondary devices. Theory does not yet allow an accurate enough prediction of the actual flow of fluid inside of a pipe to predict the performance of a flowmeter. Square-edged orifices, venturis, and nozzles are the only types for which sufficient laboratory data have been collected and analyzed to enable a user to purchase and install a meter without individual calibration. One can expect a 1% maximum uncertainty in the primary if care is taken to follow the installation requirements of standards.

The readability or resolution of the readout display is important. This is poor at the low end of the square-root scale of differential-pressure meters. Rangeability of a flowmeter, sometimes termed turndown ratio in process applications, refers to the ratio of full-scale flow to a practical minimum flow rate. A figure of 3.5:1 is normally used for square-root scale flowmeters, while linear flowmeters, such as magnetic flowmeters, vortex-shedding flowmeters, and ultrasonic flowmeters normally have figures of 10:1 or higher. Flowmeters with rotating parts in the flowing stream, although linear, may have a smaller useful range, depending on the effect of viscosity and friction at low flow rates. Mead Bradner; Lewis P. Emerson

Bibliography. N. P. Cheremisinoff, *Flow Measurement for Engineers and Scientists*, 1987; D. M. Considine and G. K. McMillan (eds.), *Process/Industrial Instruments and Controls Handbook*, 5th ed., 1999; Daniel Industries Staff, *Fluid Flow Measurement*, 1993; J. P. DeCarlo, *Fundamentals of Flow Measurement*, 1984; C. J. Ing, *Handbook of Fluid Flowmeters*, 1988; B. G. Liptak, *Flow Measurement*, 1993; R. W. Miller, *Flow Measurement Engineering Handbook*, 3d ed., 1996; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; D. W. Spitzer, *Industrial Flow Measurement*, 2d ed., 1990.

Flow of granular solids

The gravity flow of particulate solids such as ceramic powders, plastic pellets, and ores in the form of separate particles, with the particles in contact and the voids between them filled with gas, usually air.

Particulate solids are a two-phase, solid–gas system. They are compressible; their bulk density changes during flow. Since the volume of the particles changes little during this process, it is the size of the voids that is mostly affected. Changes in the size of the voids can cause changes in gas pressure and result in gas-pressure gradients across a flowing solid which tend to reduce the rate of gravity discharge. When the solid is made up of large particles, that is, its permeability is high, when the required flow rates are low, the gas-pressure gradients are not significant; the gaseous phase can be ignored and the solid treated as a one-phase, solid-only system. The theory of flow of such a system dates to the early 1960s and has since found general acceptance. Its application to storage vessels will be outlined. This theory has been extended to include effects of the gaseous phase; this will be discussed briefly.

Bins. A bin (silo, bunker) generally consists of a vertical cylinder and a converging hopper. The size of a bin can vary from that of a small press-feed hopper with a capacity of a few quarts (liters), through portable bins, to large stationary silos storing several thousand tons of material. From the standpoint of flow, there are three types of bins: mass-flow, funnel-flow, and expanded-flow.

Mass-flow bins. Mass flow occurs when the hopper walls are sufficiently steep and smooth to cause flow of all the solid, without stagnant regions, whenever any solid is withdrawn. Valleys are not permitted, nor are ledges or protrusions into the hopper. In addition, the outlet must be fully effective, that is, if

the hopper is equipped with a shut-off gate, the gate must be fully open; if it is equipped with a feeder, the feeder must draw material across the full outlet area. The range of hopper slope and friction angles leading to mass flow is shown in **Fig. 1**. Funnel flow will occur unless both the conditions for mass flow shown in this figure (the condition on θ_c in Fig. 1*b* and the condition on θ_p in Fig. 1*c*) are satisfied. The θ_c values in Fig. 1*b* also apply to conical hoppers, and the θ_p values in Fig. 1*c* also apply to wedge-shaped and chisel hoppers, provided the outlet length-to-width ratio is at least 3:1.

Mass-flow bins have advantages. Flow is uniform, and feed density is practically independent of the head of solid in the bin; this frequently permits the use of volumetric feeders for feed rate control. Indicators of low material level work reliably. Segregation is minimized because, while a solid may segregate at the point of charge into the bin, continuity of flow enforces remixing of the fractions within the hopper. Special bin design can be used for in-bin blending by circulation of the stored solid.

Mass-flow bins have a first-in-first-out flow sequence, thus ensuring uniform residence time and deaeration of the stored solid. Airlocks can often be dispensed with, provided that a critical flow rate is not exceeded; otherwise material flow will be erratic and, at high rates, material will remain fluidized and flush (pour out like a liquid) on exiting the bin.

The mass-flow type of bin is recommended for cohesive materials, for materials which degrade with time, for powders, and when segregation needs to be minimized.

Funnel-flow bins. Funnel flow occurs when the hopper walls are not sufficiently steep and smooth to force material to slide along the walls or when the outlet of a mass-flow bin is not fully effective.

In a funnel-flow bin, solid flows toward the outlet through a channel that forms within stagnant material. The diameter of that channel approximates the largest dimension of the effective outlet. When the outlet is fully effective, this dimension is the diameter of a circular outlet, the diagonal of a square outlet, or the length of an oblong outlet. As the level of solid within the channel drops, layers slough off the top of the stagnant mass and fall into the channel. This erratic behavior is detrimental with cohesive solids since the falling material packs on impact, thereby increasing the chance of material developing a stable arch across the hopper so that a complete stoppage of flow results. A channel, especially a narrow, high-velocity channel, may empty out completely, forming what is known as a rathole, and powder charged into the bin then flushes through. Powders flowing at a high rate in a funnel-flow bin may remain fluidized because of the short residence time in the bin, and flush on exiting the bin. A rotary valve is often used under these conditions to contain the material, but a uniform flow rate cannot be ensured because of the erratic flow to the valve.

Funnel-flow bins are more prone to cause arching of cohesive solids than mass-flow bins, and they therefore require larger outlets for dependable flow.

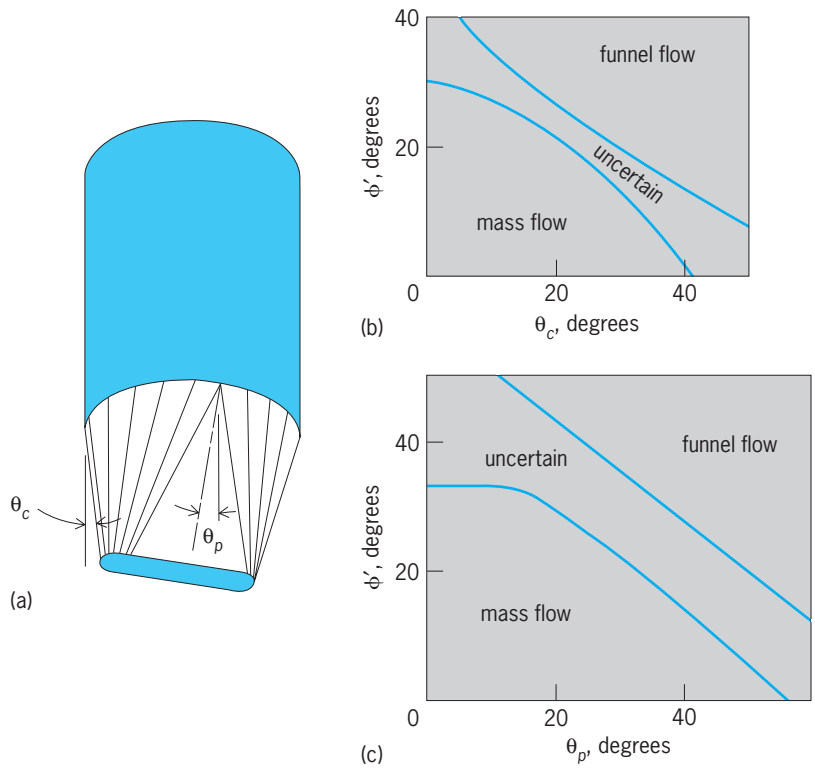

Fig. 1. Bounds on mass flow and funnel flow. (*a*) Geometry of a transition hopper, showing slope angles θ_c and θ_p. (*b*) Range of kinematic angle of friction (ϕ') and θ_c leading to mass flow and funnel flow. (*c*) Range of ϕ' and θ_p leading to mass flow and funnel flow.

These bins also cause segregation of solids (**Fig. 2**) and are unsuitable for solids which degrade with time in the stagnant regions. Cleanout of a funnel-flow bin is often uncertain because solids in the stagnant regions may pack and cake.

The funnel-flow type of bin is suitable for coarse, free-flowing, or slightly cohesive, nondegrading solids when segregation is unimportant.

Expanded-flow bins. These are formed by attaching a mass-flow hopper to the bottom of a funnel-flow bin. The outlet usually requires a smaller feeder than would be the case for a funnel-flow bin. The mass-flow hopper should expand the flow channel to a dimension sufficient to prevent ratholing.

These bins are recommended for the storage of large quantities [10,000 ft^3 (300 m^3) or more] of nondegrading solids. This design is also useful as a modification of existing funnel-flow bins to correct erratic flow caused by arching, ratholing, or flushing. The concept can be used with single or multiple outlets.

Feeders. The hopper outlet must be fully effective. If flow is controlled by a feeder, the feeder must be designed to draw uniformly through the entire cross section of the outlet. It is also essential that the feeder be either suspended from the bin itself or supported on a flexible frame so as to readily deflect with the bin.

Vibrators are suitable for materials which, free-flowing under conditions of continuous flow, cake when stored at rest, but break up into separate particles when vibrated—for example, granular sugar. Fine powders and wet materials tend to pack when

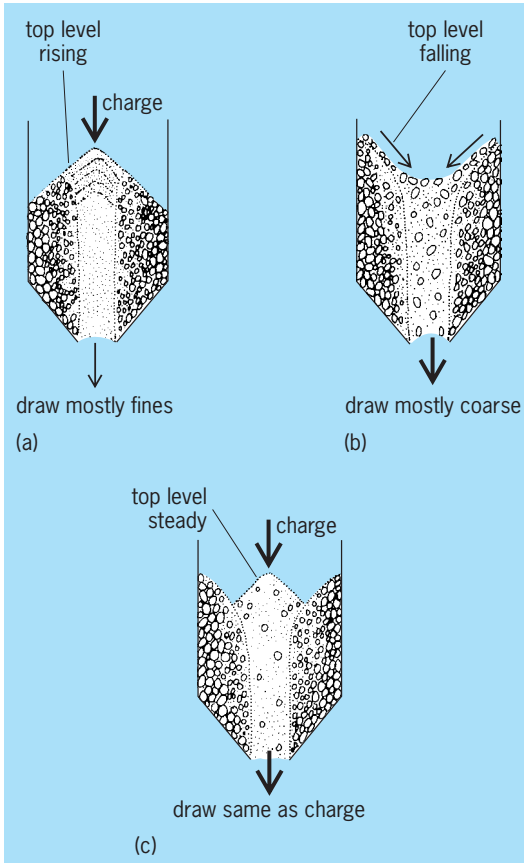

Fig. 2. Flow patterns in a funnel-flow bin, showing segregation of solids. (*a*) Charge coming in is greater than draw coming out, resulting in draw with a preponderance of fine particles. (*b*) Draw is greater than charge, resulting in draw with a preponderance of coarse particles. (*c*) Draw and charge are of same magnitude, resulting in draw with same consistency as charge. (*After A. W. Jenike, Why bins don't flow, Mech. Eng., May 1964*)

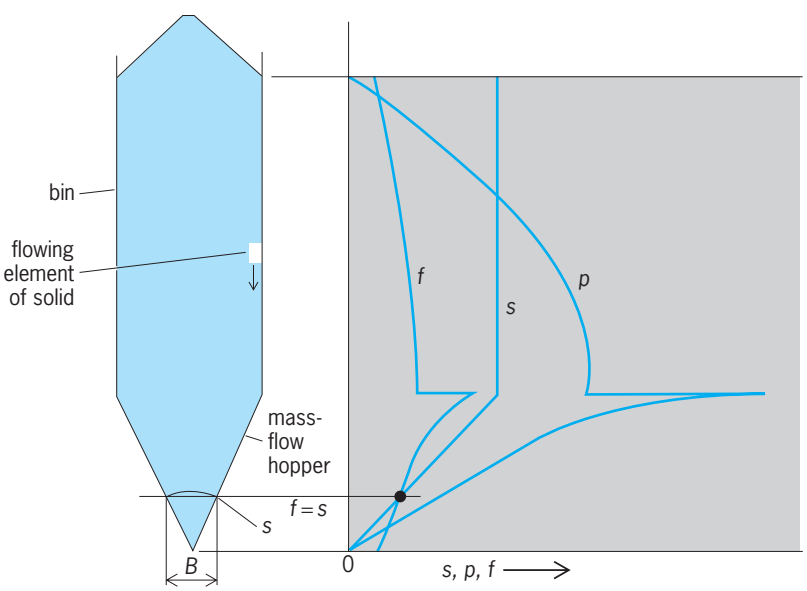

Fig. 3. Section of bin with mass-flow conical hopper, together with graph showing relative values of pressure *p*, compressive strength *f*, and stress *s*, and illustrating criterion ($f = s$) for minimum outlet dimension *B*. (*After A. W. Jenike, Quantitative design of mass-flow bins, Powder Technol., 1:237–244, 1967*)

vibrated; hence vibration is not recommended. A mass-flow bin with a screw, belt, or similar feeder is best suited to those materials.

Design for flow. The outlet dimensions of a hopper outlet must be sufficient to assure unobstructed flow. Flow may be obstructed by interlocking of large particles and by cohesive doming and ratholing across or above the outlet. Prevention of interlocking requires that the outlet be larger than several particle sizes. Prevention of cohesive obstructions requires that the outlet be sufficiently large to ensure the failure of potential cohesive obstructions.

Under the head of solid in a bin, pressure p develops within the solid. The solid consolidates and gains compressive strength f. This strength permits the development of stable domes and of stable ratholes in funnel-flow bins.

Consider a bin with a mass-flow conical hopper (**Fig. 3**). The solid is permitted to flow into it. The solid unconsolidated when deposited at the top, but, as an element of the solid flows down, it becomes consolidated under the pressure p acting on it in the bin. This pressure is shown by the p line. At first, p increases with depth, then levels off in the vertical part of the channel. At the transition to the hopper, the pressure increases abruptly and then decreases, at first exponentially and then approximately linearly toward zero at the imaginary vertex of the hopper. To each value of consolidating pressure p there corresponds a strength f of the solid, so that the strength f, generated by pressure p, similarly increases and decreases as indicated by the f line. Design for flow is based on the flow–no-flow criterion postulated as follows: a solid will flow provided the strength f which the solid develops is less than the stress s which would act in a stable obstruction to flow, $f < s$.

In a mass-flow bin, a dome across the hopper (Fig. 3) is the potential obstruction. The stress s in a dome is proportional to the span of the bin, as shown by line s. It is now observed that lines s and f intersect. Above the point of intersection, $f < s$, the flow criterion is satisfied and the solid will flow. Below that point, $f > s$, the solid has enough strength to support a dome and will not flow. The point $f = s$ determines the critical dimension B of the hopper which has to be exceeded for dependable flow.

Flowability of hoppers. As shown in Fig. 3, both the stress s in a stable dome and the consolidating pressure p are linear functions of the width of the hopper B. Since s and p are zero at the vertex of the hopper, their ratio is constant for a given hopper. This ratio is referred to as the flow factor, $ff = p/s$. It measures the flowability of a hopper. Values of ff have been computed for a wide range of hoppers.

Flowability of solids. The relation between the consolidating pressure p and the generated compressive strength f is called the flow function FF of a solid and is a measure of its flowability. The flowability of solids depends on several parameters.

Surface moisture. Solids are least free-flowing when moisture content is in the range of 70–90% of saturation. A saturated solid will usually drain in storage to

the range of minimum flowability. Flowability tests may be carried out for a range of moisture content to determine the maximum moisture at which a solid can be handled by gravity.

Temperature. Many solids, especially those with a low softening point, are affected by temperature. The critical conditions often occur when a solid is placed in storage at an elevated temperature and allowed to cool while at rest. Some solids gain maximum strength when held at a constant elevated temperature. Flowability tests indicate the maximum temperature at which a solid can be stored and whether or not it can be cooled without excessive caking.

Time of storage at rest. Many solids are free-flowing if they are kept in motion, but cake severely if stored at rest for a period of time. Flowability tests predict the maximum time that a solid can remain in storage at rest. The most widely accepted method of measuring a solid's flowability utilizes the Jenike shear tester. Using small, representative samples (usually 5 gallons or 19 liters or less), measurements are made of the solid's cohesive strength, friction, compressibility, permeability, and other relevant properties. These tests must be run under conditions that accurately simulate how the solid is handled in an industrial setting.

Effect of gaseous phase. Incidental effects of the gaseous phase are flushing of powders and limitation of discharge flow rate. In the latter, a speedup of the bin outlet feeder produces no increase in feed rate, which is limited by the hopper outflow rate. Predicting the flow effects caused by two-phase (solid–gas) interactions is difficult. Empirical approaches are often fruitless because of the many variables involved. A combined solid–gas theoretical approach can provide the needed answers.

As briefly described above, continuum mechanics has been successfully used to describe a one-phase, solid-only system. This approach can be extended to include gaseous effects. The basic equations involve equilibrium of forces, a fluid-flow rule in addition to a solids-flow rule, and continuity of the gas and solids. By incorporating boundary conditions, it is possible to predict when flow rate limitations as well as flushing will occur. Flow rate limitations can often be overcome by enlarging the bin outlet or by introducing a small amount of gas into the bin to promote flow. In addition to applying these techniques to bin flow problems, other applications involve fixed- and moving-bed reactors, lockhoppers, and standpipes.

Inserts. Various types of devices can be placed inside the hopper section of a bin to correct flow problems. These consist of inverted cones and pyramids as well as a hopper-within-the-hopper. Some of the benefits which may accrue with such devices include converting a funnel-flow bin to mass flow, achieving completely nonsegregated discharge from a mass-flow bin, or causing the bin to act as a blender as well as a storage vessel. It is important that such devices be carefully designed and fabricated both to ensure that the desired flow pattern will develop and to maintain the structural integrity of the insert and the bin.

Segregation. Particles in a mixture can segregate by one of five primary mechanisms: sifting, particle velocity on a surface, air entrainment (fluidization), entrainment of particles in an airstream, and dynamic effects. Of these, the most common mechanism for segregation is sifting of fines through a matrix of coarse particles. For this to occur, there must be at least a minimum ratio (1.5:1) of particle sizes and a minimum mean particle size (100 micrometers), the bulk solid must be free-flowing, and there must be relative motion between particles.

In order to solve a segregation problem, it is necessary to first identify the mechanism by which the particles are segregating. Then one or more of the following techniques can be considered to reduce or eliminate segregation: changing the characteristics of the material which are causing it to segregate, changing the process sequence by which the material is being handled, and changing the design of the equipment being used to handle the material.

Abrasive wear. Abrasive wear of equipment is often an important problem when storing and handling bulk granular solids. A wear tester has been developed to measure the wear rate due to solids sliding on a wearing surface as they do during flow in a bin. A dimensionless wear ratio is measured as a function of the solids normal stress applied to the wearing surface for a particular wear material and for particular bulk solid of a certain particle size distribution, moisture content, and temperature. By incorporating the principles of bulk solids flow, it is possible to predict in absolute quantitative terms the wear which will occur in any location within a bin, hopper, feeder, chute, or other bulk-solids-handling equipment. In the design or modification of such equipment, this serves as an invaluable aid to the engineer when faced with questions of placement, quantity, type, and service life of wear-resistant materials. *See* BULK-HANDLING MACHINES; PARTICULATES; WEAR.

Andrew W. Jenike; John W. Carson

Bibliography. J. W. Carson and J. Marinelli, Characterize bulk solids to ensure smooth flow, *Chem. Eng.*, pp. 78–89, April 1994; J. Marinelli and J. W. Carson, Solve solids flow problem in bins, hoppers and feeders, *Chem. Eng. Prog.*, pp. 22–27, May 1992; J. Marinelli and J. W. Carson, Use screw feeders effectively, *Chem. Eng. Prog.*, pp. 47–51, December 1992.

Flower

A higher plant's sexual apparatus in the aggregate, including the parts that produce sex cells and closely associated attractive and protective parts (**Fig. 1**). "Flower" has been used broadly enough by some botanists to include the reproductive equipment of cycads, conifers, and other plants with naked seeds. Here, however, it will be limited, as is usual, to the angiosperms, plants with enclosed seeds and the unique reproductive process called double fertilization. In its most familiar form a flower is made up

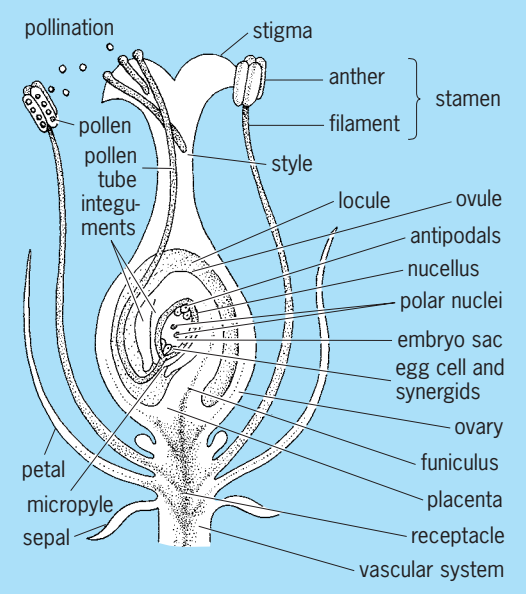

Fig. 1. Flower structure, median longitudinal section.

of four kinds of units arranged concentrically. The green sepals (collectively termed the calyx) are outermost, showy petals (the corolla) next, then the pollen-bearing units (stamens, androecium), and finally the centrally placed seedbearing units (carpels, gynoecium). This is the "complete" flower of early botanists, but it is only one of an almost overwhelming array of floral forms. One or more kinds of units may be lacking or hard to recognize depending on the species, and evolutionary modification has been so great in some groups of angiosperms that a flower cluster (inflorescence) can look like a single flower.

Floral Diversity

From the time of Carl Linnaeus, the classification of angiosperms has been based on floral features, but the diversity that makes this possible has also made it hard for botanists to develop widely applicable concepts and unequivocal terms. The old term pistil, for example, denoting the pestlelike shape of many gynoecia, could be applied to a unit or to a compound structure. It has gradually given way to carpel, which always refers to a unit. Seeking simplicity and consistency, pre-Darwinian botanists commonly reduced all the structural complexities of higher plants to two or three elements: shoot and root; axis and appendages; or root, stem, and leaf. A flower was viewed as a shortened axis bearing modified leaves. This was the "primitive" flower, the ground plan, or the archetype from which all sorts of flowers could be derived conceptually by suppressing parts (abortion), by uniting parts of the same kind (connation), or by uniting unlike parts (adnation). Interpreting the gynoecium in all its diversity was especially troublesome, and much effort went into two approaches: analyzing flower development (ontogeny) and using the position of vascular bundles to delimit the floral parts.

The terms and concepts of this era were not intended to reflect ancestor-descendant relationships. Some workers believed, rather, that regularities in plant structure would parallel regularities in crystal structure, and crystals in all their diversity had been found to be geometric modifications of a few primitive forms. Others thought the "true nature" of the flower—a notion traceable to the *physis* of pre-Socratic philosophy—could be discovered by study and contemplation. All this was brought into the evolutionary era, common ancestry replacing archetype or ground plan and "true nature" tagging along, with mixed results. Many of the floral modifications identified by pre-Darwinians were, and still are, readily seen as changes brought about by natural selection. It is now widely realized, however, that it serves no purpose to look for axes and appendages in highly altered reproductive structures and that the nature of the flower belongs to metaphysics, not science.

Though concepts have changed, the old expressions abortive, adnate, and connate survive as descriptive terms. A floral part can be called abortive if its development, when compared with that of corresponding parts in the same species or a related species, is stunted: *Ranunculus abortinus,* the abortive or kidney-leaf buttercup, has diminutive petals and carpels. Like parts of a flower are connate if they merge when followed up or down: the stamens of an orange blossom (**Fig. 2**) are basally connate, and the petals of a petunia are connate in a tubular corolla. Unlike parts are adnate if they merge: the stamens of mints, fox gloves, petunias, and many other plants are adnate to a corolla tube.

Concentric units. Most botanical terms have never been anything but descriptive, and a botanist must have a large store of them to impart the multiformity of flowers. The examples that follow are only a smattering. An extra series of appendages alternating with the sepals, as in purple loosestrife, is an epicalyx. A petal with a broad distal region and a narrow proximal region is said to have a blade and claw: the crape myrtle has such petals. The term perianth, which embraces calyx and corolla and avoids the need to distinguish between them, is especially useful for a flower like the tulip, where the perianth parts are in two series but are alike in size,

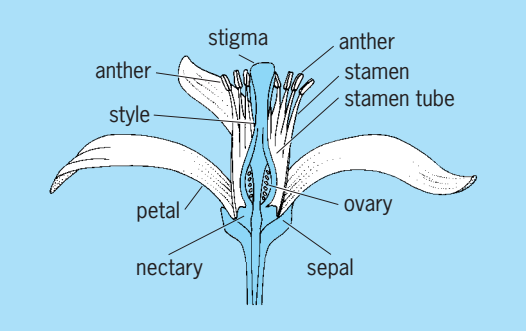

Fig. 2. Diagrammatic section of an orange blossom. (*After J. B. Free, Insect Pollination of Crops, Academic Press, 1970*)

shape, and color. The members of such an undifferentiated perianth are tepals. When the perianth has only one series of parts, however, they are customarily called sepals even if they are petallike, as in the windflower.

A stamen commonly consists of a slender filament topped by a four-lobed anther, each lobe housing a pollen sac. In some plants one or more of the androecial parts are sterile rudiments called staminodes: a foxglove flower has four fertile stamens and a staminode (**Fig. 3**). Carpellode is the corresponding term for an imperfectly formed gynoecial unit. Most stamens release their pollen by longitudinal splitting of the anther along the anther's inner face, along its outer face, or on its sides, depending on the species (introrse, extrorse, or lateral dehiscence). In others—the stamens of rhododendrons, azaleas, or potato flowers, for instance—the pollen exits from a pair of terminal pores (poricidal dehiscence).

A gynoecium is apocarpous if the carpels are separate (magnolia, blackberry; **Fig. 4**) and syncarpous if they are connate (tulip, poppy). Or the gynoecium may regularly consist of only one carpel (bean, cherry). A solitary carpel or a syncarpous gynoecium can often be divided into three regions: a terminal, pollen-receptive stigma; a swollen basal ovary enclosing the undeveloped seeds (ovules); and a constricted, elongate style between the two. The gynoecium can be apocarpous above and syncarpous below; that is, there can be separate styles and stigmas on one ovary (wood sorrel).

Every flower cited so far has a superior ovary: perianth and androecium diverge beneath it (hypogyny). If perianth and androecium diverge from the ovary's summit, the ovary is inferior and the flower is epigynous (apple, banana, pumpkin). A flower is perigynous if the ovary is superior within a cup and the other floral parts diverge from the cup's rim (cherry). A syncarpous ovary is unilocular if it has only one seed chamber, plurilocular if septa divide it into more than one. The ovules of a plurilocular ovary are usually attached to the angles where the septa meet; this is axile placentation, a placenta being a region of ovular attachment. There are other ways in which the ovules can be attached—apically, basally, parietally, or on a free-standing central placenta—each characteristic of certain plant groups (**Fig. 5**).

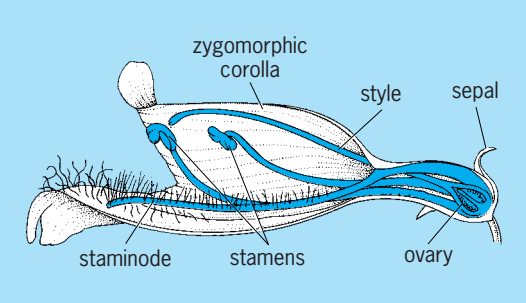

Fig. 3. Flower of *Penstemon*, a member of the foxglove family, in section. (*After C. E. Wood, Jr., A Student's Atlas of Flowering Plants: Some Dicotyledons of Eastern North America, Harper and Row, 1974*)

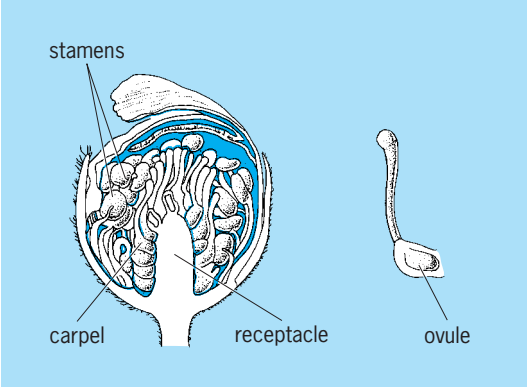

Fig. 4. Longitudinal section of a flower bud and one carpel in section of blackberry (*Rubus betulifolius*). (*After K. R. Robertson, The genera of Rosaceae in the southeastern United States, J. Arnold Arbor., 55:344–401, 1974*)

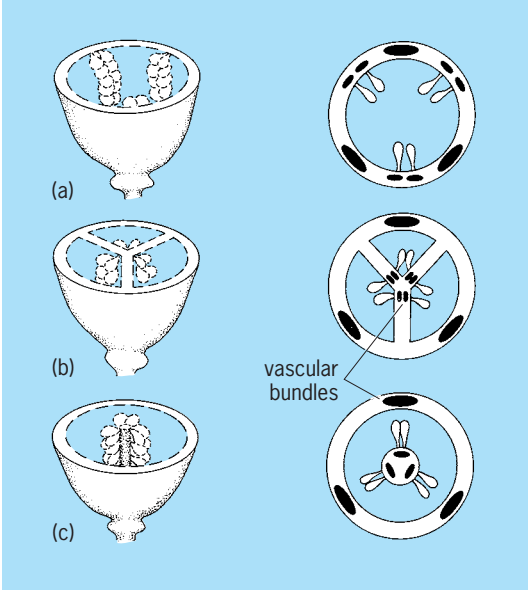

Fig. 5. Placentation: (*a*) parietal, (*b*) axile, and (*c*) free central. (*After P. H. Raven, R. F. Evert, and H. Curtis, Biology of Plants, Worth Publishers, 1976*)

The term bract can be applied to any leaflike part associated with one or more flowers but not part of a flower. Floral bracts are frequently small, even scalelike, but the flowering dogwood has four big petallike bracts below each flower cluster. The broad end of a flower stalk where the floral parts are attached is the receptacle. The same term is used, rather inconsistently, for the broad base bearing the many individual flowers (florets) that make up a composite flower like a dandelion or a sunflower.

Symmetry. A flower is actinomorphic if it is radially symmetrical, zygomorphic if bilaterally symmetrical. It is isomerous if each series has the same number of parts and it is 3-, 4-, or *n*-merous according to the number of parts in a series. Thus, an evening primrose, with four sepals, four petals, two whorls of four stamens each, and four carpels, is isomerous and 4-merous. It is also diplostemonous because its stamens are in two series; with one series it would be

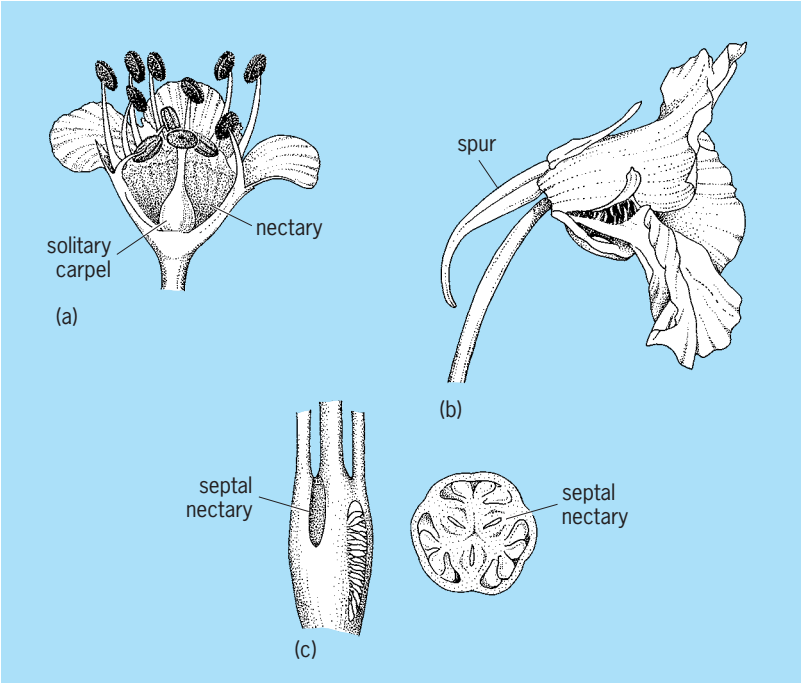

Fig. 6. Nectaries. (*a*) Cherry flower (*after W. H. Brown, The bearing of nectaries on the phylogeny of flowering plants, Proc. Amer. Phil. Soc., 79:549–595, 1938*). (*b*) Flower of nasturtium (*after B. J. D. Meeuse, The Story of Pollination, Ronald Press, 1961*). (*c*) Ovary of gladiolus (longitudinal and cross section) showing septal nectaries (*after W. H. Brown, 1938*).

haplostemonous. Flowers of dicotyledons are commonly 5-merous (carpels usually fewer), less often 3-merous, 4-merous, or 5+-merous; whereas monocot flowers are commonly 3-merous and rarely 4-merous (carpels often fewer), but never 5-merous.

Sexuality. A plant species is diclinous if its stamens and carpels are in separate flowers. A diclinous species is monoecious if each plant bears staminate and carpellate (pistillate) flowers, dioecious if the staminate and carpellate flowers are on different plants. The corn plant, with staminate inflorescences (tassels) on top and carpellate inflorescences (ears) along the stalk, is monoecious. Hemp is a well-known dioecious plant.

Nectaries. Flowers pollinated by insects or other animals commonly have one or more nectaries, regions of densely cytoplasmic tissue that secrete a sugar solution. A nectary can be nothing more than a layer of tissue lining part of a floral tube or cup (cherry; **Fig. 6***a*), or it can be as conspicuous as the secretory spur of a nasturtium (Fig. 6*b*) or a larkspur. It can be a cushionlike outgrowth at the base of a superior ovary (orange blossom) or atop an inferior ovary (parsley family). Gladiolus and a number of other monocotyledons have septal nectaries, deep secretory crevices where the carpels come together (Fig. 6*c*). Substances that give off floral odors—essential oils for the most part—ordinarily originate close to the nectar-producing region but are not coincident with it. Production by the epidermis of perianth parts is most common, but in some species the odor emanates from a more restricted region and may even come from a special flap or brush. Most insect-pollinated plants have visual cues, some of them outside the human spectral range, as well as odor to bring the pollinators to the flowers and guide them to the nectar. *See* SECRETORY STRUCTURES (PLANT).

Inflorescence. Inflorescence structure, the way the flowers are clustered or arranged on a flowering branch, is almost as diverse as floral structure. To appreciate this, one need only contrast the drooping inflorescences (catkins) of a birch tree with the coiled flowers of a forget-me-not or with the solitary flower of a tulip. In some cases one kind of inflorescence characterizes a whole plant family. Queen Anne's lace and other members of the parsley family (Umbelliferae) have umbrellalike inflorescences with the flower stalks radiating from almost the same point in a cluster. The stalkless flowers (florets) of the grass family are grouped into clusters called spikelets, and these in turn are variously arranged in different grasses. Flowers of the arum family (calla lily, jack-in-the-pulpit), also stalkless, are crowded on a thick, fleshy, elongate axis. In the composite family, florets are joined in a tight head at the end of the axis; the heads of some composites contain two kinds (**Fig. 7**), centrally placed florets with small tubular corollas and peripheral ray florets with showy, strap-shaped corollas (the "petals" one plucks from a daisy). Inflorescences differ so much with regard to patterns of branching, the order in which flowers appear, and the distribution of bracts, that a system for describing and comparing all of them is almost beyond reach. There have been a few attempts at

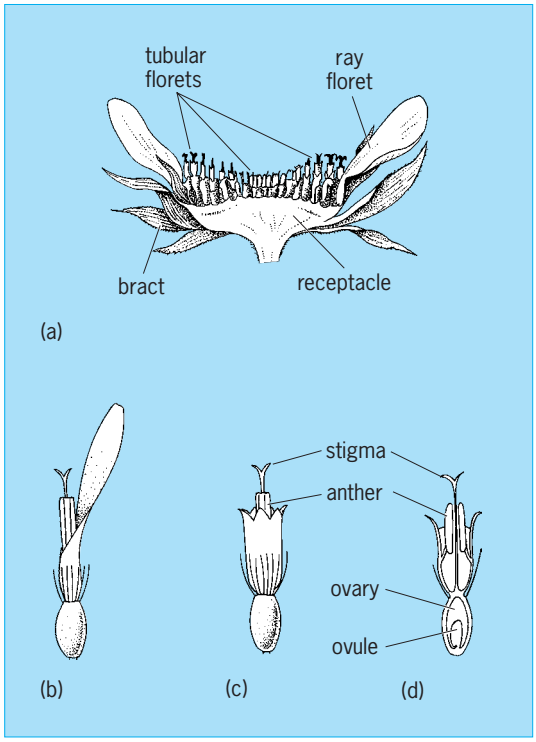

Fig. 7. Sunflower (*Helianthus annuus*) of the composite family. (*a*) Inflorescence in section. (*b*) Side view of ray floret. (*c*) Side view of tubular floret. (*d*) Section of tubular floret. (*After J. B. Free, Insect Pollination of Crops, Academic Press, 1970*)

computer modeling, with some success. *See* INFLORESCENCE.

Evolution

Inferring how flowers began is another of botany's troublesome tasks. Ancient impressions of flowers are rare and usually short on structural detail. Though flowers embedded in amber can be exquisitely preserved, none is old enough to illuminate the origin of angiosperms. Remains of dispersed pollen grains, on the other hand, are plentiful enough and old enough, and with the help of the electron microscope angiosperm pollen can be distinguished from gymnosperm pollen well enough to track the entry of angiosperms into the record. Investigations of such fossils place the beginning of flowering plants in the Early Cretaceous and support the orthodox view that they came from a common ancestral group.

Traits of early flowers. Though the ancestors of angiosperms still cannot be identified among fossil gymnosperms, the record supports the opinion that early flowers had traits seen today, in various combinations, in such living members of the order Magnoliales as the star anise (*Illicium*), the Fijian tree *Degeneria,* and *Drimys,* a genus of the Old and New World tropics. The early flowers were bisexual, with floral parts indefinite in number, free from each other, and probably arranged in low spirals. The perianth was not differentiated into calyx and corolla. The carpel may have been folded in half lengthwise along its midvein, with a stigmatic crest running along its unsealed margins. *Drimys piperita,* a magnolialean shrub of the western Pacific, has such carpels, and they are duplicated to some extent in a fossil from the Cretaceous of central Kansas.

It has been recognized since the early 1800s that all flowers can be derived conceptually from a magnolialean archetype. Twentieth-century anatomists gave the concept an evolutionary basis by pointing out, among other things, that the stamens and carpels of some geographically remote Magnoliales can be arranged as plausible vestiges of early evolutionary change. Pollen studies support the anatomical findings, demonstrating that pollen grains with single furrows, a trait found in archaic Magnoliales, are the oldest in the angiosperm record, and other forms enter the record in younger horizons as clear-cut indications of evolutionary diversification. *See* PALYNOLOGY.

Diversification and coevolution. The fossil record shows that diversification of angiosperms was rapid. It took perhaps 25,000,000–30,000,000 years for flowering plants to appear, diversify, and gain dominance over the gymnosperms, about as much time as it later took for mammals to radiate and replace reptiles and the dominant animals. It is thought that the earliest mammals, along with birds and insects, brought the flowering plants to dominance. The first angiosperms were plants of limited distribution, low abundance, and low diversity, with seeds dispersed rather inefficiently by wind, water, and reptiles. The emergence of fruit-eating and seed-eating birds and mammals enabled undigested seeds to get farther from the parent plant and its predators before germinating, thereby spreading out the species. This in turn gave insect pollination an advantage over wind pollination, which is the usual method for gymnosperms, because individuals of the same species were now often too far apart for wind-mediated outbreeding. Thus, flowering plants began to coevolve with the insects that had heretofore visited haphazardly to feed on pollen or on the flower.

Pollination. Beetles, an old group, are probably the original pollinating animals, and some of the Magnoliales have floral adaptations that perpetuate this form of breeding. In many other plants, however, pollination by beetles has given way to pollination by bees, wasps, butterflies, moths, birds, bats, or other animals. Regardless of the pollinator, the flower ordinarily offers nectar as a reward, but the way in which the reward is advertised and protected differs, depending on the plant's major pollinator. Thus, bee flowers can be bright orange, yellow, purple, or blue; but they are rarely red, a color that bees cannot see. Bird flowers, on the other hand, are often bright red, eliminating bees as competitors for the nectar. A bee flower's corolla may form a gullet above the nectary in such a way that a bee can enter but its competitors cannot; the stamens may be placed so as to release pollen onto the bee when it squeezes into the gullet; and the gullet may be so formed as to foil honeybees while admitting bumblebees (foxglove). Bee flowers that offer only pollen are constructed differently from those that also offer nectar.

Butterfly flowers and moth flowers are fragrant like most bee flowers, but they commonly have long, narrow, bee-proof tubes above the nectary that are well suited to the long, coiled lepidopterous tongue. Flowers pollinated by butterflies are brightly colored, open in daylight, and often end in a broad landing platform (petunia). Flowers pollinated by nocturnal moths are pale, and in keeping with the hovering habit of such insects they usually lack a landing platform.

Bird flowers commonly have tubular corollas that are larger in diameter than the corollas of moth flowers and butterfly flowers. They are rarely scented, since birds have little or no sense of smell. The ovary is often inferior; the nectary above it can then be probed by a bird's bill without damage to the ovules. A typical bat-pollinated flower is a large, pallid night-bloomer with many anthers and a strong odor, frequently described as musty or rancid. Bird flowers and bat flowers are all heavy nectar producers, for the high energy needs of the animals must be met to sustain the relationship.

Some flowers have coevolved with flies that eat pollen and nectar. Some of them use shiny surfaces or dangling parts to bring in the pollinators. Other flowers (or flowerlike inflorescences) attract eaters of carrion or dung by producing an odor resembling the pollinator's favorite food; they may be constructed so as to trap flies or carrion beetles for a time, thereby increasing the chance that the insect will take on some pollen before departing.

This brief overview hardly begins to convey the array of modifications that evolution has wrought in the form and physiology of flowers. In a number of lin eages the flowers have clustered so that the inflorescence functions as one flower in attracting or powdering the pollinator. Dogwoods and the huge family Compositae are examples. In some tropical tree species, flower clusters do not appear at the ends of branches but burst through the bark of the branches or trunk, an adaptation to pollination or dispersal by bats. Insect pollination has been replaced by wind pollination in some groups, including the grasses, the sedges, certain composites (ragweed), and such catkin-bearing trees as the birches and walnuts. Wind-pollinated flowers often lack color, smell, and perianth parts. They are heavy producers of pollen, and the grains are usually smooth and dry in contrast to the sticky, sculptured grains of most insect-pollinated plants, though wind-pollinated stigmas are likely to be sticky or feathery. Self-pollination, with or without the help of a pollinator, is widespread among angiosperms. There are species of violets with two kinds of flowers on each plant: the insect-pollinated ones, picked for nosegays, and smaller self-pollinated ones that never open. Though some plants have mechanisms to keep seeds from developing in a self-pollinated flower, others are constructed so that successful selfing will take place if outcrossing does not. Indeed, alternative routes to reproductive success are commonplace, and the link between a flower and a pollen vector is seldom so tight that it cannot alter with a change in selection pressure. *See* POLLINATION.

Outcrossing. Outcrossing is basic to the normal evolutionary process, so it is understandable that many lineages have evolved ways to favor or enforce it. Stamens and carpels can be kept apart in unisexual flowers (dicliny) by structural barriers within a flower or by maturing at different times (dichogamy); the barrier to selfing may also be physiological, as when pollen will not germinate except on the stigma of a genetically different individual within the species.

Formation

A plant is ready to flower when it has met complex requirements that vary from one species to another, even from one geographic race to another. There must be a prior period of growth, which is short for herbs of the desert or of alpine meadows but may be decades for a tree or a century or more for a bamboo. Some long-lived plants grow for years, flower once, and die after the fruits ripen, much as an annual plant does. Seasonal requirements are known to all gardeners; a plant may need a period of exposure to low temperature, a particular day-length, or both before it will bloom. Cabbages and some other biennial plants will not bolt—send up a flowering shoot—without first experiencing a cool period. In monsoon regions, the flowering of certain trees is triggered by the onset of the dry season. Tulips and daffodils do not need a cold period for flower formation, but the initiation of flowers in the summer and their emergence from the bulb the following spring are linked to seasonal warming. Depending on the species, plants of the wet tropics can flower continuously or in response to an internal "clock" or in response to unusual weather. Though flowering intervals vary greatly in a tropical rainforest, widespread individuals of a species frequently flower at the same time; this enables the plants to cross-pollinate and ensures that seeds will be plentiful enough for some to escape their predators.

To initiate flowers instead of leaves, a growing tip changes from an indeterminate meristem to a determinate meristem, an aggregation of dividing cells that will go through a stepwise program and then stop dividing. One of the first reactions to the flowering stimulus is an increase in ribonucleic acid (RNA) synthesis. The number of ribosomes and the rate of cell division rise, along with the relative amount of dispersed nuclear chromatin, an indication that genetic information is being transcribed. Eventually, deoxyribonucleic acid (DNA) also increases. With still more cell divisions the apex changes shape, usually becoming broader and shallower than the vegetative apex. Cells deep beneath the surface acquire bigger vacuoles, and the apex then consists of a meristematic mantle over a vacuolated core. At this stage an apex programmed to become an inflorescence looks much the same as one that will be a flower. Each floral part begins with periclinal divisions (new walls parallel to the surface), usually in the second or third cell layer. As anticlinal divisions are added, a bump forms on the apex, or several bumps for all the members of a set. The usual formative sequence is acropetal—sepals, petals, stamens, carpels—but in purple loosestrife (to cite just one exception) the inner stamens and the petals start after the gynoecium is initiated. Stamen primordia become four-lobed shortly after their initial protrusion. Carpels typically begin as protuberances, then become horseshoe-shaped before taking on the form they have when they are ready for pollination. Some syncarpous gynoecia, however, begin as circulary or elliptic ridges in which primordial carpels cannot be delimited (**Fig. 8**). Continuity of parts—connation and adnation—can come about in several ways, including interprimordial growth, extension and fusion of meristems, and initiation of a new meristem in the common base below two or more primordia. Or the surfaces of adjoining parts can unite as they grow: the tips of the paired carpels of a garden periwinkle fuse by modifying the walls of juxtaposed epidermal cells. A similar surface fusion of the margins lower down in the carpel occurs in many plants. *See* APICAL MERISTEM.

Surgical experiments show that the cells of a floral apex become increasingly committed to their final form as development progresses. Bisected early enough, an apex forms two flowers. That is, sepals and other parts form around each half. When the apex is bisected after a critical stage, however, sepals form only in the regions already committed to sepal formation, and there is another critical stage for each of the other sets of floral parts. Hormones are thought to play a major role in sequential

differentiation because floral development can be changed greatly by experimental application of hormones. Hormones can also alter sex expression in diclinous plants: in the cucumber, potentially staminate flowers can be made to develop as carpellate flowers. *See* PLANT HORMONES; VERNALIZATION.

Anatomy

Some of the general anatomical features of leaves can be found in the floral appendages. A cuticle-covered epidermis overlies a core of parenchyma cells in which there are branching vascular bundles (solitary bundles in most stamens). Sepal parenchyma and petal parenchyma are often spongy, but palisade parenchyma occurs only rarely in flowers and then only in sepals. As in other parts of the plant, color comes mostly from plastids in the cytoplasm and from flavonoids in the cell sap. Cells of the petal epidermis may have folded side walls that interlock so as to strengthen the tissue. In some species the outer walls of the epidermis are raised as papillae; apparently, this is part of the means of attracting pollinators, for the papillae are light reflectors. *See* LEAF.

Stamen. As a stamen develops, periclinal divisions in the second cell layer of each of its four lobes start a sequence that will end with the shedding of pollen (**Fig. 9**). The first division makes two cell layers. The outer daughter cells give rise to the wall of the pollen sac, and the inner ones (primary sporogenous cells) are destined to become pollen after further divisions. When mature, a pollen sac typically has a prominent cell layer, the endothecium, just below a less distinctive epidermis. The inner wall and the side walls of an endothecium cell carry marked thickenings, but the outer wall does not. Splitting of the ripe anther is due partly to the way in which these differently thickened walls react to drying and shrinking and partly to the smaller size of the cells along the line of splitting.

Pollen development. Primary sporogenous cells become pollen mother cells by enlargement or by division followed by enlargement. Through two meiotic divisions each mother cell becomes a tetrad of microspores (uninucleate pollen) with the haploid number of chromosomes. Cell walls disintegrate in the process, and a covering of callose, a polysaccharide, takes their place around the now rounded cells. Ordinarily, tetrads separate into monads before the anther opens, but some plants shed paired grains (*Scheuchzeria*); others shed tetrads (cattails, rhododendrons); still other shed polyads (certain acacias); and pollinators of many orchids pick up the agglutinated contents of a whole anther. While the sporogenous cells are changing to microspores, the pollen sac becomes lined with a densely cytoplasmic, sometimes multinucleate tissue, the tapetum, that nourishes the future pollen grains as they differentiate. Tapetal cells lose their walls in some groups of plants, and the protoplasts then move amebalike among the developing grains. A microspore becomes a pollen grain when its nucleus divides, yielding a generative nucleus and a tube nucleus. Membranes form

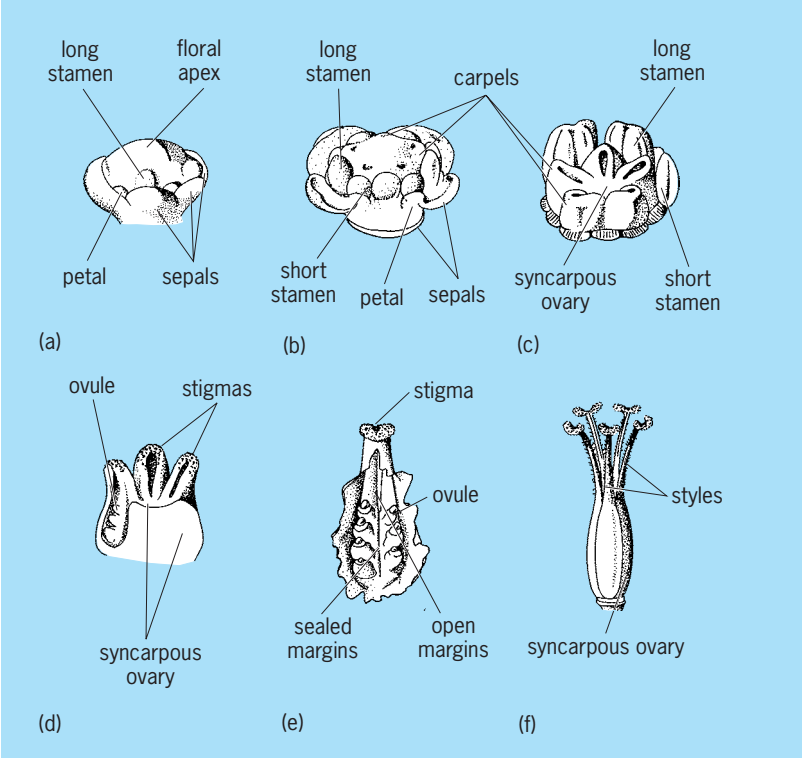

Fig. 8. Successive stages of flower formation in a wood sorrel. (*a*) Sepals and petals already initiated; the androecium begins as five bumps on the floral apex; five more will be added. At maturity, five long stamens will alternate with five short ones. (*b*) Carpel primordia taking horseshoe shape. Petal primordia lag while stamens grow. (*c*) Stamens lobed, outer parts cut away. (*d*) Gynoecium cut longitudinally. (*e*) One carpel opened from outside. (*f*) Mature gynoecium. (*After J. B. Payer, Trait d' Organogénie Comparée de la Fleur, Wheldon and Wesley, 1966*)

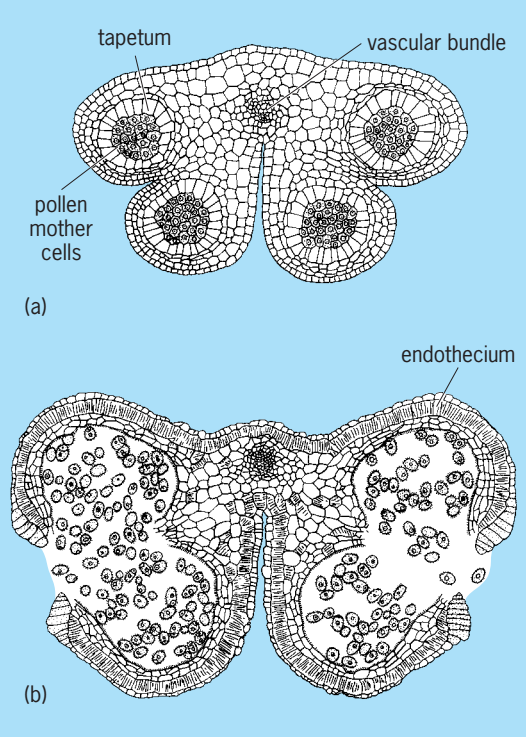

Fig. 9. Lily anther. (*a*) Cross section of young anther showing four microsporangia with sporogenous tissue. (*b*) Cross section of mature anther with two pollen sacs containing pollen. (*After A. W. Haupt, Plant Morphology, McGraw-Hill, 1953*)

around these nuclei (and in some cases a wall separates them), making the pollen grain two-celled. There are groups in which the generative nucleus cleaves almost immediately into two sperm nuclei, and the pollen is then three-nucleate, but it is more common for this final division to occur in the elongating pollen tube after the pollen germinates.

Pollen features. The pattern for outer features of the pollen grain, which vary greatly from one kind of plant to another, is set while the microspores are still encased in callose. Endoplasmic reticulum, differentially distributed under the outer membrane of each cell, determines where the germination apertures will be: no outer wall (exine) forms at the aperture sites. A temporary outer wall traversed by rods is laid down between these sites, and the addition of more material may then link the rods together. Eventually the microspores enlarge, the callose disappears, and the outer wall is permeated with sporopollenin, an oxidative polymer of carotenoids and carotenoid esters. Sporopollenin's resistance to decay and the pollen wall's range of taxonomically useful features make pollen an invaluable guide to evolution and to past environmental conditions. *See* POLLEN.

Carpel. Like other floral parts, a carpel is made up of epidermis, parenchyma, and vascular tissue. In addition, a carpel commonly has a special tissue system on which pollen germinates and through which, or along which, pollen tubes are transmitted to the ovules. Most angiosperms have solid styles, and the transmitting tissue is a column of elongate cells whose softened walls are the medium for tubal growth. The epidermis at the stigmatic end of a carpel usually changes to a dense covering of papillae or hairs; the hairs can be unicellular or pluricellular, branched or unbranched. In taxa with hollow styles, the transmitting tissue is a modified epidermis running down the stylar canal. There are two kinds of receptive surfaces, and they are distributed among the monocotyledons and the dicotyledons with taxonomic regularity. One kind has a fluid medium for germinating the pollen, and the other has a dry proteinaceous layer over the cuticle. The proteins of the dry stigmas have a role in the incompatibility reactions that encourage outbreeding.

A solitary carpel commonly has three main vascular bundles: a midvein and two marginal veins. It was realized more than a century ago that the bundles can often be used to delineate the carpels of compound gynoecia if other indicators, such as styles and placentas, are ambiguous. This works because vascular strands begin to form when the gynoecial primordia are quite small, and they can continue to differentiate as bundles while other primordial features are obscured by subsequent developmental changes. That is to say, a carpel bundle can be something of an ontogenetic marker. Some botanists built on this, as organic evolution became widely accepted, to claim that vascular bundles—all vascular bundles—are phyletically conservative; that they tend to lag in various ways as the outer form undergoes evolutionary changes; that a vestigial midvein, for instance, can mark the position of a lost appendage. And some individuals, who were unfamiliar with evolutionary theory or uncomfortable with it, or more comfortable using words from old books, saw bundles as a clue to a flower's "nature."

The view that bundles are conservative underlies most American research on floral anatomy during the 1930s, 1940s, and 1950s. It has since been questioned as a generalization. That is not to say there are no lagging or vestigial bundles: an evolving lineage might very well keep an old floral bundle in place if the bundle continues to do something useful for the plant.

Ovule. Ovule development usually takes place as the gynoecium forms, but it may be retarded when there is a long interval between pollination and fertilization (oaks, orchids). A typical ovule has a stalk (funiculus), a central bulbous body (nucellus), and one or two integuments (precursors of seed coats), which cover the nucellus except for a terminal pore (micropyle). Orientation of the ovule varies from group to group. It can be erect on its stalk or bent one way or another to differing degrees. There are also taxonomic differences in the extent to which the ovule is vascularized by branches from the gynoecial vascular system.

Fertilization

The union of an angiosperm egg and sperm takes place in an embryo sac that develops within the ovule in one of several ways.

Embryo sac. The most common way is for a large cell near the micropylar end of the nucellus to divide meiotically to form a lineup of four haploid megaspores. Three of the megaspores deteriorate; the one farthest from the micropyle enlarges and divides; and the two daughter nuclei part, one moving to the micropylar end of the cell, the other to the opposite end. Two subsequent divisions and further enlargement (at the expense of the nucellus) make an eight-nucleate embryo sac with four nuclei at each pole. Three of the nuclei at the micropylar end then differentiate further and acquire membranes to form the egg apparatus, a triplet consisting of a central egg cell with two flanking cells called synergids. Three of the four nuclei at the other end become antipodal cells. The remaining nucleus (polar nucleus) from each foursome moves to the middle of the sac, pairing with its counterpart, but cell membranes do not form. The mature sac consequently has eight nuclei but only seven cells (**Fig. 10***a*). The two polar nuclei may fuse to become a diploid secondary endosperm nucleus before fertilization. The simplified sequences just described produces the *Polygonum* type of embryo sac. There are others that differ with regard to such features as the number of megaspores and the number and distribution of the nuclei in the mature sac. These differences, as well as differences in the thickness of the nucellus and the number of integuments, can be important taxonomic characters. Seen from the standpoint of alternation of

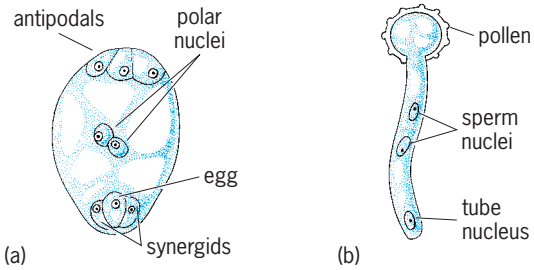

Fig. 10. Gametophyte: (*a*) embryo sac and (*b*) pollen tube. (*After O. Tippo and W. L. Stern, Humanistic Botany, W. W. Norton, 1977*)

generations, one of the unifying themes of botany, an embryo sac is the megagametophyte of the angiosperms; a pollen tube with its contents is the microgametophyte.

Pollen tube. Emergence of the tube from the pollen grain (Fig. 10*b*) occurs as soon as the grain alights on the stigma (sugarcane), after a few minutes (dandelion), or after a longer interval—as long as two days in *Garryaf*. The grain expands by absorbing liquid from the stigma, and the plastic inner layer (intine) of the pollen wall begins to bulge through one of the apertures, where there is no exine above it. If more than one tube emerges, a characteristic of some groups, only one makes progress.

Following the transmitting tissue, tubes grow at the rate of a few millimeters per hour, usually reaching the ovule within 48 h or less, but the delay between pollination and fertilization can be weeks, even months, in some plants. The tube passes through the micropyle (in most cases), penetrates the embryo sac, and releases two sperm cells from its tip. One sperm unites with the egg cell to form a zygote, which after additional divisions becomes an embryo. The other sperm fuses with the polar nuclei or with their fusion product, the secondary endosperm nucleus; subsequent division of the resulting triple fusion nucleus makes endosperm tissue (triploid if the embryo sac is of the *Polygonum* type), which nourishes the developing embryo. Ordinarily, pollination triggers the first cell divisions toward the changing of a gynoecium to a fruit, but fertilization and seed development are required for a fruit to achieve its full growth. *See* REPRODUCTION (PLANT).

Utilization

Although human life depends on the consumption of fruits and seeds, eating the flower that precedes a fruit is something of a rarity. Pickled capers and dried cloves are flower buds, and saffron is the dried stigmas of a crocus. Honey is floral nectar, concentrated and altered by its sojourn in a beehive. Flowers are used in quantity by the perfume industry, even though many fragrances can now be synthesized. The essential oils are removed by pressing, by distillation, by extraction with a solvent, or, as in the case of highly valued jasmine, they may still be extracted by enfleurage, laying the flowers in cold tallow or lard. Raising and exporting cut flowers is a major industry in the Netherlands, in some Latin American countries, and in California and Florida; several European countries and many areas in the United States have research centers for floriculture. Thus, the flower as ornament clearly has a place in today's technological world, but to prize a flower for its beauty is not new. Archeologists have found remnants of bouquets, including bachelor's buttons and grape hyacinths, in a Neanderthal grave more than 50,000 years old. *See* FRUIT; MAGNOLIOPHYTA.

Richard H. Eyde

Bibliography. K. Esau, *Anatomy of Seed Plants*, 2d ed., 1977; A. Hallam (ed.), *Patterns of Evolution as Illustrated by the Fossil Record*, 1977; A. H. Halvey (ed.), *CRC Handbook of Flowering*, 5 vols., 1985; P. H. Raven, R. F. Evert, and H. Curtis, *Biology of Plants*, 6th ed., 1998; A. J. Richards, *Plant Breeding Systems in Seed Plants*, 1986; D. L. Strider (ed.), *Diseases of Floral Crops*, 2 vols., 1985.

Fluid coating

The operation of depositing a liquid film on a solid, due to a relative motion between them (**Fig. 1**). It occurs in many situations of everyday life as well as in numerous industrial processes. Examples include a water drop sliding on a window and leaving behind a trail; an object or a person coming out wet from a bath; or the emptying of a solid full of liquid (such as a glass containing water or a porous medium filled with oil), leaving some liquid inside the solid. In all these cases, an important practical question is the thickness of the film adhering to the solid. It may be desirable to control the thickness for coating purposes, or conversely to avoid a film altogether to keep the solid dry. To achieve this control, it is important to understand the parameters affecting the existence of such a film, and the laws of deposition.

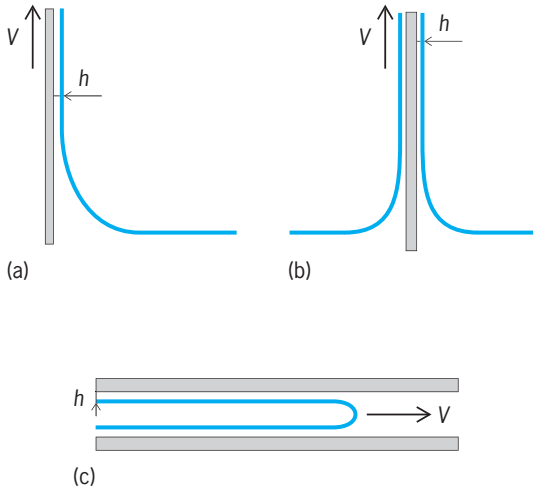

Fig. 1. Different kinds of fluid coating, according to the geometry. (*a*) Plate coating. (*b*) Fiber coating. (*c*) Film deposition in a tube. In each case, the macroscopic film coating the solid comes from a relative motion between the liquid and the solid, and the film thickness *h* can be tuned by the coating velocity *V*.

Before addressing a dynamic problem, it is often worth describing the statics. The very general situation in which a solid, a liquid, and a vapor meet together occurs, for example, in the case of a drop deposited on a solid with air all around. The different surface energies involved in such a situation (between the solid and the liquid, the liquid and the vapor, and the solid and the vapor) impose an equilibrium condition at the contact line where the three phases coexist. Since each surface tension tries to draw the contact line so as to reduce the corresponding interface, the balance of the tensions, which defines the equilibrium, fixes a contact angle close to this line. For a drop, it is the angle with which the liquid-vapor interface meets the solid. A liquid totally wets a solid if it joins it with a zero contact angle. If that is not the case (contact angle larger than zero), the corresponding situation is referred to as partial wetting. *See* SURFACE TENSION.

Total wetting can occur by itself (for example, oils on most solids) or be realized by adding wetting agents to the liquid (the most common ones being surfactants, for example, soaps in water). Drawing a solid out of such a liquid always produces a film deposition. Of course, this deposition is favored by the wetting situation, since a film would spontaneously spread from the bath onto the solid, even at rest. But such wetting films are microscopic, which is not generally the case in fluid coating (at least, for sufficiently large coating velocities): the trail of a drop is visible. The thickness of the coating film can be determined with the help of a canonical law of hydrodynamics: Because of its viscosity, a liquid in the close vicinity of a solid moves at the same velocity as the solid. This principle is the basis of fluid coating and the cause of entrainment, and calculating the film thickness requires identifying the extent of this vicinity. Different answers exist, depending on the velocity of deposition. *See* BOUNDARY-LAYER FLOW; SURFACTANT; VISCOSITY.

Low-velocity coatings. The case of low velocities was understood in the 1940s by the Russian physicists Lev Landau, Benjamin Levich, and Boris Derjaguin. The Landau-Levich-Derjaguin (LLD) picture is as follows: Drawing a solid out of a bath implies a deformation of the free surface (Fig. 1*a*), which the surface tension of the liquid opposes. The film thickness thus results from a balance between the effects of viscosity (which causes a macroscopic entrainment of liquid by the solid) and surface tension of the liquid, which resists to this deformation (and thus to the film entrainment). A dimensionless number called the capillary number compares the intensity of these antagonistic effects (defined as $\mathrm{Ca} = \eta V/\gamma$, with η and γ the liquid viscosity and surface tension, and V the deposition velocity). In the limit of low velocities, the film thickness, h, should be a function of this unique parameter (this function was found by Landau, Levich, and Derjaguin to obey a simple power law: $h \sim \mathrm{Ca}^{2/3}$), whatever the geometry of the coating. This is indeed the case, and the geometry enters only through the length that normalizes the film thickness: For plate coating, this length is the capillary length (a millimetric length that corresponds physically to the size of the static meniscus, formed when a wetting liquid meets a vertical plate), or the capillary radius for a wire or a capillary tube.

Coating at higher velocities. The hydrodynamics of fluid coating becomes more complicated at higher velocities. Then, inertia must be taken into account—in the above example, the inertia of the fluid set in motion by the solid in the bath. When the moving fluid comes out from the bath, in the tiny zone where the film thickness fixes, its inertia must be compared with the capillary force (which was found to limit the film thickness). A negligible inertia defines the LLD frame. If the inertia is not negligible, it should act to make the film thicker than predicted by the LLD picture, since it tends to pull the liquid out of the bath. It was shown experimentally in fiber (or wire) coating that this is indeed the case. The film thickness displays critical behavior as a function of velocity: Above a threshold in velocity (typically of order 1 m/s or 3 ft/s), a simple doubling of the velocity causes the film thickness to increase about tenfold (**Fig. 2**).

At still larger velocities, the film thickness stops increasing with the coating speed and even slowly decreases with it (Fig. 2). This regime concerns most industrial coating processes, where the coating speed can be of order 10–30 m/s (30–100 ft/s). At such high speeds, the contact time of a piece of solid in the bath becomes very short (it is simply $t = L/V$, with L the bath length). Thus, the problem now is to understand how much liquid this piece of solid can put in motion inside the bath—a classical question, called the viscous boundary layer problem and solved in the early twentieth century by Ludwig Prandtl. Balancing inertia with the viscous force gives the answer: The thickness of the layer moved by the solid (and thus coming out with it) increases as the square root of the contact time (it scales as $\sqrt{\eta t/\rho}$, with ρ the liquid density). Hence, the film thickness

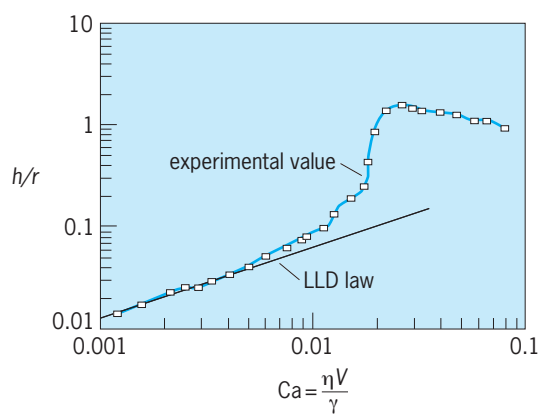

Fig. 2. Different regimes of fluid coating, according to the coating speed. These experimental data correspond to fiber coating. The film thickness *h* normalized by the fiber radius *r* is plotted as a function of the capillary number Ca. Experimental values are compared with the Landau-Levich-Derjaguin (LLD) law for low-velocity flow.

in this regime indeed slowly decreases with the coating speed (as $1/\sqrt{V}$, since t varies as $1/V$). This law expresses a balance between the viscosity (which is, of course, still responsible for the coating) and inertia—but now, the inertia of the bath, which is globally at rest and thus resists the motion. The action of the liquid tension, which was predominant at low velocity (the LLD frame), has become negligible in this high-speed regime.

Complex fluid coatings. Most liquids used in coating processes are complex fluids: emulsions (which are fine dispersions of oil in water) or water-containing surfactants. In the cases where inertial effects are dominant (high-speed regimes), the above description remains valid because the parameters fixing the coating are bulk parameters (viscosity and density). The situation is much more interesting at smaller velocities, where surface effects are important. For example, a soap bubble is drawn from soapy water with a well-defined thickness, but a solid is not necessary to make it. (Likewise, bubbles form at the liquid-air interface in a glass of beer.) Hence, a liquid-vapor interface has the same ability as a solid to entrain matter, providing that surfactants are present at this interface. This is due to the fact that in such a process the surfactants at the surface are not homogeneous in concentration, but diluted by the motion: A gradient of surfactant concentration exists there, and thus a gradient of surface tension, which induces a force that is able to entrain liquid below the surface. If a solid is drawn out of a bath containing surfactants, the film thickness can be expected to be about twice as large as is predicted by the LLD theory for pure liquids, because two interfaces (solid-liquid and liquid-vapor) are now responsible for the coating, instead of one. This is indeed what the experiments show (although more sophisticated arguments involving, for example, the kinetics of adsorption of the surfactant must be taken into account for a complete quantitative understanding of the data). *See* SURFACE COATING.

Coating threshold velocity. A solid slowly drawn out of a liquid which only partially wets it can come out dry, as is commonly observed. For example, a drop moving on a water lily leaves the leaf dry after passing on it. At rest, the liquid meets the solid with a nonzero contact angle. This angle decreases when the solid is moved, but it can remain nonzero thanks to the action of the liquid surface tension which opposes the deformation of the meniscus. Above a threshold in velocity (typically of the order of 1 cm/s or 0.4 in./s), viscous forces dominate capillary ones: A film is deposited and the LLD regime is recovered. The critical velocity of deposition is not yet fully understood, but it has been shown to depend on both the static contact angle and the liquid viscosity.

Self-running droplets. In some cases, the coating may form by itself, instead of being forced to form by an external motion. The so-called running droplets discovered separately by Colin Bain and Thierry Ondarçuhu are a remarkable example (**Fig. 3*a***). These drops contain molecular species which have

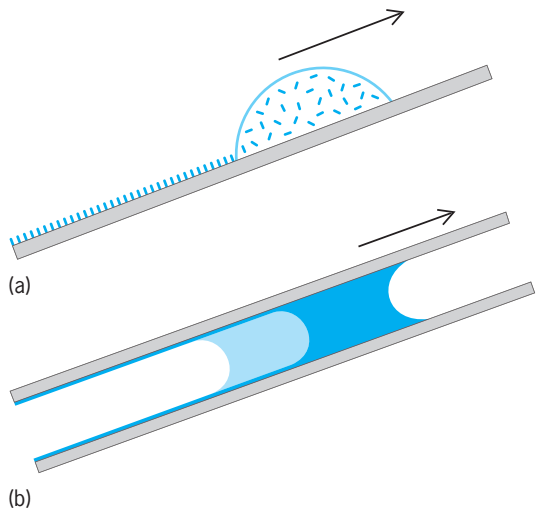

Fig. 3. Self-running droplets. (*a*) Bain-Ondarçuhu device. (*b*) Bico bi-drop.

a double property: first, they are likely to react with the solid; second, the solvent ceases to wet the solid once it is coated with these molecules. Depositing such a drop on a solid generally produces a spontaneous motion (self-running droplets). If the chemical reaction with the solid starts for some reason on one side of the drop, then the two sides sit on different solids: one on the bare solid, the other on the solid coated with the reactive species. This asymmetry provokes the motion, until the reservoir of reactives is exhausted or the whole solid is coated.

In 2005, Kenichi Yoshikawa and coworkers proposed reversing the principle: On a glass surface coated with surfactants (and immersed in soapy water), an oil drop is deposited that contains a reactive species that is likely to extract the surfactant. As the oil drop starts taking the surfactant off, it begins to move (again due to a gradient of wettability below the drop). Owing to the redeposition of the surfactant on the glass, the solid surface is regenerated once the drop has moved, allowing the drop to pass the same location many times, until the reactant in the drop is fully saturated with surfactants.

While the Bain-Ondarçuhu drops realize a dry coating of the surface, other possibilities exist if a wet coating is desired. For example, the so-called bi-drops discovered by José Bico, which juxtapose two (wetting) drops inside a tube, also produce a self-motion, again because of the asymmetry of the device (Fig. 3*b*). Each drop can leave a film behind, so that a great variety of coatings (including double coatings) can be realized. The coatings can also be used to achieve better transport properties: The film left behind by the first drop can lubricate the way for the second one. Liquids of very high viscosities can be self-transported by such a device.

In this case, it is in a sense the coating which makes the motion (instead of the motion, the coating, as is usually the case). The driving forces are capillary in essence, so that these droplets can creep upward, as in capillary rise. Research on self-running drops is active and promising, in particular in the context

of microfluidics. More sophisticated devices may be possible, where the coating can be addressed and tuned. David Quéré

Bibliography. B. V. Derjaguin and S. M. Levi, *Film Coating Theory*, 1964; S. F. Kistler and P. M. Schweizer (eds.), *Liquid Film Coating*, 1997; D. Quéré, Fluid coating on a fiber, *Annu. Rev. Fluid Mech.*, 31:347–384, 1999; K. J. Ruschak, Coating flows, *Annu. Rev. Fluid Mech.*, 17:65–89, 1985.

Fluid coupling

A device for transmitting rotation between shafts by means of the acceleration and deceleration of a hydraulic fluid. Structurally, a fluid coupling consists of an impeller on the input or driving shaft and a runner on the output or driven shaft. The two contain the fluid (**Fig. 1**). Impeller and runner are bladed rotors, the impeller acting as a pump and the runner reacting as a turbine. Basically, the impeller accelerates the fluid from near its axis, at which the tangential component of absolute velocity is low, to near its periphery, at which the tangential component of absolute velocity is high (**Fig. 2**). This increase in velocity represents an increase in kinetic energy. The fluid mass emerges at high velocity from the impeller, impinges on the runner blades, gives up its energy, and leaves the runner at low velocity (**Fig. 3**). *See* FLUID MECHANICS; HYDRAULICS.

Coupling characteristics. A fluid coupling absorbs power from an engine by means of the centrifugal pump or impeller, which increases the moment of momentum of the mass of fluid. As a consequence, the input power required to drive the coupling increases as the cube of the speed and as the fifth power of the outside diameter of the impeller. Similarly, the capacity of the coupling to absorb torque varies as the square of its input speed.

Because the efficiency of a coupling is a function of slip, couplings normally operate at 95–98% efficiency. The fluid follows a continuous toroidal helix, the pitch of which depends on the slip, or relative difference in rotational velocity, between impeller and runner. Some slip is necessary for energy transfer. As the slip increases at a given input speed, the output shaft turning slower than the input shaft, the quantity of fluid pumped between impeller and runner increases, with consequent increase in torque absorption and transmission. However, because the fluid flow cannot change abruptly, a fluid coupling absorbs rather than transmits shock loads from output to input. Torsional vibrations are similarly absorbed.

A fluid coupling is designed for a normal running load at 2–5% slip. Conservation of momentum requires that, although shaft speeds may differ, input and output torques must be equal. Thus, for example, in starting a machine from standstill, with the output from the fluid coupling locked, the engine or motor is completely unloaded initially. As the engine starts up toward running speed, the drag torque characteristic of the fluid coupling lies below the developed torque characteristic of the engine. The engine can reach running speed faster with consequent saving in fuel or electrical consumption. The slip of the fluid coupling enables the engine to operate near its maximum-torque speed while accelerating the load, although this characteristic differs from

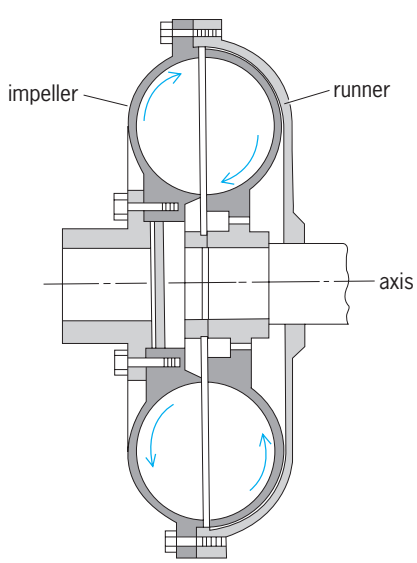

Fig. 1. Basic fluid coupling.

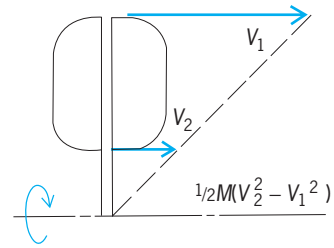

Fig. 2. Hydrokinematic principle.

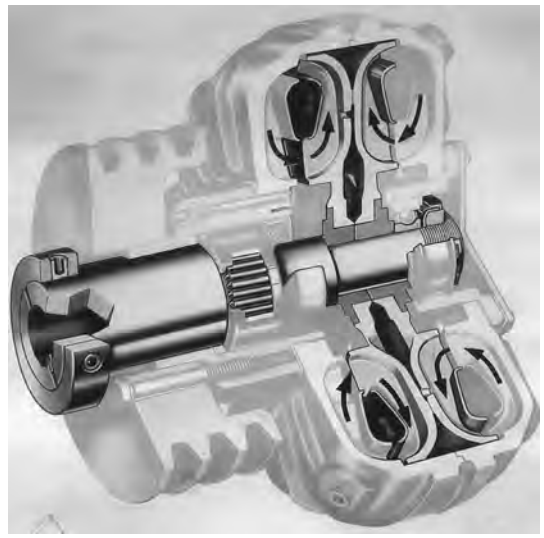

Fig. 3. Sheave with fluid coupling. Dark portion of the figure is driven directly by motor; light portion drives the load. (*Twin Disc Inc.*)

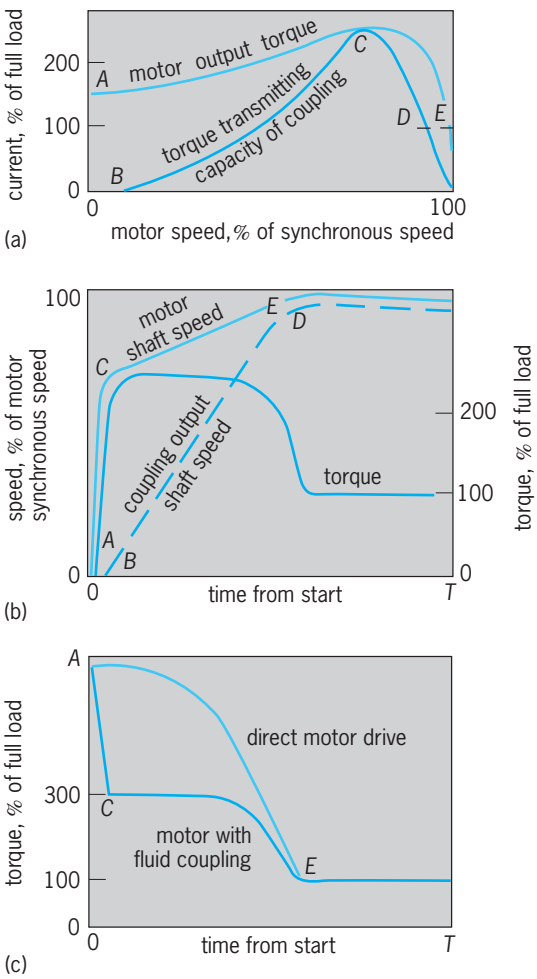

Fig. 4. Starting characteristics of electric motor with fluid coupling. (*a*) Torque speed during acceleration. (*b*) Acceleration. (*c*) Current.

that of a variable-torque transmission. *See* TORQUE CONVERTER.

Typical application. The effect of a fluid coupling can be described in terms of a specific application. If an electric motor is starting with a high inertia load that requires nearly full-motor torque, the motor accelerates slowly when connected directly to its load. However, with fluid coupling, the motor accelerates rapidly along A to C in **Fig. 4**. Meanwhile, the torque-transmitting capacity of the coupling builds up along B to C until it equals the developed motor torque. The motor then can deliver well over full load torque to accelerate the load; the torque demand falls from C to D as the load comes up to speed D. Meanwhile, the motor further accelerates C to E until steady state is reached with coupling slip represented by differential speed D to E. Use of fluid coupling improves performance by faster load acceleration and results in appreciably less power loss due to motor heating. *See* AUTOMOTIVE TRANSMISSION; COUPLING.

Henry J. Wirry

Bibliography. W. Bober and R. A. Kenyon, *Fluid Mechanics,* 1980; W. H. Crouse, *Automotive Transmissions and Power Trains*, 5th ed., 1976.

Fluid-flow principles

The fundamental principles that govern the motion of fluids. This article is concerned with these principles and the manner in which they are applied to problems of practical interest. It is assumed that a fluid can be modeled as a continuum and that the details of the molecular structure will appear only in the coefficients associated with constitutive equations or equations of state.

Fundamentals

Understanding the motion of a continuum and the laws that govern that motion is not easy; thus, it should not be surprising that the understanding of motion has its origins in the study of particles, or rigid bodies. While Newton's laws provided a basis for the study of the motion of particles, or more precisely mass points, it was L. Euler (1707–1783) who organized newtonian mechanics into a form which allows a variety of mechanical problems to be solved with relative ease. *See* CLASSICAL MECHANICS; NEWTON'S LAWS OF MOTION.

Euler's laws of mechanics. Euler represented Newton's ideas in terms of two laws of mechanics, which can be stated as follows: (1) The time rate of change of the linear momentum of a body equals the force acting on the body. (2) The time rate of change of the angular momentum of a body equals the torque acting on the body.

It is understood that the distances and velocities associated with these two laws are measured relative to an inertial frame and that the torque and angular momentum are measured relative to the same fixed point. Euler's first law contains all of Newton's three laws except for the part of the third law requiring that the actions of two bodies be directed to contrary parts. This is most often referred to as the central force law, and it is valid only for nonrelativistic systems. *See* ANGULAR MOMENTUM; MOMENTUM; RIGID-BODY DYNAMICS.

In addition to providing a concise statement of the laws of mechanics for continua, Euler postulated that these two laws hold not only for distinct bodies but also for any arbitrary body one might wish to consider. This idea is referred to as the Euler cut principle, and it provides the key to the derivation of governing differential equations from axiomatic statements.

Conservation of mass. The body referred to in Euler's laws consists of some fixed quantity of material that may move and deform in an arbitrary but continuous manner with time t. The region occupied by a body is designated by $V_m(t)$ [**Fig. 1**], and the principle of conservation of mass for a body is stated as follows: The time rate of change of the mass of a body is zero.

Density. If the mass density of the continuum is designated by ρ, the mass of the body is expressed by Eq. (1), where the right-hand side is a volume

$$\text{Mass of a body} = \int_{V_m(t)} \rho \, dV \qquad (1)$$

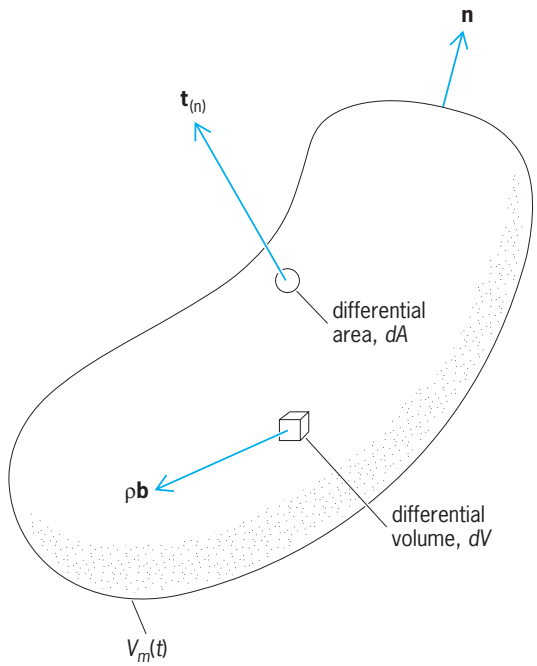

Fig. 1. A material volume.

integral over $V_m(t)$, dV being a differential volume element. The principle of conservation of mass takes the form of Eq. (2), where d/dt represents differen-

$$\frac{d}{dt}\int_{V_m(t)} \rho\, dV = 0 \tag{2}$$

tiation with respect to time, or rate of change. The density has units of mass/volume, and is nearly constant for liquids and solids. For gases, ρ depends on the temperature and pressure, and an equation of state is required in order to predict the density. *See* DENSITY; GAS; LIQUID.

To obtain a governing differential equation from the axiomatic statement given by Eq. (2), it is necessary to interchange differentiation and integration. G. W. Leibnitz (1646–1716) derived a theorem for accomplishing this, and when this theorem is applied to the time-dependent volume $V_m(t)$, it is usually known as the Reynolds transport theorem. Application of this theorem to Eq. (2) leads to Eq. (3),

$$\int_{V_m(t)} \left(\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v})\right) dV = 0 \tag{3}$$

and since the volume $V_m(t)$ is arbitrary, the integrand must be zero and continuity equation (4) results,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{4}$$

where the first term on the left-hand side represents the time rate of change of mass per unit volume, and the second term represents the net flux of mass, leaving a differential volume. In going from Eq. (3) to Eq. (4), it is assumed that the density ρ and the velocity $\mathbf{v}$ are continuously differentiable with respect to both time and space and that $V_m(t)$ is arbitrary. The velocity vector can be expressed in terms of base vectors and components as Eq. (5), and the gra-

$$\mathbf{v} = \mathbf{i}v_x + \mathbf{j}v_y + \mathbf{k}v_z \tag{5}$$

dient operator can be expressed in similar form as Eq. (6). Continuity equation (4) will be used repeat-

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z} \tag{6}$$

edly below. *See* CALCULUS OF VECTORS.

The principle of conservation of mass, as given by Eq. (2), should have obvious intuitive appeal; however, when this concept is applied to a vanishingly small volume (that is, when dV in Fig. 1 is allowed to approach zero), Eq. (4) is obtained and the physical meaning is less clear. When Eq. (4) is integrated over the control volume shown in **Fig. 2**, the macroscopic mass balance is obtained, and for steady flow in a tube it leads to the statement (7).

$$\begin{array}{l}\text{Rate at which}\\ \text{mass flows into}\\ \text{the control volume}\end{array} = \begin{array}{l}\text{Rate at which}\\ \text{mass flows out of}\\ \text{the control volume}\end{array} \tag{7}$$

Linear momentum principle. Euler's laws may be expressed in precise mathematical form. The linear momentum of the differential volume element shown in Fig. 1 is the mass ρdV times the velocity $\mathbf{v}$, and this leads to Eq. (8). The total force acting on a body

$$\begin{array}{l}\text{Linear momentum}\\ \text{of a body}\end{array} = \int_{V_m(t)} \rho \mathbf{v}\, dV \tag{8}$$

consists of body forces (gravitational, electrostatic, and electromagnetic) that act at a distance upon the mass, and surface forces that act on the bounding surface of the body. As indicated in Fig. 1, the body force per unit volume is represented by $\rho\mathbf{b}$, and the body force acting on the entire body is given by Eq. (9).

$$\begin{array}{l}\text{Body force acting}\\ \text{on the body}\end{array} = \int_{V_m(t)} \rho \mathbf{b}\, dV \tag{9}$$

The surface force per unit area is expressed by the stress vector, which is indicated in Fig. 1 by $\mathbf{t}_{(\mathbf{n})}$. The subscript $(\mathbf{n})$ is used as a reminder that the stress vector is a function of the unit outwardly directed normal vector for the surface under consideration. In terms of $\mathbf{t}_{(\mathbf{n})}$, the surface force acting on the body

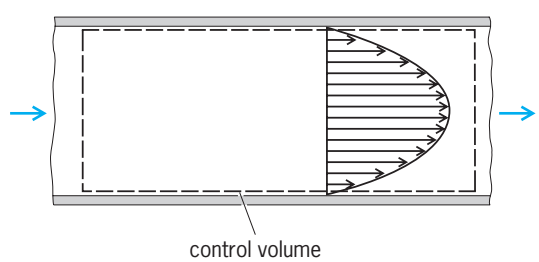

Fig. 2. Steady laminar flow in a tube.

becomes Eq. (10), in which $A_m(t)$ represents the sur-

$$\text{Surface force acting on the body} = \int_{A_m(t)} \mathbf{t}_{(\mathbf{n})}\, dA \qquad (10)$$

face of the body illustrated in Fig. 1.

Equations (8) through (10) can be used to express Euler's first law as Eq. (11), where the left-hand side is

$$\frac{d}{dt}\int_{V_m(t)} \rho \mathbf{v}\, dV = \int_{V_m(t)} \rho \mathbf{b}\, dV + \int_{A_m(t)} \mathbf{t}_{(\mathbf{n})}\, dA \qquad (11)$$

the time rate of change of momentum, and the terms on the right-hand side represent the body force and the surface force respectively.

The Euler cut principle forced the scientists of the eighteenth century to contemplate the internal state of stress that appeared when an arbitrary body was cut out of a distinct body. It was nearly a century after Euler presented his ideas on the mechanics of continua that A. L. Cauchy (1789–1857) determined the functional dependence of the stress vector on the unit normal vector $\mathbf{n}$. This is given in terms of two lemmas that can be expressed as Eqs. (12) and (13), and both of these relations are a natural con-

$$\mathbf{t}_{(\mathbf{n})} = -\mathbf{t}_{(-\mathbf{n})} \qquad (12)$$

$$\mathbf{t}_{(\mathbf{n})} = \mathbf{T} \cdot \mathbf{n} \qquad (13)$$

sequence of the form of Eq. (11). The first lemma is reminiscent of Newton's third law of action and reaction. The second lemma is not so obvious, since it introduces the doubly directed quantity known as the stress tensor. It is best to think of $\mathbf{T}$ as a function that maps the unit normal vector $\mathbf{n}$ into the stress vector $\mathbf{t}_{(\mathbf{n})}$. This idea is shown in matrix form as Eq. (14).

$$\begin{pmatrix} t_{(\mathbf{n})x} \\ t_{(\mathbf{n})y} \\ t_{(\mathbf{n})z} \end{pmatrix} = \begin{pmatrix} T_{xx} & T_{xy} & T_{xz} \\ T_{yx} & T_{yy} & T_{yz} \\ T_{zx} & T_{zy} & T_{zz} \end{pmatrix} \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix} \qquad (14)$$

See MATRIX THEORY.

Angular momentum principle. A precise mathematical statement of Euler's second law is obtained by following the procedure given by Eqs. (8) through (11) to obtain Eq. (15), where the left-hand side represents

$$\frac{d}{dt}\int_{V_m(t)} \mathbf{r} \times \rho \mathbf{v}\, dV = \int_{V_m(t)} \mathbf{r} \times \rho \mathbf{b}\, dV + \int_{A_m(t)} \mathbf{r} \times \mathbf{t}_{(\mathbf{n})}\, dA \qquad (15)$$

the time rate of change of angular momentum, and the two terms on the right-hand side represent the torque owing to the body force and the torque owing to the surface force. The mathematical analysis of this equation is extremely complex, and after a great deal of effort it can be proved that Eq. (15) yields only the result that the stress tensor is symmetric. This result can be expressed as Eqs. (16) and indicates that the stress matrix represented in Eq. (14) is symmetric. With this result from Euler's second law, (11) can be used to obtain the differential form of Euler's first law. This is given by Eq. (17), where the

$$\begin{aligned} T_{xy} &= T_{yx} \\ T_{xz} &= T_{zx} \\ T_{yz} &= T_{zy} \end{aligned} \qquad (16)$$

$$\frac{\partial}{\partial t}(\rho \mathbf{v}) + \nabla \cdot (\rho \mathbf{v}\mathbf{v}) = \rho \mathbf{b} + \nabla \cdot \mathbf{T} \qquad (17)$$

two terms on the left-hand side represent the time rate of change of momentum per unit volume, and the net flux of momentum leaving a differential volume; and the two terms on the right-hand side represent the body force per unit volume and the surface force per unit volume. Equation (17) is referred to as Cauchy's first equation or the stress equation of motion, since the surface forces are represented in terms of the stress tensor $\mathbf{T}$. The continuity equation in the form of Eq. (18) can be used in order to simplify Eq. (17) to Eq. (19). It is important to think of

$$\mathbf{v}\left(\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v})\right) = 0 \qquad (18)$$

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v}\right) = \rho \mathbf{b} + \nabla \cdot \mathbf{T} \qquad (19)$$

Eq. (4) as the governing differential equation for the density ρ and Eq. (19) as the governing differential equation for the velocity $\mathbf{v}$. Euler's second law, given by Eq. (15), provides only a constraint on the stress field, so that there are only six distinct components of the stress tensor $\mathbf{T}$. At this point there are more unknowns than equations, and in order to solve fluid-mechanical problems a constitutive equation for $\mathbf{T}$ is needed.

Fluid characteristics. A fluid is a material that deforms continuously under the action of a shear stress. Air and water are obvious examples of fluids, and their mechanical characteristics are clearly much simpler than catsup, meringue, or toothpaste. It is convenient to separate fluids into those which possess elastic characteristics (often called viscoelastic fluids) and those which do not (often called viscous fluids). In this article, only viscous fluids will be discussed, and for such fluids it is convenient to decompose the stress tensor according to Eq. (20). Here p

$$\mathbf{T} = -p\mathbf{I} + \tau \qquad (20)$$

is the thermodynamic pressure, $\mathbf{I}$ is the unit tensor, and t is the symmetric viscous stress tensor. Use of Eq. (20) in Eq. (19) leads to Eq. (21). Here a consti-

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v}\right) = -\nabla p + \rho \mathbf{b} + \nabla \cdot \tau \qquad (21)$$

tutive equation for t is needed in order to obtain an equation that can be used to determine the velocity field. *See* FLUIDS.

Viscosity. The viscosity of a fluid is a measure of its resistance to deformation, and a convenient method of deforming a fluid is illustrated in **Fig. 3**. The lower plate is fixed while the upper plate is moving at a constant velocity u_0, giving rise to a simple shear flow. Experiments of this type have shown that for a large

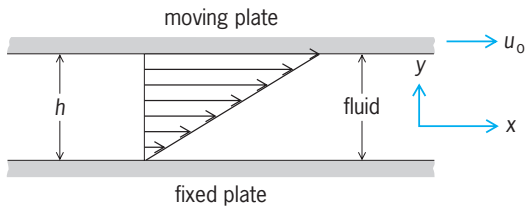

Fig. 3. Velocity distribution between two parallel plates.

class of fluids the force required to maintain the motion of the upper plate is given by Eq. (22). Here F

$$F = \mu \left(\frac{u_0}{b}\right) A \tag{22}$$

represents the force, A is the area of the surface of the plate in contact with the fluid, b is the perpendicular distance between the plates, and μ is the coefficient of shear viscosity. More commonly, μ is referred to simply as the viscosity. If the force per unit area acting in the x direction on a surface having a normal vector parallel to the y direction is designated τ_{yx}, the constitutive equation that is consistent with Eq. (22) is given by Eq. (23). Fluids which obey this relation

$$\tau_{yx} = \left(\frac{F}{A}\right) = \mu \left(\frac{\partial v_x}{\partial y}\right) \tag{23}$$

between the shear stress τ_{yx} and the rate of deformation $\partial v_x/\partial y$ are called newtonian fluids. Many fluids exhibit this type of behavior; however, other fluids do not. These other fluids can be separated into two groups: viscoelastic fluids and non-newtonian fluids. The behavior of non-newtonian fluids is illustrated in **Fig. 4**. The Bingham plastic is representative of fluids which remain rigid until a certain critical shear stress is applied. Such fluids tend to slip at solid surfaces, and in tubes they flow as a plug of toothpastelike material. Shear-thinning fluids tend to become less viscous as they are sheared more vigorously, while shear-thickening fluids become more viscous. A concentrated suspension of cornstarch in water is an example of a shear-thickening fluid. *See* NEWTONIAN FLUID; NON-NEWTONIAN FLUID; VISCOSITY.

Newton's law of viscosity. For complex flows, all six of the distinct components of the viscous stress tensor must be represented in terms of the components of the velocity gradient tensor. For linear, isotropic, viscous fluids, this representation takes the form of Eq. (24). Known as Newton's law of viscosity, it is one

$$\tau = \mu(\nabla \mathbf{v} + \nabla \mathbf{V}^T) + [(\kappa - {}^2/_3\mu)\nabla \cdot \mathbf{v}]\mathbf{I} \tag{24}$$

of a series of constitutive equations used to describe the behavior of linear, isotropic materials. Others are Ohm's law of electrical conductivity, Hooke's law of elasticity, and Fourier's law of heat conduction. In Eq. (24), κ is known as the bulk coefficient of viscosity, and it is of importance only for a special class of compressible flows, such as the damping of sonic disturbances. For a very wide class of flows, the second term in Eq. (24) is negligible, and use of this result in Eq. (21) leads to Eq. (25). Here ∇^2 represents the

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) = -\nabla p + \rho\mathbf{b} + \mu\nabla^2\mathbf{v} + \mu\nabla(\nabla\cdot\mathbf{v}) + \nabla\mu\cdot(\nabla\mathbf{v} + \nabla\mathbf{v}^T) \tag{25}$$

laplacian operator, and in rectangular coordinates it takes the form of Eq. (26). *See* LAPLACIAN.

$$\nabla^2\mathbf{v} = \frac{\partial^2\mathbf{v}}{\partial x^2} + \frac{\partial^2\mathbf{v}}{\partial y^2} + \frac{\partial^2\mathbf{v}}{\partial z^2} \tag{26}$$

Navier-Stokes equations. Since the viscosity is a function of temperature and pressure, gradients in the viscosity do exist, and $\nabla\mu$ is nonzero. However, the last term in Eq. (25) is usually small compared to the dominant viscous term $\mu\nabla^2\mathbf{v}$, and can be discarded for all but the most sophisticated analyses. In addition, the term $\mu\nabla(\nabla\cdot\mathbf{v})$ is comparable in magnitude to those terms which result from the second term on the right-hand side of Eq. (24) and can therefore be discarded. This leads to Eq. (27), where the

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{v}\right) = -\nabla p + \rho\mathbf{b} + \mu\nabla^2\mathbf{v} \tag{27}$$

two terms within the parenthesis represent the local acceleration and the convective acceleration, and the three terms on the right-hand side represent the pressure force, body force, and viscous force. These equations are known as the Navier-Stokes equations, and serve as the starting point for the analysis of a wide variety of flows. *See* NAVIER-STOKES EQUATION.

Incompressible flow. Since the density is a function of temperature and pressure, there are gradients in ρ for all flows; however, in many flows the variation in ρ can be neglected. These flows are called incompressible flows, and for such flows, a suitable approximation for the continuity equation given by Eq. (4) is Eq. (28). While there are many flows that

$$\nabla\cdot\mathbf{v} = 0 \tag{28}$$

can be treated as incompressible, that is, the density

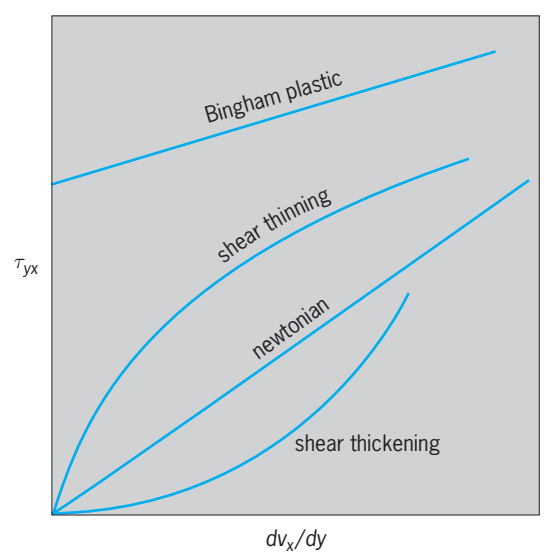

Fig. 4. Behavior of non-newtonian fluids.

can be assigned a constant value, there are no incompressible fluids. Both air and water, for example, are compressible fluids; however, there are many situations for which the flow of both these fluids can be accurately described by Eqs. (27) and (28). In general, this situation occurs when the velocity is small compared to the speed of sound (about 1150 ft/s or 350 m/s for air and about 5200 ft/s or 1600 m/s for water), but there are other situations in which density variations must be considered. The flow that occurs when a pan of water is heated on a stove is an example in which the variation of the density must be taken into account, as is the process of pumping up a bicycle tire. *See* INCOMPRESSIBLE FLOW.

Laminar and turbulent flow. Laminar flow is characterized by the smooth motion of one lamina (or layer) of fluid past another, while turbulent flow is characterized by an irregular and nearly random motion superimposed on the main (or time-averaged) motion of the fluid. The two types of flow can be observed in the heated plume of smoke shown in **Fig. 5**. The smoke originally rises in a smooth laminar stream; however, this buoyancy-driven flow is unstable, and a transition to turbulent flow takes place. This is characterized by an irregular motion as the smoke rises. The transition from laminar to turbulent flow can often be characterized by the Reynolds number, which is a dimensionless group defined by Eq. (29). Here v^* and L^* represent a characteristic

$$\mathrm{Re} = \frac{\rho v^* L^*}{\mu} \tag{29}$$

velocity and characteristic length respectively. The Reynolds number can be thought of as the ratio of inertial effects to viscous effects, but this is a simplistic interpretation, and a more detailed discussion is below. *See* LAMINAR FLOW; REYNOLDS NUMBER; TURBULENT FLOW.

Fig. 5. Laminar and turbulent flow in a smoke plume. (*From S. Whitaker, Introduction to Fluid Mechanics, Krieger Publishing, 1981*)

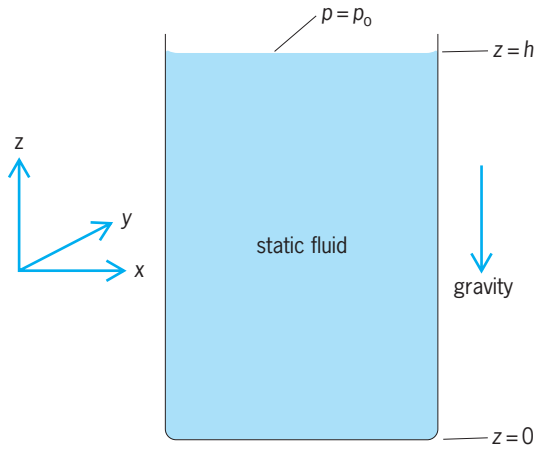

Fig. 6. Pressure variation in a static fluid.

Applications

Some applications of the Navier-Stokes equations to incompressible flows will be considered in order to illustrate the role played by the individual terms in Eq. (27).

Static fluids. When the velocity **v** is zero, Eq. (27) reduces to Eq. (30). If gravity is the only body

$$0 = -\nabla p + \rho \mathbf{b} \tag{30}$$

force under consideration, and the coordinate system shown in **Fig. 6** is used, Eq. (30) reduces to Eq. (31),

$$0 = -\frac{\partial p}{\partial z} - \rho g \tag{31}$$

where g is the acceleration of gravity. At the surface of the liquid the pressure must be equal to the ambient atmospheric pressure p_0, and under these circumstances Eq. (31) requires that the pressure be given by Eq. (32). Here it has been assumed that the

$$p = p_0 + \rho g(h - z) \tag{32}$$

density is constant, and this leads to a linear variation of pressure with position. The situation is not so simple in the atmosphere, since the density of the air depends on both the pressure and the temperature. The hydrostatic variation of the pressure that occurs in the atmosphere is not trivial, and at 10,000 ft (3000 m) above sea level it is responsible

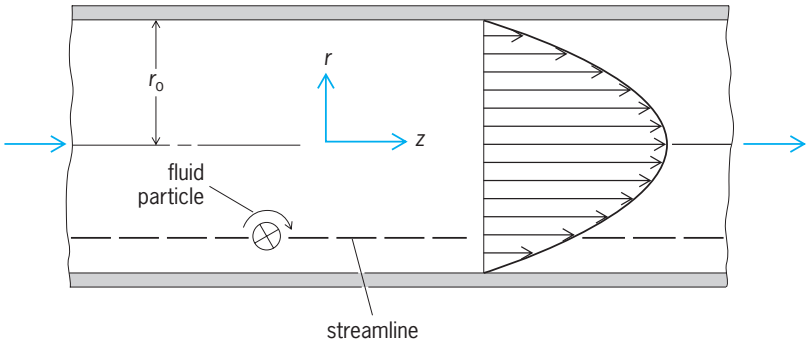

Fig. 7. Uniform flow in a tube.

for a shortness of breath and a lowering of the boiling point of water from 212°F to 193°F (from 100 to 89°C).

Archimedes' principle. The force exerted on an object immersed in a static fluid takes an especially simple form, which is generally attributed to Archimedes: A body is buoyed up by a force equal to the weight of the displaced fluid. The buoyancy force for a constant density fluid is given by Eq. (33),

$$F_{\text{buoy}} = \rho g V \tag{33}$$

where V is the volume of the object and ρ is the density of the fluid. *See* ARCHIMEDES' PRINCIPLE; BUOYANCY; HYDROSTATICS.

Uniform, laminar flow. When steady, laminar flow occurs in a conduit of uniform cross section, the velocity becomes independent of the position along the conduit after a certain distance, referred to as the entrance length. Under these circumstances, Eq. (27) reduces to Eq. (34). For flow in the cylindrical conduit shown in **Fig. 7**, there is only one nonzero component of the velocity vector, and Eq. (34) can be solved to determine the velocity profile, yielding Eq. (35). Here Δp represents the pressure drop that occurs over a distance L, and Eq. (35) indicates that the velocity profile is parabolic, as is illustrated in Fig. 7. If a small amount of dye were injected into a flow of this type, it would trace out a streamline. For this flow, all the streamlines are parallel, and one of these is indicated in Fig. 7. An arbitrarily small element of fluid that moves along this streamline rotates in the manner indicated in Fig. 7. This rotation is caused by the viscous forces, and when viscous effects are small, an irrotational flow is often encountered. When the Reynolds number is greater than 2100, the laminar-flow field becomes unstable and turbulent flow results. The details of the flow are extremely complex, and the irregular motion that accompanies turbulent flow causes the velocity profile to be nearly flat over the central portion of the tube, with very steep velocity gradients near the wall of the tube. *See* PIPE FLOW.

$$0 = -\nabla p + \rho \mathbf{b} + \mu \nabla^2 \mathbf{v} \tag{34}$$

$$v_z = \left(\frac{\Delta p}{L} + \rho b_z\right) \frac{r_0^2}{4\mu} \left[1 - \left(\frac{r}{r_0}\right)^2\right] \tag{35}$$

For turbulent flow, the relation between the volumetric flow rate Q and the pressure drop Δp must be determined experimentally; however, for laminar flow Eq. (35) can be used to obtain Eq. (36). This result is known as the Hagen-Poiseuille law in honor of G. H. L. Hagen and J. L. M. Poiseuille, who experimentally discovered this result in the early 1800s.

$$Q = \frac{\pi r_0^4}{8\mu} \left(\frac{\Delta p}{L} + \rho b_z\right) \tag{36}$$

Boundary-layer flow. Boundary-layer flow is characterized by a large gradient of the velocity in a direction perpendicular to the direction of the main flow, and a small gradient of the velocity in a direction parallel to the direction of the main flow. This type of flow is traditionally discussed in terms of the flow past a thin flat plate, as illustrated in **Fig. 8**, where u_∞ is the uniform fluid velocity (relative to the plate) at large distances from the plate, and $\delta(x)$ is the thickness of the boundary layer. The practical applications of boundary-layer theory have historical roots in the design of airfoils such as that in **Fig. 9**. The gradients of v_x in a boundary-layer flow are characterized by the inequality (37), and this inequality is not valid in the region near the leading edge of the plate, where v_x undergoes rapid changes in the x direction. *See* AIRFOIL.

$$\frac{\partial v_x}{\partial x} \tag{37}$$

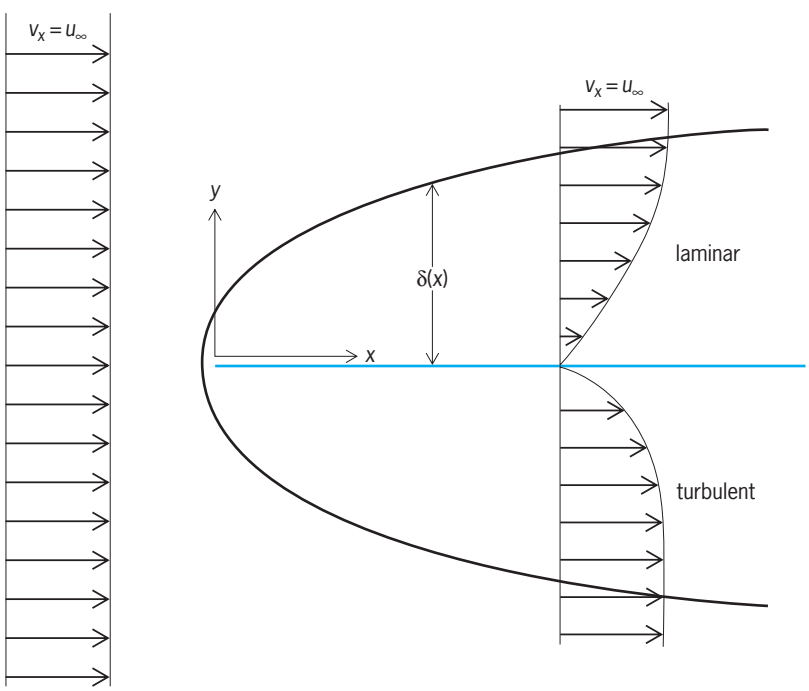

Fig. 8. Boundary-layer flow.

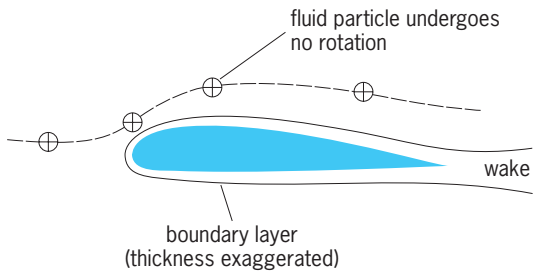

Fig. 9. Diagram of flow past an airfoil.

When the length Reynolds number, given by Eq. (38), is large compared to 1 and the flow is steady,

$$\mathrm{Re}, x = \frac{\rho u_\infty x}{\mu} \tag{38}$$

it is found that the Navier-Stokes equations are characterized by the restrictions of Eq. (39), which in-

$$0 = -\nabla p + \rho \mathbf{b} \tag{39}$$

dicates that the pressure field is hydrostatic, and inequality (40), which indicates that the flow is quasi–

$$v_x \gg v_y \tag{40}$$

one-dimensional. Under these circumstances, the Navier-Stokes equations and the continuity equation take the form of Eqs. (41) and (42). These two equa-

$$\rho\left(v_x \frac{\partial v_x}{\partial x} + v_y \frac{\partial v_x}{\partial y}\right) = \mu\left(\frac{\partial^2 v_x}{\partial y^2}\right) \tag{41}$$

$$\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} = 0 \tag{42}$$

tions are usually referred to as the Prandtl boundary-layer equations in the analysis of flow around airfoils. Equations (41) and (42) can be solved to provide values for the velocity field that are in excellent agreement with experimental measurements. In addition, the theoretical values of the velocity can be used to determine the shear stress exerted by the fluid on the flat plate, according to Eq. (43). The drag force acting

$$\tau_{yx}\Big|_{y=0} = \mu \frac{\partial v_x}{\partial y}\Big|_{y=0} \tag{43}$$

on both sides of the flat plate can then be computed from Eq. (44), in which w is the width of the plate.

$$\text{Drag force} = 2w \int_{x=0}^{x=L} \tau_{yx}\Big|_{y=0} dx \tag{44}$$

The drag force determined by Eq. (44) is also in excellent agreement with experimental values, provided the length Reynolds number is constrained by $10^3 \le \mathrm{Re}, x \le 10^5$. The reason for this constraint is that Re, x must be larger than 10^3 in order for Eq. (41) to be valid, and for values of Re, x larger than 2×10^5 the flow becomes turbulent. This gives rise to the turbulent velocity profile shown in Fig. 8, and the drag force is significantly larger than that for laminar flow.

While laminar boundary-layer theory applies only to a very limited range of the length Reynolds number, the existence of a thin layer in which viscous effects are important has a profound effect on understanding of fluid-mechanical phenomena. Certainly, the most important of these is the D'Alembert paradox, which results from the misconception that as the velocity becomes progressively larger the inertial term (or convective acceleration) in Eq. (27) should become progressively larger than the viscous term. This line of thought is supported by estimates of the type notation (45), which are based largely on dimen-

$$\rho \mathbf{v} \cdot \nabla \mathbf{v} \simeq \frac{\rho v_2}{\text{length}} \mu \nabla^2 \mathbf{v} \simeq \frac{\mu v}{(\text{length})^2} \tag{45}$$

sional arguments. If the inertial and viscous length scales are similar (and they need not be), these estimates lead to Eq. (46), and it is tempting to neglect

$$\frac{|\rho \mathbf{v} \cdot \nabla \mathbf{v}|}{|\mu \nabla^2 \mathbf{v}|} = \frac{\rho v\,(\text{length})}{\mu} = \mathrm{Re} \tag{46}$$

the viscous effects in Eq. (27) when the Reynolds number is very large compared to 1. This leads to Euler's equation (47), and under certain circum-

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v}\right) = -\nabla p + \rho \mathbf{b} \tag{47}$$

stances this result, along with Eq. (28), can provide a reasonable description of the velocity and pressure fields. However, the most interesting aspect of Eq. (47) is that it can be used to prove that the force exerted by a moving fluid on an airfoil such as the one in Fig. 9 consists of nothing but the buoyancy force. This means that if viscous effects were negligible everywhere in the flow field, the lift force on an airfoil would be zero; and this leads to the (correct) conclusion that an airplane could not fly without the aid of viscosity. Clearly, the boundary layer is of great importance in the analysis of aerodynamic flows. *See* BOUNDARY-LAYER FLOW; D'ALEMBERT'S PARADOX.

Inviscid flow. When the Navier-Stokes equations can be approximated by Euler's equations, the flow is often referred to as inviscid, since one can formally progress from Eq. (27) to Eq. (47) by setting the viscocity equal to zero, that is, $\mu = 0$. From a more realistic point of view, there are two conditions that must be met before Eq. (47) can be considered valid: (1) inertial effects are large compared to viscous effects, which is expressed mathematically by inequality (48), and (2) small causes give rise to small effects.

$$|\rho \mathbf{v} \cdot \nabla \mathbf{v}| \gg \left|\mu \nabla^2 \mathbf{v}\right| \tag{48}$$

A reasonably accurate estimate of the magnitude of the inertial term is given by Eq. (49), in which L_ρ

$$\rho \mathbf{v} \cdot \nabla v = \rho v \frac{d\mathbf{v}}{ds} \simeq \frac{\rho v \Delta v}{L_\rho} \tag{49}$$

is the inertial length scale, and Δv represents the change in the magnitude of the velocity vector that occurs over the distance L_ρ along a streamline. For situations of interest, Δv is often the same order as

v, and the estimate takes the form of notation (50).

$$\rho \mathbf{v} \cdot \nabla v \simeq \frac{\rho v^2}{L_\rho} \tag{50}$$

The magnitude of the viscous term is estimated as in notation (51), in which Δv represents the change in

$$\mu \nabla^2 v \simeq \frac{\mu \Delta v}{L_\mu^2} \tag{51}$$

the magnitude of the velocity vector that occurs over the viscous length scale L_μ. Once again, the estimate $\Delta v \simeq v$ can be used, and the ratio of inertial effects to viscous effects takes the form of notation (52).

$$\frac{\text{Inertial effects}}{\text{Viscous effects}} \simeq \frac{\rho v^2/L_\rho}{\mu v/L_\mu^2} = \left(\frac{\rho v L_\mu}{\mu}\right)\left(\frac{L_\mu}{L_\rho}\right) \tag{52}$$

In a laminar boundary layer, $L_\mu = \delta$ and $L_\rho = x$, so that notation (52) takes the form of notation (53).

$$\frac{\text{Inertial effects}}{\text{Viscous effects}} \simeq \left(\frac{\rho u_\infty \delta}{\mu}\right)\frac{\delta}{x} \tag{53}$$

From the Prandtl boundary-layer equations, it follows that viscous and inertial effects are comparable. Thus notation (53) allows an estimate of the boundary-layer thickness as given by notation (54). This is in

$$\delta(x) \simeq \sqrt{\nu x/u_\infty} \tag{54}$$

good agreement with the exact solution obtained from the Prandtl boundary-layer equations, which is given by Eq. (55). In both these results the symbol ν

$$\delta(x) = 4.9\sqrt{\nu x/u_\infty} \tag{55}$$

is used to represent the kinematic viscosity which is given by $\nu = \mu/\rho$.

In order to use notation (52) to decide when viscous effects are negligible, the Reynolds number is defined by Eq. (56) and notation (52) is used to de-

$$\mathrm{Re} = \frac{\rho v L_\mu}{\mu} \tag{56}$$

duce that inertial effects are large compared to viscous effects when inequality (57) is satisfied. When

$$\mathrm{Re}\left(\frac{L_\mu}{L_\rho}\right) \gg 1 \tag{57}$$

this inequality is satisfied and small causes give rise to small effects, (47) can be used along with the continuity equation to determine the velocity and pressure fields.

Irrotational or potential flow. When viscous effects are negligible and the velocity and pressure fields are described by Eqs. (28) and (47), the flow is often irrotational. This means that a different element of fluid moves without any rotational motion, such as is illustrated in Fig. 9 for the streamline that is outside the boundary-layer region. When the flow is irrotational, the velocity can be expressed in terms of a potential function according to Eq. (58). From continuity

$$\mathbf{v} = -\nabla \Phi \tag{58}$$

equation (28) it follows that Φ satisfies the laplacian, which can be expressed as in Eq. (59) for rectangular

$$\nabla^2 \Phi = \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = 0 \tag{59}$$

coordinates. Solutions of Eq. (59) are relatively easy to obtain for a variety of situations; thus the inviscid-flow approximation allows for the determination of complex flow fields with little mathematical effort. For the flow past an airflow shown in Fig. 9, viscous forces are important in the boundary layer, and the Prandtl boundary-layer equations can be used to analyze the flow. Outside the boundary layer, viscous effects can generally be neglected, and a solution to Eq. (59) provides the velocity field. By matching the two velocity fields at the outer edge of the boundary layer, the complete solution for the flow past an airfoil can be obtained; however, the process is not simple, and it requires considerable mathematical skill and physical insight. *See* LAPLACE'S IRROTATIONAL MOTION.

Bernoulli's equation. When the flow is steady and incompressible, and viscous effects are negligible, a particularly simple form of the Navier-Stokes equations is available. The above restrictions make it possible to write Eq. (27) as Eq. (60). If the body force

$$\rho \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla p + \rho \mathbf{b} \tag{60}$$

$\mathbf{b}$ represents a conservative force field, it can be expressed in terms of the gradient of a scalar according to Eq. (61). When $\mathbf{b}$ represents the gravity vector, ϕ

$$\mathbf{b} = -\nabla \phi \tag{61}$$

is called the gravitational potential function, and for the coordinate system shown in Fig. 6, ϕ is given to within an arbitrary constant by Eq. (62). Since the flow is incompressible, Eq. (63) is valid and Eq. (60) takes the form of Eq. (64). The velocity vector can be

$$\phi = gz + \text{constant} \tag{62}$$

$$\rho \mathbf{b} = -\rho \nabla \phi = -\nabla(\rho \phi) \tag{63}$$

$$\rho \mathbf{v} \cdot \nabla \mathbf{v} = -\nabla(p + \rho \phi) \tag{64}$$

expressed in terms of the magnitude v and the unit vector $\mathbf{l}$ illustrated in **Fig. 10**, so that the left-hand side of Eq. (64) can be expressed as Eq. (65). Here

$$\rho \mathbf{v} \cdot \nabla \mathbf{v} = \rho v \lambda \cdot \nabla \mathbf{v} = \rho v \frac{d\mathbf{v}}{ds} \tag{65}$$

s represents the arc length measured along any one

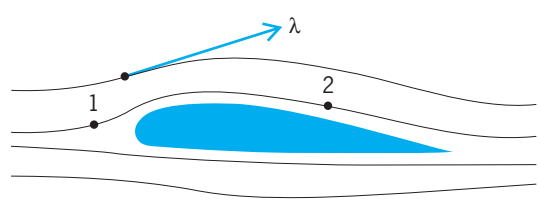

Fig. 10. Streamlines around an airfoil.

of the streamlines shown in Fig. 10, and the definition of the directional derivative given by Eq. (66)

$$\lambda \cdot \nabla = \frac{d}{ds} \tag{66}$$

has been used. Substitution of Eq. (65) into Eq. (64) yields a form that does not appear to be suitable for problem solving, Eq. (67). However, the scalar prod-

$$\rho v \frac{d\mathbf{v}}{ds} = -\nabla(p + \rho\phi) \tag{67}$$

uct (or dot product) of this equation with the unit tangent vector is Eq. (68), and a little analysis leads

$$\rho v \lambda \cdot \frac{d\mathbf{v}}{ds} = -\lambda \cdot \nabla(p + \rho\phi) \tag{68}$$

to Eq. (69). This indicates that the quantity $^1/_2\rho v^2 +$

$$\frac{d}{ds}\left({}^1/_2\rho v^2 + p + \rho\phi \right) = 0 \tag{69}$$

$p + \rho\phi$ is a constant along a streamline, and Eq. (62) can be used to express this idea as Eq. (70). This is

$$^1/_2\rho v^2 + p + \rho g z = C \tag{70}$$

known as Bernoulli's equation. For the two points identified on a streamline shown in Fig. 10, Eq. (70) can be used to derive Eq. (71). For irrotational flows,

$$^1/_2\rho v_2^2 + p_2 + \rho g z_2 = {}^1/_2\rho v_1^2 + p_1 + \rho g z_1 \tag{71}$$

it can be shown that the constant in Eq. (70) is the same for all streamlines, and most flows for which viscous effects are negligible fall into this category. Under these circumstances, Eqs. (58) and (59) are used to determine the velocity field, and Eq. (70) to determine the pressure field. *See* BERNOULLI'S THEOREM; FLUID MECHANICS; GAS DYNAMICS; HYDRODYNAMICS.

Stephen Whitaker

Bibliography. H. Aksel and O. C. Eralp, *Gas Dynamics*, 1993; I. G. Currie, *Fundamental Mechanics of Fluids*, 2d ed., 1993; H. Schlichting and J. Kestin, *Boundary Layer Theory*, 8th ed., 2000; W. G. Vincenti and C. H. Kruger, Jr., *Introduction to Physical Gas Dynamics*, 1965, reprint 1975; S. Whitaker, *Fundamental Principles of Heat Transfer*, 1977, reprint 1983; S. Whitaker, *Introduction to Fluid Mechanics*, 1968, reprint 1992.

Fluid mechanics

The engineering science concerned with the relation between the forces acting on fluids (liquids and gases) and their motions, and with the forces caused by fluids on their surroundings. A fluid subjected to unbalanced forces or stresses continues in motion as long as unbalanced forces are applied. For example, in the pouring of water from a pitcher the water velocity is very high over the lip, moderately high approaching the lip, and very low near the bottom of the pitcher. The unbalanced force is gravity, that is, the weight of the tilted water particles near the surface. The flow continues as long as water is available and the pitcher remains tilted.

In contrast, a solid deforms under forces and stresses and simply assumes a new deformed position and remains static. It does not flow.

Laypersons instinctively recognize this distinction between a flowing liquid and a statically deformed solid. In fact, not only liquids but vapors and gases are defined as fluids, because of this characteristic of continuous deformation. A fluid cannot withstand an applied shear stress, however small, without moving. Solids remain static because they resist shear. *See* FLUIDS; STRENGTH OF MATERIALS.

Types of fluids. A fluid may be a liquid, vapor, or gas. In thermodynamics the three states of matter are solid, liquid, and vapor. The term vapor denotes a gaseous substance interacting with its own liquid phase, for example, steam above water. If this phase interaction is not important, the vapor is simply termed a gas. *See* THERMODYNAMIC PROCESSES; VAPOR PRESSURE.

Gases have weak intermolecular forces and expand to fill any container. Left free, gases expand and form the atmosphere of the Earth. Gases are highly compressible; doubling the pressure at constant temperature doubles the density. *See* GAS.

Liquids, in contrast, have strong intermolecular forces and tend to retain constant volume. Placed in a container, a liquid occupies only its own volume and forms a free surface which is at the same pressure as any gas touching it. Liquids are nearly incompressible; doubling the pressure of water at room temperature, for example, increases its density by only 0.005%. *See* OPEN CHANNEL.

Liquids and vapors can flow together as a mixture, such as steam condensing in a pipe flow with cold walls. This constitutes a special branch of fluid mechanics, covering two-phase-flow. *See* CAVITATION.

Density and other properties. The physical properties of a fluid are essential to formulating theories and developing designs for fluid flow. Especially important are pressure, density and temperature.

The density of a fluid is its mass per unit volume, denoted as ρ. The most common units for density are kg/m^3 in the International System (SI) and lb/ft^3 or slugs/ft^3 in the English systems. Liquids are typically hundreds of times more dense than gases; at room temperature and pressure, $\rho = 998$ kg/m^3 (62.4 lb/ft^3 or 1.94 slugs/ft^3) for water and only 1.20 kg/m^3 (0.0075 lb/ft^3 or 0.00233 slug/ft^3) for air. Fluid density is independent of location, whether on the Earth or the Moon.

In practical problems, the specific weight of a fluid or weight per unit volume is often used, denoted as $\gamma = \rho g$, with units of newton per cubic meter (N/m^3; SI) or pound-force per cubic foot (lbf/ft^3; English). The quantity g is the acceleration of gravity, with an average value on Earth of 9.81 m/s^2 (32.2 ft/s^2). Thus, for water at standard conditions, $\gamma = 9790$ N/m^3 (62.4 lbf/ft^3). *See* DENSITY.

The pressure of a fluid, denoted as p, is the normal force per unit area exerted on any small elemental

surface in the fluid. Standard atmospheric pressure is 101,325 N/m^2 or pascals (2116 lbf/ft^2). Differences in pressure, or pressure gradients, often drive fluid flows. *See* PRESSURE.

The temperature of a fluid, denoted as T, expresses the average energy of the fluid molecules. It may be stated in convenient units (°F or °C) or in absolute units (K or °R). Standard room temperature is 20°C (293 K) or 68°F (528°R). *See* CARBOHYDRATE; TEMPERATURE.

Ideal gas. For all fluids, p, ρ, and T are related by the thermodynamic state relation for the substance. For liquids, which are nearly incompressible, ρ is assumed to be constant, independent of p or T. For gases, a convenient approximation is the ideal gas law, given in Eq. (1), where R is the gas constant. For

$$p = \rho RT \tag{1}$$

any particular gas, R is computed by dividing a universal physical constant, 8315 (m^2)/(s^2 · mole · K), by the molecular weight of the gas. For air, with an average molecular weight of 29, $R = 8315/29 = 287$ (m^2)/(s^2 · K). In Eq. (1), T must have absolute units.

The ideal gas law becomes inaccurate at very high temperatures, where chemical reactions may occur, or at lower temperatures where the gas is near its saturated liquid state. The latter case is illustrated by steam, which is near the liquid state and sufficiently nonideal for extensive tables of its (p, ρ, T) relations to have been published for engineers and scientists. *See* STEAM.

Hydrostatic condition. Since shear stresses cause motion in a fluid and result in differences in normal stresses at a point, it follows that a fluid at rest must have zero shear and uniform pressure at a point. This is the hydrostatic condition. The fluid pressure increases deeper in the fluid to balance the increased weight of fluid above. For liquids, and for gases over short vertical distances, the fluid density can be assumed to be constant. Then the pressure increases linearly with depth. If point 2 is vertically below point 1 at distance h, the two pressures are related by hydrostatic formula (2). This formula is illustrated

$$p_2 = p_1 + \rho g h = p_1 + \gamma h \tag{2}$$

in **Fig. 1** for air and water, whose densities must be inserted into Eq. (2). If the liquid volume is limited, it forms a free surface, as in Fig. 1, which is level or horizontal and at uniform pressure.

For gases over large changes in height, such as the atmosphere, Eq. (2) must be modified for varying density and results in a nonlinear formula.

The linearly varying vertical change of hydrostatic pressure in Eq. (2) gives simple results for overall pressure on a plane surface. Integration of (pressure) × (area) over the surface gives Eq. (3) for the

$$F_{\text{plane}} = p_{cg} A_{\text{plane}} \tag{3}$$

force on a plane area A_{plane} of any shape, where p_{cg} denotes the fluid pressure at the centroid of the plane

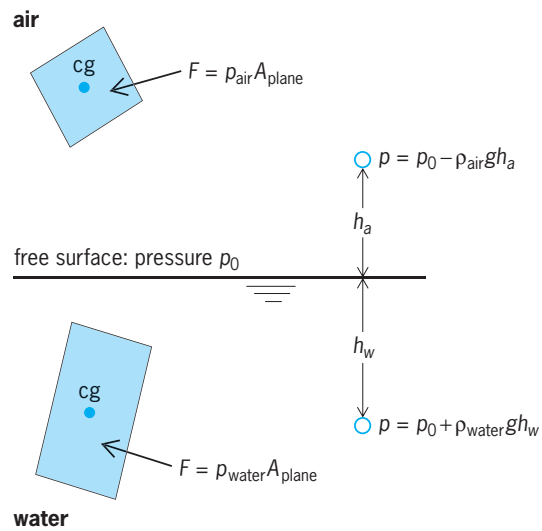

Fig. 1. Hydrostatic fluid properties. cg = centroid of the plane.

surface, as determined by Eq. (2). Equation (3) is also illustrated in Fig. 1. Formulas such as Eq. (3) are used to design submerged surfaces, including dams and submarine hulls, to withstand large hydrostatic forces.

When the pressure forces are integrated over a closed three-dimensional surface (an immersed body), Archimedes' two famous laws of buoyancy result: (1) A body immersed in a fluid experiences a vertical buoyant force equal to the weight of the fluid it displaces. (2) A floating body displaces its own weight in the fluid in which it floats. These principles are widely used in applied fluid-flow problems such as the design of ships and underwater vehicles. *See* ARCHIMEDES' PRINCIPLE; HYDROSTATICS.

Viscosity. When a fluid is subjected to shear stress, it flows and resists the shear through molecular momentum transfer. The macroscopic effect of this molecular action, for most common fluids, is the physical property called viscosity. Shear stress results in a gradient in fluid velocity; the converse is also true.

Figure 2 shows a typical case of a fluid flow whose x-directed velocity u varies with transverse distance y, $u = u(y)$. This convenient scheme of showing a row of velocity vectors with arrowheads connected by a continuous curve is called a velocity profile. At any point in the fluid, the slope of the velocity profile is a measure of local shear stress τ. For all gases and vapors and for the common liquids (water, gasoline, oil), slope and shear are proportional as in Eq. (4).

$$\tau = \mu \frac{du}{dy} \tag{4}$$

$$\mu = \mu(p, T)$$

The coefficient μ is called the viscosity of the fluid, with units of kg/(m · s), lb/(ft · s), or slug/(ft · s). The viscosity is a physical property of a fluid and is determined experimentally, for example, by shearing

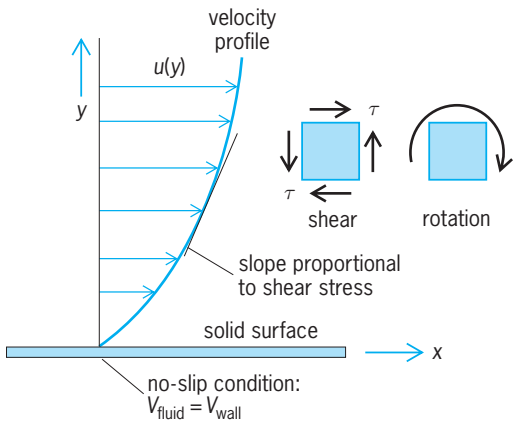

Fig. 2. No-slip condition and shear stress set up by velocity gradients.

the fluid between two plates or two concentric cylinders.

Gases have low viscosity, light liquids such as water and mercury are in the medium range, and oils have very high μ. At 20°C (68°F) and 1 atm (101 kPa), three representative examples of viscosity are air [1.8×10^{-5} kg/(m · s)], water [1.0×10^{-3} kg/(m · s)], and glycerin [1.5 kg/(m · s)]. The viscosity of liquids decreases with temperature, whereas that of gases increases. The very slight increase of fluid viscosity with pressure is usually neglected.

The common fluids for which the linear relationship of Eq. (4) holds are called newtonian viscous fluids. More complex fluids, such as paints, pastes, greases, and slurries, exhibit nonlinear or nonnewtonian behavior and require more complex theories to account for their behavior. *See* NEWTONIAN FLUID; NON-NEWTONIAN FLUID; VISCOSITY.

No-slip condition. A common characteristic of all fluids, whether newtonian or not, is that they do not slip at a solid boundary. No matter how fast they flow away from the boundary, fluid particles at a solid surface become entrapped by the surface structure. The macroscopic effect is that the fluid velocity equals the solid velocity at a boundary. This is called the no-slip condition, illustrated in Fig. 2, where the solid is fixed, so that the fluid velocity drops to zero there: $u(0) = 0$. No-slip sets up a slow-moving shear layer or boundary layer when fluid flows near a solid surface, as shown in Fig. 2. The theory of boundary-layer flow is well developed and explains many effects involving viscous flow past immersed bodies or within passages. *See* BOUNDARY-LAYER FLOW.

Kinematic viscosity. Fluid-flow problems often involve the ratio of viscosity to density, called the kinematic viscosity ν given in Eq. (5). The mass units

$$\nu = \frac{\mu}{\rho} \tag{5}$$

cancel (one meaning of kinematic is "independent of mass") and the units of ν are m²/s or ft²/s.

Although gases have low viscosity, they also have low density and their kinematic viscosity is thus often higher than for common liquids. For example, at 20°C (68°F) and 1 atm (101 kPa) air has a viscosity $\nu = 1.51 \times 10^{-5}$ m²/s, whereas for water $\nu = 1.01 \times 10^{-6}$ m²/s, or 15 times lower. As a crude estimate, this makes water 15 times more likely than air to undergo low-viscosity effects such as transition to turbulence.

Compressibility and speed of sound. All fluids are at least slightly compressible, that is, their density increases as pressure is applied. In many flows, however, compressibility effects may be neglected.

Fluid compressibility can be characterized by the isentropic bulk modulus K and the speed of sound a. The bulk modulus, given in Eq. (6), is the change

$$K = -v\left(\frac{\delta p}{\delta v_s}\right) = \rho\left(\frac{\delta p}{\delta \rho_s}\right) \tag{6}$$

in pressure for a fractional change in fluid density for an isentropic process, where v is the specific volume and s the fluid entropy. A large value of K indicates a fluid which is difficult to compress. For example, at 20°C and 1 atm (101 kPa), $K = 2.19 \times 10^9$ N/m² (4.57×10^7 lbf/ft³) for water and only 1.42×10^5 N/m² (2.97×10^3 lbf/ft²) for air.

The speed of sound a is defined as the velocity of propagation of a small pressure disturbance through the medium. Application of conservation of mass, momentum, and energy to such a disturbance results in Eq. (7). Using data given above for K and ρ at 20°C

$$a = \left(\frac{K}{\rho}\right)^{1/2} = \left(\frac{\delta p}{\delta \rho}\right)_s^{1/2} \tag{7}$$

and 1 atm (101 kPa) gives $a \simeq 1480$ m/s (4860 ft/s) for water and $a \simeq 340$ m/s (1117 ft/s) for air. For an ideal gas, Eq. (7) may be written, with the help of Eq. (1), as Eq. (8), where $k = c_p/c_v$ is the ratio of

$$a = (kRT)^{1/2} \tag{8}$$

specific heats and is approximately 1.4 for air. *See* SOUND.

A very important parameter in determining compressibility effects is the Mach number, or ratio of flow velocity V to fluid speed of sound, given by Eq. (9). For subsonic flow, Ma < 1, whereas for supersonic flow, Ma > 1. The flow is essentially incompressible if Ma < 0.3; hence for air the flow velocity is less than about 100 m/s (330 ft/s). Almost all liquid flows and many gas flows are thus treated as incompressible. Even a supersonic airplane lands and takes off in the incompressible regime. *See* MACH NUMBER.

$$\mathrm{Ma} = \frac{V}{a} \tag{9}$$

A problem involving compressibility of liquids arises when a valve is suddenly opened or closed in a piping system. Strong pressure waves may be propagated through the pipes and cause serious damage. *See* COMPRESSIBLE FLOW; WATER HAMMER.

Ideal incompressible flow. Flow analysis can be greatly simplified by neglecting friction and compressibility and assuming that the flow is steady or varying only in space. The resulting equations of motion of ideal flow may be solved analytically for many

problems. The velocity field $\mathbf{V}(x, y, z)$ may be plotted in various ways to reveal the flow pattern, and the pressure field $p(x, y, z)$ is directly related to velocity and local elevation by Bernoulli's equation for incompressible, frictionless, steady flow, Eq. (10), where c

$$\frac{p}{\rho} + \frac{V^2}{2} + gz = c \tag{10}$$

is a constant along a streamline and z is the vertical height of any point in the flow. A streamline is defined as a line everywhere parallel to the local velocity vector. If velocity and pressure are known, the force may be computed on any boundary or internal surface in the flow.

Equation (10) is a weak form of the first law of thermodynamics, balancing the work of pressure forces with changes in kinetic and potential energy. Bernoulli's equation, though widely used, is thus inaccurate if there are significant heat-transfer effects, shaft work, or viscous losses. *See* BERNOULLI'S THEOREM.

Flow visualization. Fluid flow is highly visual, and many photographs and sketchs of flow patterns have been published. Movies and video tapes have been made showing flow phenomena. Flow patterns are shown as streamlines, which are everywhere tangent to the local velocity vector; pathlines, which trace the trajectory of a single particle; and streaklines, which are the locus of particles issuing from a single point. Streamlines arise naturally in the mathematical theory of flow, whereas pathlines and streaklines arise in experimental studies using smoke, bubbles, dyes, or neutral particles. In steady fluid flow the three types of lines are identical.

Figure 3 shows the streamline pattern for ideal frictionless flow past a circular cylinder. The lines provide a qualitative feeling for the flow. Where they are equally spaced at the far left and far right, the velocity is uniform. Where they are widely spaced near the left and right of the cylinder, the velocity is low and the pressure high. Where they are closely packed above and below the cylinder, the velocity is high and the pressure low. Figure 3 has some inaccuracies due to viscous effects.

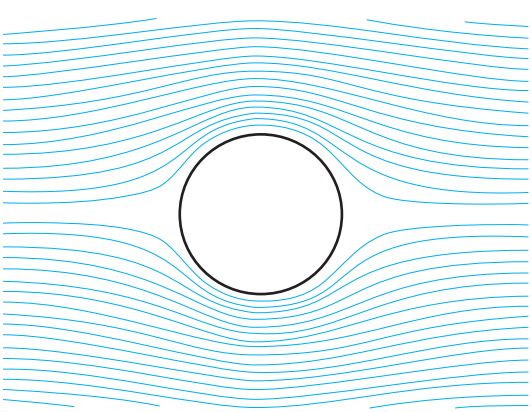

Fig. 3. Streamlines for uniform flow of a frictionless, incompressible fluid past a circular cylinder. This idealized flow pattern exhibits double symmetry.

Viscous flow effects. The assumption of frictionless flow is an idealization which may be inapplicable to the actual flow of a real or viscous fluid. In particular, where solid walls are present, the no-slip condition, as in Fig. 2, slows down the fluid near the wall and creates a boundary layer where shear effects are important. This viscous layer generally grows in thickness as it moves downstream along the wall. It may ultimately encompass a large portion of the downstream flow pattern.

In the case of flow into a pipe or duct, the shear layers grow and meet in the pipe center and thereafter fill the pipe with viscous effects. Bernoulli relation (10) is readily modified for pipe flow by adding a friction-loss term. If points 1 and 2 are upstream and downstream in the pipe flow, respectively, the modified Bernoulli relation is given by Eq. (11), where h_f

$$\frac{p_1}{\rho} + \frac{V_1^2}{2} + gz_1 = \frac{p_2}{\rho} + \frac{V_2^2}{2} + gz_2 + gh_f \tag{11}$$

is a loss term expressed in length units (m or ft). Equation (11) shows that the total flow energy decreases downstream because of friction losses. The loss term may be correlated with flow velocity, pipe size, and physical properties of the fluid. Equation (11) may not be used if there are significant heat-transfer or shaft-work effects.

In the case of flow past immersed bodies, as in Fig. 3, viscous effects are variable and involve more than just a loss term. The flow pattern is strongly dependent upon the dimensionless parameter called the Reynolds number, given by Eq. (12), where V is

$$\mathrm{Re}_D = \frac{\rho V D}{\mu} \tag{12}$$

the stream velocity and D a suitable length scale, in this case the cylinder diameter. *See* REYNOLDS NUMBER.

Figure 4 shows the real-fluid flow past a circular cylinder at a Reynolds number of 10^5. The boundary layer on the front of the body is thin, and outside this layer the flow is approximately frictionless and satisfies Bernoulli relation (10). However, as it rounds the shoulder of the body, the low-velocity boundary layer is unable to resist the rising pressure toward the rear of the body. It separates from the body and forms a broad, pulsating wake which produces a large downstream force on the body, called drag. This is a sharp contrast to the ideal theory of Fig. 3, which predicts zero drag and was thought in the nineteenth century to be a zero paradox, although actually it is simply the incorrect result of neglecting viscous effects. *See* D'ALEMBERT'S PARADOX; WAKE FLOW.

Figure 4 is a striking example of the importance of viscous effects, but other flows are not so strongly modified. Ideal-flow theory has been used to predict such phenomena as airfoil and cascade forces, flow through orifices, and surface-wave motions, and is a useful first approximation in flow analysis.

Laminar and turbulent flow. An interesting aspect of fully viscous flow within the boundary layer or inside a long duct is the existence of two types of flow.

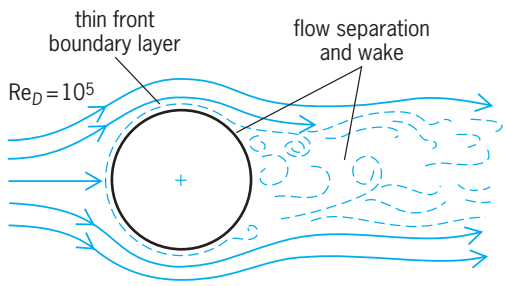

Fig. 4. Actual flow pattern for viscous flow past a circular cylinder at a moderately high Reynolds number. The outer stream is grossly perturbed by broad flow separation and wake. (*After F. M. White, Fluid Mechanics, 2d ed., McGraw-Hill, 1986*)

Although flow instrumentation was relatively crude, it was known in the early nineteenth century that flow could be either steady and orderly or fluctuating and disorderly. In 1883, O. Reynolds discovered that the controlling parameter is the Reynolds number from Eq. (12). By introducing a dye streak into a pipe flow, Reynolds was able to visually demonstrate the two flow regimes.

Figure 5 shows Reynold's sketches of the two types of flow. At low Re_D, the dye streak remains straight and smooth (Fig. 5*a*), as if the fluid particles remained in distinct laminae which merely sheared each other; this regime is now called laminar flow. At higher Re_D, the dye streak seems to grow and burst and fill the pipe with color (Fig. 5*b*). Reynold's spark photograph of this condition (Fig. 5*c*) reveals multiple eddies and fast-changing whorls in the flow; this second regime is now called turbulent flow.

Numerous pipe-flow experiments verify that the flow is laminar until Re_D reaches a critical or transition value of about 2000. The transition range, intermittently laminar and turbulent, is from 2000 to 4000. For $Re_D > 4000$, the flow becomes fully turbulent.

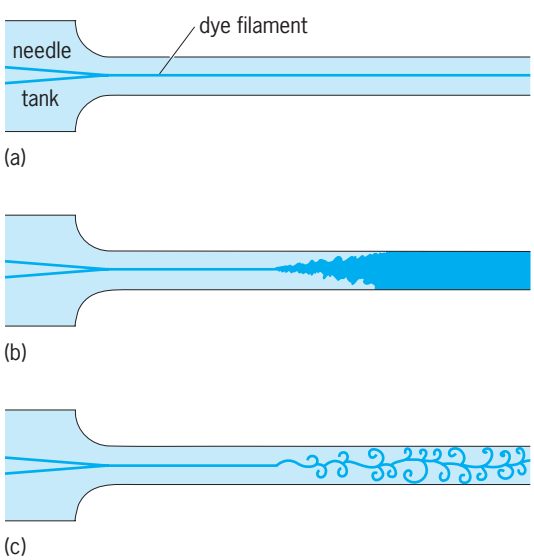

Fig. 5. Pipe-flow transition experiment. (*a*) Low Re_D, laminar flow. (*b*) High Re_D, turbulent flow. (*c*) Spark photograph of turbulent flow condition. (*After O. Reynolds, an experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous and of the law of resistance in parallel channels, Phil. Trans. Roy. Soc., London, A174:935–982, 1883*)

This same laminar-transition-turbulence sequence occurs for all types of viscous flows. The critical Reynolds number depends upon the geometry and the presence of disturbances which might trigger turbulence, such as rough walls, noisy pumps, loudspeakers, and heat-transfer effects. *See* LAMINAR FLOW; TURBULENT FLOW.

Pipe flow. Flow in a pipe is a good example of laminar versus turbulent behavior. Since the 1940s, instrumentation has been developed, such as hot wires, hot films, tiny flush-mounted piezometers, and laser-Doppler meters, which measure all the details of fluctuating turbulent velocities, temperatures, and pressures in a flow. *See* FLOW MEASUREMENT.

One of the most striking differences between laminar and turbulent flow is the shape of the mean velocity profiles (**Fig. 6**). The laminar profile (Fig. 6*a*)

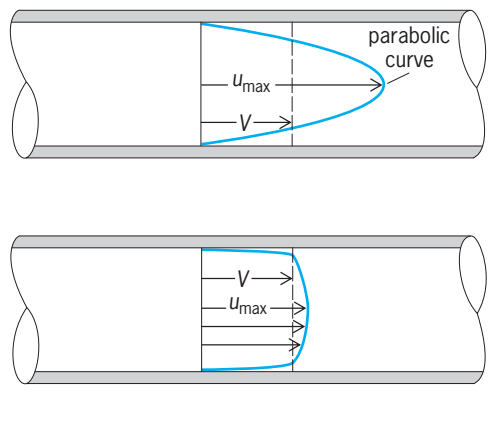

Fig. 6. Comparison of (*a*) laminar and (*b*) turbulent pipe flows at the same volume flow rate (m^3/s). V = average flow velocity; u_{max} = maximum flow velocity. (*After F. M. White, Fluid Mechanics, 2d ed., McGraw-Hill, 1986*)

is a paraboloid shape which can be predicted mathematically. Its mean value V is exactly one-half its maximum value u_{max}, and the friction loss may be exactly predicted by Eq. (13) from Eq. (11). In contrast,

$$h_{f,\text{laminar}} = \frac{32\mu LV}{\rho g D^2} \tag{13}$$

the turbulent profile (Fig. 6*b*) is flat, with extremely steep slopes at the wall. Its average velocity varies from 81 to 89% of the maximum velocity, depending upon Re_D. Its friction loss also varies with Re_D several times larger than suggested by Eq. (13), and is predicted not by theory but by empirical correlations. *See* PIPE FLOW.

Generally speaking, laminar flows can be predicted analytically and modeled accurately with digital computer simulations, whereas turbulent flows require empirical correlations and other approximate modeling tools, which make turbulence predictions, though improving with research, uncertain and lacking in generality.

Compressible flow. As the Mach number of a flow increases beyond 0.3, compressibility becomes important. The density begins to change significantly and is coupled to temperature and pressure changes. Mass and momentum conservation must be combined with the laws of thermodynamics and a state equation for the gas.

Energy relation (11) must be modified to account for internal energy changes, heat transfer, and shaft and viscous work. The result is given in Eq. (14).

$$h_1 + \frac{V_1^2}{2} + gz_1 = h_2 + \frac{V_2^2}{2} + gz_2 - q + w_s + w_v \qquad (14)$$

Here, q is the heat per unit mass transferred to the fluid, and w_s and w_v are the shaft work and viscous work, respectively, per unit mass done by the fluid, and h is the fluid enthalpy per unit mass, defined by Eq. (15), where e is the internal energy per unit

$$h = e + \frac{p}{\rho} \qquad (15)$$

mass. *See* ENTHALPY; HEAT; INTERNAL ENERGY.

The most common case is compressible flow in a duct, where viscous work is zero because of the no-slip condition. Assuming adiabatic flow, no shaft work, and negligible height changes, Eq. (14) reduces to Eq. (16), where c is a constant.

$$h_1 + \frac{V_1^2}{2} = h_2 + \frac{V_2^2}{2} = c \qquad (16)$$

See ISENTROPIC FLOW.

Equation (16) does not require frictionless flow. If Eq. (16) is combined with the ideal gas law, Eq. (1), and the definition of the Mach number, Eq. (9), the result in Eq. (17) relates temperature at any point to

$$T = \frac{T_0}{1 + 0.5(k-1)\mathrm{Ma}^2} \qquad (17)$$

the stagnation temperature T_0 at a point where the velocity is zero.

If frictionless flow is assumed, the motion is at constant entropy and the local pressure may be related to stagnation pressure, p_0, as in Eq. (18). Thus, in adi-

$$p = \frac{p_0}{\left[1 + 0.5\,(k-1)\,\mathrm{Ma}^2\right]^{k/(k-1)}} \qquad (18)$$

abatic frictionless duct flow, both the fluid temperature and pressure decrease significantly with Mach number.

A converging-diverging duct is shown in **Fig. 7**. For steady one-dimensional flow, the mass conservation relation is given in Eq. (19), where A is the area

$$\rho_1 A_1 V_1 = \rho_2 A_2 V_2 = \text{constant} \qquad (19)$$

of the duct. If the duct is fed by a source pressure p_0 and the receiver pressure p_e is gradually reduced, the throat pressure falls until the throat Mach number reaches unity, from which, for $k = 1.4$, Eq. (18) predicts that $p = 0.53p_0$. The throat is sonic, and

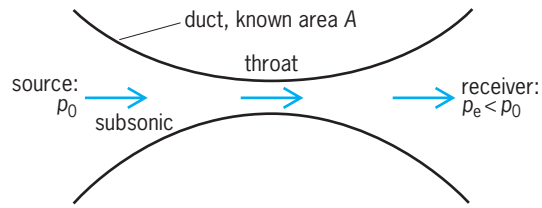

Fig. 7. Compressible duct flow through a converging-diverging nozzle. As receiver pressure drops, the throat becomes sonic and the downstream flow supersonic.

any further reduction in p_e fails to change the flow rate, because signals can no longer be propagated upstream from the throat. The nozzle is said to be choked, and the flow downstream from the throat becomes supersonic. This interesting behavior is nearly impossible in liquid flow but is easily accomplished in gases by moderate changes in boundary pressures. Other aspects of compressible flow are the formation of shock waves and the behavior of vehicles moving faster than sound. *See* GAS DYNAMICS; NOZZLE; SHOCK WAVE; SUPERSONIC FLIGHT.

Boundary-layer theory. As shown in Fig. 2, the no-slip condition creates a slow-moving boundary layer when a flow is near a solid surface. Both shear and rotation (local fluid angular velocity) within the boundary layer are large at the surface and gradually decrease to zero in the outer, nearly inviscid, stream. The boundary layer is very thin if the Reynolds number is above 10^4. A similar thermal boundary layer exists if the wall is hot or cold compared to the stream flow.

Boundary layers are laminar near the leading edge of the surface and undergo transition to turbulence as they grow thicker downstream. The theory is well developed and predicts both laminar and turbulent velocities as well as temperatures plus wall shear stress and heat transfer. The theory also predicts flow separation on the rear (or rising-pressure) region of bodies, as in Fig. 4. After the separation point, boundary-layer theory is generally no longer accurate. *See* FLUID-FLOW PRINCIPLES; HYDRODYNAMICS; RAREFIED GAS FLOW.

Frank M. White

Bibliography. R. W. Fox, A. T. McDonald, and P. J. Pritchard, *Introduction to Fluid Mechanics*, 6th ed., 2003; *Illustrated Experiments in Fluid, Mechanics*, NCFMF Book of Film Notes, 1972; I. H. Shames, *Mechanics of Fluids*, 4th ed., 2003; V. L. Streeter, E. S. Wylie, and K. W. Bedford, *Fluid Mechanics*, 9th ed., 1998; M. Van Dyke, *An Album of Fluid Motion*, 1982; F. M. White, *Fluid Mechanics*, 5th ed., 2003.

Fluidization

The processing technique employing a suspension or fluidization of small solid particles in a vertically rising stream of fluid—usually gas—so that fluid and solid come into intimate contact. This is a tool with many applications in the petroleum and chemical process industries. There are major possibilities of

even greater growth and utility in the future. Suspensions of solid particles by vertically rising liquid streams are of lesser interest in modern processing, but have been shown to be of use, particularly in liquid contacting of ion-exchange resins. However, they come in this same classification and their use involves techniques of liquid settling, both free and hindered (sedimentation), classification, and density flotation. *See* ION EXCHANGE; MECHANICAL CLASSIFICATION.

Properties of fluidized systems. The interrelations of hydromechanics, heat transfer, and mass transfer in the gas-fluidized bed involve a very large number of factors which must be understood for mathematical and physical analysis and process design. Because of the excellent contacting under these conditions, numerous chemical reactions are also possible—either between solid and gas, two fluidized solids with each other or with the gas, or most important, one or more gases in a mixture with the solid as a catalyst. In the usual case, the practical applications in multimillion-dollar plants have far outrun the exact understanding of the physical, and often chemical, interplay of variables within the minute ranges of each of the small particles and the surrounding gas phase. Accordingly, much of the continuing development of processes and plants is of an empirical nature, rather than being based on exact and calculated design.

The fluidized bed results when a fluid, usually a gas, flows upward through a bed of suitably sized, solid particles at a velocity sufficiently high to buoy the particles, to overcome the influence of gravity, and to impart to them an appearance of great turbulence, similar to that of a violently boiling liquid. Fluid velocities must be intermediate between that which would lift the particles to maintain a uniform suspension and that which would sweep the particles out of the container. The fluidized bed is in a relatively stable condition of vigorous contacting of fluid and solids, with a lower boundary at the point of fluid inlet and an upper definite and clearly marked boundary surface at which the gas disengages itself from the system. The upper boundary usually goes up or down with increase or decrease of the gas velocity.

Particle sizes are often in the range of 30–125 micrometers, but they may vary to $^1/_4$ in. (6 mm) or more. Superficial gas velocities may range 0.02–1.0 ft/s (0.006–0.3 m/s), depending on factors such as the relative densities of the gas and solid, the size and shape of solid particles, and the number of particles per cubic foot (bed density) desired.

Conditions within the bed are intermediate between that of a packed column and that of a pneumatic transporter of the particles. In pneumatic transport, the slippage of the gas past the solid particles is often relatively small, and the components of friction and inertia carry the solid particle along. In the fluidized bed, the slippage of gas past the solid particle is sufficient to balance the gravitational effect or weight, but is not sufficient to supply a continuing velocity.

An increase in fluid velocity increases the lift or upward drag on the particles. This causes them to rise, thereby increasing the bed voidage and the effective cross section available for gas passage and also decreasing the number of particles per unit volume and the density of the bed as a whole. This decreases the interstitial velocity to reduce the drag on the individual particles, which then settle somewhat until the gravitation and the fluid-dynamic forces on the particles ultimately come into balance. Further increase in fluid velocity causes further bed expansion until the bed passes beyond the limits of the containing vessel and the particles are transported.

There is excellent contact between the solid particles themselves and the particles with the gas. If there is a temperature difference, there is an excellent flow of heat between the particles themselves and between the particles and the gas. Particularly if heat is absorbed, as in drying of a solid by a hot gas, or if heat is given off because of chemical reaction, heat transfers to equate the temperatures of the gas and the solid very quickly.

Similarly, there is a major opportunity for mass transfer by molecular diffusion to or from the solid particles and between the solid particles and the gas phase. Examples are the movement of reactants in the gaseous phase toward the surface of solid particulate catalysts, and the reverse movement of products from the catalytic surface back to the body of the gas phase.

Principal applications. With such excellent opportunities for heat and mass transfer to or from solids and fluids, fluidization has become a major tool in such fields as drying, roasting, and other processes involving chemical decomposition of solid particles by heat. An important application has been in the catalysis of gas reactions, wherein the excellent opportunity of heat transfer and mass transfer between the catalytic surface and the gas stream gives performance unequaled by any other system. In the petroleum industry, for example, the cracking of hydrocarbon vapors in the presence of solid catalysts affords many advantages. Among these are a minimum tendency for overheating of one particle or one part of the catalytic bed as compared to another, and the immediate removal of the reactant products, resulting in greatly increased efficiency of the catalyst. The catalyst itself often becomes carbonized on the surface because of the coking tendency of the reacting hydrocarbons. This catalyst, as shown in the **illustration**, may then be removed by pneumatic transport or otherwise to a second fluidized reactor wherein an oxidation is accomplished, again with maintenance of relatively uniform conditions. The rejuvenated catalyst may be returned to the prime reactor by pneumatic or other transfer lines, and the process can be continued indefinitely.

Some fluid reaction systems have been standardized in design, permitting the operation of units comparable in maximum size to other types of units in the adjacent steps of processing.

From a process-development standpoint, the dense bed of particles and gas means that enough

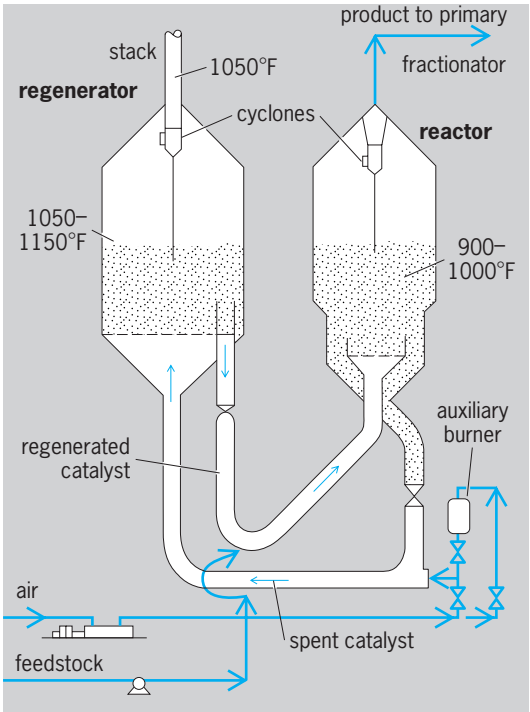

Catalytic cracking process using fluidized catalyst. °C = (°F − 32)/1.8.

catalyst can be maintained in a relatively small vessel to give a high degree of conversion with a low pressure drop causing gas flow. The mechanical design of a fluidization system and accessories is comparable, in many respects, to that of a body of liquid. However, special devices are needed to control fluid and solid handling, to introduce gas to the fluid bed, and to control the level of the solid particles, which usually are allowed to overflow the top of vertical pipes, as in maintaining the level of a true liquid. Vapor-disengaging units, filters, and cyclones, for removal of minute amounts of fines which might otherwise escape, blowers, pumps, pneumatic transport lines, and bucket elevators are also needed.

One important component in many units, besides the fluidizing chamber—often the reactor—is the standpipe, which allows a recycling of solids and builds up pressure on the solids for transfer purposes. In a large catalytic cracking unit, the rate of circulation of solids between vessels may be as much as 50 tons/min (0.75 metric ton/s).

Fluidized versus fixed beds. While catalytic particle-gas reactions have long been done in fixed beds, these cannot all be converted to fluidized operations. There are notable advantages and disadvantages of each. The advantages of the fluidized bed have been evidenced in the much better mixing, which results in greatly improved heat transfer, mass transfer, and uniformity. Thus there may often be better heat and material recovery, that is, increased yields, lower catalyst losses, and sometimes lower pressure drops. Fluidized beds allow ready additions of solids, combustion products, or direct combustion within the bed to give a much greater flexibility of process design and operation.

On the other hand, the concurrent flow in a fluidized reactor is disadvantageous, and multibed reactors are necessary to secure the advantages of countercurrent flow. There may be mixing of products with raw materials throughout the bed, and particularly when the solid is a reactant, it is difficult to obtain a high conversion and almost impossible to obtain a 100% conversion, except in a multibed system. Some attrition and thus loss of catalyst particles and changing fluidization properties with changing average size result. Recovery equipment for fines may be necessary, either because of their value or their nuisance as air pollutants if lost to the atmosphere. Similarly, there is considerable erosion of the walls, and especially of transfer lines. Because of the nature of fluidization, many reactions requiring a catalyst bed cannot be so operated, for example, those that give waxy or gummy materials which would agglomerate on the catalyst particles and also those wherein properties of the solids or of the gas stream, the necessary gas velocities, and so on, are outside of the mechanical possibilities of the fluidization operation. Thus, fixed beds are in many cases either required or most advantageous.

Individual industries. The simplest use of a fluidized bed is as a simple heat exchanger, with direct contact between solids and fluid—either liquid or gas—possibly to recover heat from a hot gas or from a hot solid by contacting with a cold solid or a cold gas, respectively. Such a heat exchanger, with no chemical reactions, is often used as a supplementary fluidized bed to one wherein there is a chemical reaction. Within a single bed, there is no possible countercurrent action, although multiple beds can be used and their particle flow arranged countercurrently to that of the fluid.

A basic example of diffusion is in the use of particles of an adsorbent, such as activated carbon, to remove vapors of solvent from a fluidizing stream of carrier gas, such as air. The air loses the vapor molecules by their diffusion first through a gas film surrounding the particles and then into the solid mass of the particles. The solvent-charged particles overflow the top of the bed to form a second bed, fluidized usually with steam to desorb the solvent. The steam-solvent mixture is separated by distillation, and the solvent-free particles are recycled to the first bed. Obviously, this example of mass transfer of molecules by diffusion to and from the particles also includes a transfer of heat in both beds.

A similar bed without chemical reaction is used as a dryer of solid particles. In this example, the moisture must diffuse to the surface of the particles, from which it is evaporated into the gas stream. It must again be noted that, because the flows in the bed are not countercurrent, a single unit cannot achieve complete dryness, except with an excess in heat consumption.

Very similar is the removal of chemically combined water and carbon dioxide, or both, in what is

usually called calcination, at temperatures about 1400–2200°F (760–1200°C). Multiple beds in countercurrent are used to minimize heat requirement. A necessary added expense is the grinding of limestone to a size which is practical to fluidize, usually not over $^1/_4$ in. (6 mm) maximum dimension. The crushed limestone descends from a top-drying fluidized bed through a series of beds to one where combustion of fuel oil is also taking place to supply the heat for calcination both in this bed and in the higher beds. The lowest bed receives and cools the burnt lime with the incoming air for combustion, which is thus preheated.

Roasting of pyrites to give sulfur dioxide is practiced throughout the world in possibly hundreds of fluidized beds; the combustion operation may be controlled better than in other types of roasting furnaces. Similarly, additional roasters are used to convert the iron oxide from hematite to magnetite, which then is heated to give pellets for steel manufacture.

Other sulfide ores are roasted in closely controlled fluidized beds to give metal sulfates, which may then be leached to recover their metal values from the gangue.

There are other examples of fluidized beds where there are more complicated chemical reactions between gases and solid. Included are the carbonization of wood and of coal to form charcoal and coke with concomitant volatile products, the combustion of pulverized coal and other solid fuels, and the production of portland cement clinker. More modern applications have been the production of titanium chloride in a bed of a ground titanium mineral, rutile, and coke fluidized with chlorine and the production of uranium compounds for nuclear fuels.

Fluid coking is an interesting cracking process, in which the undistillable residues of petroleum towers are coked in a fluidized reactor at about 1000°F (540°C) to give coke and a distillate. Coke is deposited on the fluidized particles of seed coke, prepared by fine grinding of some of the product, and these particles are built up to the desired product sizes for withdrawal and sale.

The greatest use in number and size of fluidized reactors and in volume of throughput has been in catalytic cracking and reforming of petroleum and its fractions. Scores of variations have developed the processing by particles of catalyst petroleum vapor streams. The modern petroleum industry could not otherwise operate under present conditions and costs. *See* ADSORPTION OPERATIONS; BULK-HANDLING MACHINES; CATALYSIS; CRACKING; DIFFUSION; FLUIDIZED-BED COMBUSTION; HEAT TRANSFER; UNIT OPERATIONS. Donald F. Othmer

Bibliography. J. F. Davidson (ed.), *Fluidization*, 1978; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., vol. 10, 1994; M. Kwauk (ed.), *Advances in Chemical Engineering*, vol. 20: *Fast Fluidization*, 1994; O. Levenspiel and D. Kunnii, *Fluidization Engineering*, 2d ed., 1991.

Fluidized-bed combustion

A method of burning fuel in which the fuel is continually fed into a bed of reactive or inert particles supported by upflowing air, which causes the bed to behave like a turbulent fluid. Fluidized beds have long been used for the combustion of low-quality, difficult-to-burn fuels and more recently have been developed for the clean burning of coal.

Process description. A fluidized-bed combustor is a furnace chamber whose floor is slotted, perforated, or fitted with nozzles to uniformly distribute upflowing air. The chamber is partially filled with maximum-sized particles which fluidize at the airflow rate required for combustion. When fluidization takes place, the bed of material expands (bulk density decreases) and exhibits the properties of a liquid. As air velocity increases, the particles mix more violently, and the surface of the bed takes on the appearance of a boiling liquid. If air velocity were increased further, the bed material would be blown away.

Once the temperature of the bed material is raised to the fuel-ignition temperature with start-up burners, fuel can be fed into or over the bed to achieve the desired operating temperature. Proper selection of air velocity, operating temperature, and bed material will cause the bed to function as a chemical reactor. In the case of high-sulfur coal combustion, the bed is generally limestone, which reacts with and

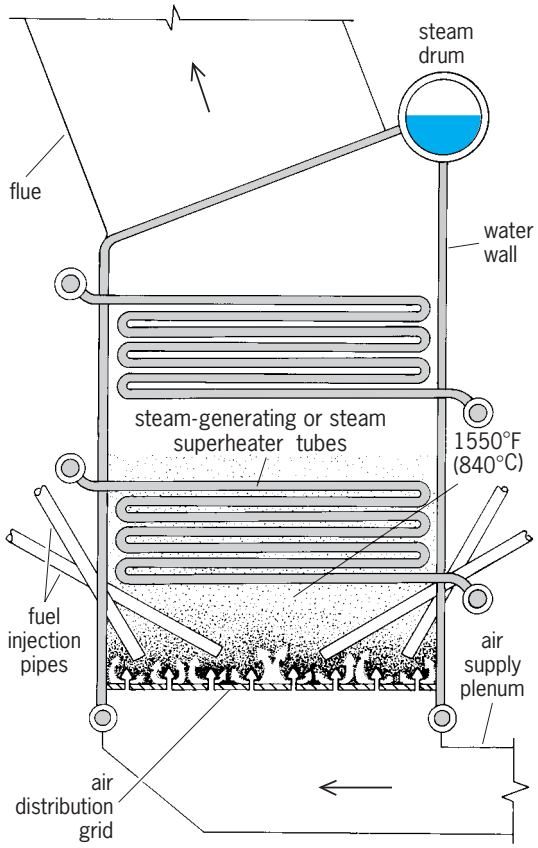

Fig. 1. Fluidized-bed steam generator.

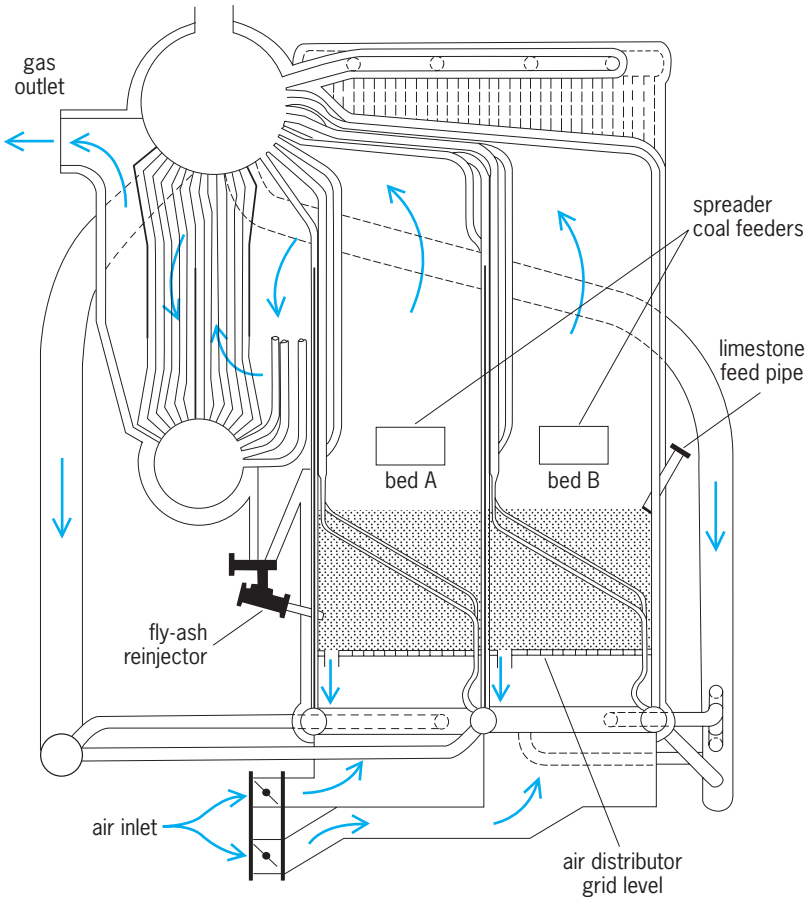

Fig. 2. Flow diagram showing the operation of a multicell fluidized-bed boiler.

absorbs the sulfer released during burning, reducing sulfur dioxide emissions. The alkali compounds in coal from the western United States can react with the sulfur dioxide to limit emissions.

Fluidized-bed combustion can proceed at low temperatures (1400–1500°F or 760–840°C), so that nitrogen oxide formation is inhibited. In the limestone-sulfur–absorbing designs, the waste product is a relatively innocuous solid, although the large volume produced does have environmental implications. Researchers in both federal and privately funded programs are exploring techniques for utilizing the waste product, for example, determining its potential value as an agricultural supplement or bulk aggregate. Low combustion temperatures minimize furnace corrosion problems. At the same time, intimate contact between the hot, turbulent bed particles and the heat-transfer surfaces results in high heat-transfer coefficients. The result is that less surface area is needed, and overall heat-transfer tube requirements and costs are lower. Intensive development of fuel-fired systems has been undertaken, particularly for pressurized operation, where the movement of solids through the pressure barrier poses some special problems. These efforts appear successful in pilot plants.

Although commonly regarded as a sulfur emissions control process, fluidized-bed combustion is applicable to low-sulfur fuels and wastes. In firing a low-sulfer coal from the western United States, for example, the bed material would simply be coal ash.

Applications. There are three areas of application: incineration, gasification, and steam generation.

Coal gasification. The original application of fluidized-bed combustion was to the gasification of coal. A gasifier burns coal in a fluidized bed with less air than that amount which is required for complete combustion. This results in a high concentration of combustible gases in the exhaust. This principle is being applied in a number of pilot plants which, for the most part, are directed at producing low-sulfur-

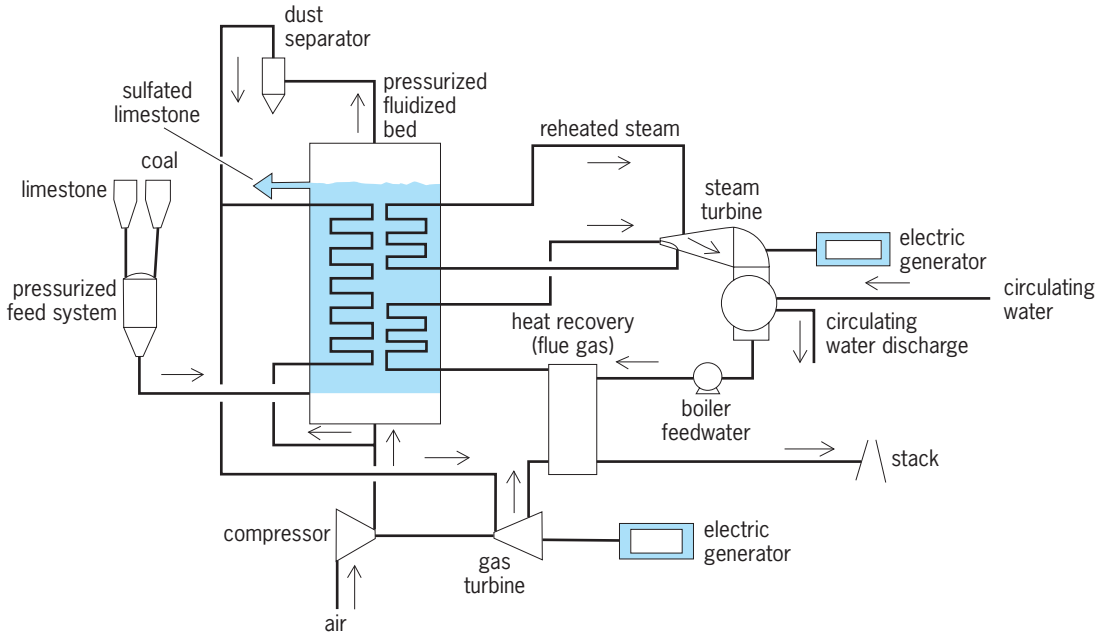

Fig. 3. Flow diagram of 600-MW combined-cycle system.

content gases of medium heating value (150–300 Btu per standard cubic foot or 5590–11,170 kilojoules per cubic meter). *See* COAL GASIFICATION.

Incineration. The use of this technology for efficient incineration of waste products, such as food-processing wastes, kraft process liquors, coke breeze, and lumber wastes, has increased rapidly. The process has the advantage of controllability, lower capital cost, more compact design, and reduced emissions.

Steam generation. Heat-transfer tubes are immersed within the hot fluidized bed, resulting in a compact steam boiler (**Fig. 1**). Combining cells of individual beds permits control of the output of larger-capacity boilers (**Fig. 2**). Beds may be arranged horizontally, vertically, or both ways in the same system.

Fluidized-bed steam generators with capacity as great as 350,000 lb/h (160,000 kg/h) are in operation throughout the world firing a variety of low-quality fuels. In Europe, emphasis has been on cogeneration systems and firing wastes. Industrial boilers firing coal and mining wastes are typical applications of the technology in the United States.

Operation and testing of a 20-megawatt pilot plant and initial steps for a 200-megawatt demonstration plant have resulted from cooperative efforts of the Electric Power Research Institute and the Tennessee Valley Authority.

In the so-called pressurized fluidized-bed combustor, operating the combustion chamber at elevated pressures permits the use of the heated exhaust gases to drive gas turbines. Tests have been conducted on a 13-MW pilot plant using this combined cycle. Designs have been made for larger plants (**Fig. 3**). *See* GAS TURBINE; STEAM-GENERATING UNIT.

Other applications. Fluidized-bed combustion techniques have had a long history of application in the chemical industries, especially in catalytic petroleum cracking, and in extractive metallurgy (ore roasters and calciners). Development of a fluidized-bed combustor for the spent graphite fuel rods from high-temperature gas-cooled nuclear reactors has been undertaken.

Competitiveness of process. Since nearly any combustible product, including such difficult fuels as oil shales, anthracite wastes, and residual oils, can be effectively burned, fluidized-bed combustion has gained worldwide commercial status. Environmentally, its competitiveness with flue-gas desulfurization, coal cleaning, and advanced technologies such as liquefaction has been established. *See* COAL LIQUEFACTION; COMBUSTION; FLUIDIZATION.

Michael Pope

Bibliography. E. P. Johnson, Fluidized-bed combustion: Better fluid-bed units to make debuts, *Chem. Eng.*, p. 39, April 1982; L. N. Rubow (ed.), *Fluidized Bed Combustion, 1993: FBC's Role in the World Energy Mix*, 1993; *Summary Evaluations of Atmospheric Pressure Fluidized Bed Combustion Applied to Electric Utility Large Steam Generators*, EPRI Report FP 308, prepared by Babcock and Wilcox Company, Alliance, Ohio, 1976.

Fluids

Substances that flow under shear stress and have a density range from essentially zero (gases) to solid-like values (liquids). Fluids comprise one of the two major forms of matter. Solids, the other form, generally deform very little when shear forces are applied, and their densities do not change significantly with pressure or temperature.

At the atomic level, the relative molecular positions in solids are well defined to very great distances, and the molecules rarely change locations. However, fluids have a uniform structure beyond a few molecular diameters. **Figure 1** schematically demonstrates the differences in molecular distribution in gases, liquids, and solids according to what x-ray and neutron scattering measurements indicate about the probability of finding molecular centers at any separation. Probabilities different from unity indicate "structure." For all phases and densities, the repulsive intermolecular forces prevent any molecules from approaching to small separations, so the probability at such separations is zero. Just beyond the repulsive barrier, the probability is higher in all phases due to attractive forces. For low-density gases, the fluid is uniform beyond the first neighbors and there is no real structure. For the denser liquids, there is only local structure, whose extent depends on the density. Fluids do not show the long-range structure observed in solids. Also, molecules in fluids change positions and orientations more frequently than do those in solids. In fact, at low densities, there is essentially no inhibition to molecular motion except for the container walls. Fluids have been conceptualized as mostly solidlike, with regions of vacuum or free individual molecules, and there even have been models, based on such ideas, that succeed in describing fluid properties. Fluid structure is not heterogeneous, as this concept implies. *See* CRYSTAL STRUCTURE; INTERMOLECULAR FORCES; NEUTRON DIFFRACTION; SOLID-STATE PHYSICS; X-RAY DIFFRACTION.

The distinction between solids and fluids is most easily seen in substances and mixtures which show a well-defined melting process. For substances with large molecules, such as polymers, ceramics, and biologicals, this distinction is less clear. Instead, there is a slow evolution of structure and of resistance to flow

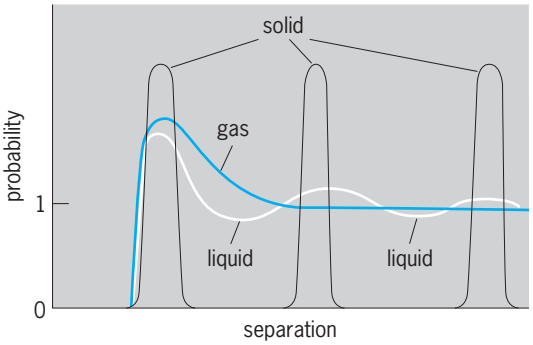

Fig. 1. Probability of finding molecular centers at various separations in gases, liquids, and solids.

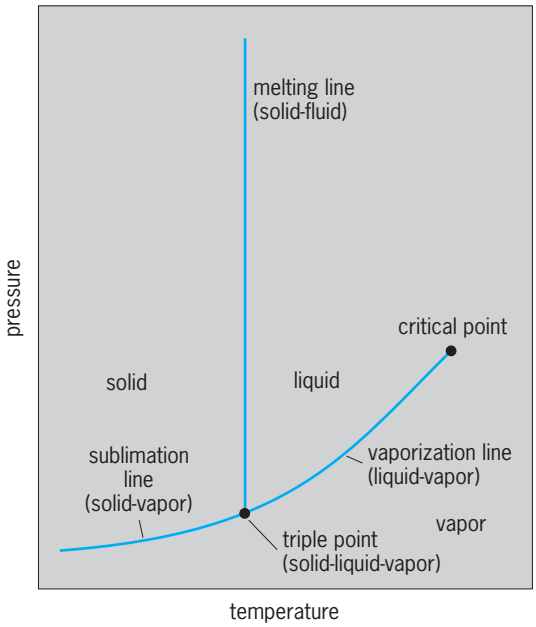

Fig. 2. Conditions of pure-component phase behavior.

as temperature or some other variable is changed. *See* GLASS TRANSITION; MELTING POINT; POLYMER.

Vapors and liquids. Molecular density varies greatly in fluids and is their most important characteristic. The distinction between vapors (or gases) and liquids is most clear for substances and mixtures that show well-defined vaporizing (boiling) and condensing processes. The high-density liquid boils to make a low-density gaseous vapor. **Figure 2** shows the pressures and temperatures for which pure substances are single phases. At the conditions of the lines between the single-phase regions, two phases can be observed to coexist. At the state of intersection of the lines (the triple point), three phases can coexist. For most substances, the triple-point pressure is well below atmospheric. However, for carbon dioxide, it is very high, so that dry ice sublimes rather than melts, as water ice does. Beyond the end of the liquid-vapor (saturation or vapor-pressure) line, vaporization and condensation cannot be observed. The state at the end of this line is called the critical point, where all the properties of the vapor and liquid become the same. There is no such end to the solid-liquid (or melting) line, because the solid and fluid structures cannot become the same. *See* CRITICAL PHENOMENA; PHASE TRANSITIONS; TRIPLE POINT.

Figure 3 shows the relationship of pressure and density of fluids at temperatures above the melting point and at pressures above the triple point. As the pressure on a vapor is increased at fixed temperature (T_1), the density changes rapidly until the condensation point is reached. During condensation, the pressure on the two phases is constant until all the vapor is gone. Then the fluid is relatively incompressible, and the density changes very little with even large changes in pressure. The region of conditions where two phases are observed is capped by the critical point. At a temperature just above the critical (T_2), the variation of density with pressure is very great, even though the system is one phase. At the highest temperatures (T_3), the pressure-density variation is more gradual everywhere. The relation between fluid-state variables of pressure P, density ρ and temperature T is often represented by a quantity called the compressibility factor, $z = P/\rho RT$ where R is a universal constant equal to 1.987 calories/mol K, 1.987 Btu/lb mole °R, or 8.3143 J/mol K and where density is in units of moles per unit volume. The behavior of z is shown in **Fig. 4**. At the lowest pressures and highest temperatures (T_3 and T_4), z approaches 1, where the fluid is called an ideal gas and the molecules behave as if they do not interact with each other. At most temperatures (T_1 and T_2), z decreases as the pressure is raised, becoming very small during the condensation process. However, because of the relative incompressibility of the liquid, z becomes large at the highest pressures. Mathematical representations of the variation of z with T, P, or ρ and composition are called equations of state, a famous form being that due to J. D. van der Waals. Many equations of state of much higher accuracy have been developed, though the most complete expressions are sufficiently complex that a computer is necessary to actually calculate the properties they can give. *See* GAS; GAS CONSTANT; THERMODYNAMIC PROCESSES; VAN DER WAALS EQUATION; VAPOR PRESSURE.

Mixtures. Mixtures of fluids show the same general density and multiphase behavior as pure fluids, but the composition is an extra variable to be considered. For example, the density differences between the vapor and the liquid phases cause them to have different relative amounts of the components. This difference in composition is the basis of the separation process of distillation, where the vapor will be richer in some components while the liquid will be richer in others. In the distillation refining of petroleum, the vapor is enriched with the components

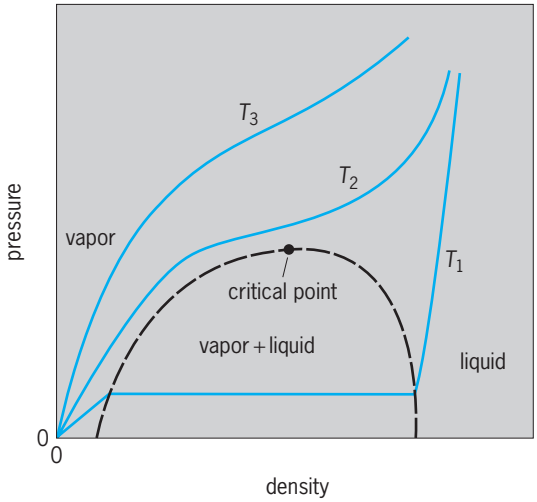

Fig. 3. Behavior of pure-component pressures with density at various temperatures.

of gasoline, while the liquid is richer in heavier oils. It is also possible for mixtures of liquids to be partially or nearly wholly immiscible, as are water and oil. The separation process of liquid extraction, used in some metal-purification systems and chemical-pollution-abatement processes, depends on different preferences of chemical solutes for one liquid phase or the other. Both vapor-liquid and liquid-liquid systems can show critical points where the densities and compositions become the same. In the single-phase conditions near the critical point, the great variation of density with small changes in pressure can be used to selectively remove components from sensitive substances such as foods and pharmaceuticals. This process (called supercritical extraction) is used to extract caffeine from coffee with carbon dioxide near its critical point. In some systems, it is possible to observe many liquid phases. A well-known example has nine liquid phases made from mixing nine liquid substances. *See* DISTILLATION; EXTRACTION; PHASE RULE.

Fluid interfaces. The usual observation of the presence of more than one fluid phase is the appearance of the boundary or interface between them. This is seen because the density or composition (or both) changes over a distance of a few molecular diameters, and this variation bends or scatters light in a detectable way. At the interface, the molecules feel different forces than in the bulk phases and thus have special properties. Energy is always required to create interface from bulk, the amount per unit area being called the interfacial tension. Water is a fluid with an extremely high vapor-liquid (or surface) tension; this surface tension allows insects to crawl on ponds and causes sprinkler streams to break up into sprays of droplets. One way to measure the surface tension of a fluid is to measure its capillary rise in small tubes. The height of the fluid rise decreases with the tube diameter and increases with the surface tension. Another effect of interfacial tension is the increase of vapor pressure of a component in a very small droplet. Liquid mercury will not evaporate in an open container with an essentially flat interface. However, if it is spilled into little drops, those too small to be easily seen will have their vapor pressure so enhanced that the mercury will vaporize. (This can be a health hazard in chemical and other laboratories.) As the conditions of temperature and composition are changed toward a critical point, the interfacial tension decreases to zero. *See* INTERFACE OF PHASES; SURFACE TENSION.

In mixtures, the molecules respond differently to the interfacial forces, so the interfacial composition is generally different from that of the bulk. This has also been the basis of a separation process. If the difference of composition is great enough and it varies with time and position because of evaporation of one or more of the components, the interfacial forces can push the fluid into motion, as can be observed on the walls of a glass of brandy (the Marangoni effect). Some substances strongly adsorb at the interface because their chemical structure has one part that prefers to be in one phase, such as water, and another part that prefers the other phase, such as oil or air. Such surfactants or detergents help solubilize dirt into wash water, keep cosmetics and other immiscible mixtures together, and form foams when air and soapy water are whipped together. *See* ADSORPTION; DETERGENT; FOAM; SURFACTANT.

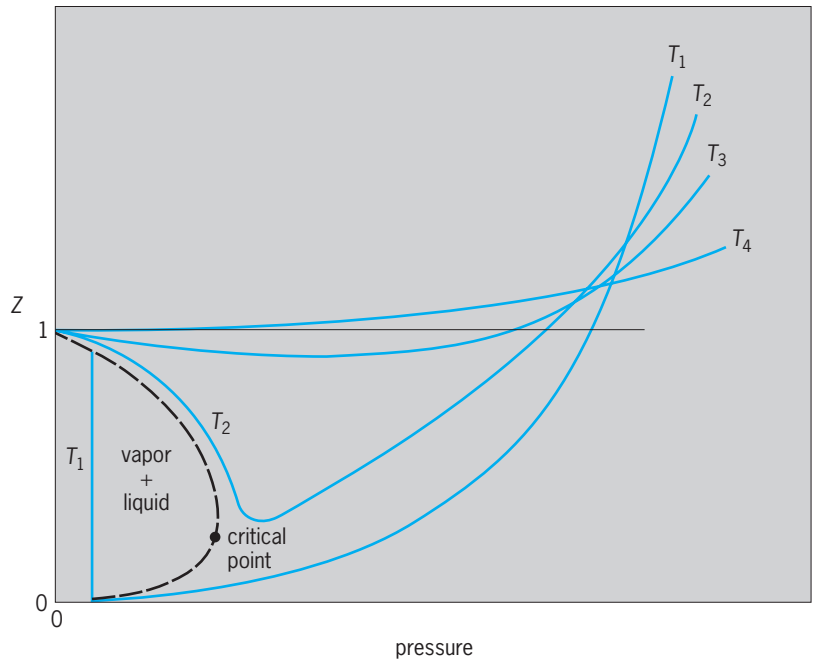

Fig. 4. Behavior of pure-component compressibility factors with pressure at various temperatures.

Transport properties. Besides the relations among pressure, density, temperature, and composition of static or equilibrium fluids, there are also characteristics associated with fluid flow, heat transfer, and material transport. For example, when a liquid or gas flows through a tube, energy must be supplied by a pump, and there is a drop in pressure from the beginning to the end of the tube that matches the rise in pressure in the pump. The pump work and pressure drop depend on the flow rate, the tube size and shape, the density, and a property of the molecules called the viscosity. The effect arises because the fluid molecules at the solid tube wall do not move and there are velocity gradients and shear in the flow. The molecules that collide with one another transfer momentum to the wall and work against one another, in a sort of friction which dissipates mechanical energy into internal energy or heat. The greater the viscosity, the greater the amount of energy dissipated by the collisions and the greater the pressure drop. The viscosity generally increases with the density. In gases, viscosity increases with temperature and decreases with molecular size and attractive forces. In liquids, the effects are opposite; the main effect is due to changes in density. If only chemical constitution and physical state are needed to characterize the viscosity, and if shear stress is directly proportional to velocity gradient, the fluid is called newtonian and the relation for pressure drop is relatively simple. If the molecules are large or the attractive forces are

very strong over long ranges, as in polymers, gels, and foods such as bread dough and cornstarch, the resistance to flow can also depend on the rate of flow and even the recent deformations of the substance. These fluids are called non-newtonian, and the relationship of flow resistance to the applied forces can be very complex. *See* NEWTONIAN FLUID; NON-NEWTONIAN FLUID; VISCOSITY.

Another fluid-transport property, thermal conductivity, indicates the ability of a static fluid to pass heat from higher to lower temperature. This characteristic is a function of chemical constitution and physical state in a similar way as is the viscosity. In mixtures, these properties may involve simple or complex dependence on composition, the variation becoming extreme if the unlike species strongly attract each other. The values of both properties increase rapidly near a critical point. *See* CONDUCTION (HEAT); HEAT TRANSFER.

Finally, the ability of molecules to change their relative position in a static fluid is called the diffusivity. This is a particularly important characteristic for separation processes whose efficiency depends on molecular motion from one phase to another through a relatively static interface, or on the ability of some molecules to move faster than others in a static fluid under an applied force. The diffusivity is highest for the smallest molecules with the weakest attractive intermolecular forces, decreasing as the size and forces increase. The diffusivity in gases increases with temperature and decreases with density; often, the product of the diffusivity and the density divided by the product of the viscosity and the absolute temperature is nearly constant. In liquids the diffusivity tends to be inversely proportional to viscosity of solvent. *See* DIFFUSION; GAS; LIQUID.

John P. O'Connell

Bibliography. P. W. Atkins, *Physical Chemistry*, 6th ed., 1997; G. M. Barrow, *Physical Chemistry*, 6th ed., 1996; I. N. Levine, *Physical Chemistry*, 5th ed., 2000; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; J. M. Prausnitz et. al., *Molecular Dynamics of Fluid-Phase Equilibria*, 2d ed., 1985; R. C. Reid, J. M. Prannitz, and B. Poling, *Properties of Gases and Liquids*, 4th ed., 1987.

Fluorescence

Fluorescence is generally defined as a luminescence emission that is caused by the flow of some form of energy into the emitting body, this emission ceasing abruptly when the exciting energy is shut off.

Definition. In attempts to make this definition more meaningful it is often stated, somewhat arbitrarily, that the decay time, or afterglow, of the emission must be of the order of the natural lifetime for allowed radiative transitions in an atom or a molecule, which is about 10^{-8} s for transitions involving visible light. Perhaps a better distinction between fluorescence and its counterpart, a longer-lasting afterglow called phosphorescence, rests not on the magnitude of the decay time per se but on the criterion that the fluorescence decay is temperature-independent. If this latter criterion is adopted, the luminescence emission from such materials as the uranyl compounds and rare-earth-activated solids would be called slow fluorescence rather than phosphorescence. The decay of their luminescence takes place in milliseconds to seconds, rather than in 10^{-8} s, showing that the optical transitions are somewhat "forbidden"; but the decay is temperature-independent over a considerable range of temperature, and it follows an exponential decay law, shown in the equation below, that is to be expected for spontaneous transitions of electrons from an excited state of an atom to the ground state when the atom has a transition probability per unit time $1/\tau_1$ (τ_1 is called the natural radiative lifetime or fluorescence lifetime). In this equation, I is the luminescence intensity at a time t, and I_0 is the intensity when $t = 0$, that is, the time at which the exciting energy is removed. *See* SELECTION RULES (PHYSICS).

$$I = I_0 \exp\left(\frac{-t}{\tau_1}\right)$$

In applying this criterion, it is necessary to take into account a restriction that arises because all luminescent systems ultimately lose efficiency or are "quenched" at elevated temperatures, each system having its own characteristic temperature for the onset of this so-called thermal quenching. Quenching sets in because increase of temperature makes available other competing atomic or molecular transitions that can depopulate the excited state and dissipate the excitation energy nonradiatively. Hence, even for a fluorescence, a temperature dependence of decay time will be observed at temperatures where thermal quenching becomes operative, because $1/\tau_{obs} = 1/\tau_1 + 1/\tau_{diss}$, where τ_{obs} is the observed fluorescence lifetime and $1/\tau_{diss}$ is the probability for a competing dissipative transition.

In the literature of organic luminescent materials, the term fluorescence is used exclusively to denote a luminescence which occurs when a molecule makes an allowed optical transition. Luminescence with a longer exponential decay time, corresponding to an optically forbidden transition, is called phosphorescence, and it has a different spectral distribution from the fluorescence. *See* PHOSPHORESCENCE.

Materials. Fluorescence can be excited in gases, liquids, or solids. Solid luminescent materials are commonly called phosphors, regardless of their decay times, which can cover a very wide range, from the order of 5×10^{-9} s for many organic crystalline materials up to 2 s for the europium-activated strontium silicate phosphor.

Applications. Fluorescent materials with decay times between 10^{-9} and 10^{-7} s are used to detect and measure high-energy radiations, such as x-rays and gamma rays, and high-energy particles such as alpha particles, beta particles, and neutrons. These agents produce light flashes (scintillations) in certain crystalline solids, in solutions of many polynuclear

aromatic hydrocarbons, or in plastics impregnated with these hydrocarbons. *See* SCINTILLATION COUNTER.

Another application of x-ray-induced fluorescence is in intensifying screens used in medical radiography. Such screens, placed in contact with radiographic film, considerably decrease the x-ray exposure needed to form an image because of the added photographic effect of the fluorescence. Blue-fluorescent calcium tungstate and rare-earth-activated phosphors are normally used.

Electronic displays. The fluorescence excited by electron-beam bombardment of a phosphor coating on a cathode-ray-tube face produces the images seen in television. For this purpose, phosphors with short decay times must be used, because each phosphor picture element is scanned very briefly by the electron beam approximately 30 times per second, and the luminescence from each excitation must decrease to a low level that will not interfere with the next signal. Phosphors with decay times as long as a few hundredths of a second can be tolerated in this application, although shorter decay times are more advantageous. *See* CATHODE-RAY TUBE; PICTURE TUBE.

Vacuum fluorescence display devices also make use of fluorescence excited by an electron beam in a flat structure resembling a triode vacuum tube. In these devices the phosphor is coated on the anodes of the tube, which are shaped to form the segments of various digital or alphanumeric figures, and the excitation of each segment is controlled by the potential applied to that anode. *See* CATHODOLUMINESCENCE; ELECTRONIC DISPLAY.

Visual effects. Laundry brighteners are another application of fluorescence. These materials are colorless organic phenyl-based dyes. They are excited by the small amount of ultraviolet light present in daylight to give a blue fluorescence, thus compensating for the slight absorption of blue light responsible for the cream color of most natural bleached textiles. The fluorescence of other dyes and inorganic phosphors applied to fabrics or other substrates is used to obtain special visual effects for advertising, theatrical, and other display purposes, under both daylight and ultraviolet excitation. *See* DYE.

Research, diagnosis, and prospecting. Fluorescence is extensively used for research and diagnosis in microbiology and medicine, where fluorescent dyes, such as fluorescein, rhodamine, and aniline blue, are used to label and trace proteins. Among other applications of fluorescence are the use of water-soluble fluorescent dyes to trace the flow of hidden currents, and as an aid in prospecting for certain minerals, sceelite (calcium tungstate) giving a blue fluorescence and willemite (zinc silicate) giving a green fluorescence under short-wave ultraviolet excitation. *See* FLUORESCENCE MICROSCOPE; IMMUNOFLUORESCENCE; SCHEELITE; WILLEMITE.

Fluorescent lamps. The so-called fluorescent lamps employ the luminescence of gases and solids in combination to produce visible light. A fluorescent lamp consists of a glass tube filled with a low-pressure mixture of argon gas and mercury vapor, coated on the inside surface with a luminescent powder or blend of such powders, and having an electrode at each end. An electrical discharge is passed through the gas between the two electrodes. Bombardment by the electrons and ions of the discharge excites the mercury atoms to higher energy states, from which they return to lower states by radiative transitions, producing both visible and ultraviolet light emission. At the low pressures employed, approximately half of the electrical energy input to the lamp is converted into the 253.7-nanometer radiation characteristic of the mercury atom.

The phosphor coating is chosen for the efficiency with which it is excited by this wavelength of ultraviolet and for the color of visible luminescence that is desired. Lamp phosphors are predominantly alkaline-earth halophosphates, symbolized by $(Ca,Sr)_5(PO_4)_3(F,C1)$, activated by antimony and manganese. The phosphors operate on the principle of sensitized luminescence; antimony, the sensitizer, is the primary absorber of the 253.7-nm radiation. Some of the excited antimony atoms fluoresce in the blue-green spectral region, while others transfer their excitation energy to manganese atoms, which fluoresce in the yellow-red region of the spectrum. The net color of the lamp fluorescence is thus some shade of white determined by the exact composition of the phosphor or blend of phosphors employed.

The small amount of visible light generated by the discharge itself is largely transmitted by the phosphor coating and adds slightly to the luminous output of the lamp. Luminous efficiencies of fluorescent lamps are considerably higher than those of incandescent lamps, which unavoidably convert into heat the major portion of the electrical energy supplied to them. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; FLUORESCENT LAMP; LUMINESCENCE.

James H. Schulman

Bibliography. B. DiBartolo (ed.), *Luminescence of Inorganic Solids*, 1978; G. G. Guilbault (ed.), *Practical Fluorescence*, 1990; K. Nassau, *Physics and Chemistry of Color*, 1983.

Fluorescence microscope

An instrument for the observation and study of microscopic specimens that absorb light and emit fluorescence. The ability of substances to emit light when excited by a strong light source is called luminescence, which encompasses the phenomena of phosphorescence and fluorescence. If the secondary emission continues for some time (milliseconds) after the light source has been turned off, it is referred to as phosphorescence. The ability to emit light only very briefly (nanoseconds) after the absorption of light is termed fluorescence. Many specimens obtained from plants, animals, and minerals have the ability to emit a weak to strong fluorescence when excited with appropriate light and without being specifically treated or stained. That phenomenon is called autofluorescence. In most

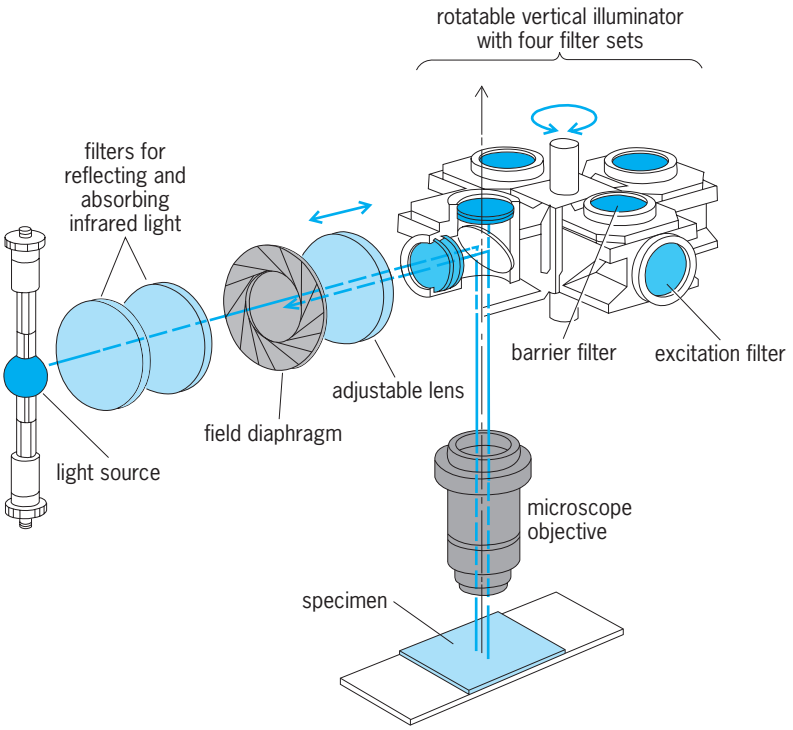

Schematic diagram of a fluorescence illuminator, the Ploemopak, with four filter sets.

cases, in order to obtain specific and meaningful fluorescence, staining with fluorescing dyes called fluorophores or fluorochromes is necessary. Fluorescence microscopy is a highly sensitive method, since often minute quantities of a fluorophore can be visualized with good microscopic contrast. In appropriate applications, brightly fluorescing images can be observed against a dark background. Individual fluorophores have different absorption and emission spectra and a different quantum efficiency (the ratio between the energy absorbed and the energy emitted), factors that must be considered for optimum fluorescence. *See* BIOLUMINESCENCE; FLUORESCENCE; LUMINESCENCE; PHOSPHORESCENCE.

Microscope components. The basics for fluorescence microscopy are the light source necessary to illuminate the specimen and the optics needed to observe the fluorescence. In addition, filters must be used to single out appropriate excitation and emission wavelengths. Excitation filters select out a limited range of excitation wavelengths from the light source that corresponds to the absorption spectrum of the fluorochrome. Barrier filters separate emitted light from unabsorbed exciting light. Fluorescence can then be observed by eye, photographed, measured by a photomultiplier, or recorded by a television camera. Two types of illumination are used in fluorescence microscopy: transmitted and incident. The earliest fluorescence microscopes relied on transmitted illumination, which generally used a dark-ground condenser to facilitate the separation of fluorescent and exciting light. With the development of epi-illumination, the exciting light reaches the preparation from above by way of a dichroic mirror and the objective, which at the same time acts as a condenser.

Epi-illumination by means of a vertical illuminator, which permits the excitation with more wavelengths, is known as a fluorescence illuminator (see **illus.**), and has become the routine instrument for fluorescence microscopy. Its dichroic has a special interference coating that reflects light shorter than a specified wavelength and transmits that of longer wavelengths. By placing the dichroic mirror at a 45° angle in the light path above the objective, the exciting, or shorter, wavelengths are reflected onto the preparation, and the emitted, or longer, wavelengths are transmitted in the direction of the eyepieces. Unlike microscopes with transmitted illumination, which require focusing of both the substage condenser and the objective, the fluorescence microscope with a fluorescence illuminator requires focusing of the objective only. In addition, a combination of epi-illumination fluorescence microscopy and transmitted microscopy is feasible.

Fluorescence illuminators have been manufactured with four different sets of dichroic mirrors, excitation filters, and barrier filters, each combination appropriate for the excitation of a specific fluorochrome. They are particularly useful in immunofluorescence microscopy, where often two or more fluorochromes are being used in one specimen. Optimum images can be obtained when low-power, high-numerical-aperture objectives are used, because the intensity of fluorescence is directly proportional to the numerical aperture of the objective (in incident illumination, to the fourth power of the numerical aperture) and image brightness is inversely proportional to the square of the total magnification. In general, simpler optics are preferred in fluorescence microscopy because they will have a higher transmission for fluorescent light and will produce less autofluorescence. Only ultraviolet exciting light demands special, quartz optics. *See* OPTICAL MICROSCOPE; OPTICAL PUMPING.

Confocal laser-scanning fluorescence microscopy. In laser-scanning microscopy, the object is not illuminated as a whole but is scanned step by step with a laser-illuminated spot. From each point the fluorescence is measured and, after analog-to-digital conversion, stored as a matrix in computer memory. The main advantage of laser-scanning microscopy over conventional methods in flourescence microscopy is the point-by-point illumination. Most stray light is avoided, and so the level of autofluorescence from optical parts and immersion oil is extremely low. The concentration of the laser light on a very small spot results in a high level of exciting energy per unit of area. That permits the investigation of very weakly fluorescing structures and the use of very low magnification (25 times), which is necessary, for example, in mapping the brain. The amount of fading is very low because of the extremely short exposure times. In the confocal mode, which uses a pinhole in the illumination pathway and a corresponding pinhole in front of the light detector, the fluorescence from above and below the selected focal plane is almost

completely eliminated from the image. That removes most of the glare experienced in conventional fluorescence microscopy of thicker specimens. Confocal fluorescence microscopy allows optical image sectioning of the specimen and, after combining the multiple images at different focal levels of a specimen, a computerized reconstruction of three-dimensional images of a section. *See* CONFOCAL MICROSCOPY; LASER.

Applications. The primary application of fluorescence microscopy is in the field of medicine, where diagnostic tests have been developed that use monoclonal antibodies to which a red, green, or blue fluorescent dye has been attached. The dyes thus reveal various components in the specimen, such as bacteria, viruses, or macromolecules like deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Many fluorescent cytochemical staining reactions have been developed to detect other components in plant and animal tissue. Outside the field of biology, fluorescence microscopy can analyze precisely the spectral wavelength range of the emitted autofluorescence of minerals and materials used in integrated electronic circuits, thereby helping identify specific substances or chemical impurities. *See* ANIMAL VIRUS; IMMUNOFLUORESCENCE; MEDICAL BACTERIOLOGY.

Johan S. Ploem

Fluorescent lamp

A lamp that produces light largely by conversion of ultraviolet energy from a low-pressure mercury arc to visible light. Phosphors, chemicals that absorb radiant energy of a given wavelength and reradiate at longer wavelengths, produce most of the light provided by fluorescent lamps. *See* PHOSPHORESCENCE.

The lamp consists of a glass tube containing two electrodes, a coating of activated powdered phosphor, and small amounts of mercury. The glass tube seals the inner parts from the atmosphere. The electrodes provide a source of free electrons to initiate the arc, and are connected to the external circuit through the ends of the lamp. The phosphor converts short-wave ultraviolet energy into visible light. The mercury, when vaporized in the arc, produces the ultraviolet radiation that causes fluorescence. An inert gas, such as argon, krypton, or neon, introduced in small quantities provides the ions that facilitate starting of the lamp (**Fig. 1**). *See* FLUORESCENCE.

Operation. To start the flow of current that forms the arc in a typical fluorescent lamp, free electrons are released into the tube by thermionic emission,

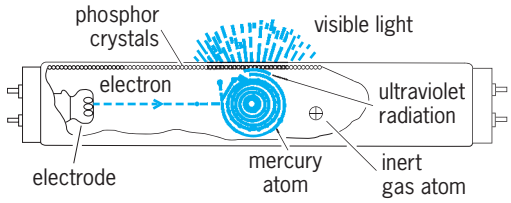

Fig. 1. Parts of a typical fluorescent lamp.

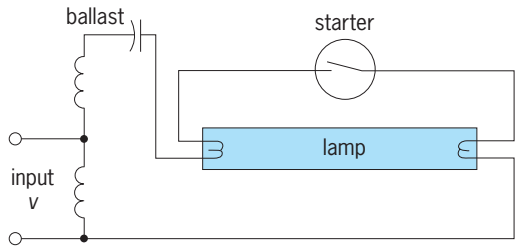

Fig. 2. Preheat or switch-start lamp circuit.

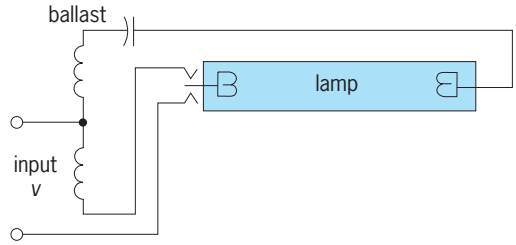

Fig. 3. Instant-start lamp circuit.

field emission, or a combination of both, causing a cathode to emit electrons.

In a preheat, rapid-start, or trigger-start type of lamp, current passed through the electrodes causes emission of electrons from the electrodes into the tube. When sufficient electrons are released and a sufficient voltage is applied, the resistance of the gap between the electrodes is low enough to permit striking the arc across the gap. *See* THERMIONIC EMISSION.

The application of a high potential difference across opposite instant-start cathodes in some lamps draws electrons from the negative electrode and attracts them toward the positive electrode. On alternating-current (ac) circuits, the electrodes are alternately negative and positive in each half-cycle, and so both electrodes emit electrons. Instant-start cathodes may be either hot or cold. The hot cathode permits greater lamp current and lower overall lighting costs; it consists of a coiled wire coated with an emissive material that yields electrons freely. The cold cathode is well suited to sign tubing and lamps of special lengths or shapes; it consists of a thimble-shaped iron cup. *See* FIELD EMISSION.

Electrodes designed to heat quickly can be used in conjunction with a moderately high voltage to attract electrons into the tube; this technique is used with rapid-start electrodes. Rapid-start electrodes are continuously heated during lamp operation.

The three types of fluorescent lamps employ either the preheat, instant-start, or rapid-start electric circuits. All include a ballast, whose primary function is to control the amount of current that flows through the lamp. Without a ballast, a fluorescent lamp would burn out instantly because there would be no impedance to limit the current. Another necessary function of the ballast is to provide the proper voltage to start the lamp. For example, a typical 40-W rapid-start ballast must provide a continuous cathode

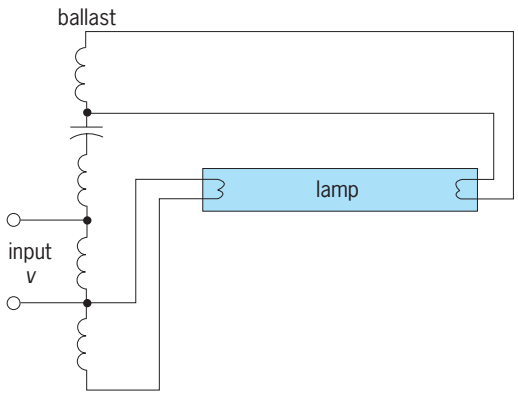

Fig. 4. Rapid-start or trigger-start lamp circuit.

heating voltage of 3.5 V during starting and operation in order to obtain optimum lamp performance. Many circuit variations are used for the operation of one or more lamps from a single ballast, with and without circuit elements for the correction of power factor.

Preheat circuit. In the preheat circuit (**Fig. 2**), a starting switch, usually an automatic starter, is used to heat the electrodes. The most common starter employs a small argon glow tube with one fixed electrode and one electrode made of a bent bimetallic strip. When the circuit is energized, a low-current glow discharge forms in the starter; the heat from this glow is sufficient to expand the bimetallic electrode until it contacts the fixed electrode, forming a short circuit through the starter. The full output voltage of the ballast then causes current to flow through the lamp electrodes, heating them and causing them to emit electrons. The starter cools because the glow is no longer present, and the switch opens, impressing full ballast voltage between the lamp electrodes. If there are enough electrons in the tube, the arc is formed; if not, the glow switch process is repeated. This process may require several repetitions, accounting for the delay and flickering at the starting of preheat lamps. When the arc has formed and full lamp current is flowing, the ballast absorbs about half its initial voltage, and there is not sufficient voltage to cause the glow switch to operate; it then becomes an inactive circuit element. *See* NEON GLOW LAMP.

Preheat fluorescent lamps are commercially available in lengths from 6 in. to 5 ft (15 cm to 1.5 m), with wattages from 4 to 90 W. Some preheat lamps have a self-contained ballast, which makes them useful as a portable light source. Longer lamps or lamps of higher wattage generally use instant-start or rapid-start circuits, which have more favorable starting characteristics. Each size of fluorescent lamp requires a ballast designed to match the lamp's requirements in terms of voltage, current, and wattage.

Instant-start circuit. In the instant-start circuit (**Fig. 3**), the ballast voltage is much higher than in the preheat circuit. Immediately upon energizing the circuit, this voltage attracts enough electrons into the tube to form the arc. For a given lamp wattage and current, the ballast must absorb a higher voltage than with preheat ballasts; hence the ballast is usually larger and dissipates greater wattage. Lampholders for instant-start circuits are connected so that the ballast is disconnected when the lamp is removed, eliminating the hazard of high voltage during maintenance operations.

Instant-start lamps for general lighting purposes are available in lengths from 24 in. to 8 ft (0.6 to 2.4 m), with wattages from about 15 to 75 W. Shorter lamps are not usually economical for general lighting service, because they require larger and more costly ballast than preheat circuits. Higher wattages would require overly large ballasts and cathodes of costly construction.

Rapid-start circuit. The rapid-start ballast, which has become the most popular type of ballast for domestic use in the United States, contains transformer windings that continuously provide the required voltage and current for electrode heating (**Fig. 4**). When the circuit is energized, these windings quickly heat the electrodes, releasing enough electrons into the tube for the lamp to arc from the voltage of the secondary windings. This combination of heat and moderately high voltage permits quick lamp starting with smaller ballasts than those for instant-start lamps, eliminates the annoying flicker associated with the starting of preheat lamps, and eliminates the starter and its socket from the lighting system.

The continuously heated cathode of the rapid-start lamp is better adapted to higher lamp currents and wattages. Therefore, rapid-start lamps are available in wattages up to 215 W for 8-ft (2.4-m) lamps. Compact

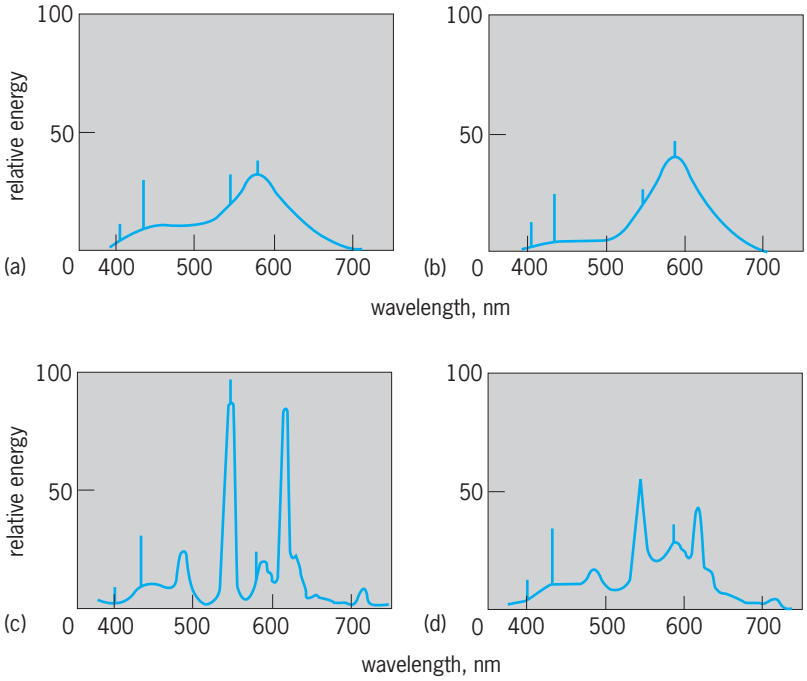

Fig. 5. Spectral energy distribution curves for fluorescent lamps with mercury emission-line spectrum. (*a*) Cool white halophosphate. (*b*) Warm white halophosphate. (*c*) Rare-earth–activated 4100 K (7000°F) triphosphor. (*d*) Triphosphor-halophosphate combination. (*General Electric Co.*)

twin tubes have been developed that are compatible with rapid-start operation. They use smaller diameters (0.4–0.6 in. or 10–15 mm), range in wattage from 5 to 40 W, and vary in length from 6 to 22.5 in. (15 to 57 cm).

Special ballast and circuits are available that permit the economical dimming and flashing of rapid-start lamps, providing a range of applications and control that was not possible with previous fluorescent lamp types. Rapid-start lamps are commonly used in flashing signs, and in residential and commercial lighting where continuously variable illumination levels are desired.

High-frequency fluorescent lighting. Fluorescent lamps are usually operated on ac circuits with a frequency of 60 Hz. However, higher frequencies permit higher-efficacy lamp operation along with ballasts of lower power dissipation per watt. Consequently, systems have been developed for the operation of fluorescent lamps at frequencies from 360 to 50,000 Hz. The most important high-frequency ballasts operate the lamps in the 25-kHz range, are lighter in weight, can have flicker-free light output, and will continue to be more cost-effective as the cost of electricity increases. Electronic technology permits these ballasts to be made with more features, such as dimming and low-cost remote controls, enhancing their overall performance value.

Lamp colors. Fluorescent lamps provide light at several times the efficacy of incandescent lamps, the exact ratio depending on the fluorescent lamp color. Lamp color is determined by the selection of chemicals used in the phosphors; various chemicals respond to the ultraviolet energy in the arc by producing different colors of light. Several types of essentially white fluorescent lamps are available commercially, as well as a range of tinted and saturated colors.

One of the earliest fluorescent whites was the so-called daylight lamp, which produced a bluish-white light of poor color rendition. The most widely used fluorescent color is called cool white; its lighted appearance is whiter than that of the daylight lamp and produces more light per watt. However, rare-earth–activated phosphors are also efficient and have substantially higher color rendition properties. **Figure 5** illustrates some typical spectral energy distributions.

Lamp shapes. The conventional fluorescent lamp is an elongated round tube. Lamp lengths vary from 6 to 96 in. (15 cm to 2.4 m). The highest-wattage lamps have indented grooves along the tube length. Compact fluorescent lamps come in circular, U-shaped, rectangular, twin-tube, quad, and other configurations.

Performance. Fluorescent lamps, commercially available since 1940, have been constantly improving in performance. Phosphor improvements have increased lumen outputs, and electrode designs now yield lamp life values of 20,000 h or more. An important measure of lamp performance is luminous efficacy, the quotient of the total luminous flux divided by the total lamp power input. The F40 system had a luminous efficacy of about 50 lumens per watt in 1950, and has reached values of 90 lumens per watt. The longer 96-in. (2.4-m) lamps, when used with the latest-technology ballasts, are expected to achieve efficacy values of about 100 lu/W. For proper low-temperature performance (−20°F or −29°C), fluorescent lamps are enclosed within jackets to assure high light output. *See* VAPOR LAMP. Edward E. Hammer

Bibliography. E. E. Hammer, High frequency characteristics of fluorescent lamps up to 500 kHz, *J. Illum. Eng. Soc.*, 16(1):52–61, 1987; E. E. Hammer and T. K. McGowan, Characteristics of various F40 systems at 60 Hz and high frequency, *IEEE Trans. Ind. Appl.*, IA-21:11–16, 1985; *Industry Applications Society Conference Record*, San Francisco, IEEE 82- CH-1817-6, 1982; C. Meyer and H. Nienhuis, *Discharge Lamps*, 1989.

Fluorine

A chemical element, F, atomic number 9, the member of the halogen family that has the lowest atomic number and atomic weight. Although only the isotope with atomic weight 19 is stable, the other, radioactive isotopes between atomic weight 17 and 22 have been artificially prepared. Fluorine is the most electronegative element, and by a substantial margin the most chemically energetic of the nonmetallic elements. *See* PERIODIC TABLE.

Properties. The element fluorine is a pale yellow gas at ordinary temperatures. The odor of the element is somewhat in doubt. Some physical properties are listed in the table. The reactivity of the element is so great that it will react readily at ordinary temperatures with many other elementary substances, such as sulfur, iodine, phosphorus, bromine, and most metals. Since the products of the reactions with the nonmetals are in the liquid or gaseous state, the reactions continue to the complete consumption of the fluorine, frequently with the evolution of considerable heat and light. Reactions with the metals usually form a protective metallic fluoride which blocks further reaction, unless the temperature is raised. Aluminum, nickel, magnesium, and copper form such protective fluoride coatings.

Fluorine reacts with considerable violence with most hydrogen-containing compounds, such as water, ammonia, and all organic chemical substances

1												13	14	15	16	17	18
1 H	2																2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

whether liquids, solids, or gases. The reaction of fluorine with water is very complex, yielding mainly hydrogen fluoride and oxygen with less amounts of hydrogen peroxide, oxygen difluoride, and ozone. Fluorine displaces other nonmetallic elements from their compounds, even those nearest fluorine in chemical activity. It displaces chlorine from sodium chloride, and oxygen from silica, glass, and some ceramic materials. In the absence of hydrofluoric acid, however, fluorine does not significantly etch quartz or glass even after several hours at temperatures as high as 390°F (200°C).

Fluorine is a very toxic and reactive element. Many of its compounds, especially inorganic, are also toxic and can cause severe and deep burns. Care must be taken to prevent liquids or vapors from coming in contact with the skin or eyes.

Natural occurrence. At an estimated 0.065% of the Earth's crust, fluorine is roughly as plentiful as carbon, nitrogen, or chlorine, and much more plentiful than copper or lead, though much less abundant than iron, aluminum, or magnesium. Compounds whose molecules contain atoms of fluorine are widely distributed in nature. Many minerals contain small amounts of the element, and it is found in both sedimentary and igneous rocks.

Uses. Fluorine-containing compounds are used to increase the fluidity of melts and slags in the glass and ceramic industries. Fluorspar (calcium fluoride) is introduced into the blast furnace to reduce the viscosity of the slag in the metallurgy of iron. Cryolite, Na_2AlF_6, is used to form the electrolyte in the metallurgy of aluminum. Aluminum oxide is dissolved in this electrolyte, and the metal is reduced electrically from the melt. The use of halocarbons containing fluorine as refrigerants was patented in 1930, and these volatile and stable compounds found a market in aerosol propellants as well as in refrigeration and air-conditioning systems. However, use of fluorocarbons as propellants has declined sharply because of concern over their possible damage to the ozone layer of the atmosphere. A use for fluorine that became prominent during World War II is in the enrichment of the fissionable isotope ^{235}U; the most important process employed uranium hexafluoride. This stable, volatile compound was by far the most suitable material for isotope separation by gaseous diffusion.

While consumers are mostly unaware of the fluorine compounds used in industry, some compounds have become familiar to the general public through minor but important uses, such as additives to toothpaste and nonsticking fluoropolymer surfaces on frying pans and razor blades (for example, Teflon).

Compounds. In all fluorine compounds the high electronegativity of this element suggests that the fluorine atom has an excess of negative charge. It is convenient, however, to divide the inorganic binary fluorides into saltlike (ionic lattice) nonvolatile metallic fluorides and volatile fluorides, mostly of the nonmetals. Some metal hexafluorides and the noble-gas fluorides show volatility that is frequently associated with a molecular compound. Volatility is often associated with a high oxidation number for the positive element.

Physical properties of fluorine

Property	Value
Atomic weight	18.998403
Boiling point, °C	−188.13
Freezing point, °C	−219.61
Critical temperature, °C	−129.2
Critical pressure, atm*	55
Density of liquid at b.p., g/ml	1.505
Density of gas at 0°C + 1 atm*, g/liter	1.696
Dissociation energy, kcal/mol	36.8
Heat of vaporization, cal/mol	1510
Heat of fusion, cal/mol	121.98
Transition temperature (solid), °C	−227.61

*1 atm = 101.325 kilopascals.

The metals characteristically form nonvolatile ionic fluorides where electron transfer is substantial and the crystal lattice is determined by ionic size and the predictable electrostatic interactions. When the coordination number and valence are the same, for example, BF_3, SiF_4, and WF_6, the binding between metal and fluoride is not unusual, but the resulting compounds are very volatile, and the solids show molecular lattices rather than ionic lattice structures. For higher oxidation numbers, simple ionic lattices are less common and, while the bond between the central atom and fluorine usually still involves transfer of some charge to the fluorine, molecular structures are identifiable in the condensed phases.

In addition to the binary fluorides, a very large number of complex fluorides have been isolated, often with a fluoroanion containing a central atom of high oxidation number. The binary saltlike fluorides show a great tendency to combine with other binary fluorides to form a large number of complex or double salts.

The fluorine-containing compounds of carbon can be divided into fluorine-containing hydrocarbons and hydrocarbon derivatives (organic fluorine compounds) and the fluorocarbons and their derivatives. The fluorine atom attached to the aromatic ring, as in fluorobenzene, is quite unreactive. In addition, it reduces the reactivity of the molecule as a whole. Dyes, for example, that contain fluorine attached to the aromatic ring are more resistant to oxidation and are more light-fast than dyes that do not contain fluorine. Most aliphatic compounds, such as the alkyl fluorides, are unstable and lose hydrogen fluoride readily. These compounds are difficult to make and to keep and are not likely to become very important. *See* FLUOROCARBON; HALOGEN ELEMENTS. Irving Sheft

Organic compounds. The carbon compounds containing fluorine belong to several classes, depending on what other substituents besides fluorine are present. The physical properties and chemical reactivity of organic molecules containing fluorine are quite different when compared to the same molecules containing other halogen atoms, such as chlorine. This is due, in part, to a unique combination of the properties of fluorine, which include its small atomic size

and high electronegativity. Stepwise replacement of several or all of the hydrogen atoms or other substituents attached to carbon is possible.

Many methods are available for creating a carbon-to-fluorine bond. A widely used method is to exchange a chlorine attached to carbon by reacting the compound with hydrofluoric acid. Elemental fluorine, which is very highly reactive, has also been used to prepare fluorine-containing compounds from a wide variety of organic compounds. The unusual property imparted to an organic molecule by fluorine substitution has led to the development of compounds that fulfill specific needs in refrigeration, medicine, agriculture, plastics, textiles, and other areas.

Fluoroolefins. These are a class of unsaturated carbon compounds containing fluorine; that is, they have a C═C in addition to other substituents. A typical fluoroolefin is tetrafluoroethylene ($F_2C{=}CF_2$). It is prepared from chlorodifluoromethane ($CHClF_2$), which loses HCl upon heating to produce $F_2C{=}CF_2$.

Many fluoroolefins combine with themselves or other olefins by the process of polymerization. Thus, polymerization of $F_2C{=}CF_2$ yields the polymer polytetrafluoroethylene (PTFE). This remarkable solid substance has outstanding physical and chemical properties. Nonstick polytetrafluoroethylene surfaces are used in kitchen utensils, bearings, skis, and many other applications. Since polytetrafluoroethylene is very viscous above its melting point, special methods have to be used for fabrication. For this reason, copolymers of tetrafluoroethylene with such olefins as ethylene have been developed. The chemical resistance of these copolymers is less than that of perfluorinated polymers. To obtain polymers with desired properties, the chemical processes to make them are carried out under rigorously controlled conditions. *See* COPOLYMER; POLYFLUOROOLEFIN RESINS; POLYMER; POLYMERIZATION.

There are many oxygen-containing fluorocarbons such as ethers, acids, ketones, and alcohols. Simple, fluorinated ethers are compounds of the type R-O-R, where R is a fluorinated alkyl group. The simple compound perfluoro ether (F_3COCF_3) is an analog of dimethyl ether. *See* ETHER.

Organofluorine chemicals offer some unique properties and solutions. In addition to the applications mentioned above, they are used in dyes, surfactants, pesticides, blood substitutes, textile chemicals, and biologically active compounds. *See* FLUOROCARBON; HALOGENATED HYDROCARBON. V. N. M. Rao

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. F. Shriver and P. W. Atkins, *Inorganic Chemistry*, 3d ed., 1999.

Fluorite

A mineral of composition CaF_2 and an important structure type with space group Fm3m, cell edge = 0.546 nm. The fluorite structure type is written RX_2, where the *R* cations are in cubic (eightfold) coordination by the *X* anions. Alternatively, the *X* atoms are tetrahedrally (fourfold) coordinated by the *R* atoms. The fluorite arrangement may occur when the ratio of *r*(*R*), the ionic radius of the *R* atoms, to *r*(*X*), the ionic radius of the *X* atoms, is equal to or greater than 0.73. The arrangement XR_2 is termed the antifluorite arrangement. Here the anions are eight-coordinated and the cations are four-coordinated. Halides of larger divalent cations, oxides of larger quadrivalent (4+) cations, some intermetallic compounds, and XR_2 alkali univalent oxides frequently belong to the fluorite structure type. Approximately 100 synthetic compounds are known to have this arrangement; distortions may result in similar arrangements but with lower symmetry, such as in ZrO_2, which is monoclinic. A vast number of complex structures can be derived from the fluorite arrangement by ordering vacancies and different cations or anions over the fluorite atomic positions.

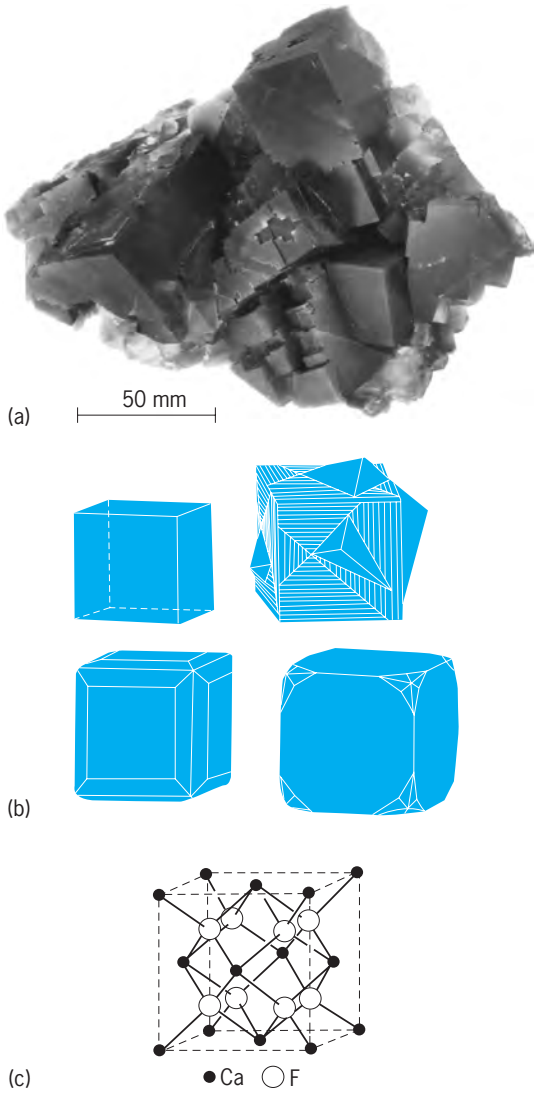

Fluorite. (*a*) Crystals from Elizabethtown, Illinois (*American Museum of Natural History specimens*). (*b*) Crystal habits (*after C. S. Hurlbut, Jr., and C. Klein, Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*). (*c*) Structure (*after C. S. Hurlbut, Jr., and C. Klein, Manual of Mineralogy, 21st ed., Wiley, 1993*).

The most abundant fluorine-bearing mineral, fluorite, occurs as cubes or compact masses and more rarely as octahedra with complex modifications (see **illus.**). The cube surfaces frequently show lineage features. Fluorite has a perfect octahedral cleavage, hardness 4 (Mohs scale), and specific gravity 3.18. The color is extremely variable, the most common being green and purple; but fluorite may also be colorless, white, yellow, blue, or brown. Colors may result from the presence of impurity ions such as rare earths or Mn^{4+}, hydrocarbon inclusions, or lattice defects. Fluorite frequently emits a blue-to-green fluorescence under ultraviolet radiation, especially if rare-earth or hydrocarbon material is present. Some fluorites are thermoluminescent; that is, they emit light when heated.

Fluorite is usually very pure, but some varieties, especially yttrofluorite, are known to carry up to 10 wt % CeF_3, the extra fluorine atoms occurring interstitially in the structure. Fluorite can be recrystallized from crushed material and used as optical glass; its more important uses are as flux material in smelting and as an ore of fluorine and hydrofluoric acid.

Fluorite occurs as a typical hydrothermal vein mineral with quartz, barite, calcite, sphalerite, and galena. Crystals of great beauty from Cumberland, England, and Rosiclare, Illinois, are highly prized by mineral fanciers. It also occurs as a metasomatic replacement mineral in limestones and marbles. Fluorite is a minor accessory mineral in many other rocks, including granites, granite pegmatites, and nepheline syenites. *See* FLUORINE. Paul B. Moore

Fluorocarbon

Any of the organic compounds in which all of the hydrogen atoms attached to a carbon atom have been replaced by fluorine; also referred to as a perfluorocarbon. Fluorocarbons are usually gases or liquids at room temperature, depending on the number of carbon atoms in the molecule. A major use of gaseous fluorocarbons is in radiation-induced etching processes for the microelectronics industry; the most common one is tetrafluoromethane. Liquid fluorocarbons possess a unique combination of properties that has led to their use as inert fluids for cooling of electronic devices and soldering. Solubility of gases in fluorocarbons has also been used to advantage. For example, they have been used in biological cultures requiring oxygen, and as liquid barrier filters for purifying air. *See* FLUORINE; HALOGENATED HYDROCARBON. V. N. M. Rao

Flutter (aeronautics)

An aeroelastic self-excited vibration with a sustained or divergent amplitude, which occurs when a structure is placed in a flow of sufficiently high velocity. Flutter is an instability that can be extremely violent. At low speeds, in the presence of an airstream, the vibration modes of an aircraft are stable; that is, if the aircraft is disturbed, the ensuing motion will be damped. At higher speeds, the effect of the airstream is to couple two or more vibration modes such that the vibrating structure will extract energy from the airstream. The coupled vibration modes will remain stable as long as the extracted energy is dissipated by the internal damping or friction of the structure. However, a critical speed is reached when the extracted energy equals the amount of energy that the structure is capable of dissipating, and a neutrally stable vibration will persist. This is called the flutter speed. At a higher speed, the vibration amplitude will diverge, and a structural failure will result. *See* AEROELASTICITY.

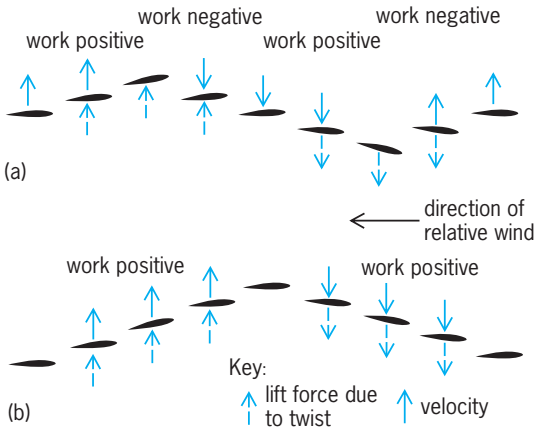

Wing displacements during a complete cycle. (*a*) Flexural and torsional displacements in phase, resulting in stable motion. (*b*) Torsional displacement lagging 90° behind flexural displacement (torsional displacement in phase with upward flexural velocity), resulting in flutter. (*After W. J. Duncan, Flutter and stability, J. Roy. Aeronaut. Soc., 53:529–549, 1949*)

The **illustration** shows examples of the displacements of a wing cross section during a complete cycle of motion. These diagrams present the wing tip in end view at intervals of one-eighth of a period, and the lift force and work done during the cycle. Part *a* illustrates the wing responding in a stable manner to a disturbance, such as a gust, below the flutter speed, and part *b* the motion during flutter. The illustration shows the physical mechanism of flutter but oversimplifies the phase relationships.

Since the first documented encounter with aerodynamic flutter of an Handley-Page bomber in 1916, the history of flutter has been determined largely by progress in unsteady aerodynamic theory, which is needed to predict both the magnitudes and phases of the aerodynamic forces acting in each of the characteristic aircraft motions. However, sufficient progress was not made until the 1930s. At that time, reliance was on flight testing for solutions to flutter problems. In 1938 a carefully planned flight flutter test of a four-engine Junkers JU 90 aircraft ended in a tragic accident in which the pilot and all the engineers aboard were killed. It became apparent that the many potential dangers associated with flutter must be anticipated before such flight tests. Since

then, greater emphasis has been placed on theoretical analysis and wind tunnel tests of dynamically scaled flutter models.

Today's aircraft manufacturers have engineering departments whose primary responsibility is flutter safety. Modern flutter analyses involve extensive computations, requiring the use of large-capacity, high-speed digital computers. Flutter engineers contribute to the design by recommending stiffness levels for the structural components and control surface actuation systems and weight distributions on the lifting surfaces, so that the aircraft vibration characteristics will not lead to flutter within the design speeds and altitudes. Examples of aircraft configuration characteristics that may be determined by flutter considerations include engine placement and fuel location and sequencing. Flutter engineers also conduct ground resonance tests to verify calculated vibration characteristics, and then monitor flight flutter tests to demonstrate adequate margins of safety for certification. Finally, these engineers specify maintenance requirements so that a satisfactory design will not deteriorate in service. Examples include rebalancing control surfaces after painting, and putting limits on hinge looseness, that is, free play, in control surfaces.

Much research involves unsteady aerodynamic theory in order to improve prediction methods. The early solutions were obtained for airfoils in two dimensions. Three-dimensional solutions, which began to appear in the 1950s, were made feasible by developments in digital computers. Accurate prediction methods are now available at subsonic and hypersonic speeds. Supersonic methods are being refined and transonic methods are being developed. Perhaps more significant is research involving structural optimization and flutter suppression. Development of filament composite materials, such as graphite/epoxy, permits tailoring of a structure to minimize its weight while providing adequate strength and stiffness to prevent flutter. Control system technology is well established for autopilot design, and has been extended to develop active and adaptive systems that will provide control surface inputs to stabilize an aircraft above the basic aircraft flutter speed. The feasibility of such automatic control systems has been demonstrated in wind tunnel tests. *See* AIRFOIL; AIRFRAME; AUTOPILOT; COMPOSITE MATERIAL; WING; WING STRUCTURE.

William P. Rodden

Bibliography. R. L. Bisplinghoff and H. Ashley, *Principles of Aeroelasticity*, 2d ed., 1962, reprint 1975; E. H. Dowell (ed.), *A Modern Course in Aeroelasticity*, 1995.

Fluvial erosion landforms

Landforms that result from erosion by water flowing on land surfaces. This water may concentrate in channels as streams and rivers or flow in thin sheets and rills down slopes. Essentially all land surfaces are subjected to modification by running water, and it is among the most important surface processes. Valleys are cut, areas become dissected, and sediment is moved from land areas to ocean basins. With increasing dissection and lowering of the landscape, the land area may pass through a series of stages known as the fluvial erosion cycle.

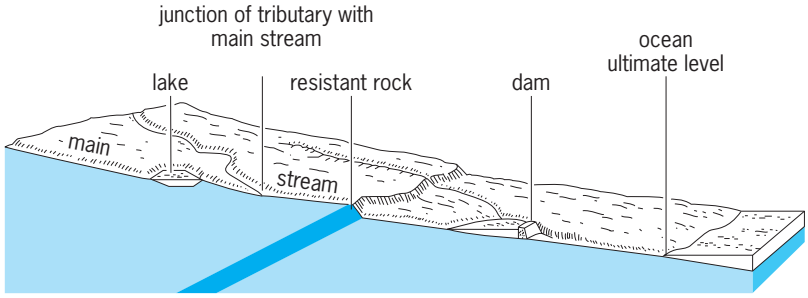

Fig. 1. Differing base levels for streams. (*After L. D. Leet and S. Judson, Physical Geology, 2d ed., Prentice-Hall, 1958*)

The lower limit of stream erosion is called base level. Utimate base level is related to sea level, and can be visualized as an imaginary surface gently inclined upward from sea level toward the interior of a land area. Local or temporary base levels exist where a resistant rock unit crosses a river, where a lake or reservoir occurs along the drainage course, and where a tributary enters a main stream (**Fig. 1**).

Fluvial system. Erosion is an integral part of the fluvial system and, along with transportation and deposition, provides the means by which rivers shape the landscape. The system is complex, and stream activity varies temporally at any one location and spatially up and down the stream course. The main characteristics of a river at a location are channel width, depth, shape and pattern, bed roughness and slope, grain size of load, and flow velocity. These are interrelated closely and are a result of the discharge and the quantity and type of load provided by the drainage basin, and the relative location of the river with respect to base level. Most rivers approach a quasiequilibrium such that the river can adjust its morphology through erosion or deposition as a result of changes in one of the controlling factors or characteristics. Such rivers are termed graded, and are both relatively stable and adjustable. During times of major adjustment the river may be either actively downcutting or aggrading. *See* RIVER; STREAM TRANSPORT AND DEPOSITION.

Fluvial erosion. Flowing water erodes material in the following ways.

1. *Direct lifting*. Loose particles on the stream bottom can be entrained and transported downstream if the stream achieves an appropriate intensity of flow. Main controlling factors are particle size, shape and density, fluid characteristics, and channel depth and slope or flow velocity.

2. *Abrasion and impact*. The solid particles entrained and transported by the stream wear down bedrock portions of the channel through impact with the bed. The bedload of the stream becomes smaller and rounder through this grinding process.

3. *Solution*. In areas of soluble rock the flowing water may dissolve portions of the channel bed. This is a relatively minor process. The enormous amounts of rock material that are carried in solution are mostly contributed to the stream by ground water.

4. *Cavitation*. This refers to the sudden collapse of vapor bubbles which may act locally as a type of hammerblow against the channel bed and be an effective erosive force. Cavitation is restricted to conditions of very great velocity and occurs mainly in waterfalls and rapids or in artificial conduits.

5. *Undercutting*. A stream that erodes laterally by one or more of the above processes may undercut the channel bank or valley wall to the extent that gravity causes the material to slump into the river, where it is moved downstream.

Once in a stream, material may be transported in solution, carried in suspension by turbulent water, or rolled, slid, or bounced along the stream bottom. It may be transported to a basin of sedimentation or, as flow conditions change, may be deposited in the channel or floodplain. *See* DELTA; FLOODPLAIN; SEDIMENTOLOGY.

Features of fluvial erosion. These may range in size from small features of the channel to large regional landscapes.

Channel and valley features. The most distinctive landform is the stream valley, which owes its origin to stream erosion. Valleys range greatly in size and shape, as do the streams that flow in them. They enlarge both through down and lateral cutting by the stream and mass wasting processes acting on the valley sides.

1. *Scour features*. The flow of water and sediment over a bedrock surface commonly leaves it smooth and partially polished as a result of abrasion and scour. The surface may contain potholes, steep-sided depressions scoured out in the bed, as well as other irregularities reflecting the current action of water or the presence of zones of weakness in the rock.

2. *Waterfalls and rapids*. Waterfalls occur where there is a sudden drop in the stream bed. This is often the case where a resistant rock unit crosses the channel and the stream is not able to erode through it at the same rate as the adjacent less resistant rock. Waterfalls also occur where a main valley has eroded down at a faster rate than its tributary valleys which are left hanging above the main stream. With time, waterfalls migrate upstream and are reduced to rapids, and eventually the steepened stream gradient is completely lost.

3. *Cut banks*. Many streams flow in a sinuous or meandering channel, and stream velocity is greatest around the outside of meander bends. Erosion is concentrated in this area, and a steep, cut bank forms. If the river meander impinges against a valley wall, the valley will be widened actively.

4. *Stream terraces*. A stream terrace (**Fig. 2**) represents a former floodplain which has been abandoned as a result of rejuvenation or downcutting by the stream. It is a relatively flat surface with a scarp slope that separates it from the current floodplain or from a lower terrace. Terraces are common features in valleys and are the result of significant changes in the stream system through time.

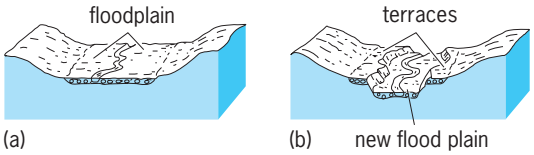

Fig. 2. Development of stream terrace. (*a*) Original floodplain. (*b*) Terrace marks level of abandoned floodplain above level of new floodplain.

Landscape features. Streams and their valleys form a drainage network which reflects the original topography and geologic conditions in the drainage basin.

1. *Drainage patterns*. A dendritic pattern, like that of a branching tree, is the most common and reflects little or no control by underlying earth materials. It develops in areas of sediment cover and flat-lying sedimentary and massive igneous rocks. Where the underlying earth materials are not uniform in resistance, streams develop in the least resistant areas, and the drainage pattern reflects the geology. If the rocks contain a rectangular joint pattern, a rectangular drainage pattern develops; if the rock units are tilted or folded, a trellis pattern of drainage is common. Topography also controls drainage development; parallel and subparallel patterns are common on steep slopes, and a radial pattern develops when streams radiate from a central high area.

2. *Slope morphology and landforms*. Slope morphology is the result of interactions among the surface processes of weathering, mass wasting and flowing water, earth materials beneath the slope, and the degree of erosional activity at the foot of the slope. Slopes in which chemical weathering and surface processes dominate are often characterized by a smooth profile with a convex upper slope and a concave lower slope with or without a straight slope element between them (**Fig. 3***a*). This type of profile is common in humid climates. If chemical weathering is relatively ineffective or erosion is active at the base of the slope, the slope will tend to be steeper and complex, and may contain a near-vertical cliff or free face (Fig. 3*b*). The latter slope element usually reflects a resistant rock unit, and is particularly common in drier areas of the world. *See* WEATHERING PROCESSES.

In areas of variable sedimentary rocks, several characteristic landforms occur depending upon the inclination or dip of the rock units. If the rocks are flat-lying, Fig. 3*b* mesas and buttes are common

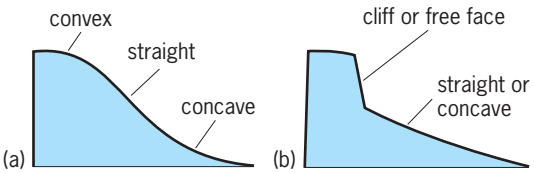

Fig. 3. Two typical slope profiles: (*a*) smooth and (*b*) steep.

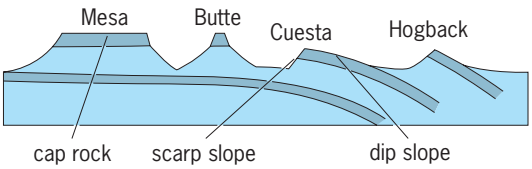

Fig. 4. Landforms developed on flat-lying and dipping sedimentary rocks. Shaded areas represent resistant rock units.

(**Fig. 4**). With a relatively small dip, an asymmetrical landform called a cuesta forms. As the dip becomes larger, the landform becomes more symmetrical and a hogback develops. These landforms may continue for long distances, forming regional escarpments. If they are dissected in a close and regular fashion by V-shaped valleys, the ridge has a triangular shape and is known as a flatiron. *See* EROSION; PLAINS. W. Hilton Johnson

Bibliography. S. A. Schumm (ed.), *Drainage Basin Morphology*, 1977; S. A. Schumm (ed.), *River Morphology*, 1972; S. A. Schumm, M. P. Mosley, and W. E. Weaver, *Experimental Fluvial Geomorphology*, 1987.

Fluvial sediments

Deposits formed by rivers. An alluvial river is one which flows within its own fluvial sediments, as distinct from one that has incised into the underlying bedrock. A river accumulates deposits because its capacity to carry sediment has been exceeded, and some of the sediment load is deposited. Rivers tend toward a state of dynamic equilibrium, in which they adjust their slope in response to changes in discharge and sediment load. The result is a channel profile that is steep in its source areas but flattens out downstream, and is graded to a slope of zero where the river discharges into a lake or the sea. Fluvial sedimentary accumulations range from temporary bars deposited on the insides of meander bends as a result of a loss of transport energy within a local eddy, to deposits tens to hundreds of meters thick formed within major valleys or on coastal plains as a result of the response of rivers to a long-term rise in base level or to the uplift of sediment source areas relative to the alluvial plain. Both these processes perturb the graded profile so that it tends to rise, creating space, or accommodation, for sediment. The same processes control the style of rivers and the range of deposits that are formed, so that a study of the deposits may enable the geologist to reconstruct the changes in controlling factors during the accumulation of the deposits. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; RIVER; STREAM TRANSPORT AND DEPOSITION.

Coarse debris generated by mechanical weathering, including boulders, pebbles, and sand, is rolled or bounced along the river bed and is called bedload. The larger particles may be moved only infrequently during major floods. Finer material, of silt and clay grade, is transported as a suspended load, and there may also be a dissolved load generated by chemical weathering. Mass movement of large volumes of sediment by sediment gravity flows, typically debris flows, may occur when rare flash floods mobilize debris that may have been accumulating in source areas for some time. Whereas the volume of sediment tends to increase downstream within a drainage system, as tributaries run together, the grain size generally decreases as a result of abrasion and selective transport. This downstream grain-size decrease may assist in the reconstruction of transport directions in ancient deposits where other evidence of paleogeography has been obscured by erosion or tectonic change. *See* MASS WASTING.

Types of river. River type may be described by two main variables, sinuosity and channel multiplicity. These variables combine to form four end-member styles, discussed below, although there are many examples of rivers showing various styles intermediate between these end members.

Braided rivers. These typically occur in areas of high sediment load and variable discharge. They consist of several or many branching, unstable channels of low sinuosity, and are characterized by abundant coarse bedload, forming bars, islands, and channel-floor deposits. The channel complex typically occupies most of the valley floor, leaving little room for a floodplain. Glacial outwash streams and ephemeral streams draining mountainous areas in arid regions are normally braided, and may form broad sheets of sand or gravel crossed by networks of shallow, shifting channels.

Meandering rivers. These are single-channel streams of high sinuosity, in which islands and midchannel bars are rare. Sediment in these rivers range from very coarse to very fine. A significant proportion of the bedload typically is deposited on the insides of meander bends, forming point bars. The channel, with its coarse deposits, may be confined to a narrow belt within an alluvial valley, flanked by a broad floodplain, upon which deposition of fine-grained sediment takes place only during flood events—seasonally or at longer intervals.

Anastomosed rivers. These develop in stable, low-energy environments or in areas undergoing rapid aggradation. They consist of a network of relatively stable, low- to high-sinuosity channels bounded by well-developed floodplains. Channels are characteristically narrow and accumulate narrow, ribbonlike sandstone bodies.

Straight channels. These are rare, occurring mainly as distributaries in some deltas.

Rivers which emerge from a mountainous catchment area into a low plain drop their sediment load rapidly. The channel may bifurcate, becoming braided in character. A distinctive landform, an alluvial fan, results.

Sedimentary facies. River deposits of sediment occur as four main types.

Channel-floor sediments. The coarsest bedload is transported at the base of the channel, commonly resulting in deposits of gravel (**Fig. 1***a*), waterlogged vegetation, or fragments of caved bank material. In

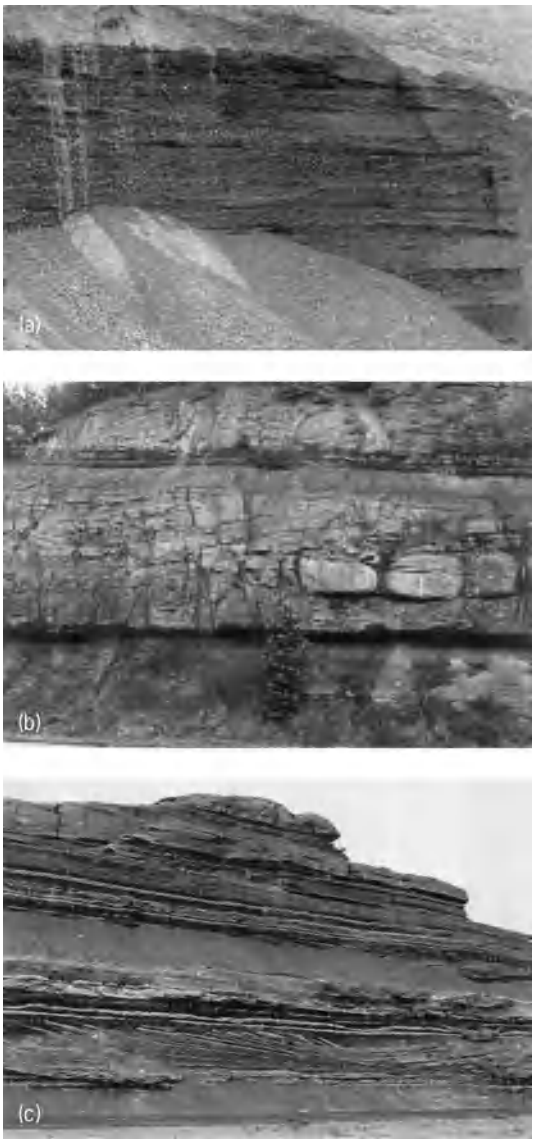

Fig. 1. Typical fluvial deposits. (*a*) Gravel and sand channel-fill and bar deposits exposed in a gravel quarry face about 40 ft (12 m) high, fluvioglacial outwash, Alberta, Canada. (*b*) Point bar, 13 ft (4 m) thick, Carboniferous, Alabama. (*c*) Typical floodplain deposits, Triassic, Arizona; outcrop is about 33 ft (10 m) high.

sand-bed rivers the channel floor commonly is covered by fields of large, sinuous-crested dunes or ripples (with amplitudes of 2 in. to 10 ft, or 5 cm to 3 m), which impart a trough–cross-bedded structure to the sand.

Bar sediments. Accumulations of gravel, sand, or silt occur along river banks and are deposited within channels, forming bars that may be of temporary duration, or may last for many years, eventually becoming vegetated and semipermanent. Bars attached to one of the channel banks are termed side or lateral bars. Those occurring on the insides of meander bends are termed point bars. They develop by lateral accretion as the meander widens or shifts in position by erosion on the outer bank of the bend (Fig. 1*b* and **Fig. 2**). Bars occurring within channels accrete by the addition of sediment on all sides, but most commonly preferentially on one side representing the inside of a bend in the adjacent channel, or at the downstream end of the bar. Such bars commonly have complex internal structures, reflecting many seasons of growth and intervals of erosion.

Channel-top and bar-top sediments. These are typically composed of fine-grained sand and silt, and are formed in the shallow-water regions on top of bars, in the shallows at the edges of channels, and in abandoned channels. Small-scale ripples, with amplitudes of less than 2 in. (5 cm), are typical sedimentary structures, together with roots and bioturbation structures.

Floodplain deposits. These are formed when the water level rises above the confines of the channel and overflows the banks (Fig. 1*c*). Much of the coarser sediment is deposited close to the channel, in the form of levees. Breaks in the channel bank, termed crevasses, permit the transportation of additional coarse sediment onto the floodplain, where it forms small deltalike bodies spreading out into the floodplain, termed crevasse splays. Much silt and mud may be carried considerable distances from the channel, forming blanketlike deposits. In swampy areas, floodplains may be the site of thick vegetation, which in time may be transformed into lignite and eventually into coal. Soils develop in response to weathering activity and plant growth, and may form distinctive brightly colored layers, termed paleosols. Nodular beds of calcium carbonate are a common component of paleosols, especially in relatively arid areas, where they form as a result of the evaporation of ground waters. *See* PALEOSOL.

Facies associations and sedimentary cycles. Fluvial sediments may be dominantly conglomeratic, sandy, or silty, depending on the nature of the sediment load of the river. This characteristic is mainly a function of slope and the proximity to sources, but is also a reflection of sediment availability. Certain source materials, such as fine-grained sediments or limestones, may yield coarse debris on erosion, but it

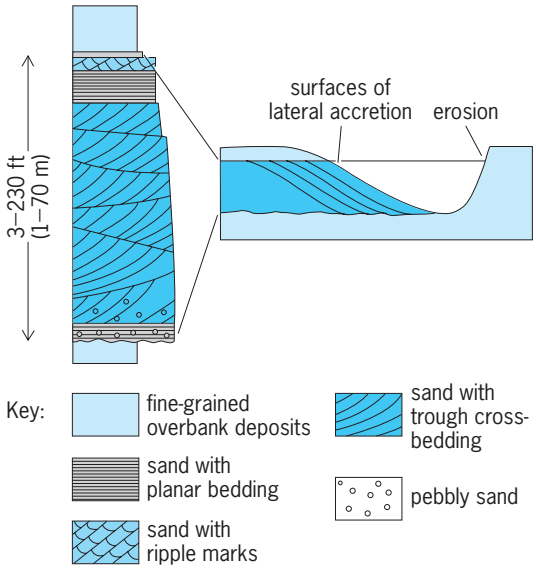

Fig. 2. Development of a fining-upward succession by lateral accretion of a point bar, such as that in Fig. 1*b*.

is likely to be broken down into fine material or dissolved on prolonged transportation. Humid climates favor chemical and biochemical weathering processes, which yield a large suspended or dissolved sediment load. Coarse detritus is more typically the product of drier climates, in which mechanical weathering processes (such as frost shattering) are dominant.

The sedimentary facies described above were listed in approximate vertical spatial order, from channel floor to floodplain. This order is one of decreasing grain size upward, a feature which may commonly be observed in ancient fluvial deposits (Fig. 2). Such deposits may consist of a series of fining-upward successions, or cycles, each a few meters to few tens of meters in thickness.

There are a variety of causes of such cycles. The first that was recognized is the mechanism of lateral accretion, whereby point bars enlarge themselves in a horizontal direction as the meander bounding them migrates by undercutting the bank on the outside of the bend (Fig. 2). The depositional surface of the point bar may be preserved as a form of large-scale, low-angle cross-bedding within the deposit, its amplitude corresponding approximately to the depth of the channel. Similar cycles may be caused by the nucleation and growth of large compound bars or sand flats within braided channel systems. The accretion surfaces, in such cases, may dip in across- or down-channel directions. Individual flood events, especially on the sand flats of ephemeral stream systems, may form sheetlike flood cycle deposits up to a meter or so thick, the upward fining corresponding to decreasing energy levels as the flood waned. The gradual choking of a channel with sediment, and the progressive abandonment of the channel, will also generate a fining-upward cycle.

Regional controls. Tectonic activity in a fluvial catchment area may cause the generation of sedimentary cycles, which either fine or coarsen upward, depending on whether relief and slope are decreased or increased, respectively. Such cycles tend to be tens to hundreds of meters thick and to extend for several to many kilometers. They may have smaller cycles formed by channel fill and migration processes nested within them.

River systems are also affected by changes in base level, that is, by a rise or fall in the level of the lake or sea into which the river drains. A fall in base level may lead to widespread incision of channels along a coastal plain as they adjust to a lower river mouth. Between channels, sedimentation may cease, with the formation of widespread, well-developed paleosols. The same effect is brought about by peneplanation, that is, long-continued subaerial erosion in the absence of tectonic rejuvenation of the river system.

A rise in base level may flood the mouths of rivers, forming estuaries. However, if the sediment supply is adequate, sedimentation may be able to keep pace with such a base-level rise, with the river changing in style in response to changes in the balance between the rate of sediment input and the rate of base-level change (**Fig. 3**). These sedimentary responses

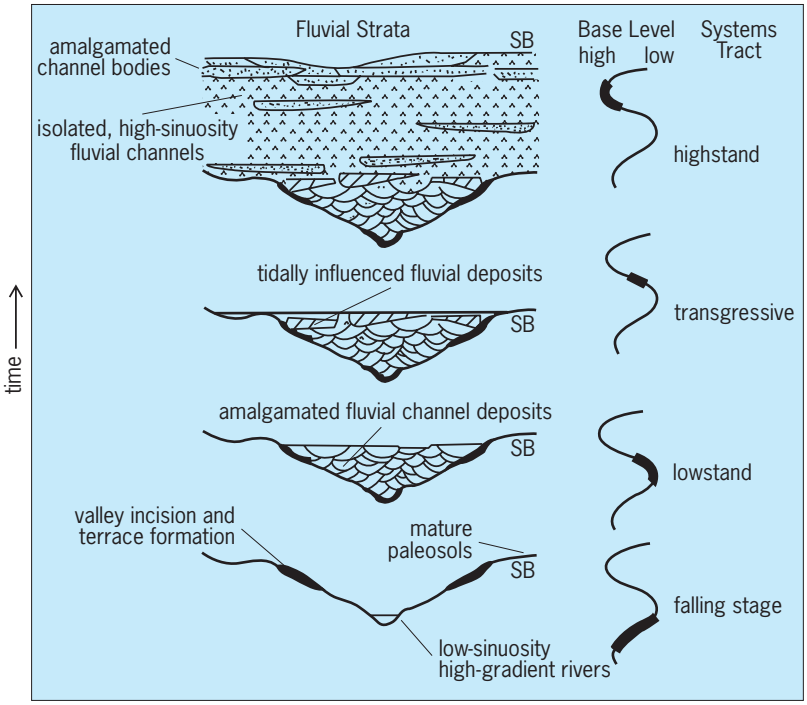

Fig. 3. Response of river systems and their deposits to a cycle of fall and rise of base level. SB = sequence boundary. (*Modified from K. W. Shanley and P. J. McCabe, Perspectives on the Sequence Stratigraphy of Continental Strata, American Association of Petroleum Geologists, vol. 78, p. 560, 1994*)

to external forcing are part of a larger story concerning the regional and global controls of sedimentation. The regional erosion surface formed at a time of falling to low base level or deep peneplanation is termed a sequence boundary. It may be cut by incised channels that are, in turn, typically filled by coarse channel sediments (in sequence stratigraphic terminology, these deposits are classified as the lowstand systems tract). At the coast a rise in base level is commonly recorded by transgression (the transgressive systems tract). The coastal rivers at this time may be of anastomosed style. As base-level rise reaches its highest level, the rate of rise slows, and the rivers typically evolve into a meandering type (highstand systems tract). These changes may be reflected in the resulting sediments by changes in the geometry and spacing of channel sandstone and conglomerate bodies. Lowstand deposits commonly consist of coarse, laterally amalgamated channel deposits. A rapid rise in base level (the transgressive systems tract) may be marked by isolated channel sands spaced out within thick floodplain units. The spacing of such channel bodies becomes closer, and more units are amalgamated, in the highstand deposits, above. *See* FLOODPLAIN; SEQUENCE STRATIGRAPHY.

Far inland, base-level changes may not markedly affect the rivers unless they persist for very long periods of time. Changes in discharge and sediment yield in response to climate change are commonly more important in controlling river style and the resulting sediment types. For example, during the Pleistocene glaciation, rapid deposition of coarse sediments occurred at the margins of continental ice caps.

Channel erosion, forming widespread surfaces of incision, occurred at times of change, between glacial and interglacial periods, because at these times river discharge tended to increase whereas sediment yield did not. Climatically driven erosion and deposition inland were therefore out of phase with the cycle of change generated by base-level change at the coast.

More than one forcing function, including climate change, base-level change, and tectonism, may be operating at any one time, resulting in complex patterns of cyclicity. The resulting sequences may be widespread. Reconstruction of this sequence stratigraphy may provide an essential mapping tool for those engaged in basinal exploration.

Tectonic setting of fluvial deposits. The thickest (up to 6 mi or 10 km) and most extensive fluvial deposits occur in convergent plate-tectonic settings, including regions of plate collision, because this is where the highest surface relief and consequently the most energetic rivers and most abundant debris are present. Some of the most important accumulations occur in foreland basins, which are formed where the continental margin is depressed by the mass of thickened crust formed by convergent tectonism. Examples include the modern Himalayan foreland basin of the Indus and Ganges valleys, the Devonian foreland basin west of the Appalachian Mountains, and the late Cenozoic foreland basin of France and Germany, north of the alpine mountain chain. *See* BASIN.

Thick fluvial deposits also occur in rift basins, where continents are undergoing stretching and separation. The famous hominid-bearing sediments of Olduvai Gorge and Lake Rudolf are fluvial and lacustrine deposits formed in the East Africa Rift System. Triassic fault-bounded basins along the North American Atlantic coast and through western Europe are an older but comparable example. Fluvial deposits are also common in wrench-fault basins, such as those in California.

Economic importance. Significant volumes of oil and gas are trapped in fluvial sandstones. Major reservoirs include those of Triassic-Jurassic age in the North Sea Basin; Triassic sandstones of the Paris Basin; Permian-Triassic sandstones of Prudhoe Bay on the Alaskan North Slope; the Lower Cretaceous reservoirs of the giant Daqing field of the Songliao Basin, China; Jurassic sandstones of interior Australia; the heavy-oil sands of the Cretaceous Athabasca and related deposits in Alberta; and numerous large to small fields in mature areas such as the Alberta Basin (Cretaceous), the southern midcontinent (Pennsylvanian), and the Gulf Coast (Cretaceous and Oligocene fields).

Placer gold, uranium, and diamond deposits of considerable economic importance occur in the ancient rock record in South Africa and Ontario, Canada, and in Quaternary deposits in California and Yukon Territory. Economically significant roll-front uranium deposits occur in the Mesozoic deposits of the American Western Interior and elsewhere, primarily in fluvial facies.

Fluvial deposits are also essential aquifers, especially the postglacial valley-fill complexes of urban Europe and North America. Much work needs to be done to investigate the internal geometry of these deposits in order to resolve problems of domestic and industrial pollution that now interfere with the use of the ground water from these sources.

Andrew D. Miall

Bibliography. P. A. Carling and M. R. Dawson (eds.), *Advances in Fluvial Dynamics and Stratigraphy*, Wiley, 1996; C. R. Fielding (ed.), Current research in fluvial sedimentology, *Sed. Geol.*, special issue, vol. 85, 1993; A. D. Miall, *The Geology of Fluvial Deposits*, Springer-Verlag, 1996.

Fluxional compounds

Molecules that undergo rapid intramolecular rearrangements among equivalent structures in which the component atoms are interchanged. The rearrangement process is usually detected by nuclear magnetic resonance (NMR) spectroscopy, which can measure rearrangement rates from 0.5 to 10,000 s^{-1}. With sufficiently rapid rates, a single resonance is observed in the NMR spectrum for a molecule that might be expected to have several nonequivalent nuclei on the basis of its instantaneous structure.

Organic structures. Within organic chemistry, degenerate Cope rearrangements represented some of the first examples of interconversions between equivalent structures, but these were relatively slow (**Fig. 1***a*). The rate of this rearrangement is rapid in more complex molecules. The epitome of degeneracy is reached in bullvalene, which has more than 1,200,000 equivalent structures (Fig. 1*b*) and rapidly interconverts among them.

Fluxional molecules are frequently encountered in organometallic chemistry, and rapid rearrangements which involve migrations about unsaturated organic rings are commonly observed. The best known (called ring-whizzers) are cyclopentadienyl and cyclooctatetraene complexes of iron (**Fig. 2**).

Inorganic structures. Inorganic structures also exhibit fluxional phenomena, and five-coordinate complexes provide the greatest number of well-known

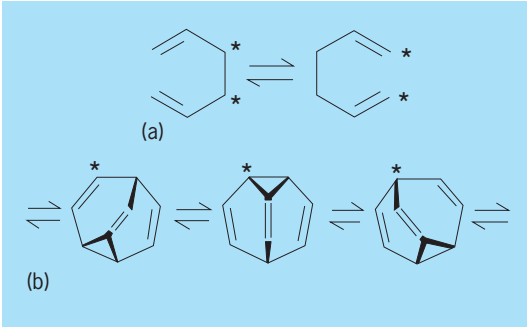

Fig. 1. Degenerate rearrangements. (*a*) Cope rearrangement. (*b*) Rearrangement of bullvalene. The asterisks indicate labels for the atoms so that they can be distinguished from one another.

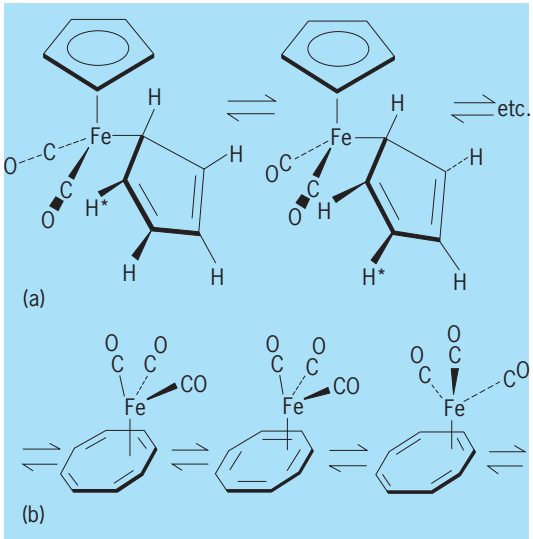

Fig. 2. Rearrangements about unsaturated organic rings. (*a*) Cyclopentadienyl iron complex. (*b*) Cyclooctate-traene iron complex.

examples. Although the static structure would suggest a significant difference between equatorial and axial fluorine nuclei, phosphorus pentafluoride rearranges rapidly and shows only a single ^{19}F resonance split by P-F spin-spin coupling even though the static structure would suggest a significant difference between equatorial and axial fluorines. The retention of this coupling during the rearrangement demands that the process be intramolecular. This distinguishes a fluxional process from other intermolecular exchange processes which might give rise to dynamic NMR spectra. Detailed line-shape analysis of the NMR spectra can also provide insight into the possible paths of the rearrangements. Thus, the ring-whizzers have been shown to rearrange preferentially by moving to an adjacent position rather than across the ring.

The rearrangement of PF_5 involves interconversion of the trigonal bipyramidal molecule (**1**) to a square pyramidal configuration (**2**) and back (**Fig. 3**). The square pyramidal structure is of sufficiently high energy that none of that structure is observed by spectroscopic techniques that operate on a faster time scale than does NMR. If two such nonequivalent structures are present in

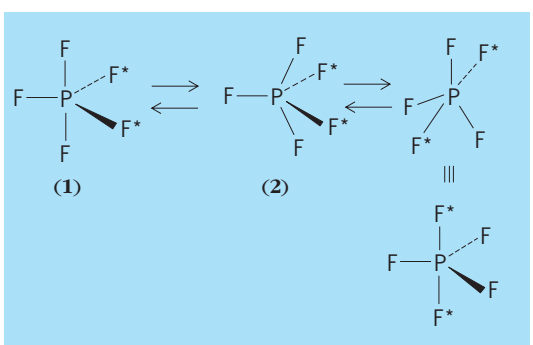

Fig. 3. Rearrangement of PF_5. Structure (2) interconverts to the equivalent structure on the right.

observable concentrations but interconvert rapidly to cause averaging in the NMR experiments, they are said to be stereochemically nonrigid. This term is generally taken to embrace all compounds that undergo rapid reversible intramolecular rearrangements. Thus, fluxional compounds are a subset of nonrigid compounds with equivalent structures. Nonequivalent structures, that is, tautomers, might be stereochemically nonrigid if they rearranged rapidly, but would not be considered fluxional.

Some workers prefer to reserve the term fluxional for molecules in which bonds are broken and reformed in the rearrangement process. Hence, of the examples above, only bullvalene and the iron complexes would be termed fluxional, whereas all would be considered stereochemically nonrigid. *See* NUCLEAR MAGNETIC RESONANCE (NMR); RESONANCE (MOLECULAR STRUCTURE); TAUTOMERISM.

J. W. Faller

Bibliography. C. J. Ballhausen, *Molecular Electronic Structures of Transition Metal Complexes*, 1979; F. A. Cotton, Fluxional organometallic molecules, *Acc. Chem. Res.*, 1:257–265, 1968; W. von E. Doering and W. R. Roth, Thermal rearrangements, *Angew. Chem.* (Internat. Ed.), 2(3):115–122, 1963; J. W. Faller, Fluxional and nonrigid behavior of transition metal organometallic π-complexes, *Adv. Organometal. Chem.*, 16:211–239, 1977; L. M. Jackman and F. A. Cotton, *Dynamic Nuclear Magnetic Resonance Spectroscopy*, 1975.

Flywheel

A rotating mass used to maintain the speed of a machine between given limits while the machine releases or receives energy at a varying rate. A flywheel is an energy storage device. It stores energy as its speed increases and gives up energy as the speed decreases. The specifications of the machine usually determine the allowable range of speed and the required energy interchange.

Theory. The energy, speed, and size of a flywheel are related by Eq. (1) for use with SI units, and by Eq. (2) for American-British engineering units. The term $n^2{}_1 - n^2{}_2$ may be replaced by the product $2c_f n^2$,

$$Mk^2 = \frac{2E}{n_1 - n_2^2} \tag{1}$$

$$Wk^2 = \frac{182.4gE}{n_1^2 - n_2^2} \tag{2}$$

where W = flywheel weight (lb)
M = flywheel mass (kg)
E = energy change (joules or ft · lb)
k = radius of gyration of flywheel (m or ft)
n_1, n_2 = maximum and minimum velocity (angular) of flywheel (radians/s or revolutions/min)
g = gravity constant (32.174 ft/s^2)

which gives Eqs. (3) and (4). The coefficient of fluc-

$$Mk^2 = \frac{E}{c_f n^2} \tag{3}$$

$$Wk^2 = \frac{91.2gE}{c_f n^2} \tag{4}$$

tuation is given by Eq. (5), and the average speed or angular velocity is given by Eq. (6).

$$c_f = \frac{\text{coefficient of}}{\text{fluctuation}} = \frac{n_1 - n_2}{n} \tag{5}$$

$$n = \frac{\text{average speed}}{\text{or angular velocity}} = \frac{n_1 + n_2}{2} \tag{6}$$

The energy interchange cycle of a flywheel may occupy any amount of angular motion about its center of rotation. Most frequently it is a cycle of one or more revolutions, but in internal combustion engines with multiple cyclinders, it may be only a small part of a revolution.

All rotating components of a machine contribute to the flywheel effect and should be considered in a complete analysis of the design. The term Mk^2 or Wk^2 in Eqs. (1)–(4) is the total flywheel effect of all rotating components and is made up of the sum of the Wk^2 or Mk^2 for all such components as shown in Eqs. (7) and (8), where subscript f refers to the

$$M_1k_1^2 + M_2k_2^2 + \cdots + M_fk_f^2 = \sum M_1k_1^2 + M_fk_f^2 \tag{7}$$

$$W_1k_1^2 + W_2k_2^2 + \cdots + M_fk_f^2 = \sum Wk^2 + M_fk_f^2 \tag{8}$$

flywheel and subscripts 1 and 2, to the other rotating components.

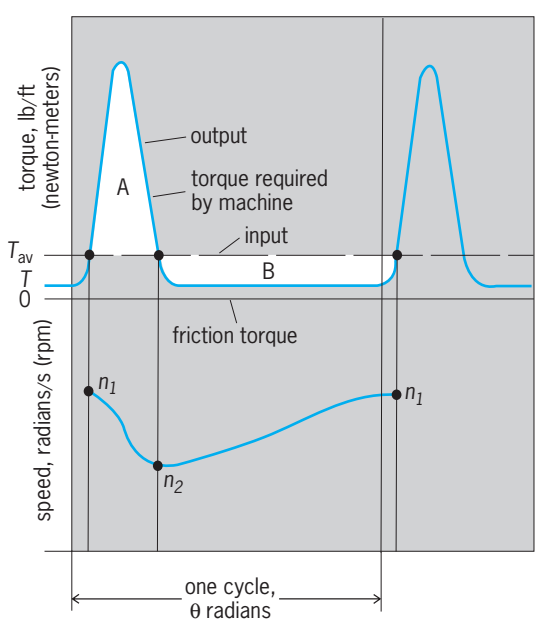

Fig. 1. Determination of energy change for a flywheel for a typical power-press torque requirement.

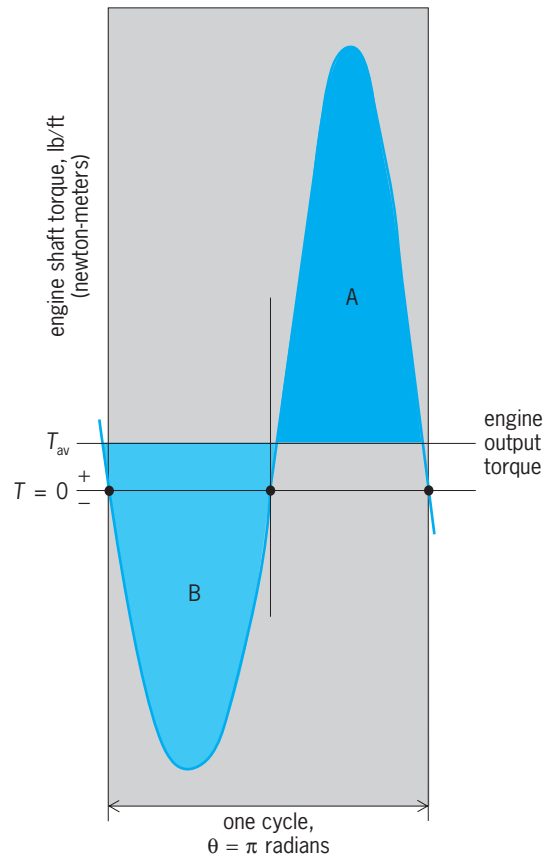

Fig. 2. Determination of energy change for a flywheel for a typical four-cylinder, four-cycle internal-combustion-engine torque requirement.

Equations (3) and (4) can now be written in more useful terms for the designer as Eqs. (9) and (10).

$$M_fk_f^2 = \frac{E}{c_f n^2} - \sum Mk^2 \tag{9}$$

$$W_fk_f^2 = \frac{91.2gE}{c_f n^2} - \sum Wk^2 \tag{10}$$

Flywheel design. In the design of a flywheel the first step is to determine the energy interchanged. This can be done best by plotting the torque required by the machine against the angular motion of the flywheel for one cycle of θ radians (**Figs. 1** and **2**). In this plot, the area under the torque curve represents the total energy requirements of the machine during the cycle. Then, average torque T_{av} defined as the total energy of the cycle E_{tot} divided by the length of the cycle θ in radians is drawn on the same torque scale. The areas between the original torque curve and the average torque curve above and below this average torque (shaded areas in Figs. 1 and 2) should be equal and are the E, or change in energy, in Eqs. (1) and (2) or (9) and (10).

The second step is to establish the speed range n_1 and n_2 of the machine or to select the appropriate coefficient of fluctuation from the **table**.

The third step is to calculate the flywheel effect $M_fk^2_f$ or $W_fk^2_f$ with Eqs. (9) and (10), calculating

Coefficients of fluctuation

Machine	C_f
Crushers, hammers, punches, shears	0.100–0.200
Machine tools	0.020–0.030
Electric generators	0.003–0.005
Compressors, pumps, and engines	0.040–0.050
Textile machinery	0.010–0.030

ΣMk^2 or ΣWk^2 of all the significant rotating components. The value of c_f is taken from the table or calculated by Eqs. (5) and (6). The value of E is established from Figs. 1 and 2: area A = area B = energy change E. The angular velocity (rotational speed) should come from the machine specifications. Having thus established numerical values for $M_f k^2_f$ or $W_f k^2_f$, the designer is now in a position to calculate the diameter and width of the flywheel to meet the flywheel effect requirements. The designer must vary all the variables that affect mass or weight to fit the space, speed, and safety requirements of the machine.

Flywheel construction. The difficulty of casting stress-free spoked flywheels leads the modern designer to use solid web castings or welded structural steel assemblies. For large, slow-turning flywheels on heavy-duty diesel engines or large mechanical presses, cast spoked flywheels of two-piece design are standard (**Fig. 3**).

Because it is difficult to calculate the stresses in a spoked flywheel, these are usually designed on the basis of a maximum allowable peripheral speed of 5000–7000 ft/min (1500–2100 m/min) for cast iron. The higher value is for rims cast by special techniques to eliminate rim porosity or blowholes. Higher rim speeds to 10,000 ft/min (3000 m/min) may be used with cast steel. Structural steel plate welded flywheels of the solid web type can be accurately analyzed for stress by applying the equations of rotating disks to the components of the flywheel and allowing a factor for the stresses caused by welding.

Unusual flywheel applications. Flywheels of appropriate design can be used to store relatively large quantities of energy for use at a later time away from their spin-driving energy source. To store large quantities of energy in a reasonably small space requires flywheel designs that can spin at very high speeds up to 20,000–30,000 revolutions per minute. They must be enclosed in a near-perfect vacuum to reduce air friction losses and be equipped with very high-efficiency transmission systems. These high-speed flywheels must also be designed with extremely high-tensile-strength materials to withstand the stresses produced at their maximum speed. *See* ANGULAR MOMENTUM; ENERGY; INERTIA; MASS; MOMENT OF INERTIA; MOMENTUM.

L. Sigfred Linderoth

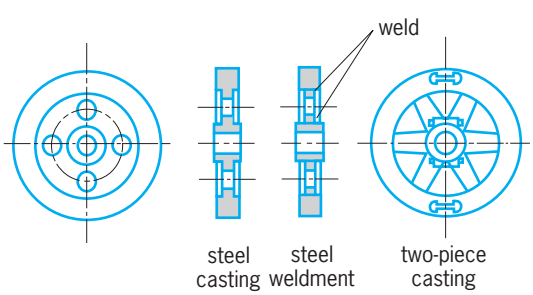

Fig. 3. Typical flywheel structures.

Bibliography. R. C. Johnson, J. E. Shigley and C. R. Mischke, *Mechanisms: A Mechanical Designer's Workbook*, 1990; A. Parrish (ed.), *Mechanical Engineer's Reference Book*, 12th ed., 1994.

Foam

A material composed of gas bubbles separated from one another by liquid or solid sheets (or lamellae). Foam densities range from nearly zero to about 43 lb/ft^3 (0.7 g/cm^3), above which gas emulsions rather than foams are formed. The bubbles are spherical when the lamellae that separate the bubbles are thick (approximately 0.04 in. or 1 mm); the bubbles are polyhedral when the lamellae are thin (approximately 0.0004 in. or 0.01 mm). Pure liquids do not foam; that is, they cannot produce liquid lamellae of any permanence. Relatively permanent lamellae are created only when a substance is present that is adsorbed at the surface of the liquid. Such substances may be in true solution in the liquid, or they may be particles of a finely divided solid which, because of poor wetting by the liquid, remain at the surface. In both cases, surface layers rich in the added substance are formed. The inability of the adsorbed substance to enter the bulk of the liquid preserves the surface and so confers stability on the foam lamellae. However, this stability is transient. Gravity drains solution out of the lamellae, and the solubility of gas in the solution, although slight, allows gas to diffuse out of the foam. These destabilizing factors tend to separate the liquid and the gas phases of the foam and so destroy it. *See* ADSORPTION; EMULSION; INTERFACE OF PHASES.

On the other hand, the stability of the foam is enhanced if the adsorbed layer is thick and gelatinous. These attributes confer two stabilizing properties on the liquid lamellae: resilience and plasticity. *See* SURFACTANT.

Resilience and plasticity of lamellae. The resilience, or elastic recovery after externally induced mechanical stress, of a liquid lamella originates in the higher surface tension of a freshly formed surface of any solution that contains a surface-active solute. This occurs because such a surface contains relatively few solute molecules. Then, as solute molecules are adsorbed, the surface tension declines. The higher dynamic surface tension of a nascent surface, if it lasts long enough, can heal any incipient breach by drawing toward it the surrounding surface of lower surface tension. Drawn along with it, by viscous drag, is a considerable amount of subsurface solution. This restoring mechanism depends on the liquid lamella being sufficiently thick to provide enough subsurface solution. Prolonged drainage impairs resilience because of extreme depletion of subsurface solution. Aged lamellae, defective in this

healing mechanism, are brittle. A lack of resilience is the cause of the ultimate collapse of foams. *See* SURFACE TENSION.

An additional factor, plasticity of the surface, can prolong stability from mere seconds to several hours. Solutions with this latter property are perfectly fluid in their flow properties as bulk liquids, but the surface acts as if it were a thin sheet of a soft plastic material. The plastic layer has a depth of 10–40 micrometers, so that a lamella, which is plastic at both surfaces, may contain very little interstitial solution that is free to drain out. The lamellae become totally plastic, and the foam is correspondingly long-lived. The presence of water-soluble proteins, such as casein or albumin, provides stability in foams produced with heavy cream, egg white, beer, and natural rubber latex. Surface plasticity adds enormously to the stability of foams, as exemplified in meringue, whipped cream, fire-fighting foams, and shaving foams. Finely divided solids may have the same effect, as exemplified in the industrial process of mineral separation by means of froth flotation.

Resilience and plasticity of lamellae are two distinct properties, corresponding to the observation that there are two distinct types of foams, which differ considerably in stability. Ephemeral or transient foams (such as champagne foam) arise from resilience alone. Such foams drain rapidly to a critical thickness of 20–30 nanometers, at which point the lamellae are brittle and soon rupture. The very stable foams (such as beer foam) begin in the same way, and instantaneous rupture is prevented by resilience. But this effect is soon replaced by the slow growth of gelatinous surface layers, which serve to preserve the lamellae. At the extreme, the lamellae may be stabilized permanently by polymerization of the medium.

Plastic foams. Plastic foams (also known as cellular, expanded, or sponge plastics) are solid foams in which the gas can be considered as a lightweight filler or extender of the plastic phase. Plastics commonly used are polyurethane, synthetic rubber, polystyrene, polyvinyl chloride, and polyolefins. In all cases the foam is produced while the medium, a monomer or a blend of monomers, is still fluid. Polymerization of the medium occurs before the foam has had time to collapse substantially. The nature of the polymer and the foam density are dominant factors in determining the physical properties of cellular plastics. Practical foam densities may vary from about 0.5 to 60 lb/ft^3 (0.008 to 0.96 g/cm^3) or almost up to the density of the solid plastic. If the foam cells are interconnected so that gas can pass freely among them, the foam is termed open-cell; if the gas is sealed in each cell, the foam is termed closed-cell. Flexible foams are usually open-cell; rigid foams are usually closed-cell. The major applications of foamed plastics are for insulation and packaging (rigid foams) and comfort products (flexible foams). *See* POLYMER.

Elimination. Although foams of exceptional stability are desired in some industrial applications, foam is a nuisance in many other situations. In the glue and paper industries, in the processing of foodstuffs, in sewage disposal, in the dehydration of oil, and in distillation and fractionation towers, foams cause undesirable complications. A common recourse is to add chemical antifoam agents, which are usually insoluble liquids of very low surface tension. When a droplet of such a liquid is sprayed onto the foam or is carried into it by mechanical agitation, it behaves as a local area that resists wetting by the medium. This results in a mechanical action, that is, the retraction of the lamella from the insoluble droplet situated within it; the process is known as dewetting. The force of its mechanical action— or, as some conceive it, the bridging of the liquid lamella by a droplet whose adhesion to the medium is too weak to hold the composite structure together—causes it to collapse. *See* COLLOID. Sydney Ross

Bibliography. K. C. Frisch and J. H. Saunders (eds.), *Plastic Foams*, 1972–1976; D. Myers, *Surfactant Science and Technology*, 2d ed., 1992; S. Ross and I. D. Morrison, *Colloidal Systems and Interfaces*, 1988.

Focal length

A measure of the collecting or diverging power of a lens or an optical system. Focal length, usually designated f' in formulas, is measured by the distance of the focal point (the point where the image of a parallel entering bundle of light rays is formed) from the lens, or more exactly by the distance from the principal point to the focal point. *See* GEOMETRICAL OPTICS.

The power of a lens system is equal to n'/f', where n' is the refractive index in the image space (n' is usually equal to unity). A lens of zero power is said to be afocal. Telescopes are afocal lens systems. *See* DIOPTER; LENS (OPTICS); TELESCOPE. Max Herzberger

Fog

A cloud comprising waterdroplets or (less commonly) ice crystals formed near the ground and resulting in a reduction in visibility to below 0.6 mi (1 km). This is lower than that occurring in mist, comprising lower concentration of waterdroplets, and haze, comprising smaller-diameter aerosol particles.

Composition. Fog results from the cooling of moist air below its saturation (dew) point. Droplets form on hygroscopic nuclei originating from ocean spray, combustion, or reactions involving trace chemicals in the atmosphere. Visibility is reduced even more when such nuclei are present in high concentrations and faster cooling rates activate a larger fraction of such nuclei. Thus, polluted fog, with more numerous smaller droplets, results in lower visibility for a given water content. *See* DEW POINT.

Haze, the precursor to fog and mist, forms at relative humidity below 100% to about 80%. It is composed of hygroscopic aerosol particles grown by

absorption of water vapor to a diameter of about 0.5 micrometer, concentration 1000 to 10,000 per cubic centimeter. Fog and mist form as the relative humidity increases just beyond saturation (100%), so that larger haze particles grow into cloud droplets with a diameter of 10 μm and a concentration of several hundred per cubic centimeter. Fog and mist are a mix of lower-concentration cloud droplets and higher-concentration haze particles. By contrast, smog is formed of particles of 0.5–1-μm diameter, produced by photochemical reactions with organic vapors from automobile exhaust. *See* ATMOSPHERIC CHEMISTRY; HUMIDITY; SMOG.

Formation. The saturated vapor pressure of water declines with falling temperature; as air containing water vapor is cooled below its dew point, excess vapor condenses to droplets. Fog forms in this way by cooling that results from several processes.

Radiation fog forms under near calm and clear nighttime conditions by cooling, which results from radiation loss to the cold sky. Such fog often appears first as shallow layers a few feet thick just above moist ground (**illus.** *a*). Both the surface of the ground and moist air cool by radiation, resulting in vapor deposits as fog droplets above the ground and, very close to the surface, vapor deposits on the ground or on vegetation as dew. Subsequently, slight air motion, resulting from the settling of denser, colder air results in local lifting of some air, with fog formation at a slightly higher level following expansion cooling. This enables the fog to build upward and reduces radiation cooling at lower levels. Such motions can also lead to fog formation by the mixing of saturated air parcels at different temperatures (analogous to a person's breath on a cold morning), leading to the phenomenon of "Arctic frost" or sea smoke, that is, shallow wispy layers of fog just above a warmer water surface. *See* DEW.

Advection fog forms when moist air moves over a colder surface, such as snow or hail that has recently fallen, or cold water, as near Newfoundland. Both radiation and advection fog build to form deeper layers of lifted fog, which are then transformed to stratocumulus clouds (illus. *b*) when mixed by increasing winds aloft or by solar heating as the Sun rises. This process ultimately leads to "burn off" of the fog over land and low clouds. *See* CLOUD.

Hill fog forms when air flows over high ground and cools by expansion; it is in many ways similar to clouds formed by convective processes.

Supercooled fog forms at temperatures below 32°F (0°C). This fog may be cleared over a localized area by artificial cloud seeding, providing the temperature is below about 25°F (−4°C). *See* WEATHER MODIFICATION.

At temperatures below about −31 to −40°F (−35 to −40°C), all waterdroplets freeze and an ice fog forms. This is particularly troublesome in communities experiencing such temperatures, because as the temperature falls, nuclei and moisture produced by combustion form water fog first and then ice fog.

Deep fogs may give a light drizzle. Viewed from above, the presence of droplets of intermediate size (50 μm) that coalesce to drizzle is seen as a fog bow, a white, almost colorless arc at the rainbow angle (42°) [illus. *c*]. *See* CLOUD PHYSICS; RAINBOW.

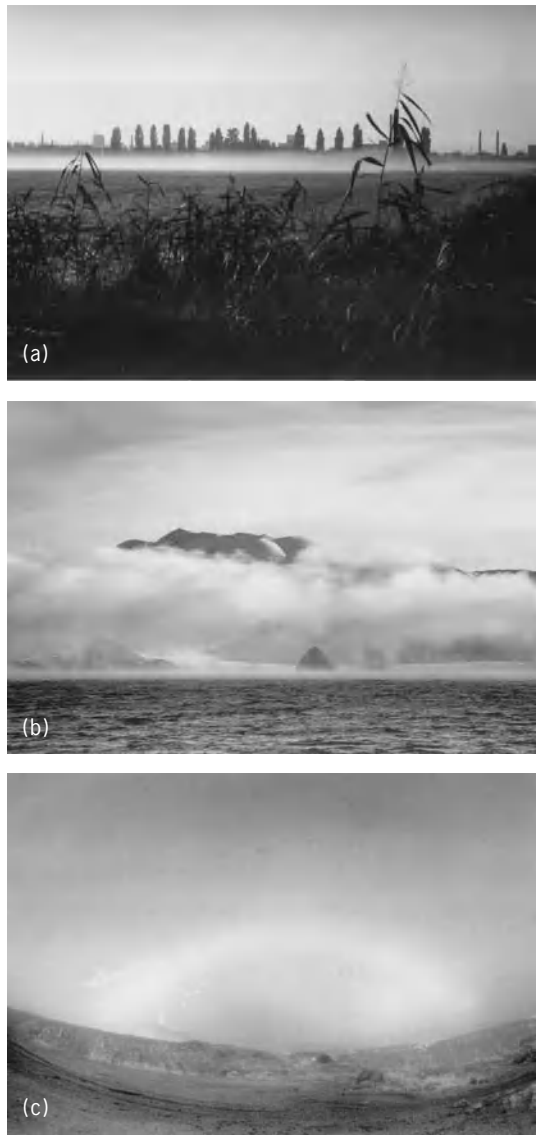

Fogs formed by various processes. (*a*) Radiation fog; appears first as shallow layers above moist ground. (*b*) Stratocumulus cloud, formed from lifted fog. (*c*) Fog bow resulting from large cloud droplets in the fog top.

John Hallett

Bibliography. F. K. Lutgens and E. J. Tarbuck, *The Atmosphere: An Introduction to Meteorology*, 6th ed., 1995; S. Twomey, *Atmospheric Aerosols*, 1977.

Fold and fold systems

Folds are recognized where layered rocks have been distorted into wavelike forms. Some folds are fractions of an inch across and have lengths measured in inches, whereas others are a few miles wide and tens of miles long.

Elements of folds. The axial surface (**Fig. 1**) divides a fold into two symmetrical parts, and the intersection of the axial surface with any bed is an axis. In

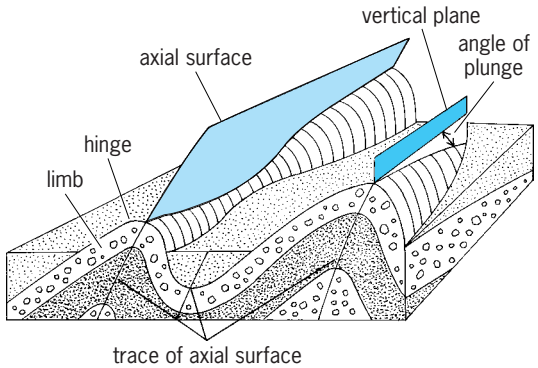

Fig. 1. Elements of folds.

general, an axis is undulatory, its height changing along the trend of the fold. Relatively high points on an axis are culminations; low points are depressions. The plunge of a fold is the angle between an axis and its horizontal projection. The limbs or flanks are the sides. A limb extends from the axial surface of one fold to the axial surface of the adjacent fold. Generally, the radius of curvature of a fold is small compared to its wavelength and amplitude, so that much of the limb is planar. The region of curvature is the hinge. *See* ANTICLINE; SYNCLINE.

Geometry of folds. The geometry of folds is described by the inclination of their axial surfaces and their plunges (**Fig. 2**). Upright folds have axial surfaces that dip from 81° to 90°; inclined folds have axial surfaces that dip from 10° to 80°; and recumbent folds have axial surfaces that dip less than 10°. Vertical folds plunge from 81° to 90°; plunging folds plunge from 10° to 80°; and horizontal folds plunge less than 10°. Auxiliary descriptive terms depend on the attitude or the relative lengths of the limbs. Overturned folds are inclined folds in which both limbs dip in the same direction; isoclinal folds are those in which both limbs are parallel; symmetrical folds have limbs of equal length; and asymmetrical folds have limbs of unequal length. The descriptions of folds consist of combinations of the above terms, for example, isoclinal upright horizontal fold, overturned plunging fold, asymmetrical inclined horizontal fold.

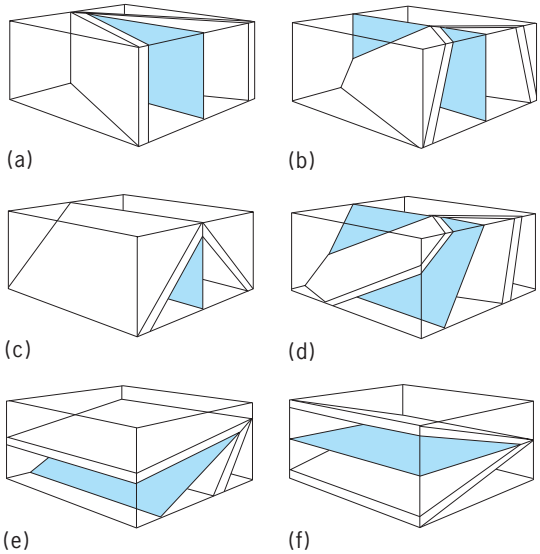

Fig. 2. Description of folds based on the attitudes of their axial surfaces and their plunges. (*a*) Vertical. (*b*) Upright plunging. (*c*) Upright horizontal. (*d*) Inclined plunging. (*e*) Inclined horizontal. (*f*) Recumbent. (*After M. J. Rickard, A classification diagram for fold orientations, Geol. Mag., 188:23,1971*)

Character of folds. In a section of folded rocks the layers possess different rheological properties; some have apparent stiffness (competency), whereas others behave less stiffly (incompetency). The most competent layers or group of layers control the folding, and the less competent units tend to conform to the fold-form of the most competent units. The least competent units take up the strain between the adjacent fold-controlling units.

The following nomenclature has been devised. The most competent units in the folded sequence are the dominant structural members. Those layers slightly less competent that conform to the fold-form of the dominant structural member are competent conforming members, and the least competent units are passive structural members.

The dominant structural members deform through much of their history of folding so that their thicknesses measured perpendicularly to their boundary surfaces are nearly constant at all points in the fold-form (**Fig. 3***a*). Such folds are said to be concentric. The passive structural members fold so that their thicknesses measured perpendicularly to their boundary surfaces vary from point-to-point in the fold-form, but thicknesses measured in the plane of their axial surfaces and perpendicularly to their axes are relatively constant (Fig. 3*b*). Such folds are called similar. The competent conforming members deform in a manner intermediate between that of concentric and that of similar folds. Under conditions of extreme strain all structural members may be flattened and take on the form of similar folds. Fold-forms when followed either upward or downward in a layered sequence tend to be disharmonic.

The wavelengths of concentric folds are related to the thickness of the dominant structural member. Studies of the relation between wavelength and thickness of the dominant structural member in folds of all sizes conclude $L \simeq 27T$, where L = wavelength and T = thickness of dominant structural member. In a layered sequence, first-cycle folds are controlled by the dominant structural member. Thin conforming competent members have, in addition to first-cycle fold-forms, second-cycle folds (minor folds) developed in the limbs of the first-cycle fold. The wavelengths of these minor folds are related to the thickness of the conforming competent member, that is, the thinner the member the shorter the wavelength.

Strain-stress relationships. Strain in rocks can be measured by comparing the shape of objects of known original form (fossils, pebbles, ooids, reduction spots) to their form after strain. Most studies of

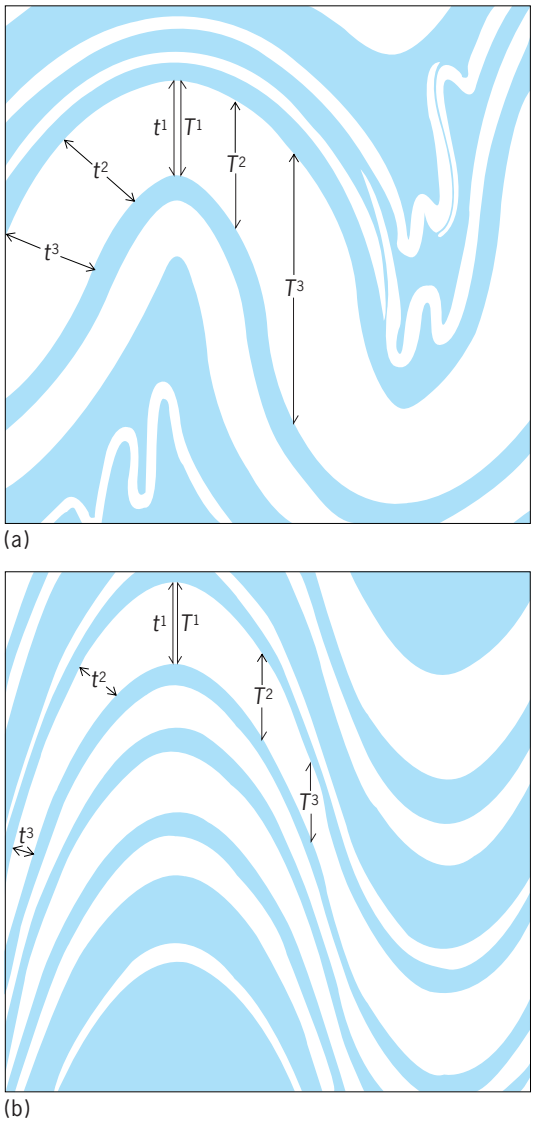

Fig. 3. Geometrical features of (*a*) concentric folds and (*b*) similar folds. Here t^1, t^2, and so on, are thicknesses measured perpendicular to boundary surface of folded layer. T^1, T^2, and so on, are thicknesses measured in axial surface and perpendicularly to fold axis. (*After J. G. Ramsay, The geometry and mechanics of "similar" type folds, J. Geol., 70:310, 1962*)

strain in folds have dealt with rocks that must be considered to be passive structural members in regions of intense folding. Where the folding has not been complicated by more than a single episode of folding, the orientations of the strain axes are as follows: the greatest direction of shortening is perpendicular to the axial surface; the greatest direction of extension lies in the axial surface and is subperpendicular to the axis; and the intermediate strain direction is perpendicular to the other two directions. Strains as great as 60 to 120% are recorded. Doubtless, detailed studies of strain in competent units would have orientations that differ with position within the unit, but such observations are generally beyond the resolution of field studies.

The measured strain orientations in folds of many regions are consistent with folds developed by axial compression of the layered sequence. Theoretical models have been derived on this assumption. Such models also assume the following: the structural members are ideally elastic or viscous; strains are infinitesimal; the thickness of the individual structural members is small relative to the thickness of the layered sequence; the axial load is constant throughout the length of the structural member; vertical normal stresses vary continuously; the fiber strains are linearly related to the distance from the center of the dominant structural member; and the layered sequence is confined by rigid boundaries. The wavelength of the multilayered sequence is given by Eq. (1), where L = wavelength; T = total thickness of multilayered sequence; t = thickness of dominant structural member and associated members; b = breadth of structural members; n = number of cycles of fold; B_h = horizontal modulus of elasticity; B_v = vertical modulus of elasticity; and I = moment of inertia. The critical load necessary for buckling with the wavelength given by Eq. (1) is expressed as Eq. (2), where P = critical axial load and G_a = modulus of shear.

$$L = 2\pi \left(\frac{T}{n\pi}\right)^{1/2} \left(\frac{B_b I}{btB_v}\right)^{1/4} \qquad (1)$$

$$P = 2\left(\frac{n\pi}{T}\right)(btB_vB_bI)^{1/2} + G_a tb \qquad (2)$$

Equations for viscous structural members are similar, but time becomes a significant parameter in them, and there is no critical axial load for buckling, although the amount of deflection is sensitive to the magnitude of the axial load. Equations based on both elastic and viscous structural members give wavelengths and numbers of dominant structural members per layered sequence in good agreement with observed characteristics of folded sequences.

Folding may be simulated with computers by programming suitable mathematical equations that relate the necessary boundary conditions, physical properties, and selected effects. These equations can be solved for any set of desired parameters. Two-dimensional maps of the distribution of stress and strain have been constructed for a single dominant structural member embedded in a viscous medium during folding (**Fig. 4**). The strain directions in simulated folds with finite amplitudes and steep limbs have a remarkable similarity to the distribution of strains measured in folded strata. *See* ROCK MECHANICS.

Minor folds and cleavage. Many first-cycle folds are accompanied by higher-cycle folds, called minor folds, and by cleavage. Empirical data indicate that, where these features formed in the same episode of deformation, the axes of minor folds are oriented approximately parallel to the axes of the first-cycle folds. Minor folds also have symmetries that vary with position in the first-cycle folds; they are asymmetrical in the limbs of the larger folds and symmetrical in their hinges.

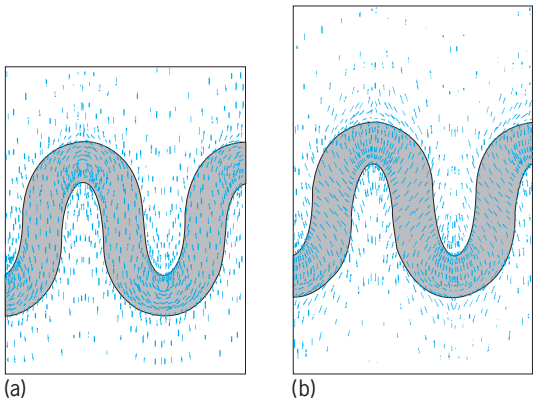

Fig. 4. Computed two-dimensional fold model after 100% shortening. (*a*) Stress relations; short lines are perpendicular to the maximum normal stress. (*b*) Strain relations; short lines are perpendicular to direction of maximum shortening.

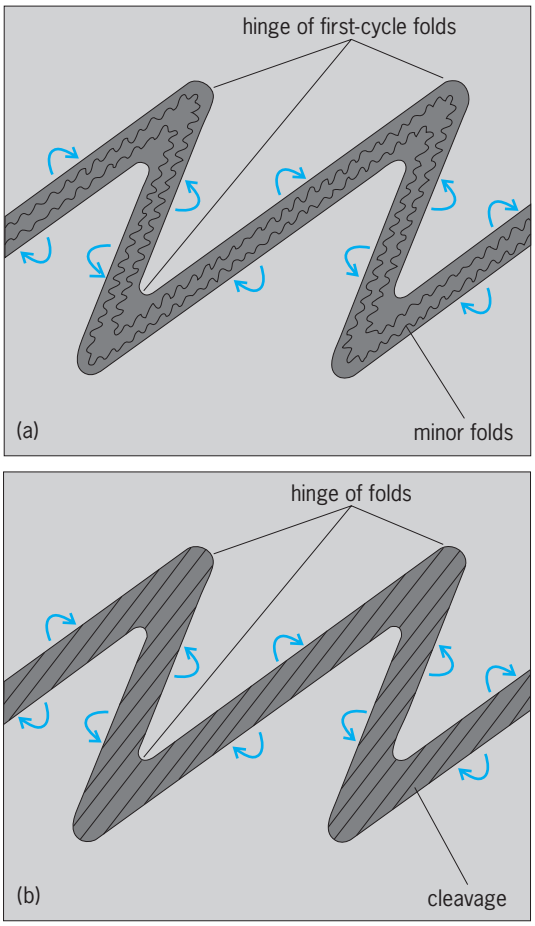

Fig. 5. Relationship between large folds and (*a*) minor folds and (*b*) cleavage; arrows indicate direction of deflection of adjacent hinge relative to limb observed.

The direction of deflection in the hinges of adjacent lower-cycle folds can be inferred from the minor folds (**Fig. 5***a*). The short limbs of the minor folds face the hinge of the larger fold to be considered, and the direction of deflection in this hinge relative to the limb observed is that given by following the long limb into the short limb of the minor folds.

Investigations also show that cleavage, which is manifest by closely spaced surfaces of rupture, approximately parallels the axial surfaces of folds, if these features were formed in the same episode of deformation. The geometric relation between cleavage and bedding is such that their intersection is approximately parallel to the axis of the fold. The direction of deflection in the adjacent hinges of folds can be inferred from the cleavage-bedding relationship (Fig. 5*b*). The acute angle between cleavage and bedding point to the hinge to be considered, and the direction of deflection in this hinge relative to the limb observed is that given by following the trace of bedding into the trace of cleavage around the acute angle between them. Generalizations of the relationships between minor folds and cleavage to larger folds allows the reconstruction of large folds from scattered observations.

Multiple folding. In many regions, layered rocks have had a long history of deformation including several episodes of folding. Because younger deformations do not completely obliterate earlier deformations, a complex array of folds develops. The relative timing of the various folds can be ascertained by their mutual relationships; the axial surfaces of older folds are folded by younger folds, and the axial surfaces of the youngest folds are not deflected (**Fig. 6**). Locally, folds of different ages may be directly superimposed, and it is at these localities that the relative ages are most readily determined. Once the relative timing of folding has been established at one or more localities, the relative ages can be applied to other localities on the basis of fold style. The orientations of superposed folds depend on the orientation of the forces producing them and on the attitude of the surfaces that are folded.

Fold systems. Folds generally do not occur singly but are arranged in festoons in mobile belts with lengths of thousands of miles and widths of hundreds of miles. The folds of these belts, or fold systems, commonly consist of great complex structures composed of many smaller folds. In this regard, a large

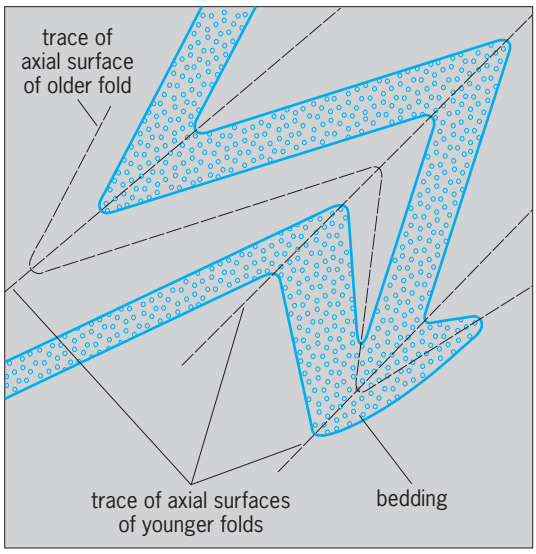

Fig. 6. Superimposed folds showing their relative ages.

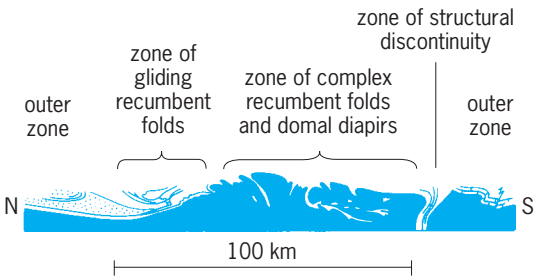

Fig. 7. Section through the Alpine Fold System. 1 km = 0.6 mi. (*After J. I. Dewey and J. M. Bird, Mountain belts and the new global tectonics, J. Geophys. Res., 75:2639, 1970*)

anticlinal structure consisting of numerous smaller folds is called an anticlinorium, and a large synclinal structure consisting of many smaller folds is termed a synclinorium. The trends of these smaller folds are more or less parallel, although the relationships may be complex. In some regions, folds belonging to several generations may be present, and locally these may have divergent trends.

Although comparisons of sections across fold systems are difficult to make because no two fold systems are identical and many are incompletely exposed for study, certain generalities emerge. Most fold systems are asymmetrical with respect to a zone of structural discontinuity (**Fig. 7**). The fold system is more extensively developed on one side of this zone than on the other, and the direction of structural transport on either side is directed toward the margins of the fold system. On the less-developed side of the zone of structural discontinuity, the folds are recumbent or asymmetrical and are accompanied by steeply dipping thrust faults. On the more extensively developed side, and adjacent to the zone of structural discontinuity, is a region of complex recumbent folds and domal massifs. On the outer side of these complex structures, recumbent folds that have moved toward the margin under the influence of gravity are extensively developed, and in the outer margin the folds are asymmetric and inclined and are associated with thrust faults having gentle dips.

The development of fold systems is closely tied to concepts of global tectonics. The favored hypothesis is that of plate tectonics, in which fold systems are formed at converging margins where continents are underriden by oceanic crust, or where collision occurs between a continent and an island arc or between two continents. Alternatively, the hypothesis of an expanding Earth, though less favored, is consistent with much of the geologic and geophysical data. According to this hypothesis, major zones of extension are the loci of gravity-driven diapirs the upper parts of which spread laterally under their own weight to give rise to fold systems. Philip H. Osberg

Bibliography. R. D. Hatcher, *Structural Geology: Principles, Concepts, and Problems*, 2d ed., 1994; D. D. Pollard and R. C. Fletcher, *Fundamentals of Structural Geology*, 2005; J. G. Ramsay and M. I. Huber, *Techniques of Modern Structural Geology*, vol. 2: *Folds and Fractures*, 1987; M. Shankar and G. W. Fisher (eds.), *Structural Geology of Fold and Thrust Belts*, 1992; R. J. Twiss and E. M. Moores, *Structural Geology*, 1992.

Folic acid

A yellow vitamin, slightly soluble in water, which is usually found in conjugates containing varying numbers of glutamic acid residues. It is also known as pteroylglutamic acid (PGA), and has the structural formula shown below.

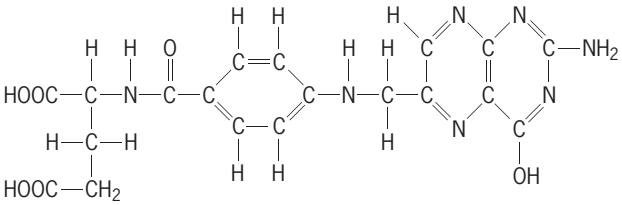

It is known that vitamins M and B_c, factors R, S, and U, the *Lactobacillus casei* factor, and a number of others are pteroylglutamates. Folic acid is usually assayed microbiologically by the use of *L. casei* or *Streptococcus faecalis*. Because of the various forms in which the vitamin appears in nature, some having different activity for the two organisms, and because of the difficulty in hydrolyzing the PGA conjugates, it is extremely difficult to determine absolute amounts of the vitamin. *See* BIOASSAY.

Biochemistry. Biologically active forms of folic acid occur as tetrahydrofolic acids. Four of these, N^5-formyltetrahydrofolic acid (the citrovorum factor, folinic acid, or leucovorum), N^{10}-formyltetrahydrofolic acid, N^5,N^{10}-methenyltetrahydrofolic acid, and formiminotetrahydrofolic acid have been found in natural products. Liver and kidney enzymes which catalyze the conversion of folic to tetrahydrofolic acids have been obtained. Ascorbic acid appears to be involved in some of these conversions. Folic acid enzymes are biological carriers of 1-carbon fragments as follows:

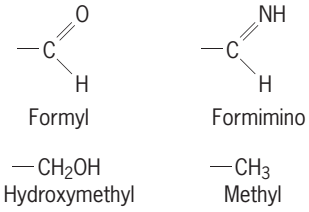

These groups are important in the synthesis of purines and pyrimidines, which are components of nucleic acids; in the synthesis of the methyl group of methionine; and in the metabolism of some amino acids, including glycine, serine, histidine, glutamic acid, and phenylalanine.

Nutritional requirements. Folic acid is required by a large number of animals and microorganisms. Many animals obtain enough folic acid from intestinal synthesis so that a dietary source is not necessary. In these species, the feeding of antibiotics or sulfa drugs can precipitate deficiency signs. Folic acid

deficiency is usually accompanied by poor growth, anemia and other blood dyscrasias, and gastrointestinal disturbances.

Folic acid is so widespread in nature and intestinal synthesis is so great that a folic acid deficiency in humans because of low dietary intake is probably not very common. With improved diagnostic methods an increasing number of folic acid deficient patients have been discovered. Deficiencies of other nutrients (particularly iron, ascorbic acid, or vitamin B_{12}) leading to impaired folic acid metabolism, difficulties in hydrolysis of folic acid conjugates by intestinal enzymes, or poor absorption may lead to a number of clinical conditions in which folic acid deficiency is involved. These include various nutritional macrocytic anemias, sprue, idiopathic steatorrhea, and pernicious anemia. The daily requirement of humans for folic acid is unknown, but on the basis of food analysis and therapeutic experience, it appears to be about 0.1 mg for the normal adult.

The anemia of pernicious anemia may be prevented by giving a person more than 0.1 mg of folic acid per day, although the neurological manifestations of this disease may progress. Sale without prescription of vitamin preparations recommending more than 0.1 mg folic acid per day has been prohibited by the Food and Drug Administration. *See* VITAMIN.

Folic acid antagonists such as aminopterin, which inhibits the conversion of folic to folinic acid, have been used in the treatment of leukemia and have in many cases brought about temporary remission of the disease. Stanley N. Gershoff

Industrial production. Folic acid is produced commercially by the simultaneous reaction of 1,1,3-tribromacetone with 2,5,6-triamino-4-hydroxypyrimidine and *p*-aminobenzoylglutamic acid. Dibrompropionaldehyde, with the assistance of an oxidizing agent, is sometimes used as a substitute for tribromacetone. Although various by-products are produced in such a reaction, selected conditions of pH and temperature give a reasonably good yield of crude folic acid. Purification of the crude folic acid through the zinc and magnesium salts gives a golden product essentially free of impurities.

Various processes can be used to make folic acid, but this one has been given because of its relative simplicity. The intermediates can be made readily by standard and familiar techniques. Most of the other processes are based on a stepwise conversion of the pyrimidine to the pteridine ring, along with additions to the benzoylglutamic molecule. Folic acid is then produced in a two-component condensation, rather than a three-component condensation. Richard J. Turner

Food

Any processed, partially processed, or unprocessed substance taken into the body primarily for providing nourishment. In addition, factors such as satisfying social needs, achieving psychological ends, and satisfying hunger govern the selection and consumption of foods. Chosen carefully, foods can provide all of the essential nutrients required for the normal functioning of the human body. In this context, food is necessary to provide energy, to provide structural components for building and repairing body tissues, and to regulate body processes. *See* METABOLISM; NUTRITION.

TABLE 1. Functions of essential nutrients

Function	Nutrients involved
Energy	Carbohydrates, fats/lipids, proteins
Structure	Proteins, lipids, minerals
Regulation of body processes	Proteins, minerals, vitamins, water

Nutrient requirements. Essential nutrients, such as carbohydrates, fats, proteins, minerals, vitamins, and water, have specific functions in the human body (**Table 1**). When energy-yielding organic compounds, including carbohydrates, fats/lipids, and proteins, are oxidized in the body, energy is captured in a chemical compound known as adenosine triphosphate (ATP), and then released slowly so that it can be used for respiration, physical activity (work), heat production, and metabolic processes. Enzymes, vitamins, minerals, and water are needed in order for these oxidation reactions to take place. Energy requirements are expressed in terms of kilocalories (kcal), calories (cal), or kilojoules (kJ). Fats and carbohydrates contain only the elements carbon, oxygen, and hydrogen. Since fats contain less oxygen than carbohydrates, they have greater potential for oxidation and thus provide more energy per gram than carbohydrates. Proteins also contain nitrogen, but it does not contribute substantially to the energy value. The physiological fuel value, or the amount of energy generated in the body, of 1 g of protein is 4 kcal (17 kJ); carbohydrate, 4 kcal (17 kJ); fat, 9 kcal (38 kJ). *See* ADENOSINE TRIPHOSPHATE (ATP); ENZYME; VITAMIN.

TABLE 2. Components in each nutrient group for which Dietary Reference Intakes (DRIs) exist

Nutrient group	Essential compounds
Fat	Linoleic acid, α-linolenic acid
Protein	Indispensable amino acids: histidine, leucine, isoleucine, lysine, methionine, phenylalanine, threonine, tryptophan, valine Conditionally indispensable amino acids: arginine, cysteine, glutamine, glycine, proline, tyrosine
Minerals	Calcium, chromium, copper, fluoride, iodine, iron, magnesium, manganese, molybdenum, nickel, phosphorus, potassium, selenium, silicon, sulfate, vanadium, zinc, sodium, chloride
Vitamins	Fat-soluble: A, D, E, K Water-soluble: C (ascorbic acid), thiamin, riboflavin, niacin, B-6, folate, B-12, pantothenic acid, biotin, choline

TABLE 3. Some functions of essential vitamins

Vitamins	Functions
Fat-soluble	
Vitamin A (retinoids, carotenes)	Night vision (preformed vitamin A); antioxidants (carotenoids); growth, reproduction, immune system integrity
Vitamin D (calciferol)	Promotes absorption and use of calcium and phosphorus; rickets prevention
Vitamin E (tocopherols)	Antioxidant, membrane integrity, heme synthesis
Vitamin K	Formation of proteins, particularly for blood clotting
Water-soluble	
Thiamin (B-1)	Coenzyme in energy metabolism, particularly from carbohydrates; antiberiberi factor
Riboflavin (B-2)	Coenzyme in energy metabolism, protein metabolism, redox reactions
Niacin	Coenzyme in energy metabolism, redox reactions; antipellagra factor
Vitamin B-6 (pyridoxine, pyridoxal, pyridoxamine)	Coenzyme for more than 100 enzymes in amino acid and protein metabolism
Folate	Coenzyme in DNA and RNA metabolism; single-carbon transfers, red blood cell formation; reduces risk of neural tube defects in infants.
Pantothenic acid	Coenzyme in fatty acid and energy metabolism
Vitamin B-12 (cobalamin)	Cofactor for methyl (carbon group) transfers; normalred blood cell formation and neurological function
Biotin	Coenzyme in fatty acid and carbohydrate metabolism
Vitamin C (ascorbic acid)	Cofactor for collagen formation; antioxidant; detoxification; scurvy-preventive vitamin
Choline	Needed for structural integrity of cell membranes, lipid and cholesterol transport

Many fats, carbohydrates (glucose and glycogen particularly), and proteins are synthesized in the body. Some of these compounds, as well as minerals and vitamins that are considered dietary essentials, or indispensable nutrients, must be provided by the food supply (**Table 2**). There are also vitamins and minerals that are essential to bodily functions (**Tables 3** and **4**). Dietary Reference Intakes (DRIs), developed by the Institute of Medicine based on scientific data, are reference values for macro- and micronutrients which can be used in dietary planning and assessment. They include four categories: Estimated Average Requirements (EARs), Recommended Dietary Allowances (RDAs), Adequate Intakes (AIs), and Tolerable Upper Intake Levels (Uls). *See* GLUCOSE; GLYCOGEN.

Water is essential in the human diet. Humans can survive for weeks without food, but only a few days without adequate water. Between 50 and 80% of the total body weight of an adult is water. Many foods supply water, but additional fluids must be ingested to meet body needs.

Nutritional guidelines. The Food and Nutrition Board of the Institute of Medicine of the National Academy of Sciences establishes RDAs (Recommended Dietary Allowances) for nutrients; that is, the quantities of certain nutrients that are thought to be adequate to meet the known nutritional needs of practically all healthy persons. To help individuals select adequate diets, food plans have been developed by U.S. Department of Agriculture (USDA) and the Department of Health and Human Services, based on the DRIs for each nutrient, including the RDAs, EARs, and Uls.

TABLE 4. Some of the functions of essential minerals

General function	Elements involved
Acid-base balance	Chloride, phosphorus, sulfate, potassium, sodium, calcium, magnesium
Water balance	Sodium, potassium, chloride
Bone and tooth formation	Calcium, phosphorus, manganese, fluoride
Catalyst for biological reaction and component of essential body compounds, such as hormones	Calcium, magnesium, zinc, copper, iron, molybdenum, cobalt, selenium, iodine, manganes, phosphorus, sulfate
Nerve function	Calcium, sodium, potassium

In 2005, new dietary guidelines, known as MyPyramid, were introduced to help individual consumers make food choices to maintain good health. The guidelines suggest eating a substantial number of servings of cereal-based foods and fruits and vegetables, and smaller amounts of animal products (meats, eggs, dairy) are recommended, compared to earlier published guidelines. The guidelines state that a healthy diet is one that emphasizes fruits, vegetables, whole grains, and fat-free or low-fat milk and milk products; includes lean meats, poultry, fish, beans, eggs, and nuts; and is low in saturated fats, trans fats, cholesterol, salt (sodium), and added sugars. Amounts of food in each group are suggested, together with descriptive information about serving sizes. For example, 2.5 cups of vegetables per day are recommended, particularly the dark-green and orange varieties, such as broccoli and carrots. In addition, balancing physical activity and food intake is emphasized to lower the risk of chronic diseases, such as cancer, diabetes, and cardiovascular disease, which are related to obesity. Reduced fat consumption, particularly saturated and trans fats (see Lipids section below), is also recommended. "Nutrition Facts" labels, on which requirements for terms used on the label to describe nutrient content are specified, aid in the selection of appropriate foods, as required by the Nutrition Labeling and Education Act.

Carbohydrate-rich foods. Plants synthesize carbohydrates by the process of photosynthesis and store them as starches (complex carbohydrates) or simpler sugars (glucose, fructose, or sucrose). The structural components of plants that are

nondigestible are called dietary fiber and include cellulose (a glucose polymer that the human body cannot digest), hemicellulose, and pectin. These, together with other complex carbohydrates (such as gums) produced by plants and bacteria, called functional fiber, are classified as total fiber. Functional fiber is beneficial physiologically, as it reduces cholesterol levels. Desirable total fiber intake can be achieved most satisfactorily by consuming fruits, vegetables, legumes, and whole-grain cereals. *See* CELLULOSE; FRUCTOSE; HEMICELLULOSE; SUCROSE.

In most of the world, about 80% of the calories consumed are carbohydrates, primarily from cereal-based foods. (Carbohydrates supply only about half of the calories in the United States.) In addition to being good sources of energy and fiber, whole grains supply B vitamins and minerals and contain small amounts of protein. The most common cereal grains consumed on a worldwide basis are wheat, rice, and corn (maize).

Some grains, such as rice, oats, barley, and buckwheat, are consumed as whole, cracked, or polished grains. Others, such as corn and wheat, undergo a series of mechanical disintegrations that reduce the grain to grits or flour. Whole or cracked grains can be consumed directly as cereals, while grits and flours are baked into bread products, made into pastas, or processed into ready-to-eat cereals. Each processing step adds to the convenience and cost of the final product. Current dietary guidelines emphasize the use of whole grains in breads and cereals. Enrichment of cereal-based foods restores some of the nutrients lost during milling. Thiamin, niacin, riboflavin, folate, and iron are required for a product to be labeled as enriched, while calcium and vitamin D are optional. It is estimated that more than 90% of the flour sold in the United States is enriched. *See* BARLEY; BUCKWHEAT; CARBOHYDRATE; CEREAL; CORN; NIACIN; OATS; RICE; STARCH; THIAMIN; VITAMIN D; WHEAT.

Protein-rich foods. Both animal and vegetable proteins provide the necessary amino acids for the synthesis of proteins in muscle, brain, blood, and other tissues in the human body, as well as hormones and enzymes. Proteins also provide about 15% of the energy in adult diets. Proteins are polymers composed of amino acids. Approximately 20 amino acids commonly are found in food proteins (Table 2), 9 of which are considered indispensable or essential in the diet because they cannot be made in the body. Six are conditionally indispensable; that is, they may be needed in the diet because they cannot be synthesized in sufficient amounts to meet a specific need. The quality of a protein is determined by how well it compares to a protein quality-scoring pattern based on the relative ratio of indispensable amino acids in the food protein and its digestibility, and to a reference protein, such as egg or milk. Legumes and nuts closely resemble meats in their amino-acid balance, but most other plant proteins are limited in at least one of the essential amino acids. The use of complementary plant proteins, such as beans and rice, or supplementation of vegetable proteins with small amounts of high-quality animal or vegetable protein provides adequate amounts of amino acids for human needs. *See* AMINO ACIDS; CHEESE; HORMONE; LEGUME; MILK; PROTEIN.

Lipids. This is the general name for the class of compounds to which fats and oils belong. Fats, chemically known as triglycerides, are composed of fatty acids attached to glycerol. A fatty acid is composed of an even-numbered carbon chain (4–22 carbons), which can be completely bonded with hydrogen (saturated) or can contain some carbon-carbon double bounds (unsaturated; that is, the adjacent carbons each lack one hydrogen). The position of the hydrogen atoms surrounding the double bonds is either *cis* (hydrogen atoms on one side of the bond) or *trans* (hydrogen atoms on opposite sides). The fatty acids of triglycerides in fats or oils are characteristic of the source of the material. All foods contain some lipids, even fruits and vegetables. In fats from animal sources, the majority of the fatty acids are saturated (for example, stearic and palmitic acid), but in vegetable fats, the fatty acids may be mono-, di-, or polyunsaturated. Cholesterol, an important component of cell membranes and precursor for steroid hormones and bile acids, is present only in foods from animal sources. *See* CHOLESTEROL; GLYCEROL; LIPID; TRIGLYCERIDE.

Saturated fats, such as butter, lard, and tallow, are solids at room temperature; polyunsaturated fats (oils) are liquids at room temperature. Saturated fats generally have low smoke points (the temperature at which the fat begins to break down) and are not as useful for frying; oils have higher smoke points. The most common vegetable oils in the United States are extracted from soybeans, cottonseed, corn, sunflower, canola, and safflower, which all have a high proportion of polyunsaturated fatty acids. Hydrogenation of oils for margarine and shortening results in formation of trans fats, which resemble saturated fats in their structure and physical properties. Labeling of trans fat content of foods was required as of January 2006, because of its connection with cardiovascular diseases. Dairy and meats naturally contain some trans fats. More than half of the dietary fat in the United States comes from animal sources; the remainder is from vegetable oils. Slightly less than 40% of the calories in the American diet are from fats and oils. *See* BUTTER; COTTON; RAPE; SAFFLOWER; SOYBEAN; SUNFLOWER.

Vegetable oils can be partially or fully hydrogenated to make solid or plastic fats. In making margarine, some liquid oil is added to hydrogenated fat, as well as color, flavoring, and liquid (usually milk) to emulate the consistency, flavor, and cooking characteristics of butter. Spreads containing lower amounts of fat than butter, or solid margarines, can be made by blending saturated (palm), polyunsaturated (soybean, canola, cottonseed, sunflower, or safflower) or monounsaturated (olive) fats and oils. By using these alternatives, cooking properties may be changed. Properties of hydrogenated shortenings are altered by fat substitution and processes that minimize trans fats in the finished products, so food companies are developing patented technologies to achieve the desired functional properties. Reduced-calorie spreads

can also be made by diluting the fat with water or incorporating air to make a whipped product. For baking and frying, shortening can be made from hydrogenated vegetable oils alone or can be blended with animal fat. Superglycerinated shortenings have mono- and diglycerides added to improve baking and emulsifying properties. *See* FAT AND OIL; MARGARINE.

Vegetables and fruits. Recent nutritional guidelines emphasize the consumption of a variety of fruits and vegetables, particularly dark-green and deep-yellow ones. These foods constitute major sources of the carotenoids (such as vitamin-A precursors), ascorbic acid, and many minerals, as well as digestible carbohydrates and dietary fiber. Fruits and vegetables also contain phytochemicals or phytonutrients, which provide a health benefit beyond those of essential nutrients. Vegetables include a variety of edible parts of plants (leaves, stems, flowers, roots, tubers, pods, and seeds) and are usually eaten cooked or raw, as a part of a meal. Fruits of certain plants (tomatoes, eggplant, squash, and peppers) are grouped with the vegetables. Fruits generally contain more sugar than vegetables. Both fruits and vegetables are very good sources of dietary fiber.

Supplies of fresh fruits and vegetables once were limited by their seasonal nature, but the development and refinement of transportation, packaging, and canning and freezing technologies have extended their availability. Freezing or canning of fruits and vegetables is done as soon as possible after harvesting, usually within hours to help preserve the optimum quality and nutritional value. The nutritional value of canned and frozen fruits can be similar to that of "retail market" fresh fruits and vegetables. Modified and controlled-atmosphere storage extends the shelf life of fresh produce. *See* FRUIT.

Foods for health. Although it is unlikely that any individual food or combination of foods can provide complete disease protection, the risk of chronic disease can be reduced by increased consumption of plant-based foods and decreased consumption of fats. Foods that provide essential nutrients, beyond quantities needed for normal growth, maintenance, and development, or contain other biologically active components are called functional foods. These include fruits, vegetables, and whole grains. Because obesity in the United States has increased significantly over the last 30 years, health, nutrition, and food-science professionals are emphasizing healthy eating and exercise to decrease the risk of obesity-related diseases, such as type-2 diabetes and cardiovascular disease.

The food industry is focusing its product development efforts to meet recommendations for weight loss and health maintenance. Formulated foods that are described as fat-free, low in fat, or reduced-calorie depend on the use of fat substitutes and mimetics, as well as high-intensity sweeteners. High-intensity artificial sweeteners, such as acesulfame-K, aspartame, saccharin, and sucralose are used in tabletop sweeteners, beverages, candies, dessert mixes, and dairy products. When these sweeteners replace sugar in food products, bulking agents (such as sorbitol, polydextrose, or cellulose) must be added. Fat substitutes and mimetics contribute fewer calories than fat because they are partially absorbed or nonabsorbed in the intestinal tract. They provide sensory characteristics in foods similar to those imparted by fats. Carbohydrate-based fat substitutes, such as cellulose, gums, dextrins, modified starches, and polydextrose, can be found in a variety of foods, including dairy products, frozen desserts, and salad dressings. Microparticulated protein is made from egg white or milk proteins. The spherical particles produced by the processes are so small that the tongue perceives them as a fluid, with a smooth, creamy texture. Microparticulated protein cannot be used in cooked foods because it forms a gel when heated. Sucrose polyesters are composed of six to eight fatty acids bound to a sucrose molecule. These fat substitutes are heat stable and have the potential for use in baked and fried foods. According to the American Dietetic Association, fat-modified foods can be used as part of a well-balanced, nutritious diet to reduce fat intake. *See* FOOD ENGINEERING; FOOD MANUFACTURING; FOOD MICROBIOLOGY; FOOD SCIENCE.

Barbara P. Klein

Bibliography. R. L. Duyff, *American Dietetic Association Complete Food and Nutrition Guide*, 2d ed., Wiley, New York, 2002; Food and Nutrition Board, Institute of Medicine, National Academy of Sciences, *Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids* (*Macronutrients*), National Academies Press, 2005; L. K. Mahan and S. Escott-Stump (eds.), *Krause's Food, Nutrition & Diet Therapy*, 10th ed., Saunders, Philadelphia, 2004; H. McGee, *On Food and Cooking: The Science and Lore of the Kitchen*, Scribner, New York, 2004; N. N. Potter and J. Hotchkiss, *Food Science*, 5th ed., Springer, 1999; G. M. Wardlaw, *Contemporary Nutrition: Issues & Insights*, 5th ed., McGraw-Hill, Boston, 2003.

Food allergy

An "altered reaction" to a food substance to which people typically do not react. Although the term "allergy" is sometimes used loosely for many different kinds of response, strictly speaking it refers to an immunological response mediated by the IgE class of immunoglobulins, or antibodies.

As many as 8% of children and 2% of adults in the United State have had food allergy symptoms. Each year about 30,000 people visit emergency departments, and an estimated 150 die from allergic reactions. Allergy to food is similar to that to other environmental allergens such as pollens, but there are additional specific risks, symptoms, and immunological mechanisms.

IgE antibodies and allergic mechanisms. IgE immunity plays an important role in protection against intestinal pathogens, primarily worms. Although everyone produces some IgE antibody to most environmental antigens, especially those that enter the body through the mucous membranes of the

respiratory and gastrointestinal tracts, people with allergies make more. This process is thought to be driven by excessive activity of helper T cells (Th) of the type 2 class (Th2) compared with Th1 cells for reasons that are both genetic and environmental. Although about 15% of the general population has some form of allergy, if both parents of an individual are allergic the individual's risk of having allergies increases to 50%. Allergies are more common in rich than poor countries and in urban than rural settings. This evidence supports the "hygiene hypothesis," the idea that exposure to environmental antigens is important in the first years of life to stimulate the Th cells and achieve an optimal balance between Th1 and Th2 cell activities. *See* ANTIBODY; CELLULAR IMMUNOLOGY.

There is very little IgE free in the blood, because as soon as it is made it binds to surface receptors on mast cells, connective tissues cells containing granules of biologically active chemicals (mediators). Thus in a person with allergies, many mast cells will be coated with IgE against one or a few antigens which, upon entering the body, will cross-link adjacent IgE molecules. This signals the mast cell to immediately release the contents of its granules, which include large amounts of histamine. Within a few hours, the mast cells also generate proinflammatory mediators, such as cytokines and leukotrienes. In response to histamine and other mediators, small blood vessels dilate and become somewhat leaky, causing local tissues to swell with fluid; bronchioles constrict; and gastrointestinal smooth muscle contracts. Histamine also causes itching by stimulating nerve endings. *See* ALLERGY; CYTOKINE; HISTAMINE; INFLAMMATION.

Triggers and symptoms. Only eight food types trigger 90% of allergies. Peanuts, tree nuts, fish, and shellfish are the most common allergens in adults. Children can be allergic to these foods, as well as to milk, eggs, soy, and wheat. For reasons that are not understood but may have to do with the maturation of Th2 cells, children frequently outgrow their allergies as they enter adolescence. Adults rarely do, though severity of symptoms may wax and wane.

The symptoms of food allergy are most common in the gastrointestinal tract but can also be seen on the skin, in the lungs, and systemically (throughout the body). They almost always occur soon (within seconds to 2 hours) after eating the trigger food. Gastrointestinal symptoms range from mild discomfort to cramping, nausea and vomiting, and diarrhea. Respiratory symptoms include coughing; wheezing; and a congested, itchy, or runny nose. The skin may display urticaria (or hives), localized itchy swellings surrounded by a flare of dilated capillaries. Occasionally, large hives may develop beneath the surface of the skin, a condition known as angioedema. The most feared symptom of allergy is anaphylaxis, a systemic allergic reaction that occurs when enough histamine is released from mast cells all over the body that generalized blood vessel dilation results in a drop in blood pressure. The patient may experience just some light-headedness or actually may go into anaphylactic shock, lose consciousness, and require emergency resuscitation. *See* ANAPHYLAXIS.

Oral allergy syndrome. A recently recognized special case of food allergy is oral allergy syndrome, an allergic reaction to certain proteins found in a variety of fruits, vegetables, and nuts that affects people with sensitivity to pollen. The reaction occurs almost immediately on putting the offending food in the mouth. The antigens seem to be able to pass through the mucous membranes of the mouth and gain rapid access to local mast cells, and eventually to mast cells at more distant sites. In most cases, these particular antigens are destroyed in the stomach; thus individuals experience only oral symptoms, including tingling of the lips and tongue, and itching and sometimes swelling of the lips. Though usually just a nuisance, oral allergy syndrome may lead to more severe and general symptoms and should not be ignored by either patients or their health care providers.

Studies have repeatedly shown that the foods that cause oral allergy syndrome contain proteins that are structurally similar to proteins in certain pollens. It is thought that the patient becomes sensitized (that is, makes IgE) in response to the pollen, which may or may not in itself trigger allergic symptoms. The patient's immune system has difficulty distinguishing between the pollen proteins and the food proteins, and triggers an allergic reaction in response to the food proteins. This is known as a cross-reaction, as may occur between inhaled birch pollen and ingested apples, nuts, and carrots; between grass pollen and green beans and peas; or between ragweed pollen and cantaloupe, honey dew melon, banana, or chamomile. The food reactions are often worse in pollen season.

Diagnosis. Allergists work to discover which food in a varied diet is triggering symptoms. A careful medical history of the patient is taken, and suspicious foods are temporarily eliminated. One by one they can be reintroduced and the patient observed. This is not done when the reaction has been severe; in these cases, skin tests may be performed (and only by specialists in the case of very severe symptoms). A skin test involves application of the antigen to the skin by injection or scratching. The development of a hive after about 15 minutes is considered a positive result. Although this does not prove the symptoms are from that antigen, it raises the index of suspicion. For most common food allergens, a completely safe blood test known as the radioallergosorbent test, or RAST, is available. The gold standard is the placebo-controlled double-blind test, in which food is placed in unmarked capsules and ingested while the patient is under careful observation in an experienced hospital or clinic setting.

Food intolerance. Food allergy is commonly misdiagnosed, often by the patient who has had some sort of unpleasant reaction to a food. Food intolerance is the generic term for such reactions, which are not mediated by IgE, and it is perhaps five times more common than true allergy. Certain food such as fish

in the tuna and mackerel family, if not properly refrigerated, can develop high enough levels of histamine to cause symptoms indistinguishable from food allergy. Some people are intolerant to lactose, a sugar found in milk, and experience bloating and diarrhea after eating dairy foods because they are deficient in the enzyme lactase, which is necessary for lactose digestion. In some people, monosodium glutamate, a flavor enhancer, causes flushing of the skin and headache; and sulfites, which are used as preservatives, can trigger asthma. Gluten intolerance is an immunological reaction to the gluten protein in wheat, though the reaction is not mediated by IgE. It is not uncommon to develop a "conditioned aversion" to a food that once made us ill, and some people experience symptoms if the food is even mentioned.

Treatment and prevention. Treatment depends on whether the problem is acute and life-threatening, or mild and chronic.

In emergency situations, in which an individual begins to experience shock and labored breathing, drugs, such as epinephrine, which constrict blood vessels and dilate bronchioles, are essential. Fluid replacement and the administration of antihistamines, oxygen, and glucocorticoids may all become necessary. Once the patient is stable, efforts must be made to identify the offending food because strict avoidance is the best preventive measure. Since people with severe allergies (such as to peanuts and shellfish) can have an anaphylactic reaction to a skin test with the trigger substances, expertise with these procedures is mandatory. Specialized centers can perform sensitivity tests in a controlled and safe manner. It is advisable for people with a history of anaphylactic reactions to wear medical alert bracelets or other identification and carry preloaded syringes containing epinephrine at all times.

For milder conditions, avoidance of the offending food is also the mainstay of treatment. This is preferable to continuous use of antihistamines, as is recommended for people with inhalant allergies. It is important for health care providers to make people with food allergy aware, without inducing panic, that mild reactions do rarely progress to severe or even anaphylactic intensity. For example, angioedema is an early warning sign of trouble, as the large hive-like swellings of the face can sometimes migrate to the throat, blocking the airway. If symptoms are progressing, it is vital to seek medical help at once.

Most children will outgrow their food allergies. Adults do not seem to do so, although severity of symptoms varies. For example, a person who can eat a few shrimp without trouble may develop diarrhea and wheezing after eating large quantities. Most people learn what works for them and can usually manage a reasonable diet without much extra vigilance. Those with severe allergy, for example to peanuts, must be very careful, read all labels, ask about ingredients, and nevertheless always be prepared for an allergic reaction.

Recently a new monoclonal antibody against IgE, called omalizumab, has been approved for treatment of severe asthma, and is being evaluated for use by peanut-sensitive patients. In early trials, patients were able to tolerate controlled doses of peanut that would otherwise have triggered anaphylaxis. *See* MONOCLONAL ANTIBODIES.

J. John Cohen

Bibliography. Articles about all aspects of food allergy, *Curr. Opin. Allergy Clin. Immunol.*, vol. 5(3), Lippincott Williams & Wilkins, June 2005; J. Brostoff and L. Gamlin, *Food Allergies and Food Intolerance: The Complete Guide to Their Identification and Treatment*, Healing Arts Press, 2000; D. Y. Leung et al., Effect of anti-IgE therapy in patients with peanut allergy, *N. Eng. J. Med.*, 348:986–993, 2003; U.S. Department of Health and Human Services, National Institute of Allergy and Infectious Diseases, *Food Allergy: An Overview*, NIH Publ. no. 04-5518, July 2004.

Food engineering

The application of engineering concepts and principles to the conversion of raw foods into safe consumer products of the highest possible quality. The entire spectrum of food engineering is associated with operation and maintenance of food processing plants as well as sophisticated research involving process design.

The applications of engineering in food handling, processing, packaging, and distribution can be described in terms of unit operations. There are many different unit operations associated with the conversion of raw food materials to consumer products. The movement of foods and other materials within the processing plant requires the use of unique equipment and processes. For example, special sanitary pumps are used to transport liquid foods, and the material-handling equipment for solid foods requires careful design for product-contact surfaces.

The importance of thermal treatments for food preservation requires that a broad range of heat-exchange equipment be used. Heat exchangers for liquids are unique in terms of sanitary design and cleanability of surfaces following thermal processing. A special component of thermal preservation is the design of thermal processes.

Several unit operations involve heat transfer in order to achieve the desired preservation even though storage stability is not the direct result of thermal treatment. An excellent example is the freezing process, where removal of thermal energy reduces product temperatures to levels where deterioration reactions are significantly inhibited. Concentration processes achieve a degree of preservation by reducing the availability of water for deterioration reactions, although the primary aim is reduction of liquid-product mass and volume. Although traditional concentration processes have used thermal energy to evaporate water, membranes of various types are now used to achieve the same results. The preservation of food products is achieved by reduction of the water content to low levels by means of dehydration processes which use thermal energy. These processes are applied to liquid foods and to products

that are naturally solid. *See* FOOD PRESERVATION.

Another series of unit operations is used to alter the product composition or structure in some manner. These include separation, mixing, and extrusion. Separation processes are designed to divide food products into two or more components. While a variety of physical or chemical properties of the product components are used in the various separation processes, two of the most important processes are filtration and extraction. Filtration, a physical process, has several applications in addition to its use for separating product components. Extraction is most often designed to remove a specific or unique product component for use in a separate operation or product formulation. After separation, the final product is obtained through the use of a mixing process which includes a variety of equipment types. Finally, the extrusion process involves the use of both thermal and flow properties to achieve product preservation as well as some specified set of structural and textural characteristics.

The importance of cleaning and sanitation must be emphasized due to direct relationships to final product quality. The required operations vary considerably depending on the type of product handled and the type of equipment used. The processes required to manage the wastes generated during food handling, processing, packaging, and distribution are all similar, and many of the waste-handling and treatment operations are the same as those used directly with the food products.

The final operation to which the product is subjected before distribution is packaging. The package barrier is important for maintaining food products at desirable quality levels. Food packaging involves the selection of equipment needed for placing the product in the package as well as the selection of packaging material needed to protect the product in an optimum manner.

An engineering input to food handling, processing, packaging, and distribution that is applied to almost all unit operations is process control. The use of instrumentation and associated electronic controls has a significant impact on the efficiency of all components of the food delivery system. *See* UNIT OPERATIONS.

Dennis R. Heldman

Pumps and Pumping

Pumps are one of the most common pieces of equipment in food processing facilities. Fluid foods with widely differing characteristics must be pumped between many different unit operations before a final product results. Some pumps are selected for a single application or duty, and others are used for various products and flow rates.

Sanitary pumps. Pumps used in handling food must be sanitary, that is, specifically designed to handle food or related biological materials, such as pharmaceutical products. Sanitary pumps are designed for easy and frequent cleaning. They are easily dismantled by removing wing nuts by hand or by using simple tools. In addition, the units are mounted to pipelines with special clamps and disconnects designed for rapid breakdown. Pumps may also be designed for cleaned-in-place systems, which reduce the need for frequent dismantling. These self-cleaning units are constructed with grooves which route sanitizers and create turbulence to help dislodge solids.

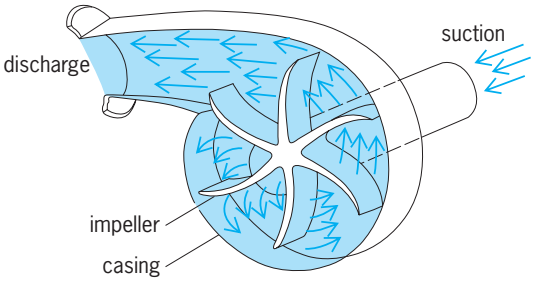

Fig. 1. Internal view of a centrifugal pump. (*After C. R. Westaway and A. W. Lommis, eds., Cameron Hydraulic Data, 16th ed., Ingersoll-Rand, 1984*)

Sanitary pumps must be manufactured from materials that are resistant to corrosion and the chemicals used for cleaning and sanitation. Pump rotors and casings are commonly available in stainless steel; rotors are also available in natural or synthetic rubber. Sanitary pumps should have no internal wear, and the lubrication systems for bearings and other moving parts should be completely sealed to prevent contact with the food being pumped. This is typically accomplished with single or double O-ring–type seals which may incorporate flushing systems for vacuum applications or when pumping hot fluids.

Centrifugal pumps. Pumps used to handle fluid foods will generally fall into one of the two classes: centrifugal or positive displacement. Centrifugal pumps differ widely in size and internal design; however, they all have the same basic principle of operation. The pump (**Fig. 1**) consists of an impeller rotating within a casing and attached to a shaft that is driven externally by an electric motor. Liquid enters the pump at the center of the rotating impeller and is accelerated to a high velocity by the movement of the impeller vanes. The resulting movement—from the center to the outer edge of the impeller—greatly increases the kinetic energy of the fluid. The fluid velocity is reduced when the material is discharged by the vanes to the casing, which causes some of the fluid kinetic energy to be converted to pressure. This increased pressure causes fluid movement from the discharge section of the pump.

Liquid is delivered from a centrifugal pump in a steady stream, that is, without pulsations. Wear is minimized because the pump has no internal rubbing parts and clearances are relatively large. Centrifugal pumps are not self-priming, and develop a limited pressure when operated at a constant speed.

Centrifugal pumps are very versatile and can handle all liquids except highly viscous ones. Large clearances make them useful in pumping liquids containing solids such as cherries, oysters, whole fish, beans, and olives. Overagitation of liquids may be a problem with centrifugal pumps if operated against high pressure. In addition, centrifugal pumps may incorporate

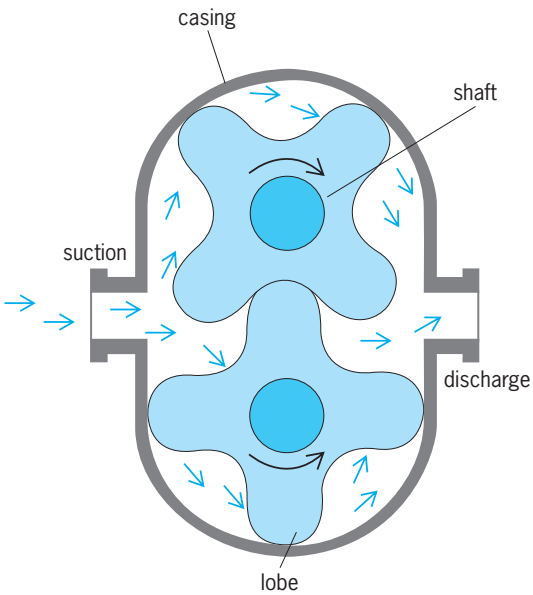

Fig. 2. Internal view of a four-lobe rotary pump.

rotor designs that provide mixing, emulsification, or dispersion as part of the pumping operation. *See* CENTRIFUGAL PUMP.

Positive-displacement pumps. Positive-displacement pumps operate with close internal clearances, and commonly use revolving gears, vanes, lobes, or screws in a fixed casing to bring about fluid movement (**Fig. 2**). In contrast to centrifugal pumps, fluid may be delivered in a pulsating flow. Lobes are attached to shafts that rotate in opposite directions. During rotation, fluid is trapped between the lobe and the casing at the suction end of the pump, then moved (or displaced) to the discharge end where volumetric displacement by the lobes causes the fluid to exit the system. A given quantity of fluid is delivered with each rotation of the shaft. Rotary pumps are self-priming and capable of producing very high pressures; however, excessive pressures can occur if the pumps are operated against a closed valve. These units are very useful in handling highly viscous fluids, and are also used for fluids with low viscosity. Due to close working clearances, rotary pumps are limited to fluids that do not contain large particles.

Rotary pumps, available for aseptic processing systems (that is, for sterile products placed in sterile containers), incorporate design features which eliminate the possibility of undesirable microorganisms entering the product through the pump. These units operate by having steam or a sterile solution circulate between double seals at the ports, cover, and shaft of the pump.

Positive-displacement pumps may combine self-priming features with the transfer capabilities of centrifugal pumps. This may be accomplished by using a flexible neoprene rotor with a pump casing having a noncircular cross section (**Fig. 3**). Rotary action of the impeller causes material to be picked up at the inlet port and transferred in discrete units to the outlet. When the rotor contacts the offset section of the casing, the rotor bends, causing the food to be squeezed from the pump. This type of design tends to minimize damage to suspended solids.

Positive-displacement pumps may be designed with a progressing cavity. One commercial design (**Fig. 4**) involves a screwlike rotor that turns within a double-thread helical stator to form a progressing cavity in the pump. Pumps based on this principle tend to be very effective in handling highly viscous, abrasive, or shear-sensitive fluids. A second design, the twin piston pump (**Fig. 5**), alternates pumping from one piston to the other. It has large, smooth passages that allow it to pump highly viscous liquids, fruit and vegetable products with large particulates, dough, and meat products. Moving sleeves and large ports minimize product damage. *See* DISPLACEMENT PUMP.

Pumping. Pumps are driven by electric motors and provide an energy input to the fluid-handling system. Power requirements for pumping are calculated by using the mechanical-energy-balance form of Bernoulli's equation. Energy inputs must increase if the potential, kinetic, or pressure energy of the fluid is increased. In addition, the pump must overcome energy losses due to friction caused by flow through straight pipe, valves, fittings, and processing equipment, such as filters and heat exchangers. Pumping fluid foods must often be given special consideration because many products exhibit non-newtonian behavior. Common problems include time dependency and the presence of a yield stress. These properties are taken into account when calculating pipeline energy losses due to friction and kinetic energy differences. In addition, many materials (such as an emulsion like mayonnaise) are sensitive to shear, which may limit engineering design options, including pump speed and pipe diameter. Product data such

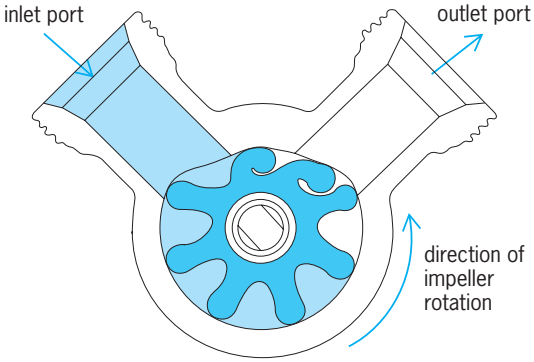

Fig. 3. Rotary pump with a flexible impeller and off-set casing.

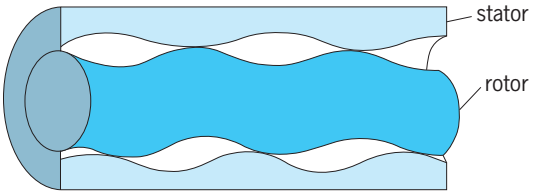

Fig. 4. Rotor and stator for pumps operating to form a progressing cavity.

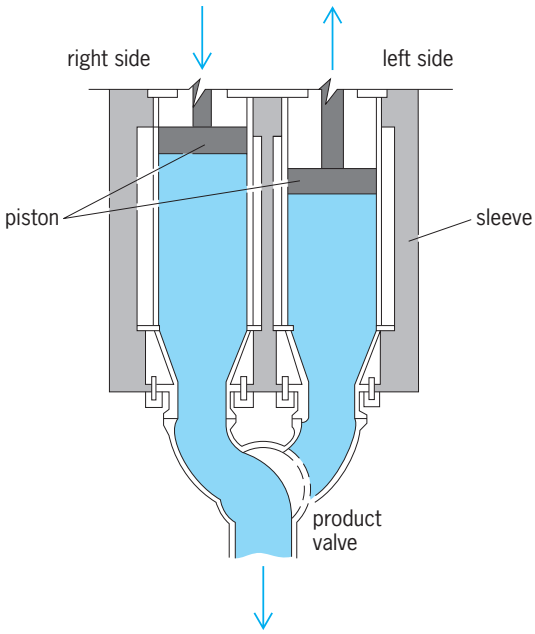

Fig. 5. Twin-piston pump. (*Marlen Research Corp.*)

as temperature, particle size, specific gravity, corrosiveness, and abrasiveness must also be considered in pump selection. *See* BERNOULLI'S THEOREM; FLUID-FLOW PRINCIPLES; FLUIDS; NON-NEWTONIAN FLUID; PUMP; PUMPING MACHINERY. James F. Steffe

Material Handling

Material handling is a term used to describe the movement of materials through various stages of processing. It includes movement of raw material from a supply location to a processing plant; movement of material through different stages of processing within a plant; and movement of finished product to a warehouse or to storage. In material handling it is imperative that the movement of a known quantity of material to a preselected location occurs at the desired time while maximizing the economy of space.

Equipment used for the mechanical handling of materials may be classified as conveyors and elevators, hoists and cranes, trucks, and pneumatic handling equipment.

Fig. 6. Belt conveyor. (*Aseeco Corp.*)

Conveyors and elevators. Conveyors and elevators carry material in a limited but continuous stream. Belt conveyors (**Fig. 6**) involve endless belts running about end pulleys and supported along the length by a series of securely mounted pulleys called idlers. Commonly the belts are made of rubber, neoprene, canvas, or stainless steel alloys. In situations where continuous cleaning is necessary, the return side of the belt may be scrubbed with roller brushes and washed with water sprays. A variety of modifications allows conveyors to be customized for special requirements. Belt conveyors with stainless steel cleats attached to rubber belts are used to move dry products. A slat-type belt may be useful to convey chunky products, such as cut-up chicken. Conveyors may also be mounted on scales for continuous weight measurements.

Screw conveyors use a helical screw that revolves and pushes the product through a trough or a box. Commonly, these conveyors consist of stainless steel ribbons wound helically around the central shaft or pipe. These conveyors may be mounted horizontally at an inclination or even vertically. They are useful in conveying dry and slightly moist products, for example, flour, cereals, and comminuted meat.

Bucket elevators consist of steel, malleable iron, or synthetic plastic buckets attached to endless chains or belts running about end sprockets or pulleys (**Fig. 7**). Usually the buckets are attached at equal spacings. Bucket conveyors raise material in a vertical plane and are used to transport dry products, such as coffee beans, sugar, and salt.

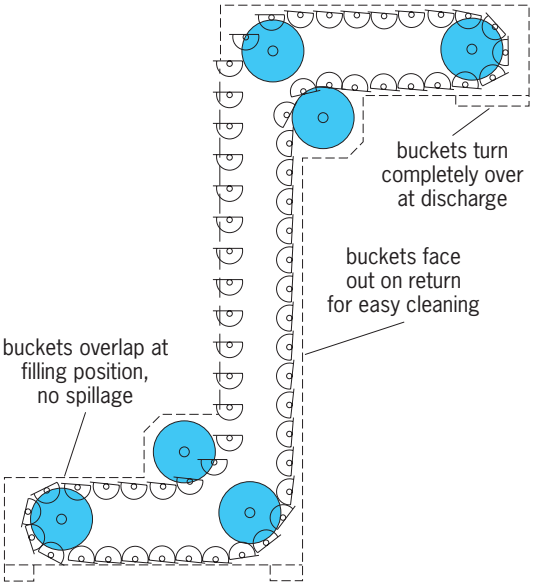

Fig. 7. Schematic of a bucket conveyor. (*Aseeco Corp.*)

To convey delicate products that must be handled gently to avoid breakage or product segregation, vibratory conveyors are used (**Fig. 8**). Vibratory conveyors employ the principle of natural frequency to move the product gently and eliminate any possibility of particle separation. The conveyor pans

are supported by springs, and the drive mechanism is tuned to resonance with the natural frequency of both springs and pans. The product moves in a "throw-and-catch" mode. During the upstroke of the pan, the product is lifted gently and pitched forward by the pan. In the following downstroke, the product is caught by the pan before it reaches the bottom of its stroke. This procedure helps in avoiding impact of the product with the pan and thus minimizes physical damage. In addition, the wear of the conveyor itself is small when it is used to convey abrasive materials. These conveyors tend to be self-cleaning and avoid buildup of material. In handling wet products, the self-cleaning feature is particularly attractive as it reduces shutdown time for cleaning.

The application of computers to control conveyor movement has become increasingly popular in plants that handle glass bottles at high speeds. For example, a programmed computer may be used to control the number of bottles on each conveyor section, and to control the speed of the conveyor belt and any associated processing machinery.

Fig. 8. Vibratory conveyor. (*Aseeco Corp.*)

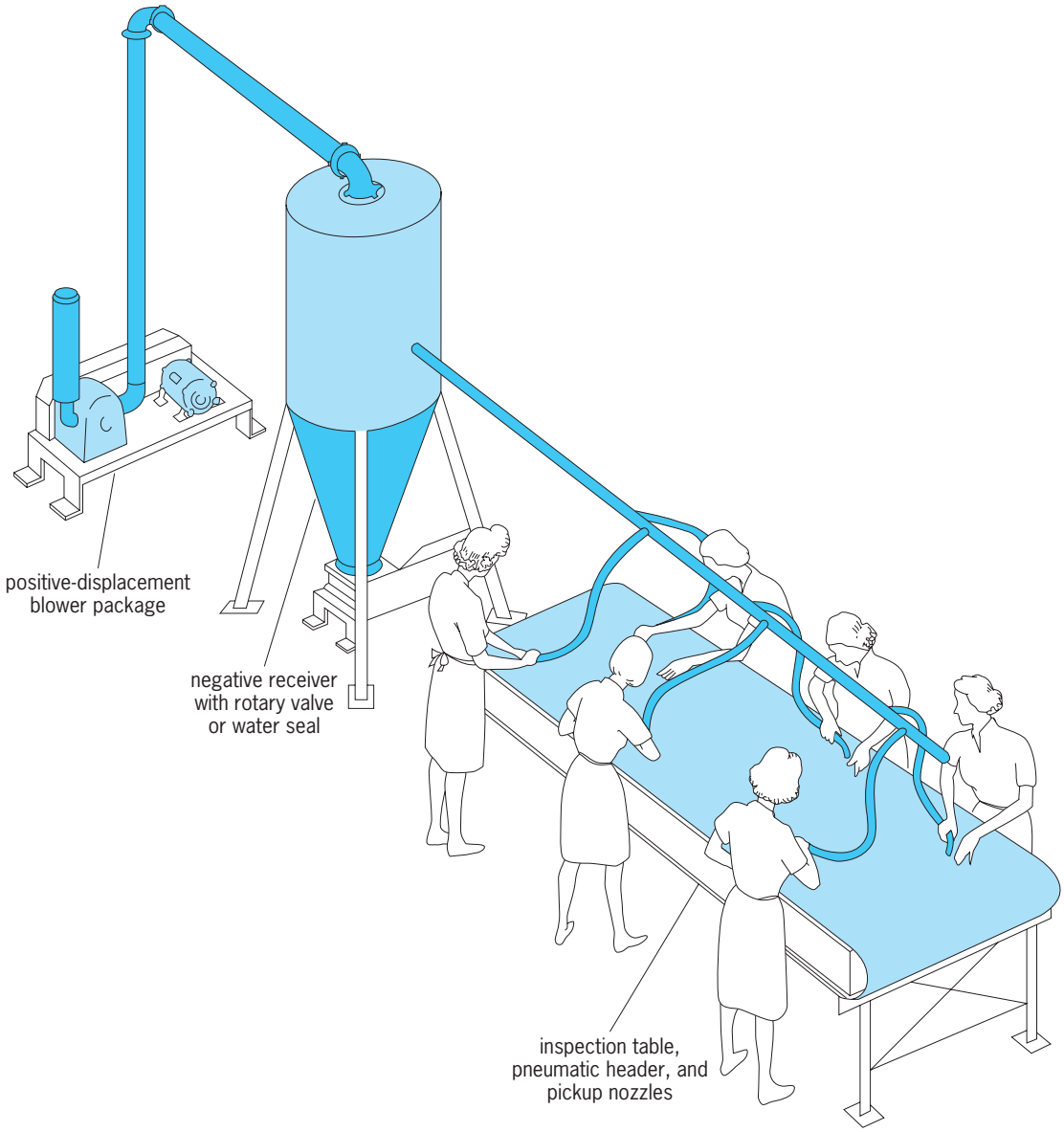

Fig. 9. Pneumatic conveying system. (*Temco, Inc.*)

Computerized control allows optimal line efficiency, lower noise level, reduced breakage by gentler handling, reduced power consumption, and reduced manual work for operators. *See* BULK-HANDLING MACHINES; CONVEYOR; ELEVATING MACHINES.

Hoisting. Hoisting towers are primarily used for unloading purposes; the hoisted material is transported according to whether the tower is of the stationary or traveling type. Simple mast rigs are used in loading and unloading loose material with buckets or scoops. Electrically operated overhead cranes are used to raise or lower heavy loads and distribute them with the trolley traveling on a bridge; the bridge itself moves on the supporting rails. *See* HOISTING MACHINES.

Industrial trucks. These are used in handling food at all stages of production. In the food industry, the term truck is used to include a large range of vehicles—from manually operated flat-pallet trucks, such as the ones used in supermarkets, to battery-powered equipment used in stacking loads of almost 3 tons (2.7 metric tons) to heights up to 30 ft (10 m) in warehouses. Most applications inside a food warehouse or processing facility preclude the use of an internal combustion engine as a power source, because it is mandatory to avoid product contamination from engine exhaust gases. Battery-powered equipment is used mostly indoors, while diesel-powered trucks are used outside. In comparison to a truck with an internal combustion engine, an electric truck is relatively slow and requires battery recharges between shifts. The trucks themselves must be protected from any hostile environment, such as contact with chemicals that may be used in the processing plant. For example, in a fish processing plant, brine can be hazardous to the truck body. Special construction features are used to reduce such damage. *See* INDUSTRIAL TRUCKS.

Pneumatic systems. These systems can be broadly classified as vacuum systems (negative-air systems) and pressurized systems (positive-displacement systems). The selection of a system depends on the product characteristics and rate of conveyance. In a negative-pressure system, vacuum is created by operating a positive-displacement blower located downstream. A rotary air-lock valve connected to the blower is also used to discharge the product. This type of system has been successfully used for removing wastes from processing plants (**Fig. 9**). Positive-airflow systems use a blower connected to a rotary air-lock valve at the feed end. The product is blown into the conveyor pipe. At the discharge end a cyclone is used to separate product from air; the product is allowed to fall by gravity into bulk bins. These systems are popular in conveying cereal grains and individually quick-frozen fruits and vegetables.

Tankers. In transporting granular and powdered products, pressurized road tankers have become more widely used. These tankers avoid the time wasted on filling bags at the warehouse and later emptying bags at the customer's location. Instead the product is loaded pneumatically into a tanker, which is then weighed and driven to its destination (the customer) where it is pneumatically emptied. Pneumatic transport offers many advantages, including reduced packing and unpacking costs, elimination of industrial injuries due to lifting of bags, less handling of product, and reduced rodent infestation. *See* MATERIALS-HANDLING EQUIPMENT. R. Paul Singh

Hydrocyclones

Hydrocyclones (also known as liquid cyclones) are stationary devices that utilize fluid dynamic forces to separate solid from liquid or two solids of differing density suspended in a liquid. The most effective use of hydrocyclones is in the separation of protein in a water suspension of starch in starch manufacturing.

In wet milling of corn, primary separation of starch and protein is performed with large disk-nozzle centrifuges. The crude starch fraction (1.5–3% protein) is then countercurrently washed (**Fig. 10**) by using small (0.4–0.6 in. or 10–15-mm diameter) hydrocyclones to remove most protein and other soluble materials (solubles). The smaller less-dense protein conglomerates remain in the overflow, while the heavier starch granules settle and exit the system in the underflow. *See* CENTRIFUGATION.

In the starch washing system, clean wash water enters from the downstream side to dilute the feed into the last-stage hydrocyclone. This step provides the necessary fluid volume for efficient hydrocyclone operation and helps to remove soluble protein and other solubles in the starch fraction. The overflow, containing protein, solubles, and a small amount of starch, is used to dilute the feed on the previous stage. Recycling of the overflow to the previous stage washes the solubles and protein upstream to the overflow of the first-stage hydrocyclone. The first-stage overflow is recycled to the primary disk-nozzle centrifuges.

The number of hydrocyclone stages in a starch washing system depends upon the ease of separation of starch from protein and the amount of wash water used. Some starches, such as potato starch, require only six to eight stages for adequate washing; other starches, such as corn or wheat, may require up to 14 stages. The number of stages can be reduced if the volume of wash water is increased. About 1.1–1.8 liters of wash water per kilogram of corn (0.13–0.22 gal per pound) is used for maize starch washing.

Because the capacity of individual hydrocyclone (often referred to as a cyclonette) is only 3.7 to 5.7 liters (0.98 to 1.5 gal) per minute, cyclonettes are arranged in parallel in a housing. They share common inlet and outlet manifolds. The most popular assembly, known as a clamshell (**Fig. 11**), contains approximately 480 cyclonettes. Other assembly arrangements are available.

Cyclonettes used in the starch industries are either 0.4 in. (10 mm) in diameter, used for small starch granules such as maize, or 0.6 in. (15 mm) in diameter, used for large starch granules such as potato. Two types of cyclonettes are used in a starch washing system. The A-type cyclonettes are 0.4 in. (10 mm) in diameter with a 0.1-in. (2.5-mm) feed inlet, a 0.1-in.

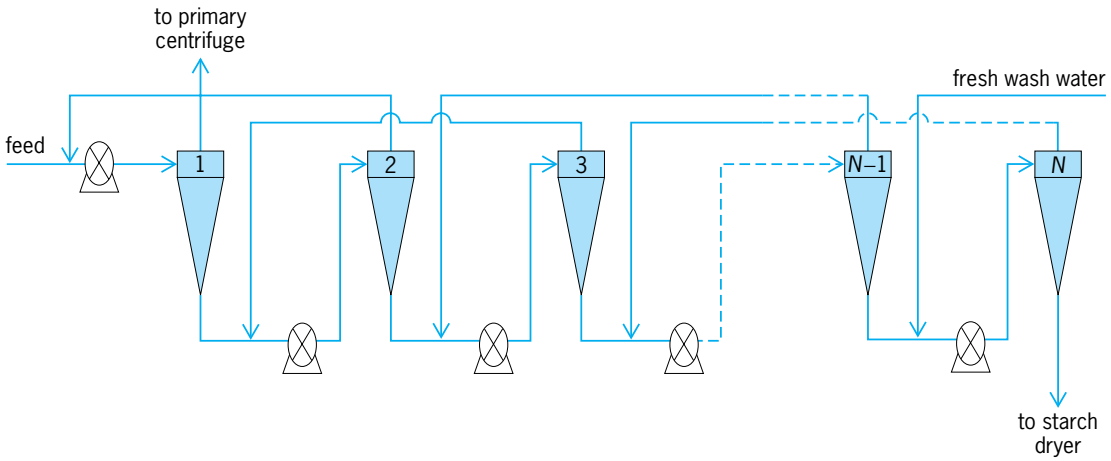

Fig. 10. Starch washing system using a number of hydroclone stages.

(2.5-mm) vortex finder opening, a 0.090-in. (2.3-mm) heavy-fraction opening, and a cone taper of 6°. This is the standard cyclonette used throughout most of the stages of starch washing. Approximately 60% of liquid flow goes to the overflow (lighter fraction) and 40% to the underflow (heavier fraction). The B-type cyclonettes are also 0.4 in. (10 mm) in diameter with similar dimensions, except that the cone taper is 8° and the heavy-fraction opening (apex) is 0.094 in. (2.4 mm) in diameter. About 70% of liquid flow goes to the light fraction, and only 30% goes to the underflow; therefore a higher starch concentration is obtained in the underflow. B cyclonettes are used in the final or last two stages of starch washing. The 15-mm (0.6-in.) cyclonettes are configured in the A or B types as described above.

Hydrocyclones/liquid cyclones have no moving parts. They rely upon slurry velocity to create adequate centrifugal forces to accelerate settling of the heavier solids. Slurry feed pressures in the range 90–140 lb/in.^{2} (620–965 kilopascals) are commonly used in the starch industry for separating starch and protein. Centrifugal separation of starch from protein in hydrocyclones depends not only on density differences but also on other starch characteristics, such as shape and size of granules.

Starch size, shape, and surface characteristics can enhance or retard separation from the smaller protein particles. Larger granules are more easily separated. Because of differences in shape, surface texture, and density of starch granules in various starches, different lower limits of starch granule size may be effectively separated from the protein among different starches. For maize, granules less than 6 micrometers in diameter will have equal probability of going out in the overflow or the underflow. This lower limit is about 1 and 5 μm for rice and wheat, respectively. The probability of a particle going out

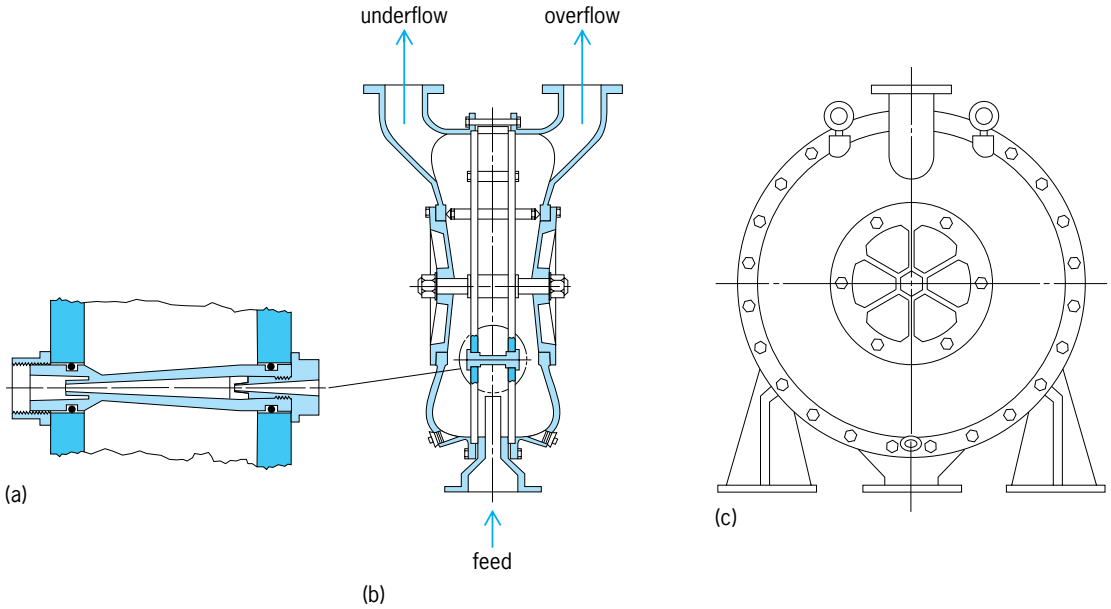

Fig. 11. Clamshell assembly of starch washing cyclonettes. (*a*) Detail showing cyclonette arrangement. (*b*) End view showing location of feed, overflow, and underflow streams. (*c*) Side view of clamshell.

with the underflow decreases with decreasing particle size for all products.

Genetic variation in the starchy raw material, environmental conditions in plant culture, and postharvest drying of the starchy raw material can also affect starch protein separation. Research on 10 United States maize hybrids representing a range of cultivars showed that although the initial and final starch granule size distributions were very similar among all hybrids, the recovery of starch varied widely. There apparently are starch density differences within a hybrid and between hybrids. Uniformity of starch granule size also influences separation. A more homogeneous starch granule distribution results in better separation.

A starch slurry containing approximately 35% starch by weight is commonly used, although the cyclone will efficiently separate more dilute concentrations. Higher concentrations of feed solids are preferred to increase production in the same size of equipment. However, very high concentration may plug the cyclonettes.

Larger hydrocyclones are used to separate germ from endosperm and fiber. Germ separation follows the first two grinding steps in the maize wet-milling process. These hydrocyclones (often referred to as germ clones) release the hydrated germ from the pericarp and endosperm. They are operated at differential pressures of 30–45 lb/in.^{2} (207–310 kPa). The germ clones are commonly 6 in. (152 mm) in diameter and 3.3–3.9 ft (1–1.2 m) long. Larger germ clones [8 in. (203 mm) or 9 in. (229 mm) in diameter] are increasing in popularity because of higher capacity, and they are less prone to plugging because of their larger overflow outlet.

Hydrocyclones may also be used in other solid-liquid separations such as recovery of suspended solids from liquid processing-plant waste effluents and separation of starch, soluble sugars, and fiber from proteins in ground oilseed slurries for protein concentration. In all hydrocyclone applications, the suspended solids must be carefully characterized so that the appropriate size of hydrocyclone and fluid flow rates for effective separation can be determined.

Steven Eckhoff

Heating and Cooling

Heat transfer is one of the most common and important unit operations in food processing. It is used to change the temperature of the food product and to change the phases present in the food (vapor, liquid, solid). Thus, preservation processes such as pasteurization, blanching, and canning are based on principles governing heat transfer, just as are drying and freezing processes. Heat-transfer operations are used to either remove or add heat in order to alter the physical, chemical, or storage characteristics of food products. An example of using heat input to change a food product is cooking to alter texture and develop color and flavor. Heat removal is used to change physical form as in the production of ice cream or candies.

Heat transfer by conduction, convection, and radiation are all employed in food processing. In the application of heat transfer, however, it is necessary to know the effects of temperature on rates of change of the physical and chemical characteristics of the food. Knowledge of food science and engineering leads to the optimum design of heat-transfer operations. *See* HEAT TRANSFER.

The food industry relies extensively on heating and cooling processes (either indirect or direct systems) to control the quality and characteristics of food products. These systems are used with a wide variety of heating and cooling media to effect temperature and phase change. In the application of these methods, the safety of the consumer is of paramount importance. Secondary criteria as to the choice of methods include production rate, cost, energy efficiency, and waste generation. Heating and cooling methods applied to foods can be classified as either indirect or direct.

Indirect methods. When these methods are used, the food product does not come in contact with the material absorbing (cooling) or supplying (heating) the heat energy. The heat is transferred to (from) the product through a physical barrier such as a stainless steel plate or a tube wall for heating (cooling). For heating processes, the heating medium is generally hot water or steam, although other heat-transfer liquids or gas flames may be used in high-temperature applications. For cooling processes, the cooling fluid is usually a commercial refrigerant cycled through a mechanical refrigeration unit.

The simplest type of indirect heat exchanger is a kettle which holds the food product. The outside of the kettle may be heated by direct flame (radiation and convection heat transfer), or the kettle may have an outer wall creating a space through which the heating medium can flow. Kettles are used in batch heating processes, and the product is usually mechanically agitated to increase heat transfer and avoid local excessive hot spots and burn-on (fouling).

For larger industrial operations, liquid food products are heated or cooled in continuous-flow heat exchangers. These take a variety of forms, including one with a tube inside another (tubular exchangers) and flow channels created by several parallel, corrugated plates (plate-and-frame heat exchanger). Milk is an example of a liquid food that is heat-treated in either type of heat exchanger prior to packaging and distribution. For more viscous food liquids, such as starch-based puddings and candies and vegetable purees (baby food), a tubular-type exchanger equipped with a rotating axial scraper blade is used. The action of the scraper blade increases heat transfer by removing the layer of liquid immediately adjacent to the wall and thus minimizes burn-on.

Direct methods. Direct heating systems are not as commonly used as indirect systems in the food industry. In cooling applications, direct systems include cryogenic freezing (using, for example, liquid nitrogen, carbon dioxide, or a commercial refrigerant), liquid immersion freezing (that is, using aqueous solutions of glycerol or sodium chloride), and cooling

with ice. The systems differ largely in their cost, rate of freezing or cooling, and effect on productbreak quality.

For heating, water, steam, and air are most frequently used as direct heating agents. In each case, the heating fluid must be of food-grade quality. For water, this presents no major problem since potable water must be used in contact with food material. For air, this frequently requires some filtering treatment prior to contact with the food. For steam, this requires that the steam be produced with FDA-approved boiler compounds. Steam which can be used in direct contact with food is called culinary steam and is frequently made by using boiler steam to make steam out of deionized potable water. When steam is used to heat food directly, the condensate that is produced dilutes the food product and it may be necessary to remove this water. Frequently this is accomplished by boiling off the water (flashing) in the cooling operation (for example, exposing the hot product to a partial vacuum). In steam heating, the steam can either be added to the product (a process called steam injection) or the product can be added to a steam atmosphere, a process called steam infusion. Steam injection is the more popular industrial method.

Heat transfer by radiation is best categorized as a direct method. Electromagnetic radiation in the infrared region (1–400 micrometers) is readily absorbed by water and many of the organic constituents of food, resulting in heat transfer. Infrared heating is used extensively in the baking and roasting industry. *See* HEAT RADIATION; INFRARED RADIATION.

Microwave and dielectric heating are accomplished by electromagnetic radiation in the radio-frequency range. In dielectric heating, the product is placed between parallel electrodes, and the charge on the plates is oscillated at 1000–3000 MHz/s. For microwave heating, electromagnetic radiation in the above frequency range is generated in a magnetron tube and transferred to the product located in a cavity where the energy is absorbed. Microwave and dielectric heating have somewhat limited applications because they are expensive. *See* DIELECTRIC HEATING; MICROWAVE.

Daryl B. Lund

Thermal Processes

The heat treatment of food is one of the important processes for the conditioning of food for preservation. This treatment accomplishes many objectives, including inactivation of microorganisms, enzymes, or poisonous compounds and production of desirable chemical or physical changes in foods.

When compared to typical engineering material such as metals or minerals, food is relatively susceptible to thermal degradation. Therefore, to produce nutritionally sound and microbially safe food products, heat treatment should be accurately controlled in order to accomplish its objectives. To assist in this control, the temperature responses of foods subjected to heat treatments are frequently estimated by using empirical or theoretical heat-balance equations.

Heat treatment. This type of processing follows different patterns. It may involve application of heat indirectly, as in a tubular heat exchanger, or it may be accomplished through the direct contact of the heating medium with the food, as in the baking of bread in a hot-air oven. The principal operations involving heat treatment of foods are blanching, preheating, pasteurization, sterilization, cooking, evaporation, and dehydration.

Blanching is a hot-water or steam-scalding treatment of raw foodstuffs to inactive enzymes which might otherwise cause quality deterioration, particularly of flavor, during processing or storage. Most vegetables and some fruits are blanched before canning, freezing, or dehydrating. The commonly used types of equipment are rotary perforated drums, rotary screw conveyors or troughs, and pipe flume blanchers. Water is the predominant heating medium in drum and flume blanchers; steam is used in the screw-type blancher. Temperature is usually 212°F (100°C) or slightly lower. The length of treatment, when preceding freezing, varies from 50 s to 10–11 min.

Preheating is a treatment used immediately before canning to ensure the production of a vacuum in the sealed container and uniform initial food temperature before further thermal treatments (pasteurization or sterilization). When the product cools, the head-space vapor condenses. Foodstuffs of liquid or slurry type are usually preheated in tubular heat exchangers which are equipped with screw conveyors if necessary. Foods of particulate type, submerged in a brine or syrup in the open-top container, are passed through a heated chamber, known as an exhaust box, in which steam is the usual heating medium.

Pasteurization is relatively mild heat treatment of food and involves the application of sufficient thermal energy to inactivate the vegetative cells of microorganisms, molds, yeasts, or enzymes that are harmful to human beings or to food quality, or both. The inactivation of bacterial spores, which are extremely resistant to heat, is not required, since pasteurized food is usually stored at refrigerator temperatures or the chemical composition of the food prohibits the germination of bacterial spores. Typical heating times and temperatures for pasteurization are 30 min at 149°F (65°C) or 15 s at 160°F (72°C). Heat sterilization requires more thermal energy than pasteurization does and is usually accomplished by heating food which is packed and sealed in containers made of metal, glass, or plastic film laminated with aluminum foil. One heat sterilization or pasteurization treatment is divided into heating and cooling phases. During the heating phase, the food is heated for the proper time by applying the heating medium, usually saturated steam or hot water maintained at the appropriate temperature. The cooling phase begins immediately following the heating phase. Since the temperature is at its maximum, or almost maximum, level at the end of the heating phase, it should be lowered as quickly as possible to avoid any unnecessary thermal destruction of the nutritional or sensory quality of the food. *See* PASTEURIZATION.

There are batchwise and continuous heat-sterilization processes which are in commercial use. Widely used for batch processing, a still retort permits food packaging in metal cans, glass jars, or polymeric film pouches with thermal processing without agitation. The heating media are normally steam for food in cans, hot water for food in glasses, and high-temperature water with air overriding pressure or steam-air mixture (typically 75% steam and 25% air) for food in pouches. Pressure above the saturation water pressure is required for pouch processing to counteract pressure buildup inside pouches, since pouch seals do not have the mechanical strength to withstand the pressure difference. Continuous processes are more economical in terms of heat energy, labor, and time than batchwise processes are, but the cost of equipment for continuous processing is generally greater. The continuous food sterilizing system, which is one of the most widely used in the United States, requires only 50% as much steam and 15–40% as much labor per unit of food processed as conventional batchwise systems need (**Fig. 12**).

The hydrostatic system (**Fig. 13**), which originated in Europe, uses about 25% less steam and water per unit of food processed than the conventional batchwise system of processing does.

The temperatures to which food is heated in conventional sterilization processes are dependent on the acidity of the food. A normal temperature range for the heat sterilization of low-acid food (food with pH 4.6 or above and with water activity equal to or above 0.80) is 221–248°F (105–120°C); and a range for high-acid food (for example, fruit) is 180–212°F (82–100°C).

Some heat sterilization methods are called high-temperature short-time (HTST) processes. An HTST process usually consists of two separate heat treatments applied in sequence. One example of these processes is a free-falling film system. Liquid food, usually milk, is preheated first to about 150°F (66°C) and carried to a falling film sterilizer by a product feed line. In the sterilizer, the food is heated as it free-falls in the form of a thin film in 520–540°F (270–280°C) culinary steam environment. The food temperature increases from about 280°F (140°C) to virtually the temperature of the steam in less than 0.3 s. In another method, liquid food with low viscosity is heated by a tubular or plate-and-frame heat exchanger. Scraped-surface heat exchangers (**Fig. 14**) are used to process highly viscous liquid foods. The heated food then passes through a holding tube for at least 3 s before cooling in a vacuum flash chamber down to 149–158°F (65–70°C) as the excess moisture is flashed off. The purpose of the first treatment is to inactivate enzymes, with most HTST processes, by heating food for 5–10 min at 149–185°F (65–85°C), the second to inactivate microorganisms, generally, by heating food for 3–30 s at 257–302°F (125–150°C). There are sound reasons for the requirement of two heat treatments in HTST processes. The rate of thermal inactivation of a biochemical or biological factor increases as the heating temperature increases. High temperatures have a great influence on microbial inactivation; but, in contrast, enzymes are less sensitive to thermal inactivation at temperatures above 248°F (120°C). Therefore, the first low-temperature and second high-temperature heatings are required for enzymic and microbial inactivations, respectively. In HTST processes, food is usually heat-sterilized before being aseptically filled into containers. Therefore, these processes are normally known as aseptic processes. The commercial application of these processes is limited mostly to liquid food since it is difficult to heat particulate foods quickly to the desired high temperatures. Another reason for such a limitation is difficulty in predicting the flow behaviors and temperatures of foods going through an aseptic processing system. However, many academic and industrial researchers are investigating ways to overcome these difficulties. Sample heat exchangers used by the researchers include scraped-surface heat exchangers (Fig. 14) and microwave heating units. Electrical, ohmic resistance heat is a basis for a heat exchanger developed in England. With this

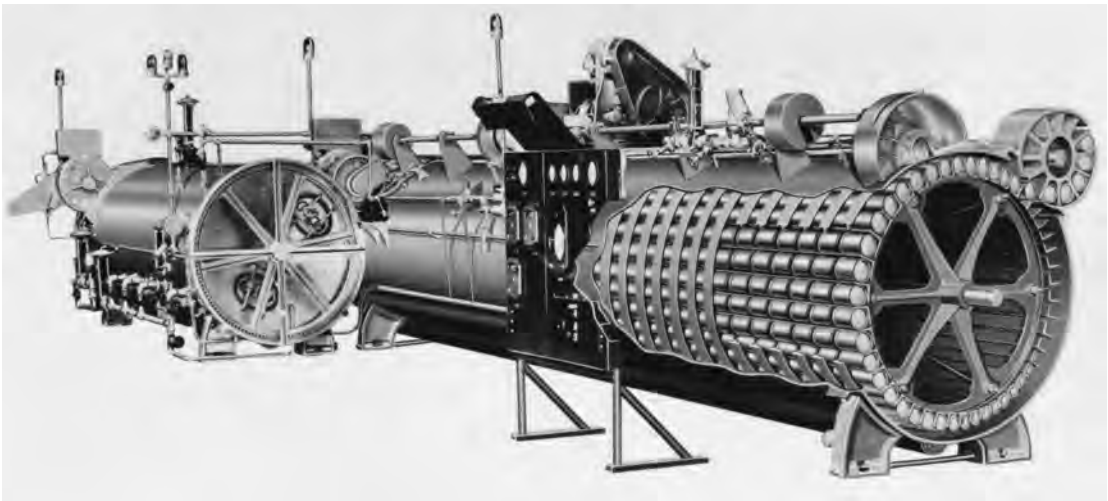

Fig. 12. Cutaway view showing a continuous-pressure food sterilizing system which is of the reel and spiral type. The smaller tank at the left is a continuous pressure cooler. (*FMC Corp.*)

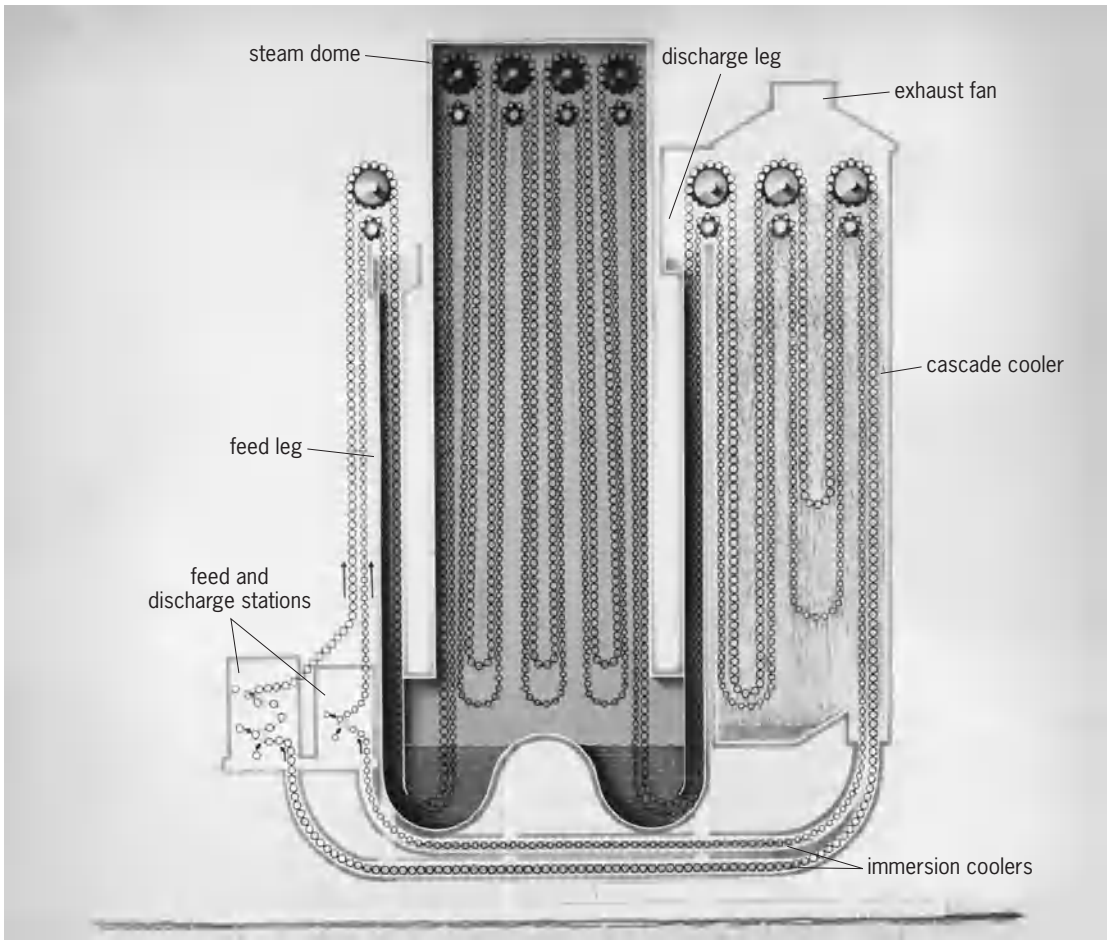

Fig. 13. Hydrostatic system of sterilizing food in sealed containers. (*FMC Corp.*)

device, an alternating current of a proper voltage is fed continuously through food passing within a plastic-lined stainless steel pipe. The pairs of electrodes are placed directly in the foodstream to provide the electric current. The claimed advantages of the ohmic heat exchanger include the ability to heat particulates quickly without disintegration.

Packaged or bulk food undergoes cooling after heating for a proper time at a proper temperature, since maintaining food at a high temperature unnecessarily causes quality deterioration. With a batch process of food packaging, the heating medium in a retort is replaced by cold water for cooling, with air overriding pressure for pouches and glass containers, or without such for most cans. In continuous-process food packaging, the cooling is accomplished by passing the packages in a cold water cooling section (Figs. 12 and 13). In a HTST process of bulk food, a water- or refrigerant-cooled heat exchanger is used. For example, a scraped surface heat exchanger is used for cooling viscous liquid exiting from a heating section.

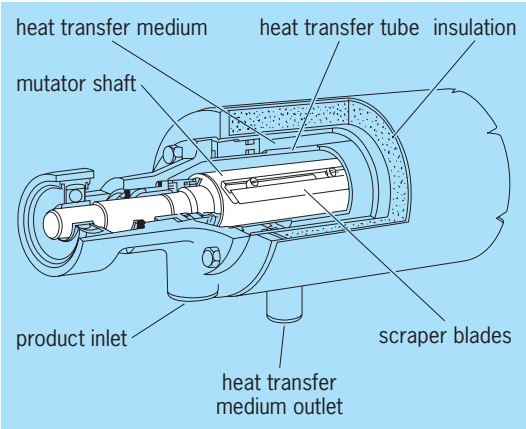

Fig. 14. Votator scraped-surface heat exchanger. (*Chemetron Corp.*)

Batch processes are flexible in operation and provide easy changeover to different products, although the processes are labor intensive. Therefore, batch processes are controlled automatically by computer-based process controllers in several thermal processing plants. Most computer controls are designed to automate starting and ending a process at a controlled temperature according to a standard sterilization process design.

A proper heating time and temperature for food sterilization may be estimated through mathematical combination of food temperature response at the slowest heating point with the thermal inactivation

TABLE 1. Commercial heat treatments for physical or chemical modification of food

Process	Heating medium	Sample food	Heating-medium temperature, °F (°C)
Baking	Direct heating with air and radiative heating with oven walls	Breads, biscuits	240–300 (120–150)
Deep fat frying	Direct heating with oil	Potato chips, noodles, doughnuts	320–365 (160–185)
Drum drying	Indirect heating with steam	Bananas, potato whey	275–293 (135–145)
Evaporating	Indirect heating with steam	Tomato paste	
Extrusion cooking	Direct heating by mechanical energy dissipation through shearing, frequently combined with direct or indirect heating by steam	Breakfast cereals, snack foods, bread substitutes*	212–230 (100–110)
Forced-air-drying	Direct heating with air	Vegetables	140–203 (60–95)
Freeze-drying	Indirect heating with heat-exchange fluid or direct heating with thermal radiation	Fish, meat	104–176 (40–80)
Microwave heating	Electromagnetic waves, sample frequencies 915 or 2450 MHz	A variety of foods requiring heating or cooking†	
Spray drying	Direct heating with air	Instant coffee, milk powders	401–590 (205–310)

*Heat is generated within food through shear (the conversion of mechanical energy to heat). Electrical heaters placed within the barrel of an extruder are used to control changes in food temperature. Typically food temperature at the inlet of an extruder is 212°F (100°C) and increases to 300–360°F (150–180°C) at the exit.

†Heat is generated within food through oscillation of bipolar molecules (for example, water molecules) at a very high rate under the influence of microwaves.

kinetic equation of target microorganisms, for example, *Clostridium botulinum* spores of low-acid food processing. Several researchers have developed estimation methods and computer programs for the estimation.

There are available commercial heat treatments whose major objectives are the physical or chemical modification of food rather than the inactivation of microorganisms or enzymes (**Table 1**).

Industrial cooling methods. Since the rate of spoilage and quality deterioration of fresh food is reduced exponentially when food temperature is lowered, food is cooled after slaughtering or harvesting. Methods for postslaughter cooling of carcasses depend on the type of animal. Beef or hog carcasses are cooled in a chilling room with an ammonia-based refrigeration system. Beef carcasses require longer cooling time because of massive size. The chilling room temperatures for beef and hog carcasses are 32–34°F and 37–39°F (0–1°C and 3–4°C) respectively, and the relative humidities are kept high, above 90%, to control weight loss during cooling. The final target temperatures are 35°F (1.7°C) for beef carcasses (average temperature) and 37–39°F (3–4°C) for hog carcasses (internal ham temperature). With beef carcasses, most cooling is done overnight (16–20 h) in a chilling room and completed during subsequent holding-room storage in about 1 day. Hog carcass cooling is completed in 8–12 h in a chilling room. Poultry carcasses are scalded for feather removal and cooled in a continuous slush-ice immersion cooler. With initial carcass temperatures of 75–95°F (24–35°C), the carcasses are cooled to 39°F (4°C) or lower in 1200–2400 s.

Fig. 15. Vacuum cooler of railroad-car size. (*Gay Engineering Corp.*)

Postharvest cooling methods for fresh fruit and vegetables depend on the type and volume of produce and the availability of cooling medium. With hydrocooling, produce is cooled by spraying or immersing in chilled water. For hauling a distance over 100 mi (160 km), hydrocooling is used for freestone peaches, clingstone peaches, tart cherries, cantaloupes, and many vegetables. Forced-air cooling is accomplished by forced flow of chilled air through the produce. For example, grapes are cooled by chilled air of 34.7°F (1.5°C) or lower and of high relative humidity with an airflow rate of 0.0027 ft^3/s lb fruit (0.00017 m^3/s kg fruit). In vacuum cooling (**Fig. 15**), water is forced to evaporate from the surfaces of vegetables or fruits by a vacuum of 29.7 in. Hg (100,292 newtons/m^2) or more created around the product. This system is especially good for products such as lettuce and spinach which have large surfaces in relation to volume. Kan-ichi Hayakawa

Freezing

Freezing is an effective method for preserving many foods. Although some foods lose significant quality

when frozen, most food products, if properly frozen and handled, will maintain a fresh condition. The shelf life of frozen foods is extended beyond that of the fresh product because the lower temperature decreases the rate of deterioration, and as the liquid water is changed to a solid, the solutes are immobilized.

Food preservation by freezing is a relatively recent phenomenon. Ice was manufactured nearly 3000 years ago by evaporative cooling and nighttime radiation cooling, but food preservation by freezing did not occur until the 1800s. Prior to the development of mechanical refrigeration in 1851, a limited amount of fish was frozen in salt-ice mixtures. Later, red meats and cheese were frozen for transport. Technical developments in the first half of the twentieth century made rapid growth of the commercial frozen food industry possible. Retail frozen foods have been significant only since 1940.

Water. Liquid water is important to many chemical and microbial deterioration processes in foods. Many foods contain 65–85% water by weight. Frequently some of this water is free to move in the food substance, but contains solutes. This water can be frozen, but the freezing point—the temperature at which the liquid water is in equilibrium with solid water or ice—is decreased by the solutes. Ice crystals contain almost no solute. As a result, when a solution is frozen by decreasing the temperature, pure water is frozen out of the system, increasing the concentration of solutes and decreasing the freezing point of the remaining solution. Thus, free water is frozen over a range of temperatures with the initial freezing point somewhat below the freezing point of pure water.

Many food systems include water that is physically bound to the solids in the food. This bound water usually is not freezable within the normal range of freezing temperatures. However, if the bonds are sufficiently weak, some of these molecules may be frozen in ice crystals (**Fig. 16**).

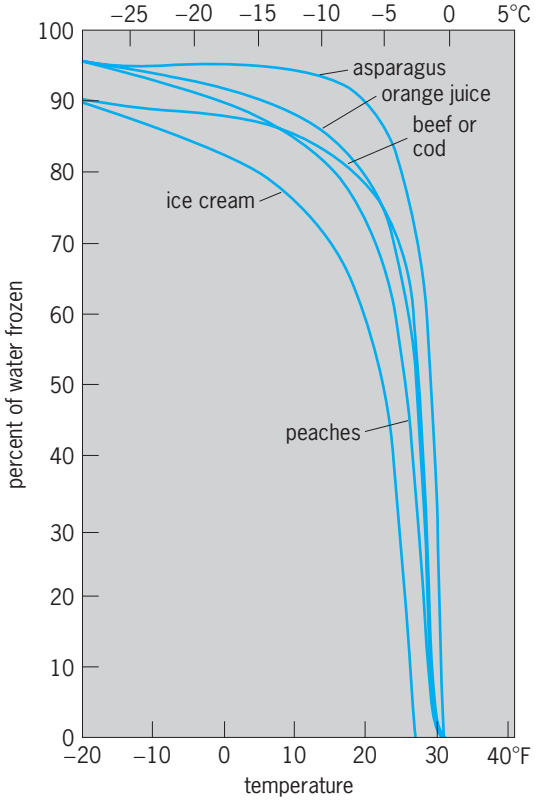

Fig. 16. Percentage of total water that is frozen at any given temperature in several food products.

Ice crystals. As liquid water turns into a solid, it forms into an organized systematic pattern or crystal. Each crystal originates as a number of molecules associating into a tiny ordered particle called a nucleus. After formation, the crystal can grow as additional molecules become associated with the original nucleus. The rates for forming nuclei and for crystal growth both increase with decreasing temperature, but the rate of nucleation is much greater than the rate increase for growth as temperature is decreased. Therefore, very rapid freezing favors the formation of numerous small crystals. During storage the crystal size and shape both change due to recrystallization, even though the product remains below the freezing point.

The location of ice crystals in tissues and cellular suspensions also depends on the rate of freezing. When the temperature decreases slowly (less than 1.8°F/min or 1°C/min), most ice crystals are formed in the locations outside the cells. Liquid water inside the cells diffuses through the cell walls to be frozen in large extracellular crystals. When the freezing occurs rapidly, both intracellular and extracellular ice crystals are formed. For most food products the formation of small ice crystals uniformly throughout the food mass results in the greatest retention of freshlike appearance and food quality.

Quality deterioration. Freezing and frozen storage are effective food preservation methods for most foods, but quality deterioration occurs during freezing, frozen storage, and subsequent thawing. The deterioration rate depends on the enzymes present, the freezing rate, the storage temperature, the consistency of the storage temperature, and the thawing method. An increase in high-quality storage life results when products are stored at lower steady temperatures (**Fig. 17**). The actual increases achieved depend on good packaging that excludes air and reduces the moisture loss from the product.

Most fresh fruits and vegetables contain enzymes that catalyze deterioration reactions during frozen storage. Many deterioration reactions can occur, but the oxidation of various food components is frequently the most important cause of quality deterioration. The naturally occurring enzymes that catalyze these reactions are generally inactivated by blanching (a heat treatment) the food before it is frozen. *See* ENZYME.

Fluctuating storage temperatures generally cause quality losses at rates greater than would be expected for the mean storage temperature. Both recrystallization and quality losses associated with chemical reactions are accelerated by varying temperatures. Small

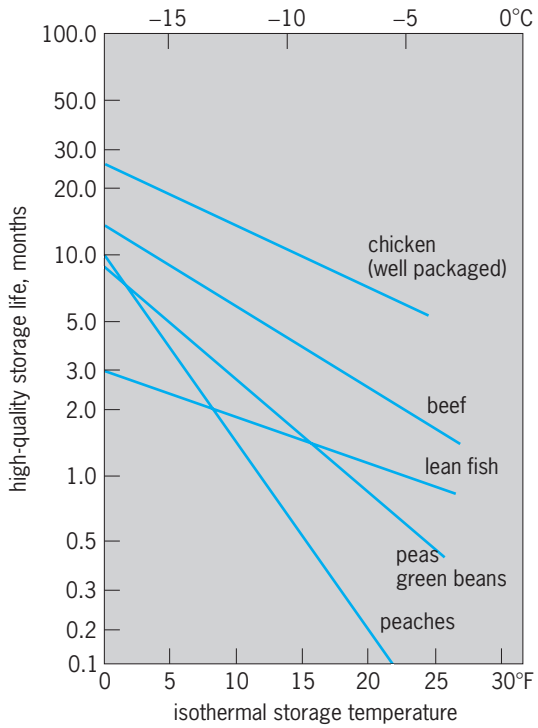

Fig. 17. Influence of isothermal storage temperatures on high-quality product life.

variations inevitably occur in any storage, but the most serious problems occur when the product is being moved from the freezer to the warehouse, between warehouses, or from the warehouses to the point of final consumption.

Thawing is inherently slower than freezing unless the thawing process is assisted by microwave energy or combined with the cooking process. Frequently the temperature gradient (ambient temperature in excess of the freezing temperature) is less than the gradient for freezing. More importantly, the rate of heat transfer through the thawed material on the outside of a food product is much slower than that through the outer frozen shell during freezing. (Of course, these comments do not apply if the product is a fluid that is allowed to flow away from the surface of the frozen core as it thaws.)

For many products there is evidence of accelerated quality loss rates when 10–75% of their water is frozen. In this range, much of the liquid water has been removed as ice crystals. Thus the remaining liquid contains more concentrated reactants that cause accelerated quality losses.

Freezing systems. Numerous commercial freezing methods are available. Each method has its own characteristics that make it more desirable for some applications and less desirable for others.

Freezing in refrigerated air. Many freezing systems use a mechanical refrigerating system to supply cold refrigerant to a set of coils that cool air that has passed through or over the food products. Thus the heat is removed from the product by cold air, and the heat is removed from the air by a mechanical refrigeration system. Freezers in homes and some commercial freezers circulate the cold air through a room or cabinet in which the food is temporarily stored. Typical air temperatures in these freezers are 0 to −22°F (−18 to −30°C). Commonly referred to as sharp freezing, it generally takes several hours to several days depending on the size of the individual food packages. In most freezers of this type the food is stationary during the freezing process, but in a few very large installations the food may move through the freezer on an intermittent or slow-speed conveyor.

Most commercial freezers using refrigerated air force the air over or through the product at a high velocity. Typical air velocities are between 100 and 3500 ft/min (30 and 1000 m/min). These freezers are generally referred to as air-blast units. In most installations the food is moved through these freezers by conveyors. The conveyors may be in a straight line, stacked above one another, or arranged as an inclined spiral. Either packaged foods or individual food particles can be frozen in these freezers. Particulate foods such as peas, beans, or shrimp are frequently frozen in a freezer in which the air is forced up through the conveyor belt and the product. The velocity is sufficiently high to create a fluidized bed, lift the product from the conveyor belt, and keep it agitated, so that each particle is individually frozen. These freezers are referred to as individually quick-frozen units. Air temperatures in air-blast freezers are generally maintained between −22 and −40°F (−30 and −40°C).

In another method of refrigerated air freezing, the air is cooled by compression, heat exchange, and expansion. The air that will remove the heat from the food is also the refrigerant. This system is designed to cool the air to −250°F (−150°C), which means it is also considered to be a cryogenic freezer. *See* CRYOGENICS.

Immersion freezers. Because the heat transfer rate between air and a solid is relatively low, many freezers are designed to contact cold liquids or solids with the food products. The product may be immersed in the cold solution, mixed with a cold powder, or sprayed with a cold solution. Salt brines and sugar solutions have been used since the early twentieth century to quick-freeze fish products and, more recently, canned citrus juice. Many novelty products are frozen in molds that are immersed in a refrigerated solution. Some products (notably, sweet corn on the cob) were frozen by a combination of spray and direct immersion in commercial chlorofluorocarbon refrigerants; these freezers are being phased out because of the environmental hazard of these refrigerants.

Both liquid nitrogen and liquid carbon dioxide are sprayed directly on food products for very rapid freezing without the capital expense of a large mechanical refrigeration system. At atmospheric pressure, liquid nitrogen boils at −320°F (−196°C). Both the vaporization of the liquid nitrogen and the sensible heat required to raise the temperature of that vapor to approximately the food temperature is used to remove heat from the food product. Carbon

dioxide exists as a liquid only under pressure. When the liquid is sprayed on a food product, it changes state to dry ice or snow at −109°F (−79°C). The snow absorbs heat and sublimes to a vapor. Both the heat of sublimation and the sensible heat to warm the vapor to near the product temperature are available for freezing the food. Freezing with either of these liquids is considered cryogenic because the phase change occurs below −76°F (−60°C).

Freezing with direct immersion or spray is generally very quick. Even for relatively large food products or packages the time is generally under an hour, and for smaller products it may require only a few minutes.

Indirect freezant-contact freezers. When freezing packaged products, it is important to maintain the intended package geometry. For example, rectangular packages tend to swell, thereby making them difficult to handle or case. By placing the unfrozen packages between refrigerated metal plates that are spaced to the intended package thickness, the dimensional characteristics can be maintained during freezing. The heat from the food is extracted through the metal surfaces and into the brine or refrigerant circulating inside the plates. Nonpackaged foods are also sometimes frozen in indirect-contact freezers using either flat plates or shaped cells. The most common indirect-contact freezers are plate freezers (with the plates stacked either vertically or horizontally and with either automatic or manual product handling), ice cream freezers (either continuous or batch style), and the novelty-product freezers (for both stick and stickless novelty items).

Heat transfer by conduction is relatively rapid, so the freezing times are typically between 15 min and 2 h, depending on the size of the item.

Refrigeration requirements. The energy that must be removed to freeze food is substantial. The refrigeration requirements include this heat plus energy losses from the freezer and refrigerating system, the heat that must be removed from any packaging materials, and heat added by product handling equipment and system inefficiencies. Typically, the food enters the freezer at a temperature above the freezing point (the temperature at which water is first crystallized in the product). Before freezing can begin, the sensible heat (the heat associated with a decrease in temperature) must be removed to bring the product to the freezing point. Typically the product temperature drops below the freezing point and then warms to the freezing point. This dip in the temperature curve (**Fig. 18**) is caused by supercooling (there are insufficient ice crystal nuclei available for growth to provide the latent heat of crystallization at the rate it is being removed by the freezer). At the supercooled temperature, additional nuclei form rapidly, and the rate of crystal growth is limited only by temperature and the rate of heat removal by the freezer. Throughout the freezing zone, the freezer is removing primarily the latent heat for crystallizing the water. There is also a small amount of sensible heat removed because of the temperature decreases associated with greater water removal (Fig. 16). In the postcooling zone, most of the freezable water has been crystallized, and the heat removal is predominantly sensible heat in the frozen product. The solid components in the food do not change phase, and therefore only the sensible heat must be removed from them.

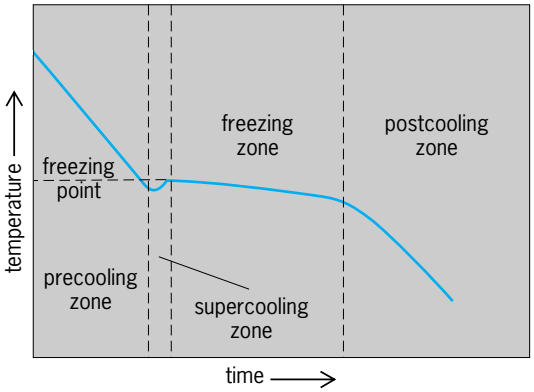

Fig. 18. Temperature profile for food freezing, showing heat transfer regimes. (*After S. D. Holdsworth, Physical and engineering aspects of food freezing, in S. N. Thorne, ed., Developments in Food Preservation, no. 4, 1987*)

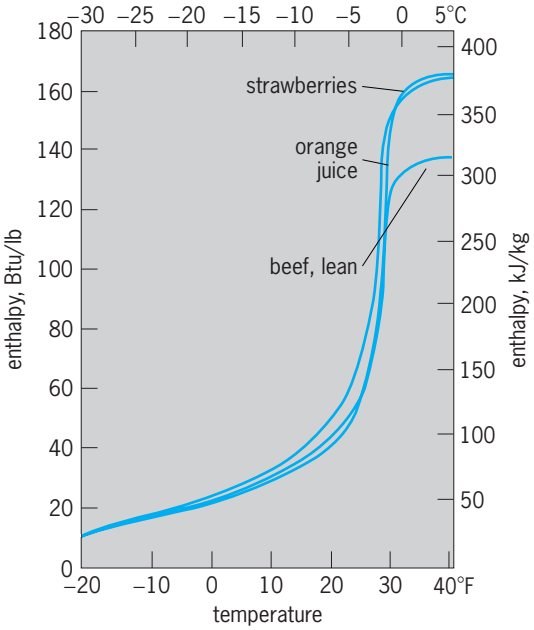

Fig. 19. Energy requirements for food freezing.

With a knowledge of the specific heats of the solid components in the food, of liquid water, and of ice, the latent heat for changing liquid water to ice, and the curve (Fig. 16) for the product being frozen, it is possible to calculate the energy requirements for freezing a product. Alternatively, the enthalpy-temperature curve (**Fig. 19**) for many products has been measured experimentally. The energy that must be removed to freeze a product is equal to the enthalpy difference between the initial and final temperatures of the product. *See* ENTHALPY; REFRIGERATION; SPECIFIC HEAT.

David R. Thompson

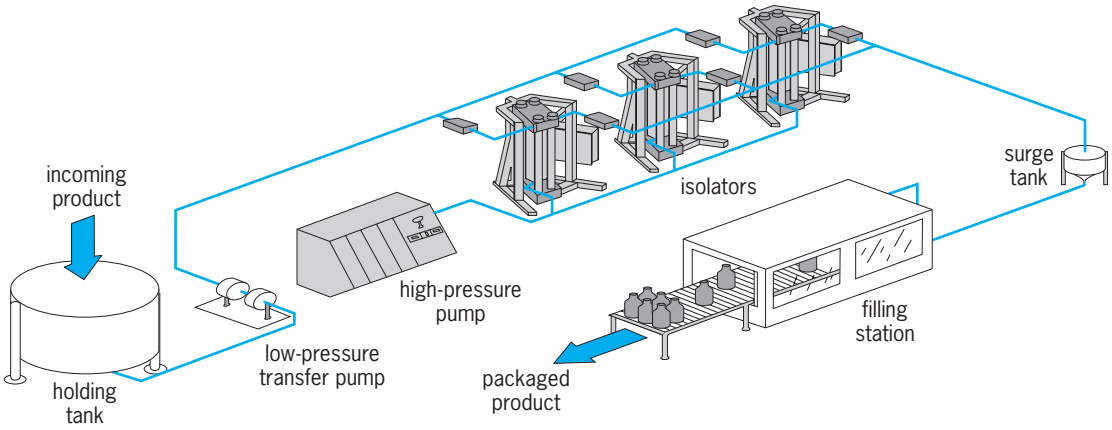

Fig. 20. Schematic diagram for a semicontinuous high-pressure operation. (*ABB Autoclave Systems, Inc.; Flow International*)

High-Pressure Processing

High-pressure (60,000–150,000 lb·in^{-2}; 400–1000 MPa) processing is a nonthermal method used to denature proteins and inactivate microorganisms and some enzymes in foods.

Advantages. High-pressure processing has many advantages over conventional methods. Since it has no effect on covalent bonds, it helps to preserve natural flavors, colors, and nutrients in foods. However, functional properties of food components, such as proteins and polysaccharides, can be altered. Due to its instantaneous and uniform distribution, it can be applied independent of the size and geometry of the product. Consequently, processing times are short, generally in the order of minutes. Once the system is under pressure, there is no loss of mechanical energy. The phase-transition behavior of water can be used to design unique freezing and thawing applications for foods. In general, liquids are used as working fluids; and unlike gases under pressure, liquids are much safer.

Disadvantages. The capital costs of high-pressure processing are considerably more than conventional methods. In addition, spores of some microorganisms are not inactivated, as well as some enzymes. Due to the compressibility of the working liquids, there is generation of heat during compression. Depending on the pressure level, this heat can be significant and void the nonthermal advantage of this process. Gases in the package are undesirable and may cause distortion of the package. At present, most commercial high-pressure operations are limited to the batch type. There is a lack of information

TABLE 2. Commercially available food products treated with ultrahigh pressure

Product	Pressure, MPa	Time, min	Temp., °C	Purpose
Fruit-based: jams (apple, kiwi, strawberries); jellies; purees; yogurts; sauces	400	10–30	20	Pasteurization, improved gelation, faster sugar penetration, residual enzyme activity
Grapefruit juice	200	10–15	5	Reduced bitterness
Mandarin juice	300–400	2–3	20	Reduced odor of dimethyl sulfide, reduced thermal degradation of methyl methionine sulfoxide, replaces first thermal pasteurization
Sugar-impregnated tropical fruits (kept at −18°C without freezing)	50–200			Faster sugar penetration and water removal
Raw ham	250	180	20	Faster maturation, faster tenderization, effect on microorganisms, improved water retention
Raw squid + salt + autolysis, raw scallops				Microbial sanitation, tenderization, control of autolysis
Fish sausages, terrines, and puddings	400			Gelation, microbial sanitation, good texture of raw gel
Raw sake (rice wine)				Yeast inactivation, fermentation stopped
Mochi rice cake, Yomogi fresh aromatic herbs, hypoallergenic precooked rice	400	10	45 or 70	Microbial reduction, enhances rice porosity and salt extraction of allergenic proteins
Ice-nucleating bacteria (used for fruit juice and milk concentration)				Inactivation of *Xanthomonas* bacteria by pressure without loss of ice-nucleating properties
Spiced mashed avocado (guacamole)	400–700			Pasteurization
Fruit juices, yogurt, and acidic foods	400–700			Pasteurization

regarding high-pressure kinetics of enzyme or microbial inactivation, and quality/functionality retention or deterioration. There is also insufficient information regarding the mechanisms involved in the application of high pressure. This includes the requirement to understand, control, and optimize pressure effects on a cellular/molecular level. In addition, more insight is needed to understand the interactions between food constituents under high pressure.

Treatment. In general, treatment with high or ultrahigh pressure (UHP) is accomplished by hydrostatic pressure; that is, the high pressure is indirectly applied to the food, sealed in a flexible or semirigid package, by the use of a working fluid. The process can be batch or semibatch.

The batch process is currently the preferred method for food treatment with high pressure, mainly to minimize problems related to hygiene, flexibility, and technical feasibility. In batch processing, the food is sealed in a hermetic flexible package and put in a pressure chamber, typically a cylindrical pressure vessel made of low-alloy steel, in which pressure is applied through the pressure-transmitting liquid. The production cycle includes the steps of loading the product into the chamber, pressurizing and depressurizing cycles, and unloading of the product. The work of compression can be delivered directly or indirectly to the working liquid. In the case of direct application of pressure, the working liquid is compressed by a piston driven into the cylindrical chamber by a low-pressure pump at one end, taking advantage of the hydraulic principle. This allows a short compression-cycle time, but the operation is limited by the high-pressure dynamic seal between the piston and the internal vessel surface. This type of operation is generally limited to small-capacity systems. The indirect compression method uses a high-pressure intensifier pump to force the working liquid into the pressure vessel until the desired pressure is reached. In this case, only a static high-pressure seal is needed.

By using several pressure vessels and alternating their cycles, a semicontinuous operation can be obtained (**Fig. 20**). In this case, while a vessel is being loaded, another vessel is kept under pressure, and still another is unloaded. Alternatively, pressure can be accomplished by the use of an aseptic continuous system. Current commercial applications of this technology are shown in **Table 2**. *See* HIGH-PRESSURE PROCESSES.

Murat O. Balaban; Roger D. Barbosa

High-Intensity Pulsed Electric Fields

Pulsed electric fields (PEFs) is an emerging nonthermal food preservation technology used to pasteurize and sterilize pumpable liquid foods (such as milk, fruit juices, liquid whole egg, and soups) without using lethal heat, ionizing radiation, or chemicals. Products are minimally affected by PEF processing. Very short pulses of high-intensity electrical fields produced by a high-voltage source can significantly reduce the level of pathogens as well as spoilage microorganisms in just a fraction of a second. PEF processing is considered superior to thermal processing because it avoids or greatly reduces the detrimental changes to the sensory, physical, and nutritional properties of food caused by thermal processing. The application of PEF technology to food preservation provides an alternative to thermal processing by producing freshlike products with high nutritional qualities, no additives, and longer shelf lives. Furthermore, PEF assists in reducing preheating and posttreatment cooling of some food products; therefore, energy consumption is considerably reduced compared to conventional thermal processes.

Process. During the PEF process, microorganisms are inactivated by irreversible structural changes in the membrane, leading to formation of channels or pores and destruction of the semipermeable barrier of the membrane. The lethality of PEFs on microorganisms is caused by the action of high-intensity electric fields, and is not due to temperature effects. The high-intensity PEFs cause disruption, permeabilization, and lysis of the cell membrane. This bactericidal effect is related to the electric field intensity, treatment time, and number of pulses, frequency, and pulse width. Microbial inactivation with PEFs is also dependent on the type of microorganisms, microbial growth stage, initial concentration of microorganisms, medium conductivity, and ionic strength. The basic components of a continuous PEF system consist of a high-voltage power supply; an energy storage capacitor; a treatment chamber; a pump to conduct food through the treatment chamber; a cooling device; voltage, current, and temperature measurement devices; and a computer to control operations (**Fig. 21**).

The types of high-voltage power supplies that may be used to charge the capacitor are generated by an ordinary source of direct current (dc). Alternating-current (ac) power from the utility line (60 Hz) is transformed into high-voltage ac power and then rectified to a high-voltage dc. Another way to generate high voltages is with a capacitor-charging power supply which uses high-frequency ac inputs and provides a command charge with higher repetitive rates

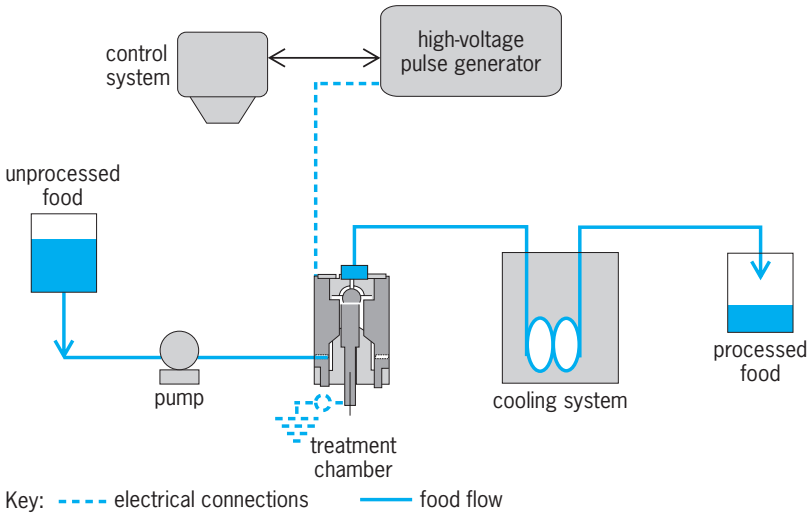

Fig. 21. Continuous pulsed electric field (PEF) processing system.

than a dc power supply. Energy from the high-voltage power supply is stored in the capacitor and discharged through the food material to generate the necessary electrical field in the food.

The energy stored in the capacitor can be discharged almost instantaneously (in a millionth of a second) at very high levels of power. The discharge is accomplished by using high-voltage switches that must be able to operate reliably at a high power and repetition rate. The type of switches can be selected from gas spark gaps, vacuum spark gaps, solid-state switches, thyratrons, and high-vacuum tubes.

The treatment chamber is used to transfer high-intensity PEFs to foods. Cooling water is recirculated through the electrodes, thus controlling the temperature of a sample in the gap formed by the two electrodes in the treatment chamber. A variable speed pump regulates the flow rate of the food during treatment. Chambers with uniform electric fields subject all bacterial cells to the same potential. This is advantageous for electroporation, where no alignment or pearl-chaining is required, and can result in high yields provided the field strength is set at an optimum value. However, since the yield might drop substantially above or below the optimum field strength, the optimum value is not always known. A practical realization of uniform field chambers is a chamber with flat parallel electrodes or a coaxial chamber with a diameter many times larger than the electrode gap. Chambers with nonuniform fields allow alignment at low cell densities with an ac field for fusion and reduce the need to have exactly the right field strength to obtain high transfection yields.

Electrical parameters such as voltage and current waveforms applied to fluid foods can be recorded via a digital data acquisition system. A digital storage oscilloscope records outputs from voltage and current monitors. The computer and oscilloscope should be placed in a shielded area to minimize electromagnetic interference, and electrical and flow parameters should be selected so that each unit volume of the food is subjected to the necessary number of pulses to create the desired inactivation of microbial cells.

After processing, the food is packaged into sterile containers. It is extremely critical to maintain a contaminant-free environment during processing and packaging to avoid cross-contamination of the product. All of the equipment and work area must be thoroughly cleaned and sanitized before processing, and the personnel involved in the operation should be properly equipped to maintain a clean working area.

Application in food preservation. The process of subjecting a food to electric fields may be accompanied by the generation of heat. If electric fields are applied to a food product in the form of short-duration pulses rather than passing electricity directly through it, minimal heat is generated and the process remains nonthermal. Since the pulse repetition rate also contributes to heat generation, the processing system generally includes a means to provide cooling of the treatment chamber, which is one of the most complicated components of the processing systems. Usually, static chambers are used for preliminary laboratory experiments, while continuous chambers are used in pilot plant and industrial operations.

Due to the advantages that PEF presents as a nonthermal process, products such as fruit juices, milk, yogurt, beaten eggs, and tap water have been under extensive research in order to implement the process at an industrial level. Flavor freshness, low energy utilization, and extended shelf life are some of the virtues attributed to PEF treatment. It has also been demonstrated that the textural and functional attributes of some products can be preserved at safe microbiological levels. **Table 3** lists food treated with PEF as well as the process conditions used.

PEF-treated beaten eggs are just one example of the appeal this technology has, since egg products processed by traditional heat treatments exhibit operational problems due to protein coagulation in heat exchangers. However, since PEF enables microbial inactivation at lower temperatures, it represents an important alternative to the egg industry. Significant shelf-life extensions with minimal changes in the physical and chemical properties of certain other foods have also been demonstrated. Experimental results have shown that, in general, the sensory properties of foods are not degraded by PEF. In fact, electric-field-treated products, such as green pea cream soup and fluid eggs, were preferred by panelists in sensory tests over at least one commercial brand product.

Electric-field treatment is also associated with energy efficiency. In the treatment of apple juice, energy utilized with the PEF technology is 90% less than the amount needed for the high-temperature short-time (HTST) processing method. *See* ELECTRIC FIELD.

Gustavo V. Barbosa-Cánovas; Juan J. Fernández-Molina; Barry G. Swanson

TABLE 3. Liquid foods processed with pulsed electric field (PEF)

Food product	Electric field kV/cm	Pulse duration, μs	Pulse number	Treatment temperature, °C
Fresh orange juice	35–60	0.85–1.46	2.5–3.3	25
Cranberry juice	39	4	50	30
Fresh apple juice	50	2	16	45
Skim milk	30–40	2	30	28
Raw milk	40	2	20	50
Green pea soup	35	2	32	53
Liquid whole eggs	40	2	50	30

Concentration

Water-rich liquid foods are concentrated to (1) provide storage stability; (2) facilitate subsequent drying; (3) produce saturated solutions and cause crystallization; (4) produce supersaturated solutions which will form glassy or amorphous solids; (5) reduce storage, shipment, and packaging volumes and costs; and (6) induce flavor and texture changes. Liquid foods are usually concentrated by evaporation; but freeze concentration, reverse osmosis, and ultrafiltration are also used.

Evaporation. Food evaporators usually contain heat-transfer tubes surrounded by a steam- or vapor-filled shell (**Fig. 22**). The liquid to be concentrated flows through the tubes. Part of the water in the liquid is vaporized by the heat that is provided by the steam or vapor (which condenses outside the tubes). The water vapor and remaining liquid separate in a settling chamber or cyclone and leave the evaporator through separate lines.

In most cases, groups of evaporators are interconnected to form multistage, multieffect systems (**Fig. 23**). Liquid feed enters the first stage, and concentrated product is discharged from the last stage. Partially concentrated liquid is pumped out of the first stage into the second stage, and similar transfers occur in the remaining stages. Steam or recompressed vapor is fed into the first-effect steam chest, and the vapor produced flows into the second-effect steam chest. Similar transfers occur in the remaining effects. The vapor from the last effect is condensed externally, and the condensate is discharged through a barometric leg. Steam jet ejectors or a liquid-ring vacuum pump are used to remove noncondensables, which interfere with heat transfer, from evaporator systems.

The evaporation in successive effects of a multieffect system occurs at progressively lower pressures and, consequently, at progressively lower temperatures. This provides temperature differences, which cause heat transfer from the condensing vapor to the evaporating liquid. The heat provided by vapor condensation generates new vapor from the evaporating liquid, which in turn generates more vapor in the following effect, and so on. In a well-designed, multieffect system, the amount of water evaporated roughly equals the amount of steam used in the first effect times the number of effects. However, if steam is used to heat the liquid feed to its boiling point, less than this number of unit weights of water will be evaporated. If the feed is preheated by using vapor from the effects (as shown in Fig. 23), the weight of water evaporated will be very close to the number of effects times the weight of steam used. In beet sugar refineries, vapor is removed from evaporator effects to provide process heating. A smaller ratio of vapor to steam used results, but overall energy consumption in the refinery is reduced.

When evaporation conditions are stable, high-pressure steam passed through an ejector can be used to compress part of the vapor produced in a multieffect system. The compressed vapor and steam provide the heat used in the first effect. When vapor from effect m in an n-effect system is compressed by using this arrangement, $n + m$ unit weights of evaporation are obtained per unit weight of steam used.

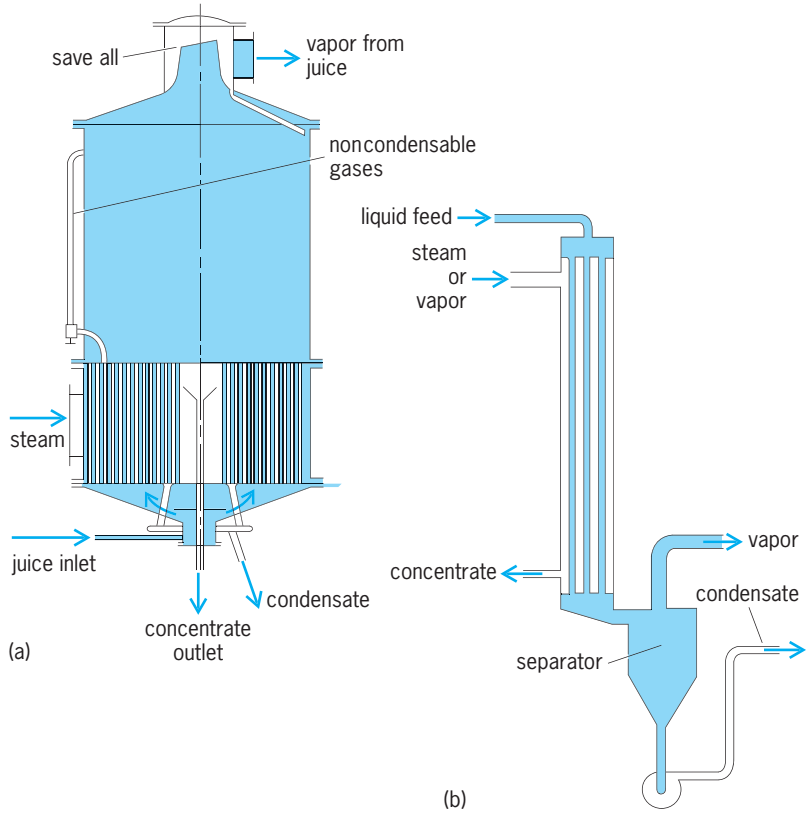

Fig. 22. Single-effect food evaporators: (*a*) short-tube type, (*b*) long-tube type.

Relatively low first-effect temperatures (for example, 162°F or 72°C for skim milk) must be used to prevent product degradation or heat-transfer surface fouling. Temperatures in the last effect, which are controlled by the condensation temperature of the vapor leaving the system, usually range between 104 and 113°F (40 and 45°C). The multieffect system is driven by the difference between the first-effect steam temperature, minus the sum of the boiling point elevations in the effects, and the evaporation temperature in the last effect. The temperature difference per effect is directly proportional to this temperature difference and inversely proportional to the number of effects used. Hence, the total heat-transfer area required in a multieffect system is directly proportional to the number of effects.

Mechanical vapor recompression. This process has been increasingly used to raise the pressures of vapors generated in multistage and multieffect systems. Either the compressed vapors are sent back to the shells of all stages in parallel (**Fig. 24*a***), or vapor from the last effect is sent back to the shell of the first effect (in Fig. 24*b*). Parallel stages are sometimes mounted in a single shell to reduce equipment costs. Very large centrifugal compressors are usually used to handle the vapor flows generated in food evaporators. The maximum temperature increase which

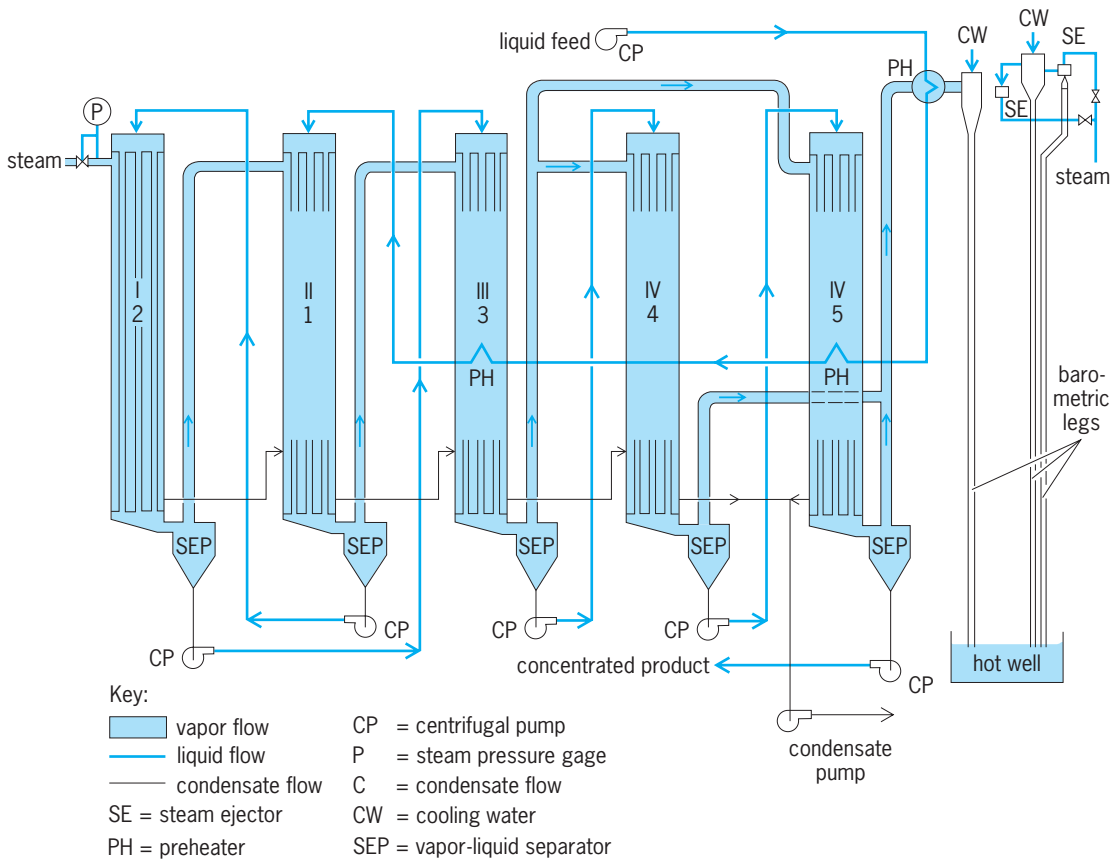

Fig. 23. Multiple-effect evaporation system. Stage numbers (arabic) denote the direction of liquid transfer, and effect numbers (roman) denote the direction of vapor transfer. The first stage is in the second effect, and the fourth and fifth stages are in the fourth effect.

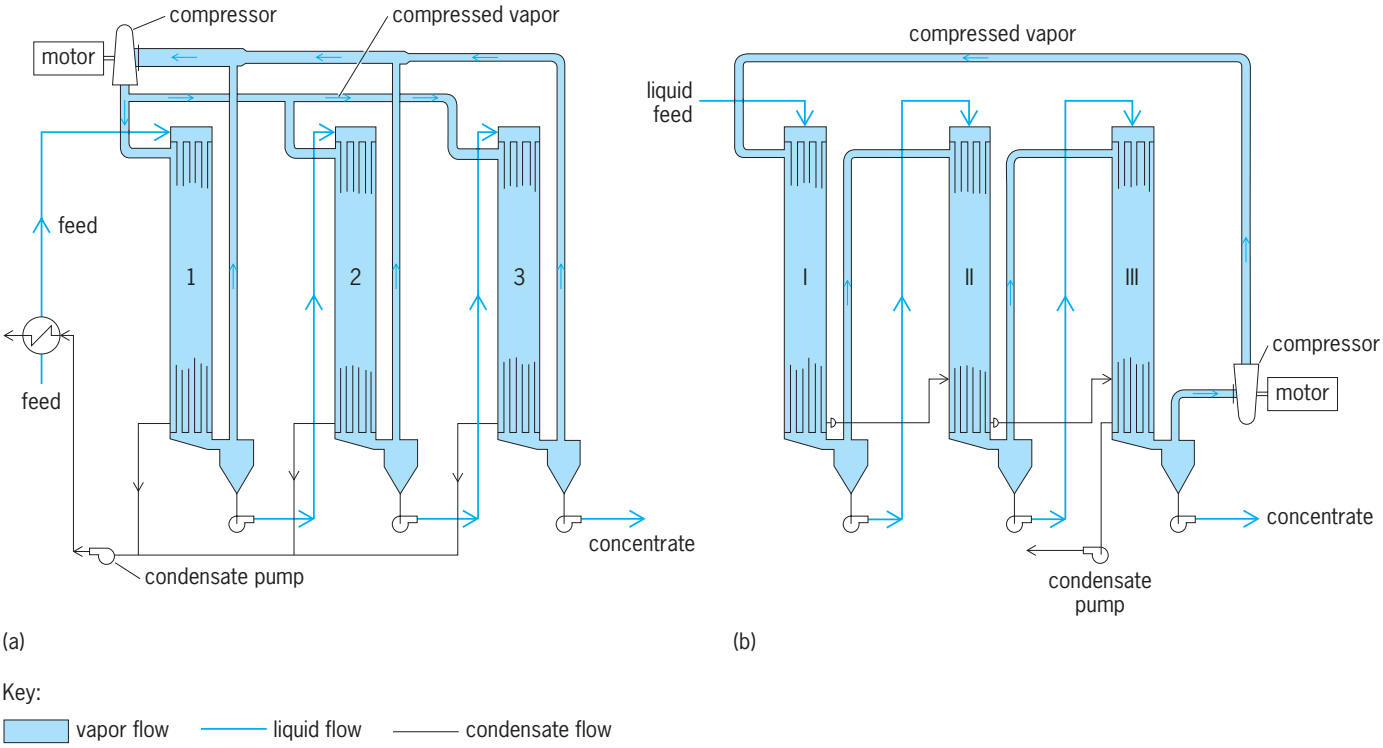

Fig. 24. Vapor-recompression evaporation systems: (*a*) multistage, parallel-mechanical type (arabic numerals are stage numbers); (*b*) multiple-effect mechanical type (roman numerals are effect numbers).

such compressors can provide ranges from 23 to 29°F (13 to 16°C). Lightweight, low-compression-ratio, centrifugal blowers that provide small temperature rises but very high thermal efficiency are used in large parallel-stage systems. In small systems, double-lobe blowers, which can provide larger temperature increases, are used. The energy used to drive a mechanical-vapor-recompression compressor providing a 23°F (13°C) temperature difference roughly equals the heat that would be used in a steam-driven, 17-effect system. However, in terms of the fuel energy required to produce the mechanical energy which drives the compressor, the mechanical-vapor-recompression system corresponds to a 7.5-effect system.

Tubes. Long, vertical tubes (Fig. 22*a*) are used to facilitate single-pass operation in most modern evaporators. Liquid feed, which is pumped into the tops of the evaporator stages, flows in falling films down the inside walls of the tubes. Single-pass, falling-film operation tends to minimize product holdup time and degradation. Falling-film operation also provides the unimpeded vapor flow needed in mechanical-vapor-recompression systems.

Arrays of short vertical tubes, which pass through a steam- or vapor-filled shell called a calandria (Fig. 22*b*), are used in most older evaporators. Many evaporators of this type are used in sugar refineries. Pools of liquid fill the bottoms of such evaporators and barely cover the tubes. Natural convection, promoted by vapor bubble formation in the tubes, causes liquid to recirculate through the tubes and a large-bore downcomer. Because of this circulation, liquid-holdup times are relatively long, and significant product degradation (for example, inversion) occurs.

Instead of vertical tubes, horizontal or tilted tubes and plate- and panel-type heat-exchange surfaces are sometimes used in evaporators. Forced, multipass circulation is used to prevent fouling and facilitate heat transfer when viscous products, such as tomato paste, are concentrated. Scraped heat-transfer surfaces are often used when very heat-sensitive or highly viscous products are concentrated. Rapidly rotating, conical heat-transfer surfaces are also used for such products. Concentrations in the 90% range can be obtained, for example, for candymaking, by using slightly pitched scrapers that both provide thin films and propel concentrates toward the discharge end of the evaporator. The concentrates obtained are so viscous that an expeller is used to pump them out of the evaporator; a large part of the heat used for evaporation is produced by friction.

Aromas and volatile flavors are invariably lost when fruit juices are concentrated by evaporation. Aromas can be recovered from vent vapor leaving the steam chest following the first effect. Carryover of aroma in vent vapor is improved by using countercurrent condensation in that steam chest. Aroma losses can be almost completely avoided by using freeze concentration instead of evaporation. *See* EVAPORATION; EVAPORATOR.

Freeze concentration. In freeze concentration units, liquid feed passes through a refrigerated, scraped-surface heat exchanger, where part of the water contained in the feed is converted into small ice crystals. This concentrates the remaining liquid. Centrifuges are sometimes used to directly separate the concentrate from the ice, but clean separations are not achieved when the slurry sent to the centrifuge contains small crystals. Therefore, the ice-concentrate slurry is usually sent to a "ripening" tank which contains larger crystals. The small crystals melt because of their excess surface energy. This subcools the concentrate with respect to the large crystals, which consequently grow larger. Ripening produces large crystals, which can be cleanly separated from concentrate. The ice-concentrate mixture is often pumped upward through a wash column. Downward flows of water, obtained by melting the ice, displace and expel the concentrate from the upward-moving bed of large crystals. When low displacement velocities are used, very clean separations are obtained.

In theory, very little energy is required for freeze concentration. Freezing and ripening are carried out at a sequence of temperatures in multistage units, and discharged ice is used as a heat sink to provide higher energy efficiency and faster processing at lower cost. Nevertheless, energy requirements for freeze concentration are still higher than for evaporation.

Concentrations obtained by freeze concentration had been limited to 40–45%, because cold viscous concentrates could not be readily separated from ice. However, centrifuges have been developed to carry out this separation, and concentrations in the 50–60% range can be obtained. This is close to the theoretical maximum concentration (roughly 65% of sugar solutions) obtainable by freezing.

Reverse osmosis and ultrafiltration. In reverse-osmosis systems, water is forced through selectively permeable membranes by using pressures which exceed the osmotic pressure of the solution being concentrated. Reverse osmosis is primarily used to purify water and to concentrate milk, whey, and dilute extracts. Apple juice and maple sap are sometimes preconcentrated by reverse osmosis prior to evaporative concentration. *See* OSMOSIS.

In ultrafiltration systems, relatively low pressures are used to force water through porous membranes which retain high-molecular-weight solutes. Any low-molecular-weight solutes which are present pass through with the water. *See* ULTRAFILTRATION.

Reverse osmosis and ultrafiltration require very little energy, but cannot provide food solute concentrations which are much higher than 20%. Freeze concentration is much more expensive than evaporation. Because of these limitations, evaporation will continue to be used to concentrate liquid foods in most cases.

Henry G. Schwartzberg

Dehydration

Drying is one of the oldest methods of food preservation; it is also one of the most difficult to model and

control. Dehydration is usually considered a supplement of drying. A substance can be dried beyond its capability of restoration, but a dehydrated product is essentially in its original state after adding water. The distinction between evaporation and drying is subtle. Evaporation generally refers to removing water vapor at its boiling point. In drying, the water is usually removed below the boiling point as vapor by air. Many techniques have been developed to dry food products, but the mechanisms involved in the process are always the same. The basis of dehydration is simple: drying transfers energy into the product and transfers moisture out. The energy supplies the necessary heat, usually the latent heat of vaporization, and is responsible for the water migration. Materials may gain or lose moisture, depending on the relative humidity and temperature of the surrounding environment.

The purpose of food dehydration is to decrease microbial growth, enzyme activity, and deteriorative reactions by removing the available water in the product. Many foods are dried to a final moisture content of about 5%, but this varies with the food, package, and intended use. Dehydrated foods are ideally suited for prolonged storage, and among other benefits, dehydration may enhance the texture and flavor of the product or preserve desirable nutrients. Dehydrated foods lower transportation costs due to the reduction in product volume, and thus are conveniently handled.

However, consumer reluctance to accept the product can be a major drawback of dehydrated foods. Drying typically results in a number of changes to the material. These include bulk density change; case hardening and toughening; heat damage, which may include browning; effects on enzyme activity, microorganisms, pigments, proteins, fats, and carbohydrates; loss of ability to rehydrate; loss of nutritive value; loss of volatiles, for example, flavor; migration of soluble constituents; and shrinkage. These changes may or may not be desirable.

Successful design of a drying process requires a complete knowledge of the changes in the product during dehydration and an understanding of the variables which may affect the drying rate. This entails detailed information of the type of food being dried and its physical properties. Water availability and the type of so-called bound water present, as well as drying-rate concepts, are other important parameters to consider.

Water availability. Dehydration is primarily a preservation technique. Dehydrated foods remain shelf-stable without sterilization due to the small amounts of available water in the product. Microorganisms, enzymes, and other biological and chemical reactions cannot function in the absence of water. The availability of water is determined by the moisture content, water molecule interactions, and food component structure. Water binding is a measure of the molecular interactions and may be estimated by various methods. The most descriptive of these determines water sorption isotherms (**Fig. 25**), a relationship between the moisture content in food and the water activity. Water availability, then, rather than moisture content, is the influencing factor when dehydrating a food product.

Water activity. The water activity a_w is defined in terms of the mole fraction of the solvent, since at equilibrium the mole fraction of water in food is usually unknown and not easily measured. In addition, most foods do not behave as ideal solutions because of capillary forces, chemically bound water, and incomplete ion dissociation. Since a_w is equal to the water concentration in the liquid phase, it is the ratio of the water vapor pressure of the solution to the vapor pressure of the pure water at the same temperature, and therefore the nonideal behavior of food will be accounted for. Water activity indicates the relative ease with which water could be removed from a product during drying, since a_w expresses the relative humidity at which products are in equilibrium with their environment. Lowering a_w generally lowers the susceptibility of food to microbial spoilage. Typically, a value $a_w < 0.90$ prevents bacteria growth, $a_w < 0.80$ inhibits enzyme activity, $a_w < 0.60$ inhibits Maillard reactions, while $a_w \approx 0.20$ is the value needed for dehydrated food. *See* VAPOR PRESSURE.

Bound water. Bound water may exist under several conditions. Unbound water is found in cell pores or fibers, where soluble constituents cause a lowering of vapor pressure. Capillary forces also tend to exert a lowered vapor pressure on free water due to the concave curvature of its surface.

In food material, three types of water are usually present. These are as follows, in order of decreasing difficulty to remove: water bound to ionic groups, water molecules that are hydrogen-bonded, and

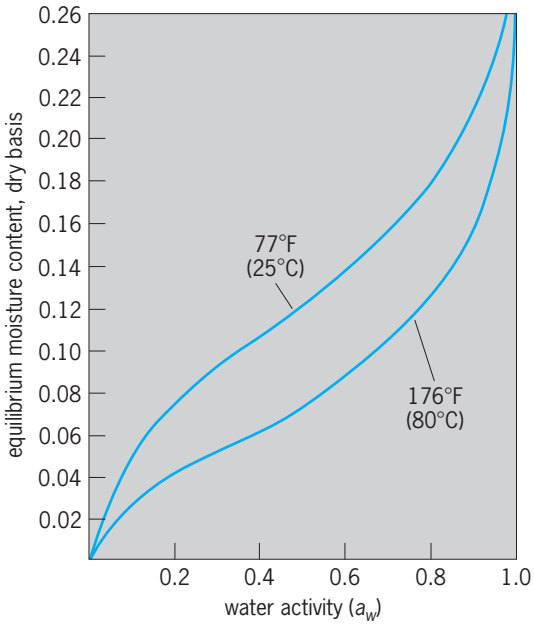

Fig. 25. Equilibrium isotherms for wheat. The values for dry basis represent unit mass of water per unit mass of dry solid. (*After M. Fortes, M. Okos, and J. Barret, Heat and mass transfer analysis of intrakernel wheat drying and rewetting, J. Agr. Eng. Res., 26:109–125, 1981*)

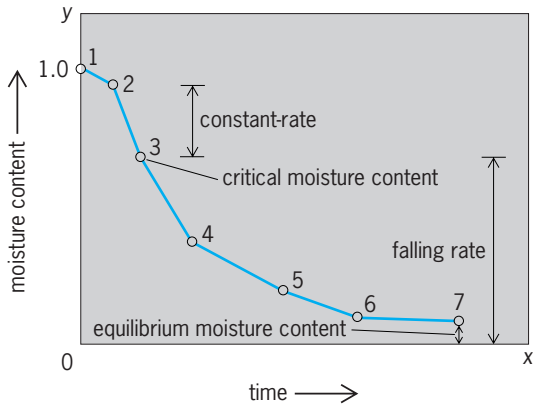

Fig. 26. Typical drying-rate curve for a biological material.

unbound free water. Certain energy requirements are associated with each level of water removal: first free water is removed, followed by hydrogen-bonded molecules, and then ionic group attachments. Since only a certain amount of energy is available in drying operations, different types of bound water are removed at different rates.

Drying rates. Criteria that hold for most dehydration processes are the characteristic drying regions of a material, which are known as the constant-rate period, the transition zone, and the falling-rate period. Each region is characterized by the type of resistance that controls the process. In the constant-rate region, external resistances control mass transfer. In the falling-rate period, internal resistances control mass transfer. The transition is the interface of the two, and is controlled by both mechanisms. It also contains the critical moisture content, defined as the moisture content where the internal resistance equals the external resistance.

In the constant-rate period a nearly linear relationship of moisture loss over time exists (**Fig. 26**). The rate of moisture removal is limited only by the rate of evaporation, and is dependent on air velocity, temperature, humidity, piece size, and bed depth. In this period, unbound, free moisture is always present on the surface of the material being dried. Moisture is able to diffuse to the surface faster than the rate at which it leaves, and the rate of evaporation is independent of the material. The drying rate is usually determined by the heat transfer. If heat addition is predominantly by forced convection, the surface temperature can usually be approximated as the wet-bulb temperature of the air stream. If the heat transfer occurs entirely by conduction, the surface temperature of the material is approximately equal to the boiling point temperature. If both conduction and convection heat transfer takes place, the surface temperature of the material will be somewhere between the wet-bulb and boiling point temperature. If heat transfer is by radiation or dielectric heating, other more complex relations must be used.

At the point of critical moisture content, the drying rate begins to proceed at a decreasing rate. After that point, an equilibrium moisture content is approached. This is the falling-rate period (Fig. 26), and there may be more than one. The drying rate is usually affected more by changes in air temperature and humidity and by piece size than by other conditions. The water within the material can no longer diffuse to the surface at a rate fast enough to keep the entire surface saturated with moisture, and the surface area decreases. The drying rate subsequently decreases as the unsaturated surface area increases. The material properties become very important and influence the drying rate. This is because moisture movement in the falling-rate period is in part determined by capillary flow, liquid diffusion, surface diffusion, and water-vapor diffusion, and is a function of the material.

Modeling the constant-rate region can be accurately accomplished by application of heat- and mass-transfer equations. The falling-rate region, however, is much more difficult to assess theoretically because the mechanism of transport is difficult to determine, it may involve a variety of mechanisms, and it is affected by external factors such as shrinkage and case hardening and toughening.

Categorization. Drying occurs by simultaneous heat and mass transfer, and may be classified in several ways. When drying is categorized according to the condition of heat transfer to the product, two general processes are considered: direct and indirect heating (**Fig. 27**). In direct drying, heat is added by direct contact to the product. Heated air is circulated in a chamber and forced over the food. The moisture saturated air leaves and removes the water vapor of the product. In indirect drying, heat is transferred indirectly to the material through an intervening medium. The product may receive heat through a plate or tray on which it is placed, or by radiation. A vacuum is usually applied to remove the water vapor, and to decrease necessary temperatures of vaporization. For instance, indirect vacuum drying would be

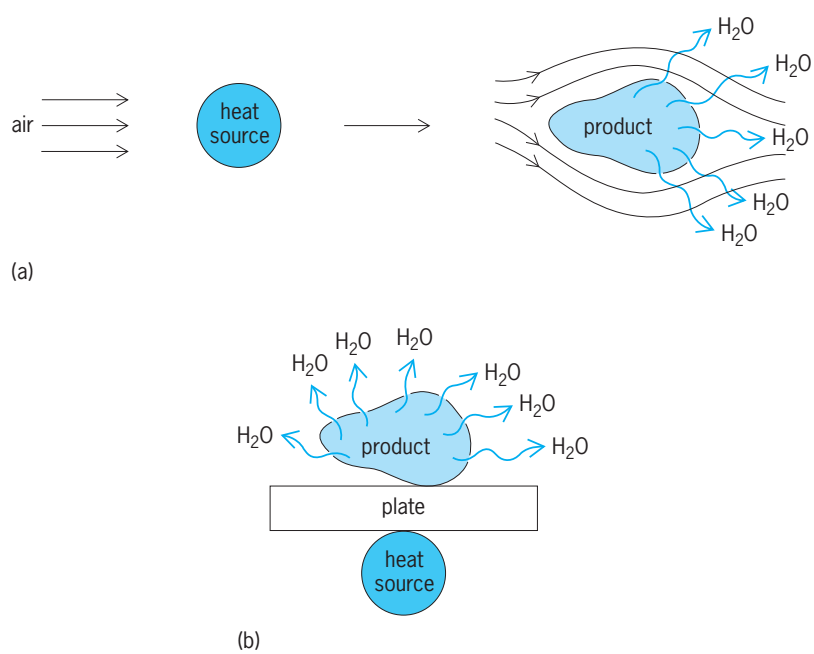

Fig. 27. Heat transfer to product by (*a*) direct and (*b*) indirect drying.

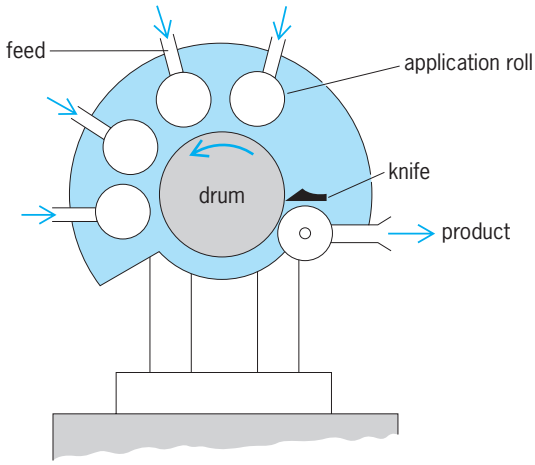

Fig. 28. Elements of a single-drum dryer.

used where a product is too sensitive to withstand exposure to a moving hot air stream.

Equipment and methods. A variety of equipment and methods is used in the drying process. The dryer chosen is dependent on product characteristics, desired quality retention, economic factors, and desired quality factors. Direct dryers include bed, belt, pneumatic, rotary, sheet, spray, through-circulation, tray, and tunnel. Indirect dryers include agitated-pan, cylinder, drum, freeze, microwave, screw-conveyor, steam-tube, vacuum, and vibrating-tray. Drum drying and spray drying are discussed below as important examples of typical drying methods.

Drum drying. Drum dryers utilize heat from the condensation of steam vapor on the inside of a drum surface to dry material that is contacted to the outside surface. Heat transfer is by conduction, and provides the necessary latent heat of vaporization. The drum rotates, and a thin layer of product is applied. Drying rates depend on film thickness, drum speed, drum temperature, feed temperature, ambient conditions, and product characteristics. The speed of rotation is adjusted so that the desired moisture content of the scraped product is obtained. Quality reduction is minimized by assuring that the product film has uniform thickness and that the dry film is removed completely. Drum dryers are very efficient, with heat loss typically 10% of the total input. However, drum dryers are not used extensively in the drying of food products due to the difficulty of applying a thin layer to the drum, and due to the long contact time at high temperatures, which may cause undesirable quality changes to the food product. Little design information exists on the subject. Most of the theoretical approaches account only for the heat transfer limitation when modeling dryer performance, when in theory drum drying is mass-transfer-limited.

The most important advantage of drum dryers is their capability of handling slurries and pastes of high moisture and viscosity. Products like mashed potatoes and powdered milk are typical applications. The ability of drum dryers to handle these difficult products makes them nearly irreplaceable in some operations.

Single-drum (**Fig. 28**) and twin-drum dryers are the most common types. Vacuum drum dryers have a vacuum tank which surrounds the drying surfaces. This feature decreases the temperature of operation.

Spray drying. Spray drying is a method of producing a dried powder out of feed in the form of pastes or slurries, or other liquids containing dissolved solids. Spray drying can be considered a two-stage process consisting of atomization and moisture removal. Atomization breaks the feed into a spray of liquid droplets. This provides a large surface area for moisture evaporation. The spray is contacted with a hot gas, usually air, and evaporation and drying produce the desired solid product particles.

Spray drying is an important preservation method in the food industry, since it removes moisture quickly and continuously without causing heat degradation. Heat damage to the food is prevented, due in part to the rapid drying cycle, the short total residence time, and cooling effects of evaporation. The disadvantages common to spray drying are the high energy needed and the requirement that the product be capable of being atomized. Furthermore, the product should not adhere to the interior of the dryer.

Spray drying is classified by the direction of product flow. There are cocurrent, countercurrent, mixed-flow, and parallel-type spray dryers. Cocurrent flow is best suited for heat-sensitive products.

In a cocurrent open cycle (**Fig. 29**) feed is pumped from the feed tank to an atomizer, where it is sprayed as tiny droplets inside the drying chamber. Air from the ambient surroundings is drawn in by a fan, filtered, and heated in an air heater. The hot air is contacted with the atomized food in the vicinity of the atomizer. Transport through the drying zone takes a matter of seconds. The powder falls to the bottom of the chamber, where it is pneumatically transported to the packaging line. The remaining product is separated in a cyclone, or a series of cyclones. The fines are recovered and transported to the pneumatic conveying system, while the air is vented to a filter and then to the atmosphere.

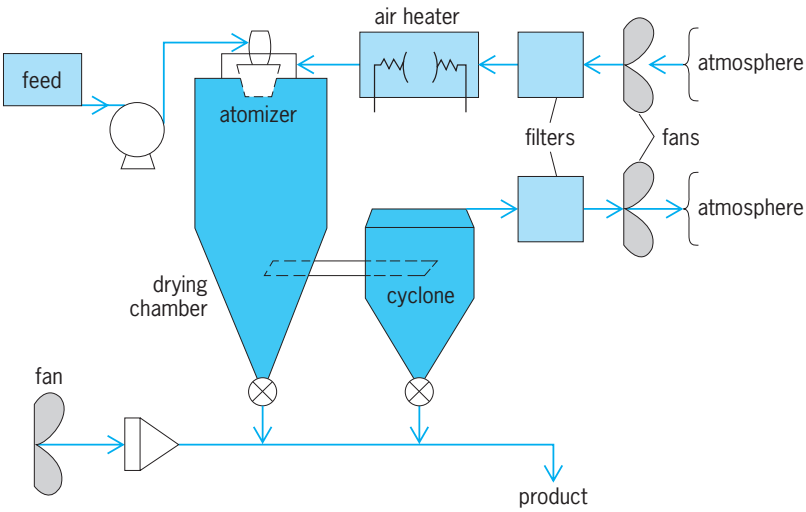

Fig. 29. Cocurrent open-cycle spray dryer layout.

There are many variations in spray dryer layout and design, depending on the requirements of the dried product. Additional equipment can be incorporated in the layout, and include wet scrubbers, fluidized-bed filters, and washing equipment. Variations in layout include open cycles as opposed to closed and mixed flow, and multiple-stage layouts that employ more than one spray dryer. A spray dryer may also use a direct-fired burner, utilizing the combustion of natural gas. Modifications of the spray dryer include a variety of atomizers (centrifugal pressure and fan-spray nozzles, two-fluid and rotary atomizers), and different shapes and sizes of the drying chamber. *See* DRYING.

Martin R. Okos

Expression

Expression is a process in which compaction caused by the action of a moving surface expels fluid from partially confined fluid-rich solids and solid-liquid mixtures, which are known as the press cake or cake. The process is used to recover oils and fats from oilseeds and rendered meat and fish scraps; reduce oil and water contents in fish, meat, bone, and poultry meals; dewater felts and mats formed from fibrous slurries; expel juice from fruits, sugarcane, and sorghum; expel pectin liquor from cooked pomace; dewater wet processing wastes; recover usable juices from such wastes; recover leaf protein; expel water or solvents from filter cakes; adjust the fat content of cocoa; dewater waste-treatment sludges; and expel whey from cheese.

The media. While most of the cake is confined, part is covered by a perforated surface, the media, or in special cases not covered at all. During compaction, expelled fluid flows out through the media or through the uncovered surface. The media often consists of light or heavy sheet metal with a uniformly spaced array of small holes, sometimes backed up by a fine-mesh screen. In some cases, a filter cloth supported by corrugated plates or grids is used as the media. When very high pressures are used, a cage made of sturdy, wedge-shaped bars separated by thin slits and backed by heavy ribs is used.

Equipment. Expression equipment includes screw presses (**Fig. 30**), in which compaction is induced by progressive reductions in the clearance between the root of the screw and the cage in which the screw is housed, or by reductions in the distance between screw flights or the clearance between adjacent twin screws. Compaction may also be induced by pressure developed by throttling the solid discharge at the end of the screw.

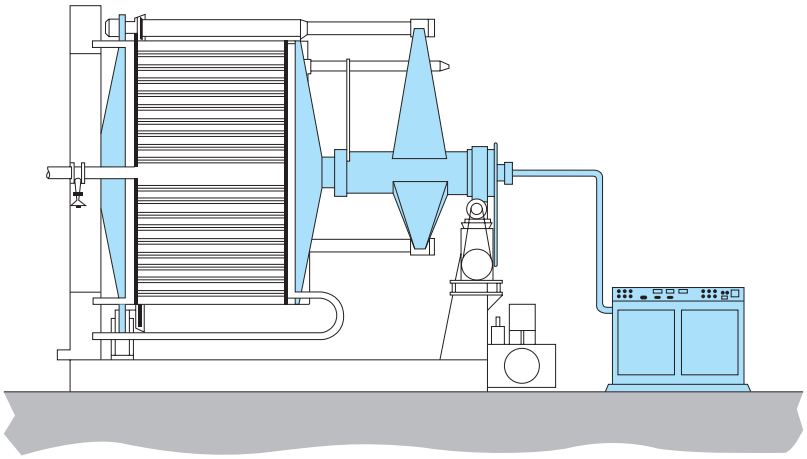

Fig. 31. Bucher-Guyer multitube press.

Cakes are also compacted inside perforated housings by solid or perforated rams or, in one case, by an inflatable rubber tube. In one widely used press, comminuted fruit is pumped into a cylinder that contains many cloth-covered, ribbed, flexible tubes attached to manifolds in the ram surface and the opposing stationary surface (**Fig. 31**). These tubes, which support the cloth media, convolute as the ram advances. Hence, the outflow area remains constant.

In rack-and-cloth presses, the cake is confined in flat pouches formed by folding over the corners of square filter cloths. These pouches are sandwiched between racks that provide mechanical support and clearance for juice drainage. Piles of the racks and pouches are compacted between platens driven together by screws or hydraulic pressure.

In belt presses, cake is deposited on a perforated moving belt and covered by a similar belt moving at the same speed. The belts are pressed together by passing them between a succession of rollers, or by placing them in tension and passing them over a rotating drum or drums. Juice yields and dewatering are improved by pulling belts over drums that provide S-like turns in the pressing path.

Sugarcane and sorghum are pressed by passing them through the nip between co-rotating, grooved rolls in three roll mills. The rolls are pressed together by hydraulic pistons. The expelled juice drains backward through solids advancing into the nip. In lighter

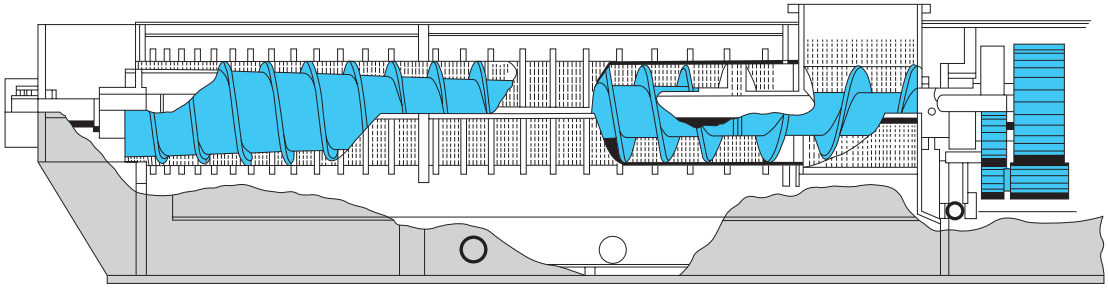

Fig. 30. Medium-weight screw press.

presses, which are used to expel juice from forage crops and dewater spent brewing grains, the solids pass through the nip between two co-rotating conical disks.

Solid-bowl, scroll centrifuges are sometimes used to press first-run grape juice from crushed, destemmed wine grapes and are often used to press olive oil and fruit water from milled olive paste. The scroll conveys solids toward the narrow, solids-discharge end of the centrifuge. Centrifugal force compacts the solids, forcing out fluid, which forms a pool over the compacted solids. The fluid discharges through holes at the wide end of the centrifuge. The process is very rapid, but compaction is not stringent. Consequently, yields tend to be low. Water is mixed with olive-paste feeds to improve yield by augmenting the volume of expellable fluid. This increases water-removal loads when expelled olive oil is subsequently separated from expelled fruit water.

Pressure and pressing rate. Pressing productivity and juice yield are governed by the relationships between the pressure used to achieve compaction, the specific volume of the presscake, and the rate of pressing. The pressure is the sum of (1) the stress developed in the solid itself or in the solid occluded-fluid ensemble; (2) compaction stress that is transferred from the presscake to adjacent surfaces by friction; and (3) the fluid pressure drop generated by fluid flow through the media and interparticle pores in the cake. The solid compaction stress depends mainly on the extent of compaction, and often increases exponentially as the cake volume decreases. It also increases, but to a lesser extent, as the rate of compaction increases. Friction reduces available compaction stress at a rate which is proportional to the compaction stress itself and the sliding-surface area.

Fluid pressure drop is proportional to the rate of pressing, the fluid viscosity, and the sum of the flow resistances of the media and cake. These resistances increase rapidly as cake volume decreases. Therefore, fluid pressure drop is often the largest compaction-pressure component at the end of pressing.

Methods have been developed for (1) measuring and correlating the forces and flow resistances encountered during expression; (2) predicting interactions between these forces and resistances and the cake specific volume; and (3) predicting juice yields as a function of pressing rate and maximum allowable pressure.

The allowable rate of pressing and the maximum allowable pressure are in turn determined by either strength limits of the expression equipment; power limits for the equipment drive train; pressures at which cake extrudes through media; pressures at which undesirable particle deformation occurs; desired extents of fluid removal; or desired uniformity of cake compaction and fluid content. Depending on which limiting factor applies, the maximum pressures used range from 2 to 5 lb/in.^{2} (14 to 35 kilopascals) for pressing soft cheeses, 6 to 12 lb/in.^{2} (41 to 83 kPa) for hard cheeses, 200 to 450 lb/in.^{2} (1.4 to 3.1 megapascals) for fruit juice expression, 500 to 1000 lb/in.^{2} (3.5 to 6.9 MPa) for dewatering spent sugarbeets, 1000 to 2000 lb/in.^{2} (6.9 to 14 MPa) for dewatering spent coffee grounds, and attain 6000 lb/in.^{2} (42 MPa) in sugarcane mills, and as much as 16,000 lb/in.^{2} (110 MPa) in some vegetable-oil screw presses.

Henry G. Schwartzberg

Filtration

Filtration processes are used for the removal of solids from a fluid (either a liquid or gas) by utilizing a physical barrier (the filter medium) containing openings or pores of the appropriate shape and size. The major force required is hydraulic pressure to pump the fluid through the filter medium. In the other mechanical separation processes, such as sedimentation and centrifugation, the force required is gravity or centrifugal force, respectively. In addition, they also require a significant difference in density between the two phases being separated, whereas filtration requires only a significant size difference.

Filtration processes cover a wide range of particle sizes (**Table 4**). The sieving or screening operations use fairly coarse barriers of perforated metal or woven screens. The screens may be vibrated or rotated to improve filtration rates (**Fig. 32**). On the other end of the spectrum is reverse osmosis (or hyperfiltration), a membrane separation process designed to filter out dissolved molecules in the ionic-size range. The middle range of solid-fluid separations involves removal of insoluble particles of 0.05 micrometer and larger, and is commonly referred to as filtration or microfiltration.

TABLE 4. Classification of filtration operations

Process	Particle size	Example
Screening, sieving	>100 μm	Protein precipitates (casein, curd particles)
Filtration	10–100 μm	Yeast, seedcoat in oilseed slurries
Microfiltration	0.05–10 μm	Bacteria, blood cells
Ultrafiltration	1–20 nm	Protein, soluble starch
Reverse osmosis	<1 nm	Salt, sugars

Equipment. Equipment for solid-liquid filtration can be classified in a number of ways: (1) whether the retained particles (the cake) or the filtrate (the clarified liquid) is the desired product; (2) batch or continuous operation; (3) driving force, that is, gravity-, pressure, or vacuum-activated; (4) mechanical arrangement of the filter medium, such as

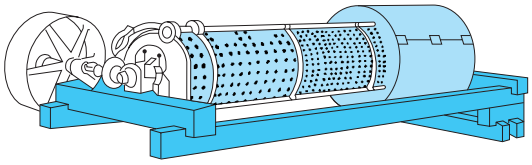

Fig. 32. Revolving screen filter. (*After W. L. McCabe and J. C. Smith, Unit Operations of Chemical Engineering, chap. 7, McGraw-Hill, 1956*)

plate-and-frame (**Fig. 33**), leaf, rotary drum (**Fig. 34**), or cartridge type; or (5) mechanism of filtration, that is, whether the separation occurs on the surface of the filter medium (as in membrane separations) or within the depths of the filter. Depth filters are usually used in dead-end-type equipment and in batch operations. An example is a clarification process in which the slurry has a low particle concentration, such as in the production of sparkling clear beverages. Surface filters are more suited for continuous operations where the retained particles have to be quantitatively recovered, such as the harvesting of microbial cells. *See* CLARIFICATION.

Filter media commonly used in industrial filtration include duckweave heavy cloth, paper, cellulose, metal cloth, nylon, and other synthetic cloths. Filter aids, such as diatomaceous earth or activated carbon, are also frequently used to improve filtration efficiency. Filter aids are applied as a precoat on the filter medium or added to the slurry before filtration. *See* ACTIVATED CARBON.

Theory. All models describing the filtration process must account for the buildup of solid particles on the filter medium and the resulting drop in filtration rate due to the increased resistance of the cake. Reasonably good mathematical approximations have been developed based on Poiseuille's law for laminar flow through channels, shown in differential form in Eq. (1), where V is the volume of filtrate in time t,

$$\frac{dV}{dt} = \frac{A^2P}{\mu\alpha CVP^s + AR_m} \qquad (1)$$

A is the filter area, P is the pressure drop across the filter, C is the solids concentration in the slurry, μ is the fluid viscosity, α is the specific cake resistance, and s is a cake compressibility coefficient. The resistance due to the filter medium, R_m, is usually much less than the cake resistance and is usually neglected (that is, $R_m = 0$). *See* LAMINAR FLOW.

Filtration is conducted in one of two modes: constant-rate or constant-pressure. Most satisfactory results are obtained when a constant-rate operation at low pressure is used at the start of the process. Due to cake buildup, however, a gradual increase in pressure will be needed to maintain constant-rate conditions. When the pressure limit of the filter or the pump is reached, the operation is conducted under constant-pressure conditions until the filtration rate drops to an uneconomically low level. The process is then stopped to wash or remove the cake and clean the filter. The equation can be integrated with the appropriate boundary conditions to describe either the constant-rate or constant-pressure operation.

Optimizing filtration variables. The approximations based on Poiseuille's law in both its differential and integrated forms are useful in evaluating and optimizing operating variables, such as pressure, temperature, solids concentration, filtration rate, and filter area. Many of the variables involved are not mutually exclusive. It is almost always necessary to conduct experimental trials for every application in order to properly design a filtration operation.

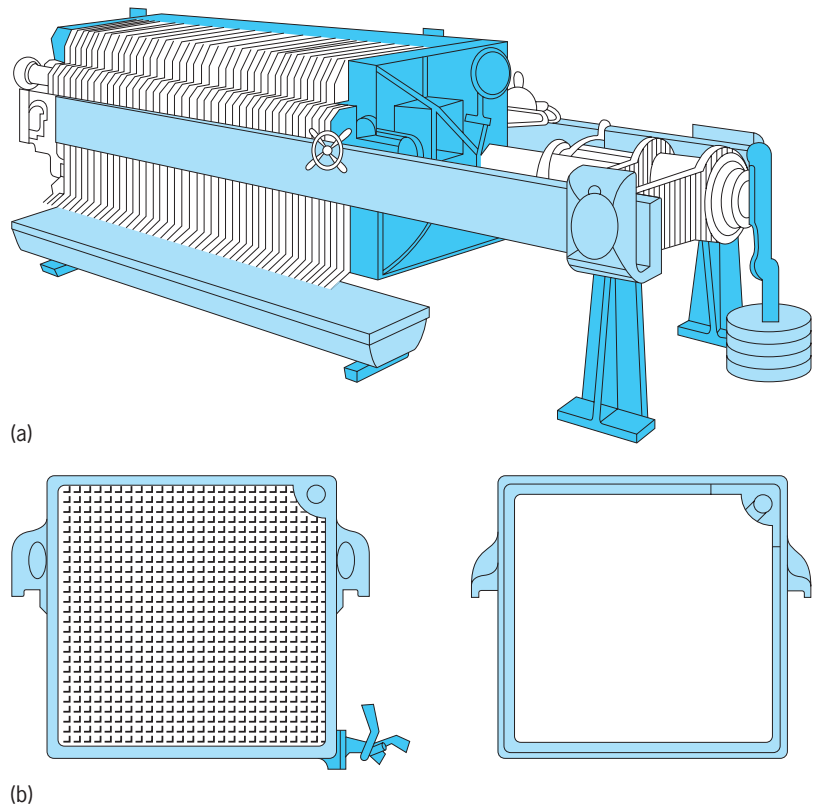

Fig. 33. Plate-and-frame filter press. (*a*) Outside view of press. (*b*) Plate and frame of open-delivery, nonwashing press. (*After W. L. McCabe and J. C. Smith, Unit Operations of Chemical Engineering, chap. 7, McGraw-Hill, 1956*)

Increasing pressure should increase filtration rate. This will be true for hard, granular cakes that are relatively incompressible (where $s = 0$ in the equation). On the other hand, if the cake is soft, gellike, and easily compressed (that is, where $s = 1$), the filtration rate will be independent of pressure, assuming $R_m = 0$. In practice, s has been found to vary between 0.1 and 0.8, and will have to be experimentally determined for each application.

Higher temperatures increase filtration rates, due primarily to its effect on lowering fluid viscosity.

Filtration rate can be increased by diluting the feed

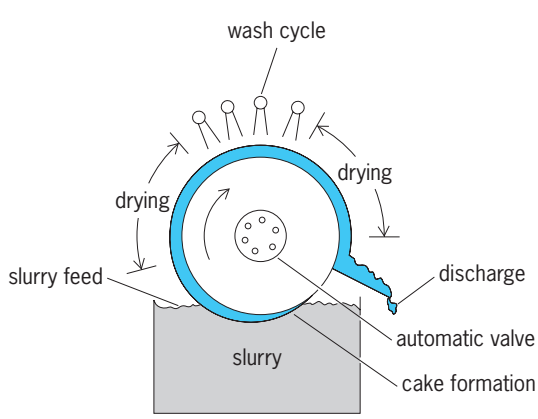

Fig. 34. Continuous rotary-drum filter. (*After C. J. Geankoplis, Transport Processes and Unit Operations, Allyn and Bacon, 1978*)

slurry. However, this must be balanced against the higher volume that must be processed.

The equation shows that the total volume filtered (V) is inversely proportional to the filtration rate (dV/dt). Higher volumes can be processed at lower filtration rates.

The filtration rate and the total volume filtered are proportional to the square of the filter area. Thus increasing filtration area may be the most cost-effective method of increasing filter throughout.

Membrane separations. Since the birth of modern membrane separations technology in 1960, a vast array of applications unmatched by any other processing technique in its variety have been developed. Desalination and water treatment by reverse osmosis is probably the earliest and best-known application. The 1970s saw increasing usage in the chemical process industries (paint, textiles, oil recovery, pulp, and paper). In the 1990s the most significant inroads of this technology were in the food and biotechnology processing industries, where membrane technology is finding increasing uses as a gentle and efficient way of fractionating, concentrating, and clarifying a variety of products.

Membrane separation processes are based on the ability of semipermeable membranes of the appropriate physical and chemical nature to discriminate between molecules primarily on the basis of size and, to a certain extent, on shape and chemical composition. The role of membrane is to act as a selective barrier, enriching certain components in a feed stream and depleting it of others. Membranes are made from a variety of polymers or inorganic materials; over 150 different materials have been used, with varying degrees of success. The food industry, however, uses only a few materials, primarily those that have met sanitation and regulatory standards. These include the cellulosics, ceramics, some polyamides,

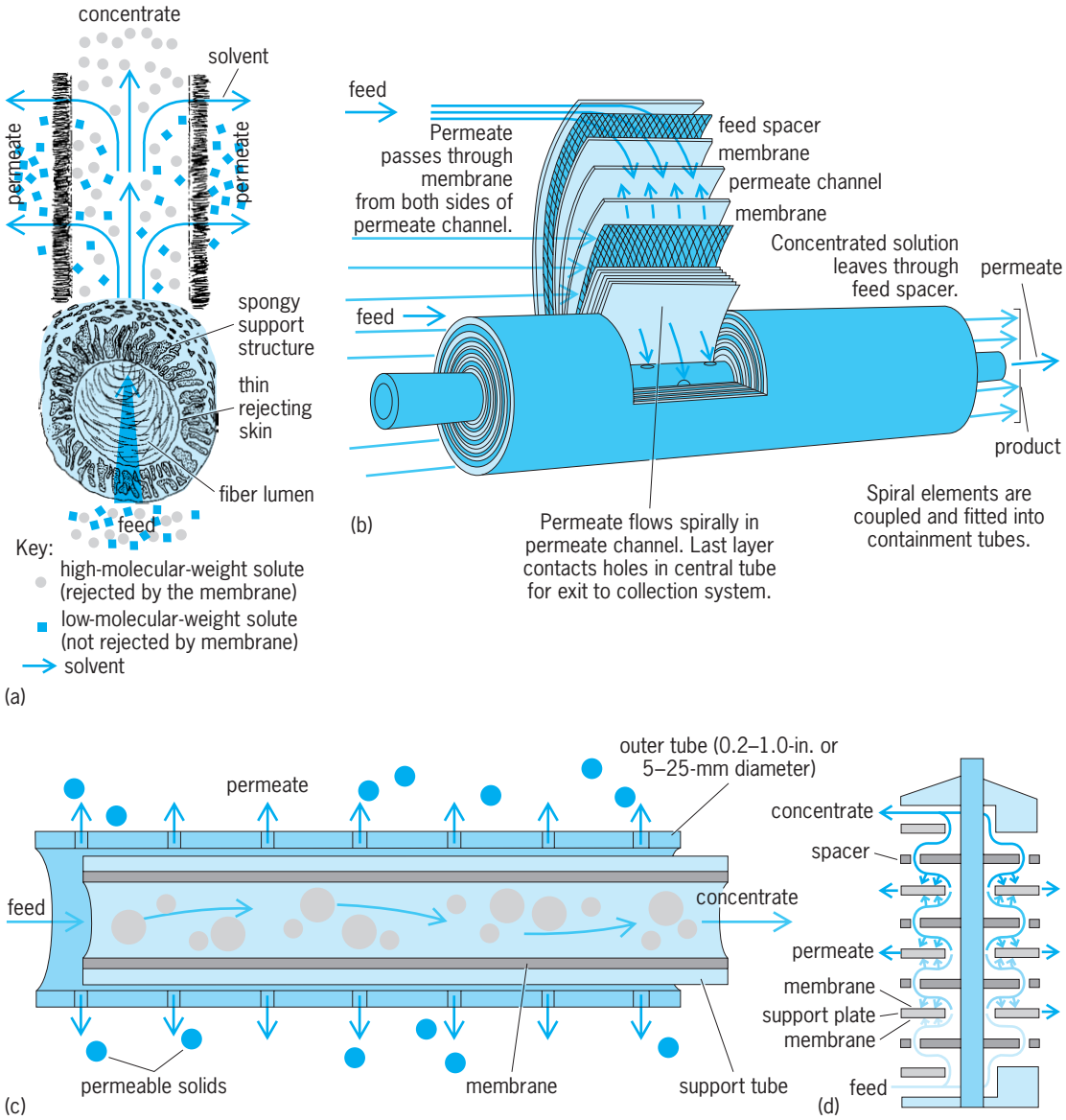

Fig. 35. Modules available for membrane filtration. (*a*) Hollow fibers. (*b*) Spiral-wound. (*c*) Tubular. (*d*) Plate system, showing internal flow.

polysulfone and its derivatives, polyvinylidene fluoride, and propylene. The chemical nature and physical properties of the membrane control which components are retained and which permeate through the membrane. For example, it is possible to concentrate or dewater by reverse osmosis, fractionate components in solution by ultrafiltration, and clarify slurries or remove suspended matter by microfiltration. Nanofiltration fits between reverse osmosis and ultrafiltration; nanofiltration membranes are designed to separate monovalent salts and undissociated organic acids from multivalent salts, sugars, and organic compounds. *See* CELLULOSE; CERAMICS; POLYMER.

One of the chief attractions of membrane technology is the low energy requirements compared to other dewatering processes. For example, open-pan evaporation may need over 0.27 kWh per pound of water removed (0.6 kWh/kg), multiple-effect or mechanical vapor recompression evaporation uses 0.020 kW/lb (0.045 kWh/kg), and reverse osmosis requires only 0.0023–0.0091 kWh/lb (0.005–0.02 kWh/kg).

There are several equipment types and module designs that, between them, will probably be able to handle any fluid stream that can be pumped (**Fig. 35**). The principal types are tubular, hollow-fiber, spiral-wound, and plate systems. Each design has its own special applications, advantages, and disadvantages (**Table 5**).

In selecting a particular membrane module, the major criteria are (1) the physical properties of the feed stream and retentate, especially viscosity; (2) particle size of suspended matter in the feed; and (3) fouling potential of the feed stream. For example, most food materials containing macromolecules display non-newtonian behavior. Their viscosity will increase dramatically above certain concentrations, making pumping difficult and reducing the mass transfer within the boundary layer. This will necessitate the use of modules that can withstand high pressure drops, eliminating most hollow-fiber and capillary modules. However, these modules have extremely high packing densities (surface area:volume ratios) and comparatively low energy consumption.

Similarly, large-bore tubular modules are more suitable for feed streams with large amounts of suspended matter such as citrus juices and animal waste streams, even though they have the lowest packing densities and highest energy consumption among all designs. Modules utilizing flat sheets (spiral-wound units, plate units, and pleated-sheet cartridges) usually have a meshlike spacer between sheets of membrane. This restricts their use to clear liquids or feed streams containing only finely suspended matter. However, spiral-would units are the lowest in capital costs and lowest in energy consumption, although this is somewhat offset by their generally lower flux with polarizing feed streams.

If there has been one single factor that has inhibited the large-scale use of the membrane technology, it is fouling of membranes. Fouling, that is, accumulation of deposits, manifests itself as a decline in performance, usually a decline in flux under constant operating conditions and a possible change in the sieving properties of the membrane. This phenomenon is frequently confused with other flux-depressing phenomena, such as change in membrane properties with time due to deterioration, changes in feed solution properties, and concentration polarization.

Minimizing membrane fouling must include consideration of the chemical nature of the membrane as well as physicochemical properties of the feed stream. For example, hydrophilic membranes (such as the cellulosics and ceramics) foul much less with protein-containing feed streams than hydrophobic membranes (such as polysulfone or polypropylene). With fermentation broths, although microbial cells may be the largest single component, it is rarely the cells themselves that constitute the main cause of fouling, but the media components, except perhaps at very high cell concentrations (>100 oz/ft^3 or 100 g/liter).

The parameter known as the cloud point is the temperature at which a dissolved solid begins to separate from a solution, imparting a cloudy appearance to the liquid. Antifoams, especially silicone-based, are particularly notorious foulants, while vegetable oils are the most benign. Antifoams with inverted cloud points (that is, cases where higher temperatures increase cloud formation) can cause dramatic reduction in flux, especially with hydrophobic membranes such as polysulfone, if operated above their cloud points.

Fouling can be minimized by (1) modifying the membrane chemically, for example, to make it more hydrophilic or to block certain functional groups that may be interacting with specific feed components; (2) by modifying the feed stream, for example, adjusting pH changes the charge or conformation of

TABLE 5. Features of membrane equipment

Features	Tubular	Hollow fiber	Plate	Spiral-wound
Channel height	>10 mm*	<3 mm*	<2.5 mm*	<1.5 mm*
Spacers	No	No	Yes/no	Yes
Prefiltration	No	Yes	Yes/no	Yes
Flow	Turbulent	Laminar	Laminar	Turbulent
Flux	High	Intermediate	Intermediate	Low
Area/volume ratio	Low	High	Intermediate	High
Energy/flux ratio	High	Intermediate	Intermediate	Low

*1 mm = 0.04 in.

a protein, which may minimize protein-membrane interactions, or it changes the solubility of salts so that they do not interact with or precipitate on the membrane; and (3) by physical or mechanical means, which may involve periodic backwashing, backpulsing, vibrating or rotating or spinning the module, pulsatile flows, maintaining a uniform transmembrane pressure with cocurrent permeate flow in the module, or electrically enhancing the removal of the polarized layer from the membrane.

On a worldwide basis, the dairy industry is probably the largest user of membrane technology in food processing, followed by the fruit juice processors. Ultrafiltration is ideal for clarification of clear juices, where the goal is to replace the conventional holding, filtration, and decanting steps, and perhaps the final pasteurization step. Higher yields; improved juice quality due to removal of pectins, polyphenol oxidase, and tannin-protein complexes; and elimination of filter-aiding and fining agents are its features. Concentration by reverse osmosis to high levels ($>42^\circ$ Brix; a concentration measurement) has been commercialized but is not widespread.

Other food applications include production of vegetable protein concentrates and isolates from soy, sunflower seeds, cottonseeds, peanuts, and so forth; removal of glucose from and partial concentration of eggs prior to dehydration; processing of animal products such as blood and gelatin; sugar refining, including desalting of molasses, purification, and preconcentration; removal of yeast from beer or alcohol distilleries by microfiltration or ultrafiltration instead of centrifuges or filters; corn refining, including clarification of dextrose and starch purification and concentration; and vegetable oil refining, primarily for degumming, hexane recovery, and deacidification.

Membrane technology possesses the necessary characteristics to meet the rigorous demands of biotechnology. The major applications are separation and harvesting of enzymes and microbial cells; continuous high-performance bioreactors for enzymatic and microbial conversion processes; tissue culture systems utilizing plant and animal cells as biocatalysts or for production of monoclonal antibodies and production of high-purity water. *See* BIOTECHNOLOGY; CHEESE; FILTRATION; MILK. Munir Cheryan

Mixing and Agitation

Mixing is carried out to provide homogeneous compositions and physical properties for food ingredients and combinations of ingredients. Mixing and agitation are also used to facilitate heat transfer and mass transfer, create dispersions, produce emulsions and foams, suspend solids, facilitate reactions, promote coalescence, and produce textural and structural changes.

Homogeneity can be characterized in terms of the smallness of local deviations from the average value of properties or component concentrations, and in terms of the domain size in which such local deviations are less than some maximum tolerable value. With liquid mixtures, homogeneity can usually be obtained over domains which involve tens of thousands of molecules; for solid mixtures, homogeneity can be provided only over domains which involve thousands or tens of thousands of particles (10^{20}–10^{21} molecules). Artificially induced circulation or agitation is frequently used for carrying out mixing.

Many different mixing and agitation methods are used in order to produce the liquid, solid, dough-like, and foamy mixtures and dispersions which are encountered in food processing. In general, these methods involve repeatedly subdividing the mixture into discrete domains which are then transported relative to each other so that the domains repeatedly encounter and exchange matter with new neighboring domains.

Liquids. Liquid mixing is usually carried out by using rapidly turning screw propellers or radial-flow turbines to circulate batches of liquid in baffled tanks (**Fig. 36**). The baffles, which are usually one-twelfth or one-tenth of the tank diameter wide, prevent the liquid from moving en masse in a vortex. Off-center and side-entering impellers can also be used to prevent vortex formation. Propellers and turbines with pitched blades (**Fig. 37***a* and *b*) provide fairly smooth axial flows which can be advantageously used for particle suspension, blending nonviscous liquids, and improving heat transfer between such liquids and heating jackets or coils. Straight-blade

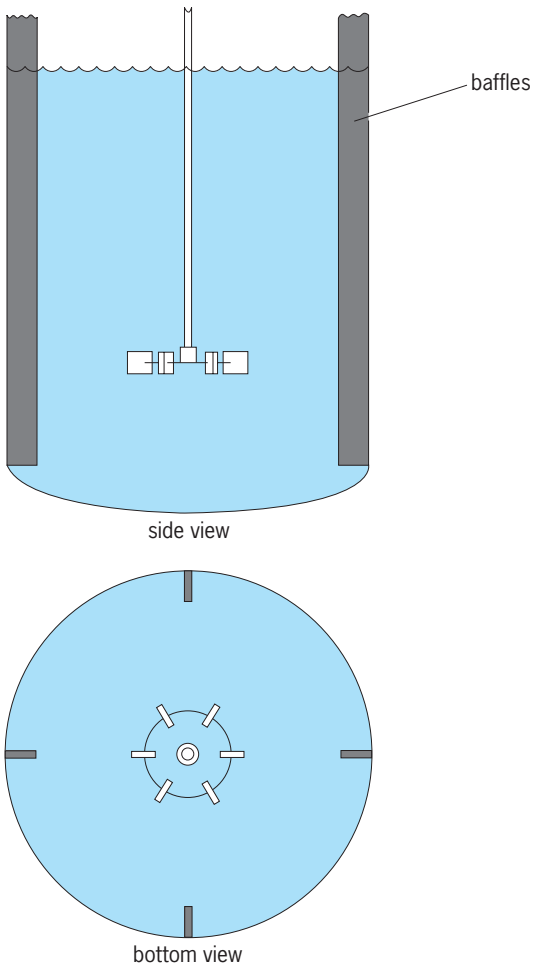

Fig. 36. Structure of a baffled mixing tank.

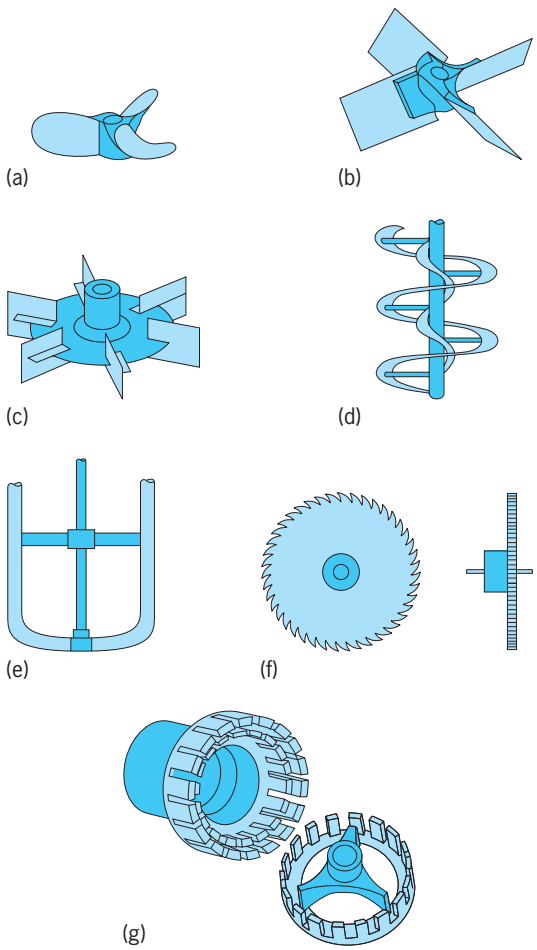

Fig. 37. Types of agitators. (*a*) Pitched-blade propeller. (*b*) Pitched-blade turbine. (*c*) Straight-blade turbine. (*d*) Double-helical impeller. (*e*) Anchor-blade impeller. (*f*) Sawlike impeller. (*g*) Slotted cagelike impeller.

turbines (Fig. 37*c*) are used to produce highly turbulent radial flows which are advantageous for producing dispersions and for facilitating mass transfer and heat transfer between suspended particles, drops, and bubbles and the liquid in which they are dispersed. Large-diameter, slow-speed, helical and double-helical impellers (Fig. 37*d*) and anchor-blade impellers (Fig. 37*e*) are used for mixing and circulating highly viscous and non-newtonian liquids. Sawlike impellers (Fig. 37*f*), slotted, cagelike impellers (Fig. 37*g*), and double counterrotating slotted cages are used to provide very high local shear rates and to create final dispersions. Various combinations of large-diameter, slow-speed helical mixers and high-speed, high-shear mixers are used to disperse and dissolve solutes which form viscous, non-newtonian solutions. *See* NON-NEWTONIAN FLUID.

The power P expended by impellers during turbulent vortex-free agitation of newtonian fluids is usually proportional to the value of the expression $C\rho N^3D^5$, where ρ is the liquid density, N is the rotational speed, and D is the impeller diameter. The constant C depends on the impeller type and tank geometry; it equals 1.0 for marine propellers and 6.0 for radial-flow turbines. The volumetric flow rate generated by impellers is proportional to ND_i^3. Heat-transfer, mass-transfer, and power-expenditure correlations for stirred tanks are frequently expressed in terms of Re_i, the impeller Reynolds number: $\mathrm{Re}_i = \rho ND^2\mu$, where μ is the viscosity. Heat-transfer coefficients for immersed surfaces in agitated tanks are usually roughly proportional to $\mathrm{Re}_i^{2/3}$. For propellers and pitched-blade turbines, fully turbulent flow occurs when Re_i is greater than 2000; for radial-flow turbines when Re_i is greater than 10,000; and for anchor-blade impellers when Re_i is greater than 100,000. Viscous flow occurs when $\mathrm{Re}_i < 10$. In the viscous range, power expenditures are proportional to μN^2D^3. *See* REYNOLDS NUMBER.

In the turbulent range, local mean velocities and fluctuating velocities in stirred tanks are proportional to $(PD/V\rho)^{1/3}$, where V is the tank volume. Mass-transfer coefficients for suspended particles, droplets, and bubbles, and the sizes of dispersed drops and bubbles are often correlated in terms of Reynolds numbers and Weber numbers based on the turbulent fluctuating velocity. Therefore, the ratio of power to volume (P/V) is often an important parameter in liquid mixing operations.

Recommended P/V values in horsepower per 1000 gallons (1 hp/1000 gallons = 0.2 kW/m^3) for typical liquid-food mixing operations are 0.25–0.33 for liquid blending; 0.12–0.25 for maintaining emulsion uniformity (for example, for preventing cream from rising in milk); 0.75–1.5 for improving heat transfer from fluids to immersed surfaces; 1.0–1.5 for dispersing starter cultures for cheese and yogurt manufacture, mixing ice cream ingredients, and dispersing and dissolving readily soluble solids; 4–10 for suspending rapidly settling particles (for example, hydrogenation catalysts), dissolving or dispersing sparingly soluble solids, and improving mass transfer between suspended particles and surrounding liquid; and 10–12 for dispersing air during aerobic fermentations. Power expenditures for suspending solids and blending liquids can be minimized by using large-diameter, slowly rotating propellers; but impeller and drive-train costs will be large when large-diameter propellers and slow rotational speeds are used. Therefore, small propellers and high rotational speeds are often used for small-scale or intermittent mixing.

Liquid mixing is usually carried out in tanks; in-line mixers, that is, agitators mounted in pipes, are also used. In-line mixers can be positioned to provide controlled holdup between mixing and subsequent processing, for example, between admixing a flocculant and separating resulting flocs. In-line mixers are also used immediately before equipment that is processing feeds that tend to separate, for example, to mix oil and aqueous liquid before entering a homogenizer. Static elements mounted in series in tubes can be used to provide in-line mixing. Elements split streams flowing through them and shift the positions of the splits. Adjacent elements create different splits and shift them in different senses. A series of staggered helical elements are often used. Successive elements turn in opposite directions and split the

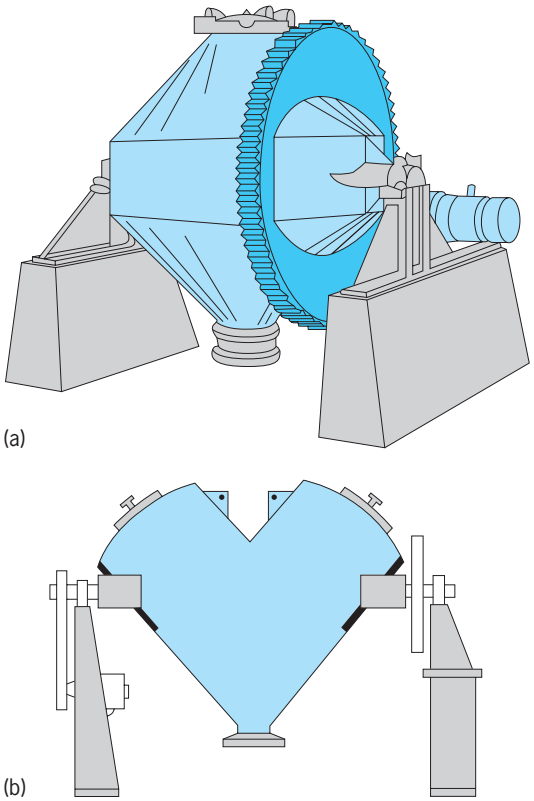

Fig. 38. Dry-powder mixers. (*a*) Double-cone blender. (*b*) Twin-shell (vee) blender.

stream 90° to the preceding split (**Fig. 38**). Mixtures split 2^n times in passing through n elements. Different flow-splitting arrangements are used in other static mixers. Static mixers are used to blend materials with widely differing viscosities, for example, frozen orange juice concentrate and water; blend shear-sensitive ingredients, for example, to mix blueberries with yogurt; carry out limited mixing, for example, to marbellize ice cream; and facilitate in-line heat exchange with viscous materials, for example, peanut butter, chocolate, and edible oils.

Solids. The mixing of solid powders is usually carried out in equipment which causes the solid mixture to repeatedly split into two or more portions, which are then recombined. In tumbling-type mixers (Fig. 38), mixing is brought about by lifting the mixture and letting it split as it falls in different increments as the mixer turns. Conveyors are used to transport part of the mixture relative to adjacent parts in ribbon blenders (where oppositely pitched, concentric pairs of ribbonlike, horizontal helices are used and in mixers where tilted, vertical screw conveyors which orbit around a conical hopper are used. Continuous, in-line mixing of solids can be provided by using twin-shell (vee) blenders mounted in series on an inclined shaft; screw conveyors with cut, cut-and-folded, or ribbonlike flights; or static mixers mounted in vertical ducts. Feed streams have to be automatically proportioned to obtain consistent uniform mixes with in-line units.

Solid mixing is characterized by diffusive processes, which tend to cause uniform spreading of the components of a mixture, and segregation processes, which tend to cause selective migration of certain components. Segregation is enhanced when the particle sizes and, to a lesser extent, the densities of components of mixtures differ. Small particles will frequently percolate through a bed of mixed particles and accumulate near the bottom of the bed, or in ribbon blenders they may be thrown into the air and temporarily adhere to the cover of the blender. In the latter case, erratic variations in mixing uniformity will occur.

Solid mixing usually results in random composition variations in local sample domains. The theoretical minimum standard deviation for the composition of a sample will be inversely proportional to the square root of the number of particles in the sample. Mixing indices based on the ratio of the theoretical minimum standard deviation to experimentally measured standard deviations are sometimes used to characterize the thoroughness of mixing and to correlate mixing rates. Maximum values for such mixing indices will frequently be in the 0.6–0.7 range—that is, substantially less than 1.0, particularly when segregation tends to occur—and are reached only after mixing times of 30 min to 1 h. By contrast, for comparable batches of liquids, very uniform mixing is usually achieved in 5–15 min. If the fine particles in a solid mixture tend to adhere to the larger particles, a more uniform type of mixing known as ordered mixing may be obtained.

The mixing of pastes and doughs is carried out in units in which the mixture is sheared and stretched between sturdy, slowly rotating elements with small interelement clearances. This process stretches and folds portions of the mixture and forces these sheared portions into other portions of the mixture. Relatively long mixing times and large power expenditures are required when doughs and pastes are mixed. Screw-based extruders are sometimes used for mixing doughlike products. Forward flow and transverse flow are induced by the shearing motion of the screw relative to the barrel, and reverse flow is induced by the pressure built up across the die through which the extruded product discharges. This combination of flows produces very intense mixing in the extruder.

The thoroughness and completeness of dough mixing can be characterized in terms of mixing indices, which are similar to those used for mixing powders. The nature and functional character of dough components—for example, gluten in wheat—and the overall rheological nature of a mixture often greatly change during mixing. In chocolate conchers, small particles of sugar are progressively broken down into very fine particles and coated with semifluid cocoa butter to produce a very smooth mixture. Textural and functional effects produced during mixing of doughs and pastes cannot be adequately measured or characterized by uniformity-based mixing indices. Power expenditure and agitator torque rise as bread doughs are mixed; the dough is considered fully developed when the torque reaches a maximum, and degrades

with further mixing. The values of P/V for dough mixing are usually 33–134 hp/1000 gallons. *See* COCOA POWDER AND CHOCOLATE. Henry G. Schwartzberg

Extraction

Extraction is one of a family of separation processes employed for selective removal of components of a mixture by contacting it with an immiscible or partially miscible solvent to form a two-phase liquid-solid or liquid-liquid system (**Table 6**). Partitioning of solute molecules between the two phases provides the driving force for extraction. Based on the difference in either the solute's affinity for the solvent or its concentration in the two phases, the partitioning is conveniently expressed as a partition coefficient (K), defined as the ratio of a solute's concentration in the lighter phase to that in the heavier phase. Analytical data on the composition of immiscible phases in equilibrium can be used to determine K. The effectiveness of separating competing solutes is indicated by a separation factor (selectivity), which is the ratio of individual solute partition coefficients. Effective separation of one solute from another requires a separation factor significantly different from unity.

The most common solvents for extraction of food components are water, organic solvents such as hexane, and supercritical fluids. A fundamental understanding of rate-controlling steps governing the extraction of the solute of interest is needed in the design of an efficient process. Like other food processing unit operations, extraction can be carried out as a batch (unsteady-state) or a continuous (steady-state) operation. While a continuous system is more efficient in a large-scale operation, batch separation operations are common for biotechnologically produced products. In continuous, large-scale systems, the two phases flow either countercurrently or cocurrently, the former being more common. When recovery of the desired component is inadequate in one extraction, the process is repeated in a cascade of separation steps, and the process is known as staged extraction. The mixing and setting time needed in a stage extraction is eliminated by differential extraction in which the two phases continuously flow past each other without attaining equilibrium.

Liquid-liquid extraction. In this type of extraction, the solution containing a desired solute and the solvent, a liquid having preferential solubility for the solute, are mixed thoroughly. The solvent acts, and an immiscible layer, called the extract, is formed. The extract is separated from the remainder of the original solution, which, along with the residual solvent, is known as the raffinate. The solute in the extract may be purified in a secondary operation by using such techniques as membrane separation, electrophoresis, crystallization, precipitation, or distillation; and the solvent is reused after purification. Liquid-liquid extraction is ideal for recovering relatively nonvolatile and heat-labile solutes. Examples include vitamins, alkaloids, and pigments from animal or vegetable extracts. *See* CRYSTALLIZATION; DISTILLATION; ELECTROPHORESIS; MEMBRANE SEPARATIONS; PRECIPITATION (CHEMISTRY).

Solid-liquid extraction. Also known as leaching, solid-liquid extraction involves separation of soluble components from a solid by dissolving in a suitable liquid solvent. The size and shape of the solid matrix affects the performance of the extraction process. Increasing the surface-to-volume ratio through grinding, cutting, agitating, and increasing the temperature is effective in enhancing the rate of solute removal. If necessary, the solids may also be conditioned to release the desired component to the solvent phase. Often, solid-liquid extraction is aimed at recovering not one particular solute but all the soluble compounds in the solid matrix. For example, in the extraction of roasted coffee beans with water the extract contains all the water-soluble coffee constituents. *See* SOLVENT; SOLVENT EXTRACTION.

Supercritical fluid extraction (SFE). A supercritical fluid is any fluid at temperature and pressure above its critical values which, for example, for carbon dioxide are 31°C (88°F) and 72.9 atm (7.4 megapascals). In its supercritical state, a fluid exhibits physicochemical properties intermediate between those of liquids and gases, enhancing its role as a solvent. Its reversible high density gives good solubilizing power, while its low viscosity and diffusivity values provide appreciable penetrating and mixing power. These unusual properties have been found useful in extracting otherwise hard-to-separate constituents from their matrices.

Supercritical fluid extraction has emerged as a unit operation that exploits the unique properties of solvents to extract soluble components from a mixture. Partly because of the increasing regulation of solvents used in conventional extraction and partly

TABLE 6. Selected examples of extraction processes

Extraction	Material	Solvent	Solute
Liquid-liquid	Vegetable extract	Methylene chloride	Steroid
	Fermentation broth	Ethyl acetate	Antibiotic
	Cottonseed oil	Propane	Oleic acid
Liquid-solid	Ground roasted coffee	Water	Soluble coffee constituents
	Pressed olive pomace	Hexane	Olive oil
	Sugarbeets	Water	Sucrose
Supercritical fluid	Green coffee beans	Carbon dioxide	Caffeine
	Milkfat	Carbon dioxide	Cholesterol
	Lemon peel	Carbon dioxide	Lemon oil
	Hops	Carbon dioxide	Alpha acids

because of the potential advantages of better product quality, supercritical fluid extraction has found application in several industries.

Although a large number of supercritical fluids are suitable for food applications, supercritical carbon dioxide has been the solvent of choice because of its safety and its moderate temperature and pressure requirements.

A system for supercritical fluid extraction consists of four basic components: a solvent compressor or pump, an extractor, a temperature/pressure-control system, and a separator or adsorber. Additionally, other equipment, including ancillary pumps, valves, back-pressure regulators, flow meters, and heater/coolers for temperature control of the fluid, are needed for proper operation of the process. *See* SUPERCRITICAL FLUID. Syed S. H. Rizvi

Extrusion

Extrusion has become a widely practiced food processing operation used for the production of snack foods, ready-to-eat breakfast cereals, soup and beverage bases, pregelatinized starches, baby foods, candy products, full-fat soy flour, and textured vegetable protein. In addition to human foods, most dry and semimoist pet foods are also produced on extruders. The ability of the food extruder to perform both cooking and forming operations within a single machine having high capacity and continuous-processing capability has been the principal reason for the increased interest in the process. Other popular aspects of the process include its energy efficiency, a capability to handle a wide variety of food ingredients, and an ability to produce many types of products while creating no undesirable waste streams.

Food extrusion began with the development of the macaroni extruder in 1935. The macaroni extruder has a blender to combine the raw ingredients consisting of semolina wheat flour and water, and a transfer auger to move the mixture to the flighted extrusion screw which thoroughly mixes the ingredients and forces them through a die to form the shaped product.

Types of extruders. There are two types of extruders commonly used in the food industry. The most common is the single-screw type having one flighted Archimedes screw rotating within a stationary barrel to accomplish the processing. The operation of the single-screw food extruder is governed by the speed of rotation of the screw and the design of the flights along its length. Screws with shallow flights impart greater shear and work, rapidly increasing the product's temperature and producing higher pressures behind the die. Such equipment is used to manufacture completely precooked products such as ready-to-eat cereals, snacks, and textured vegetable protein. Extruders with deeper-flighted screws provide a less severe processing environment, which is necessary for the forming of macaroni or chewing gum without appreciable heating of these products. Most food extrusion equipment is of the single-screw type.

The second extruder type has a twin-screw configuration (**Fig. 39**) which employs parallel

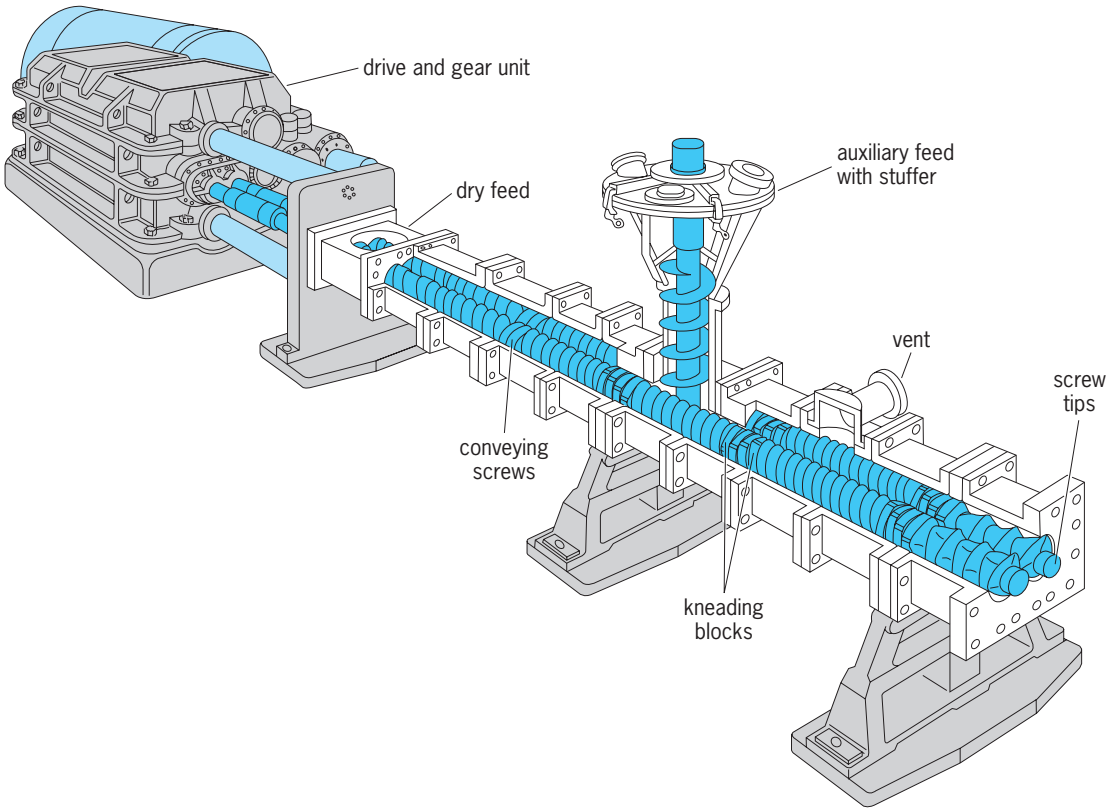

Fig. 39. Twin-screw extruder with cutaway exposing screw configuration. (*Werner and Pfleiderer Co., Ramsey, New Jersey*)

intermeshing screws which turn within a figure eight-shaped barrel opening. A typical twin-screw machine employs various screw elements for conveying, kneading (working), and pressurizing the sample. The advantages of the design are its greater ability to handle low-moisture ingredients, better control of processing conditions, and enhanced capability to convey ingredients.

Extruder operation. Major raw ingredients for extrusion processes consist of ground cereals or mill feed fractions, starch, vegetable protein meal, fat, and sugar. Minor ingredients include salt, acid or base compounds, emulsifiers, color, and flavor. Water is added to this mixture to achieve moisture contents ranging from 10 to 35% with a typical value of 25%. Ingredients may be preheated to almost 212°F (100°C) through the addition of steam to initiate cooking, which increases the capacity of the extruder and provides operating flexibility in a step called preconditioning. The preconditioner is a paddle blender capable of mixing liquid and solid food which is placed just before the extruder.

In a single-screw food extruder (**Fig. 40**) the granular raw or preconditioned ingredients fall from the feed hopper into the relatively deep pockets (channel) between the flights in the feed section of the screw. The flights on the turning screw continuously push the ingredients forward through the barrel and work them into a dough. Reducing flight heights down the screw's length decreases the area available for flow within the channels of the screw, which in turn partially restricts, the flow of the food ingredients during their passage. The restriction causes the channels of the screw to become completely filled with the food product, which increases the mechanical working of the dough. The mechanical energy input from the drive motor used to turn the screw is dissipated into heat by fluid friction and serves as the principal source of the energy necessary to cook the product. In addition to the mechanical energy heating the product, heat can be added through jackets surrounding the barrel and, in some cases, by direct injection of steam through the barrel wall. In this latter case, the action of the flights mixes all components within the screw channel to produce a relatively homogeneous mixture.

In cooking extrusion, the temperature of the food ingredients rapidly rises after they are compressed and worked to temperatures as high as 300–390°F (150–200°C) just before the die. Because the residence times in the extruder are less than 60 s, the process has been termed high-temperature short-time (HTST). High-temperature short-time processing preserves the nutritional quality of the food but inactivates enzymes, microorganisms, and other undesirable factors. An environment similar to that of a pressure cooker prevents boiling or flashing of moisture within the extruder because of the elevated pressures.

The most dramatic changes occur to the raw food ingredients in the compression section of the screw. The combination of heat and working converts the discrete particles in the raw feed into a continuous

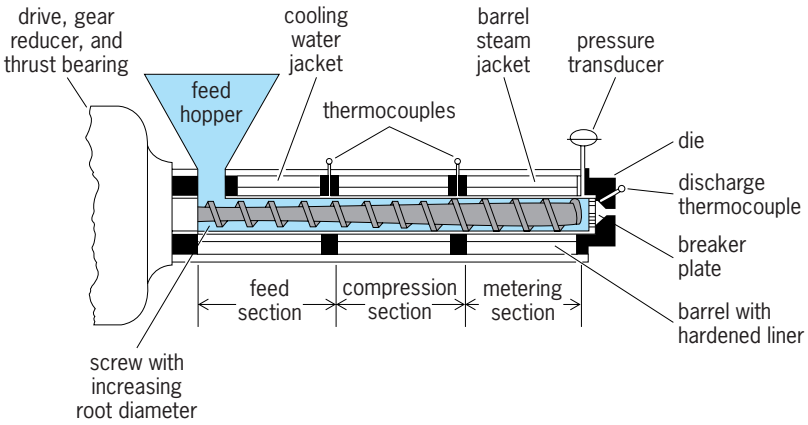

Fig. 40. Cross-sectional drawing of a single-screw food extruder. (*After J. M. Harper, Extrusion processing of foods, Food Technol., 32(7):67, 1978*)

and very viscous dough. Physical changes accompanying this transformation include the hydration and disorganization of raw starch (gelatinization), the denaturing of protein including the chemical crosslinking of adjacent molecules to form a polymerized structure, and the melting of both starch and protein to give a translucent dough when the processing is complete. Typical doughs have much higher viscosities than bread dough because of the relatively low moistures involved in the process.

The metering section of the screw has the shallowest flights, which enhances mixing to assure uniformity of the high-temperature dough. The metering section increases the pressure on the dough, which provides the force needed to push it through the die opening. The configuration of the die openings creates specific product shapes which are cut to a desired length with a rotating knife at the face of the die. Under high-temperature cooking extrusion conditions, puffing occurs once the dough exits the die and is exposed to atmospheric pressure. The puffing is the result of the expansion of the high-temperature water within the product. Loss of moisture accompanies puffing, causing the product's temperature to drop rapidly to about 140°F (60°C). Further cooling and drying are necessary to prevent adjacent pieces from sticking. Extruded ready-to-eat cereals and snacks exhibit an open cellular structure typical of extrusion-cooked and puffed cereal products. Protein flours extruded in a similar manner will puff, creating a chewy characteristic similar to meat because of the crosslinked protein molecules.

Extrusion processes. The extruder is just one of a series of pieces of equipment necessary to produce convenience food products. The sequence starts with bulk tanks that hold the raw ingredients in flour form (**Fig. 41**). These are metered into the extruder which cooks and forms them. The individual, cooked and expanded pieces are pneumatically conveyed to a dryer-cooler to reduce their moisture to less than 10%. Finally, flavor, oil, sugar, color, or vitamins and minerals are enrobed (coated) onto the surface of the pieces. To complete the process, the finished

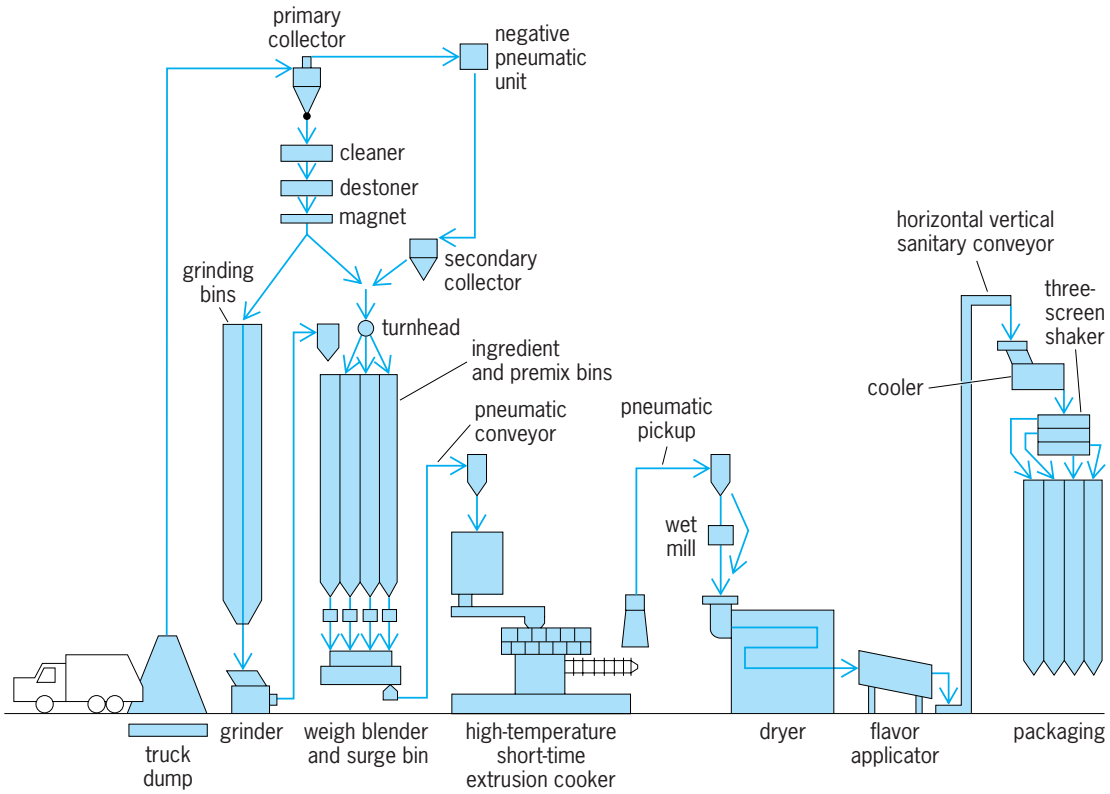

Fig. 41. Flow diagram for an extruded food process. (*After O. B. Smith and S. C. Crocco, Engineering "meat," Food Eng., 47(10):EF22, 1975*)

products are packaged in bags, boxes, or cans for distribution and consumption. Judson M. Harper

Cleaning and Sanitation

Generally speaking, the words cleaning and sanitation as used in the food industry have a similar meaning in that cleaning is recognized as being necessary for proper sanitation. Cleaning is the treatment given all product contact surfaces following each period of use in order to remove all physical evidence of residue. Typically, after cleaning is completed, it is followed by a chemical or thermal treatment to sanitize the product surfaces. In comparison, sanitation may be defined as the systematic control of environmental conditions during processing, storage, and transportation of food products to prevent contamination by animals, insects, and microorganisms.

Since the sanitation aspects of processing, storage, and transportation are specialty fields, involving many areas of expertise, this discussion will deal only with cleaning as related to processing operations in a food plant, with emphasis on clean-in-place methods.

Cleaning methods. The cleaning of product contact surfaces in a food plant can be carried out through the use of one or more techniques described in the following.

(1) In manual or hand cleaning of equipment, the equipment is disassembled, rinsed with a hose, and washed, typically using a mild wash solution and a hand brush.

(2) In semiautomatic cleaning, the equipment again is disassembled, rinsed with a hose, and then put in a clean-out-of-place tank (**Fig. 42**), which provides mechanical action of the wash solution through the utilization of a recirculation pump, and chemical action via the detergent solution which is maintained at a preset temperature.

(3) Spray or foam cleaning, utilizing spray wands, is frequently employed in the food industry

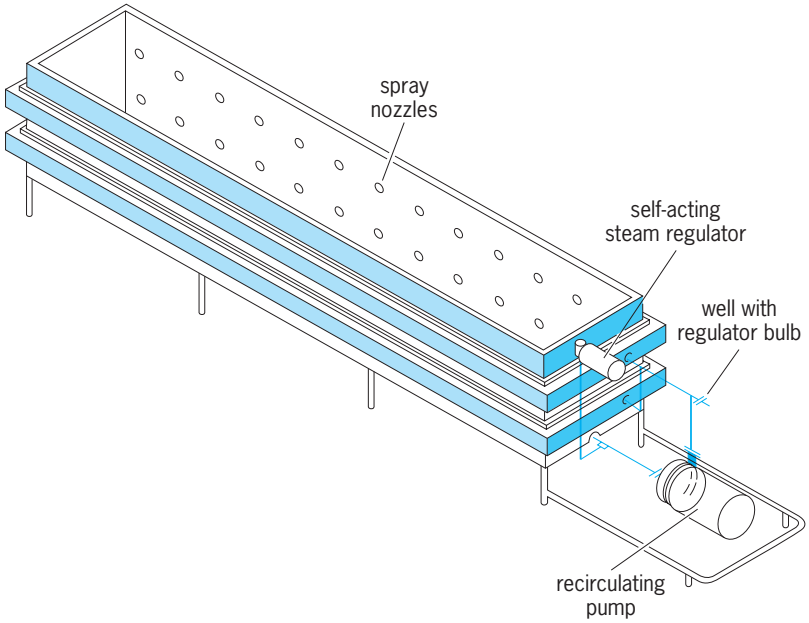

Fig. 42. Diagram of a clean-out-of-place cleaning system.

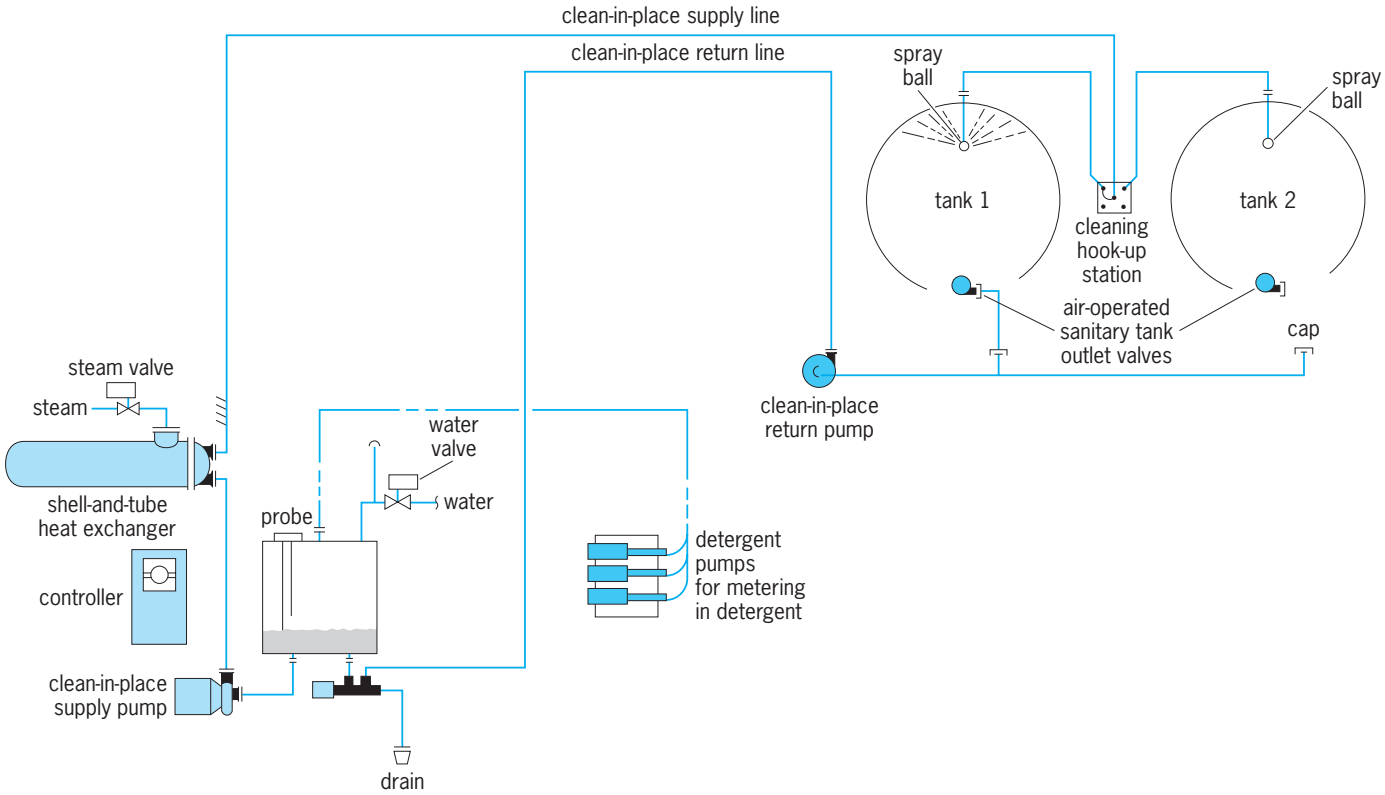

Fig. 43. Clean-in-place single-tank–single-use recirculating system.

(particularly in meat processing plants) to clean both the exterior and interior of equipment (which is not usually designed for automated cleaning) without any major disassembly of the equipment. Either low pressure (less than 100 lb/in.^{2}, gage, or 690 kPa) or high pressure (approximately 500 lb/in.^{2}, gage, or 3.5 MPa) sprays may be used for these operations.

(4) Clean-in-place is becoming the preferred method of cleaning the internal surfaces of processing equipment and the associated product pumps, sanitary automatic valves (and valve groups), and sanitary product transfer lines. The accepted definition of clean-in-place is cleaning entirely by circulation and flowing chemical detergent solutions and water rinses onto and over the surfaces that are to be cleaned by mechanical means. Cleaning-in-place techniques are utilized for cleaning both the inside surfaces of product transfer lines and the processing equipment. Generally speaking, the utilization of cleaning-in-place for product lines can be a relatively simple procedure if there are no product valves associated with the cleaning operation. However, with the incorporation of sanitary air-operated valves (such as valve groups), the proper directional flow of the clean-in-place solution through the valves can be quite complex. Cleaning-in-place for processing equipment such as storage tanks (horizontal or vertical), processors, pumps, blenders, heat exchangers (plate or tubular), evaporators, and dryers is generally complex, because not all of the equipment is specifically designed for clean-in-place applications and generally requires the addition of the proper type of spray device to assure complete coverage of all product contact surfaces, including agitator assemblies and baffles. The utilization of a permanently installed, nonrotating type of spray device of sanitary design is preferred. A simple unit is available that can be applied to cleaning operations for both product lines and processing equipment (**Fig. 43**).

Factors affecting cleaning. Regardless of the techniques utilized to clean product-contact surfaces of processing equipment, all of the food industry soils are composed of the same basic ingredients—protein, fats, carbohydrates, sugars, and mineral salts.

Soil removal through the use of detergent solutions is dependent upon four functions: (1) concentration of the detergent solution; (2) temperature of the detergent solution; (3) time that the detergent solution is in contact with the product surface; and (4) velocity or mechanical action of the detergent solution. As a rule, if the concentration of the detergent solution is increased, either the temperature of the solution or the time element can be reduced, with the same overall cleaning effect. This general relationship would also apply for increasing the temperature or the time to reduce the concentration of the detergent solution.

Detergent selection. Cleaning compounds are required to assure complete removal of all residues from product contact surfaces. Selection of a cleaning agent depends upon the composition of the soil, the nature of the soil, methods of cleaning, and

materials of construction of the equipment being cleaned. Alkaline cleaners are effective in removing residues composed of fats, carbohydrates, sugars, and proteins. High levels of alkalinity (caustic) are required for effective removal of proteins. Sodium hypochlorite, when added to an alkaline cleaning solution, can aid in the removal of proteins at lower levels of alkalinity. Mineral salts, whether a component of the food product or residues from water, require acid cleaners.

Alkaline cleaners used in clearing-in-place are generally based upon sodium hydroxide (caustic) together with water conditioning agents, dispersants, and special low-foaming surfactants. Acids detergents used in cleaning-in-place are often based upon phosphoric or combinations of phosphoric and nitric acid with or without low-foaming surfactants. Heavy soils often require both an alkaline and acid detergent applied in two steps. Acid-based sanitizers can be used to remove light mineral films and at the same time sanitize the equipment.

Detergents used for cleaning-in-place can be highly active with an alkaline pH in excess of 11.5 or acids with pH less than 2.0. Temperatures of 140–180°F (60–80°C) are typical, depending upon the equipment designed and amount of soils present. For the cleaning of non-stainless-steel equipment, such as aluminum, special noncorrosive alkaline cleaners using silicates are required.

Following cleaning and rinsing, a treatment with a chemical sanitizing agent is required to assure a sanitary environment for the production of high-quality foods. In the United States, sanitizers must be registered with the Environmental Protection Agency (EPA) and comply with the regulations of the Food and Drug Administration (FDA) as an indirect food additive. Sanitizers used in the food processing industry are based upon chlorine, iodine (iodophors), quaternary ammonium chlorides, acids, and peracetic acid. Registered sanitizers, when applied as labeled to surfaces that have been cleaned, assure a level of sanitation that protects the safety of the food supply and meets quality concerns of the producer. *See* DETERGENT; SURFACTANT.

Equipment design. There are some guidelines that must be followed to ensure proper cleaning of the food processing equipment. (1) The processing equipment must be constructed of inert, nontoxic, nonporous material such as stainless steel; plastics and other synthetic materials must be approved by the FDA or the Department of Agriculture. (2) Processing equipment should be designed according to sanitary design specifications based on standards of the Technical Committee of the Dairy and Food Industries Supply Association. (3) Processing equipment should be specified to incorporate "3-A" sanitary standards, which assures that the equipment is properly designed from a sanitation standpoint and uses only properly approved contruction materials. (4) In addition, the processing equipment, pumps, valves, and associated equipment should be designed for clean-in-place by the manufacturer. The design aspects necessary to meet this requirement can be complex, but they include such items as utilization of radius corners of $^1/_8$ in. (0.3 cm) or more rather than the more conventional 90° angle; smooth, self-draining surfaces; ease of disassembly; and ease of access to all product-contact surfaces for hands-on or visual inspection. The first attempts at in-place cleaning were carried out on the plate heat exchangers utilized in the dairy industry for high-temperature short-time pasteurization of milk. Initially, powdered detergents were manually added to the balance tank of the high-temperature short-time system, and the product pumps were used for recirculating the cleaning solution through the heat exchanger.

Early work involving in-place cleaning of storage vessels utilized specially designed spray balls which were inserted into the access hole of the storage tank to assure proper coverage of the interior surfaces. This early developmental work proved successful, and was the basis for applying clean-in-place operations to other pieces of equipment.

Generally, reference to clean-in-place applications infer that the equipment can actually be cleaned in place without any major disassembly of the equipment. With the development of larger pieces of processing equipment—particularly storage vessels—and the development of sanitary air-operated valves, sanitary products pumps, and welded product lines, the disadvantages of manual cleaning became evident.

Typical clean-in-place cycle. Basic clean-in-place operations in the food processing industry can be divided into one of three different cycles for line cleaning operations (which may include pumps, valves, or plate heat exchangers); vessel or spray cleaning operations (associated with tankers, tanks, processors, blenders, dryers, evaporators, and so forth); or the combination of the two in one program. Generally speaking, it is advantageous not to combine the cycles.

There are five basic steps associated with a typical clean-in-place operation: (1) pre-rinsing of the equipment with an ambient-temperature water to remove as much of the soil as possible, with the pre-rinse discharged to the drain; (2) recirculation of alkali or acid detergent solution at preset concentration and temperature; (3) post-rinsing with ambient temperature (or warm) water to remove wash solution residue, with the post-rinse discharged to drain; (4) rinsing of equipment with mild acidic post-rinse solution; (5) sanitization of equipment prior to use, not necessarily following clean-up of equipment.

This five-step sequence may be considered typical, but many variations are found in the field due to the type of clean-in-place recirculatory system utilized (that is, single-use, multiuse, single-tank, multitank, and so on); the type of equipment being cleaned; and the type of soil encountered. In some cases only an acid detergent may be required; in other cases, acid will be used followed by a caustic solution or such. Accordingly, for proper cleaning, consideration must be given to the use of appropriate chemicals, and the selection of a proper clean-in-place recirculating unit and associated controls, sprays, return pumps, and

so forth, as part of the total system. In addition, the equipment to be cleaned should be designed for ease of clean-in-place applications. E. F. Reinke

Waste Management

Waste disposal is a matter of great economic and technical concern for most food processing operations, as well as other industries. Both solid and liquid wastes are encountered in the food processing industry. The disposal of these wastes is regulated by the Clean Water Act of 1977 (as amended) and related environmental protection legislation, which govern the types and amounts of waste discharged into the environment. As federal, state, and local environmental regulatory agencies increase their demands for less waste discharge, it becomes essential that cost-effective methods of waste disposal be found.

Waste management is a modern approach to cost-effective food processing waste disposal. Through waste management, modifications are applied to food plant operations and manufacturing processes. These modifications reduce the amount of solid and liquid wastes, recover more product and by-products, often reduce energy consumption, and exhibit other benefits. In general, the principle is to convert waste liabilities into profitable assets.

One major objective of waste management is to eliminate or at least lessen the dependence on end-of-the-pipe sanitary engineering methods. This is achieved by reducing both the amount of waste solids generated and the volume of wastewater discharged.

Modifying plant operations. The following are examples of modifications which can be made to food plant operations: incorporating good housekeeping practices; collecting culls and other solid wastes into containers rather than discharging to the floor drain; recycling water; reusing spent process water in another plant operation; and using less or no water in plant operations that formerly used a fair to a large amount of water.

Good housekeeping practices that reduce water usage and wastes require good personnel management and employee awareness of conservation practices. Such practices as needless use of water or overloading of containers, thereby causing spillage, should be discouraged.

Recycling water in the same plant operation can be achieved by treating spent process water with activated charcoal or sand filters or by ion-exchange columns, chemical treatment, pH adjustment, temperature adjustment, pasteurization, or a combination of these and other methods. Brines are one of the more troublesome wastewater discharges to treat because of their high salt and organic content. Rather than discharging for wastewater treatment, spent olive, pickle, and cherry brines have been treated and recycled to start new batches of product (**Fig. 44**). Recycling of water has been used in other food plant operations.

Countercurrent water reuse systems can be established in many plant operations. For example,

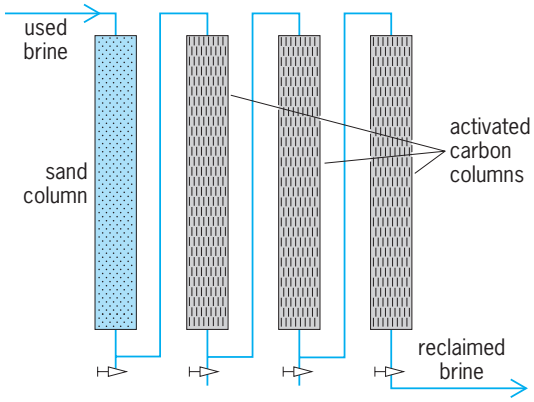

Fig. 44. Brine reclamation system used to recycle spent pickle, olive, and cherry brines for the purpose of initiating new batches of products. (*After J. H. Green and A. Kramer, Food Processing Waste Management, AVI Publishing Co., 1979*)

potable water used to rinse the finished product prior to packaging can be used a second or third time to wash the product during intermediate stages of processing and to wash the incoming commodity. The spent wash water could be used again to initiate washdown of dirty floors or to flume solid wastes away from the processing line.

Examples of methods that use less water or no water in plant operations compared with former methods are dry fluming such as mechanical or pneumatic rather than hydraulic transport systems; dry cleaning using shakers or brushes to initially clean dirt from commodities or equipment; and using high-pressure, low-volume cleaning wands for plant and equipment cleanup. The last example also reduces the amount of time, labor, hot water, and detergents needed for plant sanitation operations.

Process modifications. These often achieve the greatest reductions in water usage and solid waste discharge; they may also exhibit other benefits such as energy conservation. Process modifications require a large amount of time, technical input, and money to do the necessary research and development. Federal agencies such as the Environmental Protection Agency (EPA) and the U.S. Department of Agriculture (USDA) have funded projects through their own researchers or through universities and trade associations. Trade associations such as the National Food Processors Association (NFPA), in cooperation with industry and federal agencies, have been leaders in the research and development of food processing modifications. The return on investment has been outstanding for many of these process modification projects.

Two of the best examples of food processing modifications were the development of dry (mechanical) peeling methods (late 1960s) and steam peeling methods (1970s) for potatoes, other root crops, tomatoes, and tree fruits. For many years (prior to the late 1960s) potatoes and other root crops were peeled by immersing them in hot caustic solutions and removing the treated peel with blasts of jet-sprayed water. In the dry peeling processes, hot caustic or steam-treated peel is mechanically removed by

rubber or neoprene brushes and rotating disks. In the late 1970s, steam peeling methods were developed for root crops and tomatoes. The advantages of dry peeling are that no water is used in peel removal; the peel waste can be collected separately for by-product development; often more product (less waste) results; less volume of hot caustic peeling solutions is required; and the pollution waste loads are greatly reduced. This process modification has been shown to be very cost-effective, and it conserves energy.

Modifications of the blanching process have reduced the tremendous waste loads that often resulted from the use of large volumes of hot water in blanching tanks. Modifications to blanching include individual quick blanching, steam blanching, or hot-air blanching. Mechanical innovations such as the vibratory spiral blancher-cooler developed by USDA have increased the efficiency of steam blanching. These modifications have resulted in reducing waste loads, retaining more nutrients in the product, and conserving energy.

Other examples of process modifications include the following: deboning to recover more product from meat, poultry, and fish; ultrafiltration followed by reverse osmosis to separate cheese whey into proteins, lactose, and minerals; ultrafiltration or reverse osmosis or both used in other applications to separate components for by-product development; use of heat exchangers and other process modifications in the citrus industry to save energy, reduce pollutant loads, and recover by-products. The fishing industry has its own particular waste disposal problems. The traditional methods of fishery waste disposal have been throwing the waste overboard from fishing vessels and tossing the waste over the end of the wharf by the processing plant. These traditional methods of fishery waste disposal have probably fostered a lack of effort by this industry to find ways of utilizing fishery wastes or to find methods of improving processing efficiency. Local and national concerns have discouraged these traditional methods of fishery waste disposal to protect the environment. Solid waste losses from processing fish and shellfish have ranged from about 30 to 85% depending upon the type of fishery: fin fish filleting operations generate 30–60% losses, shrimp processing operations generate 40–80%, and crab processing operations generate 75–85%. There is a recognized need to improve resource recovery from fishery processing operations, and this has generated some specialized waste management operations for the industry.

Mechanical deboning of underutilized species, or fish processing wastes, has been carried on for many years in some segments of the industry. In the late 1970s, interest in the Japanese methods for using surimi (deboned fish flesh) was transferred to the development of a wide variety of new products from deboned fish. Thus former wasted fish flesh from underutilized species of fish, or on the frames of filleted fish, can now be converted to many new products.

The traditional fish meal industry is capital-, energy-, and technical-labor-intensive. It has been replaced in some parts of the world by the fish silage (liquefied fish) process. Industrial fish, underutilized fish, and fishery wastes can be treated with organic or inorganic acids, or combinations of both, to form fish silage, an animal feed product. It is formed by the autolytic enzyme breakdown of fish proteins. The fish silage process does not require large amounts of capital, energy, or labor and is therefore well suited for small as well as larger fish processing operations for the purpose of developing by-products from wastes. However, fish silage is about 80–85% water, and costs of transportation becomes a limiting factor for its use. If there are no local animal feed markets for fish silage, it can be combined with cereals, dried together, and shipped to distant markets.

By-product development. Recovered food-processing waste solids offer the opportunity for development of food, feed, fodder, fiber, and fuel by-products. In addition to the examples of by-product development given above, there are other innovative methods such as microbial technology used to develop single-cell proteins, fuels (for example, alcohol and methane), enzymes, and other organic chemical products from food-processing wastes and wastewaters.

Some solid wastes such as olive pits, nut shells, and bagasse (sugarcane waste) are being used as fuels. Many sugar mills in Hawaii and Brazil burn bagasse to generate steam and electric power to operate the sugar mill. Often excess electric power is generated, which is sold to the electric utility company. *See* CO-GENERATION SYSTEMS.

Shellfish wastes contain a large amount of chitin in addition to protein and calcium carbonate. Chitin, which constitutes much of the exoskeleton and some of the internal parts of shellfish, is essentially inert to most chemical reactions and therefore has little industrial use. However, shellfish wastes can now be transformed into useful products (**Fig. 45**). The wastes are treated with dilute sodium hydroxide and dilute hydrocholoric acid to produce proteins, calcium chloride, and chitin. Then a hot concentrated (40–50%) sodium hydroxide treatment deacetylizes the chitin, forming chitosan, which has a wide variety of uses and is therefore a valuable product. Shellfish processing wastes are transformed into chitosan and other by-products (protein and calcium chloride) in Japan and some other countries.

Sanitary engineering. Although the objective of waste management has been to greatly reduce or eliminate the need for end-of-pipe treatment, sanitary engineering methods are still applied to many food processing waste effluents. Both primary sedimentation and secondary biological sanitary engineering treatment methods have been applied to food processing waste effluents. Typical secondary treatment methods are: various forms of activated sludge treatment, including the oxidation ditch method; trickling filters, including the rapid and super rate modifications; anaerobic, aerobic, and aerated lagoons; and land application of wastewaters such as crop irrigation, infiltration-percolation, and overland flow methods. The activated biofiltration

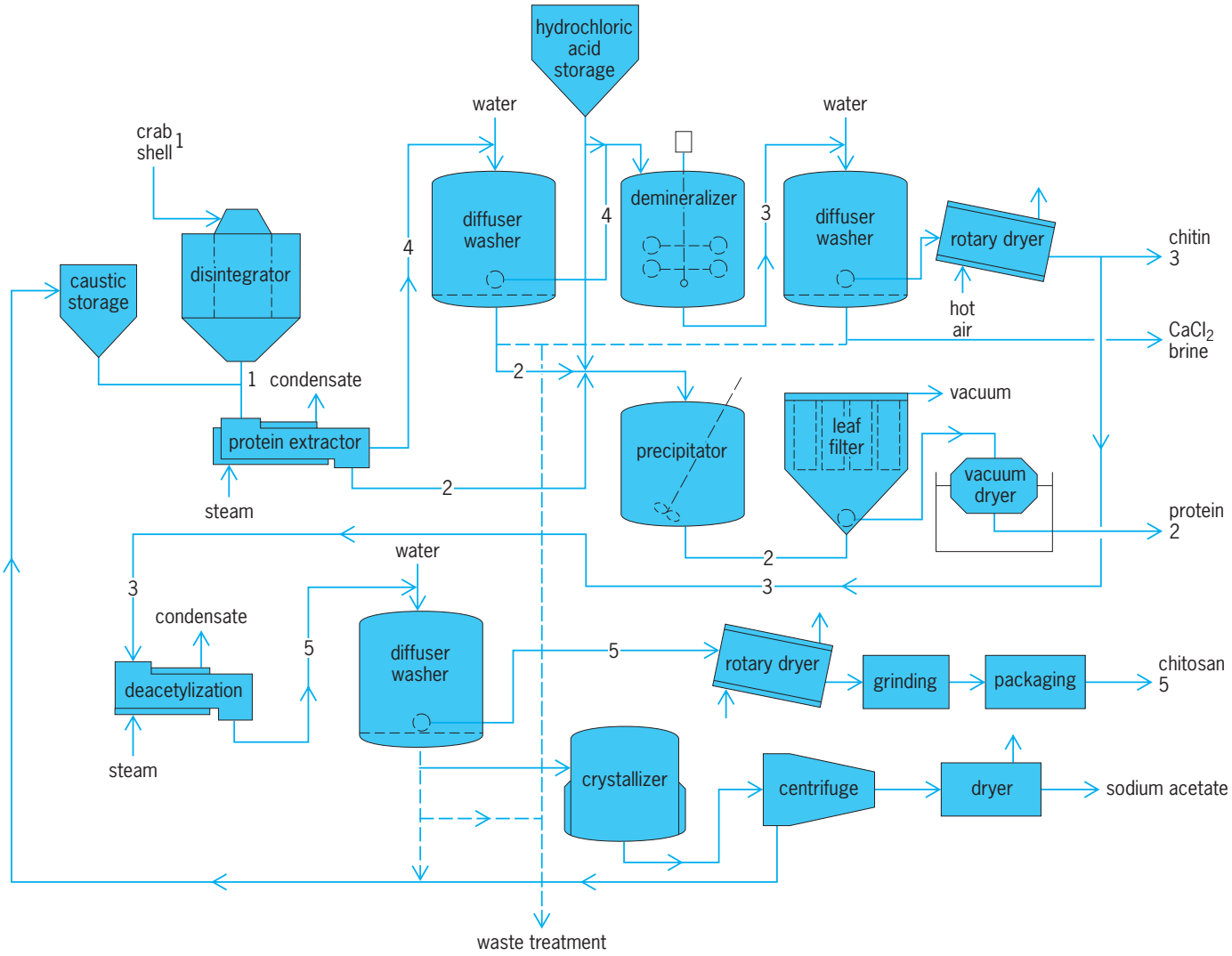

Fig. 45. Chitin-chitosan process flow chart. Process number (1) indicates waste crustacean shell, (2) protein, (3) chitin, (4) deproteinized crustacean shell, and (5) chitosan. (*From J. H. Green and A. Kramer, Food Processing Waste Management, AVI Publishing, 1979*)

method, a combination of trickling filter and activated sludge technologies, has proved to be very effective in treating fruit and vegetable wastewaters.

Many end-of-the-pipe sanitary engineering methods are both costly, in terms of capital investment and operational costs, and nonproductive. Processors that discharge into available municipal sewer systems will encounter increased service charges and probable requirements for pretreatment prior to discharge. On-site sanitary engineering installations can be modified to recover solid wastes for by-product development. Some food processors are processing sludge recovered from activated sludge treatment of food wastes into animal feed. Treated wastewater can be used for crop irrigation or for recycling as process water.

Food processors, as well as other manufacturers, had to consider tertiary or advanced wastewater treatment methods in order to meet the EPA's 1985 zero discharge effluent requirements. As an alternative to advanced wastewater treatment, many food processors are considering land application of primary or secondary treated wastewaters. Food processing wastewaters, which do not contain toxic substances, can be used to grow crops or recharge ground water. There are three methods of land application used for food processing wastewater (**Fig. 46**): application to crop land (clay to sandy, loam) to irrigate crops or orchards; applications to sandy loam or sandy soil to recharge groundwater; and overland flow to clay or clay loam soils planted with grass ground cover. In the last, the grass sets as a filter system and holds back waste solids. The harvested crops provide monetary return on the costs of waste disposal.

A food technology profession has emerged, known as waste or resource recovery management. The waste manager is closely allied with (if not the same person as) the energy conservation manager. The challenge is to find ways of reducing water usage, recovering waste solids, developing marketable by-products, conserving energy, and

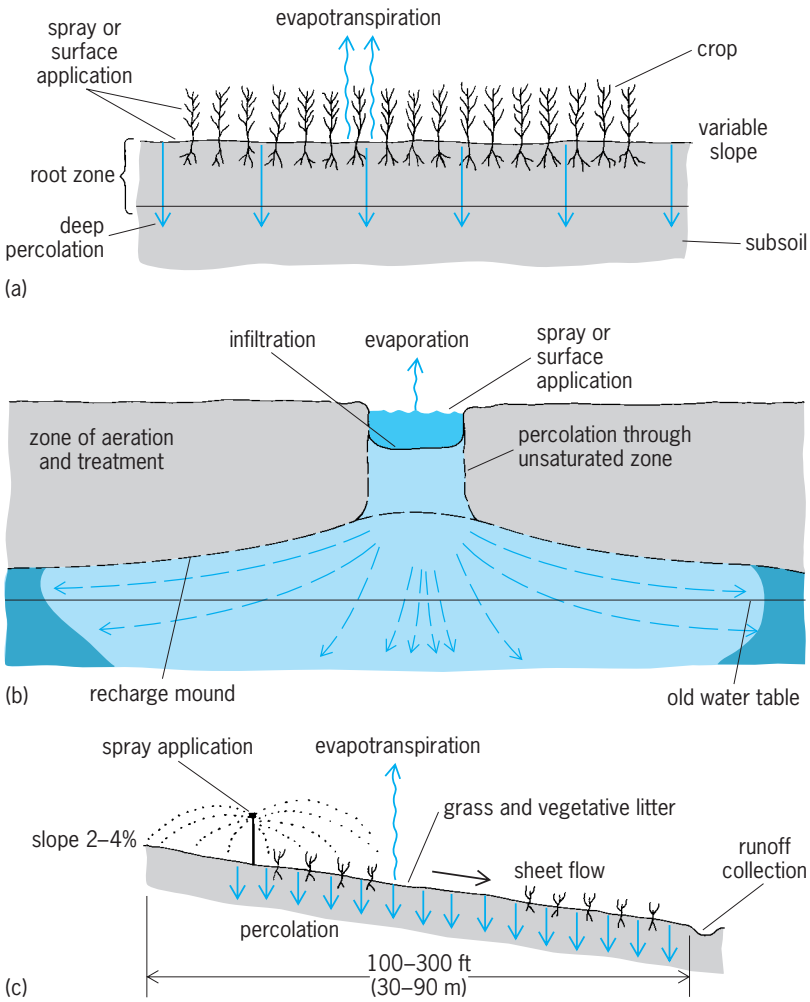

Fig. 46. Land application of pretreated food-processing wastewater. (*a*) Irrigation; sandy-loam soils to raise crops. (*b*) Infiltration-percolation; sandy soil to rechange ground water. (*c*) Overland flow; graded clay soils planted with grass or other ground cover. (*After J. H. Green and A. Kramer, Food Processing Waste Management, AVI Publishing Co., 1979*)

finding alternatives to costly, nonproductive end-of-the pipe treatment methods. *See* ENVIRONMENTAL ENGINEERING; SEWAGE TREATMENT; WATER POLLUTION.

John H. Green

Process Control

In what was probably the first food engineering application of feedback techniques, Oliver Evans in 1784 used control techniques to automate a Philadelphia flour mill. Correctly applied, control instrumentation can give a product closer to specifications than manual control with reduced waste and increased yields. Advantages of automated process control include higher efficiencies with maximum utilization of all ingredients and raw products. *See* AUTOMATION.

Control tasks. The implementation of process control is based on the particular control task. There are four fundamental control tasks in industrial food processing: sequential control, loop control, interconnected control, and adaptive control. *See* CONTROL SYSTEMS.

Sequential control. This is for the independent batching or sequence operations. Most processing of food is derived from familiar domestic patterns that are performed in sequential batch-type steps. Controlling elements in such sequential control paths are typically handled by programmable logic controllers. *See* PROGRAMMABLE CONTROLLERS.

Loop control. This control task is based on a feedback or feedforward information system. The control loop is fundamentally an algebraic summing point between a set point (desired value) and the measurement (current value). The corresponding error is used as a means of implementing a specific corrective response. The most common type of control loop is a feedback control loop. Basically the feedback controller reacts to the output from the process and regulates the process based on that output condition. Feedforward control, an alternative approach, refers to the control action in which information concerning all the conditions that can disturb the variable being controlled is converted into a specific corrective action in order to minimize deviations of the output.

Of the four general control tasks, loop control is the one most utilized within the food industry. It is a fundamental component of the interconnected and adaptive control tasks, and is usually found as an associated control system within sequential control tasks. There are four critical elements to any loop control system: (1) the process to be controlled, (2) the measurement of the current conditions using sensors, (3) the controller, (4) the final output from the controller affecting the system.

All of the techniques grouped together under the term loop control have one feature in common, and that is a continuous path from sensor through a controller and then to an output. This path can be broken intentionally, thus producing an open-loop condition for study of the controller or maintenance of the system.

Interconnected control. Under this type of control, the controller has to adapt its output, because when it controls one variable it also, by the action of changing or altering that variable, is altering a second controlled variable. A condition called coupling exists when a change meant to alter one variable also alters another variable. Interaction exists if there is coupling the other way around.

Adaptive control. A control task typical of applications in food processing is sometimes called adaptive control. Under adaptive control a model of the system is constantly tested and updated based on raw product input and processing conditions. Changes are implemented automatically in the control strategy by adjusting various parameters within the process system. There are many opportunities for this type of modeling and optimization within the food processing industry. *See* OPTIMIZATION.

Single-loop controller modes. There are basically four types of common control modes within single-loop controllers. How a decision is made concerning the implementation of a specific control mode depends on a number of factors, including economics,

precision of control required, time response of the process, processing characteristics, and safety considerations.

Two-position control. Within this classification there are three common control options. On-off control produces a control function based on a measured variable, comparing it to a set point. When the measured variable crosses the set point, the output is altered to the opposite position. The output can be in only two positions: full on or full off. The second type of two-position control is a differential gap controller, whose output is based on the time a measured variable crosses an upper and lower boundary. The third type is a time cycle system. Typically in such a controller a specific time cycle is established in which the final output must complete a full cycle; that is, if it starts in an open position, it must close and then reopen during that cycle. The time in which a valve occupies the position opposite to that it occupied at the start of a particular time cycle depends on the measurement point and whether it is above or below the set point.

Throttling control. The three other control modes are based on proportional or throttling control. Throttling control indicates any type of control system in which the final output is purposely positioned to achieve a balance between the supply and demand of the process. This basic type of throttling control is called proportional-only control. Under this type of control the position of the final output is based on the error between a set point and the measured variable multiplied by a gain factor and added to a bias. Proportional-only control attempts to return a measurement to a set point after a load upset has occurred. However, based on its design, it is impossible for a proportional controller to return the measurement exactly to the set point.

The difference between the measurement and the set point after a proportional-only controller has come to equilibrium under new load conditions is known as offset. If the offset cannot be tolerated, an additional control function must be added to the proportional-only control resulting in the third type of controller, proportional-plus-integral control. Integral action will integrate the difference between measurement and set point over time and cause the controller's output to change until that difference is brought to zero. The basic limitation of this proportional-plus-integral controller is called integral windup. This must be seriously considered on discontinuous or batch processes, because under these processes the controller output can be saturated and overshoot the set point. A mechanism must be integrated into the controller system in order to prevent this overshoot.

The fourth and final type of controller takes into account throttling control, maintenance of zero offset, and control action based on the rate of error change. This controller is known as a proportional-plus-integral-plus-derivative controller. Derivative action occurs in addition to the proportional and integral action whenever the measurement signal is changing. The derivative action is proportional to the rate of change in the measurement signal. By reacting to the rate of input change, the derivative portion of the controller allows the unit to inject more corrective action than is initially necessary in order to overcome system inertia.

Sensors. In any process control situation, accurate and reliable measurement of process conditions are critical in order to implement an adequate control strategy. A lack of suitable sensors for many applications limits the progress of automation in a number of areas. Producing sensors of sufficient accuracy, reliability, and ruggedness for operation under plant conditions is many times more difficult than designing equipment for making the same measurements under laboratory conditions. Measurements commonly made in the food industry include temperature, flow, and level. Some more specialized requirements involve measurements of viscosity, pH, conductivity, humidity, density, and color.

Process regulators. Every control system has a means for switching or regulating process equipment by some sort of final control element. There are a number of devices available for convenient on-off switching. Power switching employs relays, contactors, and solid-state relays; pneumatic and hydraulic power use solenoids. Reliable and efficient regulation of speed or flow on a continuous basis is difficult to accomplish simply with on-off devices.

Continuous variation of temperature, speed, flow rate, and other process conditions is generally achieved with one of four types of regulators: flow-controlling valve, variable-speed motors, mechanically variable pumps, and variable-power electric heating. The flow-controlling valve is the most commonly used regulator. It usually requires special sanitation precautions, and it is generally inefficient because it imposes pressure drops in the control lines. However, despite these disadvantages, it is fairly popular because in its manual form it is very simple and direct. Generally it needs only one moving part.

Pneumatic linear accuators are popular for driving valve stems. Simple types use air on one side of the diaphragm working against a spring; other types use pneumatic position feed. Pneumatic signaling with a working range of 3–15 lb/in.^{2} (21–103 kPa) has been widely used for many years in virtually all pneumatic control systems. Where computers and electronic integrated circuits are installed for performing the control function, pneumatic accuators may still be used through the use of an electronic transducer. This transducer converts digital or analog values to specific pneumatic levels within the working range. *See* SERVOMECHANISM; TRANSDUCER.

Direct digital control. A significant advancement in the microminiaturization of digital electronics accompanied by a reduction in cost has facilitated the use of digital logic in all control tasks. Computerization of process control strategies allows development to progress from sequential startup through transition into continuous operation on a virtually 100% automated basis.

The availability of systems with inherently large data processing capabilities may lead to process

automation in small plants or automation by stages, as well as total plant control. Very sophisticated small computers will permit implementation of a very diversified, distributed control strategy, with flexibility and versatility. *See* COMPUTER-INTEGRATED MANUFACTURING; DATABASE MANAGEMENT SYSTEM; DIGITAL COMPUTER.

Mark R. McLellan

Packaging Equipment

Packaging methods are influenced by the type of product and the packaging concept, and include filling, closing, labeling, multipackaging or unitizing, and coding. Primary packaging involves a close partnership between the product to be contained and the packaging material in contact with the food or beverage. Secondary and tertiary packaging encompasses collation or unitization of a number of different primary or secondary packages.

Primary packaging equipment. Packaging lines for preformed packages usually contain components such as accumulators, unscramblers, infeeds, aligners, fillers, closers, and labelers for preformed containers which include cans, glass, bottles, tubs, trays, and similar items. Prior to the infeed, the preformed containers are received from the fabricator, or from an in-plant forming operation, and introduced into a line. Before filling, the containers are cleaned to remove foreign matter resulting from fabrication and transport and to ensure their general cleanliness.

Metal cans are generally received in bulk shipment aligned in tiers on a pallet; glass jars may be received in corrugated fiberboard cases or in palletized or bulk form; and plastic jars and bottles may be received in cases, on pallets, or in dump bins. Preformed plastic trays, cups, or tubs are usually received nested in stacks. Preformed bottles or cans that are received in palletized form are removed from the pallets by mechanical depalletizers such as sweepers or vacuum cup pickers which move the containers onto accumulator tables that feed aligners. Plastic bottles or jars that are received in bulk are unscrambled, righted, and conveyed from an accumulator table to the infeed of the packaging line. Foreign matter within the containers may be removed by air-pressure cleaning or by inverting the cans or jars and subjecting them to air blasts or jets of water followed by air. Cans, bottles, or jars received in cases are mechanically removed from the cases. Trays, cups, or tubs are mechanically denested.

Paperboard folding cartons are received as precut flat or preglued blanks which must be formed or erected. Flat blanks may be erected by applying adhesive to the corners and then forcing the blanks into a mandrel to bring together the walls and tabs or edges to form an open-top or open-ended container. Another common method employs die-cut corner-locking tabs; this requires mechanical locking to erect the trays. Many paperboard cartons are received as preglued tubes which are erected by vacuum removal from a magazine and snapping open. Using preglued sleeve-style blanks, an automatic carton machine (**Fig. 47**) erects the carton, conveys the open carton for product insertion, and seals the carton ends by activating a heat-sealable coating on the paperboard surface. Typical speeds are 70–300 per minute.

In the type of double-package maker that is widely used for dump filling of dry cereals, cookies, crackers, and snack foods, an internal flexible coextruded polyolefin liner is formed around a mandrel after which a paperboard flat blank or a tube is wrapped around the liner on the mandrel. The carton with its inner liner is glued into position, and the entire preformed package is stripped from the mandrel or to a line in which it is filled and sealed.

Flexible pouches are usually formed in line with the filling equipment. Form/fill/seal equipment includes both vertical and horizontal types.

In vertical equipment, a web of flexible material is unwound and turned over a forming collar into a vertical tubular form where the edges are brought together and heat-sealed into a longitudinal tube. While the tube is in the vertical attitude, a transverse seal is formed at a base, the product is filled by gravity feed which may be assisted at the outset, and a transverse seal is made at the top. In most such units, the transverse seal for the top of one pouch constitutes the seal for the bottom of the next pouch in line. Vertical machines generally operate at cycle speeds in the range 30–40 per minute for packages of about 2–6 oz (60–170 g) contents. Output speeds may be increased by multiplying the number of tubes in parallel with each other. Output speed and reliability have been increased by incorporating drive belts on the side of the web on the tube to assist the downward motion of the tub. Vertical form/fill/seal machines are used extensively for snack foods, dry granular food products such as nuts and candies, and free-flowing products such as coffee. In multiple-lane form with face-to-face heat sealing, this system is used for very high-output unit-portion packaging of fluids such as mustard, ketchup, or mayonnaise and hot filling of high acid fluids such as tomato sauce. With chemical web sterilization, vertical form/fill/seal constitutes the principle for aseptic packaging of liquid food products such as milk and juices in brick- or block-shaped paperboard–aluminum foil–polyolefin laminations.

Horizontal form/fill/seal packaging equipment also operates from roll-stock flexible packaging material. With continuous motion, multiple filling heads in parallel with the forming, speeds of up to 700 packages per minute may be achieved, particularly for smaller-type pouches such as unit-portion sugar packets. The horizontal packaging systems are employed to package dehydrated soups, powdered drink mixes, and dry condiment and seasoning mixes, but speeds in these larger sizes are only in the range of 60–120 packages per minute, generally in intermittent motion mode.

Stand-up pouches with an added section of flexible web on the base, and gusseted sides, have been available since the 1960s, but only in the 1990s did they gain importance. Environmentalist pressures have suggested stand-up pouches as substitutes for higher-mass semirigid plastic bottles for liquids or for folding

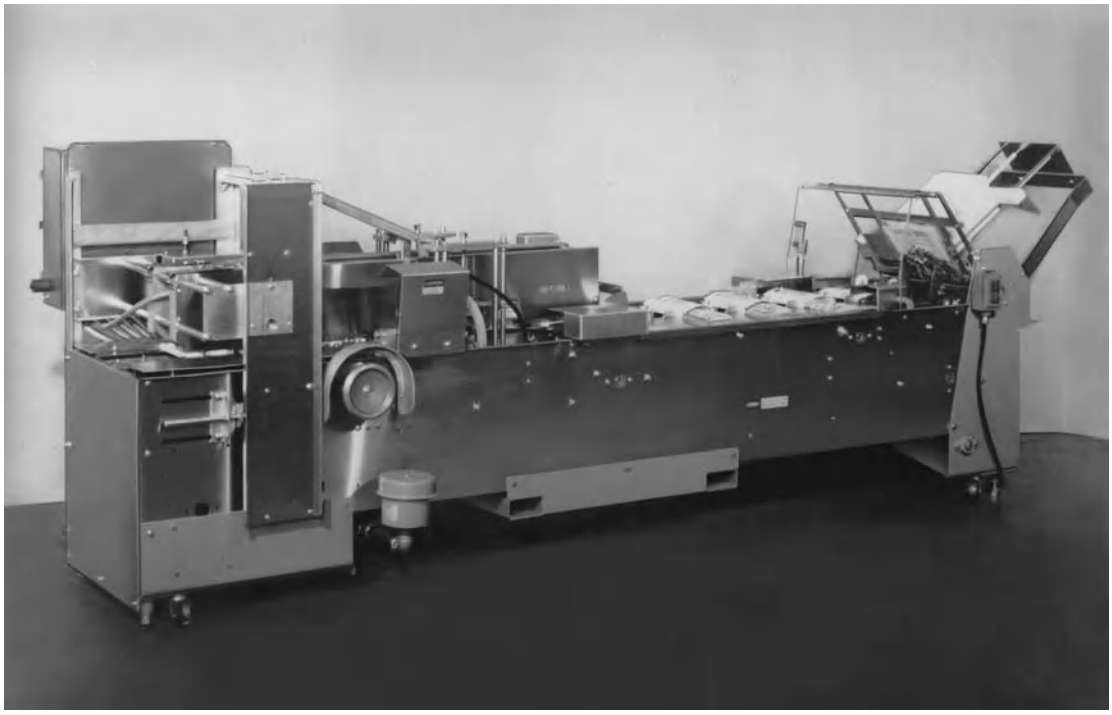

Fig. 47. End-load-style cartoning equipment using preglued sleeve style blanks. (*Kliklok Corp.*)

cartons for dry foods. Most stand-up pouches are prefabricated, with a separate web used to form the bottom gusset. Some, however, are formed on modified horizontal form/fill/seal equipment by incorporating a W-shaped fold in the base to effect the stand-up feature. Although they are accepted and used in Japan for liquid foods and in Europe for liquid household chemicals, their application in the United States has been confined to dry products such as fruit and meal-item mixes.

Fillers. Filling components are designed and engineered for specific categories (dry granular, dry solid, liquid, and viscous fluid) of food products. No single filler design is capable of handling the full range of food products. Dry-food-product fillers may have speeds of over 1000 packages per minute, provided the packages are very small size. Typical filling speeds for liquids for cans and bottles range from 60 per minute for low-speed juice fillers, through 600–900 per minute for individual unit-portion cans or bottles of juices, up to 1500 per minute for beer or soft drinks with equipment now operating at up to 2000 per minute for beer cans. Filling speeds may be increased by multiplying the number of filling heads, counting stations, weighing stations, and so forth.

Basic types of filling machines for dry products include net weighers that weigh the product prior to discharging into the package. Electronic scales operate with significantly greater accuracy and precision than older mechanical equipment (**Fig. 48**). Computerized fillers select from a multiplicity of small preweighed shots, and calculate the specific individual shots to be summed to the whole that becomes the package fill.

Gross weighing devices scale the product into the package, and employ the gross weight of the product and package to determine the quantity to be filled. Volumetric fillers are used for both liquid and dry products. The product, which has a predetermined density, is placed into a fixed volume container above the open-top package. When the receptacle has been filled, the contents are discharged into the package. Metering fillers establish a standard rate of product

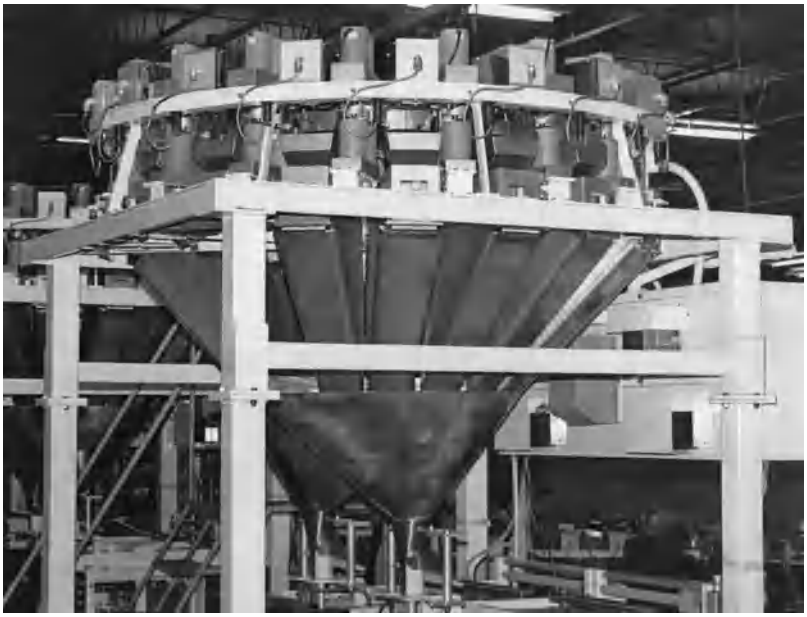

Fig. 48. Computer-controlled multiple-scale weighing device for snack foods used on vertical form/fill/seal flexible packaging machine. (*Kliklok Co.*)

flow for filling into the preformed package at specific time intervals. Auger-type fillers rotate a specified number of times to force the powdered or granular product, such as instant coffee or tea, down through an orifice and deliver by gravity the measured amount of product. Other fillers may count the number of units (such as tablets) being placed in the package.

Bottles are filled with liquid by gravity, gravity vacuum, or pressure gravity. In gravity filling, liquid flows directly by gravity from an overhead surge tank through the filling valves with the nozzle cut off at a predetermined level within the bottle. In gravity vacuum, the entire bottle neck is in a vacuum system, and the force of gravity is supplemented by mechanical vacuum drawn on the bottle interior during filling. Pressure gravity systems employ pumps to increase liquid pressure, especially for filling more viscous liquids. In vacuum filling, air is evacuated from the bottle to assist gravity, and atmospheric pressure above forces liquid from a reservoir.

Filler types for viscous fluid products are either straight-line or rotary-plunger. In either case, cylinders and reciprocating pistons are employed to accurately measure and force the product under pressure into the containers below. Positive-displacement fillers are employed for viscous products such as cottage cheese, ice cream, and juice concentrate; for filling bottles, cans, or paperboard containers with liquids such as milk and juice; and for aseptic filling of liquids in horizontal machines. A vertical form/fill/seal liquid packaging system involves a tube of preformed paperboard lamination packaging material wrapped around a liquid already present. The package is sealed through the liquid to deliver a pouch that is converted in-line into a brick shape.

Closing and sealing. Bags or pouches incorporating interior thermoplastic materials are usually sealed by the application of heat and pressure. Heat sealers include roller, band, bar, high-frequency, and impulse sealing. Induction heating requires the presence of metal in the structure to achieve sealing temperature. Most flexible materials and paperboard-foil-polyolefin laminations are closed mechanically and heat-sealed by fusing the polyolefin coatings to each other under pressure.

Paperboard cartons are closed by mechanical locking, wet glue, hot-melt adhesive, or occasionally by fusing the interior coating to itself. Metal cans are sealed almost exclusively by rotary double seamers that apply rigid metal tops, and mechanically clinch the top closure to the can body over a flange.

Can, glass bottle, and jar closers and sealers are usually off-line from the fillers and may be rotary or straight-line. Automatic capping and closing machines are employed to unscramble or align lug-type caps, closures, side-seam closures, and continuous-thread caps.

Labeling. Food cans are usually labeled in the horizontal attitude by rolling them over stacks of paper labels which have had adhesive applied. In one special method, roll-stock paper labels have been applied to cans at very high speed with die cutting between each can.

Many beverage and food bottles are prelabeled by using preprinted foamed polystyrene applied in the glass plant or by imprinting the bottles with color. Other bottles and jars require paper or other composition, for example, polypropylene labels which may be full wraparound or partial labels or partial-plus-neck labels. Continuous-motion labeling can be matched to the speed of a bottling line, that is, well in excess of 1000 per minute. Pre-die-cut labels are removed from a stack, adhesive is applied, and the label is wiped on to the bottle which is passing by at high speed with axial rotation.

Pressure-sensitive labels with preapplied adhesive are applied either mechanically or by air blowing from roll-fed stock onto the packages.

Bag-in-box. Many food products are packed in bulk bag-in-box for industrial and commercial applications. Bag-in-box had been used for many years for commercial chilled milk packaging, and this led to development of bag-in-box packaging for dry products such as nuts, fruit, frozen fruits and candy, aseptic packaging for juices and juice purees, packaging of wine, and hot filling. The bag is usually prefabricated and inserted into a corrugated fiberboard case with a pre-die-cut opening for the valve through which the product may be filled. Hot-fill pouches are generally prefabricated, filled through an open top, and closed by heat sealing.

For products requiring further protection, such as fruit purees, the bag is fabricated from a barrier material. Metallized polyester or aluminum foil are now employed to effect the oxygen barriers demanded by the contents.

Single-wind paperboard cans. The single-wind convolute paperboard can is designed for in-plant package fabrication to produce cylindrical containers capable of containing dry granular food products. The systems begin with preprinted flat blanks, thus eliminating the need to introduce preformed containers and saving on space volume. Equipment provides the means to form the containers in-line with the filling operations.

Overwrapping systems. Wrapping machines for flexible materials may be employed directly on the product or may be used to wrap other primary packaging for what is perceived by the packager or consumer to be added protection. Direct wrapping in flexible materials is common for products such as soft baked goods, fresh meats, and cheeses.

Wrapping machines may employ either reciprocating or continuous motion. In one variation, the product may be inserted into a flight section which is then elevated to the level of the precut flexible material. The material is sealed using the product as a mandrel. Typical speeds for wrapping machines vary from 60 per minute for double-point end foldup to up to 300–600 for horizontal fin seal form/fill/seal wrapping that is used for unit-portion cookies and crackers. Most overwrapping is horizontal form/fill/seal.

Thermoform/fill/seal operations employing two webs—one a semirigid or flexible plastic material and the other flexible—are widely used for hermetically packaging processed meat, cheese, jams, jellies, and liquid dairy products. In horizontal thermoform/fill/seal packaging systems, the semirigid or flexible base plastic web is first heated to soften the material and then pressure-formed into a cavity. The web containing the cavity is moved to a filling station where the product is inserted. The second flexible web is brought down over the filled cavity, and the web is heat-sealed to flanges surrounding the cavity. Cavities for processed meats and cheeses may be evacuated by vacuum or gas-flushed to extend the refrigerated shelf lives of these products.

Secondary packaging equipment. This type of equipment is designed to accumulate a number of primary packages to ensure that they are protected in the distribution cycle. Secondary packaging equipment permits handling a multiplicity of primary packages as a single unit, and thus increases the efficiency of ditributing packages. It also permits the consumer to purchase more than one primary package at a time. Usually, secondary packaging is a module integrated downstream from primary packaging.

Multiple packaging of bottles may range from the relatively simple drop-into-paperboard basket carriers to the type in which the semirigid plastic or glass bottles are collated and placed into a unit that wraps a paperboard blank around the bottles and locks or glues it. Multiple packaging of bottles may also be accomplished by collating the bottles into groupings of four, six, or eight and bringing an all-plastic clip over the top to unitize the packages.

Cans may be multipacked by aligning them in rows of two or three and bringing a series of polyolefin ring carriers over the top and parallel with the movement of the cans. At appropriate intervals, the plastic ring carriers are stretched over the top of the chimes (end seams) of the cans to capture a grouping of four, six, or eight cans and thus form a multiple.

Paperboard cartons are employed to multipack 12, 18, 20, 24, and so forth cans or bottles. The most widely employed 12/24 multiple packers use preformed paperboard tubes. The tube is removed from the magazine stack and erected in a continuous motion. Groupings of 6 or 12 are pushed in from each side to form 12 or 24, respectively. Alternatively, a flat paperboard blank may be wrapped around a preformed grouping of 12, 24, and so forth cans and then either locked or glued into place. One system groups 12 bottles and wraps and glues a flat paperboard blank around the grouping.

Printed shrink-film wrapping has increased in importance for multipacking. In continuous motion, cans are grouped into 6, 12, and so forth counts, and a sheet of printed film is cut from a web, conveyed to the grouping, and wrapped over the top. A hot-air tunnel shrinks the film snugly to the unitized cans.

Tertiary packaging equipment. Most tertiary packaging in the United States is in corrugated fiberboard cases or trays. Loading of primary or secondary packages may be by dropping into the top or side loading of preformed units, or wrapping flat blanks around a collated block of primary or secondary packages, all followed by glue closing. Shrink-film bundling used imprinted shrink film to unitize trays of 24, and so forth, cans, usually on continuous motion equipment. Reciprocating-motion shrink-film bundling machines operate at speeds of up to 30 cases per minute. There are shrink-packaging machines which operate at speeds of up to 90 cases per minute to collate and wrap groupings of cans or bottles in polyethylene shrink film.

Stretch-film palletization of cases has become very significant. After tertiary packaging, cases are generally palletized for unit distribution. The film binds the cases tightly without further damage and permits relatively safe distribution of full pallet loads of food products. *See* FOOD; FOOD MANUFACTURING.

Aaron L. Brody

Aseptic Packaging

Aseptic processing and packaging is employed in the food and drug industries to package pathogen-free products that will not spoil when stored on a dry shelf. Aseptic processing is the treatment used on the product to eliminate pathogens and other microorganisms capable of causing product spoilage during ambient storage. A packaged product that does not contain viable pathogens or other

Fig. 49. Aseptic packaging machine for forming brick-shaped packages. (*Tetra Pak, Inc.*)

microorganisms capable of product spoilage is considered commercially sterile. Aseptic packaging is the process of filling sterile product into sterile containers under conditions where viable microorganisms are prevented from recontaminating the product. Commercial aseptic packaging machines are sophisticated systems capable of sterilizing the container material, forming the container, filling the product with consistency in fill weight or volume, maintaining sterile conditions while filling, and sealing the package securely before it leaves the sterile section of the machine. These operations must be carried out at high speed for the process to be commercially viable.

Machinery. Aseptic packaging machines vary in operation and type of package produced. Some receive containers that have been sterilized at another location. These machines must permit transfer of the sterile containers into the filling section without recontamination. The inability to detect microbial incursion at any time during the transfer and filling process is a serious limitation. Thus, when packaging low-acid food products, which support the growth of the deadly microorganism *Clostridium botulinum*, packaging material sterilization and package forming, filling, and sealing, are integral parts of the same machine. Most aseptic packaging machines are preprogrammed to go through a presterilization process before packaging operations can begin, and operating conditions are monitored by using appropriate sensors for temperature, pressure in the filling section, package sterilant temperature, and flow. An automatic shutoff system, which would terminate filling when conditions indicate probable loss of commercial sterility, is an important component of aseptic packaging machines.

Sterilization. When the packaging material is heat resistant, such as metal and glass, the packages are sterilized in-line with superheated steam at 300°F (150°C) or higher. Glass must be heat sterilized with care by prolonging time to reach sterilization temperature and by cooling down slowly before filling to prevent breakage. Plastic packaging materials are sterilized by dipping in, or by spraying into the formed package, 15–35% hydrogen peroxide solution. Hydrogen peroxide is removed from the sterilized package by drying with radiant heating elements and a stream of hot air. Regulations by the U.S. Food and Drug Administration (FDA) limit the level of residual hydrogen peroxide in the product to 0.1 part per million. Hydrogen peroxide removed from the packaging materials after sterilization can create very uncomfortable conditions near aseptic packaging machinery. Machines must prevent leakage of the air–hydrogen peroxide mixtures into the surroundings. Hydrogen peroxide is removed from air by condensation before the air is vented out through a stack. Other sterilants, such as ethanol and high-intensity ultraviolet light, have been suggested and are used to a limited extent in other countries. However, in the United States hydrogen peroxide is the only chemical sterilant approved by the FDA for

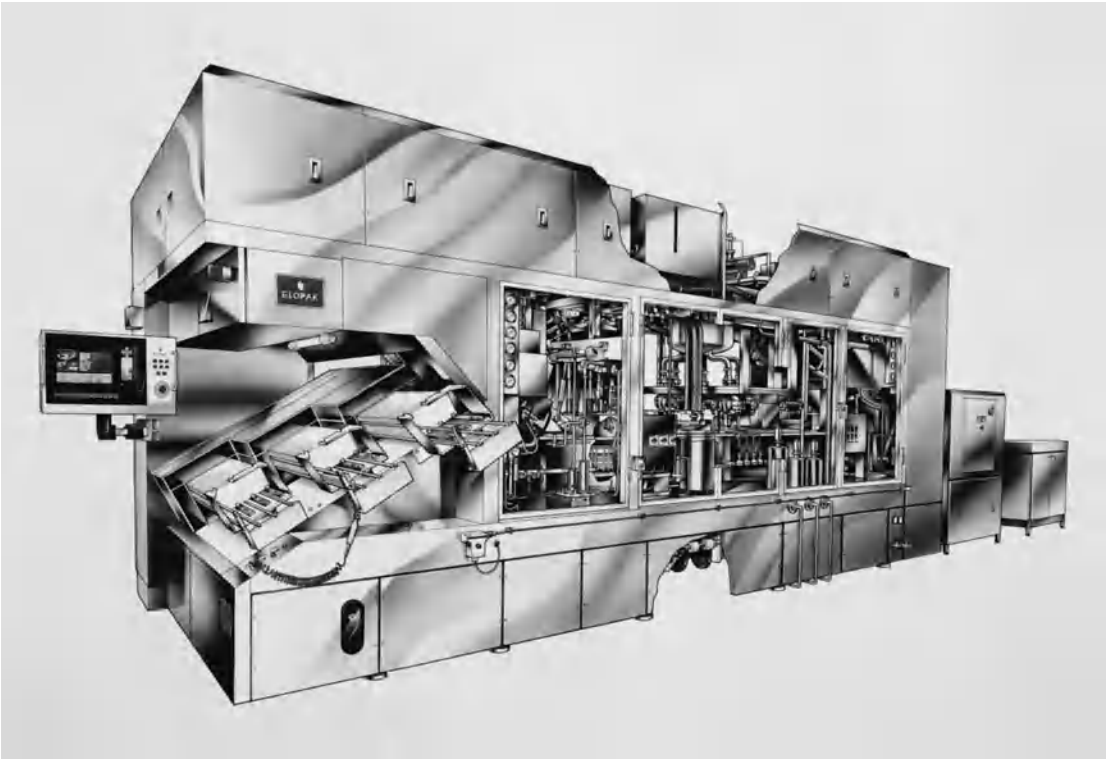

Fig. 50. Aseptic packaging machine for forming gable-top containers. (*Pure-Pak, Inc.*)

use in aseptic packaging machines. Packaging materials may be presterilized in a separate location with gamma irradiation, overwrapped to maintain sterility during transportation and storage, and transferred out of the overwrap by using an elaborate device to prevent recontamination.

Filling. Filling must be done in a presterilized filling station. Machine sterilization is accomplished by spraying hydrogen peroxide solution inside this section of the machine for a specified time, followed by drying with hot sterile air. Some machines are presterilized with superheated steam or air at very high temperature. Sterile air continuously admitted into the machine to maintain a positive pressure maintains sterile conditions during the operation. Filters capable of screening out microorganisms are used to produce the sterile air.

In the pharmaceutical industry, sterilization of clear injectable solutions is accomplished by filtration. Aseptic packaging is accomplished by using presterilized containers, and filling is done in cleanroom environments. Filling is done in an enclosed area where a stream of recirculating laminar-flow filtered air forms a screen preventing entry of microorganisms into the package during the filling process. High-efficiency particulate filters screen out microorganisms from the recirculating air. Proper worker attire and an overall low microbial load in the air near the filling equipment reduces the risk of microbial contamination of the product. Pharmaceutical preparations generally do not support microbial growth; therefore chance contamination of one viable microorganism in a package would not be a problem. However, contamination by just one viable pathogenic microorganism in a food that supports its growth will have disastrous results.

Testing. Aseptic packaging machines must be tested for their ability to maintain sterile filling conditions before they are placed into commercial operation. Typical tests include installing test strips containing indicator organisms inside the filling station to test the ability of the system to sterilize this section; and inoculating packaging material with a test organism and aseptically packaging the product to determine if the test organisms survived and spoil the product. Finally, the aseptic packaging system is used to package the food product at a speed used in commercial operations over a period of several hours. The packages are then incubated over several weeks and evaluated for signs of spoilage. Spoiled products are evidence of aseptic packaging failure.

Package designs. Aseptically packaged products are commercially available in brick-shaped packages that consist of a laminate of layers of polyethylene, foil, paper, and polyethylene; multilayered plastic film laminates in a protective cardboard box (bag-in-box); thermoformed plastic cups; and pouches of plastic film. A number of machines have been developed to produce special packages. In one type (**Fig. 49**), the machine receives roll stock of packaging material that consists of a laminate of paper, aluminum foil, and plastic and forms this into brick-shaped packages. In another type (**Fig. 50**), the machine receives paper/foil laminate blanks and forms these into gable-top containers. Packages may be in individual portion size, institutional size for food service applications, or large sizes for industrial reprocessing.

Romeo T. Toledo

Bibliography. American Society of Heating, Refrigerating, and Air Conditioning Engineers (ASHRAE), *2002 ASHRAE Handbook: Refrigeration*, 2002; N. A. Anderson, *Instrumentation for Process Measurement and Control*, 3d ed., 1997; G. V. Barbosa-Cánovas et al., *Preservation of Foods with Pulsed Electric Fields*, 1999; P. A. Belter, E. L. Cussler, and W. S. Hu, *Bioseparation: Downstream Processing for Biotechnology*, 1988; A. L. Brody, *Controlled/Modified Atmosphere/Vacuum Packaging of Foods*, 1989; A. L. Brody and K. S. Marsh, *The Wiley Encyclopedia of Packaging Technology*, 2d ed., 1997; D. Claxton, L. Svarovsky, and M. Thew (eds.), *Hydrocyclones '96*, 1996; P. Cooper (ed.), *Pumping Machinery*, 1993; C. Doona et al. (eds.), *High Pressure Processing of Foods*, 2006; D. L. Downing (ed.), *A Complete Course in Canning and Related Processes: Fundamental Information on Canning*, 13th ed., 1996; P. F. Fellows, *Food Processing Technology*, 2d ed., 2000; F. J. Francis (ed.), *Wiley Encyclopedia of Food Science and Technology*, 4 vols., 1999; C. J. Geankoplis, *Transport Processes and Unit Operations*, 4th ed., 2003; S. A. Goldblith and W. W. Rothmayr, *Freeze Drying and Advanced Food Technology*, 1975; G. W. Gould (ed.), *New Methods of Food Preservation*, 1995; R. Guy, *Extrusion Cooking: Technology and Applications*, 2001; B. Hallstorm, C. Skjodebrand, and C. Tragardh, *Heat Transfer and Food Products*, 1988; R. Hayashi (ed.), *Trends in High Pressure Bioscience and Biotechnology*, 2002; D. R. Heldman and D. B. Lund (eds.), *Handbook of Food Engineering*, 2d ed., 2006; Y. H. Hui (ed.), *Dairy Science and Technology Handbook*, 3 vols., 1994; Y. H. Hui, *Food Plant Sanitation*, 2002; A. Ibarz, *Unit Operations in Food Engineering*, 2002; K. Kulp and J. G. Ponte, Jr. (eds.), *Handbook of Cereal Science and Technology*, 2d ed., 2000; W. L. McCabe, J. C. Smith, and P. Harriot, *Unit Operations in Chemical Engineering*, 7th ed., 2004; K. R. Osborne and W. Jenkins, *Plastic Films, Technology and Packaging Applications*, 1992; R. H. Perry and D. W. Green (eds.), *Perry's Chemical Engineer's Handbook*, 7th ed., 1997; M. N. Riaz, *Extruders in Food Applications*, 2000; S. K. Sastry and B. D. Cornelius, *Aseptic Processing of Foods Containing Solid Particulates*, 2002; G. Robertson, *Food Packaging: Principles and Practice*, 2d ed., 2005; H. G. Schwartzberg and M. A. Rao (eds.), *Biotechnology and Food Process Engineering*, 1990; P. G. Smith, *Introduction to Food Process Engineering*, 2003; A. J. Stapanoff, *Centrifugal and Axial Flow Pumps: Theory, Design, and Application*, 2d ed., 1993; R. T. Toledo, *Fundamentals of Food Process Engineering*, 3d ed., 2006; P. Walstra et al., *Dairy Science and Technology*, 2d ed., 2005.

Food fermentation

Production of food with the aid of microorganisms, which may be yeasts, molds, or bacteria. Well-known examples of such foods are bread, cheese, beer, wine, vinegar, some sausages, sauerkraut, yogurt, and cultured milk. This article deals with fermented foods produced and consumed in large amounts by millions of people in Africa and Asia. In Japan, for example, soy sauce (shoyu) is consumed at the rate of 11 quarts (10 liters) per capita per year, and annual production is close to 330,900 gallons (1,252,600 liters). Production of miso and natto is in the range of 625,900 and 174,200 tons (567,800 and 158,000 metric tons), respectively.

Fermented commodities include cereals such as wheat, rice, sorghum, corn; legumes, peanuts, soybeans, pulses; red meat, sausages, pork; milk; fish, shellfish; and plant juices. Fermented food is used as a staple, such as tempeh in Indonesia; as a condiment like soy sauce; as a coloring agent like the red rice used widely in the Far East; and as the breadlike product produced in India and the Near East. The food may be a liquid like the fermented fish sauces of Indochina; a paste such as miso made from rice and soybeans in Japan; or a solid, like tempeh, made by the mold fermentation of soybeans. *See* FERMENTATION.

Production. Fermented foods were developed in prehistoric times and are widely produced all over the world. Their continued use results from the mild biochemical changes produced by enzymes. For example, starches are broken down to the sugars required for a second fermentation. In soybean fermentations, protein yields amino acids, and lipases act on the oil. Fermentations produce vitamins and simple chemicals such as lactic acid; improve digestibility, odor, and flavor; change physical properties; and prevent spoilage. *See* PROTEIN METABOLISM.

Among disadvantages of fermentation is the possibility of food poisoning due to bacteria or molds. Also, some fermentations are slow (taking up to 8–12 months), and some product is always lost. *See* FOOD MANUFACTURING.

Microorganisms. Among the molds used in food fermentations are *Aspergillus*, *Rhizopus*, *Penicillium*, *Neurospora*, *Actinomucor*, *Mucor*, *Amylomyces*, and *Monascus*. The yeasts include species of *Saccharomycopsis, Zygosaccharomyces*, and *Candida*. The bacteria are *Bacillus* and lactic acid bacteria, including *Lactobacillus*, *Streptococcus*, *Pediococcus*, and *Leuconostoc. See* YEAST.

Inoculum. The inoculum or fermentation starter may be of four types. The substrate may be moistened, heat-sterilized, and inoculated with a single organism. In the commercial production of red rice, a single species of *Monascus* is used to inoculate rice, which eventually takes on a reddish-purple hue that imparts color to meat and wine. In a second type, more than one strain of a single species is used. Koji is the mold preparation containing more than one strain of *Aspergillus oryzae* that produces enzymes on rice, wheat, or soybeans. It is used in a second fermentation to make various types of miso and shoyu. A third type of inoculum contains more than one species of microorganism. Ragi is an Indonesian culture used in various fermentations of rice or cassava. Known by different names in other countries, the dry powder contains *Amylomyces*, *Mucor*, *Rhizopus*, yeast, and bacteria. Finally, some substrates contain a complex inoculum in which many different microorganisms of unknown identity are present. An example is nuocmam, a fish sauce. Small fish are packed with salt in earthen containers and allowed to ferment at room temperature for several months. The resulting liquid is removed and used as a condiment much like soy sauce. The salt prevents the development of food-poisoning microorganisms.

Applications. There are countless varieties of food fermentations in the world, sometimes known under different names in different countries. Of the following two examples, one is quick and simple, and the other slow and complicated.

Tempeh. Tempeh is made by a simple mold fermentation, which is being introduced in the West. Only soybeans are used as substrate. Dry beans are washed and soaked overnight at about 77°F (25°C). The seed coats are removed, and the beans are boiled for 30 min, drained, cooled, and briefly dried. After inoculation with spores of a *Rhizopus* strain, the beans are incubated for 20–24 h at 88–89°F (31–32°C) in plastic bags perforated for aeration. Acid may be added to control bacterial contamination. The fermentation produces a solid cake in which the beans are bound together with mycelium. The product, after being sliced into thin pieces and fried in deep fat, is brown and has a nut flavor.

Miso. Miso fermentation is a complicated two-step process. The product is used in seasoning and as a base for soups in the Orient. Polished rice is washed, soaked for 16 h, drained, steamed, cooled, and inoculated with the mold *Aspergillus oryzae*. After about 2 days at 82°F (28°C) with frequent turning, the moldy rice is called koji. Soybeans that have been crushed, soaked in water, drained, steamed, and cooled are mixed with koji, followed by the addition of 2 kg (4.4 lb) of salt for 15 kg (33.1 lb) of the mixture and an inoculum of special yeast and bacterial cultures. The mixture is allowed to ferment for 3 months or more and is mashed to a paste for sale. There are many kinds of miso, depending on the amounts of substrates and salt and the time and temperature of fermentation. *See* SOYBEAN. C. W. Hesseltine

Bibliography. C. W. Bamforth, *Food, Fermentation and Micro-organisms*, 2005; L. R. Beuchat, *Food and Beverage Mycology*, 2d ed., 1987; Y. H. Hui, *Handbook of Food and Beverage Fermentation Technology*, 2004; G. G. Khachatourians and D. K. Arora (eds.), *Applied Mycology and Biotechnology:* vol. 2: *Agriculture and Food Production*, 2002; W. Shurtleff and A. Aoygai, *The Book of Miso*, rev. ed., 2001; K. Steinkraus (ed.), *Handbook of Indigenous Fermented Foods*, 2d ed., 1995; M. J. Waites et al. (eds.), *Industrial Microbiology: An Introduction*, 2001.

Food manufacturing

A total sequence of food operations, including the growth and selection of raw materials, harvesting, processing, preservation, and distribution. In general, the aim of all food manufacturing operations is to extend the availability of seasonal crops to year-round use.

Many of the technological operations involved in food manufacturing have a long history of use, beginning with the invention of fire and the cooking of foods, through the first large-scale food production industries in ancient Egypt and Rome, where bakeries provided consumers with bread. This was the beginning of manufactured convenience foods and, in addition, relieved consumers of the burden of making their own food products.

The products of food manufacturing differ from traditional foods of plant or animal origin which have undergone minimal treatment. For example, the quality of apples sold in the winter can be maintained, through the use of controlled-atmosphere storage, which retards the ripening process by controlling the levels of oxygen, nitrogen, and carbon dioxide in the atmosphere of the storage facility. Atmosphere control is also used to hasten ripening so that fruits may be harvested in the unripe stage for ease of handling and then ripened rapidly in storage. In other cases, the package itself allows the diffusion of only certain atmospheric gases and thus maintains quality. However, minimal treatment is not always directly correlated to quality and nutritional value. Foods that are harvested at optimum maturity, rushed to the manufacturing plant, cleaned, washed, cut, sorted, and processed by such means as rapid freezing are often higher in quality and nutrition than foods picked fresh and then stored under less than optimum conditions. Also, there are certain foods that cannot be maintained in a state close to the raw product. Tomatoes, for example, are not amenable to freezing or long periods of storage, yet they represent a major crop that must be harvested within a short period of time. If it were not for processing, tomatoes would not be available for human consumption throughout the year. Therefore, such products as heat-processed sauces, pastes, and stewed tomatoes have been developed. Other food products are even further removed from the raw product: oil is being produced from seed; and plant proteins are being used as extenders or substitutes for meat, as additives for nutritious beverages, and as bases for many formulated foods.

There are many other forms of food preservation representing both ancient and space-age technologies. The ancient operation of sundrying was first employed when it was realized that dried fruits remained wholesome and edible for long periods of time. Today, with the additional knowledge that drying, evaporation, and concentration all reduce the water activity or increase the osmotic pressure of a food to the point where bacteria will not grow, this technology is used for sophisticated products such as powdered milk and freeze-dried mushrooms. Food additives, such as salt, sugar, and other solutes, which reduce the water activity or increase the osmotic pressure, and acids, which inhibit bacterial growth, also achieve the preservation effect. Many food additives are natural in origin, and their preservative effect was noted in nature prior to their use as food additives. Freezing, heat sterilization (canning), pasteurization, fermentation, baking, and meat curing are other well-known forms of preservation. Irradiation processes for food have also been developed, and low-level irradiation has been approved in the United States by the Food and Drug Administration (FDA).

Food manufacturing, however, is not solely involved with the preservation of food but is also concerned with the production of high-quality, appealing, and wholesome food. To fulfill these goals, five broad categories of food additives are often used: flavors, coloring agents, preservatives, texturizing agents, and miscellaneous. The last category includes a variety of substances that may retain moisture, control acidity, act as leavening agents, or provide nutrients such as vitamins and minerals.

The final operation in the manufacturing process is that of packaging, which is governed by the physicochemical attributes of the food, the preservation process involved, the gaseous permeability desired, the conditions under which the product is to be stored, the desirability of viewing the product through a clear film or glass, and the expense.

Historically, metal and glass have been used to package heat-processed foods; more flexible films are used for foods which undergo less vigorous treatment. Adoption of the regulation allowing the use of hydrogen peroxide as a package sterilant has permitted the use of nonrigid flexible packages for heat-sterilized foods (aseptic packaging). This type of packaging is very cost-effective. *See* FOOD ENGINEERING; FOOD PRESERVATION. Fergus M. Clydesdale

Food Additives

Generally, a food additive is any substance used in small amounts in or on food to achieve a particular result. The terms food, ingredient, and additive are neither clearly bounded nor exclusive, and use determines how a substance is termed. Sugar in a sugar mint is a food; in a cake, it is an ingredient; in bread dough, it is a yeast food, an additive.

Some additives are direct (or intentional); that is, they are added to food to restore, enhance, or add desirable characteristics such as safety, stability, appeal, and economy, or to improve the effectiveness of processing. Others, called indirect additives, occur in food, usually in trace amounts, not because they confer any desirable property to the food itself, but because they inevitably are left as a result of migration from packaging materials, or from use in agriculture, storage, or transportation (**Table 1**).Certain direct additives may be regarded as second-order additives if they performed a useful function in some ingredient prior to its addition to a final food but lost that effective role in the final food. Thus, anticaking agents are often used in salt. If the salt

TABLE 1. Approximate numbers of direct and indirect food additives

Additive		Number
Direct		
Prior sanctioned		≈100
GRAS only		1558
Substances that are both GRAS and regulated food additives		868
Regulated food additives		312
Colors and coloring adjuvants		58
	Subtotal	2896
Indirect		
Packaging, GRAS		110
Regulated		≈3000
Pesticides		300*
	Subtotal	3410
	Total	6306†

*Active ingredients only; there are thousands of different commercial formulations.
†An overestimate; an accurate count is impractical because of duplicate listings and difficulties of classification.

was used in making mayonnaise, the anticaking agent would no longer be useful and would then be a secondary additive.

Types. The daily average composition of the American diet consists of a number of categories (**Fig. 1**). There is a certain amount of overlap, because many substances fall in more than one category. In the food manufacturing industry, dietary consumption is referred to as disappearance. However, actual consumption is less than disappearance because of loss, waste, and nonconsumptive uses such as fermentation and processing. The category "other natural constituents" (Fig. 1) is numerically enormous—probably numbering in the hundreds of thousands—and includes essential oils, resins, alkaloids, nucleotides, phospholipids, enzymes, pigments, and hormones. Some of these substances, particularly certain essential oils and their components, are used in food manufacturing as direct additives. Most of the food additives comprise major ingredients that are generally recognized as safe (GRAS), including substances such as yeast, baking powder ingredients, and common spices. Many are carbohydrates (such as invert sugar), fats (mono- and diglycerides), or proteins (albumin). The distribution of use of direct additives (**Fig. 2**) ranges from sucrose at about 44 kg (97 lb) per person per year to some flavors and micronutrients at less than 0.1 nanogram per person per year. The median direct additive is consumed at about a milligram per person per year.

Technical effects. The purpose of all direct food additives is their physical or other technical effects (**Table 2**).

Other specific technical effects can be defined, most of them related to processing, such as freezing agents, and washing and peeling aids. In addition, there are many substances that serve several technical effects. Ascorbic acid (vitamin C), for example, is a nutrient supplement, an antioxidant, or a curing adjuvant, depending on the conditions of use. Also, the boundaries between the technical

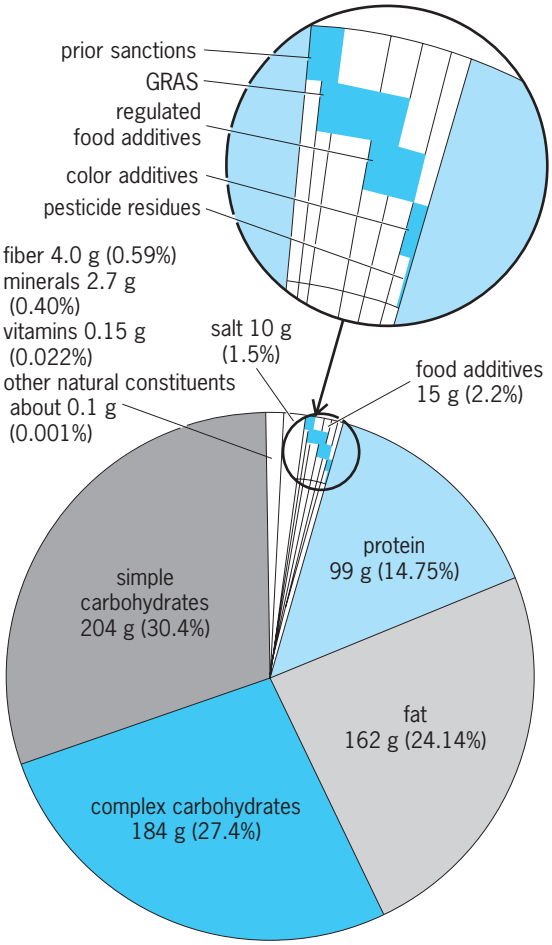

Fig. 1. Average composition of the United States diet dry weight per capita per day. Prior sanctioned substances were specifically approved by FDA or USDA for use in food prior to the passage of the Food Additives Amendment in 1958.

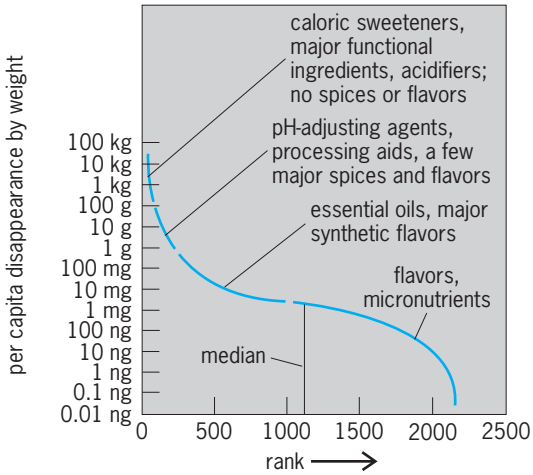

Fig. 2. Rank ordering of food ingredients by per capita annual disappearance.

TABLE 2. Technical effects of direct food additives

Category effect	Agent
Anticaking agents: prevent caking, lumping	Calcium stearate, cornstarch, sodium aluminosilicate
Antimicrobial agents (preservatives): prevent growth of microorganisms and consequent spoilage	Sodium benzoate, calcium propionate, potassium sorbate
Antioxidants: retard browning and rancidity	BHA, BHT, ascorbic acid (vitamin C), ethoxyquin
Colors and coloring adjuncts: impart, preserve, or enhance color	FD&C Blue No. 1 and other certified synthetic colors, β-carotene and other uncertified synthetic colors, beet powder, caramel and other natural colors
Curing and pickling agents: impart flavor and color to cured meats, discourage spoilage, prevent bacterial toxin formation	Sodium nitrite, salt, sodium metaphosphate
Dough conditioners, strengtheners: make dough easier handle, bread more appealing to consume	Potassium bromate, acetone peroxide, glyceryl monostearate
Drying agents: absorb moisture	Specially dried cornstarch, anhydrous dextrose
Emulsifiers: prevent separation of oil and water mixtures	Mono- and diglycerides, lecithin, sorbitan monostearate
Enzymes: promote reactions	Rennet for producing cheese curd, papain for tenderizing meat, pectinase for clarifying beverages
Fermentation and malting aids: promote rapid and proper fermentation	Gibberellic acid, potassium bromate
Firming agents: enhance crispness of plant tissues	Calcium salts, aluminum sulfate, calcium lactobionate
Flavor enhancers	Soy sauce, monosodium glutamate (MSG), disodium inosinate, disodium guanylate
Flavoring agents and adjuvants: natural and artificial flavors, fixatives, solvents	Spices, herbs, essential oils, and individual chemicals of both natural and synthetic origin
Flour-treating agents: bleach and mature flour	Acetone peroxide, benzoyl peroxide, potassium bromate
Formulation aids: carriers, binders, fillers, film formers, and so on	Starch, gum acacia, propylene glycol
Fumigants: kill undesirable organisms	Methyl bromide, ethylene oxide, phostoxin
Humectants: prevent drying	Sorbitol, propylene glycol, sodium tripolyphosphate
Leavening agents: impart light, fluffy texture	Yeast, monocalcium phosphate, sodium carbonate, and other baking powder ingredients
Lubricants and release agents	Oleic acid, hydrogenated sperm oil, mineral oil
Masticatory substances for chewing gum	Chicle, rubber, paraffin, glycerol esters of rosin
Nonnutritive sweeteners: sweeten with few or no calories	Saccharin, aspartame, acesulfame-K
Nutrient supplements: enhance or restore nutritional value	All the known essential nutrients such as vitamin A and other vitamins, iron and other trace minerals, amino acids and essential fatty acids
Nutritive sweeteners: sweetness with calories	Dextrose, fructose, sucrose
Oxidizing and reducing agents: remove or modify ingredients or contaminants that would damage quality	Glucose oxidase (enzyme) to destroy glucose in dried egg, so that it will store well; hydrogen peroxide added as a bleaching or antimicrobial agent
pH control agents: reduce or increase or stabilize acidity	Vinegar (acetic acid), sodium bicarbonate, hydrochloric acid
Processing aids: aid in clarifying, filtering, crystallizing, and so on	Charcoal, diatomaceous earth, tannic acid
Propellants, aerating agents, gases: expel a product, induce foaming, exclude oxygen	Nitrous oxide, carbon dioxide, nitrogen
Sequestrants: combine with trace metals to promote stability	Citric acid, ethylenediaminetetraacetic acid (EDTA), phosphoric acid
Solvents, vehicles: dissolve, suspend, or carry	Alcohol, propylene glycol, glycerine
Stabilizers, thickeners: increase viscosity, improve mouthfeel, prevent separation	Starch, modified food starches, carrageenan
Surface-active agents: promote rapid wetting and dispersing, enhance or prevent foam	Sodium lauryl sulfate, lactylic esters of fatty acids, dimethyl polysiloxane
Synergists: disproportionately enhance effects of other additives	Citric acid, tricalcium phosphate, and other phosphates
Texturizers: affect feel or appearance	Sodium bicarbonate, glycerine, corn syrup, modified food starch

effect categories are not always clear; that is, some flour-treating agents are oxidants, emulsifiers may act as surfactants, and some formulating and processing aids may act similarly.

While there are usually several substances available for achieving a particular technical effect, they rarely are fully interchangeable. For each application there is usually a single best choice if the requirements can be defined adequately.

Legal definition and approval for use. The definition of food additive in the Federal Food, Drug, and Cosmetic Act (FD&C Act; 21 U.S. Code 321) is far more complex than the practical definition given above. Section 201 first broadly defines the term to include all components of food, then makes exclusions: "The term 'food additive' means any substance the intended use of which results or may reasonably be expected to result, directly or indirectly, in its becoming a component or otherwise affecting the characteristics of any food . . . if such substance is not generally recognized, among experts qualified by scientific training and experience to evaluate its safety, . . . to be safe under the conditions of its intended use; except that such term does not include . . . a pesticide chemical . . . or . . . a color additive; or any substance used in accordance with a sanction

or approval granted prior to the enactment of this paragraph . . . ; or a new animal drug."

The words "generally recognized . . . to be safe" have been lifted out of the language of the act, paraphrased as "generally recognized as safe," and reduced to the acronym GRAS. Thus, GRAS substances, such as salt or pepper or vanillin (a flavor), though in a practical sense food additives, are excluded from the legal definition of food additive. Also excluded are pesticides and fumigants, colors such as beet juice or FD&C Yellow No. 5, and "prior sanctioned" materials such as sodium nitrite used in curing meats.

In Section 409, the act requires any food additive, so defined after exclusions, to be regarded as "unsafe" unless it is either exempted for investigational use, or is used following a regulation prescribing conditions for its safe use.

The information required and the specific steps involved in approval of a new food additive, in the common sense of the term, depend on the particular legal classification of the substance, although the pattern of information in all cases is generally similar.

For a substance that is legally a food additive, a petition must be filed with the FDA for a regulation permitting its use. Section 409 of the FD&C Act sets out the information required in such a petition. Included are the name, chemical composition and identity, the conditions of proposed use, the technical effect (purpose) of the additive and the amount required to produce that effect, analytical methods for determining the additive in food, and reports of all safety testing. The FDA may also require manufacturing details and samples. Neither the statute nor regulations spell out how much toxicological and related information of safety will be required. The FDA has based its requirements in each instance on an ad hoc determination of the kinds of data required to provide assurance of safe use. Expected intake, the results of previous testing, and accumulated knowledge of the toxicity of similar materials all weigh heavily. These judgments, made over years, have been collected and collated into a compilation for providing flexible general guidance called the "Red Book," published by the U.S. Department of Health and Human Services.

If the firm desiring to use the substance believes that it would be more appropriate to regard the substance as GRAS, it may petition the FDA to affirm that judgment. The data requirements are similar to those for any food additive.

The FD&C Act, however, does not define which experts shall conclude that a substance is of GRAS status. It merely requires that they be "qualified by scientific training and experience to evaluate its safety." Thus, the experts may be drawn from either the public or private sectors. Although a private determination that a substance is generally recognized as safe may seem contradictory, it is conceivable that a firm could reach this conclusion on its own. On a broader basis, the flavor industry has regularly published the judgments of a panel of independent experts on general recognition of safety, and most new flavoring substances have come into use by that route. Again, the data and judgments required are generally similar to those for all food additives.

Color additives are approved by a similar process under Section 706 of the FD&C Act, except that there is no provision for GRAS status. "Prior Sanctions" apply only to certain substances approved for use before 1958 and are no longer granted.

Pesticide residues on food are approved under Section 408 of the act, with generally similar provisions. A pesticide, however, must first be registered under the Federal Insecticide, Fungicide and Rodenticide Act (FIFRA), and the Secretary of Agriculture must also certify the pesticide's usefulness for the purposes proposed.

The Food Additives Amendment and Color Additive Amendments contain the so-called Delaney Clause. This denies regulatory approval to a substance "found to induce cancer when ingested by man or animal. . . . " With the passage of time, analytical methods have become ever more sensitive in finding minute traces of contaminants and impurities. More comprehensive and severe conditions of toxicological testing have raised both questions of possible carcinogenic potential in many materials tested and questions about the relevance of the tests to human safety. These trends have increased debate over providing more flexibility, through either administrative interpretation or legislative change, than the clause now appears to allow. Such new knowledge caused Congress to provide limited exceptions to the clause for animal drugs and nutrients. The clause does not apply to GRAS and prior sanctioned substances. *See* TOXICOLOGY. Richard L. Hall

Food Analogs

The terms food analog, structured food, fabricated food, texturized food, imitation food, and architectural food are all used to describe foods which are considered to have equal or superior characteristics to those of the foods which they are designed to imitate. These enhanced characteristics usually involve sensory properties, better quality control, superior nutritional properties, or manufacturing economies.

Food analogs usually require a high degree of design and engineering because the food product should have unique functional and marketable properties which can be distinguished by the consumer. Food analogs are usually based on soy, and include fabricated meats, seafoods, cheese, vegetables, fruits, and eggs. They have been available since about 1960 as novelty products, but have gained consumer acceptance as meat substitutes for vegetarians and individuals with special dietary restrictions, generally because they are cheaper than meat. In addition, they are used extensively in food service programs, especially to extend hamburger. *See* SOYBEAN.

Technologies. Once the structuring mechanism or process is chosen for food analogs, customized

manufacturing equipment and machinery for high-speed production conditions must be designed and constructed.

Fiber formation. Fiber formation technology is based on the chemistry and technology of proteins, particularly soy proteins. Soybeans are hulled, and the oil is removed by solvent extraction. Soybean meal contains factors that must be inactivated by moist heat. For food uses, the processing may consist of simple heating and grinding of the defatted material as in the preparation of flours and grits, or of further fractionation to increase protein content as in the production of concentrates and isolates. *See* SOLVENT EXTRACTION.

Soybean protein concentrates consist of a minimum of 70% protein plus the polysaccharides, and are made by three methods that render the major proteins insoluble while the low-molecular-weight compounds are removed. One process consists of extraction with aqueous alcohol. A second process involves extraction with dilute acid at pH 4.5 (the isoelectric point and region of insolubility of the major proteins); the acid-leached concentrate is neutralized before drying. In a third process the proteins are rendered insoluble by heat denaturation; the low-molecular-weight constituents are then washed out with water. *See* ISOELECTRIC POINT; PROTEIN.

Protein isolates are the most refined soy proteins available. They are prepared by removing all water-insoluble polysaccharides, as well as water-soluble sugars and other minor constituents. Defatted flakes and flours of high-protein solubility are extracted with dilute alkali at 120–130°F (50–55°C). After the insoluble residue (water-insoluble polysaccharides plus residual protein) is separated by screening, filtering, and centrifuging, the extract is adjusted to pH 4.5 with food-grade acid. When the major proteins are brought to their isoelectric point, they precipitate. This protein curd is filtered or centrifuged from the solubles and washed. The curd may be neutralized and spray-dried.

Isolates are the basic ingredient of the spun-fiber meat analogs. In this process, the soy protein is resolubilized in an alkali aqueous media. This liquid may be extruded through tiny holes and then recoagulated in an acid bath in the form of fibers in a manner similar to making rayon. The fibers can then be spun into ropes, with the final texture approaching the fibrous structure of chicken or beef muscle tissue. The fabricated tissue can be interlaced with fats, food flavorings, and food colors. The products also may be dehydrated, compressed, or otherwise processed. This process was first adapted from the production of synthetic textile fibers, such as rayon and nylon (**Fig. 3**).

Because the manufacturing steps are expensive, the soy meal may be processed instead directly in a heated extruder to form an expanded cooked soy matrix which can be incorporated into other foods.

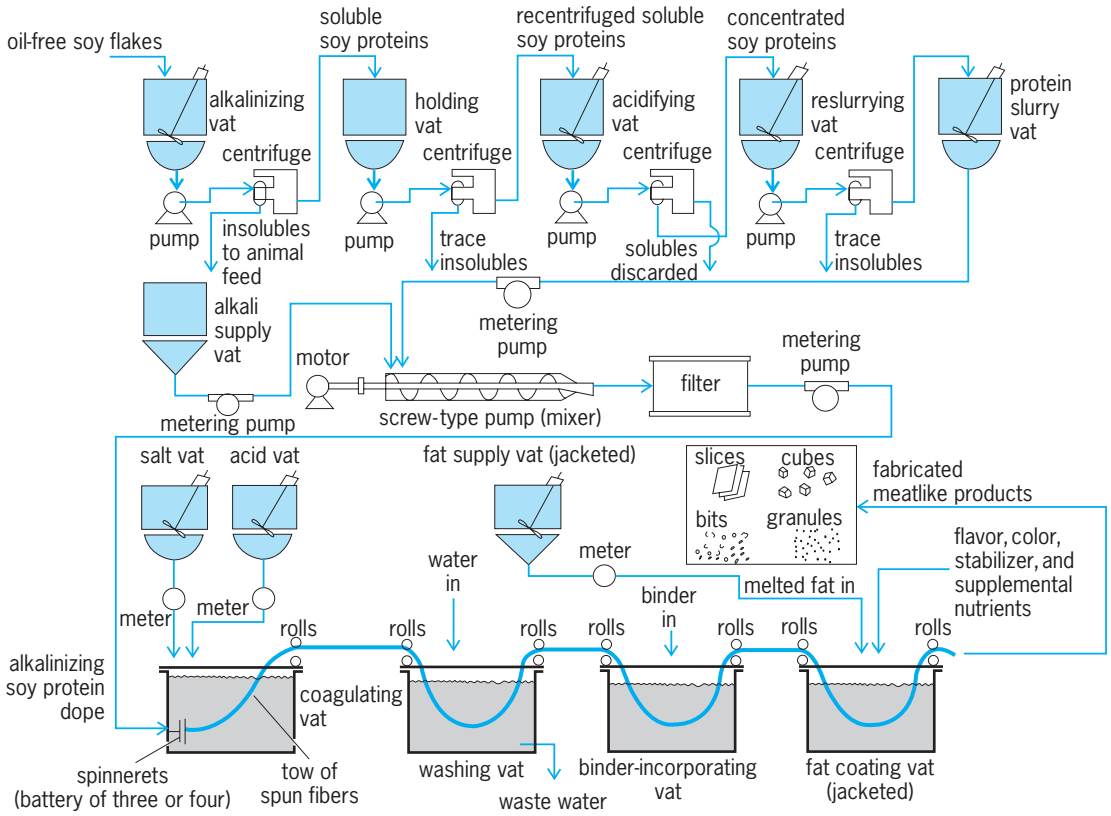

Fig. 3. Flow chart of the fiber-spinning technique used to prepare food analogs. (*After N. R. Lockmiller, Increased utilization of protein in foods, Cereal Sci. Today, 18(3):77–81, 1973*)

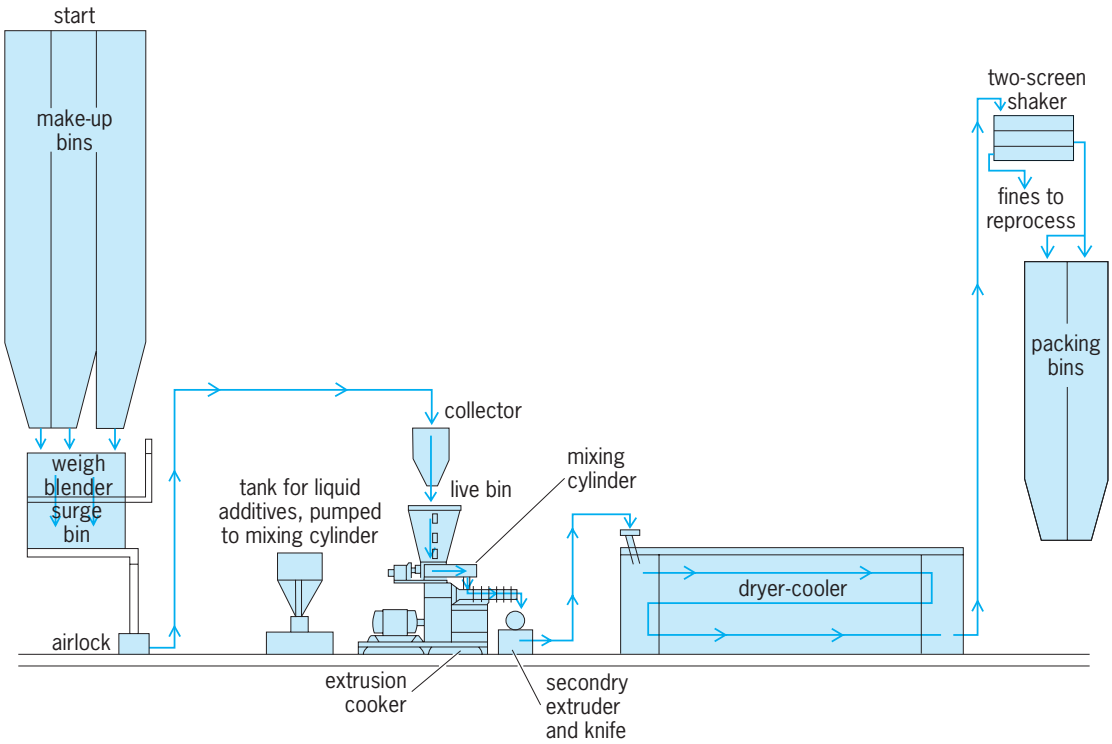

Fig. 4. Typical high-temperature extrusion process for production of a meat or seafood analog. (*After O. B. Smith, in L. D. Hill, ed., Textured Vegetable Proteins, World Soybean Research, Interstate Printers and Publishers, 1976*)

Extrusion. There are two basic types of extrusion technology, high-temperature and low-temperature. High-temperature extrusion is associated with a cooking process where a number of chemical changes take place in the dough or matrix (**Fig. 4**). Examples of these are starch formed into granules, gelatinized proteins becoming thermoplastic gels, inactivation of the antitrypsin factor in soya, and soluble proteins being denatured. Cold-temperature extrusion is a process that is dependent on compression rather than on chemical reactions. With dies designed especially for the process, many sophisticated analogs can be developed by simple combination of the raw ingredient, binders, flavors, and suitable additives.

Chemical setting. Internal setting by chemical reaction is used to prepare fruit analogs that have fairly uniform texture such as apple, peach, pear, and apricot. A two-feed process involving rapid mixing is employed. One feed, which is near pH 7, contains alginate and the calcium-ion source anhydrous dicalcium phosphate; the other feed contains fruit puree, sequestrant, and acid. Since the alginate is predissolved, no alginate hydration problems are encountered. Dicalcium phosphate is insoluble at neutral pH. The structured fruit is prepared by pumping the two mixes through a suitable high-speed mixer and then allowing the final mixture to set under shear-free conditions. The gelling reaction is brought about by the calcium ions released from the dicalcium phosphate, which dissolves as the pH is lowered on contact with the acidic puree phase. The product can next be extruded onto a slab on a moving conveyor and then be diced or cut into uniform shapes prior to further processing. This method is also used to produce pimento pieces.

There are also two other chemical processes which employ alginates—diffusion setting and setting by cooling—and these also are used in the fruit industry to produce structured fruits. *See* ALGINATE.

Types. Analog foods may be divided into three groups: dairy products, meats, seafood, and egg products; fruit and vegetable products; and farinaceous products.

Dairy products. Margarine was one of the first food analogs to be introduced into the marketplace in the United States. This term applies to certain types of shortenings as well as table spreads. The total consumption of table spreads in the United States has remained rather constant since 1970, but butter has declined in popularity while margarine has increased, mostly because of concerns over heart disease and cholesterol.

In the United States, margarine is made largely from vegetable oils that have been hydrogenated and crystallized to remove high-melting triglycerides in order to achieve the proper spreading texture. The vegetable oils may also be blended with lesser quantities of animal fats. In the United States, margarine must contain no less than 80% fat by law. To provide proper spreadability and reduce calorie consumption, whipped butters and whipped margarines are made by incorporating more water and air into them. *See* MARGARINE.

Imitation milks made from soy protein isolates and vegetable oil have met with limited consumer

acceptance. Improvement in quality, plus a greater economic advantage over cow's milk, seems necessary before these products become successful. These products have become available in paper-foil containers as a result of the development of aseptic processing of liquids. *See* MILK.

Frozen dairy novelties with artificial fruit flavors have been marketed for many years, targeted mainly for children. Structured frozen pudding products on a stick have been manufactured as well. This product is manufactured by using freezing equipment and a formula that will minimize dripping while allowing excellent textural and other sensory properties.

Many cheese analogs with a variety of characteristics are available. The key benefits of cheese analogs are linked to economics and nutrition, while some products are formulated to have improved melting qualities, such as on pizza. These analogs are prepared by comminuting and mixing, with the aid of heat, one or more types of cheese plus soy protein and certain emulsifying salts; water, seasoning, and color may be added. When certain optional dairy ingredients such as cream, skim milk, whey, or their solids are added to the cheese blend and processed, the product is called a pasteurized process cheese food or a pasteurized process cheese spread. Both products may contain less fat and more moisture than process cheese as prescribed in the Federal Definitions and Standard of Identity. In addition, the pasteurized process cheese is spreadable at 70°F (21°C). Cheese analogs are less subject to spoilage and shrinkage in weight than natural cheeses. Advantages to the consumer are a more uniform flavor, no waste, and excellent keeping quality. Structured cheese bits are also available. They have good microbiological stability and integrity and are used for incorporation into other food products, such as meats and salad dressings. *See* CHEESE.

Meats, seafood, and eggs. The use of soy proteins to manufacture simulated meats became very popular in the 1960s and was enhanced by rising beef prices. However, simulated meats (meat, poultry, and seafood analogs) have not achieved the popularity that was predicted. While extruded soy products have a cost advantage over the spun-fiber items, they contain residual flavors plus the oligosaccharides stachyose and raffinose, which appear to cause flatus when soy is ingested.

Bacon-flavored textured soy products are sold in retail stores and are marketed to hotels, restaurants, and institutions. Beef- and chickenlike products have been developed for addition to canned chili, dry-mix chicken ala king, and Spanish rice. Other uses are as replacement for part of the freeze-dried meats in dry soup mixes and for simulation of meat particles in dips, crackers, and snacks. Fine particles of textured soy flour are used as extenders in beef and pork patties, where they reduce the stickiness of meat, thereby permitting better release from patty-making equipment.

Much of the meat analog production is designed for the food service industry—hotels, restaurants, armed services, and institutions—which makes up a sizable segment of the food market. The scarcity of skilled workers in the food service industry places a premium on convenience in food preparation. Prepared foods containing meat analogs have been developed that require only thawing or reconstitution and heating.

A striplike bacon analog has been on the market for a number of years. This product is made by laying down spun isolate fibers randomly and holding them together with an edible binder. Alternating layers of red and uncolored fibers are employed to simulate the lean and fat portion of bacon. The product is shaped, set by heat, sliced, and finally frozen. For consumption, the product is merely heated. Since cooking or frying is not necessary, there is no shrinkage, as compared with regular bacon which shrinks to 25% of its original weight.

Other innovative meat analogs are the formed meats. The product is formed by compressing meat under high pressure to simulate various structures, for example, ribs or patties. The benefits are portion control and a lean and boneless product.

Examples of fabricated seafood products are the so-called surimi (compressed fish and shellfish mixtures with added flavor and binding agents) and fish rings. Fish rings of uniform size and texture can be fabricated under high-speed conditions at refrigeration or freezing temperature. This method is used for products from clam, squid, crab, shrimp, scallop, and abalone.

The first egg analogs were products that imitate scrambled eggs when cooked, but do not contain cholesterol. These products are made from various egg proteins and dairy ingredients with added flavor, color, emulsifiers, and stabilizers. The ingredients are blended together, pasteurized, homogenized, and packaged for refrigerated storage.

Fruits and vegetables. Popularity of structured fruits has grown because of uniformity in regard to quality, color, size, and taste. These products are also low in moisture content, convenient, and steady in supply because of independence of seasonal harvest. In the manufacturing of these products, stabilizers such as pectin, locust bean gum, and xanthan gum are used.

One of the largest applications for structured vegetables is in Europe, where pimento strips are manufactured for use in stuffed olives. Structured pimento strips made from alginate and guar gum allow for automatic stuffing at a reduced cost, compared to regular pimentos, which are fragile, require manual stuffing, and have an uncertain supply position.

The manufacturing process can be adapted to produce such products as structured tomatoes, beets, carrots, zucchini, green peppers, and mushrooms. The process can also be adapted to produce dehydrated vegetable pieces, such as onions, peppers, carrots, and peas, for such applications as dry soups and sauces, where uniformity and rehydration properties are important. Structured onion rings are widely used in the food service industry because they provide excellent portion control, whereas natural

onion rings vary in size, leading to serving problems as well as complicating batter and breading operations. Structured onion rings are made by producing a mash from dehydrated minced onions, flour, salt, and alginates and forming it into rings by a machine similar to that used to produce doughnuts. The rings are dropped into a calcium chloride bath where they acquire alginate skins, giving the rings sufficient rigidity to undergo the breading and frying process.

Farinaceous foods. The various possibilities for making farinaceous foods available through extrusion and compression are enormous. Products include granola bars, corn chips, potato snacks, and breakfast cereals containing soy flour as an additive to improve the amino acid balance of cereal proteins and to increase protein content. M. K. Schmidl; T. P. Labuza

Bakery Products

This category of foodstuffs includes such an extraordinarily diverse group of products that it is difficult to define. It is generally agreed that baked products contain, as a minimum, flour, water, and salt, but there are even exceptions to this rule—for example, flourless "cakes" and salt-free bread. Although the name indicates that they are cooked by baking in an oven, many varieties having similar composition, uses, and appearance are fried, for example, doughnuts and certain types of breads.

Formulations. There are numerous formulations for producing bakery products, providing a wide variety of nutritional choices.

Bread and rolls. Acceptable white bread can be made from flour, water, salt, and yeast. Italian bread is usually based on this simple combination of ingredients, and French and Vienna breads are seldom much more complicated. These basic breads tend to have a hard crust that is light in color, a coarse and rather tough interior, and a flavor that is excellent when the bread is fresh but considerably less desirable when it has been out of the oven several hours. Sweeteners, milk, and shortening are often added to improve flavor and texture; and eggs are used occasionally in specialty loaves. Barley malt syrup is very helpful in controlling fermentation. Sometimes a "sour dough" or starter consisting of a piece of dough from a previous batch that has been held a few days is used instead of, or to supplement, the yeast. Mold inhibitors and starch complexing agents (such as monoglycerides of fatty acids) can be used to improve the storage life of the bread. Yeast foods, enzymes, and emulsifiers are sometimes included to modify the processing response of the dough. Most bread is made with enriched flour, which contains several added vitamins and minerals.

Loaves are generally regarded as bread units which are more than about 8 oz (0.2 kg) in weight and are intended to be cut into chunks or slices for individual service. Rolls are considered to be portion-sized pieces, and they often have decorative conformations (as Kaiser rolls) and may be topped, as with sesame seeds, onion pieces, or poppy seeds. Both loaves and rolls may be "washed" with a fluid made from combinations of eggs, milk, and sugar before baking in order to enhance appearance of the crust. Hot rolls may be brushed with melted butter or other shortening to give a shiny crust and an attractive aroma.

There are many variations in shape, flavor, texture, and adjuncts (such as toppings and fillings). Mixtures similar in composition to those described above can serve as the basis for pizza crusts, pita bread, soda crackers, English muffins, and bagels by varying the postmixer processing steps. Variations in composition to include whole-wheat flour, and meals from rye, corn, and oats, as well as many noncereal ingredients such as potato flour and soy derivatives, yield a number of familiar items.

Sweet yeast-raised products. By increasing the levels of shortening, milk, and sweeteners in a basic bread formula, many products suitable for consumption as snacks or desserts can be prepared. These items are generally spiced and flavored and may also include adjuncts such as fruit gels, nuts, and pastry creams. Doughnuts, sweet rolls, coffee cakes, and Danish pastry are examples. Traditional Danish pastry is made from a rich dough which is sheeted out and layered with butter, after which it is repeatedly folded and sheeted until it is "made up" into fancy rolls or coffee cakes. The repeated folding and sheeting give Danish pastries a fine, silky texture with a rather flaky crust.

Chemically leavened products. This exceedingly diverse category includes all those bakery foods which rely on baking soda (sodium bicarbonate) as the source of carbon dioxide. Soda is usually added in the form of baking powder, which contains baking soda, some inert materials such as starch, and an acidic substance that will react with the soda to generate carbon dioxide. If baking soda is added separately to the mix, it is necessary to put some acidic ingredient such as molasses, sour milk, or cream of tartar into the batter. Chemically leavened products include cakes, cookies, certain kinds of crackers, cake doughnuts, most types of muffins, soda biscuits, and pancakes. Generally, these products are slightly to moderately denser than bread and rolls, and they are more crumbly due to a less elastic internal structure. In the absence of added flavors, they are also blander than yeast-leavened products, because they lack yeast and the by-products of fermentation.

On a flour-weight basis (that is, flour = 100%), the sugar content of batters for white or yellow layer cakes should fall in the range 110–160%. Shortening should be about 30–70%, while eggs should at least equal the added fats and oils. Liquid milk (whole or skim) should exceed the amount of sugar by about 25–35%. Soda should be added in an amount equal to about 1.2–2% of the flour weight, with salt 3–4%. Obviously, these general rules are often violated to secure special effects or to compensate for special processing limitations or ingredient problems.

Drop-type cookies can be made from cake batters, but commercial cookie formulas usually contain much less water and often contain less flour than

the conventional cake batter. Some cookies, such as sugar wafers and shortbreads, contain no baking soda. The viscosity of the dough or batter must be adjusted to meet the requirements of cookie-forming machines, which are often very sensitive to small changes in the material to be processed.

Miscellaneous products. Angel food cakes and most pound cakes do not contain any leavening compound. They rely on the air whipped in during mixing for their porous structure and expansion in the oven. Pie crusts are usually made of simple mixtures of flour, fat, salt, and water; since expansion of the dough is not required, leaveners are not added. Turnovers, popovers, and cream puff and eclair shells are leavened by expansion of water vapor and air in their interior as their temperature rises during baking. These batters often include relatively large percentages of shortening or eggs.

Processing equipment and methods. Although both the individual pieces of equipment and the methods of assembling them into production lines differ greatly for different kinds of products, bakery equipment functions can be generalized as measuring, mixing, holding doughs in controlled environments, separating into pieces of predetermined weight, shaping or forming, baking, combining with adjuncts, cooling or tempering, and packaging. The sequence in which these steps are performed is not necessarily in this order, and steps may be omitted or added for specialized products.

Measuring and mixing. In large commercial bakeries, the measuring of ingredients is usually done automatically by meters and scales that receive bulk materials from tanks and silos. Minor ingredients such as colors and flavors are weighed and added manually in most plants. Mixing devices and processes have many features in common with their counterparts in other food factories, but two functions peculiar to bakeries—dough development and air incorporation—require special designs.

The term development is applied to a type of mixing that yields an elastic, extensible dough. Development is accomplished by using mixer agitators that press, stretch, and fold the dough while causing relatively minor amounts of tearing and cutting. Doughs for bread and rolls are typical examples of the intermediates requiring this type of mixing. A common type of mixer suitable for dough development has a large U-shaped container enclosing three or four thick cylindrical rollers which are oriented lengthwise in the trough and move in a circular pattern around the horizontal axis of the trough during the mixing cycle (**Fig. 5**).

Air incorporation occurs in virtually all doughs and batters as a normal and expected result of the mixing process. The size and distribution of the air bubbles have a pronounced effect on the internal structure and ultimate size of the finished product. Angel food cake batter is an example of a material that is leavened entirely by the air bubbles whipped into it during mixing, and special agitators are needed to fold in the air and subdivide it into small, uniform bubbles which are optimal for these cakes.

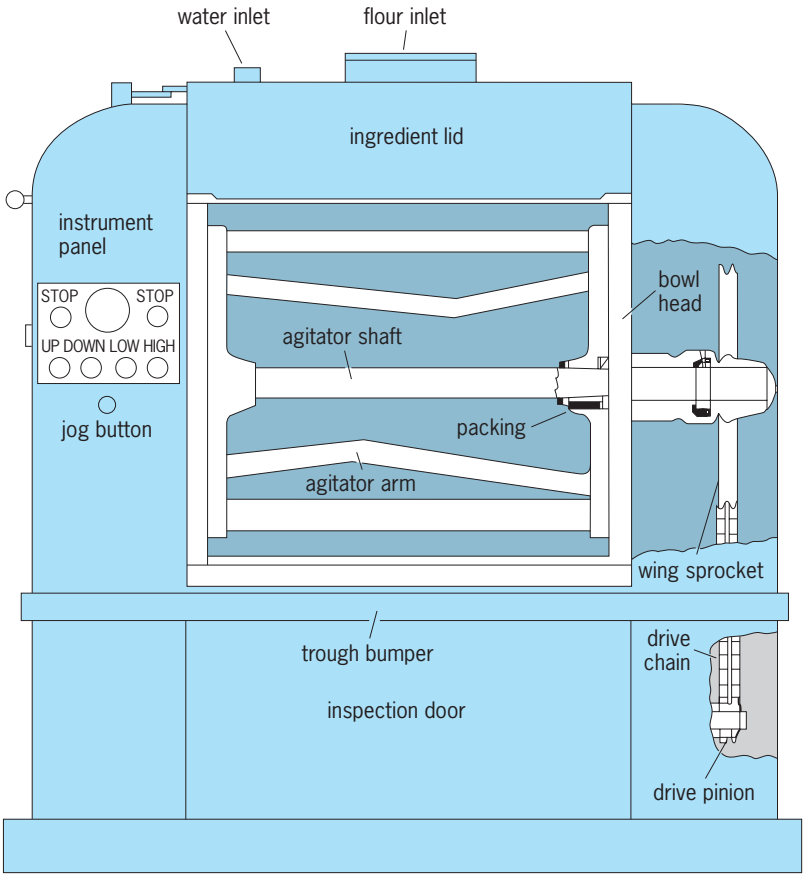

Fig. 5. Front view diagram of a horizontal dough mixer.

Fig. 6. Small integrated bread make-up system. A mass of bread dough is placed in the hopper of the divider (right) which cuts it into pieces of uniform weight. The rounder (center foreground) receives the rough pieces from the divider, forms them into spheroids, and delivers them to the overhead proofer. The dough undergoes a fermentation period while traveling through the proofer in individual pockets. The proofed dough pieces fall into the molder (center background) which sheets the dough and then curls it into a cylinder ready for depositing into baking pans. (*Adam Equipment Co.*)

Controlled environments. Virtually all yeast-leavened doughs must undergo one or more fermentation periods during which the yeast metabolizes carbohydrates and causes numerous changes in the dough. In order to control the rate of these reactions and maintain the dough in suitable condition for further processing, the dough must be held in rooms or cabinets that are maintained within narrow temperature and relative-humidity ranges. In its crudest possible form, the container will consist of a large tub or trough covered with a board. More sophisticated devices include conveyors with individual pockets for dough pieces traveling in cabinets provided with means for closely controlling the temperature and humidity. Temperatures are maintained by circulating air through steam-heated or electrically heated radiators, while high humidity is ensured by injecting water vapor into the cabinet. A high relative humidity is needed to prevent drying out, crusting, and loss of weight.

Dividing and rounding. These steps are used in processing yeast-leavened doughs. They are distinct from the final forming and shaping processes that establish the contours of the finished product. The mass of dough that comes out of a mixer (sometimes having undergone a prior fermentation stage) is dumped into the hopper of a divider, which squeezes a small amount of the dough into a chamber of adjustable volume (**Fig. 6**). The size of the dough piece is adjusted to give a baked product having the desired weight. Since the undivided dough mass continues to ferment and change in density, the divided dough pieces may slowly change in weight as the process continues, even though the divider presses much of the gas out of the dough.

From the divider, the dough pieces are carried immediately to a rounder, if bread loaves are being made. The rounder rolls the dough along a spiral guide fitted to the contours of a vertically oriented cone-shaped rotor. The cone is usually ridged or corrugated and rotates so as to mold the dough piece into a roughly spherical shape as it moves up the guide. The action develops a relatively smooth surface on the dough piece, thereby slowing down the loss of carbon dioxide and water during subsequent steps. The dough is then proofed, or held at controlled temperature and humidity for a length of time necessary to develop the desired handling characteristics and internal structure. The flavor of the finished product also depends to a considerable extent on the reactions occurring during the proofing step.

Forming the product. There are many kinds of dough-forming equipment in use of which the following are examples. Most cookies are formed in one of three ways: (1) dough is pressed into a die cavity engraved in the surface of a cylinder and then removed by suction; (2) soft, batterlike mixtures are extruded through shaped orifices and the strand cut into individual cookies by wires or pulled apart by the relative motion of the extruder and the oven band; or (3) doughs are sheeted out and stamped into individual circles, squares, or other shapes by reciprocating or rotating dies.

To form finished rolls and loaves of bread, various combinations of sheeting, cutting, curling, and

Fig. 7. Tunnel oven. Loaf-size dough pieces enter the oven at left background and are carried through the baking chamber on a continuous metal belt which serves as the hearth. The baked loaves exit at right foreground. Ordinary pan bread would require sets of pans conveyed by a slightly different mechanism, but the basic oven structure would be the same. (*Adam Equipment Co.*)

stamping are used. In some fully automatic bread production lines, a very soft cylinder of dough is extruded from the mixing chamber directly into the baking pan.

Baking. The baking process may take a few minutes for some crackers and cookies to well over an hour for large loaves of hearth breads. The usual type of oven used for the mass production of bread, cookies, and many other types of baked products is a long chamber, open at both ends, through which the products are carried on a continuous metal belt (**Fig. 7**). These tunnel ovens are commonly heated by gas flames burning inside the chamber, but some are heated by electricity, oil, or steam. The burners or other heating elements can be external to the main chamber, with hot air being carried into the oven by blowers.

Cooling or tempering. It is important to bring the temperature of a baked product to within a few degrees of normal room temperature before it is packaged. For small, low-moisture products such as crackers, cookies, and pretzels, cooling must be done slowly so that moisture remaining after the item leaves the oven has an opportunity to redistribute more or less uniformly before the piece hardens. Otherwise, the product may crack spontaneously during subsequent handling and storage. If large pieces of relatively high moisture content, such as loaves of bread, are packed while they are hot, moisture may condense on the crust, leading to more rapid spoilage from mold growth. Refrigeration is seldom used for cooling; most bakeries rely on the forced circulation of room temperature air to perform this function.

Adjuncts. Toppings, fillings, icings, and other adjuncts may be added before or after baking, depending on their characteristics and the effect desired. In most cases, the applicators are fairly simple machines which sprinkle seeds, granules, spices, or other condiments as the rolls pass beneath a hopper equipped with a rotating distributor bar at the bottom.

Packaging. Packaging of bakery products is relatively simple compared to the elaborate methods needed for many other foodstuffs, because either they are distributed by systems that take into consideration their short shelf life or they are low-moisture products that do not readily support microbiological activity. Consequently, breads and rolls are often packaged in nonrigid containers such as paper bags or (more commonly) heat-sealed plastic bags that offer moderate protection against environmental hazards such as insects, mold spores, and dust. Films used for the bags are often made of polyethylene or plastic-coated cellophane.

Samuel A. Matz

Cereal Products

Cereal grains are the primary ingredient of breakfast cereals. The processing of cereal grains into ready-to-eat or quick-cooking specialty items has been rapidly developed as a food industry, initiated only during the past century. These cereals have been promoted based on nutritional and health aspects, convenience, and variety. Consumer acceptance of breakfast cereals has been greatly influenced by product eating qualities, including size and shape, crispness, and flavor, all of which are controlled through the formulation and process sequences.

Early traditional prepared cereals were primarily precooked and dehydrated whole grains or kernel portions. Many basic cereals are prepared in near-traditional methods; however, most ready-to-eat products are prepared by blending flours (wheat, corn, oat, soy, and rice), malt, sugar, milk solids, and selected functional and nutritional additives with water, and mixing to produce a plastic dough that may be manipulated to yield distinctive products. Doughs are heated to produce a soft textured gelatinized starch and to provide desired color and flavor. Extrusion technology has been developed that permits continuous controlled blending and mixing, heating, and a broad spectrum of intricate size and shape configurations. The formed dough piece may be delivered at a critical moisture level or may require tempering prior to additional processing. Efficient extruders may be used as the complete heating and forming and expansion unit, or may be used as a specialized stage of preparation in a flaking or puffing operation. The continuous processing of a wide variety of cereal items can be accomplished only through critical control of formulation and process parameters.

Flaked cereals. The preparation of flaked cereals is initiated by relatively simple cooking of milled or whole grain.

Corn flake products. These flaked cereals are traditional. Blends of corn grits (free of germ and bran) are pressure-cooked in rotating horizontal cookers in the presence of water, malt, sugar, salt, and flavorings. This process requires an extensive cooking at about 18 lb/in.^{2} (120 kilopascals) steam pressure for 1–2 h to provide for starch hydrolysis and gelatinization, development of nonenzymatic browning, inactivation of enzymes, and infusion of moisture. Properly cooked grits will be tender and a translucent brown, possessing about 33% moisture. Cooked grits are mechanically separated and gently dried to about 20% moisture and tempered under conditions which will assure equilibration of moisture throughout each grit. Following the batch pressure cooking, a pelletizing process may be utilized to incorporate additional ingredients to enable uniform blending of a variety of grains into each flake, and to enhance uniformity and size of final toasted flakes. Tempered grits or pellets are passed between smooth cool flaking rolls. The gap (nip) between these rolls is controlled to aid flake size, thickness, and shape. Within the flaking operation, each grit is flattened and expressed to the characteristic flake shape. Tempered moisture control is important to assure uniform flake sizing. Preheating of tempered grits immediately preceding the flake rolls has been demonstrated to reduce cell rupture during flaking, resulting in a reduced rate of hydration (increased bowl crispness) in the final product. Cool moist flakes are hot air–toasted at temperatures

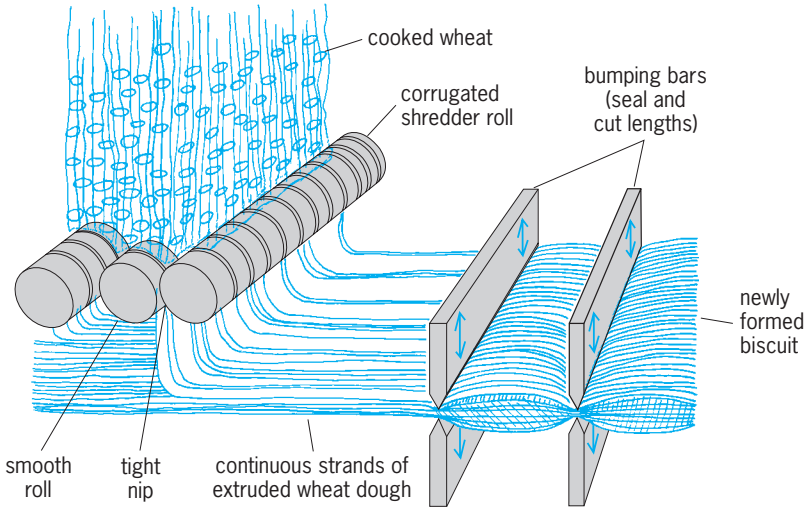

Fig. 8. Schematic for production of shredded wheat biscuits. (*After G. A. Leveille and M. A. Uebersax, Fundamentals of food science for the dietician: Wheat products, dietetic currents, vol. 7(1), pp. 1–8, Ross Laboratories, Columbus, Ohio, January–February 1980*)

exceeding 525°F (274°C) for differential times sufficient to dehydrate, crisp, brown, and blister the flake surface. The conditions of the toasting phase may be controlled to enhance flake crispness and surface blister formation. The final cooled toasted flakes may be spray-coated with an emulsion of vitamins and minerals, flavored or sugared, and dried to about 3% final moisture prior to packaging. *See* CORN.

Wheat flake products. These are prepared from whole, steam-tempered grain that is initially bumped through rollers to rupture the bran and increase moisture permeability through the endosperm. These kernels are pressure-cooked similarly to corn grits to about 50% moisture. Separated and partially dried kernels (30% moisture) undergo further air drying (20% moisture) prior to flaking. Dried cooked kernels are preheated and hot-flaked at 190°F (88°C) to yield a plasticized flake for toasting. Low-temperature (320°F or 160°C) zoned toasting enhances flake curling and surface blistering. *See* WHEAT.

Puffed cereals. The expansion and dehydration of cooked grain kernels or formed doughs is essential to impart the desirable textural properties to the final product. Control of expansion (decrease in bulk density) may be attributed to moisture, product formulation, and expansion process. Generally, expansion or puffing may be accomplished by gun puffing, or direct extrusion. Each requires controlled heating to facilitate rapid volatilization of internal moisture. Gun-puffed cereals, including puffed rice and wheat, are prepared by placing cooked grain (polished or pearled) into sealed vessels which are rapidly heated and pressurized up to 200 lb/in.^{2} (1400 kPa). Sudden release of pressure to the atmosphere expels and expands the cereal. Gun-puffed expanded cereals are also prepared from tempered dough–formed pieces. The expanded product will require additional drying or toasting. Oven puffing, commonly applied to cooked kernels or formed doughs, consists of continuous movement through a fluidized-bed heating chamber. High-velocity turbulent air lifts the cooked moist particles and rapidly heats them. The internal volatilized moisture expands the product volume, cools the surface by evaporation, and dehydrates the product. Critical control of oven design (staged heating zones), temperature, product characteristic formulation, and moisture are required to provide adequate expansion.

Extrusion systems can accommodate mixing, heating, and pressurizing ingredients in a continuous manner. The hot gelatinized dough held under high pressure will rapidly expand upon exiting the shaping die. This moist, expanded, continuously cut product requires further drying prior to packaging.

Shredded biscuits. Whole grain wheat is fully cooked and tempered to yield soft, pliable gelatinized kernels of approximately 50% moisture. This cooked wheat is formed into thin strands (0.04-in. or 1-mm diameter) by passing through high-pressure grooved rolls (**Fig. 8**). The continuous strands of pliable cooked wheat are stacked, generally up to 20 layers thick. The layered sheet is cut into biscuits by bumping bars set at designated intervals. Cut biscuits are dried and toasted (baked) to desired color and crispness in continuous-pass ovens.

Hot-served cereals. Numerous milled grains including wheat (farina), oats (regular rolled oats), and cornmeal (grits) are common. These products are steamed or toasted, dry-milled, sized, and packaged. Each requires cooking with water until palatable (generally more than 5 min). Development of rapid-cook or instant products has received much attention. Instant farina has been prepared by using selected proteolytic enzymes that will reduce cooking times to 1 min. Quick-cooking oats prepared from steel-cut thin-rolled groats are widely available and also cook within 1 min. Instant oatmeal is rendered palatable in the bowl by the addition of boiling water. This product is prepared by blending thin-rolled flakes and a hydroscopic gum to aid rehydration. Numerous fruits or flavored ingredients have been incorporated to aid in this product's popularity. *See* CEREAL.

Mark A. Uebersax

Milling cereal grains. Seeds from cereal crops and pulses provide a wide range of food products, including staples such as bread, pasta, couscous, and dhal. In their natural state, seeds are covered by a fibrous protective coating (seed coat, bran, or husk), which sometimes contains antinutritional substances that protect the seed against predators or disease-causing organisms. Removal of the outer layers of the seed (dehulling or decortication) often improves the nutritional or organoleptic quality of the food product. Dehulling of grain can be achieved by abrasion, attrition, impact, or roller milling, depending on the architecture of the seed. Dehulled grain is reduced to flour in a variety of grinding mills. Wet milling or air classification can further break seed components into starch- and protein-rich fractions, which have specialized uses in the food and chemical industries. *See* GRAIN CROPS; LEGUME.

Abrasion milling. Abrasion milling is used for grains such as sorghum, millet, or rice where the seed-coat layers adhere tightly to the endosperm. Typically, vertically or horizontally mounted, rotating grinding stones gradually rub off the outer layers of the grain, thereby exposing the seed endosperm. Small-scale abrasive dehulling machines (**Fig. 9**) have been gradually introduced in many developing countries to eliminate the time-consuming mortar-and-pestle process. After dehulling, the grain is removed from the machine by tipping over the dehuller barrel. Subsequently, the powdered seed-coat layers are separated from the dehulled grain by winnowing or sieving before grinding the latter to flour. Batch processing of grain in this type of machine ensures that a small-scale farmer in a developing country is still able to consume the grain variety that is produced on the family farm. *See* MILLET; RICE; SORGHUM.

Attrition and impact milling. These methods are used on types of grain that have thicker seed coats that are not tightly bound to the seed (for example, field peas and soybeans). An attrition mill removes hulls as the grain passes between two steel or stone disks rotating in a horizontal or vertical plane (similar to a plate mill). One of the disks may be stationary, and the surfaces of either disk may be configured in a variety of contours. In impact milling, the seeds are hurled at high velocity against an abrasive or metal surface, to crack the seed coat. *See* PEA.

Roller milling. Roller milling is commonly used on wheat to simultaneously separate the bran layers from the endosperm and to reduce the latter to flour. Roller mills are also used to produce cracked grain, as well as wheat and corn flakes. A roller mill consists of a pair of cast-iron rolls with a hard, wear-resistant surface. Adjusting the gap between the rolls produces different degrees of grinding. A wheat flour mill usually contains a series of differentially rotating break and reduction rolls. Break rolls, which are always corrugated, serve to initially cut and shear the wheat seeds into large pieces. Reduction rolls, which have smoother surfaces, further reduce the particle size and scrape the endosperm from the bran layers. Although the modern roller mill has not undergone any profound change, manufacturers have improved the drives, bearings, and feed mechanisms. The most important improvement in wheat flour milling has been in terms of automation to reduce the labor component of the cost of production. For example, in Japan a fully computerized mill with a capacity of about 440 tons (400 metric tons) of wheat per 24 h is operated by a single miller during each 8-h shift.

Combined techniques. An innovative method of combining abrasion- and roller-milling techniques has been developed as a means of improving the recovery of wheat flour in the milling process. Following tempering of the wheat with small amounts of water for a short time, about 65% of the bran is initially removed by friction and abrasion in modified rice polishers. The remaining bran, which is located in the crease of the wheat, is subsequently removed during conventional roller milling. It has been determined that semolina yields are approximately 7% higher when durum wheats are milled with this technology.

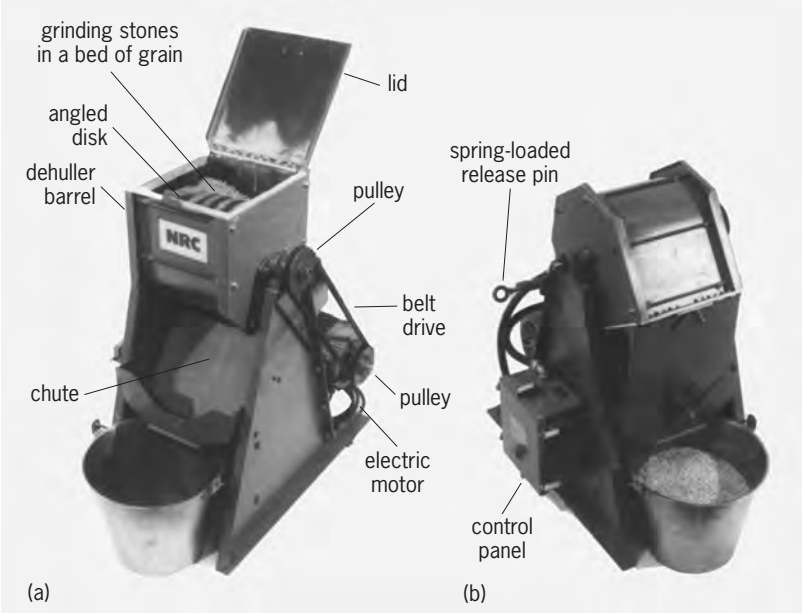

Fig. 9. Small-scale abrasive dehulling machine in (*a*) upright and (*b*) dumping positions. (*Plant Biotechnology Institute, National Research Council of Canada, Saskatoon*)

Development. Food engineers that develop and test milling equipment work closely with plant breeders to ensure that new grain varieties can be milled efficiently. Several small-scale mills have been developed which are able to accurately predict milling performance of gram quantities of early-generation plant breeding material. R. D. Reichert

Confections

It is difficult to define confections because the forms in which they appear are extremely diverse. One common factor is that they are consumed primarily for the sensory properties of appearance, flavor, and texture rather than for their nutritional value. Most of them contain large percentages of sweeteners, such as sugar, corn syrup, or honey. Generally, bakery products such as cakes and cookies are not considered to be confections, although they may have many of the same characteristics, and may be used as one of the components of candy bars (as when sugar wafer cookies are sandwiched with caramel and enrobed with chocolate).

Candies can be classified on the basis of their structure and texture as being primarily dependent on either sugar, protein, or fat for their predominant characteristics. All candies contain sugar or some other sweetener, but hard candies (glassy or amorphous forms of sugar with very low moisture content) and fondants (soft masses of very small sugar crystals surrounded by saturated syrup) rely on the characteristics of sugar for their basic structure. Proteins such as egg white, gelatin, casein, and soy protein concentrates form the matrix of caramels, marshmallows and many other types of foamed candies, and nougats. Starch-and-gum-based confectionery, with

their typical soft chewy texture, form another category. Chocolate is the prototypical form of fat-based confectionery, relying on a melt-in-the-mouth reaction rather than solubility in saliva for the pleasant textural sensation that differentiates it from the sugar and protein structures. The final form of candy depends as much on the processing technique as it does on the formula. *See* COCOA POWDER AND CHOCOLATE; SUGAR.

Hard candies. These confections consist of up to 95% sweetener, with colors, flavors, and often other characterizing additives. They vary in form from the highly colored, almost transparent fruit ball to ribbons or sticks having very complex decorations and containing fillings of contrasting texture and appearance. The sweetener, which is usually a mixture of sucrose and corn syrup, is in a glassy or noncrystalline state, and the candy must contain a very low amount of water (typically, about 1%) to prevent the sugar from crystallizing during storage.

In the original method for preparing hard candy, a concentrated solution of the sweeteners is boiled in an open bowl heated by a gas flame until most of the water is evaporated and the temperature reaches about 310–330°F (154–166°C). Great care is taken to ensure that all crystals are eliminated during cooking; if any solid material accumulates on the sides of the kettle, it is washed or brushed into the boiling candy. Any crystals left in the molten sugar will promote crystallization throughout the entire mass as it cools, ruining the texture and appearance of the finished candy. When the correct temperature is reached, the molten mass is dumped onto a cooling table, and colors, flavors, and acids are mixed in. The temperature will be around 260°F (127°C) at the start of this operation, and in the 200–220°F (93–104°C) range at the finish. Different-colored rods of material can be assembled on the cooling slab to form a pattern that will appear in reduced form in the finished candy pieces.

While it is still plastic, the hot candy block, weighing perhaps 50 or 60 lb (23–27 kg), is gradually formed into a continuous thin rod by rolling and pulling. The batch roller into which the block is placed supports the large, irregularly oval shaped lump of cooling candy on either four or six conical rollers that rotate first in one direction, then in the other. The cavity containing the rollers is heated to keep the sugar plastic. From the batch roller, the thick rope of candy is drawn through the sizer, which consists essentially of a series of heated pairs of rollers having concave and knurled surfaces. Each set encloses a smaller opening and rotates at a faster speed than the preceding pair of rollers. From the sizer, the rope passes into formers, which are available in various designs.

There are several devices for large-scale batch processing and automatic cooking of hard candy formulas. Many of these devices subject candy that has been heated to around 280°F (138°C) to a vacuum of 16 in. mercury (54 kPa) or greater to remove most of the remaining water. Quality is improved, because the less stringent heat treatment minimizes sugar degradation to give improved clarity and keeping quality. Vacuum cookers are often designed to spread the hot syrup in a thin film on steam-heated surfaces to rapidly evaporate the water, which is swept away by hot air currents. The concentrated syrup is then exhausted into a receiving vacuum chamber where the final cooking is performed. In continuous systems, the candy is discharged from the vacuum chamber through a pump or air lock into a hard candy–depositing system or a cooling conveyor. When it is deposited onto a belt, the colors and flavors are metered onto the cooked sugars, the strip is rolled and folded by plows, and the final pieces are formed substantially as described above. For automatic depositing plants, additives are placed in the molten sugar before it is fed into the pump which deposits measured quantities into metal or plastic molds. After a brief cooling period, the molds expel the finished candy pieces onto a conveyor leading to the packing equipment, and return to the depositer for filling. Although they are very efficient and produce uniform pieces of excellent quality, the continuous systems lack versatility and flexibility.

Fillings, such as chocolate, jam, or peanut butter, can be injected by a variable-speed positive-displacement pump into the candy as it is being formed into a rope on the batch roller. As the diameter of the cylinder is gradually decreased through the successive sizing rollers, the proportion of filling to coating remains approximately constant. The squeezing together of the candy rope as the mold closes around it seals the softer filling inside a casing of hard candy. When hard candy pieces are being made by the deposit methods, fillings can be introduced by using concentric nozzles, each fed by a separate metering pump. Filling is extruded through the inner nozzle at the same time that the molten sugar casing is extruded from the external nozzle.

Gums and jellies. These confections have relatively high moisture contents and, as a result, they are elastic and soft and have comparatively short shelf lives. A common variety is the orange slice, which, like most of the types in this category, is basically a starch gel sweetened with sugar and corn syrup and having a moisture content of 18–20%. A typical formula is 33% granulated sugar, 49% corn syrup [64 D.E. (dextrose equivalent), an indication of the sweetness of the syrup], 3.3% thin boiling starch, 4.9% high-amylose starch, and 9.8% water. Although open-pan cooking was the customary procedure in earlier times, continuous processing is more rapid and convenient and is used for most production in the United States. In one such method, the ingredients are first mixed and preheated to about 200°F (93°C), then passed through a jet cooker held at 335°F (168°C) to complete starch gelatization. The cooked product is exhausted into a mixer where the flavor and color are added, or these ingredients are metered continuously into the jet cooker effluent.

Pieces are usually formed by depositing the paste into molds, but the oldest method involves spreading the jelly onto a cooling table and cutting out

segments after the starch sets up. It is customary to "sand" the pieces, or cover them with coarse sugar, to reduce their tendency to stick together.

Hard gums, normally very firm, sticky candies, can be based on gum arabic as the gel agent and contain the usual sweeteners, colors, and flavors. Since starch gelatinization is not a requirement, the sweetener solution is simply boiled to around 225°F (107°C) to establish a satisfactory concentration, and then it is mixed with gum arabic that has been previously softened in water. The finished blend is deposited in starch molds and dried for several days until a satisfactory texture develops. The pieces may be either steamed or oiled to give them a glossy surface.

Pectin, agar, and gelatin can also be used with sweeteners and other ingredients to form confectionery jellies. Different textures are obtained. Gelatin tends to produce toughness, agar usually contributes a "short" texture, pectin gives soft tender jellies, while starch is more versatile. Modified starches have been developed specifically for use in confectionery. *See* AGAR; GELATIN; PECTIN; STARCH.

Products based on crystallized sugar. Fondant, a suspension of very small sugar crystals in a supersaturated sugar syrup, forms the basis of many confections such as cream centers for chocolates. Fondant can be prepared by dissolving sucrose along with some invert sugar or corn syrup (for example, 75% sugar and 25% corn syrup) in water and evaporating enough of the moisture to bring the water content to about 12%. At this point, the solution is supersaturated with respect to sugar, and caution must be taken to avoid premature crystallization. If the solution is beaten vigorously while it cools, the excess sugar will tend to be deposited in the form of minute crystals of relatively uniform size. It is generally agreed that crystals measuring more than about 35 micrometers in their largest dimension will be detected as a roughness or coarseness when the fondant is consumed—a negative organoleptic characteristic. Good fondants contain crystals falling mostly in the range 10–15 μm.

Fondant with added colors and flavors can be used as the basis for several types of confections, but it is normally too dense for optimum acceptability. A lighter, softer texture can be obtained by mixing the fondant with a frappé. This component is prepared by mixing a foam promoter, such as egg white, with sugar and corn syrup, and whipping the mixture to a meringuelike consistency. Typically, 7–10 lb (3–4.5 kg) of frappé is added to about 100 lb (45 kg) of fondant base. Whipping can be performed in the usual mixers found in bakeries, or in pressure whisks that enclose the mixing operation in a vessel containing compressed air. Continuous whipping machines are available.

Cream centers for chocolates can be shaped by extruding and cutting off in the desired piece size, but more elaborate shapes are formed by depositing the fondant mixture into cavities pressed into trays of cornstarch or into plastic molds. Liquid-center chocolates are often made by including a very small amount of the enzyme invertase in the formula. The fondant is liquefied during a period of days or weeks as the result of the hydrolysis of sucrose into the more soluble invert sugar by the enzyme. Since invert sugar is considerably more soluble than sucrose, the firm-cream center (which can be enrobed easily with chocolate) eventually changes to a thick syrup.

Butterscotch, nut brittles, toffees, and caramels depend for their characteristic textures and flavors on the interaction of sugar, dairy products, and fairly large additions of fat. Scotching is a term sometimes applied to the reactions that occur when sugar and butter are heated together until a light brown color develops. Butterscotch was probably the original confection formed in this way, and various types of toffees and caramels were developed by adding other ingredients and modifying the method. Flavor development in these products is due partly to a reaction between reducing sugars and proteins, and partly due to poorly understood reactions of the triglycerides.

Caramels can vary from soft (up to 10% moisture content) to hard (as low as 5% moisture content) and can have different degrees of graininess, which is affected by the size of sugar crystals. Fudges might be regarded as an extreme example of caramels, with high moisture content and a short texture resulting from a large amount of sugar crystallization.

Aerated products. Most of the types of candy described above can be expanded by air bubbles, forming new or modified types of confections. Some examples are nougat and marshmallow. Even hard candy, which is almost pure sugar, can be aerated by stretching and folding the hot mass [about 220–210°F (93–99°C)] through many cycles. Air is entrapped and subdivided into small bubbles by this procedure (which can be performed by hand or automatically) so as to give an opaque "satin" appearance and a more tender texture. Taffy pulling is another example of this type of processing. Heat-softened pieces of hard candy can also be expanded in vacuum chambers to give puffed balls of low density. The centers for malted milk balls are made in a similar fashion.

Nougats are made by whipping a mixture of high-boiled sugar syrup, a milk ingredient, fats, and a foaming agent such as egg albumin or modified whey. A soft, slightly chewy confection results, with the texture dependent to a large extent upon the amount of air which has been whipped into the confection. *See* ALBUMIN.

Marshmallows can vary from very soft to very firm in consistency, but they all consist of a high percentage of sweetener with some structure-forming material such as gelatin to support the foam that is established when the mixture is whipped. Most marshmallows are formed continuously by passing the mixture through pressurized mixing chambers in which assemblies of vanes rotate rapidly. Common cylindrical marshmallows are formed by extrusion and cutting, but the variety known as circus peanuts and similar versions are deposited into starch

molds and allowed to partially dry before they are packaged.

Panning. Nuts and various kinds of small candy pieces can be coated with layers of sugar in a revolving pan. Chocolate can be treated in this manner to give multicolored crunchy-coated disks of candy. The old-fashioned jawbreakers were made in this manner, with many layers of different colors being applied to each piece. The pans are bowl-shaped with an opening in front, and range in size from small lab models to production equipment 4 ft (about 1 m) in diameter. Most of them are rotated by an axle that is inclined about 30° from the horizontal and is powered by a variable-speed drive. They can be heated by steam jackets or gas flames. Up to about 200 lb (90 kg) of candy pieces are poured into the opening as the pan rotates. Small amounts of adhesive solution and powdered sugar or other coating materials are poured or sprayed into the bowl at intervals. As the candy pieces tumble and spin inside the drum, they pick up the coating material in thin, uniform layers. When the coating has been completed, the pieces can be dried by hot air currents. The pieces can be coated with a final layer of carnauba wax or the like to give a shining finish. Separate pans can be used for the finishing or polishing operation, and they may have internal metal ribs to speed up the process. When coating with chocolate, the melted chocolate is sprayed into the pan in small amounts, with each layer being solidified by cool air currents before the next portion is added. Most chocolate-covered pieces are finished by coating with a polishing syrup composed of dextrin, sugar, and corn syrup or with a lacquer dissolved in alcohol. *See* CARNAUBA WAX.

Chewing gum. This is a special kind of confection which consists of a matrix of insoluble latex or gum (natural or synthetic) containing, as a discontinuous phase, sweeteners, softeners, flavors, colors, and other additives. As this mixture is chewed, microscopic pockets of flavoring are gradually exposed to saliva and dissolved. The original gum base was chicle, obtained from the sap of a tree native to Central America. Later, jelutong from Indonesia and sorva from South America were used as replacements for chicle. In modern manufacturing, large quantities of synthetic gums, resins, and waxes are used as part or all of the gum base ingredient in formulations. They often have superior textural properties and "staying power," and are more uniform than natural latexes. Among these synthetic ingredients are poly(vinyl acetate), synthetic rubber of the polyisoprene type, polyethylene waxes, and resin esters. *See* CHICLE.

Production of chewing gum starts with the blending of chunks of gum base (often presoaked with some of the liquid ingredients) and the other formula components in a steam-jacketed heavy-duty is 19.4% gum base, 19.8% corn syrup, 59.7% powdered sugar, 0.5% glycerine, and 0.6% flavor. The mixture will usually be discharged at 113–131°F (45–55°C) and formed into loaves weighing 8–10 lb (3.6–4.5 kg). After cooling, the gum is passed through pairs of rollers in series to gradually reduce its thickness. Powdered sugar may be applied to the surface to prevent sticking. The final set of rollers scores the strip of gum into stick size, and cuts it into sheets which are placed on trays and conditioned at 59–64°F (15–18°C) and 45–50% relative humidity for 24–48 h. In the packaging line, the sheets are broken into individual sticks, wrapped, and packaged. In large factories, the process is continuous up to the conditioning stage, which may take place in a continuous tunnel or in controlled-atmosphere compartments. Bubble gum, which has a different base composition, does not require conditioning and is sized into ropes which are immediately cut into balls or other pieces.

Samuel A. Matz

Condiments

Condiments are aromatic or savory vegetable substances which contribute distinct taste or character to foods, change or enhance specific flavors of foods, and modify the appearance of foods by adding color, as with paprika, saffron, or turmeric. The individual nature of a condiment depends on the type and concentration of volatile oil and other chemical constituents.

Condiments include spices, herbs, flavoring agents, chemicals, sauces, and other materials. One way of classifying them is according to the part of the plant from which they originate: (1) roots such as garlic, ginger, sassafras, horseradish, and turmeric; (2) stems and leaves, such as basil, sage, tarragon, thyme, bay leaf, and marjoram; (3) barks such as cinnamon and cassia; (4) buds such as capers, cloves, and saffron; (5) fruits (ripe or immature) such as paprika, pepper, capsicum, juniper, and vanilla bean; (6) seeds such as mustard, cardamon, anise, celery, dill, coriander, bitter almonds, and poppy. Other condiments may be classified as chemicals (monosodium glutamate, sodium chloride, sodium citrate) and sauces (wine concentrates, vinegars, relishes).

Properties of most condiments are distinct because of their volatile characteristics or the substances of definite chemical composition therein.

With the increasing interest in international cuisine, or ethnic foods, the use of condiments has increased tremendously and should continue to do so. Many of these cuisines accent savory seasonings, spices, and penetrating tastes and aromas, such as in Italian, Cajun-Creole, Mexican, and Asian foods. With the use of savory spices which complement basic taste and flavor, the amount of salt can be reduced, a positive medical attribute. Commercial spices are frequently sold as blends. They are often irradiated prior to packaging, and frequently are marketed in concentrated form.

Standards have been established by the Federal Food and Drug Administration for the more important spices, herbs, and seeds as listed above. These consist of specifications for percentages of crude fiber, total ash, volatile oils, starches, bacteriological limits, and other basic constituents of each specific item. The purpose of such standards is for the

maintenance of uniformity of product, the insurance of freshness, and the provision of better quality control for the consumer. *See* MONOSODIUM GLUTAMATE; SALT (FOOD); SPICE AND FLAVORING.

Desserts

Desserts are generally considered to be sweet foods that are served at the end of a meal. Commercial preparation of desserts includes the following methods: canning, freeze-drying, drying, production of dried powders, and aseptic techniques. In spite of the emphasis on dieting and weight reduction, processed desserts are found in greater variety and amounts than ever. However, single portions are available to assist in dieting regimes.

Canning is used mainly for fruits. The fruit is packed in sugar syrup or in sugarless water solution—with or without synthetic sweeteners in order to reduce calories, or for dietetic purposes. The use of saccharin is being replaced by use of aspartame and acesulfame-K, which are white, odorless, crystalline products 200 times as sweet as sugar. Thickeners, such as vegetable gums, are added to simulate sugar syrup. Fruits of excellent visual appearance are often packed in glass. The canning process is also used for puddings, such as rice pudding or Indian pudding; and crepes suzettes and baba au rhum are canned by using syrups or semithick bases.

Frozen desserts include fruits such as strawberries, rhubarb, peaches, and mixed fruits, including exotic tropical fruits. Specialty items such as pies, cakes, puddings, crepes, and many types of ice cream, sherberts, and related desserts are also frozen. *See* ICE CREAM.

The development of a combination of freezing and drying, that is, freeze-drying, has produced superior fruits that are very close in taste and texture to the natural fruit.

Dried fruits, such as figs, apricots, plums, and grapes, are traditional desserts that were prepared by sundrying. This method is coupled with improved drying methods that provide controlled time, temperature, and humidity.

Gelatin desserts are made by combining sugar, plain gelatin, citric acid, sodium citrate, flavor, and color. Pudding powders, such as chocolate pudding, may be regular (requiring cooking) or instant (requiring mixing only). These contain powdered sugar (dextrose), starches (pregelatinized in instant pudding), flavor, and color. Puddings have been developed that are made with products such as tofu (bean curd) or yogurt, as well as flans that contain carrageenans (Irish moss). Additional developments include high-fiber muffins and cakes that are healthful desserts and alternatives to eating breakfast cereals. *See* AGAR; CARRAGEENAN.

By using aseptic techniques, puddings, such as rice pudding, tapioca, or chocolate, are packed in individual containers, either plastic or cans, and then they may be stored at refrigerated or room temperatures.

Another method for packaging baked goods has been developed using inert gases; this method provides flavor retention and increases shelf life. It is known as controlled-atmosphere packaging (CAP) or modified-atmosphere packaging (MAP). These baked products are usually packed in microwavable polystyrene trays.

Alfred J. Finberg

Soft Drinks

Soft drinks are nonalcoholic beverages, that is, they contain no so-called hard liquor such as whiskey or gin. Soft drinks can be divided into two classes: carbonated soft drinks (sometimes referred to as soda, soda water, and soda pop) and still beverages (also referred to as noncarbonated soft drinks, fruit-flavored drinks, and fruit punches).

Soft drinks were first offered for sale during the seventeenth century in Europe as a mixture of water, lemon juice, and honey. Their origins can be traced to the attempts by many Europeans to copy and improve the carbonated waters and beverages that were known for their therapeutic value. The first company of *Limonadiers* in Paris sold their lemonade from tanks which were carried on their backs. The production of soft drinks has evolved from this stage to highly sophisticated production lines that are capable of filling up to 2000 cans per minute. The packaging has also evolved from returnable glass bottles to lightweight nonreturnable glass bottles, and to steel, aluminum, and plastic containers. In many communities the glass bottles and aluminum cans are being segregated from other solid waste and recycled. There are studies under way involving use of plastics that are recyclable or biodegradable after being discarded.

Carbonation is the main difference between the two classes of soft drinks. Carbonated soft drinks are prepared by dissolving carbon dioxide in water which is to become part of the soft drink, or in the completely formulated soft drink itself. A volume of carbon dioxide gas will be absorbed by an equal volume of liquid at 60°F (15°C) and 14.5 lb/in.^{2} gage pressure (100 kPa); correspondingly, at 60°F (15°C) four volumes of carbon dioxide gas will be absorbed by the liquid at 45 lb/in.^{2} gage pressure (310 kPa), and five volumes at 60 lb/in.^{2} gage pressure (413 kPa).

Except for the carbonation, there is very little difference between the different classes of soft drinks. All normally contain a sweetener, acidulant, preservative (if necessary), and flavoring. The nutritive sweeteners (sugars) used may be sucrose (table sugar) derived from sugarcane or sugarbeets, 50% inverted type consisting of 42–47% sucrose with the remaining percentage equally divided with fructose and glucose, or with high-fructose corn syrup (HFCS). Regular corn syrup has not been found to be suitable for soft drinks; however, 42% high-fructose corn syrup (approximately 42% fructose and 58% glucose) and 55% high-fructose corn syrup (approximately 55% fructose and 45% glucose) are acceptable. The exact amount of dry sugar solids may vary in a product depending on the type or combination of sweeteners that is used. *See* FRUCTOSE; GLUCOSE.

The combination of sugars available to the soft drink manufacturer allows flexibility should there be a disruption in the supply of one sweetener type or if economic factors dictate the decision to purchase one over the other. There are also advantages and economy of time in the production process when using liquid sweeteners versus granulated cane or beet sugar, which must be dissolved first.

The acidulant is normally citric, malic, or phosphoric acid, depending on the product. Citrus-type beverages (for example, orange) usually use citric or malic acid. The colas as a rule use phosphoric acid. When necessary, the preservatives are either sodium benzoate or potassium sorbate, depending on the pH of the beverage. Not all beverages contain preservatives. Soft drinks are products very rigorous to degradation. They usually avoid microbiological spoilage because of their anaerobic environment and low pH. Preservatives are normally used only as an added precaution or when the product formulation warrants it. The flavoring is usually one of the ingredients added in the smallest amounts. Flavor chemistry has advanced to the stage where the flavor component is highly concentrated, and this provides savings in the costs of transportation and packaging.

Optional ingredients may also be added in such proportions as are required to accomplish their intended effects. These could include fruit juice, antioxidants, foaming agents, and so forth. Diet beverages are sweetened with saccharine alone $2.4–4.2 \times 10^{-4}$ oz/fl oz (0.24–0.41 mg/cm^3) or in combination with aspartame. Many products are sweetened with 100% aspartame alone. Both saccharin and aspartame are nonnutritive sweeteners. All ingredients must be food-grade or higher to ensure the highest quality and a uniform product. A trend has developed in the soft drink industry to use as many natural ingredients as possible in the product formulation.

Although potable water is generally taken from a municipal supply, it is processed further to eliminate or reduce impurities that could affect the quality of the finished product. In a few soft drink manufacturing plants, the water treatment equipment may simply consist of a sand filter and an activated carbon purifier. The sand filter is sized to allow a flow rate of approximately 2 gal/ft^2 (81 liters/m^2) per minute. This filtration removes minute matter. The carbon purifier removes not only chlorine and color but also tastes and odors that may be present. In most soft drink plants the water is first treated chemically by using an oxidant (such as chlorine or chlorine dioxide), lime, and ferrous sulfate. The lime serves two purposes: it reduces the water alkalinity by the removal of the calcium and magnesium bicarbonates, if necessary, and it serves to convert the ferrous sulfate to a ferric hydroxide floc. The flocculent masses entrap very minute particles, including algae and planktons. The water is then passed through the sand filter and carbon purifier. The carbon purifier is really another filter that contains activated carbon which removes color, taste, and chlorine. *See* WATER TREATMENT.

Mark J. Pietka

Sauces

The word sauce is derived from the Latin *salsa*, meaning salty or salted, and is defined as a condiment or combination of condiments in liquid or semiliquid state, eaten with food as a garnish or accompaniment for the purpose of adding flavor, piquancy, or interest. Sauces may vary in thickness from light to heavy. Some are thin liquids, with finely dispersed colloidal particles, while others may contain large pieces of solid materials.

Preparation. The use of sauces developed from the ancient method of food preservation by salting. Sauces function in three basic ways: to enhance flavor, to "atone" for its absence, or to act as a binder for other ingredients. Every sauce has a binding ingredient to hold the liquid portion and provide required texture. There are two principles in preparing sauces: the roux principle (a mixture of flour and butter) and the emulsion principle (egg yolks used to absorb and suspend particles). There are three distinct types of roux, from which mother sauces are created: white, blonde, and brown, and these give rise to hundreds of variations. The white roux is a thoroughly blended mixture of flour and butter. Blonde roux is a white roux cooked until it turns a straw color. Further cooking until nut brown produces a brown roux. Oil or another fat can replace the butter in a brown roux.

Three basic sauces prepared from blonde roux are béchamel sauce (white sauce), velouté sauce, and sauce parisienne (sauce allemand). Two basic sauces are derived from brown roux: brown sauce (sauce espagnole) and, with the addition of a meat jelly, demiglace.

There are three basic sauces based on the emulsion principle: hollandaise sauce, mayonnaise, and vinaigrette sauce, also known as French dressing.

Commercial. Many commercial sauces are available; some are processed and are stable at room temperature until used by the consumer. Others are refrigerated or frozen at the time of manufacture and must be stored at proper temperatures. The frozen products are usually more expensive but frequently are judged to be superior to products processed in jars or cans.

Packaging. The most common packaging materials are glass and plastic, since these make the product visible. Tubes, similar to toothpaste tubes, are common in Europe, but have not had wide consumer acceptance in the United States. Unit portions in individual foil or plastic packets are a convenient method of dispensing and serving sauces and dressings in fast-food chains and on planes. Advances in laminated foil packaging have increased shelf life.

Standards. The Food and Drug Administration has established standards of identity for mayonnaise, French dressing, and salad dressing. These standards state allowable ingredients and their minimum amounts. There are no official standards for

other sauces and dressings except that all ingredients must be listed on the label in order of amount.

Specially formulated dressings that are low-calorie, low-fat, or low-sugar have been developed as a result of consumer interest in preventive medicine, nutrition, and weight control. The result has been an upsurge in interest in reduced-calorie mayonnaise and dressings. To qualify for low-calorie status, the caloric content must be reduced by at least one-third, one serving supplying not more than 40 calories (167 joules).

Manufacturing. Commercially prepared sauces are either liquid (water-based) or emulsified (oil-based). The aqueous sauces, such as hot sauce, mint sauce, and Worcestershire sauce, may be stabilized to minimize separation.

Improved methods of processing have been incorporated into the handling, cooking, homogenizing, and packing of manufactured sauces. Tomato sauce and tomato ketchup are made from strained tomato pulp concentrated with vinegar, sugar, and spices to different solids levels.

In ketchup manufacture, pulped tomatoes are concentrated by boiling or in a vacuum concentration unit to the desired consistency and solids content. The condiment is made into several grades, depending upon the concentration of the tomato solids. Fancy ketchup has more than 33% of total solids present, while standard grade has as little as 25%. The prepared condiment is preserved by canning, usually in glass containers for consumer uses. Chili sauce is tomato pulp with the seeds remaining and concentrated with added sugar, vinegar, and spices to specific solids level.

Worcestershire sauce is a combination of vinegar, hydrolyzed proteins, vegetables, and fruits that is allowed to age in barrels to develop a balanced flavor blend. Soy sauces contain proteins such as wheat, corn, or soybean processed by controlled acid hydrolysis, with condiments, molasses, and caramel color added. Pepper or hot sauce is made by concentrating chili or tabasco peppers and adding vinegar and salt. Meat or barbecue sauce is a combination of tomato pulp, soy base, vinegar, and spices. Mustard sauce and prepared mustards are made by grinding or milling mustard seed and adding vinegar and spices; novelties include various blends, for example, "sweet and spicy" or "sweet and hot." Spaghetti sauces, combining tomatoes, herbs, and spices, with or without meat, are processed in cans or jars. Following the increase in popularity of ethnic foods, particularly Italian pastas, many variations of tomato sauce have been introduced.

Applesauce is a combination of cored, sliced apples cooked with sugar, water, and spices. Cranberry sauce is made by cooking cranberries, sugar, and water to specific solids level. There are many types of fruit sauces, such as those made from peaches and apricots, which are concentrated, combined with sugar and spices, and packed for use as ham glazes or with meat dishes. Wine is a component in some special sauces.

Sauces that are generally emulsified are mayonnaise, salad dressing, and related dressings. They are oil-in-water emulsions with a vegetable oil content ranging from 30 to 85%. The oils used are mainly soybean or cottonseed. These products also contain 4–10% egg yolk, with the exception of the pourable types of French dressings, which generally contain a vegetable gum or stabilizer to maintain emulsification. They are commercially manufactured with mills of colloid or homogenizing types, pressure types of homogenizers, ultrasonic homogenizers, or continuous automatic mixers combined with high-pressure units. An improvement in this field was the introduction of instantaneous cooking followed by continuous cooling of the starch phases of salad dressings. The average commercial salad dressing contains 30–40% vegetable oil. The difference is made up with cooled starch paste for salad dressings. The jet-type handling of starches has effected tremendous economies and streamlining of the entire salad dressing industry.

Quality. Digital processing and computer controls have provided further efficiency and quality control, leading to longer shelf life and product stability. The coordination of these developments has resulted in uniform final emulsification by proper dispersion due to particle-size reduction control. Commercial products such as hollandaise and bearnaise sauces have been direct results of these applications used with sterilization procedures.

Soups

A soup is generally a thin puree or liquid consisting of an extract of meat or other ingredients (such as vegetables) plus seasoning. It is prepared by boiling in a single pot. Soups normally are classified as basic consommes or bouillons, thin broths, or thicker soups. Modifications are made by thickening with tapioca, flour, cream, egg yolks, or other solids such as pureed vegetables, legumes, pieces of fish, poultry, or meat.

Originally, simple purees of tomato were the most popular commercial soups available to the public. Sophisticated consumer tastes and manufacturing skill have made soups like gazpacho, she-crab, and cream of cheddar cheese commonplace, and the variety is increasing constantly. Soups satisfy the modern consumer market needs, because they are easy to prepare and readily digested. They are economical and suitable for all ages, health conditions, and types of weight control.

Canned. Early commercial soups, typically tomato, chicken, vegetable, and meat soups, were heat-processed in cans. Usually these were in concentrated form (50% concentrate), with water, milk, or a mixture of both to be added in home preparation. Some soups were canned single-strength and served without dilution. Today many variations, such as ethnic soups, special-diet soups (for example, low in salt), and single-portion sizes are available in cans.

Because of convenience to the consumer, newer types of soups that can be opened, heated, and

served in the same container within a few minutes have become available; however, the largest number of soup products available are the canned variety.

Dry soup mixes. These are packed in flexible bags made of foil or paper boxes which are automatically sealed. The mixes are usually composed of dehydrated soup stock and may include dehydrated vegetables, meat, chicken, and spices mixed together sufficiently to obtain uniform distribution. Onion, chicken-noodle, and beef soups are typical of the varieties available.

Instant soups. These may be made from freeze-dried ingredients, spices, and precooked noodles, which rehydrate upon addition of boiling water. Unlike dry soup mixes, these require no cooking and are very useful for rapid preparation.

Frozen soups. These soups are cooked with a shorter process time than that for canned soups. After cooking, they are packaged and then frozen. Varieties best suited for freezing are fish-and-vegetable combinations such as chowders and barley-mushroom, where flavor and texture of the ingredients are better retained by short process time.

Microwavable soups. Because more people work and have less time for food preparation, this development is beneficial. These soups can go from the shelf or freezer, to the microwave oven, and to the table in less than 10 min. Some varieties are packed in sturdy high-density polypropylene bowls. Most of these soups are in frozen form, although some are shelf-stable and packed in plastic.

Compressed cubes. These were an early commercial type of soup that are still popular. They are made of salt, meat or chicken, and vegetable extracts with spices, and are slowly soluble in hot water to provide a clear bouillon.

Alfred J. Finberg

Fermentation

In the broad sense, fermentation is defined as a metabolic process in which organic compounds undergo chemical changes as a result of enzymatic actions. The participating enzymes, specialized types of proteins that function as biological catalysts in chemical reactions, are synthesized by all living organisms, including bacteria, yeasts, and molds. Since a broad spectrum of plant and animal materials readily supports the growth of fermentative microorganisms, fermentations are universal mechanisms used to prepare and preserve numerous foods and food products. *See* FERMENTATION.

Alcohol fermentation. Fermenting yeasts share a common function in the manufacture of beer, ale, and wine, namely, the production of alcohol. Hops (dried blossoms of the hop plant and an essential flavor ingredient in beer) is added to the water extracts of germinated barley grains, boiled, cooled, and inoculated with selective strains of top or bottom types of fermenting yeasts, thereby producing beer or ale, respectively. In wine production, the sugars of grape juice are readily available for direct fermentation by yeasts. Following fermentation, the liquor is removed from the sediment, aged, and filtered to achieve ultraclarity. Red and white table wines are produced from singular grape varietals or blends of varietals and usually contain 10–13% ethyl alcohol, whereas dessert wines are prepared by fortification, that is, they contain added amounts of alcohol and sugar. *See* ETHYL ALCOHOL; MALT BEVERAGE; WINE; YEAST.

Dairy fermentations. Cheeses and cultured dairy drinks (yogurts, buttermilk, and so on) are produced by the acidic fermentations of milk by lactic acid bacteria. Various types of cheeses are prepared from curds (coagula formed by acidification) that have been heated at 122°F (50°C) for varying periods of time. Soft-type cheeses (cottage, cream, Neufchatel, and so on) are prepared from curds that have been subjected to reduced cooking times, whereas increased holding times followed by extended periods of ripening at 35–60°F (2–15°C) produce the cured cheeses such as Cheddar, Parmesan, and provolone. Cofermentations, that is, processes in which mixtures of lactic and nonlactic acid bacteria or molds produce cheeses which possess unique flavors and distinct physical characteristics such as the eye or hole formation in Swiss cheese (gas-producing bacteria) or the blue-vein effects found in bleu or Roquefort cheeses (mold). Process cheeses are prepared by heating blends of cheeses and emulsifiers to a smooth mass which is subsequently dispensed in containers or packaged as sliced cheeses.

Vegetable fermentations. Fruits and vegetables (for example, cabbage, carrots, cucumbers, or olives), when treated with salt or brines, readily support lactic acid fermentations. In sauerkraut fermentation, granular salt is thoroughly mixed with shredded cabbage, 2–3 lb (0.9–1.4 kg) of salt per 100 lb (45.4 kg) of cabbage, whereas with intact fruits and vegetables, such as cucumbers or carrots, the commodities are totally immersed in the brines prior to fermentation. The levels of lactic acid generated during fermentation are governed by the amounts of fermentable sugar present in the plant materials; fermented pickles usually provide less than 1% acid, whereas more than 2% acidity is produced in the fermentation of sauerkraut. During processing, fermented sauerkraut and brines are heated to 170°F (77°C) prior to container closure, whereas pickles, upon removal from the fermentation tanks, are thoroughly washed to remove excessive amounts of salt, packed in containers containing formulated brines (spices, sugars, vinegar, and so on) and pasteurized at 165°F (73°C). Pasteurization can be circumvented by incorporating chemical preservatives into the packing brines. The latter products may be packaged in glass containers or flexible plastic bags and stored at refrigerator temperatures.

Pickling. Brines, containing vinegar (notably acetic acid), spices, and condiments, are commonly used for preserving and imparting unique and distinct flavors to many foods. Many of the fermentable foods described above can also be prepared by nonfermentative pickling processes. For example, fresh cucumbers immersed in brines overnight are

packed into containers and covered with hot brines containing 2–3% salt, vinegar (1.5% acetic acid), and spices. The acidified product is then heated at 165°F (73°C) for 15 min, followed by prompt cooling. In addition to plant materials, fish (such as herring or salmon) and cured meats (such as bacon, ham, or sausage) can also be prepared by the pickling process. John R. Stamer

Fish and Seafood Products

Fish and shellfish landed by American fishing boats are used for both industrial products and food. The industrial products are fishmeal and fish solubles used primarily as poultry feed and fish oil. The food products are fresh and frozen fish, canned fish, and cured fish (salted, smoked, and pickled). The following descriptions of processing methods emphasize American methods. *See* MARINE FISHERIES.

Industrial products. In the United States a greater volume of fish is used to make fish meal and fish solubles for animal (primarily poultry) feed and for fish oil than goes into any other human food operation. This is because the United States menhaden catch, none of which is used for human food, often is greater than that of all other finfish combined. *See* ANIMAL FEEDS.

Menhaden are rendered by the wet process (**Fig. 10**). The catch is unloaded by pumping with water, and after draining and weighing, fish pass into the raw box to await processing. The fish are cooked in a long, horizontal, cylindrical steam-jacketed cooker; often these have openings to inject steam into the cooker, which contains a screw conveyor to move the fish through. Emerging from the cooker is a mixture of solids and liquid consisting of oil and of water in which nitrogenous and other substances are dissolved. These materials go into a press and emerge as a press cake that passes into steam-jacketed dryers to produce the meal to which an antioxidant, usually butylated hydroxytoluene (BHT), is added. This retards oxidation of the residual oil to prevent overheating, which otherwise might start a fire during meal storage. The liquid phase from the press goes through centrifugal separators to give oil and stickwater, the latter being a solution of protein and other solubles. The stickwater is condensed in multiple-stage evaporators, reducing moisture content to 50% or less. The oil is purified in oil polishing centrifuges.

In addition to the very large production of meal and oil from menhaden, meal, oil, and solubles are by-products derived from tuna cannery waste. Processing procedures are very similar to the procedure used for menhaden.

Some whole fish, especially those of low oil content, are rendered into meal by the dry process; that is, the fish are dried directly without cooking or pressing.

Fresh and frozen fish. Fish to be marketed fresh or frozen usually are sold in one of the following forms (**Fig. 11**): whole, drawn, dressed, headed and gutted, steaks, or fillets. In small-scale operations, the whole or dressed fish may merely be washed, sometimes scaled, packed in boxes of crushed ice, and shipped to wholesale or retail markets.

In large-scale operations, especially if fillets are to be cut, semimechanical packing lines may be employed. The fish are first washed and sometimes scaled in cylindrical open metal washers provided with sprays of water. The fish then pass along moving belts to work stations where they are cut into fillets. The fillets next pass through a washing tank, sometimes followed by a dip in tripolyphosphate solution, especially if the fillets are to be frozen. Fillets may then pass over a translucent glass plate, lighted from below, to be inspected for the presence of parasites or defects. The fillets then go to a packing table for hand packing and sealing in moisture-vapor–proof containers. Containers are filled carefully in order to avoid air pockets, which would result in oxidation, discoloration, and off-flavor development.

Freezing of fillets often takes place on shelves made of pipes that contain the refrigerant. Sometimes fans circulate the air. In other cases the packaged fish are frozen between plates containing the refrigerant, with considerable pressure on plates and fish (plate freezing); or the packages are placed in air tunnels with rapidly blown refrigerated air.

When whole or dressed fish are frozen and stored, they are protected against dehydration and oxidation by dipping after freezing in cold water to form a protective ice glaze. If the fish are packed in paper-lined wooden boxes, they may require only infrequent reglazing; otherwise, they are frequently reglazed by spraying with water.

Fish steaks, a form used especially for salmon, halibut, or swordfish, are cut with a band saw after freezing and ordinarily are glazed before packaging.

Breaded fish products. The production of breaded fish portions is an important manufacturing operation. Fish fillets are frozen into blocks, ice-glazed, and shipped to points where the breaded product is to be prepared. The blocks are cut on band saws into either stick form or nearly square portions. The product then moves along a belt where first a batter is applied, and then they are covered with a dry breader. The final color of the product depends upon the type of batter and breader used. Commonly batters contain grained corn flour, corn meal, and nonfat dry milk solids, while breaders are often winter wheat cereals, bread, or cracker crumbs. The products are then cooked for several minutes in hydrogenated oil or liquid corn oil at 375–405°F (191–207°C). An alternate product is breaded but not cooked. The cooked products must be cooled slowly and packed while they are at 90–100°F (32–38°C).

Canned fish. Tuna and salmon are the two most common American canned fish (close to 90% of the total fish pack). Formerly tuna was packed largely in southern California. This has changed, and nearly all of the tuna is packed by American companies at canneries closer to the locations where the tuna are caught, primarily in American Samoa and Thailand in the Pacific and Puerto Rico in the Atlantic. Salmon are canned largely in Alaska, with much smaller amounts canned in Washington and Oregon.

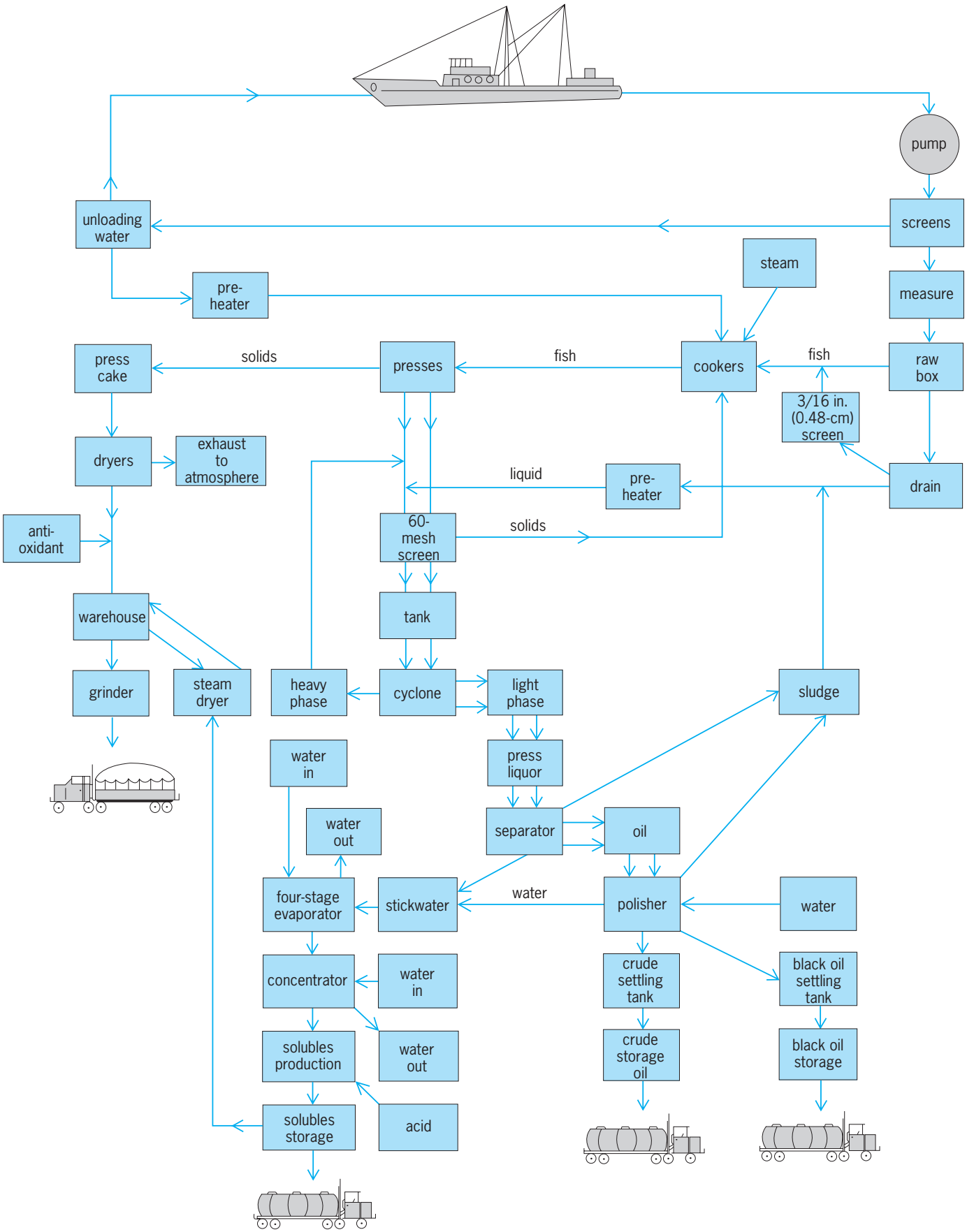

Fig. 10. Flow chart for the preparation of fish meal, condensed fish solubles, and fish oil from whole fish by using the wet process. (*Zapata Haynie Corp.*)

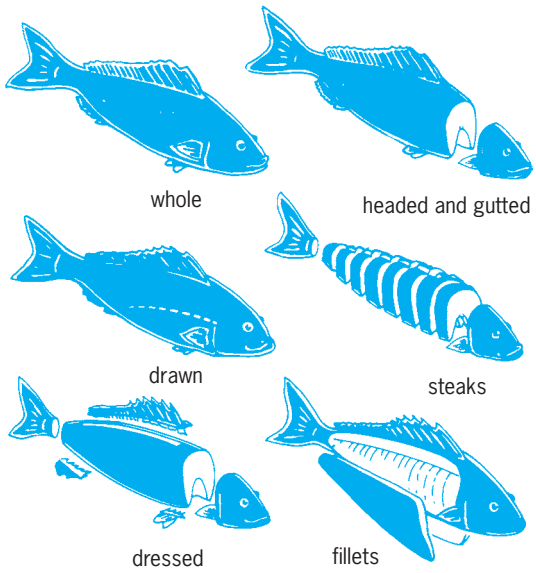

Fig. 11. Common market forms of fresh or frozen fish.

In the tuna-canning process (**Fig. 12**), the fish, most of which have been frozen on the fishing vessel and held frozen in air until approaching the tuna cannery, are completely thawed after unloading. They are then eviscerated by hand and inspected for removal of any fish showing indication of spoilage. The fish are loaded according to size in paper-covered wire baskets and precooked in ovens with steam at atmospheric pressure. The length of the precooking process varies with the size of the fish. Fish from the precooker are cooled, often overnight, and then cleaned by hand to remove skin and bones and to separate the dark from the light meat. Dark meat and other trimmings are used for pet food and for making meal and oil. The tuna loins are cut and packed by machine as either solid meat tuna or chunk pack. The filled cans pass down the line where salt and then oil—generally soya bean oil—heated to about 180°F (82°C) are added. The cans are then sealed and retorted.

Salmon are usually delivered unfrozen to the cannery, where they are sorted as to species. In the cannery the head is removed and the fish slit to remove eggs, which are exported to Japan. Dressing, which includes removal of viscera and cleaning, is done by means of a machine known as the iron chink. The fish are scaled and washed at the sliming table and then enter a cutter that chops fish into can-sized chunks that pass into the filler machine, which inserts them into cans. Salt is added by machine at levels of 1.25–1.67% salt. Cans then enter a machine which rejects those that are underfilled. At the patching table, underfilled cans have additional salmon added. Filled cans with lids loosely attached then enter the vacuum sealer. After passing under sprays of water, the cans are retorted, then cooled.

Cured fish. Unlike the situation in many other countries, curing (salting, smoking, pickling) is no longer a major fish preservation method in the United States. Salting of fish is mainly used for preparing mild cured salmon, an intermediate step toward smoking. Smoked fish (important species are salmon, sablefish, chub, herring, whitefish, and whiting) as well as small quantities of a large variety of pickled herring products are produced primarily for areas such as greater New York, Chicago, and Los Angeles. There has, however, been a steady decline in demand for all these cured fish products.

Innovations in processing. Fish flesh can be obtained in maximum yield by using special equipment such as the type used for many years by the fishing industry in Japan, where minced fish is in demand (**Fig. 13**). Such cheaply obtained minced fish can be used for products familiar to the American consuming public, such as fish fillets molded into shape from minced fish and used in the production of frozen fish blocks for the preparation of breaded fish products. Also, minced fish can be molded into shape and sold as artificial crab legs. Such uses of minced

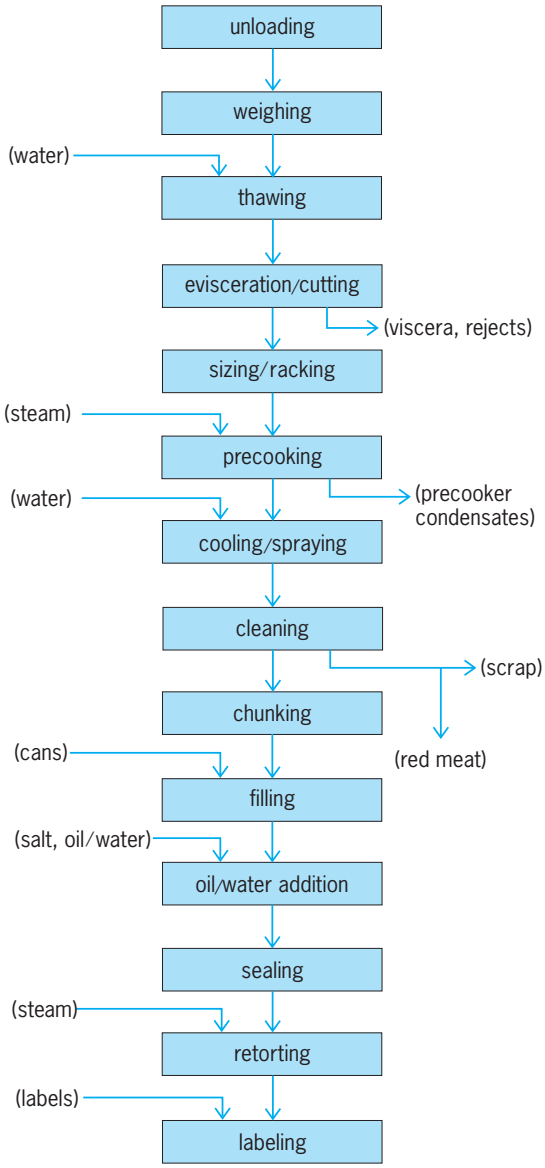

Fig. 12. Processing steps in the canning of tuna.

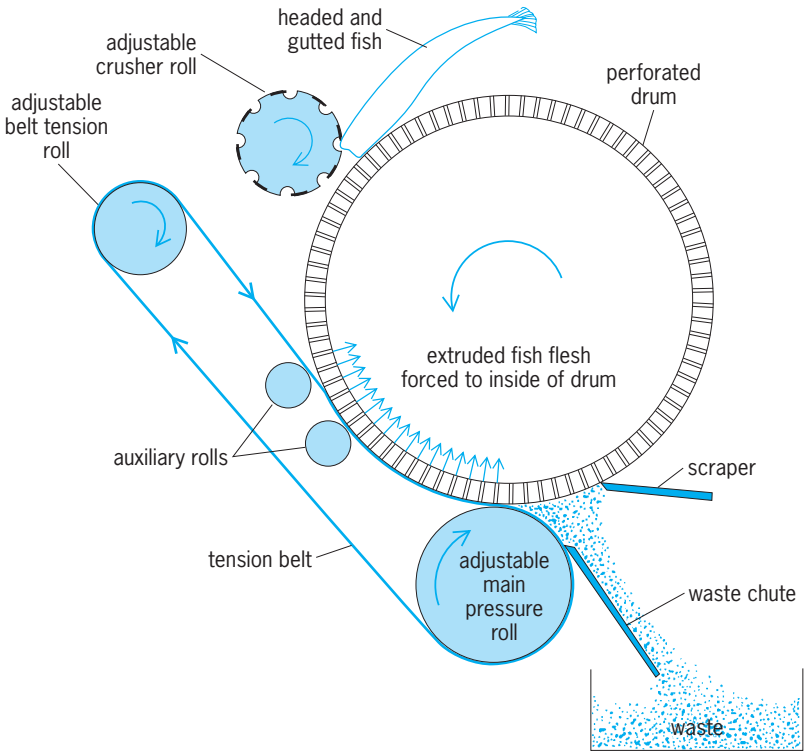

Fig. 13. Schematic sectional view of a fish flesh-separating machine to produce minced fish.

fish yields savings in production costs and, in addition, permits marketing of species unfamiliar to the public.

Another trend has developed from awareness by consumers of certain desirable nutritional properties of fish. Many fish are frozen in brine or refrigerated in chilled seawater, either of which allows penetration of varying amounts of salt into the fish. Federal regulations could require a statement of sodium content on package labels and that it be fairly uniform in amount among different samples of the particular product. The processes being used for fish preparation result in wide variations in sodium content, well beyond what would be permitted; some modification in processing methods is required.

Another nutritional property of fish relates to the findings that omega-3 fatty acids in the oil of fish seem to minimize greatly the occurrence of heart attacks. These desirable effects are due to the presence of certain highly unstable fatty acids that occur only in fish. In order for the public to gain maximum benefits from this property of fish, improved methods of packaging or treatment are being developed to ensure optimum retention of these beneficial substances.

Maurice E. Stansby

Fruit Products

Botanically, a fruit is the ripened ovary and accessory parts of a flower. In common usage of the term, a fruit implies a fleshy product typical of trees or bushes that is high in acidity and sugar and, when ripe, possesses a characteristic flavor; however, many vegetables (such as tomatoes, squash, and beans) are true fruits. Fruits, such as apples, peaches, pears, and plums, are consumed primarily for their fine texture and distinctive flavor, as well as their high nutritive levels of vitamins and minerals. Wide differentiations in flavor, color, and texture exist among fruit varieties, some of which are best suited for fresh market distribution while others are appropriate for processing. *See* FRUIT.

Handling fresh fruits. Fruits are highly metabolic and thus normally undergo relatively rapid postharvest degradation in terms of quality. Respiration, a natural biological process (see reaction below), results in consumption of sugar and oxygen and in production of carbon dioxide and heat, but will subsequently result in fruit senescence.

$$\underset{\text{Glucose}}{C_6H_{12}O_6} + \underset{\text{Oxygen}}{6O_2} \xrightarrow{\text{enzymes}} \underset{\text{Carbon dioxide}}{6CO_2} + \underset{\text{Water}}{6H_2O} + \underset{\text{Heat}}{673\ \text{kcal}}$$

Ethylene gas, which accelerates the respiration and ripening rate of many fruits, may be used under controlled conditions to enhance the storage life of that fresh fruit. Respiration rate is also dramatically influenced by temperature; thus, low temperature (41°F or 5°C) may extend high-quality shelf life. Several fruits, however, are cold-sensitive and undergo chilling injury, which reduces their quality. This moisture loss is decreased by maintaining high relative humidity (90–95%) conditions. Wilting and shriveling of fruit, caused by the process of transpiration, are a result of water loss during storage. *See* ETHYLENE.

Controlled-atmospheric storage. Significant progress in long-term storage of fruit under specified and monitored conditions has been achieved. Apples are particularly well suited for controlled-atmospheric storage. Fruit is sealed in warehouses equipped to maintain controlled temperatures, relative humidity, and gas (atmosphere) environments. Maximum storage life is attained by lowering oxygen concentrations to approximately 2% and increasing carbon dioxide (CO_2) levels to about 5%. Scrubbing the air to remove ethylene gas also improves storage conditions. High-quality shelf life, as evidenced by firm, crisp, and flavorful fruit, may be increased by many months with these controlled-atmospheric conditions.

Canning. Unlike vegetables, which require pressure-cooker processing (250°F or 121°C) to inactivate spores of *Clostridium botulinum,* fruits naturally possess sufficient acidity (low pH) to permit thermal preservation at boiling water temperature (212°F or 100°C). Fruits are prepared by washing, sorting, and trimming. Specialized preparation may be required for some fruits, such as stem and pit removal for cherries and peel and core removal for apples. Sliced or prepared fruit may require blanching to remove tissue gases prior to filling into enamel-lined cans or glass jars. Aggregate fruit products

(whole berries or sliced portions) may be canned in water or sugar syrups. Filled containers are exhausted by thermal or vacuum to remove air prior to hermetic sealing to assure removal of oxygen and adequate vacuum within the container upon cooling. Sealed containers are heated in water or atmospheric steam to achieve temperatures of about 200°F (93°C) in the container center prior to cooling. High quality of processed fruit can be achieved only through rapid handling, controlled heating, and proper selection of varieties.

Apple sauce. Peeled, cored, or whole chopped apples are heated with steam to inhibit oxidative browning and to aid in extraction of the fleshy pulp. The cooked mass may be sweetened and then passed through a fine-mesh sieve. This process, termed finishing, is typically accomplished by continuous extraction through a cylindrical screen in which rotating internal paddles express the cooked apples through 0.06-in. (1.5-mm) perforations. Large seeds and other firm portions that do not pass through the sieve exit through the end of the finisher.

Freezing. Freezing is effective in maintaining flavor and color qualities of many fruits. Rapid freezing techniques reduce ice crystal damage and improve quality. Properly prepared fruit may be frozen in sugar syrups or dry sugar packs (1 part sugar to 3–5 parts fruit) to reduce oxidation. Addition of ascorbic acid (0.2%) provides further control of pigment oxidation. Individually quick-frozen products are prepared by fluidized-bed freezing in which the fruit moves over high-velocity cold air currents to achieve rapid crystallization. They are then packaged and stored at −0.4 to −8°F (−18 to −22°C) as are all frozen fruits. Fruit purees, sugared or concentrated, are slush-frozen in a scraped surface heat exchanger, packaged as crystallized slurry, and then solid-frozen for later use in ice cream or baked goods.

Juice. Fruit juice products provide economic utilization of blemished but sound fruit. Fruits are initially washed and sorted and then prepared to facilitate maximum extraction. Fruit may be chopped and directly extracted (cold break) or heated to 180°F (82°C) prior to extraction (hot break). The latter method will increase extraction of pectic substances, inactivate pectinases (enzymes which degrade pectin), and produce a thicker juice possessing more body. Juice extraction is typically accomplished with batch hydraulic presses or continuous-screw presses which express juice and produce a dry pulp residue, termed pomace. Juice may be cloudy, that is, containing suspended solids, and may be further clarified by use of either pectinases or fining agents, such as bentonite, and centrifuged and filtered to remove the suspended material, producing a product of sparkling clarity. Juices may be pasteurized by continuous-flow heat exchangers to achieve 176–194°F (80–90°C) temperatures for 25–30 min (high-temperature–short-time). Pasteurized juice is suitable for canning in tin or glass, freezing, or refrigerated distribution. Sodium benzoate added at 0.1% is also effective in stabilizing acidic fruit juices.

Grape juice, produced by either hot or cold pressing, requires precipitation and removal of crude tartrates prior to bottling. Chilling or slow freezing will accelerate crystallization and sedimentation of tartrates. Following these pretreatments, the detartrated juice is suitable for packaging.

Citrus juices are produced by press or screw extractors, deoiled by vacuum distillation, and centrifuged, finished, deaerated, pasteurized by the high-temperature–short-time method, filled at high temperature and sealed, and rapidly cooled. Citrus concentrates utilize vacuum concentration to remove water without imparting dramatic deleterious flavor changes; however, add-back processes are used to dilute the evaporated concentrate with sufficient fresh juice to give proper final solids content and flavor profiles. Citrus concentrates are frozen in consumer packages requiring three parts water for reconstitution or are used for preparation of reconstituted, single-strength, refrigerated juice for local distribution. Advances in aseptic packaging technology have shown great potential for production of high-quality fruit juice products for use in institutional and retail markets.

Dehydration. Sundrying is traditional and accounts for a large volume of commercially dehydrated fruit. Prepared fruit is dehydrated following treatment to control the enzymes that cause oxidative browning. Blanching (steam), treatment with sulfur dioxide (sodium bisulfite solution), or drying in the presence of burning sulfur controls browning and enhances color and flavor of the dried product. Hot-air–drying is most common with fruit placed in trays for cabinet dryers (batch) or tunnel dryers (continuous) in which product moves on carts, trays, or mesh belts. Moisture is removed to about 20% to ensure microbial stability. Fruit "leathers" are prepared by drum-drying purees, while instant fruit juice powders are produced by spray-drying systems. Dehydrated fruits are packaged to control moisture absorption during storage. *See* DRYING.

Jams, jellies, and preserves. Production of pectin-acid-sugar gels from fruits is a traditional means of improving utilization of blemished but sound fruit. Jellies are prepared from extracted juices; jams from whole purees; and preserves from whole, intact fruit. Pectin (either naturally present or added) in the presence of sufficient sugar (increased hydrogen bonding), water, and acid (pH 2.9–3.4) forms a delicate gel. Standard jellies are prepared using 45 parts of fruit juice and 55 parts of sugar and concentrated to 65% soluble solids (65° Brix scale, a hydrometer scale for sugar solutions), and are hot-filled and cooled in final package (typically glass or plastic). Control of formulation and process is essential to attain a properly set gel. Dietetic jellies are produced using pectin containing limited methoxylation (pectic acid) which, in the presence of calcium, cross-links to form a gel without requiring sugar. Special care in heating and deaerating fruit for preserves and the use of a rapid-set pectin are essential to maintain distribution

of fruit in the gel, thereby eliminating the floating fruit defect. *See* GEL.

Vegetable Products

Vegetables are plant products obtained as leaves, roots, stems, flowers, or fruits. The pH of vegetables ranges from 4.5 to 7.0. Handling of fresh vegetables requires controlled low temperature and high humidity to retard desiccation. Vegetables may be classified according to their respiratory activity; high-respiring vegetables, such as peas, beans, and corn, have less postharvest stability than do low-respiring vegetables, such as potatoes, beets, or carrots. Selection of optimum harvest time (maturation stage) is essential if maximum quality is to be attained. This is frequently in opposition to maximum yield; therefore, quality grading procedures often have a function in pricing structures for harvested vegetable crops. Postharvest cooling of vegetables aids in the removal of field and respiration heat, thus improving quality. Hydrocooling in chilled water, forced-air cooling, or evaporative (vacuum) cooling are often used.

Canning. Vegetables are washed with chlorinated water to remove soil, stones, and other debris and to reduce the microbial population. Products are sorted, trimmed, peeled, or sliced according to specific needs. Peeling can be done by mechanical or chemical means. Carrots and potatoes may be peeled by mechanical abrasion or by chemical lye peeling, in which products are dipped into hot lye (5–10% sodium hydroxide at 180°F or 82°C) for a short time to dissolve the outer peel. Peels and residual lye are readily removed by water sprays.

Blanching is a relatively mild heat treatment to which a product is subjected before it is packed into the container. Blanching serves as a final wash to reduce microbial load, to remove certain raw or off-flavors, and to increase the pliability of the product, thus facilitating filling and permitting the expulsion of dissolved or occluded gases from the plant tissue. During blanching, the product is exposed to hot water or steam for several minutes. In hot-water blanching systems, the product must be submerged in water at about 194°F (90°C) for a set time. Direct steam blanching reduces nutrient losses and water requirements and has thus found wide use in the industry.

Products are put into containers immediately after being blanched. Filling systems range from simple, manual operations to highly engineered machinery. Products packed in salt brine or syrups are first put into containers and then covered with the liquid portion. Filled containers are generally exhausted by being passed through a hot water bath or steam chamber to expel gas and to expand the product before the can is sealed.

Hermetically sealed low-acid foods (vegetables) require processing temperatures in excess of 212°F (100°C) to inactivate *Clostridium botulinum* spores. These temperatures are commonly attained in pressurized steam retorts. Specific handling preparation procedures and process schedules are required for each individual vegetable product.

Commercial vegetable canneries frequently require transportation of crops from large geographic areas to enable an economically efficient operation by extending the normal local crop season. Many vegetable canneries also rely on dry beans, peas, or other dry, stable crops as an off-season production item. Dry beans require soaking (overnight in cold water or for a short time in hot water) to hydrate the seed and reduce splitting prior to filling.

Filled, exhausted, and sealed containers require pressurized cooking at 250°F (121°C) for approximately 40 min to assure health safety and produce a tender, palatable product.

Pickling. Lactic fermentation of vegetables produces a delicate flavor in a wide variety of products. Cucumbers are salt-brined and fermented by naturally occurring lactic bacteria. Salt content in the brine is measured in terms of the Salometer scale, 100° Salometer = 100% salt saturation. Salt concentrations are initiated in 1000-bu (3.5-m^3) wooden vats at 40° Salometer (around 10% salt). This concentration is reduced to about half this level in 1–2 days because of osmotic removal of water from the cucumbers. As the fermentation continues, salt concentrations are increased to 40–50° Salometer over several weeks.

Fermented salt stock is preserved by the acid and salt, and it may be stored for months prior to processing. Processed pickles require desalting by water extraction prior to other treatments, such as sugar impregnation to form sweet pickled products.

Genuine dill pickles are produced by natural lactic fermentations in the presence of spices and relatively low salt concentrations. Selected other vegetables (onions, peppers, and cauliflower) may be brined and preserved in similar fashion.

Large numbers of cucumbers are produced as fresh-packed products. Whole, cross-cut or longitudinally sliced cucumbers are filled into jars, flavored, and acidified with vinegar before steam pasteurization.

Sauerkraut is produced as a lactic fermentation of finally shredded cabbage. Salt (2.5%) is added to clean sound shredded cabbage, thoroughly mixed, and packed into fermentation vats. The distinctive lactic flavor is developed during several weeks of active fermentation in the brine produced from the cabbage tissue. Kraut is heated in its own juice, packed into glass or tin containers, and processed in boiling water.

Freezing. Block packing or individually quick freezing are suitable methods for preserving high-quality vegetables. Blanching pretreatments (176°F or 80°C water or atmospheric steam) are essential to inactivate enzymes. Improperly blanched products will undergo dramatic flavor changes during frozen storage, and the result is a grassy or an ensilage taste and odor. Rapid-freezing systems reduce ice crystal damage and enhance textural characteristics of the vegetable product.

Dehydration. Prepared vegetables (sliced, diced, or chopped) are suitable for hot-air–drying and are often used as ingredients for formulated and prepared products, such as soups, casseroles, or entrees. Cabinet, tunnel, or kiln dryers are commonly used to reduce moisture content to less than 4%. Vegetables require blanching prior to air-drying.

Mark A. Uebersax

Meat Products

The processing of meat begins with slaughtering and dressing of cattle, sheep, and hogs to produce meat for human consumption. The carcasses are normally chilled to near freezing temperatures to retard the growth of pathogenic and meat spoilage bacteria. Most carcasses are reduced into primal or subprimal cuts and distributed to supermarkets or other retail outlets, or else they are used for further processing. Beef primals and subprimals are generally vacuum-packaged in gas- and moisture-impermeable bags, packed into corrugated paper boxes, and shipped as boxed beef. This retards the growth of undesirable microorganisms and extends the shelf life of fresh meat. Although limited numbers of carcasses are distributed as halves or quarters with all cutting being done at the retail store, the amount of cutting at the retail level has steadily decreased. There has been an increase, however, in centralized cutting and packaging at warehouses that service some retail markets.

Further processing of meat involves one or more of the following procedures: freezing; curing; smoking or cooking; canning; sausage manufacturing; drying; production of sectioned and formed meat products; manufacturing of restructured meat products; enzyme tenderization; and rendering of fat. These processes decrease perishability and provide a variety of different products with distinctive flavor and texture characteristics that add to their desirability.

Frozen meats. Freezing greatly extends the storage life of meat. It is utilized not only for fresh raw meat but also for cooked meats and for many convenience items, such as precooked meat entrees, meat pies, and meat items in frozen dinners.

Although frozen meats are stable for months, their acceptability declines because of development of oxidative rancidity. In cooked products, rancidity develops much faster and is called warmed-over flavor. Development of rancidity can be delayed greatly by covering precooked meats with a sauce or gravy having antioxidant properties. Extracts from many vegetables possess natural antioxidants and can be used in the sauce or gravy.

Cured meats. Curing also greatly extends the storage life of meat products. It is accomplished by adding salt either alone or in combination with sodium or potassium nitrite. Sugar, polyphosphates, and spices may also be added to impart special properties to cured meats. The curing ingredients are commonly made into a brine and injected under pressure.

Salt acts by dehydration and altering osmosis to inhibit growth of spoilage bacteria. Nitrite retards development of rancidity and produces the characteristic cured meat flavor. It also reacts with the meat pigments to stabilize the color, and most importantly inhibits the growth of pathogenic bacteria, such as *Clostridium botulinum*, and other spoilage microorganisms. The level of nitrite allowed in curing is limited, since excess amounts may react with the natural amines of the meat and produce carcinogenic *N*-nitrosamines. Polyphosphates increase the water-binding capacity in cured meats. Sugar and spices add flavor.

Cured meats also may be smoked or cooked. However, some products (for example, corned beef and canned hams) are produced without smoking.

Smoked or cooked meats. Smoking meat is carried out for development of a smoky flavor, preservation, creation of new products, development of color, and protection from oxidation. Smoking and cooking are frequently carried out together.

Although smoke was originally generated by smoldering wood or sawdust, liquid smoke is now frequently used and this has led to a sharper demarcation between smoking and cooking. Liquid smoke is prepared by destructive distillation of wood. The particle phase is removed by washing with water, leaving the vapor phase in solution. Liquid smoke is applied by spraying, dipping, or atomization with heat. Its advantages are in uniformity of application, in ease of cleaning, and in preventing pollution.

Most smoked meats are cured and also cooked—including frankfurters and bologna, which can be eaten without further heating. However, fresh pork sausage may be smoked but requires cooking before being served.

Canned meats. Manufacturing of canned meats is accomplished by heat processing to produce either pasteurized or shelf-stable products. The former must be held under refrigeration, as is the case for most canned hams and picnics (bottom part of pork shoulders). Shelf-stable canned meats are heated for sufficient time at 250°F (121°C) to destroy most pathogens and can be stored without refrigeration.

Although more ham is canned than any other meat item, corned beef hash, beef stew, chili con carne, Vienna sausages, meat balls with gravy, potted meat, and canned luncheon meat are all common commercial products. Some canned meat products are reacted with curing ingredients before being sealed in cans and heat-processed.

Sausages. Meat products that are comminuted or emulsified with salt and seasoning are called sausages. Generally they are stuffed into cylindrical artificial or animal casings. Sausages may be coarsely or finely chopped or passed through an emulsifier to produce an emulsion, in which the particles are coated with either myosin or collagen proteins. On heating, the collagen gelatinizes to form jelly pockets, whereas myosin coagulates on cooling to form a stable emulsion. Since myosin is a contractile protein coming from skeletal muscle, an adequate quantity of lean meat is essential. Overchopping can also result in fatting-out as a result of too much surface area from the fat. *See* COLLAGEN; MUSCLE.

The most common sausage products include frankfurters and bologna, which are emulsion-type sausages; and these differ only in size. Meat loaves are also emulsion sausages but often contain added flavoring ingredients. Fresh pork sausage and smoked pork sausage are examples of uncured ground sausages, while thuringer and cervelat are semidry sausages and salami and pepperoni are dry sausages.

Dried meats. These products are produced by curing followed by drying at a low humidity. Drying must be carefully controlled, or surface case hardening will occur and retain the moisture in the center. These products include dried beef (jerky) and a number of dry sausages, such as pepperoni and salami. Their low moisture content makes them shelf-stable at room temperature.

Sectioned and formed meats. These products are produced from whole muscles or large chunks of meat that are bonded together to form a single fused piece resembling an intact cut. Binding is accomplished by (1) the myofibrillar proteins, which are extracted by tumbling or massaging to form a tacky exudate; (2) a meat emulsion, which is mixed with the intact pieces of muscle to fill the voids between pieces and to produce a sticky coating; and (3) adding a nonmeat binder, such as albumin, to coat the muscle pieces and bind them together. Salt is usually added to solubilize the binder. The meat pieces are forced into a casing to bring them into close contact. Heating then coagulates the binding material so that a single fused piece is formed during cooling.

Turkey rolls, sectioned and formed hams, and beef rolls are examples of these products. Usually sectioned and formed meat cuts are boneless, but some may contain a small amount of bone, such as mock chicken drumsticks.

Restructured meats. These products require specialized processing equipment. Three basic procedures are used: chunking and forming, flaking and forming, and shredding and forming. In chunking and forming, the meat is passed through a grinder or dicing machine to produce cubes no larger than 1.5 in.^{3} (24.6 cm^3). Salt, polyphosphates, and other seasonings are added to the meat, which are mixed or tumbled to extract the myofibrillar proteins. Conversely, a small amount of a meat emulsion may be added as a binder. The chunked meat is then stuffed into casings or formed into meat logs that are frozen, tempered, and sliced.

Flaking is accomplished by high-speed equipment that cuts the meat into thin flakes. Salt is generally added to extract the myofibrillar proteins, which serve as binders. The meat is placed in a special press that forms meat logs which are frozen, tempered, and cleaved into steaks, chops, or roasts. These products are generally frozen and sold as frozen cuts.

Shredding and forming produces a texture more like intact meat cuts. The manufacturing procedures are otherwise similar to those used in production of restructured meat.

Restructured meats often suffer from oxidative rancidity. Limiting the time in frozen storage and decreasing salt levels can help in retarding rancidity.

Enzyme-tenderized meats. Fresh meats, especially the lower grades of beef, are frequently treated with enzymes of either plant or fungal origin to improve tenderness. The cuts are either dipped or injected with enzyme solutions and are generally frozen during distribution. Upon thawing, the enzymes act upon the myofibrillar and connective tissues and enhance tenderness. This process has made steaks of acceptable tenderness available to more consumers at a lower price. A patented enzyme tenderization process involves injection of papain into the bloodstream of cattle. The enzyme is then carried to the tissues by the circulatory system before the animal is slaughtered. The enzyme fraction is active only during cooking, resulting in an increase in tenderness.

Meat fats. These are valuable by-products of the meat industry and are produced by rendering (heating), which separates the fat from the supporting connective tissues. In manufacturing lard from pork fat and tallow from beef fat, decolorization, deodorization, plasticization, interesterification, fractional crystallization, and hydrogenation may be used to alter the characteristic appearance, color, odor, flavor, and firmness.

Rendered fat is classified as edible or inedible based upon its suitability for human food. Inedible grades are used for animal feeds or for industrial purposes, such as manufacture of lubricants or soaps. *See* FAT AND OIL (FOOD); SOAP. A. M. Pearson

Poultry Products

Poultry products in the United States include chicken, turkey, duck, goose, and game birds. The processed carcasses, or parts thereof, are supplied refrigerated as ice-packed or dry-packed, or as frozen products. The nonfrozen products are often called fresh, as opposed to frozen.

Commercial production and processing of poultry is confined to specialized production units that can deliver truckload lots of live birds (8000–10,000 broiler chickens) to processing plants. This results in uniformity in birds being processed at one time. Large plants for broiler chickens convert over 300,000 live chickens to chicken products in an 8-h day.

The conversion of live birds to poultry products involves a number of processes, all of which are controlled by a U.S. Department of Agriculture veterinarian inspector to assure consumers of the safety, wholesomeness, and compliance with label claims for all products. The birds are humanely dispatched and allowed to bleed thoroughly prior to scalding to facilitate complete removal of all feathers. The inedible viscera is removed along with the head and feet. The carcasses are thoroughly washed prior to chilling to less than 41°F (5°C) in an ice water bath. Some water is absorbed during washing and chilling, but this is carefully monitored by the inspection service personnel. Many of the processes are automated so as to keep cost to a minimum.

After chilling, broilers may be packed in boxes of about 24 carcasses with ice for refrigerated distribution to retail markets and institutional feeding establishments. Dry packing is in large plastic bags with about 24 carcasses per bag. The air is evacuated and replaced with carbon dioxide gas. The bags are sealed and then placed in boxes and distributed as refrigerated food at temperatures of less than 37°F (3°C). Distribution of poultry at 28°F (−2°C) has gained rapidly as a major temperature for handling. This deep-chill temperature keeps all free water in a package in solid form, thus greatly extending the shelf life. Some birds, especially turkeys, ducks, and geese, are individually packaged and frozen for distribution. Frozen poultry should be handled at less than 0°F (−18°C).

A large quantity of poultry products is cut up at the processing plant and packaged as complete birds (all cuts from one carcass) or as parts such as breast, thigh, drumsticks, wings, backs, and necks. Labeling requirements are very explicit on cut-up parts so as to accurately describe the portions of the carcass included in the cut, such as breast with ribs, or thigh with back. These requirements aid the consumer in getting exactly the portions desired.

Breast and thigh meat are often further processed by removing the skin and bone. The boneless breast meat may then be processed further into specialty gourmet items. The thigh meat is frequently cured and smoked to produce turkey or chicken ham. The sale of poultry meat as formed products (nuggets, patties, and croquettes) is an interesting development. Boneless breast, thigh, and leg meat are converted into the formed products. The uncleaned bones and some of the bony parts remaining are then mechanically deboned to yield mechanically separated meat. In many instances, some mechanically separated meat is used in formulations for the formed products. However, the mechanically separated meat is used as a major ingredient in the production of poultry frankfurters, bologna, or other processed poultry meat products.

As channels have developed for distribution of nonfrozen uncooked meat at 28°F (−2°C), a number of poultry processors are producing precooked poultry meat products packaged in microwavable containers for distribution at this deep-chill temperature. These precooked nonfrozen items are generally available from processors located near major metropolitan areas.

Egg Products

The proteins of the whole egg contain all essential amino acids in a balance favorable for meeting human dietary needs. An excellent source of all vitamins, with the exception of vitamin C, eggs also provide significant quantities of iron and trace minerals to the diet (**Table 3**).

Most eggs to be converted to products are broken by mechanical breakers capable of operating at speeds up to 100 cases, or 36,000 eggs, per hour. The machines can either separate whites and yolks or produce whole eggs at these production rates. The egg products (albumen, yolk, whole egg, or blends), standardized to specific values (**Table 4**), are then pasteurized prior to packaging for liquid or frozen distribution or for drying. Pasteurization lowers the total microbiological population and eliminates all pathogenic microorganisms. Times and temperatures of pasteurization vary with the products, the methods employed, and the requirements in different countries. A process known as ultrapasteurization has been developed for liquid whole eggs. It uses conventional pasteurizing equipment operated at higher than usual temperatures or longer than usual times to produce a pasteurized product in an almost sterile condition. By combining this with aseptic packaging, the shelf life of the liquid egg at 41°F (5°C) has been extended from days to weeks. It is anticipated that this process will allow movement of liquid whole eggs into a much larger market area.

Following pasteurization, eggs for freezing are packaged in 30-lb (14-kg) tin containers or in

TABLE 3. Nutrient composition of egg products, in grams per 100 g of eggs or egg products*

Nutrients and units	Liquid or frozen				Dehydrated		
	Whole egg	Egg white	Egg yolk, pure	Egg yolk, commercial	Plain whole	Stabilized whites	Plain yolk
Proximate							
Solids	24.5	12.1	51.8	44.0	96.8	93.6	97.2
Calories	152	50	377	313	600	388	692
Protein	12.0	10.2	16.1	14.9	47.4	79.1	39.2
Lipids	10.9	—	34.1	27.5	43.1	—	60.8
Ash	1.0	0.68	1.69	1.49	4.0	5.3	3.3
Lipids: fatty acids							
Saturated	3.67	—	11.42	9.16	14.51	—	20.35
Monounsaturated	4.60	—	14.67	11.80	18.18	—	25.64
Polyunsaturated	1.32	—	4.20	3.37	5.22	—	7.45
Cholesterol	0.37	—	1.18	0.96	1.48	—	2.11
Lecithin	2.32	—	7.20	5.81	9.16	—	12.84
Cephalin	0.46	—	1.42	1.15	1.82	—	2.54

*Except "Calories" category. 1 g = 0.035 oz.

SOURCE: Adapted from American Egg Board, *A Scientist Speaks About Egg Products* (1981), with the exception of cholesterol values, which are calculated from 1988 USDA values based on 213 mg per large egg.

TABLE 4. Specification guide for liquid, frozen, and dried egg solids*

Category	Liquid or frozen			Dried solids				
	Whites	Plain yolk	Whole	Angel whites	Flake albumen	Free- flowing stabilized whole	Free-flowing stabilized yolk	Fortified whole egg (tex) with carbohydrates
Moisture (max.), %	88.2	57	74	8	14	3	3	4
Fat, %	0.02 max.	25 min.	11 min.	None	None	40 min.	57 min.	31 min.
pH	—	—	—	7.0 ± 0.5	5.5 ± 0.5	7.5 ± 0.5	6.5 ± 0.5	7.5 ± 0.5
Protein (min.), %	10.5	16.0	12.5	80	75	44	30	27
Carbohydrate	—	—	—	None	None	None	None	30 max.
Color	—	SOP	SOP	Creamy	Clear yellow	NEPA 2–4	NEPA 2–4	—
Viable bacteria/g (max.)	50,000	10,000	10,000	5000	10,000	10,000	10,000	10,000
Yeast (max.)	10	10	10	10	10	10	10	10
Mold (max.)	—	10	10	10	10	10	10	10
Coliform/g (max.)	1000	10	10	10	10	10	10	10
Salmonella	Neg.	Neg.	Neg.	Neg.	Neg.	Neg.	Neg.	Neg.
Granulation, %	—	—	—	100 USBS–80	—	100 USBS–16	100 USBS–16	100 USBS–16
Additives	—	—	—	SLS 0.1% max	—	—	SS 2% max	—
Performance	SOP	SOP	SOP	SOP	SOP	SOP	SOP	SOP

*SOP = specified on purchase; neg. = negative; SS = sodium silicoaluminate; SLS = sodium lauryl sulfate; NEPA = National Egg Products Association light-to-dark scale (range 2–4); and USBS–80 and USBS–16 = U.S. Bureau of Standards screens for determining particle size. 1 g = 0.035 oz.

4–10-lb (2–4.5-kg) wax-coated cardboard containers and blast-frozen at −20 to −40°F (−29 to −40°C). Egg liquid to be distributed as liquid is cooled to less than 40°F (5°C) and packaged in containers ranging from bulk tanks to the small cardboard containers used for frozen eggs. Egg liquid for drying is cooled to less than 40°F (5°C) and either dried on the premises or transported in insulated tank trucks to an egg-drying plant. Table 4 lists the properties of commercial liquid, frozen, and dried egg products.

Egg drying. Both liquid egg white and egg yolk from the breaking plants are collected at the drying plants. Prior to drying, the egg liquids undergo extensive processing to ensure retention of the functional and organoleptic properties in the finished product.

Egg whites. All egg whites for drying must be pasteurized. The small amount of free glucose present in egg white must be removed to prevent the Maillard reaction, which causes brown discoloration and off odors in the dried product. Glucose is removed by the use of bacteria or yeast capable of utilizing glucose or by enzyme systems capable of converting the glucose to a nonreactive material, such as gluconic acid. *See* MAILLARD REACTION.

Several whipping aids (most commonly sodium lauryl sulfate) have been approved by the U.S. Food and Drug Administration for addition to egg white in order to retain its high whipping qualities following drying. Following appropriate adjustment of pH with lactic, citric, or other approved food acids, the liquid egg white is spray-dried in conventional dryers. It is essential that egg white dryers be provided with dust-collecting systems. Spray-dried egg white solids are normally packed in 150-lb (68-kg) fiber drums or 25–50-lb (11–23-kg) polyethylene-lined boxes for shipment to food manufacturers.

Egg white is also air-dried in pan or tunnel systems, in which the egg white is poured in a thin layer on trays. After appropriate treatment of the liquid, as described above, these trays are subjected to temperatures up to 130°F (54°C) until a pseudocrystalline product is obtained. These crystals, or flakes, are used in this form or ground into a fine powder for greater solubility. This type of product is used primarily in the confectionery industry.

Egg yolks. Liquid egg yolk, if not already pasteurized, is held at 140°F (60°C) for 3.5 min. Egg yolk may be spray-dried directly, or it may be subjected to a stabilizing treatment to remove the free glucose, as is also necessary with egg white. The glucose-oxidase enzyme system is used for this process almost exclusively.

Egg yolk must be dried to a final product containing 95% egg yolk solids; many kinds of spray-drying equipment are suitable for this process. Egg yolk is commonly converted to a free-flowing powder by the addition of small quantities of such anticaking agents as sodium silicoaluminate (less than 2%) or silicon dioxide (less than 1%). The ability to flow freely appears to be of particular importance to mayonnaise manufacturers.

Whole egg. Whole-egg solids are manufactured in a similar manner, as are many blends of whole egg, yolk, and added ingredients, to meet specific functional requirements. Common additives are carbohydrate products, which may be derivatives of either sucrose or corn syrup solids. These carbohydrates, when added prior to spray-drying, help retain the original foaming and emulsifying abilities of the liquid egg. Whole egg may be converted to a free-flowing product, as described above for egg yolk.

Frozen products. There is a variety of frozen liquid or solid products that have egg or egg parts as major ingredients.

Liquid. These products include frozen mixes for scrambled eggs, omelets, crepes, and quiches. The scrambled egg mix includes milk products and seasonings and is packaged in waxed paperboard boxes

ranging in size from half pint (225 ml) to 2 quarts (1820 ml). A special package primarily for institutional feeding is in a flexible film allowing the eggs to be cooked in the film bag by immersion in hot water at about 194°F (90°C) for times varying with size of the package and temperature at the time of immersion.

Omelet, crepe, and quiche mixes are packed in small, waxed paperboard boxes [half pint (225 ml) to 1 quart (908 ml)]. The package is thawed at 35°F (2°C) and the desired amount poured into an omelet or crepe pan or a quiche shell. After thawing, the mixed product should be kept refrigerated and used within 4 days.

Solid. Solid products include precooked frozen French toast; omelets, either plain or seasoned; crepes, either plain or filled; quiches; souffles; egg patties; fried eggs; diced, hard-cooked eggs; and long eggs.

The French toast is white bread soaked in liquid whole egg and fried to a golden brown color. It is packaged in both retail- and institutional-size packages. Omelets are prepared with whole eggs mixed with about 10% water and seasonings. For variation, other inclusions may be cheese, ham bits, bacon bits, onions, green peppers, or almost any other product desired by the consumer.

Crepes are prepared by using a very thin pancake batter especially rich in eggs. Plain crepes are packaged for retail or institutional use, with the user providing the filler. Crepes are also filled with fruit or meat sauces prior to packaging and freezing. Quiches and souffles are cooked, packed, and frozen for retail distribution or for specialized institutional use such as airline meals.

Other frozen egg products are scrambled egg mix cooked into a patty for sandwich use. Frozen fried eggs are also available for institutional use. The patties and fried eggs are especially suited to microwave reheating at the time of serving.

Hard-cooked diced eggs are usually cryogenically frozen to minimize a toughening of the dices of egg white that can occur when the products are frozen slowly. A unique product is the long egg. This is an extruded product with a center core, about normal egg yolk size in diameter, of cooked egg yolk covered by a uniform layer of hard-cooked egg white. In this product, a small percentage, usually less than 1%, of food gums or starches is added to the albumen to minimize the problem of toughening during freezing.

Refrigerated products. Hard-cooked eggs are prepared by heating shelled eggs to an internal temperature of 175°F (65°C) in hot water. The eggs are then cooled quickly and the shells are removed. The peeled eggs are immersed in 0.5–2% citric acid solution containing about 0.2% of sorbic acid or 0.2% of sodium benzoate as a mold inhibitor. The eggs and solution are stored at 35°F (2°C) and have a microbial shelf life of over 6 weeks.

A modification of the preservative solution including about 5% acetic acid (vinegar) yields pickled eggs. Red beet juice is frequently used with the acetic acid to impart a bright red color to the albumen layer of the hard-cooked egg. Pickled eggs have a refrigerated shelf life of several months.

Another product was introduced in the United States in 1983. These are hard-cooked eggs processed and treated so as to allow assurance of a 3-month usable shelf life. The hard-cooked egg shell is lacquer-coated with a colored solution to seal the egg. Even without refrigeration, these eggs can be kept for 30 days. They are desirable as colored Easter eggs, as ready-to-eat, or for camping trips.

Dry products. The largest use of dried egg products is in bakery and confectionery products. During certain holiday seasons, eggnog mix is popular. This is a mixture of dried milk and dried whole egg with sweeteners and flavoring added. William J. Stadelman

Nonnutritive Sweeteners

Nonnutritive sweeteners are not metabolized to any significant extent and hence do not add caloric value to foods. The U.S. Food and Drug Administration (FDA) defines nonnutritive sweeteners as "substances having less than two per cent of the caloric value of sucrose per equivalent unit of sweetening capacity." Nonnutritive sweeteners include saccharin, aspartame, alitame, cyclamate, acesulfame-K, thaumatin, stevioside, dihydrochalcones, chlorosucroses, L-sugars, and PS-99 and PS-100 (both dipeptide ester compounds). Nonnutritive sweeteners are of particular appeal to those individuals on diets or with diabetes, and to manufacturers interested in producing foods with lower caloric values. Dry foods represent about 30% of the overall market for nonnutritive sweeteners, and liquids account for about 70%. Because of the large economic potential that exists for use of nonnutritive, high-intensity sweeteners in the food and beverage market, a number of materials with such properties are under investigation and development.

Saccharin has been used widely throughout the world for over 80 years. Approximately 60% of the saccharin is consumed in soft drinks, 20% in other beverages and foods, and 20% as a tabletop sweetener. It is about 300 times sweeter than sucrose, is relatively inexpensive, has excellent stability in food systems, and is noncarcinogenic. One disadvantage to saccharin is that it leaves a bitter or metallic aftertaste, particularly when used at relatively high concentrations, which many find objectionable, and which is difficult to mask successfully. In 1977, questions concerning the safety of saccharin, when used in excessively large amounts, were raised by a Canadian animal (rat) study. The FDA decided to ban most of the food uses of saccharin by revoking its approval, except for use by diabetics and the obese, but Congress placed a moratorium on such a ban as a result of public reaction. *See* SACCHARIN.

Aspartame is the methyl ester of a dipeptide derived from aspartic acid and phenylalanine, two common amino acids. It can be produced by chemical, enzymatic, or genetic engineering processes. It is about 180–200 times sweeter than sucrose. Even though aspartame is metabolized, contributing

calories, it is considered a nonnutritive sweetener under the FDA definition, since it is more than 50 times sweeter than sucrose. Aspartame has a clean sweet taste with no bitter aftertaste. It is probably the nonnutritive sweetener most closely resembling sucrose in flavor quality and profile. The major problem with aspartame is its relative instability in aqueous solutions, where it hydrolyzes below pH 3 and cyclizes above pH 6. The instability in acid solution delayed its approval for use in soft drinks in the United States. Aspartame was given limited approval for food uses in the United States in August 1981. The approval was for use in dry products such as packets for use as a tabletop sweetener, tablets for hot beverages, instant coffee and tea, cereals, chewing gum, dessert mixes, and powdered drink mixes. In July 1983, approval was extended to carbonated beverages. Aspartame has found widespread use in many of these food products. It is relatively expensive compared to saccharin or sucrose. It has been a controversial sweetener because of questions about its safety, and it must carry a warning label for phenylketonurics. Aspartame has probably been the most studied of all food additives, particularly sweeteners. A patent has been issued for a heat-stabilized composition of aspartame, which could allow its use in baked goods and baking mixes, if approved by the FDA. *See* ASPARTAME.

Acesulfame-K is a crystalline synthetic derivative of acetoacetic acid. It is about 200 times sweeter than sucrose and has excellent stability to heat and low pH. Acesulfame-K was approved in the United States in July 1988 for use in dry goods, such as sugar substitute tablets; chewing gum; dry bases for beverages, instant coffee, and instant tea; dry, free-flowing sugar substitutes in package units; dry bases for gelatins, puddings, and pudding desserts; and dry bases for dairy product analogs. FDA approval will be sought for use in liquids and baked goods. Acesulfame-K is approved for use in 20 countries and is widely used in Europe. It is expensive when compared to saccharin or sucrose.

David R. Lineback

Quality Control

Quality is an important aspect of industrial food production. Quality may be defined as the composite of those characteristics having significance in determining the acceptance of a product unit in relation to the product's design specifications and consumer expectations. Thus, quality control implies a management system for the coordination and integration of all functions of quality such as quality control, quality assurance, and quality reliability necessary to maintain attributes at levels and tolerances acceptable to the consumer while minimizing costs for the producer.

Quality attributes of food products have been divided into three major categories, quantitative, hidden, and sensory. Quantitative attributes of food quality include yield, net weight, drained weight, package integrity, and the like. Hidden attributes of food quality include nutritive value, toxic substances, adulterants, and microbial status. Sensory attributes of food quality encompass those attributes which can be perceived and judged by consumers, and include color, size, shape, consistency, viscosity, texture, mouthfeel, odor, flavor, and defects.

The term food quality control indicates the activities and functions related to controlling the three major categories of food quality during manufacture. In food manufacturing, operating under product quality controls is essential to assure that products meet both company and government (Food and Drug Administration, U.S. Department of Agriculture Food Safety and Inspection Service and the Agricultural Marketing Service, state and local health departments) quality requirements as well as consumer expectations. Controls must be used to keep operational variations in check. This is the function of the quality control department.

A food plant quality-control system should fulfill the following responsibilities:

1. In cooperation with the research and development or procurement departments, or both, written specifications must be prepared for all raw materials, packaging materials, ingredients, supplies, and finished products. Moreover, in cooperation with these departments, written operating procedure specifications must be established for each unit operation of each manufacturing process in the plant.

2. Adherence to sanitation standards and compliance with FDA "Good Manufacturing Practices" must be assured. This includes plant and storage pest control, and plant, storage, and equipment cleanliness including microbiological cleanliness. Employee personal hygiene and work habits should be controlled to assure that foreign material, food safety, and microbiological standards are met.

3. Statistically valid sampling and evaluation procedures must be developed to monitor adherence to ingredient, packaging, raw material, process, unit operation, and finished-product specifications. The results of these evaluations must be recorded on suitable forms and transmitted to management. It is most important that out-of-specification (out-of-control) material be promptly identified and segregated and action be immediately implemented to correct out-of-control conditions, to minimize the production of defective units, and to rapidly resume the production of in-specification product.

Quality-control personnel work in laboratories conducting physical, chemical, and microbiological evaluations of raw materials, ingredients, packaging supplies, work-in-process, and finished product. They also work in the plant to evaluate processes, unit operations, sanitation levels, adherence to "Good Manufacturing Practices," and all other operating specifications. It is in these two functions that considerable reliance is placed upon statistical quality control, whereby statistical techniques are applied to quality evaluations for the purposes of scientific analysis and interpretation of data. Statistical quality-control functions include the selection of appropriate sample size and frequency, the use of attributes and variables sampling techniques and control charts, and the use of correlation, analysis

of variance, and evolutionary operations techniques. The methods, procedures, and selection of instruments utilized to measure product quality attributes are all responsibilities of quality control. *See* STATISTICS.

4. In cooperation with the research and development, production, and engineering departments, the quality-control system should evaluate all plant processes for hazard analysis of critical control points (HACCP). The wholesomeness and safety of foods must be assured by identifying and evaluating (using statistical methods) all raw materials and process unit operations in order to preclude the potential for adulteration or microbial contamination which could lead to a consumer health hazard or to inordinate economic loss. It is essential to conduct these evaluations as a safeguard against public health hazards.

5. The quality-control personnel interact cooperatively with regulatory agency personnel to assure that governmental food regulations are understood and followed. Regulations related to processes, "Good Manufacturing Practices," labeling, storage, and shipping of foods must be identified, communicated, and implemented.

6. The system must be able to develop and conduct or assist in an organized program to train supervisors and operators in the quality concept of management. This training has as its ultimate goal the quality consciousness of all plant workers. Additional educational efforts are required for quality-control personnel in order to assure that quality-control workers are versed in basic concepts and perform their duties in a standard, nondeviating manner.

7. Quality control departments should also assist production in improving yields, reducing salvage losses, and improving operating efficiency as well as review and respond to (with production) consumer complaints. Development of a sound recall plan to recover adulterated or defective product from marketing channels is an additional quality-control responsibility. *See* QUALITY CONTROL.

Donald V. Schlimme; Barry C. Frey

Food Packaging

Food packaging is a comprehensive system to ensure separation and protection of the contained food product or beverage from external physical, biological, and chemical environmental influences. In synergistically engineering the package material and equipment, consideration must be given to the product contents as well as to the total distribution environment to be experienced during the life cycle of the packaged food product. The objectives of the food packaging system are to provide protection for the products, to effect communications to the consumer, to impart convenience for the consumer, and to create package units containing multiples of individual packages to facilitate distribution through the supply chain.

Regulation. The basic regulations in the United States stem from the 1938 food and drug laws, amended in 1958 to include the Delaney clause on cancer-inducing additives, and since altered to facilitate approvals. The 1964 Truth in Packaging law detailed labeling on the packaging surface. Although much has been written and spoken about packaging laws, these are the two governing laws, with the basic Food and Drug law covering general safety and the Good Manufacturing Regulations covering low-acid foods. Hazard analysis critical control points (HACCP) and nutrition labeling have been added, the former for meats and the latter for all foods.

Labeling. The Truth in Labeling Act stipulates that food packages bear a uniform and accurate representation of the weight, volume, or count of the package contents. The law further provides for accurate word descriptions of the contents, the name and address of the manufacturer or distributor, and the quantity of the contents clearly stated. Not mandated by regulation but by retailers is the requirement that all food packagers contain a universal product code, or UPC.

Indirect additives. The Delaney Amendment has been interpreted to mean that any material that is a potential cancer producer that might migrate from a packaging material in contact with the food in any quantity is regarded as an indirect additive which must be prohibited. The interpretation and implementation of the Delaney Amendment has been manifested in rulings effectively banning poly(vinyl chloride) [PVC] as a liquor packaging material in 1973 and polyacrylonitrile as a carbonated beverage packaging material in 1977. The latter restriction was rescinded in 1984, and the former continues to be under review. The FDA has established limits based on quantities to be ingested plus probabilities of hazard.

Package disposal. The environmental protection movement that began in 1970 has resulted in several states passing legislation that restricts beverage packaging by requiring that packaging for carbonated beverages, beer, and sometimes other beverages bear a deposit. The purpose is to encourage the return of all carbonated beverage and beer containers so that they may be recycled, reused, or visibly removed. A secondary purpose is to encourage the consumer to dispose of the container properly. During the mid-1980s, several states and municipalities enacted laws to encourage recycling of packaging materials, such as funded development programs that forced consumers to separate trash by recyclable and nonrecyclable categories. Some states are attempting to tax packaging. A number of municipalities have passed laws prohibiting certain specific packaging materials, such as expanded (foam) polystyrene. These materials are used more for retail food service than in packaging engineered to protect food from a factory to a store. In addition, "sustainable" packaging—packaging materials derived from renewable resources—is being used. *See* FOAM.

The concern over the potential for upper-atmosphere ozone depletion by fluorocarbons and concurrent global warming has led to legislation

prohibiting such materials in packaging such as expanded polystyrene and in aerosol cans. Most manufacturers had already removed the fluorocarbon foaming agents and propellants before the legislation. *See* GLOBAL CLIMATE CHANGE; HALOGENATED HYDROCARBON; OZONE; STRATOSPHERIC OZONE.

Levels of packaging. The levels of packaging usually employed in classification systems are primary, secondary, and tertiary.

Primary packaging is the package in intimate contact with the food product and constitutes the first and usually principal gas–liquid barrier between the food product and its external environment. Primary packaging is generally the metal can, glass jar, plastic jar or bottle, plastic tray or cup, first wrap, inner pouch, and so on into which the food product or beverage is placed prior to being packaged for distribution.

Secondary packaging may consist of a unitizer or container that supplements the function of the primary container. In a unitizer, several primary packages are assembled to permit more efficient distribution. Thus, secondary packaging may include an outer shell paperboard carton, a plastic ring for six cans, plastic shrink film, or multiple packaging for cans or bottles, particularly carbonated beverage or beer cans or bottles, or for soup or seasoning pouches.

Tertiary packaging is the distribution package. Distribution packaging unitizes primary and secondary packaging to facilitate movement through distribution channels such as warehousing and transport. Distribution packaging in the United States is generally a corrugated fiberboard shipping case, kraft paper bundle, plastic shrink film, or some combination. Downstream of tertiary packaging are pallets, pallet stretch film wraps, and so forth.

Materials. A wide range of materials is employed to provide packaging in the United States. The selection of food packaging depends upon such factors as the protection required, economics, and marketing needs. The ability to optimize and combine materials generates multi-ply materials to perform a variety of packaging functions with a single structure.

Metal. Metal cans, usually cylindrical, are rigid; have good vertical compressive strength; resist impact, shock, and vibration; and are able to contain internal pressure or vacuum. To retard the volume decline in metal cans, they are being shaped (such as rectangles, bulbous, undulating, and truncated) to distinguish them on shelves, although shaped cans are more difficult to make and process. Metal cans may be printed directly in the flat before fabrication or in the round, or they may receive an outer paper label to provide the necessary and desired graphics. Metal for food cans has evolved from hot-dipped tin, to electroplated steel, to tin-free steel with chrome–chrome oxide coatings that resist corrosion.

Aluminum is the most widely used material for carbonated beverage and beer cans because of its lightweight and ease of fabrication.

The three-piece metal can (mostly steel) is formed from a flat blank that is wrapped into a cylindrical form. During the 1980s, soldered side seams were replaced by welded side seams to reduce the potential lead contamination hazard. Circular metal disk closures are mechanically clinched to both ends. For consumer convenience, full-panel easy-open can ends increasingly are replacing conventional can ends.

The far more widely used two-piece beverage aluminum can body for carbonated beverages and beer is fabricated from a flat blank or cup. Decoration is usually in the round on the open-ended can. The open-ended can is then transported to the food or beverage packager's line, where it is filled. There the can is closed by mechanical clinching with an aluminum closure. To compete with plastic bottles, narrow-neck aluminum bottles are being offered for both carbonated and noncarbonated beverages. Ends generally may be flat steel for food cans or aluminum with a die-cut to permit easy opening for beverage cans. Diameters of the ends can be reduced by necking-in of can bodies in order to reduce the mass of materials used and to approach bottlelike structures. Finishes may be screw-on/screw-off to more closely match plastic bottles. Although overwhelmingly used for internally pressurized beverages, two-piece aluminum cans are increasingly being used for still beverages such as juices and certain foods by incorporating liquid nitrogen as an internal vacuum compensator to prevent collapse of the light wall aluminum after hot filling and cooling, with subsequent condensation of the internal steam.

Aluminum foil and occasionally steel are also used for nonround containers such as shallow flat trays used for canning. The increase in surface area increases the rate of heat penetration into or out of the package and thus reduces the thermal input required for heat sterilization. Shallow-drawn steel cans with double-seam closures are used for hotel–restaurant–institutional packs. Shallow drawing of both rigid and semirigid aluminum is employed to produce trays in which food is packaged. Aluminum has been generally replaced by plastic, usually crystallized polyester (CPET) and plastic-coated paperboard for frozen and refrigerated foods.

The lightweight and inert properties of aluminum also foster its use as a barrier lamination material in flexible and some semirigid package structures. Pouches and bags employ thin sheets of 0.00028–0.0005 in. (0.0071–0.00127 mm) gage aluminum foil as the principal barrier material in laminations. In lamination form, the aluminum foil is protected against abrasion and pinholing by a coating or laminate, and also has a sealant coating on one face. Because of the relative impermeability of aluminum foil, the metal serves as the principal barrier against the passage of water vapor, oxygen, odors, liquid, other gases, grease, fat, and light.

Vacuum metallization of less expensive substrates, such as plastic film (polyester and oriented polypropylene), often is an alternative. In vacuum metallization, micrometer thicknesses of aluminum metal are deposited on the surfaces of the substrate material to impart significant barrier properties to

the substrate. Vacuum metallized surfaces also function as microwave susceptors to convert incident microwave energy into high-temperature heat for surface browning and crisping. Glass (or more accurately, silicon oxide, SiO_x) imparts a gas or water vapor barrier while maintaining both visible-light and microwave transparency. Similar barriers may be affected by aluminum–oxide coatings. *See* PHYSICAL VAPOR DEPOSITION.

Glass. Glass is used for food packaging because of its impermeability to gas, water vapor, and flavor, and its inertness to food products. Glass in bottle or jar form has high vertical compressive strength but relatively low impact strength in the usual configurations and methods of manufacture. Glass is manufactured from very low cost raw materials, but has a high energy input because the raw materials must be melted. Glass bottles may be molded into a wide variety of shapes and sizes to permit the food product to be packaged and distributed in bottles (narrow necks) or jars (wide mouths) that are short or tall.

Reductions in glass weight have been achieved by surface strengthening by methods which include annealing, hot end coatings, and combining with other materials such as roller/dip thin plastic to reduce the potential for shock and vibration shattering and chipping. The latter is used for nonreturnable bottles of carbonated beverages. *See* GLASS.

Closures for glass containers were originally largely metal, either mechanical crown for carbonated beverages or screw-on steel for jars of food products. Major closures for modern glass containers include twist-off crown, roll-on, screw cap, lug cap, tamper-evident/resistant, and snap-ons and press-ons, most of which are available in either metal or plastic with or without internal linings to achieve a barrier with the closure.

Paper. Paper is used for only a small number of packaging applications. Its principal application is as a component of large institutional- and industrial-size multiwall bags for dry granular contents. With high tensile strength and the ability to be coated with barrier plastic materials, unbleached kraft virgin paper is a substrate for manufacture of bags to contain large quantities of sugar, flour, nuts, dry pet foods, and so on. In order to ensure the package integrity, several walls of different materials are employed. These may be combined with other materials such as plastic film or coating to improve the water-vapor barrier properties.

Bleached kraft paper is employed as a laminating material to aluminum foil lamination pouches. The paper acts to protect the aluminum foil from abrasion and puncture, and as a printing surface. Glassine or highly calendered kraft paper is still used as a fat-resistant material, although pearlized plastic, such as oriented polypropylene, is the preferred material. *See* PAPER.

Paperboard. Folding and set-up cartons are primary or secondary packages made from paperboard in a converting process in which the paperboard is printed, scored, and cut. The resulting paperboard blank may then be glued into tubular or tray form to be erected on the packager's line for subsequent filling and closing. Paperboard imparts stiffness, a good printing surface, and some compressive strength. Paperboard made from recycled fibers has been manufactured commercially for more than 100 years and has been used extensively as a secondary packaging material. Recycled paperboard has gradually increased in importance. Usually, a greater mass of recycled material is required as compared to virgin material, because the recycling process reduces fiber length and therefore strength. Recycled paperboard cartons are used on the exterior of bag- or bottle-in-box structures to contain liquid in thin-wall interior plastic film containers, with strength imparted by the exterior paperboard shell. With a water-resistant coating such as polyethylene, virgin solid bleached sulfate paperboard has liquid-resistant characteristics that permit it to be employed as a primary or secondary packaging material for refrigerated beverages or liquid frozen food products. *See* RECYCLING TECHNOLOGY.

Corrugated fiberboard. Usually manufactured from virgin or virgin plus recycled fiberboard, corrugated fiberboard consists of three layers—two outer liners and an inner fluted medium. The fluting or corrugations provide vertical stacking strength in the direction of the flutes and cushioning against vibration and impact damage. Corrugated fiberboard sheets are printed, cut, and formed into cases; or in the reverse process, are printed after fabrication. An advance in corrugated fiberboard case manufacture has been the employment of rubber- or plastic-plate flexographic printing processes. Flexographic printing permits corrugated fiberboard to be used not only for distribution packaging but also for display and consumer packaging. Increasing quantities of corrugated fiberboard case liners are being laminated to offset printed paper to provide superior graphics and the basis for consumer-oriented cartons. Increasing numbers of corrugated fiberboard cases are being made with white outer linerboard which provides a printing surface superior to traditional brown outer liner. Often such outer linerboard is preprinted by highly effective processes to deliver excellent graphics. Improved compression strength in final cases is achieved by advanced papermaking such as cross dispersion of fibers on the Foudrinier wire, by multiple laydown of fibers, or by off-line water-resistant coating techniques.

Coated and laminated paperboard. Paperboard that is extrusion-coated with polyethylene on two sides or laminated with polyethylene to aluminum foil and further extrusion-coated with polyethylene can be fabricated into liquid-tight containers. Laminated materials are being used for aseptically packaged and hot-filled/refrigerated–distribution beverages. Polypropylene as an inner coating acts as a flavor barrier. Ethylene–vinyl alcohol copolymer coating is an oxygen barrier and a paperboard flavor barrier. A wide variety of polyethylene-coated paperboard liquid containers are employed for milk, dairy drinks, and fruit beverages. Lightweight liquid-tight containers with water-repellent and water-resistant

coatings are employed for packaging dairy and delicatessen items in retail operations. These are frequently fabricated into cups, tubs, or trays from virgin bleached sulfate or sulfite paperboard, coated with polyethylene.

Composite paperboard cans. Kraft paper and recycled paper are laminated to other materials which are then fabricated by spiral winding into multi-ply paperboard-body composite cans with metal, foil lamination, or paper ends. Spiral-wound composite cans have the same shape as metal cans but have lower barrier characteristics. Spiral-wound composite cans contain frozen concentrated orange juice, refrigerated doughs, snacks, and granular dry beverage products. Spiral-wound composite paperboard cans are closed by heat-sealing aluminum foil laminations to the flange rather than by traditional mechanical seaming with rigid metal. To some degree, convolute-wound single-wall cans may be used for some applications.

Plastics. Plastic materials are made from organic resins or polymers. Nearly all plastics used for food packaging are thermoplastic and thus are sealable by heat and pressure. A small quantity of thermoset plastics is used for bottle closures. Thermoplastics are favored for packaging because of their light weight.

Cellulosic films. The first of the so-called plastic materials were derived from regenerated or otherwise chemically altered cellulose materials. Commercially, cellophane was the most important.

Cellophane is a transparent film composed of a plasticized base sheet of regenerated cellulose. Essentially greaseproof and coated for moisture barrier, cellophane was employed in applications such as primary packaging for snacks, overwraps on cartons, bags, pouches, windows on cartons, and as a component of lamination. Because of the availability of more functional and less expensive plastic materials, the use of cellophane as a packaging material has declined dramatically. *See* CELLOPHANE.

Polyethylene. The major plastic employed for food packaging is polyethylene. Low-density polyethylene with a density range between 0.915 and 0.925 g/cm^3 is largely used for making film, with the range of film gages 0.00075 in. to about 0.005 in. (0.02 to 0.127 mm). Very low density polyethylene resins have desirable properties, such as toughness and heat seal. Low-density polyethylene is a very tough, water- and water-vapor–resistant film. Gas barrier characteristics of all of the polyolefin films are not good.

Because of polyethylene's effectiveness as a water vapor barrier, it is frequently used as a bagging material to contain water vapor–sensitive products. Low-density polyethylene film is a component of a multiwall bag system with paper as another major component. It is also used as a sole component of large bags. Low-density polyethylene is used in soft baked goods bags in which short-term water vapor resistance is required.

Low-density polyethylene resin also is used extensively as an extrusion coating resin on paperboard and aluminum foil to impart both water vapor resistance and heat sealability. Polyethylene resin imparts liquid tightness to milk and juice cartons.

Linear low-density polyethylene, produced by lower-pressure processes, has been replacing low-density polyethylene. Linear low-density polyethylene permits broader heat-seal ranges and greater toughness, so that thinner films may be blown or more durable heat seals on pouches may be made.

High-density polyethylene film. High-density polyethylene is a plastic resin with density above 0.93 g/cm^3. The resulting material in film or other form has an excellent water vapor resistance but, like all polyolefins, relatively poor gas transmission. High-density polyethylene film is stiffer than low-density polyethylene and has a relatively high operating temperature. In film form, this material is used to provide high-temperature properties to paper, paperboard, or aluminum foil. It is also used for injection and blow molding to produce semirigid containers such as milk and juice bottles, jars, trays, and tubs. Many food products are packaged in high-density polyethylene bottles, provided there are no high-temperature or special oxygen barrier requirements. A tougher version, high-molecular-weight high-density polyethylene, has permitted significant down-gaging of film in bag and can or drum liner applications. With the majority of semirigid bottles and jars being manufactured from high-density polyethylene, and with increasing recovery of so-called used-bottle infrastructure, this material has become a dominant plastic in food packaging.

Polypropylene. Polypropylene film is formed by casting through a slot extruder or, less frequently, by blowing through a circular die. Cast polypropylene, however, is not readily heat-sealable and is low-temperature-sensitive. If the film is biaxially stretched and heat-set, it acquires a much improved low-temperature impact resistance and strength. With appropriate coatings such as poly(vinylidene chloride) or acrylics from coextrusion or off-line coating, oriented polypropylene (OPP) film is employed for bag and pouch making, and overwrapping. Oriented polypropylene may be printed by the methods employed for polyethylene, and so its properties as a film in common gages around 0.001 in. (0.025 mm) are satisfactory for packaging of snacks, candy, bread, and baked goods.

Polypropylene has a relatively high operating temperature, 255°F (124°C), and can be employed as a sealing medium on laminations in which high-temperature resistance is required, such as retort pouches. Polypropylene is also employed as a semirigid bottle- or tray-molding material structural component in selected applications such as retort trays or hot-filled catsup. *See* POLYOLEFIN RESINS.

Poly(vinyl chloride). Although it has been under adverse pressure by environmentalists because of its chlorine content, poly(vinyl chloride) [PVC] is used in highly plasticized form for film. The film application is to permit oxygen entry to retain the red color or bloom on the surface of fresh meat packaged in retail store backrooms.

Poly(vinylidene chloride). Poly(vinylidene chloride) [PVDC], or saran, is an outstanding oxygen, water vapor, grease, and liquid barrier. Its price, however, is high, limiting its use to very thin films or coatings to provide a water vapor and gas barrier on the surfaces of other plastic or paperboard substrates. PVDC may be coextruded with other materials to provide a high gas barrier. PVCD is copolymerized with other monomers to enhance barrier properties. *See* POLYVINYL RESINS.

Ethylene–vinyl alcohol. Ethylene–vinyl alcohol (EVOH) resins are high-gas-barrier materials that are used in coextrusion and coating for both semirigid and film applications. Generally, these resins have moisture sensitivity that increases with decreasing ethylene content or higher gas-barrier properties.

Polyester. Poly(ethylene terephthalate), also known as PET or polyester, is a high-strength family of plastic materials with outstanding dimensional stability, good transparency, good heat resistance, and modest gas and water vapor barrier properties. Injection-stretch blow-molded polyester carbonated beverage bottles are used commercially despite having only intermediate carbonation retention and oxygen hold-out characteristics. The bottle is the most rapidly accepted new package in United States packaging history. Polyester is used for packaging a variety of other bottled food products with intermediate oxygen-barrier requirements, including peanut butter and salad dressing, as well as water.

Polyester is the basis for dual ovenable (microwave or convection oven) packaging in both thermoformed forms and as extrusion coatings on paperboard. *See* POLYESTER RESINS.

Nylon (polyamide). Nylon materials are good gas and aroma barriers which are highly thermoformable. Although somewhat moisture-sensitive, they are widely used in processed meat and cheese packaging as the forming web. Amorphous nylon and other variations were proposed as a blow-molding material in high-oxygen-barrier jars and bottles but were not widely adopted. *See* POLYAMIDE RESINS.

Laminations. Although single-layer materials have been employed for packaging for many years, laminations represent a major step forward in packaging. Lamination is defined as producing film materials independently and then assembling or bonding two or more films or paper or aluminum foil together. In this manner, various functional materials can be combined to produce a single flexible web with the desired functional properties and usually at a lower cost than a single material alone.

Two basic methods of lamination are extrusion and adhesion lamination. In the former, thermoplastic material is applied to a substrate and employed as an adhesive that holds together two materials such as paper and aluminum foil. The alternate method is to apply a thin film of adhesive and then bond the adhesive layer to the second layer.

Coextruded films. Coextrusion unifies thermoplastic melts in a single extension die to intimately bond optimum quantities of plastic materials and form either sheets or film. Coextrusion often requires intermediate extrudable adhesive tie layers.

Coextruded films are used extensively for snack-food pouch packaging. Polypropylene/polyethylene coextrusions are employed for candy bar wrapping and snack food pouches.

Coextruded sheet is employed for multicolor and surface-gloss tub packaging of dairy products. Coextruded sheet containing oxygen barrier materials, such as PVDC or ethylene–vinyl alcohol copolymer, are used for packaging highly oxygen-sensitive products such as processed meats and cheeses. Plastic beer bottles are being fabricated in multiple layers by coinjection of polyester and a physical (other polymers) oxygen barrier or a chemically active layer (polymer containing oxygen scavengers).

Other comaterial plastics. Other methods of producing layers of materials include extrusion followed by blow or injection molding, extrusion followed by thermoforming and blowing, or variations thereof. The resulting products can be bottles or jars containing several layers of materials, one of which is a structural material, a second of which might be a water vapor barrier, and the third perhaps an oxygen barrier. These high-barrier bottles have found application for packaging oxygen-sensitive products such as ketchup. *See* PLASTICS PROCESSING.

Active packaging. Active packaging senses changes in the internal or external package environment and adjusts the packaging function to the change. Among the more prominent active packaging systems are moisture control, oxygen scavenging, carbon dioxide generation, gas permeability control, and antimicrobial inhibition. Intelligent packaging senses change signals such as maximum temperature, time–temperature integration, location (radio frequency identification, or RFID), or spoilage signalers. Aaron L. Brody

Package selection. The selection of packages appropriate for a food product involves decisions that satisfy a matrix of requirements. The matrix is the combination of the functions of the package and the environments in which it must perform.

Functions. The functions of a package are containment, protection, convenience, and communication. Containment simply means that the package allows the product to be handled and transported. Packages protect food products from spoilage agents, which include vermin, microorganisms, dust, mechanical forces, odors, moisture (addition or removal), and oxygen. Thus, packaging prolongs the shelf life of the product, that is, the length of time during which the product retains an acceptable quality level. Convenience, a function that frequently drives the development of new types of packages, has two aspects: apportionment, and handling and storage. Apportionment involves the selection of package capacities convenient for the consumer or ultimate user. Cheese packaged in 1-lb-capacity (0.5-kg) pouches is much more convenient for consumer usage than cheese available in 50-lb (23-kg) wheels.

The communication function of a package involves the transfer of information and the selling

of the product in the package. The label on a food product provides quantifiable information concerning grade, quantity, ingredients, and nutrition. This is a transfer of information. The label also may have a picture of the product or some other image meant to convince the consumer to purchase.

Performance environments. The environments in which food packages must function are the production (food processing) environment, the distribution environment, and the consumer or other ultimate user environment, including the disposal of the package. The properties of a food package that allow its selection for use with a particular product begin with the process requirements of the food product. If the food product is to be retorted or pasteurized in the package, the package must remain inert through this high-temperature processing. The package must also be suited to the equipment available for filling, closing, labeling, and other operations at the food processor at an acceptable cost. The distribution environment (including transport, warehousing, and often display) introduces requirements to protect against physical damage and ambient (temperature, humidity, gas permeation) spoilage. Physical damage to a package (for example, a dented can) or to a product (such as a crushed snack food) is due to impact, compression, and vibration. Impact damage occurs when a package is dropped or, in a railcar, when cars are coupled together. Compression damage generally occurs in warehouse storage or in transport, when heavy loads are placed upon a package, often by stacking pallets. Vibration damage occurs during transport because of the vibration of a truck, train, ship, or airplane. In vibration damage, packages or products are rubbed together and damaged by abrasion. Ambient conditions, such as high moisture or low oxygen, similarly cause food spoilage, either independently or in combination with physical abuse. High moisture levels in a warehouse, for example, diminish the strength of corrugated boxes, allowing the bottom boxes on a pallet to be crushed by those above. Oxygen in air causes rancidity spoilage of fats and oils. These foods need to be protected from the oxygen in the ambient environment. Other foods, such as fresh vegetables, require available oxygen for respiration. Holes are cut in their packages to allow oxygen availability.

Typical examples. The selection of a package for a particular food product begins with knowledge of the properties of the food relevant to maintenance of its quality. The pH of the food, its susceptibility to microbial spoilage, its interactions with oxygen and moisture, and the chemical reactions diminishing quality need to be considered. Thereby, some packaging materials might be ruled out. Paper and other oxygen-permeable materials would be ruled out for oxygen-sensitive foods. Food processing steps after packaging (pasteurizing, retorting) will rule out some other materials.

The selection of the appropriate polymer for flexible packaging of a particular food generally requires thorough analysis. The primary differences among resins concern cost, strength, barrier, haze, and heat sealability. The simple commodity resins (polyethylene, polypropylene) will have the lowest cost per pound, the complex barrier resins (EVOH and PVDC) the highest. Strength values between the weakest and strongest resins cover an order of magnitude; oxygen barrier values differ by 5 orders of magnitude. A slight difference in haze (1% versus 3%) is noticeable to the consumer. For each type of resin, suppliers offer several grades, each with its own combination of properties. Often, a particular grade will perform best on a particular packaging line.

Peter J. Vergano

Nutrition Labeling

The nutrition label provides consumers with standardized information on the primary nutritional characteristics of food products. The nutrition label can serve as a guide in the planning of a healthy diet. In the United States, the Nutritional Labeling and Education Act of 1990 resulted in the formulation of new nutritional labeling regulations by the Food and Drug Administration and the Department of Agriculture. Nutrition labels are required on most packaged foods. For some bulk foods at retail, a nutrition label must be displayed at the point of purchase. The use of nutrition labels is voluntary for some foods, including raw meats, poultry, fish, fruits, and vegetables.

The nutrition label in the United States is headed by the words "Nutrition Facts." The serving size, stated in both the metric system and the U.S. Customary System, and the number of servings per container are provided. The serving sizes for many foods were established in the federal regulations and must be used for the calculation of the number of servings per container. Calories per serving and the number of calories from fat per serving are provided next. Dietary guidelines suggest that consumers obtain no more than 30% of calories from fat, and the percentage can easily be calculated from these two pieces of information for each food.

A list of the nutrients and other dietary components most important to health follows. The list includes total fat, saturated fat, cholesterol, sodium, total carbohydrate, dietary fiber, sugars, and protein. The amount of each per serving (in metric units) and the percent daily value provided by a single serving for each are given on the label, with the exception of sugars and protein where only the amount per serving is provided. Certain other components of the food can be listed on a voluntary basis, including polyunsaturated fat, monounsaturated fat, potassium, soluble and insoluble fiber, sugar, alcohols, and other carbohydrates.

The percent daily value, as derived from the daily reference value (DRV), is particularly helpful in devising healthful eating strategies. On the lower part of the nutrition label, the DRVs for fat, saturated fat, cholesterol, sodium, total carbohydrate, and fiber are provided in metric weight units for two typical human daily diets, a 2000-calorie diet and a 2500-calorie diet. These amounts are established by regulation and are consistently provided on all packaged

foods bearing these nutrition labels. For example, the DRVs for a 2000-calorie diet are 65 g for total fat, 20 g for saturated fat, 300 mg for cholesterol, 2400 mg for sodium, 300 g for total carbohydrate, and 25 g for dietary fiber. For total fat, saturated fat, cholesterol, and sodium, these DRVs would represent the maximum amounts that are desirable in a 2000-calorie diet. The percent daily value that appears on the label adjacent to each nutrient or other component is calculated by dividing the amount of that component by the DRV of that component for a 2000-calorie diet. For example, a food product with 10 g of saturated fat per serving would contain 50% of the DRV for saturated fat on the basis of a 2000-calorie diet. To calculate appropriate healthful eating strategies, individual consumers should adjust the DRVs to fit their own ideal caloric intakes and calculate the percent daily value on that adjusted basis. Calculation of the percent daily value as adjusted for the consumer's ideal caloric intake will show how each individual food fits into the overall daily diet. Food choices can also be easily compared for their amounts and contributions to the DRV on a per-serving basis. Such comparisons allow consumers to make intelligent choices about the composition of their diets based on their own individual needs.

Following the list of major nutrients and other dietary components is a listing of several of the most important micronutrients in the diet—vitamin A, vitamin C, calcium, and iron. Their levels are declared as a percent daily value based upon the Reference Daily Intake (RDI), which is identical to the U.S. Recommended Dietary Allowances (RDA) as established for adults and children over the age of 4. The RDA values were established by the Food and Nutrition Board of the National Academy of Sciences, and they represent an amount of the micronutrient that will be sufficient to prevent any signs of deficiency diseases. Other vitamins and minerals can be listed voluntarily. However, their listing is required in the case of intentionally fortified food products, where a percent daily value must be given also.

The nutrition label provides, at the bottom, the comparative number of calories per gram provided by fat (say, 9 calories per gram), carbohydrates (4 calories per gram), and protein (4 calories per gram). This information explains how calories per serving can be calculated by knowing the amounts of fat, carbohydrates, and protein per serving.

Most packaged foods display nutrition labels containing all of the information described above. In a few cases, simplified labels are allowed. Examples include foods in small packages and foods with insignificant levels of some components, for example, soft drinks.

Steven L. Taylor

Bibliography. E. D. Aberle et al., *Principles of Meat Science*, 4th ed., 2001; M. Ash and I. Ash (eds.), *Handbook of Food Additives*, 1995; A. Brody, *Active Packaging Beyond Barriers*, 2002; A. Brody and J. Lord, *Developing New Food Products for a Changing Marketplace*, 2000; A. Brody and K. Marsh, *The Wiley Encyclopedia of Packaging Technology*, 2d ed., 1997; R. P. Carpenter, D. H. Lyon, and T. A. Hasdell, *Guidelines for Sensory Analysis in Food Product Development and Quality Control*, 2d ed., 2000; H. Charley, *Foods: A Scientific Approach*, 3d ed., 1997; F. Clydesdale and F. J. Francis, *Food, Nutrition, and Health*, 1985; J. J. Connell (ed.), *Control of Fish Quality*, 4th ed., 1995; P. M. Davidson et al. (eds.), *Antimicrobials in Foods*, 3d ed., 2005; R. B. Fast and E. F. Caldwell (eds.), *Breakfast Cereals and How They Are Made*, 1990; Food and Drug Administration (U.S.), *Food Labeling Guide*, 1999; W. A. Gould, *Tomato Production, Processing and Technology*, 1992; W. A. Gould and R. W. Gould, *Total Quality Assurance for the Food Industries*, 3d ed., 2001; R. C. Griffin, S. Sacharow, and A. L. Brody, *Principles of Package Development*, 2d ed., 1993; B. Holland (ed.), *Fish and Fish Products*, 1994; J. M. Jay, M. J. Loessner, and D. A. Golden, *Modern Food Microbiology*, 7th ed., 2005; W. A. Jenkins and J. P. Harrington, *Packaging Foods with Plastics*, 1991; J. M. Jones, *Food Safety*, 1992; M. J. Kirwan (ed.), *Paper and Paperboard Packaging Technology*, 2005; A. Lopez, *A Complete Course in Canning*, Books 1 and 2: *Basic Information on Canning*, 12th rev. ed., 3 vols., 1987; M. McWilliams, *Food Fundamentals*, 8th ed., 2006; S. A. Matz, *Chemistry and Technology of Cereals as Food and Feed*, 2d ed., 1991; S. A. Matz, *Snack Food Technology*, 1992; M. Mathlouthi, *Food Packaging and Preservation*, 1994; N. H. Mermelstein, A new era in food labeling, *Food Technol.*, 47(2):81–86, 94, 96., 1993; N. N. Potter and J. H. Hotchkiss, *Food Science*, 5th ed., 1995; H. Reuter, *Aseptic Packaging of Food*, 1989; G. L. Robertson, *Food Packaging Principles and Practice*, 2d ed., 2005; D. B. Sauer (ed.), *Storage of Cereal Grains and Their Products*, 4th ed., 1992; S. Selke, *Understanding Plastics Packaging*, 1997; R. P. Singh and D. R. Heldman, *Introduction to Food Engineering*, 3d ed., 2001; J. Smith (ed.), *The Technology of Reduced-Additive Foods*, 2d ed., 2004; W. J. Sultan, *Practical Baking*, 5th ed., 1990; R. T. Toledo, *Fundamentals of Food Process Engineering*, 3d ed., 2006; D. H. Watson (ed.), *Food Chemical Safety: Contaminants*, 2001; D. H. Watson (ed.), *Food Chemical Safety*, vol. II: *Additives*, 2002; R. B. H. Wills, *Postharvest: An Introduction to the Physiology and Handling of Fruit and Vegetables*, 4th ed., 1998.

Food microbiology

A subdiscipline in the field of microbiology concerned with the study of bacteria, fungi, and viruses that grow in or are transmitted by foods. While bacteria are frequently associated with food spoilage and food poisoning, some species preserve foods through fermentation or produce food ingredients. Food microbiology is a broad field that can include not only microbiology but also sanitation, epidemiology, biochemistry, engineering, statistics, and mathematical modeling.

Pathogens and spoilage organisms. Some people dismiss food poisoning as a minor annoyance. In

reality, the suffering and economic losses stemming from food-borne pathogens are substantial, but they are often hidden. Annual economic losses from food-borne pathogens are extremely high. *Salmonella*, which cause an average of 40,000 cases yearly and 2000–3000 deaths in the United States, are responsible for about a third of these losses. Individual outbreaks of food-borne diseases can affect thousands of people. Many outbreaks are predictable and preventable through good sanitation, preservatives, thermal processing, and refrigeration. More than half, however, are of unknown etiology are poorly understood, and may be caused by so-called new pathogens. *See* FOOD POISONING.

Historical pathogens. In the 1960s, most food-related illnesses were attributed to one of five major groups of pathogenic bacteria. These were associated with particular foods, commodities, or processes and were classified as infectious or toxin-producing. These five groups, described below, remain major causes of food-borne illness. *See* BACTERIA.

Salmonella and Shigella. The primary infectious bacterium associated with foods is *Salmonella*. These organisms cause gastroenteritis with symptoms of fever, diarrhea, and vomiting 12–36 h after ingestion. Salmonellosis is usually self-limiting, but it can be fatal in the old, young, or medically compromised individuals. *Salmonella* are commonly found on meats, especially poultry and eggs. *Salmonella* are easily killed by cooking. However, items contacted by the contaminated raw meat can transfer the *Salmonella* to food that is ready to eat (cross contamination) and cause illness. The seasonal increase in *Salmonella* isolations illustrates how food-borne illness increases in warm summer months.

Shigella are related organisms which produce a similar infectious syndrome. They are usually transmitted by a fecal-oral route or through feces-contaminated water rather than through foods.

Clostridium botulinum. The most dreaded toxin-producing organism is *Clostridium botulinum*. It excretes a potent neurotoxin that causes weakness, double vision, slurred speech, paralysis, and often death if ingested. The vegetative reproductive form of *C. botulinum* is heat-sensitive, lives only in the absence of air, does not compete well with other bacteria, and is rarely a problem in fresh foods. *Clostridium botulinum* spores are killed only through severe heating, such as in canning.

Historically, botulism has been associated with foods canned at home. If canned foods receive inadequate heat processing, competing bacteria are killed, air is expelled, and the botulinal spores germinate. Fortunately, botulinal toxin is often destroyed by heat when the food is cooked before serving; hence the standard advice is to boil home-canned foods before eating.

Modern commercial canning is designed to destroy *C. botulinum* spores. Reported outbreaks of botulism caused by pot pies, potatoes, and fried onions have been caused by temperature abuse, that is, the holding of foods at warm temperatures that promote bacterial growth. *Clostridium botulinum* can also be a problem in processed meats, such as hams and sausages. In this case, its growth is controlled through the use of nitrite, salt, and refrigeration. One type of *C. botulinum* is associated with fish. *See* BOTULISM; TOXIN.

Clostridium perfringens and Bacillus cereus. These are spore-forming, toxin-producing bacteria that cause illness when foods are heated enough to kill competing bacteria but not enough to kill the spores. When large volumes of foods are prepared, cooked, and then kept warm until they are served, spores can germinate. In the case of *C. perfringens*, which is associated with meats, the ingested cells release toxin in the digestive tract, resulting in cramps and diarrhea. *Bacillus cereus*, found in meats, dried foods, and rice, produces two different types of toxins: the diarrheal toxin, which has an etiology similar to *C. perfringens,* and the emetic (vomiting) toxin, which causes symptoms similar to those produced by staphylococcal toxins.

Staphylococcus aureus. This bacterium produces toxins that are very resistant to heat. *Staphylococcus aureus* is found in the nose and throat of many healthy people and is transferred to food by inadequate hygiene. When foods are temperature-abused, the bacteria grow and produce toxin. Subsequent heating of the food kills the bacteria but does not inactivate the toxin. The toxin causes severe vomiting and diarrhea from $^1/_2$ to 4 h after ingestion. The microorganism grows well at salt and sugar concentrations that inhibit many competing bacteria. Foods high in protein, such as cured meats, custards, and cream-filled bakery goods, pose special hazards for staphylococcal food poisoning. *See* STAPHYLOCOCCUS.

Microbial ecology of foods. Modern food microbiology views foods as habitats where different organisms compete for survival. The fact that there are 250 genera of bacteria and that only 25 of these (8 pathogenic) are found in foods suggests that foods provide unique ecological niches. Viruses do not reproduce in foods and are not competitors in this sense (the food acts only as a carrier). Yeasts and molds usually grow more slowly than bacteria and are rarely a problem in foods that support bacterial growth. *See* FUNGI; YEAST.

Bacteria reproduce by binary fission; it takes only 20 doublings for one cell to yield more than 1 million cells. In environments where the doubling time is short, this occurs quite rapidly. Many preservation methods alter foods' environmental conditions in order to slow microbial growth. *See* BACTERIAL GROWTH.

Temperature. The most important environmental condition is temperature. Most food-borne pathogens are mesophiles; that is, body temperature is optimal for growth. With a doubling time of 20 min at 98.6°F (37°C), one bacterium generates 1 million progeny in less than 7 h; at 32°F (0°C) the doubling time increases to 1200 min and the 1 million cell count is not reached for 16 days. Keeping hot foods hot (>145°F or 63°C) and cold foods cold (<45°F or 7°C), combined with rapid heating and cooling to

get rapidly beyond the growth-promoting temperature range (45–145°F or 7–63°C), prevents most foodborne illnesses. Psychrophylic (cold-loving) bacteria such as *Listeria monocytogenes* are exceptions.

Acidity. A food's acidity, quantified as pH, is another major environmental factor. The pH range for bacterial growth is 4–9, with fastest growth at neutrality (pH 7). Changing a food's acidity can change the rate of bacterial growth. Meats, fish, poultry, and most dairy products are near pH 7, which is ideal for bacterial growth; fermented foods and fruits have pH less than 4. Many yeasts and molds grow in acidic environments and spoil acidic foods. The pH value of 4.6 has special significance because *C. botulinum* can grow and produce toxin above this value. Canned foods with pH above 4.6 are legally classified as low-acid and must be processed in retorts under steam at 240–280°F (116–138°C) to kill *C. botulinum* spores. Foods with pH below 4.6 are legally high-acid and are processed in open pans of boiling water. In this case, *C. botulinum* need not be killed because it cannot grow at low pH. *See* PH.

Water activity. The amount of water available for microbial growth, that is, water activity (a_w), is the third major factor influencing microbial competition. Water activity is the equilibrium relative humidity generated by a food in a closed chamber divided by 100 to give a 0 to 1.00 scale. Salad dressings and honey, which both contain 50% water, are microbiologically quite different. The dressing separates into a 100% free-water phase ($a_w = 1$) and supports bacterial growth, while the sugar in honey binds water so tightly that it is unavailable for microbial growth. Most bacteria grow only at $a_w = 0.90$–1.00. Fresh meats, vegetables, fruits, and perishable foods have water activity in this range. Most yeasts can grow at slightly lower values. *Staphylococcus aureus* is the pathogen most insensitive to water; it grows at $a_w = 0.86$. Since no pathogenic bacteria grow below $a_w = 0.85$, this value has special significance in the regulations defining low-acid foods. Foods having an a_w value below 0.85 are legally considered high-acid, regardless of their pH. Most molds grow at a_w values as low as 0.8 and compete well in foods such as flour, cakes, beans, rice, and cereals. Some xerophilic molds and yeasts grow at a_w values as low as 0.6. Dehydrated foods, with even less available water, are completely recalcitrant to microbial spoilage.

Oxygen. Oxygen can be favorable, neutral, or inhibitory to bacterial growth. In one process, foods are first vacuum-packed to inhibit aerobic spoilage organisms and are then partially cooked. This environment is perfect for anaerobic spore-forming microorganisms. However, the Food and Drug Administration prohibits the use of this process because of the potential botulinal hazard.

Preservatives. Chemical preservatives also render food environments unsuitable for microbial growth. The oldest preservative is common table salt; at very high levels, it produces water activities that are inhibitory to microbial growth, although many organisms are inhibited by as little as 3% salt. Nitrites are used in cured meat as anticlostridial agents. Acetic, lactic, citric, benzoic, and propionic acids and sodium diacetate can also be added to foods as microbial inhibitors. Considering the low levels used, the long history of safe use, and the consequences of microbial growth, the risk:benefit ratio associated with chemical preservatives is very low.

Multiple barriers. Consumer preferences for fresh and natural foods make it difficult to alter any one environmental factor enough to inhibit microbial growth, and so the trend is to use multiple barriers, or hurdles. This approach employs several inhibitors at suboptimal levels. For example, clostridia may be inhibited by 7% salt at pH 7.0 or 0% salt at pH 4.6, but a meat treated this way would be unacceptable, tasting either salty or acidic. However, 3% salt at pH 6 in the presence of nitrite at a concentration of 125 parts per million provides multiple barriers sufficient to inhibit the bacteria and not impair flavor.

Emerging pathogens. The demand for longer shelf life in refrigerated foods combined with their increased popularity has caused renewed concern about psychrophilic pathogens, such as *Yersinia enterocolitica*, and enterotoxigenic organisms, such as *Escherichia coli* and *Listeria monocytogens*. These bacteria grow most rapidly at 59–86°F (15–30°C) and, at refrigerated temperatures, can succesfully compete with the normal mesophilic bacteria, thus limiting the shelf life of refrigerated foods. *Listeria* cause special concern because they infect women and their unborn children preferentially. *See* ESCHERICHIA; LISTERIOSIS; YERSINIA.

Campylobacter jejuni is a pathogen that is responsible for more illnesses than *Salmonella* and *Shigella* combined. Ingestion of relatively small numbers can cause diarrhea, cramps, and nausea. This organism is microaerophilic (requires 5–10% oxygen), is relatively fragile, and shows a seasonal pattern of outbreaks similar to *Salmonella*. It is associated with raw meats and unpasteurized milk, and can be controlled by pasteurization, heating, and good sanitation.

Analytic approaches. Microbial analysis of foods frequently requires "zero defects" in the absence of 100% testing. Legally, ready-to-eat foods must be free of *Salmonella*. This demands that the food microbiologist be able to detect one *Salmonella* among millions of innocuous bacteria in a pound of food. Moreover, all of the food cannot be tested because microbial analysis is destructive. Therefore, statistical sampling plans determine how many samples must be tested to have confidence that the whole lot is free of *Salmonella*.

In the classical methods for counting microorganisms, a food or its hemogenate is highly diluted so that only 30–300 cells are transferred to growth media. After 2–10 days, each cell grows into a colony, and these are counted and multiplied by the dilution factor to estimate the number of cells in the food. Automated methods have been developed that measure growth products, bacterial deoxyribonucleic acid (DNA), or specific toxins; these methods dramatically reduce the analysis time and are rapidly replacing the petri-dish method.

A procedure known as hazard analysis critical control points (HACCP) can replace much postproduction testing. This technique examines a food, its ingredients, and its processing to identify points critical to safety. These points are then heavily monitored during production; if they are maintained, a safe product results.

Beneficial food-borne organism. When certain bacteria grow in foods, they produce desirable flavors and textures, and may also inhibit pathogenic organisms. Most of these bacteria belong to the genera *Streptococcus*, *Lactobacillus*, *Leuconostoc*, *Pediococcus*, or *Micrococcus*. They are used to make fermented dairy products, meats, and vegetables, and to preserve food by converting the sugars needed by competing microbes to lactic acid, which inhibits their growth. These lactic acid bacteria are unusually tolerant of acidic environments. *Acetobacter* and *Gluconobacter* are used in the production of vinegar. Yeasts, usually *Saccharomyces*, which produce ethanol and carbon dioxide, are used in the processes of brewing and baking.

Lactid acid coagulates casein, the major protein in milk, and this process is used to manufacture cheese. During the aging of cheese, bacterial enzymes generate characteristic flavors, allowing a wide variety of products to be made by using many different bacteria. The bacteria used can be indigenous to the milk, derived from a previous fermentation, or added as pure cultures. *See* CHEESE; FERMENTATION.

Until the late 1960s, staphylococcal food poisoning was a major problem in certain meat products, such as bologna, pepperoni, and salami. Since acid production by indigenous bacteria is often unpredictable, it is now recommended that defined starter cultures be used to ensure that sufficient acid is produced early enough to prevent staphylococcal growth.

A novel use for starter cultures of lactic acid bacteria is to prevent botulinal growth in bacon. A small amount of culture and sugar are added to the cured meat; if the bacon is temperature-abused, the lactic acid bacteria grow and produce acid to inhibit botulinal growth.

Certain vegetables are preserved by fermentation. Pickles are made by fermenting cucumbers; olives and many oriental foods are also fermented. During sauerkraut fermentation, the addition of 2.5% salt (by weight) to shredded cabbage selects for the growth of *Leuconostoc mesenteroides*, which stop growing when acid levels reach about 0.067%. This environment favors *Lactobacillus plantarum*, which produces acid to levels of 1.25%, which is tolerated by *L. brevis*, and this bacterium brings the product to a final acidity of 1.7%.

Biotechnology. Advances in molecular biology have generated interest in applications to food processing.

The most important contribution of biotechnology to food microbiology is the production of probes that detect pathogenic organisms much faster than conventional methods. For example, conventional methods require 5 days to confirm the presence of *Salmonella* in foods; probes that detect *Salmonella*-specific DNA or antigens can give similar results in 2 days.

The dairy industry has benefitted from advances in biotechnology by acquiring the ability to determine the genetic basis for the bacterial metabolism of lactose in milk and to stabilize it. In addition, enzymes that accelerate the aging of cheese have become commercially available, making it possible to produce a cheese with the taste of 9-month-old cheddar in just 3 months. *See* BIOTECHNOLOGY; ENZYME; FOOD ENGINEERING; FOOD MANUFACTURING; FOOD PRESERVATION; GENETIC ENGINEERING; VIRUS.

Thomas J. Montville

Bibliography. W. C. Frazier and D. C. Westhoff, *Food Microbiology*, 4th ed., 1988; International Commission on Microbiological Specifications for Foods, *Microbial Ecology of Foods*, vol. 1: *Factors Affecting Life and Death of Microorganisms*, 1980; J. M. Jay, *Modern Food Microbiology*, 6th ed., 2000; T. J. Montville (ed.), *Food Microbiology,* vol. 2: *New and Emerging Technologies*, 1987; J. L. Oblinger (ed.), Bacteria associated with foodborne diseases, *Food Technol.*, 42(4):181–200, 1988; B. P. Wasserman, T. J. Montville, and E. L. Korwek, Food biotechnology, *Food Technol.,* 42(1):133–146, 1988.

Food poisoning

An acute gastrointestinal or neurologic disorder caused by bacteria or their toxic products, by viruses, or by harmful chemicals in foods.

Bacteria may produce food poisoning by three means: (1) they infect the individual following consumption of the contaminated food; (2) as they grow they produce a toxin in food before it is consumed; or (3) they produce toxin in the gastrointestinal tract after the individual consumes the contaminated food.

Infectious bacteria associated with food poisoning include *Brucella*, *Campylobacter jejuni*, enteroinvasive *Escherichia coli*, enterohemorrhagic *E. coli*, *Listeria monocytogenes*, *Salmonella*, *Shigella*, *Vibrio parahaemolyticus*, *Vibrio vulnificus*, and *Yersinia enterocolitica*. These organisms must be ingested for poisoning to occur, and in many instances only a few cells need be consumed to initiate a gastrointestinal infection. *Salmonella* and *C. jejuni* are the most prevalent causes of food-borne bacterial infections. *See* YERSINIA.

Staphylococcus aureus and *Clostridium botulinum* are bacteria responsible for food poisonings resulting from ingestion of preformed toxin. *Staphylococcus aureus* produces heat-stable toxins that remain active in foods after cooking. Staphylococcal poisoning accounts for 20–40% of food-borne outbreaks reported in the United States. *Clostridium botulinum* produces one of the most potent toxins known. Botulinal toxin causes neuromuscular paralysis, often resulting in respiratory failure and death. *See* BOTULISM; STAPHYLOCOCCUS; TOXIN.

Food-poisoning bacteria that produce toxin in the gastrointestinal tract following their ingestion include *Bacillus cereus*, *Clostridium perfringens*, enterotoxigenic *E. coli*, and *Vibrio cholerae*. *Bacillus cereus* and *C. perfringens* are spore-forming bacteria that often survive cooking and grow to large numbers in improperly refrigerated foods. Following ingestion, their cells release enterotoxins in the intestinal tract. Enterotoxigenic *E. coli* is a leading cause of travelers' diarrhea. *See* DIARRHEA; ESCHERICHIA.

Viruses that cause food-borne disease generally emanate from the human intestine and contaminate food through mishandling by an infected individual, or by way of water or sewage contaminated with human feces. Hepatitis A virus and Norwalk-like virus are the preeminent viruses associated with food-borne illness. *See* HEPATITIS.

Chemical-induced food poisoning is generally characterized by a rapid onset of symptoms which include nausea and vomiting. Foods contaminated with high levels of heavy metals, insecticides, or pesticides have caused illness following ingestion. *See* FOOD MICROBIOLOGY; MEDICAL BACTERIOLOGY; TOXICOLOGY.

Mike Doyle

Bibliography. F. L. Bryan, *Diseases Transmitted by Foods*, 2d ed., Centers for Disease Control, Atlanta, 1982; M. P. Doyle, *Foodborne Bacterial Pathogens*, 1989; M. P. Doyle and E. M. Foster, Bacterial agents of foodborne disease, *Wis. Med. J.*, 80:24–27, 1981; Y. H. Hui, M. Gorham, and Cliver (eds.), *Foodborne Disease Handbook: Diseases Caused by Hazardous Substances*, 1994.

Food preservation

The branch of food science and technology that deals with the practical control of factors capable of adversely affecting the safety, nutritive value, appearance, texture, flavor, and keeping qualities of raw and processed foods. Since thousands of food products differing in physical, chemical, and biological properties can undergo deterioration from such diverse causes as microorganisms, natural food enzymes, insects and rodents, industrial contaminants, heat, cold, light, oxygen, moisture, dryness, and storage time, food preservation methods differ widely and are optimized for specific products. Apart from application of a single food preservation method, a combination of methods may be used to improve the safety and storage stability of foods. This concept is currently referred to as hurdle technology; it is also known as food preservation by combined methods, combination preservation, or combination techniques.

Food preservation methods involve the use of heat, refrigeration, freezing, concentration, dehydration, radiation, pH control, chemical preservatives, and packaging applied to produce various degrees of preservation in accordance with the differing use patterns and shelf-life needs of unique products.

Perishability of many food materials was somewhat controlled long before the principles of modern food preservation were understood. Cheese and other fermented milk products, wine, sauerkraut and pickles, smoked meats and fish, dried and sugared fruits, and numerous other foods had their beginnings in attempts to extend the storage life of the basic commodities from which they were derived, but results were often disappointing.

Optimum food preservation must eliminate or minimize all of the factors that may cause a given food to deteriorate, without producing undue adverse effects. This can be especially difficult since the components of foods may be more sensitive to preservation treatments than the highly resistant bacterial spores and natural food enzymes targeted for destruction. Many nonbiological causes of food deterioration must be prevented also; these include oxygen, light, and loss of moisture.

While traditional methods of food preservation, including heating, freezing, drying, refrigerating, and acidifying, are still widely used in the food industry, new food preservation methods are actively being researched. These include physical methods such as high hydrostatic pressure, pulsed electric fields, ohmic heating, and ultrasound. Such methods have been shown to be effective in preserving food with minimal effects on nutrients and sensory quality that might otherwise be destroyed by traditional food preservation methods such as heating, drying, and acidification.

Heat. Thermal processes to preserve foods vary in intensity. True sterility to ensure total destruction of the most heat-resistant bacterial spores in nonacidic foods may require a treatment of at least 250°F (121°C) of wet heat for at least 15 min, or its lethal equivalent, to be delivered throughout the entire food mass. Such a treatment would be damaging to most foods. The term commercial sterility refers to a less severe condition that still assures destruction of all pathogenic organisms, as well as organisms that, if present, could grow in the product and produce spoilage under normal conditions of handling and storage. Most of the canned food supply that is stable at room temperature is commercially sterile. This is commonly achieved in canning retorts with steam under pressure at temperatures and for times that vary, depending upon container size and chemical and physical properties of the food, which can affect heat-transfer rates and the thermal resistance of organisms. *See* STERILIZATION.

Many foods are subjected to still less severe heating by methods that produce pasteurization to assure destruction of pathogens and extend product shelf life by inactivating food enzymes and reducing the number of spoilage organisms. Pasteurization of milk is achieved with a temperature of 145°F (63°C) for 30 min, or its thermal lethal equivalent. Since significant numbers of nonpathogenic bacteria survive, storage life is extended by refrigerating the pasteurized milk. Beer, wine, fruit juices, and other foods are commonly pasteurized, but at different temperatures. Heat blanching is a kind of pasteurization applied to vegetables to inactivate enzymes when such products are to be frozen, since frozen storage of

itself does not stop enzyme activity. *See* PASTEURIZATION.

The lethality of heat always depends upon temperature and time. Higher temperatures for shorter times can be as effective as lower temperatures for longer times, and appropriate combinations can be selected for thermal lethal equivalency. Time–temperature combinations with equivalent microbial lethality, however, are not equal with respect to the damaging effects these can have on color, flavor, texture, and nutritive values of foods. In this regard, higher temperatures for shorter times will yield products superior to those produced with lower temperatures for longer times.

Advances in thermal processing incorporate the high-temperature short-time (HTST) principle whether pasteurization or commercial sterilization is the goal. The application of high-temperature short-time processing is more easily accomplished with liquid foods or liquids containing small particulates than with solid foods, since the former can easily be heated and cooled rapidly by passing them in thin layers through specially designed heat exchangers. This is done in the process of aseptic canning, where products prepared to commercial sterility standards are heated to temperatures as high as 302°F (150°C) for 1 or 2 s and as quickly cooled, and then sealed in previously sterilized containers within an aseptic environment.

Cooling and freezing. The slowing of biological and chemical activity with decreasing temperature is the principle behind cooling (refrigeration) and freezing preservation. In addition, when water is converted to ice, free water required for its solvent properties by all living systems is removed. Even severe freezing, however, will not destroy large numbers of microorganisms or completely inactivate food enzymes; these can resume rapid activity, unless inhibited by other means, when food is removed from cold or frozen storage.

Most microorganisms grow best in the range of about 60–100°F (16–38°C). Psychrotrophic bacteria thrive at low temperatures and can grow slowly at temperatures down to 32°F (0°C) and below if free water exist. Most pathogens cannot grow below 40°F (4°C). Home refrigerators commonly operate in the range of about 40–45°F (4–7°C). Some fruits and vegetables store best at temperatures of about 50°F (10°C), and commercial refrigerated storage may be optimized for specific products. Refrigerated storage life of many foods can be extended by the use of packaging that minimizes moisture loss and controls gas atmospheres within packages.

Highest-quality frozen foods depend upon very fast rates of freezing. Slow freezing leads to the growth of large irregular ice crystals capable of disruption of delicate food textures. Slow freezing also increases the time during which food constituents can react adversely with solutes that become concentrated by liquid water changing to ice as freezing progresses. Thus, rapid freezing has been the goal of advanced freezing processes.

Commercial freezing methods utilize refrigerated still air; high-velocity air, which is faster and more efficient; and high-velocity air made to suspend particulate foods, such as peas, as in a fluidized-bed fast freezer. Indirect-contact freezing utilizes hollow flat plates chilled with an internally circulated refrigerant to freeze solid foods, or with refrigerated tubular heat exchangers that rapidly slush-freeze liquids. Immersion freezing involves direct contact of the food or its container with such refrigerants as cold brine, a glycol approved for food, or a fast-freezing cryogenic liquid, such as liquefied carbon dioxide or liquid nitrogen. Liquid nitrogen has a temperature of −320°F (−196°C). *See* COLD STORAGE.

Concentration and dehydration. When sufficient water is removed from foods, microorganisms will not grow, and many enzymatic and nonenzymatic reactions will cease or be markedly slowed. Free water that can enter into biological and chemical reactions is more important than total water, since some water may be bound and unavailable to support deteriorative processes. Free water exerts vapor pressure and possesses water activity (that is, provides water for bacterial growth), which must be decreased below critical levels if foods are to be preserved. Sugar syrups are concentrated foods whose water activity is below that required to support microbial spoilage. Sugar added to fruit juice will bind water, lower the juice's water activity, and, in sufficient concentration, yield a jelly that does not undergo microbial spoilage at room temperature. Concentration preservation, therefore, can be achieved by physically removing water, as by boiling or with lower-temperature vacuum evaporation, or by binding water through the addition of sugar, salt, or other solutes.

Foods preserved by dehydration contain considerably lower water activity and less total water than concentrated foods. Sun-dried cereal grains contain about 14% total water. Most dehydrated foods such as dried milk, instant coffee, and dehydrated potato flakes or granules contain less than 10% total water, and some, such as fruit juice crystals, contain less than 2%.

Most dehydration methods utilize heat to vaporize and remove water. This is most efficiently achieved when a food can be highly subdivided to produce a large surface area for rapid heat transfer into the food and rapid moisture transfer out. Liquid foods and purees commonly are atomized into a heated chamber (as in spray drying), spread thinly over the surface of a revolving heated drum from which they are continuously scraped (as in drum drying), and sometimes thickened or foamed and cast on belts that move through a tunnel oven. Solid foods may be diced to uniform piece size for more even drying and dried with heated moving air in cabinets, on belts, or within rotating cylinders to provide tumbling action. Fluidized-bed dryers use high-velocity air to suspend particulates for still faster drying.

The heat and oxygen sensitivity of many foods necessitates vacuum dehydration for high quality. Under vacuum, water can be removed at reduced temperature, and oxidative changes are minimized.

Solid foods tend to shrink and undergo shape distortion when they are dried. This can be overcome by freeze-drying whereby foods are frozen quickly and placed in a chamber under high vacuum. Vacuum and temperature conditions are regulated to promote sublimation of water vapor from the ice phase without the ice melting. The food's structure remains rigid as it goes directly from the frozen state to dryness. Because of its gentleness, freeze-drying is also used to dehydrate liquid foods such as coffee. A disadvantage of freeze-drying, however, is that it is more costly than other drying methods. *See* DRYING; SUBLIMATION.

Irradiation. X-rays, ultraviolet light, and ionizing radiations (including gamma and beta rays) belong to the electromagnetic spectrum of radiations and differ in frequency, wavelength, penetrating power, and the effects upon biological and nonbiological systems. Ionizing radiations may be obtained from radioactive isotopes, such as cobalt-60, or from electron accelerator machines. These radiations penetrate foods and exert their major effects by producing free radicals from water and other substrates. Depending upon dose intensity, these radiations can inhibit the sprouting of tubers, destroy insects, inactive some enzymes, and kill microorganisms to the point of pasteurization or sterilization. Preservation of food using sterilizing irradiation doses is done for various purposes, such as for use by astronauts during space missions and for immune-compromised persons who are highly susceptible to microbial infection. *See* FREE RADICAL.

Food irradiation remains highly controversial, partly because of fears that the safety of products and processes cannot be adequately regulated. In the United States, treatment of spices, raw or frozen beef, pork, or poultry meat to destroy microbial contamination is among the very few applications that are permitted. Several other countries permit wider use of food irradiation, including low-dose irradiation pasteurization to extend the storage life of highly perishable fruits and vegetables, poultry, and seafoods.

pH control. Hydrogen-ion concentration affects the rate and course of a great variety of chemical reactions. Microbial growth and metabolism and the activities of food enzymes exhibit pH optima and can be controlled to various degrees beyond these optima. The natural acids of certain fruits and vegetables, acid added as a chemical, and acid produced by fermentation can inhibit or partially inhibit several pathogenic and spoilage organisms. *Clostridium botulinum*, the most heat-resistant pathogen found in foods, will not grow and produce toxin at a pH of 4.6 or below. Therefore, foods with a pH in this range do not constitute a health hazard from this organism, and they do not require heat processing as severe as that required for more alkaline foods. Further, acid enhances the lethality of heat, often permitting milder heating conditions.

The pH of acidic foods, however, is rarely sufficiently low to assure long-term preservation from acid alone. Many acidic and fermented foods further depend upon prior pasteurization of their ingredients, the addition of salt and other chemicals, and refrigeration. *See* PH.

Chemical preservative. The U.S. Food and Drug Administration and comparable agencies in various countries vigorously regulate the chemicals that may be added to foods as well as the conditions of their use. Chemical preservatives and similar substances include antimicrobials, such as sodium benzoate, sorbic acid, and sodium nitrite; enzyme inhibitors, such as sulfur dioxide, to control browning of fruits and vegetables; and antioxidants, including butylated hydroxyanisole (BHA) and butylated hydroxytoluene (BHT), to help control fat rancidity. New chemicals must undergo rigorous testing to be approved, and approved chemicals may subsequently be prohibited when new information relative to safety warrants such action. There is much pressure to remove chemicals from the food supply, especially where their effects can be achieved by other means.

Packaging. Preservation methods cannot be effective without adequate packaging. Packaging protects foods from contamination, moisture gain or loss, flavor loss and odor pickup, the adverse effects of light, physical damage, and intentional tampering. Packaging enables food to be stored under vacuum, inert gases, or carefully selected gases that can control respiration of fruits and vegetables, biological changes in meat, and growth of microorganisms. Packages and packing materials must be carefully chosen to withstand the stresses of heating, freezing, and other operations since many products are processed within their final containers. Ultimately, a food product's quality and storage life are determined largely by its package. *See* FOOD ENGINEERING; FOOD MANUFACTURING; FOOD MICROBIOLOGY.

Combined preservation methods. The application of combined food preservation methods is called hurdle technology. This approach to food preservation is based on the use of a combination of several preservative factors (called hurdles) such as high or low temperature, reduced water activity, acidity, and antimicrobial food additives. These hurdles inhibit or prevent growth of food-borne microorganisms to enhance food safety and reduce food spoilage. The application of hurdle technology permits mild but very effective preservation of safe, stable, nutritious foods with desirable sensory quality because only a relatively low level of each preservation factor is usually required to improve the storage stability of foods.

Aubrey F. Mendonca; Norman N. Potter

Bibliography. G. V. Barbosa-Canovas et al., *Nonthermal Preservation of Foods*, Marcel Dekker, New York, 1998; L. Leistner and G. Gould, *Hurdle Technologies: Combination Treatments for Food Stability, Safety and Quality*, Kluwer Academic/Plenum, New York, 2002; D. G. Olson, Irradiation of food, *Food Technol.*, 52:56–62, 1998; N. N. Potter and J. H. Hotchkiss, *Food Science*, 5th ed., Aspen Publishers, Gaithersburg, 1998; M. Shafiur Rahman (ed.), *Handbook of Food Preservation*, Marcel Dekker, New York, 1999; P. Zeuthen and L. Bogh-Sorensen, *Food Preservation Techniques*, CRC Press, Boca Raton, 2003.

Food science

A science dealing with the physical, chemical, and biological properties of foods, in addition to the factors affecting them and their ultimate effects upon the sensory, nutritional, and storage properties and the safety of foods. Food properties are influenced by growing, harvesting, and slaughtering practices, preservation and preparation methods, processing and storage conditions, and packaging. Food science and its applications must be further concerned with economics and marketing; food preferences of various populations; quality assurance and control; regulatory aspects dealing with safety, wholesomeness, and honest representation; and the production of affordable, quality food on a worldwide basis. Therefore, food science interfaces with and draws upon many disciplines, including chemistry, physics, mathematics, the plant and animal sciences, biochemistry, enzymology, microbiology, genetics, engineering, statistics, computer science, nutrition, toxicology, psychology, and law.

Food science is concerned with food composition, which includes not only the kinds and amounts of food constituents (such as proteins, fats, carbohydrates, vitamins, minerals, emulsifiers, acids, oxidants, antioxidants, enzymes, pigments, flavors, and water) but also their physical and physicochemical arrangements. These constituents determine the properties of color, texture, flavor, nutritive value, and keeping quality of food. Each property may be grossly changed by processing. That is, processing may alter a food's cellular integrity and the states of colloidal dispersion, emulsification, or solution of its constituents (as well as its constituent interactions), and the susceptibility to deterioration by air, light, temperature, relative humidity, microorganisms, and natural food enzymes.

Food science deals with many food commodities and thousands of derived products. These commodities are processed for many reasons, including preservation, creation of new product forms, improvement of sensory and nutritional qualities, convenience, and removal of natural toxicants. Common processes are heating, cooling, freezing, concentration, dehydration, fermentation, sometimes irradiation, and packaging. Each process can be further classified by method; dehydration, for example, may employ drum drying, spray drying, cabinet drying, tunnel drying, fluidized-bed drying, vacuum drying, foam drying, freeze drying, and osmotic drying. Each process utilizes unique equipment that requires optimization of variables to maximize product quality and minimize costs, and is capable of yielding product with distinct characteristics. Further, commodities are increasingly utilized for their constituents which are separated, extracted, chemically and physically modified, and then recombined into an endless variety of formulated and engineered foods that are indistinguishable from their source materials. This often requires the use of highly specific food additives to improve processing properties and acceptance factors, including nutritional quality. Nutrient levels and nutrient availability from foods may be decreased or increased by handling and processing practices.

Food science is also concerned with all aspects of food safety, including natural food toxicants, industrial contaminants, misuse of food additives, and microbiological contamination, as well as methods for the detection, exclusion, inactivation, removal, and regulation of harmful substances. Consumers also must be protected against economic deception and misrepresentation.

Undergraduate and graduate degrees in food science are offered by universities around the world. Food scientists, food technologists, and professionals from related fields are employed throughout the food industry and by suppliers of equipment, ingredients, packaging materials, and other services, as well as by government and academia. Food scientists belong to many professional organizations; predominant among them is the Institute of Food Technologists. This organization, whose membership is international, is an excellent resource for information on all aspects of food science. *See* FOOD ENGINEERING; FOOD MANUFACTURING. Norman N. Potter

Bibliography. A. H. Ensminger et al., *Foods and Nutrition Encyclopedia*, 2d ed., 1993; Y. H. Hui, *Encyclopedia of Food Science and Technology*, 4 vols., 1991; N. N. Potter, *Food Science*, 5th ed., 1998.

Food web

A diagram depicting those organisms that eat other organisms in the same ecosystem. In some cases, the organisms may already be dead. Thus, a food web is a network of energy flows in and out of the ecosystem of interest. Such flows can be very large, and some ecosystems depend almost entirely on energy that is imported. A food chain is one particular route through a food web.

A food web helps depict how an ecosystem is structured and functions. Most published food webs omit predation on minor species, the quantities of food consumed, the temporal variation of the flows, and many other details.

Example. Pitcher plants (belonging to the family Nepenthaceae) are tropical species with a structure that holds water into which insects fall and drown. The plant gains nutrients from the decaying insects. Many other species thrive in these tiny ponds, grabbing the energy from the decaying insects before the plant can do so (see **illus.**).

At the base of the food web in the pitcher plant are drowning insects, dead insects, and older organic debris. Some of the species that feed on the decaying material have predators, others do not. In addition, some species may receive energy from more than one level. One predator, a spider (genus *Misumenops*), feeds on another predator, a fly (*Lestodiplosis*), and both these species feed on another fly (*Endonepenthia*). Thus, species are not always arranged into neat, clear food chains.

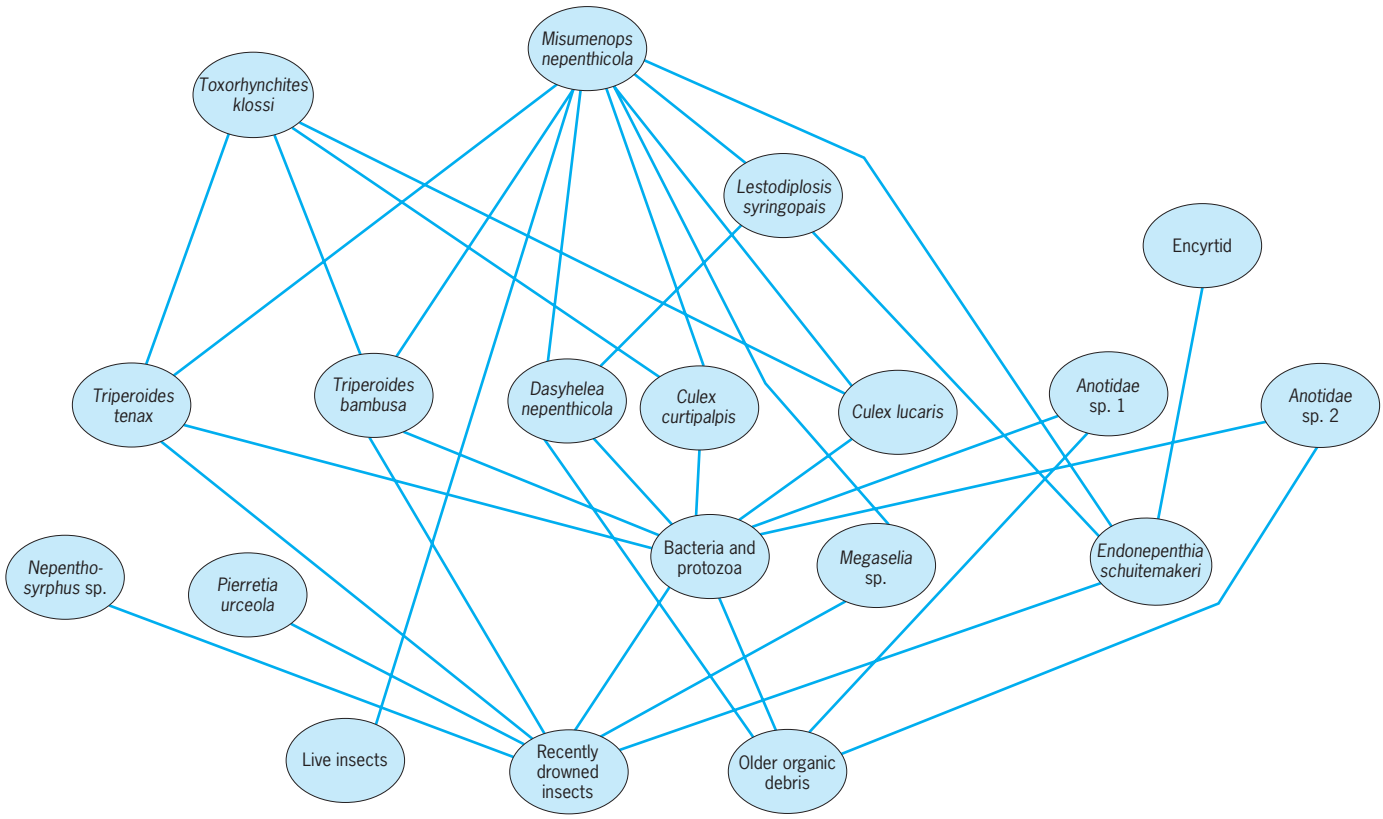

Food web of the insects found in the pitcher plant, *Nepenthes albomarginata*, in West Malaysia. Each line represents a trophic linkage; predators are higher in the web than their prey.

Food chains. Along a simple food chain, A eats B, B eats C, and so on. For example, the energy that plants capture from the sun during photosynthesis may end up in the tissues of a hawk. It gets there via a bird that the hawk has eaten, the insects that were eaten by the bird, and the plants on which the insects fed. Each stage of the food chain is called a trophic level. More generally, the trophic levels are separated into producers (the plants), herbivores or primary consumers (the insects), carnivores or secondary consumers (the bird), and top carnivores or tertiary consumers (the hawk).

Food chains may involve parasites as well as predators. The lice feeding in the feathers of the hawk are yet another trophic level. When decaying vegetation, dead animals, or both are the energy sources, the food chains are described as detrital.

Patterns. Food webs have a long history in ecology, including C. Darwin's famous description of a "tangled bank" at the end of *Origin of Species*. This analogy to tangles implies that even when food webs ignore some details, they demonstrate how complex nature can be. This complexity raises questions: First, are food web patterns simply random, or is there a pattern to their tangles? Second, what are the consequences of these patterns for the dynamics of the ecosystem?

Food chains are usually short. The shortest food chains have two levels. For example, in the illustration there are the drowning insects in the water, and the spider *Misumenops*. The longest chains have five levels [old organic debris, the bacteria and protozoa that feed on it, two species of *Culex* (mosquitoes), *Dasyhelea* (biting midges), and *Misumenops*]. One way to describe and simplify various food chains is to count the most common number of levels from the top to the bottom of the web. Thus, although *Misumenops* sits atop chains of length two and five, the most common chains in the web are those of length three, which is quite typical. Most food chains are three or four trophic levels long (if parasites are excluded), though there are longer ones.

Energetics. There are several possible explanations for why food chains are generally short. The first involves energy. In many ecosystems there are more plants than insects, more insects than insectivorous birds, and more insectivorous birds than hawks. These ecological pyramids reflect an underlying energetic constraint. The first law of thermodynamics states that when energy is converted from one form to another the amount of energy remains constant. An approximate statement of the second law is that the amount of useful energy decreases at each conversion. When insects eat plants, they convert energy locked in plant tissues into insect tissue. Yet, they pay an energetic cost in doing so. Mammals and birds pay an even greater cost. Only about 1% of the energy of the food consumed goes to produce new tissues. The rest is lost as heat, both for warmth and as a by-product of the conversion process. *See* ECOLOGICAL ENERGETICS.

Thus, between each trophic level, much of the energy is lost as heat. As the energy passes up the food chain, there is less and less to go around. There may not be enough energy to support a viable population of a species at trophic level five or higher.

This energy flow hypothesis is widely supported, but it is also criticized because it predicts that food chains should be shorter in energetically poor ecosystems such as a bleak arctic tundra or extreme deserts. These systems often have food chains similar in length to energetically more productive systems.

Recovery from disaster. Another hypothesis about the shortness of food chains has to do with how quickly particular species recover from environmental disasters. For example, in a lake with phytoplankton, zooplankton, and fish, when the phytoplankton decline the zooplankton will also decline, followed by the fish. The phytoplankton may recover but will remain at low levels, kept there by the zooplankton. At least transiently, the zooplankton may reach higher than normal levels because the fish, their predators, are still scarce. The phytoplankton will not completely recover until all the species in the food chain have recovered. Mathematical models can expand such arguments. These models show that the longer a food chain, the longer it will take its constituent species to recover from perturbations. (The phytoplankton could recover quickly in the example if they were the only trophic level.) Species atop very long food chains may not recover before the next disaster. Such arguments predict that food chains will be longer when environmental disasters are rare, short when they are common, and will not necessarily be related to the amount of energy entering the system.

Consequences for species dynamics. The number of trophic levels a food web contains will determine what happens when an ecosystem is subjected to a short, sharp shock—for example, when a large number of individuals of one species are killed by a natural disaster or an incident of human-made pollution—and how quickly the system will recover.

The food web will also influence what happens if the abundance of a species is permanently reduced (perhaps because of harvesting) or increased (perhaps by increasing an essential nutrient for a plant). For example, the population of a fertilized plant species may not expand because herbivores consume the increased plant production, and then the more numerous herbivores may be consumed by their predators, and so on up the food chain.

Some species have redundant roles in an ecosystem so that their loss will not seriously impair the system's dynamics. Therefore, the loss of such species from an ecosystem will not have a substantial effect on ecosystem function. The alternative hypothesis is that more diverse ecosystems could have a greater chance of containing species that survive or that can even thrive during a disturbance that kills off other species. Highly connected and simple food webs differ in their responses to disturbances, so once again the structure of food webs makes a difference. *See* ECOLOGICAL COMMUNITIES; ECOSYSTEM; POPULATION ECOLOGY.

Stuart Pimm

Bibliography. S. R. Carpenter, *Complex Interactions in Lake Communities*, 1988; J. E. Cohen, F. Briand, and C. M. Newman, *Community Food Webs*, 1990; S. L. Pimm, *Food Webs*, 1982; S. L. Pimm, J. H. Lawton, and J. E. Cohen, Food web patterns and their consequences, *Nature*, 350:669–674, 1991.

Foot-and-mouth disease

A highly contagious viral disease of domesticated and wild cloven-hoofed animals with the potential to cause enormous economic losses. It is characterized by the formation of vesicles on the feet, in and around the mouth (**Fig. 1**), and on the mammary gland. At the acute stage there is high fever, depression, lameness, and reduced appetite. Milking animals show a sudden reduction in production. The mortality in adult animals is usually less than 3%, but in young animals it can exceed 90%.

Properties of virus. The causative virus is a member of the *Aphthovirus* genus of the family Picornaviridae. The virion ranges 20 to 30 nanometers in diameter and contains a positive-stranded ribonucleic acid (RNA) genome. The nonenveloped viral capsid is icosahedral in shape and consists of 60 copies of each of the four structural proteins VP1–4 (**Fig. 2**). The virus is acid-labile and inactivated outside the pH range 6.0–8.0.

Seven serotypes are recognized: O, A, C, SAT 1, SAT 2, SAT 3, and Asia I. There is no cross-immunity between serotypes. Each serotype contains strains of varying antigenic and genomic characteristics.

Pathogenesis. The virus can infect by different routes and mechanisms. Ruminant animals are usually infected through the upper respiratory tract following the inhalation of virus associated with droplets or aerosol particles. Pigs are most often infected through the alimentary tract following the ingestion of virus in contaminated foodstuffs (such as meat or milk). In cattle infected by the respiratory route, the virus replicates initially in the mucosae and lymphoid tissues of the pharyngeal region and associated lymph nodes. It then enters the bloodstream and infects glandular organs and lymph nodes

Fig. 1. Mouth of a steer with foot-and-mouth disease lesions on the tongue, along the lower gum, and inside the lower lip.

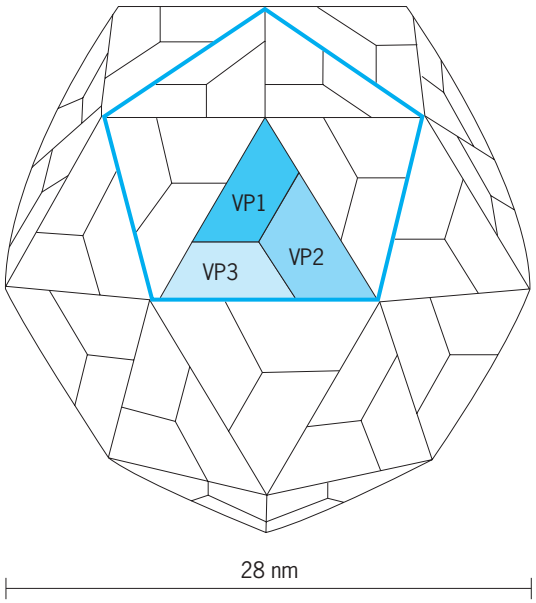

Fig. 2. Schematic diagram of the capsid of a foot-and-mouth disease virion. The four structural proteins, VP1-4 (also called 1A, 1B, 1C, and 1D), make up a protein subunit or protomer. Five protomers join to form a pentamer, and a capsid is an assembly of 12 protomers. VP4 (1A) is located internally beneath the other three proteins.

throughout the body. Other predilection sites, such as the epithelial tissues of the mouth, feet, and mammary gland, are infected during this phase, leading to vesicle formation. The virus has a high affinity for the cardiac tissue of young animals, in which it can cause a severe pathology leading to heart failure and death.

During the acute phase of disease, large amounts of virus occur throughout the tissues and organs of the infected animal and in its excretions and secretions. After 4–5 days, circulating antibody appears, and there is rapid clearance of virus from all sites and fluids. An exception is the pharyngeal region of ruminants, where virus can persist in a proportion of animals for months or years. The duration of infection is species-related; the African buffalo can remain infected for 5 years, cattle for $3^1/_2$ years, and sheep for 9 months.

Epidemiology. The movement of infected animals and the transmission of virus by contact to susceptible animals in a new herd or flock is by far the most common mechanism of spread. Next in frequency is the movement of contaminated animal products, such as meat and milk—most likely to infect pigs through ingestion. Once a few pigs have been infected, the quantity of virus will be greatly amplified, and spread by direct contact will be possible. More rarely, spread results from the mechanical transfer of virus by people, by vehicles, or on fomites (any article that has been in contact with a virus and is capable of transmitting it). Another infrequent mechanism is dispersal of the virus by wind. This generally requires a series of special circumstances: infected pigs at source, a flat terrain, a low-to-moderate wind speed, low atmospheric turbulence, a high humidity, and a high density of cattle downwind. The most dramatic instances of airborne spread of foot-and-mouth disease have been recorded in northwestern Europe, where windborne spread is believed to have taken place over distances of 60 km (37 mi) over land and 300 km (186 mi) across the sea.

Immune response. Immunity to the disease is primarily mediated through neutralizing antibodies directed against the structural proteins of the virus, especially VP1. Antigen-specific T cells play an essential part in the induction and regulation of the circulating antibody response. Specific antibodies are detectable by bioassay 4–5 days after infection, at which time viremia ceases and there is a progressive reduction in virus excretion as the lesions heal. Recovered animals are immune to challenge by the homologous strain of virus for over 4 years.

Laboratory diagnosis. The detection of the viral antigen is sufficient for a positive diagnosis. The specimen of choice is epithelial tissue from an unruptured or freshly ruptured vesicle. Normally, viral antigen is detected and serotyped by enzyme-linked immunosorbent assay (ELISA). In cases where it is not possible to collect epithelial tissue, the demonstration of an increased titre in sequentially collected paired serum samples from nonvaccinated animals is sufficient for a positive diagnosis. Virus neutralization and ELISA are used for serological tests. *See* SEROLOGY.

Control. Countries free from foot-and-mouth disease take stringent measures to prevent the entry of virus. Normally they import animals and animal products only from other countries free of the disease. If they decide to import from a country of lower health status, strict safeguards are imposed to manage the risk. When an outbreak occurs in a country free of the disease, the usual procedure is to apply the stamping-out policy. This involves the slaughter of all affected and in-contact cloven-hoofed animals on the infected premises and the disposal of the carcasses by burning or burial. This step is followed by a thorough cleansing and disinfection of the premises. A movement standstill is enforced in the surrounding area until it is evident that virus is no longer circulating.

Control in countries where the disease is endemic or sporadic is mainly by vaccination. The vaccines contain inactivated viral antigens of one or more serotypes, depending on the prevailing disease situation. Immunity following primary and booster doses of vaccine lasts for around 6 months. *See* ANIMAL VIRUS; EPIDEMIOLOGY.

Alex I. Donaldson

Bibliography. G. R. Thomson, Foot-and-mouth disease, in J. A. W. Coetzer, G. R. Thomson, and R. C. Tustin (eds.), *Infectious Diseases of Livestock with Special Reference to Southern Africa*, vol. II, Oxford University Press, 1994.

Foot disorders

Musculoskeletal, neurological, or dermatological abnormalities of the foot. They may be of developmental or acquired origin.

Disorders of bones and joints. Musculoskeletal foot problems may be congenital or secondary to

neurological disorders, joint laxity, infection, arthritis, or tendinopathy (tendon injury). Poorly fitting footwear may exacerbate symptoms from a deformed foot but is not in itself a cause of foot deformity. *See* JOINT (ANATOMY); TENDON.

Clubfoot. Clubfoot is a congenital deformity that is also known as talipes. Talipes calcaneovalgus, in which the sole of the foot is turned upward and outward, is usually a mobile deformity and has a good prognosis even if untreated. Talipes equinovarus, in which the sole of the foot is turned downward and inward, requires early treatment, especially when the deformity is rigid. The incidence in the United States is 1 in 1000 live births but is lower in some countries such as China and much higher in other areas, especially Polynesia. Almost half of cases are bilateral.

If untreated the deformity persists so that the child has to walk on the top of the foot. The results of early manipulation and use of a serial plaster cast are good. If this fails to give good correction, surgical release of the contracted tendons and ligaments is necessary to prevent the development of secondary bony deformity. *See* CONGENITAL ANOMALIES.

Flatfoot. Flatfoot, or pronated foot, is a condition in which the longitudinal arch of the foot is lost. Although it gives rise to little disability in childhood, it is a common cause of maternal concern, and it used to be a reason for rejection from military service. In this condition, which is not seen until a child starts to walk, the heel of the foot is seen to go into a valgus position (turned outward) and this corrects well when the child stands on tiptoe. The deformity is usually hypermobile and associated with excessive generalized joint laxity. However, fixed flat foot deformity may have more serious underlying causes such as congenital coalition of the hindfoot bones.

The prognosis for flatfoot is good. Most cases are treated by an insole (orthosis), and surgery is rarely indicated.

A suddenly occurring flatfoot deformity in the adult may occur in middle-aged females and is usually due to tendinopathy or rupture of the tibialis posterior tendon on the inside of the foot and ankle. If the tendon is ruptured, early surgical treatment is essential.

Bunion. A bunion (hallux valgus) is a deformity in which the big toe is deviated away from the midline, and the first metatarsal bone moves away from adjacent second metatarsal with splaying out of the forefoot. The ensuing prominence of the first metatarsal head at the base of the big toe gives rise to the apparent bony deformity or bunion. Wearing tight, narrow, high-heeled shoes makes this a painful condition in many middle-aged females. Pressure of the shoe against this prominent bone may give rise to a protective sac of fluid (adventitious bursa). There may be an associated pronation of the foot, and treatment with an orthosis may be advised in mild cases.

Surgical treatment is indicated in cases with persistent pain, especially if the condition does not respond to modifications in footwear. Simple excision of the bunion is not usually successful, and a variety of surgical procedures to realign the first metatarsal bone and correct the position of the joint at the base of the big toe are used.

Hammertoe. This clawing of the toes may present in small children and is usually thought to be associated with a mild neuromuscular imbalance. Severe cases may occur in children with spina bifida, poliomyelitis, and Charcot-Marie-Tooth disease. Fixed deformity may develop, and in severe cases surgical treatment by release or transfer of tendons becomes necessary.

In the adult, hammertoe deformity may develop in the smaller toes without any obvious underlying neurological abnormality. In this deformity, the proximal phalanx of the toe (the one nearest to the rest of the foot) becomes extended upward, the middle phalanx flexed down, and the distal phalanx (the one farthest from the rest of the foot) hyperextends to remain parallel with the floor. Pain is due to pressure on the joint between the proximal and distal phalanges and may be relieved by fitting a large toe box. In severe cases, surgical treatment to fuse this joint (arthrodesis) may be advised.

Heel spur and plantar fasciitis. The plantar fascia is a strong ligament that maintains the longitudinal arch of the foot and acts like a bow string with attachment to the heel (calcaneus) and forefoot. It may become inflamed at the calcaneal attachment, especially in athletes and overweight patients, when excessive forces are applied to this area. In some cases, there is an outgrowth of bone in this area (calcaneal spur), although a calcaneal spur may be an incidental finding on x-ray in someone without pain.

Treatment may involve a variety of measures, including stretching exercises for the Achilles tendon, use of an orthosis, local steroid injection, and extracorporeal shock wave therapy. If these measures fail, surgery to release the attachment of the plantar fascia may be indicated.

Calcaneal apophysitis. This condition, also known as Sever's disease, causes heel pain from the apophysis (or posterior center of ossification of the calcaneus) in children, and may represent a stress fracture of the calcaneus. It may occur as a result of excessive athletic activity. Most cases settle with rest and use of an orthosis. Immobilization in a plaster cast may also be effective.

Neurological disorders. Neurological disorders may affect the foot by causing muscle imbalance and sensory loss. For example, an increasingly common problem seen in diabetic patients is peripheral neuropathy (damage to the peripheral nervous system resulting in the loss of sensation) causing deformity due to neuropathic arthropathy (degeneration of the stress-bearing portion of a joint) and painless ulceration of the sole of the foot. Morton's neuroma and tarsal tunnel syndrome are examples of entrapment neuropathy in the foot. *See* DIABETES; NERVE; NERVOUS SYSTEM (VERTEBRATE).

Morton's neuroma. A neuroma is a benign tumor of nerve cells. Morton's neuroma, most common in female patients between the ages of 50 and 60, is characterized by the thickening of the third intermetatarsal nerve over a period of years, causing

a small benign fusiform (spindle-shaped) tumor to eventually form, usually in the space between the third and fourth toes. Pain is felt in the space and may be associated with pain and pins and tingling radiating into the third and fourth toes (or second and third toes if the neuroma lies between the second and third metatarsal heads). A fusiform tumor may be diagnosed by an ultrasound scan.

Conservative treatment is advised and involves the use of an orthosis in the shoe and a local steroid injection in the space between the metatarsal heads. Entrapment of the nerve may be relieved surgically by dividing the intermetatarsal ligaments to perform decompression. More commonly, the benign tumor is excised, although there is a risk of recurrent symptoms due to the risk of a traumatic neuroma developing in the stump of the nerve.

Tarsal tunnel syndrome. The posterior tibial nerve runs into the sole of the foot from behind the inner aspect of the ankle joint and into a channel known as the tarsal tunnel. Entrapment of the nerve within this tunnel gives rise to tarsal tunnel syndrome, characterized by pain, and tingling and numbness radiating into the sole of the foot along the cutaneous distribution of the nerve. In some cases, entrapment is due to a bony swelling near the inside of the ankle or to a cystic swelling such as a ganglion.

The initial treatment consists of placement of an orthosis inside the shoe or local steroid injection into the tarsal tunnel. Measurement of conduction velocity in the posterior tibial nerve may be used to confirm the diagnosis; if the changes are severe, surgical treatment may be necessary to decompress the nerve.

Dermatological disorders. The skin of the foot undergoes stress from pressure and friction. In response, it protects itself with defensive modifications that themselves cause pain.

Corns and calluses. A corn, also known as a heloma, is a small, sharply circumscribed, conically shaped and deep-seated area of skin. Calluses are similar lesions but are flatter and more evenly distributed. Both are likely to occur in the sole of the foot in response to excessive pressure and consist largely of the protein keratin. This pressure is usually the result of a bony deformity, for example, a metatarsal head prolapsed into the sole of the foot giving rise to a painful callosity beneath the fibular condyle of the metatarsal head known as intractable plantar keratosis. An adventitious bursa may also form in response to excessive pressure.

Shaving the corn or callus does not address the underlying problem. Orthotic treatment should be advised to reduce localized pressure. If this fails, surgery to correct the underlying bony deformity may become necessary.

Warts. Warts are benign lesions of the skin induced by the human papilloma virus and are secondary in prevalence to corns and callosities. They can occur elsewhere in the body but are most common in the sole of the foot (plantar warts). Transmission from person to person may occur but depends on individual immunity. The virus lies in the epidermis and stimulates this superficial layer of the skin to produce excessive keratin. This causes pressure, further stimulating keratin production and irritating pain receptors in the underlying dermis.

Treatment involves applying agents such as salicylic acid to the skin with an occlusive dressing. Surgical removal may be effective but runs the risk of painful scarring.

Ingrown toenail (onychocryptosis). This penetration of the edge of the nail into the surrounding soft tissues may be due to hyperplasia (increase in size) of the nail groove or deformity of the nail plate; the latter may be associated with an osseous (bony) malformation of the big toe or a fungal infection. This condition is most often seen in teenage boys and may respond to improved attention to hygiene and correct cutting of the toe nail. Initially, there is pain, redness, and swelling on either or both borders of the nail. If chronic infection supervenes, removal of all or part of the nail may be necessary. The nail bed may be treated by phenol or by surgical excision to prevent further growth of the ingrowing border and subsequent recurrence. Radical procedures to perform a terminal amputation of the big toe should not be necessary.

T. W. D. Smith

Forage crops

Grasses and legumes that make up grasslands and are used as forages for livestock. The grasslands represent an ancient renewable natural resource. They form 25% of the world's vegetation and occupy the largest area of any single plant type. They benefit humanity indirectly by providing food for both wild animals and domesticated livestock, some of which are ruminants that, because their digestive systems contain microorganisms, are able to digest fibrous forage material. Thus, the prime value of grassland areas lies in the meat, milk, or work produced by the livestock that graze on them. *See* GRASSLAND ECOSYSTEM.

Grasslands have other attributes as well. In order to withstand the tug and pull of the grazing animal, a forage plant must have an extensive root system, and this makes a contribution to soil fertility. When the plant is grazed or cut, the photosynthetic area that remains is not large enough to provide sugars to maintain root respiration, and so part of the root system dies and adds to the organic material in the soil and to soil structure. Such is the basis of many highly productive crop rotations that maintain soil fertility without expensive fertilizer applications. Most forage plants are long-lived perennials that can be defoliated repeatedly during a growing season. *See* AGRICULTURAL SOIL AND CROP PRACTICES.

Forages also protect the soil from erosion by both wind and water. In fact, where row crop farming has led to soil deterioration, as it did when dust bowl conditions prevailed in North America, forages are used to restore and stabilize the land. Many of grassland areas make an important contribution to water catchment. They stabilize land not only in high-rainfall

zones but also on the banks of rivers and canals that bring water to areas of high population. Moreover, the green and pleasant areas covered by forage have an esthetic appeal to people. Grasses are used to beautify parks, airfields, playgrounds, and highways. *See* EROSION; LAWN AND TURF GRASSES.

Plant groups. The grasses and the legumes, both of which are used as forages, are divided into cool-season or temperate species and warm-season or tropical plants.

Grasses. The grasses are very numerous, herbaceous, monocotyledonous plants. They are the result of over 30 million years of coevolution between plants and the animals that graze upon them. During this time, certain mouthparts of animals were modified to permit grazing on different plant types. For example, sheep are able to graze very close to the ground, while cattle can use only plant growth over 2–3 in. (5-7.5 cm) tall. Grasses have developed low growing points, well below the bite level of grazing animals, so that the plant can regrow after defoliation. The plant's inflorescence is extruded rapidly and is vulnerable to the animal for only a minimum time. Grass seeds have awns, spikes, and hooks. Many are disseminated by becoming attached to an animal's coat. At maturity, some grass heads are unattractive to grazing animals, and so the seed is not eaten. When the seed is eaten, it may be capable of passing through the animal's digestive system without damage. The cool-season grasses are lower-yielding but provide forage material of higher quality than do the warm-season type. *See* GRASS CROPS.

Some commonly grown grasses are the following:

Cool-season grasses
- Bentgrasses (*Agrostis* spp.)
- Bluegrasses (*Poa* spp.)
- Smooth bromegrass (*Bromis inermis*)
- Fescues (*Festuca* spp.)
- Meadow foxtail (*Alopecurus pratensis*)
- Orchardgrass (*Dactylis glomerata*)
- Ryegrasses (*Lolium* spp.)
- Wildryes (*Elymus* spp.)
- Timothy (*Phleum pratense*)
- Wheatgrasses (*Agropyron* spp.)

Warm-season grasses
- Bahiagrass (*Paspalum notatum*)
- Bermudagrass (*Cynodon dactylon*)
- Dallisgrass (*Paspalum dilatatum*)
- Johnsongrass (*Sorghum halepense*)
- Rhodegrass (*Chloris gayana*)

Legumes. The legumes have been used by farmers for over 6000 years. The reason they are so valuable to humanity is that their root systems are capable of forming a symbiotic relationship with bacteria of the genus *Rhizobium*. The plant provides the bacteria with sugars for growth and metabolism. The bacteria are able to fix atmospheric nitrogen and make it available to the plant in the form of nitrates. The plant builds these salts first into amino acids and then into proteins. Thus, a nitrogen atom from the air has passed along a short food chain from bacteria to plant and on to the animal. *See* LEGUME FORAGES; NITROGEN FIXATION.

The most commonly grown forage legume is alfalfa, which is used in both the temperate and subtropical areas. The warm-season legumes are important since, grown in a mixture with warm-season grasses, they improve forage quality. *See* ALFALFA.

Some of the legumes are the following:

Temperate-season legumes
- Alfalfa (*Medicago* spp.)
- Birdsfoot trefoil (*Lotus corniculatus*)
- True clovers (*Trifolium* spp.)
- Sainfoin (*Onobrychis viciaefolia*)
- Sweetclover (*Melilotus* spp.)

Warm-season legumes
- Crimson clover (*Trifolium nicaenatum*)
- Kudzu (*Pueraria lobata*)
- Subterranean clover (*Trifolium subteraneum*)

Forage quality. There are three ways in which forage quality may be measured. Farmers use a number of "visual" indices like color and odor which indicate that the hay or silage was well made and is still in good condition. For a pasture that is to be grazed, the quality criteria are the species present and their stage of maturity, for as plants age quality declines. The second indices are obtained from chemical analysis of forage material. These attempt to determine the undigestible part of the forage, usually cellulose fiber coated with lignin. This fraction is called the acid detergent fiber (ADF). A chemical analysis can also be used to determine elemental nitrogen. This is usually expressed as crude protein (CP), which contains several nonprotein compounds as well as protein (see **table**). The third types of forage quality indices are the most reliable and the most expensive to determine. These are obtained from the performance of the animals themselves, but unfortunately this is very individualized. The size, age, species, and the individual likes and dislikes of the animal will influence intake and production. Thus, large animal populations (10–20 animals) have to be fed over several months to determine animal intake and live weight gain. Intake contributes 70% to live weight gain, while digestibility accounts for 30%. The important component of digestibility is the rate of passage through the digestive system. Good-quality forages are rapidly digested, and the animal will then start to eat again, thereby increasing intake.

Establishment. All forage seeds are very small and so, in contrast to large-seeded row crops like corn

Crude protein (CP) and acid detergent fiber (ADF) values for common forages

Forage type	CP,%	ADF,%
Alfalfa hay	17.4	38.1
Timothy hay	8.1	37.1
Barley silage	9.7	33.8
Oats silage	8.4	38.0

or wheat, need a carefully prepared seedbed. For forages, the seedbed should be firm and have a fine soil structure. Firm soil ensures that neither root nor shoot grow through air pockets since, having few food reserves, the small seedling might not survive. A fine structure ensures good contact between soil and seed and facilitates water uptake as well as water movement in the soil. Given these conditions, the seed should be placed at a very shallow depth in the soil. Small seedlings are unable to reach the surface if planted deeply. The young forage plant has very few nutrient requirements during the first year, and fertilizers are banded, that is, they are placed in the soil directly below the seed. Forage seed establishment requires precision and specialized forage seeding equipment.

Forages crops are usually grown in species mixtures that contain a legume and one or two grasses. The legume provides high protein levels, and the grass is persistent and high-yielding. A mixture will also make production more uniform over the growing season. Forage plants do not yield consistently throughout the year and, by mixing those that give high early yields with those that do well later, pasture production is stabilized. Forage may be conserved for winter use as hay (that is, dried) or as silage. In making silage the sugars from the forage are used to produce lactic acid, which pickles the forage.

Fertilizing. When a forage crop is harvested for hay or silage, the whole plant, with the exception of the crown and the root, is removed. This involves a substantial loss of soil nutrients and may be contrasted with grazing when much of the nutrient material in the forage is returned directly to the pasture in the animal's excrement. A grass and legume hay crop will remove about 80 kg (176 lb) of nitrogen, 8 kg (17 lb) of phosphorus, and 33 kg (73 lb) of potassium, if the annual yield level is about 4 metric tons per hectare (8 tons/acre). Fertilizer applications for such a crop are most important if soil fertility is not to be drastically reduced. *See* FERTILIZER.

Applications of nitrogen are essential for the grasses grown alone. Grasses require large amounts of nitrogen, which they take up in the form of nitrate ions. For legumes, which fix their own nitrogen, potassium is an essential element. It enhances legume yield, quality, longevity, and winter survival on many soils.

Plant health. Forage plants are affected by diseases, insects, and, most seriously, weeds. Weeds play a dominant role because they compete for light, water, and nutrients, and also because they harbor insects and disease organisms. Since pastures typically are mixtures of perennial broad-leaved and grassy plants, they present many unique weed control problems. The destruction of weeds by cultivation, which takes place every spring and fall in arable fields, is a rare and very expensive event on grazing lands, and so selective herbicides are commonly used. These substances are plant hormones that kill weeds and leave the forage unharmed when applied at appropriate rates. However, the most desirable weed control is obtained by a close soil cover of forage plants. Then the weeds have no opportunity to establish. *See* HERBICIDE.

The insects that attack forage plants may feed on their foliage, suck out their sap, or, in the form of underground grubs, eat their roots. Control may be chemical or biological. Concern for the environment has led to the use of other insects or diseases to control pests and to breeding resistant cultivars. *See* INSECT CONTROL, BIOLOGICAL.

Bacteria, fungi, viruses, and mycoplasma all cause diseases in legumes and grasses. While chemical control methods exist, forages are not highly priced crops, and the most desirable disease protection is by breeding resistant cultivars. The most serious diseases of forages are those that attack at the seedling stage. At that time, plant populations may be so seriously reduced that the whole field has to be replanted. The application of a fungicide to the seed before planting frequently gives protection during establishment.

Pasture management. Skill and experience in the evaluation of field situations are important to the pasture manager. There are two major tools at the disposal of a manager: the system to be used (number of fields and duration of use) and the stocking rate (number of animals per unit of land). Most grazing systems include a period when livestock are excluded from the pasture and the plants are allowed to regrow. There must be compromises in the choices the manager makes. Pasture production varies during the growing season. In spring, plant growth may exceed animal needs. In midsummer and fall it is slow. The animal needs must always be met without overgrazing or damaging the pasture. Also, optimum animal weight gains per area of land are not achieved by the same stocking rates as are the most favorable gains per individual animal. These factors must be balanced against a background of large variations in forage yield due to climatic changes within and between seasons. P. D. Walton

Bibliography. B. N. Chatterjee and P. K. Das, *Forage Crop Production*, 1989; A. Dorrat, Y. Cohen, and A. Goldman, *Irrigated Forage Production*, 1993; G. Fahey et al. (eds.), *Forage Quality, Evaluation, and Utilization*, 1994; M. E. Heath, R. E. Barnes, and D. S. Metcalfe, *Forages: The Science of Grassland Agriculture*, 4th ed., 1985; P. D. Walton, *Principles and Practices of Plant Science*, 1988.

Foraminiferida

An order of Granuloreticulosia in the class Rhizopodea. Foraminiferans are dominantly marine protozoa (single-celled animals), with a secreted or agglutinated shell, or test, enclosing the continually changing ameboid body that characterizes this and other orders of the superclass Sarcodina. Their unique combination of long geologic history, ubiquitous geographic distribution, and exceptional diversity of test composition, form, and structure make the foraminiferans the most useful of all marine fossils for stratigraphic correlation, geologic age

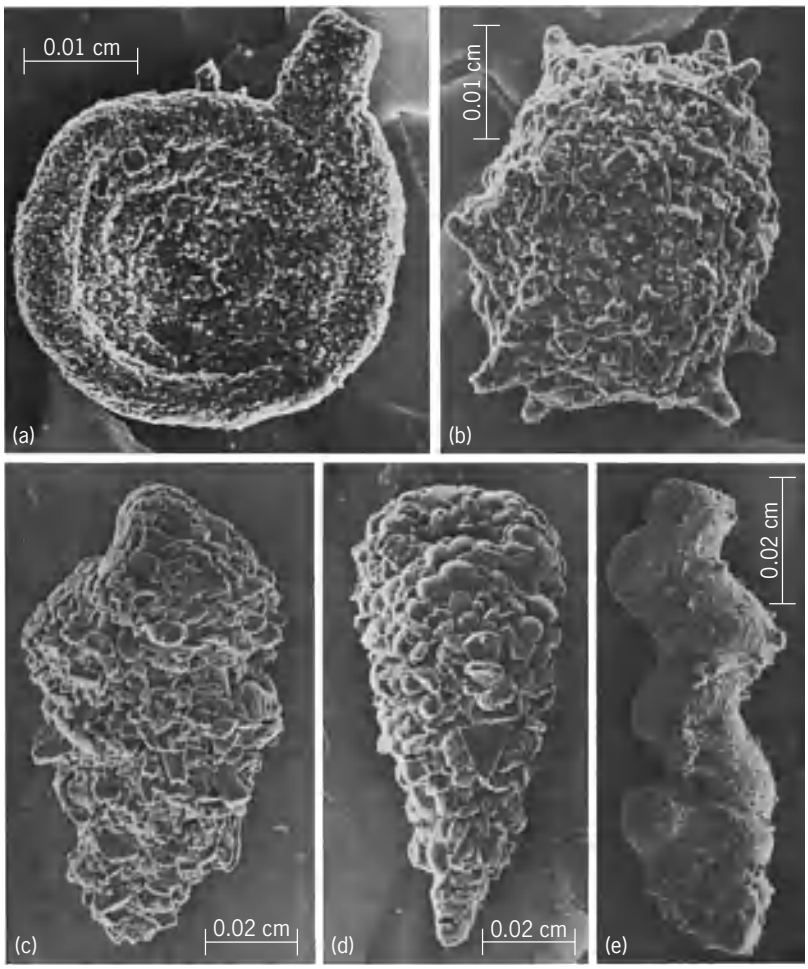

Fig. 1. Scanning electron micrographs of foraminiferans of suborder Textulariina. Superfamily Ammodiscacea: (*a*) *Ammodiscus*, a planispirally enrolled, fine-grained agglutinated test, from Silurian of Missouri. Superfamily Astrorhizacea: (*b*) *Thurammina*, a single globular chamber with apertures on short necks, from Silurian of Missouri. Superfamily Textulariacea: (*c*) *Textularia*, from Holocene of Timor Sea; side view, and (*d*) edge of biserial test coarsely agglutinated of shell fragments. Superfamily Verneuilinacea: (*e*) *Gaudryinella*, with triserial base followed by biserial and zigzag uniserial chambers, wall finely agglutinated, from Upper Cretaceous of Texas. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

dating of sediments, and paleoecologic interpretation. Although small in size, commonly less than 0. 04 in. (1 mm), with a range between 20 micrometers and 4. 8 in. (12 cm), foraminiferans may be abundant, more than 2600 living specimens having been recorded on 1. 5 in.^{2} (10 cm^2) of sea floor. Their tests accumulated in great numbers and are recoverable from small quantities of sediment, rock outcroppings, well cores or cuttings, or ocean dredging and submarine coring.

Living animal. Temporary extrusions (pseudopodia) from the ameboid body form a delicate anastomosing network 2–10 times the test diameter. The pseudopodial net may arise solely from the apertural region of the test in those with tectinous, porcelaneous, and agglutinated walls, or may radiate in all directions through many tiny perforations of the hyaline test wall. The pseudopodia variously serve in capture, ingestion, and digestion of food, in test and temporary cyst construction, for anchorage, and for locomotion. The finely granular cytoplasm consists of darker-colored endoplasm surrounded by a thin, transparent outer layer of ectoplasm. The characteristic granular streaming of the continuously moving cytoplasm differentiates the Foraminiferida from the other Sarcodina orders: the Amoebida, Arcellinida, and Gromiida.

Food. Diatoms are the major food of foraminiferans, although such other organic material as bacteria, phytoflagellates, fragments of larger algae, ciliates, radiolarians, small crustaceans, nematode worms, and other foraminiferans may be used. The food may be carried into the test or may be digested by the external pseudopodia and only the dissolved nutrients ingested. Symbiotic algae have been reported in both benthic and planktonic species. Foraminiferans are in turn used as food by various detritus feeders on the sea floor.

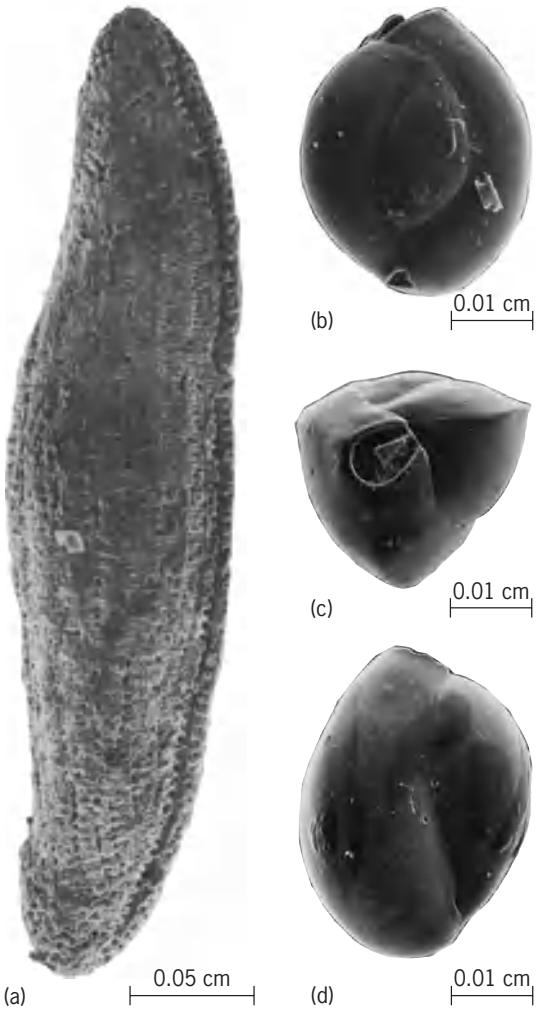

Fig. 2. Scanning electron micrographs of foraminiferans of suborder Miliolina. Superfamily Alveolinacea: (*a*) *Alveolinella*, an isomorph of the microgranular fusulines, but with porcelaneous wall; abraded surface shows the many tiny chamberlets; apertural face at right with numerous apertural pores; from Holocene of Timor Sea. Superfamily Miliolacea: *Triloculina*, from Holocene of Adriatic Sea: (*b, d*) views of opposite sides, each chamber one-half coil in length with coiling in three planes, and (*c*) top view, aperture on final chamber, with bifid tooth projecting from one margin. Wall is porcelaneous and imperforate. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

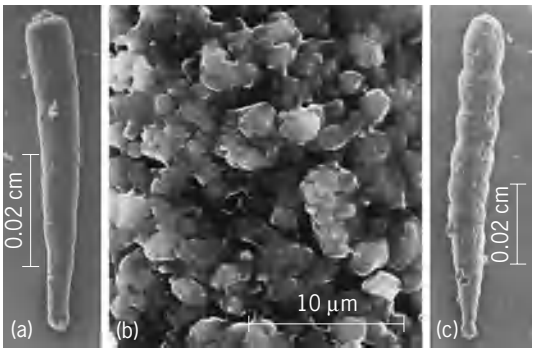

Fig. 3. Scanning electron micrographs of foraminiferans of suborder Fusulinina. Superfamily Eardlandiacea: *Earlandia*, from Pennsylvanian of Texas: (*a*) side view of elongate, undivided tubular test; (*b*) surface of wall, which is composed of tiny angular calcite grains randomly packed. Superfamily Nodosinellacea: (*c*) *Earlandinita*, showing uniserial chambers; from Pennsylvanian of Texas. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

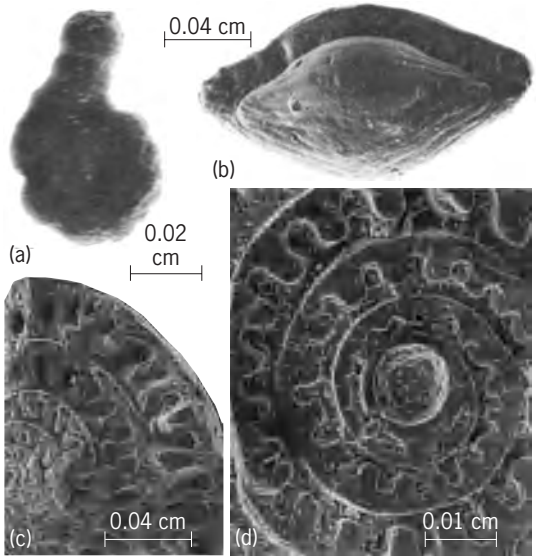

Fig. 4. Scanning electron micrographs of foraminiferans of suborder Fusulinina. Superfamily Endothyracea: (*a*) *Endothyranella*, early chambers enrolled, later ones uncoiled, from Pennsylvanian of Texas. Superfamily Fusulinacea: *Triticites*, from Pennsylvanian of Texas: (*b*) exterior of fusuline test, showing elongate axis of coiling; (*c*) part of transverse weathered half-section showing alveolar wall, proloculus at left; and (*d*) early portion of transverse section enlarged, showing globular proloculus in center, thick wall, and many small chambers per whorl. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

Reproduction and dimorphism. Two methods of reproduction normally alternate, an asexually produced haploid generation (gamont) giving rise by sexual reproduction to a diploid generation (agamont or schizont), which in turn by asexual multiple fission immediately after meiosis produces the new haploid generation. Distinct but morphologically similar haploid and diploid generations occur in the animal kingdom only in the foraminiferans. The entire parent individual is involved in both asexual schizogony and in gamogony; nothing but an empty test remains of the old generation after reproduction.

The separate generations also can be differentiated by test characters. In fact, dimorphism was observed in fossils long before its ontogenetic basis was known. The uninucleate gamont generation normally has a relatively large (megalospheric) first test chamber, or proloculus, whereas the multinucleate agamont generation has a smaller (microspheric) proloculus but a larger, more differentiated adult test. All nuclei may be alike in form and function (homokaryotic), or vegetative and generative nuclei may be differentiated (heterokaryotic), with only the generative nuclei taking part in reproduction. Megalospheric and microspheric tests may be otherwise nearly identical in appearance or sufficiently different that some have been described as separate species. Dimorphism is most pronounced in the more highly evolved calcareous species; the distinctive float chamber formed prior to gametogenesis by *Rosalina* gamonts led to description of a separate genus, *Tretomphalus*.

Modifications of the usual asexual reproduction include shell fragmentation and simple binary fission in some agglutinated types (*Bathysiphon*) or budding and fission (*Halyphysema*). Either asexual schizogony or sexual gamogony may be repeated successively before intervention of the other process to complete the cycle. In sexual reproduction, countless tiny biflagellate gametes typically are released into the water (*Boderia*, *Iridia*, *Peneroplis*, *Elphidium*). Gametes of the Miliolina have a distinctive axostyle. *Glabratella* (superfamily Glabratellacea) has triflagellate gametes, produced by two parent individuals that attach by their umbilical surfaces (plastogamy) to form a joint "brood chamber" in which fertilization occurs. A few large ameboid gametes are produced by two or more parent individuals

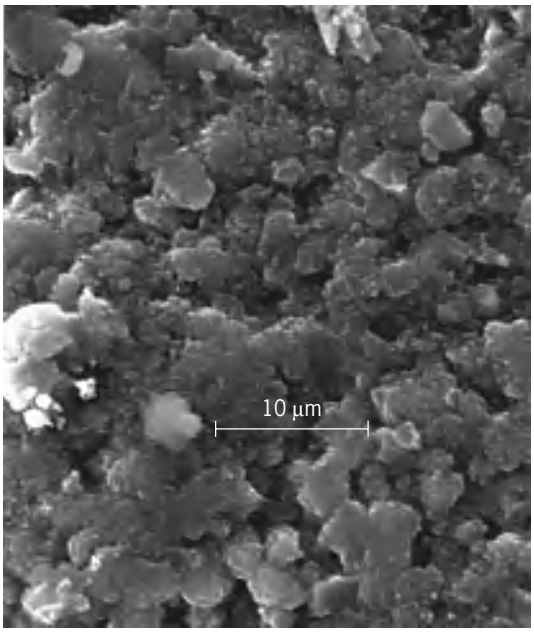

Fig. 5. Enlarged portion of Fig. 4*b* showing surface of microgranular wall constructed of fine, equidimensional calcite granules. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

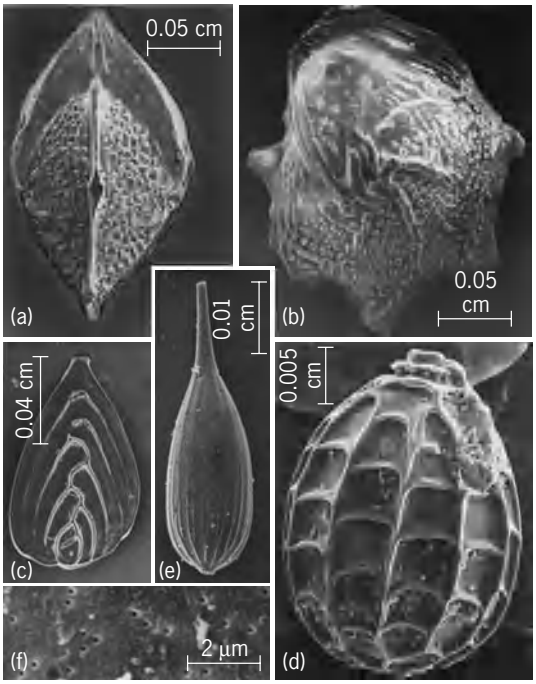

Fig. 6. Scanning electron micrographs of foraminiferans of suborder Lagenina, superfamily Nodosariacea. *Lenticulina*, from Pliocene of Italy: (*a*) edge view showing radiate aperture at top of figure; (*b*) side view of planispirally enrolled test, with surface ribs, nodes, and peripheral keel. (*c*) *Kyphopyxa*, flattened test, with early biserial chambers and later chevron-shaped uniserial ones; aperture at top of figure; from Upper Cretaceous of Texas. (*d*) *Favulina*, single-chambered test with surface costae, from Pleistocene of California. *Laqena*, from Pleistocene of California: (*e*) single-chambered test with elongate apertural neck; (*f*) surface view, showing very small pores in the hyaline wall. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

that in association form a joint reproductive cyst (*Spirillina*, *Patellina*). The large gametes and hence large zygote result in a larger proloculus in the schizont, in contrast to the usual smaller microspheric schizont.

Many other variations in the basic life cycle are known in the approximately two dozen species thus far studied, representing only a few families or genera; other modifications also are probable, for nothing is yet known as to sexual reproduction for many superfamilies, and even some entire suborders (Lagenina, Silicoloculinina, or Robertinina: *Rubratella* may belong to the Robertinina, but no information is yet available as to its wall characters.)

Test. The foraminiferan constructs its own test. Growth of the individual may cease after test construction, or the test may enlarge by continued growth in one or more directions (tubular or branching tests) or by periodic formation of separate but always interconnected chambers. Growth may be either continuous or periodic and may be in a straight line, a planispiral or trochospiral coil, a cycle, or a zigzag. The latter zigzag is expressed by chambered forms in biseriality. Variations and combinations of these growth patterns occur repeatedly in different lineages (isomorphism).

Continual circulation of protoplasm within and without the test is through test openings. A major opening on the final chamber, the aperture, is characteristic in form and position. Supplemental apertures, complex canal and stolon systems, and tiny wall perforations also may occur. The test wall may be thin and delicate or thick and firm, may be simple and homogeneous or have two or more similar or differing layers, and may have a lamellar or labyrinthic structure. The wall may be organic, agglutinated, calcareous, or rarely siliceous.

Tectinous and agglutinated walls. The suborder Allogromiina has an organic (tectinous) test of protein and an acid mucopolysaccharide. Agglutinated walls, characteristic of the suborder Textulariina, consist of foreign particles, quartz and other mineral sand or silt grains, sponge spicules, and tiny shells or shell fragments held in an organic or calcareous cement (**Fig. 1**). Some show selectivity in the foreign material used for test construction and in its arrangement, whereas others utilize a variety of materials in a heterogeneous arrangement. The organic cement is a protein-carbohydrate, at least in part an acid mucopolysaccharide ("tectin"), and morphologically may consist of closely packed strands covering the surface of the cemented grains and filling the intergrain spaces. In others the grains are cemented by a fibrous meshwork, or the organic material may form a foamlike mass of organic bubbles. An inner

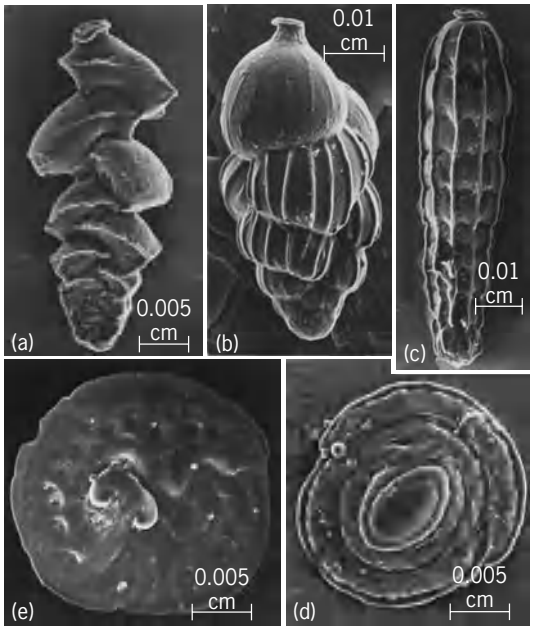

Fig. 7. Scanning electron micrographs of foraminiferans of suborders Rotaliina and Spirillinina. Superfamily Eouvigerinacea: (*a*) *Eouvigerina*, elongate test, early stage biserial, later uniserial, aperture terminal with everted lip; from Upper Cretaceous of Texas. Superfamily Buliminacea: (*b*) *Uvigerina*, elongate triserial test; from Miocene of California. (*c*) *Siphogenerinoides*, reduced early biserial stage, later uniserial; from Upper Cretaceous of Texas. Suborder Spirillinina: *Patellina*, from Pleistocene of California; (*d*) umbilical view of low conical test, showing aperture; (*e*) spiral view showing oval proloculus, enrolled tubular chamber, and later biserial chambers with secondary partitions of later chambers resulting in a crenulate margin. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

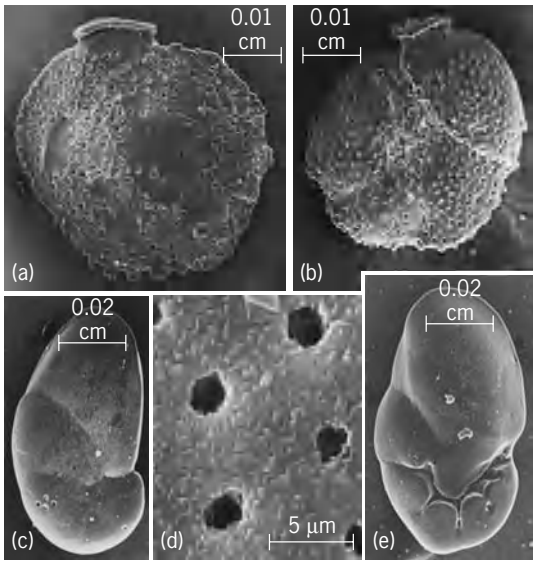

Fig. 8. Scanning electron micrographs of foraminiferans of suborder Rotaliina, superfamily Siphoninacea. *Siphonina*, from Upper Eocene of Mississippi: (*a*) spiral and (*b*) umbilical views of coarsely perforate, trochospiral test with terminal aperture and everted lip. Superfamily Discorbacea: *Cancris*, from Pliocene of Italy: (*c*) spiral and (*e*) umbilical views of finely perforate right- and left-coiled specimens, respectively, the latter showing large imperforate flap that extends over the basal aperture; (*d*) an enlarged part of the spiral side, showing pores and the surface of small subrhombic calcite crystals. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

organic lining separates the test from the cytoplasm, and some also have an outer organic layer.

Although ferruginous, siliceous, and calcareous cements have been reported, the ferruginous material was shown to have consisted of coatings on the foreign particles before their incorporation in the wall, and siliceous cement has not yet been proven to occur. Calcareous cement does occur, particularly in members of the Textulariacea, where it may appear as tiny (0.1-μm) equidimensional grains of low magnesium calcite that may be aligned in rods or bundles of rods that in turn are randomly arranged in the wall. The calcite crystals are enveloped by organic material, suggesting intracellular calcification and later transport with invesicles to the final position of the crystals in the wall. Many of those with calcitic cement (Eggerellidae, Textulariidae, Pseudogaudryinidae, Valvulamminidae, Valvulinidae, Chrysalidinidae, Glaucoamminidae) also have wall pores or canaliculi, which may be closed at each end by an organic plug or lining or may be obscured at the surface by agglutinated material but do not permit passage of the cytoplasm.

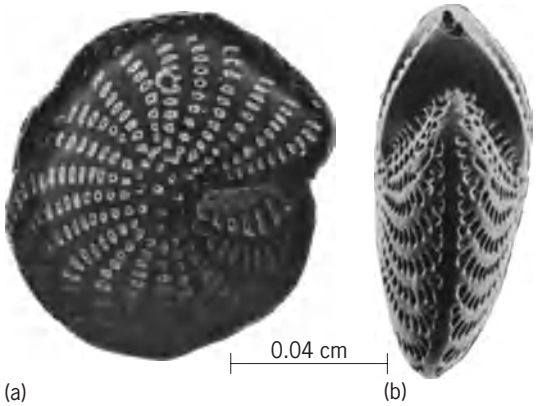

Fig. 10. Scanning electron micrographs of *Elphidium*, suborder Rotaliina, superfamily Rotaliacea, from Holocene of Adriatic Sea. (*a*) Side view showing large sutural pores that open into canal system. (*b*) Edge view with series of apertural pores at lower margin of final chamber face and imperforate peripheral band. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

Porcelaneous walls. The imperforate porcelaneous test (suborder Miliolina) is constructed of randomly arranged needles of high-magnesium calcite, containing considerable organic matter (**Fig. 2**). The three-dimensionally random arrangement of the crystals refracts light in all directions, giving the test a milky opacity. In transmitted light it appears brown. A

Fig. 9. Scanning electron micrographs of *Ammonia*, suborder Rotaliina, superfamily Rotaliacea, from Pliocene of Italy. (*a*) Spiral, (*b*) umbilical, and (*c*) edge views of trochospiral test showing the umbilical and the sutural fissures and canals, the umbilical pillars, and the basal aperture of the last chamber. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

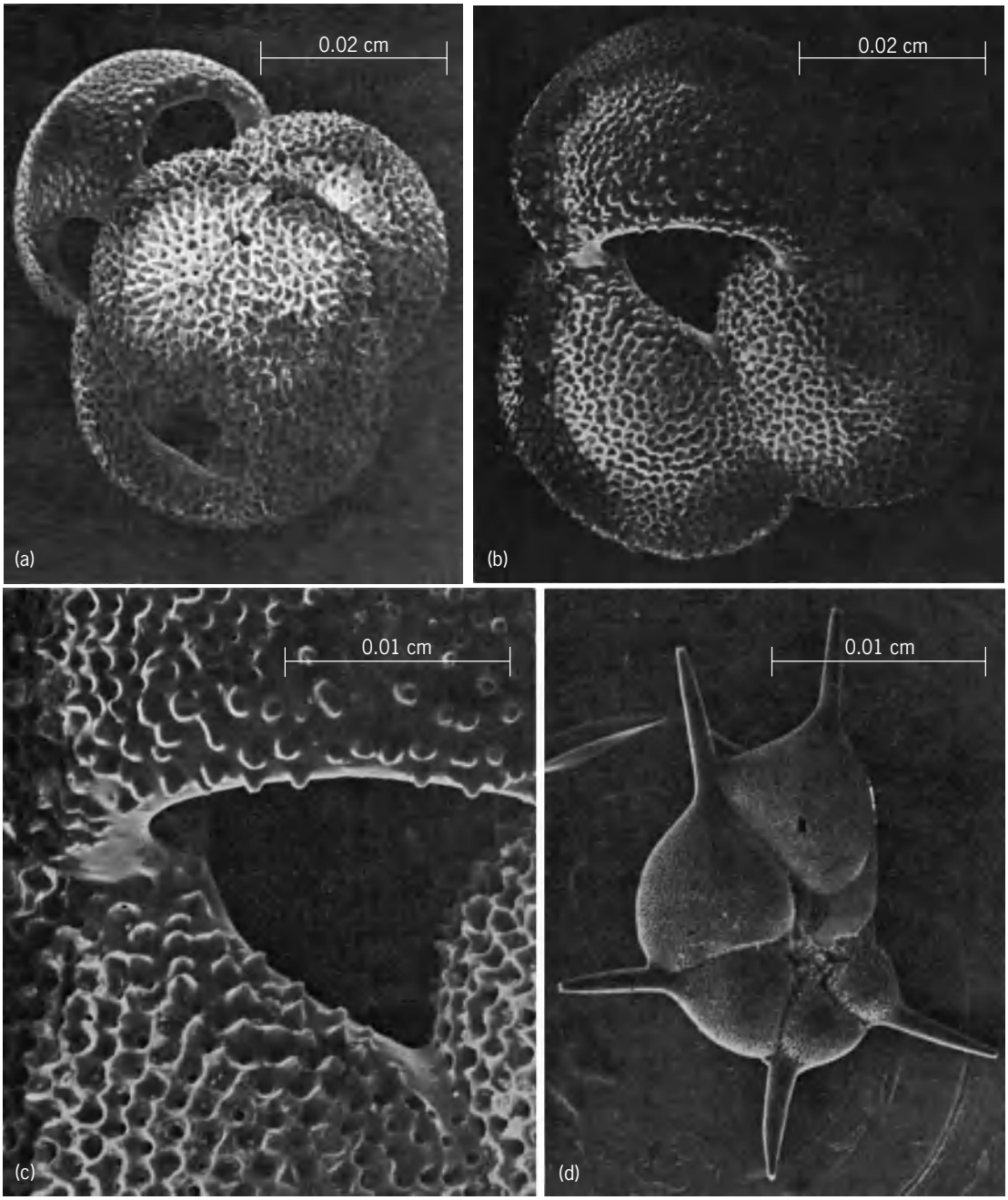

Fig. 11. Scanning electron micrographs of foraminiferans of suborder Globigerinina, superfamily Globigerinacea. *Globigerinoides*, from Holocene of Caribbean Sea: (*a*) spiral and (*b*) umbilical views of the thin-walled, coarsely perforate test of the planktonic species; (*c*) enlargement of apertural region of *b* showing imperforate apertural margin and coarse perforations elsewhere; short, blunt spines are the broken ends of very fine elongate delicate spines that project radially during life. Superfamily Hantkeninacea, from Upper Eocene of Mississippi. (*d*) *Hantkenina*, showing perforate chamber wall and imperforate tubulospines and apertural flap of last chamber at center right. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

surface coating may have parallel groups of calcite laths. True pores are lacking, but pseudopores may occur in the surface layer; these end in the randomly oriented layer below and do not completely penetrate the wall.

Microgranular calcite. The Fusulinina have a microgranular wall of tiny, tightly packed equidimensional subangular calcite crystals (**Figs. 3–5**). The granules may be regularly aligned, producing a fibrous appearance; as many as four separate granular and fibrous wall layers may occur in various combinations. Some Fusulinacea have perforated and even alveolar walls. Other characteristic structures result from solution and secondary deposits formed in the enlarging test.

Lamellar hyaline tests. Hyaline calcareous walls occur in the suborders Lagenina, Spirillinina, Robertinina, Globigerinina, and Rotaliina (**Figs. 6–15**). Successive laminae are added with growth. A thick organic lining, probably a polysaccharide, is just within the calcareous wall; constructed first, it gives form

Fig. 12. Scanning electron micrograph of *Spiroplecta*, suborder Globigerinina, superfamily Heterohelicacea, showing biserial test and inflated chambers and weakly striate surface; from Upper Cretaceous, Texas. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

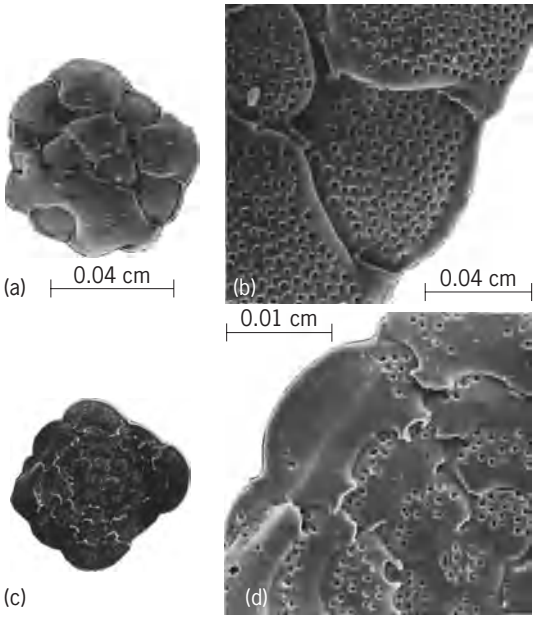

Fig. 13. Scanning electron micrographs of *Planorbulina*, suborder Rotaliina, superfamily Planorbulinacea, test with trochospiral early stage and proliferating later stage; from Pliocene of Italy. (*a*) Umbilical view; (*b*) enlargement, showing large pores on surface and two or more large apertures at the edge of each chamber of the outer whorl; (*c*) view of attached spiral side; (*d*) enlargement of *c* showing fewer perforations and small apertural projections from each later chamber. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

and shape to the newly added chamber. Electron microscopy has detected a proteinaceous basal membrane just outside the lining that serves as the organic matrix for calcification. As each new chamber is added, its new basal membrane also extends over the previously formed test; a new calcareous lamina is added to all exposed surfaces of previous chambers as the newly formed one is calcified. The test lamellae are not merely juxtaposed calcite layers, but the calcite layers are interspersed with fine separating membranes. Foraminiferans with numerous chambers in each whorl develop very thick walls, layer by layer with each chamber addition (as in *Nummulites*), whereas those with few chambers per whorl remain thin-walled (*Globigerina*). The inner organic lining is also added against the previous apertural face and, within earlier chambers, continually thickening with new chamber additions. The resultant thick inner lining of early chambers may give them a distinct brown coloration in the living test, as in *Rosalina*.

Lamellar septa. Depending upon the method of its formation, the new test extension or new chamber wall may consist of a single layer (Textulariina, Involutinina, Spirillinina, Miliolina, Lagenina) or be double, because of calcification at both sides of the organic membrane serving as a template for the new chamber. As the next chamber is added, the previous apertural face becomes an intercameral septum, retaining the original monolamellar or bilamellar character, whereas the outer test wall may be enveloped by an additional lamella at each instar. The originally bilamellar septa of the Rotaliacea and Nummulitacea are additionally thickened by a septal flap attached to the previous apertural face as the new chamber is formed. Spaces left between this septal flap and the septum may form a complex system of radial canals, umbilical cavities, and intraseptal and subsutural canals in these groups.

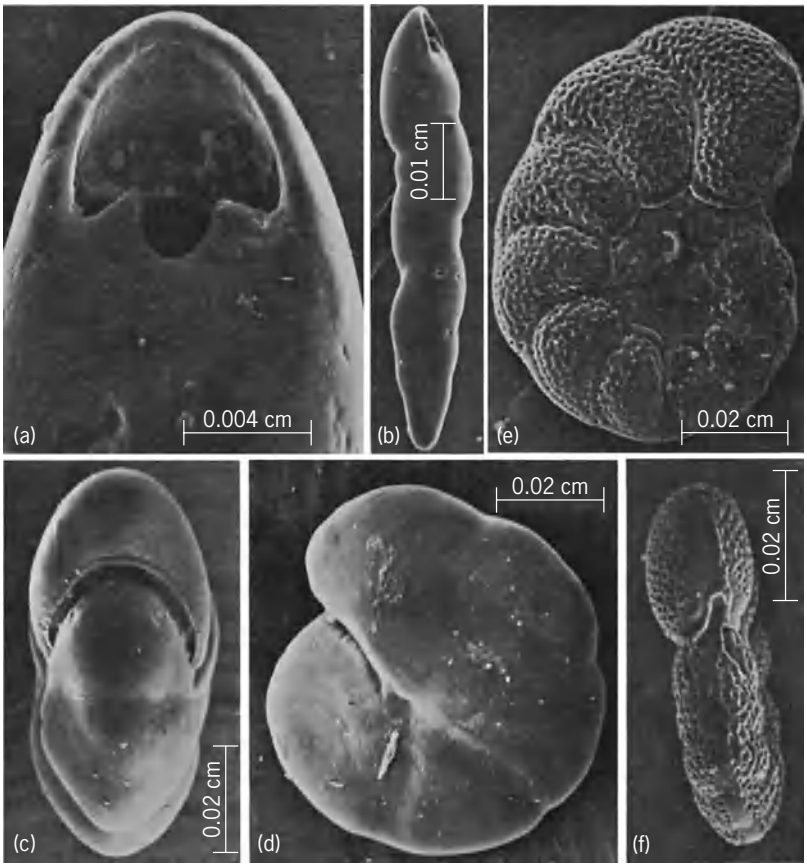

Fig. 14. Scanning electron micrographs of foraminiferans of suborder Rotaliina. Superfamily Pleurostomellacea: *Pleurostomella*, from Lower Cretaceous of England: (*a*) apertural view of final chamber showing two projecting teeth; (*b*) side view of biserial to uniserial test with perforate granular wall and hooded aperture. Superfamily Nonionacea: *Pullenia*, from Holocene of Antarctica: (*c*) apertural and (*d*) side views of planispiral finely perforate test with granular wall. Superfamily Chilostomellacea: *Holmanella*, from Miocene of California: (*e*) spiral and (*f*) edge views of bievolute, planispiral test, with a coarsely perforate optically granular wall, imperforate peripheral margin, and slitlike aperture extending up the terminal face. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

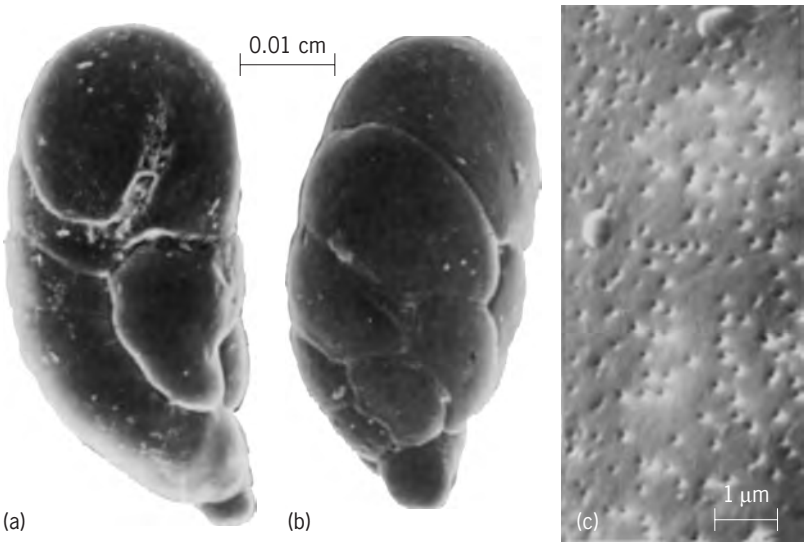

Fig. 15. Scanning electron micrographs of *Robertina*, suborder Robertinina, superfamily Robertinacea, from Pleistocene of California. (*a*) Apertural and (*b*) opposite sides of test with median subdivisions of the chambers; (*c*) surface with very fine pores in optically radial aragonitic wall. (*Courtesy of R. B. MacAdam, Chevron Oil Field Research Co.*)

Pores. Wall perforations of the hyaline foraminiferans may be extremely tiny or relatively large (Figs. 6*f*, 8, 10, 11, 13, 14, and 15). Two sizes may be interspersed; small pores may enlarge, coalesce, or become closed with lamellar growth; and some hyaline species seem completely lacking in perforations. Pores may be randomly spaced or restricted to certain areas of the chambers. Apertural lips, tooth plates, and surface ribs and carinal bands are commonly imperforate. Pores may contain organic plugs at the level of successive wall lamellae; perforated plugs are regarded as having a filtering function, but others are solid. Pores also may penetrate the calcite wall but not the organic lining and hence be nonfunctional as passages. Pores are also organically lined and may be only a remnant of the organic matrix, strengthening the test by binding the organic membranes of successive lamellae.

Radial and granular structure. Two main types of hyaline microstructure occur, termed optically radial (or hyaline radial) and optically granular (hyaline oblique), in reference to the orientation of the calcite or aragonite crystals in the wall. In the optically radial structure, each growth lamella consists of many layers of subrhombic calcite crystals, whose preferred orientation, with *c* axes normal to the test surface, gives a characteristic extinction pattern (black cross with concentric rings of color) under crossed Nicol prisms in the polarizing microscope. This occurs in the Spirillinina, Lagenina, Globigerinina, and many Rotaliina. Walls of the Robertinina are similarly constructed of the orthorhombic aragonite form of calcium carbonate. Crystals in the wall of Lagenina have irregular outlines, the surface ultrastructure showing irregularly formed interlocking units, with boundaries marked by a honeycomblike organic matrix. Crystal boundaries are irregular, but their preferred orientation is like that of other hyaline radial walls. In Spirillinina the calcite of the entire test is in optical continuity, acting as a single crystal in polarized light.

The optically granular or hyaline oblique wall (Fursenkoinacea, Delosinacea, Pleurostomellacea, Annulopatellinacea, Chilostomellacea, most Nonionacea, some Loxostomatacea, Cassidulinacea, Turrilinacea, and Siphoninacea, and a few Rotaliacea) contains smaller calcite grains, with preferred crystal orientation on a rhombohedral face, so that the *c* axis is oblique to the surface instead of perpendicular. In the polarizing microscope the granular crystals show only tiny flecks of color instead of a black cross.

In addition, two other types of wall structure have been described, an intermediate structure and a compound structure. The intermediate structure occurs in some Planorbulinacea (*Cibicides*, *Planorbulina*) and Chilostomellacea (*Gyroidina*). In this structure the crystal units are small and do not extend continuously through the wall. Near the outer surface the *c* axes are nearly normal to the wall surface as in the radial structure, but in the interior the crystal units are oblique to the surface as in the granular type. Hence, under polarized light the extinction pattern shows characteristics of both radial and granular walls.

The compound structure found in *Lobatula* (superfamily Planorbulinacea) consists of a wall with separate layers of differing type. An inner layer has *c* axes perpendicular to the test surface as in the radial type, but the outer primary layer has *c* axes and one edge of each microcrystal parallel to the test surface, the combined layers resulting in a granular extinction pattern under polarized light.

Spicular structure. The wall of the suborder Carterinina contains many elongate rodlike or fusiform spicular crystals whose *c* axes lie lengthwise of the spicules. The spicules or aligned bundles of spicules are held in a groundmass of similar smaller spicules, an inner layer being aligned normal to the test periphery and a surface layer being aligned parallel to the test surface and periphery. Later chambers have progressively larger spicules, whose lamellar structure indicates their further enlargement within the test.

Classification. As is true of most protists with skeletons or tests, systematic differentiation and classification of foraminiferans is based on test composition, microstructure, and gross morphology. Information available concerning cytoplasmic characters, life cycles, and so on has shown good agreement with this classification, although the function and origin of many shell characters believed to be of systematic importance (canal systems, pores, septal doubling, and apertural tooth plates) are not all understood. Characters of suborders and superfamilies are given in the **table**.

Evolutionary development. Foraminiferans have shown a remarkable diversification and rapid evolutionary development over their 550-million-year known history. Some with uncomplicated tests have been reported to have a long geologic range, but most with diagnostic shell structure and morphology had relatively short histories, each successively

Characteristics of foraminiferan suborders and superfamilies

Superfamily	Test wall	Test form	Distribution
Suborder Allogromiina	Organic		
Lagynacea (5 families, 56 genera)	Membranous to tectinous; surface may be agglutinated or ferruginous	Single ovid, tubular, or irregular chamber, rarely multilocular; free or attached; one or more apertures	Marine, brackish, fresh water; benthic, a few planktonic; Cambrian to Holocene
Suborder Textulariina	Agglutinated with organic or calcareous cement		
Astrorhizacea (12 families, 93 genera)	Simple to labyrinthic; with organic lining and organic cement	Irregular, rounded, tubular or branching nonseptate chamber, or cluster of chambers; one or more apertures	Marine, benthic; Cambrian to Holocene
Komokiacea (2 families, 7 genera)	Silt particles in organic cement; organic lining	A system of branching tubules; aperture consists of wall pores	Marine, benthic; abyssal and hadal; Holocene
Hippocrepinacea (3 families, 18 genera)	Of quartz particles and sponge spicules; organic cement	Proloculus and undivided tubular or flaring second chamber; free or attached; aperture terminal, single	Marine, benthic; Ordovician to Holocene
Ammodiscacea (1 family, 31 genera)	Arenaceous; organic cement	Proloculus and enrolled to uncoiled, tubular, nonseptate second chamber; aperture terminal, single	Marine, benthic; Cambrian to Holocene
Rzehakinacea (1 family, 11 genera)	Finely agglutinated; organic cement	Enrolled, two to three chambers per whorl; planispiral or coiled in varied planes	Marine, or brackish to saline lakes; benthic; Cretaceous to Holocene
Hormosinacea (6 families, 32 genera)	Firmly cemented and agglutinated; thin with single layer of grains to very thick; may be perforated by coarse tubuli	Multilocular, chambers uniserial to branching; free or attached; one or more apertures	Marine, bathyal to abyssal; Ordovician to Holocene
Lituolacea (9 families, 62 genera)	Agglutinated on organic base; no pores	Multilocular; free or attached; planispiral to uncoiled; aperture single to multiple, basal to terminal	Brackish to marine, benthic; neritic to abyssal; Devonian to Holocene
Haplophragmiacea (7 families, 25 genera)	Agglutinated, solid to alveolar, rarely with canaliculate septa	Streptospirally enrolled to uncoiled, may branch; chambers simple or with inner partition; aperture basal to terminal	Brackish to marine, benthic; neritic to bathyal; Jurassic to Holocene
Biokovinacea (3 families, 11 genera)	Finely agglutinated, imperforate outer layer; inner part may be perforate or canaliculate	Planispiral or may uncoil; aperture single to multiple, basal to areal	Benthic, marine, on calcareous substrate; Jurassic to Upper Cretaceous
Coscinophragmatacea (2 families, 5 genera)	Coarsely perforate to alveolar; pores closed by organic inner lining of wall	Attached, coiled to uncoiled or branching; terminal aperture, rounded to irregular	Benthic, marine; Triassic to Holocene
Cyclolinacea (2 families, 10 genera)	Microgranular, calcareous, agglutinated; with imperforate epidermis	Discoidal; early stage coiled, later peneropliform to cyclic chambers; may have secondary partitions or pillars; aperture multiple, may result in radial or oblique stolon system	Benthic, marine; in calcareous sediments; Jurassic to Upper Cretaceous
Loftusiacea (6 families, 47 genera)	Agglutinated, with imperforate outer layer and alveolar or reticulate subepidermal layer	Multilocular, coiling planispiral to streptospiral, rarely trochospiral; later may uncoil or become biserial; interior simple, or with pillars, beams, and rafters; aperture single to multiple, basal to areal, or terminal	Marine, benthic; Triassic to Holocene
Spiroplectamminacea (5 families, 27 genera)	Solid and noncanaliculate, agglutinated with organic cement	Early planispiral to streptospiral stage, rarely biserial; later may be biserial or uniserial; chambers simple, undivided; aperture single to multiple, basal to areal or terminal	Marine, benthic; upper Carboniferous to Holocene
Pavonitacea (2 families, 7 genera)	Noncanaliculate	Early stage coiled, triserial or biserial, later biserial or uniserial; interior may have vertical beams, rarely rafters; aperture basal to terminal, single to multiple	Marine, benthic; Upper Cretaceous to Pliocene
Trochamminacea (2 families, 32 genera)	Agglutinated, organic cement, noncanaliculate	Low trochospiral coil, or uncoiling; secondary partitions due to infolding of wall; aperture basal to areal, single to multiple, may have supplementary umbilical openings	Benthic, brackish to marine; upper Carboniferous to Holocene
Verneuilinacea (4 families, 39 genera)	Agglutinated, solid, noncanaliculate	Elongate; early stage trochospiral, triserial, or biserial, later may be uniserial and rectilinear; aperture basal, areal, or terminal on neck	Benthic, marine; upper Carboniferous to Holocene
Ataxophragmiacea (8 families, 61 genera)	Noncanaliculate, agglutinated; may be alveolar or have imperforate epidermis and reticulate subepidermal layer	Trochospiral, later reduced to biserial or uniserial, or with increased number of chambers per whorl; internally may have secondary partitions, buttresses, or pillars; aperture basal to areal, simple to cribrate	Benthic, marine; Triassic to Holocene
Orbitolinacea (1 family, 41 genera)	Noncanaliculate, agglutinated	Low to high conical; early stage trochospiral or with protoconch and deuteroconch; later chambers uniserial, discoidal, numerous, with radial and transverse partitions and vertical pillars	Benthic, marine; Jurassic to Oligocene
Textulariacea (7 families, 65 genera)	Agglutinated, calcareous cement, canaliculate	Trochospiral, triserial or biserial in early stage; later biserial or uniserial; chambers simple or with internal partitions or pillars; aperture basal to areal, single to multiple	Benthic, marine; Jurassic to Holocene

(*cont.*)

Characteristics of foraminiferan suborders and superfamilies (cont.)

Superfamily	Test wall	Test form	Distribution
Suborder Fusulinina	Nonlamellar, microgranular calcite, of equidimensional crystals a few micrometers in diameter		
Parathuramminacea (10 families, 44 genera)	Homogeneous, one to three layers	Single globular to irregular chamber or cluster of chambers; free or attached; aperture not recognizable, or multiple, on necks or conical projections	Benthic, marine; Silurian to Permian
Earlandiacea (3 families, 11 genera)	Simple, rarely with some agglutinated material	Globular proloculus and undivided straight or enrolled tubular second chamber; aperture terminal	Benthic, marine; Silurian to Permian
Archaediscacea (2 families, 21 genera)	One or more layers, one with radial structure, one compact; one or both sides may have axial thickenings	Proloculus followed by planispiral, streptospiral, or trochospiral tubular second chamber; aperture at open end of tubular chamber	Benthic, marine; Carboniferous
Moravamminacea (3 families, 12 genera)	Single-layered	Proloculus followed by enrolled or rectilinear tubular second chamber; incipient or partial septa present; aperture terminal	Benthic, marine; Silurian to lower Carboniferous
Nodosinellacea (2 families, 7 genera)	Simple, or with outer granular layer and fibrous inner layer	One or more distinct uniserial chambers, rectilinear to arcuate; aperture terminal	Benthic, marine; Silurian to Permian
Geinitzinacea (2 families, 8 genera)	Two layers, outer one radially fibrous and inner one dark and microgranular; secondarily nonlamellar or partially lamellar	Uniserial; rounded to flattened in section; aperture terminal	Benthic, marine; Devonian to Permian
Colaniellacea (1 family, 4 genera)	Two layers, outer one vitreous, inner one finely granular	Uniserial; chambers strongly overlapping; with internal radial, vertical interseptal partitions; aperture terminal	Benthic, marine; Devonian to Permian
Ptychocladiacea (1 family, 2 genera)	Microgranular, banded, possibly due to transverse tubuli	Attached; uniserial, branching or spreading; no distinct aperture	Benthic, marine; Devonian to upper Carboniferous
Palaeotextulariacea (3 families, 22 genera)	One or two layers; may have agglutinated coating	Biserial, enrolled biserial, or uniserial; advanced ones with secondary vertical partitions; aperture basal to terminal, single to multiple	Benthic, marine; Devonian to Permian
Tournayellacea (2 families, 48 genera)	One or two layers; may have agglutinated inclusions; may have secondary deposits	Proloculus and tubular planispirally to streptospirally coiled second chamber, or may uncoil; incipient septa may produce teardrop-shaped chambers; single to cribrate aperture	Benthic, marine; Devonian to upper Carboniferous
Endothyracea (2 families, 59 genera)	Two to three layers; may have agglutinated particles also; supplementary deposits may occur	Planispiral to streptospiral coiling; later may be uncoiled and rectilinear; true septa present; basal to areal aperture, single or rarely cribrate; may have supplementary sutural openings	Benthic, marine; Devonian to Permian
Tetrataxacea (4 families, 7 genera)	One or two layers	Conical; trochospiral, evolute spiral side and involute umbilical side; few to many chambers per whorl; numerous secondary partitions; aperture umbilical	Benthic, marine; lower Carboniferous to Permian
Fusulinacea (8 families, 167 genera)	One to four layers; resorption may produce foramina or tunnels; secondary deposits form axial fillings, chomata, parachomata, and tectoria	Globular, lenticular, or fusiform; tightly coiled or may later uncoil; many chambers per whorl; septa plane to fluted; advanced ones with secondary septula; aperture a basal slit to multiple openings	Benthic, marine; Devonian to Permian
Suborder Involutinina	Homogeneous, calcareous originally aragonite but commmonly recrystallized		
(4 families, 30 genera)	Perforate, radiate; may be recrystallized to homogeneous microgranular structure; lamellar thickening or pillars on one or both sides	Proloculus followed by planispiral to trochospiral enrolled, undivided tubular second chamber; aperture at open end of tubular chamber	Benthic, marine; in calcareous deposits; Permian to Upper Cretaceous
Suborder Spirillinina	Optically single calcite crystal; formed by marginal accretion		
(2 families, 11 genera)	Of calcite crystals, which may show preferred orientation; pseudopores of umbilical surface filled with organic matter; marginal calcification, not on organic template	Planispiral to high trochospiral coil; proloculus and coiled undivided second chamber or with two to three chambers per whorl; may have secondary radial partitions; aperture at open end of tube on periphery or directed toward umbilicus	Benthic, shallow marine; Triassic to Holocene
Suborder Carterinina	Rodlike spicules secreted by the organism		
(1 family, 1 genus)	Inner organic lining, rodlike or fusiform spicules, each a single low-magnesium calcite crystal oriented with respect to periphery, held in organic matrix	Attached; trochospiral early chambers becoming crescentic and finally irregular, later ones with secondary radial beams; single umbilical aperture, later becoming multiple	Benthic, marine, shallow tropical water; Eocene, Holocene

Characteristics of foraminiferan suborders and superfamilies (cont.)

Superfamily	Test wall	Test form	Distribution
Suborder Miliolina	Imperforate porcelaneous, of fine randomly oriented rodlike crystals of high-magnesium calcite, containing organic material, may have surface veneer of tabular crystals		
Squamulinacea (1 family, 2 genera)	As in suborder	Single globular to subcylindrical chamber; free or attached	Benthic, marine; Upper Cretaceous, Holocene
Cornuspiracea (7 families, 73 genera)	As in suborder	Free or attached; planispiral, streptospiral, trochospiral, uncoiled, or irregular, spreading to discoidal; proloculus and undivided second chamber, later chambers may be tubular, rounded, or cyclic, and may be subdivided; aperture terminal or row of pores	Benthic, marine; Lower Carboniferous to Holocene
Miliolacea (6 families, 113 genera)	As in suborder; wall may have pseudopores	Enrolled, two or rarely three tubular chambers per whorl, added in various planes around elongate axis; may have longitudinal or transverse secondary subdivisions; aperture terminal, single or multiple	Benthic, marine, mostly warm water; Triassic to Holocene
Alveolinacea (3 families, 32 genera)	As in suborder; may have basal wall thickening, subepidermal partitions, or pillars	Enrolled about elongate axis; proloculus and flexostyle followed by planispiral, streptospiral or milioline stage; adult may uncoil; may have internal septulae or pillars; aperture multiple, rarely single	Benthic, marine, shallow tropical water; Lower Cretaceous to Holocene
Soritacea (6 families, 59 genera)	As in suborder; early stage coarsely perforate or pitted, later imperforate; some extinct taxa perforate throughout	Planispiral to uncoiling, flabelliform or discoid; may have internal interseptal partitions or pillars; aperture single or more commonly multiple	Benthic, marine, warm shallow water; Permian to Holocene
Suborder Silicoloculinina (1 family, 1 genus)	Secreted rods of opaline silica Imperforate; inner and outer organic membranes, with median layer of randomly arranged siliceous rods bordered above and below by rods that parallel surface	Enrolled, proloculus followed by chambers one-half coil in length, added in varying planes as in miliolines	Benthic, abyssal at high latitudes; Miocene to Holocene
Suborder Lagenina	Optically and ultrastructurally radiate hyaline calcite wall, crystal *c* axis perpendicular to surface, with organic membrane around crystal units		
Robuloidacea (4 families, 30 genera)	Monolamellar, finely perforate; secondary lamination absent or partial in more advanced taxa	Elongate, enrolled, rectilinear, or arcuate; proloculus followed by undivided tubular chamber, or with pseudosepta or true septa and uniserial, or spiraling about long axis; terminal aperture round, slitlike or radiate	Marine, benthic; lower Carboniferous to Lower Cretaceous
Nodosariacea (6 families, 191 genera)	As above, but with secondary lamination resulting from continued growth	Free-living, or rarely attached; single chamber or multilocular, inrolled, arcuate, or rectilinear, or spiraling about long axis; aperture single or multiple, round, slitlike or radiate, and may have entosolenian tube	Marine, benthic; Triassic to Holocene
Suborder Robertinina	Hyaline, of optically and ultrastructurally radial aragonite		
Duostominacea (3 families, 9 genera)	Appears granular; probably was aragonitic but commonly recrystallized; nonlamellar; may include some foreign particles	Planispiral or low to high trochospiral coil, few chambers per whorl; aperture basal, single or double	Marine, benthic; Triassic to Jurassic
Ceratobuliminacea (2 families, 27 genera)	Finely perforate, optically radial aragonite; with primary and secondary lamination, septa primarily bilamellar; surface smooth, pustulose or costate	Trochospiral, few chambers per whorl, chambers with internal partition; aperture basal or areal, may be closed as next chamber is added, and foramen is formed by resorption; may have slitlike peripheral opening	Marine, benthic; Jurassic to Holocene
Conorboidacea (1 family, 3 genera)	As above	Low to high trochospiral coil, later with fewer chambers per whorl to uniserial; aperture basal in trochospiral taxa, terminal in uniserial ones, with hemicylindrical columellalike toothplate	Marine, benthic; Lower Cretaceous to Upper Cretaceous
Robertinacea (1 family, 9 genera)	As above; surface smooth	Low to high trochospiral, transverse inner partition due to infolding of wall produces chamberlets; aperture basal or high loop or both, part remaining as septal foramen; may have supplementary openings	Marine, benthic; Paleocene to Holocene

(*cont.*)

Characteristics of foraminiferan suborders and superfamilies (cont.)

Superfamily	Test wall	Test form	Distribution
Suborder Globigerinina	Perforate hyaline radial calcite wall, *c* axis normal to surface, primarily bilamellar septa		
Heterohelicacea (3 families, 21 genera)	Primary and secondary lamination; surface smooth or striate; may have pore mounds	Trochospiral, triserial or biserial, may reduce to uniserial or show chamber proliferation; aperture a basal arch, terminal in uniserial stage; may have supplementary or accessory openings	Marine, planktonic; Lower Cretaceous to Holocene
Planomalinacea (3 families, 8 genera)	As above, surface smooth, pitted, or with inflational pustules	Planispiral or tending to trochospiral; aperture equatorial, basal, part may remain as relict openings	Marine, planktonic; Lower to Upper Cretaceous
Rotaliporacea (4 families, 17 genera)	As above; surface smooth, pitted, or with inflational pustules; may have imperforate carinal band or keel	Trochospiral; primary basal aperture umbilical-extraumbilical; may also have secondary sutural openings	Marine, planktonic; Lower to Upper Cretaceous
Globotruncanacea (2 families, 17 genera)	Surface smooth, pitted, or with inflational pustules; may have imperforate carinal band	Trochospiral, chambers globular to subangular; primary aperture umbilical, umbilical area with tegilla; may have accessory apertures	Marine, planktonic; Upper Cretaceous
Globorotaliacea (6 families, 40 genera)	Finely lamellar surface smooth, pitted, or cancellate, with inflational pustules or muricae; nonspinose	Trochospiral, chambers rounded to radial elongate; periphery rounded to carinate; primary aperture basal, umbilical-extraumbilical, with imperforate lip; may have supplementary sutural apertures and bullae with accessory openings	Marine, planktonic at medium depths; Paleocene to Holocene
Hantkeninacea (3 families, 8 genera)	Surface smooth to finely pustulate	Planispiral to enrolled biserial; globular to elongate chambers, may have tubulospines; single or multiple equatorial aperture, basal to areal; may have relict secondary openings	Marine, planktonic; Paleocene to Miocene
Globigerinacea (2 families, 24 genera)	Surface smooth or costellate, with nonlamellar, monocrystalline elongate spines set in holes in wall; crystal *c* axes lengthwise of spine	Trochospiral, planispiral, streptospiral, or enveloping; aperture basal and umbilical, umbilical-extraumbilical, or equatorial; may have secondary sutural openings, and may have bullae	Marine, planktonic in surface waters; Eocene to Holocene
Suborder Rotaliina	Perforate hyaline lamellar calcite, deposited at each side of organic lamella; optically radial or granular, according to normal to surface or oblique crystal orientation		
Bolivinacea (2 families, 12 genera)	Optically radial; surface smooth, striate or costate	Biserial, or may later become uniserial; aperture a high loop or slit, basal to areal, with internal toothplate	Benthic, marine; Upper Cretaceous to Holocene
Loxostomatacea (3 families, 7 genera)	Optically radial or granular; surface smooth, pustulose, or costate	Biserial, or may tend to uniserial; aperture single to cribrate, basal to terminal, no toothplate	Benthic, marine; Upper Cretaceous to Holocene
Bolivinitacea (1 family, 2 genera)	Optically radial; surface smooth, striate, reticulate or nodose	Biserial, flattened, truncate margins; one border of high opening bent inward to form toothplate	Benthic, marine; Miocene to Holocene
Cassidulinacea (2 families, 20 genera)	Optically radial or granular; surface smooth, papillate or cancellate	Biserial or derived from such; biserial axis enrolled in planispiral or rarely trochospiral coil, or may uncoil; basal to terminal aperture may have toothplate; may have sutural secondary openings	Benthic, marine; Paleocene to Holocene
Eouvigerinacea (2 families, 6 genera)	Optically radial, surface smooth, hispid or pustulose	Biserial or uniserial, may have early planispiral coil; aperture with hemicylindrical internal toothplate	Benthic, marine; Lower Cretaceous to Eocene; Pliocene
Turrilinacea (3 families, 19 genera)	Optically radial or granular; surface smooth, pustulose, longitudinally costate, striate, or reticulate	Trochospiral or triserial, later may reduce to biserial or uniserial; aperture a basal arch to areal, single to cribrate, may have simple toothplate	Benthic, marine; Jurassic to Holocene
Buliminacea (8 families, 57 genera)	Optically radial; surface smooth, longitudinally costate, hispid, or laterally carinate	High trochospiral or triserial, may reduce to biserial or uniserial; aperture a basal loop to terminal and rounded, internal toothplate attached to apertural margin; rarely with internal subdivisions	Benthic, marine; Upper Cretaceous to Holocene
Fursenkoinacea (2 families, 11 genera)	Optically granular, surface smooth or hispid in early stage	Triserial or twisted biserial; aperture loop-shaped with internal toothplate; may have sutural supplementary openings	Benthic, marine; Upper Cretaceous to Holocene
Delosinacea (3 families, 10 genera)	Optically granular, surface smooth or with sparse basal spines	Trochospiral or triserial, later may be biserial or uniserial; aperture a high basal slit with toothplate; may have sutural pores	Benthic, marine; Lower Cretaceous to Holocene
Pleurostomellacea (1 family, 16 genera)	Optically granular, surface smooth or rarely striate or costate	Triserial or biserial, may become uniserial; straight to curved slit aperture excentric, may have hood, or may be terminal and cribrate; may have internal siphon	Benthic, marine; Lower Cretaceous to Holocene

Characteristics of foraminiferan suborders and superfamilies (cont.)

Superfamily	Test wall	Test form	Distribution
Suborder Rotaliina (cont.)			
Stilostomellacea (1 family, 5 genera)	Optically radial, surface smooth, weakly striate to costate, or with sparsely hispid margins	Uniserial, arcuate to rectilinear; terminal aperture with phialine lip and toothlike projection	Benthic, marine; Upper Cretaceous to Holocene
Annulopatellinacea (1 family, 1 genus)	Optically granular, surface smooth	Conical; proloculus and reniform second chamber, then uniserial with discoid chambers; radial tubules open as pores at surface; plastogamic sexual reproduction	Benthic, shallow marine; Miocene to Holocene
Discorbacea (15 families, 80 genera)	Optically radial, finely to coarsely perforate, imperforate around aperture; surface smooth to pustulose, may have inflational peripheral spines or keel	Low trochospiral coil, rarely uncoils; may have secondary septula; basal aperture on umbilical side or areal and cribrate; may have secondary sutural openings	Benthic, may have planktonic stage; marine; Jurassic to Holocene
Glabratellacea (3 families, 22 genera)	Optically radial, surface smooth to pustulose, may be carinate; umbilical side with radial striae, costae or nodes that aid plastogamy	Low to high trochospire, aperture basal and umbilical; sexual reproduction plastogamic	Benthic, marine; Eocene to Holocene
Siphoninacea (1 family, 5 genera)	Optically radial or granular; commonly with imperforate fimbriate keel	Trochospiral to planispiral, may uncoil, become biserial or uniserial; aperture slitlike to rounded, on neck with phialine lip	Benthic, marine; Upper Cretaceous to Holocene
Discorbinellacea (4 families, 22 genera)	Optically radial, sparsely perforate or lacking pores; surface smooth	Trochospiral, rarely has internal partitions subdividing chamber; aperture basal, an equatorial, low arch, slit up face, or areal; may have supplemental apertures	Benthic, marine; Paleocene to Holocene
Planorbulinacea (6 families, 60 genera)	Optically radial or with intermediate structure, inner and outer layers of calcite crystals normal to the surface and median layer oblique; coarsely perforate	Free or attached; discoidal, lenticular, or conical; trochospirally coiled, uncoiled, rectilinear or biserial, or with chamber proliferation or irregularly added chambers; aperture basal, extraumbilical to equatorial, subterminal if uncoiled; may have supplementary or accessory openings	Benthic, marine; Lower Cretaceous to Holocene
Acervulinacea (2 families, 11 genera)	Optically radial wall; coarsely perforate	Free or attached; may be encrusting, or may grow erect, spiral early stage followed by spreading or irregular chambers; aperture of mural pores or pores in cribrate upper surface of chambers	Benthic, marine warm water; Paleocene to Holocene
Asterigerinacea (7 families, 39 genera)	Optically radial, surface smooth or granulose	Trochospiral to nearly planispiral; chambers with inner partition that attaches to outer wall and forms small umbilical chamberlet; may have distinct equatorial and lateral layers of chambers; aperture basal to areal, on umbilical side to equatorial, may have sutural supplementary openings; may have stolon system	Benthic, marine; Upper Cretaceous to Holocene
Nonionacea (3 families, 30 genera)	Optically granular, rarely optically radial; finely to coarsely perforate	Planispirally coiled to asymmetrical, involute or evolute; aperture simple to cribrate; equatorial, basal to areal, and may have additional peripheral or sutural openings	Benthic, marine; Upper Cretaceous to Holocene
Chilostomellacea (11 families, 69 genera)	Optically granular; finely to coarsely perforate; surface smooth to pustulose	Free or attached; trochospirally coiled, may uncoil or have enveloping chambers; may have inner partition related to toothplate; aperture basal, a vertical slit, umbilical to peripheral, or terminal in uncoiled forms, may have lip orflap, may have secondary sutural openings	Benthic, marine; Lower Cretaceous to Holocene
Orbitoidacea (3 families, 23 genera)	Optically radial, surface smooth or pustulose; may have fissures	Large, discoidal to lenticular, dimorphism prominent, microspheric test coiled or biserial at first, megalospheric test with bilocular embryo; later stage with similar or differentiated equatorial and lateral chambers, intercameral openings modified as stolons, but no canal system	Benthic, marine; Upper Cretaceous to Upper Eocene
Rotaliacea (6 families, 79 genera)	Optically radial, or rarely optically granular, finely to coarsely perforate, surface smooth, pustulose or costate, may have large coarse spines; septal flap attaches to former apertural face as new chamber is added, enclosing system of canals and cavities	Trochospiral or planispiral coiling, evolute to involute, chambers small and numerous, may be subdivided by radial plates or rods; primary aperture single or multiple, basal to areal; sutural openings connect to canal system, not to chamber interior	Benthic, marine; Upper Cretaceous to Holocene
Nummulitacea (4 families, 22 genera)	Optically radial; surface smooth to pustulose; septal flap present as in Rotaliacea	Planispiral, evolute to involute, dimorphism prominent; has many small chambers, may be subdivided into chamberlets, equatorial and lateral chambers may be differentiated, later chambers may be in annular series; aperture single to multiple; may have subsutural canals, spiral marginal cord, spiral canal system, and intraseptal canals	Benthic, marine; Paleocene to Holocene

being replaced by others. The oldest-known and presumably primitive tests are the morphologically simple tectinous and agglutinated ones of the Cambrian and Ordovician, with globular, tubular, branching, or irregular form. By the mid-Ordovician, chambered tests were present (*Reophax*), and in the Silurian, tests with microgranular calcareous walls also had appeared. Many early Paleozoic foraminiferans have been obtained from acid residues or thin sections of rock.

Microgranular tests diversified greatly in the late Paleozoic, true chamber development being followed by many wall and chamber modifications, layered and alveolar walls, labyrinthic chambers, crenulated septa, secondary septa, and chamberlets. Similar wall and chamber modifications were later repeated in other foraminiferal lineages.

Beginning in the Mississippian, various groups of foraminiferans successively were so abundant that their discarded tests accumulated as foraminiferal limestones, particularly noteworthy being the Mississippian endothyrids, Pennsylvanian and Permian fusulines, Cretaceous and Tertiary miliolids, and Eocene nummulites. Although the Fusulinina were abundant and varied in the Paleozoic, they then disappeared whereas the few porcelaneous and hyaline calcareous species that appeared in the late Paleozoic survived to expand greatly in the Mesozoic.

Planktonic foraminiferans became abundant in the Cretaceous and Tertiary; extensive pelagic limestones and chalks represent earlier accumulations like the present *Globigerina* and *Orbulina* oozes on the sea floor.

Ecology. Most foraminiferans are benthic, living upon the sea floor, within the upper few centimeters of ooze, or upon benthic algae or other organisms. They occur from the intertidal zone to oceanic depths, in brackish, normal marine, or hypersaline waters, and from the tropics to the poles. Some modern Lagynacea live in fresh water, but none are known as fossils. Assemblages vary widely in response to local conditions, with the greatest diversity occurring in warm, shallow water. A smaller number, the Globigerinina, are planktonic, living at various depths in the water column from the surface to the bottom, being most numerous between 18 and 90 ft (6 and 30 m). Vertical migration may be diurnal and may occur during ontogenetic development. The preferred depth range of a species may vary geographically in response to temperature differences or to changes in water density.

Abundant distributional data for living species have made possible the extensive use of foraminiferans for paleoecologic interpretation and age dating, their abundance permitting use of statistical methods. Paleoclimatic interpretations are based on temperature-related ratios of right- and left-coiling specimens of trochospiral planktonic species, the latitudinal gradation of species diversity, and the varied oxygen isotope ratios of the test carbonate. Faunal abundance and variety, test size, presence of trace elements, ratios of planktonic to benthic species, and ratios of calcareous to agglutinated tests, as well as gross morphology, have been the basis for other studies of such ecologic parameters as depth of water, distance from shore, sea-level fluctuations, paleocurrents, and relative oxygenation or stagnation of the bottom, reworked or displaced faunas, salinity, pH, turbidity and turbulence, nature of the substrate, rate of detrital influx, and rate of accumulation of sediments. Unsurpassed as stratigraphic markers for petroleum exploration or as tools for paleoecologic interpretation of ancient sediments, foraminiferans have proved to be equally valuable in oceanographic studies of the more recent events in the Earth's history.

Helen Tappan Loeblich

Bibliography. S. Bellemo, *Bull. Geol. Inst. Univ. Upsala*, new series, vol. 6, 1974; H. Bender and C. Hemleben, *Abhandlungen der Geologischen Bundesanstalt*, vol. 41, 1988; R. H. Hedley and C. G. Adams (eds.), *Foraminifera*, vol. 3, 1978; A. R. Loeblich, Jr., and H. Tappan, *Foraminiferal Genera and Their Classification*, 2 vols., 1987.

Force

Force may be briefly described as that influence on a body which causes it to accelerate. In this way, force is defined through Newton's second law of motion. *See* ACCELERATION.

This law states in part that the acceleration of a body is proportional to the resultant force exerted on the body and is inversely proportional to the mass of the body. An alternative procedure is to try to formulate a definition in terms of a standard force, for example, that necessary to stretch a particular spring a certain amount, or the gravitational attraction which the Earth exerts on a standard object. Even so, Newton's second law inextricably links mass and force. *See* MASS.

Many elementary books in physics seem to expect the beginning student to bring to his study the same kind of intuitive notion concerning force which Isaac Newton possessed. One readily thinks of an object's weight, or of pushing it or pulling it, and from this one gains a "feeling" for force. Such intuition, while undeniably helpful, is hardly an adequate foundation for the quantitative science of mechanics.

Newton's dilemma in logic, which did not trouble him greatly, was that, in stating his second law as a relation between certain physical quantities, he presumably needed to begin with their definitions. But he did not actually have definitions of both mass and force which were independent of the second law. The procedure which today seems most free of pitfalls in logic is in fact to use Newton's second law as a defining relation.

First, one supposes length to be defined in terms of the distance between marks on a standard object, or perhaps in terms of the wavelength of a particular spectral line. Time can be supposed similarly related to the period of a standard motion (for example, the rotation of the Earth about the Sun, the oscillations of the balance wheel of a clock, or perhaps a particular vibration of a molecule). Although applying

these definitions to actual measurements may be a practical matter requiring some effort, a reasonably logical definition of velocity and acceleration, as the first and second time derivatives of vector displacement, follows readily in principle.

Absolute standards. Having chosen a unit for length and a unit for time, one may then select a standard particle or object. At this juncture one may choose either the absolute or the gravitational approach. In the so-called absolute systems of units, it is said that the standard object has a mass of one unit. Then the second law of Newton defines unit force as that force which gives unit acceleration to the unit mass. Any other mass may in principle be compared with the standard mass (m) by subjecting it to unit force and measuring the acceleration ($\mathbf{a}$), with which it varies inversely. By suitable appeal to experiment, it is possible to conclude that masses are scalar quantities and that forces are vector quantities which may be superimposed or resolved by the rules of vector addition and resolution.

In the absolute scheme, then, Eq. (1) is written

$$\mathbf{F} = m\mathbf{a} \tag{1}$$

for nonrelativistic mechanics; here boldface type denotes vector quantities. The quantities on the right of Eq. (1) are previously known, and this statement of the second law of Newton is in fact the definition of force. In the absolute system, mass is taken as a fundamental quantity and force is a derived unit of dimensions MLT^{-2} (M = mass, L = length, T = time).

Gravitational standards. The gravitational system of units uses the attraction of the Earth for the standard object as the standard force. Newton's second law still couples force and mass, but since force is here taken as the fundamental quantity, mass becomes the derived factor of proportionality between force and the acceleration it produces. In particular, the standard force (the Earth's attraction for the standard object) produces in free fall what one measures as the gravitational acceleration, a vector quantity proportional to the standard force (weight) for any object. It follows from the use of Newton's second law as a defining relation that the mass of that object is $m = w/g$, with g the magnitude of the gravitational acceleration and w the magnitude of the weight. The derived quantity mass has dimensions FT^2L^{-1}.

Because the gravitational acceleration varies slightly over the surface of the Earth, it may be objected that the force standard will also vary. This may be avoided by specifying a point on the Earth's surface at which the standard object has standard weight. In principle, then, the gravitational system becomes no less absolute than the so-called absolute system. *See* GRAVITATION.

Composition of forces. By experiment one finds that two forces of, for example, 3 units and 4 units acting at right angles to one another at point 0 produce an acceleration of a particular object which is identical to that produced by a single 5-unit force inclined at arccos 0.6 to the 3-unit force, and

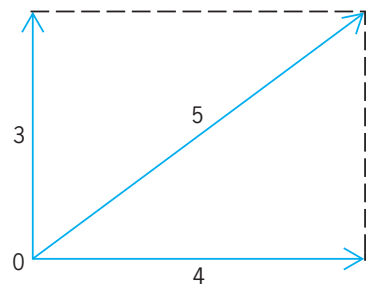

Vector addition of forces.

arccos 0.8 to the 4-unit force (see **illus.**). The laws of vector addition thus apply to the superposition of forces.

Conversely, a single force may be considered as equivalent to two or more forces whose vector sum equals the single force. In this way one may select the component of a particular force which may be especially relevant to the physical problem. An example is the component of a railroad car's weight along the direction of the track on a hill.

Statics is the branch of mechanics which treats forces in nonaccelerated systems. Hence, the resultant of all forces is zero, and critical problems are the determination of the component forces on the object or its structural parts in static equilibrium. Practical questions concern the ability of structural members to support the forces or tensions. *See* STATICS.

Specially designated forces. If a force is defined for every point of a region and if this so-called vector field is irrotational, the force is designated conservative. Physically, it is shown in the development of mechanics that this property requires that the work done by this force field on a particle traversing a closed path is zero. Mathematically, such a force field can be shown to be expressible as the (conventionally negative) gradient of a scalar function of position V, Eq. (2).

$$\mathbf{F} = -\nabla V \tag{2}$$

A force which extracts energy irreversibly from a mechanical system is called dissipative, or nonconservative. Familiar examples are frictional forces, including those of air resistance. Dissipative forces are of great practical interest, although they are often very difficult to take into account precisely in phenomena of mechanics.

The force which must be directed toward the center of curvature to cause a particle to move in a curved path is called centripetal force. For example, if one rotates a stone on the end of a string, the force with which the string pulls radially inward on the stone is centripetal force. The reaction to centripetal force (namely, the force of the stone on the string) is called centrifugal force. *See* CENTRIFUGAL FORCE; CENTRIPETAL FORCE.

Methods of measuring forces. Direct force measurements in mechanics usually reduce ultimately to a weight comparison. Even when the elastic distortion of a spring or of a torsion fiber is used, the calibration of the elastic property will often be through a balance which compares the pull of the spring with

a calibrated weight or the torsion of the fiber with a torque arising from a calibrated weight on a moment arm. *See* BALANCE.

In dynamic systems, any means of measuring acceleration—for example, through photographic methods or radar tracking—allows one to calculate the force acting on an object of known mass.

Units of force. In addition to use of the absolute or the gravitational approach, one must contend with two sets of standard objects and lengths, the British and the metric standards. All systems use the second as the unit of time. In the metric absolute system, the units of force are the newton and the dyne. The newton, the unit of force in the International System (SI), is that force which, when applied to a body having a mass of 1 kilogram, gives it an acceleration of 1 m/s^2. The poundal is the force unit in the British absolute system, whereas the British gravitational system uses the pound. Metric gravitational systems are rarely used. Occasionally one encounters terms such as gram-force or kilogram-force, but no corresponding mass unit has been named. *See* UNITS OF MEASUREMENT. George E. Pake

Bibliography. F. Bueche and D. A. Jerde, *Principles of Physics*, 6th ed., 1995; D. Halliday, F. Resnick, and K. Krane, *Physics*, 4th ed., 1992; F. J. Keller, W. E. Gettys, and M. J. Skove, *Physics*, 2d ed., 1993; C. Kittel, W. D. Knight, and M. A. Ruderman, *Berkeley Physics Course*, vol. 1: *Mechanics*, 2d ed., 1973.

Force analysis

An analysis yielding the respective forces acting at any point of any member, or part of a member, of a mechanism, obtained by using relationships for dynamic equilibrium in a plane rigid body subject to external forces within this plane and to internal forces due to its motion in this plane. The following treatment is a careful selection and condensation of essential features of the reasoning.

The basic relationships for equilibrium are that the forces acting upon a body, or their resultants, must act through a point (their force polygon is closed); and that the couples acting on the body also reduce to zero (the moment of the existing forces on any member about any point is zero). If there are no more than three unknowns, a solution can be found, utilizing the three conventional equations summing forces in two perpendicular directions to zero and summing moments to zero. If there are more unknowns, further information must be acquired from other members.

The method of approach may, in practice, be purely algebraic, for instance, using the three mentioned equations of force and moment equilibrium; or it may be graphical, using diagrams adapted to yielding quantitative results. It is often advantageous to combine the two methods where some of the quantities used graphically are cumbersome and can be more easily computed.

Free-body diagram. A fundamental tool in solving problems is the free-body diagram. In this, each part of the mechanism is isolated and the forces acting upon each part are established. It is then often possible to piece this knowledge together until the entire solution is known. In **Fig. 1***a*, a four-bar linkage with frictionless pins and weightless members is acted upon by two forces *S* and *P* on members 3 and 4, respectively. The forces on all pins and the couple on member 2 are sought. Figure 1*b* shows the three moving members isolated in their respective free-body diagrams. Using these diagrams, if one starts with 4, one can successfully work backward for the

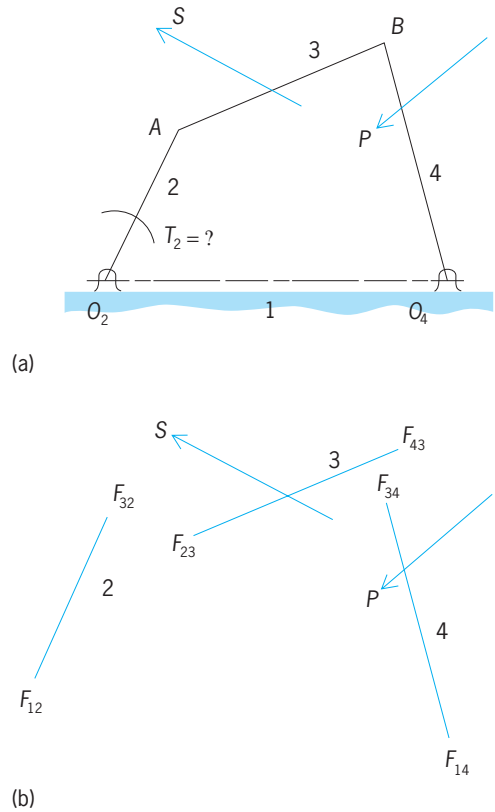

Fig. 1. Construction of free-body diagram. (*a*) Forces *S* and *P* acting on members 3 and 4 of a four-bar linkage. (*b*) Free-body diagrams of the three moving members of *a*.

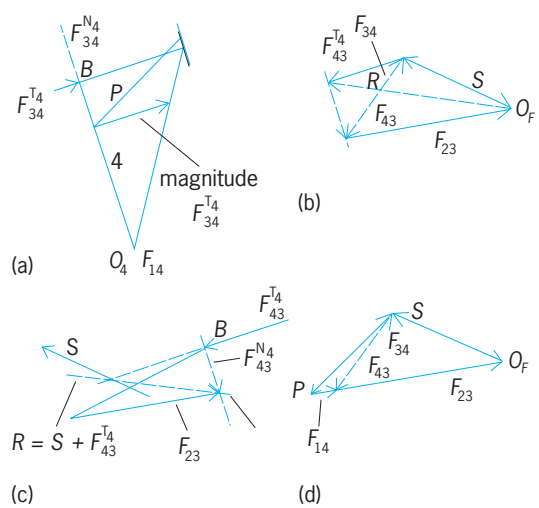

Fig. 2. Successive steps (*a*-*d*) in solution of forces.

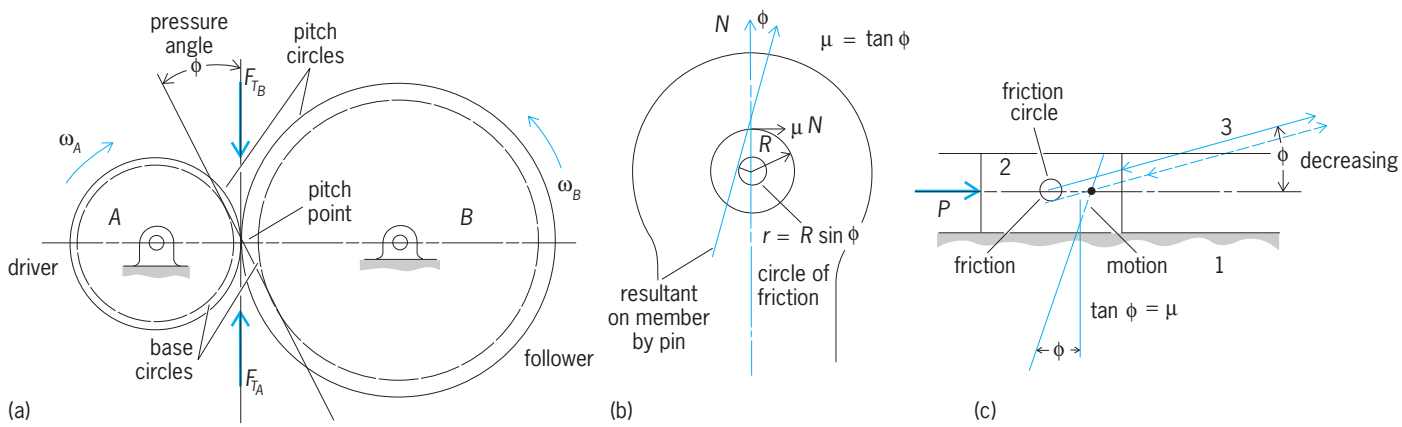

Fig. 3. Force diagrams. (*a*) Force between two gears. (*b*) Force on a pivot or pin of radius *R*. (*c*) Force on the upper or lower side of a slider.

entire solution. In **Fig. 2***a*, a moment diagram with fulcrum at O_4 permits solution for F_{34}^{T4}, the normal force component of member 3 on member 4 at pin *B*. In Fig. 2*b* (F_{34}^{T4}, *S*, and *R* are now known), F_{23} and F_{43}^{N4} can be found. In Fig. 2*c*, knowledge of F_{43}^{N4} yields the direction of *R*. Figure 2*d* thus reveals the direction and magnitude of F_{14} on pin O_4, as well as the pin force (equal and opposite to F_{23}), and hence the couple on 2. *See* FOUR-BAR LINKAGE.

Force diagram. Force diagrams for many friction and other devices common to mechanisms may be helpful. The force between two gears is one through the pitch point in the direction of the common tangent to the two base circles and inclined at the pressure angle to the common tangent to the pitch circles (**Fig. 3***a*). The force acting on a pivot or pin of radius *R* will act along the center line and through the center point unless there is friction in the pin (Fig. 3*b*), in which case the force will lie at the friction angle where the center line meets the moving head, at angle ϕ ($\tan\phi = \mu$). Hence it is tangent to a small circle, the friction circle, $r = R\sin\phi$ (Fig. 3*b*). The force on the upper or lower side of a slider will vary in direction with forward and return strokes, but will lie at the friction angle to the normal and through the point of concurrency of other forces (Fig. 3*c*). *See* GEAR.

In **Fig. 4**, a press mechanism with massless members but with pin and slider friction is treated. There is a known resistance force of *P* units at output slider

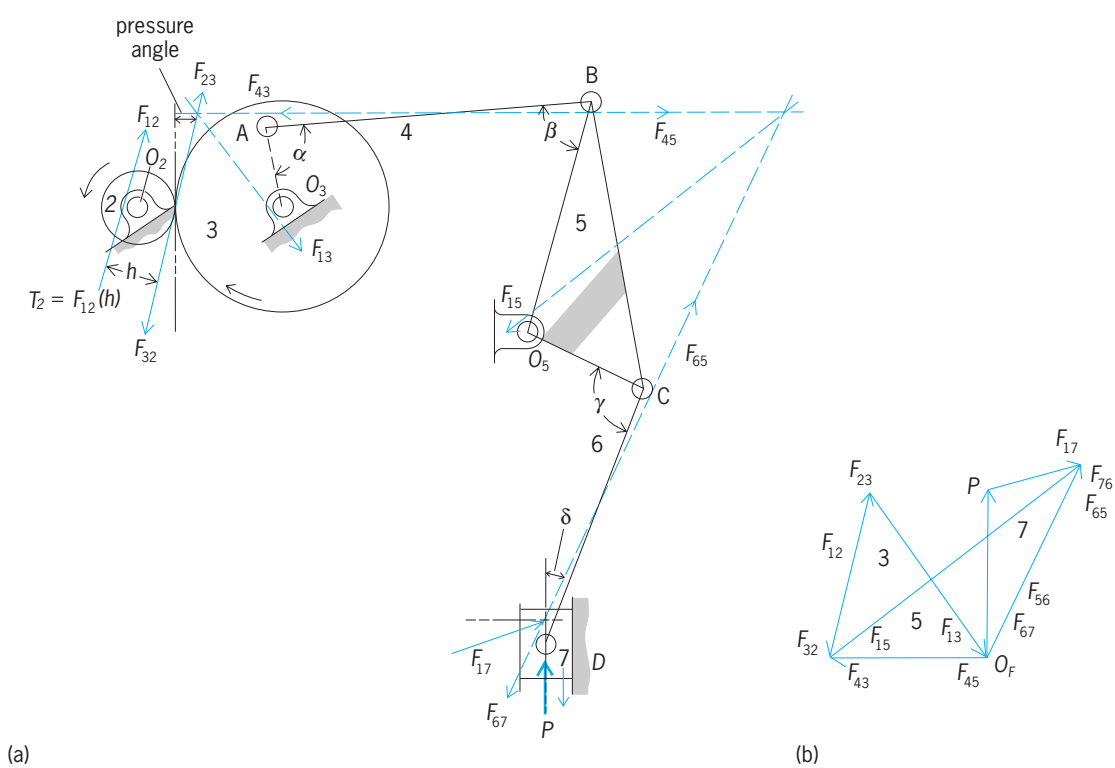

Fig. 4. Solution of (*a*) force problem by (*b*) force triangles.

block 7. The solution polygon showing all forces appears in Fig. 4*b*. All joint frictions have been assumed equal, with gear friction not being treated. The general method of solution in Fig. 4*a* observes that polygon triangles can be evaluated backward, starting with the three forces acting on 7. Respective part directions must first be ascertained followed by general force directions under impending motion. Next, in order to fix the placement and direction of friction forces, it is necessary to decide whether the angles are respectively increasing or decreasing. This can be done by inspection, by diagrams, or by computations. With the direction of each of the forces at the pins known, it can be determined upon which side of the pin the force must pass to induce a rotation opposite to that due to joint friction. Thus at point A, the geometry shows that angle α is expanding; hence friction must be acting clockwise (to oppose it), and hence F_{43} must act on the upper side of the friction circle to turn it counterclockwise, as shown. Corresponding patterns apply for angles β, γ, and δ. Thus at D, since δ is diminishing under the impending motion, pin motion is counterclockwise, pin friction is clockwise, and the force F_{67} must yield a counterclockwise effect. At D there is a backward, upward friction force on the left side of the slider, establishing the friction force.

Once the positions of the forces are known, as in Fig. 4*a*, three basic force triangles are successively solved; in each, the direction and magnitude of one force and the directions of two others are known. Thus with P known, F_{65} is ascertained for the triangle of forces involved with body 7. For the force triangle for body 5, F_{56} has just been found, and so F_{45} can be found in direction and magnitude. For the force triangle for body 3, F_{45} is known in position and magnitude. Hence, with knowledge of the directions of F_{32} and F_{13}, the complete triangle is known, and F_{12} can be ascertained in direction and magnitude.

When inertia forces and moments are involved, as occurs in moving parts due to their masses, the situation can be met by regarding the forces momentarily as static forces in equilibrium. By using such devices, the existing forces (such as compression forces or shear forces) and couples on mechanisms can frequently be found at points of interest. It is assumed that this mass is concentrated at the known center of gravity and that its linear and angular accelerations have been learned from kinematic studies. Kinematics is the inevitable background ingredient of every preparation for a force study where mass is present.

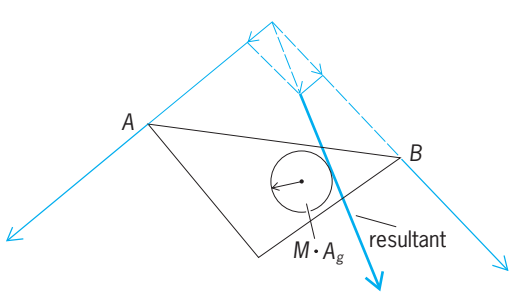

Fig. 5. Representation of the forces acting on a mass to cause the resultant force.

The force $F = MA_g$ is thus known in direction and magnitude, and Eq. (1) holds. Here I is the mass

$$T = I\alpha \tag{1}$$

moment of inertia (the inertia of rotation) and T is the turning force, torque, or inertia moment. Then Eq. (2), or Eq. (3), can be written. Here k is the radius of gyration.

$$MA_gH = I\alpha \tag{2}$$

$$h = \frac{I\alpha}{MA_g} = \frac{k^2\alpha}{A_g} \tag{3}$$

Equations (1)–(3) can also be shown in circle form (Fig. 4*a*), where the force acting upon the mass (to produce the known acceleration) is represented as the resultant of the other forces upon the piece (**Fig. 5**). *See* MOMENT OF INERTIA. Douglas P. Adams

Bibliography. E. A. Dijksman, *Motion Geometry of Mechanism*, 1976; H. H. Mabie and F. W. Ocvirk, *Mechanisms and Dynamics of Machinery*, 4th ed., 1987; B. Paul, *Kinematics and Dynamics of Planar Machinery*, 1979.

Force fit

A means for holding mating mechanical parts in fixed position relative to each other. In the case of a force fit of cylindrical parts, the inner member has a greater diameter than the hole of the outer member; that is, the metals of the two parts interfere. *See* ALLOWANCE.

Light drive fits require assembly pressure and produce a more or less permanent assembly of thin sections or long fits. They are suitable for cast iron external members. For a tighter fit, medium or heavy drive fits are used, although such fits may require more than direct mechanical pressure for assembly. In a true force fit, the parts are highly stressed, the interference amounting to 0.002 or 0.003 in. (0.0051 or 0.0076 cm) for parts with a basic diameter of 1 in. (2.5 cm). *See* SHRINK FIT. Paul H. Black

Forced oscillation

A response of a mechanical or electrical system in reaction to an external signal.

Electrical systems. A simple *RLC* circuit (**illus.** *a*) consists of a resistor with resistance R (measured in ohms), an inductor with inductance L (measured in henrys), and a capacitor with capacitance C (measured in farads). The dynamics relating the input voltage, $u(t)$, to the current, $y(t)$, passing through the resistor are described by Eq. (1). Equation (1)

$$L\frac{dy}{dt} + Ry + \frac{1}{C}\int y\,dt = u(t) \tag{1}$$

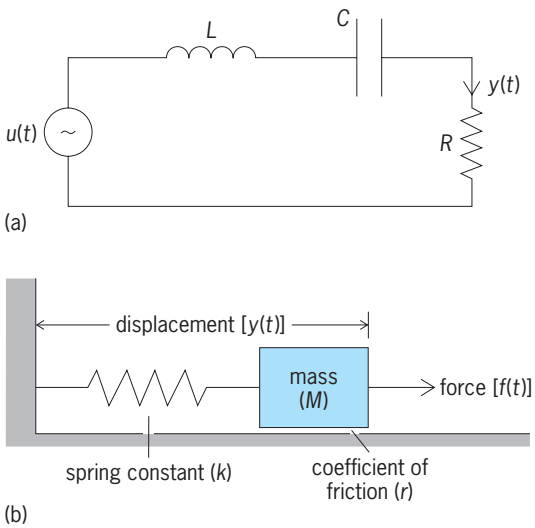

Examples of forced oscillation. (*a*) Electrical system, an *RLC* circuit. (*b*) Analogous mechanical system, a spring-mass-damper system.

states that the input voltage is equal to the sum of the voltage across the inductor, the voltage across the resistor, and the voltage across the capacitor, where the voltage across the inductor is the product of its inductance (L) and the rate of change of the current through the inductor; the voltage across the resistor is the product of its resistance (R) and the current passing through it; and the voltage across the capacitor is the integral over time of the current through the capacitor (that is, the charge on the capacitor plates) divided by the capacitance (C).

A fundamental property of differential equations states that the response of a differential equation to a periodic input can be decomposed as a sum of two responses. The first one, called the zero-input response or free oscillation, is due to initial energy stored in the circuit and decays eventually to zero. The second one, due to the voltage input $u(t)$, converges to a periodic signal with the same frequency as $u(t)$. The latter is referred to as the forced oscillation or the steady-state response. The decaying rate of the free oscillation depends on the time constant of the circuit which is determined by the values of R, L, and C and the structure of the circuit. *See* TIME CONSTANT.

Impedance. If the input voltage $u(t)$ is a sinusoidal function with magnitude A (measured in volts) and angular frequency ω (measured in rad/s), then the forced oscillation of $y(t)$ is also a sinusoidal function with the same angular frequency, whose magnitude (B) and phase (θ) with respect to $u(t)$ depend on a quantity called the complex impedance (Z). The complex impedance is a complex number whose real part is just the resistance (R) and whose imaginary part equals $\omega L - 1/(\omega C)$. The magnitude of the forced oscillation (B) is equal to the magnitude of the input voltage (A) divided by the modulus of the complex impedance ($|Z|$). The phase (θ) of the forced oscillation with respect to the input voltage is just the negative of the argument of the complex impedance. Since the complex impedance is a complex number, its modulus is the square root of the sum of the squares of its real and imaginary parts, and its argument is the inverse tangent of the ratio of its imaginary part to its real part. *See* ALTERNATING-CURRENT CIRCUIT THEORY; COMPLEX NUMBERS AND COMPLEX VARIABLES; ELECTRICAL IMPEDANCE; PHASE (PERIODIC PHENOMENA).

Damping. The free oscillation is independent of the input voltage $u(t)$. It can take the form either of an exponentially decaying function or of an exponentially decaying sinusoidal function, depending on the damping coefficient, $R/(2L)$, and the undamped resonant frequency, $1/\sqrt{LC}$. If the damping coefficient is greater than the undamped resonant frequency, then the free oscillation takes the form of an exponentially decaying function; and if the damping coefficient is less than the undamped resonant frequency, then the free oscillation assumes an exponentially decaying sinusoidal function. In contrast, the forced oscillation is independent of the damping coefficient and the undamped resonant frequency. It always has the same frequency as the input signal $u(t)$. *See* DAMPING.

Mechanical systems. Similarly, an analogous mechanical system, a simple spring-mass-damper system (illus. *b*), consists of a body with mass M, which is attached to a wall by a spring with spring constant k, and rests on a horizontal surface over which it moves with friction coefficient r. The dynamic equation that relates the force applied to the body, $f(t)$, to the body's displacement, $y(t)$, is given by Eq. (2).

$$M\frac{d^2y(t)}{dt^2} + r\frac{dy(t)}{dt} + ky(t) = f(t) \qquad (2)$$

Equation (2) states that the force applied to the body equals the sum of the three quantities: the product of the body's mass and its acceleration, the negative of the frictional force, and the negative of the force exerted by the spring. Here, the negative of the frictional force is the product of the coefficient of friction and the body's velocity, and the negative of the force exerted by the spring is the product of the spring constant and the body's displacement. Moreover, the body's velocity is the first derivative of its displacement with respect to time, and its acceleration is the second derivative of its displacement with respect to time.

Analogous to the *RLC* circuit case, application of a sinusoidal force $f(t)$ results eventually in a forced oscillation of the displacement $y(t)$ that is also a sinusoidal function. The magnitude and the phase of the displacement $y(t)$ depends on the complex mechanical impedance that is a function of the mass (M), the spring constant (k), and the friction coefficient (r). The exact evaluation is similar to the *RLC* circuit case. *See* MECHANICAL IMPEDANCE; MECHANICAL VIBRATION; VIBRATION.

Resonance. In both examples, the magnitude of the forced oscillation depends on the complex impedance Z. The imaginary part of Z will be zero if ωL equals $1/(\omega C)$. The value of the angular frequency

ω that satisfies this equation is given by Eq. (3). At

$$\omega_o = \frac{1}{\sqrt{LC}} \tag{3}$$

this value of ω, the impedance becomes a real number and the maximum magnitude of the oscillation occurs. The frequency ω_o is called the resonant frequency, and the system at this frequency is said to be in resonance. *See* RESONANCE (ACOUSTICS AND MECHANICS); RESONANCE (ALTERNATING-CURRENT CIRCUITS).

Quality factor *Q*. A closely related concept is the quality factor Q. In the context of the RLC circuit, Q is defined by Eq. (4). It can be shown that the ratio of

$$Q = \frac{\omega_o L}{R} = \frac{1}{\omega_o RC} \tag{4}$$

the magnitude of the output voltage $y(t) \cdot R$ to that of $u(t)$ is given by Eq. (5). This equation shows that

$$\frac{BR}{A} = \frac{1}{\sqrt{1 + Q^2\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right)^2}} \tag{5}$$

if Q is large, the ratio is small unless ω is close to ω_o. In other words, high Q corresponds to narrow bandwidth and low Q corresponds to wide bandwidth.

The quality factor Q also has a different interpretation, which can be explored via an energy analysis. In fact, Q is 2π times the ratio of the maximum energy stored in the system to the energy dissipated per cycle. There is a continuous exchange of energy between the magnetic field of the inductor and the electric field of the capacitor. During each half cycle, the energy stored in the capacitor's electric field varies from a maximum value to zero and back to the maximum value again. The energy stored in the inductor's magnetic field varies in a similar manner. The way in which energy exchange takes place is as follows: During one-quarter cycle the capacitor takes in energy as it is released by the inductor, and during the following one-quarter cycle the inductor absorbs energy as the capacitor discharges it. The total energy stored in the circuit is a constant, though the energy stored in either capacitor or inductor varies with time. Similarly, in the mass-spring system, there is a continuous exchange between the kinetic energy of the mass and the potential energy of the spring, but the total energy of the system is constant. *See* HARMONIC MOTION; HARMONIC OSCILLATOR; OSCILLATION.

Er-Wei Bai

Bibliography. R. C. Hibbeler, *Engineering Mechanics*: *Dynamics*, 8th ed., 1997; J. D. Irwin and R. M. Nelms, *Basic Engineering Circuit Analysis*, 8th ed., 2005.

Forensic anthropology

The application of physical anthropology theory and techniques to answering questions for the law. Most human remains are soon discovered and thus available for analysis. Medical evidence and biological techniques used on the living can be applied to these intact remains, as can extensive internal examinations and histological preparations. However, burned, decomposed, mutilated, or fragmented bodies contain less soft tissue evidence and are therefore more difficult for the medical examiner or forensic pathologist to analyze. In these cases, the information provided by the skeleton becomes more important, and may yield the only information available about the identity of the individual and the circumstances surrounding death. Age, sex, ancestry, stature, muscularity, handedness, habits, occupational activities, disease, and injuries can be ascertained from elements of the skeleton, and in such cases the services of a forensic anthropologist may be a valuable adjunct to the work of the forensic pathologist or coroner.

Anthropology is traditionally divided into three major areas: cultural anthropology, which deals with the social study of human cultures; archeology, which deals with past human societies; and physical (or biological) anthropology, which concentrates on humans and their primate relatives. Physical anthropology includes the variation found in the anatomy of the human body and how the body has evolved through time, how it varies in existing populations around the world, how it adapts to different environments, and how it changes during life. Not all physical anthropologists are specialists in the human skeleton, and in fact, even those who are human osteologists may not be experienced in all aspects of skeletal identification and analysis. Training is required in such areas as skeletal variation in current populations or injuries from modern weapons. Many physical anthropologists, however, are increasingly involved in the forensic applications of their expertise, a few on a full-time basis. Recently, archeologists have been asked to aid in the excavation of human remains from clandestine graves, and so the field of forensic archeology has arisen, but the term "forensic anthropology" usually refers to the biological aspects. *See* ANTHROPOLOGY; ARCHEOLOGY; PHYSICAL ANTHROPOLOGY.

In the field. Qualified forensic anthropologists must be actively involved at the scene of the incident to locate, document, map, and recover scattered, buried, decomposed, burned, or dismembered remains. Accurate documentation of the location and condition of remains is important in the identification of the remains and in the reconstruction of the incident. Whether in death investigations involving one body or in a mass fatality incident, the investigator must always assume that litigation will occur and must be prepared to document every move in preparation for a trial. Anthropologists with extensive training in archeological location and documentation techniques (sometimes with forensic archeologists) are extremely valuable at a crime or other recovery scene, and often discover and document materials and context (the relative location of materials to each other within a scene) overlooked by untrained law enforcement personnel. Recovery of

crime scene materials should always be performed with a crime scene technician, however.

Human versus nonhuman. This determination is usually simple for the forensic anthropologist but is critical. An investigation can begin or end with this decision. A wide variety of nonhuman remains have been mistakenly identified as human, including burned rabbits, bear paws and other bones, goats, turtle shells (mistaken as skull fragments), and chicken bones.

Identification of scattered remains. The ability to recognize human remains as they are found can assist in many immediate decisions concerning the location and nature of additional remains, and aid in the identification of the circumstances surrounding death. Since the more than 200 bones from an adult skeleton may be scattered over several acres, it is necessary for the forensic anthropologist to determine what bones have been found and what portions are still missing. The relative locations of various skeletal elements and associated evidence give clues as to the site of initial deposition of the body and the agents responsible for scattering the remains.

Recovery of buried remains. It is certainly possible to train crime scene investigators or other personnel to recognize intact human bones and to conduct proper search and excavation of human remains. Still, the forensic anthropologist may be able to offer much better documentation of the exact location of objects in the burial environment, as well as to better interpret the date and shape of the initial excavation at the time of the burial of the body. Many anthropologists or archeologists have the highly developed skills of excavation that can even determine the tool used to dig the original grave. In complex situations, a forensic anthropologist, archeologist, and trained crime scene team are recommended.

Burned bodies and mass fatality incidents. Scenes such as airplane crashes and building explosions and fires can be complicated and may overwhelm the local authorities and facilities. Forensic anthropologists are increasingly members of teams of identification specialists (such as odontologists, pathologists, fingerprint experts, and radiologists) called to help identify victims of mass fatality incidents. Recognition of bones becomes more difficult in these instances, because when burning has destroyed the soft tissue, the bones may be black, gray, or even white. Also, burned bone is extremely fragile and can disintegrate with little movement, and the burning process breaks, warps, and shrinks bones, sometimes up to 20%.

In the morgue or laboratory. It is important to involve forensic anthropologists, odontologists, pathologists, and radiologists as a team working together to identify remains. If the anthropologist is allowed to see only those remains unidentified by the other team members, it will be more difficult to reassociate body parts to identified bodies. An initial examination of the remains is ideally conducted jointly by the forensic pathologist and forensic anthropologist. As sharp instruments are used by the forensic pathologist in the autopsy process, the anthropologist is often present to document this additional damage to the remains (to avoid confusion in the analysis of perimortem trauma later). At this point, evidence (such as insect remains or degree and type of decomposition) for the determination of time since death is collected or observed before forensic examination causes further alteration of the remains. Maceration is the removal of all soft tissue from the bones. Usually the forensic anthropology laboratory will macerate the remains if necessary.

Minimum number of individuals. When incomplete skeletons are found, the number of individuals represented must be determined with care by checking duplication of skeletal elements, symmetry between sides of paired bones, and the size of bony articulations. Commingling is always considered a possibility.

Determination of sex. The sex of the remains can be determined by the morphology of many skeletal elements, particularly the pelvis. The expert can determine sex from more or less complete skeletons with a high degree of reliability, except in the case of subadult skeletons. Overall size and markings of muscle attachments are useful in almost all bones. Measurements and statistical tests can also be used to support conclusions.

Determination of age. Age can be determined from the human skeleton by a great many indicators. Which indicators are appropriate depends upon the general age of the deceased and the condition of the remains. A very young individual normally shows a wealth of maturation information associated with growth and development, including dental development and eruption, and appearance of centers of bone growth. Differences in skeletal age and dental age are frequently observed, particularly in females. The age estimates are usually precise within a small range. During the middle to late teen years, growth nears completion and age estimates may have more variability. A few areas of skeletal maturation often persist until the beginning years of the fourth decade (30 to 32). Skeletal remains of the mature adult still exhibit changes related to age, but most are probably degenerative in nature. Age estimates based upon maturation are usually much more precise than those derived from degenerative changes.

Determination of ancestry. The cranial and facial features are the best biological indicators of ancestry. These may be demonstrated by statistical techniques and metric data from the cranium. However, determination of ancestry is one of the most difficult diagnoses to make in forensic anthropology, in part because there is more variation within a race than between races. In fact, one cannot take one or two traits and assign every skeleton in the world to one of the commonly defined races. Moreover, while the forensic anthropologist is diagnosing ancestry by using biological features, an individual is classified (or classifies himself or herself) to an ancestral or ethnic group based on social as well as biological features.

Stature estimation. Measurements of individual long bones are the best source for stature estimates.

Formulas specific for males and females and various ancestries are available for each of the six major long bones. When very complete skeletons are available, individual measurements of all bones that contribute to stature can yield good estimates with several techniques. Measurements of incomplete long bones can also yield stature estimates, but they are obviously less precise.

Idiosyncrasies. Occasionally, the skeleton shows unusual variants, such as a sternal foramen, nonclosed posterior arches of the sacrum, bifurcated ribs, or cervical or lumbar ribs. Very often, asymmetry of the face, limbs, or jaws will be present. These may aid in the identification process.

Disease, trauma, and time of death. Many diseases and injuries suffered by the individual during life will leave evidence in the skeletal system. These include arthritis, bone infections (including tuberculosis, syphilis, and osteomyelitis), cancers, developmental disorders, healed fractures, orthopedic surgical procedures, or surgically implanted devices that might be embedded in the bones or surrounding soft tissue. The active diseases present in the victim around the time of death, or injuries received at the time, are always important considerations in any death investigation. Conclusions about the cause and manner of death must be made by the pathologist, or in some jurisdictions, the coroner; but it is the role of the forensic anthropologist to describe and interpret any evidence present in the skeletal remains that may be useful for the pathologist to reach informed conclusions. Without clinical experience, the anthropologist normally would not offer testimony or opinions concerning amount of pain to the victim, how long consciousness would have been maintained, or whether an injury would have been lethal. Some cases may be exceptional, such as a shotgun wound resulting in massive destruction of the cranial vault. Likewise, injuries to the now absent soft tissue may be inferred by the damage to the bones. For instance, the wounds which caused cut marks from the sharp edge of a weapon on the anterior surfaces of the bodies of the cervical vertebra would have likely damaged the vascular structures in the neck.

Forensic anthropologists can often interpret postmortem damage to the remains and use this information to help diagnose the elapsed time following death by taking into consideration the environmental conditions (including climate, depth of deposition, and coverings over the body) and the degree of decomposition, skeletonization, scavenging, and weathering of the remains. Very fresh bodies can yield precise estimates of the time of death, but as time passes, less precision can be achieved. Estimates of time since death for decomposed remains are broad, but they reflect the variety of reactions that the remains can have to environmental conditions. Narrow estimates will produce more difficulties in the investigation or prosecution than broad estimates.

Facial approximation. Publication of photographs of three-dimensional sculpting over the skull and two-dimensional drawings based on photographs of the skull is a way of seeking information from the public about the identity of victims. This technique is a last attempt to identify a victim, and is used when the investigators do not know of any possible victims, or when all known possible victims have proven to be incorrect. The names proposed by the public as a result of the reconstructions can then be checked by the investigators. If one appears to be correct after further investigation, then one of the above techniques should be used to confirm the identification. Facial reconstructions do not confirm identifications, but they are means of seeking possible victims. Many forensic anthropologists do not do facial reconstructions.

Photo superimposition. This relatively new technological application is used to compare skulls with antemortem photographs of possible victims. The technique usually employs two video cameras, one imaging the skull and the other the photograph of the person when alive. A modified version of this means of narrowing the identities of individuals involves the use of photographs taken at automated teller machines, trap camera, or bank cameras compared with photographs taken of living suspects or victims.

Positive identification. Final identification of the remains as those of a particular person may be made by a number of means. The most readily accepted means of positive identification ("to the exclusion of all others") include dental identification (comparing the dental records before and after death), fingerprint analysis, and DNA. Forensic anthropologists often aid in the collection of evidence for these means of identification, but most forensic anthropologists are not qualified to make the final identifications. If enough individualizing characteristics are discovered in skeletal elements (through radiographic or gross comparison of information before and after death), positive identification may be made in court from those alone. *See* ANTHROPOMETRY; FORENSIC CHEMISTRY; FORENSIC MEDICINE. Diane L. France

Forensic biology

The scientific analysis of biological evidence to provide objective information on legal matters or those that pertain to criminal and civil law. Biological evidence such as bodily fluids or tissues that may be found at crime scenes can be analyzed through deoxyribonucleic acid (DNA) typing. Typing requires detection and screening of the biological evidence, such as blood, semen, or saliva, extracting the DNA from a specimen, amplifying specific regions of the DNA using the polymerase chain reaction (PCR), and typing the resulting PCR products to determine a DNA profile. The DNA profile from the evidence is then compared to known profiles from suspects, victims, or database samples to determine the significance of the result. Samples containing mixtures require additional interpretation to infer individual donor allele designations. Forensic biologists must also assess the statistical significance of their results,

write reports, and testify in court. *See* CRIMINALISTICS; DEOXYRIBONUCLEIC ACID (DNA); FORENSIC EVIDENCE.

The establishment of a United States national DNA database, the Combined DNA Index System or CODIS, has facilitated the ability to compare DNA profiles from unknown biological crime scene evidence to DNA databases of known convicted criminals ("cold hits") or to DNA left at other crime scenes, resulting in the ability to link cases. As of 2005, CODIS contained over 2.5 million profiles. Many other countries, such as the United Kingdom and Germany, have DNA databases with some of the same genetic markers, permitting international database searches. In the United States alone, over 25,000 investigations have been aided through CODIS.

By comparison of the DNA profile from crime scene samples to known samples, the results can serve to link victims and suspects with the crime scene or can exclude a suspect from association with that crime. Additionally, scientific analysis of biological evidence may provide unbiased information to substantiate case circumstances, corroborate or refute an alibi, or identify a weapon used in a crime. Cases may include nonhuman samples such as botanical, fungal, entomological, or zoological specimens, which can also be used to link victims and suspects with each other or to the crime scene. Forensic biology may also be applied to homeland security, in which crimes may include weapons of mass destruction such as pathological microbes. The forensic biologist may be able not only to identify the microbe (such as anthrax) but to type the microbial DNA and link the strain of the microbe to a strain produced at a certain location or to the original progenitor strain. *See* FORENSIC BOTANY; FORENSIC ENTOMOLOGY; MICROBIOLOGY; PATHOLOGY.

Two main principles permit the use of DNA in forensics. First, no two individuals have the same DNA, with the exception of identical twins. Second, the DNA from any source (such as blood, hair, or skin) of a particular individual will be the same.

Evidence detection and screening. Detection starts with evaluating the investigative information to understand the nature of the case and samples. Visual examination of stains with alternative light sources (such as ultraviolet or infrared light and lasers) and chemical enhancement reagents may be performed. Forensic biologists can determine the nature of the biological stain (for example, blood, semen, or saliva), and whether it is human through presumptive and confirmatory tests. These tests consist of analytical procedures including microscopy, chemical tests, and immunological assays. *See* FORENSIC CHEMISTRY; FORENSIC MICROSCOPY.

DNA extraction. Following the detection and screening of samples from a crime scene, DNA must be extracted. Among the methods used are the following:

1. *Organic extraction*. This method consists of lysis of the cells in a detergent-based buffer followed by one or more rounds of purification using an organic phase separation and concentration using column centrifugation or ethanol precipitation.

2. *Chelex resin extraction*. This method utilizes a fast, simple extraction of small amounts of sample in the presence of a chelating resin. It results in a somewhat crude extract but is usually adequate for PCR amplification of the forensic genetic loci.

3. *Solid-phase extraction*. These methods utilize a membrane that act as a capture device for the DNA. Samples are spotted onto the membranes, and the subsequent washes remove the impurities.

4. *Silica-based extraction*. Nucleic acids are first adsorbed to the silica in the presence of chaotropic salts (salts that weaken hydrophobic associations). Polysaccharides and proteins do not adsorb and are removed. Next, following a wash in low salt, pure nucleic acids are released. This method has been automated using robotic stations and is being used in several laboratories. *See* EXTRACTION.

DNA quantification. Assessing the quantity and quality of the sample is the next step. Several methods are being utilized in crime laboratories. These include agarose gel electrophoresis in the presence of quantification standards (samples with known quantities of DNA), known as yield gel electrophoresis; slot blot hybridization, using known DNA standards immobilized on a membrane followed by hybridization to a human/higher primate–specific DNA probe; homogeneous plate assays, using a DNA fluorescent dye and scanning in a plate reader; and more recently, real-time detection using quantitative PCR (QPCR). Real-time QPCR has several advantages over the other methods in that it is extremely accurate and sensitive over a broad dynamic range, and it occurs in a closed-tube system, reducing the potential for carryover contamination. Using this technique, a forensic biologist can monitor and quantify the accumulation of PCR products during log phase amplification.

DNA amplification. PCR is a fast in-vitro DNA synthesis process that can provide up to a billion (10^9) copies of a given target sequence. Specific DNA markers can be targeted for duplication by a DNA polymerase. Primers are designed to hybridize to the specific markers along the length of the DNA template during the cycling of temperatures. In the thermal cycle, DNA strands are separated, primers bind to the template, and then a special DNA polymerase that is heat-stable is used to copy and amplify the genetic markers. Through a process of 28–32 heating and cooling cycles, the DNA is increased so that it can be analyzed. The thermal cyclers contain many sample wells, permitting the amplification of multiple samples simultaneously; as many as 96 samples may be amplified in under 3 hours. *See* FORENSIC MEDICINE; GENE; GENE AMPLIFICATION; MOLECULAR BIOLOGY.

Multiplex PCR allows several different loci to be simultaneously amplified in a single tube. This enables typing from a single aliquot of the extracted genomic DNA, reducing sample consumption. Recently the ability to analyze as many as 15 autosomal short tandem repeats (STRs) simultaneously has

been reported using DNA from a very small amount of degraded sample.

DNA separation and detection. Separation and detection of the amplified products is required following PCR. There are many different methods to achieve typing. These include polyacrylamide gel electrophoresis (PAGE) followed by silver staining or, if the primers are fluorescently tagged, detection by fluorescent gel scanners; and capillary electrophoresis (CE) with laser-induced fluorescence. The latter method has become the most commonly utilized method of detection, as it is highly automated (there is no gel to pour and load), samples can be easily be reinjected (robotically), and since the DNA traverses the entire length of the capillary, the resolution of the higher-molecular-weight loci is usually better than in the PAGE methods. *See* ELECTROPHORESIS.

Forensic genetic markers. The genetic markers currently being typed in most forensic biology laboratories include autosomal short tandem repeats (STRs), mitochondrial DNA, and Y chromosome STRs. *See* GENETIC MAPPING.

Short tandem repeats. Short tandem repeats consist of regions of 2–7 base pairs that are repeated consecutively. Individuals may vary in the number of repeats and/or the content of the repeats. The variation in the content of the repeats occurs in either a change in the base within a repeat unit or as a deletion in the repeat unit. STRs used in forensics are either tetranucleotide or pentanucleotide repeats. STRs are highly abundant and well studied in the human genome, and their small size and the alleles' small size range facilitate typing from highly degraded, small quantities of starting material. There are 13 CODIS core loci that are being uploaded into the national DNA database.

Mitochondrial DNA. Mitochondrial DNA (mtDNA) is useful for forensic DNA in that it exists in high copy numbers in each cell, and therefore has a better chance of being detected in small samples. Moreover, mtDNA is maternally inherited, so any individual within the maternal lineage may provide an mtDNA reference sample. Finally, since the size of the amplicons (the DNA segments to be generated by PCR) is small, mtDNA can be typed from degraded DNA. Mitochondrial DNA hypervariable regions I and II are the most commonly sequenced targets in forensic DNA laboratories. The most commonly used method of mtDNA typing is fluorescent Sanger's dideoxy sequencing.

Y chromosome markers. Y chromosome markers, including Y STRs, have been recently used in many forensic DNA crime laboratories. The interest in Y chromosome markers is well supported for the following reasons:

1. The total number of male cells that are present at a rape scene may be very small in the case of rapists who are azoospermic (having no sperm) or oligospermic (having a low sperm count).

2. The total number of male cells may be low due to loss of sample or degradation.

3. Multiple semen donors may need to be identified in a multiple rape case.

4. In criminal paternity or mass disaster victim identification, determination of the haplotype of a missing individual may be conducted by typing a male relative.

5. In sexual assaults, the time-consuming and sometimes inefficient differential extraction procedure for the separation of sperm and nonsperm fractions may be bypassed.

6. Y chromosome typing may provide increased statistical discrimination in mixture or kinship analysis cases in which the discrimination obtained from autosomal markers is insufficient for identification purposes.

Interpretation. Once amplified and typed, the results need to be interpreted. Forensic biologists must have a clear understanding of the molecular methods utilized, with in-depth knowledge of the basis of typing (that is, the cell biology, technology, and genetics of the loci), the specific loci and amplification parameters (determined both externally and internally in developmental and internal validation studies), the empirically derived limitations of the system, the protocols and quality control measures implemented in their own laboratories, the instrument validations, the analytical software utilized in determining the DNA profiles, and the statistical and population genetics databases and software used.

In single-source samples, the interpretation requires setting a threshold of detection. That is, alleles that are at or above a certain fluorescent threshold may be designated. In highly degraded DNA samples, the interpretation requires more analysis, as higher-molecular-weight loci generally degrade more rapidly than lower-molecular-weight loci. In this case, the alleles are expected to appear in decreasing amounts as the size of the alleles increases (see **illustration**). In mixtures, the interpretation requires considering all combinations of alleles that are present, allele ratios within and among loci, and the sample type and condition.

Future technologies. Many technologies applicable to forensic biology are currently under investigation, including the following:

Hand-held, microcapillary STR typing. Development of hand-held DNA typing devices has obvious advantages for forensics in that samples might be processed quickly at the crime scene. The potential to rapidly determine the DNA type of samples left at a crime scene coupled to the growing national DNA database provides a powerful tool to law enforcement, as "cold hits" might provide leads early in the investigation. Devices have already been developed that can detect microbial pathogens and can be used out in the field. Moreover, under development are miniaturized microcapillary devices for typing human STRs.

Single nucleotide polymorphisms. Single nucleotide polymorphisms (SNPs) are single base sites that vary between individuals and as such can be used in forensic DNA typing. They have already been utilized in forensic DNA tests. The PCR-based DNA typing systems HLA-DQ Alpha and Polymarker loci were based on SNPs. The detection was by reverse dot blot

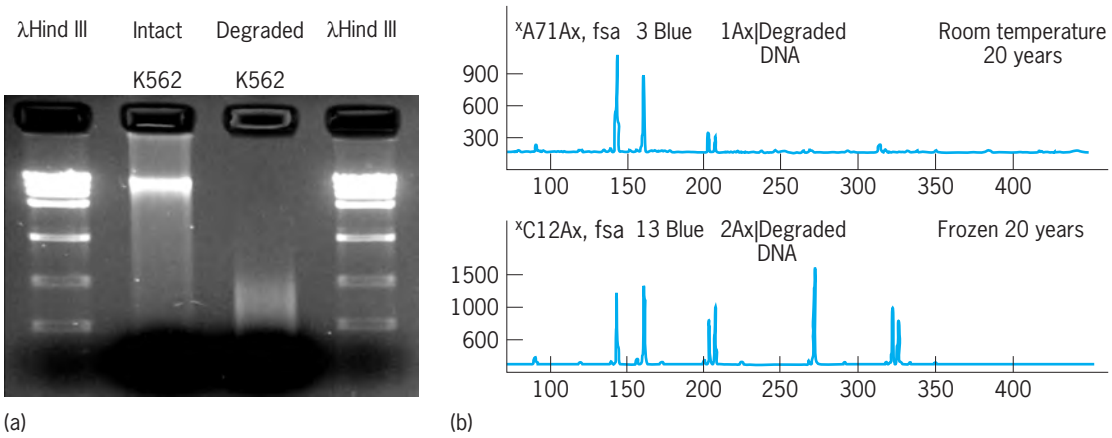

Results of degrading DNA by mechanical disruption. (*a*) Agarose gel electropherogram of intact DNA and degraded DNA, electrophoresed in lanes 2 and 3 through a 1% agarose gel, stained with ethidium bromide, and then photographed with UV light. Molecular-weight ladders of Lambda Hind III were placed on either side of the samples in lanes 1 and 4. The sample in lane 3 has been highly degraded with no detectable high-molecular-weight band as is apparent in lane 2. (*b*) Capillary electropherogram of degraded DNA stored at room temperature versus frozen. The decrease in PCR product amount as base pair size increases, as shown in the top panel, is the expected result for a degraded sample. The result demonstrates the importance of proper (frozen) sample storage.

hybridization, where the known samples were immobilized on the blots. Today there is interest in the detection of mtDNA and Y SNP typing using a variety of approaches, including primer extension assays, taq man assays, microarrays, liquid bead arrays (Luminex), and pyrosequencing.

Forensic biometrics. The ability to determine the physical characteristics of an individual by typing genetic markers has been called forensic biometrics. Inferring population of origin from DNA evidence using Y chromosome SNPs has been recently reported. The ability to determine the potential genetic origin of a perpetrator has obvious benefits to law enforcement. However, it also has obvious ethical and legal implications, as well as significant limitations, in that the genetic boundaries between races are not clear. Other applications may be in determining the age of a suspect using genetic markers.

Tissue typing using mRNA profiles. Different tissues have different genetic expression patterns. Recently the use of ribonucleic acid (RNA) has been reported for body fluid identification. The potential use of molecular technology was also reported in determining the age of a bloodstain, which could be useful in establishing the time of the crime, using analysis of messenger RNA (mRNA):ribosomal RNA (rRNA) ratios. Advantages of the mRNA-based approach, versus conventional biochemical tests, include greater specificity, simultaneous and semiautomatic analysis, rapid detection, decreased sample consumption, and compatibility with DNA extraction methodologies.

Low-copy-number amplification (LCN). Biological evidence is often found with an extremely low number of starting templates (1–15 diploid cells, <100 picograms). Analysis of LCN samples requires the development of novel strategies for its collection, analysis, and interpretation. Also known as touch sample amplification, low-copy-number amplification may require additional cycles of amplification and/or additional cleaning or sampling, but permits the analysis of extremely small amounts of sample such as might be found on the grip of a gun.

Steven Lee

Bibliography. B. Budowle et al., Forensics and mitochondrial DNA: Applications, debates, and foundations, *Annu. Rev. Genomics Hum. Genet.*, 4:119–141, 2003; B. Budowle et al., Public health: Building microbial forensics as a response to bioterrorism, *Science*, 301(5641):1852–1853, 2003; J. Butler, *Forensic DNA Typing: Biology and Technology Behind STR Markers*, 2d ed., Academic Press, San Diego, 2005; A. Carracedo and P. Sanchez-Diaz, Forensic DNA typing technologies: A review, in A. Carracedo (ed.), *Forensic DNA Typing Protocols: Methods in Molecular Biology*, pp. 1–297, Humana Press, Totowa, NJ, 2004; M. A. Jobling and P. Gill, Encoded evidence: DNA in forensic analysis, *Nat. Rev.*, 5:739–751, 2004; E. L. Jones, Jr., The identification of semen and other body fluids, in *Forensic Science Handbook, volume II*, 2d ed., pp. 329–399, Pearson Education, New Jersey, 2005; N. Rudin and K. Inman, *Forensic DNA Analysis*, 2d ed., CRC Press, Boca Raton, FL, 2001; J. A. Siegel, G. C. Knupfer, and P. J. Saukko (eds.), *Encyclopedia of Forensic Sciences*, vols. 1–3, Academic Press, 2000.

Forensic botany

The application of plant science for purposes of civil or criminal law. Forensic botany has opened aspects of the field of botany to novel avenues within criminal investigations. Today, several subdisciplines of plant science are being applied successfully in criminal investigations—for example, plant systematics including identification of plant species, plant anatomy, and plant ecology. Often, forensic applications come from combinations of these areas in a given case.

Plant systematics. Plant systematics is the study of evolutionary relationships among plants. Certain knowledge in this field is of considerable use in criminalistics and crime scene work. Being able to attach the proper scientific name to a plant is important in several situations. These items should be gathered carefully and notations made concerning exact location of the plant materials. Collection of plant material evidence is especially important in crime scene investigations, but often is overlooked or collected in an improper manner. The identification of plant materials can involve whole plants, plant fragments, and seeds. There are guides to plant identification called floras. These books are most useful when whole plants or whole flowers are to be identified. In a given setting, the more local the plant identification guide (flora), the better. For example, if dealing with a crime scene at or near Yellowstone National Park, the best source would be the Park's flora. *See* CRIMINALISTICS; FORENSIC EVIDENCE; PLANT TAXONOMY.

In some cases, plant fragments may be the only material available, for which there may be published guides. There are several picture guides to plant seeds. These are more difficult to use than the floras because distinctions among seeds can be very subtle. There also are guides for leaf fragments of trees and branch tips from woody shrubs and trees. With seeds and plant fragments, botanical expertise from a local college or university may be needed.

Applications of plant systematics. Sometimes homicide victims may be found in places far from the murder site. In one case, a blanket-wrapped body was found in rural grassland with assorted plants attached to the victim's clothes and the blanket. The plant fragments could not have come from that grassland because they were from ornamental plants associated with formal gardens. The person had been killed in the city and the corpse transported many miles into the country. The ornamental plants were in agreement with those found in a suspect's yard, linking the victim to the suspect, and perhaps to the crime scene.

In another case, a kidnapping victim was grabbed from the street and released somewhere in the mountains above Boulder, Colorado. The investigators suspected useful evidence might be found at the release site. A van used in the crime had been recovered. The victim identified the van, and her DNA was found in it. As investigators inspected the vehicle, they collected plant fragments and seeds from the seats and floor, the floor pads and pedals, the tire treads, the outside window wells, and the undercarriage of the van. The plant samples were collected with forceps and placed in coin envelopes. Once the plants pieces were assembled and identified, the release site was pinpointed and further evidence was found on site. Investigators are urged to collect plant material with forceps rather than with the special vacuum sweepers found in many departments, because vacuuming shatters dry plant fragments and complicates the identification process.

Plant anatomy. Plant anatomy is the study of plant cells. Applications of plant anatomy in forensic science have been based in large part on a unique characteristic of plants—their cellulose cell walls. Cellulose is a complex carbohydrate that is highly resistant to ordinary decomposition processes. For example, wood is made up of cellulose that has undergone a process known as lignification. Plant cells can pass through the human digestive tract with their walls intact, maintaining the same size and shape they possessed before ingestion. Because the human diet in any culture is made up of a relatively few kinds of plant foods, seldom more than 75 different kinds, and because each food plant has cells of distinctive sizes, shapes, cell arrangements, and cell inclusions such as crystals, it is possible to identify the plants consumed in someone's last meal (see **illustration**). This use of plant anatomy has been employed especially to postmortem the stomach contents. *See* PLANT ANATOMY.

A further use of stomach contents can be an estimation of time of death because the valve allowing stomach contents to proceed to the small intestine

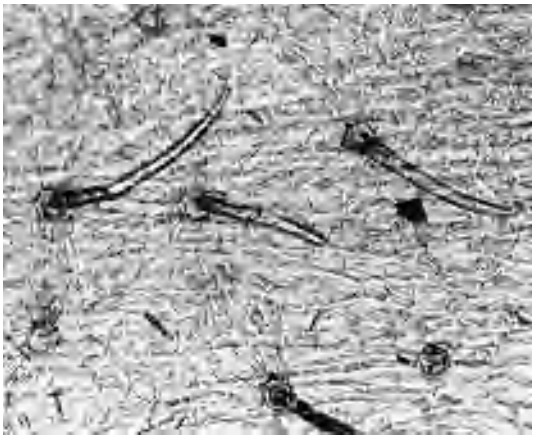

(a)

(b)

Surface of a green bean pod, showing its spines using (*a*) light microscope (*courtesy of David O. Norris*) and (*b*) scanning electron microscope (*courtesy of Meredith Lane*).

closes at death. If the individual was healthy and had eaten an average-sized meal, the material would remain in the stomach for only 1–3 h under ordinary circumstances and then pass to the intestines.

Fecal analysis shows identifiable plant cells as well, a technique used with great success by anthropologists in their studies of ancient cultures.

People are sometimes stabbed with wooden instruments. One of the properties of cellulose is that it is birefringent (glows) under light microscopy. This allows determination that the stabbing was carried out with a wooden instrument, and may be used as well to link the instrument with a specific wound.

Plant anatomy applications. Sometimes the contents of the last meal can play a role in criminal investigations. In a "black widow" case, the suspect claimed to have fed the victim his last meal (supper), giving her an alibi for the estimated time of his murder. But analysis of his stomach contents showed his last meal differed greatly from her menu. He was observed earlier in the day eating his true last meal (breakfast) at a restaurant. This occurred at a time for which the "black widow" had no alibi.

In another case, the vegetables that had been selected for a pizza order helped the delivery person recall the individuals in the room where the delivery was made. Several persons in the room were killed, and the investigators thought the perpetrator might also have been present when the pizza arrived, so the deliverer's recall was significant.

Plant ecology applications. Plant ecology is the study of the relationships between plants and their physical and biological environments. The body of a missing person was discovered along a ditch in midsummer in the Great Plains wheat country. An attempt had been made to hide the corpse by covering it with giant sunflowers growing along the roadside. The victim had been reported missing only a few days before the body was discovered, but the crime scene investigators suspected the person had been dead for a longer time due to the state of decay and evidence from forensic entomology. The suspect had left the area about 10 days before the body was discovered and well before the missing person's report was filed. The sunflowers used for covering the body were wilted. In order to determine how many days it took for an uprooted sunflower to reach this particular wilted condition, several living sunflowers were collected from the site where the body was found, exposed to the environmental conditions at the site, and observed daily until the same level of wilting occurred. The appropriate level of wilting took 8 days and was maintained at that state up to 15 days, providing further evidence that the homicide had occurred before the person was reported missing and before the suspect had left the area. *See* ECOLOGY.

A federal agency suspected a kidnapping had occurred based on a single photograph it had received. This photo was circulated widely at forensic science meetings and elsewhere. This photo was all the evidence the agency had that a crime had occurred, and workers were anxious to locate the crime scene. It was concluded that the picture had been taken in the upper montane ecological zone of the southern Rocky Mountains, in either northern Colorado or southern Wyoming, based upon the Engelmann spruce trees and logs, the metamorphic rocks, and certain shrubs shown in the photo. But it was not possible to provide a more specific location.

NecroSearch International Ltd. is a nonprofit organization made up of people from many disciplines that have forensic relevance. Their purpose is to locate clandestine graves. Sequential changes in plant species known as plant succession patterns are distinct over the remains of buried bodies and often remain distinct from surrounding ground cover for many years. Evidence of plant succession patterns has been used in locating such graves. *See* ECOLOGICAL SUCCESSION.

Summary. Forensic botany is a tool in criminal investigations. Its techniques are simple for the most part and are readily accepted in court testimony. For much of the work, qualified consultants can be found at local colleges and universities among the biology faculties trained in plant science.

Jane H. Bock; David O. Norris

Bibliography. J. H. Bock and D. O. Norris, Forensic botany: An under-utilized resource, *J. Forensic Sci.*, 42(1):364–367, 1997; J. H. Bock and D. O. Norris (eds.), *Handbook of Forensic Botany*, Humana Press, 2004; M. J. Crawley (ed.), *Plant Ecology*, 3d ed., Blackwell Science, Oxford, England, 1997; W. C. Dickenson, *Integrative Plant Anatomy*, Academic Press, London, 2000; S. Jackson, *No Stone Unturned*, Kensington Publication Corp., Boston, 2002; D. O. Norris and J. H. Bock, Method for examination of fecal material from a crime scene using plant fragments, *J. Forensic Invest.*, 52(4):367–377, 2001; J. S. Sachs, *Corpse: Nature, Forensics, and the Struggle To Pinpoint Time of Death*, Perseus Publishing, Cambridge, MA, 2001.

Forensic chemistry

The application of chemistry to the study of physical materials or theoretical problems, the results of which may be entered into court as technical evidence. Forensic chemistry is part of the broader field of forensic science. The boundaries of forensic chemistry are not sharply defined, and topics that are not entirely chemical in nature are included. The scope of the analyses performed depends on the criminal, civil, and regulatory statutes of a given jurisdiction. A forensic chemist may be asked to examine debris from a fire scene for the presence of accelerant residues, to determine the species of an imported animal product for ad valorem tax classification, to demonstrate the presence of a narcotic in a pill, or to determine if a certain tool made a scratch on an illegally entered door.

A forensic laboratory may be part of a governmental agency or may be within private industry. The extent of its service varies considerably. Full-service laboratories analyze a wide range of evidentiary items.

Regional laboratories usually screen cases, performing basic examinations, while the more complex tests are done in a central laboratory. Specialty laboratories, such as a medical examiner's lab or a regulatory lab, have a narrower range of services covering a few specified areas. The crime laboratory is the best known of the three types.

Evidence types. Every material in existence could become evidence at some time. Some of the items most often encountered in crime laboratories, and the information sought, are (1) body fluids and viscera to be analyzed for poisons, drugs, or alcohol, the quantitation of which may assist in determining the dosage taken or the person's behavior prior to death; (2) licit and illicit pills, vegetable matter, and pipe residues for the presence of controlled substances; (3) blood, saliva, and seminal stains, usually in dried form, to be checked for species, type, and genetic data (DNA); (4) hairs, to establish animal or human origin (and if human, the race, body area of origin, and general characteristics); (5) fibers, to determine type (animal, vegetable, mineral, or synthetic), composition, dyes used, and processing marks; (6) liquor, for alcoholic proof, trace alcohols, sugars, colorants, and other signs of adulteration; (7) paint, glass, plastics, and metals, usually in millimeter-sized chips, to classify and compare to known materials; (8) inks on documents, to determine type, dye content, or possible age; also chemical obliterations and restoration of charred documents; (9) damaged floppy disks or hard drives, for recovery of erased computer files; (10) debris from a fire or explosion scene, for the remains of the accelerant or explosive used; (11) materials seized from a clandestine laboratory, to determine what illicit drugs were being made; (12) broken headlight filament, to determine whether it was "on" or "off" at the time of an accident.

Regulatory labs receive any product that has defined legal specifications. These may include chemical composition, physical performance, or the absence of hazardous chemicals. Private laboratories act as consultants for either side in a legal contest by providing general and specific chemical knowledge.

Legal aspects. Forensic chemists are concerned with more than just the technical aspects of an analysis. They are also involved with the handling of evidence before and after it enters the laboratory. An item considered to be evidence goes through five stages: recognition, collection, preservation, analysis, and testimony. At each stage there are legal requirements that must be met if the item is to be judged admissible in court, and chemical precautions which must be observed if it is to be analyzed properly. One of the legal requirements is that the chain of evidence, that is, the origin, transfer, possession, and analysis of each item, must be recorded as it occurs. All packages and seals should remain intact. Detailed notes are necessary.

At some crime scenes the chemist, with the advantage of specialized knowledge, is aware of the potential use of certain materials involved and can recognize, recover, and preserve evidence so that its chemical usefulness is not ruined. Communication and cooperation between field assistants and laboratory analysts allow the problems to be solved rapidly and ensure that the evidence is received in good condition.

For court testimony, a forensic chemist must be able to explain complex instrumental techniques to a nonspecialist jury in an understandable manner, and must have a thorough knowledge of each instrument and chemical property in order to withstand cross-examination.

Purposes of analyses. Before beginning an analysis, the forensic chemist must understand what the investigator or prosecutor or defense is seeking from it. To prove possession of a controlled substance, the chemist must identify a compound according to the definition in the relevant statute. If a drug statute specifies the difference between possession of cocaine base (crack) and cocaine as a salt (coke), then the analytical procedure must be able to identify each species separately. In general this is referred to as establishing an element of a crime. Another purpose that rules the analytical process is proving or disproving a statement made by a suspect, victim, or witness. Most of these analyses are of the "either-or" type. If a red car is struck during a hit-and-run incident and the police locate a car with red smears, the question is, "Could the red smears on the suspect's vehicle have come from the victim's car?" If the suspect should offer an alibi that the red paint came from the trim of his garage door, then a second question arises, "Can the laboratory show this statement to be true or false?"

Databases. Sometimes the laboratory is asked for investigative help: "What can the physical evidence tell an investigator that will help him find a suspect?" The answer may be a simple description of the evidence such as that the suspect vehicle has a metallic green top coat and shows damage to the right front headlight and grille area. In other cases, searchable databases are consulted to identify individuals or to link crimes.

In the United States, the National DNA Index System (NDIS) contains DNA profiles from convicted offenders. All states collect DNA samples from convicted sex offenders; some states have broadened their collections to include all convicted felons. The Federal Bureau of Investigation (FBI) manages the national index, which is a combination of all the state data. DNA from blood, semen, saliva, skin cells, and so on can be profiled and compared to the national index to identify a suspect. In addition, serial crimes can be linked by a common profile. The agencies involved can share their case information to help identify a suspect.

The Integrated Automated Fingerprint Identification System (IAFIS) allows laboratories to search a single latent print against state and national (FBI) collections of 10-print cards obtained during arrest procedures. *See* FINGERPRINT.

Drugfire and IBIS (Integrated Ballistics Identification System) are systems that allow for searching of breech markings on cartridge casings and striations

on bullets. They can link shooting incidents even without access to the weapon itself.

Theory of analysis. Forensic examinations are designed for one of two purposes: identification or comparison. In the former case, a particular chemical species, such as blood alcohol level, the amount of a drug in a tissue, or the amount of barium and antimony in gunshot residues, is identified and often quantitated. DNA analyses by short tandem repeats (STR) have now reached a level of discrimination such that people can be uniquely identified (except identical twins).

Comparative examinations determine consistencies or inconsistencies between the chemical and physical properties of two samples in an attempt to show or exclude a common origin. Many of the materials encountered in a forensic laboratory are manufactured in large batches that have a homogeneous nature. A thousand gallons of paint may be used on many cars. A sample from any of these cars will be consistent with any other sample. If a car is involved in an accident, the paint can be matched to it, but not to the exclusion of the other cars. Inconsistencies are positive, since they exclude the possibility of a common origin.

Thus, for each item of evidence, an analytical scheme can be devised which leads to identification or comparison. An example of identification involves drug analysis, where it is necessary to identify the controlled substance and quantitate it. A small portion of a pill is submitted to a series of color tests. The results indicate the class of drug presence (barbiturate, alkaloid, and so on), but not which member of the class. The drug is then extracted from the excipients in the pill, and individual members of the class are separated by a chromatographic method. Confirmation of the identity is typically obtained from infrared spectroscopy and quantitation by gas chromatography. For purposes of comparison, the paint from a suspect car is compared to paint from a hit-and-run victim. The color, and number and relative thickness of the layers are studied; the chemical behavior toward solvents and spot-test reagents is observed; the inorganic pigments are identified and the polymer film characterized. If both paints have similar results in all tests, they are consistent; but the report must state that another source could exist.

In the analytical scheme, modification may be made as needed. When the sample is small, nondestructive tests are done first. These tests include microscopic examination, index of refraction, density, x-ray diffraction, x-ray fluorescent elemental analysis, and in some cases, infrared spectroscopy.

Equipment and instrumentation. A modern forensic laboratory has a vast array of sophisticated instruments. Each has its own purpose, and is used for several different evidence types. One of the simplest tests is the spot test, which gives a color change, a precipitate, gas evolution, or a crystal when it reacts with a particular class of chemicals. Spot tests are not unique, but allow for the rapid screening of many samples. Other simple tests are solubility in water and organic solvents, density, index of refraction, melting and boiling point, and flash point.

Microscopy. Probably the most frequently used piece of equipment is the microscope. Low-power stereomicroscopes are used to examine the surfaces of opaque materials such as paints, soils, fibers, and paper. Examples of the uses of optical transmission microscopes include the examination of the internal structure of hairs, measurement of the index of refraction of small chips of glass, checking for spermatozoa in assault cases, and the study of tissue samples in pathology laboratories. Optical transmission microscopes may also have accessories for luminescent studies, dispersion staining, and the measurement of optical properties with polarized light. Comparison microscopes allow the chemist to view two samples side by side on a split screen. Metallographic microscopes are used to study fractured metal parts in liability investigations. *See* MICROSCOPE; OPTICAL MICROSCOPE.

The scanning electron microscope (SEM) with an elemental analyzer also has found forensic applications. Magnifications up to 100,000 times can be achieved with a greater depth of field than in optical microscopes. Minute marks on fired casings, scratches on tumblers, and the composition of very small paint layers have been compared using the scanning electron microscope. Research has shown that minute particles (less than 0.001 mm) of a certain shape and elemental composition are unique to gunshot residue, and the scanning electron microscope can be used to examine material from the hands of individuals for such residues. *See* SCANNING ELECTRON MICROSCOPE.

Spectroscopy. A second common group of instruments are in the field of spectroscopy. The infrared spectrum of a pure compound is a chemical fingerprint. This method is used to positively identify a substance, and can also classify materials present in mixtures. Fourier transform infrared (FTIR) instruments have supplanted dispersion instruments. Ultraviolet-visible spectrophotometry is a screening technique that characterizes materials that possess color.

Elemental analyses are done by using either a graphite furnace atomic absorption spectrophotometer (AAS) or ion coupled plasma (ICP) emission spectrograph, each capable of determining the major or trace elements in a sample, with ultimate detection levels in the parts per billion being common. Applications include the identification of metallic poisons such as lead and arsenic, trace pollutants, and gunshot residue.

The mass spectrometer, used in drug analysis since it can positively identify a volatile chemical substance, is often coupled to a gas chromatograph as a specific detector. Toxicology labs rely upon these instruments for rapid screening at low levels of occurrence. Special methods have been developed for the identification of explosives. Trace chemists use mass spectrometry to identify petroleum products used as accelerants and as a detector on an ICP emission spectrometer. *See* FORENSIC TOXICOLOGY; MASS SPECTROMETRY; SPECTROSCOPY; TOXICOLOGY.

Chromatography. Since most evidence is a complex mixture containing both relevant and irrelevant materials, forensic chemists depend upon separation techniques. The classical methods of extraction, distillation, and precipitation are sometimes used, but they have been replaced to a large degree by chromatographic methods. Chromatography not only separates but also quantitates at the same time.

The use of specific detectors has broadened the use of these methods. Thin-layer chromatography is a rapid screening technique for drugs, explosives, poisons, and ink dyes, while gas chromatography is both a screening and quantitating technique. These methods are applied to drugs, arson accelerants, pesticides, and hazardous chemicals. Liquid chromatography, which is a relatively new forensic tool, has the potential to replace many thin-layer and gas chromatography procedures, since fewer preparations of the sample are required, more versatile detectors can be used, and nonvolatile samples can be examined. Drugs, inks, explosives, and carcinogenic materials have been detected by liquid chromatography. *See* CHROMATOGRAPHY.

Electrophoresis. Electrophoresis separates large molecules such as proteins and enzymes. Forensic serologists can study these species in dried blood stains and compare them to standard samples from a suspect. Most crime laboratories have stopped performing conventional blood group analyses and have switched to DNA analyses. One procedure uses capillary electrophoresis to separate alleles. Since the frequency with which each allele occurs in a population is known, the probability of two persons having the same DNA profile can be calculated. *See* ELECTROPHORESIS.

X-ray techniques. These techniques are widely used in forensic laboratories. X-ray diffraction can positively identify any crystalline material, even in mixtures. Paint pigments can be run without separation from the film, since the film is noncrystalline and does not diffract the x-rays. Other uses are for soils, minerals, household chemicals, and drugs. X-ray fluorescence techniques can be used to determine the amounts of elements heavier than sodium, simultaneously, in 100 seconds. Such a rapid method has found acceptance in many laboratories. *See* X-RAY DIFFRACTION; X-RAY FLUORESCENCE ANALYSIS.

In addition to the above instrumental techniques, many less common physical and chemical examinations are conducted by forensic laboratories. It is incumbent upon the analyst to choose those which will satisfy the legal requirements for a specific case. *See* ANALYTICAL CHEMISTRY; CLINICAL PATHOLOGY; FORENSIC MEDICINE.

Michael J. Camp

Forensic engineering

Engineering applied toward the purposes of law. In forensic engineering, the majority of investigations are carried out in the context of civil litigation; for example, the cause of a plane crash is often due to defective design, which is a civil litigation consideration. However, the crash could be due to a bomb or terrorist activity in the cockpit, which is a criminal consideration. Often the cause of an accident is initially unknown or misattributed. *See* CRIMINALISTICS.

Accident reconstruction. Accident reconstruction is the a posteriori scientific process of using the facts of an accident and the appropriate natural laws of science to determine the possible circumstances and causes, consistent with the available data. The quality and reliability of the reconstruction depend on the amount and quality of the available data, as well as the skill and knowledge of the investigator.

Consider, for example, a two-vehicle, right-angle, intersection collision (see **illustration**). The two vehicles had proceeded through an intersection, and then an impact occurred. Each vehicle had skidded prior to impact, and the final positions of the vehicles are indicated relative to the point of impact. For an analysis, the input data required (distances and angles as shown along with vehicle weights) would have to be obtained from a field investigation and from databases of vehicle properties. The principle of conservation of linear momentum, which follows from a straightforward application of Newton's laws, where all calculation details are given, enables the engineer to calculate the speeds of both vehicles at impact (v_1^i, v_2^i) and at the onset of preimpact skidding (v_1, v_2), based solely on the data (independent of driver or witness statements). Depending on the actual input data and the calculated speeds, the results may be used in a variety of ways. Driver 2, for example, may claim that he stopped at the stop sign, and the accident occurred because driver 1 was speeding through the intersection. Based on the input data for the case illustrated, this claim cannot be true: the accident occurred because driver 2 did not stop at the stop sign, and driver 1 was not speeding. *See* NEWTON'S LAWS OF MOTION.

Product liability. In a product liability investigation, engineers are asked to determine whether a product is defective and whether the defect was causally related to any injuries that may have occurred. This is a challenging area of investigation and analysis. The three types of product defects recognized by law are (1) design defects (the product lacks those elements necessary for its safe and foreseeable uses), (2) manufacturing defects (the product was not made according to the manufacturer's specifications), and (3) failure to warn or instruct (the product was not accompanied by adequate warnings or instructions for its proper and safe use). In a lawsuit, if a forensic engineer has been retained by the plaintiff (the person bringing the lawsuit), the engineer's role is to objectively analyze the product for defects, which may or may not be found, and to causally relate the plaintiff's injuries to the product defect. If a forensic engineer has been retained by the defendant, the engineer's role will be to objectively defend the product, if it can be done, in view of the defect allegations made by the plaintiff. As part of

the defense, the engineer may opine that the alleged defect does not exist or that the injuries sustained by the plaintiff are not causally related to the product design. Product liability investigations can be very intense and can involve considerable analysis and calculations, testing, and computer modeling, often requiring the expenditure of large sums of money by both the plaintiff and defendant.

Design defects. As a simple but meaningful example of a design defect, consider cigarette lighters. Children are naturally inquisitive and may pick up lighters and try to light the flame. Some lighters have bright colors and shapes that make them particularly attractive to children. If the lighter does not have a child-proof or -resistant mechanism, which prevents children from lighting the flame, the lighter is defectively designed. The plaintiff's engineer in this case would show that at the time of manufacture it was practical and feasible to have provided a child-proof mechanism, and that the device may have even been in existence on other lighters, which would have prevented the accident.

Manufacturing defects. The fact that a person is injured using a product does not necessarily mean the product is defective. In addition, defective products do not cause injuries every time they are used. For example, a knife is designed to have a sharp edge in order to cut. If the user is cut, it does not necessarily follow that the knife is defective. However, suppose that a person is using a knife to carve a turkey and the blade fractures, causing the person to lose balance, to fall forward, and to be stabbed by the fractured blade. After the accident, a metallurgical analysis of the knife blade reveals that the steel was not properly heat-treated and that there were inclusions in the blade, which weakened it and caused the failure; this determination indicates that the knife was defectively manufactured.

Failure to warn. Car air bags can be life-saving devices if properly designed and used. When an air bag deploys in an accident, its front can move at up to 200 mi/h (320 km/h). A person sitting too close to the air bag can be seriously injured or killed by the impact to the head or upper torso. Small adults and children are particularly susceptible, as well as infants in rearward-facing child seats. Automobiles carry prominent warnings that occupants should not sit too close to air bags, that small children should sit in the rear, and that rearward-facing child seats should never be placed in the front seat of an air bag–equipped vehicle. Without such a warning, the vehicle would be defective by virtue of a failure to warn. The plaintiff's biomechanical engineer in this case would show by tests and/or calculations, or even using the automobile manufacturer's own data, that a person or child sitting too close to the deploying air bag would be subjected to injurious levels of impact forces.

Biomechanics of injury and death. Engineering science and methodology is being applied to injury and death investigation, a subject rooted in human biology. Biomechanical engineers working in the field of injury and death investigation are generally interested in applying engineering principles to the understanding of injury causation mechanisms in accidents, and in developing the criteria and conditions for when injury is likely to occur. For an accident, the kinematics of movement of the human body must be calculated and then correlated with any impacts and injuries that the body may have sustained. In an automobile accident, the movement of an occupant with respect to the vehicle (occupant kinematics) is often correlated with the acceleration-time history of the occupant, the occupant's change in velocity (delta-v), and the resulting injuries. This is a critical step if the occupant alleges an injury because of the accident or a vehicle defect. A very active research field in biomechanics is the determination of criteria that delineate the conditions in which injury may occur in a particular accident. *See* BIOMECHANICS.

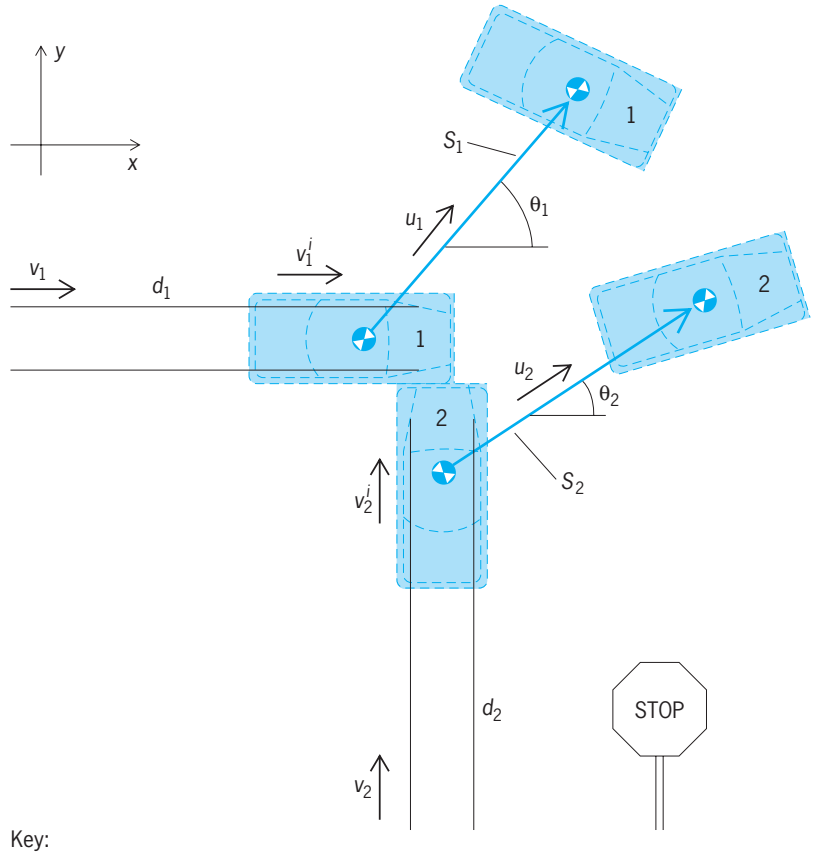

Key:

d = preimpact skidding distance
S = postimpact movement
v = velocity at beginning of skid
v^i = impact velocities
u = postimpact velocities
θ = angles as shown

Intersection collision.

Computer forensics. This rapidly changing area of forensic engineering refers to methods for solving crimes committed using computer technology. Such crimes include extortion, violations of Internet security, hacking into supposedly secure Web sites and computers, computer theft of sensitive information and proprietary files (including national security

data and financial data), and creation and propagation of computer viruses. Computer forensics uses techniques and procedures for recovering magnetic data, tracking hackers, and analyzing audio tapes, videotapes, and photographs for possible tampering and alteration. *See* DIGITAL EVIDENCE; INTERNET.

Cause and origin. This area refers to investigations to determine the cause and origin of fires and explosions, and to determine if the cause was accidental (electrical shorting, product defect, cooking mishap, and so on) or intentional (arson or planting an explosive device). An understanding of heat transfer, burn patterns, combustion rates, pressure-wave propagation, and analysis of explosive residues is essential for cause and origin investigators.

Structural collapse, blast loading. This area refers to the catastrophic failure of structures either during construction or after the structure is in service. For example, the Hyatt Regency Skywalk collapse in Kansas City, Missouri, on July 17, 1981, was caused by a simple design error and resulted in the loss of 114 lives. The collapse was caused by an improperly designed connection, which did not meet Building Codes, where the vertical hanging rods passed through the box beams that supported the pedestrian walkways. Forensic civil engineers are also concerned with understanding the effect of blast and seismic loading on structures in order to ensure that future structural designs are safer and more resistant to dynamic loading arising as a result of explosions and earthquakes. *See* EARTHQUAKE ENGINEERING.

Significant structural lessons were learned from the bombing of the Murrah Federal Building in Oklahoma City on July 19, 1995, and its resulting collapse; these lessons (outlined in the Proceedings of the 1st Congress of the American Society of Civil Engineers in 1997) will be used to mitigate blast loading effects on future reinforced concrete building designs. Forensic engineers are becoming more actively involved in developing counterterrorism tactics and procedures, and this activity will likely grow as a separate area of forensic engineering. *See* STRUCTURAL DESIGN.

Steven C. Batterman; Scott D. Batterman

Bibliography. S. C. Batterman, *Education of Forensic Engineers*, 20(6), American Academy of Forensic Sciences, 1990; S. C. Batterman and S. D. Batterman, Accident Investigation/Motor Vehicle (accident reconstruction and biomechanics of injuries), in J. A. Siegel et al. (eds.), *Encyclopedia of Forensic Sciences*, Academic Press, 2000; S. C. Batterman and S. D. Batterman, Forensic engineering, chap. 12 in Y. H. Caplan and R. S. Frank (eds.), *Medicolegal Death Investigation: Treatises in the Forensic Sciences*, 2d ed., Forensic Science Foundation Press, 1999; F. P. Beer and E. R. Johnston, *Vector Mechanics for Engineers: Statics and Dynamics*, McGraw-Hill, 1977; D. T. Greenwood, *Classical Dynamics*, McGraw-Hill, 1977; K. L. Rens (ed.), *Forensic Engineering: Proceedings of the 1st Congress*, American Society of Civil Engineers, 1997; H. Yeh and J. I. Abrams, *Principles of Mechanics of Solids and Fluids: Particle and Rigid Body Mechanics*, McGraw-Hill, 1960.

Forensic entomology

The application of insect evidences or damage for purposes of civil or criminal law. Forensic entomology includes three subdisciplines: stored product entomology, structural entomology, and medicolegal entomology. Stored product entomology deals with damage to stored products, such as food materials or clothing, caused by insect activities. Structural entomology deals with damage to buildings and other structures by pest species such as termites or carpenter ants. Both these areas typically involve civil actions when there is a need to assess monetary damages resulting from insect activity. Medicolegal entomology deals with insects as evidence in criminal events, most frequently homicides. This is the area most commonly recognized by entomologists, law enforcement, and the public as forensic entomology. The application of entomological evidence to criminal investigations is not new, with the first recorded use coming from twelfth-century China. Beginning in the early 1980s, there has been a resurgence of interest in the field, and new applications of insect evidence are being exploited. *See* CRIMINALISTICS.

Insect evidence. Insects can serve as evidence in criminal investigations in several different ways. They can provide data for estimating the time since death (postmortem interval), detecting possible movement of the body following death and circumstances of the crime scene, and assessing antemortem versus postmortem wounds on the body. In addition, insect evidence can provide alternative specimens for toxicological analyses in highly decomposed bodies, yield DNA material linking a suspect to a victim for ectoparasitic taxa, and help assess periods of abuse or neglect of children and the elderly. Of these, the most frequent application is the estimation of the postmortem interval. *See* DEATH.

A dead body provides a changing and ephemeral food source for a wide variety of organisms, ranging from bacteria to vertebrate scavengers. Among these organisms, insects are major factors because they arrive at decomposing remains in a predictable pattern and develop in known patterns. Insects arrive shortly following death, often within minutes. The first arrivals are most frequently flies in the families Calliphoridae (blowflies and bottle flies) and Sarcophagidae (flesh flies). The female flies arrive and begin to investigate the natural body openings of the head, anus, and genitals. Wounds present on the body provide another area for activity. The female flies take a meal of blood or other exudates from the body and then deposit their eggs or, in the case of the Sarcophagidae, larvae into these body openings. This action starts a biological clock which is stopped and interpreted by the forensic entomologist when the body is discovered and the insects are collected. The use of insects to estimate the postmortem interval requires an understanding of the insect's life cycle, the relationship of the insect to the remains, and the relationship of the remains to the habitat in which they are discovered.

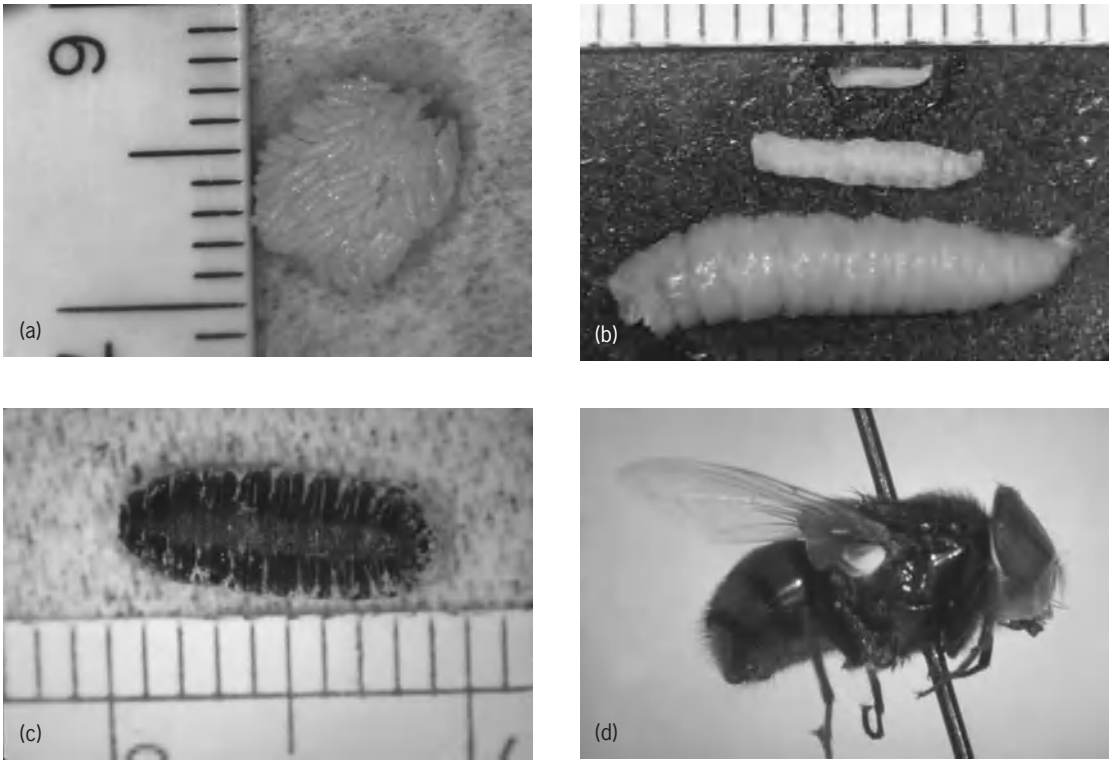

Chrysomya megacephala **(family Calliphoridae). (*a*) Egg mass. (*b*) Larval instars: first (top), second (middle), and third (bottom). (*c*) Puparium. (*d*) Adult female.**

Insects pass through a number of distinct life-cycle stages. Using the blowfly in the family Calliphoridae as an example, the female fly will arrive at the body and deposit eggs (**illus.** *a*) in the natural body openings or in wounds. These eggs will hatch into the larvae (maggots) which will feed on the decomposing tissues. There are three distinct larval stages (illus. *b*), called instars, with a molt between each stage. Once the maggot has reached complete development, it will cease feeding and move away from the remains to enter the pupal stage (illus. *c*), an inactive stage during which the larval tissues are reorganized to produce the adult fly (illus. *d*).

The insects encountered on a corpse in any given habitat will consist of some species unique to that particular habitat and some species having a wider distribution. The former group may be restricted to a particular geographic area or a particular habitat type within a given geographic area. For example, in Hawaii there are taxa that are restricted to a rainforest habitat, while others are specific to a more arid tropical habitat. Taxa having wider distributions are frequently encountered in several different habitat types and are typically highly mobile species, such as those taxa that are found both in rainforests and on arid beaches. Many of those taxa closely tied to carrion show this wider pattern of distribution. In estimating the postmortem interval, taxa from both groups may, under given circumstances, provide essential information concerning the history of the corpse. *See* INSECTA.

Species analysis. Of those insect taxa having a direct relationship to the corpse, there are four generally recognized relationships:

1. Necrophagous species actually feed on the corpse. Included are many of the Diptera (Calliphoridae and Sarcophagidae) and Coleoptera (families Silphidae and Dermestidae). Species in this group may be the most significant isolatable taxa for use in postmortem interval estimates during the earlier stages of decomposition, defined here as days 1–14. *See* DIPTERA.

2. Parasites and predators of necrophagous species constitute the second most significant group of carrion-frequenting taxa. Many of the Coleoptera (Silphidae, Staphylinidae, and Histeridae), Diptera (Calliphoridae and Stratiomyidae), and Hymenoptera parasites of Diptera larvae and puparia are included. In some instances, Diptera larvae, which are necrophages during the early portions of their development, become predators during their later larval development. *See* COLEOPTERA.

3. Omnivorous species include taxa such as ants, wasps, and some beetles, which feed on both the corpse and associated arthropods. Large populations of these may severely retard the rate of carcass removal by depleting populations of necrophagous species. *See* ANT; ARTHROPODA; WASP.

4. Adventive species include those taxa which use the corpse as an extension of their own natural habitat, as in the case of the Collembola, spiders, and centipedes. Acari in the families Acaridae, Lardoglyphidae, and Winterschmidtiidae, which feed on molds and fungi growing on the corpse, may be included in this category. Of less certain association are the various Gamasida and Actinedida, including the

Macrochelidae, Parasitidae, Parholaspidae, Cheyletidae, and Raphignathidae, which feed on other acarine groups and nematodes. In these investigations, what is being estimated is actually a period of insect activity rather than an actual time since death.

Interpretation. During the early stages of decomposition, the estimate of the postmortem interval is most frequently based on the development rates of individual species of Diptera, most frequently the flies in the family Calliphoridae. The most mature specimens collected from the body are preserved and identified. Given the particular species and stage of development represented, the entomologist can determine the period of time required to reach that stage. As the insects are dependent on environmental conditions, particularly temperature, for their rate of development, this must be factored into the estimate. Data obtained from detailed laboratory studies of life cycles conducted under controlled conditions can be correlated with conditions at a crime scene using weather data from National Oceanic and Atmospheric Administration (NOAA) stations and the concept of accumulated degree hours or accumulated degree days, a technique originally developed to predict pest outbreaks in agriculture. This calculated time period will represent the minimum postmortem interval. It is important to note that the time period is the period of insect activity and not the actual postmortem interval. Under most conditions, these will be quite similar, but there may be factors that delay the onset of insect activity, such as wrapping or concealment of the body.

After the first 2–3 weeks, depending on the location and environmental conditions, those flies in the initial wave of invaders will have completed their development and departed the body. At this point, the estimation of the postmortem interval is based on the ecological succession of insects onto the body. Those insects initially colonizing the body change the body by their activities. These changes make the body attractive to another group of insects, which feed and change the nature of the body, thus making it attractive to yet another group of insects. This process continues until the resource of the body has been completely exhausted. By making complete collections of the insects and other organisms present on a decomposing body and comparing these taxa with results of detailed decomposition studies conducted in similar habitats and areas, periods of time during which the death most probably occurred can be determined. Care must be taken in applications of results from decomposition studies as even geographically proximate localities may support different arthropod populations. These differences in species composition and relative abundance may serve to alter the successional picture. Generally speaking, the closer to the time of death, the more accurate the estimated postmortem interval will be. The estimate begins in terms of hours and proceeds to days, months, and finally seasons of the year.

Entomological evidence is not present in all cases and, when present, may not ultimately prove to be of major significance. When insect evidence is present and properly interpreted, it may prove to be a powerful tool for the solution of the case. M. Lee Goff

Bibliography. J. H. Byrd and J. L. Castner (eds.), *Forensic Entomology: The Utility of Arthropods in Legal Investigations*, CRC Press, Boca Raton, 2000; E. P. Catts and N. H. Haskell (eds.), *Entomology and Death: A Procedural Manual*, Joyces Print Shop, Clemson, SC, 1990; M. L. Goff, Estimation of postmortem interval using arthropod development and successional patterns, *Forensic Sci. Rev.*, 5:81–94, 1993; F. Introna and C. P. Campobasso, *Entomolgia Forense: II ruolo dei ditteri nelle indagini medico legali*, Essebiemme Editore, Italy, 1998.

Forensic evidence

Scientific, technical, or other specialized knowledge applied in courts of justice. Courts are finding that some areas of knowledge that have been considered forensic sciences for decades are not based completely on scientific knowledge under contemporary legal and academic standards, but rather are based on technical or other specialized knowledge that must be evaluated under standards not clearly defined. Indeed, some of the most interesting issues facing forensic evidence arise in cases where the defining characteristics of "scientific," "technical," and "other" specialized knowledge are at issue.

Federal law in the United States now requires trial courts to assure that forensic evidence is reliable and valid—and not merely generally accepted by others in the witness's field—before it is admitted in court. The highest courts of several states have also adopted this federal standard.

Historical context. Several areas of forensic specialization matured in the United States under the "general acceptance" standard announced in *Frye v. United States*, 293 F 1013 (D.C. Cir., 1923). This standard allowed scientific evidence to be admitted in court if the evidence was generally accepted as reliable by others in the expert's field. In 1993 in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, the U.S. Supreme Court rejected the *Frye* standard as the sole test for admissibility in federal courts, and interpreted the Federal Rules of Evidence to require proof that scientific evidence was valid before it would be found sufficiently reliable to be admitted into evidence. The Court noted that "scientists typically distinguish between 'validity' (does the principle support what it purports to show?) and 'reliability' (does application of the principle support consistent results?)," but stated that "our reference here is to *evidentiary* reliability—that is, trustworthiness. . . . In a case involving scientific evidence, *evidentiary reliability* will be based upon *scientific validity*." In 1999, in *Kumho Tire Co. v. Carmichael*, 526 U.S. 137, the Supreme Court made it clear that *Daubert*'s requirements of validity and reliability apply not only to scientific evidence but to all types of forensic evidence, and stated that forensic evidence must have "a reliable basis

in the knowledge and experience of the witness's discipline."

Many types of forensic evidence used in criminal proceedings were developed to meet the practical needs of law enforcement and prosecutorial functions of government, and not specifically to address theoretical or empirical questions. Courts are finding that many areas of forensic specialization are not grounded in science or scientific method, and are just beginning to formulate questions that may allow empirical investigation of the validity and reliability of their fields of knowledge. Areas of forensic knowledge at this stage of development are easy targets for challenges to reliability and validity in litigation.

Areas of forensic specialization that are grounded in science and scientific method are open to challenge on any aspect of theory or practice that may affect the reliability or validity of the evidence. Testimony of a witness or a peer that an established theory or technique has been applied in a manner that is generally accepted in the field is no longer sufficient to allow the evidence to be admitted. The trial court may consider general acceptance in evaluating the reliability and validity of the evidence, but independent proof of reliability and validity is also required.

Validity and reliability. The shift from "general acceptance" to "validity" and "reliability" is guided primarily by three elements of Federal Rule of Evidence 702 and state counterparts. This rule requires that (1) admissible forensic evidence must consist of "scientific, technical, or other specialized knowledge" (2) that will "assist the trier of fact to understand the evidence or to determine a fact in issue," and (3) must be provided by a witness who is qualified as an expert by "knowledge, skill, experience, training or education."

Specialized knowledge. *Daubert* defined the "specialized knowledge" element to require "scientific validity" as a basis for "evidentiary reliability." Applied to nonscientific evidence, *Kumho Tire* requires forensic evidence to have "a reliable basis in the knowledge and experience of the witness's discipline." Both *Daubert* and *Kumho Tire* identified four factors that trial courts may use to evaluate the validity and reliability of forensic evidence: (1) whether the knowledge or technique can be (and has been) tested, (2) whether it has been subjected to peer review and publication, (3) the technique's known or potential error rate, and (4) whether it is generally accepted in the witness's community. The Supreme Court cautioned in *Daubert* that this was not a "checklist" and that trial courts could consider anything that addressed the validity and reliability of the evidence. Later, in *Kumho Tire*, the Supreme Court stated that the trial judge has discretion to decide which, if any, of the four factors to employ in evaluating the reliability and validity of forensic evidence. The range of possible challenges to reliability and validity under these standards is limited only by the creativity of litigants and proponents of areas of specialized knowledge in our increasingly complex society.

Intradisciplinary knowledge. A forensic expert's knowledge of the scope and limits of the area of specialized knowledge claimed by the witness is, of course, subject to cross-examination under any or all of the four *Daubert* factors, and other factors raised by counsel and accepted by the court. A witness should therefore be knowledgeable of authoritative texts, treatises, standards, and training materials that constitute the knowledge base of the witness's claimed area of expertise. Opposing counsel may produce these materials and ask the witness to identify those deemed authoritative in the witness's field, and then request the witness to explain any discrepancies between the witness's testimony or practices and the content of the authoritative treatises. The witness may also be asked to provide a rational, logically and substantively valid explanation of the theories, techniques, standards, customs, and practices of the witness's field of specialization, and how these were applied to the facts of the case to reach the witness's opinions or conclusions. The purpose of this type of cross-examination is to identify any acts, omissions, biases, prejudices, or other matters that may affect the validity and reliability of the witness's evidence, and to provide a basis for counsel to assure that the evidence offered by the witness is within the scope of the witness's expertise.

Extradisciplinary knowledge. More importantly, a witness should attempt to be prepared for cross-examination concerning areas of knowledge that may be beyond the witness's area of specialization but are relevant to the determination of the validity and reliability of the witness's opinions or conclusions. *Daubert*'s expansion of the judicial inquiry beyond the realm of general acceptance in the witness's community supports this type of challenge, and opposing counsel may rely on this attribute of *Daubert* to test the limits of a witness's knowledge.

For example, the definition of "science" is at issue in an increasing number of cases. A witness from a field of specialized knowledge that is claimed to be a science may be challenged not only concerning the validity and reliability of evidence as viewed from within the witness's field of knowledge, but also concerning the validity and reliability of the witness's evidence as viewed from the field of science. Cross-examination may, for example, address how the witness's evidence accords with definitions of "science" and "scientific method," or with definitions of particular areas of forensic science or forensic specialization. *See* SCIENCE; SCIENTIFIC METHODS.

Criteria for "science" may be found in the vast literature in this field. For example, nine guideposts that courts and attorneys may use to prepare and challenge scientific evidence have been proposed [B. Black and P. Lee, 1997]: (1) the explanatory power of the evidence to explain, clarify, and predict relationships in general and in the case at hand; (2) logical consistency of the evidence; (3) falsifiability; (4) the variety and severity of testing of hypotheses on which the evidence is based; (5) precision and objectivity of the reasoning and conclusions, rather

than testimony consisting of broad generalizations and subjective impressions; (6) post-hypothesis testing, that is, whether the testimony is based on an ad hoc reinterpretation of existing data or on knowledge obtained through successful predictions and observations; (7) consistency of the opinion with accepted theories, that is, general acceptance in the witness's and related communities of specialization; (8) application and use of the knowledge of the witness by the scientific community outside the context of litigation; and (9) peer review and publication of the knowledge base or, where proprietary information is involved, other internal review.

Though not required under *Daubert* or *Kumho Tire*, these and other authoritative criteria for determining whether a field of knowledge is science may be invoked by counsel as a basis for establishing or challenging an area of specialized knowledge that is claimed to be a science. Such knowledge-based inquiries into the validity and reliability of a particular field of specialization are supported (and arguably required) by *Daubert*'s focus on validity and reliability. Challenges to forensic evidence from fields of knowledge outside the witness's field of knowledge may become the rule, rather than the exception, for cross-examination of witnesses in fields of forensic knowledge that have matured under the protection of general acceptance by only their guilded peers.

Assistance to trier of fact. Whether the forensic evidence will assist the trier of fact to understand the evidence or determine a fact in issue has been defined as whether there is a "fit" between the evidence and the issues in the case. Valid, reliable evidence offered by a qualified witness will be inadmissible if it does not assist the trier of fact. The Supreme Court's example of this element makes the point very well.

"A study of the phases of the moon, for example, may provide valid scientific 'knowledge' about whether a certain night was dark, and if darkness is a fact in issue, the knowledge will assist the trier of fact. However (absent creditable grounds supporting such a link), evidence that the moon was full on a certain night will not assist the trier of fact in determining whether an individual was unusually likely to have behaved irrationally on that night. Rule 702's 'helpfulness' standard requires a valid scientific connection to the pertinent inquiry as a precondition to admissibility."

Qualifications. The threshold inquiry into the witness's qualifications may address any and all formal education the witness has or has not completed. Education in the area of specialized knowledge in which the witness claims to be proficient will be most relevant. Any publications or other educational materials authored or edited by the witness may be reviewed and addressed on cross-examination, particularly if they make statements inconsistent with any statements made by the witness in the present case. Professional or technical training in the area of specialization, the witness's performance in that training, and certifications or other credentials related to the areas of specialization may also be addressed.

The witness's experience in the field of specialized knowledge may also be covered, together with any issues related to the witness's competent performance in the field. Performance evaluations, disciplinary proceedings, and any other evidence relevant to the witness's experience may be scrutinized. Testimony of the witness in other court proceedings may also be addressed if it may be used to undermine the reliability or validity of the witness's opinion in the case at hand.

Testimony. Though each deposition and court appearance is a unique experience, forensic witnesses may reasonably expect that direct and cross examination will cover at least the key areas addressed above. Direct examination will progress in a manner that allows the witness to favorably state the witness's education, training, and experience in the witness's field of specialized knowledge; the facts of the case on which the witness relied in preparing any reports, summaries, or formulating any opinions for the case; any theories, techniques, methods, or procedures applied by the witness in understanding and interpreting the facts; and any conclusions or opinions the witness may have reached as a result of this process. Direct examination is typically a smooth, comfortable exposition of the witness's qualifications, reasoning, and opinion.

Cross-examination by opposing counsel is typically more challenging. The primary goal of cross-examination is to identify any weaknesses that may undermine the relevance, reliability, and/or validity of the witness's evidence. Weaknesses will be sought in the witness's qualifications; knowledge of the scope, limits, validity, and reliability of the witness's area of specialization; application of the area of specialized knowledge to the facts of the case; and/or relevance of the evidence to the issues in the case. The forensic witness must, therefore, be prepared for intense scrutiny in any of these areas.

The intensity of cross-examination across the range of areas that may affect validity and reliability of forensic evidence has occasionally been interpreted by some forensic witnesses as a personal attack. Personalities aside, the forensic witness may be better prepared to face intense cross-examination by recalling that the forensic expert is, by definition, present to serve the court, and counsel is present to serve the court and a client. Counsel has a duty to assure that evidence presented against the client is valid and reliable; the forensic witness may request assistance from the court if cross-examination interferes with the witness's ability to serve the court's fact-finding function.

Court-appointed experts. Trial courts have the authority to appoint independent experts to consult with the court concerning issues presented in a case, but they rarely use this authority. *Daubert*'s instruction that the trial judge serve as "gate-keeper" may increase the frequency of appointment of experts by courts. The American Association for the Advancement of Science has launched a program to

assist federal judges in identifying appropriate court-appointed experts in civil cases. *See* FORENSIC CHEMISTRY; FORENSIC MEDICINE. Stephen A. Brunette

Bibliography. B. Black and P. Lee (eds.), *Expert Evidence: A Practitioner's Guide to Law, Science, and the FJC Manual*, 1997; J. Cecil and T. Willging, *Court-Appointed Experts: Defining the Role of Experts Appointed Under Federal Rule of Evidence 706*, 1993; D. Faigman et al., *Modern Scientific Evidence: The Law and Science of Expert Testimony*, 1997; Federal Judicial Center, *Reference Manual on Scientific Evidence*, 1994; P. Giannelli and E. Imwinkelried, *Scientific Evidence*, 2d ed., 1993; A. Moennsens et al., *Scientific Evidence in Civil and Criminal Cases*, 4th ed., 1995.

Forensic medicine

The branch of medicine concerned with the resolution of legal issues by the application of scientific medical knowledge. The issues may be of criminal or civil nature. The specialty of forensic medicine is often referred to as legal medicine in Europe and in Spanish-speaking countries. The development of forensic medicine began in Europe in the early nineteenth century. Although forensic medicine has its basis in the specialty of pathology, physicians in other specialties, such as psychiatry, orthopedics, cardiology, and neurology as well as specialists in other disciplines, may be called on to resolve legal questions posed by judges, attorneys, investigators, and hearing boards.

In death investigations, a number of forensic specialists may work together. A forensic odontologist may be called in to identify the deceased person through dental examination, or the perpetrator of bite marks left on the deceased's body. A forensic anthropologist may be needed to identify skeletal remains; a forensic toxicologist for the identification of poisons or drugs; and a criminalist for investigation at the scene and collection of evidence, or for study of trace evidence such as blood stains, hair, paints, and seminal fluid. In cases of equivocal suicides, forensic psychiatrists and psychologists may be called in. Recently the biomechanical engineer has been added to the roster of forensic specialists, to test and study injury patterns to determine how the injury came about. The forensic medical specialist may express an opinion in writing or may be required to testify in person in the courtroom. As with all legal proceedings, the physician witness must be cognizant of issues such as the degree of proof, the chain of custody of specimens or evidence, competency of the witness, and court procedures. *See* FORENSIC ANTHROPOLOGY; FORENSIC EVIDENCE.

Medicolegal investigative system. In many countries other than Canada and the United States, medical schools have departments of forensic or legal medicine, and medical students take courses in forensic medicine. In these countries, in cases of medicolegal disputes, the professor of legal medicine at the local medical school is asked to examine the patient or conduct the forensic autopsy, and to offer opinion to the courts, attorneys, or investigators. In the United States, in earlier times, the autopsy service was provided by a local physician hired by the coroner's office, which handled the investigation and determined the cause of death. In many of the states and larger metropolitan areas, beginning with the Commonwealth of Massachusetts and New York City, this system evolved into the medical examiner's system. The medical examiner's office has jurisdiction either statewide or countywide. A federal medical examiner's office has been established in Washington, DC. The American Board of Pathology sets the standards, establishes the qualifications, and certifies the competency of the forensic pathologists.

Most board-certified forensic pathologists work as medical examiners for a state or county. Their work involves death investigation, conducting forensic autopsy to determine the cause and manner of death, and often testifying in court. Few maintain full-time private consultation service. A well-staffed medical examiner's office has toxicological and forensic laboratories; a team of investigators; consultants in such forensic specialties as odontology, anthropology, and entomology; and medical consultants in radiology and psychiatry.

Clinical forensic medicine. Forensic medical specialists also deal with medicolegal issues involving living persons. Psychiatrists may assist with the evaluation of insanity pleas by defendants; a cardiologist may deal with a worker's claim for disability due to heart disease; and an orthopedic surgeon may evaluate the degree of a patient's disability for a court or for attorneys prior to settlement. In the United Kingdom and former commonwealth countries, and in the Far East, a system of police surgeons is set up to provide medical advice to the local police. A physician appointed as police surgeon is called on to visit the jail to examine a detainee, or is called to the scene of death. In the United States, clinical evaluation of jail inmates is provided by individual practicing physicians, and the setup varies with the locality.

Authority and practice. In the United States, deaths that are sudden and unexpected or violent are usually investigated by the medical examiner's office, except in areas where the coroner still has jurisdiction. Even in those jurisdictions, most autopsies are conducted by board-certified forensic pathologists. Throughout the world, such deaths are investigated by an official agency. In continental Europe and the Far East, the police conduct the investigation, with an autopsy authorized by a magistrate. In the United Kingdom and former commonwealth countries, the coroner, who is also the magistrate, conducts the inquest. The medical examiner in the United States is a board-certified forensic pathologist, who serves as an appointed rather than an elected official. In all states there are state or countywide medical examiner's offices. Their purpose is to determine by investigation the cause and manner of death. Typical cases include gunshot wounds, stab wounds, blunt-force injuries, traffic accidents, drowning, hanging, asphyxiation, and drug and alcohol abuse. Any death

suspected of being an accident, suicide, or homicide is investigated.

The medical examiner's authority has been expanded to cover such incidents as death in nursing homes, sudden and unexpected death during surgery, or death of a prisoner in custody. Furthermore, the medical examiner's office usually takes jurisdiction over cases of undiagnosed epidemic diseases that constitute a public hazard. In some jurisdictions, any proposed cremation of the remains is usually subjected to inquiry by the medical examiner's office prior to the cremation.

The medical examiner's office not only certifies the cause and the manner of death, but also helps prevent accidental, suicidal, and homicidal deaths by maintaining and disseminating biomedical information and by participating in professional and public education. *See* PUBLIC HEALTH.

With the advent of transplants as therapeutic procedures, some medical examiners have become involved in the administration of organ and tissue donation programs and related ethical matters. This necessitated a new legal definition of death. Moreover, ethical problems have arisen concerning the patient's right to refuse certain kinds of care. When a patient is dying or irreversibly comatose, such decision-making is a troubling matter. In some states, a surrogate can be authorized to make decisions about the patient's care (power of attorney for health care). Forensic medicine plays a vital role in the enactment of appropriate legislation to handle these and other ethical matters. *See* DEATH.

Drug and poison detection. Drug and alcohol abuse have been social problems for centuries. Before World War II, alcohol, barbiturates, heroin, aspirin, cyanide, and arsenic were the chemicals most commonly causing death. In the 1960s, stimulants and hallucinogenic drugs arrived. In the 1970s, phencyclidine (PCP) and lysergic acid diethylamide (LSD) caused major problems and deaths, and in the 1980s cocaine became the most widely publicized illegal drug involving fatalities. The medical examiner determines the cause and manner of death due to drugs and alcohol, but also assesses drug and alcohol involvement in deaths due to traffic accidents, homicides, and suicides, and maternal and perinatal deaths.

The rapid improvement in the accuracy and sensitivity of analytical instrumentation, coupled with the availability of sophisticated computers, now makes it possible to detect drugs, poisons, and their metabolites in samples as small as 10^{-12} g. Although the incidence of drug overdose leading to deaths has declined, alcohol, multiple therapeutic drugs, and illegal substances continue to plague society. *See* FORENSIC CHEMISTRY; FORENSIC TOXICOLOGY.

Scanning electron microscope. As part of the medicolegal investigation, the medical examiner performs microscopic examination of tissues. The analysis of trace materials, including gunshot residue, fibers, glass, and serological material, may greatly add to the understanding of the circumstances surrounding death. Microscopic examination, in particular with the scanning electron microscope (**Fig. 1**), which permits viewing of minute objects with magnifications up to 100,000 times, is useful for the analysis and identification of minute fragments of objects left on human bone, cartilage, and soft tissue, such as the crushing marks of an assault tool on the skin and serrated markings on bone in the case of stabbing. Another example is electrocution, where the skin may show the burn of the electric wire and a deposit of copper. Scanning electron microscopy coupled with energy-dispersive x-ray spectroscopy can effectively identify chemicals deposited on the wound and may be used to identify gunshot residues, which are composed principally of lead, antimony, and barium. Atomic absorption spectroscopy and neutron activation analysis are also used for this purpose. *See* ANALYTICAL CHEMISTRY; SCANNING ELECTRON MICROSCOPE.

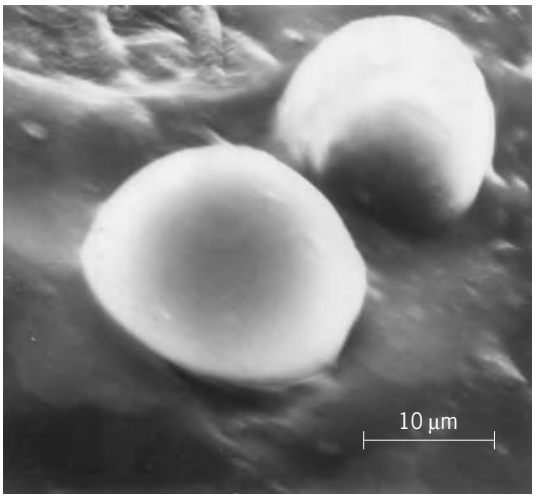

Fig. 1. Scanning electron microscopic image. The two round objects represent molten primer compound emitted from a recently fired cartridge. Energy-dispersive x-ray spectroscopy produced the spectrum characteristic of gunshot residue. (*Office of Chief Medical Examiner-Coroner, County of Los Angeles*)

Identification of bite marks. During the investigation of a death, a forensic pathologist looks for bruises and marking on the victim's skin. The assailant may leave a bite mark on the skin, which can be used to identify the person. Bite marks are second only to fingerprints in their forensic value for identifying murderers, particularly serial killers.

Genetic analysis. Another important development is the forensic application of genetic analysis. Every individual has a unique genetic content determined by deoxyribonucleic acid (DNA) sequences. The DNA present in forensic samples such as hair, bloodstains, and seminal fluid can identify a suspect. One analysis technique is called DNA fingerprinting; it can be used to compare the DNA sequence of a suspect or a defendant with material evidence at the crime scene. The DNA is broken into unique fragments by restriction endonucleases and then separated by electrophoresis. The pattern seen after hybridization with specific probes is called the DNA fingerprint, and like an ordinary fingerprint is

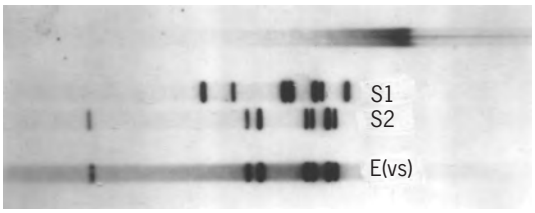

Fig. 2. DNA sequence of two rape-murder suspects, S1 and S2, are compared with the DNA sequence of the seminal fluid found in the vaginal smear from the victim. Suspect sample S2 matches the DNA sequence of the seminal fluid E(vs). A guilty verdict was rendered against suspect S2. (*Cellmark Diagnostics, Division of ICI Americas Inc.*)

characteristic of the individual (**Fig. 2**). DNA analysis is also used in identification of an unknown person or of parts of a human body. Paternity exclusion by ABO blood typing has been replaced by DNA fingerprinting. *See* DEOXYRIBONUCLEIC ACID (DNA).

Thomas T. Noguchi

Bibliography. K. A. Bennett, *A Field Guide for Human Skeletal Identification*, 2d ed., 1993; S. B. Karch, *The Pathology of Drug Abuse*, 3d ed., CRC Press, Boca Raton, 2001; N. Rudin and K. Inman, *An Introduction to Forensic DNA Analysis*, CRC Press, Boca Raton, FL, 2d ed., 2001; W. U. Spitz, *Spitz and Fisher's Medicolegal Investigation of Death: Guidelines for the Application to Crime Investigation*, 3d ed., Charles C. Thomas, Springfield, IL, 1993; C. V. Wetli, R. E. Mittleman, and V. J. Rao, *An Atlas of Forensic Pathology*, ASCP (American Society of Clinical Pathologists) Press, Chicago, 1999.

Forensic microscopy

The application of microscopy techniques for purposes of civil or criminal law. Edmond Locard (1877–1966) stated that every contact leaves a trace. This is known as the Locard exchange principle, the basis for much of forensic microscopy. Since the early 1800s, the microscope has been used to help solve crimes. Today, it remains one of the most used tools in the crime laboratory. Criminal evidence ranging in scale from micrometer-sized particles to hair and paint chips can be found and removed to a laboratory for microscopic analysis, where its history may be deduced to help solve the crime. Microscopy can provide insight into the identity and origin of a material, what has happened to it and when, and the routes it may have taken between a crime victim, suspect, and crime scene (**Fig. 1**). *See* CRIMINALISTICS.

Stereomicroscope. The low-powered stereomicroscope is probably the most used microscope in the crime lab. The magnification typically ranges from 7× to 40×, and the image is upright (not reversed) and three-dimensional. With additional lenses, the magnification can range from below 3× to above 100×. This type of microscope can be used with a variety of lighting techniques such as transmitted, fluorescence, diffuse, coaxial, and oblique illumination. *See* FLUORESCENCE MICROSCOPE; MICROSCOPE; OPTICAL MICROSCOPE.

The stereomicroscope is used in drug analysis to help identify marijuana and to evaluate powders and chunks of material that may need to be separated before analysis. In examining firearms and tool marks, it is used to evaluate striations or imprint marks, look for trace evidence, evaluate bullet holes for muzzle blast, and search for unburned gunpowder particles. In addition, when a bullet passes through an object, some of that object may be transferred to the bullet and detected microscopically. In forensic biology, the stereomicroscope is used to find and measure small bloodstains, to evaluate a bloodstain to determine if it came from the inside or outside of a garment, and to remove stains from hidden places. When examining documents, it can be used to evaluate typefaces, examine strikeovers and ink lines that cross over, and find erasures or alterations. For trace evidence, the stereomicroscope is the tool of choice for finding and characterizing hairs, fibers, paint, glass, soil, and building material. *See* FORENSIC BIOLOGY.

Biological microscope. The biological microscope is used for examining biological fluids, human tissue, hairs, fibers, drug crystals, food, and other stains. A typical biological microscope has 4×, 10×, 20×, and 40× objectives on a revolving nosepiece. It is normally equipped with 10× eyepieces giving a final magnification of from 40× to 400×. Many additional lenses can be added to this style of microscope to increase the range of magnification from less than 20× to greater than 1500×. In addition, these microscopes can be modified to include darkfield, phase contrast, fluorescence, dispersion staining, darkfield epi-illumination, and reflected coaxial illumination. When darkfield epi-illumination or reflected coaxial illumination is used, the biological microscope is transformed into a metallurgical microscope, which

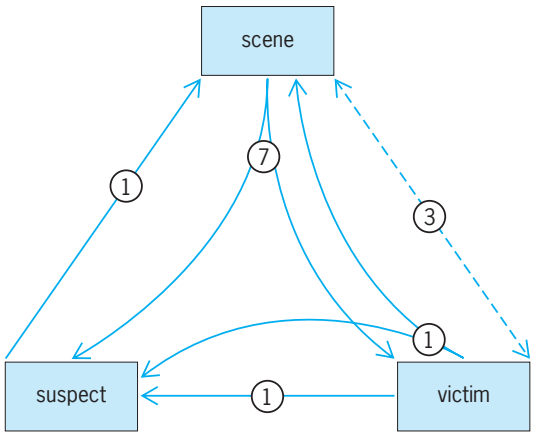

Fig. 1. Hypothetical flow of trace evidence between a crime scene, victim, and suspect. Seven types of trace evidence were transferred from the crime scene to both the suspect and victim. Three types of trace evidence were found at both the crime scene and on the victim. One type of trace evidence was transferred from the victim to the suspect and the crime scene. One type of evidence was transferred from the victim to the suspect. One type of evidence was transferred from the suspect to the crime scene.

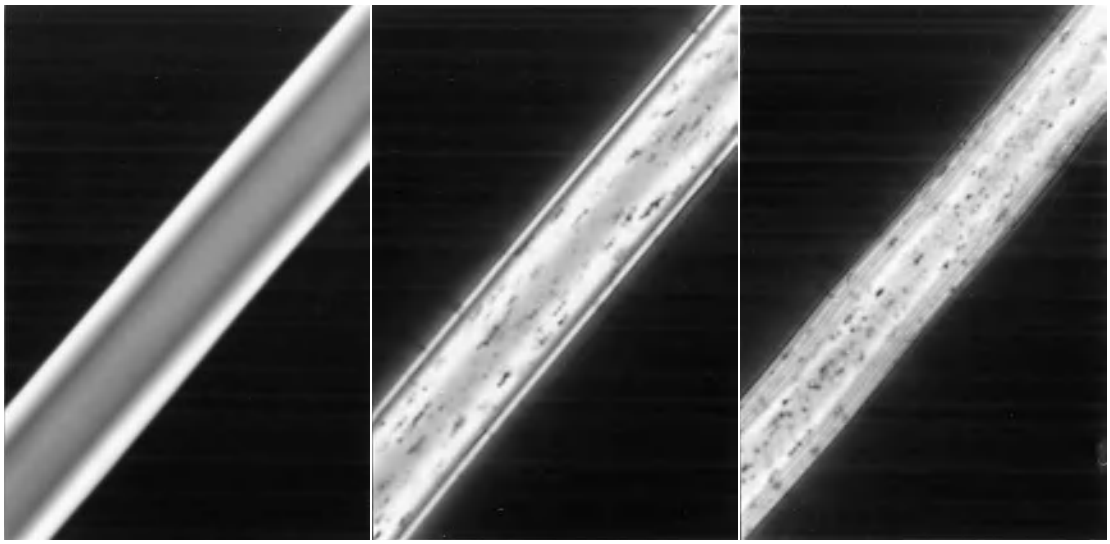

polypropylene nylon polyester

Fig. 2. Crossed polarizers showing interference colors (represented here in black and white) of three colorless synthetic fibers. These interference colors can be used to calculate the birefringence, which is an optical property used to identify these fibers. The birefringence of polypropylene is medium, nylon is high, and polyester is very high.

can be used to examine surface detail and opaque objects. *See* PHASE-CONTRAST MICROSCOPE; REFLECTING MICROSCOPE.

A forensic pathologist typically uses the biological microscope to examine tissue samples from injuries to determine if the injuries occurred before, during, or after death. In the forensic biology section of the laboratory, the microscope is typically used to identify sperm in rape cases. This process involves differential staining and examining the slide at greater than 200×. When sperm are identified from samples collected from a victim, this is proof that sexual activity has taken place. The legal issue of consent is not addressed by the presence of sperm. Other body fluids such as vaginal secretions, saliva, and feces might have been mixed with semen, and microscopy can give vital clues as to the origins of these other body fluids.

Fig. 3. Fifteen anagen hair roots sent for DNA typing.

Some crime labs use crystal tests to identify drugs such as phencyclidine, amphetamines, cocaine, barbiturates, heroin, morphine, and codeine. In a typical crystal test, a small amount of the drug and a liquid reagent are placed on a microscope slide. After a short time, unique crystals for that drug can be recognized at 100× with a biological microscope.

Polarized light microscope. The polarized light microscope is a biological microscope with special modifications, including a polarizer, analyzer, rotating stage, accessory slot, Bertrand lens, and a flip-up condenser. This microscope is typically used to identify synthetic fibers, minerals, glass, starch, and many other types of particles. It has the advantage of measuring optical properties without altering the sample. Six major synthetic fiber types (acrylic, acetate, rayon, olefin, nylon, and polyester) can be quickly differentiated with the polarized light microscope (**Fig. 2**). A significant amount of training and experience, as well as reference standards, is necessary to master the use of this instrument, since most universities do not teach this subject. The McCrone Research Institute in Chicago has trained many forensic scientists in the use of the polarized light microscope. *See* BIREFRINGENCE; POLARIZED LIGHT MICROSCOPE.

Comparison microscope. A comparison microscope consists of two identically equipped compound microscopes joined together with an optical bridge, giving a split-screen image. The comparison microscope is used to compare bullets, cartridge cases, and toolmarks. This microscope normally uses low-powered objectives and reflected, diffused, or oblique illumination. The comparison microscope is also used to compare trace evidence such as hairs and fibers. It uses the same range of magnifications as the biological microscope.

One such case illustrating the comparison of hair, as well as DNA analysis, involved an elderly male

victim who was stabbed in a robbery at a hamburger stand. Clumps of long dark hair that appeared to have been pulled out during a struggle were found in several locations at the crime scene. Microscopic examination of that hair showed more than 50 anagen roots. Anagen roots are those still attached to the scalp with a blood vessel and a nerve, indicating that the clumps of hair were forcibly removed during the struggle. The hair from the crime scene had a maximum length of 17 in. The maximum length of the victim's gray hair was less than 1 in. The police collected and the crime lab compared hair from several long-haired suspects with negative results. Finally, a female drug user was identified by a relative as the murderer. A hair sample was collected and submitted to the hair examiner, who testified that the hairs from the crime scene probably came from the suspect. A blood sample from the suspect along with 15 roots from the crime scene hairs were submitted to a DNA laboratory (**Fig. 3**). The DNA matched, and the suspect was convicted. This case became the first California DNA case to successfully get through the Appeals Court.

Scanning electron microscope. The scanning electron microscope with energy dispersive x-ray spectrometer (SEM/EDX) is routinely used to identify the elemental content of very small samples. The most common sample analyzed with the SEM/EDX is gunshot residue. This is done by dabbing the shooter's hands with sticky tape attached to an SEM stub (a substrate for mounting specimens). The stub is then examined for gunshot residue (GSR) particles (usually 1–6 micrometers) that contain lead, barium, and antimony. Many of the SEM/EDX microscopes are connected to computers that have programs that can search automatically for GSR particles. *See* SCANNING ELECTRON MICROSCOPE.

Digital forensic microscopy. Most microscope systems can be modified to capture digital images, which can be sent by e-mail to forensic scientists for examination and comment. An automated system for glass refractive index measurement uses a phase-contrast microscope, hot stage, video system, and computer to compare glass samples. The National Integrated Ballistics Identification Network (NIBIN), sponsored by the Federal Bureau of Investigation (FBI) and Bureau of Alcohol, Tobacco and Firearms (ATF), uses digitally captured images (through the microscope) of cartridge cases or bullets to link shooting incidents with guns taken from suspects in distant jurisdictions.

The infrared spectrophotometer with a microscope attached is a common tool in crime laboratories. It is used to perform spectrophotometric analysis on very small samples such as paints, plastics, fibers, and many other substances. The microspectrophotometer is used to measure the spectrum of the visible colors, ultraviolet radiation, and fluorescence that are present in microscopic objects. An automated version of the microspectrophotometer using a remote control stage and computer can search tape lifts taken from suspects, victims, and crime scenes. First a fiber is chosen and entered into the apparatus. The tape lifts are then searched for that target fiber.

Another digital system is a portable microscope with fiber-optic cables that is used to acquire high-quality microscopic images such as a paint transfer on a bumper or hood of a vehicle without dismantling.

Edwin L. Jones, Jr.

Bibliography. P. De Forest, Foundations of forensic microscopy, in R. Saferstein (ed.), *Forensic Science Handbook*, Prentice Hall, Upper Saddle River, NJ, 2002; B. Fisher, *Techniques of Crime Scene Investigation*, CRC Press, Boca Raton, 1993; J. Houde, *Crime Lab: A Guide for Nonscientists*, Calico Press, Ventura, 1999; S. Palinik, Microscopy, in J. A. Siegal et al. (eds.), *Encyclopedia of Forensic Sciences*, Academic Press, San Diego, 2000; R. Saferstein, *Criminalistics: An Introduction to Forensic Science*, Prentice Hall, Upper Saddle River, NJ, 2001.

Forensic physics

The application of physics for purposes of civil or criminal law. Forensic science has been dominated by the fields of chemistry (mainly analytical chemistry) and biology (serology and DNA profiling). Indirectly, physics has contributed to forensic science via the invention of the microscope, the electron microscope, the mass spectrometer, and optical spectrometers; but directly, its role in forensic science has been minimal. Forensic physics has traditionally involved the measurement of density (soil and glass examination), index of refraction, and birefringence (fiber analysis, glass examination). In the last 25 years, the use of the photoluminescence phenomenon for physical evidence examination has emerged, with latent fingerprint detection the most notable application. In criminalistics, fingerprint detection is important because it provides absolute identity, and does not suffer from the contamination problems to which DNA profiling is prone. *See* CRIMINALISTICS.

Photoluminescence. Light emission by substances can have various causes, such as heat (incandescence), electricity (electroluminescence), or a chemical reaction (chemiluminescence), as in the reaction of luminol (3-aminophthalhydrazide) with blood, a much-used criminalistics procedure. The most important origin of luminescence in forensic science is the prior absorption of light by the substance and subsequent emission of light at longer wavelength, called photoluminescence. The photoluminescence phenomenon was used in various science fields long before its quantum-mechanical underpinnings were understood. Early on, two types of photoluminescence were distinguished on the basis of photoluminescence lifetime, namely fluorescence (of short lifetime) and phosphorescence (of long lifetime). On practical grounds, this distinction suffices. Fluorescence techniques are presently more widely used in criminalistics than phosphorescence techniques, simply because intense fluorescence is generally encountered

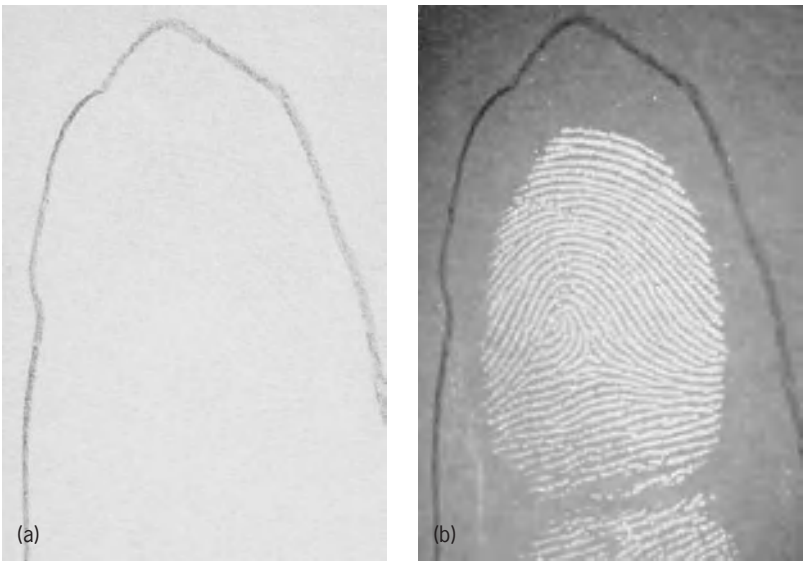

Fig. 1. Fingerprint on paper (*a*) after standard ninhydrin development and (*b*) after subsequent processing by zinc chloride and laser photoluminescence excitation.

in nature more frequently than phosphorescence. Phosphorescence approaches, however, are beginning to show much promise, especially in instances in which fluorescence approaches fail. *See* LUMINESCENCE; PHOTOLUMINESCENCE.

Fluorescence techniques. In the past, a white surface would typically be dusted with a black powder to reveal a latent fingerprint. This mode of visualization involves absorption/reflectance phenomena. In the case of a weak print, to which only few of the absorbing black powder particles adhere, the print would reflect only slightly less than the surroundings. This inherently limits sensitivity because of the small difference between the two relatively large light signals. The photoluminescence approach detects a small signal, rather than a small difference between large signals and is therefore inherently more sensitive. Thus, photoluminescence-based techniques have generally supplanted absorption/reflectance techniques in most scientific fields. The principal application of photoluminescence in forensic science is fingerprint detection. Fingerprint residue left on an article typically has a mass of about a tenth of a milligram, and is composed primarily of water (about 98–99%), which soon evaporates, with the remaining residue made up about equally of inorganic material (mainly salts, largely useless for detection purposes) and a myriad of organic compounds. Thus, only nanograms of organic material are available to be made visible to the naked eye. Not only does the material have to be detected, but also its spatial distribution has to be visualized in fine detail. Thus, lasers, by virtue of their power (at the right wavelength to excite luminescence) were the first choice of illumination for purposes of photoluminescence detection of fingerprints (**Fig. 1**). Filtered lamps have since seen wide use, but sacrifice sensitivity in favor of price. Other photoluminescence applications in criminalistics include fiber analysis (especially in finding otherwise elusive fibers), document examination (detecting altered and erased writing), detection of body fluids (mostly semen and blood), trace evidence detection, authentication of currency and passports, explosives origin specification, and DNA profiling (with the radioactive tags of the past replaced by fluorescent tags). *See* FINGERPRINT; FLUORESCENCE.

Phosphorescence techniques. In fingerprint detection, one often encounters background fluorescence from the substrate on which the fingerprint is located, and this background fluorescence, which typically has a lifetime on the order of a nanosecond, cannot always be eliminated by optical filtering. Detection sensitivity is therefore reduced. This is a general problem with analytical fluorescence methods.

There has emerged in recent years a procedure whereby fingerprints are treated so that they exhibit phosphorescence, and then time-resolved imaging techniques are applied to suppress background fluorescence, to detect otherwise elusive prints. In such techniques, the imaging device is turned on and off in synchronization with the photoluminescence excitation source, which is also periodically turned on and off. Recording by the imaging device is delayed with respect to the excitation source turn-off, such that the background fluorescence has decayed by the time the imaging device turns on. Likewise, the imaging device turns off before the onset of the next illumination pulse (**Fig. 2**). Typically, a charge-coupled device (CCD) camera equipped with a proximity-focused microchannel plate image intensifier is used as the imaging device. The time-resolved detection technique is not yet routinely used in crime laboratories, but operational prototype instruments have been in existence in United States and Canadian university forensic science laboratories since the mid-1990s. *See* CHARGED-COUPLED DEVICES; PHOSPHORESCENCE.

A number of fingerprint treatments that lend themselves to time-resolved imaging, mostly based on lanthanide compounds that display millisecond

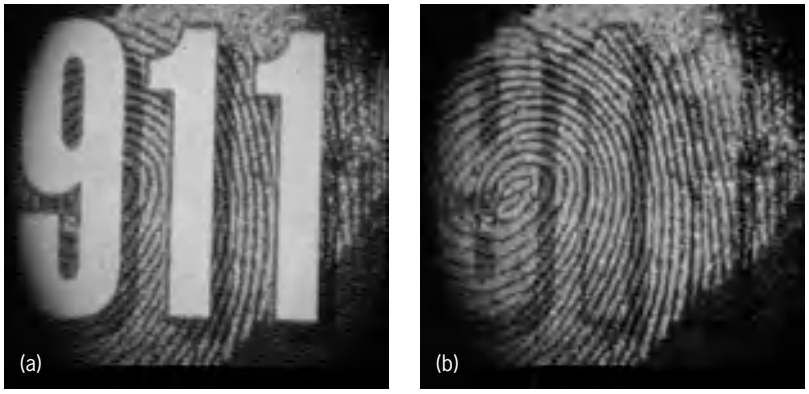

Fig. 2. Fingerprint (*a*) after standard photoluminescence development and (*b*) after time-resolved development. Note the suppression of the background fluorescence from the numbers 911.

luminescence lifetimes, have been developed. Time-resolved imaging techniques typically require phosphorescence lifetimes longer than microseconds. This lifetime requirement places considerable restrictions on the procedures that can be used to treat latent fingerprints.

Treatments that result in fingerprint luminescence lifetimes of 10–1000 nanoseconds can be used if phase-resolved imaging is implemented. Instead of turning the illumination source on and off (by means of a light chopper), the source intensity is modulated sinusoidally. This allows for much higher modulation frequencies than is possible with the abrupt on-off modulation of time-resolved techniques. In response to the sinusoidal luminescence excitation, the luminescence will also be sinusoidally modulated in intensity, but will be phase-shifted by an angle ϕ with respect to the excitation according to $\tan \phi = \omega\tau$, where ω is the modulation angular frequency and τ is the luminescence lifetime. If the gain of the imaging device (CCD camera equipped with a proximity-focused microchannel plate image intensifier) is modulated with the appropriate phase delay versus the modulation of the light source, a suppression of background fluorescence (small ϕ) with respect to the fingerprint luminescence is achieved (optimally ϕ is roughly 45° because phase shifts approaching the maximum of 90° lead to large demodulation of the luminescence intensity). The general scheme here is reminiscent of the lock-in amplifier, long used as a frequency and phase selective device. Phase-resolved imaging has been implemented in cell microscopy, for instance, but instruments have yet to be designed for forensic applications, perhaps because the corresponding fingerprint treatments have yet to reach maturity. There are, however, promising fingerprint treatments in the offing that appear suitable for phase-resolved imaging. The luminescence excitation sources needed for time- and phase-resolved imaging will have to be lasers.

Nanoparticles. Semiconductor material, when fabricated in size on the order of nanometers, may exhibit luminescence intermediate between what would be obtained from the bulk material and would be obtained from individual ions. The photoluminescence lifetimes of such nanoparticles tend to be 10–1000 ns. The color of the luminescence and the luminescence lifetime can be tailored by particle size. The broad absorption spectrum of the particle is typical of semiconductors, and thus amenable to optical luminescence excitation spanning a wide range of wavelengths; whereas the particle's luminescence is sharply defined, akin to molecular fluorescences, and occurs at the (long-wavelength) absorption band edge. Nanoparticles are ideally suited for fingerprint work in terms of their spectroscopic properties, especially in their ability to suppress background fluorescence.

Presently, nanoparticles are used as selective taggants mostly in the biosciences, but the tagging chemistries should apply to fingerprint detection, making this an active area of current forensic research. Forensic application at present focuses on cadmium sulfide nanoparticles incorporated into tree-shaped polymers called dendrimers. The dendrimers have terminal functional chemical groups that selectively tag fingerprints. Cadmium selenide nanocrystals are also used. These crystals are typically encapsulated in zinc sulfide (or silica). The encapsulant protects the nanocrystal, such that its optical properties are not compromised, and functions as the attachment site for molecules (conjugating ligands), which chemically react with the material to be detected (**Fig. 3**). *See* NANOPARTICLES.

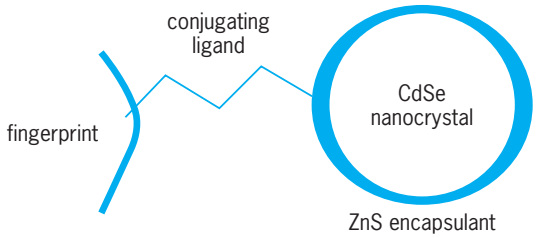

Fig. 3. General scheme for fingerprint detection with nanocrystals.

Given that nanotechnology is a very active field today, with much bioscience application, it would seem to be only a matter of time before it finds its way to routine forensic science work. To date, financial considerations and the conservative nature of the forensic science community itself have been obstacles to its implementation. *See* NANOTECHNOLOGY.

E. Roland Menzel

Bibliography. H. C. Lee and R. E. Gaensslen, *Advances in Fingerprint Technology*, Elsevier, 1991; E. R. Menzel, *Fingerprint Detection with Lasers*, 2d ed., Marcel Dekker, 1999; E. R. Menzel, *Laser Spectroscopy: Techniques and Applications*, Marcel Dekker, 1995; E. R. Menzel, Photoluminescence detection of latent fingerprints with quantum dots for time-resolved imaging, *Fingerprint Whorld*, 26(101):119–123, 2000; E. R. Menzel et al., Photoluminescent CdS/dendrimer nanocomposites for fingerprint detection, *J. Forensic Sci.*, 45(4):770–773, 2000; E. R. Menzel et al., Photoluminescent semiconductor nanocrystals for fingerprint detection, *J. Forensic Sci.*, 45(3):545–551, 2000; R. Saferstein (ed.), *Forensic Science Handbook*, Prentice Hall, 1982; R. Saferstein, *Criminalistics*, 6th ed., Prentice Hall, 1998; D. A. Tomalia et al., A new class of polymers: Starburst-dendritic macromolecules, *Polymer J.*, 17(1):117–132, 1985.

Forensic toxicology

An interdisciplinary science dealing with the adverse effects of drugs and chemicals on various biological systems in a medical-legal context. The forensic toxicologist may work with a medical examiner or coroner in order to determine the role that a particular chemical compound may have played in a death. The forensic toxicologist's activities may also

involve assessing emergency room patients, helping to determine suitability of an individual for employment or promotion, screening for performance-altering drugs in athletes, and working with law enforcement agencies, for example, performing tests to determine if a driver operated a motor vehicle under the influence of drugs. The forensic toxicologist is involved not only in the analysis of body fluids and tissue for drugs and poisons but also in the interpretation of the resulting information in a judicial context.

Specimens and samples. Because of the medical-legal context of forensic toxicology, specimens must be maintained under a "chain of custody." Each sample must be accompanied by a paper trail allowing for the tracing of the specimen to its origin. The external portion of the chain of custody should contain information such as the collector, date, time, storage (place and conditions), courier, and time and date of delivery to the laboratory. Once the specimen has been accepted by the toxicology laboratory, an internal chain of custody tracks the persons having access to the samples and the procedures applied to those tissues. Each analysis must then be validated. Validation is an exhaustive accumulation of evidence that a given procedure will produce accurate results.

Sources of samples. Death unattended by a physician or occurring under violent, unusual, or sudden unexplained conditions happens in approximately 20% of the population and requires a thorough medical-legal investigation. The forensic toxicologist provides for the isolation of chemicals from biological samples and the subsequent analysis to determine whether a chemical agent played a role in the cause of death. At autopsy, the forensic pathologist collects postmortem specimens (such as blood, brain, liver, kidney, stomach contents, and bile, and on occasion, lung, bone, hair, fingernails, and fat). Following the extraction of chemicals from tissues, these substances are identified and quantified. The distribution of chemicals in the body provides information about the mode of exposure (for example, ingestion, injection, or inhalation) and the time of exposure relative to death. Of growing importance to the field of forensic toxicology is the analysis of metabolites related to the parent drugs. The forensic toxicologist working in a drug abuse program, emergency room, clinical laboratory, or workplace program is limited to specimens such as blood, urine, and vomitus, which are more easily collected from living patients. Recent publications have explored the use of hair, sweat, and oral fluids as potential forensic specimens.

Analytical techniques. Immunoassays for various drugs and classes of drugs have become commercially available, and equipment has been perfected for automating these procedures. Therefore, immunoassay tests may be used by the forensic toxicologist as the first line of testing. Traditionally, the toxicologist has used a number of isolation techniques such as steam distillation, selective solvent extraction, and microdiffusion, depending on the tissue and the analyte. The use of solid-phase extraction has become commonplace and is applicable to semiautomated procedures. Once the substance has been isolated from the biological specimen, techniques such as chromatography (thin-layer, paper, gas, or liquid), spectrophotometry (colorimetric, ultraviolet, visible, and infrared), and immunoassays can be applied to qualitatively and quantitatively determine the drug. The combination of gas chromatography (and more recently, high-performance liquid chromatography) with mass spectrometry is presently accepted as the "gold standard" of analysis.

Drugs, driving, and traffic safety. One application of forensic toxicology that emerged during the 1990s, concomitant with increased emphasis on traffic safety, is the determination of the role of drugs and alcohol on the ability to operate a motor vehicle. Working with law enforcement agencies, the forensic toxicologist has assumed the vital role of providing scientific confirmation of the conclusions reached by the trained drug recognition expert. Analytical methods have been developed and applied to provide unquestioned identification of drugs and their metabolites as well as dependable quantification. The law enforcement officer may be able to recognize the signs and symptoms of a particular drug; however, the forensic toxicologist can best relate the drug at a given concentration to human psychomotor performance.

Interpretation. Once the analytical information has been collected, the forensic toxicologist must interpret the finding with reference to a cause of death or another set of medical-legal circumstances, such as workplace testing or accusation of driving under the influence of alcohol or drugs. Toxicological data can be obtained from pharmaceutical manufacturers, published literature, or various registries. The greatest challenge facing the forensic toxicologist is the interpretation of cases involving combinations of drugs and chemicals and their complex interactions. *See* FORENSIC CHEMISTRY; FORENSIC MEDICINE; TOXICOLOGY.

Jesse H. Bidanset

Bibliography. R. C. Baselt, *Disposition of Toxic Drugs and Chemicals in Man*, 7th ed., Biochemical Publications, 2004; H. Brandenberger and R. A. A. Maes (eds.), *Analytical Toxicology for Clinical, Forensic and Pharmaceutical Chemists*, W. de Gruyter, Berlin, 1997; J. Chamberlain, *Analysis of Drugs in Biological Fluids*, 2d ed., CRC Press, 1995; A. S. Curry, *Poison Detection in Human Organs*, 4th ed., Charles C. Thomas, 1988; M. J. Ellenhorn et al. (eds.), *Ellenhorn's Medical Toxicology: Diagnosis and Treatment of Human Poisoning*, 2d ed., Williams & Wilkins, 1997; L. R. Goldfrank et al., *Goldfrank's Toxicological Emergencies*, 7th ed., McGraw-Hill, 2002; C. D. Klaassen, *Casarett and Doull's Toxicology: The Basic Science of Poisons*, 6th ed., McGraw-Hill, 2001; B. Levine (ed.), *Principles of Forensic Toxicology*, 2d ed., American Association for Clinical Chemistry, 2003; A. C. Moffat (ed.), *Clarke's Isolation and Identification of Drugs in Pharmaceuticals, Body Fluids and Post-Mortem Material*, 2d ed., Pharmaceutical Press, London, 1986.

Forest and forestry

A plant community consisting predominantly of trees and other woody vegetation, growing closely together, is a forest. Forests cover about one-fourth of the land area on Earth. The trees can be large and densely packed, as they are in the coastal forests of the Pacific Northwest, or they can be relatively small and sparsely scattered, as they are in the dry tropical forests of sub-Saharan Africa.

Forests are complex ecosystems that also include soils and decaying organic matter, fungi and bacteria, herbs and shrubs, vines and lichens, ferns and mosses, insects and spiders, reptiles and amphibians, birds and mammals, and many other organisms. All of these components constitute an intricate web with many interconnections. Thus, a bird may be dependent upon the upper branches of a tree for nesting, while the tree may be dependent upon the fungi surrounding its roots to obtain water and nutrients. *See* FOREST ECOSYSTEM.

Forests have important functions, such as cleansing the air, moderating the climate, filtering water, cycling nutrients, providing habitat, and performing a number of other vital environmental services. They also supply a variety of valuable products ranging from pharmaceuticals and greenery to lumber and paper products. *See* FOREST TIMBER RESOURCES; LUMBER.

Forest types. There are many ways to classify forests, as by (1) location (for example, temperate zone forests, tropical zone forests); (2) ownership (for example, public forests, private forests); (3) age or origin (for example, old-growth forests, second-growth forests, plantation forests); (4) important species (such as Douglas-fir forests, redwood forests); (5) economic and social importance (for example, commercial forests, noncommercial forests, urban forests, wilderness); (6) wood properties (for example, hardwood forests, softwood forests); (7) botanical makeup (for example, broadleaf forests, evergreen forests); or (8) a combination of features (such as moist temperate coniferous forests, dry tropical deciduous forests). The last approach tends to be the most descriptive because it often integrates several dominant characteristics related to climate, geography, and botanical features. Some examples of the major forest types are:

1. Northern coniferous forests which span the cold, northern latitudes of Canada and Europe. They are dominated by vast expanses of spruce, fir, and larch, although species of birch, aspen, and pine often are present following fire and other disturbances.

2. Temperate mixed forests which occupy the eastern United States, southeastern Canada, central Europe, Japan, and East Asia, and parts of the Southern Hemisphere in Chile, Argentina, Australia, and New Zealand. They comprise a variety of broadleaf species (such as oak, hickory, beech, and maple in the Northern Hemisphere, and eucalyptus and southern beech in the Southern Hemisphere); subregions such as the southeastern United States also can have large components of pine or mixed pine-broadleaf forests.

3. Temperate rainforests which are situated along moist, coastal regions of the Pacific Northwest, southern Chile, southeastern Australia, and Tasmania. They are frequently dominated by large conifers (such as redwood, Douglas-fir, spruce, cedar, and hemlock) in the Northern Hemisphere and broadleaf species (such as eucalyptus, acacia, casuarina, and araucaria) in the Southern Hemisphere.

4. Tropical rainforests which are found in the equatorial regions of Central and South America (for example, Costa Rica, Brazil, and Ecuador); on the west coast of Africa (for example, Congo, Ivory Coast, and Nigeria); and Southeast Asia (for example, Thailand, Malaysia, and Indonesia). They are typified by a rich diversity of broadleaf trees (such as teak, mahogany, lauan, and meranti) and attendant vegetation (shrubs, vines, and other flora) arrayed in a multilayered canopy.

5. Dry forests which occur in the southwestern United States, the Mediterranean region, sub-Saharan Africa, and semiarid regions of Mexico, India, and Central and South America. They are characterized by relatively sparse distributions of pine, juniper, oak, olive, acacia, mesquite, and other drought-resistant species growing in scrub woodland, savanna, or chaparral settings.

6. Mountain forests which are characteristic of mountainous regions throughout the world. The vegetation varies greatly depending on location, but tends to follow an elevational gradient from open woodland at lower elevations, to mixed conifer-broadleaf forest at middle elevations, to relatively pure conifer forest at subalpine elevations; there may be a wet side and a dry side to the mountain chain due to orographic weather patterns, with characteristic flora associated with each.

Characteristics. Although forests take a variety of forms, they have several features in common that allow them to develop in their respective environments.

Resilience. Forests generally contain a broad array of species, each of which is well adapted to the environmental conditions of the region. This biodiversity and adaptability help the forests cope with natural (and in some cases human-caused) forces of destruction, including wildfire, windstorms, floods, and pests. This built-in resiliency also allows the periodic extraction of wood and other products without jeopardizing the long-term health and productivity of the ecosystems—provided such harvesting operations are performed with care.

Succession. Forests are dynamic—they are constantly changing at both landscape and smaller scales of resolution. This natural propensity to change and develop over time is called forest succession. For example, a forest recovering from a wildfire may be dominated by colonizing species such as grasses, forbs, sprouting shrubs, and fast-growing pines during the first few years or decades. As the forest ages, more shade-tolerant, longer-lived species such as oak, maple, or hemlock may gradually encroach and

assume dominance. During the latter stages of forest succession, small patchlike openings may occur in the canopy that allow yet a third wave of plants to become established. In a natural forest, this sequence of events may be repeated in numerous configurations, resulting in a diverse mosaic of vegetation until the next major disturbance occurs. The interval between major disturbances may be as short as a few decades to as long as several centuries. *See* ECOLOGICAL SUCCESSION.

Under managed forest conditions, the process of succession can be modified to favor one stage of forest development over another. In the southeastern United States, it is possible to accelerate the establishment and growth of commercially important pine species and to retard the encroachment of late-successional broadleaf species through a variety of silvicultural practices.

Environmental influences. Forests have a mitigating influence on the environment. This characteristic not only facilitates their own survival and development but also moderates the surrounding climate. Thus, the air temperature in and near forests is lower in the summer and higher in the winter than over open land. Effects of wind and rainfall are ameliorated by the physical structure of forests, thereby reducing wind speed, wind chill, evaporation, soil erosion, soil compaction, and other parameters. Many animal species and humans seek the protection provided by forests in one capacity or another. There also is evidence that removing forests over large areas can affect regional weather conditions.

Ecological processes and hydrologic cycle. Forests play a vital role in ecological processes. From a global perspective, they help convert carbon dioxide in the atmosphere to oxygen, thereby facilitating life for aerobic organisms.

Forests can also capture, store, convert, and recycle a variety of nutrients such as nitrogen, phosphorus, and sulfur. These abilities are based on complex symbiotic relationships among forest plants, microorganisms, soil particles, and other components of the ecosystem. They allow forests to essentially fertilize themselves, thereby maintaining their inherent productivity. This propensity to handle a variety of chemical substances helps forests deal with several industrial pollutants that might otherwise accumulate in the environment to toxic levels. Thus, emissions of sulfur and nitrous oxides can be effectively scrubbed from the atmosphere, provided concentrations are below thresholds toxic to the forest ecosystems themselves. *See* AIR POLLUTION; FORESTRY, URBAN.

Forests also play a critical role in the hydrologic cycle. The canopy of a forest shields the forest floor and soil from the erosive and compacting effects of falling raindrops. In some cases the canopy acts as a large condenser, trapping moisture from fog and low clouds that would otherwise be lost from the area. The spongelike nature of the forest floor acts as a natural filter, cleansing the water and reducing the likelihood of flash floods and erosion. The relatively large storage capacity of most forest soils allows much of the water to be returned to the atmosphere via transpiration from the foliage. *See* FOREST SOIL.

Finally, forests play a crucial ecological role in the habitat that they provide for countless organisms. The need to maintain the health, productivity, and integrity of forest ecosystems is vital to the welfare of humans, now and in future.

Forestry and forest management. The Society of American Foresters defines forestry as the science, the art, and the practice of managing and using for human benefit the natural resources that occur on and in association with forest lands. Natural resources have traditionally entailed major commodities such as wood, forage, water, wildlife, and recreation. However, the concept of forestry has expanded to encompass consideration of the entire forest ecosystem, ranging from mushrooms to landscapes. The practice of forestry requires in-depth knowledge of the complex biological nature of the forest. It also requires an understanding of geology and soils, climate and weather, fish and wildlife, forest growth and development, and social and economic factors.

Foresters, wildlife managers, park rangers, and other natural resource specialists are trained in biology, physics, chemistry, mathematics, statistics, computer science, communications, economics, and sociology. Specialized areas may involve subjects such as plant identification, wildlife biology, forest ecology, forest physiology and genetics, soil science, forest engineering and surveying, hydrology, fire protection, meteorology, biometrics and modeling, wood technology and utilization, and silviculture. Training in forest pest management, wildlife conservation, watershed management, recreation resource management, forest policy and decision making, and business management is crucial.

By combining formal education with practical training and experience, foresters and other natural resource specialists are able to manage forest resources for a variety of purposes. The forest environment can be protected, rehabilitated, or enhanced depending on the circumstances. Effectively manipulating forest stands and landscapes is a prerequisite to sustaining the flow of products, services, and amenities derived from forest ecosystems. *See* FOREST MANAGEMENT.

Silviculture. Silviculture is the art, science, and practice of controlling the establishment, composition, and growth of a forest. It entails the use of both natural and induced processes to foster forest development. For example, reforestation of a harvested or burned-over area can be accomplished by natural seeding from nearby trees or by planting seedlings. *See* REFORESTATION.

Where planting is done, care must be taken to ensure that the trees are adapted to the site. Seed may be collected from mature trees in the area ahead of time. Alternatively, the seed may come from a seed orchard that contains genetically improved trees derived from the general area to be reforested. In either case, the seed are sown in a nursery or greenhouse and carefully cultivated to produce vigorous planting

stock. Survival and early growth of the outplanted seedlings can be enhanced by reducing weed competition and by protecting them from animal pests.

Where natural fires have been excluded from forest ecosystems for long periods of time, it is sometimes necessary to reintroduce fire in order to restore forest health and diversity. Foresters can use prescribed burning as well as carefully contained natural fires to accomplish such objectives. In other instances, low-intensity understory burns can be conducted to stimulate forage production for wild-life and to reduce the risk of catastrophic crown fires. *See* FOREST FIRE.

Several other silvicultural techniques are used to foster stand development and maintain forest health. Periodic thinning of the trees helps maintain stand vigor and promotes rapid growth of the forest. Damage by forest insect pests and diseases often can be prevented or reduced by using direct and indirect control measures. Soil fertility can be enriched by favoring nitrogen-fixing plants that occur naturally in the ecosystem or by applying fertilizers. Trees designated for harvest can be pruned to reduce knots, thereby improving wood quality. Finally, a variety of harvesting practices ranging from clearcutting to selective tree removal can be used to provide wood products, wildlife habitats, stream improvements, recreation areas, and combinations thereof. *See* SILVICULTURE.

Forest ownership and administration. Both public and private entities are involved in forest ownership. Public ownership ranges from small municipal forests to large holdings managed by the state and federal governments. Private ownerships range from individually owned woodlots to large corporate holdings. In the United States, for example, the categories of ownership are nonindustrial private (55%), forest industry (15%), federal government (20%), and state and local government (10%). Ownership can vary dramatically by region. For example, two-thirds of the forest land in the western United States is controlled by the federal government, whereas most of the forest land in the eastern United States is owned by individuals.

The goals of forest owners vary greatly, too. For example, the owner of a small woodlot in New England may manage the land for personal enjoyment and occasional revenue. An agroforester in Africa may be more interested in providing fuelwood, food, and watershed protection for families and livestock. An industrial landowner in Brazil may need to grow short-rotation crops of eucalyptus to supply a local pulpmill. Or, a tree farmer in the Pacific Northwest may produce Christmas trees and sawlogs. National Forests in the Rocky Mountains may serve multiple uses, ranging from outdoor recreation to livestock grazing and timber production. National Parks such as Yellowstone and natural wilderness areas are designed to preserve outstanding geological, biological, and cultural features of the landscape. *See* LANDSCAPE ECOLOGY.

Laws and policies. Regardless of ownership types and goals, the management of forest land in the United States is regulated by numerous laws and policies. Federal agencies must comply with laws such as the National Forest Management Act (1976), the Forest and Rangeland Renewable Resources Planning Act (1974), and the National Environmental Policy Act (1969). Other public and private forest landowners generally must comply with state regulations or guidelines designed to promote sound forest stewardship in their respective regions. In some instances, federal laws may apply to both public and private lands; examples are the Endangered Species Act (1973), Clean Air Act (1968), and Water Pollution Control Act (1972). In other instances, public agencies and private owners may adopt self-imposed standards even more stringent than those required by federal and state law in order to meet special needs. Many other countries have developed or are developing analogous laws, regulations, and guidelines governing their forests.

Sound forest policy is a precursor to the regulatory framework protecting the integrity of forest lands and human communities that depend on them. Policy makers in government, industry, environmental organizations, and the private sector strive to balance the multitude of interests surrounding forest resources. Input from the public as well as resource managers and specialists is a crucial ingredient in the process.

Utilization of forest resources. Forests are often focused on particular uses. For example, plantation forests are generally designed to produce wood and fiber products. Conversely, public forests are increasingly devoted to nonconsumptive purposes such as the preservation of biodiversity, natural conditions, and scenic vistas. However, all forests can provide multiple benefits, including harvestable products, watershed protection, recreation opportunities, wildlife habitat, and ecological services. *See* FOREST GENETICS; FOREST HARVEST AND ENGINEERING.

Sustainable forest management. The 1992 United Nations Conference on Environment and Development—the Earth Summit—brought worldwide attention to the issues of forest management, practices, and government policy. The non-legally-binding "Forest Principles" endorsed by 144 countries at the Earth Summit promote the sustainable management of forestlands, which includes conservation of biodiversity, wildlife, and other important environmental values, and the use of forest products for economic and social development. Following the Earth Summit, both governmental and nongovernmental organizations increasingly focused on forestland uses and forest management practices, recognizing that forestlands must be conserved as forests and that forest management practices must be improved to ensure ecological benefits and sustainability over time.

While sustainable forest management was recognized as the goal, there was no internationally accepted formal or informal definition of it. Groups of countries came together to work on these definitions. As a result, the components or elements of

sustainable forest management have been identified at the international level for the European countries (Pan-European Process); the non-European temperate and boreal forest countries (Montreal Process), including the United States; the tropical countries (the International Tropical Timber Organization); the African countries (African Timber Organization); and others. Currently there is no legally binding international convention or treaty governing forests or forest management policy.

Forest management certification. Programs to certify reforestation success have been in practice for a number of years. However, consumers in a number of European countries raised concerns about the source of wood products, and questioned whether the wood was coming from forests managed on a sustainable basis, especially wood from tropical rainforests. This phenomenon, along with growing public concern in other developed countries about the environmental impacts of forest management practices, led to development by nongovernmental environmental organizations and the timber industry of new programs to promote, evaluate, verify, or certify the management of forestlands on a voluntary basis by landowners. While the early focus of certification efforts was on tropical forests in developing countries, increasingly the focus has turned to programs in temperate developed countries, including the United States, European countries, and Canada.

Forest management standard-setting and certification programs have been developed by nongovernmental organizations both internationally and within the United States. The forest products industry in the United States has implemented the Sustainable Forestry Initiative^{sm} (SFI^{sm}) program, which sets out principles and standards for forest management by member companies and licensees of the American Forest & Paper Association. The Forest Stewardship Council (FSC) is an international organization that has established broad principles and guidelines for forest management, to be supplemented by regionally developed standards. The Canadian Standards Association has developed a program for use in Canada, and the United Kingdom has developed the UK Forestry Standard with the intent to be compatible with a range of other certification programs. Not all of these programs mandate certification of forest management practices. Some allow for verification by first, second, and third parties as desired by the landowner.

Several programs include the option of labeling wood products as certified under a specific program. Labeling of wood products is a complicated process and, in many countries, is regulated by the government. Claims made by the producers of wood or the manufacturers of wood products are varied from the Forest Stewardship Council claim of wood from "well-managed forests" to Indonesian claims of wood from "sustainably managed" forests. *See* FOREST MANAGEMENT; WOOD PRODUCTS. Mary J. Coulombe

Bibliography. J. G. Laarman and R. A. Sedjo, *Global Forests: Issues for Six Billion People*, 1992; D. W. MacCleery, *American Forests: A History of Resiliency and Recovery*, 1993; K. Ramakrishna and G. M. Woodwell (eds.), *World Forests for the Future*, 1993; N. P. Sharma (ed.), *Managing the World's Forests: Looking for Balance Between Conservation and Development*, 1992; G. W. Sharpe et al., *Introduction to Forest and Renewable Resources*, 1995; L. C. Walker, *Forests: A Naturalist's Guide to Trees and Forest Ecology*, 1990; J. Westoby, *Introduction to World Forestry*, 1989; R. A. Young and R. L. Giese (eds.), *Introduction to Forest Science*, 2d ed., 1990.

Forest ecosystem

The entire assemblage of organisms (trees, shrubs, herbs, bacteria, fungi, and animals, including people) together with their environmental substrate (the surrounding air, soil, water, organic debris, and rocks), interacting inside a defined boundary. Forests and woodlands occupy about 38% of the Earth's surface, and they are more productive and have greater biodiversity than other types of terrestrial vegetation. Forests grow in a wide variety of climates, from steamy tropical rainforests to frigid arctic mountain slopes, and from arid interior mountains to windy rain-drenched coastlines. Forests with deciduous and evergreen trees are found in the tropics, in the temperate zone, and in the taiga of high latitudes. Forests are found growing on top of moving glaciers; in areas with little rainfall that survive on impacted fog droplets; and in areas with standing fresh and even salt water. The type of forest in a given place results from a complex of factors, including frequency and type of disturbances, seed sources, soils, slope and aspect, climate, seasonal patterns of rainfall, insects and pathogens, and history of human influence.

Ecosystem concept. Because ecosystem boundaries are arbitrarily set, primarily as a research tool, they can be defined at many scales, from a cup of soil to the entire planet. Often forest ecosystems are studied in watersheds draining to a monitored stream. The structure of forested ecosystems in watersheds is then defined in vertical and horizontal dimensions. Usually the canopy of the tallest trees forms the upper ecosystem boundary, and plants with the deepest roots form the lower boundary. The horizontal structure is usually described by how individual trees, shrubs, herbs, and openings or gaps are distributed. Wildlife ecologists study the relation of stand and landscape patterns to habitat conditions for animals.

Photosynthesis. Photosynthesis is the most important metabolic process on Earth, converting solar energy into the food that fuels the food chain, producing gaseous oxygen (O_2) that animals breathe, and controlling how carbon (C), water, and nutrients are cycled, and even how fast rocks are converted to soil. Photosynthesis is, in effect, the basis of life on Earth. Only autotrophs (green plants and some microbes) can produce chlorophylls, pigments that

capture sunlight and convert it into chemical energy. The energy of the Sun's photons excites these pigments in a way that leads to the conversion of carbon dioxide gas (CO_2) from the air and water from the soil to simple sugars (photosynthates). These sugars sustain plant growth, creating more leaves, branches, wood, roots, and seeds. Heterotrophs are organisms, including herbivores, saprophytes, and carnivores, that rely entirely on energy captured by autotrophs. *See* PHOTOSYNTHESIS.

Structure. Woody trees and shrubs are unique in their ability to extend their branches and foliage skyward—more than 100 m (330 ft) in some conifers—and, as a stand of individuals, to capture carbon dioxide and most of the incoming photosynthetically active solar radiation. Only about half of the Sun's light-energy spectrum falls in the photosynthetically active radiation spectrum (400–700 nanometers); some light is reflected back to the atmosphere and some passes through leaves to the ground (infrared light). High rates of photosynthesis require lots of water, and many woody plants have deep and extensive root systems that tap stored ground water between rain storms. Root systems of most plants are greatly extended through a relation between plants and fungi, called mycorrhizal symbiosis. In one type called ectomycorrhizae, the fungus forms a mantle around small rootlets. Fungal mycelia grow into the root's cortical cells and receive plant sugars. These sugars help the fungus to extend mycelia 10–100 cm or more into the surrounding soil. Because the hyphae (branches of the mycelia) can be as small as 5 micrometers, and roots are typically greater than 60 μm, the mycelia extend the effective volume of soil from which plant roots obtain nutrients and water. *See* ROOT (BOTANY).

Nutrients. Photosynthesis and growth require many different mineral elements as constituents of metab-olites, as enzyme cofactors, and for cell walls and membranes. About 95% of plant dry weight is made of carbon, oxygen, and hydrogen. Among the remaining essential elements, phosphorus is used in electron-transfer reactions; nitrogen is part of all amino acids and proteins; sulfur is part of some amino acids; calcium occurs in cell walls and membranes; magnesium is a constituent of chlorophyll; and potassium is necessary for proper functioning of leaf stomates (pores in leaves). Micronutrient elements (iron, copper, zinc, manganese, molybdenum, cobalt, boron, chlorine, and possibly others) are required as well. Most of these elements are taken up by plant roots from the soil. A small portion of what plants require comes from elements dissolved in rain or blown onto plants as dust. Uptake of nitrogen from the atmosphere by nitrogen-fixing bacteria, and the plants with which they form symbiotic relations, are exceptions.

Biomass and production. The biomass of a forest is defined here as the mass of living plants, normally expressed as dry weight per unit area. Confusion can result from variable definitions that may include the mass of other living organisms and the mass of dead organisms (detritus). Living microbes and insects in soil have a significant weight per unit area, but higher animals usually make up far less than 1% of the total mass. Forests usually have a lot of nonliving biomass, such as soil humus and litter and logs on the ground, and most of the heartwood of large trees is nonliving tissue. Thus, biomass needs to be carefully defined.

Biomass production is the rate at which biomass is accrued per unit area over a fixed interval, usually one year. If the forest is used to grow timber crops, production measures focus on the biomass or volume of commercial trees (one common measure of volume is the board foot (12 × 12 × 1 inches, or 144 cubic inches). Likewise, if wildlife populations are the focus of management, managers may choose to measure biomass or numbers of individual animals. Ecologists interested in the general responses of forest ecosystems, however, try to measure net primary production (Npp), usually expressed as gross primary production (Gpp) minus the respiration of autotrophs (Ra), where Ra can be thought of as the cost of doing business:

$$\mathrm{Npp} = \mathrm{Gpp} - \mathrm{Ra}$$

The Gpp is extremely difficult to measure directly, although eddy correlation gas diffusion techniques are being tried in large towers on a few forests. The most common approach is to measure cumulative accrual of organic matter in all plants per unit area per year. This approach is hindered by the difficulty in measuring plant organic matter accruals belowground. Another response commonly of interest is net ecosystem production (NEP) usually expressed as

$$\mathrm{NEP} = \mathrm{Gpp} - (\mathrm{Ra} + \mathrm{Rh}) = \mathrm{Npp} - \mathrm{Rh}$$

where Rh is respiration of heterotrophs. Pragmatic strategies include measuring changes in soil organic matter and soil respiration, which includes Rh and root Ra. *See* BIOMASS.

Forest productivity. Productivity is the change in production over multiple years. Monitoring productivity is especially important in managed forests to ensure that practices are not undermining the capacity of the forest soil to regenerate and grow subsequent timber crops or to provide animal habitat or other desired goods and services from the forest through time. Changes in forest productivity can be detected only over very long periods. In addition to annual fluctuations in climate, most forest stands take several hundred years to develop through successional stages, and meaningful changes in productivity may occur only after several successional or management cycles. Concerns over long-term productivity, therefore, mostly focus on changes in the status of soil nutrients, soil organic matter, structure, and other factors thought to control biomass production. *See* BIOLOGICAL PRODUCTIVITY.

Cycles. Forested ecosystems have great effect on the cycling of carbon, water, and nutrients, and these effects are important in understanding long-term productivity. Cycling of carbon, oxygen, and

hydrogen are dominated by photosynthesis, respiration, and decomposition, but they are also affected by other processes. Forests control the hydrologic cycle in important ways. Rain falls on forests and drips through the canopy or along tree stems to the soil. Wind-driven fog-water droplets can also be captured by finely divided leaves and needles.

Photosynthesis requires much more water than is required in its products, called photosynthates. Water is lost back to the atmosphere (transpiration) simply to maintain stomates open and to transport nutrients from the soil through the xylem. Water on leaf and branch surfaces also evaporates under warm and windy conditions. Water not taken up or evaporated flows into the soil and eventually appears in streams, rivers, and oceans where it can be reevaporated and moved back over land, completing the cycle.

Forest plants and animals alter soil characteristics, for example, by adding organic matter, which generally increases the rate at which water infiltrates and is retained. Nutrient elements cycle differently from water and from each other.

Rocks already contain most elements required for plant growth, except nitrogen. Elements such as phosphorus, calcium, and magnesium are released from primary minerals in rocks through chemical weathering. Plants can speed weathering, and as a result release nutrients that support plant growth. Elements incorporated into biomass are returned to the soil with litterfall and root death; these elements become part of soil organic matter and are mineralized by decomposers or become a component of secondary minerals. Some elements are held on exchange sites where they can be taken up by subsequent generations of plants. Some elements, such as potassium, calcium, and magnesium, are readily dissolved in water and can move to streams in soil solutions. All elements can leave ecosystems through erosion of particles and then be transported to the oceans and deposited as sediment. Deeply buried sediments undergo intense pressure and heat that reforms primary minerals. Volcanoes and plate tectonic movements eventually distribute these new minerals back to land.

Nitrogen differs by being rare in rocks but the most common gas (molecular N_2) in the atmosphere (78% by volume). Only certain bacteria can form a special enzyme (nitrogenase) which, with substantial consumption of carbohydrates, breaks apart N_2 and combines with photosynthates to form amino acids and proteins. In nature, free-living N_2-fixing microbes and a few plants that can harbor N_2-fixing bacteria in root nodules (called symbiotic N_2-fixing plants, such as legumes and alders) play important roles controlling the long-term productivity of forests limited by nitrogen supply. Because nitrate (NO_3^-) is an easily soluble anion, it moves freely in soil and can be lost by leaching when plant roots are not active. Bacteria that convert ammonium ion (NH_4^+) to NO_3^- (nitrifiers) and bacteria that convert NO_3^- back to N_2 (denitrifiers) are important in nitrogen cycling as well. *See* FOREST SOIL; HYDROLOGY; NITROGEN CYCLE.

Forest disturbance and succession. Changes in the plant species of a forest over 10 to 100 years or more are referred to as succession. Changes in forest structure are called stand development; changes in composition, structure, and function are called ecosystem development. Simplified models of succession and development have been created and largely abandoned because the inherent complexity of the interacting forces makes model predictions inaccurate.

An example model is that the greatest changes occur immediately after the death of the dominant trees through disturbances such as fire, windthrow, flooding, insects, disease, or harvesting. A typical sequence would be: After disturbance, a reorganization phase begins with pioneer species such as herbs, shrubs, and rapidly growing hardwoods, usually along with some plants or sprouts from the previous stand becoming established. This phase is followed by a stem-exclusion or aggradation phase where canopy dominance shifts to trees that grow up through the pioneers' canopy, initially in partial shade. Especially in conifer forests, as these trees dominate and grow, the pioneer species are overtopped and shaded out. Over time, a transition or understory reinitiation phase develops where individual trees die from windthrow, ground fires, or other causes, creating gaps. When gaps between living trees are big enough to allow enough light through, new trees begin to form a secondary, lower canopy. In the so-called old-growth or steady-state phase, stands tend to have scattered large-diameter, old trees forming an upper canopy and one or more lower canopies of younger trees. Standing snags and rotting logs on the ground are also found in most old-growth stands.

Many problems with this general model have been noted. Most important is that disturbances tend to be variable in their nature, extent, and frequency. For example, individual windthrows, ground and crown fires, and harvesting disturbances create nearly unique conditions and result in different species assemblages and growth trajectories. Further confounding the problem, herbivores and diseases may alter species composition, seeds of important species may or may not be present, and disturbances vary unpredictably through time. The result is a complex, and likely unreproducible, set of conditions at any one time. *See* ECOLOGICAL SUCCESSION; FOREST FIRE.

Environmental issues. Forests provide people with many goods and services. Globally, wood is mostly used for heating and cooking. Wood is the dominant structural material for housing and fiber for paper. Alternatives to wood are gaining wider use, often with much greater environmental consequences. Fossil fuel—used in heating and cooking and in the manufacture of concrete, brick, and steel—can replace wood, but it is not a renewable resource. Mining and manufacturing of alternative materials have greater air and pollution side effects than do the forest operations producing a similar quantity of material. Recycling of paper products has greatly

helped to reduce demand for raw materials. Clean water is becoming a dominant forest resource in many places, and the forest's role in cleaning pollutants out of air and water is becoming more widely recognized. As people increasingly populate Earth, forests are also becoming a last refuge to experience nature through hunting and gathering, hiking, fishing, nature appreciation, and spiritual renewal. Many people in cities often support protecting forests for their and their children's future use. Concern over biodiversity currently drives legal strategies to limit extractive uses in the United States and elsewhere.

Ecosystem-based management. Forest managers on United States public lands and around the world have struggled to meet competing and changing societal values and needs. Before 1980, many public forests were systematically converted into managed plantations with the primary goal of producing timber, but not without concerns about other resource values, where possible. As the importance of nontimber values gained public support, intense conflicts emerged. Many public agencies responded by adopting loosely defined concepts of ecosystem management. At its core is the idea that ecosystems are complex assemblages of organisms interacting with their environment and changing in complex ways over time, and that science-based knowledge of how ecosystems work is important to maintaining their biodiversity and long-term productivity. The first step in implementing ecosystem management has often been to reallocate or rezone forests to meet new primary objectives. Concepts of joint production, however, are emerging that attempt to find ways to simultaneously manage for multiple objectives by determining compatible groupings of objectives, where possible.

A complementary concept, disturbance-ecology-based management, is also being tried. This strategy centers on the premise that organisms are more adapted to the historical disturbance patterns than to specific successional states, and that management could more closely emulate natural disturbances and ecosystem responses to disturbance as a way to maintain diversity and long-term productivity, and at the same time permit limited resource extraction. The concept is to manage for a distribution of disturbance types, intensities, and frequencies rather than for static successional stages. To the extent that management-induced disturbances begin to mimic natural disturbances, concerns for maintaining biodiversity (especially of soil species unknown at present) and for maintaining important ecosystem processes can be reduced.

Likewise, globally, loosely defined concepts of sustainable development and ecosystem sustainability have gained acceptance. The principal idea is that although forests should be managed for the welfare of local peoples, management should not preclude similar opportunities for future generations. The interpretation of these general concepts as applied to local situations has frequently proved problematic because people's demands are growing and understanding of ecological limits remains inadequate. *See* GREENHOUSE EFFECT.

Bernard T. Bormann

Streams

One of the products of an undisturbed forest is water of high quality flowing in streams, which are a biologically diverse and productive ecosystem. The ecological integrity of the stream is a reflection of the forested watershed that it drains. When the forest is disturbed (for example, by cutting or fire), the stream ecosystem will also be altered.

Streams are dependent upon the surrounding forested watershed in at least six ways: (1) The nature and extent of vegetation on the watershed alters the amount of water flowing in the channel and the timing and intensity of floods. (2) Uptake of nutrients by the forest vegetation and soils alters the chemistry of stream water and influences the concentrations of nutrients (for example, ammonium, nitrates, and phosphates essential for life). (3) The presence of vegetation alters the amount of light reaching the stream, influencing both stream temperature regimes and algal productivity. (4) The amount of sediment entering the channel reflects watershed activities. (5) Dead wood falling in the stream alters channel form and provides essential habitat for stream fauna. (6) Organic matter produced in the forest falls into the stream and provides a food resource for stream fauna.

Detritus in stream food webs. Primary productivity in streams flowing through forests is often limited by light because of the intense shade provided by the forest. Because of limited availability of organic matter produced within the stream by algae, stream food webs are also dependent on detritus (that is, nonliving organic matter) which comes from the surrounding forest. Detritus enters the food web both as particles (for example, falling leaves) and as substances dissolved in water. The two sources of detritus have somewhat different fates in the stream food web.

When a leaf falls into a stream, its water-soluble constituents are initially leached. It is then colonized by fungi and bacteria that break down the structural matrix. After the leaf has been partly decomposed, it is the preferred food for a group of aquatic insect larvae called shredders because they tear apart the leaves in the process of feeding on them. Shredders are taxonomically diverse and include stoneflies, caddisflies, and craneflies as well as other groups. The feces of shredders are fine particles of organic matter that combine with fine particulate organic matter from other sources and are further colonized by bacteria. Shredders are extremely important in accelerating the breakdown of leaves in streams and as a source of fine particulate organic matter. When shredders are experimentally removed from the stream, leaf decay rates are less and the amount of fine particulate organic matter generated decreases dramatically.

Some of the fine particulate organic matter generated by shredders is transported downstream in the water column (called seston). It serves as the food

source for another group of aquatic insects called filtering collectors. These include insects such as caddisflies that spin silken nets to capture the drifting organic matter as well as any insects that may have been caught in the current. The fine particulate organic matter that remains on the bottom of the stream provides food for a group of insects called collector-gatherers. This group includes taxa like midges and mayflies.

In addition to providing leaves, the surrounding forest is a source of dissolved organic matter to the streams. Dissolved organic matter includes compounds such as dissolved sugars and amino acids, although it is dominated by larger, less biologically available molecules such as humic substances. Dissolved organic matter is material that is leached out of the soil and carried by water to the channel. In the stream, it is taken up by bacteria or is chemically bound to fine particles. Hence some dissolved organic matter becomes fine particulate organic matter and is consumed by the organisms described above. Many bacteria in streams are attached to hard substrates such as rocks or wood and form an active biofilm that contains both algae and bacteria, and extracellular material that cements the entire community on the rock surface. Very small consumers such as protozoa, rotifers, nematodes, and copepods are also a part of the biofilm community. A group of aquatic insects called scrapers feed on this biofilm. Common scrapers include mayflies, beetles, and snails. *See* HUMUS.

The food web in a forest stream is complex. It is supported by small amounts of algae produced in the stream and larger amounts of dissolved and particulate organic matter produced in the forest. The various groups of aquatic insects described above provide the food for predators, which include other aquatic insects such as stoneflies, true bugs, and helgrammites. Vertebrate predators such as salamanders and fishes are also important predators in stream food webs. *See* FOOD WEB.

Woody debris. The surrounding forest contributes structure to the channel in the form of fallen wood that forms debris dams. These debris dams slow the inevitable downstream movement of organic matter, retaining it in the stream so that the fauna can use it as a food resource. Debris dams also create pools in the channel that provide a critical habitat for stream organisms such as fish. Streams in the western United States show increases in stocks of salmon associated with restoring debris dams to forested streams from which wood had previously been removed. In sand-bottom streams, woody debris offers the only available stable substrate and is heavily colonized by aquatic insects.

Flowing water is the most visible component of a stream. Yet there is flowing water that cannot be seen beneath the surface of the sediments, and there is a community of organisms living deep within the sediments. That environment is called the hyporheic zone of (literally, "beneath the flow"). There is exchange of water, nutrients, and fauna between the hyporheic and surface flowing waters. In addition, much of the water entering the stream channel from the watershed flows through the hyporheic zone and, hence, processes occurring in the hyporheic zone influence what is observed in surface stream water.

Watershed disturbance. There are tight linkages between forests and the streams flowing through them. Thus, streams are susceptible to disturbances occurring in the surrounding forest, particularly if the disturbance is in the part of the watershed immediately adjacent to the stream channel (the riparian zone). Forest streams are altered by any practices (for example, lumbering) or chemical input (such as acid deposition) that alter forest vegetation, by the introduction of exotic species, and by the construction of roads that increase sediment delivery to streams. *See* ACID RAIN; STREAM POLLUTION; STREAM TRANSPORT AND DEPOSITION.

Judy L. Meyer

Vertebrates

Forest animals are the consumers in forest ecosystems. They influence the flow of energy and cycling of nutrients through systems, as well as the structure and composition of forests, through their feeding behavior and the disturbances that they create. In turn, their abundance and diversity is influenced by the structure and composition of the forest and the intensity, frequency, size, and pattern of disturbances that occur in forests. Forest vertebrates make up less than 1% of the biomass in most forests, yet they can play important functional roles in forest systems.

Energy flow. In forests, energy used by animals usually is acquired by energy flow through decomposers, because wood that dominates the biomass of forests is largely indigestible by most animals. Fungi and some invertebrates can digest the complex molecules of cellulose and lignins that constitute wood. These organisms provide the basis for food for many forest vertebrates (a detrital-based system). The animals that eat the fungi (fungivores) or invertebrates (insectivores) become food for other predators through a trophic system. Animals use the energy that they gain for survival, movement, reproduction, and in the case of birds and mammals, for maintaining a constant body temperature. As energy is used, it is lost as heat. Consequently, proceeding from one trophic level to the next involves a release of heat and a net loss in the energy available to the next trophic level. Because mass and energy are generally interchangeable, there will be less mass produced per unit of time of high-level carnivores in a system than in the herbivores or lower-level carnivores that they eat.

If a disturbance such as a severe fire occurs in a forest, the early stages of succession may be dominated by foliage rather than wood. At this time, the predominant flow of energy is through the herbivores to the carnivores (a grazing-based system) rather than through the decomposers to the carnivores, unless large amounts of dead wood persist after the fire. As succession proceeds, detrital-based systems again predominate until the next disturbance. Late in ecological succession, tree-fall gaps provide small-scale

disturbances within which both detrital and grazing-based systems can operate. However, both detrital- and grazing-based systems operate to some degree in all stages of ecological succession. *See* ECOLOGICAL ENERGETICS; TROPHIC ECOLOGY.

Nutrient cycling. In some situations, animals can greatly influence the cycling of nutrients within systems as well as the transport of nutrients among systems. For instance, animals that burrow in the soil (for example, moles) mix nutrients throughout the soil layers and influence water and air flow through the soil. Birds that feed in one system, such as agricultural lands, and roost communally in forests (for example, blackbirds) can deposit great amounts of nitrogen to the forest floor through their feces. Although these dramatic examples of nutrient movement within and among patches are rare, some level of nutrient movement occurs at all times.

Animal habitat. Each animal species has its own habitat requirements, and its own geographic range which is structured by global climate patterns and soil types, which influence forest vegetation patterns and subsequently food and cover resources for each species. Some species occur over continents (for example, white-tailed deer), but others live only within a few counties (for example, the del Norte salamander). Within a geographic range, each individual within a population occupies a home range area within which it can secure the resources that it needs to survive. Because mass and energy are interchangeable, animals of large mass generally have larger home ranges than smaller animals because large animals require more food for survival. Carnivores have larger home ranges than herbivores because carnivores must have a large enough supply of herbivores to eat if they are to survive, and because there is decreasing energy available from one trophic level to another. Within the home range of each animal, food and cover resources are usually distributed within patches, so that an animal must move among these patches to secure these resources. Patches can occur horizontally throughout a home range as well as vertically within a forest. Forest-dwelling vertebrates have home ranges that usually are volumes rather than areas. Indeed, it is the vertical distribution and use of resources that can influence the diversity of animals that can occur in a forest.

Plant-animal interactions. As animals move among patches, they take resources with them from one patch to another. For instance, some plants produce fruits or flowers that are attractive to animals (for example, cedar waxwings and hummingbirds), while others may produce seeds that adhere to the fur or feathers of an animal. The animals move the seeds or pollen from patch to patch, allowing plants to disperse to new areas. Similarly, some species of voles eat the fruits of mycorrhizal fungi, below the soil surface. The voles disperse the ingested fungal spores by depositing their feces in their underground runways. This allows the fungi to spread to new areas. These species of fungi grow on the roots of some trees and shrubs, thereby enabling the plants to be more efficient in acquiring nutrients from the soil. Not all plants rely on animals for dispersal, however; some seeds disperse from wind or water. *See* POPULATION DISPERSAL.

Although animals may influence forest structure and composition through their movement of nutrients and plant propagules, they also can play a role as a form of forest disturbance. Herbivores that select particular plant species for food can nearly eliminate these plant species from systems that the plant would dominate if the herbivores were not present. In the eastern United States, the grazing habits of white-trailed deer can greatly influence the structure and composition of the deciduous forests. Another animal that alters the structure and function of a forest system is the beaver. Beavers change plant species composition by selectively feeding on certain plant species and by altering water levels of forest streams. In addition, they change flowing streams into ponds which cause sediments to accumulate upstream. Continual sediment accumulation causes the pond to fill, eventually forming meadow. As a result, a different animal community will occupy a beaver-influenced forest stream system. Species such as beaver that constitute a small part of the biomass of the system but greatly influence its structure and function are called keystone species.

Some plant species have defense mechanisms against herbivory by animals, and so can persist in a forest. Spines and thorns are physical deterrents against some large mammals. Other plants produce chemicals in their foliage that either can be distasteful (such as saponins) or can decrease the availability of certain ingested nutrients to the herbivores (such as phenols). However, black-tailed deer have a protein in their saliva that can bind with the phenols, inactivating the chemical and allowing access to the nutrients in the plant. Other animals may not contain this protein. The production of phenols in plant foliage in the Pacific Northwest seems generally to increase with increasing amounts of light available to the plants, and therefore can be influenced by disturbance in the forest. *See* HERBIVORY; PLANT-ANIMAL INTERACTIONS.

Forest disturbance. Physical disturbances to forests, such as fire and wind, can influence the structure and composition of the animal community in the system. An animal community is a collection of populations of species that occur in an area. The size, frequency, severity, and pattern of disturbances in a forest can influence each species in a forest differently. If disturbances are small, are infrequent, cause little change, and are widely dispersed (such as an individual tree fall), the effect of each disturbance on a species with a large home range (for example, deer) is probably minimal. However, that same disturbance may greatly influence resource availability for a species with a small home range, such as a salamander that may use the fallen tree as cover. The accumulation of many tree-fall gaps throughout the home range of a species such as an elk can improve food and cover availability for that species.

Frequent disturbances reduce the opportunity for recovery of the system and hence the predominance of grazing- rather than detrital-based energy flow through the system. Disturbance severity can also influence the availability of dead wood or shrubs persisting after the disturbance, and hence opportunities for food or cover for various species. For example, dead trees (snags) that persist following a fire are used by cavity-nesting birds and mammals which would not inhabit that area otherwise.

Global patterns of environmental stability and disturbance frequency seem to influence the diversity of animals. The most diverse animal communities occur in moist tropical forests where severe disturbances are very infrequent, plant species diversity is very high, and forests are tall and provide a large volume of habitat for different species to occupy. Such environmentally stable systems provide long periods of time (many generations) for plants and animals to coevolve. Moving toward the Poles and from hardwood to conifer forests, environmental stability declines (for example, stronger seasonal fluctuations), disturbance severity, frequency, and size may increase (for example, glaciation or fires), and plant species diversity is reduced. There generally is a decrease in animal diversity from the Equator to the Poles, dependent on forest site productivity and plant diversity.

Forest management. Humans use forests for a variety of reasons and so can alter habitat quality for some animal species, often by altering the characteristics of forest disturbance. Disturbances imposed by humans that are more frequent or severe than those which occurred in the past and to which animals are adapted, can lead to decline and endangerment of species in the area. For example, even-aged forest management that completely removes all standing living and dead trees usually is more severe than a disturbance caused by fire or wind. The living and dead trees that remain following a fire or wind disturbance are used by a wide variety of species that otherwise would be absent. Management that retains some living and dead trees may allow some of these species to persist. Similarly, very frequent cutting of forests may not allow the forest to recover to a point that provides some of the large dead trees and logs and fungi necessary for some species to occur. A less frequent harvest may be preferable for these species. Finally, deforestation for agricultural purposes can cause a much more severe impact on species persistence within a region than forest management where stands are allowed to regrow.

Forests managed for timber production, if consideration is given to severity, frequency, size, and pattern of cutting, may allow each species some habitat in a region at all times and opportunities for movement among patches as some stands are cut and others are regrown. Practices that do not provide sufficient patch sizes (home ranges or more), or that are isolated from one another for long periods, such as conversion of forests to urban or agricultural areas, can result in decline or elimination of species from large areas and replacement of these species with others adapted to the new habitat that is created. All forests, fields, and cities are habitats for some species; however, it is the amount and distribution of each of these that will allow species to persist over time and continue to function as a part of these ecosystems. *See* FOREST MANAGEMENT.

Invertebrates

Invertebrates are major components of forest ecosystems, representing most of the biological diversity and affecting virtually all forest processes and uses. Many species are recognized as important pollinators and seed dispersers that ensure plant reproduction. However, some invertebrates that feed on plants and decaying organic matter have been viewed as detrimental to tree growth or commercial production of forest products. Nevertheless, there are indications that even so-called pests may be instrumental in maintaining ecosystem processes critical to soil fertility, plant productivity, and forest health. Consideration of invertebrate diversity and the variety of roles that invertebrates play in forest ecosystems is essential to balanced assessments of their effects and management of forest resources.

Roles. Invertebrates affect forests primarily through the processes of herbivory and decomposition. They are also involved in the regulation of plant growth, survival, and reproduction; forest diversity; and nutrient cycling. Typically, invertebrate effects on ecosystem structure and function are modest compared to the more conspicuous effects of plants and fungi. However, invertebrates can have effects disproportionate to their numbers or biomass. Litter-feeding invertebrates are instrumental in fragmenting organic matter and infusing it with fungi and microorganisms that hasten its decay. As few as 10,000 southern pine beetles (a tiny phloem-feeding insect) can kill an apparently healthy mature pine tree; some seed predators can destroy virtually the entire reproductive effort of a particular plant species.

Changes in population size also affect the ecological roles of invertebrates. For example, small populations of invertebrates that feed on plants may maintain low rates of foliage turnover and nutrient cycling, with little effect on plant growth or survival, whereas large populations can defoliate entire trees, alter forest structure, and contribute a large amount of plant material and nutrients to the forest floor. Different life stages also may represent different roles. Immature butterflies and moths are defoliators, whereas the adults often are important pollinators.

Populations of a particular invertebrate species fluctuate as environmental conditions vary seasonally or annually. Typically, invertebrate populations are small (especially in diverse forests), but some species can increase rapidly under favorable conditions and affect forest structure and function dramatically.

Climatic conditions. Herbivorous invertebrates are affected by climatic conditions and by changes in plant abundance or nutritional quality. All plants produce various physical and chemical defenses (such as

tough leaves and toxic compounds) that deter or reduce feeding by herbivores overall, but volatile compounds often attract adapted herbivore species. Differences in chemical composition distinguish plant species and are responsible for preferences by herbivores adapted to these defenses. Chemical composition can vary in a particularly plant as a result of changes in soil moisture or fertility or as a result of injury or crowding. Stressed plants tend to reduce production of defensive chemicals in order to meet more immediate metabolic needs, thereby becoming more susceptible to herbivores. Densely spaced plants of a particular species also are more vulnerable to increased populations of herbivores because of the ease of moving from one host to the next.

Most soil and litter invertebrates depend on warm, moist conditions for survival and may virtually disappear during dry seasons or years, or as a result of canopy removal and exposure of the forest floor to solar heating. Because dead organic matter often contains plant defensive compounds and usually is nutritionally poor (especially in nitrogen), detritivores often have developed mutualistic associations with microorganisms that enhance the nutritional quality of litter resources.

Plant growth, survival, and reproduction. Elevated populations of plant-eating invertebrates can reduce plant growth, survival, and reproduction dramatically. Reduced growth can result from the conspicuous loss of foliage to defoliating species, including moths or sawflies; from less conspicuous loss of buds, young shoots, or roots to budworms or shoot and root borers; and from the loss of soluble photosynthetic compounds to sap-sucking species, such as aphids, spider mites, and root nematodes. Some plant-eating invertebrates are also vectors for plant pathogens, including viruses and bacteria transmitted by sap-sucking invertebrates, and pathogenic fungi and nematodes transmitted by bark- and wood-boring beetles. Severe plant injury or disease often results in death of the host plant. In sufficient numbers several bark beetle species are capable of killing relatively healthy conifers by introducing pathogenic fungi to exhaust tree defenses.

Plant reproduction is affected by invertebrates involved in pollination, seed dispersal, or seed predation. Insects are primary pollinators and seed dispersers for many trees and understory plants, especially in the tropics, where a greater diversity of plants requires more efficient pollen and seed transport than wind. Many ant species are primarily seed predators whose moist, nutrient-rich storage chambers provide ideal locations for seed germination. Seedling growth on abandoned ant colonies typically is more rapid than on surrounding areas. However, seed predation can often destroy the entire seed crop for a particular plant species.

Although the ability of some plant-eating invertebrates to interfere with production of forest products has been the basis for past suppression efforts in forests, growth loss and death of plants are not inevitable results of feeding by invertebrates. Surviving plants often show enhanced growth. Defoliated trees may grow faster for 20–30 years following defoliation than do nondefoliated trees.

This compensatory growth may reflect improved water or nutrient conditions. For example, improved water balance and increased nutrient availability enhance plant survival during drought. Canopy opening increases precipitation penetration through the canopy and reduces evapotranspiration. In some instances, the amount of water reaching the forest floor during the dry season doubles at 20% defoliation, compared to nondefoliated trees. Foliage loss also reduces water use by drought-stressed trees. Whereas desert plants often shed foliage at the onset of dry conditions, water-stressed forest trees may benefit from foliage removal by invertebrates. Alleviation of drought or nutrient stress through defoliation also has been shown to reduce plant vulnerability to pathogens that infect stressed trees.

Litter invertebrates also contribute to growth of forest plants. Worms and other burrowing invertebrates are critical to nutrient release from litter and to the porosity of woody litter and soil. Porous woody litter and soil facilitate root penetration and water-holding capacity.

Forest structure. Plant-eating invertebrates influence forest structure through selective herbivory on abundant or stressed hosts. Suppression of an abundant or stressed plant species opens the canopy and releases space and resources for nonhost plant species or plant life stages, often representing earlier successional stages. Because different plant species use water and nutrients differently, a diverse vegetation competes less for resources and maximizes the probability that sufficient plants will survive disturbance to maintain canopy integrity and ecosystem function.

Plant-eating invertebrates often interact with fire to facilitate nutrient turnover (especially on more arid regions), and maintain vegetation structure and diversity, often by promoting plants tolerant of drought and fire (**Fig. 1**). Fire suppression in managed forests is associated with dense, stressed trees vulnerable to outbreaks of plant-eating invertebrates. Defoliation or plant mortality stimulates turnover of plant tissues and increases fuel accumulation, leading to increased likelihood of fire, and ensures survival of the fire-adapted tree species.

Invertebrates may affect local and regional climates through their effects on forest structure. Forests with closed canopies create relatively moderate environmental conditions because of shading and interception of precipitation and wind. Defoliation or tree mortality affects climate by altering albedo (reflectance), increasing soil temperature, reducing evapotranspiration and relative humidity, increasing canopy surface roughness and wind turbulence, and stimulating emission (through decomposition) of water, carbon dioxide, and nitrous oxide (major greenhouse gases) to the atmosphere. *See* CLIMATE HISTORY; PLANT RESPIRATION.

Nutrient cycling. Invertebrates contribute substantially to nutrient cycling processes critical to soil

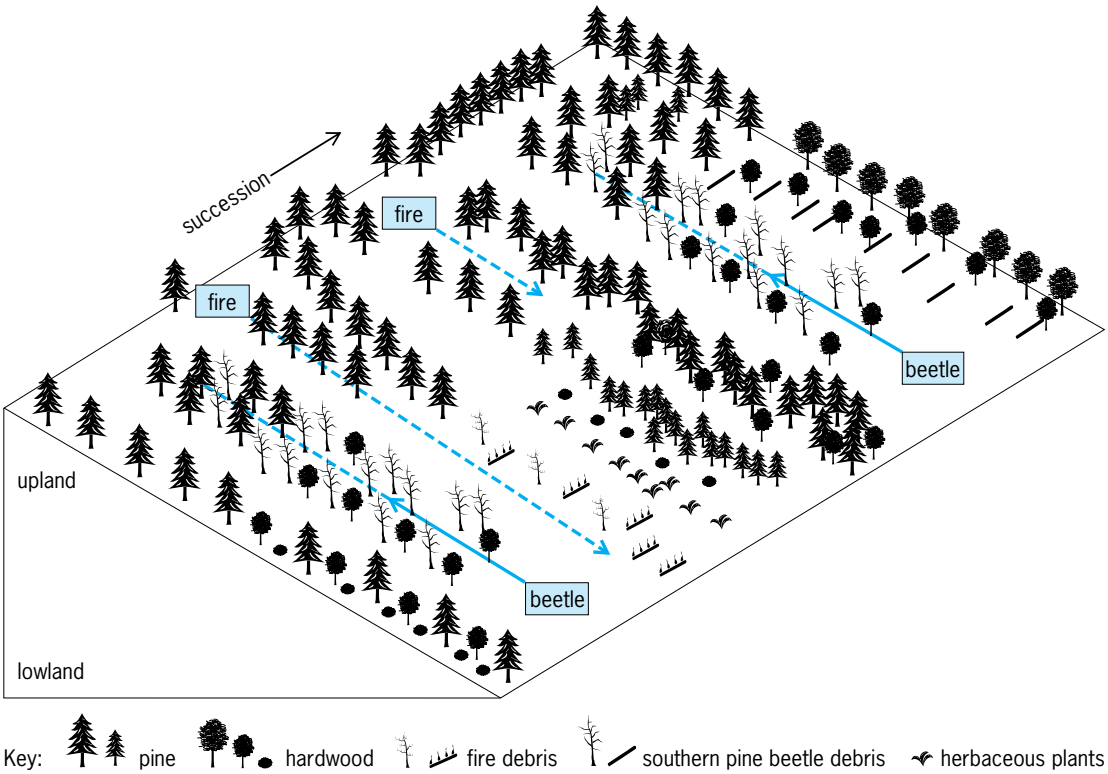

Fig. 1. Southeastern coniferous forest (upland and lowland), illustrating interactions between the southern pine beetle and fire. (*After T. D. Schowalter, R. N. Coulson and D. A. Crossley, Jr., Role of Southern pine beetle and fire in maintenance of structure and function of the southeastern coniferous forest, Environ. Entomol., 10(6):821–825, 1981*)

fertility and forest productivity (**Fig. 2**). First, herbivorous invertebrates are major regulators of nutrient turnover from plant biomass. Pruning and thinning are established horticultural methods for promoting plant growth and fruit production. Invertebrates similarly stimulate plant growth by removing plants or plant parts and thereby reducing competition for limited plant resources. Consumption and turnover of plant tissues throughout the growing season provides a more constant flow of organic matter and nutrients to the forest floor, compared to bulk litterfall prior to cold or dry seasons. More constant organic matter and nutrient supply provides resources for organisms in soil and litter and contributes to soil fertility. Partially digested plant material passed through herbivore guts is fragmented and infused with enzymes and microorganisms that facilitate chemical breakdown. Plant defensive compounds may be partially or wholly metabolized, enhancing subsequent decomposition on the forest floor. Defoliation can stimulate microbial nitrogen fixation and nitrification processes on the forest floor, processes critical to supply of this important nutrient. *See* NITROGEN CYCLE.

Invertebrates that feed an wood, especially beetles, are instrumental in initiating decomposition and nutrient turnover from dying trees and woody litter by penetrating bark and inoculating wood with saprophytic and nitrogen-fixing microorganisms. The winding galleries of wood-eating beetles and termites ensure rapid inoculation of microorganisms throughout logs. Termites host nitrogen-fixing bacteria or protozoa in their guts and increase the availability of nitrogen in the vicinity of their colonies.

Comminution (fragmentation) of dead plant or animal matter is a necessary precursor to litter decomposition and nutrient release (mineralization). A diverse array of saprophagous soil and litter invertebrates are responsible for the fragmentation and infusion of ingested food by microorganisms. Litter usually has lower nutritional value than living tissues because of the retrieval by plants of essential nutrients, especially nitrogen. Therefore, decomposition usually does not progress substantially until litter has been preconditioned by passage through invertebrate guts. Large invertebrates such as millipedes and annelid worms are responsible for fragmenting large litter material; smaller invertebrates such as oribatid mites and springtails process finer material. Most of these invertebrates also are coprophagous, reingesting fecal material after microbial enrichment.

Invertebrates also concentrate major nutrients several orders of magnitude above plant and soil-litter concentrations. Defoliators are particularly rich sources of potassium, calcium, and magnesium, and termites are rich in nitrogen. These nutrients become available to other organisms more readily than do nutrients in plant tissues.

The process of ecosystem recovery from disturbance, as affected by insects and pathogens, also contributes to nutrient balance in forest ecosystems.

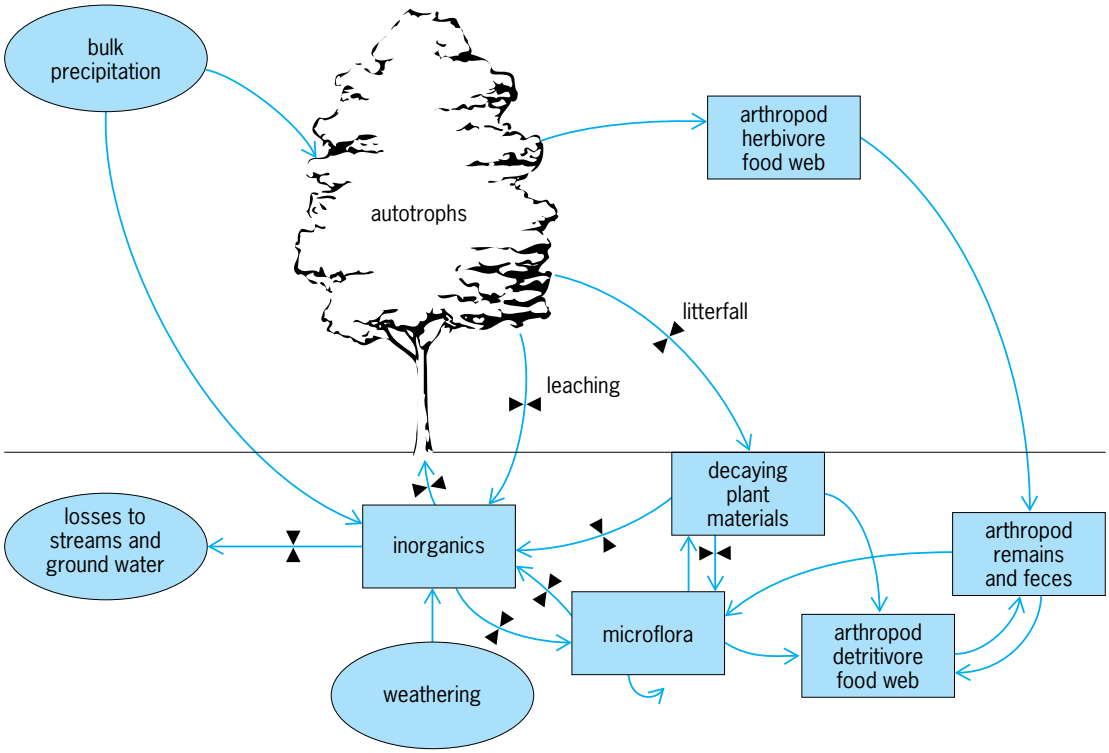

Fig. 2. Conceptual model of elemental cycling in a terrestial ecosystem, emphasizing the presence and activities of arthropod consumers. The indirect regulation of elemental movement by arthropods is indicated by the hourglass-shaped valves on these flows. (*After T. R. Seastedt and D. A. Crossley, Jr., The influence of arthropods in ecosystems, BioScience, 34(3):157–161, 1984*)

Recovery is facilitated by fast-growing early successional species that incorporate nutrients into biomass. As these species grow and later successional species become established, crowding leads to competitive stress, eventually triggering population outbreaks of plant-eating invertebrates and pathogens on the most stressed plants. Thus, invertebrates are instrumental in accelerating the successional transition to more tolerant, nutrient-conserving species, as well as opening the canopy (often over large areas) and promoting earlier successional species. Tim Schowalter

Microorganisms

Microorganisms, including bacteria, fungi, and protists, are the most numerous and the most diverse of the life forms that comprise any forest ecosystem. Their size, prolific reproduction, and varied nutritional strategies allow them to colonize most surfaces above and below ground. The structure and functioning of forests are dependent on microbial interactions. Four processes are particularly important: nitrogen fixation, decomposition and nutrient cycling, pathogenesis, and mutualistic symbiosis.

Nitrogen fixation. Nitrogen fixation is crucial to forest function. While atmospheric nitrogen is abundant, it is unavailable to trees or other plants unless fixed, that is, converted to ammonia (NH_4), by either symbiotic or free-living soil bacteria. Symbiotic nitrogen fixation takes place in root nodules on leguminous plants such as acacia, and on a few other plants including, in the forest, alder (*Alnus*) and the shrub *Ceanothus*.

Lichens are a symbiotic association between algae or cyanobacteria and fungi. Lichens with cyanobacteria fix nitrogen and provide a significant portion of the nitrogen input to old-growth coniferous forests, for example. They are very sensitive to air pollutants and are used to monitor air quality. However, their diversity and abundance decreases with proximity to sources of air pollution. *See* AIR POLLUTION; NITROGEN FIXATION.

Decomposition and nutrient cycling. Most microorganisms are saprophytic decomposers, gaining carbon from the dead remains of other plants or animals. In the process of their growth and death, they release nutrients from the forest litter, making them available once again for the growth of plants. Their roles in carbon, nitrogen, and phosphorus cycling are particularly important. Fungi are generally most important in acid soils beneath conifer forests, while bacteria are more important in soils with a higher pH. Bacteria often are the last scavengers in the food web and in turn serve as food to a host of microarthropods (**Fig. 3**).

Microorganisms reduce the mass of forest litter and, in the process, contribute significantly to the structure and fertility of soils as the organic residue is incorporated. Basidiomycete fungi are particularly important because they possess enzyme systems that allow them to obtain energy from decomposition of the complex carbohydrates (cellulose) and phenolic

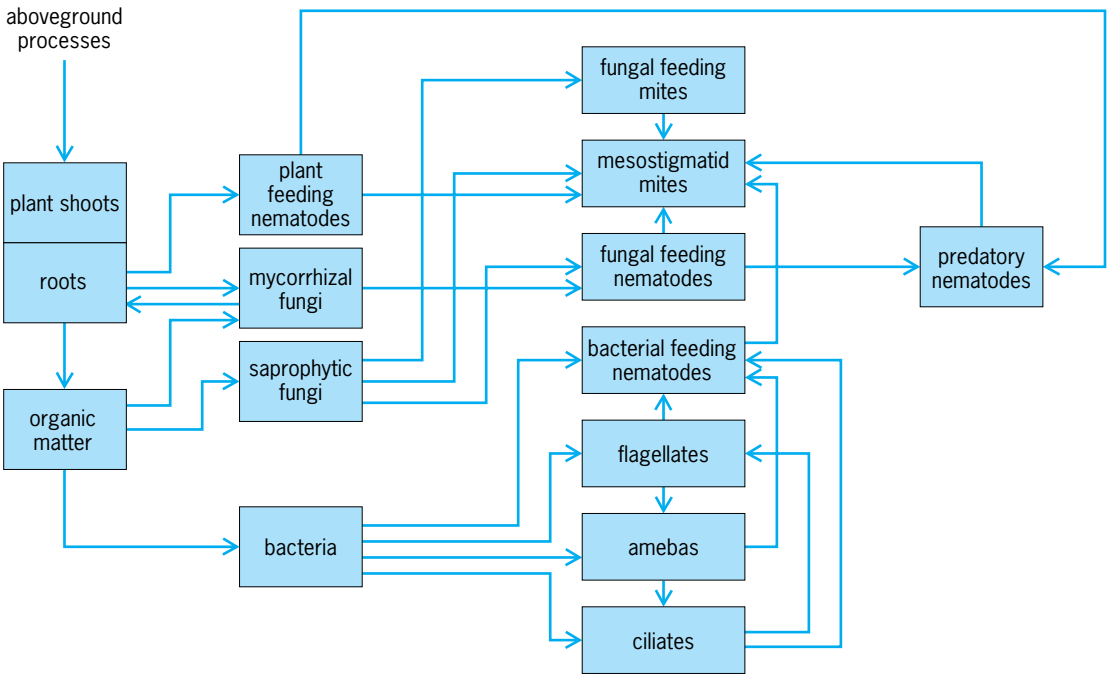

Fig. 3. Belowground food web. (*Courtesy of Elaine Ingham*)

compounds (lignin) of wood. Some fungi that can decompose lignin (white rot fungi) are used to remove or detoxify a variety of organic wastes in industrial processes.

Plant pathogens. Some bacteria and many fungi are plant pathogens, obtaining their nutrients from living plants. The species of plant-pathogenic fungi are specialized in many ways. Some are opportunists, successful as saprophytes, but capable of killing weakened or wounded plant tissues. Others require a living host, often preferring the most vigorous trees in the forest. Pathogenic fungi usually specialize on roots or stems or leaves, on one species or genus of trees. Some pathogens kill their hosts outright, but most attack only a portion of the tree, perhaps making it vulnerable to some secondary agent such as a bark beetle. Plant pathogens gain entry through wounds or by forcing or dissolving their way through the natural barriers of bark and cuticle. They secrete toxins that kill cells, allowing nutrients to escape and be absorbed, and enzymes that dissolve cell walls and allow further colonization by the fungus. Their strategies are diverse, but all benefit by colonizing living plant tissues, before the numerous saprophytes can become established.

Pathogenic fungi are important parts of all natural forest ecosystems. The forest trees evolved with the fungi, and have effective means of defense and escape, reducing the frequency of infection and slowing the rates of tissue death and tree mortality. However, trees are killed, and the composition and structure of the forest is shaped in large part by pathogens.

Pathogens remove weak or poorly adapted organisms from the forest, thus maintaining the fitness of the population. Decay fungi that kill parts of trees or rot the heartwood of living trees create an essential habitat for cavity-nesting birds and the other animals dependent on hollow trees.

By killing trees, pathogens create light gaps in the forest canopy. The size and rate of light gap formation, whether from the death of individual trees by a canker fungus or from the death of expanding groups of trees by a root rot pathogen, and the relative susceptibility of the tree species present on the site determine the ecological consequences of mortality. Succession is often advanced as shade-tolerant trees are released in small gaps. For example, in the Douglas-fir forest the late successional hemlock can reach the canopy only when the seral fir is killed, usually by root rot and the associated bark beetles. Gaps allow the growth of herbaceous plants in the island of light, creating habitat and food diversity for animals within the forest. In many forests, pathogens are the most important gap formers and the principal determinants of structure and succession in the long intervals between stand-replacing disturbances such as wildfires or hurricanes. *See* PLANT PATHOLOGY.

Mycorrhizae. The fungus roots of trees, and indeed most plants, represent an intimate physical and physiological association of particular fungi and their hosts. Mycorrhizae are the products of long coevolution between fungus and plant, resulting in mutual dependency. Two main types of mycorrhizae are known, representing distinct groups of fungi and host plants.

Ectomycorrhizae are composed of a fungus sheath around the outside of root tips, with individual hyphae penetrating between the cortical cells of the root to absorb photosynthates. Outside the root, the hyphae extend into the soil matrix, absorbing and transferring water and nutrients to the tree. In endomycorrhizae (also called vesicular-arbuscular

mycorrhizae), the fungus colonizes inside the roots, forming distinctive absorptive structures within cortical cells. Again the hyphae extend into the surrounding soil.

Mycorrhizae are particularly important to trees because they enhance the uptake of phosphorus from soils. This mineral is relatively immobile in the soil and soon is depleted in the immediate vicinity of growing roots. Mycorrhizal fungi greatly extend the absorptive surface of roots through the network of external hyphae. *See* ECOSYSTEM; FOREST AND FORESTRY. Everett Hausen

Bibliography. J. R. Barnes and G. W. Minshall (eds.), *Stream Ecology*, 1983; F. H. Bormann and G. E. Likens, *Pattern and Process in a Forested Ecosystem*, 1979; G. C. Carroll and D. T. Wicklow, *The Fungal Community: Its Organization and Role in the Ecosystem*, 1992; C. C. Grier, K. J. Elliott, and D. G. McCullough, Biomass distribution and productivity of *Pinus edulis-Juniperus monosperma* woodlands of north-central Arizona, *For. Ecol. Manag.*, 50:331–350, 1992; R. T. T. Forman and M. Godron, *Landscape Ecology*, 1986; M. L. Hunter, *Wildlife, Forests, and Forestry: Principles of Managing Forests for Biological Diversity*, 1990; J. P. Kimmins, *Forest Ecology*, 1987; J. L. Meyer, A blackwater perspective on riverine ecosystems, *BioScience*, 40:643–651, 1990; C. D. Oliver and B. C. Larson, *Forest Stand Dynamics*, 1990; T. D. Schowalter, W. W. Hargrove, and D. A. Crossely, Jr., Herbivory in forested ecosystems, *Annu. Rev. Entomol.*, 31:177–196, 1986; J. A. Standford and J. V. Ward, The hyporheic habitat of river ecosystems, *Nature*, 335:64–66, 1985; E. O. Wilson, *The Diversity of Life*, 1992.

Forest fire

The term wildfire refers to all uncontrolled fires that burn surface vegetation (grass, weeds, grainfields, brush, chaparral, tundra, and forest and woodland); often these fires also involve structures. In addition to the wildfires, several million acres of forest land are intentionally burned each year under controlled conditions to accomplish some silvicultural or other land-use objective or for hazard reduction (**Fig. 1**).

Fig. 1. Controlled fire conditions. (*a*) Closed jack pine (*Pinus banksiana*) cones, which require heat in order to open. (*b*) Prescribed burning after seed tree cutting so that jack pine cones can open for natural regeneration. (*Photographs by W. R. Beaufait*)

Nature of fire. Fire is a natural phenomenon that reverses the creative process of photosynthesis, which produces all wood fiber and other plant materials. In simple terms, photosynthesis uses solar energy to turn carbon dioxide and water into cellulose (woody fiber) and oxygen: $CO_2 + H_2O +$ solar energy $\rightarrow$ cellulose + O. When the temperature near woody or other plant material is raised to the kindling point by a lighted match, burning cigarette, or other means, combustion takes place. As the material burns, carbon dioxide and water are released, along with heat which can spread the fire to nearby combustible plant materials. If dead woody or plant material does not get burned, it decays. In decay the same process produces the same end products of carbon dioxide, water, and heat at a much slower pace. The heat of oxidation is usually not noticeable, and the debris of decaying organic material accumulates to form a thick layer, as on the forest floor. During dry seasons this layer may dry out to such an extent as to be a serious potential fire hazard. *See* COMBUSTION.

Forest fire causes. Most wildfires are caused by human beings, directly or indirectly. In the United States less than 10% of all such fires are caused by lightning, the only truly natural cause. In the West (the 17 Pacific and Rocky Mountain states) lightning is the primary cause, with smoking (cigarettes, matches, and such) the second most frequent. Combined they account for 50 to 75% of all wildfires. In the 13 southern states (Virginia to Texas) the primary cause is incendiary. This combined with smoking and debris burning make up 75% of all wildfires. Miscellaneous causes of wildfires are next in importance in most regions. The other causes of wildfires are machine use and campfires. Machine use includes railroads, logging, sawmills, and other operations using equipment.

Fire behavior. The manner in which fuel ignites, flame develops, and fire spreads and exhibits other phenomena constitutes the field of fire behavior. Factors determining forest fire behavior may be considered under four headings: attributes of the fuel, the atmosphere, topography, and ignition.

Fuel. Live and dead vegetation in the natural forest complex constitutes the fuel for forest fires. The amount of moisture in relation to the heat content of the fuel determines its flammability. Moisture in living vegetation varies with species and with time of year. The moisture content of dead fuels, the more important component of the fuel complex, responds to precipitation, relative humidity, and to a lesser extent temperature. It shows important daily fluctuations. *See* WOOD PROPERTIES.

The rate of combustion and the heat output are also determined by the quantity of fuel, the fuel arrangement, and the thermal and chemical properties of individual fuel particles.

Atmosphere. Major direct effects of the atmosphere on fire behavior are through variations in oxygen supply and in flame angle caused by wind, and the stability of the atmosphere. Secondary effects on

Fig. 2. Characteristic fire behavior. (*a*) Crown fire. (*b*) Surface fire. (*U.S. Forest Service*)

combustion rate are through air and fuel temperature, and the relative humidity of the air. *See* ATMOSPHERE; HUMIDITY.

Topography. Heated air adjacent to a slope tends to flow up the slope, creating a chimney effect. Unburned fuels above a fire advancing up a slope are closer to approaching flames because of the angle of the slope. Both of these effects act similarly to wind on flat ground in increasing the rate of fire spread.

Ignition. An igniting agent of sufficient intensity and duration is necessary to start a forest fire. The agent may be human-caused, such as a match or lighted cigarette, or natural, such as lightning. Once a fire has started, its spread is by successive ignition of adjacent unburned fuel. This fuel is brought to ignition temperature by heat radiated and convected from the flame front. Generally fuels must be very close to, or even enveloped by, flame or superheated gases in the convection column before igniting. This mechanism results in more or less continuous advancement of a fire. Fire spread may also be caused by spotting, that is, ignition by burning brands carried outside the fire area by wind and turbulence.

Types. A forest fire may burn primarily in the crowns of trees and shrubs—a crown fire (**Fig. 2***a*); primarily in the surface litter and loose debris of the forest floor and small vegetation—a surface fire (Fig. 2*b*); or in the organic material beneath the surface litter—a ground fire. The most common type is a surface fire.

Indexes. The U.S. Forest Service has developed a National Fire Danger Rating System (NFDRS) to provide fire-control personnel with numerical ratings to help them with the tasks of fire-control planning and the suppression of specific fires. The system includes three basic indexes: an occurrence index, a burning index, and a fire load index. Each of these is related to a specific part of the fire-control job. These indexes are used by dispatchers in making decisions on setting up fire-fighting forces, lookout systems, and so forth.

Seasons. Because of annual cycles in precipitation and in growth of vegetation, most sections of the United States have a well-defined fire season during which most forest and wild-land fires occur. This fire season varies widely in duration, from two short seasons in the northeastern states, in spring and fall, to frequently a year-long season in southern California. **Figure 3** shows the normal peak fire seasons for the United States.

Prevention of wildfire. This means keeping all unwanted human-caused wildfires from starting. Actually prevention covers a wide range of activities, and must be patterned to the causes of fires in a particular area. The best tools for prevention are education, elimination of hazards, and fire law enforcement. Risks refer to the presence of people and to their activities, such as cigarette smoking and starting brush fires. Hazards are the various types, conditions, volumes, and locations of fuels, such as grass, brush, and forest fuels.

In areas of intensive use, hazard reduction includes many forestry activities aimed at reducing fuel accumulations in order to reduce the channel of a starting fire or to reduce the severity of one that might be started. These activities include roadside burning, slash disposal, controlled burning of forest areas to reduce fuel accumulations, and some silvicultural practices. *See* SILVICULTURE.

Prevention activities should be conducted in line with a carefully planned and written program. The

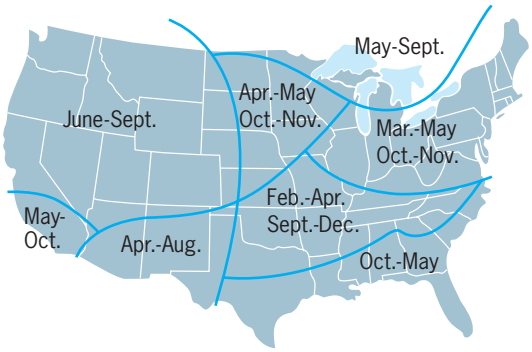

Fig. 3. Normal peak fire seasons in United States.

keystone of such a program is a wildfire prevention plan that includes at least the following elements: (1) basis of the plan (consisting of maps and statistics covering past fires, maps showing fire risks, industrial operations, and hazards, and maps showing signs, radio and television stations, and newspaper offices); (2) summary of objectives, problems, and action to be taken; (3) prevention contact plan; (4) public education; (5) closures and restrictions; (6) industrial operations; (7) reduction of physical hazards; (8) sign-posting plan; (9) prevention training; and (10) legislative needs and law enforcement.

Detection. Although many forest fires are detected and reported by local residents and transient forest users, primary reliance for detection is placed on specialized detection systems. Most fire-control agencies use a combination of lookout towers with trained observers during the fire season and regularly scheduled aerial patrols. Use of aircraft for routine detection patrol has increased greatly, paralleling the increase in aircraft use in other protection activities.

Infrared scanners developed originally by the military for battlefield applications have been adapted for use in detecting forest fires. Aircraft-mounted detectors are routinely used to detect fires started by intense lightning storms. Such scanners have the capability to detect even campfires from an altitude of 12,000 ft (3.7 km). Other electronic equipment, including television, radar, and sferics (to detect lightning discharges and thus storms), has been used experimentally.

Communication. Rapid and accurate communication is essential not only to the detection network but also to the forces engaged in actual suppression work. Land-line telephones are being rapidly supplanted or supplemented by radios. In many organizations all motorized units are equipped with two-way radios. As the size of portable, self-contained equipment is reduced, more and more supervisory personnel on actual fires are also being equipped.

Presuppression. In spite of careful plans and efforts to prevent wildfires, they do occur. Then they must be detected and reported promptly to the suppression forces responsible. These will be prepared to act if adequate suppression steps have been taken. Aside from prevention, the steps include all the preparation, organization, development, and maintenance of equipment, as well as planning, cooperation, mutual aid arrangements, and training.

Generally, public agencies are responsible for presuppression planning and suppression action. State and federal forestry or forest fire organizations have a clearcut understanding and lines of authority as to which areas are their primary responsibility. Each should have a well-defined organizational structure and a specific plan to prevent fires and to detect and suppress any fires that do occur. This plan includes maps of the type described for a fire-prevention program, an inventory of equipment, and a program of training and a program for securing additional personnel equipment if needed. The program for additional personnel should be established through the use of cooperative agreements well before each fire season begins.

Transportation. Emphasis on fast attack of discovered fires has steadily increased over the years. Many fire-control organizations, both urban and forest, divide their protection area into zones of allowable attack times. These are based on times shown by experience as necessary for successful control according to acceptable standards. Location of suppression forces and methods of transportation are planned to meet these objectives.

Because of the great distances involved, forest fire-control organizations have placed special emphasis on increasing the speed and mobility of their forces. Most personnel and equipment arrive at the majority of fires by a combination of ground-transportation methods, including the most elementary of all, walking. Therefore in areas of high value and high fire incidence, an extensive network of roads and trails is usually maintained. The development of reliable, lightweight gasoline engines has made possible the use of trail scooters and motorized equipment carriers. These require trails intermediate in quality between horse and foot trails and the more expensive truck trails.

In remote areas with relatively low and widely scattered fire incidence, the airplane and helicopter are the primary means of delivering both personnel and material to the scene of a fire. Even in areas with good road and trail networks, it is often cheaper and faster to supply suppression forces by air. Successful parachute jumping by fire fighters, started by the U.S. Forest Service in 1940, demonstrated the feasibility of wartime use of paratroops. Postwar development by the Forest Service of techniques and equipment for delivery of personnel and material by parachute and by helicopters led to increased use of airborne fire fighters in the forests of the western United States (**Figs. 4** and **5**).

Mapping. Specialized heat-sensitive infrared mapping devices are used operationally to map fire perimeters from the air. These are somewhat similar to television systems that are sensitive to radiation in the far infrared rather than to visible light. They have the great advantage of being able to see through the smoke of a fire and locate the regions of active combustion on the ground. The image, which can be photographed, indicates the temperature of the fire in shades from black to white and thus enables fire-control personnel to map the fire edge, determine the intensity of the fire along the perimeter, and locate spot fires outside the main fire at night or when the fire itself is completely obscured by smoke.

Such sophisticated systems of mapping are likely to be applied primarily in the West, where fires may involve extensive areas or the most valuable stands of virgin timber or second growth. For most ground fires that are common in the eastern states, fire maps can be quickly plotted directly on prints of aerial photographs, or made from compass and pacing surveys.

Suppression methods. Fire can be suppressed by robbing it of its fuel or oxygen supply, or enough heat

Fig. 4. Fire fighters rappelling to the scene of a fire.

may be removed to stop the chain reaction, and the fire goes out. Combustion may also be suppressed by the addition of small amounts of certain chemicals to the combustion zone.

Fuel removal. One of the most common methods of controlling a forest fire is by creating a fuelless barrier. This may be done by digging a trench to mineral soil ahead of an approaching fire of sufficient width to prevent the fire from crossing (**Fig. 6***a*). Since this may require a line of considerable width (some evidence is available that this width is approximately equal to the square of the height of the flames), it is customary for the unburned fuel between the prepared line and the approaching fire to be intentionally fired, thereby effectively widening the line.

Techniques of constructing fire lines vary with terrain, soil character, and type of vegetation. Organized crews with hand tools, such as axes, shovels, rakes, and other specially constructed tools, are used in many areas, especially those too rugged or remote to permit use of mechanized equipment (Fig. 6*b*). Specially developed motor-driven flails, similar to garden-tractor rotary tillers, are sometimes employed, although their usefulness is limited to terrain and soil types on which they can be maneuvered satisfactorily. Bulldozers are used extensively (Fig. 6*c*), and tractor-drawn plows are standard line-construction equipment in the relatively flat terrain of the southeastern states.

Occasionally, existing roads or previously prepared firebreaks are used as control lines, usually as a line from which a counterfire is set. Natural firebreaks, such as rocky ridges or rivers, are used similarly.

Heat removal. The second major method used in suppressing fires is to remove enough heat to break the chain reaction. Water, with its high specific heat and heat of vaporization, is an ideal substance for this purpose. Long the favorite with urban fire-control organizations, water is becoming more popular with forestry organizations as techniques and equipment for handling it are improved. An extraordinarily large amount of burning material can be extinguished with a small amount of water when it is used with maximum efficiency.

Hand-operated pumps with a 5-gal (14-liter) tank carried on the back are most satisfactory when used on fires in light fuels, for suppressing small spot fires, and for mop-up, the tedious work of extinguishing burning fuel after the fire has been contained within a fireproof boundary.

Gasoline-operated pumps are used extensively wherever water supplies are convenient or road networks permit the use of tank trucks (Fig. 6*d*). Portable pumps are available that can be carried by one person (Fig. 6*e*), making hose lays of several thousand feet (1 ft = 0.3 m) or even several miles (1 mi = 1.6 km) feasible. Techniques have been developed for laying hose rapidly from helicopters,

Fig. 5. Parachutes used for the delivery of fire fighters and materials to fire areas. (*a*) Two fire fighters descending. (*b*) Air drop of fire-fighting equipment, showing drop location marker (X). (*U.S. Forest Service*)

Fig. 6. Land fire suppression methods. (*a*) Digging a trench ahead of an approaching fire. (*b*) Constructing a fire line to stop the advancing fire. (*c*) Bulldozer making fire lane near the edge of the fire. (*d*) Using a tank truck to suppress fire. (*e*) Fighting fire with hose attached to a portable pump. (*U.S. Forest Service*)

cutting the time required to lay a mile of hose from hours to minutes, even seconds. For use in inaccessible regions, small units including a pump, a 50- or 100-gal (179- or 379-liter) tank, and several hundred feet of hose have been developed that can be delivered by helicopter to the scene of a fire.

Chemicals that inhibit the combustion process are used regularly in the suppression of large forest fires. These are generally prepared in the form of water solutions or slurries, and are dropped from aircraft or applied to fires from ground-based tank trucks. More than 15,000,000 gal (57×10^6 liters) of such fire-retardant chemicals are used annually in the United States.

Although surplus World War II aircraft were originally used in the development of airborne delivery systems (**Fig. 7***a*), newer aircraft have permitted more flexible and efficient release systems. Four-engine cargo planes with pressurized tanks permit the delivery of as much as 3000 gal (11,000 liters) of retardant mixture to a fire. The retardant is released at a low altitude, either directly on the burning material at or near the fire edge, or some distance in front of an advancing fire. A single pass may cover an area several yards wide and several hundred yards long. Helicopters are also used, especially in initial attack and mop-up operations (Fig. 7*b*). While helicopters carry smaller amounts of retardant materials (50 to 100 gal or 179 to 379 liters), their great accuracy makes them excellent for strategic use of expensive materials.

Retardant drops are used primarily to cool down hot spots and help hold a fire pending arrival of ground forces, and to control spot fires outside the main fire perimeter. The air tanker has limited value in attacking the head of a rapidly moving crown fire.

Little direct use is made of the third possibility of removing the oxygen supply to a fire, although water has this effect to some extent and soil is sometimes used to smother flames.

Suppression strategy. The manner in which the various suppression techniques are utilized on a fire is determined by the fire boss, who is comparable to the fire chief in urban organizations. It is the responsibility of the fire boss to determine the strategy appropriate to the particular fire. Depending on the size of the fire, the fire boss may have a few people or a few hundred, and occasionally more than a

Fig. 7. Aerial fire suppression. (*a*) Mixing and loading fire retardants in preparation for an air drop. (*b*) A helicopter with a detachable 40-gal (151-liter) tank drops borate slurry on a pinpoint target. (*U.S. Forest Service*)

thousand, to assist in carrying out the suppression strategy.

When a wildfire has been suppressed, it is essential that it be covered by a report. The fire boss should have someone specifically charged with report duties, or make the report personally. Usually there are printed forms available which include at least the following: (1) name of the fire or district number in which the fire occurred, (2) location of the fire, (3) size of the fire, (4) cause of the fire, (5) estimate of damage in dollars, and (6) date and times when the fire started and when it was suppressed or brought under control. *See* FOREST AND FORESTRY.

Willard S. Bromley

Bibliography. A. A. Brown and K. P. Davis, *Forest Fire: Control and Use*, 1973; T. T. Kozlowski and C. E. Ahlgren, *Fire and Ecosystems*, 1974; W. E. Reifsnyder, *Proceedings of the 10th Annual Tall Timbers Fire Ecology Conference*, 1970; T. Van Nao, *Forest Fire Prevention and Control*, 1982.

Forest genetics

The subdiscipline of genetics concerned with genetic variation and inheritance in forest trees. The study of forest genetics is important because of the unique biological nature of forest trees (large, long-lived plants covering 30% of the Earth's land surface) and because of the social and economic importance of forests. Knowledge of forst genetics is also important. Forest genetics is the basis for conservation, maintenance, and management of healthy forest ecosystems; and development of programs which breed high-yielding varieties of commercially important tree species.

Variation in natural forests. The outward appearance of a tree is called its phenotype. The phenotype is any characteristic of the tree that can be measured or observed such as its height or leaf color. The phenotype is influenced by (1) the tree's genetic potential (its genotype) determined by genes residing in every living cell in the tree; and (2) the environment in which the tree grows as determined by climate, soil, diseases, pests, and competition with other plants. Thus, a tree may be taller than its neighbor because it has the genetic potential for faster growth or because it is growing in a slightly better environment with more light or better soil (or both).

No two trees of the same species have exactly the same phenotype, and in most forests there is tremendous phenotypic variation among trees of the same species (for example, trees of different size, shape, branching patterns, and leaf color; **Fig. 1**). It is only on a very homogeneous site planted with trees of the exact same genotype (called a clone) that the trees may appear to have little phenotypic variability, and even then phenotypic differences are found. Forest geneticists often question whether the observed phenotypic variation among trees in forests is caused mostly by genetic differences or by differences in environmental effects. In other words, which is more important, nature or nurture?

It is impossible to look at the outward appearance of a tree (its phenotype) and know anything about its underlying genotype, because the genetic and environmental effects are combined in the phenotype. For example, a tree can be tall because of its genetic potential or its environment or both. Therefore, common garden tests are often used to hold the environment constant and therefore isolate the

Eucalyptus grandis **is shown growing in two different conditions that exhibit different levels of phenotypic variability. (*a*) In a natural stand in Australia with large phenotypic variability caused by both genetic and environmental differences among the trees (*photo by K. Eldridge*). (*b*) In a plantation of Mondi Forests in South Africa where the trees are more uniform both because the site is uniform and because all trees are the same genotype or same clone (*photo by T. White*).**

genetic and environmental effects on phenotypic variability. To accomplish this, seed is collected from many trees of the same species growing in natural forests, and these seed are planted in randomized, replicated experiments in one or more common field locations (called common gardens). Under these conditions, if the seedling offspring (called progeny) from some parent trees consistently perform better than progeny from other parents, these average differences must be due to genetic causes since the progeny from all parents are being compared in the same common environments.

As an example, suppose that the heights of 1000 trees are measured in a forest of pure loblolly pine, *Pinus taeda*. Each tree has its own genotype, and each has experienced its own microclimate and microsite in the unique part of the stand where it grows. Thus, the measurements of 1000 phenotypes produce a wide range of tree heights, from short to tall. This is the phenotypic variation for the particular trait of height growth. A common-garden experiment to assess how much of this phenotypic variation is caused by genotypic differences among the 1000 trees might be established as follows: (1) Collect seed from the tallest five trees and the shortest five trees in the stand. (2) Keep the seed from each tree in separate bags so that there are ten bags or ten treatments. (3) Plant the seedlings produced from the ten different treatments in randomized, replicated designs in several forest locations (these new field sites are the common gardens). (4) After several years, measure the heights of all seedlings in all replications on all sites. (5) Condense the data by calculating ten averages, one for each of the ten treatments corresponding to the ten parent trees in the original stand. If the progeny seedlings originating from the five tall trees in the parent stand are taller, on average, than the progeny from the five short parents, then at least a portion of the original phenotypic variability among the ten parents must have been caused by genetic differences. This genetic portion of the phenotypic superiority for height growth of the taller parents was passed on to the offspring, as proven by the superior growth of their progeny in common-garden environments.

Geographic variation in natural forests. The term "provenance" refers to a specific geographical location within the natural range of a tree species. Genetic differences among provenances can often be quite large, especially for wide-ranging species occupying many diverse climates. For example, Norway spruce (*Picea abies*) has an extremely wide natural distribution in Europe and Asia that spans many different countries, elevations, climates, and soil types. Natural selection during the course of evolution has adapted each provenance to its particular local environment. This means that there are large genetic differences among provenances growing in different environments. Provenances originating from colder regions (farther north or higher elevations) tend to grow more slowly, begin height growth earlier each spring, end height growth earlier each fall, and have narrower crowns with flatter branches better adapted to the dry snow and types of frosts in colder climates. To demonstrate that these differences among provenances are genetic in origin, common-garden tests, called provenance tests, have been planted. That is, seed has been collected from several provenances and planted for comparison in randomized, replicated studies in various forest locations.

Most wide-ranging species of forest trees (such as Douglas-fir, *Pseudotsuga menziesii*, in the western United States, and loblolly pine, *Pinus taeda*, in the eastern United States) have large genetic differences among provenances from different parts of the range. However, common-garden experiments have revealed that natural selection also acts on a smaller scale within any given region. For example, Douglas-fir stands at lower elevations in the Cascade Mountains of Oregon are genetically distinct from higher-elevation provenances only 50 mi (80 km) apart. Also, Douglas-fir trees from the south and north side of the same mountain can sometimes

differ genetically, presumably in response to natural selection on the colder north-facing slopes.

Some tree species show relatively little geographic variation among provenances. Sometimes this is because the native range is relatively uniform in its climate and soils, such that there have not been strong differences in natural selection pressures in different parts of the range (such as slash pine, *Pinus elliottii*, in the southeastern United States).

The study of geographic variation through provenance tests should be a first step in the genetic research of any tree species. The information gained about the amount and patterns of genetic variation among provenances within the species' native range is important for (1) learning about the interplay and importance of evolutionary forces that caused the observed patterns; (2) making reforestation decisions about how far seed can safely be transferred from the site of seed collection to a distant planting site and still ensure adequate adaptation; (3) determining the most suitable provenances to use to initiate a tree breeding program for the species; and (4) formulating gene conservation strategies that capture the natural genetic diversity that exists within the species.

Genetic variation within forest stands. In addition to genetic differences among provenances, there is usually substantial genetic variation among trees within the same provenance and even within the same forest stand. There are two reasons for this genetic diversity: (1) Different trees have different genotypes in most natural stands. (2) Each tree is heterozygous for many genes, meaning that a given tree has multiple forms (called different alleles) of many genes. Population genetics studies patterns of genetic diversity in populations (such as forest stands). Results of many studies have shown that most forest tree species maintain very high levels of genetic diversity within populations.

Most tree species do not respond well to self-pollination or other forms of inbreeding which would lead to a reduction in genetic diversity. Rather, most species of forest trees have evolved one or more mechanisms to promote outcrossing (that is, cross-pollination between different trees) and maintain genetic diversity: (1) In many conifers (such as pines), the female flowers (that subsequently mature into pine cones) are located in the tree tops, and the pollen-producing catkins (male parts) are located farther down in the tree crown, and this spatial separation helps avoid self-pollination. (2) In some tree species, the male and female parts are ripe for pollination at different times, and this temporal separation of several days helps avoid self-pollination. (3) Even after self-pollination, there are sometimes molecular and physiological mechanisms that favor the germination, growth, and survival of outcross pollen, thus tending to inhibit self-fertilization.

While most tree species maintain large amounts of genetic diversity, there are some exceptions. One example is Torrey pine, *Pinus torreyana*, from southern California, which shows almost no genetic tree-to-tree variation; that is, all trees are nearly the same genotype. In the evolutionary past of these species with little genetic diversity, there have been severe reductions in the total number of living trees of the species (such as due to advance of glaciers during the ice ages). These severe reductions in population size eliminate much of the genetic diversity of the species. Even if the native range of the species today is substantial with many millions of trees, all trees are descendants of the few trees left after the reduction; hence, the genetic diversity is low. *See* POPULATION GENETICS.

Forest tree breeding. Beginning with the natural genetic variation that exists in an undomesticated tree species, tree breeding programs use selection, breeding, and other techniques to change gene frequencies for a few key traits of the chosen species. As with agricultural crops, tree breeders produce genetically improved, commercial varieties that are healthier, grow faster, and yield better wood products. After an existing forest stand is harvested, a new stand of trees is planted to replace the previous stand in the process called reforestation. Use of a genetically improved variety for this reforestation means that the new plantation will grow faster and produce wood products sooner than did the previous stand.

Large-scale tree breeding programs began in earnest in the 1950s. Today, tree breeding programs are very widespread in the world. There are breeding programs for all forest tree species that are planted in substantial quantity. Depending on the species and country, some programs are conducted by single private companies, others by government agencies, and still others by many organizations in which several companies and agencies pool their efforts. In the southeastern United States alone, there are more than 30 private companies and state forestry agencies working cooperatively to breed improved varieties of loblolly and slash pine. Each year these organizations produce more than 1 billion genetically improved seedlings that are used to reforest more than 1 million harvested acres (400,000 hectares). This is more than 90% of the total annual reforestation with pines in the southeast.

Due to the long time frames involved and their recent initiation, tree breeding programs have not advanced as far as breeding programs for most agricultural crops. This means that the new improved tree varieties are usually not discernibly different to the naked eye from the undomesticated, natural species. Only a few traits of key importance (such as growth rate, disease resistance, and wood properties) are targeted for genetic improvement in the tree breeding program, and there is little change in most other traits. Nevertheless, these new varieties benefit the organizations involved by producing more wood products in a shorter time. On a global scale, the overall benefit to society of tree breeding programs is to increase the yields from the world's forest plantations. This helps to meet the rising demand for forest products, which in turn reduces the need to meet this demand from natural forests. Thus, more natural forests can be conserved or used for other social objectives.

Forest biotechnology. Several laboratory techniques of molecular genetics and biotechnology promise to make major contributions to forest genetics and tree breeding in the near future: (1) Somatic embryogenesis is a technique to duplicate (or propagate) selected trees asexually from their vegetative (somatic) cells, and this allows the best trees to be immediately propagated commercially as clones without sexual reproduction (that is, no seed is involved). (2) Genetic mapping of some important tree species is well under way, and these maps will be useful in many ways to understand the genetic control of important traits, such as disease resistance. (3) Marker-assisted selection is the use of some kinds of genetic maps to help select excellent trees at very early ages based on their deoxyribonucleic acid (DNA) genotype as assessed in a laboratory (instead of growing trees in the field and selecting based on performance in the forest). (4) Functional genomic analysis is an exciting new field of genetics that aims to understand the function, controlled expression, and interaction of genes in complex traits such as tree growth. (5) Genetic engineering or genetic modification is the insertion of new genes into trees from other species, and as one example the gene for resistance to the herbicide glyphosate has been inserted into species of *Populus* and *Eucalyptus*.

Conservation of genetic resources. For commercially important species of forest trees, gene conservation is practiced by tree breeding programs to sustain the genetic diversity needed by the program. However, conservation of genetic diversity is a major global concern and is important for all forest species to (1) maintain the health and function of forest ecosystems, and (2) sustain the genetic diversity of noncommercial species that may eventually have economic value. There are many current medicinal and other nonwood products produced by trees (especially trees in tropical forests such as in the Amazon), and there are many more future products as yet undiscovered. For example, the Pacific yew (*Taxus brevifolia*) native to the Pacific Northwest of the United States was once considered a weed. However, this species recently became valuable for its production of a natural chemical, taxol, used in fighting certain types of cancer.

There are two broad categories of gene conservation programs. In situ programs conserve entire forest ecosystems in forest reserves, national parks, wilderness areas, or other areas set aside for conservation purposes. Ex situ programs obtain a sample of the genotypes from different provenances of a single tree species and collect seed or vegetative plant material from each genotype to store in a separate location (such as a seed bank in a refrigerated room). Both types of programs are important for conserving the world's forest resources. *See* BREEDING (PLANT); FOREST AND FORESTRY; FOREST ECOSYSTEM; GENETIC ENGINEERING; PLANT PROPAGATION. Timothy L. White

Bibliography. A. K. Mandal and G. L. Gibson (eds.), *Forest Genetics and Tree Improvement*, CBS Publishers, New Delhi, 1998; T. L. White, W. T. Adams, and D. B. Neale, *Forest Genetics*, CAB Publishers, London, 2001; B. J. Zobel and J. Talbert, *Applied Forest Tree Improvement*, John Wiley, New York, 1984.

Forest harvest and engineering

Application of engineering principles to the solution of forestry problems, such as those dealing with harvesting, forest transportation, materials handling, and mechanical silviculture, with regard to long-range environmental and economic effects.

The work that forest engineers perform varies widely throughout the United States. In the Northeast and Southeast, tasks include mechanization of harvesting, site preparation, planting, and product handling. Development and modification of machinery is an important part of this job, as is the improvement of the ability of workers to use machines in the woods, an activity known as work science. In the West, the terrain changes the job, as does the size of the trees usually harvested. Planning, design, and construction of road systems are major operational and environmental challenges in the West. Harvesting with mechanical fellers or yarding with ground skidders is often limited by tree size and terrain. Cable systems are commonly used to overcome these constraints, and the design and layout of these cable logging units require a great deal of engineering skill if logging is to be done in an economic and environmentally acceptable manner. In short, the skills possessed by forest engineers in the eastern United States parallel most closely those of the mechanical or agricultural engineer. In the West, the skills are more closely aligned with civil engineering. *See* AGRICULTURAL ENGINEERING; CIVIL ENGINEERING; MECHANICAL ENGINEERING.

Education. The education of the forest engineer in the United States and in most of the universities teaching this subject matter throughout the world is most complex. The profession is comparatively new; it grew out of the need for professional engineers in forestry. It has been described as a mixture of the traditional forest management and civil engineering curricula. The forest engineering curriculum had its beginning along these lines, but since the early 1950s, at least in the United States, it has developed its own curriculum, including a number of courses that are oriented toward the objectives of the forest engineering profession. Obviously, curricula vary from school to school, but a representative university curriculum would include courses on general, agricultural, and civil engineering; forest engineering; forest management and forest products; physical, biological, and mathematical sciences; social sciences; economics; business; and the humanities and the arts.

Most curricula are based on four years of study for the bachelor's degree, with options for a five-year double degree in forest engineering–forest management, forest engineering–civil engineering, and so on. Students can proceed to the master's degree and doctor's degree in forest engineering. Summer field

work or a summer camp is required at most schools of forestry.

In several states, forest engineering graduates are eligible to take the state land surveying-in-training and the engineer-in-training tests upon graduation. If they pass, they are eligible to take the state logging test, and some qualify for the civil engineer's test as well.

Since forest engineering is very specialized, most of the graduates elect to stay within the field. The majority find private employment, while the remainder find positions in the U.S. Forest Service, Bureau of Land Management, and the various state and county forestry services.

Logging practices. Forest engineering encompasses many important and diverse subjects. Logging is one of the most important aspects. It is very critical that the logging engineer not only log at the lowest cost possible but also consider other items, such as safety, water and soil protection, ergonomics, esthetics, machine design and development, and systems development.

Generally, the definition of logging includes the planning stage, felling and bucking, yarding, loading, hauling, and unloading. The most spectacular of these and usually the most expensive is the yarding or skidding phase. Yarding is moving a log by some method, usually a cable system, that applies some lift to the leading end of the log; skidding accomplishes the same result, but very little or no lift is applied. A typical example of skidding would be horse skidding, during which the log is in contact with the ground at all times. Yarding or skidding methods fall into the following categories: animal skidding, cable yarding, surface-vehicle yarding or skidding, and under exceptional conditions helicopter or balloon yarding. *See* LOGGING.

Forest engineering problems. The specialized training and subsequent experience of qualified forest engineers are called upon when they may be required to do the following:

1. Compare the costs of owning and operating different units and different combinations of logging equipment. Such comparisons are usually made by dividing cost per unit of time by the production rate— the volume of board-feet, cords, or other volume produced for the same unit of time.
2. Recommend the standards of the roads that are to be used in transporting forest products from the woods.
3. Determine the most efficient spacing of logging and other roads with a given set of logging equipment, or with a given plan for managing all the forest resources.
4. Survey and designate the location of logging and other roads.
5. Locate and design the type of culverts, bridges, and other construction features of the road system recommended.
6. Assist with the location, design, and construction of recreational trails, buildings, shelters, camps, and picnic and playground facilities.
7. Assist with the selection and operation of the most efficient machinery to be used in such silvicultural programs as forest tree nursery operations, site preparation, tree planting, and management of problem vegetation.
8. Be responsible in all planning and operation to conform with environmental regulations and all laws affecting the equipment and programs under the forest engineer's recommendations or direction.
9. Be conscious of and provide for the safety of the persons operating the equipment or using the structures recommended or designed.

Environmental aspects. Protecting the forest environment is increasingly important to the forest engineer. The "cut and get out" attitude toward forest properties is a thing of the past. Today's engineer has an obligation to the employer as well as the public to preserve forests, whether they are publicly or privately owned. Sometimes, this creates a conflict among interested parties in various phases of forest management. Environmental issues include water quality, soil disturbance, preservation of scenic values and recreation and wilderness areas, presence of wildlife, controversial logging methods like clearcutting, smoke due to slash burning, and forest pest control. *See* FOREST AND FORESTRY; FOREST MANAGEMENT.

John E. O'Leary

Bibliography. F. C. Ford-Robertson and R. K. Winters (eds.), *Terminology of Forest Science, Technology, Practice and Products,* rev. ed., 1983; B. H. Maddison, *Handbook of Timber Engineering Design,* 1982; K. A. Staaf (ed.), *Tree Harvesting Techniques,* 1984.

Forest management

The planning and implementation of sustainable production of forest crops and other forest resources and uses. Key decisions include land allocation to different uses or combination of uses, silvicultural method and practices, intensity of management, timber harvest scheduling, and environmental protection.

Private owners. Nearly three-fourths (360 million acres or 140 million hectares) of all commercial forest land in the United States is privately owned by farmers, forest investment groups, other types of nonindustrial owners, or industrial firms engaged in the business of growing and harvesting timber for conversion to wood products. The objectives and practices of private owners are extremely diverse. A small, nonindustrial owner may practice extensive forestry, harvest timber infrequently, emphasize recreational activity, and treat forest assets as an immediate source of cash. A large public corporation has a financial responsibility to maximize the value of the forest and any associated facilities. Monetary objectives are therefore important in making forest management decisions: Timber is commonly harvested annually on a sustained yield basis; harvested sites are planted with genetically improved trees; competition may be removed; and growing stands may be fertilized and thinned periodically. Recreational activity such as hunting and fishing may

be permitted and generate significant income. Unpriced resources may be of secondary importance to the owner.

Many states and local governments have enacted laws that regulate the practice of private forestry. Historically, the primary objectives were to assure long-term timber supply, protect streams, control soil erosion, minimize wildfire, and provide for compatible land uses. Since the 1980s, wildlife habitat and biological diversity have become important regulatory objectives in some states. Therefore, management planning for a specific property requires a detailed review of the owner's objectives, resources, and any legal constraints regarding land uses or choice of management practices in the local area.

Public agencies. One-fourth (about 120 million acres or 49 million hectares) of the commercial forest in the United States is administered by federal and state agencies. The National Forest System, managed by the U.S. Forest Service, is particularly important in the western United States, where it includes nearly one-half of the commercial forest. Congress requires that federal lands be managed for multiple uses under long-term sustained yield principles. Each national forest is required by federal law to develop a long-term land management plan, involve the public in evaluating alternatives, project future practices and outputs, and identify methods to mitigate adverse environmental impacts. These plans provide for timber harvesting, wilderness management, watershed protection, wildlife habitat, and other services in combinations that vary from forest to forest. The recovery plans of many endangered species occur primarily on public forest lands.

Land-use plans. Planning is a process for making decisions about objectives, allocating land among alternative uses, selecting management practices, determining timber harvest levels, and providing for environmental protection. Computer-based mathematical models can be used to project and evaluate the potential outcomes of alternative programs. However, the decision-makers must exercise considerable professional judgment because of uncertainty, necessary assumptions, and because resource values, for example those associated with wildlife and scenery, are not readily quantified.

A forest land-use plan is the outcome of a planning process. The complexity of the process and the plan itself will depend upon the owner's situation. Farmers and other small, private forest owners may retain professional foresters, such as extension foresters, company foresters, or private consultants, to develop plan. They may also use intuitive, informal methods. Few develop formal written plans. In contrast, federal and state agencies are required by law to develop written land management plans. Many large industrial firms also develop formal plans. Written plans vary in degree of detail and resource use issues. Generally, the ideal plan includes two parts: a map that illustrates the location of land uses, access system, and site development; and a written narrative that establishes objectives and constraints, describes and evaluates alternatives, estimates the expected flow of services including timber harvests, describes and schedules the management activities needed to implement the plan, compares costs and benefits, and discusses environmental considerations. *See* LAND-USE PLANNING.

Resource assessment. Planning and project implementation requires information about the physical, vegetative, and developmental characteristics of forest resources within the management unit. Aerial photography, satellite imagery, and statistically designed ground surveys are commonly used to obtain the necessary information. To facilitate development of management prescriptions, estimation of timber growth and yield, and environmental analyses, the unit should be divided into relatively homogeneous land classes based on such attributes as soil type, vegetation cover type, erosion hazard, slope, and age- and size-class of forest stands. Maps showing the geographical distribution of each class, existing road access, any legally restricted land uses, and boundaries of the planning unit are essential. Computer-operated software for preparing such maps is widely used. *See* GEOGRAPHIC INFORMATION SYSTEMS; REMOTE SENSING.

The resource assessment should normally include estimates of timber volume classified according to species, age or size class, quality, location, and other attributes that affect value in the local market or have relevance to decision-making. Typically, statistically designed sampling procedures using plots or strips are employed by specially trained technicians to estimate volume. Periodic surveys at 5-, 10-, or 15-year intervals provide information for monitoring changes in inventory, timber growth, regeneration following harvest, and effects of logging on environmental conditions. *See* FOREST MEASUREMENT.

Information about nontimber resources such as wildlife, streams and lakes, fisheries, and historical and archeological sites may also be required, depending upon the owner's objectives and local forest practice regulations. The identification of species classified legally as threatened or endangered is critically important to the manager, as their presence may affect land use and choice of management prescriptions. Assessment methods normally utilize both ground and aerial surveys, and professional specialists employed by the owner or by outside consultants.

Forest structure, growth, and yield. For timber management purposes, stands of trees may be classified as either even aged or uneven aged. The classification determines the method for regulating growth and yield.

In even-aged stands, all trees were established at about the same time, whether by natural (for example, germination of seeds following destruction of the previous stand by fire) or artificial means (planting of trees or seeding following a commercial timber harvest). As the stand ages, competition may eventually produce a mix of tree sizes. An even-aged stand has a life cycle with a definite beginning and ending. The period of time between these points is called the rotation. Stands which are being managed for

biomass production (for example, wood to be used for the production of paper pulp), may be planted at a high density, receive no thinning, and be harvested at a point near to the stand's maximum sustainable yield. Intensively managed hybrid poplar (*Populas*) or willow (*Salix*) stands may have rotations as short as 7 years while low productivity Jack Pine (*Pinus banksiana*) stands may use a 50-year rotation. Stands which are managed for solid wood production (for example, lumber or veneer) will be maintained at a lower density in order to target growth, receive one or more thinnings, and have a longer rotation (as short as 20 years in the U.S. South to over 100 years for quality hardwoods). The sequence of practices in the second and subsequent cycles may differ from the first if conditions, including the manager's objectives, have changed.

A forest can be structured to produce a regular periodic or annual yield of timber by developing and maintaining a series of even-aged stands growing toward harvest age in a regular sequence. For example, a 10,000-acre (4050-ha) forest that is regulated under a 50-year rotation to produce an annual harvest would have 200 acres (80 ha) in each 1-year age class. Each year, the oldest age class is harvested and immediately regenerated. Theoretically, over time, total growing stock (inventory) and timber harvest remain constant. Sustainable annual harvest $H = A/t \times Y(t)$, where A is the forest area in acres, t is rotation in years, and $Y(t)$ is yield of timber per acre at age t.

Uneven-aged stands typically have a mix of trees of different ages, sizes, and species. Such stands do not normally have a specific beginning or ending because the site is continuously stocked with growing trees. Uneven-aged stands may be managed in a continuing series of partial cuttings. The period between such harvests is the cutting cycle. During each cycle, a portion of the inventory is harvested, leaving a reserve stock of trees to grow over the next cycle.

Regulation of timber harvest under an uneven-aged management system is based on the cutting cycle in a manner similar to the use of a rotation period in the even-aged system. Each acre in a 10,000-acre (4050-ha) forest managed under a 1-year cutting cycle would be partially cut each year. In practice, this may not be feasible because one-year's growth in per-acre units is usually too small for economically efficient harvesting. Cutting cycles of 10–20 years are more commonly followed. To achieve a regular annual yield, the forest area is divided by the length of the cutting cycle to determine the number of cutting units. For example, for a forest of 10,000 acres (4050 ha) managed under a 10-year cutting cycle, there would be 10 cutting units of 1000 acres (405 ha) each. Each year, one unit is selectively harvested; over the 10-year cutting cycle, each unit is harvested once. The sequence may be continued perpetually, producing a regular annual yield and maintaining reserve growing stock at a constant level.

In practice, because the demand for timber fluctuates in response to the general business cycle, under either the even- or uneven-aged management systems it may be necessary, or financially desirable, to cut more in some years and less in others. *See* FOREST TIMBER RESOURCES; REFORESTATION; SILVICULTURE.

Management scheduling. A forest typically will have complex structure. There may be a range of soil types, slopes, and aspects that differ in potential productivity. The distribution of stands and trees according to age and size classes may be irregular rather than balanced. Some stands may be understocked, while others are too dense for optimal growth. Federal or state laws may require preservation of habitat for an endangered species or stream protection.

To facilitate planning for such complex situations, optimization methods may be used to schedule management activities over time, determine the timber harvest level, allocate the land base among alternative uses, and calculate benefits and costs. A common method, linear programming, requires that the manager specify both a linear objective function to be maximized and a set of linear equations that describe management constraints. The constraints normally will include land area stratified according to vegetation-soil characteristics, beginning inventory, timber harvest-flow policy, regeneration requirements, and environmental protection requirements. Computer-based software is available to solve such problems, which typically include an automated procedure for generating the system of equations that defines the manager's objectives, alternatives, and constraints; a linear programming package; and a report writer that prepares the results in a form useful to the manager. *See* LINEAR PROGRAMMING.

Multiple use. Modern forest management usually involves the production of multiple services of value to the owner or to society. Determining the best mix of services requires technical information on trade-offs between the different outputs, costs and values, and pertinent legal constraints. Subjective evaluations may be required in the case of unpriced services such as wildlife, water, and scenery.

Forest structure has a major impact on the mix of outputs. On a forest of fixed size managed on an even-aged basis, increasing the planning rotation beyond a certain age may reduce the level of sustainable timber production but provide for a higher level of biological diversity. The populations of some wildlife species may be favored by short rotations, while others require old-growth-like conditions associated with long rotations. The technical relationships that define such trade-offs are often not available or are poorly understood. Thus, a great deal of subjective judgment by the manager is required in making multiple-use decisions.

Forest outputs or services can generally be classified as complementary or competitive. Services are complementary if an increase in the output of one is accompanied by an increase in the output of the other. They are competitive if effort to increase the output of one results in a reduction in the output of the other. Services may be complementary over one range of outputs and competitive over another.

In response to concerns about landscapes, such as maximum opening sizes, spatial diversity, and

wildlife habitat, new scheduling methods are being developed and used (for example, tabu search, simulated annealing, and heuristics). Unlike linear programming, these methods do not guarantee optimal solutions. However, they do provide quality solutions that consider the complexity of modern forest management and can be implemented on the ground. *See* FOREST AND FORESTRY.

Dennis E. Teeguarden

Bibliography. M. D. Bowes and J. V. Krutilla, *Multiple-Use Management: The Economics of Public Forest Lands*, 1989; F. W. Cubbage, J. O'Laughlin, and L. S. Bullock III, *Forest Resource Policy*, 1993; L. S. Davis and K. Norman Johnson, *Forest Management*, 4th ed., 2000; W. A. Duerr, *Introduction to Forest Resource Economics*, 1993; W. A. Duerr et al., *Forest Resource Management: Decision-Making Principles and Cases*, 2d ed., 1987; W. D. Klemperer, *Forest Resource Economics and Finance*, 1996.

Forest measurement

The science and practice of measuring the volume, growth, and development of trees individually and collectively and estimating the products obtainable from them. Foresters use quantitative sciences such as mathematics and statistics for these measurements.

Regardless of the land management objectives—timber, wildlife, recreation, watershed, or a combination of these resources—the tree overstory (the forest canopy) must be quantified for informed decision making. Forest cover is an important part of wildlife habitat, and the understory component is related to the overstory characteristics. The recreation potential of wildland is a function of many variables, including the size and number of trees present. Water yields are related to the composition and density of the tree canopy. The measurement principles discussed here are applicable to all forest resource management situations that require quantitative information about the tree component of the land base.

Primary product measurement. Products cut from trees are usually measured in different units, depending on local customs and end use. The most common units are cubic volume, stacked wood measure, board-feet, and weight. Tree sections 8 ft (2.4 m) or more in length are called logs; shorter pieces are referred to as sticks or bolts.

Cubic volume. Cubic volumes of logs or bolts are computed from the length and the cross-sectional area, which may be taken at the midpoint of the log or averaged from the areas of the ends or from the weighted average of the small ends and the midsection. Although tree cross sections rarely form true circles, they are assumed to be circular for purposes of computing cross-sectional areas. Areas and volumes are computed in U.S. Customary units (square feet or cubic feet) or in metric units (square meters or cubic meters).

Stacked wood. Pulpwood and firewood are commonly sold by stacked wood measure. Stacked wood is generally measured in terms of cords. A standard cord of wood measures 4 ft × 4 ft × 8 ft (1.2 m × 1.2 m × 2.4 m); its volume is 128 ft^3 (3.6 m^3) of wood, bark, and air space. The cord content of any stack of roundwood is given by Eqs. (1).

$$\text{Cord content} = \frac{\text{width (ft)} \times \text{height (ft)} \times \text{stick length (ft)}}{128} \tag{1a}$$

$$\text{Cord content} = \frac{\text{width (m)} \times \text{height (m)} \times \text{stick length (m)}}{3.6} \tag{1b}$$

Cordwood customers are primarily interested in the solid-wood volume rather than the total space occupied. The solid content of a stack of wood varies with species, method of piling, diameter, length and straightness of sticks, and length of branch stubs.

Board-foot log rules. Logs that will be converted into lumber are generally scaled for their estimated board-foot content. Scaling involves determining the log diameter inside the bark at the small end and the log scaling length (in multiples of 1 or 2 ft; or 0.3 or 0.6 m). The log scaling dimensions are then used to enter a log rule, which is a table or formula showing the estimated volume, usually in board-feet, for logs of various diameters and lengths. The board-foot is equivalent to a plank 1 in. (2.5 cm) thick and 1 ft (30 cm) square; it contains 144 $in.^3$ (2360 cm^3) of wood. Although the board-foot has been useful for measuring sawed lumber, it is an inconsistent unit for log scaling.

The board-foot content of sawed lumber is given by Eqs. (2).

$$\text{Bd-ft} = \frac{\text{thickness (in.)} \times \text{width (in.)} \times \text{length (ft)}}{12} \tag{2a}$$

$$\text{Bd-ft} = \frac{\text{thickness (cm)} \times \text{width (cm)} \times \text{length (m)}}{23.6} \tag{2b}$$

In the United States, three log rules are commonly used: the Scribner log rule, the Doyle log rule, and the International log rule. The Scribner log rule was derived by diagramming 1-in. (2.5-cm) boards drawn to scale within circles of various sizes that represent the small end of logs to be scaled. Allowance was not made for taper, and so the rule ignores all volume outside scaling cylinders projected from small ends of logs.

The Doyle log rule is based on a mathematical equation. This rule tends to underestimate small logs and overestimate large ones, but it nevertheless remains in wide use, particularly in the southern United States.

The International log rule, based on a reasonably accurate formula, includes a fixed taper allowance of $^1/_2$ in. (1.27 cm) per 4 ft (1.2 m) of log length. The International $^1/_8$-in. rule allows for saw kerf of $^1/_8$ in. (0.32 cm), whereas the International $^1/_4$-in. rule allows for saw kerf of $^1/_4$ in. (0.64 cm). The Scribner,

TABLE 1. Board-foot log rules for 16-ft (4.9-m) logs

Log diameter, in. (cm)	Scribner, bd-ft (m^3)	Doyle, bd-ft (m^3)	International $^1/_4$ in., bd-ft (m^3)
8 (20.3)	32 (0.076)	16 (0.038)	40 (0.094)
12 (30.5)	79 (0.186)	64 (0.151)	95 (0.224)
16 (40.6)	159 (0.375)	144 (0.340)	180 (0.425)
20 (50.8)	280 (0.661)	256 (0.604)	290 (0.684)
24 (61.0)	404 (0.953)	400 (0.944)	425 (1.003)
28 (71.1)	582 (1.373)	576 (1.359)	585 (1.380)
32 (81.3)	736 (1.737)	784 (1.850)	770 (1.817)
36 (91.4)	923 (2.178)	1024 (2.416)	980 (2.312)
40 (101.6)	1204 (2.841)	1296 (3.058)	1220 (2.879)

Doyle, and International $^1/_4$-in. log rules for 16-ft (4.9-m) logs are compared in **Table 1.**

As noted in Table 1, the different log rules provide varying estimates of board-foot content. Further, one would not expect any given log rule estimate to coincide exactly with the mill tally (actual lumber volume sawed). If the lumber tally is greater than the log scale, the excess is called overrun. When log scale values are larger than sawed output, the deficit is termed underrun. The overrun or underrun can be expressed by Eq. (3).

$$\text{Overrun or underrun (percent)} = \frac{\text{mill tally} - \text{log scale}}{\text{log scale}} \times 100 \quad (3)$$

Weight scaling. Much of the pulp and paper industry buys pulp wood by weight scaling rather than linear measure. Price/weight equivalents are usually established by making thousands of paired weighings and stick scalings of loads of freshly cut wood. A weight-per-cord conversion is stated, and prices are quoted on a per-cord basis.

Sawlogs are also bought by weight, with suitable adjustments made to reflect the fact that yields of usable lumber per unit weight are less in smaller or defective logs. *See* LOGGING; PAPER; WOOD PRODUCTS.

Measuring standing trees. Standing trees are commonly measured for diameter, height, and age. The diameter and height measurements are used to estimate the volume (or weight) and value of individual trees; ages are used in assessing site quality and predicting growth.

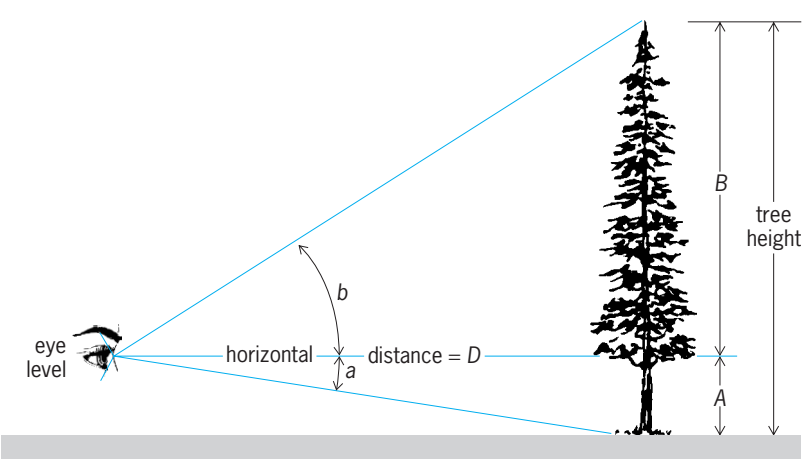

Fig. 1. Determination of tree height using common sighting instruments. *A* = tan angle *a* × *D*. *B* = tan angle *b* × *D*. Tree height = *A* + *B*. (*After T. E. Avery and H. E. Burkhart, Forest Measurement, 4th ed., McGraw-Hill, 1994*)

Diameters. The tree measurement made most frequently by foresters is diameter at breast height (dbh), defined in the United States as the stem diameter outside the bark at a point 4.5 ft (1.3 m) above the average ground line on the uphill side of the tree. The instruments used to measure tree diameters are called dendrometers. A diameter tape actually measures circumference but is graduated to provide direct readings of tree diameter. Calipers are convenient for measuring small trees. The Biltmore stick is a straight wooden slat, like a ruler, but with a specially graduated scale for direct readings of dbh.

Upper-stem diameters are occasionally obtained directly with the aid of ladders or climbing irons to reach the point of measurement, but this is time-consuming and may be hazardous. Upper-stem dendrometers have been developed to allow measurement to be made from the ground. These devices include calipers on a pole, stadia-based instruments, split-image rangefinders (such as the Barr and Stroud optical dendrometer), and the pentaprism caliper.

Heights. Height-measuring devices are called hypsometers and usually embody the basic trigonometry principle illustrated in **Fig. 1.** The observer stands at a fixed horizontal distance from the tree. Tangents of angles to the top and base of the tree are multiplied by the predetermined horizontal distance to give the height of each measured section of the tree. For accurate results, trees must not lean more than 5° from the vertical, and the fixed horizontal distance must be carefully determined. *See* SURVEYING.

For less precise work, a Merritt hypsometer may be used. This has a linear scale, often printed on one face of a Biltmore stick and based on the principle of similar triangles. It is normally used for determining merchantable log heights rather than total tree heights.

For smaller trees, direct linear measurement of tree height may be made by using height poles, which consist of joined tubing with colored stripes at regular (1- or 2-ft, or 0.3- or 0.6-m) intervals.

Foresters generally measure either total tree height or merchantable height, although in some instances both are measured. Total tree height is the linear distance from ground level to the upper tip of the tree crown. The tip of the crown is easily discerned when trees have conical shapes, but it may not be readily determined for deciduous trees with irregular

TABLE 2. Merchantable volume* of ponderosa pine trees, in cubic feet (cubic meters)

Diameter at breast height, in. (cm)	Total height above ground		
	40 ft (12 m)	60 ft (18 m)	80 ft (24 m)
6 (15.2)	3.2 (0.091)		
8 (20.3)	5.8 (0.164)	8.9 (0.252)	
10 (25.4)	9.1 (0.258)	13.8 (0.391)	
12 (30.5)	13.1 (0.371)	19.8 (0.561)	26.6 (0.753)
14 (35.6)	17.7 (0.501)	26.8 (0.759)	35.9 (1.017)
16 (40.6)	23.1 (0.654)	34.8 (0.985)	46.6 (1.320)
18 (45.7)	29.1 (0.824)	43.9 (1.243)	58.7 (1.662)
20 (50.8)	35.9 (1.017)	54.0 (1.529)	72.2 (2.044)
22 (55.9)	43.3 (1.226)	65.2 (1.846)	87.0 (2.464)
24 (61.0)		77.3 (2.189)	103.3 (2.925)
26 (66.0)		90.6 (2.565)	120.9 (3.423)
28 (71.1)		104.8 (2.968)	139.9 (3.961)
30 (76.2)		120.1 (3.401)	160.3 (4.539)

*Inside-bark volume, from a stump height of 1 ft (0.3 m) to a 4-in. (10-cm) top diameter.

or round-topped crowns. Consequently, total height is most often applied when measuring coniferous trees.

Merchantable tree height refers to the usable portion of the tree stem, and is defined for single-stemmed trees as the length from an assumed stump height to an arbitrary upper-stem diameter. If the main trunk does not extend as a single stem to the tree tip, the upper limit of merchantability is generally dictated by the branching characteristics.

Age. Many tree species found in the northern temperate zone grow in diameter by adding each year a single, distinctive layer of wood. This layer consists of a band of early or spring wood and a band of less porous, darker late or summer wood. The two bands make up an annual ring, which appears in a cross section of the stem as a series of concentric rings that can be counted to give the age of the tree.

Complete cross sections give the most reliable ring counts but are often unavailable. Ages of standing trees are therefore often determined by extracting a radial core of wood with an instrument called an increment borer. The reliability of annual ring counts depends on the species and the growth conditions. Prolonged drought or defoliation by insects followed by favorable growing conditions can create "false rings," which are often incomplete and distinguishable only in full cross sections. Heavily suppressed forest trees can also have cross sections with more than one year's growth represented by an annual ring. Such "missing rings" may be evident along only part of the tree length. *See* DENDROLOGY; TREE GROWTH; WOOD ANATOMY.

Volumes and weights. A volume table gives the average contents for standing trees of various sizes. The most commonly employed units of volume are board-feet, cubic feet, cubic meters, and cords. Volumes may be listed for some specified merchantable portion of the stem or for the total stem. In modern practice, equations are generally used to predict tree volumes rather than obtaining the values from tables. However, volume table has persisted in forestry usage as a generic term meaning tabulations or equations that show the contents of standing trees. Tree weight tables are analogous to volume tables except that weights (green or dry) of standing trees are predicted rather than volumes.

Volume tables based on diameter at breast height and tree height are referred to as standard or multiple-entry tables (**Table 2**). Local or single-entry volume or weight tables are based solely on diameter at breast height.

The most commonly used equation for deriving standard (multiple-entry) volume or weight tables is the so-called combined variable equation (4), where

$$V = b_0 + b_1 D^2 H \qquad (4)$$

V is the tree volume or weight, D the diameter at breast height, and H a measure of total or merchantable tree height. The constants b_0 and b_1 are based on sample data.

Tarif tables. Among local or single-entry volume tables, tarifs are of special interest. Composed of tabulated information, they are based on the assumption of proportionality between the volume of a tree and its basal area (the cross-sectional area at the diameter at breast height). The constant of proportionality varies with the species, age, height, and other factors. Tarif numbers are averaged from height and diameter-at-breast-height data of a stand of trees and used to construct a local volume table based solely on diameter at breast height.

Biomass estimation. With increasing emphasis on complete tree utilization and use of wood as an energy source, methods have been developed for determining the weights of trees and their components (bole wood, bark, branches, foliage). The results are expressed in terms of dry weight and incorporated in biomass tables, which may include only the aboveground portion or the entire tree, including roots. Like tree volume and weight, biomass is estimated from diameter at breast height and tree height.

Timber inventory. The purpose of a timber inventory, or timber cruise, is to estimate the volume (or value) of standing trees in a given area. This is usually a sampling process. Other things being equal, the intensity of sampling tends to increase as the size of

the tract decreases and as the value of the timber increases. Foresters commonly conduct timber inventories for land acquisition, timber sales, management planning, and special assessments, such as evaluating regeneration and insect or disease outbreaks.

Sampling methods vary. They may, for example, use continuous strips of uniform width spaced uniformly through the forest (**Fig. 2**). Another technique is to use a set of fixed-area plots located throughout the area of interest. The sample plots can be located by any random or systematic sampling design. However, the most common procedure is to space the plots at intervals on lines separated by a fixed distance (**Fig. 3**). This is termed a line-plot cruise.

Forest monitoring. In addition to inventories aimed at determining current conditions, land managers need trend data. Monitoring consists of collecting information over time, generally on a sample basis by measuring change in key indicator variables, in order to determine the effects of management treatments in the long term. These data, along with research results, can be used to modify management on a continuing basis to ensure that objectives are being met. The sampling design for monitoring generally involves repeated measurements on the same sample plots or individuals.

Geographic information systems. Forest inventory information is commonly stored, updated, and retrieved through geographic information systems (GISs). A GIS is a computerized database for storing, manipulating, and displaying map (spatial) data and tabular (attribute) information. In a GIS, forest inventory information can be stored in a computer and directly linked to associated forest maps, making it easier and faster to analyze and graphically display the results of forest inventories.

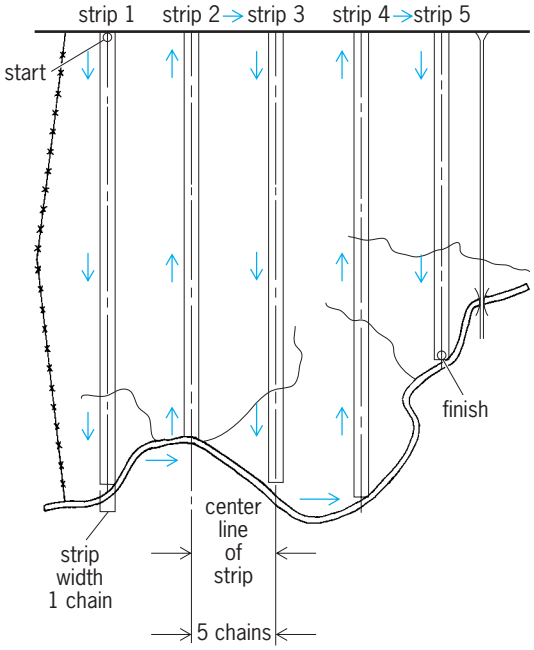

Fig. 2. Plan for a nominal 20% systematic strip inventory. Sample strips 1 chain (66 ft or 20 m) wide are spaced 5 chains (330 ft or 100 m) apart. (*After H. E. Burkhart et al., Timber inventory, in K. F. Wenger, ed., Forestry Handbook, 2d ed., John Wiley, 1984*)

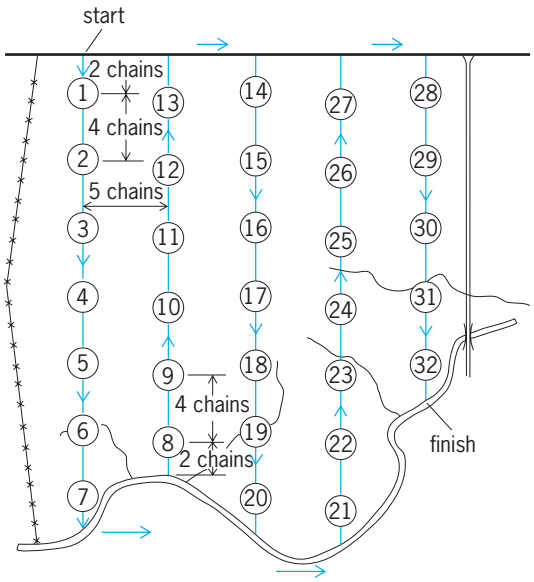

Fig. 3. Plan for nominal 10% systematic plot inventory. Sample plots are circular, $^1/_5$ acre (0.08 ha) in size, spaced at 4-chain (264-ft or 80-m) intervals on lines 5 chains (330 ft or 100 m) apart. (*After H. E. Burkhart et al., Timber inventory, in K. F. Wenger, ed., Forestry Handbook, 2d ed., John Wiley, 1984*)

GISs can make forest inventory information more powerful by allowing forest resource managers to integrate it with other data commonly needed to make management decisions. For instance, managers may need to know the location of roads, streams, or sensitive soils when developing management plans. Combining forest inventory data with other land resource information allows managers to make more informed decisions.

In a GIS, the basic types of data are separated into individual maps called themes (**Fig. 4**). The locations of each feature in each theme are stored in the computer database, and the characteristics associated with that feature are linked to it. For example, the locations of stand boundaries are stored in the stands theme of the GIS, along with the inventory results (for example, number of trees, volume) for each stand. Similarly, the locations and attributes (such as the flow rate, pH) of streams, soil types with their characteristics (such as, soil series, drainage class, texture, erodability class), and other relevant resources are entered into the GIS. Each of these themes can be examined individually or in combination with others.

Site quality. Foresters estimate site quality to assess present and future forest productivity and to provide a frame of reference for land management. Many parameters that affect productivity are difficult or impossible to measure directly, and as a consequence site quality is determined indirectly.

TABLE 3. Yield table for thinned stands of yellow-poplar, in total cubic-foot volume per acre (m^3/ha)*

Age, years	Site index,† ft (m)	Basal area		
		50 ft^2/acre (11.5 m^2/ha)	90 ft^2/acre (20.7 m^2/ha)	130 ft^2/acre (29.8 m^2/ha
20	90 (27.4)	1032 (72.2)		
	110 (33.5)	1266 (88.6)		
	130 (39.6)	1459 (102.1)	2582 (180.7)	
40	90 (27.4)	1789 (125.2)	3166 (221.5)	
	110 (33.5)	2195 (153.6)	3884 (271.8)	5552 (388.5)
	130 (39.6)	2529 (177.0)	4475 (313.1)	6395 (447.5)
60	90 (27.4)	2149 (150.4)	3804 (266.2)	5436 (380.4)
	110 (33.5)	2637 (184.5)	4666 (326.5)	6669 (466.6)
	130 (39.6)	3037 (212.5)	5375 (376.1)	7682 (537.5)
80	90 (27.4)	2356 (164.9)	4169 (291.7)	5958 (416.9)
	110 (33.5)	2890 (202.2)	5114 (357.8)	7309 (511.4)
	130 (39.6)	3329 (232.9)	5891 (412.2)	8420 (589.2)

*Yields are for wood and bark of the total stem for all trees with a diameter at breast height of 4.5 in. (11.4 cm) and larger.
† Base age 50 years.

Most commonly, site quality is evaluated from tree height in relation to age. Theoretically, height growth is sensitive to differences in site quality, little affected by varying stand-density levels, relatively stable under varying thinning intensities, and strongly correlated with volume. Thus height has been found to be a practical, consistent, and useful indication of site quality. Site index is defined as the average height of the dominant trees (those in the upper portion of the main canopy) at an arbitrary index age (typically 25, 50, 100 years). As generally applied, site index is estimated by determining the average total height and age of dominant and codominant trees in well-stocked, even-aged stands. When these two variables (height and age) have been ascertained for a given species, they are used as coordinates for interpolating site index (height at index age) from a specially prepared set of curves, or for substitution into a site index equation.

Stand density and growth. Quantitative measures of stand density are used when deriving silvicultural prescriptions and predicting growth and yield. The two most commonly used measures of stand density are tree basal area per unit area and number of trees per unit area. Stand basal area is the cross-sectional area at diameter at breast height of all stems, or some specified portion of the stand, expressed on a per-acre (or per-hectare) basis. Similarly, trees per acre may be determined for all stems or for some specified portion of the stand. *See* SILVICULTURE.

Although growth rates vary widely among different tree species, the general pattern of growth is remarkably consistent. When expressed as a cumulative function, all measures of growth follow a characteristic sigmoid curve like that shown for height in **Fig. 5**.

Forest management decisions require information about both current and future resource conditions. Inventories taken at one instant in time provide data on current volumes and related forest characteristics. However, forests are dynamic biological systems that are continuously changing, and information on change is also needed. Stand dynamics (that is, growth, mortality, and reproduction and related changes in the stand) are predicted with growth and

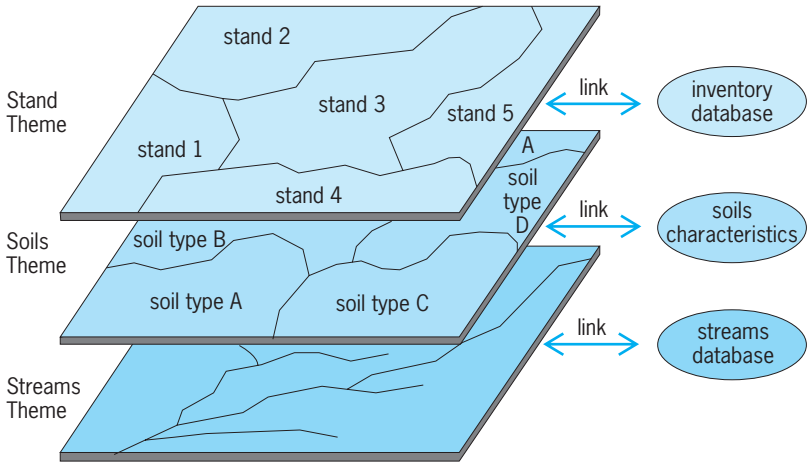

Fig. 4. Data from diverse sources can be stored, manipulated, and displayed by use of geographic information systems. (*After T. E. Avery and H. E. Burkhart, Forest Measurements, 4th ed. McGraw-Hill, 1994*)

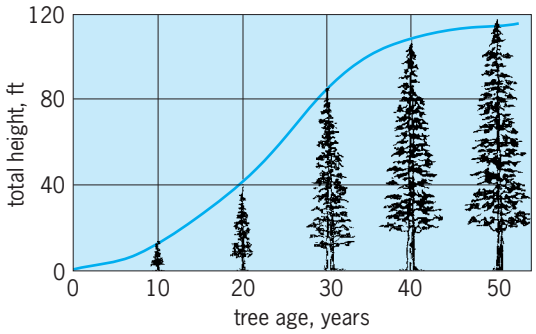

Fig. 5. The characteristic sigmoid growth curve of trees is illustrated by the cumulative height-growth pattern of some coniferous species (1 ft = 0.3 m). (*After T. E. Avery and H. E. Burkhart, Forest Measurements, 4th ed., McGraw-Hill, 1994*)

yield models. These models are quantitative representations of complex biological processes.

Growth is the increase (increment) over a given period of time. Yield is the total amount available for harvest at a given time, that is, the summation of the annual increments. The factors most closely related to growth and yield of forest stands of a given species composition are the point in time in stand development, the site quality, and the degree to which the site is occupied. For stands of uniform age, these factors can be expressed quantitatively by stand age, site index, and stand density, respectively. For a given site index and initial stand-density level, volume per unit area (yield) plotted over stand age results in a sigmoid curve. Yields for thinned stands of yellow-poplar are shown in **Table 3**. *See* FOREST AND FORESTRY; FOREST MANAGEMENT. Harold E. Burkhart

Bibliography. T. E. Avery and H. E. Burkhart, *Forest Measurements*, 4th ed., McGraw-Hill, New York, 1994; J. F. Bell and J. R. Dilworth, *Log Scaling and Timber Cruising*, O.S.U. Book Stores, Corvallis, Oregon, 1997; M. S. Philip, *Measuring Trees and Forests*, 2d ed., CAB International, Wallingford, UK, 1994; H. T. Schreuder, T. G. Gregoire, and G. B. Wood, *Sampling Methods for Multiresource Forest Inventory*, John Wiley, New York, 1993; K. F. Wenger (ed.), *Forestry Handbook*, 2d ed., John Wiley, New York, 1984.

Forest recreation

The use of forests for recreation is increasing. For many people, outdoor recreation involves direct contact with forests in various activities ranging from walking in the woods to wilderness backpacking. The primary suppliers of forest recreation opportunities in the United States are federal land management agencies such as the U.S. Department of Agriculture (USDA) Forest Service, the National Park Service, and the Bureau of Land Management; state recreation, park, wildlife, and forest departments; and private landowners and corporations (ski areas, industrial forests, resorts). The federal land management agencies administer over 650 million acres (265 million hectares) open for outdoor recreation, while there are more than 11 million acres (4.5 million hectares) of state lands and over 180 million acres (73 million hectares) of private lands open or leased for outdoor recreation.

About 95% of all Americans 16 years and older participated in outdoor recreation activities in 1994–1995, and participation has been increasing for some time. In 1960, for example, 13 million people 12 years and older participated in camping, and by 1995 there were 58.5 million campers. Over the same period, participation in snow skiing increased from 2.6 million to 26.2 million. Projections of participation in many activities for the next 50 years suggest continued dramatic growth in forest recreation. Outdoor recreation has become so popular in certain areas that some national parks and forest areas have weekend use within their boundaries that equals the population of small cities.

Management. Such growth in outdoor recreation participation requires fairly sophisticated management attention and techniques. In the early decades of the twentieth century, recreation management was mostly custodial management, keeping areas free from litter and pollutants and providing fire protection. As recreation use increased dramatically in the 1950s, a concept of activity-based management emerged with a focus on the numbers of users, the activities in which they participated, and the sites necessary for participation. In 1958 the U.S. Congress empaneled a commission to assess the situation and make recommendations for the future of outdoor recreation. It was called the Outdoor Recreation Resources Review Commission. The commission's work set the stage for significant investment in recreation facilities and areas, in recreation education and research, and in policy development focused on outdoor recreation at national and state levels. *See* FOREST MANAGEMENT.

As outdoor recreation use of forests grew and as research began to focus on behavior of participants, the activity-based management concept evolved into a more sophisticated concept of experience-based management that focused on the achievement of satisfying experiences from recreation engagement. Opportunities for experiences became the objective of management, and sites and areas are now managed to provide opportunities for a variety of experiences. Three important concepts that have come from experience-based management are a definition of recreation that is behaviorally based, the Recreation Opportunity Spectrum (ROS) approach to recreation classification and management, and the Limits of Acceptable Change (LAC) planning system. Experience-based management stems from the concept that recreation as a human experience is an intrinsically rewarding voluntary engagement during nonobligated time. The key change that this concept brought to recreation management was that attention to what recreationists experience is critical to their satisfaction. It suggests that management of places (sites and areas) and situations (numbers and behaviors of other recreationists, level of facility development) influences the experiences realized by participants.

In its most popular form, the Recreation Opportunity Spectrum visualizes the array of recreation opportunities possible, ranging from modern-urban to those described as primitive and unconfined. Most forest recreation occurs in the part of the spectrum that includes categories such as roaded-modified, roaded-natural, semiprimitive motorized and nonmotorized, and primitive, thus characterizing recreation that takes place in forest areas where environments have been modified, to recreation in wilderness and other remote and relatively unmodified areas. Using specific criteria to define different types of outdoor recreation, the Recreation Opportunity Spectrum allows managers to characterize the kind of recreation

to be offered and to guide management decisions. It also is used to assess the impacts of recreation and proposed recreation on other uses such as timber management.

The Limits of Acceptable Change has replaced an earlier notion of recreation-carrying capacity that dealt with how much use is too much, and has focused attention on acceptable conditions of both the forest and the social situation. It incorporates the Recreation Opportunity Spectrum and explicitly makes monitoring and evaluation of use and impacts a part of planning and management. While the Limits of Acceptable Change has been mostly applied to wilderness recreation, it and its various agency-specific versions are applicable to virtually all forest recreation situations. The Limits of Acceptable Change is important in integrating forest recreation issues and concerns into multiple-use land management.

Quite recently, management has evolved to focus on the potential benefits that recreationists, society, and communities might realize from recreation opportunities and participation. This has become known as benefits-based management, and expected benefits are used to guide design of recreation sites and their access systems, to develop recreation information programs, and to integrate recreation with other forest uses. Benefits-based management is an outgrowth of experience-based management that extends the consideration of impacts on recreationists (the kinds of experiences they expect and realize) to impacts on communities and society as well.

Effects on ecosystems. The benefits-based management paradigm and the Recreation Opportunity Spectrum framework for recreation planning and management allow consideration of the relationship of forest recreation to ecosystem management. Ecosystem management deals with large-scale, collaborative and integrated management, and the spatial features of the Recreation Opportunity Spectrum ideally fit such a landscape view of management that cuts across ownership boundaries and vegetative and other characteristics of landscapes. It permits the assessment of the impact of recreation on the environment and on its other potential uses, as well as permitting assessment of the impacts of these other uses on potential recreation opportunities. Through its evaluative factors such as type of access, degree of management intrusion, and remoteness from sights and sounds of human activity, current recreation opportunities can be characterized, and the management actions necessary for moving toward desired recreation opportunities can be determined. Likewise, the impacts of changes in these factors due to natural or human-caused actions can be assessed for their effect on recreation opportunities.

Recreation is not neutral in its impact on ecosystems. The educational, physical, and mental health effects on individuals engaged in recreation on forest lands can have positive effects on sustainability of environmental and natural resources. Likewise, recreation can be detrimental. Wildlife can be adversely affected, as in some instances of harassment by off-road vehicle users, and of traffic on recreational road. Campsites and campgrounds can be expanded, resulting in the loss of vegetation and soil permeability. Trees and other vegetation can be chopped down for firewood. Trails and roads can be impacted by use such that they fail and lead to problems of soil erosion and stream and lake pollution. Campers can introduce soap and other pollutants into lakes and streams. The sheer volume of recreational use means that people cover much of the landscape with recreational activity, and without careful and considered use, significant impacts to the environment can occur. A major challenge for forestland managers is to help people achieve their recreational goals, but in ways that minimize negative impacts on ecosystems. Many activities, from educational programs to the design of facilities, can be effective in this work. The benefits-based management paradigm helps focus on the benefits that are to be realized and the appropriate supportive management techniques.

Future. New equipment, new activities, and changes in available time will impact outdoor recreation use in the twenty-first century. Growth is expected in participation in most activities, but for a few such as hunting, participation is expected to decrease. High-adventure activities and extreme sports are expected to increase dramatically. The use of outfitters and others to package recreation opportunities for those with limited time for planning also is likely to increase. Finally, forest recreation managers will develop ways to more fully integrate recreation with other land management activities to meet the needs of recreationists. *See* FOREST AND FORESTRY; FOREST ECOSYSTEM. Perry J. Brown

Bibliography. H. K. Cordell, *Outdoor Recreation in American Life: A National Assessment of Demand and Supply Trends*, Sagamore Publishing, Champaign, IL, 1999; R. E. Manning, *Studies in Outdoor Recreation: Search and Research for Satisfication*, 2d ed., Oregon State University Press, Corvallis, 1999; G. W. Sharpe et al., *Introduction to Forest and Renewable Resources*, 6th ed., McGraw-Hill, New York, 1995.

Forest soil

Weathering of underlying rock, climatic factors (such as rain,wind, and temperature fluctuations), living organisms, topography, and disturbances (such as glacial action, flooding, erosion, or agriculture) interact over time to produce soil. Soil helps sustain the productivity and diversity of terrestrial life by serving as the source of nutrients and water for plants, by physically anchoring plants, and by serving as the habitat for soil organisms. Far more numbers and species of organisms live belowground than aboveground. Because soil usually develops quite slowly, soil conservation is important for maintaining the

productivity of the land. Soil also acts as a water purifier; organic substances and most pollutants are arrested in upper soil horizons, purifying water before it enters streams or aquifers. Soil is the repository of life; as plants and animals die, they become part of the soil. Soil profiles reveal much about the soil history, including plant and animal remains, disturbances such as flooding, volcanic ash deposits, windthrow events, rotting wood, charcoal from fires, and past human activities.

Soil types. Variations in the bedrock and other types of parent material, climate, types of plants, topography, and disturbance history result in a wide range of soil conditions. The primary minerals that initially constitute bedrock are created deep below the surface under high pressure and temperature. As these rocks are exposed to the atmosphere through geological processes, they become thermodynamically less stable. Physical and chemical breakdown (weathering) creates products, mostly organic matter and secondary minerals, common as soil constituents. Plants can speed weathering and soil development by producing acids and organic chelates (organic compounds that speed weathering by reacting with end products to form organomineral complexes) that degrade primary minerals. With abundant rainfall and anion carriers such as bicarbonate and nitrate ions, some products of weathering move in solution through soils to stream water.

Forest soils differ from agricultural soils mainly by history and types of disturbance and topography and climate. Plowing, which homogenizes the upper soil layers, is feasible only on nonmountainous terrain. Areas too dry to support forests are sometimes farmed by irrigating soils. Soil disturbances not found in agriculture but common in forests include windthrow, landslides, and log-yarding practices. These disturbances vary much more in time and space than do agricultural disturbances, making forest soils much more heterogeneous. Erosion can be extreme in steep terrain, especially after fire removes plant cover and heavy rain follows. In general, though, disturbances of forest soil are far less frequent than of agricultural soil, allowing distinctive horizons to develop. *See* AGRICULTURAL SOIL AND CROP PRACTICES.

Four common forest-soil horizonation patterns are spodosols (podzols or mor soils), mollisols (brown or mull soils), inceptisols, histisols (organic soils), and oxisols (weathered tropical soils), although many others are recognized by various classification systems. Spodosols are found mostly at mid to high latitudes under cool, moist climates, most often with conifers. They are characterized usually by clearly visible horizons: a thick organic (O) layer of partly to highly decomposed plant litter, a highly leached, light-colored (E) layer at the top of the mineral soil overlying several (such as Bh or Bs) layers of accumulated products of weathering, and overlying unaltered parent material (C). Acids move downward, weathering the upper E horizon and depositing the products of weathering in the B horizons. Slightly warmer conditions; deciduous plants and grasses; frequent mixing of the organic and upper mineral soil layer by earthworms, insects, and other soil animals; and other disturbances can create mollisols. Inceptisols are soils with less organic matter accumulation than in podzols or mollisols, because of climate, parent material, and mixing. Lack of mixing and very wet conditions can lead to histosols, where plants root only in deep organic horizons. Oxisols are very old soils with little erosion, found on flat terrain in the tropics, where weathering proceeds so far that even silicates are lost, leaving mainly iron and aluminum oxides. These soils can become quite infertile, although some tropical trees can root to distances of more than 30 m (100 ft) to less weathered layers.

Cycling of carbon, nutrients, and water. Carbon, nutrient elements, and water move and are chemically changed in highly interrelated and cyclical patterns. These cycles can be rapid or slow, and over small to large spatial scales. As forest plants, microbes, and animals grow, they alter the soil. In the carbon cycle, autotrophs (green plants), through photosynthesis, use solar energy to capture carbon dioxide (CO_2) from the atmosphere and convert it to sugars and carbohydrates (especially cellulose and lignin). Soil heterotrophs (fungi, bacteria, and soil fauna) live by consuming live roots or decomposing dead litter. As organic matter is decomposed, it passes through many forms, usually ending as carbon dioxide released back to the atmosphere, and some quite stable carbon compounds, referred to as humus. As organic matter in mineral soil increases, generally so does plant growth. Organic matter increases water infiltration, water-holding capacity, and available nutrients in exchange sites and reduces erosion. Thus, carbon cycling influences how fast rain can infiltrate into soil and how long it can be held there before being adsorbed by plants and then transpired back to the atmosphere.

Most nutrient elements, except nitrogen and sometimes sulfur, originate as part of primary minerals in silicate rocks. Elements are released from primary minerals through physical and chemical weathering. Physical weathering results, for example, from glaciers that grind rocks beneath them, and freeze-thaw cycles that shatter rocks. Chemical weathering is in part driven by carbon dioxide, released by respiring roots and heterotrophs in soil, and dissolves in water to form carbonic acid. Certain plants and microbes greatly speed weathering by exuding organic acids and chelates. Some of the nutrient elements released by weathering are held on exchange sites where they are available for plant uptake. Elements incorporated into plant biomass are then cycled with carbon, although many plants extract nutrients out of senescing leaves before they fall to the ground. Elements such as potassium, calcium, and magnesium tend to be exported easily with anions in solution; others such as iron and aluminum tend to be less mobile and accumulate over time. Eventually, most elements are dissolved or carried as sediment into rivers and oceans; they are deposited on the

ocean bottom and eventually reformed into primary minerals.

Nitrogen differs from other nutrients by being rare in rocks and common as a gas (molecular N_2) in the atmosphere. Only certain bacteria can form a special enzyme (nitrogenase) which, with substantial consumption of carbohydrates, breaks apart N_2 to form amino acids and proteins. In nature, free-living N_2-fixing microbes and a few plants that can harbor N_2-fixing bacteria in root nodules (called symbiotic N_2-fixing plants, such as the legumes and alders) play very important roles in controlling the long-term productivity of forests limited by nitrogen supply. Because nitrate (NO_3^-) is an easily soluble anion, it moves freely in soil and can be lost by leaching when plant roots are not active. Bacteria that convert ammonium (NH_4^+) to NO_3^- [known as nitrifiers] and bacteria that convert NO_3^- back to dinitrogen (N_2) [known as denitrifiers] are important in nitrogen cycling as well.

Shrubs and minor species can have especially important effects on carbon and nutrient cycling. Shrubs, for example, tend not to be able to accumulate carbon in woody tissues as well as trees do. Therefore, more carbon fixed by photosynthesis is returned as litter. Deciduous trees also tend to return more litter to the soil than do evergreen trees. By doing so, shrubs and hardwoods speed many nutrient cycles and may help build soil organic matter. Different species result in different amounts and types of soil organic matter. Conifers tend to develop thick organic horizons and are more often associated with podzol-type soils; hardwoods are more often found on mull soils with thin litter layers. Plants that establish after soil is disturbed are often called pioneering or early-successional plants. By fixing N_2 and increasing the rate of weathering, they tend to alter soils more quickly than do late-successional plants. Other plants have adapted to wet soils and can tolerate and grow in low-O_2 soils. A few can even oxygenate their root zones.

Rates of soil formation. Soil forming on hard bedrock such as granite takes thousands of years. Lichens and bacteria slowly etch the rock surface, creating a weathering rind. Specially adapted higher plants become established in cracks, and eventually a soil is created. Soils can develop on newly exposed sand dunes, glacial soils, mudflows, and landslides remarkably fast, forming distinctive horizons in less than 100–200 years. Soils can also develop on some sedimentary bedrock quickly in the humid tropics. Disturbances such as windthrow and landslides and deep-rooting plants can speed weathering and create deeper soils over time. Erosion actually increases plant productivity in older soils by removing highly weathered layers, giving plants access to unweathered minerals and the nutrients they contain.

Soil-plant interface. Many of the carbon and nutrient cycling reactions are at the soil-root interface. Plants or microbes, to maintain ionic balance, must release a proton or hydroxyl ion into solution to be able to take up a similarly charged nutrient ion. Some plants and microbes release organic acids that act to speed weathering. Plants and microbes can also release enzymes, such as phosphatase, that can react with soil organic matter, releasing phosphorus into soil solution. Higher plants differ from lower plants and free-living bacteria and fungi by possessing the ability to extend roots deep into the soil, creating a larger surface area for nutrient and water uptake. A special symbiosis between most plants and some fungi called mycorrhizae allows plants to access an even larger surface area from which they draw nutrients and water. One form, ectomycorrhizae, creates a fungal sheath around small roots and has mycelia that extend inward between cortical root cells (where carbon substrates are provided by the plant) and outward into the soil where mycelia absorb nutrients and water. Mycelia and their hyphae are much finer than roots and can grow into 5-micrometer crevices, expanding the effective volume of soil explored by plant roots. Plants also exude carbohydrates into the soil and alter the microbial composition and activities in soil adjacent to roots, in a zone called the rhizosphere. *See* MYCORRHIZAE; ROOT (BOTANY).

Effects of forest disturbances. Disturbance of forest stands affects soils to different extents. Some stand-replacing disturbances where new soil is deposited, such as volcanoes, glaciers, flooding, and massive landslides, can create new soil surfaces. These soil surfaces may be poorly productive at first because of little organic matter and available nutrients, but can quickly become productive and even exceed productivity of older, undisturbed soils. Fires that lead to deep erosion of upper soil layers, exposing poorly weatherable bedrock such as granite, may take thousands of years to develop a new soil. Landslides exposing more weatherable sedimentary bedrock in the wet tropics can become soil in less than a hundred years. Soil mixing by windthrow, landslide, flooding, or dust deposit is likely to increase forest productivity by preventing nutrient immobilization in wet organic horizons.

Harvesting trees affects soils by reducing evapotranspiration and water flow through soil, and decreasing nutrient uptake. Nutrients, especially NO_3^-, can be quickly lost to stream water when plant roots are not active. Harvesting of stems and other tree parts also removes nutrients from the site. Short crop rotations and whole-tree harvesting techniques are likely to lead to nutrient exports exceeding the capacity of subsequent trees to obtain nutrients, and fertilization or crop rotation with N_2-fixing plants may be required to maintain growth rates. Hot crown and slash fires can lead to massive volatilization of nitrogen and other nutrients, especially if followed by heavy rains. Generally, light underburning of the litter layer removes far fewer nutrients and may even increase the availability of nutrients tied up in organic horizons. Heavy harvesting and yarding equipment may also cause compaction and erosion, which can reduce tree growth. This concern has led to elaborate cable-yarding systems to avoid ground-disturbing equipment, especially on steep slopes. Quick reestablishment of pioneer species as well

as crop trees is important to reestablish biotic regulation of carbon, nutrient, and water cycles. *See* ACID RAIN; FOREST AND FORESTRY; FOREST ECOSYSTEM; SOIL; SOIL ECOLOGY. Bernard T. Bormann

Forest timber resources

The timber produced by temperate, boreal, and tropical forests around the world. Timber is included in a wide variety of industrial and other commodities that are traded in national and international markets.

United States resources. Of the total land area of the United States, forests cover one-third, that is, about 737 million acres (295 million hectares). Approximately two-thirds of the total forest area in the United States is classified as timberland, that is, forestland capable of producing at least 20 ft^3 (0.6 m^3) of industrial wood per acre per year, and not reserved for uses incompatible with timber production. The remaining forest area meets the growth criteria for timberland, but has been set aside for parks and wilderness areas or is incapable of producing a sustained crop of industrial wood.

Most nontimberland is in public ownership, including 36 million productive forest acres (15 million hectares) legally withdrawn for such uses as national parks, state parks, and national forest wilderness areas. Of the remaining nontimberland, about 107 million acres (43 million hectares) are in Alaska (part of the Pacific Coast region), and much of this acreage is boreal forest as opposed to the temperate forest types prevalent in most of the United States.

Between 1952 and 1995, the area of timberland had a net decrease of 4%. Much of the decline was the result of withdrawals of public timberland for wilderness or for other land uses that do not permit timber harvest. About half of the decline was offset by conversions of abandoned cropland to softwood timber types, especially in the southern United States. From 1985 to 1995, the timberland area remained fairly stable. There was, however, some reclassification from timberland to other forest areas as a result of reevaluation of site productivity, and some locally significant loss of timberland to urbanization. In the southern United States, the loss of timberland to agricultural uses was stemmed in part by reductions in agricultural production.

Forest types. Most timberland in the United States is in the temperate zone. The broad distinction made among temperate tree species is softwoods (the conifers, or cone-bearing trees, such as pine, fir, and spruce) versus hardwoods (the broad-leafed trees, such as maple, birch, oak, hickory, and aspen). The five largest timber types are oak-hickory, southern pine, maple-birch-beech, Douglas-fir, and oak-pine.

In the east, hardwood types comprise nearly two-thirds of the forest area. Important eastern timber types are the oak-hickory, maple-birch-beech (found throughout the New England, Middle Atlantic, and Great Lakes regions), the oak-gum-cypress (primarily in the Mississippi Delta and other southern river bottoms), and the aspen-birch of the Great Lake states (relatively short-lived species that follow logging and fires). These forests supply much of the quality hardwoods in the United States.

Softwood types dominate the western forests. Douglas-fir and true firs and spruce constitute nearly half of the western commercial timberland. The western softwood types are a major source of lumber and plywood in the United States. Nearly all the commercial forest area of coastal Alaska is of the hemlock–Sitka spruce type. Hardwoods (mostly in Washington and Oregon) occupy only a small percentage of the commercial forest area of the western United States. *See* PLYWOOD.

Growth. Growth on United States timberlands increased 55% between 1952 and 1995, being evenly divided between softwoods and hardwoods. The fastest growth rates were in the South for both hardwoods and softwoods. In 1992, growth exceeded removals in all areas, except for softwoods in the South, where harvests increased after the late 1980s to offset reduced harvests in the Pacific Coast states because of the strategies adopted to protect the habitat of the northern spotted owl.

Timber inventory. The growing-stock inventory contained a vast amount of sound wood at the beginning of 1992. Timber inventories grew at 0.6% per year between 1955 and 1995; most of this growth was in the hardwood component.

Forestland management in the United States has produced positive results in the sense of sustaining forest resources. However, there are concerns about biodiversity, lack of certain seral stages, and loss of what is publicly perceived as nonrenewable forest resources.

The average diameter of harvested timber is a key indicator of timber utilization. For example, between 1976 and 1992 the average diameter of harvested timber fell by 35% for the Pacific Coast states, 44% for the Rockies, and 18% for the South. In the northern United States, the average diameter increased over the same period by 5% for both softwoods and hardwoods.

Ownership. Nearly three-quarters of timberland is privately owned. The remainder is held by various public owners, and comprises national forests and state, county, and municipal holdings. Timberland under Native American sovereignty accounts for about 1% and is included with other private timberlands. Private owners can be distinguished as industrial, farm, or miscellaneous. Approximately 14% of the commercial forestland is in the hands of industry, the remainder being divided between miscellaneous private owners and farmers.

Some of the most productive forestlands of the United States are in private industrial holdings. These lands are generally managed to produce high volumes of wood fiber for both the solid-wood-products and pulp and paper industries. Many timberland forest industries are concentrated in the southern states.

The miscellaneous category, which holds 34% of the privately owned forestland, consists of a great variety of individuals, groups, and non-forest-products companies. These owners range from those who are interested in producing timber to those whose objectives are nontimber. Most farm forests are managed for timber production. As such, the condition of farm-owned woodlands has been a source of disappointment, discussion, and considerable action on the part of public conservation agencies.

A diverse array of public agencies hold large forest areas, the most important being the national forests. These comprise 89 million acres (36 million hectares) of commercial forestland that is managed and administrated by the U.S. Forest Service, a bureau of the Department of Agriculture. Most public holdings are managed under multiple-use principles. There has been a dramatic shift in focus away from wood as the principal product of the national forests to sustaining ecosystems where wood is just one of the outputs.

Inventories on public timberlands have been stable since 1955—largely because withdrawals for wilderness and other uses offset inventory gains attributed to increases in growth. Private inventories, however, grew by 51% over the same period. Softwood inventories are roughly evenly divided between public and private timberland owners, but the hardwood inventory is concentrated on private timberland, especially nonindustrial ownerships.

World forest resources. Forests cover just 30% of the land area of the Earth. Included are closed and open natural forest, forest plantations, and other woodland. Shrub, scrub, and brushland cover about one-fifth of the forest area.

Half of the world's forests are in Asia, Latin America, and Africa. These areas comprise predominantly developing countries, and the forests are mostly tropical. About one-quarter of the world's forests, nearly half of the tropical forests, and more than half of the wet tropical forests (rainforest and moist lowland forest) are in Latin America. Just over half of the world's forests are in temperate zones; the boreal region in the temperate zone accounts for more than one-third of the world's forests.

Between 1980 and 1990, the worldwide area of forest declined by 4%. Deforestation across all regions occurred at rates estimated to be 0.3–0.6% per year between 1975 and 1995. An increase in the area of temperate forest over the period 1980–1990 was much more than offset by a reduction in the area of tropical forest. Lowland tropical forests account for the majority of tropical forest area and tropical deforestation; however, the rate of deforestation is most severe in the dry tropical forests, where population pressure and environmental stress are combined.

About two-thirds of the world's forests are publicly owned. Although slightly more than half of temperate forests are in public ownership, nearly all tropical forests are in public or communal ownership. The United States, Japan, and the Nordic countries have relatively higher proportions of private ownership than other developed countries. Ownership patterns affect the mix of forest outputs and the type of management practices. Although public ownership is typically thought to best assure provision of nonmarket goods and intangible forest benefits, private management is recognized as efficiently providing commodity-based forest outputs.

Plantation forests. Although plantations are estimated to be little more than 3% of world's forest area, they are an important component of industrial timber production in the temperate zone, and are a source of both industrial timber and nonindustrial products in the tropical zone.

Temperate-zone plantations are predominantly coniferous species; most are managed for industrial wood products, and are the consequence of reforestation rather than afforestation. That is, these plantations generally do not expand existing forest area. Afforestation of idle crop or grazing land (as in the United States and United Kingdom) is an exception. Native species account for the majority of the plantation area in most Northern Hemisphere temperate-zone countries. Growth rates may be somewhat higher than growth rates for natural forests as a result of the use of improved planting stock or more intensive management practices. *See* REFORESTATION.

In the temperate zone of the Southern Hemisphere, most plantations are fast-growing, exotic species, with a shorter production cycle than native species. During 1985–1995, production from exotic softwood plantations in New Zealand, Chile, and Australia increased significantly; these plantations account for a small percentage of world coniferous industrial roundwood production.

India, Indonesia, and Brazil account for the majority of plantations in the tropical zone. More than half of the plantations in the tropical zone were established for purposes other than timber production; these include plantations for erosion control and for agroforestry (that is, joint production of trees and agricultural crops). Tree plantations established for annual crop production in the tropical zone (such as rubber, coconut, and oil palm) also have contributed to the supply of wood and wood products. *See* FOREST AND FORESTRY; FOREST MANAGEMENT.

Richard W. Haynes; David J. Brooks

Forestry, urban

The planning and implementation of actions to establish, protect, restore, and maintain trees and forests in cities and smaller communities. When it began in the 1960s, the field of urban forestry focused on individual trees along streets and adjacent to homes and buildings, and on groups of trees in specific spatial areas such as parks. Since then, it has evolved toward more holistic consideration of the structure, processes, and functions of urban ecosystems at a larger scale. Scientists have come to realize that trees and forests play a critical role in maintaining

healthy urban ecosystems, providing ecological services such as filtering water and air pollution, reducing stormwater flows, sequestering carbon emissions, conserving energy, and reducing soil erosion, as well as providing human health benefits, recreation, esthetics, and fish and wildlife habitat. These ecological services can be translated into economic values worth billions of dollars by comparing them to the costs of achieving the same benefits with technology and human-made infrastructure. Urban trees, forests, and related vegetation are often referred to as green infrastructure in order to compare and contrast them with human-made or hard infrastructure, such as buildings and roads. *See* ECOLOGICAL COMMUNITIES; ECOLOGICAL SUCCESSION.

Urban trees and forests also provide significant social and economic benefits. By making downtown and residential areas more livable, urban trees and forests can increase real-estate values and attract businesses; in turn, this helps to create additional employment opportunities and contributes to the overall economic viability of the community. More livable urban environments strengthen the social fabric of communities, creating more opportunity for citizen involvement in community activities, greater pride and safety in neighborhoods, and conditions that lead to greater mental and physical health within the community. Engaging community groups in urban forestry efforts can also build a stronger sense of community, empower inner-city residents to improve neighborhood conditions, and promote environmental stewardship.

Just as communities depend on healthy urban forests for ecological services, urban forests depend on socially and economically healthy communities for their care. Urban areas that lose trees and forests often go into decline, becoming impoverished environmentally, socially, and economically. Where urban areas flourish, community groups have developed a sense of civic pride and the social and economic capacity to maintain healthy trees and forests.

New technological tools. New satellite remote sensing and geographic information system tools are enabling communities to assess changes in their green infrastructure and to develop plans for restoring and maintaining urban trees and forests. City planners and policymakers use the same tools to address other urban infrastructure needs, such as transportation and housing, which allows them to integrate green infrastructure information into plans for other infrastructure development.

Recent assessments that use tree cover as an indicator of ecosystem health show that urban forests are an overlooked and underappreciated community resource. Since the early 1970s, three major metropolitan areas in the United States (Seattle, Baltimore/Washington, and Atlanta) have lost one-third or more "heavy tree cover" (50% or more). These losses, and the parallel expansion of areas with "low tree cover" (less than 20%), result in billions of dollars of lost ecological services. In the 3.9-million-acre Puget Sound area, heavy tree cover dropped by 37% between 1972 and 1996, whereas low tree cover more than doubled. When translated into economic values, this loss in tree cover and increase in impervious surface area resulted in $2.4 billion in lost stormwater retention capacity and $95 million annually in the capacity of trees to filter air pollution. Cities that recognize the economic values associated with green infrastructure can slow and reverse losses in urban trees and forests by incorporating green infrastructure into urban planning and development efforts, and by investing in plans and programs to restore and maintain urban trees and forests. Incentive-based policies may encourage private investment in urban trees and forests.

Urban and rural links. As urban forestry has expanded to include larger ecosystems, spatial boundaries between urban and rural areas have begun to blur. Their ecological, social, and economic links have come into view. Watershed connections have become a key consideration, while other social and economic links related to movements of people, businesses, goods, and services have also become clearer. Rural lands, both public and private, provide an array of ecological services to nearby cities, including drinking water, recreational opportunities, fish and wildlife habitat, and agricultural and forest products.

Sprawling urban development—the effects of people and businesses moving to areas surrounding cities—is a major policy concern with significant implications for urban forestry. Because of the expanded interpretation of urban forestry, policymakers should consider the different impacts of sprawl across the landscape. Older inner-city areas, newer communities on the urban fringe, and nearby public and private rural lands face different challenges to their green infrastructure:

1. Mature parts of cities are experiencing the movement of people and public and private financial resources to rapidly growing areas. Urban forestry efforts focus on protecting the often-limited green infrastructure and restoring degraded or abandoned (such as brownfields) sites with trees and forests. Older cities look for ways to invest in rebuilding green infrastructure to improve ecosystem functioning and social and economic well-being.

2. Newer parts of cities, often suburbs or urban fringe areas, are experiencing the greatest loss of trees and forests from sprawl because of massive public and private investment in new residential neighborhoods, public infrastructure, and private businesses. Urban forestry efforts focus on retaining critical parts of existing green infrastructure. Newer cities seek planning tools and policy mechanisms to retain significant green infrastructure in development plans and actions.

3. Sprawl has major impacts on public and private rural landscapes. Individual decisions and uncoordinated actions affect these landscapes by reducing the amount of open space, curtailing opportunities for traditional farm and forest management, and

threatening overall ecosystem health. Of particular concern is forest fragmentation, caused by the conversion of private farms and forests to other land uses and the breakup of large private ownerships into multiple smaller parcels. Policy efforts focus on retaining private lands as open space for public environmental benefits and on maintaining working lands for ecological, social, and economic benefits. As urban growth nears public forest lands (generally federal or state lands), the management of these lands often changes to meet the use preferences of urban people. National forests near major metropolitan areas have come to be called "urban national forests" and generally emphasize recreational uses and the protection of ecological services, a shift from traditional timber-oriented management.

Management. Urban forests are owned by a diverse array of public and private entities, including individual homeowners, private businesses, and federal, state, and local governments. Their management and use is closely tied to the homes, buildings, transportation systems, utilities, and other infrastructure in an urban area. A key challenge is coordinating the efforts of the large number of individuals, groups, and organizations involved in urban forestry, each with its own information, resources, and objectives. *See* LAND-USE PLANNING.

Generally, city governments guide urban forestry efforts, although their direct management responsibility focuses on city-owned lands. They receive assistance from federal and state governments and provide assistance to other landowners in the city. Because many local government agencies have jurisdiction over activities that impact the urban forest, including environmental protection, transportation, housing, and public works, there is a great need for local governments to coordinate agency planning and development activities. Most cities have some narrow policies and programs focused on planting and protecting trees, but most have not recognized the ecological, social, and economic benefits of a healthy green infrastructure through policies and programs that integrate the efforts of various local agencies.

The interest of national policymakers in urban forestry and green infrastructure has grown in recent years, as demonstrated by increasing federal support and policy initiatives. In many cities, however, financial support for urban forestry has fluctuated and often declined as local governments and policymakers have dealt with a broad array of social and economic issues. Until the general public and local policymakers understand the benefits of green infrastructure and the interdependence of healthy forests and communities, urban forestry will have difficulty building support as a policy priority of city governments.

Many professionals are involved in urban forestry efforts, including foresters, arborists, landscape architects, horticulturists, engineers, planners, wildlife biologists, and hydrologists. Planning and implementing urban forestry projects in difficult city environments requires a variety of technical and people-oriented skills. Educational programs should incorporate training in policy, social sciences, and communication, as well as the natural sciences, planning, and engineering.

The involvement of community groups and citizen volunteers is an essential element of urban forestry. Many groups plant and maintain urban trees and forests. These efforts might focus on neighborhood enhancement, stream restoration, or the transformation of brownfields into urban gardens or parks. Volunteers and community groups working in partnership with agencies and other professionals can improve urban ecosystems, community livability, and social capacity.

Outlook. Three major challenges presently face urban forestry: (1) to develop better understanding of the ecological services that urban trees and forests provide, (2) to integrate green infrastructure into city governments' planning and development efforts, and (3) to stimulate public and private investment in actions to restore and maintain urban trees and forests. To achieve these goals, the general public will need to become more aware of the interdependence between urban forests and communities, and more people will need to become involved in urban forestry activities. *See* FOREST AND FORESTRY.

Gerald J. Gray; Gary A. Moll

Bibliography. *A National Research Agenda for Urban Forestry in the 1990's*, International Society of Arboriculture, Urbana, IL, 1991; J. F. Dwyer, The human dimensions or urban forest ecosystem management, *Proceedings of the 7th National Urban Forest Conference*, pp. 44-46, American Forests, Washington, DC, 1995; G. J. Gray, Sustainable communities: Exploring uran-rural linkages, *Proceedings from the 1999 National Urban Forest Conference*, pp. 86-89, American Forests, 1999; G. Moll and S. Young, *Growing Greener Cities*, Living Planet Press, Washington, DC, 1992; *Regional Ecosystem Analysis: Chesapeake Bay Region and the Baltimore-Washington Corridor*, American Forests, 1999; C. Summers, Incorporating natural systems into urban infrastructure, *Proceedings of the 7th National Urban Forest Conference*, pp. 54-56, American Forests, Washington, DC, 1995.

Forging

The plastic deformation of metals, usually at elevated temperatures, into desired shapes by compressive forces exerted through a die. Forging processes are usually classified either by the type of equipment used or by the geometry of the end product. The simplest forging operation is upsetting, which is carried out by compressing the metal between two flat parallel platens. From this simple operation, the process can be developed into more complicated geometries with the use of dies. A number of variables are involved in forging; among major ones are properties of the workpiece and die materials, temperature,

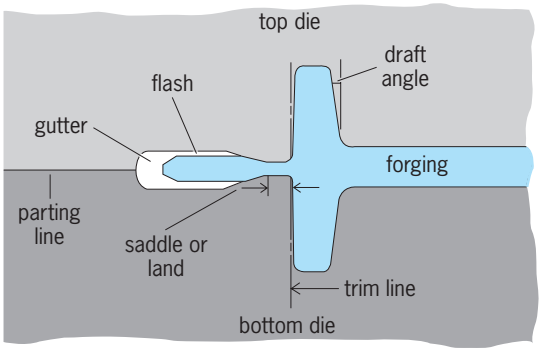

Fig. 1. Closed-die forging terminology.

friction, speed of deformation, die geometry, and dimensions of the workpiece. One basic principle in forging is the fact that the material flows in the direction of least resistance.

Friction between dies and workpiece (and, in hot forging, the more rapid cooling of the surfaces of the workpiece that are in contact with relatively cooler dies) causes barreling. Barreling can be reduced by lowering friction and also by carrying out the forging process in a more isothermal condition.

Forgeability. In practice, forgeability is related to the material's strength, ductility, and friction. Because of the great number of factors involved, no standard forgeability test has been devised. For steels, which constitute the majority of forgings, torsion tests at elevated temperatures are the most predictable; the greater the number of twists of a round rod before failure, the greater is its ability to be forged. A number of other tests, such as simple upsetting, tension, bending, and impact tests, have also been used.

In terms of factors such as ductility, strength, temperature, friction, and quality of forging, various engineering materials can be listed as follows in order of decreasing forgeability: aluminum alloys, magnesium alloys, copper alloys, carbon and low-alloy steels, stainless steels, titanium alloys, iron-base superalloys, cobalt-base superalloys, columbium alloys, tantalum alloys, molybdenum alloys, nickel-base superalloys, tungsten alloys, and beryllium. *See* METAL, MECHANICAL PROPERTIES OF.

Forging dies. Design and selection of die materials for forging involve considerable experience. Both static and impact strength are required, in addition to resistance to abrasion and to heat checking. Dies are generally heated to below 1000°F (540°C) in hot forging.

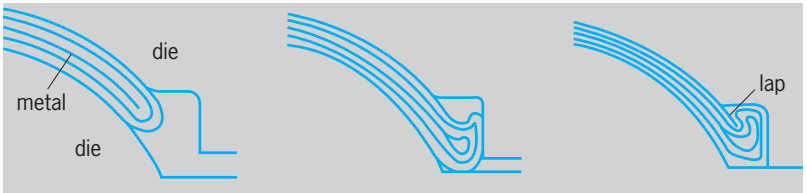

Fig. 2. Formation of a defect in closed-die forging.

Some of the terminology in forging is shown in **Fig. 1**. Draft angles facilitate the removal of the forging from the die cavity; they usually range between 3 and 10°. The purpose of the saddle or land in the flash gap is to offer resistance to the lateral flow of the material so that die filling is encouraged. Die filling increases as the ratio of land width to thickness increases up to about 5; larger ratios do not increase filling substantially and are undesirable due to increased forging loads and excessive die wear. The purpose of the gutter is to store excess metal. The flash is removed either by cold or hot trimming or by machining.

Filling of the die cavities is an important aspect and requires a careful balance of die geometry, temperature, lubrication, and speed of forging. Lubrication is important not only for reducing friction but also in its role as a parting agent, in addition to serving as a heat barrier between the die and the hotter forging. The commonest lubricant is graphite suspended in grease, oil, or water; other lubricating materials are molybdenum disulfide, talc, mica, sawdust, salt water, and glasses. Some metals, such as tungsten and molybdenum, form soft oxides upon heating, and these serve as lubricants.

Because of the many factors involved, great variations in strength and ductility have been observed in the forging of the more complex shapes. Incorrect design of dies or improper operation can result in forged products with internal defects such as laps (**Fig. 2**).

Forging design. There are in general four types of forging designs: blocker type, commercial, close tolerance, and precision. The blocker type involves the use of large radii and draft angles, smooth contours, and generous allowances. Commercial designs have more refined dimensions, such as standard draft angles of 5–7°, and smaller radii and allowances. In close-tolerance designs draft angles are in the order of 1–3°, tolerances are less than half of those for commercial designs, and there is little or no allowance for finish. Precision designs usually involve close dimensional tolerances; they require special tooling and equipment and are used mainly for aluminum and magnesium forging.

Forging equipment. A number of methods produce the necessary force and die movement for forging. Two basic categories are open-die and closed-die forging. Drop hammers supply the energy through the impact of a falling weight to which the upper die is attached. Two common types of drop hammers are the board hammer and the steam hammer. In the former the ram is attached to wooden boards which slide between two rollers; after the ram falls freely on the forging, it is raised by friction between the rotating rollers. In the steam hammer the ram is raised and lowered by a double-acting steam cylinder; the forging force is a combined effect of gravity and steam pressure.

Another type of forging equipment is the mechanical press, varieties of which are crank or eccentric, knuckle, percussion, and toggle. For large forgings the hydraulic press (with existing capacities of

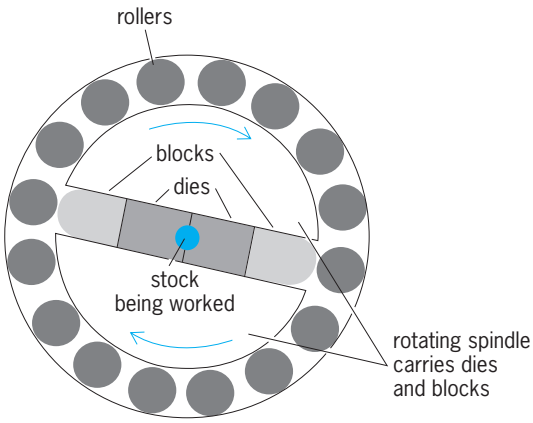

Fig. 3. Operation principle of swaging machine.

50,000 tons or 45,000 metric tons and future capacities in the order of 200,000 tons or 180,000 metric tons) is the only equipment with sufficient force. However, the speed for such presses is about one-hundredth that of hammers. Other types of forging equipment are counterblow hammers (vertical) and impacters (horizontal).

Swaging. Swaging or, as it is also called, rotary swaging consists of two or four dies which are activated radially by blocks in contact with a series of rollers (**Fig. 3**). The rotation of the block-die assembly causes the curved ends of the blocks (either circular or modified sine curve) to be in contact with the rollers; relative motion between the blocks and the roller housing then gives a reciprocating motion to the dies. The hammering action on the outer surface of the stock reduces its diameter; the stock is generally prevented from rotating.

A variety of geometries is obtained in this process: (1) reduction of wall thickness of tubes; (2) sinking of tubes, whereby both the outside and inside diameters are reduced, the final wall thickness being controlled by a mandrel; (3) tapered or stepped sections; and (4) a variety of internal geometries, such as fins or hexagonal or spiral grooves, each with the use of a specially designed mandrel. Cold swaging produces excellent surface finish and dimensional accuracy, together with improved mechanical properties. *See* METAL FORMING. Serope Kalpakjian

Bibliography. B. Avitzur, *Metal Forming*, 1980; National Safety Council, *Forging Safety Manual*, 1990; E. G. Nisbett, *Steel Forging*, STP903, 1986.

Formaldehyde

The simplest aldehyde, formula HCH═O. Because of its extreme reactivity, even with itself, it cannot be readily isolated or handled in the pure state. Therefore, it is produced and marketed as an aqueous solution (usually 37–50% formaldehyde by weight), sometimes known as Formalin. It is also sold as the solid hydrated polymer known as paraformaldehyde or paraform. Approximately 60% of all the formaldehyde produced is for captive use. *See* ALDEHYDE.

Properties. Formaldehyde undergoes many of the general reactions which are typical of aldehydes. Some interesting reactions which are somewhat specific for formaldehyde are (1) and (2). Bis-

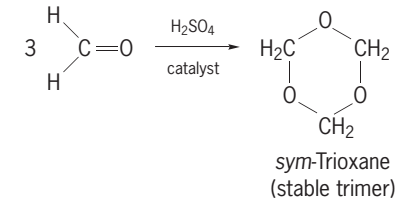

(1)

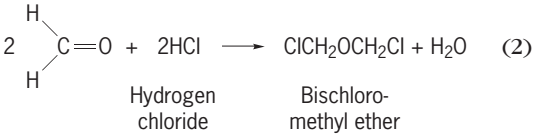

(2)

chloromethyl ether has been classified by the U.S. Occupational Safety and Health Administration as a carcinogen.

The Mannich reaction (3) involves the reaction of

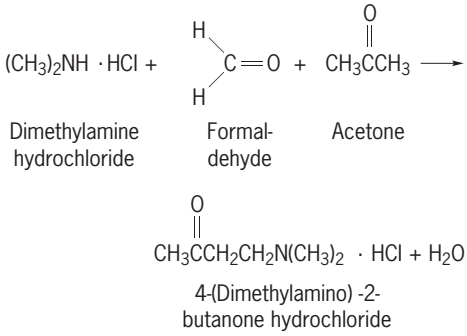

(3)

an amine hydrochloride with formaldehyde and a ketone. Reaction of formaldeyde with carbon monoxide is shown in reaction (4). Reaction (5) represents

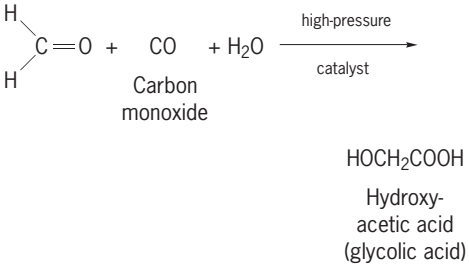

(4)

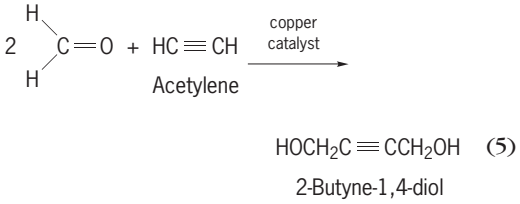

(5)

the Reppe ethynylation reaction.

The Prins reaction (6) involves the reaction of formaldehyde with olefinic aromatic hydrocarbons in the presence of an organic acid and a mineral acid catalyst.

The reaction of formaldehyde with phenols, urea, and melamine to form resins proceeds first with

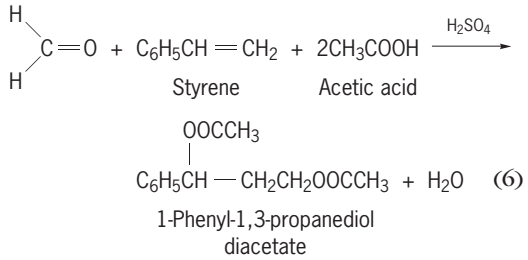

formation of methylol (—CH_2OH) derivatives followed by intermolecular dehydration to produce methylene (CH_2) linkages. This reaction is catalyzed by either alkali or acid, and the extent of resinification is controlled by temperature, reaction time, and concentration. *See* PHENOLIC RESIN.

Production. Most of the formaldehyde is manufactured by the oxidation of methanol with air over a metal catalyst in the temperature range of 450–650°C (840–1200°F). Catalysts are either silver, copper, or an iron-molybdenum mixture. The reaction is shown as (7).

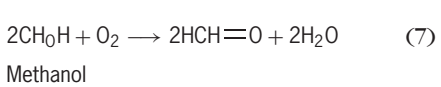

Increasing quantities of formaldehyde are produced by the partial oxidation of hydrocarbons from natural gas. Air is used as the source of oxygen, excess hydrocarbon to minimize complete oxidation, and steam to quench the reaction. The yields of formaldehyde are low, but the raw materials are plentiful and low in cost. Since other products, such as alcohols, acids, ketones, and other aldehydes, are formed, the purification of formaldehyde produced by this process may involve fractionation, liquid-liquid extraction, azeotropic distillation, and extractive distillation.

Uses. Formaldehyde is used principally to produce synthetic resins and adhesives by reaction with phenols, urea, and melamine. This use accounts for about 75% of the total production. Approximately 15% is used in the manufacture of textiles, dyes, drugs, paper, leather, photographic materials, embalming agents, disinfectants, and insecticides. It is also used for the industrial production of pentaerythritol and hexamethylenetetramine, reactions (8) and (9).

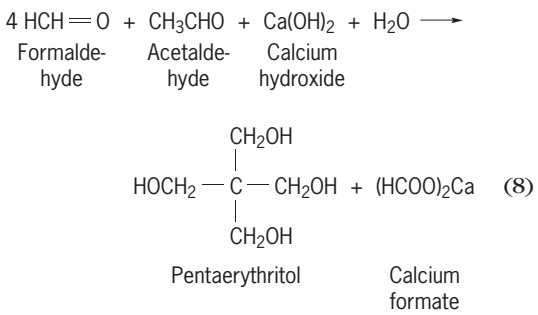

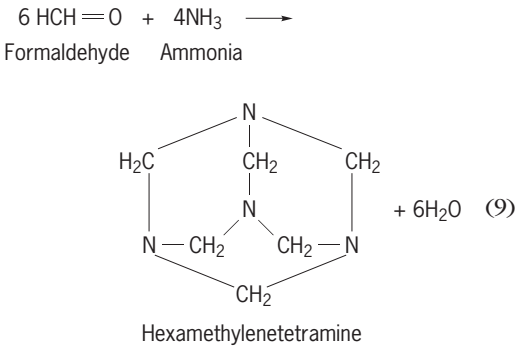

L. Madestau

Formation

A fundamental geological unit used in the description and interpretation of layered sediments, sedimentary rocks, and extrusive igneous rocks. A formation is defined on the basis of lithic characteristics and position within a stratigraphic succession. It is usually tabular or sheetlike, and is mappable at the Earth's surface or traceable in the subsurface (for example, between boreholes or in mines). Examples are readily recognized in the walls of the Grand Canyon of northern Arizona (see **illus.**). Each formation is referred to a section or locality where it is well developed (a type section), and assigned an appropriate geographic name combined with the word formation or a descriptive lithic term such as limestone, sandstone, or shale (for example, Temple Butte Formation, Hermit Shale). This usage of "formation" by geologists differs from its informal lay usage for stalactites, stalagmites, and other mineral buildups in caves. *See* STRATIGRAPHY.

Distinctive lithic characteristics used to designate formations include chemical and mineralogical composition, particle size and other textural features, primary sedimentary or volcanic structures related to processes of accumulation, fossils or other organic content, and color. Lithology may be interpreted also in boreholes from cores and cuttings produced by drilling, and from electrical, radioactive, seismic, or other properties measured by instruments lowered into the holes. Contacts or boundaries between formations are chosen at surfaces of abrupt lithic change or within zones of gradational lithic character. Commonly, these contacts correspond with recognizable changes in topographic expression, related to variations in resistance to weathering. For example, massive carbonate rocks and sandstones, such as the Redwall Limestone and Coconino Sandstone in the Grand Canyon, tend to be cliff-forming, whereas shales such as the Bright Angel form ledges or areas of low topography because clay minerals are easily weathered. Formations normally possess some degree of internal lithic homogeneity, but they may be composed of repetitions of two or more different kinds of sediments or rocks, or extreme heterogeneity may itself constitute a distinctive feature.

Different formations in a sedimentary succession may also be lithologically similar, but useful to distinguish because they are located at different stratigraphic levels. *See* SEDIMENTARY ROCKS; SEDIMENTOLOGY.

Formations tend to be tabular because most sedimentary layers are characterized by gradual lateral changes in lithic character. Exceptions are found at the seaward margins of ancient continental shelves and carbonate platforms, where abrupt lateral changes in paleowater depth are reflected by abrupt changes in lithology; and along the faulted margins of some sedimentary basins, where coarse-grained alluvial gravels or conglomerates pass laterally into comparatively fine-grained lacustrine or marine deposits.

Mappability is an essential characteristic of a formation because such units are used to delineate geological structure (faults and folds), and it is useful to be able to recognize individual formations in isolated outcrops or areas of poor exposure. Well-established formations are commonly divisible into two or more smaller-scale units termed members and beds (for example, subdivisions of the Redwall Limestone). In other cases, formations of similar lithic character or related genesis are combined into composite units termed groups and supergroups (for example, Supai Group). The rank of a named unit may vary from one area to another (for example, from group to formation) according to whether or not subunits are readily mappable. Changes in rank are also justified in the light of new geological knowledge.

Formations vary in thickness from a feather edge at a depositional or erosional limit to several kilometers. They tend to be thin where successions are characterized by abundant and easily recognized variations in lithology, and in areas that have been the focus of many decades or even centuries of geological investigations (for example, southern England). Formations tend to be thicker in successions that are characterized by limited lithic variability or that have been subjected only to reconnaissance geological mapping [for example, basinal shales, successions in remote areas, and strata of Precambrian age (older than 543 million years)].

Although commonly used as a framework for interpreting geological history, formations and related units are conceptually independent of geological time. They may represent either comparatively short or comparatively long intervals of time. Accumulation of a particular unit may have begun earlier in some places than in others, and the time span represented by a unit may be influenced by later erosion. In some cases, a formation cropping out at one locality may be entirely older or younger than the same lithic unit at another locality. Normal practice is to avoid defining formations to include regional erosion surfaces, but recent efforts to locate such discontinuities in sedimentary successions have inevitably led to the recognition of breaks in sedimentation within many of the formations defined in earlier studies.

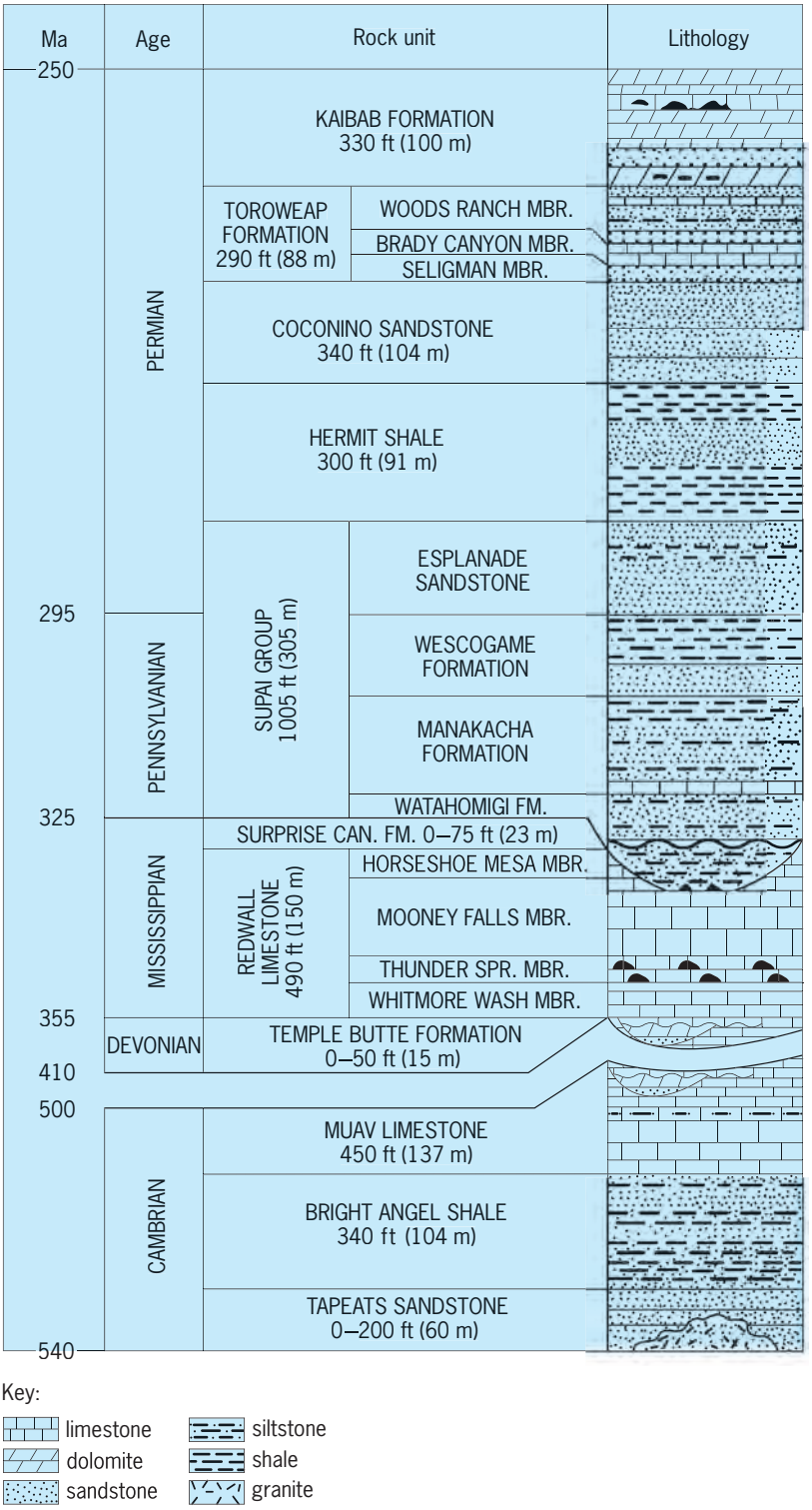

Geological units of Paleozoic age recognized in the walls of the Grand Canyon, Arizona. (*Modified from S. S. Beus, Geology along the South Kaibab Trail, eastern Grand Canyon, Arizona, in S. S. Beus, ed., Geological Society of America Centennial Field Guide, vol. 2: Rocky Mountain Section, Geological Society of America, pp. 371–378, 1987*)

Although the concept of time plays no role in the definition of a formation, evidence of age is useful in the recognition of lithologically similar units far from their type localities. Nicholas Christie-Blick

Bibliography. North American Commission on Stratigraphic Nomenclature, North American Stratigraphic Code, *Amer. Ass. Petrol. Geol. Bull.*, 67:841–875, 1987; A. Salvador (ed.), *International Stratigraphic Guide*, 2d ed., Geological Society of America, 1994.

Fossil

A record of earlier life buried in rock. Originally meaning any distinctive object that has been dug up (from Latin *fodio*, dig), the term "fossil" soon came to refer particularly to things resembling animals and plants. These objects were known before it was recognized that life on Earth has a long history. They were generally interpreted as having somehow grown in the rock or been placed there by a creator, or as being remains of organisms living somewhere else at the time in undiscovered seclusion.

In the eighteenth and nineteenth centuries, the realization grew that fossils indeed represent a record of older life, and the success of Darwin's theory of evolution provided the possibility for scientific investigation of life's history through paleontology. Fossils are now known from almost the whole time span of Earth's sedimentary record, and they provide a panoramic picture of how the biosphere has changed through time, gradually as well as in bursts of evolutionary activities and catastrophic extinctions.

Fossilization. A widespread conception is that a fossil is a shell or skeleton turned into stone. This picture is not false, but it represents only a few of the ways in which life can leave its trace in Earth's accreting skin of sedimentary rocks. Any part of an organism can get preserved, not only hard parts but also soft tissues, cells, and organelles. Even when nothing of the original organisms is preserved, impressions and traces in the sediment give important information about their former presence, activities, and ecological roles. Also, fossilization can imply everything from preservation of the almost unaltered original tissues to their complete replacement with sediment or minerals growing in place. Organic molecules can be preserved, though in a more or less degraded state.

The biosphere normally recycles all organic and inorganic matter produced by organisms; fossils represent dead individuals that to some degree escaped that process. Most decomposition is by aerobic scavengers, fungi, and bacteria, and so a prerequisite for fossilization is that the dead body is quickly and permanently subjected to an environment in which decomposers cannot be active. A combination of anoxic water and rapid sedimentation is a typical condition favorable to fossilization, though other conditions, such as extreme temperatures, salinity, poisonous environment, desiccation, or rapid mineralization, are also known to promote fossilization. Some kind of microbial activity, however, seems to be a prerequisite for many types of fossilization, particularly of soft tissue. *See* EDIACARAN BIOTA.

Fig. 1. Skull of sabre-toothed cat, *Amphimachairodus*, from the Miocene Epoch. (*Palaeontological Museum, Uppsala University*)

Shells and skeletons. Mineralized hard parts, such as shells, spicules, and bones, are by far the most common type of fossil (**Fig. 1**). Although they typically contain a substantial proportion of organic material, their mineral phase usually ensures that they are more resistant than soft tissues to biological decomposers. The most common skeletal minerals are opal (hydrated silica), apatite (calcium phosphates), and calcite and aragonite (calcium carbonates). Apatite, the common constituent of bone, is particularly stable under normal conditions; and it is often preserved crystallographically intact in fossils, though replacement and infilling of the other components of the hard tissue with minerals (permineralization) has usually taken place. Calcite is also often preserved with its original crystallography, whereas aragonite and opal tend to be recrystallized to calcite and quartz, respectively, or replaced altogether.

Even when none of the original hard tissue is preserved, its former presence may promote fossilization through its initial resistance to degradation. Shells are frequently preserved as molds or casts; for example, lithified infillings (internal molds) of mollusk shells are common fossils.

Some hard or protective tissues may not contain any appreciable mineral phase but are nevertheless resistant to degradation and may commonly be fossilized. Arthropod cuticle, cnidarian perisarc, hemichordate periderm, leaf cuticle, and spore exine are examples of such outer protective tissues. What gives them their resistance is usually tanned proteins, polysaccharides, or waxes.

Exceptional preservation. Hard or otherwise resistant tissues are thus responsible for the overwhelming majority of all fossils. Uncritical reliance on such skeletal fossils leads to grave biases with regard to the nature and anatomy of individual organisms and to the composition of the once-living biota. Only about one-tenth of the taxa living today have tissues and life habitats that make them readily fossilizable. For this reason, paleontology is increasingly dependent

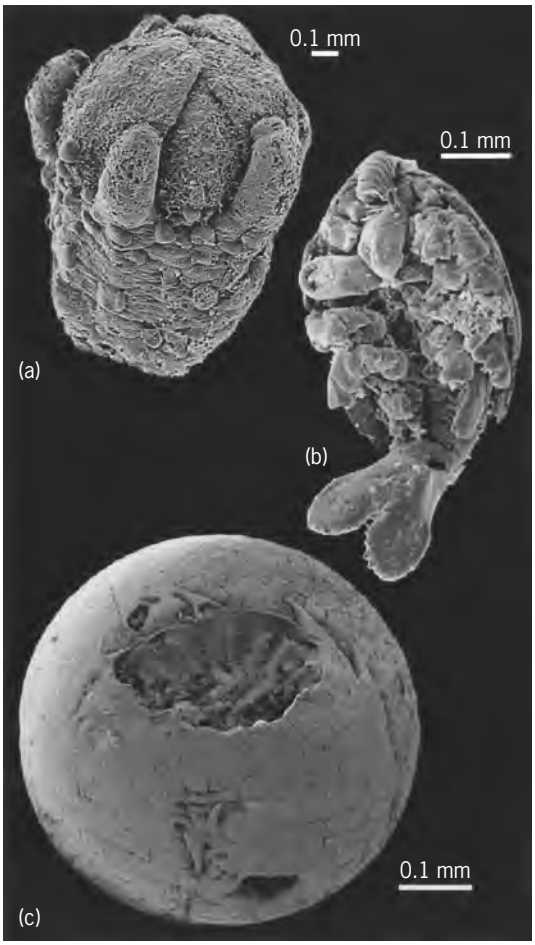

Fig. 2. Fossils with exceptional preservation of soft parts. (*a*) Flower from the Cretaceous Period, *Silvianthemum*, preserved by charring due to a forest fire. (*b*) Crustacean arthropod, *Rehbachiella*, and (*c*) embryo of segmented worm, *Markuelia*, both from the Cambrian Period and preserved by impregnation with calcium phosphate. (*Part a from E. M. Friis, Silvianthemum suecicum gen. et sp. nov., a new saxifragalaen flower from the Late Cretaceous of Sweden, Biologiske Skrifter, Det Kongelige Danske Videnskabernes Selskab, 36, plate 1, fig. 1, 1990; b from D. Walossek, The Upper Cambrian Rechbachiella and the phylogeny of Branchiopoda and Crustacea, Fossils and Strata, 32, plate 32, fig. 2, 1993; c reprinted with permission from S. Bengtson and Z. Yue, Fossilized metazoan embryos from the earliest Cambrian, Science, 277:1645–1648, copyright 1997, American Association for the Advancement of Science*)

on sites with unusual conditions of preservation, allowing for the fossilization also of soft parts and of more complete samples of the total biota (**Fig. 2**).

Early silicification of sediments may trap and preserve biota; this is the main process responsible for knowledge of the microbially dominated biosphere of the Archean and Proterozoic eons, up to about 550 million years ago. The oldest known fossils on Earth are bacterialike structures from Archean sediments, dated to approximately 3.5 billion years. *See* PALEOECOLOGY; SEDIMENTOLOGY.

Another process known to promote exquisite preservation is impregnation with calcium phosphate during early diagenesis (physical and chemical changes occuring in sediments between deposition and solidification); this has been known to preserve soft tissues, even to cellular detail. Small arthropods have been preserved in profusion and in detail by phosphatization of the soft cuticular integument. Even early embryonic stages of animals are preserved through phosphatization; the record of animal embryos now goes back to the late Proterozoic.

Seemingly destructive forest fires may result in excellently preserved plant tissues through coalification. Wood fragments, seeds, and even delicate structures such as flowers can be thus preserved by charring.

Amber, fossilized tree resin, is well known for its capacity to trap and fossilize insects and other small animals and plant parts. Insects in amber are known from at least the last 100 million years.

Freezing has yielded spectacular finds of soft-tissue preservation of, for example, mammoths. The dependence on permanent low temperatures for maintaining the fossils, however, limits this kind of preservation to the most recent fossil biotas.

Various types of fine-grained shales and mudstones have for a long time been known to be potential sites of exquisitely preserved fossils. Examples are the Cambrian Burgess Shale (British Columbia) and Chengjiang mudstones (south China), the Devonian Hunsrück Shale, the Jurassic Holzmaden Shale [= Posidonienschiefer] and Solnhofen Lithographic Limestone (Germany), the Jurassic-Cretaceous Chaomidianzi shales of Liaoning (northeast China), and the Eocene Messel Oil Shale (Germany). These rocks are more or less compressed by the weight of overlying sediment, and so the fossils are not preserved in as full relief as in the other types of extraordinary preservation mentioned earlier. However, the shaley deposits are capable of preserving much larger fossils (such as the ichthyosaurs of Holzmaden and the feathered dinosaurs and birds of Liaoning) than most of the other processes. *See* BURGESS SHALE.

Other types of exceptional fossil preservation are known, though they are more incidental and may be restricted to a short stratigraphic interval. The late Proterozoic Ediacara biota of large, probably multicellular organisms is preserved in coarse-grained sediments spanning a relatively short interval of time; a suggested explanation of this pattern is that the preservation was dependent on the presence of ungrazed microbial mats on the sea floor and that these disappeared when grazing and burrowing animals became common.

Trace fossils (marks of animal activities in sediment) and coprolites (fossilized feces) generally give less information than body fossils about the anatomy of the ancient organisms, but they are important sources of ecological and behavioral information (**Fig. 3**). *See* TRACE FOSSILS.

Time indicators. Fossils were used to order and date rocks even before the principle of evolution was generally accepted, and throughout time the most important practical use of fossils has been biostratigraphy (the identification, ordering, and correlation of fossil-containing sedimentary rocks). Rocks from each slice of time represent a unique but diverse biosphere, and the preservation of the organisms is subject to complex interacting factors; therefore

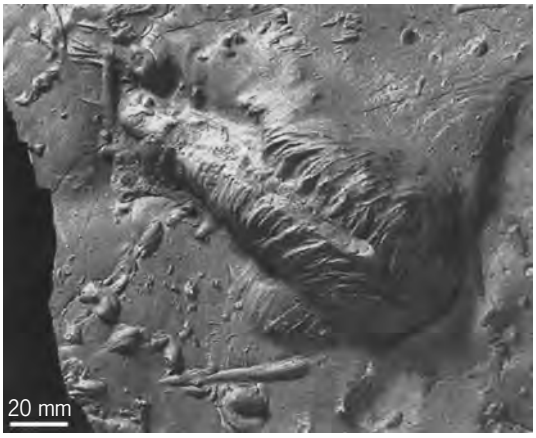

Fig. 3. Trace fossil, *Cruziana*, formed by an arthropod (probably a trilobite) digging for a wormlike animal in soft mud during the Cambrian Period. The picture shows the lower side of a sandstone bed casting of the original markings in the mud (now vanished). (*Swedish Geological Survey; from S. Jensen, Trace fossils from the Lower Cambrian Mickwitzia sandstone, south-central Sweden, Fossils and Strata, vol. 42, 1997*)

biostratigraphy combines analyses of systematics, evolution, ecology, sedimentology, and biogeography.

Biostratigraphy is further becoming integrated with other methods of correlating rocks, for example isotope geology, geomagnetism, and sequence stratigraphy. A central question in modern (bio) stratigraphy concerns what time resolution is possible, how much confidence can there be in known stratigraphical ranges of certain fossils, and how large are the errors in the dating of biological and geological events. *See* GEOLOGIC TIME SCALE; GEOLOGY.

Environmental indicators. Every organism has its range of tolerable environmental conditions, and so a key use for fossils is as indicators of parameters such as temperature, salinity, light intensity, and nutrition. Such paleoecological inferences are most reliable in the younger parts of the geological column, where there are organisms closely related to those living today, with known ecological preferences. In older rocks, inferences must be based more on analogs, functional morphology, and the character of the entombing sediment. Trace fossils are particularly important as environmental indicators, because they record activities that are tightly bound to the particular spot studied (no possibility of postmortem transport).

The chemistry of fossils is frequently used as environmental indicators; their isotopic composition can be used to infer, for example, sea-water temperature ($^{18}O/^{16}O$ ratios in shells) and diet ($^{13}C/^{12}C$ and $^{15}N/^{14}N$ ratios in bone); and trace elements in shells that grew by accretion may provide a proxy for changes in environmental chemistry during the life of the animal. *See* PALEOCLIMATOLOGY.

Global importance. The process of fossilization has a wider significance than just providing a record of past life. A considerable part of Earth's sedimentary cover is of biological origin, that is, fossils; and the permanent or temporary withdrawal of, for example, carbonate and organic matter from biogeochemical cycling by burial in sediment has a direct and profound influence on the composition of atmosphere and oceans and thus on the conditions for life. Stefan Bengtson

Fossil record. The fossil record is the sequence of life preserved as fossils in rocks of all ages in all parts of the Earth. The term is also used in a more restricted sense for the preserved history of a particular taxonomic group, as in "the fossil record of vertebrates"; the preserved history from a particular interval of time, as in "the Precambrian fossil record"; or the preserved history of a particular ecosystem, as in "the fossil record of hydrothermal vent communities."

Quality. The fossil record is vast, extending over 3.5 billion years and preserving representatives of all living kingdoms in all major ecosystems. But the record is also incomplete and knowledge of it is imperfect, despite more than two centuries of scientific investigation. Approximately 250,000 named animal species have been collected from the marine fossil record. The percentage that this represents of the total number of marine animals that ever existed can be estimated from the average duration of species and the approximate number of species extant at any given time. Average durations vary with taxonomic group, but 4 million years is probably a good average. The first animal appeared about 600 million years ago, and today there are about 250,000 species in the oceans. The path of increase in species numbers, or diversity, has been irregular, but can be roughly approximated as a straight line for the sake of calculation (**Fig. 4**). Given this assumption, the total number of marine animal species that ever lived is $^1/_2 \times 2.5 \times 10^5$ species $\times$ (600×10^6 years $\div$ 4×10^6 years) $= 188 \times 10^5$ species, and the percentage that have been described as fossils is 100 $\times$

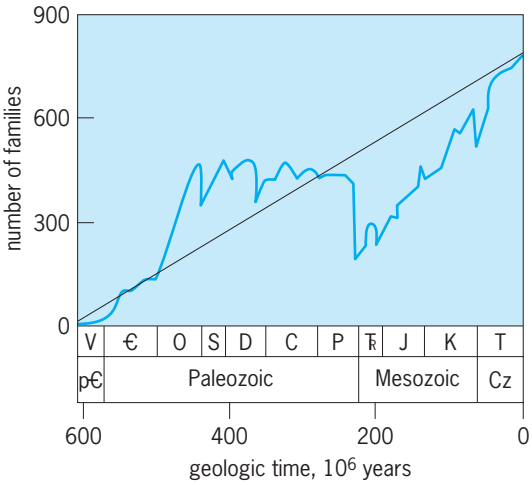

Fig. 4. Diversity of taxonomic families of animals in the marine fossil record, showing the irregular path of increase from the Vendian (that is, latest Precambrian) to the present. The increase of species diversity is thought to follow a similar path. The straight line shows a rough approximation of this diversity increase, with excess diversity after evolutionary radiations canceled by reduced diversity after mass extinctions. Geologic eras and periods are Precambrian (p-€), Cenozoic (Cz), Vendian (V), Cambrian (-€), Ordovician (O), Silurian (S), Devonian (D), Carboniferous (C), Permian (P), Triassic (-Ꞧ), Jurassic (J), Cretaceous (K), and Tertiary (T). (*After N. Eldredge, ed. Systematics, Ecology, and the Biodiversity Crisis, Columbia University Press, 1992*)

(2.5×10^5 species $\div$ 188×10^5 species) = 1%. Thus, the known sample of the fossil record is small and will not permit a detailed reconstruction of the phylogeny of all life. However, for statistical and many other purposes, this sample is large, especially in comparison to political and consumer surveys.

Unfortunately, the sample from the fossil record is not random and unbiased. The quality varies greatly, from nearly complete sequences of populations of certain Miocene lake fish, to scattered, poorly preserved fossils of soft-bodied worms, to absence of recognized fossils for many bacterial groups. The fossil record is best for invertebrate animals with robust skeletons composed of calcium carbonate (for example, corals, clams, snails, and brachiopods) and poorest for small organisms without skeletons or durable cuticles (for example, many bacteria, protists, and fungi and various wormlike animals such as flat worms and nematodes). Between these extremes there is a wide spectrum of preservation that varies with the type of organism and its geologic age, environment, and biogeography.

In general, the fossil record is better known for younger eras because more sedimentary rock has survived the ravages of erosion, subduction, and metamorphism and because there has been more economic incentive for study. The vast majority of fossils have been obtained from Phanerozoic rocks, 0–545 million years in age. Archean rocks, 2.5–3.8 billion years in age, have so far yielded fewer than 10 deposits with verifiable fossil.

The Phanerozoic fossil record is more complete for marine environments, which are sites of net deposition, than for continental environments, which are sites of net erosion. Within the marine realm, the fossil record tends to be better for the continental shelf than for the deep sea, where most sediments older than 150 million years have been subducted or metamorphosed. On land, the record is much better for lowland areas, especially floodplains of coastal rivers, than for upland areas such as high plains, hills, and mountains where sedimentary deposits rarely are shielded from erosion for any significant length of time.

Even with preservation, sampling of the fossil record is quite variable. Organisms with multielement skeletons that dissociate after death, such as many echinoderms, are plentiful in the fossil record but are often very difficult to study (exceptions include many mammals and plants with easily identifiable teeth and pollen or spores, respectively). Preservable organisms that were rare, geographically restricted, or of short geologic duration tend to be sampled less frequently than abundant, widespread organisms; and organisms useful for stratigraphic correlation, such as foraminifers, ammonites, graptolites, conodonts, and pollen-producing plants, have been more intensively studied than organisms with limited utility, such as cyanobacteria, calcareous algae, and sponges. Finally, sampling of the fossil record has varied with political geography, with more intensive study in the industrialized world than in developing nations.

Overview of Phanerozoic record. Species are the basic units of evolution and the subject of study in detailed paleontologic investigations. However, most generalizations about the fossil record are made at higher taxonomic levels, particularly the genus and family. This is because complete data are easier to compile for the less numerous genera and families (approximately 40,000 and 4000, respectively, in the Phanerozoic marine record) and because sampling of higher taxa is more complete (only one species is needed to document the presence of a genus, family, and such). Figure 4 illustrates the record of diversity of animal families in the oceans and shows that in general the fossil record has become richer toward the Recent. **Figures 5** and **6** display summaries of family diversities within taxonomic classes for most marine and continental organisms. Diversities in these two figures are plotted as spindle diagrams that show the number of families through time graphed symmetrically about a central axis. These spindle diagrams show that the fossil record of taxonomic classes is highly skewed in terms of size, duration, and time of origin. Among marine animals (Fig. 5), most classes are small, and more than half of the recognized families are contained in only seven major classes, the Cephalopoda, Articulata (Brachiopoda), Crinoidea, Osteichthyes, Gastropoda, Bivalvia, and Trilobita (listed in order of decreasing size).

Approximately 60% of marine classes appear in the first quarter of the illustrated time (Vendian and Cambrian), whereas only 5% first appear in the last half (Permian to Pleistocene). Many of the early classes remain small and become extinct before the end of the Paleozoic Era. In contrast, those that radiate and attain substantial diversity during the Paleozoic tend to persist to the present day (the Trilobita being the most notable exception). However, there are many classes that have long histories but few fossil families. Most of these comprise soft-bodied animals (for example, Nemertina, Priapulida, Sipunculida) that are known from only a few extraordinary fossil deposits, termed *Lagerstätten*. In the marine record, 20% of fossil classes are known from only three *Lagerstätten*: the Cambrian Burgess Shale, the Devonian Hunsrück Shale, and the Carboniferous Francis Creek Shale of the Mazon Creek area.

Classes in the continental fossil record (Fig. 6) are similarly skewed. More than half of the 39 animal classes contain fewer than 7 fossil families, and more than 40% of the 2500 recognized families belong to a single class, the Insecta. From the Carboniferous to present, this class contains more families than all others in the continental fossil record, including the sum of all vertebrate classes. As in the marine record, exceptional views of the diversity of continental life are provided by a few *Lagerstätten*, notably the Oligocene Baltic ambers which have yielded an unprecedented number of nonmarine arthropods.

The continental fossil record of animals begins in the Silurian, later than the marine record, which begins in the Vendian. As in the marine record, first appearances of classes tend to be early, with two-thirds appearing before the end of the Carboniferous. The record for terrestrial plants (Fig. 6, bottom) is similar: most classes appear early (except

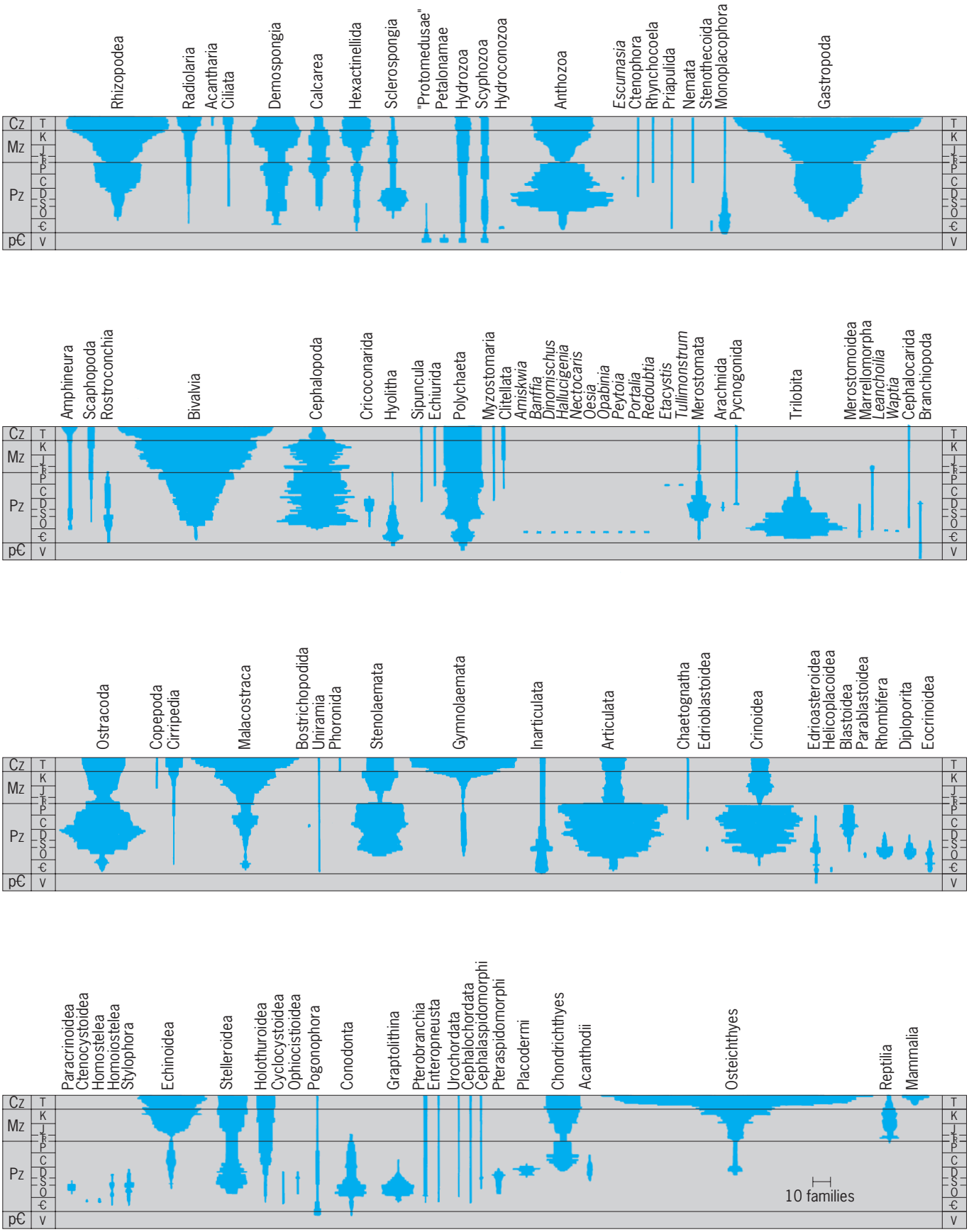

Fig. 5. Spindle diagrams displaying the diversity of families within animal classes in the marine fossil record. Time runs vertically in each strip from Vendian to Pleistocene. (Modern families without fossil representatives are not included.) Geologic eras and systems are Precambrian (p -€), Paleozoic (Pz), Mesozoic (Mz), Cenozoic (Cz). (*After J. W. Valentine, ed., Phanerozoic Diversity Patterns: Profiles in Macroevolution, Princeton University Press, 1985*)

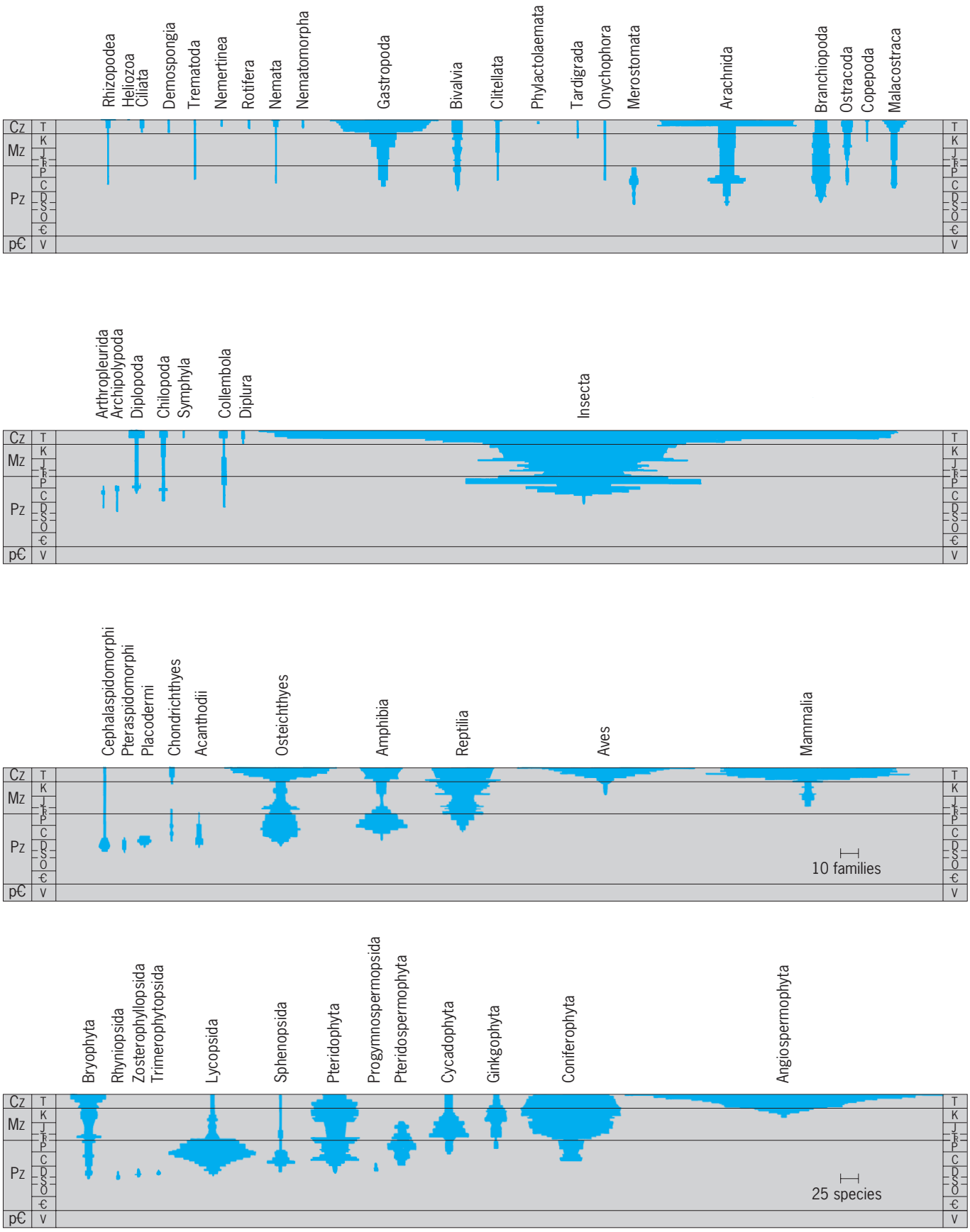

Fig. 6. Spindle diagrams displaying the diversity of families within animal classes and of species within plant classes in the continental (that is, nonmarine) fossil record. Plotting conventions are the same as in Fig. 5. (*After J. W. Valentine, ed., Phanerozoic Diversity Patterns: Profiles in Macroevolution, 1985*)

the Magnoliophyta = angiosperms), and only a few classes are large (notably the Pinophyta and Magnoliophyta). *See* PALEOBOTANY; PALEONTOLOGY.

J. John Sepkoski, Jr.

Bibliography. R. C. Moore et al. (eds.), *Treatise on Invertebrate Paleontology*, 1953–1998; D. R. Prothero, *Bringing Fossils to Life: An Introduction to Paleobiology*, 1998; D. M. Raup and S. M. Stanley, *Principles of Paleontology*, 2d ed., 1978; G. G. Simpson, *The Major Features of Evolution*, 1953; S. M. Stanley (ed.), *Macroevolution: Pattern and Process*, 1979; J. W. Valentine (ed.), *Phanerozoic Diversity Patterns: Profiles in Macroevolution*, 1985.

Fossil apes

Apes and humans are closely related primates in the superfamily Hominoidea. The living hominoids are subdivided into the families Hylobatidae and Hominidae. The hylobatids or lesser apes (genus *Hylobates*) are represented by approximately nine species found throughout Southeast Asia. Humans and the great apes—the orangutan (*Pongo pygmaeus*), the gorilla (*Gorilla gorilla*), the common chimpanzee (*Pan troglodytes*), and the pygmy chimpanzee or bonobo (*Pan paniscus*)—are grouped in the Hominidae. In the past, the great apes were included in a separate family, the Pongidae, but subsequent anatomical and molecular studies showed that the African apes (*Gorilla* and *Pan*) are more closely related to humans than they are to the Asian orangutan.

The classification of fossil and living apes, along with their geological ages and geographical regions, follows. The asterisk indicates extinct apes.

Superfamily Hominoidea
- Family Hylobatidae
  - *Hylobates* (Pleistocene to Recent, Southeast Asia)
- Family Hominidae
  - Subfamily Kenyapithecinae
    - *Griphopithecus** (middle Miocene, Turkey and Austria)
    - *Kenyapithecus** (middle Miocene, East Africa)
  - Subfamily Dryopithecinae
    - *Dryopithecus** (middle to late Miocene, Europe)
    - *Oreopithecus** (late Miocene, Italy)
  - Subfamily Ponginae
    - *Ankarapithecus** (late Miocene, Turkey)
    - *Gigantopithecus** (late Miocene to Pleistocene, Asia)
    - *Lufengpithecus** (late Miocene, China)
    - *Pongo* (Pleistocene to Recent, Southeast Asia)
    - *Sivapithecus** (middle to late Miocene, Indo-Pakistan)
  - Subfamily Homininae
    - *Ardipithecus** (early Pliocene, Ethiopia)
    - *Australopithecus** (Pliocene, Africa)
    - *Gorilla* (Recent, Africa)
    - *Graecopithecus* (= *Ouranopithecus*)* (late Miocene, Greece)
    - *Homo* (late Pliocene to Recent, cosmopolitan)
    - *Pan* (Recent, Africa)
    - *Paranthropus** (Pliocene to Pleistocene, Africa)
    - *Samburupithecus** (late Miocene, Kenya)
  - Uncertain family status
    - *Morotopithecus** (early Miocene, Uganda)
    - *Otavipithecus** (middle Miocene, Namibia)
- Uncertain superfamily status
  - Family Proconsulidae
    - *Afropithecus** (early Miocene, Kenya)
    - *Dendropithecus** (early Miocene, East Africa)
    - *Heliopithecus** (early Miocene, Saudi Arabia)
    - *Kalepithecus** (early Miocene, East Africa)
    - *Kamoyapithecus** (late Oligocene, Kenya)
    - *Limnopithecus** (early to middle Miocene, East Africa)
    - *Mabokopithecus** (middle Miocene, Kenya)
    - *Micropithecus** (early and middle Miocene, East Africa)
    - *Nyanzapithecus** (early and middle Miocene, East Africa)
    - *Proconsul** (early and middle Miocene, East Africa)
    - *Rangwapithecus** (early Miocene, East Africa)
    - *Simiolus** (early Miocene, Kenya)
    - *Turkanapithecus** (early Miocene, Kenya)

Evolutionary history. The evolutionary history of the extant hominoids is poorly known, with the notable exception of humans, which have a relatively complete fossil record extending back more than 4 million years. The earliest fossil apes that can be definitively linked to the modern hylobatids are known from sites in China dated to less than 1.5 million years ago (Ma), while the fossil record for the African apes is entirely unknown. The evolution of the orangutan is, by comparison, much better documented. Fossil teeth from cave sites in Asia dating back more than 1 million years show that orangutans in the past were considerably larger than they are today, and that unlike their living relatives, which are found only on Sumatra and Borneo, they once had a wider distribution in Southeast Asia that extended as far north as China. In contrast to the paucity of fossils available to trace the evolutionary history of hominoids over the past 5 million years, there is a wealth of evidence from the Miocene Period (23–5 Ma). This evidence shows that apes were once much more common and more diverse than they are today.

Proconsulids. The remains of apelike fossil primates, commonly known as proconsulids, have been recovered from sites in Kenya, Uganda, and Saudi Arabia dating to the early Miocene (23–16 Ma). There are 15 species of proconsulids, ranging in size from the tiny *Micropithecus clarki* (4 kg; 9 lb) to the chimpanzee-sized *Proconsul major* (50 kg; 110 lb). Comparisons of their teeth, jaws, and skeletons indicate that proconsulids exhibited a wide diversity of dietary and locomotor behaviors, but they were typically arboreal and ate various combinations of soft fruits and young leaves. The best-known species are

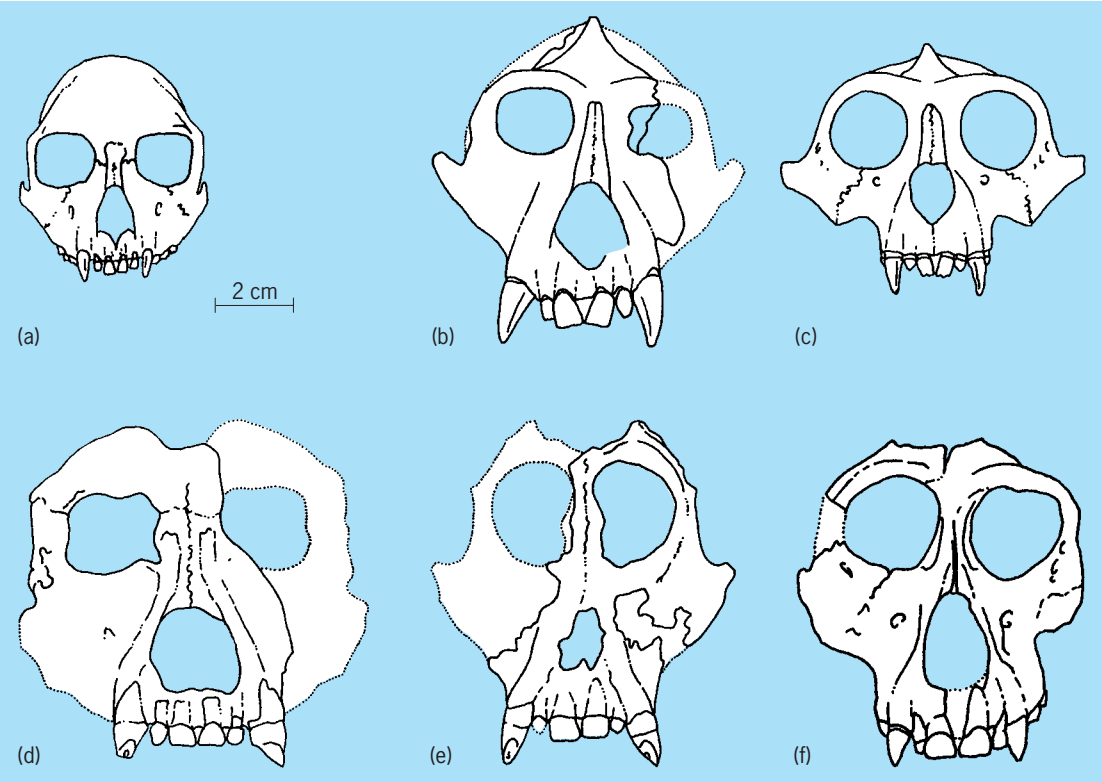

Fig. 1. Reconstructions and partial reconstructions of the skulls of fossil apes: (*a*) *Proconsul*, (*b*) *Afropithecus*, (*c*) *Oreopithecus*, (*d*) *Graecopithecus*, (*e*) *Sivapithecus*, and (*f*) *Ankarapithecus*.

Proconsul heseloni, *Turkanapithecus kalakolensis*, and *Afropithecus turkanensis* (**Fig. 1**). Studies have shown that proconsulids represent either the earliest known hominoids or primitive stem catarrhines (the group which gave rise to both Old World monkeys and apes). They are certainly more primitive than any of the living apes, retaining generalized skulls and teeth, and monkeylike postcranial skeletons. However, during the early Miocene there was at least one species of hominoid living in East Africa, *Morotopithecus bishopi*, which had already acquired some of the unique features of modern apes. This species, from Moroto in Uganda (dated to more than 20 Ma), had specializations of the lumbar vertebrae and scapula, not found in the contemporary proconsulids, that indicate that it had developed the stiff-backed, partially upright posture and suspensory forelimbs that are characteristic of modern hominoids.

During the middle Miocene (16–10 Ma), conditions in East Africa became somewhat drier, cooler, and more seasonal, and open woodland habitats replaced the humid tropical forests that were typical of the early Miocene. These ecological changes coincided with the appearance in East Africa of a more advanced type of hominoid, *Kenyapithecus*. This ape is derived relative to the proconsulids in having thickened enamel on its cheek teeth, more robust jaws, and relatively larger upper premolars. The limb bones indicate that *Kenyapithecus* was more terrestrially adapted than proconsulids and exhibited a number of specialized features that link it more closely to extant hominoids. Another intriguing fossil ape from this time period is *Otavipithecus* from Namibia, dated to approximately 13 Ma, and the only Miocene hominoid recorded from southern Africa. Until recently, it was known only by a single jaw fragment, so its relationships to other fossil and extant hominoids have been difficult to establish. However, additional cranial and postcranial specimens of *Otavipithecus* have now been discovered, and it is hoped that these will shed further light on its affinities.

Eurasian hominoids. Until the middle Miocene, hominoids were restricted to Africa, but during this period they migrated into Eurasia. The earliest Eurasian hominoid, dated to 16–15 Ma, is *Griphopithecus*, which is known from Turkey and central Europe. The teeth and jaws are similar to those of *Kenyapithecus* from East Africa, to which it is probably closely related. Once in Eurasia, hominoids became established over a wide geographical region, extending from Spain in western Europe to eastern China, and they became increasingly diversified during the middle and late Miocene (16–5 Ma). The best-known fossil Eurasian hominoids are *Dryopithecus* (western and central Europe), *Oreopithecus* (Italy), *Graecopithecus* or *Ouranopithecus* (Greece), *Ankarapithecus* (Turkey), *Sivapithecus* (Indo-Pakistan), and *Lufengpithecus* (China) [Fig. 1]. Of these forms, *Sivapithecus* is evidently closely related to the living orangutan, but the relationships of the other Eurasian Miocene hominoids remain contentious. A number of alternative hypotheses about the interrelationships of Eurasian Miocene hominoids have been proposed: (1) that they form

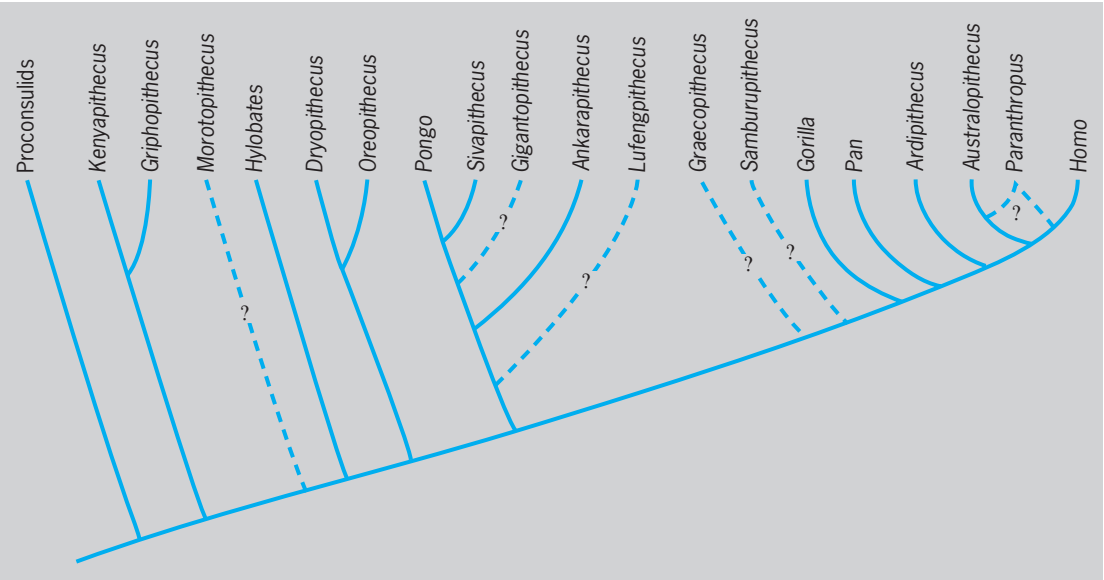

Fig. 2. Cladogram of the evolutionary relationships of fossil and living apes. Broken lines indicate uncertain relationships.

a closely related group with the living orangutan, all being derived from a common ancestor that migrated into Europe from Africa sometime during the middle Miocene; (2) that some of the Eurasian hominoids, such as *Dryopithecus* and *Graecopithecus*, are more closely related to the African apes and humans than they are to *Sivapithecus* and the orangutan; and (3) that they represent a diverse group containing primitive hominoids, as well as forms belonging to the orangutan and African great ape lineages (**Fig. 2**).

An ecological shift from moist temperate woodlands to drier, more seasonal habitats during the later Miocene coincided with a sharp decline in the diversity of hominoids in Eurasia. The only survivor in Europe toward the end of the Miocene was *Oreopithecus*, a highly specialized relative of *Dryopithecus* that was isolated on a group of islands in the northern Mediterranean which today form part of Italy (Fig. 1). *Lufengpithecus* and *Sivapithecus*, along with *Gigantopithecus*, are found in the late Miocene of Asia. The ape *Gigantopithecus* was the largest known hominoid, with massive jaws and teeth specialized for eating tough, fibrous vegetation, and with an estimated body weight that may have exceeded 200 kg (440 lb) [living male gorillas average only 170 kg or 375 lb]. All of these Eurasian hominoids became extinct by the close of the Miocene, except for *Gigantopithecus*, whose remains have been recovered from Pleistocene cave sites in China dated to less than 1 Ma.

Hominoids also became extremely rare in Africa during the late Miocene. A large hominoid, *Samburupithecus*, known only by a single maxilla from Kenya (dated to 10–8 Ma), may represent a close relative of the African apes and humans. A few isolated teeth and a lower jaw fragment of fossil hominoids from the late Miocene sites of Lukeino and Lothagam (dated to 7–5 Ma) in northern Kenya have been suggested to be the earliest known occurrence of the human lineage, but the remains are too scrappy to be certain of their taxonomic affinities. The earliest definitive record of fossil hominoids that are more closely related to humans than they are to the African great apes is known from the Pliocene (5.2–1.6 Ma) with the appearance of *Ardipithecus ramidus* from Ethiopia (4.4 Ma), *Australopithecus anamensis* from Kenya (4.2–3.9 Ma), and *Australopithecus afarensis* from Ethiopia and Tanzania (4–3 Ma). *See* APES; FOSSIL HUMANS; MAMMALIA; MONKEY.

Terry Harrison

Bibliography. D. R. Begun, C. V. Ward, and M. D. Rose, *Function, Phylogeny, and Fossils: Miocene Hominoid Evolution and Adaptations*, 1997; R. L. Ciochon and R. S. Corruccini, *New Interpretations of Ape and Human Ancestry*, 1983; J. G. Fleagle, *Primate Adaptation and Evolution*, 2d ed., 1998; R. G. Klein, *The Human Career*, 1989; B. Wood, L. B. Martin, and P. Andrews, *Major Topics in Primate and Human Evolution*, 1986.

Fossil fuel

Any naturally occurring carbon-containing material which when burned with air (or oxygen) produces (directly) heat or (indirectly) energy. Fossil fuels can be classified according to their respective forms at ambient conditions. Thus, there are solid fuels (coals); liquid fuels (petroleum, heavy oils, bitumens); and gaseous fuels (natural gas, which is usually a mixture of methane, CH_4, with lesser amounts of ethane, C_2H_6, hydrogen sulfide, H_2S, and numerous other constituents in small proportions).

Some differentiation must be made between the natural fossil fuels and those synthetic fuels derived by incurring major changes in the character of the fossil fuel. Examples are the liquid products made from coal by any one of a variety of potential process options; such materials are synthetic liquid fuels.

Heating values of representative fuels

Fossil fuel	Btu/lb	Heating value Btu/ft^3	MJ/k^3	MJ/m^3
Natural gas		900		33.5
Petroleum	19,000		44.1	
Heavy oil	18,000		41.8	
Tar-sand bitumen	17,800		41.3	
Coal				
Lignite	8,000*		18.6	
Subbituminous	10,500*		24.4	
Bituminous	15,500*		36.0	
Anthracite	15,000*		34.8	

*Representative values are given because of the spread of subgroups with various heating values.

See COAL GASIFICATION; COAL LIQUEFACTION; SYNTHETIC FUEL.

Occurrence. Coal is a naturally occurring organic rock that originated from material which accumulated from the prolonged (millions of years) decay of woody plants. Coal origins are typically represented as beginning in a peat swamp in which trees, ferns, and the like are deposited, and buried by sand, silt, and mud. As a result of temperature and pressure effects, metamorphosis of the woody material occurs to produce the various types of coal. A large proportion of existing coal originated from materials deposited during the Carboniferous Period some 2.8–3.5 $\times$ 10^8 years ago. *See* COAL.

Petroleum occurs in highly porous rock reservoirs that are enclosed by an impermeable basement rock and an impermeable cap rock. The petroleum originated from the deposition of dead microscopic plant and animal organisms in ancient marine sediments. Geological effects such as temperature, time, and the pressure of overburden brought about chemical changes resulting in the production of hydrocarbon materials. In addition, it is generally recognized that the reservoir rock and the source rock are not the same. Migration occurred, for example, because of overburden pressures, from the source rock (now usually shales and clays) into the more coarse-grained reservoir rock (usually sandstones or carbonate rocks). The source rocks range in age up to about 5 $\times$ 10^8 years, but it is believed that conversion of the source material occurred during a period of 1–3 $\times$ 10^6 years. *See* PETROLEUM; PETROLEUM GEOLOGY.

Heavy oils and tar-sand bitumens (natural asphalts) are a poorer-quality petroleum that has come into prominence because of the dwindling supplies of conventional petroleum. Both had origins similar to that of conventional petroleum but are believed to be younger (less mature), and occur in sandstone or carbonate reservoirs. *See* ASPHALT AND ASPHALTITE; BITUMEN; OIL SAND; OIL SHALE.

Like conventional petroleum, heavy oils occur in fields, but unlike conventional petroleum, heavy oils usually require additional (thermal) stimulation for recovery at the surface. Tar-sand bitumens occur as deposits in which the ore body (reservoir rock) may contain 1–20% by weight bitumen. The bitumen is immobile at the formation conditions and occurs as a thick, black, near-solid natural asphalt. The bitumen can be recovered by mining followed by removal from the associated sand (hot-water separation process). Nonmining (in-place) techniques require severe thermal or chemical treatments to stimulate bitumen mobility in the reservoir and recovery at the surface.

Like petroleum, natural gas also occurs in subterranean porous reservoir rocks either alone or in association with petroleum. The traps also consist of impermeable basement and cap rocks. Natural gas also probably resulted from the decomposition of dead microscopic organisms which had been deposited in aqueous environments. It may also be produced as an integral part (or by-product) of the maturation process in which source material is converted to petroleum. *See* NATURAL GAS.

Composition. Fossil fuels are composed of carbon (C) and hydrogen (H) with varying amounts of nitrogen (N), oxygen (O), and sulfur (S). On an atomic basis, natural gas has the highest atomic H/C ratio and coal the lowest; petroleum has an atomic H/C ratio between the two. There are also varying H/C ratios within the coal and petroleum series. The atomic H/C ratios for various fossil fuels are compared in the **illustration**.

Coals vary in elemental composition over a wide range from about 70% C and about 5% H in the lignitic coals to about 93% C and about 4% H in the anthracite coals. The differences are made up of N (< 1%) and S (< 2%) with the remainder occurring as O. In contrast, most petroleums fall into the range 83–87% C

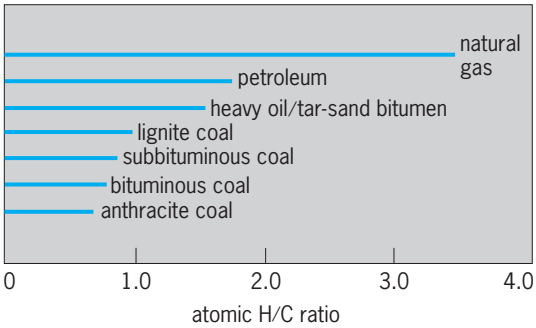

Comparison of the atomic H/C ratios for various fossil fuels.

and 10–14% H with 2–5% S and the remainder occurring as N (< 1%) and O (< 1%). Natural gas which is mostly methane has about 75% C and about 25% H.

Heating value. One important aspect of the fossil fuels is the heating value of the fuel, which is measured as the amount of heat energy produced by the complete combustion of a unit quantity of the fuel. For solid fuels and usually for liquid fuels the heating value is quoted for mass, whereas for gaseous fuels the heating value is quoted for volume. The heating values are commonly expressed as British thermal units per pound (Btu/lb). In SI units the heating values are quoted in megajoules per kilogram (MJ/kg). For gases, the heating values are expressed as Btu per cubic foot (Btu/ft^3) or as megajoules per cubic meter (MJ/m^3). The **table** gives heating values of representative fuels. *See* ENERGY SOURCES; HEAT.

James G. Speight

Bibliography. W. Francis and M. C. Peters, *Fuels and Fuel Technology*, 2d ed., 1980; A. Markuszewski and B. D. Blaustein (eds.), *Fossil Fuels Utilization: Environmental Concerns*, 1986; R. A. Meyers (ed.), *Coal Handbook*, 1981; H. H. Schobert, *Coal: The Energy Source of the Past and Future*, 1987; J. G. Speight, *The Chemistry and Technology of Coal*, 2d ed., 1994; J. G. Speight, *The Chemistry and Technology of Petroleum*, 3d ed., 1999.

Fossil humans

All prehistoric skeletal remains of humans that are archeologically earlier than the Neolithic (necessarily an imprecise limit), regardless of degree of mineralization or fossilization of bone, and regardless of whether the remains may be classed as *Homo sapiens sapiens* (anatomically modern humans). In this sense, the term "humans" is used broadly to mean all primates related to living people since the last common ancestor of people and African apes, thus all species currently included in the genera *Homo*, *Australopithecus*, *Ardipithecus*, and *Paranthropus* (and potentially others discussed below).

Discoveries began early in the nineteenth century, although their meaning and antiquity were not recognized before the finding of the Neanderthal specimen in 1856, combined with Darwin's demonstration of the reality of evolution soon afterward. Fossil human remains have come principally from Europe, western Asia, China, Java, and Africa. Because of the rather late entry of humans into the New World, all American remains are of relatively recent origin and recognizable as *H. sapiens*.

Dating fossils. The human lineage was once thought to have come into existence only at the beginning of the Pleistocene, the geological time interval from about 1.8 million years ago to the present. It is now known to have been fully separate from the ancestry of the apes at least as far back as the Early Pliocene or perhaps even the later Miocene (4–7 million years ago, or Ma). This view was widely held by anthropologists, on theoretical grounds, in the early part of the twentieth century. It was abandoned in the 1940s because of lack of supporting evidence. Discoveries in Pliocene and Miocene deposits since then have led to a reappraisal of the evidence (**Fig. 1**).

Dating within the Pleistocene (and earlier) is accomplished by the methods of conventional paleontology and geology, by association with human implements, by several chemical and physical tests for relative age (such as the fluorine test for accumulation of this element in bone), and by geochronometric methods which provide an age in years. The latter usually involve measuring the presence of radioactive substances such as radiocarbon and radiopotassium as well as other radioactivity-based methods such as electron spin resonance, thermoluminescence, and fission-track dating. All of thesemethods are broadly termed radiometric methods, as opposed to approaches such as amino acid racemization or dendrochronology (tree-ring dating). *See* DATING METHODS; RADIOCARBON DATING.

One of the most important methods of correlation between dated geological sequences is paleomagnetism, which does not itself provide a date because it is basically a boolean system with two states (so-called normal and reversed polarity of geomagnetism) of which all occurrences are indistinguishable. But if the pattern of reversals at any site can be matched against the global paleomagnetic time scale and even approximately dated by means of geochronology or paleontology, a sequence of dates may result. *See* PALEOMAGNETISM.

Prehuman ancestry. Humans are catarrhine primates, part of a group including Old World monkeys, apes, and various extinct forms. Most evidence from both comparative morphology and molecular studies of proteins shows that humans' closest living relatives are the African apes: the chimpanzee and the gorilla. Less close is the Asian orangutan, and most distinctive of all apes are the gibbons. A classification which conforms to these relationships within the Hominoidea (apes, humans, and close extinct relatives) recognizes the family Hylobatidae for the gibbons and Hominidae for humans and great apes. The latter family is divided into Ponginae (orangutans and extinct relatives) and Homininae. Many workers, however, continue to reject this view and place humans in the Hominidae as contrasted with the paraphyletic (multi-origined) apes in Pongidae. A few others have suggested that in fact orangutans are the closest living relatives of humans. No fossils of any modern ape are known, other than Pleistocene gibbon and orangutan teeth from southern Asia, and a few Kenyan teeth identified as chimpanzee which were recognized in 2005. *See* FOSSIL APES; FOSSIL PRIMATES; MOLECULAR ANTHROPOLOGY; MONKEY; PRIMATES.

The oldest certain representatives of the Catarrhini are fossils from the Fayum beds of northern Egypt dated around 34 Ma. The best known is *Propliopithecus* (=*Aegyptopithecus*) *zeuxis*, a species near the common ancestry of apes, humans, and Old World monkeys. Hominoids originated in Africa, probably during the Late Oligocene and Early Miocene (26–17 Ma), as suggested by the

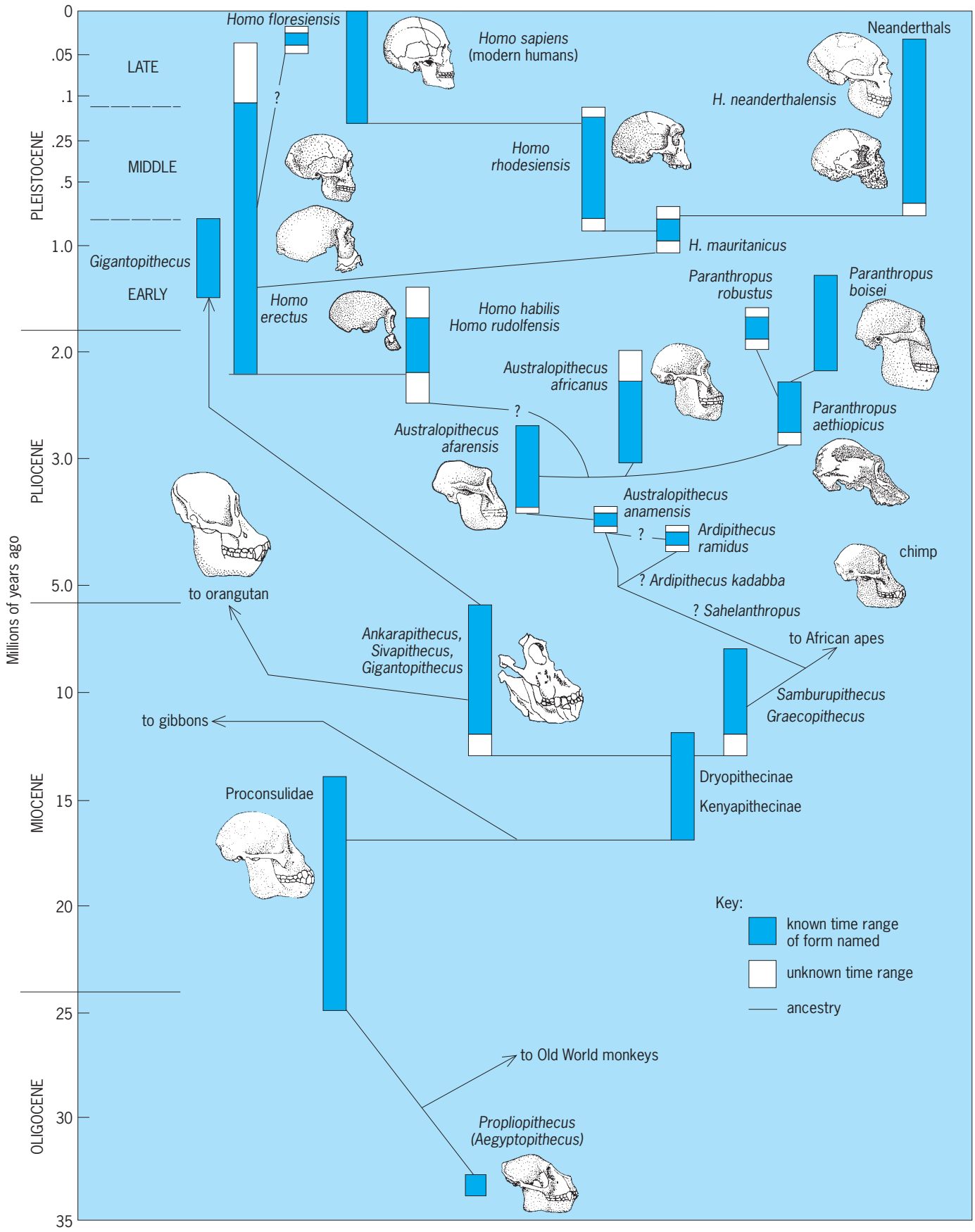

Fig. 1. Human phylogeny from the Oligocene to the present time, showing the skulls of the major known fossil relatives and possible ancestors of modern humans.

genera *Kamoyapithecus* and *Proconsul*, placed in the family Proconsulidae. However, some scholars have questioned the view that these forms are already hominoids, proposing instead that they were conservative catarrhines that predate the evolutionary split between hominoids and Old World Monkeys. Between 20 and 15 Ma, there are several more derived ("advanced") African hominoid genera (for example, *Morotopithecus*, *Afropithecus*, and *Kenyapithecus*) and one Eurasian form (*Griphopithecus*). If these are more closely related to living great apes than are gibbons, as has been proposed in the past, they may represent early members of the Hominidae. However, if they are conservative ("primitive," plesiomorphic) by comparison with the gibbons, then they should not be included in the hominid family.

In the latter case, several workers have argued that the origin of modern ape (and human) lineages must have occurred in Eurasia, where a number of derived genera occur between 14 and 8 Ma. These include *Dryopithecus* and *Pierolapithecus* (discovered in 2004), which were probably close to the common ancestor of all great apes and humans; *Ankarapithecus* and *Sivapithecus*, which appear to represent two early stages in the orangutan lineage (Ponginae); and *Graecopithecus* (**Fig. 2**; also termed *Ouranopithecus*), which has been suggested as a member of Homininae, that is on the lineage leading toward African apes and humans. Another early hominine is *Samburupithecus*, known only from an upper jaw fragment found in the Samburu Hills of Kenya and (like *Graecopithecus*) dated to about 9.5 Ma. A possible alternative interpretation is that a form such as *Morotopithecus*, with relatively derived but poorly known postcranial elements, might have given rise both to the early Eurasian hominids and to an as yet unknown African line which culminated in *Samburupithecus*. In either case, it is not until the latest Miocene (7–5 Ma) that we see the first possible fossil representatives of the Hominini (hominins), the group including modern humans and all our relatives since the last common ancestor shared with chimpanzees. Molecular anthropological studies of DNA sequences and other lines of evidence have suggested that chimpanzees are more closely related to humans than either is to gorillas. This would appear counterintuitive, given the morphological similarity between the two African apes, both of which are also apparently derived in their knuckle-walking locomotion. But similarities may be merely holdovers from the common ancestor of all three forms, while chimps and humans shared an intermediate common ancestor after the gorilla lineage split away. Paleontological evidence is sparse in Africa between 12 and 5 Ma, but it is possible that *Samburupithecus* (or *Graecopithecus*) represents a species near the base of that three-way split or already on the gorilla line. Their ages of 10–9 Ma fit reasonably well with the "molecular clock" estimates of 10–7 Ma for that divergence. The human–chimp split is molecularly estimated at 8–5 Ma or even less, but the human fossil record refutes a date younger than perhaps 5 Ma at the minimum. *See* FOSSIL APES; FOSSIL PRIMATES; MOLECULAR ANTHROPOLOGY; MONKEY; PRIMATES.

Fig. 2. Face (cast) of *Graecopithecus freybergi* from Xirochori, Greece, about 9.5 million years old. (*Courtesy of and © E. Delson; photo by L. Meeker*)

Identification of early hominins. When paleoanthropologists find early fossils that might be on the human lineage, they must identify them as such by locating bony features which distinguish the earliest hominins from their ape relatives or ancestors.

In order to find such features in fossils, we must begin by determining what makes modern humans different from living apes. We can then discern the sequence in which such features appeared in the human fossil record. However, not all changes happened at the same time. This pattern is known as mosaic evolution, referring to mosaic pictures made up of numerous small parts which fit together to make a complete image.

One major difference involves the mode of locomotion used by hominins versus apes. Hominins are habitually bipedal animals (walking on two legs rather than four), whereas our ape ancestor likely included quadrupedal, suspensory and/or knuckle-walking behaviors in its locomotor repertoire. The human skeleton has been modified in many ways to accommodate this change. For example, hominins have a foot with a longer big toe that is in line with the other toes, rather than a grasping ape-like foot with a divergent big toe. The foot has also developed two arches, one longitudinal and one transverse (along the long and short axes of the foot, respectively) that act as shock absorbers when walking bipedally. While the spine of apes is C-shaped, hominins evolved thoracic and lumbar curvatures, resulting in an S-shaped spine that acts to balance the torso over the pelvis and legs. This spine also connects to the skull more vertically (and centrally) underneath rather than toward the back of the skull (as in apes), which can be seen by the more

forward placement of the foramen magnum (opening where the spinal cord enters the skull). The pelvis and pelvic muscles have also gone through a major reorganization: the pelvis is less tall and more bowl-shaped, and the gluteus maximus muscle is more developed to help extend the leg backward in walking. The femur (bone of the thigh) is more angled in at the knee in order to center the feet beneath the body, and the knee is capable of straightening out completely. Modern humans also have quite long legs and only moderately long arms compared to body size. All of these features aid in more stable and efficient bipedal locomotion in hominins than apes. The fossil record suggests that relatively modern-looking spinal curvature, pelvic shape, and femoral angling appeared early, while foot arching, perhaps big toe position, and especially fully modern limb proportions did not evolve until later in the hominin lineage. The human skull has also undergone changes not directly related to bipedal locomotion, such as an increase in brain size (relative to body size) and changes in the teeth, including a decrease in canine size. While some early hominins have relatively larger brains than chimpanzees of similar body weight, major brain expansion probably did not occur until the earliest member of our own genus *Homo* evolved.

The next question then is: When in the fossil record do we first see evidence of these changes? Recently discovered fossils from Chad (termed *Sahelanthropus*; **Fig. 3**) may include the oldest hominin cranium with smaller canines and a forwardly positioned foramen magnum (which may in turn suggest some bipedalism). Limb bone fossils from Kenya (named *Orrorin*) are said to have a number of characteristics consistent with bipedalism, though some debate surrounds these claims. Both sets of fossils probably date from 7–6 Ma, making them the oldest possible hominins in existence. Younger finds (~4.4 Ma) from Ethiopia (*Ardipithecus*) have much smaller canines and may also have a foramen magnum that is positioned more forward than in apes. Canine reduction may have begun even before hominins appeared, as seen in the fossil ape *Graecopithecus*. Further decreases in size eventually led to the very small canines present in modern humans. It has been suggested that smaller canines (with no significant difference between those of males and females) reflect a different social organization, because large canines are associated in apes (and monkeys) with a high degree of competition among males for mates. Fossil australopiths (see below) show clear evidence of bipedalism both in their skeletons and in fossilized footprints found at the site of Laetoli, Tanzania. These hominins still display many apelike features in their skull, showing that habitual bipedalism evolved before modern skull anatomy.

Miocene Hominini. In 2002, French and Chadian researchers led by Michel Brunet reported the discovery of new fossils from Chad in Central Africa. Based on the associated mammalian fossils, the site of Toros-Menalla was estimated to date about 7–6 Ma. Previously, almost all early human remains had been found in eastern or southern Africa, although a few fragments from Chad dating to about 3.5 Ma had been known since the 1990s. The new fossils (named *Sahelanthropus tchadensis*) included a fairly complete though crushed cranium (the skull without the lower jaw) of what might be the oldest hominin ever found. Recent computer-based analysis has permitted a "virtual" reconstruction of what the cranium might have looked like before it was crushed (Fig. 3). Although the brain size is no bigger than that of a similarly sized ape, its teeth are more like a hominins, especially in having an apparently reduced canine. The relatively anterior position of the foramen magnum suggested to the discoverers that this species may have been at least partly bipedal, but others have questioned this view. *Sahelanthropus* is likely near to the chimpanzee–human split, based on both its morphology and the molecular divergence date discussed above.

Fig. 3. Cranium of *Sahelanthropus tchadensis* from deposits ca. 6–7 million years old in Chad, original fossil (left) and virtual reconstruction. (*Courtesy of and © M. P. F. T.*)

In late 1999, Paris-based Martin Pickford and Brigitte Senut working in Kenya discovered a small collection of fragmentary fossils, mostly limb elements, which they also claimed as the earliest known hominin. The fossils were first termed "Millenium Man," as they were announced to the public just before 2000, but they have been formally named *Orrorin tugenensis*. These remains are better dated than those of *Sahelanthropus*, to about 6 Ma, by radiometric techniques. It has been claimed that the femur shows features indicating bipedalism, but most researchers have questioned this interpretation, and some have suggested that *Orrorin* is actually the long-sought fossil of an early African ape ancestor. Unfortunately, dental and cranial remains are too fragmentary to be certain either way.

Still younger fossils are known from sites in the Middle Awash Valley in Ethiopia. Recovered between 1997 and 2004 and first described in 2001, these were named *Ardipithecus kadabba* in 2005. The fossils range in age between 5.8 and 5.2 Ma, at the very end of the Miocene epoch. Most are fragmentary jaws, but a single toe bone has been claimed to indicate bipedalism, although (as usual), some researchers have questioned this interpretation.

Pliocene Hominins: the australopiths. Pliocene humans have been grouped in various ways, but it now seems that four main types can be distinguished. Three of these, dating from roughly 4.5–1 Ma, have previously been assigned to the genus *Australopithecus* and can informally be termed australopiths (here they are placed in the genera *Ardipithecus*, *Australopithecus*, and *Paranthropus*). The fourth group includes early species of *Homo*, beginning about 2.5 Ma. The australopiths have previously been divided into gracile and robust varieties, but the former term really is not accurate and is not used here. All australopith species appear to share a number of basic characteristics (see above) distinguishing them from living and fossil apes and also from later humans, although clearly linking them to the latter. The species of australopith also broadly share a smaller body size than modern humans with surprisingly little difference among the known forms. Estimates of body size based on analysis of weight-bearing joint surfaces and bone lengths suggest average weights around 130 lb (59 kg), with females often under 77 lb (35 kg) and males over 155 lb (70 kg). This quite high sexual dimorphism is also typical of australopiths, as are back teeth that are large for the estimated body size, although the actual sizes and tooth proportions are among the features distinguishing the several species from one another. In light of the diversity now recognized, many researchers accept a division into the genera *Australopithecus* and *Paranthropus*, as well as *Ardipithecus*, but others continue to recognize only one or two of these genera. *See* AUSTRALOPITHECINE.

The fossils of these early humans were first found in South Africa in 1924, but the most recent major discoveries and the best evidence of their age come from East Africa. The South African fossils come from six main sites, which are the remains of ancient cave systems. The australopiths did not live in caves, but their carcasses may have been washed or dropped into crevices near the ground surface by leopards or other carnivores. In two sites especially, many fossils are known from relatively short spans of time (perhaps less than 200 thousand years at each), but the evidence for dating is not definite.

In 1959 Mary Leakey and L. S. B. Leakey discovered a nearly complete australopith skull at Olduvai Gorge, Tanzania, and colleagues dated it at about 1.75 Ma, far older than previously thought. Since then, American and joint American-French and American-Ethiopian expeditions have found hundreds of human fossils in the Omo and Afar areas of Ethiopia, while a Kenyan team has worked around the shores of Lake Turkana (formerly Lake Rudolf), Kenya, and various studies have continued at Olduvai and nearby Laetoli. These regions have yielded smaller numbers of specimens at many separate subsites, but the age of each site can usually be estimated closely by potassium-argon and paleomagnetic dating. Also, many specimens are more complete and show less distortion than their contemporaries from South Africa. Until 1993, no definite examples of *Australopithecus* or *Paranthropus* were known outside these areas, although some claims have been made. The recovery of australopith specimens from Chad (and of *Paranthropus* and early *Homo* from Malawi) demonstrates that new productive regions remain to be explored. It is not clear whether australopiths made stone tools, although they were probably collectors of plant foods, using unpreserved wooden sticks and skin bags, and perhaps scavengers of small game animals.

Ardipithecus: still a question. The oldest known likely human species, *Ardipithecus ramidus*, is known from a small group of fossils found at the Aramis locality in the Middle Awash Valley, Ethiopia. (Additional remains of the species have been reported from the nearby Gona locality as well.) The Aramis fossils have been recovered since 1993 from rocks just above a layer dated to 4.4 Ma, which also yielded animal and plant fossils suggesting a relatively wooded environment. The remains of *Ard. ramidus* include teeth, part of a skull base, and the almost complete long bones of a forelimb, as well as many pieces of a fragmented adult skeleton, including parts of almost all regions of the body. As of early 2006, this skeleton had not yet been formally described or illustrated, but the skull is said to be crushed flat and fragmented into hundreds of tiny pieces which have not yet been reconstructed. The elements described so far document a mosaic pattern combining features similar to those of younger humans (for example, the anterior position of the foramen magnum and nearby structures, implying upright posture and presumably bipedalism; and partial reduction of the canines) with others reflecting retention of apelike conditions (for example, thin enamel covering on molars and incisors; lack of a second cusp, or metaconid, on the lower anterior premolar; relatively large size of canines; and apelike shape of the lower anterior deciduous premolar or "milk molar"). It has

been suggested that their presence in a wooded environment implies that australopiths may have differentiated from apelike ancestors in forests rather than open savannahs. But an alternative view suggests that this species represents a "failed" human lineage, one which returned to the forest and secondarily developed thinner enamel convergently with African apes. The remaining bony features of *Ard. ramidus* are all reasonably interpreted as ancestral conditions, to be expected in an ancient human ancestor. Analysis of the partial skeleton should permit determination of the locomotor abilities of this species. Other specimens from Kenya might be referable to this species but are too fragmentary to be sure yet.

Earliest definite humans: Australopithecus. The first habitual bipeds, *Australopithecus*, appear in the fossil record in quantity around 4.2 Ma, during the Early Pliocene. The earliest representatives of *Australopithecus*, named *Aus. anamensis*, have been found since 1994 at sites in the southern Lake Turkana region of Kenya. From Kanapoi, in a layer dated to 4.2–4.1 Ma, come several jaws that differ in shape from those of other *Australopithecus* species. A partial ankle joint from a younger layer confirms upright walking around 4 Ma. It seems likely that *Aus. anamensis* may be close to the ancestry of later species and a possible descendant of *Ard. ramidus*.

Fossils from sites in Ethiopia and Tanzania reveal far more details about a still younger species, *Aus. afarensis* (**Fig. 4**). The most complete material is known from the Ethiopian site of Hadar, about 50 km (30 mi) north of Aramis, where deposits yielded fossils dating between 3.4 and 2.9 Ma. In 1974 a partial skeleton was found and identified as a female by its pelvic bones (and small size compared to other fossils) and nicknamed Lucy. This individual would have stood only 3.5 ft (106 cm) tall and weighed perhaps 65 lb (30 kg). The leg bones of this skeleton indicate that Lucy's legs would have been rather short for a modern human of comparable body weight, but they were in proportion for her estimated stature; this observation supports other evidence that most australopiths were more robustly built than modern humans. Larger males, perhaps 130–175 lb (59–80 kg) in weight, are known from more fragmentary remains, although a nearly complete skull (not known for Lucy) was found in the 1990s and analyzed in detail in 2004 (Fig. 4). Additional specimens, including partial skulls, continue to be found at Hadar. In addition, mandibles and postcranial elements from Maka (across the Awash River from Aramis) and Laetoli (Tanzania) dating around 3.4–3.7 Ma confirm the sexually dimorphic but taxonomically unified nature of the species, while a frontal bone from Belohdelie (3.8 Ma, near Maka) and teeth from the Lake Turkana Basin (3–2.7 Ma) define its temporal range. A fragmentary lower jaw from Chad is estimated to date to a similar time interval by comparing the animal fossils associated with it to those from East Africa; it has been given a new species name but could be a western representative of *Aus. afarensis*.

Fig. 4. Cranium of male *Australopithecus afarensis* from Hadar, Ethiopia, dated to about 3.1 million years ago. (*Courtesy of and © Institute of Human Origins*)

Lucy's pelvis and leg bones, as well as remarkably preserved footprints from Laetoli, clearly demonstrate that upright bipedal walking was well developed by 3.6 Ma, along with a brain somewhat larger than in modern apes of similar body size. Brain size in *Aus. afarensis* might have been between 350 and 450 ml, in a body weighing some 65–154 lb (30–70 kg) as compared to 365 ml in a 100-lb (45-kg) chimpanzee, 500 ml in a 300-lb (135-kg) gorilla, or 1400 ml in a 150-lb (67-kg) living human. Arguments as to the priority of brain expansion or walking ability in human evolution thus have yet to be fully resolved.

Australopithecus afarensis combines both of these advanced, human characteristics with numerous other features reminiscent of later Miocene hominids and modern apes. The lower face is rather projecting (prognathic), the canines project slightly beyond the level of the neighboring teeth, the anterior lower premolar sometimes has only one major cusp, the tooth row is elongated and nearly parallel-sided, and the forehead is low and retreating. However, the mastoid region (below and behind the bony ear opening in the skull) projects inferiorly more than in either living apes or modern humans. Moreover, the long curved fingers and toes support the idea that it often climbed trees, perhaps to sleep, although there is still some controversy over the degree of posible arboreal behavior in this species.

The geologically youngest of the early australopiths is actually the first to have been recognized. *Australopithecus africanus* was named by R. Dart in 1925 on the basis of a juvenile specimen from Taung (South Africa). It appears to have lived from 3 to 2.3 Ma, but so far it is known only from four sites in South Africa; by far the most numerous specimens come from the middle levels (Member 4) of Sterkfontein cave, dated about 2.8–2.6 Ma (**Fig. 5**). It may have stood 4–5 ft (120–150 cm) tall, weighed 65–150 lb (30–68 kg), and had an average brain size of some 450 ml. The skull seems more lightly built than in *Aus. afarensis*, with a rounded vault but more projecting face. The teeth are more humanlike as well, especially in the presence of two cusps on all anterior lower premolars (like all younger hominins) and less projecting canines. Although the absolute tooth size of *Aus. africanus* is nearly equal to that of small gorillas, the proportions are human, with a smooth decrease in size from molars through

Fig. 5. Cranium of adult (female?) *Australopithecus africanus* from Sterkfontein, South Africa, about 2.5 million years old. (*Courtesy of I. Tattersall*)

incisors. Its postcranial skeleton appears quite similar to that of *Australopithecus afarensis*, with a fully bipedal (if by no means modern) locomotor adaptation. Thus, it was perhaps little changed from its putative ancestor, but those few changes are roughly in the direction of later humans. Some workers have suggested that it might lie close to the ancestry of the genus *Homo*, but pending clearer evidence for its phyletic position, that step is not taken here.

In 1995 researchers described bones of a partial foot from the lower levels (Member 2) of Sterkfontein, perhaps contemporaneous with *Aus. afarensis*. These bones were interpreted as indicating a foot partially adapted to grasping and perhaps tree climbing rather than only to bipedal walking. In late 1998 the find of additional parts of the same skeleton was reported, including both legs, some arm bones, and what appears to be most of the skull. All of these elements must be fully removed from the encasing rock and reconstructed before they can be accurately analyzed, but much more of the skeleton may be preserved, in which case it might surpass Lucy in its completeness. Dates between 4 and 3 Ma have been suggested by different techniques (and some scholars even think it might be younger than 2.5 Ma), and it has not yet been determined if it should be identified as *Aus. africanus*, *Aus. afarensis*, or a new species of australopith.

In 1999, Ethiopian and American researchers described *Aus. garhi* from deposits in the Awash valley of Ethiopia, dated to about 2.5 Ma. A partial face and uncertainly associated limb bones were said to represent a new species close to the origin of *Homo*, but other workers have suggested alternative interpretations, such as that these fossils represented a late population of *Aus. afarensis* or female individuals of *Par. aethiopicus*. Stone tools and animal bones with stone tool cut marks were also found in the vicinity, raising the possibility that australopiths may have manufactured stone tools. On the other hand, early members of our own genus, *Homo*, first appear in the fossil record at other sites around this time period and may have been responsible for making and using these tools.

Robust varieties. Until 1986, robust australopiths were known from two forms, usually accepted as distinct species and increasingly given generic status as *Paranthropus: P. robustus* in South Africa and the more extreme *P. boisei* in East Africa. These species lived between about 2.3 and 1.4 Ma (the age of *P. robustus* is known with less certainty, perhaps 1.9–1.6 Ma) and are distinguished from the other australopiths by their larger size and craniodental specializations. They may have been 4 ft, 6 in. to 5 ft, 9 in. (135–175 cm) tall, weighed 80 to 190 lb (36 to 86 kg), had a somewhat heavy and muscular body build, and a brain size of about 525 ml. The skull is robust, with deep cheekbones and thick lower jaw and often a slightly raised sagittal (midline) crest in the middle part of the skull roof from back to front. These features indicate strong chewing muscles and perhaps a diet of tough foods. The teeth themselves are distinctive: the back teeth (molars and premolars) are large to huge; the front teeth (incisors and canines) are quite small and run nearly straight across the front of the mouth. This difference from other hominins, including apes (which generally have large front teeth and small back teeth), combined with a low forehead and a concave, nearly upright face, further suggests adaptation to powerful chewing. The anterior teeth were probably used as much for grinding as for cutting. The scrappy postcranial elements reveal no major adaptive differences from those of *Australopithecus*.

An even more ancient species, (*P. aethiopicus*), presents a combination of many of these *Paranthropus* features plus others that appear to be holdovers from an *Aus. afarensis*–like ancestry. Here, large molars lie behind sockets for rather large incisors and canines; a concave upper face with low forehead sits above a projecting snout; and the sagittal crest extends back to meet the large nuchal (neck muscle) crest (**Fig. 6**). The brain size was small, near 400 ml. The combination indicates that this species emphasized both large front and back teeth, occupying an evolutionary position intermediate between *Aus. afarensis* and the two previously known robust australopiths; its large mastoid area also links it to *Aus. afarensis*. It is also intermediate in time, ranging from 2.7 to 2.3 Ma.

Australopith relationships. The widely accepted view before 1978 was that *Aus. africanus* represented the common ancestor of the robust forms and *Homo*. The discovery of *Aus. afarensis* led to placement of this species as the basal hominin, with several alternative views of its descendants. In the 1990s the newly recovered *Aus. anamensis* was usually ranked as a still older common ancestor, with *Ard. ramidus* held off to the side, in suspense. Early analyses of *P. aethiopicus* considered it as close to the common ancestor of *P. boisei* and *P. robustus*, but some later studies argued that the three robust species were "only" linked by their common possession of a heavily built chewing apparatus, which might have evolved convergently in two or more disparate lineages. These studies proposed (as had earlier workers on other grounds) that *Aus. africanus* might have been ancestral to *P. robustus* in South Africa or that *P. aethiopicus* might have been an early experiment unrelated to later robust species. Such views seem

Fig. 6. Cranium of adult (male?) *Paranthropus aethiopicus* from West Turkana, Kenya, dated to about 2.5 million years ago. (*Courtesy of A. Walker*)

poorly founded, for if the three robust species are in fact linked by the adaptive complex of "heavy chewing" composed of a number of closely similar elements in each form, it is far more likely that they developed from a single common ancestor than that they were convergently comparable but unrelated. Here the robust clade or lineage is recognized as an evolutionary unit, the genus *Paranthropus*, which was thus long-lived but seems to have had no later descendants.

No known australopith clearly shows features that link it to later human species of the genus *Homo*, but many workers have suggested that *Aus. africanus* might represent the closest approach to such an ancestry yet recovered. However, some have thought that several shared features of the skull point to a common ancestry of *Homo* and *Paranthropus* to the exclusion of *Aus. africanus* (see also below). If *Aus. afarensis* is eventually determined to have been (close to) the common ancestor of both *Paranthropus* and *Homo* (plus *Aus. africanus*?), then it will probably be wise to provide a new genus name for *afarensis* (and also *anamensis*), but for the moment all three species are retained in the genus *Australopithecus*. *See* AUSTRALOPITHECINE.

The rise of *Homo*. The only other genus of the Hominini is *Homo*, true humans, into which all later forms are placed. To summarize in advance, in this essay we suggest that after the brief occurrence of two or three early species of *Homo* in Africa, *H. erectus* evolved about 2 Ma and soon spread into Eurasia east of continental Europe; its far-flung populations were connected by periodic population movements spreading genetic modifications and keeping the species relatively unified. *Homo erectus* persisted, at least in eastern Asia, until 0.2 Ma or later, but farther west around 1 Ma, a geographically restricted species arose near the Mediterranean, with possible representatives in southern Europe and northern Africa. This new species in turn gave rise to at least two major lineages: one in Europe led to the Neanderthals, while a second in Africa continued with little change for half a million years. By about 200 Ka (Ka = thousand years ago) our own species *Homo sapiens* arose in Africa and slowly spread across that continent and into Eurasia, where it eventually replaced Neanderthals and all other human forms.

Early Homo. The identification of the earliest specimens of *Homo* is a subject of debate among paleoanthropologists. In the late 1970s the scientific pendulum had swung back to an idea proposed on less secure grounds by L. S. B. Leakey and colleagues in 1964. They named the species *H. habilis*, based on several finds from Olduvai. Especially significant was the discovery of the remains of a juvenile's lower jaw, with teeth much like those of *Aus. africanus*, and its partial skull, with an estimated cranial capacity of about 685 ml, dated about 1.8 Ma. After much argument over the "reality" and distinctiveness of the new species, it was made clear from additional finds at Olduvai, Lake Turkana, and probably a younger level at Sterkfontein (**Fig. 7**) that a relatively small-brained (510–700 ml) and small-toothed *Homo* was present during 2.0–1.5 Ma. This was thought to be younger than *Australopithecus*, older than (most) *H. erectus*, and contemporaneous with *P. boisei*. A partial skeleton discovered at Olduvai in 1986 has similar teeth and an estimated body size comparable to Lucy; this small size was reported as remarkable but should have been expected, given the similarity in skull size to those of the smaller australopiths.

Several fossils, especially from the Lake Turkana region, appeared to represent a different "morph" or structural pattern. These were typified by skull KNM-ER 1470 (its catalog number in the Kenya National Museum) which has a relatively flat but protruding face, a brain size of about 750 ml, a high rounded vault, and probably large teeth (the crowns are broken off). This (and a few more fragmentary specimens) were at first suggested to represent male individuals, while the smaller individuals were considered females of the same species. Leg bones suggesting a size of 4 ft, 9 in. (155 cm) and 110 lb (50 kg) were originally thought to go with these cranial parts, but it later became clear that early *H. erectus* overlapped in time with the more archaic fossils and had comparable leg bones. Unassociated

Fig. 7. Partial cranium of probable *Homo habilis* from Sterkfontein, South Africa, about 1.9 million years old. (*Courtesy of the late A. R. Hughes*)

postcranial elements have had to be set aside pending the recovery of more bones clearly linked to identifiable crania.

One species or two? These two sets of early human fossils pose an active problem in paleoanthropology: either *H. habilis* had greater sexual dimorphism (especially in brain size) than any known primate, or two (rather homogeneous) species were masquerading under a single name, with the smaller set (the supposed females) most similar in facial morphology to both *Aus. africanus* and later *Homo*. There are at least two opposing solutions to this problem. One view holds that all the known specimens from Olduvai, the Turkana Basin, and South Africa represent the single species *H. habilis*, a larger-brained and bigger creature than *Aus. africanus*, but one with essentially the same dental apparatus—another example of mosaic evolution. It would have had as much sexual dimorphism as did *Aus. afarensis*, perhaps at least as much as in modern gorillas or orangutans, in both form and size of teeth, face, and brain.

Another suggestion is that almost all of the Olduvai fossils, the smaller Turkana region specimens and some from South Africa, are recognized as *H. habilis*, while the 1470 specimen and other larger (non-*Paranthropus*) individuals from Turkana are considered as *H. rudolfensis*. This two-species view is accepted here. In 1992 a partial mandible was found at Uraha in Malawi that is extremely similar to a Turkana *H. rudolfensis* specimen. Not only is Malawi geographically intermediate between the known eastern and southern African site regions, but preliminary age estimates based on associated fossil mammals place the find at about 2.5–2.3 Ma, making it one of the oldest representatives of the genus *Homo*. A temporal bone (the lower middle part of the cranium, where the mandible hinges and the ear is housed) of similar age from central Kenya might also represent this species, and some authorities have even suggested that a few Olduvai specimens belong here as well. One new fossil (the lower part of a face with teeth) described in 2003 from the western part of Olduvai shares certain similarities with both the 1470 fossil and the Olduvai lower jaw first included in *H. habilis*. The describers of this new specimen suggested that the name *H. habilis* would thus best be applied to the larger-sized group (including what had previously been called *H. rudolfensis*), requiring a new name for the remaining smaller fossils; the jury is still out. To further complicate matters, some researchers have suggested that 1470 and its closest relatives might be linked to another Kenyan cranium found in 2001 and dated close to 3.5 Ma. Termed *Kenyanthropus platyops* ("flat-faced Kenya man"), this fossil is somewhat "squashed" and its surface bone broken into many small, slightly separated pieces. It might represent a new lineage of Pliocene hominins; that lineage might include 1470, renamed as *K. rudolfensis*; or the new fossil might just be a badly damaged representative of a known species, such as *A. afarensis*. For the moment, it seems wisest to accept just two named early species of *Homo*, with other options awaiting stronger supporting evidence.

Although australopiths were habitually bipedal, limb bones questionably attributed to *H. rudolfensis* suggest a wider pelvis, perhaps to permit the birth of larger-headed (because larger-brained) infants, which resulted also in the leg bones themselves appearing more modern. It is this combination of apparently significant change in the two major human adaptations of locomotion and intelligence that leads most authors to classify these species as *Homo*, but some have placed both of them in *Australopithecus*. Those researchers take the position that the anatomy of *H. habilis* and *H. rudolfensis* suggests their "adaptive strategies" were more like that of *Australopithecus* than other *Homo* species. The problem with this argument is that most other studies support a stronger evolutionary relationship between *H. habilis*, *H. rudolfensis*, and the other members of the genus *Homo*, including *H. erectus*, Neanderthals and modern humans, than with *Australopithecus*. While the former idea has not taken hold widely in the paleoanthropological community, it does suggest that a fundamental shift in adaptive strategies took place after these early *Homo* species.

It is not clear which species of early *Homo* might have been ancestral in turn to *H. erectus*. Both early forms have been found in the 2.0–1.6 Ma time range in the Turkana Basin and probably Olduvai, where they overlapped with *P. boisei* and *H. erectus*. But neither *H. habilis* nor *H. rudolfensis* as known are morphologically very close to early *H. erectus*. The former species differs not only in small brain and body size but also in limb proportions, although its facial morphology is potentially acceptable in an ancestor of *H. erectus*. The brain of *H. rudolfensis* is large, as might be expected in an ancestor of *H. erectus*, but its teeth are large and distinctive, its face is unique, and its limb bones have not been satisfactorily identified. One wonders if there might not be an as yet undiscovered early species of *Homo* which combined smaller teeth and face with a larger brain. Perhaps the upper jaw from Makaamitalu (found in 1994 high in the Hadar sequence) dated about 2.35 Ma might be a step toward fulfilling this prediction.

Early Homo technology. Evidence of archeological activity has been found with or near some of these early *Homo* specimens, mostly flakes and choppers or "pebble tools" of the Oldowan stone-tool industry (see **table**), and also the remains of small animal prey. It is not possible to tell which types of humans made the tools, but at present the oldest evidence is from Ethiopia, in the Gona region (near Hadar and Aramis) and the Omo valley, about 2.5 Ma. The close time and space concordance of Oldowan tools and early *Homo* suggests that at least one of these species was the actual toolmaker; whichever it was may have been an active hunter of small game, collector of plant and insect food items, and scavenger of larger mammals. *See* PALEOLITHIC; PREHISTORIC TECHNOLOGY.

Homo erectus. While *H. habilis* and *H. rudolfensis* apparently were short-lived and relatively rare African species, their likely successor, *H. erectus*, was common, widespread, and long-surviving. The

Old World archeological variants before 10,000 years ago

Technological mode	Description	Paleolithic subdivisions	Selected regional industries*
4	Blade cores, to produce scrapers, points, burins; bone harpoons, art	Late (Upper) Paleolithic	Later Stone Age [sub-Saharan Africa, 40 to < 5 Ka (mode 5, microlithic)] Upper Paleolithic of Europe, including Perigordian Solutrean, Magdalenian Aurignacian (Europe, SW Asia?, > 40 to 27 Ka)
3	Widespread use of prepared cores to obtain variety of flake forms, used as scrapers and projectile points	Middle Paleolithic	Chatelperronian (western Europe, 36–30 Ka) Aterian (northern Africa, 100 to 30 Ka) Mousterian (many variants; western Eurasia and northwestern Africa, 200 to 35 Ka) Middle Stone Age (many variants; sub-Saharan Africa and southern Asia, 250 to 40 Ka)
2	Large bifaces (handaxes and cleavers), also simple cores and retouched flakes	Early (Lower) Paleolithic	In late stages, use of prepared cores and wooden throwing spears Acheulean (Africa and southwestern Asia, 1.5 to <0.2 Ma; eastern Asia, 0.8 to <0.2 Ma; Europe and southern Asia, 0.5? to <0.2 Ma)
1	Simple cores and "casual" flakes	Early (Lower) Paleolithic	Zhoukoudian (China, 1 to 0.25 Ma) Various (Europe, >1 to <0.4 Ma) Oldowan (Africa, 2.5 to <1.5 Ma)

*Ka = thousand years ago; Ma = million years ago.

first fossils were found in Java in 1893 and termed *Pithecanthropus erectus*. Most of the later finds in China and across Africa were given distinctive generic and specific names, but all are now usually considered local variants or subspecies of the single species *H. erectus*. Some scholars still recognize additional species, especially *H. ergaster* for the earlier East African populations, but here only one species is accepted. The major anatomical characteristics of this form are the following: a body of nearly modern form and proportions below the neck, topped by a low and slightly elongated skull with cranial capacity averaging 1100 ml (with a range of about 800–1300 ml), smaller teeth in a less projecting face than *Aus. africanus* or *H. rudolfensis*, large solid brow ridges, thick cranial bones, and no chin.

African populations. The earliest specimens are probably from East Africa, dating to as much as 1.9 Ma. These populations might have continued to make simple Oldowan artifacts, as no other form of stone tools is known at this time. There is evidence of the use of fire as early as 1.8 Ma and in scattered sites thereafter. Acheulean bifaces (handaxes and cleavers) appear by 1.5 Ma, though only rarely in direct association with *H. erectus* fossils; evidence for group hunting of large animals is rarer still and has been questioned by some scholars. In addition, the eventual extinction of *P. boisei* in this time interval may have been caused by direct or indirect competition with the more advanced *H. erectus*.

The most important African *H. erectus* find of recent decades is a mostly complete skeleton of an adolescent male (its sex judged by pelvic bones as well as by relative cranial robusticity), recovered on the west side of Lake Turkana in 1984 (**Fig. 8**). Judging by the pattern of dental eruption, he was probably 9–10 years old at death, although an age of 11–12 has also been suggested. One still ambiguously answered question is whether this species underwent the adolescent growth spurt which occurs in modern children of ages 12–14. Using formulas based on limb bone lengths of modern human adults and children, it was estimated that the "Turkana boy" might have reached an adult height of nearly 6 ft (183 cm) and a body weight of 150 lb (68 kg). Moreover, his body shape would have been slender or elongate, much as seen in living people of the same region, as a result of similar adaptation to a warm and dry climate. The development of brow ridges and other cranial structures in the adolescent male and a larger skull from Olduvai contrasts markedly with other Turkana skulls, indicating the persistence of strong sexual dimorphism.

A number of interesting new fossils thought to be *H. erectus* were found in East Africa in 1995–2002. A relatively complete cranium without a face from Lake Turkana and cranial fragments from another Kenyan site represent quite small individuals. The latter was found in the same stratigraphic layer as many bifacial Acheulean tools, good evidence that they were indeed the toolmakers. The first potential *H. erectus* fossils from Ethiopia and Eritrea have also been recently described, both from around 1 Ma. Measurements of the Ethiopian fossil indicate its similarity to both African and Asian *H. erectus* fossils. The Eritrean fossil, on the other hand, shows an increase in width of the back of the skull that is also seen in modern humans. Together, these new fossils extend the known variation in African *H. erectus* considerably.

Eurasian representatives. Homo erectus (perhaps as a result of increasing population size) spread from Africa into Eurasia through the Middle East, perhaps earlier than has previously been thought. New dates for some long-known Javanese fossils suggested ages around 1.7 Ma, but these have been questioned.

Fig. 8. Partial skeleton of adolescent male *Homo erectus* from West Turkana, Kenya, about 1.5 million years old. (*Courtesy of and © National Museums of Kenya*)

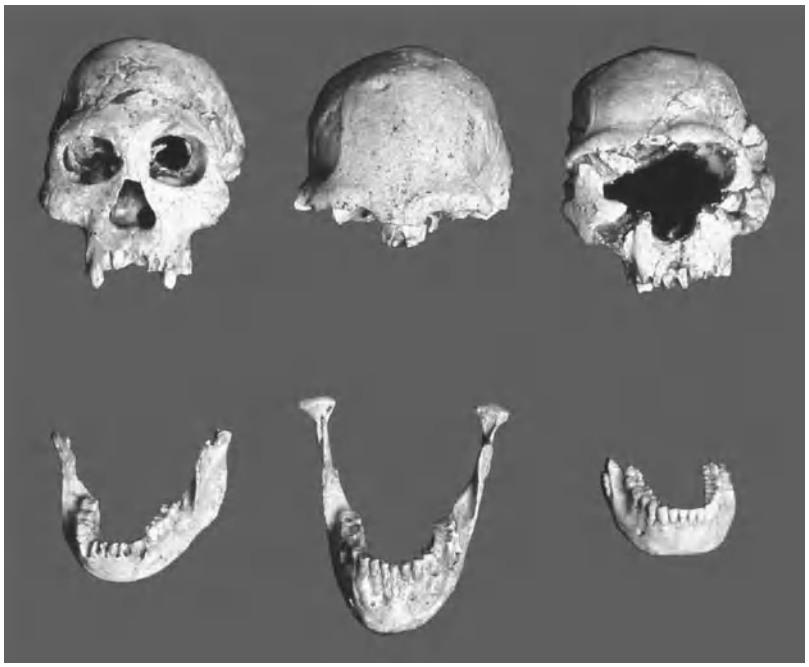

Fig. 9. Crania and mandibles (in front view) of *Homo erectus* from Dmanisi, Georgia, reveal diversity in a single sample dated ca 1.7 Ma. (*Courtesy of and © Georgian National Museum*)

The most exciting new finds have come from the site of Dmanisi (Georgia), mainly from 1999 to 2005 (**Fig. 9**). There are currently five crania, four mandibles and a large number of postcranial elements along with Mode 1 (Oldowan-like) tools and animal bones dated to 1.8–1.6 Ma. Mode 1 and early Mode 2 artifacts from Israel (as yet not definitively found with human fossils) are slightly younger. Thus, *H. erectus* in the broad sense accepted here must have been the first human species to leave Africa in large numbers, probably about 1.8 Ma, soon after it first appeared in the fossil record.

Fossils of this species may extend in mainland Asia to nearly 200,000 years ago, mostly associated with fauna from the warmer intervals in this time of alternating glacial climate. From central Java, Indonesia, come a series of cranial and dental remains spanning from perhaps 1.7 Ma to possibly 30 Ka. The earlier range of time yielded several groups of skulls known popularly as Java man, including the first found members of this species; these specimens (and also those from China) tend to be more derived or extreme in their morphology than most African specimens, including the common presence of a thickened "keel" of bone down the midline of the braincase and a more projecting browridge. A dozen younger partial skulls from Ngandong, Java were long thought to be quite recent in age and once were placed in *H. sapiens*, but later study indicated close similarity to earlier Indonesian *H. erectus*, as they are now classified. Their age remains uncertain, with some evidence suggesting a range between 400 and 100 Ka, while a dating attempt in 1996 arrived at the surprisingly young age of 50–30 Ka, which has been questioned on several grounds. If this is indeed a correct date for these fossils, this would be the first demonstration of the coexistence of *H. erectus* with anatomically modern humans, known elsewhere in Indonesia and Australia at that date. *Homo erectus* specimens continue to be recovered from Java, and one partial cranium was even recognized in 1999 at a New York City shop selling natural history specimens (**Fig. 10**). It was determined that it had been found in Indonesia some years earlier and illegally removed from that country; the shop's owner returned it to Indonesian authorities.

A fascinating find was made in 2003 on the Indonesian island of Flores in a cave called Liang Bua, which has an archaeological record preserved from 200 Ka to the present. The fossils found there, including a well-preserved skull, additional lower jaws, and postcranial bones, have been dated to about 18 Ka. Dubbed "Hobbits" in the popular press, the remains are those of a small-brained (380 ml) and small-bodied (3 ft 3 in. or 1 m) hominin with a skull that resembles that of a small *H. erectus*. This led the discovers to conclude that a population of *H. erectus* somehow reached the island over water and became "dwarfed" after isolation, a phenomenon observed in other large mammals on islands; they named the fossils *H. floresiensis*. Other scientists have suggested that the remains belonged to pathological modern humans, which

Fig. 10. Partial cranium of ?female *Homo erectus* from Sambungmacan, Java, Indonesia, age uncertain (probably between 500,000–100,000 years ago). This specimen was identified in a natural history store in New York City, years after it was found by local workers in Indonesia. (*Courtesy of and © E. Delson; photo by C. Tarka*)

seems unlikely given the number of individuals of differing ages recovered. Further finds published in late 2005 highlighted the long arms and other anomalous postcranial features of this species, which are difficult to explain.

The first Chinese *H. erectus*, once called Peking man, was found at Zhoukoudian, near Beijing, where they occupied a large cave during most of the period between 500 and 250 Ka. Additional fossils of this form are known from Yuanmou in the southwest, Lantian and Yunxian in the center, and Hexian and Tongshan (Nanjing) in eastern China; some may be as old as 1–1.2 Ma. Archeological remains found with these fossils are mainly of the simple Mode 1 variety (no stone tools are unambiguously known from layers with human fossils in Java). However, more complex bifacial tools have been found in the Bose Basin of southern China, from about 0.8 Ma. These tools are comparable in overall form and complexity to Acheulian (Mode 2) tools found in Africa at this time and predate Mode 2 tools in Europe. This contradicts the long held hypothesis that only simple Mode 1 tools were made in East Asia while more complex Acheulean tools were restricted to Africa and Europe at this time. Although there have been claims, no definite *H. erectus* fossils are yet known from Europe, nor are archeological remains or more modern humans unambiguously documented there older than about 800,000 years. The Dmanisi specimens from Georgia are the most ancient human representatives in the region, but the lack of further remains over the succeeding million years may indicate that *Homo* required further physical or cultural adaptations in order to survive in Europe proper.

Middle Pleistocene Homo. It has been suggested that the increased rigor of the glacial climate in Europe at this time was the impetus leading to the evolution of humans who seem to be physically more modern in several ways than Afro-Asian *H. erectus*. These people were often termed early or archaic *H. sapiens*, or sometimes placed in their own species, *H. heidelbergensis*. This view was acceptable so long as the most ancient African representatives of this group were poorly dated or younger, but studies in the later 1990s suggested some modification. The earliest human fossils in Europe were long thought to date to about 500 Ka in England (a tibia from Boxgrove) and Germany (the mandible from Mauer near Heidelberg found in 1908). The long-known human fossils from Tighenif (previously called Ternifine, in Algeria), dated to 800–700 Ka, were transferred out of *H. erectus* by some workers because they present at least one derived feature of the lower jaw. Moreover, the partial skull from Bodo (Ethiopia) was also dated older than 600 Ka, and similar South African fossils were estimated to be of comparable age. These dates suggested that the earliest representatives of these more derived humans may have lived in the northern half of Africa and perhaps evolved there from local *H. erectus* populations early in the Middle Pleistocene (780–125 Ka).

Then, in 1995, fragmentary human fossils and associated stone tools from the older levels at Atapuerca (Spain) were dated to about 800 Ka, implying that archaic *H. sapiens* appeared at about this date all over the western Old World. These specimens were named *H. antecessor* in 1997, and it was suggested that they represented the common ancestor of all later human varieties. Meanwhile, in 1996, a cranium lacking the face was described from Ceprano, Italy, in a context suggestive of a date of about 700 Ka or more. However, this date is not particularly secure. A revised reconstruction claims strong similarity to *H. erectus*, otherwise unknown in Europe. In late 2005, stone tools (of both Mode 1 and Mode 2 types) and animal fossils were described from Pakefield in eastern Britain, a site estimated to date to 700 Ka; no human fossils are yet known, but they are eagerly awaited.

All of these finds, combined with theoretical arguments about the best way to recognize and delimit species in the fossil record, have led to competing interpretations of the number of species of *Homo* known in the past million years. Some workers continue to place all post-*erectus* fossils in "archaic *Homo sapiens*," sometimes recognizing a variety of temporal and geographic subspecies (such as the Neanderthals and anatomically modern humans). A few have gone so far as to include *H. erectus* within an over-enlarged *H. sapiens*. At the other extreme, some researchers accept between five and seven species in the same time period: *H. antecessor*, *H. cepranensis* (for Ceprano) *H. heidelbergensis* (either restricted to Europe or extended to Africa and even East Asia), *H. rhodesiensis* (for African Middle Pleistocene populations), *H. neanderthalensis*, *H. sapiens* (restricted to anatomically modern humans), and perhaps others. A possible middle ground would be to (1) include the earliest of these African and European populations in one named group [for example, *antecessor*, or perhaps *mauritanicus* (the name originally given to the Tighenif fossils), as suggested by J.-J. Hublin]; (2) combine all post-500 Ka nonmodern European fossils in *neanderthalensis* (including *heidelbergensis* here); (3) group

nonmodern African fossils younger than Tighenif in *rhodesiensis*; and (4) restrict *sapiens* to anatomically modern humans worldwide. These "groups" could be given specific or subspecific status, depending on the theoretical model used. Recent work discussed below supports full species status for Neanderthals based on several lines of evidence, and all four groups are here interpreted as full species. While both fossil and genetic evidence supports an African origin for modern humans, the evolutionary relationships among all of these groups remains unclear.

Homo rhodensis and early representatives of *H. neanderthalensis* occur in Africa and Europe, respectively, between 650 and 250 Ka, thus contemporaneous with *H. erectus* populations in eastern Asia. They share somewhat larger brains (for body size), smaller teeth, more expanded facial sinuses and occiput (rear of the skull), but less robustness than in *H. erectus*. All these features are found in more extreme form in modern humans and in the late "classic" Neanderthals (see below). In most areas, these people still used Acheulean tools, but perhaps with greater efficiency. It is likely that these "intermediate-grade" humans spread gradually eastward across the Old World, replacing late-surviving populations of *H. erectus* everywhere by 200 Ka. These geographic variants were not only distinct from *H. erectus* but also from each other to a greater degree than is true among living varieties or "races" of anatomically modern humans. In southern Africa, one cranium was found at Broken Hill, now Kabwe, Zambia (formerly Northern Rhodesia, hence the name Rhodesian man), and broadly similar specimens are known in South Africa (Saldanha and Florisbad), Tanzania (Ndutu), Ethiopia (Bodo), and Morocco (Salé and Thomas quarries). These people made Acheulean or equivalent Mode 2 tools and apparently hunted big game across most of Africa. In early 2006, a new partial cranium probably belonging to this species from Gawis (Ethiopia, near the Gona Pliocene tool site) was reported by Ethiopian and U.S. researchers led by Sileshi Semaw. Its age is as yet uncertain but probably in the 500–250 Ka interval.

Fig. 11. Crania in frontal view of male *Homo neanderthalensis* from La Ferrassie, France (on right) and Amud, Israel, ca 50,000–60,000 years old. (*Courtesy of and © Israel Antiquities Authority (Jerusalem), Musée de l'Homme (Paris) and E. Delson; photo by C. Tarka*)

Rare specimens from China appear to be younger, mainly dating to about 250–150 Ka. These include a nearly complete cranium from Dali, in central China, and a partial skull and skeleton from Jinniushan (or Yingkou), in the northeast, as well as scattered, less complete remains. All of these fossils, especially Dali, are broadly similar to the African specimens just mentioned, as well as to some of the earliest European *H. neanderthalensis*. Other Chinese specimens, such as Maba (from the southeast), and the central Indian Hathnora (or Narmada) fossil, are partial crania which are both younger (perhaps about 150–75 Ka) and more derived morphologically, although not in the direction of either early anatomically modern people or the contemporaneous Neanderthals.

Neanderthals. The best known of the archaic varieties are the Neanderthals, from Europe and western Asia. It now seems likely that this group evolved locally in Europe from early Middle Pleistocene *Homo* via intermediate populations previously called *Homo heidelbergensis* such as those known from England (Swanscombe, ?Boxgrove), Spain (Atapuerca), France (Arago, Montmaurin), Germany (Mauer, Steinheim), and Greece (Petralona). They became adapted to the cold climates of glaciated Europe, with prototypical Neanderthal anatomy well established by about 200 Ka. During the warm interval about 120 Ka, they may have spread into the Near East and central Asia. In the cold glacial phase between 110 and 35 Ka, "classic" (or extreme), cold-adapted Neanderthals were abundant in cold northern parts of western and central Europe, while less extreme forms (perhaps more like their immediate predecessors) inhabited areas to the south and east. They were essentially stocky humans, but had long, low skulls with a projecting occipital region, large faces, teeth, and brow ridges; and brains averaging 1500 ml in volume (**Fig. 11**). Their limbs and trunks were heavily muscled, indicating great strength, but many bones were broken and healed during life. They made Mousterian tools (a variant of Middle Paleolithic or Mode 3 flake-based tool kits), often lived in caves or wooden shelters where they controlled fire, hunted big game, and had primitive religious beliefs, including burial of the dead with grave goods. There is intense argument among paleoanthropologists as to how modern the Neanderthals were behaviorally, in terms of their stoneworking and hunting techniques and modes of foraging, whether planned or merely ad hoc. For example, the frequent broken bones may have resulted from hunting large game at close proximity, rather than using projectile technology from a distance. Such controversies feed back into the question of whether the Neanderthals are a distinct species or a distinctive subspecies of *H. sapiens*. A related question is whether the Neanderthals were in any way ancestral to anatomically modern humans, especially of Europe.

Recognition of a separate Neanderthal species implies an almost absolute reproductive isolation and

lack of genealogical continuity, while the opposite is true for most interpretations of Neanderthals as members of *H. sapiens*. The interpretation accepted here is that *H.neanderthalensis* is considered to have been geographically and culturally isolated from early anatomically modern humans and their ancestors, two independent but closely related lineages (species) evolving in parallel until they finally met, after which the former group soon became extinct. Indeed, current research suggests that differences in the skulls between Neanderthals and modern humans appeared early in individual growth, meaning that Neanderthal children looked quite different from modern human children. Neanderthals also differ in developing more rapidly than modern humans. In other words, they achieved their adult anatomy faster than modern humans do. In addition, work by K. Harvati and colleagues published in 2004 shows that the differences in skull anatomy between Neanderthals fossils and modern humans are at the level of separate species when compared to other primates. Furthermore, when scientists extracted mitochondrial DNA from the bones of several Neanderthal fossils they found that Neanderthals have some unique sequences of DNA when compared to modern humans. Neanderthals, possibly due to local climatic differences, were probably isolated from other hominins and as a result became quite distinct in their development as well as their skull shape compared to modern humans. *See* NEANDERTHALS.

Spread of modern humans. One of the major foci of recent paleoanthropological research is the clarification of the area of origin and the early history of modern humans, *H. sapiens*. The skull of this form is characterized by a small, nonprotruding face; small teeth and brow ridges; a chin; and a high, rounded braincase. There are no specimens of this type known (or even hinted at) anywhere in the world earlier than about 200–150 Ka. New fossils representing early modern humans were recovered from Herto, Ethiopia, in 1997 (published 2003) and dated 164–150 Ka. Aside from having larger faces than recent humans, these are quite modern looking. Moreover, in 2005, the long-known fossils from Kibish (Omo Valley, Ethiopia) were dated close to 200 Ka. This confirms previous suggestions that about 200–100 Ka some fossils from eastern and southern Africa (for example, one from Kibish and one from a late horizon at Laetoli) suggest the persistence of a "Rhodesian-like" morphology, while others (for example, another from Kibish and several from Djebel Irhoud, Morocco) appear to be nearly modern. Two somewhat younger sites in South Africa have produced important additional modern human fossils. At Border Cave, a partial cranium and other fragments may date to nearly 90 Ka; they are clearly modern in form, but their date is questionable. The Klasies River Mouth caves, on the southern coast, have yielded a sequence of layers with good dates and archeological context; the human remains dated about 100 Ka are scrappy but appear modern, with a chin, small brow ridges, and overall gracility. Such gracilization, especially of the face, is a major feature of the evolution of (and within) *Homo sapiens*. In combination, these remains and other, less complete fossils indicate that early moderns were living in sub-Saharan Africa by 150 Ka. Archeological remains dated 100–75 Ka in South Africa and D. R. Congo (ex-Zaire) indicate that at least some of these people were making Mode 3 (MSA) tool kits with elements (such as bone harpoons, engraved bone, and minerals and shell beadwork) that do not appear in Europe until after 30 Ka. *See* EARLY MODERN HUMANS; PALEOLITHIC; PREHISTORIC TECHNOLOGY.

From such a possible sub-Saharan origin, anatomically modern *H. sapiens* appear to have spread across the Old World, differentiating into local races by 80–50 Ka. This view of human dispersal has received support from studies of the distribution pattern of human mitochondrial DNA haplotypes (variants) and other genetic evidence. The majority of these studies suggest that the major dichotomy in modern human population genetics is between Eurasians and Africans. Such results fit well with the fossil evidence for African versus Eurasian divergence about 100 Ka. Moreover, dates on early anatomically modern remains from Israel (Djebel Qafzeh and Skhul) documented the presence of the ancestors of Eurasians outside Africa by about 110–90 Ka. This is especially intriguing because most Israeli Neanderthals have been dated to about 65–45 Ka, which is significantly younger than the early moderns and implies that they may have overlapped in this region of the world. Even more complexity is implied by the strong similarity of Mousterian tool kits associated with such diverse human forms as the Neanderthals of Europe and Israel (and farther east), the early moderns from Israel, and the "premoderns" from Jebel Irhoud (Morocco), but the implications of this cultural similarity are as yet unclear.

Found in the late 1970s, the youngest known Neanderthal skull comes from southern France and is associated with tools of the Châtelperronian industry, a Mode 3 or 4 variety previously thought to have been made by *H. sapiens*. This specimen dated to 34 Ka has been alternatively interpreted as the maker of these tools (possibly after contact with Late Paleolithic moderns) or as evidence for direct Neanderthal ancestry of moderns. In some cases, more fossils do not answer questions but create new ones. Sites in southern Spain and Portugal have yielded less complete Neanderthal fossils and Mousterian tools dated about 30 Ka, after which modern *H. sapiens* was the sole form of human to be found anywhere. One reason for the success of *H. sapiens* may have been their greater tool-making efficiency, as documented by the Late (or Upper) Paleolithic Mode 4 blade-and-burin industries. These people included large quantities of worked bone in their tool kits (using chisel-like burins to carve and engrave the bone), constructed dwellings of wood or of already fossilized animal bone, hunted large game, fished with harpoons, and in general behaved much like their living descendants. In many parts of the world, they also wore personal ornamentation (jewelry) and engaged in artistic pursuits, including carving small animal statues and perhaps calendars, as well as painting on the

walls of rock overhangs and deep caves. *See* PALEOLITHIC; PREHISTORIC TECHNOLOGY.

In 1999, Portuguese and American paleoanthropologists described the remains of a 4-year-old child, buried near Lagar Velho, Portugal, about 25 Ka. The burial pattern and most of the child's morphology indicated links to the Gravettian culture, made by early anatomically modern people elsewhere in Europe at this time. However, features of the lower leg bone (tibia) and lower jaw (the cranium was crushed and later reconstructed) suggested similarity to Neanderthals. The describers hypothesized originally and in a major publication in 2005 that this individual might have been the result of hybridization between Neanderthals and moderns, but other researchers argued that the morphology was not that different from what could be expected in a robust anatomically modern child and that hybridization would have resulted in features intermediate between Neanderthals and moderns, not clear features of each.

Many names have been given to early modern humans, especially in Europe, but these indicate only minor differences. The term Cro-Magnon derives from several skeletons found in 1868 in Les Eyzies, France. They gave their name to a "race" said to occur either just in France or across most of Europe. In fact, Cro-Magnon people were already essentially Europeans, while early Africans are known from sites in eastern and southern Africa. Australasia was colonized over water after about 70 Ka, with important finds at Keilor and Lake Mungo. New World Indians likely originated from Siberia, by means of crossing a land bridge over what is now the Bering Strait. Many human fossil remains are known in the Americas as far back as 12 Ka, but some dates as old as 35 Ka have been obtained on archeological sites, indicating that perhaps several crossings of the land bridge occurred.

In contrast to the above "Out of Africa" view of human dispersal (based on the idea that modern humans evolved in sub-Saharan Africa more than 100,000 years ago from populations of archaic *Homo*) accepted here, a minority view (the "Multiregional" hypothesis) interprets the fossil record to document the nearly parallel origin of modern humans in different regions of the Old World from a *H. erectus* ancestry. Each regional variety is said to present morphological characteristics linking archaic to modern populations, while gene flow (migration and interbreeding) between regions kept the geographical varieties united in a single species at any one time. While most researchers reject a significant contribution from Neanderthals or *H. erectus* to the modern human gene pool, there may still have been some interbreeding among the species. *See* EARLY MODERN HUMANS; PALEOLITHIC; PREHISTORIC TECHNOLOGY.

Eric Delson; Karen L. Baab

Bibliography. E. Delson (ed.), *Ancestors: The Hard Evidence*, 1985; E. Delson et al. (eds.), *Encyclopedia of Human Evolution and Prehistory*, 2d ed., 2000; J. G. Fleagle, *Primate Evolution and Adaptation*, 2d ed., 1998; W. C. Hartwig, *The Primate Fossil Record*, 2002; R. G. Klein, *The Human Career: Human Biological and Cultural Origins*, 2d ed., 1999; J. H. Schwartz and I. Tattersall, *The Human Fossil Record*, vols. 1–4, 2001–2005; A. Walker and R. Leakey (eds.), *The Nariokotome Homo erectus Skeleton*, 1993; P. F. Whitehead, W. K. Sacco and S. B. Hochgraf, *A Photographic Atlas for Physical Anthropology*, 2005.

Fossil primates

Extinct members of the order of mammals to which humans belong. All current classifications divide the living primates into two major groups (suborders), but zoologists differ as to whether the tarsier (*Tarsius*) should be classified with the lower primates (lemurs, lorises, and bushbabies) or the higher primates (New and Old World monkeys, greater and lesser apes, and humans).

All primates have a common origin that, however, is not reflected in the universal possession of a suite of diagnostic features. The order as a whole has been characterized in terms of showing a group of progressive evolutionary trends, notably toward predominance of the visual sense, reduction of the sense of smell and associated structures, improved grasping and manipulative capacities, and enlargement of the higher centers of the brain. Among the extant primates, the lower primates more closely resemble forms that evolved relatively early in the history of the order, while the higher primates represent a group more recently evolved (**Fig. 1**).

A classification of the primates, as accepted here, is shown on the next page.

Early primates. The earliest primates are placed in their own semiorder, Plesiadapiformes (as contrasted with the semiorder Euprimates for all living forms), because they have no direct evolutionary links with, and bear few adaptive resemblances to, any group of living primates. However, the chewing teeth and the locomotor anatomy of these fossil forms sufficiently resemble those of later primates to demonstrate the common origin of the two groups. Best known from the Paleocene Epoch, around 65–55 million years ago (Ma), and found in both the Old World and the New World, most plesiadapiforms retained clawed hands and feet. However, in 2002 it was shown that at least one species, *Carpolestes simpsoni*, possessed an opposable great toe with a nail, as in euprimates (**Fig. 2**). Plesiadapiforms also had rather small brains compared to their body size, possessed large specialized front teeth, were probably arboreal in habit, and may have been adapted for terminal branch feeding, as recently suggested. These animals are known from fossil deposits as far north as Ellesmere Island, in Arctic Canada, which, during the Paleocene, was covered by the subtropical forest stretching continuously from western North America across a landlocked North Atlantic into western Europe.

Eocene primates. There is no known plesiadapiform that is a satisfactory candidate for the ancestry of the fossil primates of modern aspect typical of the succeeding epoch, the Eocene (55–34 Ma). Often

Primates
- Semiorder Plesiadapiformes (archaic extinct primates)
  - Superfamily: Paromomyoidea (paromomyids, picrodontids)
    - Plesiadapoidea (plesiadapids, carpolestids)
- Semiorder Euprimates (modern primates)
  - Suborder Strepsirhini (toothcombed prosimians and extinct allies)
    - Infraorder Adapiformes (extinct early strepsirhines)
      - Superfamily Adapoidea (extinct early strepsirhines)
    - Infraorder Lemuriformes (modern strepsirhines)
      - Superfamily: Lemuroidea (typical lemurs)
        - Indrioidea (indris, aye-ayes, and subfossil relatives)
        - Lorisoidea (lorises, bush-babies, mouse and dwarf lemurs)
  - Suborder Haplorhini (tarsiers and higher primates)
    - Hyporder Tarsiiformes (tarsiers and extinct relatives)
      - Superfamily: Tarsioidea (tarsiers and close relatives)
        - Omomyoidea (extinct early haplorhines)
    - Hyporder Anthropoidea (higher primates)
      - Infraorder Paracatarrhini (archaic anthropoids)
        - Family: Oligopithecidae (archaic protoanthropoids)
          - Parapithecidae (extinct Oligocene monkeys)
      - Infraorder Platyrrhini (New World anthropoids)
        - Superfamily Ateloidea (New World monkeys)
          - Family: Atelidae (howler, spider, saki, titi, and owl monkeys)
            - Cebidae (squirrel, capuchin, and marmoset monkeys)
      - Infraorder Catarrhini (Old World anthropoids)
        - Parvorder Eocatarrhini (archaic catarrhines)
          - Family: Propliopithecidae (extinct common ancestors of hominoids and cercopithecoids)
            - Pliopithecidae (extinct early catarrhines)
        - Parvorder Eucatarrhini (advanced catarrhines)
          - Superfamily Hominoidea (apes and humans)
            - Family: Proconsulidae (extinct early apes)
              - Hylobatidae (gibbons, lesser apes)
              - Hominidae (great apes, humans, and extinct relatives)
          - Superfamily Cercopithecoidea (Old World monkeys)
            - Family Cercopithecidae (Old World monkeys)

termed euprimates, they are divided broadly into lemurlike forms, usually grouped into the superfamily Adapoidea, and tarsierlike forms (Omomyoidea), although this division may prove to be oversimplified. Eocene primates of both the Old and New Worlds already display the trends that mark modern primates as a whole: These arboreal animals possessed grasping hands and feet in which sharp claws were replaced by flat nails backing sensitive pads; the snout was reduced, suggesting a deemphasis of smell, while the bone-ringed eyes faced more forward, producing stereoscopic vision and suggesting primary reliance on the sense of sight; and the brain was somewhat enlarged relative to body size when compared to those of other mammals of the period. A new fossil skull from China reported in 2004 (**Fig. 3**) documented the presence of omomyoids in the earliest Eocene of eastern Asia, and its relatively small orbits suggested that it was diurnal.

It is possible that the origins of some specific groups of extant lower primates may be traced back to or through certain known primate genera of the Eocene. In 2003, two distinct primate dentitions bearing the toothcomb characteristic of all modern strepsirhines were documented in Egypt to date back to the Eocene epoch, around 40 Ma. It has been suggested that one of these forms is related to the living bushbabies and the other to the lorises. Similarly, some dental fossils reported in 2001 from 30 Ma deposits in Pakistan have been identified as lemuroids. Even if this is not the case, however, it is universally accepted that the antecedents of the living lower primates are to be sought somewhere within the Eocene primate radiation, though the details of this ancestry still remain unclear. In North America and Europe, lower primates had nearly disappeared by the close of the Eocene, while virtually all known fossil lower primates of later epochs in Africa and Asia are closely related to modern primates of these areas.

Modern lower primates. The extant lower primates are allocated to the suborder Prosimii if *Tarsius* is included, and to the suborder Strepsirhini if this strange primate is excluded, as is provisionally done here. All extant strepsirhines possess dental scrapers (toothcombs), in which the lower front teeth are elongated, closely approximated, and forwardly projecting. These specialized structures are used both

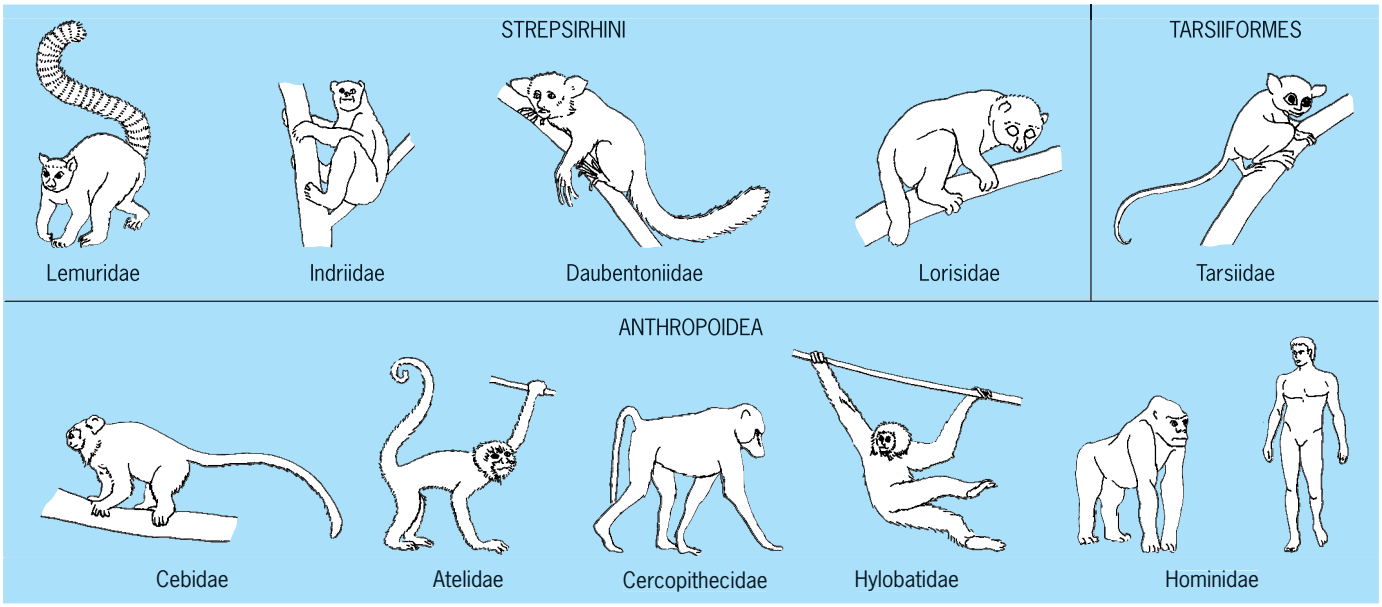

Fig. 1. Representatives of living primate families.

in feeding and in social grooming. Additionally, all living strepsirhines retain a moist, naked rhinarium (wet nose, as in a dog) and related structures, reflecting a greater reliance on the sense of smell than is typical of higher primates. All are also united by possessing a toilet claw (used for self-cleaning) on the second digit of the foot. Although all strepsirhines possess grasping hands and feet, their manual dexterity is generally inferior to that of higher primates; their brains are also relatively small.

There is no ancient (older than 10 Ka) primate fossil record in Madagascar, home of the most diverse group of modern lower primates, but fossils of now-extinct species document a much wider adaptive radiation before the arrival of humans on the island less than 2000 years ago. Notable among those extinct lemurs are *Palaeopropithecus*, a large-bodied under-branch hanger, somewhat slothlike in its locomotion (**Fig. 4**); the semiterrestrial and dentally specialized members of the subfamily Archaeolemurinae, related to the living indriids but adapted somewhat like baboons; and *Megaladapis*, a very large, vertically postured, short-limbed, and probably slow-moving climber with adaptive resemblances to the koala of Australia.

Tarsiers. The tiny *Tarsius*, which lives today in Southeast Asia, represents a link between the toothcombed strepsirhines and the anthropoids

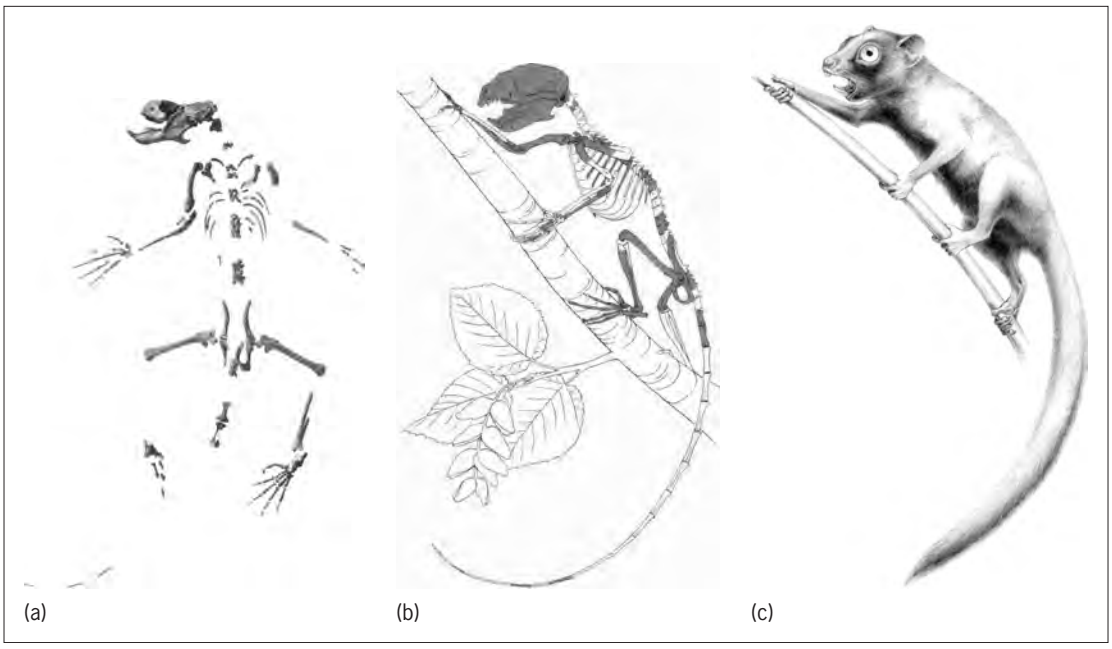

Fig. 2. ***Carpolestes simpsoni*, a 55.5-million-year-old plesiadapiform from Wyoming. (*a*) Skeleton as recovered, laid out in approximate anatomical position. (*b*) Reconstruction in life stance on tree branch, elements in dark gray are known. (*c*) External restoration. (*Courtesy of and © Jon Bloch and Doug Boyer, artwork by D. Boyer*)**

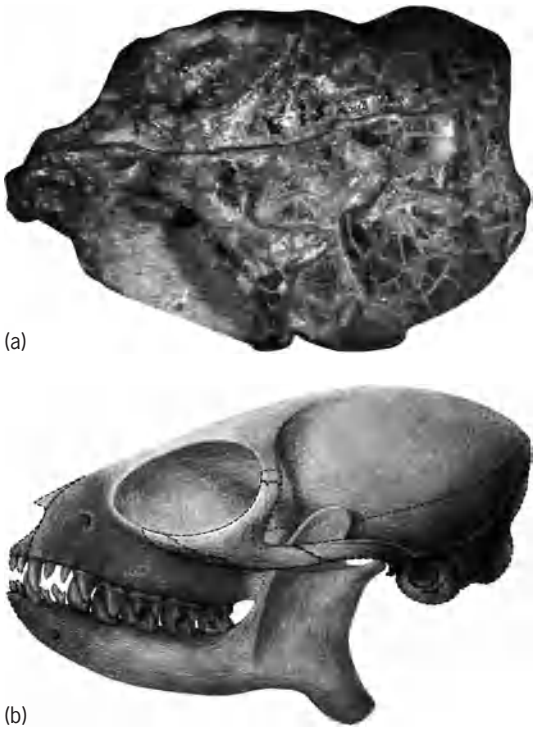

Fig. 3. Cranium of the omomyoid tarsiiform *Teilhardina asiatica*, from deposits dated ca 55 Ma in China. (*a*) Dorsal (top) view as found. (*b*) Left lateral reconstruction. (*Courtesy of and © Xijun Ni*)

(monkeys, apes, and humans). In many respects (such as lack of toothcomb, dry nasal area, molecular biology, and placentation), it is related to the anthropoids, but dentally it is usually thought primitive, although some authors have found similarities to some strepsirhines. The extinct Eocene omomyoids are close to tarsiers skeletally; they are often considered broadly ancestral to the anthropoids for that reason and also because some have monkeylike front teeth. Tarsierlike fossils have been recovered in Egypt and China from deposits of early Oligocene and middle Eocene age, respectively (30 and 45 Ma); in 2006, new material of the Chinese form supported a close link to living tarsiers. Another Chinese Eocene form, *Eosimias*, has been claimed to be an early anthropoid (higher primate), but some researchers argue that is also a tarsier relative, if more distant than the two preceding forms. Living tarsiers eat insects and small vertebrates, and are vertical clingers and leapers; their nocturnal habit is shown by their relatively huge eyes. They appear to live mainly in pairs, but they may form larger associations in some regions.

Higher primates. The anthropoids include three main groups of living animals and their extinct relatives; their divergence from a possibly omomyoid stock probably took place some 50 Ma. The monkeys of the New World and those of the Old World are of similar evolutionary grade, but the latter share a much more recent common ancestry with apes and humans, with which they are grouped in the infraorder Catarrhini.

The earliest known anthropoids may be represented by a few teeth (of *Algeripithecus*) from Algeria about 45 million years old, but better-preserved forms appear in the Fayum region of Egypt between 37 and 30 Ma following new finds and a revision of the geological dating in late 2005 and early 2006. Although now a desert, at the time this was a lush tropical forest through which meandered the sluggish proto-Nile river. Three main groups of early higher primates are represented: the oligopithecids, the parapithecids, and the propliopithecids. The former two appear to be only distantly related to any of the living monkeys or apes, although they may be termed monkeys in the broad sense, as their adaptations are similar to those of modern monkeys. The propliopithecids (species of *Propliopithecus*, including *Aegyptopithecus*) may be close to the common ancestry of later catarrhines (Old World anthropoids). These arboreal animals were the size of small cats, with apelike teeth, small brain, and limbs similar to those of the acrobatic South American atelines. Representatives of modern lineages begin to occur in the fossil record by about 28–20 Ma in both hemispheres. Most scholars subscribe to a colonization of South America by Fayum-age African anthropoids that crossed a then-narrower South Atlantic ocean on rafts of natural vegetation to reach the isolated island continent. *See* MONKEY.

New World primates. The platyrrhine or ateloid monkeys of South and Central America are divided into two families, of which the contents vary according to author. Here, the small marmosets and the common squirrel and capuchin monkeys are grouped into the family Cebidae, while the generally large-bodied spider-howler and saki-uakari groups are linked to the smaller titis and (probably) owl monkeys in the Atelidae. These two families differ in the relative robustness of their jaws and reduction of last molars. All living (and well-known extinct)

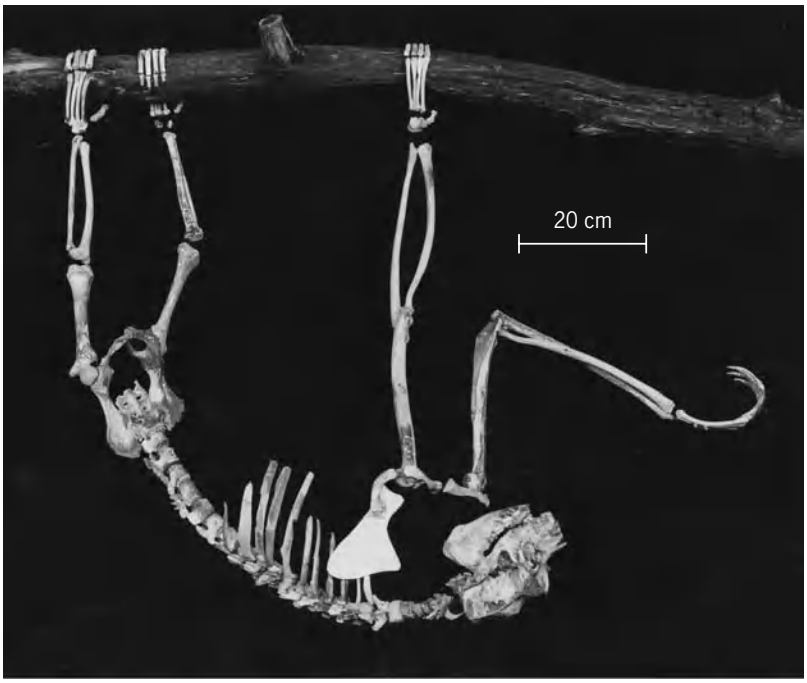

Fig. 4. Anjohibe specimen of *Palaeopropithecus ingens*, mounted in probable habitat posture. (*Photo courtesy of American Geological Institute*)

forms are arboreal and occupy forested areas between Argentina and southern Mexico. A distinctive evolutionary pattern observed in this group is the antiquity of the extant lineages as reflected by the close relationships of most of the few known fossils to modern genera.

The earliest fossil platyrrhine, 26-million-year-old *Branisella* from Bolivia, is as yet known only by teeth and jaw fragments. A probable size reduction of its last molars may suggest an evolutionary relationship to the Cebidae. It has also been suggested that *Branisella* might have lived in a savannah-like environment and been adapted to life at least partly on the ground, as opposed to all later platyrrhines. Two slightly later (about 20 Ma) species from Patagonia and Chile are also probable cebids, close to the ancestry of the living *Saimiri*, while a new species described early in 2006 is younger (16.4 Ma) and shows precocious brain size enlargement in the cebine lineage. Several forms dated 18–17 Ma represent early atelids, close to the living pitheciines (the saki-uakari and titi-owl monkey groups).

The largest number of fossil platyrrhines comes from the La Venta beds of Colombia, dated about 13 Ma. These include one or two species in each subfamily of Cebidae (Callitrichinae and Cebinae), one relative each of owl and howler monkeys, and four early members of the saki-uakari group. The last 100,000 years saw another flowering of extinct platyrrhine lineages. A cave site in eastern Brazil has yielded partial skeletons of "giant" relatives of the howler and spider monkeys, while several localities in the Caribbean produced controversial fossils perhaps related to howlers, saki-uakaris, and squirrel monkeys (or possibly representing a distinct lineage whose members came to resemble those other groups). *See* MONKEY.

Old World monkeys. The living Cercopithecidae are divided into two subfamilies, Colobinae and Cercopithecinae. The oldest cercopithecids are found in Africa, with a few fossil forms such as *Victoriapithecus* of 20–15 Ma probably predating the divergence of the modern subfamilies. The cercopithecines include a wide variety of forms, all of which share cheek pouches for temporary food storage and usually large incisors reflecting a fruit diet; colobines, by contrast, are more restricted in morphology, range, and behavior pattern, and all are leaf eaters with a complicated digestive tract to facilitate the low-nutrition diet. All cercopithecids possess ischial callosities (rump-sitting pads), which have been lost to a greater or lesser extent in apes and humans.

The earliest members of the two living subfamilies also are mainly African. One colobine jaw is known by 9 Ma, and from 7 to 5 Ma species of both cercopithecines and colobines become more abundant. Large collections of Old World monkey fossils have been recovered from East and South African sites (often in association with early human remains) in the 4–1.5 Ma interval. Cercopithecines are represented first by *Parapapio*, which was a semiterrestrial animal probably closer to the common ancestor of later forms, and then by several extinct species of the highly terrestrial living *Theropithecus*. At least three types of large colobine monkeys were also common, ranging from arboreal to terrestrial in adaptation. Many of these species were significantly larger than their living relatives: The biggest colobine may have weighed close to 50 kg (110 lb) [**Fig. 5**] while a *Theropithecus* dated about 1 Ma has been estimated nearly twice that large. Some of these large and slow-moving species may have been hunted and perhaps driven to extinction by early humans.

Cercopithecids entered Eurasia from Africa. *Mesopithecus pentelicus*, an 11–8-million-year-old colobine known in a geographical range from Germany through Afghanistan, is the best-represented Eurasian fossil monkey, with dozens of individuals recovered from sites in Greece and Bulgaria. It was semiterrestrial, perhaps like the living hanuman langur *Semnopithecus entellus*, sleeping in trees and feeding on the ground along watercourses. A possible descendant is *Dolichopithecus*, which lived in Europe from about 4.5 to 2.5 Ma and perhaps in Japan, Mongolia, and Siberia as well. It was even more terrestrially adapted, more so than any living colobine. The living macaques (*Macaca*) are widespread across eastern Asia and in North Africa, and their fossil record adds to that large range. Scattered specimens are known from North Africa after 7 Ma, and populations have been recovered across Europe from 5.5 Ma to about 100,000 years ago. Large and highly terrestrial relatives, such as *Paradolichopithecus*, inhabited Europe, central Asia, and eastern Asia from about 3 to 1.5 Ma.

Hominoids. The most humanlike of all primates are the apes, which form a group distinguished by generally large body size, relatively large brain, lack of an external tail, and advanced placentation pattern. Living forms include the lesser apes, or gibbons (*Hylobates*), placed in their own family Hylobatidae, and the several great apes: orangutan (*Pongo*), chimpanzee (*Pan*), and gorilla (*Gorilla*). The great apes

Fig. 5. Partly reconstructed skeleton (casts) of *Paracolobus chemeroni*, 3.2-million-year-old colobine monkey from Kenya; on exhibit at American Museum of Natural History, New York. (*Courtesy of and © Eric Delson*)

and humans, along with some extinct relatives, are grouped as the Hominidae by some authors (as here), while others place only humans in the Hominidae and class all great apes in the Pongidae.

One of the earliest probable members of Hominoidea is *Proconsul*, of the East African Miocene, 23–14 million years old; a few teeth of a similar form date to 26 Ma. *Proconsul* is well known by most of its skeleton (**Fig. 6**). Several species ranged in size from small chimpanzee to small gorilla, with a somewhat chimplike skull, large projecting canine teeth, and limb bones seemingly adapted to quadrupedal running. However, *Proconsul* has few of the defining features of the ape group. It lacks the mobile shoulder characteristic of all living species, and there is some controversy over its tail: some workers believe that tail bones have been found but not always recognized as such. Thus, several researchers have removed *Proconsul* from the Hominoidea and placed it and several relatives as an advanced archaic catarrhine. For the present, *Proconsul* is retained as a hominoid belonging to a distinct archaic family of its own. Two other groups of roughly contemporaneous species (the Eurasian Pliopithecidae and the African "*Dendropithecus-group*") are clearly more primitive than *Proconsul*, although they have at times wrongly been included in Hominoidea, usually as purported relatives of the gibbons. The oldest fossil gibbons date only to about 1 Ma; presumably, their ancestors lived after *Proconsul* and entered Eurasia about the same times as did early Hominidae.

Several fossil apes from Africa dating to 20–14 Ma are often placed in the Hominidae because they appear to share "advanced" dental features with the modern great apes. But most of these species have recently been questioned as hominids because they lack apelike features of the postcranial skeleton (in the elbow, shoulder, hip, and knee joints, especially). Surprisingly, the oldest of these, *Morotopithecus*, may be the most modern, although only fragmentary postcranial elements are known. The youngest, *Kenyapithecus*, is similar to the first Eurasian hominoid, *Griphopithecus*, found in Turkey, Germany, Slovakia, and Austria between 16 and 13 Ma. None of these fossils are specially related to any living great ape, and they are often placed in the hominid subfamily Kenyapithecinae or even outside Hominidae.

Although interpretations vary, there appear to be three groups of Eurasian hominids between 13 and 7 Ma. *Dryopithecus* is characteristic of the Dryopithecinae, which may include the common ancestors of all later great apes (and humans). This animal has been known since the 1850s, but only in the 1990s were partial crania and skeletons described. These show several similarities to living apes in general but not to any specific forms.

Spread of modern ape ancestors. Most scientists today agree that, of the great apes, the orangutan is evolutionarily farthest from humans. This view is based mainly on the results of numerous studies of proteins and DNA sequences, which reveal great similarity among the genes of gorillas, chimpanzees, and humans, with orangutans, gibbons, and monkeys successively farther from that close-knit group. As a result, orangutans and their extinct relatives are here placed in the subfamily Ponginae, while African apes, humans, and their relatives are included in Homininae. The living orangutans (*Pongo*) inhabit a small range in the deep forests of Borneo and Sumatra; less than 1 Ma, the same species was also present as far north as southern China. *See* APES.

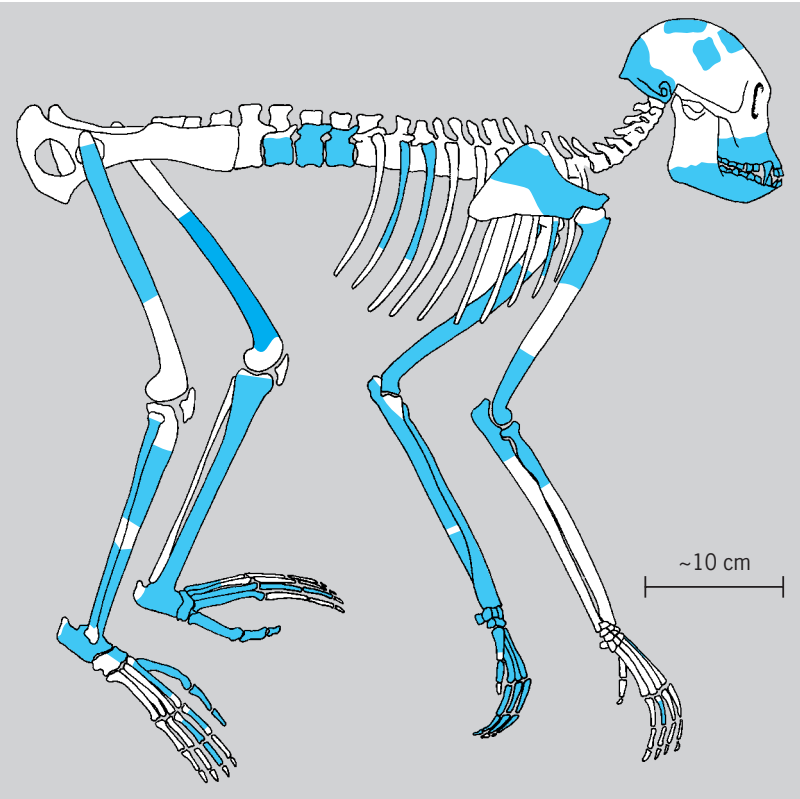

Fig. 6. Reconstructed skeleton of Proconsul africanus, an early hominoid (23-14 Ma) from East Africa. Known parts are shaded. (*After R. L. Ciochon and R. S. Corruccini, eds., New Interpretations of Ape and Human Evolution, Plenum Press, 1983*)

The orangutan lineage is, however, the oldest well-documented one among all catarrhines. Fossils from Pakistan and India known as *Sivapithecus* show facial, palatal, and dental architectures clearly specialized in the orang direction as far back as 13–12 Ma. Although slightly younger in time, *Ankarapithecus* from Turkey at approximately 10 Ma appears intermediate between *Sivapithecus* and *Dryopithecus* in facial morphology. These three genera seem to form a transformation series representing the approximate manner in which orang ancestry evolved from dryopithecines, and both *Ankarapithecus* and *Sivapithecus* are usually included in Ponginae. But whereas the living great apes (and human ancestors) share a complex of arm-bone features related to suspensory behavior and forelimb flexibility, the upper arm of *Sivapithecus* lacks at least one part of this complex. Thus, it is possible either that this lack denies *Sivapithecus* close relationship with orangs, or that Asian and African great apes independently evolved these features long thought to document their shared ancestry. Alternatively, the lack may merely represent a secondary specialization of *Sivapithecus* which happens to be similar to what is seen in more archaic apes such as the kenyapithecines.

This last option makes the most sense, as *Sivapithecus* in other ways appears to be somewhat more adapted to quadrupedal life on the ground than is the living orangutan.

Two larger species which probably belong to the Ponginae are placed in the genus *Gigantopithecus*: One dates to about 9–6 Ma from India and Pakistan, the other lived about 1.5–0.5 Ma in China and perhaps Vietnam. Hundreds of specimens, mostly isolated teeth, are known from China, and these document a species which was probably the largest primate that ever lived (perhaps weighing 200–400 kg or 440–880 lb, although teeth are not the best body part from which to estimate weight). *Gigantopithecus* was once thought to be near the ancestry of humans, but today it is usually considered a collateral descendant of *Sivapithecus*.

The pongines probably evolved in Asia from an arboreal dryopithecine (or even kenyapithecine) ancestry which expanded into moist forested environments. Such a habitat would have provided an abundance of fruit plus gritty and tough food objects (like fibrous and thick-skinned fruits or nuts), to which this group's dentition appears adapted. It was at first thought that these creatures might have spent significantly more time walking and feeding on the ground than modern or earlier apes, but partial limb bones suggest overall similarity of structure, and thus probably of behavioral function, to living arboreal and partly ground-living apes. They share with orangutans and some early humans a complex of dental-related features, including a thick molar enamel covering (to prolong tooth life with heavy wear), large molars in relation to estimated body size, strongly buttressed jaws (to take up stresses of heavy chewing), and some reorientation of the front teeth. A major factor in human dental evolution was facial shortening combined with reduction of the canine teeth and their transformation into incisorlike elements, but no Miocene fossils yet known show this feature clearly.

The origin of the Homininae is more problematic. The fossil ape *Graecopithecus* (also termed *Ouranopithecus*) is known from several Greek localities estimated to date between 9.6 and 8.7 Ma. Well-preserved facial material of this animal and of *Dryopithecus* recovered or reanalyzed in the 1990s has led different workers to suggest that one or both forms may lie near the split between Ponginae and Homininae or already on the hominine lineage, effectively close to the common ancestor of African apes and humans. This would agree well with the date of about 13–12 Ma for the earliest pongines on the assumption that both branches of the split would have developed their derived features contemporaneously. *Graecopithecus* shows more evidence of the downwardly bent face typical of African apes (a pattern termed klinorhynchy) and also very thick molar enamel (as expected in an early hominine), while *Dryopithecus* has thin enamel. Some of the differences between them may just reflect the far larger size of *Graecopithecus*, but at present this genus appears more derived in the direction of later hominines. At about the same time in Africa, the only known ape fossil is a single upper jaw which was named *Samburupithecus* in 1997 and which may also represent an early, gorillalike member of Homininae. These fossils imply that the hominines may have evolved in Eurasia and then returned to Africa after about 10 Ma. Previous workers often thought that the hominine lineage could be traced purely within Africa, but that hypothesis now appears less likely. The ancestry of the African apes is still a mystery: until mid-2005 no fossils had been found which clearly represent their lineage before or after separation from humans. However, a few teeth have now been reported from a Kenyan archeological site which appear to represent a form of chimpanzee, but so far these say little about the history of African ape adaptations.

Human evolution. During the last decade, three new African genera have been described that have been claimed to throw new light on the origin of the Hominini, the lineage including living humans but postdating the last common ancestor of humans and apes. Potentially the oldest of these is *Sahelanthropus*, found in 2002 in Chad, and dated in the 7–6 Ma range. Slightly younger is *Orrorin*, from Kenya, at about 6 Ma; and the youngest is *Ardipithecus*, reported from sites in Ethiopia dating between 5.8 and 4.3 Ma. Each of these is known from essentially different skeletal elements, and each has been disputed in one way or another, reflecting the fact that we do not as yet have an adequate conception of what the earliest hominins should be expected to look like. All, however, have been claimed, on the basis of tenuous evidence, to have been upright bipeds.

Several species of the genus *Australopithecus* are known from sites dating between 4.2 and 2.3 Ma in South Africa, Chad, and the Rift Valley from Tanzania through Ethiopia. They have crania with internal volumes ("brain size") of 400–500 ml, in the range of living gorillas; body weight, however, is far smaller, about 30 kg for females, up to 60 kg for males. The jaw musculature was small enough not to require strong crests on the outer surface of the skull. The dentition had moderately large incisors, relatively small and partly incisivized canines, thick-enameled molars, and a broadly "humanlike" arcade shape. The postcranial skeleton was clearly adapted to upright bipedalism (also indicated by footprint trails), although it differed in detail from modern humans, with relatively short legs. The diet of these forms was probably omnivorous, including fruits, seeds, and perhaps scavenged animal protein; there is no definitive evidence for toolmaking. *See* AUSTRALOPITHECINE; FOSSIL HUMANS.

The "robust" australopiths, genus *Paranthropus*, were generally younger (about 2.6–1.4 Ma) and more derived craniofacially. Their postcanine teeth were greatly enlarged and adapted to crushing and grinding hard food items, while the canines and incisors were strongly reduced except in the earliest species, *P. aethiopicus*. Sagittal and nuchal crests served as anchors for the large temporal musculature covering a vault enclosing an only slightly larger brain than in *Australopithecus*; the postcranial skeleton is poorly known, but body structure and size were probably not much different from that of its predecessor.

The earliest toolmakers probably belonged to species included in the living genus *Homo*, first appearing 2.5–2.2 Ma in Malawi, Kenya, and Ethiopia, soon after the oldest stone tools. There is some debate over the number of early *Homo* species, but it seems likely that at least three were contemporaneous if not actual neighbors in eastern Africa around 2 Ma. Soon afterward, *Homo erectus* (or according to some authors, a closely related species named *H. ergaster*) emigrated from Africa to eastern Eurasia where *H. erectus* persisted until at least 300 Ka (thousand years ago). Continental Europe may not have been colonized until about 1 Ma; almost all younger populations there appear to be part of a clade that included the Neanderthals, known from 200 to 30 Ka between Spain and central Asia. African populations of the last half million years appear to have evolved less directly into anatomically modern humans (*Homo sapiens*), some of whom dispersed via southwestern Asia to colonize the rest of the Earth after 150 Ka. *See* FOSSIL HUMANS; MAMMALIA.

Eric Delson; Ian Tattersall

Bibliography. E. Delson et al. (eds.), *Encyclopedia of Human Evolution and Prehistory*, 2d ed., 2000; J. G. Fleagle, *Primate Adaptation and Evolution*, 2d ed., 1998; W. C. Hartwig, *The Primate Fossil Record*, 2002; S. V. Nelson, *The Extinction of Sivapithecus*, 2003; J. H. Schwartz and I. Tattersall, *The Human Fossil Record*, vols. 1–4, 2001–2005; F. S. Szalay and E. Delson, *Evolutionary History of the Primates*, 1979.

Fossil seeds and fruits

Seeds, ovules containing a fertilized egg and ready to be shed from the plant, are reproductive organs characteristic of both gymnospermous and angiospermous plants. In angiosperms (Magnoliophyta) an additional structure, the matured ovary, encloses one or more seeds to form a fruit. *See* MAGNOLIOPHYTA.

Seeds and fruits are less commonly found as fossils than are vegetative remains. They may be preserved structurally as casts, or as compressions which are sometimes found with leaf compressions. Seeds and fruits often occur in lignites. *See* LIGNITE.

The oldest known seed plants are of Mississippian age. Carboniferous seed plants include the extinct Cordaitales, probable conifer ancestors, and Pteriodospermae, seed plants with fernlike foliage. Both groups had similar seeds, basically like those of modern cycads. No embryos have been found in Paleozoic seeds. *See* CORDAITALES; PTERIDOSPERMS.

During the Mesozoic Era, all major modern groups of seed plants were represented, along with members of the declining cordaitalean and pteridospermous stocks. Among the most completely known Mesozoic seeds are those of the cycadeoids, extinct cycad relatives. Their silicified, beehive-shaped trunks may include shoots bearing numerous small seeds between thick scales. The Mesozoic Caytoniales, whose small seeds were borne in fleshy enclosures, were thought at one time to be ancestral angiosperms. Now they are recognized as gymnospermous forms, probably related to the pteridosperms. *See* CAYTONIALES; CYCADEOIDALES.

Angiosperm fruits are rare in Lower Cretaceous beds; Upper Cretaceous fruits are known from northern Africa, Long Island, New York, and elsewhere. Tertiary fruits and seeds have been found in numbers in the United States in the Brandon lignite of Vermont (see **illus.**) and in the Clarno Formation of central Oregon. Silicified cones of araucarian conifers are known from Patagonia. The best-known European Tertiary fruits and seeds are from the brown coals of

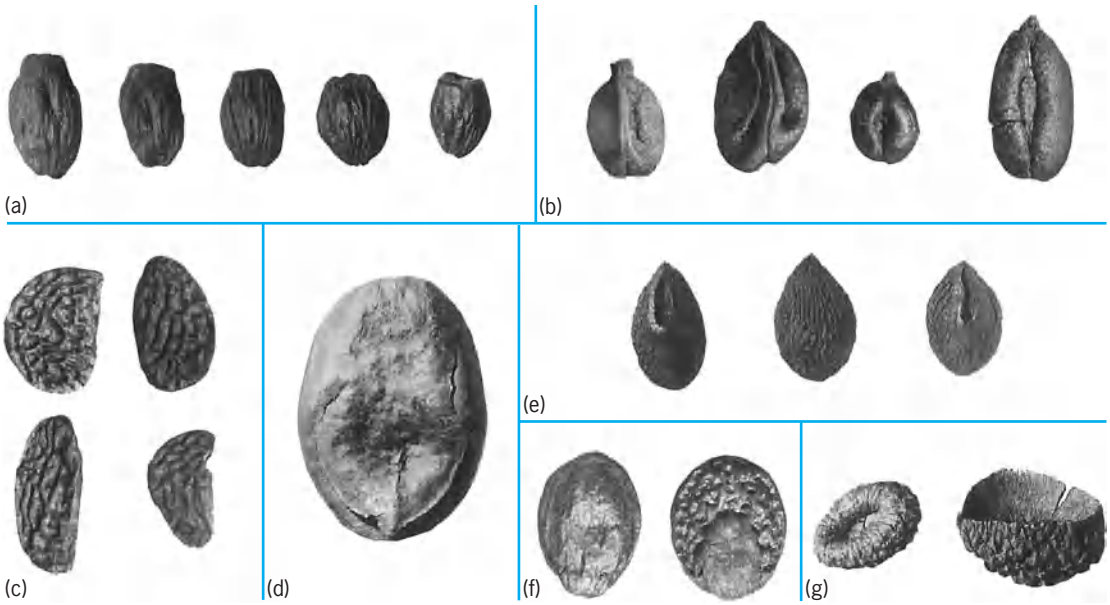

Early Tertiary fruits and seeds which were found in Brandon lignite of Vermont. (*a*) Endocarps of *Symplocos*. (*b*) Seeds of grape (*Vitis*). (*c*) Seeds of *Rubus*. (*d*) Endocarp of extinct species of *Nyssa*. (*e*) Seeds of fossil genus of family Rutaceae, which are related to *Phellodendron*. (*f*) Endocarps of a species of *Alangium*. (*g*) Acorn cups, *Quercus* sp. (*Courtesy of Elso S. Barghoorn*)

Germany and the Eocene London Clay Formation of England.

Important paleobotanical findings resulting from the study of fossil seeds and fruits include the knowledge obtained of the independent evolution of the seed habit in unrelated groups; the discovery that much Carboniferous fernlike foliage was borne on seed plants rather than on ferns; and the discovery that *Glossopteris*, an important plant in widespread Permian floras of the Southern Hemisphere, was a seed plant. Pyritized fruits from the London Clay Formation reveal the presence of many extinct genera along with modern genera in early Tertiary time. Morphological changes in herbaceous angiosperm seeds from sequences of Tertiary beds furnish data on rates of evolution in plants. Because plant classification is based primarily upon reproductive structures, fossil seeds and fruits provide highly reliable evidence for identification and interpretation of fossil plants. *See* PALEOBOTANY. Richard A. Scott

Bibliography. H. N. Andrews, *Studies in Paleobotany*, 1961; E. M. Reid and M. E. J. Chandler, *The London Clay Flora*, 1933; R. A. Scott, Fossil fruits and seeds from the Eocene Clarno Formation of Oregon, *Palaeontographica,* 96(B):66–97, 1954; A. C. Shukla and S. P. Misra, *Essentials of Paleobotany*, 2d ed., 1982; T. N. Taylor, *Paleobotany*: *An Introduction*, 1981.

Foucault pendulum

A pendulum or swinging weight, supported by a long wire, by which J. B. L. Foucault demonstrated in 1851 the rotation of Earth on its axis. Foucault used a 62-lb (28-kg) iron ball suspended on about a 200-ft (60-m) wire in the Pantheon in Paris. The upper support of the wire restrains the wire only in the vertical direction. The bob is set swinging along a meridian in pure translation (no lateral or circular motion). In the Northern Hemisphere the plane of swing appears to turn clockwise; in the Southern Hemisphere it appears to turn counterclockwise, the rate being 15 degrees times the sine of the local latitude per sidereal hour. Thus, at the Equator the plane of swing is carried around by Earth and the pendulum shows no apparent rotation; at either pole the plane of swing remains fixed in space while Earth completes one rotation each sidereal day. *See* DAY; INERTIAL GUIDANCE SYSTEM; PENDULUM; SCHULER PENDULUM. Frank H. Rockett

Bibliography. G. O. Abell, D. Morrison, and S. C. Wolff, *Exploration of the Universe*, 6th ed., 1991; C. Kittel, W. D. Knight, and M. A. Ruderman, *Mechanics*, Berkeley Physics Course, vol. 1, 2d ed., 1973; K. R. Symon, *Mechanics*, 3d ed., 1971.

Foundations

Structures or other constructed works are supported on the earth by foundations. The word "foundation" may mean the earth itself, something placed in or on the earth to provide support, or a combination of the earth and the elements placed on it. The foundation for a multistory office building could be a combination of concrete footings and the soil or rock on which the footings are supported. The foundation for an earth-fill dam would be the natural soil or rock on which the dam is placed. Concrete footings or piles and pile caps are often referred to as foundations without including the soil or rock on which or in which they are placed. The installed elements and the natural soil or rock of the earth form a foundation system; the soil and rock provide the ultimate support of the system. Foundations that are installed may be either soil-bearing or rock-bearing. The reactions of the soil or rock to the imposed loads generally determine how well the foundation system functions. In designing the installed portions, the designer must determine the safe pressure which can be used on the soil or rock and the amount of total settlement and differential settlement which the structure can withstand. For a discussion of soil as a part of the foundation system *see* SOIL MECHANICS.

The installed parts of the foundation system may be footings, mat foundations, slab foundations, and caissons or piles, all of which are used to transfer load from a superstructure into the earth. These parts, which transmit load from the superstructure to the earth, are called the substructure (**Fig. 1**).

Footings. Footings or spread foundations are used to spread the loads from columns or walls to the

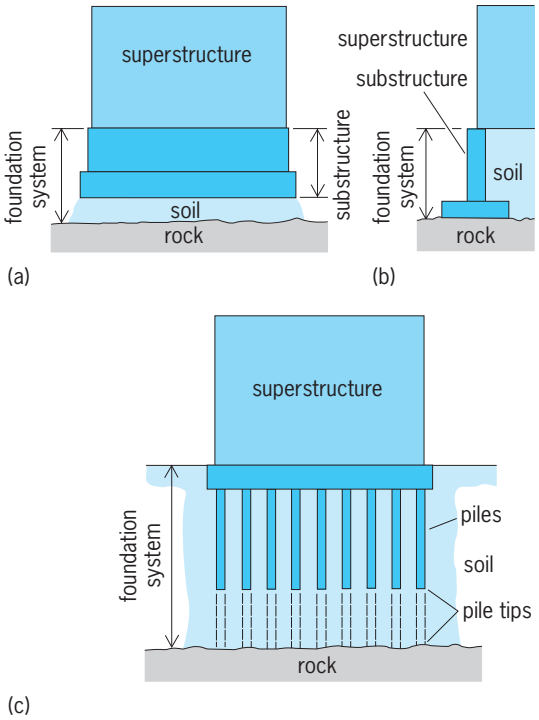

Fig. 1. Examples of foundation systems. (*a*) Structure supported on a foundation bearing on soil. (*b*) Structure supported on a foundation bearing on rock. (*c*) Structure supported by a pile foundation. The piles may be installed so that the pile tips terminate in soil and all support is derived from the soil, or they may be installed so that the tips penetrate to rock.

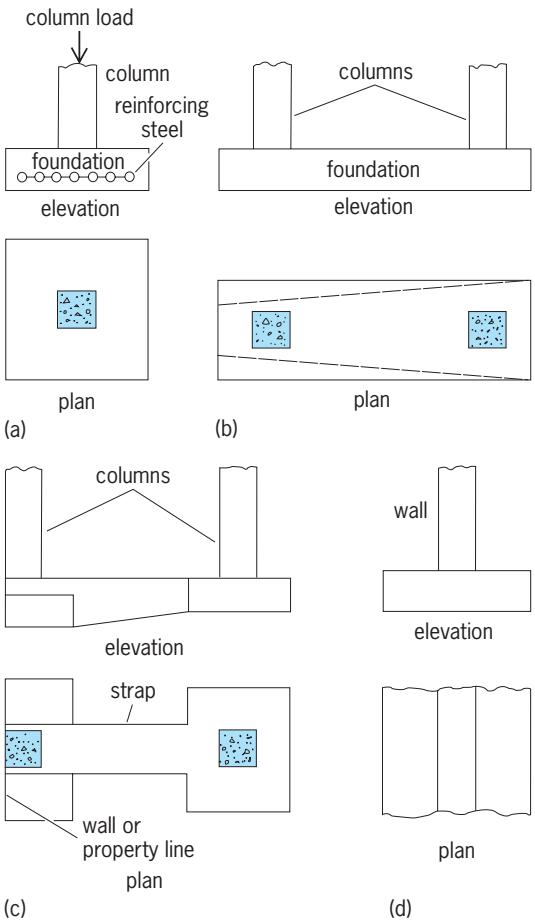

Fig. 2. Types of footings. (*a*) Column footing. (*b*) Combined footing, which can be either rectangular or trapezoidal. (*c*) Cantilevered footing. (*d*) Wall footing.

underlying soil or rock. Normally, footings are constructed of reinforced concrete; however, under some circumstances they may be constructed of plain concrete or masonry. When each footing supports only one column, it is square. Footings supporting two columns are called combined footings and may be either rectangular or trapezoidal. Cantilever footings are used to carry loads from two columns, with one column and one end of the footing placed against a building line or exterior wall. Footings supporting walls are continuous footings (**Fig. 2**).

The sizes of footings are determined by dividing the loads to be imposed at the base of the footings by the allowable bearing pressure which can be imposed on the soil or rock of the earth. Most building codes and textbooks on foundations contain tables listing allowable bearing pressures for various types of soil and rock; however, these tables give only general classifications and descriptions of the soil or rock and must be used with caution. More specific information about the soil or rock is normally obtained by drilling test borings, extracting soil or rock samples, performing laboratory tests on the samples, and making engineering analyses to determine suitable bearing pressures. In addition to bearing pressure, consideration must be given to the amount of settlement which may occur and the capability of the structure to withstand such settlement. If settlement is a problem it may be necessary to use an alternate foundation type rather than footings or to enlarge the footings and decrease the bearing pressure.

Grade beams may be used between exterior column footings to support walls, with the beams transferring the weight of the walls to the column footings. Beams are also used between interior column footings to act as braces or ties or to support interior walls. Retaining walls are thoses walls subjected to horizontal earth pressures due to the retention of earth behind them. The foundations for these walls must have sufficient frictional resistance with the soil or rock on which they rest so that they will not slide when subjected to the horizontal earth pressure. In addition, retaining walls must be designed so they will not overturn (**Fig. 3**). In frost-susceptible areas, footings must be placed below the frost line.

Mat foundations. Mat or raft foundations are large, thick, and usually heavily reinforced concrete mats which transfer loads from a number of columns or columns and walls to the underlying soil or rock. Mats are also combined footings, but are much larger than a footing supporting two columns. They are continuous footings and are designed to transfer a relatively uniform pressure to the underlying soil or rock. Mats are rigid and will act as a bridge over discontinuities in the soil or rock on which they are founded. Mats founded several meters below the ground surface, when combined with external walls, are termed floating foundations (**Fig. 4**). The weight

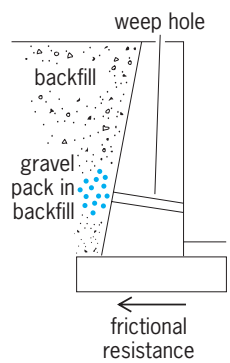

Fig. 3. Retaining wall. Gravel pack behind the wall prevents the weep hole from clogging with soil.

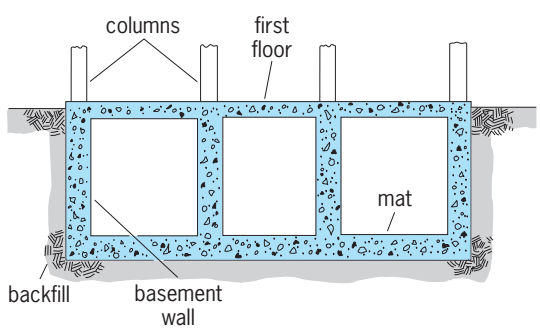

Fig. 4. Mat foundation, sometimes called a floating foundation, generally heavily reinforced in both directions.

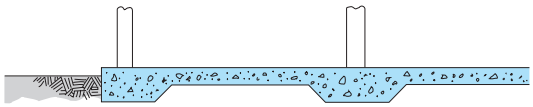

Fig. 5. Slab foundation.

of the soil excavated from the ground surface to the bottom of the mat may be equal to or approach the total weight of the structure. In this case, little or no new load is applied to the underlying supporting soil, and settlements of a structure may be minimal after construction.

Slab foundations. Slab foundations are used for light structures wherein the columns and walls are supported directly on the floor slab. The floor slab is thickened and more heavily reinforced at the places where the column and wall loads are imposed (**Fig. 5**).

Special problems. Ground water is a major problem in connection with the design and installation of foundations where a substructure is to be placed below the ground-water level. Well points, pumping from deep wells, or pumping from sumps are methods used to dewater construction sites during foundation installation. Other methods which are less often used are freezing, removal of water by electroosmosis, and the installation of cutoff walls made of piling or grout around the periphery from within the excavation. If dewatering operations are performed in an area surrounded by existing structures, precautions must be taken to protect them, as the lowering of the ground-water may cause the soil on which they are supported to subside.

If a basement is partially or totally below the ground-water level its walls must be designed to withstand the hydrostatic pressures of the water on the outside in addition to the pressure from the soil backfill. An alternate procedure is to install a permanent system to remove water outside the walls. Some substructures below ground-water level may at times be subjected to hydrostatic uplift forces which are greater than the downward forces imposed by the structures. In these cases, provisions must be made to anchor the structures to prevent them from floating upward.

Ground water also causes problems by infiltrating through basement walls, slabs, and joints into the basement itself. This can be prevented or reduced by providing an external permanent drainage syatem that carries water away from the basement, by encasing the walls and slabs in an inpermeable plastic membrane, or by coating the external walls with an asphaltic mastic to lower their permeability (**Fig. 6**). Combinations of the foregoing are also used. Retaining walls and abutments often have weep holes in the lower sections through which water accumulating behind the walls or abutments can escape. The water pressure behind the walls is relieved as the water flows through the walls into an open external drainage system.

Foundations placed on expansive soils are often subjected to distressing movements unless special precautions are taken. Expansive soils are those which swell and contract excessively with varying amounts of moisture. Problems can be overcome by installing foundations below the zone of significant change in moisture content and backfilling with non-expansive materials, by altering the soil with an admixture such as lime or cement so that volume changes do not occur, or by providing flexibility in the structure to accommodate movements.

Underpinning of foundations is often necessary, and it may either remedial or precautionary. Remedial underpinning is used to correct defects in existing foundations which may have settled excessively. If the structure is to be saved or returned to its original state, additional foundation support must be provided. Precautionary underpinning is used when new structures are to be installed adjacent to or beneath existing structures, as in the construction of city subways. Underpinning of foundations is a specialized construction technique. The work is generally performed in a confined space, such as the basement of a building, or in small pits excavated outside a building area. It is necessary to provide support for the loads of the existing structure while new foundations are installed. The new foundations may be footings which are placed deeper in the ground than the original foundations, or they may be piles or caissons.

Underpinning of a wall footing may be performed by excavating pits adjacent to and beneath existing foundations. The pits are small, some 3 ft (0.9 m) wide by 4 ft (1.2 m) long. Horizontal sheeting is placed in the pits as excavation proceeds to prevent caving of the walls and undermining of the structure being underpinned. When the new bearing stratum is reached, forms are placed in the pit, and concrete is poured from the new bearing stratum up to within 3 in. (76 mm) of the bottom of the old footing. After

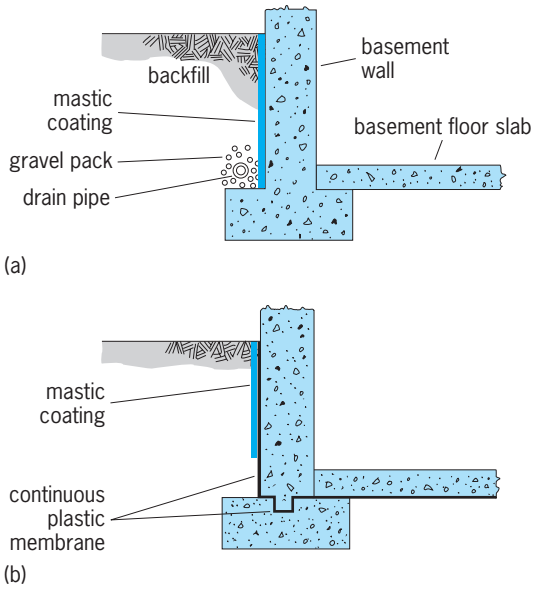

Fig. 6. Methods of reducing water infiltration into basements. (*a*) External drain pipe. (*b*) Walls, floor slab, and joints encased in a plastic membrane.

the new concrete has hardened, the 3-in. (76-mm) space is packed by hand with a mixture of sand, cement, and a small amount of water. Called grout, the mixture is packed very tightly into the space between the top of the new footing and the underside of the old footing. The pit underpinning process is repeated throughout the entire length of the wall footing. The resulting new foundation may be a continuous wall or intermittent piers. *See* BRIDGE; BUILDINGS; CAISSON FOUNDATION; DAM; PILE FOUNDATION; RETAINING WALL. Gardner M. Reynolds

Bibliography. R. Ahlvin and V. A. Smoots, *Construction Guide for Soils and Foundations*, 2d ed., 1988; J. E. Bowles, *Foundation Analysis and Design*, 5th ed., 1996; H. Y. Fang, *Foundation Engineering Handbook*, 2d ed., 1991; S. Hansbo, *Foundation Engineering*, 1994; L. ZeeVaert, *Foundation Engineering for Difficult Subsoil Conditions*, 2d ed., 1982.

Four-bar linkage

A basic linkage mechanism used in machinery and mechanical equipment. The term has been applied to three types of linkages: plane, spherical, and skew.

The plane four-bar linkage (**Fig. 1**) consists of four pin-connected links forming a closed loop, in which all pin axes are parallel. The spherical four-bar linkage (**Fig. 2**) consists of four pin-connected links forming a closed loop, in which all pin axes intersect at one point. The skew four-bar linkage (**Fig. 3**) consists of four jointed links forming a closed loop, in which crank 2 and link 4 are pin-connected to ground 1 and the axes of the pins are generally nonparallel and nonintersecting; coupler 3 is connected to crank 2 and link 4 by ball joints.

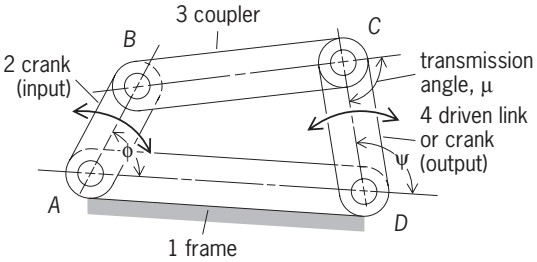

Fig. 1. Plane four-bar linkage.

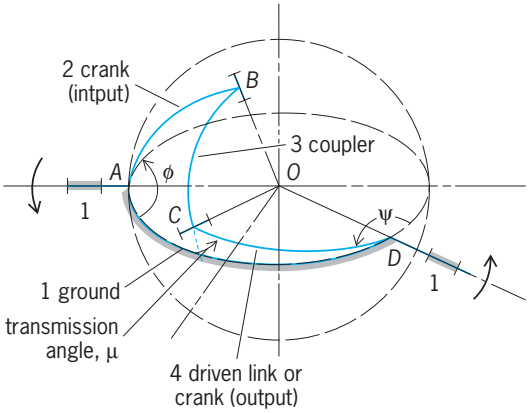

Fig. 2. Spherical four-bar linkage.

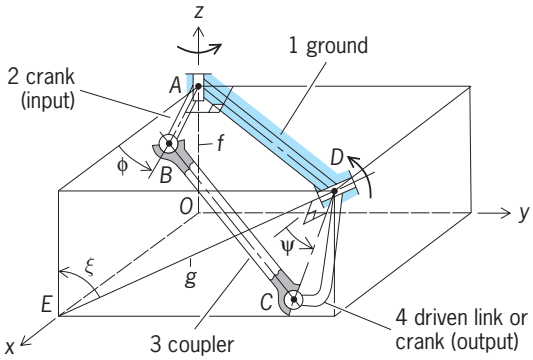

Fig. 3. Skew four-bar linkage. *OA* = *f*; *ED* = *g*; *OE* = common perpendicular between axes of pin joints at *A* and *D*.

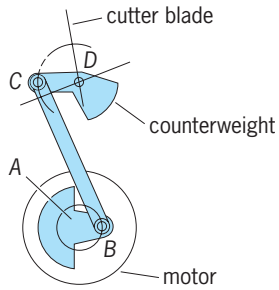

Fig. 4. Plane crank-and-rocker as used in an electric razor.

Mechanical function. Four-bar linkages are most frequently used to convert a uniform, continuous rotation (the motion of crank 2) into a nonuniform rotation or oscillation (the motion of link 4). Depending on whether the axes of the connected shafts (attached to links 2 and 4) are parallel, intersecting, or skew, a plane, spherical, or skew four-bar linkage is used. In instrument applications the primary function of the linkage is the conversion of motion, while in power applications both motion conversion and power transmission are fundamental. Four-bar linkages also are used to impart a prescribed motion to the coupler, as well as for other purposes.

Types of motions. Each of the above linkages can be proportioned for three types of motion, or linkage types: crank-and-rocker, drag, and double-rocker.

Crank-and-rocker linkages. In this motion the crank (link 2) is capable of unlimited rotation, while the output link (link 4) oscillates or rocks through a fraction of one turn (usually less than 90°). This is the most common form of the plane and the skew four-bar linkage, and is used in machinery and appliances of all types, for example, in an electric razor (**Fig. 4**). The most common spherical crank-and-rocker motion is known as the wobble plate. In wobble-plate motion the angle between the pin axes on the crank is less than 90°, while for each of the other links the angle between the pin axes is equal to 90°. The wobble motion is that of the coupler. It has been used in axial engines and elsewhere, for example, in dough kneading (**Fig. 5**), in which the dough is stirred by

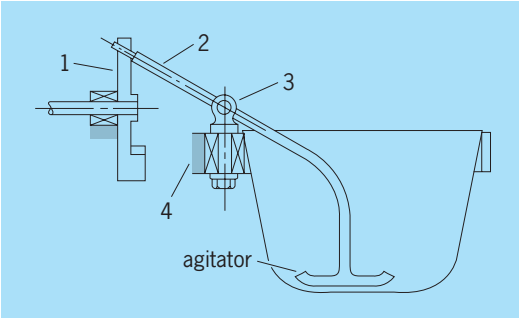

Fig. 5. Dough-kneading mechanism.

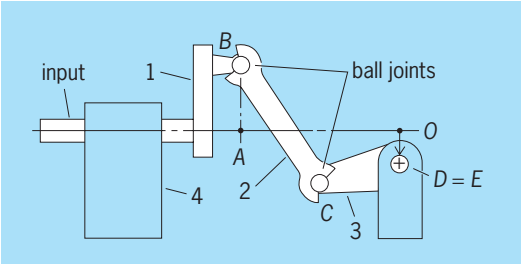

Fig. 6. Skew crank-and-rocker linkage.

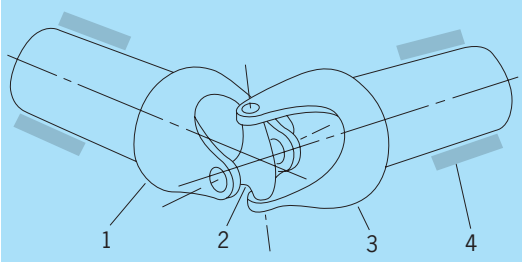

Fig. 7. Hooke joint, a universal joint with spherical drag linkage.

an agitator integral with the coupler. A skew four-bar linkage of crank-and-rocker proportions is shown in **Fig. 6**; the angle between the axes of crank 2 and rocker 4 is a right angle. Such motions have been used in textile machinery, tobacco-processing machinery, and generally for "going around corners."

Drag linkages. In drag linkages the motions of cranks 2 and 4 are both capable of unlimited rotations. The plane drag linkage has been used for quick-return motions, but care must be taken to avoid mechanical interference of the cranks with the coupler. The skew drag linkage is rarely used. The most common drag linkage is the spherical drag linkage. One such linkage is the Hooke-type universal joint, or Hooke joint, in which the angle between the pin axes of each moving link is a right angle (**Fig. 7**). The Hooke joint is used to connect intersecting shafts, for example, in automotive drive lines. The angular-velocity ratio between the connected shafts depends on the shaft angle (which is usually less than or equal to 30°) and is not constant. *See* UNIVERSAL JOINT.

Double-rocker linkages. In such linkages (**Fig. 8**) neither crank 2 nor 4 is capable of complete rotations. Such motions occur in hand tools and mechanical equipment in which only limited rotations are required, such as the oscillating motion in certain household fans.

Grashof's inequality. In the case of plane four-bar linkages the proportions which determine the type of motion are governed by Grashof's inequality: the sum of the lengths of the longest and shortest links is less than the sum of the lengths of the two intermediate links. If the inequality is satisfied and either link 2 or link 4 is the shortest link, the linkage is a crank-and-rocker (the shortest link functioning as the crank). If the equality is satisfied and the frame is the shortest link, the linkage is a drag linkage. In all other cases the linkage is a double-rocker linkage or a folding linkage. A folding linkage occurs when Grashof's inequality becomes an equality. Such linkages include those having parallelogram proportions, such as occur in the parallel and antiparallel equal-crank linkages. The latter can be used only, however, if provision is made for a positive drive while passing through the folded position or positions.

For spherical four-bar linkages a simple extension of Grashof's inequality is available, but for skew four-bar linkages the criteria governing motion types are more involved.

Transmission characteristics. For the transfer of power and motion in crank-driven four-bar linkages, the linkages should be proportioned to avoid lockup positions. In such positions a turning effort applied to the crank will not produce a turning effort applied to the output link. In plane and spherical four-bar linkages, a lockup position occurs when the pin axes on the coupler and output link are coplanar. The transmission angle μ (Figs. 1 and 2) is then equal to 0° or 180°. For efficient force transmission in plane and spherical four-bar linkages the maximum deviation Δ of the transmission angle μ from 90° should be minimized (usually $\Delta \leq 30°$, at least in high-speed operation). Analogous but more complicated conditions apply to skew four-bar linkages. The Hooke joint and wobble plate are well proportioned for force transmission, as are the "centric" plane crank-and-rocker linkages for which $(AB)^2 + (AD)^2 = (BC)^2 + (CD)^2$. The determination of crank-and-rocker proportions yielding optimum force transmission as a function of

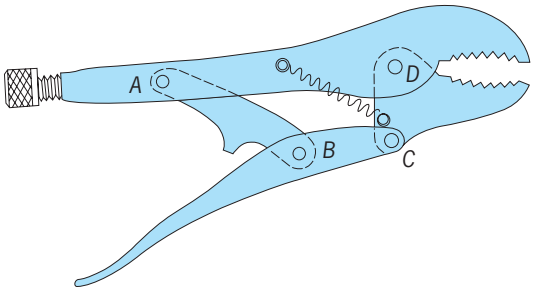

Fig. 8. Linkage-type pliers.

rocker swing angle and corresponding crank rotation has been solved as well.

Angular displacements. Equation the angular displacements of crank 2 and link 4 for the plane four-bar linkage and for a special case of the skew four-bar linkage—the "simply skewed" ($f = g = 0$; Fig. 3).

$$f(t,u) = \{[(a-b-d)^2 - c^2] + [(a+b-d)^2 - c^2]t^2\}u^2 - 8\lambda bdtu + [(a+b+d)^2 - c^2]t^2 + [(a-b+d)^2 - c^2] = 0 \quad (1)$$

Equation (2) gives the displacements of crank 2 and link 4 for the spherical four-bar linkage.

$$f(t,u) = \{[\cos(a-b-d) - \cos c] + [\cos(a+b-d) - \cos c]t^2\}u^2 + 4tu \sin b \sin d + [\cos(a-b+d) - \cos c] = 0 \quad (2)$$

These equations are useful not only in determination of the angular displacements of links 2 and 4, but in the determination of their angular velocities and accelerations, which follow by differentiation.

Link sizes are expressed in ordinary length units, such as inches (for the plane and skew four-bar linkages), and in great circular arcs (in degrees, say, for the spherical four-bar linkage). Let *a, b, c, d* denote the lengths (linear of arc) of ground (*AD*), crank, coupler, and output link, respectively, and let ξ denote the angular offset between pin axes of links 2 and 4 of the simply skewed four-bar linkage. $\lambda = 1$ (in a plane four-bar linkage) or $\cos \xi$ (in a simply skewed four-bar linkage). If ϕ and ψ denote the angular positions of links 2 and 4, with $t = \tan \phi/2$ and $\mu = \tan \psi/2$, the displacement equations, $f(t,u) = 0$, giving the relationship between angular displacements of links 2 and 4, are as in Eqs. (1) and (2).

Design. This involves proportioning the linkages for given motion and power requirements and considering kinematics, dynamics, strength, lubrication, wear, balancing, tolerances, actuation, control and other factors. At low speeds the links can often be treated as rigid bodies, but at high speeds the elasticity needs to be included. The design procedures can be analytical, graphical, experimental, or a combination of these.

Due to the nonlinearity of the motion of four-bar linkages, their design is often time-consuming and difficult. However, the determination of linkage proportions for certain standard design requirements has been worked out in an efficient manner. These include transmission-angle optimization, the coordination of the rotations of links 2 and 4, and the generation of a particular coupler motion or point path (**Fig. 9**). In addition, general-purpose computer codes for the kinematic design and dynamic analysis of linkages, including four-bar linkages, have been developed.

Design trends emphasize high-speed operation and increasingly sophisticated motion and power-transmission requirements. *See* MECHANISM; SLIDER-CRANK MECHANISM; STRAIGHT-LINE MECHANISM.

Ferdinand Freudenstein

Bibliography. A. G. Erdman, *Modern Kinematics: Developments in the Last Forty Years*, 1993; A. G. Erdman and G. N. Sandor, *Mechanism Design: Analysis and Synthesis*, vol. 1, 3d ed., 1996; B. Paul, *Kinematics and Dynamics of Planar Machinery*, 1979; S. B. Tuttle, *Mechanisms for Engineering Design*, 1967.

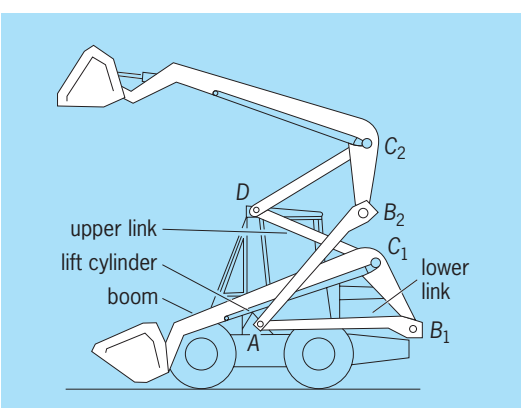

Fig. 9. Front loader with the bucket mounted on coupler *BC* of the four-bar linkage *ABCD*.

Fourier series and transforms

Mathematical tools for the analysis of functions through decomposition into sinusoids. In 1822, J. Fourier proposed that ordinary mathematical functions could be represented by a sum of sinusoids, even though at first sight it might seem that the blandness and repetitive nature of the sinusoid ill suited it to accommodate the variety of functions in general. The advantage of such a representation is this: A differential equation that is difficult to solve under the untidy given external conditions that often arise in technology may very well be soluble under an external condition that is simply sinusoidal. If solutions can be obtained for each of the constituent sinusoids that make up the given external function, then perhaps the required solution will be the sum of the separate simple solutions. This was indeed so in the case studied by Fourier, who was interested in the differential equation governing the diffusion of heat in homogeneous solids; in fact, it is often the case that solving for sinusoids permits synthesis of the required solution. When this does happen, the solution is itself composed of sinusoids and there is said to be a sinusoidal response to sinusoidal input; the differential equation will also exhibit linearity combined with shift invariance. *See* DIFFERENTIAL EQUATION; HARMONIC ANALYZER; LINEARITY; TRIGONOMETRY.

Fourier series. A Fourier series is a sum of a constant and any number of sinusoidal functions, say of t, with the property that the frequencies of the component functions belong to an arithmetic sequence. An example is expression (1), where ω_0 is the fundamental frequency. The phase of each term is arbitrary, as is the amplitude, and in particular the amplitude associated with the fundamental term may be zero. Allowing an amplitude c_n for the nth term, and a phase α_n, the most general time-dependent Fourier series is given by expression (2). An equivalent expression,

$$2 + \cos(\omega_0 t + 0.4) + {}^1/_2 \cos(2\omega_0 t + 0.5) + {}^1/_4 \cos(3\omega_0 t - 0.2) \quad (1)$$

$$\sum_{n=0}^{\infty} c_n \cos(n\omega_0 t + \alpha_n) \quad (2)$$

which uses two amplitudes a_n and b_n instead of one amplitude and a phase, is (3), where the coefficients in the two expressions are related by Eqs. (4). If the

$$\sum_{n=0}^{\infty} (a_n \cos n\omega_0 t + b_n \sin n\omega_0 t) \quad (3)$$

$$\begin{aligned} a_n &= c_n \cos \alpha_n \\ b_n &= -c_n \sin \alpha_n \end{aligned} \quad (4)$$

fundamental frequency is unity, the Fourier series in its simplest form results in expression (5).

$$a_0 + (a_1 \cos t + b_1 \sin t) + (a_2 \cos 2t + b_2 \sin 2t) + \cdots \quad (5)$$

Fourier series have wide application in physics because systems that vibrate in response to a stimulus often do so periodically; a violin string offers an example. The reason that the vibrations of a clamped uniform string are closely periodic is that the velocity of low-amplitude transverse waves on a stretched uniform string is dependent, to a good approximation, on only the string tension and the string mass per unit length. *See* VIBRATION.

Not all natural vibrations are periodic; examples are furnished by drums, bells, bars, and horns, where the more general trigonometric series applies, namely expression (6). The frequencies ω_n in such

$$\sum_{n=0}^{\infty} c_n \cos(\omega_n t + \alpha_n) \quad (6)$$

cases are not necessarily in arithmetic sequence.

No term in the series (5) is altered in value if t is replaced by $t + 2\pi$; consequently if the series has a finite sum, that sum will be a function of t that is periodic in t with period 2π. In the discussion of Fourier series, attention may therefore be restricted to periodic functions. A real function $f(t)$ defined on the closed interval $[-\pi, \pi]$ and periodic with period 2π is said to have a Fourier series (5), where the coefficients are given by Eq. (7), and for $n > 0$, by Eqs. (8) and (9).

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\, dt \quad (7)$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos nt\, dt \quad (8)$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt\, dt \quad (9)$$

Existence conditions. When $f(t)$ is not suitably behaved, one or more of the integrals for the Fourier coefficients may not exist. However, the functions encountered in everyday experience are not apt to cause difficulty; even multiple discontinuities and poles are not sufficient to exclude a function. However, the existence conditions have been in contention since the time of Fourier and were responsible for major developments in the theory of functions of a real variable. The flavor of the discussion is given by noting that in the case of $f(t) = \cos(1/t)$, a function that has an infinite number of maxima in any finite interval around $t = 0$, no matter

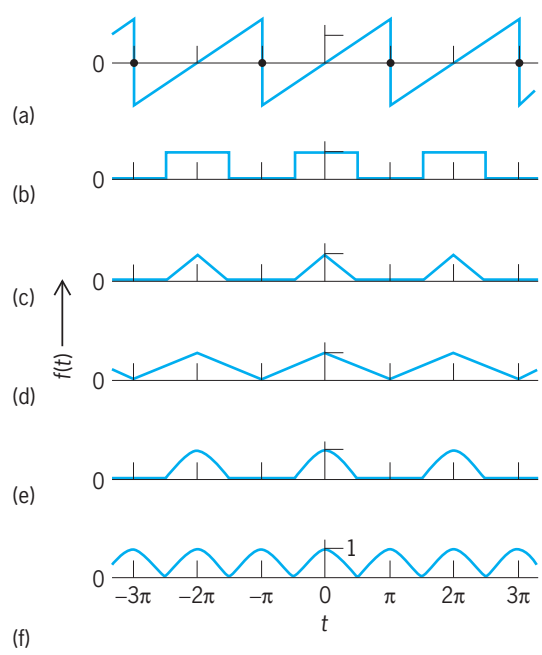

Examples of (*a*–*c*) macrodasyid and (*d*–*f*) chaetonotid gastrotrichs. (*a*) *Xenodasys*. (*b*) *Dolichodasys*. (*c*) *Thaumastoderma*. (*d*) *Draculiciteria*. (*e*) *Halichaetonotus*. (*f*) *Aspidiophorus*. (*part a after C. Schoepfer-Sterrer, Chordodasys riedli gen. nov., spec. nov., a macrodasyoid gastrotrich with a chordoid organ, Cah. Biol. Mar., 10:391–404, 1969; b after G. D. Gagne, Dolichodasys elongatus n.g., n. sp., a new macrodasyid gastrotrich from New England, Trans. Amer. Microsc. Soc., 96:19–27, 1977; c after A. Remane, Klassen und Ordnungen des Tierreichs, vol. 4(2), Akad. Verlagsges., Leipzig, 1936; d after P. Luporini et al., Contribution à la connaissance des Gastrotriches des côtes de Toscane, Cah. Biol. Mar., 12:433–455, 1971; e after H. Mock, Chaetonotoidea (Gastrotricha) der Nordseeinsel Sylt, Mikrofauna Meersebodens, 78:1–107, 1979; f after W. D. Hummon, Gastrotricha from Beaufort, North Carolina, U.S.A., Cah. Biol. Mar., 15:431–446, 1974*)

how narrow, the a_n do not exist. However, coefficients do exist for the function $t \cos (1/t)$, which also has an infinite number of maxima and minima. This situation can be covered by saying that the coefficients will exist if the arc length of $f(t)$ is finite. Neither of the above functions can be graphed, and the discussion belongs to a realm outside everyday experience. Furthermore, there are functions with infinite arc length that do have Fourier coefficients. Even if the coefficients exist, the sum of the Fourier series need not equal the original function. The rising importance of finite mathematics, largely associated with computers, means that, increasingly, users of Fourier series are concerned only with entities that can be represented by finite sets of data having finite values; the remainder of this article addresses such applications. *See* REAL VARIABLE; SERIES.

Examples. A simple periodic function is the periodic sawtooth waveform which equals $^1/_2 t$ in the interval $-\pi < t < \pi$. The coefficients a_n, including a_0, are all zero because of odd symmetry, and the values of b_n are easily obtained by integration to yield the series (10). This particular series converges to $^1/_2 t$ in the

$$\sin t - {}^1/_2 \sin 2t + {}^1/_3 \sin 3t - {}^1/_4 \sin 4t + \cdots \qquad (10)$$

open interval $(-\pi, \pi)$ and had been summed by L. Euler before 1800. For other values of t (except odd multiples of π) it converges to the sawtooth generated by replication at multiples of 2π. Where $t = \pm\pi, \pm 3\pi, \pm 5\pi, \ldots$ the series also converges; the sum is zero. It is a general property that at a point of simple discontinuity, where the periodic function jumps, the series sums to the midpoint between the limiting values approached from left and right (**illus.** *a*).

The Fourier series for some simple waveforms (see illus.) given in Eqs. (11). The sinc function is defined by $\operatorname{sinc} x = (\pi x)^{-1} \sin \pi x$.

Sawtooth:

$$a_0 = 0 \quad a_n = 0 \quad b_n = (-1)^{n-1} n^{-1} \qquad (11a)$$

Rectangles:

$$a_0 = {}^1/_2 \quad b_n = 0 \quad a_n = \operatorname{sinc} {}^1/_2 n \qquad (11b)$$

Spaced triangles:

$$a_0 = {}^1/_4 \quad b_n = 0 \quad a_n = {}^1/_2 \operatorname{sinc}^2 {}^1/_4 n \qquad (11c)$$

Adjacent triangles:

$$a_0 = {}^1/_2 \quad b_n = 0 \quad a_n = \operatorname{sinc}^2 {}^1/_2 n \qquad (11d)$$

Spaced cosinusoidal half-cycles:

$$\begin{aligned} &a_0 = 1/\pi \quad b_n = 0 \\ &a_n = {}^1/_2 \operatorname{sinc}\left({}^1/_2 n + {}^1/_2\right) + {}^1/_2 \operatorname{sinc}\left({}^1/_2 n + {}^1/_2\right) \end{aligned} \qquad (11e)$$

Adjacent cosinusoidal half-cycles:

$$\begin{aligned} &a_0 = 2/\pi \quad b_n = 0 \quad a_1 = 0 \\ &a_n = -4 \cos\left({}^1/_2 n\pi\right)/[\pi(n^2 - 1)] \text{ for } n > 1 \end{aligned} \qquad (11f)$$

Results for waveforms formed by replicating other pulse shapes can be obtained by integration; alternatively, the coefficients can be obtained from the Fourier transform of a single pulse, as explained below.

Fourier transforms. The integrals in Eqs. (8) and (9), which are Fourier integrals, have been of great historical importance. In modern times they have gained additional importance for Fourier analysis of aperiodic functions and the corresponding spectral analysis or diffraction of electromagnetic radiation (especially for chemical analysis by ultraviolet, visible, and infrared light, and by x-rays), microwaves, seismic waves, sound waves, ocean waves and tides, and vibrations. The advent of fast computing has also given importance to algorithms such as the fast Fourier and fast Hartley transforms, which have facilitated progress in plasma physics and fluid dynamics, where numerical multidimensional analysis was previously infeasible. *See* ACOUSTIC SIGNAL PROCESSING; DIFFRACTION; HOLOGRAPHY; INFRARED SPECTROSCOPY; INTERFEROMETRY; OCEAN WAVES; SEISMOLOGY; X-RAY CRYSTALLOGRAPHY.

A function $f(x)$ that is not periodic does not analyze into a discrete set of terms, each of which has finite energy carried by a single frequency, but rather into a spectral continuum. This continuum is analogous to the spectrum of sunlight, whose energy spectrum is describable in terms of the energy per hertz or other unit of frequency. A function $f(x)$ and its amplitude spectrum or Fourier transform $F(s)$ are related by Eqs. (12). If x represents time in seconds,

$$\begin{aligned} F(s) &= \int_{-\infty}^{\infty} f(x) e^{-i2\pi sx} dx \\ f(x) &= \int_{-\infty}^{\infty} F(s) e^{i2\pi sx} ds \end{aligned} \qquad (12)$$

then s is frequency in cycles per second, or hertz; while if x represents distance, then s is spatial frequency in cycles per unit of x. In two dimensions the Fourier transform is given by Eq. (13), where u

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-i2\pi(ux+vy)} \, dx \, dy \qquad (13)$$

and v are spatial frequencies measured in the x and y directions respectively. *See* LAPLACE TRANSFORM.

A few special functions, defined in Eqs. (14), (15), and (16), are useful in specifying basic Fourier

$$\operatorname{rect} x = \begin{cases} = 1, & |x| < 0.5 \\ = 0, & |x| > 0.5 \end{cases} \qquad (14)$$

$$\Lambda(x) = \begin{cases} 1 - |x|, & |x| < 1 \\ 0, & |x| > 1 \end{cases} \qquad (15)$$

TABLE 1. Some Fourier transform pairs

$e^{-\pi x^2}$	$e^{-\pi s^2}$
sinc x	rect s
sinc^2 x	$\Lambda(s)$
1	$\delta(s)$
$\cos \pi x$	$\frac{1}{2}\delta(s + \frac{1}{2}) + \frac{1}{2}\delta(s - \frac{1}{2})$
$\sin \pi x$	$\frac{1}{2}i\delta(s + \frac{1}{2}) - \frac{1}{2}i\delta(s - \frac{1}{2})$
III(x)	III(s)

$$\mathrm{III}(x) = \sum_{n=-\infty}^{\infty} \delta(x - n) \tag{16}$$

transform pairs (**Table 1**). The unit impulse function, or Dirac delta function, is represented by $\delta(x)$, while the infinite train of unit impulses at unit spacing is represented by the shah function III(x), an entity that is its own Fourier transform. Sample values of a function $f(x)$ defined on continuous x are expressible by the product III(x)$f(x)$, while a periodic function derived by replicating $f(x)$ is expressible by the convolution III(x) $* f(x)$. The asterisk notation for convolution is defined by Eq. (17).

$$f(x) * g(x) = \int_{-\infty}^{\infty} f(x - u)g(u)\,du \tag{17}$$

Generation of Fourier pairs by theorem. A number of general theorems for the Fourier transform (**Table 2**) can be used to generate numerous additional Fourier transform pairs. For example, knowing that $f(x) = \exp(-\pi x^2)$ has Fourier transform $F(s) = \exp(-\pi s^2)$, it is not necessary to resort to further integration to find the Fourier transform of, say, $\exp(-x^2)$. The result can be deduced from the similarity theorem, which states that $f(ax)$ has Fourier transform $|a|^{-1}F(s/a)$. Hence $\exp(-x^2)$ has Fourier transform $\pi^{1/2}\exp(-\pi^2 s^2)$.

Periodic functions. A periodic function $p(x)$ of period T can be composed by replicating a given function $f(x)$ that is zero where $|x| > T/2$. Then $p(x)$ can be expressed as the convolution of $f(x)$ with a train of unit impulses at spacing T, namely $T^{-1}\mathrm{III}(x/T)$. The factor T^{-1} is required because an impulse when stretched becomes stronger, a property that is expressed by $\delta(x/a) = |a|\ \delta(x)$. Thus, $p(x)$ is given by Eq. (18).

$$p(x) = T^{-1}\mathrm{III}(x/T) * f(x) \tag{18}$$

Application of the convolution theorem yields the Fourier transform $P(s)$ as a simple sampling product, Eq. (19), which is equal to zero except where

$$P(s) = \mathrm{III}(Ts)F(s) \tag{19}$$

$Ts = n$ (an integer). The only information retained about $F(s)$ is the set of sample values, $F(n/T)$, $-\infty < n < \infty$. These values must be equivalent to the Fourier coefficients; thus $a_0 = T^{-1}$ and, for $n > 0$, the relationship is given by Eq. (20).

$$a_n - ib_n = 2T^{-1}F(n/T) \tag{20}$$

Discrete Fourier transform. In numerical practice, the equivalent of a given function $f(x)$ of the continuous independent variable x is a finite number N of data values associated with a serial number τ, customarily taken as running from 0 to $N - 1$. Then there is a discrete Fourier transform $F(\nu)$ defined by Eq. (21). The quantity ν is an integer running from 0

$$F(\nu) = N^{-1}\sum_{\tau=0}^{N-1} f(\tau)e^{-i2\pi(\nu/N)\tau} \tag{21}$$

to $N - 1$. The factor N^{-1} ensures that $F(0)$ is the mean value of the N data values and is thus analogous to a_0, the Fourier coefficient that represents the mean value over one period. For $1 \le \nu \le N/2$, the quantity ν/N is analogous to frequency as measured in cycles per spacing of τ values. Values of ν exceeding $N/2$ represent negative frequencies $(\nu - N)/N$. The lowest positive frequency, or fundamental, is $1/N$, and the highest frequency is $\frac{1}{2}$. The shortest period representable by data is necessarily two units of τ, since at least two samples per period are required. The purpose of computing the discrete transform may be to present the power spectrum $|F(\nu)|^2$. In other cases the purpose may be to manipulate the spectrum by removing low frequencies, enhancing high frequencies, or removing high frequencies representing unwanted noise. Having done this, one returns to the data domain. Given a discrete transform $F(s)$, the corresponding data-domain function $f(\tau)$ is retrieved by using Eq. (22).

$$f(\tau) = \sum_{\nu=0}^{\infty} F(s)e^{i2\pi(\nu/N)\tau} \tag{22}$$

See INFORMATION THEORY; Z TRANSFORM.

Hartley transform. Where a data set consists of real values and there is no point in going to complex representation, the Hartley transform offers an alternative to the Fourier transform, especially in its

TABLE 2. Theorems for the Fourier transform

Theorem	Function, $f(x)$	Fourier transform, $F(s)$
Similarity	$f(ax)$	$\lvert a\rvert^{-1}F(s/a)$
Addition	$f(x) + g(x)$	$F(s) + G(s)$
Shift	$f(x - a)$	$e^{-i2\pi as}F(s)$
Convolution	$f(x) * g(x)$	$F(s)G(s)$
Modulation	$f(x)\cos 2\pi f_0(x)$	$\frac{1}{2}F(s + f_0) + \frac{1}{2}F(s - f_0)$
Autocorrelation	$\int_{-\infty}^{\infty} f^*(u)f(u + x)du$	$\lvert F(s)\rvert^2$
Derivative	$f'(x)$	$i2\pi sF(s)$
Rayleigh's	$\int_{-\infty}^{\infty} \lvert f(x)\rvert^2 = \int_{-\infty}^{\infty} \lvert F(s)\rvert^2 ds$	

discrete form. The definition is Eq. (23), and the inversion formula is Eq. (24), where cas $\theta = \cos\theta + \sin\theta$. For computing purposes it is convenient not to handle imaginaries, especially with multidimensional applications such as images, and there is a simplicity about calling a Hartley transform subprogram which, unlike the Fourier, is the same whether the operation performed is going to the transform domain or returning from it. The power spectrum, which is computed from the Fourier transform by summing the squares of the real and imaginary parts, is obtained as $[H(\nu)]^2 + [H(N-\nu)]^2$. The complex Fourier transform of real data contains a twofold redundancy, since $F(N-\nu)$ is not independent of $F(\nu)$, and this redundancy imposes a time penalty on computation. Nonreciprocal transforms accepting complex input and delivering real output and others converting real to complex can be combined into packages to avoid the waste time, but the elegant Hartley technique has made such devices unnecessary.

$$H(\nu) = N^{-1}\sum_{\tau=0}^{N-1} f(\tau)\,\mathrm{cas}\,(2\pi\nu\tau/N) \qquad (23)$$

$$f(\tau) = \sum_{\nu=0}^{N-1} H(\nu)\,\mathrm{cas}\,(2\pi\nu\tau/N) \qquad (24)$$

Fast Fourier transform. In applications of harmonic analysis the discrete Fourier transform may be computed directly from Eq. (21) but, where speed is essential, less time will be required if the data values are split into two sets of length $N/2$ and two separate (but shorter) discrete Fourier transforms are computed and suitably combined. A further time saving occurs if the shorter data sets are themselves halved, and so on. Because of this, data sets are often set up so that, $N = 2^P$, where P is some power, so that halving can continue down to length 2. Around 1805 Gauss was already practicing this technique and stated that an advantage could be obtained, wherever there were factors such that $N = n_1 n_2$, by dividing the data sets into lengths n_1 and n_2. An algorithm that uses this idea is called a fast Fourier transform. Significant further time saving results when $N = 4^P$, in which case the algorithm is called a radix-4 fast Fourier transform. When the number of data values given falls short, for example, if $N = 730$, it may be advantageous to append 294 zeros to the data in order to bring N up to $4^5 = 1024$. It is true that in principle the factorization $2 \times 5 \times 73 = 730$ could lead to time saving, but it is inconvenient to maintain special computer code for prime factors such as 73. Furthermore, extending with zeros means that the discrete Fourier transform values, when computed, will be packed a little closer together than is strictly necessary, which leads to a smoother graph, in the event that a power spectrum, such as an infrared spectrum, is to be displayed. In addition, the speed of computers has repeatedly doubled, whereas improvements in fast algorithms have slackened. Consequently, radix-4 algorithms can be recommended for general-purpose harmonic analysis. When data values are real, as is usually the case, the fast Hartley transform algorithm, which is strictly analogous to the fast Fourier transform, will be twice as fast as the comparable complex fast-Fourier transform computer code. Ronald N. Bracewell

Bibliography. R. N. Bracewell, *The Fourier Transform and Its Applications*, 3d ed., 1999; R. N. Bracewell, *The Hartley Transform*, 1986; R. N. Bracewell, *Two-Dimensional Imaging*, 1994; J. W. Goodman, *Introduction to Fourier Optics*, 2d ed., 1996; J. S. Walker, *Fourier Analysis*, 1988.

Fractals

Geometrical objects that are self-similar under a change of scale, for example, magnification. The concept is helpful in many disciplines to allow order to be perceived in apparent disorder. For instance, in the case of a river and its tributaries, every tributary has its own tributaries so that it has the same structural organization as the entire river except that it covers a smaller area. The branching of trees and their roots as well as that of blood vessels, nerves, and bronchioles in the human body follows the same pattern. Other examples include a landscape with peaks and valleys of all sizes, a coastline with its multitude of inlets and peninsulas, clouds, lightning, the mass distribution within a galaxy, the distribution of galaxies in the universe, and the structure of vortices in a turbulent flow. The rise and fall of economic indices has a self-similar structure when plotted as a function of time.

Mathematical versus natural fractals. A triadic Koch curve (see **illus.**) is a good example of how a fractal may be constructed. The procedure begins with a straight segment. This segment is divided into three equal parts, and the (single) central piece is replaced by two similar pieces (illus. *a*). The same procedure is now applied to each of the four new segments (illus. *b*), and this is repeated an infinite number of times. The curve is self-similar, because a magnification by 3 of any portion will look the same as

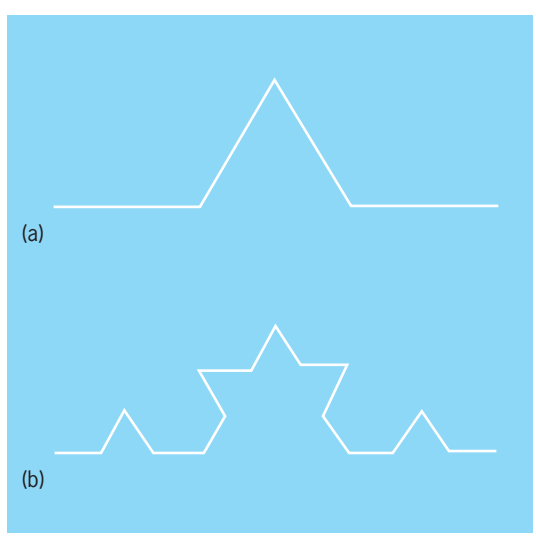

Koch curve, (*a*) first and (*b*) second stages.

the original curve. As a geometrical curve, the Koch curve has many unsettling properties. For instance, it has an infinite number of zigzags between any two points, and its length between any two points is infinite. The curve was constructed in the late nineteenth century to demonstrate that continuous curves are not necessarily differentiable and rectifiable. The invention of many such curves around that time marks the origin of fractals as a branch of mathematics.

Fractals came into natural sciences when it was recognized that natural objects are often random versions of mathematical fractals. They are self-similar in a statistical sense; that is, given a sufficiently large number of samples, a suitable magnification of a part of one sample can be matched closely with some member of the ensemble. Unlike the Koch curve, which must be magnified by an integral power of 3 to achieve self-similarity, natural fractal objects are usually self-similar under arbitrary magnification. Also, the scale factor need not be the same in every direction. The landscape is self-similar when it is magnified by one factor in the horizontal and another factor in the vertical direction. Such objects, including rough solid surfaces, are said to be self-affine fractals.

Natural fractals are self-similar or self-affine within a finite range of length scales. A coastline is limited from above by the size of the continent, and it cannot be subdivided beyond a grain of sand. Fractal properties of solid surfaces cannot be extended beyond the size of an atom. It is believed that natural fractals, like the Koch curve, are created by stacking up some basic building blocks—sand grains or atoms—following a set of simple rules under random conditions. There are many examples to illustrate this principle, but finding the rules of construction for a given fractal object is not always straightforward.

The multitude of zigzags of the Koch curve mimics the irregularities of a coastline with protrusions and indentations of sizes ranging from hundreds of kilometers down to a grain of sand. The length of such a curve depends on the scale of measurement, which in practice is the limit of resolution. For example, if the scale of the Koch curve in the illustration is the length of the original line segment, the central part of the original segment (illus. *a*) will not be visible. An improvement of the resolution by a factor of 3 makes visible this entire structure (illus. *a*) but not the finer details (illus. *b*). The curve now appears longer by a factor of $4/\sqrt{3}$. By repeating this procedure, the length increases by a factor of $4/\sqrt{3}$ every time the resolution is improved by a factor of 3. Similarly, the length of a coastline measured from a satellite picture is considerably shorter than that measured from the ground.

Fractal dimension. A magnification of a line segment by a factor of 3 results in an increase in length by the same factor. For a square, a magnification of each side by a factor of 3 results in an increase in area by a factor of 9 or 3^2. For a cube, the increase in volume is proportional to the third power of the magnification factor of each side. In every case the dimension of the object appears as the power of the length scale factor. In contrast, the length of the Koch curve as defined in the preceding paragraph increases by a factor of 4 when it is magnified by a factor of 3. An analogous definition of the dimension, $4 = 3^D$, leads to the fractal or Hausdorff dimension of the curve, $D = (\ln 4)/(\ln 3) \simeq 1.26$. The fractal dimensions of a typical coastline is 1.2. Self-affine fractals have well-defined dimensions when one length scale dominates. Landscapes and catalytically active substrates have $D \simeq 2.1$ compared with $D = 2$ for smooth surfaces. Thus, D is a quantitative measure of how much more detail is visible in an object when it is magnified.

Some fractal objects appear less substantial when examined in more detail. For instance, the distribution of matter in a galaxy is such that more empty space is apparent under higher resolution. The dimensions of such objects are less than the spatial dimension in which they are embedded, for example, 1.2 for typical galaxies. Polymer molecules in solutions form loosely packed spheres with $D \simeq 1.7$. On the other hand, space-covering fractals have dimensions equal to the spatial dimension. The river system has $D = 2$, meaning that a small body of water drains the entire land. The blood vessel system, with $D = 3$, allows a small amount of blood to carry nutrients to every cell of the body.

Physical properties. In a diffusion process in uniform space, the mean-square-average distance covered by the diffusing substance is proportional to the lapsed time. The same process on a fractal network, for example, a percolation cluster, gives a very different result. A percolation network is a lattice constructed of a random mixture of conducting and nonconducting links. At low concentrations of conducting links, the lattice is an insulator because there is no connected path between two opposite ends of the lattice. At a critical concentration, a fractal connected path is formed and the cluster becomes a conductor. In a diffusion process on a two-dimensional model of the percolation cluster, the mean-square-average distance is proportional to the time raised to the power 0.7. This is one illustration of the unusual dynamical properties of fractal objects. A fractal cluster consisting of masses connected by springs has an unusual vibrational spectrum in that the distribution of normal modes is fractal. The relations between fractal geometry, turbulent flow, and chaotic motion of nonlinear systems are under active investigation. *See* CHAOS; DIFFUSION.

General applications. Physicists have used the concept of fractals to study the properties of amorphous solids and rough interfaces and the dynamics of turbulence. It has also been found useful in physiology to analyze the heart rhythm and to model blood circulation, and in ecology to understand population dynamics. In computer graphics it has been shown that the vast amount of information contained in a natural scene can be compressed very effectively by identifying the basic set of fractals therein together with their rules of construction. When the fractals are reconstructed, a close approximation of the original scene is reproduced. *See* AMORPHOUS SOLID; CARDIOVASCULAR SYSTEM; COMPUTER GRAPHICS; TURBULENT FLOW.

S. H. Liu

Applications in physics. As noted above, natural fractal objects show structures at all scales and are therefore intrinsically irregular. Their properties cannot be described by the usual mathematical concepts that focus instead on regularity and analyticity. Fractal geometry has provided a mathematical tool to characterize these structures in a quantitative way.

Before the application of fractals to physics, scale invariance had already appeared in the study of the critical properties of continuous phase transitions. This concept was essential in the development of the renormalization group theory that has provided an understanding of the origin of scale invariance in these systems. However, in critical phenomena fractal properties appear only at the critical point and, in order to be observed, they require an extremely accurate fine tuning of an external parameter such as the temperature. *See* CRITICAL PHENOMENA; RENORMALIZATION.

By contrast, most of the fractal structures observed in nature arise from irreversible dynamical processes and do not require any fine tuning. These differences are of great conceptual importance and imply that new theoretical concepts are necessary for the understanding of these phenomena. In particular, it is essential to focus on the temporal evolution and the dynamical aspects of these phenomena, while in the study of phase transitions only the static configurations are considered.

Fractal growth models. The first model of irreversible fractal growth based on a well-defined physical process is the diffusion-limited aggregation model introduced in 1981 as a model of smoke aggregation and colloid formation. The process starts with a seed particle, and a second particle is introduced into the system which moves with random-walk dynamics (brownian motion) until it touches the first particle. At this point it stops and becomes part of the structure. A new particle is then introduced, and the process is repeated iteratively. The result of this simple process is a random fractal structure of great complexity. *See* BROWNIAN MOVEMENT.

A generalization of diffusion-limited aggregation, the dielectric breakdown model, shows that a system which obeys the Laplace equation and follows a stochastic iteration scheme spontaneously configures its boundaries into a highly irregular fractal structure. Such growth models can describe the essential fractal properties of a great variety of phenomena such as electrochemical deposition, dielectric breakdown, viscous fingering, the propagation of fractures, colloids, and polydisperse solutions. *See* LAPLACE'S DIFFERENTIAL EQUATION; STOCHASTIC PROCESS.

Unsolved problems. Nevertheless, in many other systems well-defined fractal properties have been identified either from experiments or by computer simulations, but theoretical understanding is lacking. An important example is fully developed turbulence, for which there is empirical and phenomenological evidence that energy dissipation occurs on a highly irregular, possibly fractal structure, but the connection of this property to the Navier-Stokes equations is still unknown.

Another important problem is represented by the large-scale structure of the universe. Visible matter is aggregated in galaxies that are part of groups that are aggregated into clusters and superclusters, and so on. The analysis of galaxy and cluster catalogs with the methods of modern statistical mechanics shows that visible matter has fractal properties up to the present limits of observation. This situation is in contrast with the usual assumption of large-scale homogeneity and leads to a theoretical challenge. *See* COSMOLOGY; GALAXY, EXTERNAL; UNIVERSE.

Disordered media. A more natural and direct application of fractal concepts is in the field of disordered systems and porous media, mentioned above. Often these systems show well-defined fractal behavior, so it is important to study the physical properties of fractal structures such as their vibrations (fractons) and their electronic and transport properties.

L. Pietronero

Bibliography. M. Barnsley, *Fractals Everywhere*, 2d ed., 1993; P. H. Coleman and L. Pietronero, The fractal structure of the universe, *Phys. Rep.*, 213: 311–391, 1992; H. Ehrenreich and D. Turnbull (eds.), *Solid State Physics*, vol. 39, 1986; B. B. Mandelbrot, *The Fractal Geometry of Nature*, 1988; L. Pietronero and E. Tosatti (eds.), *Fractals in Physics*, 1986; T. Vicsek, *Fractal Growth Phenomena*, 1992.

Frame of reference

A base to which to refer physical events. A physical event occurs at a point in space and at an instant of time. Each reference frame must have an observer to record events, as well as a coordinate system for the purpose of assigning locations to each event. The latter is usually a three-dimensional space coordinate system and a set of standardized clocks to give the local time of each event. For a discussion of the geometrical properties of space-time coordinate systems *see* SPACE-TIME; RELATIVITY.

Newtonian reference frames. In the ordinary range of experience, where light signals, for all practical purposes, propagate instantaneously, the time of an event is quite distinct from its space coordinates, since a single clock suffices for all observers, regardless of their state of relative motion. The set of reference frames which have a common clock or time is called newtonian, since Isaac Newton regarded time as having invariable significance for all observers.

Inertial reference frames. In these a body moves with constant velocity when free of impressed forces. It follows that one inertial frame cannot be accelerating with respect to another. If x_1, y_1, z_1 and x_2, y_2, z_2 are the coordinates of a particle P in parallel inertial frames 1 and 2, where 2 is moving with respect to 1 with velocity components v_x, v_y, v_z, then Eqs. (1) hold. (Newtonian frames are understood

$$\begin{aligned} x_2 &= x_1 - v_x t \\ y_2 &= y_1 - v_y t \\ z_2 &= z_1 - v_z t \end{aligned} \tag{1}$$

here so that $t_2 = t_1$.) This is called a galilean transformation. Since the vs are independent of time, the component accelerations of P in 1 and 2 are the same, that is, Eq. (2) is valid.

$$A_{x_2} = \frac{d^2x_2}{dt^2} = \frac{d^2x_1}{dt^2} = A_{x_1} \text{ etc.} \tag{2}$$

Accelerated reference frames. A noninertial frame of reference is called an accelerated frame. A point fixed in such a frame will be accelerated with respect to inertial frames and so must be held by a force. An observer in an accelerated frame will see accelerations that are not proportional to the impressed forces, since a force is required to keep the particle at rest. Newton's second law does not hold without modification. *See* ACCELERATION.

Rotating reference frames. A rotating frame of reference is one whose coordinate axes are rotating with respect to some inertial system. A point P fixed in the rotating frame moves uniformly about a circle in the inertial frame and is accelerated toward the axis of rotation. There is an additional acceleration (with respect to the inertial system) if P moves in the rotating system. For example, if the point moves away from the axis along a moving radius, it gains circular speed also. *See* CENTRIPETAL FORCE; CORIOLIS ACCELERATION.

Newton's second law can be modified to take account of these two accelerations that are not observed in the rotating system. Let A_{Ac}, A_C, A_{Co} be the observed, centripetal, and Coriolis accelerations, respectively. The total acceleration in the inertial frame is given by Eq. (3), where the accelerations must

$$A = A_{Ac} + A_C + A_{Co} \tag{3}$$

be added vectorially. Defining the centrifugal force, $Fc = -mAc$, and the Coriolis force, $F_{Co} = -mA_{Co}$, permits Newton's second law to be written as Eq. (4).

$$mA_{Ac} = F + F_C + F_{Co} \tag{4}$$

In rotating systems a consistent mechanical theory is obtained if the two additional forces are added to the impressed forces. Conversely, the presence of such forces will enable the observer to determine the motion of his reference frame.

Astronomical reference frames. These are fixed relative to celestial objects. The most convenient one for observational purposes is, of course, a geocentric system. The Earth's axis of rotation is used as a coordinate axis. A second reference direction is chosen as the intersection of the Earth's equatorial plane and the plane of the Earth's orbit (the ecliptic). The basic time unit is the Earth's rotation period, the sidereal day. The positions and times of distant astronomical events are determined indirectly. *See* ASTRONOMICAL COORDINATE SYSTEMS.

Reference frames are regarded as more basic as more phenomena appear regular with respect to them. Thus the planetary motions are described naturally in an ecliptic coordinate system centered on the Sun, and the motions of stars as viewed from the Earth appear regular in a coordinate system with its origin at the center of the galaxy.

Reference frames based on the mean positions of large enough celestial objects, for example, a group of galaxies, approximate inertial frames closely. Between two such reference frames, the distances and relative velocities may be so large that transmission of light signals takes considerable time. Newtonian reference frames with a single universal time are no longer meaningful, and each reference frame must have its own system of time. Inertial frames of this type are called galilean reference frames, a somewhat misleading name since the transformation of the coordinates of an event from one such frame to another is not the Galilean transformation given earlier (with $t_1 = t_2$) but one which involves the time directly. It is known as a Lorentz transformation. Galilean reference frames are often called Lorentz frames. *See* LORENTZ TRANSFORMATIONS. Bernard Goodman

Bibliography. A. Einstein, *The Meaning of Relativity*, 5th ed., 1956; D. Halliday, R. Resnick, and K. S. Krane, *Physics*, 2 vols., 5th ed., 2002; J. B. Hartle, *Gravity*, 2003; C. Kittel, W. D. Knight, and M. A. Ruderman, *Mechanics*, Berkeley Physics Course, vol. 1, 2d ed., 1973; D. Morrison, S. C. Wolff, and A. Fraknoi, *Abell's Exporation of the Universe*, 7th ed., 1995.

Francisella

A genus of very small, coccoid to ellipsoidal, pleomorphic, nonmotile and nonsporulating, gram-negative, rod-shaped bacteria. It was named for Edward Francis, who studied the tularemia bacterium. Fastidious and strictly aerobic, it grows at 98.6°F (37°C) within several days, but only in enriched media such as coagulated egg yolk or glucose-cysteine-blood-agar. The genus is catalase-negative. Acid is produced from several carbohydrates; indole, urease, gelatinase, and nitrate reductase are not formed. The genome deoxyribonucleic acid contains 32–36 mol % guanine plus cytosine. The organisms occur in natural waters of the Northern Hemisphere, and can be parasitic and pathogenic in birds, anthropods, and mammals, including humans.

Francisella tularensis was first isolated in 1911 from ground squirrels with a plaguelike disease in Tulare County, California. It has since been found in many wild animals throughout North America, continental Europe, and Asia. Tularemia in humans is acquired via transmission by blood-sucking arthropods or by contact with infected animals, most frequently by hunters and butchers. *See* TULAREMIA.

Francisella novicida was isolated only once from a water pond, and various other *Francisella*-like organisms have been isolated from patients but so far have remained unclassified. *See* ANTIBIOTIC; MEDICAL BACTERIOLOGY. Walter Mannheim

Bibliography. N. R. Krieg and J. G. Holt (eds.), *Bergey's Manual of Systematic Bacteriology*, 9th ed., vol. 1, 1984; M. P. Starr (ed.)., *Prokaryotes: A Handbook on Habitats, Isolation, and Identification of Bacteria*, 1989.

Francium

A chemical element, Fr, atomic number 87, an alkali metal element falling below cesium in group 1 of the periodic table. Distinguished by nuclear instability, francium exists only in short-lived radioactive forms, the most durable of which has a half-life of 21 min. The chief isotope of francium is actinium-K, an isotope of mass 223, which arises from the radioactive decay of the element actinium. From the properties of the known isotopes, it is reasonably certain that no long-lived form of element 87 will ever be found in nature or synthesized artificially. *See* PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

The chemical properties of francium can be studied only on the tracer scale. The element has all the properties expected of the heaviest alkali element. With few exceptions, all the salts of francium are water-soluble. *See* ACTINIUM; ALKALI METALS.

Earl K. Hyde

Bibliography. G. Choppin et al., *Radiochemistry and Nuclear Chemistry*, 2001; G. Friedlander et al., *Nuclear and Radiochemistry*, 3d ed., 1981; E. K. Hyde, *Radiochemistry of Francium*, U.S. Atomic Energy Commission, 1960.

Franck-Condon principle

The generalization that the transition from one energy state of a molecular system to another occurs so nearly instantaneously that the nuclei of the atoms involved can be considered as stationary during the process. The Franck-Condon principle is closely related to the Born-Oppenheimer approximation, in which the various motions (electronic, nuclear vibrations and rotations) are considered to be separable, and in which the electrons respond to the instantaneous vibrations of the system whereas the system responds only to the average position of the electrons. The principle, proposed by J. Franck in 1925 and developed quantum-mechanically by E. U. Condon in 1928, is important in discussing systems of more than one atom. It is therefore valuable in molecular spectroscopy and in the interpretation of the optical properties of liquids and solids.

The principle may be illustrated by considering the excitation of photoluminescence in molecules or in special luminescent sites (centers) in solids. In the state of lowest electronic energy (ground state), the interatomic distance in a diatomic molecule or the distance of a given atom in a center from its nearest neighbors (a distance called a configuration coordinate) so adjusts itself that the total energy of the system assumes a minimum value. Within this ground electronic state, however, the system may occupy many vibrational energy states, depending on the motion of the atoms composing the molecule or center. In a higher electronic state, where the forces involved are different from those in the ground state, the system has a different set of vibrational levels and a different minimum-energy interatomic distance or configuration coordinate.

If the luminescent system is in its ground state and absorbs light which raises it to an electronically excited state, the Franck-Condon principle asserts that immediately after the transition the nuclei are still in the ground-state equilibrium configuration. This configuration does not represent the condition of minimum energy for the excited state, but usually corresponds to this state in a condition of high vibrational excitation. Since the vibrational frequency of molecules or centers is about 10^{13}/S whereas the transition probability for an allowed transition emitting visible light is about 10^{8}/S, the molecule or center has ample time to relax to the new minimum configuration characteristic of the electronically excited state. It does this by giving off the excess vibrational energy as heat.

A similar process occurs in the luminescence transition back to the electronic ground state, where heat energy is again given off when the luminescent center readjusts to the ground-state equilibrium configuration and the center returns to its initial state. However, the absorption transition is achieved by the action of light energy only, whereas the system returns to the ground state by giving up heat energy twice and emitting the remaining energy as luminescence. As a result, the luminescent emission is of lower energy than the absorbed light, and since the wavelength varies inversely as the energy, the emission is of longer wavelength than is the absorbed light. This general result, known as the Stokes shift in luminescence, is thus explained on the basis of the Franck-Condon principle. *See* LUMINESCENCE; MOLECULAR STRUCTURE AND SPECTRA.

James H. Schulman; Clifford C. Klick

Bibliography. F. Argullo-Lopez and C. R. Catlow, *Point Defects in Materials*, 1988; J. H. Crawford and L. M. Slifkin (eds.), *Point Defects in Solids*, vol. 1, 1972; R. J. Elliott and A. F. Gibson, *An Introduction to Solid State Physics*, 1974.

Fraunhofer lines

Dark absorption features in the solar spectrum. They are named in honor of J. Fraunhofer, who first studied them in 1814. They are found from the ultraviolet at about 180 nanometers to the infrared at 20 micrometers. Each line represents the net

Fraunhofer lines

Wavelength, nm	Name*	Species†	Cycle variability,‡ %	Comment§
279.54		Mg II	10	Ultraviolet emission, high chromosphere
280.23				
388.36	(CN bandhead)	CN	3	High photosphere
393.36	K	Ca II	15	Chromosphere
396.85	H			
486.13	F (Hβ)	H I		Chromosphere
517.27	b_2	Mg I	—	Low chromosphere
518.36	b_1			
525.02		Fe I	0.3	Photosphere, magnetic fields ($g = 3$)
538.03		C I	0.0	Low photosphere
589.00	D_2	Na I	—	Upper photosphere, low chromosphere
589.59	D_1			
630.25		Fe I	—	Photosphere, magnetic fields ($g = 2.5$)
656.28	C (Hα)	H I	6	Chromosphere
849.80		Ca II	1	Low chromosphere
854.21				
866.22				
868.86		Fe I	—	Photosphere, magnetic fields ($g = 1.7$)
1083.03		He I	200	High chromosphere
1281.8	Paschen β	H I	—	Chromosphere
1564.85		Fe I	—	Photosphere, magnetic fields ($g = 3$)
4665		CO	—	High photosphere
12,320		Mg I	—	High photosphere, magnetic fields ($g = 1$)

* J. Fraunhofer's original designation, if any.
† Atom or molecule responsible.
‡ Any temporal variability observed in integrated light.
§ Origin and diagnostic value of the line for solar physics research.

absorption of light by a particular atom or molecule. Most form in the Sun's atmosphere, although the Earth's telluric spectrum contributes lines of molecular oxygen (O_2), carbon monoxide (CO), and other molecules. Some lines such as Fraunhofer's C line in the red (hydrogen-alpha) can be seen with a pocket spectroscope. Powerful research instruments reveal millions of lines, most of which are weak and blended together in an almost inextricable tangle. The **table** lists some interesting Fraunhofer lines.

A spectrum line is caused by the absorption of photons of light that excite the atom from a lower to a higher energy level. Spontaneous decay back to the atom's lower level then follows, accompanied by the isotropic emission of light at the wavelength of the line. The result is a loss of light in the Sun-Earth direction. The probability of a given photon being absorbed depends on the quantum state of the atom.

The study of the Fraunhofer spectrum is the principal means of learning about physical conditions in the solar atmosphere. On the resolved solar disk, variations in line strength form point to point convey information about temperature, Doppler shifts of the lines reveal gas motions, and line splitting from the Zeeman effect maps magnetic fields. Because each line represents a chemical element, the composition of the solar atmosphere can be deduced. *See* DOPPLER EFFECT; SOLAR MAGNETIC FIELD; SUPERGRANULATION; ZEEMAN EFFECT.

Profiles. A high-dispersion spectrograph shows that Fraunhofer lines are not sharp; most have a gaussian profile with extended wings. This broadening results mainly from the Doppler motions associated with thermal and unresolved turbulence, together with collisional damping, wherein the atoms interact with their neighbors during the excitation-deexcitation process. Like line intensity and wavelength shifts, line profiles can yield further information about the solar surface, for example, the gas density and pressure. *See* SPECTROGRAPH.

Variability. The strength of the weak line of carbon at 538.0 nm is a sensitive indicator of surface temperature. The cores of chromospheric calcium H and K lines at 396.8 nm and 393.3 nm, which originate in active regions, correlate with the Sun's ultraviolet output. A variability of the Sun that might affect the Earth can be looked for by measuring the temporal variation of these lines in integrated sunlight. *See* ASTRONOMICAL SPECTROSCOPY; SOLAR CONSTANT; SUN.

William C. Livingston

Bibliography. A. Bhatnagar and W. C. Livingston, *Fundamentals of Solar Astronomy*, 2005; P. V. Foukal, *Solar Astrophysics*, 2d ed., 2004; L. Golub and J. M. Pasachoff, *Nearest Star: The Surprising Science of Our Sun*, 2001; K. H. Schatten and A. Arking (eds.), *Climate Impact of Solar Variability*, 1990.

Free-electron theory of metals

The treatment of a metal as containing a gas of electrons completely free to move within it. The theory was originally proposed in 1900 to describe and correlate the electrical and thermal properties of metals. Later, quantum mechanics became the basis for the theory of most of the general properties of simple metals such as sodium, with one free electron per atom, magnesium with two, and aluminum with three. Transition metals, such as iron, have partially

filled electronic d states and are not treated by the free-electron model.

Development. Three years after J. J. Thomson's 1897 discovery of the electron, P. Drude suggested that the transport properties of metals might be understood by assuming that their electrons are free and in thermal equilibrium with their atoms. This theory was made more quantitative by H. A. Lorentz. Assuming that the mean free path of electrons was limited by collisions, he was able to derive Ohm's law for the electrical conductivity and obtain the ratio of thermal to electrical conductivity in excellent agreement with experiment. This ratio, divided by the absolute temperature, is called the Wiedemann-Franz ratio and had been observed to be universal 50 years earlier. *See* CONDUCTION (ELECTRICITY); CONDUCTION (HEAT); ELECTRICAL CONDUCTIVITY OF METALS; KINETIC THEORY OF MATTER; MEAN FREE PATH; OHM'S LAW; THERMAL CONDUCTION IN SOLIDS; WIEDEMANN-FRANZ LAW.

The theory, however, had two major shortcomings. First, it predicted a large component of the specific heat of a metal, not present in insulators, which was not observed. Second, comparison of the theory with experiment indicated that the mean free path of the electrons became extremely large at low temperatures; the model offered no justification.

In 1928 A. Sommerfeld revised Lorentz's treatment by using quantum statistics, which removed the difficulty of the specific heat without losing the successful description of transport properties. The resulting theory remains the basis for the understanding of most transport properties of metals and semiconductors. At about the same time, W. V. Houston and F. Bloch solved the quantum-mechanical wave equation for electrons in a regular periodic structure, finding that they could indeed have arbitrarily large mean free paths if there were no defects in the periodicity, thereby putting the free-electron theory on a firm basis. *See* BAND THEORY OF SOLIDS; BLOCH THEOREM; FERMI-DIRAC STATISTICS; SPECIFIC HEAT OF SOLIDS; STATISTICAL MECHANICS.

Equilibrium properties. In the classical theory of Drude and Lorentz, the state of an electron can be specified by giving its position $\mathbf{r}$ and momentum $\mathbf{p}$. The corresponding state-space, the six-dimensional $\mathbf{r}$, $\mathbf{p}$ space, is called phase space. For the statistical analysis of the electron gas, it was convenient to divide the phase space into cells and specify the number of electrons in each cell as a function of time. In quantum theory the electrons are also treated as waves, satisfying appropriate conditions at the boundaries of the metal. This limits the wavelengths λ, and therefore the momentum values $p = h/\lambda$, where h is Planck's constant. It is easily shown that there is one allowed state for each cell of volume h^3 in phase space. The Pauli exclusion principle allows each state to contain only a single electron with each of the two spins. Thus in quantum theory the number of electrons in each cell is replaced by the probability $f(\mathbf{p}, \mathbf{r}, t)$ that the cell at the momentum $\mathbf{p}$ and position $\mathbf{r}$ is occupied by an electron of spin-up, usually equal to the probability of occupation by an electron of spin-down. The function $f(\mathbf{p}, \mathbf{r}t)$ is called the distribution function for the electrons. *See* EXCLUSION PRINCIPLE.

For a free-electron gas in thermal equilibrium, the distribution function does not depend upon time and is the same at all $\mathbf{r}$. Furthermore, it does not depend upon the direction of motion and is thus a function only of the magnitude of the momentum p, or the energy $E = p^2/2m$, where m is the electron mass. For a given number of electrons and a given total energy, there are many ways in which the electrons can be distributed in phase space. Statistical mechanics provides the distribution function $f_0(E)$, for a given number of electrons and given total energy, that can be obtained in the most ways. The subscript zero indicates the equilibrium distribution, which is the Fermi-Dirac distribution function given by Eq. (1). Here E_F is called the Fermi energy

$$f_0(E) = \frac{1}{e^{(E-E_F)/k_BT} + 1} \tag{1}$$

and for any system takes a value giving the correct total number of electrons; k_B is the Boltzmann constant; and the temperature T determines the total energy.

If the number of free electrons is very small, E_F is far below the energy of the lowest state, $E = p^2/2m = 0$, and the exponential is so large that the number one in the denominator can be neglected, and the classical Boltzmann distribution proportional to e^{-E/k_BT} is obtained, corresponding to Drude's theory. However, in real metals the density of electrons is so high that states must be occupied up to energies of 5–10 eV, whereas k_BT at room temperature is 0.025 eV, and the opposite limit is appropriate. Then E_F becomes of the order of 5 or 10 eV; $f_0(E)$ is near one for E up to that energy and then drops, over a range of a few hundredths of an electronvolt, to approximately zero. Consistent with the Pauli exclusion principle, almost all levels are occupied up to about the Fermi energy, and almost none above. Changing the temperature has only a small effect on the total energy; the large classical contribution, $^3/_2\, k_B$ per electron, to the specific heat is eliminated. In fact, only those electrons near the Fermi energy, a fraction of order k_BT/E_F, can take up thermal energy; the specific heat is $\pi^2 k^2_B T/2E_F$ per electron (a $\pi^2/3$ factor was derived from the full analysis), consistent with experiment. *See* BOLTZMANN STATISTICS.

An understanding of magnetic susceptibility in free-electron theory requires quantum mechanics. The equilibrium distribution for electrons in the presence of a magnetic field is needed. However, since the field deflects electrons in classical physics without energy change, the equilibrium distribution is unaffected; the energy is independent of field and the susceptibility is zero. In quantum theory the deflected electrons form orbits, which are quantized as are orbits in the atom. When these quantized levels are statistically occupied, the total energy does depend upon the field and a diamagnetic susceptibility is predicted. This contribution, along with a contribution from the partial alignment of electron spins,

describes the observed susceptibility of metals. *See* DIAMAGNETISM; MAGNETIC SUSCEPTIBILITY.

Transport theory. When electric fields are applied, the electrons are no longer in equilibrium. For a steady voltage applied to a long wire, the distribution function is still independent of time and position along the wire, but depends upon the direction of the electron momentum, $f(\mathbf{p})$. In free-electron theory the distribution function can be determined by following $f[\mathbf{p}(t)]$ as the momentum changes under the influence of the electric field $\mathscr{E}$ according to Eq. (2), where $-e$ is the electronic charge. The proba-

$$\frac{d\mathbf{p}}{dt} = -e\mathscr{E} \tag{2}$$

bility of the occupation of the state would not change with time except for electron collisions which can scatter the electron to another momentum, or scatter another electron to this momentum; thus Eq. (3) holds. The distribution function is assumed

$$\frac{df[\mathbf{p}(t)]}{dt} = \left.\frac{\partial f(\mathbf{p})}{\partial t}\right|_{\text{collisions}} \tag{3}$$

to be close to the equilibrium distribution written in Eq. (4), where $f_1(\mathbf{p})$ is small and proportional to the electric field $\mathscr{E}$.

$$f(\mathbf{p}) = \frac{f_0 p^2}{2m} + f_1(\mathbf{p}) \tag{4}$$

Thus evaluation of the left side of Eq. (3) in terms of patial derivatives produces a term $df_0(E)/dE$ $(dF/d\mathbf{p}) \cdot (-e\mathscr{E})$ and a term proportional to f_1 and to $\mathscr{E}$, and therefore negligible. The collisions are expected to relax any deviation $f_1(\mathbf{p})$ from the equilibrium distribution back toward zero, more rapidly with larger deviation. Then the right side of Eq. (3) is approximated by $-f_1(\mathbf{p})/\tau$. This is called the relaxation-time approximation, and τ is the relaxation time, usually assumed to be independent of $\mathbf{p}$, $\mathbf{r}$, and t. Then Eq. (5) may be used to solve for $f_1(\mathbf{p})$,

$$f_1(p) = e\mathbf{v} \cdot \mathscr{E}\tau \frac{df_0(E)}{dE} \tag{5}$$

noting that $dE/d\mathbf{p}$ is the electron velocity $\mathbf{v}$. *See* RELAXATION TIME OF ELECTRONS.

This analysis is the derivation and solution of the linearized Boltzmann equation in the relaxation-time approximation. The distribution having been found, the current contribution $-e\mathbf{v}$ can be summed for the state in each cell, multiplied by the probability that it is occupied, and by two for spin, and divided by the volume to obtain the current density. The sum over cells is written in Eq. (6) as an integral over phase

$$\mathbf{j} = \frac{2}{h^3\Omega}\Omega \int d^3p\,(-e\mathbf{v})f_1(p) = \frac{Ne^2\tau}{m}\mathscr{E} \tag{6}$$

space, $(1/h^3)\Omega\int d^3p$, where Ω is the volume. The last step requires a partial integration to obtain a form like that for the electron density N. The proportionality factor $Ne^2\tau/m$ is the predicted conductivity σ and is the same as that obtained for the Boltzmann distribution assumed by Drude and Lorentz.

In a similar way a temperature gradient is accommodated by allowing the T in Eq. (1) to vary along the wire and summing the flow of energy $\mathbf{v}p^2/2m$ for each occupied state to obtain the heat current. With an electric field strong enough to prevent any electric current, the heat current yields the thermal conductivity, as in Eq. (7). The Wiedemann-Franz ratio is a

$$\kappa = \frac{\pi^2}{3}\frac{k_B^2}{e^2}T\sigma \tag{7}$$

constant, $\kappa/(T\sigma) = \pi^2 k_B^2/(3e^2)$, as predicted by the classical theory, but with a different value; the factor $\pi^2/3$ arises because only electrons near the Fermi energy contribute.

The field that arises from the temperature gradient is a thermoelectric effect. It gives rise to the voltage measured from a thermocouple and is related thermodynamically to the Peltier heat, which is generated or absorbed when current passes through a thermocouple. These equilibrium and transport properties constitute the essence of the free-electron theory of metals. The most drastic approximation may be the relaxation-time approximation, but within free-electron theory more accurate approximations are possible for the scattering term. *See* THERMOELECTRICITY.

Electron-lattice interactions. The free-electron theory neglects the interactions between the electrons and the atoms making up the metallic crystal, except for the effects of collisions in the transport theory. This was first justified by applying Floquet's theorem to the quantum-mechanical wave equation in a perfectly periodic lattice to show that the solutions can be taken in the form of perfectly periodic functions of position $u(\mathbf{r})$, modulated by free-electron waves. (The modulation is of the form of the free-electron wave function, $e^{i2\pi\mathbf{p}\cdot\mathbf{r}/h}/\sqrt{\Omega}$.) These waves indeed propagate through the perfect crystal without scattering, although the presence of a $u(\mathbf{r})$ that is not constant can modify the velocity of propagation. Such corrections could be incorporated in the free-electron theory by replacing the electron mass by an effective mass m^*, but in fact the corrections are so small in the simple metals that it is ordinarily adequate to retain the true electron mass. *See* CRYSTAL STRUCTURE.

Then it is only the deviations from periodicity, such as impurity atoms or lattice vibrations, which produce the scattering of electrons assumed in transport theory. It is then to be expected that in pure materials at low temperatures the mean free paths become extremely long. As expected, the resistivity, $\rho = 1/\sigma$, is proportional to concentration of impurities. At elevated temperatures where the lattice-vibration amplitudes are proportional to $\sqrt{k_BT}$, the scattering rate is predicted to be proportional to temperature, as observed. Thus these interactions give significance to the parameters of free-electron theory, and do not invalidate it. *See* CRYSTAL DEFECTS; ELECTRICAL RESISTIVITY; LATTICE VIBRATIONS.

It was recognized in the 1950s that the effects of the periodic lattice on the electron energies are

remarkably small. The recognition was based on the observation that the momentum values $\mathbf{p}$ for the electrons that had the Fermi energy differed very little from those of the sphere in wavenumber space, $p^2/2m = E_F$, which free electrons would have. This surface in wavenumber space is called the Fermi surface. Although the potential from each atom felt by the electron is strong, its effect, called the pseudopotential, is weak. The pseudopotential is in fact so weak that it can be said only to affect electrons that satisfy the conditions for diffraction. The corresponding restructuring of the Fermi surface by diffraction leads to the nearly free-electron Fermi surfaces, which were accurately confirmed experimentally. From the point of view of free-electron theory, there may be a small number of electrons at any moment involved with diffraction, but the overwhelming majority hardly experience any potential at all and are therefore truly free. *See* FERMI SURFACE.

Interactions among electrons. Even in the context of a free-electron gas, there are strong Coulomb interactions between electrons which are frequently neglected in the free-electron theory of metals. This neglect was justified in the late 1950s by L. D. Landau, who asserted that, even with strong electron–electron interactions, there is a one-to-one correspondence between the excited states, called quasiparticle states, of the real system and the one-electron excitations from the ground state (the $T = 0$ equilibrium state) of the noninteracting electron gas. Thus, all of the formulation given here for free-electron theory still follows, but perhaps with modifications of parameters such as mass. Subsequent theory indicates that indeed these modification due to the electron–electron interaction are extremely small for the low-energy excitations present in thermal equilibrium, and so again the simplest theory succeeds for many properties, although substantial modifications are required for the higher-energy excitations caused by light. There are additional corrections, which are much larger than those from the electron–electron interaction, arising from the interaction between electrons and phonons, the quantum-mechanical term for lattice vibrations. In many metals these vibrations reduce the electron velocities by factors of as much as 2, increasing the electronic specific heat although they turn out not to modify the conductivity itself. *See* PHONON.

Another feature of the electron–phonon interaction is a resulting interaction among electrons, which is attractive and tends to cancel or exceed the repulsive electron–electron interaction. At low temperatures the net attraction binds electrons in pairs in a superconducting state. The theory of J. Bardeen, J. R. Schrieffer, and L. N. Cooper (the BCS theory of superconductivity), which first explained this phenomenon, is also a free-electron theory, but assumes that the free electrons have such a net attractive interaction. In contrast, it is generally believed that the high-temperature superconductors discovered in 1986 are very far from free-electron in character, and most workers do not believe that phonons are primarily responsible for the attractive interaction. *See* SOLID-STATE PHYSICS; SUPERCONDUCTIVITY.

Walter A. Harrison

Bibliography. J. A. Alonso and N. H. March, *Electrons in Metals and Alloys*, 1989; N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; W. A. Harrison, *Solid State Theory*, 1970, reprint 1980; C. M. Hurd, *Electrons in Metals*, 1975, reprint 1982; G. Lehman and P. Ziesche, *Electronic Properties of Metals*, 1991.

Free energy

A term in thermodynamics which in different treatments may designate either of two functions defined in terms of the internal energy U or enthalpy H, and the temperature-entropy product TS.

The function $A = U - TS$ is called the Helmholtz energy (formerly the work function or the Helmholtz free energy). The analogous function $G = H - TS$ is called the Gibbs energy (formerly the Gibbs free energy, or simply the free energy). *See* WORK FUNCTION (THERMODYNAMICS).

Theory. For a closed system (no transfer of matter across its boundaries), the work which can be done in a reversible isothermal process is given by the series shown in Eq. (1). For these conditions, $T\Delta S$

$$\begin{aligned} W_{\text{rev}} &= -\Delta A = -\Delta(U - TS) \\ &= -(\Delta U - T\Delta S) \end{aligned} \tag{1}$$

represents the heat given up to the surroundings. Should the process be exothermal, $T\Delta S < 0$, then actual work done on the surroundings is less than the decrease in the internal energy of the system. The quantity $(\Delta U - T\Delta S)$ can then be thought of as a change in free energy, that is, as that part of the internal energy change which can be converted into work under the specified conditions. This then is the origin of the name free energy. Such an interpretation of thermodynamic quantities can be misleading, however; for the case in which $T\Delta S$ is positive, Eq. (1) shows that the decrease in "free" energy is greater than the decrease in internal energy. *See* CHEMICAL THERMODYNAMICS.

For constant temperature and pressure in a reversible process the decrease in the Gibbs energy G for the system again corresponds to a free-energy change in the above sense, since it is equal to the work which can be done by the closed system other than that associated with its change in volume ΔV under the given constant pressure P. The relations shown in Eq. (2) can be formed since $\Delta H = \Delta U + P\Delta V$.

$$\begin{aligned} \Delta G &= -(\Delta H - T\Delta S) = W_{\text{net}} \\ &= W_{\text{rev}} - P\Delta V \end{aligned} \tag{2}$$

Each of these free-energy functions is an extensive property of the state of the thermodynamic system. For a specified change in state, both ΔA and ΔG are independent of the path by which the change is accomplished. Only changes in these functions can be measured, not values for a single state.

The thermodynamic criteria for reversibility, irreversibility, and equilibrium for processes in closed systems at constant temperature and pressure are expressed naturally in terms of the function G. For any infinitesimal process at constant temperature and pressure, $-dG \geqq \delta w_{\text{net}}$. If δw_{net} is never negative, that is, if the surroundings do no net work on the system, then the change dG must be negative or zero. For a reversible differential process, $-dG > \delta w_{\text{net}}$; for an irreversible process, $-dG > \delta w_{\text{net}}$. The Gibbs energy G thus decreases to a minimum value characteristic of the equilibrium state at the given temperature and pressure. At equilibrium, $dG = 0$ for any differential process taking place, for example, an infinitesimal change in the degree of completion of a chemical reaction. A parallel role is played by the Helmholtz energy A for conditions of constant temperature and volume. Because temperature and pressure constitute more convenient working variables than temperature and volume, the Gibbs energy is more common in thermodynamics.

Partial molal quantities. For a particular homogeneous phase in the absence of surface, gravitational, and magnetic forces, the Gibbs energy G depends on the amount of constituents present, the temperature T, and the pressure P. Let Ω represent the total number of constituents, n_i the amount of constituent i, and designate by subscript n constant composition, by subscript n_j constancy of the amounts of all constituents except n_i; then Eq. (3) is formed.

$$dG(T,P,n_1,\ldots,n_\Omega) = \left(\frac{\partial G}{\partial T}\right)_{P,n} dT + \left(\frac{\partial G}{\partial P}\right)_{T,n} dP + \sum_{i=1}^{\Omega}\left(\frac{\partial G}{\partial n_i}\right)_{T,P,n_j} dn_i \qquad (3)$$

In Eq. (3) the term $(\partial G/\partial n_j)_{T,P,nj}$ is the chemical potential μ_i of the ith constituent. It is identical to the partial molal Gibbs energy $\overline{G}_i$. It then follows that Eq. (4) holds.

$$dG = -S\,dT + V\,dP + \sum_{i=1}^{\Omega}\mu_i\,dn_i \qquad (4)$$

See SOLUTION.

Because the chemical potentials at constant T, P are intensive variables whose values are fixed, like that of the density, by relative amounts of the various constituents present, and are independent of the total mass of the phase, this equation can be integrated for constant T, P and relative composition starting from $n_i = 0$ to obtain Eq. (5). This procedure yields Eq. (6). Consistency with the expression for dG in Eq. (4) requires that Eq. (7) hold. This is

$$G(T,P,n_1,\ldots,n_\Omega) = \sum_{i=1}^{\Omega} n_i\mu_i \qquad (5)$$

$$dG = \sum_{i=1}^{\Omega}\mu_i\,dn_i + \sum_{i=1}^{\Omega} n_i\,d\mu_i \qquad (6)$$

$$S\,dT - V\,dP + \sum_{i=1}^{\Omega} n_i\,d\mu_i = 0 \qquad (7)$$

the Gibbs-Duhem equation. For constant temperature and pressure, this relation imposes a condition on the composition variation of the set of chemical potentials.

Heterogeneous systems. The Gibbs energy of a closed, heterogeneous system is the sum of the Gibbs energies of its various phases. In the absence of such a constraint as provided by the subdivision of the system by a rigid, semipermeable membrane, the general thermodynamic criterion of equilibrium requires that the temperature and pressure be uniform throughout the system and that the chemical potential of each constituent have a common value for all phases in which it is present. Further, if any of the constituents can be formed from others, the chemical potentials of the reactants and products are related in accordance with the stoichiometry of the reaction equation. Thus, for the reaction in Eq. (8), at equilibrium Eq. (9) can be formed. Expressing each

$$\mathrm{A} + 2\mathrm{B} \rightleftharpoons 3\mathrm{C} + 4\mathrm{D} \qquad (8)$$

$$\mu_{\mathrm{A}} + 2\mu_{\mathrm{B}} = 3\mu_{\mathrm{C}} + 4\mu_{\mathrm{D}} \qquad (9)$$

chemical potential μ_i in terms of the standard value μ_i° and its associated activity term $RT \ln a_i$ results in Eq. (10). In Eq. (10) $\Delta_r G^\circ$ is called the standard Gibbs

$$RT \ln\left(\frac{a_{\mathrm{C}}{}^3 a_{\mathrm{D}}{}^4}{a_{\mathrm{A}} a_{\mathrm{B}}{}^2}\right)_{\text{equil}} = -(3\mu_{\mathrm{C}}{}^\circ + 4\mu_{\mathrm{D}}{}^\circ - \mu_{\mathrm{A}}{}^\circ - 2\mu_{\mathrm{B}}{}^\circ) = -\Delta_r G^\circ \qquad (10)$$

energy for the reaction. Its value depends on the standard states chosen, but for a given temperature and pressure, it is a constant characteristic of the reaction involved. A true equilibrium constant K then results as shown by Eqs. (11) and (12). If the pressure for

$$K = \left(\frac{a_{\mathrm{C}}{}^3 a_{\mathrm{D}}{}^4}{a_{\mathrm{A}} a_{\mathrm{B}}{}^2}\right)_{\text{equil}} \qquad (11)$$

$$RT \ln K = -\Delta_r G^\circ \qquad (12)$$

each standard state is fixed and independent of the pressure of the reaction system, $\Delta_r G^\circ$ and hence K are functions of temperature only. This is the conventional approach in treating gas-phase equilibria, but not the approach that is used ordinarily for condensed phases. *See* ACTIVITY (THERMODYNAMICS).

Since the activities can be correlated with partial pressures or concentrations through fugacity coefficients or activity coefficients, this thermodynamic approach eliminates the uncertainties otherwise associated with equilibrium calculations that are based on the law of mass action. *See* CHEMICAL EQUILIBRIUM; FUGACITY.

The prediction of equilibrium constant then requires the calculation of $\Delta_r G^\circ$ for the reaction. The so-called third-law method involves calculation for the reaction at 25°C (77°F) of the value of $\Delta_r H^\circ$, the standard enthalpy of reaction, from tabulated standard enthalpy of formation data and of $\Delta_r S^\circ$ from tabulated third-law entropies. These are combined in the sense of $\Delta_r G^\circ = \Delta_r H^\circ - T\Delta_r S^\circ$ to permit calculation of the equilibrium constant for 25°C (77°F).

This in turn is used for evaluation of the integration constant in the integration of the relation in Eq. (13).

$$\frac{d \ln K}{dT} = \frac{\Delta_r H^\circ}{RT^2} \tag{13}$$

The integration requires expression of $\Delta_r H^\circ$ as a function of temperature, which necessitates a knowledge of the heat capacities $C^\circ_{P(i)}$ for the various reactants over the temperature range involved.

Alternatively, if values of the free-energy function $(G^\circ - H^\circ{}_{298K}/T)$ are available, either from experimental measurement or from statistical thermodynamical computations, they can be combined with the standard heat of reaction at 25°C (77°F) to give the desired result, Eq. (14).

$$\Delta_r G^\circ = \Delta_r \left(\frac{G^\circ H^\circ{}_{298K}}{T} \right) + \frac{\Delta_r H^\circ{}_{298K}}{T} \tag{14}$$

See ENTROPY; HEAT CAPACITY; THERMOCHEMISTRY.

Paul Bender

Bibliography. P. W. Atkins, *Physical Chemistry*, 1998; G. Carrington, *Basic Thermodynamics*, 4th ed., 1994; K. G. Denbigh, *Principles of Chemical Equilibrium*, 4th ed., 1981; K. S. Pitzer and L. Brewer, *Thermodynamics*, rev. ed., 1961; P. A. Rock, *Chemical Thermodynamics*, 1983; K. Wark, *Thermodynamics*, 5th ed., 1988.

Free fall

The accelerated motion toward the center of the Earth of a body acted on by the Earth's gravitational attraction and by no other force. If a body falls freely from rest near the surface of the Earth, it gains a velocity of approximately 9.8 m/s every second. Thus, the acceleration of gravity g equals 9.8 m/s^2 or 32.16 ft/s^2. This acceleration is independent of the mass or nature of the falling body. For short distances of free fall, the value of g may be considered constant. After t seconds the velocity v_t of a body falling from rest near the Earth is given by Eq. (1).

$$v_t = gt \tag{1}$$

If a falling body has an initial constant velocity in any direction, it retains that velocity if no other forces are present. If other forces are present, they may change the observed direction and rate of fall of the body, but they do not change the Earth's gravitational pull; therefore a body may still be thought of as freely "falling" even though the resultant observed motion is upward. In the case of an initial downward velocity v_0, the downward velocity v_t after t seconds is given by Eq. (2). In the case of a body projected

$$v_t \text{ (down)} = v_0 + gt \tag{2}$$

upward with an initial velocity v_0, the resultant upward velocity after $t$seconds is given by Eq. (3), and

$$v_t \text{ (up)} = v_0 - gt \tag{3}$$

the body loses upward velocity; that is, it is accelerated downward at the rate of free fall, just as it would be if it were falling. When its resultant upward velocity has been reduced to zero, the rising body has reached its maximum height.

For a body falling a very large distance from the Earth, the acceleration of gravity can no longer be considered constant. According to Newton's law of gravitation, the force between any two bodies varies inversely with the square of the distance between them; therefore with increasing distance between any body and the Earth, the acceleration of the body toward the Earth decreases rapidly. The final velocity v_f, attained when a body falls freely from a very large distance h, involves integration of the force equation as a function of distance giving as a result Eq. (4), where R is the radius of the Earth.

$$v_f = \sqrt{\frac{2ghR^2}{R^2 + Rh}} \tag{4}$$

In free fall from an infinite distance to the surface of the Earth, Eq. (4) reduces to $v_f = \sqrt{2gR}$, which gives a numerical value of 11.3 km/s or 7 mi/s. This is consequently the "escape velocity," the initial upward velocity for a rising body to completely overcome the Earth's attraction.

Because of the independent action of the forces involved, a ball thrown horizontally or a projectile fired horizontally with velocity v will be accelerated downward at the same rate as a body falling from rest, regardless of the horizontal motion (**Fig. 1**). If an observer could travel with the same initial horizontal velocity as the ball or the projectile, the observer would remain above it and in the observer's frame of reference the observed motion would be only that of free fall downward.

At a sufficiently large horizontal velocity, a projectile would fall from the horizontal only at the same rate that the surface of the Earth curves away beneath it (**Fig. 2**). The projectile would thus remain at the same elevation above the Earth and in effect become an earth satellite. The critical horizontal velocity v_b at which this would occur is that for which the required inward acceleration v^2/r for a body to

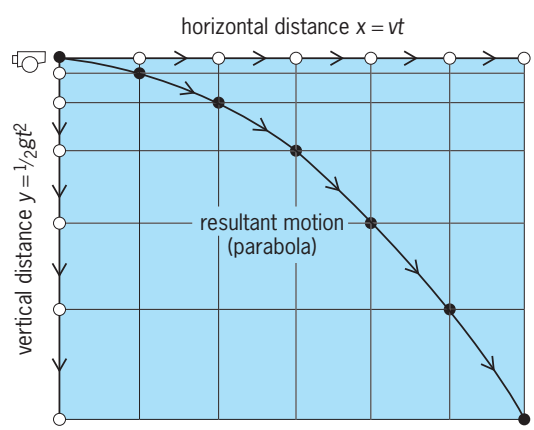

Fig. 1. Path of projectile fired horizontally, illustrating independence of horizontal and vertical components of motion.

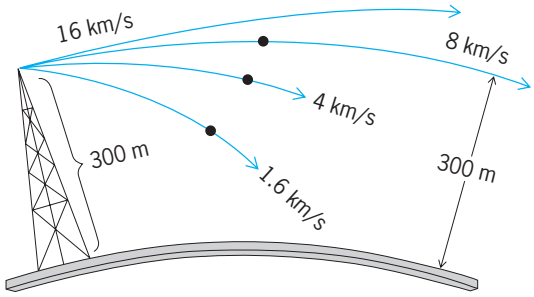

Fig. 2. Paths of projectiles fired horizontally at high velocities over the curved surface of the Earth. A projectile fired horizontally with a velocity of 8 km/s (4.8 mi/s) would circle the Earth like a satellite at constant altitude. 1 m = 3.3 ft; 1 km = 0.6 mi.

be in uniform circular motion at velocity v about a circle of radius r is exactly supplied by the free-fall acceleration of gravity g at that elevation. Thus, for the special case of horizontal motion near the Earth's surface, where $v_b = v$ and $r = R$ (the radius of the Earth), the required orbital velocity for the projectile to go into satellite motion is given by Eq. (5), and is

$$v_b = \sqrt{gR} \tag{5}$$

close to 8 km/s or 5 mi/s. At a higher velocity the satellite would rise to a higher orbit.

It was Newton's perception that for the Moon to be a satellite of the Earth the Moon must be in continual free fall toward the Earth which led him to propose his universal law of gravitation. *See* EARTH, GRAVITY FIELD OF; GRAVITATION.

Rogers D. Rusk

Free radical

Any molecule or atom that possesses one unpaired electron. This definition does not include transition- metal ions or molecules with more than one unpaired electron, such as dioxygen (O_2). Most chemists accept this definition; however, some spectroscopists employ a looser definition, in which they define any transient species (atom, molecule, or ion) in the gas phase to be a free radical.

Free radicals can be very reactive chemically (for example, the methyl radical, $CH_3\cdot$) or they can be very stable entities (for example, nitric oxide, NO). Often a free radical is designated by a dot (·). The modifier "free" probably arose because organic chemists often used to refer to a substituent (for example, hydroxyl or carboxyl) as a radical. However, this usage of the term radical is no longer prevalent, and increasingly the terms radical and free radical are used synonymously.

The term biradical is used for a molecule that contains two unpaired electrons so that the two are clearly confined to different parts of the molecule and do not interact significantly. That is, a biradical corresponds to two free radicals that are chemically bonded to each other. If the two unpaired electrons interact significantly, the entity is referred to as a triplet state. *See* TRIPLET STATE.

Free radicals were frequently, and incorrectly, postulated throughout the nineteenth century. Avogadro's hypothesis was not taken too seriously by the early organic chemists, and substances such as C_2H_6 and C_4H_{10} were frequently described as CH_3 and C_2H_5, respectively. However, by the end of that century this situation had been cleared up, and the "impossibility" of the independent existence of free radicals appeared to be well established. This situation was completely upset by Moses Gomberg's discovery of the triphenyl methyl radical during an attempt to synthesize hexaphenylethane. He found that the solution, which was expected to be colorless, was in fact yellow. Furthermore, on heating, the yellow color intensified. Also, molecular-weight determinations gave values lower than expected, and the new species was very reactive toward O_2, I_2, and NO at room temperature. This evidence provided strong support for the hypothesis that hexaphenylethane in solution is dissociated according to reaction (1).

$$(C_6H_5)_3C{-}C(C_6H_5)_3 \rightleftharpoons 2(C_6H_5)_3C\cdot \tag{1}$$

In spite of this evidence, Gomberg's claim met with years of doubt and denial. Since then, many similar radicals have been discovered, and such entities are now freely postulated in organic reaction mechanisms. *See* ORGANIC REACTION MECHANISM.

In 1929, the first evidence for very reactive free radicals was obtained by F. Paneth. He observed that when tetramethyllead [$Pb(CH_3)_4$] vapor, with a carrier gas, was passed very rapidly over a heated tube, a lead mirror was formed some distance downstream from the point of heating, and ethane was detected as the principal gaseous product. If the tube was then heated upstream of the mirror with tetramethyllead passing through, the old mirror was observed to disappear. This was interpreted in terms of indirect evidence that methyl radicals ($CH_3\cdot$) were present as transient free radicals in the gas phase.

Free radicals can be grouped into three major classes: atoms (for example, H·, F·, and Cl·), inorganic radicals (for example, OH·, CN·, NO·, and $ClO_3\cdot$), and organic radicals [for example, methyl ($CH_3\cdot$), allyl ($H_2C{-}CH\cdot{-}CH_2$), and phenyl ($C_6H_5\cdot$)]. Such radicals are of great importance because they have been shown to appear as intermediates in both thermal and photochemical reactions. Free radicals are also known to initiate and propagate polymerization and combustion reactions. *See* COMBUSTION; PHOTOCHEMISTRY; POLYMERIZATION; REACTIVE INTERMEDIATES.

Production. Most often, free radicals are formed by the homolytic rupture of a bond in a stable molecule with the production of two fragments, each with an unpaired electron. The resulting free radicals may participate in further reactions or may combine to reform the original compound. In the gas phase, equilibria such as reaction (2) may be established,

$$R_1{-}R_2 \rightleftharpoons R_1\cdot + R_2\cdot \tag{2}$$

especially at elevated temperatures. Because in most cases recombination [the leftward reaction of the

equilibrium shown as reaction (2)] occurs at nearly every collision of R_1 and R_2, the composition of the reaction mixture under ordinary circumstances indicates only a minute amount of decomposition into radicals. Even though radical partial pressures may be less than 10^{-6} torr (1.33×10^{-4} pascal) and their lifetimes less than 1 millisecond, these radicals do play an important role in reaction kinetics. The transitory existence of such radicals has been confirmed by spectroscopic study. *See* CHEMICAL DYNAMICS.

Free radicals may be generated by a wide variety of methods, such as thermal decomposition, electric discharge, microwave discharge, photochemical decomposition, electrolysis at an electrode, rapid mixing of two reactants, and radiolysis with gamma-ray, x-ray, or electron-beam irradiation. In thermal methods, a stable molecule is decomposed at an elevated temperature. In exceptional circumstances, the dissociation into radicals at equilibrium may be considerable. For example, hydrogen atoms may be produced by heating hydrogen gas (H_2) to a very high temperature, as in reaction (3). At 1900 K (2960°F),

$$H_2 \rightleftharpoons 2H\cdot \tag{3}$$

this equilibrium corresponds to 1% dissociation into atoms when the pressure is 1 atm (100 kilopascals). In a few cases, substances in solution are considerably dissociated into radicals at room temperature. It is then possible to obtain radicals of apparent long life in high concentrations. Examples are hexaphenylethane [reaction (1)], which in a 2–3% solution in benzene at 5°C (41°F) is about 3% dissociated into triphenylmethyl radicals; hexa(*p*-biphenyl) ethane [$(C_6H_5—C_6H_4)_3C—C(C_6H_5—C_6H_4)_3$] is virtually 100% dissociated under similar circumstances. However, this type of equilibrium is the exception rather than the rule, as most thermal decomposition reactions are irreversible under the conditions of the reaction. Most gaseous organic substances decompose wholly or in part by a mechanism involving an initial split into radicals. *See* PYROLYSIS.

Free radicals may also be produced by passing a gas through an electrical discharge at high speed. Atoms, such as hydrogen, oxygen, or nitrogen, are often produced in this manner and then subjected to further chemical reactions which often produce other radicals. A gentler method is to use a microwave discharge.

Photochemical methods of producing free radicals are very common and have been studied extensively. When subjected to ultraviolet light with wavelengths less than 200 nanometers, most gaseous organic molecules decompose photochemically into free radicals. Certain classes of molecules (for example, carbonyls) will react photochemically with ultraviolet light from 200 to 350 nm wavelength, while a few molecules, such as NO_2 and Cl_2, will decompose with visible light. When photolyzed in the range 250–310 nm, acetone splits, as in reaction (4). This

$$CH_3COCH_3 + h\nu \longrightarrow CH_3\cdot + CH_3CO\cdot \tag{4}$$

reaction has been one of the most frequently used sources of methyl and acetyl radicals in gas-phase photochemistry.

Chlorine decomposes on irradiation with visible light and near-ultraviolet light to give chlorine atoms [reaction (5)]. This is often the starting step in the

$$Cl_2 + h\nu \longrightarrow 2Cl\cdot \tag{5}$$

chlorination of hydrocarbons, reactions which are particularly important in the synthesis of herbicides and pesticides.

The free radical $NO_2\cdot$ decomposes via reaction (6)

$$NO_2\cdot + h\nu \longrightarrow NO\cdot + O \tag{6}$$

with visible and near-ultraviolet light. This reaction is of considerable importance in the chemistry of air pollution.

Hydrogen peroxide dissociates photochemically with ultraviolet light of wavelengths less than 300 nm according to reaction (7). This reaction is used industrially in advanced oxidation processes, which utilize powerful ultraviolet lamps (up to 30 kW each) to destroy organic pollutants in contaminated ground and industrial process waters. *See* AIR POLLUTION; ATMOSPHERIC CHEMISTRY.

$$H_2O_2 \xrightarrow{h\nu} 2\cdot OH \tag{7}$$

A photochemical production of free radicals via an electron transfer reaction from a chlorophyll entity to a quinone is one of the primary photochemical reactions in photosynthesis as carried out by green plants, algae, and certain photosynthesis bacteria. *See* ELECTRON-TRANSFER REACTION; PHOTOCHEMISTRY; PHOTOSYNTHESIS.

The field of phototherapy has shown considerable promise in cancer therapy. Here the patient is injected with a dye (for example, hematoporphyrine) that has a marked affinity for tumor cells. After a couple of days, when most of the dye has concentrated in the tumor, the patient is irradiated with red light, usually with a laser. This light is absorbed by the dye, raising it to an excited state. In this excited state, the dye generates free radicals that attack the tumor cells and destroy them. *See* ONCOLOGY.

Electrolysis often produces radicals at one of the electrodes. For instance, the electrolysis of anthracene in dimethylformamide, in the absence of oxygen and water, leads to the generation of the anthracene negative radical ion at the cathode and the positive radical ion at the anode. When an anthracene anion and cation diffuse together, there is enough energy to generate the excited anthracene singlet state. This is the basis of electroluminescence. *See* ELECTROLUMINESCENCE; ELECTROLYSIS.

Rapid mixing of two suitable reactants can lead to moderate steady-state concentrations of unstable radical intermediates. For example, rapid mixing of titanium trichloride ($TiCl_3$), and of hydrogen peroxide (H_2O_2), results in the formation of the OH· radical, presumably by reaction (8).

$$Ti^{3+} + H_2O_2 \longrightarrow Ti^{4+} + OH\cdot + OH^- \tag{8}$$

High-energy radiation such as gamma rays and x-rays, and high-energy particles given off by radioactive nuclei and high-energy electron beams can cause extensive disruption of molecules leading to production of radicals and ions in a process known as radiolysis. For example, the radiolysis of water produces hydrogen atoms (H·), hydroxyl radicals (OH·), and hydrated electrons ($e^-_{H_2O}$). Each of these radical intermediates can initiate many possible secondary reactions with solutes. Radiolysis is an important technique in cancer therapy. Pulsed radiolysis, as generated by very short pulses in an electron beam, has proven to be a very important technique in the study of the mechanism of radiolytic reactions. *See* CHARGED PARTICLE BEAMS; MOLECULAR MECHANICS.

Detection and estimation. The earliest methods of detection involved the chemical properties of the radicals, such as the Paneth method for the detection of methyl radicals. Later and more reliable methods used absorption spectroscopy, mass spectrometry, and especially electron paramagnetic resonance spectroscopy.

Certain chemicals have proved effective as radical traps. For example, if iodine is added to a gas-phase photochemical system, rapid reactions (9) and (10) occur. Under appropriate conditions, all radicals

$$R\cdot + I_2 \longrightarrow RI + I\cdot \qquad (9)$$

$$R\cdot + I\cdot \longrightarrow RI \qquad (10)$$

may be removed from the system and trapped as iodides. In solution, nitrones such as 5,5′-dimethyl-1-pyrroline-1-oxide (DMPO) have proved to be effective spin traps via reaction (11). The resulting

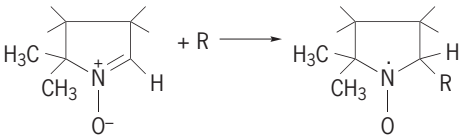

(11)

spin adduct is usually a stable nitroxide free radical which may be detected by electron paramagnetic resonance spectroscopy. In this manner, both quantitative and qualitative information can be obtained about the nature of the reactive radical R·. For example, hydroxyl radicals (OH·) have been detected by this method in the ultraviolet photolysis of aqueous suspensions of the anatase crystal form of titanium dioxide (TiO_2). Photolysis of such suspensions has been shown to catalyze the photodegradation of chlorohydrocarbon pollutants. *See* PHOTODEGRADATION.

Absorption spectroscopy is a simple method of detecting radicals in the gas phase and has been used extensively for many years. For example, the following transient radicals (under the spectroscopists' definition) have been identified spectroscopically: OH, NH, NH_2, CN, CF_2, C_2, CH, CHO, C_3, CH_3O, C_2H_5O, NCO, NCS, HNO, PH_2, and CH_3; many others have been identified with varying degrees of certainty. The knowledge of the absorption wavelengths of these radicals has permitted astronomers to detect many of these entities in interstellar space. The flash-photolysis technique, by which large amounts of light energy may be absorbed in a short period of time, has greatly advanced the detection of radicals by absorption spectroscopy and their involvement in radical mechanisms.

Another popular technique for radical detection is matrix isolation. Here the radicals are generated (for example, by photolysis) in the gas phase in an inert gas such as argon. Immediately after generation, the gas mixture is frozen onto a finger, a protusion into the gas stream that is maintained at a very low temperature (less than 20 K or −423°F). Alternatively, molecules may be isolated in a frozen argon matrix and then photolyzed to generate the radicals. Once isolated in the argon matrix by this method, the radical intermediate may be studied at will by using ultraviolet, visible, or infrared spectroscopy. *See* LUMINESCENCE ANALYSIS; MATRIX ISOLATION; SPECTROSCOPY.

An ingenious method for the detection of gas-phase radicals has been developed using mass spectrometry. A rapid-flow system is used in which radicals are produced thermally or protochemically. The gas stream is sampled through a leak into the mass spectrometer. An electron energy is selected so that radicals are ionized, but stable molecules are not dissociated. Any radical ions detected in the mass spectrometer must come from radicals already present in the gas mixture, and a direct and unequivocal detection of radical intermediates is possible. Under favorable circumstances, the method can be used for the quantitative estimation of radical concentrations. This technique has proved to be particularly valuable in the study of combustion reactions. *See* MASS SPECTROMETRY.

Electron paramagnetic resonance spectroscopy is by far the best technique available for the detection and characterization of free radicals in solids, liquids, or gases. The technique is very sensitive (10^{-9} mole per liter free radicals can be detected under favorable conditions) and, furthermore, the technique is specific to molecules that contain one or more unpaired electrons. Generally, each radical has a characteristic electron paramagnetic resonance spectrum which can be used both for identification and for deducing structural information about the radical. For example, the methyl radical $CH_3\cdot$ was shown to have a planar structure by analyzing its electron paramagnetic resonance spectrum, whereas the corresponding $CF_3\cdot$ radical was shown to have a pyramidal structure. Analysis of the electron paramagnetic resonance spectra of radicals such as the anthracene negative ion yields information concerning the distribution of the unpaired electron over the molecule. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY.

Free-radical mechanisms. In well-investigated reactions, it is frequently possible to establish that the reaction proceeds via a multistep mechanism involving free-radical intermediates. For example, a 50:50 mixture of chlorine and hydrogen will explode on being subjected to a flash from a photoflash unit. This

reaction is thought to proceed via the mechanism outlined in reactions (12)–(16). This is an example of

$$Cl_2 + h\nu \longrightarrow 2Cl\cdot \quad (12)$$

$$Cl\cdot + H_2 \longrightarrow HCl + H\cdot \quad (13)$$

$$H\cdot + Cl_2 \longrightarrow HCl + Cl\cdot \quad (14)$$

$$H\cdot + HCl \longrightarrow H_2 + Cl\cdot \quad (15)$$

$$Cl\cdot + X \longrightarrow \text{nonradical products} \quad (16)$$

a radical chain reaction since reactions (13) and (14) can cycle many times generating the product HCl before the chain is terminated by reaction (16) [X in reaction (16) is any substance which can remove Cl· radicals from the system]. The system explodes because reactions (13) and (14) are extremely fast and the reaction generates a considerable amount of heat, which further accelerates the reaction. Many thermal decomposition reactions are also thought to proceed via free-radical chain reactions. A typical example is the thermal decomposition of ethane, reactions (17)–(23). General mechanisms of this type,

$$C_2H_6 \longrightarrow 2CH_3\cdot \quad (17)$$

$$CH_3\cdot + C_2H_6 \longrightarrow CH_4 + C_2H_5\cdot \quad (18)$$

$$C_2H_5\cdot \longrightarrow C_2H_4 + H\cdot \quad (19)$$

$$H\cdot + C_2H_6 \longrightarrow C_2H_5\cdot + H_2 \quad (20)$$

$$H\cdot + C_2H_5\cdot \longrightarrow C_2H_6 \quad (21)$$

$$2C_2H_5\cdot \longrightarrow C_4H_{10} \quad (22)$$

$$2C_2H_5\cdot \longrightarrow C_2H_4 + C_2H_6 \quad (23)$$

the Rice mechanisms, have been well established in many cases by detailed investigation of the kinetic behavior. With the aid of computer simulation, it is possible to predict the behavior of very complex reactions over a wide range of conditions. This approach has been particularly valuable in the study of gas-phase chemistry in the upper atmosphere and in polluted air. *See* CHAIN REACTION (CHEMISTRY); COMPUTATIONAL CHEMISTRY. James R. Bolton

Bibliography. A. Carrington and B. Thrush (eds.), *The Spectroscopy of Molecular Ions*, 1988; B. Halliwell, *Free Radicals in Biology and Medicine*, 3d ed., 1999; H. Kurreck, *Electron Nuclear Double Resonance of Radicals in Solution*, 1988; W. A. Pryor, *Frontiers of Free Radical Chemistry*, 1980; J. A. Weil, J. R. Bolton, and J. E. Wertz *Electron Paramagnetic Resonance*, 1994.

Frequency (wave motion)

The number of times which sound pressure, electric field, or other quantities specifying a wave vary (or oscillate) from their equilibrium value through a complete cycle in a specified time. The most common unit of frequency is the hertz; 1 Hz is equal to 1 cycle per second. In one cycle there is a positive variation from equilibrium, a return to equilibrium, then a negative variation, and return to equilibrium. This relationship is often described in terms of the sine wave, and the frequency referred to is that of an equivalent sine-wave variation in the parameter under discussion. *See* SINE WAVE.

The time between consecutive variations or oscillations of the wave is the period T of the wave. The frequency f is the reciprocal of the period:

$$f = 1/T$$

Frequency is a convenient means for describing the various ranges of interest in wave motion. For example, audible sound is between approximately 20 and 20,000 Hz. Infrasonic frequencies are below approximately 20 Hz; sound having frequencies above the audible is termed ultrasonic. *See* INFRASOUND; SOUND; ULTRASONICS.

Ultrasonic waves having frequencies as high as 200 gigahertz have been generated and detected in solids and have been used to generate images (1 GHz = 10^9 Hz). Typical approximate frequency ranges for electromagnetic waves are: AM radio in the United States, 550 to 1700 kHz; FM radio, 88 to 108 MHz; Global Positioning System (GPS) microwaves, 1.2276 GHz and 1.57542 GHz; visible light, 4×10^{14} to 7.5×10^{14} Hz. *See* ELECTROMAGNETIC RADIATION; FREQUENCY MEASUREMENT; WAVE MOTION. William J. Galloway; Philip L. Marston

Frequency comb

A device that facilitates and greatly simplifies the measurement of very high frequencies, particularly in the optical range from 10^{14} to 10^{15} Hz. Extremely stable and reproducible optical frequencies f_{opt} in this range are produced by trapped-atom and trapped-ion atomic clocks. Frequency-comb technology is based on the train of short-duration pulses of optical radiation emitted at a stable and measurable repetition rate f_R of 100 MHz to 1 GHz by a femtosecond laser. Such a pulse train can be equivalently represented as a continuously produced "comb" of frequencies. The "teeth" of the comb are discrete phase-related frequencies spaced by f_R from one another, and the frequency of the mth tooth is mf_R, where m is an integer (see **illustration**). The frequencies produced typically span a rather narrow wavelength range of 30 nm to around 800 nm, but this span can be increased to a factor of 2 by passing the pulsed light through an optical fiber made from a highly nonlinear material, enabling a comb frequency to be matched to many different optical frequency standards. *See* ATOMIC CLOCK.

The frequencies of the comb as a whole are offset by f_0 because the optical frequency of the femtosecond laser in general differs by a frequency f_0 from an exact multiple of f_R. One of the comb frequencies, $mf_R + f_0$, can be mixed with f_{opt}, and the output from the mixing diode used as feedback to stabilize this comb frequency and confer on it the full frequency stability of f_{opt}. The frequency f_0 can be measured by frequency doubling a frequency in the left-hand

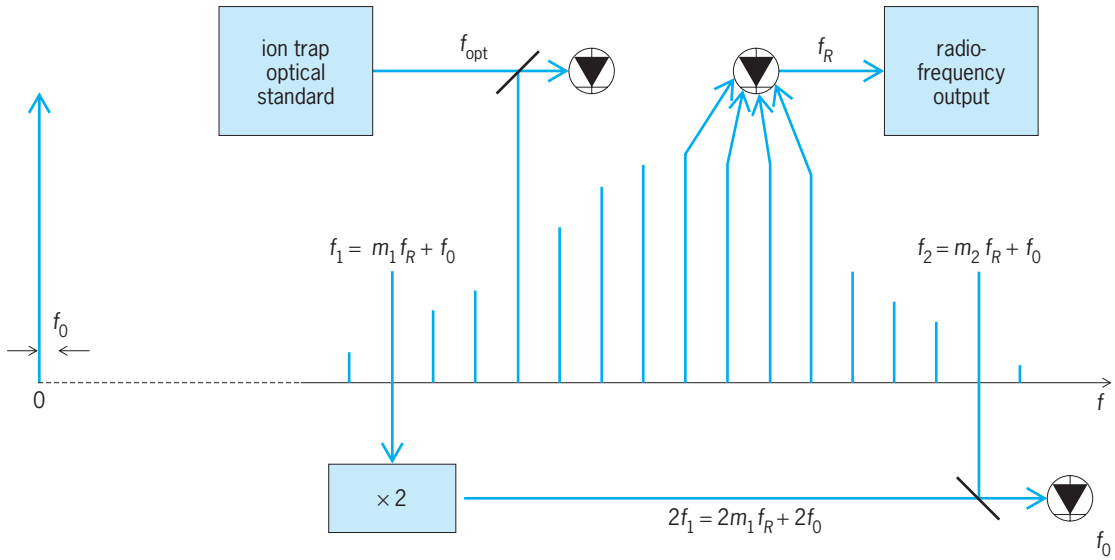

Operation of the frequency comb, illustrating the self-referencing technique used to measure the offset frequency f_0. (*Hugh Klein*)

(infrared) end of the comb (by passing the light through a nonlinear optical medium) to generate a frequency $2m_1f_R + 2f_0$, and mixing this frequency with that of a tooth of frequency $m_2f_R + f_0$ from the right-hand end of the comb. If the chosen tooth has $m_2 = 2m_1$, the beat frequency is f_0 and can be measured. *See* BEAT.

Provided the integer m is known for the tooth frequency nearest to f_{opt} (from an approximate knowledge of f_{opt} and f_R), f_R can be calculated as $f_{opt}/m - f_0/m$. In this way the long-term stability of optical frequency standards can be subdivided to a radio frequency f_R for extremely accurate frequency and time measurement. The fractional stability of f_R is the same as that of the optical frequency standard to which it is locked and this is now of the order of 1 in 10^{15}, equaling that of the very best cesium clocks. This stability will undoubtedly improve even further. Then it is anticipated that the SI second will be redefined in terms of an optical frequency standard. *See* FREQUENCY MEASUREMENT; PHYSICAL MEASUREMENT; TIME.

Bryan P. Kibble

Bibliography. *Proceedings of the First ESA Optical Clock Workshop*, 2005; F. Riehle, *Frequency Standards*, Wiley-VCH, 2004; Special issue: 50 years of atomic time-keeping: 1955 to 2005, *Metrologia*, 42(3):S1–S153, June 2005.

Frequency counter

An electronic instrument used to precisely measure the frequency of an input signal. Frequency counters are commonly used in laboratories, factories, and field environments to provide direct frequency measurements of various devices. The most common applications for frequency counters are measurement and characterization of oscillator and transmitter frequencies. *See* OSCILLATOR; RADIO TRANSMITTER.

There are several classes of frequency counters. Basic frequency counters provide measurement of frequency only. Universal counter-timers are two-channel instruments that provide measurement of frequency, period, phase, totalize (the total number of pulses generated by some type of event over the duration of an experiment), ratio (of frequencies on two channels), and time intervals such as pulse width or rise time. Microwave counters are an extension of basic frequency counters offering coverage of microwave frequency ranges to 40 GHz and beyond. *See* MICROWAVE.

Architectures. The three main architectures are conventional counting, reciprocal counting, and continuous counting. Conventional counting is the oldest and simplest but has the lowest performance and the least measurement flexibility. A conventional counter uses a simple register to count each cycle of the input signal during a 1-s measurement gate time. Since the count register contains a direct count of the number of cycles in one second (hertz), its contents can be sent to a digital display as the measurement result with no further processing.

Reciprocal counting is the most common architecture (see **illus.**). It provides improved performance and flexibility over conventional counting. The main gate is set by the user and determines the nominal time over which the measurement is to be made (measurement gate time). The input signal is applied to the input conditioning circuits and routed to the event counter. The event counter simply counts the exact number of cycles of the input signal during the measurement gate time. The reciprocal technique synchronizes the opening and closing of the main gate to the input signal. This provides increased precision but requires that the actual measurement gate time also be measured precisely. This is the purpose of the time counter. The time counter counts the number of cycles of the internal time base during the measurement gate time.

At the end of the measurement gate time, the values in the time and event counters are read by the arithmetic circuits. The arithmetic circuits then

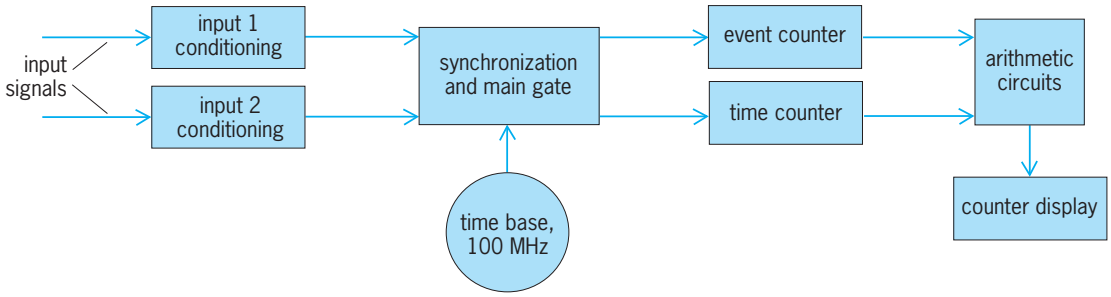

Block diagram of a reciprocal frequency counter.

compute frequency as the number of cycles per unit time. The event counter represents the number of cycles of the input signal counted, and the exact measurement time is given by the number of counts of the precision time base multiplied by the time-base period. The result is then converted to hertz and sent to the counter display.

Continuous counting is based on the reciprocal technique, but employs high-speed digital circuits to continuously sample the contents of the count registers. These continuous samples can be digitally processed to provide improved resolution to as many as 12 digits in 1 s of measurement time. Continuous counting is also employed in frequency and time-interval analyzers, which provide graphical displays of continuous frequency versus time.

Range extension. The typical upper range for a direct frequency counter is 200 MHz. Prescalers are often used to extend the range of frequency counters to 3 or 5 GHz. A prescaler effectively divides the input signal frequency down into the range of the digital counter. The prescaler is placed after the input conditioning section (see illus.). Counting is then performed as described above with the prescale factor multiplied back in by the arithmetic circuits to display the correct answer.

Harmonic heterodyne converters are typically used to extend the measurement range to 40 GHz and beyond. In its simplest form, the microwave input signal is mixed in a sampler with a frequency generated by a synthesizer to provide an intermediate frequency that is in the direct measurement range of the digital counter. Modern microwave frequency counters integrate the harmonic heterodyne converter into their design to automatically downconvert and measure microwave frequencies. *See* FREQUENCY MEASUREMENT; HETERODYNE PRINCIPLE.

Barry Dropping

Bibliography. C. F. Coombs, Jr., and G. D. Sasaki, *Electronic Instrument Handbook*, 3d ed., 1999.

Frequency divider

An electronic circuit that produces an output signal at a frequency that is an integral submultiple of the frequency of the input signal.

Several information-processing and transmitting techniques require frequency division. In television, for example, it is essential to maintain a precise relationship between the horizontal-scanning frequency and the vertical-scanning frequency. Frequency division can be easily accomplished in two ways, digital division and division by triggering a subharmonic.

Digital division. Many circuits are available to count pulses. A bistable or flip-flop circuit produces one output pulse for every two input pulses. By cascading successive flip-flops, any desired degree of division can be obtained. Division by powers of 2 can be achieved simply by monitoring the output of the proper stage in the cascade. Division by other numbers can be achieved by gating to obtain the proper set of flip-flop conditions. Division by 10, or powers of 10, is readily obtained by using decade rings. *See* DIGITAL COUNTER; GATE CIRCUIT.

Reduction of the input to pulses can be accomplished when needed by amplifying, clipping, and differentiating the original signal. To provide an output that is a desired shape, the signal from the frequency divider can be used to synchronize a multivibrator to obtain a square wave and additional power. The square wave can be fed into a tuned circuit or filtered to obtain a sine wave, if that is needed. In many applications the input and output signals need no alteration. *See* WAVE-SHAPING CIRCUITS.

Subharmonic triggering. Any circuit which has a characteristic resonance responds to certain types of input energy by ringing, that is, by going through one or more cycles of electrical activity caused by the nature of the circuit rather than by the nature of the input. This characteristic can be used to accomplish frequency division, provided the input frequency does not vary over any extensive frequency range. A triggered multivibrator can be used in this manner. Either a monostable or a synchronized astable type will serve. (An astable type can be used because, while it is normally free-running with no stable state, it can be triggered to provide synchronized pulse rates slightly faster than it would otherwise generate.)

The component values of the frequency-dividing multivibrator are selected so that triggering cannot occur during a fixed number of cycles of the input signal. In operation, each time the multivibrator returns to the state where it is sensitive to an input trigger, the next cycle of the input signal initiates a new output pulse. Hence a new cycle of the multivibrator is initiated every n input cycles. The frequency of the resulting output signal is therefore an exact submultiple of the frequency of the input signal. Because the input may go through a large number of cycles for each output cycle, greater division per stage can be realized by this method than by digital division.

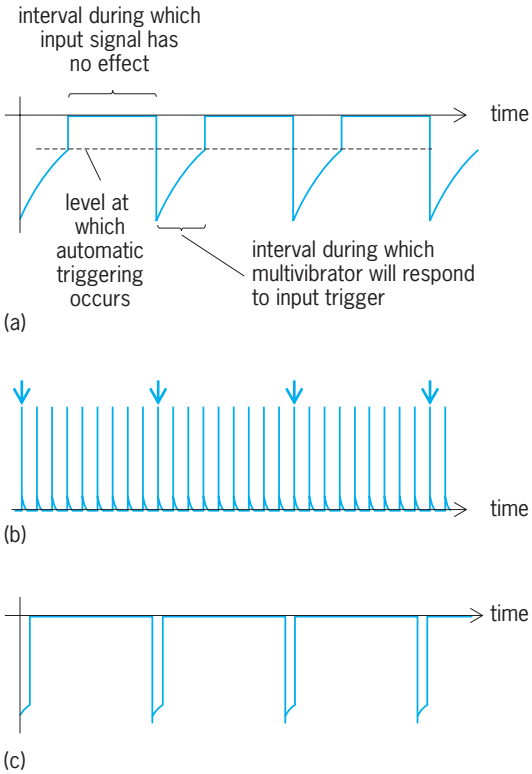

Waveform of frequency-dividing astable multivibrator. (*a*) Waveform at base terminal of free-running astable multivibrator, showing how multivibrator would perform if there were no input. (*b*) Input signals. Arrows indicate triggering signals. (*c*) Output signal, synchronized to every *n*th input cycle.

The input–output relationship can be most readily perceived when the input consists of short pulses (see **illus.**), but the effect is not dependent on the shapes of the input signal.

Other resonant or tuned circuits can be induced to supply submultiple outputs in a similar manner. *See* MULTIVIBRATOR; RESONANCE (ALTERNATING-CURRENT CIRCUITS).

William W. Snow

Frequency measurement

The determination of the number of cycles of a periodically varying quantity occurring in unit time. Many physical systems demonstrate cyclic behavior; that is, one or more of their properties vary in a characteristic fashion before returning to the initial value and then repeating the cycle. Examples are the angular positions of the planets and satellites in the solar system, the pressure in a cylinder in a reciprocating engine, and the heights and fields associated with surface, acoustic, and electromagnetic waves. The duration of a single cycle, the period, may vary widely, from 10^{-27} s for the electromagnetic field associated with a cosmic gamma ray to 10^{8} years for the rotation of a galaxy in space. The frequency, which is the inverse of the period, is the number of cycles, including fractions, occurring in unit time. The unit of frequency is the hertz (Hz), named after Heinrich Hertz, who investigated the nature of electromagnetic radiation. Measurement of the characteristic frequencies of a system, and their variation with time or under changing conditions, yields valuable information on its properties and behavior. Together with temperature and voltage, frequency ranks as one of the quantities most often measured in modern science and technology. *See* CELESTIAL MECHANICS; CERENKOV RADIATION; ELECTROMAGNETIC RADIATION; GAMMA-RAY ASTRONOMY; MILKY WAY GALAXY; WAVE MOTION.

Analog techniques. The measurement of an unknown frequency requires a standard producing a fixed, stable, and known frequency, and a system or technique for the comparison of the unknown frequency with the standard. In the past, a wide variety of analog techniques and material standards have been employed. An example is the use of a tuning fork to adjust a musical instrument, usually a piano. The combination of notes from the instrument and the resonating fork has an amplitude modulation at the difference between their frequencies, which is detected as a slow audible beat. Similarly, two sinusoidal electrical signals close in frequency applied to the X and Y plates of an oscilloscope produce an ellipse on the screen which rotates at the beat frequency. If this frequency is small, less than 1 Hz, it may be measured by timing the period with a stopwatch. The more complicated Lissajous figures are produced if the two frequencies are (nearly) harmonically related. *See* BEAT; LISSAJOUS FIGURES; TUNING FORK.

Analog frequency measurement techniques possessed two major disadvantages: The frequency of the standards depended upon the material properties and dimensions of critical components, which meant that they were prone to drift and affected by variations in the ambient temperature. In addition, optimum accuracy was achieved only when the unknown and standard frequencies were close or harmonically related.

Developments in electron-tube and, later, solid-state electronics improved matters. These included the quartz crystal oscillator, in which a thin slice of crystalline quartz acts as the resonant element in an electronic feedback circuit. As a result of the sharpness of the resonance and the stability of the properties of the quartz, this device provides a stable frequency in the range from 10 kHz to 100 MHz and remains the most common secondary frequency standard in use. In addition, a range of circuits were developed to generate more complex harmonic and subharmonic frequencies from a standard source. This led ultimately to the frequency synthesizer which, with an array of phase-locked loops, could be set to produce one of a very wide range of output frequencies. In use, however, it was still necessary to measure the beat or heterodyne frequency from the unknown frequency. *See* FEEDBACK CIRCUIT; FREQUENCY DIVIDER; FREQUENCY MULTIPLIER; OSCILLATOR; PHASE-LOCKED LOOPS; PIEZOELECTRICITY; QUARTZ CLOCK.

Digital techniques. Fast, inexpensive solid-state digital circuits have replaced analog frequency measurement techniques and many of their associated standards. The underlying principle of the digital

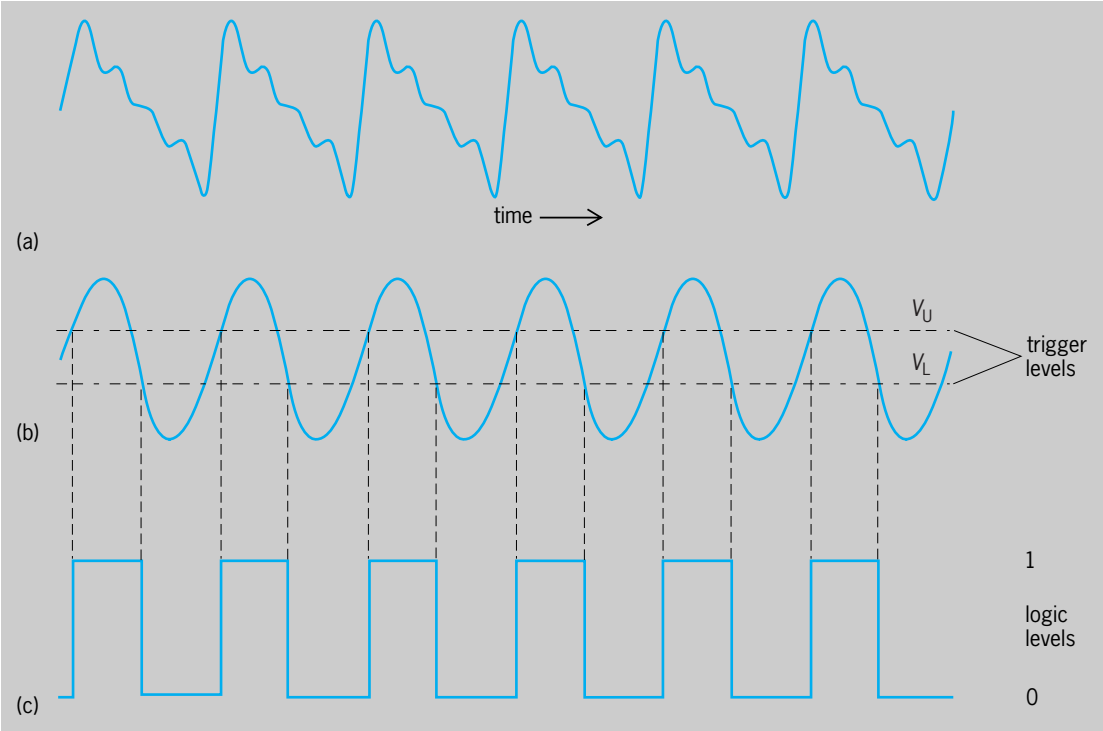

Fig. 1. Waveforms in frequency measurement. (*a*) Arbitrary cyclic variation. (*b*) Sine wave at the fundamental frequency. (*c*) Digital signal for counting.

technique is simple: the electrical signal (**Fig. 1***a*) from the sensor or transducer observing the physical system under test generally contains, from Fourier analysis, the fundamental frequency (Fig. 1*b*) and components at integral harmonics of this frequency. It is filtered to select the fundamental and converted into a rectangular waveform (Fig. 1*c*), representing transitions between the binary logic levels 0 and 1. A frequency measurement then consists of counting the number of positive- or negative-going transitions between the two levels in a known time. *See* FOURIER SERIES AND TRANSFORMS.

In parallel with the production of counters capable of operating at frequencies up to around 1 GHz, frequency standards based upon selected atomic transitions rather than the properties of bulk materials have been developed. These have the advantage that the frequency produced from a particular transition is in principle universal; that is, it is largely independent of the design of the standard and the materials used in its construction, and of changes in the ambient conditions. The combination of high-speed digital counters and of very stable atomic reference sources allows a wide range of frequencies to be determined simply, inexpensively, and very accurately.

As a result, much work has been carried out on the definition and measurement of other physical quantities in terms of frequency. Clearly, time and frequency are closely related; not only are the measurement, calibration, and dissemination techniques largely interchangeable, but any frequency standard may be converted into a standard of time, that is, a clock, by adding an appropriately designed counter. The unit of time, the second, is itself defined as the duration of 9 192 631 770 cycles of the electromagnetic radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom. The primary standards of voltage and current are now also realized in terms of frequency using the superconducting Josephson effect in conjunction with the von Klitzing (quantum Hall) effect. *See* ATOMIC TIME; ELECTRICAL UNITS AND STANDARDS; RESISTANCE MEASUREMENT; TIME; VOLTAGE MEASUREMENT.

General-purpose frequency counters. In a frequency counter, sometimes referred to as a frequency meter or timer-counter (**Fig. 2**), the input signal enters a direct-coupled broadband amplifier covering the frequency range of the instrument, whose gain can be switched to accommodate different signal levels. Protective elements may be placed at the input to prevent damage from excessive voltage levels. The amplified signal is then converted into a rectangular waveform with a bistable logic element, for example a Schmitt trigger. When the signal rises through a preset dc voltage level (trigger level V_u, Fig. 1*b*), the circuit switches very rapidly, in a few nanoseconds or less, from the 0 to the 1 state. The trigger level can be varied to select an appropriate point on the input waveform, preferably where it is varying rapidly. The trigger circuit is designed to show some hysteresis; that is, the reverse transition from 1 to 0 takes place at a lower voltage level (trigger level V_L) to reduce the possibility that noise might produce additional and unwanted transitions. *See* DIRECT-COUPLED AMPLIFIER; HYSTERESIS; MULTIVIBRATOR.

The output from the trigger circuit is passed to a counter through an electronic gate, which is opened for the required measurement period. The gating

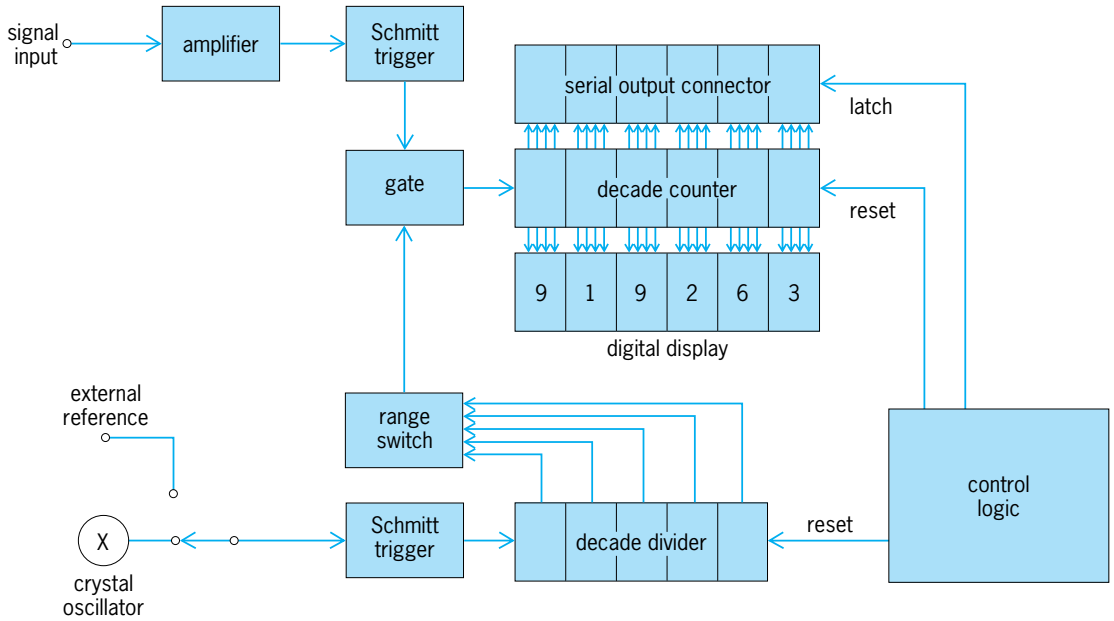

Fig. 2. Frequency-counter block diagram.

waveform is generated by dividing down a reference frequency, usually from a quartz crystal oscillator at 1, 5, or 10 MHz, with a decade divider. The waveforms from different stages of the divider may be used to provide various measurement times, for example from 1 ms to 10 s. For critical applications, an external frequency reference with high stability and accuracy may replace the internal crystal oscillator. *See* GATE CIRCUIT.

The selected transitions are counted by a decade counter, which may contain up to 10 stages, each of which consists of four solid-state bistable elements arranged to count in binary-coded decimal (BCD) code. The count is displayed on the front panel of the instrument using a light-emitting diode or fluorescent digital display. Liquid-crystal displays are sometimes found on small portable counters where battery life is important. The final count at the end of each measurement cycle is transferred (latched) into solid-state memory so that it is available for transmission to other instruments, either in parallel BCD code or after conversion to serial form. *See* DIGITAL COUNTER; ELECTRONIC DISPLAY.

The basic operations required of the control logic are to reset the counters at the start of each measurement cycle, to transfer the result to the output memory register at its conclusion, and to arrange an adjustable delay before the start of the next cycle if repeated measurements are required. Many modern frequency counters contain a microprocessor which supervises these functions, and adds others to increase the versatility of the instrument (**Fig. 3**). *See* FREQUENCY COUNTER; MICROPROCESSOR.

Measurement uncertainties. The errors in the readings displayed on the counter consist of avoidable errors arising from unwanted components in the input signal, and those inherent in the operation of the counter. Excessive noise in the signal, at a level greater than the hysteresis in the trigger circuit, gives rise to additional transitions and hence a high reading. This occurrence is usually obvious, as the number of additional counts is random and the reading varies with time. A stable reading is not, however, a guarantee that correct operation is taking place. The presence of harmonics in the signal (Fig. 1*a*) may lead to additional transitions within each cycle which add a constant proportion to the reading. On the other hand, the presence in the input of a large component at a low frequency, usually from power-line pick-up, may reduce the reading by shifting some cycles outside the trigger levels. Fortunately, each of these possibilities is easily detected by observing the input signal on an oscilloscope; it should consist mostly of the fundamental harmonic of the frequency to be measured, with the random noise and hum levels well below the hysteresis in the trigger circuit. *See* OSCILLOSCOPE.

In the absence of these gross errors, the uncertainty in the reading is the sum of two contributions. Because the counter effectively measures the ratio of

Fig. 3. General-purpose frequency counter. (*Hewlett-Packard Co.*)

the unknown frequency f to that of the reference f_0, any error Δf_0 in the latter appears proportionally in the result, as in Eq. (1), where Δf is the measurement.

$$\frac{\Delta f}{f} = -\frac{\Delta f_0}{f_0} \tag{1}$$

The frequencies generated by quartz crystal oscillators depend upon the ambient temperature and drift steadily with time. The temperature-compensated crystal oscillator (TCXO) found in less expensive counters contains a temperature-sensitive network in the oscillator circuit to minimize thermal effects. The associated uncertainty would typically be around 10^{-6} for a year after calibration. For better performance the crystals are aged, selected for good stability, and operated in a temperature-controlled oven. A stability of 10^{-7} or better over a year may then be achieved, limited by the aging rate.

The second uncertainty arises from the resolution of the counter itself. Because the phases of the unknown and reference frequencies are in general independent, the times of the signal transitions are unrelated to the start and end of the gate waveform. As a result, the count may vary by one in the last digit. The resultant uncertainty is given by Eq. (2), where

$$\frac{\Delta f}{f} = \frac{1}{\tau f} \tag{2}$$

τ is the gate or measurement time in seconds. For a given frequency, this uncertainty may be reduced only by increasing τ.

Measurements at low frequencies. Frequency counters often not only count the number of complete cycles within the gate period but also estimate the contribution from the partial cycles at each end. A simpler alternative uses a counter in the time interval mode, measuring the duration of a known number of cycles of the unknown frequency. Essentially, the input and reference signals (Fig. 2) are interchanged, so that the gate waveform is obtained by dividing down the unknown frequency while the reference frequency f_0 is counted for this period. As f_0 is typically between 1 and 10 MHz, the resolution uncertainty is then comparable with that of the reference frequency.

The unknown frequency is calculated from the time interval. The gate period is affected by the presence of noise in the input signal. An approximate calculation of the additional uncertainty arising from this source gives Eq. (3), where S_N is the input signal-

$$\frac{\Delta f}{f} = \frac{1}{\tau f S_N} \tag{3}$$

to-noise ratio. This is of the same form as the resolution uncertainty, but reduced by S_N.

Microwave frequency counters. The finite transition time of counting circuits limits their operation to frequencies less than about 1 GHz. Prescalars, very fast bistable elements arranged to divide the input frequency by a factor of 2 or 4, can extend the counter operating range up to about 3 GHz. To retain the advantages of the direct counting method, a method of transforming even higher frequencies into the working range in an accurate and predictable manner is required. Two mechanisms are available for this purpose, beat-frequency and harmonic-frequency generation. If two signals at frequencies f and $f + \Delta f$ are combined, the difference or beat frequency Δf appears as an amplitude modulation of the result. For electrical signals, the beat frequency may be isolated by mixing the waveforms in a nonlinear device such as a rectifying diode and removing the higher-frequency components with a filter circuit. *See* HETERODYNE PRINCIPLE; MIXER.

Because nonlinear devices distort the signal waveform, they also produce harmonics from a clean sinusoidal input. Step recovery diodes, for example, which have a sharp nonlinearity with a transition time of a few picoseconds, can produce a comb (so called from the appearance on a spectrum analyzer) of harmonics extending across the microwave region. It is possible to combine harmonic generation and beating in a single device to produce a measurable difference frequency Δf from the standard f_0 and a much higher unknown frequency f, given by Eq. (4).

$$\Delta f = \pm[f - nf_0] \tag{4}$$

To calculate the unknown frequency f, the appropriate sign in Eq. (4) and the harmonic number n must he determined. A microwave frequency counter based upon this principle of harmonic heterodyne down-conversion (**Fig. 4**) employs a synthesizer producing a high frequency f_S phase-locked to the internal crystal reference. The synthesizer can be programmed by a microprocessor controller to vary f_S in steps over a range, typically from 200 to 300 MHz. The frequency f_S is multiplied and mixed with the microwave frequency in the harmonic mixer forming the first stage of the counter. The measurement cycle involves reducing f_S from its maximum value until the filtered signal from the broadband intermediate-frequency amplifier contains a beat frequency Δf_1 within the range of the decade counter. This is measured, f_S is reduced by a small amount Δf_S, and a second reading Δf_2 of the beat frequency is taken. Bearing in mind that the harmonic number n must be a positive integer, Eq. (5) enables the sign and n to he determined ex-

$$n \approx \pm\left[\frac{\Delta f_2 - \Delta f_1}{\Delta f_S}\right] \tag{5}$$

actly, so that the microwave frequency can be calculated and displayed.

Optical frequency measurements. Above the microwave frequency range, for electromagnetic radiation in the submillimeter, infrared, and optical regions of the spectrum, frequency measurements became feasible only with the development of lasers producing narrow-bandwidth coherent radiation. Early work demonstrated that the stability of the laser frequency could be greatly improved by locking it to the centers of molecular absorption lines. By reflecting the

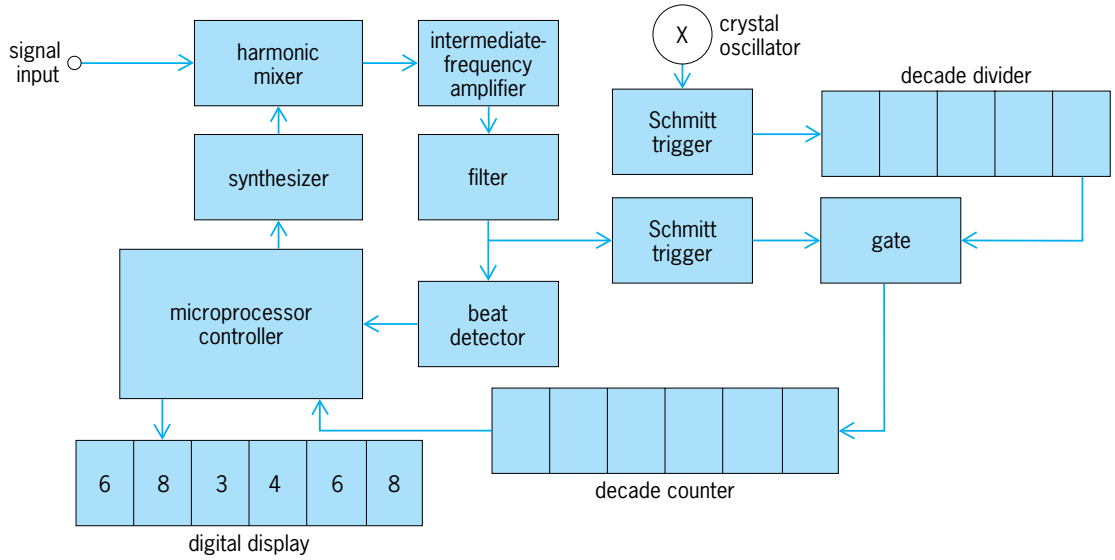

Fig. 4. Microwave frequency counter.

laser beam back through a cell containing a suitable gas at low pressure, it is possible to eliminate the line broadening from the random motion of the molecules, that is, the first-order Doppler effect. This technique of saturated absorption gives sub-Doppler linewidths of 1 MHz or less, and enables stabilities of around 1 part in 10^{12} to be achieved. *See* LASER; LASER SPECTROSCOPY.

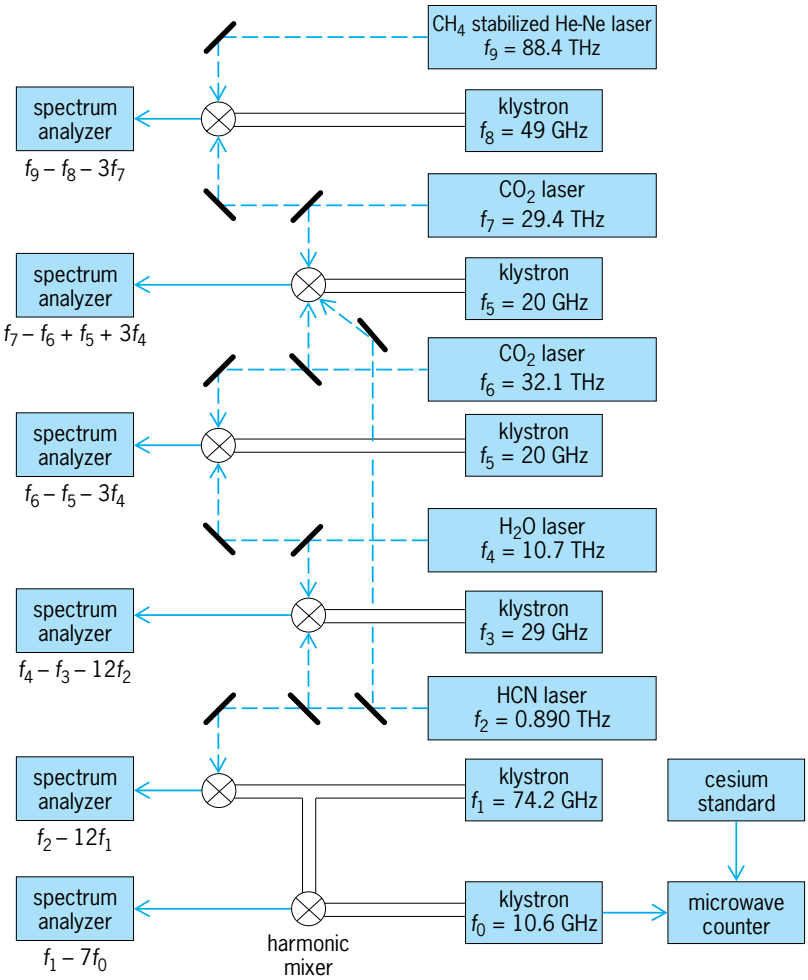

Fig. 5. Laser frequency measurement chain.

The equipment and techniques required for optical frequency measurement are complex and of value only in scientific and metrological research. A major reason for the early work was the possibility of obtaining a more accurate value for the speed of light c_0 from the frequency ν and vacuum wavelength λ_0 of a stabilized laser, using the relationship $\nu\lambda_0 = c_0$. The first precise determination of c_0 by this method was made in 1972 through the measurement of the frequency of a methane-stabilized helium-neon laser at 88.4 THz, corresponding to a wavelength of 3.39 micrometers (**Fig. 5**). The harmonic mixers in this and later work were point-contact devices bearing a close resemblance to the "cat's whisker" detectors from the early days of radio receivers. Here the whisker also serves as an antenna to couple the laser radiation into the mixer. Although the stabilized lasers were selected so that the best frequencies between their harmonics were small, microwave radiation from a stabilized klystron was also necessary to down-convert to a radio frequency within the range (30–1500 MHz) of the spectrum analyzers which monitored and measured the beats. All frequencies were referred to that (f_0) from an X-band klystron which was calibrated against a cesium atomic beam standard. *See* KLYSTRON; SPECTRUM ANALYZER.

The value obtained for c_0, 299 792 458 m · s^{-1}, was assigned an uncertainty of 4×10^{-9}, which arose mainly from the difficulty of comparing the laser wavelength to that of the krypton-86 emission line then forming the standard of length. In 1983 the General Conference of Weights and Measures adopted this value to define c_0; as a result, several stabilized lasers can now act as length standards with much greater accuracy than the historical krypton-86 lamp. *See* LIGHT; PHYSICAL MEASUREMENT; WAVELENGTH STANDARDS.

Characteristics and performance of frequency standards

Characteristic	Ovened quartz oscillator	Rubidium gas cell	Cesium beam	Hydrogen maser
Frequency f_T, MHz	0.01–100	6834.682614	9192.631770	1420.495752
Q factor	10^6	10^7	10^7	10^9
Reproducibility	Must be set	10^{-10}	2×10^{-12}	10^{-12}
Aging rate per year	10^{-8}	$<10^{-10}$	$<10^{-13}$	10^{-12}
Stability per day	10^{-11}	3×10^{-13}	2×10^{-13}	10^{-15}
Temperature coefficient				
Device	10^{-8}	2×10^{-10}	$<10^{-13}$	10^{-11}
Instrument	10^{-10}	10^{-12}	$<10^{-13}$	10^{-15}

The frequency chain shown in Fig. 5 is now mainly of historical interest. Later work has been directed toward the extension of measurements into the visible region of the spectrum, approximately 400–700 THz. The critical requirement to select a set of laser wavelengths which are nearly harmonically related, so that the beat frequencies are within the range of microwave counter systems, has been relaxed by the development of the optical-frequency comb generator. This consists of an electrooptic crystal, coated with reflecting surfaces to produce a Fabry-Perot cavity tuned to the input laser frequency and placed in a microwave resonator. It generates a comb of several hundred sidebands around the laser frequency with separations equal to the microwave frequency. As in the case of the microwave counter, the precise harmonic which produces the required beats for the counting system may be determined by shifting the microwave frequency by a small amount and observing the change in the count. *See* CAVITY RESONATOR; ELECTROOPTICS; FREQUENCY COMB; INTERFEROMETRY.

Frequency standards. To calibrate the internal quartz oscillators in frequency counters, and to enable frequency measurements to be made at the highest accuracies, up to and occasionally beyond 10^{-12}, standards laboratories require a selection of very stable frequency standards. The four types in common use are the temperature-stabilized or ovened quartz crystal oscillator, the rubidium gas cell, the cesium atomic beam standard, and the hydrogen maser. Their performance (see **table**) depends essentially upon the quality factor Q—the ratio of the resonant or transition frequency f_T to its half-bandwidth—and the sensitivity of f_T to changes in the properties of materials or in the ambient and operating conditions. A high value of Q means that the system acts as a narrow filter, reducing the effects of random noise processes and improving the short- to medium-term stability. *See* Q (ELECTRICITY).

The long-term stability is determined by the variation of f_T with time. In only one type of standard, the quartz crystal oscillator, is f_T determined directly by bulk physical properties; as a result, this has the poorest long-term stability of the four. However, quartz oscillators are employed in most of the atomic standards to reduce the short-term noise and to provide a convenient output frequency (usually 10 MHz). In these, f_T is set by atomic transitions whose properties are in principle fixed and universal. In practice, small interactions with the containment system and the operating conditions mean that this ideal is not completely realized. In the rubidium gas cell, the transition is perturbed by collisions with other buffer gas atoms whose temperature and composition may change in time; in the hydrogen maser, collisions of the hydrogen atoms with the inert coating inside the containing bulb produce the so-called wall shift, which depends upon the condition of the coating. Atoms in the cesium beam standard are very well isolated from each other and the container, and this is reflected in the low drift rates and temperature coefficients observed. *See* ATOMIC CLOCK.

The effects of these factors are well illustrated by a comparison (**Fig. 6**) of the long-term performances against Universal Coordinated Time of a high-stability cesium standard and a good-quality hydrogen maser. The cesium frequency stability is limited by random noise processes as result of its lower Q, but shows little evidence of long-term drift. The maser, which has a very high Q, is very stable in the short term but possesses a linear frequency drift. Small adjustments to the maser operating conditions, as made around day 800, can alter the drift rate significantly.

Because drift rates can be measured and corrections made, the factors limiting the accuracy achieved with a frequency standard are the type and magnitude of the various internal noise sources. Plotted as a function of the measurement time (**Fig. 7**), the stability of a frequency standard shows a general boat-shaped curve. At short times, where white phase noise dominates, the stability improves as the measurement time increases. It eventually flattens

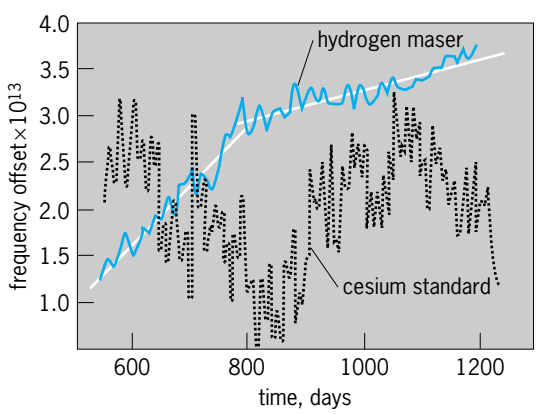

Fig. 6. Long-term stabilities of a cesium standard and a hydrogen maser. Straight lines indicate a linear drift of the maser.

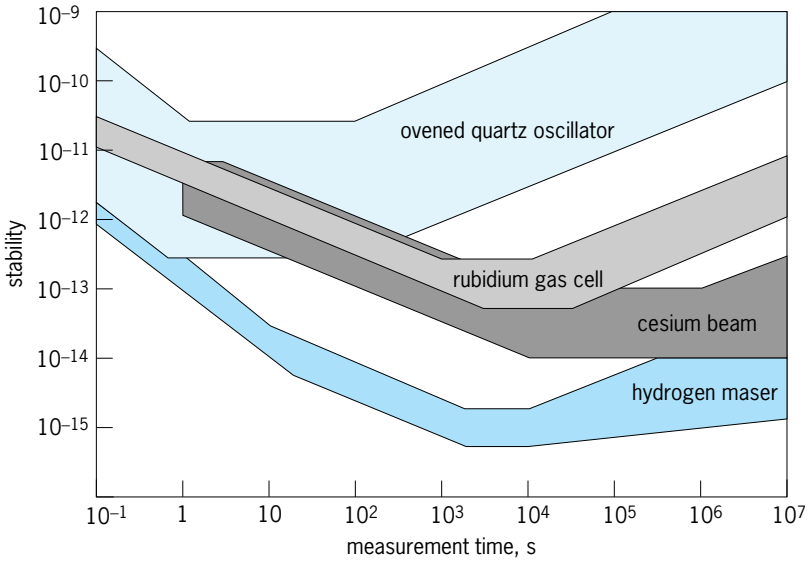

Fig. 7. Stability of frequency standards as a function of measurement time. The spread for each type indicates the range of values expected from good commercial standards. The graph is idealized to emphasize the change, with measurement time, of the type of noise that limits performance.

off to an optimum value, when flicker noise is the major component, before rising again as random-walk phase noise sources become important. *See* ELECTRICAL NOISE.

Comparison techniques. The comparison of frequency standards benefits from the fact that the signals are close to ideal: they are clean and stable, and the differences from the nominal value, usually 1, 5, or 10 MHz, are extremely small. The beat frequency is often too low to be measured directly; for example, a difference of 10^{-10} at 10 MHz corresponds to a frequency of 0.001 Hz. The resolution may be improved by measuring the phase of the beat frequency or by employing a frequency difference multiplier.

The relative phase θ of two sine waves may be measured by converting them to square waves and measuring the overlap of the waveforms with AND gates or J-K bistable logic elements. Digital methods have several advantages over analog techniques: the output varies linearly with θ and is independent of the signal amplitudes, there is no ambiguity in the sign of the frequency difference, and the constant of proportionality between output voltage and phase is easily determined. The main disadvantage of the simpler digital circuits is that there is often a dead zone around zero phase difference due to the finite transition time of the logic circuits. This may be removed by mixing a third frequency with both the unknown and reference signals. *See* LOGIC CIRCUITS.

In a frequency difference multiplier (**Fig. 8**), the unknown frequency is multiplied by a factor of 10, while the reference frequency is multiplied by 9. When the two are subtracted in a mixer, the frequency difference Δf has been increased by a factor of 10. Several stages can be placed in series to increase the multiplication factor. The output signal is measured with a conventional counter. As Δf is superimposed upon the reference frequency, there is no ambiguity as to its sign. For a four-stage multiplier and an unknown frequency of approximately 1 MHz, the resolution on the highest range is 10^{-10} for a measurement time of 1 s. It may be improved to below 10^{-13} for longer times as long as the signals are clean and stable. Ultimately the practical limit to the multiplication factor is set by noise in the multiplication and mixing circuits.

Sine-wave signals contain a large measure of redundancy; for the 10-MHz output from a standard with an accuracy of 10^{-12}, the number of complete cycles is known for periods of up to a day. Converting continuous frequency signals into timing pulses allows concentration on the small differences accumulating over long periods. An additional advantage, particularly important for signals being transmitted or broadcast over long distances in the presence of interference, is that signal continuity is no longer necessary.

Comparisons with remote standards. For measurements and calibrations at the highest level of accuracy, working standards must be referred to the national or international standards of time and frequency. The International Atomic Time (TAI) scale provides the

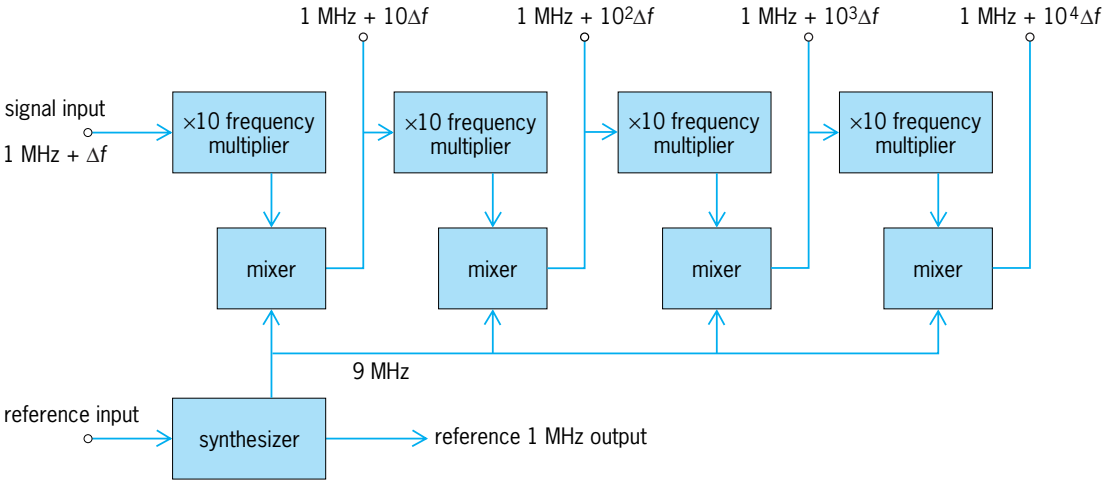

Fig. 8. Frequency difference multiplier. The unknown frequency of the signal input is approximately 1 MHz.

ultimate reference in this respect. It is generated at the International Bureau of Weights and Measures (BIPM) in Paris from the readings of a worldwide set of commercial cesium clocks, long-beam cesium primary standards, and hydrogen masers. The results are issued in the form of Universal Coordinated Time (UTC), which contains adjustments to TAI in the form of leap seconds added every 1–2 years, which keep it close to the astronomical time scale based upon the rate of rotation of the Earth. Because TAI and UTC are calculated "paper" scales with values for each month issued 1 month in arrears, the national standards laboratories in many countries generate real-time approximations to UTC. These are "steered" so that the difference from UTC is not more than 100 ns in time and around 10^{-13} in frequency.

At one time, laboratory standards could be compared with national or international standards through traveling cesium clocks. This system is time-consuming, is expensive, and allows only infrequent comparisons. It has therefore largely been replaced by day-to-day monitoring against a remote reference standard using suitable transmissions, for example, from very low frequency (VLF) or low frequency (LF) standard broadcasts, the Loran-C navigational chain, television signals, and the satellite Global Positioning System (GPS). Suitable transmissions are also monitored by national laboratories, which publish tables of corrections to enable users to refer their measurements to the national standards. The main uncertainty arises from variations in the propagation delays between the transmitter and different laboratories. *See* LORAN; SATELLITE NAVIGATION SYSTEMS.

Phase measurement has been applied to the monitoring of laboratory standards using the VLF or LF services provided by some national standards laboratories. A phase-tracking receiver detects, amplifies, and filters the signal. A higher frequency, usually 1 MHz, is then generated by a phase-locked loop, and the phase difference from the corresponding frequency from the laboratory standard is monitored. The phase stability of the received signal is sensitive to interference of the ground-wave component by the variable sky wave, limiting the useful range to around 600 mi (1000 km) from the transmitter. Measurements are best carried out around the middle of the day, when the sky-wave effects are at a minimum. A resolution equivalent to a mean frequency change of 10^{-12} from one day to the next can be achieved, although propagation errors in the middle of winter may be several times larger.

The transmissions from the satellites forming the GPS are particularly convenient for this purpose, as they contain timing information from the atomic standard present in each satellite, and the signal may be received at any point on the Earth's surface. The pseudo-random-noise-coded signals may be analyzed by a suitable receiver to give time with respect to the U.S. Naval Observatory's realization of UTC. Although UTC(USNO) is very stable in the long term, frequency measurements may be affected by short-term noise. Common-view operation, in which the same schedule of GPS satellites is observed by the user and a national standards laboratory, allows cancellation of the short-term differences. In this case, frequency measurements with respect to the national standard accurate to 10^{-13} over one day may be achieved. The comparison process has been simplified by the development of GPS-disciplined oscillators, that is, instruments in which a quartz or rubidium standard is controlled by reference to the signals from GPS satellites to improve the long-term performance.

Peter B. Coates

Bibliography. A. E. Bailey (ed.), *Microwave Measurements*, 2d ed., 1989; K. M. Baird, Frequency measurement of optical radiation, *Phys. Today*, 36(1):52–57, January 1983; Chronos Group, *Frequency Measurement and Control*, Micowave Technology 7, 1994; V. E. Bottom, *Introduction to Quartz Crystal Units Design*, 1982; J. Jesperson and J. Fitz-Randolph, *From Sundials to Atomic Clocks: Understanding Time and Frequency*, 2d ed., 1999; P. Kartaschoft, *Frequency and Time*, 1978; M. Ohtsu (ed.), *Frequency Control of Semiconductor Lasers*, 1996; J. Vanier and C. Audoin, *The Quantum Physics of Atomic Frequency Standards*, 2 vols., 1989.

Frequency modulation

A technique used in telecommunications transmission systems whereby the instantaneous frequency of a carrier signal is determined by the characteristics of an information signal, the so-called modulating signal. Frequency modulation (FM) is a form of angle modulation. For systems in which the modulating signal is digital in nature, the term frequency-shift keying (FSK) is usually employed. *See* ANGLE MODULATION.

The mathematical properties of FM signals were first presented by J. R. Carson in 1922. Carson established the fact that frequency-modulated carrier signals generally occupy a larger transmission bandwidth than would be the case for amplitude-modulated carriers. Further investigation by E. H. Armstrong demonstrated the improved noise suppression properties of FM as a benefit of the expanded bandwidth. Thus, by trading increased transmission bandwidth for improved signal-to-noise ratio it is possible to reduce transmitter power requirements for a given transmission range or desired performance. This property is of special importance in applications such as satellite communications, where satellite transmitter power and weight are important considerations. *See* BANDWIDTH REQUIREMENTS (COMMUNICATIONS); COMMUNICATIONS SATELLITE; SIGNAL-TO-NOISE RATIO.

Additional advantages of FM over other forms of modulation are due to the constant amplitude of the modulated carrier, independent of the modulating signal. For example, this feature provides partial immunity to the presence of nonlinearities in the signal processing equipment and allows the reduction of

amplitude variations due to signal fading and additive noise in the transmission channel.

Constant-amplitude signals also exhibit constant levels of peak and average power. This property simplifies the design and implementation of transmitters, amplifiers, and automatic volume control circuits, and mitigates against problems associated with circuit nonlinearities.

Carrier signal specification. In the usual definition of FM, the carrier signal is taken to be a periodic function of time represented mathematically as Eq. (1).

$$c(t) = A\cos(2\pi f_c t + \theta) \tag{1}$$

Here the quantities A, f_c, and θ are constants known as the amplitude of the carrier signal in volts, its frequency in hertz, and its phase angle in radians, respectively. *See* CARRIER.

If θ is allowed to be a function of time $\theta(t)$, $c(t)$ will not be periodic and the concept of frequency would, strictly speaking, not apply. In this case it is meaningful to define the concept of instantaneous frequency $f_i(t)$, given by Eq. (2).

$$f_i(t) = f_c + \left(\frac{1}{2\pi}\right) d\theta(t)/dt \tag{2}$$

Clearly this quantity varies with the time rate of change of the carrier signal's phase. When time is measured in seconds, instantaneous frequency is again measured in hertz.

The difference between $f_i(t)$ and f_c is known as the frequency deviation Δf and is given by Eq. (3).

$$\Delta f = f_i(t) - f_c = \left(\frac{1}{2\pi}\right) d\theta(t)/dt \tag{3}$$

Carrier modulation. Let $m(t)$ be the information signal modulating the carrier and assume its maximum value to be m_p and its bandwidth to be B hertz. In frequency modulation $m(t)$, which is assumed to be varying continuously with time, is impressed upon the instantaneous frequency of the carrier signal according to Eq. (4), so that the frequency deviation is given by Eq. (5).

$$f_i(t) = f_c + km(t) \tag{4}$$

$$\Delta f = km(t) \tag{5}$$

Here k is the so-called deviation constant. The maximum frequency deviation is then given by Eq. (6).

$$\Delta f_{\max} = km_p \tag{6}$$

Combining Eqs. (1), (2), and (4) shows that the modulated carrier takes the form given by Eq. (7).

$$s(t) = A\cos\left[2\pi f_c t + 2\pi k \int m(\tau)d(\tau)\right] \tag{7}$$

For the case of frequency-shift keying, $m(t)$ is a discrete waveform that can assume only a finite number of amplitudes that can change only at discrete points in time. These amplitudes are impressed onto the carrier signal by shifting its frequency by predetermined amounts. In binary FSK the modulated carrier signal corresponding to the two amplitudes takes the form given by Eqs. (8). Here Δf is known as the frequency deviation.

$$s_1(t) = A\cos[2\pi(f_c + \Delta f)t] \tag{8a}$$

$$s_2(t) = A\cos[2\pi(f_c - \Delta f)t] \tag{8b}$$

Figure 1 shows examples of frequency-modulated periodic carrier signals for which the modulating signal $m(t)$ is either a ramp function (Fig. 1*a*) or is itself a periodic function of time (Fig. 1*b*).

For a sinusoidal modulating signal of the form $m(t) = a\cos(2\pi f_m t)$, where a is a constant, the modulating index β is defined by Eq. (9), and represents the ratio of maximum frequency deviation relative to the frequency of the modulating signal. For binary FSK the modulation index is defined by Eq. (10).

$$\beta = ka/f_m \tag{9}$$

$$\beta = \Delta f/B \tag{10}$$

Spectrum and power characteristics. Theoretically, the energy spectrum of a frequency-modulated signal exhibits an infinite number of sidebands and so extends over the entire frequency spectrum. Practically, however, only a finite number of sidebands contain significant energy. Although the computation of the effective bandwidth is quite complicated, certain approximations can be made that lead to simple formulas. Two cases are typically distinguished. In so-called narrowband FM, the maximum frequency deviation is assumed to be quite small, in which case it can be shown that the effective bandwidth of the FM signal is given by Eq. (11).

$$\mathrm{BW}_{\mathrm{NBFM}} = 2B \tag{11}$$

See SIDEBAND.

The case where the above assumption does not hold is referred to as wideband FM. For large values of frequency deviation the bandwidth is essentially determined by the maximum frequency deviation, and is given by Eq. (12).

$$\mathrm{BW}_{\mathrm{WBFM}} = 2\Delta f_{\max} \tag{12}$$

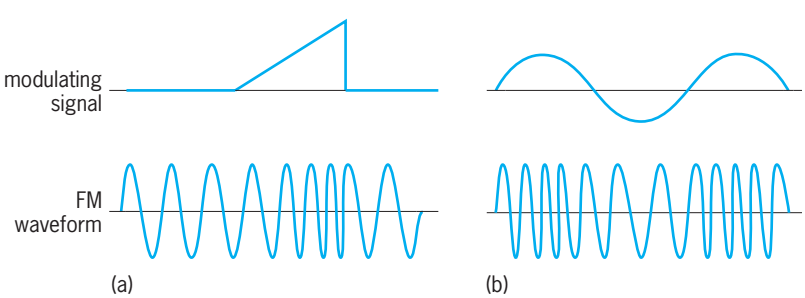

Fig. 1. Examples of modulating signals and corresponding frequency-modulated waveforms. (*a*) Ramp modulating signal. (*b*) Periodic modulating signal.

These two special cases can be combined into the general formula (13) which is known as Carson's

$$\mathrm{BW}_{\mathrm{FM}} = 2(\Delta f_{\max} + B) \tag{13}$$

rule. A similar equation holds for the bandwidth of a binary FSK-modulated carrier.

The average power of the modulated signal, as well as its peak power, are constant, and relative to a 1-ohm resistance they are equal to $A^2/2$ watts and A^2 watts, respectively. Both powers are apportioned between the unmodulated carrier and the modulating signal. Thus, as the power in the modulating signal increases, the power in the unmodulated carrier decreases correspondingly.

Modulation devices. There are several techniques for generating a frequency-modulated signal.

The direct method of modulation typically employs a voltage-controlled oscillator whose instantaneous frequency can be changed linearly in accordance with a voltage applied at its input. Such devices usually include a resonant *LC* circuit with either a varactor diode or a reactance tube whose capacitance or inductance can be changed in direct proportion to an applied voltage, thereby changing the resonant frequency of the circuit (**Fig. 2**). In some cases a certain amount of frequency multiplication may be necessary to produce the required amount of frequency deviation. *See* FREQUENCY MULTIPLIER; RESONANCE (ALTERNATING-CURRENT CIRCUITS).

In the indirect method of generating a frequency-modulated signal, a two-step process is followed. First, the modulating signal $m(t)$ is passed through an integrator whose output is multiplied with the unmodulated carrier signal.

Adding to that product the carrier signal phase-shifted by 90° results in the so-called narrowband FM signal. In the second step, the frequency of the narrowband signal is increased through one or more steps of frequency multiplication and spectrum translation (**Fig. 3**). *See* FREQUENCY MODULATOR.

Demodulation. As shown by Eq. (4), demodulation of the frequency modulated signal $s(t)$ in Eq. (7) requires the determination of its instantaneous frequency $f_i(t)$. One commonly used method of extracting $f_i(t)$ involves the intermediate generation of an amplitude-modulated signal whose amplitude is proportional to $f_i(t)$, and the subsequent use of an AM

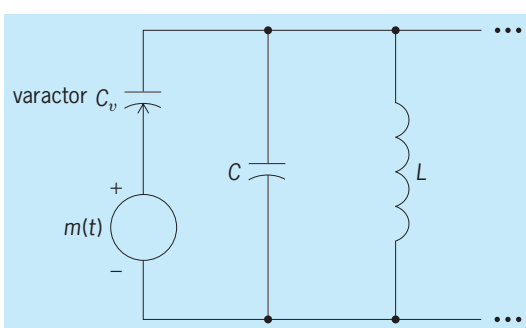

Fig. 2. Example of direct method of frequency modulation.

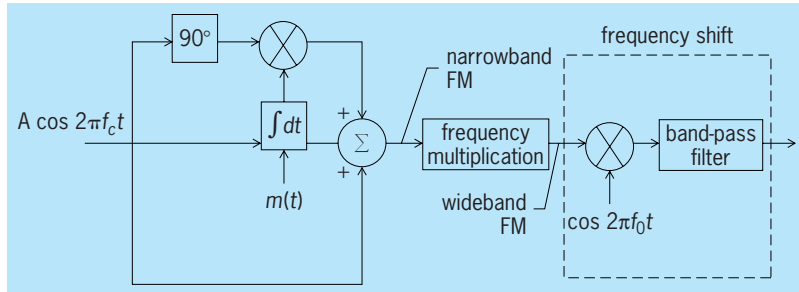

Fig. 3. Example of indirect method of frequency modulation.

demodulator to recover the original modulating signal $m(t)$.

The FM-to-AM conversion is generally performed by devices known as discriminators, of which several types are in common use. One of them performs the differentiation of $s(t)$ with respect to t, which results in Eq. (14).

$$ds(t)/dt = -2\pi A[f_c + km(t)] \times \sin\left[2\pi f_c t + 2\pi k \int m(\tau)d(\tau)\right] \tag{14}$$

This quantity represents a frequency-modulated carrier whose amplitude varies in proportion to $f_i(t)$. In another method $s(t)$ is processed through a filter whose gain characteristic is linear over the range of frequencies covered by the signal. Such linear gain characteristics can be achieved in circuits known as balanced discriminators. *See* ELECTRIC FILTER.

Standard amplitude demodulation techniques such as envelope detection can be used on the output of the discriminator to recover $m(t)$. *See* AMPLITUDE-MODULATION DETECTOR.

As an alternative to FM-to-AM conversion, FM demodulation can be achieved through the use of phase-locked loops or other so-called feedback demodulators. These generally are of greater complexity, but perform better in low signal-to-noise ratio environments. In such feedback systems, the difference between the output of the demodulator and a reference signal is applied at the input, which in turn increases or decreases the output, depending on whether the feedback is positive or negative. The result is generally a decrease of distortion, especially for narrowband signals exposed to wideband channel noise or fading. *See* PHASE-LOCKED LOOPS.

Instantaneous frequency can also be determined through the technique of measuring the rate at which the amplitude of the modulated carrier changes its polarity. Devices operating on this basis are known as zero-crossing detectors. *See* FREQUENCY-MODULATION DETECTOR.

Performance. Since the information in a frequency modulation system is contained in the frequency of the transmitted signal, distortions of the signal during transmission that affect its amplitude can be removed prior to demodulation through a process known as amplitude clipping. This technique is particularly effective in the presence of additive noise. Since

clipping tends to produce higher-order harmonic frequencies, low-pass filtering after demodulation is frequently employed.

Wideband FM provides a tradeoff between the bandwidth occupied by the transmitted signal and the effective postdemodulation signal-to-noise ratio. As an example, it can be shown that for additive channel noise with a constant power spectral density and a large ratio of carrier power to modulating signal power, the signal-to-noise ratio increases as the square of the transmission bandwidth. This effect, however, no longer holds when the signal-to-noise ratio falls below a certain limit. In that case, the receiver locks unto the noise as the primary signal and suppresses the information signal. This phenomenon is known as noise capture and is sometimes referred to as the threshold effect. The signal-to-noise ratio at which it occurs is known as the threshold. Under certain assumptions it can be shown that the onset of noise capture occurs at a signal-to-noise power ratio of 10.

Preemphasis and deemphasis. An additional improvement in the signal-to-noise ratio beyond the expansion of the signal bandwidth can be achieved by taking advantage of the fact that the noise power at the receiver output increases parabolically with increasing frequency, even if the receiver input noise is spectrally flat. The high-frequency components of the information signal are therefore more affected by the noise than the low-frequency components. By artificially boosting the high-frequency components of the information signal prior to modulation and restoring them to their original value after demodulation, a signal-to-noise ratio advantage can be gained. Circuits that perform these operations are known as preemphasis and deemphasis filters.

Applications. Due to its enhanced performance over amplitude modulation, wideband FM is widely used in commercial radio broadcasting, although at the cost of an increase in occupied bandwidth to 200 kilohertz, compared to 10 kHz for commercial AM broadcasting. The additional bandwidth allows for the transmission of audio program material over a range of 15–20 kHz compared to only 5 kHz for AM. It is also possible to transmit audio signals in stereo through the multiplexing of two audio channels, and to include additional channels of monophonic or stereo programming material for special purposes, all within the same 200-kHz bandwidth. *See* FREQUENCY-MODULATION RADIO; STEREOPHONIC RADIO TRANSMISSION.

Wideband FM is also used in the transmission of the audio portion of commercial television programs, and finds widespread application in satellite communications for telephony, single-channel-per-carrier systems, and wide-area commercial television distribution. *See* TELEVISION.

Narrowband FM is primarily employed in telemetry and mobile communication systems. Its major advantage over AM lies in its better response at very low frequencies and its ability to reject large additive noise pulses through the technique of amplitude limiting. *See* MOBILE RADIO; MODULATION; TELEMETERING.

Hermann J. Helgert

Bibliography. B. P. Lathi, *Modern Digital and Analog Communication Systems*, 3d ed., Holt, Rinehart and Winston, 1998; J. Proakis and M. Salehi, *Fundamentals of Communication Systems*, Pearson Prentice Hall, 2005; M. Schwartz, *Information Transmission, Modulation, and Noise*, 4th ed., McGraw-Hill, 1990.

Frequency-modulation detector

A device for the detection or demodulation of a frequency-modulated (FM) wave. Frequency-modulation detectors operate in several ways. In one class of detector, known as a discriminator, the frequency modulation is first converted to amplitude modulation, which is then detected by an amplitude-modulation detector. Another type of FM detector employs a phased-locked oscillator to recover the modulation. A still different type converts the frequency modulation to pulse-rate modulation, which can be converted to the desired signal by use of an integrating circuit. Examples of these types will be discussed here. *See* AMPLITUDE-MODULATION DETECTOR.

Discriminator. An amplitude-modulation (AM) detector demodulates an FM wave, provided the detector is detuned so that the carrier frequency is on one side of the passband instead of in the center of the passband. Frequency variations are then converted into amplitude variations or modulation (**Fig. 1**), and an AM detector produces the desired output signal. This detection technique has serious limitations, however. Low distortion is realized if the amplitude variation is linearly related to the frequency variation or deviation of the FM wave. This requirement confines the frequency deviation to a linear or straight-line portion of the tuned-circuit response curve. This restriction seriously limits the permissible deviation, and thus nullifies the advantages of FM that result from high deviation ratios. However, amateur radio operators sometimes use this detuning technique to receive narrow-band FM on an AM receiver. *See* FREQUENCY MODULATION.

The technique of converting the frequency variations of an FM signal into amplitude variations is used in the Foster-Seely discriminator and the ratio

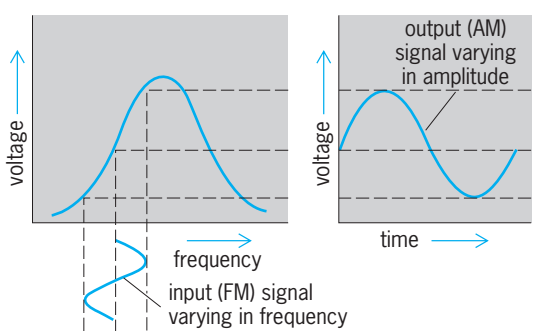

Fig. 1. Conversion of frequency modulation to amplitude modulation by a detuned resonant circuit.

detector. In these circuits the primary voltage of the transformer used to couple the final intermediate-frequency (i-f) amplifier or limiter is coupled to the center tap of the tuned secondary, which is connected to two detector diodes in series. These discriminators performed well but were not easily adaptable to integrated circuits and have, therefore, been essentially eliminated from modern radio technology.

Quadrature detectors. Frequency-modulated waves may be demodulated by a phase-shifting network, an analog multiplier, and a low-pass filter (**Fig. 2**). The *RLC* circuit in the phase shifter is tuned to the carrier frequency of the FM wave, and the reactance of the series capacitor C_1 is large in comparison with the resistive impedance of the tuned circuit at the carrier frequency. Then the voltage v_2 is in phase with the capacitive current, which is approximately 90° out of phase with the input voltage v_i at the carrier frequency. Therefore, if v_i is a sine wave, v_2 is a cosine wave. When the input signal is v_i is modulated, the input frequency changes by $\delta\omega$; thus the instantaneous phase is $(\omega_O + \delta\omega)t$. At this higher frequency the tuned circuit is no longer resonant but is capacitive, because the capacitive current exceeds the inductive current and the voltage v_2 lags the capacitive current and its phase leads the phase of v_i by less than 90°. If the relative phase shift is proportional to the frequency deviation $\delta\omega$, then the result of multiplying the two points is prefilter output v'_0, which is the sum of a low-frequency and a high-frequency term. The high-frequency term is eliminated by the low-pass filter, leaving an output voltage v_0 which differs from the modulating voltage only in amplitude. *See* ELECTRIC FILTER.

The assumption that the relative phase shift of the phase-shifting network is directly proportional to the frequency deviation of the modulated wave is not strictly true. However, the distortion resulting from this lack of linearity may be limited to about 1% if the bandwidth of the resonant circuit in the phase-shifting network is at least five times as wide as the maximum deviation of the FM wave.

An elementary multiplying circuit is given in **Fig. 3**. This circuit is essentially the same as the unbalanced differential, or emitter-coupled, amplifier except that the base of the current-source transistor T_1 is capacitively coupled to an input terminal. Therefore, the collector current of T_1 is controlled by

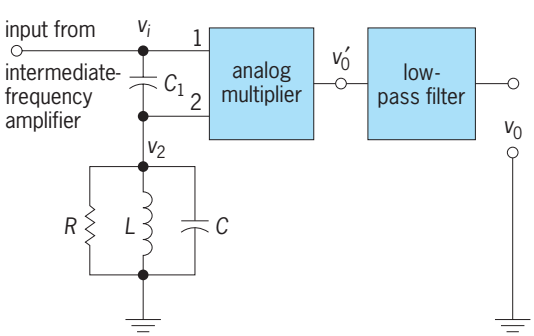

Fig. 2. Quadrature detector.

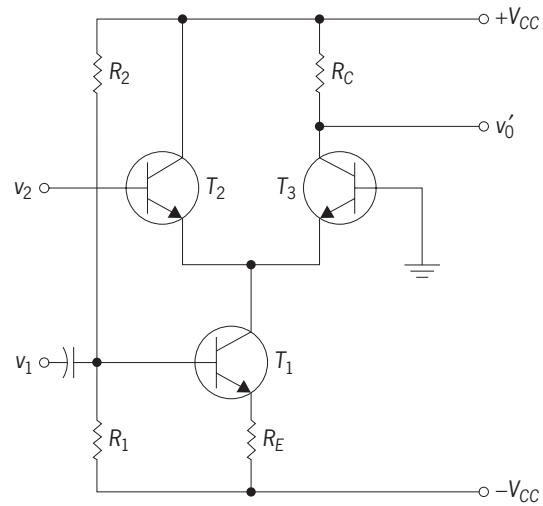

Fig. 3. Elementary multiplier circuit.

the input voltage v_1. The signal component of this current is essentially v_1/R_E. On the average, one-half of this current flows through T_2 and the other half through T_3. The collector current of T_3 due to a signal v_1 is therefore $v_1/2R_E$, and the voltage gain of a common-base or common-emitter transistor amplifier is g_mR_L, where g_m is the transconductance and R_L is the load resistance of the capacitor, here equal to R_C. For a transistor at normal temperature, it may be shown that $g_m = 40I_C$, leading to Eqs. (1) and (2).

$$\frac{v'_0}{v_2} = g_mR_C = 40I_CR_C = \frac{40v_1R_C}{2R_E} \qquad (1)$$

$$v'_0 = \frac{20R_C}{R_Ev_1v_2} \qquad (2)$$

See AMPLIFIER; DIFFERENTIAL AMPLIFIER.

Integrated four-quadrant multipliers based on the above principle are used extensively in various types of demodulators. These multipliers accept signal voltages of either polarity at both input ports.

The analog multiplier may be replaced by an AND gate in the quadrature detector. The gate is then called a coincident circuit. When the carrier is unmodulated, the 90° phase shift causes the two gate inputs to be positive and the gate output to be high for about 90° out of each cycle with average output voltage of approximately $V_{CC}/4$, where V_{CC} is the collector supply voltage. When the modulation causes the frequency to increase, the relative phase decreases and the two inputs are high, or coincident, for more than 90° out of each cycle; the average output increases in proportion to the frequency deviation toward a limiting voltage of $V_{CC}/2$. On the other hand, as the frequency deviation is downward, the coincidence time decreases toward zero as the relative phase approaches 180° and the output voltage decreases toward zero. The relative phase shifts must, however, be small in comparison with these limiting values in order to achieve acceptable distortion levels. The no-modulation phase shift in the phase-shifting network may differ considerably from the

ideal 90° without serious degradation of the detector output. *See* LOGIC CIRCUITS; SWITCHING CIRCUIT.

Locked-oscillator detector. Another type of FM detector is the locked oscillator, in which a local oscillator follows, or is locked to, the input-signal frequency. The phase difference between the local oscillator and the input signal is proportional to the frequency deviation of the input signal. A phase detector and low-pass filter are used to produce an output voltage that is proportional to the phase difference and therefore the frequency deviation of the input signal. The output voltage is also used to control the local-oscillator frequency and maintain its lock to the input signal over the desired frequency range. This arrangement is commonly known as a phase-locked loop (PLL).

A block diagram of a phase-locked loop is given in **Fig. 4**. The phase detector may be a multiplier or AND gate, as described above. The voltage-controlled oscillator (VCO) is basically a frequency modulator. The voltage-controlled oscillator frequency is adjusted to the carrier frequency of the desired signal. As this frequency approaches this desired signal to the point where the difference frequency is within the bandwidth of the low-pass filter, the voltage-controlled oscillator pulls into lock and is said to be captured. It maintains lock with the input signal as long as their phase difference does not exceed 90°, which is set by the limits of the phase detector. This lock range is proportional to the loop gain of the phase-locked loop. This gain, and therefore the lock range, may be controlled, or set, by inserting an amplifier between the loop filter and the output or between the output and the voltage-controlled oscillator. The choice depends upon the desired output voltage. The lock range must be greater than the maximum frequency deviation of the FM signal but should not exceed the total bandwidth allotted to the FM station. *See* FREQUENCY MODULATOR.

The phase-locked loop is a widely used demodulator for FM signals. It requires no *LC* (inductance-capacitance) tuned circuits, which are cumbersome to incorporate into integrated circuits. Furthermore, the only source of distortion in a properly locked loop is the lack of linearity of the voltage-controlled oscillator, and integrated voltage-controlled oscillators that provide distortions well under 1% are readily available. However, the phase-locked loop is more complex and therefore more costly than the simpler quadrature detectors. *See* PHASE-LOCKED LOOPS.

Integrating detector. A frequency-modulated wave can be converted to a pulse-rate-modulated wave from which the original modulating signal can be recovered by the use of an integrator. The FM input signal is first converted to a low intermediate frequency by mixing the input signal with a local oscillator signal in a nonlinear amplifier. *See* AMPLITUDE MODULATION.

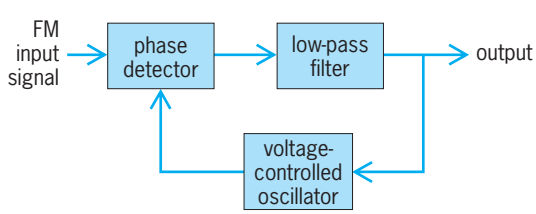

Fig. 4. Diagram of phase-locked loop.

The local oscillator frequency is nearly equal to the carrier frequency, and so the difference, or intermediate, frequency is low compared with the carrier frequency. Then the frequency deviation is a high percentage of the intermediate frequency. For example, the maximum frequency deviation of a standard broadcast FM signal at 100 MHz may be 150 kHz, or 1.5 parts per thousand. Then, if the intermediate frequency is only 200 kHz, the frequency deviation of the intermediate frequency is 750 parts per thousand. Thus, the intermediate-frequency amplifier is normally an untuned low-pass amplifier. The intermediate-frequency amplifier is followed by a limiter which clips the signal peaks and converts the sinusoidal intermediate frequency to a nearly square wave. This square wave is differentiated to obtain pulses which trigger the bistable multivibrator. The multivibrator then produces a rectangular pulse of constant amplitude and width for each input pulse. Therefore, the pulse rate is equal to the input frequency. If the frequency deviation is such that the input frequency is minimum, the pulses are widely spaced for a low-frequency input; on the other hand, the pulses are closely spaced when the deviation produces maximum input frequency. The integrator is basically a low-pass *RC* filter, which produces an output voltage proportional to the average value of the input voltage and filters out the high-frequency components of the pulses. Therefore, the output voltage is proportional to the pulse rate. The simple *RC* integrator can be replaced by an operational-amplifier-type integrator for improved filtering, linearity, and output voltage. *See* LIMITER CIRCUIT; MULTIVIBRATOR; OPERATIONAL AMPLIFIER.

Charles L. Alley

Bibliography. R. G. Best, *Phase-locked Loops: Simulation, Design and Applications*, 4th ed., 1999; H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, 1980; C. A. Schuler, *Electronics: Principles and Applications*, 5th ed., 1998.

Frequency-modulation radio

Radio transmission accomplished by variation of the frequency aspect of a radio-frequency (RF) carrier signal by an input signal (called the modulating signal). The frequency-modulation (FM) technique was invented by Major Edwin A. Armstrong in the 1930s. FM radio broadcasting offers a number of advantages over amplitude-modulation (AM) radio broadcasting, most significantly higher audio fidelity and greater immunity to atmospheric and human-made sources of interference.

When an RF carrier is frequency-modulated, a theoretically infinite number of "sidebands" are created which surround the carrier (that is, are located above and below the carrier, in frequency), although in practical implementations the extent of these

sidebands is limited to a finite bandwidth (called the frequency deviation). The sidebands formed during the modulation process are separated from one another by an amount proportional to the frequencies present in the audio signal. The number of significant sidebands depends on the amplitude of the modulating signal, while the amplitude of these sidebands diminishes progressively (though not monotonically) as the sidebands occur farther and farther from the carrier frequency.

Different FM systems use different frequency deviations, also called deviation ratios, depending upon the application. The greater the deviation ratio, the larger the bandwidth required for the transmission. Deviation ratios are standardized for various classes of service for optimum use of the radio-frequency spectrum by a large number of stations. For example, FM broadcast signals typically have a 75-kHz deviation frequency, resulting in a transmitted signal with a bandwidth of approximately 150 kHz (twice the deviation frequency).

Because an FM signal is constant-envelope (that is, it has a constant amplitude versus time), FM radio systems are highly resistant to the amplitude-related interference problems that plague AM radio-frequency systems. Most FM receivers are designed to remove completely any amplitude variation existing on the received signal, since these amplitude variations are by definition due to degradations and perturbations impressed upon the transmitted FM signal as it traveled from the transmit to the receive antenna. Then, the FM signal is demodulated using a frequency-discriminator detector circuit that recovers the original modulating signal, which is unaffected by the amplitude variations experienced during transmission. Another property of FM receivers is the relative freedom from interference between distant and local stations using the same channel; only the strongest signal is received, even if the wanted signal is only 3–6 dB stronger than the interfering one. This characteristic is known as the capture effect. In contrast, AM radio signals differing in strength by 35 dB result in noticeable interference. *See* AMPLITUDE MODULATION; FREQUENCY-MODULATION DETECTOR.

Since most FM receivers are equipped with a muting circuit (squelch) that silences the audio channel when no signal is present, an irritating hiss noise is not emitted from the receiver when the frequency (or channel) being monitored is not in use.

Frequency modulation is used mainly for transmissions above 25 megahertz (MHz). Typical uses are in broadcasting; television sound; mobile radio telephony; radio paging systems; space telemetry; intercity microwave relaying of all classes of public traffic including voice channels, teleprinting, facsimile, broadcast network programs, and television and computer data; and intercontinental telecommunications via satellite. Frequency modulation is used for both analog and digital communications, and phase modulation (variation of carrier phase as a function of the modulating signal) as well as frequency modulation is used (**Fig. 1**). The result, with regard to output waveform, detection, and signal-to-noise improvement, is essentially the same whether phase-modulation (PM) or FM techniques are employed. *See* PHASE MODULATION; SIGNAL-TO-NOISE RATIO.

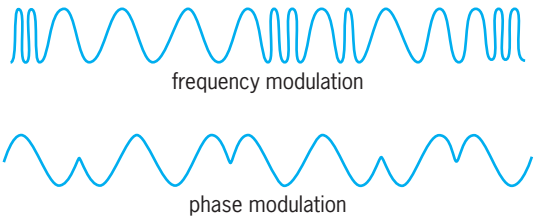

Fig. 1. Modulated waveforms.

FM broadcasting. The frequency band 88–108 MHz is allocated to FM broadcasting worldwide by international agreement. In the United States, this band is divided into 100 channels, with 200 kHz allocated for each channel. Each channel will support transmission of an analog audio signal with an audio frequency range of 50–15,000 Hz. Federal Communications Commission (FCC) rules allow for a maximum deviation of 75 kHz, and the emitted spectrum is twice this value. Because of the relatively small signal power in the modulating frequencies above 4 kHz, the received signal-to-noise ratio is improved substantially by preemphasis of the audio signal in transmission, necessitating complementary deemphasis in the receiver to restore natural program balance. In fact, preemphasis produces a sort of hybrid form of modulation, being pure frequency modulation at the lowest audio frequencies, and gradually changing to phase modulation at the highest. *See* FREQUENCY MODULATION.

Stereophonic broadcasting. Most FM stations broadcast two-channel stereophonic sound (that is, auditory perspective), where two microphones (left and right) are used in the recording of the audio source material. Stereophonic transmission must also provide normal reception for nonstereophonic (monaural) FM receivers. To accomplish this, the main channel modulation consists of combining, by additive polarities, the sum of the outputs from the left and right microphones. For stereophonic reception, a multiplexed channel is created which contains the L–R signal, amplitude-modulated onto a 38-kHz carrier signal. A 19-kHz pilot signal (38 kHz divided by 2) is sent with these L+R and L–R signals to allow for demodulation of the L–R signal (**Fig. 2**). Because the L+R and L–R signals are antiphased, high-volume peaks occur in one channel when the other is at its lowest. Each channel has access to almost half the total frequency deviation when its signal peaks to its highest volume. The available power is shared between the two, and the nonstereophonic receiver is unaffected by the multiplexed (L–R) channel and pilot.

FM subcarriers. The large channel bandwidth of FM broadcast stations permits simultaneous transmission of other services (in addition to the main-channel audio service) by use of so-called FM subcarriers. These are additional signals that are

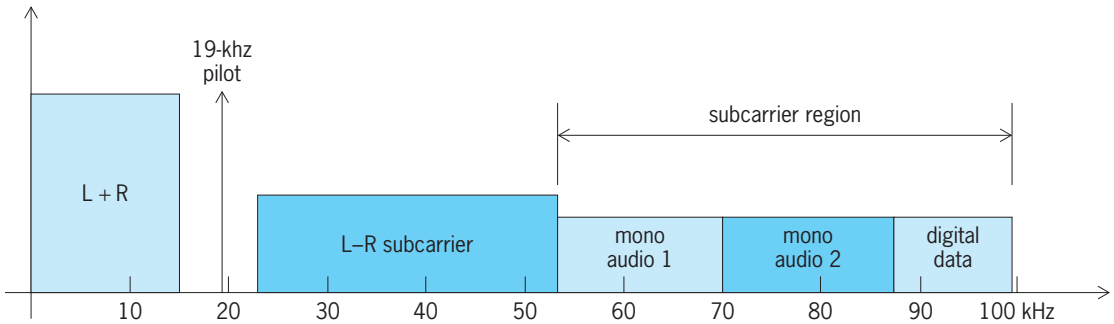

Fig. 2. FM subcarrier services.

frequency-division-multiplexed along with the main-channel audio, modulating at frequencies above the highest program frequencies. This is accomplished by frequency-division multiplexing; an example is shown in Fig. 2. FCC regulations require that the modulating signal input to an FM modulator (also called an FM exciter) be limited to the frequency band of 0–99 kHz, referred to as the FM baseband. A typical FM-baseband input is shown in Fig. 2, consisting of a main-channel audio signal (stereo, consisting of L+R, 19-kHz pilot, and L–R subcarrier) and, in this example three additional FM subcarriers, occupying frequencies from approximately 53 kHz to 99 kHz in the FM baseband. These subcarriers can be used for subscription or other special services. Normal FM broadcast receivers will not respond to these subcarrier signals, but special receivers are employed to select at least one of them, to the exclusion of the normal broadcast program. These multiplexed special emissions are permitted by the FCC under special communication authorization (SCA). *See* RADIO BROADCASTING; STEREOPHONIC RADIO TRANSMISSION.

In-band/on-channel (IBOC) digital radio. FM radio broadcasters in the United States were granted permission by the FCC in October 2002 to begin transmitting a digital radio signal along with their analog FM signal using a technique known as hybrid in-band/on-channel (IBOC) digital radio broadcasting (**Fig. 3**).

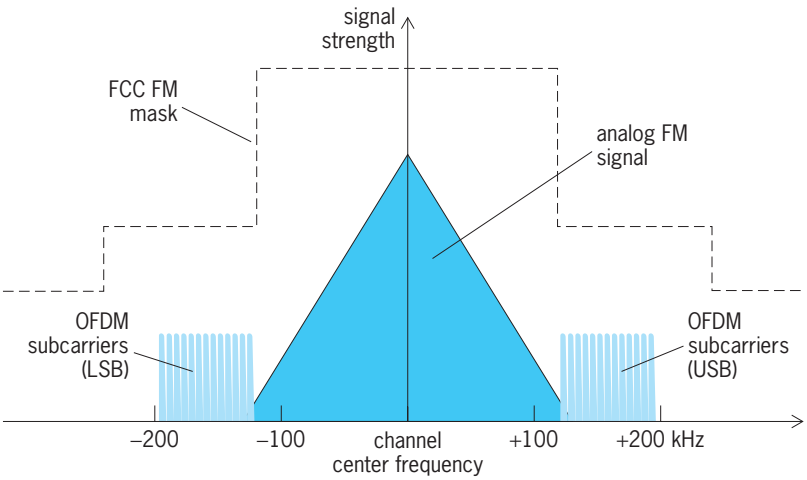

Fig. 3. IBOC digital radio signal (FM hybrid).

In an FM IBOC signal, the main-channel audio signal is broadcast (simulcast) in both the analog and digital portions, and the receiver uses the analog signal as a backup for the digital signal, which is less powerful and cannot be received as far away from the transmitter as can the analog FM signal. The digital portion of the IBOC signal uses a digital modulation technique called orthogonal frequency-division multiplexing (OFDM), which is a very robust technique favored for use by terrestrial digital broadcasting systems.

FM mobile transmission. Millions of land, maritime, aeronautical mobile, and portable FM transceivers are used by police, fire fighters, public safety agencies, industrial and commercial enterprises, private citizens, and radio amateurs who desire the benefits of en route telephony. Prior to the widespread use of commercial cellular and personal communications services, the intensity of private two-way radio use was growing exponentially, mainly because of the availability of reliable, small, low-power-consumption solid-state equipments that are economical, and also because of the public realization of the benefits of having such communications. The rapid increase in such usage became a major concern at the time to managers of the radio-frequency spectrum when it became apparent that all available channels would be assigned in major city areas.

The channel width assigned for mobile radio telephony has been reduced in some bands to 15 kHz (12.5 kHz in the United Kingdom), from the wide-band frequency modulation initially employed, and state-of-the-art receivers provide intermediate-frequency (IF) selectivities of -100 dB for ± 15-kHz channel spacings. Additional bands have been allocated for mobile radio (800 MHz), and new techniques are being investigated as an alternative to FM, such as single-sideband (SSB), spread-spectrum, stored speech, and the virtual elimination of speech by the use of data transmission in those applications where standard forms of message predominate. SSB transmission, employing linear compression and expansion to combat the adverse effects of fading on SSB systems, has demonstrated that a minimum channel separation of about 6 kHz is possible compared with 15 kHz for FM transmission. *See* SINGLE SIDEBAND.

Typically, mobile telephony is intermittent, each station transmitting only to send. Receivers are open

to receive incoming calls on their channel, or the receiver is activated only when certain signaling (touch or subaudible) tones are received. Simplex transmission is the rule; that is, by means of a push-to-talk button on the handset or microphone, the transmitter is switched on to talk. Many stations can share a channel, each being informed of when the channel is clear.

Mobile transmitters have output powers of 10–50 W. Hand-held units employing powers of 1–6 W are also used. Base stations employ powers up to 250 W. Vertical polarization is used for FM radio telephony, since a mobile whip, which uses the vehicle as a ground plane, is the most practical antenna configuration. The geographical range of such stations is a function of the transmitting and receiving antenna heights. The communications range of mobile units with antenna heights of 6 ft (1.8 m) is on the order of 6 mi (3.7 km), with the range increasing to 17 mi (27 km) for mobile communications with a base station with an antenna 100 ft (30.5 m) above ground level. Favorable (flat) terrain will tend to increase communications range, while hilly or mountainous terrain and heavily treed areas (for example, tropical forests) will tend to decrease range. The use of repeaters favorably located on the top of high-rise buildings, on the top of hills, mountains, or tall towers increases the useful range for mobile-to-repeater communications to distances of 60 mi (100 km) and beyond. In the ideal situation, therefore, mobile-to-mobile communication by way of a repeater station to twice this distance is possible. The repeater receives the input signal on one frequency and rebroadcasts it on another frequency, with higher transmitter power or high antenna gain, or both. Amateur-radio 2M repeaters employ a frequency spacing of 600 kHz between input and output frequencies, necessitating high-Q cavity duplexers so that the receiver and transmitter can employ a common high-gain antenna. The transmission signal must be attenuated by more than 100 dB, even with adequate receiver IF selectivity, since the receiver must respond to signals as small as 0.2 microvolt or so in the presence of a transmission signal of tens of volts.

With increasing occupancy of the mobile radio bands by many nearby users operating on equally spaced channels, problems of intermodulation are a major limitation to further expansion in some congested areas. Intermodulation occurs when two or more frequencies combine external to, but usually in, the receiver front end to produce a difference frequency that is the same as the desired frequency. Typically, communication receivers experience intermodulation problems when the combination frequencies are about 80 dB above the receiver threshold, although improved performance is possible with state-of-the-art design.

Radio repeaters have become very sophisticated, with tone burst, subaudible tones which can be used for remote control of the repeater functions, linking, and so on, and push-button autopatch (automatic connection and push-button dialing into the public switched telephone system). The numbers of amateur-radio repeaters (simple or complex) have grown at a rate that exceeds development of commercial repeaters, and there are one or more repeaters serving most cities and surrounding areas in the United States and Canada. *See* MOBILE RADIO.

Cellular systems. Cellular systems (both analog and digital) employ equipment operating in the 800-MHz and 2000-MHz bands. The mobile user can move from cell to cell, and these movements are kept track of, whether the user is monitoring, initiating, or receiving a radio-to-telephone interconnect call. Such systems must be oriented toward microprocessor control if all the functions proposed are to be implemented. Cellular and trunking radio systems have a great deal in common. Analog modulation schemes like AMPS (advanced mobile phone system) and TACS (total-access communications system) use frequency modulation, while digital schemes like TDMA (time-division multiple access) and CDMA (code-division multiple access) use variants of phase modulation.

Radio paging service. These paging transmitters, which employ high power, a 600-W centrally located transmitter, or several 60-W transmitters frequency-synchronized to within a few hertz throughout the service area, so that reception is possible employing shirt-pocket-sized receivers and inefficient antennas inside buildings and automobiles. Tone, numeric, and text pagers are used. The sender activates the pager by telephone, modem link, or the Internet. Tone pagers signal the user to take a prearranged action. In numeric and text paging systems, the sender transmits a callback telephone number or a short text message to the user. The portable receiver or pager, selectively called by a digital or dual-tone-multifrequency (DTMF) code sequence, receives the message. The numeric or text message is displayed on a liquid crystal display (LCD). *See* ELECTRONIC DISPLAY; RADIO PAGING SYSTEMS.

Radio relaying. Frequency modulation is used for long-haul microwave radio relaying over land, over water, and to great distances using satellites, sometimes carrying thousands of simultaneous telephone conversations or several television channels. The advent of requirements for short-haul services, local distribution networks within cities, television relay, and a wide variety of optical communication services—including high-speed computer communications, electronic mail, data transmission—and other services, where it may be cost-effective to avoid the local telephone loop, presents another application for FM radio relaying.

Frequency modulation is of special importance in this application because changes to the signal amplitude along the propagation path, due to fading, do not affect reliability as significantly as with some other modulation schemes. Although the background noise increases as the transmissions fade, the signal level at the output of the receiver limiter remains steady, a highly essential characteristic in a communication system. High-efficiency amplifiers

can be used on FM signals since linearity is not critical as with amplitude modulation (AM) and digital modulation.

In microwave communications, frequencies above 1 GHz and to about 40 GHz are used for terrestrial and satellite relaying. By these means, hundreds of telephone, teleprinter, television, and computer data channels of communications can be transmitted on a single microwave carrier. In large systems, many carriers operating in parallel over the same route are used to achieve increased capacity.

High communications capacities are obtained by multiplexing, using large bandwidths. Two types of frequently employed carrier modulation are analog frequency-division multiplexing-frequency modulation (FDM/FM), which is more common, and digital phase-shift keying (PSK), which is growing in use. In digital communications the analog signal is sampled at some predetermined rate, and the amplitude is quantized at this sample interval by an analog-to-digital converter. The main advantage of digital transmission consists in the possibility of overcoming noise and interference that are caused by human activities by means of special coded signals and by reception techniques which are optimum for the given conditions. In uncoded systems the signal-to-noise ratio increases linearly with increased signal bandwidth, whereas in coded systems the signal-to-noise ratio rises exponentially with increased signal bandwidth. This can be explained as follows: Doubling of the spectrum width means doubling the number of pulses (or switch transitions) in the code combination, which, according to the rules that govern combinations, squares the number of signals which can be transmitted. *See* ANALOG-TO-DIGITAL CONVERTER.

When PSK or FSK (frequency-shift keying) are used with coherent reception, high frequency and phase synchronization are essential. This is a complicating requirement for this form of modulation, particularly since the PSK signal does not normally contain a component at the carrier frequency, and thus some method of synchronization is necessary.

Radio and television broadcasting using digital communications channels is possible. Although such systems are based on the principles of speech transmission, the system parameters are different, owing to the wider band of continuous signals and the necessary increase in the number of quantizing levels. Sophisticated high-speed digital communications equipments employ not just two-level PSK (0 to 180°) but quadraphase-shift-keying (QPSK) modulation (0 or 180° as well as 90 or 270°) and octal phase-shift keying (45° between states), and thus provide nearly double or quadruple the transmission capacity.

Most systems employ analog FDM/FM. For example, a normal telephone signal occupying a range of 300–3400 Hz is used to modulate a suppressed carrier, commonly between 60 and 108 kHz. By means of electronic filters, only a single sideband is selected, this being the original signal translated to a higher frequency. These subcarriers are spaced at 4 kHz, so that 12 conversations can fill the 60–108-kHz range. In turn, the 12 channels are multiplexed with other groups of 12, to form a supergroup of 60 channels, similarly transposed. Master groups of 600, 980, 1200, 1800, or 2700 channels are further extensions of the principle.

Finally, the carrier is frequency-modulated by this aggregate of individual channels to occupy a frequency band that may reach 8 MHz. The carrier peak-deviation ratio is fractional for overland relaying. Typical deviation per telephone channel is 2 kHz. Larger per-channel deviations are used for troposcatter systems and in satellite communications. The Canadian Anik satellite system has a maximum capacity of 4800 two-way voice circuits or 10 television channels. The American *Intelsat IVA* satellite was designed to have a capacity of about 10,000 two-way voice circuits or 20 television channels. *See* COMMUNICATIONS SATELLITE; TELEPHONE SERVICE.

Video recording. In VHS video recording, the black-and-white (luminance) portion of the analog video signal is frequency-modulated. The response of magnetic video tape is nonlinear over the 3-MHz VHS video bandwidth, and FM minimizes the effects of distortion introduced during the recording process.

Telegraphy. Telegraphy, including teleprinting and binary digital data transmission, is based on frequency-shift keying (FSK), a form of FM modulation. In FSK, the modulating signal shifts the carrier frequency or its phase between two limiting values, one of which represents a mark signal and the other a space signal. This frequency shift (or phase shift) is a form of FM signaling used over a wire, cable, or radio. A common approach is to use multiplexed audio tones, each deviated a few tens of cycles (or shifted in phase between two discrete levels), to carry individual messages. A normal voice channel can carry 16–24 such telegraph channels operating at standard teleprinter speeds. High-speed data transmission requires larger bandwidths and more elaborate equipment, but nevertheless is based on the same principle of frequency or phase shift. *See* TELEGRAPHY.

Facsimile. Black-and-white images (line drawings and typed copy) can be transmitted by employing the principles used in FM telegraphy; one limit frequency corresponds to black, and the other to white, on the image to be transmitted. It is necessary only that the original document be scanned (for example, by a photosensitive cell), that the scan at the recorder be synchronized with that at the transmitter, and that some means be employed—for example, using an electrosensitive or electrostatic (dry) process—to make the recording paper black when a black signal is received and leave it white for a white signal. A continuous gray scale can be transmitted and recorded if, instead of just two frequencies, a continuous frequency shift is employed between some low frequency (say, 1500 Hz) and some higher frequency (say, 2700 Hz), the exact frequency at any instant being proportional to the gray level of the image. The recorder employs a frequency- amplitude transducer, the response of which is the inverse of

that at the transmitter. More complicated modulating arrangements allow color images to be scanned and transmitted to receivers capable of color reproduction. The audio-frequency tones can be sent over telephone lines; these tones can modulate an FM radio transmitter, or they can modulate an SSB transmitter. *See* FACSIMILE.

Telemetry. Prior to digital communications, frequency modulation was the preferred method for transmission of information or data from a remote or inaccessible location such as a rocket vehicle in flight. Each condition to be remotely observed would actuate one subchannel, which, when multiplexed with other channels reporting other status conditions, modulated the radio carrier by frequency modulation. Telemetry is not much different in principle from FM and PM telegraphy, except that it is a one-way transmission system, with each channel containing a transducer which converts sensed phenomena into electrical signals. The signals may be either digitized coded pulses or continuous (analog) changes. At the receiver, each channel is suitably instrumented to record directly in real time the information reported by the transducer on that channel; or the FM signal can be frequency-translated and recorded directly on magnetic tape for subsequent analysis. Frequency or phase modulation is very useful for digital data recording, since, for example, phase modulation is effective against varying oxide thicknesses on the magnetic tape and varying signal strength. Telemetry systems exist in all degrees of sophistication, up to hundreds of data channels. *See* TELEMETERING.

John S. Belrose; David H. Layer; Robert W. Denny, Jr.

Bibliography. E. Armstrong, Radio Signaling System, U.S. Patent No. 1,941,066, Dec. 26, 1933; A. B. Carlson, P. B. Crilly, and J. C. Rutledge, *Communications Systems*, 4th ed., 2001; J. G. Proakis and M. Salehi, *Communication Systems Engineering*, 2d ed., Prentice Hall, 2001; M. Schwartz, *Information Transmission, Modulation and Noise*, 4th ed., 1990; H. Taub, *Principles of Communication Systems*, 2d ed., 1986; W. Tomasi, *Electronic Communication Systems: Fundamentals Through Advanced*, 3d ed., 1997; M. E. Van Valkenburg and W. M. Middleton (eds.), *Reference Data for Engineers: Radio, Electronics, Computer, and Communications*, 9th ed., 2002.

Frequency modulator

An electronic circuit or device producing frequency modulation. This device changes the frequency of an oscillator in accordance with the amplitude of a modulating signal. If the modulation is linear, the frequency change is proportional to the amplitude of the modulating voltage.

High-frequency oscillators usually employ either *LC* (inductance-capacitance) tuned circuits or piezoelectric crystals to establish the frequency of oscillation. This frequency can be controlled by changing the effective capacitance or inductance of the tuned circuit in accordance with the modulating signal. A simple example of this type of modulator is a capacitor microphone used as a capacitive tuning element in an oscillator. The acoustic waves that strike the diaphragm of the microphone change the capacitance, which in turn changes the oscillator frequency. The acoustic pressure are usually so small that the capacitance change and the frequency change, or frequency deviation, are linearly related to the acoustic pressures. However, this modulation method is cumbersome and seldom used. Practical circuits usually employ a varactor diode to change the oscillator in accordance with a modulating voltage.

Varactor modulator. The oscillators in high-frequency electronic systems, such as frequency-modulating (FM) transmitters, usually employ piezoelectric crystals for precise control of the carrier frequency. These crystals are equivalent to a series *LC* tuned circuit with an extremely high *Q*. The crystal holder has a small capacitance which is in parallel with the crystal and therefore causes parallel resonance at a slightly higher frequency than the series resonant frequency of the crystal. The actual oscillator frequency is between these two resonant frequencies and is controllable by the parallel capacitance. *See* PIEZOELECTRICITY; Q (ELECTRICITY).

The junction capacitance of a semiconductor diode varies with the diode voltage, and a reverse-biased diode may be used to control the oscillator frequency to produce frequency modulation. Low-loss diodes designed for this service are known as varactor diodes and have trade names such as Varicaps or Epicaps. A basic varactor modulating scheme is shown in **Fig. 1**. In this circuit, the transistor that drives the varactor modulator provides reverse bias as well as the modulating voltage v_m. The radio-frequency (rf) choke provides very high impedance at the oscillator frequency to isolate the transistor amplifier output impedance from the oscillator circuit but to allow the modulating signal to pass through with negligible attenuation. Only the frequency-determining part of the oscillator is shown. The symbols C_c, L_c, and R represent the electrical equivalents of the compliance, mass, and loss,

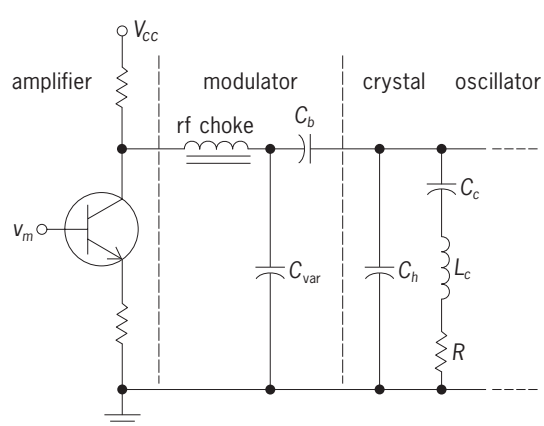

Fig. 1. Basic varactor modulator circuit. V_{CC} = collector supply voltage.

respectively, of the crystal; C_b is the crystal-holder capacitance and C_b is a dc blocking capacitor. *See* SEMICONDUCTOR DIODE; VARACTOR.

The capacitance of a typical varactor diode is inversely proportional to the square root of the reverse-bias voltage. This relationship may be expressed as in Eq. (1), where C_{var} is the varactor capacitance, $C_{-1\text{V}}$

$$C_{\text{var}} = \frac{C_{-1\text{V}}}{(-v_d^{1/2})} \tag{1}$$

the varactor diode capacitance at 1 V reverse bias, and v_d the bias voltage of the diode. Since v_d is negative for reverse bias, the quantity in parentheses is positive. This relationship is nonlinear, and even if it were linear the frequency deviation would not be linearly related to the modulating voltage, because the oscillator frequency is not linearly related to the tuning capacitance. The modulation distortion may be low for small frequency deviations. Fortunately, the distortion of the modulation system may be reduced to desirable values at normal frequency deviations by using negative feedback in the system. This feedback may be accomplished by coupling a linear FM demodulator such as a phase-locked loop (PLL) to the oscillator and feeding the voltage output back to an appropriate point in the driving amplifier where the feedback voltage has opposite polarity to the modulating-signal voltage. *See* FEEDBACK CIRCUIT; PHASE-LOCKED LOOPS.

The varactor-diode modulator is also commonly used to control the frequency of local oscillators in radio receiving equipment where programmed, push-button, or remote tuning is desirable. In these applications, conventional *LC* tuned circuits may be used.

Voltage-controlled oscillator modulator. Oscillators having a linear relationship between their output frequency and a controlling voltage have been developed for phase-locked loops. Known as voltage-controlled oscillators (VCOs), they are available in integrated circuit form as a separate chip or as a separately usable section of an integrated phase-locked loop.

The voltage-controlled oscillator is similar to an astable multivibrator. The main difference is that emitter-circuit resistances of a typical emitter-coupled multivibrator have been replaced by current sources (**Fig. 2**), which provide charge and discharge currents to the timing capacitor C_f. They are usually transistors with fixed values of collector current. In order to understand how the circuit operates, it may be assumed that transistor T_1 is cut off and I_1 is therefore the only current flowing through the timing capacitor C_f. Then the voltage across C_f changes linearly with time t at the rate given by Eq. (2). The emitter

$$v_{cf} = \frac{1}{C_f}\int_0^t I_1 dt = \frac{I_1}{C_f}t \tag{2}$$

of T_1 becomes negative with respect to the emitter voltage of T_2, which is fixed by the voltage divider R_3 and $(R_{c1} + R_4)$ between V_{CC} and the control voltage. During this time the base potential of transistor

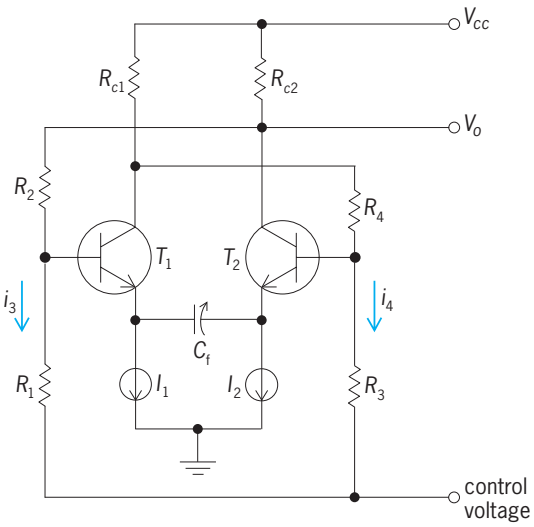

Fig. 2. Simple voltage-controlled oscillator circuit.

T_1 is low because transistor T_2 is either at or near saturation, the current i_3 is small, and the base potential of T_1 is near the control voltage, holding T_1 in cutoff condition. Eventually, the emitter potential of T_1 becomes lower than the base potential and T_1 begins to conduct. Therefore, the current through C_f decreases and the collector potential of T_1 drops while the collector potential of T_2 rises, thus causing the base potential of T_1 to rise while the base potential of T_2 falls toward the control voltage. This action cuts off transistor T_2 while T_1 is either at or near saturation and current I_2 flows through C_f from left to right, causing the emitter potential of T_2 to decrease linearly until T_2 begins to conduct and the switching begins. The switching action takes place very quickly and the output voltage is essentially a square wave; the voltage at each emitter is a triangular wave. As the control voltage becomes more positive, the voltage change across C_f required to cause the switching action is reduced. Therefore, the frequency of the voltage-controlled oscillator increases in direct proportion to the change in control voltage as shown in Eq. (2). *See* FREQUENCY MODULATION; MULTIVIBRATOR; OSCILLATOR. Charles L. Alley

Bibliography. H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, 1980; A. P. Malvino, *Electronic Principles*, 6th ed., 1998; C. A. Schuler, *Electronics: Principles and Applications*, 5th ed., 1998.

Frequency multiplier

An electronic circuit that produces an output frequency which is an integral multiple of the input frequency. There are three basic types of frequency multipliers. The first type is a nonlinear amplifier that generates harmonics in its output current and a tuned load that resonates at one of these harmonics. The second type uses the nonlinear capacitance of a junction (semiconductor) diode to couple energy from the input circuit, which is tuned to the

fundamental frequency, to the output circuit, which is tuned to the desired harmonic. The third type of multiplier uses a frequency divider in a locked-oscillator circuit.

Nonlinear amplifier. Harmonic-generating amplifiers are usually operated class C, biased so that the output current flows for only a small part of the input cycle, perhaps 90°. The output is then richer in harmonics, and higher efficiency is attained than with normal class C bias. The output circuit may be tuned to any harmonic, but the higher harmonics are weaker and the power output may be less than the power input if the selected harmonic is higher than the third or fourth. *See* AMPLIFIER; HARMONIC (PERIODIC PHENOMENA).

A highly efficient doubler can be devised by using two amplifiers, such as transistors, with their inputs driven with opposite polarity, obtained from opposite ends of a center-tapped coil, and their outputs connected in parallel. These amplifier-type frequency multipliers are often used in crystal-controlled high-frequency transmitters. The crystal oscillator is desirable because of its excellent frequency stability, but crystals are available for only a limited frequency range. Therefore, a moderate crystal-oscillator frequency is multiplied to a high output frequency.

Nonlinear coupler. Diode-type multipliers are usually used in very high-frequency or ultrahigh-frequency solid-state transmitters. The frequency limit at which a transistor has useful power gain is highly dependent upon the power output capability of the transistor. Varactor-diode multipliers are often used to provide an output frequency for the transmitter that is several times as high as the frequency limit of the power transistor or transistors in the transmitter. *See* TRANSISTOR.

Efficiency is of prime importance in this type of multiplier because the transmitter power output is the power output of the final transistor amplifier minus the power loss in the multiplier. Fortunately, the efficiency of varactor-diode multipliers may be as high as 80–90% if the circuit and diode Qs are high and the multiplication in each stage is no greater than two or three. The multipliers can be advantageously connected in cascade if higher multiplication ratios are desired. A typical varactor-diode multiplier circuit is shown in **Fig. 1**. Coil L_1 and capacitor C_1 resonate near the input frequency, while coil L_2 and capacitor C_2 resonate near the output frequency. *See* SEMICONDUCTOR; VARACTOR.

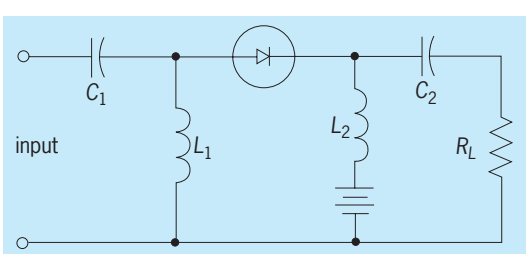

Fig. 1. Varactor-diode multiplier.

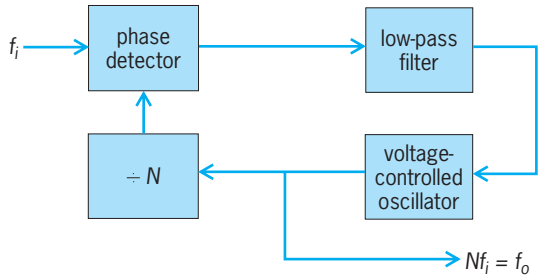

Fig. 2. Locked-oscillator frequency multiplier.

Locked-oscillator type. A convenient technique for multiplying a signal frequency by any reasonable integer N uses a frequency divider in a locked-oscillator circuit (**Fig. 2**). This circuit may consist simply of a phase-locked loop and a digital counter, both of which are available off-the-shelf items for output frequencies that do not exceed the capability of available high-speed counters and phase-locked loops. The input signal usually comes from a crystal-controlled oscillator because of the frequency stability of these oscillators.

Since the signal frequencies must be the same at the two phase-detector inputs (Fig. 2), the frequency of the voltage-controlled oscillator must be N times as high as the input-signal frequency f_i. Therefore the voltage-controlled oscillator provides the output f_o for the multiplier. This versatile type of frequency multiplier is commonly used in radio transmitters and receivers where the digital counter may be used to provide push-button selection of the communication-channel frequency. *See* DIGITAL COUNTER; FREQUENCY DIVIDER; PHASE-LOCKED LOOPS; RADIO RECEIVER; RADIO TRANSMITTER.

Charles L. Alley

Bibliography. H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, 1980; A. P. Malvino, *Electronic Principles*, 6th ed., 1998; C. A. Schuler, *Electronics: Principles and Applications*, 5th ed., 1998.

Freshwater ecosystem

A physiologically interactive system within which growth, adaptation, species composition, nutrient cycling, biological productivity, and energy flows among inland aquatic microbial, plant, and animal communities are integrated with the environment. These inland waters include lakes, reservoirs, rivers, streams, and wetlands. The discipline termed limnology evaluates how these relationships regulate, and are regulated by, natural and anthropogenic variations in the physical, chemical, and biotic environments. The standing-water (lentic) lake or reservoir ecosystem is intimately coupled with the land in its surrounding drainage area and with its running (lotic) waters that change chemical components from the land as they are transported to the lake. Often the land-water interface supports

TABLE 1. Water in the biosphere*

	Volume, 10^3 km^3	Percent of total	Renewal time
Oceans	1,370,000	97.61	3100 years†
Polar ice, glaciers	29,000	2.08	16,000 years
Ground water (actively exchanged)‡	4,000	0.295	300 years
Freshwater lakes	125	0.009	1–100 years§
Saline lakes	104	0.008	10–1000 years§
Soil and subsoil moisture	67	0.005	280 days
Rivers	1.2	0.00009	12–20 days¶
Atmospheric water vapor	14	0.0009	9 days

*Modified from R. G. Wetzel (1983, 1999) after Kalinin and Bykov. Slightly different values are given by Shiklomanov, but ratios are similar.
†Based on net evaporation from the oceans.
‡Kalinin and Bykov estimated that the total ground water to a depth of 5 km in the Earth's crust amounts to 60×10^6 km^3. This is much greater than the U.S. Geological Survey estimate of 8.3×10^6 km^3 to a depth of 4 km. Only the volume of the upper ,actively exchanged ground water is included here.
§Renewal times for lakes vary directly with volume and mean depth, and inversely with rate of discharge. The absolute range for saline lakes is from days to thousands of years.
¶Twelve days for rivers with relatively small catchment areas of less than 100,000 km^2; 20 days for major rivers that drain directly to the sea.

productive wetlands, that is, areas of shallow water often found at the edge of lakes (littoral zone) or along rivers (floodplain). *See* ECOLOGY; LIMNOLOGY.

Ecosystem structure. Most water on the Earth's surface is saline in oceans (97.61%), and most of the remainder is polar ice (2.08%) and ground water (0.29%). Only 0.009% exists in freshwater ecosystems as temporary storage reservoirs on land with renewal times much shorter than those of the oceans (**Table 1**). The volume of water flowing from these reservoirs to the sea in rivers is very small (0.00009%) with a mean residence time of about 2 weeks. During high flow periods, surface water often recharges adjacent ground-water aquifers. Return flows from ground water to rivers occur during periods of low flow or drought, and usually maintain a base flow in river channels. *See* RIVER.

Water storage and retention times are altered by shifts in the balance between water inputs and losses. Lakes receive water from precipitation directly on the surface, from surface inflows from the drainage basin, and from subsurface ground-water seepage. Lakes lose water by flow from an outlet (drainage lakes), seepage through the basin walls into the ground water (seepage lakes), evaporation, and evapotranspiration from higher aquatic plants. Saline lakes occur in closed basins with no outflow except by evaporation.

Lakes are concentrated in the subarctic and temperate regions of the Northern Hemisphere, and reservoirs are concentrated in the subtemperate and subtropical zones. Although some 40% of the total volume of fresh water is contained in seven great lake basins of Siberia, North America, and eastern Africa, most lakes and reservoirs are very much smaller. Most of the millions of lakes and reservoirs are small and relatively shallow, usually less than 15 m (50 ft) in depth (**Fig. 1**). Mean depths are even shallower (less than 5 m or 17 ft), and as a result often over half of the sediments occur within the photic zone; therefore much of the sediments of lakes and reservoirs is potentially capable of supporting photosynthesis. Similarly, major flowing-water ecosystems occur in low-gradient tropical and subtropical regions, and they are posited to support high photosynthetic productivity in extensive floodplain and land-water interface regions. *See* LAKE; PHOTOSYNTHESIS.

The productivity and internal metabolism of aquatic ecosystems are driven and controlled by energy from solar radiation acquired by photosynthesis. Inland waters receive organic products of photosynthesis directly from their aquatic flora, and indirectly from their drainage basins as particulate and dissolved organic matter from terrestrial and wetland plants imported by stream water, storm runoff, and ground water.

In addition, solar energy is absorbed and dissipated as heat. As surface waters warm, the upper layers

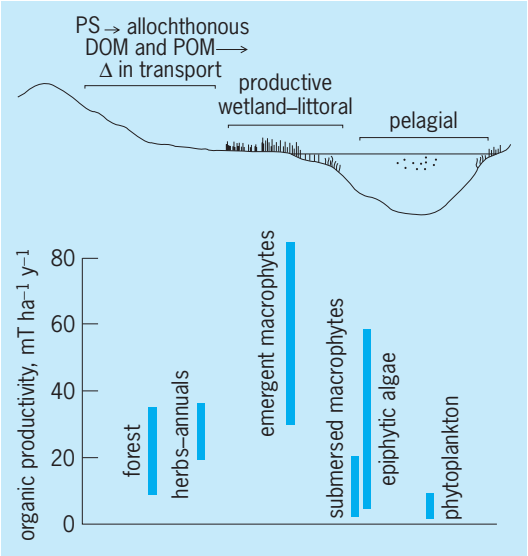

Fig. 1. Lake ecosystem showing the drainage basin with terrestrial photosynthesis (PS) of organic matter, movement of nutrients and dissolved (DOM) and particulate organic matter (POM) in surface and ground-water flows toward the lake basin, and chemical and biotic alteration of these materials en route, particularly as they pass through the highly productive and metabolically active wetland-littoral zone of the lake per se (net organic productivity in metric tons per hectare per year). (*Modified from R. G. Wetzel, 1983*)

TABLE 2. Mean composition of river waters of the world (mg L^{-1})*

	Ca^{+2}	Mg^{+}	Na^{+}	K^{+}	CO_3^{-2} (HCO_3^{-})	SO_4^{-2}	Cl^{-}	NO_3^{-}	Fe (as Fe_2O_3)	SiO_2	Sum
North America	21.0	5.	9.	1.4	68.	20.	8.	1.	0.16	9.	142
South America	7.2	1.5	4.	2.	31.	4.8	4.9	0.7	1.4	11.9	69
Europe	31.1	5.6	5.4	1.7	95.	24.	6.9	3.7	0.8	7.5	182
Asia	18.4	5.6	5.5	3.8	79.	8.4	8.7	0.7	0.01	11.7	142
Africa	12.5	3.8	11.	—	43.	13.5	12.1	0.8	1.3	23.2	121
Australia	3.9†	2.7	2.9	1.4	31.6	2.6	10.	0.05	0.3	3.9	59
World	15.	4.1	6.3	2.3	58.4	11.2	7.8	1.	0.67	13.1	120
Cations (μeq L^{-1})	750	342	274	59	—	—	—	—	—	—	1425
Anions	—	—	—	—	958	233	220	17	—	—	1428

*From R. G. Wetzel (1983, 1999) after data from several sources.
†Values of calcium are likely less, on the average, than sodium and magnesium in Australian surface waters.

are less dense and overlie deeper, cooler water (the maximum density of water occurs at a temperature of 4°C or 39°F; warmer water is less dense). This stratification generally persists until seasonal insolation (exposure to the Sun's rays) is reduced and water temperatures and densities of the different strata equalize. Density-induced stratification influences physical, chemical, and biological properties of lakes and reservoirs, including the thermal structure, water mass stratification, and hydrodynamic movements. Heat is distributed and altered by the physical work of wind energy, currents and other water movements, basin morphometry, and water losses. These characteristics of warmer, less dense water overlying cooler, more dense strata for long periods of time (months) structure the aquatic habitats and have marked attendant effects on all chemical cycles, metabolic rates, and population dynamics and their productivities.

The most abundant dissolved ions, constituting total salinity, in inland waters are usually four major cations [calcium (Ca^{+2}), magnesium (Mg^{+2}), sodium (Na^{+}), potassium (K^{+})] and four anions [bicarbonate (HCO_3^{-}), carbonate (CO_3^{-2}), sulfate (SO_4^{-2}), chloride (Cl^{-})]. The salinity of freshwaters has a world average concentration of about 120 mg/L, but varies among continents and with the lithology of land masses (**Table 2**).

The composition of cellular protoplasm of organisms is dominated by six elements (carbon, hydrogen, oxygen, nitrogen, phosphorus, sulfur). Biogeochemical cycles control the availability of phosphorus, nitrogen, and several minor nutrients, and these elements often limit biotic development. Phosphorus, in comparison with other macronutrients required by biota, is least abundant and commonly the first element to limit biological productivity. When phosphorus is in adequate supply, nitrogen availability invariably limits productivity (**Table 3**). Standing inland waters have been categorized into

TABLE 3. General ranges of primary productivity of phytoplankton and related characteristics of lakes of different trophic categories*

Trophic type	Mean primary productivity, mg C m^{-2} day^{-1}†	Phytoplankton density, cm^3 m^{-3}	Phytoplankton biomass, mg C m^{-3}	Chlorophyll *a*, mg m^{-3}	Dominant phytoplankton	Light extinction coefficents, η m^{-1}	Total organic carbon, mg L^{-1}	Total P, μg L^{-1}	Total N, μg L^{-1}	Total inorganic solids, mg L^{-1}
Ultraoligotrophic	<50	<1	<50	0.01–0.5		0.03–0.8		<1–5	<1–250	2–15
Oligotrophic	50–300		20–100	0.3–3	Chrysophyceae Cryptophyceae	0.05–1.0	<1–3			
Oliogomesotrophic		1–3			Dinophyceae Bacillariophyceae			5–10	250–600	10–200
Mesotrophic	250–1000		100–300	2–15		0.1–2.0	<1–5			
Mesoeutrophic		3–5						10–30	500–1100	100–500
Eutrophic	>1000		>300	10–500	Bacillariophyceae Cyanobacteria	0.5–4.0	5–30			
Hypereutrophic		>10			Chlorophyceae Euglenophyceae			30 to >5000	500 to >15,000	400–60,000
Dystrophic	<50–500		<50–200	0.1–10		1.0–4.0	3–30	<1–10	<1–500	5–200

*From R. G. Wetzel (1983, 1999) after many authors and sources.
†Referring to approximately net primary productivity, such as measured by the ^{14}C method.

various trophic scales on the basis of ranges of major nutrients and algal productivity (Table 3). *See* BIOGEOCHEMISTRY; BIOLOGICAL PRODUCTIVITY.

Organic matter is either soluble or particulate; once organic matter enters an aquatic ecosystem, it tends to remain there. Oxygen, however, is quite insoluble in water, and most oxygen produced in an aquatic system will tend to diffuse to the atmosphere. When organic matter in lakes is utilized and biochemically respired by heterotrophic organisms, the amount of organic matter to be oxidized can be much larger than the oxygen available for oxidation. In large lakes, inputs of organic matter from internal photosynthesis and from external sources are usually small relative to the large amount of dissolved oxygen in the water. In contrast, small lakes have a greater proportion of water supporting photosynthesis in wetland and littoral (pertaining to shores of lakes, seas, or oceans) areas, import proportionally more organic materials from terrestrial habitats, and have a smaller pool of dissolved oxygen. Thus, respiration can exhaust oxygen dissolved in the lake water in lower strata that are poorly connected hydrodynamically with the atmosphere. The exhaustion of dissolved oxygen in large portions of lake ecosystems is exacerbated by the process of eutrophication (the increased photosynthetic production of organic matter in response to excessive loading of nutrients, particularly phosphorus).

Once the dissolved oxygen is consumed, other dissolved ions [nitrate (NO_3^{-1}), sulfate (SO_4^{-2}), iron (Fe^{+2}), manganese (Mn^{+2}), and organic compounds] can function as alternative electron acceptors under anoxic conditions. These processes are less efficient energetically than oxidative decomposition, and the rates and efficacy of decomposition decline under anaerobic conditions, such as in anoxic lower strata of productive lakes.

Diversity of habitats. Freshwater ecosystems consist of entire drainage basins. The nutrient and organic matter content of drainage water from the catchment area is modified as water moves downgradient to and within the lake or reservoir per se (**Fig. 2**). Autotrophic photosynthetic productivity is generally low to intermediate in the terrestrial components, highest in the wetland interface region between the land and water, and lowest in the open water (Fig. 2). Along the gradient from land to river channels, the greatest productivity is also in the marginal floodplain regions. Autotrophic productivity in river channels is generally low as in the pelagic (open water, nonlittoral) regions of lakes. Most of the organic matter utilized in running waters is imported from floodplain and terrestrial sources.

The land-water interface region of aquatic ecosystems is always the most productive per unit area along the gradient from land to open water of both lakes and reservoirs. Because most aquatic ecosystems occur in geomorphologically mature terrain of gentle slopes and are small and shallow, the wetland-littoral components usually dominate in productivity and in synthesis of organic matter. The region of greatest productivity is the emergent macrophyte zone. Emergent aquatic plants have structural and physiological adaptations that not only tolerate the hostile reducing anaerobic sediments but exploit the high nutrient and water availability of this habitat. Nutrients entering the zone of emergent macrophytes tend to be assimilated by microflora of the sediments and detritus particles, and are then recycled to the emergent macrophytes. Export from the emergent zone is dominated by dissolved organic compounds released from decomposition of plant detrital material. Nutrients, particularly nitrogen and phosphorus, tend to be conserved and retained in the wetland-littoral regions.

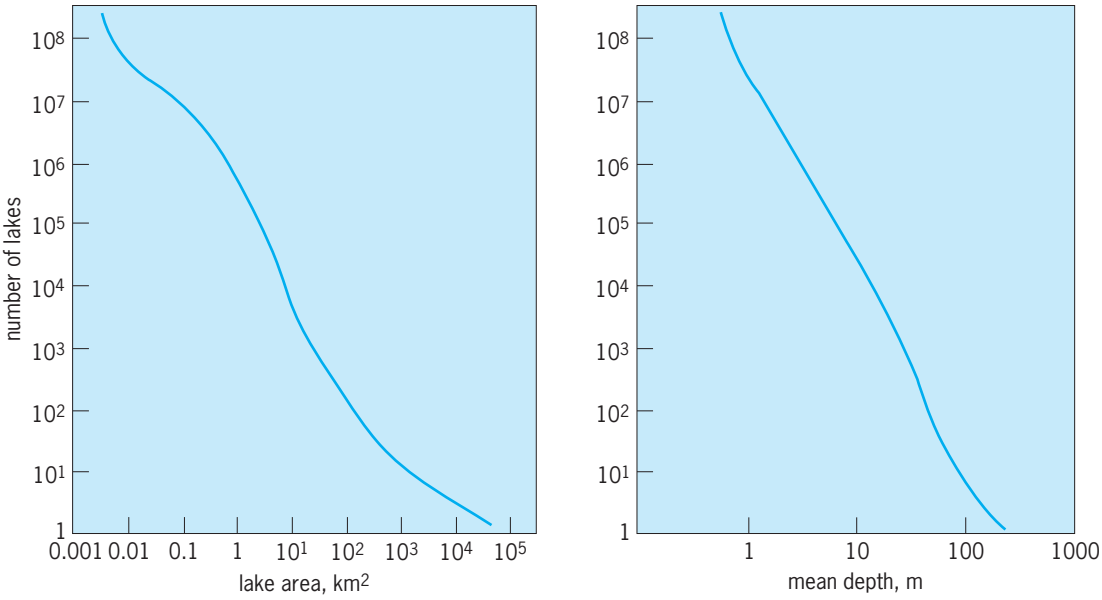

Fig. 2. Approximate number of lakes of the world in relation to lake area and approximate mean water depth. (*After R. G. Wetzel, 1990*)

Of the total light energy entering water, a portion is scattered, and the remainder is exponentially attenuated by absorption of the water molecules themselves, dissolved organic compounds (particularly humic substances), and suspended particulate matter. Submersed macrophytes are limited physiologically by reduced availability of light underwater and by slow diffusion rates of gases and nutrients in water. Internal recycling of resources, both of gases [carbon dioxide (CO_2), oxygen (O_2)] of metabolism and of nutrients, are important to the abilities of submersed plants to function and grow as well as they do under light and gas limitations.

The second most productive component of the wetland-littoral community is the microflora attached to aquatic plants and other surfaces. The surfaces provided by aquatic plants in lakes and rivers can be extensive, often exceeding 25 m^2 per square meter of bottom. The physiology and growth of attached microflora are intimately coupled to the physical and metabolic dynamics of the living substrata upon which they grow. High sustained growth of attached microflora results from their recycling of essential gases (CO_2, O_2) and dissolved nutrients. Nutrient uptake is directed primarily to the high net growth of attached microflora, and it is responsible for the high capacity of wetland-littoral areas to improve the quality of water passing through these communities.

The wetland-littoral complex, including the marginal floodplains of many rivers, produces the major sources of organic matter and energy of many fresh-water ecosystems. Most of the particulate organic matter is decomposed within these interface regions. Organic matter is exported predominantly as dissolved organic matter to the recipient lake or river (**Fig. 3**).

The deep-water pelagic zone of lakes is the least productive region along the gradient from land to water, regardless of nutrient availability. Growth of phytoplanktonic algae of the pelagic zone is limited by sparse distribution in a dilute environment where nutrient recycling is restricted by the sinking of senescent phytoplankton below the depth of photosynthesis. When nutrient recycling and availability is increased, greater phytoplankton cell densities attenuate underwater light and reduce the volume of water in which photosynthesis occurs. Despite low productivity per unit area, pelagic productivity can be collectively important in large lakes and for higher trophic levels that depend upon this organic matter.

Among biological communities, functionally similar organisms can be grouped into trophic levels based on similarities in patterns of organic matter production by photosynthesis and consumption by heterotrophic microbes and animals. Energy is transferred and nutrients are cycled within an overall ecosystem trophic structure. The productivity of each trophic level is the rate at which energy enters the trophic level from the next lower level. Because organisms expend considerable energy for maintenance and since death of an organism routes much energy and nutrients into the detrital pool of dead organic matter, only a portion of the energy of one trophic level (5–15%) is available for transfer and use by higher trophic levels. Available energy and efficiency of energy transfer decreases progressively at higher trophic levels, so that rarely can more than five or six trophic levels be supported.

A second trophic level consists of zooplankton (dominated by four major groups of animals: Protozoa/Protista, Rotifera, and the crustaceans Cladocera and Copepoda) and benthic invertebrates. In the pelagic zone these herbivorous organisms are consumed by small fishes, fry of larger fishes, and predatory zooplankton, which constitute a third trophic level (primary carnivores). A fourth trophic level may consist of medium-sized piscivorous fishes; and a fifth, large piscivorous fishes. Higher trophic levels are rare in freshwater ecosystems.

The species composition of the higher trophic levels affects the pathways of energy utilization from lower trophic levels. For example, efficiency

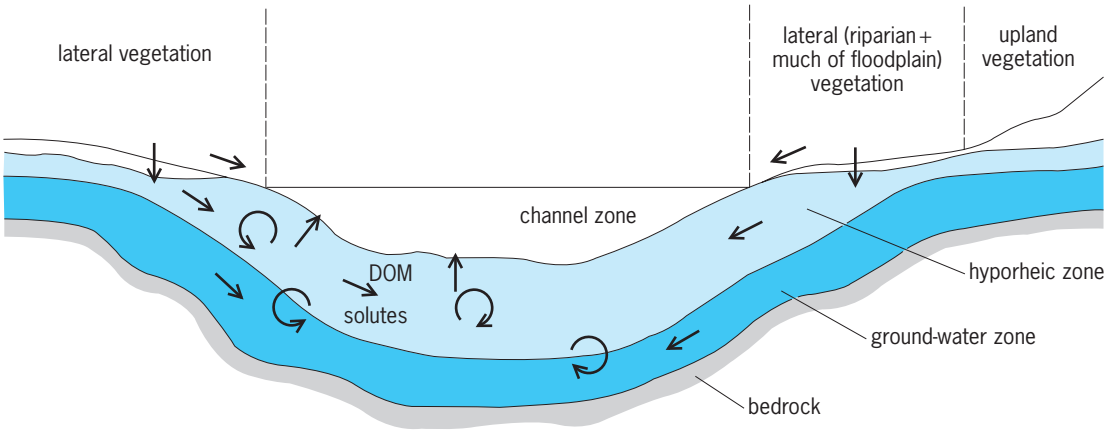

Fig. 3. Lateral and vertical boundaries of flowing-water ecosystems. The stream ecosystem boundary is defined as the hyporheic–ground-water interface and thereby includes a substantial volume beneath and lateral to the main channel. Vegetation rooted in the hyporheic zone is therefore part of stream ecosystem production. Arrows indicate flow pathways of dissolved organic matter and inorganic solutes derived from plant detritus within the stream ecosystem. (*Modified from R. G. Wetzel and A. K. Ward, 1992*)

of consumption of primary production by zooplankton is often appreciably greater in the absence of zooplankton-feeding fishes than in their presence. The community structure of phytoplankton responds variably to grazing impacts in concert with their available resources (light, nutrients, organic constituents) and may or may not be able to compensate for grazing losses in overall primary production. *See* COPEPODA; POPULATION ECOLOGY; ROTIFERA; ZOOPLANKTON.

Interactive regulatory mechanisms. Most of the organic carbon of aquatic ecosystems, and hence energy available for system operation, exists as dead organic matter (detritus), in both dissolved and particulate forms. Dissolved and particulate organic carbon moves with the water; residual particulate organic carbon is ultimately deposited at the bottom of static water. Dissolved organic carbon can also sediment if adsorbed to particulate matter or if polymerization occurs. Most heterotrophic decomposition occurs in benthic regions where organic carbon is often largely displaced from sites of production to the dominant benthic sites of decomposition at the sediment-water interface.

Although the specific composition of organic matter varies greatly, detritus inevitably carries most of the energy of the ecosystem from its points of photosynthetic origins to places of transformation. Most of that heterotrophic transformation does not occur by digestion in predatory metazoan animals, but by microbes and often under anaerobic conditions in sediments. Productivity of microbes is maximized within the resources available from dead organic matter and nutrients. Microbial heterotrophic utilization is highly dynamic and changes with sufficient rapidity to allow many generation turnovers before predators with much slower turnover rates can respond reproductively. Viral mortality of microbes as well as ingestion of bacteria by protists (heterotrophic flagellates and protozoa of the microbial loop) diverts much organic carbon from animal trophic levels. That productivity, much of which is not ingested by higher metazoan animals, is of major importance and usually dominates the energy flows within the pelagic region and completely dominates energy flows of the complete fresh-water ecosystems (wetlands, littoral, benthic, and pelagic compartments). Microbial heterotrophy is of importance to higher trophic levels in feedback processes, both positively (such as nutrient recycling and utilization by primary producers) and negatively (such as oxygen consumption and production of toxic fermentative metabolic end products). The abundance, distribution, and microbial decomposition of dissolved and particulate detritus both regulate and stabilize energy metabolism and nutrient availability in aquatic ecosystems. *See* FOOD WEB.

Relation to society. Freshwater is the pivotal resource upon which agriculture, industry, and domestic life depend. Natural or artificially created basins contain water for various periods of time. These bodies of water function in multiple ways in human activities—as sources of water, sources of biota for nutrition, and sources of recreational and esthetic value to human well-being. These water bodies are also used as purification systems, intentionally or fortuitously, for various waste products of human activities. Many of the purposes and uses to which lakes, reservoirs, and rivers are put are incompatible.

Nearly all problems of alteration of surface waters and ground-water aquifers focus on (1) introduction of abnormal amounts of alien substances or biota, or (2) modification of retention or distribution of the water by excessive use. Eutrophication of surface water results from loading of nutrients, particularly phosphorus, to lakes and reservoirs which leads to excessive production of certain algae and higher aquatic plants, and often a decrease in biodiversity. Excessive atmospheric loadings of sulfur and nitrogen oxides from fossil fuel combustion have resulted in high loadings of strong acids in liquid (rainfall) and particulate (dust) precipitation. The resulting acidification of surface and ground waters has radically altered the water chemistry and biota, with a marked reduction in biodiversity. Pollutants, such as heavy metals (mercury, copper, lead) and toxic organic compounds, continue to contaminate many surface and ground waters to the extent that not only has biodiversity been reduced but many of these water reservoirs cannot be used for human activities. In some cases, introduction of exotic species has similarly resulted in competitive exclusion of many endemic species and has also altered chemical conditions of the water. In all of these cases of environmental alterations, paleolimnological analyses of the mineralogy and structure of sediments, their organic and inorganic chemical constituents, and the fossil remains of organisms preserved in the sediments permit interpretations about past states and changing conditions of the ecosystem.

A great diversity of terrain, geomorphology, climate, and hydrology occurs in every country. This natural diversity is overlain with cultural histories in which humans have systematically and often ingeniously modified hydrological patterns, particularly of surface waters, to accommodate agricultural and water supply needs. Numerous political and human conflicts originated or were incited by the need for or mismanagement of freshwaters. The present global warming associated with increased loading of carbon dioxide and other pollutant gases to the atmosphere will certainly alter regional climatic patterns, hydrological regimes, and the availability and quality of fresh-water ecosystems. *See* ACID RAIN; EUTROPHICATION; WATER POLLUTION. Robert G. Wetzel

Bibliography. P. L. Brezonik, *Chemical Kinetics and Process Dynamics in Aquatic Systems*, Lewis Publishers, Boca Raton, FL, 1994; G. A. Cole, *Textbook of Limnology*, 3d ed., 1988; G. D. Cooke et al., *Restoration and Management of Lakes and Reservoirs*, Lewis Publishers, Boca Raton, FL, 1993; P. H. Gleick (ed.), *Water in Crisis: A Guide to the World's Fresh Water Resources*, Oxford University Press, New York, 1993; D. Gunnison (ed.), *Microbial Processes in Reservoirs*, 1985; G. E. Hutchinson, A *Treatise on Limnology: I. Geography, Physics, and Chemistry*

(1957), *II. Introduction to Lake Biology and the Limnoplankton* (1967), *III. Limnological Botany* (1975), *IV. The Zoobenthos* (1994), John Wiley, New York; O. T. Lind, *Handbook of Common Methods in Limnology*, 2d ed., 1985; H. Loffler, *Paleolimnology*, 1987; M. Straskraba and A. H. Gnouck, *Freshwater Ecosystems*: *Modelling and simulation*, Develop. Environ. Model. 8, 1985; R. Thomas et al. (eds.), *Ecological Effects of In-Situ Sediment Contaminants*, 1987; R. G. Wetzel, Death, detritus, and energy flow in aquatic ecosystems, *Freshwater Biol.*, 33:83–89, 1995; R. G. Wetzel, *Limnology: Lake and River Ecosystems*, 3d ed., Academic Press, San Diego, 2000; W. D. Williams, *Life in Inland Waters*, 1983.

Friction

Resistance to sliding, a property of the interface between two solid bodies in contact. Many everyday activities like walking or gripping objects are carried out through friction, and most people have experienced the problems that arise when there is too little friction and conditions are slippery. However, friction is a serious nuisance in devices that move continuously, like electric motors or railroad trains, since it constitutes a dissipation of energy, and a considerable proportion of all the energy generated by humans is wasted in this way. Most of this energy loss appears as heat, while a small proportion induces loss of material from the sliding surfaces, and this eventually leads to further waste, namely, to the wearing out of the whole mechanism. *See* WEAR; WORK.

In stationary systems, friction manifests itself as a force equal and opposite to the shear force applied to the interface. Thus, as in **Fig. 1**, if a small force S is applied, a friction force P will be generated, equal and opposite to S, so that the surfaces remain at rest. P can take on any magnitude up to a limiting value F, and can therefore prevent sliding whenever S is less than F. If the shear force S exceeds F, slipping occurs. During sliding, the friction force remains approximately equal to F (though often it is smaller by about 20%) and always acts in a direction opposing the relative motion. The friction force is proportional to the normal force L, and the constant of proportionality is defined as the friction coefficient f. This is expressed by the equation $F = fL$.

The "laws" of friction are essentially statements about the friction coefficient, and have been worked out over the past five centuries by a number of distinguished engineers and scientists, among them Leonardo da Vinci and Charles de Coulomb. These laws, which are approximately, but never perfectly, obeyed by typical sliding systems are as follows: (1) The friction coefficient is independent of the normal force. This is another way of saying that the friction force is proportional to the normal force. (2) The friction coefficient is independent of the sliding velocity. However, as was stated above, the friction for surfaces at rest is often about 20% greater than for the same system when sliding. Such systems are often spoken of as having larger static friction than kinetic friction. (3) The friction coefficient is independent of the apparent area of contact. Thus, flat surfaces and surfaces with contacting asperities but of the same projected area give the same friction coefficient. (4) The friction coefficient is independent of surface roughness.

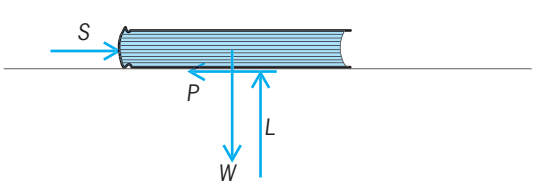

Fig. 1. The forces acting on a book resting on a flat surface when a sheer force *S* is applied. The friction from *P* is equal to *S* (up to a limiting value *F*), while *L*, the normal force, is equal to the weight *W* of the book.

Actually, the friction coefficient is primarily a property of the contacting materials and the contaminants or lubricants at the interface.

Closely related to the friction coefficient is the concept of friction angle. The friction angle θf is the largest angle relative to the horizontal at which a surface may be tilted, so that an object that is placed on the surface does not slide down. Leonhard Euler was the first to show that tan θ_f is equal to the friction coefficient.

Mechanism. A number of processes occur at the interface between two solids that tend to inhibit sliding or to use up energy during sliding. All these contribute to friction. Over the last three centuries, there has been much discussion over which one of these processes is the most important and can thus be considered the main cause of friction. Until about 1940 it was considered that surface roughness was the main cause of friction, and that work had to be done during sliding to lift one sliding surface over the high spots on the other one. Modern work has largely discounted the importance of surface roughness: first, from the standpoint of theoretical considerations, because the work done in sliding up a high spot on a surface is largely recovered on the down side, and second, because experimental testing shows that for most sliding systems friction coefficients are largely independent of the roughness of the surfaces. Very smooth surfaces, like cleaved mica, which is smooth to within one atomic diameter, give friction at least as great as that of ordinary surfaces. For reasons that are not clear, most people find it very hard to accept the fact that smooth surfaces give as much friction as rough ones.

The modern theory of friction attributes friction to adhesion between surface atoms. These surface atoms form bonds, similar in strength to those that hold the solids themselves together, and friction represents the force required to break the bonds. Atomic forces are known to be of very short range (about 10^{-10} m), and accordingly the bonds are appreciable only at those small patches (called junctions) at which the surface atoms come within 10^{-10} m of each other. All the friction originates at these

junctions, which make up the real area of contact. Over the rest of the apparent area of contact the separation of the surfaces is generally many hundreds or thousands of atomic diameters, and essentially none of the friction originates there. In a typical sliding system involving relatively large surfaces, the real area of contact is less than 1% of the apparent area.

According to the adhesion theory, lubricants at the surface reduce friction because they form bonds of reduced strength. Since hard solids generally form strong bonds while liquids and soft solids form weak bonds, it comes as no surprise to discover that good lubricants are generally either liquids or soft solids.

Besides the adhesion mechanism, which generally accounts for more than 90% of the total friction force, there are a few minor mechanisms that act so as to increase the friction. One is a plowing mechanism, when a hard surface slides against a soft one, and plows out a series of grooves. A second one, with very rough surfaces, consists of a component of the motion perpendicular to the interface. (This is the roughness effect, which at one time was thought to be the main component of friction.) Third, there is elastic hysteresis, as result of the fact that there is elastic compression and then relaxation of the material near the contacting interface, and not all the energy is recoverable. Fourth, there is an electrostatic attraction between the surfaces, but this is only significant when electrically insulating materials are used.

Almost all the frictional energy appears as heat at the interface between the sliding surfaces. This frictional heat was used by humans in prehistoric times to light fires, and this use survives today in the striking of matches, with chemical combustion initiated by the temperature rise produced by sliding. However, frictional heat usually is a nuisance, and sliding surfaces must often be cooled to prevent heat damage.

Friction coefficient values. Values of the friction coefficient are required in analyzing many problems in mechanics, in order to estimate the frictional forces at interfaces and to compute their consequences. **Figure 2** shows typical friction coefficient values, plotted as a function of the state of cleanliness or of lubrication of the surfaces. As will be seen, friction coefficients range from about 1.5 to about .07 depending on the sliding conditions. Unlubricated metals give higher friction than nonmetals, but well-lubricated metals give less. In consequence, when sliding mechanisms giving low friction are desired, metals are used when a lubricant is available, but nonmetals are preferred when the surfaces must operate in an unlubricated condition.

It should be emphasized that the friction coefficient values shown in Fig. 2 are typical ones, and that in any specific case friction values differing by as much as 30% are quite likely to be found. Indeed, it is not yet possible, even for an expert in the friction field, to be able to estimate the friction coefficient in any given case to within 10% or better.

A number of sliding systems are exceptional, in that they give friction values that are rather different from those shown in Fig. 2, and much use is

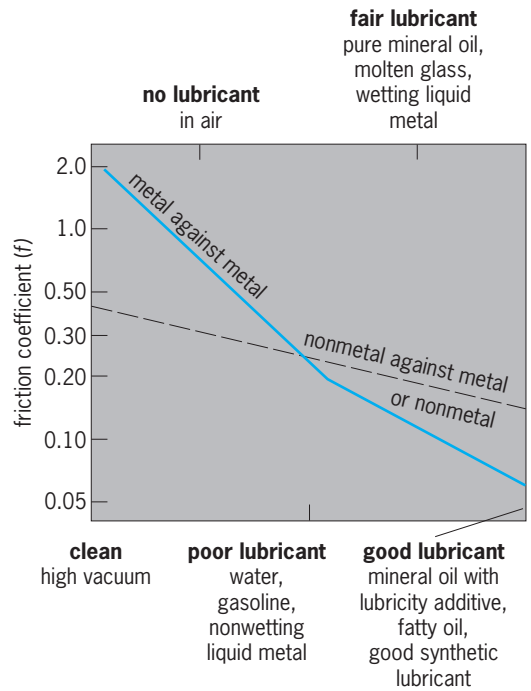

Fig. 2. Typical friction coefficient values.

made of their exceptional friction properties. Elastomeric materials like rubber give friction values that are about twice as great as those of other nonmetals, and this explains the use of rubber in shoe soles and heels, in automobile tires, and in other cases in which good traction is required. Solids that give exceptionally low friction, about half the values for other nonmetals, include graphite, Teflon, and ice. Teflon and graphite are used in low-friction coatings, while the low friction of ice is made use of in skating and skiing. Perhaps the lowest friction values are found in mammalian joints (like the human knee joint), where it has been found that friction coefficient values are generally as low as .02. It is not clear how such a low value is achieved.

Two special cases of low friction should be mentioned. One involves the use of rolling rather than sliding action. This generally produces much lower friction than is possible during sliding. For example, ball bearings generally give friction coefficients in the range .002 to .005. The other involves fluid-lubricated sliding systems at high sliding speeds, in which hydrodynamic effect allows separation of the surfaces by a full fluid film. Here friction coefficient values of .001–.003 are common.

Stopping distances. For a sliding object moving with velocity v, the distance required to bring it to a stop through friction action is governed by the equation $s = v^2/2gf$, where g is the acceleration of gravity. The distance is independent of the weight of the object. The typical stopping distances given in the **table** assume an initial speed of 60 mi/h (100 km/h). The distances are minimal, assuming in the case of rolling vehicles that the brakes are applied severely enough to cause locking of the wheels, or a braking action just short of that.

Stopping distances at 60 mi/h (100 km/h)

Typical system	Friction coefficient	Stopping distance, ft (m)
Car on dry road	.8	151 (46)
Car on wet road	.5	242 (74)
Train on dry rails	.3	403 (123)
Train on wet rails	.2	605 (184)
Car on icy road	.10	1210 (369)

Many traffic accidents at level railroad crossings occur because drivers of cars are not aware of the great length of stopping distances for trains. Road accidents frequently take place because drivers used to high friction coefficient values under dry road conditions do not realize that weather conditions have changed and that lower friction values of wet surfaces apply.

Human interactions with friction. For normal walking, humans require friction coefficient values of .20 or above. When surfaces with lower friction coefficient values are encountered, people instinctively slow down, shorten their stride, and are able to walk satisfactorily when the friction is as low as .10. Slipping generally occurs when someone walking normally on a high friction surface encounters a low-friction spot (like a banana peel), and there is no opportunity to adjust stride length.

When lifting an object by hand, a neurophysiological mechanism comes into play so that the object is gripped with a force that is only a little larger than would allow it to slip.

Frictional oscillations. It was noted above that the laws of friction are not perfectly obeyed. This generally means that if one of the variables (for example, load or apparent area) is varied by a factor of 10, the friction coefficient is changed by less than 10%. Only one of these departures from constancy has practical significance, namely, if the friction goes down when the sliding speed goes up (a so-called negative characteristic), frictional vibrations may be produced. At slow sliding speeds the vibrations take the form of relaxation oscillations (creaking, chattering, and groaning noises), while at high speeds they are harmonic oscillations (squeaking). As a general rule, unlubricated and poorly lubricated surfaces give a negative characteristic and show a tendency toward oscillation, while well-lubricated surfaces give a positive characteristic and are immune to oscillations.

Friction oscillations are generally referred to as "stick-slip," but strictly speaking this term applies only to relaxation oscillations, because during part of each oscillation cycle the sliding surfaces actually stick and have zero relative velocity. There is always some interfacial slip during harmonic oscillations.

Frictional oscillations can generally be avoided by changing one or more of the parameters. Helpful changes include reducing the normal load, increasing the sliding speed, introducing damping into the system, stiffening all components acting as springs, changing materials, or introducing good lubricants.

Frictional oscillations can be useful. They are the mechanism of music production in bowed stringed instruments, and they have been used occasionally to detect the entry of intruders into a home. More significantly, in many sliding systems they act as a warning signal that the lubricant has disappeared, that excessive wear damage is imminent, and that relubrication is in order.

Significant friction problems. In prehistoric and early historic times, humans' main interest in friction was to reduce the friction coefficient, to reduce the labor involved in dragging heavy objects. This led to the invention of lubricants, the first of which were animal fats and vegetable oils. A great breakthrough was the use of rolling action—in the form of rolling logs and then wheels—to take advantage of the lower friction coefficients of rolling systems.

In modern engineering practice available materials and lubricants reduce friction to acceptable values. In special circumstances when energy is critical, determined efforts to minimize friction are undertaken. Friction problems of practical importance are those of getting constant friction in brakes and clutches, so that jerky motion is avoided, and avoiding low friction in special circumstances, such as when driving a car on ice or on a very wet road. Also, there is considerable interest in developing new bearing materials and new lubricants that will produce low friction even at high interfacial temperatures and maintain these properties for long periods of time, thus reducing maintenance expenses. Perhaps the most persistent problem is that of avoiding frictional oscillations, a constant cause of noise pollution of the environment. *See* LUBRICANT.

Ernest Rabinowicz

Bibliography. American Society for Metals, *Friction, Lubrication, and Wear Technology*, 1992; F. P. Bowden and D. Tabor, *Friction: An Introduction to Tribology*, 1973 reprint 1982; M. Grunze and H. J. Kreuzer (eds.), *Adhesion and Friction*, 1989; I. M. Hutchings, *Tribology: Friction and Wear of Engineering Materials*, 1989.

Friedel-Crafts reaction

Originally a substitution reaction, catalyzed by aluminum chloride, in which an alkyl (R—) group or an acyl (RCO—) group replaces a hydrogen atom of an aromatic nucleus. Later, the term came to imply any of a group of electrophilic reactions in which the production of an intermediate cation or positively polarized complex species was catalyzed, either directly or indirectly, by the presence of a Lewis acid or a Brønsted-Lowry acid. The broadened definition embraces some addition, isomerization, elimination, and cracking reactions as well as the synthetically important alkylation or acylation of aromatic systems. *See* ACID AND BASE.

Alkylation. In the classical alkylation reaction, an alkyl halide (RX) serves as the alkylating agent. In reaction (1), an excess of the hydrocarbon benzene

$$R{-}Cl + C_6H_6 \xrightarrow{AlCl_3} R{-}C_6H_6 + HCl \qquad (1)$$

(C_6H_6) is the solvent. In case the aromatic hydrocarbon is not so readily available, a smaller quantity may be used in an inert solvent such as carbon disulfide or petroleum ether.

Alkylations usually need less than 1 mole of aluminum chloride ($AlCl_3$) per mole of alkyl halide, and a molar ratio of 1:10 is common. As might be predicted from the relative ease of ionization, tertiary halides (R_3CX) are superior to secondary (R_2CHX) or primary (RCH_2X) halides for the alkylation reaction. As a consequence of the possible rearrangement of the carbocationic intermediate, the alkyl group which becomes attached to the aromatic ring may differ in structure from that of the alkyl halide employed. For example, with *n*-propyl chloride ($CH_3CH_2CH_2Cl$), a large part of the product would be isopropylbenzene, $C_6H_5CH(CH_3)_2$, rather than the expected *n*-propylbenzene, $C_6H_5CH_2CH_2CH_3$. Since alkyl groups attached to an aromatic nucleus facilitate further substitution, it is difficult to prevent di- or polysubstitution during the reaction. The structure of the dialkylation and trialkylation products cannot be predicted simply from a knowledge of the rules of substitution, since aluminum chloride is a catalyst for the intermolecular and intramolecular migration of alkyl groups. Alkenes (C_nH_{2n}) may be substituted for alkyl halides in the Friedel-Crafts reaction, and the industrial synthesis of ethylbenzene and isopropylbenzene (cumene) has been carried out in this way.

Acylation. For acylation of aromatic hydrocarbons, acyl halides [reaction (2)] have proved most valuable although acid anhydrides [reaction (3)] have also

$$R{-}COCl + C_6H_6 \xrightarrow{AlCl_3} R{-}COC_6H_6 + HCl \quad (2)$$

$$(RCO)_2O + C_6H_6 \xrightarrow{AlCl_3} RCOC_6H_5 + RCOOAlCl_2 + HCl \quad (3)$$

been used. For each mole of acid chloride present, slightly more than 1 mole of $AlCl_3$ is commonly required, whereas with acid anhydrides slightly more than 2 moles of $AlCl_3$ are necessary. The acyl group deactivates the benzene ring toward further substitution. Therefore, it is usually easy to obtain good yields of monosubstitution products. In contrast to the alkyl group, the substituted acyl group does not undergo rearrangement, making acylation a more reliable synthetic tool than the alkylation reaction.

An important application of the Friedel-Crafts acylation has been in cyclizations via an intramolecular reaction. Reaction (4), the cyclization

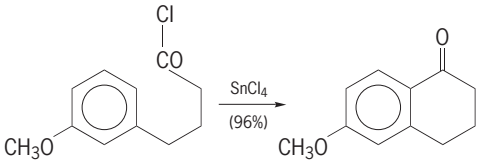

(4)

of 4–(3-methoxyphenyl)butanoyl chloride to yield 6-methoxytetralone, is an example. *See* ACYLATION.

Catalysts. In addition to aluminum halides, a wide variety of Lewis acids, chiefly acidic halides, have been used as catalysts. Some important examples are boron trifluoride, zinc chloride, stannic chloride, and antimony pentafluoride; arenetricarbonylmolybdenum has also been recommended as a Friedel-Crafts catalyst. Of the Brønsted type, H_2SO_4, HF, and trifluoroacetic acid are important, but much the same function can be served by solid acids, such as acidic cation exchange resins or aluminum silicates. *See* AROMATIC HYDROCARBON; BENZENE; CATALYSIS; SUBSTITUTION REACTION.

Charles K. Bradsher

Bibliography. R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., vol. 10, 1994; G. A. Olah, *Friedel-Crafts Chemistry*, 1973.

Fringe (optics)

The part of optics using the light or dark bands produced by interference or diffraction of light. Distances between fringes are usually very small, because of the short wavelength of light. Fringes are clearer and more numerous when produced with light of a single color.

Diffraction fringes are formed when light from a point source, or from a narrow slit, passes by an opaque object of any shape. The Fraunhofer diffraction fringes produced when the light approaches and leaves the obstacle in essentially plane waves are especially important in the theory of optical instruments. *See* DIFFRACTION; RESOLVING POWER (OPTICS).

Interference fringes are obtained by bringing together two or more beams of light that have originated from a common source. This is usually accomplished by means of an apparatus especially designed for the purpose called an interferometer, although interference fringes may also be seen in nature. Examples are the colors in a soap film and in an oil film on water. When the fringes are controllable, for example, by changing the paths traversed by the two beams in an interferometer, they are valuable for accurate measurement of small distances and of slight differences in refractive index. *See* INTERFERENCE OF WAVES; INTERFEROMETRY.

If the light from a laser is passed through an interferometer, features such as coherence and phase fluctuations may be checked by observing the sharpness and drift of the fringes. *See* LASER.

F. A. Jenkins; W. W. Watson

Front

An elongated, sloping zone in the troposphere, within which changes of temperature and wind velocity are large compared to changes outside the zone. Thus the passage of a front at a fixed location is marked by rather sudden changes in temperature and wind and also by rapid variations in other weather elements such as moisture and sky condition. *See* TROPOSPHERE.

In its idealized sense, a front can be regarded as a

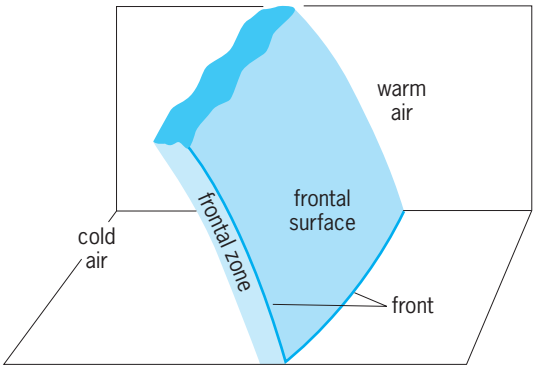

Fig. 1. Schematic diagram of the frontal zone, the angle with Earth's surface much exaggerated.

sloping surface of discontinuity separating air masses of different density or temperature. In practice, the temperature change from warm to cold air occurs mainly within a zone of finite width, called a transition or frontal zone. The three-dimensional structure of the frontal zone is illustrated in **Fig. 1**. In typical cases, the zone is about 0.6 mi (1 km) in depth and 60–120 mi (100–200 km) in width, with a slope of approximately 1/100. The cold air lies beneath the warm in the form of a shallow wedge. Temperature contrasts generally are strongest at or near the Earth's surface, the frontal zone usually being narrowest near the ground and becoming wider and more diffuse with height. Frontal zones seldom extend more than several miles above the Earth's surface. There are, however, sharp, narrow fronts of limited extent often present in the middle and upper troposphere, as discussed at the end of this article. With the exception of that section, the following concerns fronts at the Earth's surface.

The surface separating the frontal zone from the adjacent warm air mass is referred to as the frontal surface, and it is the line of intersection of this surface with a second surface, usually horizontal or vertical, that strictly speaking constitutes the front. According to this more precise definition, the front represents a discontinuity in temperature gradient rather than in temperature itself. The boundary on the cold air side is often ill-defined, especially near the Earth's surface, and for this reason is not represented in routine analysis of weather maps. In typical cases about one-third of the temperature difference between the Equator and the pole is contained within the narrow frontal zone, the remainder being distributed within the warm and cold air masses on either side. *See* AIR MASS; WEATHER MAP.

The wind gradient, or shear, like the temperature gradient, is larger within the frontal zone than on either side of it. In well-developed fronts the shift in wind direction often is concentrated along the frontal surface, while a more gradual change in wind speed may occur throughout the frontal zone. An upper-level jet stream normally is situated above a well-defined frontal zone. *See* JET STREAM.

A front moves approximately with the speed of the wind component normal to it. The strength of this component varies with season, location, and individual situation but generally lies in the range of 0–50 mi/h (0–80 km/h); 25 mi/h (40 km/h) is a typical frontal speed.

Frontal waves. Many extratropical cyclones (the lows on a weather map) begin as wavelike perturbations of a preexisting frontal surface. Such cyclones are referred to as wave cyclones. The life cycle of the wave cyclone is illustrated in **Fig. 2**. In stage I, prior to the development, the front is gently curved and more or less stationary. In stage II the front undergoes a wavelike deformation, the cold air advancing to the left of the wave crest and the warm air to the right. Simultaneously a center of low pressure and of counterclockwise wind circulation appears at the crest. The portion of the front which marks the leading edge of the cold air is called the cold front. The term warm front is applied to the forward boundary of the warm air. During stage III the wave grows in amplitude and the warm sector narrows. In the final stage the cold front overtakes and merges with the warm front, forming an occluded front. The center of low pressure and of cyclonic rotation is found at the tip of the occluded front, well removed from the warm air source. At this stage the cyclone begins to fill and weaken. *See* STORM.

Many frontal waves are weak, shallow, and fast-moving, appearing more like stage II or III (Fig. 2), rather than progressing to stage IV. Also, cases have been documented in which the occluded structure depicted in panel IV (Fig. 2) forms in a different manner from that described above. In such cases, sometimes referred to as pseudo-occlusions, the low-pressure center is observed to retreat into or form within the cold air, and frontogenesis takes place along a line joining the low center and the tip of the warm sector. Cloud observations from meteorological satellites have provided visual evidence of this process. Since the classical occlusion process,

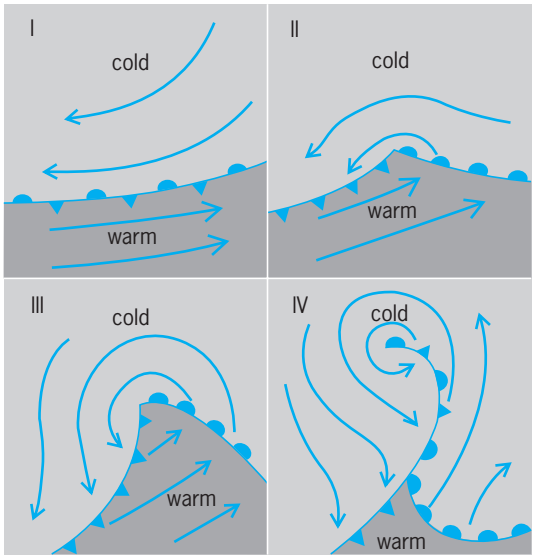

Fig. 2. Life cycle of the wave cyclone, surface projection. Arrows denote airflow. Patterns depicted are for the Northern Hemisphere while their mirror images apply in Southern Hemisphere.

in which the cold front overtakes and merges with the warm front, has never been adequately verified, it is possible that most occlusions form in this other way.

Cloud and precipitation types and patterns bear characteristic relationships to fronts, as depicted in **Fig. 3**. These relationships are determined mainly by the vertical air motions in the vicinity of the frontal surfaces. Since the motions are not unique but vary somewhat from case to case and, in a given case, with the stage of development, the features of the diagram are subject to considerable variation. In general, though, the motions consist of an upgliding of the warm air above the warm frontal surface, a more restricted and pronounced upthrusting of the warm air by the cold front, and an extensive subsidence of the cold air to the rear of the cold front. Fast-moving cold fronts are characterized by narrow cloud and precipitation systems. Smaller, more intense cells or rainbands usually are embedded within the more extensive regions of warm and cold frontal precipitation. *See* CLOUD; CLOUD PHYSICS; PRECIPITATION (METEOROLOGY).

When potentially unstable air is present in the warm sector, as often occurs in the spring and summer over North America, the main weather activity often breaks out ahead of the cold front. Complexes of cumulonimbus clouds, having horizontal dimensions on the order of 60 mi (100 km) and collective lifetimes of perhaps a day, may be present. Convective precipitation may also occur in the form of elongated lines of cumulonimbus clouds called squall lines, as represented by the tall cumulonimbus in Fig. 3*b*. *See* SQUALL; SQUALL LINE.

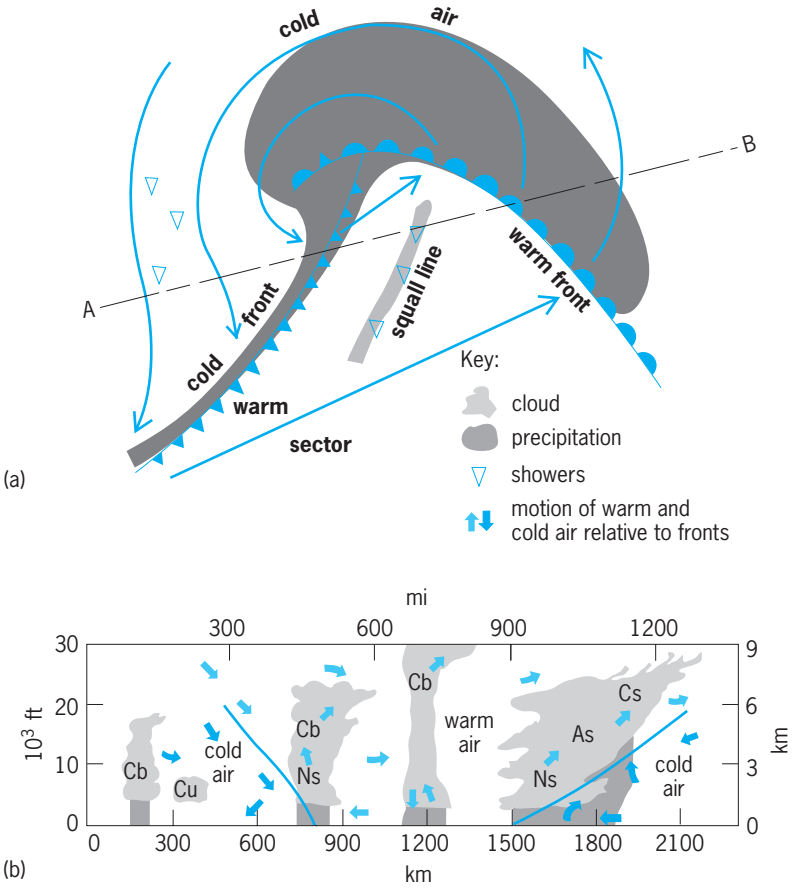

Fig. 3. Relation of cloud types and precipitation to fronts. (*a*) Surface weather map. (*b*) Vertical cross section (along A-B in diagram *a*). Cloud types: Cs, cirrostratus; As, altostratus; Ns, nimbostratus; Cu, cumulus; and Cb, cumulonimbus

In the above conceptual model, the movement of the frontal zones associated with the cyclone's evolution is primarily responsible for the cloudiness and precipitation as the warm air flows up and over the warm front and as the cold front wedges under the warm air sector. Thus, the largely mechanical wedge action of the frontal zones produces the vertical motion associated with the clouds. This conceptual model (known as the polar-front cyclone model) was based on the idea that most cyclones formed on preexisting surface fronts.

Since about 1940, it has been realized that extratropical cyclones often develop in regions where no significant surface frontal zones exist. The development of the cyclones is often more related to the three-dimensional structure of the atmosphere over a large volume than to the location of any frontal zones. More modern conceptual models of cyclone development (which can be referred to as baroclinic-wave models) emphasize the flow of airstreams relative to the evolving cyclone, rather than the distortion of air masses by the moving fronts. Yet the relative distribution of cloudiness and precipitation, as well as the evolution of frontal zones with respect to the cyclone, is still quite similar in these later models to that in the older model discussed above. Thus, Figs. 2 and 3 are consistent with the newer concepts. *See* DYNAMIC INSTABILITY.

Polar front. A front separating air of tropical origin from air of more northerly or polar origin is referred to as a polar front. Frequently only a fraction of the temperature contrast between tropical and polar regions is concentrated within the polar frontal zone, and a second or secondary front appears at higher latitudes. In certain locations such a front is termed an arctic front.

The polar front can also be defined in a climatological sense as a region in which the horizontal temperature gradient tends to be large compared to surrounding areas; within this region frequent formation of sharp frontal zones occurs. This definition applies for the remainder of this section.

In winter the major or polar frontal zones of the Northern Hemisphere extend from the northern Philippines across the Pacific Ocean to the coast of Washington, from the southeastern United States across the Atlantic Ocean to southern England, and from the northern Mediterranean eastward into Asia. An arctic frontal zone is located along the mountain barriers of western Canada and Alaska. In summer the average positions of the polar frontal zones are farther north, the Pacific zone extending from Japan to Washington and the Atlantic zone from New Jersey to the British Isles. In addition to a northward-displaced polar front over Asia, an arctic front lies along the northern shore and continues eastward into Alaska.

The polar frontal zone of the Southern Hemisphere lies near 45°S in summer and slightly poleward of that latitude in winter. Two frontal bands, more pronounced in winter than in summer, spiral into the main zone from subtropical latitudes. These bands, originating east of the Andes and northeast of New Zealand, merge with the main frontal zone after making a quarter circuit of the globe.

Frontogenesis. When extratropical cyclones form in the absence of preexisting fronts, strong frontogenesis (initial formation of a front) can occur in conjunction with the cyclone's development. The formation of a frontal zone requires an increase in the temperature gradient and the development of a wind shift.

The transport of temperature by the horizontal wind field can initiate the frontogenesis process as is shown for two cases in **Fig. 4**. The two wind fields shown would, in the absence of other effects, transport the isotherms in such a way that they would become concentrated along the A-B lines in both cases. (Both of these types of wind fields, or combinations of them, are found in developing extratropical cyclones, and sometimes elsewhere.) Since the temperature gradient is inversely proportional to the spacing of the isotherms, it is clear that frontogenesis would occur along the A-B lines. Vertical air motions will modify this frontogenesis process, but these modifications will be small near the ground where the vertical motion is small. Thus, near the ground the temperature gradients will continue to increase in the frontal zones as long as the horizontal wind field does not change. As the temperature gradient increases, a circulation will develop in the vertical plane through C-D in each case. The thermal wind relation is valid for the component of the horizontal wind which blows parallel to the frontal zone. The thermal wind, which is the change in the geostrophic wind over a specific vertical distance, is directed along the isotherms, and its magnitude is proportional to the temperature gradient. As the frontogenesis process increases the temperature gradient, the thermal wind must also increase. But if the thermal wind increases, the change in the actual wind over a height interval must also increase (since the thermal wind closely approximates the actual wind change with height). This corresponding change in the actual wind component along the front is accomplished through the action of the Coriolis force. A small wind component perpendicular to the front is required if the Coriolis force is to act in this manner. This leads to the circulation in the vertical plane that is shown in **Fig. 5**. *See* CORIOLIS ACCELERATION; CYCLONE; GEOSTROPHIC WIND; THERMAL WIND.

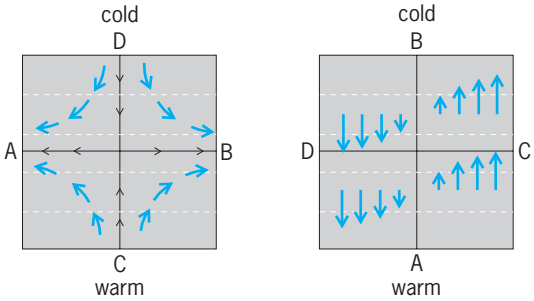

Fig. 4. Two horizontal wind fields which can cause frontogenesis. Broken lines represent isotherms, and colored arrows show wind directions and speeds (length of arrow shaft is proportional to the speed).

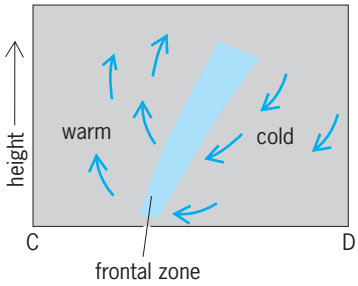

Fig. 5. Circulation in the vertical plane through C-D for both cases of Fig. 4.

This circulation plays an important role in the frontogenesis process and in determining the frontal structure. The rising motion in the warm air and the sinking motion in the cold air are consistent with observed cloud and precipitation patterns. The circulation helps give the front its characteristic vertical tilt which leaves the relatively cooler air beneath the front. Near the ground the circulation causes a horizontal convergence of mass. This speeds up the frontogenesis process by increasing the rate at which the isotherms move together. The convergence field also carries the momentum lines together in the frontal zone in such a way that a wind shear develops across the front. This wind shear gives rise to the wind shift which is observed with a frontal passage. Eventually the front reaches a quasi-steady state in which the turbulent mixing balances the frontogenesis processes. Other frontogenesis effects are important in some cases, but the mechanism presented above appears to be the predominant one.

Upper-level frontal zones. In addition to and independent of the surface frontal zones described above, elongated sloping zones of strong temperature contrast and wind shear can form in the upper troposphere in conjunction with jet stream intensification. These upper-level frontal zones are characterized by low humdity, air sinking within them. Their development is often accompanied by a downward folding of the tropopause and the incorporation of stratospheric air into the upper portion of the zones. The strong wind shear within these frontal zones often produces severe turbulence, a potentially dangerous situation for jet aircraft, as the zones cannot be detected visually due to a lack of cloudiness in the dry, descending air. Upper-level frontal zones thus are significant (but not the only) sources of clear-air turbulence (CAT). *See* CLEAR-AIR TURBULENCE; UPPER-ATMOSPHERE DYNAMICS. Stephen E. Mudrick

Bibliography. L. K. Lutgens and E. J. Tarbuck, *The Atmosphere: An Introduction to Meteorology*, 7th ed., 1997; S. E. Mudrick, A numerical study of frontogenesis, *J. Atm. Sci.*, 31:869–892, 1974;

P. Ray (ed.), *Mesoscale Meteorology and Forecasting*, American Meteorological Society, 1987; J. M. Wallace and P. V. Hobbs, *Atmospheric Science*: *An Introductory Survey*, 1977.

Frost

The formation of ice on the ground, on plants, or on structures near the Earth's surface as a result of the drop in air temperature to lower than 32°F (0°C) and the presence of water vapor or liquid water in the air. The causes of the cooling, and the types of ice that form, can be quite varied, depending on the existing meteorological factors. Frost can have serious impact, with the main effect being in agriculture, but transportation and other areas are also affected. *See* AIR TEMPERATURE.

Causes. There are two principal causes of the cooling necessary for frost formation: heat loss by radiation, and the advection of cold air. Other factors that have a significant influence on frost formation are wind, atmospheric humidity, surface wetness, and topography.

Radiational heat loss. The Earth's surface loses heat by long-wave, infrared radiation. At night, when there is no incoming radiation from the Sun, and especially with clear skies, when there is no returned infrared radiation from clouds, the heat loss can lead to rapid cooling of the surface. In turn, the cold surfaces cool the lowest layers of air as well. In the absence of wind to stir the air, a steep gradient of temperature can become established, so that the temperature within inches of the ground may be many degrees colder than the air several feet above the ground. The depth of this inversion layer increases and the temperature near the surface steadily decreases during the night. *See* TEMPERATURE INVERSION.

Wind tends to cause turbulence and this mixes the coldest air from near the surface with warmer air higher up. Cloudy skies and, to a lesser degree, humidity in the air reduce the rate of heat loss and the risk of frost. Also, conduction of heat from the subsurface soil helps to reduce the loss of heat to some degree. The total amount of cooling that can develop may be calculated by using semiempirical schemes, but the errors in such calculations are fairly large. Furthermore, in hilly terrain the coldest air tends to sink to the lowest elevation, and even small differences in terrain height can lead to substantial differences in temperature. As a result of radiational heat loss, up to 36°F (20°C) cooling during the course of a night is possible.

Heat loss by advection. The movement of cold air, such as that which accompanies the advance of a cold front, is the second major class of situations that can lead to frost formation. In these situations, the air is colder then the ground and, usually, the temperature decreases with altitude. The severity of the drops in temperature can approach those due to radiation cooling, but the most dramatic cooling events are those where both processes contribute. However, fronts are often associated with clouds and usually with winds, so that conditions tend to be contrary to those producing the strongest radiational cooling. *See* FRONT.

Ice types. While low temperatures themselves might be considered as frost, the manifestations and the most serious impacts of frost are in the formation of ice that results from the low temperatures when water is present. The water may be held in the soil or in the plants, or it may come from vapor or droplets in the air.

In order for the water within the surface soil or in the tissues of plants to freeze, a small amount of supercooling is required, typically a few degrees; thus ice formation will not begin (the process is called nucleation) until those temperatures are reached. Nucleation depends on the presence of specific substances. Continued freezing may be limited by the release of latent heat associated with the freezing; in fact some plants have developed protection against extensive freezing damage by having large internal reservoirs of water. Wind velocity is a crucial factor in this regard, as air motion greatly increases the removal of heat from the plant.

Water vapor from the air deposits on surfaces whose temperature is lower than the ice point corresponding to the ambient humidity. These deposits grow as feathery crystals, known as hoarfrost; the crystals are quite beautiful and can attain a size greater than 0.5 in. (1.3 cm). The formation of hoarfrost is analogous to the formation of dew at temperatures higher than 32°F (0°C). *See* DEW; DEW POINT.

If condensation of droplets takes place in the air (that is, a fog forms) coincident with the cooling of the ground and plants, some fog droplets impinge and freeze onto the objects, forming rime. Rime is much denser than hoarfrost. The total mass (weight) deposited by rime can also be much higher than the deposit of hoarfrost. Occasionally, the two forms of ice form in combination. *See* FOG.

If moisture is deposited as liquid water prior to the lowering of temperatures below the freezing point, the water deposit (dew or rain) may subsequently freeze, forming a transparent (black) ice coating known as glaze. Very heavy and damaging glaze can form if rain falls from warmer air aloft into a surface layer of cold air. *See* PRECIPITATION (METEOROLOGY).

Climatology. Frost climatology is widely available for agricultural areas. The dates of first and last frosts for 32, 28, 24, 20°F (0, −2.2, −4.4, −6.7°C), or other sets of temperatures, are usually tabulated, as well as the probabilities of frosts for given calendar days. Tables of the numbers of frost days per year are also available. These data represent the average conditions over many years. Deviations from the average can always be expected. Also, local variations are very important. Other factors, such as humidity or water content in the soil, add further variability to the actual frequency and impact of frosts.

Forecasting and prevention. General advances in the 1- to 2-day synoptic-scale meteorological forecasts are providing improved information for the

generation of frost forecasts. As a result, frost warnings of several hours are quite reliable, but forecasts more than a day ahead are still beset with inaccuracies. The incorporation of small-scale local effects places further demands on the forecaster. Good forecasts of frost are of great value to growers who operate expensive frost-protection systems. Prevention measures are decided on the basis of the forecasts and the susceptibility of the plants at that moment. Common techniques of frost prevention are heating of the air with burners, smoke production to reduce radiation heat loss, spraying of the plants with water to utilize the latent heat of freezing as a heat source, and treatments to reduce the temperature threshold for nucleation. The last-mentioned method is experimental, but it has the potential of giving protection down to 24°F (−4.4°C). *See* AGRICULTURAL METEOROLOGY.

Gabor Vali

Froude number

The dimensionless quantity $U(gL)^{-1/2}$, where U is a characteristic velocity of flow, g is the acceleration of gravity, and L is a characteristic length. The Froude number can be interpreted as the ratio of the inertial to gravity forces in the flow. This ratio may also be interpreted physically as the ratio between the mean flow velocity and the speed of an elementary gravity (surface or disturbance) wave traveling over the water surface. *See* WAVE MOTION IN LIQUIDS.

In an open channel, the characteristic length is taken to be the hydraulic depth, which by definition is the cross-sectional flow area divided by the width of the free surface. For rectangular sections, the characteristic length equals the flow depth. *See* OPEN CHANNEL.

When the Froude number is equal to one, the speed of the surface wave and that of the flow is the same. The flow is in the critical state. When the Froude number is less than one, the flow velocity is smaller than the speed of a disturbance wave traveling on the surface. Flow is considered to be subcritical (tranquil flow). Gravitational forces are dominant. The surface wave will propagate upstream and, therefore, flow profiles are calculated in the upstream direction. When the Froude number is greater than one, the flow is supercritical (rapid flow) and inertial forces are dominant. The surface wave will not propagate upstream, and flow profiles are calculated in the downstream direction.

The dimensionless wave drag and power required for a surface ship are strong functions of its Froude number. Therefore, tow-tank models are tested at the same Froude number as the prototype ship. The Reynolds number of the model does not match the prototype, and viscous drag is generally estimated in tow-tank tests. *See* REYNOLDS NUMBER; SHIP POWERING, MANEUVERING, AND SEAKEEPING; TOWING TANK.

The Froude number is useful in calculations of hydraulic jump, design of hydraulic structures, and ship design, where forces due to gravity and inertial forces are governing. In these cases, geometric similitude and the same value of the Froude number in model and prototype produce a good approximation to dynamic similitude. *See* DIMENSIONAL ANALYSIS; DIMENSIONLESS GROUPS; HYDRAULICS; SHIP DESIGN.

Raymond M. Wright

Bibliography. N. H. C. Hwang and R. J. Houghtalen, *Fundamentals of Hydraulic Engineering Systems*, 3d ed., Prentice Hall, 1996; V. L. Streeter, E. B. Wylie, and K. W. Bedford, *Fluid Mechanics*, 9th ed., McGraw-Hill, 1998; T. Szirtes and P. Rozsa, *Applied Dimensional Analysis and Modeling*, McGraw-Hill, 1997; R. B. Zubaly, *Applied Naval Architecture*, Cornell Maritime Press, 1996.

Fruit

A matured carpel or group of carpels (the basic units of the gynoecium or female part of the flower) with or without seeds, and with or without other floral or shoot parts (accessory structures) united to the carpel or carpels. Carpology is the study of the morphology and anatomy of fruits. The ovary develops into a fruit after fertilization and usually contains one or more seeds, which have developed from the fertilized ovules. Parthenocarpic fruits usually lack seeds. Fruitlets are the small fruits or subunits of aggregate or multiple fruits. Flowers, carpels, ovaries, and fruits

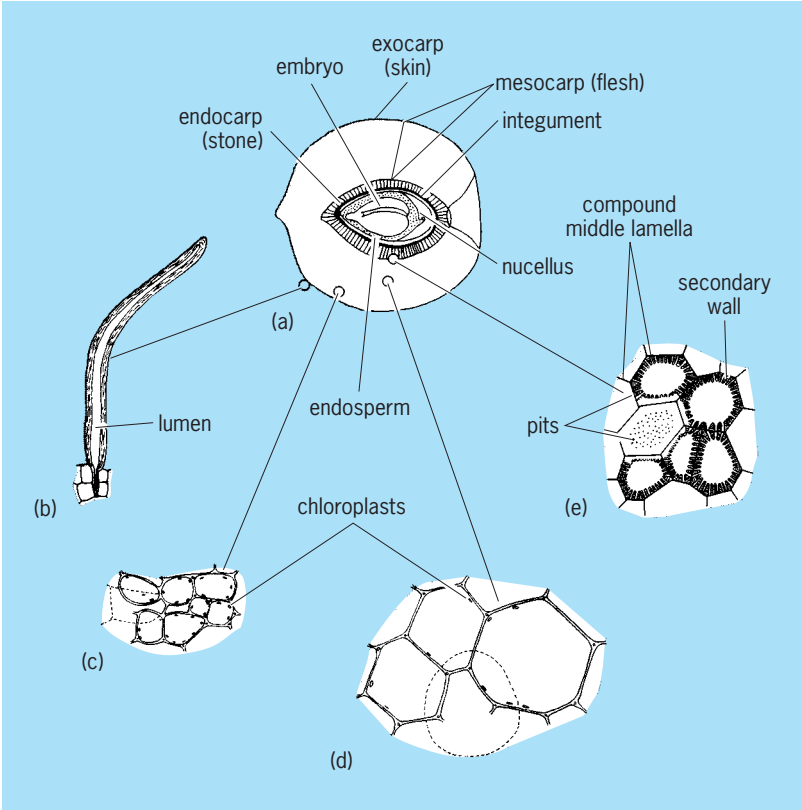

Fig. 1. Tissue components of fruit of *Prunus persica* (peach). (*a*) Longitudinal section of fruit, a drupe. Circles indicate positions of sections shown in *b–e*, which are representative of a young fruit about 1 in. (2.5 cm) in diameter. (*b*) Epidermal hair of exocarp. (*c*, *d*) Parenchyma of fleshy mesocarp. (*e*) Sclereids of stony endocarp. (*After K. Esau, Plant Anatomy, 2d ed., John Wiley and Sons, 1967*)

are, by definition, restricted to the flowering plants (angiosperms), although fruitlike structures may enclose seeds in certain other groups of seed plants. The fruit is of ecological significance because of seed dispersal. *See* FLOWER; SEED.

Morphology. A fruit develops from one or more carpels. Usually only part of the gynoecium, the ovary, develops into a fruit; the style and stigma wither. Accessory (extracarpellary or noncarpellary) structures may be closely associated with the carpel or carpels and display various degrees of adnation (fusion) to them, thus becoming part of the fruit. Such accessory parts include sepals (as in the mulberry), the bases of sepals, petals, and stamens united into a floral tube (apple, banana, pear, and other species with inferior ovaries), the receptacle (strawberry), the pedicel and receptacle (cashew), the peduncle (fleshy part of the fig), the involucre composed of bracts and bracteoles (walnut and pineapple), and the inflorescence axis (pineapple).

True versus accessory fruits. A fruit derived from only carpellary structures is called a true fruit or, because it develops from a superior ovary (one inserted above the other floral parts), a superior fruit (**Fig. 1**; corn, date, grape, plum, and tomato). Fruits with accessory structures are called accessory (or inaptly, false or spurious) fruits (pseudocarps) or, because of their frequent derivation from inferior ovaries (inserted below the other floral parts), inferior fruits (**Fig. 2**; banana, pear, squash, and walnut). True fruits invariably develop from the superior ovaries of hypogynous or perigynous flowers; accessory fruits usually develop from the inferior or part-inferior ovaries of epigynous or semiepigynous flowers. Exceptions include the accessory fruits of the rose and strawberry, which contain achenes (true fruitlets) that develop from the carpels of perigynous flowers. The fruit of the rose (the so-called hip) represents a floral tube enclosing the achenes. A strawberry is a large, fleshy receptacle bearing small achenes. True and accessory fruits are distinguishable on the basis of developmental anatomy.

Simple and aggregate versus multiple fruits. Fruits can be characterized by the number of ovaries and flowers forming the fruit. A simple fruit is derived from one ovary (Figs. 1–5), an aggregate fruit from several ovaries of one flower (**Fig. 3***a* and *b*; magnolia, rose, and strawberry). A multiple (collective) fruit is derived from the ovaries and accessory structures of several flowers consolidated into one mass (Fig. 3*c*;

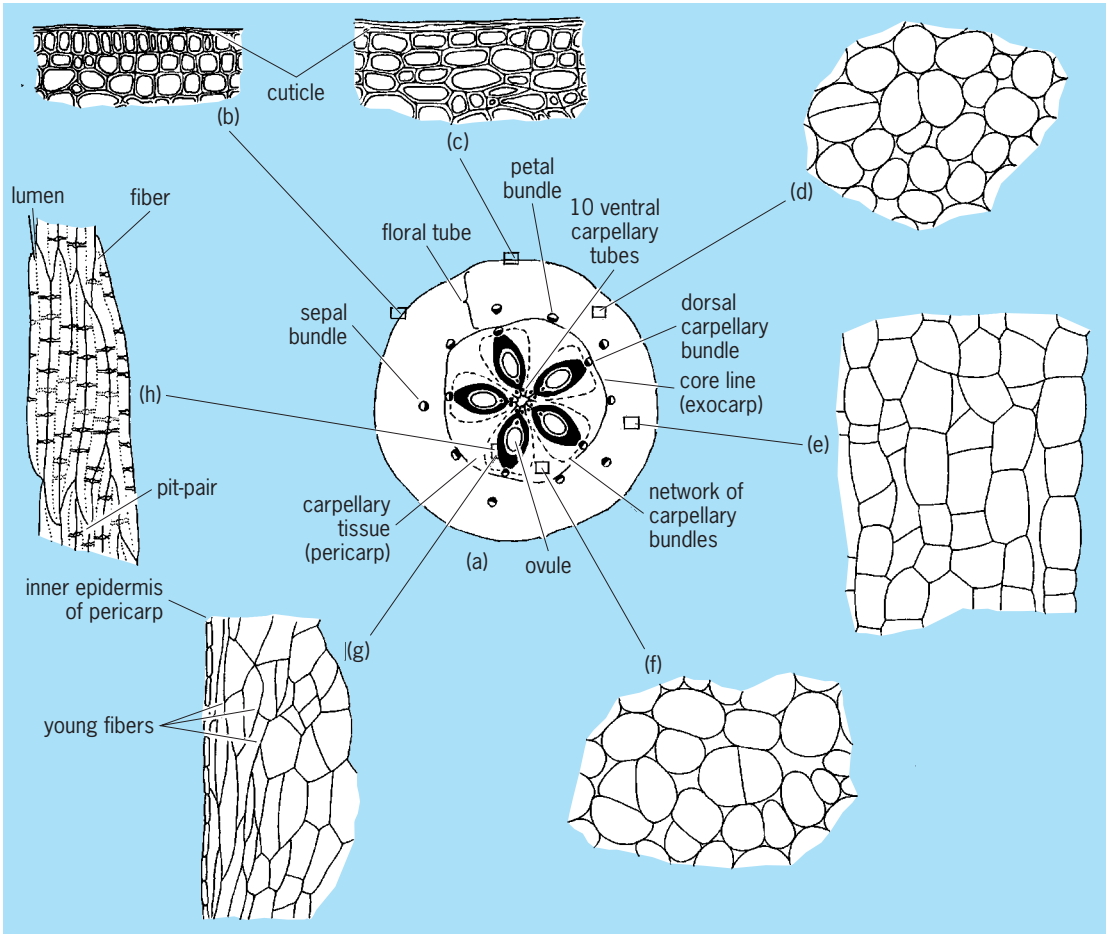

Fig. 2. Tissue components of fruit of *Malus sylvestris* (apple). (*a*) Transverse section of fruit, a pome. Rectangles indicate positions of sections shown in *b–h*; *b* and *d–g* are from a young fruit 0.4 in. (1 cm) in diameter; *c* and *h* are from a mature fruit. (*b*) Epidermis and subjacent collenchyma from young fruit and (*c*) from mature fruit. (*d*, *e*) Parenchyma of floral-tube part of flesh. (*f*) Parenchyma of mesocarp part of flesh. (*g*) Fibers of cartilaginous endocarp from young fruit and (*h*) from mature fruit. (*After K. Esau, Plant Anatomy, 2d ed., John Wiley and Sons, 1967*)

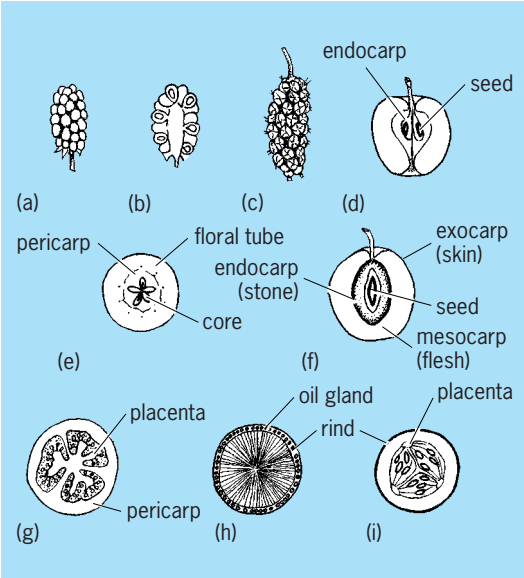

Fig. 3. Fleshy fruits. (*a*) Aggregate fruit (blackberry) in surficial view and (*b*) in sectional view. (*c*) Multiple fruit (mulberry). (*d–i*) Simple fruits: (*d*, *e*) pome (apple), (*f*) drupe (cherry), (*g*) berry (tomato), (*h*) hesperidium (orange), (*i*) pepo (cucumber); *d* and *f* are longitudinal sections, *e* and *g–i* are transverse sections.

fig, pandan, pineapple, and sweet gum). Simple fruits are either true (Fig. 1) or accessory (Fig. 2) fruits. However, aggregate and multiple fruits are always accessory fruits. Their individual units or fruitlets are characterized by the terminology of simple fruits given below. In an aggregate fruit a fruitlet is the individual carpel of a flower with separate carpels, whereas in a multiple fruit a fruitlet is the ovary and accessory structures of a flower. Simple, aggregate, an multiple fruits may be fleshy or dry, and dehiscent or indehiscent.

Dry versus fleshy, and dehiscent versus indehiscent fruits. The fruit wall at maturity may be fleshy or, more commonly, dry. Fleshy fruits range from soft and juicy to hard and rough. Dry fruits may be dehiscent, opening to release seeds, or indehiscent, remaining closed and containing usually one seed per fruit. Dehiscent fruits have areas of opening called sutures and separating segments called valves (**Fig. 4**). Fleshy fruits are rarely dehiscent.

Fruit wall versus pericarp. The pericarp is the fruit wall developed from the ovary. In true fruits, the fruit wall and pericarp are synonymous (Fig. 1), but in accessory fruits the fruit wall includes the pericarp plus one or more accessory tissues of various derivation (Fig. 2). Some persons inappropriately broaden the term pericarp to include noncarpellary tissue. The pericarp is usually divisible anatomically into the exocarp, mesocarp, and endocarp. Besides the fruit wall, a fruit contains one or more seed-bearing regions (placentae) and often partitions (septa).

Classification. Fruits have been classified on the basis of numerous criteria, including dehiscence versus indehiscence, dry versus fleshy texture, morphology (form), development (ontogeny), relationship of the ovary to other reproductive parts, and number of carpels and seeds. The terminology and classification presented here is the traditional one developed by taxonomists and temperate botanists and is biased toward fruits of economic importance. Although these fruit types are widely accepted, the classification is artificial since numerous transitional types occur and since morphological, histological, developmental, ecological, and functional criteria are used rather indiscriminately. Nevertheless, the numerous attempts at developing a natural system of fruit classification have not been very successful.

Dehiscent dry fruits. Dehiscent fruits, sometimes uncritically called pods, usually have a dry texture at maturity. The following types (Fig. 4) are representative: follicle, containing one carpel, opening along one suture (milkweed and fruitlet of magnolia); legume, with one carpel, but opening along two sutures (common garden bean and many other members of the legume family); loment, a special type of legume constricted between seeds and breaking crosswise into one-seeded segments (beggar's-ticks); silique, composed of two elongated carpels with two sutures, the valves opening from below upward to expose the seed-bearing partition (replum) still attached to the stalk (many members of the mustard family); silicle, like a silique, but short and broad (shepherd's-purse); schizocarp, splitting into one-seeded usually indehiscent units (mericarps or sometimes "cocci"), each a derivative of a carpel—sometimes regarded as intermediate between dehiscent and indehiscent dry fruits (mallow and members of the unbel family); capsule, containing two or more carpels, and usually with a like number of vertical valves or sutures (jojoba and lily). Capsules have diverse modes of dehiscence. Thus, in the pyxis subtype dehiscence is circumferential, with the top separating as a lid (purslane).

Indehiscent dry fruits. Representative types (**Fig. 5**) include caryopsis or grain, with the fruit wall fused to the single seed (corn and other members of the grass family); achene (sometimes akene), with the fruit wall readily separable from the single seed (clematis and the fruitlets of a strawberry; cypselae are

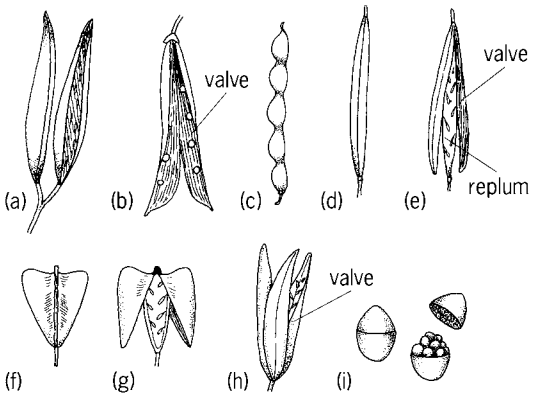

Fig. 4. Dehiscent dry fruits: (*a*) Follicles (milkweed). (*b*) Legume (bean). (*c*) Loment (beggar's-ticks). (*d*, *e*) Silique, closed and open (mustard). (*f*, *g*) Silicle, closed and open (shepherd's-purse). (*h*) Capsule (lily). (*i*) Pyxis, closed and open (purslane).

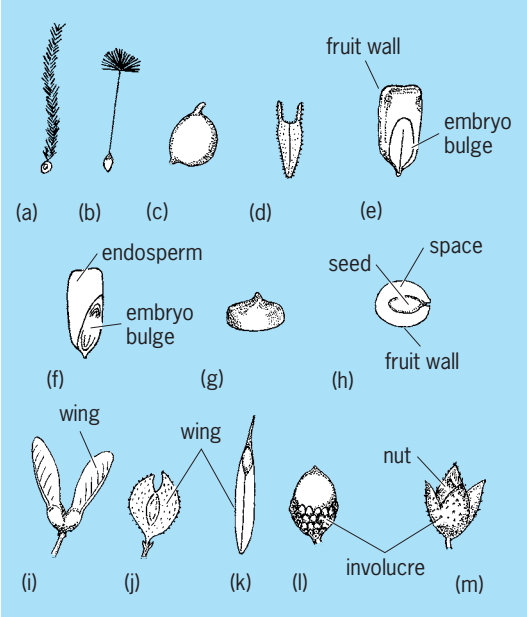

Fig. 5. Indehiscent dry fruits. (*a–d*) Achenes; (*a*) clematis, (*b*) dandelion, (*c*) buttercup, and (*d*) sticktight. (*e*) Caryopsis or grain (corn) in surficial view and (*f*) in sectional view. (*g*) Utricle (goosefoot) in surficial view and (*h*) in sectional view. (*j–k*) Samaras: (*i*) maple, (*j*) elm, and (*k*) ash. (*l*) Nut and involucre (acorn) of oak. (*m*) Nut of beech.

achenes derived from inferior ovaries, as in members of the sunflower family); utricle, like an achene, but with the fruit wall thin and bladdery (beet and goosefoot); samara, like an achene, but with a wing for wind dispersal (ash, elm, and maple); nut, like an achene but larger, with the fruit wall hard throughout, derived from several carpels, and usually one-seeded and often at least partly enclosed in an involucre (beech, hazelnut, walnut, and the acorn of oak).

Fleshy fruits. Fleshy fruits are usually indehiscent, but there are exceptions, such as squirting cucumber. The chief types (Fig. 3) of simple fleshy fruits include pome, with a thin, cartilaginous inner part (endocarp) and an fleshy outer part (Fig. 2; pear and pomegranate, the latter a leathery pome or balausta); drupe (stone fruit), with an outer exocarp (skin) in the fruit wall, a central fleshy mesocarp (flesh), and an inner stony endocarp or stone (Fig. 1; cherry); berry, with the entire fruit fleshy (avocado, banana, grape, kiwi, and tomato); hesperidium, a modified berry with a thick, leathery separable rind (citrus fruits); pepo, a modified berry with a hard or tough inseparable rind (cucumber and other members of the gourd family).

Anatomy. Anatomically or histologically, a fruit consists of dermal, ground (fundamental), and vascular systems and, if present, one or more seeds. After fertilization the ovary and sometimes accessory parts develop into the fruit; parthenocarpy is fruit production without fertilization. The fruit generally increases in size and undergoes various anatomical changes that usually relate to its manner of dehiscence, its mode of dispersal, or protection of its seeds. The economically important, mainly fleshy fruits have received the most histological and developmental study. *See* COLLENCHYMA; EPIDERMIS (PLANT); PARENCHYMA; SCLERENCHYMA.

Fruit growth. Size increase of fruits is hormonally controlled and results from cell division and especially from cell enlargement. Cell number, volume, and weight thus control fruit weight. Cell division generally is more pronounced before anthesis (full bloom); cell enlargement is more pronounced after. Fruit growth may result only from cell enlargement, all cells of the mature fruit being present at anthesis (corinth grape and raspberry). More often the fruit develops by cell enlargement accompanied by cell division, which may continue briefly (tomato, and peach and other stone fruits) or until maturity (avocado and citrus). Cells may enlarge 10–15 times (kiwi) or more. Expansion of intercellular spaces may also contribute substantially to fruit growth, for instance, 25% of the total volume in apples. The growth curve may be a single sigmoid (apple, pea, and tomato) or double sigmoid (grape, kiwi, and stone fruits). Usually not all parts of the fruit grow at the same rate. Cell layers present at anthesis or during early fruit development may later become obscured.

Fruit wall and pericarp. As defined above, the pericarp or fruit wall of true fruits represents the ovary wall, whereas the fruit wall of accessory fruits includes the pericarp plus accessory noncarpellary tissue. The pericarp is subdivided into the exocarp, mesocarp, and endocarp, but these designations are purely descriptive and have no real developmental basis. The pericarp layers may merge imperceptibly into one another, and only endocarp and exocarp may be present. In accessory fruits the pericarp may be anatomically indistinguishable from noncarpellary tissue (lettuce). However, in some apple cultivars (pink pearl) these regions are distinguishable: the pericarp is white, and the accessory tissues are pink. Anatomically, fruit walls of indehiscent dry fruits often resemble seed coats (in the achene of sunflower and caryopses of grasses). Structures other than the fruit wall may constitute a principal part of a fruit, for example, expanded placentae (Fig. 3*g*; tomato) or fleshy seed coats (pomegranate). True fruits may be slightly (kiwi) or appreciably (green ash) surrounded by a persistent calyx.

Epidermis and parenchyma are always present, and vascular tissue nearly always present in the fruit wall (fruits of *Balanophora* completely lack vascular tissue), whereas collenchyma and especially sclerenchyma—fibers or sclereids or both—are often present, sometimes forming rather late in fruit development (**Figs.** 1, 2, **6**, and **7**). The flesh in mature fleshy fruits is due to succulent parenchyma cells, which contain water and solutes mainly in their vacuoles. Immature dry fruits also exhibit some fleshiness. Parenchyma cells may be tightly packed (Figs. 1 and 6) or loosely arranged with small or often large intercellular spaces (Fig. 2). Mealiness or mushiness of overripe fruits results from pronounced cell separation (apple). Complete cell separation occurs in the flesh of some fruits (cantaloupe). Collenchyma

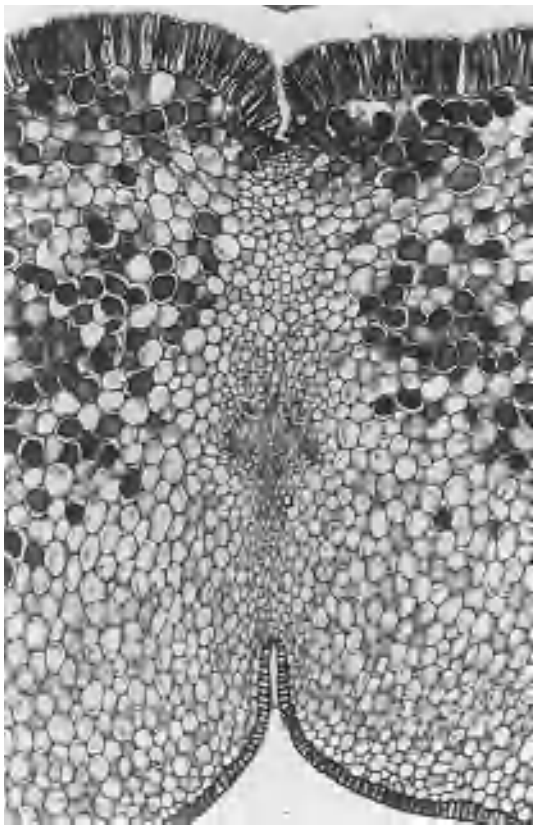

Fig. 6. Transverse section of wall of immature fruit of *Simmondsia chinensis* (jojoba) at point of potential dehiscence (central smaller cells). The outer epidermis (at the top) consists of tannin-filled macrosclereids. (*From R. Schmid, Floral and fruit anatomy of jojoba, Simmondsia chinensis, Mem. II. Conf. Int. Jojoba Aprovech., Ensenada, Mexico, pp. 143–148, 1978*)

and sclerenchyma occur in both dry and fleshy fruits, though usually in greater abundance in dry fruits. Pear flesh represents noncarpellary tissue consisting of parenchyma with vascular tissue and islands of sclereids. Dry fruits are generally more complex than fleshy ones, in part because many dry fruits dehisce whereas fleshy fruits rarely do. Tannins (Fig. 6), crystals, starch grains, latex, oils, fats, proteins, chloroplasts, pigments (anthocyanins and carotenoids) regularly occur in many fruit walls. Both true fruits (kiwi and sapote) and accessory fruits (apple and pear) may be russeted from cork development in some of the outer cell layers. Functional nectaries (trumpet creeper) and hydathodes sometimes occur on fruits.

Exocarp. In accessory fruits the exocarp (sometimes called the epicarp) may merge imperceptibly with the noncarpellary tissue of the other fruit wall, although some multivars of apple exhibit an exocarp or core line (Fig. 2) of smaller cells. In true fruits the exocarp may consist only of epidermis (Fig. 6; pea) or may be multilayered and thus thick and distinct (Fig. 1; the skin of plum). A thin or thick cuticle is invariably present. The exocarp may be smooth (Fig. 6; avocado and grape) or pubescent, with unicellular or multicellular hairs, or both (Fig. 1; apricot and kiwi). Stomata may be present (citrus and kiwi) or absent (grape). In mature fruits lenticels may replace stomata (horse chestnut). The epidermal cells may develop into macrosclereids (Fig. 6). Often beneath the epidermis is a hypodermis or a subepidermal region (or both) one to several layers thick and containing pigments, tannins, crystals, starch grains, oil cells, oil cavities, collenchyma, fibers, or sclereids. Vascular tissue is usually absent. The exocarp may be fleshy or dry, and its cells may have thin or thick walls. The outer layers of noncarpellary tissue of accessory fruits may have the same structure as exocarps of true fruits. As noted above, russeting sometimes results from cork formation.

Mesocarp. The mesocarp (sometimes called the sarcocarp) occurs between the exocarp and endocarp. Other than the vascular tissue that is typically present, the mesocarp may contain only homogeneous parenchyma (Figs. 1, 2, and 6; grape and kiwi),

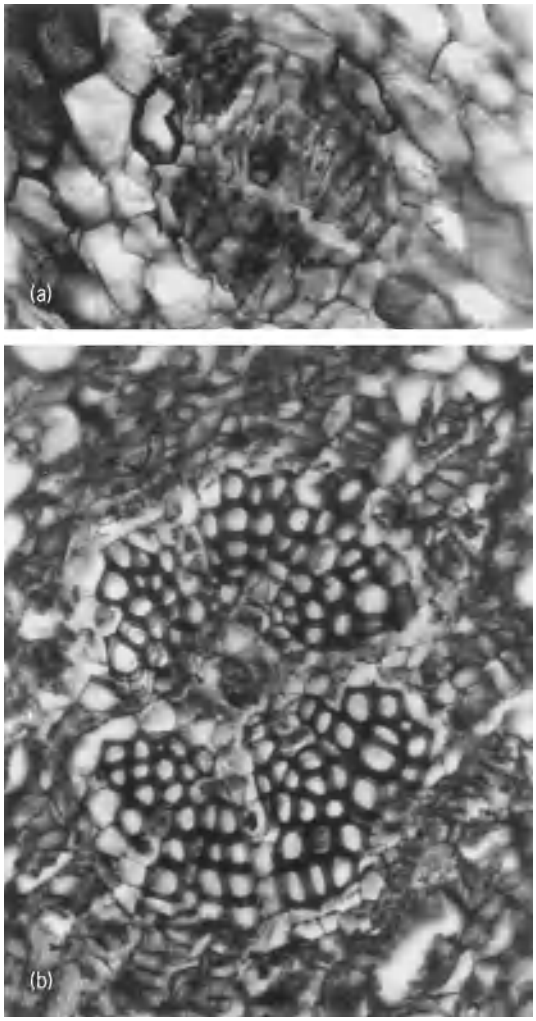

Fig. 7. Transverse sections of placental vascular bundles of (*a*) female flower bud and (*b*) mature fruit of *Actinidia chinensis* (kiwi). Both bundles are from comparable parts of the flower or fruit, both are mesarch amphicribral, and both are at the same magnification. The fruit bundle in *b* has a parenchymatous center and apparent secondary growth. During fruit development the large ovary bundles increase in size 40–80%. (*From R. Schmid, Reproductive anatomy of Actinidia chinensis (Actinidiaceae), Bot. Jahrb. Syst., 100:149–195, 1978*)

or intermixed parenchyma and sclerenchyma, especially fibers (legumes), or intermixed parenchyma and collenchyma (legumes). With maturity some parenchyma may become disorganized and show less wall thickening (peach and tomato). Sometimes mesocarp and endocarp are indistinguishable from each other.

Endocarp. The endocarp may be simply a thin- or thick-walled epidermis next to the ovary locule and seed (Fig. 6), but usually it is more complex (Figs. 1 and 2). The endocarp may be parenchymatous and anatomically like the mesocarp, but usually the endocarp is distinct, as in the highly sclerenchymatous stone and pome fruits (Figs. 1 and 2), or in legumes with layers of fibers often oriented crosswise. Collenchyma is sometimes present (legumes), and vascular tissue is frequently so. The epidermis often becomes sclerenchymatous and may bear a cuticle, hairs, stomata, or, in citrus fruits, outgrowths that are the juice vesicles.

Vascular system. The fruit wall, especially of large fruits, often contains much vascular tissue, either few large bundles or many small ones, or both. Frequently additional vascular bundles are differentiated to supplement the basic vascular pattern of the flower. Vascular tissue may occur throughout the pericarp and, in accessory fruits, also in noncarpellary tissue (Fig. 2). Often vascular bundles anastomose to form a complex network in the fruit wall (Fig. 2; kiwi). Fruit vascular bundles commonly are mesarch and amphicribral, with phloem surrounding xylem (Fig. 7), or endarch and collateral, with increased amounts of phloem (Fig. 2). Some bundles may be strictly phloic, whereas others may increase in diameter during fruit development by primary or secondary growth, or both (Fig. 7). *See* PHLOEM; XYLEM.

Abscission. Unless anatomically united with the carpel or carpels, floral parts usually abscise (separate and fall) during fruit development. Thus the ovary frequently is the only floral structure remaining at maturity. Sometimes perianth parts, especially sepals, are persistent. Fruits usually abscise from the plant when ripe and are dispersed (coconut, peach, and sycamore). The fruit and its stalk (peduncle or pedicel) may separate as a unit (apple and pear), or the stalk may remain on the branch when the fruit abscises (citrus and peach). The abscission zone, if present, may involve either the formation of a definite layer of corklike cells across the stalk or at the base of the fruit, or cytological changes in existing cells, especially in and between their cell walls. Some cells show no changes at abscission but are easily broken mechanically. Some fruits abscise without a definable abscission zone. *See* ABSCISSION.

Functional aspects. Most fruits increase in size compared to the flower, sometimes dramatically so (300,000 times in avocado). Large fruits generally require additional anatomical modifications for nutrition or support or both. The extra phloem in fruit vascular bundles and the often increased amount of vascular tissue in the fruit wall and septa supply nutrients to the developing seeds and, especially in fleshy fruits, to the developing walls. Large, especially fleshy fruits (apple, gourd, and kiwi) usually contain proportionally more vascular tissue than small fruits (achenes and follicles). Vasculary tissue also serves for support and in lightweight fruits may be the chief means of support. Collenchyma and particularly sclerenchyma are common supportive and protective tissues. Often the long axes of cells are oriented crosswise to give a mechanically strong structure.

Crystals, tannins, and oils commonly occur in fruits and may protect against pathogens and predators. The astringency of tannins, for example, may be repellent to organisms. With fruit maturation, tannin content ordinarily decreases, so the tannin repellency operative in early stages is superseded in fleshy fruits by features (tenderness, succulence, sweetness through odor and increased sugar content, and so on) attractive to animal dispersal agents. Many fruits are dispersed by hairs, hooks, barbs, spines, and sticky mucilage adhering the fruit to the surface of the dispersal agent. Lightweight fruits with many air spaces or with wings or plumes may be dispersed by wind or water. Gravity is always a factor in dispersal of fruits and seeds. Dehiscent fruits release seeds by varied and often complicated mechanisms. For example: increased turgor pressure may eventually cause the fruit to explode (squirting cucumber and touch-me-not); water loss by crossed superposed layers of cells, especially fibers, may set up antagonistic forces eventually causing the fruit to open (legumes), or ordinary cell enlargement may effect the same (jojoba). Small thin-walled cells often mark where the fruit wall actually splits (Fig. 6). *See* PLANT ANATOMY; REPRODUCTION (PLANT). Rudolf Schmid

Bibliography. B. G. Coombe, The development of fleshy fruits, *Annu. Rev. Plant Physiol.,* 27:207–228, 1976; A. Fahn and E. Werker, Anatomical mechanisms of seed dispersal, in T. T. Kozlowski (ed.), *Seed Biology,* vol. 1, 1972; I. Roth, Fruits of angiosperms, in Linsbauer's *Handbuch der Pflanzenanatomie*, vol. 10(1), 2d ed., 1977; R. Schmid, Functional and ecological interpretations of floral and fruit anatomy, *Bot. Rev.*, vol. 47, 1981.

Fruit, tree

Tree fruits include temperate, subtropical and tropical zone species. Most temperate zone fruits are deciduous, that is, they lose their leaves in the autumn. They are grown principally in regions protected from prolonged summer heat and severe winter cold (above −10 to −15°F or −23 to −26°C). The principal deciduous tree fruits grown in the United States are apple, peach, pear, plum, apricot, sweet cherry, tart cherry, and nectarine. Tree nuts, such as almond, pecan, walnut, and filbert, are sometimes classified as deciduous tree fruit crops.

Most subtropical fruit trees are evergreen. They will withstand temperatures somewhat below freezing during their dormant or semidormant season, but not the extreme temperatures tolerated by the temperate zone crops. Major subtropical fruits in the

United States are the citrus group (for example, orange, grapefruit, lemon), olive, avocado, fig, and others of lesser importance. The Japanese persimmon might be considered as a borderline case between the temperate and subtropical zone groups, and avocado is often classed as a tropical fruit.

Tropical fruits are killed or severely injured by temperatures below freezing and some are sensitive to cold temperatures above freezing. For example, papayas will survive in the milder areas of the subtropics but fail to mature fruits of good quality. The principal tropical fruits grown in the continental United States are date and mango, while banana and papaya are important in its island possessions. See separate articles on the common-name fruits and nuts.

Areas of production. California produces over 50% of the total United States deciduous fruit tonnage and is the number one state in production of peaches, pears, plums, apricots, and nectarines, as well as an important producer of apples and sweet cherries. Washington is first in production of apples and sweet cherries, while Michigan ranks first in tart cherry production. California also ranks first in production of lemons, olives, avocados, figs, dates, almonds, walnuts, grapes, and several other orchard crops. Florida is first in production of oranges and grapefruits.

Propagation. Nearly all tree fruits, whether deciduous or tropical, are propagated asexually by budding or grafting the desired variety onto the selected rootstock. The rootstock is usually a young seedling or a clonal stock. *See* PLANT PROPAGATION.

Most deciduous tree fruits are grown on seedlings of selected species. Trees grown on seedling stocks attain normal size, as opposed to dwarfing or semidwarfing stocks which have gained favor, particularly with the apple. Seedlings of peach, apricot, and several plum species (such a mariana and myrobalan) are used as stocks for peach, apricot, and plum. Mazzard and mahaleb cherry seedlings are used for most sweet and tart cherries. Apple and pear seedlings are used for standard-sized trees of these two fruits, while selected dwarfing clones of apple are used for dwarf apple trees, and quince is used as a pear dwarfing stock. Several citrus species are used as rootstocks for the various citrus fruits.

Cultivation. The tree fruit crops are grown commercially in orchards or groves, usually in single rows which permit necessary cultural operations for each tree. Nearly all tree fruit crops have similar requirements of training, pruning, spraying to control diseases and insects, and cultivation or chemical control of weeds. Most tree fruits receive applications of nitrogen and other nutrients as required. In areas of low rainfall, orchards are irrigated.

Most orchard crops are harvested by hand, although mechanical harvesting has been successful for some processing crops, including tart cherries and plums for drying (prunes). *See* AGRICULTURAL MACHINERY.

R. Paul Larsen

Diseases. Tree fruits have diseases that may spoil the appearance of the fruits or render them unfit for human consumption; diseases may also lower yield or shorten the productive life of the trees. Most of the tree-fruit diseases are highly specialized, attacking specific hosts. Only infrequently will one tree-fruit disease attack a wide range of fruits.

Fig. 1. Apple scab on young fruit.

Temperate-zone fruits. Diseases of temperate-zone fruits (sometimes called deciduous fruits) are due to fungi, bacteria, viruses, and adverse environmental conditions.

Fungi are the most frequent causes of temperate-zone tree-fruit diseases. For example, a perennial threat to most commercial and home apple orchards is the apple scab disease (**Fig. 1**). Caused by the fungus *Venturia inaequalis*, scab may cause premature defoliation, June drop of young fruits, and unsightly blemishes on ripe apples. A similar disease attacks most varieties of pears and is caused by *V. pyrina*. *See* FUNGI.

Stone fruits, such as peaches, plums, cherries, apricots, and nectarines, are frequently damaged extensively by brown rot caused by the fungus *Monilinia fructicola* (**Fig. 2**). This disease attacks the flowers (causing blossom blight), infects terminals and twigs with twig blight, and is widely recognized by its destruction of ripe fruit at harvest time.

Many diseases caused by fungi attack and destroy leaves of tree fruits: cherry leaf spot, *Coccomyces hiemalis;* peach leaf curl, *Taphrina deformans;* apple cedar rust, *Gymnosporangium juniperi-virginianae;* leaf blight of pears, *Mycosphaerella sentina;* and apple powdery mildew, *Podosphaera leucotricha*.

Bacteria also attack tree fruits, causing extensive damage in some regions. For example, fireblight of apples and pears, *Erwinia amylovora*, is a common disease of pomaceous fruits in most seasons (**Fig. 3**). Bacterial leaf spot of stone fruits, *Xanthomonas pruni*, is often the cause of serious damage to leaves,

Fig. 2. An example of brown rot on plums.

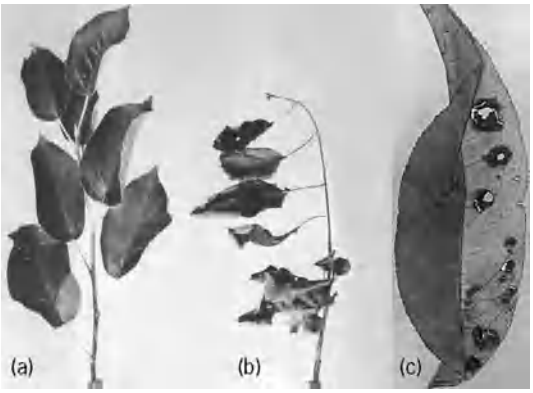

Fig. 3. Several examples of bacterial plant diseases. (*a*) Healthy apple twig and (*b*) fireblight-diseased apple twig. (*c*) Bacterial leaf spot of peach.

twigs, and fruits of peaches, apricots, almonds, cherries, plums, and nectarines. Crown gall, *Agrobacterium tumefaciens*, may infect tree fruits in the nursery row, causing the formation of large galls on roots and stems of young apple, peach, and pear trees. *See* BACTERIA.

Diseases caused by viruses, such as yellows, rosette, and phony, are especially prevalent on peaches, plums, cherries, apricots, and almonds. *See* PLANT VIRUSES AND VIROIDS.

The growth of fruit trees and the production of fruit may also be affected by unfavorable environmental conditions such as infertile soil, low temperatures, drought, and poor drainage. Eric G. Sharvelle

Citrus diseases. Citrus is subject to a great number of diseases, most of which are distributed throughout the citrus-producing countries of the world. Infectious diseases of citrus are caused by bacteria, fungi, and viruses and may be grouped according to the parts of trees affected, that is fruit, leaves and twigs, and trunks and roots.

Numerous rots, internal derangements, and external markings result from fungal and bacterial infections. Fruit decay develops chiefly during storage and transit from infections established largely in surface wounds during growing and from picking, packing, or marketing operations. Citrus fruit requires careful handling to avoid economic loss from decay caused by such fungi as the common green mold, *Penicillium digitatum;* blue mold, *P italicum* (**Fig. 4**); brown rot, *Phytophthora citrophthora;* and cottony-rot, *Sclerotinia sclerotiorum*. Some decays are described as contact rots because the causal fungi can spread through an entire box of fruit from one infected fruit.

Some of the fruit-rotting fungi and others cause leaf-blight and dieback of twigs. A bacterium, *Xanthomonas citri*, is the cause of citrus canker which attacks leaves, twigs, and fruits. Introduced on citrus plants sent from Japan to Florida in 1913, it threatened to destroy the citrus industry there. Its eradication, requiring destruction by burning of 4,000,000

Fig. 4. Blue-green molds (*Penicillium italicum* and *P. digitatum*) which take a heavy toll in storage, transit, and market. The illustration shows how these decay-causing molds can spread by contact from a single infected citrus fruit. (*Photograph by L. J. Klotz*)

Fig. 5. Gummosis, a disease caused by the fungus *Phytophthra citrophthora*, which damages or kills many citrus trees. Infection by this fungus, a water mold, is favored by excessive moisture. (*Photograph by L. J. Kotz*)

orchard and nursery trees, is the most remarkable achievement in plant disease control.

Some of the fruit-rotting fungi, particularly species of *Phytophthora* and *Diplodia*, are responsible for root rots and gumming diseases of trees (**Fig. 5**).

Citrus is subject to several destructive virus diseases. One of these, psorosis, causes development of bark lesions and eventual death of trees. Since it was learned that psorosis is a bud-perpetuated disease that does not spread naturally, it has been possible to avoid it by propagating from psorosis-free sources. Another virus disease, tristeza, causes rapid decline or death of trees of sweet orange, grapefruit, and tangerine propagated on certain susceptible rootstock varieties. Tristeza virus is known to be spread by three species of aphids. It virtually destroyed the citrus industries of Argentina and Brazil in the 1930s. Damage from tristeza in the United States has been less because more of the citrus is on resistant rootstocks and the two species of the genus *Aphis* that spread it here are less efficient vectors than the one species common in South America. James M. Wallace

Fucales

A large and diverse order of conspicuous brown algae (Phaeophyceae) including such well-known seaweeds as *Fucus* and *Sargassum*. Definitive features include apical growth, production of sex organs in cavities, and a life history with only one somatic phase, which is diploid.

General structure. The mature thallus consists typically of a holdfast from which arise several axes with dichotomous or radial primary branching (**Fig. 1***a*). Growth is initiated by an apical meristem that results from the stimulation of neighboring cells by one or more apical cells located in a depression at the tip of each branch. An apical cell has lateral and basal cutting faces. Lateral derivatives of an apical cell continue to divide, producing a thin, pigmented, surface layer (meristoderm) and a thicker, colorless layer (cortex). The innermost lateral derivatives and the basal derivatives give rise to the medulla which, because of stretching, becomes filamentous at maturity. Vegetative thalli range in size, shape, and complexity from a small top-shaped structure, as in *Himanthalia*, to densely branched axes several meters long, differentiated into stems, leaves, and flotation organs (bladders or vesicles), as in *Sargassum*.

Reproduction. Sexual reproduction is oogamous, with oogonia and antheridia produced in flask-shaped cavities called conceptacles (Fig. 1*b*, *c*). Usually the conceptacles are localized at the swollen tips of vegetative branches or in special branches that range from long, dichotomously branched straps emerging from the vegetative thallus (*Himanthalia*) to small clusters of slender cylindrical axes borne in the axils of leaves (*Sargassum*).

The thallus in Fucales is diploid, with meiosis taking place within the sex organs. From one to eight eggs are produced in each oogonium, and usually

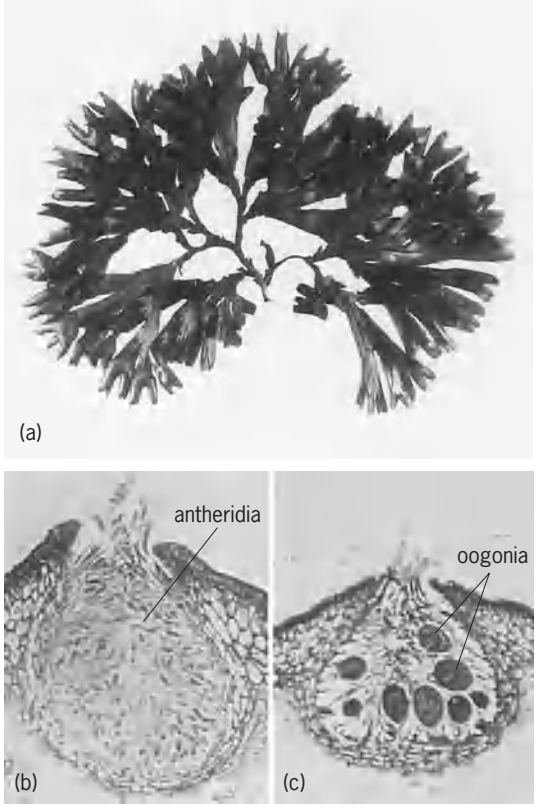

Fig. 1. ***Fucus*, a typical member of the order Fucales. (*a*) Plant with basal disk or holdfast and dichotomously branched frond with inflated tips containing reproductive cavities or conceptacles. (*b*) Section through male conceptacle lined with branched hairs bearing antheridia. (*c*) Section through female conceptacle lined with hairs and oogonia. (*Parts b and c from H. J. Fuller and O. Tippo, College Botany, rev. ed., Holt, 1954*)**

64 antherozoids (sperm) in each antheridium. Oogonia and antheridia occur on the same thallus or on different thalli. The antherozoid, like all motile reproductive cells in Phaeophyceae, is biflagellate but in some genera it departs from the usual situation by having the smooth, posteriorly directed flagellum longer than the hairy, anteriorly directed flagellum. Numerous antherozoids attempt to penetrate an egg, but the initial entry triggers the immediate formation of a fertilization membrane, which inhibits further entry. The union of gametes is followed by cell divisions that occur in a fairly precise order and result in a free-living embryo. The development of the embryo, which is easily observed and controlled in culture, has been studied extensively to provide information on the establishment of polarity by environmental factors. At the tip of the embryo there is a tuft of hairs, each with a meristematic basal region. The upper portion of one or more of these hairs abscises, while the lowermost meristematic cell metamorphoses into an apical cell.

The fucalean life history has been interpreted in two ways. In the first interpretation, the events are taken at face value: the thallus is considered a diploid gametophyte. In the second interpretation, the life history is compared to that of higher plants, in which the gametophytes are reduced to

the cytological level. The oogonia and antheridia are considered megasporangia and microsporangia, the nuclei resulting from meiosis are considered megaspores and microspores, while subsequent nuclear divisions leading to the gametes supposedly represent gametophytes that have been lost in the course of evolution. Support for this interpretation is given by the observation of walls within the oogonium in certain genera.

Important families and species. In the family Fucaceae (rockweeds), the thallus typically is composed of a discoid or conical holdfast with a cluster of erect fronds that are branched dichotomously in one plane. In most genera, conceptacles are borne on swollen tips of branches. *Fucu* is common throughout the North Atlantic and North Pacific. Vesicles, when present, are often paired, with one on either side of the midrib. *Pelvetia*, similar to *Fucus* but with segments that are narrower, without a midrib, and usually without vesicles, is also common in the Northern Hemisphere, but is not found on the American side of the Atlantic. *Ascophyllum* occurs abundantly in the North Atlantic, both in exposed sites and in estuaries far from the open sea. It is characterized by large vesicles on the midline of the axis, and special swollen reproductive branches borne laterally. Loose, pinnate branching is superimposed on the primary dichotomous branching.

In subtropical and tropical regions of all oceans, the family Sargassaceae is common, dominating the biomass of benthic seaweeds (**Fig. 2**). The holdfast is discoid, conical, hapteroid (with rootlike outgrowths), or rhizomatous. The primary axis is condensed or percurrent (running from base to apex) and bears radially arranged secondary axes, which in turn produce radially or bilaterally arranged laterals. The primary axis is usually perennial, while the secondary axes are renewed annually. Vegetative branches as well as fertile branchlets are borne in the axils of subtending leaves.

Fig. 2. ***Sargassum*, a member of the family Sargassaceae. Plant shows percurrent axis, dentate leaves, spherical vesicles or bladders, and fertile branchlets borne in axils of leaves.**

Sargassum is by far the largest and most important genus in the Sargassaceae. Hundreds of species and varieties have been described. Its taxonomy is complicated by morphological variation related to season, age, and habitat. In that portion of the North Atlantic between the West Indies and the Azores, an abundance of seaweed floats on relatively calm water. This region is called the Sargasso Sea because *Sargassum* is the primary constituent of the floating vegetation. The species of *Sargassum* in the Sargasso Sea reproduce solely by fragmentation.

Perhaps the best-known species of *Sargassum* is *S. muticum*, which is modest and well behaved in its native land (Japan) but noxiously intrusive elsewhere. It is a weed in the sociological sense of a plant growing outside its original range and in such a manner as to be unwanted. It was introduced to northwestern North America, presumably in shipments of oyster spat, just before World War II, and now occurs in protected waters from British Columbia to Baja California. It was detected in southern England in 1973 and has subsequently spread to nearby shores on the European continent. There is fear that it will eventually displace native species, and the long dense fronds are known to interfere in the operation of small boats in harbors.

In *Turbinaria*, a much smaller genus of Sargassaceae, the thallus is constructed on the same basic plan as *Sargassum*, but it is condensed into a structure resembling a gymnosperm cone. The leaves are modified into top-shaped or obpyramidal structures that partly conceal the fertile branchlets, like cone scales concealing seeds.

The family Cystoseiraceae, closely related to the Sargassaceae, is found in all temperate waters, but is especially well represented in southern Australia, where a consortium of members forms a nearly continuous belt in the sublittoral fringe.

The family Hormosiraceae comprises the single genus *Hormosira*, with probably only one species (*Hormosira banksii*) represented by several growth forms. This alga is common in shallow waters of New Zealand and the cool-temperate portion of Australia. The thallus has a striking appearance, being composed of swollen segments arranged like beads in a necklace. The conceptacles are scattered over the surface of the segments.

Rockweeds are well known to coastal peoples of the temperate regions of the Northern Hemisphere. Traditionally, the seaweed was either plowed into the soil as a green manure or burned to an ash that was spread on the fields. In modern chemical industry, *Ascophyllum* is an important source of alginate, especially in Norway. In a pulverized form, it is added to livestock and poultry feed for nutritive purposes. On the east coast of the United States, it is used

to pack bait worms for shipment, and clumps are frequently encountered in San Francisco Bay, where they have been discarded from fishing boats, but so far the alga has not established itself there. Rockweeds are important ecologically because they protect a host of intertidal animals from desiccation during low tides. In warm waters, *Sargassum* is an especially significant constituent of the ecosystem because of its abundance and biomass. It is a minor source of alginate. *See* ALGAE; ALGINATE; PHAEOPHYCEAE. Paul C. Silva; Richard L. Moe

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae: Structure and Reproduction*, 2d ed., 1997; J. B. Jensen, Morphological studies in Cystoseiraceae and Sargassaceae (Phaeophyceae) with special reference to apical organization, *Univ. Calif. Publ. Bot.*, 68:1–61, 1974.

Fuel cell

An electrical cell that converts the intrinsic chemical free energy of a fuel directly into direct-current electrical energy in a continuous catalytic process. As in the classical definition of catalysis, the fuel cell itself should not undergo change; that is, unlike the electrodes of a battery, its electrodes ideally remain unchanged. For most fuel–oxidant combinations, the available free energy of combustion is somewhat less than the heat of combustion. In a typical thermal power conversion process, the heat of combustion of the fuel is turned into electrical work via a Carnot heat-engine cycle coupled with a rotating electrical generator. Since the Carnot conversion rarely proceeds at an efficiency exceeding 40% because of heat source and sink temperature limitations, the efficiency of conversion in a fuel cell can be greater than in a heat engine, especially in small devices. *See* CARNOT CYCLE; CATALYSIS.

Principle of operation. The fuel cell reaction usually involves the combination of hydrogen (H) with oxygen (O) [reaction (1)], as shown in the **illustration**.

$$H_2(g) + {}^1/_2O_2(g) \rightarrow H_2O(l) \qquad (1)$$

Under standard conditions of temperature and pressure, 25°C (77°F) and 1 atm (100 kilopascals), the reaction takes place with a free-energy change $\Delta G = -237$ kilojoules per mole of liquid water. Since the formation of water involves two electrons, this value corresponds to −1.23 electronvolts (1 eV = 96.5 kJ/equivalent). Thus, at thermodynamic equilibrium (zero current), the cell voltage should theoretically be 1.23 V, yielding a theoretical efficiency based on the higher heating value heat of combustion [ΔH for $H_2O(l) = -1.48$ eV] of 83.1%. If gaseous water at 100°C (212°F) is the product, the corresponding ΔG and lower-heating value ΔH are −1.16 eV and −1.25 eV. *See* FREE ENERGY.

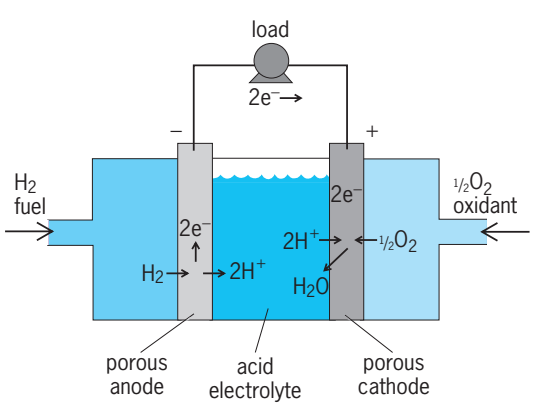

Principle of operation of a fuel cell. (*After A. J. Appleby and F. R. Foulkes, Fuel Cell Handbook, Van Nostrand Reinhold, 1989*)

At a net (nonzero) current, all cells show losses in cell voltage (V). In low-temperature fuel cells, these are due largely to the kinetic slowness (irreversibility) of the oxygen reduction reaction, which requires the breaking of a double bond with transfer of four electrons per molecule in a complex sequence of reactions. Such cells show deviations from the theoretical 1.23-V open circuit. Typical values are 0.9 to 1.0 V. In high-temperature systems, oxygen reduction losses are less significant, since the reaction rate increases with temperature. However, the available free energy then decreases, falling to a value corresponding to about 1.0 V at 1000°C (1832°F). A further thermodynamic loss results from high cell fuel (or oxidant) conversion to avoid waste, so that the effective reversible potential is displaced from the standard state. At high temperature, the major loss is thermodynamic, which tends to compensate for the irreversible oxygen electrode losses at low temperature. As a result, cell voltages under typical loads vary from about 0.65 V for simple terrestrial cells to 0.9 V for aerospace cells. Cell voltage falls with increasing current per unit area. Since cell thermal efficiency based on the higher heating value of hydrogen (HHV, liquid water product) is given by $V/1.25$, cell performance is a compromise between relative cost (that is, kilowatts available per unit area) and fuel efficiency, to give the lowest cost of electricity for a given application. *See* OXIDATION-REDUCTION.

Fuels. While any chemically suitable fuel, including metals such as lithium (Li), sodium (Na), aluminum (Al), and zinc (Zn), may be used in a fuel cell, hydrocarbons (for example, natural gas) will not react at a significant rate in low-temperature fuel cells. They will crack thermally before reacting electrochemically if injected directly into high-temperature fuel cells. Practical hydrogen-bearing fuels have reactivities in low-temperature cells that increase in the order hydrocarbons << ammonia < methanol << hydrazine < hydrogen, provided that suitable anode catalysts are used. Simple low-power units operating directly on methanol at ambient temperature do find some use, and liquid-fueled hydrazine cells have also found specialized applications. However, the high manufacturing energy requirement for hydrazine, together with its high cost and hazardous nature, leaves hydrogen the only suitable general high-performance fuel candidate. *See* AMMONIA; HYDRAZINE; HYDROCARBON; HYDROGEN; METHANOL.

As oxidant, oxygen from air or as a cryogenic liquid for use in space is the first choice, although liquid oxidants (for example, hydrogen peroxide or nitric acid) have occasionally been used in specialized applications.

Because of the low solubility of H_2 and O_2 in electrolytes, a large area of contact at the electrode–electrolyte interface is required to achieve high reaction rates (0.2–0.8 A/cm^2 at atmospheric pressure, depending on the technology, and up to several amperes per square centimeter at elevated pressure). The large contact area is required because the reactions take place at the electrode–electrolyte interface. This large area is obtained by using porous catalytic materials with high surface areas that provide contact between electrolyte and gas, catalyze the reaction, maintain the electrolyte as a very thin layer at the surface of the electrode, and conduct electrons at the external circuit.

For practical fuel cells, hydrogen can be produced from readily available fuels, such as clean light distillate (for example, naphtha) usually by steam reforming, or from coal via gasification at high temperature (the direct use of coal or carbon has been abandoned). In the high-temperature cells under certain conditions, internal steam reforming of simple hydrocarbons and alcohols (for example, methane and methanol) can take place by the injection of the fuel with steam, which avoids cracking. Since methanol fuel reacts only slowly at low temperature, it is also steam-reformed to hydrogen. Methanol reforming takes place at only about 350°C (660°F), giving mixtures of hydrogen and carbon dioxide (CO_2) with a small amount of carbon monoxide (CO). In contrast, steam reforming of higher-molecular-weight alcohols or clean light distillates requires temperatures in excess of 700°C (1290°F). This favors mixtures of hydrogen and carbon monoxide, as in coal synthesis gas. *See* COAL GASIFICATION; SYNTHETIC FUEL.

Classification. Since the electrolyte that is used determines the chemical environment, hence stability, of the electrodes and other components, fuel cells are conveniently classified by electrolyte. There are three broad groups: aqueous, molten salt, and high-temperature solid ionic conductor.

Aqueous. To avoid excessive polarization, the major conducting ion in the electrolyte must be produced in the cell reaction at one electrode and must be consumed in the other. A good example is the aqueous alkaline fuel cell, using potassium hydroxide (KOH) as the electrolyte, with electrode reactions (2) and (3). In these reactions, the hydroxide ion (OH^-) is

$$2H_2 + 4OH^- \rightarrow 4H_2O + 4e^- \quad \text{(anode)} \qquad (2)$$

$$O_2 + 2H_2O + 4e^- \rightarrow 4OH^- \quad \text{(cathode)} \qquad (3)$$

the conducting ion. In acid cells, the conducting ion is hydrogen ion (H^+), produced from H_2 at the anode and consumed to give water at the cathode. Alkaline cells are capable of very high performance, and they have a benign corrosion environment that can allow the use of inexpensive (non-noble-metal) cathode catalysts. However, they cannot tolerate carbon dioxide and are not suitable for direct use with reformed carbonaceous fuels. Because of their efficiency, they have been used in space, operating at 260°C (500°F) with sintered nickel oxide cathodes used in the Apollo Program, and at about 85°C (185°F) in the high-performance lightweight system in the space shuttle orbiter. Both used cryogenic hydrogen and oxygen supplied at about 4 atm (400 kPa) pressure. The orbiter cell uses gold alloy cathodes and platinum alloy anodes bonded with nonwetting Teflon to maintain a stable interface. Low-capital-cost alkaline systems operating on "green" hydrogen were attracting attention in 2005.

Acid cells reject carbon dioxide and hence are suitable for use on hydrogen–carbon dioxide mixtures. However, carbon monoxide is an anode catalyst poison above certain critical levels that depend on cell operating temperature and anode catalyst composition. Unlike alkaline cells, acid cells require noble-metal catalysts, though they can be supported economically on nano-size carbon black or graphitic particles in low loadings (3 g or 0.1 oz per kilowatt or less). A high-performance system (the polymer electrolyte membrane or proton exchange membrane fuel cell, PEMFC) that has found favor for small units uses a stable fluorinated sulfonic acid polymer as the electrolyte. It is normally used with pure hydrogen fuel. Units vary in size from a few watts (operating directly on methanol as fuel) for portable power applications to greater than 100 kW (for experimental vehicle propulsion using pure hydrogen fuel). They operate at temperatures up to about 70–80°C (160–175°F), limited by the requirement for the presence of liquid water to ensure efficient proton conduction in the electrolyte. This means that they have very low (parts per million) carbon monoxide tolerance.

Throughout the 1990s, the most important acid fuel cell was the phosphoric acid fuel cell (PAFC) that used phosphoric acid as the electrolyte at 190–205°C (375–400°F), with Teflon-bonded low-loading platinum-based electrodes and graphite intercell separator plates. The high operating temperature was possible because self- ionizing phosphoric acid does not require liquid water for proton conduction. It made the anode catalysts tolerant to up to 1.5% CO input, simplifying the cleanup of the hydrogen-carbon dioxide feedstocks produced by steam reforming. It also simplified the cooling system, giving waste heat for cogeneration, and produced steam for reforming so no extra fuel was required for this purpose. This integrated system approach increased overall efficiency. Exhaust emissions levels are very low, allowing the power plant to be used where required, even in cities. *See* COGENERATION; ELECTRIC POWER GENERATION.

Molten salt. The only fuel-cell technology using molten salt electrolyte is the molten-carbonate fuel cell (MCFC), using molten carbonate (typically Li-K or Li-Na) in the temperature range 600–700°C (110–1300°F). The cells require that gaseous CO_2, recycled from the anode, exit to the cathode, as shown

Cell and system thermal efficiencies for different fuel-cell technologies and operating conditions

Type	Temperature, °C (°F)	Pressure, atm (kPa)	Current density, mA/cm^2	Voltage, V	System thermal efficiency (ac output/natural gas HHV), %
Alkaline	70 (158)	1 (100)	200	0.8	32
Phosphoric acid	195 (374)	1 (100)	330	0.65	36
PEM	75 (167)	1 (100)	400	0.7	35
Molten carbonate (internal reforming)	650 (1202)	1 (100)	160	0.78	45.6
Solid oxide (integrated reforming)	1000 (1832)	1 (100)	200	0.68	43.3

in the electrode reactions (4) and (5). MCFCs are attractive because they use relatively inexpensive materials. They are particularly suited for use with reformed hydrocarbon fuels, since they can consume CO directly via the internal water–gas shift process;

$$H_2 + CO_3^{2-} \rightarrow H_2O + CO_2 + 2e^- \quad (4)$$

$$CO_2 + {}^1/_2O_2 + 2e^- \rightarrow CO_3^{2-} \quad (5)$$

and they can be designed to reform natural gas internally if the stack (fuel cells connected in series) gas-flow channels incorporate a reforming catalyst. This has the engineering advantage of adsorbing 40% of the stack waste heat, greatly simplifying cooling by reactant gas streams. This is the principle of the direct fuel cell (DFC). Cell and system performance is higher than that of the PAFC, and they may attain over 50% system efficiency, or higher values, if they use a gas or steam turbine bottoming cycle powered by their high-quality waste heat. *See* COGENERATION.

MCFCs were first developed as stacks in the early 1980s and were scaled up to the 100-kW demonstrator stage by 1989. During the 1990s, a 2-MW DFC unit was demonstrated. Two other 2-MW units and a 1-MW unit operating on digester gas followed. The standard product is a 250–300-kW DFC unit with 0.78 m^2 (8.4 ft^2) active area cells.

High-temperature solid ionic. The solid-oxide fuel cell (SOFC) operates between 800 and 1000°C (1470 and 1830°F) and uses thin-layer yttria-stabilized zirconia as the electrolyte. Conducting oxides matched for thermal expansion are used as the cathode and to interconnect the cells. The anode is a nickel cermet. One such system uses tubular 2 m long by 2.2 cm diameter cells, 1.5 m in active length (834 cm^2 active area). The cells are made from porous cathode material and are supplied by air through a central alumina tube that acts as a ceramic heat exchanger. Partially reformed natural gas and steam are passed between the cells from the direction of the closed ends. Cells are linear-interconnected to the inner cathode extending through the electrolyte and anode, and insulated from the latter. The surface of the interconnect is connected to the anode of the next tube via a nickel felt. *See* CERMET.

Unused reformate (15%) burns with unused air in an alumina plenum chamber at the top. The system offers high-quality waste heat. Internal reforming at the operating temperature is so rapid that too much cell cooling occurs, so separate integrated prereforming is employed, using part of the waste heat. The remainder can be used in cogeneration, or in a 3-atm (300-kPa) pressurized system the heat can be supplied to a turbine, which drives the compressor and produces additional power. Scale-up to 250 kW took place in the 1990s. Several 250-kW units, including one pressurized system, have been delivered.

Elsewhere, researchers are concentrating on planar cells, in which one major problem is sealing. Research is aimed at reducing the SOFC operating temperature by the use of thinner components and/or the use of different families of conducting materials. This approach will allow the use of metal heat exchangers and treated metal bipolar plates.

Performance. The **table** shows cell and system thermal efficiencies for different fuel-cell technologies and operating conditions using hydrogen derived from natural gas via steam reforming. System efficiencies are based on the higher heating value (HHV) of natural gas and include parasitic power losses and dc/ac conversion. The HHV is used because the gas is purchased based on this value. For rotating machinery, the lower heating value (LHV) efficiencies are usually used. These are a factor of 1.1 higher. The thermal efficiency of the alkaline system is low because pure hydrogen must be separated and fuel must be burned for reforming. The two high-temperature systems are assumed to have no bottoming cycle. Assumptions used are SOFC fuel utilization 85%, MCFC 78%, inverter efficiency 93%, with auxiliaries 7% of ac output. They reflect the efficiency of heat recovery in reforming and upgrading methane ($\Delta H = -1.04$ eV for gaseous water product) to hydrogen ($\Delta H = -1.25$ eV) internally within the cell and system itself. These show the concern with fuel-cell power-system design, rather than simply emphasizing the fuel cell itself. In practice, systems would probably be operated at higher current density and lower efficiency to reduce capital cost and, therefore, the cost of electricity. A. John Appleby

Bibliography. A. J. Appleby, Fuel cell technology, status and future prospects, *Energy*, 21(7–8):521–653, Elsevier Science, Amsterdam, 1996; A. J. Appleby and F. R. Foulkes, *Fuel Cell Handbook*, Van Nostrand Reinhold, 1989; R. L. Busby, *Hydrogen and Fuel Cells: A Comprehensive Guide*, PennWell, 2005; P. Hoffman, *Tomorrow's Energy, Hydrogen,*

Fuel Cells, and the Prospects for a Cleaner Planet, MIT Press, 2002; K. Kordesch and G. Simader, *Fuel Cells and Their Applications*, Wiley-VCH, 1996; J. Larminie and A. Dicks, *Fuel Cell Systems Explained*, 2d ed., Wiley, Chichester, U.K., 2003.

Fuel gas

A fuel in the gaseous state whose potential heat energy can be readily transmitted and distributed through pipes from the point of origin directly to the place of consumption. The development and use of fuel gases is closely associated with the progress of civilization. Such gases have become prominent in the industrial development period since 1900.

The types of fuel gases are natural gas, LP gas, refinery gas, coke oven gas, and blast-furnace gas. The last two are used in steel mill complexes. Typical analyses of several fuel gases are presented in the **table**. Since these analyses are based on dry gases, the heating value of gases saturated with water vapor would be slightly lower than the values shown.

Most fuel gases are composed in whole or in part of the combustibles hydrogen, carbon monoxide, methane, ethane, propane, butane, and oil vapors and, in some instances, of mixtures containing the inerts nitrogen, carbon dioxide, and water vapor.

Natural gas. The generic term "natural gas" applies to gases commonly associated with petroliferous geologic formations. As ordinarily found, these gases are combustible, but nonflammable components such as carbon dioxide, nitrogen, and helium are often present. Natural gas is generally high in methane. Some of the higher paraffins may be found in small quantities.

The olefin hydrocarbons, carbon monoxide, and hydrogen are not present in American natural gases. The term "dry natural gas" indicates less than 0.1 gal (1 gal, U.S. = 0.003785 m^3) of gasoline vapor occurs per 1000 ft^3 (1 ft^3 = 0.028 m^3); "wet natural gas" indicates more than 0.1 gal/1000 ft^3. "Sweet" and "sour" are terms that indicate the absence or presence of hydrogen sulfide.

There is no single composition which might be termed typical natural gas. Methane and ethane constitute the bulk of the combustible components; and CO_2 and nitrogen, the inerts. The net heating value of natural gas served by a utility company is often 1000 to 1100 Btu/ft^3 (3.8 to 4.2 × 10^7 joules/m^3).

Natural gas is an ideal fuel for heating because of its cleanliness, ease of transportation, high heat content, and the high flame temperature. *See* LIQUEFIED NATURAL GAS (LNG); NATURAL GAS.

LP gas. This term is applied to certain specific hydrocarbons, such as propane, butane, and pentane, which are gaseous under normal atmospheric conditions but can be liquefied under moderate pressure at normal temperatures. *See* LIQUEFIED PETROLEUM GAS (LPG).

Oil gas. This term encompasses a group of gases derived from oils by exposure of such oils to elevated temperatures. Refinery oil gases are those obtained as by-products during the thermal processing of the oil in the refinery. They are used primarily for heating equipment in the refinery. Gas made by thermal cracking of oil was formerly very important as an urban fuel gas, but has been almost totally displaced by natural gas, and the equipment for its production dismantled. A typical oil gas consists of saturated and unsaturated hydrocarbons and has a heating value of 1300 to 2000 Btu/ft^3. Methane, ethane, propane, butane, ethylene, and propylene are the main constituents.

Coal gas. Until about 1940, gas produced from coal was an important part of the energy mix in the United States. These gases were rapidly replaced by natural gas in the distribution systems of gas utilities serving the residential, commercial, and industrial markets. During 1971 it became clear that the supply of natural gas would probably be insufficient to meet the ever-increasing demand. Therefore, it is expected that gas will once again be produced from coal. The older processes will be replaced by modern methods capable of very high production rates at improved efficiency and reduced cost.

Coal gasification can be accomplished in a large number of ways, including pyrolysis or partial oxidation with air or oxygen and steam. Various processes operate as fixed beds or fluidized beds or with the coal entrained. The pressure may vary from near atmospheric to 1000 pounds per square inch gage (1 lb/in.^{2} = 6895 pascals) or more. All operate at high temperatures (1200–3000°F; 650–1645°C). The direct products of gasification vary in heating value from 120 to 150 Btu/ft^3 or 4.6 to 5.7 × 10^3 J/m^3

Typical gas analyses*

Type	Analysis, % vol								Specific gravity	Btu/ft³†	
	CO_2	O_2	CO	H_2	CH_4	C_2H_6	C_3H_8 and C_4H_{10}	N_2		Gross	Net
Dry natural gas	0.2				99.2			0.6	0.56	1007	906
Propane (LP)						2.6	97.3	0.1	1.55	2558	2358
Refinery oil gas		0.2	1.2	6.1	4.4	72.5	15.0	0.6	1.00	1650	1524
Coke oven gas	2.0	0.3	5.5	51.9	32.3		3.2	4.8	0.40	569	509
Blast-furnace gas	11.5		27.5	1.0				60.0	1.02	92	92
Producer gas	8.0	0.1	23.2	17.7	1.0			50.0	0.86	143	133

*From L. Shnidman (ed.), *Gaseous Fuels,* 2d ed., American Gas Association, 1954.
†1 Btu/ft^3 = 3.8 × 10^4 joules/m^3.

(low-Btu gas), to 300 Btu/ft^3 or 11.4×10^7 J/m^3 (medium-Btu gas or synthesis gas), to as high as 600 Btu/ft^3 or 2.28×10^7 J/m^3. Low-Btu gas can be used as fuel for industrial processes or for production of electrical power by electric utilities. Medium-Btu gas, which consists principally of carbon monoxide and hydrogen, can be used directly as a fuel, or it can be upgraded by catalytic methanation to essentially pure methane, which is, for all practical purposes, identical to natural gas. Such a product gas is commonly called pipeline-quality gas or substitute natural gas (SNG). *See* COAL GASIFICATION.

Low-Btu gas. Producer gas is made by oxygen-deficient combustion of coal or coke, in which process a mixture of air and steam is blown upward through a thick hot bed of coal or coke. The gas is high in nitrogen introduced in the air. Its heating value is low, its specific gravity is high, and the percentage of inerts is high. Producer gas, which contains 23–27% carbon monoxide, is used as it comes from the generators after some preliminary purification. It was once the cheapest form of industrial gas, and could again become important; it may be possible to produce it in modernized equipment on a large scale.

Blast-furnace gas is a by-product from the manufacture of pig iron. Like producer gas, it is derived from the partial combustion of coke. Some of the combustibles in the gas are used to reduce the iron ore; thus the final gas contains about 27% carbon monoxide and more than 70% of inert gases (CO_2 and N_2), giving it the lowest heating value, less than 100 Btu/ft^3 or 3.8×10^6 J/m^3, of any usable fuel gas. It is used for the operation of gas engines, heating by-product coke ovens, steel plant heating, steam raising, and crude heating. *See* FOSSIL FUEL; PYROLYSIS.

Medium-Btu gas. Coke oven gas is the only important gas that has an intermediate heating value. The gas is made by destructive distillation of a packed bed of coal out of contact with air. The process results in the formation of coke, which is used in the blast furnace. The gas is utilized totally within the steel-making complex.

The combination of the 1974 oil embargo, the energy crisis, and restrictions on air pollution resulted in the proposal of a large number of modern processes. Nearly all these processes use a mixture of steam and oxygen to combust coal. Less oxygen than is required for complete combustion is used; therefore, the products are primarily carbon monoxide and hydrogen. In some cases, depending on the end use of the gas, moderate amounts of methane and liquid products are produced. *See* COKE.

SNG. Gas which is to be used within the modern natural gas transmission and distribution system must be essentially methane. The medium-Btu gas processes can be used as precursors to produce substitute natural gas (SNG), as noted above. Other new processes which operate at very high pressure (1000 lb/in.^{2} gage) are being developed to produce very large quantities of gas (2.5×10^8 ft^3/day of 1000 Btu/ft^3 gas or 7×10^6 m^3/day of 3.8×10^7 J/m^3 gas) to supplement the declining supplies of natural gas. Some of these processes use the partial combustion of oxygen as a direct source of heat, and others use air combustion in a variety of indirect modes.

Nonfossil sources. In view of the immediate shortage of natural gas and the longer-range (Af5D. 2030–2060) forecast that all fossil fuels will be consumed, there are several projects which may ensure a perpetual supply of fuel gas. These include conversion of waste materials to gas by pyrolysis or by anaerobic digestion. Large efforts have been initiated to convert solar energy to fuel gas. This would be accomplished by growing plants on land or sea and converting the harvest to gas in processes similar to those used on wastes. Finally, there has been a large effort to develop a "hydrogen economy." Hydrogen would be produced by thermochemical processes using heat from various sources such as atomic energy reactors.

Jack Huebler

Bibliography. M. Crow et al.,*Synthetic Fuel Technology Development in the United States: A Retrospective Assessment*, 1988; R. A. Myers, *Handbook of Synfuels Technology*, 1984; R. F. Probstein and R. E. Hicks, *Synthetic Fuels*, 1990.

Fuel injection

The pressurized delivery of a metered amount of fuel into the intake airflow or combustion chambers of an internal combustion engine. Metering of the fuel charge may be performed mechanically or electronically. In a diesel engine, the fuel is injected directly into the combustion chamber (direct injection) or into a smaller connected auxiliary chamber (indirect injection). In the spark-ignition engine, the fuel is injected into the air before it enters the combustion chamber by spraying the fuel into the airstream passing through the throttle body (throttle-body injection) or into the air flowing through the port to the intake valve (port injection). *See* COMBUSTION CHAMBER.

Diesel fuel injection. The diesel engine must be supplied with fuel from the injection nozzle at a pressure of 1500–5000 lb/in.^{2} (10–35 megapascals) for indirect-injection engines, and up to 15,000 lb/in.^{2} (100 MPa) or higher for direct-injection engines. The high pressure is necessary to deliver fuel against the highly compressed air in the engine cylinders at the end of the compression stroke, and to break up the fuel oil which has low volatility and is often viscous. Extremely accurate fuel metering is necessary, with the start of injection occurring within a precision of up to 1° of engine crankshaft angle. A timing device in the injection pump automatically advances the start of of fuel delivery as engine speed increases to optimize the start of combustion.

The intake air is not throttled in a diesel engine, with load and speed control accomplished solely by controlling the quantity of fuel injected. The mean effective pressure developed by combustion is controlled by the volumetric capacity of

the injection pump. To prevent an unloaded diesel engine from increasing in speed until it destroys itself, a governor is required to limit maximum engine speed. *See* DIESEL CYCLE; DIESEL ENGINE; GOVERNOR; INTERNAL COMBUSTION ENGINE; MEAN EFFECTIVE PRESSURE.

In a typical diesel fuel-injection system, a low-pressure supply pump delivers fuel from the fuel tank to the injection pump, which pressurizes, times, and meters the fuel charge and then delivers it through tubing to an injection nozzle that sprays the pressurized fuel into the combustion chamber. Either all fuel lines between the injection pump and nozzles are made of equal length, or a high-pressure manifold or fuel rail is used. The high fuel-injection pressure is produced by an engine-driven reciprocating piston pump fitted to very close tolerances, either in the injection-pump assembly or in a unit injector.

There are two types of injection pump: the in-line pump and the distributor pump. The in-line pump has two or more pumping elements that pressurize the fuel and may also time and meter it. Each pumping element, which is the combination of a pump plunger and its barrel, serves only one engine cylinder. As the plunger moves from the bottom to the top of its stroke, it closes off a spill port and begins to compress the fuel in the high-pressure chamber. The increasing pressure opens the delivery valve in the pump, allowing a pressure wave to travel at the speed of sound toward the injection nozzle. When the nozzle-opening pressure is reached, the nozzle needle opens against spring force, and then fuel is injected until the pump plunger opens the spill port. This relieves the pressure and the delivery valve closes.

In the distributor pump (**Fig. 1**), one pumping element is used to provide metered delivery to all cylinders. The plunger stroke delivers the fuel to a rotating distributing device, which then directs the fuel to the appropriate engine cylinder. The number of plunger strokes during one revolution of the injection-pump driveshaft corresponds to the number of engine cylinders.

Instead of an injection pump, some engines have a unit fuel injector for each cylinder. It receives fuel under supply-pump pressure and is then actuated by an engine or electronic mechanism to meter and inject the fuel charge into the combustion chamber at high pressure and at the proper time.

Gasoline fuel injection. On automotive spark-ignition engines, the carburetor has largely been replaced by a gasoline fuel-injection system with either mechanical or electronic control of fuel metering. Many of the systems are of the speed-density type, in which the mass airflow rate is calculated based on cylinder displacement and the measured intake-manifold absolute pressure (engine load), engine speed, intake-manifold air temperature, and theoretical volumetric efficiency. When the feedback signal from an exhaust-gas oxygen sensor is included, these systems allow the engine air-fuel ratio to be maintained near the stoichiometric ratio (14.7:1) during normal operating conditions. This minimizes exhaust emissions. *See* CARBURETOR.

In the typical gasoline fuel-injection system (**Fig. 2**), an electric fuel pump provides a specified fuel flow at the required system pressure to one or

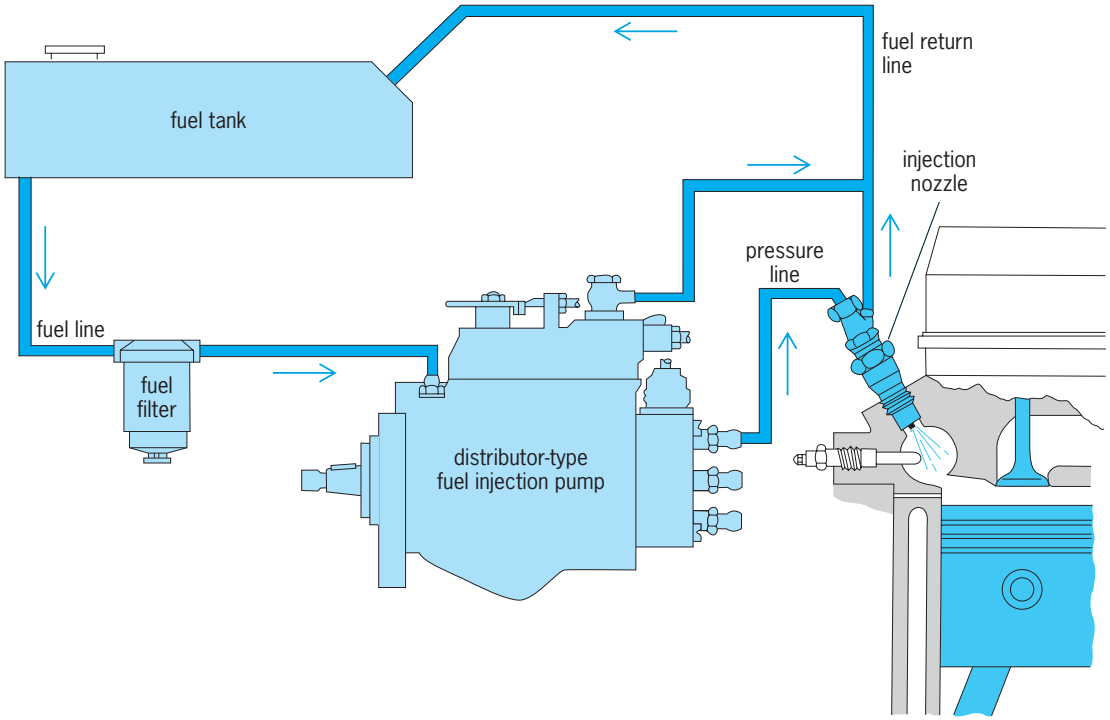

Fig. 1. Diesel-engine fuel-injection system using a distributor-type fuel-injection pump. (*Robert Bosch Corp.*)

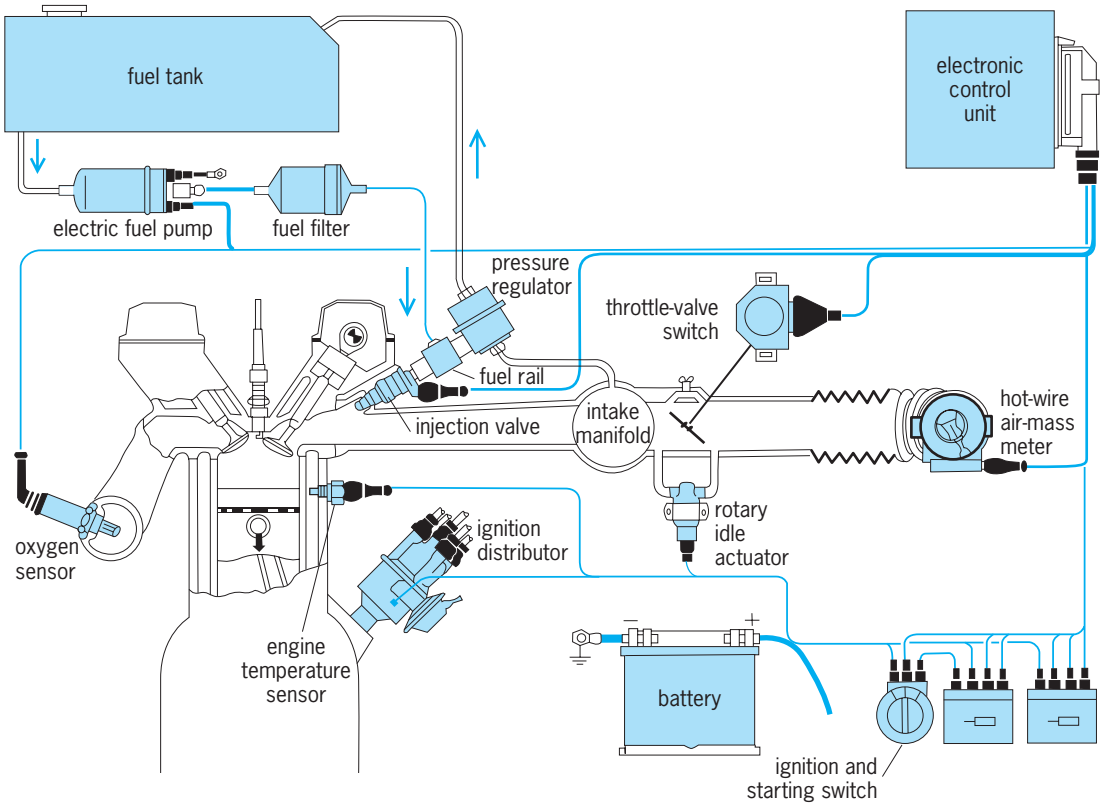

Fig. 2. Electronic port fuel-injection system for a spark-ignition engine. (*Robert Bosch Corp.*)

more fuel-injection valves, or fuel injectors. The gasoline fuel injector is an electromagnetic (solenoid-operated) or mechanical device used to direct delivery of or to meter pressurized fuel, or both. A fuel-pressure regulator maintains a controlled fuel pressure at each injector, or a controlled differential pressure across the injector.

The three types of gasoline fuel injection are throttle-body (central or single-point) injection, port (multipoint) injection (Fig. 2), and continuous injection. Throttle-body injection is an electronically controlled system in which one or two fuel injectors are located in the throttle body, usually above the throttle valves. The quantity of fuel delivered is determined by an electronic control unit whose output signal controls injector open time. This is based on information received by the electronic control unit from several sensors that monitor the operating conditions of the engine.

In electronic port injection systems, the fuel is delivered to a fuel manifold or rail that distributes it to an individual fuel injector located in the intake port for each cylinder. The injectors may be triggered either in groups or sequentially. A continuous injection system may be controlled either electronically or mechanically while fuel flows continuously from the injectors independent of cylinder events. Fuel flow rate is controlled by a variable orifice or fuel pressure. *See* FUEL PUMP; FUEL SYSTEM. Donald L. Anglin

Bibliography. *Bosch Automotive Electric/Electronic Systems*, 1986; *Bosch Automotive Handbook*, 1986; W. H. Crouse and D. L. Anglin, *Automotive Fuel, Lubricating, and Cooling Systems*, 1981.

Fuel oil

Any of the petroleum products which are less volatile than gasoline and are burned in furnaces, boilers, or other types of heaters. The two primary classes of fuel oils are distillate and residual. Distillate fuel oils are composed entirely of material which has been vaporized in a refinery distillation tower. Consequently, they are clean, free of sediment, relatively low in viscosity, and free of inorganic ash. Residual fuel oils contain fractions which cannot be vaporized by heating. These fractions are black and viscous and include any inorganic ash components which are in the crude. In some cases, whole crude is used as a residual fuel.

Uses. Distillate fuel oils are used primarily in applications where ease of handling and cleanliness of combustion are more important than fuel price. The most important use is for home heating. Distillate fuel oils are also used in certain industrial applications where low sulfur or freedom from ash is important. Certain types of ceramics manufacture are examples. Increasing amounts of distillate fuel oils have been burned in gas turbines used for electricity generation. *See* ELECTRIC POWER GENERATION; GAS TURBINE.

Residual fuel oils are used where fuel cost is an important enough economic factor to justify additional

investment to overcome the handling problems they pose. They are particularly attractive where large volumes of fuel are used, as in electric power generation, industrial steam generation, process heating, and steamship operation.

Fuel oils are frequently made for specific uses, and these may be distillates, residuals, or mixtures of the two. Therefore, terms such as domestic fuel oils, diesel fuel oils, and heavy fuel oils are being used in classification schemes.

Combustion. Fuel oils are burned efficiently by atomizing them into fine droplets, about 50–100 micrometers in diameter, and injecting the droplet spray into a combustion chamber with a stream of combustion air. Small home heating units use a pressure atomizing nozzle. Larger units use pressure, steam, or air atomizing nozzles, or else a spinning cup atomizer. Older units employ natural draft to induce combustion air, but new units usually use forced draft, which gives better control of air-to-fuel ratio and increased turbulence in the combustion chamber. This permits more efficient operation and clean, smoke-free combustion. *See* OIL FURNACE.

Tests and specifications. Fuel oils are blended to meet certain tests and specifications which ensure that they can be handled safely and easily, that they will burn properly, and that they will meet air-pollution regulations. These tests include (1) flash point, which determines that the fuel can be stored and handled without danger of explosion; (2) pour point, which determines that it will not solidify under normal handling conditions; (3) carbon residue, which determines that it will not coke during handling or form undue amounts of carbon particulates during combustion; (4) sediment, which determines that it will not clog pumps and nozzles; (5) sulfur content, which determines how much sulfur dioxide will be emitted during combustion; (6) ash content; and (7) viscosity or resistance to flow.

Viscosity. Viscosity is an important factor in determining the grade of a fuel oil. Distillate fuels and light grades of residuals have a low viscosity so that they can be handled and atomized without heating. Heavier grades have high viscosity and must be heated in order to be pumped and atomized. The grade classification of the fuel oil determines what viscosity range it is blended to, and therefore the temperature level required in the preheat system.

Ash content. Fuel ash content may be important in residual fuels, even though ash content is negligible compared with coal, because certain ash components may cause slagging and corrosion problems. Vanadium and sodium are frequently present in oil ash. During combustion, they are converted to materials which tend to accumulate on boiler and superheater tubes. Under certain conditions, they may cause corrosion of these tubes.

Sulfur content. Sulfur content of fuel oil is carefully controlled in order to meet air-pollution regulations. This has required a drastic reduction in sulfur content of residual fuel oils in many areas. To achieve the reduction, extensive processing and changes in blending procedure were required, which caused major increases in the price of fuel oil.

A detailed list of specifications for fuel oils is given in the *Annual Book of ASTM Standards*, published by the American Society for Testing and Materials, Philadelphia. *See* PETROLEUM PRODUCTS.

C. W. Siegmund

Fuel pump

A mechanical or electrical pump for drawing fuel from a storage tank and forcing it to an engine or furnace. The type of pump chosen for a given fuel depends to a great extent on the volatility of the liquid to be pumped. In a gasoline engine the fuel is highly volatile at ambient temperature. Therefore, the fuel line is completely sealed from the tank to the carburetor or fuel-injection system to prevent escape of fuel and to enable the pump to purge the line of vapor in the event of vapor lock—a condition in which the fuel vaporizes owing to abnormally high ambient temperature. *See* CARBURETOR; FUEL INJECTION; FUEL SYSTEM.

Most carbureted gasoline engines use a spring-loaded diaphragm-type mechanical pump which is normally actuated by a rocker arm or pushrod that rides on an eccentric on the engine camshaft (**illus.** *a*). Electric motor–driven and solenoid-operated diaphragm pumps and plunger pumps are also available that can be mounted near the main fuel tank to minimize vapor lock in the fuel lines. *See* VAPOR LOCK.

Many gasoline-engine vehicles have a submersible electric fuel pump, which serves as the main supply pump, located in the fuel tank. In some fuel-injection systems, the in-tank pump is used as the supply pump for a high-pressure fuel-injection pump. The in-tank pump may be of the gear, plunger, sliding-vane, or impeller type (illus. *b*).

Diesel engines normally use a gear, plunger, or vane-type pump to supply fuel to the injection pump. In the diesel engine, where fuel is injected at high pressure through an injection nozzle into the highly compressed air in the combustion chamber, a plunger or piston serves as its own inlet valve and as the compression member of the injection pump. When the required high pressure is reached in the injection nozzle, a spring-biased needle valve opens and fuel sprays into the combustion chamber. In an oil-fired furnace, although nozzle pressures need not be so high as in diesel engines, a piston pump is also used to provide positive shutoff of the fuel line when the pump stops. *See* DIESEL ENGINE.

In a liquid-fueled rocket with pump feed, both fuel pump and oxidizer pump are typically centrifugal or turbine pumps, usually driven by an accessory turbine engine that burns a small quantity of pumped fuel and oxidizer. The pumps must withstand the corrosive liquids and have close clearances, preferably with no leakage. Flow rates may be limited by cavitation at the pump inlet, especially because the fuel or cryogenic fluid, such as liquid oxygen, may be

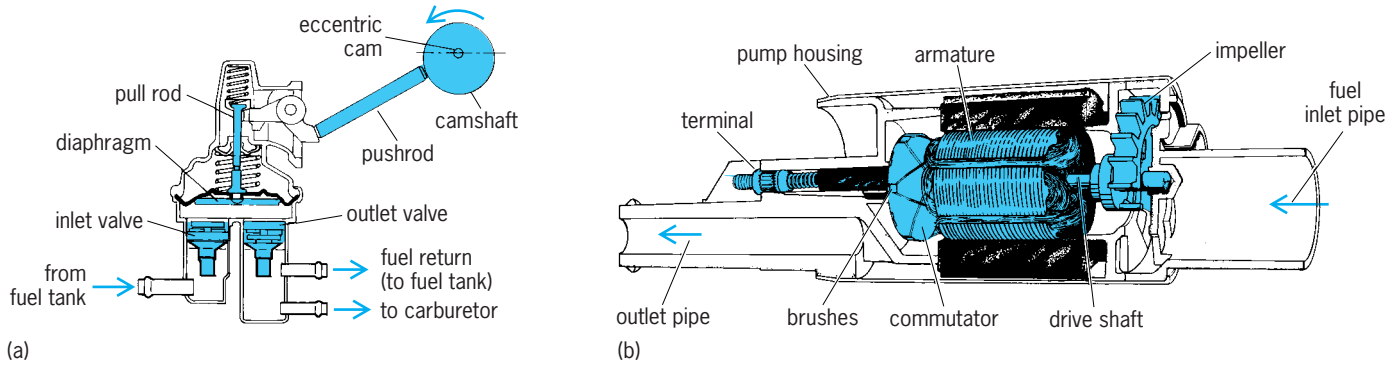

Fuel pumps for gasoline engines. (*a*) In-line mechanical diaphragm pump (*Chrysler Corp*). (*b*) In-tank electric impeller pump (*AC-Delco Division of General Motors Corp.*).

near its vapor pressure because of its high velocity. *See* CAVITATION.

Fuel pumps must operate before the engine or furnace that they serve can function. Consequently, an auxiliary or starter motor is usually an essential adjunct to a fuel pump. Fuel pumps should be easily accessible and mounted so that the outlet line can be disconnected and a pressure gage installed to check operating pressure. In diaphragm-type pumps, the diaphragm material must be capable of withstanding degradation by aromatic-type fuels. Donald L. Anglin

Bibliography. *Bosch Automotive Electrics/Electronics*, 1988; *Bosch Automotive Handbook*, 1986; W. H. Crouse and D. L. Anglin, *Automotive Fuel, Lubricating, and Cooling Systems*, 1981.

Fuel system

The system that stores fuel for present use and delivers it as needed to an engine; includes the fuel tank, fuel lines, pump, filter, vapor return lines, carburetor or injection components, and all fuel system vents and evaporative emission control systems or devices that provide fuel supply and fuel metering functions. Some early vehicles and other engines had a gravity-feed fuel system, in which fuel flowed to the engine from a tank located above it. Automotive and most other engines have a pressurized fuel system with a pump that draws or pushes fuel from the tank to the engine. *See* CARBURETOR; FUEL INJECTION; FUEL PUMP.

Automobile. The commonly used components for automobile and stationary gasoline engines are fuel tank, fuel gage, filter, electric or mechanical fuel pump, and carburetor or fuel-injection system (**Fig. 1**). In the past, fuel metering on automotive engines was usually performed by a carburetor. However, this device has been largely replaced by fuel injection into the intake manifold or ports, which increases fuel economy and efficiency while lowering exhaust gas emissions. Various types of fuel management systems are used on automotive engines, including electronically controlled feedback carburetors, mechanical continuous fuel injection, and sequential electronic fuel injection. *See* AUTOMOBILE.

Most carbureted engines use a diaphragm-type mechanical fuel pump that is operated by the rotation of an eccentric on the engine camshaft. Many gasoline fuel-injected engines use an impeller-type electric fuel pump located in the fuel tank (Fig. 1). This helps prevent problems with vapor lock, particularly during short engine stops on hot days, which results when the volatile fuel boils in the supply line and pumping chamber. *See* VAPOR LOCK.

Electric fuel pumps may have plunger-, diaphragm-, or impeller-type pumping elements, and are designed so that their speed of operation and current consumption depend upon the rate of fuel use. Safety devices prevent fuel delivery if the vehicle rolls over of if the engine stops running. *See* AUTOMOTIVE ENGINE. Donald L. Anglin

Aircraft. The presence of multiple engines and multiple tanks complicates the aircraft fuel system. Also, the reduction of pressure at altitude necessitates the regular use of boost pumps, submerged in the fuel tanks, which are usually of the centrifugal type and electrically driven. These supplement the engine-driven fuel supply pumps, which are usually of the gear or eccentric vane type.

Components of a typical aircraft fuel system include one main and two auxiliary tanks with their gages; booster, transfer, and engine-driven pumps; various selector valves; and a fuel jettisoning or defuel valve and connection, which is typical also of what would be needed for either single-point ground

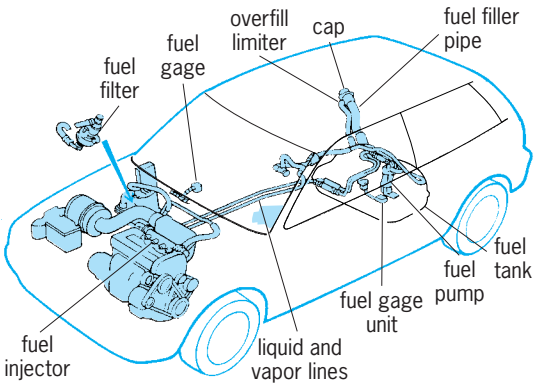

Fig. 1. Components of an automobile fuel system. (*Chrysler Corp.*)

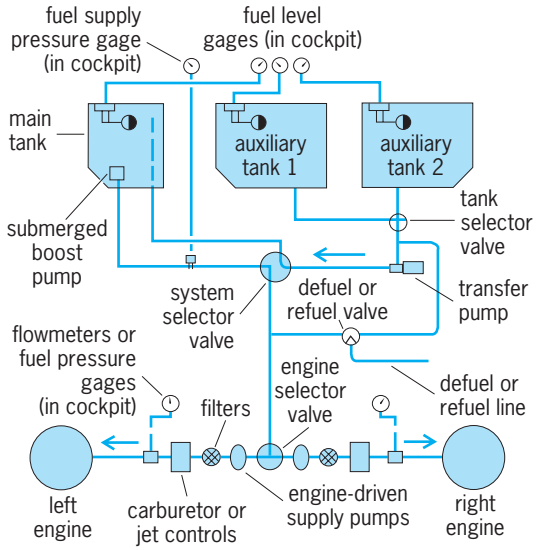

Fig. 2. Diagram of a typical aircraft fuel system.

or flight refueling (**Fig. 2**). The arrangement is usually such that all the fuel supply will pass to the engines by way of the main tank, which is refilled as necessary from the auxiliary tanks. In case of emergency, the system selector valve may connect the auxilary tanks to the engines directly. Tank vents, not shown, are arranged so that overflow will go safely overboard.

Because of the large quantities of fuel used, aircraft fuel systems are often contaminated with dirt, metal chips, and lint; adequate filtration is therefore essential to reliable service. *See* AIRCRAFT ENGINE.

J. A. Bolt

Fugacity

A function introduced by G. N. Lewis to facilitate the application of thermodynamics to real systems. Thus, when fugacities are substituted for partial pressures in the mass action equilibrium constant expression, which applies strictly only to the ideal case, a true equilibrium constant results for real systems as well.

The fugacity f_i of a constituent i of a thermodynamic system is defined by Eq. (1) [where μ_i is the

$$\mu_i = \mu_i^* + RT \ln f_i \tag{1}$$

chemical potential and μ_i^* is a function of temperature only], in combination with the requirement that the fugacity approach the partial pressure as the total pressure of the gas phase approaches zero. At a given temperature, this is possible only for a particular value for μ_i^*, which may be shown to correspond to the chemical potential the constituent would have as the pure gas in the ideal gas state at 1 atm pressure. This definition makes the fugacity identical to the partial pressure in the ideal gas case. For real gases, the ratio of fugacity to partial pressure, called the fugacity coefficient, will be close to unity for moderate temperatures and pressures. At low temperatures and appropriate pressures, it may be as small as 0.2 or less, whereas at high pressures at any temperature it can become very large.

The fugacity concept is not restricted to gaseous systems, however. Because of its relation to the chemical potential, the basic thermodynamic criterion of equilibrium requires that the fugacity of a constituent have the same value at equilibrium for every phase in which it is present. This permits the indirect determination of the fugacity for a condensed phase through the calculation of the value for the equilibrium vapor phase, for which the fugacity may be computed routinely if the dependence of its volume on temperature, pressure, and composition is known. Thus, results are readily obtained for a pure gas, but because of the more extensive data required, accurate calculations have been made for very few mixtures.

For an ideal solution, the fugacity is given by the mole fraction of the constituent multiplied by the fugacity of the pure constituent at the temperature and pressure of the solution. For liquid solutions, this is the thermodynamic counterpart of Raoult's law, but the relation applies also to ideal gaseous solutions and can serve for the prediction of the properties of real gas mixtures. Where no equation of state data are available for pure gases, their fugacities can be estimated by means of the generalized fugacity coefficient chart. At sufficiently high dilution in liquid solution, the fugacity of a nondissociating solute will become proportional to its concentration, the proportionality constant depending on the concentration scale used; this is the thermodynamic statement of Henry's law.

The dependence of the fugacity on temperature at constant pressure and composition is given by Eq. (2). Here H_i^* is the enthalpy per mole for the

$$\left(\frac{\partial \ln f_i}{\partial T}\right)_{P,\text{comp}} = \frac{H_i^* - \overline{H}_i}{RT^2} \tag{2}$$

constituent in the gas phase at very low pressure, and $\overline{H}_i$ its contribution per mole to the enthalpy of the system for the state under consideration. *See* ACTIVITY (THERMODYNAMICS); CHEMICAL EQUILIBRIUM; CHEMICAL THERMODYNAMICS; GAS; PHASE EQUILIBRIUM; SOLUTION.

Paul Bender

Bibliography. K. S. Pitzer and L. Brewer, *Thermodynamics*, rev. ed., 1961; C. E. Reid, *Chemical Thermodynamics*, 1990; P. A. Rock, *Chemical Thermodynamics*, 1983.

Fullerene

A hollow, pure carbon molecule in which the atoms lie at the vertices of a polyhedron with 12 pentagonal faces and any number (other than one) of hexagonal faces. The fullerenes were discovered as a consequence of astrophysically motivated chemical physics experiments that were interpreted by using geodesic architectural concepts. Fullerene chemistry, a field that appears to hold much promise for

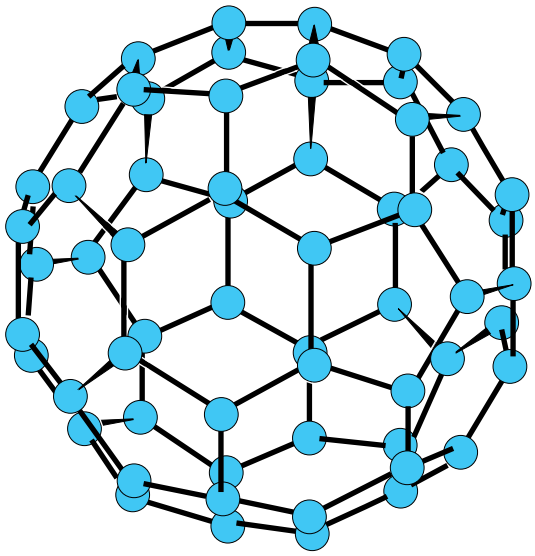

Fig. 1. Structure of fullerene-60 (C_{60}).

materials development and other applied areas, was born from pure fundamental science.

Buckminsterfullerene (C_{60} or fullerene-60) (**Fig. 1**) is the archetypal member of the fullerenes. Other stable members of the fullerene family have similar structures (**Fig. 2**). The fullerenes can be considered, after graphite and diamond, to be the third well-defined allotrope of carbon. Macroscopic amounts of various fullerenes were first isolated in 1990, and since that time it has been discovered that members of this class of spheroidal organic molecules have numerous novel physical and chemical properties. The fullerenes promise to have synthetic, pharmaceutical, and industrial applications. Derivatives have been found to exhibit fascinating electrical and magnetic behavior, in particular superconductivity and ferromagnetism. *See* CARBON; DIAMOND.

Structures. In the fullerene molecule an even number of carbon atoms are arrayed over the surface of a closed hollow cage. Each atom is trigonally linked to its three near neighbors by bonds that delineate a polyhedral network, consisting of 12 pentagons and n hexagons. (Such structures conform to Euler's theorem for polyhedrons in that n may be any number other than one, including zero.) All 60 atoms in fullerene-60 are equivalent and lie on the surface of a sphere distributed with the symmetry of a truncated icosahedron. The 12 pentagons are isolated and interspersed symmetrically among 20 linked hexagons; that is, the symmetry is that of a modern soccerball. The molecule was named after R. Buckminster Fuller, the inventor of geodesic domes, which conform to the same underlying structural formula. Three of the four valence electrons of each carbon atom are involved in the sp^2 sigma-bonding skeleton, and the fourth p electron is one of the 60 involved in a pi-delocalized molecular-orbital electron sea that covers the outside (exo) and inside (endo) surface of the molecule. The resulting cloud of pi-electron density is similar to that which covers the surface of graphite; indeed, the molecule can be considered a round form of graphite. The smallest member of the family, fullerene-20, is dodecahedrene, whose fully hydrogenated analog ($C_{20}H_{20}$) has been created by rational synthesis (production of a compound by using a sequence of chemical reaction steps strategically chosen for the purpose). Several members of the family have been isolated, in particular C_{70}, C_{76}, C_{78}, and C_{84}. Giant fullerenes with at least 600 atoms also seem to form, and appear to have interesting polyhedral shapes. Some species have highly symmetric icosahedral structures, such as C_{240} and C_{540} (**Fig. 3**). *See* ELECTRON CONFIGURATION; GRAPHITE.

General properties. In benzene solution, fullerene-60 is magenta and fullerene-70 red. Fullerene-60 forms translucent magenta face-centered cubic (fcc) crystals that sublime. The ionization energy is 7.61 eV and the electron affinity is 2.6–2.8 eV. The strongest absorption bands lie at 213, 257, and 329 nanometers. Studies with nuclear magnetic resonance spectroscopy yield a chemical shift of 142.7 ppm; this result is commensurate with an aromatic system.

Solid C_{60} exhibits interesting dynamic behavior in that at room temperature the individual round molecules in the face-centered cubic (fcc) crystals are rotating isotropically (that is, freely) at around 10^8 Hz. At around 260 K (8.3°F) there is a phase transition to a simple cubic (sc) lattice accompanied by an abrupt lattice contraction. Rotation is no longer free, and the individual molecules make rotational jumps between two favored (relative) orientational configurations—in the lower-energy one a double bond lies over a pentagon, and in the other it lies over a hexagon. At 90 K (−300°F) the individual molecules stop rotating altogether, freezing into an orientationally disordered crystal involving a mix of the two configurations. *See* CRYSTAL STRUCTURE.

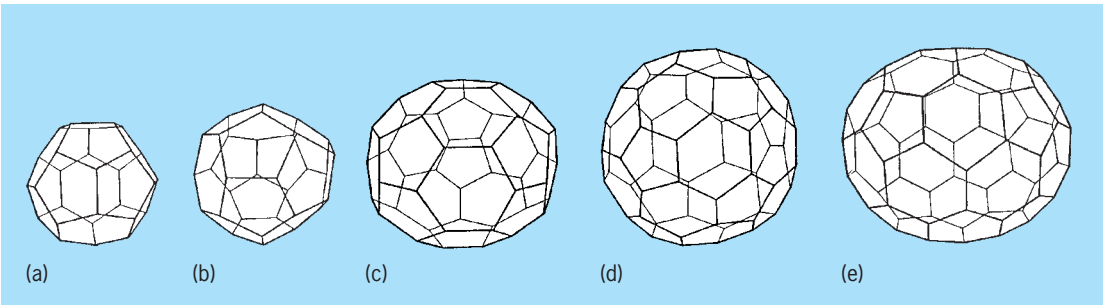

Fig. 2. Some of the more stable members of the fullerene family. (*a*) C_{28}. (*b*) C_{32}. (*c*) C_{50}. (*d*) C_{60}. (*e*) C_{70}.

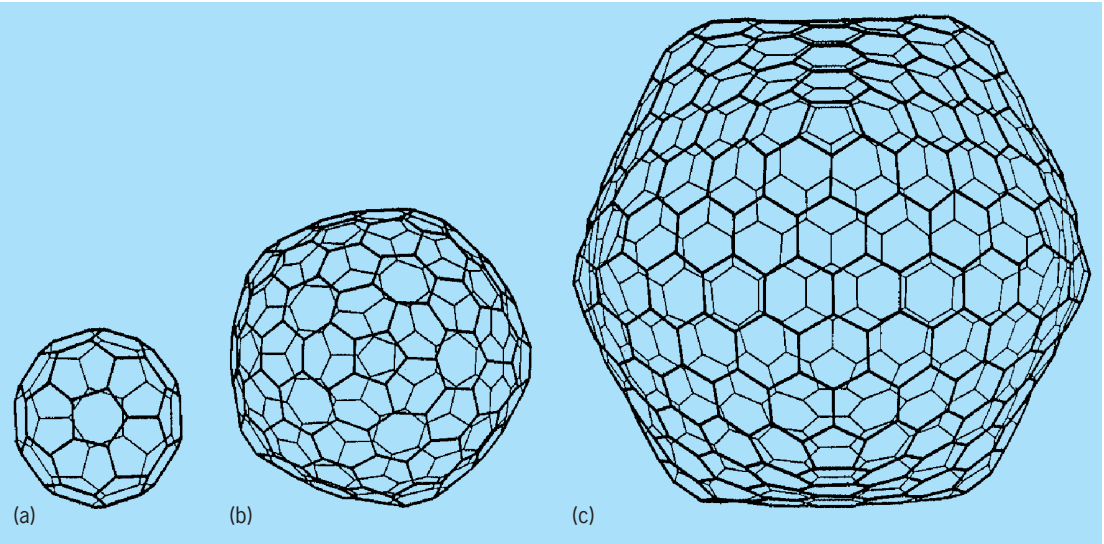

Fig. 3. Comparison of the structure of fullerene-60 with that of two giant fullerenes. (*a*) C_{60}. (*b*) C_{240}. (*c*) C_{540}.

Chemistry. Fullerene-60 behaves as a soft electrophile, a molecule that readily accepts electrons during a primary reaction step. It can accept three electrons readily and perhaps even more. The molecule can be multiply hydrogenated, methylated, ammonated, and fluorinated. It forms exohedral complexes in which an atom (or group) is attached to the outside of the cage, as well as endohedral complexes in which an atom is trapped inside the cage. Evidence for oxygen addition compounds, such as $C_{60}O$ and $C_{70}O$, has also been reported. Fullerene-60 appears to catalyze the formation of singlet oxygen in the presence of light. This ability may be related to the observation that under certain conditions solutions of fullerene-60 decompose in the presence of oxygen and light.

Perhaps the most significant chemical result to have been revealed since macroscopic amounts of fullerenes became available in 1990 is that, for synthetic purposes, the C_{60} molecule behaves as though it has only a single resonance form—one in which the 30 double bonds are localized in the bonds that interconnect the pentagons. This is a key factor, as addition to these double bonds is the most important reaction as far as the application of C_{60} in synthesis is concerned.

C_{60} (and other fullerenes) behaves like an electrophilic polyene, and it is susceptible to relatively facile addition by nucleophiles. The pattern of reactivity of the surface network and the way in which specific fullerene derivatives may be produced are gradually unfolding. Addition derivatives with halogens, hydroxyl, phenyl, cyclopentadienel, sugars, and other moieties have been produced:$C_{60}Cl_6$, $C_{60}Br_6$ (**Fig. 4**), $C_{60}Br_8$, $C_{60}Br_{24}$, $C_{60}F_{36}$, $C_{60}Ph_6$, $C_{60}Ph_8$, $C_{60}HPh_5$, $C_{60}(CpH)_6$, and so on. Free-radical additions to produce stable $C_{60}R_n$ radicals ($R = CH_3$ and $n = 1$–5, and so on) are particularly facile. Transition-metal complexes such as $C_{60}O_2OsO_2$ (4-*t*-butylpyridine)$_2$, $C_{60}Ir(CO)Cl(PPh_3)_2$, and $C_{60}[Pt(PEt_3)_2]_6$ have been isolated. Host-guest crystal complexes form with cyclodextrin, ferrocene, hydroquinone, and sulfur.

Some progress has been made on the production of endohedral (inside cage) complexes; in particular $La@C_{82}$ and $U@C_{28}$ appear to be stable. It is customary with endohedral complexes in which an atom or molecule is trapped inside a fullerene cage to use nomenclature such as $M@C_{60}$, where @ signifies that M is inside the cage—in this case C_{60}. Helium (He) atom may, under very high pressure, be forced inside the C_{60} cage to form $He@C_{60}$. The He atom can be expelled on heating. Helium can also be injected into the cage by high-velocity C_{60}/He beam collision experiments.

The possible hazards associated with working with fullerene-60 have not been determined. However, common sense and the ability to catalyze

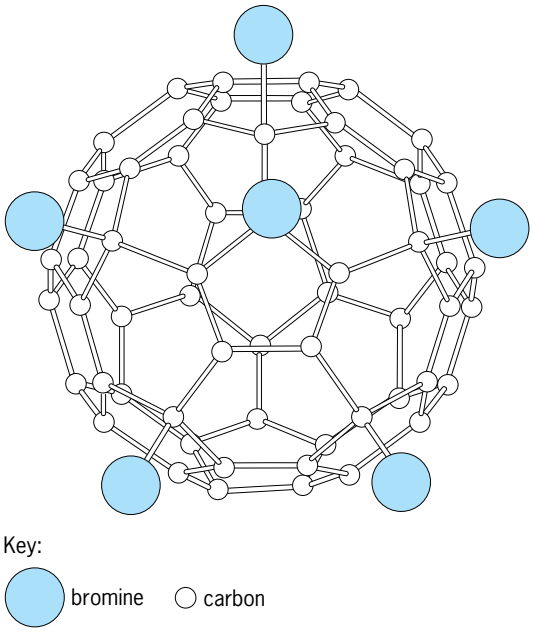

Fig. 4. Structure of the molecule $C_{60}Br_6$.

formation of singlet molecules suggest that some care should be exercised when working with it until more is known about possible toxicity.

Superconductivity. On exposure of C_{60} to certain alkali and alkaline earth metals, exohedrally doped crystalline materials are produced that exhibit superconductivity at relatively high temperatures (10 to 33 K or −440 to −400°F). The first to be discovered was the ionic material $K_3{}^+C_{60}{}^{3-}$, which is superconducting at a critical temperature (T_c) of 19 K (−390°F). The highest T_c so far is 33 K (−400°F), exhibited by the mixed intercalate $CsRb_2C_{60}$. Previously, only metallic and ceramic materials exhibited superconductivity at temperatures much greater than a few kelvins. This discovery has opened up the field of superconductivity to a different range of substances—in this case, molecular superconductors. The C_{60} molecule has a triply degenerate lowest unoccupied molecular orbital (LUMO), which in the superconducting materials is half filled, containing three electrons. Other ionic phases, such as M_nC_{60} ($n = 1, 2, 4, 6$, where M is the intercalated metal atom), exist but are not superconducting—they appear to be metallic or semiconductor/insulators. *See* MOLECULAR ORBITAL THEORY; SUPERCONDUCTIVITY.

Nanoparticles and nanotubes. The discovery that graphene networks (single sheets of hexagonally interconnected carbon atoms) can close readily has revolutionized general understanding of certain types of graphitic materials. It has revealed that when isolated aggregates with as many as 10 million carbon atoms anneal, the most stable structure formed is a spheroidal onionlike graphite particle consisting of sets of concentric graphene shells. The discovery of C_{60} itself, together with the nanoparticle studies, clearly shows that dynamic edge effects operating during the graphite formation process govern the resulting intrinsic structures and, in the case of isolated aggregates, result in closed spheroidal structures.

Carbon microparticles can spontaneously rearrange at high temperatures to form onionlike structures in which the concentric shells are fullerene or giant fullerenes. This phenomenon reveals the dynamics of carbon "melting."

Most importantly, carbon nanofibers consisting of concentric graphene tubes (**Fig. 5**) can form. These structures are essentially elongated giant fullerenes and apparently form quite readily. They will be significant in the production of carbon-fiber composite materials.

Applications. The properties of fullerene materials that have been determined suggest that there is likely to be a wide range of areas in which the fullerenes or their derivatives will have uses. The facility for acceptance and release of electrons suggests a possible role as a charge carrier in batteries. Thin films of C_{60} are transparent at low light intensities and opaque at high intensities, a behavior ideal for optical limiter applications. The fast response means that a thin layer of C_{60} molecules may be used to protect the eye from high transient light pulses. The properties of graphite suggest that lubricative as well as tensile and other mechanical properties of the fullerenes are worthy of investigation. Liquid solutions exhibit excellent properties of optical harmonic generation. The high temperature at which superconducting behavior is observed suggests possible applications in microelectronics devices, as does the detection of ferromagnetism in other fullerene derivatives.

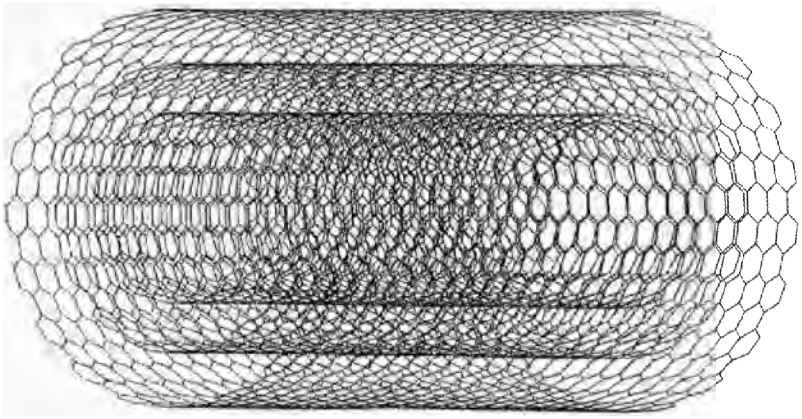

Fig. 5. A section of a carbon nanotube showing the concentric graphene tubes composing it.

Fullerene formation and implications. Molecular beam experiments indicate that fullerene cages can grow by accretion of carbon, and combined mass spectrometric–synthetic studies have made some progress in developing understanding of the basic fullerene creation process. This process appears to involve the gas-phase aggregation of cumulene/polyyne chains and rings, which then undergo intraaggregate ring cyclizations to produce the closed polycyclic 5/6-ring carbon cage networks.

Perhaps the most important aspect of the fullerene discovery is that the molecule forms spontaneously. This fact has important implications for understanding the way in which extended carbon materials form, and in particular the mechanism of graphite growth and the synthesis of large polycyclic aromatic molecules. It has become clear that as far as pure carbon aggregates of around 60–1000 atoms are concerned, the most stable species are closed-cage fullerenes. A flat, pure graphite carbon sheet has dangling bonds at the edge and these can be satisfied by closure, so that the energy of the system is reduced. In the case of bulk graphite, which consists of stacks of hexagonally arrayed carbon sheets, the interlayer interactions and size effects appear to stabilize the flat sheets. Also, it may be possible to isolate some of the more stable smaller fullerenes, those with fewer than 60 carbon atoms, such as fullerene-50.

Astrophysical implications. Fullerene-60 was discovered as a direct result of physiochemical investigations that simulated processes occurring in stars and in space. Consequently, the likelihood that fullerenes, in particular fullerene-60, and analogs are present in space is a fascinating conjecture. Particularly likely are the exohedral protonated form, $C_{60}H^+$, and exohedral metallated species such as $C_{60}M^+$, where M is an abundant interstellar atom such as sodium, potassium, calcium, or oxygen.

Jonathan Hare; Harold W. Kroto

Bibliography. R. F. Curl and R. E. Smalley, Probing C_{60}, *Science*, 242:1017–1022, 1988; W. Kraetschmer et al., Solid C_{60}: A new form of carbon, *Nature*, 347:354–358, 1990; H. W. Kroto, Space, stars, C_{60}, and soot, *Science*, 242:1139–1145, 1988; H. W. Kroto et al., C_{60}: Buckminsterfullerene, *Nature*, 318:162–163, 1985; H. W. Kroto, W. Allaf, and S. C. Balm, C_{60}: Buckminsterfullerene, *Chem. Rev.*, 91:1213–1235, 1991; K. Prassides and H. W. Kroto, Fullerene physics, *Phys. World*, 5:44–49, 1992; R. Tayor and D. R. M. Walton, The chemistry of fullerenes, *Nature*, 363:685–693, 1993.

Fuller's earth

Any natural earthy material that decolorizes mineral and vegetable oils and has high sorbent capacity for water and oil. The term fuller's earth has no genetic or mineralogic significance. However, the most common earthy materials classed as fuller's earth are calcium montmorillonites and palygorskites (attapulgites) and sepiolites. The term originated in England, where in ancient times raw wool was cleaned by kneading it in water with clay materials that adsorbed dirt and lanolin. The process was known as fulling, and the clay or earth became known as fuller's earth. *See* ADSORPTION; MONTMORILLONITE; SEPIOLITE.

Commercial sources. Several clay deposits in the world are mined and processed for their absorbent, adsorbent, and decolorizing or bleaching properties. Some clays have a high natural decolorizing ability; however, in most instances a clay, normally a calcium montmorillonite, is acid-activated to enhance its bleaching or decolorizing properties. Sulfuric acid is commonly used, and in the treatment process sodium, calcium, magnesium, and iron that occupy the cation exchange sites on the clay surface are removed by the acid and replaced by hydrogen. Also, some aluminum, iron, or magnesium is removed from the mineral structure, increasing the negative charge on the clay surface. These highly charged surfaces covered with hydrogen ions selectively absorb the color bodies and other impurities in the oil. *See* CLAY; CLAY MINERALS.

The largest applications for fuller's earth are as sorbents, and by far the biggest market is pet-litter production. Other large sorbent applications are as carriers for insecticides, pesticides, and fertilizers used in agriculture and as absorbers of oil and water spills on the floors of machine shops, factories, service stations, and other manufacturing plants for safety purposes. All of these sorbent applications use granular particles that are usually of coarse sand size. The particles are dried at high heat to completely drive off absorbed and interlayer water from the clay material and to harden the particles so they do not slake readily in the presence of water or oil.

Occurrence. Fuller's earth is mined in many areas of the world. The most important deposits are located in Georgia, Florida, Mississippi, Tennessee, Illinois, and Missouri in the United States, and in England, Germany, Spain, Russia, Senegal, China, and India.

Haydn H. Murray

Bibliography. H. H. Murray, *Applied Clay Minerology Today and Tomorrow*, 1999; R. H. S. Robertson, *Fuller's Earth: A History*, 1986.

Fulminate

A compound containing the —ONC group and derived from fulminic acid, HONC. Fulminates are isomeric with cyanates; that is, cyanates have the same atoms but in different arrangement, —OCN.

The fulminates are commercially important because of the use of mercury fulminate, $Hg(ONC)_2$, in priming compositions and as an initial detonating agent. Mercury fulminate is very sensitive to impact, friction, and heat, but lead azide is replacing mercury fulminate as a detonating agent. *See* AZIDE; CYANATE.

E. Eugene Weaver

Fumigant

A pesticidal chemical or chemical formulation that functions in a gaseous state. Chemical formulations are designed to increase toxicity, reduce flammability, give off warning odors, and provide for sorption at different rates.

Types. Physical types of fumigants include gases, liquids, and solids. Gases, such as methyl bromide and hydrogen cyanide, exist in a gaseous state at normal room temperatures and pressures and must be stored in cylinders under pressure until used. Liquids, such as the 8:20 mixture of carbon tetrachloride and carbon disulfide, volatilize upon exposure to air. Solids, such as aluminum phosphide and calcium cyanide, produce hydrogen phosphide and hydrogen cyanide upon exposure to moisture in air. Most fumigants except hydrogen cyanide are heavier than air.

There are several chemical types of fumigants. These include: halogenated hydrocarbons, such as carbon tetrachloride and ethylene dibromide; sulfur-containing compounds, such as carbon disulfide and sulfur dioxide; cyanides, such as hydrogen cyanide and calcium cyanide; and others, such as phosphine and ethylene dioxide. Although many chemicals are registered as fumigants, extensive use is limited to a few.

Properties. In additon to volatility and toxicity to insects, essential properties of fumigants include diffusion and sorption, all of which affect their use. Heavy fumigant gases diffuse more slowly than lighter gases and do not penetrate as well. All gases diffuse more rapidly at higher temperatures. Heavier-than-air gases diffuse by gravity. The gas molecules may adsorb to the surface of commodities such as grain or they may be absorbed or chemisorbed into the commodities. The extent of sorption is determined by weight and volatility of the gas, the temperature and moisture of the area or commodity, and the presence of extraneous materials such as dust

and chaff. Desorption follows at variable rates. An effective fumigant application occurs when toxic concentrations are maintained long enough to be lethal to the pests.

Uses. Fumigants are used in space fumigations to disinfest food-processing plants, warehouses, grain elevators, boxcars, shipholds, stores, and households, and in spot fumigations within those structures. They are used in atmospheric vaults and vacuum chambers and are applied extensively to stacked bags of grain or stored foods under polyethylene sheets, to areas of land to destroy weeds, soil-infesting insects, and nematodes.

Disadvantages. Fumigants have certain inherent disadvantages: Reinfestations may develop as the gas dissipates; off-flavors and odors may result; some are flammable and explosive; some leave toxic residues; certain liquid fumigants are organic solvents; and others are corrosive to metals. All fumigants are toxic to humans; danger exists before, during, and after application. However, fumigants can be applied safely and effectively by trained applicators equipped with safety equipment. For safety reasons, applicators should never work alone.

Action. Fumigant molecules enter the insect's respiratory tubes or penetrate the cuticle and act as respiratory, nerve, or protoplasmic poisons. Some insects close their respiratory tubes after contact with the gas, thus requiring prolonged toxic concentrations. Considerable variability in toxicity exists depending on the fumigant employed and the pest involved. *See* INSECTICIDE; PESTICIDE.

Donald A. Wilbur, Sr.

Bibliography. Institution of Chemical Engineers (England), *A User Guide to Dust and Fume Control*, 1981; H. A. U. Monro, *Manual of Fumigation for Insect Control*, 2d ed., rev. 1969; National Fire Protection Association, *Standards on Fumigation*, 1973.

Funariales

An order of the true mosses (subclass Bryidae). The Funariales, often annuals or biennials and sometimes ephemeral, are for the most part characterized by a uniformity of gametophytic structure in contrast to a variability of sporophytic characters (see **illus.**). Reduction in sporophytic characters is often associated with disturbed habitats and a shortened life cycle.

The plants are generally small or minute, growing scattered or tufted mostly on soil. The stems are erect, and the leaves, inserted in numerous rows, are usually broadly oblong-lingulate and rounded-to-acuminate at the apex, with a single costa. The leaf cells, are lax, thin-walled, and smooth, or less commonly papillose because of projecting upper ends. The sporophytes are terminal (except that in the Funariaceae the sporophyte is formed at the apex of a leading shoot that forms below a basal perigonial bud), with variable seta length. The capsules are generally exserted, erect, and symmetric but sometimes inclined and strongly asymmetric, ovoid to pyriform,

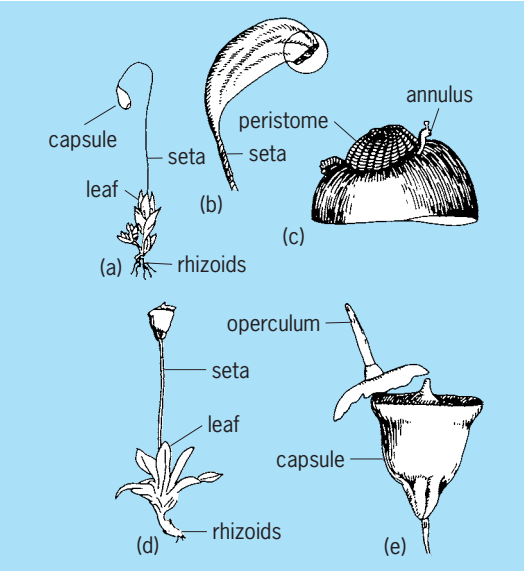

Funariales sporophytes. (*a*) *Funaria hygrometrica*, entire plant (*after H. S. Conard, How To Know the Mosses, Jaques, 1944*); (*b*) deoperculate capsule; (*c*) capsule tip enlarged. (*d*) *Physcomitrium turbinatum* (*P. pyriform*), entire plant; (*e*) capsule (*after W. H. Welch, Mosses of Indiana, Indiana Department of Conservation, 1957*).

and often with a noticeably differentiated sterile neck. The capsules are often inoperculate and lack a peristome, but more often the peristome is present and single or double, the 16 teeth sometimes having endostome segments opposite the teeth and sometimes fused with them. The spores may be spherical or tetrahedral, while the calyptra may be cucullate or mitrate. Chromosome numbers are 9, 11, 12, 13, and 14 or multiples.

The Funariales include five families, and are heterogeneous in composition. The neck of the capsule is well differentiated, and endostome segments are opposite the peristome teeth. The tendency for stomata to consist of a central slit in a single guard cell gives some unity to the Funariaceae, Gigaspermaceae, and Disceliaceae. The uniformity in gametophytic characters contrasts with a striking complexity and diversity of sporophytic features. The Pseudoditrichaceae have narrowly subulate leaves and long cells, but the segments of the inner peristome are opposite the teeth. The Splachnobryaceae have a peristome consisting essentially of endostome segments and an exostome resembling a very low preperistome (actually a scarcely differentiated primary peristomial cell layer). The Ephemeraceae are minute ephemerals with much reduced gametophytes and sporophytes; the capsules are inoperculate and eperistomate. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA.

Howard Crum

Function generator

An electronic instrument which generates periodic voltage or current waveforms that duplicate various types of well-defined mathematical functions. The simplest function generator usually generates a

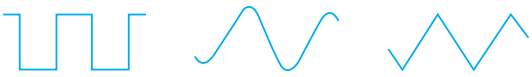

Fig. 1. Waveforms generated by a simple function generator.

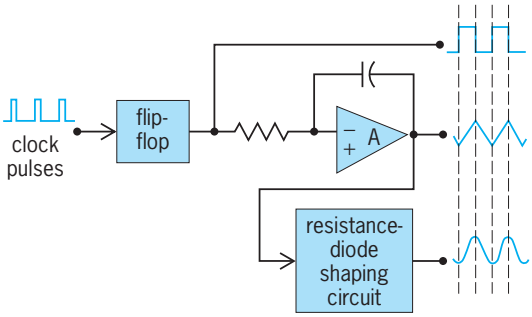

Fig. 2. Elementary function generator.

combination of square waves, triangular waves, and sine waves (**Fig. 1**).

One electronic circuit approach to the design of a simple function generator is to begin with a bistable multivibrator or "flip-flop" controlled in time by a succession of clock pulses which generates the square wave (**Fig. 2**). The triangular waveform is obtained by integrating the square wave through the use of the operational amplifier integrator. The sine wave is obtained by applying the triangular wave to a shaping circuit consisting of a combination of resistors and diodes. *See* AMPLIFIER; MULTIVIBRATOR; WAVE-SHAPING CIRCUITS.

Alternatively the sine wave may be generated by a sinusoidal oscillator (**Fig. 3**). From this output, the square wave may be obtained by amplication, limiting, and clipping of the sine wave. Then the triangular wave may be obtained using an integrator as before. *See* LIMITER CIRCUIT.

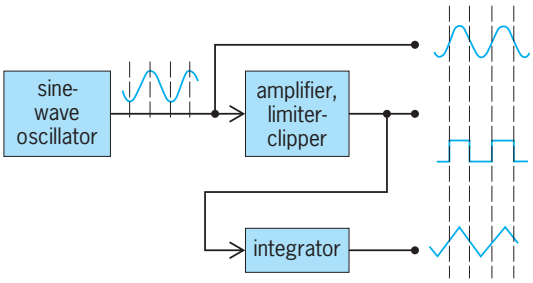

Fig. 3. Alternate form of a function generator.

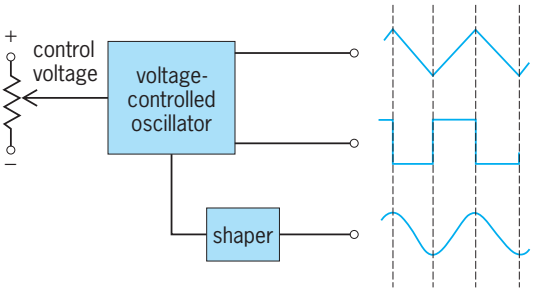

Fig. 4. Voltage-controlled oscillator as a function generator.

A still different configuration of a function generator consists of a circuit, which simultaneously generates triangular waves and square waves, whose rate is controllable by a direct-current level (**Fig. 4**). Such a circuit is identified as a voltage-controlled oscillator (VCO). One of the outputs can be applied to a resistor-diode shaping circuit to obtain the sine wave output. *See* FREQUENCY MODULATOR; PHASE-LOCKED LOOPS.

More sophisticated function generators have circuits built in which will allow many other waveforms to be generated, such as parabolic waveforms and logarithmic functions, using combinations of resistor-diode combinations. Such function generators are widely available as electronic test equipment for general-purpose applications. *See* ELECTRONIC TEST EQUIPMENT.

A combination of counters, programmed read-only memories (PROMs), and a digital-to-analog converter can be used as a function generator, generating almost any function desired to almost any degree of accuracy. *See* COMPUTER STORAGE TECHNOLOGY; DIGITAL-TO-ANALOG CONVERTER. Glenn M. Glasford

Bibliography. H. H. Chiang, *Electronic Wave Forming and Processing Circuits*, 1986, reprint 1989; G. M. Glasford, *Analog Electronic Circuits*, 1986; G. M. Glasford, *Digital Electronic Circuits*, 1988; J. Millman and A. Grabel, *Microelectronics*, 2d ed., 1987; R. B. Northrop, *Analog Electronic Circuits: Analysis and Applications*, 1990.

Functional analysis and modeling (engineering)

The discipline that addresses the activities that a system, software, or an organization must perform to achieve its desired outputs; that is, what transformations are necessary to turn the available inputs into the desired outputs. Additional considerations include the flow of data or items between functions, the processing instructions that are available to guide the transformations, and the control logic that dictates the activation and termination of functions. Functional analysis diagrams have been developed to capture some or all of these concepts.

Functional analysis is performed in systems engineering, software systems engineering, and business process reengineering as a portion of the design process. These design processes typically involve the steps of requirements definition and analysis, functional analysis, physical or resource definition, and operational analysis. This last step of operational analysis involves the marriage of functions with resources to determine if the requirements are met. *See* REENGINEERING; SOFTWARE ENGINEERING; SYSTEMS ENGINEERING.

The concept of examining the logical architecture via functional analysis concurrent with the development of the physical architecture has become a

well-accepted principle in the related fields of systems engineering, software engineering, and business process reengineering. Experience in these fields has shown that rushing to a physical representation of a system or organization is likely to produce a set of resources that cannot produce some of the outputs desired by the system. The resulting systems often suffer from gaps in capability, insufficient feedback and control to achieve the objectives desired, and inefficient processing. Functional analysis and modeling is thus a critical activity for the design and analysis of systems and organizations.

Elements. There are four elements to be addressed by any specific functional analysis approach. First, the functions are represented as a hierarchical decomposition, in which there is a top-level function for the system or organization. The top-level function is partitioned into a set of subfunctions that use the same inputs and produce the same outputs as the top-level function. Each of these subfunctions can then be partitioned further, with the decomposition process continuing as often as it is useful.

Second, functional analysis diagrams can represent the flow of data or items among the functions within any portion of the functional decomposition. As the first and subsequent functional decompositions are examined, it is common for one function to produce outputs that are not useful outside the boundaries of the system or organizations. These outputs are needed by other functions in order to produce the needed and expected external outputs.

The **illustration** is a functional analysis diagram for the functional decomposition of an automated teller machine (ATM), showing the preservation of inputs, controls, and outputs. The top-level function, A0, which provides automated bank transactions for customers, is partitioned into three subfunctions: Subfunction A1 services the customer, A2 provides periodic status reports and identifies system problems, and A3 provides access to the ATM for the bank employees and security for the customers.

Processing instructions are a third element that appear in some functional analysis diagrams. These instructions contain the needed information for the functions to transform the inputs to the outputs.

The fourth element is the control flow that sequences the termination and activation of the functions so that the process is both efficient and effective. Sample questions that this information answers are (1) can these functions work serially or must they be processed concurrently, (2) are these functions activated once or a series of times, and (3) what circumstances dictate that one function be activated rather than another function. Included in this control flow are the completion criteria that dictate that a function has finished, and the activation logic that dictates when a function begins.

Approaches. The major functional analysis approaches are structured analysis, data and control flow diagrams, function flow block diagrams and N^2 charts, and behavior diagrams.

The development of the functional hierarchy can be approached in a top-down (decomposition), bottom-up (composition), or middle-out fashion. Decomposition, often referred to as top-down structuring, begins with the top-level system function and partitions that function into several subfunctions. This decomposition process must conserve all of the inputs to and outputs from the system's top- or zeroth-level function. To conserve inputs or outputs means to use or produce all of them and add no new ones (see illus.). Next, each of the several first-level functions are decomposed (partitioned) into a second level set of subfunctions. Not every function must be decomposed; only those for which additional insight into the production of outputs is needed should be partitioned.

The opposite approach, composition, is a bottom-up approach. With composition, the analyst starts by identifying the simple low-level functions associated with simple scenarios involving only one of the outputs of the system. Each set of low-level functions is a sequence of input, function, output-input, function, output-input, . . . , function, output-input, function, output. These functions usually show up in third, fourth, or even lower levels of the hierarchy. For complex systems this initial step is a substantial amount of work. After the many sets of low-level functions have been defined, the analyst begins the process of grouping the functions into similar groups. These groups are, in turn, aggregated into larger similar groups; this process continues until a hierarchy is formed from bottom to top. The composition method dates back to the 1960s and 1970s when systems engineering was in its infancy; many systems engineers continue to prefer this approach. There is no empirical evidence that either the composition approach or decomposition approach is better.

Ultimately, using a combination of decomposition and composition approaches is wisest. This is sometimes referred to as middle-out. Often, the analyst makes use of a set of low-level functions associated with specific scenarios defined in the operational concept to establish a "sense" of the system, and then posits a top-level decomposition that is likely to meet the needs of the functional analysis.

Feedback and control. Feedback plays an important role in functional analysis and modeling. Feedback and control is the comparison of the actual characteristics of an output with desired characteristics of that output for the purpose of adjusting the process of transforming inputs into that output. Open-loop control processes may or may not make this measurement, but in either case make no adjustments to the process once started. Closed-loop control processes use measurements of the output as feedback for the purpose of adjusting or controlling the transformation process.

Closed-loop control processes contain at least four subprocesses: comparison of current and desired output characteristics, control adjustments to the process based upon the comparison, the transformation process for turning inputs into outputs, and a sensing process for turning the output into measured dimensions that can be compared to the desired

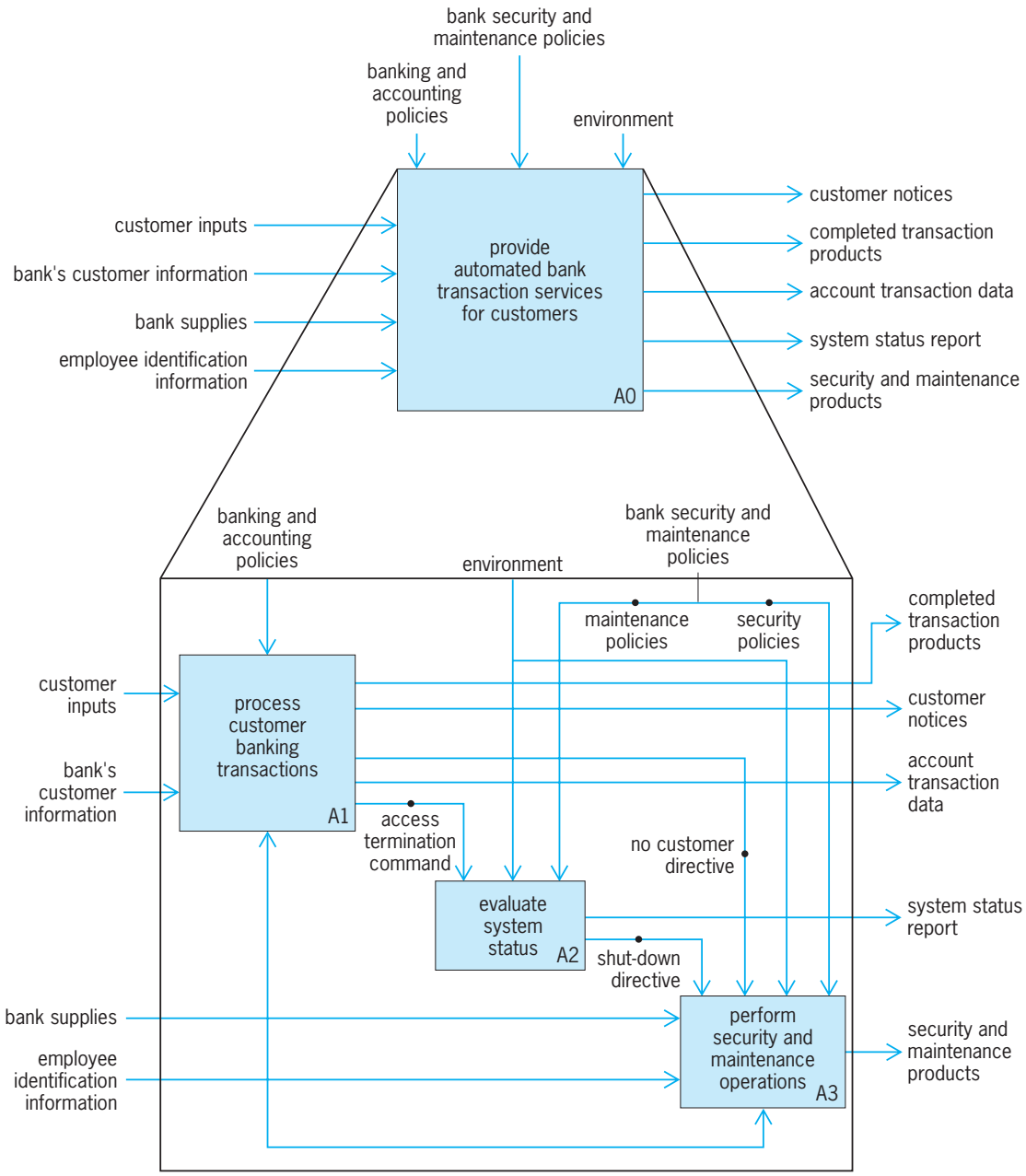

Functional decomposition of an automated teller machine's functions in an IDEF0 model, showing the preservation of inputs, controls, and outputs.

output. The first element is the comparison process in which current values of key variables are compared with desired values of those variables. The comparison process requires definition in advance for what elements of the state of the process are going to be compared. This comparison inevitably introduces a time lag into the process. This element of the feedback process is trivial but at the same time is the cornerstone of the process. The second element is the control process for deciding what to do about the difference between the current value of the output and the desired value of the output. The third element of the feedback process is the transformation process that is being controlled by the feedback process. This process dictates a great deal about how a successful feedback process should be created and is often adapted by the feedback process as part of the correction activity. Sensing the output of the process being controlled is the final element of the feedback process. *See* CONTROL SYSTEMS. Dennis M. Buede

Bibliography. D. M. Buede, *The Engineering Design of Systems*, Wiley, New York, 2000; D. J. Hatley, P. Hruschka, and I. A. Pirbhai, *Process for System Development: A Practical Guide to Requirements and Modeling*, Dorset House, New York, 2000; D. W. Oliver, T. P. Kelliher, and J. P. Keegan, Jr., *Engineering Complex Systems with Models and Objects*, McGraw-Hill, New York, 1997; R. J. Wieringa, Combining static and dynamic modeling methods: A comparison of 4 methods, *Comput. J.*, 38(1):17–30, 1995.

Fundamental constants

The set of physical quantities that play a fundamental role in basic physical theories, primarily at the atomic and subatomic levels, including also those quantities that provide the connection between microscopic physics and the macroscopic world. Examples of fundamental constants are the velocity of light in vacuum c; the Planck constant h; the universal gravitation constant G; the electron charge e; the masses of the proton, the neutron, and the electron, m_p, m_n, and m_e; the Boltzmann constant k; and the Avogadro constant N_A; as well as any physical constant such as the fine-structure constant α that is a combination of these constants.

These quantities typify the scope of the field. The speed of light c and the gravitation constant G provide the scale for Einstein's theory of relativity; Planck's constant h is the fundamental magnitude of quantum mechanics. The fine-structure constant α and the masses of the fundamental particles, together with c and h, are the basic quantities that define the structure of atoms and molecules and provide the foundation for chemistry and atomic physics. The Avogadro constant N_A is the connection between the atomic and macroscopic scales. *See* ATOMIC STRUCTURE AND SPECTRA; AVOGADRO'S NUMBER; PLANCK'S CONSTANT; QUANTUM MECHANICS; RELATIVITY.

As considered here, the rubric of fundamental constants combines the fields of metrology, statistics, and data analysis to determine the best estimates of their values. Why the physical constants have the values that they do will not be discussed. For one thing, that question strictly can be asked only of dimensionless quantities (quantities of dimension 1), such as m_p/m_e, α, or Gm_p^2/hc, since the value of a quantity with dimensions depends on the arbitrary units used to measure it. As an example, since 1983 the velocity of light c has been defined to be 299 792 458 m/s, thus defining the meter as the distance light travels in a specified time as determined by an ideal clock. "Why is c so much larger than 'ordinary' speeds?" is a different kind of question, and finds its answer in atomic physics and quantum mechanics. *See* LIGHT.

Some fundamental physical quantities can be measured directly in terms of the basic units of measurement, but in most cases an experiment yields only a combination such as m_p/m_e, h/e, N_Ah, or e^2/h. The fine-structure constant α is proportional to e^2/h and may be considered to be a more fundamental constant than e itself. In the unrationalized (three-dimensional) centimeter-gram-second (cgs) system of units, $\alpha = e^2/\hbar c$ (where $\hbar c = h/2\pi$); in the modern rationalized, four-dimensional International System of Units (SI), this becomes $\alpha = [e^2/(4\pi\epsilon_0)]/\hbar c = \mu_0 ce^2/(2h)$, where ϵ_0 is the permittivity of vacuum and $\mu_0 = 1/(\epsilon_0 c^2) \equiv 4\pi \times 10^{-7}$ N/A^2 is the permeability of vacuum.

No relationship, such as the foregoing one for α, is known between the gravitational constant G and other physical constants, and one of the major challenges in theoretical physics is grand unification, which attempts to combine gravitation, electromagnetism, and nuclear forces into a single entity, presumably establishing a relationship between G and α, h, c, and the masses of elementary fundamental particles. *See* GRAND UNIFICATION THEORIES; GRAVITATION.

Importance of measurement. The interest in precise measurements is due to the existence of quantities in nature that are invariable. Without the development of quantum mechanics at the beginning of the twentieth century, there could be no basis for the assumption that the masses of all protons were exactly the same. Whereas the mass of a baseball is easily measured to within a fraction of a percent, the measurement applies only to that one ball, and there is little need to be more precise, since the masses of legal baseballs may vary by as much as $\pm 2.5\%$. The mass of a proton or an electron, however, is known to within 1 part in 10^7, and because of the quantization of physical quantities this applies to individual particles, not merely to the mean mass of a large number of particles. Since the properties of atoms and molecules are the same wherever and whenever they are measured (and there is no evidence to contradict the assumption that this holds throughout the universe), the atomic constants provide a means for preserving and reproducing standards of measurements. Except for the unit of mass, the world's measurement standards are derived from the speed of light and the properties of atomic systems rather than from arbitrary artifacts. *See* PHYSICAL MEASUREMENT.

A second important reason for determining the values of the atomic constants with all possible precision and accuracy (precision referring to the reproducibility of the result of repeated measurement, and accuracy referring to the degree to which the measurement gives the true or correct value) is that the only test of the validity of physics is the ability to predict the outcome of experiment. If two determinations of a physical property do not agree within the uncertainties of the measurements, either the experimental measurement or the underlying theory of at least one of the determinations is wrong.

An early example of precise measurement leading to new physics was Lord Rayleigh's attempt in 1895 to improve the precision of the determination of the atomic weight of nitrogen. He found that after the oxygen, carbon dioxide, and water vapor had been removed from air the density of the remaining gas (presumed to be pure nitrogen) was greater by 0.5% than that of nitrogen prepared from chemical sources such as ammonia. This observation, and the conviction that the measurements were sufficiently precise that the discrepancy could not be the result of experimental error, led to the identification of the previously unknown gas argon and ultimately to a whole new family of noble gases.

Another example is the discovery of the Lamb shift. According to the Dirac theory of atomic spectra, the S and $P_{1/2}$ states in atomic hydrogen should have the same energy, but in 1947 W. Lamb, Jr., and

R. C. Retherford, in a measurement of microwave absorption spectra, showed that the $2S$ level was 1.3 parts per million (ppm) higher than the $2P_{1/2}$ level. The search for the explanation of this shift led to the development of quantum electrodynamics (QED). Indications of such a shift had actually been seen in the optical spectrum as much as 20 years earlier, but the precision of those measurements was sufficiently low that it was not generally recognized as a real effect. *See* QUANTUM ELECTRODYNAMICS.

History of measurements. The attempts to determine the numerical values of physical constants date back to the measurements of c and G in the seventeenth and eighteenth centuries. Following the development of the Bohr theory of the atom and quantum mechanics in the first three decades of the twentieth century, precise measurements were made of the elementary charge e and the ratios e/h and e/m_e. Technical developments in spectroscopy led to successively more accurate determinations of the Rydberg constant R_∞, while x-ray spectroscopy and crystallography led to determinations of the Avogadro constant N_A. Microwave technology developed during and after World War II greatly broadened the capability for precise measurement and made possible microwave spectroscopy of the fine structure and hyperfine structure of hydrogen which, with a sufficiently refined theory, provide values of the fine-structure constant α. In addition, nuclear magnetic resonance techniques yielded measurements of the ratios of electron and proton magnetic moments and cyclotron frequencies, which give m_p/m_e. The proton gyromagnetic moment γ_p combined with the proton magnetic moment ratio μ_p/μ_N provides a determination of the Faraday constant, given in Eq. (1), that is free of

$$F = N_A e = \frac{\gamma_p M_p}{\mu_p/\mu_N} \tag{1}$$

the problems associated with the classical electrochemical determination of that quantity. (Here μ_p is the proton magnetic moment, μ_N is the nuclear magneton, and M_p is the proton molar mass.) *See* ELECTROCHEMICAL EQUIVALENT; ELECTRON; ELECTRON SPIN; FINE STRUCTURE (SPECTRAL LINES); GYROMAGNETIC RATIO; HYPERFINE STRUCTURE; MAGNETISM; MAGNETON; MICROWAVE SPECTROSCOPY; NUCLEAR MAGNETIC RESONANCE (NMR); RYDBERG CONSTANT.

The study of atomic constants received an unexpected stimulus from solid-state physics in 1967. The Josephson effect in superconductors provided the basis for a determination of the ratio $2e/h = K_J = nf/V(n)$, where f is the frequency of the microwave field on the junction and $V(n)$ is the potential difference of the nth step. Until this discovery, and the surprising demonstration of the precision achievable in the measurement of this ratio, the determination of e/h was based on measurements of the photoelectric effect at optical wavelengths or its equivalent, the short-wavelength limit of the continuous x-ray spectrum. These measurements were inherently much less precise than the Josephson effect, which determined the voltage-wavelength product (or the voltage/frequency ratio) in the microwave region of the electromagnetic spectrum. The value of $2e/h$ could be combined with the proton gyromagnetic ratio to determine a value of the fine-structure constant that was independent of atomic spectroscopy and of any uncertainties of quantum electrodynamics. *See* JOSEPHSON EFFECT.

An analysis of the atomic constants in 1963 had to deal with two incompatible values of the fine-structure constant determined from the microwave spectrum of hydrogen, one of which included a contribution dependent on the structure of the proton. If the proton is not an elementary particle but has an internal structure, its center of charge in a hydrogen atom would be affected by the electromagnetic field of the orbiting electron and might not be identical with its center of mass. This recoil effect had been estimated to produce a shift in the hyperfine interval that might be as large as 35 ppm. This uncertainty of the contribution required that the measurement be excluded from the analysis. However, by 1969 additional theoretical studies had indicated that these corrections were in fact no more than 2–3 ppm. The value of α deduced from the Josephson effect and the proton gyromagnetic ratio was then in excellent agreement with the hyperfine-structure value. The fine-structure value of α was thus shown to be too low by 20 ppm. More significantly, since the solid-state value α was independent of quantum electrodynamics, the validity of that theory could now be tested unambiguously by comparing theory and experiment without using those experiments at the same time to determine α. *See* PROTON.

This is an example of the unity of science: a low-temperature solid-state measurement gave information that resolved a problem in atomic physics (fine-structure spectrum of hydrogen, characteristic energy ~0.1 meV) and verified a calculation associated with an important question in high-energy physics (proton structure, characteristic energy >10 GeV).

Another connection between solid-state and atomic physics, discovered by K. von Klitzing in 1980, is the quantization of the ratio of the transverse potential (V) to the longitudinal current (I; Hall coefficient) in a two-dimensional electron gas at a semiconductor heterojunction at high magnetic field and sufficiently low temperature.

Specifically, when the longitudinal potential approaches zero, this ratio is given by Eq. (2), where

$$\frac{V}{I} = \frac{R_K}{i} = \frac{h}{e^2 i} \tag{2}$$

i is an integer. The von Klitzing effect provides a determination of the fine-structure constant as well as a reproducible and drift-free resistance standard that is much more convenient than that provided by the calculable capacitor. *See* HALL EFFECT.

The level of precision of measurements has improved steadily. In one of the earliest careful studies in this field, by R. T. Birge in 1929, the uncertainty

of the atomic constants was typically 1 part in 500 to 1 part in 1000 (except for c with a precision of 20 ppm and R_∞ with 0.15 ppm). By 1950 the level of precision of the constants had improved by a factor of roughly 50, to the order of 30 ppm. In the 1998 analysis the level of uncertainty had generally been reduced to levels of 0.1 ppm to 0.01 ppm. Thus, over a period of 70 years the precision of the determinations has improved by a factor of 5 every 10 years.

Atomic standards for electrical units. The improvements in precision of the fundamental constants have had an important influence on electrical engineering. Units for electrical quantities are defined from the laws of electromagnetism (Ampère's law, Biot-Savart law, and Ohm's law) and the requirement for coherence between the units of mechanical and electrical energy. As a practical matter, however, the units of electrical potential and electrical resistance had been maintained by using electrolytic cells and resistance coils, respectively. These artifact standards must be calibrated in terms of the SI definitions, but when the measurements of the fundamental constants had a relative precision of 10^{-4} to 10^{-3}, the uncertainty in the consistency of the laboratory electrical standards with their definitions was unimportant. When the relative precision improved to the level of 10^{-5} and better, these calibrations had to be included with the other pertinent calibration quantities in the analysis of the values of physical constants. *See* AMPÈRE'S LAW; BIOT-SAVART LAW; OHM'S LAW.

Prior to 1950, electrical units could be evaluated only by using a current balance to compare (at a level of the order of 1–10 ppm) electrical and mechanical forces on a current-carrying coil. The development of the Thompson-Lampard calculable capacitor made it possible to calibrate a laboratory standard of resistance directly, depending only on geometrical measurements and the value of c, while the Josephson effect provided a means to monitor the constancy of voltage standards. Beginning in 1973, many national laboratories adopted a representation of the volt using the definition $(2e/h)V_{\mathrm{LAB}} = 483{,}594$ GHz, where V_{LAB} is the volt standard. (The United States had used the frequency 483,593.42 GHz since July 1972 to maintain its standard of the volt.) It then became evident, based on such calibrations and on international intercomparisons, that resistance coils and standard cells can drift in value by parts per million per decade. *See* CAPACITANCE MEASUREMENT.

Based on the results of the 1986 adjustment of the physical constants and on subsequent additional measurements of both the Josephson effect and the von Klitzing effect, new standards for the volt ($2e/h = 493{,}597.9$ GHz/V_{90}) and the ohm ($R_H = 25{,}812.807\ \Omega_{90}$) became effective on January 1, 1990. A realization of the volt in terms of the Josephson effect reproduces the SI volt with an

Recommended values (2002) of selected fundamental physical constants

Quantity	Symbol	Numerical value*	Units†	Relative uncertainty (standard deviation)
Speed of light in vacuum	c	299792458	m/s	(defined)
Permeability of vacuum	μ_0	$4\pi \times 10^{-7}$	N/A^2	(defined)
Permittivity of vacuum	ϵ_0	8.854187817. . .	10^{-12} F/m	(defined)
Constant of gravitation	G	6.6742 (10)	10^{-11} m^3/(kg · s^2)	1.5×10^{-4}
Planck constant	h	6.6260693 (11)	10^{-34} J · s	1.7×10^{-7}
Elementary charge	e	1.60217653 (14)	10^{-19} C	8.5×10^{-8}
Magnetic flux quantum, $h/(2e)$	Φ_0	2.06783372 (18)	10^{-15} Wb	8.5×10^{-8}
Fine-structure constant,	α	7.297352568 (24)	10^{-3}	3.3×10^{-9}
$\mu_0 ce^2/(2h)$	α^{-1}	137.03599911 (46)		3.3×10^{-9}
Electron mass	m_e	9.1093826 (16)	10^{-31} kg	1.7×10^{-7}
Proton mass	m_p	1.67262171 (29)	10^{-27} kg	1.7×10^{-7}
Neutron mass	m_n	1.67492728 (29)	10^{-27} kg	1.7×10^{-7}
Proton-electron mass ratio	m_p/m_e	1836.15267261 (85)		4.6×10^{-10}
Rydberg constant, $m_e c\alpha^2/(2h)$	R_∞	10973731.568525 (73)	m^{-1}	6.6×10^{-12}
Bohr radius, $\alpha/(4\pi R_\infty)$	a_0	5.291772108 (18)	10^{-11} m	3.3×10^{-9}
Compton wavelength of the electron, $h/(m_e c) = \alpha^2/(2R_\infty)$	λ_c	2.426310238 (16)	10^{-12} m	6.7×10^{-9}
Classical electron radius, $\mu_0 e^2/(4\pi m_e) = \alpha^3/(4\pi R_\infty)$	r_e	2.817940325 (28)	10^{-15} m	1.0×10^{-8}
Bohr magneton, $eh/(4\pi m_e)$	μ_B	9.27400949 (80)	10^{-24} J/T	8.6×10^{-8}
Electron magnetic moment	μ_e	−9.28476412 (80)	10^{-24} J/T	8.6×10^{-8}
Electron magnetic moment/Bohr magneton ratio	μ_e/μ_B	−1.0011596521859 (38)		3.8×10^{-12}
Nuclear magneton, $eh/(4\pi m_p)$	μ_N	5.05078343 (43)	10^{-27} J/T	8.6×10^{-8}
Proton magnetic moment/nuclear magneton ratio	μ_p/μ_N	2.792847351 (28)		1.0×10^{-8}
Avogadro constant	N_A	6.0221415 (10)	10^{23}	1.7×10^{-7}
Faraday constant, $N_A e$	F	96485.3383 (83)	C/mol	8.6×10^{-8}
Molar gas constant	R	8.314472 (15)	J/(mol · K)	1.7×10^{-6}
Boltzmann constant, R/N_A	k	1.3806505 (24)	10^{-23} J/K	1.8×10^{-6}

*As recommended by the CODATA Task Group on Fundamental Constants. The digits in parentheses represent the one-standard-deviation uncertainties in the final two digits of the quoted value.

†A = ampere, C = coulomb, F = farad, J = joule, kg = kilogram, K = kelvin, m = meter, mol = mole, N = newton, s = second, T = tesla, Wb = weber.

overall relative uncertainty of 4 parts in 10^7 (one standard deviation), while a realization of the ohm in terms of the von Klitzing effect reproduces the ohm with an uncertainty of 2 parts in 10^7. The watt balance measures the force on a current-carrying coil in a constant magnetic field. This force is equal to BIl, where B is the magnetic field, I is the current, and l is the length of the coil, and it is balanced by a gravitational force mg, where m is the mass of the coil and g is the acceleration of gravity. Then, with the current source removed, the coil is moved at constant speed v in the same field, and the induced potential difference $V = Blv$ between its ends is measured. Hence it is possible to equate the electrical and mechanical power, $VI = mgv$. Since the electrical units are maintained in term of $R_K = h/e^2$ and $K_J = 2e/h$, these measurements allow a determination of the Planck constant h. Such measurements not only provide a means to monitor possible changes in the mass of the International Prototype of the kilogram but can ultimately lead to its redefinition in terms of invariant atomic standards. *See* ELECTRICAL UNITS AND STANDARDS; RESISTANCE MEASUREMENT; VOLTAGE MEASUREMENT; WATT BALANCE.

Analysis of experimental data. The various experimental measurements of the atomic constants are not independent. For example, the fine-structure constant may be determined from atomic spectroscopy of hydrogen; from the quantum Hall resistance (von Klitzing effect); by combining R_∞, the Josephson constant $2e/h$, the proton gyromagnetic moment, and the ratio of the proton nuclear magnetic resonance frequency to the electron cyclotron frequency; or from a measurement of the anomalous moment of the electron (coupled with a calculation of its theoretical expression). Similarly, the Avogadro constant can be determined directly by measuring the density of a crystal, its atomic lattice spacings, and its molar mass, and indirectly by combining measurements of the von Klitzing constant, the Josephson constant, and the Faraday constant; while the Faraday constant, measured directly by the electrolysis of silver, is also indirectly deduced from measurements of the proton gyromagnetic moment, using Eq. (1).

Because of such relationships among the measurements of the atomic constants, all of the data, including the calibration factors relating the international standards of the ohm and the volt to their SI definitions, must be analyzed as a single entity. In determining a "best" value of a single quantity from repeated measurements, the average is generally used, often with statistical weights so that the more reliable data are given more importance in the average. The formalism of least squares is an appropriate generalization of this process when more than one quantity is to be determined from a set of data. *See* LEAST-SQUARES METHOD.

Since 1969 the adjustment of the fundamental constants (see **table**) has been carried out under the auspices of the CODATA Task Group on Fundamental Constants. CODATA, the Committee on Data for Science and Technology, is an interdisciplinary committee of the International Council of Scientific Unions.

E. Richard Cohen

Bibliography. G. Cohen-Tannoudji, *Universal Constants in Physics*, Horizons of Science Series, McGraw-Hill, New York, 1992; W. H. McCrea and M. J. Ree (eds.), *The Constants of Physics* (Proceedings of a Royal Society discussion meeting held May 25–26, 1983, London), 1985; P. J. Mohr and B. N. Taylor, CODATA recommended values of the fundamental physical constants: 1998, *Rev. Mod. Phys.*, 72(2):351–495, 2000; P. J. Mohr and B. N. Taylor, CODATA recommended values of the fundamental physical constants: 2002, *Rev. Mod. Phys.*, 77(1):1–107, 2005; B. W. Petley, *The Fundamental Physical Constants and the Frontier of Measurement*, Adam Hilger, Boston, 1988.

Fundamental interactions

Fundamental forces that act between elementary particles, of which all matter is assumed to be composed.

Properties of Interactions

At present, four fundamental interactions are distinguished. The properties of each are summarized in the **table**.

Gravitational interaction. This interaction manifests itself as a long-range force of attraction between all elementary particles. The force law between two particles of masses m_1 and m_2 separated by a distance r is well approximated by the newtonian expression $G_N(m_1m_2/r^2)$, where G_N is the newtonian constant, equal to $6.6742 \pm 0.0010 \times 10^{-11}\ \mathrm{m^3 \cdot kg^{-1} \cdot s^{-2}}$. The dimensionless quantity $(G_N m_e m_p)/(\hbar c)$ is usually taken as the constant characterizing the gravitational interaction, where m_e, and m_p are the electron and proton masses, $2\pi\hbar$ is Planck's constant, and c is the velocity of light. *See* GRAVITATION.

Electromagnetic interaction. This interaction is responsible for the long-range force of repulsion of like, and attraction of unlike, electric charges. The dimensionless quantity characterizing the strength of electromagnetic interaction is the fine-structure constant, given by Eq. (1) in SI units, where e is the electron charge and ϵ_0 is the permittivity of empty space.

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} = \frac{1}{(137.03599911 \pm 0.00000046)} \tag{1}$$

At comparable distances, the ratio of gravitational to electromagnetic interactions (as determined by the

Properties of the four fundamental interactions

Interaction	Range	Exchanged quanta
Gravitational	Long-range	Gravitons (g)
Electromagnetic	Long-range	Photons (γ)
Weak nuclear	Short-range $\approx 10^{-18}$ m	W^+, Z^0, W^-
Strong nuclear	Short-range $\approx 10^{-15}$ m	Gluons (G)

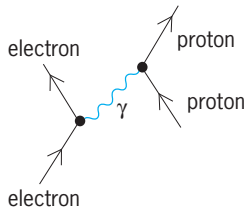

Fig. 1. Feynman diagram of electromagnetic interaction between an electron and a proton.

strength of respective forces between an electron and a proton) is given by the quantity $4\pi\epsilon_0 G_N m_e m_p/e^2$, which is approximately 4×10^{-37}. *See* COULOMB'S LAW; ELECTROSTATICS.

In modern quantum field theory, the electromagnetic interaction and the forces of attraction or repulsion between charged particles are pictured as arising secondarily as a consequence of the primary process of emission of one or more photons (particles or quanta of light) emitted by an accelerating electric charge (in accordance with Maxwell's equations) and the subsequent reabsorption of these quanta by a second charged particle. The space-time diagram (introduced by R. P. Feynman) for one photon exchange is shown in **Fig. 1**. A similar picture may also be valid for the gravitational interaction (in accordance with the quantum version of A. Einstein's gravitational equations), but with exchanges of zero-rest-mass gravitons (g) rather than zero-rest-mass photons. (The existence of the graviton, however, has not yet been experimentally demonstrated.) *See* FEYNMAN DIAGRAM.

In accordance with this picture, the electromagnetic interaction (to one photon exchange approximation) is usually represented by reaction (2), where

$$e + P \rightarrow (e + \gamma) + P \rightarrow e + (P + \gamma) \rightarrow e + P \quad (2)$$

γ is the photon, emitted by the electron and reabsorbed by the proton. For this interaction, and also for the gravitational interaction represented by reaction (3), the nature of the participating parti-

$$e + P \rightarrow (e + g) + P \rightarrow e + (P + g) \rightarrow e + P \quad (3)$$

cles (electron e and proton P) is the same, before and after the interaction, and the exchanged quanta (γ or g) are electrically neutral. *See* GRAVITON; LIGHT; MAXWELL'S EQUATIONS; PHOTON; QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; QUANTUM GRAVITATION; QUANTUM MECHANICS.

Weak nuclear interactions. The third fundamental interaction is the weak nuclear interaction, which is responsible for the decay of a neutron into a proton, an electron, and an antineutrino. Its characteristic strength for low-energy phenomena is measured by the Fermi constant G_F, which is equal to $1.026 \times 10^{-5}\ m_p^{-2}\hbar^3/c$. Unlike electromagnetism and gravitation, weak interactions are short-range, with a force law of the type $e^{-M_W cr/\hbar}$, the range of the force ($\hbar/M_W c$) being of the order of 10^{-18} m. Until 1973, the only known weak interactions were those which changed the nature of the interacting particles (unlike electromagnetism and gravity). For example, consider reactions (4), where P is the proton, N is the

$$P + e^- \xrightarrow{\text{weak}} N + \nu_e \quad \text{(this reaction is equivalent to } \beta \text{ decay of the neutron } N \rightarrow P + e^- + \overline{\nu}_e) \quad (4a)$$

$$P + \mu^- \xrightarrow{\text{weak}} N + \nu_\mu \quad \text{(muon capture by a proton with the emission of a neutrino)} \quad (4b)$$

$$\mu^- + \nu_e \xrightarrow{\text{weak}} \nu_\mu + e^- \quad \text{(this reaction is equivalent to muon decay } \mu^- \rightarrow e^- + \overline{\nu}_e + \nu_\mu) \quad (4c)$$

neutron, μ^- is the negatively charged muon, ν_e and ν_μ are the electronic and muonic neutrinos, and $\hbar_e$ and $\hbar_\mu$ are the corresponding antineutrinos. In reaction (4a), the weak interaction transforms a proton into a neutron and at the same time an electron into a neutrino.

An important question was finally answered in 1983: Is the weak interaction similar to electromagnetism in being mediated primarily by intermediate objects, the W^+ and W^- particles. If this is the case, then reactions (4a) and (4c), for example, would in detail be represented as reactions (5a) and (5b).

$$P^+ + e^- \rightarrow (N^0 + W^+) + e^- \rightarrow N^0 + (W^+ + e^-) \rightarrow N^0 + \nu^0 \quad (5a)$$

$$\mu^- + \nu_e^0 \rightarrow (\nu_\mu^0 + W^-) + \nu_e^0 \rightarrow \nu_\mu^0 + (W^- + \nu_e^-) \rightarrow \nu_\mu^0 + e^- \quad (5b)$$

[The superscript on each particle gives its electrical charge (+, 0, −) in units of the proton's charge.] The experimental answer (discovered at the CERN laboratory at Geneva) is that W^+ and W^- do exist, with a mass m_W of 80.9 ± 2.0 GeV/c^2. Each carries a spin of magnitude $\hbar$ just as does the photon (γ). The mass of these particles gives the range [$\hbar/(m_W c) = 10^{-18}$ m] of the weak interaction, and is also related to its strength G_F, as discussed below. *See* INTERMEDIATE VECTOR BOSON.

Another crucial discovery in weak interaction physics was the neutral current phenomenon in 1973, that is, the discovery of new types of weak interactions where (as in the case of electromagnetism or gravity) the nature of the interacting particles is not changed during the interaction, as in reactions (6).

$$\nu_\mu + e^- \xrightarrow{\text{weak}} \nu + e^- \quad (6a)$$

$$\nu_\mu + P \xrightarrow{\text{weak}} \nu_\mu + P \quad (6b)$$

$$\nu_e + N \xrightarrow{\text{weak}} \nu_e + N \quad (6c)$$

$$e^- + P \xrightarrow{\text{weak}} e^- + P \quad (6d)$$

See NEUTRAL CURRENTS.

The 1983 experiments at CERN also gave evidence for the existence of an intermediate particle Z^0 which is believed to mediate such reactions. Thus

reaction (6*a*), expressed in detail, is reaction (7).

$$\nu_\mu + e^- \rightarrow (\nu_\mu + Z^0) + e^- \rightarrow \nu_\mu + (Z^0 + e^-) \rightarrow \nu_\mu + e^- \quad (7)$$

The mass m_Z of the Z^0 has been found to be 93.0 ± 2.0 GeV/c^2. The magnitudes of the W^+, W^-, and Z^0 masses had been predicted by the unified theory of electromagnetic and weak interactions (the electroweak interaction, discussed below), 16 years before the experiments which discovered them.

In contrast to gravitation, electromagnetism, and strong nuclear interactions, weak interactions violate left-right and particle-antiparticle symmetries. *See* PARITY (QUANTUM MECHANICS); SYMMETRY LAWS (PHYSICS); WEAK NUCLEAR INTERACTIONS.

Strong nuclear interaction. The fourth fundamental interaction is the strong nuclear interaction between protons and neutrons, which resembles the weak nuclear interaction in being short-range, although the range is of the order of 10^{-15} m rather than 10^{-18} m. Within this range of distances the strong force overshadows all other forces between protons and neutrons, with a characteristic strength parameter of the order of unity (compared with the electromagnetic strength parameter $\alpha \approx 1/137$).

Protons and neutrons are themselves believed to be made up of yet more fundamental entities, the up (u) and down (d) quarks ($P = uud$, $N = udd$). Each quark is assumed to be endowed with one of three color quantum numbers [conventionally labeled red (r), yellow (y), and blue (b)]. The strong nuclear force can be pictured as ultimately arising through an exchange of zero rest-mass color-carrying quanta of spin $\hbar$ called gluons (G) [analogous to photons in electromagnetism], which are exchanged between quarks (contained inside protons and neutrons), as in reaction (8). Since neutrinos, electrons, and muons

$$\begin{aligned}\text{Quark} + \text{quark} &\rightarrow (\text{quark} + \text{gluon}) + \text{quark}\\ &\rightarrow \text{quark} + (\text{gluon} + \text{quark})\\ &\rightarrow \text{quark} + \text{quark} \qquad (8)\end{aligned}$$

(the so-called leptons) do not contain quarks, their interactions among themselves or with protons and neutrinos do not exhibit the strong nuclear force. There is indirect experimental evidence for the existence of the gluons and of their spin being $\hbar$. *See* COLOR (QUANTUM MECHANICS); GLUONS; LEPTON; QUANTUM CHROMODYNAMICS; QUARKS; STRONG NUCLEAR INTERACTIONS.

Gauge interactions. Three of the four fundamental interactions (electromagnetic, weak nuclear, and strong nuclear) appear to be mediated by intermediate quanta (photons γ; W^+, Z^0, and W^-; and gluons G, respectively), each carrying spin of magnitude $\hbar$. This is characteristic of the gauge interactions, whose general theory was given by H. Weyl, C. N. Yang, R. Mills, and R. Shaw. This class of interactions is further characterized by the fact that the force between any two particles (produced by the mediation of an intermediate gauge particle) is universal in the sense that its strength is (essentially) proportional to the product of the intrinsic charges (electric, or weak-nuclear, or strong-color) carried by the two interacting particles concerned.

The fourth interaction (the gravitational) can also be considered as a gauge interaction, with the intrinsic charge in this case being the mass; the gravitational force between any two particles is proportional to the product of their masses. The only difference between gravitation and the other three interactions is that the gravitational gauge quantum (the graviton) carries spin $2\hbar$ rather than $\hbar$. As discussed below, it is an open question whether all fundamental interactions are gauge interactions. *See* GAUGE THEORY.

Unification of Interactions

Ever since the discovery and clear classification of these four interactions, particle physicists have attempted to unify these interactions as aspects of one basic interaction between all matter. The work of M. Faraday and J. C. Maxwell in the nineteenth century, which united the distinct forces of electricity and magnetism as aspects of a single interaction (the gauge interaction of electromagnetism), has served as a model for such unification ideas.

Gravitation and electromagnetism. The first attempt in this direction was made by Einstein who, having succeeded in understanding gravitation as a manifestation of the curvature of space-time, tried to comprehend electromagnetism as another geometrical manifestation of the properties of space-time, thus achieving a unification between these forces. In this attempt, to which he devoted all his later years, he is considered to have failed. *See* RELATIVITY.

Electroweak interaction. A unification of weak and electromagnetic interactions, employing the gauge ideas discussed above, was suggested by S. Glashow and by A. Salam and J. C. Ward in 1959. This followed a parallel between these two interactions, pointed out by J. S. Schwinger in 1957. Assuming that (the then known) weak interactions (4) were mediated by exchanges of (the then hypothetical) W^+ and W^- particles, it could be shown from the empirical properties of weak interaction phenomena, that if the W's existed, they must carry an intrinsic spin of magnitude $\hbar$, just as does the photon, the gauge quantum of electromagnetism. If a bold unifying assumption was made that this magnitude of spins $\hbar$ for W^+, W^-, and the photon γ connotes a gauge character for a unified electroweak interaction, and that the intrinsic coupling strength of weak interactions is universally the same as that for electromagnetism (that is, $\alpha = 1/137$), then it could be shown that the masses of the W^+ and W^- particles must be in excess of the quantity given in Eq. (9).

$$\sqrt{\pi\alpha\hbar^3/(\sqrt{2}G_F c)} = 37.4 \text{ GeV}/c^2 \qquad (9)$$

Following this initial attempt, Glashow (and independently Salam and Ward) noted that such a unification hypothesis is incomplete, inasmuch as electromagnetism is a left-right symmetry-preserving

interaction, in contrast to the weak interaction, which violates this symmetry. A gauge unification of such disparate interactions could be effected only if, additionally, new weak interactions represented by reactions (5) are also postulated to exist. Equivalently, there must exist a new electrically neutral intermediate weak-quantum Z^0 besides the (hypothetical) W^+ and W^-.

Spontaneous breaking and renormalization. There were two major problems with this unified electroweak gauge theory considered as a fundamental theory. Yang and Mills had shown that masslessness of gauge quanta is the hallmark of unbroken gauge theories. The origin of the masses of the weak interaction quanta W^+, W^-, and Z^0 (or equivalently the short-range of weak interactions), as contrasted with the masslessness of the photon (or equivalently the long-range character of electromagnetism), therefore required explanation. The second problem concerned the possibility of reliably calculating higher-order quantum effects with the new unified electroweak theory, on the lines of similar calculations for the "renormalized" theory of electromagnetism elaborated by S. Tomonaga, Schwinger, Feynman, and F. J. Dyson around 1949. The first problem was solved by S. Weinberg and Salam and the second by G. t'Hooft and by B. W. Lee and J. Zinn-Justin. *See* RENORMALIZATION.

Weinberg and Salam considered the possibility of the electroweak interaction being a "spontaneously broken" gauge theory. By introducing an additional self-interacting Higgs-Englert-Brout-Kibble particle into the theory, they were able to show that the W^+, W^-, and Z^0 would acquire well-defined masses through the so-called Higgs mechanism, these masses being given by Eqs. (10), where 37.4 GeV/c^2

$$m_W = \frac{37.4\ \text{GeV}/c^2}{\sin\theta_w} \qquad m_Z = \frac{37.4\ \text{GeV}/c^2}{\sin\theta_w \cos\theta_w} \tag{10}$$

is the combination of constants given by Eq. (9). Here θ_w is a weak mixing parameter for electromagnetism and weak interactions. The constant $\sin^2\theta_w$ can be determined from experiments which give the ratios of cross sections of Z^0-mediated reactions (6) to the W^+ and W^--mediated reactions (4). The best available value, calculated from all low-energy experiments, is given by Eq. (11).

$$\sin^2\theta_w \approx 0.230 \tag{11}$$

See SYMMETRY BREAKING.

The predicted theoretical mass values of the W and Z particles deduced by substituting Eq. (11) into Eqs. (10) are in good accord with the experimental values found by the CERN 1983 experiments. The existence of the W and Z particles and this accord with regard to mass values give support to the basic correctness of the electroweak unification ideas, as well as to the gauge character of the electroweak interaction.

Prior to this direct evidence, indirect evidence for the existence of the characteristic reactions (6), predicted by the electroweak theory, had existed since

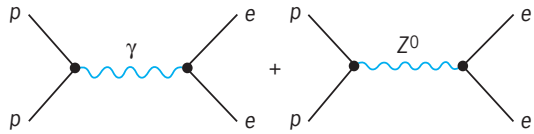

Fig. 2. Feynman diagrams corresponding to the Stanford (1978) experiment, which showed the interference effects between mediating photons and Z^0 exchanges in electron-proton scattering.

1973. The most crucial experiment in this respect, carried out at Stanford during 1978, exhibited interference effects between the photon (γ) and the Z^0 particle in the scattering of polarized electrons from protons (**Fig. 2**). These effects were established through observing the characteristic weak left-right symmetry violation in the reaction $e^- + d \rightarrow e^- + d$. The findings of this experiment provided indirect but quantitative confirmation of the predictions of the electroweak theory.

Higgs particle. The Weinberg-Salam electroweak theory contains an additional neutral particle (the Higgs) but does not predict its mass. A search for this particle will be seriously undertaken when the electron-positron accelerators LEP (Large Electron-Positron storage ring) at CERN and SLC (Stanford Linear Collider) come into commission. Meanwhile, there has been theoretical speculation on whether the Higgs particle is a composite object held together by a new, fundamental type of very strong interaction, the so-called technicolor interaction. This suggestion has the possible merit of eliminating the need for introducing a fundamental nongauge (self-) interaction among Higgs particles. *See* ELECTROWEAK INTERACTION; HIGGS BOSON; PARTICLE ACCELERATOR; STANDARD MODEL.

Electronuclear interaction. The gauge unification of weak and electromagnetic interactions, which started with the observation that the relevant mediating quanta (W^+, W^-, Z^0, and γ) possess intrinsic spin $\hbar$, can be carried further to include strong nuclear interactions as well, if these strong interactions are also mediated through quanta (gluons) carrying spin $\hbar$. The resulting theory, which appears to explain all known low-energy phenomena, is called the standard model. (It is a model based on three similarly constituted generations of quarks and leptons plus the mediating quanta W^+, W^-, Z^0, photons, and gluons plus the Higgs particle.) A complete gauge unification of all three forces (electromagnetic, weak-nuclear, and strong-nuclear) into a single electronuclear interaction seems plausible. Such a (so-called grand) unification necessarily means that the distinction between quarks on the one hand and neutrinos, electrons, and muons (leptons) on the other, must disappear at sufficiently high energies, with all interactions (weak, electromagnetic, and strong) clearly manifesting themselves then as facets of one universal gauge force with a primitive universal strength equal to $\alpha/\sin^2\theta_w$. The fact that at low energies presently available, these interactions exhibit vastly different effective strengths is ascribed to differing renormalizations due to

successive spontaneous symmetry breakings. A startling consequence of the eventual universality and the disappearance of distinction between quarks and leptons is the possibility, first discussed by J. C. Pati and Salam within their electronuclear model, of protons transforming into leptons and pions. Contrary to the older view, protons would therefore decay into leptons and pions and not live forever. A somewhat different model elaborated a year later by H. Georgi and Glashow (called the grand unifying theory), predicts a lifetime of the order of 10^{29} years for the proton P, with decay principally through the mode $P \rightarrow e^+ + \pi^0$, where π^0 is the neutral pion and e^+ is the positron. Experiments carried out during 1983 to search for this mode of proton decay gave negative evidence for protons decaying at this rate, although other types of decay modes may have been observed in other experiments. The discovery of proton instability (with decays into leptons or antileptons) would be an epic discovery and a direct confirmation of the electronuclear (grand) unification. *See* GRAND UNIFICATION THEORIES; PROTON.

Consequences of symmetry breaking. Spontaneous symmetry breaking of gauge interactions has the characteristic that symmetry breaking is a phase phenomenon and disappears in a high-temperature environment. This implies that, at temperatures T in excess of 10^{15}K (T greater than $m_z c^2/k$, where k is the Boltzmann constant), that is, up until 10^{-12} s after the outset of the big bang, there was no spontaneous breaking of the symmetry of electroweak interactions, and the W and the Z particles were massless, like the photons and the gluons. The onset of such phase transitions plays a crucial role in modern cosmological theories of the early universe, resolving some old dilemmas. For example, proton decay, and left-right and particle-antiparticle symmetry violations, provide a natural explanation for the fact that the present universe contains a preponderance of protons and neutrons rather than of their antiparticles. However, the existence of such phase transitions also poses some new dilemmas, such as the prediction of the existence of heavy magnetic monopoles (in the early universe), with abundances surviving into the present epoch, for which there is no experimental evidence. To remedy this, it is necessary to postulate an inflationary epoch having occurred in the universe's history somewhere about 10^{-33} s after the onset of the big bang. *See* ANTIMATTER; BIG BANG THEORY; COSMOLOGY; INFLATIONARY UNIVERSE COSMOLOGY; MAGNETIC MONOPOLES; PHASE TRANSITIONS; UNIVERSE.

Prospects for including gravity. Research in unification theories of fundamental interactions is now concerned with uniting the gauge theories of gravity and of the electronuclear interactions. One promising approach is the extension of space-time to more than four dimensions, following ideas developed by T. Kaluza and O. Klein in the 1920s. Remarkably, the formal expression for Einstein's gravitational interaction in a space-time of dimensions higher than four, is equivalent to the standard Einstein theory of spin-$2\hbar$ gravitons in four dimensions plus a Yang-Mills theory of spin-$\hbar$ particles (that is, a theory describing the electronuclear type of gauge interactions) when the extra dimensions are contracted down to less than 10^{-35} m. No realistic model of such a compactified unified theory has emerged, though Einstein-like supersymmetric theories in 10-space and 1-time (a total of 11 dimensions) are the favored candidates. (Supersymmetry is the principle which treats gauge and Higgs particles on a par with quarks and leptons.) *See* SUPERGRAVITY; SUPERSYMMETRY.

The most promising approach appears to be that of superstring theories. Such theories appear to describe the only possible theory of gravity which is finite and suffers from no ultraviolet infinities. A closed string is a (one-dimensional) loop which may exist in a d-dimensional space-time (where d must equal 10 to completely eliminate all ultraviolet infinities). The quantum oscillations of the string correspond to particles of higher spins and higher masses, which may be strung on a linear trajectory in a spin-versus-mass^2 (Regge) plot. Among these are the zero-mass gravitons and the gauge mesons. The theory has a unique built-in gauge symmetry.

So far, it has not been possible to go down from $d = 10$ dimensions to $d = 4$ conventional space-time dimensions and to produce the emergence of the standard model, although there is hope that this may be accomplished. If these string ideas are successful, they may help lead to one single theory which unites all known low-energy phenomena. *See* SUPERSTRING THEORY.

Abdus Salam

Bibliography. R. A. Carrigan, Jr., and W. P. Turner, *Fundamental Particles and Forces*, 1989; S. Glashow, A. Salam, and S. Weinberg, Nobel Lectures in Physics, *Rev. Mod. Phys.*, 52:515–543, 1980; G. Kane, *Modern Elementary Particle Physics: The Fundamental Particles and Forces*, 1993; C. Quigg, Elementary particles and forces, *Sci. Amer.*, 252(4): 84–95, April 1985; C. Rubbia, Experimental observation of the intermediate vector bosons W^+, W^- and Z^0, Nobel Lecture, December 8, 1984, *Rev. Mod. Phys.*, vol. 57, no. 3, part 1, pp. 699–722, 1985.

Fungal biotechnology

All aspects of cultivating fungi together with products and processes derived from such cultures. Fungi, including yeasts, are an extremely diverse group of heterotrophic organisms with an enormous impact, both beneficial and detrimental, on the environment, on animal and plant life, and on the global economy and human social development. Even though they are unable to synthesize their own food, they exhibit a wide range of biosynthetic and biodegradative activities. Since fungi can bring about chemical change in almost any natural or synthetic organic molecule, many species have been selected and propagated in pure culture specifically for applications in biotechnology and industry, for example in food and beverage production.

While the fermentation industry remains the largest and economically most important user of fungal cultures, fungi are also utilized by the pharmaceutical, cosmetic, chemical, agricultural, food, enzyme, wood product, and waste treatment industries. In the United States, two prominent examples of fungal products are citric acid, with an annual production of 350,000 tons (160,000 kg), and beta-lactam antibiotics. Yeasts are the most commercially exploited microorganisms. They have been used extensively in baking, brewing, winemaking, and distilling, and in making various metabolic products. Several million tons of fresh yeast are produced each year, mostly for the baking industry. *See* FERMENTATION; FOOD FERMENTATION.

Fungi stand at the forefront of biotechnology. The technology for large-scale fermentation and downstream processing is well understood. Newer techniques, such as protoplast fusion and genetic engineering, along with immobilized molecules and powerful fermentation processes, will lead to further utilization of the biosynthetic potential of fungi. Moreover, any fungus that can be grown in culture is a prospective host for the production of economically important, biologically active proteins native to higher forms.

Alcohols. The production of alcohols by fungi through the fermentation of sugar is the oldest industrial fermentation process. Of the hundreds of known metabolic products of fungi, ethanol is perhaps the most widely used. Ethanol is an ideal fuel and the basis for countless chemical transformations. It serves as a solvent for different materials and as the raw material from which other compounds are synthesized. As a solvent, it is employed in the manufacture of such diverse materials as dyes, drugs, soaps, plastics, resins, and polishes; and as a raw material, in the synthesis of compounds such as ether, esters, acetic acid, hypnotics, and synthetic rubber.

Ethanol (ethyl alcohol) from the yeast *Saccharomyces cerevisiae* is the largest fungal fermentation product in both volume and value. It is used primarily in the production of alcoholic beverages such as beer, wine, distilled spirits, cider, sake, and liqueurs. Fuel ethanol produced by yeast fermentation of sugarcane juice and molasses is used as an octane enhancer and replacer in fuel (gasoline) for internal combustion engines in the majority of automobiles in Brazil. Other industrially useful alcohols produced by yeast fermentation are higher alcohols, and polyhydric alcohols such as glycerol, xylitol, sorbitol, arabinitol, erythritol, and mannitol. Glycerol is widely used in the production of synthetic resins, pharmaceuticals, cosmetics, and toothpastes. However, the current production of glycerol by fermenation is insignificant compared with chemically produced glycerol. Xylitol from *Candida* species finds increasing uses in the food industry as a sweetener and as a sugar substitute for insulin-dependent diabetics. Arabinitol from *Candida entomaea* is used as a low-calorie bulking agent, in biosurfactants, as an edible coating, and as a carrier for pharmaceutical agents. *See* ALCOHOL; ALCOHOL FUEL; ETHYL ALCOHOL; FOOD MANUFACTURING.

Amino acids. Amino acids are the building blocks of proteins. While bacteria are more likely to be considered as amino acid–producing organisms because of their rapid growth rate, fungi represent the simplest eukaryotic model system that possesses the full capacity of amino acid synthesis. Fungi are actually superior to bacteria in their ability to grow on a variety of polysaccharides, wood wastes, or other cheap materials. Their potential to produce amino acids may become even more important if such carbon sources are used.

Amino acids from fungi are used as food supplements and precursors for products. So far, however, only tryptophan, lysine, and methionine have been produced commercially. Phenylalanine is used in the synthesis of the artificial sweetener aspartame. L-Lysine is known to be produced by *Apiotrichum humicola*, *Aspergillus ustus*, *Candida lipolytica*, *Candida pelliculosa*, *Cryptococccus humicolus*, *Cryptococccus laurentii*, *Kloeckera africana*, *Sphacelotheca sorgii*, *Trichosporon beigelii*, and *Ustilago maydis*; L-methionine by *Aspergillus oryzae*, *Candida lipolytica*, *Candida sake*, and *Hanseniaspora valbyensis*; L-phenylalanine by *Rhodotorula glutinis*; and L-tryptophan by *Candida sake*, *Candida zeylanoides*, *Claviceps fusiformis*, *Claviceps purpurea*, *Hansenula anomala*, *Pichia anomala*, and *Pichia pastoris*. *See* AMINO ACIDS.

Enzymes. Fungal sources currently provide the largest assortment of bulk commercial enzymes with the widest variety of applications. In some processes the crude enzymes secreted by cells to act upon their substrates in an external environment can be exploited without the preparation of refined enzyme extracts.

Fungal enzymes range from bulk technical, detergent, and food grade (typically proteases and carbohydrate-hydrolyzing enzymes available at low cost and in less than two dozen general types) to the specialty enzymes used for research or analytical needs. The latter are available in many varieties and forms and are high in purity as well as cost.

The food industry alone consumes more than half the world's output of industrial enzymes. The majority of food enzymes are hydrolases, such as proteases from *Aspergillus oryzae* for bread-making and cheese-making, pectinases from *Penicillium* and *Aspergillus* for fruit processing, and amylases from *Aspergillus* for the brewing industry. *Trichoderma viride* and *Aspergillus niger* are commercial sources of cellulases for treating various feedstuffs and food modifications. *See* CHEESE; ENZYME; MALT BEVERAGE.

For the manufacture of paper pulp, lignin can be removed from wood without destroying the cellulose and hemicellulose by the lignin-degrading enzymes of certain white-rot fungi, such as *Phanerochaete chrysosporium*. (Brown-rot fungi utilize only cellulose and cannot degrade lignin.) Biological delignification has the advantages of low

capital cost, low energy input, greater specificity, and high yield without generating polluting by-products. In addition, nonspecific lignases of *Pichia chrysosporium* can degrade a variety of phenol-containing compounds, such as pesticides, and wastewater discharges containing trinitrotoluene. The wastewater from industries that use molasses as a raw material contains a dark brown pigment that is not decolorized by usual biological treatments. The phenol-oxidizing enzymes of *Aspergillus oryzae*, *Pleurotus*, or *Auricularia* can decolorize such fermentation wastes. Decolorization of sugarcane molasses is also accomplished in the food industry by fungi. Among the yeasts, *Lipomyces* can degrade the herbicide paraquat, *Rhodotorula* can convert benzaldehyde to less toxic benzyl alcohol, and *Candida* can degrade formaldehyde. Species of *Pleurotus* are able to break down pentachlorophenol (PCP) and DDT. The filamentous fungi *Aspergillus niger* and *Chaetomium cupreum* degrade the tannins in tannery effluents. White-rot fungi, such as the mushroom *Coriolus versicolor*, are effective in solubilizing low-grade coal, which has a structure similar to lignin. For retting flax fibers (loosening the fibers from the woody tissue), the polysaccharide-degrading enzymes of the mushroom *Pleurotus* can be used instead of boiling with sodium hydroxide for removal of noncellulosic materials. Enzyme-treated fibers produce high-quality yarns. A lipase from *Thermomyces lanuginosus* is used to remove fatty stains such as those made by olive oil and lard from clothes. A commercial fungal cellulase is used to remove the microfibers during washing without degrading the cotton fibers, making the cotton fibers regain the smooth appearance characteristic of new fabrics and making the fabric stay soft and bright in color longer than when washed without the cellulase. *See* PAPER.

Organic acids. Fungi produce over 40 different organic acids. The compounds may accumulate as the result of carbohydrate metabolism, they may be terminal products of glycolysis, or they may result from the incomplete oxidation of sugars. However, only citric acid, itaconic acid, and gluconic acid are produced in significant quantities by large-scale commercial processes. Citric acid is widely used in the food industry for fruit drinks, confectionery, jams, and preserved fruit, and in the cosmetic industry for astringent lotions. The ability of citric acid to sequester metal ions in acidic media makes it useful in the formulation of shampoos and electrolyte solutions, in the tanning of animal skins, and even for the clearing of pipes in the oil industry. It is produced industrially almost exclusively by *Aspergillus niger* and *Aspergillus wentii*. Gluconic acid from *Aspergillus niger* is used in medicine as a diet supplement. Itaconic acid from *Aspergillus terreus* is used in the production of fibers, paint adhesives, and thickening agents. *See* CITRIC ACID.

Due to the acids they produce, fungi are successful bioleaching agents. For example, *Aspergillus niger*, *Scopulariopsis brevicaulis*, and *Penicillium expansum* are capable of extracting aluminum from aluminosilicates. *Penicillium* can leach out nickel and insoluble phosphates from ore. Yeasts have been used to leach zinc, copper, and lead from filter dusts from copper-process converters.

Polysaccharides. Fungi produce a number of nonprotein polymers. Almost all are polysaccharides composed of simple sugars. Industrial applications for the polymers are numerous and varied: stabilizers for such food products as syrups and ice cream; pigment suspension in paint of high viscosity; dispersion of pigments for fabric printing and dying; additives to drilling mud in the oil exploration industry; extraction of crude oil from oil-bearing shales; improvement of the viscosity of cosmetic lotions and dentifrices. They are also used as insoluble matrices for immobilizing enzymes and cells. An advantage of using fungi for the production of polymers is their ability to grow at low pH, limiting most bacterial contaminations. Thus, growth conditions for large-scale production do not have to be rigorously aseptic. *See* POLYSACCHARIDE.

Pullulan produced by *Aureobasidium pullulans* is a neutral water-soluble polysaccharide. It is colorless, odorless, tasteless, nontoxic, low-calorie, completely biodegradable, and highly soluble in water; it adheres to surfaces such as paper, wood, glass, or metal. One application is as an adhesive in the manufacture of wood-based building products. It can also be formed into films, fibers, or shaped articles.

Some of the more recently isolated substances are metabolic regulators, antitumor drugs, and immunomodulators that enhance host resistance against cancer, immunodeficiency disease, or generalized immunosuppression. Four commercially available polysaccharide preparations are currently used in treating various cancers: lentinan from *Lentinus edodes*, schizophyllan (SPG) from *Schizophyllum commune*, the polysaccharide-polypeptide PSK (krestin) or PSP from *Coriolus versicolor*, and scleroglucan from *Sclerotium rolfsii* and *Sclerotium glucanicum*. The yeasts *Hansenula*, *Pichia*, and *Pachysolen* produce extracellular phosphorylated mannans useful as thickening, stabilizing, and dispersing agents. Levans from yeasts and *Aspergillus sydowii* and *Aspergillus versicolor* can substitute for gum arabic. Chitosan extracted from *Mucor rouxii* and *Absidia coerulea* is used as a flocculating agent for clarifying sewage, in oil recovery, as a paper-sizing agent, and as a chelating agent for metal ions.

Food proteins. Fungi can grow on almost anything that contains carbohydrates, including vegetable and distillery wastes and wood-processing effluents. Because of their rapid growth rate and high protein content, fungi are an ideal source of protein for animals and humans. Modern production of single-cell protein from yeasts, such as *Candida utilis*, is making fungal protein competitive with other more traditional animal feeds. Mycoprotein, a product from filamentous fungi, is used as a meat substitute.

At present, mushroom culture represents the only major process in biotechnology that can successfully convert industrial waste, wood waste, and urban refuse into a product that is directly edible by humans. The most common mushroom grown for commercial purposes in the United States is the

white button species, *Agaricus bisporus*. In Asian countries, shiitake (*Lentinula edodes*), the padi straw mushroom (*Volvariella volvacea*), and the oyster fungus (*Pleurotus ostreatus*) are popular and profitable crops. *See* MUSHROOM.

Flavor, aroma, and coloring compounds. While fungi have long been utilized in the food industry, their potential for the production of natural flavor and color compounds has been recognized only recently. In comparison to expensive extraction from plants or chemical synthesis, microbial production of aromatic compounds offers definite advantages, particularly for consumers wishing to avoid chemical additives in food and other household products. Certain filamentous fungi have been used to produce desirable flavors in cheeses, such as *Penicillium roquefortii* and *Penicillium camembertii*, and in the manufacture of fermented foods, particularly in Asian countries. Examples of popular fermented foods are soy sauce and miso, which utilize the *Aspergillus flavus-oryzae* group of fungi, and tempeh from the action of *Rhizopus*. A variety of yeasts and filamentous fungi, such as *Penicillium notatum*, can produce sweet, fruity, coconutlike, buttery, or nutlike flavors. In addition, a wide assortment of compounds with unique and potent flavors have been isolated as naturally occurring metabolites of edible and wood-decaying mushrooms. The filamentous fungus *Monascus purpureus* has been used in the production of red color in rice wine. Both yeasts and filamentous fungi can produce orange-colored β-carotene, which is a popular, nontoxic additive to such foods as butter, ice cream, orange juice, and candies. The carotenoid astaxanthin from the red yeast *Phaffia rhodozyma* is used as a pigment in salmonid feeds in fish farms, and is efficiently absorbed in the flesh. *See* SPICE AND FLAVORING.

Vitamins and plant growth factors. Fungi can produce a variety of vitamins, including biotin, folic acid, pantothenic acid, pyridoxine, riboflavin, thiamine, and cyanocobalamin, but only the production of riboflavin has been commercialized. Riboflavin from *Acremonium diospyri*, *Ashbya gossypii*, *Candida guilliermondii*, *Eremothecium ashbyi*, *Nematospora gossypii*, and *Pachysolen tannophilus* competes well with chemical synthesis. In addition, *Aspergillus giganteus*, *Blakeslea trispora*, *Mortierella ramanniana*, and *Phycomyces blakesleanus* produce β-carotene, a precursor of vitamin A, and *Candida norvegensis* produces ascorbic acid, vitamin C. *See* VITAMIN.

Plant growth regulators are chemical compounds with the capability of coordinating the growth and morphogenesis of plants. Gibberellin is the only one produced commercially by means of fungal fermentation; *Fusarium moniliforme* and *Gibberella fujikuroi* are used. Under the right conditions, some strains of *Gibberella fujikuroi* yield only one gibberellin (gibberellic acid, or gibberellin A3), which in malting is sometimes applied to the germinating barley to increase the yield of amylases. Gibberellic acid has received much experimental attention in horticulture for the induction of flowering of plants independently of photoperiodism. Its practical use so far has been limited to a few applications, such as obtaining seedless grapes. *See* GIBBERELLIN.

Antibiotics and drugs. Antibiotics are among the most important compounds produced by fungi. The timely discovery and successful use of penicillin derived from *Penicillium chrysogenum* during World War II provided an early model to guide production of new antibiotics. Penicillin is not one chemical compound but a mixture of several, each with varying therapeutic values. Cephalosporin C, a structurally related antibiotic produced by *Acremonium*, is the parent compound for all of the cephalosporins. Fusidic acid, which was first isolated from *Fusidium coccinellum*, is an antibiotic effective against gram-positive infections, particularly penicillin-resistant staphylococci. Griseofulvin is an antifungal antibiotic produced by *Penicillium griseofulvin*. It is the only effective antibiotic available for the systemic treatment of fungal infections of skin, hair, and nails. *See* ANTIBIOTIC.

The cyclosporins from *Tolypocladium inflatum* and *Trichoderma polysporum*, first identified as antifungal antibiotics, also exhibit immunosuppressive activity. Their use has greatly reduced morbidity and increased survival rates in human transplant patients. Gliotoxin from *Aspergillus fumigatus* has antibiotic and immunosuppressive properties.

More than 40 alkaloids produced by *Claviceps purpurea* have been isolated from ergot, and hundreds of derivatives have been chemically prepared. Alkaloids have found use in the treatment of migraine headaches, in the control of uterine motor activity, and in the therapy of Parkinson's disease. Asperlicin from *Aspergillus allicaceus*, an antagonist of cholecystokinin, is used in the treatment of disorders of the gastrointestinal system that involve cholecystokinin. Lovastatin from *Aspergillus terreus* and *Monascus ruber* is a very potent inhibitor of cholesterol biosynthesis and is used for lowering cholesterol levels. Mevastatin (compactin) from *Hypomyces chrysospermus*, *Penicillium brevicompactum*, *Penicillium citrinum*, *Trichoderma longibrachiatum*, and *Trichoderma pseudokoningii* and paravastatin from *Mucor hiemalis* also lower plasma cholesterol levels. *See* ALKALOID.

Recombinant proteins. Because they are capable of secreting large quantities of certain proteins in liquid culture, fungi have proven to be useful as cloning hosts for the production of recombinant proteins of fungal and human origin. Fungi commonly used as transformation agents are *Saccharomyces cerevisiae*, *Aspergillus awamori*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, and *Pichia pastoris*. Technologies have been developed to scale up the production of new or novel products. *Aspergillus nidulans* has been designed to produce many human therapeutic proteins, including growth factors and protein hormones. The yeasts *Saccharomyces cerevisiae* and *Pichia pastoris* have been utilized for the expression of human interferon and serum albumin. Yeast chromosomes are also being employed in the mapping of the human genome. *See* GENETIC ENGINEERING.

Shung-Chang Jong

Mycopesticides. Manufactured products which contain living fungi and are used to control pests are called mycopesticides. They are utilized to control weeds, harmful insects, nematodes (roundworms), or even other fungi. Although formerly confined to experimental settings, mycopesticides are increasingly available as commercial products, especially for the agricultural market. They may impact pest populations through direct parasitism, secretion of antibiotics, competition for nutrients, or a combination of these effects, and may be used alone or in combination with chemical pesticides.

The fungi used in formulating mycopesticides are usually selected from natural populations for their ability to survive environmental conditions, control the target pest, and avoid damage to nontarget organisms. However, experimental genetic manipulation has been employed to enhance one or more desirable properties of these fungi, such as rendering the incorporated fungus resistant to a particular fungicide.

Any commercially available mycopesticide must be formulated to withstand conditions of storage, transport, and application; must be effective against the target pest; and must be thoroughly tested for safety for humans and other nontarget organisms. Mycopesticides containing genetically modified organisms are subject to additional regulatory controls. In the United States, the principal regulatory authority for pesticides is the Environmental Protection Agency. If genetically modified microorganisms are released in the field, the Animal and Plant Health Inspection Service (U.S. Department of Agriculture) has additional oversight.

Examples of fungi that have been utilized in mycopesticide production include *Colletotrichum gloeosporioides* f.sp. *aeschonomene* (against northern jointvetch), *Puccinia chondrillina* (against skeletonweed), and *Entyloma ageratinae* (against mistflower); *Metarhizium anisopliae* (against cockroaches and termites), *Beauveria bassiana* (against grasshoppers, locusts, and beetles), *Verticillium lecanii* (against whiteflies and thrips), and *Paecilomyces fumosoroseus* (against whiteflies); *Arthrobotrys superba* (against plant-parasitic nematodes); *Gliocladium virens* (against the fungi *Rhizoctonia* and *Pythium*), and *Trichoderma harzianum* (against wood rot fungi). *See* FUNGI; FUNGISTAT AND FUNGICIDE; PESTICIDE.

Frank M. Dugan

Bibliography. A. Altman (ed.), *Agricultural Biotechnology*, Marcel Dekker, New York, 1998; D. K. Arora, R. P. Enlander, and K. G. Mukerji, *Handbook of Applied Mycology*, vol. 4: *Fungal Biotechnology*, 1992; M. A. Levin and E. Israeli (eds.), *Engineered Organisms in Environmental Settings: Biotechnological and Agricultural Applications*, CRC, New York, 1996; R. G. Van Driesche and T. S. Bellows, Jr., *Biological Control*, Chapman & Hall, New York, 1996; M. Wainwright, *An Introduction to Fungal Biotechnology*, 1992; G. M. Walker, *Yeast Physiology and Biotechnology*, 1998; D. T. Wicklow and B. E. Söderström (eds.), *The Mycota*, vol. 4: *Environmental and Microbial Relationships*, Springer-Verlag, Berlin, 1997.

Fungal ecology

The subdiscipline in mycology and ecology that examines community composition and structure; responses, activities, and interactions of single species; and the functions of fungi in ecosystems. The fungi, as they have been historically classified, comprise organisms from two distinct evolutionary lineages: the true fungi, or Eumycota; and the Oomycota, which are allied with chrysophyte algae in the separate kingdom Chromista. The oomycetes resemble the true fungi in having a filamentous growth form and occurring as plant parasites or free-living saprobes (organisms living on decaying organic matter) in soil or decaying vegetation. The slime molds, or Mycetozoa, are another distinct group in the kingdom Protista. These organisms display an extraordinary diversity of ecological interactions and life history strategies, but are alike in being efficient heterotrophs. In all ecosystems, producers (generally plants and some bacteria) synthesize cell components by using solar energy to chemically reduce carbon dioxide and assemble carbohydrates into the complex polymers of cell walls. Heterotrophs consume and disassemble, obtaining energy from the catabolism of photosynthetically formed carbohydrates and other molecules. Fungi and bacteria are the primary decomposers, facilitating the flow of energy and the cycling of materials through ecosystems. Both classes of organisms occur wherever there is organic matter. With an estimated 1.5 million species worldwide, of which only about 5% have actually been named, the fungi rank just below the insects in biodiversity. *See* ECOLOGICAL ENERGETICS; FUNGI.

Habitats. Fungi occur in many different habitat types—on plant surfaces; inside plant tissues; in decaying plant foliage, bark and wood; and soil—generally changing in abundance and species composition through successional stages of decomposition. Fungi are also found in marine and aquatic habitats; in association with other fungi, lichens, bacteria, and algae; and in the digestive tracts and waste of animals. Some fungi grow in extreme environments: rock can harbor free-living endolithic fungi or the fungal mutualists of lichens; thermotolerant and thermophilic fungi can grow at temperatures above 45°C (113°F); psychrophilic fungi can grow at temperatures to below −3°C (27°F). Xerotolerant fungi are able to grow in extremely dry habitats, and osmotolerant fungi grow on subtrates with high solute concentrations. Most fungi are strict aerobes, but species of the chytrid *Neocallimastix*, which inhabit the rumen of herbivorous mammals, are obligate anaerobes. Several aquatic fungi are facultative anaerobes.

Although fungi are ubiquitous in terrestrial habitats, their abundance and diversity in these various domains is highly variable. Many fungi occur as free-living saprobes, but fungi are particularly successful as mutualistic, commensal, or antagonistic symbionts with other organisms. Fungi are ecologically and economically important as the principal pathogens of vascular plants. However, they are

also beneficial mutualistic mycorrhizal associates of roots, which aid in nutrient acquisition. Fungi are the mutualistic partners with algae or cyanobacteria in lichens, and endobiotic fungi in certain grasses produce toxins that inhibit insect herbivores. Fungal associations with nematodes, protozoa, mites, and insects also range from predation to mutualism. Fungi are cultivated and consumed as food by leaf-cutting ants, termites in the family Macrotermitinae, and wood-boring ambrosia beetles. Several groups of fungi have evolved a variety of specialized means for capturing or parasitizing nematodes and protozoa. The oomycete *Haptoglossa* pierces the body of rotifers with a harpoonlike organelle. Predacious fungi capture rotifers and nematodes by adhesive knobs and nooselike constricting hyphal rings. Commensal trichomycetes inhabit the digestive tracts of arthropods. Fungal parasitism is widespread among several groups of fungi specializing on various insects, such as *Cordyceps* on Lepidoptera, *Entomophthora muscae* on houseflies, and *Septobasidium* on scale insects. Laboulbeniales, an order with over 1800 species, are specialized insect parasites that occur on several insect families, especially beetles. Fungi also are the principal or exclusive diet of species, genera, and sometimes families of protozoa, nematodes, mites, collembola, and insects. Truffles and mushrooms are an important component of the diets of several small mammals. *See* POPULATION ECOLOGY.

Biomass. Fungal hyphae are pervasive in soil and decomposing plant litter. The lengths of hyphae in deciduous forest materials have been estimated at about 1300 ft (400 m) per gram of dry soil and up to 20 times that length in the humus or litter above. In coniferous forest systems, the comparable figures are 1–1.6 mi (1500–2500 m) per gram of dry soil [37 mi (60,000 m) in mycorrhizal mats] and about 41.5 mi (67,000 m) per gram of dry organic horizon. Fungal biomass in tropical soils may be twice that in temperate zone soils. Fungal individuals, as a contiguous, genetically distinct hyphal network (thallus), are among the largest and oldest organisms. A single fungus of the tree root colonizer *Armilliaria gallica* was estimated to span 2083 ft (635 m), to weigh over 22,000 lb (10,000 kg) and to be over 1500 years old. In contrast, genets (individuals that are genetically distinct) of parasitic or saprobic fungi may consist of only a few cells.

Significant features. Fungi possess unique features that affect their capacity to adapt and to function in ecosystems:

1. Fungi are composed of a vegetative body (hyphae or single cells) capable of rapid growth. Hyphae are linear strands composed of tubular cells that are in direct contact with the substrate. The cells secrete extracellular enzymes that degrade complex polymers, such as cellulose, into low-molecular-weight units that are then absorbed and catabolized. Many fungi also produce secondary metabolites such as mycotoxins and plant growth regulators that affect the outcomes of their interactions with other organisms.

2. Filamentous fungi are able to mechanically penetrate and permeate the substrate.

3. Fungi have an enormous capacity for metabolic variety. Fungal enzymes are able to decompose highly complex organic substances such as lignin, and to synthesize structurally diverse, biologically active secondary metabolites. Saprotrophic fungi are very versatile; some are able to grow on tree resins and even in jet fuel.

4. Structural and physiological features of fungi facilitate absorption and accumulation of mineral nutrients such as phosphorus, nitrogen, and potassium, as well as toxic elements such as arsenic, selenium, and mercury. The capacity of fungi to absorb, accumulate, and translocate is especially significant ecologically where hyphal networks permeate soil and function as a link between microhabitats. For example, mycorrhizal fungi serve as a link between soil and their plant symbionts, an interconnected network allowing movement of carbon and minerals.

5. Fungi have the capacity for indeterminate growth, longevity, resilience, and asexual reproduction. The vegetative cells of Eumycota are often multinucleate, containing dissimilar haploid nuclei. This combination of features gives the fungi an unparalleled capacity for adaptation to varying physiological and ecological circumstances and ensures a high level of genetic diversity.

6. Many species of fungi have a capacity to shift their mode of nutrition. The principal modes of nutrition in fungi are saprotrophy (the utilization of dead organic matter) and biotrophy, which is characteristic of parasitic, predacious, and mutualistic fungi (including mycorrhizae and lichen fungi). Many plant and animal parasites, and some mycorrhizal fungi, are facultative or interim saprotrophs that are able to persist in soil or decaying organic matter in the absence of a suitable host. *See* BIODEGRADATION; FUNGAL GENETICS; MYCORRHIZAE.

Communities. Communities and assemblages of fungi occur in all terrestrial and many aquatic and marine habitats. Regardless of domain, species composition and abundance are closely regulated by environment. Because fungi are heterotrophs, they are nutritionally dependent on the plants, animals, and microorganisms with which they coexist. Thus, the composition of fungal communities is strongly influenced by other biota. In the various biomes of the world, the fungal communities are as different and distinctive as are the plants and animals of those biomes. In temperate zone biomes, high species diversities for plants and microfungi are correlated, but the total species of fungi far exceeds that of plants.

As microorganisms, fungal communities can vary considerably in scale, from the surface inhabitants of a single leaf to a grassland soil. Individuals and species in fungal communities are functionally interconnected through their interactions. The ecological principles that govern distribution, abundance, and dynamics of species of plants and animals apply equally to fungal communities. Resource limitation,

competition, niche specialization, trophic interactions (predation and parasitism), and antagonistic interactions all operate and affect species distribution and abundance in fungal communities. In fungal communities, as in communities of other organisms, disturbance can have a cascading effect leading to extinctions, changes in abundance, and appearance of previously undetected species.

Functions. Fungi interact with all organisms in ecosystems, directly or indirectly, and are key components in ecosystem processes. As decomposers, fungi are crucial in the process of nutrient cycling, including carbon cycling as well as the mineralization or immobilization of other elemental constituents. As parasites, pathogens, predators, mutualists, or food sources, fungi can directly influence the species composition and population dynamics of other organisms with which they coexist. Fungi may act both as agents of successional change or as factors contributing to resilience and stability. Introduced fungal pathogens have had a profound effect on the species composition of forests worldwide during the past century. Nitrogen-fixing lichens are important components of the nitrogen cycle. Fungi also form an essential link interconnecting functional components of food webs in many habitats.

Mycorrhizal fungi function as an interface between plant and soil, and are essential to the survival of most plants in natural habitats. Due to their fine hyphae and ability to permeate the substrate, mycorrhizal fungi enable their plant symbionts to extract nutrients from soil more efficiently. Mycorrhizal fungi extract essential nutrients from the inorganic pool, which are translocated to the plant symbiont. Generally, mycorrhizal plants are more vigorous than nonmycorrhizal plants. Approximately 70–90% of the vascular plant species worldwide are mycorrhizal in their natural habitats. Most mycorrhizal fungi have limited ability to decompose organic matter and rely on their plant symbionts for carbon nutrition, although some mycorrhizal fungi accomplish both decomposition and direct nutrient transfer.

Mycorrhizal fungi also link individuals and permit complex species interactions in ecosystems. The achlorophyllous vascular plant *Monotropa* obtains nutrients by digesting mycorrhizal hyphae that in turn receive nourishment from their tree associates. In the old-growth forests of the Pacific Northwest, truffles (fungi that form belowground sporocarps) in the genus *Hysterangium* are mycorrhizal with Douglas-fir. Most truffles produce odors attractive to rodents, which dig them up, eat them, and aid in their dispersal. *Hysterangium* and other mycorrhizal truffles are an important component of the diet of the northern flying squirrel, which in turn is preyed on by the northern spotted owl. A similar interdependence is known from the southwestern United States between ponderosa pine, the tassel-eared squirrel, and the mycorrhizal truffle *Rhizopogon*. Thus, the fungi form a crucial link in complex food webs.

Root pathogens of forest trees, such as *Phellinus* and *Armillaria*, can occur as large infection centers that expand radially, spreading by root contact and killing trees on the outer perimeter of the infestation. These infection centers contribute to age structure and heterogeneity in forests. Dying trees within the infection center lead to the formation of forest gaps where different, more disease-resistant tree species invade. As the infection centers age and expand, the trees nearer the center grow larger and a typical successional development occurs. Infection centers can be several hundred meters in diameter and over 1000 years old. Fungal plant pathogens, whether native or introduced, can profoundly affect the composition of the dominant vegetation. *See* HERBIVORY.

Applied ecology. There are several economically important areas that benefit from application of the principles of fungal ecology: biotechnology, biological control, bioremediation, agriculture, forestry, and land reclamation. Plant pathologists employ the principles of fungal ecology to study disease epidemics and control fungal pathogens in agricultural systems. Fungal biological control agents are used against weeds, pathogenic fungi, and insect pests. Northern jointvetch in the southeastern United States is successfully controlled by a *Colletotrichum*; *Heterobasidion* root disease of conifers is prevented from infecting cut stumps by the fungus *Phanerochaete gigantea*; the insect pathogen *Entomophaga maimaga* shows promise for control of the gypsy moth in the northeastern United States. Many fungi that attack insects are host-specific and thus can be used against specific pests without harming nontarget species.

Inoculation of tree seedlings for reforestation and landscape use with mycorrhizal fungi greatly improves their early establishment and survival. The increasing use of mycorrhizal fungi in agricultural production helps increase crop yields with less fertilization. The mycorrhizal fungus *Pisolithus* greatly improves survival of pines planted for reclamation of mine spoils. The ability of fungi to degrade complex substrates and tolerate high levels of toxic elements has led to their utilization for bioremediation of contaminated soil.

The observation that fungi produce substances that inhibit other microorganisms has led to discoveries of numerous antibiotics and other important pharmaceuticals, such as penicillin and cephalosporin from *Penicillium* and *Emericellopsis*, and cyclosporin from *Tolypocladium*. Mycotoxins such as aflatoxin in foodstuff are a human health concern. Fungal enzymes are used in numerous industrial processes; citric, itaconic, gallic, and gluconic acids are produced by *Aspergillus*; and several foods are produced through fungal fermentations. With only a small fraction of the total species known, the fungi offer a rich potential for bioprospecting, the search for novel genetic resources with unique, useful biochemical properties. *See* ECOLOGY; ECOSYSTEM; FUNGI; FUNGAL BIOTECHNOLOGY.

Martha Christensen; Jeffrey K. Stone

Ruminal gastrointestinal tract. A system of four chambers makes up the foregut of the ruminants. The rumen, reticulum, and omasum are pregastric

chambers that are formed from modifications of the esophagus and the stomach. The rumen and reticulum together form a large fermentation vessel that is approximately 10 liters (3 gal) in sheep and 100–150 liters (26–40 gal) in cows. The rumen, which is the largest pregastric chamber, is a highly specialized environment. It contains heterogeneous plant material and microorganisms in a fluid amorphous and dynamic condition similar to a continuous culture chamber. Stratification may occur in the rumen as a result of particle size and specific gravity; a gas space exists above the partially digested feed. The pH of the rumen fluctuates between 5.8 and 6.8, depending on the frequency of feeding and feed type. However, a copious amount of saliva, which constantly enters with the feed, contains bicarbonate and phosphate buffers to maintain the pH. As a result of the animal's metabolism and heat from microbial fermentation, the temperature of the rumen is maintained between 38 and 40°C (100 and 104°F), depending on frequency of feed or water intake. The composition of the gases in the rumen is approximately 65% carbon dioxide (CO_2), 35% methane (CH_4), 7% nitrogen (N_2), 0.6% oxygen (O_2), 0.2% hydrogen (H_2), and 0.1% hydrogen sulfide (H_2S). The redox potential of the rumen ranges from −250 to −450 mV because facultative anaerobic microbes such as bacteria and yeast quickly utilize air that accompanies the feed into the rumen. *See* ANIMAL FEEDS.

Bacteria, protozoa, and fungi are the three major groups of microorganisms that inhabit the gastrointestinal tract. These microorganisms are suspended in the liquid phase, or attached to plant fragments, or attached to the lining of the rumen epithelium. Of the three groups of microbes, the fungi were most recently discovered. Three types of fungi occur in the gastrointestinal tract of herbivores. The first group includes facultatively anaerobic and aerobic fungi that are described as transient because they continually enter the rumen in feed and may not be able to grow under the anaerobic conditions of the rumen environment. The second group consists of two fungal species that parasitize ciliate protozoa. The third group consists of the obligately anaerobic zoosporic fungi that are saprophytic on plant material.

Interactions of fungi and microorganisms. Due to the complex nature of the rumen ecosystem, interactions between the different types of microorganisms occur. Interactions are known to occur between anaerobic fungi and methane-producing (methanogenic) bacteria, between anaerobic fungi and nonmethanogenic bacteria, and between anaerobic fungi and protozoa.

Methanogens. Anaerobic rumen fungi have synergistic relationships with methanogens. Cultures of an anaerobic fungus, *Neocallimastix frontalis*, degrade filter paper more extensively in the presence of methanogens than in their absence. Fermentation shifts from the formation of electron sink products such as lactate and hydrogen toward more reduced end products such as carbon dioxide and methane and acetate in the coculture. *Neocallimastix frontalis* alone causes 53% dry-matter loss in vitro, while cocultures of the fungus with *Methanosarcina barkeri* and *Methanobrevibacter* sp. lead to losses of 69% and 87%, respectively. A triculture of all three organisms results in a 98% loss of filter-paper dry weight, with the end products being carbon dioxide and methane.

Methanogens attach to the surfaces of the zoospore or thalli of the fungi when grown in coculture with fungi. Hydrogenosomes, enzymes involved in the production of hydrogen and carbon dioxide in anaerobic fungi, are located near the surface of the plasma membrane. It has been suggested that methanogens are able to readily utilize the products of the hydrogenosomes for growth. When *Piromyces* sp. and *Caecomyces* sp. are grown in cocultures with methanogens, there is an increase in acetate production. *See* METHANOGENESIS (BACTERIA).

Nonmethanogenic bacteria. Interactions between nonmethanogenic bacteria and anaerobic zoosporic fungi may be synergistic, symbiotic, or competitive. An example of a synergistic relationship is the growth of *Bacteroides ruminicola* in coculture with *N. frontalis* because degradation of xylan in the coculture is more than threefold greater than in monocultures of either the bacterium or fungus. *Selenomonas ruminantium*, a bacterium that consumes hydrogen, succinate, and lactate, increases cellulose digestion by rumen fungi. *Streptococcus bovis* and *Lachnospira multiparus* suppress xylan utilization by *N. frontalis*.

Protozoa. A two- to fourfold increase in zoospores and zoosporangia of anaerobic fungi in sheep is observed when protozoa are removed. Defaunation may be influenced by an animal's diet, since greater populations of zoospores are found in animals fed high-fiber diets than in those fed low-fiber or high-concentrate diets. The interaction between anaerobic fungi and rumen protozoa may be complementary rather than competitive in nature. *See* MICROBIAL ECOLOGY.

Daniel A. Wubah

Bibliography. L. R. Batra, *Insect-Fungus Symbiosis*, John Wiley, New York, 1979; G. C. Carroll and D. T. Wicklow, *The Fungal Community*, 2d ed., Marcel Dekker, New York, 1992; M. Christensen, A view of fungal ecology, *Mycologia*, 81:1–19, 1989; R. C. Cooke and A. D. M. Rayner, *The Ecology of Saprotrophic Fungi*, Cambridge University Press, 1984; N. J. Dix and J. Webster, *Fungal Ecology*, Chapman and Hall, London, 1995; S. Isaac, *Fungal Plant Interactions*, Chapman and Hall, London, 1992; D. A. Wubah, D. E. Akin, and W. S. Borneman, Biology, fiber degradation and enzymology of the anaerobic zoosporic fungi, *Crit. Rev. Microbiol.*, 19:99–115, 1993.

Fungal genetics

The study of gene structure and function in fungi. Genetic research has provided important knowledge about genes, heredity, genetic mechanisms, metabolism, physiology, and development in fungi, and in higher organisms in general, because in

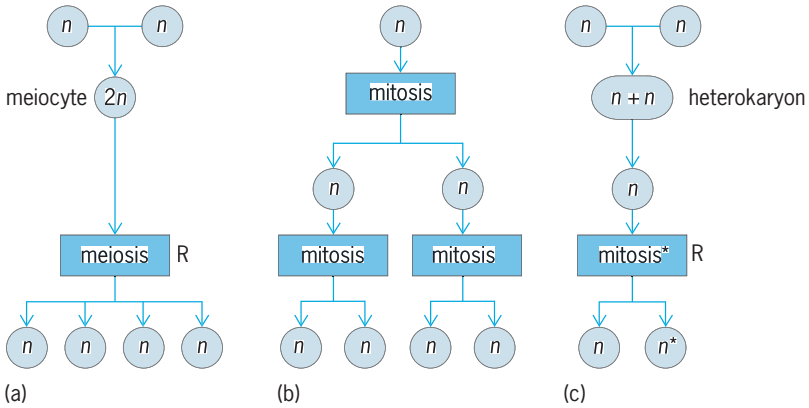

Fig. 1. Three different kinds of reproduction occurring in fungi, each of which provides opportunities for genetic analysis. (*a*) Sexual reproduction leads to recombination (R) of genes at meiosis. (*b*) Asexual reproduction (shown here in a typical haploid fungal cell) usually reproduces the gene set faithfully. (*c*) Parasexual reproduction derives from an atypical mitotic division of an unstable cell that produces haploid cells and other aneuploid unstable intermediates.

certain respects the fungal life cycle and cellular attributes are ideally suited to both mendelian and molecular genetic analysis. Fungal nuclei are predominantly haploid; that is, they contain only one set of chromosomes. This characteristic is useful in the study of mutations, which are usually recessive and therefore masked in diploid organisms. Mutational dissection is an important technique for the study of biological processes, and the use of haploid organisms conveniently allows for the immediate expression of mutant genes.

Reproduction in fungi can be asexual, sexual, or parasexual (**Fig. 1**). Asexual reproduction involves mitotic nuclear division during the growth of hyphae, cell division, or the production of asexual spores. Sexual reproduction is based on meiotic nuclear divisions fairly typical of eukaryotes in general. In ascomycetes and basidiomycetes, the spores, containing nuclei that are the four products of a single meiosis, remain together in a group called a tetrad. The isolation and testing of the phenotypes of cultures arising from the members of a tetrad (tetrad analysis) permit the study of the genetic events occurring in individual meioses; this possibility is offered by virtually no other eukaryotic group. In other groups, genetic analysis is limited to products recovered randomly from different meioses. Since a great deal of genetic analysis is based on meiosis, fungal tetrads have proved to be pivotal in shaping current ideas on this key process of eukaryotic biology. *See* EUKARYOTAE; MEIOSIS.

Because their preparation in large numbers is simple, fungal cells are useful in the study of rare events (such as mutations and recombinations) with frequencies as little as one in a million or less. In such cases, selective procedures must be used to identify cells derived from the rare events. The concepts and techniques of fungal asexual and parasexual genetics have been applied to the genetic manipulation of cultured cells of higher eukaryotes such as humans and green plants. However, the techniques remain much easier to perform with fungi.

Heterokaryons. In many fungi, notably the Fungi Imperfecti, there is no true sexual cycle based on meiosis. However, such fungi utilize a parasexual cycle which incorporates a unique reductive division more similar to mitosis than meiosis. The parasexual cycle uses the natural tendency of fungal cells to fuse, thus promoting nuclear and cytoplasmic mixing. The resulting mixed strains are called heterokaryons. Although the different nuclear types in a heterokaryon remain separate and haploid, rare nuclear fusions occur to create a transient diploid nucleus containing a chromosome set from each original strain. These diploid nuclei are unstable and promptly undergo atypical mitotic divisions leading to haploids and other unstable intermediate aneuploids. The derived haploid cells are commonly genetic recombinants. *See* DEUTEROMYCOTA; RECOMBINATION (GENETICS).

Heterokaryons are useful in other ways. They provide a convenient method of testing whether two recessive mutations are alleles by combining the two mutations in a heterokaryon. If the heterokaryon maintains the mutant phenotype of the two individual strains, the mutations can be assigned to the same gene; that is, they are alleles. Heterokaryons also provide a test for cytoplasmic inheritance. If heterokaryosis demonstrates that a novel phenotype is transmitted by cytoplasmic contact alone, then that phenotype must be based on cytoplasmic genes, probably those in the mitchondria. *See* ALLELE.

Genetics of mating type. The sexual cycle of heterothallic fungi requires fertilization by strains of different mating types. Generally, no outward difference is evident between these mating types. In bipolar species, mating type is determined by the alleles of one gene. For example, in *Saccharomyces cerevisiae* there are two alleles of the mating-type locus MAT, called a and α, whereas in the filamentous ascomycete *Neurospora crassa*, the alleles of the mt locus are called A and a. In tetrapolar species, such as many basidiomycetes, mating type is determined by two separate loci, A and B, and multiple alleles are found in nature. Sexual development occurs only if the mates differ for each locus: $A_1B_1 \times A_2B_2$ or $A_1B_2 \times A_2B_1$. $A_5B_1 \times A_1B_3$ are sexually compatible, but $A_5B_1 \times A_5B_3$ or $A_5B_3 \times A_1B_3$ are not.

In *Schizophyllum commune*, approximately 300 A and 100 B alleles are known. Recently the A loci were shown to contain pairs of genes encoding homeodomain proteins. The proteins derived from different A alleles form complexes within the fusion cells. These complexes are thought to bind to and regulate expression of genes essential for sexual development. The B loci have been shown to encode pheromones and serpentine transmembrane receptors. The types of molecules produced from both A and B loci are common types of regulators found in eukaryotic organisms. *See* BASIDIOMYCOTA; YEAST.

Mapping of genes and centromeres. In some ascomycetes, such as *Neurospora crassa*, the meiotic products in the tetrad (or octet, as in some fungi the postmeiotic nuclei soon undergo a mitosis) occur in a linear array within the ascus (*AAAAaaaa* or

AAaaAAaa). Thus, both the position of the meiotic products and the phenotypes which they display can be considered together. The frequencies of various phenotypes (*AB*, *Ab*, *aB*, or *ab*) and their positions within the octet allow estimates of the relative distance between two genes or between a gene and the contromere. *See* GENETIC MAPPING; MENDELISM.

Fungi lend themselves to gene mapping in two additional ways. Because fungal chromosomes are small, they may be separated by size when suspended in a gel matrix and then subjected to an electric field. Once the chromosomes have been separated by this pulsed-field gel electrophoresis, they may be visualized by staining, and subsequently transferred to a membrane by blotting. When this membrane is bathed with a solution containing a deoxyribonucleic acid (DNA) probe, made from a single, previously isolated gene, the probe binds to the chromosomes at the site of that gene. If the probe had been made radioactive, a photographic image of the blot can be visualized. In this way, specific genes may be mapped to particular chromosomes. Another way to map (and discover genes) is to determine the entire DNA sequence of the organism. Because fungi commonly have small genomes in comparison to other eukaryotes, it is feasible to sequence the entire genome. In fact, the first eukaryotic organism whose genome was completely sequenced was the baker's yeast, *S. cerevisiae.* The results demonstrated that yeast contains many genes that are also common to other eukaryotes, such as humans, but that the function of many genes remains unknown. This type of analysis has given rise to the new field of genomics, the study of genes and their functions.

Gene conversion and crossing-over. Within tetrads, occasional exceptions to Mendel's law of equal segregation of genes are seen; these rare exceptions have provided insight into the mechanism of crossing-over. The major exceptional types are *AAAAAaaa* (5:3 ratio) and *AAAAAAaa* (6:2 ratio). Both are examples of gene conversion, in which one or two copies of one allele have converted to the other allele, thus skewing in the 4:4 ratio. The regular observation of these types resulted in the heteroduplex theory of crossing-over, the prevailing theory for most organisms (**Fig. 2**). *See* CROSSING-OVER (GENETICS).

Mitochondrial inheritance. Inheritance determined by mitochondria may be suggested by uniparental inheritance or the heterokaryon test. The simplest type of uniparental inheritance is maternal, and one of the first examples was shown by the slow-growing phenotype, poky, in *Neurospora*. When a cross was made between a poky female and a normal male, all the progeny were poky, whereas in the reciprocal cross between a normal female and poky male, all the progeny were normal. The term female is used here, as commonly applied in fungal genetics, to identify the mate which contributes nearly all of the cytoplasm to the mating. Thus, poky was determined to be a cytoplasmically determined trait. Poky results from a mutation to the small circular DNA molecule carried within the mitochondrion.

Study of fungal mitochondrial DNA has led to an understanding of the structure and function of these highly specialized organelles. Some surprises have resulted, including the discovery of a novel mitochondrial genetic code and a unique splicing mechanism of group I introns. *See* MITOCHONDRIA.

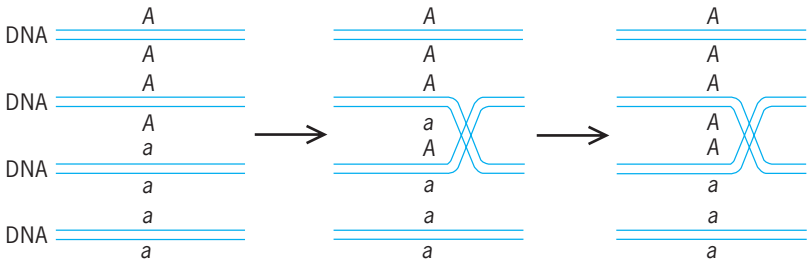

Fig. 2. Genetic crossing-over at the molecular level. As part of the process, heteroduplex DNA is formed (center). If the heteroduplex is corrected, as shown here from *a* to *A*, gene conversion is explained. Either or both heteroduplexes can be converted to *A* or to *a*, and hence a variety of aberrant ascus ratios are obtained; a 5:3 ratio is used here.

Genes and enzymes. The fact that each enzyme is coded by its own specific gene was first recognized in fungi and was of paramount importance because it showed how the many chemical reactions that take place in a living cell could be controlled by the genetic apparatus. The discovery arose from a biochemical study of nutritional mutants in *Neurospora*. The mutants were selected on the basis of their need for arginine in the growth medium (normal strains make their own arginine). Some mutants (type 1, for example) had an absolute requirement for arginine, whereas others (type 2) could also respond to a related compound, citrulline. A third type (3) responded to either citrulline or another related compound, ornithine. Those results can be explained if the mutations are in genes that code for enzymes that catalyze the sequential conversion of intermediates, including ornithine and citrulline, into arginine, as in the reaction below.

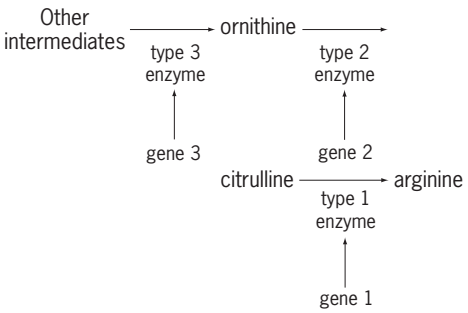

The enzyme deficiencies could be circumvented only by adding compounds normally synthesized after the block in each case. *See* ENZYME.

Genetic transformation. In genetically transformed organisms, the genome has been modified by the addition of DNA, a key technique in genetic engineering. Fungi were the first cells in which this key technique was shown to be feasible in eukaryotic organisms. In fungi, the cell wall is temporarily removed; exogenous DNA is then taken up by cells and the cell wall is restored. The incorporation of DNA must be detected by a suitable novel genetic marker

included on the assimilated molecule in order to distinguish transformed from nontransformed cells. The fate of the DNA inside the cell depends largely on the nature of the vector or carrier. Some vectors can insert randomly throughout the genome. Others can be directed to specific sites, either inactivating a gene for some purpose or replacing a resident gene with an engineered version present on the vector. A third kind of vector remains uninserted as an autonomously replicating plasmid. The ability to transform fungal cells has contributed greatly to understanding the genetics, physiology, and development of fungi. Furthermore, the technique has permitted the engineering of fungi with modified metabolic properties for making products of utility in industry. *See* GENETIC ENGINEERING.

Fungal transposons. A surprising development in the molecular biology of eukaryotes was the discovery of transposons, pieces of DNA that can move to new locations in the chromosomes. Although transposons were once known only in bacteria, they are now recognized in many eukaryotes. The transposons found in fungi mobilize by either of two processes: one type via a ribonucleic acid (RNA) intermediate that is subsequently reverse-transcribed to DNA, and the other type via DNA directly. In either case, a DNA copy of the transposed segment is inserted into the new site and may contain, in addition to the transposon itself, segments of contagious DNA mobilized from the original chromosomal site. Because of the rearrangements which transposons may produce, they have been important in the evolution of the eukaryotic genome. *See* FUNGI; GENETICS; TRANSPOSONS. A. J. F. Griffiths; R. Ullrich

Bibliography. J. W. Bennett and L. L. Lasure (eds.), *Gene Manipulations in Fungi*, 1985; J. W. Bennett and L. L. Lasure (eds.), *More Gene Manipulations in Fungi*, 1991; J. H. Burnett, *Mycogenetics*, 1975; J. R. S. Finchman, P. R. Day, and A. Radford, *Fungal Genetics*, 4th ed., 1979.

Fungal genomics

The study of the entire genetic content of organisms. It is a new discipline, developed out of advances in automated nucleic acid sequencing and computer technology. In the case of fungal genomics, the objective of the deoxyribonucleic acid (DNA) sequencing and analysis is the determination of the genomes of species in the kingdom Fungi. Fungal genomics seeks to establish the set of genes that makes a fungus specifically a fungus, as well as to determine what these species have in common with other organisms. Current genomics projects try to determine the entire genetic makeup of a given species by sequencing many overlapping fragments of DNA (at least 100,000 for a fungus), piecing these fragments together by computer, looking at the consensus sequence to identify genes and other regions of interest, and attempting to work out what these genes might do. Ultimately, this information is integrated into models of transcription (transcriptomics), protein expression (proteomics), and metabolic pathways (metabolomics). Sometimes the transcription, protein expression, and metabolic modeling stages of a genome project are called functional genomics. The deciphering of the resultant information has engaged computer scientists, mathematicians, and statisticians as well as biologists and chemists in a new discipline that is called bioinformatics. *See* DEOXYRIBONUCLEIC ACID (DNA); NUCLEIC ACID.

Fungal kingdom. The fungi are eukaryotic organisms that were once classified as plants but are now usually placed in their own kingdom. Four divisions (phyla) are distinguished within the fungal kingdom: Chytridiomycota, Zygomycota, Ascomycota, and Basidiomycota. In addition, fungi lacking a known sexual cycle are classified as Deuteromycota (Fungi Imperfecti). Taxonomists have reclassified the Oomycota, once called the lower fungi, into the kingdom Stramenopila. These so-called lower fungi include well-known species such as the causative agent of potato blight (*Phytophthora infestans*) that historically have been studied by mycologists. Out of tradition, the study of oomycete genomics usually is included in discussions of fungal genomics.

All true fungi are characterized by spore formation, an absorptive (osmotrophic) mode of nutrition and, with the exception of the chytrids, the absence of a motile stage in their life cycles. The best-known fungi are macroscopic organisms easily recognized by their distinctive fruiting bodies, for example, mushrooms, shelf fungi, and truffles. The majority of fungal species, however, are microscopic. Their common names—blight, mold, rot, and smut—reflect their capacity for causing both biodeterioration and disease. Some fungi are single-celled, of which the yeasts are the best known. The budding yeast, *Saccharomyces cerevisiae*, and the fission yeast, *Schizosaccharomyces pombe*, have been the targets of extensive genome projects. However, most macroscopic and microscopic fungi are filamentous, composed of threadlike cells called hyphae. The hypha is a distinctive tubular structure that endows filamentous fungi with cellular and genetic dynamics not found in other eukaryotic kingdoms. *See* FUNGI.

Economic importance. Fungi have enormous economic impact. They not only are major recyclers of organic materials in ecological cycles but also are important plant and animal pathogens. All fungi secrete degradative enzymes and organic acids into the environment. Many also make biologically active small molecules, some of which are toxic (for example, mushroom poisons, mycotoxins). Others have pharmaceutical value and are produced by industrial-scale fermentations for use in the manufacture of drugs such as penicillin, cephalosporin, and cyclosporin. Fungi also play a vital role in food fermentation (bread, wine, certain cheeses, soy sauce). In general, laboratory models or economically important fungi have been selected for fungal genome projects.

Genomic analysis. A typical genome project seeks to determine the complete DNA structure of a given organism and to identify and map all of its genes.

Shotgun phase. Today most sequencing centers use some version of the "shotgun" method, in which the genomic DNA of an organism is fragmented into small pieces, usually using high-frequency sound waves (sonication), and cloned to make a genomic library. The DNA from random clones is isolated and its sequence determined. Many steps in this process are automated, involving robots and expensive machines called DNA sequencers. Assembly involves putting together these random sequenced stretches of DNA into contigs (groups of cloned DNA representing overlapping segments of a particular chromosomal region) by using computer programs that compare individual sequences and find overlaps. Then genome assemblers put the fragments together into larger pieces. For large genomes, some of the most powerful computer arrays in the world are required for these computationally intense tasks. During this random or whole-genome shotgun phase, the idea is to sequence every bit of the genome at least six to ten times. However, some genome regions cannot be sequenced for various technical reasons. Often because of time and money constraints, genome-sequencing projects do not go beyond the whole-genome shotgun phase.

Finishing phase. The final "finishing" phase requires human input to join together the contigs. This process involves generating additional sequences to cover gaps and sequencing regions of the genome that were not cloned into the genomic libraries in the first place. The assembly and finishing steps work best when genetic maps, good karyotypes (organized arrays of chromosome pairs from a single cell), and/or reference genomes are available. For example, assembled sequences may be positioned to chromosomes using in-situ hybridization techniques, which enable identification of particular nucleic acid sequences while they remain in their original location in the nucleus. The last, labor-intensive stages of producing a finished sequence consume a disproportionate amount of time and money compared with the earlier, highly automated phases. For example, a draft of the human genome was announced early in 2000, but the complete version was not released until April 2003. *See* GENETIC MAPPING.

Annotation. Vast amounts of data are generated through automated sequencing. The ultimate goal is to associate DNA sequences with their functional roles in the cell. The most basic function-prediction method is the automated annotation of sequence. Programs such as Artemis, Genefinder, Genmaker, Glimmer M, and Grail are used to predict open reading frames (that is, genes are found using uninterrupted series of DNA triplets coding for amino acids without any stop codons), splice junctions (where coding and noncoding DNA meet), and other informational DNA. Once extracted, predicted genes can be subjected to sequence similarity (homology) searches to find homologs in nucleic acid or protein sequence data, usually using a member of the Basic Local Alignment Search Tool (BLAST) family of programs. Defining protein function is complicated because each gene can encode several distinct proteins. In general, when new protein coding regions are annotated, about half the genes are completely unknown. *See* GENETIC CODE; NUCLEIC ACID.

Postgenomics. Even annotated sequence information is not useful unless it leads to experimental analyses. After whole genome sequences are available, attention shifts from genome content to protein functions and systems biology. Linking unknown genes to the corresponding biologic function is variously called postgenomics or functional genomics. Genome data can be used to produce DNA microarrays, in which thousands of gene segments are deposited on a glass slide or chip and then used to simultaneously analyze the expression levels of the corresponding genes. For example, fungal molecules of ribonucleic acid (RNA) can be isolated from different time points in the sporulation and sexual cycle, reverse-transcribed to form copy DNA (cDNA), labeled with fluorescent chromophores, and the cDNAs tested against genomic microarrays. These data then are analyzed for expression patterns.

Fungal genome projects. The number of fungal genome projects is growing quickly, so it is difficult to give a comprehensive listing of the species that have been sequenced or are in the process of being sequenced. However, some of the major projects are described below.

Model organisms. The first eukaryotic organism with a completed genome was the budding yeast *S. cerevisiae*, completed in 1996. The genome size is slightly larger than 12 megabasepairs, with approximately 6000 identified genes. (Fungal genomes tend to be in the 10–50-Mb range, larger than bacteria but much smaller than mammals, insects, and most green plants.) Among all eukaryotes, *S. cerevisiae* provides the most sophisticated model for how cells work. However, its nonfilamentous nature reduces the value of comparisons with filamentous fungi.

The first filamentous fungus to reach the whole-genome shotgun phase, with its genome data in the public domain, was *Phanerochaete chrysosporium*, an organism sometimes called the model white rot. It has been studied extensively for its role in the degradation of lignin (the complex carbohydrate that makes wood hard) and for its ability to degrade a number of toxic environmental contaminants such as trinitrotoluene (TNT), atrazine, and benzo[a]-pyrene. Two other genetic models have also been sequenced: *Neurospora crassa*, the pink bread mold, and *Aspergillus nidulans* (also known as *Emericella nidulans*), a green mold that was used to elucidate the parasexual cycle. *See* FUNGAL GENETICS.

Economically important species. In addition to model organisms, fungal genome projects have targeted species of economic importance, in particular pathogens that cause plant and animal diseases. Medically important species that are the focus of genome projects include *Candida albicans*, an opportunistic pathogen that is commonly involved in vaginal infections and is the causative agent of thrush (whitish oral lesions); *Cryptococcus neoformans*, causative agent of cryptococcosis, a mycosis (fungal

infection) that commonly spreads to the nervous system; *Histoplasma capsulatum*, causative agent of the most common human pulmonary mycosis; and *Pneumocystis carinii*, an unusual fungus that causes pneumonia, especially in patients with acquired immune deficiency syndrome (AIDS). Finally, *Aspergillus fumigatus*, the main agent associated with the disease aspergillosis, which is the biggest cause of infectious death in bone marrow transplant patients, is the target of an international genome project.

Plant pathogens with genome projects underway include *Ashbya gossypii*, a cotton pathogen; *Ustilago maydis*, cause of corn smut disease; and *Magnaporthe grisea*, causative agent of rice blast. Industrially important species such as *Aspergillus oryzae* (used in soy sauce and sake fermentations), *A. niger* (producer of enzymes and citric acid), *A. terreus* (source of the class of drugs called statins, which reduce blood pressure), *Trichoderma reseii* (maker of enzymes that break down the complex plant carbohydrate cellulose), and *Penicillium chryosgenum* (penicillin producer) are also the focus of genome projects in various stages of completion.

Applications. Comparative genomics is the study of differences and similarities in the content and organization of genomes among organisms. Genomic data coupled with microarray technology facilitate the study of networks of genes and pathways. Entire fungal genomes may be compared for the presence or absence of homologous pathways. A number of proprietary genome projects are also being conducted by pharmaceutical and agricultural companies to isolate new targets for drug screening. When new targets are discovered, libraries of small molecules can be screened to look for their ability to perturb particular fungus-specific pathways, in the hope of finding agents with narrow antifungal activity. Further, genomic analysis can provide information on organisms that are not easily cultivated in the laboratory.

Because genomics provides a new way of finding commonalities and differences among organisms, it facilitates analysis of comparative overviews of biological systems. Genome data can be compared with the genomes of selected eukaryotic models (for example, the fruitfly *Drosophila melanogaster* and the nematode *Caenorhabditis elegans*) and with the genomes of other fungi.

The challenge is to come to an understanding of complex traits such as the ability of organisms to cause disease and produce toxic molecules. In fungi, regulation is often tied to biological circuits for carbon metabolism, nitrogen metabolism, and pH sensing. All of these regulatory circuits are proving to be more complex and more nuanced than many researchers expected. *See* FUNGAL BIOTECHNOLOGY; MOLECULAR BIOLOGY. J. W. Bennett

Bibliography. J. W. Bennett and J. Arnold, Genomics for fungi, in R. Howard and N. Gow (eds.), *The Mycota VIII: Biology of the Fungal Cell*, pp. 267–297, Springer, 2001; A. Kumar et al., An integrated approach for finding overlooked genes in yeast, *Nat. Biotech*, 20:58–63, 2002; R. A. Prade and H. J. Bohnert (eds.), *Genomics of Plants and Fungi*, Marcel Dekker, New York, 2003; C. W. Sensen (ed.), *Essentials of Genomics and Bioinformatics*, Wiley-VCH, Weinheim, 2002.

Fungal infections

Invasion of the body caused by pathogenic fungi. Fungi are microorganisms that fill many niches in nature. Several thousand have been described, but fewer than 100 are routinely associated with invasive diseases of humans. In recent years there has been an alarming increase in the number of exotic and common species implicated in diseases of humans, particularly in those who have acquired immune deficiency syndrome (AIDS) and those who are being treated with broad-spectrum antibiotics or immunosuppressive agents. In general, healthy humans have a very high level of natural immunity to fungi, and most fungal infections are mild and self-limiting. Intact skin and mucosal surfaces and a functional immune system serve as the primary barriers to colonization by these ubiquitous organisms, but these barriers are sometimes breached (**Table 1**).

Unlike viruses, protozoan parasites, and some bacterial species, fungi do not require human or animal tissues to perpetuate or preserve the species. Virtually all fungi that have been implicated in human disease are free-living in nature. However, there are exceptions, including various *Candida* spp., which are frequently found on mucosal surfaces of the body such as the mouth and vagina, and *Malassezia furfur*, which is usually found on skin surfaces that are rich in sebaceous glands. These organisms are often cultured from healthy tissues, but under certain conditions they cause disease. Only a handful of fungi cause significant disease in healthy individuals. Once established, these diseases can be classified according to the tissues that are initially colonized (**Table 2**).

Superficial mycoses. Four infections are classified in the superficial mycoses. Black piedra, caused by *Piedraia hortai*, and white piedra, caused by *Trichosporn beigleii*, are infections of the hair. The skin infections include tinea nigra, caused by *Exophiala werneckii*, and tinea versicolor, caused by *M. furfur*. Where the skin is involved, the infections are limited to the outermost layers of the stratum corneum; in the case of hairs, the infection is limited to the cuticle. In general, these infections cause no physical discomfort to the patient, and the disease is brought to the attention of the physician for cosmetic reasons.

Cutaneous mycoses. The cutaneous mycoses are caused by a homogeneous group of keratinophilic fungi termed the dermatophytes. Species within this group are capable of colonizing the integument and its appendages (the hair and the nails). In general, the infections are limited to the nonliving keratinized layers of skin, hair, and nails, but a variety of pathologic changes can occur depending on the etiologic agent,

TABLE 1. Types of Diseases caused by fungi

Type	Description
Allergies	Hypersensitivity states acquired through exposure to a particular substance (allergen), for example, hypersensitivity pneumonitis
Mycotoxicoses	Infections caused by accidental or recreational ingestion of toxic fungal metabolites, such as mushroom poisoning
Infections	Colonization and invasion of living tissue by fungi, for example the mycoses Primary infection—the result of direct entry of the fungus into healthy susceptible tissues or organs such as the skin and lung Secondary infection—colonization of fungi in active or healed cavitary lesions such as those of tuberculosis, histoplasmosis, or infarcts, such as "fungus ball" Opportunist infections—growth of fungi in tissues of organs of immunosuppressed or otherwise debilitated patients, for example candidiasis

site of infection, and immune status of the host. The diseases are collectively called the dermatophytoses, ringworms, or tineas. They account for most of the fungal infections of humans.

This group of fungi is classified into the genera *Trichophyton*, *Microsporum*, and *Epidermophyton*. They elicit several well-defined clinical syndromes collectively called the tineas. The same species may be involved in different expressions of disease depending on the anatomic site involved. Examples are tinea capitis (ringworm of the scalp), tinea manuum (ringworm of the hand), tinea corporis (ringworm of the glabrous skin), tinea cruris (jock itch or ringworm of the groin), tinea unguium (ringworm of the nails), and tinea pedis (athlete's foot or ringworm of the foot).

There appears to be a delicate balance between host and parasite in dermatophyte infections. Some of these fungi show an evolutionary development toward a parasitic existence. Those that have achieved a high level of coexistence with humans also exhibit a degree of specificity toward tissues that are colonized. As versatile as these fungi seem, they do not appear to invade organs other than the keratinized layers of skin, hair, and nails.

Subcutaneous mycoses. The subcutaneous mycoses include a wide spectrum of infections caused by a heterogeneous group of fungi. The infections are characterized by the development of lesions at sites of inoculation, commonly as a result of traumatic implantation of the etiologic agent. The infections initially involve the deeper layers of the dermis and subcutaneous tissues, but they eventually extend into the epidermis. The lesions usually remain localized or spread slowly by direct extension via the lymphatics, for example, subcutaneous sporotrichosis. The type of pathology evoked is a result of the interplay between host and pathogen (**Table 3**).

TABLE 2. Categories of fungal infections

Type	Description
Superficial mycoses	Infections limited to the outermost layers of the stratum corneun or the hair shaft
Cutaneous mycoses	Infections of the epidermis, fingernails, toenails, and cortex of the hairs
Subcutaneous mycoses	Infections involving the dermis, subcutaneous tissues, fascia, or bone
Systemic mycoses	Infections originating in the lungs, sometimes spreading to other tissues by hematogenous or direct extension

Systemic mycoses. The initial focus of the systemic mycoses is the lung. The vast majority of cases in healthy, immunologically competent individuals are asymptomatic or of short duration and resolve rapidly, accompanied in the host by a high degree of specific resistance. However, in immunosuppressed patients the infection can lead to life-threatening disease.

Histoplasmosis. This mycosis, caused by *Histoplasma capsulatum*, is the most common respiratory fungus infection in the world. The organism grows in soil with a high nitrogen content, in particular, soils contaminated with the fecal droppings of bats and birds (especially chickens and starlings). Bats are naturally infected but birds are not. This disease is sometimes referred to as spelunker's disease because epidemics have occurred in individuals who have explored bat-infested caves and disturbed contaminated soil, resulting in exposure to the fungus. The disease is distributed throughout temperate, tropical, and subtropical areas of the world and is endemic in the Ohio and Mississippi valley regions of the United States, and scattered areas of Mexico and Central and South America. It is a mold in nature and a yeast in tissues of infected patients.

Blastomycosis. This disease is caused by *Blastomyces dermatitidis*. The organism has been isolated from soil and from woody specimens taken from a beaver dam that was the focus of an epidemic. The geographic distribution is limited to portions of North America and Africa. In the United States, the infection overlaps that for histoplasmosis and includes the Ohio and Mississippi valley rivers and the Missouri River basin. The organism is a mold in nature and a yeast, characterized by a broad-based attachment to the budding cell, in infected tissue.

Paracoccidioidomycosis. This disease, caused by *Paracoccidioides brasiliensis*, has a high incidence in Brazil, with endemic foci in Venezuela and Colombia. Epidemiologic surveys indicate that men and women have the same degree of exposure (50:50), but clinical disease is seen more often in males than females

TABLE 3. Subcutaneous mycoses

Disease	Clinical disease	Etiologic agents
Sporotrichosis	Lymphocutaneous sporotrichosis is the disease most commonly associated with this fungus. The lesions start as a painless nodule at the site of injury, followed by other nodules that develop along the lymphatics that drain the site. They become fluctuant and ulcerate. The infection rarely extends beyond the regional lymphatics.	*Sporothrix schenkii*
Chromoblastomycosis	The most common form of chromoblastomycosis consists of warty, vegetative lesions which take on the appearance of a cauliflower floret.	Various dematiaceous (darkly pigmented) fungi belonging to the genera *Fonsecaea, Wangiella, Phialophora*, and *Cladosporium*
Mycetoma	Localized indolent, deforming, swollen lesions are composed of abscesses, and draining sinuses with the presence of "grains" are microcolonies of the fungus.	*Pseudallescheria boydii, Madurella grisea, Acremonium killense, Leptosphaeria senegalensis*, and so on
Zygomycosis	The disease can run a rapid acute course ending in death or become chronic.	Organisms belonging to the phylum Zygomycota including species of *Mucor, Absidia, Cunninghamella*, and so on

(90:10). The organism is a mold in nature and a multipolar budding yeast in infected tissues.

Coccidioidomycosis. Caused by *Coccioides immitis*, this is truly a disease of the New World. Natural infections occur only in North, Central, and South America. Areas of high endemicity include the central San Joaquin valley in California, the southwestern United States, and certain regions of Argentina. It is a mold in nature and a spherule (sporangium) in tissues of the infected host.

Cryptococcosis. The etiologic agent for this disease, *Cryptococcus neoformans*, has been recovered in large numbers from the excreta of pigeons and their roosting areas. However, the organism does not infect these birds. It is the most common cause of fungal meningitis. It is an encapsulated yeast in tissues and culture. The disease is worldwide in distribution and is frequently associated with patients who are immunosuppressed or debilitated, but many patients may have no significant underlying immune or metabolic defect. *See* FUNGI; MEDICAL MYCOLOGY.

George S. Kobayashi

Bibliography. L. Ajello and R. J. Hay, *Medical Mycology*, vol. 4 of *Topley and Wilson's Microbiology and Microbial Infections*, ed. by L. Collier, A. Balows, and M. Sussman, Arnold, London, 1998; K. J. Kwon-Chung and J. E. Bennett, *Medical Mycology*, Lea & Febiger, Philadelphia, 1992; P. R. Murray et al., *Medical Microbiology*, 3d ed.,C. V. Mosby, St. Louis, 1998.

Fungal virus

Any of the viruses that infect fungi (mycoviruses). In general these viruses are spheres of 30–45-nanometer diameter composed of multiple units of a single protein arranged in an icosahedral structure enclosing a genome of segmented double-stranded ribonucleic acid (dsRNA). Viruses are found in most species of fungi, where they usually multiply without apparent harm to the host. Most fungal viruses are confined to closely related species in which they are transmitted only through sexual or asexual spores to progeny or by fusion of fungal hyphae (filamentous cells). Some fungal strains are infected with multiple virus species.

Characteristics. There are exceptions to each of the above generalizations. One mycovirus has been described that is greater than 45 nm in diameter and similar in structure to the deoxyribonucleic acid (DNA) of an animal virus, herpesvirus. This implies that it would have genomes of DNA instead of dsRNA. Some mycoviruses have genomes of single-stranded RNA instead of dsRNA. In some fungal species, virtually every specimen tested has viruses, whereas in other species they are rarely found.

Early studies of fungal viruses focused on their potential for biological control of important plant pathogenic fungi. This was prompted by Michael Hollings's observation that mushrooms killed by a disease known as dieback contained abundant virus-like particles. Subsequently scientists in the drug industry examined the potential effect of virus infection on the secretion of drugs such as penicillin from the fungi which produced them. Although many viruses were found, none appeared to affect the secretion of drugs. However, studies of killer toxins in *Saccharomyces cerevisiae* and *Ustilago maydis* showed an association between the presence of dsRNA mycoviruses and toxin production. Although hundreds of virus-containing fungi have been reported, very few have been studied in significant detail. Three families of mycoviruses are recognized by the International Committee on Taxonomy of Viruses. The most thoroughly studied mycoviruses are in the family Totiviridae.

Totiviridae. This family of viruses has strains in which the entire genetic sequence is contained in one strand of dsRNA. The *S. cerevisiae* virus has only two genes. One gene codes for the polymerase protein responsible for the replication of the virion within the fungal cells, and the second gene codes for the protein that makes up the capsid (shell) of the virus. *Saccharomyces cerevisiae* and *U. maydis* viruses are indistinguishable in appearance by cryo-electron microscopy at a resolution of 26 angstroms (2.6 nm). Their capsids are uniquely

arranged in 60 pairs of protein subunits in icosahedral symmetry.

Saccharomyces cerevisiae is the yeast of commerce, and *U. maydis* is the corn smut fungus responsible for the annual loss of at least 5% of the world corn crop. Both of their viruses have strains that produce killer proteins (KP strains) that are toxic to other members of the same fungal species. Their potential importance to the yeast fermentation industry and to agriculture may be substantial. One *U. maydis* virus toxin (P4) has been crystallized and reconstructed on the computer to a resolution of 1.9 A (0.19 nm). It has 109 amino acids and structural similarity to scorpion toxin. The P4 toxin blocks calcium channels in experiments with mammalian cells, and there is evidence that it may kill fungal cells by interfering with calcium channel activity. Other *U. maydis* and *S. cerevisiae* KP toxins are known, and their nucleic acid coding sequences have been defined.

Partitiviridae. This family is composed of viruses in which genomes are segmented and the critical viral genes (polymerase and capsid) occur in two or more capsids. One of these viruses, *Penicillium stoloniferum* virus S, has been found in three distinct genera of fungi, and most have been found in various molds or plant pathogens. Since they have not been reported to affect human health or agriculture, they have not been studied as thoroughly as members of the Totiviridae.

Hypoviridae. This family has definite economic importance. It contains a single genus, *Hypovirus*, whose numerous strains have specific dsRNA segments that severely reduce the ability of the host fungus to cause disease (hypovirulence). This fungus, *C. parasitica*, is responsible for the decline of the American and European chestnut trees, once an important forest species and a source of wood for houses, furniture, fences, and fires. During the first half of the twentieth century, chestnut blight killed the standing part of most chestnut trees; however, roots survived. As sprouts from these roots were produced, most were also killed by the fungus; however, in recent years stands of chestnut trees in Europe and the United States have begun to recover, and all are infected with strains of *C. parasitica* carrying the hypovirulent strains of the virus. The hypoviruses are unusual in that they do not have a capsid enclosing the dsRNA genome. It appears that the dsRNA may be enclosed in a type of vacuole.

Plant disease. The above viruses are those for which sufficient biological or molecular data have been accumulated to place them in a taxonomic classification. Many other fungal viruses have been studied, but their placement in the classification system or their relationship to specific biological phenomena, such as plant disease, is tentative. The most prominent of these are the mushroom viruses.

The mushroom viruses are strongly associated with a serious disease known as dieback that occurs sporadically in commercial mushrooms throughout the world. There are at least three partially defined mushroom viruses, ten dsRNA segments characterized by size, and three dsRNA segments completely sequenced. Yet it is not clear which capsids and dsRNAs constitute a specific virus since the diseased mushrooms almost always have a complex of more than one virus. Nine of the dsRNA are closely associated with diseased mushrooms, but one commonly appears in healthy mushrooms. Most investigators feel that the mycoviruses are the cause of the dieback disease, but Koch's rules of proof are yet to be demonstrated.

Other plant diseases, in which there is a tentative association of a mycovirus with disease severity, include scab of potatoes caused by *Rhizoctonia solani*, the take-all disease of wheat caused by *Gaeumannomyces graminis*, and Dutch elm disease caused by *Ophiostoma ulmi*. *See* FUNGI; MYCOLOGY; PLANT PATHOLOGY; VIRUS; VIRUS CLASSIFICATION.

R. F. Bozarth

Bibliography. B. Chen and D. L. Nuss, Infectious cDNA clone of hypovirus CHV1-Euro7: A comparative virology approach to investigate virus-mediated hypovirulence of the chestnut blight fungus *Cryphonectria parasitica*, *J. Virol.*, 73:985–992, 1999; S. A. Ghabrial, New developments in fungal virology, *Adv. Virus Res.*, 43:303–388, 1994; F. Gu et al., Structure and function of a virally encoded fungal toxin from *Ustilago maydis:* A fungal and mammalian Ca^{2+} channel inhibitor, *Structure*, 3:805–814.

Fungi

Nucleated, usually filamentous, sporebearing organisms devoid of chlorophyll; typically reproducing both sexually and asexually; living as parasites in plants, animals, or other fungi, or as saprobes on plant or animal remains, in aquatic, marine, terrestrial, or subaerial habitats. Yeasts, mildews, rusts, mushrooms, and truffles are examples of fungi.

Phylogeny

Fungal phylogeny traditionally has been based on morphology, and there have been many schemes, each supported by different morphological characters or interpretations. Nucleotide sequences of genes coding for ribosomes have been compared to infer phylogenies for many organisms, including fungi. The fungal phylogeny most strongly supported by molecular information is also consistent with many basic morphological and biochemical features, so fungal phylogenetics is now much simpler. Organisms in the kingdom Fungi are mostly haploid, use chitin as a structural cell-wall polysaccharide, and synthesize lysine by the alpha amino adipic acid pathway; and their body is made of branching filaments (hyphae). The fungi arose about 1 billion years ago along with plants (including green algae), animals plus choanoflagellates, red algae, and stramenopiles. Ribosomal comparison indicates that the closest relatives to the fungi are the animals plus choanoflagellates.

There are significant life-cycle differences between the early diverging fungal divisions. In both of the early groups, haploid individuals mate to produce zygotes that typically serve as resting spores; the diploid nuclei in these spores meiose before

germination. In the terminal groups, haploid individuals mate but the nuclei do not fuse immediately; as a result, zygote formation is delayed, leaving pairs of unfused nuclei in the hyphae. When the zygote is made, it does not usually serve as a resistant body; instead meiosis occurs directly, followed by spore formation.

Ascomycetes are the most numerous fungi (75% of all described species), and include lichen-forming symbionts. The group has traditionally been divided into unicellular yeasts and allies with naked asci, and hyphal forms with protected asci. However, ribosomal gene sequences indicate that some traditional yeasts and allied forms diverged early (early ascomycetes), at about the time ascomycetes were diverging from basidiomycetes. Hyphal ascomycetes protect their asci with a variety of fruiting bodies; the earliest fruiting bodies may have been open cups (Discomycetes), while in more recent groups they are flask shaped (Pyrenomycetes and Loculoascomycetes) or are completely closed (Plectomycetes). Ascomycetes lacking sexual structures have been classified in the Fungi Imperfecti, but molecular comparisons now allow their integration with the ascomycetes. John W. Taylor

General Characteristics

The mycelium, generally the vegetative body of fungi, is extremely variable. Unicellular forms, thought to be primitive or derived, grade into restricted mycelial forms; in most species, however, the mycelium is extensive and capable of indefinite growth. Some are typically perennial though most are ephemeral. The mycelium may be nonseptate, that is, coenocytic, with myriad scattered nuclei lying in a common cytoplasm, or septate, with each cell containing one to a very few nuclei or an indefinite number of nuclei. Septa may be either perforate or solid. Cell walls are composed largely of chitinlike materials except in one group of aquatic forms that have cellulose walls. Most mycelia are white, but a wide variety of pigments can be synthesized by specific forms and may be secreted into the medium or deposited in cell walls and protoplasm. Mycelial consistency varies from loose, soft wefts of hyphae to compact, hardened masses that resemble leather. Each cell is usually able to regenerate the entire mycelium, and vegetative propagation commonly results from mechanical fragmentation of the mycelium.

Reserve foods are stored in various fungi as fats and glycogen (animal starch); volutin granules are the typical reserve material of yeasts.

Asexual reproduction. Asexual reproduction, propagation by specialized elements that originate without sexual fusion, occurs in most species and is extremely diverse. The most common and important means of asexual reproduction are unicellular or multicellular spores of various types that swim, fall, blow, or are forcibly discharged from the parent mycelium (**Fig. 1**). The more important types of spores are the following.

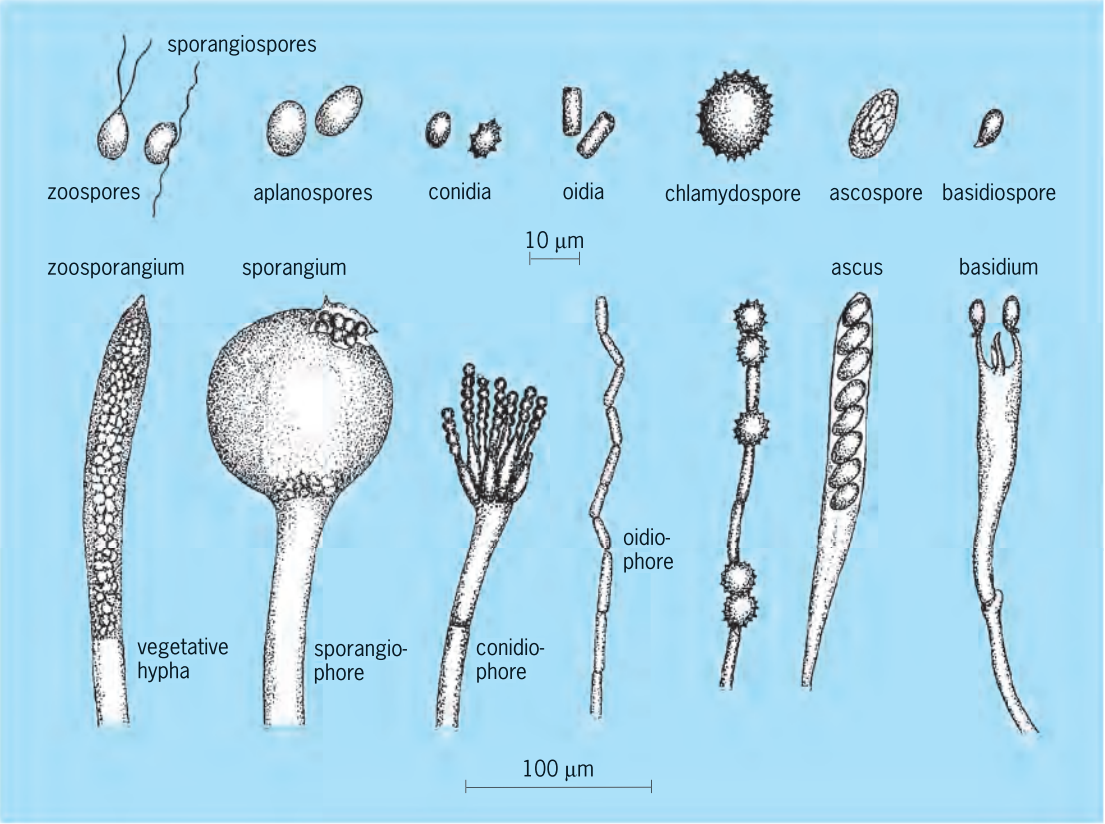

Fig. 1. Spore types and structure of parent mycelia in fungi.

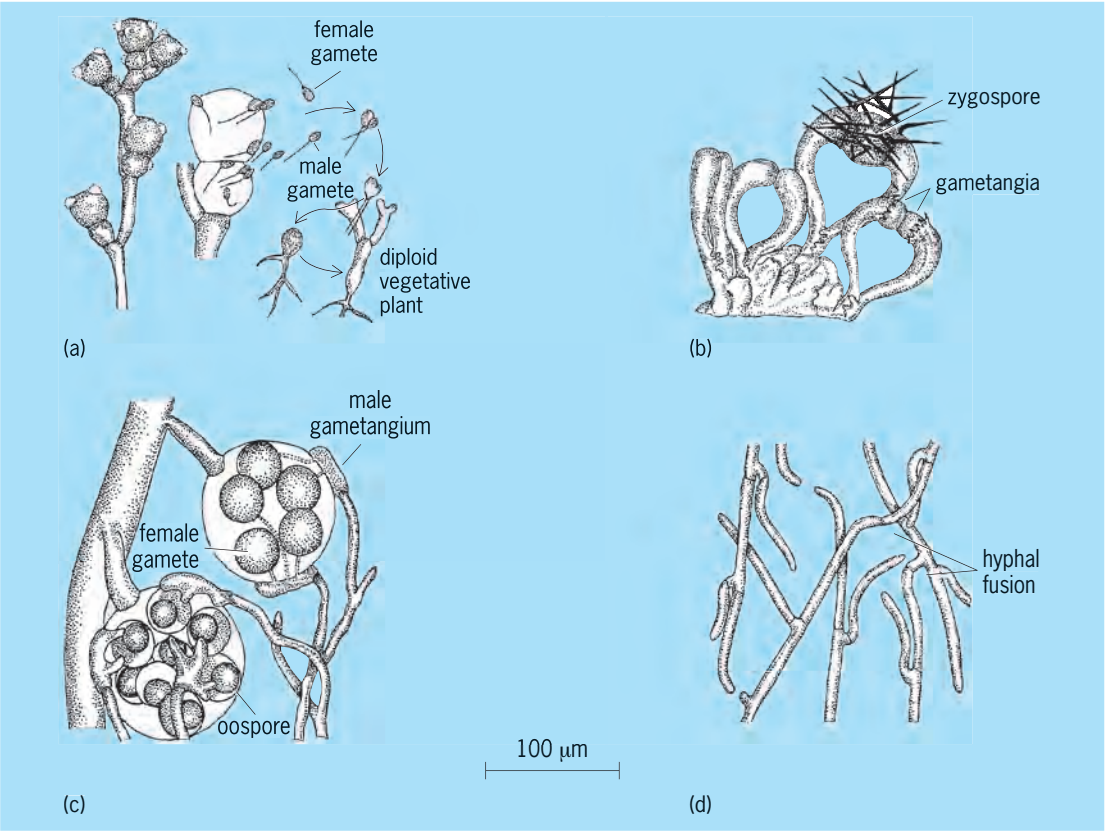

Fig. 2. Sexual mechanisms in fungi. (*a*) Gametic copulation. (*b*) Gametangial copulation. (*c*) Gamete-gametangial copulation. (*d*) Somatic copulation.

Sporangiospores are borne in unicellular sacs termed sporangia. Sporangia originate either by differentiation of vegetative cells or as more specialized, newly formed structures, frequently at the end of elongated stalks or sporangiophores. Sporangiospores are of two types: motile zoospores, equipped with flagella, and nonmotile aplanospores. The number of spores per sporangium varies from a few to thousands.

Conidia resemble nonmotile sporangiospores in shape, size, and structure but are produced externally upon a conidiophore which may be simple or quite elaborate.

Oidia, small, thin-walled cells that usually have flat ends, are produced by the autofragmentation of the vegetative hyphae.

Chlamydospores are thick-walled, nondeciduous spores interposed along vegetative hyphae.

Numerous other spore forms represent variants of the basic types above. Various pigments may be deposited in spore walls and account for mold colors in most cases. Asexual spores typically germinate by germ tubes that develop directly into vegetative mycelia.

Sexual reproduction. Sexual reproduction occurs in a majority of species of all classes. Juxtaposition and fusion of compatible sexual cells are achieved by four distinct sexual mechanisms:

Gametic copulation is the fusion in pairs of differentiated, uninucleate sexual cells or gametes formed in specialized sporangialike gametangia (**Fig. 2***a*). In isogamy, the two members of the fusion pair are alike; in anisogamy, they are morphologically different.

Gametangial copulation is the direct fusion of gametangia without actual differentation of gametes themselves (Fig. 2*b*).

Gamete-gametangial copulation is the fusion of a differentiated gamete of one sex with a gametangium of the other sex (Fig. 2*c*); the differentiated gamete may be either female or male.

Somatic copulation is the sexual fusion of undifferentiated vegetative cells (Fig. 2*d*).

Pattern of sexuality. This defines the underlying relationship between the two sexual elements that fuse in sexual reproduction; this relationship may be homothallism, which is fusion of elements of the same mycelium, or heterothallism, which is fusion only between elements of different mycelia. Obligatory cross-mating in heterothallic species is imposed by a number of distinct genetic devices that determine the mating competence of individuals with respect to sexual sign, that is, male and female, or to incompatibility class, that is, mating type. Incompatibility involves no morphological differentiation and in different groups may determine only two, hundreds, or even thousands of distinct self-sterile, but cross-fertile, mating types.

Product of sexual fusion. This and the time of nuclear fusion vary widely in the different classes

and subgroups. Immediate nuclear fusion and the production of resting diploid spores, such as oospores and zygospores, or a diploid vegetative phase can also occur in lower forms. In more advanced forms, sexual fusion leads to the establishment of nuclear pairs, each a dikaryon, in hyphal systems capable of limited growth or indefinite growth. The dikaryotic mycelium can be parasitic within the closed, flask-shaped, or cup-shaped fruiting body known as the cleistothecium, perithecium, and apothecium, respectively, and its apical cells are differentiated as asci. Or, the dikaryotic mycelium can be capable of independent, indefinite vegetative growth and constitute the fruiting body of mushrooms and puffballs upon, or within, which the basidia are borne. Asci and basidia are saclike cells within which nuclear fusion and meiosis occur prior to production of endogenous ascospores and exogenous basidiospores, respectively.

Heterokaryosis. In heterokaryosis the association of genetically dissimilar nuclei in vegetative mycelia occurs.

Parasexuality. This is a vegetative process involving nuclear fusion and recombination of parental characters in the absence of sexual organs or fruiting structures.

John R. Raper

Physiology

The functions and activities that fungi use to obtain food from the environment and the chemical modifications that they make to the materials taken up relates to fungal physiology. Fungi obtain organic substances (food) from their environment which have been produced through the (photosynthetic) activities of green plants, since fungi do not contain chlorophyll and are unable to manufacture their own food. Fungi are able to digest food externally by releasing enzymes into their environment which have the capacity to break down large molecules into smaller ones. These smaller molecules can be absorbed into the fungal body and transported to various locations where they can be used for energy or converted into different chemicals to make new cells or to serve other purposes. Some of the by-products of fungal metabolism may be useful to humans.

Fungi are multicellular organisms that obtain their food by absorption through cell walls. This characteristic distinguishes them from animals, which ingest their food, and from plants, which manufacture their food. Most fungi use nonliving plant material for food, but a few use nonliving animal material and therefore are called saprophytic organisms. In nature the decomposition of dead plant material is an important function of fungi, as the process releases nutrients back into the surrounding ecosystem where they can be reused by other organisms, including humans.

Decomposition. Fungal decomposition of plant material is very important in nature, as plant material contains high percentages of celluloses, lignins, and hemicelluloses which must be broken down to release the carbon, hydrogen, and oxygen. Fungi biodegrade the substrate, and as much as 70% of the dry organic matter is released as carbon dioxide and water through the process of respiration. Therefore, these fungi are highly involved in the carbon cycle of nature and influence the rate at which it proceeds. Fungi are the most effective decomposers of plant components, and evidence indicates that lignin decomposition is completed almost entirely by fungi. One gram of fertile soil may contain 10,000–100,000 fungal cells, many of which may be involved with different stages of the breakdown of plant material in the soil. *See* BIODEGRADATION; CELLULOSE; FUNGAL ECOLOGY; LIGNIN.

Pathogens. A few fungi have the physiological capability to grow on living plants and may cause diseases such as wheat rust or corn smut on these economically important plants. Some fungi can grow on grains and may produce substances known as aflatoxins which can be detrimental to animals or humans. A few species of fungi have the ability to grow and acquire their food from skin or hair on living animals such as cats, horses, and humans. The disease known as ringworm may result. It is not caused by a worm but by an expanding circular growth of a fungus which has the physiological capability to use the components of skin or hair as the food source. *See* AFLATOXIN; MEDICAL MYCOLOGY; PLANT PATHOLOGY.

The most frequently encountered fungal disease in humans is candidiasis, which is caused by one of the few fungi that is normally found associated with humans (*Candida albicans*). This organism does not ordinarily become a problem until the physiology of the body is changed or compromised in some way. Changes in hormonal level, general health, or immune status of the individual may alter the body physiology. The immune status may be altered by disease, prolonged broad-spectrum antibiotic treatment, immunosuppressant drugs associated with organ transplant, steroid therapy, or the prolonged use of drugs.

Industrial applications. Certain species of fungi have been used by humans since early times in the preparation of foods such as leavened bread, cheeses, and beverages. Additional by-products of fungal physiology are being used in industrial applications such as antibiotics, solvents, and pharmaceuticals. *See* FUNGAL BIOTECHNOLOGY; INDUSTRIAL MICROBIOLOGY; YEAST.

Nutritional requirements. While fungi share many common nutritional factors, groups or individual species of fungi may have specific requirements or capabilities to utilize different or unique food sources. The moisture level must be high for fungal growth to occur, and a relative humidity of about 100% usually supports maximum growth; however, most fungi do not require free water for growth. A few species can grow where relative humidity levels are as low as about 65%, but growth is usually not optimal.

Fungi usually require an organic source of carbon and a source of nitrogen which can generally be from an inorganic source. Carbon sources used by fungi are often carbohydrates, with the most common being sugars. Glucose is the sugar that most

fungi have the ability to utilize at optimal rates. Minerals required by fungi are essentially the same ones that are required by green plants and other organisms.

Fungi are often able to make their needed vitamins; however, certain fungi may require an outside source of one or more vitamins. The need for a vitamin may be absolute (that is, growth is not possible without the vitamin) or partial, in which case, as more vitamin is added, the rate of growth increases. The most commonly reported vitamin deficiency in fungi is of the B-vitamin thiamine.

pH. The measure of the hydrogen-ion concentration in aqueous solutions is a measure of the acidity or alkalinity which may affect fungal growth. In general, fungi are able to grow over a relatively broad pH range, but a pH in the acidic range (below 7) is usually optimal. When growing in culture, fungi often change the pH of the medium. *See* PH.

Some fungi are facultative anaerobes, that is, they can live where there is no free oxygen, but can grow when oxygen is present. Most fungi are obligate aerobes and require some free molecular oxygen in their surroundings to support growth. Some species of fungi grown in shaken liquid cultures may have a growth rate as much as five times that of stationary cultures. Increased oxygenation of the culture medium seems to be a primary factor responsible for the increased growth.

The carbon dioxide concentration of air is approximately 0.03% (by volume); however, the concentration may increase because of cellular respiration in a wet soil or other confined environment. The concentration of carbon dioxide generally does not affect the growth of fungi until it becomes greater than the 10–15% range, and then it may become inhibitory. The concentration at which inhibition becomes measurable varies with different species of fungi.

Reproduction. Many fungi produce nonsexual spores as well as sexual spores in the reproductive process. Production of these spores is often affected by physiological process, including the nutritional status, presence of certain nutrients or metabolites, light, or the ratio of elements, such as carbon to nitrogen. The nutritional and environmental factors required for sporulation are often more specific than those required for vegetative growth, and may also demonstrate a linear relationship; that is, the greater the amount of the required factor, the greater the number of spores.

Physiological associations. Certain species of fungi produce chemicals that permit them to gain entrance to internal plant locations where they may grow and gain food from the host plant. Many of these fungi may become pathogenic and harm the host plant. A number of fungal species are able to enter plant roots and develop an association that may be beneficial to the plant under natural field conditions. This association of a higher plant root and a fungus that does not produce a disease is called a mycorrhiza. This fungal association with the plant root may permit the plant to live under soil conditions where it may not otherwise survive because of an excess of acid in the soil or a lack or excess of certain nutrients. While the presence or absence of this fungal association should be considered in the study of the nutritional status of plants, it is often overlooked. Consequently, under- or overestimation of the function of root systems has resulted. *See* MYCORRHIZAE.

Some species of fungi (belonging to the class Trichomycetes) are found growing within the intestinal tract of aquatic insect larvae (for example, the mosquito, midge, and blackfly) and other arthropods. Most of these species do not grow when they are removed from the intestinal tract and placed in a fungal food source. They have not been found growing elsewhere and, therefore, have an obligate association with the host, even though the host can do quite well without the fungus. They are apparently dependent on a nutritional component or combination of components most likely found only in the intestinal tract of the host. This specific physiological requirement is in contrast to other fungal species which are able to grow on a wide range of food sources. *See* FUNGAL GENETICS; FUNGAL VIRUS; FUNGISTAT AND FUNGICIDE; INSECT PATHOLOGY.

Marvin C. Williams

Bibliography. G. C. Ainsworth, F. K. Sparrow, and A. S. Sussman (eds.), *The Fungi*, vols. 7A and B, 1973; C. J. Alexopoulos and C. W. Mims, *Introductory Mycology*, 4th ed., 1995; L. Bobby (ed.), *Nitrogen, Phosphorous and Sulphur Utilization by Fungi*, 1989; T. D. Bruns, T. J. White, and J. W. Taylor, Fungal molecular systematics, *Annu. Rev. Ecol. Syst.*, 22:525–564, 1991; M. O. Garraway and R. C. Evans, *Fungal Nutrition and Physiology*, 1984, reprint 1991; D. L. Hawksworth, B. C. Sutton, and G. C. Ainsworth, *Ainsworth and Bisby's Dictionary of the Fungi*, 8th ed., 1996; B. Kendrick, *The Fifth Kingdom*, 2d ed., 1992; E. S. Luttrell, The function of taxonomy in mycology, *Mycologia*, 50:942–944, 1958; E. Moore-Landecker, *Fundamentals of the Fungi*, 4th ed., 1996; J. Webster, *Introduction to Fungi*, 2d ed., 1980.

Fungistat and fungicide

Synthetic or biosynthetic compounds used to control fungal diseases in animals and plants. A fungistat prevents the spread of a fungus, whereas a fungicide kills the fungus.

Agricultural fungicides. Chemical compounds are used to control plant diseases caused by fungi. Fungicides now used include both inorganic and organic compounds. Agricultural fungicides must have certain properties and conform to very strict regulations, and of thousands of compounds tested, few have reached the farmer's fields. Special equipment is needed to apply fungicides. *See* PLANT PATHOLOGY.

Requirements and regulations. The stipulations that must be met by manufacturers before they can sell fungicides and other pesticides have been drastically revised. In 1972 the Federal Environmental Protection and Control Act (FEPCA) substantially changed

the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) of 1947. The FEPCA made the U.S. Environmental Protection Agency responsible for regulating the sale and use of pesticides within the country. The law states that all pesticides shall be classified according to their degree of toxicity to humans and other nontarget organisms: safer materials for "general use," and more toxic materials for "restricted use." All pesticides must still be proven effective before they may be registered for sale, and their labels must state all legal uses and the conditions of use. The law also requires that eventually each state must certify the qualifications of those who apply pesticides commercially. Fortunately, most agricultural fungicides are generally safer than the compounds used as insecticides or herbicides. *See* HERBICIDE; INSECTICIDE; PESTICIDE.

Because fungicides must now be labeled for specific uses on specific crops, the cost of development largely limits their use to such major crops as peanuts, citrus, potatoes, and apples. To fill the gap, the U.S. Department of Agriculture has established the IR-4 program to finance and facilitate label clearance for pesticide use on minor crops.

Formulation. The manner in which these compounds are applied, that is, as wettable powders, dusts, or emulsions, is often essential to the success of agricultural fungicides. Raw fungicides must be pulverized to uniform particles of the most effective size, mixed with wetting agents, or dissolved in solvents. These carriers or diluents must not degrade the fungicides or injure the plants.

Seed and soil treatments. Seeds and seedlings are protected against fungi in the soil by treating the seeds and the soil with fungicides. Seed-treating materials must be safe for seeds and must resist degradation by soil and soil microorganisms. Some soil fungicides are safe to use on living plants. An example is pentachloronitrobenzene, which can be drenched around seedlings of cruciferous crops and lettuce to protect them against root-rotting fungi. Other soil fungicides, such as formaldehyde, chloropicrin, and methyl isothiocyanate, are injurious to seeds and living plants. These compounds are useful because they are volatile. Used before planting, they have a chance to kill soil fungi and then escape from the soil.

Volatile fatty acid fungicides. Formic acid, acetic acid, and propionic acid up through pelargonic acid and capric acid (the C_1–C_{10} volatile fatty acids) possess significant fungicidal activity. Many of them are present in natural foodstuffs that are resistant to fungal attack. Use of the volatile fatty acids and their salts in bread to prevent ropy mold is widespread. A. I. Virtanen was awarded the Nobel prize in 1945 for his discovery and development of these lower volatile fatty acids to prevent fungal growth and so preserve the nutritious quality of cattle fodders. Formic acid, acetic acid, propionic acid, and isobutyric acid are used extensively for this purpose. Since these volatile fatty acids stop fungal growth, they prevent mycotoxin generation and lessen the risk of cancer from exposure to mycotoxins. *See* ANIMAL FEEDS.

Application methods. Dusters and sprayers are used to apply foliage fungicides. Conventional sprayers apply 300–500 gal/acre (2800–4700 liters/hectare) at pressures up to 600 lb/in.^{2} (4.1 megapascals). This equipment ensures the uniform, adequate coverage necessary for control. Among later developments in spray equipment are the mist blower and the low-pressure low-volume sprayer. The mist blower uses an air blast to spray droplets onto foliage. Mist blowers have been successful for applying fungicides to trees but are less satisfactory for applying fungicides to row crops. The low-pressure, low-volume sprayers are lightweight machines that apply about 80 gal of concentrated spray liquid per acre (240 liters per hectare) at a pressure of about 100 lb/in.^{2} (6.9 MPa). These have been successfully used to protect tomatoes and potatoes against diseases caused by fungi. A refinement of this method is ultra-low-volume (ULV) spraying. With ULV spraying, which employs special spinning cage micronizers, growers can protect certain crops by applying as little as 0.5–2 gal of spray liquid per acre (5–19 liters per hectare). Fungicides are also applied from aircraft.

James G. Horsfall; Saul Rich

Mode of action of fungicides. Two types of fungicides are used to control plant diseases: (1) Surface protectants remain on the plant surface and exert their toxic action on fungi before they have penetrated into plant tissue. (2) Systemic fungicides move into plant tissue and exert their toxic action on fungi which have already penetrated internally. These fungicides can also provide surface protection by acting on fungi before they have penetrated the plant.

Surface protectants. Prior to 1970, almost all fungicides used to control plant disease lacked the biochemical specificity necessary for therapeutic action within plant tissue; disease control by chemicals was limited essentially to surface protection. Before 1945, inorganic copper compounds such as Bordeaux mixture and copper carbonate and inorganic sulfur in the form of elemental sulfur and lime sulfur were the main fungicides available for plant protection. Copper is a multisite inhibitor that interferes with many different enzymes in fungal cells. Injury to plants is avoided by regulating the amount of soluble copper present on the plant surface. Elemental sulfur accepts electrons from the mitochondrial electron transport system and is reduced to hydrogen sulfide. This action interferes with energy generation and seems responsible for the fungitoxicity of sulfur.

After 1945, organic surface protectants became increasingly available for disease control. Among these fungicides are the dialkyldithiocarbamates, which include the fungicides thiram, ferbam, and ziram. These multisite inhibitors are capable of several kinds of toxic action in fungal cells, including metal chelation, mixed disulfide formation, and transport of heavy metals across membranes. Dialkyldithiocarbamates inhibit a multitude of enzymes, but the pyruvic dehydrogenase system is particularly sensitive to these fungicides.

The mode of action of ethylenebisdithiocarbamates, which include zineb, maneb, metiram, and mancozeb, differs from that of the dialkyldithiocarbamates. They undergo transformation to ethylenediisothiocyanate, which inactivates thiol groups of enzymes and metabolites in fungal cells. The phthalimides captan, folpet, and captafol and the nitrile fungicide chlorothalonil are important multisite inhibitors that also inactivate thiol groups of enzymes and metabolites in fungal cells.

The triphenyltins are surface protectants that interfere with oxidative phosphorylation in fungal mitochondria. They act at the same site as the antibiotic oligomycin.

Surface protectants are inexpensive and are rarely overwhelmed by fungal resistance. Therefore, they are still widely used for plant disease control even though many newer, more potent systemic fungicides have been developed.

Systemic fungicides. Most agricultural fungicides developed since 1965 are systemic compounds that act at a single target site in fungal cells. These fungicides comprise several mode-of-action groups.

The carboxamides, which include carboxin, oxycarboxin, mebenil, pyracarbolid, benodanil, mepronil, flutolanil, and furmecyclox, control basidiomycete-type fungi. Studies of carboxin indicate that the toxicity of the carboxamides is due to specific inhibition of succinic dehydrogenase.

The benzimidazoles benomyl, carbendazim, and thiabendazole and the phenylcarbamate diethofencarb specifically interfere with the formation of microtubules which function in a variety of cellular processes, including mitosis and maintenance of cell shape. These fungicides bind specifically to protein subunits called tubulin, and prevent their assembly to form microtubules.

The largest and most important group of systemic fungicides used to control fungal diseases of plants and humans are inhibitors of ergosterol and other desmethyl sterol biosynthesis. These fungicides are often referred to as ergosterol biosynthesis inhibitors or sterol biosynthesis inhibitors. The ergosterol biosynthesis inhibitors can be divided into three groups based on their specific site of action in the sterol biosynthetic pathway. (1) Squalene epoxidase inhibitors, which include the fungicides tolnaftate, naftifine, and terbinafine, have been developed for medical uses. (2) Inhibitors of the sterol C-14 demethylase include numerous fungicides (see **table**). (3) Inhibitors of the sterol $\Delta^8 \rightarrow \Delta^7$ isomerase, where the double bond is shifted from the 8(9) position to the 7(8) position, include the morpholine fungicides tridemorph, dodemorph, aldimorph, trimorphamide, and fenpropimorph. The morpholine fungicides also inhibit Δ^{14}-reductase, the enzyme that reduces the double bond at position 14(15) in the sterol structure to saturation, in some fungi. Morpholines are used to control plant fungal diseases but are not used for medicinal purposes. The action of ergosterol biosynthesis inhibitors results in a deficiency of the appropriate sterol (usually ergosterol) necessary for proper membrane structure and function and causes cell death or inhibition of growth.

Fungicides that inhibit sterol C-14 demethylase as their mode of action

Chemical class and name	Use*	Chemical class and name	Use*
Imidazole		*Triazole*	
Imazalil	P	Triadimefon	P
Fenapanil	P	Triadimenol	P
Prochloraz	P	Prociconazole	P
Miconazole	M	Etaconazole	P
Ketoconazole	M	Diniconazol	P
Clotrimazole	M	Bitertanol	P
Econazole	M	Diclobutrazol	P
Isoconazole	M	Flutraifol	P
		Fusilazol	P
Pyrimidine		Myclobutanil	P
		Penconazole	P
Fenarimol	P	Itraconazole	P
Nuarimol	P	Terconazole	M
Pyridine		*Piperazine*	
Buthiobate	P	Triforine	P
Pyrifenox			

*P = used or developed to control fungal disease of plants. M = used or developed for medicinal purposes.

The dicarboximide fungicides iprodione, procymidone, vinclozolin, chlozolinate, and metomeclan are particularly valuable for the control of plant diseases caused by species of *Botrytis*, *Sclerotinia*, *Monilinia*, *Alternaria*, *Sclerotium*, and *Phoma*. The modes of action of those compounds remain uncertain, although they are known to inhibit certain flavin enzymes. That action is believed to result in the production of reactive oxygen radicals that destroy membranes by peroxidation of fatty acids in phospholipids and attack nucleic acids and proteins. The dicarboximides are believed to act like the aromatic hydrocarbon fungicides chloroneb, quintozene, and dicloran because fungi that develop resistance to dicarboximides usually are resistant also to the aromatic hydrocarbons.

The phenylamide fungicides metalaxyl, furalaxyl, benalaxyl, ofurace, cyprofuram, and oxadixyl are used to control fungal diseases of plants. Their toxicity is limited primarily to fungi in the order Peronosporales and is due to selective inhibition of ribosomal ribonucleic acid synthesis.

Several antibiotics produced by microorganisms are used to control fungal diseases of plants or animals. The polyene antibiotics amphotericin B and nystatin are widely used to control fungal infections of humans. They bind to ergosterol in fungal membranes and disrupt selective membrane permeability, which leads to cell death.

Cycloheximide, a compound widely used as a specific inhibitor of protein synthesis in eukaryotic organisms, is a potent antifungal agent. Because of its toxicity to higher plants, it is used only to a limited extent for the control of fungal diseases of plants. The antibiotic inhibits protein synthesis by blocking polymerization of amino acids by fungal ribosomes.

The antibiotics blasticidin S and kasugamycin are also specific inhibitors of protein synthesis. They are used primarily in Japan to control rice blast and other fungal diseases of plants.

The polyoxin antibiotics inhibit chitin synthetase, leading to an interference with fungal cell wall synthesis and the distortion and rupture of germ tubes. Polyoxins are used in Japan to control rice blast and several fungal diseases of fruits and vegetables.

Most fungicides control disease by direct toxicity to the growth of fungal pathogens in their saprophytic as well as parasitic phases, but diseases can be controlled by inhibiting activities related specifically to parasitism. The melanin biosynthesis inhibitors tricyclazole and pyroquilon are not toxic to saprophytic growth of the rice blast fungus, *Pyricularia oryzae*, but block the fungus from penetrating the plant epidermis and prevent it from becoming established as a parasite. That is accomplished by inhibiting reactions in the polyketide pathway to melanin. This pathway is induced after the germinated fungus spores have formed appressoria (penetrating structures) on the plant epidermis. It remains unclear why blocking melanin biosynthesis prevents epidermal penetration by appressoria, but evidence points to decreased rigidity of appressorial walls, to a weakened adhesion of appressoria to the plant surface, or to an accumulation of toxic melanin precursors.

Resistance to agricultural fungicides. There have been very few instances of fungal resistance to surface protectant fungicides like inorganic copper compounds, dithiocarbamates, and phthalimides, even though they have been used for more than 35 years. The absence of resistance of target fungi exposed to those fungicides is attributed to their action at multiple sites within fungal cells. However, most systemic fungicides that act at a single site have generated serious problems with fungal resistance. A single gene mutation in a fungus can lead to loss of effectiveness of all fungicides in a particular mode-of-action group. Experience has shown that frequent, uninterrupted use of a fungicide increases the risk for development of a resistant strain of the target organism. That risk can be reduced by alternating use of fungicides with different modes of action or by using them in mixtures. The emergence of a fungicide resistance problem and its persistence in the absence of fungicide use are dependent on the fitness of the resistant strain. If the resistant strain is as fit as the wild-type strain, the former may persist with little reduction in the fungal propulation for many years, whether or not the fungicide is used. Such has often occurred with benzimidazole-resistant fungal strains. On the other hand, if fitness of the resistant strain is lower than that of the sensitive, wild-type strain, the resistant strain may decline to a low percentage of the population in one or two seasons, and the fungicide may again be effective for a limited number of treatments.

The mode of action of a systemic fungicide is also a factor in the risk for resistance development. Benzimidazoles and phenylamides are high-risk fungicides, whereas the ergosterol biosynthesis inhibitors are low- or moderate-risk fungicides. Most mutants resistant to ergosterol biosynthesis inhibitors appear to have a reduced fitness.

The principal mechanism of resistance is decreased affinity of the target sites for the fungicide. This mechanism of resistance has been observed for benzimidazoles, phenylamides, carboxamides, cycloheximide, and kasugamycin. In cases of resistance to benzimidazoles and carboxamides, new fungicides have been designed to fit the modified target sites, and these are highly toxic to the resistant strains. An example is the phenylcarbamate diethofencarb, which was designed to control benzimidazole-resistant strains of *Botrytis cinerea*; it is used in vineyards in France. *See* FUNGI.

Hugh D. Sisler

Bibliography. C. J. Delp (ed.), *Fungicide Resistance in North America*, 1988; Y. L. Nene and P. N. Thapliyal, *Fungicides in Plant Disease Control*, 3d ed., 1993; D. Pimentel (ed.), *Handbook of Pest Management in Agriculture*, 2d ed., 1990; M. R. Siegel and H. D. Sisler (eds.), *Antifungal Compounds*, 1977; A. J. P. Trinci and J. F. Ryley (eds.), *Mode of Action of Antifungal Agents*, 1984.

Furan

One of a group of organic heterocyclic compounds containing a diunsaturated ring of four carbon atoms and one oxygen atom. Furan (**1**) is a typical member of the group. Furfural (**2**) and some of its close rela-

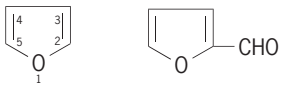

(**1**) (**2**)

tives, such as furfuryl alcohol, tetrahydrofurfuryl alcohol, and tetrahydrofuran, are important chemicals of commerce. *See* HETEROCYCLIC COMPOUNDS.

Properties and preparation. Furan (**1**) is a colorless, volatile liquid, bp 31.4°C (88.5°F), n_D^{20} 1.42140, density (20/4) 0.9378, which is stable to alkali but not to mineral acid. Its water solubility is approximately 1% at room temperature. On exposure to air, furan decomposes very slowly by autoxidation. Substituted furans, particularly negatively substituted furans, are much less sensitive. The furan system is aromatic, with resonance energies of 17–25 kcal/mole reported. Nitration, halogenation, acylation, mercuration, and sulfonation reactions occur with relative ease. The extreme sensitivity of furan itself to strong acid precludes many direct electrophilic substitutions. The presence of negative groups stabilizes the ring. Unoccupied 2 or 5 positions are favored as the site of substitution.

Addition reactions, in contradistinction to substitution reactions, have also been observed. Furan, for example, condenses with dienophiles in the Diels-Alder reaction. *See* DIELS-ALDER REACTION.

The most versatile general furan synthesis is

the acid-catalyzed cyclization of 1,4-dicarbonyl compounds (**3**). The R groups in (**3**) may be hydro-

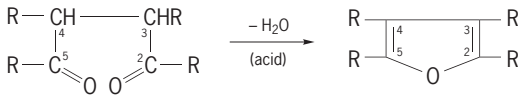

(**3**)

gen, aryl, alkyl, or carbethoxyl. Another synthesis (Treibs) transforms an α,β-unsaturated ketone (**4**) to an unsaturated δ-sulfone (**5**), which on pyrolysis loses

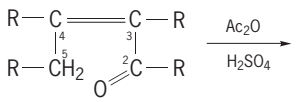

(**4**)

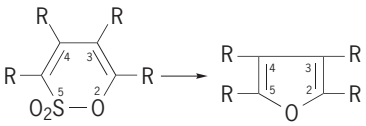

(**5**)

sulfur dioxide and forms the furan. Furans are also formed in the acid-catalyzed dehydration of carbohydrates, the most important example being the production of furfural from the polysaccharides present in annually renewable agricultural waste materials (**6**). Furan is obtained from furfural either directly by catalytic decarbonylation or by oxidation to furoic acid (**7**) followed by decarboxylation.

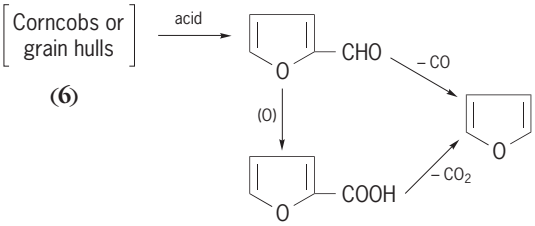

(**6**)

(**7**)

Important derivatives. Oxidation of furans generally disrupts the system, although some ring oxidations give useful products. Ring reduction is possible, the standard method being catalytic rather than chemical. Several products derived from furfural directly or indirectly by hydrogenation are of considerable industrial importance.

Reaction of furans with halogens, although not altogether free of complications, can be used for the synthesis of halofurans. Furansulfonic acids are prepared by sulfonation with sulfuric acid, or with a sulfur trioxide–pyridine complex. Nitration of furan compounds generally makes use of fuming nitric acid in acetic anhydride solvent. The synthesis of the bacteriostatic and bacteriocidal agent nitrofurazone, furacin (**9**), involves such a nitration of derivative (**8**).

Furan aldehydes, prepared by formylation with hydrogen cyanide and hydrogen chloride, resemble benzaldehyde in their functional group reactions. Furyl ketones, prepared by Friedel-Crafts acylation,

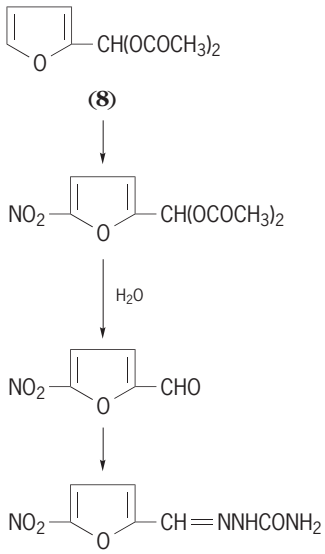

(**8**)

(**9**)

react normally. All the furan mono- and polycarboxylic acids are known. Removal of the carboxyl group from the 2 or 5 positions by decarboxylation is easier than from the 3 or 4 positions.

Tetrahydrofuran (**11**) is prepared either by decarbonylation of furfural to furan followed by catalytic hydrogenation, or by the Reppe process from butynediol (**10**). Tetrahydrofuran is a water-miscible,

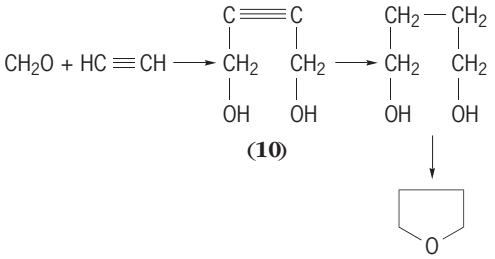

(**10**)

(**11**)

colorless liquid, bp 66°C (151°F), density (20/4) 0.8880, and n_D^{20} 1.4073. It is an industrial solvent, and also has served as an intermediate in the synthesis of butadiene for polymerization, adiponitrile (**12**) for nylon, and pyrrolidone (**13**) for the blood extender polyvinylpyrrolidone.

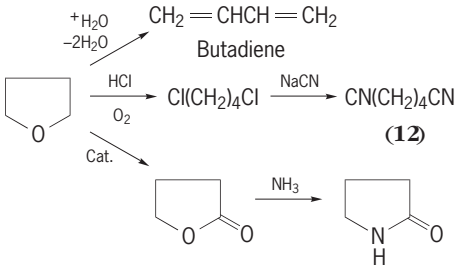

(**12**)

(**13**)

Walter J. Gensler

Bibliography. A. P. Dunlop and F. N. Peters, *The Furans*, Amer. Chem. Soc. Monogr. no. 119, 1953; G. P. Ellis, *Synthesis of Fused Heterocycles*, 1987; R. J.

Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; T. L. Gilchrist, *Heterocyclic Chemistry*, 3d ed., 1997.

Furnace construction

A furnace is an apparatus in which heat is liberated and transferred directly or indirectly to a solid or fluid mass for the purpose of effecting a physical or chemical change. The source of heat is the energy released in the oxidation of fossil fuel (commonly known as combustion) or the flow of electric current through adjacent semiconductors or through the mass to be heated. In recent years, scientific and engineering effort has been made to utilize nuclear and solar energy for heating purposes. Therefore, according to the source of heat and method of its application, there are four categories of furnaces; combustion, electric, nuclear, and solar, in the order of their present commercial or industrial importance.

Furnaces employing combustion vary widely in construction, depending upon the application of the heat released, whether direct or indirect. Direct heat transfer is used, in regenerative refractory-type heaters, in flow systems in which reactants are injected into the combustion gases. Indirect heat transfer is employed in heaters in which the mass to be heated is kept separate from the combustion gases and made to flow in tubes which absorb and transmit the heat to the fluid to be heated.

Furnaces developed for indirect heat transfer can be divided into two classes. One class of heaters is used solely for general utility purposes, such as all types of boilers; and the second class is applied in the petroleum and chemical industries as an essential unit operation in refining or processing plants. The second category of furnaces will be discussed in this article. *See* BOILER.

Directly fired furnaces are employed in oil refineries and chemical process units whenever the temperature level to which a fluid must be heated is above that attainable with utility steam. Furnaces as heat-transfer apparatuses generally cost more per thermal unit of heat transferred than conventional tubular heat exchangers. According to the kind of service, furnaces in process units can be divided into two classes:

1 Those which perform solely a heating duty, that is, raising the temperature of a fluid and effecting essentially no change of state or of chemical composition. Furnaces which perform this one duty may be termed conventional heaters.

2 Those which handle a fluid undergoing a change during heating. The physical or chemical changes constitute an essential performance requirement. Typical applications of this order are associated with such processes as distillation or preheating of temperature-sensitive materials, pyrolysis of hydrocarbons or organic chemicals, and catalytic steam-gas reforming for the production of synthesis gas.

The conventional tubular heater came into industrial use, especially in oil refineries, about 1925 and was termed a tube still to distinguish it from the shell still, a horizontal cylindrical vessel mounted on top of a firebox. The design of these tube stills or furnaces, which is essentially the same as originally conceived by analogy with developments in steam-generating equipment, consists of pipes connected by 180° return bends forming a continuous coil and arranged in a refractory furnace setting, partly in the combustion chamber. Heat is absorbed mainly by radiation and partly in a confined flue-gas passage from the combustion chamber in which heat is absorbed mainly by convection. The fluid flow through the coil is generally countercurrent to the flow of combustion gases, first through convection tubes and then through the radiant-tube section; thus a reasonable thermal efficiency can readily be attained by providing convection tube surface to an economically justifiable extent. *See* ELECTRIC FURNACE; FIRED HEATER; PETROLEUM PROCESSING AND REFINING; STEAM-GENERATING FURNACE.

The furnaces provided in modern process units vary considerably in outer shape, arrangement of tubes, and type and location of burners, depending mainly upon the desire of the designer or manufacturer to be identified with a specific model. Some distinctive furnace designs are shown in **Fig. 1**.

Design. The furnace design for a given performance or thermal efficiency is usually evolved by the following procedure: (1) determination of the composition of the combustion products and the amount of the liberated heat which must be utilized to meet the postulated thermal efficiency; (2) allocation of heat to be absorbed by the heating elements located in the combustion or radiant chamber and in the convection section; (3) determination of the heat-transfer rate and heating surface area in the radiant section; and (4) determination of the heat-transfer rate and tube surface area in the convection section or sections.

Source of heat. The combustion products vary in composition according to the type of fossil fuel burned and the excess air used in the oxidation process. The carbon and hydrocarbon content of the fuel governs the combustion-gas composition, and many formulas relating it to the elemental, oxidizable constituents of the fuel have appeared in the literature. Complete combustion is a prerequisite for high thermal efficiency; to ensure this, air is used in excess of the minimum or stoichiometric requirements, depending upon the type of fuel and the combustion equipment or system (**Table 1**).

The type of fuel, its heating value, and the excess air applied in combustion determine the theoretical flame temperature which would prevail if the oxidation were instantaneous. However, combustion is a rate process requiring time, and the actual flame temperature is considerably lower because of the radiation from the combustion zone. The theoretical flame temperature can readily be determined from the heating value of the fuel and the amount of combustion products evolved, including the excess air. Accurate enthalpy data on the combustion gases are

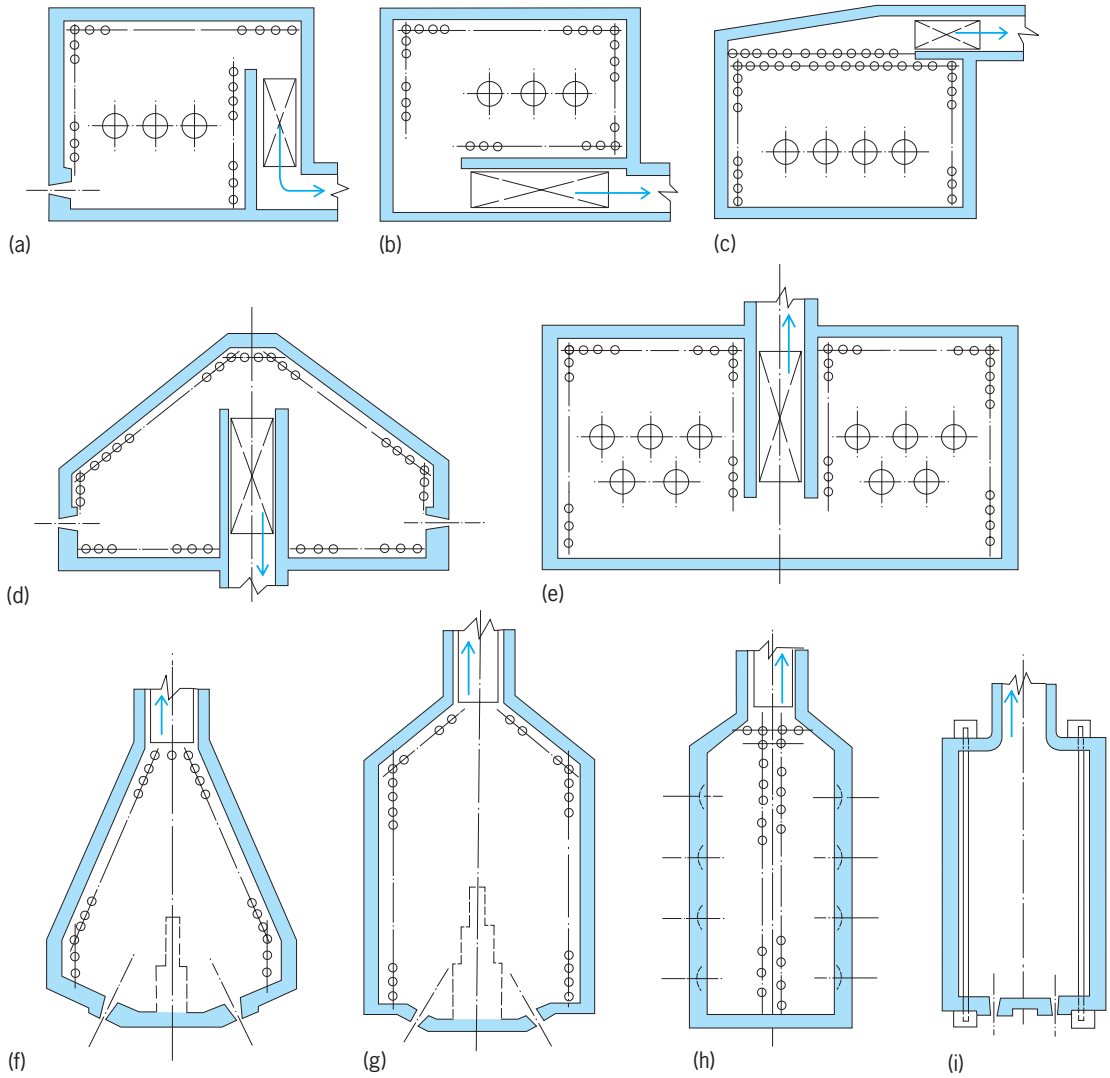

Fig. 1. Typical process furnace designs and characteristic features. (*a*) Radiant, updraft, vertical-convection bank. (*b*) Radiant, downdraft, horizontal-convection bank. (*c*) Radiant, updraft, with high convection effect in radiant roof tubes. (*d*) Large refractory heater, dual radiant sections, sloping roofs. (*e*) Large conventional refinery heater, dual radiant sections. (*f*) Updraft furnace, slanting walls, two parallel coils. (*g*) Updraft furnace, vertical walls, two parallel coils. (*h*) Updraft furnace, tubes in center, burners in wall, and with good heat-intensity distribution. (*i*) Updraft circular, all-radiant furnace, with the burners in the floor.

needed and are given in the literature. *See* COMBUSTION; FOSSIL FUEL.

In order to meet a required thermal efficiency, the heat losses must be appraised. There is a certain loss through the furnace setting which varies with the surface area of the enclosure, the heating capacity, and the atmospheric environment. The wall construction of most types of furnaces, that is, application of insulating material, is such that the specific heat loss, Btu/(ft^2)(h) of outer wall area, is practically the same for all operating temperature levels; hence the heat loss will vary principally only with the furnace capacity. **Table 2** shows the average values that will be encountered in industrial furnace constructions. The setting heat loss ΔH_{sl} can be calculated from the relation $\Delta H_{sl} = \lambda H_{fl}$ where λ is the fraction of heat released, and H_{fl} is the enthalpy of the combustion gas at the theoretical flame temperature.

TABLE 1. Fuel-air mixtures

Type of fuel	Type of burner	Normal range of excess air, %
Natural or refinery gas	Air and fuel gas premixed	5–15
Natural or refinery gas	Air induced by natural draft	10–20
Distillate fuel oils	Steam atomization air by natural draft	15–25
Distillate fuel oils	Mechanical atomization air by natural draft	15–30
Residual fuel oils	Steam atomization air by natural draft	20–35

TABLE 2. Heat loss from setting

Heat liberated, Btu/h*	Loss as a fraction of heat released, λ $\Delta H_{SI} = \lambda H_{fl}$
15,000,000	0.030
30,000,000	0.0275
60,000,000	0.0225
75,000,000+	0.020

*1 Btu = 1.055 kJ.

The flue-gas effluent temperature from the furnace can then be determined in accordance with the postulated thermal efficiency η, as shown in the equation below, where H_{st} is the enthalpy of the combus-

$$H_{st} = H_{fl} \times (1 - \eta) - \Delta H_{st}$$

tion gas at the convection-bank exit temperature.

Heat load. The distribution of heat load between the radiant and convection sections of conventional heaters is associated with the manufacturer's design approach and service requirements. There are all-radiant heaters, and furnaces with tubes arranged mostly for heat absorption by convection. The amount and disposition of heating surface in the combustion or radiant chamber has a distinct influence upon the radiant-heat absorption and the supplementary heat to be recovered by convection for the required thermal efficiency. Therefore, the allocation of heating duties to the radiant and convection sections is more or less an empirical matter. Conventional process heaters are generally designed for approximately 75% thermal efficiency in the United States, where fuel costs are relatively low. Most box-type heaters are designed for moderate radiant-heat-transfer rates; approximately 65–75% of the heat load is carried by the radiant section and 25–35% by the convection bank. In some cases, very high radiant-heat intensities are employed as a matter of design principle, and in these furnaces, the heat absorbed in the radiant section may be only 50% of the total duty.

Radiant section. The present procedure for determining the radiant-heat-transfer rate and heating surface rests on a long series of scientific investigations and evaluation of operating data.

The combustion gases radiate to the solid-body environment, tubes, and refractories. Carbon dioxide and water vapor are the principal constituents with emissive power. Their emissivity depends upon the partial pressure and the thickness of the gas layer, and at 1900°F (1040°C) has a value of approximately 25% of the blackbody radiation intensity. The terms luminous- and nonluminous-flame radiation used in scientific treatises are somewhat misleading; the mere luminosity or light has no significance as far as heat radiation is concerned, because practically all emissivity, and certainly carbon dioxide and water radiation, is in the infrared wavelength region, even at flame temperatures encountered in process heaters. The maximum gas radiation for an infinite value of the product of the partial pressure and thickness of the gas layer, termed black gas, amounts to only 21–33% of the complete blackbody emissivity at temperatures of 2500 and 1100°F (1400 and 590°C), respectively. The mean radiating temperature in the combustion chamber lies within this range.

The consideration of heat transmission by radiation alone, from the combustion gas and refractory walls by reradiation, has not resulted in the derivation of a universally applicable formula for the heat-transfer rate in the radiant section. This is because heat is also transmitted to the tubes by convection, which is on the order of 10% to as high as 35% of the heat absorbed in the radiant section, depending upon (1) the shape of the combustion chamber, (2) the arrangement of tubes, and (3) the type and location of burners. The convection effect is a linear function of the temperature difference (not the difference of temperatures to the fourth power) and can be varied for the same furnace design by the amount of excess air employed. *See* HEAT RADIATION.

Convection heating surface. The rate of heat transfer solely by convection, with flow of gas at a right angle to tubes, increases with the gas velocity and the temperature level, and is greater for small than for large tubes.

There are radiation effects in the convection-tube bank augmenting the heat absorption, especially when the flue-gas temperatures or the temperature differences between tube wall and flue gases are large. These include (1) radiation to the front two rows of tubes, which can be appreciable if they are exposed to the combustion chamber; (2) radiation from the hot gases surrounding the tubes; and (3) the radiation from the refractory enclosure to the tube bank as a whole. *See* CONVECTION (HEAT).

Process furnaces. The design of process furnaces, although following generally the procedure outlined earlier for conventional heaters, requires careful consideration of the transitory state of the fluid being heated.

Vacuum distillation of a temperature-sensitive fluid or the stripping of solvent from the raffinate and extract of a selective solvent-separation process are typical cases in which temperature limitations are encountered in the heating operation. Partial vaporization of the fluid must take place in the tubes under relatively mild heat-transfer rates and tube-wall temperatures, not greatly in excess of the bulk-flowing temperature. Aside from the knowledge and distribution of the combustion-gas side, the proper design of the furnace coil involves exacting pressure-drop calculation for mixed vapor-liquid flow in conjunction with phase-equilibrium determination of the amount and composition of vapor and liquid at any given point in the coil. The pressure drop for mixed vapor-liquid flow is reasonably well established, and vapor-liquid equilibrium relationships are known for many binary and ternary systems. The determination of temperature, pressure, and transition state of the fluid throughout the coil involves stepwise trial-and-error calculations which can best be executed by a modern digital computer for various coil dimensions and heat-intensity patterns. The coil design that satisfies all technical criteria and cost aspects can then

be selected from the computer evaluations. *See* DISTILLATION.

Gas reforming is another industrially important heating process involving a chemical change of the fluid. Methane or natural gas and steam are decomposed in the presence of excess steam in tubes filled with a catalyst. The principal products are hydrogen, H_2, carbon monoxide, and carbon dioxide, and maximum conversion of hydrocarbons is attempted. The chemical transformation proceeds at approximately 1400–1450°F (760–790°C) and is highly endothermic, having a heat of reaction of 87,500 Btu/(lb)(mole of methane converted). The tubes are located in the center of the radiant or combustion chamber and are heated from two sides to effect an even circumferential heat-intensity distribution. The average rate of heat transfer is in the order of 18,000 Btu/(ft^2)(h).

Pyrolysis of hydrocarbons, the thermal conversion of ethane, propane, butane, and heavier hydrocarbons to produce olefins and diolefins for the petrochemical industry, is another example of a process furnace. The modern pyrolysis furnace is an outgrowth of the oil-cracking heater, the mainstay of motor-fuel production from petroleum before the advent of catalytic cracking. The design of a pyrolysis furnace must take thermochemical equilibria and reaction rates into consideration in order to realize a high conversion and a product distribution most favorable with respect to yield of olefins. The conversion of higher-molecular-weight hydrocarbons which have a high rate of decomposition involves the problem of attaining the coil outlet temperature dictated by equilibrium relationships. A shorter coil-residence time and correspondingly higher heat-transfer rates are means to this effect, but still the basic principle of moderating the heat intensity in the pyrolysis coil zone, in which the reactant is already in a high state of conversion, must be observed. **Figure 2** gives a typical temperature-pressure and heat-flux pattern for a pyrolysis coil cracking a naphtha boiling in the range 190–375°F (88–191°C) for maximum ethylene yield. The design of such a pyrolysis coil can be executed with greatest precision or refinement by the use of a modern digital computer, permitting the rapid evaluation of all design variables. *See* CRACKING; PYROLYSIS.

Mechanical construction. The design of the furnace setting and application of refractories and insulating materials generally follow standard practice prevailing in boiler-furnace construction. Fully suspended wall construction and the use of lightweight firebrick with bulk densities of 30–65 lb/ft^3 or 480–1040 kg/m^3(as compared with 125–150 lb/ft^3 or 2000–2400 kg/m^3 for fireclay bricks) are generally preferred for process heaters to reduce the heat storage in the setting, start-up, and shutdown time. *See* REFRACTORY.

The burners are carefully selected and must offer great flexibility with respect to firing rate and excess air. The use of a multiplicity of small-capacity burners is rather common for specific process heaters such as gas-reforming and pyrolysis furnaces.

The design of the heating coil for high-temperature

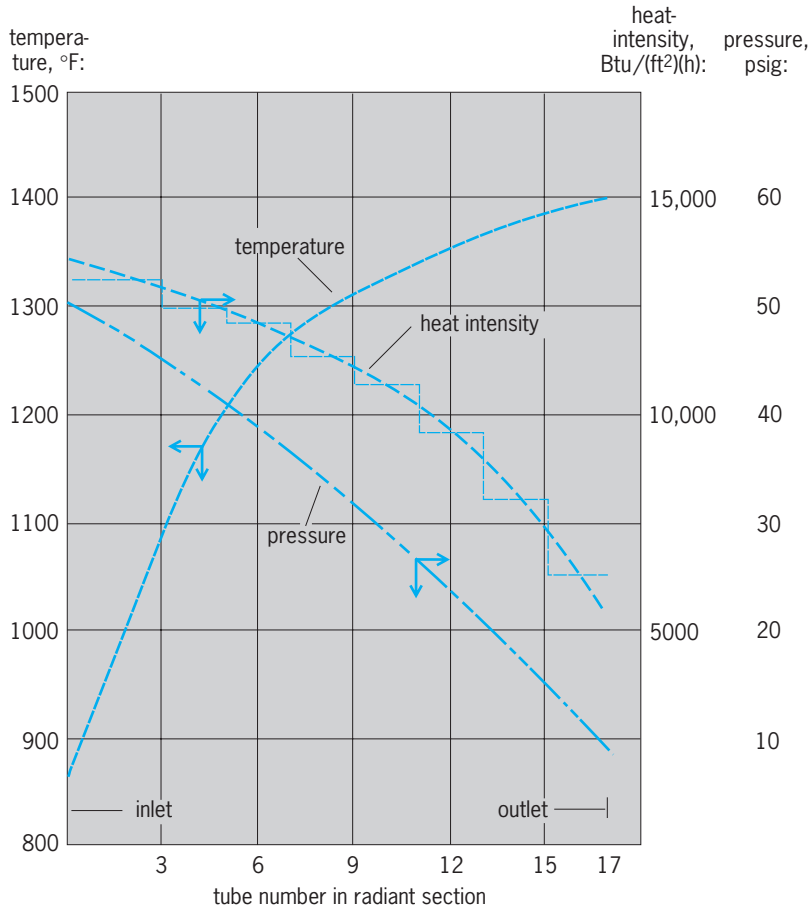

Fig. 2. Typical temperature and heat-flux pattern for naphtha pyrolysis coil. °C = (°F−32)/1.8. 1 Btu/(ft^2)(h) = 3.15 W/m^2. 1 psig = 6895 Pa.

service involves the application of alloy steels. Tube materials are usually selected according to the schedule in **Table 3**. Tubes of carbon and low-chrome steel are usually connected by forged-welding return bends, and those of stainless steel by cast 25/12 chrome-nickel alloy steel bends, also weldable to the tubes.

The flue-gas exhaust is normally carried to the atmosphere by a stack of sufficient height to induce a draft equivalent to the pressure drop in the flue-gas passage. For highest heat economy a process furnace can readily be equipped with an air preheater or waste-heat boiler usually requiring air and flue-gas exhaust blowers. Preheated combustion air

TABLE 3. Heating-coil materials

Tube-wall temperature range, °F (°C)	Tube material*
875 (470) and under	Carbon steel, A-161
875–1150 (470–620)	Low-chrome steel, A-200 (1.25–2.25% Cr, 0.5% Mo)
1150–1550 (620–840)	1/√100 stainless steel, A-271, type 304
1550–1800 (840–980)	2/√520 stainless steel, A-271, type 310

*American Society for Testing and Materials specifications.

results in higher flame and mean effective radiation temperatures which aid in attaining the high radiant-heat-transfer rates required in certain pyrolysis operations. These furnace accessories are provided in the same manner as in boiler design practice. *See* CHIMNEY; GAS FURNACE; HEAT EXCHANGER; OIL FURNACE; PYROMETALLURGY. H. C. Schutt; H. L. Beggs

Bibliography. J. D. Gilchrist, *Fuel, Furnaces, and Refractories*, 1977; D. Q Kern, *Process Heat Transfer*, 1950; R. D. Reed, *Furnace Operations*, 3d ed., 1981.

Fuse (electricity)

An expendable protective device that eliminates overload on an electric circuit. The fuse is connected in series with the circuit being protected. The components of a typical low-voltage high-power fuse are a fuse element or wire, an insulating material support and housing, two metal end fittings, and a filler (**Fig. 1**).

The fuse element is a silver strip or wire that melts when the current is higher than the rated value. The melting of the wire generates an electric arc. The extinction of this arc interrupts the current and protects the circuit. The fuse element is connected to the metal end fittings which serve as terminals. *See* ARC DISCHARGE.

The filler facilitates the arc extinction. The most commonly used filler is sand, which surrounds the fuse element. When the fuse element melts, the heat of the arc melts the sand near the element. This removes energy from the arc, creating a channel filled with the mixture of melted sand and metal. The metal particles from the melting fuse wire are absorbed by the melted sand. This increases the channel resistance, which leads to the gradual reduction of the current and the extinction of the arc. The insulating support and the tubular housing holds the fuse elements and the filler, which also serves as insulator after the fuse has interrupted the current.

The interruption time is the sum of the melting and the arcing time. It is inversely proportional to the current, that is, a higher current melts the wire faster (**Fig. 2**). The fuse operates in a time-current band between maximum interruption time and minimum meeting time. It protects the electric circuit if the fault current is interrupted before the circuit elements are overheated. The arc extinction often generates overvoltages, which produce flashovers and damage. A properly designed fuse operates without overvoltage, which is controlled by the shape of the fuse element and by the filler.

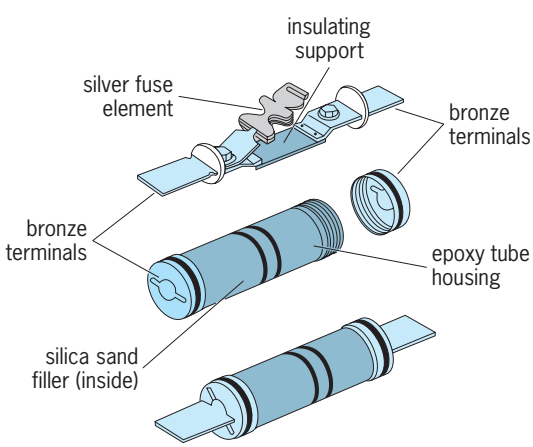

Fig. 1. Power fuse assembly. (*After A. J. Pansini, Electric Power Equipment, Prentice Hall, 1988*)

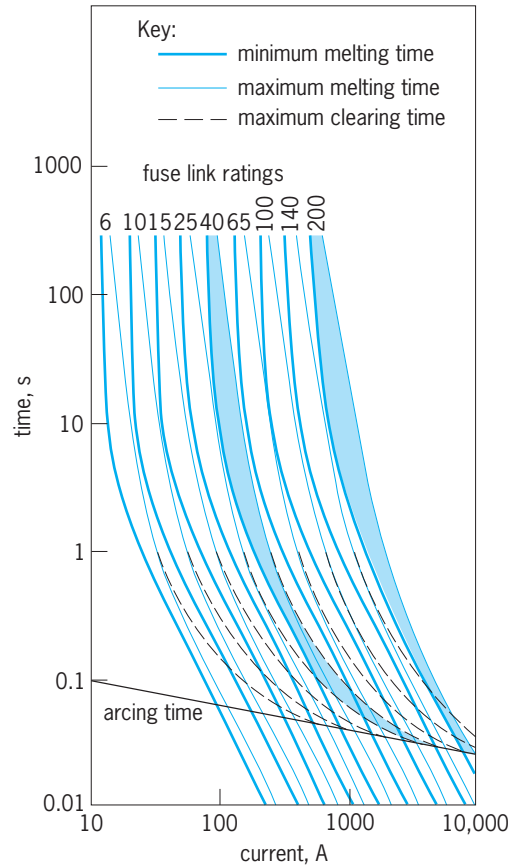

Fig. 2. Minimum and maximum time-current curves for type-K fuse links. (*McGraw Edison Co.*)

Ratings. Fuses protect every kind of electric equipment from the low-voltage direct-current circuits of a motor vehicle to alternating-current circuits of a power distribution system. The fuses are rated to operating voltage and current. The operating voltage is the system voltage; it appears across the fuse after circuit interruption. The voltage rating is between a few volts to 220,000 V. The fuse can carry the rated current continuously. Typical values range from milliamperes to a few thousand amperes. The fusing or melting operation starts around two times the rated current, although the fuse is able to interrupt several thousand amperes of fault current. As an example, the National Electric Manufacturers Association (NEMA) requires a minimum interrupting capacity of 10,000 A from low-voltage ac fuses. Some fuses are able to interrupt 100,000 and even 200,000 A.

Many applications require the connection of fuses in series. A typical example is a low-voltage bus-bar

with five outgoing lines each rated 40 A. The bus-bar is protected by a 200-A fuse, and the outgoing lines are protected by 40-A fuses. A line fault should be cleared by the line fuse before the bus-bar fuse starts to melt. This is achieved by coordinating the fuses, which requires that the fuse characteristics are different. Figure 2 shows that the 40-A fuse interrupts any fault current before the 200-A fuse starts to melt. Safety requires at least a 0.4-0.6-s difference in operation time. *See* BUS-BAR.

Classification. Fuses are classified according to whether they carry direct or alternating current and the level of voltage.

Low-voltage direct-current. Typical dc fuses are used to protect circuits in an automobile or in electronic equipment. The modern automobile uses a blade-type fuse that has a ribbon fuse element mounted on a flat substrate.

Electronics equipment uses a cartridge fuse, which consists of a glass tube with two metal caps and a fuse wire or blade between the caps. Both of these fuses are mounted between spring-loaded contacts or held in place by a screwed tubular housing.

In a special category are fuses designed to protect semiconductor devices such as diodes, transistors, and thyristors. These fuses operate extremely fast and interrupt the current before the sensitive semiconductor is damaged.

Low-voltage alternating-current. Typical examples are the ac fuses used to protect domestic electric systems. The fusible element, a silver strip, is mounted in a holder similar to the base of the standard light bulb (**Fig. 3**). Melting starts at the indent or notch on the fusible element. Burned-out fuses can be changed without switching off the supply, because the top of the holder is made of insulating material.

High-power industrial systems are protected by cartridge-type fuses (Fig. 1). The expendable fuse is bolted on two insulated metal blocks or held by spring-loaded contacts.

The starting of an induction motor causes a short overcurrent, three to eight times the motor rated current. Repeated start of the motor should not melt the fuse, but a short circuit in the motor or the blocking of the rotor should melt the fuse immediately. Motor protection requires time-delayed fuses, with longer operation time for moderate overcurrents. This is achieved by increasing the thermal capacity of the fuse element by attaching a blob of solder to the center of an oversized silver wire. The wire delays the melting of the fuse for short durations of overcurrent. Sustained overcurrent melts the solder, which dissolves the silver and initiates fuse operation. Different time-delay fuses are available for motor protection, but none provide full protection, for which a combination of a fuse and a magnetic circuit breaker is required. The fuse provides protection for short circuits and the circuit breaker for overloads.

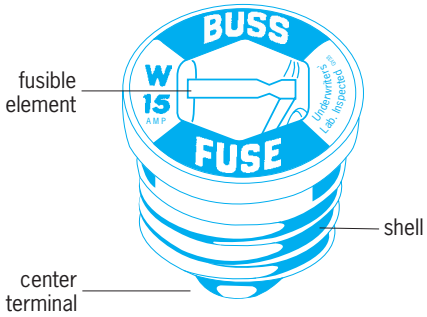

Fig. 3. Common plug-type fuse. (*Bussman Manufacturing Division, McGraw Edison Co.*)

Short-circuit current and a high-power network can be very high. It may produce both mechanical and thermal damage. Special very fast fuses are used for limiting peak current. They interrupt the current before the peak value is reached. Fast operation is achieved by cutting notches or holes in the fusible element. This results in the simultaneous melting of the element at each notch and produces several arcs connected in series. These arcs melt the sand, which removes energy from the circuit faster than the single arc found in the ordinary fuse. The performance of current-limiting fuses can be improved by covering the notches with solder.

The detection of the melted fuse is difficult, but can be facilitated by equipping fuses with mechanical devices that provide visual indication of fuse melting in the form of a colored flag.

High-voltage. High-voltage fuses are extensively used by utilities for protecting distribution systems, and have limited use for transmission systems. The voltage rating is up to 220 kV, and the interrupting capacity may be over 40 kA.

High voltage requires long fuses; and the most frequently used is the expulsion type. The fuse element is installed in an open-ended unfilled tube cartridge. The melting of the element and the arc generates high pressure. The expansion of the high-pressure gases through the open end cools the arc and interrupts the current.

Liquid-filled tubes and short spring-loaded fusible elements are also used. The melting of the fusible element activates the spring, which pulls the terminals apart, extends the arc, and interrupts the current. Some fuses use boric acid, bone fiber, or a liquid solution as a filler. The arc generates water vapor, which helps to cool and quench the arc.

Pole-mounted fuse cutouts are used for the protecting distribution systems. These fuses, mostly of the expulsion type, are mounted in such a way that operation causes the fuse to drop to an open position. This permits the easy detection of the burned-out elements from the ground. **Figure 4** shows a 15-kV fuse cutout, which can be mounted on distribution line poles and used for both line and transformer protection. An additional advantage of the fuse cutout is that in no-load condition the lineman can open the circuit by an insulated stick (hot stick) for repair. The open fuse cutout provides visual proof that the circuit is open.

For the deenergization of a loaded circuit, manufacturers offer a current-interrupting tool with the fuse cutout. The lineman connects the current interrupter in parallel with the fuse. The device interrupts

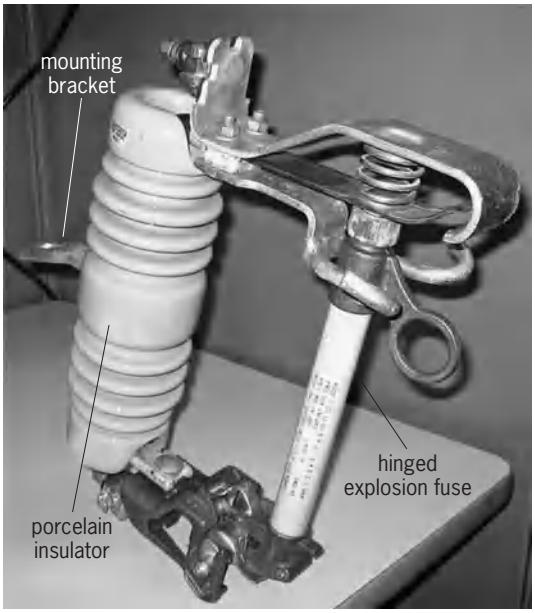

Fig. 4. Typical pole-mounted fuse cutout (15 kV, 100 A).

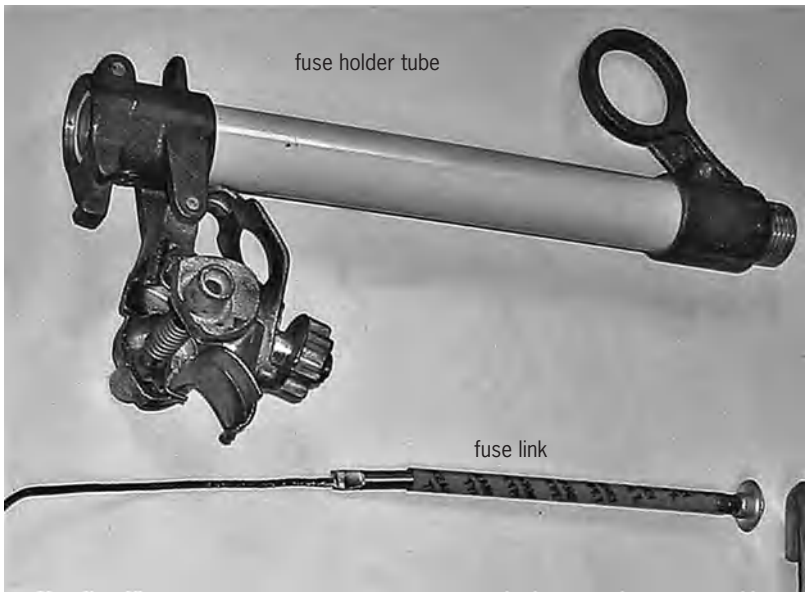

Fig. 5. Fuse holder tube and fuse link.

the load current, and the lineman pulls the hinged fuse to the open position using a hot stick.

The major components of a fuse cutout are mounting bracket, porcelain insulator, and hinged explosion fuse. The fuse contains a fuse holder tube and a fuse link. The fuse link is installed in the plastic fuse holder tube. **Figure 5** shows the fuse holder tube and fuse link.

The fuse link contains one or more fusible elements. **Figure 6** shows the cross section of a typical fuse link. The fusible element is installed in a plastic tube and shunted by a thin wire. A flexible copper wire is connected to the fusible element. The end of the plastic tube is open. The fuse link is installed in the fuse holder. The flexible connecting wire is attached to a spring-loaded lever at the bottom of the fuse holder.

The high short-circuit current melts the fusible element and produces an arc. The arc burns the plastic tube and heats the inner lining of the fuse holder. The high temperature produces high-pressure gases, which exhaust at the bottom of the fuse holder. The high-speed gas jet extinguishes the arc. Simultaneously the spring-loaded connecting wire pulls the burned fusible element out and triggers the fuse-holder opening mechanism. The pivoted fuse holder rotates to the open position.

Overload current melts the thin wire, which produces gases by heating the thin plastic tube surrounding the fuse link. The generated gas extinguishes the arc. The spring-loaded connecting wire pulls the fusible link out and triggers the fuse-holder opening mechanism. The fuse holder swings down to the open position. Most fuses of this type interrupt the current at a zero crossing.

In addition, current-limiting fuses are used for the protection of medium-voltage (15-kV) distribution transformers. These fuses interrupt current safely above three times the rated current but start to melt around one and a half times the rated current. This fuse limits the large short-circuit current because it interrupts the current before the current reaches its peak value.

In order to assure safe current interruption in the full operating range, a type of fuse has been developed with two fuse elements in one housing: a current-limiting element and an expulsion element connected in series. The expulsion element operates at low current, and the current-limiting element operates at high current. Another solution is the use of gas-emitting material in the current-limiting fuse. The long-duration, low-current arc produces gas, which

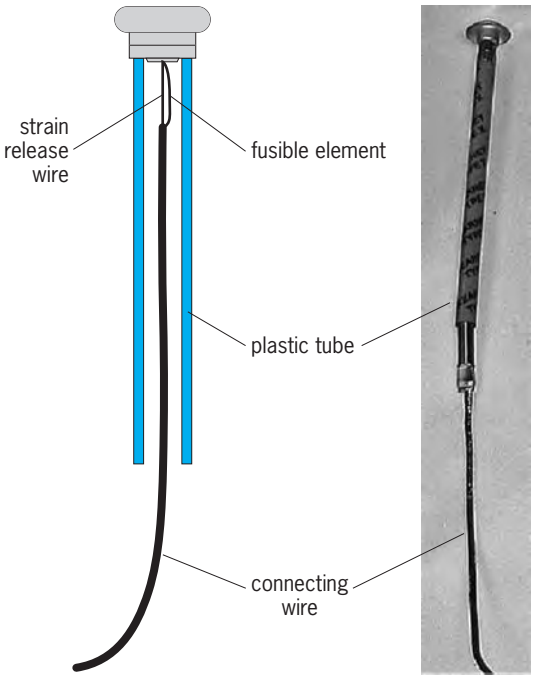

Fig. 6. Cross section and photo of the fuse link.

expands, blows out the arc, and interrupts the current. *See* CIRCUIT BREAKER; ELECTRIC PROTECTIVE DEVICES.

George G. Karady

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; P. Kiameh, *Electrical Equipment Handbook*, McGraw-Hill, 2003; R. E. Loyd, *Electrical Raceways & Other Wiring Methods: Based on the 2005 National Electric Code*, 5th ed., Thomson Delmar Learning, 2005; A. J. Pansini, *Electric Power Equipment*, 1988; B. Peoples, AC versus DC fusing, *Electrical Insulation Conference and Electrical Manufacturing & Coil Winding Conference: 2001 Proceedings*, pp. 239–243, Oct. 16–18, 2001; R. Peters, *Electrical Basics*, Sterling Publishing, 2003; A. Wright and P. G. Newberry, *Electric Fuses*, 3d ed., 2004.

Fused-salt phase equilibria

Conditions in which two or more phases of fused-salt mixtures can coexist in thermodynamic equilibrium. Phase diagrams of these equilibrium conditions summarize basic knowledge about fused salts. Numerous advances in the technologies which are based on high-temperature chemistry have become possible through the increase in knowledge about fused salts. The increasingly significant role of fused salts in industrial processes is evident in the widening application of these materials as heat-transfer media, in extractive metallurgy, in nonaqueous reprocessing of nuclear reactor fuels, and in the development of nuclear reactors which create more fuel than they consume (breeder reactors); moreover, these technologies are all direct outgrowths of research and development with fused salts. *See* NUCLEAR REACTOR; PHASE EQUILIBRIUM.

Fused-salt media were of great interest before 1900 for the electrolytic reduction of cations to metals. The bulk of research on fused-salt phase equilibria has stemmed from interest in electrometallurgical processes from that time until the mid-twentieth century. In the 1950s and 1960s the U.S. Atomic Energy Commission's programs showed that molten-salt mixtures may serve successfully as nuclear reactor fuels, fuel solvents, coolant fluids, moderators, nuclear reactor breeder blankets for the preparation of plutonium or uranium-233, high-temperature bearing lubricants, and fuel-reprocessing media.

Fused-salt mixtures find application in technology when the need arises for liquids which are stable at high temperatures. For most applications, suitably low melting temperatures and low vapor pressures are primary considerations. To some extent these requirements are conflicting, because salts which are useful in obtaining low freezing temperatures often tend to have appreciable covalent character and therefore to exhibit unfavorably high vapor pressures.

As a special class of liquids, one which is composed entirely of positively and negatively charged ions undiluted by weak-electrolyte supporting media, fused salts are used in many different types of research. For example, advances in solution theory, thermodynamics, and crystal chemistry have come about through studies of fused-salt systems. *See* FUSED-SALT SOLUTION.

A close connection between fused-salt phase diagrams and geochemistry exists and stems from the model principle developed by V. M. Goldschmidt, who noted that isomorphic structures are assumed by ions of the same proportionate size and stoichiometric relations but of different charge. Thus the fluorides of beryllium, calcium, and magnesium, for example, are structural models for silicon dioxide (SiO_2), titanium dioxide (TiO_2), and zirconium dioxide (ZrO_2). The fluoride structures are referred to as weakened models because of the smaller electrostatic forces resulting from smaller ionic charges; they have been useful for comparisons with oxide and silicate systems. According to Goldschmidt's interpretation, saltlike materials were derived from components such as water (H_2O), carbon dioxide (CO_2), sulfur trioxide (SO_3), chlorine (Cl_2), and fluorine (F_2), which were volatilized from molten magmas as they crystallized. Crystallization equilibria in fused-salt systems therefore provide a convenient way to study the mechanisms occurring in the formation of igneous rocks. Such investigations were once pursued vigorously, producing the basic techniques which have been employed routinely in phase studies, essentially since the enunciation of the phase rule by J. W. Gibbs. *See* IGNEOUS ROCKS; LITHOSPHERE; PHASE RULE.

Methods of investigation. For many years scientists in the United States and elsewhere expended a great deal of effort in attempts to predict the phase relationships of uninvestigated systems; however, the necessary thermochemical data which were required, such as heat capacities of the constituents, the variation of this property with temperature, the heats of fusion, and the free energies of mixing, were not available. It soon became evident that the acquisition of these data was at least as arduous a task as the determination of the actual equilibrium behavior, and that quantitatively the phase relationships predicted from the thermochemical data were too imprecise to make the method attractive. As a result, it has not yet become routine to estimate with practical accuracy the behavior of untested systems. Accordingly, the determination of equilibrium phase diagrams has remained principally in the region of experimental science. To be sure, numerous correlations are available which can assist the experimentalist in minimizing tasks such as determining the interrelationships of crystal structure, thermodynamic quantities, and the predictable degree of ideality of some ionic systems. Such information makes it possible for phase studies to be conducted with greater efficiency and rapidity, and therefore tends to maintain the field as experimental. The largely experimental character of fused-salt phase equilibria is even more a consequence of the development of experimental techniques which permit the delineation of

phase relationships in fused-salt systems far more accurately and rapidly than was previously possible.

Innovations in the original means of study have brought about the application of a wide variety of experimental techniques to studies of heterogeneous phase equilibria. These include differential thermal analysis, visual methods for direct observations of melting and freezing behavior at high temperatures, filtration of partially molten mixtures, high-temperature centrifugation, hot-stage microscopy, and high-temperature x-ray diffraction. Electromotive-force measurements and vapor-pressure determinations have also been used with success in the study of heterogeneous equilibria in salt systems.

One of the most important developments among the experimental devices used for solid-state phase studies was the automatic x-ray diffractometer for single crystal structure studies. With automatic diffraction equipment, the average period for determination of new structures is about one-tenth that required previously. With the assistance of computerized equipment, it is commonplace to measure automatically the intensities associated with several thousand reflections from crystal planes and to deduce the structures with significantly greater accuracy than was obtained previously. This has a special function in phase studies: By concurrently determining the equilibrium relations in a system and the structure of the crystalline solids which are found in the system, the stoichiometry, compositional variability, coordination chemistry, and so on, are established unequivocally.

Types of equilibria. In use, fused salts are generally employed under conditions where temperature and composition are varied extensively. Drastic changes in equilibrium phase relationships are thus commonplace. In order to understand how these salts behave chemically, it is necessary to gain a rudimentary knowledge of heterogeneous equilibria. Pragmatically, this approach need not entail a complete nor even extensive study of the systems involved; it will be useful as a framework which prescribes the conditions that can be employed in use and in related investigations with fused salts. Characterization of salt phase equilibria is simple and straightforward with binary and ternary systems, but may become increasingly intricate and of decreasing utility as the effect of additional components must be considered. Considerations here are therefore limited to condensed equilibria in binary and ternary systems.

Binary systems. Consider first the binary system A-B. As drawn (**Fig. 1**) the phase diagram indicates that the components A and B react to form the intermediate compounds A_2B and AB. Limited miscibility in the solid state occurs when mixtures of A and A_2B are present at equilibrium at temperatures between T_1 and T_2, and T_3 and T_4. The compound AB melts incongruently to A_2B and liquid. Crystals of AB are dimorphic; at temperatures above T_5 they exist in a crystalline state which transforms, on cooling below temperature T_5, to a less symmetric (lower-energy form) crystalline state. Liquids formed from mixtures of AB and B are immiscible at compositions such as C_1 and C_2, which separate into immiscible liquids of composition described by horizontal tie lines in the two-liquid region shown in Fig. 1.

Ternary systems. The temperature-composition relations in condensed ternary systems are generally represented by a polythermal projection of the liquidus surfaces. This is a projection parallel to the temperature axis on the triangular composition plane; it shows, therefore, the various parts of the liquidus surface or surfaces as viewed from the direction of high temperature. Such diagrams indicate the compositions of liquids saturated with one or more solids. At the invariant reaction points, liquids are in equilibrium with three solids. The direction of temperature change is shown by arrows on boundary curves, and liquidus temperatures are shown by isothermal contours.

The hypothetical ternary system shown in **Fig. 2** displays behavior which is typically described by three-component phase diagrams. Assume that invariant equilibrium points in the phase diagram A-B-C will ultimately be located as shown in Fig. 2 and that there are no intermediate compounds formed from components B and C; a single intermediate compound AC_2, melting congruently, occurs in the system A-C; the congruently melting compound AB_2 and the incongruently melting compound A_2B are formed from components A and B. A careful examination of several crystalline mixtures cooled from the liquid state can be expected to furnish much information about equilibria in the system A-B-C. Some examples are given below.

1. All crystallized specimens having compositions within the triangle AB_2-B-C contain only the pure solid phases AB_2, B, and C. These results indicate that the compounds AB_2, B, and C form a subsystem, that AB_2 probably melts congruently, and that the section

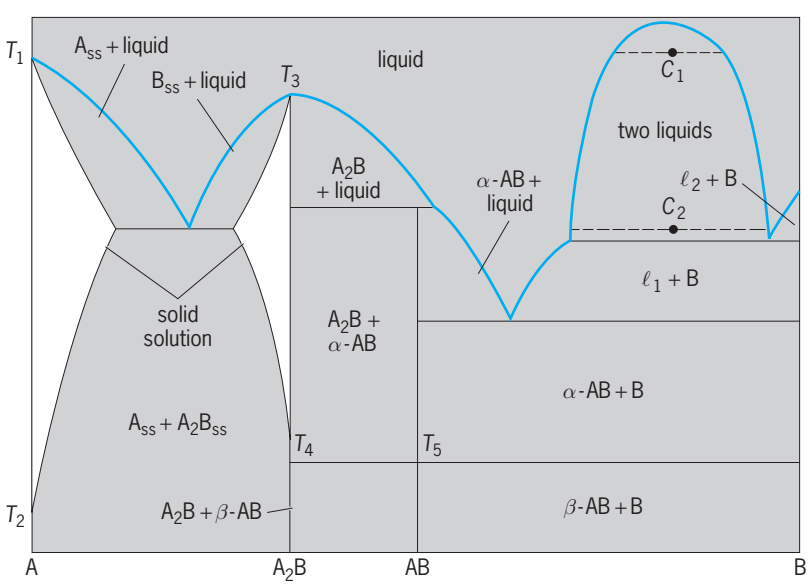

Fig. 1. Diagram of typical phase equilibrium relationships in condensed binary systems of molten salts. Here ℓ is liquid and ss is solid solution.

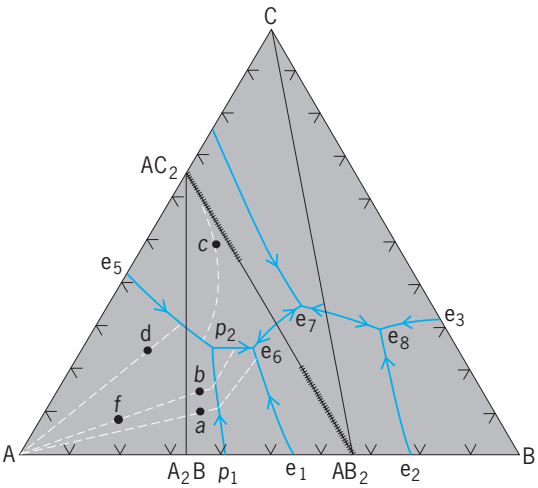

Fig. 2. Typical phase equilibrium relationships in condensed ternary systems of molten salts.

AB_2-C is a quasi-binary system. Under these circumstances, a single invariant point, the eutectic e_8, can occur in AB_2-B-C. It may be possible to estimate its approximate composition by examination of the crystallized mixtures. Morphological data, obtained from petrographic or metallographic methods, may reveal the domains of the primary phases AB_2, B, and C, and the composition of the eutectic.

2. Assume that AC_2 and AB_2 form extensive but limited solid solutions exhibiting maximum solubility at the solidus, as shown in Fig. 2. In general, lattice constants of AB_2 and AC_2 will change as a function of solute concentration. Therefore, x-ray data obtained from crystallized specimens having compositions intermediate between AB_2 and AC_2 can be examined for lattice-spacing changes and correlated employing the parametric method. Some exsolution (precipitation of crystals from the solid solution) is to be expected as temperatures are lowered. The exsolution curve often exhibits considerable slope and precludes the possibility of obtaining data which indicate concentrations of maximum solubility of AB_2 and AC_2 except by static methods.

3. Occurrence of solid solutions of AB_2 and AC_2 indicates that inferences regarding the subsystem AB_2-AC_2-C can be made from phase analyses of mixtures having compositions in triangle AB_2-B-C in addition to those in AB_2-AC_2-C. Some mixtures in triangle AB_2-AC_2-C which crystallize primary-phase AC_2 solid solutions will solidify before the liquid fraction attains the composition e_7 and may not contain detectable concentrations of AB_2 or C. Mixtures which crystallize primary-phase AB_2 solid solutions will behave similarly, producing possibly undetectable amounts of AC_2 or C. The identity of subsystem AB_2-AC_2-C may thus be equivocal without the conclusion based on other data that AB_2-C is a quasibinary.

4. Much information about phase behavior with the triangles A_2B-AC_2-AB_2 and A-A_2B-AC_2 can be obtained from x-ray analyses of specimens crystallizing within triangle A_2B-AC_2-AB_2. The solid-solution formation along AB_2-AC_2 and the melting relations of A_2B will cause mixtures with A_2B-AB_2-AC to undergo a variety of crystallization reactions. Depending on the character and extent of the solid solution and the melting-freezing behavior of A_2B, x-ray analysis may furnish very informative phase data. Crystallization in polycomponent systems often results in the formation of final products which coexist in an unstable equilibrium. Identification of the crystalline phases in the final crystallized mixture provides useful inferences as to the nature of the equilibrium phase diagram. *See* EUTECTICS.

Applications in research and development. Two general categories encompass the principal activities in phase equilibrium investigations, those which center about technological applications and those which are essentially fundamental science.

The principal applications in which fused salts are employed in technology and research are listed below. As is implied in the foregoing paragraph, the division between technology and research for a number of the listed categories is artificial.

Technology

1. Reductive extraction processes for metals production
   a. Electrochemical processes: (1) Sacrificial an ode methods: tantalum (Ta), niobium (Nb), tungsten (W), and molybdenum (Mo). (2) Electrowinning methods: lanthanides, aluminum (Al), transition metals.
   b. Pyrochemical processes: lanthanides actinides
2. Working fluids for molten-salt breeder reactor development
   a. Fuels, based on uranium fluoride (UF_4), thorium fluoride (ThF_4), plutonium fluoride (PuF_3)
   b. Blankets for breeder reactors
   c. Coolant salts
3. Solvents for use in volatility processes for recovery of uranium from spent fuel elements
4. Nonnuclear heat-exchange media: nitrates, fluoroborates, carbonates
5. Application to regenerative fuel cells
6. Metal alloys–salt compatibility
7. Development of analytical chemical procedures

Research

1. Coordination chemistry
   a. Lanthanide-actinide model relationships
   b. Complex ions in solutions
   c. Electrochemical research
   d. Transition-metal chemistry
   e. Structure studies
2. Thermodynamics of solution behavior
   a. Freezing-point depression
   b. Liquid-liquid immiscibility
   c. Reciprocal systems
3. Optically active crystals
   a. Lasers
   b. Electronic theory

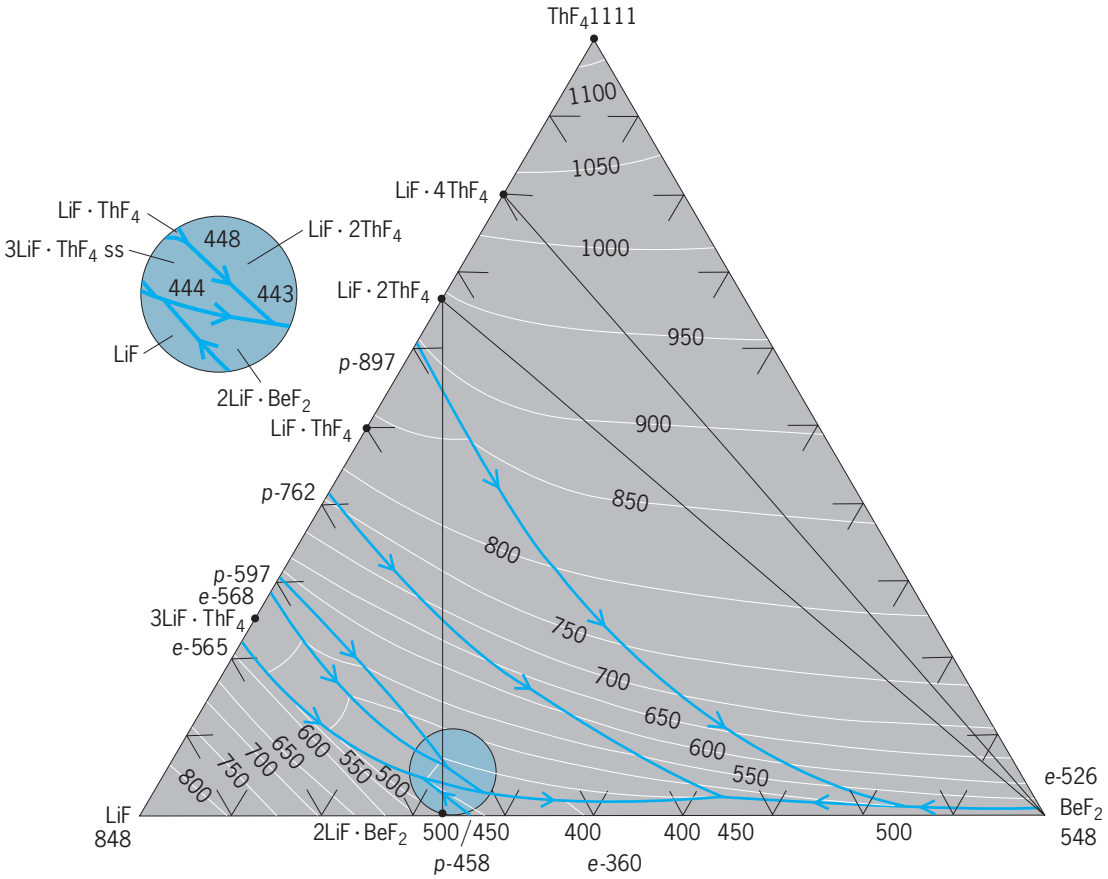

Fig. 3. The system LiF-BeF_2-ThF_4. Temperature is in °C. °F = (°C × 1.8) + 32. (*After R. E. Thoma et al., Phase equilibria in the systems BeF_2-ThF_4 and LiF-BeF_2-ThF_4, J. Phys. Chem., 64:865–870, 1960*)

4. Crystal growth
   a. Single crystals of oxides and fluorides for structure studies
   b. Zone refining
   c. Phonon scattering
5. Metal-metal halide systems
6. Source of new types of molten-salt research

Perhaps the principal factor which distinguishes technological and research applications of molten-salt phase equilibrium studies is that the materials problems in technology often necessarily involve multicomponent systems, whereas the number of components in fundamental research studies is limited by choice. In technology the processes are frequently highly complex reactions involving simultaneously metal, gas, salt, and oxide or other phase reactions as well as oxidation-reduction equilibria. One application, involving a variety of fused-salt systems, has been in molten-salt breeder reactor technology.

In the United States during the 1960s, efforts to develop advanced-design nuclear reactor technology included those devoted to a single-region fluid-fuel reactor. Advances in the chemistry of processing molten salts to remove protactinium and the lanthanides indicate that development of a large, 1000-MW (electrical), single-fluid molten-salt breeder reactor would be feasible. Thermal energy generated in the reactor fluid would be transferred to a secondary coolant-salt circuit, which couples the reactor to a supercritical steam cycle. The results of design studies showed that the salt system should be composed of lithium fluoride (^{7}LiF), beryllium fluoride (BeF_2), ThF_4, and $^{235}UF_4$, with concentrations optimized by nuclear considerations and constraints imposed by phase equilibria and physical properties. The phase diagram of the ^{7}LiF-BeF_2-ThF_4 system (**Fig. 3**) then dictated the choices which were available for selecting the optimal composition of the salt mixture. For minimum liquidus temperatures, minimum viscosity of the molten salt, and optimal nuclear performance, the composition ^{7}LiF-BeF_2-ThF_4-$^{233}UF_4$ (68:20:11.2:0.2 mole %) was selected provisionally. Although development of the technology was terminated, its feasibility remains viable. Any future development of the fuel and reprocessing technology will entail optimizations that consider the phase behavior of fused-salt systems that are candidates for application. *See* NUCLEAR POWER. R. E. Thoma

Bibliography. J. Braunstein et al., *Advances in Molten Salt Chemistry*, vols. 1–4, 1971–1981; L. Eyring (ed.), *Progress in the Science and Technology of the Rare Earths*, vol. 2, 1966; D. Inman and D. G. Lovering (eds.), *Ionic Liquids*, 1981; G. J. Janz, *Molten Salts Handbook*, 1967; I. M. Kolthoff (ed.), *Treatise on Analytical Chemistry*, pt. 1: *Theory and Instrumentation*, 1961; J. Lumsden,

Thermodynamics of Molten Salt Mixtures, 1966; G. Mamantov et al. (eds.), *Advances in Molten Salt Chemistry*, vols. 5 and 6, 1983–1987; G. Mamantov and R. Marrasi (eds.), *Molten Salt Chemistry: An Introduction and Selected Applications*. 1987.

Fused-salt solution

A nonaqueous solvent system particularly useful in the production of reactive metals and in the study of the coordination chemistry of the elements. Fused salts are a large class of liquids that are composed largely of ions. Many simple inorganic salts melt at high temperatures (greater than 600°C or 1100°F), forming liquids which have high specific ionic conductivities (1–6 ohm^{-1} cm^{-1}). There are, however, a number of exceptions; for example, the electrical conductivity of $AlCl_3$ decreases sharply upon melting due to the formation of the molecular liquid Al_2Cl_6. Mixing molecular fused salts with truly ionic salts results in the formation of so-called potential electrolytes. Each molten salt system constitutes its own solvent system, and needs the establishment of its own electromotive force series and requires, in some cases, the development and use of special reference electrodes. *See* ELECTROLYTIC CONDUCTANCE; FUSED-SALT PHASE EQUILIBRIA; HIGH-TEMPERATURE CHEMISTRY; SOLVENT.

The use of binary or ternary melt compositions results in liquids that typically have much lower melting temperatures than do the starting constituents and somewhat lower specific ionic conductivities than do pure salts. The choice of melts for use as solvents frequently is based on considerations such as the availability and cost, the lowest melting temperature attainable, and the ease of purification of the solvent, as well as the width spanning the electrochemical stability and the spectroscopic transparency of the melt. Several molten-salt solvents, such as the LiCl-KCl eutectic and the equimolar $NaNO_3$-KNO_3 melt, have been extensively studied, and many of their physical and chemical properties are well known. Fused salts used in the industrial production of reactive metals are usually multicomponent systems, with each component added to the melt in order to improve a specific property. Many melt systems, particularly ternary and compositions that are more complex, have been only partially characterized; their physical properties are estimated via modeling by using the data available for less complex systems, such as the binary component melts.

Applications. Fused-salt solutions have been used extensively as electrolytes in the production and refining of aluminum (oxyfluoride solutions), magnesium (chloride solutions), lithium (chlorides), and calcium (chlorides) metals. Fused salts have also been used as solvents in the production by thermite (reduction) reaction of high-melting metals such as titanium, zirconium, niobium, tantalum, rare earths, actinides, and alkali metals. Some useful high-temperature new materials, such as titanium diboride and lanthanum hexaboride, also have been synthesized in high purity using fused-salt solutions. One of the largest commercial applications of fused salts is for the heat treatment of metals and alloys, where fused salts are used as baths for their heat dissipation capacity.

Emerging applications. Fused salts have been investigated for use in rechargeable batteries and in fuel cells because they allow the use of higher currents than aqueous systems.

Potential applications. Fused-salt systems are well suited for the environmentally sound treatment of wastes and recycling processes. Fused salts also have been investigated for the formation of synthetic carbonate rocks as a potential permanent storage form of the greenhouse gas carbon dioxide. In addition, low-temperature fused-salt systems, called room-temperature molten salts, can provide the advantages of both the fused salts and the aqueous systems. Room-temperature fused salts consist of a mixture of metal halide salts and organic halides. The most thoroughly studied systems contain aluminum trichloride and alkylpyridinium halides or alkylimidazolium chlorides. Room-temperature molten-salt systems are suitable for electroplating reactive metals.

Chloroaluminate systems. For some melt systems, such as those containing aluminum halides, the melt composition can be adjusted to change the acid-base properties of the solvent, which determine the chemistry of the dissolved solutes. The $AlCl_3$-NaCl system has been studied extensively and will be used to illustrate the concepts involved. An acid (Lewis acid) in this system is defined as a chloride ion acceptor and $pCl^- \equiv -\log [Cl^-]$. (This concept is similar to pH.) The modified Lewis acidity is frequently used in discussing these systems. Raman spectral measurements have shown that the $AlCl_4^-$ ion is the predominant anion in the equimolar $AlCl_3$-NaCl melt. As the $AlCl_3$/NaCl molar ratio is made greater than one, the melt becomes acidic and the concentration of the $Al_2Cl_7^-$ ion increases. Near the equimolar point, the dominant equilibrium is that shown in reaction (1).

$$2AlCl_4^- \rightleftharpoons Al_2Cl_7^- + Cl^- \qquad (1)$$

The mole fraction equilibrium constant for this equilibrium has been determined from potentiometric measurements as 1.1×10^{-7} at 175°C (347°F). The most basic sodium chloroaluminate system corresponds to a NaCl-saturated melt in which the pCl^- is constant in the presence of solid NaCl and is equal to 1.1 at 175°C (347°F). Two other equilibria [reactions (2) and (3)] become important as the melt acidity is

$$3Al_2Cl_7^- \rightleftharpoons 2Al_3Cl_{10}^- + Cl^- \qquad (2)$$

$$2Al_3Cl_{10}^- \rightleftharpoons 3Al_2Cl_6 + 2Cl^- \qquad (3)$$

increased by increasing the $AlCl_3$ content. Reaction (1) is analogous to the dissociation of water [reaction (4)]. Similar equilibria have been proposed for other

$$2H_2O \rightleftharpoons H_3O^+ + OH^- \qquad (4)$$

melt systems containing a covalent halide, such as $FeCl_3$ or $SbCl_3$.

A major difference between water (and related solvents) and the chloroaluminate solvent system is the ability of the latter to stabilize unusually low oxidation states of metals in acidic melts. Thus ions such as Cd_2^{2+}, Bi^+, and Hg_3^{2+} are quite stable in $AlCl_3$-rich melts. This unusual chemistry has been attributed to the small tendency for disproportionation of the low oxidation state. Such disproportionation is favored by large concentrations of halide ions which complex preferentially the higher oxidation state. For example, reaction (5) is favored by the presence of chloride ions; the chloride concentration can be made small by adding a Lewis acid such as $AlCl_3$. The chemistry in molten chloroaluminates bears considerable similarity to superacid systems, such as HSO_3F-SbF_5. Thus sulfur, selenium, and tellerium in both media form homopolyatomic cations such as S_8^{2+} and Se_4^{2+}, as well as poorly characterized radical species. *See* CHEMICAL EQUILIBRIUM; SUPERACID.

$$Cd_2^{2+} \rightleftharpoons Cd + Cd^{2+} \qquad (5)$$

It should be stressed that small or moderate changes in melt composition can have a drastic effect on the chemistry of solutes in the chloroaluminate melts. For example, Ti(II) is a stable oxidation state in melts quite rich in $AlCl_3$, such as $AlCl_3$-NaCl (65–35 mole %). In less acidic melts, Ti(II) forms a precipitate, while no evidence for the formation of divalent titanium has been obtained in melts that contain an excess of NaCl. Similarly in alkali fluoride melts, only Ti(IV) and Ti(III) exist; the reduction of trivalent titanium leads directly to titanium metal. The use of fluoride media results in greater stability of the high oxidation states because many metal ions in these oxidation states form very stable complexes with fluoride ions.

It is interesting to contrast the reduction of refractory metal ions, such as those of niobium, tantalum, and tungsten, in the basic alkali fluoride media—for example, the LiF-NaF-KF eutectic (Flinak)—and in the acidic chloroaluminate melts. In Flinak, the reduction proceeds through several steps (some of which are irreversible) to the metal, which may be obtained in coherent nondendritic form. In $AlCl_3$-rich chloroaluminate melts, metal clusters, such as W_6Cl_{12}, are readily formed and the formation or plating of the metal may not be possible. On the other hand, the cluster species that is formed may be of interest in other applications, such as catalysis. It has been reported that $Ir_4(CO)_{12}$ in acidic $AlCl_3$-NaCl melts catalyzes the methanation reaction (formation of CH_4 from CO and H_2).

Catalysts. One use of fused salts is as media for organic reactions. Not only does the fused-salt environment provide for better thermal control of the reaction (heat dissipation is readily possible), but it also may serve as a catalyst. For example, molten $SbCl_3$ and $ZnCl_2$ have been found to be effective hydrocracking catalysts for coal. It has been found that polycyclic hydrocarbons, such as anthracene, undergo several types of reactions in $SbCl_3$ melts, including formation of radical cations and protonated species that react further to form condensed systems, such as anthra[2.1-*a*]aceanthrylene. Other examples are the chlorination of benzene in NaCl-ACl_3-$FeCl_3$, the formation of methyl chloride in KCl-$ZnCl_2$, the synthesis of ethyl chloride in $AlCl_3$-KCl-NaCl, and the preparation of silicon hydride in KCl-LiCl. *See* CATALYSIS.

Batteries and fuel cells. Another technological area of great interest in which molten salts play a key role is that of advanced batteries and fuel cells. LiCl-KCl eutectic is the solvent in the rechargeable Li(Al)-FeS (or FeS_2) battery, and sodium polysulfide melts are used in the sodium/sulfur battery which operates at about 350°C (660°F). The rechargeable cell Na/Na^+ conductor/S in $AlCl_3$-NaCl has an open circuit voltage of 4.2 V at 200–250°C (390–480°F) as well as high energy densities. Molten carbonates, such as the ternary Li_2CO_3-Na_2CO_3-K_2CO_3 melt, are used in fuel cells that use H_2/CO mixtures and oxygen as electrochemical fuels. *See* BATTERY; FUEL CELL.

Georges J. Kipouros; Gleb Mamantov

Bibliography. R. J. Gale and D. G. Lovering (eds.), *Molten Salts Techniques*, vols. 1–4, 1991; D. G. Lovering (ed.), *Molten Salt Technology*, 1982; J. Plambeck, *Encyclopedia of Electrochemistry of Elements X: Fused Salt Systems*, 1976; P. Wasserscheid and T. Welton (eds.), *Ionic Liquids in Synthesis*, 2003.

Fuselage

The component of an aircraft that provides the payload containment and the structural connection for the wing and the empennage (tail assembly). The fuselage and the wing are major structural components of an aircraft. The fuselage is the counting structure for the horizontal and tail surfaces that provides stability as well as the means of introducing pitch and yaw control to the aircraft. For some aircraft like fighter and private aircraft, the fuselage houses the engine or engines. The nose or tail gear and the main landing gear are often attached to the fuselage structure. *See* WING.

The fuselage is the main payload structure for all aircraft. The payload can consist of only crew of pilot, electronics, and armament for a fighter, or can contain crew, passenger attendants, passengers, cargo, and baggage for a commercial passenger aircraft. All high-performance jet aircraft obtain optimum performance at high altitudes. This requires a pressurized fuselage, whether it is only a pilot compartment or the total fuselage of a commercial passenger aircraft. The pressurization requirements, the payload support, the wing, empennage, and landing gear connections, and the flight regime establish the structural loading environment for the design of the structural elements of the fuselage. For each aircraft configuration or type, the aircraft's functional requirements will dictate the location, shape, and efficiency of these structural elements.

Configuration. The fuselages of fighter aircraft are usually designed to provide a minimum envelope for the pilot, engines, fuel, electronics, and armament, with the additional requirement of easy access for combat maintenance. By maintaining a minimum external envelope, the fuselage structural weight and aerodynamic drag are minimized. A similar minimum envelope is usually maintained for bomber aircraft, with additional internal volume being required for bomb payload. In both these aircraft the exteriors of the fuselage are tailored carefully to maximize high-speed aerodynamic performance. This fuselage aerodynamic tailoring results in complex external contours and internal structure which is usually more expensive to produce than the structure found in cargo, commerical, or private aircraft. *See* MILITARY AIRCRAFT.

Cargo aircraft fuselages are designed to maximize internal volume and to provide maximum convenience in cargo loading. For military cargo aircraft, the fuselage configurations are basically all similar to those shown in **Fig. 1**.

The fuselage configurations for all large commercial passenger aircraft are established by passenger payload requirements. For those cargo aircraft which are derivatives of commercial passenger aircraft, the fuselage configurations are also controlled by the passenger requirements. The passenger aircraft fuselage configurations are generally divided into two separate sections: the passenger compartment above the floor, and a lower belly cargo compartment. The design objective for private aircraft fuselages which are not pressurized is to provide a simple envelope which is easy to manufacture at minimum cost.

Material and fabrication. The history of the construction of aircraft fuselages has evolved through the early wood truss structural arrangements to the current metal semimonocoque shell structures. Currently a majority of aircraft fuselages are fabricated from aluminum alloys and are produced by a process of automatic machining of the skins and stringers, with much of the assembly being done by automatic drilling, countersinking, and fastener installation. In some areas, adhesive bonding is used as a means of attaching doublers to reinforce skin panels. The Fokker F-27 and F-28 commercial aircraft use extensive adhesive bonding in fuselage construction. Adhesive bonding is also used in the fabrication of the honeycomb sandwich structure, which has been used in the fuselage of General Dynamic's F-111 fighter.

In many of the high-performance aircraft, such as fighters and bombers, extensive use is made of titanium and high-strength steel. The use of numerical machining has allowed the application of these materials to complex bulkheads, fittings, major structural elements, and attachment points. In the fabrication of titanium, there are methods which use diffusion bonding and superplastic forming to reduce the cost

Fig. 1. Military cargo aircraft, showing fuselage configurations. (*a*) C-130 Hercules, (*b*) C-141 Star Lifter. (*c*) C-5A Galaxy. (*Lockheed Corp.*)

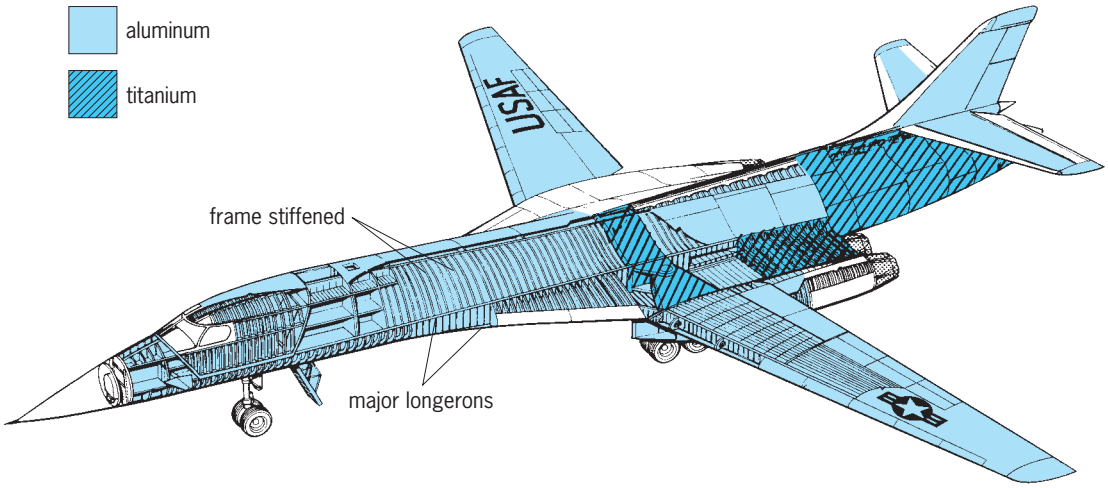

Fig. 2. B-1 bomber fuselage structural configuration, based on frame-stiffened skins with major longerons.

of titanium part fabrication. Other manufacturing techniques include the use of castings, extrusions, forgings, and welding.

Structural configuration. The fuselage structural configuration of modern aircraft is dominated by two generic types of shell structures. One of these consists of a frame-stiffened skin construction with major longerons supporting the bending loads. The other structural configuration is composed of stringer-stiffened skins supported by frames. These two configurations are utilized separately or in combination to produce the most efficient structural configuration.

Fighters and bombers are usually configurated of complex aerodynamic shapes which must provide maximum access to equipment and armament. The best structural design of these fuselages, for functional purposes, is the frame-supported skins with a few major longerons. For this structure configuration the skins carry shear, while the major bending material is located in either the major longerons or the keel-beam-type support structure (**Fig. 2**).

The second type, the semimonocoque structural configuration in which the stringer-stiffened skin shells are supported by frames, is the most efficient for the large near-circular fuselage bending loads (**Fig. 3**). *See* AIRFRAME.

Flight regime requirements. Once the aircraft's configuration has been established and the external contours of the aircraft defined, the designer needs information on the flight regime requirements such as the flight loads, pressurization level, and structural temperature range, the last associated with either the flight regime or the installation of equipment or engines.

Flight loads. The requirements of flight loads applied to fuselages can be established by a static balance of lift, drag, and balancing tail loads along with the loads imposed by payload or equipment distributions. The loading diagram in **Fig. 4** illustrates the loads applied to the aircraft and how they impose simple beam loads on the fuselage. The flight load requirements vary in magnitude depending on whether they are introduced by aircraft maneuvers, gust loads, or landing loads, and each must be considered in order to design all the structural elements of the fuselage. Concentrated loads are introduced at the attachment points of wings, empennage, engines, or landing gear. There are also flight control, equipment, and passenger or pilot accommodations equipment attachment loads.

Pressurization. Once the structural elements have been designed for either inertia or equipment-induced loads, the other major load consideration is the pressurization requirements. Pressurization loads are generated by the differential between the cabin pressure employed to sustain a comfortable atmosphere for the pilot or passengers and the ambient

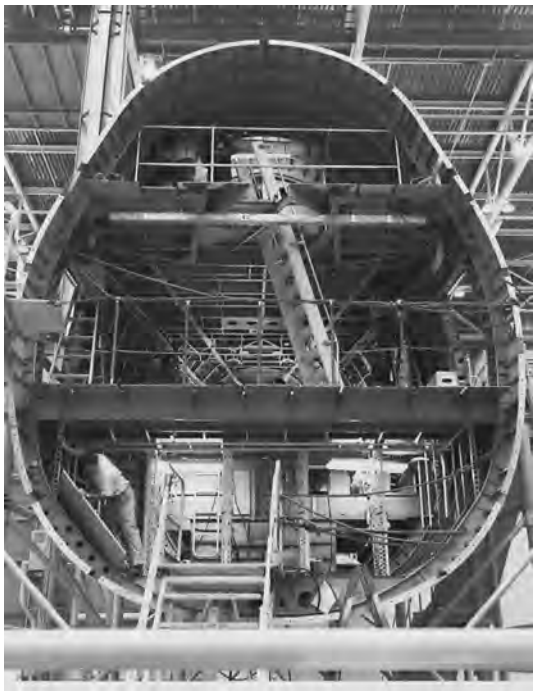

Fig. 3. Boeing 747 fuselage with stringer-stiffened skin supported by frames. (*Boeing Co.*)

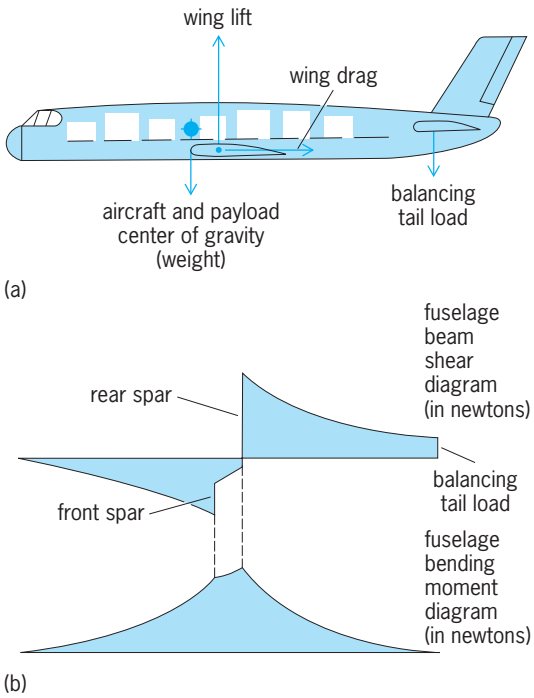

Fig. 4. Flight loads. (*a*) Aircraft static balance. (*b*) Fuselage simple beam loading.

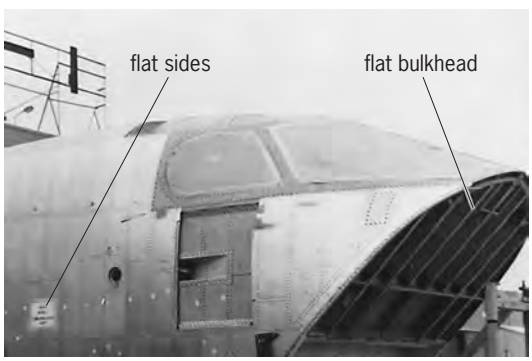

Fig. 5. Pressurized crew compartment of B-1 bomber.

altitude pressure associated with the flight regime. Since fighter and bomber fuselages are configurated for a minimum envelope and maximum aerodynamic efficiency, their shape may not be optimum for carrying pressure. A pressurized compartment containing flat sides may impose significant local internal structural loads since it is the most difficult structural configuration for sustaining pressure loads (**Fig. 5**). Fuselage configurations for large commercial transports are circular or near circular since this configuration is optimum for pressure containment. In the thin-shell fuselage structures this allows the skin to carry the pressure loads in hoop tension (**Fig. 6**) without introducing any significant loads in frames.

Temperature. The temperature requirements are associated with two potential sources of heat. One is the temperature change imposed on the structure by equipment such as engine exhaust, air-conditioning ducts, or armament discharge areas. The other structural temperature consideration is due to the ground environment and flight regime for either subsonic or supersonic performance requirements. The typical temperatures considered for the subsonic aircraft are those of ambient ground temperature plus solar heating and the −65°F (−54°C) temperature of high-altitude flight. For fighters and bombers and the British and French Concorde, the flight regime imposes aerodynamic heating temperatures up to 300°F (150°C). For this elevated-temperature supersonic flight regime, special design considerations must be taken into account when selecting the material to be used in the fuselage construction, as well as in the design of the structural details, to accommodate the temperature-induced deformation. *See* THERMAL STRESS.

Material selection. The above design requirements imposed by the flight and equipment loads, pressure containment, and the operational temperature range are all considerations in selection of the material to be used in a particular structural detail of the fuselage. The designer must consider the basic characteristics of those materials available for current aircraft construction. Aluminum, titanium, higher-strength steel, and some advanced composites such as graphite/epoxy have been used for construction of fighter and high-performance aircraft. A major portion of most cargo and commercial aircraft consists of aluminum, with highly loaded members, such as landing gear beams and engine mounts, of steel or titanium. Control surfaces of commercial aircraft are designed with graphite/epoxy materials. *See* COMPOSITE MATERIAL.

The basic structural characteristics of these materials are that aluminum provides a specific strength-to-weight density ratio on the order of 0.6×10^6 in. (1.5×10^4 m), while steels and titanium are of a similar strength-density ratio. The advanced composites

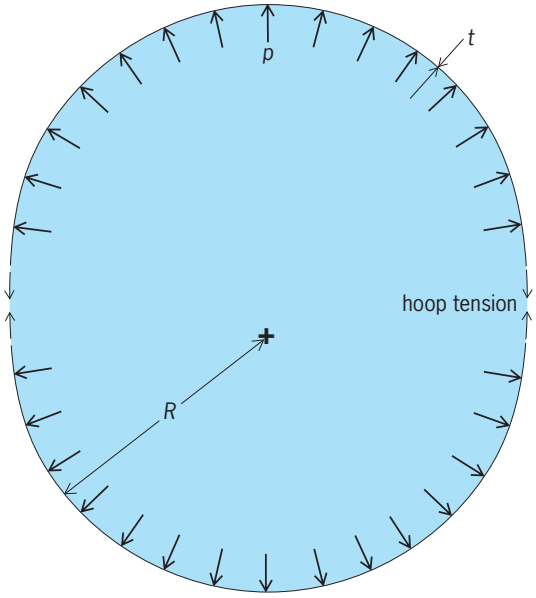

Fig. 6. Hoop tension from pressurization in thin skin shell. Here p = pressure, R = radius, t = skin thickness; hoop tension stress $f_t = pR/t$.

(graphite/epoxy or kevlar/epoxy) offer the advantages of higher strength-density ratios (1.2×10^6 in. or 3.0×10^4 m). *See* ELASTICITY; STRESS AND STRAIN.

Stability and strength. The configuration using the frame-stabilized major longeron and shear panel concept requires the designer to review each structural element for both strength and stability considerations. The frames provide the support, from a strength point of view, for all the local loads introduced by the attachment of wings, empennage, engines, landing gear, and armament, and the local loads induced by compartment pressurization. The longerons provide the main bending stiffness for the fuselage (Fig. 3). The major longerons act as columns between the frames; therefore, the stability of those columns must be a design consideration. *See* BEAM COLUMN.

In the fuselage structural configuration utilizing stringer-stiffened skin acting with frames, the frames stabilize or break up the skin stringer panel lengths. The stiffeners, with the skin, provide the structural stability and bending stiffness for the fuselage shell. This configuration is typical for commercial passenger and cargo aircraft (**Fig. 7**).

Private aircraft are often a combination of the above two types of structure. Major longerons and simple shear panels are used where large cutouts for passenger or cargo access are required. Stringer-stiffened skins are used for the continuous structure connecting the forebody to the empennage.

As discussed above, jet aircraft require structural elements which must support pressure loads. Those aircraft composed of major longerons and shear panels often have flat pressure-containment panels, which use diaphragm action or plate bending to support the pressure loads. In the circular fuselages, the hoop stresses induced in the shell provide pressure containment. Both of these pressure-containment configurations contribute to the fatigue and damage tolerance characteristics of the fuselage.

Fatigue. The fatigue considerations require a review of the repeated loads that are applied to the fuselage structure. These repeated loads are produced by the aircraft maneuvers, gusts, and typical landing and taxi loads. For pressurized areas, an additional requirement of the ground-air-ground pressure cycle must be introduced in the fatigue analysis. This ground-air-ground pressure cycle is often the dominant fatigue consideration.

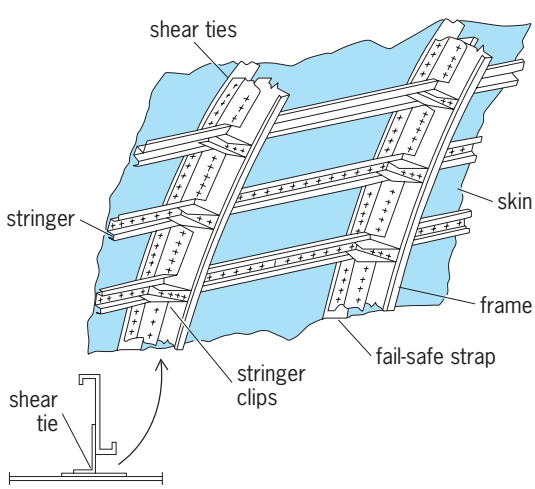

Fig. 7. Typical transport-aircraft aluminum fuselage panel.

Fatigue produced by this cyclic loading is most critical in the complex details (such as joints, splices, fittings, and cutout areas) where stress concentrations may occur at fastener holes, pad-up areas, and load discontinuities. The key considerations for fatigue are (1) the load type and its frequency, (2) the load-induced stress level at the structural detail, (3) the configuration of the detail or the magnitude of the stress concentration, (4) the fatigue characteristics of the material, and (5) the service life requirements. These life requirements vary dramatically, from the shorter usage life requirements of a fighter aircraft (about 6000 flight hours) to the long service life of a commercial passenger aircraft (about 60,000 flight hours). The magnitude of the loads encountered in the maneuver spectrum of a fighter aircraft are much higher than for commercial aircraft because of the fighter's combat maneuver requirements.

Once the elements of fatigue consideration have been assessed, a fatigue analysis is performed based on a damage accumulation rule. This damage accumulation rule is then used to assess the fatigue life of the aircraft.

Since many fatigue parameters are not finitely determinable, fatigue testing of the fuselage structural details and often of the total aircraft is part of the development of any aircraft fuselage. Since fatigue of metallic aircraft parts initiates crack growth in the material, this growth must be understood and its criticality assessed to establish the structure's damage tolerance *See* AIRCRAFT TESTING; METAL, MECHANICAL PROPERTIES OF.

Damage tolerance. Developments in fracture mechanics have provided design and analysis tools adequate to assure the damage tolerance of modern aircraft. Damage tolerance is the ability of an aircraft to tolerate fatigue-initiated flaws as well as external damage. This external damage may be due to such things as engine disintegration or foreign-object damage, and for fighter aircraft may be due to enemy weapons fire. This tolerance to damage provides the safety in modern commercial aircraft which makes them one of the safest means of transportation.

The approach to damage tolerance involves an understanding of the types of flaws that will occur, the rate at which these flaws will grow, the length at which the flaw will become critical for a specific material and structural configuration, and the length at which the flaw will become detectable. A fracture mechanics parameter describing the roughness of a material was developed, called the stress intensity factor. The application of this stress intensity factor allows the development of designs which can tolerate some predetermined-size flaw either by controlling the rate of crack growth or by providing a means of crack arrest. Whatever the means of containment,

the designs must also take into consideration the ability to detect these flaws during inspections for maintenance of the aircraft. The inspection intervals and inspection methods must be established to ensure that a flaw is discovered prior to its reaching critical length.

The flaw arrest procedure generally has been used for the pressurized fuselages of commercial aircraft. Fail-safe straps or shear-tied frames are utilized to contain the crack in one or two bays between the fuselage frames (Fig. 7).

Other design considerations. While the structural requirements for strength, stability, fatigue durability, and damage tolerance are being met in the design of the fuselage, other configuration considerations must be taken in account by the designer, namely access to the interior of the aircraft or its contained equipment, the noise environment associated with the placement of the engines on, around, or in the aircraft fuselage, and maintenance and repair of the fuselage.

Interior access. In fighter fuselages, where the engines are buried in the main fuselage structure, the ability to access and remove these engines is a key consideration in the structure's functional design. In commercial aircraft, loading and unloading of the passengers and the passenger window accommodations are obviously very important.

Engine noise. Many commercial aircraft have engines mounted on the fuselage. The engine noise can induce sonic fatigue loads in the fuselage structural elements. The noise design requirements are met by reducing the noise level at the engine and by taking the noise loadings into consideration in the fatigue design of the fuselage. In addition, appropriate soundproofing is applied to the fuselage shell of commercial aircraft to limit the passenger annoyance.

Inspection, maintenance, and repair. Gross external damage must be detected during walkaround inspections or noted at the time of the occurrence during flight. Specially identified critical areas can be more thoroughly inspected by nondestructive inspection methods to estimate the extent of visible and nonvisible damage. These inspections will also cover those areas most likely to be damaged by mechanical process or atmospheric corrosion. The critical areas requiring special inspection instructions are identified at the time of the aircraft's design. Special inspections may include the removal of internal and external panels and a thorough inspection conducted with the aid of portable nondestructive inspection equipment, such as ultrasonics and x-ray. However, the most easily inspected designs provide visual access to both the external and internal sides of the fuselage shell. *See* NONDESTRUCTIVE EVALUATION.

The fuselage or its subassemblies are not removed from the aircraft for repair, as is often the case for ailerons, rudders, and elevators, or even in some cases empennages. Access to all parts of the fuselage structure for repair therefore is a necessity. Structural configuration compromises must be considered to facilitate repair and maintenance along with the other structural design requirements.

Weight and part reduction. For cargo and commercial aircraft the wing and fuselage are the two largest contributors to the aircraft structural weight, with each contributing 34–37% of the total structural weight. The fuselage weight of large aircraft has grown faster than the wing weight, which suggests that major opportunities for weight saving are in new fuselage design concepts (**Fig. 8***a*). Also because of wing bending-moment relief due to weight in the wing, greater payload improvement is achieved for fuselage weight reduction than for wing weight reduction (Fig. 8*b*).

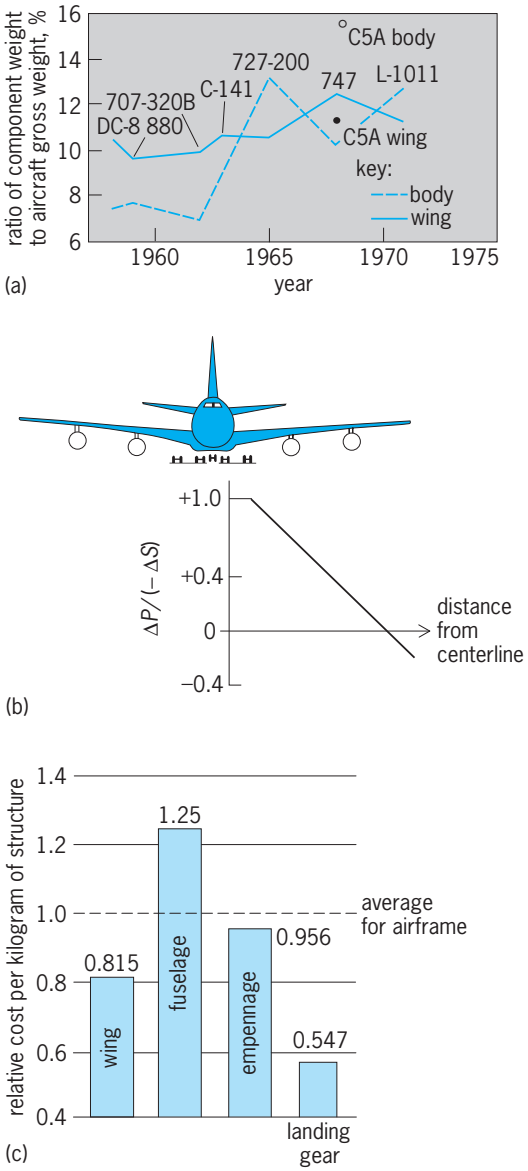

Fig. 8. Fuselage weight reduction. (*a*) Wing and fuselage weight history. (*b*) Plot of ratio of increase in payload ΔP to corresponding decrease in structural weight, $-\Delta S$, as a function of distance from aircraft centerline, showing centerline payload-weight effect. (*c*) Comparison of relative costs of aircraft components per kilogram of structure. (*After J. E. McCarty, Application of advanced structures and materials to transport aircraft fuselages, AIAA Pap. 74-340, AIAA/ASME/SAE 15th Structures, Structural Dynamics and Materials Conference, Las Vegas,1974*)

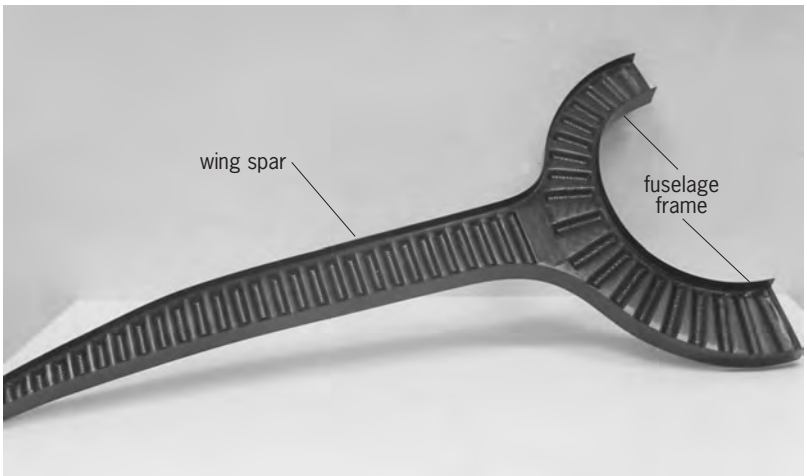

Fig. 9. Graphite/epoxy integral wing spar-fuselage frame for fighter aircraft. (*Boeing Research and Development*)

In addition to weight savings, Fig. 8*c* shows that the fabrication cost per pound of transport aircraft fuselage structure is 25% higher than the airframe average. Since the fuselage is approximately 37% of the airframe structural weight, high leverage in reducing airframe cost is available. New materials, configuration changes, and construction methods can reduce weight and cost of aircraft fuselages.

The material that dominates aircraft construction is aluminum; however, advanced composites are being introduced into empennage and wing structures of fighter aircraft, and advanced composites are to be applied to major aircraft fuselage structures. For circular fuselages, filament winding of composites may provide aircraft designers with the means to reduce weight and fabrication cost of large fuselages.

In addition to the application of advanced composite structures to aircraft fuselages, certain configuration adjustments are being made to high-performance fighter aircraft. Integral wing spar/frame designs allow more freedom of engine placement and better external aerodynamic contours, and provide a minimum weight configuration (**Fig. 9**). *See* WING STRUCTURE.

Fig. 10. Bonded honeycomb sandwich fuselage panel. (*After J. E. McCarty, Application of advanced structures and materials to transport aircraft fuselages, AIAA Pap. 74-340, AIAA/ASME/SAE 15th Structures, Structural Dynamics and Materials Conference, Las Vegas,1974*)

The application of adhesive-bonded aluminum fuselage structure in the configuration of bonded honeycomb sandwich construction can be utilized to improve the fatigue- and damage-tolerant characteristics of pressurized fuselages (**Fig. 10**). The adhesive-bonding assembly methods also reduce the construction cost of the fuselage, by reducing part count and mechanical fastening requirements. *See* AIRPLANE.

John E. McCarty

Bibliography. J. Cutler, *Understanding Aircraft Structures*, 3d ed., 1999; B. K. Donaldson, *Analysis of Aircraft Structures: An Introduction*, 1992; B. C. Hoskin and A. A. Baker (eds.), *Composite Materials for Aircraft Structures*, 1986; T. H. Megson, *Aircraft Structures for Engineering Students*, 3d ed., 1999.

Fusulinacea

An extinct group of marine organisms making up the superfamily Fusulinacea of the phylum Protozoa, order Foraminiferida. Fusulinaceans first appeared late in the Mississippian Period, and in the succeeding 100,000,000 years evolved into more than 125 genera and 6000 species before becoming extinct near the close of the Permian Period. Among protozoa fusulinaceans became giants, many reaching 0.4 in. (1 cm) in length and some even attaining 4 in. (10 cm) in length. The group is characterized by rapid evolution of distinctive morphological features, which have enabled paleontologists to establish a detailed fusulinacean phylogeny and biostratigraphic zonation. These fossils are widespread in Upper Mississippian, Pennsylvanian, and Permian marine strata in the major Eurasian and American geosynclinal belts and adjacent marine shelves, and have been reported from all continents except Australia and Antarctica, as indicated in **Fig. 1**.

Morphology. The structure of the families of the Fusulinacea is shown in **Fig. 2**. The calcareous shell of fusulinaceans is constructed on a simple plan. The initial chamber, the proloculus, is a small calcareous sphere with an opening on one side. As the individual grew and the volume of protoplasm increased, chambers were added successively in a planispiral coil around the proloculus. In most genera these chambers overlap the ends of the previous chambers and become elongate, parallel to the axis of coiling. The spiral wall (spirotheca), as in Fusulinidae, is composed of several layers. The initial wall is formed of a thin, dark organic-rich layer (tectum), beneath which is a thicker translucent layer (diaphanotheca) formed of calcium carbonate.

In several families of the Fusulinacea a secondary veneer of calcareous material (tectorial deposits) coats the interior of chambers and thus covers the floor, sides, or roof, and sometimes all three, of the chambers.

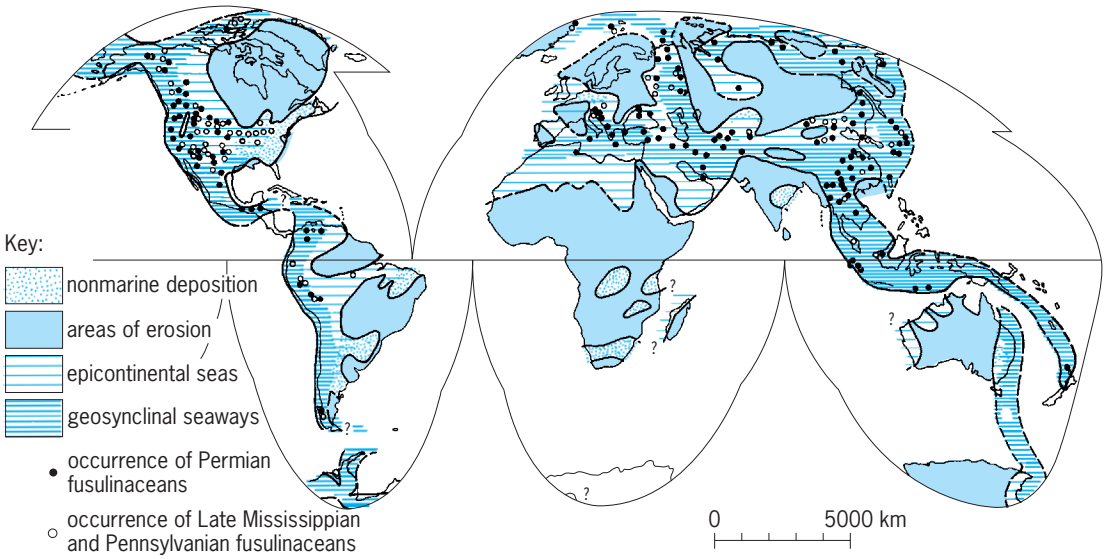

Fig. 1. Occurrences of Fusulinacea and distribution of sediments in the Pennsylvanian and Permian periods. 1 km = 0.6 mi.

Very small pores (mural pores), and 2 micrometers in diameter, penetrate the tectum and extend through the spiral wall. In highly evolved families, such as the Verbeekinidae and Schwagerinidae, a reduction in the number of layers in the spiral wall is accompanied by a joining together in the translucent layer of several mural pores to form larger tubes (alveoli), which gives the layer a honeycombed construction and modifies it to one distinctive layer, the keriotheca.

The frontal wall of each chamber forms a septum which divides the interior of the shell. In primitive genera septa are nearly planar, and only near the axial poles are they gently folded. In many of the more advanced genera the septa commonly become strongly folded across the entire length of the chamber. Larger pores (septal pores), about 10 μm in diameter, are distributed across the face of the septa.

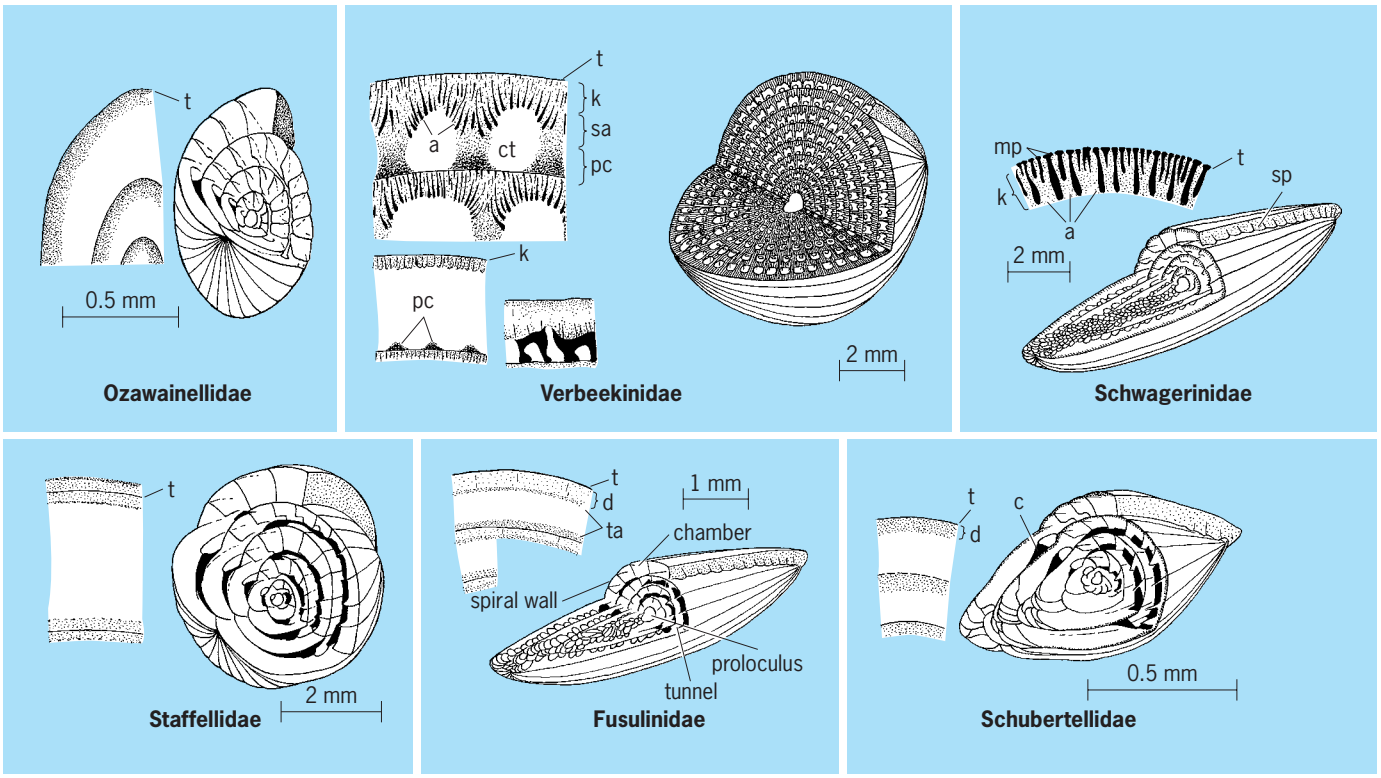

Fig. 2. The six families of Fusulinacea. Representatives shown in cutaway diagrams, with wall structures of each greatly enlarged; a = alveoli, c = chomata, ct = chamberlets, d = diaphanotheca, k = keriotheca, mp = mural pores, pc = parachomata, sa = septula, sp = septal pores, t = tectum, and ta = tectorial deposits.

Within the shell of fusulinaceans, a spiral tunnel is formed by resorption of part of the septa near the midplane. In many genera secondary deposits form ridges (chomata) along the floor of the chambers adjacent to the tunnel, as in Schubertellidae. This tunnel probably served as a passageway, along which the nucleus of the cell moved from the center of the shell toward the outer chambers. Most genera have single tunnels. A few have multiple or auxiliary tunnels and thus fusulinaceans may have developed multiple nuclei. In some advanced genera additional passageways between chambers (cuniculi) were formed by resorption of septa where opposing folds of adjacent septa meet. In the Verbeekinidae, secondary transverse and longitudinal partitions (septula) project downward from the spiral wall, and transverse

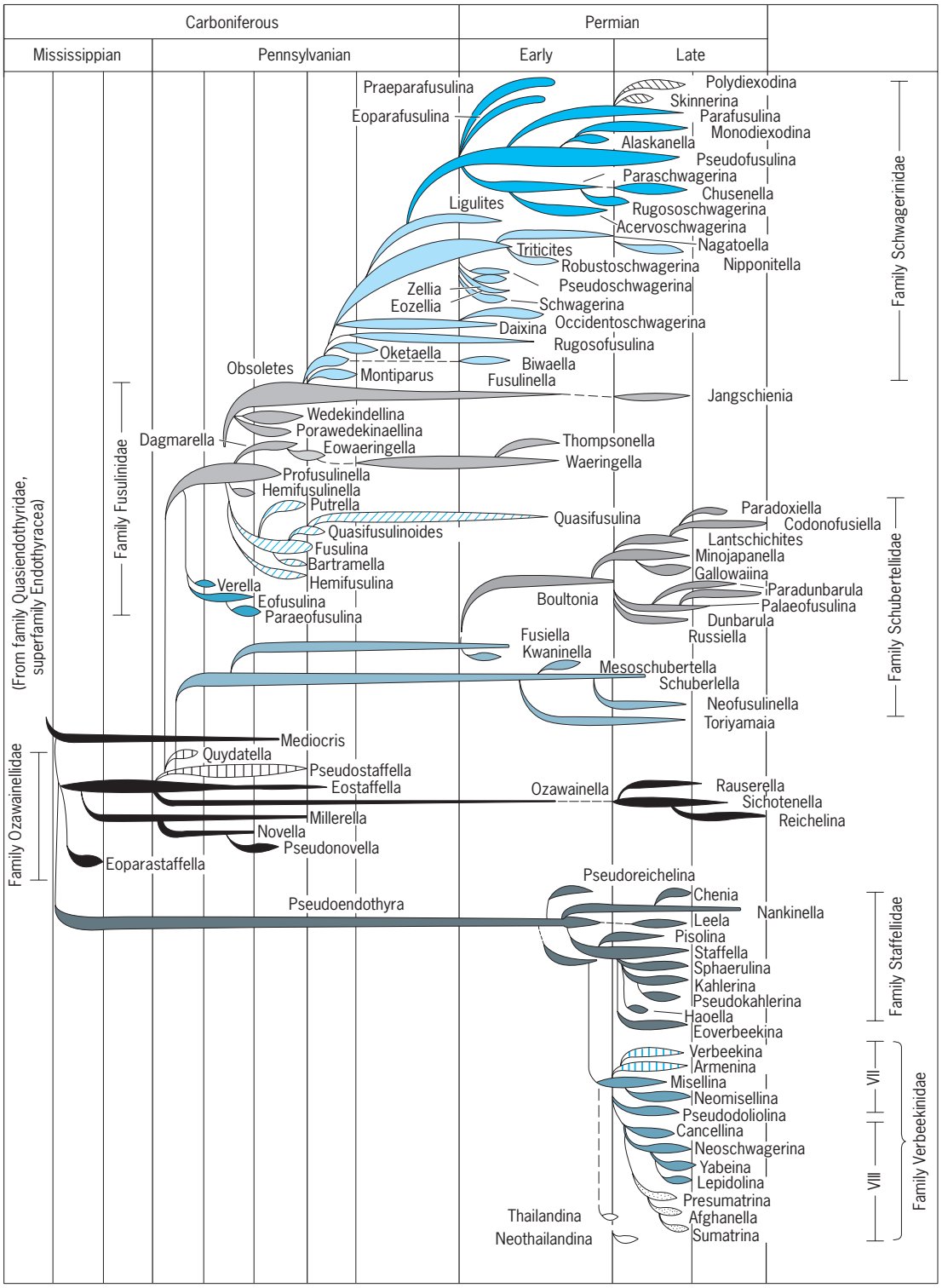

Fig. 3. Evolutionary relationships of major genera and families of fusulinaceans.

deposits (parachomata) form ridges on the floor of chambers to subdivide the chambers into chamberlets.

Evolution. Six evolutionary lineages of fusulinaceans (**Fig. 3**) corresponding to the six families, are differentiated on the basis of general shell structures and structural changes of the shell. The families Ozawainellidae, Schubertellidae, and Staffellidae persisted as small, simple, and apparently unspecialized lineages for most of the Pennsylvanian and Early Permian. During the Middle and Late Permian, new genera evolved in these three families and many of these genera were adapted to specialized ecological conditions, probably shallow, warm marine water having slightly above-normal salinity, conditions which appear to have persisted locally until near the end of the Permian Period.

The family Fusulinidae had great evolutionary diversification during the Early and Middle Pennsylvanian, and then within a remarkably short time came close to extinction before the beginning of the Late Pennsylvanian. Only a few genera of Fusulinidae survived, and this much-reduced lineage is traceable into the Late Permian.

The family Schwagerinidae arose from the Fusulinidae at about the end of the Middle Pennsylvanian, evolved slowly during the Late Pennsylvanian, and produced a burst of nearly 20 new genera during the Early Permian. The Schwagerinidae gradually, and then more rapidly, dwindled during the later part of the Early Permian, and finally became extinct near the middle of the Late Permian.

The sixth family, the Verbeekinidae, evolved from the Staffellidae in the Early Permian, underwent a rapid and distinctive evolutionary development, and with almost the same rapidity became extinct near the middle of the Late Permian. The Verbeekinidae and several genera in the other families had geographic distributions that were largely restricted to a Permian seaway that extended from northern Africa eastward through south-central Asia and southeastern Asia to New Zealand, and through China, Japan, and eastern Siberia into western North America as far south as California, but with rare exceptions not into other parts of Eurasia and North America and not into South America. *See* FORAMINIFERIDA. Charles A. Ross

Bibliography. K. Kanmera, K. Ishii, and R. Toriyama, *Geol. Paleontol. Southeast Asia*, 17:129–154, 1976; A. R. Loeblich and H. Tappan, *Treatise on Invertebrate Paleontology*, 2 vols., 1964; R. C. Moore, C. G. Lalicker, and A. G. Fischer, *Invertebrate Fossils*, 1952; C. A. Ross, *J. Paleontol.*, 41:1341–1354, 1967; S. E. Rozovskaya, *Trudy Paleontol. Inst. Akad. Nauk U.S.S.R.* (in Russian), vol. 149, 1975; R. R. Shrock and W. H. Twenhofel, *Principles of Invertebrate Paleontology*, 2d ed., 1953.

Fuzzy sets and systems

A fuzzy set is a generalized set to which objects can belong with various degrees (grades) of memberships over the interval [0,1]. In general, fuzziness describes objects or processes that are not amenable to precise definition or precise measurement. Thus, fuzzy processes can be defined as processes that are vaguely defined and have some cognitive uncertainty in their description. The data arising from fuzzy systems are, in general, soft, with no precise boundaries. Examples of such systems are large-scale engineering complex systems, social systems, economic systems, management systems, medical diagnostic processes, and human perception. *See* SET THEORY.

Function. The mathematics of fuzzy set theory deals with the uncertainty and fuzziness arising from interrelated humanistic cognitive types of phenomena such as subjectivity, thinking, reasoning, cognition, and perception. This type of uncertainty is characterized by structures that lack sharp (well-defined) boundaries. This approach provides a way to translate a linguistic model of the human thinking process into a mathematical framework for developing computer algorithms for computerized decision-making processes. The theory has grown very rapidly. Many fuzzy algorithms have been developed for application to process control, medical diagnosis, management sciences, engineering design, and many other decision-making processes where soft data are generated. Thus, fuzzy mathematics provides a modeling link between the human reasoning process, which is vague, and computers, which accept only precise data.

For example, in the design of many engineering systems, process information is not available both because it is difficult to understand precisely the complexity of the phenomena and because human reasoning is inexact and is based upon subjective perception. However, by virtue of knowledge and experience, which are inexact, it is possible to build increasingly good systems. In fact, fuzziness in thinking and reasoning processes is an asset since it makes it possible to convey a large amount of information with a very few words. However, in order to emulate this experience and these reasoning processes on a computer, for example, for intelligent robotics applications and medical diagnosis, a mathematical precision must be given to the vagueness of the information so that a computer can accept it. This is done by using the theory of fuzzy sets. Probability theory deals with the uncertainty or randomness that arises in mechanistic systems, whereas fuzzy set theory has been created to deal with the uncertainty that arises in human cognitive processes. *See* PROBABILITY.

The premise of fuzzy set theory is that the key elements in human reasoning processes are not numbers but labels of fuzzy sets. The degree of membership is specified by a number between 1 (full membership) and 0 (full nonmembership). An ordinary set is a special case of a fuzzy set, where the degree of membership is either 0 or 1. By virtue of fuzzy sets, human concepts like small, big, rich, old, very old, and beautiful can be translated into a form usable by computers.

Elements of fuzzy set theory. A fuzzy set A in the universe X is a set of ordered pairs $A = \{x,\mu_A(x)\}$,

where x is a member of X ($x \in X$) and $\mu_A(x)$ is the grade of membership of x in A. Thus, μ_A is a function of X over the interval [0,1], as stated symbolically in Eq. (1). A fuzzy set is said to be normal if and only

$$\mu_A : X \rightarrow [0, 1] \tag{1}$$

if the maximum value of $\mu_A(x)$ [max $\mu_A(x)$] equals 1 (**Fig. 1**).

The calculus of the fuzzy sets is based on the following operations:

1. Inclusion. A fuzzy set A is said to be included in a fuzzy set B if Eq. (2) is valid.

$$\mu_A(x) \leq \mu_B(x) \qquad \text{for all } x \in X \tag{2}$$

2. Intersection. The intersection of the fuzzy sets A and B is defined by Eq. (3). The intersection corresponds to the logical connective AND. Thus, Eq. (4) holds.

$$\begin{aligned}\mu_{A\cap B}(x) &= \min[\mu_A(x), \mu_B(x)] \\ &= \mu_A(x) \wedge \mu_B(x)\end{aligned} \tag{3}$$

$$A \text{ AND } B = A \cap B \tag{4}$$

3. Union. The union of fuzzy sets A and B is defined by Eq. (5). The union corresponds to the connective OR. Thus, Eq. (6) follows.

$$\begin{aligned}\mu_{A\cup B}(x) &= \max[\mu_A(x), \mu_B(x)] \\ &= \mu_A(x) \vee \mu_B(x)\end{aligned} \tag{5}$$

$$A \text{ OR } B = A \cup B \tag{6}$$

4. Complement. The complement $\bar{A}$ of a fuzzy set A is defined by Eq. (7). The operation of complementation corresponds to the negation NOT. Thus, Eq. (8) holds. An example of this operation is shown in Fig. 1*a*.

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \tag{7}$$

$$\bar{A} = \text{NOT } A \tag{8}$$

5. Product. The product of two fuzzy sets is defined by Eq. (9).

$$\mu_{A\cdot B}(x) = \mu_A(x) \cdot \mu_B(x) \tag{9}$$

6. Algebraic sum. The algebraic sum is defined by Eq. (10).

$$\mu_{A+B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x) \cdot \mu_B(x) \tag{10}$$

In contrast to ordinary set theory, the union of a set and its complement is not necessarily the universe X; and the intersection of a set and its complement is not necessarily the null set ϕ. These two important properties are expressed in Eqs. (11) and (12).

$$A \cup \bar{A} \neq X \text{ (not necessarily)} \tag{11}$$

$$A \cap \bar{A} \neq \phi \text{ (not necessarily)} \tag{12}$$

An example that illustrates these operations is provided by two fuzzy sets A and B, defined over the universe X, where X is given by Eq. (13), and the sets

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6\} \tag{13}$$

A and B are defined by Eqs. (14) and (15), where numerical values over the interval [0,1] indicate the

$$A = \{(0.6/x_1), (0.4/x_2), (0.3/x_3), (0.9/x_4), (1.0/x_5), (0.4/x_6)\} \tag{14}$$

$$B = \{(0.3/x_1), (0.8/x_2), (0.4/x_3), (0.7/x_4), (0.6/x_5), (1.0/x_6)\} \tag{15}$$

grade of the membership μ. The OR, AND, and COMPLEMENT operations of the fuzzy sets A and B are

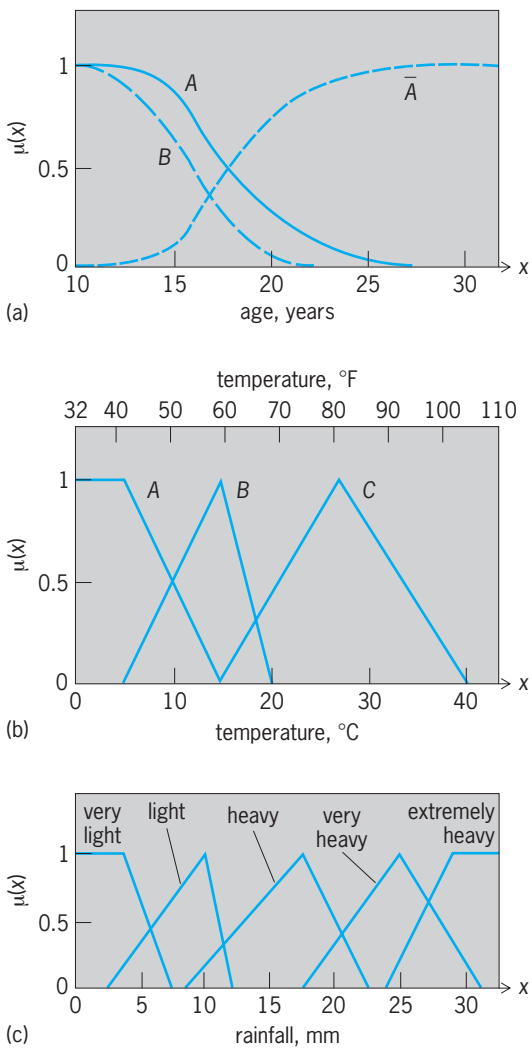

Fig. 1. Examples of fuzzy sets. (*a*) Fuzzy sets of age groups of high school children: young (*A*); not young ($\bar{A}$); very young (*B*). (*b*) Typical triangular representations of fuzzy labels of temperature: less than 6°C [43°F] (*A*); about 15°C [59°F] (*B*); and about 27°C [81°F] (*C*). These fuzzy labels may also be interpreted as: *A* = very cold; *B* = Cold; *C* = warm. (*c*) Graphs of membership used for classifying the rainfall in a region. 1 mm = 0.04 in.

given by Eqs. (16), (17), and (18) respectively.

$$A \cup B = \{(0.6/x_1), (0.8/x_2), (0.4/x_3), (0.9/x_4), (1/x_5), (1/x_6)\} \quad (16)$$

$$A \cap B = \{(0.3/x_1), (0.4/x_2), (0.3/x_3), (0.7/x_4), (0.6/x_5), (0.4/x_6)\} \quad (17)$$

$$\bar{A} = \{(0.4x_1), (0.6/x_2), (0.7/x_3), (0.1/x_4), (0/x_5), (0.6/x_6)\} \quad (18)$$

Other fuzzy set operations. The concept of partial membership makes possible other operations on fuzzy sets which are not found in the ordinary set theory. These operations are very important in modeling complex systems, for emulating the humanlike reasoning process in computers and in computer vision, and in modeling the human cognitive and perception process. Using abbreviated notation, some further important operations will be defined. *See* COGNITION; COMPUTER VISION; PERCEPTION.

The bounded product, bounded sum, and bounded difference of fuzzy sets A and B are defined by Eqs. (19), (20), and (21), respectively.

Bounded product:

$$\mu_{A \odot B} = O \vee (\mu_A + \mu_B - 1) \quad (19)$$

Bounded sum:

$$\mu_{A \oplus B} = 1 \wedge (\mu_A + \mu_B) \quad (20)$$

Bounded difference:

$$\mu_{A \ominus B} = O \vee (\mu_A - \mu_B) \quad (21)$$

The concentration operation provides a relatively sharp boundary to a fuzzy set, and is defined by Eq. (22).

$$\mu_{\mathrm{CON}(A)} = \mu_A^{\alpha} \qquad \alpha > 1 \quad (22)$$

The dilation operation provides a relatively flexible boundary to a fuzzy set, and is defined by Eq. (23).

$$\mu_{\mathrm{DIL}(A)} = \mu_A^{\beta} \qquad 0 < \beta < 1 \quad (23)$$

The contrast intensification operation increases membership if it is greater than 0.5, and decreases if it is less than 0.5. It is defined by Eq. (24).

$$\mu_{\mathrm{INT}(A)} = \begin{cases} \mu_{\mathrm{CON}(A)} & \text{for } \mu(x) < 0.5 \\ \mu_{\mathrm{DIL}(A)} & \text{for } \mu(x) \geq 0.5 \end{cases} \quad (24)$$

The blurring operation is the inverse of intensification and is defined by Eq. (25). Thus, the blurring

$$\mu_{\mathrm{BLR}(A)} = \begin{cases} \mu_{\mathrm{DIL}(A)} & \text{for } \mu(x) < 0.5 \\ \mu_{\mathrm{CON}(A)} & \text{for } \mu(x) \geq 0.5 \end{cases} \quad (25)$$

operation, unlike the intensification operation, decreases membership if it is greater than 0.5, and increases it if less than 0.5.

For the fuzzy set A of the above example, the membership vector is given by Eq. (26). Thus, for $\alpha = 2$

$$\mu_A = \{0.6, 0.4, 0.3, 0.9, 1.0, 0.4\} \quad (26)$$

and $\beta = {}^1/_2$, the concentration, dilation, intensification, and blurring operations of the fuzzy set A are given by Eqs. (27)–(30).

$$\mu_{\mathrm{CON}(A)} = \mu_A^2 = \{0.36, 0.16, 0.09, 0.81, 1.0, 0.16\} \quad (27)$$

$$\mu_{\mathrm{DIL}(A)} = \mu_A^{1/2} = \{0.77, 0.63, 0.55, 0.95, 1.0, 0.63\} \quad (28)$$

$$\mu_{\mathrm{INT}(A)} = \{0.77, 0.16, 0.09, 0.95, 1.0, 0.16\} \quad (29)$$

$$\mu_{\mathrm{BLR}(A)} = \{0.36, 0.63, 0.55, 0.81, 1.0, 0.63\} \quad (30)$$

Systems modeling and control. Fuzzy set theory is a useful tool for modeling complex systems and for generating decision (control) rules. For example, a complex dynamic process with $x(t)$ as an input and $y(t)$ as an output (**Fig. 2**) can be modeled by using certain input and output observations. For example, a typical linguistic model would be given by Eq. (31),

$$\sum_i \quad \text{IF } X_i \text{ THEN } Y_i \quad (31)$$

which represents a collection of IF . . . THEN rules on labels of fuzzy input variables X_i and fuzzy output variables Y_i.

Control (decision) rules also can be generated for a feedback control process (**Fig. 3**) by using IF . . . THEN rules. The fuzzifier (F) converts the physical output $y(t)$ into fuzzy variables Y_f. The fuzzy logic controller then generates the fuzzy control action U_f using IF . . . THEN rules given by Eq. (32). The

$$\text{IF } Y_f \text{ THEN } U_f \quad (32)$$

defuzzifier (DF) converts the fuzzy control action U_f into a set of crisp values $u(t)$ which in turn controls the dynamic system.

Based upon such fuzzy control rules, several industrial controllers are in use for process control and similar operations. Also, very large scale integrated circuit (VLSI) hardware is being developed to implement these fuzzy logics. *See* CONTROL SYSTEMS; DECISION THEORY; INTEGRATED CIRCUITS.

Process control applications. A process operator, having an assigned control goal, observes the process state, controller, and process output, and intuitively assesses the variables and parameters of the process. From a subjective assessment of these quantities, the

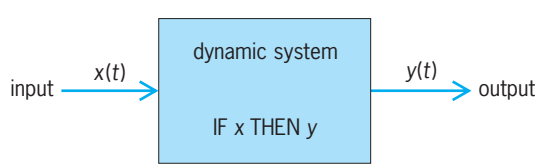

Fig. 2. Dynamic system with IF . . . THEN rules.

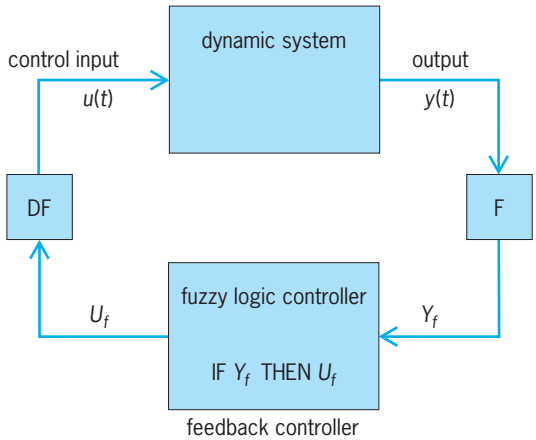

Fig. 3. Fuzzy logic controller using IF . . . THEN control rules. F = fuzzifier; DF = defuzzifier.

operator makes a decision and performs a manual alteration of the system control value so as to achieve an assigned control goal. It is possible to make a description of control strategy, a description of process behavior, and finally a subjective operational model of the operator's procedure. Such a control algorithm is very flexible and adequate to the actual process situation, and, as a rule, its performance is much better than that of algorithms obtained by using modern control theory. The linguistic algorithm comprises all the so-called metaphysical skills of the operator, such as intuition, experience, and intelligence, which cannot be dealt with by conventional mathematics. The operator's mind and knowledge of the controlled technological process can be formalized mathematically, to some extent, by using the theory of fuzzy sets and logic. *See* ALGORITHM; ARTIFICIAL INTELLIGENCE; EXPERT SYSTEMS.

For example, in observing an error E and error change ΔE, a process operator might carry out manual control of the process according to the following verbal description:

If E = big and ΔE = small, then U = zero

also

If E = medium and ΔE = zero, then U = big.

also

⋮

In order to calculate control U, when given fuzzy relation R, error E, and error change ΔE, the compositional rule of inference given by Eq. (33) would be used, where $\circ$ is the max-min composition.

$$U = E \circ \Delta E \circ R \tag{33}$$

There are four ways to obtain control rules: (1) they may be taken directly from the operator's experience; (2) they may be derived from a fuzzy model of the process, where the process characteristics are expressed in a form similar to control rules; (3) the operator's actions may be monitored and a control derived by using a fuzzy model of the operator's behavior; or (4) the rule may be learned by the controller.

Applications of fuzzy control theory include the fuzzy control of steam engines, heat exchangers, sinter plants, vehicular traffic, pressurized tanks, motors, ship autopilots, automobile speed, water cleaning processes, room temperature, pump operations, robots, prosthetic devices, warm-water plants, pressurized tank liquids, batch chemical reactors, nuclear fuel extraction processes, cement kilns, activated-sludge wastewater treatment, maintenance scheduling in transportation systems, stirred tanks, the process of changes of methane concentration, canvas production, fermentation processes of antibiotic synthesis, hydraulic servo systems, pulp plants, electric networks, converter steel-making processes, water-tank systems, robot arc welders, automatic train operation using predictive logic, automobile operation using self-learning, aircraft flight, and electrical trains, as well as fuzzy algorithms for path selection in autonomous vehicle navigation, and fuzzy adaptive control of continuous casting plants.

Fuzzy control of dc motor. An example of fuzzy control is provided by a direct-current (dc) motor (**Fig. 4**). The objective is to maintain constant motor speed despite the effect of a disturbing moment $M(t)$. The system may be described by the fuzzy variables S (error sum), I (motor current), and U (control action). The values of U corresponding to various values of S and I may be specified in a table

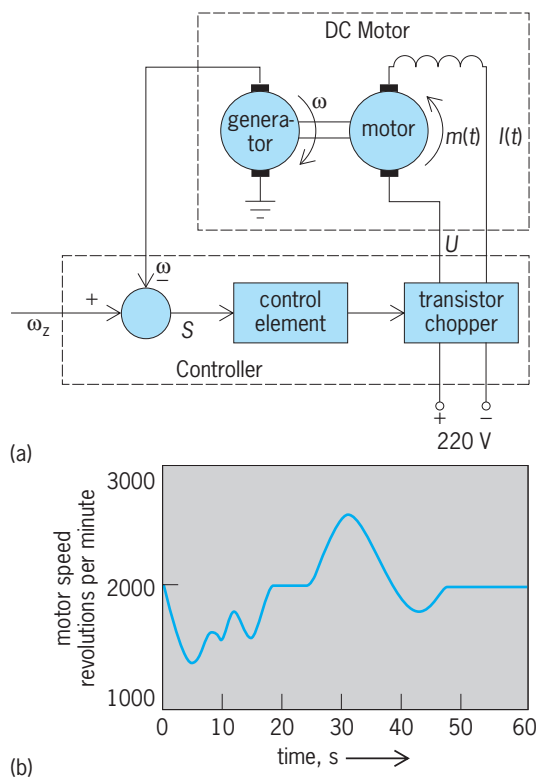

Fig. 4. Fuzzy control of a direct-current (dc) motor. (*a*) Elements of motor and controller. ω = motor angular speed, ω_Z = reference angular speed, *M(t)* = disturbing moment, *S* = error sum, *I(t)* = motor current, *U* = control action. (*b*) Motor speed with fuzzy control.

Relation of fuzzy variables in control of a dc motor showing values of control action (*U*)

	Motor current (*I*)			
Error sum (*S*)	Zero	Small	Medium	Big
Super-big			Zero	
Big			Zero	Small
Medium		Null		Small
Small	Null	Null	Zero	Zero
Zero		Zero		Medium
Null		Zero		Big
Negative-null	Null		Zero	
Negative-zero	Null	Zero	Big	
Negative-small	Zero	Big		
Negative-medium	Null	Big	Big	
Negative-big	Zero	Big		
Negative-super-big	Small	Medium		

(see **table**) or in a list of relations, such as

$$\text{If } S = \text{null and } I = \text{small, then } U = \text{zero.}$$

These values satisfy the fuzzy relation determined by Eq. (34). The results of the experiment

$$\mu_R(s, i, u) = \vee\{\mu_S(s) \wedge \mu_I(i) \wedge \mu_U(u)\} \qquad (34)$$

(Fig. 4*b*) are very satisfactory. *See* ADAPTIVE CONTROL; INTELLIGENT MACHINE; PROCESS CONTROL; ROBOTICS.

Madan M. Gupta

Bibliography. B. M. Ayyub and M. M. Gupta (eds.), *Uncertainty Analysis in Engineering and Sciences: Fuzzy Logic, Statistics, and Neural Networks Approach*, 1997; M. M. Gupta et al., *Approximate Reasoning in Expert Systems*, 1985; M. M. Gupta and N. K. Sinha, *Intelligent Control Systems*, 1996; M. M. Gupta and T. Yamakawa, *Fuzzy Computing: Theory, Hardware and Applications*, 1988; M. M. Gupta and T. Yamakawa (eds.), *Fuzzy Logic in Knowledge-Based Systems, Decision and Control*, 1988; A. Kandel and G. Langholz (eds.), *Fuzzy Control Systems*, 1993; A. Kaufmann and M. M. Gupta, *Fuzzy Mathematical Models in Engineering and Management Science*, 1988; A. Kaufmann and M. M. Gupta, *Introduction to Fuzzy Arithmetic: Theory and Applications*, 1985; R. Kruse, J. Gebhardt, and F. Klawonn, *Foundations of Fuzzy Systems*, 1994; W. Pedrycz, *Fuzzy Control and Fuzzy Systems*, 1993.

Fuzzy-structure acoustics

Large structures such as ships and airplanes can undergo a variety of complicated vibrations. Such structures typically consist of an outer body made of metal plating (for example, the hull of a ship) or perhaps a massive metallic frame (for example, the chassis of a truck), and a large variety of internal objects that are connected to either the plating or the frame. In designing such structures, it is highly desirable to have some method for predicting how they will vibrate under various conditions. The radiation of sound caused by these vibrations, either into the environment or into the empty portions of the structure, is also of interest because this sound is often either unwanted noise or a means of inferring information about the details of the structure or the excitation. Fuzzy-structure acoustics refers to a class of conceptual viewpoints in which precise computationally intensive models of the overall structure are replaced by nonprecise analytical models, for which the initial information is said to be fuzzy. *See* FUZZY SETS AND SYSTEMS.

External and internal structures. Fuzzy-structure theories divide the overall structure (**Fig. 1**) into a master structure and one or more attached structures, the latter being referred to as the fuzzy substructures, the internal structures, or the internals. (An example of a master structure is the hull and major framework of a ship.) The master structure is presumed to be sufficiently well known at the outset that its vibrations or dynamical response could be predicted if the forces that were exerted on it were known. Some of the forces are exerted on it by the substructures at the points at which they are attached. Such forces can be very complicated; nevertheless, there is some hope that a satisfactory approximate prediction of the vibrations of the master structure itself can be achieved with a highly simplified model.

The fuzzy substructures can be regarded as structures that are not known precisely. The most information-demanding type of fuzzy-structure model has significant drawbacks. Perhaps the most serious disadvantage is that the exploration of such models requires Monte Carlo simulations that are often infeasible or prohibitively time consuming. Consequently, there is some impetus to consider fuzzier models of the fuzzy substructures.

Forces and impedances. The fuzzier models that have been developed in the past few years rest on derivable consequences of the science of mechanics. The substructures are passive, so the displacements of attachment points on a substructure can be related to the forces exerted on them. The same relations, because of Newton's third law, have to hold

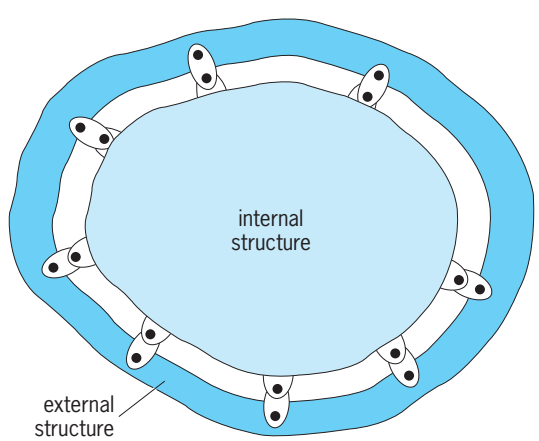

Fig. 1. Mechanical structure conceptually divided into an internal structure (fuzzy substructures) and an external structure (master structure).

between the forces exerted on the master structure and the corresponding displacements, except for a simple change in sign. The influences of the substructures on the master structure are analogous to those of springs connecting the master structure to rigid walls, but here the springs are more complicated than those described by Hooke's law (in which force is directly proportional to extension) in that displacements at distant attachment points may affect a given force, the forces are not necessarily in phase with the displacements, and the proportionality and the phase shift vary with the frequency with which the master structure is oscillating.

Nevertheless, there is still a surprisingly simple form for the general relations between the set of forces and the set of attachment-point displacements. Typically, workers in acoustics and vibrations express such relations as a rectangular array, any one element providing a ratio of a force experienced at one attachment point to the velocity of another attachment point, given that the displacements of all the other attachment points are held to zero, that the oscillations are sinusoidal (fixed frequency), and that complex numbers are used for the forces and velocities (so that differences in phases can be accounted for). Ratios of complex force amplitudes to complex velocity amplitudes are termed impedances, and an array of ratios of different forces to different velocities is termed an impedance matrix. *See* ACOUSTIC IMPEDANCE; MATRIX THEORY; MECHANICAL IMPEDANCE.

Modes and natural frequencies. The branch of mechanics known as linear vibration theory leads to the prediction that each of the impedance matrix elements can be written as a sum of terms, each term being similar to what would be expected from shaking a simple oscillator (**Fig. 2**), consisting of a mass, a spring, and a dashpot at an attachment point. Such a term, representing the ratio of a force to a velocity, has a frequency-dependent denominator that would go to zero in the absence of damping at an angular frequency (referred to as the resonance frequency or the natural frequency) equal to the square root of the spring constant divided by the mass. For very weak damping, the shaking force and the displacement are in the same direction when the shaking frequency is less than the natural frequency, but oppositely directed when the shaking frequency is higher than the natural frequency. *See* FORCED OSCILLATION; RESONANCE (ACOUSTICS AND MECHANICS).

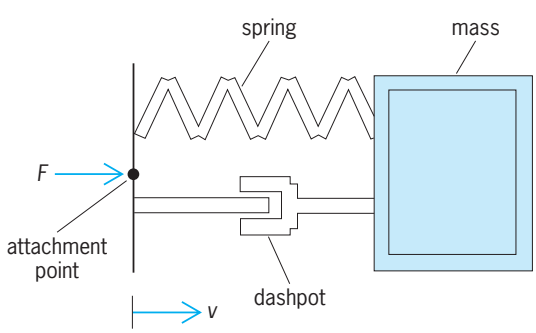

Fig. 2. Simple oscillator driven at its attachment point. *F* is the driving force, and *v* is the velocity at the attachment point. The ratio of the complex amplitude of *F* to the complex amplitude of *v* gives the driving-point impedance.

Typical substructures, when regarded in terms of the impedance matrix that they present to the external environment, behave as a set of independent simple oscillators, each having its own natural frequency. Corresponding terms in the series for the individual impedance matrix elements have the same denominator; consequently they go to zero, in the limit of zero damping, at the same natural frequency. These natural frequencies represent the frequencies at which the substructure would freely vibrate indefinitely in the absence of internal damping, when all of the attachment points are held motionless but the rest of the substructure is free. The overall manner in which the substructure is oscillating when undergoing vibrations under such circumstances is called a natural mode of vibration.

Modal masses. The numerators of the corresponding modal terms in the separate elements of an impedance matrix lead to the identification of a modal mass matrix for each natural mode. Each element of the modal mass matrix has the units of mass. For the special case where there is only one attachment point and the points in the substructure all move back and forth in the same direction, the matrix for each mode is just a one-by-one matrix, so there is a single mass associated with each mode. If there are a number of different attachment points but the points in the substructure still are constrained to move in the same direction, then the elements of the modal mass matrix for a given mode sum to a quantity identified as the modal mass for that mode. Finally, if substructures are considered where the points can move in all three directions and the definitions of the forces and velocity are generalized so that each cartesian component of each force is regarded as a separate force, then the sum of all the modal mass matrix elements for any given mode has a value that is three times as large as what would be regarded as the modal mass associated with that mode. In each such case, the modal masses when summed over all natural modes yield the total mass of the substructure.

Modal mass per unit frequency. Recently developed theories of fuzzy structures lead, after various plausible idealizations, to a formulation that requires only a single function, this being the modal mass per unit frequency bandwidth. Such a function could be regarded as a quasi-limit. A band of frequencies is chosen that is of narrow bandwidth but nevertheless contains a large number of natural frequencies. A sum is taken over all the modal masses for all of the modes that have frequencies in that band, this sum is then divided by the bandwidth, and the resulting value is associated with the center frequency of the band. This concept is similar to that of modal density, which is the number of resonance frequencies per unit frequency bandwidth—only here there is a weighting by modal mass.

This function can be estimated without the necessity of calculating each resonance frequency and

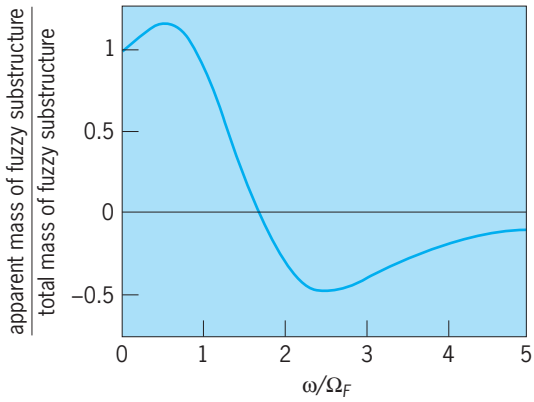

Fig. 3. Sample calculation of the frequency dependence of the apparent mass per unit area added to a master structure (rectangular plate) by a distribution of fuzzy structures when the master structure is oscillating at angular frequency ω. Ω_F is a characteristic frequency for the set of fuzzy substructures. ***(After A. D. Pierce, V. W. Sparrow, and D. A. Russell, Fundamental structural-acoustic idealizations for structures with fuzzy internals, J. Vibrat. Acous., 117:339–348, 1995)***

each modal mass separately, with very little detailed knowledge of the design, configuration, or material properties of the substructure itself. Examples considered so far suggest that, for any general category of substructure, a simple functional form can be identified for the dependence on frequency with the necessity to specify only a small number of parameters. A judicious sampling of blueprints or parts lists, in conjunction with a few simple field measurements, might be sufficient to estimate these parameters.

Predictions. In the formulation alluded to above, the influence of fuzzy substructures attached to the master structure tends to resemble that of an added frequency-dependent mass attached to the master structure in parallel with a frequency-dependent dashpot connecting the master structure to a hypothetical rigid wall. The added mass is a frequency-weighted integral over the modal mass per unit natural frequency, the weighting being such that the natural modes whose natural frequencies are less than the driving frequency have a positive contribution, while those for which the natural frequencies are greater than the driving frequency have a negative contribution. The master structure can seem to be less massive than it actually is (**Fig. 3**) when the bulk of the substructure mass is associated with resonant frequencies less than the excitation frequency.

The apparent dashpot constant, equal to the ratio of the apparent reacting force to the velocity at the attachment point, has a frequency-dependent value that is directly proportional to the mass per unit natural frequency bandwidth evaluated at the driving frequency, with an added multiplicative factor of the square of the driving frequency. Because the passive nature of the attached substructures precludes their acting as an energy source, the dashpot constant is always positive.

The prediction that the actual value of the dashpot constant is independent of the values of any parameters associated with energy dissipation (or damping) within the substructures is only approximate, but it does hold in the limit of sufficiently weak damping. There may be some relevant mechanism by which energy is absorbed within the substructure, but the actual amount of energy absorbed is asymptotically independent of the mechanism's strength. Similar phenomena, in which the amount of energy absorbed seems independent of the strength of the absorbing mechanism, occur in other branches of physics. For example, the energy loss at a shock front is dependent only on the pressure jump at the front in the limit when the shock front is thin and the absorption mechanism is weak. For the fuzzy-structure case, the evidence so far suggests that most internal structures of interest are sufficiently resilient that the asymptotic limit independent of damping mechanisms in the internals is of practical interest. This limit can be realized when (1) the number of resonances is very large and the resonance bandwidth of any given mode, although narrow compared with the resonance frequency itself, is sufficiently wide that there is considerable overlap between the resonance peaks of different modes; or (2) when the excitation of the master structure, while having a narrow bandwidth compared to the central frequency, is sufficiently wide that it includes a large number of natural frequencies of the substructure; or (3) when the force excitation is a short-duration pulse of nearly constant frequency, but the duration of the pulse is short compared to the reverberation time of the substructure, the latter being roughly equal to the reciprocal of the bandwidth of a resonant frequency of the substructure.

One implication of the newly emerging fuzzy-structure theories is that, insofar as there is concern with the vibrations of only the master structure, it is possible to drastically curtail the estimation or measurement of any parameters within the substructures that are associated with internal damping. *See* MECHANICAL VIBRATION; VIBRATION. Allan D. Pierce

Bibliography. A. D. Pierce, Resonant-frequency-distribution of internal mass inferred from mechanical impedance matrices, with application to fuzzy structure theory, *J. Vibrat. Acous.*, 119:324–333, 1997; A. D. Pierce, V. W. Sparrow, and D. A. Russell, Fundamental structural-acoustic idealizations for structures with fuzzy internals, *J. Vibrat. Acous.*, 117:339–348, 1995; C. Soize, A model and numerical method in the medium frequency range for vibroacoustic predictions using the theory of structural fuzzy, *J. Acous. Soc. Amer.*, 94:849–865, 1993; M. Strasberg and D. Feit, Vibration damping of large structures induced by attached small resonant structures, *J. Acous. Soc. Amer.*, 99:335–344, 1996.

G

Gabbro — Gentianales

Gabbro

The plutonic equivalent of its more abundant extrusive equivalent, basalt. Because it crystallized from a magma intruded deep within the crust, gabbro has a grain size visible to the naked eye with approximately equal amounts of calcic plagioclase (with 50% or more anorthite, the calcium aluminium feldspar) and pyroxene. Olivine is common as an early crystallized mineral, but either nepheline or quartz could be a late-stage crystallization product found in the matrix. Hornblende or biotite is commonly formed as an alteration product of pyroxene during the late stages of the magmatic crystallization, when water becomes enriched in the residual magma. *See* BIOTITE; HORNBLENDE.

Gabbro is found in diverse tectonic environments, ranging through oceanic ridges, convergent plate boundaries, stable continents, and rifts. The forms of the intrusive gabbro bodies include dikes, sills, pipes, laccoliths, stocks, batholiths, and large layered intrusive complexes.

Nomenclature. The detailed terminology of gabbro is complex, involving two triangular diagrams and many varietal names; however, it can be simplified by utilizing one triangular diagram with three end members: calcic plagioclase and clinopyroxene, calcic plagioclase and orthopyroxene, and calcic plagioclase and olivine (see **illus.**). *See* FELDSPAR; OLIVINE; ORTHORHOMBIC PYROXENE; PYROXENE.

If the rock contains any hornblende or nepheline or quartz, these minerals become the modifier for the rock name. For example, if quartz is found with labradorite and clinopyroxene, the rock is called a quartz-bearing gabbro or a quartz gabbro. It is rare, however, to find quartz in gabbro containing visible magnesium-rich olivine because a chemical reaction takes place between the two to form the equilibrium mineral pyroxene. Thus, the term quartz gabbro is not used for a gabbro containing magnesium-rich olivine. In the rare cases where quartz is found with magnesium-rich olivine, the quartz usually will be rimmed with pyroxene, indicating that it was picked up accidentally by the gabbro magma and that it is clearly not in equilibrium. However, nepheline is stable with olivine. Hence, nepheline gabbro or nepheline troctolite, though rare, is not unusual. Nepheline and orthopyroxene, however, are mutually exclusive for chemical reasons again; hence, nepheline gabbronorite is not an appropriate term. *See* LABRADORITE; NEPHELINE; QUARTZ.

If the proportion of any of the three main phases increases over the other, new names are utilized. When the pyroxene exceeds plagioclase by a ratio of more than 7:3, the rock is called a melanogabbro or a melanonorite. When plagioclase is less than 10% of the rock, the rock is no longer a gabbro, and a classification scheme for ultramafic rocks must be adopted. However, when plagioclase exceeds 90% of the rock, the felsic (plagioclase feldspar–rich) rock is called an anorthosite. In the following discussion,

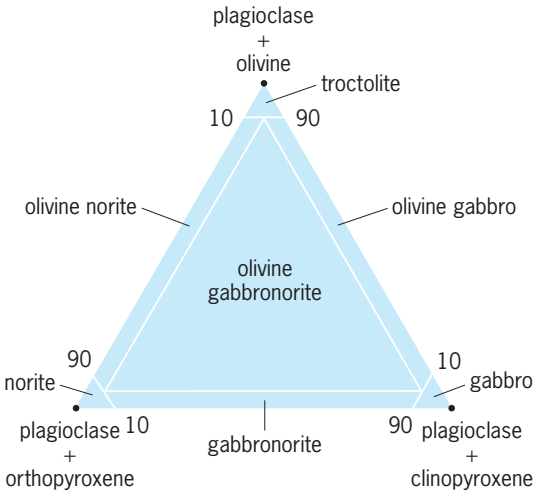

Triangular diagram showing end members and the nomenclature of gabbroic rocks.

gabbro will be used in a general sense to represent all seven names. *See* ANORTHOSITE.

Texture and structure. The grain size of gabbroic rocks ranges from a millimeter to centimeters. The finer-gained gabbro is commonly referred to as a diabase (or dolerite in the United Kingdom) that usually has small granular pyroxene interstitially enclosed by randomly oriented laths of calcic plagioclase. This diabasic texture results from the faster cooling of the magma due to its injection as small dikes and sills in shallow crust. In coarser-grained gabbro found in large plutons injected in deeper crust, the pyroxenes are larger and enclose partially (a texture called subophitic) or fully (a texture called ophitic) randomly oriented labradorite. Gabbros with mineral gains larger than 2 cm (0.8 in.) are rare, but such rocks are referred to as gabbro pegmatites. At contacts with country rock, the gabbro is commonly very fine grained or glassy because of the fast chilling of the magma. In some plutons, especially near their margins, gabbroic minerals are aligned perpendicular to the walls, yielding comb layering. *See* DOLERITE; MAGMA; PLUTON.

Chemical composition. Except for volatile compounds such as water, carbon dioxide, and sulfur molecules, which are largely lost during the eruption of basaltic magma, gabbros are chemically identical to basalts. Their silica (SiO_2) contents range 45–53% by weight, with 15–17% alumina (Al_2O_3). Gabbros have a high content of magnesium oxide (MgO), calcium oxide (CaO), and iron oxide (FeO) commonly ranging about 7–10 wt %, but the soda (Na_2O) and potash (K_2O) usually total less than about 3–4 wt % with the soda/potash ratio averaging less than 3. As with the basalts, some gabbros are alkali poor in relation to their silica content. As a consequence, saturated (with respect to silica) gabbros are specifically gabbronorites with two pyroxenes (both ortho and clinopyroxenes) and possibly quartz. These lack the undersaturated minerals such as nepheline. This variety is most common in the layered intrusions. The alkali-rich gabbros are undersaturated with respect to silica and, consequently, have only one pyroxene (clinopyroxene) and can have nepheline as an associated mineral.

Occurrence and origin. Unlike basalts, which are found in surface or near-surface environments, gabbros are found as shallow-to-deep intrusive bodies. Large blocks measuring more than 10 km (6 mi) thick, with a suite of rocks including serpentinite, pillow lavas, and chert, contain gabbro. This suite, called ophiolites, is commonly found on continents along convergent plate boundaries; it is thought to have been tectonically emplaced by thrusting onto the continental margins. It consists of a suite of rocks that are believed to represent the oceanic crust and upper mantle. This suite, exposed in places such as Newfoundland in Canada, Point Sal in California, Oman, and Cyprus (Troodos Massif), has a characteristic stratigraphy with the following sequence. From presumed top to bottom, there is an upper layer of oceanic sedimentary rocks (cherts) overlying basaltic pillow lavas that were fed from below by a diabasic sheeted dike complex. These rocks are above a layered gabbro sequence that overlies layered ultramafic rocks. Finally, at the bottom is an unlayered and metamorphosed ultramafic rock (serpentinite) representing the upper mantle. The sequence from the lower layered gabbro and ultramafic rocks to the sheeted dikes and the pillow lavas is believed to have resulted from fractional crystallization of a basaltic magma formed by partial melting of mantle rock. *See* CHERT; LAVA; OPHIOLITE; SERPENTINITE; STRATIGRAPHY.

Gabbros occur in minor abundance in the large granite batholiths found at convergent plate boundaries along the western Cordillera of the United States, including the Southern California, Sierra Nevada, Idaho/Boulder, and Coast Range (British Columbia) batholiths. The most extensive occurrence of gabbro (14% of total exposed area) is found in the Southern California batholith. Such occurrences are thought to be formed by fractional crystallization of the dominant granodioritic parent magma.

In Precambrian terranes, large batholiths (called massifs) of anorthosites (an igneous rock with greater than 90% plagioclase) commonly have gabbronorite at their margins. Such an occurrence is most common in the Grenville tectonic province (1000 to 900 million-year-old rocks) of eastern Canada and in Norway. The anorthosites and the gabbronorites are believed to have been formed by igneous processes and emplaced as a crystal-liquid mush into shallower regions of the crust.

Large layered intrusions such as the Stillwater Complex in Montana, the Skaergaard Intrusion in Greenland, the Muskox in Northwest Territories of Canada, and the one in the Bushveld Complex in South Africa have a layered gabbro/norite sequence similar to that found in the ophiolite suites. However, the homogeneous lower layer of serpentinite, the sheeted dike complex, the pillow lavas, and the chert are missing, and are replaced by a zone of silica-rich granophyre (quartz enclosed in sodic and potash feldspar, plus iron-rich olivine and pyroxene rock). As in ophiolites, an ultramafic layered sequence underlies the layered gabbro/norite sequence. The gabbronorite layers are continuous for hundreds of meters, yet are only a few centimeters to meters thick and alternate commonly with anorthosite layers. Within this layer the compositions of the pyroxenes and plagioclases become respectively iron-rich and sodium-rich from the bottom to the top. This layering is thought to originate in a manner similar to that in an ophiolite sequence by a fractionation process involving the settling of the denser grains from the liquid magma. *See* BASALT; IGNEOUS ROCKS.

Albert M. Kudo

Bibliography. A. Hall, *Igneous Petrology*, 2d ed., 1996; D. W. Hyndman, *Petrology of Igneous and Metamorphic Rocks*, 2d ed., 1985; A. R. McBirney, *Igneous Petrology*, 2d ed., 1993; A. R. Philpotts, *Principles of Igneous and Metamorphic Petrology*, 1990.

Gadiformes

An order of fishes containing the cods, codlings, codlets, grenadiers, and hakes. The Gadiformes plus the orders Ophidiiformes (pearfishes, cusk-eels, and brotulas), Batrachoidiformes (toadfishes), and Lophiiformes (anglerfishes), collectively called the Anacanthini, are thought by some authors to be a monophyletic lineage but the concept is doubted by others. The group is characterized by the absence of a myodome (a cavity in the postorbital region of the skull in which lodge the muscles of the eye; actinopterygian fishes typically have a well-developed myodome); absence of parapophyses (singular parapophysis; a long, transverse process arising from the abdominal vertebral centrum, which serves to support epipleural ribs and, in Gadidae, the gas bladder; also called transverse process) on at least the first three vertebrae; and insertion of the first few pairs of ribs in cavities of the vertebral centra rather than on the parapophyses.

General characteristics. Members of the Gadiformes share the following characteristics: The pelvic fins are inserted below (thoracic) or anterior to (jugular) the pectoral fins, but rarely behind the pectoral fin (in some species in the family Macrouridae); there are no true spines in the fins; the dorsal and anal fins are long, with one, two, or three lobes; the scales are usually cycloid, rarely ctenoid; the premaxilla, which is protractile in some species, forms the entire margin of the upper jaw; the orbitosphenoid and basisphenoid bones are absent; the branchiostegal rays number six to eight; and the swim bladder is physoclistic. Some of the gadiforms have a complex glycoprotein antifreeze that protect them from internal ice formation at subzero water temperatures.

Classification and distribution. Concepts of the composition of Gadiformes have varied over the years. An early classification recognized only four families: Muraenolepididae, Moridae, Bregmacerotidae, and Gadidae. Then more families were recognized, including the Ophidiidae, Carapidae, and Zoarcidae, of which the first two are now placed in the order Ophidiiformes and the last in Perciformes. Currently 9 families, 75 genera, and 555 species are recognized, only one of which is limited to freshwater. About one-quarter of the world's commercial marine fishery is composed of gadiform fishes, primarily cod, haddock, pollock, and hake. In North American waters there are six families (Macrouridae, Moridae, Bregmacerotidae, Phycidae, Merluccidae, Gadidae) and 50 species.

Euclichthyidae (eucla cod). This is a monotypic family. The species *Euclichthys polynemus* is benthopelagic off New Zealand and Australia. The second dorsal and anal fins are continuous and not indented; pelvic fins are jugular, each of four long free rays, the longest reaching posteriorly to the anus; no chin barbel; maximum length 35 cm (14 in.).

Macrouridae (grenadiers). The grenadiers are classified in four subfamilies, 27 genera, and 350 species, which range in deep water from the Arctic to the Antarctic. Most of the species are benthopelagic, at depths between 200 and 2000 m (660 and 6600 ft), in tropical and subtropical waters. The second dorsal and anal fins are long and low and continuous with the tail, which tapers to a sharp point; the pelvic fins are thoracic but are behind the pectoral fins in some species; the caudal fin is absent as well as most of the caudal skeleton that supports it; the scales are small; a light organ may be present on the midline of the abdomen just in front of the anus; maximum length is 80 cm (32 in.).

In the subfamily Bathygadinae the second dorsal fin rays are longer than the anal rays; the mouth is wide and terminal; a chin is barbel present or absent; the snout is rounded; and the maximum length is 65 cm (26 in.). These fishes occur primarily from 200 to 2700 m (660 to 8900 m) in tropical to subtropical seas, but they are absent from the eastern Pacific.

The subfamily Macrouroidinae consists of two monotypic genera, one lacking a pelvic fin, the other with a small pelvic fin of five rays. The head is large, the mouth subterminal, the single dorsal fin is low; chin barbel absent; and the maximum length is 40 cm (16 in.). They are bentho- to bathypelagic fishes that occur worldwide in tropical to subtropical sea.

The subfamily Trachyrincinae consists of two genera, one lacking a chin barbel, the other with a small barbel. The second dorsal fin rays are somewhat longer than the anal fin rays, the mouth is wide and subterminal, the snout is long and pointed, the body scales are spinous, and the maximum length is 60 cm (24 in.). They occur practically worldwide in temperate waters.

The subfamily Macrourinae comprises 32 genera and 255 described species plus a number of undescribed species. The second dorsal fin rays are much shorter than those of the anal fin, the mouth is terminal to subterminal, a barbel is present or absent, many species have a ventral light organ, and the maximum length is 150 cm (59 in.) in one species but much less in the other species. They are practically worldwide but absent in high Arctic latitudes.

Moridae (morid cods). This family comprises 18 genera and about 105 species, which occur worldwide in deep waters. Morid cods have one or two, rarely three, dorsal fins and one or two anal fins; a chin barbel is present or absent; the swim bladder is connected with the auditory capsule (called otophysic connection); and some species have a light organ (see **illustration**).

Melanonidae (pelagic cods). Only two species in one genus, *Melanonus*, comprise this bathypelagic family known from the Atlantic, Indian, and Pacific oceans, as well as subantarctic waters. They have a single, long-based dorsal fin, reaching near but free from a small caudal fin; a barbel is absent; and there is no otophysic connection.

Bregmacerotidae (codlets). The one genus and 15 (probably more) species of this family are marine (rarely estuarine) fishes of tropical and subtropical seas. The first dorsal fin, consisting of a single elongate ray, is on the nape and separate from the second dorsal by

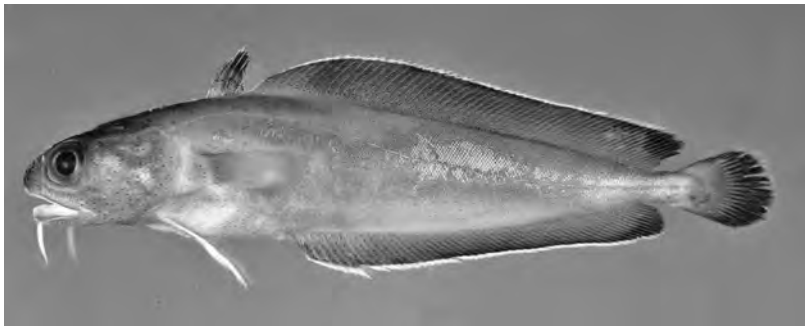

Beardie (*Lotella physcis*), a species of morid cod. (*Photo © John E. Randall*)

a long gap. The second dorsal and anal fins, long and with a wide notch in the middle, are mirror images; the pelvic fins are jugular, with five rays, the outer three of which are free filaments; and the maximum length is 12 cm (4.7 in.).

Muraenolepididae (eel cods). The four species of eel cods, in a single genus, *Muraenolepis*, occur primarily in Antarctic waters. The first dorsal fin has only two rays, the second dorsal and anal fins are continuous with the caudal fin; the pelvic fins are inserted under the gill openings, which extend upward just to the base of the pectoral fins; a chin barbel is present; and the maximum size is 40 cm (16 in.).

Phycidae (phycid hakes). Two subfamilies are recognized, Gaidropsarinae (rocklings) and Phycinae (phycine hakes). The rocklings have three closely set dorsal fins, the first a single ray, the second of low rays in a fleshy cover, and the third long and continuous to near the caudal fin; the anal fin is single and similar to the third dorsal; the snout has two to four prominent barbels plus additional rudimentary barbels in some species, and a single barbel is on the chin. The phycine hakes differ in having two dorsal fins, the first relatively short, the second long and similar to a smaller single anal fin; and snout barbels are absent but a chin barbel is present. The rocklings, comprising of three genera and 15 species, are found primarily in the northern Atlantic and to a lesser extent in the Southern Hemisphere off New Zealand and Tasmania. The phycine hakes, in two genera and 10 species, are limited to the Atlantic, including the Gulf of Mexico and the Mediterranean Sea.

Merlucciidae (merluccid hakes). This family is represented by three subfamilies. The subfamily Merlucciinae (merluccid hakes), comprising 13 species in a single genus, *Merluccius*, occurs in the western and eastern Atlantic (including the Mediterranean and Black seas), off the southern tips of India and New Zealand, and in the eastern Pacific from British Columbia to the tip of South America. There are two dorsal fins, the second deeply notched, giving the appearance of three dorsal fins; the anal fin is almost a mirror image of the second dorsal; the mouth is large and terminal and equipped with long teeth; the pelvic fins are thoracic. Commonly applied names of various species of *Merluccius* are hake and whiting; however, whiting is also used for some other gadiforms, and the genus *Urophycis* in the family Phycidae are also called hake.

The subfamily Macruroninae (southern hakes) has seven extant species in three genera. These are Southern Hemisphere fishes, primarily off the southern shores of Africa, South America, New Zealand, and Australia. The body is elongate and strongly compressed (flattened laterally); the pectoral fins are high on the body, above the axis of the body; usually the upper jaw is equipped with two rows of teeth, with those in the outer row caninelike.

The subfamily Steindachneriinae (luminous hake) is monolytic. The one species, *Steindachneri argentea*, is unique among the gadiforms in having the anus displaced forward to between the pelvic fins, leaving the urogenital pores in their normal place just in front of the anal fin. The caudal fin and skeleton are extremely reduced and the tail tapers to a fine point, and an elaborate light organ system is present on the head. The luminous hake occurs in the Gulf of Mexico, Caribbean Sea, and the Guianan shelf off northeastern South America.

Gadidae (cods). Three subfamilies are usually recognized, although many authors rank the subfamilies as families. A unifying character is a swim bladder that is not connected with the auditory capsule.

The subfamily Lotinae (cuskfishes) comprises three genera and five species. One species, the burbot (*Lota lota*), is limited to freshwater. It inhabits large deep rivers and lakes in the northern United States and throughout Canada and northern Eurasia. Some freshwater populations of another gadid, the Atlantic tomcod (*Microgadus tomcod*) of the Atlantic coast of North America, are known. The remaining three species are strictly marine in the Arctic and northern Atlantic, including the Mediterranean Sea. Lorinae are characterized by one or two dorsal fins and one anal fin; always a chin barbel, never a barbel on the snout; a rounded caudal fin; and eggs with an oil globule.

The subfamily Gadinae (cods and haddock), comprising 12 genera and about 25 species, occur in the Arctic, Atlantic, and Pacific oceans. These fishes have three dorsal fin and two anal fins, a chin barbel is usually present but no barbel on the snout; a truncate or slightly forked caudal fin; eggs without an oil globule; and maximum length 1.8 m (5.9 ft). Gadinae is considered the most advanced group in the order.

The Ranicipitidae (tadpole cod), a subfamily of Gadidae, is monotypic, with the single species *Raniceps rantus*, which occurs in shallow coastal waters of the northeastern Atlantic. The second dorsal and anal fins are continuous and not indented; pelvic fins are thoracic; a chin barbel is present; maximum length is 30 cm.

Herbert Boschung

Bibliography. D. M. Cohen, Gadiformes: Overview, in H. G. Moser et al. (eds.), *Ontogeny and Systematics of Fishes*, Spec. Publ. no. 1, pp. 259–265, Amer. Soc. Ichthyol. Herp., 1984; D. M. Cohen (ed.), *Papers on the Systematics of Gadiform Fishes*, Nat. Hist. Mus. Los Angeles City, Sci. Ser. 32, 1989; D. M. Cohen et al., *FAO Species Catalogue*, vol. 10: *Gadiform Fishes of the World (Order Gadiformes): An Annotated and Illustrated Catalogue of Cods,*

Hakes, Grenadiers and Other Gadiform Fishes Known to Date, FAO Fish. Synop. no. 10 (125), 1990; G. J. Howes, Anatomy, phylogeny and taxonomy of the gadoid fish genus *Macruronus* Günther, 1873, with a revised hypothesis of gadoid phylogeny, *Bull. Brit. Mus. Nat. Hist. (Zool.)*, 51(1):77–110, 1991; G. J. Howes, Biogeography of gadoid fishes, *J. Biogeog.*, 18:595–622, 1991; G. J. Howes, Phylogenetic relationships of macrouroid and gadoid fishes based cranial myology and arthrology, in D. M. Cohen (ed.), *Papers on the Systematics of Gadiform Fishes*, Nat. Hist. Mus. Los Angeles City, Sci. Ser. 32, pp. 113–128, 1989; G. J. Howes, The syncranial osteology of the southern eel-cod family Muraenolepididae with comments on its phylogenetic relationships and on the biogeography of sub-Antarctic gadoid fishes, *Zool. J. Linn. Soc.*, 100:73–100, 1990; D. F. Markle, Aspects of character homology and phylogeny of the Gadiformes, in D. M. Cohen (ed.), *Papers on the Systematics of Gadiform Fishes*, Nat. Hist. Mus. Los Angeles City, Sci. Ser., 32, pp. 59–88, 1989; J. S. Nelson, *Fishes of the World*, Wiley-Interscience, 1976; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, 2006; C. Patterson and D. E. Rosen, The Paracanthopterygii revisited: Order and disorder, in D. M. Cohen (ed.), *Papers on the Systematics of Gadiform Fishes*, Nat. Hist. Mus. Los Angeles City, Sci. Ser. 32, pp. 5–36, 1989.

Gadolinium

A metallic chemical element, Gd, atomic number 64 and atomic weight 157.25, belonging to the rare-earth group. The naturally occurring element is

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

composed of eight isotopes. It is named in honor of the Swedish scientist J. Gadolin. The oxide, Gd_2O_3, in powdered form is white, and solutions of the salt are colorless. Gadolinium metal is paramagnetic and becomes strongly ferromagnetic below room temperatures. The Curie point, where this transition occurs, is about 16 K. *See* PERIODIC TABLE; RARE-EARTH ELEMENTS.

Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner, Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

Gages

Devices for determining the relative size or shape of objects. The function of gages is to determine whether parts are within or outside of the specified tolerances, which are expressed in a linear unit of measurement. Gages are the most widely used production tools for controlling linear dimensions during manufacture and for assuring interchangeability of finished parts. A gage may be an indicating type that measures the amount of deviation from a mean or basic dimension, or it may be a fixed type that simply accepts parts within tolerance and rejects parts outside tolerance. *See* TOLERANCE.

Gage tolerance. If parts are to be interchangeable, it is necessary that, when one device is used to control the accuracy of another, the controlling device itself ordinarily be accurate to the next significant figure. Thus, if a dimension of a part being manufactured is to be controlled accurately in thousandths of an inch, the gages to do this must be accurate in ten-thousandths of an inch. The gages must be set or measured by instruments accurate in hundred-thousandths of an inch. The instruments must be set by gage blocks accurate in millionths of an inch. The size of the gage blocks is verified by an interferometer reading in ten-millionths of an inch.

The tolerance allowed the gage maker and its effect on the size of the part can be illustrated as follows (**Fig. 1**). If the diameter of a 2.000-in. (5.080-cm) shaft is to be controlled during manufacture to 0.004-in. (0.01016-cm) tolerance, the "go" ring gage would have a wear allowance of 0.0002 in. (0.000508 cm) and a gage maker's tolerance of 0.0002 in. (0.000508 cm). Its size would be 1.9998 − 0.0002 in. (5.079492 − 0.000508 cm). The "not go" gage would have no wear allowance, but a gage maker's tolerance of 0.0002 in. (0.000508 cm). Its size would be 1.9960 + 0.0002 in. (5.06984 + 0.000508 cm). Thus, because of the necessity for a gage tolerance, if the part tolerance must be maintained absolutely, something less than the full part tolerance is available as a working tolerance (in this case 85–95%), depending upon the actual sizes of the "go" and "not go" gages.

Gage blocks. The standard of lineal measurement for most manufacturing processes is the gage block (**Fig. 2**). Gage blocks are used for setting gages, for setting machines, and for comparative measurements.

Gage blocks are made of chrome steel and may be chrome-plated or have carbide wear faces. Individual blocks have two surfaces which are flat and parallel. The parallel distance between the surfaces is the size marked on the block to a guaranteed accuracy of 0.000002, 0.000005, or 0.000008 in. (0.0508, 0.125, or 0.203 micrometer) at an ambient temperature of 68°F (20°C). Blocks are sold in sets so that combinations of individual blocks, when wrung firmly together, will produce a new end standard equal to the sum of the sizes of the individual blocks. With a set of 81 blocks, any length expressed in four decimal places from 0.2000 in. (0.5080 cm) to approximately

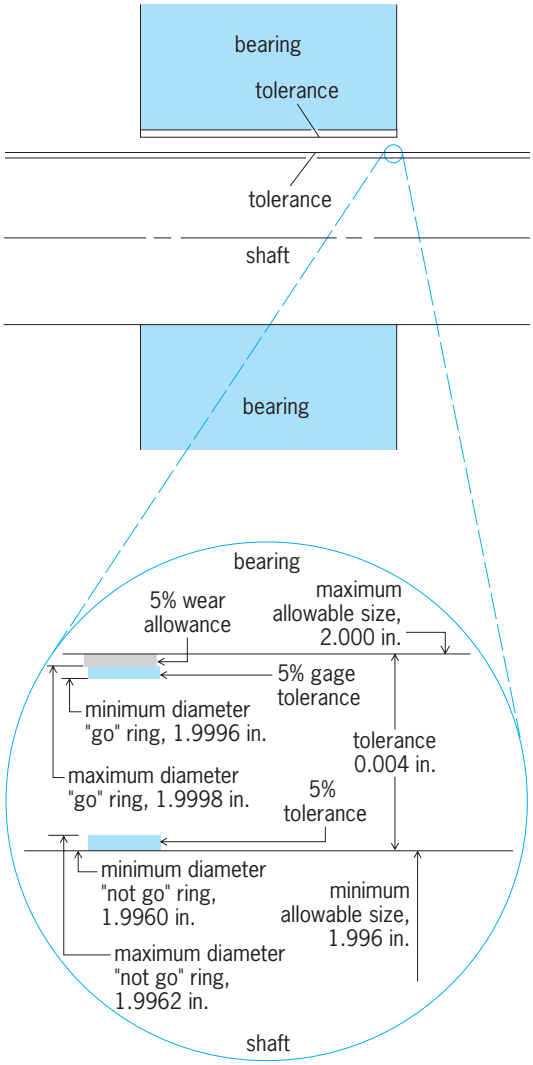

Fig. 1. Makeup of gage tolerances. 1 in. = 2.5 cm.

12.0000 in. (30.4800 cm) can be built up. Between 0.3000 in. and 4.0000 in. (0.7620 and 10.1600 cm), three to four gage groups can be built at the same time from a set to an identical length expressed in four decimal places.

Gage blocks are measured by interferometry. Their actual sizes are determined by comparison with the wavelength of red light in the color spectrum, the ultimate standard of lineal measurement being the wavelength of light. Gage blocks translate this basic lineal standard into a practical form for shop use. *See* OPTICAL FLAT.

Ring gages. Cylindrical rings of steel whose inside diameters are ground and then honed or lapped to gage tolerance are used for checking the diameter of a cylindrical object such as a shaft. The gaging surface is sometimes chrome-plated or made of carbide for wear resistance. A single gage controls one limit of tolerance of the diameter. The "go" gage will pass over the largest acceptable part but will not pass over any larger part. The "not go" gage will not pass over any part within tolerance but will pass over an undersize part. A ring gage is superior for checking the "go" dimension because it checks an infinite number of diameters simultaneously and, to the extent of its length, straightness too. Conversely, it is not preferred as a "not go" gage; an out-of-round or bent part could have many diameters undersize but still be "accepted" because the gage would "not go" if the part were bent or if even one diameter were up to the low limit. Ring gages are made to order.

Plug gage. A cylinder of steel whose diameter is ground, then honed or lapped to gage tolerance, a plug gage is used for checking the diameter of a cylindrical cavity such as a bolt hole or bearing bore. The materials, dimensions, and tolerances of plug gages are comparable to those of ring gages. A single plug gage controls one limit of tolerance of an inside diameter. Plug gages are usually used in pairs, a "go" gage and a "not go" gage. Plug gages are also made to order.

Screw-thread gage. This type of gage may be used to inspect the pitch diameter, major diameter, minor diameter, lead, straightness, and thread angle of a screw thread (**Fig. 3**). Screw-thread gages may be plug gages, ring gages, snap gages, or indicating gages. A "go" plug or ring gage inspects all these

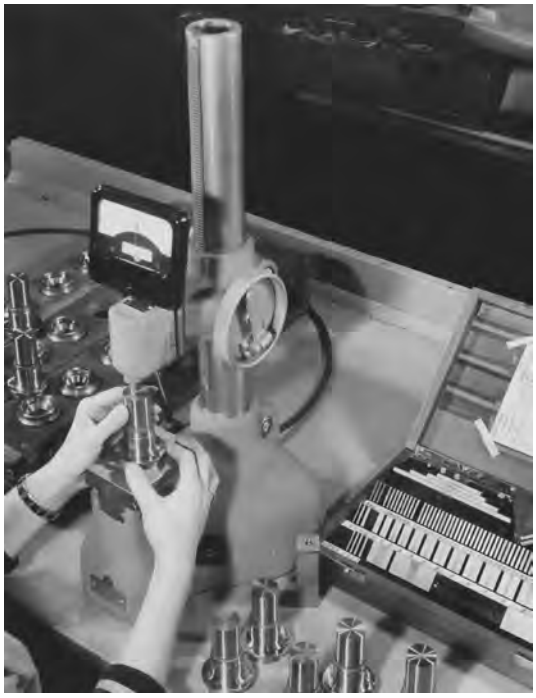

Fig. 2. General-purpose comparator of the electrical type being set with gage blocks (in box to right). (*Pratt and Whitney Co.*)

Fig. 3. Thread plug gage. (*Pratt and Whitney Co.*)

Fig. 4. Indicating gage of the electronic type being used to inspect a part. (*Federal Products Corp.*)

characteristics. A "not go" plug or ring gage checks only pitch diameter and lead. Snap or indicating gages check pitch diameter and one or more of the other characteristics. Ring gages are adjustable and can be set with a "setting" plug.

Taper pipe thread gage. Standards for pitch diameter and taper of a pipe thread were first set to control threads used in petroleum service, where safety depended on accuracy of the thread dimensions. Standards were set jointly by the American Petroleum Institute and the U.S. Bureau of Standards. Size is determined by measuring the distance that a plug gage will screw into a coupling or that a ring gage will screw onto a pipe. A manufacturer of pipe or couplings has a master plug and master ring gage certified by an approved laboratory. These are used to control working gages. As gages wear, the distance of engagement changes and limits are reset as long as thread form remains satisfactory.

Snap gage. A snap gage has two surfaces that are flat and parallel; these are spaced to control one limit of tolerance of an outside diameter or a length. A progressive snap gage has the "not go" surface just behind the "go" surface so that in a single motion both limits of tolerance can be checked. Adjustable snap gages have adjustable gaging surfaces so that the distance between gaging surfaces can be set with gage blocks and used to gage any dimension falling within the limits of the frame size.

Snap gaging is faster than ring gaging, but it checks only one diameter or length at a time. It is frequently used as a "not go" gage only, in conjunction with a ring gage for "go" limit to detect out-of-round condition in cylindrical objects. Adjustable gages have an advantage over fixed gages because allowances for wear and gage tolerance may be disregarded in their setting.

Receiving gage. A fixed gage designed to inspect a number of dimensions and also their relation to each other is termed a receiving gage because of its similarity to the cavity which receives the part in actual service. The gage checks on "go" limits. A typical example is a chamber gage which resembles the chamber of a rifle but which is made to the largest limits of size of a cartridge. Thus any cartridge which fits the receiving gage will fit any rifle chamber of that caliber.

Indicating gage. An indicating-type gage has contact points that move as they contact the part being inspected (**Fig. 4**). The movement is amplified on an indicator whose scale designates the limits of tolerance. Initial setting is done with a master or check gage. An indicator often may be substituted for a fixed or adjustable contact in which a single dimension is being checked. Snap gages are commercially available with indicators instead of fixed contact points. Indicators are also used for inspecting out-of-roundness, centerline runout, or taper, by rotating the part on one axis or diameter and indicating the "run out" of another axis or diameter. Indicating gages do not require allowances for wear or gage tolerance in their setting. *See* INSPECTION AND TESTING.

Rush A. Bowman

Bibliography. C. W. Kennedy and S. D. Bond, *Inspection and Gaging*, 6th ed., 1987; A. L. Window and G. S. Holister, *Strain Gauge Technology*, 1994.

Gain

The relative change in quantity, such as signal voltage or power, between two locations in a linear system, where the quantity at the second location depends on the quantity at the first location. It is usually expressed as the ratio of the magnitude of an output quantity to the magnitude of the associated input quantity. Equation (1) expresses the gain G as

$$G = \frac{|X_{\text{out}}|}{|X_{\text{in}}|} \tag{1}$$

illustrated in **Fig. 1**, where $|X_{\text{out}}|$ and $|X_{\text{in}}|$ denote the magnitude of the output and input signals, respectively. Since signal gain values can be very large, they are often expressed in decibels (dB), as in Eq. (2). An operational amplifier providing a voltage gain of 100,000 has a voltage gain of 100 dB. If the input and output quantities have units of power, the decibel definition is 10 $\log_{10} G$. *See* AMPLIFIER; DECIBEL; OPERATIONAL AMPLIFIER.

$$\text{Signal gain in decibels} = 20 \log_{10} G \tag{2}$$

The gain function that relates the input and output quantities often depends upon the frequency of the input excitation. Bode diagrams are commonly used to represent the response of a linear system to sinusoids with varying frequency. For a given sinusoidal input signal to a linear system, the output differs in

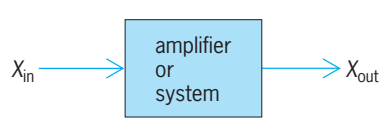

Fig. 1. Gain by an amplifier or other system.

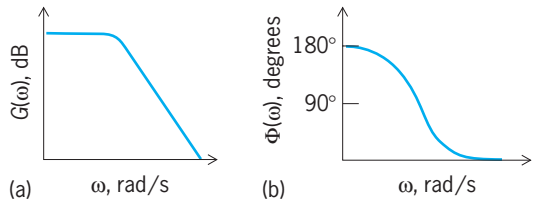

Fig. 2. Bode diagrams. (*a*) Gain plot. (*b*) Phase plot.

amplitude and phase (shift). Assuming the system in Fig. 1 is linear, if Eq. (3) holds, then Eq. (4)

$$X_{\text{in}} = A \cos(\omega t) \tag{3}$$

$$X_{\text{out}} = AG(\omega) \cos[\omega t + \Phi(\omega)] \tag{4}$$

follows, where ω is the radian frequency, $G(\omega)$ is the frequency-dependent gain, and $\Phi(\omega)$ is the frequency-dependent phase shift. The Bode diagrams show the gain $G(\omega)$ and the phase shift $\Phi(\omega)$ as functions of sinusoidal input frequency ω. **Figure 2** shows typical Bode gain and phase plots.

Gain functions and their graphical representations (Bode plots) are widely used in the analysis and design of mechanical and electrical systems. Besides gain and phase shift, information on system stability and other features is readily obtained from Bode plots. Kirk D. Peterson

Bibliography. T. G. Beckwith, J. H. Lienhard, and R. D. Marangoni, *Mechanical Measurements*, 5th ed., 1993; B. C. Kuo, *Automatic Control Systems*, 7th ed., 1995; M. E. Van Valkenburg, *Network Analysis*, 3d ed., 1974.

Galaxy, external

One of the large self-gravitating aggregates of stars, gas, and dust that contain a large amount of the visible baryonic matter in the universe. Typical large galaxies have symmetric and regular forms, are about 50,000 light-years (3×10^{17} mi or 5×10^{17} km) in diameter, and are roughly 3×10^{10} times more luminous than the Sun. The stars and other material within a galaxy move through it, often in regular rotation, with periods of a few hundred million years. The characteristic mass associated with a large galaxy is a few times 10^{12} solar masses. (The solar mass is 4.4×10^{30} lb or 2×10^{30} kg.) Galaxies often occur in groups or clusters containing from a few to many thousands of individual galaxies and ranging in size from a few hundred thousands to tens of millions of light-years. The nearest galaxy to the Milky Way Galaxy, the Sagittarius Dwarf Galaxy, is about 80,000 light-years (4×10^{17} mi or 6×10^{17} km) away; the farthest, more than 1×10^{10} light-years (6×10^{22} mi or 1×10^{23} km). Galaxies are the landmarks by which cosmologists survey the large-scale structure of the universe. *See* BARYON.

Composition. Most of the hundreds of billions of stars making up a galaxy are not generally individually observable because they are too faint and distant. Only the brightest stars in the nearest galaxies can be observed directly with large telescopes. Such stars include young, blue, and massive hydrogen-burning (that is, their energy comes from the nuclear fusion of hydrogen into helium) main-sequence stars; red, helium-burning giant stars near the end of their life cycle; evolved, luminous variable stars; and exploding (nova and supernova) stars. Although such stars may contribute most of the total visible light from a galaxy, they are few in number; most of the stars are in the form of much fainter and lower-individual-mass hydrogen-burning main-sequence stars and in the faint burned-out remnants of dead stars. The faint stars in galaxies can be studied only indirectly through the properties of their combined light. Furthermore, most of a galaxy's mass is in the form of "dark matter," a component of unknown nature that is detectable only through the effects of its gravitational pull on the stars of a galaxy. *See* DARK MATTER; NOVA; STELLAR EVOLUTION; SUPERNOVA.

The stars in a galaxy include many mixes of types of stars, depending on the evolutionary history of the galaxy. In the Milky Way Galaxy there are two general types of stellar populations: One type (population I) is characterized by the presence of young stars and by ongoing star formation. It is usually associated with the presence of gas. The second type (population II) shows an absence of gas and young stars as well as other indications that star formation ceased long ago. The Sun is a population I star. *See* STAR; STELLAR POPULATION.

Galaxies contain gas (mostly un-ionized hydrogen) in amounts varying from essentially zero up to a considerable fraction of their total mass. Dust in galaxies, although small in mass (typically 1% of the gas mass), is often dramatic in appearance because it obscures the starlight. *See* INTERSTELLAR MATTER.

Galaxies the size of the Milky Way Galaxy and larger have been confirmed to be dominated by dark matter, a form of matter whose nature is still unclear and whose existence has been confirmed only by gravitational effects on the surrounding visible matter. The dark matter distribution in galaxies increases radially toward the center of the galaxy and falls off steeply toward the outskirts of the galaxy. The existence of dark matter in galaxies smaller than the Milky Way Galaxy, so-called dwarf galaxies, has also been confirmed in several cases, especially the nearby low-surface-brightness dwarf galaxies in the constellations Leo and Draco.

Form and size. Galaxies generally display strikingly regular forms. The most common form is a disk with a central bulge. The disk is typically 100,000 light-years (6×10^{17} mi or 1×10^{18} km) in diameter and only about 1000 light-years (6×10^{15} mi or 1×10^{16} km) thick. Its appearance is characterized by radially decreasing brightness with a superposed spiral or bar pattern, or both (**Figs. 1** and **2**). The central bulge may vary in size from hundreds to many thousands of light-years. Such galaxies are classified as spirals (S) and subclassified a, b, or c (for example, Sa) to distinguish increasingly open spiral structure

and small bulge size. The disks of these galaxies are dominated by population I stars, while their bulges contain mainly population II stars. The Milky Way Galaxy is an Sb type. *See* MILKY WAY GALAXY.

Many spiral and irregular galaxies have a nearly linear feature in their central regions. Called barred galaxies, these objects can otherwise fit into the general scheme of galaxy types. The letter B is added after the S in the classification of spiral galaxies that contain conspicuous barlike features (for example, M61 is an SBc galaxy). It is likely that almost all spiral galaxies have at least some stars moving in barlike orbits.

Another common type of galaxy is an ellipsoid with radially decreasing brightness. These galaxies are classified as ellipticals (E) and subclassified according to their axial ratios by a number from 0 (E0 = round; **Fig. 3**) to 7 (E7 = 3-to-1 axial ratio). Additionally, ellipticals are subclassified according to the shape of their lines of constant brightness on images, so-called isophotes. Very luminous E galaxies show isophotes with boxlike shapes (boxy E) while less luminous E galaxies show isophotes with disklike shapes (disky E). Ellipticals may vary in size from thousands to several hundred thousand light-years. They are most commonly found in clusters of galaxies and often contain a hot gas halo and in about half of all cases appreciable amounts of dust. The brightest galaxies are usually ellipticals. They are dominated by population II stars.

Other, rarer forms of galaxies include a transition class called S0 that has a disk superimposed on an otherwise elliptical type of light distribution, and an irregular (Irr) class composed of galaxies with chaotic forms (Fig. 2) and generally low total luminosity.

Exotic galaxy types. Some galaxies lie outside the normal range of morphologies. Most of these galaxies have suffered some disturbing event, such as a gravitational encounter, a merger, or violent nuclear activity. Other such galaxies seem to have started out with anomalous characteristics, such as the low-surface-brightness galaxies. Those with activity in their nuclei are described below.

Starbursters. One of the more spectacular examples of exotic galaxies is the starbursters, galaxies that are presently manufacturing stars at an unusually vigorous rate. It is now known that some gravitational impulse has triggered the unusual star formation activity in at least most cases. The burst is a temporary condition and the galaxies now bursting must have spent most of their lives in a more quiet condition. *See* STARBURST GALAXY.

Low-surface-brightness galaxies. Another type of exotic galaxy is the low-surface-brightness galaxies, star systems that have such a low spatial density of stars that they are almost invisible. The Sculptor dwarfs have many characteristics similar to the globular star clusters (very old stars, cluster-type variable stars, and smooth stellar distribution) but are millions of times less dense, and are dominated by dark matter. Other low-surface-brightness galaxies have a composite population of stars. For example, at least two star-forming events occurred in the Carina dwarf, one about 13×10^9 years ago and one about 7×10^9 years ago. *See* STAR CLUSTERS.

Another type of low-surface-brightness galaxy includes extreme irregular and spiral galaxies, which have some of the structural properties of the normal examples of these types but are so faint that it is difficult to detect them against the sky's brightness. A significant fraction of the mass of the universe may be in the form of these nearly invisible galaxies.

Fig. 1. Great Spiral Galaxy in Andromeda (M31, NGC 224) and its two small elliptical companions (NGC 205 and 221), photographed with the 48-in. (122-cm) Schmidt telescope. (*California Institute of Technology/Palomar Observatory*)

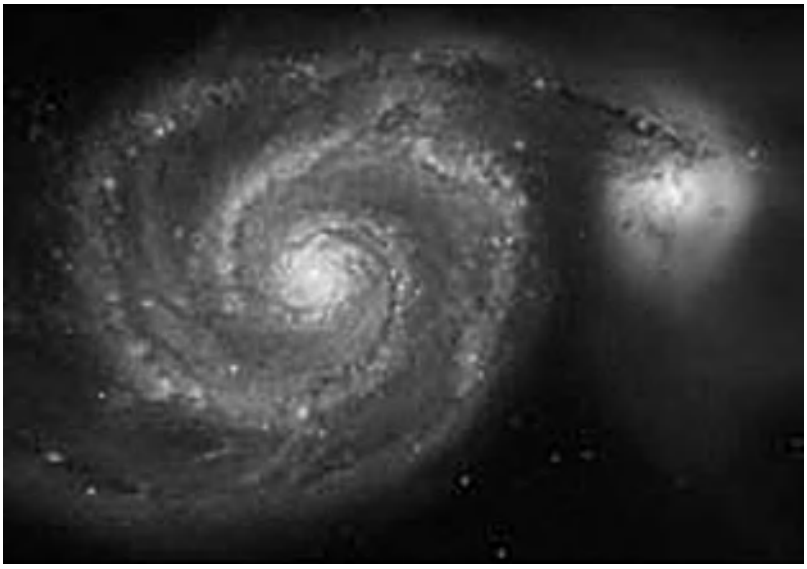

Fig. 2. "Whirlpool" Galaxy (NGC 5194), type Sc, and a companion irregular satellite (NGC 5195). [*NASA, ESA, S. Beckwith (Space Telescope Science Institute), The Hubble Heritage Team (Space Telescope Science Institute/AURA)*]

Fig. 3. E0 galaxy M87 (NGC 4486) in the Virgo cluster, which is in the constellation of Virgo. This galaxy is a source of radio emission, and it has an active nucleus. (*California Institute of Technology/Palomar Observatory*)

Internal motions. The stars and interstellar matter in a galaxy revolve around the center of mass of the galaxy, which is often a bright nucleus. In a spiral galaxy such as the Milky Way there are generally two types of stellar orbits. The stars in the flat plane tend to have nearly circular orbits, while the stars in the bulge and halo have more highly elliptical orbits. All of the motions are the result of the gravitational interactions of the stars with each other. The galaxies are supported against gravitational collapse by these motions in the same sense in which the planets of the solar system are kept from falling into the Sun by their orbital motions. Usually, the inner regions in the plane of a spiral galaxy undergo nearly solid-body rotation (velocity proportional to radius), while the outer regions rotate differentially (velocity constant). The total velocity of the material within galaxies varies from one to several hundred miles per second. Typical orbital periods for stars are several hundred million years.

The distribution of kinetic energy into randomly oriented and circular rotational motions varies with the galaxy type. The disk of a spiral galaxy may have only about 1% of its total kinetic energy in random motions, while boxy E galaxies in contrast to disky E galaxies may have most of their kinetic energy in their random component.

Luminosities. The number of galaxies with total luminosities L is roughly proportional to L^{-y}, where y is between 1 and 1.5 for luminosities less than about 3×10^{10} solar luminosities. The number is exponentially cut off for higher luminosities. The brightest observed galaxies are fainter than 2×10^{11} solar luminosities; the faintest, brighter than about 1×10^{6}. The Milky Way Galaxy's total luminosity is roughly 1×10^{10} times greater than the Sun's.

The distribution of luminosities is such that while there are very many faint galaxies, they do not contribute a large fraction of the total light given off by galaxies. Only the brighter galaxies, visible for great distances through space, can be observed easily and in great numbers. It is generally found that the masses of galaxies are roughly proportional to their luminosities.

At each luminosity the fraction of different galaxy types strongly varies with luminosity. The most luminous galaxies are mainly ellipticals and the less luminous mainly spirals. Additionally the number of elliptical galaxies at all luminosities increases in environments where the density of galaxies is high.

Clustering. Although galaxies are scattered through space in all directions for as far as they can be observed, their distribution is not uniform or random. Most galaxies are found in associations (**Fig. 4**) containing from two to hundreds of individual bright galaxies and at least 10 times as many fainter dwarf galaxies. The E and S0 galaxies tend to be concentrated in large clusters more strongly than spirals. The Milky Way Galaxy and the Andromeda Nebula (Fig. 1) are members of a cluster called the Local Group. *See* ANDROMEDA GALAXY; LOCAL GROUP.

Like the stars within a galaxy, the galaxies within a cluster move about under the influence of their mutual gravitational attraction. The motions are generally randomly oriented and the cluster shows little evidence of rotation. Typical velocities range from about a hundred up to a thousand miles per second or more.

On scales larger than individual small groups and rich clusters, the distribution of galaxies through space is still not random. This very large scale structure in the galaxy distribution is usually referred to as superclustering to indicate that it involves the higher-order clustering of the individual first-order associations of galaxies. The largest structures generally do not have the shape of spherical clusters, but rather the concentrations of galaxies tend to be filamentary in nature. Between filaments there are often large-scale voids, in which few galaxies are found.

This very large scale structure is generally attributed to the hierarchical clustering that resulted from the gravitational instability from primordial fluctuations immediately following the big bang. These would have systematically formed small galaxies first, followed by larger and larger concentrations of more massive galaxies. Evidence for the existence of the density fluctuations that are inferred to have seeded structure formation has come from the discovery and mapping of the relic temperature fluctuations in the cosmic microwave background. Indeed, computer simulations of conditions in the early universe indicate that the observed large-scale distribution of galaxies and galaxy clusters is a natural result of the evolution of the universe after the big bang, when the earliest structures began

to form in what was then a much denser universe. *See* COSMIC BACKGROUND RADIATION; COSMOLOGY; UNIVERSE.

Active nuclei. In the very central regions (sizes at least as small as a light-year, 6×10^{12} mi or 1×10^{13} km) of galaxies, violent behavior is often observed. This activity is manifested in many ways, including the high-velocity outflow of gas, strong nonthermal radio emission (implying relativistic particles and magnetic fields), intense and often polarized and highly variable radiation at infrared, optical, ultraviolet, and x-ray wavelengths, and ejection of jets of relativistic material. In the most extreme cases the energy in the nuclear activity surpasses that in the rest of the galaxy combined. These phenomena are generically referred to as nuclear activity, and the objects that exhibit them are called active galactic nuclei.

One of the most prominent characteristic features of active galactic nuclei is the ejection of large masses of high-temperature gas at great velocities. Characteristic temperatures and velocities are in the range of tens to hundreds of thousands of kelvins and thousands to over 10,000 mi/s (16,000 km/s). Total gas masses exceeding 1×10^6 times that of the Sun may be involved. These powerful gas flows reveal themselves as bright and broad spectral lines.

A second characteristic feature of active galactic nuclei is the emission of radiation over a wide range of different wavelength bands, from the radio band to gamma rays. An object such as an ordinary star, which emits radiation because it is hot, does so in a characteristic wavelength band. Such radiation is called thermal, and the characteristic wavelength is determined by the object's temperature. The typical active galactic nucleus emits a quite different sort of radiation, called nonthermal radiation, implying that it is produced by a quite different mechanism. *See* HEAT RADIATION.

Another indication that the radiation is of an unusual, nonthermal origin is the dramatic brightness changes of active galactic nuclei over relatively (by astronomical standards) short periods of time. Most active galactic nuclei show moderate changes over time scales of months to years, and some show dramatic variations (by factors of 2 to 10 or more) over times ranging down to a few hours. Since sources of radiation cannot generally change their brightnesses in times much shorter than that required for light to travel across them, these variations imply that the radiation from an active galactic nucleus arises in a very small region, in extreme cases no larger than the solar system. This is a fantastically small volume, considering that the total radiative power output can rival or exceed that of an entire galaxy many tens of thousands of light-years across.

Another important and peculiar phenomenon exhibited by active galactic nuclei is the emission of highly directional jets of relativistic plasmas and magnetic fields. These jets contain elementary particles, particularly electrons, moving at velocities near the speed of light, and often extend far outside the nucleus and even beyond the whole body of the galaxy. *See* PLASMA (PHYSICS).

Fig. 4. Clustering characteristics of galaxies. This association of galaxies in the constellation Hercules includes many different types of galaxies. (*California Institute of Technology/Palomar Observatory*)

There are a variety of classes of active galactic nuclei. The Seyfert galaxies display the broad emission lines produced by the rapid outflow of hot gas but frequently do not exhibit much radio-wavelength emission. Another complementary class shows strong radio emission but weak or absent emission lines. Yet another class (BL Lac objects, often referred to as blazars) also shows only weak emission lines but is often extremely variable. When active galactic nuclei achieve such great luminosities that they dominate that of the rest of the galaxy, they are sometimes referred to as AGNs. Quasars are the most extreme sort of active galactic nuclei, having emission so intense that the ordinary galaxy in which they exist is entirely lost in the glare of the nuclear emission. It appears that active galactic nuclei in general, and certainly quasars, were much more common during the early history of the universe than they are at present. If so, most or all large galaxies may contain the burned-out remnant of an active nucleus in their centers, although only about 1% still show detectable activity. *See* QUASAR.

Perhaps the most intriguing question concerning

active galactic nuclei is that of the nature of the energy source that drives all of their diverse phenomena. While there is no certain answer to this question, there is a widely accepted best guess or consensus model that appears consistent with all that is known about active galactic nuclei. The basic idea of this model is that active galactic nuclei are powered by the energy released when matter falls into a massive black hole occupying the center of a galaxy. These black holes are found to have masses in the rough range of 10^6–10^9 solar masses and to have formed because of the high density of material expected to accumulate at the center of a galaxy due to its gravitational field. Such a black hole will continue to accrete any gas that finds its way into the vicinity. As such gas falls toward the black hole, its angular momentum will cause it to take up a nearly circular orbit in a disk of material surrounding the black hole. This disk (called an accretion disk) will slowly inject gas into the black hole. As the gas approaches the black hole, the latter's enormous gravitational field will compress and heat the gas to very high temperatures, causing it to radiate. Intense jet radio emission, powered by energy released during infall onto the central black hole, is ejected along the minor axis of the accretion disk. Depending on the viewing angle of the observer, the resulting morphology can account for a wide variety of active galactic nuclei, radio galaxies, and quasars. A given mass of gas can release 10 or more times as much energy in this way as it could if it were used as nuclear fuel in a star or a reactor. A gas infall rate onto the central engine of several solar masses of gas per year suffices to power the most luminous active galactic nuclei in the universe. There is ample gas available in the interstellar medium to act as a fuel supply. Gas may be driven into the central regions of the galaxy, perhaps following a merger, where it can supply and activate the nucleus. The mechanisms that covert the thermal radiation generated in this way into the nonthermal radiation and relativistic plasmas observed in active galactic nuclei are not well known. *See* ASTROPHYSICS, HIGH-ENERGY; BLACK HOLE.

Paul Hodge; Edwin L. Turner; Joyce B. Turner

Correlations. A remarkable number of well-established correlations exist between structural properties of elliptical and spiral galaxies that are indicative of their origin. As mentioned above, spiral galaxies rotate like a solid body in their inner parts and differentially in their outer parts. A maximum in the rotation velocity usually occurs at the transition point between these two regimes. This maximum rotation velocity is strongly correlated to the luminosity of the spiral galaxy in such a way that it increases with luminosity. A similar correlation has been found for elliptical galaxies between the amount of random motion (the so-called velocity dispersion) of their stars and their luminosity. More luminous ellipticals show larger random motion of their stars.

The integrated color of the stars in spirals and ellipticals is redder in the center and becomes bluer going radially away from the center. The reason for this behavior is twofold. Stars that grow old become redder and stars that are enriched with elements more heavy than helium tend to be redder, also. The interpretation of the data therefore suggests that the stars in the center are older and that the percentage of elements heavier than helium increases toward the center. This latter influence is verified by the strength of spectroscopic features in the galactic starlight that measure the abundance of iron and other heavy elements.

Additionally the global percentage of heavy elements in a galaxy of any type increases with its luminosity. This is explained by the larger gravitational potential generated by the larger amount of mass in more luminous galaxies. This potential is very effective in keeping the heavy elements produced during the life cycles of stars that subsequently are ejected into the interstellar medium by supernovae. The heavy elements are subsequently incorporated in the next generation of new stars and cause them to be redder. This and an old generation of stars are responsible for the observed correlation that more luminous galaxies are redder.

Elliptical galaxies over a wide range of luminosities are compared according to their central surface brightness μ_0, central velocity dispersion σ_0, and effective radius R_0, which is the radius from within which half the light of the galaxy is emitted. The comparison shows that they do not occur in all possible combinations but occupy a two-dimensional plane, called the fundamental plane, in this three-dimensional parameter space. Thus, only certain combinations of these three parameters are realized by nature. This correlation is usually attributed to the condition that the galaxies are in an energetically relaxed state.

It is widely believed that bulges of both spirals and elliptical galaxies harbor supermassive black holes with masses up to several times 10^9 solar masses at their centers. A striking correlation between the central velocity dispersion of ellipticals and bulges of spirals with the central black hole mass has been established, which hints towards a connection in the formation of bulges and supermassive black holes. Spheroids with larger velocity dispersion have more massive black holes.

Sadegh Khochfar; Joseph Silk

Bibliography. J. Binney and M. Merrifield, *Galactic Astronomy*, Princeton University Press, 1998; J. Binney and S. Tremaine, *Galactic Dynamics*, 1987; D. M. Elmegreen, *Galaxies and Galactic Structure*, Prentice Hall, 1998; J. H. Krolik, *Active Galactic Nuclei: From the Central Black Hole to the Galactic Environment*, Princeton University Press, 1999; A. Sandage and J. Bedke, *The Carnegie Atlas of Galaxies*, 2 vols., 1994; A. Sandage, *The Hubble Atlas of Galaxies*, Carnegie Institution, 1961; S. van den Bergh, *Galaxy Morphology and Classification*, Cambridge University Press, 1998; S. van den Bergh, *The Galaxies of the Local Group*, Cambridge University Press, 2000; W. H. Waller and P. W. Hodge, *Galaxies and the Cosmic Frontier*, Harvard University Press, 2003.

Galaxy formation and evolution

How the diverse array of galaxies that are now observed originated and evolved into their present form is a topic of intense speculation.

Evidence from structural properties. Some clues can be discerned in certain structural properties of galaxies. The most plausible explanation for the smooth and round light distribution of an elliptical galaxy is that the stars formed out of a collapsing gas cloud. The rapidly changing gravitational pull experienced by different stars as the collapse proceeds has been shown by computer simulations to rearrange the stars into the observed shape.

The highly flattened disks of spiral galaxies must have formed during a similar collapse, but it is believed that most star formation did not occur until the rotating gas cloud had already flattened into a pancakelike shape. Had the stars formed at an earlier stage of the collapse, their rapid motions would have led to the formation of an elliptical galaxy. However, if the cloud stays gaseous until it flattens, much of the kinetic energy of the collapse is radiated away by gas atoms. Subsequent star formation is found to maintain a highly flattened, disklike shape, characteristic of spiral galaxies.

The flattening occurs in part because of the centrifugal forces in the rotating cloud. A confirmation of this picture has come from the discovery that elliptical galaxies rotate much less rapidly than spiral galaxies. This raises the question of the origin of the rotation itself. A natural explanation seems to lie in the action of the gravitational torques exerted by neighboring protogalaxies.

Evidence from composition. Another aspect of galaxies that has evolutionary significance is their composition, and, in particular, the distribution of heavy elements. The amount of heavy-element enrichment can be inferred from the color of the starlight, blue stars being metal-poor. Galaxies are found to be significantly bluer in their outermost regions and redder toward their central nuclei. The explanation seems to be that galaxies formed out of collapsing gas clouds that formed stars in a piecemeal fashion. As stars formed, they evolved, underwent nuclear reactions, produced heavy elements, and eventually shed enriched material (some stars even exploding as supernovae). Successive generations of stars formed out of the debris of earlier stars, and in this way the stellar content of galaxies systematically became enriched. The greatest enrichment would naturally occur toward the center of a galaxy, where the gaseous stellar debris tended to collect. *See* NUCLEOSYNTHESIS.

Interactions. Observations reveal many systems of interacting galaxies and close pairs of galaxies, which are possible candidates for later interactions. As galaxies move about within clusters, they will occasionally pass very near one another or even collide directly. Possible outcomes include loss of material from the galaxies' outer regions, transfer of material from one galaxy to another, merger of the two galaxies, modification of the galaxies' forms by tidal perturbations, and loss of gas and dust due to collisional heating. Whether a merger occurs mainly depends on the relative velocity difference of the two galaxies. If they pass each other too fast, the gravitational drag between them will not be efficient enough to change their trajectory and the passage does not result in merging.

Observations show an increase of interactions in the last few 10^9 years. The random velocity of a galaxy inside an association of galaxies increases with the mass connected to that association. Once the association forms and the galaxies within it start moving under its gravitational influence, the relative velocities in encounters of galaxies will become too high for mergers to occur and the number of mergers decreases. Looking back in time, the increase of interactions is stronger in clusters of galaxies than in environments with few galaxies. This is a consequence of the higher number density of galaxies, which increases the probability of having an encounter.

Galaxies undergoing mergers experience dramatic morphological changes. Due to tidal forces the merging galaxies start deforming and develop very prominent features, most notably the so-called tidal arms. Mergers can trigger periods of intense star formation. Gas which was available in the disks of the progenitor spiral galaxies will be driven to the centers of the remnant galaxies and start forming stars in a starburst. It has been proposed that ultraluminous infrared galaxies like the Antennae galaxies (**Figs. 1** and **2**) are just galaxies undergoing mergers in which an extensive starburst occurs. Theoretical considerations suggest that the outcome of the interaction

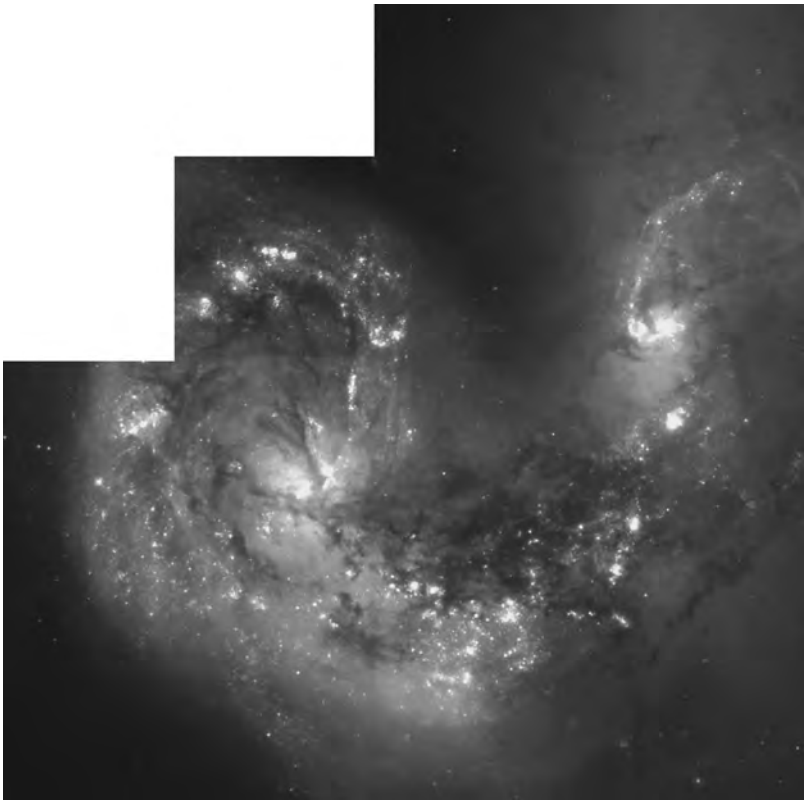

Fig. 1. Centers of the Antennae Galaxies (NGC 4038 and NGC 4039) seen in visible light. This pair of interacting (colliding) galaxies shows many sights of prominent star formation. [*B. C. Whitmore Space Telescope Science Institute), NASA*]

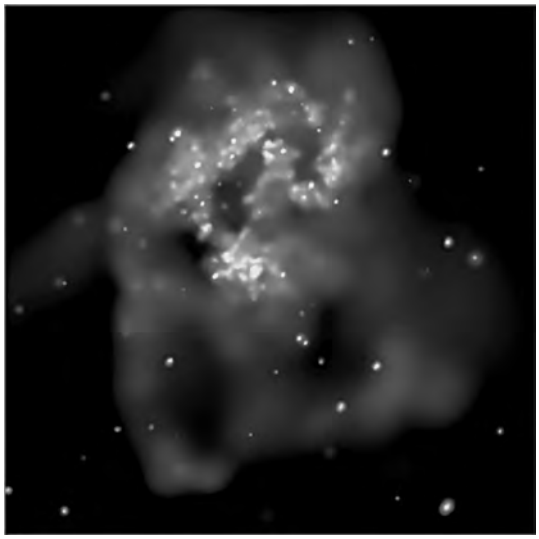

Fig. 2. Antennae Galaxies seen in x-rays. Many bright spots are seen which could be connected to black holes. A very prominent halo of hot gas is also seen. (*NASA, Chandra X-ray Observatory; Smithsonian Astrophysical Observatory; G. Fabbiano et al.*)

between two galaxies of similar size will be an elliptical galaxy. This is one of the most favored formation mechanisms for elliptical galaxies. *See* STARBURST GALAXY.

Origin. An outstanding and unresolved issue concerns the origin of the primordial gas clouds out of which the galaxies evolved. The cosmic background radiation yields a glimpse of the universe prior to the epoch of galaxy formation. The universe is now completely transparent to this radiation. However, about 500,000 years after the big bang, the radiation was sufficiently hot that matter was ionized, and the matter was also sufficiently dense to render the universe completely opaque to the radiation. To observe the background radiation now (some 10^{10} years later) is to see back to this early epoch, known as the decoupling epoch: at earlier times, matter and radiation were intimately linked, and subsequently the radiation propagated freely until the present time. *See* BIG BANG THEORY; COSMIC BACKGROUND RADIATION.

Theory of formation from fluctuations. The cosmic background radiation is very uniform, but fluctuations were discovered in 1992 by the *Cosmic Background Explorer (COBE)* satellite. These are at a level of 1 part in 10^5, and are on angular scales of several degrees. The precursor fluctuations of the primordial inhomogeneities that gave rise to the observed structures in the universe would be on scales of degrees, for the largest superclusters and voids, to arcminutes, for galaxy clusters. A definitive measurement is not yet available for these angular scales, but without such primordial fluctuations galaxies could not have formed. The mutual action of gravity exerted between these infinitesimal fluctuations results in their gradual enhancement. Eventually, great gas clouds develop that will collapse to form galaxies. The required amplitude for these primordial seed fluctuations must be of the order of 1 part in 10^5, precisely what is measured on larger scales, to within uncertainties of at most a factor of 2.

Numerical simulations of galaxy clustering have enabled the spectrum of fluctuation length scales and amplitudes to be inferred. Galaxies are not randomly distributed, as would be "white" noise; rather, they are correlated. Given a galaxy at an arbitrary position, at a distance away, there is an excess probability, above random, of finding another galaxy. These correlations are large on scales less than 5 megaparsecs (1.0×10^{20} mi or 1.5×10^{20} km), and are measured out to 20 Mpc (4×10^{20} mi or 6×10^{20} km). The parent fluctuations that gave rise to the galaxies must be similarly correlated, although the amplitude of the effect was much less.

A theory of the very early universe, first proposed in 1980, provides an explanation of the distribution of amplitudes of the fluctuations with scale, but accounts only qualitatively for their strength. According to this theory, the initial stages of the big bang were characterized by a period of rapid inflation during the first 10^{-35} s of the expansion. One consequence of an inflationary epoch is that quantum-statistical fluctuations are amplified up to scales of galaxies and of clusters of galaxies. The predicted distribution of fluctuations is initially the same, on all scales, from that of galaxy clusters to the observable universe. On the largest scales, where there has been little time to develop deviation from the initial conditions, the distribution of cosmic microwave background fluctuations measured by *COBE* in 1992 and in many subsequent experiments, especially the *Wilkinson Microwave Anisotropy Probe (WMAP)*, is approximately consistent with that predicted by the simplest inflationary cosmological model. Thus, the most simple of cosmologies may contain the nascent seeds of future galaxies. *See* INFLATIONARY UNIVERSE COSMOLOGY; WILKINSON MICROWAVE ANISOTROPY PROBE.

Isothermal and adiabatic fluctuations. The possible fluctuations in the early universe can be categorized into distinct varieties. Of particular importance for galaxy formation are density fluctuations that are found to generally be a combination of two basic types: adiabatic and isothermal.

Primordial adiabatic fluctuations are analogous to a compression of both matter and radiation. They are generic to almost all models of the early universe. In the absence of weakly interacting dark matter, the diffusive tendency of the radiation tends to smooth out the smaller adiabatic fluctuations. This process remains effective until the decoupling epoch, and galaxy formation occurs only relatively recently. However, the dominant presence of dark matter that does not interact with the radiation other than by gravity allows fluctuations to survive on all scales in the weakly interacting dark matter. The theory of fluctuation origin does not specify the strength of the fluctuations. However, the observations of the cosmic microwave background demonstrate that some 300,000 years after the big bang when the radiation was last scattered by the matter, the

amplitude of the density fluctuations amounted to only a few parts in 10^4 on galaxy cluster scales. This means that massive galaxies and galaxy clusters formed relatively recently, although small galaxies could have formed when the universe was just a tenth of its present size. Galaxies form when the gaseous matter cools and condenses in the gravity field of the dark matter, forming gas clouds that subsequently fragment into stars. *See* DARK MATTER.

In some variations of the standard model for structure formation, primordial isothermal fluctuations were also present in the very early universe. These consist of variations in the matter density, without any corresponding enhancement in the radiation density. Consequently, in the radiation-dominated early phase of the big bang, isothermal fluctuations neither grow nor decay, as the uniform radiation field prevents any motion. Once the universe becomes transparent, the matter fluctuations respond freely to gravity and grow if they are above a certain critical size. The smallest isothermal fluctuations that can become enhanced and form gas clouds contain about 10^6 solar masses. Galaxy formation occurs very early in this case.

Role of dark matter. The presence of weakly interacting dark matter is almost universally accepted by astronomers in order to account for the rotation curves of galaxies. The baryonic component of matter in the universe is known from calculations of the abundances of the light elements to amount to about 3% of the critical density for closing the universe. There is at least 10 times as much dark matter, which constitutes at least 90% of the mass of the universe. Consequently, dark matter dominates the growth of the primordial density fluctuations. The gravitational influence of this dark matter greatly aids this growth. Baryon fluctuation growth is suppressed by interactions with the radiation prior to the epoch of decoupling of matter and radiation, whereas weakly interacting particles are able to cluster freely as long as the dominant form of density is ordinary matter rather than radiation. Since the density in the very early universe was dominated by radiation, fluctuation growth in the presence of dark matter is enhanced by about a factor of 10, equivalent to the expansion factor between the epochs of ordinary matter dominance when fluctuation growth first commences and the last scattering of the radiation. The associated fluctuations in the cosmic microwave background required in order to form structures by a given epoch are reduced by a corresponding factor. The detection of cosmic microwave background temperature fluctuations at a level of about one part in 10^5, initially by the *COBE* satellite and subsequently by more than 20 experiments, means that the precursor fluctuations of the largest structures, such as galaxy clusters, have been identified, in a statistical sense, in the sky. Dark matter plays an essential role in reconciling the level of the observed fluctuations with the limited growth period available since the universe was first matter-dominated, approximately 10,000 years after the big bang. This matter consists of massive weakly interacting particles whose existence is predicted by the theory of supersymmetry. *See* BARYON; SUPERSYMMETRY; WEAKLY INTERACTING MASSIVE PARTICLE (WIMP).

The implications of particle dark matter for structure formation are considerable. If the particles are massive, they are slowly moving at the onset of fluctuation growth, when the universe is first matter-dominated. Such particles are called cold dark matter. As structure develops, cold dark matter clusters, first on the smallest scales, then on progressively larger scales. This leads to a bottom-up scenario of hierarchical clustering. Unique predictions are made for both the microwave background fluctuations and the density fluctuations that are measured in large-scale structure studies.

Reconciliation of large-scale structure and galaxy formation with cold dark matter has proven remarkably successful on the largest scales. Most data point to a universe in which the density of cold dark matter is about 30% of the critical value. The cosmic microwave background temperature fluctuations demonstrate that the universe is at critical temperature to account for the locations of the observed angular peaks in the fluctuation distribution observed on the microwave sky. In a universe that has a near-Euclidean geometry, the positions of these peaks are displaced because of the bending of the light rays that have traversed the universe since the epoch when the microwave background photons were last scattered by the matter. A Euclidean geometry requires that the universe must be at critical density, if most of the energy density is in the form of the vacuum energy that is associated with the cosmological constant term introduced by Albert Einstein. One prediction of such a cosmological model is that the expansion of the universe is currently accelerating as a consequence of the nature of the additional energy. Data from use of distant supernovae to measure the deceleration of the universe suggest that the universe is indeed accelerating. Large-scale structure still provides a severe constraint on the nature of the dark matter. *See* ACCELERATING UNIVERSE; COSMOLOGICAL CONSTANT; DARK ENERGY.

In a universe with a critical density of dark matter, excessively strong clustering of galaxies occurs. Dark energy, which is uniform and smooth, does not participate in gravitational clustering. Three independent observational results contribute to make a strong case for a standard model of the modern universe. These are the cosmic microwave background temperature fluctuation peaks, the acceleration of the universe as inferred from the distances to remote supernovae, and the large-scale structure of the galaxy distribution. The standard model of the universe consists of 30% dark matter, 65% dark energy, and 5% baryons. Variations in the standard model introduce a component of hot dark matter. This matter consists of neutrinos that are assumed to have a small mass, sufficient to account for about one-quarter of the critical density. However, the resulting mixture of hot and cold dark matter gives poor agreement with the astrophysical data on fluctuations at all scales. *See* NEUTRINO.

According to a much less accepted alternative viewpoint, the dark matter is entirely baryonic. It consists of very low mass stars or of burnt-out stars such as white dwarfs. In this case, the matter density of the universe is only about one-tenth of the critical density, with the rest of the critical density being made up of dark energy. One then finds that adiabatic fluctuations in a baryon-dominated universe, supplemented by a subdominant admixture of hot dark matter, can result in temperature fluctuations that agree with the observational constraints. A further possible advantage of this interpretation is that by reducing the Hubble constant to about two-thirds of the currently preferred value, it is also possible to dispense with dark energy if one ignores the evidence for acceleration of the universe inferred from distant supernovae. The universe would then contain a critical density of baryons, along with some hot dark matter. Another scenario appeals to warm dark matter, for which a massive neutrino is the expected candidate. However, this option requires an admixture of isothermal fluctuations in order to allow early structure formation. The isothermal fluctuations are consistent with the cosmic microwave background data, provided they are subdominant. In this case, a complex baryon genesis scenario is required, in which matter is created with spatial inhomogeneities in the number of baryons relative to the number of photons. By far the simplest model is one in which the observed structure is seeded by primordial adiabatic density fluctuations generated during inflation. *See* COSMOLOGY; UNIVERSE.

Joseph Silk; Sadegh Khochfar

Bibliography. J. Binney and S. Tremaine, *Galactic Dynamics*, 1987; P. Coles and F. Lucchin, *Cosmology: The Origin and Evolution of Cosmic Structure*, 2d ed., 2002; S. Dodelson, *Modern Cosmology*, 2003; M. S. Longair, *Galaxy Formation*, Springer-Verlag, 1998; T. Padmanabhan, *Theoretical Astrophysics*, vol. 3: *Galaxies and Cosmology*, 2002; P. J. Peebles, *Principles of Physical Cosmology*, 1993; J. Silk, *On the Shores of the Unknown: A Short History of the Universe*, 2005; J. Silk, *The Big Bang*, 3d ed., 2000.

Galilean transformations

The family of mathematical transformations used in Newtonian mechanics to relate the space and time variables of uniformly moving (inertial) reference systems.

In the simple case of two similarly oriented cartesian reference frames, moving along their common (x, x') axis, the transformation equations can be put in the form of Eqs. (1), where x, y, z and x', y', z'

$$x' = x - vt \qquad y' = y \qquad z' = zt' = t \qquad (1)$$

are the space coordinates of a given particle, and v is the speed of one system relative to the other.

The transformation equations for cartesian reference frames having arbitrary displacements and orientations take the more general form ($x_1 = x$, $x_2 = y$, $x_3 = z$) of Eq. (2), where a_1, a_2, a_3, a_4 and υ_1, υ_2, υ_3

$$x'_j = \sum_{k=1}^{3} c_{jk}(x_k - a_k - v_k t) \qquad t' = t - a_4 \qquad (2)$$

are arbitrary real numbers and the coefficients (c_{jk}) are constants. The matrix $C = [c_{jk}]$ is real and orthogonal, so that it satisfies the condition $C^{-1} = C_t$, where C^{-1} and C_t are the inverse and transposed matrices of C.

The galilean transformations form a 10-parameter group which can be generated from translations of the space and time coordinates, rotations of the space coordinate frame, and transformations to moving reference frames. *See* FRAME OF REFERENCE; SPACE-TIME.

E. L. Hill

Gallbladder

A hollow muscular organ, present in humans and most vertebrates, which receives dilute bile from the liver, concentrates it, and discharges it into the duodenum. It also participates in the enterohepatic (re)circulation of bile, and in secretion and removal of conjugated xenobiotics, including radiopaque substances taken orally or intravenously. Although not a vital organ, it stores bile, regulates biliary tract pressures, and, when diseased, enhances precipitation of various constituents of the bile as gallstones.

Extrahepatic biliary tract. The system of bile ducts lying outside the liver is known as the extrahepatic biliary tract.

In humans (**Fig. 1**) right and left hepatic ducts empty into the common hepatic duct, which

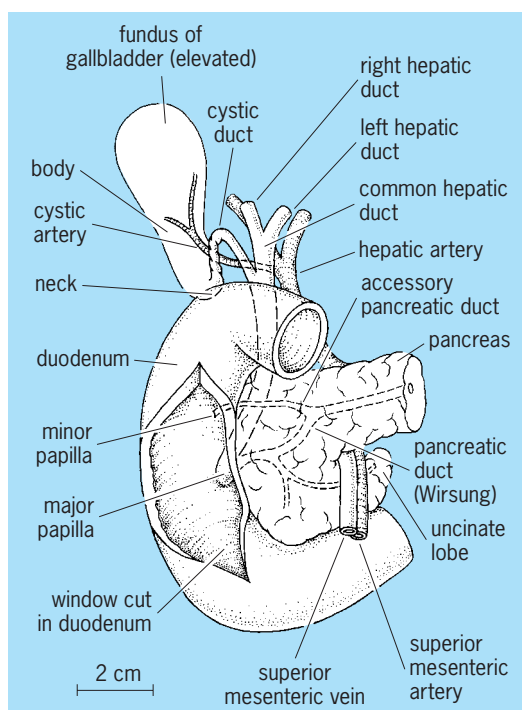

Fig. 1. Extrahepatic biliary tract in humans.

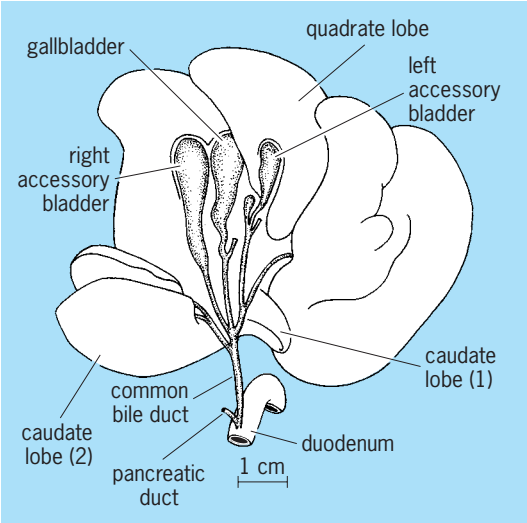

Fig. 2. Liver and accessory gallbladders of the ductule type in the cat. (*After E. A. Boyden, Amer. J. Anat., 38:177–231, 1926*)

continues to the duodenum as the common bile duct, or ductus choledochus. The gallbladder and cystic duct thus appear to be accessory organs and therefore are removable. However, they are converted into main-line structures by the presence of a sphincter at the choledochoduodenal junction. Tonic contraction of this sphincter between meals forces the bile to back up into the gallbladder. *See* LIVER; PANCREAS.

In most other vertebrates (**Fig. 2**) essentially similar relations exist except when the gallbladder is absent, but there is considerable variation in proportion and arrangement of ducts, including the pancreatic ducts. *See* PANCREAS.

Absence of gallbladder. In certain species, and as a rare anomaly in other, the gallbladder may be absent. The organ first appears in cyclostomes. It is present in all fishes, in amphibians, in reptiles, and in most birds. It is completely absent in ostriches and several other species of bird families which usually have a gallbladder (for example, it is absent in New World parrots but present in cockatoos as well as in other Australian parrots). Primitive mammals and their derivative groups—the insectivores, bats, flying lemurs, carnivores, and primates—have a gallbladder. It is absent in many later-originating herbivorous types, such as the horse, tapir, rhinoceros, elephant, peccary, deer, and camel. It is also absent in dolphins and whales. Absence of the gallbladder in the giraffe and hippopotamus is not uncommon. Conditions are most chaotic in the Eocene-developed group of rodents; among these it is missing in the common rats, castor, kangaroo rat, some porcupines, and some gophers. In humans, the gallbladder is usually embedded in the liver; it can, however, float free in approximately 4% of the normal population. Another anomaly, a folded fundus, can be found with an incidence of 2%. Additionally, in humans, agenesis (incomplete and imperfect development) and atresia (closure of normal openings) of the gallbladder occur, as well as its duplication. The cystic duct may be absent or duplicated. Moreover, the cystic duct may be abnormally long and spiral-shaped, joining the common bile duct anteriorly or posteriorly, close to its termination, or it may connect and enter the common bile duct close to its termination, or it may join the common hepatic duct; sometimes an accessory hepatic duct enters the gallbladder itself. The common bile duct can be duplicated or may develop accessory ducts. While ectopic gallbladders can be found associated with other anomalies and diseases, four different accessory gallbladders also occur in mammals, tending to be characteristic for a given species, as those of the human, cow, sheep, pig, and cat (Fig. 2).

Aberrant but vestigial bile ducts (of Luschka), originating in the hepatic tissues but not connected with the gallbladder lumen, can sometimes be observed in the adventitial layer.

Development of extrahepatic biliary tract. The way in which this system develops in humans is responsible for certain untoward clinical episodes. Both liver and ventral pancreas arise as outgrowths of a diverticulum of the embryonic gut (**Fig. 3**). Because the ventral pancreas is carried out on the bile duct, the latter becomes embedded in the head of the pancreas. Cancer of this gland can therefore obstruct the flow of bile, causing it to back up into the capillaries of the liver. Another consequence is the creation of a terminal chamber, the ampulla of Vater, into which bile and pancreatic ducts empty. Stones acting as a ball valve may lodge there; or spasms of the sphincter ampullae may result in reflux of bile into the pancreatic duct, a factor in pancreatitis. The left lobe of

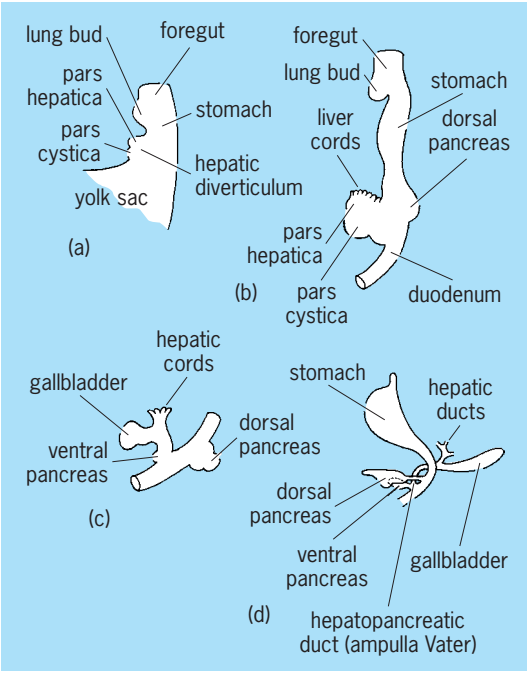

Fig. 3. Development of gallbladder (pars cystica) in human embryos. (*a*) 26 days after ovulation (horizon XII). (*b*) 28 days (XIII); 5-mm embryo. (*c*) 32 days (XV); 7-8 mm. (*d*) 35 days (XVII); 13-mm. Roman numerals indicate developmental horizons of Streeter.

the ventral pancreas may encircle the gut (annular pancreas) with consequent obstruction of the duodenum. In cats such aberrant lobes may grow up alongside the gallbladder, giving rise to a pancreatic bladder. Because of the oblique passage of the ducts through the intestinal wall, herniations or duodenal diverticula of the mucous membrane may develop. *See* LIVER DISORDERS; PANCREAS DISORDERS.

Choledochoduodenal junction. The oblique pathway through the duodenal wall traversed by the common bile duct and pancreatic duct is the choledochoduodenal junction. In humans this is surrounded by a musculus proprius which is subject to the same hormonal control as the gallbladder. This sphincter of Oddi, developing from embryonic connective tissue at a later period than the intestinal muscle, differentiates into a sphincter ductus choledochus, a sphincter ductus pancreaticus, and a sphincter ampullae, if an ampulla persists. At the point where the ducts enter the duodenum, this sheath receives auxiliary fibers from the intestinal muscle which tie the ducts to the intestine and assist in the erection of the papilla (**Fig. 4**). Because of this connection, some authors have denied the independence of the sphincter. It can act independently, however, and its spasms may result in one type of biliary colic. Physiologically its most significant component is the sphincter ductus choledochus. Tonic contraction of this sphincter causes filling of the gallbladder.

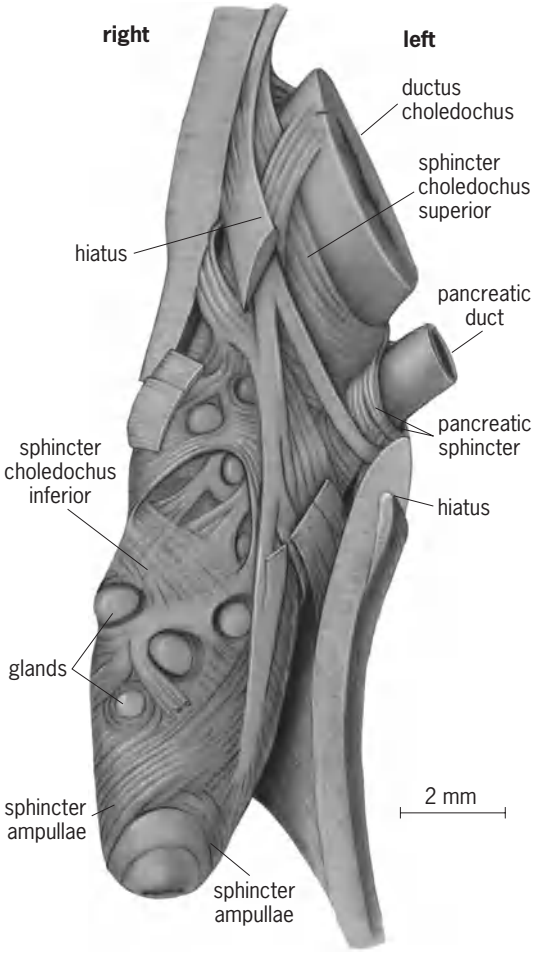

Fig. 4. Human sphincter of Oddi; upper and lower margins of window in muscle layer of duodenum through which bile and pancreatic ducts enter major duodenal papilla (Fig. 1). (*After E. A. Boyden, Surg. Gynecol. Obstet., 104(6):641–652, 1957*)

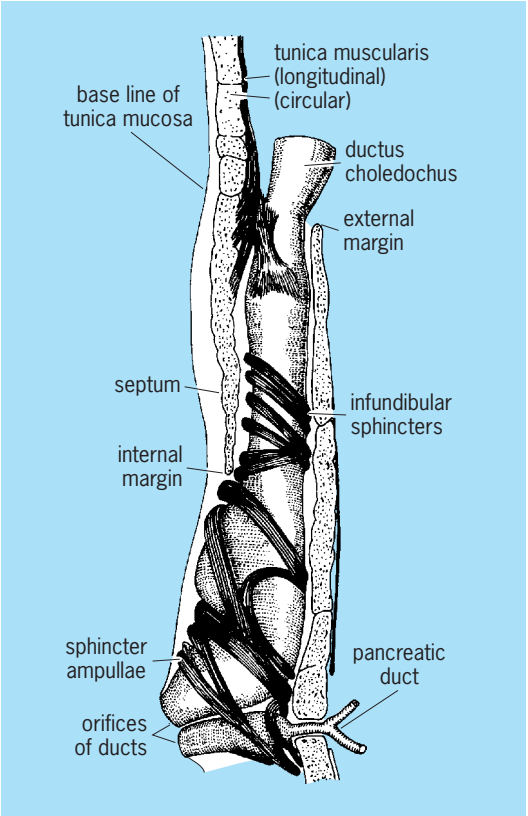

Fig. 5. Choledochoduodenal junction through wall of duodenum in newborn puppy. (*After E. P. Eichhorn and E. A. Boyden, Amer. J. Anat., 97:431–459, 1955*)

In other mammals the anatomy of the junction has been worked out in only a few species, namely the cat, dog, chimpanzee, and monkey. The first two differ from primates in that the bile duct passes through a funnel in the intestinal muscle before it enters the subterminal area where the special sphincters lie (**Fig. 5**). Therefore, bile flow in such species is regulated partly by control of intestinal activity. Experimental findings cannot be transferred from such species to humans without reservations.

Microscopic structure. The wall of the gallbladder consists of five layers: (1) a mucosal layer which lines the hollow viscus; (2) a vascular lamina propria (also named submucosal layer); (3) a fibromuscular layer containing entwined loose connective tissue and smooth muscle fiber bundles; and, where the organ is not attached to the liver, (4) a subserosal connective layer, rich in nerves, blood, and lymph vessels, (5) a serosal layer comprising a mesothelium or peritoneum. In areas where the gallbladder wall contacts the liver tissues, the subserosal layer is referred to as the adventitial layer. The mucosal layer is a simple columnar epithelium consisting of tall epithelial cells (pale and dark cells) and the subjacent lamina propria which contains a rich vascular supply, including fenestrated capillaries and lymphatics.

Both epithelium and lamina propria constitute a mucous membrane. In addition, like the intestinal lining, the mucosal epithelial cells contain lysosomal bodies and secrete mucus glycoproteins which cover the microvilli and contribute to protecting the cell surface from the potential detergent actions of the bile. This membrane concentrates the bile 7–10 times. If the membrane is diseased, these functions are lost or impaired; and the condition may be diagnosed by the Graham-Cole test, which depends on the concentration by the gallbladder of radiopaque substances administered orally. The surface area of the gallbladder mucous membrane is increased grossly by rugose folds and microscopically by countless minute cellular processes, the microvilli, each coated by a rich glycocalyx (**Fig. 6**). The surface lining epithelium is responsible for removing water and inorganic salts from the bile. These substances are transported into the basolateral spaces, then passed into the vascularized lamina propria which fills the rugae, and are evacuated by the subserosal or adventitial drainage. A rich nerve plexus originating from intramural ganglia provides numerous intraepithelial nerve endings. The surface epithelium invaginates deeply into the lamina propria and contributes to formation of pocketlike invaginations and extended rugae or folds. However, in case of inflammation, the latter may extend into and through the fibromuscular layer as pseudodiverticula of Rokitansky-Ashoff. Small mucous glands are found only in the neck of the gallbladder and within the wall of the cystic duct.

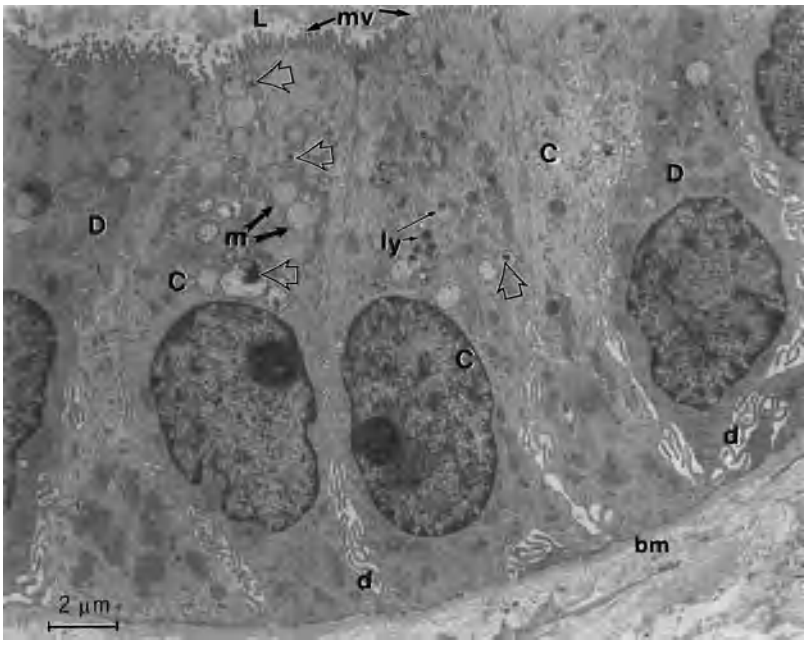

Fig. 6. General aspect of a normal epithelium from a sample of human cholecystic gallbladder. Columnar undifferentiated dark (D) and differentiated clear (C) cells produce mucus secretory granules (m) in their Golgi and subapical zones. Most granules contain proteinaceous electron-dense cores (large open arrows). Note the apical surface, lining the organ lumen (L), covered by numerous microvilli (mv); d, basolateral digitations; bm, basement membrane; ly, lysosomal bodies. (*After J. Gilloteaux et al., Amer. J. Anat., 184:323–335, 1989*)

Comparative physiology. In humans, evacuation of the gallbladder is accomplished by a trigger mechanism which is set off by the presence of fatty foods, meat, and hydragogue cathartics in the duodenum and upper jejunum. Absorption of these substances by the mucous membrane results in the release of cholecystokinin (CCK), a hormone which rapidly circulates in the bloodstream and simultaneously produces contraction of the gallbladder and relaxation of the sphincter of Oddi. The most effective food is egg yolk which contains certain *l*-amino acids. By calculating the changing volumes of the gallbladder in a series of x-rays, it has been shown that the first phase of contraction is the most important, because two-thirds to three-fourths of the contents is expelled within the first 40 min (**Fig. 7**). Resorption of bile salts by the intestine stimulates secretion of bile for hours after a meal. Before puberty, the male gallbladder empties faster than the female but it is slower afterward. Pregnancy and pernicious anemia in females markedly slow the rate. In peptic ulcer patients the rate of emptying is accelerated because initial food escapes faster from the stomach. After section of both vagus nerves, the gallbladder merely resumes its normal rate. For this and other reasons, it is believed that evacuation of the biliary tract in humans is controlled primarily by hormones.

In laboratory mammals the situation is more complicated. In animals below primates, gallbladder and sphincter are under control of both hormones and the autonomic nervous system. In cats contraction of the gallbladder can be inhibited within 1 min by inducing a spasm of the gut. The resulting reflex, from gut to gallbladder, can be blocked by sectioning the splanchnic nerves. Comparable experiments with human beings, in which medical students swallowed electrodes into the duodenum, failed to inhibit the gallbladder. Section of the right vagus in cats triples the time required to empty the gallbladder. Partial vagotomy in humans does not reduce the rate of emptying, and total vagotomy merely returns it to normal.

Several human clinical and comparative studies confirm and complete the aforementioned observations because the gallbladder appears to possess a rich ganglionated nerve plexus where circulating hormones, inflammation-related factors, and several co-localized neuropeptides can affect or modulate the actions of the acetylcholine release of the main vagal preganglionic gallbladder innervation. Humoral cholecystokinin has been confirmed to be the main postprandial stimulus for gallbladder contraction; it can readily access these gallbladder ganglia since there is no blood-ganglionic barrier. As a result of this lack of barrier, cholecystokinin can act presynaptically within the gallbladder ganglia, facilitating the acetylcholine release which in turn initiates gallbladder contraction via stimulation of the intramural intrinsic nerve plexus on the organ musculature. Norepinephrine and prostaglandin E_2 are known to antagonize the acetylcholine action.

The sphincter of Oddi also displays a ganglionated nerve plexus which contains 5-hydroxytryptamine (5-HT) neurons and fibers, suggesting that 5-HT plays a role in the sphincter regulation. The recent findings that duodenal sensory neurons project to the

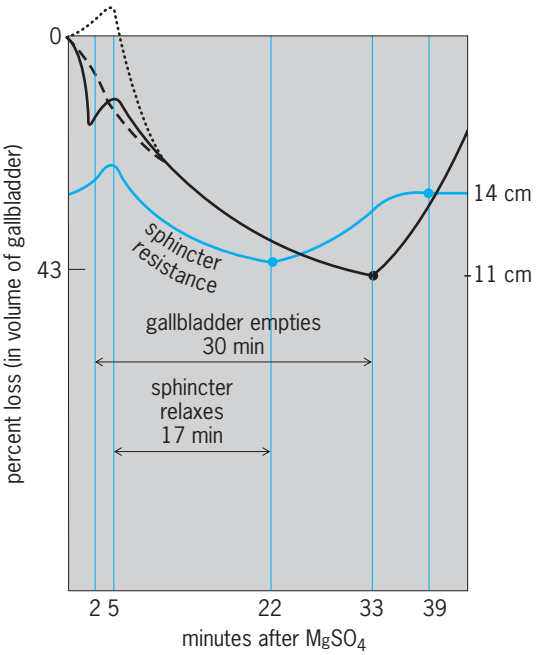

Fig. 7. Reaction of gallbladder and sphincter of Oddi to magnesium sulfate in humans. A mean drop from 14 to 11 cm Hg in pressure within common duct in 12 cholecystectomized patients; 43%, mean percent of bile expelled by gallbladder in 10 male students. When egg yolk was given, the times remained the same, but sphincter resistance fell from 14 to 7 cm of water, and gallbladder expelled 71% of its contents. (*After E. A. Boyden, G. S. Bergh, and J. A. Layne, Surgery, 13:723–733, 1943*)

sphincter of Oddi ganglia made plausible that these connections may detect the postprandial release of cholecystokinin from the mucosa and therefore provide the signal to decrease sphincter of Oddi resistance, and hence facilitate the flow of bile into the duodenum.

Functions in different species. Animals without a gallbladder, such as the pocket gopher, pigeon, horse, and rat, do not compensate for its absence by secreting a more concentrated bile. However, the liver of such animals secretes a relatively large quantity of bile. Conversely, animals secreting a relatively small quantity of bile, such as humans, the mouse, dog, and cat, have gallbladders which concentrate well. The gallbladders of the sheep, goat, cow, and pig appear to manifest little, if any, power of concentrating the bile. Correspondingly, the sphincter offers little or no resistance to the flow of bile. In humans and dogs, the gallbladder has both a storage and pressure-regulatory function; in the rabbit, sheep, cow, goat, and pig, only a pressure-regulatory function; and in still others, such as the guinea pig, no apparent function because its removal results in no grossly demonstrable changes. *See* DIGESTIVE SYSTEM.

Edward A. Boyden; Jacques Gilloteaux

Bibliography. P. W. Brunt and A. E. Read (eds.), *The Liver and Biliary System*, 1984; W. Ganong, *Physiology: A Study Guide*, 2d ed., 1987; J. Gilloteaux, The biliary tract, the gallbladder and gallstones, *Microsc. Res. Tech.*, 38(6) and 39(1), 1997; A. P. Gokin et al., *Amer. J. Physiol.*, 271:G876–G883, 1996, and 272:G1365–G1371, 1997; K. Hillsley and G. M. Mawe, *Amer. J. Physiol.*, 275(5, pt 1):G1018–G1027, 1998; W. P. Hopman et al., *Gastroenterology*, 89:1242–1247, 1985; A. S. Romer and T. S. Parsons, *The Vertebrate Body*, 6th ed., 1986; K. Schmidt-Nielsen et al. (eds.), *Comparative Physiology*, 1978.

Gallbladder disorders

Although not essential to life or health, the gallbladder is the site and source of appreciable suffering and disease in humans. With its cystic duct, the gallbladder constitutes a blind-ended, lateral extension of the common bile duct. It is roughly the size and shape of a small pear, and is adherent to the outer surface of the liver. Besides acting as a reservoir for bile, the gallbladder concentrates and otherwise alters the composition of bile. *See* GALLBLADDER.

Gallstones. Gallstones are round, oval, or faceted concretions formed within the gallbladder from the salts and pigment of bile. The size may be upward of 1 in. (2.5 cm) in diameter. Such stones may also be formed in any of the bile ducts within or outside the liver, but the incidence in the bile ducts is low compared to the number originating in the gallbladder. The occurrence of gallstones increases with age, and 20–30% of all elderly adults have one or more such stones. They are particularly common in women, blacks, and individuals with diabetes. Pure stones account for 10% of all gallstones, and are composed of cholesterol, calcium bilirubinate, or calcium carbonate, in decreasing frequency. Mixed types make up 80% of all gallstones and are most often associated with cholecystitis. The remaining 10% are the combined forms that usually begin as pure stones, with other basic materials being subsequently deposited.

Although the mechanism and reason for their formation are not clearly understood, the major predisposing factors are stasis (prolonged retention of bile in the gallbladder), abnormal composition of the bile (excessive amounts of cholesterol, bilirubin, or calcium), and infection. Passage of a gallstone through the ducts into the duodenum usually produces severe pain, called biliary colic. If a stone causes obstruction of the ducts, the result may be damage to the liver, pancreas, biliary system, and related structures either directly or through concomitant inflammation. Gallstones are rare in animals, although they have been found in nearly all species, especially in cattle. *See* BILIRUBIN; CHOLESTEROL; CIRRHOSIS.

Cholecystitis. Cholecystitis, or inflammation of the gallbladder, is a common disease in humans; as many as 70% of all elderly patients show anatomical evidence of having had one or more attacks. It is nearly always associated with gallstones and is particularly common in obese middle-aged women. It is rare in animals. Most cases are thought to be the result of chemical irritation caused by excessively concentrated bile, which is in turn the result of partial or complete obstruction to the outflow of bile. The gallbladder that is irritated and inflamed in this fashion is susceptible to secondary bacterial infection, and this

in turn may spread throughout the bile ducts into the liver tissue. The inflamed gallbladder is usually enlarged because of edema of its wall and the presence of pus in its lumen. Perforation may occur, leading to the development of peritonitis or to abnormal communications with the intestine. Prolonged or recurrent episodes of inflammation result in chronic cholecystitis, characterized by thickening and scarring of the wall, contraction, and impairment of normal function. The latter may be detected clinically by administering a radiopaque dye which in a normal gallbladder is concentrated sufficiently to produce a shadow on the x-ray film. The absence of such a shadow indicates a failure of this concentrating ability and suggests past or present disease.

Tumors. Malignant tumors are less common in the gallbladder than in several other parts of the intestinal tract, such as the stomach, colon, and pancreas, but make up as much as 3% of all malignant tumors in humans. They are almost invariably associated with the presence of gallstones. Because they produce little in the way of symptoms, and because they very soon invade the liver, these tumors are rarely curable by surgical therapy at the time they are discovered. Similar carcinomas arise in the common bile duct and in the ampulla of Vater, but there the prognosis is somewhat better. In the large bile ducts, tumors usually cause biliary obstruction and jaundice early in their course and may be surgically resectable if the formidable technical problem of reestablishing biliary and pancreatic drainage can be surmounted. Benign tumors of the gallbladder and ducts are rare in humans, and in animals both benign and malignant tumors of these sites are extremely uncommon. *See* LIVER DISORDERS. Milton R. Hales

Bibliography. T. C. Jones and R. D. Hunt, *Veterinary Pathology*, 6th ed., 1997; L. Okoliesanyi et al. (eds.), *Assessment and Management of Hepatobiliary Disease*, 1987; M. H. Sleisenger and J. S. Fordtran, *Gastrointestinal Disease: Pathophysiology, Diagnosis, and Management*, 4th ed., 1989; S. Sherlock, *Diseases of the Liver and Biliary System*, 10th ed., 1997.

Galliformes

A large order of birds containing the gallinaceous, or chickenlike, birds. They are found worldwide, although several of the subgroups have a more limited distribution. Part of the superorder Neognathae, the galliforms are closely related to the anseriforms and are often placed with them in a separate group, the Galloanserae. The gallinaceous birds include the most important domesticated birds, the chicken (*Gallus gallus*) and the turkey (*Meleagris gallopavo*); the rapidly breeding Japanese quail (*Coturnix japonica*) is an important species for much scientific work. The hoatzin of South America is often classified as a galliform suborder, but here it will be treated as a distinct order, the Opisthocomiformes. *See* ANSERIFORMES; AVES; OPISTHOCOMIFORMES.

Classification. The Galliformes are divided into two suborders with two families.

Order Galliformes
- Suborder Cracoidea
  - Family: Megapodiidae (megapodes or brush turkeys; 7 species; Australasia to southeast Asia)
  - Cracidae (curassows, guans, chachalacas, 49 species; New World, mainly Neotropics)
- Suborder Phasianoidea
  - Family Phasianidae (grouse, pheasants, quail, guinea fowl, turkeys; 212 species)
    - Subfamily: Tetraoninae, (grouse; Holarctic; temperate regions of the Northern Hemisphere)
    - Phasianinae (pheasants, peafowl, chicken; Old World to Australia)
    - Odontophorinae (New World quail)
    - Numidinae (guinea fowl; Africa)
    - Meleagidrinae (turkeys; Mexico to North America)

Although the galliforms are usually divided into two suborders, Cracoidea and Phasianoidea, and the Megapodiidae and Cracidae are considered primitive for the entire order, these conclusions must still be documented carefully.

Characteristics. The fowls are small to large birds with a short, arched bill for pecking seeds, insects, and vegetable matter. Their wings are short and rounded, and their flight muscles are large but myoglobin-poor (white); hence they fatigue quickly. Their flight is explosive but of short duration and is used mainly for purposes of escape. Although the fowls fly infrequently, no Recent species is flightless. Their legs and feet are strong, as they spend most of the time on the ground in search of food. Males of many species, especially pheasants, have colorful plumage and elaborate wing and, especially, tail feathers that are used in courtship displays. The train of the male peafowl, with its many complex ocelli (round, light-sensitive eyespots) and lacy appearance, is only an extreme of the trend seen in other species of galliform birds; moreover, this train is composed of upper covert feathers, not tail feathers. Courtship and pair bonding are brief; several species court communally in courtship display areas known as leks. The females are responsible for incubation and care of the young, which are downy and leave the nest immediately after hatching.

Cracoidea. The cracids (chachalacas and curassows) are restricted to South and Middle America and are the most arboreal members of the order; they have never hybridized with members of other families and may represent the most primitive group of galliform birds.

The megapodes (mound birds) do not build a nest or incubate the eggs. Rather, the eggs are buried in a mound of decaying vegetation, in warm sand, or in

White-tailed ptarmigan (*Lagopus leucurus*), Glacier National Park, Montana. (*Photo by Gerald and Buff Corsi, © 2001 California Academy of Sciences*)

earth heated by thermal vents. The young develop extensively, hatch fully feathered and self-sufficient, and can fly soon after hatching. The megapodes of Australia, and New Guinea to the Philippines and the Nicobar Islands are the only birds that do not exhibit parental care.

Phasianoidea. Grouse are Holarctic with colorful fleshy areas and pouches on the head and neck that are used in courtship. Some species eat conifer needles during the winter and/or burrow in the snow during the night to conserve heat. The ptarmigans (*Lagopus*), which have a brown summer plumage, have evolved a special white winter plumage that serves as camouflage, a characteristic that is unique in birds (*see* **illustration**).

Found in forests, turkeys are large galliforms with dark plumage that is metallic-colored in males. The skin of the head and neck is bare, and in males becomes red and blue during courtship. Males have a breast "brush" of unique specialized epidermal appendages, not of modified feathers.

The common quail of Eurasia (*Coturnix coturnix*) is notable because it is one of the very few migratory members of this order and it is differentially toxic to humans, which has been known since biblical times but is still poorly understood.

Fossil record. The earliest galliform fossil is *Gallinuloides* from the lower Eocene of Wyoming and may be an early cracid. A scattered fossil record of younger forms exists for all major groups of galliforms, but it does not clarify the relationships or evolutionary history of the order. Megapode fossils are interesting in that several giant, flightless forms are known from the Quaternary, including *Sylviornis* from New Caledonia, which was originally believed to be a ratite bird. *See* STRUTHIONIFORMES.

Economic significance. The galliforms have been one of the most economically important groups of birds to humans. Almost all species have been hunted for food, and many are still important as game birds for hunters. Most species are kept in captivity by aviculturists and are perhaps the most important group of domesticated birds, particularly chickens and turkeys. Chickens were domesticated from red jungle fowl by primitive humans in southeast Asia, rapidly spread over the Old World, and arrived in the New World with European explorers. Divergent breeds of chickens have been developed for egg and meat production. The wild turkey, which is found throughout most of the United States and Mexico and was domesticated by Mexican Indians in pre-Columbian times, was brought to Europe by early explorers. It is used exclusively for meat production. The Indian peafowl (*Pavo cristatus*) has been kept in captivity as an ornamental bird, and a white variety has been developed. One species of guinea fowl, *Numida meleagris*, has been domesticated on a much smaller scale. The ring-necked pheasant (*Phasianus colchicus*) of Asia has been introduced in many parts of the world as a game bird. Several other species have also been introduced as game birds, but with less success. Some domesticated gallinaceous birds have escaped and established themselves as feral birds.

Many species of galliforms have become greatly reduced in numbers and some threatened because of overhunting and habitat destruction. Some species, especially the forest pheasants of Asia, are seriously endangered. *See* ENDANGERED SPECIES.

Walter J. Bock

Bibliography. T. M. Crowe, The evolution of the Guinea-fowl, *Ann. S. Afr. Mus.*, 76:43–136, 1978; J. Delacour and D. Amadon, *Curassows and Related Birds*, 2d ed., Lynx Edicions, 2004; P. A. Johnsgard, *The Grouse of the World*, University of Nebraska Press, 1983; P. A. Johnsgard, *The Pheasants of the World: Biology and Natural History*, 2d ed., Smithsonian Institution, 1999; P. A. Johnsgard, *The Quail, Partridges, and Francolins of the World*, Oxford University Press, 1988; D. N. Jones, R. W. R. J. Dekker, and C. S. Roselaar, *The Megapodes*, Oxford University Press, 1995; S. Madge, S. McGowan, and P. McGowan, *Pheasants, Partridges and Grouse*, Princeton University Press, 2002.

Gallium

A chemical element, Ga, atomic number 31, atomic weight 69.72. Gallium is a member of group 13 and the fourth period of the periodic table (IUPAC). *See* PERIODIC TABLE.

The major commercial sources of gallium are bauxite, containing gallite ($CuGaS_2$), and zinc and germanium sulfides. Normal ore-grade deposits usually

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb

actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

contain substantially less than 0.1% gallium. In the United States the bauxite deposits in Arkansas and the zinc deposits in Oklahoma are the main sources of domestic production. Much of the gallium used in the United States is imported from Switzerland and Germany, with lesser amounts from Canada and France. *See* BAUXITE.

Gallium is a unique element in that it possesses the largest liquid range of any element. Its normal freezing point of 29.78°C (85.60°F) is lower than any metal except mercury and cesium. Its boiling point is in the vicinity of 2420°C (4388°F), although there is some uncertainty owing to the reactivity of gallium with the container material at this temperature.

The valence-electron notation of gallium corresponding to its ground-state term is [Ar, $3d^{10}4s^24p^1$], which accounts for the maximum oxidation state of III in its chemistry. Compounds of formal oxidation state II and I are also known.

Approximately 95% of the gallium consumed in the United States and presumably in the world is used in the electronics industry. Minor quantities have been used or studied for use in thermometers, low-melting solders, as a heat-transfer fluid, in arc lamps, batteries, vanadium-gallium superconductors, and in catalyst mixtures.

The most important gallium semiconductors are gallium arsenide (GaAs) and gallium phosphide (GaP). The magnitude of the energy gap in GaAs favors its use in transistors. The electron mobility in GaAs is very much higher than the hole mobility; in contrast, the electron and hole mobility in GaP are of similar magnitude and very much lower than in GaAs. By doping with the appropriate elements, these properties can be altered. Electron transport (*n*-type) GaP semiconductors are used in rectifiers, hole transport (*p*-type) in light sources and photocells. *n*-Type GaAs semiconductors are used in injection lasers and *p*-type GaAs in electroluminescent transistors. *See* HOLE STATES IN SOLIDS; SEMICONDUCTOR.

GaN is prepared by the reaction of metallic gallium or Ga_2O_3 at elevated temperature with ammonia, and the other semiconductors by direct reaction with the elements or Ga_2O at high temperature. *See* ALUMINUM; INDIUM; THALLIUM. Edwin M. Larsen

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; A. J. Downs (ed.), *Chemistry of Aluminium, Gallium, Indium and Thallium*, 1993; W. F. Gale and T. C. Totemeier, *Smithells Metals Reference Book*, 8th ed., 2004; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; K. T. Tsen (ed.), *Ultrafast Physical Processes in Semiconductors: Semiconductors and Semimetals*, vol. 67, 2000.

Galvanizing

The generic term for any of several techniques for applying thin coatings of zinc to iron or steel stock or finished products to protect the ferrous base metal from corrosion; more specifically, the hot dipping that is widely practiced with mild steel sheet for garbage cans and corrugated sheets for roofing, sheathing, culverts, and iron pipe; and with fencing wire. During dipping, molten zinc reacts with the steel to form a brittle zinc-iron alloy. Control of temperature and the addition of aluminum reduce formation of the alloy, resulting in a more ductile coating. For marine use, magnesium is added.

An electrolytic process (also called cold galvanizing or electrogalvanizing) is also used for wire, as well as for applications requiring deep drawing. An alloy layer does not form, hence the smooth electroplated coating does not flake in the drawing die. *See* METAL COATINGS. Frank H. Rockett

Bibliography. R. M. Burns and W. Bradley, *Protective Coatings for Metals*, 3d ed., 1975; J. M. Hornsby, *Hot Dip Galvanizing: A Guide to Process Selection and Galvanizing Practice*, 1994; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., vol. 25, 1997; F. Lowenheim, *Modern Electroplating*, 3d ed., 1974.

Galvanomagnetic effects

Electrical and thermal phenomena occurring when a current-carrying conductor or semiconductor is placed in a magnetic field. The galvanomagnetic effects are closely related to the thermomagnetic effects. Both sets of effects yield important information on the band structure of metals and semiconductors and on the nature of the conductivity process. *See* THERMOMAGNETIC EFFECTS.

Let the electric current density j be transverse to the magnetic field H_z, for example, along x. Then the following transverse-transverse effects are observed: (1) Hall effect, an electric field along y, as shown in Eq. (1), where R is the Hall coefficient. According

$$E_y = Rj_xH_z \tag{1}$$

to experimental conditions, R may be the adiabatic or the isothermal Hall coefficient. (2) Ettingshausen effect, a temperature gradient along y, as in Eq. (2),

$$\frac{\partial T}{\partial y} = Pj_xH_z \tag{2}$$

where P is the Ettingshausen coefficient. Also the following transverse-longitudinal effects are observed: (3) Transverse magnetoresistance, an electrical potential change along x. (4) Nernst effect, a temperature gradient along x. *See* HALL EFFECT; MAGNETORESISTANCE.

Let the electric current density j be along H. Then, the most important effect is longitudinal magneto resistance, or an electrical potential change along H.

These effects arise from the Lorentz force on moving charges in a magnetic field. They are complicated by interaction between these charges and the lattice in which they move. For this reason, the signs of R and P may be negative or positive, depending upon

composition and in many cases on temperature; the magnitudes show tremendous range.

Elihu Abrahams; Frederic Keffer

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; R. E. Prange and S. M. Girvin (eds.),*The Quantum Hall Effect*, 2d ed., 1990; H. Weiss, *Structure and Application of Galvanomagnetic Devices*, 1969.

Galvanometer

A device for indicating very small, usually direct, electric currents. Although the deflection of a galvanometer results from current in the moving coil, the voltage in a closed circuit producing this current is frequently the quantity of interest to the user. In this mode, galvanometers are used to detect a null or an unbalanced condition in a bridge or potentiometer circuit. Galvanometers have played a pivotal role in the evolution of electrical measurements but electronic instruments that employ amplifying circuits to achieve sensitivities approaching the nanovolt level are now more frequently used instead.

D'Arsonval galvanometer. This is the most common type and is widely used. Its indicating system consists of a light coil of wire suspended from a copper or gold ribbon a few thousandths of an inch wide and less than 0.001 in. (0.025 mm) thick. The coil, free to rotate in the radial field between the shaped pole pieces of a permanent magnet, is attached to a pointer moving over a scale. In the most sensitive instruments, the "pointer" is a beam of light that is reflected off a mirror attached to the coil onto a fixed scale (see **illustration**). Current is conducted to and from the coil by the suspension ribbons. The torque which deflects the indicating element is produced by the reaction of the coil current with the magnetic field in which it is suspended, and is opposed by the twisting of the suspension ribbon. *See* MAGNETISM.

Sensitivity is optimized by choosing a coil resistance that matches that of the circuit as measured between the points to which the galvanometer is connected. Oscillations of the coil on applying or changing the current are damped out by the circuit resistance or by winding the coil on a short-circuited metal former.

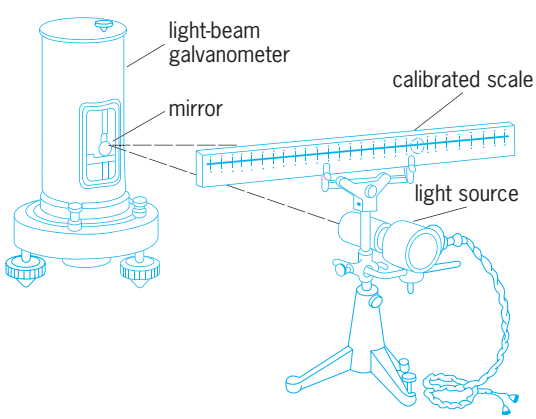

Light-beam galvanometer, showing the detached scale and the light source as used in laboratory work. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

The sensitivity of galvanometers ranges up to 0.04 in. (1 mm) of deflection, on a scale 40 in. (1 m) distant from the mirror, for a current of 0.00001 microampere. Such a galvanometer may have a coil resistance of 800 ohms and a critical damping resistance of 100,000 Ω. The voltage response of this instrument amounts to 0.04 in./microvolt (1 mm/μV) at critical damping. A galvanometer by the same maker, designed for voltage sensitivity, has a coil resistance of 20 Ω, a critical damping resistance of 30 Ω, and a response of 0.04 in. (1 mm) for 0.05 μV in a critically damped circuit. It will be seen from these examples that a large response to current is associated with large coil resistance and high critical damping resistance, whereas voltage response is associated with low coil resistance and low critical damping resistance.

Portable galvanometers. Portable and some wall-type instruments have their coils suspended between upper and lower taut ribbons so that the motion is restricted to rotation about the coil axis. Such taut-suspension systems, however, are stiffer than those having the lower suspension loosely coiled, and such galvanometers are usually less sensitive. Pivoted galvanometers also have their motion confined to rotation. In both of these types the moving system is mechanically balanced around the axis of rotation, and the effect of mechanical disturbances is much reduced.

Ballistic galvanometer. The conventional galvanometer may also be used ballistically to integrate a transient current or voltage pulse of short duration. The electric charge in a capacitor may be measured by allowing it to discharge through a galvanometer and observing the magnitude of the first swing of the moving system from its rest position. A galvanometer's response to current of short duration is called its coulomb sensitivity. If the current pulse is short, the deflection of the moving system is proportional to the time integral of the current, that is, to the electric charge in the pulse. Similarly, a voltage pulse rising from a change in a magnetic field linking a coil may be integrated by a galvanometer. The time integral of voltage induced in the coil is proportional to the ballistic deflection of the galvanometer which is a measure of the total change in coupling between field and coil. This response defines the emf-time (or flux/linkage) sensitivity of the galvanometer. *See* AMMETER; CURRENT MEASUREMENT; ELECTRICAL MEASUREMENTS; ELECTROMAGNETIC INDUCTION; ELECTROMETER; VOLTAGE MEASUREMENT; VOLTMETER.

Forest K. Harris; Ronald F. Dziuba; Bryan P. Kibble

Bibliography. D. A. Bell, *Electronic Instrumentation and Measurements*, 2d ed., 1997; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; F. K. Harris, *Electrical Measurements*, 1952, reprint 1975; A. D. Helfrick and W. D. Cooper, *Modern Electronic Instrumentation and Measurement Techniques*,

1990; P. Kantrowitz et al., *Electronic Measurements*, 1979; L. M. Thompson, *Electrical Measurements and Calibration*, 2d ed., 1994.

Game theory

The theory of games of strategy can briefly be characterized as the application of mathematical analysis to abstract models of conflict situations. The first such models analyzed by the theory were parlor games such as chess, poker, and bridge. Since then, models arising from the behavioral sciences such as economics, sociology, and political science have been analyzed. Game theory is used in or closely connected to other areas such as linear programming, statistical decisions, management science, operations research, and military planning. In certain areas, the language and concepts of the theory are sometimes used even though the corresponding mathematics is not. *See* LINEAR PROGRAMMING; OPERATIONS RESEARCH.

Games in extensive form. Parlor games are specified by a list of rules. To be playable, a game must have the following properties: (1) There must be a way of starting the game. (2) There must be a well-defined list of legal (that is, permissible) moves for any possible situation that can be reached in the game. (3) At each move, exactly one of the players must be assigned to make the choice, or else the choice is made by a chance device (such as rolling dice or a spinning pointer). (4) After a finite number of moves, the game is terminated; a winner is declared or payments are exchanged among the players, or both. Any other conflict situation that satisfies mathematical axioms similar to these rules can also be considered a game and the analysis of game theory can be applied to it.

The players of a game are called persons, and such a person may actually consist of one or more people (for instance, in bridge the pairs of partners, east-west and north-south, each make up a player in the game; it is considered a two-person game). Thus a person in a game may be considered synonymous with a team. Chance moves occur when hands are dealt from a shuffled pack, dice are rolled, or pointers are spun. One says that all chance moves are allotted to the chance player—a fiction that is useful in abstracting properties of games.

When specified by a list of rules, a game is said to be in extensive form. For mathematical purposes, it is convenient to have games in normalized form, and for that, the idea of a pure strategy is needed.

A pure strategy for a player (not the chance player) is a complete list of choices of legal moves that he or she will make for every possible situation that can occur during the game. This is a much more complete list of decisions than that commonly called a strategy. For instance, in tic-tac-toe a small part of a single pure strategy would be, "If he moves to the center, I will move to the upper right-hand corner; but if he moves to the upper right-hand corner, I will move to the center . . . ; if he moves first to the center, and I to the upper right-hand corner, and he to the lower right-hand corner, then I move to the upper left-hand corner. . . ." The number of pure strategies that are possible in a game can be astronomical even for childish games such as tic-tac-toe, as any reader who tries to enumerate them will discover. Because of the enormous number of pure strategies, the actual applications of game theory even to parlor games have been severely limited by computational difficulties. Simplified versions of the games have been developed for which computations have been completely carried out.

Games in normalized form. After players have chosen pure strategies in a game, they need not physically play the game. Instead they could hand their strategies to a neutral person, or umpire, who could then carry out their instructions and make the moves they would have made. Their intuitively obvious idea leads naturally to the normalized form of the game.

Assume for the moment there are no chance moves in the game, that is, that there are n real players but no chance player. Denote by $s_1, s_2, \ldots, s_n$ specific pure strategies for players $1, 2, \ldots, n$, respectively. Given these, the game must be played in exactly one way and a unique outcome will result. Let $P_i(s_1, s_2, \ldots, s_n)$ be the monetary outcome to player i for this play of the game. (In case only a winner is announced, but no money is exchanged, it can be arbitrarily required that the losers pay one unit each and the winners divide the losers' payments equally, and in this way make a monetary assignment.)

Before the effect of the chance player can be introduced, the important concept of mathematical expectation must be explained. Suppose that $O_1, O_2, \ldots, O_k$ are mutually exclusive monetary outcomes of a chance event, and suppose further that they happen with probabilities $p_1, p_2, \ldots, p_k$ where $p_i > 0$ and $p_1 + p_2 + \cdots + p_k = 1$. Then the mathematical expectation E of the chance event is defined to be the sum $E = p_1O_1 + p_2O_2 + \cdots + p_kO_k$. *See* PROBABILITY.

As an example, suppose that one tosses a coin and is paid \$3 if heads turn up and loses \$1 if tails turn up. If it is assumed that the coin is balanced, each event will happen with probability 1/2 and the expectation is then $3(1/2) - 1(1/2) = 1$. Note that the expectation is not equal to either of the payoffs that can actually be obtained on a single toss of the coin. It can be interpreted as the average amount obtained if the experiment is repeated many times.

If there are chance moves in the game, a set of pure strategies, one for each player, will not determine a unique outcome of the game but merely a set of possible outcomes. These outcomes will be mutually exclusive and have probabilities depending on the chance moves associated with their occurrence. Hence in this case one can let $P_i(s_1, s_2, \ldots, s_n)$ be the expected payoff to player i for each $i = 1, 2, \ldots, n$.

Now the normalized form of a game is defined as the list of all expected payoffs to each player for every possible combination of pure strategies. In the case of two-person games it is most convenient to list these in tables called matrices.

Matching pennies. Two players R and C compare pennies, with R trying to match C. Each player can choose either heads or tails. Each has two pure strategies, "choose heads," H, or "choose tails," T. The game can be summarized as in tabulation (1). Here

	C chooses	
R chooses	H	T
H	1	−1
T	−1	1

(1)

the entries represent the payments by C to R. Thus, if both R and C choose heads, R matches C and receives one penny from C; but if R chooses heads and C tails, then R receives − 1 from C; that is, R pays one penny to C. The other two entries are determined similarly. The solution of this game is given below.

The Mississippi gambler. A passenger on a Mississippi river boat was approached by a flashily dressed stranger (the gambler) who offered to play the following game: "You take a red ace and a black deuce and I'll take a red deuce and a black trey; we will simultaneously show one of the cards; if the colors match I'll pay you and if the colors don't match you pay me; moreover if you play the red ace we will exchange the difference of the numbers on the cards; but if you play the black deuce we will exchange the sum of the numbers. Since if you lose you will pay me either \$2 or \$4 and if I lose I will pay you either \$1 or \$5, the game is obviously fair."

The proposed game is most easily summarized as in tabulation (2). The payoffs listed are those

	Gambler chooses	
Passenger chooses	Red 2	Black 3
Red 1	1	−2
Black 2	−4	5

(2)

received by the passenger. It is shown below that this game is not fair, and that the gambler has a substantial advantage.

Simplified poker. The deck used in simplified poker has only two kinds of cards, in equal numbers, high (H) and low (L). For instance, a bridge deck could be used with red cards high and black cards low. There are two players R and C and each player antes an amount a and is dealt one card which is his hand. It is easy to see that there are only four ways to deal the hands in the game, namely, (H, H), (H, L), (L, H), and (L, L), where R gets the first card of the pair and C the second.

After the deal, R has two alternatives, namely, to see or to raise by adding an amount b to the pot. If R sees, the higher hand wins the pot or equal hands split the pot equally. If R raises, C can fold or can call by adding b to the pot. If C folds, player R wins the pot (without revealing his hand). If C calls, the higher hand wins the pot or equal hands split the pot equally. These are all the rules.

Pure strategies for the players tell them exactly what to do in every situation. A pure strategy for R is to raise if a high card is dealt and to see if a low card is dealt. This can be abbreviated simply to raise-see. It is easy to see that R has four pure strategies, raise-raise, raise-see, see-raise, and see-see. Similarly, C has four pure strategies, fold-fold, fold-call, call-fold, and call-call.

Given a choice of a pure strategy for each player there are exactly four ways the game can proceed, depending on which of the four deals is made. if the mathematical expectation of each is computed, tabulation (3) results.

High		Fold	Fold	Call	Call
	Low	Fold	Call	Fold	Call
See	See	0	0	0	0
See	Raise	$\frac{3a}{4}$	$\frac{2a}{4}$	$\frac{a-b}{4}$	$\frac{-b}{4}$
Raise	See	$\frac{a}{4}$	$\frac{a+b}{4}$	0	$\frac{b}{4}$
Raise	Raise	$\frac{4a}{4}$	$\frac{3a+b}{4}$	$\frac{a-b}{4}$	0

(3)

Since a and b are positive numbers, R is better off choosing the fourth row rather than the second, and the third row rather than the first. Similarly, C, who is minimizing, prefers the third column to either of the first two. Therefore the maxtrix of tabulation (3) can be reduced to that of tabulation (4).

Play			Conservative	Bluffing
	High		Call	Call
		Low	Fold	Call
Conservative	Raise	See	O	$\frac{b}{4}$
Bluffing	Raise	Raise	$\frac{a-b}{4}$	0

(4)

In tabulation (4) the raise-see strategy is labeled as conservative for R and the raise-raise strategy as bluffing because these correspond to the usual notions. Similarly, call-fold is conservative and call-call is bluffing for C.

This game will be solved for various values of a and b in a later section.

Classification of games. A game is called zero-sum if, for every possible n-tuple of pure strategies $s_1, \ldots, s_n$, the sum of the payoffs to all players is zero, that is, Eq. (5) holds. If this sum is not zero for some n-tuple

$$P_1(s_1, \ldots, s_n) + P_2(s_2, \ldots, s_n) + \cdots + P_n(s_1, \ldots, s_n) = 0 \quad (5)$$

of pure strategies, the game is called nonzero-sum. Examples of zero-sum games are matching pennies

and poker, and also a wide class of games, called matrix games, discussed below. If money payments are assigned in bridge according to the points obtained by players, then bridge becomes a nonzero-sum game.

A game in extensive form is said to have perfect information if every player knows and remembers each move of each of her or his opponents as they are made. Examples of games with perfect information are board games such as checkers, chess, or backgammon. A game is said to have perfect recall if each player knows and remembers everything she did (but not necessarily what her opponent did). Poker is a game with perfect recall because each player knows what is in her own hand. However, bridge does not have perfect recall, because the personality of a player (= team) is divided between two actual persons and, for instance, north does not see what is in south's hand (except when south is dummy) even though they are members of the same team.

Games can now be classified according to the number of players they have, whether or not they are zero-sum, and whether or not they have perfect information or perfect recall. The games that will be discussed in more detail are (1) games in extensive form; (2) two-person zero-sum games or matrix games; (3) one-person nonzero-sum games, frequently called statistical games; (4) two-person nonzero-sum games; and (5) games involving three or more persons.

Signaling and behavior strategies. Some aspects of the theory can be studied only when games are specified in extensive form.

By a theorem of E. Zermelo and J. von Neumann, any game having perfect information has a pure strategy solution. Thus all board games can be solved in pure strategies. For chess this means that either (1) one of the players has a pure strategy that guarantees winning regardless of what the opponent does, or (2) each player has a strategy that will force the game to a draw or a win regardless of what the other player does. N. Dalkey has extended the theorem and has characterized the games that have pure strategy solutions.

H. W. Kuhn has shown that every game with perfect recall can be solved by means of behavior strategies. A behavior strategy is a set of probability weights assigned by a player to the alternatives at each of the choice points in the game. (One part of a behavior strategy for poker, for instance, would be to bet or fold with certain probabilities at a given turn in the game.) A consequence of this theorem is that poker can be solved by means of behavior strategies.

G. L. Thompson has shown that every game in extensive form can be solved by a combination of signaling strategies and behavior strategies. He has constructed a simplified model of bridge for which signaling strategies (partial pure strategies) are needed for the solution.

Matrix games. A matrix is a rectangular array of numbers such as those of tabulations (1) to (4). The matrices in tabulations (1) and (2) are called two-by-two matrices because they have two rows and two columns. The matrix in tabulation (3) is four-by-four. In general, a matrix is called m-by-n, if m is the number of its rows and n the number of its columns. Any two-person zero-sum game can be written in normalized form as a matrix game. The matching pennies and simplified poker games discussed above are examples.

Conversely, any matrix with numerical entries can be considered a two-person game between player R who chooses a row of the matrix and player C who chooses a column. After making their choices, C pays R the entry in the common row and column they have chosen (in case the entry is negative, R must pay C the amount).

A matrix then provides an abstract mathematical model for the study of two-person zero-sum games. What is meant by the solution to such a game? Consider the matrix of tabulation (6).

	C chooses	
R chooses	1	2
	−3	5

(6)

Player R (the row player) will receive the positive amounts of the matrix and must pay out the negative ones. This is the maximizing player. Similarly, C (the column player) must pay out the positive amounts and will receive the negative ones, so this is the minimizing player. In tabulation (6) player R would like to get the 5 entry, but to do so would have to choose the second row. And if R did this, player C would choose the first column and R would lose \$3. Hence, R plays the first row and C, who wants to minimize payments, will still choose the first column. The result is that R will receive \$1 from C.

The example of tabulation (6) is that of a strictly determined game, that is, a game in which there is a matrix entry that is simultaneously the minimum entry in its row and the maximum entry in its column. Such an entry is called a saddle-point entry of the matrix. Any matrix game that has a saddle-point entry has a solution, namely, that the players should choose rows or columns that contain such saddle-point entries. The value v of such a game is the saddle-point entry itself. In a strictly determined game the row player can be assured of at least v by playing an optimal strategy. To establish this, note that if this player chooses a row whose minimum entry is maximum, then the least payoff he can get is v, the minimum entry of this row. By a similar argument, the column player can be assured of not losing more than v by playing an optimal strategy.

However, the matrix games of tabulations (1) and (2) do not have saddle-point entries. How can such games be solved? Perhaps the matching pennies example of tabulation (1) suggests an answer. It is quite common for players of this game to randomize their choice of heads or tails by flipping the coin in the air before choosing. In this way they do not know themselves exactly what choice they will make but leave it to a chance device. Hence, their opponent

cannot discover in advance their choice and cannot take advantage of such knowledge. Extensions of this idea give a solution concept to any game, strictly determined or not.

By a mixed strategy for a player in a matrix game is meant a set of probability weights, that is, nonnegative numbers that add up to 1. There are exactly as many probability weights as the player has pure strategies. Thus, in an m-by-n matrix game, a mixed strategy for player R is a set of numbers $p_1, p_2, \ldots, p_m$ with $p_1 \geqq 0$ and $p_1 + p_2 + \cdots + p_m = 1$. Similarly, the set of numbers $q_1, q_2, \ldots, q_n$, with $q_j \geqq 0$ and $q_1 + q_2 + \cdots + q_n = 1$, is a mixed strategy for player C. In using a mixed strategy, the player operates a chance device that chooses among the alternatives with probabilities equal to the corresponding weights. For instance, the device of flipping a coin in a matching pennies game has already been mentioned. In other games, the random choice of a pure strategy may be made by rolling dice, spinning pointers, or choosing random numbers.

A mixed strategy has the desirable property that a player's opponent cannot find out exactly what the player's choice will be, because the player does not know until the chance device has been operated. But even more is true. The use of a mixed strategy by a player introduces a chance move into the game and eliminates all personal moves. Hence, its effect in the game can be accounted for by computing each player's expected value, given a choice of mixed strategy for each. Let p stand for player R's mixed strategy and q for player C's. Let $E(p, q)$ be the expected payoff to R for these choices; then $-E(p, q)$ is C's expected payoff. Thus, if only the payoff $E(p, q)$ is considered, it is evident that R wants to choose p to maximize this amount and C wants to choose q to minimize it.

To analyze the game further, consider two modified games, one in which player C makes a choice of mixed strategy after R has announced a mixed strategy and the other in which R makes a choice after C's announcement. If C chooses after R, C will select a q that minimizes the result; because R knows that C will do this, R selects p initially to maximize the eventual result, that is, in this case p and q are chosen so that expression (7) is obtained. In the other game,

$$\max_p \min_q E(p, q) \tag{7}$$

in which R chooses after C, p and q will be chosen so that expression (8) is obtained. It is not hard to

$$\min_q \max_p E(p, q) \tag{8}$$

show that relation (9) holds for any matrix game. It is

$$\max_p \min_q E(p, q) \leq \min_q \max_p E(p, q) \tag{9}$$

obvious that, if R chooses a strategy according to the max-min rule, R is playing very conservatively and is being protected from the worst that C can do, and similarly for C and the min-max strategy.

The celebrated min-max theorem of von Neumann, proved in 1928, states that even more than relation (9) is true; namely, that for any matrix game there exist mixed strategy vectors $\mathbf{p}^0$ and $\mathbf{q}^0$ so that equality in relation (9) holds. In other words, relation (10) holds. It is this theorem that makes the study

$$\begin{aligned} E(\mathbf{p}^0, \mathbf{q}^0) &= \max_p \min_q E(p, q) \\ &= \min_q \max_p E(p, q) = v \end{aligned} \tag{10}$$

of games a mathematical theory rather than just a descriptive, intuitive theory. The strategies $\mathbf{p}^0$ and $\mathbf{q}^0$ are called optimal mixed strategies for the players R and C, respectively, and $v = E(\mathbf{p}^0, \mathbf{q}^0)$ is called the value of the game. The min-max theorem states, in words, that every matrix game has a value and at least one optimal strategy for each player. Thus every matrix game has a solution.

Many proofs of this important theorem exist. As it stands, it is an existence theorem, that is, it assures the existence of a solution to every matrix game, but it does not tell how to find it. Methods have been worked out for finding on electronic computers solutions to very large games.

Solution of examples. Although it is not possible here to go into the details of finding solutions to games, it is easy to verify solutions when stated.

Matching pennies. The solution to this game is for each player to choose pure strategies with equal probabilities, that is, $\mathbf{p}_1{}^0 = \mathbf{p}_2{}^0 = 1/2$ and $\mathbf{q}_1{}^0 = \mathbf{q}_2{}^0 = 1/2$. In other words the coin-flipping strategy usually employed by players of this game is actually optimal! The value of the game is 0 and it is called a fair game.

Mississippi gambler. Here the optimal strategies are $\mathbf{p}_1{}^0 = 3/4$, $\mathbf{p}_2{}^0 = 1/4$, $\mathbf{q}_1{}^0 = 7/12$, $\mathbf{q}_2{}^0 = 5/12$, and the value of the game is $-1/4$. In other words the passenger can expect to lose an average of a quarter each time the game is played!

Simplified poker. Two special cases for selected values of a and b will be solved.

1. $a = 4$ and $b = 8$. The matrix of tabulation (4) becomes tabulation (11), which is a strictly determined game. Each player should play conservatively, and the value of the game is 0.

	Conservative	Bluffing
Conservative	0	2
Bluffing	−1	0

(11)

2. $a = 8$ and $b = 4$. The matrix of tabulation (4) becomes tabulation (12). Here the value of the game is 1/2 meaning that the game is biased in favor of player R. Optimal strategies are that each player should bluff

	Conservative	Bluffing
Conservative	0	1
Bluffing	−1	0

(12)

half the time and play conservatively half the time. In ordinary parlance, bluffing is regarded as aggressive play, but note that, to play optimally, it is necessary to bluff half the time.

Statistical games. One of the first applications of game theory to another field was its use in the theory of statistical decision functions by the statistician Abraham Wald. In such statistical games, the row player is interpreted to be the statistician and the column player, Nature. The difference between Nature and a real-life player is that there is no known bias on her part. She may sometimes choose to help the statistician and sometimes not. Wald himself chose to apply the theory of games to this problem directly and gave the solution for the statistician as the minmax strategy in the matrix game. This solution has been objected to by others as being too pessimistic, and several other solutions have been proposed.

For example, in the Doctor's Dilemma, the statistician is a doctor who is testing two drugs A and B for the treatment of a certain disease. Drug A is known to be somewhat effective for treatment and never has been associated with aftereffects. Drug B has been spectacularly effective in certain cases but there have been undesirable aftereffects that, however, cannot be definitely attributed to the drug alone. The doctor analyzed the situation to be as in tabulation (13).

Doctor's choice	Nature's choice: A better than B	Nature's choice: B better than A
Uses A rather than B	10	5
Uses A rather than A	−5	20

(13)

She assigned these payoffs by associating a positive monetary outcome to early recovery from the disease and a negative outcome to undersirable effects. How shall the doctor make her choice?

To protect herself from the worst that can happen using pure strategies, she would choose the first row, because it is the row whose minimum entry is maximum. If she permits herself mixed strategies, she would use the optimal strategy that selects the first row with probability 5/6 and the second row with probability 1/6. (For instance, she could roll a die and choose drug B if and only if an ace turns up.)

For many people the use of chance devices in making important decisions such as these seems repugnant. Yet the theory of games shows that, if the goal is to maximize mathematical expectation, and the analysis of the situation is correct, then such a decision rule is optimal.

Nonzero-sum games. By far the most satisfactory part of the theory of games consists of the zero-sum two-person cases, that is, in matrix games. Applications of the theory to such areas as economics, sociology, and political science almost invariably lead to many-person nonzero-sum games. Although no universally accepted theory has been developed to cover these games, many interesting and useful attempts have been made to deal with them.

When more than two persons are involved in a conflict situation, the important feature of the game becomes the coalition structure of the game. A coalition is a group of players who band together and, in effect, act as a new player in the game. There are two extremes to be considered. One is the noncooperative game in which such coalitions are banned by some means. Equilibrium-point solutions, discussed below, provide reasonably satisfactory solutions to such games. The other extreme is that in which all the players join together in a coalition to maximize jointly their total payoff. A game in which coalitions are permitted is called a cooperative game.

In the noncooperative game, each player is solely interested in his own payoff. By an equilibrium point in such a game is meant a set of mixed strategies $s_1, \ldots, s_n$ such that Eq. (14) holds for each $i =$

$$P_i(s_1, \ldots, s_i, \ldots, s_n) = P_i(s_1, \ldots, s_i', \ldots, s_n) \quad (14)$$

$1, \ldots, n$ for all strategies S_i of player i. What this means is that no player can, by changing strategy and assuming that the other strategies stay fixed, improve the payoff. By a theorem of J. Nash, every game has at least one equilibrium-point solution (commonly there are several).

For example, in the gas war, Jones and Smith own gasoline stations on opposite sides of the street. No other stations are nearby. There are only two prices they can charge, high and low, and each day they must decide which price they will charge. They are not permitted to change prices during the course of the day. Suppose their daily gross receipts are as in tabulation (15). In each pair of numbers enclosed in

Jones	Smith: High	Smith: Low
High	(10, 10)	(6, 16)
Low	(16, 6)	(7, 7)

(15)

parenthesis the first number is Jones' receipts and the second number is Smith's receipts.

To interpret these numbers, observe that if both charge the high price, each receives \$10. If both charge the low price (gas war), their income is cut to \$7 for the day. Finally, if one charges the high price and the other low, the low-priced one can not only draw business away from the immediate competitor but also from elsewhere; and the volume of business yields \$16, whereas the opponent gets only \$6.

The unique equilibrium point is (7, 7) as can easily be checked. In other words the gas war prevails. If either player changes his strategy, he will only

cut his own income, so that this solution is stable in the sense that neither can, by himself, profitably change to another strategy.

Cooperative solutions. To most people the equilibrium-point solution to the gas-war example seems irrational, because at the (10, 10) point each player is better off than at the equilibrium point. But the (10, 10) point is not stable, because either player can, by switching to the low price, improve the take from 10 to 16 (assuming that the opponent does not change). It is also true that at the (6, 16) or (16, 6) points the total amount the players take in is greater than the total amount they take in at the (10, 10) point, and so either of these points is preferable. Yet, again, neither of these points is stable in the above sense.

The von Neumann-Morgenstern cooperative solution to the gas-war game is that the players should join together and cooperate to maximize their total intake; they should divide it in such a way that each player gets at least 7, the amount that can be received in the noncooperative solution. In the example, each player would get 7 plus some part of the surplus profit of 8, but the theory does not show further how division of the profit should be made.

Nash and H. Raiffa have given rationale that indicates that the surplus should be divided equally and the net payoff to each player should be $11. The division of the surplus could be carried out by means of a payment of one player to the other or else by having Jones charge high and Smith low on alternate days, and on the other days let the reverse be true. Then their average payment will be $11 each.

M. Shubik and G. L. Thompson have proposed another solution to the game. Suppose that Jones is richer than Smith and suppose that at the (7, 7) point both players actually lose money. Because the game is played over and over again every day, one could imagine Jones making a threat to Smith as follows: "If you do not joint with me in charging the high price, the day after you lower your price I shall lower mine and keep it down until you are bankrupt." Because Jones has more capital than Smith, the threat is effective and, if carried out, would result in the closing of Smith's business. If Smith tried to make a similar threat, he would simply ruin himself, and his threat is therefore suicidal. Here Jones is able to enforce the (10, 10) point and has the dominant position, in what otherwise appears to be a symmetric game, simply by being richer.

Recently, W. F. Lucas has shown that not every n-person game has a solution in the sense originally stated by von Neumann and Morgenstern. Shapley and Shubik have given economic interpretations to this result. They are also applying the theory of games to a series of economic problems.

Simple games. An important class of n-person games for applications to political behavior are the so-called simple games. Each coalition in such a game can be either winning, losing, or blocking. For instance, a winning coalition may be a set of voters who can elect their candidate, or a group of lawmakers who can pass their bill. The players not in a given winning coalition form a losing coalition. Finally, a coalition is blocking if neither it nor the players not in it can enforce their wishes.

It is a common observation that the power of an individual in such a voting situation is not always proportional to the number of votes he controls. This results in part from personality differences, and in part from the voting structure itself. The power index, proposed by L. S. Shapley, assigns to each player an a priori number that indicates the power that the voting structure gives to that player. It is illustrated in the following example.

Green, Jones, and Smith are on a committee. On any issue Green can cast 1 vote, Jones 2 votes, and Smith 3 votes. If it takes 4 votes to carry a measure, what are the winning, losing, and blocking coalitions and what is the relative voting power of each man? The winning coalitions are (Green, Smith), (Jones, Smith), and (Green, Jones, Smith). Losing coalitions are (Green) and (Jones). Blocking coalitions are (Green, Jones) and (Smith).

Suppose that the men enter the committee room singly. A man is pivotal when his entry into the room changes the set of people in the room from a losing or blocking coalition into a winning coalition. There are 6 orders in which they can enter the room. The pivotal man's name is listed in **boldface**. Since smith is pivotal 4 out 6 times he is assigned power $4/6 = 2/3$. Similarly, Green and Jones are pivotal once each and each are assigned power 1/6. Note that the sum of the powers adds up to 1. Note also that Smith controls only one-half of the votes but has two-thirds of the power. Also Jones controls 1/3 of the votes but has only 1/6 of the power:

Green, Jones, **Smith**
Green, **Smith**, Jones
Jones, Green, **Smith**
Jones, **Smith**, Green
Smith, **Green**, Jones
Smith, **Jones**, Green

This model can be used to compare the actual power of a voting member with his index as given above. Thus personality factors can be separated from voting structure factors by means of the model. Generalizations of the model have also been made.

Continuous games. If, in the normalized form of the game, each player is permitted to have a continuous range of pure strategies and the payoff function is permitted to be a function of the two real variables that range over each player's strategies, the result is a continuous game. The model most frequently used is to let player R choose a number x between 0 and 1 and C choose a number y between 0 and 1, and let the payoff function $f(x, y)$ be defined on the unit square. Such games are frequently labeled "games on the square" for obvious reasons. There exist such games without solutions, but with some reasonable assumption as to properties of $f(x, y)$, it can be shown that the min-max theorem holds for these games. The same theorem has been generalized still further by permitting the strategy spaces

of the players to be the more general spaces of analysis.

Game-playing machines. One of the first applications of large electronic computers to numerical problems was in solving large matrix games. Several methods have been devised for finding such solutions. One such method is the simplex method. It has been programmed to solve games with $m + n \geqq 4000$. So-called decomposition methods can extend these methods to certain problems having thousands or even millions of variables and constraints. The principal application of the method is to solve linear programming problems, which can be shown to be equivalent to matrix games.

Computers have also been programmed to play board games such as checkers and chess. Strictly speaking they do not use the theory of games at all at present, but instead use some of the game sense of the people who devised the codes. *See* ARTIFICIAL INTELLIGENCE; DECISION THEORY; INFORMATION THEORY; MATRIX THEORY; STOCHASTIC PROCESS.

Gerald L. Thompson

Bibliography. J. W. Friedman, *Game Theory with Applications to Economics*, 2d ed., 1990; J. McDonald, *Strategy in Poker, Business and War*, 1989, reissued 1996; P. Morris, *Introduction to Game Theory*, 1994; M. J. Osborne and A. Rubinstein, *A Course in Game Theory*, 1994; M. Shubik, *Mathematics of Conflict*, 1983; P. D. Straffin, *Game Theory and Strategy*, 1993; S. Vajda, *Mathematical Games and How to Play Them*, 1992.

Gametogenesis

The production of gametes, either eggs by the female or sperm by the male, through a process involving meiosis.

Animals. In animals, the cells which will ultimately differentiate into eggs and sperm arise from primordial germ cells set aside from the potential somatic cells very early in the formation of the embryo. In some cases, the area of the uncleaved egg which will be included as the cytoplasm of the primordial germ cells is recognizable by its content of organelles, called polar granules. After the embryo gastrulates, prospective germ cells, which are distinguishable from prospective somatic cells by their staining properties, often migrate to sites such as the extraembryonic membranes, where they remain until the gonads are constructed by somatic cells. The primordial germ cells then migrate to and populate the gonads. Evidence indicates that they make no contribution to somatic tissues, and that if they are lost during migration no gametes can be formed.

The final products of gametogenesis are the large, sedentary egg cells, and the smaller, motile sperm cells. Each type of gamete is haploid; that is, it contains half the chromosomal complement and thus half as much deoxyribonucleic acid (DNA) as the somatic cells, which are diploid. Reduction of the DNA content is accomplished by meiosis, which is characterized by one cycle of DNA replication followed by two cycles of cell division. The primordial germ cells are designated either oogonia or spermatogonia when they have arrived in the gonads. In mammals, oogonia divide extensively until birth, while spermatogonial divisions continue throughout the life of the male. At the end of the gonial divisions, cells in both sexes enter first meiotic prophase, during which the chromosomes assume a filamentous appearance and homologous chromosomes pair and may exchange segments. In oogenesis a "lampbrush'' morphology often characterizes the chromosomes of the meiotic prophase. *See* CHROMOSOME; MEIOSIS.

The production of sperm differs from that of oo- cytes in that each primary spermatocyte divides twice to produce four equivalent spermatozoa (**Fig. 1**) which differ only in the content of sex chromosomes (in XY sex determination, characteristic of mammals, two of the sperm contain an X chromosome and two contain a Y). The morphology of sperm is highly specialized, with distinctive organelles forming both the posterior motile apparatus and the anterior acrosome, which assists in penetration of the oocyte at fertilization. *See* SPERMATOGENESIS.

The cytoplasm of the primary oocyte increases greatly during the meiotic prophase and often contains large quantities of yolk accumulated from the blood. Meiotic divisions in the oocyte are often set in motion by sperm entry, and result in the production of one large egg and three polar bodies (**Fig. 2**). The polar bodies play no role, or a very

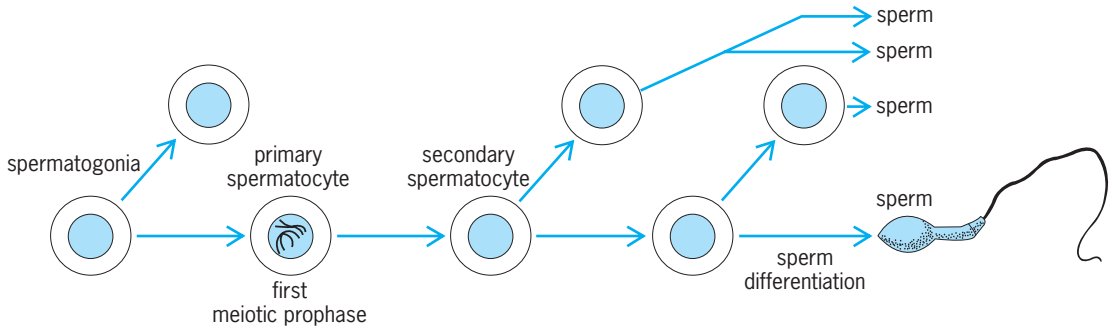

Fig. 1. Generalized diagram of spermatogenesis in animals.

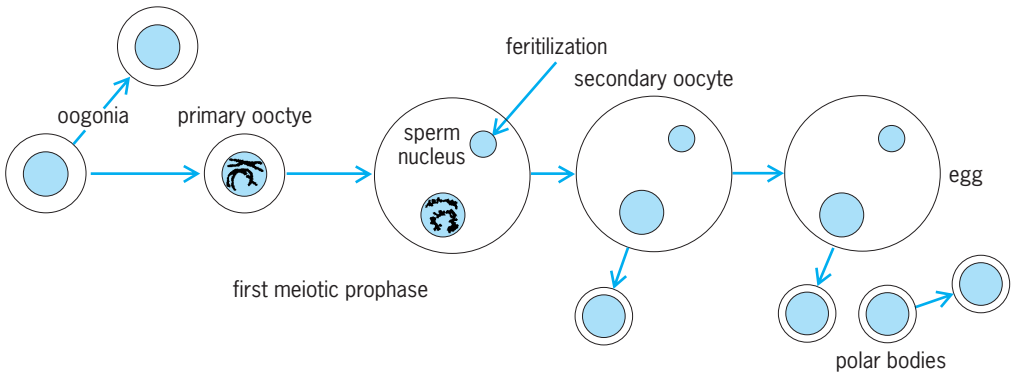

Fig. 2. Generalized diagram of oogenesis in animals.

subordinate one, in the formation of the embryo. *See* OOGENESIS.

After fertilization and the formation of the polar bodies, the haploid sperm and egg nuclei (pronuclei) fuse, thus restoring the normal diploid complement of chromosomes. *See* FERTILIZATION (ANIMAL); REPRODUCTION (ANIMAL). Spencer J. Berry

Flowering plants. Meiosis in flowering plants, or angiosperms, is essentially similar to that in animals. However, the cells produced after meiosis are spores, and these do not develop directly into gametes. Female spores (megaspores) and male spores (microspores) develop into gametophytes, that is, female and male haploid plants that bear within them the egg and sperm, respectively.

Female gametophyte. Development of the female gametophyte takes place within a developing ovule in the ovary of the pistil, the female structure of the flower (**Fig. 3**). One of the cells in the nucellus of the ovule differentiates into a megaspore mother cell, or megasporocyte, which undergoes the two meiotic divisions to form a linear tetrad of haploid (1*N*) megaspores. Three of the megaspores degenerate; the fourth enlarges and the nucleus divides

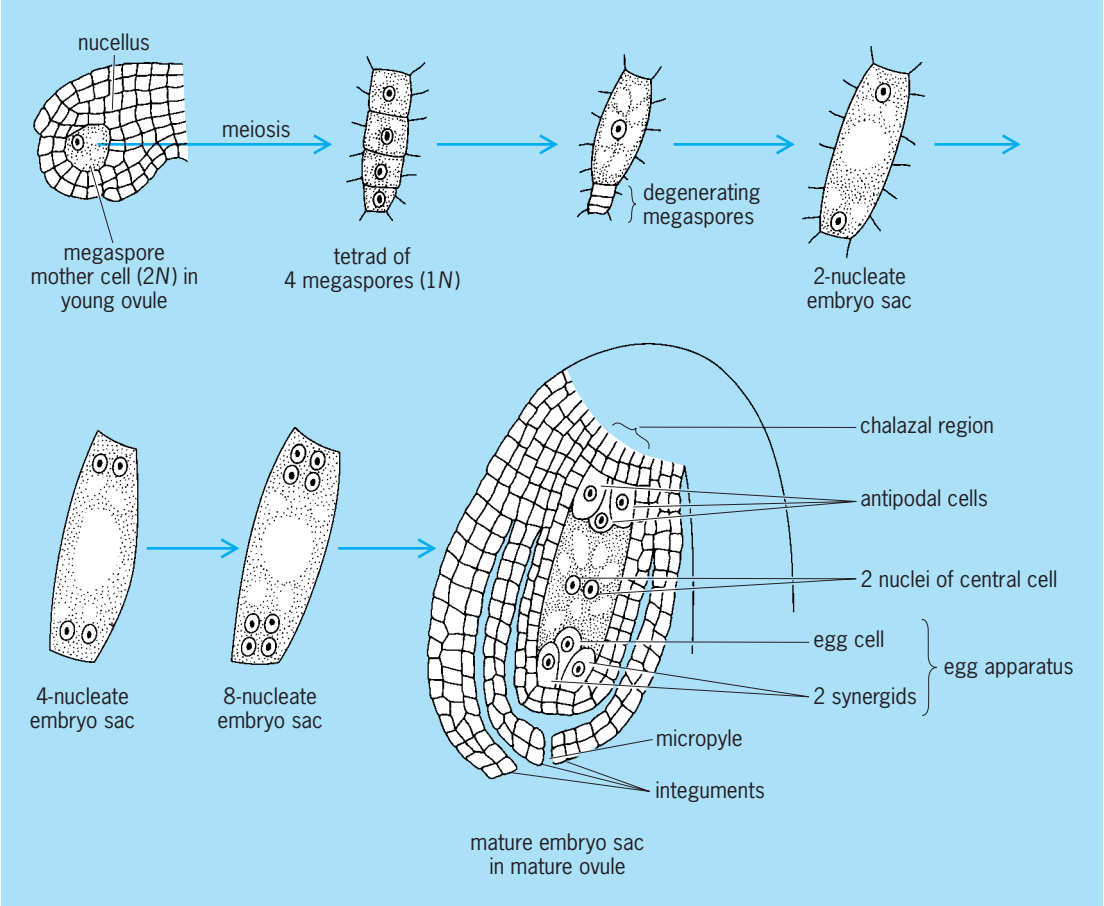

Fig. 3. Stages in the development of the embryo sac and female cells involved in fertilization in angiosperms.

mitotically, forming a two-nucleate embryo sac. Each nucleus divides mitotically, forming a four-nucleate embryo sac, and then each of these nuclei divides, giving rise to an eight-nucleate cell. At this stage, one nucleus from each group of four moves to the center of the cell to form a two-nucleate central cell. The cell is consequently organized into an embryo sac with seven cells through delimitation by cell membranes and cell walls around the remaining six nuclei with their associated cytoplasm. The embryo sac is now the female gametophyte. The three haploid cells at the chalazal end of the embryo sac are the antipodal cells, and these do not play any direct role in fertilization. The three haploid cells at the micropylar end compose the egg apparatus and consist of two synergids and an egg cell. The two nuclei of the large central cell may or may not fuse before fertilization, depending on the plant species. This is the most common type of embryo sac development found in flowering plants. Several other types of development and organization of the embryo sac, however, are known to occur in different plant species.

Male gametophyte. The development of the male gametophyte (**Fig. 4**) occurs within the anther, the male structure of the flower. Each microspore mother cell (microsporocyte) in the anther undergoes meiosis, producing a tetrad of four microspores which enlarge rapidly. The microspore nucleus divides mitotically but unequally to form the young male gametophyte, or pollen grain, which consists of two cells, the vegetative cell and the generative cell. The generative cell is small and lies within the cytoplasm of the vegetative cell. In the pollen of several plant species the generative cell divides in the pollen grain, producing two sperm cells. In most plants, however, the generative cell completes its division much later when the mature pollen grains are released from the anther and are deposited on the stigma of the pistil. After reaching the stigma, each pollen grain germinates, that is, extrudes a pollen tube, and the generative cell completes its division during the growth of the pollen tube but prior to the time it enters the embryo sac. The contents of the pollen tube are discharged into one of the synergids, and double fertilization occurs: One sperm cell fuses with the egg, giving rise to the zygote, while the other fuses with the central cell, giving rise to a 3*N* primary endosperm cell. Flowering plants are unique in nature in their requirement of double fertilization for the formation of a seed which consists of an embryo and the endosperm.

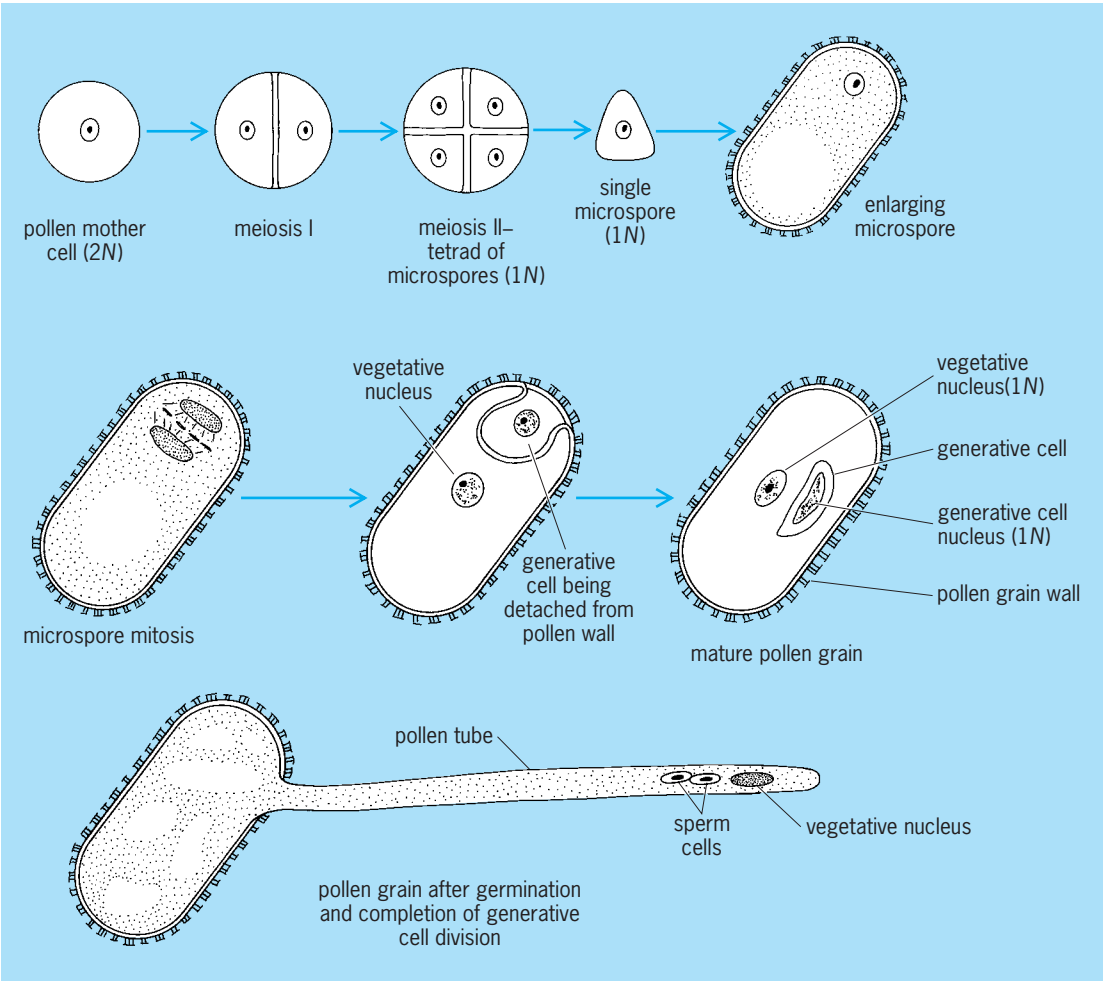

Fig. 4. Stages in the development of the pollen grain and sperm cells.

Other plants. Gametogenesis in angiosperms, the most advanced plants, has been considered in some detail, but there is a wide range in the details of development and structure of gametes among the different groups of plants.

Gymnosperms. In pines, for example, at the time of pollination, the male gametophyte consists of 4 cells, two prothallial cells, a generative cell and a tube cell. During pollen tube growth the generative cell divides to form a sterile cell and a spermatogenous cell. A short time before fertilization the spermatogenous cell divides to form two male gametes which are unequal in size. After meiosis during female gamete development, a single functional haploid megaspore is produced. The female gametophyte developing from the megaspore consists of about two thousand cells. Depending on the species of pine, within each female gametophyte 1 to 6 archegonia, or female gametangia, develop. Each multicellular archegonium produces only a single, nonmotile egg which at the time of fertilization is situated at the base of the archegonial canal. The pollen tube enters between the neck canal cells of the archegonium discharging its contents near the egg cell, enabling fertilization to occur. The larger of the two male gametes fuses with the egg cell. The smaller male gamete does not participate in fertilization.

Lower plants. In all lower vascular plants such as ferns, *Selaginella*, etc., no structure equivalent to the pollen tube is produced; the sperm are motile, flagellated cells that require free water to reach the passive, nonmotile egg. The mechanisms of gametogenesis and fertilization in plants are thus complex, exhibiting great variation between the different groups of plants. *See* DEVELOPMENTAL BIOLOGY; FERTILIZATION (ANIMAL); FLOWER; POLLINATION; REPRODUCTION (PLANT); SEED.

Joseph P. Mascarenhas

Bibliography. L. W. Browder, *Developmental Biology*, 3d ed., 1997; K. Esau, *Anatomy of Seed Plants*, 1977; A. S. Foster and E. M. Gifford, *Comparative Morphology of Vascular Plants*, 1974; P. D. Niewkoop and L. A. Sutasurya, *Primordial Germ Cells in the Chordates*, 1979; S. D. Russell and D. D. Cass, Ultrastructure of the sperms of *Plumbago zeylanica* L.: Cytology and association with the vegetative nucleus, *Protoplasma*, 107:85–107, 1981.

Gamma function

A particular mathematical function that can be used to express many definite integrals. There are, however, no significant applications where the gamma function by itself constitutes the essence of the solution. Instead it occurs usually in connection with other functions, such as Bessel functions and hypergeometric functions.

A special case of the gamma function is the factorial $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$ (for example, $1! = 1$, $2! = 2$, $3! = 6$, $4! = 24$). It is defined only for integral positive values of n. The factorial occurs, for instance, in the expansion

$$\exp z = 1 + \frac{z}{1!} + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots$$

The binomial coefficient (N/n) can be expressed in terms of factorials as $N!/[n!(N-n)!]$. Many occurrences of the factorial are found in combinatorial theory (for instance, $n!$ is the number of permutations of n different elements), in probability theory, and in the applications of this theory to statistical mechanics. For large values of n, the factorial can be easily, although only approximately, computed with Stirling's formula,

$$n! \approx \left(\frac{n}{e}\right)^n (2\pi n)^{1/2}$$

This formula is in error by a factor, which lies, for all $n \geqq 1$, between 1 and $1/(11n)$. For $n = 10$, Stirling's formula gives approximately 3598695.6 . . . instead of the accurate value $10! = 3628800$.

The gamma function can be considered as a certain interpolation of the factorial to nonintegral values of n. It is defined by the definite integral

$$\Gamma(z) = \int_0^\infty u^{z-1} e^{-u} du$$

This definition is applicable to all complex values of z with positive real part. The value of $\Gamma(n+1)$ coincides with $n!$ if $n = 1, 2, 3, \ldots$. The definition of $\Gamma(z)$ can be extended to all complex values of z by the principle of analytic continuation.

Other notations are in use: $z!$ or $\Pi(z)$ for $\Gamma(z+1)$. Sometimes the phrase factorial function is used for $z!$ or $\Pi(z)$ when z is arbitrary.

From its definition follow the fundamental properties of the gamma function:

1. $\Gamma(z+1) = z\Gamma(z)$ (difference equation or functional equation).
2. $\Gamma(z)\Gamma(1-z) = \pi/\sin(\pi z)$, $\Gamma(^1/_2 + z)\Gamma(^1/_2 - z) = \pi/\cos(\pi z)$.
3. $\Gamma(z)\,\Gamma(z + {}^1/_2) = \pi^{1/2} 2^{1-2z}\Gamma(2z)$ (Legendre's duplication formula).
4. $\Gamma(z)$ is an analytic function in the complex z plane except for the points $z = -m$ ($m = 0, 1, 2, \ldots$), where it goes to infinity, as does $[(-1)^m m!\,(z+m)]^{-1}$, if z is close to $-m$.
5. As shown in the **illustration**, $\Gamma(z)$ is real for real values of z, $\Gamma(1) = 0! = 1$, $\Gamma(1/2) = \pi^{1/2}$.

In the angular domain $-\pi < \arg z < \pi$, there exists an expansion of the gamma function, namely,

$$\Gamma(z+1) = z^{z+1/2} e^{-z} (2\pi)^{1/2} \left(1 + \frac{1}{12z} + \frac{1}{288z^2} + \cdots\right)$$

which is not convergent but nevertheless useful for the numerical computation of $\Gamma(z)$ for large z. It gives approximate numerical values of the gamma function for large values of z if one takes only a finite number of terms into account. In particular for real positive values of z, the error is less than the last term taken into account. This expansion generalizes Stirling's formula.

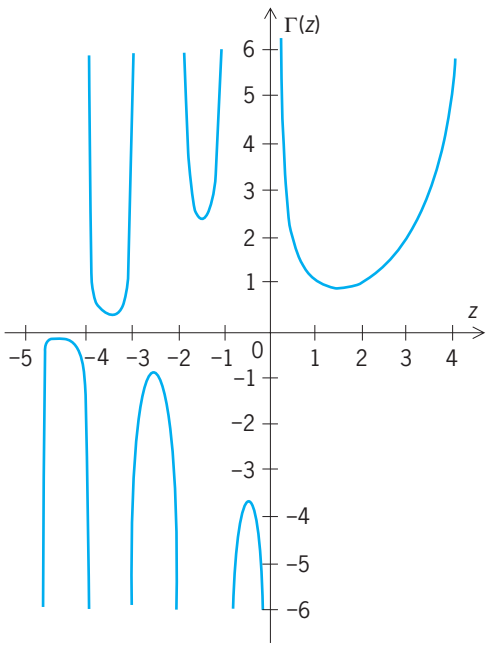

Depiction of gamma function Γ(*z*) for real values of *z*.

Psi function. The logarithmic derivative of the gamma function $d \ln \Gamma(z/dz = \Gamma'(z)/\Gamma(z)$ is denoted by $\psi(z)$. This function has the properties:

$$\psi(z+1) - \psi(z) = \frac{1}{z}$$

$$\psi(1) = -C = -0.5772157\ldots$$

$$\psi(n+1) = -C + 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$$

$$\text{for } n = 1, 2, 3, \ldots$$

$$\psi({}^1/_2) = -C - 2\ln 2$$

$$\psi(z) = -C + \sum_{n=0}^{\infty}\left(\frac{1}{1+n} - \frac{1}{z+n}\right)$$

Beta function. The beta function generalizes the binomial coefficient (N/n) to noninteger values of N and n. The beta function $B(x, y)$ is by definition equal to $\Gamma(x)\Gamma(y)/\Gamma(x+y)$. Therefore one has $(N/n)^{-1} = (N+1)B(n+1, N-n+1)$.

Some of the definite integrals, whose values can be expressed by the gamma function, the psi function, and the beta function are

$$\int_0^{\infty} \exp(-at^n)t^z\,dt = (z+1)^{-1}a^{-(z+1)/n}\Gamma\left(\frac{z+n+1}{n}\right) \qquad (\mathrm{Re}\, a > 0,\ \mathrm{Re}\, z > -1)$$

$$\psi(z) = \int_0^{\infty}\left(\frac{e^{-t}}{t} - \frac{e^{-zt}}{1-e^{-t}}\right)dt = -C + \int_0^1 \frac{1-t^{z-1}}{1-t}\,dt \qquad (\mathrm{Re}\, z > 0)$$

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt = 2\int_0^{\pi/2} \sin^{2x-1}\zeta\,\cos^{2y-1}\zeta\,d\zeta$$

See BESSEL FUNCTIONS. Josef Meixner

Bibliography. L. C. Andrews, *Special Functions of Mathematics for Engineers*, 2d ed., 1998; E. Artin, *The Gamma Function*, 1964; A. Erdelyi (ed.), *Higher Transcendental Functions*, vol. 1, 1953; Z. X. Wang and D. R. Guo, *Special Functions*, 1989; E. T. Whittaker and G. N. Watson, *Course of Modern Analysis*, 4th ed., 1940, reprint 1997.

Gamma-ray astronomy

The study of gamma rays of cosmic origin. This vast spectral domain extends from an energy of 0.01 MeV (wavelength of 10^{-10} m), the adopted boundary between x-ray and gamma-ray photons, to 10^{11} MeV (10^{-23} m), an experimental barrier imposed by the extreme scarcity of ultrahigh-energy photons. *See* GAMMA RAYS.

Low-energy (or soft) gamma-ray astronomy (up to a few megaelectronvolts) deals mainly with processes in dense media, such as plasmas confined close to neutron stars and black holes. It also concerns cosmic sites where monoenergetic photons are released either by deexcitation of atomic nuclei (nuclear lines) or by positron annihilation (the 0.511-MeV line). Gamma-ray astronomy at higher energies relates to emissions induced by relativistic particles throughout the entire interstellar medium, as well as in the vicinity of some neutron stars and in the powerful jets beamed by active galactic nuclei. The penetration power of gamma-ray photons enables exploration of regions that are hidden at other wavelengths, such as the galactic center region, as well as the first stages of the universe, since the cosmos is particularly transparent to gamma rays (with the exception of ultrahigh-energy photons).

Because photons in the gamma-ray regime are completely absorbed by the Earth's atmosphere, gamma-ray detectors are placed on board high-altitude balloons, or, better still, artificial satellites. Ground-based telescopes, making use of the upper-atmospheric layers as a detector, operate successfully in the very high energy gamma-ray band. The first generation of gamma-ray telescopes included the U.S. *SAS-2* satellite launched in 1972 and the European *COS-B* satellite launched in 1975. Missions with much greater capabilities, including the French SIGMA telescope on board the Russian spacecraft *GRANAT* launched in 1989, and the large U.S. *Compton Gamma-Ray Observatory*, launched from the space shuttle in 1991, have obtained major scientific results.

At present, the gamma-ray sky in the band from 0.1 to 10 MeV is investigated by the European Space Agency's *INTEGRAL* (*International Gamma-Ray Astrophysics Laboratory*) satellite, launched in 2002. The *GLAST* (*Gamma-Ray Large-Area Space Telescope*) mission, a collaboration of the National Aeronautics and Space Administration (NASA) and the Stanford Linear Accelerator Center (SLAC), to explore the sky at higher energy, is scheduled for launch in 2007. Low-energy gamma-ray focusing

experiments (*Simbol-X*, *MAX*) are foreseen in the more distant future.

Instrumentation

The information obtained by gamma-ray instruments is limited by various astronomical and instrumental factors.

Astronomical limitations. The astronomical limitations result from the weak intrinsic fluxes of gamma rays and from background problems. A typical gamma-ray observation requires long exposures, up to 1 month in duration, to obtain significant data. The main background originates when cosmic-ray particles interact with the instrument to give secondary gamma rays. Because the observations are limited by background noise, sensitivity improves only in proportion to the square root of the detector area.

Instrumental limitations. In other regions of the electromagnetic spectrum, sensitivity is increased by the straightforward method of gathering large numbers of photons and concentrating them to form an image, by means of arrangements of reflectors or lenses. Such a method has not been applied to gamma-ray telescopes, since reflection or refraction of gamma-ray photons is nearly impossible. Use of grazing-incidence mirrors involving classical reflecting surfaces would seem impracticable beyond the lower limit ascribed to the gamma-ray domain. There remains a possibility of increasing the reflection capability of metallic surfaces, for example, by means of a multilayer coating that would be able to extend the use of reflecting concentrators to energies of about 0.1 MeV. This is the principle of the US/Danish *HEFT* and the French *Simbol-X* projects. Telescopes based on Laue diffraction are also feasible, and their signal-to-noise ratio could be dramatically improved at energies up to a few megaelectronvolts. *See* X-RAY TELESCOPE.

Detection of gamma-ray photons. This involves measurements of the photon energy loss in the mass of the detector. In most cases, it is the kinetic energy imparted to charged particles by the gamma rays that is lost and measured. In photoelectric absorption, the main process at energies of less than 0.2 MeV, and in the Compton scattering process, which dominates in the 0.2–10–MeV band, part or all of the photon energy is converted to kinetic energy of secondary electrons. At higher energies, gamma rays are annihilated in the field of detector atom nuclei, producing electron-positron pairs that carry most of their incident energy. *See* COMPTON EFFECT; ELECTRON-POSITRON PAIR PRODUCTION.

Semiconductors and crystal scintillators are widely used as gamma-ray detectors. In these materials, the secondary electrons lose energy by ionization and produce electron-hole pairs. In semiconductors, such pairs are collected to provide an electric signal; scintillation detectors require a further conversion of charged particles energy to the production of visible photons. Scintillation detectors are generally made of sodium iodide (NaI) or cesium iodide (CsI) crystals, but the energy resolution of these instruments is quite poor, making difficult the detection of spectral lines. Semiconductor germanium detectors offer very good energy resolution but require cooling to less than 100 K (−280°F). Semiconductor cadmium telluride (CdTe) or cadmium zinc telluride (CdZnTe) detectors, affording noticeable spectrometric performances combined with high detection efficiency, can operate at room temperature. Advanced position sensitive detectors, in operation in the soft gamma-ray band in the IBIS telescope on *INTEGRAL* and on NASA's *Swift* satellite (which was launched in 2004 and is devoted to the study of gamma-ray bursts), include arrays of semiconductor CdTe/CdZnTe detectors. *See* GAMMA-RAY DETECTORS; JUNCTION DETECTOR; SCINTILLATION COUNTER.

Low-energy gamma-ray telescopes. Telescopes operating in the soft gamma-ray domain, where focusing techniques have long been considered impracticable, initially used the combination of wide angle (a few degrees) collimation and on-off source chopping to yield source fluxes and location. Because of the poor angular resolution of these techniques, the firm identification of discovered sources with known astronomical objects was often possible only when the emission had a clear time signature.

Coded-aperture telescopes. A telescope for soft gamma-rays can be improved by using the coded-aperture technique to actually image celestial sources. A coded mask is a pattern of tungsten blocks that absorb gamma-ray photons and are arranged so that a given point source at infinity projects on a position-sensitive detector a pattern that is characteristic of the direction of arrival of the photons (**Fig. 1**). The position of the source in the sky is determined by comparing the observed pattern with all possible projection patterns. This technique maintains the angular resolution of a single pinhole camera, while increasing the overall effective area of the instrument. Moreover, the simultaneous measurement of the sky-plus-detector background minimizes systematic effects that result from temporal variations in the background. Both main telescopes

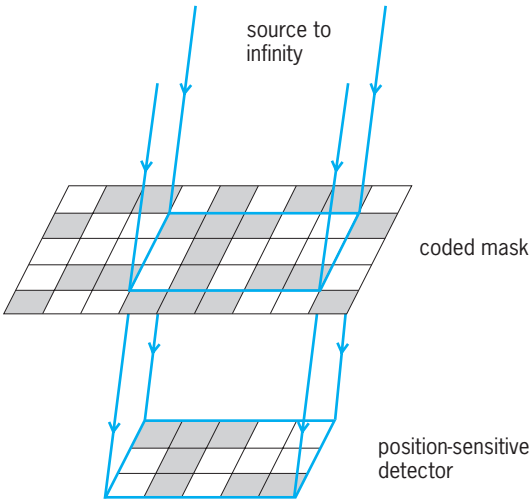

Fig. 1. Principle of a coded-mask aperture telescope whose mask is built from a 3 × 5 array extended on both axes by cyclic permutation.

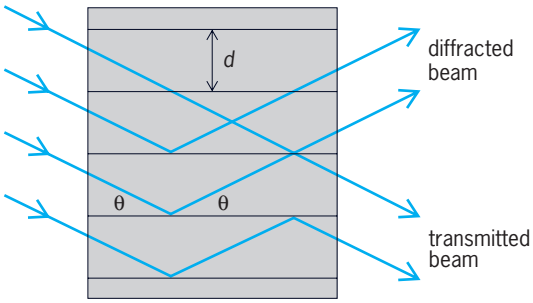

Fig. 2. Principle of Laue diffraction. A beam of very short wavelength radiation λ, entering a plane-parallel crystal at an angle θ, emerges at the same angle θ such that $\sin\theta = n\lambda/2d$, where n is an integer and d is the distance between the crystal reticular planes.

on *INTEGRAL*, IBIS and SPI, use the coded-mask principle.

Compton telescopes. Based on the detection of a Compton-scattering interaction of a gamma ray in a first detector, followed by the detection of the Compton-scattered photon in a second detector, such telescopes allow for the simultaneous determination of photon direction and energy. At megaelectronvolt energies, a set of two detectors measuring the Compton kinetics can be used in conjunction with coded-mask telescopes, in order to realize the good position resolution of coded-mask telescopes together with the low background of Compton detectors.

Concentrators. A plane-parallel crystal, intercepting a beam of very short wavelength radiation, acts as a three-dimensional diffraction array which deflects radiation by a process known as Laue diffraction (**Fig. 2**). The most advanced project of gamma-ray concentrators based on Laue diffraction, *MAX*, employs a series of concentric rings, over which are mounted several hundred germanium crystals oriented so that each deflects the incident radiation toward the same focal point. *See* X-RAY DIFFRACTION.

High-energy gamma-ray telescopes. High-energy gamma-ray observations are performed with devices derived from particle-physics detectors. High-energy gamma-ray photons interact almost exclusively via electron-positron pair production; thus that process forms the basis for the measurement principle by providing a unique signature for gamma rays and allowing a determination of the incident photon directions via the reconstruction of the electron-positron pair trajectories. After passing undetected through an anticoincidence shield sensitive to charged particles, a photon is converted to an electron-positron pair in one of the conversion foils (**Fig. 3**). Trajectories of the resulting electron and positron are measured by particle-tracking detectors. The energy deposited in an underlying calorimeter is used with the trajectory data to determine the arrival direction and total energy of the gamma rays.

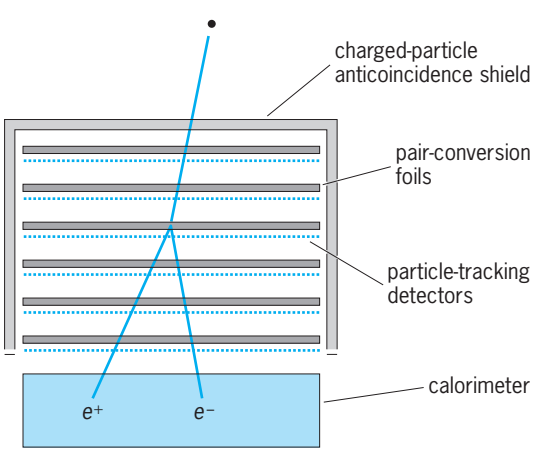

Fig. 3. Principle of a pair-conversion telescope.

Following EGRET, which operated on the *Compton/Gamma-Ray Observatory* in the 1990s, the next high-energy telescope based on these principles is the Italian pathfinder mission *AGILE* (*Astro-rivelatore Gamma a Immagini L'Eggero*), to be launched in 2005, followed by the *GLAST* mission, in 2007.

Very high energy gamma-ray telescopes. As the gamma-ray energy approaches 10^5 MeV, the intensities of celestial gamma rays become too low for them to be seen with space telescopes within a reasonable observation time. However, above a few thousand megaelectronvolts, a gamma-ray photon induces in the upper atmosphere a shower of secondary relativistic particles whose propagation through the air produces a narrow beam of Cerenkov visible light that can be detected on the ground by a large parabolic mirror. Detailed studies of the Cerenkov light beam enable the determination of the arrival direction of the generating gamma ray to within 0.1° while discriminating gamma-ray-induced events from much more numerous events induced by interactions of very high energy cosmic-ray protons and nuclei. The HESS (High-Energy Stereoscopic System) telescope, for instance (**Fig. 4**), which has operated in Namibia since 2003, is a stereoscopic telescope system, where multiple telescopes view the same air shower. The initial four HESS telescopes (Phase I) are arranged in the form of a square with a side length of 120 m (400 ft) to provide multiple stereoscopic views of air showers. *See* CERENKOV RADIATION; COSMIC RAYS.

Stellar Sources

Other than the Sun, which is very near, the only stellar sources of gamma rays that are luminous enough to be observed from Earth relate to massive stars in their final stages of evolution.

Sun. Gamma-ray observations of the Sun have the objective of investigating high-energy processes that take place in the Sun's atmosphere and their relationship to the basic problems of solar activity. Solar gamma-ray emission can be detected only during solar flares. Gamma-ray continuum emissions are produced by processes, such as bremsstrahlung, involving particles accelerated during the flare. Series of sporadic gamma-ray lines are also detected. Other than the line at 2.223 MeV from deuterium formation and the line at 0.511 MeV from electron-positron annihilation, all lines are attributable to interactions

Fig. 4. HESS (High-Energy Stereoscopic System) very high energy gamma-ray telescope in Namibia.

involving energetic protons. The 2.223-MeV line implies the presence of a significant thermal neutron flux. *See* BREMSSTRAHLUNG; SUN; THERMAL NEUTRONS.

Supernovae. Major contributions to the theoretical understanding of explosive nucleosynthesis have come from the data obtained on supernova SN 1987A, which appeared on February 24, 1987, in the Large Magellanic Cloud, a nearby galaxy. Supernova explosions are considered particularly efficient in the synthesis of various radioactive nuclei. The nuclear gamma rays produced by the decay of cobalt-56 are thought to be the main power source of the supernova in the first months following the explosion. Beginning in August 1987, as the envelope of material ejected by the supernova explosion expanded and became more tenuous, the gamma-ray lines from the decay of cobalt-56 were observed (slightly preceded by the soft gamma rays of the corresponding continuum spectrum of Compton-scattered photons). In 1988, balloon-borne germanium detectors carried out fine spectroscopic studies of these lines. The evolution of the lines with time gives unprecedented constraints on explosive nucleosynthesis models and on the mixing of freshly synthesized material deep inside the supernova envelope. Gamma-ray line emissions from long-lived isotopes synthesized in supernova explosions (such as titanium-44) have also been observed in the three-century-old Cassiopeia A supernova remnant. However, a deep survey made in the galactic center region by *INTEGRAL* at the titanium-44 line energy did not reveal any sources, although ten had been expected. This discrepancy remains an unresolved issue in understanding the nucleosynthesis of titanium-44 in supernovae. *See* NUCLEOSYNTHESIS; SUPERNOVA.

Neutron stars and stellar black holes. Virtually all types of compact stars have significant high-energy emission luminosity. Many gamma-ray sources have been firmly identified with neutron stars (isolated and in binary systems) and stellar black holes.

Spin-powered neutron stars. Attested by the discovery of radio pulsars in 1967, the capability of neutron stars to accelerate relativistic electrons was confirmed a few years later by the discovery of the gamma-ray emission of the Crab and Vela pulsars. The energy spectrum of the Crab pulsar, as observed up to 10^4 MeV, suggests that the gamma radiation results from the synchrotron emission of relativistic electrons in the intense magnetic fields (of the order of 10^8 tesla or 10^{12} gauss) which prevail in the close vicinity of newly formed neutron stars. It is generally agreed that the ultimate source of the radiated energy is the rotational energy of the neutron star. Electrons are accelerated to very high energies by the huge electric fields induced by the pulsar rotation, as implied by the gamma-ray emission detected above 5×10^5 MeV in the direction of the Crab Nebula by atmospheric Cerenkov devices. It has been suggested that such a very high energy gamma-ray emission results from inverse Compton processes where ultrarelativistic electrons accelerated by the Crab pulsar energize soft photons within the Crab nebula. *See* CRAB NEBULA; SYNCHROTRON RADIATION.

Nevertheless, much of the basic workings of pulsars remains unknown: How and where does the particle acceleration take place? What is the shape of the particle beam, and how is its energy converted to photons? High-energy gamma-ray studies provide an essential tool for answering these questions. The stage has been set by the seven pulsars detected by the *Compton Gamma-Ray Observatory*, which are being studied in more detail by *INTEGRAL* (**Fig. 5**). These include the brightest persistent sources in the 10^2–10^4-MeV band: the Crab and Vela pulsars and Geminga, which was also found in 1992 to be a gamma-ray pulsar. Geminga has been identified with a very weak visible star whose measured parallax indicates that it is one of the nearest known neutron stars, 450 light-years (4.3×10^{15} km or

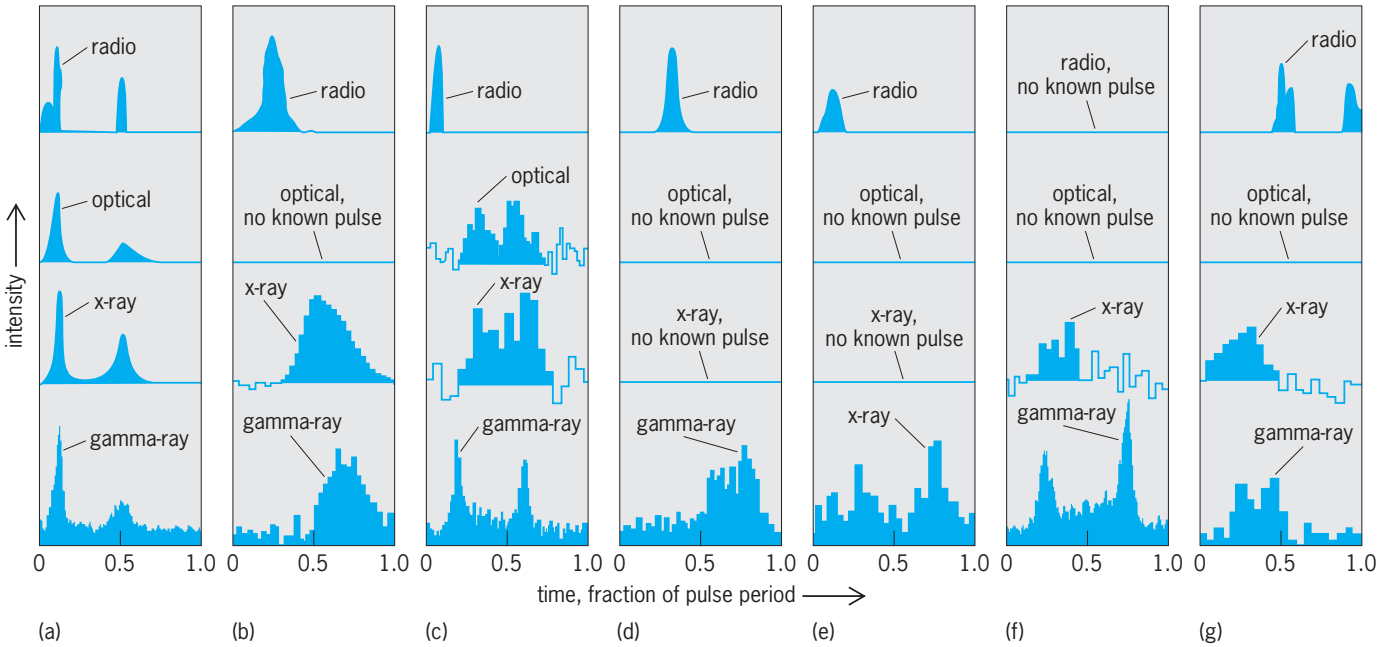

Fig. 5. Light curves of seven isolated pulsars detected in the gamma-ray range by the telescope EGRET onboard the *Compton Gamma-Ray Observatory*. Light curves are shown in the radio band, the optical band, x-rays, and the gamma-ray energy domain. (*a*) Crab pulsar, period ∼33 ms. (*b*) PSR 1509-58, period ∼150 ms. (*c*) Vela pulsar, period ∼89 ms. (*d*) PSR 1706-44, period ∼102 ms. (*e*) PSR B1951+32, period ∼39 ms. (*f*) Geminga, period ∼237 ms. (*g*) PSR 1055-52, period ∼197 ms. (*Courtesy D. Thomson, NASA/GSFC*)

2.6×10^{15} mi) from the Earth. The Crab and Vela pulsars have also been detected at very high energy, 10^5–10^6 MeV, by several Cerenkov experiments (including Whipple, CAT, and CANGAROO). *See* GEMINGA; NEUTRON STAR; PULSAR.

Accreting compact stars. The compact remnant of an exploded star belonging to a close binary system exerts such an attraction on the external layers of its companion star that matter transfers may occur. This transfer results in a thick plasma disk, the accretion disk, which encircles the collapsed star, itself enshrouded in a dim corona. Heated by viscous stresses, the inner part of the disk releases intense thermal radiation in the x-ray band. Gamma rays could be produced by the Compton upscatter process, in which hot (fast moving) electrons near the compact object scatter and kick the abundant x-ray photons to higher energies. The electrons, which gain energy while falling toward the compact object, lose energy during this scattering process, so that an equilibrium temperature is obtained close to the compact object surface. This temperature is particularly high if the collapsed star is a black hole, and hence lacks a solid surface that could reemit a lot of soft x-ray photons, which are particularly efficient in cooling the electron population.

Almost all known accreting black-hole systems produce strong and variable gamma-ray emission. In systems with low-mass (less than one solar mass) companion stars, huge outbursts called x-ray novae are observed, such as Nova Muscae 1991, which for one week was the brightest source in the soft gamma-ray sky. On January 20–21, 1991, it underwent an eruptive episode that appeared in the form of a spectral feature centered around 0.5 MeV (**Fig. 6**), probably related to electron-positron annihilation processes.

Several accreting black holes have been observed in the central region of the Milky Way Galaxy (**Fig. 7**), while the galactic nucleus itself has been detected at a surprisingly weak level in the soft gamma-ray regime by *INTEGRAL*. This weakness may be due to the different type of accretion occurring close to the supermassive black hole sleeping in the galactic center. Accurate radio observation of the source 1E 1740.7-2942, located only 50 arc-min from the nucleus, and of a few similar black-hole candidates, has revealed that accreting stellar black holes could

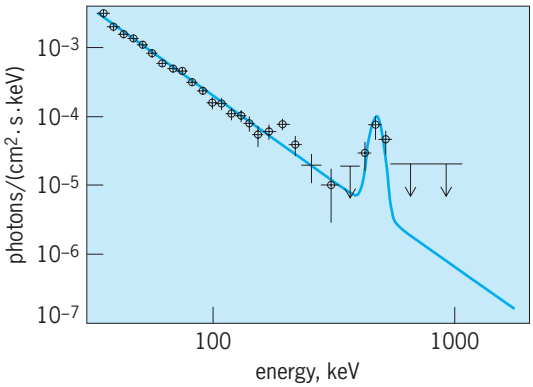

Fig. 6. Energy spectrum of Nova Muscae 1991 as obtained by the SIGMA telescope on board the *GRANAT* satellite on January 20–21, 1991 (crosses). The most straightforward model (solid line) consists of superposing a gaussian-profile line centered at 0.480 MeV onto a power-law spectrum of spectral index −2.4. (*After A. Goldwurm et al., SIGMA observations of the x-ray nova in Musca, Astron. Astrophys. Suppl., 97:293–297, 1993*)

Fig. 7. Image in the 0.020–0.060-MeV band of a $90^\circ \times 30^\circ$ region around the galactic center as seen by the IBIS telescope on board the *INTEGRAL* satellite. The map is in galactic coordinates; the galactic plane lies along the horizontal line at galactic latitude $= 0^\circ$. More than 90 sources can be seen on this map, some of them being black holes. (*F. Lebrun et al., Compact sources as the origin of the soft γ-ray emission of the Milky Way, Nature, 428:293–296, 2004*)

generate powerful bipolar jets of relativistic particles. Featuring two basic properties of quasars (black holes and radio jets), such sources are sometimes referred to as microquasars. *See* ASTROPHYSICS, HIGH-ENERGY; BLACK HOLE; MILKY WAY GALAXY.

Other Galactic Sources

A large fraction of the cosmic gamma-ray photons originates in interstellar sites, occasionally so widely spread out that they are well resolved by gamma-ray devices, in spite of their limited angular resolution.

Cosmic-ray-induced interstellar emission. This emission results mostly from the interaction of cosmic rays (electrons, and protons) with the interstellar gas.

Relativistic electrons (of energies above a few megaelectronvolts) radiate gamma rays via the bremsstrahlung process when they are deviated by interstellar nuclei; fast protons (of energies above a few hundred megaelectronvolts) indirectly generate gamma rays when they collide with protons at rest, through the creation of π° particles (neutral pions), which decay almost instantaneously into two 68-MeV gamma rays. *See* ELEMENTARY PARTICLES.

The high-energy gamma-ray sky is dominated by radiation from the galactic plane (**Fig. 8**). Its spatial distribution and intensity can be reliably modeled from knowledge of the interstellar gas distribution. An important advance in understanding cosmic rays was the detection by the *Compton Gamma-Ray Observatory* of the broad spectral feature at 68 MeV, due to π° decay. This feature is the signature of the presence of accelerated particles throughout the galactic disk. *See* INTERSTELLAR MATTER.

Localized sources. Pointlike sources have been systematically searched for on top of the large-scale emission. Above a few megaelectronvolts, this search is particularly difficult as the telescopes operating in this energy range have a poor localization power. Indeed, the third EGRET catalog of 271 sources at energies above 100 MeV includes 170 specimens that have no established counterparts at other wavelengths. On the other hand, a number of supernova remnants are reported to be coincident with EGRET sources. These high-energy gamma-ray emissions may result from the interaction of particles accelerated in the shocks produced by collisions of the supernova remnant with the surrounding dense interstellar clouds. Although these results are somewhat ambiguous, recent observations of the remnant of SN 1006 (a supernova seen in the year 1006) performed in the very high energy band by ground-based Cerenkov telescopes demonstrate that at least electrons are accelerated to very high energies in the shocks of supernova remnants.

A large fraction of the sources observed at medium galactic latitude may be related to the local interstellar medium, and more specifically to the giant cloud complexes of Gould's Belt. In addition to the pointlike sources, two positive identifications of larger-scale features are a concentration of clouds near the star ρ Ophiuchi and the Orion cloud complex. In the

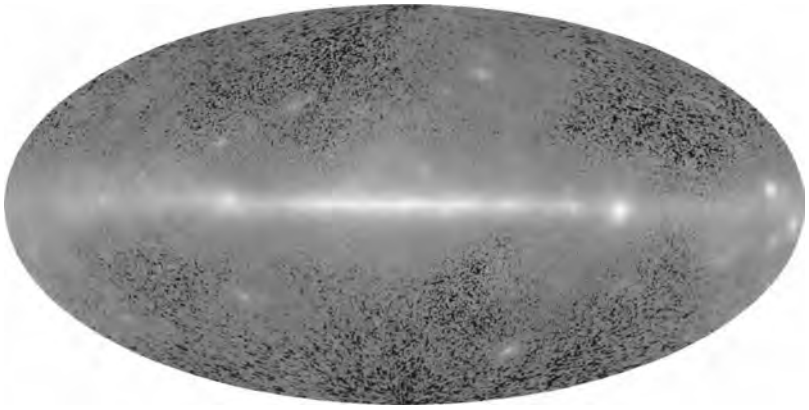

Fig. 8. Intensity map of the gamma-ray sky at energies greater than 100 MeV as measured by the EGRET experiment on the *Compton Gamma-Ray Observatory*. The map is in galactic coordinates, with the galactic plane along the abscissa. The broad band along the galactic plane is interstellar emission from cosmic-ray interactions in the Milky Way. The brightest sources close to the plane are gamma-ray pulsars. The bright sources away from the plane are active galactic nuclei or yet unidentified. (*Courtesy of the EGRET team*)

constellation Aquila, the detection of high-energy gamma-ray emission led to the discovery of interstellar clouds, later confirmed by radio observations of the carbon monoxide molecule. *See* ORION NEBULA.

Interstellar line emissions. Spectroscopic observations of the whole galactic center region, especially by the spectrometer on board *INTEGRAL*, SPI (**Fig. 9**), have demonstrated the presence of a large-scale component of 0.511-MeV radiation due to the annihilation of positrons in the interstellar medium. Such interstellar positrons were thought to be produced through the β^+ decay of radioactive nuclides produced by novae, red giants, Wolf-Rayet stars, and supernovae. However, the SPI observations imply large numbers of produced positrons, which seem inconsistent with the standard models. Alternative explanations, involving more exotic objects, such as hypernovae or dark-matter particles, may be implied by the SPI measurements. *See* POSITRON.

The discovery and confirmation of 1.809-MeV-line emission from the galactic plane (**Fig. 10**) was the first direct detection of a cosmic radioactive nucleus (aluminum-26) from the characteristic signature of its decay. As the half-life of aluminum-26 (716,000 years) is very short compared to the time scale of galactic chemical evolution, the 1.809-MeV-line observations clearly demonstrate that nucleosynthesis is currently taking place in the Milky Way Galaxy. The celestial distribution of the 1.809-MeV-line emission suggests that advanced stages of massive stars, such as Wolf-Rayet stars and supernovae, are the major sources of aluminium-26. *See* RADIOACTIVITY; WOLF-RAYET STAR.

Extragalactic Gamma-Ray Sources

With the exception of the Large Magellanic Cloud, which harbors both SN 1987A and a diffusely distributed source of high-energy gamma rays, similar to that related to the whole Milky Way Galaxy, and with the exception of the gamma-ray bursts, all localized extragalactic sources of gamma rays are active galactic nuclei, among the most energetic and distant objects in the universe. *See* GAMMA-RAY BURSTS.

Active galactic nuclei. These include radio sources such as Seyfert galaxies and quasars, all with different properties depending on the observing wavelength. The ultimate source of active galactic nuclei activity is believed to be massive (10^6–10^9-solar-mass) black holes, accreting 10–100 solar masses per year to account for their overall luminosity. An accretion disk with a collimated perpendicular jet is the favored model for explaining the luminous, broadband radiation emitted from the central engines of active galactic nuclei. While ultraviolet, x-ray, and soft gamma-ray photons are emitted somewhat isotropically from the accretion disk (several Seyfert galaxies have been detected in the soft gamma-ray band), photons at higher energies are produced by radiation mechanisms that focus gamma rays in the jet direction.

Only BL Lacertae objects as well as highly variable and strongly polarized quasars (collectively referred as blazars) have been detected at high energies, because blazar jets are favorably aligned along the line of sight. A total of 66 high-confidence identifications of blazars are indexed in the third EGRET catalog, including very distant objects (up to redshift $z = 2.3$). In addition, 27 lower-confidence potential blazar identifications are noted. A few nearby blazars have been detected at very high energies by ground-based Cerenkov telescopes. Their spectra extending to energies above 10^6 MeV, variability time scales measuring less than a day, and apparent luminosities often exceeding 10^{15} solar luminosity are evidence that extremely energetic sites of efficient relativistic particle acceleration must be involved. High-energy gamma rays are thought to result from Compton scattering of soft photons by energetic nonthermal electrons in the jet. The source of the soft photons is controversial: synchrotron self-Compton models, employing

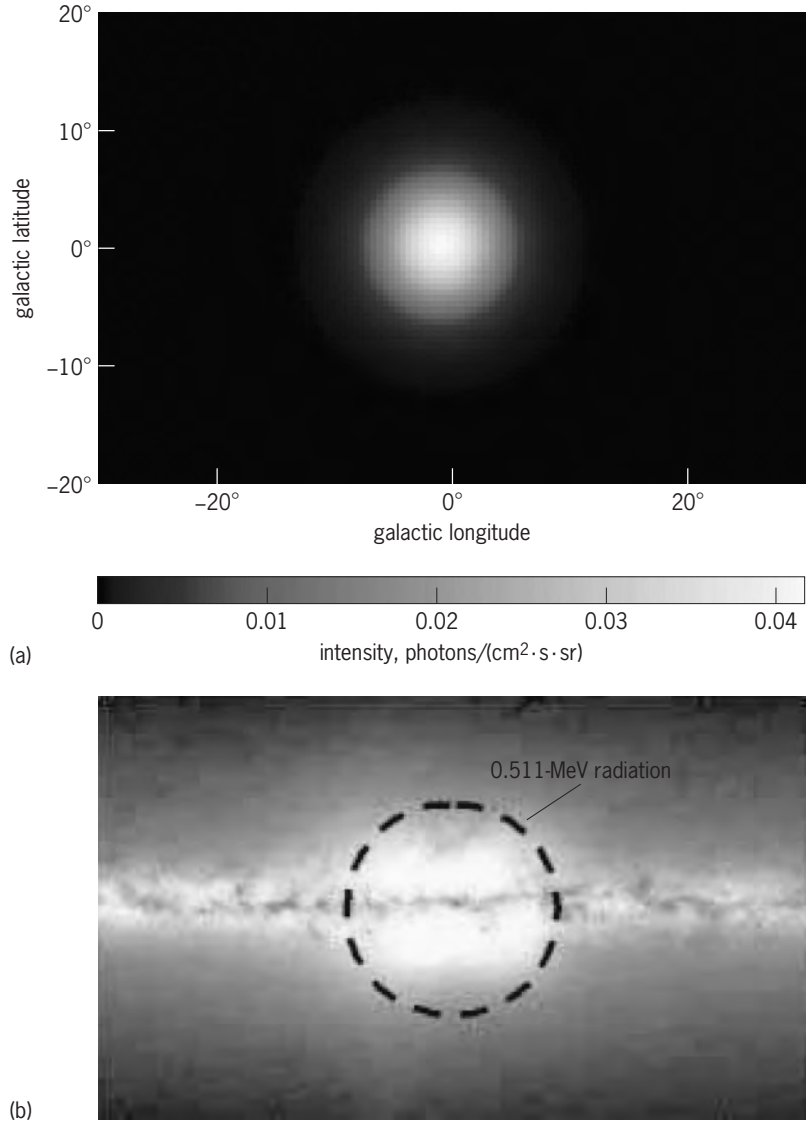

Fig. 9. Gamma-ray line emission at 0.511 MeV from electron-positron annihilation from the direction of the galactic center. (*a*) Intensity map of 0.511-MeV emission as measured by the SPI telescope on the *INTEGRAL* satellite. The map is in galactic coordinates, with the galactic plane along the abscissa. (*From G. Weidenspointner et al., SPI observations of positron annihilation radiation from the 4th galactic quadrant: Sky distribution, in Proceedings of the 5th INTEGRAL Workshop on The INTEGRAL Universe, Munich, pp. 133–139, ESA SP-552, 2004*) (*b*) Distribution of 0.511-MeV radiation superposed on an infrared view of the Milky Way, showing that the distribution coincides with the galactic bulge. (*CESR/Toulouse*)

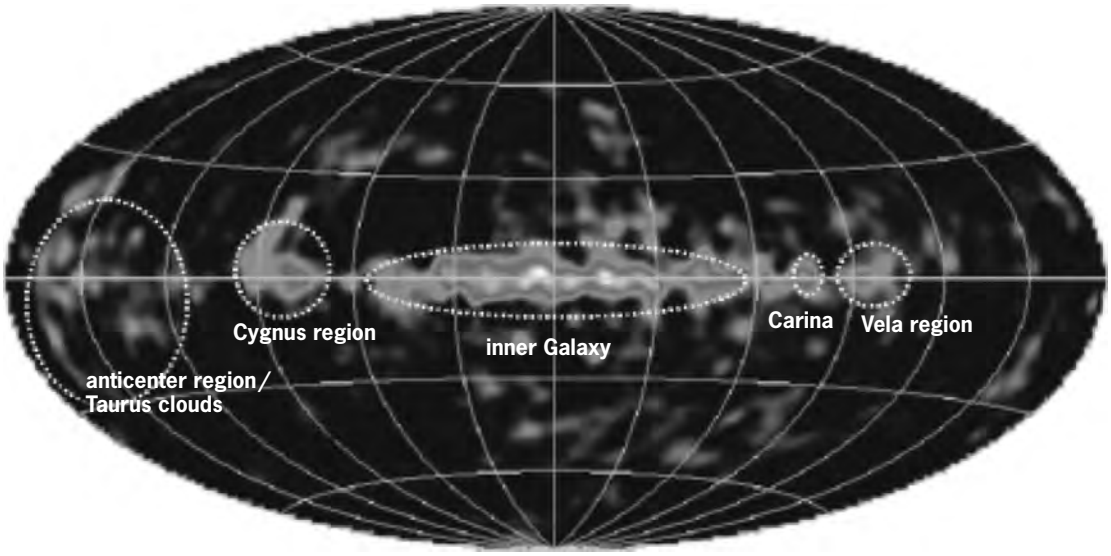

Fig. 10. Intensity map of the gamma-ray line emission at 1.809 MeV from aluminum-26 decay as measured by the COMPTEL experiment on the *Compton Gamma-Ray Observatory.* The map is in galactic coordinates, with the galactic plane along the abscissa. (*After U. Oberlack, Ph.D. thesis, Technische Universität München, 1997*)

the electrons' own synchrotron radiation, are one possibility. *See* GALAXY, EXTERNAL; QUASAR.

Cosmic background radiation. An apparently isotropic, presumably extragalactic, radiation has been measured in the gamma-ray regime up to about 150 MeV. The whole cosmic gamma-ray background could be explained in term of composite light from a large number of faint active galactic nuclei such as Seyfert galaxies at low energy and blazars at high energy. Other more exotic possibilities are relics from some yet-unknown high-energy process in the early universe.

Philippe Laurent

Bibliography. N. Gehrels and J. Paul, The new gamma-ray astronomy, *Physics Today*, pp. 26–32, February 1998; R. C. Hartman et al., The Third EGRET Catalog of High-Energy Gamma-Ray Sources, *Astrophys. J. Suppl.*, 123:79–202, 1999; Special Letters Issue: First science with *INTEGRAL, Astron. Astrophys.*, 411:L1–L460, 2003.

Gamma-ray bursts

Short cosmic blasts of very high energy electromagnetic radiation (gamma radiation) that are recorded at an average of about once per day by detectors placed above the Earth's atmosphere. At distances of billions of light-years, the energy emitted in a gamma-ray burst (GRB) is more than a billion billion (10^{18}) times the energy emitted each second by the Sun. The mechanism for the origin of gamma-ray bursts remains a focus of current research. Gamma-ray bursts longer than 2 seconds appear to be associated with the supernova explosions of massive stars in distant galaxies. Recent evidence suggests that shorter gamma-ray bursts are the result of merging neutron stars or black holes. There are indications that both types of gamma-ray bursts lead to the birth of black holes. *See* GAMMA RAYS.

History of observations. In 1963, the United States and other nations signed the Nuclear Test Ban Treaty, which forbade the testing of nuclear weapons in or above the atmosphere. In order to ensure compliance with this treaty, the United States launched the *Vela* series of Earth-orbiting satellites to look for violations which would produce blasts of gamma radiation. No violations of the Nuclear Test Ban Treaty were ever reported. However, in 1973 Ray W. Klebesadel, Ian B. Strong, and Roy A. Olson announced the discovery of the first gamma-ray bursts and provided evidence of their cosmic origin based on the data from the *Vela* satellites.

During the next 20 years, an international array of satellites slowly built up a catalog of information about the brightness and energies of the gamma rays in scores of gamma-ray bursts. Some of these satellites were located in the solar system at great distances from Earth, while others were in near Earth orbits. The combined data from this Interplanetary Network (IPN) allowed the positions of some of the bursts to be located with greater precision. However, until the notable event that occurred on March 5, 1979, the identification of the gamma-ray bursts with familiar objects that radiate at, for example, visible wavelengths of light, proved elusive.

March 1979 event. On March 5, 1979, at least nine satellites in the Interplanetary Network were swamped with a blast of gamma radiation, 100 times stronger than any burst detected previously. Because of the large number of detections converging from different angles, the Interplanetary Network was able to precisely locate the burst. Surprisingly, it emanated from a supernova remnant in our neighboring galaxy, the Large Magellanic Cloud, at a distance of about 150,000 light-years. A sinusoidal signal was seen in the gamma-ray light as the intensity decayed, followed hours later by another weaker burst, with many more bursts occurring over the

next month. Studies of gamma-ray bursts for the next decade assumed that the March 5, 1979, event was a prototype for all gamma-ray bursts, and many scientists searched for sinusoidal oscillations and repeated weaker bursts at the positions of seemingly similar events. However, it is now known that the March 5, 1979, event was not a "classic gamma-ray burst" but the defining member of a new class of cosmic explosions—the soft Gamma repeaters (SGRs). These seemingly similar detonations result from starquakes on the surfaces of rotating neutron stars which possess extremely strong magnetic fields, with strengths up to 10^{11} tesla. *See* MAGELLANIC CLOUDS; NEUTRON STAR.

CGRO observations. The *Compton Gamma-Ray Observatory* (*CGRO*), launched into Earth orbit by the National Aeronautics and Space Administration (NASA) on April 5, 1991, had a complement of four gamma-ray-detecting instruments, including one designed specifically to study gamma-ray bursts: the Burst And Transient Source Experiment (BATSE). During its 9-year lifetime, BATSE detected and located over 2700 gamma-ray bursts; **Figure 1** shows their positions and fluxes on a map of the sky. There is no obvious clustering of the bursts near the plane of the Milky Way Galaxy (horizontally across the center of the figure) or anywhere else. This result surprised almost all astronomers, as it had been widely believed that the gamma-ray bursts must arise from within the Milky Way Galaxy.

Two possible interpretations were advanced to explain the lack of any pattern in the gamma-ray burst skymap: either the bursts were very close to the solar system, or they were very far away, at cosmic distances. In this latter case, however, the energies required to power the bursts were so large that scientists had trouble explaining the cause of the bursts. The evidence that gamma-ray bursts were truly located across the universe did not appear until 1997.

Discovery of afterglows and distances. On February 28, 1997, a gamma-ray burst was observed by detectors on board *BeppoSAX*, an Italian-Dutch satellite. Eight hours later, after maneuvering the satellite to use a different set of detectors, the *BeppoSAX* observing team discovered an x-ray source that was located at the position of the initial blast of gamma-rays. A second observation 3 days later showed that the source was fading—they had discovered the first x-ray "afterglow" of a gamma-ray burst, radiation emitted by the cooling embers of the much hotter initial explosion. The *BeppoSAX* team then alerted observers at ground-based telescopes, leading to the discovery of the first visible-light afterglow. On May 8, 1997, the *BeppoSAX* team located another x-ray afterglow, and quick work resulted in the first determination of a distance to a gamma-ray burst—about 7×10^9 light-years—providing the first real evidence that gamma-ray bursts are located at cosmic distances. *See* X-RAY ASTRONOMY.

The distances to the gamma-ray bursts cannot be obtained solely by using gamma-ray data, but result from the measurement of red-shifted spectral lines in the visible-light spectrum of the afterglow. (Spectral lines are redshifted due to the expansion of the universe. As the universe expands, light waves traveling great distances are stretched by the expansion. Longer-wavelength light is "redder," hence the term redshift. With a few additional assumptions about the expansion history of the universe, a measured redshift can then be converted into a distance.) As of 2006, distances had been measured to over 70 gamma-ray bursts: the farthest, detected on September 4, 2005, by NASA's *Swift* satellite, is at a distance of about 12.8×10^9 light-years (redshift of 6.3), and the average distance is 9.8 billion light-years (redshift of 1.7). For comparison, the most distant galaxy seen so far is a quasar at about 13×10^9 light-years (or a redshift of 6.5). Evidence has accumulated that bursts with afterglows are located in regions of distant galaxies that are actively forming stars. *See* COSMOLOGY; GALAXY, EXTERNAL; QUASAR; REDSHIFT.

Observations with burst alert systems. As BATSE accumulated observations of gamma-ray bursts, ground-based observers became eager to help unravel the mystery of the bursts' origins. To facilitate these efforts, in 1993 Scott Barthelmy created a system for distributing information about the bursts, including crude positions, to interested observers. Originally called BACODINE (BAtse COordinates DIstribution NEtwork), the system used telephones, pagers, and eventually the Internet to notify interested astronomers of burst arrivals within seconds. In 1998, the system was renamed the Gamma-ray Burst Coordinates Network (GCN). It now sends out burst alert notices from several satellites, and also provides a repository for the ground- and satellite-based observers to post their scientific results. In many cases the burst alerts are directly transmitted into the control systems for robotic telescopes linked through the Internet. The telescopes then autonomously repoint to observe the sky locations of the bursts.

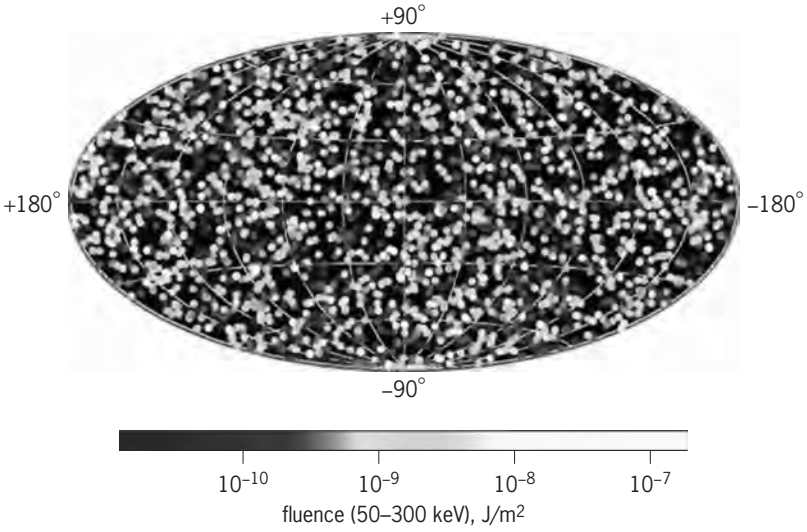

Fig. 1. Equal-area projection map of the locations of over 2700 gamma-ray bursts on the sky as measured by the BATSE detector on the *Compton Gamma-Ray Observatory*. Different shades indicate the total energy emitted (fluence) for each burst. The plane of the Milky Way Galaxy spans the Equator. (*NASA/CGRO/BATSE Team*)

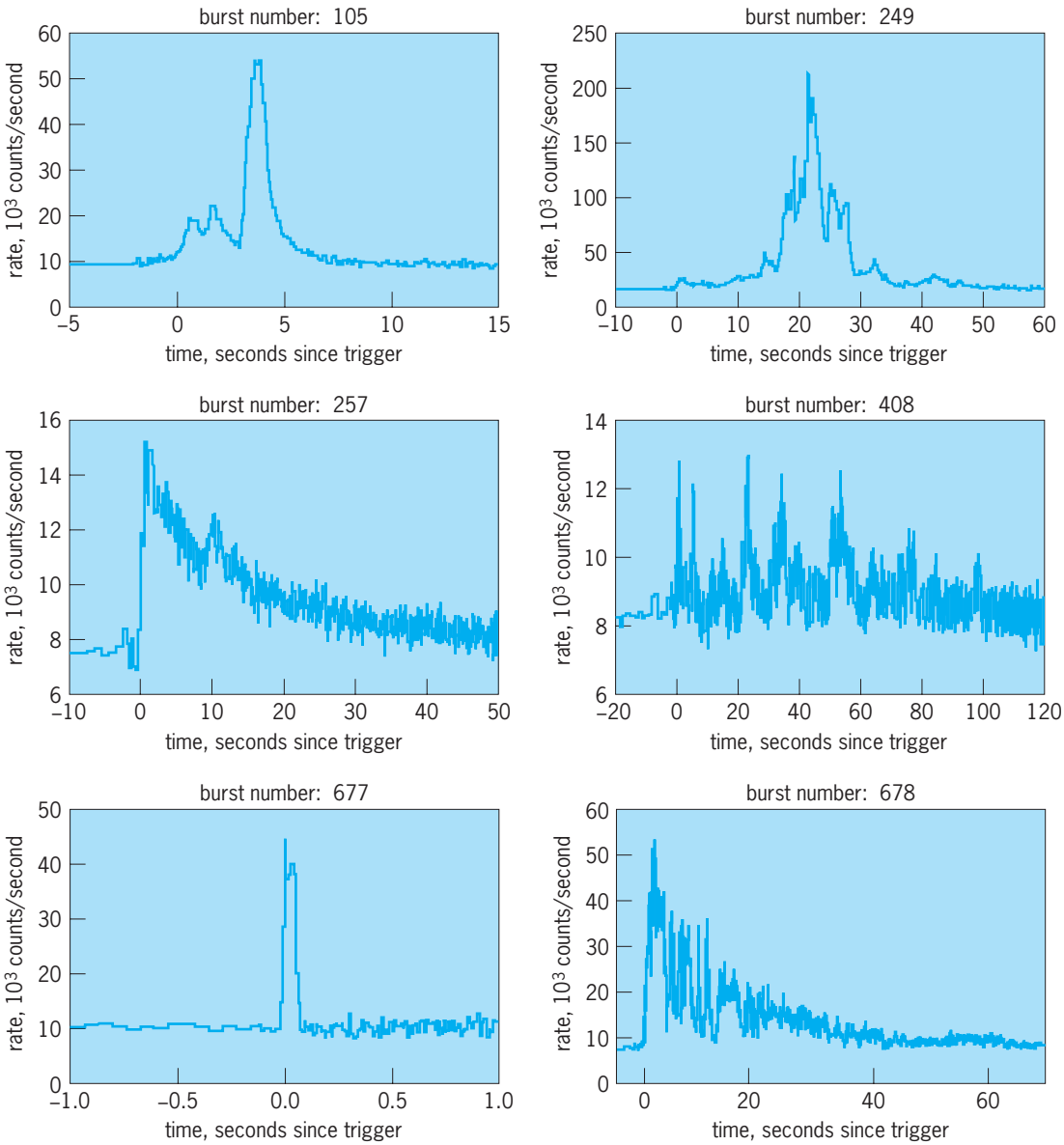

Fig. 2. Time histories of various gamma-ray bursts detected by BATSE on the *Compton Gamma-Ray Observatory*, showing a broad range of time scales. (*NASA/CGRO/BATSE Team*)

The Robotic Optical Transient Search Experiment (ROTSE) was the first autonomous telescope to be connected to the Internet-driven burst alert system. After over 5 years of monitoring the burst alert notices, on January 23, 1999, ROTSE quickly repointed to acquire a series of images beginning 22 seconds after the initiation of a burst reported by BATSE. It was the first burst "caught in the act." ROTSE's images of the rapidly fading "prompt" optical emission (rather than the afterglow) were so bright (9th magnitude) that the optical flash could have been seen by an observer using binoculars. It was realized that, with a sufficiently quick response, it would be possible to observe optical light that was emitted during the burst itself, as well as the afterglows. Following the launch in November 2004 of NASA's *Swift* satellite, the hunt for the origin of gamma-ray bursts intensified as many additional ground-based observers joined the burst chase. One of the first *Swift* gamma-ray bursts (on January 17, 2005) was observed using onboard x-ray, ultraviolet, and optical telescopes while flaring activity continued for hundreds of seconds. *Swift*'s rapid response has allowed many ground-based robotic telescopes to study bursts in their early phases, as well as the compiling of detailed observations of afterglow emission across the electromagnetic spectrum.

Observational properties. The time histories of the brightnesses of gamma-ray bursts show a remarkable range of variation. (**Fig. 2**). Some bursts last only a few milliseconds, while others show multiple peaks that last for minutes. These widely varying characteristics pose a difficult challenge to scientists seeking to understand the origin of the bursts.

There are two classes of gamma-ray bursts that are primarily distinguishable by their durations and spec-

tral characteristics: short bursts, with durations less than 2 seconds and more energetic (harder) spectra; and bursts longer than 2 seconds, with less energetic (softer) spectra. Almost all the afterglow measurements that have been obtained so far are from the longer, softer class of gamma-ray bursts. To date, there are only a handful of short, hard bursts with observed afterglows: on May 9, 2005, NASA's *Swift* satellite detected the first x-ray afterglow, and on July 7, NASA's *HETE-2* satellite detected a short burst that led to the first optical afterglow, seen in data collected with ground-based telescopes.

Possible origins. With the knowledge that the longer-duration gamma-ray bursts are being created at vast distances, in star-forming regions of galaxies, speculation has increased that they are somehow related to supernova explosions of massive stars that have reached the end of their relatively short, nuclear-fueled lives. Such massive stars preferentially form in denser regions in galaxies where there is an abundance of material to seed the star-formation process, and burn very brightly, exhausting all their fuel within a few million years. When a massive star exhausts its nuclear fuel, it will collapse inwardly while flinging out its outer layers in a supernova explosion. If a large piece, greater than 3 solar masses (6×10^{30} kg or 1.3×10^{31} lb), of the inner core remains intact, it can collapse directly into a black hole. *See* BLACK HOLE; STELLAR EVOLUTION; SUPERNOVA.

There is direct evidence that two gamma-ray bursts are associated with supernovae. On March 29, 2003, an extremely bright gamma-ray burst was detected by NASA's *HETE-2* satellite, and its position was relayed by the Gamma-ray Burst Coordinates Network. This well-studied burst produced an afterglow that was detected by visible-light observers at ground-based telescopes around the world. A detailed study of the afterglow light revealed the characteristic signature of a supernova, located only 2×10^9 light-years away—one of the closest gamma-ray bursts ever detected. Further studies of this supernova proved conclusively that it was related to the gamma-ray burst, providing indisputable proof that the two events were connected. On February 18, 2006, *Swift* detected a gamma-ray burst precursor to a supernova, which occurred later that week, as astronomers worldwide trained their telescopes on the amazing event. This burst, at a distance of 440 million light-years, was the second-closest burst on record, offering another clue to the gamma-ray burst–supernova connection.

In the supernova scenario—named the "collapsar" or "hypernova" model—an extremely massive star, greater than 100 times the mass of the Sun (2×10^{32} kg or 4×10^{32} lb), explodes, and then its inner core collapses to form a black hole. The newly born black hole then accretes some of the swirling and collapsing matter to form jets. These jets carry matter and energy outward, plowing through the collapsing layers of matter, and creating internal shocks that produce gamma rays (**Fig. 3**). The interaction of the jets with the collapsing material produces a wide variety of different gamma-ray time histories, forming

Fig. 3. Conception of a hypernova explosion with gamma-ray emitting jets. (*Dave Ambrecht, General Dynamics*)

gamma-ray bursts with different numbers of peaks and durations—depending on the amount and distribution of the matter that is being shocked. When the jets impact the more tenuous matter blown off earlier by the star as it evolves, additional shocks create x-rays, and then visible and finally radio light as the matter cools. These lower-energy emissions are seen as the afterglow, which varies depending on the density and temperature of the ambient material in the nearby vicinity. Radio observations, at the Very Large Array have shown that the radio emissions from gamma-ray bursts are highly collimated into beams or jets, thus reducing the energy needed to power the bursts. (Since the gamma rays that we see are in narrow jets aimed in our direction, rather than emitted uniformly in all directions, the energy inferred in the gamma-ray burst is lowered by a factor of around 100.) *See* RADIO TELESCOPE.

Important clues to the origin of the shorter gamma-ray bursts—those with duration less than 2 s—were revealed in 2005 with the discovery of afterglows by *Swift* and *HETE-2*. These observations are consistent with the orbital decay of two compact objects in a binary system: two neutron stars, black holes, or perhaps one of each. Einstein's theory of gravity (general relativity) predicts that binary orbits will decay by the emission of gravitational radiation. Several binary pulsar (neutron star) systems are now known (including the original system, discovered by Russell Hulse and Joseph Taylor). As the two objects finally gravitate together, their death spiral will lead to the formation of a black hole, accompanied by a relatively small disk of debris. The formation of jets from the accretion of this disk onto the rapidly spinning black hole can release a tremendous surge of gamma rays (**Fig. 4**). In this case, there is little nearby material to create afterglow emission, and the gamma rays can escape the stellar envelope more quickly, as not much obscuring matter is present. Hence the gamma-ray bursts that are formed in

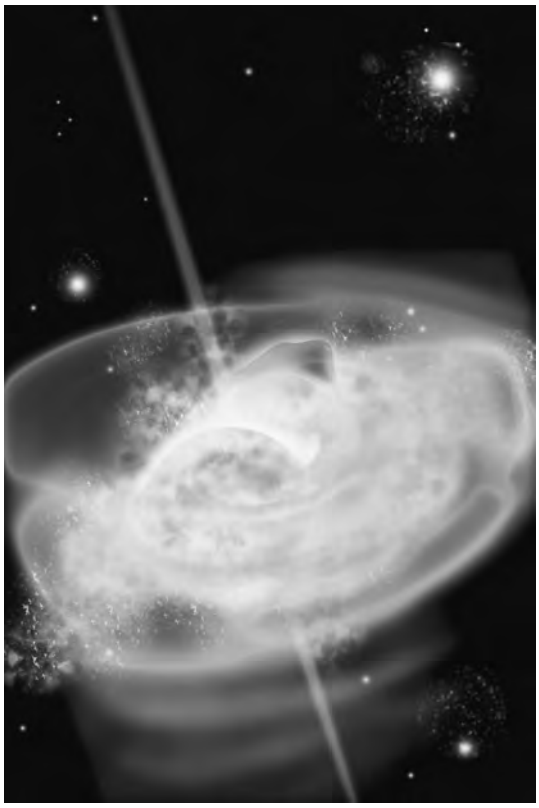

Fig. 4. Conception of merging neutron stars. (*Aurore Simonnet, Sonoma State University*)

the merger scenario are much shorter, and should not be accompanied by long-lived afterglows. The x-ray and optical afterglows from the short bursts seen with *Swift* and *HETE-2* have confirmed this picture. The afterglows are short-lived, with no evidence for any supernovae. In addition, the locations in which bursts occurred are older, elliptical galaxies, or galactic regions that are not actively forming stars. In both the collapsar and merger scenarios, accounting for the longer and shorter gamma-ray bursts respectively, the bursts appear to signal the creation of black holes. *See* GRAVITATIONAL RADIATION; PULSAR; RELATIVITY.

Present and future studies. In mid-2006, eight satellites were observing gamma-ray bursts, and participating in the Interplanetary Network: NASA's *HETE-2*, *RHESSI*, *Mars Odyssey*, and *Swift* spacecraft; the *Ulysses* mission of NASA and the European Space Agency (ESA); NASA's *Wind* experiment on the Russion *Konus* mission; the Japanese *Suzaku* satellite; and ESA's *INTEGRAL* mission. Future missions that are planned to study gamma-ray bursts include the *Gamma-ray Large Area Space Telescope* (*GLAST*), scheduled for launch in 2007.

Lynn Cominsky

Bibliography. N. Gehrels et al., The *Swift* gamma-ray burst mission, *Astrophys. J.*, 611:1005–1020, 2004; N. Gehrels, L. Piro, and P. J. Leonard, The brightest explosions in the universe, *Sci. Amer.*, 287(6):84–92, December 2002; C. Kouveliotou, R. C. Duncan, and C. Thompson, Magnetars, *Sci. Amer.*, 288(2):34, 2003; F. Reddy, The star-splitting jets of supernovae, *Astronomy*, 32(10):44, October 2004; D. Reichart, The gamma-ray burst/supernova connection, *Mercury*, 32(5):15, September/October 2003; G. Schilling, Catching gamma-ray bursts on the wing, *Sky & Telesc.*, 107(3):32–40, March 2004; G. Schilling, *Flash! The Hunt for the Biggest Explosions in the Universe*, Cambridge University Press, 2002; J. Wise, How to outsmart a gamma-ray burst, *Pop. Sci.*, October 2004.

Gamma-ray detectors

Instruments that register the presence of gamma (γ) radiation. Such detectors convert some or all of the energy of gamma radiation into an electrical signal. Most instruments are capable of detecting individual gamma-ray photons (or quanta), each of which produces a short (0.1–5-microsecond) current pulse at the detector output. The output pulses may be made visible on an oscilloscope, made audible through a speaker (such as the familiar Geiger counter), or be electronically processed further, depending on the application. *See* GAMMA RAYS; OSCILLOSCOPE.

Gamma-ray detectors range from hand-held devices capable of giving some indication of the intensity of a radiation field, to devices that accurately measure the energy and event time of individual photons reaching detectors assembled into a single complex instrument. These diverse detectors are widely used in industry, medicine, and research.

Response to gamma radiation. In common with most radiation detectors, gamma-ray detectors respond not to the radiation but to its secondary effects, in this case energetic electrons. Photons have neither mass nor charge and pass through matter rather easily. In so doing, they lose energy by (1) elastic scattering of electrons (Compton effect), (2) electron-positron ($\beta^+\beta^-$) pair production, and (3) at lower energies by photoabsorption. In these processes the energy of the photon is converted to the kinetic energy of the few electrons with which it interacts. Since electrons are much less penetrating than gamma-ray photons, their energy is largely trapped within the detector, where their ionizing effect creates a response convertible to an electrical output. In a gas-ionization device, such as Geiger counter, this occurs by the production of ion–electron pairs and in a solid-state device, such as a germanium detector, by production of electron-hole pairs. In a scintillation device, for example, a sodium iodide (NaI) detector, the response is caused by the emission of optical photons from atoms excited by the passage of energetic electrons. *See* COMPTON EFFECT; ELECTRON-POSITRON PAIR PRODUCTION; GEIGER-MÜLLER COUNTER; IONIZATION CHAMBER; SCINTILLATION COUNTER.

In accurate instruments the magnitude of the current pulse created by a single gamma-ray photon is closely proportional to the energy within the detector volume. However, gamma radiation is so penetrating that any particular event may involve only partial absorption of the photon. For example,

a single Compton scattering may be followed by the escape of the scattered photon (now reduced in energy) from the detector, leaving behind only the energy of the scattered electron.

Photon interactions in matter. The three principal interactions of gamma rays with matter depend on the photon energy and the atomic number (Z) of the absorber. In photoabsorption, a photon is absorbed by collision with an atom, its energy and momentum being conserved by the ejection of an electron. This mechanism is important to gamma-ray detection since it is the only one of these interactions in which the photon energy is completely captured by ionization. This mechanism dominates at low photon energies and for absorbers of high Z; for high-energy photons, pair production dominates (**Fig. 1**). This process cannot occur unless the photon energy exceeds twice the rest-mass of an electron (2×511 keV), the minimum requirement for creating an electron-positron pair ($\beta^+ \beta^-$). Although this positron loses energy rapidly in a solid, it ultimately annihilates an electron to liberate two new gamma photons, each of energy 511 keV. To register the full energy, both photons must be contained within the detector.

Desirable detector characteristics are therefore (1) high Z, enhancing the domain of photoabsorption, (2) reasonable size, giving photons a chance to convert energy to ionization in more than one encounter, and (3) high density, maximizing the number of electrons encountered in the distance traveled.

Gamma-ray spectra and detector characteristics. An experiment may be imagined in which a detector is placed adjacent to a point source such as a cobalt-60 source that emits gamma rays of energies 1173 and 1332 keV. The output current pulses (one per photon detected) are amplified electronically and digitized. The resulting stream of data is stored in the form of a histogram showing numbers of events on the ordinate and output pulse amplitude on the abscissa, and called a spectrum by analogy with optical spectra. Since the source has only two photon energies (or wavelengths, photon energy E being related to wavelength λ by $\lambda = hc/E$, where h is Planck's constant and c is the speed of light), two lines are found in the spectrum (**Fig. 2**).

For events at the peak energy (or photopeak), all the photon energy is trapped within the sensitive volume. The background of partial-energy events is called the Compton platform. Such spectra help to define the three characteristic parameters of gamma-ray detectors: (1) resolution, the width of a peak at its half-height point, which measures the minimum energy separation at which two different photon peaks of comparable intensity are distinguishable; (2) total detection efficiency, the number of events registered in the spectrum versus the number of photons known to have struck the detector (by calculation from the source emission rate); and (3) peak-to-total ratio, the number of events in the photopeak versus the total number detected at that and all lower energies. These characteristics vary with photon energy, which must therefore be specified with the performance figures.

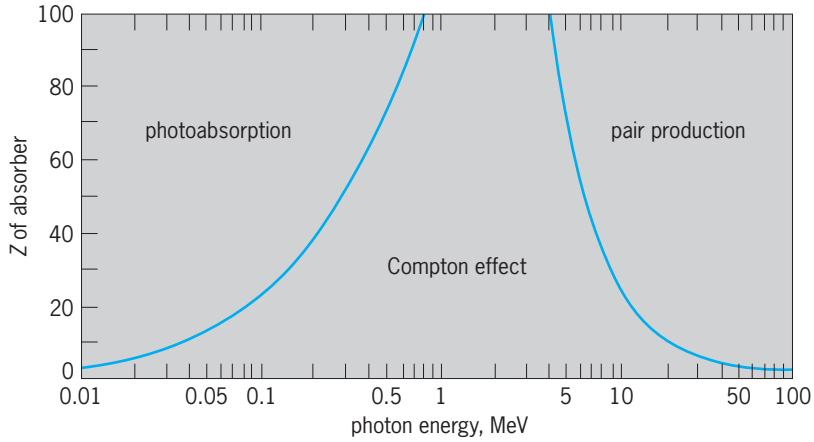

Fig. 1. Domains in photon energy versus atomic number (Z) of an absorber. Each of the three mechanisms by which gamma rays lose energy to ionization is of dominant importance in its domain. *(After R. D. Evans, The Atomic Nucleus, McGraw-Hill, 1955)*

Gas-ionization detectors. These detectors are simple, cheap, and reliable but have very low efficiency because of the low density of their sensitive volume. Their principal application is in monitoring strong radiation fields for industry and radiation protection. The individual output pulses are electronically integrated; the instrument gives a reading of count rate or field intensity, often calibrated directly in dose

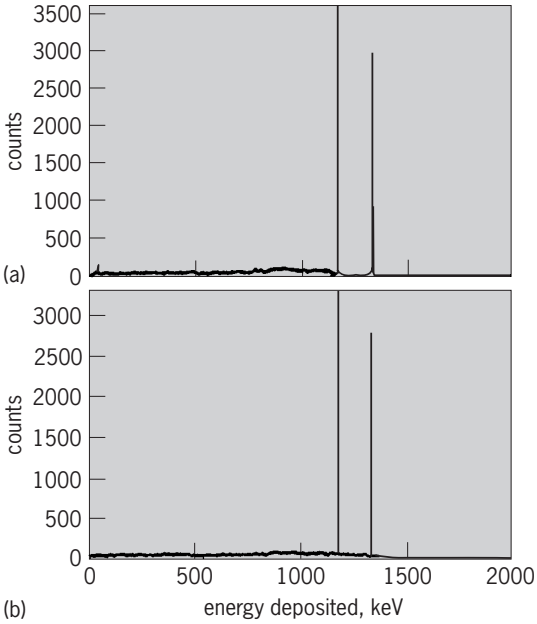

Fig. 2. Gamma-ray spectra of a cobalt-60 radiation source. (*a*) Spectrum from a high-purity germanium detector of 25% efficiency. The two peaks represent events in which the original photon energy (1173 and 1332 keV) was completely contained within the crystal. The platform at lower energies corresponds to events in which a photon lost some energy to the crystal but eventually escaped. (*b*) Spectrum that results from activating a bismuth germanate scintillation detector surrounding the high-purity germanium crystal so that the event is not recorded if both detectors give an output. The platform of degraded events seen in the high-purity germanium detector is substantially suppressed.

rate. The amplification occurs in the gas itself as a discharge develops under the intense electric field surrounding the anode wire. This process is virtually noise-free and results in a gain of about 10^6 on the original number of electrons liberated.

Scintillation detectors. These detectors employ certain materials which are nearly transparent and convert a significant amount of ionization energy into light emitted in the visible or near-visible region. The material is viewed by one or more photomultiplier tubes, which collect the light and convert it into an electrical pulse with very low noise and a typical gain of 10^8. The most common scintillator for gamma-ray application is sodium iodide with a thallium doping, NaI(Tl). Crystals can be produced in virtually any shape or size up to 15 × 15 in. (38 × 38 cm) cylinders; the most common laboratory detector is 3 × 3 in. (7.6 × 7.6 cm). For small crystals, the resolution is typically 6.5% at 662 keV (cesium-137 radiation) and increases approximately with the inverse square root of the energy. *See* PHOTOMULTIPLIER.

Other alkali halide scintillators in common laboratory use are barium fluoride (BaF_2) and cesium iodide (CsI). These have a higher density and average atomic number than sodium iodide, and for a given detector size give a higher efficiency and higher peak-to-total ratio, although their resolution is slightly poorer. The barium fluoride scintillator is used where a fast response is important, as with very high counting rates or fast electronic timing.

Another class of scintillator used for gamma-ray detection comprises the heavy-metal complex oxides, especially bismuth germanate ($Bi_4Ge_3O_{12}$). It has a very high density and offers the strongest gamma-radiation absorption of any known practical scintillator. In resolution it is appreciably poorer than sodium iodide; a good performance for a 3 × 3 in. (7.6 × 7.6 cm) crystal is 11.5% at 662 keV.

Solid-state detectors. For general application, the germanium detector is the most useful. The first successful ones, called Ge(Li) detectors, depended on lithium compensation of crystal imperfections. Manufacturing techniques have been improved, and detectors based on high-purity intrinsic germanium (HPGe), are now readily available. Germanium detectors function at liquid-nitrogen temperature. High-purity germanium detectors have the following advantages over Ge(Li) detectors: they may be thermally cycled, and radiation damage, such as that caused by neutrons, can be annealed cheaply and easily. *See* JUNCTION DETECTOR.

Germanium detectors have better resolution than most other gamma-ray detectors, which scales differently with energy than the resolution of scintillation detectors: a typical performance is 1 keV resolution at 100 keV, 2 keV at 1300 keV, and 3 keV at 2000 keV. Historically, the photopeak efficiency of germanium detectors has been defined in terms of a 3 × 3 in. (7.6 × 7.6 cm) NaI(Tl) detector at the same distance from a radiation source, usually cobalt-60. On this standard, an efficiency of 25% is typical for high-purity germanium, but detectors of up to 70% are routinely available, and a 100% detector has been demonstrated.

Silicon detectors have less general applicability. They must also be cooled to liquid-nitrogen temperature, and have significantly better resolution than germanium detectors but are less efficient. These detectors are useful in the spectroscopy of very low-energy radiation (below 50 keV) where the efficiency is equal but the resolution (typically 0.6 keV) nearly two times better than that of germanium.

Detector systems. Powerful instrumentation has been created by assembling separate gamma-ray detectors into a single unit. For example, Fig. 2*b* shows the performance gained by suppressing escaped radiation from a high-purity germanium detector. In this case the active crystal was surrounded by a bismuth germanate scintillator detector except in the direction of the radiation source. Electronic circuits processed the output pulses from both detectors so that if they were simultaneous (within 5×10^{-8} s) the event was rejected. In some instruments a spherical shell of scintillation detectors (typically 100 elements) surrounds the radiation source. In one application the outputs of all the detectors are summed whenever an event occurs. In this way the instrument acts as a gamma-ray calorimeter, registering an output proportional to the energy of gamma radiation emitted in the event. For example, the fusion of two heavy nuclei is typically accompanied by a burst of up to 30 gamma rays with a total energy of perhaps 30 MeV. Photon energies exceeding 100 MeV are associated with the decay of subnuclear systems such as the J/ψ particle. *See* ELEMENTARY PARTICLE; NUCLEAR FUSION; PARTICLE DETECTOR.

David Ward

Bibliography. D. A. Bromley (ed.), Detectors in nuclear science, *Nucl. Instrum. Meth.*, vol. 162, 1979; J. Cerny (ed.), *Nuclear Spectroscopy and Reactions*, pt. A, 1974; K. Debertin and R. G. Helmer, *Gamma and X-Ray Spectrometry with Semiconductor Detectors*, 1988; H. Ejiri and M. J. De Voigt, *Gamma-Ray and Electron Spectroscopy in Nuclear Physics*, 1989; R. M. Singro, *Introduction to Experimental Nuclear Physics,* 1975.

Gamma rays

Electromagnetic radiation emitted from excited atomic nuclei as an integral part of the process whereby the nucleus rearranges itself into a state of lower excitation (that is, energy content). For the theory of gamma emission *See* NUCLEAR STRUCTURE; RADIOACTIVITY.

Nature of gamma rays. The gamma ray is an electromagnetic radiation pulse—a photon—of very short wavelength. The electric (**E**) and magnetic (**H**) fields associated with the individual radiations oscillate in planes mutually perpendicular to each other and also the direction of propagation with a frequency ν which characterizes the energy of the radiation. The **E** and **H** fields exhibit various specified phase-and-amplitude relations, which define the character

of the radiation as either electric (EL) or magnetic (ML). The second term in the designation indicates the order of the radiation as 2^L-pole, where the orders are monopole (2^0), dipole (2^1), quadrupole (2^2), and so on. The most common radiations are dipole and quadrupole. Gamma rays range in energy from a few kiloelectronvolts to 100 MeV, although most radiations are in the range 50–6000 keV. As such, they lie at the very upper high-frequency end of the family of electromagnetic radiations, which include also radio waves, light rays, and x-rays. *See* ELECTROMAGNETIC RADIATION; MULTIPOLE RADIATION; PHOTON.

Wave-particle duality. The dual nature of gamma rays is well understood in terms of the wavelike and particlelike behavior of the radiations. For a gamma ray of intrinsic frequency ν, the wavelength is $\lambda = c/\nu$, where c is the velocity of light; energy is $E = h\nu$, where h is Planck's constant. The photon has no rest mass or electric charge but, following the concept of mass-energy equivalence set forth by Einstein, has associated with it a momentum given by $p = h\nu/c = E/c$. *See* LIGHT; QUANTUM MECHANICS; RELATIVITY.

While gamma rays and x-rays are usually labeled by their energies (in kiloelectronvolts or megaelectronvolts), an equivalent useful specification in terms of either wavelength or frequency is easily obtained from the relationships given above. For less energetic members of the electromagnetic family, it is customary to classify radio waves in terms of wavelength (in meters) or frequency (in hertz). Photons in the visible and near-visible regime are labeled alternately by energy (in electronvolts) or wavelength (in nanometers). However, the general considerations on propagation, polarization, and Doppler shifts are quite similar for all. One of the early techniques for precise measurements of low-energy nuclear gamma rays utilized a bent-crystal spectrometer, which measures wavelength; the corresponding energies were subsequently obtained from the equivalence relations given above. *See* DOPPLER EFFECT; NUCLEAR CHEMISTRY; POLARIZATION OF WAVES.

Origin. From a historical perspective, the term "gamma ray" was introduced to label the electromagnetic radiations emanating from nuclear deexcitations (as distinct from alpha, beta, and x-rays). Early observations of nuclear giant-dipole resonance phenomena dealt with gamma rays in the range up to about 30 MeV. High-energy radiations are also observed in cosmic events; for example, the pi-zero meson (π^0) annihilates by the simultaneous emission of two gamma rays with total energies in the 100-MeV range. *See* GAMMA-RAY ASTRONOMY; GIANT NUCLEAR RESONANCES.

Other sources of very high energy electromagnetic radiations (or gamma rays) have become of increasing interest. Electron beams in the gigaelectronvolt range can produce gamma rays in the range of hundreds of megaelectronvolts via brehmsstrahlung from a fixed target, or alternatively from a collision between high-energy beam electrons and an opposing photon beam from a simple laser (all carried out in vacuum). *See* BREMSSTRAHLUNG; LASER; PARTICLE ACCELERATOR.

For example, if an argon laser [wavelength (λ) $\approx$ 350 nm; energy (E) $\approx$ 3.5 eV] is focused opposite to the electron beam direction, low-energy photons backscattered from the electrons (in Compton scattering) will achieve an energy boost proportional to the electron momentum, resulting in scattered photons with energies of around 500 MeV. With this technique, simultaneous observation of the scattered electron (direction and momentum) also defines the precise gamma-ray energy. Since the Compton scattering process conserves polarization, the polarization of the gamma-ray flux is also known. This production process has been used to map out the region of the nuclear delta resonance (at around 300 MeV), which lies well above the giant dipole resonance. *See* COMPTON EFFECT; DELTA RESONANCE.

Nuclear gamma rays. One of the most frequently utilized sources of nuclear gamma rays is ^{60}Co (that is, the cobalt isotope of $N = 33$ neutrons, $Z = 27$ protons, and thus of atomic mass number $A = N + Z = 60$). The decay process (**Fig. 1**) begins when ^{60}Co (in its ground state, or state of lowest possible excitation) decays to ^{60}Ni ($N = 32$, $Z = 28$) by the emission of a β^- particle. More than 99% of these decays lead to the 2506-keV level of ^{60}Ni; this level subsequently deexcites by an 1173-keV gamma transition to the 1332-keV level, which in turn emits a 1332-keV gamma ray leading to the ^{60}Ni ground state.

The gamma rays from ^{60}Ni carry information not only on the relative excitation of the ^{60}Ni levels, but also on the quantum-mechanical nature of the individual levels involved in the gamma decay. From the standpoint of nuclear physics, the levels of a given nucleus can be described most simply in terms of their excitation energies ($\mathbf{E}_x$) relative to the ground state, and in terms of the total angular momentum ($\mathbf{J}$) and parity (π) quantum numbers given as $\mathbf{J}^\pi$. For a gamma-ray transition from initial state i to final state f, one obtains $E^i{}_x - E^f{}_x = E'_\gamma$ is the measured gamma energy after small (second-order) corrections

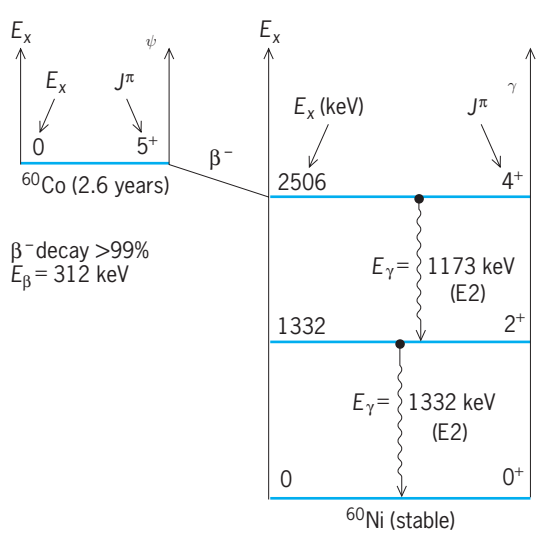

Fig. 1. Energy-level diagram illustrating the gamma decay of levels of ^{60}Ni resulting from beta decay of ^{60}Co.

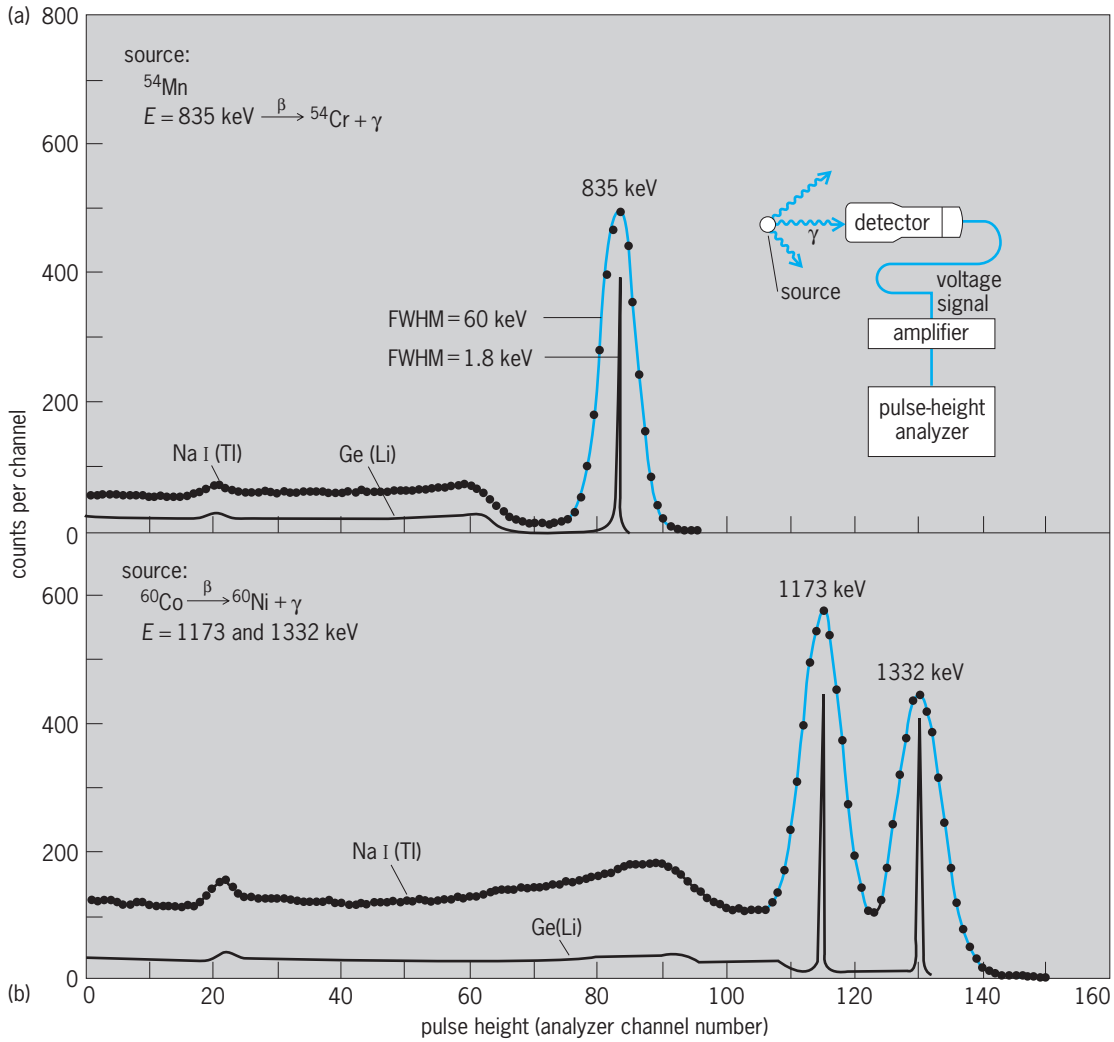

Fig. 2. Gamma-ray spectra from radioactive sources as measured with both NaI(Tl) and Ge(Li) detectors. Inset shows the components of detector apparatus. (*a*) ^{54}Mn source. (*b*) ^{60}Co sources.

for nuclear recoil and relativistic effects. Nuclear selection rules restrict the multipole character of the radiation according to the change in the quantum numbers J^π of the initial and final states. In Fig. 1, for example, the transitions must be electric quadrupole (E2), since they connect states of similar parity ($\pi = +$) by radiation of order $L \geq J_i - J_f = 2$. *See* NUCLEAR ISOMERISM; NUCLEAR SPECTRA; PARITY (QUANTUM MECHANICS); SELECTION RULES (PHYSICS); SPIN (QUANTUM MECHANICS).

Use as nuclear labels. Various nuclear species exhibit distinctly different nuclear configurations: the excited states, and thus the gamma rays which they produce, are also different. Precise measurements of the gamma-ray energies resulting from nuclear decays may therefore be used to identify the gamma-emitting nucleus, that is, not only the atomic number Z but also the specific isotope as designated by A. This has ramifications for nuclear research and also for a wide variety of more practical applications.

The two most widely used detectors for such studies are the NaI(Tl) detector and the Ge(Li) detector. **Figure 2** shows typical gamma spectra measured for sources of ^{60}Co and ^{54}Mn. Full-energy peaks are labeled by the gamma-ray energy, given in kiloelectronvolts. The figure of merit for these detectors, defined for a given gamma energy as the full-width-at-half-maximum (FWHM) for the full energy peak, is indicated. Although the more efficient NaI(Tl) detector can clearly distinguish the ^{60}Co and ^{54}Mn gamma rays, it is evident that the Ge(Li) detector, having a line width of only 1.8 keV, is more appropriate for complex nuclei, or for studies involving a greater number of source components. *See* GAMMA-RAY DETECTORS.

Applications to nuclear research. One of the most useful studies of the nucleus involves the bombardment of target nuclei by energetic nuclear projectiles, to form final nuclei in various excited states. For example, ^{48}Ca bombarded by ^{16}O makes ^{60}Ni strongly via the ^{48}Ca(^{16}O, $4n$)^{60}Ni reaction, as well as numerous other final species. Ge(Li) measurements of the decay gamma rays are routinely used to identify the

various final nuclei according to their characteristic gamma rays, that is, the 1332- and 1173-keV gamma rays of ^{60}Ni, for example.

Precise measurements of the gamma energies, together with intensity and time-coincidence measurements, are then used to establish the sequence of gamma-ray decay, and thus construct from experimental evidence the nuclear level scheme. Angular correlation and linear polarization measurements determine the radiation character (as M1, E1, E2, or M2 or mixed) and thus the spin-parity of the nuclear levels. These studies provide a very useful tool for investigations of nuclear structure and classification of nuclear level schemes.

Practical applications. In these applications, the presence of gamma rays is used to detect the location or presence of radioactive atoms which have been deliberately introduced into the sample. In irradiation studies, for example, the sample is activated by placing it in the neutron flux from a reactor. The resultant gamma rays are identified according to isotope by Ge(Li) spectroscopy, and thus the composition of the original sample can be inferred. Such studies have been used to identify trace elements found as impurities in industrial production, or in ecological studies of the environment, such as minute quantities of tin or arsenic in plant and animal tissue. *See* ACTIVATION ANALYSIS.

In tracer studies, a small quantity of radioactive atoms is introduced into fluid systems (such as the human bloodstream), and the flow rate and diffusion can be mapped out by following the radioactivity. Local concentrations, as in tumors, can also be determined. *See* RADIOACTIVE TRACER.

Doppler shift. If $E_{\gamma 0} = h\nu_0$ is the gamma ray energy emitted by a nucleus at rest, then the energy $E_\gamma = h\nu$ emitted from a nucleus moving with velocity v at angle θ (with respect to the direction of motion) is given by Eq. (1) where c is the velocity of light. In

$$E_\gamma = E_{\gamma 0}\left(1 + \frac{v}{c}\cos\theta\right) \qquad (1)$$

terms of the frequency ν, this expression is entirely analogous to the well-known Doppler shift of sound waves. Experimental measurements of the Doppler shift are used to determine the velocity of the nucleus and, more importantly, to shed light on the lifetime of the nuclear gamma-emitting state. A major advantage of this technique is that the same nuclear reaction which produces the excited nuclear states can also be employed to impart a large velocity to the nucleus.

For example, the velocity of ^{60}Ni nuclei produced via the ^{48}Ca(^{16}O,$4n$)^{60}Ni reaction at $E(^{16}\text{O}) = 50$ MeV is $v/c = 0.00204$, and instead of $E_\gamma = 1332$ keV one should observe $E_\gamma = 1359$ keV. The extent of the shift is clearly within the resolving power of the Ge(Li) detector, which may therefore be used to measure E_γ and thus infer v. In most nuclear reactions, v is a known function of time t [that is, $v = v(t)$], and one therefore obtains a distribution of E_γ's whose precise

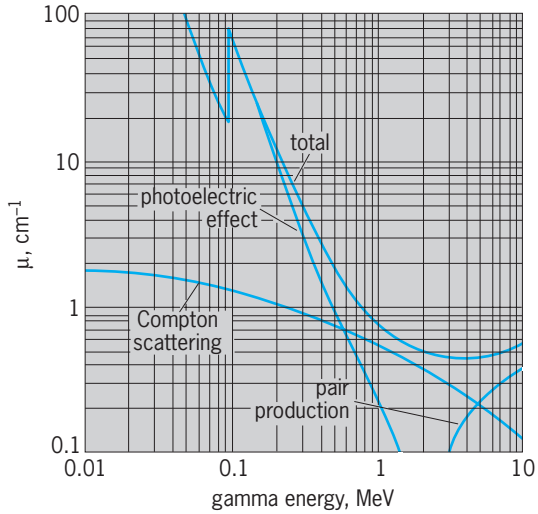

Fig. 3. Graphic representation of partial and total attenuation coefficients for lead as a function of gamma energy. (*National Institute of Standards and Technology*)

shape may be related to the lifetime of the nuclear state.

Doppler-shift measurements of gamma rays from recoil nuclei produced in nuclear reactions have been routinely used since the mid-1960s to measure nuclear lifetimes of 10^{-9} to 10^{-14} s, a range previously considered inaccessible to study.

Interaction with matter. For the three types of interaction with matter which together are responsible for the observable absorption of gamma rays, namely, Compton scattering, the photoelectric effect, and pair production, *See* COMPTON EFFECT; ELECTRON-POSITRON PAIR PRODUCTION; PHOTOEMISSION.

The energy of a photon may be absorbed totally or partially in interaction with matter; in the latter case the energy of the photon is reduced and its direction of motion is changed. Photons are thus absorbed not gradually, but in discrete events, and one interaction is sufficient to remove a photon from a collimated beam of gamma rays. The intensity I of a beam decreases exponentially, as in Eq. (2), where x is the

$$I = I_0 e^{-\mu x} \qquad (2)$$

path length, I_0 is the initial intensity, and μ is the linear attenuation coefficient, which is characteristic of the material and the gamma energy.

The dependence of the attenuation coefficient on gamma-ray energy is shown in **Fig. 3** for a lead absorber. For different absorbers, the attenuation is greater for the more dense materials. Most attenuation coefficients are tabulated as mass attenuation coefficients μ/ρ where ρ is the material or elemental density.

J. W. Olness

Bibliography. J. Cerny (ed.), *Nuclear Spectroscopy and Reactions*, pt. C, 1974; H. A. Enge and R. P. Redwine, *Introduction to Nuclear Physics*, 2d ed., 1995; R. D. Evans, *The Atomic Nucleus*, 1955, reprint 1982; K. S. Krane, *Introductory Nuclear Physics*, 1987.

Ganglion

A group of nerve cell bodies, usually located outside the brain and spinal cord. A ganglion located inside the central nervous system is called a nucleus.

The dorsal root ganglia are rounded clusters of cell bodies and fibers, surrounded by a connective tissue covering, located on the dorsal, or sensory, root of each spinal nerve. They lie just outside the cord and contain the cell bodies of sensory nerves whose fibers originate in some peripheral sensory receptor.

Other ganglia are given specific names which indicate their function or location, such as acoustic, cardiac, carotid, jugular, celiac, and sympathetic ganglia. Sympathetic ganglia, lying on either side of the vertebral column, unite by fiber strands to form a sympathetic chain. Preganglionic axons of the sympathetic system run along this path until they reach their terminal point in these ganglia, where they synapse with the postganglionic neurons. *See* SYMPATHETIC NERVOUS SYSTEM.

The term ganglion may be applied to a tumorlike, often cystic growth found on tendons, joints, and other connective tissues, but this usage is rare. *See* BRAIN; SPINAL CORD.

Walter Bock

Gangrene

A form of tissue death, or necrosis, usually occurring in an extremity and due to insufficient blood supply.

Dry gangrene. If no bacterial contamination is present, the part becomes dry, greenish-yellow, and finally turns brown or black. This is known as mummification. A sharp inflammatory border marks the edge of the adjacent viable tissue. This dry gangrene is seen most often in small portions of the extremities, such as the fingers and toes. Senile gangrene is the form caused by deterioration of blood supply in the elderly, usually as the result of progressive arteriosclerosis. Similar types are often present in diabetes mellitus, Reynaud's disease, and Buerger's disease (thromboangiitis obliterans).When tissue loss has progressed beyond the point of salvage, amputation may be required. *See* ARTERIOSCLEROSIS; DIABETES.

Moist or wet gangrene. When bacterial infection intervenes, putrefaction ensues, thus producing the moist or wet type of gangrene. Moist gangrene may occur in the body anywhere the blood supply is blocked and bacterial contamination occurs. The size of the area varies from a small, asymptomatic, local lesion to the destruction of an entire organ, such as the gall bladder or appendix. The parts most commonly affected by dry gangrene may also become contaminated, with the subsequent production of wet gangrene.

Gas gangrene. If bacterial contamination is present and the invading microorganisms are gas-producing *Clostridium* species, rapidly progressive gas gangrene is the result. This localized but rapidly spreading necrotizing wound infection is characterized by extensive edema with gas production and discoloration of the tissue, and is often contaminated by a putrefactive odor. The disease commonly arises following septic abortion (an abortion complicated by acute infection of the endometrium) or dirt contamination of deep wounds. The microbial flora of the gangrenous wound usually comprises one or more species of toxigenic anaerobic bacteria mixed with nontoxigenic anaerobic species, aerobic species, or both. The following gram-positive, spore-forming species are generally considered the primary causal agents of gas gangrene: *C. perfringens* (*C. welchii*), *C. septicum* (*Vibrion septique*), *C. novyi* (*C. oedematiens*), *C. histolyticum, C. fallax*, and *C. bifermentans*. Anaerobic streptococci may also cause gas gangrene.

Edward G. Stuart; N. Karle Mottet; Wilbert S. Aronow

Spread. Gangrene does not necessarily follow the presence of these organisms in a wound, as initiation of the disease depends on the virulence of the organisms and other factors relating to the resistance of the host. The virulence of the gangrene-producing organisms depends on the toxins produced, and the same species may produce several immunologically different toxins. Some of the toxins have been identified as specific enzymes, such as hyaluronidase, collagenase, and lecithinase. *See* TOXIN; VIRULENCE.

Diagnosis and treatment. The rapidly spreading nature of the disease precludes extensive laboratory diagnostic aids because therapeutic measures, possibly including amputation, usually must be instituted before laboratory results are available. Procedures should begin, however, for isolation of all species present and determination of toxin type and drug sensitivity. Initial thorough surgical debridement (removal of foreign material and devitalized tissue) of all contaminated wounds is essential. Polyvalent antitoxins are available for prophylactic and therapeutic use, and penicillin and tetracycline antibiotics are valuable in some cases. Formol toxoids and alum-precipitated formol toxoids for active immunization against the most important species have been produced but are not routinely used. The effectiveness of antibiotic therapy seems to depend on the species involved in the infection and the elapsed time between injury and treatment. Clostridial sepsis should be treated with intravenous clindamycin and high-dose penicillin. Hyperbaric oxygen therapy, which involves the administration of oxygen under greater-than-atmospheric pressure, is beneficial for individuals who are diagnosed early. *See* ANTIGEN; BIOLOGICALS; ENZYME; HYPERBARIC OXYGEN CHAMBER; TOXIN; TOXIN-ANTITOXIN REACTION.

Leland S. McClung; Wilbert S. Aronow

Bibliography. W. S. Aronow, Management of peripheral arterial disease, *Cardiol. Rev.*, 13:61–68, 2005; A. L. Bisno and D. L. Stevens, Streptococcal infections in skin and soft issues, *NEJM*, 334:240–245, 1996; N. Jallali, S. Withey and P. E. Butler, Hyperbaric oxygen as adjuvant therapy in the management of necrotizing fasciitis, *Amer. J. Surg.*, 189:462–466, 2005;

B. Lorber, Gas gangrene and other *Clostridium*-associated diseases, in G. L. Mandell et al. (eds.), *Principles and Practice of Infectious Diseases*, 5th ed., pp. 2549-2560, Churchill Livingstone, New York, 2000; B. N. Perry and W. E. Floyd III, Gas gangrene and necrotizing fasciitis in the upper extremity, *J. Surg. Orthop. Adv.*, 13:57–68, 2004; R. W. Titball, Gas gangrene: An open and closed case, *Microbiology*, 151(pt. 9):2821–2828, 2005.

Gantt chart

A graphic device that depicts tasks, machines, personnel, or whatever resources are required to accomplish a job on a calendar-oriented grid. Charts may be provided for various managerial levels and responsibilities, but detailed planning occurs at the lowest organizational level. Performance may be monitored and controlled throughout the organization.

Developed by Henry L. Gantt in 1917, the Gantt chart was the first scheduling model ever devised. The original intent of the chart was to assign and control jobs on factory machines, thereby minimizing delays and conflicts in scheduling. Machine loading could thus be planned weeks or months in advance.

Bar chart. An initial step in development of a Gantt chart may be to specify the tasks or activities making up a project, as shown in the **illustration**. The amount of time required for each activity is represented as a horizontal bar on the chart, with open triangles designating original start and finish dates in this example. The open start triangle is changed to a filled triangle upon inauguration of the activity, and the bar is filled in with vertical lines to indicate progress and completion. The open finish triangle is also filled upon completion. Slippage times are documented on the chart by broken lines, and the diamond symbols are employed to indicate rescheduled work. The vertical line on the chart is the current-date indicator and indicates present and future status of the project as of that date.

Updating of a Gantt chart will reveal difficulties encountered in the conduct of a project. Possible solutions include rescheduling, overtime, multishift operations, use of additional equipment and facilities, and changes in method.

Color coding may be used to isolate or identify specified tasks. Charts may designate individual workers, machines, and facilities rather than the tasks shown in the illustration. The time scale may be subdivided into days, weeks, or other units.

Computer programs have become available for drawing and updating Gantt charts. Large and elaborate commercial display boards using a pegboard design are available, some employing revolving disks, tape, or plastic slides to portray the status of planned and current work.

The Gantt chart is an effective tool for planning and scheduling operations involving a minimum of dependencies and interrelationships among the activities. The technique is best applied to activities for which time durations are not difficult to estimate, since there is no provision for treatment of uncertainty. On the other hand, the charts are easy to construct and understand, even though they may contain a great amount of information. In general, the charts are easily maintained provided the task requirements are somewhat static.

Milestone chart. An outgrowth of the bar chart technique is the milestone chart. A milestone is an important activity in the sequence of project completion. The most significant activities may be designated major milestones. The primary difference in utilization of this concept is the graphic display, since the method and collection of data are the same.

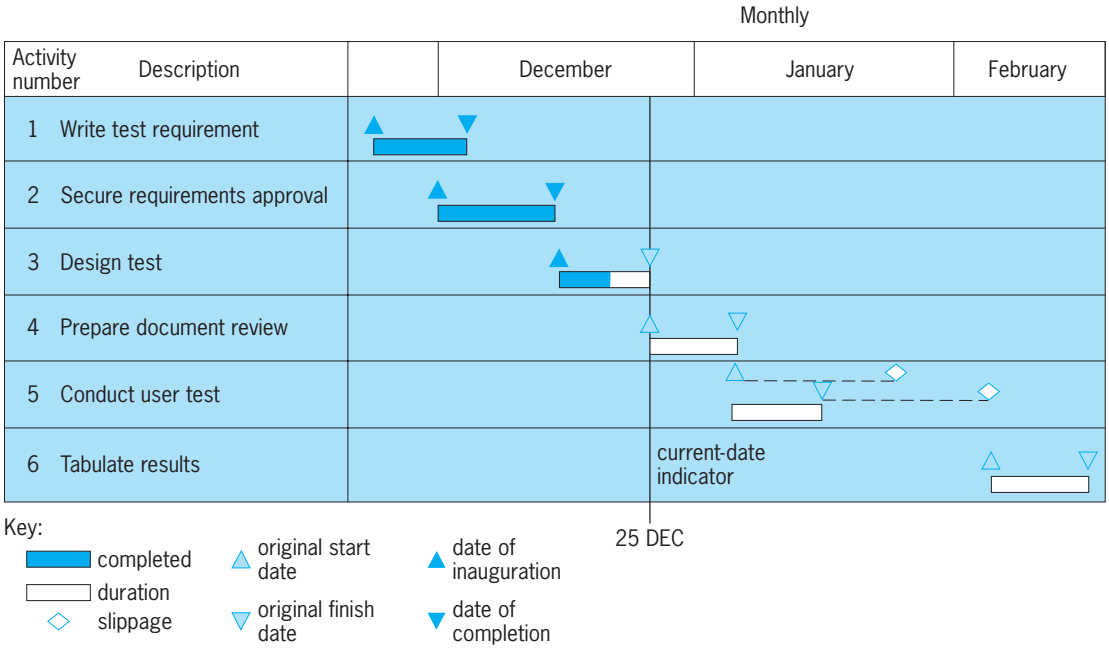

Example of a Gantt bar chart.

The milestone approach offers no intrinsic improvement over the basic Gantt chart but provides a means for focusing resources on critical items. *See* PERT.

Lawrence S. Hill

Bibliography. Harvard Business Review Staff, *Project Management*, 1991; E. Menipaz, *Essentials of Production and Operations Management*, 1984; G. Salvendy, *Handbook of Industrial Engineering*, 2d ed., 1991.

Garlic

A hardy perennial, *Allium sativum*, of Asiatic origin and belonging to the plant order Liliales. Garlic is grown for its pungent bulbs, segments of which are used primarily for seasoning. Europeans have grown it for more than 200 years. Propagation is commonly by bulb segments, sometimes called cloves; seeds are seldom produced. Cultural practices are similar to those used for onions. Popular varieties are Italian, Tahiti, and Creole or Mexican. Harvest of the mature dry bulbs is 7–8 months after planting. Garlic salt is made from dehydrated cloves. California is the most important producing state; smaller acreages are planted to garlic in Louisiana and Texas. *See* LILIALES; ONION.

J. John Carew

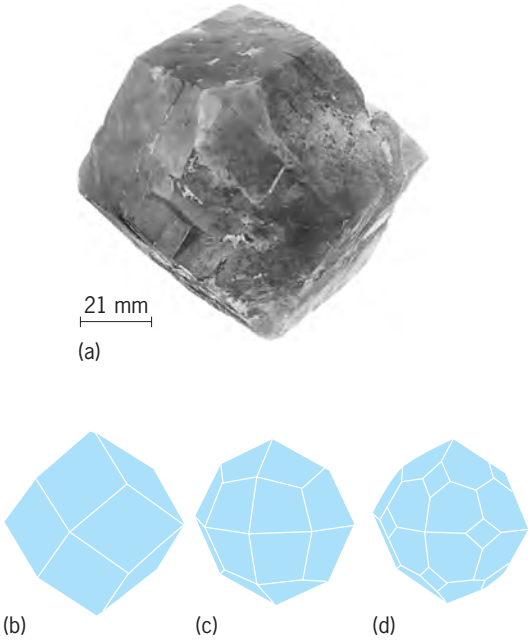

Garnet. (*a*) Grossular, Lake Jaco, Mexico (*American Museum of Natural History specimens*). (*b–d*) Crystal habits (*after C. Klein and C. S. Hurlbut, Jr., Dana's Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*). (*b*) Dodecahedron. (*c*) Trapezohedron. (*d*) Combined dodecahedron and trapezohedron.

Garnet

A hard, dense silicate mineral which occurs as crystals of cubic symmetry in a wide range of geologic environments. The general chemical formula of the silicate garnet group is $A_3B_2(SiO_4)_3$ where, in natural occurrences, the A cations are dominantly Fe^2, Mn^2, Mg, and Ca and the B cations are Al, Fe^3, and Cr^3.

The garnet crystal structure is based on isolated (SiO_4) tetrahedra with the A cations surrounded by eight oxygens at the corners of a cube and the B cations in octahedral oxygen coordination. Crystals are usually well developed with the most common forms being the dodecahedron and the trapezohedron (see **illustration**).

Chemical compositions. The garnet mineral group is generally divided into a number of individual species on the basis of chemical composition. The names of the more common of these species and the idealized compositions to which they refer are as follows:

A series	*B series*
Pyrope	Grossular
$Mg_3Al_2(SiO_4)_3$	$Ca_3Al_2(SiO_4)_3$
Almandine	Andradite
$Fe_3Al_2(SiO_4)_3$	$Ca_3Fe_2(SiO_4)_3$
Spessartin	Uvarovite
$Mn_3Al_2(SiO_4)_3$	$Ca_3Cr_2(SiO_4)_3$

Wide ranges of solid solution exist between the end members of series A and within series B, with specific garnets generally being named after the dominant species they contain. Solid solution between the two series is, however, restricted in nature partly due to the instability of such garnets and partly because of the scarcity of rocks with compositions appropriate to their formation. All compositions between pyrope and grossular and between almandine and grossular may be synthesized stably at elevated pressures (20–40 kilobars or 24 gigapascals). All of the pure end members with the exception of pyrope may be synthesized at low pressure; pressures on the order of 12 kb (1.2 GPa) are required to stabilize pyrope.

Extensive substitutions involving ions other than those listed above are not common, but there are occurrences of $(OH)_4$ replacing (SiO_4) tetrahedra in grossular (hydrogrossular) and of Ti substitution in andradite. In the latter case there is complete solid solution between andradite and $Ca_3Ti_2Fe_2SiO_{12}$ (schorlomite). *See* SOLID SOLUTION.

Physical properties. Garnets are substantially denser than most chemically analogous silicates, with specific gravities ranging between 3.58 (pyrope) and 4.32 (almandine). They also have high refractive indices (1.71–1.89) and hardness, on Mohs scale, of $6^1/_2$ to $7^1/_2$. The relative hardness, coupled with the absence of cleavage, has led to the use of garnet as an abrasive.

The color of garnet is primarily controlled by its chemical composition. Common garnets of series A are red, reddish brown, or black, while grossular-andradite mixtures are white or yellowish brown to black depending on iron content. Uvarovite is emerald green. Gem varieties of garnet are generally clear, deep red pyrope, but some darker almandines and andradites are also used as gemstones. Demantoid is a rare grass green to emerald green gem variety of

andradite. Garnets are normally optically isotropic, although some grossular-andradites exhibit biaxial optical anomalies of a sectoral nature. *See* GEM; SILICATE MINERALS.

The cubic space group of garnet is Ia3d with unit cell edge of between 1.146 and 1.205 nanometers depending on composition. Cell dimensions vary linearly with composition within the two main series. The complete range of stable solid solubility between Mg and Ca end members and between Fe^2 and Ca end members is unique among the major silicate mineral groups, limited miscibility being the general rule. The region of negative deviation from linearity at low calcium concentration, coupled with thermodynamic data, indicates that the substitution of Ca for Mg and Fe^2 may be stabilized by the formation of an ordered Ca(Mg,Fe) compound of lower symmetry than Ia3d.

Occurrence. Garnets are widespread in their occurrence, particularly in rocks which formed at high temperatures and pressures. Among the A series, pure pyrope is stable only at depths greater than about 24 mi (40 km), and garnets rich in this component are thus confined to rocks of subcrustal origin. Pyropic garnets occur in ultrabasic rocks, particularly those which occur as xenoliths in kimberlite, in serpentinized peridotites, and in eclogites. Almandine and spessartine are stable at much lower pressures and hence occur in crustal rocks. In particular, Fe^2-rich garnets are present in many gneissic and schistose metamorphic rocks formed by the recrystallization of argillaceous sediments at depths greater than about 6 mi (10 km). Such garnets typically contain about 75% almandine together with small amounts of spessartine, pyrope, and grossular. During progressive metamorphism of argillaceous rocks, garnet appears at the expense of chlorite when temperatures reach about 930°F (500°C). Because of the large, readily identifiable crystals which form, the first appearance of garnet (at the garnet "isograd") is commonly used by geologists as an index of the intensity, or grade, of metamorphism. Spessartine-rich garnets occur in metamorphosed sedimentary manganese deposits and in some granite pegmatites and acid igneous rocks. In the latter case the garnets may also be rich in almandine component.

Grossularite and andradite are typical minerals of the low-pressure metamorphism of impure calcareous rocks. Solid solutions between these end members are also common constituents of metasomatic skarns, where they are typically associated with magnetite and iron-bearing pyroxenes. Uvarovite-rich garnets are rare, occurring chiefly in serpentinites, where they are typically associated with chromite.

Garnets are strongly resistant to weathering and alteration and are hence widespread constituents of sands and sediments in areas of garnetiferous primary rocks. *See* METAMORPHIC ROCKS; SILICATE MINERALS.

B. J. Wood

Bibliography. L. G. Berry and B. Mason, *Mineralogy*, 2d ed., 1983; C. Klein and C. S. Hurlbut, *Dana's Manual of Mineralogy*, 21st ed., 1993, revised 1998; R. C. Newton, T. V. Charlu, and O. J. Kleppa, Thermochemistry of high pressure garnets and clinopyroxenes in the system $CaO\text{-}MgO\text{-}Al_2O_3\text{-}SiO_2$, *Geochim. Cosmochim. Acta*, 41:369–377, 1977.

Gas

A phase of matter characterized by relatively low density, high fluidity, and lack of rigidity. A gas expands readily to fill any containing vessel. Usually a small change of pressure or temperature produces a large change in the volume of the gas. The equation of state describes the relation between the pressure, volume, and temperature of the gas. In contrast to a crystal, the molecules in a gas have no long-range order.

At sufficiently high temperatures and sufficiently low pressures, all substances obey the ideal gas, or perfect gas, equation of state, shown as Eq. (1),

$$p\overline{V} = RT \tag{1}$$

where p is the pressure, T is the absolute temperature, $\overline{V}$ is the molar volume, and R is the gas constant. Absolute temperature T expressed on the Kelvin scale is related to temperature t expressed on the Celsius scale as in Eq. (2).

$$T = t + 273.16 \tag{2}$$

The gas constant is

$$\begin{aligned} R &= 82.0567 \text{ cm}^3\text{-atm/(mole)(K)} \\ &= 82.0544 \text{ ml-atm/(mole)(K)} \end{aligned}$$

The molar volume is the molecular weight divided by the gas density.

Empirical equations of state. At lower temperatures and higher pressures, the equation of state of a real gas deviates from that of a perfect gas. Various empirical relations have been proposed to explain the behavior of real gases. The equations of J. van der Waals (1899), Eq. (3), of P. E. M. Berthelot (1907), Eq. (4), and F. Dieterici (1899), Eq. (5), are frequently used.

$$\left(p + \frac{a}{\overline{V}^2}\right)(\overline{V} - b) = RT \tag{3}$$

$$\left(p + \frac{a}{T\overline{V}_2}\right)(\overline{V} - b) = RT \tag{4}$$

$$pe^{a/\overline{V}RT}(\overline{V} - b) = RT \tag{5}$$

In these equations, a and b are constants characteristic of the particular substance under considerations. In a qualitative sense, b is the excluded volume due to the finite size of the molecules and roughly equal to four times the volume of 1 mole of molecules. The constant a represents the effect of the forces of attraction between the molecules. In particular, the internal energy of a van der Waals gas is $-a\overline{V}_i$. None of these relations gives a good representation of the compressibility of real gases over a wide range of temperature and pressure. However, they reproduce qualitatively the leading features of experimental pressure-volume-temperature surfaces.

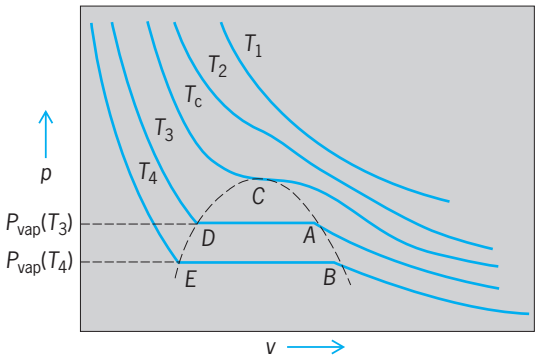

Fig. 1. Schematic isotherms of a real gas. *C* is the critical point. Points *A* and *B* give the volume of gas in equilibrium with the liquid phase at their respective vapor pressures. Similarly, *D* and *E* are the volumes of liquid in equilibrium with the gas phase.

Schematic isotherms of a real gas, or curves showing the pressure as a function of the volume for fixed values of the temperature, are shown in **Fig. 1**. Here T_1 is a very high temperature and its isotherm deviates only slightly from that of a perfect gas; T_2 is a somewhat lower temperature where the deviations from the perfect gas equation are quite large; and T_c is the critical temperature. The critical temperature is the highest temperature at which a liquid can exist. That is, at temperatures equal to or greater than the critical temperature, the gas phase is the only phase that can exist (at equilibrium) regardless of the pressure. Along the isotherm for T_c lies the critical point, C, which is characterized by zero first and second partial derivatives of the pressure with respect to the volume. This is expressed as Eq. (6).

$$\left(\frac{\partial p}{\partial \overline{V}}\right)_c = \left(\frac{\partial^2 p}{\partial \overline{V}}\right)_c = 0 \qquad (6)$$

At temperatures lower than the critical, such as T_3 or T_4, the equilibrium isotherms have a discontinuous slope at the vapor pressure. At pressures less than the vapor pressure, the substance is gaseous; at pressures greater than the vapor pressure, the substance is liquid; at the vapor pressure, the gas and liquid phases (separated by an interface) exist in equilibrium.

Along one of the isotherms of the empirical equations of state discussed above, the first and second derivatives of the pressure with respect to the volume are zero. The location of this critical point in terms of the constants a and b is shown below; p_c and $\overline{V}_c$ are the pressure and volume at the critical temperature.

	van der Waals	*Berthelot*	*Dieterici*
p_c	$\dfrac{a}{27b^2}$	$\left(\dfrac{aR}{216b^3}\right)^{1/2}$	$\dfrac{a}{4e^2b^2}$
$\overline{V}_c$	$3b$	$3b$	$2b$
T_c	$\dfrac{8a}{27Rb}$	$\left(\dfrac{8a}{27Rb}\right)^{1/2}$	$\dfrac{a}{4Rb}$
$\dfrac{p_c\overline{V}_c}{RT_c}$	0.3750	0.3750	0.2706

Some typical values of $p_c\overline{V}_c/RT_c$ for real gases are as follows: 0.30 for the noble gases, 0.27 for most of the hydrocarbons, 0.243 for ammonia, and 0.232 for water. The van der Waals and Berthelot equations of state, Eqs. (3) and (4), cannot quantitatively reproduce the critical behavior of real gases because no substance has a value of $p_c\overline{V}_c/RT_c$ as large as 0.375. The Dieterici equation, Eq. (5), gives a good representation of the critical region for the light hydrocarbons but does not represent well the noble gases or water.

At temperatures lower than the critical point, the analytical equations of state, such as the van der Waals, Berthelot, or Dieterici equations, give S-shaped isotherms as shown in **Fig. 2**. From thermodynamic considerations, the vapor pressure is determined by the requirement that the cross-hatched area DEO be equal to the cross-hatched area AOB. Under equilibrium conditions, the portion of this isotherm lying between A and D cannot occur. However, if a gas is suddenly compressed, points along the segment AB may be realized for a short period until enough condensation nuclei form to create the liquid phase. Similarly, if a liquid is suddenly overexpanded, points along DE may occur for a short time. For low temperatures, the point E may represent a negative pressure corresponding to the tensile strength of the liquid. However, the simple analytical equations of state cannot be used for a quantitative estimate of these transient phenomena. Actually, it is easy to show that the van der Waals, Berthelot, and Dieterici equations give poor representations of the liquid phase since the volume of most liquids (near their freezing point) is considerably less than the constant b.

Principle of corresponding states. In the early studies, it was observed that the equations of state of many substances are qualitatively similar and can be correlated by a simple scaling of the variables. To describe this result, the reduced or dimensionless

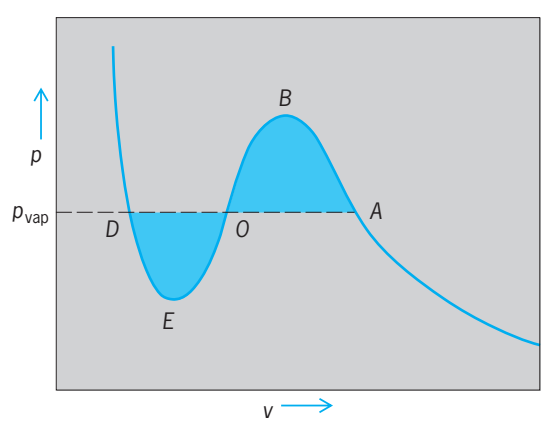

Fig. 2. Schematic low-temperature isotherm as given by van der Waals, Berthelot, or Dieterici equations of state. Here the line *DOA* corresponds to the vapor pressure. The point *A* gives the volume of the gas in equilibrium with the liquid phase, and *D* gives the volume of the liquid. The segment of the curve *DE* represents overexpansion of the liquid. The segment *AB* corresponds to supersaturation of the vapor. However, the segment *EOB* could not be attained experimentally.

variables, indicated by a subscript r, are defined by dividing each variable by its value at the critical point. These variables are given in Eqs. (7)–(9).

$$p_r = \frac{p}{p_c} \tag{7}$$

$$T_r = \frac{T}{T_c} \tag{8}$$

$$\overline{V}_r = \frac{\overline{V}}{\overline{V}_c} \tag{9}$$

In its most elementary form, the principle of corresponding states asserts that the reduced pressure, p_r, is the same function of the reduced volume and temperature, v_r and T_r, for all substances. An immediate consequence of this statement is the statement that the compressibility factor, expressed as Eq. (10),

$$z = \frac{p\overline{V}}{RT} \tag{10}$$

is a universal function of the reduced pressure and the reduced temperature. This principle is the basis of the generalized compressibility chart shown in **Fig. 3**. This chart was derived from data on the equation-of-state behavior of a number of common gases.

It follows directly from the principle of corresponding states that the compressibility factor at the critical point z_c should be a universal constant. It is found experimentally that this constant varies somewhat from one substance to another. On this account, empirical tables were developed of the compressibility factor and other thermodynamic properties of gases as functions of the reduced pressure and reduced temperature for a range of values of z_c. Such generalized corresponding-states treatments are very useful in predicting the behavior of a substance on the basis of scant experimental data.

Theoretical considerations. The equation-of-state behavior of a substance is closely related to the manner in which the constituent molecules interact. Through statistical mechanical considerations, it is possible to obtain some information about this relationship. If the molecules are spherically symmetrical, the force acting between a pair of molecules depends only on r, the distance between them. It is then convenient to describe this interaction by means of the intermolecular potential $\varphi(r)$ defined so that the force is the negative of the derivative of $\varphi(r)$ with respect to r. *See* INTERMOLECULAR FORCES.

Two theoretical approaches to the equation of state have been developed. In one of these approaches, the pressure is expressed in terms of the partition function Z and the total volume V of the container in the manner of Eq. (11). Here k is

$$p = kT\frac{\partial \ln Z}{\partial V} \tag{11}$$

the Boltzmann constant or the gas constant divided by the Avogadro number N_0, $k = R/N_0$. For a gas

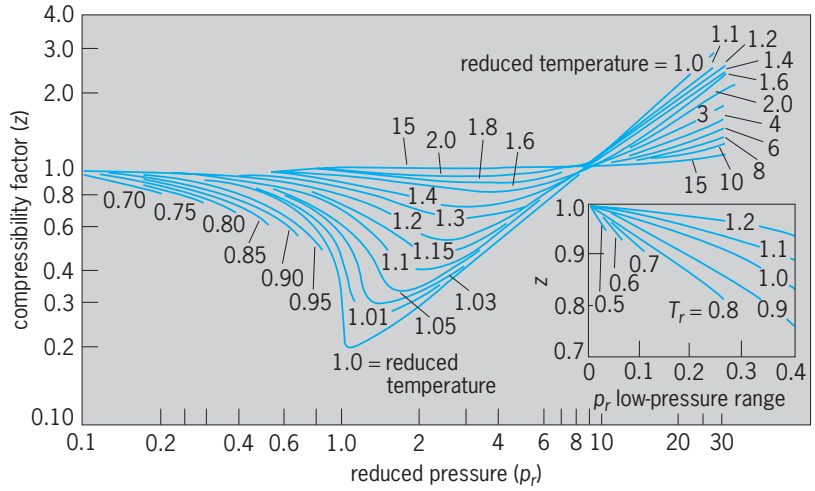

Fig. 3. The compressibility factor $p\overline{V}_i/RT$ as a function of the reduced pressure $p_r = p/p_c$ and reduced temperature $T_r = T/T_c$. (*After O. A. Hougen, K. M. Watson, and R. A. Ragatz, Chemical Process Principles, pt. 2, Wiley, 1959*)

made up of spherical molecules or atoms with no internal structure, the partition function is given as Eq. (12) In this expression, φ_{ij} is the energy of in-

$$Z = \frac{1}{N!}\left(\frac{2\pi mkT}{h^2}\right)^{3N/2} \times \int \exp\left(-\sum_{i>j}\frac{\varphi_{ij}}{kT}\right) d\overline{V}_1 d\overline{V}_2 \cdots d\overline{V}_N \tag{12}$$

teraction of molecules i and j and the summation is over all pairs of molecules, h is Planck's constant, N is the total number of molecules, and the integration is over the three cartesian coordinates of each of the N molecules. The expression for the partition function may easily be generalized to include the effects of the structure of the molecules and the effects of quantum mechanics.

In another theoretical approach to the equation of state, the pressure may be written as Eq. (13), where

$$p = \frac{NkT}{V} - \frac{2\pi N^2}{3V^2}\int g(r)\frac{d\varphi}{dr}r^3\,dr \tag{13}$$

$g(r)$ is the radial distribution function. This function is defined by the statement that $2\pi(N_2/V)g(r)r_2dr$ is the number of pairs of molecules in the gas for which the separation distance lies between r and $r + dr$. The radial distribution function may be determined experimentally by the scattering of x-rays. Theoretical expressions for $g(r)$ are being developed.

The compressibility factor $z = pV/NkT$ may be considered as a function of the temperature, T, and the molar volume, $\overline{V}$. In the virial form of the equation of state, z is expressed as a series expansion in inverse powers of $\overline{V}$, as in Eq. (14). The coefficients

$$z = 1 + \frac{B(T)}{\overline{V}} + \frac{C(T)}{\overline{V}^3} + \cdots \tag{14}$$

$B(T)$, $C(T)$, . . . , which are functions of the temperature, are referred to as the second, third, . . . , virial coefficients. This expansion is an important method

of representing the deviations from ideal gas behavior. From statistical mechanics, the virial coefficients can be expressed in terms of the intermolecular potential. In particular, the second virial coefficient is Eq. (15). If the intermolecular potential is known,

$$B(T) = 2\pi N_0 \int (1 - e^{-\varphi/kT})\, r^2\, dr \qquad (15)$$

Eq. (15) provides a convenient method of predicting the first-order deviation of the gas from perfect gas behavior. The relation has often been used in the reverse manner to obtain information about the intermolecular potential. Often $\varphi(r)$ is expressed in the Lennard-Jones (6-12) form, Eq. (16), where ϵ and σ

$$\varphi(r) = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] \qquad (16)$$

are constants characteristic of a particular substance. Values of these constants for many substances have been tabulated. In terms of these constants, the second virial coefficient has the form of Eq. (17), where

$$B(T) = \frac{2}{3}\pi N_0 \sigma^3 B^* \frac{kT}{\epsilon} \qquad (17)$$

$B^*(kT/\epsilon)$ is a universal function. If all substances obeyed this Lennard-Jones (6-12) potential, the simple form of the law of corresponding states would be rigorously correct. *See* CHEMICAL THERMODYNAMICS; KINETIC THEORY OF MATTER.

C. F. Curtiss; J. O. Hirschfelder

Bibliography. I. B. Cohen (ed.), *Laws of Gases*, 1981; J. O. Hirschfelder et al., *Molecular Theory of Gases and Liquids*, 1964; R. Holub and P. Vonka, *The Chemical Equilibrium of Gaseous Systems*, 1976; I. N. Levine, *Physical Chemistry*, 4th ed., 1995; R. Mohilla and B. Ferencz, *Chemical Process Dynamics*, 1982.

Gas absorption operations

The separation of solute gases from gaseous mixtures of noncondensables by transfer into a liquid solvent. This recovery is achieved by contacting the gas stream with a liquid that offers specific or selective solubility for the solute gas or gases to be recovered. The operation of absorption is applied in industry to purify process streams or recover valuable components of the stream. It is used extensively to remove toxic or noxious components (pollutants) from effluent gas streams.

Examples of application of gas absorption are the removal of hydrogen sulfide and mercaptans from natural gas, the recovery of carbon monoxide in petrochemical syntheses, the removal of sulfur oxides from power plant stack gases, and the recovery of silicon tetrafluoride and hydrogen fluoride from fertilizer production stack gases.

The absorption process requires the following steps: (1) diffusion of the solute gas molecules through the host gas to the liquid boundary layer based on a concentration gradient, (2) solvation of the solute gas in the host liquid based on gas-liquid solubility, and (3) diffusion of the solute gas based on concentration gradient, thus depleting the liquid boundary layer and permitting further solvation. The removal of the solute gas from the boundary layer is often accomplished by adding neutralizing agents to the host liquid to change the molecular form of the solute gas. This process is called absorption accompanied by chemical reaction.

Factors affecting the kinetics of the gas absorption process

Zone	Factor
Gas	Gas velocity
	Molecular weights and sizes of host and solute gases
	Temperature
	Concentration gradient
Solvation	Solubility in host liquid
Liquid	Molecular weights and sizes of host liquid and solute gas
	Temperature
	Viscosity of liquid
	Liquid surface velocity
	Surface renewal rate
	Concentration gradient of solute gas
	Concentration gradient of neutralizing reagent

The factors affecting the rate of absorption or the kinetics of transfer from the host gas to the host liquid are indicated in the **table**.

Historically, the kinetics of gas absorpion have been represented by statistical average overall coefficients in the form of Eq. (1), where N_A is the rate of

$$N_A = K_{G\alpha} V (P - P^*) \qquad (1)$$

mass-transfer quantity of solute gas transferred per

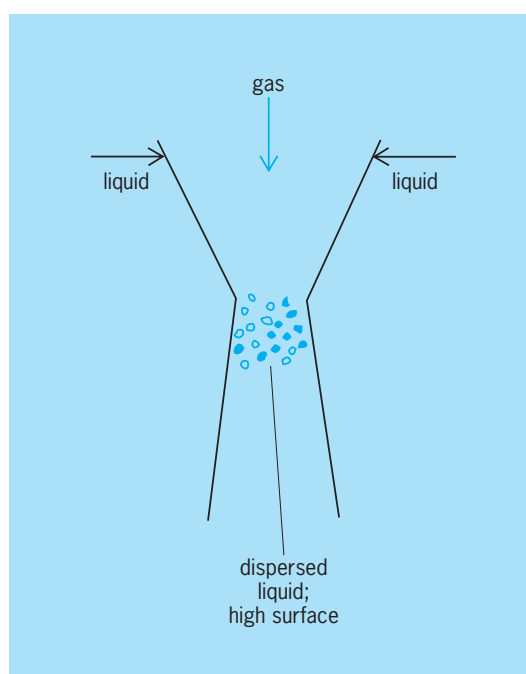

Fig. 1. Operation of a venturi system.

unit time; $K_{G\alpha}$ is the mass-transfer coefficient based on partial pressure gradient from the host-liquid bulk concentrations, quantity per unit time per volume of the transfer system per unit partial pressure gradient; V is volume of the transfer system; P is partial pressure of the solute gas in the bulk of the host gas; and P^* is partial pressure of the host gas in equilibrium with the bulk of the host liquid.

An alternative method of sizing the system is the transfer unit concept expressed in Eqs. (2)–(4),

$$N_{oG} = \int \frac{dp}{P - f(P^*)} \frac{1 - f(P_f)}{1 - P} \tag{2}$$

$$H_{oG} = \frac{Gm}{K_{G\alpha\pi}} \frac{1 - f(P_f)}{1 - P} \tag{3}$$

$$H = N_{oG} H_{oG} \tag{4}$$

where H is height of the absorber; N_{oG} is number of transfer units; H_{oG} is height of a transfer unit; $K_{G\alpha}$, P, and P^* are as in Eq. (1) ; P_f is partial pressure of the solute gas in the interfacial film between gas and liquid; π is total pressure of operation; and Gm is molar gas flow rate.

The effect of the relative resistances in the liquid and gas phases are established by the relationship shown in Eq. (5), where $K_{G\alpha}$ is as in Eq. (1); $k_{G\alpha}$ is

$$K_{G\alpha} = \frac{1}{\dfrac{1}{k_{G\alpha}} + \dfrac{m}{k_{L\alpha}}} \tag{5}$$

mass-transfer coefficient in the gas phase only, quantity per unit per volume of transfer system per unit partial pressure gradient; $k_{L\alpha}$ is mass-transfer coefficient in liquid phase only, quantity per unit time per unit volume per unit concentration gradient; and m is partial pressure–liquid phase concentration relationship of solute gas.

The gas-phase and liquid-phase mass-transfer coefficients are established as Eqs. (6) and (7), where D_G

$$k_G = \frac{D_G \pi}{RT \chi P_{BM}} \tag{6}$$

$$k_L = Z \left(\frac{D_L \phi}{\pi t} \right)^{1/2} \tag{7}$$

and D_L are diffusivities in the gas and liquid phases, respectively; R is the universal gas constant; T is the absolute temperature; χ is the effective thickness of the gas boundary layer; P_{BM} is the partial pressure of the inert gases; ϕ is the surface renewal rate; and t is time of contact.

It has been established that except for unique situations of very high solubility, the transfer in the liquid phase tends to become the controlling factor in rate of absorption.

The emphasis in design of systems has been to increase the surface renewal rate and to increase the effective liquid diffusivity by neutralizing the solute gas in the liquid phase as rapidly as possible. In the case of neutralization or total removal, the mass-transfer coefficient in the liquid phase follows the relationship shown in Eq. (8), where k'_L is liquid-phase mass-

$$\frac{k'_L}{k_L} = 1 + \left(\frac{q}{C_{ai}} \frac{D_B}{D_L} \right)^n \tag{8}$$

transfer coefficient when chemical reaction occurs; q is concentration of unreacted reagent in the bulk of the host liquid; C_{ai} is concentration of solute gas in the liquid phase (physically absorbed) in equilibrium with P; D_B and D_L are diffusivities of the solute

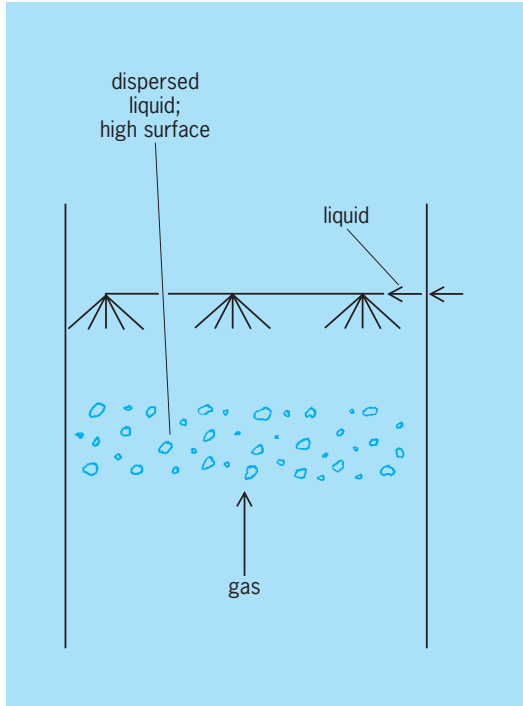

Fig. 2. Operation of a spray tower.

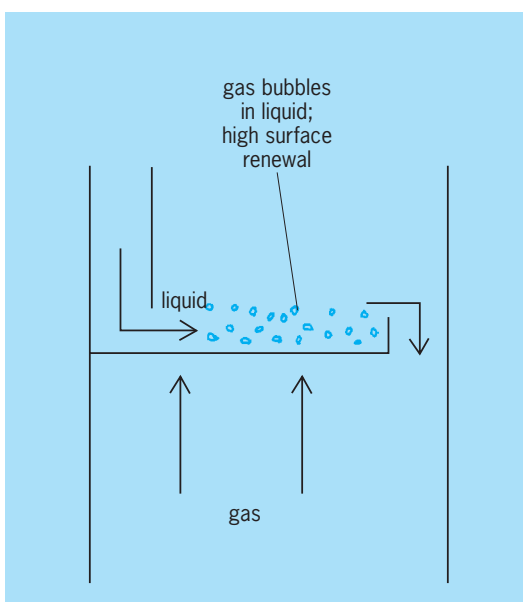

Fig. 3. Operation of a tray tower.

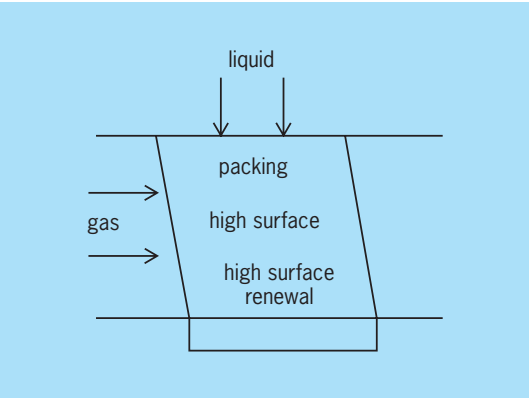

Fig. 4. Operation of a packed tower.

gas and reagent in the host liquid, respectively; and n is an exponent, generally less than 1. The transfer systems require effective contact of liquid and gas, with maximum surface contact and surface renewal, and minimum consumption of energy. The types of equipment used for this service are venturis (high surface, no surface renewal); spray towers (high surface, no surface renewal); tray towers (high surface renewal); and packed towers (high surface, high surface renewal).

The venturi system (**Fig. 1**), requiring high energy input, is often used where large quantities of particulates are present in the gas or the scrubbing liquid has large concentrations of suspended solids. The spray tower (**Fig. 2**) is used where the transfer requirements are low. The tray tower (**Fig. 3**) is not used extensively because of poor efficiency in low-concentration gas systems. The packed tower (**Fig. 4**) is the predominant system used in gas absorption because of its inherent differential mechanism, low power consumption, and capability for handling low concentrations of particulates in gas streams. It is often used in conjunction with a solids recovery unit such as a venturi.

The packed column volume generally is randomly filled with a commercially available geometric shape, ranging from 1 to 4 in. (2.5 to 10 cm) in the largest dimension, forms that provide either large surfaces or capability for surface renewal, or both.

The gas absorption equipment, usually called scrubbers, generally use water or solutions of reagents in the water as the host liquid, although hydrocarbon liquids are used for specific applications in the chemical industry. The efficiency of absorption can be uniquely high, especially in the case of absorption accompanied by chemical reaction. Reduction of the solute gas concentration to approximately 1–5 ppm is often required for air-pollution control. These effluent concentrations have been achieved in air-pollution control in the steel, fertilizer, glass, and pulp and paper industries. The gas absorption equipment range in capacity from 100 cubic feet per minute (4.7×10^{-2} m^3/s) to 2,000,000 cfm (940 m^3/s). *See* ADSORPTION; DISTILLATION; HUMIDIFICATION.

Aaron J. Teller

Gas and atmosphere analysis

Qualitative identifications and quantitative determinations of substances essential for the evaluation of the air quality in the ambient air and in the industrial workplace.

The terms air pollution and gas analysis are used here in the broad sense. Air pollution refers to unwanted air contaminants in both the ambient air and the work environment, since many of the same pollutants are found at both places and are assessed using similar procedures. These pollutants may exist in the gaseous state or as aerosols. Aerosols are liquid droplets or solid particles small enough (0.01 to 100 micrometers) to remain dispersed in the air for extensive periods of time. Gas analysis refers to the analysis of both gases and vapors. The term vapor is used for the gaseous phase of substances that are liquids or solids at standard temperature and pressure. Thus the gaseous states of gasoline, mercury, and water are examples of vapors, whereas methane, hydrogen, and ozone are gases. Many important pollutants are vapors that have arisen from the volatilization or decomposition of organic materials. *See* AEROSOL; AIR POLLUTION; INDUSTRIAL HEALTH AND SAFETY.

Qualitative Identification

The qualitative identification of air pollutants can be extremely complex, and may require the use of several instruments which provide complementary information about composition and structure. Since the entire sample is often limited to milligram or microgram quantities, the classical identification methods, such as boiling point and refractive index determinations, functional group tests, combustion analyses, and derivative preparations, have been largely replaced by instrumental methods. Information for identification purposes is now generally obtained from instruments such as mass, nuclear magnetic resonance, infrared, and ultraviolet spectrometers that rely upon the response of a molecule to an energy probe.

Organic compounds. These are identified by means of mass spectroscopy and gas chromatography.

Mass spectroscopy. This is probably the single most powerful technique for the qualitative identification of volatile organic compounds, and has been particularly useful in the identification of many environmental contaminants. When a sample is introduced into the mass spectrometer, electron bombardment causes the parent molecule to lose an electron and form a positive ion. Some of the parent ions are also fragmented into characteristic daughter ions, while other ions remain intact. All of the ions are accelerated, separated, and focused on an ion detector by means of either a magnetic field or a quadrupole mass analyzer. Using microgram quantities of pure materials, the mass spectrometer yields information about the molecular weight and the presence of other atoms, such as nitrogen, oxygen, and halogens, within the molecule. In addition, the fragmentation pattern often provides a unique so-called fingerprint

of a molecule, allowing positive identification. If the gas is a mixture, interpretation of the mass spectral data is difficult since the fragmentation patterns are superimposed. However, interfacing the mass spectrometer to a gas chromatograph provides an elegant solution to this problem. *See* MASS SPECTROMETRY.

Gas chromatography. A gas chromatograph is essentially a highly efficient apparatus for separating a complex mixture into individual components. When a mixture of components is injected into a gas chromatograph equipped with an appropriate column, the components travel down the column at different rates and therefore reach the end of the column at different times. The mass spectrometer located at the end of the column can then analyze each component separately as it leaves the column. In essence, the gas chromatograph allows the mass spectrometer to analyze a complex mixture as a series of pure components. More than 100 compounds have been identified and quantified in automobile exhaust by using a gas chromatograph–mass spectrometer combination. *See* GAS CHROMATOGRAPHY.

Metals. A powerful technique for identifying metals is inductively coupled plasma spectrometry. An inductively coupled plasma spectrometer is an emission spectrometer capable of operating at 6000–8000 K (10,000–14,000°F). Emission spectrometers raise the temperature of the sample to a sufficiently high temperature that chemical bonds are broken and the electrons are raised from their stable energy levels to higher energy levels. The excited atoms lose a part of the excess energy by emitting a characteristic wavelength of light. Sensors placed around the plasma can identify and quantify 35 metals in a mixture. *See* EMISSION SPECTROCHEMICAL ANALYSIS.

Quantitative Analysis

Once a qualitative identification of an important pollutant has been established, further interest often centers on quantifying the levels of the pollutant as a function of time at various sites.

Methods. The methods employed chiefly for quantification can be classified for convenience into direct and indirect procedures. Direct-reading instruments are generally portable and may analyze and display their results in a few seconds or minutes, and can operate in a continuous or semicontinuous mode. Indirect methods are those involving collection and storage of a sample for subsequent analysis. Both direct and indirect methods have inherent advantages and disadvantages. By using indirect methods, samples with several pollutants can be simultaneously collected from a number of different sites with relatively inexpensive collection devices and later analyzed at a central laboratory. On the other hand, direct methods may require one instrument for each pollutant at each sampling site, and thus may become prohibitively expensive. However, reduction of the delay before the results are available may be the basis for selecting a direct over an indirect method, or the pollutant in question may not be stable under the conditions of storage. For example, if a worker needs to enter an enclosure with an atmosphere potentially hazardous to life because of an oxygen deficiency, the presence of an explosive mixture, or the presence of a high concentration of a toxic gas, it is essential to have the analytical results available in a few minutes.

Direct methods. These consist of methods utilizing colorimetric indicating devices and instrumental methods.

Colorimetric indicators. Three types of direct-reading colorimetric indicators have been utilized: liquid reagents, chemically treated papers, and glass tubes containing solid chemicals (detector tubes). The simplest of these methods is the detector tube. Detector tubes are constructed by filling a glass tube with silica gel coated with color-forming chemicals. For use, the ends of the sealed tube are broken and a specific volume of air, typically 6 in.^{3} (100 cm^3), is drawn through the tube at a controlled rate. Detector tubes often utilize the same color-forming chemical to detect several different gases, and therefore may be nonspecific for mixtures of these gases. Temperature, humidity, age, and uniformity of packing also influence the performance. Detector tubes for analyzing approximately 400 different gases are commercially available. Accuracy is sometimes low, and detector tubes for only 25 gases meet the National Institute for Occupational Safety and Health (NIOSH) accuracy requirement of ±25%. For some gases, semicontinuous analyzers have been developed that operate by pulling a fixed volume of air through a paper tape impregnated with a color-forming reagent. The intensity of the color is then measured for quantification. Phosgene, arsine, hydrogen sulfide, nitric oxide, chlorine, and toluene diisocyanate have been analyzed by indicating tapes. *See* COLORIMETRY.

Direct-reading instruments. With the availability of stable and sensitive electronics, direct-reading instruments capable of measuring gases directly at the parts-per-billion range were developed. Most direct-reading instruments contain a sampling system, electronics for processing signals, a portable power supply, a display system, and a detector. The detector or sensor is a component that is capable of converting some characteristic property of the gas into an electrical signal. While there are dozens of properties for the bases of operation of these detectors, the most sensitive and popular detectors are based on electrical or thermal conductivity, ultraviolet or infrared absorption, mass spectrometry, electron capture, flame ionization, flame photometry, heat of combustion, and chemiluminescence. Many of these detectors respond to the presence of 10^{-9} g quantities, and even to 10^{-12} g levels. In addition to improved accuracy, precision, and analysis time, another advantage is that most instruments produce an electrical signal which can be fed into a computer for process control, averaging, and record keeping. Rapid fluctuations and hourly, daily, and yearly averages are readily obtained. Several important instruments are based on the following principles.

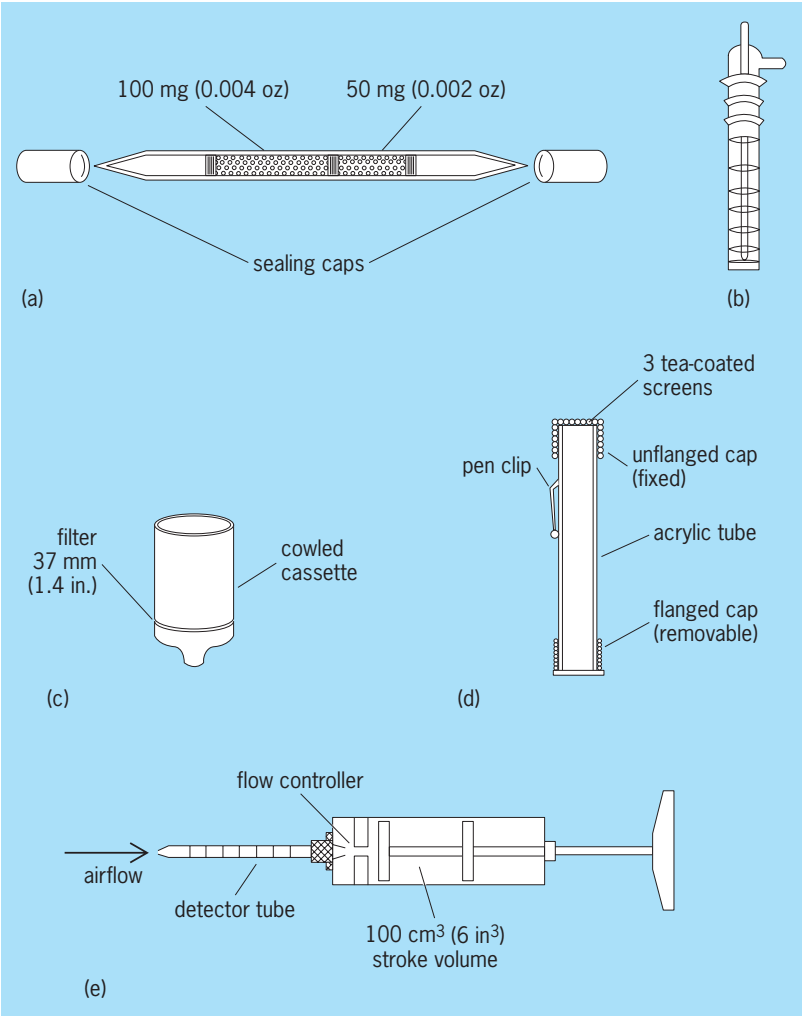

Fig. 1. Collection devices for monitoring air in the workplace. (*a*) Solid sorbent (charcoal) sampling tube. (*b*) Midget impinger. (*c*) Cowled cassette and its filter. (*d*) Passive monitor, used for nitrogen dioxide. (*e*) Piston-type detector tube pump.

1. *Heat of combustion*. Many portable direct-reading meters for explosive atmospheres are based on the principle of catalytic or controlled combustion on a heated filament. The filament is usually one arm of a Wheatstone bridge circuit. The resulting heat of combustion changes the resistance of the filament (usually platinum), and the resulting imbalance is related to the concentration of the gas. A meter displays the results. This method is nonspecific, and gives an indication of all combustible gases present in the range from about 100 parts per million (ppm) to a few percent.

2. *Chemiluminescence*. The phenomenon of chemiluminescence is employed for the determination of levels of ozone, oxides of nitrogen, and sulfur compounds. Chemiluminescence is an emissive process that occurs when all or part of the energy of a chemical reaction is released as light rather than heat. A familiar example is the "cold light" of fireflies. Ozone levels in the range from 1 to 100 ppb can be determined by measuring the emission at 585 nanometers, which occurs when ozone is mixed with excess ethylene. Similarly, nitric oxide (NO) levels from 10 ppb to 5000 ppm can be measured by a chemiluminescence method. The analysis of nitrogen dioxide (NO_2) is performed by reducing NO_2 to NO in a catalytic converter and measuring the NO emission. Mixtures of NO and NO_2 can be analyzed by measuring the NO level by chemiluminescence, reducing the NO_2 to NO, and again measuring the chemiluminescence level. The NO_2 level is obtained by difference. Measurement of the chemiluminescence produced by thermal means in a hydrogen-rich flame allows the detection of sulfur compounds from 1 to 1000 parts per billion (ppb). With the use of appropriate light filters, interferences by other gases can be reduced.

3. *Infrared absorption*. Measurement of carbon monoxide utilizes a variety of other principles, including infrared absorption at a fixed wavelength, nondispersive infrared absorption, measurement of the heat of reaction in oxidizing carbon monoxide to carbon dioxide, and gas chromatography.

4. *Gas chromatography*. The gas chromatograph is widely used both for the analysis of collected samples and as a semicontinuous direct-reading instrument. The availability of many different detectors capable of measuring the effluent from the gas-chromatograph column and the development of valves for automatic injection of samples and for directing sample flow have further extended the versatility of this instrument. Accessories are available that can collect samples from about 12 different distant sites through sampling lines for sequential injection directly into a gas chromatograph for analysis. Another application involves the use of valves to inject an air sample into a gas chromatograph equipped with a flame ionization detector to measure the methane content. A second air sample is injected directly into the detector to measure the total hydrocarbon content. The methane signal is subtracted from the total hydrocarbon content to provide an indication of the smog potential.

5. *Light scattering*. Portable dust monitors using light-scattering principles can directly measure concentrations to 0.01 mg/m^3. The sampling period can be varied but is typically about 1 min. The results are displayed as the average value for the sampling period. Some instruments can distinguish between dust particles and fibers. These have had some success in measuring airborne asbestos fibers.

6. *Electrochemical sensors*. Gases such as oxygen, carbon monoxide, carbon dioxide, sulfur dioxide, and hydrogen sulfide can be measured by using electrochemical sensors. Usually a separate sensor is required for each gas, but several can be combined into a single instrument.

7. *Ionization detectors*. Flame ionization detectors and photoionization detectors are capable of measuring many organic compounds in the parts-per-million range. These detectors produce ions, and a current is produced in proportion to the type of molecule and the number of the molecules. Flame ionization detectors use a hydrogen flame to produce the ions, and the photoionization detectors use ultraviolet light to produce ions. *See* CHEMILUMINESCENCE; LIGHT-SCATTERING PHOTOMETRY.

Indirect methods. For indirect methods, the main collection devices are freeze traps, bubblers, evacuated bulbs, plastic bags, and solid sorbents. Because of their convenience, solid sorbents dominate collection procedures. NIOSH developed a versatile method for industrially important vapors, based on the sorption of the vapors on activated charcoal or, to a lesser extent, on other solid sorbents such as silica gel and porous polymers. Typically, in this technique a few liters of air are pulled through a glass tube containing about 0.004 oz (100 mg) of charcoal. The charcoal tube is only 7 cm × 6 mm (3 in. × 0.2 in.), and has the advantage that it can be placed on the worker's lapel. A battery-operated pump small enough to fit into a shirt pocket is connected by a plastic tube to the collecting device, so that the contaminants are continuously collected from the breathing zone of the worker. Many solvent vapors and gases are efficiently trapped and held on the charcoal. The ends of the sample tube are then capped, and the tube is returned to a laboratory for analysis. In the laboratory the tube is broken open, and the charcoal poured into carbon disulfide to desorb the trapped vapors. Following desorption, a sample of the solution is injected into a gas chromatograph for quantification.

This technique has been highly successful for several classes of compounds, such as aromatics, aliphatics, alcohols, esters, aldehydes, and chlorinated compounds. Sulfur- and nitrogen-containing compounds can also be analyzed by using a gas chromatograph which is equipped with a sulfur- or nitrogen-sensitive detector.

Applications. The major applications of gas and atmosphere analysis involve industrial hygiene, indoor air quality, hazardous waste sites and spills, and sampling of ambient air and stationary sources.

Industrial hygiene. The Occupational Safety and Health Administration (OSHA) has established exposure standards for about 600 substances. Because a number of factors affect exposure, sampling at the breathing zone of the worker is preferred. Filters, sorbent tubes, impingers, and passive monitors are used to collect the samples (**Fig. 1**). Both OSHA and NIOSH have developed a number of sampling and analytical methods (**Table 1**).

Indoor air quality. Complaints about the indoor air quality became prevalent in the 1970s. NIOSH

TABLE 1. Examples of NIOSH analytical methods

Chemical*	Flow rate, liters/min	Minimum volume, liters	Maximum volume, liters	Analytical technique	Sampler	Accuracy† (±), %	Validation status‡
Ketones (6, for example, acetone)	0.01–0.2	0.5	3	Gas chromatography–flame ionization detector	Coconut shell charcoal tube	6–8	Full
Inorganic acids (6, for example, hydrofluoric)	0.2–0.5	3	100	Ion chromatography	Silica gel	12–23	Full
Asbestos and other fibers	0.5–16	400	Variable	Phase contrast microscopy	Mixed cellulose ester fiber	46	Full (special criteria)
Metals (26, for example, cadmium)	1–4	Varies ∼5	Varies ∼200	Inductively coupled plasma–atomic emission spectroscopy	Mixed cellulose ester fiber	Varies with metal	Partial
Carbon dioxide	0.02–0.1	0.5	80% of bag capacity	Gas chromatography–thermal conductivity detector	Bag	5.3	Full
Cresol	0.01–0.1	1	24	Gas chromatography–flame ionization detector	Porous polymer	Not determined	Partial
Chlorine	0.3–1	2	90	Ion chromatography–conductivity detector	Silver membrane filter	14.8	Full
Ethylenimine	0.2	1	48	High-performance liquid chromatography–ultraviolet	Bubbler	15.5	Full
Formaldehyde	0.2–1	1	100	Visible absorption spectrophotometry	Impinger	18	Full
Nitroglycerin	0.2–1.0	3	100	Gas chromatography–electron capture detector	Porous polymer	20.3	Full
Nitrogen dioxide	0.025–0.2	1.5	6	Visible absorption spectrophotometry	Treated solid sorbent	14.6	Full
Organophosphorous pesticides	0.2–1	12	Varies ∼240	Gas chromatography–flame photometric detector	Treated solid sorbent for pesticides	Varies ∼18	Full
Particulates, nuisance dust	1–2	7	133	Gravimetric	Polyvinyl chloride filter	11.0	Full
Silica, crystalline, respirable	1.7	400	1000	X-ray diffraction	Cyclone and polyvinyl chloride filter	18	Full
Toluene	Diffusion controlled	(15 min)	(8 h)	Gas chromatography–flame ionization detector	Passive	7.4	Full

*Methods are shown for specific chemicals or families of chemicals. The number of chemicals for which the method is applicable is sometimes given.
†The overall accuracy, including the sampling and analytical error, is expressed as the percent at the 95% confidence level.
‡Validation may be expressed as Full, Partial, or Unrated. Full validation means that the method has been tested with generated samples and is within ±25% of the true value.
SOURCE: *After NIOSH Manual of Analytical Methods*, 4th ed., August 1994.

TABLE 2. EPA sampling and analytical methodologies

Compound	Sampling methodology	Analytical procedure
Total particulate matter	Filtration	Gravimetric
Respirable particulate matter	Cyclones with filtration	Gravimetric
	Cascade impaction with filtration	Gravimetric
Nitrogen oxides	Absorption	Spectrophotometric or ion chromatographic
Sulfur dioxide	Direct analysis absorption	Chemiluminscence Titration
	Direct analysis	Ultraviolet, nondispersive infrared or fluorescence
Carbon monoxide	Flexible or rigid container	Nondispersive infrared or spectrophotometric
Acid gases	Direct analysis absorption	Specific ion electrode, ion chromatography, or colorimetric
Heavy metals	Filtration with absorption	Atomic absorption spectrophotometric
Volatile organics	Flexible/rigid containers, adsorption, absorption, cryogenic sampling, direct analysis	GC/FID, GC/MS GC/PID, GC/ECD, GC/FPD, GC/MS*
Semivolatile organics	Filtration with adsorption	GC/FID, GC/MS GC/PID, GC/ECD, GC/FPD, GC/MS*

*CEM = cellulose ester membrane. GC = gas chromatography. ICAP = inductively coupled argon plasma. FID = flame ionization detector. PID = photoionization detector. MS = mass spectrometry. FPD = flame photometric detector. ECD = electron capture detector.

conducted over 500 investigations and reported the following causes: building materials contamination (4%), microbiological contamination (5%), outside contamination (10%), inside contamination (15%), ventilation inadequate (53%), and unknown (13%). Ventilation accounted for over 50% of the complaints. In addition to checking the ventilation system, NIOSH found that the concentration of carbon dioxide could be used as an indicator of the adequacy of the ventilation. Carbon dioxide is exhaled by humans; although the level of carbon dioxide is not the cause of the complaints, it serves as an indicator of the buildup of all air pollutants emitted by people, such as perfumes and shaving lotions. When the levels of carbon dioxide increase from the outside levels (250–300 ppm) to above 800 ppm, complaints about the indoor air quality begin. At 1200 ppm, many people will complain. *See* AIR POLLUTION.

Ambient air and stationary source sampling. In many sampling situations, regulatory requirements dictate the type of sampling and analytical methodologies required during ambient air and stationary source sampling programs (**Table 2**). However, even in those situations, variables such as sampling location, sampling time, sampling procedures, gas stream characteristics, chemical properties, analytical detection limits, analytical interferences, and test program accuracy and precision requirements must be evaluated to ensure that applicable quality assurance objectives are satisfied. Sampling techniques must provide for representative samples with minimal deterioration and contamination. In addition, the collected sample must be of sufficient quantity to ensure that the concentration is above the detection limit of the analytical method.

The procedures employed for ambient air and source sampling are essentially the same procedures as described for industrial hygiene. The main differences are that larger volumes of the air are sampled and higher flow rates are used. One other major difference that is not encountered in industrial hygiene sampling is isokinetic sampling, which is required for particulate matter in stacks.

Gas sampling may be accomplished with evacuated flasks (**Fig. 2**), bags, canisters, an impinger sampling system, and sorbent tube sampling techniques. This method is limited by the size of

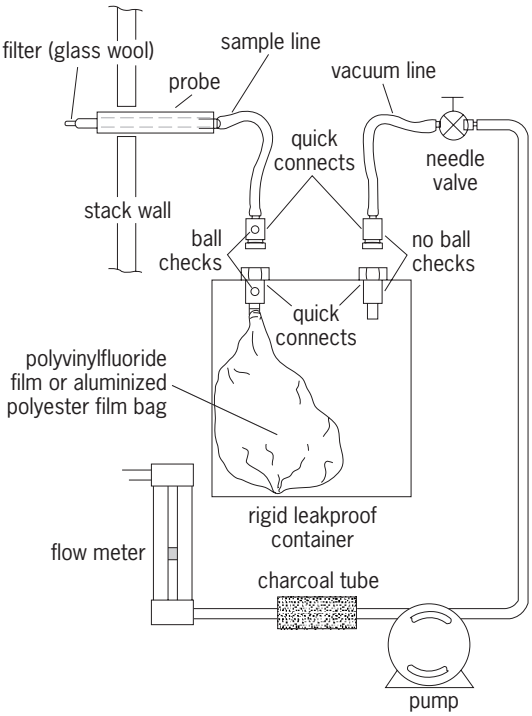

Fig. 2. Evacuated-flask sampling system.

sample collected and possible nonintegration of sample.

William R. Burg; Michael Glowatz, Jr.; Laurence S. Bernson

Bibliography. American Conference of Governmental Industrial Hygienists, *Air Sampling Instruments for Evaluation of Atmospheric Contaminants*, 9th ed., 2001; M. J. Boss and D. W. Day, *Air Sampling and Industrial Hygiene Engineering*, 2000; National Institute for Occupational Safety and Health, *Manual of Analytical Methods*, vols. 1 and 2, DHHS Pub. 84–100, 1993; G. D. Wight, *Fundamentals of Air Sampling*, 1994; E. D. Winegar, *Sampling and Analysis of Airborne Pollutants*, 1993.

Gas chromatography

A method for the separation and analysis of complex mixtures of volatile organic and inorganic compounds. Most compounds with boiling points less than about 250°C (480°F) can be readily analyzed by this technique. A complex mixture is separated into its components by eluting the components from a heated column packed with sorbent by means of a moving-gas phase. A typical chromatogram is shown in **Fig. 1**.

Classification. Gas chromatography may be classified into two major divisions: gas-liquid chromatography, where the sorbent is a nonvolatile liquid called the stationary-liquid phase, coated as a thin layer on an inert, granular solid support, and gas-solid chromatography, where the sorbent is a granular solid of large surface area. The moving-gas phase, called the carrier gas, is an inert gas such as nitrogen or helium which flows through the chromatographic column packed with the sorbent. The solute partitions, or divides, itself between the moving-gas phase and the sorbent and moves through the column at a rate dependent upon its partition coefficient, or solubility, in the liquid phase (gas-liquid chromatography) or upon its adsorption coefficient on the packing (gas-solid chromatography) and the carrier-gas flow rate. Open tubular glass or stainless steel capillary tubes of 0.005–0.02 in. (0.1–0.5 mm) inside diameter and length often as great as 300 ft (90 m), coated on the inside walls with a nonvolatile stationary liquid phase, are also widely used in the separation of complex mixtures.

Gas-solid chromatography is historically the older of the two forms of gas chromatography but was formerly limited primarily to the separation of permanent gases or relatively nonpolar solutes of low molecular weight, whereas gas-liquid chromatography has been the more popular and more versatile approach to the separation of a wide range of higher-molecular-weight compounds, owing to the large choice of liquid phases available. Research in gas-solid absorption chromatography, however, has shown considerable progress in decreasing objectionable peak tailing, thus providing higher separation efficiencies than were formerly possible. It is entirely possible that gas-solid chromatography will achieve greater status in future years. *See* ADSORPTION.

Basic apparatus. The apparatus used in gas chromatography consists of four basic components: a carrier-gas supply and flow controller, a sample inlet system providing a means for introduction of the sample, the chromatographic column and associated column oven, and the detector system. A schematic diagram of a working instrument is shown in **Fig. 2**.

Carrier-gas and flow monitor. Although the carrier gas is most commonly nitrogen or helium, other gases such as carbon dioxide, argon, xenon, and hydrogen are occasionally used. If the detector is a thermal conductivity detector, then helium is the carrier gas of choice. This is because helium has the largest thermal conductivity of all gases with the exception of hydrogen, a flammable and explosive gas. Since the

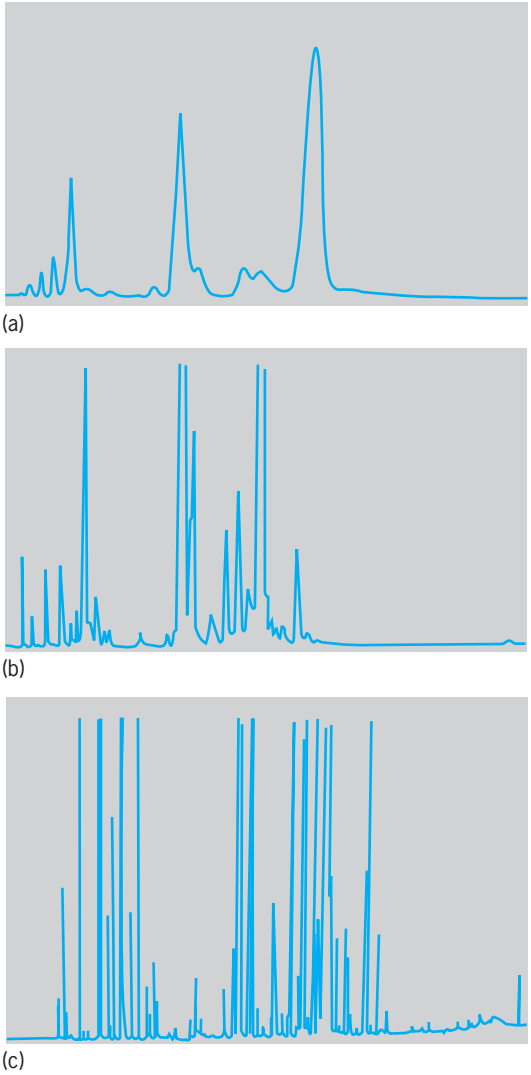

Fig. 1. Three generations in gas chromatography. Peppermint oil separated on (*a*) a 0.25 in. × 6 ft (6 mm × 1.8 m) packed column; (*b*) a 0.03 in. × 500 ft (0.8 mm × 150 m) stainless steel capillary column; and (*c*) a 0.98 in. × 164 ft (0.25 mm × 50 m) glass capillary column. (*After W. Jennings, The uses of glass capillary columns for food and essential oil analysis, J. Chromatog. Sci., 17:636–639, 1979*)

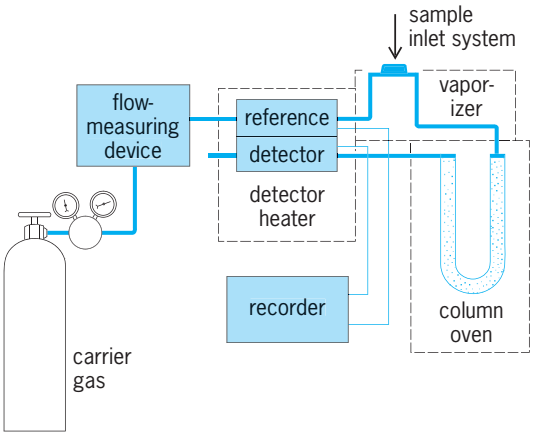

Fig. 2. Basic apparatus used for gas chromatography.

sensitivity of the thermal conductivity detector is a function of the difference in thermal conductivity between the pure carrier gas and that of a mixture of the carrier gas with solute, helium gives high sensitivity. Use of a carrier gas of higher molecular weight will improve column efficiency. Therefore nitrogen or perhaps a gas of even higher molecular weight is preferred to helium if a detector other than thermal conductivity is being used.

A rotometer may be used in the carrier-gas system to give an approximate indication of flow rate. A rotometer consists of a graduated tube with slowly increasing inside diameter and a glass or metal ball that is suspended in the gas flow within the tube at a height dependent upon the flow rate. Since the position of the ball is a function of both the flow rate and the column back pressure when positioned at the column inlet, a rotometer can be used only for rough approximations of flow rate. A soap-bubble flowmeter is used for more accurate measurements.

Sample inlets. These are of two general types depending upon whether the sample is gaseous, liquid, or solid. Liquid samples are generally injected by means of a calibrated hypodermic syringe through a silicon rubber septum into a metal or glass-lined metal injection port, while gaseous samples are introduced by means of a valve and sample loop system. Liquid sample sizes commonly range from 0.5 to 5 microliters. Thus, only very small amounts of material are required for gas chromatographic analysis, 1 microliter being approximately equal to one-fiftieth of a normal-sized drop. Gaseous samples of 0.5–1 milliliter are often used. Injection techniques somewhat similar to those used for liquid samples are used for solids. In order to produce sharp chromatographic peaks with minimum peak overlap, solid and liquid samples must be vaporized rapidly upon injection by maintaining the injection port at a temperature greater than the boiling point of the sample. With some gas chromatographic detectors, the sensitivity is so great that an inlet splitter must be used to reduce the sample amount to one-tenth or one-hundredth of the amount injected. Sample splitters are commonly used with capillary columns, owing to the relatively small amount of liquid phase used with these columns and the much reduced sample capacity which results.

Column. This is the heart of the gas chromatograph, and separation of components on packed columns depends more on the choice of liquid phase than on any other factor. Typically the column is a glass or metal tube of 0.125 or 0.25 in. (3 or 6 mm) in diameter and 4–6 ft (1.2–1.8 m) in length, packed with an inert diatomaceous earth support coated with a nonvolatile liquid to 3–20% by weight. In open tubular or capillary column technology, the support for the thin film of liquid phase is the wall of the capillary itself. Support-coated open tubular columns are also sometimes used, the sample capacity of the columns being increased by the presence of very loosely packed support or by a roughening of the capillary walls.

In choosing a liquid phase for a particular separation, several factors are considered, including solute-solvent interactions (for example, hydrogen bonding), which if present gives special selectivity for a particular solute; temperature limitations of the liquid phase; and the possibility of rare irreversible reactions on the column. Common liquid phases include high-molecular-weight hydrocarbons such as Apiezon L and the natural product, squalane, silicone gum rubbers such as SE-30 and methyl phenyl silicones, poly(ethylene glycol)s such as Carbowax 20M, and certain esters of low volatility such as triethylene glycol succinate. The general rule "like dissolves like" is useful in liquid-phase selection, since it is generally true that polar liquid phases make good solvents for polar solutes while nonpolar liquid phases are often the choice for separation of nonpolar solutes.

In packed column gas chromatography, the liquid phase is adsorbed as a uniform, thin film over the surface of an inert solid support. A good solid support should have the following characteristics: chemical inertness, mechanical strength, relatively large surface area per unit volume, low pressure drop, and thermal stability. Materials with surface areas of 300–1800 ft^2/oz (1–6 m^2/g) are available which fulfill most of these requirements. The most commonly used support in gas chromatography is diatomaceous earth (kieselguhr), which consists of agglomerates of the siliceous skeletons of diatoms treated to produce materials known variously as Chromosorb-P and -W, Celite, C-22 firebrick, Dicalite, Sil-O-Cel, Embacel, and Sterchamol. Chromosorb-G is a diatomaceous earth treated to produce an inert solid, similar in properties to Chromosorb-W but much less friable. To reduce chemical reactivity, the solid support is frequently acid-washed, rinsed, dried, and treated with a silanizing reagent such as trimethylchlorosilane or dimethyldichlorosilane. The supports are generally size-graded to fall in the range 60–80 or 80–100 mesh.

Detectors. The two most popular detectors are the flame ionization detector and the thermal conductivity detector. The flame ionization detector is the more sensitive, with a limit of detection of 0.1–0.001 that of the thermal conductivity detector. The flame

ionization detector has a large linear dynamic range, the range in sample size over which the detector responds in a linear manner. In this detector the sample is burned in a small hydrogen flame as it is eluted from the column. Upon combustion of the organic material, charged species are formed in the flame, and these species are collected by a polarized collector electrode. The signal is then fed to an electrometer, a signal being observable for samples as small as 10^{-11} g. Although the detector is sensitive to most carbon compounds, it gives little or no signal for CO_2, CO, CS_2, H_2O, O_2, N_2, NH_3, the rare gases, or oxides of nitrogen and sulfur. The detector has enjoyed widespread use in gas chromatographic analysis and is obviously attractive in pollution studies involving trace organic compounds in air and water owing to its insensitivity toward the air and the water.

The thermal conductivity cell was one of the first gas chromatographic detectors developed. It consists of a metal filament or a thermistor, heated by an electric current from a power supply, and positioned in the gas stream at the end of the column. The filament increases in temperature when organic materials are eluted, since organic compounds have a thermal conductivity less than that of the helium carrier gas, and the heat is therefore conducted away less rapidly.

Wire filaments increase their resistance about 0.4% per degree while thermistors decrease their resistance by about −4% per degree. Thus thermistor thermal conductivity detectors are more sensitive than filaments at low temperatures but become less sensitive at higher detector temperatures. The change in resistance is sensed by making the detector filament a part of a Wheatstone bridge. The advantages of the thermal conductivity detector are that it is sensitive toward all components except the carrier gas being used, and it is nondestructive. It also has a large linear dynamic range, but the limit of detection is less than that of the flame ionization detector.

Other popular gas chromatographic detectors include the flame photometric detector, the electron capture detector, and the thermionic detector. The flame photometric detector may be made selective toward individual metals in organometallic compounds, volatile metal chelates, and volatile metal halides, as well as toward compounds of sulfur and phosphorus. An inexpensive monochromator system or a filter system may be used to isolate the emission wavelength of a particular element, thus making the detector highly selective. The electron capture detector is one of the most sensitive detectors for compounds with electron-capturing properties, such as halogenated compounds and highly conjugated polynuclear aromatic hydrocarbons. It is insensitive to compounds with little or no electron-capturing properties. This detector has proved of particular value in the analysis of pesticides, many of which are chlorinated compounds. If certain alkali metal salts (such as Rb_2SO_4 and CsBr) are placed in the flame of the flame ionization detector and the flow of hydrogen and air is properly adjusted, it is found that the detector becomes particularly sensitive to nitrogen and phosphorus compounds. This orientation is called the alkali flame or thermionic detector.

Quantitative and qualitative analysis. Qualitative information in gas-liquid chromatography is obtained from retention measurements of the peak, and quantitative information is obtained from the area under a peak or, with control of certain variables, from the peak height. Retention time is constant for a given substance on a given column at a given column temperature and carrier-gas flow rate. Thus, a comparison of the retention time of an authentic sample with a component of an unknown mixture gives useful qualitative evidence. Combination of gas chromatography with mass spectrometry provides the ultimate in qualitative information and has been used extensively in research.

The area under a chromatographic peak is measured by means of an electromechanical integrator, an on-line computer system, a planimeter, triangulation approximations (such as multiplication of peak height by the peak width at half height), or sometimes crudely by cutting out the recorded chromatogram and weighing on a balance. Calibration curves may be made by plotting the area or peak height as a function of concentration or amount of sample.

Nonanalytical applications of gas chromatography include measurement of the heat of solution, adsorption, and other thermodynamic quantities; the surface area of powders; the vapor pressure of materials; chemical kinetics; and equilibrium constants. *See* MASS SPECTROMETRY; QUALITATIVE CHEMICAL ANALYSIS.

Important equations. A 4-ft (1.2-m) packed gas chromatography column typically has the equivalent of 600–2000 theoretical plates, N. Capillary columns of 1,000,000 plates have been prepared. N is a measure of the column efficiency and is given by Eq. (1) where x is the distance to the peak maximum

$$N = 16\left(\frac{x}{y}\right)^2 \tag{1}$$

from the point of injection, and y is the peak width measured at the base. Column efficiency and peak broadening may be evaluated by Eq. (2) where H is

$$H = 2\lambda d_p + \frac{2\gamma D_g}{u} + \frac{8}{\pi^2}\frac{k}{(1+k)^2}\frac{d_l^2}{D_l}u \tag{2}$$

the average height equivalent to a theoretical plate ($= L/N$); L is the column length; λ and γ are related to the uniformity in size and shape of the particles in the column; d_p is the average particle diameter; D_g and D_l are the diffusivity of the solute in the gas and liquid phases, respectively; u is the average linear gas velocity; k is the partition ratio, related to the thermodynamic partition coefficient; and d_l is the average thickness of the liquid layer adsorbed on the solid support. This equation shows that optimum column efficiency (smallest value of H) is obtained by using particles of small diameter, uniform in size; a carrier gas of high molecular weight at an optimum velocity

of generally 1–4 in./s (2.5–10 cm/s); a nonvolatile liquid phase of low viscosity which will retain the solute well on the column; and a uniform distribution of the liquid phase in a thin layer over the surface of a solid support possessing a large specific surface area. *See* CHROMATOGRAPHY; GAS AND ATMOSPHERE ANALYSIS. Richard S. Juvet, Jr.

Bibliography. R. L. Grob and E. F. Barry (eds.), *Modern Practice of Gas Chromatography*, 4th ed., 2004; W. Jennings et al., *Analytical Gas Chromatography*, 2d ed., 1997; H. M. McNair and J. M. Miller, *Basic Gas Chromatography*, 1998; W. M. A. Niessen (ed.), *Current Practice of Gas Chromatography-Mass Spectrometry*, 2001.

Gas constant

The universal constant R that appears in the ideal gas law, Eq. (1), where P is the pressure, V the volume, n the amount of substance, and T the thermodynamic (absolute) temperature. The gas constant is universal in that it applies to all gases, providing they are behaving ideally (in the limit of zero pressure). The gas constant is related to the more fundamental Boltzmann constant, k, by Eq. (2), where N_A is the Avogadro constant (the number of entities per mole). The best modern value in SI units is $R = 8.314\ 472\ (15)$ J/K · mol, where the number in parentheses represents the uncertainty in the last two digits. *See* BOLTZMANN CONSTANT; GAS.

$$PV = nRT \tag{1}$$

$$R = N_A k \tag{2}$$

According to the equipartition principle, at a temperature T, the average molar energy of each quadratic term in the expression for the energy is $(1/2)RT$; as a consequence, the translational contribution to the molar heat capacity of a gas at constant volume is $(3/2)R$; the rotational contribution of a linear molecule is R. *See* KINETIC THEORY OF MATTER.

Largely because R is related to the Boltzmann constant, it appears in a wide variety of contexts, including properties unrelated to gases. Thus, it occurs in Boltzmann's formula for the molar entropy of any substance, Eq. (3), where W is the number of arrangements of the system that are consistent with the same energy; and in the Nernst equation for the potential of an electrochemical cell, Eq. (4), where E° is a standard potential, F is the Faraday constant, and Q is a function of the composition of the cell. The gas constant also appears in the Boltzmann distribution for the population of energy levels when the energy of a level is expressed as a molar quantity. *See* BOLTZMANN STATISTICS; ELECTRODE POTENTIAL; ENTROPY. P. W. Atkins

$$S = R \ln W \tag{3}$$

$$E = E^\circ - (RT/nF) \ln Q \tag{4}$$

Bibliography. P. W. Atkins, *Physical Chemistry*, 6th ed., 1998.

Gas discharge

A system made up of a gas, electrodes, and an enclosing wall in which an electric current is carried by charged particles in response to an electric field, the gradient of the electric potential, or the voltage between two surfaces. The gas discharge is manifested in a variety of modes (including Townsend, glow, arc, and corona discharges) depending on parameters such as the gas composition and density, the external circuit or source of the voltage, electrode geometry, and electrode material. A gas discharge can also be inductively coupled to an alternating-current (ac) circuit, eliminating the need for electrodes. Gas discharges are useful both as tools to study the physics existing under various conditions and in technological applications such as in the lighting industry and in electrically excited gas lasers. *See* LASER; VAPOR LAMP.

The study of gas discharges encompasses a broad range of topics in physics and engineering, only a few of which are mentioned here. New applications in gas insulation, in high-power electrical switching, and in materials reclamation and processing will assure a continuing effort to better understand all aspects of gas discharges. *See* GAS TUBE.

Sources of electrons. Electrons, rather than ions, are the main current carriers in gas discharges because their mass is smaller and their mobility is correspondingly much higher than that of ions. Electrons are produced by ionization of the gas itself, or they originate at the electrodes present in the system. Gas ionization can be accomplished in several ways, including electron impact ionization, photoionization, and associative ionization. Bombardment by photons, energetic ions or electrons, and excited neutral particles can cause secondary emission from the electrodes. A high-energy-per-unit electrode surface area can induce thermionic or field emission of electrons. Each means of producing electrons leads to a different response of the gas discharge as a circuit element. *See* ELECTRON EMISSION; FIELD EMISSION; IONIZATION; PHOTOEMISSION; SECONDARY EMISSION; THERMIONIC EMISSION.

Townsend discharge. A nonself-sustained discharge is one in which an agent external to the circuit is necessary to replenish the electrons lost from the system in order to maintain a flow of current; an example is a Townsend discharge. A Townsend discharge, considered here in a parallel plate electrode geometry, is a low-current (less than 10^{-5} A) discharge in which the primary source of current carriers is, for example, photoemission resulting from illumination of the cathode. The electrons gain energy from the electric field between the electrodes and lose energy to the gas through elastic and inelastic collisions with the neutral gas particles. Except in regions very near

the electrodes, the energy gained by the electrons from the field balances that lost in collisions, and the discharge characteristics can be parametrized simply by the neutral gas composition and by the ratio of the field strength E to the neutral gas density n. Townsend discharges have been widely used to study the physics of the interaction of electrons with neutral gas particles because external control of E/n provides a convenient means for adjusting the electron energy. *See* TOWNSEND DISCHARGE.

Gas breakdown. For E/n greater than about 10^{-19} V · m^2 (100 townsends) some of the electrons have sufficient energy to produce additional ionization in collisions with the neutral gas particles. The succeeding generation of electrons may also produce further ionization, and if the volume rate of ionization exceeds the volume losses due to diffusion and attachment, an exponential increase in the electron density results. This cascade is termed an electron avalanche. If the avalanche grows to the point where the gas loses its insulating properties and becomes highly conducting, gas breakdown is said to have occurred. The luminosity accompanying gas breakdown is due to electron impact excitation of the gas particles which then radiate. Lightning is a familiar example of the effect. *See* ELECTRICAL BREAKDOWN.

In determining the minimum voltage required for a self-sustained discharge, electron emission from the electrodes must be considered as well as the cascade ionization; that is, electrons originating from secondary emission at the electrodes must be considered in determining the minimum applied voltage for which the electron sources in the discharge balance the losses. The Paschen curve shows that this minimum voltage, or sparking potential, is a function of the product of the gas pressure p and the electrode separation d. The sparking potential at a given value of pd depends on gas composition, electrode material and geometry, and applied field frequency, but a minimum in the sparking potential always occurs at a finite value of pd. *See* BREAKDOWN POTENTIAL; ELECTRIC SPARK.

Glow discharge. At pressures lower than about 100 torr (13 kilopascals), but depending on the gas composition, a succession of overlapping avalanches between two electrodes leads to the very stable glow discharge, where space-charge fields adjust themselves so that the volume production of ionization, and the resulting secondary emission from the electrodes, balances the loss of the current carriers. The glow discharge is characterized by moderate current (10^{-5} to 10^{-2} A) and voltage (on the order of 100 V), a particular field distribution due to the space-charge distortion of the applied field, and a well-defined luminosity pattern between the electrodes. In addition, the average electron energy is considerably higher than the average kinetic or internal energies of the neutral particles. In the normal glow discharge, the discharge voltage and the current density are nearly constant. An increase in current is achieved by reducing the external circuit impedance or by increasing the source voltage, which causes an increase in the electron-emitting area on the cathode surface. Once the glow covers the cathode surface, an additional increase in current is associated with an increase in the discharge voltage and in the current density. This regime is termed the abnormal glow. *See* GLOW DISCHARGE.

Arc. By lowering the circuit resistance, or increasing the source voltage, a further increase in current can be realized. Past a certain current level, however, this increase is accompanied by an abrupt change in the discharge characteristics, and this high-current discharge mode is termed an arc. An arc is distinguished operationally from a glow by the large currents and the small voltage drop. A more fundamental distinction can be made by considering the sources of the ionization. A low-pressure arc is sustained by thermionic or field emission of electrons from the electrode surfaces; secondary electron emission and electron impact ionization are still present but are not very important as electron sources. At higher pressures, volume sources of electrons dominate, but the sources involve multistep, cumulative ionization through excited states, rather than the electron impact ionization predominantly from ground-state neutral particles, which is characteristic of most glow discharges. The range of applications of arcs is enormous. *See* ARC DISCHARGE; ARC LAMP; ARC WELDING.

Glow-to-arc transition. At pressures higher than about 100 torr (13 kPa), depending on the gas composition, the glow discharge is not stable and has a tendency to constrict to form an arc. The details of the mechanisms for the glow-to-arc transition are not clearly established, but involve changes in the kinetic and internal energy of the gas due to the large amounts of energy deposited in the gas. In a high-pressure arc, the gas is hot and highly excited, and collisions between electrons or neutral particles and excited gas particles produce most of the ionization; that is, the system approaches local thermodynamic equilibrium described by temperatures on the order of 10^3 to 10^4 K (2×10^3 to 2×10^4 °F). Diffusion at the lower pressures tends to dissipate the energy in the gas and to stabilize the discharge in the glow mode, but the lower diffusion at higher pressures leads to nonuniform gas heating and excitation in the center of the discharge and the associated highly conducting arc. Any other nonuniformities in the system, either in the gas or on the electrodes, can also provide a seed for arc formation at high pressures. Arcs are undesirable in applications that depend on uniform volume excitation of the gas such as electrically excited gas lasers.

Corona discharge. This discussion has implicitly focused on parallel plate electrode geometries. Discharges in irregular geometries where the applied field confines the current density to localized regions are called corona discharges. These relatively low-current, high-voltage discharges occur in Townsend or glow discharges with point-plane electrode configurations or around conductors with a small radius of curvature, and are often intermittent. Corona discharges can represent significant power

losses in high-voltage transmission lines. *See* CORONA DISCHARGE; ELECTRICAL CONDUCTION IN GASES.

Leanne C. Pitchford

Bibliography. R. W. Crompton et al. (eds.), *Gaseous Electronics and Its Applications*, 1991; M. N. Hirsh and H. J. Oskam (eds.), *Gaseous Electronics*, vol. 1: *Electrical Discharges*, 1978; E. E. Kunhardt and L. H. Luessen (eds.), *Electrical Breakdown and Discharges in Gases*, 2 vols., 1983; J. M. Meek and J. D. Craggs (eds.), *Electrical Breakdown in Gases*, 1978; Y. P. Raizer and J. E. Allen, *Gas Discharge Physics*, 1997.

Gas dynamics

The study of gases in motion. In general, matter exists in any of three states: solid, liquid, or gas. Liquids are incompressible under normal conditions; water is a typical example. In contrast, gases are compressible fluids; that is, their density varies depending on the pressure and temperature. The air surrounding a high-speed aircraft is an example. *See* GAS; LIQUID.

Scope of subject. Gas dynamics can be classified in a variety of ways. The following discussion deals with gases as a continuum. The structure of gases on the particle level is called rarefied gas dynamics which is discussed in the last section of the article. *See* AEROTHERMODYNAMICS; COMPRESSIBLE FLOW; FLUID FLOW.

Fundamental relations. Gases in motion are subject to certain fundamental laws. These are the laws of the conservation of mass, momentum, and energy. They apply as long as the velocity does not approach the speed of light, when relativistic effects must be taken into consideration. This regime is outside the scope of this discussion. The conservation laws are represented by the continuity-of-mass equation, the conservation-of-momentum equation, and the conservation-of-energy equation. In other words, the fundamental conservation principles of mechanics and dynamics constitute the theoretical basis of gas dynamics. In the case of the dynamics of incompressible fluids, it is usually sufficient to satisfy only the laws of conservation of mass and momentum. This distinction constitutes the fundamental difference between high-speed gas dynamics and hydrodynamics. If irreversibilities are involved, a fourth equation called the entropy balance equation may be considered. Whereas mass, momentum, and energy are conserved, the entropy is not. Real problems are irreversible; that is, losses such as friction are involved. However, as a first approximation such effects are generally not considered. *See* CONSERVATION LAWS (PHYSICS); CONSERVATION OF ENERGY; CONSERVATION OF MASS; CONSERVATION OF MOMENTUM.

The mass, momentum, and energy equations are higher-order, nonlinear equations that have no general solution, and only a limited number of special cases can be solved. Another approach is to resort to numerical solutions using high-speed digital computers. While this approach has proven to be very useful, it limits the degree to which flow phenomena can be conceptualized. Accordingly, it is frequently permissible to write the equations in one-dimensional form. By one-dimensional flow is meant that the properties of gas such as its velocity and density are assumed to be constant in the direction perpendicular to the direction of the gas flow. Generally, the one-dimensional approach gives excellent insights into understanding the physical behavior of the flow. It is also very useful in setting up the computer algorithm for numerical solutions. *See* COMPUTATIONAL FLUID DYNAMICS; DIFFERENTIAL EQUATION.

Continuity equation. One other matter must be considered, namely whether the flow is steady or unsteady. In steady flow, the flow characteristics do not vary with time, whereas unsteady flow implies that the flow assumes different configurations over time. Thus, unsteady flow is broader in scope. In this case the continuity equations for conservation of mass may be written as Eq. (1). In this equation, the first

$$\frac{\partial(\rho V)}{\partial x} + \frac{\partial \rho}{\partial t} = 0 \tag{1}$$

term defines the mass-flow changes with respect to the space coordinates, whereas the second term indicates the changes with time. Here, ∂ is the partial differential operator; x denotes the space coordinate, in this case the direction of flow; ρ is the gas density; V is the gas velocity; and t is the time.

If the flow is steady, there is no time-dependent term, and hence the continuity equation can be written in integrated form as Eq. (2), where A denotes

$$\rho V A = \text{constant} \tag{2}$$

the area in the direction perpendicular to the flow direction.

Momentum equation. The momentum equation is the mathematical representation of the law of conservation of momentum. It is a statement of the forces acting on the gas. Different types of forces must be recognized. Body forces, such as gravitation and electromagnetic forces, act at a distance. The so-called surface forces may assume different forms, such as normal stresses and viscosity. The simplest form of the momentum equation relevant to gas dynamics, which applies to steady flow without friction, is Eq. (3). This equation demonstrates that, where the

$$-dp = \rho V dV \tag{3}$$

pressure increases along the flow streamlines, the velocity decreases. For gases at very low speed (nearly constant density), Eq. (3) may be integrated to give the familiar Bernoulli equation (4), where the con-

$$\frac{p}{\rho} + \frac{V^2}{2} = \text{constant} \tag{4}$$

stant can be reduced by friction. *See* BERNOULLI'S THEOREM.

Energy equation. The energy equation expresses the first law of thermodynamics and accounts for the changes in energy as the gas moves about its path.

It can also take into consideration energy exchanges between the gas and its environment, such as radiation. The simplest form of the energy equation for compressible gas flow, which applies along streamlines in the absence of heat transfer, is Eq. (5), where

$$c_pT + \frac{V^2}{2} = \text{constant} \tag{5}$$

c_p denotes the specific heat at constant pressure and T is the gas temperature. Equation (5) shows that where temperature rises the speed decreases and vice versa. The constant in Eq. (5) can be increased by heat transfer to the gas.

The above review intentionally presented the governing equations of gas dynamics in their elementary forms. Incorporating all possible effects requires elaborate discussions that can be found in textbooks.

Speed of sound. The speed of sound or the acoustic velocity is a very important term in gas dynamics because it serves as a criterion to identify flow regimes. Being able to do so is crucial because the designer must know the conditions that the gas will generate or, conversely, experience. The acoustic velocity (a) is defined by Eq. (6), where γ is the specific-heat

$$a = \sqrt{\gamma RT} \tag{6}$$

ratio and R is the gas constant.

In prescribing flow regimes, the flow velocity of the gas is compared with the acoustic velocity. This ratio, called the Mach number (M), is defined by Eq. (7). Using the Mach number the following flow

$$M = \frac{V}{a} \tag{7}$$

regimes are described:

$M < 1$	subsonic flow
$M = 1$	sonic flow
$0.9 < M < 1.1$	transonic flow
$M > 1$	supersonic flow
$M > 5$	hypersonic flow

High-speed aircraft are categorized by the Mach number. The Mach number is significant otherwise. Squaring both sides of Eq. (7) yields Eq. (8), which

$$M^2 = \frac{V^2}{a^2} \tag{8}$$

can be recognized as the ratio of the kinetic energy to the thermal energy. This concept is very helpful in the manipulation of gas-dynamic equations because gas dynamics is really a combination of dynamics and thermodynamics. *See* MACH NUMBER; SOUND; THERMODYNAMIC PRINCIPLES.

Wave phenomena. Consider a pulsating pressure source moving with velocity V. This source starts pulsating at time $t = 0$. In a time interval t_1, the source travels a distance Vt_1, while the signal reaches the surface of a sphere of radius at_1 (the surface of this sphere constitutes the wavefront at time t_1). At a later time t_2, the source point is at Vt_2, while the signal front is at at_2. Thus, as long as $V/a = M < 1$, the source point is always inside the outermost wavefront.

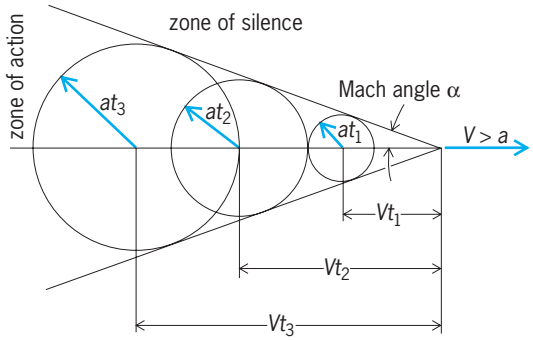

Fig. 1. Rule of forbidden signals from a point source moving at supersonic velocity. (*After A. B. Cambel and B. H. Jennings, Gas Dynamics, McGraw-Hill, 1958*)

Still another use of the term wave refers to a wave envelope. If the point source moves so fast that $V > a$, the wavefront spheres will no longer contain the source (**Fig. 1**). The envelope to this family of spheres is a cone, known as the Mach cone, and the Mach angle is such that $\sin \alpha = 1/M$.

The Mach line constitutes a demarcation (Fig. 1). The fluid outside the Mach line will receive no signal from the source. T. von Kármán has appropriately called this phenomenon the rule of forbidden signals and designated the region ahead of the Mach line the zone of silence and the region inside the Mach line the zone of action. *See* HYPERSONIC FLIGHT; SUPERSONIC FLIGHT; TRANSONIC FLIGHT.

Shock waves. In the same manner in which a Mach wave is the envelope of infinitesimal disturbances, a shock wave is the envelope of finite disturbances. The steady conditions on either side of a standing shock wave can be obtained by applying the conservation laws expressed by Eqs. (9), (10), and (11).

$$\rho_1 V_1 = \rho_2 V_2 \qquad \text{(continuity)} \tag{9}$$

$$p_1 + \rho_1 V_1^2 = p_2 + \rho_2 V_2^2 \qquad \text{(momentum)} \tag{10}$$

$$h_1 + \frac{V_1^2}{2} = h_2 + \frac{V_2^2}{2} \qquad \text{(energy)} \tag{11}$$

In Eq. (11), $h = u + p/\rho$, where u is the internal energy. The quantity h is termed enthalpy, and for a perfect gas it is given by c_pT. By a simple rearrangement of these equations, the approximate expressions are obtained for the shock Mach number in the case of a perfect gas, Eq. (12) for weak shocks and Eq. (13) for strong shocks.

$$M_s \approx 1 + \frac{\gamma + 1}{4\gamma}\frac{p_2 - p_1}{p_1} \tag{12}$$

$$M_s \approx \left(\frac{\gamma + 1}{2\gamma}\frac{p_2}{p_1}\right)^{1/2} \tag{13}$$

In a sound wave, $p_2 - p_1 \simeq 0$ and, therefore, $M_s = 1$; for $p_2/p_1 = 4$, the shock speed is roughly twice the

	M_1	M_2	p_2/p_1	V_2/V_1	ρ_2/ρ_1	T_2/T_1
detonation	>1	≶1	>1	<1	>1	>1
deflagration	<1	<1	<1	>1	<1	>1

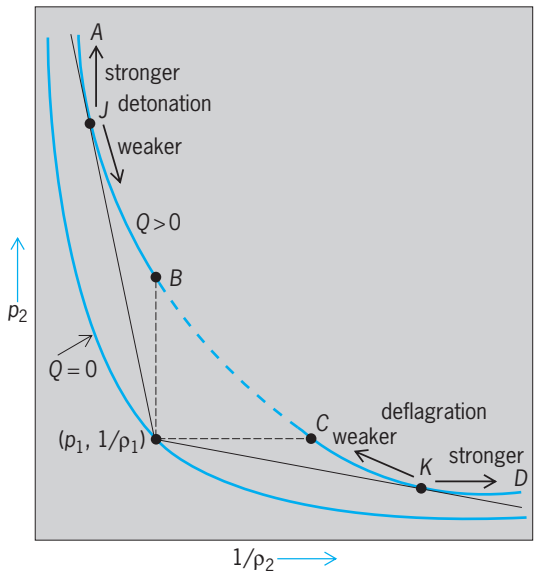

Fig. 2. Hugoniot diagram and Chapman-Jouget conditions. (*After A. B. Cambel and B. H. Jennings, Gas Dynamics, McGraw-Hill, 1958*)

speed of sound, as shown in Eq. (13). *See* ENTHALPY; SHOCK WAVE.

Detonation and deflagration waves. Other interesting gas-dynamic waves are characterized by the same continuity and momentum equations. The energy equation, however, is modified to include a term which accounts for chemical heat release (Q); such waves are either detonations or deflagrations.

Eliminating the kinetic energy from the energy equation by the use of the continuity equation and the momentum equation yields Eq. (14).

$$h_1 - h_2 + Q = \frac{1}{2}(p_1 - p_2)\left(\frac{1}{\rho_2} + \frac{1}{\rho_1}\right) \tag{14}$$

For a given Q, zero or nonzero, and given p_1 and ρ_1 (which through the equation of state gives h_1) and one additional variable behind the wave, for example, V_2 or h_2, the locus of all possible combinations of ρ_2 and p_2 can be plotted on a so-called Hugoniot diagram (**Fig. 2**). The lines OJ and OK are tangent to the Hugoniot curve; O is the point p_1, $1/\rho_1$. Points J and K separate strong and weak waves. Flows corresponding to J and K are characterized by the fact that all the thermodynamic and fluid-mechanic variables have an extremum. Transitions from B to C or vice versa involve a decrease in entropy and are therefore forbidden. *See* ENTROPY.

Mach number functions. In many applications it is reasonable to assume that the gas is perfect—both thermally and calorically—and that the flow is adiabatic. It then becomes very useful to express all the dependent variables in terms of the Mach number.

Starting with the energy equation, a stagnation temperature T_0 is defined by expression (15).

$$\frac{V^2}{2} + c_p T = c_p T_0 \tag{15}$$

Physically, T_0 represents the temperature the gas would have if all its kinetic energy were transformed into thermodynamic enthalpy. The stagnation enthalpy can be measured by a thermometer immersed in a gas stream.

From Eq. (14), a simple derivation yields Eq. (16),

$$\frac{T_0}{T} = 1 + \frac{\gamma - 1}{2}M^2 \tag{16}$$

where T is now called the static temperature. This temperature is measured with a thermometer at rest with respect to the gas. The stagnation temperature is constant in any adiabatic flow, even through a shock, and thus provides an excellent reference parameter. By using the isentropic relation $p = \rho^\gamma$ and the perfect gas law, Eq. (17), a reference stagnation

$$P = \rho RT \tag{17}$$

pressure and density can be defined by means of Eqs. (18) and (19).

$$\frac{p_0}{p}\left(1 + \frac{\gamma - 1}{2}M^2\right)^{\gamma/(\gamma-1)} \tag{18}$$

$$\frac{\rho_0}{\rho}\left(1 + \frac{\gamma - 1}{2}M^2\right)^{1/(\gamma-1)} \tag{19}$$

When $M^2 \ll 1$, Eq. (18) may be expanded by the binomial theorem to obtain Eq. (20).

$$p_0 = p + \frac{1}{2}\rho V^2\left(1 + \frac{M^2}{4} + \cdots\right) \tag{20}$$

The deviation from Bernoulli's equation, Eq. (4), due to compressibility at $M = 0.5$ is only 6%.

Flows can be classified as internal flow and external flow. Internal flow refers to the cases where the gas is constrained by a duct of some sort. Characteristically external flow is flow over an airplane or missile.

Internal one-dimensional flow. Internal flows are conveniently characterized by (1) the shape of the duct and its variation, (2) the heat transfer through the walls of the duct and internal heat sources, and (3) frictional effects. By varying one of these characteristics at a time, the essential features of internal flow can be discussed most simply.

Variable area flow. A device to accelerate the flow of a gas or liquid is termed a nozzle. In most engineering applications the contour of the nozzle is first converging and then diverging; it thus has a minimum cross section, called a throat. *See* NOZZLE.

For isentropic flow in a convergent-divergent nozzle in which the flow is supersonic in the divergent section, the velocity at the throat is sonic, that is, $M^* = 1$. The throat pressure is then said to be

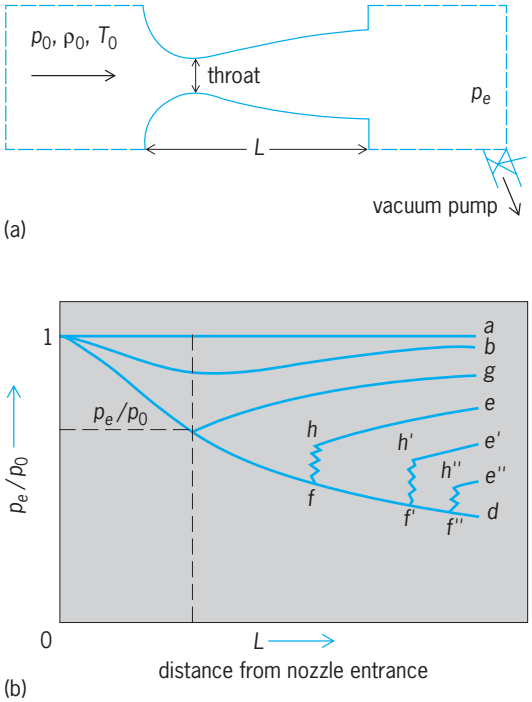

Fig. 3. Convergent-divergent nozzle between two reservoirs. (*a*) Diagram of nozzle. (*b*) Pressure distribution in nozzle for given reservoir pressure (p_0) and various values of exit pressure (p_e) (*After A. B. Cambel and B. H. Jennings, Gas Dynamics, McGraw-Hill, 1958*).

critical p^* and is given by Eq. (21). Velocity and pressure are related in this case by Eq. (22).

$$p^* = p_0\left(\frac{2}{\gamma+1}\right)^{\gamma/(\gamma-1)} \tag{21}$$

$$V = \frac{2\gamma R}{\gamma-1}T_0[1-(p/p_0)^{(\gamma-1)/\gamma}] \tag{22}$$

A convergent-divergent deLaval nozzle can be inserted between two reservoirs (**Fig. 3***a*), and the pressure can be plotted as a function of the distance from the nozzle entrance (Fig. 3*b*). There will be no flow if the ratio of exit pressure to reservoir pressure is $p_e/p_0 = 1$, as represented by curve *a*. If p_e is reduced so that it is slightly less than the entrance pressure, the nozzle will act like a conventional venturi, as represented by curve *b*. For this case, the flow is always subsonic and resembles incompressible flow. When the exit pressure is reduced further, the critical pressure can be reached at the throat, as curve *g* shows. In this case, the velocity is sonic at the throat but is never supersonic within the nozzle, even though the pressure at the throat corresponds to the critical. The minimum pressure which can exist in the nozzle outlet is depicted by point *d*, for which the pressure at the throat will be the critical. Here the velocity in the converging section is subsonic, in the diverging section it is supersonic, and at the throat it is sonic. For the range of exit pressures from p_d to p_g, the rate-of-flow remains constant for a given reservoir pressure p_0. The flow rate reaches a maximum value and remains there over this wide range of exit pressures.

Even in the absence of friction, isentropic flow can exist only for the range of exhaust pressures from p_a to p_g and at the pressure reached along curve *d*, but not at intermediate pressures. The pressures p_g and p_d are the significant design pressures for a given nozzle. For exhaust pressures in the range between p_d and p_g, shocks will occur in the nozzle, raising the pressure from *f* to *h* (or *f'* to *h'*), followed by a pressure rise after the shock points to an exit pressure such as p_e. If p_e is less than p_d, the jet leaving the nozzle is said to be underexpanded and will drop in pressure after leaving the mouth of the nozzle. The velocity V_1 in front of a shock is supersonic, and the velocity V_2 behind a normal shock (in contrast to an oblique shock) is always subsonic. For a normal shock $V_1V_2/a^{*2} = 1$, where a^* is the critical speed of sound corresponding to $M = 1$. Thus for $V_1 > V_2$, $V_1/a^* > 1$ and therefore $V_2/a^* < 1$. *See* ISENTROPIC FLOW.

Diabatic flow. Heat exchangers and combustion chambers are devices in which heat transfer occurs. The equations describing nonadiabatic or diabatic processes are complicated; consequently, certain limiting assumptions are usually required to make possible analytical solutions of the equations.

These assumptions are that (1) the flow takes place in a constant-area section, (2) there is no friction, (3) the gas is perfect and has constant specific heats, (4) the composition of the gas does not change, (5) there are no devices in the system which deliver or receive mechanical work, and (6) the flow is steady.

Equations which conform to these requirements are called Rayleigh equations, and the associated flow is designated as Rayleigh flow. Designating by $Q_{1\to2}$ the quantity of heat introduced between stations 1 and 2, Eq. (23) is obtained for the energy equation.

$$\begin{aligned} Q_{1\to2} &= c_p(T_2 - T_1) + \frac{V_2^2 - V_1^2}{2} \\ &= h_2 - h_1 + \frac{V_2^2 - V_1^2}{2} \end{aligned} \tag{23}$$

If the stagnation enthalpy is introduced, then $Q_{1\to2}$ is given by Eq. (24), which can be expressed in terms of stagnation temperatures by Eq. (25).

$$Q_{1\to2} = h_{02} - h_{01} \tag{24}$$

$$Q_{1\to2} = c_p(T_{02} - T_{01}) \tag{25}$$

Because $Q_{1\to2} \neq 0$ for diabatic flow and because $c_p > 0$ always, it follows that $T_{02} \neq T_{01}$. This inequality states that in diabatic flow the stagnation temperature is not solely determined by the reservoir conditions, as is the case with adiabatic flow. Heating raises the stagnation temperature; cooling lowers it. *See* ADIABATIC PROCESS.

The locus of points of properties during a constant-area, frictionless flow with heat exchange is called the Rayleigh line. By definition, along the Rayleigh line the continuity equation and the

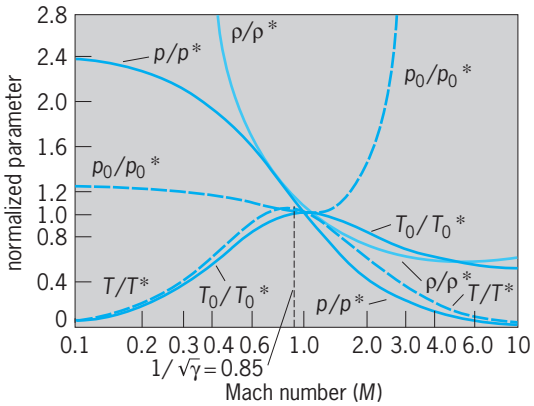

Fig. 4. Diabatic flow parameters for specific heat ratio of 1.4. Asterisk indicates value of parameter when $M = 1$. (*After A. B. Cambel and B. H. Jennings, Gas Dynamics, McGraw-Hill, 1958*)

momentum equation must apply. Equation (9) applies to steady flow in a constant-area duct. Mass velocity by definition is $G = \rho V$ and, from Eq. (10), the momentum relation is $p + \rho V^2 = C$, where C is a constant. Consequently, Eq. (26) is obtained, which

$$p + \frac{G^2}{\rho} = C \tag{26}$$

is one of the many Rayleigh-line equations.

The variations of pressure, temperature, and density with Mach number may be plotted for Rayleigh flow (**Fig. 4**). The fact that the curve for T_0/T_0^* reaches a maximum at a Mach number of unity indicates that it is impossible to pass from one flow domain into the other by the same heat-transfer process. Thus, if heat is added to a subsonic flow, the flow can be accelerated only until its Mach number becomes unity. Further addition of heat will not further accelerate the gas but will result in choking of the flow. As a consequence, the flow must readjust itself, which it will do by lowering its initial Mach number (**Table 1**). *See* CHOKED FLOW.

Flow with friction. In long pipes the effects of friction may result in a significant pressure drop. Over a length dx, this pressure drop dp is given by the Fanning equation (27), where f is a friction factor that

$$dp = -f\frac{\rho V^2}{2r_h}dx \tag{27}$$

must be determined experimentally and r_h is the hydraulic radius of the pipe. To solve friction-flow problems analytically, certain simplifying assumptions are made and the resulting hypothetical flow is called Fanno flow. The Fanno flow assumptions are the same as those for Rayleigh flow except that the assumption that there is no friction is replaced by the requirement that the flow be adiabatic. Numerous Fanno flow equations may be written by combining the energy and the continuity equations in accordance with these assumptions. The variation of properties during Fanno flow may be plotted (**Fig. 5**), and the trends of the most important properties during subsonic and supersonic flow may be summarized in tabular form (**Table 2**).

The Fanno and Rayleigh lines may be plotted for the same constant mass velocity $G = \rho V$ (**Fig. 6**). The Rayleigh and Fanno lines have two points of intersection, denoted by a and b; a normal shock connects these two points. The flow through a shock wave is irreversible; thus, associated with it is an increase in entropy. Point b, thus, always lies to the right of point a.

There is another interesting point about the Fanno curve. If frictional flow continues along the subsonic portion of the Fanno line, the Mach number tends to increase toward unity, whereas if it continues along the supersonic portion, the Mach number decreases toward unity. As in the case of Rayleigh flow, it is impossible, by virtue of the second law of thermodynamics, to pass from one flow regime to the other (subsonic into supersonic or conversely) unless the mass velocity is readjusted.

External flow. Boundary layers and wakes are the centers of interest in external flows. Here the effects of compressibility are substantially more difficult to analyze than in internal flows, if for no other reason than the inapplicability of a one-dimensional approach.

Ballistic missiles and space vehicles enter the Earth's atmosphere with velocities typically of 4 mi/s (6–7 km/s). The corresponding Mach number, depending on the altitude, is of the order of 20. The energy that maintains the bow shock and the work done to overcome the viscous shear stresses, Φ, reduce the kinetic energy of the vehicle, and it decelerates. This loss of kinetic energy, which is the work of the drag forces, reappears in part as the increased enthalpy and temperature of the fluid near the vehicle surface. The rate of heat transfer driven by the large enthalpy gradient $(h_g - h_s)/\Delta$, where h_g is the gas enthalpy at some suitably defined distance Δ from the surface and h_s is the enthalpy right at the

TABLE 1. Variation of flow properties for Rayleigh flow

	Heating		Cooling	
Property	$M > 1$	$M < 1$	$M > 1$	$M < 1$
T_0	Increases	Increases	Decreases	Decreases
p	Increases	Decreases	Decreases	Increases
ρ_0	Decreases	Decreases	Increases	Increases
V	Decreases	Increases	Increases	Decreases
T	Increases	Increases when $M < 1/\sqrt{\gamma}$ Decreases when $M > 1/\sqrt{\gamma}$	Decreases	Decreases when $M < 1/\sqrt{\gamma}$ Decreases when $M > 1/\sqrt{\gamma}$

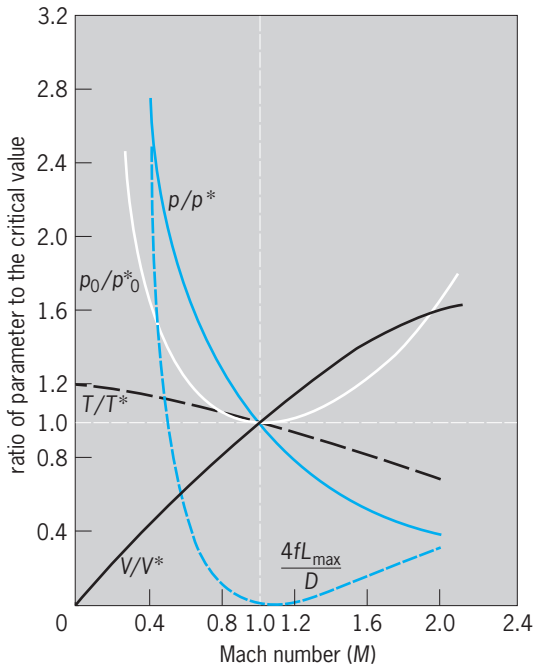

Fig. 5. Functions for constant-area flow with friction ($\gamma = 1.4$). Asterisk indicates value of parameters when $M = 1$. f = friction factor; L_{max} = maximum length of pipe without occurrence of choking; D = pipe diameter. (*After A. B. Cambel and B. H. Jennings, Gas Dynamics, McGraw-Hill, 1958*)

TABLE 2. Fanno flow phenomena

Property	Initial flow is subsonic	Initial flow is supersonic
M	Increases	Decreases
V	Increases	Decreases
p	Decreases	Increases
T	Decreases	Increases
ρ	Decreases	Increases

surface, is so large that special protection must be afforded the vehicle. To this purpose, the vehicle can be covered with a material designed to char, melt, or gasify and in so doing absorb much of the heat that would otherwise penetrate into the structure. This "ablation" process, as it is called, introduces large amounts of material (some of it chemically active, some of it ionized, and some of it radiating) into what ordinarily would be called a boundary layer. Unfortunately, many of the assumptions that make it possible to introduce the boundary-layer simplifications are violated here. Very complex computer programs have been developed, however, that yield reasonably accurate estimates of these effects. *See* ATMOSPHERIC ENTRY; BOUNDARY-LAYER FLOW.

Joshua Menkes; Ali B. Cambel

Rarefied gas dynamics. Rarefied gas dynamics is that branch of gas dynamics dealing with the flow of gases under conditions where the molecular mean free path is not negligibly small compared to some characteristic dimension of the flow field. Rarefied flows occur when the gas density is extremely low, as in the cases of vacuum systems and high-altitude flight, but also when gases are at normal densities if the characteristic dimension is sufficiently small, as in the case of very small particles suspended in the atmosphere.

The dimensionless parameter which describes the degree of rarefaction existing in a flow is the Knudsen number, $\mathrm{Kn} = \lambda/L$, defined as the ratio of the mean free path λ to some characteristic dimension L of the flow field. Depending on the situation, L might be chosen, for example, as the diameter of a duct in a vacuum system, the wavelength of a high-frequency sound wave, the diameter of a suspended submicrometer-size particle, the length of a high-altitude rocket, or the thickness of a boundary layer or a shock wave. The mean free path λ, which is the average distance traveled by a gas molecule between successive collisions with other molecules, is equal to the molecular mean speed, given by Eq. (28), divided by the collision frequency ν_c: thus, Eq. (29)

$$\bar{C} = \sqrt{\frac{8}{\pi}RT} \tag{28}$$

$$\lambda = \frac{\bar{C}}{\nu_c} \tag{29}$$

is satisfied. However, it is often more convenient in evaluating the Knudsen number to use the viscosity-based mean free path given by Eq. (30), where ν is

$$\lambda \simeq \frac{2\nu}{\bar{C}} \tag{30}$$

the kinematic viscosity. *See* VISCOSITY.

Flow regimes. It is convenient to divide rarefied flows into three flow regimes, according to the range of values of the appropriate Knudsen numbers. The regime of highly rarefied flow, which obtains for Kn much greater than 1 (typically greater than 10), is called collisionless or free-molecule flow, while the regime of slight rarefaction, where Kn is much less than 1 (typically less than 0.1), is called slip flow. Flows at Knudsen numbers intermediate to these limiting values are termed transition flows. The phenomena and methods of analysis associated with the

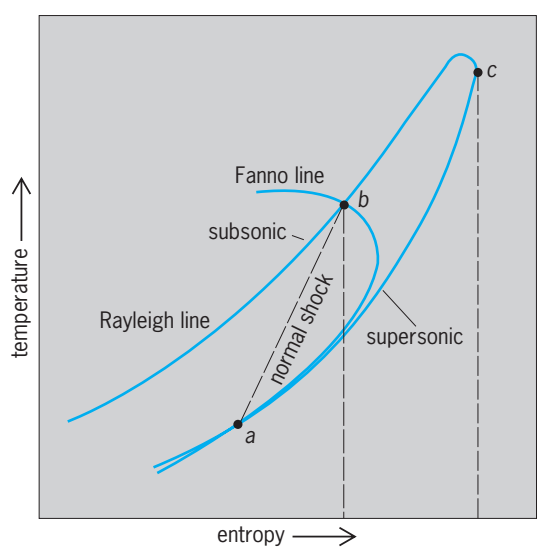

Fig. 6. Rayleigh and Fanno lines. (*After A. B. Cambel and B. H. Jennings, Gas Dynamics, McGraw-Hill, 1958*)

three regimes are in general quite dissimilar, so the classifications are helpful.

Collisionless flow. Collisionless flows are ones for which intermolecular encounters are very rare. Thus, in the case of a vehicle in high-altitude flight, such as a satellite, molecules impinging on a surface will travel, after reflection, very far before colliding with other molecules, with the consequence that the incoming flux of molecules is unperturbed by the reflected flux. For many applications, such as high-altitude aerodynamics, it can be assumed that the molecular velocity distribution f_i of the incident molecules is a drifting maxwellian given by Eq. (31),

$$f_i = n\left(\frac{m}{2\pi kT}\right)^{3/2} e^{-mc^2/2kT} \tag{31}$$

where m is the molecular mass, n the number density, k Boltzmann's constant, T the free stream temperature, and the peculiar molecular velocity $\mathbf{C}$ is the difference between the absolute molecular velocity $\boldsymbol{\xi}$ and the mean gas velocity $\mathbf{V}$, which in the case of an aerodynamic body in steady flight would be its flight velocity. With this assumption, the incident fluxes of mass, momentum, and energy on a unit area of a body can be evaluated readily in terms of appropriate moments of f_i. *See* BOLTZMANN STATISTICS; KINETIC THEORY OF MATTER; STATISTICAL MECHANICS.

The net fluxes to the surface element are determined by the nature of the velocity distribution f_r of the reflected molecules. Unfortunately, f_r is generally not known, and recourse must be had to empirical models of the molecular reflection process. One limiting case is specular reflection, in which the molecular velocity component normal to the surface is reversed, and the tangential velocity component remains unchanged. However, this is not characteristic of real gas-solid surface interactions. Another limiting case is diffuse reflection, for which the velocities of the molecules after reflection are independent of their incident velocities, and are in a half-range maxwellian distribution at the wall temperature T_W. The degree to which an actual surface behaves in a diffusive manner is customarily measured by three surface interaction parameters, or accommodation coefficients.

Calculation of the incident and reflected fluxes of energy and momentum leads to evaluation of the local pressure, shear stress, and energy transport to a surface element of a body. Then by integration over the entire body surface, overall lift, drag, and heat-transfer characteristics are obtained. The calculations are straightforward for simple convex shapes such as flat plates, spheres, and cylinders.

The analysis of collisionless flows becomes more complex if molecules can undergo multiple surface interactions, as in flows in ducts and over concave surfaces.

Slip flow. This is the regime of slight rarefaction, and is manifested initially as an alteration of the boundary conditions associated with the basic flow equations, the Navier-Stokes equations. The typical length scale L defining the Knudsen number is the boundary-layer thickness δ, although it could be the diameter of a pipe in the case of an internal flow. Boundary layers in the slip flow regime are usually laminar, but may exhibit strong compressibility and heat-transfer effects in high-speed flows. Slip relations are applicable only under conditions of small rarefaction, typically $\lambda/L \lesssim 0.1$.

The phenomena of slip and of temperature jump arise from the fact that molecules arriving at a stationary wall from a mean-free-path distant region of moving gas will carry with them the mean energy (temperature) and velocity characteristics of that region. When these are averaged with the properties of the molecules reflected from the wall, a finite-bulk gas slip velocity at the wall results, in contrast with the zero-slip velocity boundary condition employed with the Navier-Stokes equations in the continuum regime. Likewise, the temperature of the gas immediately adjacent to the wall is found to be different from that of the wall itself; thus the terminology temperature jump. The wall layer where these effects occur, the Knudsen layer, has been the subject of many detailed kinetic theory analyses based on the Boltzmann equation or models of it. These analyses have yielded Eqs. (32) and (33) for the fluid-

$$u_0 = C_m\lambda_0\left(\frac{\partial u}{\partial y}\right)_0 + \frac{C_s v_0}{T_0}\left(\frac{\partial T}{\partial x}\right)_0 \text{ (velocity slip)} \tag{32}$$

$$T_0 - T_w = C_t\lambda_0\left(\frac{\partial T}{\partial y}\right)_0 \text{ (temparature jump)} \tag{33}$$

solid surface boundary conditions at a plane wall ($y = 0$), with x and velocity u in the direction of the flow, where the subscript 0 indicates that the quantities are evaluated in the gas at $y = 0$. Currently accepted kinetic theory values for the coefficients are $C_m = 1.14$, $C_s = 1.1.7$, and $C_t = 2.18$, for diffuse reflection in a monatomic gas. The first term in Eq. (31) is the velocity slip due to finite wall velocity gradient, while the second is the thermal creep term arising from a temperature gradient in the gas in the direction of flow. Thermal creep is responsible for the phenomenon of thermophoresis, where small suspended particles experience a force in the direction opposite to a temperature gradient (although the phenomenon persists throughout the entire range of rarefaction, to the collisionless limit). Slip and temperature jump have the effects of reducing skin friction and heat transfer, but these effects are often obscured by other low Reynolds number and compressibility effects, and it has been found in general that slip flow corrections to the Navier-Stokes equation have a very limited range of applicability.

Transition flow. Numerous attempts have been made to develop higher-order continuum equations to replace the Navier-Stokes equations for regimes of

moderate rarefaction. The results, which for the most part have been based on expansions of the Boltzmann equation, such as the Chapman-Enskog method, have proved not to be measurably superior to the Navier-Stokes equations. More success has been achieved through the use of simple interpolation between the slip flow and near-collisionless flow limits, or by solving the Boltzmann equation or models of it, or by direct simulation. This last method appears to be the most powerful and versatile of those mentioned, and will be described briefly.

The direct-simulation method models a real gas flow by computing the trajectories and positions of thousands of simulated molecules, and storing their positions and velocities as a function of time. With each time step, the molecules are advanced along their current trajectories, representative collisions with other molecules or with solid surfaces are calculated, the new velocities and positions are stored, and the process is repeated for the next time step. The flow is always unsteady, since at time $t = 0$ the molecules are introduced into the flow field in an arbitrary fashion, but a steady flow field is obtained as the long time average limit of the unsteady flow. The method has proved successful in modeling shock-wave structure, the free jet expansion of a gas into a vacuum, multidimensional flows about bodies, and polyatomic and chemically reacting gas flows. *See* SIMULATION. Lawrence Talbot

Bibliography. M. H. Aksel and O. C. Eralp, *Gas Dynamics*, 1993; J. D. Anderson, *Modern Compressible Flow: With Historical Perspective*, 3d ed., 2003; G. A. Bird, *Molecular Gas Dynamics and the Simulation of Gas Flows*, 1994; A. B. Cambel and B. H. Jennings, *Gas Dynamics*, 1958; S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases*, 3d ed., 1970, reprint 1990; J. E. A. John, *Gas Dynamics*, 2d ed., 1984; A. H. Shapiro, *The Dynamics and Thermodynamics of Compressible Flow*, vol. 1, 1953, vol. 2, 1954, reprint 1983; W. G. Vincenti and C. H. Kruger, *Introduction to Physical Gas Dynamics,* 1965, reprint 1975.

Gas field and gas well

The term gas field refers to a geographical area which is underlain by one or more commercial reservoirs of petroleum. This commercially valuable gas, primarily methane but with smaller amounts of ethane, propane, and butanes and in some cases containing significant concentrations of carbon dioxide, hydrogen sulfide, and nitrogen, is produced through wells which penetrate subterranean reservoirs, composed of porous rock. The gas in the reservoir may be free or dissolved in crude oil.

About one-third of the gross production of natural gas in the United States is produced from reservoirs in which there is no substantial amount of crude oil (mixture of higher-molecular-weight pentanes plus hydrocarbons) in contact with the gas. The gas in such reservoirs is called nonassociated gas. Associated gas is that produced from reservoirs in which there is a substantial contact of free gas (in the gas cap) with crude oil or gas dissolved in crude oil.

Nonassociated gas is produced as a result of volumetric expansion of the gas as the reservoir pressure is decreased from its original value down to the economically determined abandonment pressure. The recovery efficiency from such volumetric reservoirs is very high. The flow of associated gas through a reservoir and ultimately into the bore hole of a producing well is complicated by the fact that the two fluids are flowing simultaneously and essentially competing to reach the well bore. In fact, it is the release of gas from solution and the subsequent expansion of the free gas within the reservoir which provides the energy to drive the oil from solution gas reservoirs. It is usually necessary to restrict the production of natural gas from associated gas reservoirs in order to maximize oil production. *See* NATURAL GAS; OIL AND GAS FIELD EXPLOITATION; OIL AND GAS WELL DRILLING; PETROLEUM; PETROLEUM GEOLOGY; PETROLEUM RESERVOIR ENGINEERING. Todd M. Doscher

Bibliography. N. J. Hyne, *Nontechnical Guide to Petroleum Geology, Exploration, Drilling and Production*, 2d ed., 2001; W. C. Lyons, *Standard Handbook of Petroleum and Natural Gas Engineering*, 2d ed., 2004.

Gas furnace

An enclosure in which a gaseous fuel is burned. Domestic heating systems may have gas furnaces. Some industrial power plants are fired with gases that remain as a by-product of other plant processes. Utility power stations may use gas as an alternate fuel to oil or coal, depending on relative cost and availability. Some heating processes are carried out in gas-fired furnaces. *See* STEAM-GENERATING FURNACE.

Among the gaseous fuels are natural gas, producer gas from coal, blast furnace gas, and liquefied petroleum gases such as propane and butane. Crude industrial heating gases carry impurities that corrode or clog pipes and burners. Solid or liquid suspensoids are removed by cyclones or electrostatic precipitators; gaseous impurities are removed chemically. The cleaned gas may be mixed with air in the furnace, in the burners, or in a blower before going to the burners. The gas and air may be supplied at moderate pressure, or one or both at high pressure. The high-pressure component may serve to induce the other component into the furnace. The burner may be a single center-fire type or a multispud type with numerous small gas parts, depending on how the heat is to be concentrated or distributed in the furnace. Crude uncleaned gases are fired through burners with large ports, the burners being removable for cleaning. *See* FUEL GAS. Ralph M. Hardgrove

Gas thermometry

A method of measuring temperatures with gas as the thermometric fluid. Gas thermometry is the primary source of information about a fundamental physical parameter, temperature, over the range from about 3 to 900 K (−454 to 1160°F).

In principle, gas thermometry consists of using the ideal gas law, Eq. (1), where P is the pressure, V the

$$PV = nRT \tag{1}$$

volume, n the number of moles of gas, R the molar gas constant, and T the thermodynamic temperature, to evaluate an unknown temperature by reference to the single defining temperature of the Kelvin thermodynamic temperature scale, namely the triple-point temperature of water. This reference temperature, achievable within about one part in 10^7 by standard laboratory practice, has the value 273.16 K (0.01°C or 32.018°F). Determination of the unknown temperature requires two sets of measurements of the pressure P and the volume V of n moles of an ideal gas. One set of measurements usually is performed while the gas is maintained at the water triple-point temperature; the second set is obtained with the gas at the unknown temperature. The unknown temperature can readily be evaluated by rewriting Eq. (1) as $n = (PV)/(RT)$, which has the same value for both sets of measurements. The desired result is given by Eq. (2), where the primed quantities refer to the un-

$$T' = \frac{273.16(P'V')}{(PV)} \tag{2}$$

known temperature. In Eq. (2) the precise values of the gas constant R and of n need not be considered. *See* GAS; TRIPLE POINT.

Unfortunately, no real substance obeys the ideal gas law over any substantial range of temperature. Furthermore, the measurement of gas pressure and gas volume at the part-per-million level requires painstaking effort, as does the successful manipulation of measured quantities of high-purity gases.

Several different techniques have been developed to minimize the uncertainties arising from gas nonideality and from difficult measurements. These techniques include constant-volume gas thermometry, constant-pressure gas thermometry, constant-bulb-temperature gas thermometry, and isothermal gas thermometry.

Constant-volume gas thermometry. This is by far the most commonly used of the gas-thermometry methods. The name is somewhat misleading because no gas bulb truly exhibits constant volume over any substantial temperature range. Several steps are involved: inserting a fixed mass of a working gas into a rigid container or bulb; determining the pressure of the gas at the triple-point temperature of water; heating or cooling the container to a new temperature whose value is to be determined; and measuring the gas pressure at the new temperature.

Uncertainties. For best accuracy, many details in the construction and operation of the gas thermometer must be carefully pursued. These include accurate measurement of the working pressure; correction for the deviation from ideality of the working gas; minimization of impurities in the working gas, determination of the nature and concentration of the remaining impurities, and correction for their effects; determination of the change of volume of the gas bulb between the two operating temperatures; correction for the quantity of working gas that is not maintained at the working temperature (the so-called dead-space correction); correction for hydrostatic pressure effects; correction for the thermomolecular pressure effect; and evaluation of nonuniformity of the working temperature. Whenever mercury manometry is employed for gas-thermometer pressure measurements, as is often the case, additional uncertainties must be considered, particularly concerning the density of the mercury and the acceleration due to gravity at the measuring site. *See* MANOMETER.

In the determination of temperatures above the triple-point temperature of water, uncertainties arising from most sources are exaggerated approximately in the ratio of the Kelvin temperatures involved. The reverse is true for gas thermometry below room temperature.

The working gas used in most modern gas thermometry is helium, isotopically enriched to nearly pure ^{4}He. This gas is essentially inert chemically. Its departure from ideal thermodynamic behavior can be described by the virial equation (3). For work at

$$PV = nRT(1 + BP + CP^2 + \cdots) \tag{3}$$

the moderate pressures of 100 kilopascals (1 atmosphere) or less, usually used in gas thermometry, only B, the second virial coefficient, need be known. *See* VIRIAL EQUATION.

Temperature evaluation. The evaluation of unknown temperatures by constant-volume gas thermometry generally proceeds by equating the total number of moles of working gas at the two temperatures, as in Eq. (4). Here, the term with subscript b' ac-

$$n' = \left[\frac{PV}{RT[1+BP]}\right]_{b'} + \left[\frac{PV}{RT[1+BP]}\right]_{d'}$$
$$= n = \left[\frac{PV}{RT[1+BP]}\right]_{b} + \left[\frac{PV}{RT[1+BP]}\right]_{d} \tag{4}$$

counts for the gas in the bulb at the unknown temperature. The term with subscript d' describes gas contained in the dead space during the unknown-temperature measurements; usually the dead-space gas is distributed over a range of temperatures between the laboratory temperature and the unknown temperature. The term with subscript b refers to gas contained in the gas bulb at 273.16 K. The term with subscript d refers to gas contained in the dead space during the triple-point measurements. Equation (4) indicates the inclusion of nonideality and dead-space effects; besides these, the effects of gas-bulb thermal

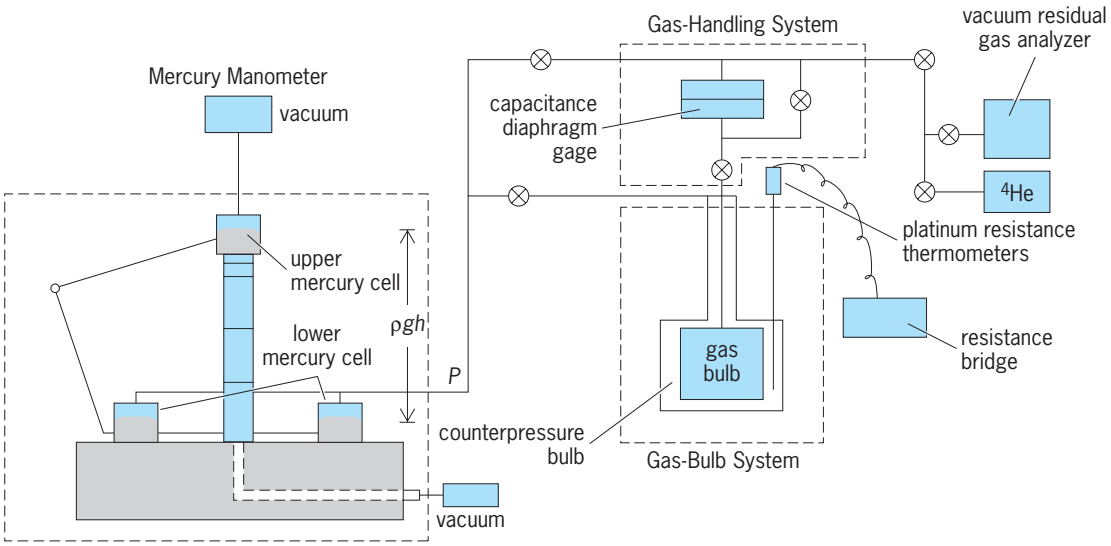

Typical constant-volume gas thermometer. (*After J. F. Schooley, Thermometry, CRC Press, 1986*)

expansion and the other problems mentioned above must be made part of the calculations.

Instrumentation. A typical constant-volume gas thermometer (see **illus.**) includes a mercury manometer, used for the measurement of pressure; a gas-bulb system; and a gas-handling system. The part-per-million accuracy achievable in manometric pressure measurements is central to the thermometric accuracy of this instrument. The entire manometer is operated in a temperature-controlled environment. The distance between the surface of the mercury in an upper cell and mercury surfaces in two lower cells is measured by the use of wrung stacks of calibrated end gage blocks. Axial holes through the gage blocks permit detection of the quality of the wringing process by measurement of the internal vacuum of the stack. The pressure exerted by the column of mercury is given by the product ρgh, where ρ is the density of the mercury, g is the acceleration due to gravity at the manometer, and h is the height of the gage-block stack.

The gas bulb is completely enclosed by a second bulb in which a so-called counterpressure of helium gas equal to the gas-bulb pressure is maintained at all times. The counterpressure gas minimizes pressure-induced changes in the gas-bulb volume and helps to reduce contamination of the working gas from impurities in the gas-bulb thermostat.

Isothermal gas thermometry. It is possible to determine the thermodynamic temperature of a gas bulb by repeatedly adding measured quantities of a working gas to it and measuring the resulting pressures. The repeated measurements at a single temperature are known as an isotherm. The virial equation (3) can be used to obtain both the unknown gas-bulb temperature and the value of the second virial coefficient.

Two variations on the technique can be used, the absolute isotherm and the relative isotherm. In absolute isothermal gas thermometry, measured quantities of working gas are introduced into a gas bulb at the unknown temperature from a known volume that is maintained at 273.16 K. In the relative isotherm method, the working gas is added stepwise to a gas bulb while it is maintained at another, more convenient reference temperature. *See* LOW-TEMPERATURE THERMOMETRY; PHYSICAL MEASUREMENT; TEMPERATURE MEASUREMENT.

James F. Schooley

Bibliography. T. D. McGee, *Principles and Methods of Temperature Measurement*, 1988; T. J. Quinn, *Temperature*, 2d ed., 1991; J. F. Schooley, NBS/NIST gas thermometry from 0 to 660°, *J. Res. Nat. Inst. Stand. Tech.*, 95:255–290, 1990; J. F. Schooley (ed.), *Temperature, Its Measurement and Control in Science and Industry*, vol. 5, 1982; J. F. Schooley, *Thermometry*, 1986.

Gas tube

An electron tube, usually a diode (see **illus.**) or triode, containing a gas. The gas contributes two significant advantages: (1) The plasma, or gas of electrons and ions, that is formed in the gas during conduction allows space-charge effects to be overcome; and (2) the gas acts to reduce power dissipation at electrodes. These factors result in tubes that can conduct high currents and remain in use for long periods. Gas tubes are particularly important for high-power applications. Two important types are the thyratron and the gas-filled rectifier.

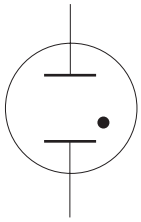

Symbol for a gas-filled diode.

Voltage and current characteristics. The voltage and current characteristics of a gas tube differ from those of a vacuum tube. In a vacuum tube, quasifree charge, or space charge, is present. The power dissipated in a vacuum tube (which is the product of the voltage drop times the current) is usually larger because of the electric fields produced by space charge. This is described by the Child-Langmuir law, which states that the current in a vacuum diode is proportional to the voltage to the $^3/_2$ power. In addition, the charges between electrodes acquire significant energy from the applied voltage and release this energy when they strike the electrode. These two means of energy loss are reduced in a gas tube for the following reasons. As electrical breakdown occurs in a gas tube, a plasma is formed. A plasma is a mixture of electrons and ions that may be electrically neutral overall, because the electrons are provided by the ions, and thus the net charge is balanced. When a potential is applied, the positive charges move in a direction opposite to that of the electrons. Thus, it is possible to overcome the obstacle of the large electric fields due to space charge, and provide more charges besides, by producing a plasma. Most current is carried by the electrons, because their mass is much less than even the lightest of positive ions, and they move much faster than a heavy ion in the same potential field. The gas also acts to slow down electrons because of collisions between electrons and un-ionized atoms or molecules within the plasma. This reduces energy loss at the electrodes. *See* ELECTRICAL CONDUCTION IN GASES; SPACE CHARGE; VACUUM TUBE.

Most of the potential fall in a tube occurs at the interface between the plasma and cathode. This region includes the cathode fall, which is a thin sheath that has a potential drop that corresponds to the energy required to ionize the gas in the tube. Thus, electrons are emitted from the cathode (requiring energy), and accelerated through a potential until they have sufficient energy to begin ionizing the gas that is present. This energy varies, depending on the ionization potential of the gas. *See* IONIZATION POTENTIAL.

Choice of gas. The gas is chosen with regard given to hold-off voltage and current-handling capabilities, as well as other factors related to specific applications, such as rate of rise of current (turn-on time), which is important for some pulsed applications. For example, mercury and xenon are used for high-current applications, and hydrogen is used for high-voltage applications.

Thermionic cathode. Many gas tubes use a thermionic or hot cathode, that is, a cathode that is heated and produces current through emission from a low-work-function material such as barium oxide. The development of improved cathodes remains an important technological problem. A good thermionic cathode allows a tube to operate for long periods with excellent repeatability and good current-handling capability. *See* THERMIONIC EMISSION.

Applications. Solid-state devices have replaced vacuum and gas tubes for many applications. Among the important remaining applications for gas tubes are high-power switching and pulsed-power applications requiring high power and rapid closing, or turn-on. For example, the hydrogen thyratron is still used as a radar modulator switch, and is also applied to the switching of high-power lasers. *See* LASER; RADAR.

Types of tubes. The hydrogen thyratron is a triode that operates at hydrogen pressures between 300 and 900 millitorr (40 and 120 pascals). A plasma is formed, and a glow discharge may be seen in a glass-enclosed tube during the closed or on phase. Hydrogen thyratrons are rated for stand-off voltages up to about 50 kV, and peak currents up to about 5000 A (off the shelf). They also close, or turn on, rapidly. They use a thermionic cathode and have a reservoir of hydrogen within the tube. They achieve a high stand-off voltage through an interesting "trick." The grid is positioned very near the anode—about 0.1 in. (2 or 3 mm)—rather than away from it. The high stand-off voltage occurs because the average distance that an accelerated electron travels is comparable to the grid-anode separation, thus reducing the ionization of the gas, and therefore preventing breakdown. The stand-off voltage in one hydrogen thyratron may be 20 kV, but that tube can be switched on with only 100 to 200 V, applied between the grid and cathode. *See* BREAKDOWN POTENTIAL; ELECTRICAL BREAKDOWN; GLOW DISCHARGE.

Other gas-phase switches include the ignitron, where a pool of mercury serves as a cathode and mercury vapor is the conducting gas; the liquid-metal plasma tube, where a small amount of mercury is distributed over a metal cathode; the crossed-field tube, using a magnetic field to bend the paths of electrons; spark gaps; mercury rectifying tubes; and gas voltage regulators, which take advantage of a constant voltage across a glow discharge to clamp the voltage of an erratic source.

The spark gap has applications for very high voltages and currents. It operates in an arc mode, rather than a glow-discharge mode. The gas may be at atmospheric pressure, and gases such as nitrogen and helium are used. The spark gap tends to suffer from electrode degradation caused by arcing, but is able to stand off higher voltages and conduct higher currents than other gas-phase switches. *See* ARC DISCHARGE; SPARK GAP.

Outlook. Gas tubes are expected to retain their importance for higher-power applications, for example, voltages exceeding roughly 5 or 10 kV, along with high currents and fast turn-on. They are resilient to most perturbations. However, the continuing development of solid-state devices is expected to have increasing effect in this area as well: the thyristor (a *pnpn* semiconductor) has already made significant inroads. *See* SEMICONDUCTOR RECTIFIER.

Martin Gundersen

Bibliography. R. W. Crompton et al. (eds.), *Gaseous Electronics and Its Applications*, 1991; Institute of Electrical and Electronics Engineers, *IEEE Guide for the Application of Gas Tube and Air Gap Arrester Low-Voltage (Equal to or Less than 1000 Vrms*

or 1200 Vdc) Surge-Protective Devices, 1993; J. A. Kunc and M. Gundersen, A fundamental theory for high-power thyratrons, I and II, *Lasers and Particle Beams*, vol. 1, pt. 4, pp. 395–425, 1983.

Gas turbine

One of a class of heat engines which use fuel energy to produce mechanical output power, either as torque through a rotating shaft (industrial gas turbines) or as jet power in the form of velocity through an exhaust nozzle (aircraft jet engines). The fuel energy is added to the working substance, which is gaseous in form and most often air, either by direct internal combustion or indirectly through a heat exchanger. The heated working substance, air comixed with combustion products in the usual case of internal combustion, acts on a continuously rotating turbine to produce power. The gas turbine is thus distinguished from heat engine types where the working substance produces mechanical power by acting intermittently on an enclosed piston, and from steam turbine engines where the working substance is water in liquid and vapor form. *See* INTERNAL COMBUSTION ENGINE; STEAM TURBINE.

All the variations of the gas turbine discussed below involve the transfer of heat to and from the gaseous working substance for the purpose of enhancing overall performance.

Air cycle. Gas turbine engines depend on the principle of the air cycle, just as their piston engine counterparts do, but they use a different succession of processes. In the basic simple cycle (**Fig. 1**), ambient air is first compressed to a maximum pressure level, at which point fuel heat energy is added to raise its temperature, also to a maximum level. The air is then expanded from high to low pressure through a turbine. The expansion process through the turbine extracts energy from the air, while the compression process requires energy input.

As the air moves through the engine, the turbine continuously provides energy sufficient to drive the compressor. In addition, because the turbine expansion process starts from a high temperature that comes from the fuel energy released by combustion, surplus energy beyond that required for compression can be extracted from the air by further expansion. At the point where the turbine has provided sufficient energy to power the compressor, the air pressure remains higher than the outside ambient level. This higher pressure represents available energy in the air that can be turned into useful output power by a final expansion process that returns the air pressure to ambient. The exhaust air leaves the engine with pressure equal to the outside, but at a higher temperature. As with any heat engine, the high exhaust temperature represents wasted energy that will dissipate into the outside atmosphere. *See* COMPRESSOR.

From an energy accounting standpoint, the sequence of processes acting on the air from front to rear constitutes a full cycle. It starts with the outside air entering at its initial state, and is completed when the air returns again to both ambient pressure and temperature levels. The series of cycle processes includes the final outside dissipation of the wasted exhaust energy, inevitable for every heat engine according to Carnot's principle. The ideal version of the gas turbine cycle is known as the Brayton cycle. *See* BRAYTON CYCLE; CARNOT CYCLE.

For any completed cycle, the total energy added from the fuel sources will always be equal to the sum of the useful output energy and the wasted exhaust energy. The thermal efficiency, which is the ratio of net output energy to fuel input energy for the cycle, measures the engine's ability to minimize wasted energy. A thermal efficiency of 60% means that for every 100 units of added energy 60 units will be available as useful output while 40 units will leave the engine as high-temperature exhaust.

Another performance measure is the specific power, which is the ratio of output power to quantity of working substance mass flow rate. Gas turbine engines, in comparison with other types of heat engines, are characterized not only by high levels of efficiency but also by very high levels of specific power. They are especially useful for applications that need compact power.

Simple cycle engines. By far the most common mechanical arrangement for the gas turbine is an in-line axial flow positioning of all components (**Fig. 2**). In the ground-based engine, the inlet at the front guides the incoming air into the compressor, which in turn delivers high-pressure air into the combustor section. The combustor burns the injected fuel at a high reaction temperature, using some of the air itself as an oxygen source. The combustion products in the combustor mix with the remaining unused air to reach a uniform equilibrium temperature, still high but diluted down from the reaction temperature. The hot, high-pressure combustor exit air enters the compressor drive turbine, where it expands down in pressure toward, but staying higher than, ambient level. This expansion process results in output shaft power that can be delivered directly

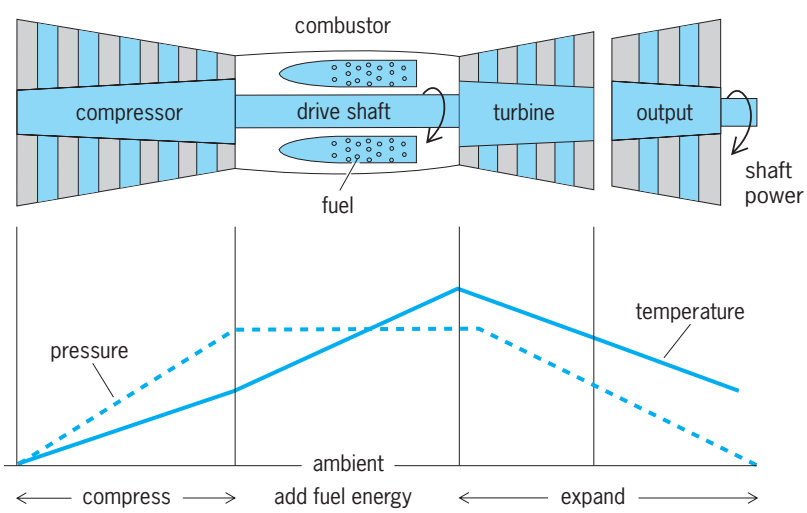

Fig. 1. Basic gas turbine cycle.

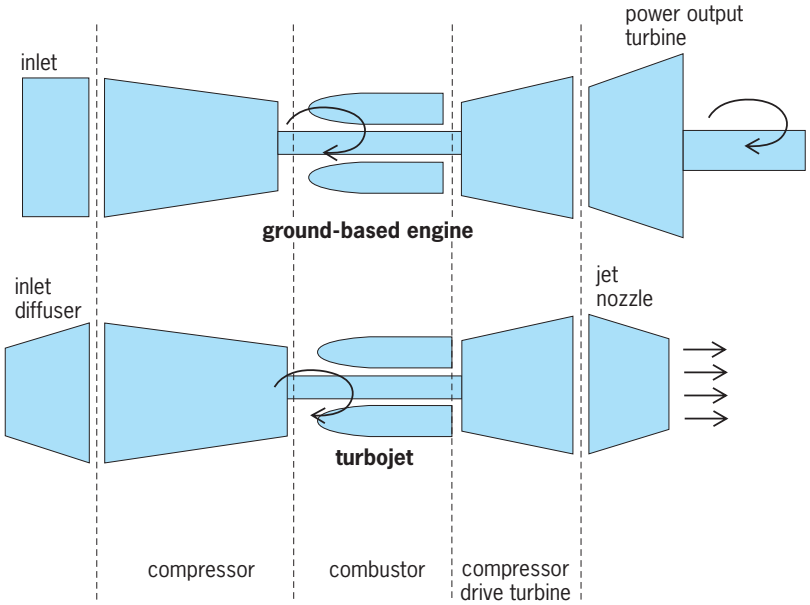

Fig. 2. Simple gas turbine component arrangements.

to the compressor through a connecting rotating shaft. Starting from the exit of the compressor drive turbine, net output power remains available. This power can be realized through the process of further pressure expansion completely down to the ambient level. For ground-based applications, the final expansion takes place through a power turbine whose output shaft is connected to the external load. In the single-spool arrangement the power turbine and compressor drive turbine are indistinguishably combined into one unit which, together with the compressor and the output load, is connected to a common shaft. For aircraft applications, either a power turbine extracts useful power to drive a propeller through a separate shaft (turboprop), or the expansion process takes place through a nozzle which acts to convert some of the thermal energy into velocity energy to be used for jet propulsion. *See* AIRCRAFT ENGINE; JET PROPULSION.

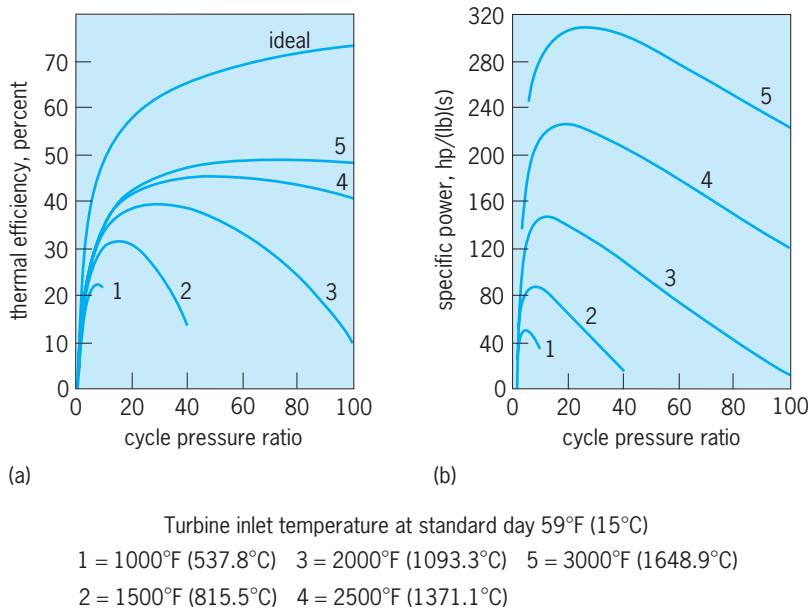

Fig. 3. Simple cycle performance. (*a*) Thermal efficiency versus cycle pressure ratio and (*b*) specific power versus cycle pressure ratio, which is the ratio of maximum pressure to ambient pressure.

Design performance. For the cycle composed of ideal (100% efficient) components with no friction losses, thermal efficiency increases continuously with increased maximum pressure, expressed as a ratio to ambient pressure. Real components, however, temper this ideal trend. When cycle pressure ratio exceeds some value, the deleterious effects of component aerodynamic inefficiencies begin to predominate. This causes thermal efficiency to reach a maximum level at some pressure-ratio value and then fall off, depending on the levels of both the maximum cycle (turbine inlet) temperature and the component efficiencies. Increasing the level of maximum turbine temperature acts to offset the effect of energy waste due to component inefficiencies, as manifested by a shift in trend curves toward the ideal (**Fig. 3***a*). Specific power output increases directly with turbine entrance temperature and peaks with increasing pressure ratio, depending on temperature level (Fig. 3*b*).

Component elements. **Figure 4** shows a typical gas turbine engine in exploded view to reveal the individual component elements in a dual-spool arrangement.

Compressors. Axial-flow compressors are composed of a series of peripheral vanes designed to impart a swirling velocity to the incoming air as air density is increased. The individual vanes are shaped as airfoils to most efficiently turn the air away from the axial direction. They are arrayed in rows surrounding, and held in place by, rotating discs. The simultaneous turning and pressurizing of the air exerts a torque resistance that is overcome by an input shaft connected to the disc. Behind each row of rotating vanes, a row of stationary blades is emplaced that acts to straighten the swirling flow while increasing its pressure. This action of converting swirl velocity energy to air-pressure increase by means of a mechanical rotor enables continuously flowing air to be compressed. It is the central operating principle that makes feasible the gas turbine engine as a continuous power-producing machine.

Compressors are usually made up of several rows of alternating rotor vanes and stator blades, known as stages. By using multiple stages in tandem, the pressure of the air can be gradually increased by small increments, resulting in the most aerodynamically efficient process for achieving high overall pressure ratios. At the very highest of pressure ratios, the compressor is split into two or even three separate inline compressors. Each is connected to its own shaft, nested one inside the other, and driven by individual turbines. The separate compressor and turbine set with its drive shaft is known as a spool. Placing two or more spools in series enables each group of compressor stages to operate near its own most aerodynamically efficient rotating speed. Progressing from front to rear, individual stages inherently

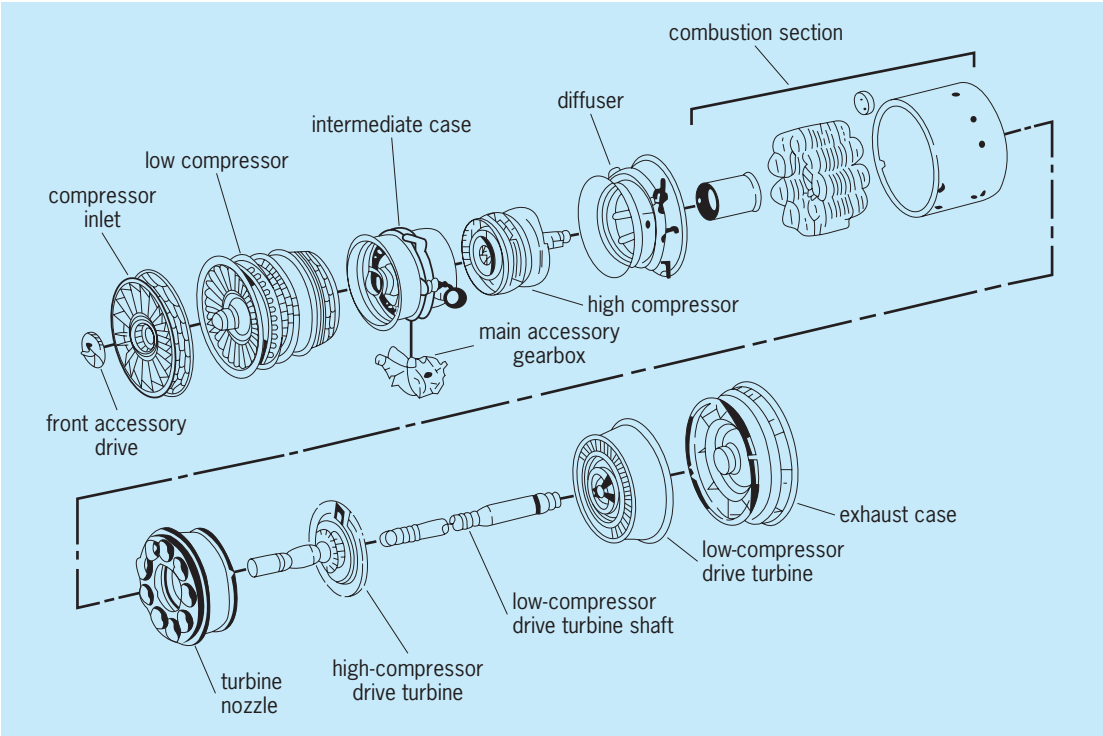

Fig. 4. Typical dual-spool gas turbine components.

function much better when free to travel at increasingly faster speeds.

Radial-flow compressors are most often found in very small gas turbine engines. In contrast to axial flow, the radial-flow, or centrifugal, compressor achieves air compression by imparting swirl to the air as it moves from the center of the rotor to its periphery. At the outer portion the swirling air energy is converted to increased pressure by means of fixed vanes that act to diffuse the air to lower velocity. Inherently, because the swirl process takes place while the air moves outward in radius, radial-flow compressors tend to present much greater frontal diameters for the same airflow quantity. As an offsetting advantage, they are simpler in design and less expensive to manufacture. This is especially important in the smaller sizes where radial-flow vanes are inherently less sensitive to those aerodynamic efficiency losses that must be controlled by stringent manufacturing tolerances in axial-flow airfoils.

A few gas turbines employ hybrid, or mixed-flow, compressors that, on a single spool, attach an axial-flow compressor directly to a radial-flow compressor.

Combustors. The gas turbine combustor is designed for simplicity and compactness while providing for efficient, stable combustion with low emissions. Clean combustion products always come with efficient burning. Combustion stability and efficiency are achieved primarily by significantly reducing the air velocity from the relatively fast level at which it leaves the compressor to a very slow level with high turbulence. An increasing-area diffusion duct placed between the two components acts to reduce air velocity. At the end of the diffuser, atomizing fuel nozzles spray a fuel mist into the low-velocity airstream where a high temperature flame front is established. To feed the combustion zone with oxygen, metal liners, pierced by a succession of small holes, direct the air inward toward the center of the combustion chamber where the spray nozzles are located. The rest of the air, entering at the cooler compressor exit temperature around the outside of the combustion zone, is forced radially inward by another series of small liner holes designed for thermal mixing. The process of mixing cold with hot continues downstream until the overall gas temperature distribution is sufficiently devoid of hot peaks to enter the turbine. *See* COMBUSTION CHAMBER.

Turbines. Axial-flow turbines are most commonly employed. Turbine operation is the reverse of that of compressors. Turbines are also composed of a series of stationary vane and rotating blade pairs, or stages, designed in the form of airfoils. However, the stationary turbine vanes act first to convert the thermal energy of the air into rotating swirl velocity as density decreases. The swirling air leaving the vanes impinges on the traveling airfoil blades and imparts a resisting torque to the vanes. The rotating airfoils change the swirl back to the axial direction while reducing air pressure in the process. Their torque travels through the discs holding the blades to the connecting output shaft that is attached to the resisting, power-absorbing load on its other end. In their most elementary version, turbines act just like simple pinwheels.

Radial in-flow turbines are usually coupled with radial-flow compressors. This accommodates the

large diameters inherent to both in a back-to-back arrangement. Radial-flow turbines reverse the operating principle of their compressor counterparts. Surrounding the periphery, high-temperature air is expanded to high velocity in a series of nozzles that directs swirling air inward toward the rotating vanes. As the air moves toward the center, the vanes expand the air to lower pressure while its swirl decreases. The air with reducing swirl exerts a tangential force on the vanes that appears as torque acting on the connecting shaft.

Critical to the progression of gas turbine technology, turbine airfoils must withstand increasingly high temperatures along with strong centrifugal and bending stresses. They are made from metal alloys with protective coatings, both especially developed for aircraft engines, and they employ interior cooling. Complex internal passages circulate cooler air drawn from the compressor upstream through the inside of the turbine airfoil. Small, laser-drilled holes lead the inside air through blade surfaces into the outside main stream. *See* TURBINE.

Diffusers/nozzles. Air that travels through ducting responds in velocity and pressure to changes in area along the duct. If the area gets smaller, pressure goes down while velocity speeds up; if the area gets bigger, pressure increases while velocity slows down. Area-increase conduits, called diffusers, are integral gas turbine components. A diffuser is used to slow down the compressor air entering the combustor. In aircraft gas turbines a diffuser is also placed in front of the compressor to slow down the captured outside air from forward-flight velocity to a velocity that is aerodynamically compatible with the front compressor airfoils. The consequent pressure increase in the inlet diffuser, or ram pressure, contributes to the overall cycle pressure ratio. This means that to achieve the same maximum pressure ratio, flight gas turbines powering high-speed airplanes do not need as high a pressure ratio from the compressor as do low-speed ones. However, supersonic aircraft must be equipped with complex inlet diffusers to achieve efficient ram-pressure increases.

Area-reduction conduits are called nozzles. Nozzles in flight gas turbines find use in the exhaust section. The pressure level exiting the compressor drive turbine expands through the engine exhaust nozzle down to ambient pressure while the jet velocity increases to provide jet propulsion.

Heat exchangers. The components described above are found in all gas turbines from the simple to the very complex. The more complex cycles use heat exchangers either to remove heat to cool the inside air, using an outside cooling source, or to add heat from a hotter outside source. In gas turbines the external cooling sink usually is a nearby body of water such as a lake or a river. In the water-air case a shell-and-tube arrangement is most practical and economical. In the case of recuperators or regenerators (see below), the heat exchange normally takes place between two separate streams of air. In this instance the preferred heat exchanger would be the plate-fin type. As a rule, cycles that employ heat exchangers to improve output and efficiency are limited to only those special power plant applications that warrant their use on the basis of overall economics. Air, even in a compressed state, has such relatively low density that the size and cost of the heat exchanger alone can outweigh any performance advantages. *See* HEAT EXCHANGER.

Gas turbine variations. A number of gas turbine engines have been developed that use variations of the simple cycle system.

Closed versus open systems. A gas turbine's working substance can either pass through the engine continuously with its exiting exhaust products replaced by the incoming air (open), or it can recycle the air (closed). Because there will always be wasted exhaust energy, causing the gas to leave the turbine at the rear with a higher temperature than that of the air entering the compressor at the front, closed cycles require a cooling heat exchanger to be placed between those two components. This heat exchanger serves to dissipate the wasted exhaust heat into the outside atmosphere. It normally uses circulating water as the coolant and a cooling tower to dispose of the water's heat to the outdoors. The gas-cooled nuclear reactor system represents a prime example of a closed system. Here the working substance is helium gas. Helium passing through the reactor is impervious to radioactivity. An advantage of all closed systems is their ability to pressurize the fully contained gas. The high density reduces the size of all components.

Regenerative cycle. With a recuperating heat exchanger placed behind the turbine, some of the wasted exhaust energy can be recovered. This is accomplished by directing the compressor exit air through one side of a heat recovery unit while passing the turbine exhaust through the other side. The compressed air is externally ducted to the rear of the engine, receives the exhaust heat from the turbine side in a heat exchanger, then is returned to reenter at the front of the combustor. The higher-temperature air entering the combustor requires less fuel to reach the same maximum cycle temperature and output power (**Fig. 5**).

Intercooled cycle. Advantage can be taken of a basic physical principle governing the air-compression process; that is, the power needed to compress is directly proportional to the entrance temperature. When a cooling heat exchanger is placed in the middle of the compressor to lower air temperature at that point, the net work needed to power the back part of the compression process is accordingly reduced. This net reduction in compressor work becomes directly available as useful power that can be added to the output turbine (Fig. 5).

Intercooled-regenerative cycle. The principles that govern regenerative and intercooled cycles can be combined into a single cycle. Here, the two separate principles reinforce each other in such a way that the overall thermal efficiency becomes higher than for either of the two separate cycles (Fig. 5).

Reheat cycle. Strictly for the purpose of increasing output specific power, adding a second combustor

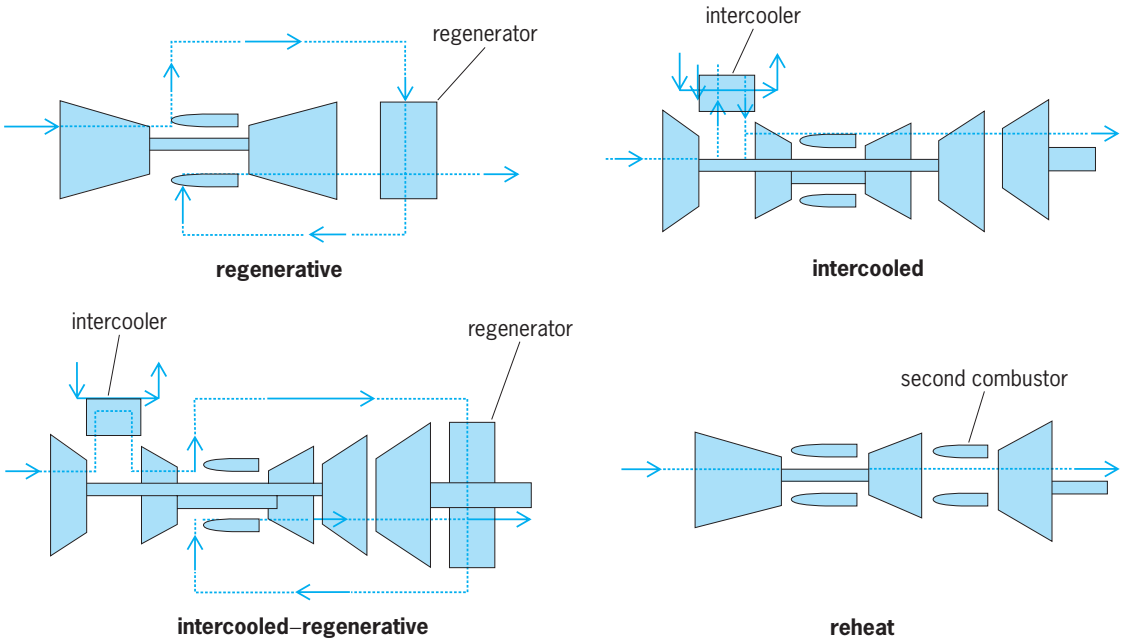

Fig. 5. Gas turbine variations.

behind the turbine allows for extra fuel energy input within the limitations of the turbine to withstand excessive temperature. When the heat addition point is placed behind the compressor drive turbine, the power output turbine can be conveniently located behind the second combustor as a single unit. This arrangement avoids having an internal drive shaft penetrate the hot combustor section from either end. In aircraft applications the second combustor is known as an afterburner. The afterburner is particularly valuable for jet fighters that need short, sudden bursts of thrust power (Fig. 5).

Externally fired system. Although not common, in certain circumstances processes other than internal combustion can be the source of a gas turbine's energy addition (firing). The nature of the fuel source most often dictates the need for external combustion, with a heat exchanger to transfer heat indirectly to the air. The nuclear reaction, for which the air must be shielded from radioactivity, is one example. Coal, unacceptable for internal combustion in a gas turbine due to its ash residue, is another.

Design performance of simple cycle variants. For all of the variants as well as the basic simple cycle, both thermal efficiency and specific power are related strictly to the maximum turbine-inlet temperature (related to local ambient temperature as a ratio) and the maximum pressure (related to local ambient pressure as a ratio). **Figure 6** compares typical efficiency and specific power trends as a function of these key variables. The trends depicted are also related to the level of efficiency and pressure losses and turbine cooling air penalties associated with each component. Figure 6 assumes typical levels.

Industrial applications. The industrial gas turbine industry started in the late 1950s, following the development of aircraft gas turbines. The three applications of industrial gas turbines are generation of electric power, mechanical drive (principally gas or oil pipeline compression), and marine propulsion. Over 90% of gas turbine applications, as measured in megawatts (MW) of power produced, are in electric power generation. The majority of the navies of the world use gas turbines to propel most of their surface ships.

In electric power generation service, gas turbines offer the combined advantages of lower capital cost, shorter installation time, high efficiency, and low environmental impact as compared to steam

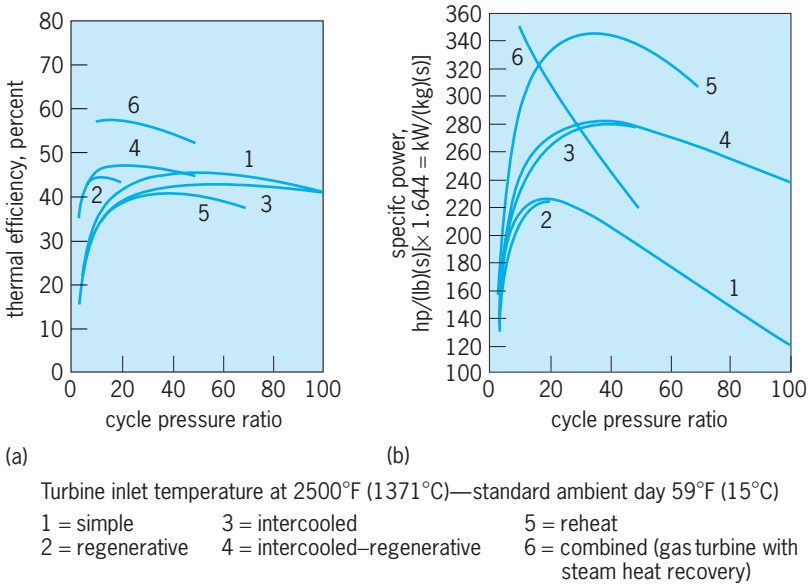

Fig. 6. Typical performance for gas turbine variants, where cycle pressure ratio is the ratio of maximum pressure to ambient pressure.

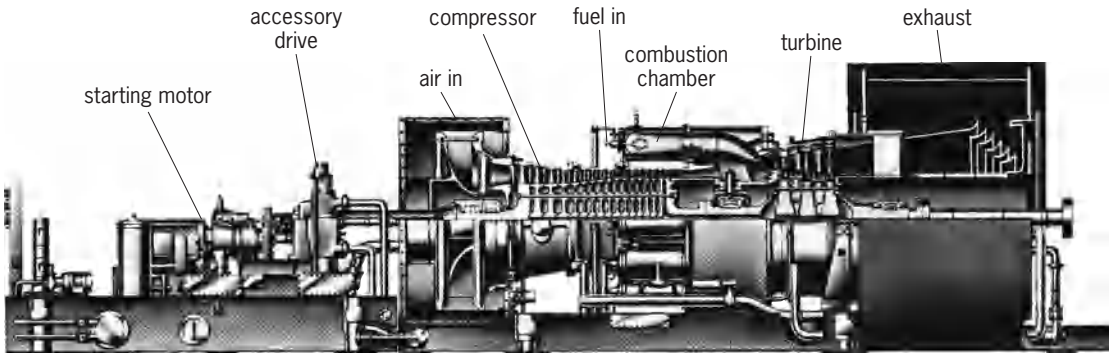

Fig. 7. Industrial gas turbine.

turbine-based power plants. As a result, gas turbine-based power plants have captured a majority of the market for new power generation equipment. The U.S. Energy Information Agency, in its Annual Energy Outlook for 1998, projected that 85% of new power generation equipment to be added worldwide by the year 2020 would be gas turbine-based.

Compliance with emissions regulations, particularly that for nitric oxide, is a major application consideration. In the late 1970s and the 1980s, water or steam was injected into the combustion reaction zone to decrease the flame temperature, which is the principal parameter affecting nitric oxide emission. Present technology is producing combustion system designs that feature premixed air and fuel in lean mixtures to reduce the flame temperature without any outside diluent.

Industrial design. Most gas turbines with outputs above 1–2 MW have multistage axial-flow compressors and turbines. Lower-power units tend to have single-stage centrifugal compressors and radial-inflow turbines to minimize weight, size, and cost. Thermal efficiency improves with larger size, as friction and tip leakages become a smaller percentage of power produced, and multistage axial-flow components become more efficient than radial stages. **Figure 7** shows typical design features for a large frame-type gas turbine designed specifically for industrial application.

The same major components are evident in **Fig. 8**, which shows an aeroderivative gas turbine, that is, one which was derived from an aircraft jet engine. The engine depicted has a twin-spool (dual-compressor) arrangement. Hot gases exiting the gas generator portion drive the power turbine, which is connected to the external load. Typical aeroderivatives differ from frame-type gas turbines by being lighter in weight and having higher design pressure ratios. Their outputs are limited to a maximum of about 50 MW because of the size of the aircraft jet engines from which they are derived. The largest frame-type gas turbines have outputs of over 200 MW. Aeroderivatives tend to produce higher simple-cycle efficiencies, whereas frame types produce the highest combined-cycle efficiencies. For a given output, frame-type gas turbines tend to be somewhat lower in cost than the aeroderivatives.

For applications where the output shaft speed varies, as in a compressor drive, a separate power turbine is needed. This requires a multiple-shaft gas turbine such as shown in Fig. 8. A single-shaft gas turbine has a typical effective operating range of 85–105% of rated speed. With a separate power turbine to drive the load, the output shaft speed range is typically about 50–105% of rated speed.

Typical performance. Component efficiencies, airflow, pressure ratio, and turbine inlet (or firing) temperature are the major factors affecting gas turbine output and efficiency. Typical multistage axial component efficiencies are in the 86–93% range. Material developments and turbine cooling techniques permit turbine rotor inlet temperatures to exceed 2500°F (1370°C) for the most advanced units. The **table** shows typical thermal efficiencies and pressure ratios for various cycles representing engines manufactured and in actual service.

Aircraft engines. Invented simultaneously in England and Germany in the late 1930s expressly for airplane power, the aircraft gas turbine as a jet engine has grown to become the dominant power plant for both

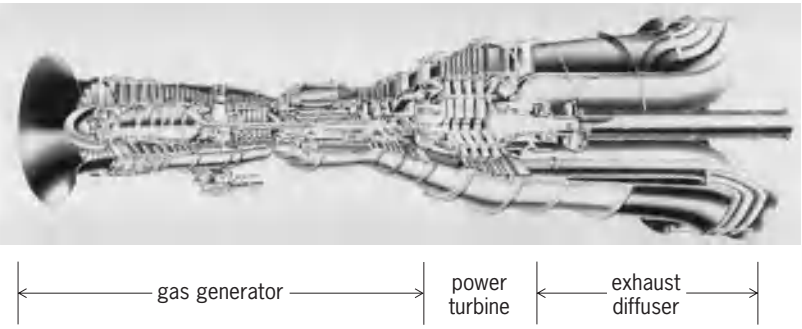

Fig. 8. Aeroderivative gas turbine. (*Pratt & Whitney*)

Typical pressure ratios and thermal efficiencies

Cycle	Pressure ratio	Thermal efficiency, %
Simple	15–30	32–42
Intercooled	40–60	45–49
Regenerative	8–10	37–40
Combined cycle	15–23	50–60

commercial and military aircraft. Its use for ground-based applications started in the 1950s when, with their recognized compact power and quick start-up capability, these engines began to be adapted directly from aircraft engine designs. The use of jet propulsion to replace piston-driven propeller planes allowed airplanes to take an incremental step in flight speed. Jet engines offered two significant advantages. First, the gas turbine's power output for a given size far surpasses that of the piston engine. Second, and most important for flight, jet propulsion overcomes the propeller's most severe limitation. As flight speeds approach the sonic velocity, airplane propellers drastically lose aerodynamic efficiency. Because of their consequent superior speed capability, jet fighters, as they evolved from the early experimental versions in the latter part of World War II, soon replaced the propeller-driven planes with their inherent speed limitation. Subsequently, this became true for commercial passenger flight as well. Even use of the compact gas turbine to power propellers in what is called the turboprop engine is no longer widespread. With the further development of supersonic flight, the jet engine became the only choice as the power plant for these aircraft. *See* TURBOJET.

Jet propulsion is achieved when exhaust gases are ejected at high velocity. As an opposite reaction in force to the engine's pushing the air backward, the air pushes the engine forward. The gas turbine develops the high exit velocity as a result of expansion of the gases in the exhaust nozzle. Leaving the compressor drive turbine at a pressure level higher than ambient, the exiting gases further expand to ambient pressure in the rear nozzle by a process that converts the latent available exhaust thermal energy directly to energy of motion, that is, kinetic energy.

As an evolution of the original simple turbojet, most jet engines are of the bypass turbofan variety. Turbofans more efficiently convert available power to propulsive power. In a sense, the turbofan is a hybrid combination of pure turbojet and propeller engines. After extraction of turbine power sufficient to drive the compressor, the remaining useful energy of the working air is divided in two. First, some additional energy is extracted by further expansion in an additional turbine. This added turbine drives a fan that, like a propeller, adds velocity to air that surrounds and bypasses the engine. By propelling the bypass air rearward, the fan provides a reactive forward thrust force. The remainder of the available energy then further expands in the jet nozzle to produce thrust power. The major difference between the fan and the propeller lies in the external ducting which surrounds the bypass air. The turbofan's ducting contains the inlet air as it is slowed by diffusion to a much lower velocity than the flight velocity. For the air entering the bypass fan as well as the front compressor, this internal diffusion avoids the efficiency degradation that the propeller suffers with high aircraft speeds. *See* TURBOFAN.

Finally, small gas turbines with power turbines connected directly to the lifting rotor are frequently used as power plants for helicopters.

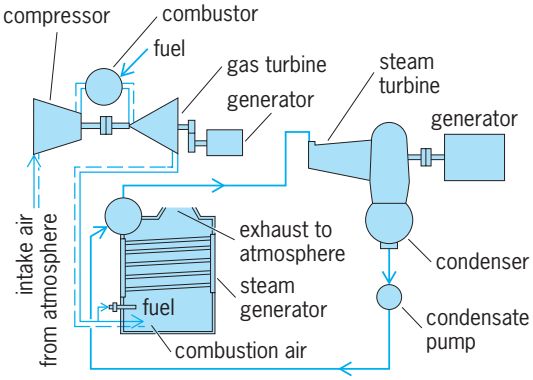

Fig. 9. Combined cycle.

Automotive use. Some of the major automobile manufacturers have experimented with gas turbines for cars and trucks. However, any advantages over the ubiquitous piston engine have not proven sufficient to displace the worldwide infrastructure of garages, gas stations, etc. already in place to support the existing automobile piston engine. However, the U.S. Army's Main Battle Tank (M1A1, Abrams) is powered by a small gas turbine engine. To accommodate the need for variable shaft torque at widely different traveling speeds and climb gradients, these vehicles require a gearing system between the engine's output shaft and the ground traction system. Such would be the case for any road vehicle.

Hybrid power-plant systems. In addition to variations of the simple-cycle gas turbine described in the previous section, hybrids of gas turbines and other power-plant equipment are used in the electric power generation industry. The most common of these is the combined cycle, in which the exhaust gas from the gas turbine is passed through a boiler to make steam for a steam turbine (**Fig. 9**). In some cases, the boiler has supplemental heat from a separate fuel source, as indicated in the figure.

Another industrial variant of the simple cycle is cogeneration. The term refers to the use of the exhaust heat from the gas turbine for some process other than electric power generation. The most common arrangement for cogeneration is to make steam from the exhaust heat, as in a combined cycle, but the steam is sent to the process instead of to a steam turbine. Examples of processes to accept the steam can be heating or domestic hot-water systems for buildings or industrial processes in chemical plants. *See* COGENERATION SYSTEMS.

Much work has been done to adapt gas turbines to coal fuel. Coal gasification, utilized to remove contaminants, is particularly attractive since the gas turbine combined cycle can be integrated into the gasification process for improved efficiency. Air extracted from the gas turbine can be used as the source of oxidant for the gasification process, and the steam produced in the process can be expanded through a steam turbine. **Figure 10** shows how various components relate in an integrated gasification combined

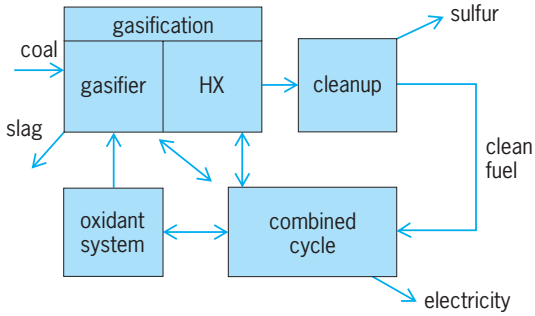

Fig. 10. Integrated coal gasification system.

cycle (IGCC). Petroleum coke is also used as feedstock for an IGCC. In some cases, biomass (for example, wood chips) has been used as fuel for an IGCC.

Industrial fuels. Aviation gas turbines use special jet fuels that come directly from petroleum refineries. Ground-based gas turbines widely use natural gas. It is both clean-burning and easy to transport, if a supply pipeline is nearby. Lower-cost liquid fuels, typically number 2 home heating oil or diesel oil, can be used in industrial applications. Most engines employed for electric power generation are equipped with dual fuel capability, operating on either natural gas or liquid fuel as needed. Such is the case in regions where, during periods of cold weather, home heating takes highest priority for the available natural gas supply, or where fluctuating local natural gas prices cause the use of liquid fuel to be variably more or less economical.

The other two applications for industrial gas turbines use one fuel each: gas from the pipeline for pipeline compressor drives, and liquid fuel for marine propulsion. Some use of various solid fuels has been accomplished in gas turbines, as described above in the section covering hybrid types.

Controls and part-load operating characteristics. Gas turbines characteristically produce smooth and linear throttle response over their entire operating range. Rotor speeds normally vary continuously over this range without the need for the gear shifting and clutch mechanisms found in piston engines. The governing fuel control senses rotor speeds, pressures, and temperatures to maintain stable, steady power or thrust output and, when needed, ensure rapid accelerations and decelerations. The control is programmed, normally by electronic input, to guard against harming the engine during throttle changes by governing the appropriate fuel input rate. Most important, during throttle transients the control functions to prevent turbine overheating, burner blowout, and compressor surge.

The gas turbine output shaft speed can be flexibly adapted to the speed characteristics of a wide variety of types of external loads. While many engines mate with their external loads through a separate shaft, or free power turbine, the flexibility feature holds true even for single-shaft designs. The free turbine arrangement can most efficiently adapt to the widest range of load characteristics. However, in the extreme, single-shaft industrial engines used for electric power generation can directly connect to and drive alternating-current electric generators that must maintain constant speed from idle to full power.

Part-load efficiency normally falls off from the design condition in varying amounts. This fall-off occurs least in the regenerative cycle variant which, for applications such as marine propulsion and automobiles, can offer a compelling advantage. Both ships and ground transportation vehicles, where regeneration is common, tend to consume the most fuel when cruising for long periods of time well below their maximum accelerating power.

John H. Lewis; William H. Day

Gasket

Deformable material used to make a pressure-tight joint between stationary parts, such as cylinder head and cylinder, that may require occasional separation. Gaskets are known as static seals, as compared with packing or dynamic seals. In packings the parts are frequently in motion, as in piston rods and valve stems. *See* PRESSURE SEAL.

Gaskets are made of sheet materials such as natural or synthetic rubber, cork, vegetable fiber such as paper, asbestos and plastic pastes, or of soft metallic materials such as lead and copper. Rubber in the form of O-rings is used for light pressure.

Paul H. Black

Bibliography. P. H. Black and O. E. Adams, Jr., *Machine Design*, 3d ed., 1968; R. C. Juvinall, *Fundamentals of Machine Component Design*, 3d ed., 2000; J. E. Shigley and C. R. Mischke, *Mechanical Engineering Design*, 2d ed., 1989; Society of Automotive Engineers, *SAE Handbook*, revised annually.

Gasoline

A mixture of hydrocarbons whose boiling point is below 200°C (390°F), obtained in the fractional distillation of petroleum. Gasoline is a liquid at ambient temperature, but it volatilizes readily in air to form a flammable mixture. The hydrocarbon fuel is used to power the internal combustion engine. Gasoline is composed primarily of the alkanes (paraffins) hexane, heptane, and octane, plus smaller amounts of higher-boiling alkanes. *See* ALKANE; INTERNAL COMBUSTION ENGINE.

Historical development. The major force behind the existence of the petroleum industry is the constant consumer demand for products such as gasoline and other liquid fuels. There have been many changes in emphasis on product demand since petroleum first came into use some 5000–6000 years ago. Most of the changes relate to the establishment of the modern refining industry prompted by the discovery of petroleum in the United States in 1859. The main

Component streams for gasoline

Stream	Producing process	Boiling range °C	Boiling range °F
Paraffinic			
Butane	Distillation Conversion	0	32
Isopentane	Distillation Conversion Isomerization	27	81
Alkylate	Alkylation	40–150	105–300
Isomerate	Isomerization	40–70	105–160
Straight-run naphtha	Distillation	30–100	85–212
Hydrocrackate	Hydrocracking	40–200	105–390
Olefinic			
Catalytic naphtha	Catalytic cracking	40–200	105–390
Steam-cracked naphtha	Steam cracking	40–200	105–390
Polymer	Polymerization	60–200	140–390
Aromatic			
Catalytic reformate	Catalytic reforming	40–200	105–390

demand at that time was for the kerosine fraction, used as an illuminant. Thereafter, the development of the internal combustion engine brought about a demand for a new product. *See* KEROSINE.

Production. In the early days of the industry, gasoline requirements were met by distilling crude oil, that is, by maintaining the oil at high temperatures until the different oil components (fractions) reach boiling point and vaporize. The fraction obtained was referred to as straight-run gasoline. However, the heavier (more viscous) crude oils do not contain any gasoline, and the equivalent fraction (naphtha) is now produced by thermal decomposition (cracking), which involves decomposition of the higher-molecular-weight constituents under temperatures greater than 450°C (840°F).

Gasoline is usually produced by catalytic cracking or by reforming processes. In catalytic cracking, the petroleum (or petroleum-derived feedstock) is fed into a reaction vessel containing a catalyst. In reforming, naphtha (refined or unrefined) is heated with hydrogen in the presence of a catalyst. Reforming causes a rearrangement of the structures of the molecular constituents and creates a gasoline product. *See* CRACKING; PETROLEUM PROCESSING AND REFINING; REFORMING PROCESSES.

Composition. The hydrocarbon constituents in the boiling range of gasoline are those that have 4–12 carbon atoms in their molecular structure. Thus, gasoline can vary widely in composition; even gasolines with the same octane number may be quite different. For example, low-boiling distillates with high (above 20%) aromatics contents can be obtained from some crude oils. The variation in aromatics content as well as the variation in the content of normal paraffins, branched paraffins, cyclopentanes, and cyclohexanes is dependent upon the characteristics of the petroleum feedstock, and influence the octane number of the gasoline. *See* AROMATIC HYDROCARBON; DISTILLATION.

Into the first decade of the twentieth century, the gasoline produced was that which was originally present in crude oil or which could be condensed from natural gas (so-called natural gasoline). However, it was soon discovered that if the heavier portions of petroleum (such as the fraction that boils higher than kerosine, for example, gas oil) was heated to more severe temperatures, thermal degradation (or cracking) occurred to produce smaller molecules that were within the range suitable for gasoline. Therefore, gasoline that was not originally in the crude petroleum could be manufactured. *See* NATURAL GAS.

Blends. The differences in composition of gasoline dictate that, in order to produce a uniform product, blending of the products from several component streams (see **table**) is necessary. The properties of each stream may vary considerably, significantly affecting the product gasoline. The blending process is relatively straightforward, but the determination of the amount of each component to include in a blend is much more difficult. The operation is carried out by simultaneously pumping all the components of a gasoline blend into a pipeline that leads to the gasoline storage. The pumps adjust for the correct proportion of each component, while baffles in the pipeline are often used to mix components as they flow to the storage tank.

The low octane number of many paraffinic naphtha fractions can limit their use as gasoline components, although other properties may make them desirable. Other complications arise through the compositional changes induced in the component streams by processing variations. For example, an increase in cracking temperature will produce a higher-octane cracked naphtha (often due to the presence of aromatic constituents), and before this naphtha can be included in a blend, adjustments must be made in the proportions of the other hydrocarbon components.

Volatility. Volatility is an important property of gasoline and is a necessity to ensure engine starting in cold weather. In winter, volatility is raised and the flash point is lowered by adding the more volatile

butanes and pentanes. To prevent vapor lock in warm weather, the amounts of the more volatile constituents are reduced to produce mixtures that will not vaporize in the fuel lines.

Aviation gasoline, used in light aircraft as well as certain types of civil aircraft, has a narrower boiling range [38–170°C (100–340°F)] than automobile gasoline [0–200°C (30–390°F)]. The narrower boiling range ensures better distribution of the vaporized fuel through the more complicated induction systems of aircraft engines. Since aircraft operate at altitudes where the prevailing pressure is less than the pressure at the surface of the Earth, the vapor pressure of aviation gasolines must be limited to reduce preliminary vaporization (boiling) in the tanks, fuel lines, and carburetors. *See* AIRCRAFT ENGINE; AIRCRAFT FUEL.

Additives. Additives are incorporated into commercial gasoline blends, for example to inhibit oxidation and gum formation during storage. Dyes may be added for identification purposes. Alcohol and surfactants are used to reduce carburetor icing and corrosion. Detergent additives remove from the engine and fuel injector some of the deposits produced by gasoline combustion.

Octane number. Before they are ignited by a spark plug, the hydrocarbons in a gasoline blend may ignite spontaneously under the high temperature and pressure conditions inside an engine cylinder. This preignition causes a characteristic engine knock.

The octane number is a measure of the ability of a hydrocarbon fuel to resist preignition. It is obtained by comparing the antiknock performance of the gasoline with that of a mixture of isooctane and heptane: a gasoline blend with an octane number of 90 equals in performance a mixture of 90% octane and 10% heptane. The octane number of a gasoline can be increased by the use of reforming techniques and by alkylation, where gasoline components are recombined to build a larger molecule with a high octane number. *See* OCTANE NUMBER.

Gasohol. Gasohol is a mixture of gasoline and ethyl alcohol that is used as a motor fuel. Although ethyl alcohol is a petroleum product, the necessary commercial quantities for gasohol production can be obtained by fermentation of agriculture waste. Methyl alcohol is also suitable for gasohol mixtures.

Standard automobiles can use gasohol containing up to about 17% ethyl alcohol, although 10% is usually recommended because of environmental concerns. To use higher percentages, an automobile engine must have modifications made to its carburetor, spark plugs, engine timing, and seals. Gasohol burns more slowly, coolly, and cleanly than gasoline, providing greater octane and fewer pollutants. A liability of gasohol is the greater solvent properties of ethyl alcohol, which in concentrated amounts can damage rubber seals and diaphragms and certain finishes. *See* ALCOHOL; ALCOHOL FUEL.

Environmental aspects. For many years tetraethyllead was added to gasoline to increase octane ratings and improve fuel efficiency. It was found, however, that the extremely high levels of airborne lead in urban areas were directly traceable to the combustion of leaded gasoline. In 1985 the U.S. Environmental Protection Agency ruled that gasoline lead content be reduced by over 90%, and by the early 1990s unleaded gas had become the standard throughout the United States. As a result, there was a dramatic drop in airborne lead emissions. To control the levels of other polluting emissions from automobile exhausts, engines must be equipped with various emission-control devices. The most effective is the catalytic converter, which can operate only with unleaded gasoline. *See* AIR POLLUTION; CATALYTIC CONVERTER; PETROLEUM.

James G. Speight

Bibliography. J. G. Speight, *Gas Processing: Environmental Aspects and Methods*, 3d ed., 1999.

Gasteromycetes

An artificial class of fungi in the phylum Basidiomycota in which basidiospores are produced in a mass (gleba) and enclosed within a membrane called the peridium. Such enclosed fruit bodies evolved many times; hence the class is heterogeneous. Most Gasteromycetes have lost the ability to discharge spores ballistically off basidia directly into air. Since internal sporulation prevents normal spore drop into air currents, these fungi have developed other dispersal mechanisms. Puffballs (Lycoperdales) typically form powdery spores in a perforated peridium that acts like a bellows when hit by rain. Earthstars (*Geastrum*) disperse similarly, but the peridia are elevated on arched rays. *Bovista* scatters spores while tumbling in the wind; *Calvatia* cracks open in place. In the gleba, cottony threads called capillitia slow the rate of dispersal. Stalked puffballs (Tulostomatales and Podaxales) characteristically form underground, emerging on stalks after spore maturation; they are adapted to desert life by their protected hypogeous maturation. False puffballs (Sclerodermatales) typically have thicker peridial walls than true puffballs.

The highly evolved Nidulariales have funnel-shaped structures that hold seedlike, splash-dispersed, sporulating disks called peridioles. These fruit bodies resemble miniature bird nests. The related *Sphaerobolus* violently everts an inner membrane, tossing a sticky glebal ball several centimeters. Stinkhorns (Phallales) mature their sticky, foul-smelling glebas in a hypogeous egg from which erupt showy phallic, floral or cagelike receptacles designed to attract insects. Flies or beetles ingest the spores and disperse them. Hymenogastrales (and possibly Melanogastrales and Gauteriales) are dispersed like truffles (Ascomycetes) as spores in feces after excavation and ingestion by animals lured to these false truffles by characteristic odors. *See* ASCOMYCOTA.

Gasteromycetes are either saprophytic or mycorrhizal. Historically, mature dry spores of puffballs were used to clot wounds. Young puffballs are edible, but false puffballs are poisonous. Because of their phallic shape, stinkhorns have been associated

with witchcraft. *See* BASIDIOMYCOTA; EUMYCOTA; FUNGI. Scott A. Redhead

Bibliography. M. B. Ellis and P. J. Ellis, *Fungi Without Gills (Hymenomycetes and Gasteromycetes): An Identification Handbook*, 1990; D. S. Hibbett et al., *Proc. Natl. Acad. Sci.*, 94:12002-12006, 1997; O. K. Miller and H. H. Miller, *Gasteromycetes: Morphological and Developmental Features*, 1988.

Gasterosteiformes

An order of teleost fishes composed of about 257 species, most of which have a small mouth at the end of a long tubular snout and body armor of dermal plates or bony rings. They are further identified by the absence of supramaxillary, orbitosphenoid and basisphenoid bones; and a pelvic girdle that is not connected to the cleithra. Gasterosteiforms are carnivorous predators, feeding chiefly on crustaceans, which they siphon from the water with their tubelike mouthparts. Large species such as trumpetfishes and cornetfishes are primarily piscivorous. Sticklebacks, pipefishes, and seahorses are popular aquarium fishes.

Gasterosteoidei (Thoracostei). The members (three families and about 10 species) of this suborder are identified by a protractile upper jaw, presence of circumorbital, lachrymal, nasal, and parietal bones, absence of the postcleithrum, and kidneys that produce a gluelike substance that is used by males to construct a nest from plant material.

Hypoptychidae (sand eel). A monotypic family that is known from marine waters of Japan and Korea north to the Sea of Okhotsk.

Aulorhynchidae (tubesnouts). Two coastal marine species comprise this family, one on the coasts of Japan and Korea, the other from Alaska to California.

Gasterosteidae (sticklebacks). There are five genera and no less than seven species of sticklebacks, the exact number awaiting studies of the various species complexes. Sticklebacks, so called because of a row of 3 to 16 free (not attached to each other by membranes), stiff, sharp spines that precede the soft dorsal fin. Additional identifying characteristics are a single spine in the anal and pelvic fins (the latter fin fails to develop in several species in certain localities); the mouth is not at the end of a tubular snout and the small jaws are equipped with teeth; a narrow caudal peduncle; and elaborate courtship behavior, which has been the subject of numerous studies. The body of marine species is armed with heavy bony plates, whereas the body of freshwater species is virtually naked. The maximum length is about 18 cm (7 in.). Sticklebacks inhabit marine, brackish, and freshwaters of the north temperate zones of the Northern Hemisphere. The male constructs a nest and provides parental care. *See* STICKLEBACKS.

Syngnathoidei (Solenichthyes). The fishes of this suborder have a small mouth at the end of a long tubular snout and the upper jaw is not protractible. Additionally, the pelvic fins when present are abdominal; the circumorbital bones, other than the lachrymal, are absent; ribs are absent; and the anterior three to six vertebrae are elongate. There are eight families and about 247 species in the suborder.

The following three families comprising the infraorder Syngnatha have uniquely shaped gill filaments. The filaments are reduced to small rosettelike tufts attached to rudimentary arches. This is called the lophobranch pattern (from the Greek *lopho*, "tuft," and *branchos*, "gill", and the fishes are referred to as the Lophomorphs.

Pegasidae (seamoths). This is a family of tropical marine shorefishes (up to 150 m (495 ft) in depth) of the Indo-West Pacific. The body is broad and depressed and encased in body plates; the mouth is beneath a long flattened rostrum; the pectoral fins are large, horizontal and supported by unbranched rays. There are two genera and five species, the largest of which attain 14 cm (possibly 18 cm) in length (5.5, possibly 7.1 in.).

Solenostomidae (ghost pipefishes). This is a family of only one genus and two species, each of which is known from tropical marine waters of the Indo-West Pacific. The family is distinguished by a short, compressed body with two separate dorsal fins, the first five weak spines, the second of small and posteriorly placed opposite the anal fin. The female has a brood pouch formed by the large pelvic fins. Unlike male pipefishes and seahorses, female ghost pipefishes care for the eggs. Maximum size is 16 cm (6.3 in.).

Syngnathidae (pipefishes and seahorses). The body of syngnathids in this family is elongate and encased in bony rings; the gill openings are small; the supracleithrum is absent; a kidney is present only on the right side; the pelvic fins are absent; dorsal, pectoral, and anal fins are usually present, the last is very small, and one or more of the three fins may be absent in adults of some species; the caudal fin is absent is some species and when absent the tail is prehensile. The male cares for the eggs, which are attached by the female to a brood organ (often a pouch) on the underside of his trunk or tail. Most species occur in the temperate and tropical shore waters of the Atlantic, Indian, and Pacific oceans; however, some New World species range into cool waters from southwestern Alaska to Tierra del Fuego. Syngnathids are primarily marine and brackish-water species; however, some species are known only from freshwater.

Two subfamilies are named, Syngnathinae (pipefishes) and Hippocampinae (seahorses); however, some genera are morphological intermediates, perhaps evolutionary links, between pipefishes and seahorses. Examples of intermediate forms are the seadragons of Australia, which are large and have ornate leaflike appendages. The body is not straight as in typical pipefishes, neither is it contorted to the extent seen in seahorses. Some of the intermediate forms are called pipehorses (see **illustration**).

There are about 190 species of pipefishes in

Alligator pipefish, *Syngnathoides biaculeatus*. (*Copyright © J. E. Randall*)

51 genera and only 25 species of seahorses, all in the genus *Hippocampus*. Seahorses differ from pipefishes by having the head bent 90° downward, a prehensile tail, no caudal fin, and they are limited to a marine environment. The pipehorses, in having a caudal fin, are placed with the pipefishes.

Indostomidae. This family, in the infraorder Indostomoida, is represented by a single species, *Indostomus paradoxus*, from southeastern Asia, where it inhabits swamps with soft bottoms and dense vegetation. It has a slender body covered with bony plates; the gill filaments are lobe shaped as opposed to comblike or tufted; the dorsal and anal fins each have six rays and are mirror images; the dorsal fin is preceded by five isolated spines, the opercle is armed with six spines; and the caudal peduncle is very long and slender.

Infraorder Aulostomoida. The following four families, comprising the infraorder Aulostomoida, have normal comblike gill filaments and the postcleithrum is present. Species of the families Aulostomidae and Fistulariidae share the following characteristics: The four anteriormost vertebrae are elongate; well-developed bones (nuchal plates) are medially placed at the nape; and the pelvic fins usually have six soft rays. Species of the families Macroramphosidae and Centriscidae have five or six elongate anterior vertebrae, and pelvic fins with one spine and four soft rays.

Aulostomidae (trumpetfishes). In this family trumpetfishes are tropical marine fishes, most commonly found on reefs of the Atlantic and Indo-Pacific oceans. They have an elongate, scaly, and compressed body (flattened laterally), a fleshy barbel at tip of the lower jaw, rounded caudal fin, and 8 to 12 free spines preceding the dorsal fin, which is a mirror image of the anal fin. These fishes lie at odd angles, most often in a head-down posture, as they wait to attach prey from ambush. A length of 80 cm is the maximum of the three species, all in the genus *Aulostomus*.

Fistulariidae (cornetfishes). Cornetfishes are also tropical marine fishes of reefs and shallow shore waters of the Atlantic and Indo-Pacific oceans. They have an elongate and depressed body (flattened dorsoventrally), the tubular snout is exceptionally long, the skin is either naked or with minute prickles and a linear series of scutes, no barbel at tip of the lower jaw, forked caudal fin with trailing elongate filament, no dorsal spines, and the dorsal fin a mirror image of the anal fin. Cornetfishes are piscivorous, feeding in open water as well as in coral reefs. The largest of the four species in the one genus, *Fistularia*, attains a length of 1.8 m; the remaining three species are usually less than 1 m.

Macroramphosidae (snipefishes). This family, composed of three genera and about 12 species found in the tropical and subtropical Atlantic, Indian, and Pacific oceans, is characterized by a deep and compressed body, usually with bony plates on the back, and the second of four to eight dorsal spines very long. Additionally the long and slender snout ends in a very small mouth, the jaws of which lack a barbel. The snipefishes are bathydemersal (on the bottom in deep water) marine occupants of the continental shelf and slopes to 1000 m (3300 ft), where they feed on benthic invertebrates.

Centriscidae (shrimpfishes). The shrimpfish family, composed of two genera and four species, are the most bizarre fishes among the gasterosteiforms, if not all fishes of the world. They are extremely compressed with the body almost entirely encased by thin bony plates that are expansions of the vertebral column. The elongate body, so encased and with an exceptionally sharp ventral keel, suggests a razor. The spikelike first dorsal spine is at the very end of the body, and it is inline with the flat dorsal contour of the body, displacing two short spines and closely spaced soft dorsal and caudal fins to a ventral position at the posterior end of the fish. Shrimpfishes, limited to marine and brackish waters of the Indo-Pacific, swim in a head-down vertical position, in schools or small groups among long-spine sea urchins and branching coral, to depths of 20 m to 100 m, depending on the species. They feed primarily on planktonic crustaceans and reach a maximum length of 15 cm (6 in.).

Herbert Boschung

Bibliography. M. A. Bell and S. A. Foster (eds.), *The Evolutionary Biology of the Threespine Stickleback*, Oxford University Press, 1994; C. E. Dawson, Family Syngnathidae, in *Fishes of the Western North Atlantic*, Mem. Sears Found. Mar. Res. Mem. 1(pt. 8), pp. 1–172, 1982; P. C. Heemstra, Macrorhamphosidae, pp. 459–461, in M. M. Smith and P. C. Heemstra (eds.), *Smith's Sea Fishes*, Springer-Verlag, Berlin, 1986; E. S. Herald, From pipefish to seahorse—a study of phylogenetic relationships, *Proc. Calif. Acad. Sci.*, Ser. 4, vol. 29, pp. 465–473, 1959; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, 1994; T. W. Pietsch, Evolutionary relationships of the sea moths (Teleostei: Pegasidae) with a classification of gasterosteiform families, *Copeia*, 1978(3):517–529, 1978; R. J. Wootton, *The Biology of the Sticklebacks*, Academic Press, 1976; R. J. Wootton, *A Functional Biology of Sticklebacks*, University of California Press, 1984.

Gastornithiformes

An order of extinct flightless birds known from Paleocene and Eocene deposits in North America, Europe, and Asia. They are believed to have dispersed across the North Atlantic via a land bridge when those continents were connected in the Eocene. These were giant, 2-m-tall (7-ft) birds with large heads, huge laterally compressed bills, broad cervical vertebrae, wide pelves, massive legs with relatively short tarsi and heavy toes suggestive of a slow-moving gait, and reduced wings that were evidently too short to permit flight.

Mode of life. The mode of life of gastornithiforms is a matter of controversy. They have often been depicted as predators that evolved to fill the bipedal carnivore niche vacated by theropod dinosaurs at the end of the Cretaceous Period. A biomechanical analysis of their powerful jaw apparatus, suggesting considerable biting strength, has been used to support the idea that they were carnivorous and capable of feeding on both live prey and carrion. According to an alternative view, their jaw and hindlimb anatomy indicates that they were primarily herbivorous and incapable of sustained rapid running. Various skull and postcranial characters suggest that gastornithiforms are related to the waterfowl (Anseriformes) and, more distantly, to the true fowl (Galliformes). *See* ANSERIFORMES; GALLIFORMES.

Fossil record and classification. Gastornithiformes were first described from the Eocene of France in 1855, with the genus *Gastornis*. In 1876, a comparable giant bird from the Eocene of New Mexico was described as *Diatryma*. The original material from both Europe and North America was fragmentary, and *Gastornis* was first reconstructed in a highly inadequate fashion. When much more complete material from North America was discovered in the early twentieth century, it seemed that the North American form, *Diatryma*, was significantly different from the European *Gastornis*. Later, new specimens from Europe were described as *Diatryma*. A reexamination of the original French specimens and descriptions of new finds from Europe show that *Gastornis* and *Diatryma* were in fact very similar and should be placed in a single genus, *Gastornis* (the earlier name having priority).

Gastornis is known from the Paleocene of Germany, France, and Belgium; the Lower Eocene of France, England, the United States, and Canada; and the Middle Eocene of Germany. Two species (*Gastornis giganteus*, *Diatryma regens*) are recognized from the New World, and there may be three from the Old World (*G. parisiensis*, *G. russelli*, and *G. sarasini*). A second genus, *Zhongyuanus*, was described from the Lower Eocene of China on the basis of scanty material and remains poorly known. *See* EOCENE; PALEOCENE.

Habitat. Gastornithiforms are most abundant in alluvial sediments of the Bighorn Basin of Wyoming—where the most complete known skeleton, the holotype of *Diatryma steini* Matthew and Granger [=*Gastornis giganteus* (Cope); see **illustration**], was discovered; in fluvial deposits in eastern France; and in lake and swamp deposits of the Geiseltal of Germany. Associated sediments, floras, and faunas indicate that the Gastornithiformes inhabited well-vegetated, temperate-to-subtropical coastal lowlands and inland floodplains, much like the marshy habitats now occupied by their vegetarian relatives, the Anhimidae or screamers (Anseriformes), in South America.

Gastornis giganteus **(reconstruction).** ***(American Museum of Natural History)***

Resemblance to modern birds. Modern avifaunas contain no exact counterparts to the gastornithiforms. According to the herbivorous interpretation, the closest analogs probably are the takahe, *Porphyrio mantelli* (Rallidae), and kakapo, *Strigops habroptilus* (Psittacidae) of New Zealand, which are flightless and herbivorous and are, respectively, the largest rail and parrot. Large body size and flightlessness in gastornithiforms, *Porphyrio*, and *Strigops* seem related to adoption of a low-energy diet of leaves, which in turn requires a large alimentary canal, forfeiture of the energetic expense of flight, and the evolution of a jaw apparatus specialized for feeding on vegetation. Supporters of the carnivorous hypothesis stress resemblances with some modern scavenging birds, such as the marabou stork (*Leptoptilus crumeniferus*) or the African carrion crow (*Corvus crassirostris*); however, these birds are not flightless. According to this interpretation, possible extinct analogs may be the South American phorusrhacids, which were giant flightless birds for which a predatory lifestyle is generally accepted. *See* AVES.

Allison V. Andors; Eric Buffetaut

Bibliography. A. V. Andors, *Diatryma* among the dinosaurs, *Nat. Hist.*, 104(6):68–71, 1995; A. V. Andors, Paleobiology and relationships of the giant groundbird *Diatryma* (Aves: Gastornithiformes),

Acta XX Congressus Internationalis Ornithologici, vol. I, pp. 563–571, 1991; A. V. Andors, Reappraisal of the Eocene groundbird *Diatryma* (Aves: Anserimorphae), *Science Series*, Los Angeles, no. 36, pp. 109–125, 1992; L. D. Martin, The status of the Late Paleocene birds *Gastornis* and *Remiornis*, *Science Series*, Los Angeles, no. 36, pp. 97–108, 1992.

Gastrointestinal tract disorders

Malfunctions of the organs of digestion from the esophagus to the rectum. The gastrointestinal tract usually functions unnoticed, intruding into consciousness only in the form of specific sensations of hunger, intestinal motion, or the need to defecate. However, it is subject to a wide array of pathologic states.

Esophagus. The esophagus is the site of two principal disorders: disturbance of the muscular contractions, and mechanical interference with the passage of food caused by a narrowed lumen. Carcinoma of the esophagus produces mechanical obstructions to swallowing; the long-term prognosis is poor. Excessive acid reflux or regurgitation from the stomach may result in inflammation of the lower esophagus with heartburn and indigestion. Changes in diet and eating habits may bring relief.

Achalasia. The prime example of muscular or neuromuscular disorder is achalasia, a condition of unknown etiology characterized by degeneration of the nervous mechanisms that coordinate contractions of the esophagus and relaxation of the lower esophageal sphincter. Swallowing solids and liquids is difficult, and nutritional consequences usually result. The condition is normally painless, but some persons have active contractions of the body of the esophagus, a condition known as vigorous achalasia. Treatment is directed at reducing the pressure in the lower esophageal sphincter to permit food to pass by force of gravity. Weakening of the sphincter is accomplished surgically by cutting the outer circular muscle layer of the esophagus or by forcible pneumatic dilation with a balloon which ruptures the circular muscular fibers. Both techniques are reasonably effective. Medications such as calcium channel blockers improve motility and reduce pressure in the lower esophageal sphincter.

Connective tissue disease. Other disorders that affect the esophagus include connective tissue diseases such as scleroderma or dermatomyositis. Scleroderma is characterized by weakening of the smooth musculature, which is present in the lower two-thirds of the esophagus, and the lower esophageal sphincter. Muscle contraction is weakened, and the sphincter fails to prevent the reflux of gastric acid; reflux esophagitis frequently ensues. Dermatomyositis affects the striated musculature present in the upper third of the esophagus. There are no specific treatments for connective tissue diseases affecting the esophagus. However, in the case of scleroderma, acid reduction and antireflux measures are often helpful; surgery can reduce the hazard of acid reflux. *See* CONNECTIVE TISSUE DISEASE.

Mechanical disorders. Mechanical problems include stricture or scarring of the lower esophagus, which usually develops from acid reflux. A stricture forms slowly, usually preceded by symptoms of esophagitis, that is, heartburn and discomfort in the lower chest, particularly at night or in recumbency. Stricture tends to be progressive, with swallowing becoming increasingly difficult. It is diagnosed with a barium meal x-ray or by direct visualization with upper gastrointestinal endoscopy. Strictures can be dilated with cylindrical instruments called bougies, which are introduced orally in progressively larger diameters. Results are good with acid reduction and antireflux measures. Medications such as bethanechol and metoclopramide improve the tone of the lower sphincter and aid esophageal emptying. In severe cases, surgery may be required to reduce acid reflux.

Swallowing is also impeded mechanically by a lower esophageal, or Schatzki, ring, which is a mucosal constriction at the junction of the esophagus and stomach in the presence of a hiatus hernia. Reflux esophagitis may favor the development of such a ring, which tends to stabilize; the resulting problem is not progressive but intermittent and dependent upon the size of the food bolus. It may produce the steak house syndrome, in which a person who is eating, and possibly under the influence of alcohol, swallows a lump of food that sticks at the Schatzki ring. Treatment is based primarily on patient education, with occasional dilation or surgical correction of the tight ring. *See* HERNIA.

Cancer. Carcinoma of the esophagus causes mechanical obstructions to swallowing. The cancerous growth, whether arising in the squamous epithelium of the esophagus or in the columnar epithelium of the upper stomach, may progressively reduce the size of the lumen and lead to increasingly severe dysphagia. Ultimately, even swallowing fluids becomes impossible. Carcinoma of the esophagus or the upper region of the stomach is usually diagnosed by means of fiber-optic endoscopy and biopsy. The long-term prognosis for all forms of esophageal cancer is poor, but encouraging results have been achieved in Japan with high-dose radiation followed by extensive surgical resection. Palliative treatments are available: radiation therapy can temporarily shrink squamous cell cancer of the esophagus.

Reflux esophagitis. Reflux esophagitis is an inflammation of the lower esophagus that results from excessive acid reflux or regurgitation from the stomach. The common symptoms of heartburn and indigestion are produced by reflux of acid. The major cause seems to be a weakness in the lower sphincter, which ordinarily serves as a physiological barrier to stomach acid. Pressure within the sphincter may be low in the presence of a hiatus hernia; however, reflux esophagitis can occur even in the absence of hiatus hernia or other disease. Scleroderma may impair the lower esophageal sphincter and lead

to reflux esophagitis. Other factors that may contribute include inadequate clearing of the acid from the esophagus, and the composition of gastric contents that reach the esophagus. Symptoms vary from occasional heartburn to severe distress caused by regurgitation of gastric acid into the throat, known as water brash, or even expectoration of acidic fluid. Treatment is directed at neutralizing gastric acid or reducing its secretion in the stomach, or at improving the tone of the lower esophageal sphincter. The individual can reduce acid reflux by elevating the upper body during sleep, eating smaller meals, retiring on an empty stomach, losing weight, and avoiding caffeine, tobacco, and foods found to cause distress. Those measures are usually successful; surgery is rarely required unless serious complications occur, such as bleeding from a highly inflamed esophageal lining or stricture and scarring.

Stomach. The stomach is the site of secretion of the highly concentrated hydrochloric acid. The lining of the stomach and duodenal both need to be protected against the effects of the acid.

Peptic ulcer disease. Duodenal ulcer is one of the most common maladies, but for unknown reasons its incidence seems to be declining. It is associated with high concentrations of hydrochloric acid caused by excessive gastrin secretion. The role of genetic or environmental influences is not known. In persons with a gastric ulcer, acid concentrations are usually normal or low. Gastric ulcer occurs in persons of middle age or older, whereas duodenal ulcer tends to be a disease of young adults. The cause of gastric ulcer is not known but may involve increased reflux of bile and other constituents of the small intestine. Bile salts are capable of breaking the gastric mucosal barrier, which normally keeps hydrochloric acid in the lumen. When the barrier is broken, hydrogen ions diffuse into the cells and cause inflammation. The barrier can also be broken by aspirin and other nonsteroidal anti-inflammatory drugs. The combination of aspirin and alcohol seems to be particularly deleterious: alcohol causes vascular congestion, and aspirin breaks the mucosal barrier. Together they increase the risk of bleeding in the upper gastrointestinal tract.

Peptic ulcer disease is typically diagnosed with an upper gastrointestinal barium meal or fiber-optic endoscopy. When an infectious agent, particularly *Campylobacter pylorii*, is suspected, it may be identified by biopsy of the lining of the stomach antrum. However, it is a difficult organism to culture, and methods of eradication are still being studied. Duodenal ulcer tends to be recurrent and even chronic; once the diagnosis is confirmed, further testing is usually not warranted. The presence of a gastric ulcer should be firmly established, since it can occasionally mask cancer. Highly effective medications are available for the treatment of duodenal ulcer. Antacids are useful in large doses. Histamine receptor blockers inhibit the secretion of hydrochloric acid from the parietal cell. Some drugs act at the site of the ulcer to provide a protective barrier against pepsin. Pharmaceuticals that inhibit the proton pump, which takes part in the final stage of parietal cell functioning, would dramatically reduce gastric acid secretion.

Surgery, once frequently performed for peptic ulcer disease, is now uncommon, and removal of part of the stomach is indeed rare. Surgery, when required, usually involves sectioning (cutting) of the vagal nerves or their branches to reduce gastric acid secretion. Such an operation is known as a vagotomy or selective vagotomy and is usually performed when noninvasive medical treatment has failed.

Gastritis. Inflammation of the stomach lining, known as gastritis, ranges in severity from acute to chronic and can ultimately lead to atrophy of the tissues and loss of function. Chronic forms of gastritis may occur in association with other diseases such as diabetes mellitus or hypothyroidism. When gastric atrophy is severe, gastric acid and pepsin are no longer secreted. In the most severe forms of gastritis, intrinsic factor is no longer secreted, and pernicious anemia with vitamin B_{12} deficiency develops. Immunologic factors seem to play an important role in pernicious anemia; antibodies to both parietal cells and intrinsic factor are detectable in the blood. *See* ANEMIA; DIABETES; THYROID GLAND DISORDERS.

Tumors. The incidence of gastric cancer has been declining more rapidly than any other cancer. At one time it was among the most common causes of death but is now less common than either colon or pancreatic cancer. Some of the decline has been attributed to improved refrigeration methods and decreased exposure to nitrites, which are thought to be carcinogenic. Nitrites are found in partially decomposed meats, smoked and cured meats, and cigarette smoke. Nonetheless, cancer of the stomach remains a serious problem. In Japan, where the incidence is high, screening programs have detected cancers in very early stages, when cure rates are high. The typical symptoms of gastric cancer include early satiety after eating small amounts of food, abdominal pain, anemia from blood loss, and difficulty in swallowing. Benign stomach tumors are usually asymptomatic but may cause bleeding.

Small intestine. The small intestine is normally resistant to disease. When nutrition is adequate, structure and function remain intact into old age with only slight changes if no specific diseases occur.

Malabsorption syndrome. A prototypical disease of malabsorption is gluten enteropathy, or nontropical sprue. Probably due in part to genetic abnormalities, gluten enteropathy is characterized by inflammation and loss of the normal architecture of the small intestine following ingestion of some proteins. In the extreme form, villus atrophy, the fingerlike projections from the intestinal wall are absent. The causative substance is contained in gluten, the protein of cereal grains; a gluten fraction, gliadin, contains the toxin that is responsible. A gluten-free diet excludes wheat, barley, rye, and oats, although the toxicity of oat protein has been questioned. Several genetic markers are evident in patients with gluten enteropathy, and antibodies to wheat protein fractions can be found in their blood. Diagnosis of the condition can be based

on the response to gluten restriction and on a small-intestine biopsy performed with an endoscope or a swallowed capsule. A gluten-free diet, once it has been adopted, is usually maintained for life. In children the condition is called celiac disease; the diet must be observed throughout life despite lessening of the symptoms with advancing age.

Other conditions may also interfere with the absorptive function of the small intestine. Tropical sprue, found in tropical and subtropical climates, resembles gluten enteropathy in its manifestations. It may, however, be an infectious diarrhea; it can be treated with antibiotics and high doses of folic acid. A gluten-free diet does not cure the condition.

Infections within the small intestine, particularly those of the parasite *Giardia lamblia*, may also cause malabsorption. Bacterial infection by *Escherichia coli* causes most cases of traveler's diarrhea. Among the many other diseases that may lead to malabsorption is regional enteritis, or Crohn's disease. *See* INFLAMMATORY BOWEL DISEASE.

Tumors. Of all the body's organs, the small intestine is one of the most resistant to neoplastic disease. Benign tumors are rare, as are cancer and lymphoma. (In some familial conditions, however, benign polyps may develop.) An understanding of that resistance might have value in cancer research.

Colon. The colon, or large intestine, is subject to malfunction of widely varying degrees, ranging from mild irritation to life-threatening diseases.

Altered bowel function: constipation and diarrhea. Disturbance of bowel function is among the most common complaints. Constipation probably predominates, particularly with advancing age, but loose stool and diarrhea are also frequent and troublesome. A familiar condition associated with altered function is irritable bowel syndrome, a nonspecific label for an intestinal disturbance without known anatomic cause.

Constipation may be defined as bowel movements less frequent than once every 3 days or as hard stool that is difficult to pass. Until recently, the typical Western diet contained low amounts of fiber compared with that common in underdeveloped nations and probably contributed to the incidence of constipation. Fiber in the diet, particularly from the indigestible component of grains known as bran, frequently gives relief. In some cases, a severe problem seems to worsen with age and the attendant loss of bowel motility. Constipation may be a symptom of other diseases such as hypothyroidism and diabetes, but after the age of 50 a change in bowel habit suggests the possibility of cancer of the colon or rectum.

Diarrhea is less common than constipation as a long-term complaint. It may result from intolerance to certain foods, notably milk. If milk sugar (lactose) is incompletely digested, it draws water osmotically into the colon and is acted upon there by bacteria to produce gas and lactic acid, which contribute to diarrhea. Among adults who are Asians, blacks, or Ashkenazi Jews, lactose intolerance is found in over 60%. The incidence in populations of northern European descent is only about 15%. Abstention from milk and milk products is curative.

Other food substances may also contribute to chronic diarrhea. Fructose, a sugar found in fruits and used as a sweetener, is inadequately absorbed in some individuals and can lead to diarrhea. The sugar sorbitol is used as a low-calorie substitute for sucrose or glucose because it is poorly absorbed, but diarrhea may result if it is consumed in large quantities.

Diarrhea can, of course, occur as a symptom of serious disease such as regional enteritis or ulcerative colitis. It may also result from infection, particularly by the parasites *Giardia lamblia* or *Entamoeba histolytica*. Bacterial infections caused by *Salmonella*, *Shigella*, and *Campylobacter* may cause short-term diarrhea. Diabetes mellitus, hyperthyroidism, and other systemic diseases can also be associated with diarrhea. *See* BACILLARY DYSENTERY; DIARRHEA; INFANT DIARRHEA.

Irritable bowel syndrome. Among the most common problems in adults is the irritable bowel syndrome. The term is a general one that incorporates such labels as spastic colitis, functional bowel disease, and psychophysiologic gastrointestinal reaction. Population studies have found the incidence of the condition to be as high as 15% in adults; women are more likely to be affected. Anatomical disease is not present, and good health is retained with no long-term consequences. The symptoms include recurrent abdominal pain, constipation or diarrhea that may alternate, and sensation of gaseousness and bloating. The cause has not been identified, but medication and dietary changes, especially the addition of fiber, may offer relief.

Diverticulosis. In diverticulosis, small pouches appear along the wall of the colon. These represent herniations of the mucosa through the muscular layers. In most cases the intestinal musculature is thickened. The disease is more prevalent with increasing age; barium enema radiology or colonoscopy reveals an incidence of 30% or higher among persons in their seventies.

Diverticulosis is associated with high pressures within the colon and particularly its narrowest segment, the sigmoid colon. The pressure pushes the mucosa through the weakest points in the wall, leading to formation of the diverticuli. The symptoms are usually mild, with little or no discomfort or with episodes of pain that are usually experienced in the lower left quadrant of the abdomen. Perforation of the pouches may cause an inflammatory process known as diverticulitis or peridiverticulitis. Erosion of a small artery supplying blood to the colon wall may lead to bleeding.

A high-fiber diet is the standard treatment for diverticulosis, but some authorities also advocate restricting the ingestion of seeds and other foods that might become impacted within a diverticulum. Surgery may be required in cases of peridiverticulitis or hemorrhage.

Cancer. The second most common internal cancer is that of the colon and rectum; more than 130,000 cases are diagnosed each year in the United States.

Cancers frequently arise in benign polyps of the colon, which are stalklike or leafed structures (villous adenomas). A high intake of red meat and low fiber and low vitamin A consumption may be contributing factors; a genetic predisposition has been suggested. Early detection offers the best hope of recovery.

Annual testing for the presence of blood in the stool after age 50 is widely recommended. Detection is possible by means of a barium enema radiological examination or colonoscopy. Early cancers can be removed with a good prognosis; in later stages the chance for recovery is reduced. Small cancers can be removed by colonoscopy; otherwise surgery is required.

Hemorrhoids. Hemorrhoids are veins near the anus that have become distended or occluded. The cause is not fully understood but may be related to straining at stool. Hemorrhoids may be treated with a high-fiber diet, stool softeners, and applications of heat. Severe cases may require surgery.

Appendix. Appendicitis, that is, inflammation of the vermiform appendix, was once common among young people but is now relatively rare; the reason for the sharp drop in frequency is unclear. Appendicitis seems to result when a hardened piece of fecal matter known as a fecalith becomes trapped within the appendix and causes infection that can lead to gangrene. If the condition persists, the appendix may perforate, resulting in widespread infection in the abdomen. Symptoms usually begin with vague discomfort in the midabdomen, occasionally associated with nausea and vomiting; pain and tenderness tend to localize in the right lower quadrant, and then a low-grade fever is likely to develop as well. An inflamed appendix should be removed before it ruptures. *See* APPENDICITIS; CANCER (MEDICINE); DIGESTIVE SYSTEM; TUMOR. Leonard A. Katz

Bibliography. D. A. Coffman et al., *Gastrointestinal Disorders*, 1987; M. Feldman et al., *Sleisenger and Fordtran's Gastrointestinal Disease: Pathophysiology, Diagnosis, and Management*, 2 vols., 6th ed., 1998; J. Norton, *Gastrointestinal Disorders: A Pathophysiologic Approach*, 1986.

Gastrolith

Any of the pebbles swallowed by animals and retained for a time in the gizzard or stomach, where they serve to grind up the food and in so doing become rounded and highly polished. Birds generally use such pebbles, as do some living reptiles, notably the crocodile and certain lizards. Some of the Mesozoic reptiles also used gastroliths. Articulated skeletons of the Plesiosaurs are generally associated with such highly polished cobbles, even when embedded in formations such as the Niobara chalk, which is otherwise free of gravel. In some instances as many as a half bushel of such stones have been found within the rib cage. Among the dinosaurs the great sauropods likewise used gizzard stones.

Similar highly polished stones found in the arid Southwest of the United States are so abundant and widespread, however, as to indicate that they are due to natural weathering. No satisfactory criteria are known for distinguishing them from gastroliths. Carl O. Dunbar

Gastropoda

The largest and most varied class in the phylum Mollusca, possibly numbering over 74,000 species and commonly known as snails. *See* MOLLUSCA.

Functional Morphology

The shell is in one piece that, in the majority of forms, grows along a turbinate (equiangular) spiral (**Fig. 1**) but is modified into an open cone in various limpets or is secondarily lost in various slugs.

Effects of torsion. All gastropods, at some time in their phylogeny and at some stage in their development, have undergone torsion. The process does not occur in any other mollusks. It implies that the visceral mass and the mantle shell covering it have become twisted through 180° in relation to the head and foot. As a result of torsion, all internal organs are twisted into a loop. Similarly in gastropods, the mantle cavity (the semi-internal space enclosed by the pallium or mantle) containing the characteristic molluscan gills (ctenidia) has become anterior and placed immediately above and behind the head. The most primitive gastropods, like stem stocks in the other molluscan groups, retain a pair of aspidobranch (bipectinate or featherlike) gills, each with alternating ctenidial leaflets on either side of a ctenidial axis in which run afferent and efferent blood vessels (**Fig. 2**). Lateral cilia on the faces of the leaflets create a respiratory water current (toward the midline and anteriorly) in the direction opposite to the flow of blood through the gills, to create the physiological efficiency of a countercurrent exchange system. *See* COUNTERCURRENT EXCHANGE (BIOLOGY).

As also occurs in other mollusks, the ctenidia form a gill curtain which functionally divides the mantle cavity into an inhalant part (lateral and below) with characteristic osphradia as water-testing sense organs, and an exhalant part into which the anus and renogenital ducts discharge. In zygobranch (two-gilled) gastropods, various shell slits or complex openings have been developed (Fig. 2*c* and *d*) to deal with the problems of sanitation created by the mantle cavity being brought anteriorly above the head. In various ways, these accommodate the exhalant stream of deoxygenated water and the feces and genital and kidney products which accompany it, so that this exhalant discharge is not directly over the head. In the majority of gastropods this problem is resolved by the reduction of the gill pair to a single ctenidium (**Fig. 3**) or to a pectinibranch ctenidium (a one-sided, comb-shaped "half-gill").

Shell coiling and asymmetry. In a typical snail the shell is turbinate or helicoid, growing as a steadily expanding conic tube along the "screw spiral" of a

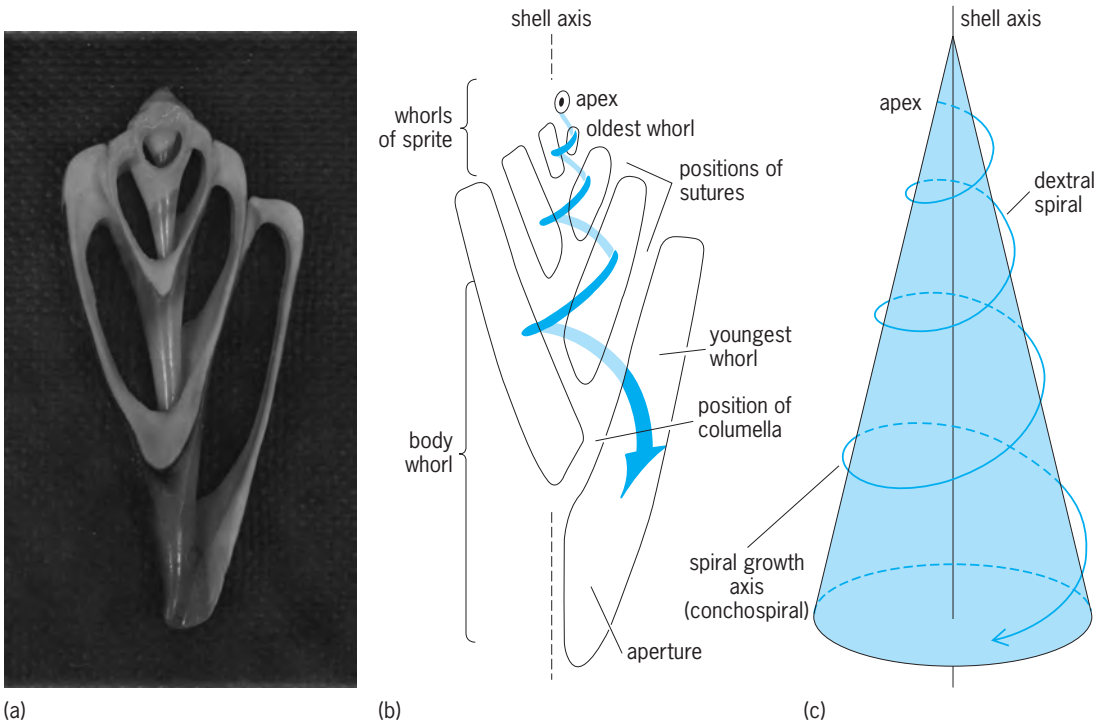

Fig. 1. Typical turbinate shell of gastropods. (*a*) Longitudinal section ground through the shell of a specimen of *Conus purius* to reveal the central columella and spiral of whorls expanding to the aperture. (*b*) Diagram of the shell to show the relationship of the whorls sectioned. (*c*) Stylized representation of the conchospiral line, as viewed with the shell apex tilted toward the observer. This logarithmic or equiangular line is the center line of a steadily expanding conic tube secreted by the mantle edge (at the shell aperture) during growth. (*After W. D. Russell-Hunter, A Life of Invertebrates, Macmillan, 1979*)

logarithmic or equiangular line on the surface of a hypothetical cone (Fig. 1). Only a minority of gastropods have planospiral shells or the open limpet form. The majority with turbinate shells cannot have similar left and right halves, and almost all snails are anatomically asymmetric. Even in those gastropods in which the shell is lost and there has been a secondary return to a bilateral symmetry of external features, there remain marked asymmetries of internal anatomy. The majority of turbinate-shelled gastropods have dextral coiling, with a consequent reduction of the pallial organs of the right side, so that the left ctenidium persists. Internal organs are similarly reduced, so that "higher" gastropods have only the left auricle of the heart (the monotocardiac condition), and the left kidney and a single gonad opening to the exterior by the originally right-hand renogenital coelomoduct.

Mantle cavity evolution. From the most primitive forms which retain most clearly the basic effects of torsion (Fig. 2), the further evolution of the gastropods has involved increasing asymmetry of

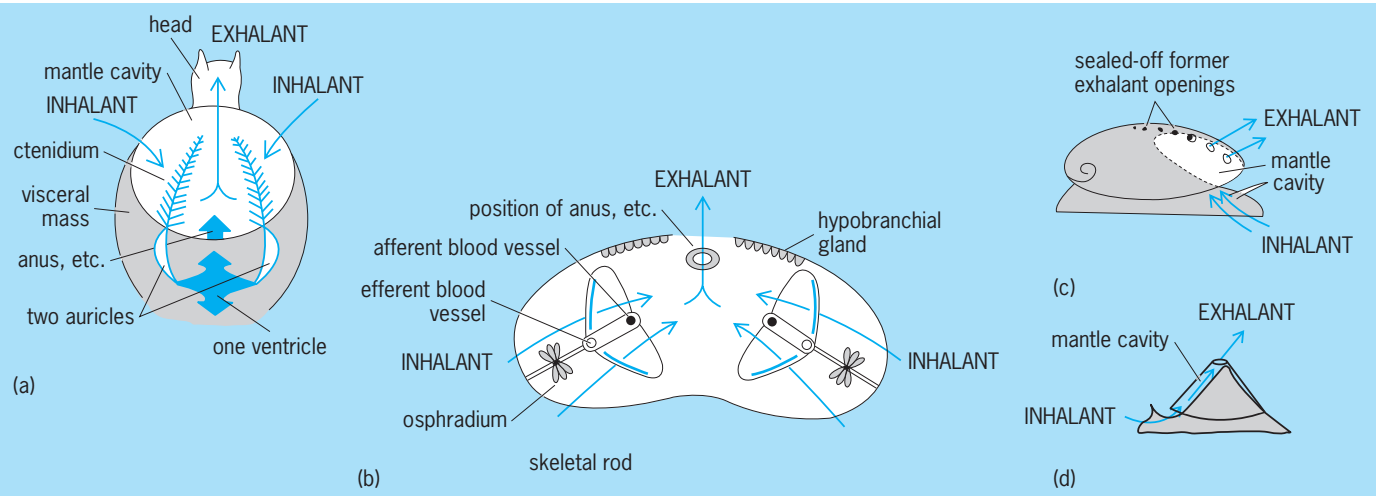

Fig. 2. Mantle cavity and associated organs in a primitive zygobranch gastropod, with a symmetrical pattern of two aspidobranch gills (*a*) in dorsal view and (*b*) in cross section. This arrangement of mantle cavity is found in (*c*) abalones and (*d*) keyhole limpets. (*After W. D. Russell-Hunter, A Life of Invertebrates, Macmillan, 1979*)

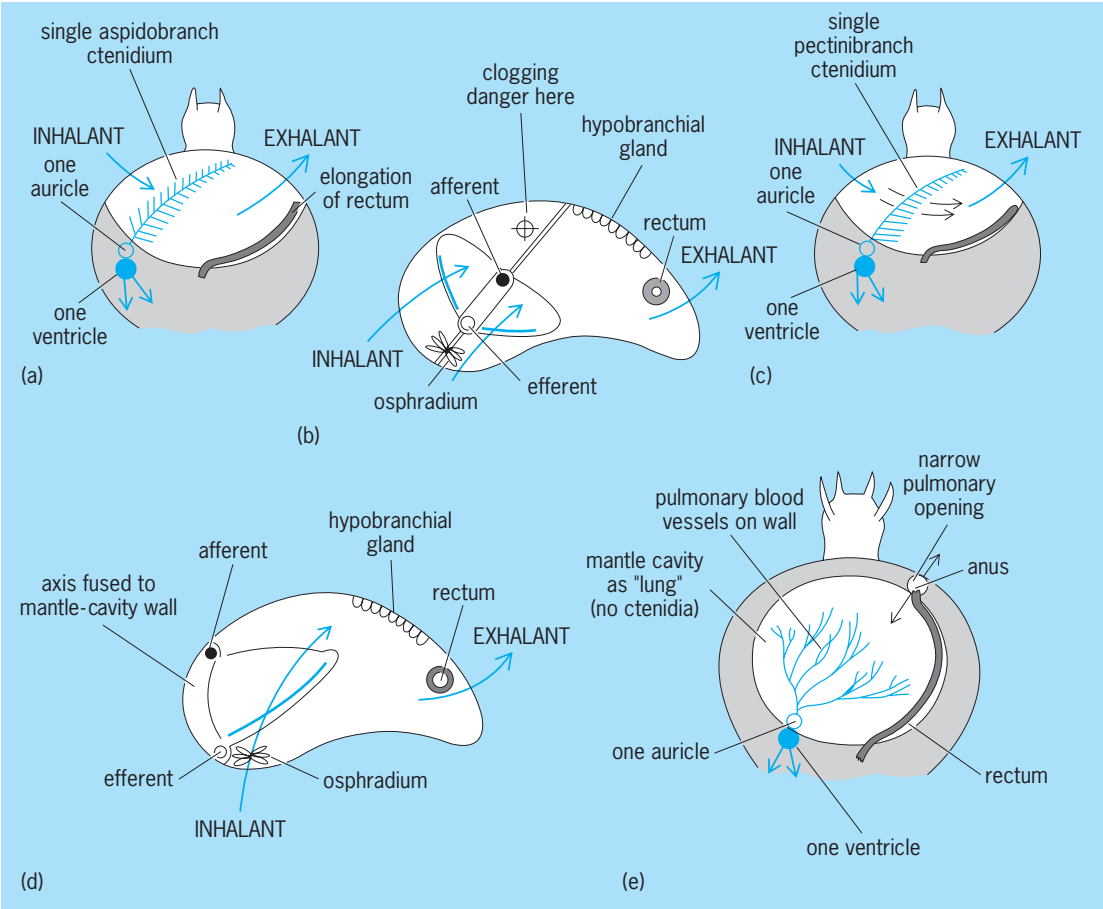

Fig. 3. Further evolution of the mantle cavity in gastropods. The mantle cavity and pallial complex is shown (*a*) in dorsal view and (*b*) in cross section for an asymmetric snail with a single aspidobranch gill, and (*c*, *d*) for an asymmetric snail with a "half-gill" or pectinibranch ctenidium. (*e*) The mantle cavity of a pulmonate land snail in dorsal view shows that there is now no gill and the mantle has become vascularized as a lung wall. (*After W. D. Russell-Hunter, A Life of Invertebrates, Macmillan, 1979*)

the mantle cavity organs. Relatively few living gastropods are zygobranch with two ctenidia (and the diotocardiac condition, Fig. 2*a*). These include keyhole limpets such as *Fissurella* and abalones (*Haliotis*).

In most gastropods the right ctenidium is completely lost, and the pallial water flow has become inhalant from the left side of the snail, with the anus and renogenital openings moved over to the right (now exhalant) side (Fig. 3*a* and *b*). A minority of these single-gilled snails retain an aspidobranch gill, with some danger of particulate material clogging the dorsal part of the mantle cavity. Such snails include the limpet genus, *Tectura* (=*Acmaea*), common on both Atlantic and Pacific coasts, the big worldwide group of top shells (*Trochus* and its allies), and the tropical littoral genus *Nerita*. All these forms are ecologically limited to relatively clean water over hard substrata and are unable to invade areas of the sea bottom or seashore covered with mud or silt.

By far the most successful marine gastropods (without such ecological limitation) are those in which the pallial structures are further reduced (Fig. 3*c* and *d*), with a pectinibranch ctenidium whose axis is fused to the mantle wall, resulting in greater hydrodynamic efficiency. Note that this "half-gill" pattern still presents the same functional relationships of ciliary water currents and blood vessels in a countercurrent system. Most familiar snails of the seashore, including such genera as *Busycon* (whelks), *Nassarius* (mud snails), *Littorina* (periwinkles), *Euspira* (=*Polinices*) [moon snails], and very many others, have pectinibranch gills and this highly asymmetric arrangement of the mantle cavity.

A final reduction of pallial structures is found in the pulmonate snails and slugs (Fig. 3*e*), in which the mantle cavity is an air-breathing lung and there are no ctenidia. The other more specialized subclass, the Opisthobranchia, shows detorsion after loss of the shell and a variety of secondary (neomorphic) gills, especially in the extremely beautiful sea slugs or nudibranchs. *See* NUDIBRANCHIA; PULMONATA.

Diversity and Classification

More than half of all molluscan species are gastropods, and they encompass a range from the marine limpets, which can be numbered among the most primitive of all living mollusks, to the highly evolved terrestrial air-breathing slugs and snails. Pulmonates and certain mesogastropod families are the only successful molluscan colonizers of land and

freshwaters. The anterior mantle cavity (resulting from torsion), as opposed to the posterior mantle cavity of all other major molluscan stocks, remains diagnostic of the class despite a diversity of functional morphology unequaled by any comparable group in the entire animal kingdom.

Presently classification of gastropods is rather unsettled, even though the phylogeny of this group is being examined vigorously. Consequently, an older arrangement of gastropod taxa, in practice prior to the onset of phylogenetic systematics, is presented here. While most of these taxa are now considered to be artificial, they are useful to consider in that each one demonstrates parallel morphologies and habitats.

The older systematic arrangement of the class Gastropoda involves three somewhat unequal subclasses: Prosobranchia, Opisthobranchia, and Pulmonata.

Prosobranchia. The largest and most diverse subclass is Prosobranchia, which is made up largely of marine snails, all retaining internal evidence of torsion. The prosobranchs have traditionally been divided into three orders: Archaeogastropoda, Mesogastropoda, and Neogastropoda. The primitive Archaeogastropoda are characterized by the presence of one or two aspidobranch (bipectinate) ctenidia, auricles, metanephridia (kidneys), and osphradia (Figs. 2, 3*a* and *b*). Members of the polyphyletic Mesogastropoda, which comprises almost 100 families, have only one pectinobranch (monopectinate) ctenidium (left), auricle, and nephridium (Fig. 3*c* and *d*). Members of the Neogastropoda, which appears to be monophyletic, also exhibit these characters, but their shell has a canal or notch that holds a tubular extension of the mantle (siphon).

In some current classifications, most families of the Mesogastropoda and Neogastropoda (and some Archaeogastropoda) constitute a taxon known as the Caenogastropoda, whereas most of the archaeogastropods are placed within the Patellogastropoda (true limpets) and Vetigastropoda (abalones, top and turban snails, keyhole limpets). *See* MESOGASTROPODA; NEOGASTROPODA; PROSOBRANCHIA.

Opisthobranchia and Pulmonata. The other two subclasses (Opisthobranchia and Pulmonata) are each considerably more uniform than the subclass Prosobranchia and, in both, the effects of torsion are reduced or obscured by secondary processes of development and growth. Even so, in newer classifications they are often combined into a single group (Heterobranchia) along with a few groups (sundials, pyramidellids) formerly classified as prosobranchs.

The marine subclass Opisthobranchia consists largely of sea slugs, in which the shell and mantle cavity are reduced or lost and there is a bilaterally symmetrical adult with a variety of secondary gill conditions. *See* OPISTHOBRANCHIA.

The final subclass, Pulmonata, consists of gastropods with the mantle cavity modified into an air-breathing lung and with no ctenidia (Fig. 3*e*). There are a few littoral marine forms, but the order Basommatophora is mainly made up of the freshwater lung-snails, and the order Stylommatophora consists of the successful land snails such as *Helix* plus a few shell-less families of land slugs. *See* PULMONATA; BASOMMATOPHORA; STYLOMMATOPHORA.

W. D. Russell-Hunter; Robert E. Knowlton

Bibliography. V. Fretter and A. Graham, *British Prosobranch Molluscs*, 2d ed., 1994; F. W. Harrison and A. J. Kohn (eds.), *Microscopic Anatomy of Invertebrates*, vol. 5, 1994, vol. 6B, 1997; R. N. Hughes, *A Functional Biology of Marine Gastropods*, 1986; R. M. Linsley, Shell form and the evolution of gastropods, *Amer. Sci.*, 66:432–441, 1978; W. D. Russell-Hunter, *A Life of Invertebrates*, 1979; J. Taylor (ed.), *Origin and Evolutionary Radiation of the Mollusca*, 1996; E. R. Trueman and M. R. Clarke, *The Mollusca*, vol. 10, 1985.

Gastrotricha

A phylum of minute metazoan animals (sometimes placed in the aschelminth group, in the Cycloneuralia, or in other combinations of phyla), mostly less than a millimeter in length. It presently numbers 600 described species worldwide. Some 380 species have been reported from the marine habitat, with new ones being described every year.

Gastrotrichs comprise two orders, the Macrodasyida and the Chaetonotida. The term Gastrotricha refers to the ventral locomotor cilia by which the animals glide gracefully over the substratum or through its interstices; unlike many other ciliated animals, they cannot move in reverse. Gastrotrichs have a complete digestive tract, with a triradiate muscular sucking pharynx, a simple intestine with a wall only a single cell thick, and a ventral anus. They appear to be selective feeders on bacteria, very small protozoa, and yeasts. The musculature, of obliquely striated fibers, form diverse patterns (including spirals or helices) that reach throughout the body, and allow a wide range of nonlocomotor contortive motion. Most have biflagellate protonephridia, accounting in part for the broad salinity tolerances of many species. The sensory organs are quite sophisticated for an animal of this size. Dual gland adhesive tubes (one with glue, the other with releaser) can occur in groups on the front, side, and rear of the animals; in most chaetonotids these are reduced to a single pair in the rear or entirely lost. The nonchitinous cuticle may have one to many ultrastructural layers; its surface may be naked and translucent or elaborated into thickenings, scales, spines, or hooks (**Figs. 1** and **2**).

Reproduction and development. Gastrotrichs are bisexual, either simultaneously or sequentially male and female. Some macrodasyids pass a spermatophore from donor to recipient animal, with sperm penetrating the body cuticle and moving internally to the eggs; others cross-copulate, some after first moving sperm from a midbody duct to a rear transfer organ. Reproduction in chaetonotids is unique, egg-to-adult ratio being among the greatest in the animal kingdom; a period of asexual reproduction by parthenogenesis is followed by

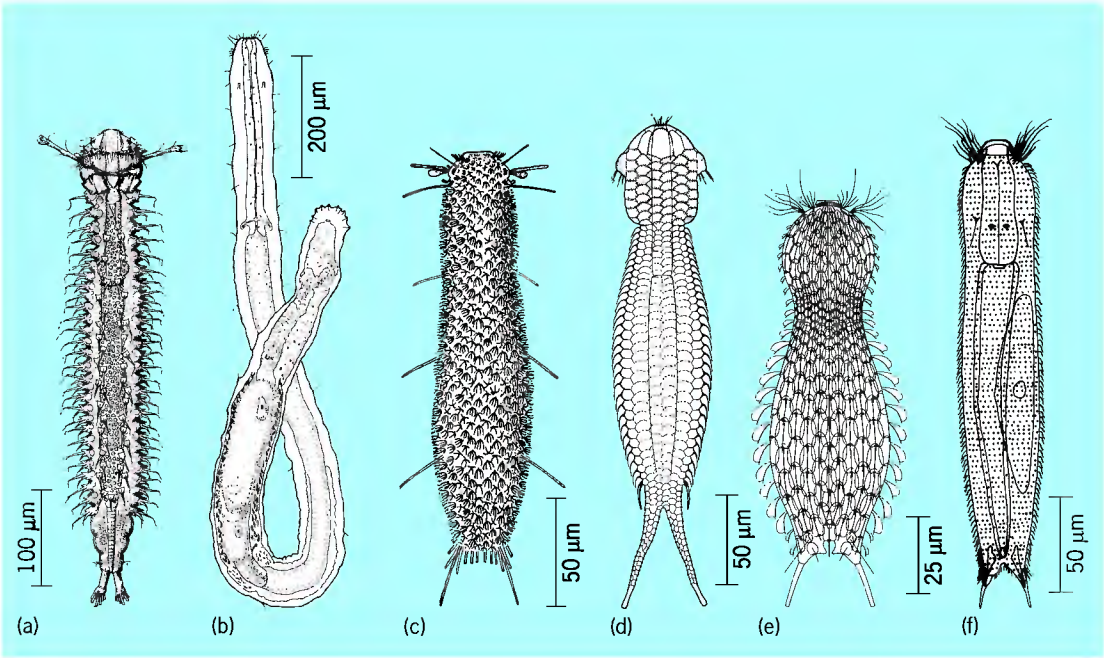

Fig. 1. Diagrams of (*a–c*) macrodasyid and (*d–f*) chaetonotid gastrotrichs. (*a*) *Xenodasys*. (*b*) *Dolichodasys*. (*c*) *Thaumastoderma*. (*d*) *Draculiciteria*. (*e*) *Halichaetonotus*. (*f*) *Aspidiophorus*. (*Part a after C. Schoepfer-Sterrer, Chordodasys riedli gen. nov., spec. nov., a macrodasyoid gastrotrich with a chordoid organ, Cah. Biol. Mar., 10:391–404, 1969; b after G. D. Gagne, Dolichodasys elongatus n.g., n. sp., a new macrodasyid gastrotrich from New England, Trans. Amer. Microsc. Soc., 96:19–27, 1977; c after A. Remane, Klassen und Ordnungen des Teirreichs, vol. 4(2), Akad. Verlagsges., Leipzig, 1936; d after P. Luporini et al., Contribution à la connaissance des Gastrotriches des côtes de Toscane, Cah. Biol. Mar., 12:433–455, 1971; e after H. Mock, Chaetonotoidea (Gastrotricha) der Nordseeinsel Sylt, Mikrofauna Meersebodens, 78:1–107, 1979; f after W. D. Hummon, Gastrotricha from Beaufort, North Carolina, U.S.A., Cah. Biol. Mar., 15:431–446, 1974).*

cross-fertilization between reciprocal hermaphrodites. Eggs then rupture through the body cuticle or exit through a ventral pore and are deposited on the surface or in crevices of substratum particles. Reproductive biology and RNA phylogenies are areas of active research.

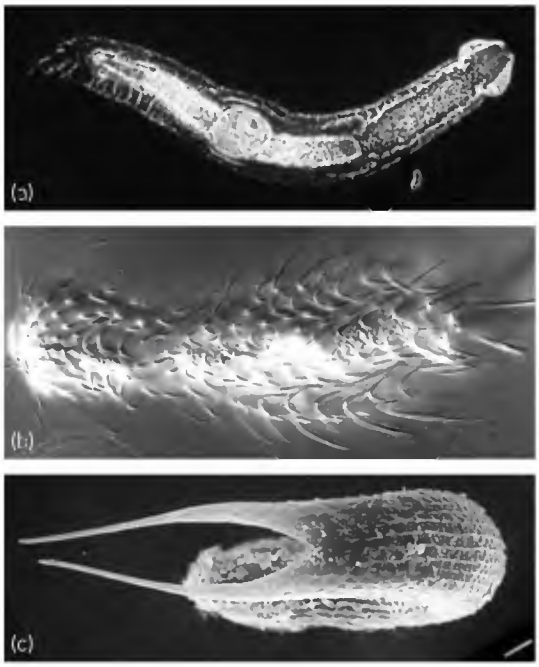

Fig. 2. Photographic images of (a) *Cephalodasys littoralis*, 670 m, via phase contrast microscopy; (*b*) *Chaetonotus siciliensis*, 300 m, via differential interference contrast microscopy; and (*c*) *Musellifer sublitoralis*, 330 m, via scanning electron microscopy, (scalebar 10 m). (*Parts a and c after W. D. Hummon, c after M. Balsamo*)

Most gastrotrich eggs are elliptical and measure 40 × 60 micrometers; those of macrodasyids are smooth, of chaetonotids mostly rough or spiny. Usually, only one egg is laid in each clutch. Cleavage is modified radial with blastomeres being subequal in size. Embryogenesis at 20°C (68°F) takes 10–12 days for macrodasyids or 1–3 days for chaetonotids. Development is direct: juveniles being about 100 μm in length and growing to adulthood in 30 days, in macrodasyids, or 2–5 days, in chaetonotids. The life span of macrodasyids can be up to 6 months, probably averaging less than a month, for chaetonotids up to a month, averaging less than a week; populations can increase rapidly under suitable conditions. Chaetonotids have fewer than a thousand cells, among the least among metazoans; they are cell-constant and show eutely (all adults have the same number of cells). Chaetonotids may lay eggs that resist temperature extremes and drying, some being able to hatch even after a decade. *See* CELL CONSTANCY.

Biogeography and evolution. Gastrotrichs appear to be regionally cosmopolitan, with 20–30% having broad distributions within continents, and 10–15% between continents; endemism probably does not exceed 20%. Biogeographical research has broadened from Europe, India, and North America to include South America, Australia, and the Far East.

In having species with monociliated epithelial cells, the phylum Gastrotricha is among the most primitive in the aschelminth group of phyla.

Gastrotrichs and nematodes probably share a common ancestor, which in turn was descended from a stock that included gnathostomulids and turbellarianoid animals. *See* CHAETONOTIDA; GNATHOSTOMULIDA; MACRODASYIDA; NEMATA; ROTIFERA; TURBELLARIA.

William D. Hummon

Bibliography. K. G. Adiyodi and R. G. Adiyodi (eds.), *Reproductive Biology of Invertebrates*, 6 vols., 1983–1993; R. P. Higgins and H. Thiel (eds.), *Introduction to the Study of Meiofauna*, 1988; C. Nielsen, *Animal Evolution: Interrelationships of the Living Phyla*, 2d ed., 2001; S. D. Rundle, A. L. Robertson, and J. M. Schmid-Araya (eds.), *Freshwater Meiofauna: Biology and Ecology*, 2002; E. E. Ruppert, Gastrotricha, in F. W. Harrison (ed.), *Microscopic Anatomy of Invertebrates*, vol. 4, 1991; W. H. Thorp and A. P. Covich (eds.), *Ecology and Systematics of North American Freshwater Invertebrates*, 1991.

Gastrulation

A set of cellular and epithelial movements that transform the blastula stage of embryonic development into a more complex arrangement of three germ layers. The outer germ layer, which will be exposed to the external environment, is the ectoderm. It gives rise to the epidermis and nervous system of the developing organism. The inner layer of cells, the endoderm, forms the inner lining of the digestive tract and its derivatives. An intermediate layer, the mesoderm, develops into a diverse array of structures, including the skeleton, muscles, kidneys, and circulatory system. The ectoderm and endoderm are usually epithelial, consisting of closely packed sheets of cells connected by tight junctions. The mesoderm is sometimes epithelial, but at other times it is mesenchymal, forming a loose arrangement of cells surrounded by extracellular material. The three germ layers are formed in nearly all animal embryos, while the movements that generate them vary. *See* CLEAVAGE (DEVELOPMENTAL BIOLOGY); GERM LAYERS.

Beginning of morphogenesis. Gastrulation marks the onset of dramatic cellular and epithelial movements that are collectively called morphogenesis because they generate the embryo's body plan. During morphogenesis, individual cells move relative to each other, while entire epithelia spread, fold, and change shape. The amazing complexity and order of these movements are best appreciated in time-lapse movies, which show how cells divide, wedge themselves between other cells, and change between epithelial and mesenchymal states. These cellular behaviors are based on the orchestrated synthesis of proteins that control the shape, adhesion, and movements of cells. *See* EMBRYONIC INDUCTION; MORPHOGENESIS.

Gastrulation in sea urchins. Sea urchins illustrate gastrulation with almost schematic clarity. The sea urchin blastula is a sphere of epithelial cells surrounding a blastocoel filled with a gelatinous fluid (**Fig. 1***a*). A group of cells with long cilia marks the blastula's animal pole; the opposite pole is the vegetal pole. Gastrulation begins when the vegetal cells become tall and columnar, forming the vegetal plate (Fig. 1*b*). Next, some cells of the vegetal plate begin a morphogenetic movement known as ingression: They break the epithelial junctions that have sealed them to neighboring cells and move into the blastocoel as single cells (Fig. 1*c*). Called primary mesenchyme cells, they later form the skeleton of spicules supporting the arms of the sea urchin larva.

A major event in gastrulation is the formation of the primitive gut, or archenteron. It begins

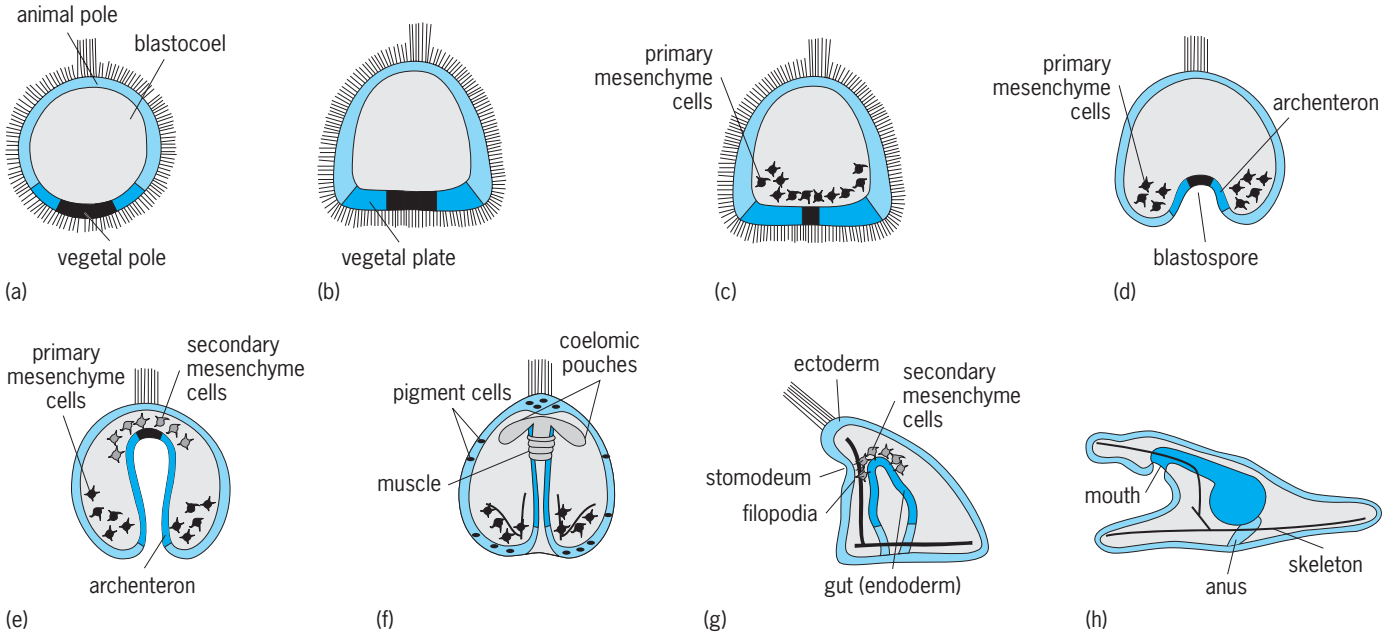

Fig. 1. Sea urchin gastrulation. (*Reprinted with permission from K. Kalthoff, Analysis of Biological Development, 2d ed., McGraw-Hill, 2001*)

with a common type of epithelial movement called invagination, in which an epithelium buckles inward to create a depression (Fig. 1*d*). The opening of the archenteron, called the blastopore, marks the site where endodermal and mesodermal cells move inside the embryo. The short archenteron rudiment then undergoes another typical epithelial movement known as convergent extension, that is, an elongation in one dimension with simultaneous shrinkage in one or two other dimensions (Fig. 1*e*). This is achieved by two cellular activities. First, individual archenteron cells elongate in a coordinate way. Second, the same cells slide past one another, so that the number of cells counted along the archenteron's length increases while the number of cells in its circumference decreases.

While archenteron elongation is still underway, a set of secondary mesenchyme cells become motile at the archenteron tip (Fig. 1*e*). Some of them form pouches of mesodermal epithelium, which displace the blastocoel by generating a secondary body cavity called the coelom (Fig. 1*f*). Other secondary mesenchyme cells give rise to pigment cells, muscles, and blood cell equivalents. While still attached to the archenteron tip, secondary mesenchyme cells also initiate the final phase of archenteron elongation by extending thin cytoplasmic processes called filopodia. They probe the inside of the blastocoel wall, attach at suitable sites and then contract, orienting the archenteron tip toward the ventral side of the embryo (Fig. 1*g*). Here, the archenteron tip is met by the stomodeum, a depression in the oral surface of the ectoderm. Fusion of the stomodeum with the archenteron tip generates the mouth of the larva, while the blastopore becomes the anus (Fig. 1*h*).

Gastrulation in humans. In humans, the inner cell mass of the blastocyst forms a disc consisting of two layers called epiblast and hypoblast. In a morphogenetic movement called delamination, the epiblast splits into the embryonic epiblast and the amnion (**Fig. 2**). The amniotic cavity between them will expand and contain the amniotic fluid that surrounds the growing fetus until birth. The following description focuses on the embryonic epiblast, called simply epiblast from here on, which will give rise to the embryo proper.

Gastrulation begins with formation of the primitive streak, a pair of ridges with a groove between them, extending along the epiblast's midline (**Fig. 3**). The thickened anterior end of the streak is known as Hensen's node (like the corresponding structure in the chicken embryo, from which much of our knowledge on human gastrulation has been gleaned). The primitive streak becomes the major gastrulation site, similar to the blastopore in the sea urchin embryo, where future endoderm and mesoderm cells move inside. In humans, epiblast cells converge to the primitive streak. Here they undergo a morphogenetic movement known as involution, which is defined as an inward movement of an expanding epithelium around an edge. As soon as the epiblast cells have involuted around the ridges of the primitive streak and disappeared into the groove, they

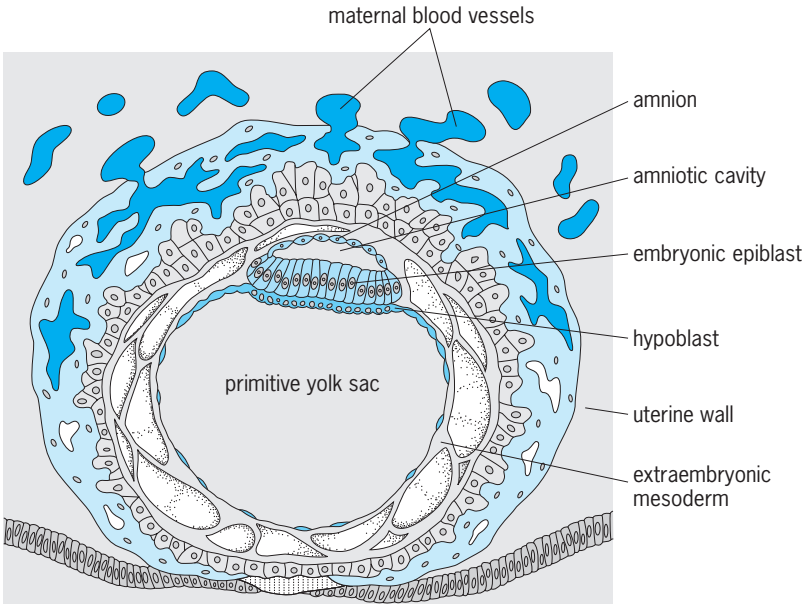

Fig. 2. Human embryo at day 12. (*Reprinted with permission from K. Kalthoff, Analysis of Biological Development, 2d ed., McGraw-Hill, 2001*)

become mesenchymal and ingress into the space between the epiblast and hypoblast (**Fig. 4**).

The fate of the ingressing cells depends on the time of ingression and the direction of their further movement. The cells that ingress first form the embryonic endoderm, which will line the inside of the digestive tract. Subsequently, Hensen's node recedes to the posterior, thus shortening the primitive streak. Later, ingressing cells that stay along the midline, in the wake of the receding Hensen's node, become the axial mesoderm, which will form the precursor of the embryonic backbone. Cells ingressing to more lateral positions form paraxial and lateral mesoderm, which make trunk musculature, kidneys, blood vessels, heart, and blood. Epiblast cells that do not involute remain on the outside and form the ectoderm,

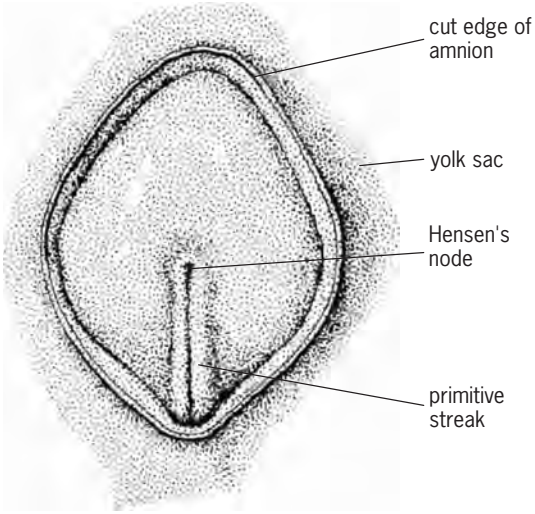

Fig. 3. Human epiblast in dorsal view at day 16. The amnion has been cut away. (*Modified with permission from B. M. Carlson, Human Embryology and Developmental Biology, Mosby, 1994*)

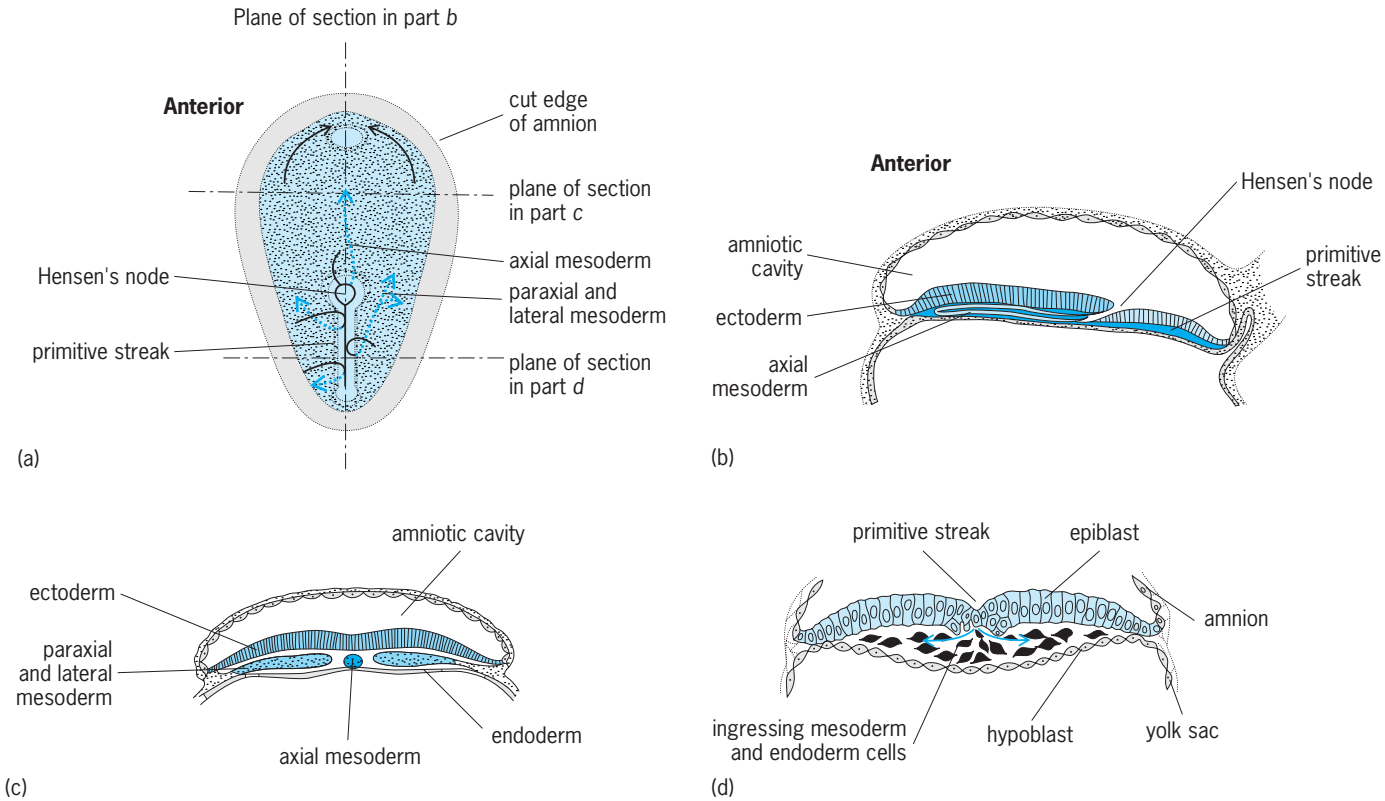

Fig. 4. Gastrulation in the human embryo. (*Reprinted with permission from K. Kalthoff, Analysis of Biological Development, 2d ed., McGraw-Hill, 2001*)

which gives rise to the epidermis, brain, and spinal cord.

In broad terms, the sea urchin gastrula is a three-layered sphere, with ectoderm outside, endoderm inside, and mesoderm between. In the human, the same germ layers form a triple disk that will be transformed into a cylindrical embryo by subsequent morphogenetic movements. Klaus Kalthoff

Bibliography. B. M. Carlson, *Human Embryology and Developmental Biology*, Mosby, 1994; K. Kalthoff, *Analysis of Biological Development*, 2d ed., McGraw-Hill, Boston, 2001.

Gate circuit

An electronic circuit that consists of elements, which may be transistors, diodes, or resistors, combined in such a manner that they perform a logic operation. Gate circuits are the most basic building blocks of a digital system. These circuits have one or more inputs and one output which is a boolean function of the inputs. The input and output signals can have only two discrete values, low (for example, 0 V) and high (for example, 3.3 V). These values are usually represented as 0 and 1, or "false" and "true," respectively.

Whereas the early gate circuits consisted of diodes, resistors, and transistors, the majority of gate circuits nowadays are built exclusively with transistors. The dominant technology for fabricating gate circuits is the metal-oxide-semiconductor (MOS) silicon method, followed by the silicon bipolar and gallium arsenide (GaAs) techniques. The manufacturing process has become so sophisticated that transistors smaller than 1 square micrometer can be fabricated, allowing the placement of millions of gate circuits on a silicon chip the size of a fingernail. The main advantage of MOS technology is that it gives rise to very low power circuits that can still operate at relatively high clock speed. It is these characteristics

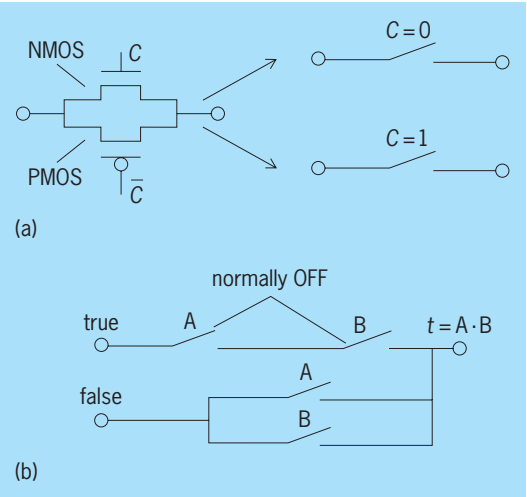

Fig. 1. Switch circuit architecture. (*a*) CMOS transmission gate, with an NMOS and a PMOS transistor in parallel, clocked at opposite clock signals (*C* and $\bar{C}$). (*b*) Switch circuit for two-input logic AND function, with two normally-OFF switches in series and two normally-ON switches in parallel. Each switch represents a transmission gate.

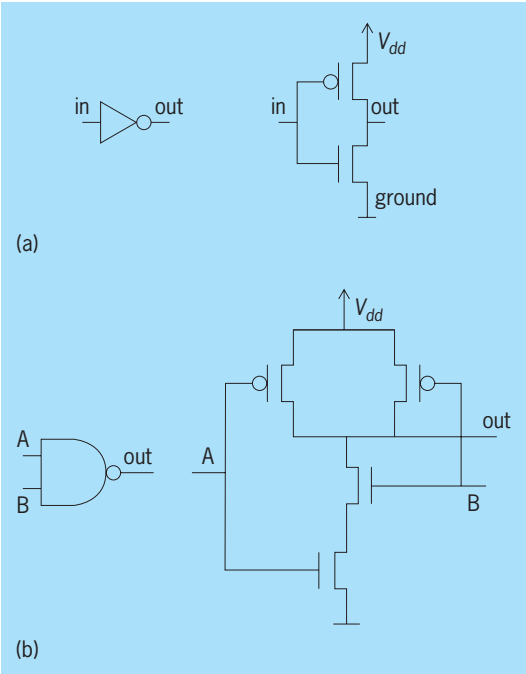

Fig. 2. Logic symbols and corresponding gate circuits. (*a*) CMOS inverter circuit built with a PMOS and an NMOS transistor. (*b*) Two-input NAND circuit.

that have allowed the fabrication of very complex digital systems such as microprocessors and memories. *See* INTEGRATED CIRCUITS; MICROPROCESSOR; SEMICONDUCTOR MEMORIES.

The transistors in a gate circuit are used as ON-OFF switches. By combining these transistors in a certain way, it is possible to realize logical, arithmetical, and memory functions. There are two type of MOS transistors: NMOS and PMOS field-effect transistors (FETs), corresponding to a normally-OFF or normally-ON switch, respectively. Circuits in which both types of transistors are used are called CMOS (complementary MOS) circuits. CMOS circuits now constitute the majority of gate and logic circuits.

Transmission gate or switch circuits. In switch circuits an MOS transistor is used to pass or block the flow of information in a similar fashion to a mechanical switch. By placing these switches in a network, it is possible to realize different logic functions. A transistor used in this fashion is often called a pass-transistor. In order to improve the switching characteristics, an NMOS switch and a PMOS switch are placed in parallel, each clocked at opposite clock signals. Such a combination is called a CMOS transmission gate (**Fig. 1***a*). Several transmission gates can be combined to form a logic AND circuit (Fig. 1*b*). Each switch represents a transmission gate composed of two transistors in parallel (Fig. 1*a*). The two switches in series are normally-OFF, while the two switches in parallel are normally-ON switches. The network of switches (Fig. 1*b*) is such that when both inputs A and B are true the two series switches are ON and will pass the information. As soon as one of the inputs is false, the corresponding series switch will be OFF and the parallel switch will be ON, making the output false, thus implementing an AND function. Transmission gates are also widely used to implement more complex circuits such as delay elements, flip-flops, and other transmission gate logic structures. *See* SWITCHING CIRCUIT.

Logic gates. An alternative way to realize logic functions is to make use of logic gates. The simplest gate circuit, the inverter, takes an input signal and presents the inverted signal at the output (**Fig. 2***a*). This circuit works as follows: When 0 V (corresponding to a logic 0 or "false" signal) is applied to the input, the NMOS transistor (bottom one) will be OFF, while the PMOS transistor will be ON, connecting the output node to the voltage V_{dd} (corresponding to a logic 1). In contrast, a logic 1 (that is, V_{dd}) applied to the input will switch the NMOS transistor ON and the PMOS OFF, thus connecting the output node to the ground (logic 0).

One of the two transistors in this circuit will always be OFF when the output is stable. This implies that there is no current flowing in the transistors so that the quiescent power dissipation goes to zero. Only when the transistors are switching from the ON to the OFF state will both be ON for a short time, giving rise to dynamic power dissipation. This is one of the main advantages of CMOS circuits and is the reason why the majority of integrated circuits are fabricated in CMOS.

An example of a NAND circuit is shown in Fig. 2*b*. The operation of this circuit can be easily understood. Since the NMOS transistors are normally-OFF and the PMOS transistors are normally-ON, the output will be connected to V_{dd} (logic 1) as soon as one of the input signals is low because at least one of the PMOS transistors will be ON. Only when both inputs are high will the two NMOS transistors be ON, thus making the output low. By using additional transistors, it is possible to implement more complex functions such as arithmetic or memory elements.

The above examples are all CMOS gate circuits. Similar logic functions can be realized with bipolar or

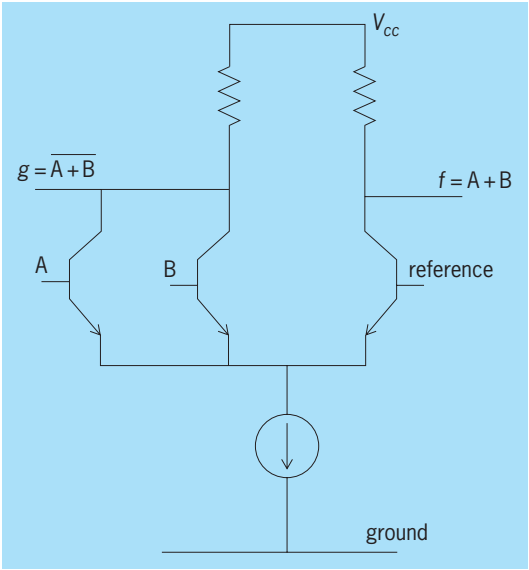

Fig. 3. Emitter-coupled logic (ECL) two-input OR/NOR circuit, consisting of a differential pair.

gallium arsenide transistors. The transistor-transistor logic (TTL) circuits were very popular before the CMOS technology came to maturity. An alternative bipolar technology is the emitter-coupled logic (ECL) circuits. These are used mainly to implement very high speed logic circuits such as those found in the core of processors of supercomputers. A two-input OR/NOR emitter-coupled logic circuit (**Fig. 3**) consists of a differential pair. As soon as one of the inputs A or B is high, the output g will go low and f will go high, implementing the OR and NOR functions, respectively. *See* LOGIC CIRCUITS; TRANSISTOR.

Jan Van der Spiegel

Bibliography. S.-M. Kang and Y. Leblebici, *CMOS Digital Integrated Circuits*, 2d ed., McGraw-Hill, New York, 1998; J. Rabaey, *Digital Integrated Circuits*, Prentice Hall, Upper Saddle River, NJ, 1996.

Gauge theory

The theoretical foundation of the four fundamental forces of nature, the electromagnetic, weak, strong, and gravitational interactions. Gauge symmetry lies at the heart of gauge theory. A gauge symmetry differs from an ordinary symmetry in two important respects:

1. Gauge symmetry is a local symmetry rather than a global symmetry. For a local symmetry, the element of the symmetry group (G) that acts on the fields of a theory at a space-time point ($\mathbf{x}$, t) depends on the position $\mathbf{x}$ and time t, whereas for a global symmetry a fixed group element acts on fields at different space-time points.

2. A gauge transformation leaves a physical state invariant. Gauge symmetry reflects a redundancy in the variables used to describe a physical state. By contrast, a global symmetry acting on a physical state in general produces a new, distinct physical state. *See* FUNDAMENTAL INTERACTIONS; SPACE-TIME.

Electromagnetism. The simplest example of a gauge theory is electromagnetism. In classical electrodynamics, gauge invariance reflects the arbitrariness that exists in choosing the potentials $\mathbf{A}(\mathbf{x}, t)$ and $\phi(\mathbf{x}, t)$ to represent the electric and magnetic fields, $\mathbf{E}$ and $\mathbf{B}$, according to Eqs. (1), where c is the

$$\begin{aligned} \mathbf{E}(\mathbf{x}, t) &= -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \\ \mathbf{B}(\mathbf{x}, t) &= \nabla \times \mathbf{A} \end{aligned} \tag{1}$$

speed of light. If $\Lambda(\mathbf{x}, t)$ is an arbitrary scalar field, Eqs. (2) define a gauge transformation. The poten-

$$\begin{aligned} \phi \to \phi' &= \phi - \frac{1}{c}\frac{\partial \Lambda}{\partial t} \\ \mathbf{A} \to \mathbf{A}' &= \mathbf{A} + \nabla\Lambda \end{aligned} \tag{2}$$

tials $\mathbf{A}'$ and ϕ' may equally well be used to represent the electromagnetic fields $\mathbf{E}$ and $\mathbf{B}$. By using units such that $c = 1$, a more compact, manifestly Lorentz covariant notation may be adopted, in which there is a single 4-vector potential $A_\mu = (\phi, -\mathbf{A})$, so that the field strengths are components of a single antisymmetric, second-rank 4-tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, whose components are given by Eq. (3).

$$\begin{aligned} F_{\mu\nu} &= \begin{pmatrix} F_{00} & F_{01} & F_{02} & F_{03} \\ F_{10} & F_{11} & F_{12} & F_{13} \\ F_{20} & F_{21} & F_{22} & F_{23} \\ F_{30} & F_{31} & F_{32} & F_{33} \end{pmatrix} \\ &= \begin{pmatrix} 0 & +E_x & +E_y & +E_z \\ -E_x & 0 & -B_z & +B_y \\ -E_y & +B_z & 0 & -B_x \\ -E_z & -B_y & +B_x & 0 \end{pmatrix} \end{aligned} \tag{3}$$

See MAXWELL'S EQUATIONS; POTENTIALS; RELATIVISTIC ELECTRODYNAMICS.

In nonrelativistic quantum mechanics, gauge invariance is realized as follows. The Schrödinger equation for a particle with an electromagnetic charge q and mass m is Eq. (4), where $\hat{H}$ is the hamilto-

$$\begin{aligned} \hat{H}\Psi &= \frac{1}{2m}\left(\hat{\mathbf{p}} - \frac{q}{c}\mathbf{A}\right)^2 \Psi + q\phi\Psi \\ &= \frac{1}{2m}\left(-i\hbar\nabla - \frac{q}{c}\mathbf{A}\right)^2 \Psi + q\phi\Psi \\ &= i\hbar\frac{\partial}{\partial t}\Psi \end{aligned} \tag{4}$$

nian operator, $\hat{p}$ is the momentum operator, Ψ is the wave function of the particle, and $\hbar$ is Planck's constant divided by 2π. Equation (4) is invariant under the gauge transformation given by Eqs. (5).

$$\begin{aligned} \Psi \to \Psi' &= \exp\left[\frac{iq\Lambda}{\hbar c}\right] \cdot \Psi \\ \phi \to \phi' &= \phi - \frac{1}{c}\frac{\partial \Lambda}{\partial t} \\ \mathbf{A} \to \mathbf{A}' &= \mathbf{A} + \nabla\Lambda \end{aligned} \tag{5}$$

Electromagnetism is a $U(1)$ gauge symmetry [where U(1) is the one-dimensional unitary group, which is represented by the complex numbers $e^{i\varphi}$ with $0 \le \varphi < 2\pi$] because a gauge transformation rotates the phase of the wave function in a space-time–dependent manner and adjusts the potentials $\mathbf{A}$ and ϕ accordingly. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; SCHRÖDINGER'S WAVE EQUATION; UNITARY SYMMETRY.

In classical electrodynamics, gauge invariance plays a more formal role, because classically it is possible to formulate electromagnetism entirely in terms of the gauge-invariant, physically observable field strengths $\mathbf{E}$ and $\mathbf{B}$ without recourse to the potentials $\mathbf{A}$ and ϕ. Although at the classical level charged particles interact only with the field strengths $\mathbf{E}$ and $\mathbf{B}$ through the Lorentz-force equation, this is not true quantum mechanically. Any quantum-mechanical description of the electromagnetic interaction of charged particles must involve the potentials $\mathbf{A}$ and ϕ, as illustrated by the Aharonov-Bohm effect. *See* AHARONOV-BOHM EFFECT.

Geometric interpretation. Gauge invariance may also be regarded in a more geometric way by using the language of connections and fiber bundles. Although the ordinary derivatives of the wave function

$\nabla\psi$ and $\partial_t\psi$ are not gauge invariant, the covariant derivatives, defined by Eqs. (6), acting on the wave

$$\mathbf{D} = \nabla - \frac{iq}{\hbar c}\mathbf{A} \qquad D_t = \partial_t + \frac{iq}{\hbar}\phi \tag{6}$$

function $\psi(\mathbf{x}, t)$ give gauge-invariant quantities. The idea of a covariant derivative is based on the notation of parallel transport. In electromagnetism the gauge field $A_\mu(x)$ defines a parallel transport for quantum phases (that is, how quantum phases at different space-time points should be compared). If C is a path connecting the space-time point x to the point y, then the quantity $U(x, y; C)$ defined by Eq. (7) is

$$U(x, y; C) = \exp\left(\frac{iq}{\hbar c}\int_{C_{x\to y}} dx^\mu A_\mu\right) \tag{7}$$

the operator used to parallel-transport the phase of a field of charge q from x to y along the path C. (In this and subsequent equations, whenever an index such as μ is repeated, it is to be summed over its possible values, in this case 0, 1, 2, 3.) The expression given by Eq. (8) is then gauge invariant. In general,

$$\Psi(y)^\dagger U(x \to y;\ C)\Psi(x) \tag{8}$$

the parallel transport depends on the choice of path C connecting the end points. This path dependence leads to the notion of curvature. Upon parallel transport around a closed path C, starting and ending at x_0, the wave function acquires a phase given by Eq. (9),

$$U = \exp\left[\frac{iq}{\hbar c}\oint_C dx^\mu A_\mu\right] = \exp\left[\frac{iq}{\hbar c}\int_S dS^{\mu\nu}F_{\mu\nu}\right] \tag{9}$$

where S is a two-dimensional surface in space-time whose boundary ∂S is C. The antisymmetric tensor $F_{\mu\nu}$, also known as the gauge curvature, is the electromagnetic field strength. On an infinitesimal level, the gauge curvature reflects the lack of commutation of the covariant derivatives, according to Eq. (10). Here,

$$F_{\mu\nu} = \frac{-i}{(q/\hbar c)}[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{10}$$

$[D_\mu, D_\nu]$ stands for $D_\mu D_\nu - D_\nu D_\mu$, called the commutator of D_μ and D_ν. *See* FIBER BUNDLE; RIEMANNIAN GEOMETRY.

Quantum electrodynamics. The action for an electromagnetic field coupled to a free scalar field is given by Eq. (11). Here $D_\mu = (\partial_\mu + ieA_\mu)$, where

$$S = \int d^4x\left[\frac{-1}{4}F_{\mu\nu}F_{\mu\nu} + [D_\mu\Psi]^\dagger[D^\mu\Psi]\right] \tag{11}$$

e is the charge of the electron and units are chosen such that $\hbar = c = 1$. When the fields $A_\mu(x)$ and $\phi(x)$ are taken to be quantum fields rather than classical fields, the resulting theory is called quantum electrodynamics (QED). This theory cannot be solved exactly; however, QED may be formulated perturbatively, as a power series in the fine-structure constant $\alpha = e^2/\hbar c$. Originally, perturbative QED was fraught with difficulties, because higher powers in α gave infinities. A successful program of regularization and renormalization (to render finite these infinities and to relate the bare parameters appearing in the action to physically observable quantities) was developed independently by R. Feynman, J. Schwinger, and S. Tomonaga in the late 1940s, making QED a theory consistent to all orders of perturbation theory. *See* ACTION; QUANTUM ELECTRODYNAMICS; RENORMALIZATION.

One effect of higher-order quantum corrections in QED is vacuum polarization, which causes running of the coupling constant α. At the classical level (to lowest order in α) the strength of the electromagnetic interaction is the same at all length scales (that is, the inverse square law is valid). However, radiative corrections cause α to increase as increasingly smaller length scales are probed, because of vacuum polarization. Heuristically, this may be understood as screening. On length scales shorter than the electron Compton wavelength, a polarized cloud of virtual electron-positron pairs shields bare charge because of the strong electric field, in much the same way as a charge in a dielectric medium is partially shielded because of the polarization of the dielectric medium. On smaller scales, more charge is seen, so that the effective strength of the coupling seems to increase.

Nonabelian gauge theory. The concept of a gauge theory may be generalized to larger, nonabelian Lie groups, such as $G = \mathrm{SU}(2)$, $\mathrm{SU}(3)$, $\mathrm{SU}(5)$, or $\mathrm{SO}(10)$. A nonabelian Lie group G has generators τ_a ($a = 1, \ldots, d$), with the commutation relations $[\tau_a, \tau_b] = f^c_{ab}\tau_c$. A function $\Omega(x)$, which depends on the space-time point x and has values in the group G, will be used to generate a gauge transformation. This function can be written in the form $\Omega(x) = \exp[i\omega^a(x)\tau_a]$, at least locally, where the quantities ω^a are generalizations of the quantity Λ introduced in electrodynamics. The theory has matter fields, transforming according to the representations $\Gamma^{(i)}$ of G, respectively. Each representation consists of a set of hermitian matrices satisfying the commutation relations $[\Gamma^{(i)}_a, \Gamma^{(i)}_b] = f^c_{ab}\Gamma^{(i)}_c$. The gauge field A_μ takes values in the Lie algebra and may be expanded in components as $A_\mu = A^a_\mu\tau_a$. Under a gauge transformation, these fields transform according to Eqs. (12), where the

$$\Psi^{(i)} \to \Psi^{(i)\prime} = \Omega^{(i)}\Psi^{(i)} \qquad A_\mu \to A'_\mu = \Omega^{(a)}A_\mu[\Omega]^{(a)-1} + i[\Omega]^{(a)-1}\left(\partial_\mu\Omega^{(a)}\right) \tag{12}$$

superscript (a) indicates the adjoint representation. The covariant derivative is $D_\mu = (\partial_\mu + igA^a_\mu\Gamma_a)$, where the representation of Γ_a corresponds to that of the field on which D_μ acts. The gauge curvature is given by Eq. (13), and transforms under gauge

transformations according to Eq. (14).

$$F^a_{\mu\nu} = -[D_\mu, D_\nu]^a = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^a_{bc} A^b_\mu A^c_\nu \quad (13)$$

$$F_{\mu\nu} \to F'_{\mu\nu} = \Omega^{(a)} F_{\mu\nu} [\Omega^{(a)}]^{-1} \quad (14)$$

It is customary to absorb the coupling constant g into a rescaling of the gauge field, so that the action has the form given by Eq. (15), where tr $[F_{\mu\nu} F^{\mu\nu}]$ is

$$S = \int d^4x \left[\frac{-1}{4g^2} \mathrm{tr} \left[F_{\mu\nu} F^{\mu\nu} \right] + \sum_{(i)} \left(D_\mu \Psi^{(i)} \right)^\dagger \left(D_\mu \Psi^{(i)} \right) \right] \quad (15)$$

the trace of the product of the matrices $F_{\mu\nu}$ and $F^{\mu\nu}$. *See* GROUP THEORY; LIE GROUP; MATRIX THEORY.

Electroweak and strong interactions. In 1954 C. N. Yang and R. L. Mills suggested gauging the SU(2) isospin symmetry, thus developing the first nonabelian gauge theory, also known as Yang-Mills theory. In 1971 G. 't Hooft demonstrated the renormalizability of nonabelian gauge theory. Nonabelian gauge theory is the foundation of the electroweak and strong interactions. In the electroweak theory, formulated by S. Weinberg and A. Salam in 1967, the gauge group $G = \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ is spontaneously broken to $\mathrm{U}(1)_Q$, the gauge group of ordinary electromagnetism, by the condensation of a Higgs field. There are four kinds of particles that mediate the gauge interactions, called gauge bosons: two massive charged weak gauge bosons, the W^+ and W^-, with a mass of 80.6 GeV; a neutral gauge boson, the Z, with a mass of 91.2 GeV; and finally a massless photon (γ). The $W^\pm$ and the Z are responsible for the charged and neutral weak interactions, respectively, and the photon is the gauge particle responsible for the electromagnetic interaction. In quantum chromodynamics (QCD) the symmetry group is $\mathrm{SU}(3)_{\mathrm{color}}$, which remains unbroken. There are eight gauge bosons called gluons. Quarks, which come in three colors, carry a color charge, and the quark field forms an SU(3) triplet. *See* ELECTROWEAK INTERACTION; SYMMETRY BREAKING.

Asymptotic freedom and confinement. One remarkable property of quantum chromodynamics is that color charge is antiscreened rather than screened, a feature arising from the nonabelian character of the gauge symmetry. This situation is opposite to that in QED. Consequently, as shorter distance scales are probed, the coupling constant α_{QCD} decreases, so that at very short distances QCD approaches a free (noninteracting) field theory. This feature, pointed out in 1973 by D. Gross and F. Wilczek and by H. D. Politzer, is known as asymptotic freedom and provides the theoretical foundation for the parton model. Although on large scales QCD is a strongly coupled theory, at very small scales, which may be probed by scattering at very high energies, the constituents of hadrons (quarks and gluons) behave almost as if they were free particles, originally called partons. Similarly, at larger distances scales α_{QCD} becomes larger, and perturbation theory breaks down. This increase in the coupling constant at large distances leads to a phenomenon known as confinement, which prevents colored objects (such as quarks or gluons) from being isolated. *See* GLUONS; QUANTUM CHROMODYNAMICS; QUARKS.

Gravity. Gravitation, as defined by the general theory of relativity formulated by A. Einstein in 1915, is also a gauge theory. In gravity the gauge transformations are the general coordinate transformations. Perturbatively, the gravitational interaction may be interpreted as the result of graviton exchange. The graviton is the spin-2 gauge particle associated with general coordinate invariance. Unfortunately, perturbative gravity is a nonrenormalizable theory, and it has not proved possible to formulate a consistent quantum theory of gravity. String theory shows some promise. By replacing point particles with strings, the divergences in higher orders of perturbation theory are softened. The fact that superstring theories are perturbatively renormalizable and have spin-2 excitations suggests that string theory might provide a correct theory of quantum gravity. However, at this point the understanding of string theory has not progressed far enough to determine whether this hope will be realized. To formulate the correct theory of quantum gravity remains one of the great challenges of theoretical physics. *See* GRAVITATION; QUANTUM GRAVITATION; RELATIVITY; SUPERSTRING THEORY.

Symmetry. One of the beautiful aspects of gauge theory is the overarching role that symmetry plays. Gauge symmetry is more than a simplifying principle that places restrictions on the possible dynamics of a theory; rather, it is the symmetry itself that determines the interactions. *See* ELEMENTARY PARTICLE; QUANTUM FIELD THEORY; SYMMETRY LAWS (PHYSICS).

Martin Bucher

Bibliography. E. Commins and P. Bucksbaum, *Weak Interactions of Leptons and Quarks*, 1983; C. Itzkyson and J.-B. Zuber, *Quantum Field Theory*, 1980; A. Pais, *Inward Bound*, 1986; C. Quigg, *Gauge Theories of Strong, Weak, and Electromagnetic Interactions*, 1983, reprint 1997; J. Schwinger (ed.), *Selected Papers on Quantum Electrodynamics*, 1958; F. Wilczek, Quantum chromodynamics: The modern theory of the strong interaction, *Annu. Rev. Nucl. Part. Sci.*, 32:177, 1982.

Gauss' theorem

The assertion, under certain light restrictions, that the volume integral through a volume V of the divergence of vector function $\Im(x, y, z)$ is equal to the surface integral of the exterior normal component of $\Im$ over the boundary surface S of V, or in symbols $\iiint \nabla \cdot \Im \, dV = \iint \Im \cdot \nu \, dS$, with ν the unit exterior normal to S. This theorem is also known as the divergence theorem and as Green's theorem.

A simple illustration is provided by the radius vector $\mathfrak{r}$ ($\mathfrak{r} = xi + yi + zf$), and the sphere $S(O, a)$ (center at O, radius a). Because $\nabla \cdot \mathfrak{r} = 3$ and on S, $\nu = \mathfrak{r}a$, $\mathfrak{r} \cdot \nu = a$, it follows that $\iiint \nabla \cdot \mathfrak{r} \, dV = 3(4\sqrt{3}\pi a^3)$, while $\iint \mathfrak{r} \cdot \nu \, dS = a(4\pi a^2)$.

A common method of proof consists of the conversion of $\int\int\int, \nabla \cdot \Im \, dV$ into a sum I of three double integrals with domains in the coordinate planes followed by the conversion of I to $\int\int \Im \cdot \nu \, dS$. The function $\Im$ may be assumed to be of class C' and the volume V to be convex with smooth boundary pieces. The class of admissible volumes may be extended by the introduction of protuberances and cavities of the previous type.

The divergence theorem plays an important role in a variety of subjects such as electricity and magnetism, mechanics of continuous media (including fluid dynamics), heat flow, partial differential equations, and potential theory. *See* CALCULUS OF VECTORS; GREEN'S THEOREM. Homer V. Craig

Gavial

The name of two species of reptiles which form the family Gaviidae in the order Crocodilia. The Indian gavial (*Gavialis gangeticus*) is confined to the Ganges River and its tributaries and may attain a length of 30 ft (9 m), while the Malayan gavial (*Tomistoma schlegeli*) is found in Borneo and Sumatra and has a maximum length of about 15 ft (4.5 m).

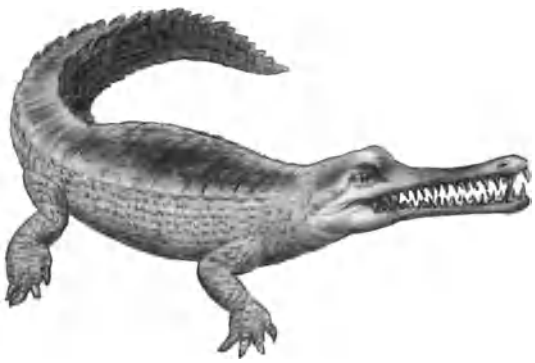

Indian gavial (*Gavialis gangeticus*).

The gavial is distinguished from other members of the order by its extremely long, slender snout (see **illus.**). The tip of the snout is enlarged, and fleshy elevations surround the openings of the nostrils. The mouth cavity is lined with long, stout, sharp-edged teeth, the upper set interlocking with the lower. These animals feed almost exclusively on fish. They are timid animals and, when alarmed, move with great rapidity into the nearby water.

The females lay about 30–40 eggs in nests which they have prepared on the riverbanks. The newly hatched young are about 1 ft (0.3 m) long and quite active. *See* ALLIGATOR; CROCODILE; CROCODYLIA; REPTILIA. Charles B. Curtin

Gaviiformes

A small order of aquatic birds that contains a single living family, the Gaviidae (loons), with five species restricted to the Northern Hemisphere, and a fossil family, the Colymboididae. *See* AVES.

Common loon (*Gavia immer*), Jammer Lake, Minnesota. (© *DonGettyPhoto.com.*)

Phylogeny and fossil record. Relationships of the loons to other birds is uncertain; suggested affinities with the charadriiforms, the fossil hesperornithiforms, or the grebes are disputed. The earliest fossil loons may be *Polarornis* from the Cretaceous of Antarctica. Later forms are the small, less specialized *Colymboides*, known from the Eocene and Miocene of Europe and placed in a distinct family, the Colymboididae. Modern loons of the genus *Gavia* are known from the lower Miocene. *See* CHARADRIIFORMES; HESPERORNITHIFORMES; PODICIPEDIFORMES.

Characteristics. Loons, or divers, are large swimming and diving birds that catch fish with their strong bills in underwater pursuits. They are foot-propelled swimmers, with short legs placed far back on the body and webbed feet; land locomotion is awkward. Loons take wing with difficulty but are strong fliers. Sexes are similar in plumage, which is white below and checked black and white above (see **illustration**). Loons are monogamous, have a strong pair bond, and engage in elaborate courtship rituals that include loud, quavering calls often given in flight. Breeding is solitary and takes place on freshwater lakes. The young are downy, leave the nest immediately after hatching, and are cared for by both parents. Because all flight feathers are molted simultaneously, the birds are rendered temporarily flightless. Loons migrate south to wintering areas, which are mostly along ocean coasts.

Status. Common loons (*Gavia immer*; see illustration) in the southern part of their breeding range have been decreasing in number because of direct disturbance from boating on lakes and the effects of mercury poisoning as a consequence of chemical action of acid rain on deposits in lakes. *See* ENDANGERED SPECIES. Walter J. Bock

Bibliography. C. Carbonear, Order Gaviidae (divers), in J. del Hoyo et al. (eds.), *Handbook of*

the Birds of the World, vol. 1, pp. 162–172, Lynx Edicions, 1992.

Gear

A machine element used to transmit motion between rotating shafts when the center distance of the shafts is not too large. Toothed gears provide a positive drive, maintaining exact velocity ratios between driving and driven shafts, a factor that may be lacking in the case of friction gearing which is subject to slippage. While the motion transmitted to mating gears is kinematically equivalent to that of rolling surfaces identical with the gear pitch surfaces, the action of one gear tooth on another is generally a combination of rolling and sliding motion. When the distance between shafts is large, other methods of transmission are used. *See* BELT DRIVE; CHAIN DRIVE; ROLLING CONTACT.

The application of gears for power transmission between shafts falls into three general categories: those with parallel shafts, those for shafts with intersecting axes, and those whose shafts are neither parallel nor intersecting but skew (see **table**).

Principal features. Figures 1 and **2** illustrate the principal features of toothed gears. Such terms as pitch circle, addendum circle, and root circle, being geometrical, are defined by the diagrams.

Definitions for other commonly used terms for describing gears are addendum, the radial distance between the pitch circle and the addendum circle; dedendum, the radial distance between the pitch circle and the root circle; face, the tooth surface outside the pitch circle; flank, the tooth surface between the pitch circle and the root circle; clearance, the amount by which the dedendum exceeds the addendum of the mating gear; whole depth, addendum plus dedendum; and working depth, whole depth minus clearance; circular pitch, the distance from a point on one tooth to the corresponding point on the next, measured along the pitch circle; and base pitch, the similar measurement along the base circle.

Terms descriptive of a pair of spur gears in mesh are shown in Fig. 2. Where the pitch circles are tangent is pitch point *P*. Each pair of teeth always has its point of contact on the pressure line. The angle between this pressure line and the common tangent to the pitch circles through *P* is the pressure angle. As Fig. 2 illustrates, the pressure angle at which two gears actually operate depends on the distance between centers at which they are mounted. Initial contact between each pair of teeth occurs at point *C*, where the addendum circle of the driven gear crosses the pressure line; and final contact is at *D*, where the driver's addendum circle crosses the pressure line. Distance *CD* is the length of the path of contact. It must exceed the value of the base pitch if a pair of teeth is to come into contact before the previous pair has gone out of contact. The ratio of the length of the path of contact to the base pitch is called the contact ratio, which can be viewed as approximately the average number of pairs of teeth in contact. It is usual to design for a contact ratio of 1.4 or above to ensure smooth, continuous tooth action.

Chief types of application of gears for power transmission

Shaft relationship	Type of gearing
Parallel	Spur
	Helical
	Herringbone
	Double helical
Intersecting	Bevel
Skew	Helical
	Worm
	Hypoid

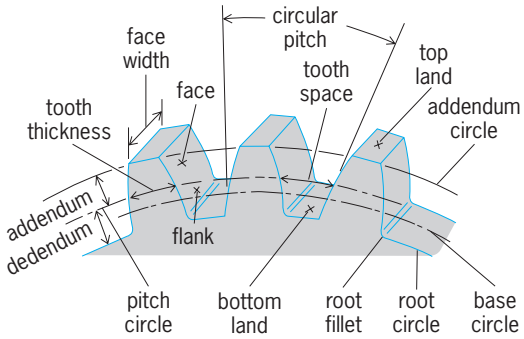

Fig. 1. Principal features of gear teeth.

Gear tooth sizes are designated by diametral pitch, which is the number of teeth per inch of diameter of the pitch circle. Pitch circle is, in turn, the circle whose periphery is the pitch surface, or surface of an imaginary cylinder that would transmit by rolling contact the same motion as the toothed gear. A gear with a 20-in. (51-cm) pitch diameter and teeth having a diametral pitch of 2 has 40 teeth. Diametral pitch *P* times circular pitch *p* (Fig.1) equals π. The diametral pitches most widely used are: 1, 11/2, 2, 21/2, 3, 4, 5, 6, 8, 10, 12, 16, 20, 24, 32, 40, 48, 64, 72, 96, 120, 160.

The smaller of two gears in mesh is the pinion. It has the fewer teeth and is the driving gear in a speed reduction unit. The minimum number of teeth an involute pinion can have and still run without interference between its flanks and the tips of the mating gear teeth is fixed by the tooth system. Smaller pinions are possible only if the pinion's flanks are undercut.

Backlash. The amount of which the tooth space of a gear exceeds the tooth thickness of the mating gear at the pitch circle is the backlash. It can be determined in the plane of rotation or, for helical gears, in the plane normal to the tooth face.

If mating gears have zero backlash, gears and mountings need to be dimensionally perfect. To retain zero backlash with varying operation conditions, all parts need exactly the same thermal expansion characteristic. Because of the difficulty of meeting these requirements and for lubrication, freedom-backlash is provided between gear teeth. The usual practice is to reduce the tooth thickness of each gear by an amount equal to half the desired

backlash. However, in the case of a gear and a small pinion it is customary to reduce the tooth thickness on the gear only, leaving the pinion with standard tooth thickness. Backlash can also be adjusted by slight changes in the center distance between gears. Except for a small change in the pressure angle, the action of involute gear teeth is not affected by backlash or center distance adjustment.

In the case of precision gearing for control systems and similar applications, backlash results in a nonlinear relation between input and output. Several methods of reducing backlash are in use. One method is to place two identical spur gears on the same shaft, one fixed to the shaft, one free. The loose gear is attached to the fixed one by springs, which keep the composite gear in positive contact with the pinion at all times. A second method is to use tapered-tooth gearing (beveloid) with adjustment along the shaft to eliminate excessive backlash.

Gear action. A principal function of gears is to change the speed of rotation. This action is described by the velocity ratio of the gears in mesh. Ratio VR is the number of revolutions N_1 of the driving gear divided by the number of revolutions N_2 of the driven gear in the same time interval. For gears with teeth T_1 and T_2, respectively, VR is expressed as in Eq. (1).

$$VR = \frac{N_1}{N_2} = \frac{T_2}{T_1} \tag{1}$$

When two curved surfaces, such as the mating surface of two gear teeth, are in driving contact, there is a definite velocity ratio between the bodies. The angular velocities of the two bodies are inversely proportional to the segments into which their line of centers is divided by a line passing through their point of contact and normal to their surfaces at this point. Thus a constant angular velocity ratio between bodies in driving contact demands that the common normal to the profiles at the point of contact cut the center line at a fixed point, the pitch point. This latter statement is frequently referred to as the fundamental law of gear tooth action. Pure rolling contact between gear teeth occurs only when they are in contact at the pitch point. At all other positions the teeth have some sliding with the maximum sliding at the first and last instants of contact. Although there are a number of tooth forms that will satisfy the fundamental law, only two of them have been used to any great extent. The cycloidal tooth predominated until the late 1800s but has been replaced to a great extent by the involute gear tooth. Cycloidal teeth are still found in instruments, watches, clocks, and, occasionally, in cast and cut gears.

Gears are said to be interchangeable when any gear of the set will run with any other gear of the same set. Actually there is little need today for gears that are interchangeable. In the manufacture of the overwhelming proportion of machines pairs or groups of gears are designed to mesh with each other and no others. Standardization in manufacture is common, however. Each of the gears is made from one of a set of standardized cutters. Thus it is the standardization of the production tools rather than of the gears themselves that is of most importance.

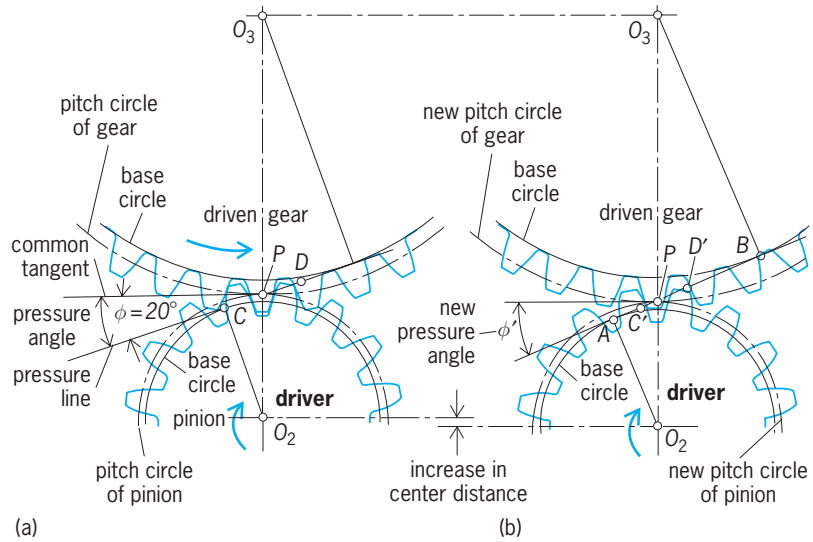

Fig. 2. Action of involute gearing. (*a*) Gear and pinion mounted at normal center distance. (*b*) Gear and pinion mounted at greater than normal center distance. (*After J. E. Shigley, Theory of Machines, McGraw-Hill, 1961*)

Involute gear teeth. An involute tooth is laid out along an involute (**Fig. 3**), which is the curve generated by a point on a taut wire as it unwinds from a cylinder. The generating circle is called the base circle of the involute. Proportions of the tooth are fixed by the gear tooth system and the diametral pitch. The involute curve establishes the tooth profile outward from the base circle. From the base circle inward, the tooth flank ordinarily follows a radial line and is faired into the bottom land with a fillet. The basic rack form of the involute tooth has straight sides, an important property from the manufacturing standpoint.

Gear teeth may interfere with each other, especially where pinions have a relatively small number of teeth. In Fig. 2, point of initial contact C occurs on the pressure line to the right of (after) the point of tangency of the pressure line and the pinion's base

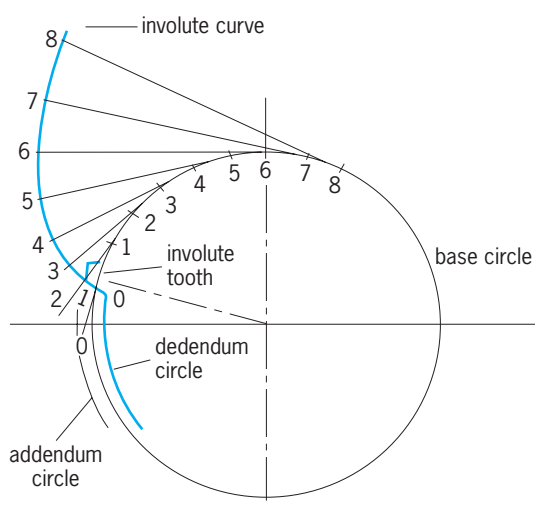

Fig. 3. Method of generation for face of involute tooth.

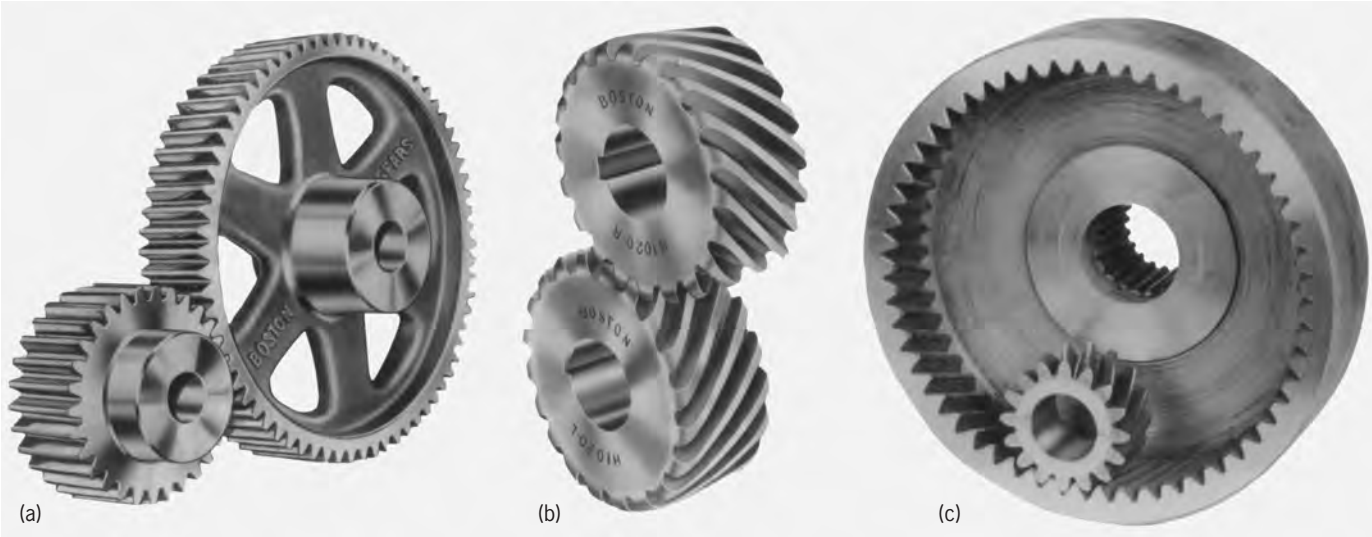

Fig. 4. Spur gears. (*a*) External spur gear and pinion, the commonest type of spur gear. (*b*) Single helical gearset (*Boston Gear Works*). (*c*) Helical internal gear and pinion (*Fellows Gear Shaper Co.*).

circle. Here there is no interference. But if point C were to the left of the tangency point, this would indicate premature contact between the teeth, a contact occurring on the noninvolute portion of the pinion tooth flank below the base circle. The tip of the gear tooth, in such a case, digs into the flank of the pinion tooth.

There are several ways to eliminate interference: (1) Reduce addendum of the gear. (2) Increase pressure angle. (3) Increase backlash by increasing center distance between gears. (4) Undercut flank of the pin ion. (5) Relieve or modify face of the gear tooth.

Spur gears. In the truest sense, spur gears are only those that transmit power between parallel shafts and have straight teeth parallel to the gear axis (**Fig. 4***a*). It is common practice, however, to group helical gears that have parallel shafts under the heading of spur gears (Fig. 4*b*). The pitch surfaces of gears with parallel axes are rolling cylinders, and the motion these gears transmit is kinematically equivalent to that of the rolling pitch cylinders. Spur gears are classified as external, internal, and rack and pinion. External spur gears, the most common, have teeth which point outward from the center of the gear. Internal or annular gears have teeth pointing inward toward the gear axis (Fig. 4*c*). A rack may be considered as a gear having an infinite pitch circle radius. Thus its pitch surface is a plane. A rack and pinion running together transform rectilinear motion into rotary motion, or vice versa.

To standardize manufacture of gears, the American Gear Manufacturers Association has adopted three basic types of spur-gear tooth, with pressure angles of 20° and 25° for coarse-pitch gears (with a diametral pitch under 20) and 20° for fine-pitch gears (diametral pitch of 20 and above).

An internal gear has the positions of the addendum and dedendum reversed from those of an external gear. This results in a different tooth action and less slippage than with an equivalent external spur gear. For a given tooth ratio, the arc of action of an internal gear is slightly greater than that of an external gear of the same size and the tooth is stronger. The nature of an internal gear makes it especially suited to closer center distances than could be used with

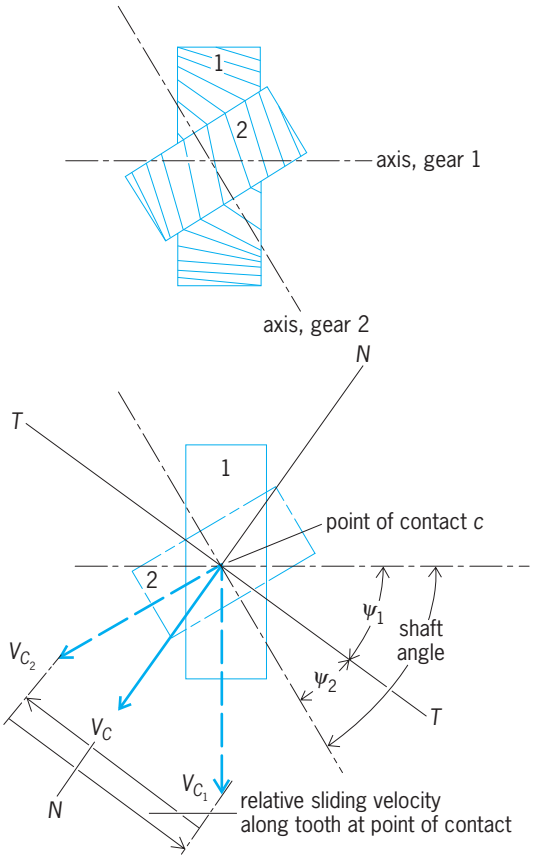

Fig. 5. Action of crossed helical gears.

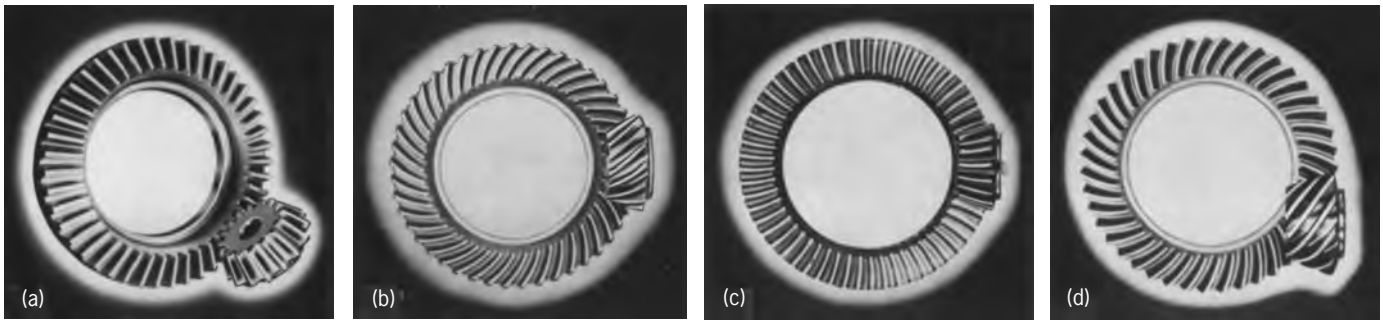

Fig. 6. Bevel gears with pinions. (*a*) Straight. (*b*) Spiral. (*c*) Zero. (*d*) Hypoid. (*Gleason Gear Works*)

an external gear of the same size. When it is necessary to maintain the same sense of rotation for two parallel shafts, the internal gear is especially desirable because it eliminates the need for an idler gear. These conditions make the internal gear highly adaptable to epicyclic and planetary gear trains.

Noncircular gears are used to obtain velocity ratios that vary in a precise manner. Elliptical gears are an example of noncircular gears. They provide a convenient method of obtaining a quick return for machines that do most of their work during only a portion of the drive shaft revolution. Noncircular gears are used in computing mechanisms and other devices where a prescribed varying output function is to be obtained using a linear input.

Helical gears. Gears running on parallel axes and with teeth twisted oblique to the gear axis are essentially spur gears. Because of the twist, contact is progressive across the tooth surface, starting at one edge and proceeding across the face of the tooth. The action results in reduced impact and quieter operation, particularly at high speed. Herringbone gears are equivalent to two helical gears of opposite hand placed side by side. They are especially suited for high-speed operation and eliminate the axial thrust produced by single helical gears. Double helical gears have a central groove for tool runout, making it possible to finish the teeth by a shaving operation. They can be run at even higher speeds than herringbone gears. Dimensions of a helical gear are determined on both the plane of rotation and on the plane perpendicular to the helix angle of the tooth. Thus, a helical gear has a circular pitch measured on the plane of rotation and a normal circular pitch measured on the normal plane.

Crossed helical gears. Where shafts cross obliquely, motion is transmitted by crossed helical gears (**Fig. 5**). The teeth are helical but differ from the teeth of worm gears in that no one tooth (thread) makes a complete turn on the pitch circle. Pitch surfaces of crossed helical gears are cylindrical as with spur gears. However, with crossed shafts, the oblique teeth have point contact rather than the line contact that occurs with parallel shafts. Analysis of the gears is based on equal components of the pitchpoint velocity on each mating gear along common normal N–N. Sliding occurs in direction T–T of the tooth elements. As with spur gears, the revolutions per unit time are inversely proportional to the numbers of teeth. Velocity ratio VR may also be expressed as in Eq. (2), where D is pitch diameter and ψ helix angle.

$$VR = \frac{N_1}{N_2} = \frac{D_1 \cos \psi_1}{D_2 \cos \psi_2} \tag{2}$$

Helix angle is between the shaft axis and a line tangent to the tooth through the pitch point.

Helical gears are referred to as right- or left-hand in the same manner as screw threads, a right-hand gear being one on which the teeth twist clockwise as they recede from an observer looking along the axis.

Bevel gears. Where shafts intersect, bevel gears transmit the motion. Such gears may be used only to change the shaft axis direction or to change speed as well as direction. Two bevel gears with equal numbers of teeth and running together with their shaft axes intersecting at 90° are called miter gears. Several forms of bevel gears are in use, including straight-tooth, spiral, and skewed bevel gears (**Fig. 6**).

External bevel gears have pitch angles less than 90° (**Fig. 7*a***). Internal bevel gears have pitch angles greater than 90°, hence their pitch cones are inverted (Fig. 7*b*). A crown gear is one having a pitch angle of 90° (Fig. 7*c*). Thus its pitch surface is a plane, and the crown gear corresponds in this respect to a rack in spur gearing.

Straight bevel gears. The simplest form of bevel gear has straight teeth which, if extended inward, would come together at the intersection of the shaft axes. This point is also the apex of the rolling cone, which forms the pitch surface of the gear. Much of the terminology applied to bevel gears is the same as that used for spur gears. Additional terms used in reference to bevel gears are given in **Fig. 8**. Diametral

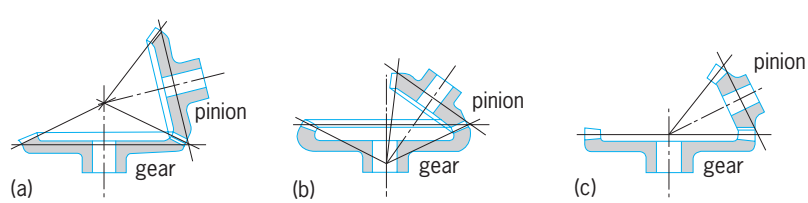

Fig. 7. Bevel gears. (*a*) External. (*b*) Internal. (*c*) Crown. (*After G. L. Guillet and A. H. Church, Kinematics of Machines, 5th ed., Wiley, 1950*)

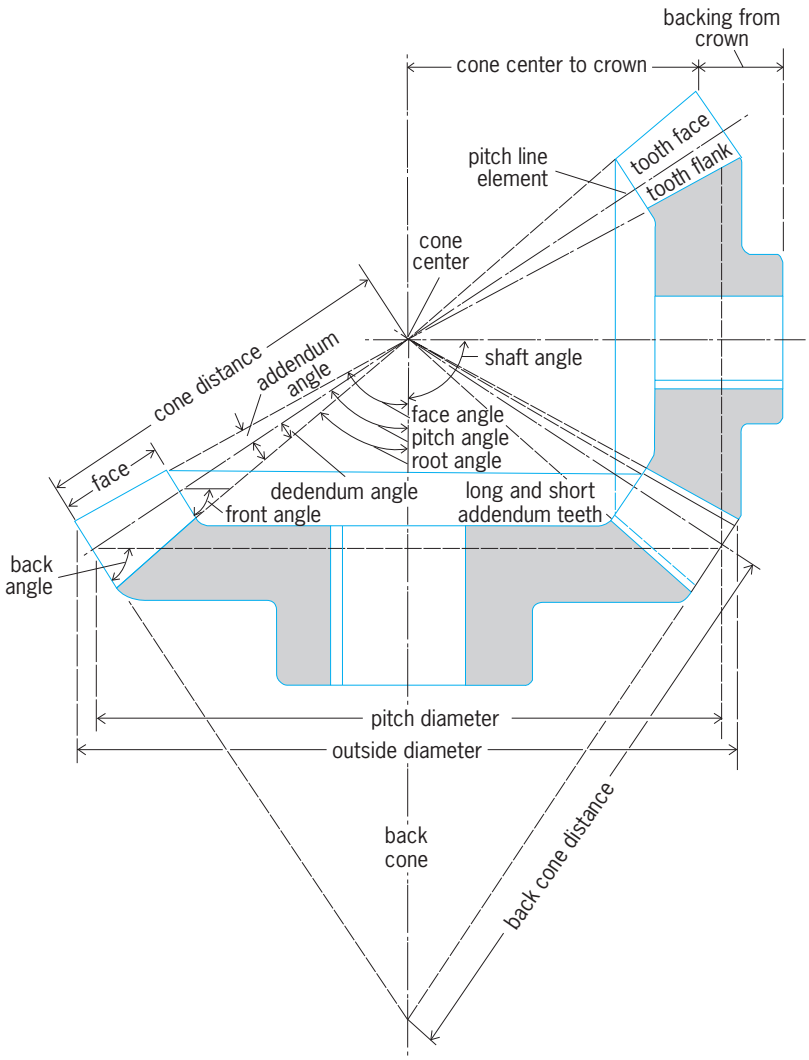

Fig. 8. Characteristics of bevel gears. (*After C. W. Ham, E. J. Crane, and W. L. Rogers, Mechanics of Machinery, 4th ed., McGraw-Hill, 1948*)

pitch of a bevel gear is not constant across the full width of the tooth. The diametral pitch at the pitch diameter is used in fixing tooth proportions. The formative number of teeth in a bevel gear is the number of teeth that would be a spur gear whose pitch radius equaled the back cone distance of the bevel gear.

Speeds of the shafts of bevel gears (velocity ratio) are inversely proportional to the numbers of teeth on the gears or to the sines of the pitch angles, but not to the formative number of teeth. Use of straight bevel gears is limited to low-speed operations, ordinarily below 1000 surface feet per minute or, in the case of small gears, 1000 rpm.

Because each point on a straight tooth bevel gear remains a fixed distance from the pitch cone apex, there is no sliding along the tooth as it engages. Contact across the full tooth face occurs instantaneously—as with spur gears—as the teeth come into mesh.

Spiral bevel gears. To provide a gradual engagement, as contrasted to the full line engagement of straight bevel gears, the teeth of spiral bevel gears are curved and oblique. Theoretically the curve is a spiral but, to facilitate manufacture, the curve is actually a circular arc which, within the tooth face width, closely approximates a spiral. This tooth inclination brings more teeth in contact at any one time than with an equivalent straight-tooth bevel gear. The result is smoother and quieter operation, particularly at high speeds, and greater load-carrying ability than with straight bevel gears of the same size.

Spiral bevel gears are used in sewing machines, motion picture equipment, machine tools, and other applications where quiet, smooth operation is essential. They should, in general, be mounted on antifriction bearings because of the axial thrust due to the oblique teeth. In the past they have been used extensively in the rear axle drives of automobiles, but are being replaced by hypoid gears.

Zero bevel gears. A special form of bevel gear has curved teeth with a zero-degree spiral angle. Thus the teeth are not oblique as is the case with spiral bevel gears. Rather, the teeth lie in the same general direction as those of an equivalent straight-tooth bevel gear, and so the gears are usually used in the same types of drives as the straight-tooth gear. As with straight bevels, they produce no axial thrust and, therefore, may be used without thrust bearings. The face that they may be produced on the same equipment as spiral, bevel, and hypoid gears makes them economically desirable.

Tooth proportions for bevel gears follow the standards established by the Gleason Works and adopted as the recommended standard of the American Gear Manufacturers Association. Tooth proportions are a function of the velocity ratio. Thus, bevel gears are not as interchangeable as spur gears.

Hypoid gears. To connect nonparallel, nonintersecting shafts, usually at right angles, hypoid gears are used. They are similar to spiral bevel gears in their general appearance. The axis of the hypoid pinion may be offset above or below the axis of the gear. The shape of the tooth is similar to that of the spiral bevel gear and gives progressive contact across the tooth. In operations these gears run even more smoothly and quietly than spiral bevel gears. To maintain line contact of the teeth, with the offset shaft, the pitch surface of the hypoid gear is a hyperboloid of revolution rather than a cone as in bevel gears.

One of the first uses of hypoid gears was in the rear-axle drive of Packard automobiles. The operating smoothness of hypoid gears, along with the lower body lines made possible by the offset pinion shaft, has made them extremely popular for automotive use. Industrial applications of the hypoid gear also take advantage of the pinion offset, which allows the mounting of any number of pinions on a single continuous shaft, a feat not possible with bevel gears. The shaft arrangement of the hypoid gear and pinion enables bearings to be placed on both sides of the gear and of the pinion. The offset axis results in a larger and, consequently, stronger hypoid pinion tooth than on an equivalent straight-tooth or spiral bevel gear. Expressed in terms of pitch

diameter, a hypoid pinion has fewer teeth than a spiral bevel pinion of the same pitch diameter. It is possible to use hypoid pinions having 7, 8, or 9 teeth in contrast to a minimum of 12, 13, or 14 teeth—depending on the velocity ratio—on a spiral bevel pinion. Lubricants must withstand the higher loading and the sliding that occurs along the teeth of hypoid gears.

Hypoid gears are suitable for large velocity reductions; reduction ratios of 60:1 and higher are entirely feasible. In general, shaft offset should not exceed 40% of the equivalent bevel gear back cone distance and, when the loading is heavy, as in truck and tractor drives, the offset should be nearer 20% of this distance. Direction of the offset, above or below center, must be specified for any given installation. The gears should, in general, be mounted on antifriction bearings in an oil-tight case. Thrust bearings must be provided. Because of the sliding tooth action, the efficiency of hypoid gears is somewhat less than that of equivalent bevel gears.

Worm gears. A chief way to connect nonparallel, nonintersecting shafts that are at right angles is through a worm gear. The worm, ordinarily the driver, is similar to a crossed helical gear except that it has at least one complete tooth (thread) around the pitch surface. The mating gear is the worm wheel or worm gear. Worm gearing is generally used to obtain large velocity reductions with the worm as the driver and the worm wheel as the driven gear, although occasional applications, for example cream separators, have the worm wheel as the driver. The pitch surfaces of straight worms are cylinders and the involute teeth have point contact. Because the appearance of the worm is similar to that of a screw, the teeth are called threads.

The pitch of the worm is the axial distance from any point on one tooth to the corresponding point on the next tooth (**Fig. 9**). This must equal the circular pitch of the mating worm wheel. Lead is the axial distance the worm helix advances in one complete revolution around the pitch surface. A single thread worm has the pitch and lead equal; one revolution of such a worm will, for shafts at right angles, advance the worm wheel $1/N$ revolutions if N is the number of teeth on the worm wheel. A double-threaded worm has the lead equal to twice the pitch and will then advance the worm wheel $2/N$ revolutions per turn of the worm. Thus, worm gears follow the general rule of angular velocity ratio inversely proportional to the ratio of the numbers of teeth. Worms are right- and left-hand in the same sense as helical gears. Changing hand of the worm reverses the relative rotation of the worm wheel.

Improved load-carrying capacity and wear characteristics are made possible by increasing the contact between the worm and wheel (**Fig. 10**). Line contact is obtained by making the worm wheel surface concave to conform to the tooth profile of the worm. Still greater contact is obtained by using a concave worm as well. Known as a cone-drive or Hindley worm, this design permits greater contact surface and allows more teeth to be in contact at one time. *See* GEAR TRAIN; MECHANISM; PLANETARY GEAR TRAIN.

John R. Zimmerman

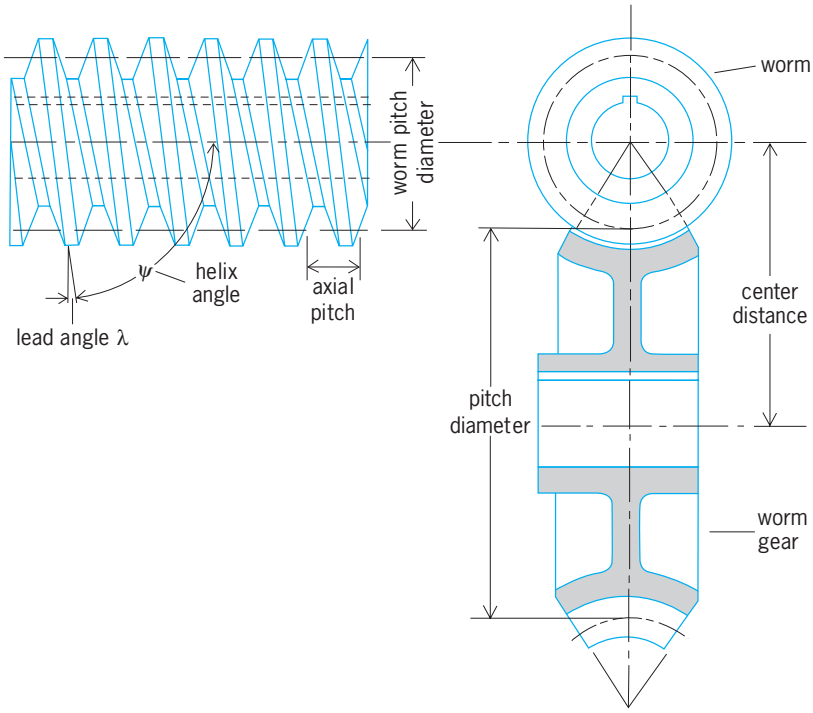

Fig. 9. Nomenclature of a single-enveloping worm gearset. Teeth of worm gear are called threads. (*After J. E. Shigley, Theory of Machines, McGraw-Hill, 1961*)

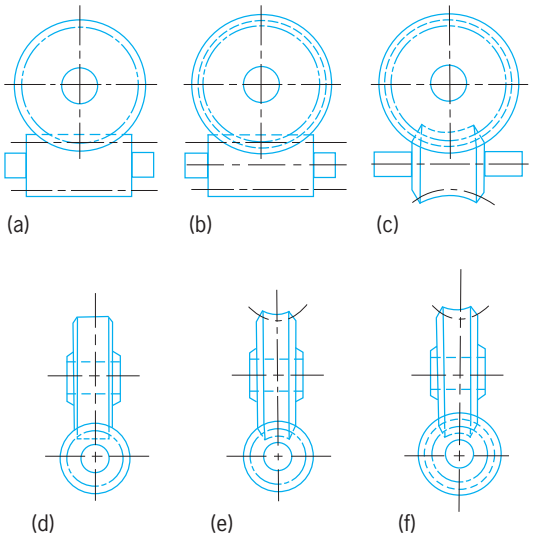

Fig. 10. Nature of contact for worm gears. (*a*) Nonthroated. (*b*) Single-throated. (*c*) Double-throated on cone. (*d*) Point contact. (*e*) Line contact. (*f*) Area contact. (*Michigan Tool Co.*)

Bibliography. N. P. Chironis (ed.), *Gear Design and Application*, 1967; N. P. Chironis, *Mechanisms and Mechanical Devices Sourcebook*, 3d ed., 2001; J. R. Colbourne, *The Geometry of Involute Gears*, 1987; J. W. Patton, *Kinematics*, 1979; B. A. Shtipelman, *Design and Manufacture of Hypoid Gears*, 1978; D. W. South, *Encyclopedic Dictionary of Gears and Gearing*, 1994.

Gear loading

The power transmitted or the contact force per unit length of a gear. If gear speed, size, and tooth contour are fixed, the power-handling capacity may be increased by increasing the axial length of the gear. However, where space is restricted (as in aircraft and many ground vehicles), the gear is usually loaded to its safe limit.

To obtain satisfactory tooth-surface durability from highly loaded gears, experience indicates that several items of the gear set must be properly designed and manufactured, namely, the tooth profile which must be properly modified from a true involute to suit the operating conditions; index of teeth and parallelism of teeth, which must be held within close limits; gearing, which must be mounted so the teeth will not deflect out of line; and gear tooth surfaces, which must be of sufficient hardness and proper finish and which should have good lubrication, particularly on start of initial operation.

Gear design is a compromise between tooth strength and surface durability. Larger teeth give more tooth strength but less surface durability; smaller teeth result in less tooth strength but more surface durability.

Surface distress on the teeth becomes particularly important at loadings in the order of thousands of pounds per inch of face.

In designing the gear tooth surfaces, one mesh of teeth is assumed to carry the load through the action, and the action is balanced so that the product of Hertz maximum compressive pressure P in pounds per square inch and sliding velocity V in feet per second is the same at each end of the assumed single-mesh action. This product, known as the PV value, may be 3,000,000 or higher in gears with high pitch-line speeds for which the designer may use PVT values. The design may produce a smaller contact ratio than is conventional, but this result indicates that contact ratio alone is an unreliable measure of gear capacity.

A highly loaded tooth of adequate rigidity deflects about a point in the middle of the rim, bending as a rigid body under load rather than as a nonuniform beam only. Relief or other modification of tooth profile provides clearance so as to avoid excessive loading at teeth tips due to deflection of the preceding mesh, and ramps at the tooth tips assure that first contact does not extend to the tips.

This design refinement necessitates corresponding care in production, as in minimizing distortion during carburizing. Teeth can be held parallel within 0.0003 in. (7.5 micrometers) in the width of the tooth, and the index can be maintained within 0.0002 in. (5 μm) between adjacent teeth of a gear.

To assure removal of all possible error in grinding when this method of finishing is used and to achieve a more satisfactory working surface on the gear tooth face, grinding wheels are dressed to produce a finish which is coarser than usual. Slight surface roughness (profilimeter readings between 15 and 37) produces less surface distress, or scuffing, of highly loaded steel gears than do smoother finishes, although a steel gear running with a bronze one should be smoother.

Quiet operation is an indication of efficient operation free from abrupt tooth engagement and severe scuffing of one tooth face by another. Helical gears provide smooth transition of effort from tooth to tooth. The gear is designed with a modified involute to provide soft contact at the start of engagement and end of engagement to reduce the pressure on the teeth when the sliding of the tooth surfaces is highest. For quietest operation they should be cut and shaved accurately for lead, profile, and index. A helical contact ratio of 2 or more is desirable. This ratio can be obtained by increasing the width or the helix angle, although effort expended to increase helical contact ratio reduces deflection and PV values more than does the same effort expended to increase involute contact ratio. As tooth size and resultant sliding increase, this proportioning of effort becomes more important. Helix angles in the order of 45° on constant mesh gears require more expensive mountings but improve quietness. Rigid and accurate mounting is essential to achieve best operating condition.

With other gears subject to deflection, such as hypoid rear-axle gearing, conventional practice is to crown the tooth and modify the tooth profile to keep contact away from the tooth ends and tips. The gear set is usually designed for stiffness rather than for strength, and the shafts and mountings should also be designed on this criterion so that little crowning is needed. *See* GEAR. Forest R. McFarland

Bibliography. E. Buckingham, *Analytical Mechanics of Gears*, 1988; D. L. Landau, E. M. Lifshitz, and A. M. Kosevich, *Theory of Elasticity*, 3d ed., 1986, revised 1995; A. Stokes, *Gear Handbook*, 1992; J. P. Smith, *Gears and Their Vibration: A Basic Approach to Understanding Gear Noise*, 1983; D. Townsend, *Dudley's Gear Handboook*: *Design, Manufacture, and Application*, 1992.

Gear train

A combination of two or more gears used to transmit motion between two rotating shafts or between a shaft and a slide. In theory two gears can provide any speed ratio in connecting shafts at any center distance, but it is often not practical to use only two gears. If the ratio is large or if the center distance is relatively great, the larger of the two gears may be excessively large. Moreover, an additional gear may be necessary simply to give the proper direction to the output gear. Belt, rope, and chain drives are frequently used in conjunction with gear trains. *See* BELT DRIVE; CHAIN DRIVE; PLANETARY GEAR TRAIN.

Classification. Gear train classifications include simple, compound, reverted, epicyclic (planetary), and various combinations. The most important distinction is that between ordinary and epicyclic gear trains. In ordinary trains (**Fig. 1***a*), all axes remain stationary relative to the frame. But in epicyclic trains (Fig. 1*b*), at least one axis moves relative to the frame.

In Fig. 1*b* gear *B*, whose axis is in motion, is called a planet. The gears *A* and *C* are sun gears.

An ordinary gear train is a single degree of freedom mechanism: A single input, such as an input to gear *A* of the train in Fig. 1*a*, suffices to control the motions of the other moving members. But an epicyclic gear train (Fig. 1*b*) has two degrees of freedom: Two inputs are necessary. In the epicyclic train of Fig. 1*b*, the two input members are gear *A* and the planet carrier, link *D*. Only if both these members are controlled by external agencies can the motions of gears *B*and *C*be predicted. Frequently one gear of an epicyclic train is fixed. This then is one of the input members with a velocity of zero revolutions per unit time.

A simple gear train is one in which each gear is fastened to a separate shaft (Fig. 1*a*). If at least one shaft has two or more gears fastened to it (**Fig. 2**), the train is compound. The train of Fig. 2 is also a reverted gear train, because the input and output shafts are in line. If the input shaft is not in line with the output shaft, the train is nonreverted.

The 1000-hp (750,000-W) mill drive and pinion stand in **Fig. 3** are an example of a large industrial ordinary gear train. The first speed reduction is with the opposed single helical gears, and the second is through the herringbone gears. This train is compound and nonreverted.

Ratios. For an ordinary gear train, the ratio of the angular velocity of the last driven gear to that of the driving gear is known as the train value. The ratio of the driver's velocity to that of the last driven gear is the velocity ratio. By these definitions train value and velocity ratio are reciprocal quantities. The train value for the ordinary train of Fig. 1*a* is −0.25. The output gear *C*turns at one-fourth the speed of the input gear *A*and in the opposite direction.

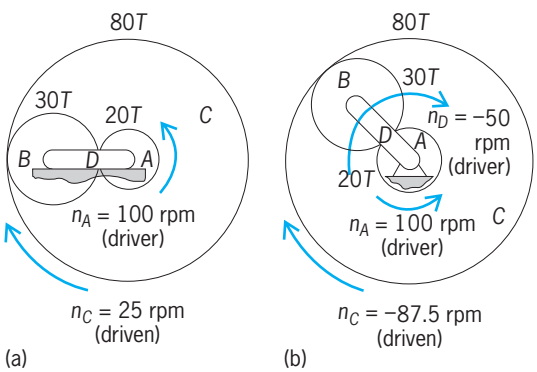

Fig. 1. Gear trains. (*a*) Ordinary. (*b*) Epicyclic.

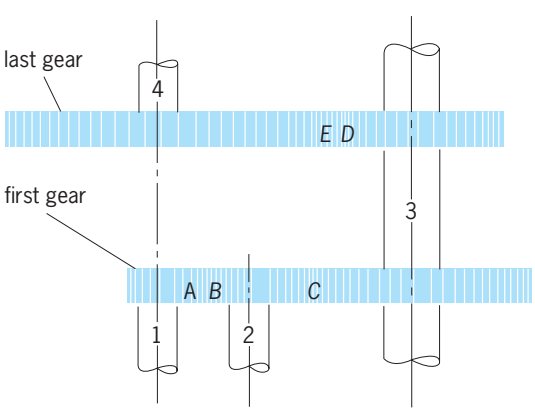

Fig. 2. Compound reverted gear train.

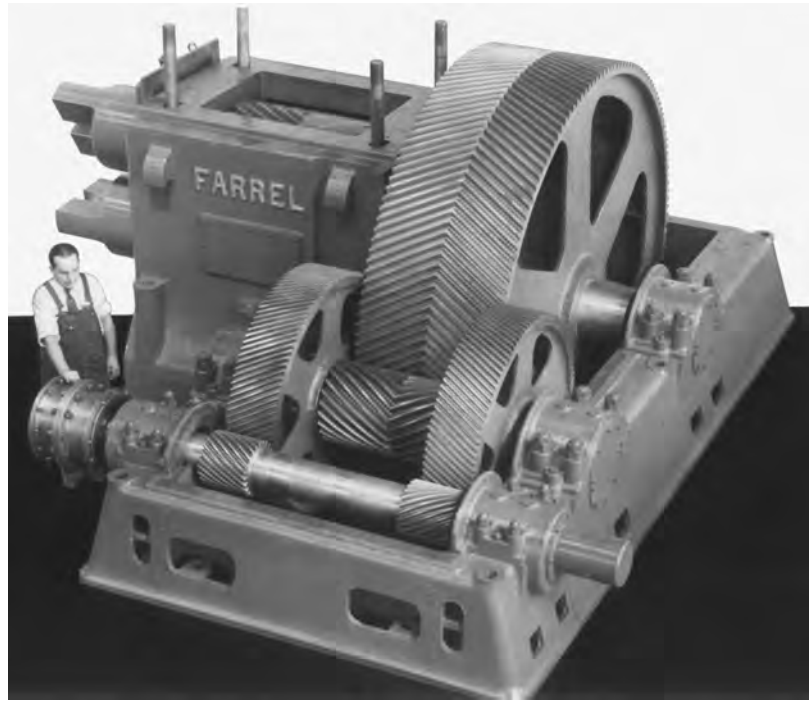

Fig. 3. Helical gear speed reducer. (*Farrel-Birmingham Co.*)

For epicyclic trains the situation can be a little more complicated. Since in general the output velocity is dependent on the two input velocities, the term train value is nonspecific. However, if one gear of an epicyclic train is fixed, the output velocity is some multiple of the velocity of the nonfixed input member; the term train value then applies.

Direction of rotation. The direction of rotation of any gear in an ordinary gear train can be determined by inspection. Mating external gears have opposite directions of rotation, while an internal gear has the same direction as its mating gear.

Ordinary train. Train value is by definition Eq. (1)

$$E = \frac{\text{angular velocity of the last driven gear}}{\text{angular velocity of the first driving gear}} \tag{1}$$

when the angular velocities are measured with respect to the frame supporting the gears. Referring to the ordinary train of Fig. 2, the train value is (ignoring sign) Eq. (2), where n is the speed in revolutions

$$E = \frac{\text{rpm of shaft 4}}{\text{rpm of shaft 1}} = \frac{n_4}{n_1} \tag{2}$$

per unit time. This can be expanded as an identity for the entire train to Eq. (3). Each of the ratios on

$$E = \frac{n_4}{n_1} = \frac{n_2}{n_1} \times \frac{n_3}{n_2} \times \frac{n_4}{n_3} \tag{3}$$

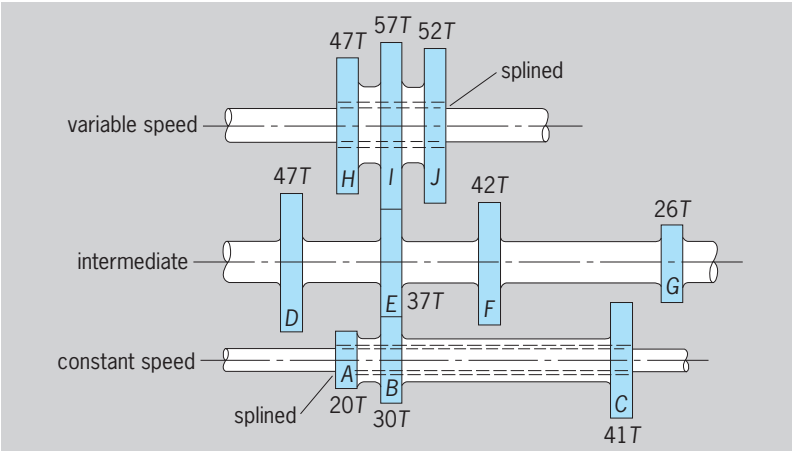

Fig. 4. Compact 9-speed transmission utilizing 10 gears and 3 shafts. (*After J. R. Zimmerman, Elementary Kinematics of Mechanisms, Wiley, 1962*)

the right side of Eq. (3) is the train value for a pair of meshing gears. Since for any two gears in mesh the speeds vary inversely as the numbers of teeth, the train value for the whole train must be (again ignoring sign) Eq. (4), where N is the number of teeth. In Eq. (4) the number of teeth appearing in the nu-

$$E = \frac{n_4}{n_1} = \frac{N_A}{N_B} \times \frac{N_B}{N_C} \times \frac{N_D}{N_E} \tag{4}$$

merator of the expression on the right are those for driving gears, while those appearing in the denominator are the numbers of teeth on the driven gears. Thus, the general expression for the magnitude of the train value for an ordinary gear train is Eq. (5).

$$E = \frac{\text{product of numbers of teeth on driving gears}}{\text{product of numbers of teeth on driven gears}} \tag{5}$$

Because any two gears in mesh must have the same diametral pitch, expressions for the train value can also be written in terms of pitch diameters instead of numbers of teeth.

In the expression for the train value of the train of Fig. 2 the number of teeth on gear B cancels out. Gear B is an idler; it is both a driver and a driven gear. Its size has no effect on the train value's magnitude. It does affect, however, the sign of the train value. Idlers also are useful where a relatively large center distance must be spanned.

Transmissions. If a machine such as a machine tool or motor vehicle must be operated at any one of several output speeds, a multiple-speed gearbox, or transmission, may be used as a component part. The speed of the output shaft of a transmission can be varied by sliding gears in and out of contact or by connecting gears in continual mesh to shafts by means of clutches. A compact nine-speed transmission utilizing 10 gears and three shafts is shown in **Fig. 4**. Gears A, B, and C slide as a unit on the constant-speed shaft, providing three ways in which the constant speed and the intermediate shafts can be connected (A and D, or B and E, or C and G). Similarly gears H, I, and J slide as a unit on the variable-speed shaft, providing three ways in which it can be connected to the intermediate shaft (D and H, or $E$and I, or F and J). Consequently nine possible combinations are available. For example, the lowest output speed is attained when A and D mesh and E and I mesh. The train value is then as in Eq. (6).

$$E = \frac{20 \times 37}{47 \times 57} = 0.276 \tag{6}$$

The highest speed occurs when C and G mesh and D and H mesh. The train value for this combination is shown in Eq. (7).

$$E = \frac{41 \times 47}{26 \times 47} = 1.577 \tag{7}$$

See AUTOMOTIVE TRANSMISSION.

Epicyclic train. An epicyclic train is named for the path described by any point on the pitch circle of a planet gear as it rolls on a sun gear. The term epicyclic refers only to the motion and is not related to the gear tooth form, which may be involute or any other form satisfying the law of gearing. An epicyclic train can be used to obtain a considerably greater velocity ratio than would be possible with an ordinary gear train of the same size.

In the compound and reverted epicyclic gear train of **Fig. 5**, gear A is fixed to the frame, but not to shaft 1. The planet carrier, arm or link E, is fastened to shaft 1 and carries planet gears B and C, which are fastened to a common shaft. Planet B rolls on sun gear A, while planet C rolls on sun gear D. Gear A is one of the input members; it has a velocity of 0 rpm. The other input could be either the carrier E or gear D. In the example to follow it will be assumed that the arm carrier E is the input member.

To determine the train value of an epicyclic train with a fixed gear, the following special but very convenient technique can be used. The net motion is divided into two parts. In the first, the train is locked with no relative motion between any of its components. For this step gear A is released from the frame and locked to gear B. The locked train is rotated one revolution in an arbitrary direction. Counterclockwise rotation is assumed positive, clockwise negative. In the second step, the assembly is viewed as an ordinary gear train with the planet carrier regarded as fixed. While the fixed gear is returned to its original position by one revolution in the sense opposite that used in the first step, the consequent rotation

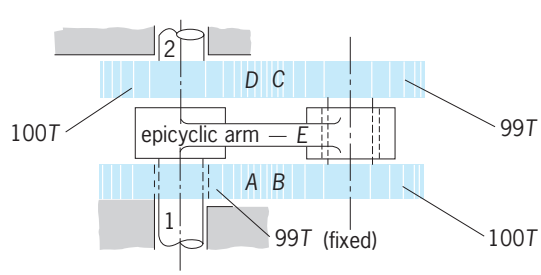

Fig. 5. Reverted epicyclic gear train.

Algebraic sum of revolutions for a reverted epicyclic gear train

Category	Gear A	Gear B	Gear C	Gear D	Arm E
Locked train	+1	+1	+1	+1	+1
Locked arm	−1	$+\frac{99}{100}$	$+\frac{99}{100}$	$+\frac{99}{100}\times\frac{99}{100}$	0
Net motion		$\frac{199}{100}$	$\frac{199}{100}$	$\frac{199}{10{,}000}$	+1

of the other gears is noted. For any one gear, the algebraic sum of the revolutions during the two parts is the net motion of that gear for one revolution of the planet carrier. To illustrate, consider the **table** for the train of Fig. 5, with the number of teeth as marked for each gear. The table shows that, for each revolution of the carrier, gear *D* will turn 199/10,000 of a revolution in the same direction.

For epicyclic trains with no fixed gears, a modification of the two-step method is needed. In the first step, the locked train is given the same rotation as the known motion of the planet carrier. In the second, with the carrier stationary, the gear whose velocity is known is rotated sufficient turns in the proper direction to make its net motion equal to the known value.

Probably the most common epicyclic train with no fixed gear is the automobile differential. When the automobile is moving along a straight path, there is no relative motion between the bevel differential gears fastened individually to the right and left axles. But as the car makes a turn, the differential gears move relative to one another in the manner of an epicyclic bevel gear train with no fixed gears. *See* DIFFERENTIAL; GEAR; PLANETARY GEAR TRAIN. John R. Zimmerman; Donald L. Anglin

Bibliography. H. H. Mabie and C. F. Reinholtz, *Mechanisms and Dynamics of Machinery,* 4th ed., 1987; G. H. Martin, *Kinematics and Dynamics of Machines*, 2d ed., 1982; J. F. Shannon, *Marine Gearing*, 1981; J. E. Shigley and J. J. Uiker, *Theory of Machines and Mechanisms*, 2d ed., 1994.

Gecko

The name for about 300 species of reptiles that form the family Gekkonidae in the order Squamata. They are small lizards, primarily arboreal and nocturnal, which occur in the warm regions of the world. The body is flattened and never exceeds 14 in. (35 cm) in length. Most species have five digits, while some have only four; the toes often have adhesive pads to assist the animal when climbing on smooth surfaces.

Geckos feed on small animals, especially insects, and all species have a long sensitive tongue to aid in capturing their prey. In some species, the tail is used as a fat-storage organ. These animals move with agility and speed, and if caught by the tail, they can break it off by a spontaneous muscular contraction and rapidly regenerate a new one. Most species of gecko are oviparous; although they lay their eggs in protected areas, no particular care is given to them. *See* REGENERATION (BIOLOGY).

The largest, most aggressive species is the orange-spotted Tokay (*Gekko gecko*), which measures 14 in. (35 cm) and is indigenous to Southeast Asia. Another species is the flying gecko (*Ptychozoon homalocephalum*), which is about 8 in. (20 cm) long. It is essentially arboreal and is well adapted to leaping and gliding because it has folds of skin on either side of the body that can be opened out to form a planing surface. The banded gecko (*Coelonyx variegatus*) of the southwestern United States is about 3–4 in. (7.5–10 cm) long (see **illus.**). It is one of the few species

Banded gecko (*Coelonyx variegatus*).

with movable eyelids, and with claws instead of pads on the toes. *See* SQUAMATA. Charles B. Curtin

Geiger-Müller counter

A detector of ionizing radiation. When a fast-moving charged particle traverses a Geiger-Müller counter, an electrical impulse is produced. These impulses can readily be counted by electronic circuits. Geiger-Müller (GM) counters, usually referred to simply as Geiger counters, are widely used to indicate the presence and intensity of nuclear radiations.

Construction and operation. A Geiger counter consists of a gas between two electrodes (see **illus.**). One electrode, usually cylindrical and hollow, is the cathode. The other electrode, a fine wire stretched along the axis of the cylinder, is the anode. A potential of about 1000 V is placed on the wire. When an atom of the gas between the two electrodes is ionized by collision with a charged particle passing through the gas, the electron produced in the collision is drawn toward the central wire. The electron

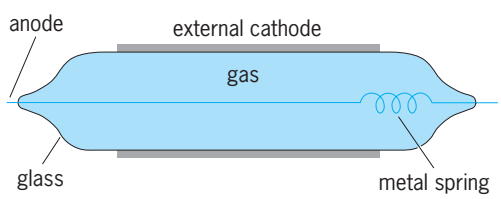

Cylindrical external-cathode Geiger-Müller counter, with thin soda glass and central wire of 0.003-in.-diameter (0.012-mm) tungsten. Metal spring keeps central wire taut.

then collides with the atoms of the gas. Near the central wire the electric field is very intense, and the electron may acquire enough energy between two collisions to allow it to ionize another atom. A second electron is then set free, and by successive collisions, an avalanche of electrons is produced which is then collected as charge on the central wire. This charge produces an electrical impulse which in typical cases may be 50 V.

A Geiger counter is a nonproportional particle detector; that is, the pulse produced is independent of the nature and energy of the particle that liberated the electrons within the gas. The operation of a Geiger counter depends critically on the voltage placed on the central wire. If the voltage is too low, the electron avalanche never builds up, and the counter operates only as an ionization chamber (a device which gives the total ionization produced in the gas) or a proportional counter, in which the output pulses are much smaller. If the voltage is too high, additional electrons produced by the positive ions start additional discharges, and the counter discharges continuously. Starting at the threshold for Geiger counting and continuing up to the region of multiple pulsing is the plateau region of the Geiger counter. The threshold is rather sharp, and occurs at an anode potential of 800–1000 V in typical counters. The plateau of a good Geiger counter extends 100–500 V above threshold. *See* IONIZATION CHAMBER; PARTICLE DETECTOR.

The pulses coming from a Geiger counter are amplified electronically and are then counted by an electromechanical register or as clicks in a loudspeaker. If the counting rate is so high that the mechanical register cannot follow the pulses fast enough, the pulses are fed into a scaling circuit, which divides the number of input pulses by a known scaling factor before they are fed to the register. The pulses may also be sent into a counting-rate meter, which contains a condenser that charges a certain amount each time a pulse reaches it. The accumulated voltage on the condenser indicates the counting rate. This voltage can be measured if a resistance is placed across the condenser to discharge it slowly.

Coincidence counting. Geiger counters are often used in coincidence. When a single particle passes through two or more Geiger counters, the pulses from each counter are practically time-coincident. The pulse from each counter is then sent to a coincidence circuit which indicates pulses coincident in time. Coincidence counting is used to eliminate much background count due to local radioactivity. Arrays of Geiger counters in coincidence also select particles going in a given direction. This technique is used, for example, to measure the angular distribution of cosmic rays.

Types and dimensions. Construction of Geiger counters requires careful control of both cleanliness of the electrode surfaces and purity of the gas. An inert gas, such as argon, is often used, mixed with a small amount of organic vapor to a total pressure of about 0.14 atm (2.1 lb/in.^{2} or 14 kilopascals). The organic vapor quenches the discharge and prevents multiple pulsing. However, some of the vapor is used during each pulse and such tubes thus have a limited life, usually 10^8–10^9 counts. Geiger tubes quenched by halogen gas are also available, and these have much longer life. The central wire of a Geiger counter is usually made of tungsten, 0.003–0.005 in. (75–125 micrometers) in diameter. The outer electrode may be copper, brass, or aluminum. For certain applications, external cathode counters, such as the one shown in the illustration, are used. The cathode is formed by spraying a conducting coating on the thin glass envelope surrounding the central wire. The electrical conductivity of soda glass, while small, is sufficient to permit the current to pass through the glass. Such counters have a long recovery time, but are useful in cosmic-ray research and are much cheaper than metallic counters. They have the additional advantage of not being destroyed if an overvoltage is accidentally applied.

Geiger counters may be made in various sizes, depending on the use for which they are intended. Counters 0.5 in. (13 mm) in diameter and 3 in. (75 mm) long, with thin aluminum walls, are used for detection of beta particles and cosmic rays. Counters of diameter 2 or 3 in. (50 or 75 mm) have been made, with lengths up to 3 ft (0.9 mm). However, if a large area is to be covered by GM counters, it is customary to use several connected electronically in parallel. Point Geiger counters have a somewhat different geometry than cylindrical counters. A sharp point is used as the anode, and a flat plane or sphere is the cathode.

Applications. Geiger counters are widely used in industry, medicine, mineral exploration, and scientific research.

Industry. A typical industrial use is in the measurement of thickness of sheet material. A source of beta rays is placed on one side of the material and a Geiger counter on the other. Some of the beta rays are absorbed in the material, and the counter responds to changes in absorption caused by changes in the thickness. Continuous accurate measurement of the thickness can be made even if the sheet material is moving at high speed. Another industrial application is in the detection of traces of radioactive elements in research on the wear of bearings. Such counters are used in connection with x-ray machines to search for flaws in large castings.

Medicine. Radioactive sources used in medicine are often detected with Geiger counters, which can locate the position or distribution of the radioactive material after it has been administered. In medical applications it is important to determine the total radiation dosage and to monitor the radiation; Geiger counters are used for this purpose. They have also been valuable in locating expensive and potentially dangerous sources of radioactivity.

Geology. Geologists use Geiger counters in mineral exploration. Minerals containing uranium are radioactive, and can be detected by Geiger counters. Simple lightweight portable Geiger detectors are available for prospectors. *See* GEOPHYSICAL EXPLORATION; RADIOACTIVE MINERALS.

Research. Chemists, physicists, and biologists use Geiger counters in various types of scientific research. Radioactive tracers used in chemical research are followed through complicated chemical reactions by Geiger counters. The counters give information on the quantity of material present and also on the lifetimes of radioactive products produced in the transmutation of elements. Geiger counters can be designed to employ a radioactive gas instead of the gas normally used. In such counters, the wall thickness is not important because the ionizing particle originates within the counter. Applications of Geiger counters in high-energy nuclear physics are confined to low-counting-rate experiments, such as in cosmic rays, because of the long dead time of Geiger counters. The counters are rarely used near particle accelerators, for they jam in a large flux of ionizing particles. William B. Fretter

Bibliography. C. F. Delaney and E. C. Finch, *Radiation Detectors: Physical Principles and Applications*, 1992; K. Klein Knecht, *Detectors for Particle Radiation*, 1988; G. F. Knoll, *Radiation Detection and Measurement*, 3d ed., 1999; W. J. Price, *Nuclear Radiation Detection*, 2d ed., 1964.

Gel

A continuous solid network enveloped in a continuous liquid phase; the solid phase typically occupies less than 10 vol % of the gel. Gels can be classified in terms of the network structure. The network may consist of agglomerated particles (formed, for example, by destabilization of a colloidal suspension; **Fig. 1***a*); a "house of cards" consisting of plates (as in a clay) or fibers (Fig. 1*b*); polymers joined by small crystalline regions (Fig. 1*c* and **Fig. 2**); or polymers linked by covalent bonds (Fig. 1*d*).

In a gel the liquid phase does not consist of isolated pockets, but is continuous. Consequently, salts can diffuse into the gel almost as fast as they disperse in a dish of free liquid. Thus, the gel seems to resemble a saturated household sponge, but it is distinguished by its colloidal size scale: the dimensions of the open spaces and of the solid objects constituting the network are smaller (usually much smaller) than a micrometer. This means that the interface joining the solid and liquid phases has an area on the order of 1000 m^2 per gram of solid. As a result, the properties of a gel are controlled by interfacial and short-range forces, such as van der Waals, electrostatic, and hydrogen-bonding. Factors that influence these forces, such as introduction of salts or another solvent, application of an electric field, or changes in pH or temperature, affect the interaction between the solid and liquid phases. Variations in these parameters can induce huge changes in volume as the gel imbibes or expels liquid, and this phenomenon is exploited to make mechanical actuators or hosts for controlled release of drugs from gels. For example, a polyacrylamide gel (a polymer linked by covalent bonds) shrinks dramatically when it is transferred from a dish of water (a good solvent) to a dish of acetone (a poor solvent), because the polymer chains tend to favor contact with one another rather than with acetone, so the network collapses onto itself. Conversely, the reason that water cannot be gently squeezed out of such a gel (as from a sponge) is that the network has a strong affinity for the liquid, and virtually all of the molecules of the liquid are close enough to the solid-liquid interface to be influenced by those attractive forces. *See* HYDROGEN BOND; INTERMOLECULAR FORCES.

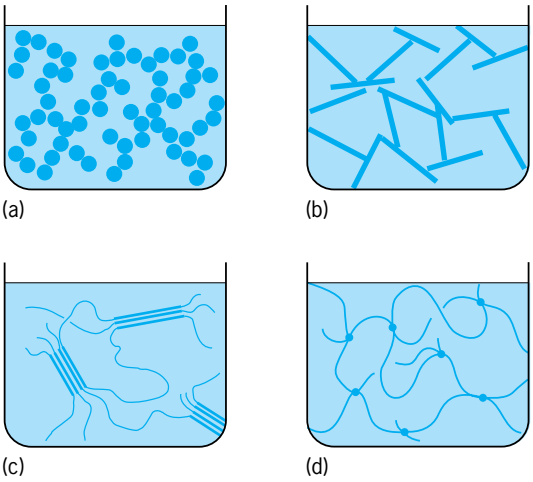

Fig. 1. Gel structures. (*a*) Agglomerated particles. (*b*) Framework of fibers or plates. (*c*) Polymers linked by crystalline junctions. (*d*) Polymers linked by covalent bonds. (*After M. Djabourov, Architecture of gelatin gels, Contemp. Phys., 29(3):273–297, 1988*)

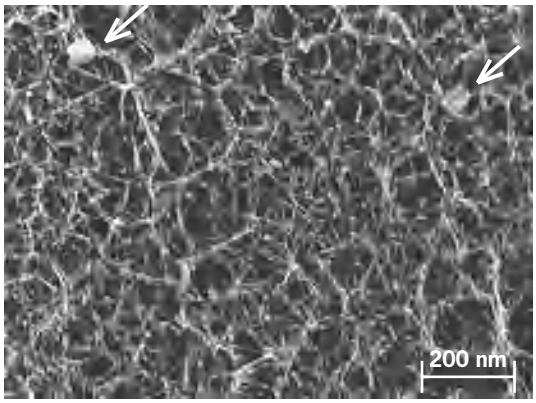

Fig. 2. Electron micrograph of a gelatin gel, where the solid phase occupies ~2% of the total volume. The arrows indicate latex particles with diameter of 90 nm that were included in the solution before gelation to serve as markers. The water was removed from the gel by vitrification (rapid quenching causing freezing without crystallization of the water) and sublimation. (*From M. Djabourov, Architecture of gelatin gels, Contemp. Phys., 29(3):273–297, 1988*)

Gelation. The most striking feature of a gel, which results from the presence of a continuous solid network, is elasticity: if the surface of a gel is displaced slightly, it springs back to its original position. If the displacement is too large, gels, except those with polymers linked by covalent bonds, may suffer some permanent plastic deformation, because the network is weak. The process of gelation, which transforms a liquid into an elastic gel, occurs with no

change in color or transparency and no evolution of heat. It may begin with a change in pH that removes repulsive forces between the particles in a colloidal suspension, or a decrease in temperature that favors crystallization of a solution of polymers or the initiation of a chemical reaction that creates or links polymers. *See* PH.

When the particles in a colloid are attracted to one another, they aggregate into a cluster. If the particles form weak bonds when they collide, each particle will tend to detach from the cluster and reattach at another spot, seeking a place where it can bond to several other particles at once. This process leads to dense crystalloids or precipitates that settle out of solution. However, if each collision between particles results in an irreversible bond, the random nature of the collisions leads to a loose, disorganized aggregate with a great deal of occluded liquid. Collisions between clusters produce even less dense structures, until finally the whole volume of the original colloid is crowded with wispy aggregates jostling against one another, and gradually linking into a single giant cluster that becomes the network of the gel. In the early stage of this process, all of the bonds are formed within clusters that are free to diffuse about, so the colloid still flows, and is properly called a liquid. (This stage is accurately described by the theory of cluster-cluster aggregation.) As large clusters begin to link together, the viscosity of this liquid rises drastically. The gel point is reached when enough bonds have formed between clusters that a continuous network extends throughout the volume of the original colloid, arresting flow and imparting elasticity to the system. (The evolution of structure and properties in the vicinity of the gel point is properly described by percolation theory.) *See* AMORPHOUS SOLID; COLLOID; PRECIPITATION (CHEMISTRY).

Aging and drying. In many gelling systems the bonds are strong but not permanent, so that the structure of the gel changes slowly over time. This process is called aging. If the solid phase is soluble in the liquid, coarsening (termed Ostwald ripening) will occur, and sometimes an amorphous gel network will crystallize. In polymeric systems, flexible branches of the network may come into contact and form new bonds, leading to a gradual contraction of the network (known as syneresis).

In many cases, both the liquid and solid phases of a gel are of practical importance (as in timed release of drugs, or a gelatin dessert). More often (as in preparation of catalyst supports, chromatographic columns, or desiccants) it is only the solid network that is of use. If the liquid is allowed to evaporate from a gel, large capillary tension develops in the liquid, and the suction causes shrinkage of the network. The porous network remaining after evaporation of the liquid is called a xerogel, and it usually retains high porosity (>25%) and enormous surface area (150–900 m^2/g), along with very small pore size (1–10 nm). To maximize the porosity of the dried product, the gel can be heated to a temperature and pressure greater than the critical point of the liquid phase, where capillary pressures do not exist. The fluid can then be removed with little or no shrinkage occurring. The resulting solid is called an aerogel, and the process is called supercritical drying. Such materials are useful as Cerenkov radiation detectors, thermal insulation, and catalytic substrates. *See* DESICCANT; DRUG DELIVERY SYSTEMS; DRYING; GELATIN; HETEROGENEOUS CATALYSIS.

Sol-gel processing. Sol-gel processing comprises a variety of techniques for preparing inorganic materials by starting with a sol, then gelling, drying, and (usually) firing. Many inorganic gels can be made from solutions of salts or metallorganic compounds, and this offers several advantages in ceramics processing: the reactants are readily purified; the components can be intimately mixed in the solution or sol stage; the sols can be applied as coatings, drawn into fibers, emulsified or spray-dried to make particles, or molded and gelled into shapes; xerogels can be sintered into dense solids at relatively low temperatures, because of their small pore size. Moreover, hybrid materials can be made by combining organic and inorganic components in the gel. Many hybrids have such compliant networks that they collapse completely during drying, leaving a dense solid rather than a porous xerogel; therefore, they can be used as protective coatings (on plastic eyeglass lenses, for example) without heat treatment. Hybrid gels show great promise for active and integrated optics, because optically active organic molecules retain their activity while encapsulated in the gel matrix.

The most important application of gels continues to be for the preparation of thin films (on the order of a micrometer in thickness). Dense films, which must be sintered, typically at 500–700°C (932–1292°F), are used for antireflective or infrared reflective coatings, and protection against corrosion or abrasion. Unfired porous films, usually with encapsulated organic molecules, are being studied for use as sensors. Xerogel membranes are of interest for filtration and gas separation.

Catalytic substrates and chromatographic media have been made by sol-gel processing for decades, but improvements in control of microstructure should provide substantially improved performance. By growing the gel in the presence of surfactants, ordered mesoporous microstructures (consisting of regular cylindrical tubes, uniformly spaced layers, and even three-dimensional networks) have been prepared. These structures can be compared to zeolites, which have uniform pores on the order of 0.3–1 nm and which allow reactions to be performed exclusively on the molecules that fit into the pores; the mesoporous gels have pores an order of magnitude larger than zeolites, so size- and shape-selective chemistry can be extended to much larger molecules. Control over the microstructure of chromatographic media has been achieved by inducing phase separation in a solution of polymers and metalorganic compounds. After the inorganic gel network is formed, the polymer is removed by washing or thermal decomposition, leaving a uniform interconnected network of relatively large pores (up to

the micrometer scale) within the micro- or mesoporous gel. The large pores provide rapid transport into the interior of the gel body, while the micropores perform the chromatographic separation.

The preparation of aerogels by supercritical drying was first reported in 1931, and there has been recurrent interest in their use as catalysts and as thermal insulation, but the cost of fabrication has so far prevented widespread use. It has been demonstrated that materials with the same physical structure as aerogels can be made without the expensive and hazardous supercritical processing. For example, if a silica gel is treated with chlorosilanes, it shrinks during conventional drying but then springs back to its original dimensions; alternatively, if a gel is aged in a solution containing additional silicate monomers, it becomes so rigid that it does not shrink during drying, so a low-density xerogel is obtained. *See* RHEOLOGY; SINTERING; SOL-GEL PROCESS. George W. Scherer

Bibliography. C. J. Brinker and G. W. Scherer, *Sol-Gel Science*, 1990; D. DeRossi et al. (eds.), *Polymer Gels: Fundamentals and Biomedical Applications*, Plenum, 1991; P. J. Flory, *Principles of Polymer Chemistry*, 1953; K. Nakanishi, Pore structure control of silica gels based on phase separation, *J. Porous Mat.*, 4:67–112, 1997; Proceedings of the 5th International Symposium on Aerogels, *J. Non-Crystalline Sol.*, vol. 225, 1998; Royal Society of Chemistry, *Faraday Discussions*, no. 101, *Gels*, 1995; Special issues of *J. Sol-Gel Sci. Technol.*, Organic/Inorganic Hybrids, 5(2), 1995, 7(3), 1996, and Non-linear Optical Materials, 9(2), 1997, and Optical Solids, 11(3), 1998; T. Tanaka, Gels, *Sci. Amer.*, 244:124–138, 1981; R. Zallen, *The Physics of Amrophous Solids*, 1983.

Gel permeation chromatography

A separation technique involving the transport of a liquid mobile phase through a column containing the separation medium, a porous material. Gel permeation chromatography (GPC), also called size exclusion chromatography and gel filtration, affords a rapid method for the separation of oligomeric and polymeric species. The separation is based on differences in molecular size in solution. It is of particular importance for research in biological systems and is the method of choice for determining molecular weight distribution of synthetic polymers.

Procedure. The separation medium is a porous solid, such as glass or silica, or a cross-linked gel which contains pores of appropriate dimensions to effect the separation desired. The liquid mobile phase is usually water or a buffer for biological separations, and an organic solvent that is appropriate for the sample and is compatible with the column packing for synthetic polymer characterization. Solvent flow may be driven by gravity, or by a high-pressure pump to achieve the desired flow rate through the column. The sample to be separated is introduced at the head of the column (**Fig. 1**). As it progresses through the column, small molecules can enter all pores larger than the molecule, while larger molecules can fit into a smaller number of pores, again only those larger than the molecule. Thus, the larger the molecule, the smaller is the amount of pore volume available into which it can enter. The sample emerges from the column in the inverse order of molecular size; that is, the largest molecules emerge first followed by progressively smaller molecules. In order to determine the amount of sample emerging, a concentration detector is located at the end of the column. Additionally, detectors may be used to continuously determine the molecular weight of species eluting from the column. The volume of solvent flow is also monitored to provide a means of characterizing the molecular size of the eluting species.

As the sample emerges from the column, a concentration detector signal increases above the baseline, passes through a maximum, and returns to the baseline (**Fig. 2**). This signal provides a relative concentration of emerging species and is recorded as a function of elution volume. Molecular-weight detectors based on light scattering or viscometry produce a signal, independent of molecular weight, that increases with increasing molecular weight for a given sample concentration. The molecular-weight chromatograms therefore are skewed with respect to the concentration detector signal. Typically, constant-volume pumps are used, and the time axis is transformed into a volume axis. Devices such as siphons that empty after a known volume of solvent has collected, or electronic circuitry that measures the transit time of a thermal pulse between two points in the

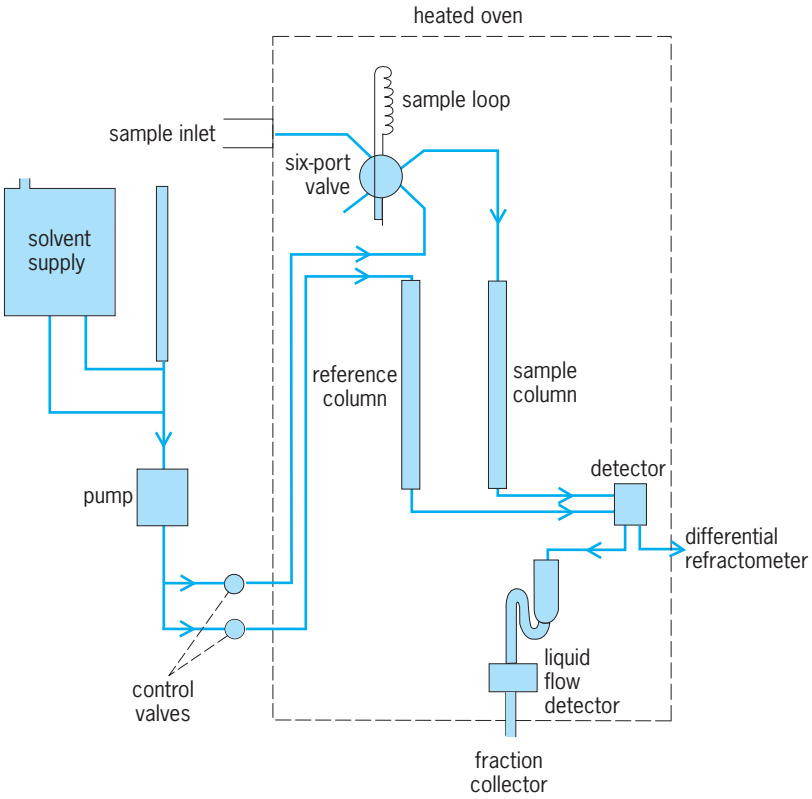

Fig. 1. Gel permeation chromatograph.

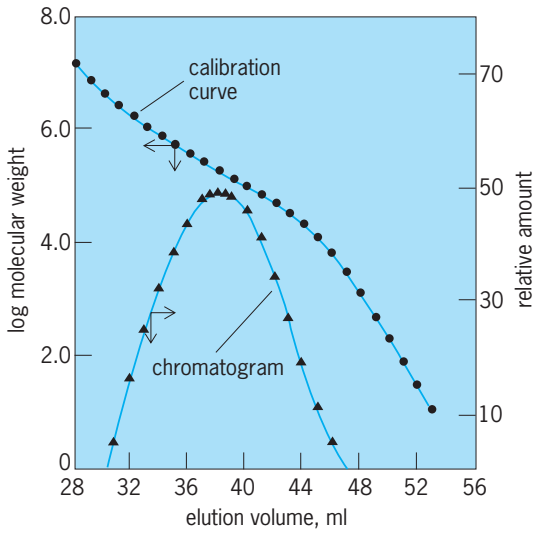

Fig. 2. Typical gel permeation chromatogram with a calibration curve.

flowing solvent, have been used to provide a measurement of the volumetric flow rate.

Calibration procedure. In the absence of direct molecular-weight detection, a calibration curve (Fig. 2) is required to convert elution volume into molecular weight. This is normally prepared by injecting standard samples of a polymer with a narrow molecular-weight distribution, so that the molecular weight at the maximum in the chromatogram may be unambiguously assigned. It must be assumed that elution volume depends only on molecular weight; that is, the only operative separation mechanism is size exclusion. The molecular weights of the standards are determined by primary molecular-weight methods, such as light scattering of dilute solutions or membrane osmometry, which determine the weight-average and number-average molecular weights, respectively. Secondary molecular-weight methods such as dilute-solution viscometry and vapor-pressure osmometry are also commonly used.

The procedure described above requires a set of characterized molecular-weight standards for each polymer type being characterized. Two approaches have been pursued to circumvent this problem. One approach involves an extensive search for a "universal" calibration curve, that is, for one on which a wide variety of molecular structures could be represented by a single line. It has been established that a plot of the logarithm of the intrinsic viscosity times the molecular weight versus elution volume produces such a curve. Many high-molecular-weight materials can be characterized from this curve, which is established by using only one type of material. This quantity, intrinsic viscosity times molecular weight, is proportional to the hydrodynamic volume. It appears, therefore, that the separation is based mainly on molecular size in solution. There are other contributing factors such as diffusion and nonsteric effects that affect separation, but these seem to be relatively unimportant compared to the size factor.

The other approach has been to develop detectors that continuously measure the molecular weight or some parameter that is proportional to molecular weight. The former involves the use of laser light scattering to measure the weight-average molecular weight. The latter is obtained from a continuously recording viscometer, measuring the intrinsic viscosity, which is proportional to the viscosity-average molecular weight. If the polymer eluting from the column at any time is essentially monodisperse, these two averages will be identical. The intrinsic viscosity may be converted into molecular weight by use of the Mark-Houwink relationship of intrinsic viscosity to molecular weight. Both approaches require the concurrent use of a concentration detector to measure the concentration of species eluting at a given time. A typical chromatogram obtained by using both light scattering and viscometric detectors in addition to a concentration detector has three curves (**Fig. 3**).

Light-scattering molecular-weight detectors allow calibration curves to be determined for a given column set by using either a set of narrow molecular-weight distribution fractions or a broad molecular-weight distribution polymer of undetermined molecular-weight distribution. Viscometric detectors allow a calibration curve and the universal calibration curve to be determined from either a set of narrow molecular-weight distribution fractions or a broad molecular-weight distribution polymer if the Mark-Houwink relationship is known. Similarly, the universal calibration curve may be determined by using a broad molecular-weight distribution polymer or narrow molecular-weight distribution fractions of a polymer of unknown chemical composition; or the Mark-Houwink relationship may be used if light scattering and viscometric detectors are used simultaneously. *See* VISCOSITY.

Once a calibration curve relating elution volume to molecular weight or the universal calibration curve is available, it is comparatively simple to convert the raw data from the chromatogram, that is, relative concentration versus elution volume, to a molecular-weight distribution. If molecular-weight detectors

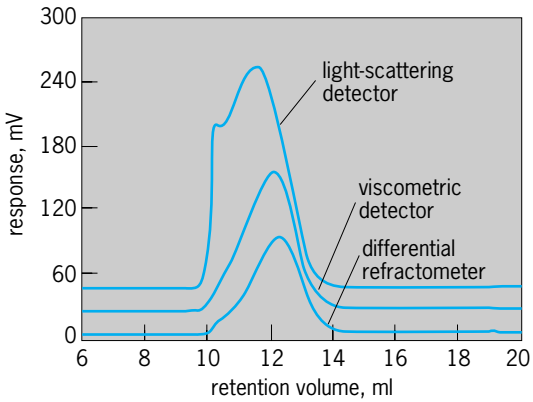

Fig. 3. Typical triple chromatogram obtained by using both light scattering and viscometric detectors in addition to a concentration detector.

are used, a direct relationship of molecular weight versus relative concentration is obtained. Such calculations are performed automatically by the data acquisition computer. The usual parameters calculated are the cumulative and differential molecular-weight distributions. The various molecular-weight averages may be calculated from these distributions.

Other parameters such as inhomogeneity factors and descriptions of the breadth of the distribution may be computed if desired. Thus, the formerly laborious process of determining precise molecular-weight distributions becomes very simple with the use of gel permeation chromatography. *See* MOLECULAR WEIGHT.

Column packings. A wide variety of porous materials have been used for column packing. For aqueous applications, cross-linked dextrans and polyacrylamide have found widespread use. In addition, cross-linked hydrophilic gels and inorganic materials with appropriate surface treatments having greater mechanical strength have been used. For organic solvents, cross-linked polymeric gels such as polystyrene and various other cross-linked copolymers are frequently employed. Rigid porous materials such as glass and silica that have high strength and extremely high solvent resistance have also been shown to produce effective separations. The trend has been to produce columns having column packings with particle diameters of a few micrometers. These columns tend to be fairly short to avoid large pressure drops across them, which would crush or distort the packing materials. Faster analyses are achieved, and the beneficial side effect is lower solvent use.

Detectors. In analytical applications, the concentration of solute is extremely small in order to maintain high-resolution separations and to avoid sample viscosity effects. Sample elution concentration is continuously monitored by using differential refractometry, infrared and ultraviolet/visible spectroscopy, colorimetric detectors, and various combinations of these. Multiple detectors are required if the samples are not chemically homogeneous. Some of these detectors are extremely sensitive to temperature and flow rate, so variations must be kept to a minimum. *See* COLORIMETRY; REFRACTOMETRIC ANALYSIS; SPECTROSCOPY.

Associated equipment. Pumping systems range from columns using gravity flow to high-pressure precision pumping systems that can be regulated to give a constant flow rate with very good accuracy. Soft gel column packings will not tolerate high pressures, and slower analyses result. Samples may be injected by hypodermic syringe or from calibrated volume injection loops attached to multiport valves. Columns may be metal or glass and have diameters from a few millimeters to several centimeters for preparative columns, and lengths from a few centimeters to several meters. In general, column efficiency increases directly with column length, while band spreading increases as the square root of column length. Thus, in principle, with a long enough column it should be possible to separate very difficult resolvable materials. The time for analysis, however, increases linearly with column length.

Flow rate. As the analysis time is inversely proportional to flow rate, it is advantageous to use the highest permissible flow rate that will give the required resolution. The effect of flow rate has been extensively studied. Column performance may be defined in terms of the number of plates per unit length as measured with monodisperse (essentially homogeneous) substances. In general, it is found that efficiency decreases with increasing flow rate. The decrease, however, is not large, and in typical analytical columns of approximately 0.2 in. (5 mm) inside diameter where flow rates are of the order of 1–5 milliliters per minute, a decrease of only a few percent is noted. At higher flow rates, the decrease in efficiency becomes more pronounced. Flow-rate limitations are sometimes imposed by the pressure at which a given column set may be operated. High flow rates should be avoided for high-molecular-weight samples because of viscosity effects and the possibility of shear degradation of the sample.

Sample size. Normally, the sample size for analytical gel permeation chromatography is 1 mg or less. As the molecular weight increases, the sample size should be reduced.

Solvent and temperature. Selection of the solvent and temperature is normally based on two factors. The first is that the sample must be soluble and nonassociated in solution. The second is compatibility with the detector system. For the differential refractometer and light-scattering detectors, a large difference between the refractive index of the sample and solvent is desirable. For spectrophotometric detectors, the absence of solvent absorption in the desired wavelength range is required. For viscometric detectors, use of a thermodynamically good solvent will cause a higher signal at a given molecular weight. A thermodynamically good solvent is one in which the polymer coil is expanded and is reflected in a higher value of a certain constant in the Mark-Houwink relationship.

Unlike most chromatographic separations, temperature has very little effect on separability, as the separation is based on the hydrodynamic volume of the molecule. There are minor effects on this volume with changes in temperature. Normally, however, all the species will change in the same direction and approximately the same amount so there is little overall effect on the relative separation. There are minor changes in efficiency due principally to changes in viscosity. Lower viscosity of the solvent improves mass transfer and gives a faster approach to equilibrium. The changes due to viscosity are usually also minor; therefore, the temperature chosen is normally just above ambient so that adequate temperature control of the solvent can be achieved to permit steady baselines from the detectors used. For some synthetic polymers, for example, the polyolefins, of difficult solubility, it is necessary to operate at elevated temperatures, and usually antioxidants must be added to the solvent to prevent oxidative degradation of sample and solvent. Temperatures ranges

from below ambient to 150°C (300°F) have been reported.

Applications. Gel permeation chromatography has had widespread applications. For determining the molecular weight of synthetic polymers, at least 50 types of polymers have been characterized. These include alkyd resins, natural and synthetic rubbers, cellulose esters, polyolefins, polyamides, polyesters, polystyrenes, polyacrylates, uncured epoxy, urethane and phenolic resins, and a wide variety of oligomeric materials. Additionally, the ability to determine the molecular-weight distribution and changes in distribution has led to many applications in areas such as blending distributions, chain-length studies in semicrystalline polymers, interactions in solution, radiation studies, mechanical degradation studies, mechanisms of polymerization research, polymerization reactor control, and evaluation of the processing of polymers. In the field of natural and biological polymers, numerous systems have been separated and analyzed. Among these are acid phosphatases, adrenalin, albumin, amino acids and their derivatives, enzymes, blood group antibodies, collagen and related compounds, peptides, and proteins.

Additionally, gel permeation chromatography is capable of making separations of low-molecular-weight compounds. This is particularly important when both low- and high-molecular-weight species are present in the same sample. *See* CHROMATOGRAPHY; POLYMER. Anthony R. Cooper

Bibliography. S. Mori and H. G. Barth, *Size Exclusion Chromatography*, 1999; B. H. Stuart, *Polymer Analysis*, 2002; C.-S. Wu (ed.), *Handbook of Size Exclusion Chromatography and Related Techniques*, 2d ed., 2003.

Gelatin

A protein extracted after partial hydrolysis of collagenous raw material from the skin, white connective tissue, and bone of animals. It is a linear polymer of amino acids, most often with repeating glycine-proline-proline and glycine-proline-hydroxyproline sequences in the polypeptide linkages. Gelatin gives positive test for hydroxyproline, and yields amino acid components upon complete hydrolysis. *See* AMINO ACIDS; COLLAGEN; PROTEIN.

Characteristics. The unique characteristics of gelatin are reversible sol-to-gel formation, amphoteric properties, swelling in cold water, film-forming properties, viscosity-modifying properties, and protective colloid properties. The last is exemplified by a very low Zsigmondy gold number. Gelatin solutions are levorotatory; the specific optical rotation, which greatly increases at lower temperatures, is $-137°$ at 589 nanometers and 45°C (113°F). *See* OPTICAL ACTIVITY.

Most physical properties of gelatin are measured on aqueous solution, in either the sol or gel state. The physical rigidity or gel strength of gelatin is measured at 6.67% concentration with a Bloom gelometer and is expressed in grams. Gelatin from first extracts yields 250–325 Bloom gelatin, whereas the last extract might yield 50–100 Bloom gelatin. Viscosity of 6.67% gelatin solution at 140°F (60°C) can range from 25 to 65 millipoises, with most commercial gelatin being in the 40–50 millipoise range. The pH is usually adjusted to between 5.2 and 6.5, depending on the end-use specifications. Gelatin contains molecular species from 15,000 to 400,000, with the average molecular weight between 50,000 and 70,000 for 200–300 Bloom gelatin.

Chemical composition. Gelatin contains 26.4–30.5% glycine, 14.8–18% proline, 13.3–14.5% hydroxyproline, 11.1–11.7% glutamic acid, 8.6–11.3% alanine, and in decreasing order arginine, aspartic acid, lysine, serine, leucine, valine, phenylalanine, threonine, isoleucine, hydroxylysine, histidine, methionine, and tyrosine.

Absence of only two essential amino acids—tryptophan and methionine—makes gelatin a good dietary food supplement. Most of the active groups in gelatin molecules are either terminal or part of side groups. An α chain with a molecular weight of about 96,500 behaves in a solution like a random-coil polymer, whereas the gel form may contain as much as 70% helical conformation. With the same gelatin, the rigidity of the gel is approximately proportional to the square of the gelatin concentration. There are two types of α chain, with slight chemical differences, designated α_1 and α_2. The original collagen triple helix contains two α_1 and one α_2 chain. The dimer made up from either $\alpha_1 + \alpha_2$ or $\alpha_1 + \alpha_1$ in gelatin is called β, and a trimer with a molecular weight of 298,500 is called γ. Homogeneous α-chain gelatin has been prepared by pretreating collagen with pronane in the presence of 0.4 *M* calcium chloride ($CaCl_2$). Enzyme-pretreated gelatin has the isoionic point of type A gelatin.

Manufacturing process. Collagen, the precursor of gelatin, is pretreated for 10–30 h with 1–5% mineral acid for type A gelatin production, or 35–90 days with a lime slurry for type B gelatin production. Type A gelatin exhibits the isoionic point at pH 7.0–9.5, whereas type B gelatin, due to deamination in the liming process, exhibits the isoionic point at 4.8–5.2. Under alkaline conditions, arginine is slowly hydrolyzed to ornithine by liberating urea. A short, 1–3-day alkaline process using 10–15% solution of sodium hydroxide (NaOH) and sodium sulfate (Na_2SO_4) has been patented in Great Britain. Type A gelatin is manufactured from frozen and fresh edible-grade pig skins or from bone ossein after demineralization. A type A gelatin made from hides has also become available commercially. Bovine hides and skins (known as splits) and trimmings of raw or salted dehaired hide pieces are used for type B gelatin. Some type B gelatin is also made from pig skins in Japan and in China. Most of type B gelatin, however, comes from bones. The bone pieces (0.5–2 cm) are first demineralized in 4–7% hydrochloric acid for 7–14 days before liming. After pretreatment with either acid (type A) or lime (type B), the materials are washed and then are subjected to four or five 4–8-h extractions at increased temperatures

of 131–212°F (55–100°C). The extracts, containing 3–7% gelatin, are filtered, concentrated in vacuum evaporators, chilled and extruded as noodles, and dried at 86–140°F (30–60°C). The dry gelatin is then ground and blended to specifications. The viscosity of consecutive extracts from bone ossein increases, whereas, the trend in type A gelatin from pig skins is reversed.

Uses. The principal uses of gelatin are in foods, pharmaceuticals, and photographic industries. Other uses of industrial gelatin are in the field of microencapsulation, health and cosmetics, and plastics.

Foods. The food industry is the largest user of gelatin. The formulations use exclusively either water or aqueous polyhydric alcohol solutions. The principal uses are in gelatin desserts, marshmallows, candies, meat products such as canned hams and luncheon meats, canned soups, bakery icings and pie fillings, sour cream, and ice cream. Flavors and vitamins can be encapsulated in gelatin to control release or to protect against aerial oxidation. Gelatin is also recommended to counteract fingernail defects. Most edible gelatin is of type A, but type B gelatin is also used.

Photography. Gelatin has been used in photography for over 100 years as a binder in light-sensitive and backing layers. After 1960 the photographic industry switched to so-called inert type B gelatin, where chemical sensitizers and restrainers have been removed and the gelatin no longer controls the silver halide crystal growth or its final light sensitivity. The function of gelatin remains as a protective colloid, as a bromine acceptor, and as a matrix which allows coating and processing of modern photographic formulations. Manufacture and quality control of inert photographic gelatin is very critical, since impurities of the order of parts per million can have a pronounced photographic effect. The largest-volume photographic products are x-ray films, phototypesetting paper, color paper, lithographic films, and microfilm. Derivatized gelatins such as phthalated gelatin also find wide use in the photographic market. Type A gelatin use in photography is limited to graphic arts products. *See* PHOTOGRAPHIC MATERIALS.

Pharmaceuticals. The largest use of gelatin in pharmaceuticals in the United States is in the manufacturing of hard and soft capsules for medications. The formulations are made with water or with aqueous polyhydric alcohols. The gelatin for soft capsules is a low-Bloom type A (6.0–6.3 oz or 170–180 g), type B (5.3–6.2 oz or 150–175 g), or a mixture of the two. The material is encapsulated between two plasticized gelatin sheets, and shaped and sealed in a rotary die. Hard capsules consist of two tightly fitting halves prefabricated from medium- to high-Bloom type A (8.8–9.9 oz or 250–280 g), type B (7.9–8.8 oz or 225–250 g), or the mixture of the two types. After the capsule-tampering scare in the 1980s, a different form of gelatin-coated medication referred to as gelcaps was introduced in addition to caplets. Tamperproof capsules have been introduced also. A special pyrogen-free gelatin grade is used as a blood plasma substitute. Glycerinated gelatin is used in suppositories. Enteric capsules are made with a cross-linked or coated gelatin. Absorbable gelatin sponge (Gelfoam) is gelatin foam partially insolubilized by cross-linking, and is used for arresting hemorrhage during surgery.

Technical applications. In the plastics industry gelatin can be used to control polymerization rate and particle size. Electroplating in gelatin solution produces a bright finish to electroplated surfaces of silver, gold, copper, brass, and zinc. Gelatin coating can protect copper, zinc, and aluminum surfaces against corrosion, and is used as a pickling inhibitor for stainless steel. Gelatin can be a source for nitrogen, as well as time-release coating for fertilizer components. In animal feed, gelatin is used to stabilize and pelletize vitamins and minerals otherwise susceptible to aerial oxidation or degradation. A different application of gelatin is its use for simulating body tissue consistency in the study of ballistics. A great deal of research has been done in this field with ordnance-type gelatin. *See* ANIMAL FEEDS; ELECTROPLATING OF METALS; FERTILIZER; POLYMERIZATION. Felix Viro

Bibliography. Gelatin Manufacturers Institute of America, Inc., *Gelatin*, 1983; L. Stryer, *Biochemistry*, 4th ed., 1995.

Gem

A mineral or other material that has sufficient beauty for use as personal adornment and has the durability to make this feasible. With the exception of a few materials of organic origin, such as pearl, amber, coral, and jet, and inorganic substances of variable composition, such as natural glass, gems are lovely varieties of minerals. *See* JET (GEMOLOGY).

Natural Gems

A mineral is defined as an inorganic substance with a characteristic chemical composition and usually a characteristic crystal structure. Each distinct mineral is called a species by the gemologist. Two stones that have the same essential composition and crystal structure but that differ in color are considered varieties of the same species. Thus ruby and sapphire are distinct varieties of the mineral species corundum, and emerald and aquamarine are varieties of beryl. Two or more minerals that have the same structure and are related chemically are called groups. For example, almandite, pyrope, rhodolite, grossularite, andradite, and some other minerals form the garnet group. *See* MINERAL; MINERALOGY.

Most gemstones are crystalline (that is, they have a regular atomic structure) and have characteristic properties, many of which have a direct bearing on the gem's beauty or durability.

Durability. Each mineral has a characteristic hardness (resistance to being scratched) and toughness (resistance to cleavage and fracture).

Cleavage. Cleavage is the tendency of a crystalline substance to separate parallel to a crystal face or a

possible crystal face. The term is also applied to such a separation visible in a crystalline substance.

Fracture. Fracture is a break or a tendency to break in a direction that bears no relationship to the atomic structure of the mineral.

Hardness. The hardness of a gem material determines its resistance to being scratched and thus, in conjunction with resistance to cleavage and fracture, its practicability for use as a jewelry stone.

The hardness of minerals is measured with reference to an empirical scale known as Mohs scale, which consists of 10 minerals, ascending from talc (hardness 1) to diamond (hardness 10). The scale was formulated by determining which of the various well-known minerals scratched others. Gypsum (2) scratches talc, calcite (3) scratches gypsum, and so on up the scale. However, the hardness of diamond is much greater in relation to corundum (9) than is corundum in relation to topaz (8). With few exceptions, the most important gemstones are those at the top of the hardness ladder; for example, diamond is 10, ruby and sapphire are 9, topaz and beryl (emerald and aquamarine) are 8, and the quartz family has a hardness of 7. *See* HARDNESS SCALES.

Beauty. Optical properties are particularly important to the beauty of the various gem materials. The important optical properties include color, dispersion (or "fire"), refractive index, and pleochroism. Gemstones usually are cherished for their color, brilliancy, fire, or one of the several optical phenomena, such as the play of color of a fine opal or the star effect in a sapphire.

Brilliancy. Brilliancy depends on the refractive index, transparency, polish, and proportions of a cut stone. Refractive index is a measure of a gem's ability to reduce the velocity of light that passes through it. It is defined either as the ratio of the velocity of light in air to the velocity of light within a gemstone, or as the ratio of the sine of the angle of incidence to the sine of the angle of refraction. Although diamond does not have the highest refractive index known, it has one of the highest to be found among transparent materials. This quality, combined with its unsurpassed hardness (which permits it to be given a magnificent polish) and great transparency, gives diamond the highest brilliancy potential of any mineral or manufactured material. Although synthetic rutile (titania, or titanium dioxide) has a higher refractive index, its lower transparency, relative softness (which makes a superior polish impossible), and other factors, such as its large birefringence, combine to reduce its brilliancy well below that of diamond. *See* REFRACTOMETRIC ANALYSIS.

In order for a gemstone to display its latent brilliancy to best advantage, it must be cut to the proper angles for its refractive index. For example, the optimum angle of the pavilion (bottom portion) of a brilliant-cut diamond is 41°, whereas the stones of lower refractive index should be cut to slightly greater angles (measured from the plane of the girdle).

Refractive indices. The refractive indices of the important gems vary from a low of approximately 1.45 for opal to a high among transparent stones of 2.6–2.9 for synthetic rutile. The value for diamond is 2.42. Refractive index is a rough measure of brilliancy; in other words, the brilliancy of two colorless stones of equal transparency and cutting quality is approximately in proportion to their refractive indices. The refractive indices of other important gems are as follows: ruby and sapphire, 1.76; spinel, 1.72; chrysoberyl, 1.74; topaz and tourmaline, 1.62; beryl, 1.58.

Dispersion. Dispersion is the breaking up of white light into its component colors. It is measured by the difference in the refractive indices of certain of the red and blue wavelengths of light as they pass through a gem. In order to give comparable figures for the dispersive powers of different gemstones, two specified wavelengths in the red and blue ends of the Sun's spectrum are used. For example, the refractive indices for these two wavelengths in spinel are 1.710 and 1.730, respectively; therefore, the amount of dispersion is expressed as 0.020. A comparable figure for synthetic rutile is 0.330.

Selective absorption. Although diamond and colorless zircon are valued for their brilliancy and prismatic fire, gemstones are more often prized for the loveliness of their colors. The color of most gemstones is caused by selective absorption. Selective absorption refers to a gem's ability to absorb or transmit certain wavelengths of light more readily than others. The wavelengths transmitted with least absorption are those that give a gem its color. Absorption is usually caused by very small percentages of metallic oxides that are present as impurities.

Pleochroism. Some gemstones exhibit pleochroism, the property of some doubly refractive materials of absorbing light unequally in the different directions of transmission, resulting in color differences. If two different colors are exhibited, the result is called dichroism (displayed by ruby, sapphire, and emerald); if three colors are seen, the result is called trichroism (displayed by alexandrite and tanzanite). *See* DICHROISM; TRICHROISM.

Asterism. Some gemstones depend on unusual optical effects for their beauty. Perhaps the most important of these are the star stones, which display the phenomenon called asterism. In a star, the reflection of light from lustrous inclusions is reduced to sharp lines of light by a domed cabochon style of cutting. The usual star effect is the six-rayed star seen in star sapphires and rubies. Four-rayed stars occur in garnets and some spinels. Other six-rayed stars are occasionally seen in such stones as quartz and beryl. In corundum, the most important species in which asterism is a significant feature, the phenomenon is caused by reflection from three sets of needlelike inclusions arranged in planes perpendicular to the sides of the hexagonal (six-sided) crystal, with each of the three sets of inclusions in planes parallel to a pair of faces of the crystal. *See* CORUNDUM.

Chatoyancy. Another important optical phenomenon is chatoyancy, or a cat's-eye effect. The most important gem in which it is a prominent feature is the variety of the mineral chrysoberyl known as precious

cat's-eye. This effect, when seen in chrysoberyl, is usually much more silkily lustrous than in any other gemstone. As a result, the term precious cat's-eye has come to be applied to the finer specimens of the chrysoberyl variety. Other kinds of gemstones in which cat's-eyes are sometimes encountered include tourmaline, quartz, beryl, scapolite, diopside, and some of the other rarer gem minerals.

Play of color. The gem mineral opal is one of the most attractive of the phenomenal gems. The colorful display called play of color results from light interference caused by diffraction effects from tiny spheres of material joined to make up the body of the opal.

Adularescence. Another optical phenomenon that produces an interesting gem variety is adularescence, the billowy light effect seen in adularia or moonstone varieties of orthoclase feldspar. In its most attractive form, it has a slightly bluish cast. Moonstone and opal are among the softer of the fairly common gem minerals and have hardnesses of approximately 6 on Mohs scale.

Precious and semiprecious stones. For many years gemstones have been commonly designated as precious or semiprecious. This is a somewhat meaningless practice, however, and often misleading, since many of the so-called precious gem varieties are inexpensive and many of the more attractive varieties of the semiprecious stones are exceedingly expensive and valuable. For example, a cabochon of fine-quality jadeite may be valued at approximately 100 times the price per carat of a low-quality star ruby. Fine black opals, chrysoberyl cat's-eyes, and alexandrites are often much more expensive than many sapphires of certain colors. *See* CARAT.

Occurrence. Gemstones occur under a variety of conditions in nature. A number of the highly prized minerals occur as primary constituents in igneous rocks or in alluvial deposits. The only known occurrences of diamonds in primary deposits are in either of the ultrabasic rocks (those with a low percentage of silica—less than 40%) known as kimberlite, a type of peridotite; or in lamprophyre. Although most primary diamond deposits are volcanic necks or conduits (so-called pipe deposits), dikes and sills of kimberlite or lamprophyre also may contain diamonds. Diamond-bearing pipes are known in South Africa, Tanzania, the two Congos, Angola, Sierra Leone, India, Lesotho, Russia, Australia, Canada, and the United States (Arkansas). Extensive alluvial deposits are mined in Ghana, Angola, Zaire, Sierra Leone, Guinea, Ivory Coast, Namibia, South Africa, Brazil, Guyana, and Venezuela. *See* IGNEOUS ROCKS; LAMPROPHYRE; PERIDOTITE.

Perhaps the most important types of primary deposits in which gemstones other than diamond are found are pegmatite dikes, contact metamorphic deposits, and weathered lava flows containing corundum. The intrusion of molten rock into impure limestone often brings about the crystallization of calcite to form a coarse-grained marble, plus a broad assemblage of other minerals. Under ideal conditions, some of the minerals produced in the contact zone include gem-quality crystals; for example, corundum and spinel. The Mogok area of Myanmar is the foremost supplier of the finest rubies and very fine sapphires. Pegmatite dikes are found in a number of places throughout the world, but only a few of them produce gem-quality crystals. The most important gem pegmatite areas include Brazil, Madagascar, Mozambique, Zambia, Namibia, Nigeria, and southern California. Minerals recovered from pegmatite dikes include aquamarine (and the other varieties of beryl other than fine emerald), topaz, tourmaline, and the kunzite variety of spodumene. Fine emeralds occur in calcite veins in black shale near Muzo, Colombia. Elsewhere most emeralds occur in mica schist, sometimes with other beryllium minerals, such as chrysoberyl. *See* PEGMATITE.

Most of the important colored-stone mining is done in alluvial deposits. For example, the island of Sri Lanka, long known as the "island of gems," has produced a variety of gemstones from alluvial deposits for many years.

Identification. Since most gems are varieties of mineral species, and since mineral species have fairly constant chemical compositions and characteristic crystal structures, their physical and optical properties vary within rather narrow limits. Thus, identification is a matter of measuring these properties and other characteristics, the most important of which are refractive index, optic character, and specific gravity. Because less than 10% of all known mineral species have varieties that are useful as gems, comparatively few tests are needed to separate them; only the first three tests are required in the majority of identifications. However, the ready availability of synthetic gemstones, plus a number of different enhancement techniques to improve the appearance of natural gemstones, complicates this task.

Origin. The properties of synthetic gem materials are identical to those of their natural counterparts; therefore, it is usually necessary to study suspected stones under magnification to determine whether their growth characteristics are natural or the result of synthetic production. For example, the accumulation lines in the form of color banding in natural sapphires are arranged in a hexagonal pattern, whereas those in many synthetic materials are curved. Spherical gas bubbles are characteristic of synthetics, but the inclusions in natural sapphires are angular. These same conditions also hold true for ruby and synthetic ruby. To make this determination, binocular magnification with dark-field illumination is the most effective (**Fig. 1**).

Refractive index. The relative refractivity of a gemstone is measured on a refractometer, an instrument that measures the critical angle between an optically dense hemisphere and the gemstone being tested. The instrument is calibrated directly into refractive indices and, as a result, readings from the scale are not in angles but in refractive indices. Doubly refractive minerals show two readings if the birefringence is sufficiently pronounced (birefringence is the difference between the two indices of a doubly refractive material). Thus, a characteristic birefringence

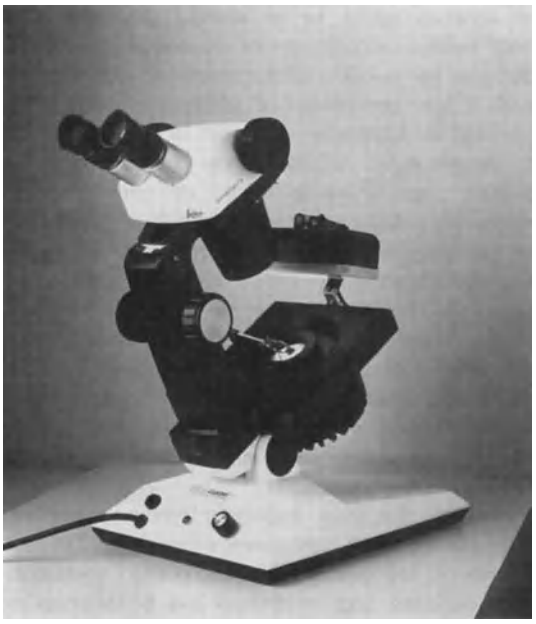

Fig. 1. Gemolite, a binocular magnifier with dark-field illuminator. (*Gemological Institute of America*)

Fig. 2. Polariscope to distinguish between single and double refraction. (*Gemological Institute of America*)

may be read as well as a characteristic refractive index. For example, tourmaline shows two readings (one at 1.624, another at 1.644), which alone are sufficient to identify this gem mineral. If monochromatic sodium light is used instead of ordinary incandescent illumination, it is possible to read indices accurately to the third decimal place. When two readings are seen, the necessity for the use of the polariscope, the instrument used to distinguish between single and double refraction, is unnecessary. The polariscope (**Fig. 2**) utilizes two polaroid plates, set so that their vibration directions are at 90° to one another, and the gemstone is examined between the two plates. A singly refractive gemstone rotated between the plates remains uniformly dark, whereas a doubly refractive stone becomes alternately light and dark. Diamonds, garnets, and glass remain dark in the polariscope, whereas emerald, ruby, tourmaline, topaz, and zircon become alternately light and dark.

Specific gravity. The density of a gem material compared with an equal volume of water is measured either by heavy liquids or by weighing the stone in air and in water on a sensitive balance adapted for this purpose. Gem materials vary in specific gravity from amber, at about 1.05, to hematite, at 5.20.

Fluorescence characteristics in ultraviolet light, spectroscopy, x-ray diffraction, chemical composition, and hardness and other tests may be helpful in identification (see **table**).

Gem materials. More than 100 natural materials have been fashioned at one time or another for ornamental purposes. Of these, however, only a relatively small number are likely to be encountered in jewelry articles.

Hardness, specific gravity, and refractive indices of gem materials

Gem material	Hardness	Specific gravity	Refractive index
Amber	2–2½	1.05	1.54
Beryl	7½.–8	2.67–2.85	1.57–1.58
Synthetic emerald	7½–8	2.66–2.7	1.56–1.563 to 1.57–1.58
Chrysoberyl and synthetic	8½	3.73	1.746–1.755
Corundum and synthetic	9	4.0	1.76–1.77
Diamond			
Synthetic cubic	10	3.52	2.42
Zirconia	8½	5.80	2.15
Feldspar	6–6½	2.55–2.75	1.5–1.57
Garnet			
Almandite	7½	4.05	1.79
Pyrope	7–7½	3.78	1.745
Rhodolite	7–7½	3.84	1.76
Andradite	6½–7	3.84	1.875
Grossularite	7	3.61	1.74
Spessartite	7–7½	4.15	1.80
Hematite	5½–6½	5.20	
Jade			
Jadeite	6½–7	3.34	1.66–1.68
Nephrite	6–6½	2.95	1.61–1.63
Lapis lazuli	5–6	2.4–3.05	1.50
Malachite	3½–4	3.34–3.95	1.66–1.91
Opal	5–6½	2.15	1.45
Pearl	3–4	2.7	
Peridot	6½–7	3.34	1.654–1.690
Quartz			
Crystalline and synthetic	7	2.66	1.54–1.55
Chalcedony	6½–7	2.60	1.535–1.539
Spinel and flux synthetic	8	3.60	1.718
Synthetic spinel, flame	8	3.64	1.73
Spodumene	6–7	3.18	1.66–1.676
Topaz	8	3.53	1.61–1.63
Tourmaline	7–7½	3.06	1.62–1.64
Turquois	5–6	2.76	1.61–1.65
Zircon	7½	4.70	1.925–1.98
Metamict	7	4.00	1.81
Zoisite (tanzanite)	6–7	3.35	1.691–1.70

Amber. Amber is a fossil resin that is usually yellow to brown, but it is occasionally found in other colors. In gem quality it is used mostly for beads or carvings. Amber is also used for pipe stems, cigarette holders, and other items. *See* AMBER.

Beryl. Beryl is a beryllium-aluminum silicate. In gem quality it is usually transparent and faceted, but translucent beryl may be carved into attractive forms used in jewelry. When green, it is known as emerald; when light blue, as aquamarine; when pink to light red, as morganite. It also occurs in yellow, light green, brown, orange, medium red, and colorless forms. *See* BERYL; EMERALD.

Chrysoberyl. Chrysoberyl is a transparent to translucent beryllium aluminate. Its varieties, known as cat's-eye and alexandrite, are among the important gemstones. Cat's-eye exhibits a band of light across the dome of a cabochon, which is caused by reflections from parallel silklike inclusions. The color is greenish yellow to yellowish green. Alexandrite changes color from green in daylight to garnet red in candlelight. *See* CHRYSOBERYL.

Coral. The gem material coral is an assemblage of colonies of the tiny marine animal coral. It is usually an orange to flesh color, but may be white or deep brownish red, and semitranslucent to nearly opaque.

Corundum. Corundum is transparent to translucent aluminum oxide. Ruby is the orange-red to violet-red variety; all other colors are called sapphire. The principal colors of sapphire other than blue are yellow, colorless, orange, pink, purple (amethystine), and green. Both ruby and sapphire occur with a beautiful six-rayed star effect. *See* RUBY; SAPPHIRE.

Diamond. Diamond, the transparent form of carbon, is the most important gemstone. In addition to the popular colorless form, it occurs in yellow, brown, pink, blue, and green. Bombardment by subatomic particles (followed by heat treatment for some colors) yields green, yellow, orange, and brown; blue is also possible. *See* DIAMOND.

Feldspar. The gemstone of importance in the feldspar group is moonstone. It is a semitransparent form of potassium feldspar that exhibits a floating light effect. Less important gems in the feldspar family include the amazonite variety of microcline, sunstone (a spangled gold-colored variety resembling moonstone, except for color), and the iridescent, translucent gray labradorite. Since several minerals are included in the feldspar group, the properties vary from species to species. *See* FELDSPAR; LABRADORITE; MICROCLINE; ORTHOCLASE.

Garnet. Within the garnet group are several distinct minerals that share the same crystal structure but differ in physical properties and chemical composition. Two of the garnet species, almandite and pyrope, have the dark red color commonly ascribed to garnet. A mixture of the almandite and pyrope compositions in about a 3:1 ratio produces a distinctively colored stone that is given the separate name rhodolite. It is a lighter, more transparent gemstone, with a distinctive violet-red color. Among other species of garnet are andradite, of which the lovely green demantoid is a variety; grossularite, of which the orange-brown hessonite and the green tsavorite are varieties; and spessartite, which has attractive orange and orange-red varieties. *See* GARNET.

Hematite. The principal ore of iron, hematite, sometimes occurs in dense, hard masses that take a very high polish; these are often fashioned into intaglios and cameos. Hematite is metallic grayish black in color. *See* CAMEO; HEMATITE; INTAGLIO (GEMOLOGY).

Jade. The material popularly known as jade may be a variety of either jadeite, a member of the pyroxene group, or nephrite, a combination of amphibole group minerals. Pyroxenes and amphiboles are major rock-forming minerals. The jades are characterized by a felted structure, creating exceptional toughness, which makes even delicate carvings relatively durable. In their gem qualities, they are also prized for their attractive coloring and translucency. Because it may occur in more saturated colors and may be more nearly transparent, green jadeite potentially is the more expensive of the jade minerals. It is semitransparent to semitranslucent and may be intense green, white with green streaks or patches, mauve, brown, yellow, orange, or violet. Nephrite is translucent to opaque and usually is a darker, less intense green than jadeite; it may also be off-white, gray, or black. *See* JADE.

Lapis lazuli. Lapis lazuli is an opaque, vivid deep-blue gemstone flecked with golden-yellow pyrite. It has been used for many centuries and is undoubtedly the stone to which the term "sapphire" was first applied. *See* LAZURITE.

Malachite. The colorful, opaque mineral malachite is often banded in two or more tones of green and may show a radial fibrous structure. It is always green and is often accompanied by the deep violet-blue azurite, another copper mineral. Malachite is used mostly for cameos and intaglios and for inexpensive scarab bracelets. *See* AZURITE; MALACHITE.

Opal. The fabled beauty of opal is due to diffraction and light interference in which patches of intense colors are seen against either a white or a nearly black background. There is also a transparent orange to red variety called fire opal. Since it is a relatively fragile material, opal must be treated with care. *See* OPAL.

Pearl. Oriental pearls are those found as lustrous, nacreous concretions in one of the three species of the salt-water mollusk genus *Pinctada*. Concretions in edible oysters are without pearly luster and are valueless. Pearls found in several genera of fresh-water clams are called fresh-water pearls, to distinguish them from Oriental pearls. Pearls usually occur in white, cream, or yellow colors with rose or other overtones, although black and gray are also very desirable. *See* PEARL.

Peridot. Peridot is a yellowish-green to green variety of the mineral group olivine; it is a magnesium-iron-aluminum silicate. In gem quality, it is always transparent and olive green. *See* OLIVINE.

Quartz. In its two major types, single crystal and cryptocrystalline (or chalcedonic), quartz, the most common mineral, has more gem varieties than any

other mineral species. The important varieties of crystalline quartz are amethyst, citrine, aventurine, and tiger's-eye; the most important varieties of cryptocrystalline quartz are carnelian, sard, chrysoprase, bloodstone, agate, and onyx. Amethyst is purple to violet and transparent. Citrine is yellow to brown and also transparent; it is more commonly known as topaz-quartz and, unfortunately, is often sold as a topaz. Aventurine is usually green with lustrous or colored spangles; it is translucent. Tiger's-eye is a translucent, fibrous, broadly chatoyant, yellow-brown stone that may be dyed other colors. Carnelian is red to orange-red, and sard is a darker brownish red to red-brown; both are translucent. Chrysoprase is light yellowish green. Bloodstone is dark green with red spots. Agate is a translucent chalcedony showing curved bands in a variety of colors. Onyx is similar to agate, except that the bands are straight. *See* AGATE; AMETHYST; ONYX; QUARTZ.

Spinel. This gem mineral is of particular interest because of the strong resemblance of many of its varieties to comparable colors of corundum. In general, red spinel is less intense in color than ruby; similarly, blue spinel is less intensely colored than blue sapphire. On the other hand, some varieties of spinel are lovely in their own right. Flame spinel is an intense orange-red and is a very attractive gemstone; spinel also occurs in green and amethystine colors. All gem varieties of spinel are transparent, with the exception of the rare black star spinel. *See* SPINEL.

Spodumene. Spodumene is a member of the pyroxene group of minerals; in contrast to jadeite, however, it is fragile. The principal variety of spodumene for gem purposes is kunzite, which is a lovely light-red to light-purple transparent stone. *See* SPODUMENE.

Topaz. Topaz is best known in its yellow to brown variety, but the red, pink, and blue varieties are also attractive and desirable. Colorless topaz can be irradiated in a nuclear reactor and then heat-treated to produce an attractive light blue color. *See* TOPAZ.

Tourmaline. Although the color range of tourmaline is as wide as that of any gem material known in nature, its best known varieties are red (rubellite) and dark green; it may also be colorless, yellow, blue, black, brown, or other colors. *See* RUBELLITE; TOURMALINE.

Turquoise. Turquoise is an opaque gemstone with an intense light-blue color. Its intense color has attracted people from the earliest times. *See* TURQUOISE.

Zircon. Zircon is best known as a transparent colorless or blue gemstone. The colorless variety has been used principally as an inexpensive substitute for diamond. It also occurs in green, yellow, brown, red, and flame colors. The properties of zircon vary rather widely. *See* ZIRCON.

Zoisite. A mineral that had been associated with rather drab varieties achieved a gem quality when, in 1967, in Tanzania, a variety dubbed tanzanite was discovered. It has an exceptionally pleasing vivid blue color, reminescent of the finest so-called Kashmir sapphires from the Himalaya area.

Manufactured Gems

Some minerals or other materials that have sufficient beauty and durability for use as a personal adornment can be manufactured. The term "manufactured," as used here, does not include such processes as shaping, faceting, and polishing, but only the processes that affect the material from which the finished gem is produced. These processes are (1) those that change the mineral in some fundamental characteristic, such as color, called a treated gem; (2) those by which a material is made that is identical with the naturally occurring mineral, called a synthetic gem; and (3) those that produce a simulated material with the appearance but not both the composition and structure of the natural gem, called an imitation gem. *See* PRECIOUS STONES.

Treated gems. Gem treatment to improve color or to conceal or make flaws less obvious has become increasingly pervasive. Although the U.S. Federal Trade Commission mandates full disclosure, treatments are not revealed. Some minor treatments are undetectable, such as gentle heating to change or make a color more uniform. Major treatments have a profound effect on apparent value. If the separations on a heavily flawed diamond are filled with a foreign substance with a high refractive index, the fractures may become all but invisible. Later heating during a prong retipping, or ring sizing may make them more visible than before the treatment, or cause additional damage.

In a different context, many colored stones, including most green tourmaline, aquamarine, and colorless and flame-colored zircon and most pink topaz and blue zircon have been heated to improve their color. This tends to be regarded in the gem industry as being as much a part of normal market preparation as cutting and polishing. This does not apply to intense heating of close to the melting point such as is done to impart color by diffusion of a coloring agent into sapphire, for example.

There are a number of different types of treatment, including dyeing or staining, impregnation by paraffin, plastic, or some other material; filling in fractures or cavities to hide them; bleaching; laser drilling to reach ugly inclusions in order to make them less obvious; heat treatment; diffusion of color into the near surface of a gem too light in color to be high in value; and irradiation by subatomic particles. In some efforts to make a stone more attractive, two or more treatments are employed.

Dyeing and staining. This method of coloration was one of the earliest to have been adopted widely. The most frequently used gem subjected to dye is agate (cryptocrystalline quartz, or chalcedony). Colored substances are introduced into the minute interstices between the microscopic quartz crystals. Because of the varying porosity in the individual layers, the difference in color between adjacent layers is intensified. The best method of coloring a given specimen is usually determined by trial-and-error

treatment of a small sample. In general, the process of dyeing an agate consists of first drying the specimen thoroughly; second, soaking it in an appropriate liquid dye until the interstices are penetrated; and finally, fixing the dye within the pores. In the processing of cultured pearls eosine dyes are often used.

Fracture and cavity filling. By a process first developed in Israel, diamonds with prominent cleavages or fractures are subjected to a treatment by which a liquid with a high refractive index is introduced into the separation. When it solidifies, the high refractivity of the transparent in-filling reduces the visibility of the separation. Thus, the clarity appears often to be improved greatly. Cavities in colored stones of high per-carat value are occasionally filled with a glass of similar refractive index as the host material. Many more emeralds contain inclusions than most other gemstones. Almost as long as emeralds have been used as gems, miners have presented them to buyers after soaking in some type of oil, which tends to improve apparent clarity.

Bleaching. Oxidizing agents or acids are sometimes used to remove unwanted colors. In the processing of cultured pearls employing bead nuclei, the first deposition is likely to be brown. After harvesting, this unwanted color gradually fades, but the processor speeds the fading by using a bleaching step to make matching for necklaces possible more quickly, and with less chance of later change in appearance.

Laser drilling. A diamond with a large and obvious internal inclusion that does not reach the surface may be enhanced in apparent clarity grade by employing a laser to drill a tiny hole to the feature, so that a bleaching or dissolving agent may reach it to reduce its visibility. Dramatic improvement in appearance is possible. *See* LASER.

Impregnation. The gemstone most often subjected to impregnation has been turquoise, but jade has also been bleached to remove an unattractive color, and then impregnated with plastic to restore durability. Since the demand for fine-quality natural turquoise greatly exceeds the supply, porous, chalky, poor-quality turquoise is soaked in paraffin or impregnated with plastic. The soaking deepens and improves color, and a plastic impregnation improves both color and durability.

Heat treatment. Heating in one or more different atmospheres is often used to improve the natural color of a gem or to develop new colors. As found in nature, most aquamarine is light to medium yellowish green, a color much less in demand that the lovely blue of a fine aquamarine. Light heating effects a permanent color change to the blue associated with aquamarine. Some amethyst becomes more evenly colored with gentle heating. When heated to about 450°C (842°F), most amethyst changes to a yellow- to reddish-brown citrine color, but the amethyst from a few localities changes to a green color when heated to this temperature, and the brown color of smoky quartz may often be changed to yellow or reddish-yellow citrine. Some yellow or brownish zircons change to blue, golden, or red when they are heated in a reducing atmosphere, or become colorless or rich yellow if heated in an oxidizing atmosphere. Over the years, heating techniques and controls have become increasingly sophisticated, with temperature controlled to less than 1°C (1.8°F), and with the type of atmosphere also controlled.

Diffusion. After effecting atmospheric and temperature controls to improve heat-treatment consistency, a logical next step was to introduce coloring agents. This practice has become very common, particularly in Bangkok. Faceted colorless or pale sapphires are heated to close to the melting point in a bed of the appropriate metallic oxide chromophore or chromophores. When successful, enough of the chromophore migrates into the lattice to impart color. This process may be repeated several times to increase the penetration.

Radiation. Irradiation of different types induces many color changes in gems. The induced color varies with the original material and with the kind and intensity of radiation used. X-rays and radiation from radium, radioactive isotopes, electrons, fast neutrons, and other sources may produce striking changes. Diamond subjected to bombardment in a cyclotron or a nuclear reactor first changes to a green color, and, with more radiation, the green passes to black with an accompanying breakdown in the crystal structure. If the irradiation is not carried too far, the change can be reversed to a rich yellow or brown by annealing at high temperatures. After annealing, the color changes appear to be permanent. Many changes induced by irradiation are short-lived. The term metamict is used to describe crystals in which structural changes have been introduced by intensive natural radiation. *See* METAMICT STATE; PLEOCHROIC HALOS.

Synthetic gems. The U.S. Federal Trade Commission has restricted the term synthetic gems to manufactured materials that have the same chemical, physical, and optical properties as their naturally occurring counterparts. Of the gem materials, including diamond, that have been synthesized, many have been produced in such small sizes or poor quality that they are unsatisfactory as gemstones. Some attempts to make gemstones have resulted in producing substances not known in nature, a few of which are of great importance industrially. Others have resulted in significant improvements in existing processes.

In 1891 E. G. Acheson, while attempting to synthesize diamond, produced SiC (carborundum), the third hardest known substance after diamond, and artificially produced borazon, a boron nitride. Knowledge gained in attempting to make ruby and sapphire, gem varieties of the mineral corundum (Al_2O_3), aided in the development of nongem crystals of corundum, which, with SiC, form the basis of the abrasives industry. The quantity of synthetic ruby used as jewel bearings in watches and electrical instruments greatly exceeds that used as gems.

In 1957 began the production and sale of an industrial-diamond grit for abrasive use, averaging

0.004 in. (0.1 mm) in diameter, in competition with natural diamond grit, which is made by crushing low-grade diamond crystals. The abrasive use of diamond grit for the manufacture of bonded-diamond-grinding wheels greatly exceeds in volume, but not in value, the amount of diamond used for other industrial applications and for gemstones. In 1971, the experimental production of cuttable-quality diamonds in sizes of over 1 carat was announced. These were grown under conditions of such extreme temperatures and pressures that the equipment costs made the product more costly than comparable natural diamonds. Since then, larger cuttable-quality crystals have been grown. In 1993, Russian-grown synthetic diamonds in cut stones of up to a carat were made available.

Verneuil process. This process uses the high temperatures produced by an oxyhydrogen flame. In 1902 A. V. L. Verneuil, a French chemist, announced the production of a synthetic ruby of gem quality made by a process that still bears his name; it is also called the flame-fusion method.

1. *Sapphire and ruby*. Synthetic ruby and sapphire are produced by the fusion of very finely divided and highly pure (Al_2O_3) in an apparatus that resembles an oxyhydrogen torch, with the flame directed downward into an insulating chamber. The powder passing through the flame fuses into droplets, which initiate the development of a cylindrical boule when the flame impinges on a fire-clay support. The diameter of the boule is controlled by lowering the fire-clay support while regulating the gas flow and the amount of Al_2O_3 introduced. The upper surface of the boule remains molten. Various colors are produced by the addition of appropriate metallic oxides. It is essential to exclude traces of sodium, which prevent the formation of clear boules. The average commercial Verneuil boule is about $^1/_2$ in. (13 mm) in diameter and 1–2 in. (25–50 mm) long, and weighs about 125 carats. Crystallographically oriented boules may be produced with a properly oriented seed crystal. Rods of transparent Al_2O_3 as small as 0.04 in. (1 mm) in diameter and up to 18 in. (46 cm) long may be made by a semiautomatic process. After being ground to uniform diameter, several rods are held firmly together in a bundle and sawed perpendicular to their length with a diamond saw. Each saw cut produces a circular blank from each rod, which is processed into a jewel bearing. From mushroomlike boules as large as 4 in. (10 cm) in diameter, circular disks are cut and polished. Because of the resistance of aluminum oxide to high temperatures, these disks are used as windows in furnaces. They are also highly conductive in the infrared region of the spectrum; and because of this conductivity, the rods may be inserted into furnaces to transmit the infrared radiation into heat-controlling apparatus on the outside of the furnace.

Ruby is used for jewel bearings in preference to clear sapphire because colored jewels are easier for workers to see. Ruby is made by adding a small percentage of chromium oxide (Cr_2O_3) to the alumina. Where extreme hardness is not essential, the more easily worked spinel is preferred. Colored varieties are easier to make and give clearer crystals than corundum for all colors except red. The hardness of spinel, 8 on Mohs scale, is adequate for most gem substitute purposes. Synthetic spinel boules are elongated cubes with rounded corners. Synthetic star rubies are made by adding an excess of titania (titanium oxide; TiO_2) to the alumina powder and then heating the finished boules at a temperature above 1830°F (1000°C). The titania recrystallizes as needles oriented by the host crystal, in the plane of the basal pinacoid, in a hexagonal pattern that is controlled by the crystal structure of the alumina.

2. *Rutile*. Another mineral synthesized by the flame-fusion process is rutile. For the production of synthetic rutile (titania), the flame-fusion apparatus is modified to oxidize the titanium more completely by the addition of a third tube at the orifice surrounding the other two tubes through which oxygen passes. The boules, as produced, are black and must be heated in a furnace in an oxygen atmosphere at 2550–2730°F (1400–1500°C) to remove the color. Unless metallic coloring oxides have been added, the boules retain a faint tinge of yellow after the oxygen-furnace treatment, because of an absorption band in the blue. The unoxidized titanium atoms in undercolorized boules make them semiconductors whose conductivity is proportional to the amount of color remaining. Synthetic titania boules form elongated tetragonal prisms with rounded corners. The boules are clear and transparent, larger than natural rutile crystals, and much superior in quality. Natural rutile is always brownish, even when transparent. Since the introduction of synthetic cubic zirconia, synthetic rutile is used rarely as a diamond substitute. *See* RUTILE.

3. *Strontium titanate*. Another once popular diamond substitute is strontium titanate. Strontium titanate made by the flame-fusion method has, since 1955, been a successful imitation of diamond. It has an index of refraction nearly the same as that of diamond and a somewhat higher dispersion, which gives it a brilliancy approaching, and a degree of fire somewhat greater than, diamond.

4. *Other melt methods*. These include the Czochralski process, the floating zone process, and the skull melt. Czochralski, or crystal pulling technique, melts the appropriate material, such as aluminum oxide in an inert crucible, inserts a seed crystal barely into the melt and withdraws it gradually as the melted material crystallizes on the retracting seed. By this method, large fine crystals of synthetic ruby, sapphire, alexandrite and other gem-quality crystals are produced. With modifications of this technique, very large crystals have been produced.

In the floating zone process of synthesis, a rod is prepared by sintering the essential ingredients of a gem material. The rod is held at both ends, and one end is subjected to a narrowly focussed convergence of infrared rays that bring the mixture to the melting point. The ends of the rod are rotated slowly in opposite directions. The heat is applied to move slowly

down the rod, and it slowly becomes a single crystal as crystallization takes place from top to bottom. This is used to create synthetic rubies, sapphires, and the alexandrite variety of chrysoberyl.

The ingenious skull melt method for growing crystals of highly refractory materials owes its name to the appearance of the container. A series of water cooled tubes make up a container that is filled with the purified powder of the refractory material. The tubes are encircled by coils attached to a radio-frequency generator that heats the center portion of the powder to a very high temperature. The powder next to the cooling tubes insulates the tubes. This is the method by which huge quantities of zirconium oxide, the cubic zirconia (diamond substitute) is grown.

Flux growth. This is another fundamental method of synthetic crystal manufacture. Flux fusion, or flux melt, employs a material with a relatively low melting point in which the material to be synthesized will dissolve much more readily than in an aqueous solution. Slow cooling brings about the crystallization of the desired product.

Hydrothermal process. Synthesis by crystallization from aqueous solutions has been employed in numerous laboratories experimentally. It was first used on a large scale by the Bell Laboratories to produce large, colorless quartz crystals. Later, they made synthetic ruby experimentally. The first major gem synthesis by this method was announced by the Linde Company in 1965 in the making of synthetic emeralds. Hydrothermal synthetic emeralds are manufactured commercially.

Assembled stones. Once one of the most important imitations, the doublet of garnet and glass is of minor importance today. A combination of two parts of synthetic spinel, plus a colored cement, is still used extensively to imitate ruby, emerald, and several other gemstones. As assemblage of natural and synthetic ruby or sapphire is used to simulate star ruby or star sapphire. The only other important assembled stone is made of a thin piece of opal cemented to a backing of quartz or common opal and covered with a protective layer of transparent quartz.

Imitation gems. Since prehistoric times glass has been the most widely used gem imitation. Since World War II colored plastics have replaced glass to a great extent in the least expensive costume jewelry.

Glass. Brilliants and rhinestones made of flint glass are often silvered on the back to give them a high brilliancy. Glass with a high dispersion, which is often called paste diamond, has a high lead-arsenic content. The term paste owes its origin to the practice of wet mixing the ingredients before firing. It is somewhat softer than ordinary glass.

Plastics. Modern plastics have made costume jewelry very popular, not only because plastics are cheap but because they are lighter in weight than glass and are easily molded and worked into various forms. Thermosetting plastics, such as Bakelite, were first used in jewelry early in the twentieth century. Unlike thermoplastics, which can be resoftened by heating after molding, thermosetting plastics become permanently hard after the initial molding. *See* PLASTICS PROCESSING.

Identification of substitutes. Materials made by a flame-fusion process almost always contain gas bubbles, which are usually spherical or nearly so. Those with medium to dark tones of color often show color banding, or striae, with a curvature corresponding to that of the top of the boule. Natural gem materials are characterized by angular inclusions and straight color bands, if any are present. Often, there are differences in absorption spectra, fluorescence, or transparency to ultraviolet or other properties sufficient to enable a gemologist to identify them.

Most flux-fusion synthetic emeralds have distinctly lower refractive indices and specific gravities than natural emeralds and are characterized by wisplike or veillike flux inclusions. Most show a red fluorescence under ultraviolet light, whereas most natural emeralds are inert. Two- or three-phase inclusions (liquid and gas or solid, liquid, and gas enclosed in the same space) are typical of natural emerald.

Synthetic flux rubies have veillike inclusions and a transparency to shortwave ultraviolet slightly greater than that of natural rubies.

Hydrothermally made synthetic emeralds have properties similar to many natural emeralds, but their inclusions differ and they are characterized by a very strong red fluorescence under ultraviolet.

The cheaper forms of glass are cast in molds and, under a hand lens, show rounded edges at the intersections of facets; the facets are often concave. The better grades, known as cut glass, have been cut and polished after first being molded approximately into the desired form. Cut glass has facets that intersect in sharp edges. Both types may contain gas bubbles or have a roiled appearance in the heart or have both, in contrast to most of the colored stones they imitate.

Assembled stones are detected usually by observing the plane where the parts are joined together. In a faceted stone, these parts may be seen more readily if the stone is immersed in a transparent liquid and viewed parallel to the girdle.

Gem Cutting

The polishing of rough gem materials into faceted or rounded forms for use in jewelry is called lapidary or gem cutting. The term lapidary is limited in application to the cutting of colored stones. One who fashions colored stones is called a lapidary, lapidarist, or lapidist.

Commercial colored stone and diamond cutting are separate and distinct fields. The problems of colored-stone cutting are different from diamond cutting, because all colored stones are soft enough to be shaped readily by using silicon carbide as an abrasive, although diamond may be used for faster results. In contrast to diamond, they may be sawed or polished in any direction. There are major differences between the cutting of the most valuable emerald and ruby rough material, and the less expensive materials. Weight loss in fashioning is usually greater in colored stone cutting, because diamond shapes are usually more adaptable to the round

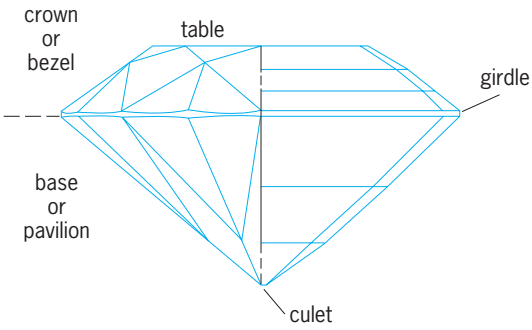

Fig. 3. Arrangement of faceted cuts, showing brilliant cut (left side) and step cut (right side).

brilliant that is the usual cut. Given the pleochroism of most valuable colored stones, hexagonal prisms are less likely to be adaptable to a high weight retention. Thus, when cutting the most valuable ruby and emerald crystals, planning is vital. Face-up appearance has to be balanced with weight retention potential. With less expensive rough material, mechanical, computer-controlled cutting may make it possible to cut dozens of exactly sized, almost cloned, stones at one time. In medium to expensive rough, an experienced lapidary may set by eye the angles to which facets should be cut in relation to the girdle plane or the top facet. In addition, the cutter creates by eye the perfect symmetry seen when the finished stone is viewed from above. In this respect, commercial gem cutting differs markedly from that done by hobbyist gem cutters, for the hobbyist usually controls the angles and the symmetry by the use of various mechanical faceting heads.

The first step in the usual cutting operation is to saw the crystal or piece of massive gem material to obtain the size and general shape desired. If the material is relatively inexpensive, it may be ground to shape rather than reduced by cutting. Next, if the stone is soft enough, it is usually ground into the approximate final shape against a silicon carbide wheel; if it is 9 on Mohs scale, a diamond-charged metal lap, or wheel, is used.

Facets are ground and polished on other laps. The commercial cutter often employs what is called the jamb-peg method. In this method, a square or rectangular board containing rows of partly drilled holes is clamped into a position at right angles to the spinning lap. The stone that is being faceted is mounted with adhesive on the end of a stick, called a dopstick, the opposite end of which is pointed. The cutter holds the pointed end of the stick in a hole chosen by the cutter as one representing the proper angle for the facet being ground. The end of the stick on which the stone is mounted is held against the spinning lap. The term dop is used both in diamond and colored-stone cutting.

Facet grinding is usually accomplished on copper, iron, or lead laps, using diamond or silicon carbide powder for ruby and sapphire or their synthetic counterparts, and silicon carbide powder for the softer materials. Polishing is accomplished on tin, wood, or plastic laps, with the wooden laps usually covered with leather or cloth. The polishing agents used include tin oxide, cerium oxide, alumina, and diamond powder. Both polishing and grinding laps are operated at rates of speed that are more rapid for hard materials than for soft ones.

Not all lapidaries appreciate the importance of the direct relationship between the proportions of the cut stone and its beauty. Rough gem material is so expensive that the natural tendency to retain every milligram possible in the final product often overcomes consideration of potential beauty. Since diamonds of fine quality are valued for their brilliancy and prismatic fire, the need for cutting to precise angles and proportions is appreciated by most diamond cutters. However, many depart somewhat from ideal figures to obtain larger stones that they believe will be more readily sold at a lower per-carat price. Since

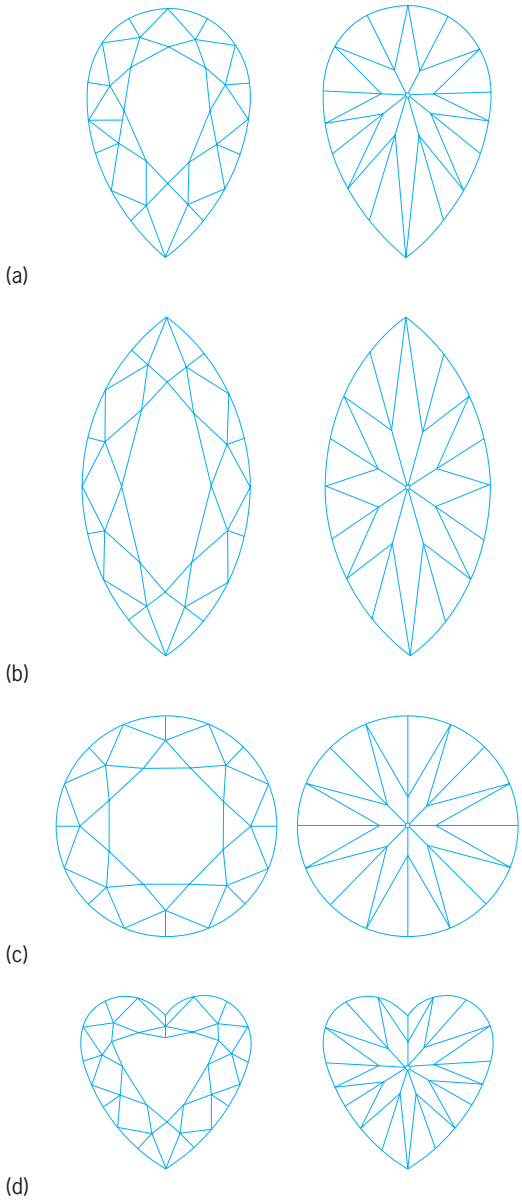

Fig. 4. Examples of brilliant-cut facets. (*a*) Pear-shape cut. (*b*) Marquise or navette cut. (*c*) Round brilliant. (*d*) Heart-shape cut. (*Gemological Institute of America*)

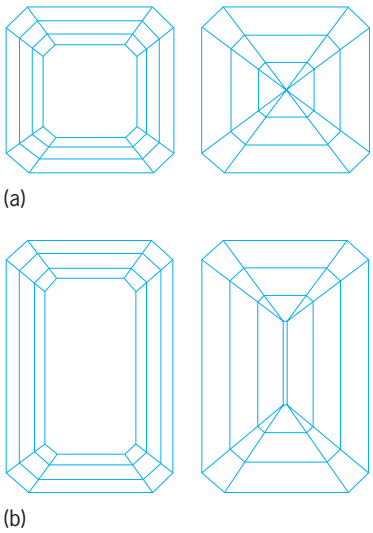

Fig. 5. Examples of step-cut facets. (*a*) Square emerald cut. (*b*) Emerald cut. (*Gemological Institute of America*)

colored stones are valued for their beauty of color, some lapidaries believe that any symmetrical cutting is satisfactory. Actually, the lower refractive indices of the important colored stones reduce both their potential brilliancy and their selective absorption, which produces the color. Many colored stones are so poorly cut that their reflective quality is seriously impaired.

There are many different styles to in which gems are cut. The two basic types are brilliant and step cutting. Both are characterized by a large facet (table) topping the crown facets, which slope to the periphery of the stone (girdle), and by a base portion, which slopes from the periphery to a tiny facet (culet) at the lowest point (**Fig. 3**). The two styles differ in that the facets on step cuts are parallel to the table and the girdle, so that they are all trapeze-shaped; whereas brilliant-cut facets, other than the table and culet, are triangular or kite-shaped. The brilliant style (**Fig. 4**) is usually used for stones with round or curved girdle outlines, and the step cut (**Fig. 5**) for those with square, rectangular, or other straight-sided girdle outlines.

Gem Mounting

Gemstones are usually set in rings or other jewelry pieces made of one of the precious metals. Those metals most widely used for jewelry purposes are yellow gold, white gold, platinum, and palladium.

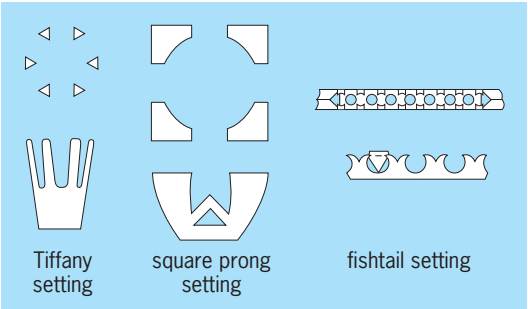

Fig. 6. Gemstone settings. (*Gemological Institute of America*)

Setting. When a stone is set in a ring, the portion in which the stone is placed is called the setting (**Fig. 6**). The stone may be held by small beads of metal, which are pushed over the edges by the stonesetter; by prongs, which are also bent over the edges of the stone to hold it securely; or by a rim of metal, which the stonesetter turns over the edges all the way around the stone.

Seat. In any type of setting in which the stone is held in place by metal, it is essential to the preservation of a stone subjected to normal wear that the "seat" into which it is placed be even. The seat is usually cut into the metal of the setting, after which the stone is pushed into position and the metal beads, prongs, or metal bezel are forced down around the edge of the stone. If the seat is uneven, the pressure on the stone is applied unevenly, and even a diamond may cleave or fracture when subjected to the light blows that it is certain to receive in the course of normal wear. Richard T. Liddicoat, Jr.

Bibliography. Gemological Institute of America, *Gem Reference Guide*, 1993; C. S. Hurlbut and R. C. Kammerling, *Gemology*, 2d ed., 1991; R. T. Liddicoat, *Handbook of Gem Identification*, 12th ed., 1987, reprint 1993; R. Webster and B. W. Anderson, *Gems: Their Sources, Description, and Identification*, 5th ed., 1994.

Geminga

A nearby neutron star that emits pulsed x-rays and gamma rays, steady optical radiation, and possibly radio and optical pulsations. Since the 1970s, it has been studied at a level of detail very unusual for a neutron star. As a result, not only are its nature and distance well known, but also a good understanding has been reached of the physical processes responsible for its multiwavelength emission.

Neutron stars. Understanding neutron stars remains a fundamental challenge for modern astronomy. They represent the final state of evolution for many normal stars. Thus, their study also holds potentially great rewards for basic physics because of the special state of the matter which composes them. They lack the radiating atmosphere of stars like the Sun, and thus their only thermal emission arises from the planckian (blackbody) radiation of their surfaces. The solid crust of a neutron star, probably made out of a layer of iron nuclei, is heated by its neutron superfluid interior initially at very high temperatures by the supernova explosion that generated the star. However, the tiny dimensions of the surface (with a radius of 10 km or 6 mi), as well as its rapid cooling (down to temperatures less than 10^6 K after 10^5 years), limit the overall luminosity of the neutron star to a small fraction of that of the Sun. As a result, the classical channel of optical astronomy has been largely unable to provide information about these stars. *See* HEAT RADIATION; SUPERNOVA.

Owing to their large, rapidly spinning magnetic fields, however, neutron stars also emit significant nonthermal luminosity, that is, radiation produced by the interaction of accelerated particles and electromagnetic fields, or by the spinning fields themselves. Historically, this radiation was first detected at radio frequencies with the discovery of pulsars in 1968. Subsequently, notably with the launch of *COS-B* in 1975 and of the *Compton Gamma-Ray Observatory* in 1991, came the observation of gamma-ray pulsars. These are radio pulsars that emit a large flux of high-energy (100-MeV) gamma-ray photons. At the same time, a gamma-ray source was discovered that had all the markings of being connected with a radio pulsar but had no such object in its optical positional error box in the sky. *See* GAMMA-RAY ASTRONOMY; RADIO ASTRONOMY; SATELLITE (ASTRONOMY).

Nature and radiation. The object, named Geminga (an acronym for the constellation Gemini, and a gamma-ray source therein but also derived from a phrase in Milanese dialect meaning "it does not exist"), is indeed a rotating neutron star. It has, however, a number of unique properties: It was the first unidentified gamma-ray source in the sky, and subsequently became the first isolated neutron star to be discovered through its x-ray and gamma-ray emission, without the help of radio astronomy. High-precision positional measurements using the *Hipparcos* satellite and the *Hubble Space Telescope* allowed, for the first time, the absolute astrometry of an object this faint. Through a delicate space-time correlation, this work has made it possible to count every single rotation of the neutron star (at the rate of about 4 per second) throughout a period of over 25 years. *See* ASTROMETRY; HUBBLE SPACE TELESCOPE; X-RAY ASTRONOMY.

Optical and ultraviolet observations, carried out from the ground and from space, have allowed the measurement of Geminga's annual parallax, defining its distance to be about 150 parsecs (450 light-years), and thus determining with accuracy the star's absolute energy output at each wavelength. X-ray observations with the *XMM/Newton* telescope have shown that the thermal radiation from the neutron star surface has two components: a "cool" one (0.5×10^6 K) coming from a large surface and a "hot" one (10^6 K) coming from a football-field-size hot spot, probably close to the polar region and heated by high-energy particles impinging on the neutron star surface. Since both thermal emission components are modulated by the neutron star rotation, for the first time it is possible to clearly "see" the rotating surface of Geminga down to detail as minute as a 100-m^2 (1000-ft^2) portion of the surface. The particles needed to heat the hot spot are similar to those required to produce the copious gamma-ray flux from Geminga, thus providing the missing link between the x-ray and gamma-ray emission from Geminga. *See* PARALLAX (ASTRONOMY).

Wide-band data have shown that both thermal and nonthermal processes are at work in this neutron star, which is approximately 350,000 years old. Such processes originate on its surface as well as in the extremely intense stellar magnetic field. They may be another unique feature of Geminga, a very promising astrophysical laboratory.

Interaction with interstellar medium. A diffuse emission of x-rays has been discovered surrounding Geminga with the *XMM/Newton* and *Chandra* x-ray observatories. This emission has a complex, elongated shape, precisely aligned with the object's proper motion (see **illustration**). The diffuse x-rays have a hard, nonthermal spectrum, suggesting that they originate from the interaction of extremely energetic pulsar electrons with the interstellar magnetic field.

The shape of the diffuse emission traces the bow shock created by Geminga as it moves supersonically through the interstellar medium. Interestingly, the lifetime of the radiating electrons in the diffuse emission corresponds precisely to the pulsar displacement time over the size of the emission region. This agreement clinches the case both for high-energy electron acceleration by pulsars and for active pulsar interaction with their surrounding interstellar medium. *See* SHOCK WAVE.

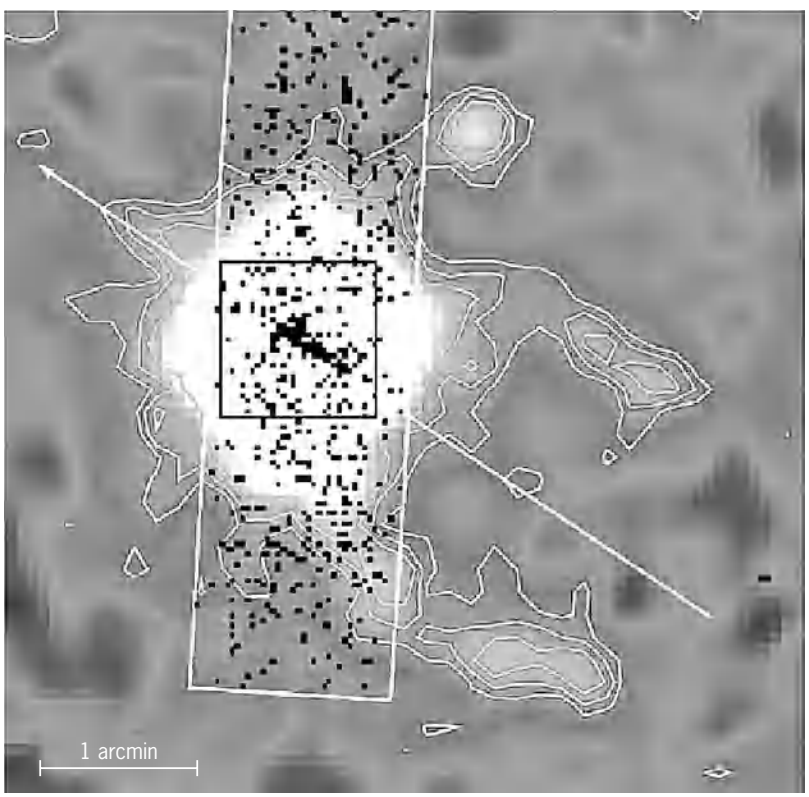

Superposition of x-ray images of Geminga. North is up, east is left. Shadings and contours show inner part of image obtained by the EPIC/MOS1 and MOS2 instruments on the *XMM/Newton* x-ray telescope (0.3–5.0-keV spectral range). Small black squares show image obtained by the ACIS instrument on the *Chandra X-ray Observatory* (0.45–7 keV, 0.492 arcsec pixel size), rebinned to a pixel size of 2 arcsec. The ACIS field of view is marked by the elongated white rectangle. The square box identifies a 1 × 1 arcmin area. In the *XMM/Newton* image, two elongated tails of diffuse emission are seen to emerge from the source, and show a remarkable symmetry with respect to the pulsar proper motion direction, marked by the arrow. Although insufficient in exposure and coverage to complement the *XMM/Newton* data, the *Chandra* exposure does show the existence of a thin 25-arcsec-long tail trailing the source and aligned to the object's motion, like the larger *XMM/Newton* tails. (*From A. DeLuca et al., On the complex x-ray structure tracing the motion of Geminga, Astron. Astrophys., 455:L9–L13, 2006*)

Closer to the central pulsar, high-resolution *Chandra* data offer additional information. Thanks to an x-ray optics that allows for 10 times sharper images than the *XMM* ones, *Chandra* detects a short tail of emission trailing Geminga, again along its proper motion direction. Its origin is probably different from that of the large-scale diffuse emission seen by *XMM/Newton*. Closer to the pulsar, accelerated electrons lose their energy in the pulsar-generated magnetic field, which is much stronger than the diffuse interstellar one. With a total of 50 photons, a clear indication of a hard spectrum is present in the *Chandra* data, pointing to this pulsar-wind electron synchrotron mechanism. Unfortunately, the limited field of view of the *Chandra* observation does not cover the large-scale *XMM* tails, thus preventing a direct comparison of the results obtained by the two x-ray observatories. While *XMM/Newton* is more sensitive to large-scale features, *Chandra's* superior resolution can explore the immediate surroundings of the neutron star.

The resulting combined Geminga image represents the most complete description of a magnetized neutron star in its environment so far. Geminga is the first neutron star showing both an outer, large-scale bow shock and an inner comet-like pulsar tail. *See* CHANDRA X-RAY OBSERVATORY; NEUTRON STAR; PULSAR. Giovanni F. Bignami

Bibliography. G. F. Bignami, Astronomy of the invisible, *Phys. World*, 16(9):27–30, September 2003 (http://physicsweb.org/articles/world/16/9/7/1); G. F. Bignami and P. A. Caraveo, Geminga: Its phenomenology, its fraternity, and its physics, *Annu. Rev. Astron. Astrophys.*, 34:331–381, 1996; P. A. Caraveo et al., Geminga's tails: A pulsar bow shock probing the interstellar medium, *Science*, 301:1345–1347, 2003; A. DeLuca et al., On the complex x-ray structure tracing the motion of Geminga, *Astron. Astrophys.*, 455:L9–L13, 2006.

Gemini

The Twins, a winter zodiacal constellation (see **illustration**). It is marked by a pair of bright stars, Castor and Pollux, usually pictured in the heads of the twins. This pair is easy to pick out with the unaided eye since they are of similar brightness and are separated by less than a fist's width when held at the end of the observer's outstretched arm. *See* ZODIAC.

In Greek mythology, Castor was a mortal, the son of the king of Sparta, and his half-brother Pollux was a son of Zeus. Both were sons of Leda, the queen of Sparta. After Castor died and went to Hades, the brothers could not bear the separation, and were reunited permanently in the heavens.

Castor is actually a multiple star containing two double stars (known to be double from their spectra) and two additional stars. Gemini includes part of the Milky Way, so it is rich in star clusters. *See* BINARY STAR; STAR CLUSTERS.

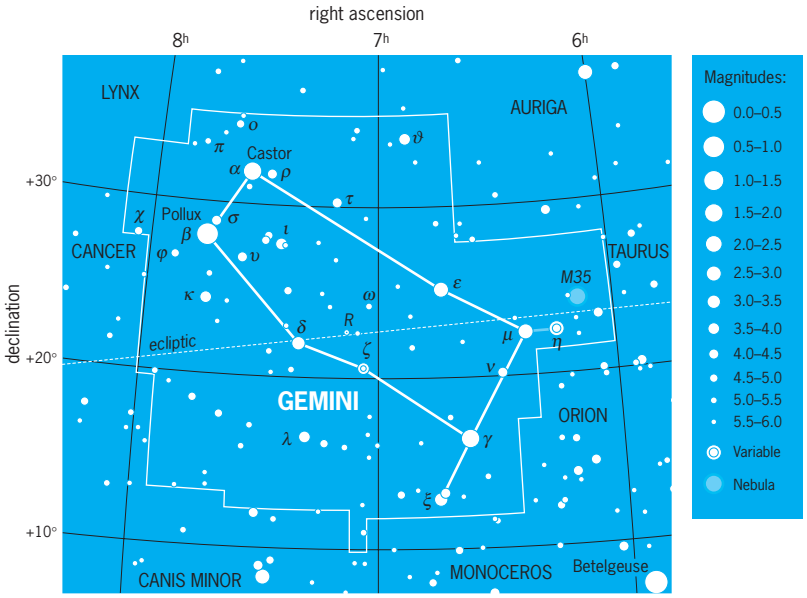

Modern boundaries of the constellation Gemini, the Twins. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION. Jay M. Pasachoff

Gemology

The science of those minerals and other materials which possess sufficient beauty and durability to make them desirable as gemstones. It is concerned with the identification, grading, evaluation, fashioning, and other aspects of gemstones.

For many years the little systematic attention given the subject was divided between jewelers and mineralogists. Mineralogy itself gained the status of an exact science only in the nineteenth century. Gemology was given impetus by the accelerating developments in synthetics, imitations, various means of color alteration practiced to defraud, and a growing need for developing methods of detecting these products. *See* GEM. Richard T. Liddicoat, Jr.

Gene

The basic unit in inheritance. There is no general agreement as to the exact usage of the term, since several criteria that have been used for its definition have been shown not to be equivalent.

The nature of this difficulty will be indicated after a description of the earlier position. The facts of mendelian inheritance indicate the presence of discrete hereditary units that replicate at each cell division, producing remarkably exact copies of

themselves, and that in some highly specific way determine the characteristics of the individuals that bear them. The evidence also shows that each of these units may at times mutate to give a new equally stable unit (called an allele), which has more or less similar but not identical effects on the characters of its bearers. *See* ALLELE; MENDELISM.

These hereditary units are the genes, and the criteria for the recognition that certain genes are alleles have been that they (1) arise from one another by a single mutation, (2) have similar effects on the characters of the organism, and (3) occupy the same locus in the chromosome. It has long been known that there were a few cases where these criteria did not give consistent results, but these were explained by special hypotheses in the individual cases. However, such cases have been found to be so numerous that they appear to be the rule rather than the exception. For a more complete discussion *See* MUTATION; RECOMBINATION (GENETICS).

The term gene, or cistron, may be used to indicate a unit of function. The term is used to designate an area in a chromosome made up of subunits present in an unbroken unit to give their characteristic effect. It is probable that with increasing knowledge of the nature and properties of deoxyribonucleic acid (DNA) it will become possible to reach a more generally acceptable solution to the problems of terminology. *See* DEOXYRIBONUCLEIC ACID (DNA); NUCLEIC ACID.

Alfred H. Sturtevant

Molecular biology. Every gene consists of a linear sequence of bases in a nucleic acid molecule. Genes are specified by the sequence of bases in DNA in prokaryotic, archaeal, and eukaryotic cells, and in DNA or ribonucleic acid (RNA) in prokaryotic or eukaryotic viruses. The flow of genetic information from DNA, to messenger RNA (mRNA), to protein is historically referred to as the central dogma of molecular biology; however, this view required modification with the discovery of retroviruses, whose genetic flow goes from RNA to DNA by reverse transcription and then to mRNA and proteins. The ultimate expressions of gene function are the formation of structural and regulatory RNA molecules and proteins. These macromolecules carry out the biochemical reactions and provide the structural elements that make up cells. *See* RETROVIRUS.

Flow of genetic information. The goal of molecular biology as it applies to genes is to understand the function, expression, and regulation of a gene in terms of its DNA or RNA sequence. The genetic information in genes that encode proteins is first transcribed from one strand of DNA into a complementary mRNA molecule by the action of the RNA polymerase enzyme. The genetic code is nearly universal for all prokaryotic, archaeal, and eukaryotic organisms. Many kinds of eukaryotic and a limited number of prokaryotic mRNA molecules are further processed by splicing, which removes intervening sequences called introns. In some eukaryotic mRNA molecules, certain bases are also changed posttranscriptionally by a process called RNA editing. The genetic code in the resulting mRNA molecules is translated into proteins with specific amino acid sequences by the action of the translation apparatus, consisting of transfer RNA (tRNA) molecules, ribosomes, and many other proteins. The genetic code in an mRNA molecule is the correspondence of three contiguous (triplet) bases, called a codon, to the common amino acids and translation stop signals (see **table**); the bases are adenine (A), uracil (U), guanine (G), and cytosine (C). There are 61 codons that specify the 20 common amino acids, and 3 codons that lead to translation stopping; hence the genetic code is degenerate, because certain amino acids are specified by more than one codon. *See* INTRON.

The sequence of amino acids in a protein is determined by the series of codons starting from a fixed translation initiation codon. AUG and GUG are the

Genetic code, showing the correspondence between triplet codons in mRNA and the 20 common (L-isomer) amino acids inserted into polypeptides by the translation apparatus

First position in codon	Second position in codon				Third position in codon
	U	C	A	G	
U	Phenylalanine	Serine	Tyrosine	Cysteine	U
	Phenylalanine	Serine	Tyrosine	Cysteine	C
	Leucine	Serine	Translation stop	Translation stop	A
	Leucine	Serine	Translation stop	Tryptophan	G
C	Leucine	Proline	Histidine	Arginine	U
	Leucine	Proline	Histidine	Arginine	C
	Leucine	Proline	Glutamine	Arginine	A
	Leucine	Proline	Glutamine	Arginine	G
A	Isoleucine	Threonine	Asparigine	Serine	U
	Isoleucine	Threonine	Asparigine	Serine	C
	Isoleucine	Threonine	Lysine	Arginine	A
	Methionine	Threonine	Lysine	Arginine	G
G	Valine	Alanine	Aspartate	Glycine	U
	Valine	Alanine	Aspartate	Glycine	C
	Valine	Alanine	Glutamate	Glycine	A
	Valine	Alanine	Glutamate	Glycine	G

major translation start codons of prokaryotic genes. AUG is almost always the translation start codon of eukaryotic genes. The bacterial start AUG and GUG codons specify a modified form of methionine, whereas AUG or GUG codons internal to reading frames specify methionine or valine, respectively. *See* GENETIC CODE.

The translation apparatus reads the next codon in the mRNA and attaches the specified amino acid onto methionine through a peptide bond. In most cases, this linear process of moving to the next codon and attaching the corresponding amino acid continues until one of the translation stop codons is encountered. Meanwhile, the nascent polypeptide chain folds by itself, or with the assistance of proteins called chaperones, into the functional protein. The biochemical rules that govern protein folding are often referred to as the second genetic code. In addition, some eukaryotic and prokaryotic proteins undergo protein splicing, which removes internal polypeptide segments called inteins. Some RNA transcripts are not translated; instead, they are cut and processed to form structural RNA molecules, such as tRNA and the three large RNA molecules associated with proteins in the ribosomes. *See* PROTEIN; RIBONUCLEIC ACID (RNA).

Isolating genes. In many cases, only genes that mediate a specific cellular or viral function are isolated. The recombinant DNA methods used to isolate a gene vary widely depending on the experimental system, and genes from RNA genomes must be converted into a corresponding DNA molecule by biochemical manipulation using the enzyme reverse transcriptase. The isolation of the gene is referred to as cloning, and allows large quantities of DNA corresponding to a gene of interest to be isolated and manipulated.

After the gene is isolated, the sequence of the nucleotide bases can be determined. The goal of the large-scale Human Genome Project is to sequence all the genes of several model organisms and humans. The sequence of the region containing the gene can reveal numerous features. If a gene is thought to encode a protein molecule, the genetic code can be applied to the sequence of bases determined from the cloned DNA. The application of the genetic code is done automatically by computer programs, which can identify the sequence of contiguous amino acids of the protein molecule encoded by the gene. If the function of a gene is unknown, comparisons of its nucleic acid or predicted amino acid sequence with the contents of huge international databases can often identify genes or proteins with analogous or related functions. These databases contain all the known sequences from many prokaryotic, archaeal, and eukaryotic organisms. Putative regulatory and transcript-processing sites can also be identified by computer. These putative sites, called consensus sequences, have been shown to play roles in the regulation and expression of groups of prokaryotic, archaeal, or eukaryotic genes. However, computer predictions are just a guide and not a substitute for analyzing expression and regulation by direct experimentation. *See* HUMAN GENOME PROJECT; MOLECULAR BIOLOGY.

Prokaryotic gene. In eubacteria, cyanobacteria (blue-green algae), and many bacteriophages, the genetic material is double-stranded DNA. However, in some bacteriophages the genetic material is single-stranded DNA or even RNA. As in other organisms, bacterial genes specify structural and regulatory RNA molecules, which do not encode proteins, or mRNA molecules, which do encode proteins. Some sites that play important cellular roles but are not copied into RNA molecules are also considered genes, such as the origin for bacterial chromosome replication. *See* BACTERIOPHAGE; CYANOBACTERIA.

Structure and expression. The arrangement of prokaryotic genes varies from simple to complex. Bacterial genes are delineated by sites in the DNA of the bacterial chromosome and in the RNA of transcripts. Transcribed genes start with promoters, which are binding sites of RNA polymerase. RNA polymerase melts the DNA near one end of each promoter region and then begins copying one of the DNA strands into an RNA molecule with a complementary sequence of bases. The starting end of bacterial RNA molecules initially contains a triphosphate group, which is not modified or capped as in eukaryotic cells. The first base is usually a purine (adenine or guanine). As RNA polymerase adds nucleotides to a growing RNA chain, it moves down the DNA molecule in one direction only until it encounters a signal to terminate the transcription process. Promoters face in both directions around the bacterial chromosome, and the orientation within promoters directs RNA polymerase one way or the other. Bacterial genes do not usually overlap extensively, and only one DNA strand is transcribed in most regions of bacterial chromosomes.

Genes that specify structural RNA molecules, such as tRNA or ribosomal RNA (rRNA) required for protein synthesis, are initially transcribed into long precursor molecules. Mature structural RNA molecules are cut from these precursor transcripts by the concerted activities of specific ribonucleases, which are enzymes that break the chemical bonds in the phosphodiester backbone of RNA. For genes that specify proteins, the nascent mRNA is translated as soon as a ribosome-binding site clears the transcribing RNA polymerase. Simultaneous synthesis and translation of mRNA molecules is a fundamental property of prokaryotic cells, because they lack the nuclear boundary that separates transcription from translation in eukaryotic cells. Segments of mRNA between the transcript beginning and the first ribosome-binding site are called leader regions and are often shorter than 100 nucleotides; however, longer leaders are found that play roles in gene regulation.

The ribosome binding site consists of two parts. The translation start codon of bacterial genes is usually AUG or GUG, which specifies a modified form of the amino acid methionine that is often removed from the final protein product. However, not every AUG or GUG in a transcript directs ribosome binding.

AUG and GUG start codons in mRNA are preceded by another short segment of nucleotides, called the Shine-Delgarno sequence. The bases in the Shine-Delgarno sequence pair with complementary bases in 16S rRNA molecules in the ribosome and properly position the start codon for translation. *See* RIBOSOMES.

Following initiation, the ribosome usually moves down the mRNA, reading one triplet (three-base) codon at a time. The amino acid corresponding to each codon in the genetic code is attached to the preceding amino acid by formation of a peptide bond. The ribosome continues synthesis of the polypeptide chain until one of the three translation stop codons is encountered, and the polypeptide and mRNA are released from the ribosome. The polypeptide chain is folded and sometimes binds to other folded polypeptides to form an enzymatically or structurally active protein. Some time after synthesizing the translation stop codon, RNA polymerase encounters a signal to stop transcription and to release the mRNA and DNA from the enzyme. Two kinds of transcription stop signals are used in bacteria. Factor-independent termination involves formation of a folded structure preceding a run of uracil residues in the nascent RNA chain. Factor-dependent termination involves interaction between a protein called Rho and RNA polymerase.

Expression level and regulation. Bacteria are ideally suited for survival as single cells in many environmental conditions. Part of this survivability involves controlling the expression of genes to optimize bacterial metabolism in response to environmental changes. For example, when a bacterium is presented with the amino acid tryptophan, a series of regulatory events are set off that turn off the genes encoding the enzymes that synthesize tryptophan. However, lack of tryptophan triggers the synthesis of these biosynthetic enzymes.

The expression level and regulation of a bacterial gene is influenced by five different processes. First, expression level depends on how frequently RNA polymerase transcribes a bacterial gene. The rate of transcription initiation depends on the relative intrinsic strength of a gene's promoter and whether transcription initiation is activated or repressed by additional protein regulatory factors that bind to DNA at or near a promoter in response to environmental factors. In addition, bacterial cells contain several kinds of RNA polymerase molecules that recognize different DNA consensus sequences as promoters in response to changing environmental conditions. Second, gene expression level depends on whether a transcribing RNA polymerase encounters a transcription termination signal called an attenuator that precedes the translated regions of certain genes. The frequency of termination at attenuator sites is sometimes controlled by environmental factors. Third, gene expression level depends on the stability of an mRNA molecule, because the longer its lifetime, the more frequently an mRNA species can be translated. Generally, bacterial mRNA molecules have chemical half-lives of only about 2 min, which is short compared to most eukaryotic mRNA. Fourth, gene expression level depends on the efficiency of translation of a given mRNA molecule. This efficiency depends on the relative intrinsic strength of a ribosome binding site and whether access of ribosomes to a binding site is regulated by protein factors or folded structures in the mRNA transcript. Finally, gene expression level depends on the relative stability of the gene product RNA or protein molecules. Stable RNA and protein molecules will accumulate in cells, compared to ones that are rapidly degraded. *See* BACTERIAL GENETICS; PROKARYOTAE. Malcolm E. Winkler

Eukaryotic gene. Eukaryotic genes are arranged in a linear array on one set of chromosomes (haploid germ cells) or two sets of chromosomes (diploid somatic cells). There are about 100,000 genes in the mammalian genome, located within chromatin at specific sites in the nucleus. Eukaryotic organelles (such as mitochondria or chloroplasts) also contain genomes that encode a much smaller number of proteins.

Transcription. In eukaryotes, transcription and translation are compartmentalized, respectively, in the nucleus and in the cytoplasm. Eukaryotic genes are transcribed by three different RNA polymerases that use the ribonucleotide triphosphates as substrates and the DNA as the template: polymerase I transcribes the larger rRNA genes; polymerase II generates mRNA by transcribing genes with open reading frames that encode proteins (the enzyme also generates certain small, nuclear RNAs that complex with nuclear proteins and play important roles in splicing); polymerase III transcribes the genes for tRNA and small rRNA as well as those for other small nuclear RNAs of mostly unknown functions. The initiation of transcription of mRNA takes place in eukaryotic genes at multiple sites, usually 30–100 base pairs downstream from a short adenine thymine–rich sequence referred to as the TATA box.

Eukaryotic genes are usually mosaics of coding (exons) and noncoding (introns) sequences. Since the pattern of removal of the introns (splicing) from the transcripts of a single gene may vary, a gene can encode several related proteins. When this happens, it is at variance with the long-standing one-gene-one-protein hypothesis. Mature, spliced mRNA (and in rare cases proteins) generates encoded proteins that lack a colinearity between the sequence of the DNA and its encoded protein. *See* EXON; TRANSPOSONS.

The polymerase II transcript is posttranscriptionally modified by capping of its 5′ end, by cleavage and polyadenylation of the 3′ end, and by splicing before the mature mRNA is transported to the cytoplasm to be translated.

Translation. Translation of eukaryotic mRNAs almost always initiates at the first ATG codon, probably controlled by ribosomes and associated initiation factors that scan the mRNA from its 5′ end. Prokaryotic protein initiation takes place at internal ATGs, depending on the location of the Shine-Delgarno sequence complementary to a rRNA sequence.

Operons have not been detected in eukaryotes. Genes that are not contiguous but coordinately

regulated are members of regulons. Both eukaryotes and prokaryotes (for example, *GAL* in yeast and *arg* in *Escherichia coli*) have some of their genes organized in this manner. The expression of some genes is inducible, responding to specific changes in cell environment; others are constitutively expressed, at a uniform level. Genes that carry out specific functions may be expressed in a tissue-specific manner and regulated developmentally.

The regulation of eukaryotic genes with open reading frames is more complex than that of prokaryotic genes. Several cis-acting upstream elements (called promoters) are involved; these serve as binding sites for multiple positive and negative trans-acting gene-specific and general transcription factors. These factors, together with cofactors, interact with polymerase II to regulate its activity. Also, the regulation of some eukaryotic genes may be controlled by chromatin structure and by DNA upstream or downstream enhancer sequences that act independently of their distance (some located up to several thousand base pairs away) and/or orientation. *See* EUKARYOTAE; GENETIC CODE; GENETIC ENGINEERING; GENETICS. Julios Marmur

Bibliography. E. A. Birge, *Bacterial and Bacteriophage Genetics: An Introduction*, 4th ed., 2000; M. Karin (ed.), *Gene Expression: General and Cell-Type-Specific*, 1993; B. Lewin, *Genes V*, 1994; S. R. Maloy, J. E. Cronan, Jr., and D. Freifelder, *Microbial Genetics*, 1994; V. A. McKusick, *Mendelian Inheritance in Man: Catalogs of Autosomal Dominant, Autosomal Recessive, and X-Linked Phenotypes*, 2 vols., 12th ed., 1998; P. J. Russell, *Fundamentals of Genetics*, 1994; J. D. Watson et al., *Molecular Biology of the Gene*, 1997; E. Wingender, *Gene Regulation in Eukaryotes*, 1993.

Gene amplification

The process by which a cell specifically increases the copy number of a particular gene to a greater extent than it increases the copy number of genes composing the remainder of the genome (all the genes which make up the genetic machinery of an organism). It is therefore distinguished from duplication, which is a precise doubling of the genome preparatory to cell division, and endoreduplication, which leads to endopolyploidy.

Gene amplification results from the repeated replication of the deoxyribonucleic acid (DNA) in a limited portion of the genome, in the absence of or to a much greater extent than replication of DNA composing the remainder of the genome. Thus is formed a cell in which the genes composing a limited portion of the genome are present in relatively high copy number, while the genes composing the remainder of the genome are present in approximately normal copy number. *See* DEOXYRIBONUCLEIC ACID (DNA).

Genes coding for ribosomal RNA. Ribosomes are the site of cellular protein synthesis. They are particles which are composed of a specific type of ribonucleic acid (RNA), ribosomal RNA (rRNA), and some 30 specific proteins called ribosomal proteins. The genes coding for rRNA are multicopy genes (there are many identical or nearly identical copies in the genome of most organisms).

Gene amplification was first described in the oocytes of certain amphibians. The frog oocyte contains approximately 400,000 times as much rRNA as a normal frog liver cell in spite of the fact that the amount of DNA in the oocyte nucleus is approximately the same as that of the liver cell. It has been estimated that, by using the normal genome number of rRNA genes, it would take the frog oocyte approximately 1000 years to synthesize the amount of rRNA found in the oocyte. By amplifying these genes approximately to the level known to occur (1000- to 2000-fold), the oocyte can synthesize this amount of rRNA in 6–9 months.

The developmental significance of such amplification lies in the fact that no rRNA is synthesized during early embryogenesis in amphibians. Proteins synthesized during early embryogenesis are synthesized on maternal ribosomes that were produced by the developing oocyte. Amplification of rDNA in the amphibian oocyte occurs extrachromosomally by a rolling-circle mechanism of DNA synthesis similar to that which occurs when certain bacteriophage viruses replicate their DNA.

Amplification of rDNA also occurs in the macronucleus (the vegetative nucleus) of certain ciliated protozoans. The macronucleus grows by polyploidization, but during macronuclear growth the genes coding for rRNA are replicated to a far greater extent than is the DNA composing the remainder of the genome; that is, the rRNA genes are amplified. *See* RIBONUCLEIC ACID (RNA); RIBOSOMES.

Single-copy genes. Since gene amplification increases the copy number of a specific region of the genome without altering the copy number of genes composing the remainder of the genome, it would appear to offer an alternative method for developmental control of gene expression. By increasing the number of copies of a particular gene, the number of gene copies available for transcription could thereby be increased.

In the ovary of insects, follicle cells produce the proteinaceous egg shell (or chorion). In the ovary of the fruit fly, *Drosophila melanogaster*, the chorion is produced during the final 5 h of egg-chamber development. Production of the chorion involves the synthesis of messenger RNAs (mRNAs) for several chorion proteins within a very brief period. The genes coding for the various chorion proteins are clustered within the genome, each gene present in single-copy number in germ-line cells. Amplification results in a 15- to 50-fold increase in the number of chorion gene copies per haploid genome. Chorion gene amplification appears to enable the genes to produce the amounts of mRNA necessary to build the egg shell.

Although the mechanism of chorion gene amplification has not been completely elucidated, the data suggest that each chorion gene cluster contains a

specific origin of DNA replication. Multiple rounds of replication beginning at this origin give rise to a multiforked structure in which one strand branches into two, each of these branch again and again, and the branched structures contain copies of the chorion gene.

Another single-copy gene which appears to be amplified during development is one of the genes coding for the muscle protein actin during development of muscle cells (myogenesis) in the chick. The actin genes constitute a small family of genes which are present in single or very low copy numbers in the germ line. During myogenesis (formation of skeletal muscle cells), the actin genes coding for skeletal muscle cell actin are amplified approximately 85-fold.

Of course, not all single-copy genes which are responsible for the synthesis of a large amount of cell product at a particular stage of development are amplified. The silk gland of the larval silk moth, *Bombyx mori*, for instance, produces copious amounts of silk fibroin protein. Analysis of DNA from silk glands indicates that the silk fibroin genes are not amplified above the level of ploidy expressed by the silk glands. Similarly, amplification of globin genes is not a factor in the ability of the reticulocyte (precursor cell of the red blood cell) to synthesize large amounts of the protein hemoglobin.

Other examples. In the above-mentioned examples of gene amplification, the amplification phenomenon appears to be developmentally regulated, and the amplified copies of the gene are subsequently lost from the cell. Studies on cells in culture have demonstrated "amplification" of genes involved in resistance to specific drugs. Resistance to the chemotherapeutic drug methotrexate by cultured mouse cells is associated with amplification of the dihydrofolate reductase gene. In this case the amplified genes are passed to daughter cells at the time of cell division. Such gene amplification can be either retained by cells or lost from cells when they are grown in the absence of methotrexate. *See* DRUG RESISTANCE.

Polymerase chain reaction. The polymerase chain reaction (PCR) is a technique that amplifies DNA sequences in laboratory cultures. Beginning with a sample that may contain only one DNA molecule, a selected sequence within that DNA molecule can be amplified millions or even billions of times.

To amplify a given segment of DNA, its entire sequence need not be known, but only short sequences near the ends of the segment. Short oligonucleotides, complementary to the end sequences and designed so that their 3′ ends orient toward the interior of the segment, are synthesized by conventional chemical means. The DNA in the sample is denatured by heating and then is cooled in the presence of the oligonucleotides, which anneal to the separated DNA strands. The annealed oligonucleotides are then utilized by DNA polymerase as primers for DNA synthesis, replicating the target DNA segment. These simple steps are repeated over and over again, doubling the amount of the target segment in each cycle (see **illus.**). The use of thermostable DNA polymerase (isolated from bacteria that live naturally in hot springs) is an important refinement. Since the polymerase is not denatured when the DNA is heated, it need not be added at every cycle. Amplification of DNA by the polymerase chain reaction is carried out conveniently by inexpensive machines that produce the amplified product DNA in a few hours. Target sequences between a few hundred and 10,000 base pairs in length are readily amplified. After 20 cycles, the target DNA is amplified more than a millionfold; after 30 cycles the amplification is over a billionfold.

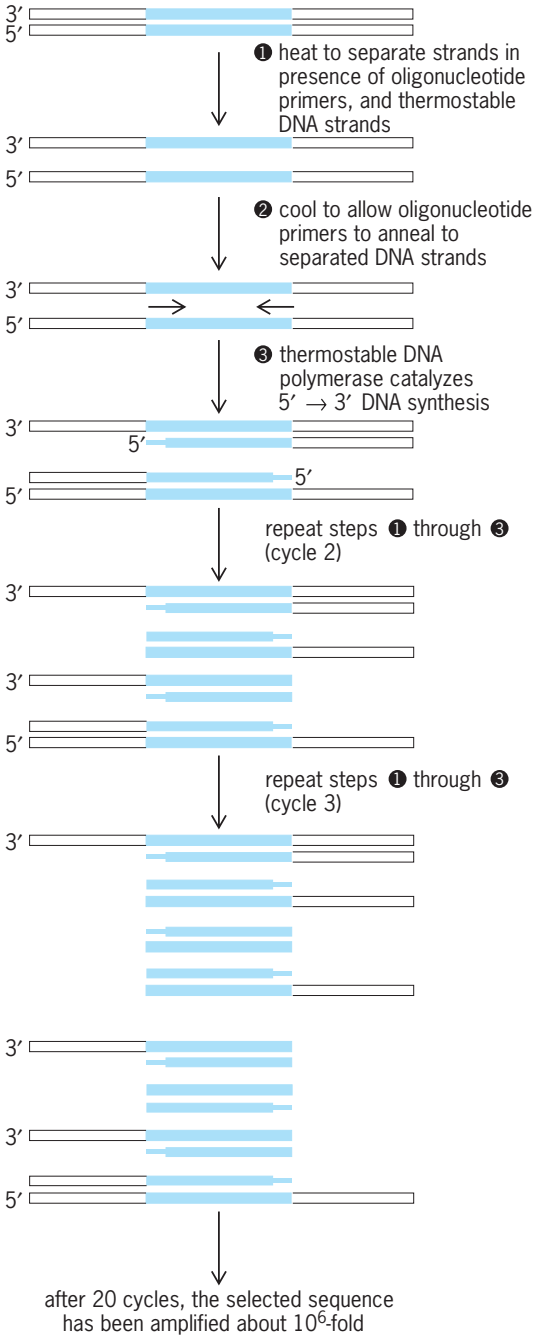

Amplification of DNA using the polymerase chain reaction. The colored region is the target DNA to be amplified. The shorter arrows represent the oligonucleotide primers.

The utility of polymerase chain reaction can be illustrated with a few examples. (1) For molecular biologists, this technique simplified the cloning of known genes, and led to the development of powerful methods to alter DNA sequences in laboratory cultures and detect alterations in DNA in organisms. (2) Certain parts of the human genome are sufficiently variable that they can be used to identify a particular individual unambiguously. With the aid of the polymerase chain reaction, DNA from samples as small as a single human hair can be amplified and analyzed to place an individual at the scene of a crime. Similar techniques are used to accurately establish paternity. Human DNA extracted frm centuries-old burial sites and amplified by this technique is being used to trace prehistoric human migrations. (3) As the molecular genetic basis for more genetic diseases becomes known, polymerase chain reaction–based diagnostic tests for these diseases are possible. For example, prenatal diagnosis of genetic diseases can be carried out with minute samples of fetal tissue. (4) Polymerase chain reaction has also been used to amplify and study DNA from extinct species such as the quagga and the woolly mammoth. In addition, very small amounts of DNA added to the cargo hold of an oil tanker can be used to trace the origins of an oil spill. *See* FORENSIC MEDICINE; GENE; GENETIC ENGINEERING; MOLECULAR BIOLOGY; PRENATAL DIAGNOSIS.

Michael M. Cox

Bibliography. W. Bloch, A biochemical perspective of the polymerase chain reaction, *Biochemistry*, 30(11):2635–2647, 1991; H. A. Erlich and N. Arnheim, Genetic analysis using the polymerase chain reaction, *Annu. Rev. Genet.*, 26:479–506, 1992; S. Paabo, R. G. Higuchi, and A. C. Wilson, Ancient DNA and the polymerase chain reaction: The emerging field of molecular archaeology, *J. Biol. Chem.*, 264(17):9709–9712, 1989; R. Schimke (ed.), *Gene Amplification*, Cold Spring Harbor Laboratory, 1982; A. C. Spradling and A. P. Mahowald, Amplification of genes for chorion proteins during oogenesis in *Drosophila melanogaster*, *Proc. Nat. Acad. Sci. U.S.A.*, 77:1096–1100, 1980; R. J. Steffan and R. M. Atlas, Polymerase chain reaction: Applications in environmental microbiology, *Annu. Rev. Microbiol.*, 45:137–161, 1991.

General aviation

All aircraft activity not associated with major airlines or the military. Among all classifications of aviation in the United States, general aviation consists of the largest number of aircraft and pilots and accounts for the largest number of flying hours.

General aviation is an integral part of the transportation system of the United States, serving both business and pleasure travel. Business travel, which accounts for more flying hours than any other branch of general aviation, employs aircraft ranging in size from single-engine vehicles flown by one person to jet equipment with a professional crew of two. Approximately 70% of all general aviation is associated with some commercial activity, such as business travel, construction, aerial application of fertilizer and pesticides, or flight instruction. Except for business jets, the average general aviation aircraft is much slower than the typical airliner, but the estimated 4.28×10^9 mi (6.89×10^9 km) flown by general aviation in 1986 exceeded by over 1×10^8 mi (1.6×10^8 km) the distance flown by scheduled carriers. General aviation departures numbered 4.76×10^7 during 1986, more than five times the number for the major airlines.

Aircraft. A general aviation aircraft can range in size from the smallest pleasure craft designed and built by an individual to the largest commercial air transport. Surplus military aircraft owned by individuals or businesses are also included in the category of general aviation, a classification based on use rather than size or original purpose.

Of the more than 220,000 active general aviation aircraft registered by the Federal Aviation Administration (FAA), almost 80% are single-engine vehicles powered by intermittent-combustion engines. Although many of these aircraft carry fewer than three passengers at speeds below 150 mi/h (240 km/h), most can carry four to six passengers at speeds up to 200 mi/h (320 km/h). Approximately 11% of the general aviation fleet are twin-engine aircraft, powered by intermittent-combustion engines and capable of cruising at speeds of 180–250 mi/h (290–400 km/h) with six to ten passengers. *See* AIRPLANE; RECIPROCATING AIRCRAFT ENGINE.

In addition, approximately 6000 multiengine aircraft are powered by turboprop engines and 4500 by turbojet or turbofan power plants. Jet aircraft employed within general aviation usually carry from five to fifteen passengers, depending on the model, with a crew of two pilots. Turboprop aircraft may also be flown by two pilots, although this is not required for all models; passenger capacity typically ranges from six to nine. Cruising speeds for a turboprop aircraft are about 300 mi/h (480 km/h), but some can exceed 400 mi/h (640 km/h) and one model can reach 460 mi/h (740 km/h) in level flight. Turboprop aircraft usually fly at altitudes between 20,000 and 40,000 ft (6000 and 12,000 m). Jet aircraft designed for general aviation typically cruise between 430 and 520 mi/h (690 and 835 km/h) and at altitudes up to 45,000 ft (13,700 m), but some are capable of cruising at speeds of 560 mi/h (900 km/h) at 51,000 ft (15,500 m), which is above the altitude approved for all commercial airliners except the Concorde. *See* JET PROPULSION; TURBOFAN; TURBOJET; TURBOPROP.

Rotorcraft number about 7000, of which approximately 4000 are powered by turboshaft engines and the rest by intermittent-combustion engines. There are also approximately 7000 gliders in the general aviation fleet. *See* GLIDER; HELICOPTER.

Certification standards. Nearly all general aviation craft weighing more than 12,499 lb (5669 kg) are certified by the federal government under Federal Air Regulation (FAR) Part 25, the same standard applied to the large commercial transports flown by major airlines. Lighter airplanes are certified under

FAR Part 23, which is only slightly less stringent in its safety specifications. Helicopters must meet the requirements of FAR Part 27 or FAR Part 29, depending upon their weight and purpose.

Pilots. General aviation pilots may hold any federal license, such as airline transport pilot, commercial, private, student, or certified flight instructor. A license, however, is valid only in conjunction with a current medical certificate of the appropriate class. At the minimum a student license is needed to pilot an aircraft, but to carry passengers a more advanced license is required, together with the successful completion every 24 months of a biennial flight review (BFR) of the pilot's knowledge and skill.

In addition to a license, current medical certificate, and biennial flight review, pilots must possess a rating for the category and class of vehicle they wish to fly. Category refers to a broad classification of aircraft, such as airplane, rotocraft, glider, or lighter-than-air balloon. Class refers to the classification of aircraft within a category having similar operating characteristics, such as single-engine, multiengine, land, sea, helicopter, gyroplane, airship, or free balloon. To operate aircraft weighing more than 12,499 lb (5669 kg) and for all aircraft powered by a turbojet or fan-jet engine, pilots must posses a specific type rating affixed to their license.

To fly an aircraft during poor weather conditions, nominally defined as visibility less than 3.0 mi (5.0 km) and a cloud ceiling of less than 1000 ft (305 m), a pilot must possess an instrument rating or an airline transport pilot license; one of these is held by most general aviation pilots who fly professionally or for business purposes. *See* AIR-TRAFFIC CONTROL.

Limitations of scheduled airlines. The demand for business transportation by general aviation was stimulated by the oil embargoes of 1973 and 1979 and by the airline deregulation of October 1978. Fuel prices increased nearly 1000% between 1974 and 1981, and even the most fuel-efficient airliners are now cost-effective only when flown with high load factors, that is, with most seats occupied by paying passengers.

Deregulation permitted airlines to concentrate on heavily traveled routes such as those between major cities; locations where the demand for travel was low were abandoned. Accordingly the major airlines developed a system of hub cities, where flights would connect either with small commuter lines that brought passengers from outlaying locations or with larger airliners flying between major hubs. This system became known as the hub-and-spoke system because the passenger traffic appeared to flow into the hub city along spokes of a wheel, with few if any flights between points on the circumference of the wheel. For example, flying by scheduled airlines from Wichita, Kansas, to Rochester, Minnesota, required intermediate stops at Memphis, Tennessee, and Minneapolis, Minnesota. The return flight stopped in Waterloo, Iowa, and St. Louis, Missouri. Such a trip would last nearly 14 h, but a direct flight by general aviation would take only 3.5–5 h. *See* AIR TRANSPORTATION.

Business flying. Because of the characteristics of hub-and-spoke systems, business people have increased their demand for general aviation, which also offers more control over travel arrangements. Although the major airlines provide scheduled service to approximately 394 locations and unscheduled flights to an additional 312 airports, over two-thirds of all flights were between only 30 cities. This service is significantly supplemented by general aviation, which utilizes the nearly 5800 public-use landing places and many of the approximately 12,000 limited-use landing places in the United States.

With a small single-engine aircraft accommodating four or five passengers, a business person with a private license and perhaps an instrument rating can escape many of the limitations of airline schedules based on the hub-and-spoke system. Some pilots prefer small or medium-size twin-engine aircraft for business travel, particularly in night and instrument flying. Some companies own their aircraft and employ a professional crew; others prefer chartering from an aircraft taxi or management firm. Business jets that fly nonstop more than 5000 mi (8100 km) are operated by many multinational corporations, offering security as well as efficiency.

Fuel consumption. General aviation aircraft are relatively fuel-efficient. A six-passenger single-engine airplane can deliver more than 80 passenger miles per gallon of fuel (34 km/liter); typical cabin-class twin-engine aircraft generate about 50 passenger miles per gallon (21 km/liter). General aviation consumes less than 0.5% of the gasoline used in the United States and 6.5% of the turbine fuel, accounting for about 0.75% of the output of United States oil refineries and representing an efficient use of energy for transportation. *See* AIRCRAFT FUEL.

Commercial applications. Commuter and regional airlines supplement the hub-and-spoke scheduling typical of commercial air travel since deregulation. Unlike business flying, which operates under FAR Part 91, commercial activities within general aviation fall under FAR Part 135. All airlines flying aircraft with more than 30 passenger seats operate under FAR Part 121, which covers both air taxis and scheduled commuter and regional airlines. Commuter and regional airlines offer transportation over prescribed routes as specified in a published schedule, whereas air taxis fly passengers and freight on demand over random routes.

Commuter and regional airlines. With the growth of hub-and-spoke air travel, small airlines offer the only scheduled service in some communities, and often cooperate with major airlines by sharing code designations used in computer reservation systems and by travel agencies. As a result of code-sharing, which became widespread in 1984, travelers often book tickets on commuter or regional carriers without realizing it. These carriers have become more sophisticated and meet many of the same standards as the major airlines.

Commuter and regional airlines might be compared with the local service airlines in the period before deregulation, which operated under FAR Part

121 with equipment smaller than that used by the major carriers.

However, since the fuel crises of 1973 and 1979 and airline deregulation in 1978, commuter and regional air carriers have experienced the same problems as the major airlines. They must choose routes that generate high load factors. Thus, they have concentrated on hub locations and moved into routes previously flown only by the major carriers. Consequently, aircraft operated by companies or individuals often offer the quickest and most cost-effective transportation between locations that are outside the routes connecting major hubs.

As a result of these changes, there is a question as to whether commuter and regional airlines should still be classed with the general aviation, or whether they should have a category of their own, or whether they really belong with the major airlines. The same FAA field organization that monitors commuter and regional operations also audits the major airlines, and the National Transportation Safety Board now places commuter and regional airline accidents within a category separate from the commercial segment of general aviation. In the United States, however, the same manufacturing, maintenance, and personnel infrastructure supports general aviation and the commuter and regional airlines.

Air taxis. Air taxis provide on-demand air transportation for hire in a wide variety of aircraft, operating under FAR Part 135. They do not fly on a schedule over established routes. The demand for increased service is being met by operators ranging from small flying services with one four-passenger single-engine airplane for occasional charter work to large transportation companies with a fleet of business jets. Under FAR Part 91, the air taxi segment of general aviation manages and operates aircraft owned by a business or shared by more than one business.

Agriculture. General aviation is widely used in agriculture; aerial seeding, fertilizing, and spraying are efficient and widely used by farmers. Over 1.8×10^8 acres (7.3×10^7 hectares) of farmland are treated in the United States each year in this way. For example, about 90% of the United States rice crop is seeded by air. A day's work with a tractor is accomplished by an aircraft in 1 h with a 90% fuel saving. Chemicals are applied more effectively by air than by tractor. *See* AGRICULTURAL AIRCRAFT.

Flight instruction. Flight instruction represents about 15% of all general aviation activity and is the principal source of professional pilots. Most flight instruction and air taxi services are provided by fixed-base operators at an airport. A typical establishment rents aircraft and offers instruction as well as taxi or charter flying. It may also sell and service new or used equipment; it usually also sells gas and oil. The number of fixed-base operator chains and franchises has increased as the service segment of general aviation has matured.

Private flying. About 30% of general aviation involves the personal use of small aircraft for transportation and leisure purposes as well as some of the proficiency flying required by the Federal Aviation Administration of pilots who wish to maintain their licenses. Most personal flying employs small single-engine craft.

The use of gliders, sports aircraft, balloons, and home-built vehicles is considered personal unless it involves flight instruction for profit. Home-built aircraft cannot be used for any commmercial purpose and can serve only in personal flying. *See* ULTRALIGHT AIRCRAFT.

Research applications. General aviation has often been the testing ground for innovation. Most of the pioneer aviators flew equipment that could be considered the general aviation vehicles of their day. General aviation manufacturers produced the first nonmilitary aircraft to be built entirely from composite materials and approved for routine use.

Voyager. Over a 6-year period, a team of individuals led by pilots D. Rutan and J. Yeager and designer B. Rutan conceived and implemented a plan to circumnavigate the globe by air without refueling, a feat that had never been accomplished. Their aircraft, *Voyager* (**Fig. 1**), was designed and constructed for the mission. With wings longer than a Boeing 727 airliner's, *Voyager* carried nearly 9000 lb (4000 kg) of fuel in a structure that was built of advanced composite materials and weighed less than 1000 lb (450 kg) when empty. In effect, the twin-engine aircraft was a flying fuel tank with minimal space for pilot and copilot. Its fuselage, wings, and twin booms contained aviation gasoline for a liquid-cooled engine that was located in the aft fuselage and used for cruising flight and for a larger, air-cooled engine that was located in the aircraft's nose and used in combination with the rear engine for takeoff. Once in cruising flight, the forward engine was stopped and its propeller positioned in a feathered or knife-edge orientation to minimize wind resistance.

Voyager took off from Edwards Air Force Base, California, on December 14, 1986, and landed there 9 days later without refueling. In spite of challenging weather and several mechanical failures, the crew completed their global circumnavigation without mishap or physical injury.

Human-powered flight. The dream of flight using only the aviator's physical strength was unobtainable until advanced materials and computer-aided design

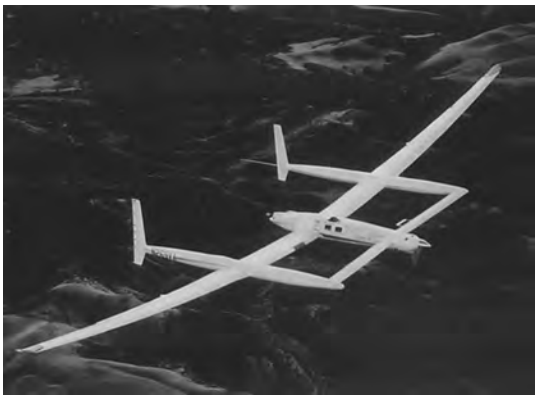

Fig. 1. ***Voyager*** **after takeoff on world flight. (*Jeffrey Vock/Visions*)**

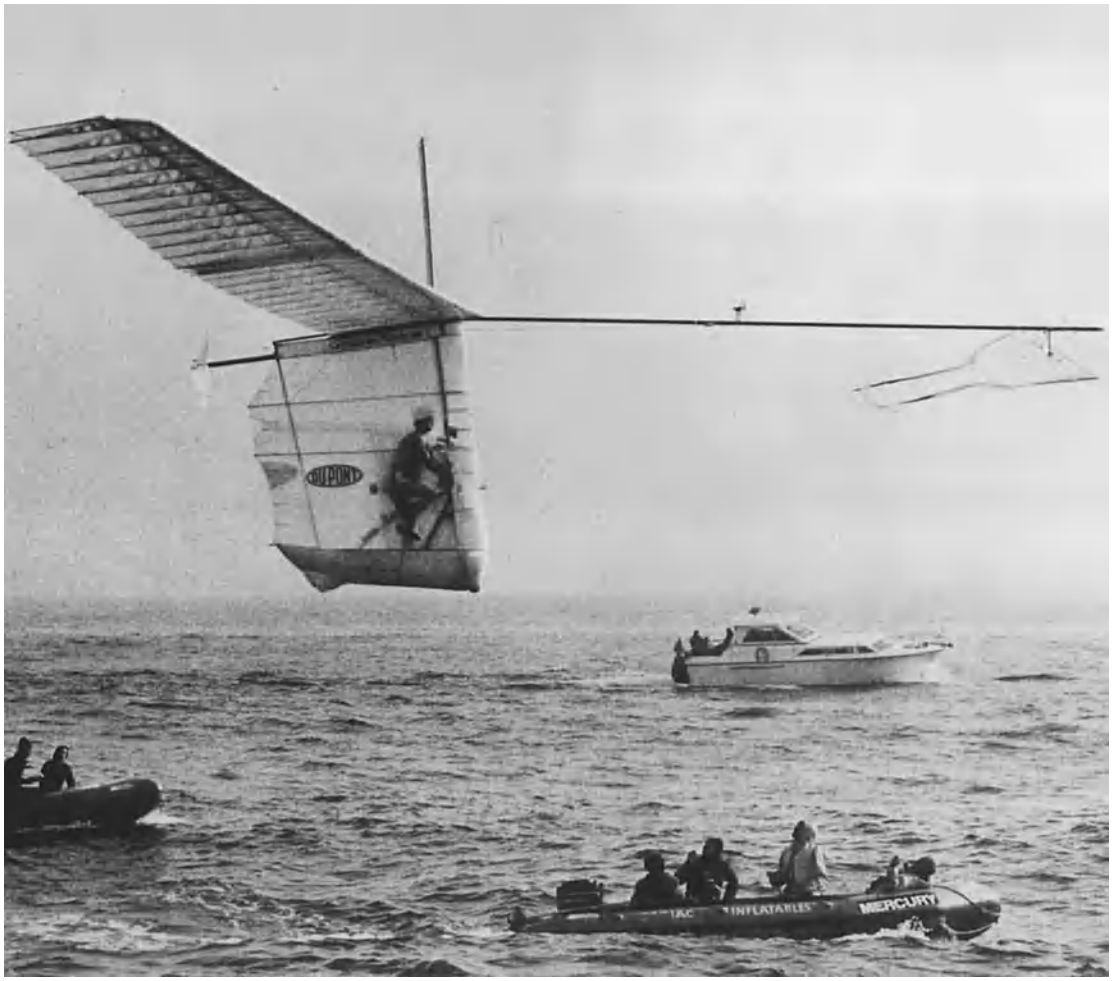

Fig. 2. Human-powered aircraft *Gossamer Albatross* in flight. (*Photograph by Don Monroe*)

provided the tools needed to design aircraft with sufficient strength-to-weight characteristics and efficiency to be powered entirely by a human. *See* COMPOSITE MATERIAL; COMPUTER-AIDED DESIGN AND MANUFACTURING.

Modern attempts to achieve human-powered flight were stimulated in 1959 when H. Kremer posted a prize for the first British aircraft to fly solely by means of human power, without the aid of power storage or buoyancy, around a figure-eight course. (The prize was later opened to all nationalities.) The aircraft had to be powered and controlled by a crew throughout its flight, and only one crewmember could assist during the takeoff. Still air, defined as a wind not to exceed 10 knots (5.1 m/s), was required for the flight in order to claim the prize.

The first practical human-powered aircraft flew in 1961 but without sufficient performance to win the Kremer prize, which went unclaimed until 1977 when it was won by the *Gossamer Condor* aircraft, which was designed by P. D. MacCready, Jr., and flown by B. Allen. The team of MacCready and Allen subsequently accomplished many significant events in human-powered flight, including the 1979 flight of the *Gossamer Albatross* (**Fig. 2**) across the English Channel.

Safety. General aviation aircraft flown in corporate operations by salaried crews have a safety record rivaling that of the major airlines. For all general aviation, including pleasure flying, the accident rate was 7.95 per 100,000 flight hours, and the fatal accident rate was 1.49 per 100,000 flight hours, in 1988. Safety has been improving steadily in general aviation. *See* AVIATION.

John W. Olcott

Bibliography. General Aviation Manufacturers Association, *General Aviation*, 1989; General Aviation Manufacturers Association, *General Aviation Statistical Databook*, 1988; *National Transportation Safety Board Annual Review of Aircraft Accident Rates*, annually; U.S. Department of Transportation, *Census of U.S. Civil Aircraft*, annually; U.S. Department of Transportation, Federal Aviation Administration, *FAA Statistical Handbook of Aviation*, annually.

Generator

A machine in which mechanical energy is converted to electrical energy. Generators are made in a wide range of sizes, from very small machines with a few watts of power output to very large central-station

generators providing 1000 MW or more. All electrical generators utilize a magnetic field to produce an output voltage which drives the current to the load. The electric current and magnetic field also interact to produce a mechanical torque opposing the motion supplied by the prime mover. The mechanical power input is equal to the electric power output plus the electrical and mechanical losses.

Generators can be divided into two groups, alternating current (ac) and direct current (dc). Each group can be subdivided into machines that use permanent magnets to produce the magnetic field (PM machines) and those using field windings. A further subdivision relates to the type of prime mover and the generator speed. Large generators are often driven by steam or hydraulic turbines, by diesel engines, and sometimes by electric motors. Generator speeds vary from several thousand rotations per minute for steam turbines to very low speeds for hydraulic or wind turbines. *See* DIESEL ENGINE; HYDRAULIC TURBINE; MOTOR; PRIME MOVER; STEAM TURBINE; WIND POWER.

Theory of operation. The production of the output voltage is based upon Faraday's law, which states that when the magnetic flux across a coil of wire changes, the resulting induced voltage is proportional to the number of turns in the coil multiplied by the rate of change of the flux, as stated in Eq. (1), where e is

$$e = N\frac{d\phi}{dt} \tag{1}$$

the instantaneous induced voltage, N the number of turns in the coil, ϕ the flux in webers, and t the time in seconds. *See* ELECTROMAGNETIC INDUCTION.

Typically, voltage is induced in the armature windings of a generator by mechanically driving one member relative to the other so that the active coils (the armature winding) move through the magnetic field. In this case the flux linking the armature coils changes because of the relative motion, and the resulting voltage is called a motional or speed voltage. When the magnetic field is constant, as is usual, the motional voltage can be expressed by Eq. (2), where B is the

$$e = 2NBlv \tag{2}$$

magnetic flux density in teslas, l the length of armature conductors in meters, and v the relative velocity of the armature conductors with respect to the magnetic field in meters per second. This form of Faraday's law is applicable to most conventional generators.

When the generator supplies an electrical load, an electric current exists in the armature windings. This current interacts with the magnetic field to produce a mechanical force f which is in oppositon to the relative motion and proportional to the current and the magnetic field, as shown by Eq. (3), where i is

$$f = 2NBli \tag{3}$$

the current in the armature conductors. The power balance in the generator can be demonstrated by observing that the electrical power given by the product of voltageand current in Eq. (4) is equal to the

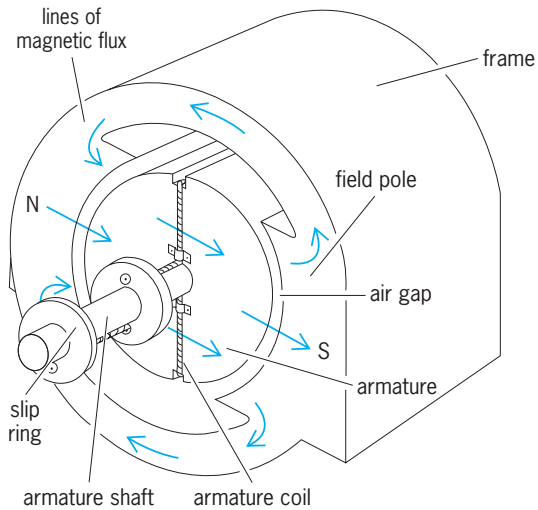

Fig. 1. Elementary generator.

$$ei = 2NBlvi \tag{4}$$

mechanical power expressed by the product of force and velocity in Eq. (5).

$$fv = 2NBliv \tag{5}$$

See MAGNETIC FIELD; POWER.

Basic design. **Figure 1** shows an elementary generator with a stationary field and a single rotating armature coil. If the field is uniformly distributed along the field pole, the voltage generated as the coil moves through the field is nearly constant over each half revolution according to Eq. (2). The polarity of the voltage reverses once per revolution as the coil sides come under the influence of first a north pole and then a south pole. The resulting rectangular-shaped voltage wave has the appearance of the wave shown in **Fig. 2***a*. If the slip rings shown in Fig. 1 are replaced by a two-segment commutator, the polarity of the output can be maintained in one direction to produce the pulsating dc voltage shown in Fig. 2*b*. *See* COMMUTATOR; DIRECT-CURRENT GENERATOR.

To produce a sinusoidal shaped ac output voltage, the field poles can be modified to produce a flux distribution in which the field strength builds up gradually to a peak at the center of the pole. If the shape of the field distribution were made sinusoidal, the resulting voltage wave would, according to Eq. (2), vary sinusoidally (Fig. 2*c*). This voltage can be extracted via the slip rings shown in Fig. 1. With additional armature coils and slip rings, polyphase voltages can be produced. *See* ALTERNATING-CURRENT GENERATOR.

Although Fig. 2 illustrates a machine with only two magnetic poles, machines with a larger, even number of poles are also possible. Since one electrical cycle is generated each time an armature coil passes two magnetic poles, the rotational speed necessary

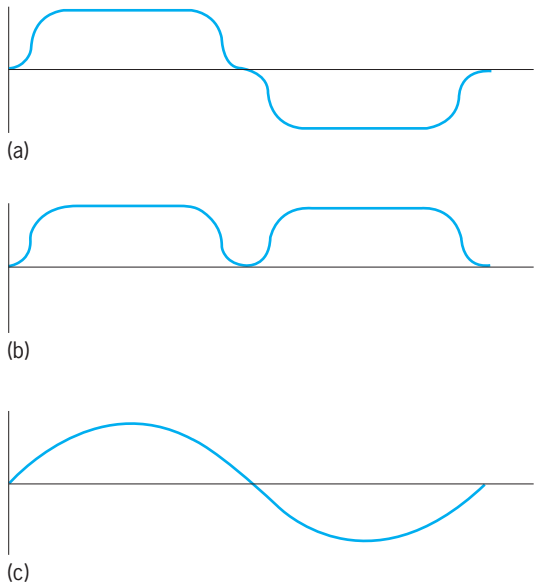

Fig. 2. Output voltage waveforms of elementary generators. (*a*) At slip-ring terminals, with uniform flux distribution. (*b*) At commutator terminals, with uniform flux distribution. (*c*) At slip-ring terminals, with sinusoidal flux distribution.

to produce a particular output frequency is given by Eq. (6), where RPM is the rotational speed in revo-

$$\text{RPM} = \frac{120f}{P} \tag{6}$$

lutions per minute (rpm), f is the frequency of the generated voltage in hertz (Hz), and P is the number of poles. An ac generator producing 60 Hz can have speeds of 3600 rpm for two poles, 1800 rpm for four poles, 1200 rpm for six poles, and so forth.

Losses, rating, and efficiency. The conversion of energy from one form to another is never a process without loss. The principal losses in generators are the mechanical losses associated with bearing friction and windage, the electrical losses resulting from the resistance of the conductors (proportional to the current squared), and the magnetic losses in the iron cores of the armature and field structures. These losses all produce heat. The allowable temperature rise of the generator system limits the total losses that can be tolerated. This, in turn, determines the rating of the generator in volt-amperes for ac generators or watts for dc generators. The rating denotes the electrical load that the generator can supply continuously without damage caused by overheating. *See* CORE LOSS; ELECTRICAL RESISTANCE; FRICTION.

The efficiency of a generator is the ratio of the electrical power output to the mechanical power input. Without power losses, the efficiency would be 100%. Large generators can reach efficiencies as high as 97% or more, but small generators are less efficient. A generator of a few hundred watts rating, such as that used to charge an automobile battery, may have an efficiency of only 60–70%. *See* EFFICIENCY.

Construction. The field structure of a generator establishes the magnetic flux needed for energy conversion. In small generators, permanent magnets can be used to provide the required magnetic field. In large machines, dc field windings are more economical and permit changes in the magnetic flux and output voltage. This allows control of the generated voltage, which is important in many applications. Electromechanical energy conversion does not take place in the field windings; the input electrical energy is all converted to heat.

In dc generators the field structure must be stationary to permit a rotating mounting for the commutator and armature windings. However, since the field windings require low voltage and power and have only two lead wires, it is convenient to place the field on the rotating member in ac generators. The field winding is more easily insulated and supported against rotational forces and is better suited to sliding contact slip rings than the higher-voltage, higher-power armature windings. Usually the field windings are mounted to encircle the protruding field poles, creating a salient pole generator (Fig. 1), but in some cases, especially for high-speed ac generators, the field windings are placed in slots in the surface of a cylindrical field structure. In larger generators a second, smaller generator called an exciter supplies the dc power required to energize the field windings. In most modern large generators the exciter is an ac generator whose output is electronically converted to dc. If the electronic rectifier components rotate with the exciter and field windings, no slip rings are needed and the combination is called a brushless exciter.

The armature and field cores are constructed of steel to minimize the field current (or size of permanent magnets) needed to establish the magnetic flux. The air gap between the cores is kept small since it is difficult to magnetize compared to the steel cores. Those portions of the core in which the magnetic flux is constant can be solid steel. This includes the field poles and outer yoke of dc generators and most of the field structure of ac generators. The armature core, which is subject to time-varying flux, is always constructed from thin sheets of high-grade electrical steel to reduce eddy-current losses. *See* EDDY CURRENT; MAGNET.

The armatures of all generators are cylindrical and do not have the protruding poles that are common for the field structure. The windings are placed in slots along the periphery of the armature structure and are electrically insulated from the core. In ac generators these slots are found along the inner surface of the stationary armature core. In dc generators the slots are found on the outer surface of the rotating armature core (Fig. 1). The slots have two important functions. First, the teeth between them provide a path for the flux. Second, most of the mechanical force which would be exerted on the windings if they were placed in the air gap transfers to the core itself. There is thus no need to support the windings to transmit the forces required in large generators. *See* WINDINGS IN ELECTRIC MACHINERY.

Power electronics in generators. The availability of power electronic converters capable of modifying the voltage and frequency of large amounts of

electric power has resulted in the appearance of a number of generator systems offering reduced cost, unique properties, or both. An example, involving millions of units, is the battery-charging automotive alternator, consisting of an ac generator and a rectifier. In this case the rectifier serves to replace the commutator of the dc generator used in earlier times. In addition to lower cost, the system offers a wider useful speed range and simpler control. On a totally different scale, most modern wind turbine generators employ an inverter to adjust the frequency to a constant 60 Hz while allowing the wind turbine speed to vary to optimize the turbine efficiency. Although small in number, the largest of these generators are in the 1–5-MW size range. Many other examples of power electronics applied to generator systems exist, and their number is expected to grow. *See* ELECTRIC POWER GENERATION; ELECTRIC ROTATING MACHINERY. Donald W. Novotny

Bibliography. A. R. Bergen and V. Vittal, *Power Systems Analysis*, 2d ed., 2000; J. J. Cathey, *Electric Machines: Analysis and Design*, 2001; S. J. Chapman, *Electric Machinery Fundamentals*, 4th ed., 2005; A. E. Fitzgerald, C. Kingsley, Jr., and S. D. Umans, *Electric Machinery*, 6th ed., 2003.

Genetic algorithms

Search procedures based on the mechanics of natural selection and genetics. Such procedures are known also as evolution strategies, evolutionary programming, genetic programming, and evolutionary computation. Originating in the cybernetics movement of the 1940s and the 1950s, genetic algorithms are increasingly solving difficult search, optimization, and machine-learning problems that have previously resisted automated solution.

Motivation. Just as natural selection and genetics have filled a variety of niches by creating genotypes (attributes of chromosomes) that result in well-adapted phenotypes (attributes of organisms), so too can genetic algorithms solve many artificial problems by creating strings (artificial chromosomes) that result in better solutions. Users ultimately turn to genetic algorithms for robustness, that is, for algorithms that are broadly applicable, relatively quick, and sufficiently reliable. This emphasis on robustness contrasts starkly with the philosophy of operations research, where new algorithms must be tailored to specific problems. The need to invent a new method for each new problem class is daunting, and users look for methods that can solve complex problems without this requirement. *See* GENETICS; OPERATIONS RESEARCH; ORGANIC EVOLUTION.

Mechanics. There are many variations on genetic and evolutionary algorithms. For concrete exposition, the discussion is limited to a simple genetic algorithm that processes a finite population of fixed-length, binary strings. In practice, bit codes, k-ary codes, real (floating-point) codes, permutation (order) codes, program codes, and others have all been used with success.

A simple genetic algorithm consists of three operators: selection, crossover, and mutation.

Selection. Selection is the survival of the fittest within the genetic algorithm. The key notion is to give preference to better individuals. For selection to function, there must be some way of determining what is good. This evaluation can come from a formal objective function, it can come from the subjective judgment of a human observer or critic, or it can come from the co-evolution of one's competitors or collaborators. There are many ways to achieve effective selection, including ranking, tournament, and proportionate schemes. For example, in two-party tournament selection, pairs of strings are drawn randomly from the parental population, and the better individual places an identical copy in the mating pool. If a whole population is selected in this manner, each individual will participate in two tournaments, and the best individual in the population will win both trials, the median individual will typically win one trial, and the worst individual wins not at all. As this example makes clear, the primary requirement of selection is for a partial ordering.

Crossover. If genetic algorithms were to do nothing but selection, the trajectory of populations could contain nothing but changing proportions of the strings in the original population. To do something more sensible, the algorithm needs to explore different structures. A primary exploration operator used in many genetic algorithms is crossover. Simple, one-point crossover begins when two individuals are chosen from the population by using the selection operator, and these two structures are considered to be mated. A cross site along the string length is chosen uniformly at random, and position values are exchanged between the two strings following the cross site. For example, starting with the two strings $A = 11111$ and $B = 00000$, if the random choice of a cross site turns up a 3, the two new strings $A' = 11100$ and $B' = 00011$ are obtained following crossover; these strings would be placed in the new population. This process continues pair by pair until the new population is complete, filled with so-called offstrings that are constructed from the bits and pieces of good (selected) parents. There are many other variants of crossover, but the main issue is whether the operator promotes the successful exchange of necessary substructures.

Selection and crossover are surprisingly simple operators, involving nothing more complex than random-number generation, string copying, and partial string exchanges, yet their combined action is responsible for much of a genetic algorithm's effectiveness. To understand this statement intuitively, it is useful to think in terms of human process of innovation. Innovation or creativity is often based on combining notions that work well in one context with notions that work well in another context to form new, possibly better ideas of how to attack the problem at hand. Similarly, genetic algorithms juxtapose many different, highly fit substrings (notions) through the combined action of selection and crossover to form new strings (ideas).

Mutation. In a binary-coded genetic algorithm, mutation is the occasional (low-probability) alteration of a bit position, and with other codes a variety of diversity-generating operators may be used. By itself, mutation induces a simple random walk through string space. When it is used with selection alone, the two combine to form a parallel, noise-tolerant, hill-climbing algorithm. When used together with selection and crossover, mutation acts both as an insurance policy against losing needed diversity and as a hill climber.

Capabilities. Genetic algorithms can be attractive for a number of reasons: they can solve hard problems quickly and reliably, are easy to interface to existing simulations and models, are extensible, and are easy to hybridize.

One of the primary reasons to use genetic algorithms is that they are efficient and effective over a broad class of problems. Empirical work has long suggested this characteristic, but theory is catching up, and it appears that genetic algorithms can quickly solve problems that have many difficult-to-find optima. There is now a considerable body of genetic algorithm theory from application design to fully rigorous nonlinear difference equations and Markov chains. Application design theory gives useful, yet simple, algebraic relations for the principled selection of population sizes and other algorithm parameters, eliminating unnecessary trial and error. This understanding when properly used to design and size scalable genetic algorithms, sometimes called competent genetic algorithms, results in procedures that appear to solve a large class of difficult problems in times that grow as a subquadratic function of the number of decisions variables in a probably approximately correct sense. *See* STOCHASTIC PROCESS.

Because genetic algorithms require very little problem-specific information, they are remarkably easy to connect to extant application code. Many algorithms require a high degree of integration between solver and objective function. For most users, the difficult part of optimization is getting a good model, and once it is tested and calibrated, the genetic algorithm can be interfaced quite directly without additional difficulty.

Even simple genetic algorithms can be broadly capable, but real problems can pose unanticipated difficulties. When these problems arise, there is often a solution from nature available to solve the problem. For example, many problems in artificial intelligence have search spaces that are highly multimodal and solution sets that may contain multiple global solutions. In these cases, it is desirable to have a population converge to multiple optima simultaneously. By borrowing the biological notions of niche and speciation, genetic algorithms can be designed to converge to a diverse set of highly fit solutions, thereby solving multimodal problems in a single run.

When no solution from nature is available, it is often possible to use problem-specific information to help make a hybrid or knowledge-augmented genetic algorithm. Effective hybrids can be constructed by combining the global perspective of the genetic algorithm with the local efficiency of a problem-specific heuristic. There are also a number of ways that problem-specific information can be built into the operators or the codings. Hybrids are one effective means of enhancing the efficiency of genetic algorithms. Researchers have been exploring effective parallelization, time utilization, and evaluation relaxation as well. Applicable design theory may be used in each of these areas to construct principled approaches to acceptable solutions in a reasonable time.

Applications. Applications have fallen into one of two categories: search and optimization, and machine learning. Genetic algorithms have also inspired other forms of computation and novel scientific approaches.

Search and optimization. Some of the notable commercial applications of genetic algorithms in search and optimization include gas and steam turbine design, expansion of fiber-optic communications networks, and automatic drawings of criminal suspect faces from witness recollection. Prospective applications include financial time-series prediction, protein folding, aerospace composite structural design, and integrated circuit design. Two important trends in the application of genetic algorithms are the use of multiobjective and interactive genetic algorithms. Multiobjective genetic algorithms exploit the existence of a population to search for nondominated (Pareto) solutions along two or more objectives. Interactive genetic algorithms combine human judgment with computerized innovation to permit the solution of problems where the creation of an objective function is difficult or impossible. Interactive genetic algorithms have been used to create pleasing art, music, and advertising, to name a few examples. *See* OPTIMIZATION.

Machine learning. Some of the more interesting, and difficult, applications of the ideas of evolutionary computation come under the heading of genetics-based machine learning (GBML). Perhaps the most ambitious of GBML systems is the learning classifier system, a type of learning expert system that combines a primitive expert system shell, genetic and other learning algorithms, and a means of apportioning credit to alternative rules. *See* EXPERT SYSTEMS.

A more direct approach to solving GBML problems has arisen in genetic programming systems. In genetic programming, modified simple genetic algorithms are used to recombine computer programs, which are then rated in their ability to solve some problem. The more highly fit programs are recombined with others to form new, possibly better programs. A broad array of problems have been tackled by using the methods of genetic programming. Such progress is not surprising since natural genotypes are dynamic structures more akin to computer programs than to static sets of decision-variable values. *See* PROGRAMMING LANGUAGES.

Computational and scientific influences. Genetic algorithms have been instrumental for creating other forms of biologically inspired computation. For example, ant colony optimization, artificial immune

systems, and particle swarm optimization use different metaphors as motivation, but their population orientation and detailed mechanics have much in common with, and have drawn heavily from, evolutionary computation.

Genetic algorithms have also influenced a number of important scientific directions. Agent-based computational economics, artificial life, and other forms of complex adaptive systems inquiry draw heavily from the example, theory, and lore of genetic algorithms. These approaches now are being used to create computational laboratories in scientific fields where experimentation is otherwise difficult. *See* ALGORITHM; ARTIFICIAL INTELLIGENCE; INTELLIGENT MACHINE.

David E. Goldberg

Bibliography. T. Baeck, D. B. Fogel, and M. Michalewicz (eds.), *Handbook of Evolutionary Computation*, 1997; K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*, 2001; D. E. Goldberg, *The Design of Innovation: Lessons from and for Competent Genetic Algorithms*, 2002; D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, 1989; J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, 1992; J. H. Holland, *Hidden Order: How Adaptation Builds Complexity*, 1996; C. L. Karr and L. M. Freeman (eds.), *Industrial Applications of Genetic Algorithms*, 1998; J. Koza et al., *Genetic Programming 3: Darwinian Invention and Problem Solving*, 1999; W. B. Langdon et al., *Foundations of Genetic Programming*, 2001.

Genetic code

The rules by which the base sequences of deoxyribonucleic acid (DNA) are translated into the amino acid sequences of proteins. Each sequence of DNA that codes for a protein is transcribed or copied (**Fig. 1**) into messenger ribonucleic acid (mRNA). Following the rules of the code, discrete elements in the mRNA, known as codons, specify each of the 20 different amino acids that are the constituents of proteins. In a process called translation, the cell decodes the message in mRNA. During translation (**Fig. 2**), another class of RNAs, called transfer RNAs (tRNAs), are coupled to amino acids, bind to the mRNA, and in a step-by-step fashion provide the amino acids that are linked together in the order called for by the mRNA sequence. The specific attachment of each amino acid to the appropriate tRNA, and the precise pairing of tRNAs via their anticodons to the correct codons in the mRNA, form the basis of the genetic code. *See* DEOXYRIBONUCLEIC ACID (DNA); PROTEIN; RIBONUCLEIC ACID (RNA).

Universal genetic code. The genetic information in DNA is found in the sequence or order of four bases that are linked together to form each strand of the two-stranded DNA molecule. The bases of DNA are adenine, guanine, thymine, and cytosine, which are abbreviated A, G, T, and C. Chemically, A and G

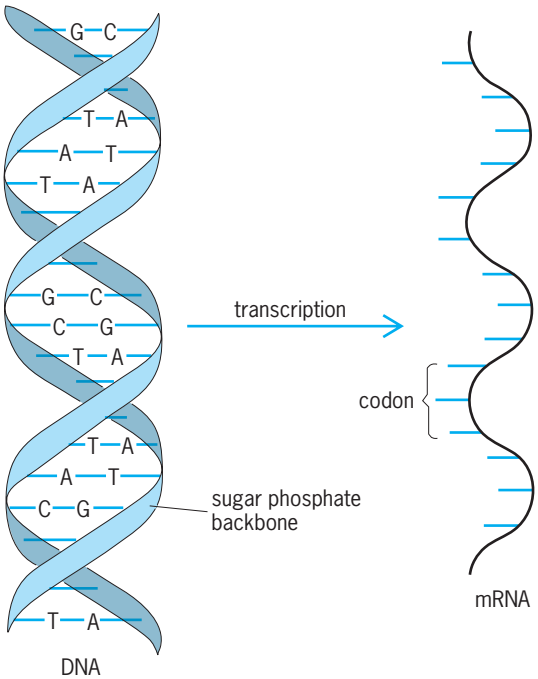

Fig. 1. DNA double helix and mRNA copy made from it by transcription.

are purines, and C and T are pyrimidines. The two strands of DNA are wound about each other in a double helix that looks like a twisted ladder (Fig. 1). The two posts of the ladder consist of sugar (deoxyribose) and phosphate chains that link the bases together in a directional manner. In DNA, the two strands are said to be antiparallel, because they run in opposite directions. Each rung of the ladder is formed by two bases, one from each strand, that pair with each other by means of hydrogen bonds. For a good fit, a pyrimidine must pair with a purine; in DNA, A bonds with T, and G bonds with C. *See* PURINE; PYRIMIDINE.

Ribonucleic acids such as mRNA or tRNA also comprise four bases, except that in RNA the pyrimidine uracil (U) replaces thymine. During transcription a

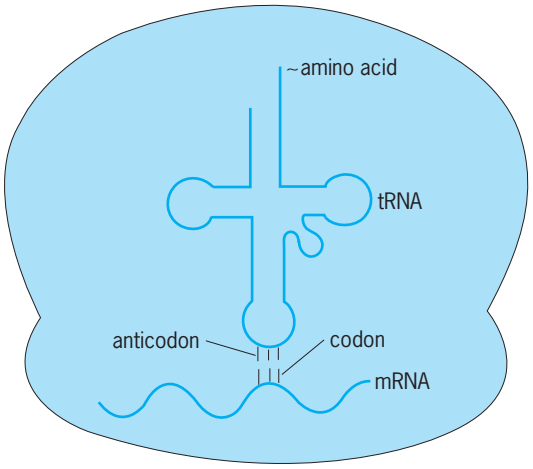

Fig. 2. Translation process in the ribosome, showing the tRNA with attached amino acid interacting with the mRNA via the codon-anticodon interaction.

single-stranded mRNA copy of one strand of the DNA is made. That mRNA copy of the DNA is bound to a large particle known as the ribosome and decoded according to the rules of the genetic code (Figs. 1 and 2). *See* RIBOSOMES.

If two bases at a time are grouped together, then only 4×4 or 16 different combinations are possible, a number that is insufficient to code for all 20 amino acids that are found in proteins. However, if the four bases are grouped together in threes, then there are $4 \times 4 \times 4$ or 64 different combinations. Read sequentially without overlapping, those groups of three bases constitute a codon, the unit that codes for a single amino acid.

The 64 codons can be divided into 16 families of four (**Fig. 3**), in which each codon begins with the same two bases. With the number of codons exceeding the number of amino acids, several codons can code for the same amino acid. Thus, the code is degenerate. In eight instances, all four codons in a family specify the same amino acid. In the remaining families, the two codons that end with the pyrimidines U and C often specify one amino acid, whereas the two codons that end with the purines A and G specify another. Furthermore, three of the codons, UAA, UAG, and UGA, do not code for any amino acid but instead signal the end of the protein chain.

Decoding mRNAs with tRNAs. The nucleic acid code of an mRNA is converted into an amino acid sequence with the aid of tRNAs. These RNAs are relatively small nucleic acids that vary from 75 to 93 bases in length. Some of the bases in a tRNA pair with each other, so that a tRNA is folded into a pattern that can be represented in two dimensions

	U	C	A	G
U	UUU, UUC — Phe UUA, UUG — Leu	UCU, UCC, UCA, UCG — Ser	UAU, UAC — Tyr UAA, UAG — *Stop*	UGU, UGC — Cys UGA — *Stop* UGG — Trp
C	CUU, CUC, CUA, CUG — Leu	CCU, CCC, CCA, CCG — Pro	CAU, CAC — His CAA, CAG — Gln	CGU, CGC, CGA, CGG — Arg
A	AUU, AUC, AUA — Ile AUG — Met	ACU, ACC, ACA, ACG — Thr	AAU, AAC — Asn AAA, AAG — Lys	AGU, AGC — Ser AGA, AGG — Arg
G	GUU, GUC, GUA, GUG — Val	GCU, GCC, GCA, GCG — Ala	GAU, GAC — Asp GAA, GAG — Glu	GGU, GGC, GGA, GGG — Gly

Fig. 3. Universal (standard) genetic code. Each of the 64 codons found in mRNA specifies an amino acid (indicated by the three-letter abbreviation) or the end of the protein chain (stop). The amino acids are phenylalanine (Phe), leucine (Leu), isoleucine (Ile), methionine (Met), valine (Val), serine (Ser), proline (Pro), threonine (Thr), alanine (Ala), tyrosine (Tyr), histidine (His), glutamine (Gln), asparagine (Asn), lysine (Lys), aspartic acid (Asp), glutamic acid (Glu), cysteine (Cys), tryptophan (Trp), arginine (Arg), and glycine (Gly).

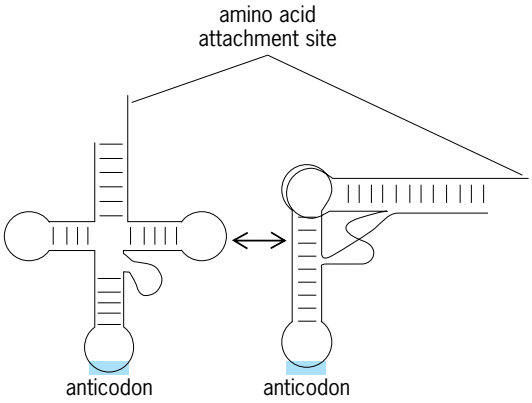

Fig. 4. A transfer RNA in the cloverleaf representation is shown with a two-dimensional projection of how the structure is folded into an L-shaped molecule in three dimensions. The short horizontal and vertical lines represent base pairs that form within the structure. (*Courtesy of Dr. Douglas Buechter*)

as a cloverleaf (Fig. 2). This pattern is further folded in three dimensions to form an L-shaped molecule (**Fig. 4**), to which an amino acid can then be attached. At the other end, three bases in a single-stranded loop are free to pair with a codon in the mRNA. These three bases of a tRNA constitute the anticodon.

Each amino acid has one or more tRNAs, and because of the degeneracy of the code, many of the tRNAs for a specific amino acid have different anticodon sequences. However, the tRNAs for one amino acid are capable of pairing their anticodons only with the codon or codons in the mRNA that specify that amino acid. The tRNAs act as interpreters of the code, providing the correct amino acid in response to each codon by virtue of precise codon–anticodon pairing. The tRNAs pair with the codons and sequentially insert their amino acids in the exact order specified by the sequence of codons in the mRNA. That takes place in the ribosome.

Codon-anticodon interaction. The anticodon of the tRNA pairs with the codon in the mRNA in an antiparallel fashion so that the first base of the codon pairs with the third base of the anticodon, the second base of the codon pairs with the second base of the anticodon, and the third base of the codon pairs with the first base of the anticodon. For the first two bases of the codon, pairing follows the usual rule: A pairs with U, and G with C. However, at the third base of codon, pairing occurs according to a set of rules first suggested by F. Crick and known as the wobble hypothesis:

First base of anticodon	*Third base of codon*
C	G
U	A or G
G	C or U
I	U, C, or A

The purine base inosine (I) is often found instead of A in the first position of the anticodon. There, inosine can pair with any of three bases in the third

position of the codon, thus allowing a single tRNA to read three out of the four codons for a single amino acid. However, inosine is not found in tRNAs in which it would allow pairing with a codon that specifies another amino acid. For example, the histidine tRNA does not contain inosine in the first position of the anticodon because that would allow pairing with the CAA codon which codes for glutamine, not histidine.

Variations. The rules of the genetic code are virtually the same for all organisms, but there are some interesting exceptions. In the microorganism *Mycoplasma capricolum*, UGA is not a stop codon; instead it codes for tryptophan. This alteration in the code is also found in the mitochondria of some organisms. In contrast, UGA is the sole termination codon in some ciliated protozoa, where UAA and UAG are translated as glutamine. A special case involves the unusual amino acid selenocysteine. This metal (selenium)–containing amino acid has a unique tRNA that inserts selenocysteine in response to a UGA codon, in a limited number of proteins in bacteria and mammals. Because these organisms also use UGA as a stop codon, it is thought that the base sequence of the mRNA around selenocysteine-encoding UGAs influences reading to favor amino acid insertion instead of termination. Other specific changes in the meaning of codons that normally encode isoleucine, leucine, and arginine have also been observed in the mitochondria of some organisms.

In addition to changes in the meanings of codons, a modified system for reading codons that requires fewer tRNAs is found in mitochondria. Following the wobble rules, a minimum of two tRNAs are normally needed to read each family of four codons. For example, two tRNAs, with I or C in the first position of the anticodon, could read all four members of a family. In contrast, each of the eight four-codon families that codes for a single amino acid in mitochondria is read by a single tRNA with U in the first position of the anticodon. An explanation for this apparent relaxation of the wobble rules is that the presence of a chemical modification of U in the universal code restricts pairing of U to A or G. In mitochondrial tRNAs, U is unmodified and is therefore able to pair with codons ending in U, C, A, or G. In those cases where two codons in a family specify one amino acid and the other two members specify another amino acid or a stop codon (Fig. 3), the U in the first position of the anticodon is modified and thus pairs only with A and G, as in the universal system. *See* MITOCHONDRIA.

Aminoacyl-tRNA synthetases. The specific attachment of amino acids to tRNAs is catalyzed by a class of enzymes known as aminoacyl-tRNA synthetases. Each amino acid has a different aminoacyl-tRNA synthetase, which catalyzes a reaction known as aminoacylation. Each enzyme attaches its particular amino acid to all of the tRNAs that have anticodons corresponding to that amino acid. Each aminoacyl-tRNA synthetase must distinguish between the various tRNAs so that only the proper ones are aminoacylated. Thus, the accurate recognition of tRNAs by these enzymes establishes the rules of the genetic code.

Aminoacyl tRNA synthetases are divided into two classes of 10 synthetases each. The classification of a synthetase is based on which of two distinct three-dimensional structures it has for the active-site region that catalyzes attachment of an amino acid to its cognate tRNA. There is no evidence that the two distinct structures for the active-site region came from a common ancestor, and it is believed that the two classes arose independently. Also, the occurrence of two classes of synthetases appears to be universal; that is, there is no known example of an organism where all synthetases are of the same class. The historical rationale for the two classes of synthetases is a question of great interest. An attractive hypothesis that is gaining support postulates that, in an early environment, synthetases from opposite classes paired themselves on a single tRNA. One member from each class bound to opposite sides of a given tRNA to form a two-synthetase–one-tRNA complex. From this perspective, the distinct architectures of the two classes of synthetases are seen as needed in order to fit two synthetases to distinct sides of the tRNA.

Because synthetases are responsible for the genetic code and are present in all living organisms, it is likely that they are among the oldest enzymes on Earth. The active-site region of a synthetase is thought to be the primordial, or earliest, form of the synthetase. In addition to the active-site-containing region that is conserved among all members of the same class, each synthetase has a region that is referred to as a nonconserved domain. This domain is more or less unique to the synthetase, even within a particular class.

Recognition of tRNAs. Aminoacyl tRNA synthetases engage in a process called recognition of tRNAs, analogous to the recognition of different buildings or trees by the human eye. Each tRNA has one or more distinguishing characteristics that are determined by the nucleotide sequence and associated three-dimensional structure. The distinguishing features are recognized by the appropriate synthetase that catalyzes the attachment of its cognate amino acid.

One feature that distinguishes a particular tRNA is the anticodon trinucleotide. Were the anticodon trinucleotide the only feature that synthetases use to distinguish among tRNAs, then the interaction of a tRNA with a synthetase would establish a direct connection between a particular amino acid and its complementary trinucleotide of the genetic code. However, there are synthetases that do not interact with the anticodon and yet still correctly identify their tRNA. In these cases, the relationship between a particular trinucleotide and a given amino acid is indirect, and suggests the existence of another code, namely, an operational RNA code for amino acids that was the precursor to the genetic code.

Second genetic code. In the case of the tRNA specific for alanine with an anticodon GGC, a single G-U base pair, at a specific position far removed from the anticodon, is needed for aminoacylation with alanine. This base pair is located in a portion of the tRNA molecule relatively close to the site of amino acid

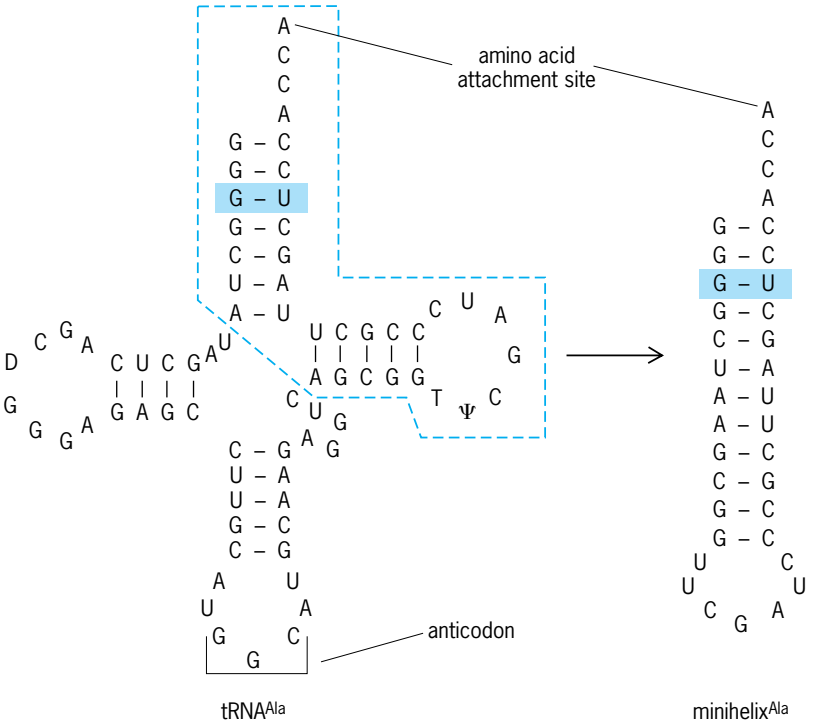

Fig. 5. Cloverleaf representation of the nucleotide sequence, comprising four monomer units (A = adenine, G = guanine, C = cytosine, and U = uracil), of an alanine tRNA with a GGC anticodon and of a synthetic minihelix based on a limited portion of the whole tRNA (enclosed by a broken line). D, T, and Ψ are modified uracil bases in the tRNA, and T and Ψ have been replaced by U in the minihelix. The critical G-U base pair recognized by the alanine tRNA synthetase is highlighted in both structures. (*Courtesy of Dr. Douglas Buechter*)

attachment (**Fig. 5**). The part of the tRNA structure that is near the amino acid attachment site is known as the acceptor stem, and consists of a short helix joined to a single-stranded piece that terminates at the site of amino acid attachment. A truncated version of alanine tRNA that contains the critical base pair together with the adjacent structure near the amino acid attachment site can be synthesized as a minihelix. This tRNA fragment is recognized by alanine tRNA synthetase.

Even smaller RNA fragments that recapitulate the acceptor ends of several tRNAs are aminoacylated with the correct amino acid (**Fig. 6**). Because aminoacylation is specific and depends on the order of bases in the short fragments, the relationship between the structure-sequence of RNA fragments and each amino acid can be viewed as an operational RNA code, referred to also as the second genetic code.

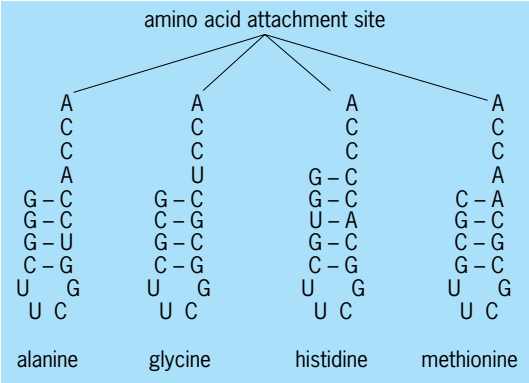

Fig. 6. The smallest RNA oligonucleotides that have been shown to be specifically aminoacylated. The amino acid which can be attached to each RNA oligonucleotide is indicated. (*Courtesy of Dr. Barry Henderson*)

Several synthetases recognize the anticodon as one of the elements needed for identification of the tRNA, but even in these cases nucleotides in the acceptor stem are still important for precise and efficient recognition of the whole tRNA and small RNA fragments. The class-defining and active-site-containing domain of aminoacyl tRNA synthetases is largely responsible for interactions with the acceptor end of a tRNA and, therefore, for interpretation of the second genetic code. After establishment of the second genetic code in a primitive world, RNA fragments with signals for attachment of specific amino acids may have become incorporated into RNA structures which eventually became transfer RNAs. That step, in turn, gave birth to the modern genetic code. In addition, the synthetases developed editing functions that are tRNA-dependent. The development of these editing functions was one of the events that linked the evolution of tRNAs with that of the synthetases. These editing functions enable the synthetases to achieve high accuracy of aminoacylation. During evolution, the earliest tRNA synthetases apparently acquired additional, unique domains which allowed for editing. *See* GENE; GENE ACTION; GENETICS.

P. Schimmel; K. Ewalt

Bibliography. J. M. Berg, J. L. Tymoczko, and L. Stryer, *Biochemistry*, 5th ed., 2002; S. Cusack, RNA-protein complexes, *Curr. Opin. Struc. Biol.*, 9:66–73, 1999; A. J. Griffiths, *An Introduction to Genetic Analysis*, 6th ed., 1999; P. Schimmel and L. Ribas de Pouplana, Formation of two classes of tRNA synthetases in relation to editing functions and genetic code, *Cold Spring Harbor Symp. Quantitative Biol.*, 66:161–166, 2001; P. Schimmel and L. Ribas de Pouplana, Transfer RNA: From minihelix to genetic code, *Cell*, 81:983–986, 1995.

Genetic engineering

The artificial recombination of nucleic acid molecules in the test tube, their insertion into a virus, bacterial plasmid, or other vector system, and the subsequent incorporation of the chimeric molecules into a host organism in which they are capable of continued propagation. The construction of such molecules has also been termed gene manipulation because it usually involves the production of novel genetic combinations by biochemical means. *See* NUCLEIC ACID.

Genetic engineering provides the ability to propagate and grow in bulk a line of genetically identical organisms, all containing the same artificially recombinant molecule. Any genetic segment as well as the gene product encoded by it can therefore potentially be amplified. For these reasons the process has also

been termed molecular cloning or gene cloning. *See* GENE.

Basic Techniques

The central techniques of such gene manipulation involve (1) the isolation of a specific deoxyribonucleic acid (DNA) molecule or molecules to be replicated (the passenger DNA); (2) the joining of this DNA with a DNA vector (also known as a vehicle or a replicon) capable of autonomous replication in a living cell after foreign DNA has been inserted into it; and (3) the transfer, via transformation or transfection, of the recombinant molecule into a suitable host.

Isolation of passenger DNA. Passenger DNA may be isolated in a number of ways; the most common of these involves DNA restriction.

DNA restriction. DNA purified from an organism can be prepared for cloning only after it has been cut into smaller molecules. Restriction endonucleases make possible the cleavage of high-molecular-weight DNA. Although three different classes of these enzymes have been described, only type II restriction endonucleases have been used extensively in the manipulation of DNA. Type II restriction endonucleases are DNAases that recognize specific short nucleotide sequences (usually 4 to 6 base pairs in length), and then cleave both strands of the DNA duplex, generating discrete DNA fragments of defined length and sequence. *See* RESTRICTION ENZYME.

A number of restriction enzymes make staggered cuts in the two DNA strands, generating single-stranded termini. Restriction enzyme Eco R1 cleaves the sequence GAATTC (where G = guanine; A = adenine; T = thymine; C = cytosine) between the G and A residues, generating protruding single-stranded ends having 5′ termini.

The various fragments generated when a specific DNA is cut by a restriction enzyme can be easily resolved as bands of distinct molecular weights by agarose gel electrophoresis. Specific sequences of these bands can be identified by a technique known as Southern blotting. In this technique, DNA restriction fragments resolved on a gel are denatured and blotted onto a nitrocellulose filter. The filter is incubated together with a radioactively labeled DNA or RNA probe specific for the gene under study. The labeled probe hybridizes to its complement in the restricted DNA, and the regions of hybridization are detected autoradiographically. Fragments of interest can then be eluted out of these gels and used for cloning. Purification of particular DNA segments prior to cloning reduces the number of recombinants that must later be screened.

Mechanical shearing of DNA. Another method that has been used to generate small DNA fragments is mechanical shearing. Intense sonification of high-molecular-weight DNA with ultrasound, or high-speed stirring in a blender can both be used to produce DNA fragments of a certain size range. Shearing results in random breakage of DNA, producing termini consisting of short, single-stranded regions.

Other sources. Another source of DNA to be replicated is DNA complementary to poly(A) RNA, or cDNA, which is synthesized in the test tube. The cDNA synthesis involves the use of the enzyme, reverse transcriptase, which uses poly(A) RNA as a template for the synthesis of a complementary DNA strand. The RNA strand in the resulting RNA–DNA hybrid is removed by alkaline hydrolysis, and the remaining single-stranded DNA is converted into the double-stranded form in a second reaction involving DNA polymerase I. *See* ENZYME.

In addition, short oligonucleotides, synthesized chemically, can be used for cloning purposes. *See* OLIGONUCLEOTIDE.

Joining DNA molecules. Once the proper DNA fragments have been obtained, they must be joined. This is accomplished by the following methods.

Cohesive end ligation. When cleavage with a restriction endonuclease creates cohesive ends, these can be annealed with a similarly cleaved DNA from another source, including a vector molecule. When such molecules associate, the joint has nicks a few base pairs apart in opposite strands. The enzyme DNA ligase can then repair these nicks to form an intact, duplex recombinant molecule, which can be used for transformation and the subsequent selection of cells containing the recombinant molecule (**Fig. 1**).

Cohesive ends can also be created by the addition of synthetic DNA linkers to blunt-ended DNA molecules. These linkers are short DNA duplexes containing the recognition site for a restriction enzyme which produces cohesive termini. Linkers are ligated to blunt-ended passenger DNA molecules by DNA ligase encoded by the phage T4. After digesting the product with the restriction enzyme that cleaves the linkers, the products can be ligated to any vehicle DNA via the complementary termini and then cloned.

Homopolymer tailing. Another method for joining DNA molecules involves the addition of homopolymer extensions to different DNA populations followed by an annealing of complementary homopolymer sequences. For example, short nucleotide sequences of pure adenine can be added to the 3′ ends of one population of DNA molecules and short thymine blocks to the 3′ ends of another population. The two types of molecules can then anneal to form mixed dimeric circles that can be used directly for transformation. Single-stranded gaps that may remain in the two strands at each connection will be repaired in the transformed cell.

Blunt-end ligation. T4 DNA ligase carries out the intermolecular joining of DNA substrates at completely base-paired ends; such blunt ends can be produced by cleavage with a restriction enzyme or by mechanical shearing followed by enzyme treatment.

Transformation. The desired DNA sequence, once attached to a DNA vector, must be transferred to a suitable host. Transformation is defined as the introduction of foreign DNA into a recipient cell. Transformation of a cell with DNA from a virus is usually referred to as transfection.

Transformation in any organism involves (1) a method that allows the introduction of DNA into

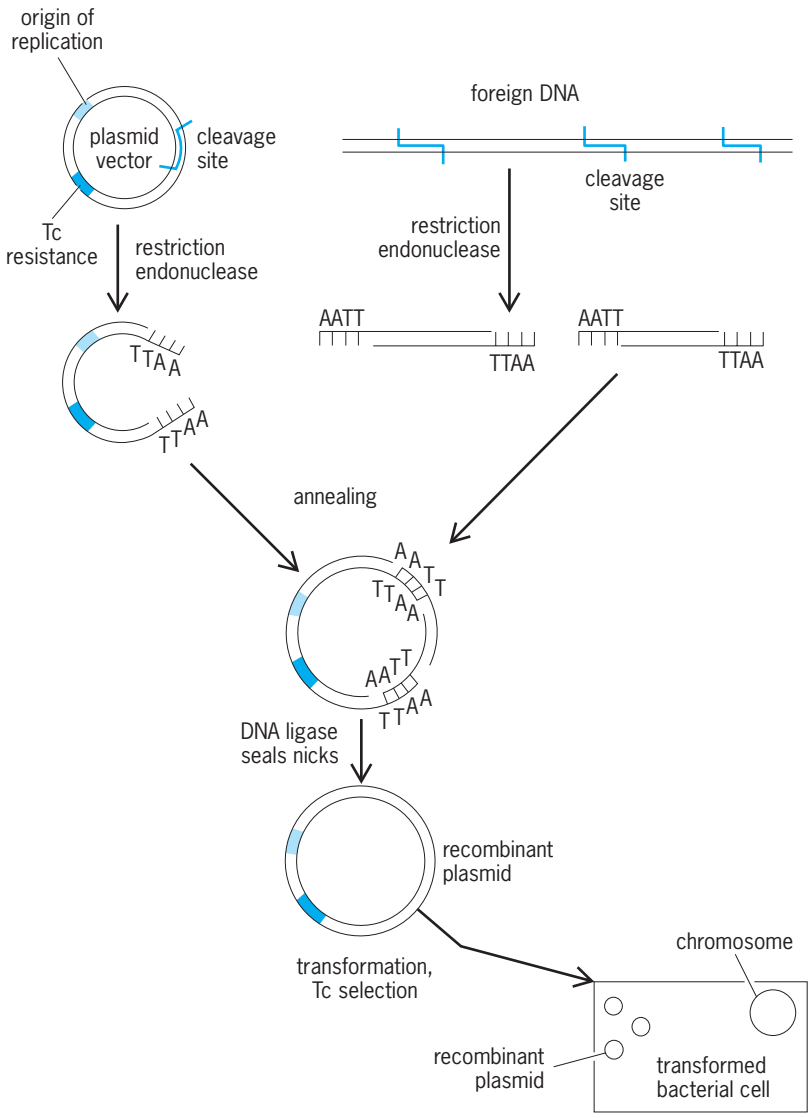

Fig. 1. Use of DNA ligase to create a covalent DNA recombinant joined through association of termini generated by Eco R1. (*After R. W. Old and S. B. Primrose, Principles of Gene Manipulation, 3d ed., Blackwell Scientific Publications, 1985*)

the cell and (2) the stable integration of DNA into a chromosome, or maintenance of the DNA as a self-replicating entity. *See* TRANSFORMATION (BACTERIA).

Prokaryotes and lower eukaryotes. *Escherichia coli* is usually the host of choice for cloning experiments, and transformation of *E. coli* is an essential step in these experiments. *Escherichia coli* treated with calcium chloride are able to take up DNA from bacteriophage lambda as well as plasmid DNA. Calcium chloride is thought to effect some structural alterations in the bacterial cell wall.

In addition to *E. coli*, other prokaryotes have also been used as hosts for cloning experiments, among them the nonenteric bacteria, *Bacillus* species, and actinomycetes. An efficient method for transformation in *Bacillus* involves poly(ethylene glycol)–induced DNA uptake in bacterial protoplasts and subsequent regeneration of the bacterial cell wall. Actinomycetes can be similarly transformed. Transformation can also be achieved by first entrapping the DNA with liposomes followed by their fusion with the host cell membrane. Similar transformation methods have been developed for lower eukaryotes such as the yeast *Saccharomyces cerevisiae* and the filamentous fungus *Neurospora crassa*. *See* LIPOSOMES.

Animal cells. Several methods are available for the transfer of DNA into cells of higher eukaryotes. Specific genes or entire viral genomes can be introduced into cultured mammalian cells in the form of a coprecipitate with calcium phosphate. DNA complexed with calcium phosphate is readily taken up and expressed by mammalian cells. DNA complexed with diethylamino-ethyl-dextran (DEAE-dextran), or DNA trapped in liposomes or erythrocyte ghosts may also be used in mammalian transformation. Alternatively, bacterial protoplasts containing plasmids can be fused to intact animal cells with the aid of chemical agents such as poly(ethylene glycol) [PEG]. Finally, DNA can be directly introduced into cells by microinjection. The efficiency of transfer by each of these methods is quite variable.

Higher plants. Introduction of DNA sequences by insertion into the transforming (T)-DNA region of the tumor-inducing (Ti) plasmid of *Agrobacterium tumefaciens* is a method of introducing DNA into plant cells and ensuring its integration. The Ti plasmid has been used to introduce selectable marker genes, including bacterial genes coding for antibiotic resistance, into plant cells. The control sequences used to express these chimeric genes are derived from plant genes. Many foreign genes have been expressed in the transformant plant cells and have conferred selectable properties into plant cells.

Although the Ti plasmid can be used as a gene transfer vector, its use is limited in that *A. tumefaciens*, the bacterium that transmits Ti plasmids, normally infects primarily dicotyledonous plants. Monocotyledonous plants, including economically important cereals, cannot be infected by the Ti plasmid.

Because of the limitations of the host range of *A. tumefaciens*, alternative transformation systems are being developed for gene transfer in plants. They include the use of liposomes, as well as induction of DNA uptake in plant protoplasts. With those methods, however, transformation frequencies are low. Plant DNA viruses, such as the cauliflower mosaic virus, might also be used as vectors for DNA transfer.

Foreign DNA has been introduced into plant cells by a technique called electroporation. This technique involves the use of electric pulses to make plant plasma membranes permeable to plasmid DNA molecules. Plasmid DNA taken up in this way has been shown to be stably inherited and expressed.

Cloning Vectors

There is a large variety of potential vectors for cloned genes. The vectors differ in different classes of organisms.

Prokaryotes and lower eukaryotes. Three types of vectors have been used in these organisms: plasmids, bacteriophages, and cosmids.

Plasmids. Plasmids are extrachromosomal DNA sequences that are stably inherited. *Escherichia coli* and its plasmids constitute the most versatile type of host–vector system known for DNA cloning.

Several natural plasmids, such as ColE1, have been used as cloning vehicles in *E. coli*. In addition, a variety of derivatives of natural plasmids have been constructed by combining DNA segments and desirable qualities of older cloning vehicles. These qualities, which facilitate manipulation of cloned DNA and limit the ability of recombinants to survive in the environment beyond the laboratory, include (1) reduced molecular weight of the plasmid; (2) construction of only one site for a given restriction enzyme located in a region not essential for plasmid replication; (3) introduction of one or more phenotypic markers into the plasmid to permit selection of cells containing the plasmid and to aid in the identification of plasmids that have incorporated the foreign DNA. It is also desirable to use plasmids that can be obtained in high yields.

The most versatile and widely used of these plasmids is pBR322. This plasmid has been completely sequenced, providing a detailed restriction enzyme cleavage map and DNA sequence information (**Fig. 2**). The pBR322 plasmid carries the genes coding for resistance to the antibiotics ampicillin and tetracyline.

Transformation in yeast has been demonstrated using a number of plasmids, including vectors derived from the naturally occurring 2μ plasmid of yeast. These vectors contain DNA fragments from the 2μ yeast plasmid, yeast nuclear DNA, and the *E. coli* vector, pMB9. Other yeast vectors are recombinants consisting of *E. coli* vectors and yeast DNA fragments.

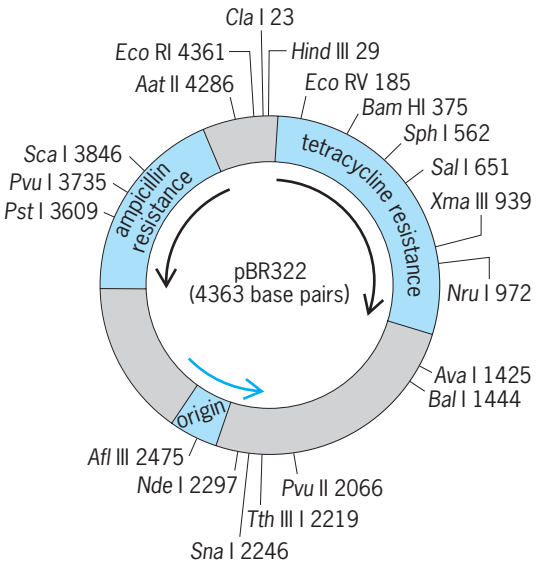

Fig. 2. Structure of pBR322 showing the unique cleavage sites. The numbers relate to the coordinate of the base which corresponds to the 5′ nucleotide of each recognition sequence. The black arrows show the directions of transcription of the Ap^R and Tc^R genes. The colored arrow shows the direction of DNA replication. (*After R. W. Old and S. B. Primrose, Principles of Gene Manipulation, 3d ed., Blackwell Scientific Publications, 1985*)

Bacteriophages. Bacteriophage lambda is a virus of *E. coli*. Several lambda-derived vectors have been developed for cloning in *E. coli*, and for the isolation of particular genes from eukaryotic genomes. These lambda derivatives have several advantages over plasmids: (1) Thousands of recombinant phage plaques can easily be screened for a particular DNA sequence on a single petri dish by molecular hybridization. (2) Packaging of recombinant DNA in laboratory cultures provides a very efficient means of DNA uptake by the bacteria. (3) Thousands of independently packaged recombinant phages can be easily replicated and stored in a single solution as a "library" of genomic sequences. *See* BACTERIOPHAGE.

Cosmids. Plasmids have also been constructed that contain the phage *cos* DNA site, required for packaging into the phage particles, and ColE1 DNA segments, required for plasmid replication. These plasmids have been termed cosmids. After packaging laboratory cultures, the vector is used to infect a host. The recombinant cosmid DNA is injected and circularizes like phage DNA but replicates as a plasmid. Transformed cells are selected on the basis of a vector drug resistance marker. Cosmids provide an efficient means of cloning large pieces of foreign DNA and have therefore been used for constructing libraries of eukaryotic genomic fragments.

Animal cells. In contrast to the wide variety of plasmid and phage vectors available for cloning in prokaryotic cells, relatively few vectors are available for introducing foreign genes into animal cells. In the latter case, the most commonly used vectors are derived from simian virus 40 (SV40). *See* ADENO-SV40 HYBRID VIRUS.

The SV40 genome consists of a small covalently closed circular DNA molecule for which the entire nucleotide sequence is known. Normal SV40 cannot be used as a vector, since there is a physical limit to the amount of DNA that can be packaged into the virus capsid, and the addition of foreign DNA would generate a DNA molecule too large to be packaged into a viral particle. However, SV40 mutants lacking portions of the genome can be propagated in mixed infections in which a "helper" virus supplies the missing function.

Propagating recombinants in eukaryotic cells using SV40 vectors imposes some limitations, such as the severe limitation on the size of gene segments that can be cloned in the small viral genome and the fact that only cells permissive for SV40 replication can be used.

Bovine papilloma virus is another DNA virus that has been used as a vector in eukaryotic animal cells. The rat insulin gene inserted into the viral DNA has been used successfully to transform mouse cells.

Another system for the stable transfer of genes into mammalian cells involves cotransformation of cells with two physically separate sets of genes. For example, the introduction of the thymidine kinase (tk) gene from Herpes simplex virus to mouse tk^- cells results in stable and efficient transformation of cells that express the viral gene.

In *Drosophilia*, transposable DNA elements, that is, DNA segments that are capable of moving from one position to another within the genome of a cell, have provided the basis for development of an efficient and controlled system of gene transfer. DNA segments of interest can be transposed into germ-line chromosomes along with a transposable element to which the segment has been previously ligated. *See* TRANSPOSONS.

Plant cells. Two systems for the delivery and integration of foreign genes into the plant genome are the Ti plasmid of the soil bacterium *Agrobacterium* and the DNA plant virion cauliflower mosaic virus.

The Ti plasmid is a natural gene transfer vector carried by *A. tumefaciens*, a pathogenic bacterium that causes crown gall tumor formation in dicotyledonous plants. A T-DNA segment present in the Ti plasmid becomes stably integrated into the plant cell genome during infection. This property of the Ti plasmid has been exploited to show that DNA segments inserted in the T-DNA region can be cotransferred to plant DNA. *See* CROWN GALL.

Caulimoviruses belong to a group of plant DNA viruses that contain double-stranded DNA. Of these, cauliflower mosaic virus is the best studied for its cloning vehicle potential. The genomes of over 30 strains of cauliflower mosaic virus have been mapped by using restriction enzymes. Purified preparations of the cauliflower mosaic virus DNA can infect plant cells with relatively high efficiency. Cauliflower mosaic virus DNA is also transformable after gene manipulation, which indicates that it might be possible to use cauliflower mosaic virus as a plant vector.

Cloned Gene Expression

It is sometimes the aim of the genetic engineer to promote the expression of a cloned gene not only for the analysis of gene structure and function but also for amplification of the synthesis of a desirable gene product. DNA cloning methods have enabled the genetic modification of bacteria and unicellular eukaryotes, rendering them capable of producing virtually any gene product of animal or plant cells.

The aim of gene transfer experiments is to examine the mechanisms of gene regulation and expression in the context of normal cell differentiation. DNA sequences cloned by recombinant DNA technology have been transferred into fertilized mouse, *Drosophila*, and *Xenopus* oocytes by microinjection techniques. In the mouse and *Drosophila* this system has allowed study of the expression and regulation of transferred gene sequences in specific tissues during normal cellular differentiation, and their transmission from one generation to the next. The *Xenopus* oocyte has been used mainly for analysis of transcription and translation of injected sequences.

In the mouse, foreign DNA is injected into a pronucleus of a fertilized egg and subsequently becomes incorporated into the chromosomes of the diploid zygote. The injected eggs are then transferred to a foster-mother mouse, where normal embryonic development occurs. Some of the progeny mice come to contain the foreign DNA in their cells, and the expression of the introduced DNA can be examined during development of these transgenic mice.

A variety of recombinant DNAs have been injected into the mouse oocyte, including plasmids containing SV40 sequences, the Herpes simplex tk gene, a human alpha interferon cDNA, and beta-globin genes from a variety of animals. The injected DNA has been shown to be integrated into the genomes of some of the transformants, and in many cases transmission of DNA from the original transgenic mouse to its progeny has also been demonstrated.

Applications. Genetic engineering technology has made possible a wide variety of practical applications, several of which are discussed below.

Isolation of specific genes. Recombinant DNA technology has permitted the isolation and detailed structural analysis of a large number of prokaryotic and eukaryotic genes. This contribution is especially significant in the eukaryotes because of their large genomes. The methods outlined above provide a means of fractionating and isolating individual genes, since each clone contains a single sequence or a few DNA sequences from a very large genome. Isolation of a particular sequence of interest has been facilitated by the ability to generate a large number of clones and to screen them with the appropriate "probe" (radioactively labeled RNA or DNA) molecules.

Analysis of gene structure and function. Genetic engineering techniques provide pure DNAs in amounts sufficient for mapping, sequencing, and direct structural analyses. Furthermore, gene structure–function relationships can be studied by reintroducing the cloned gene into a eukaryotic nucleus and assaying for transcriptional and translational activities. The DNA sequences can be altered by mutagenesis, before their reintroduction in order to define precise functional regions.

Such analyses have provided insights into the complexity of the eukaryotic genome. For example, many eukaryotic genes contain noncoding sequences (introns) interspersed in the coding sequences (exons) of genes. The primary RNA transcripts are spliced to remove the noncoding sequences as the mRNA matures. Exon–intron boundaries from diverse sources have been sequenced to reveal a common sequence pattern at the opposite ends of different introns. These sequences may be part of a mechanism that brings opposite ends of introns together prior to splicing. *See* EXON; INTRON.

Cloning and sequencing of increasing numbers of eukaryotic genes have brought to light similarities in specific sequences near the 5′ and 3′ ends of different genes. These sequences are suspected to be involved in the regulation of transcription or translation.

Molecular probes for a number of eukaryotic genes have revealed the existence of clustered gene families. Some families, such as the histones, have genes with nearly identical sequences which are arranged tandemly along the chromosomal DNA and which function simultaneously to synthesize their gene

products. Members of other gene families, however, though clustered on the chromosome, are not identical; such genes encode closely related proteins. Embryonic and adult globin polypeptides are examples of such a gene family. Within the globin gene cluster are also found "globinlike" sequences that are nonfunctional because of deletions in essential regulatory regions. These segments of DNA may represent vestiges of globin genes that were once functional but, during evolution, have lost their physiological role. *See* HEMOGLOBIN; NUCLEOPROTEIN.

Polypeptide production. Genetic engineering methodology has provided means for the production of polypeptides and proteins. Although this expression of heterologous genes is a function of a variety of complex factors, maximizing the expression of cloned sequences has been under intense and rapid development. It is now possible to produce a wide variety of foreign proteins in *E. coli*. These range from enzymes useful in molecular biology to a vast range of polypeptides with potential human therapeutic applications, such as insulin, interferon, growth hormone, immunoglobins, and enzymes involved in the dynamics of blood coagulation.

Biotechnology on a large scale has evolved as a result of genetic engineering in the laboratory. Human insulin manufactured through the use of genetically engineered *E. coli* has been used safely in diabetes therapy. Other polypeptides such as human growth hormone and interferon have been in production. Research has been under way to manufacture vaccines and many other therapeutic and diagnostic agents by using the new technology. *See* BIOTECHNOLOGY.

Medical applications. The screening of appropriately restricted human DNA with suitable molecular probes, that is, true genes or simply nonrepeated sequences of human DNA cloned in a vector, has disclosed differences between individuals with respect to the size of the restricted DNA fragment containing DNA sequences homologous to the probe. These types of variants within a population, called restriction-fragment-length polymorphisms, can be used to map the human genome. *See* GENETIC MAPPING.

Restriction-fragment-length polymorphisms are inherited by mendelian segregation and are distributed in populations as classical examples of common genetic polymorphisms. If such a DNA variant is located close to a defective gene (which cannot be tested for directly), the DNA variant can be used as a marker to detect the presence of the disease-causing gene. It is estimated that detection of 150 to 300 different DNA markers of this type distributed randomly throughout the human genome would allow detection of any disease-producing gene. Such information would be important in determining individual susceptibility to diseases and in taking measures to prevent them. *See* POLYMORPHISM (GENETICS).

In cases where the specific mutation in DNA has been identified, such as sickle cell anemia, certain restriction enzymes that recognize the abnormal DNA sequence at the mutant site can be used to demonstrate the mutation without study of the affected and unaffected family members (**Fig. 3**). By using restriction enzymes in molecular hybridization experiments, changes in DNA sequences that occur in other hemoglobinopathies (genetically determined defects of function or synthesis of hemoglobins) may also be recognized directly by using the appropriate probes. Such changes include deletions, mutations of restriction sites, and deletions of restriction sites. This kind of information provides means of antenatal diagnosis of human diseases by analysis of fetal blood samples.

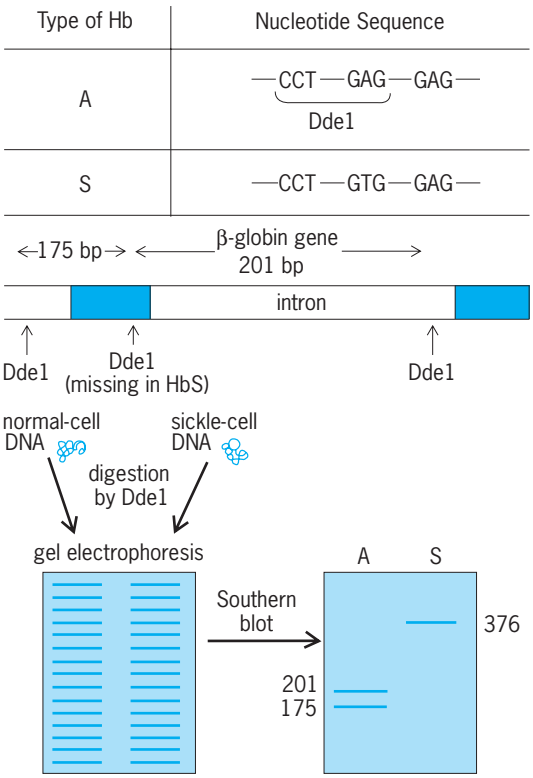

Fig. 3. Detection of the sickle cell globin gene by Southern blotting. The base change (A to T) that causes sickle cell anemia destroys a Dde1 site that is present in the normal β-globin gene. This difference can be detected by digesting sickle cell and normal DNA with Dde1 and performing Southern blot hybridization. Normal DNA will produce two Dde1 fragments of 201 and 175 base pairs, whereas sickle cell DNA will produce a single fragment of 376 base pairs. (*After J. D. Watson et al., Recombinant DNA: A Short Course, W. H. Freeman, 1983*)

A number of human diseases are caused by the abnormal function of a major single gene. Various thalassaemias—conditions characterized by reduced or absent synthesis of alpha- or beta-globin chains (the protein components of hemoglobin)—are examples of such diseases. If a normal copy of the defective gene could be introduced and was expressed in the appropriate cells of an individual suffering from such a genetic disorder, the normal gene product would be able to alleviate the problems associated with the presence of the defective gene. Since hemoglobin is produced by certain bone marrow cells that can easily be retrieved, thalassaemia might be treated by the addition of normal isolated genes (in the form of DNA) to the individual's abnormal

erythropoietic marrow cells. After the normal genes have been taken up, the manipulated marrow cells could be returned to the person for proliferation and production of normal hemoglobin. *See* HEMATOLOGIC DISORDERS; HEMOGLOBIN.

Genetic modification of plants. Experiments showing the successful transfer and expression of foreign DNA in plant cells using the Ti plasmid, as well as the demonstration that whole plants can be regenerated from cells containing mutated regions of T-DNA, indicate that the Ti plasmid system may be an important tool in the genetic engineering of plants. Such a system will help in the identification and characterization of plant genes as well as provide basic knowledge about gene organization and regulation in higher plants. Once genes useful for crop improvement have been identified, cloned, and stably inserted into the plant genome, it may be possible to engineer plants to be resistant to environmental stress, to pests, and to pathogens. *See* BREEDING (PLANT); GENE; GENE ACTION; SOMATIC CELL GENETICS.

Pamela K. Mulligan

Bibliography. H. Chen, *Medical Genetics Handbook*, 1988; J. Edwards-Moulds and L. C. Lasky (eds.), *Clinical Applications of Genetics Engineering*, 1987; J. Evans and A. Hollaender (eds.), *Genetic Engineering of Animals: An Agricultural Perspective*, 1986; B. R. Glick and J. J. Pasternak, *Molecular Biotechnology: Principles and Applications of Recombinant DNA*, 2d ed., 1998; C. P. Hollenberg and H. Sahm (eds.), *Microbial Genetical Engineering and Enzyme Technology (Biotech 1)*, 1987; D. S. Nicholl, *An Introduction to Genetic Engineering*, 1994; R. Old and S. Primrose, *Principles of Gene Manipulation: An Introduction to Genetic Engineering*, 5th ed., 1994.

Genetic mapping

The resolution of the elements of the genetic material of an organism into a linear array. In order of decreasing size, these elements are chromosomes, genes, and the nucleotide sequence of DNA. A complete description of the genetic material of an organism would consist of its nucleic acid sequence. *See* CHROMOSOME; DEOXYRIBONUCLEIC ACID (DNA).

The construction of a genetic map proceeds through a series of successive approximations. First the chromosomal constitution (karyotype) of the organism is established. Genetically linked loci on each chromosome are grouped and ordered within linkage groups. This process is continued until each chromosome is covered by a set of genetically linked loci to form a complete linkage map. Such maps are used to establish the position of other loci or genes on the chromosome. DNA segments encoding genes on the map are isolated and further analyzed by using the techniques of molecular biology. Finally, the organization and primary DNA structure of genes are determined by DNA sequencing. *See* MOLECULAR BIOLOGY.

The first efforts to sequence DNA, pioneered by Walter Gilbert and Fred Sanger in the 1970s, decoded stretches of DNA measuring a few hundred bases long. The sequencing of the first viral genome (measuring 5000 bases) provided unique insights into the structure and function of genes and its genome organization. The potential to generate vast amounts of information about an organism from its genetic code inspired efforts toword the automation of DNA sequencing and the creation of the Human Genome Project.

The Human Genome Project was conceived in 1985 and implemented in 1990. Its principal goals were to achieve the complete mapping and sequencing of the human genome by 2005. The research program was divided into three principal parts: construction of dense genetic maps, construction of a physical map, and sequencing of the whole genome using the information gained from the first two goals. In 2001 the first draft of the human genome sequence became avaliable to the scientific community, and its completion was announced in April 2003. Since then, the genomic sequences of more than 15 model organisms have also been obtained. *See* HUMAN GENOME PROJECT.

The genetic mapping process has benefited enormously from the resources created through the Human Genome Project. The identification of disease genes has been aided by the availability of dense genetic maps, physical maps, and the genome sequence. The ultimate goal of genetic mapping is to identify disease genes. The isolation and analysis of such genes will provide better tools for early detection of disease susceptibility and eventually could lead to the development of preventive and personalized therapeutic strategies.

Karyotyping and chromosome assignment. The karotype describes the number of complete sets of chromosomes (ploidy) within a cell and the number of chromosomes within each set. For example, human cells are diploid, containing 22 pairs of autosomes (numbered 1 to 22) and one pair of sex chromosomes (XX in females, XY in males). The karyotype of an organism is determined by observation of stained chromosome preparations under a light microscope. Homologous pairs of chromosomes are identified on the basis of size, centromeric position, and the banding pattern produced by such dyes as the Giemsa stain.

Traditionally genes or genetic loci were assigned to a particular human chromosome with the aid of somatic cell hybrids. Fusion of human cells with those of another species, typically a rodent, initially results in hybrid cells containing a complete set of chromosomes from both species. In subsequent cell divisions, human chromosomes are preferentially and randomly lost from such human–rodent hybrids. This results in a series of hybrid cells, each of which contains a subset of the original human chromosome complement. By analyzing a number of such hybrid cells, it is possible to correlate the presence or absence of a particular chromosome with the presence or absence of a particular gene. *See* SOMATIC CELL GENETICS.

The presence of a particular gene within a hybrid cell can be determined in any of a number of ways. In some cases, the human gene may be able to complement a mutation in the recipient cell, or the protein products of equivalent donor and recipient genes can often be distinguished on the basis of size, charge, or enzymatic activity. If the gene of interest encodes proteins present on the cell surface, antibodies raised against these proteins can be used to identify and separate hybrid cells that express such genes from those that do not. This process has been simplified by the use of antibodies with fluorescent labels and a fluorescence-activated cell sorter. *See* ANTIBODY; IMMUNOFLUORESCENCE.

The presence of any DNA sequence can also be detected in hybrid cells by nucleic acid hybridization. Under appropriate conditions, sequence-specific annealing of two complementary single-stranded DNA molecules occurs. A sequence is tested for by the isolation and immobilization of genomic DNA from the hybrid cell on a nylon membrane. Incubation of the membrane with a radioactively labeled single-stranded DNA fragment (probe) in aqueous solution allows the formation of double-stranded molecules between the probe and complementary hybrid DNA sequences. Such duplexes can be detected by autoradiography with x-ray film. Nucleic acid hybridization permits the mapping of any DNA sequence in hybrid cells, including polymorphic DNA markers, independent of the expression of a human gene in hybrid cells. *See* AUTORADIOGRAPHY; NUCLEIC ACID.

If the sequence of the gene is known, a polymerase chain reaction (PCR) assay is used to type the cell hybrid for the relevant sequence. Today, this can also be applied to mapped clones used for sequencing the human genome. In addition, the sequence of the gene in question can be directly compared to and mapped to the human genome. Short sequences can be identified rapidly using the computer program e-PCR.

Longer sequences can be found using the Basic Local Alignment Search Tool (BLAST) or other rapid sequence-similarity searching algorithms.

Isolated DNA sequences can also be mapped to whole chromosomes by in-situ hybridization. At first, in-situ hybridization experiments used radioactive-labeled DNA probes and photographic emulsion for detection. Today, the preferred method is fluorescence in-situ hybridization (FISH), which uses fluorescent detection. Compared with radioactive detection methods, fluorescence detection is safer and faster and can map more than one DNA sequence simultaneously and at a higher resolution.

Somatic cell genetics contributed significantly to our ability to analyze and understand the human genome in the pregenomic era. The mapping methodologies used, such as in-situ hybridization, were applied to the mapping of microbial cloning systems (such as yeast artificial chromosomes, bacterial artificial chromosomes, cosmids, and plasmids), which were used as the bases of modern physical maps.

Linkage mapping. The order and distance between genetic loci on the same chromosome can be determined by the frequency of recombination events between them. Recombination between homologous chromosomes occurs during meiosis and results in the exchange of chromosomal regions. The probability of a recombination event occurring between two loci is proportional to the distance between them. Thus, the frequency of recombination between two loci provides an estimate of the distance separating them. Positioning of a third locus can be achieved by determining the frequency of recombination which occurs between it and the other two loci. The stepwise analysis of many genetic loci in this manner can be used to generate a linkage map of each chromosome. *See* RECOMBINATION (GENETICS).

Accurate determination of linkage between two loci requires the analysis of large numbers of individuals from multigeneration families for recombination events. In human populations these are seldom available, and it is thus necessary to combine the information obtained from many smaller families by using so-called lod (log odds) scores. The lod score gives the probability that a particular distribution and combination of traits would have occurred if the two loci were linked with a particular recombination fraction (σ), and is defined by the following equation:

$$\text{lod} = \log_{10} \frac{\text{probability of linkage between two loci assuming } \sigma = \sigma}{\text{probability of the two loci being unlinked } (\sigma = 0.5)}$$

Lod scores from individual families can be summed to give a cumulative score. A lod score of 3.0 is considered good evidence for linkage, and one of −2.0 is indicative of its absence.

The observation of recombination events is dependent upon the possibility of distinguishing between equivalent loci on homologous chromosomes. Only loci with a minimum of two alleles can be used in linkage analysis, and these alleles must be present in the population at a reasonable frequency to be informative. Although some genes, such as those for blood groups, have suitable allelic frequencies, their number is limited.

Restriction-fragment-length polymorphism markers. Polymorphisms (genetic variations among individuals in a population) at the DNA sequence level, called restriction-fragment-length polymorphisms, have provided an unlimited source of codominantly inherited markers ("landmarks") ideal for the construction of linkage maps. Restriction enzymes, isolated from bacteria, recognize and hydrolyze DNA at specific 4–8-bp sequences. More than 200 restriction enzymes are known, each of which recognizes a particular sequence. Any sequence change within a restriction enzyme recognition site prevents cleavage there. In the absence of such a site, cleavage with the enzyme produces a larger DNA fragment. This absence can be detected by Southern blotting (**Fig. 1**), in which genomic DNA is cleaved by the restriction enzyme and fractionated according to size

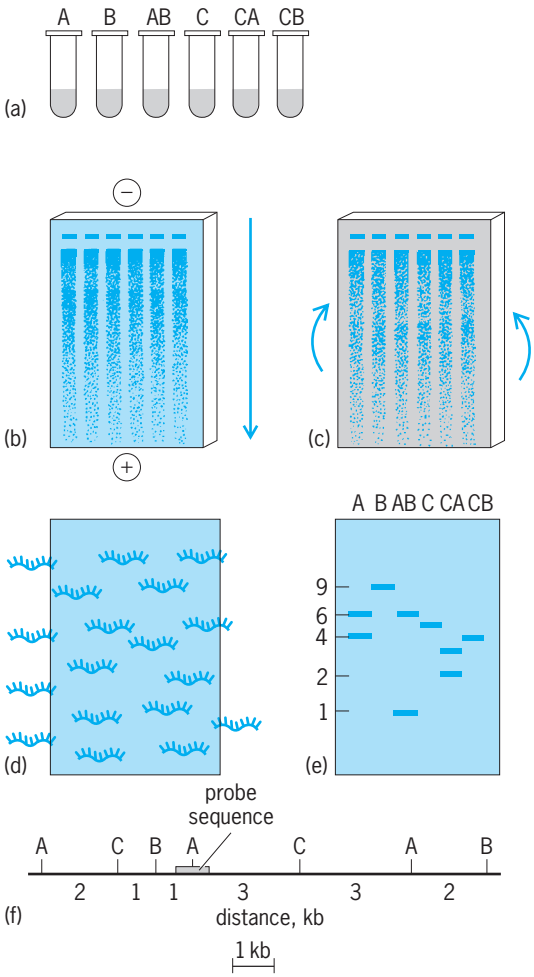

Fig. 1. Southern blotting. (*a*) Genomic DNA is digested singly and doubly with restriction enzymes A, B, and C. (*b*) Digested DNA is fractionated by size-dependent migration through a gel matrix in an electric field. (*c*) Denatured DNA is transferred and immobilized onto a nylon membrane. (*d*) Filter is incubated with a radiolabeled single-stranded probe in aqueous solution. (*e*) Duplexes formed between the single-stranded probe and genomic DNA are detected by autoradiography, showing the calculated size of DNA fragments. (*f*) Comparison of the products obtained from single and double digestion allows the construction of a restriction map (1 kb = 1000 base pairs).

by electrophoresis through a gel matrix. The DNA is transferred to a membrane and probed with a DNA segment near the polymorphic site (**Fig. 2**). *See* RESTRICTION ENZYME.

Short tandem repeat (microsatellite) markers. The development of short tandem repeat markers, or microsatellites, spanning the genome greatly facilitated human linkage analysis, and the era of positional cloning came to the forefront. Short tandem repeats are repetitive DNA sequences of two to five nucleotides that are scattered throughout the genome in abundance. They are often used as genetic markers because they tend to vary between individuals and can be easily typed using a conjuction of PCR assays and get electrophoresis. One of the first efforts of the Human Genome Project was to identify and map at least one Short Tandem Repeat marker per centimorgan (cM). Currently, more than ten thousand Short Tandem Repeats have been mapped at high density to produce genetic maps accessible through a web interface. *See* POLYMORPHISM (GENETICS).

Applications. Linkage maps provide the starting point for gene mapping. By establishing linkage between a gene and known polymorphic markers, the position of that gene can be determined. That technique is particularly useful for mapping of genes without a known biochemical function, such as those involved in diseases. Furthermore, closely linked markers can be used for prenatal diagnosis in families with a hereditary disease. The cosegregation of a linked marker with a disease gene permits the prediction of whether an offspring has the disease gene simply by following the inheritance of the marker. Linkage analysis also permits the simultaneous mapping of multifactorial traits to many chromosomal regions. *See* HUMAN GENETICS.

Association studies. Linkage analysis has been very effective in identifying rare, single-gene disorders (or Mendelian disorders) but has been less successful at locating genes responsible for common complex diseases, such as cardiovascular disease or cancer. In common diseases, each gene variant contributes only modestly to disease risk, and thus detection of these small effects is difficult unless the number of families studied is very high.

A complementary approach to linkage mapping is to identify disease genes by testing for an association between a specific polymorphic variant and the disease, that is, to conduct an association study. This is done by comparing the frequency of the specific polymorphism in a group of affected individuals with the frequency of the selected polymorphism in a group of unaffected controls.

There are two approaches to association studies: the "direct approach," in which each putative causal variant is tested for correlation with the disease and the "indirect approach," in which the polymorphic variants are used as genetic markers to detect the association between a particular genomic region and the disease.

In order to provide the tools needed to conduct association studies in an efficient and cost-effective manner, the international HapMap project was created.

The aim of this project is to determine the common patterns of DNA sequence variation in the human genome by the characterization of sequence variants, their frequencies, and the correlations between them in DNA samples from populations of

Fig. 2. DNA sequence polymorphisms. Restriction-fragment-length polymorphisms. A polymorphic restriction site R* is detected by the probe. A 3-kb fragment is produced when the site is absent, and a 2-kb fragment when it is present.

African, Asian, and European descent. The tools generated will allow the application of the indirect association approach to any functional candidate gene, any region suggested by linkage studies, or ultimately to whole genome scans searching for disease risk factors.

The polymorphic DNA markers of choice for association studies are single nucleotide polymorphisms. Single nucleotide polymorphisms are the most abundant DNA variation in humans (occurring on average in 1/300 bases), and they are caused by a single base change in the genome sequence. They are found both in genes and noncoding sequences, and some of them are expected to be the cause for common diseases.

A variety of single nucleotide polymorphism genotyping technologies have been developed recently. They vary in their assay formats, detection methods, and throughput capabilities. Some of the methods are capable of generating from several thousand to well over one million single nucleotide polymorphism genotypes per day, making the analysis of whole genomes possible. These technologies combine robotics, multiplex assays, and specialized software to deliver high-confidence and highly accurate genotypes.

DNA sequencing. Recombinant DNA technology has led to the mapping and analysis of genes at a molecular level. Fundamental to this has been the ability to produce large quantities of highly purified DNA sequences for further analysis by using molecular cloning techniques. Restriction enzymes can be used to insert DNA sequences into vectors propagated in bacterial hosts. Plasmid vectors consist of small, self-replicating circles of DNA, which can be used to clone segments of DNA of up to 10,000 bp (10 kilobases, or kb). Larger segments can be propagated in vectors derived from bacteriophage λ. Part of the λ genome is not essential for lytic growth and can be replaced with up to 20 kb of exogenous DNA. This recombinant-DNA molecule is incorporated into phage particles in the test tube and used in multiple rounds of infection of *Escherichia coli* cells. The DNA segments of 45 kb can be cloned into cosmid vectors, which contain only the sequences from bacteriophage λ necessary for incorporation in phage particles. However, these can be used only for a single round of infection and are often unstable.

In order to clone large genomic regions, two main cloning systems were developed: yeast artificial chromosomes and bacterial artificial chromosomes. Yeast artificial chromosomes clones contain DNA sequence inserts of more than 1 Mb (megabase) in size. They were used to constuct the first human and mouse physical maps. However, yeast artificial chromosomes are not very stable and have a tendency to rearrange their insert content. Bacterial artificial chromosomes clones have smaller DNA sequence inserts (100–200 kilobases), but they are more stable and have proven extremely useful for constructing second-generation physical maps of large genomes. Bacterial artificial chromosomes became the cloning vector of choice for the mapping of the human genome and the mapping of the genomes of the model organisms that followed.

The order of nucleotides in a DNA fragment can be determined by DNA sequencing. In the Sanger method (also known as the dideoxy chain termination sequencing method), a complementary sequence to a single-stranded template is synthesized by the enzyme DNA polymerase. Synthesis of the complementary strand is terminated at specific residues by the incorporation of a modified nucleotide. A separate reaction is performed for each nucleotide and results in a family of molecules, each of which terminates at that nucleotide. These products can be fractionated on the basis of size by electrophoresis through an acrylamide gel, and the sequence can be read from a ladder of fragments of different size (**Fig. 3**). *See* ELECTROPHORESIS.

During the sequencing efforts of the Human Genome Project, crucial enhancements were made toward increasing the accuracy and efficiency of the Sanger method. Advances in laser instrumentation, allowing for better automated detection of fluorescently labeld DNA molecules; the development of capillary-based sequencing instruments capable of analyzing 96 samples in parallel; and improvements

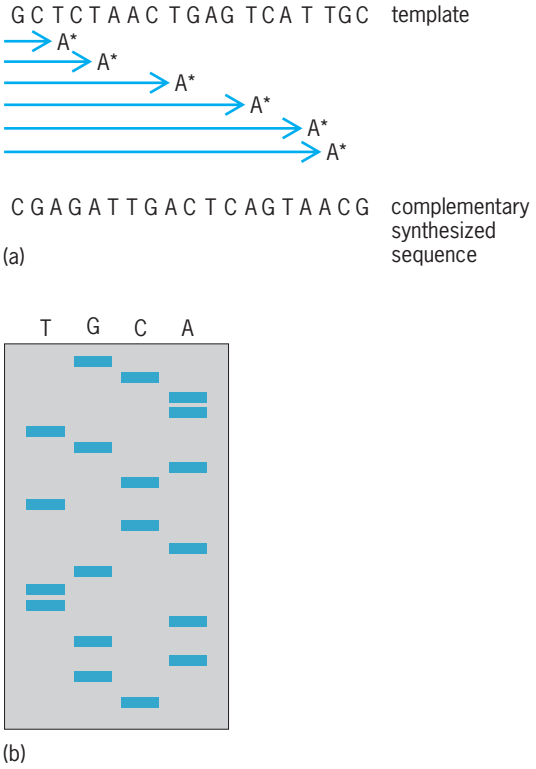

Fig. 3. DNA sequencing. (*a*) Complementary strand of DNA is synthesized to a single-stranded template by the enzyme DNA polymerase. Chain growth is terminated upon the incorporation of a modified nucleotide, in this case A*, which results in a family of fragments each of which terminates in the residue A. A similar reaction is performed for each of the other three nucleotides, T, G, and C. (*b*) The products of each reaction are fractionated according to size by electrophoresis. The order of nucleotides is read from the bottom of the gel (the smallest fragment) to the top.

in the biochemistry of the reaction made possible the dramatic increase in throughput (500–800 quality bases per reaction; 1.5 million bases per 24 h) and at least a 100-fold reduction in costs. In addition, software systems were developed for the analysis and quality assessment of the primary sequence data, and viewers were created for inspection and editing of the resulting sequence assemblies.

DNA sequencing continues to evolve toward faster, more affordable, and more sensitive technologies. In particular, nanotechnology approaches—capable of a high level of miniaturization, automation, parallelization, and integration—are making progress toward this end. These improvements will not only allow sequencing of entire new genomes at a reasonable cost but will also open the possibility of sequencing additional human genomes. The ability to decode the genomic sequence of virtually any organism will provide unparalleled access to the fundamentals of biology. The ability of resequencing the human genome without restrictions will have an enormous impact on understanding human genetic variability and the current approaches to manage and cure disease.

Sequence mapping. The expression of a gene is dependent upon the synthesis of a messenger RNA (mRNA) molecule and its translation into the amino acid sequence of a protein by ribosomes in the cytoplasm. The DNA copies of mRNA (complementary DNA, or cDNA) from cells are synthesized by using the viral enzyme reverse transcriptase, and cloned for further analysis. Comparison of cDNAs with genomic sequences reveals that the DNA encoding many eukaryotic genes is discontinuous. The blocks of coding sequences (exons) of a gene are brought together in the mRNA molecule by the splicing out of intervening sequences (introns). The presence of introns results in genes being distributed over a much larger stretch of DNA than would be expected from the analysis of their mRNA or protein sequences. The size and number of exons and introns vary widely between genes. For instance, the β-globin gene contains only two introns of 14 and 331 bp and encodes a 620-bp mRNA. In comparison, the gene disrupted in Duchenne's muscular dystrophy encodes a 16-kb mRNA that comprises at least 60 exons and extends over 2×10^6 bp of DNA. *See* EXON; INTRON; RIBONUCLEIC ACID (RNA).

Before the Human Genome Project, mapping of the intron-exon structure of a gene was approached by using restriction enzymes. Restriction sites present in the genomic DNA but absent in the cDNA indicate the presence of some introns. However, precise mapping of the boundaries of the intron-exon structure can be achieved only by obtaining the genomic DNA sequence.

In the past, this required the isolation of large sections of genomic DNA by chromosome-walking techniques. A library for this is constructed by partially digesting genomic DNA with a frequently cutting restriction enzyme. This produces a series of overlapping DNA fragments that cover the entire genome. These are cloned into bacteriophage or cosmid vectors, and the resulting library is screened by hybridization with a radioactive probe. Positively reacting clones that contain genomic DNA surrounding the probe are purified and restriction-mapped. New probes are isolated from the ends of these clones and used to obtain a second set of clones that overlap the original clone. Repetition of this process results in the isolation of a series of clones that contain an ordered set of overlapping genomic fragments. *See* BACTERIOPHAGE; GENETIC ENGINEERING.

The isolation of genomic sequences surrounding a gene by chromosome walking may result in the mapping of gene clusters. For example, the human β-globin gene family stretches over 50 kb on chromosome 11. It contains five related functional genes that are active at different stages in human development. All of the genes are composed of three exons and probably arose from a single ancestral gene. Duplication of the ancestral gene by unequal crossing-over followed by sequence divergence over time has resulted in the evolution of five related but distinct genes. A homologous but functionally inactive β-globin pseudogene was also mapped within this region. Pseudogenes are genes that were presumably once active but have become defunct from the accumulation of mutations. The comparison of sequences surrounding genes has led to the identification of sequences involved in the regulation of genes. *See* PARASEXUAL CYCLE.

With the map and sequence of the human genome available, most of the steps described above can be circumvented, and the mRNA sequence can be compared with the genomic sequence without any additional experimental work. The exon-intron boundaries can be defined using sequence homology software, which compares the mRNA sequence to the publicly available human genome sequence.

Mapping molecular mutations. Many mutations are the result of small, subtle changes within the DNA sequence of a gene. The substitution of a single base pair within the reading frame of a gene may result in the incorporation of a different amino acid in the protein or premature termination of mRNA translation. These changes can eliminate or drastically alter the activity of a protein. For example, the reduced affinity of hemoglobin for oxygen in sickle cell anemia is caused by a single amino acid substitution resulting from a single base change in the β-globin gene. These mutations, as well as others, are mapped by comparing the DNA sequence of the mutant gene with that of the wild type. The sequence of specific regions of the human genome in large populations can be ascertained by the selective amplification of those regions by using PCR. Oligonucleotides flanking the target sequence are used as primers for repetitive cycles of DNA synthesis of that sequence in the test tube (**Fig. 4**). The resulting amplified product can be directly subjected to DNA sequence analysis. Alternatively, the presence or absence of a known polymorphic restriction site within the product can be tested directly by enzyme digestion

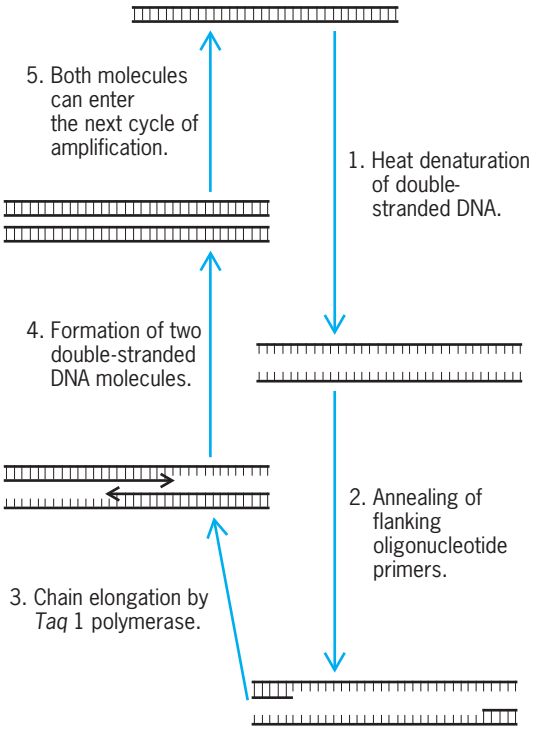

Fig. 4. Polymerase chain reaction. Regions of DNA defined by flanking oligonucleotides are amplified by successive cycles of denaturation, primer annealing, and chain elongation. Ideally, each round of amplification would double the target sequence. The resulting product can be analyzed with restriction enzymes and by DNA sequencing.

and gel electrophoresis. This simplifies the typing of individuals for restriction-fragment-length polymorphisms in linkage analysis. Furthermore, the magnitude of DNA amplification by polymerase chain reaction permits the analysis of single sperm cells. Each sperm cell is the product of a single meiosis. The amount of recombination occurring between two polymorphic loci can be accurately measured by simultaneously amplifying and typing those loci in a large number of individual sperm cells. Although this analysis is restricted to DNA sequence polymorphisms, it may provide a much simpler and more accurate method of measuring the genetic distance separating such loci than is offered by extensive family studies. *See* GENE AMPLIFICATION; MEIOSIS.

Reverse genetics (or positional cloning). Traditionally, the mapping and isolation of genes was dependent upon prior knowledge of the activity or amino acid sequence of the respective protein products. Before 1980, very few human disease genes had been identified, because for the majority of genes, including those involved in genetic disorders, this information was not available. Such genes are isolated by reverse genetics, also called positional cloning. This methodology identifies disease genes based solely on their approximate chromosomal location. The first stage in positional cloning involves linkage mapping of the disease to the smallest possible region of the genome in affected families. Next novel candidate genes are identified in the region of linkage. The final step is to show that patients have mutations in a gene that are not present in control subjects. Before these steps could be completed, researchers had to produce high-resolution genetic maps and build clone contigs in order to identify and sequence the genes of interest. This process is now easier to implement due to the resources—maps, clones, sequence, expression and phenotypic data—provided by the Human Genome Project as well as other genome research projects. Thus, the experimental part is reduced to linkage mapping of the disease gene to a region of the genome and to the selection and screening of the candidate genes for mutations in patients.

Candidate genes can be identified by the presence of unmethylated CpG-rich stretches of DNA, the conservation of DNA sequence with other animal species, or the identification of regions that are transcribed into mRNA. The protein sequence of these genes can be obtained by the isolation and sequence analysis of their cDNAs, providing a starting point for examining the gene function. Often the disease gene can be identified by its disruption of activity in affected individuals. Many genes, including those involved in Duchenne's muscular dystrophy, sex determination, and chronic granulomatosis disease, have been mapped and isolated by reverse genetics. *See* GENETIC CODE; MOLECULAR BIOLOGY.

Matthijs Smith; Leroy Hood

Bibliography. R. DeSalle and M. Yudell, *The Genomic Revolution: Unveiling the Unity of Life*, Joseph Henry Press, 2002; B. Lewin, *Genes*, 2004; G. S. Stent and R. Calender, *Molecular Genetics: An Introductory Narrative*, 1978; R. K. Saiki et al., Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase, *Science*, 239:487–491, 1988; T. Strachon and A. P. Read, *Human Molecular Genetics*, 3d ed., 2003; J. D. Watson et al., *The Molecular Biology of the Gene*, 5th ed., 2000; T. J. White, N. Arnheim, and H. A. Erlich, The polymerase chain reaction, *Trends Genet.*, 5:185–194, 1989.

Genetics

The science of biological inheritance, that is, the causes of the resemblances and differences among related individuals.

Genetics occupies a central position in biology, for essentially the same principles apply to all animals and plants, and understanding of inheritance is basic for the study of evolution and for the improvement of cultivated plants and domestic animals. It has also been found that genetics has much to contribute to the study of embryology, biochemistry, pathology, anthropology, and other subjects. *See* BIOCHEMISTRY; EMBRYOLOGY; PATHOLOGY; PHYSICAL ANTHROPOLOGY.

Genetics may also be defined as the science that deals with the nature and behavior of the genes, the fundamental hereditary units. From this point

of view, evolution is seen as the study of changes in the gene composition of populations, whereas embryology is the study of the effects of the genes on the development of the organism. *See* POPULATION GENETICS.

Study and analysis. The geneticist uses various methods of study, nearly all of which rest on the presence of differences between individuals. These techniques do not make it possible to analyze the portions of the hereditary makeup of individuals for which differences cannot be found. If differences occur, it is also necessary that they be transmitted to later generations; and in general the full analysis will also require the presence in the material of sexual reproduction or some analogous process that allows the recombination of inherited properties from different individuals. In brief, what the geneticist usually does is to cross diverse individuals and study the descendants (or in such material as humans, analyze existing pedigrees). Such a study must usually be carried through at least two successive generations and requires enough individuals to establish ratios between the classes present, since conclusions are usually based on these. The mechanism of heredity is such that its analysis requires the use of probability theory, although for much of the subject only very simple statistical concepts are needed. *See* BIOMETRICS; BREEDING (ANIMAL); BREEDING (PLANT); CHROMOSOME; CHROMOSOME ABERRATION; HETEROSIS; HUMAN GENETICS; MENDELISM. A. H. Sturtevant

Molecular genetics. A major discipline within the field of genetics that describes the basis of inheritance at the molecular level, molecular genetics had its beginnings in the 1940s. At that time, work on simple, single-celled organisms revealed that the genetic material in most organisms is deoxyribonucleic acid (DNA), a very long polymer made up of four different nucleotide monomers: adenosine (A), guanosine (G), cytosine (C), and thymidine (T). A landmark discovery was the demonstration in 1953 by James Watson and Francis Crick that the genetic material consists of two DNA strands wrapped around each other to form a double helix. The two strands are held together by specific noncovalent, weak interactions between the nucleotides: A always binds to T, and G always binds to C. These long, double-stranded DNA molecules make up chromosomes, which are divided into thousands of different functional units called genes. The genes specify the structure and function of an organism. *See* ALLELE; DEOXYRIBONUCLEIC ACID (DNA); GENE; NUCLEIC ACID.

The field of molecular genetics focuses on two general questions: how do genes specify the structure and function of organisms, and how are genes replicated and transmitted to successive generations? Both questions have been answered. Genes specify organismal structure and function according to a process described by the central dogma of molecular biology: DNA is made into ribonucleic acid (RNA), which specifies the structure of a protein, a polymer of 20 different amino acids that carries out a particular function. In the first step of gene expression, DNA is transcribed into messenger RNA (mRNA)—a single-stranded polymer of nucleotides very similar to DNA—by the enzyme RNA polymerase. In the second step, the mRNA molecule serves as a template for protein synthesis, which is carried out by a complex machinery that comprises a particle called a ribosome and special adapter RNA molecules called transfer RNA. *See* RIBONUCLEIC ACID (RNA); RIBOSOMES.

The specific sequence of nucleotides in the DNA (A, T, G, C) specifies gene function in two ways. First, particular sequences signal the RNA polymerase to begin and end synthesis of the mRNA. In addition, specific DNA sequences ensure that RNA polymerase transcribes certain genes only in the correct cell type at the proper time. A major goal of molecular genetics is to learn how DNA sequence determines the regulation of gene expression. The second way in which the nucleotide sequence specifies gene function is by determining the structure of the encoded protein. The sequence of the mRNA is read by the ribosome in groups of three nucleotides, each of which is called a codon. The four different nucleotides are arranged in 64 codons that specify the 20 amino acids in proteins. Therefore, the DNA sequence determines the structure and function of an organism. Molecular geneticists have learned how to determine large amounts of DNA sequence with relative ease. Large-scale projects have been undertaken to determine the complete DNA sequence of several organisms, including humans. Although it is clear how the sequence of DNA determines the sequence of protein, a major challenge remaining for molecular geneticists is to explain how the sequence of a protein determines its particular function. *See* GENETIC MAPPING.

The structure of DNA provides a simple mechanism for genes to be faithfully reproduced: the specific interaction between the nucleotides means that each strand of the double helix carries the information for producing the other strand. That transfer is accomplished by several enzymes, including DNA polymerase. It copies each DNA strand into two molecules of double-stranded DNA that are apportioned to each new cell during cell division. Rarely, DNA polymerase makes a mistake during replication, or the DNA becomes damaged, and an altered gene results. Such mutations can change the function of the encoded protein and thereby lead to evolutionary change. In addition, DNA molecules can recombine through complex processes that involve the breakage and rejoining of two molecules and lead to different gene combinations that can alter the structure and function of the organism. *See* GENETIC CODE; GENETIC ENGINEERING; MOLECULAR BIOLOGY; MUTATION. Mark Johnston

Bibliography. M. Levitan, *Textbook of Human Genetics*, 3d ed., 1988; B. Lewin, *Genes*, 3d ed., 1987; G. M. Malacinski and C. Simonsen, *Molecular Genetics of Mammalian Cells: A Primer in Development Biology*, 1986; G. S. Stent and R. Calender, *Molecular Genetics*: *An Introductory Narrative*, 1978; J. D. Watson and N. H. Hopkins, *Molecular Biology of the Gene*, 4th ed., 1988.

Genomics

The comprehensive analysis of an organism's genome (the full complement of inherited information-bearing genetic material residing in an organism's chromosomes), including, but not limited to, its genes and the signaling codes regulating gene expression. As a discipline, genomics focuses on an organism's entire set of genes rather than on individual genes. The term was coined in 1986 to name a journal oriented to the mapping, sequencing, and analysis of genomes.

Typically, genetic material is synonymous with DNA, although some viruses use RNA rather than DNA to encode genomic information. DNA is a polymer consisting of four nucleotide building blocks, also referred to as bases, called A (adenine), C (cytosine), G (guanine), and T (thymine). During the 1950s and 1960s, seminal advances in understanding DNA's double-helical structure, its mechanism of replication, and its code for specifying the amino acid content of proteins led to a significant insight: "Sequence"—the arrangement of nucleotides (bases) in some particular order along a DNA strand—is the key to deciphering a genome's information content (**Fig. 1**). This means that the information in the genome is digital; that is, it consists of strings of discrete letters whose order provides meaning or information just as the 0s and 1s in computer programs provide meaning. The field of genomics, therefore, is directed to the informative features embedded in a given genome's DNA sequence. In practice, a global genomic analysis requires an empirically determined DNA sequence called a "reference genome" for a species of interest. The best-known example of a comprehensive DNA sequence-based genomic analysis is the Human Genome Project. It is fair to say that no other single project has so transformed the practice of both biology and medicine. Owing to spectacular advances in DNA sequencing technology resulting from the Human Genome Project, genomes from hundreds, soon to be thousands, of species of medical, agricultural, industrial, and phylogenetic significance have been made available for biologists to explore.

Importance. Genomic information is crucial to understanding biology. In essence, a species' genes and regulatory circuitry determine possibilities for, and set constraints on, the structure and function of its individual members. Genomic programming allows an acorn to develop into an oak tree and a caterpillar into a butterfly. An acorn will not, and cannot, produce a butterfly, however, because there are significant differences in specialized genes and regulatory circuitry between plant and insect genomes. Genomic information allows biologists to catalogue and compare "parts lists" for different species. Parts lists might include DNA sequences identified as protein-coding and regulatory RNA genes. In addition, there are short sequence stretches called promoters, enhancers, and silencers that, with the help of proteins called "transcription factors," control whether, when, and how much a particular gene is transcribed. Parts lists also contain sequences that serve as origins of DNA replication and otherwise affect the functioning of chromosomes. In essence, differences within and between species at the nucleotide sequence level provide the informational foundation for a detailed molecular understanding of cellular function and organism development.

Scientific approach. Genomics aims to develop large-scale, high-throughput, cost-effective strategies for the analysis of genomes. These strategies are used first to decipher what a given genome's DNA sequence is and then to determine or predict what information content that sequence contains. Thus,

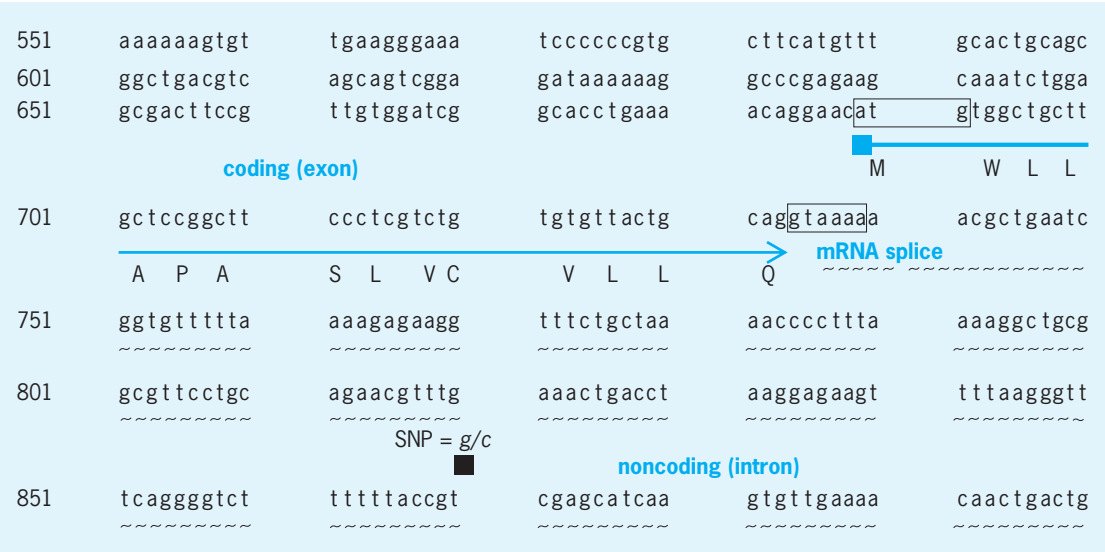

Fig. 1. Informative features within the DNA sequence. This stretch of sequence represents a 350-base portion of a pufferfish toll-like receptor gene containing the first exon and a portion of the first intron. The black square depicts the location of a sequence variation (polymorphism). The color arrow indicates the protein-coding sequence, with the amino acids depicted in capital letters.

the methodologies of genomics encompass two distinct activities:

1. *Sequence determination*: Using best-practice technologies to delineate the order of nucleotide building blocks in an organism's chromosomes or genome.

2. *Sequence interpretation*: Using computational and wet-laboratory procedures to extract meaningful biological information from a genomic sequence or sequences.

A detailed and comprehensive analysis of a species' genome requires an accurately assembled reference sequence for each of the chromosomes. Once the reference sequence has been determined, the location and frequency of polymorphic variations can be assessed. "Polymorphism" is the name given to the small changes in DNA sequence that distinguish a genome of one member of a species from the genomes of other members. Polymorphisms are identified by sequencing genomes, or portions of genomes, of two or more members of the species. From a technical standpoint, obtaining the initial reference sequence, called "de novo sequencing," is considerably harder than selectively "resequencing" genomes for discovering polymorphisms.

Ideally, a reference genome sequence will be of "finished" quality. In a finished genome, the accuracy of each base is greater than 99.99%. Perhaps more importantly, the sequence consists of gap-free contiguous strings of correctly ordered As, Cs, Gs, and Ts that correctly represent the organism's chromosomes. Several eukaryotic genomes have been finished—for example, human, fruitfly, worm, and baker's yeast—and hundreds of microbial genomes have been completely sequenced. The ease of obtaining a finished sequence for any given genome is a function of its size, complexity, and sequence composition. Some genomes are harder to finish than others because they contain multiple gene duplications, extremely high or low GC content, or regions that are difficult to propagate with the commonly used cloning vectors (agents into which small fragments of DNA are inserted for amplification prior to sequencing). Unfinished genomes are provided in "draft" form. Although their quality is lower than finished sequences, draft sequences nonetheless provide valuable information about a species' genomic content.

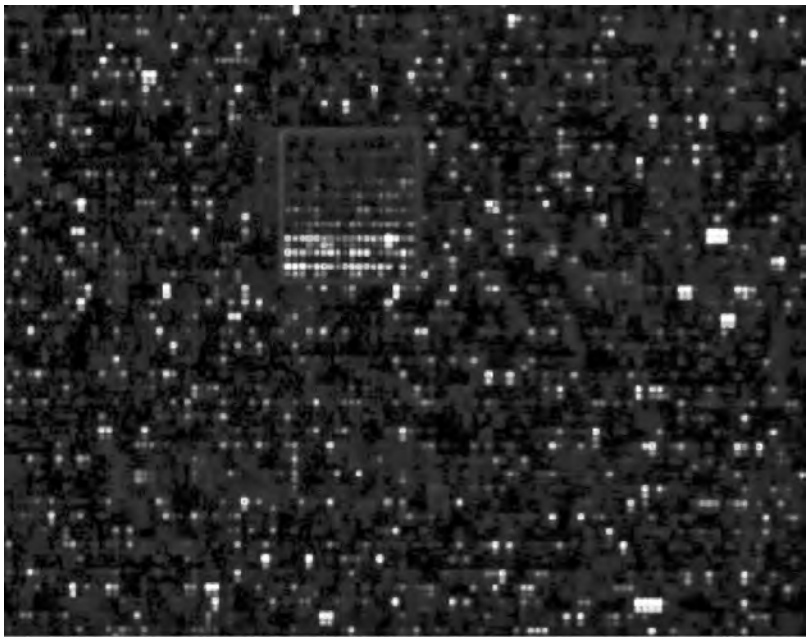

Fig. 2. Image of an Affymetrix U133 Plus 2.0 Genechip hybridized with biotin-labeled target derived from prostate luminal cells. The array contains probe sets that detect 47,000 transcripts that represents the human transcriptome. The intensity of signal in the image represents the abundance of the mRNA which the probe set targets. Signal intensity is depicted as a gradient from gray to white, where gray is the lowest signal and white is the greatest signal. (*Asa Oudes, Institute for Systems Biology*).

Once a reference genome sequence has been obtained, a variety of computational and experimental strategies are employed to identify its significant features. Analyses include but are not limited to the following types of information:

1. *Genomic landscape features*: Delineating gene models, order of genes along a chromosome, gene duplications, common repeats, hot and cold spots of recombination, and features related to chromosome organization and function.

2. *Regulatory signals*: Identifying sequence motifs that affect the occurrence and extent of transcription of specific genes or groups of genes under varying cellular conditions.

3. *Cross-species comparisons*: Identifying blocks of synteny (used here to mean regions where the order of the genes is approximately the same in the genomes or two or more species) using other reference genomes, relating a genome's information content to that of other species, and discovering highly conserved blocks of noncoding sequence that may be indicative of new and important types and functions of sequence information.

4. *Sequence variations*: Applying resequencing technologies to genomes of individuals within a species to catalogue single nucleotide substitutions, insertion-deletion polymorphisms, gene-copy-number polymorphisms, and chromosomal rearrangements.

5. *Disease association studies*: Identifying candidate genes whose malfunctions cause deleterious effects.

A large-scale international effort, called the ENCODE project (Encyclopedia of DNA elements) is under way to characterize and identify functional features of the human genome sequence. It is anticipated that approaching the problem from a variety of experimental and computational perspectives will illuminate the information in the human genome in ways that cannot currently be foreseen.

Genomics and systems biology. Genomics has stimulated the development of numerous systems biology-related technologies aimed at examining gene expression profiles, gene/protein interactions, and regulatory networks. These high-throughput technologies perform multiple measurements simultaneously so that interactions can be assessed in a

comprehensive manner. For example, microarrays take advantage of the ability of RNA to hybridize to known stretches of complementary DNA placed in defined locations on a chip. Microarrays are used to compare, under varying conditions, the increases and decreases in expression of thousands of gene transcripts (mRNAs) in parallel. Transcripts from samples of cells or tissues can also be identified and assessed quantitatively based on high-throughput parallized measurements (that is, measurements made in parallel) of short, unique signature sequences. DNA microarrays are also being used to identify regions of sequence that are bound by transcription factor proteins under experimentally specified conditions. The results reveal sequence signals important to regulating biological functions or disease pathology (**Fig. 2**). Application of these technologies and others in a concerted fashion is illuminating genetic, metabolic, signaling, and regulatory networks. These networks govern a cell's responsiveness to its internal developmental programming and external environmental perturbations. Studies of such networks are termed systems biology.

Moving forward. Advances in sequencing technology have vastly increased throughput since the Human Genome Project was begun in 1990. Significant improvements in base-call accuracy, read length, automation, machine capacity, instrument run time, and assembly algorithms have made de novo genomic sequencing both cheaper and easier to do. However, in terms of future needs, enormous reductions in cost are required. To this end, several "new generation" sequencing technologies based on highly parallel sequencing (millions of samples at one time) or single molecule sequencing are being developed. These aim to improve the efficiency and throughput of de novo sequencing, polymorphism detection, and transcriptional profiling while greatly decreasing cost and making affordable personal human genome sequences a possibility. Individual human genome sequences will be a key for predictive medicine of the future.

Inspired by this optimism, an ambitious project surpassing the scale of the Human Genome Project was launched at the beginning of 2006. This effort, named "The Cancer Genome Atlas" (TCGA), has the goal of identifying key mutational events responsible for oncogenesis and metastasis in various types of cancerous cells. Because cancerous cells accumulate thousands of mutations during their rounds of propagation, it is a daunting task to identify those key changes that account for differences in malignancy among the different cancer types. However, a more detailed molecular and systems understanding of how cellular networks are perturbed in various types of cancer will undoubtably lead to more effective drug targets and therapies.

As a logical extension of The Human Genome Project, intensive efforts are under way to construct detailed "haplotype maps." These maps delineate polymorphic variations that group together in "blocks or haplotypes"of sequence shared by differing human subpopulations. Advances in technologies for identifying polymorphisms and genotyping individuals have promoted an increased interest in "personalized medicine"—the right treatment for the right patient at the right time. Personalized medicine relies on the ability to coordinately correlate the sequence polymorphisms of multiple genes with specific effects, such as susceptibilities to diseases, the efficacy of pharmacological interventions, and the severity of side effects.

The explosion of sequence information produced by genomic technologies demands progress on two additional fronts: development of statistical algorithms and computational modeling tools for interpreting the information; and examination of the social, legal, and ethical implications of genomic research. *See* DEOXYRIBONUCLEIC ACID (DNA); GENE; GENE ACTION; GENE AMPLIFICATION; GENETIC CODE; HUMAN GENOME PROJECT; MOLECULAR BIOLOGY.

Lee Rowen; Leroy Hood

Bibliography. L. Gonick and M. Wheelis, *The Cartoon Guide to Genetics*, rev. ed., Harper-Collins, New York, 1983, updated 1991; L. Hood et al., *Biological Information and the Emergence of Systems Biology*, Roberts and Company, 2006; The sequence of the human genome, *Science*, 291, no. 5507, Feb. 16, 2001; G. Smith, *The Genomics Age: How DNA technology Is Transforming the Way We Live and Who We Are*, Amacom, New York, 2005; Initial sequencing and analysis of the human genome, *Nature*, 409:860–921, 2001.

Gentianales

An order of flowering plants (angiosperms) in the euasterid I group of the asterid eudicotyledons. The order consists of five families and approximately 17,500 species. The circumscription of Gentianales has been widened to include Rubiaceae, largely on the basis of deoxyribonucleic acid (DNA) sequence data. Circumscription of Loganiaceae has also changed on the basis of these data, with many taxa, including the ornamental *Buddleja*, being transferred to Lamiales. The order is characterized by opposite leaves, frequent occurrence of alkaloids, and the presence of internal phloem in the wood (not in Rubiaceae).

Apocynaceae (including Asclepiadaceae), approximately 5000 species, are mostly tropical and subtropical and have a well-developed latex system, superior ovary, thickened and apically modified style, and carpels that are usually united only by the common style or stigma. *Asclepias* (milkweed), *Catharanthus* (rosy periwinkle, the source of the powerful anticancer drugs vincristine and vinblastine), *Hoya* (wax flower), *Nerium* (oleander), and *Stapelia* (carrion flower) are well-known genera. Rubiaceae, approximately 11,000 species, are cosmopolitan, but particularly important in the tropics. They have stipulate leaves and an inferior ovary. The widespread occurrence of alkaloids makes the family economically

significant; the major products are coffee (*Coffea*) and the antimalarial drug quinine (*Cinchona*). Other genera, including *Gardenia*, *Ixora*, and *Pentas*, are widely used in horticulture. Gentianaceae, approximately 1200 species, are principally temperate and subtropical. They often contain bitter iridoid compounds, and some are used medicinally or in bitter alcoholic beverages (such as *Gentiana*, gentians). Several genera are widely grown ornamentals, including *Exacum* (Arabian violet), *Gentiana*, and *Lisianthus* (prairie gentian). The remaining two families are relatively small but contain some genera of economic importance. *Strychnos* (Loganiaceae) is the source of strychnine and the arrow poison curare, and Gelsemiaceae include the ornamental *Gelsemium* (allspice jasmine). *See* ASTERIDAE; MAGNOLIOPSIDA; PLANT KINGDOM; RAUWOLFIA; STROPHANTHUS; STRYCHNOS.

Mark W. Chase

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

8 GEO-HYS

www.MHEST.com

G

Geochemical prospecting — Gyrotron

Geochemical prospecting

Exploration for mineral deposits by chemically analyzing sampled rock, soil, vegetation, and other natural materials for trace amounts of the principal and associated elements. By now, most of the easily discovered mineral deposits have been found, making the search for new deposits ever more challenging. In spite of this, worldwide demands are rising for energy, metallic, and nonmetallic mineral resources. Search continues in regions of known mineralization and in areas where deposits are covered by sediment or obscured by vegetation and ice. Geochemical techniques have become an important part of almost all exploration programs. *See* MINERAL; PROSPECTING.

Types of mineral deposits. Mineral deposits occur in two broad categories: syngenetic and epigenetic. Syngenetic deposits form at the same time and by the same processes as the rocks in which they occur. They are a natural product of rock formation, and an understanding of how the rocks form is crucial to knowing where to look for important mineral concentrations. Epigenetic deposits are superimposed on rocks and may be completely unrelated to the processes that formed the surrounding rock (country rocks). Training and experience are important in evaluating both syngenetic and epigenetic mineralization. This is especially true if deposits have been exposed to surface weathering, or if they are covered by vegetation or the debris left behind when glaciers melted. *See* FORMATION; ORE AND MINERAL DEPOSITS.

Syngenetic and epigenetic mineral deposits often are exposed at the surface through prolonged physical and chemical weathering and erosion, and many deposits are completely destroyed as rock systems evolve naturally. The presence of syngenetic deposits can be deduced because the mineralization was part of a larger geologic setting, while epigenetic deposits may leave no clue to their former existence. *See* EROSION; WEATHERING PROCESSES.

Compared to the environment in which mineral deposits formed, the overall target size for explorationists is greater since they are looking not just for the mineralization but also for associated minerals and features that could be identified as important accompanying mineralization. Placer deposits are the exception to the rule. They are concentrations of chemically and mechanically durable, high-specific-gravity minerals that may not be associated in meaningful ways with the surrounding geology. Even in such settings, explorationists may be able to predict the presence of such deposits by looking at the regional setting of stream sediments, glacial deposits, and so on.

Once formed, both syngenetic and epigenetic deposits may be affected by tectonic adjustments (such as faulting, uplift, tilting) and by weathering and erosion as they are exposed at the surface. Thus, explorationists are forced to rely not only on their knowledge of the geologic setting of deposits but also on how mineralization may be affected by surface processes. Deposits may be covered by soil and vegetation or by residual weathering products in regions where soils are not well developed. The cover may be many meters thick, and both soil and residual material may be covered by surface debris and glacial deposits, and sometimes even by ice.

As the postmineralization history of deposits lengthens, the complexity and thickness of weathering products may increase, as does the chance that deposits will be covered by sediment and vegetation. It is increasingly difficult to "see through" the cover as the time of exposure increases, and the training, experience, and insight of the explorationist becomes more important. Augmented by technology, exploration in such areas can still be fruitful.

Brief history. The search for mineral deposits goes back to when humans began using resources. Just as today, early individuals likely noted the coloration of rocks and of soils near rocks that carried mineralization. In some cases, they may even have associated coloration with faults and fractures in the

rocks, even to the point of correlating significant mineralization with intersections of linear features. Recognition of the relationship between leached and oxidized caprocks (gossans) and concealed mineralization at depth could not have been far behind.

References in very early literature to the use of stream water and water from springs and seeps as guides to ore attest to the curiosity and intelligence of explorationists as early as the sixteenth century. Most notable is *De Re Metallica* by Georgius Agricola, published in 1556.

By the latter half of the nineteenth century, workers had recognized the association of mineral deposits with centers of igneous activity. The regular zoning of metals with respect to pressure and temperature was becoming obvious. At about this time, wall rock alteration—changes in the mineral composition and appearance of rocks adjacent to mineralization—was recognized in bordering veins. Alteration later would include pervasive changes in the mineralogy and chemistry of large volumes of rock that completely encompass deposits. *See* IGNEOUS ROCKS.

The modern era of geochemical prospecting began in the early 1930s in Europe, specifically in the Soviet Union, and shortly thereafter in Scandinavia. The Soviet Union had a large geological survey for research on sampling protocols and sample survey design, which coincided with the development there of rapid multiple-element emission spectrographic analysis. Sample media included soil, rock, and vegetation. The analyses were semiquantitative, with an error of as much as 30% of the value. But in the regional context of many of the surveys, this would highlight broad chemical anomalies, even specific anomalies in many instances. *See* EMISSION SPECTROCHEMICAL ANALYSIS.

Area selection. Explorationists look first for direct evidence of mineralization, such as fragments of mineralization or rocks that are cut by mineralization. As the obvious deposits were mined out, searchers were forced to consider mineralization processes, including the important characteristics of deposits, metals associated with specific minerals, and types of rocks hosting deposits, as well as how the country rocks were affected by the mineralizing processes. Explorationists focus on genetically significant features, regardless of the type of mineralization. Characteristics that may be interesting but not particularly significant are eliminated, and efforts are concentrated on conceptual and quantitative information until broad targets, known as area selection, emerge. Careful selection of areas minimizes the cost of regional surveys by focusing on targets with the greatest potential.

Anomaly identification. As defined in the *Glossary of Geology* (1987), an anomaly is "a geological feature, especially in the subsurface, distinguished by geological, geophysical, or geochemical means, which is different from the general surroundings and is often of potential economic value." In exploration, especially regional exploration, investigators are looking for any hint that mineralization may be present, and always are alert to "good luck spikes" where samples come directly from zones of mineralization. Such spikes are not required if the sample program has been properly designed, and even with well-designed programs it is possible to miss mineralization.

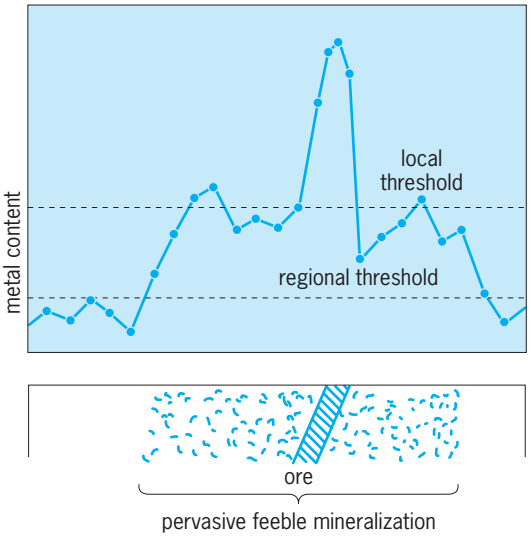

Fig. 1. Recognition of regional and local anomaly thresholds is the basis for geochemical exploration. (*From A. W. Rose et al., Geochemistry in Mineral Exploration, 2d ed., Academic Press, London, 1979*)

Figure 1 illustrates a fundamental characteristic of mineral deposits; that is, mineralization almost always lies within genetically related altered and mineralized rocks, ranging from centimeters thick for some veins to hundreds of meters around some deposits. Such zoning reflects the intensity, complexity, and duration of the mineralization processes. Knowledge of hydrothermal alteration and metal zoning can be used to great advantage, and information on structural orientation, host rock type, and episodes of hydrothermal activity can dramatically affect the size and shape of the target, or regional threshold (Fig. 1). Elements that are associated genetically with the principal target may be more mobile and thus more widely dispersed than the target within the hydrothermal environment (regional threshold). *See* HYDROTHERMAL ORE DEPOSITS.

In geochemical prospecting, indicator elements are the principal targets of exploration efforts, and the pathfinder elements are elements that occur with the indicator elements but often are present in greater quantities and/or exhibit greater mobility and thus greater dispersion within the host rocks. Pathfinder elements will be present in the regional threshold and perhaps within the local threshold (Fig. 1).

Pathfinder elements are likely to be more widely dispersed in weathering environments than indicator elements. They may be liberated more easily during weathering and more widely dispersed than indicator elements. Dispersion during weathering may distort the bedrock pattern considerably (for

example, downslope and downstream), and the signature can become too diffuse to be recognizable, if the distance is too great.

If a regional threshold is discernible, the local threshold can be located with follow-up sampling (Fig. 1). Knowledge of the geology and the characteristics of the suspected target become critical as follow-up sampling is done. Sample frequency, perhaps even the configuration and type of samples being collected, may change at this point; for example, going from broadly spaced sediment samples at stream intersections to soil and bedrock samples in a region of iron-oxide-stained bedrock.

Vectoring in on targets. The goal of all exploration is to "zero in" on economic mineralization. Explorationists often call this process "vectoring in," in which geologic knowledge is used together with geochemical data and perhaps geophysical clues to point in the direction where additional data should be gathered, including downward when all factors have vectored in on specific locations.

Well-designed sample programs will include orientation information, by which a guess is made as to the type of deposit being sought, the samples that will have to be collected, and the spacing of samples (for example, grid size for soil samples). A lot of information is available in the literature and company files that can be used to design sampling and analytical protocols, and to help establish a background against which regional and threshold concentrations may be compared. The geology and mining history of a region often can be accessed in the library, to which may be added Landsat information and multispectral data from airborne and satellite platforms. Regional geophysical data on rock magnetism and density, plus electrical conductivity and radioactivity, also may be available. *See* GEOPHYSICAL EXPLORATION; REMOTE SENSING.

Sampling large areas ($\times 10^5$ km^2) usually involves collecting sediments from streams that drain the region, to which may be added data from mineral springs and seeps, even sediment from lakes that are oriented across suspected mineral trends. The confluences of streams are sampled as far upstream as practical. A sample is taken below the intersection or a suite of samples is collected, with one sample from each stream above an intersection and perhaps a third sample from below the intersection. The pH of the water may be recorded and the water may even be sampled, but the sediment is the most reliable medium to test for mineralization. Large fragments are screened out, and the remaining sediment is panned down to a heavy concentrate. The concentrate will provide an indication of mineralization in the drainage, but sieved fractions will allow the data to be refined. For example, iron oxide and manganese oxide coatings are known to scavenge heavy metals that concentrate in the fine fractions of stream sediments.

Land position. Early in the evaluation process, it is necessary to obtain a land position if the preliminary indications of mineralization are promising. Because of the capricious nature of mineral occurrences, it is always a good idea to secure enough land to account for heterogeneities in occurrence. It also may be a good idea to make sure the best available target is being leased or claimed, as well as to acquire every spot that exhibits the same characteristics as the one of initial interest, since mineral occurrences often are clustered together.

Target-specific exploration. If a reconnaissance study has identified a region of interest and a land position has been established, the explorationist will use techniques that are designed to highlight specific mineral targets. Geologic studies, including preliminary mapping, are appropriate if outcrops are available. In addition, low-level airborne geophysical and ground-based geophysical investigations may help vector in on targets. *See* GEOLOGY; GEOPHYSICS.

Once an area that may contain mineralization has been located, explorationists focus on the deposit type and metals of interest. Sample protocols may be modified at this time, and reconnaissance geology is done, mapping and sampling rocks, with samples in this case consisting of perhaps a kilogram of rock chips collected randomly within 2–3 m (6–10 ft) of outcrops.

Where rock outcrops are limited, small excavations can be made using a rock hammer or a shovel to turn up rock chips from the underlying bedrock. If the region is largely soil-covered, a soil-sample survey may be appropriate. In such cases, the spacing of the grid should be less than the maximum size of the suspected target. It would be a shame to be in the vicinity of mineralization but miss the target by choosing the wrong sample spacing. At a grid interval of 0.25 km, 25 samples collected from 1 km^2 would adequately cover a region if large, porphyry-scale mineralization was the target, but it likely would miss most vein occurrences. Interestingly, many of the historic mining districts of the world were developed in vein deposits, and porphyry mineralization often spawns such veins, so they cannot be ignored as sample plans are being developed. *See* PORPHYRY.

The types of soil being sampled must be well defined to ensure reliable results. The characteristics of soils and the mobility of metals within soils may vary considerably according to the parent rock, the topographic relief, and the climate under which the soil developed. These simple variables determine the depth, type, and stability of soils, as well as the residence time, Eh (oxidation-reduction potential) and pH characteristics of ground water, and the types and activity of soil bacteria. All of these variables influence the extent to which metals may be mobilized from source materials and the location within the soil profile where elements of interest may be concentrated. *See* SOIL; SOIL CHEMISTRY.

Detailed geochemical studies, perhaps accompanied by ground-based geophysical studies, cover less ground than reconnaissance studies, and are more costly but provide very good data, especially when combined with geology. High-resolution soil sampling uses a much closer spacing, and a biogeochemical evaluation may be useful at this time. Plants flux large quantities of water and dissolved components

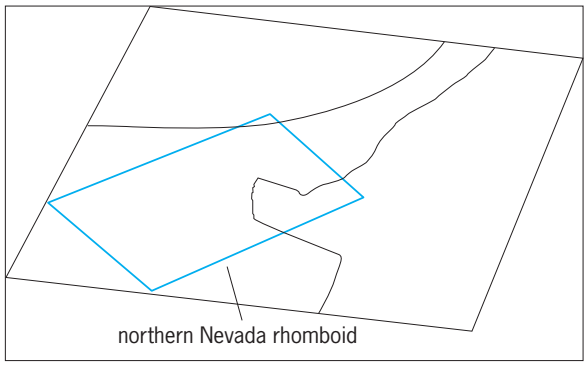

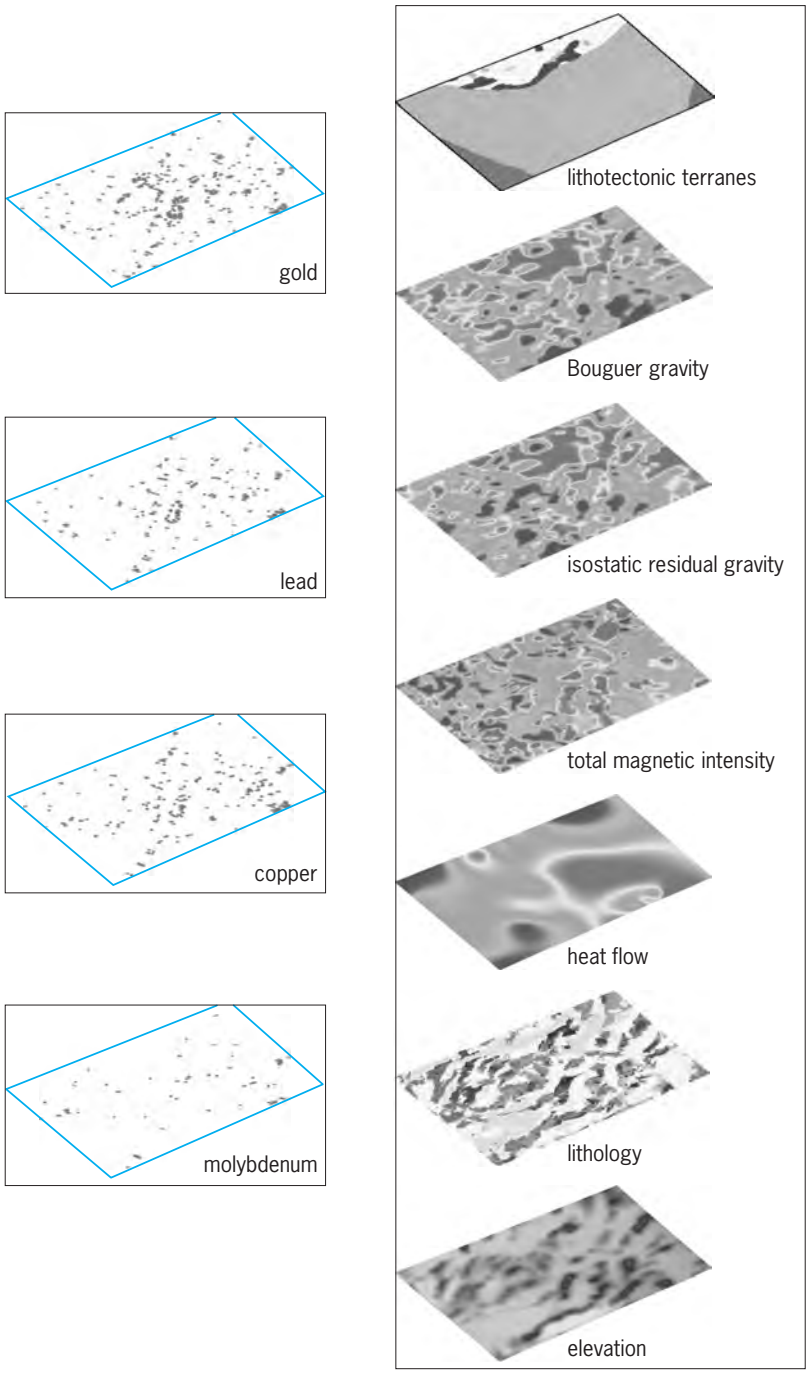

through their root systems and through new leaves, stems, and reproductive cells. Indigenous plants in desert regions have very well developed root systems that may reach tens of meters below the surface to sample mineralized rocks. Organic debris concentrates in soil and soil litter, as well as in bogs and lake sediments. These media probably should be sampled and interpretation must be done with care, as false anomalies can occur in such settings. Soils contain a large number and variety of bacteria, and these now are studied by well-prepared explorationists. For example, N. L. Parduhn and J. R. Waterson have described a location in Alaska where *Bacillus cereus* were easier to culture from soils above a vein system containing gold and quartz. *See* BIOGEOCHEMISTRY.

Metals and metal compounds with high vapor pressures may "bleed" upward through soil and overburden to reveal anomalous occurrences. Mercury, for example, can be detected with "sniffers" consisting of an inverted cup containing a platinum wire coated with activated carbon. Buddingtonite ($NH_4AlSi_3O_8$), a type of feldspar that occurs with gold and silver deposits, releases ammonia, which can be detected with a sniffer similar in design to the mercury device. *See* FELDSPAR.

Vectoring in the third dimension. Augering, or digging, pits to bedrock, perhaps trench sampling with a bulldozer or backhoe, is expensive, but it provides significant data on rock structure and on hydrothermal alteration and mineralization. The best way to test for mineralization at depth is by drilling. Short of mining, drilling is the only reliable way to determine what is present at depth. Even at this point, the explorationist's work is not done. Chemical data, together with the characteristics of the country rocks and of alteration and mineralization, will define whether the vectoring process has been successful.

Samples of rock cuttings from augering and rotary drilling, and of drill-core samples, normally are analyzed for multiple elements which, together with the rock properties, provide explorationists an evolving picture of the mineral system with depth. Any interval can be chosen, but homogenized samples of 1.5–3-m (5–10-ft) core segments is a good place to start. As many as 75 elements can be determined from samples, with analytical techniques that range from fire assay and atomic absorption spectrometry to inductively coupled plasma spectrometry and mass spectrometry (ICP-MS). Major, minor, and rare-earth element concentrations provide important clues to the nature of the rock systems and to the events that have affected them, but the "big 11" elements provide the most reliable guides to mineralization. From

Fig. 2. Gold, lead, copper, and molybdenum with respect to a GIS for the northern Nevada rhomboid. GIS provides topography, geology, and geophysical background for the metal occurrences. The rhomboid is shown in all frames. Boundaries in the top panel are between igneous rocks with oceanic (NW) and continental (E) affinities. The rhomboid lies largely within a region of mixed oceanic and continental affinities. (*From C. M. Tremper and D. E. Pride, in Window to the World: Geological Society of Nevada Symposium Proceedings, vol. 2, 2005*).

low to high temperature (near surface to deep), they are (1) gold and (2) silver, plus (3) mercury, (4) arsenic, and (5) antimony, (6) lead, and (7) zinc, often with (8) copper, and copper and (9) molybdenum, sometimes accompanied by (10) tin and (11) tungsten. Anomalous concentrations of one or more of these elements signal success, and provide the basis upon which decisions are made that involve time, money, and reputation. *See* ANALYTICAL CHEMISTRY; DRILLING, GEOTECHNICAL.

Data manipulations. A large amount of data routinely is available to explorationists, in addition to mineralization models and personal experience. Data provide both a challenge and an opportunity, with tools available to accept the former and maximize the latter. The isolation of authentic anomalies (the goal in exploration) is accomplished most efficiently when various sources of information are integrated. Mathematical and statistical packages, together with geographic information systems (GIS) technology, provide the tools to combine, overlay, integrate, and manipulate data to highlight the best targets. *See* GEOGRAPHIC INFORMATION SYSTEMS; MATHEMATICAL SOFTWARE.

For grid soil surveys, and for rock surveys in which there is a reasonable distribution of samples, information may be gleaned by contouring element concentrations. The locations of important samples and clusters of samples are easily identified in this way. Since hundreds of samples and dozens of elements may be involved, the efficacy of doing this by computer is obvious.

Moving-average calculations, and evaluations of regional trends in element concentrations, can identify broad anomalies with respect to background, and the "residuals," or samples that vary by more than can be accounted for mathematically, provide strong clues to the presence of anomalies. The data package itself can be "mined" statistically and mathematically by looking for multiple element correlations and ways elements may be related to each other or to other factors that are not obvious in the data. These methods are especially helpful as data packages become larger and contain larger numbers of elements. In such an environment, the ability to analyze and manipulate data becomes very important. Bias enters the picture only from the explorationist, whose experience and training make one set of data and geologic parameters somehow more interesting than a similar data package.

To make decisions, it often is necessary to overlay many types of information. Throughout history, this has been the time-honored approach to exploration, but today we are able to do this with geographic information systems (GIS) technology (**Fig. 2**). Using GIS, the land position, geology, topography, and geophysical parameters can be integrated with data layers of element concentrations, combinations of elements, perhaps ratios of element concentrations, to highlight areas with a high potential for discovery.

Douglas E. Pride

Bibliography. Agricola, Georgius, *De Re Metallica*, transl. by H. C. Hoover and L. H. Hoover, Dover, 1912; R. L. Bates and J. A. Jackson, *Glossary of Geology*, 3d ed., 1987; N. L. Parduhn and J. R. Waterson, *Preliminary Studies of Bacillus cereus Distribution near a Gold Vein and a Disseminated Gold Deposit*, U. S. Geol. Surv. OFR 84-509, 1984; A. W. Rose, H. E. Hawkes, and J. S. Webb, *Geochemistry in Mineral Exploration*, 2d ed., Academic Press, London, 1979; C. M. Tremper and D. E. Pride, Mineralization and the northern Nevada rhomboid, in H. N. Rhoden, R. C. Steininger, and P. G. Vikre (eds.), *Window to the World: Geological Society of Nevada Symposium Proceedings*, vol. 2, 2005.

Geochemistry

A field that encompasses the investigation of the chemical composition of the Earth, other planets, and indeed the solar system and universe as a whole, as well as the chemical processes that occur within them. The discipline is large and very important because basic knowledge about the chemical processes involved is critical for understanding subjects as diverse as the formation of economically valuable ore deposits, safe disposal of toxic wastes, and variations in the Earth's climate.

Geochemistry is rooted in geology, although many of its early contributors were chemists. Much of the initial work in geochemistry was descriptive, as practitioners attempted to determine the exact chemical composition of rocks, minerals, ocean and river water, the atmosphere, and meteorites. Current emphasis is on understanding geochemical processes. V. M. Goldschmidt is usually considered to be the father of modern geochemistry. His investigations focused on the basic "laws" underlying the observed distribution of the chemical elements in the Earth. He and his students conducted systematic element-by-element analytical studies of minerals and rocks in order to provide the basis for such understanding. "Goldschmidt's Rules" predict how the elements will behave, particularly in igneous rocks, based on their size and electrical charge. Like Goldschmidt's work, modern geochemistry is based on accurate analytical studies combined with investigations of the chemical processes that lead to the observed chemical state of the Earth, the solar system, and their component parts.

As in any broad interdisciplinary field, a number of specialties have become subdisciplines. However, their boundaries are by no means rigid, and substantial overlap occurs. A brief description of some of the more important subfields follows.

Isotope geochemistry. This subdiscipline is based on the fact that the isotopic compositions of various chemical elements may reveal information about the age, history, and origin of terrestrial and extraterrestrial materials. Isotopes of an element share the same chemical properties but have slightly different nuclear makeups and therefore different masses. During chemical reactions or processes such as diffusion or evaporation, mass differences between isotopes may lead to enrichment or depletion, and in

such cases the isotopic compositions of the materials serve to identify and quantify the processes. Furthermore, the degree of enrichment or depletion is sometimes temperature-dependent, permitting interpretation of the isotopic composition of a sample in terms of past temperatures (paleothermometry).

Some naturally occurring isotopes are radioactive and decay at known rates to form daughter isotopes of another element; for example, radioactive uranium isotopes decay to stable isotopes of lead. Radioactive decay is the basis of geochronology, or age determination: the age of a sample can be found by measuring its content of the daughter isotope.

Both radioactive decay and the processes that enrich or deplete materials in certain isotopes cause different parts of the Earth and solar system to have different, characteristic isotopic compositions for some elements. These differences serve as fingerprints for tracing the origins of, and characterizing the interactions between, various geochemical reservoirs. *See* DATING METHODS; ELEMENTS, GEOCHEMICAL DISTRIBUTION OF; ISOTOPE; LEAD ISOTOPES (GEOCHEMISTRY).

Cosmochemistry. Cosmochemistry deals with nonearthly materials. Typically, cosmochemists use the same kinds of analytical and theoretical approaches as other geochemists but apply them to problems involving the origin and history of meteorites, the formation of the solar system, the chemical processes on other planets, and the ultimate origin of the elements themselves in stars. Cosmochemistry is important because the geochemical makeup and history of the Earth can only be properly understood within the framework of a general knowledge of the chemical composition and evolution of the solar system. Cosmochemistry has some unique aspects. For example, some of the radioactive isotopes created in the interiors of stars and spewed out into space during supernovae explosions have quite short half-lives and have now completely decayed away. Only in meteorites that have remained undisturbed since the time of formation of the solar system does evidence for these isotopes remain. *See* COSMOCHEMISTRY; METEORITE; SOLAR SYSTEM.

Organic geochemistry. Organic geochemistry deals with carbon-containing compounds, largely those produced by living organisms. These are widely dispersed in the outer part of the Earth—in the oceans, the atmosphere, soil, and sedimentary rocks. Organic geochemistry is important for understanding many of the chemical cycles that occur on Earth because biology often plays a major role. An example is the carbon cycle, during which carbon dioxide (CO_2) is consumed by plants, and carbon is stored by burial of plant and animal matter, transformed into coal and petroleum, then released again to the atmosphere as CO_2 when the fossil fuel is burned. Organic geochemists are also active in investigating such areas as the origin of life, the formation of some types of ore deposits that may be biologically mediated, and the origin of coal, petroleum, and natural gas. *See* BIOGEOCHEMISTRY; COAL; NATURAL GAS; ORGANIC GEOCHEMISTRY; PETROLEUM; PREBIOTIC ORGANIC SYNTHESIS.

Other areas. Low-temperature geochemistry is a broad field distinguished by its concern with processes that occur at approximately the ambient temperatures of the Earth's surface. For the most part, this branch of geochemistry deals with chemical reactions and processes involving the Earth's natural waters. Cycles of elements in the oceans, weathering processes on the continents, and sedimentation in lakes and seas are typical areas of interest. Marine chemistry and environmental geochemistry are aspects of low-temperature geochemistry that deal with the chemistry of the oceans and (for the most part) with the geochemical impact of humans on the Earth, respectively. Atmospheric geochemistry is also much concerned with the impact of humans, in this case on the atmosphere. *See* ATMOSPHERIC CHEMISTRY; HYDROSPHERE; MARINE SEDIMENTS; SEAWATER; WEATHERING PROCESSES.

In recent years there has been widespread application of geochemical techniques to problems in paleoclimatology and paleoceanography. In this approach, ocean sediments, sedimentary rocks on land, ice cores, and other continuous records of the Earth's history are analyzed for fossil chemical evidence of past climates or seawater composition. As in most areas of geochemistry, precise and accurate analytical methods for determining the isotopic and elemental composition of the samples are critical. *See* EARTH SCIENCES; PALEOCEANOGRAPHY; PALEOCLIMATOLOGY.

J. D. Macdougall

Bibliography. A. H. Brownlow, *Geochemistry*, 2d ed., 1995; G. Faure, *Principles and Applications of Inorganic Geochemistry*, 2d ed., 1997; K. B. Krauskopf, *Introduction to Geochemistry*, 3d ed., 1994.

Geode

A roughly spheroidal hollow body, lined on the inside with inward-projecting small crystals (see **illus.**). Geodes are found most frequently in limestone beds but may occur in some shales. Typically, a geode consists of a thin outer shell of dense chalcedonic silica and an inner shell of quartz crystals, sometimes beautifully terminated, pointing toward the hollow

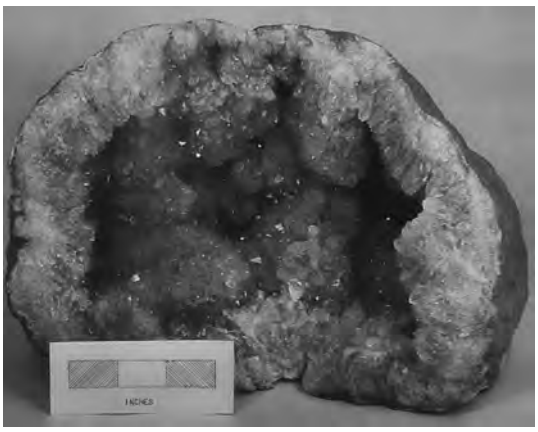

Geode, lined with quartz crystals, Keokuk, Iowa. 1 in. = 2.5 cm. (*Brooks Museum, University of Virginia*)

interior. Many geodes are filled with water; others, having been exposed for some time at the surface, are dry. Calcite or dolomite crystals line the interior of some geodes, and a host of other minerals are less commonly found. In some geodes there is an alternation of layers of silica and calcite, but almost all geodes show some banding suggestive of rhythmic precipitation. *See* CHALCEDONY.

The origin of geodes lies in the presence of an original cavity, in many cases a void within a fossil shell, from which the geode originally grew. The geode grows by expansion, the layer of chalcedonic silica being the hardened equivalent of an original silica gel. The expansion is due to osmotic pressure from original seawater trapped inside the silica gel shell and fresh water on the outside of the gel. The projecting quartz crystals are precipitated from later ground waters infiltrating the already hardened, hollow spheroid. *See* SEDIMENTARY ROCKS.

Raymond Siever

Geodesic dome

A curved lattice grid dome that utilizes the equilateral triangle as the basis of its surface grid geometry. Buckminster Fuller, the inventor and champion of the geodesic dome, obtained a patent in 1954 that described a method of dividing a spherical surface into equilateral triangles. The realization that a triangular modular grid could be imposed on a spherical surface had great impact on the design philosophies of architects in Fuller's time.

Fuller found that if a regular (equal-sided and -angled) pentagon was used as the base of a pentagonal pyramid and if the inclined sides of the pyramid were the same length as the pentagon's sides, the apex of the pyramid would be at the center of the pentagon and on the surface of a sphere, passing through all the vertices of the pentagon.

Based on this geometry, two geometric regular polyhedra (all of whose edges are the same length and all of whose faces are regular, identical polygons) are used in generation of the geodesic dome's grids. The two regular polyhedra that can be inscribed in a sphere are the dodecahedron (12 faces, each of which is a regular polygon; **illus.** *a*) and the more utilized icosahedron (20 faces, each of which is an equilateral triangle; illus. *b*). *See* POLYHEDRON.

The geodesic dome has been used for everything from great exhibition spaces and halls to outdoor tent supports and jungle gyms. Use of the equilateral triangle as the basic modular grid has led to use of the geodesic dome where prefabrication, speed of erection, and dismantling are major considerations. Geodesic domes have been built up to 143 m (469 ft) in diameter. The earliest example is the 30-m diameter (98-ft) dome designed by Fuller and built in 1952. Since 1955, the U.S. Army has used geodesic domes to enclose radar installations (radomes) throughout Canada and Alaska, because the design lends itself to prefabrication of units that can be easily erected.

Prefabrication of the geodesic dome received impetus from the development of aluminum struts to interconnect the structure, and diamond-shaped aluminum sheets for the covering. The first aluminum dome was built in 1957 in Hawaii; it had a floor diameter of 44 m (145 ft) and a height of 15 m (49 ft). The geodesic dome that received the most publicity was erected in 2 weeks in Moscow in 1959 for the U.S. Technical Exhibition; it had a diameter of 61 m (200 ft).

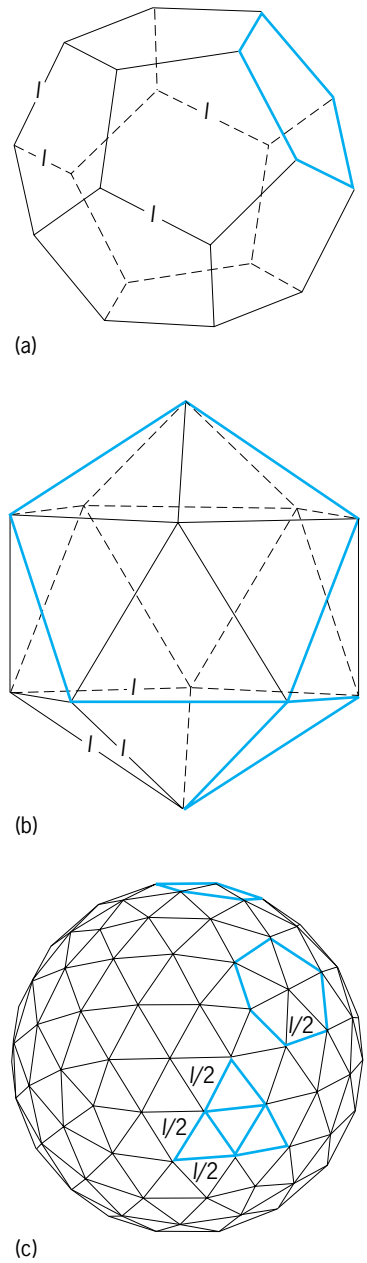

Geometry of geodesic domes. (*a*) Dodecahedron: a regular pentagon is typical of each face; every point is an apex because all apexes are on the sphere; each strut (*l*) is the same length. (*b*) Icosahedron: an apex is above the center of each polygon and on the surface of the sphere; the equilateral triangle typical of each face is highlighted; each strut (*l*) is the same length. (*c*) Larger dome based on the icosahedron: subdivision is formed by connecting midpoints of struts (*l*) of equilateral triangles (each half strut is labeled *l*/2); the original pentagon is shown at the top, and a formed hexagon is also shown.

By utilizing the icosahedron as the basic building block of the geodesic dome, larger domes are possible with additional triangular subdivisions. This subdivision is known as the frequency. The first frequency is to interconnect the projected midpoints of the struts of each equilateral triangle of the icosahedron as they will project on the spherical surface. The result is four almost equilateral triangles where there was one before. The resulting lattice has similar but not exactly equilateral triangles if the grid is to remain on the spherical surface. This subdivision process can continue. The resulting grids have both triangular and hexagonal grids as a by-product within the basic geodesic dome geometry, with pentagons around the apex of the basic underlying icosahedron framework (illus. *c*). I. Paul Lew

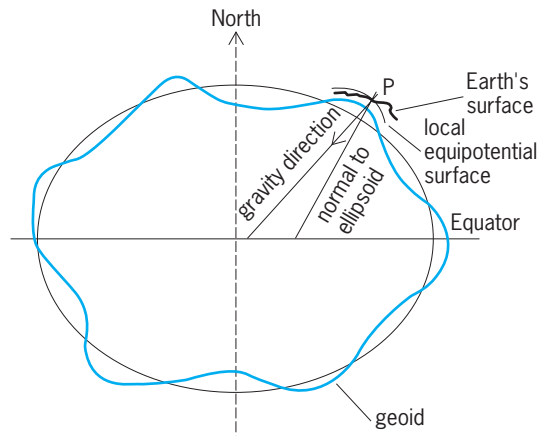

Fig. 1. Determination of geodetic and astronomic latitudes. P represents the location at which the latitude and longitude are to be defined.

Geodesy

The science of measuring the size, shape, and gravity field of the Earth. Geodesy supplies positioning information about locations on the Earth, and this information is used in a variety of applications, including civil engineering, boundary demarcations, navigation, resource management and exploration, and geophysical studies of the dynamics of the Earth. *See* EARTH.

The conventional measurement systems in geodesy are triangulation and trilateration for determining horizontal positions, and leveling for determining heights. These techniques depend on the Earth's gravity field, and so a major part of geodesy has been not only position determination but also the measurement of the Earth's gravity field.

Two major measurement systems were developed in the late 1970s and early 1980s: satellite laser ranging (SLR) systems, which could measure the distance from the ground to a satellite equipped with special corner-cube mirrors; and very long baseline interferometry (VLBI), which could measure the difference in arrival times between radio signals from extragalactic radio sources. With these systems it is possible to measure accurately (within a few millimeters) the distances between points located on different continents, making possible the creation of truly global coordinate systems. Both systems were deployed around the world to measure not only the positions of locations but also the changes in those positions; and thus it was confirmed that the Earth is not a static but a highly dynamic body, with much of this dynamism causing catastrophic events such as earthquakes and volcanic eruptions. The more recent Global Positioning System (GPS) offers much of the capability of SLR and VLBI. *See* RADIO ASTRONOMY.

Earth models. For much of the history of geodesy, measurements were made locally, and this is reflected in the ways that geodetic data are interpreted. When local measurements are made, the only direction that is known is the direction of the local gravity vector. A bubble in a liquid in a curved tube will be perpendicular to this direction, and a plumb line lies along this direction. When an astronomical latitude and longitude are determined, it is the direction of the gravity vector that is determined. When height differences between points are measured by using spirit leveling, it is the heights above an equipotential surface (that is, a surface that has a constant gravitational potential or that is always orthogonal to the local direction of gravity) that is determined.

The direction of gravity is affected by the mass distribution within the Earth; therefore coordinate systems based on gravity are difficult to use for precise determinations of the positions, because the coordinates do not vary in a geometrically predictable fashion. For horizontal positions such as latitude and longitude, it is possible to convert triangulation and trilateration measurements (the measurement of angles or a series of distances between points on the surface of the Earth) to geometric coordinates, that is, ellipsoidal coordinates. However, there is a difference between ellipsoidal and astronomical latitude (**Fig. 1**). For heights, conversion to geometric coordinates requires knowledge of the difference in height between an ellipsoidal shape and an equipotential surface known as a geoid. A geoid is the equipotential surface coinciding approximately with mean sea level, and deviates by up to 330 ft (100 m) from an ellipsoid shape. If mean sea level is assumed to form an equipotential surface, this definition can be used to globally define a height datum along all coastlines around the world. Spirit leveling can then be used to determine the heights at points away from the coast. Geodetic systems such as SLR and VLBI are able to determine the positions of points geometrically, and thus their natural coordinate system is an ellipsoidal one for both height and latitude and longitude. *See* LATITUDE AND LONGITUDE; SURVEYING; TOPOGRAPHIC SURVEYING AND MAPPING.

Techniques. Conventional geodesy uses the techniques of triangulation and trilateration to determine latitude and longitude. In triangulation, the angles in triangles are measured and converted to position

by knowing the length of at least one side of the triangle. By linking the sides of the triangles, it is possible to determine all the lengths of all the sides of a linked set of triangles by measuring directly just a few distances in the set. Trilateration uses the direct measurements of the sides of the triangles, and became possible only in the 1950s after the development of radar. Distances are measured effectively by measuring the time it takes for a radio or light signal to travel along the side of a triangle. Individual triangles measured with these systems can be only as large as the eye can see. Linking networks of triangles together makes measurements across a whole country possible.

Height measurements are made using leveling in which, by measuring the height difference between nearby points, the height difference between distant points is obtained by summing the height differences of all linked pairs of points between them. The height differences are measured by setting a telescope perpendicular to the direction of gravity and reading the height difference from rulers (called staffs), each 7–10 ft (2–3 m) long, set vertically above the two points being measured. The telescope and two points need to be close together with separations of usually less than 164 ft (50 m). Long lines of these points form a leveling network that can run for thousands of kilometers and can take many years to measure. The absolute heights of the points in the network are found by having some of the points near the ocean, and the zero of the height scale is set to coincide with mean sea level measured on tide gauges over many years.

Modern geodetic measurements, referred to as a spaced-based system, use time-delay measurements that are far less affected by atmospheric refraction than angle measurement systems. Both SLR and VLBI allow measurement of positions of points separated by thousands of kilometers with accuracies of few millimeters. However, both systems are complex and expensive to operate. Satellite laser ranging uses pulsed laser beams, and the round-trip travel time to a satellite equipped with corner-cube reflectors can be measured to a few millimeters. Very long baseline interferometry records signals from extragalatic radio sources on high-density tape. When these recordings from different stations are cross-correlated, the difference in arrival times of the signals from the extragalatic radio source can be measured to a few millimeters of equivalent light-travel time. *See* LASER.

The use of SLR and VLBI has been restricted to about 200 locations around the world, and they have been accessible only to large government-run measurement programs. However, based on the experience with SLR and VLBI, a geodetic measurement system, the Global Positioning System, was developed, originally to allow world positioning to a few meters' accuracy in real-time, with lightweight portable equipment. The system consists of 24 GPS satellites in orbits that are about 12,000 mi (20,000 km) above the surface of the Earth. The satellites transmit encoded signals at two frequencies (1575.42 and 1227.60 MHz) in the L-band of the radio spectrum. The coding on the signals includes the time when a signal was transmitted and information about the location of the satellites. Each satellite transmits by using a different code sequence, called a pseudo random number sequence, which allows signals from different satellites to be separated in a GPS receiver, which receives signals from all visible satellites (usually about eight). A receiver, by measuring the time on its own clock when signals from different satellites are received and having determined where the GPS satellites are located, is able to determine its position in a global reference frame. The best GPS receivers can measure the time of arrival of the signals to about 1 nanosecond, which is equivalent to 12 in. (30 cm) of light travel time, thus allowing them to determine their positions to about 3 ft (1 m). The receiver is able to make this measurement in about one-fiftieth of a second, and therefore the receiver can continuously monitor its position at this level of accuracy. The GPS revolutionized navigation around the world, allowing aircraft, ships, and other moving vehicles to continuously monitor their positions. *See* ELECTROMAGNETIC RADIATION; SATELLITE NAVIGATION SYSTEMS.

Prior to January 2000, there were restrictions on access to the full accuracy of GPS signals: the transmission times of signals from the satellites were corrupted in a pseudorandom fashion, such that standalone nonauthorized receivers could be positioned only to 330 ft (100 m). This degradation of the positioning accuracy was called selective availability (SA) and could be circumvented by making measurements continuously at a known location and transmitting to all nearby GPS receivers the errors in the times of the GPS transmissions. This differential GPS (DGPS) system is still widely used for precise navigation in harbors and near airports where accuracy of better than 33 ft (10 m) is required. The use of DGPS greatly reduces positioning errors due to errors in modeling the locations of the GPS satellites and signal delays induced by the refraction of the Earth's atmosphere. The other known purposeful corruption of GPS signals, antispoofing (AS), is the addition of an extra code on the signal that denies access to the precise positioning codes that are written on the GPS signals at both frequencies. For very accurate applications, measurements at both GPS frequencies are needed to eliminate the error caused by the propagation of the signals through the Earth's ionosphere. With specialized receivers used in the highest-accuracy applications of GPS, it is possible to make measurements using both GPS frequencies, even in the presence of AS. The generation of GPS satellites launched starting in 2005 includes known codes on both GPS frequencies, making dual frequency measurement much easier.

The most accurate applications of GPS do not use directly the time information on the GPS signals, but the measurement of the phase of the radio signal carrier on which the GPS information is transmitted. The use of this phase information allows accurate measurements in the millimeter range to be made

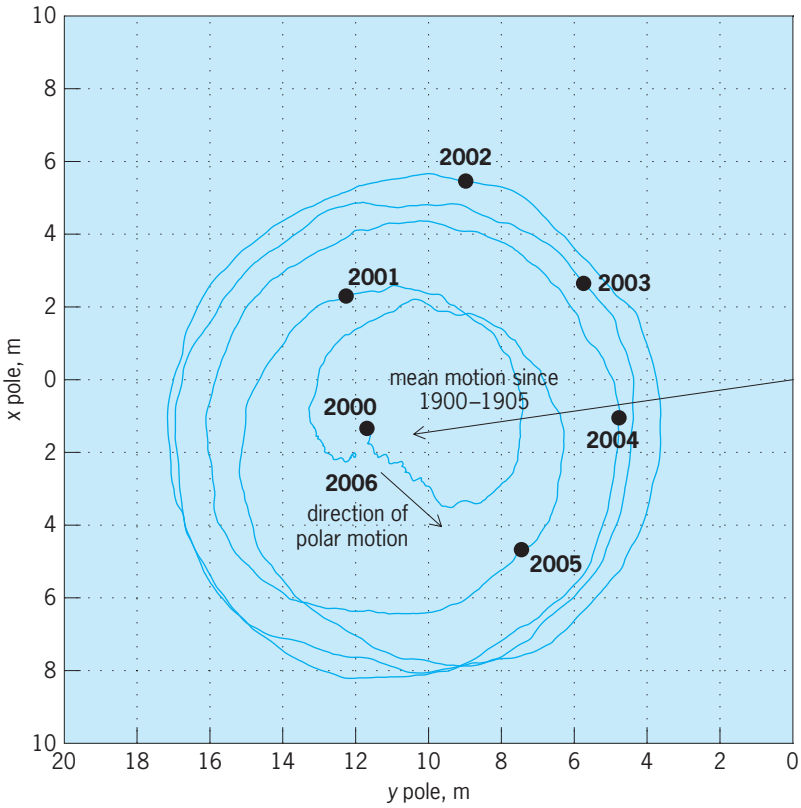

Fig. 2. Motion of the position of the Earth's rotation axis as determined by geodetic methods. The line shows daily measurements of the position of the pole with dots labeled for years 2000–2006. The *x* axis points along the Greenwich meridian, and the *y* axis points in the direction of longitude 90°W.

over thousands of kilometers. Such measurements have been used to study the dynamics of the Earth. For example, 4-in.-accuracy (10-cm) position measurements allow railroads to determine definitely on which track a train is located; vehicle collision avoidance systems are possible with systems possessing this accuracy; and such data can be used in large-scale construction projects.

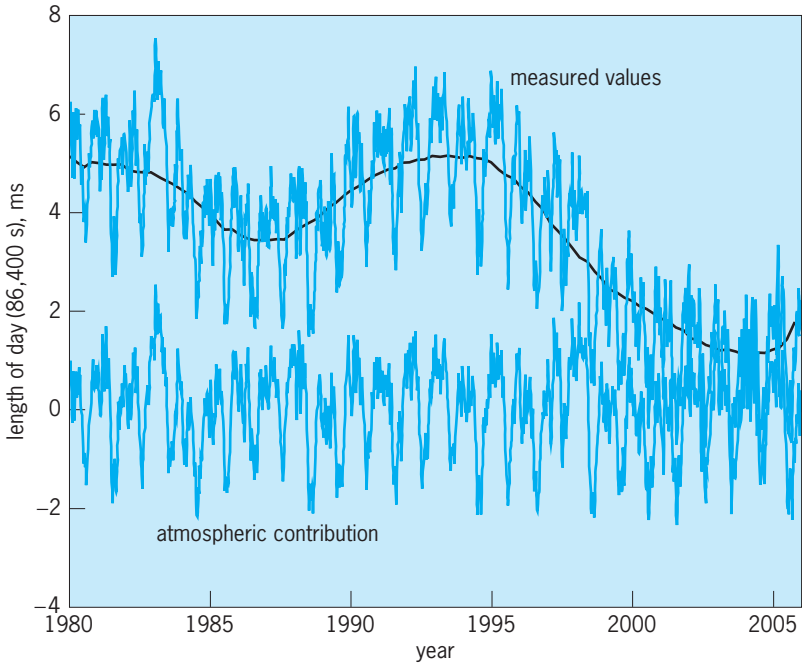

Fig. 3. Changes in the length of day for a 26-year interval, 1980–2006, measured with geodetic methods and inferred from variations in atmospheric angular momentum. The decadal time-scale differences are believed to arise from fluid motions in the Earth's core.

The most recent development in geodetic techniques is interferometric synthetic aperture radar (InSAR). This technique is used to measure heights of the topography or, if the topography is already known, the changes in the topography between two synthetic aperture radar (SAR) images. Heights measured with InSAR are far less accurate than normal geodetic height measurements, but since InSAR is an imaging system, large areas can be measured easily. If the InSAR instrument is on an orbiting spacecraft, global topography can be measured. The measurement of changes in topography with InSAR has been widely used to measure the surface displacements after earthquakes (by comparing before and after SAR images) and for monitoring volcanic deformations. *See* RADAR.

Geophysical applications. Some of the major impacts of modern geodetic measurements have been in the study of the dynamics of the Earth. The measurement systems enable the observation of many of the minute motions of the Earth, such as those associated with plate tectonics and other geophysical processes, and changes in the rotation of the Earth.

Earth deformations. Early evidence for plate tectonics was based on the geologic record, which showed that whole continents moved by tens to hundreds of kilometers per million years, but modern geodesy has allowed these motions to be observed on time scales of 10 years. The tectonic motions of thousands of locations around the world have been measured with VLBI, SLR, and GPS. The measurements clearly reveal the pattern of plate tectonic motions. These results show that plate tectonics is a continuously operating process when the sites are well away from regions of active volcanism and earthquakes. Even more remarkable is the agreement between the expected changes of this distance based on a million years of geology and the measurements made over a 10-year interval. In general, the motions of locations away from regions of active deformations agree within a few millimeters per year with geologic estimates. However, this is not always the case, and it is never the case in regions of active deformation. The study of regions where geology and geodesy disagree is one of the prime applications of modern geodesy. The western United States is one region that has been extensively measured with all forms of geodetic systems. A large motion as the west coast is approached is expected because of the motion of the Pacific tectonic plate relative to North America. However, the precise partitioning of a motion across California can be determined only with modern geodetic systems. The broad width of the motions, going all the way into the Mojave Desert, had not been expected before these measurements were made. *See* CONTINENTAL DRIFT; PLATE TECTONICS.

Earth motions. Measurements of changes in the rotation of the Earth have long played an important role in understanding the dynamics of the Earth.

Changes in the position of the Earth's rotation axis are of two types: changes in direction in inertial space (precession and nutation) and changes with respect to the crust of the Earth (polar motion). Precession is a secular motion of the rotation axis that causes the rotation axis to trace out a cone in space every 26,000 years. Nutations are the smaller-magnitude nodding of rotation axis about the precessional path of the rotation axis. Both precession and nutation are caused by the gravitational torque applied to the equatorial bulge of the Earth, and their values are a direct measure of the dynamic flattening of the Earth. The amplitudes of the nutations are also greatly influenced by the presence of the fluid core and the elastic properties of the Earth. Measurements of the nutations made with VLBI have yielded the most accurate estimate available of the flattening of the fluid core because of its influence on the nutations. Precession and nutations are determined from the apparent changes in the positions of stars and extragalactic radio sources. The major impact of VLBI has been that the extragalactic radio sources observed have very small motions, unlike those of stars. *See* EARTH ROTATION AND ORBITAL MOTION; NUTATION (ASTRONOMY AND MECHANICS); PRECESSION; TORQUE.

Polar motion is measured by apparent changes in the latitudes and longitudes of positions on the Earth's surface when these are determined by astronomical observations. The Earth has a natural period for the oscillations of the rotation axis about a mean position as determined by the direction of the maximum moment of inertia of the Earth; this oscillation mode is known as the Chandler wobble. It takes about 433 days for the axis to complete one cycle in this wobble mode. An example of the motion of the rotation axis can be shown as the position near the North Pole where the rotation axis pierces the Earth (**Fig. 2**). In one determination of this type, the conventional definition of the position of the North Pole was set to the mean position of the Earth rotation axis between 1900 and 1905; it was found that the average position of the rotation axis has changed by about 33 ft (10 m). Most of this motion is thought to arise from the relaxation of the Earth after the deformations caused by the Laurentide glaciation 10,000 years ago. Superimposed on the Chandler wobble is an annual motion of the rotation axis due to atmospheric mass redistribution between the Northern and Southern hemispheres between seasons. It is not known what process keeps the Chandler wobble excited to is current amplitude, but it is clear that changes in atmospheric wind and pressure and in ocean currents that seem to cause most of the smaller fluctuations are the primary mechanism. *See* ATMOSPHERIC GENERAL CIRCULATION.

The rotation rate of the Earth also varies on time scales of days to millions of years. A variety of processes contribute to changes in the rotation rate of the Earth. Some of the major ones are a slow decrease in the rotation rate due to energy loss to the lunar orbit through dissipation of tide energy and exchanges of angular momentum between the atmosphere and the fluid core and the solid Earth. For example, the dominant role of the atmosphere on short-period (less than several years) changes in the rotation rate of the Earth can be demonstrated by measuring changes in the length of day over a 2-year interval with modern geodetic systems, and predicting the changes based on calculation of the angular momentum of the atmosphere from wind speeds in the atmosphere and pressure measurements (**Fig. 3**). The variations are approximately 1 millisecond. The 2-ms mean value of the length of day is due to the slowing of the rotation since the average length of day was defined at the turn of the century. *See* EARTH, GRAVITY FIELD OF.

Thomas Herring

Bibliography. G. Bomford, *Geodesy*, 4th ed., 1980; B. Hofmann-Wellenhof and H. Moritz, *Physical Geodesy*, 2005; K. Lambeck, *Geophysical Geodesy: The Slow Deformations of the Earth*, 1988; A. Leick, *GPS Satellite Surveying*, 3d ed., 2003.

Geodynamics

The branch of geophysics that studies the processes leading to deformation of planetary mantle and crust and the related earthquakes and volcanism that shape the structure of the Earth and other planets. On the largest scale, these processes are a consequence of the transfer of heat out of planetary interiors due to cooling at their surfaces. Rock contracts as it cools, so that its density increases. The cool surface layer is heavier than the interior and has a tendency to sink into it. At the same time, cooling and solidification of the metallic core heats the deepest portion of the surrounding rocky mantle, causing it to become buoyant. The resulting flow of the mantle causes deformation at the surface. Volcanism arises from the partial melting of hot mantle that rises toward the surface from the deeper interior, in response either to buoyancy or to surface deformation. Surface deformation also results from external loads, such as the distribution of ice and water, tidal loads due to the gravitational attraction of nearby planetary bodies, and meteor impacts. *See* EARTH, CONVECTION IN; EARTH, HEAT FLOW IN; GEOPHYSICS; VOLCANO.

A planet's response to its internal heat flow depends largely on the rheology of deforming rock. At low temperatures, near the surface, rock behaves as a brittle-elastic material, allowing the propagation of seismic waves and the support of surface loads by elastic stresses. Deformation occurs by the formation of cracks or faults. On geologic time scales and at the higher temperatures of the deeper interior, thermally activated creep allows the solid, rocky mantle to flow like a viscous fluid. But even at these high temperatures, rock behaves elastically on short time scales so that elastic shear waves propagate through the slowly flowing mantle. In the case of the Earth in its current stage of evolution, plate tectonics describes how the surface behaves: large, cold, relatively rigid plates move laterally across the surface while the deeper mantle flows by creep. In the cold plates, deformation is largely confined to boundary faults between the plates. Faults slip with a stick-slip

behavior, giving rise to large earthquakes that occur primarily on the plate boundaries. *See* EARTHQUAKE; PLATE TECTONICS; RHEOLOGY; ROCK MECHANICS.

Geodynamics deals with phenomena that occur on a wide range of time and length scales. In an earthquake, stored elastic strain energy is released in times on the order of seconds. In contrast, significant flow of the deeper mantle occurs on time scales of tens to hundreds of million years. Tidal and glacial loads occur on intermediate time scales ranging from hours to thousands of years. The time- and temperature-dependent rheology of rock is controlled by processes that occur on the millimeter scale of individual mineral grains and the boundaries between them. It is remarkable that processes occurring on this small scale control the behavior of the Earth and planets on scales of thousands of kilometers, comparable to the dimensions of the tectonic plates.

Given the difficulty of direct observation and the wide range of scales involved in phenomena of interest, multiple approaches are needed to understand geodynamic processes. Laboratory experiments on relatively small samples of rock are used to characterize the rock's physical properties, such as its rheology, at high pressures and temperatures. The rate of deformation due to creep in nature is much too slow to measure directly in the laboratory. Thus a theoretical understanding of deformation mechanisms is needed to extrapolate from the much higher deformation rates of laboratory measurements to naturally occurring conditions.

Field studies of rocks once deep in the interior and brought to the surface by uplift and erosion provide evidence of the processes that have affected them. But the interior of the Earth where the processes of interest are actually occurring is not directly accessible for study. Thus geodynamicists must design large-scale observational experiments that allow them to create conceptual and physical images of the interior, using combinations of seismic, gravitational, electromagnetic, and heat-flow measurements. Variations in global gravity and corresponding surface topography can be remotely sensed from orbiting spacecraft. This is particularly important in the case of planets other than the Earth where other types of geophysical measurements have not yet been made. These data seldom have a unique interpretation, but numerical experiments with theoretical/computational models that describe relevant processes can determine which of these are consistent with available observations. Such analyses, in turn, suggest new observations that might further refine understanding. *See* EARTH CRUST; EARTH INTERIOR; GEODESY; GEOMAGNETISM; SEISMOLOGY.

One long-standing controversy has been whether plates that sink into the mantle at convergent plate boundaries descend into the deep interior or remain in the upper part of the mantle. This question of how deep the mantle circulation implied by plate tectonics extends has important implications for the thermal and chemical evolution of the Earth. A recent accomplishment of seismic tomography has been to image mantle elastic-wave speed variations associated with sinking plates. Tomographic images clearly suggest that plates do, at least sometimes, descend into the deep mantle (see **illus.**).

The questions and problems that geodynamicists seek to answer encompass important concerns for humanity. Understanding where essential energy and material resources occur in the Earth, how they form, and the physical processes involved in their recovery is the basis of the current standard of living. For example, the flow of fluids through cracks in rocks is important in hydrocarbon extraction, nuclear waste isolation, and geothermal energy. Understanding the mechanisms that produce magma at depth, how magma is extracted and transported to the surface, and its episodic and sometimes spectacular eruptions is key to mitigating the hazards that volcanoes pose to human life and property. Geodynamicists are also concerned with environments, such as deep-sea hydrothermal systems, in which primitive forms of life originate and evolve on the Earth, and possibly on other planetary bodies. *See* DEEP-SEA FAUNA; MID-OCEANIC RIDGE; PETROLEUM GEOLOGY; VOLCANOLOGY.

E. M. Parmentier

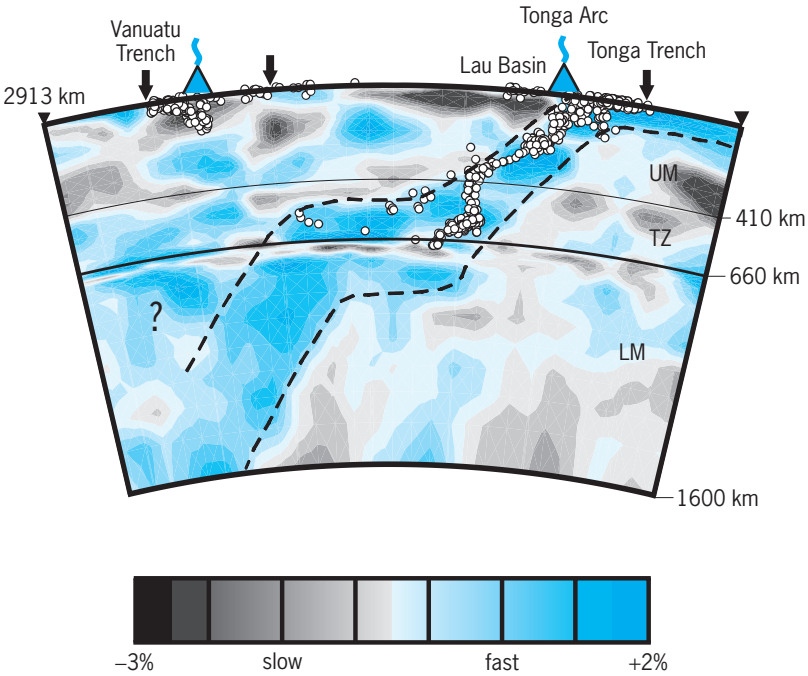

Seismic (P-wave) velocity in the mantle beneath the northern Tonga convergent plate boundary. The Pacific plate approaches the Tonga trench from the right and sinks into the mantle beneath the volcanic Tonga arc. Circles show hypocenters of earthquakes that occur in the sinking plate. The tint bar is the key to seismic velocity anomaly. Seismic velocities in the plate are higher than in the surrounding mantle. Although earthquakes in the sinking plate extend only to 660 km depth, the seismic velocity anomaly continues to greater depth. UM, upper mantle; TZ, transition zone; LM, lower mantle. (*Adapted from R. van der Hilst, Complex morphology of subducted lithosphere in the mantle beneath the Tonga trench, Nature, 374:154–157, Macmillan Magazines Ltd., 1995*)

Bibliography. C. M. R. Fowler, *The Solid Earth: An Introduction to Global Geophysics*, Cambridge University Press, 1990; R. van der Hilst, Complex morphology of subducted lithosphere in the mantle beneath the Tonga trench, *Nature*, 374:154–157, 1995.

Geodynamo

The mechanism thought to be responsible for the generation of the Earth's magnetic field through the convection of conducting fluids in the Earth's core.

Paleomagnetic measurements suggest that the Earth has possessed a magnetic field for at least 3.5 billion years. Geophysicists generally accept that the ambient magnetic field measured at the Earth's surface is due to electric currents flowing in its liquid iron core (**Fig. 1**). In the absence of electromotive forces, like those of chemical batteries, electric currents will decay as magnetic energy is converted to heat. Without some regenerative process to offset such natural ohmic dissipation in the Earth's core, any electric currents and the associated magnetic field would vanish in about 15,000 years. Regeneration of the field is necessary. In the Earth it is thought that the magnetic field is maintained by dynamo action, whereby the kinetic energy of convective motion in the Earth's liquid core is converted into magnetic energy. Since this process operates without an external energy source, the geodynamo is said to be self-sustaining. *See* GEOELECTRICITY; GEOMAGNETISM; PALEOMAGNETISM.

Many issues in geodynamo theory remain unresolved, since the nonlinear equations of magnetohydrodynamics are difficult to solve. These equations, describing the processes operating in the core, are akin to the equations of oceanography and meteorology, but with the additional complication presented by the magnetic field through the Lorentz force (magnetomotive force). In the laboratory, mechanical dynamos—moving machines consisting of particular arrangements of metal—have been made since the days of Michael Faraday. Still, it is not obvious how a simply connected conducting fluid body, like the Earth's core, functions as a dynamo without the induced currents simply short-circuiting and eliminating field generation. In fact, the electric current in a dynamo and the magnetic field that it sustains cannot be too simple; a theorem, due to T. G. Cowling, says that no axisymmetric, or even two-dimensional, dynamo magnetic field can exist. Although the magnetic north and south poles usually are nearly coincident with the geographic poles, indicating that the rotation arising from the Coriolis force plays an important role in the core's dynamics, it is no accident that the compass does not point toward true north everywhere on the Earth's surface—an inherent lack of symmetry. As a result, theoretical progress has been slow since scientists often take advantage of symmetry, should it be present, when solving mathematical equations. *See* CORIOLIS ACCELERATION; MAGNETOHYDRODYNAMICS.

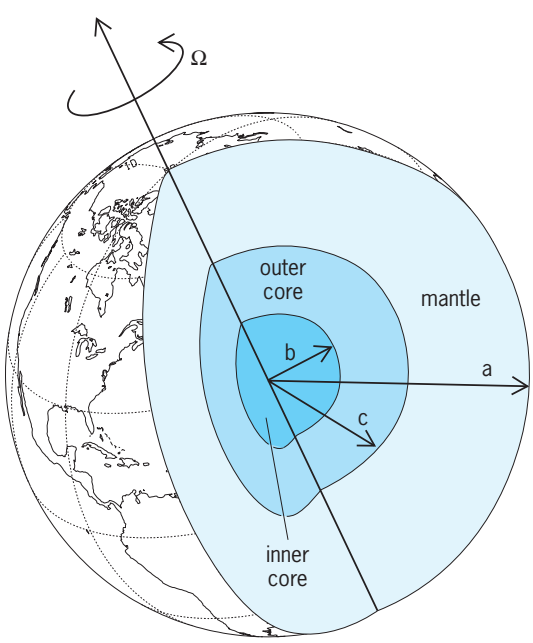

Fig. 1. Anatomy of the Earth. The rocky mantle has a radius $a = 6371$ km (3959 mi), the liquid iron outer core has a radius $c = 3485$ km (2165 mi), and the solid inner core has a radius $b = 1215$ km (755 mi). The Earth's rotational vector is Ω.

Geophysicists do, however, have a good qualitative understanding of how the geodynamo works. In the 1940s and 1950s, W. M. Elsasser and E. N. Parker first elucidated the so-called α-ω (alpha-omega) mechanism, by which core fluid motion can act as a dynamo if it consists of a combination of differential rotation and convective helical motion (**Fig. 2**). Since then it has been shown mathematically that dynamo regeneration can arise from the turbulent motion of a rotating fluid. Although the α-ω mechanism probably describes how the field is amplified, it is the dynamics that ultimately governs the strength of the field, which would grow until a rough balance between the Coriolis force and the Lorentz forces is attained.

With respect to the energy sustaining the fluid convection in the outer core, there are two possible sources of buoyancy—thermal and compositional. Thermal convection is perhaps most familiar, with heat sources, such as radioactive potassium, distributed over the volume of the outer core. With sufficient internal heating, the fluid is gravitationally unstable and, as a result, convection is sustained. Compositional convection is currently favored by most geophysicists as the energy source of the geodynamo. Although the core is primarily of iron, there are probably light impurities, such as sulfur. Due to the effects of pressure, as the Earth slowly cools, iron solidifies at the inner-core boundary. This causes the inner core to grow and leaves the lighter constituents behind in the fluid at the base of the outer core, supplying the buoyancy that drives the convection. *See* CONVECTION (HEAT).

Although dynamo action is an inherently asymmetric process, researchers are currently trying to resolve whether or not the the time-averaged morphology of the geomagnetic field is asymmetric, showing either persistent longitudinal variation or asymmetry under reflection through the equatorial plane. It is an issue of importance because persistent asymmetry in the geometric form of the field would indicate that the core and mantle are dynamically coupled. Such a possibility was first proposed by R. Hide, who was inspired by the Taylor-Proudman theorem of rotating

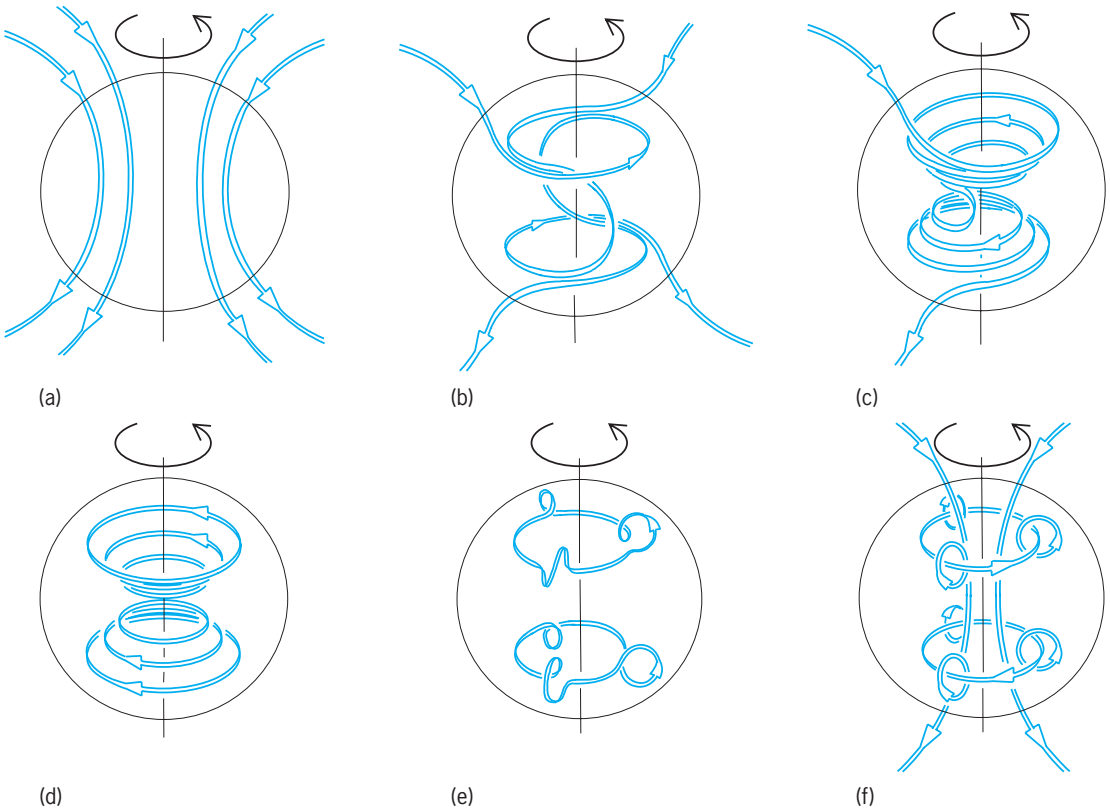

Fig. 2. The α-ω dynamo mechanism. Conventional geodynamo theory presupposes (*a*) an initial, primarily dipolar, poloidal magnetic field. The ω-effect consists of (*b*, *c*) differential rotation, wrapping the magnetic field around the rotational axis, thereby creating (*d*) a quadrupolar toroidal magnetic field. Symmetry is broken, and dynamo action maintained, by the α-effect, whereby (*e*) helical upwelling creates loops of magnetic field. (*f*) These loops coalesce to reinforce the original dipolar field, thus closing the dynamo cycle. (*After J. J. Love, Reversals and excursions of the geodynamo, Astron. Geophys., 40:6.14–6.19, 1999*)

fluid mechanics which states that small bumps on the boundary of a container of fluid can have a dramatic effect on the fluid's motion, even far from the boundary itself. Since the mantle convects very slowly compared to the core (their overturn times are approximately 10^8 and 10^3 years, respectively), the relatively steady nature of the conditions established by the mantle at the core-mantle boundary could also affect the pattern of core flow, thereby producing persistent asymmetric features in the magnetic field. Coupling between the core and mantle could arise from topography on the core-mantle boundary, or from thermal or electrical conductivity variations at the base of the mantle. *See* BOUNDARY LAYER FLOW; FLUID MECHANICS.

With an understanding of the boundary conditions applicable to the core, it should, in principle, be possible to solve the nonlinear magnetohydrodynamic equations that govern the geodynamo. The dynamo equations describe deterministic chaos, and data show that the magnetic field can exhibit highly aperiodic variation. Interestingly, the equations are symmetric under change in sign of the magnetic field, and thus there is no reason to expect that the Earth's field should have one polarity or the other. Moreover, a geomagnetic polarity reversal can occur with only a slight change in core's fluid motion. With respect to the cause of polarity transitions, mathematical models of hypothetical magneto-mechanical systems, in particular the Rikitake disc-dynamo system, exhibit irregular reversing behavior. So, it is not necessary to invoke external random factors such as ice ages or meteorite impacts to account for the irregular occurrence of polarity reversals.

The Sun is a familiar dynamo, and it reverses regularly almost every 11 years. So, why does the Earth's magnetic field not display such regularity? The difference is thought to be due to the presence in the Earth of a solid electrically conducting inner core, where the magnetic field can change only rather slowly by diffusion. Recent calculations suggest that because the inner core is electromagnetically coupled to the outer core, its presence acts to stabilize the magnetic field, so that only particularly large fluctuations of the field in the outer core are sufficient to overcome the damping effect of the inner core. *See* DIFFUSION; ELECTROMAGNETISM; GEOPHYSICS; MAGNETIC REVERSALS.

J. J. Love

Bibliography. F. H. Busse, Recent developments in the dynamo theory of planetary magnetism, *Annu. Rev. Earth Planet Sci.*, 11:241–268, 1983; J. A. Jacobs, *Deep Interior of the Earth*, Chapman & Hall, New York, 1994; J. J. Love, Reversals and excursions of the geodynamo, *Astron. Geophys.*, 40:6.14–6.19, 1999; E. N. Parker, Magnetic fields in the cosmos, *Sci. Amer.*, 249(2):44–54, 1983.

Geoelectricity

Electromagnetic phenomena and electric currents, mostly of natural origin, that are associated with the Earth. Geophysical methods utilize natural and artificial electric currents to explore the properties of the Earth's interior and to search for natural resources (for example, petroleum, water, and minerals). Geoelectricity is sometimes known as terrestrial electricity. All electric currents (natural or artificial, local or worldwide) in the solid Earth are characterized as earth currents. The term telluric currents is reserved for the natural, worldwide electric currents whose origins are almost entirely outside the atmosphere. Geoelectromagnetism is a more comprehensive term than geoelectricity. Time variations of any magnetic field are associated with an electric field that induces electric currents in conducting media such as the Earth.

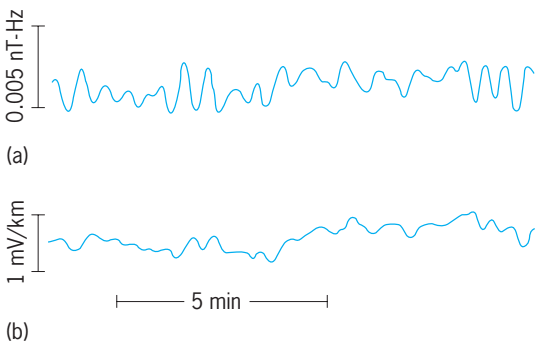

Fig. 1. Time variations of the horizontal orthogonal components of the natural (*a*) magnetic and (*b*) electric fields, simultaneously measured at one site at the surface.

Magnetic fields, electric fields, and electric currents are the constituents of electromagnetism, and are related by Maxwell's equations. For instance, **Fig. 1** shows the time variations of the natural magnetic and electric fields simultaneously measured at one location at the surface of the Earth. These two traces are related to each other, not only by Maxwell's equations but also by the physical properties of the subsurface rocks in the vicinity of the measuring site. Either one of the two traces may be computed synthetically from the other if the properties of the subsurface rocks are known. Conversely, the two traces together can yield geologic information; this is a form of geophysical exploration or prospecting. Thus, the terms geoelectricity, geomagnetism, and geoelectromagnetism are essentially interchangeable, although each one may have a somewhat different emphasis. For example, the term geomagnetism is sometimes used for the study of the Earth's quasi-stationary main magnetic field. *See* GEOMAGNETISM.

Measurements of electric and magnetic fields. A component of the electric field in a desired direction is measured by planting two electrodes (for example, metal stakes or special nonpolarizable electrodes) aligned in that direction. The electrodes are connected by an insulated wire, and voltage difference between them is measured with a voltmeter of high input impedance (for example, 10 megohms). The average electric field between the electrodes is expressed in units of volts per meter. Since this unit is very cumbersome for measuring the Earth's field, it is customary to use millivolts per kilometer. (Figure 1*b* and **Fig. 2** show the time variations of such components.) To obtain the total horizontal electric field, two orthogonal components, north–south (N–S) and east–west (E–W), are measured by means of an L-shaped electrode array. The trajectory of the head of the electric field vector is traced by feeding the two components into an oscilloscope or a paper X–Y recorder (**Fig. 3**). The magnetic field is measured by magnetometers. The cryogenic magnetometer has a resolution of better than 1 picotesla, 1 part in 50,000,000 of the Earth's total magnetic field. The nanotesla (nT) or gamma γ is used in practice. Figure 1*a* shows the time variations of one horizontal component of the Earth's natural magnetic field, measured with a coil-core magnetometer whose output is the time derivative of the magnetic field, with the scale given in terms of nanoteslas times frequency. Worldwide studies of natural electromagnetic phenomena are made by monitoring primarily the magnetic field rather than the electric field, which is much more affected by local geology.

Electric earth currents. These may be local or worldwide.

Local. Such currents can be natural or caused by human activities. The latter (called stray, industrial, or cultural currents) may be caused by electric trains, rural water pumps, and pipelines. Natural local

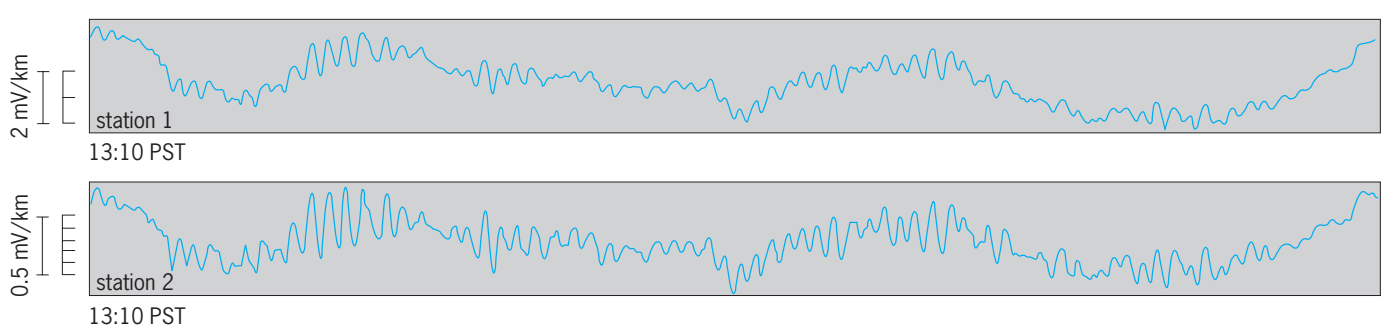

Fig. 2. Two tellurograms (stations 1 and 2, San Joaquin Valley, California) representing the time variations of the natural electric field (micropulsations), N60°E components, simultaneously recorded over a time interval of 30 min. The recording sites are separated by 27 mi (43 km) in the direction of the components. PST = Pacific Standard Time.

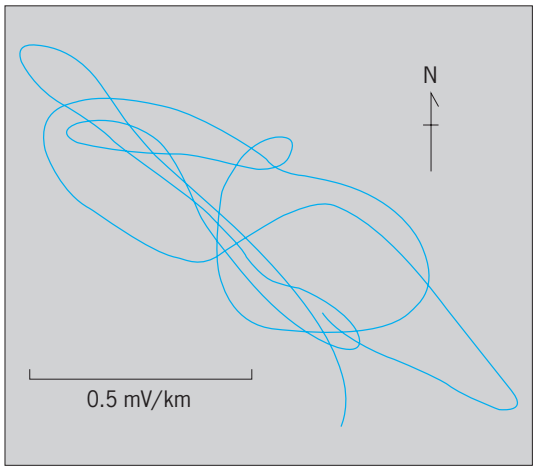

Fig. 3. A vectogram representing a few minutes of recording of the natural electric field vector. The band-pass filter peaked at the 20-s period (0.05 Hz).

currents represent the phenomena of spontaneous potentials or self-potentials. Some deposits in the Earth, such as certain metallic sulfides and graphite, constitute buried natural electric cells because of their high electrical conductivity and also because of oxidation and reduction processes associated with ground water. Thus, a hidden ore body, such as a copper ore deposit, can be discovered by measuring the electric field at the surface of the Earth, which may be as large as 1 V over a distance of 300 ft (100 m). Two other sources of spontaneous potentials are ground water movements and topographic elevation changes.

Worldwide. Telluric currents are of natural origin. There are various types, sources, and frequencies (or periods) of the worldwide natural electromagnetic fields which are associated with electric currents in the Earth (**Fig. 4**). The time variations of these electromagnetic fields are simply called variations.

Secular variations. The Earth's main magnetic field is thought to be caused by motions in the electrically conducting fluid core of the Earth, which acts as a kind of dynamo, creating electric currents which in turn create the magnetic field. This field is not stationary, but has time variations with periods ranging from about 30 to 300 years per cycle, which are the secular variations. Electric currents at the surface of the Earth associated with the main field and its secular variations have not been monitored effectively because of the difficulties involved in separating them from other effects, such as electrode potentials and tidal potentials. *See* GEOMAGNETIC VARIATIONS.

Diurnal (daily) variations. The air layers of the ionosphere, from a height of about 60–200 mi (100–300 km), are ionized by solar radiation, while air below the ionosphere is practically nonconducting. The ionization (electrical conductivity) in the ionosphere is renewed daily. Tidal oscillations of the ionosphere in the presence of the Earth's main magnetic field constitute an atmospheric dynamo, inducing electric currents in the Earth. They are thought to be driven primarily by the thermal effects of the Sun and partially by the attraction of the Moon (**Fig. 5**). *See* IONOSPHERE.

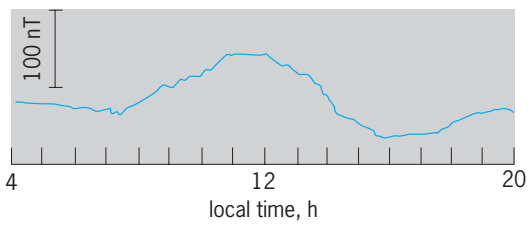

Fig. 5. Diurnal variations (solar plus lunar) of the horizontal component of the magnetic field, December 21, 1933, Huancayo, Peru. (*After S. Chapman and J. Bartels, Geomagnetism. 2 vols., Oxford University Press, 1962*)

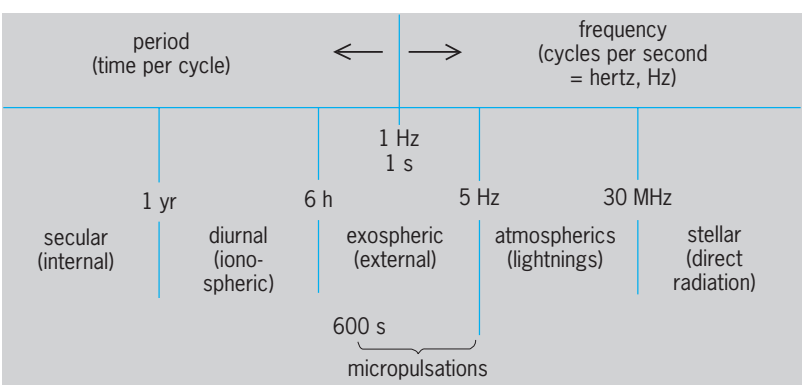

Fig. 4. Approximate and schematic frequencies and origins of the natural electromagnetic fields.

Exospheric-origin variations, or micropulsations. Short time fluctuations of the Earth's magnetic field (micropulsations) that fall within the approximate period range of 0.2–600 s per cycle (5–0.0017 Hz) occur almost continuously as a background noise. Amplitudes depend on latitude, solar activity, frequency, local time, season, and local geology, with worldwide and long-term statistical amplitudes of the order of a few millivolts per kilometer and a few tenths of a nanotesla. While the mechanism of their generation is not completely understood, it appears that micropulsations are generated by the magnetohydrodynamic effect through the interaction of the solar wind with the main magnetic field and atmosphere of the Earth. Study of the exospheric-origin electromagnetic phenomena constitutes a branch of geophysics called aeronomy. Figure 1 is a record of micropulsations measured at stations located in a sedimentary basin. The magnetic field trace (Fig. 1*a*) is called a magnetogram; the electric field trace (Fig. 1*b*) a tellurogram. Figure 2 shows two tellurograms simultaneously measured at two stations 27 mi (43 km) apart and in the direction of station separation. These measurements represent normal, usual activity on a quiet day. **Figure 6** shows the amplitude spectra of the tellurograms shown in Fig. 2. The differences between the two tellurograms, and consequently between the two spectra,

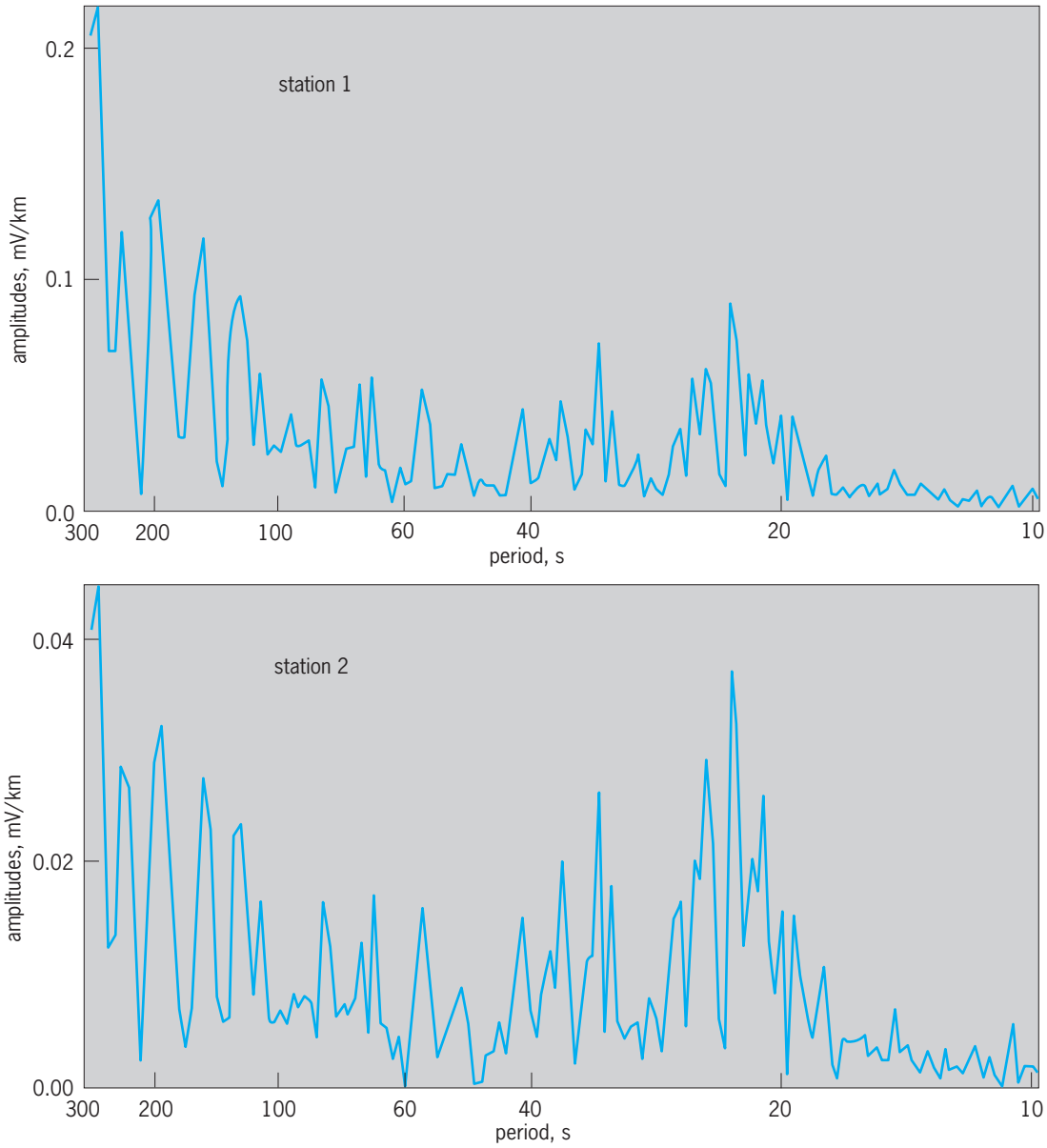

Fig. 6. Amplitude spectra of the tellurograms shown in Fig. 2.

are almost totally due to the differences in the geologic conditions at the two measuring sites (stations). Such measurements can be used for geologic exploration of the subsurface. *See* MAGNETOHYDRODYNAMICS; SEISMIC STRATIGRAPHY; SOLAR WIND; UPPER-ATMOSPHERE DYNAMICS.

Magnetic storms are very intense disturbances of long duration that occur about once a month on the average. Caused by large-scale bursts of solar wind associated with sunspots and solar flares, they usually commence suddenly and almost instantaneously (within about 0.5 min) throughout the world. Their amplitudes may reach hundreds of nanoteslas and hundreds of millivolts per kilometer, disrupting radio and telegraph communications. It is interesting to note that they cause fish to migrate into deeper waters. **Figure 7** shows the records of a magnetic storm. Magnetic storms are frequently associated with aurorae polares (northern or southern lights), which are seen as spectacular luminous formations at ionospheric heights. *See* AURORA.

Atmospherics. The major cause of the variations within the frequency range of about 5–10 kHz is the lightning occurring almost continuously in Central Africa and in the Amazon region. While audio-frequency variations are included in atmospherics, lightning itself is a concern of meteorology. *See* LIGHTNING; SFERICS.

Stellar variations. Above the frequency of 30 MHz, these originate predominantly from the direct radiation of electromagnetic waves propagated by the Sun.

Subsurface geophysical exploration. Electrical methods, more properly called electromagnetic methods, are used to explore the subsurface from depths of a few inches (for example, popular coin detectors

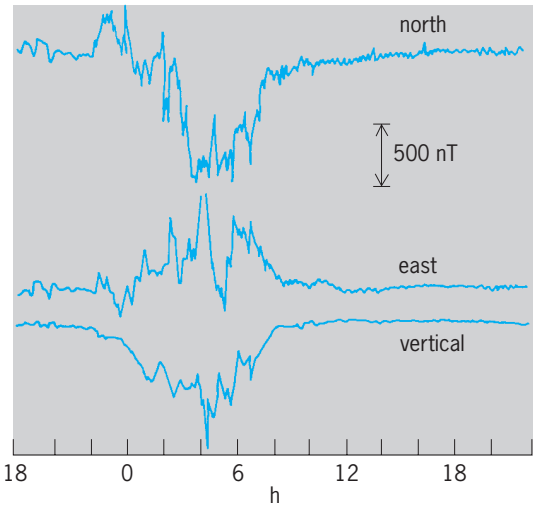

Fig. 7. Three components of the magnetic field of a magnetic storm of May 14, 1921, Potsdam, Germany. (*After S. Chapman and J. Bartels, Geomagnetism, 2 vols., Oxford University Press, 1962*)

or mine detectors) down to depths of hundreds of miles. In general, these methods require an input into the Earth, either an artificial direct or alternating electric current, or a natural electromagnetic field, such as micropulsations or diurnal variations. This input is a source signal coupled with the Earth, which behaves as a filter whose response is measured in terms of electric or magnetic fields. It is analogous to measuring the input and output of an electronic filter to determine its characteristics, which in this case is the geologic information sought. These methods supply only the electrical properties of the subsurface (mainly the electrical conductivity). Different rocks have, in general, different conductivities. For instance, limestones usually have much lower conductivities than clay-rich shales. A knowledge of the conductivity distribution in the subsurface, combined with other geologic information, allows interpretation of the rock-type distribution. Artificial direct-current methods involve feeding a current into the Earth with a pair of electrodes and measuring the resulting electric field with another pair of electrodes. The alternating-current methods use magnetometers to measure the magnetic field created by inducing currents in the Earth. The two most popular methods employing the natural electromagnetic fields are the magnetotelluric and telluric methods. The magnetotelluric method requires simultaneous measurements of the electric and magnetic fields at one site. Figure 1 represents such a data set. The telluric method requires only the measurements of the electric field, made simultaneously at two or more sites (Fig. 2). These electromagnetic (or electrical) methods are unlike the magnetic methods of geophysical exploration, which deal only with the magnetization of rocks due to the main magnetic field of the Earth. *See* GEOPHYSICAL EXPLORATION.

S. H. Yungul

Bibliography. G. V. Keller and F. C. Frischknecht, *Electrical Methods in Geophysical Prospecting*, 1966; S. H. Yungul, The telluric methods in the study of sedimentary structures: A survey, *Geoexploration*, 15:207–238, 1977; M. S. Zhdanov and G. V. Keller, *The Geoelectrical Methods in Geophysical Exploration*, 1993.

Geographic information systems

Computer-based technologies for the storage, manipulation, and analysis of geographically referenced information. Attribute information and spatial information are integrated in geographic information systems (GIS) through the notion of a data layer, which is realized in two basic data models—raster and vector (**Fig. 1**). The major categories of applications comprise urban and environmental inventory and management, policy decision support, and planning; public utility and business (including agribusiness) applications, engineering, and defense applications; and scientific analysis and modeling.

Characteristics. A geographic information system differs from other computerized information systems in two major respects. First, its information is geographically referenced (geocoded). Geocoding is usually achieved by recording the geographical coordinates of the objects of interest (which may be the corner points of a plot on a cadastral map, or the location of cities of more than 100,000 inhabitants at the global scale). Alternatively, the location of any point of interest within an area can be inferred on the basis of a number of reference points registered for that area. Second, a geographic information system has considerable capabilities for data analysis and scientific modeling, in addition to the usual data input, storage, retrieval, and output functions.

A geographic information system can answer arbitrarily complex queries about things on or near the surface of the Earth, their attributes, and the spatial relationships (such as distance, direction, adjacency, and inclusion) among them. Basic kinds of queries supported by this system include (1) location questions, to determine the attributes of a given place (for example, What tree types are found in a given forestry tract?); (2) condition questions, seeking to find locations fulfilling certain conditions (for example, Where are vacant lots larger than 5 acres and within 1 mile of a paved road?); (3) trend questions, seeking to determine changes in place attributes over time (for example, How much has the population of a specific census tract grown between 1980 and 1990?); (4) routing questions, especially useful in situations where vehicles need to be dispatched from a place of origin to a destination (for example, Which is the shortest, quickest, or safest route from an ambulance station to the site of an emergency?); and (5) pattern questions, whereby scientists or managers can investigate the spatial distribution of some phenomenon for diagnostic purposes or in the course of exploring some scientific hypothesis (for example, Is the density of diseased trees greater around campgrounds? Are pedestrian traffic

casualties higher than expected in low-income neighborhoods?).

Elements. A geographic information system is composed of software, hardware, and data. Some authors also include the users and their institutional context. Originally used on mainframe computers, geographic information systems did not become popular until the 1980s, when software packages that could run on workstations and desktops became widely available. Some of the most powerful systems now run on minicomputers, but the field is dominated by packages developed for workstation platforms in the case of the larger systems and desktops for the rest. The trend in the late 1990s was toward flexible, portable, user-friendly systems, increasingly capable of interoperating, increasingly geared toward the Internet, and in many cases tailored to specific kinds of applications. Often these special-purpose geographic information systems are coupled with other digital systems or devices such as global positioning systems (GPS); intelligent vehicle highway systems (IVHS); navigational devices for vehicles, ships, or aircraft; or devices for the automatic delivery of agricultural fertilizers as a function of soil quality. *See* HIGHWAY ENGINEERING.

The notion of data layer (or coverage) and overlay operation lies at the heart of most software designed for geographic information systems. The landscape is viewed as a collection of superimposed elementary maps, each storing information pertinent to one aspect or attribute of the landscape: relief, soils, hydrology, vegetation, roads, land use, land ownership, and so forth. By combining the corresponding information on several layers, the composite properties of any object of interest can be deduced. (This approach is the electronic version of a traditional manual method involving transparent sheets.) Two fundamental data models, the vector and raster models, embody the overlay idea in geographic information systems (**Fig. 2**). In a vector geographic information system, the geometrical configuration of a coverage is stored in the form of points, arcs (line segments), and polygons, which constitute identifiable objects in the database. In a raster geographic information system, a layer is composed of an array of elementary cells or pixels, each holding an attribute value without explicit reference to the geographic feature of which the pixel is a part. Both models have strengths and drawbacks. *See* COMPUTER GRAPHICS; ELECTRONIC DISPLAY; IMAGE PROCESSING.

A data layer or coverage integrates two kinds of information: attribute and spatial (geographic). The functionality of a geographic information system consists of the ways in which that information may be captured, stored, manipulated, analyzed, and presented to the user. Spatial data capture (input) may be from primary sources such as remote sensing scanners, radar, or global positioning systems, or from scanning or digitizing images and maps derived from remote sensing. Output (whether as a display on a cathode-ray tube or as hard copy) is usually in map or graph form, accompanied by tables and reports linking spatial and attribute data. The critical data management and analysis functions fall into four categories: retrieval, classification, and measurement (for example, measurement of the area inside a polygon); overlay functions; neighborhood operations; and connectivity functions. *See* DATABASE MANAGEMENT SYSTEM; DIGITAL COMPUTER; REMOTE SENSING.

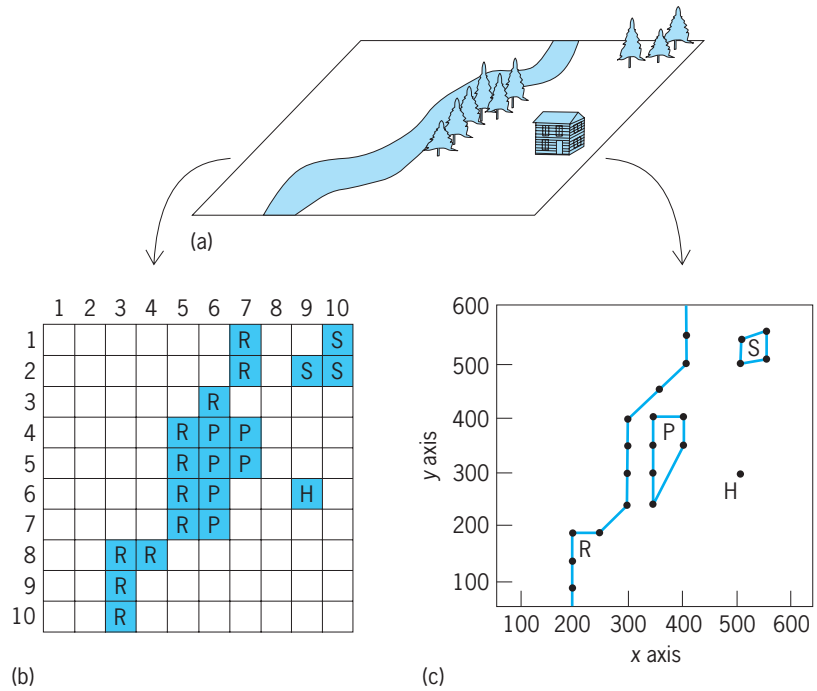

Fig. 1. Comparison of the raster and vector models. The pine forest stand (P) and spruce forest stand (S) are area features. The river (R) is a line feature, and the house (H) is a point feature. (*a*) Landscape. (*b*) Raster representation. (*c*) Vector representation. (*After S. Aronoff, Geographic Information Systems: A Management Perspective, WDL Publications, 1989*)

Uses. Geographic information systems find application mainly in three areas: business, engineering, and defense; management, planning, and policy; and research in various disciplines.

Business, engineering, and defense. Business applications are increasingly widespread and include market analysis (for example, identifying a customer base in a given area), store location (for example, identifying the most competitive location for a new chain store outlet), and agribusiness (for example, determining the correct amount of fertilizers or pesticides needed at each point of a cultivated field). Engineers use geographic information systems when modeling terrain, building roads and bridges, maintaining cadastral maps, routing vehicles, drilling for water, determining what is visible from any point on the terrain, integrating intelligence information on enemy targets, and so forth. Such applications have been facilitated through the integration of geographic information systems with global positioning systems, involving the automatic determination of latitude and longitude coordinates and elevation based on the processing of signals from a network of satellites. Because global positioning systems allow the determination of position with unparalleled ease, speed, and accuracy, mobile units coupling geographic

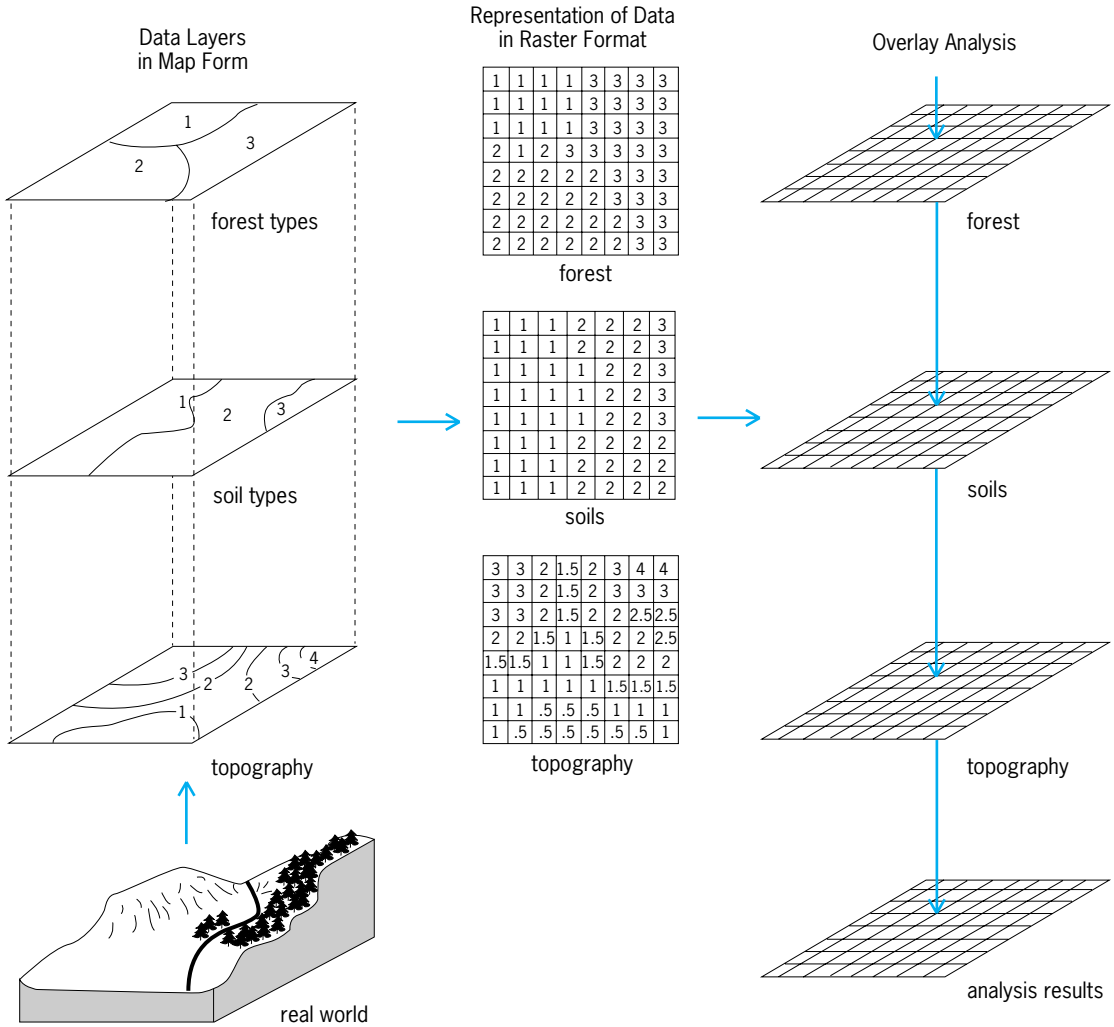

Fig. 2. Overlay analysis using raster data files. (*After S. Aronoff, Geographic information Systems: A Management Perspective, WDL Publications, 1989*)

information and global positioning systems have greatly increased the range of applications, especially for work in the field and navigation. *See* COORDINATE SYSTEMS; SATELLITE NAVIGATION SYSTEMS.

Management, planning, and policy. There are many management applications of geographic information systems. Among the earliest and still most widespread applications of the technology are land information and resource management systems (for example, forest and utility management). Other common uses of geographic information systems in an urban policy context include emergency planning, determination of optimal locations for fire stations and other public services, assistance in crime control and documentation, and electoral and school redistricting.

Nonurban applications include ecological area management (for example, estuarine or National Parks management), hazardous waste facility location and management, and biodiversity preservation projects. A significant development is the introduction of geographic information systems in third-world environmental management and planning. This development was made possible in the 1980s through the wide availability of inexpensive desktop computer platforms for geographic information systems, by the increasing ease of use of such computers, and by the diffusion of educational opportunities in geographic information. Major applications in developing countries include natural resource management, soil conservation and irrigation projects, and the siting of health and other services.

More sophisticated developments involve a new class of software known as spatial decision support systems, which can aid the policy-making and decision process whenever location is part of the problem context. A typical system of this type presents the different options, facilitates the comparison of costs and benefits, helps explore the consequences of changing the weights attributed to different decision criteria, and allows unanticipated alternatives (for example, solutions proposed by third parties) to be easily formulated and evaluated. The power of spatial decision support systems lies primarily in their flexible, interactive nature and in the visualization of the complex issues debated. *See* DECISION SUPPORT SYSTEM.

Other applications. Uses of geographic information systems have spread well beyond geography, the source discipline, and now involve most applied sciences, both social and physical, that deal with spatial data. These sciences include sociology, archeology and anthropology, urban studies, epidemiology, ecology, forestry, hydrology, and geology. The nature of the applications of geographic information systems in these areas ranges from simple thematic mapping for illustration purposes to complex statistical and mathematical modeling for the exploration of hypotheses or the representation of dynamic processes (such as the spread of fire on a vegetated landscape or the investigation of trends in crime patterns in an urban area). *See* ANTHROPOLOGY; ARCHEOLOGY; ECOLOGY; EPIDEMIOLOGY; FOREST AND FORESTRY; GEOLOGY; HYDROLOGY.

Conceptual and scientific issues. While the geographic information system is still considered largely a technology, its coming of age has been marked by an increasing emphasis on its conceptual and scientific aspects. There are a number of fundamental research questions concerning geographic information systems. They involve (1) the nature of spatial data, including questions of measurement, sampling, uncertainty, error, scale, and the nature of geographic space itself (Is it a container of geographical objects, or is it defined through the relations among objects?); (2) the digital representation of geographic space and phenomena, and in particular questions of database structures, models, and semantics; (3) functional issues, including the appropriate basic operations the geographic information system should support, the propagation of errors and uncertainties in spatial databases, and the appropriate query languages; (4) display issues, such as questions of visualization, user interfaces, and cartographic design; and (5) operational issues, dealing with quality controls and standards for both the systems themselves and their products, with legal and management issues, and with any other issues affecting the fitness of geographic information systems for their numerous kinds of current and potential applications. A major recent research theme involving all these issues concerns the interoperability of geographic information systems, or the possibility to move or share data, functions, and operations across systems and platforms. Helen Couclelis

Bibliography. J. C. Antenucci et al., *Geographic Information Systems: A Guide to the Technology*, 1991; T. Bernhardsen, *Geographic Information Systems: An Introduction*, 3d ed., 2002; P. A. Burrough and R. A. McDonnell, *Principles of Geographical Information Systems*, 2d ed., 1998; M. N. DeMers, *Fundamentals of Geographic Information Systems*, 3d ed., 2002; M. Fischer and P. Nijkamp (eds.), *Geographic Information Systems: Spatial Modeling and Policy Evaluation*, 1993; R. C. Frohn, *Remote Sensing for Landscape Ecology: New Metric Indicators for Monitoring, Modeling, and Assessment of Ecosystems*, 1998; P. Longley et al., *Geographical Information Systems: Principles and Applications*, 2d ed., 1998.

Geography

The study of physical and human landscapes, the processes that affect them, how and why they change over time, and how and why they vary spatially. Geographers determine the factors that influence landscapes or landscape features and seek to explain the interactions among them. This often involves explaining landscapes or environments from a spatial perspective by analyzing patterns or the lack of patterns. Geographers are also interested in the spatial arrangements of natural and human phenomena, as well as the relationships among these phenomena. Consequently, a key element of geography is areal differentiation, the process of defining regions with common unifying characteristics.

Geography is a broad-based discipline that integrates many aspects of the social and physical sciences. The discipline has been described as consisting of four traditions: Earth science, area studies, spatial analysis, and human–environment interactions. Earth science includes studies of the atmosphere, hydrosphere, lithosphere, and biosphere. Area studies refer to the detailed geographic knowledge and understanding of particular regions of the world. Spatial analysis deals with the measurement, analysis, prediction, and explanation of spatial organization, as well as the interactions, patterns, and factors that influence change in these landscape elements. Human–environment interaction is concerned with the relationships between humans and their environment. Generally, this includes the ways that humans modify and influence their physical environment to meet certain needs, as well as the ways that humans modify their behavior and surroundings to adapt to particular environmental circumstances. A majority of studies in geography concern at least one of these four traditions. *See* ATMOSPHERE; BIOSPHERE; HYDROSPHERE; LITHOSPHERE.

Geographers consider, to varying degrees, both natural and human influences on the landscape, although a common division separates physical and human geography. Physical geographers study landforms (geomorphology), water (hydrology), climate and meteorology (climatology), biotic environments (biogeography), and soils (pedology). Human geographers focus on cultural geography, economic geography, political geography, location analysis, and spatial analysis of ethnic or gender issues, as well as transportation, area studies, and urban, regional, or environmental planning. *See* BIOGEOGRAPHY; CLIMATOLOGY; GEOMORPHOLOGY; HYDROLOGY; PEDOLOGY; PHYSICAL GEOGRAPHY.

Most professional geographers specialize in one or two subdisciplines. In addition, many are involved in developing techniques and applications that support spatial analytical studies or the display of spatial information and data. Maps—printed, digital, or conceptual—are the basic tools of geography. Geographers are involved in map interpretation and use, as well as map production and design. Cartographers supervise the compilation, design, and development

of maps, globes, and other graphic representations. *See* CARTOGRAPHY.

The techniques, tools, and methods used in geographic and spatial analytical work have undergone a revolution since the advent of relatively inexpensive computer graphics systems and technologies such as satellite imaging. Remote sensing specialists, often geographers, work to interpret and enhance digital and photographic images, particularly those gathered from aircraft or satellites. A geographic information system (GIS) combines the advantages of computer-assisted cartography with those of spatial database management, and has widespread applications. A geographic information system facilitates the storage, retrieval, and analysis of spatial information in the form of digital map "overlays," each representing a different landscape component such as terrain, hydrologic features, roads, vegetation, soil types, or any mappable factor. Each of these data layers can be fitted digitally in any combination to the same map scale and map projection, permitting the analysis of relationships among combinations of environmental variables. *See* COMPUTER GRAPHICS; GEOGRAPHIC INFORMATION SYSTEMS; MAP PROJECTIONS; REMOTE SENSING; SCIENTIFIC AND APPLICATIONS SATELLITES.

Many geographers are applied practitioners, solving problems using a variety of tools, particularly computer-assisted cartography, statistical methods, remotely sensed imagery, the Global Positioning System (GPS), and geographic information systems. Today, nearly all geographers, regardless of their focus, use some or all of these techniques in their professional endeavors. *See* SATELLITE NAVIGATION SYSTEMS.

James F. Petersen

Bibliography. R. Abler, M. Marcus, and J. Olson, *Geography's Inner Worlds: Pervasive Themes in Contemporary American Geography*, 1992; G. Gaile and C. Willmott (eds.), *Geography in America at the Dawn of the 21st Century*, 2005; Geography Education Standards Project, *Geography for Life: National Geography Standards*, 1994; P. Gould, *Becoming a Geographer*, 1999; National Research Council, *Rediscovering Geography: New Relevance for Science and Society*, 1997.

Geologic thermometry

The measurement or estimation of temperatures at which geologic processes take place. Methods used can be divided into two groups, nonisotopic and isotopic. The isotopic methods involve the determination of distribution of isotopes of the lighter elements between pairs of compounds in equilibrium at various temperatures, and application of these data to problems of the temperature at which these compounds (commonly minerals) form in nature.

Nonisotopic Methods

Earth temperatures can be measured directly by surface and near-surface features or indirectly from various properties of minerals and fossils.

Direct measurement. Temperatures can be measured directly in hot springs, fumaroles, flows of lava, and artificial openings such as mines, boreholes, and wells.

Hot springs. The temperatures of hot springs range from slightly above the mean annual temperature of the region in which they occur to the boiling point of water at the elevation of the outlets. In other words, in temperate regions and at moderate altitudes the temperatures of hot springs range from about 20 to 100°C (68 to 212°F).

Fumaroles. At temperatures above its boiling point, water issues as steam from vents called fumaroles: temperatures of these fumaroles have been measured up to 560°C (1040°F) near Vesuvius, and up to 645°C (1193°F) in the Valley of Ten Thousand Smokes (Alaska). Fumaroles in lavas may reach temperatures of 700–800°C (1290–1470°F).

Lava. Lava is molten rock flowing onto the Earth's surface. Its temperature on extrusion ranges from about 700 to 900°C (1290 to 1650°F) for andesitic and dacitic lava and to 1200°C (2190°F) for basaltic lava. The viscosity of lava increases with decreasing temperature, and basaltic lava ceases to flow when it cools to 700–800°C (1290–1470°F). Intrusive magmas of similar compositions are probably intruded at similar temperatures, as indicated by their effects on the coal beds into which they are intruded, and the forms and assemblages of the first minerals to crystallize. *See* LAVA.

The temperature to which rocks around an intrusion (country rocks) are heated depends on many factors: temperature of the magma; temperature, composition, and structure of the country rock; abundance and nature of solutions given off by the intrusion; and size of the intrusion. In general, the temperature of the country rocks at the contact will be much lower than that of the magma. For example, an intrusive sheet of dolerite at 1100°C (2010°F) may heat the contact rocks to 600–700°C (1110–1290°F). *See* MAGMA.

Mines, boreholes, and wells. Temperatures have been measured in enough artificial openings in the Earth's crust so that the temperature distribution is well known in many areas to depths of several thousand feet. Measurements of gradients range from about 12 to 52 m/°C (22 to 94 ft/°F) in nonvolcanic areas. The highest temperature yet encountered in such an area is 154°C (309°F) from a well 20,521 ft (6255 m) deep in Sublette County, Wyoming.

Gradients in volcanic areas are much higher near the surface (up to 0.4 m/°C or 0.7 ft/°F), but at depths of 750–1000 ft (230–300 m) a temperature of about 250°C (482°F) is commonly reached and persists to much greater depths (to at least 5500 ft or 1.7 km in Tuscany).

Calculations and extrapolations give greatly different pictures for temperature distribution from the zone of measurements to the center of the Earth, depending on the assumptions made. Estimates of the temperature at the center of the Earth range from 1600 to 76,000°C (2900 to 137,000°F).

Indirect methods. Some indirect nonisotopic methods appear to give estimates with the same accuracy as direct measurements; others place the temperature within a certain range; still others indicate only that a given process took place above or below a certain temperature.

As phase-equilibrium relationships in systems analogous to those in nature become more accurately known, it will be possible to make more exact estimates of geologic temperatures. The relationships commonly employed are listed below. *See* PHASE EQUILIBRIUM.

Melting points. The melting point of a mineral, corrected for the pressure under which it was formed, gives a maximum temperature of formation for the assemblage in which it grew because other substances, especially water, lower the crystallizing temperature for a mineral. For example, if realgar, AsS, occurs in a vein with other minerals in such a way that they must have crystallized simultaneously, then the whole assemblage must have formed at a temperature lower than 320°C (608°F), the melting point of realgar.

Transformation temperatures (inversions). Many minerals have two or more crystalline modifications which form, or exist, in different temperature ranges. For example, under certain conditions marcasite forms at temperatures below 300°C (570°F), but at about 450°C (840°F) it transforms to pyrite at an appreciable rate. Therefore, a coprecipitated mineral assemblage including marcasite was certainly formed below 450°C (840°F) and probably below 300°C (570°F).

Other pairs of minerals transform in either direction at a definite temperature; for example, low (α) quartz changes to high (β) quartz when it is heated to 573°C (1060°F), and high quartz changes to low when it is cooled below 573°C (1060°F). Therefore, phenocrysts of high quartz in lavas were formed above 573°C (1060°F), crystals of low quartz in veins, below 573°C (1060°F).

Dissociation and decomposition temperatures. Many minerals break up when they are heated. If one of the products is a gas, the temperature of decomposition changes rapidly with pressure; the pressure at the time of formation must be known or estimated before such a mineral can be used as a geologic thermometer. For example, under atmospheric pressure, calcite ($CaCO_3$) dissociates into lime and carbon dioxide (CO_2) at 885°C (1620°F), but under a sufficiently high pressure of CO_2 (1025 atm or 104 megapascals) it melts at 1339°C (2442°F) without decomposing.

When only solids and liquids are involved, however, pressure is relatively unimportant and can be neglected for processes that take place near the Earth's surface. For example, danburite decomposes to two liquids at about 1000°C (1830°F), so this mineral and others immediately associated with it in pegmatities must have formed at temperatures below 1000°C (1830°F).

Solid solutions and exsolution pairs. Many pairs of minerals with similar structures form homogeneous solid solutions at high temperatures, but on cooling separate (exsolve) into lamellae of the two minerals. This process commonly produces a characteristic texture that can be recognized. When two such minerals occur in a rock or ore in an exsolution relationship, this is evidence that the temperature of formation was above their temperature of homogenization. *See* SOLID SOLUTION.

Some common exsolution pairs and their temperatures of homogenization are magnetite-spinel, 1000°C (1830°F); ilmenite-hematite, 600–700°C (1110–1290°F); chalcopyrite-bornite, 500°C (930°F); chalcopyrite-cubanite, 450°C (840°F); bornite-tetrahedrite, 275°C (527°F); and bornite-chalcocite, 225°C (437°F). These temperatures depend upon the composition of the host phase, but the ones given here are for compositions commonly encountered in nature.

This method has been used to estimate the temperatures of formation of many ore deposits. Some examples are sulfide replacement in Gilman, Colorado, 150–300°C (300–570°F); sulfide mineralization at Pine Vale, Queensland, 475–500°C (807–930°F); scheelite veins, Australia, 350–600°C (660–1110°F); and sphalerite-stannite-chalcopyrite mineralization, Tasmania, about 600°C (1110°F).

In some mineral pairs there is a limited amount of solid solution: the amount depends upon the temperature of formation. For example, in the system iron sulfide- zinc sulfide (FeS-ZnS) the amount of FeS in sphalerite is a function of the temperature of formation, if there is an excess of FeS (pyrrhotite) present when the sphalerite crystallizes. Sphalerite formed at 200°C (390°F) under these conditions will contain about 7 mol % FeS; at 500°C (930°F), 18 mol %. Other associations that can be used to indicate temperature of formation in this way are scandium in biotite (in a given petrologic province), titanium dioxide (TiO_2) in magnetite (with coexisting ilmenite), iron and magnesium in coexisting olivines and pyroxenes, albite in potassium feldspar (with coexisting plagioclase), and magnesium carbonate ($MgCO_3$) in calcite (with coexisting dolomite).

By using the FeS-ZnS relationship, the temperature of formation of sphalerite in various settings has been estimated: in graphite schist near a granite contact (Norway), 440°C (820°F); replacement deposits, Gilman, Colorado, 380–600°C (710–1110°F) and Broken Hill, Australia, 600°C (1110°F); and in uranium deposits, Colorado Plateau, 138°C (280°F). TiO_2 in the magnetite of the northwestern Adirondacks indicates a temperature of formation of 475–600°C (807–1110°F). The composition of alkali feldspars in granite indicates a final consolidation temperature of about 600°C (1110°F). A metamorphic calcite with 10% $MgCO_3$ forms at about 650°C (1200°F); with 20%, at 830°C (1520°F).

Eutectics. When two or more minerals crystallize simultaneously at a eutectic, a so-called eutectic texture may be produced. However, it is difficult to be certain that a natural intergrowth of minerals was produced by eutectic crystallization. Therefore, this method has been little used thus far

in geologic thermometry. It can be used in the same way that melting points can, that is, to indicate a temperature that cannot have been exceeded during crystallization of the assemblage. Some eutectic temperatures of common minerals are iron-nickel, 1435°C (2615°F); anorthite-diopside, 1270°C (2320°F); albite-nephelite, 1068°C (1954°F); orthoclase-albite, 1070°C (1958°F); orthoclase-silica, 990°C (1814°F); quartz-orthoclase-albite, 937°C (1719°F); silver-copper, 785°C (1445°F); chalcocite-galena-argenite, 400°C (750°F); and sulfur-selenium, 100°C (212°F). Some of these temperatures are affected markedly by volatile components of a magma. They may be lowered hundreds of degrees by 1000 atm (101 MPa) or more of water vapor pressure. *See* EUTECTICS.

Mineral assemblages. Certain types of mineral assemblages, such as eutectics and exsolution pairs, have been discussed. Other types that do not belong to one of these classes can also give indications of temperatures of formation. These indications may be based on syntheses, including hydrothermal experiments, known stability ranges of the individual minerals of the assemblage, and effect of pressure (where volatiles such as H_2O and CO_2 occur).

For example, hydrothermal experiments have shown that analcite forms at temperatures of about 100–380°C (212–716°F) under moderate water vapor pressures, but in runs of the same composition albite forms above 380°C (716°F). Therefore, in mineral assemblages that formed near the Earth's surface, that is, in cavities in volcanic rocks or shallow intrusions, the presence of analcite can be taken to indicate formation temperatures below about 400°C (750°F) and the presence of albite to indicate higher temperatures. Good crystals of potassium feldspar (variety adularia) have been formed hydrothermally at 245°C (473°F), and quartz crystals have been formed down to about 200°C (390°F). Well-formed crystals of these minerals, even in sedimentary rocks, indicate growth temperatures of at least 100°C (212°F) and probably 200°C (390°F) or higher.

Clay minerals can also be important indicators of temperatures of formation. Their stability ranges are affected by such factors as pH of the solutions, water vapor pressure, and concentrations of constituent cations in the solutions, but some useful generalizations can be made without quantitative evaluation of these variables. Kaolin forms in acid solutions up to about 350°C (660°F) if aluminum is high and potassium is low, but when the pH is significantly above 7, montmorillonite forms from the same compositions over the same temperature range. Coprecipitated gels of alumina and silica (neutral) produce kaolin up to a little over 300°C (570°F), dickite at about 345–360°C (650–680°F), and beidellite at 360–390°C (680–730°F). *See* CLAY MINERALS.

Sepiolite and attapulgite are decomposed hydrothermally at temperatures at least as low as 200°C (390°F); thus the presence of either of these in a mineral assemblage indicates temperature of formation not over 200°C (390°F) and possibly below 100°C (212°F). The lower limit of formation of these minerals is not known.

Similar experiments give analogous results for other minerals and mineral groups. Sericite, for example, forms at about 200–525°C (390–977°F) in slightly basic to somewhat more acid solutions if aluminum and potassium are both high. Pyrophyllite forms at about 300–500°C (570–1020°F) if aluminum and potassium are both low. Serpentine cannot form above about 500°C (930°F) even under very high pressures of water vapor, and most varieties of chlorite crystallize below 500°C (930°F). Muscovite is stable from about 400 to 800°C (750–1470°F) at pressures likely to be involved in rocks formed at depths small enough so they can be brought to the surface by erosion. Phlogopite forms from about 800 to 1100°C (1470 to 2010°F) in the same pressure range. Talc does not form above about 825°C (1510°F).

If two or more minerals have been formed in equilibrium, it may be possible to narrow the temperature range considerably. Though talc is stable at temperatures up to 825°C (1510°F), and in equilibrium with enstatite at moderate pressures, the assemblage is stable only from about 670 to 800°C (1230 to 1470°F). Likewise, serpentine can form in equilibrium with brucite only below 450°C (840°F) at moderate pressures or below 400°C (750°F) at low pressures.

So many mineral associations have been studied with respect to conditions of formation that only a few interesting examples can be mentioned in the brief review below. They range from very low temperature environments of clay minerals and carbonates through hydrothermal mineral deposits to very high temperature and pressure assemblages of kimberlites, peridotites, and other deep-seated rocks.

1. *Low-temperature processes*. (a) Montmorillonite loses water at about 100°C (212°F) and changes to a combination of mixed layer clay and illite. (b) Heating acanthite below 177°C (350°F temperature of inversion to the high-temperature form, argentite) produces twinning, which is not, therefore, evidence for inversion from an original high temperature form. (c) Mineral assemblages in the Green River evaporites, from invariant points and such in equilibrium diagrams, seem to have formed in the range of about 20–60°C (68–140°F). (d) Molar sodium (Na), potassium (K), and calcium (Ca) in natural waters indicate last equilibrium with rocks at 4–340°C (39–640°F). (e) Concentrations of noble gases in ground water indicate temperatures of equilibrium with the atmosphere when the water was on the surface; indicated temperatures are up to 63°C (145°F). Ratios of pairs of the gases [especially xenon/neon (Xe/Ne) and krypton/neon (Kr/Ne)] appear to give better results than concentrations of individual gases, of which Xe and Kr appear to be the best. (f) The glaucophane-lawsonite association indicates temperatures approximately 200°C (390°F) and pressures of approximately 8 kilobars (800 MPa).

2. *Intermediate temperatures*. The temperature of inversion of low quartz increases 3.6°C for each part per million (ppm) of Al accepted into the lattice.

Manganese:iron (Mn:Fe) ratios in wolframite indicate that higher Mn varieties were formed at higher temperatures and nearer the source of the mineralizing solutions. Magnesium (Mg) and Fe in coexisting garnet and biotite indicate temperatures of formation of 300–610°C (570–1130°F) in a series of rocks studied. Mg partition between staurolite and garnet in a series of metamorphic rocks suggests temperature of formation to be 515–570°C (960–1060°F). Compositions of coexisting muscovite and paragonite, interpreted from an experimental determination of the muscovite-paragonite solvus, suggest a temperature of formation of 550–570°C (1020–1060°F).

3. *High temperatures*. Numerous studies of mineral assemblages in kimberlites, lherzolites (diallage-bronzite peridotite), and other mantle-derived rocks and their included nodules have allowed workers to assign values to the temperatures and pressures (and therefore depths) of formation of these interesting assemblages.

(a) Among those that have been used are the following: Garnet-spinel (for example, 22 kilobars or 2.2 GPa at 1240°C or 2260°F). High chromium (Cr) spinels in kimberlite at 40–55 kilobars (4.0–5.5 GPa) and 950–1050°C (1740–1920°F). Volcanic spinels at 5–10 kilobars (0.5–1.0 GPa) and 800–900°C (1470–1650°F). Low-calcium clinopyroxenes (cpx) and calcic orthopyroxenes (opx) equilibrated at $T >$ 1300°C. Cr-poor cpx and opx; T = 1025–1310°C (1880–2390°F). Aluminum oxide (Al_2O_3) in equilibrium assemblage of cpx-pl-opx [for example, 9 kilobars or 0.9 GPa and 900°C (1650°F); pl = plagioclase]. Nickel (Ni) partitioning between ol, opx, and cpx in basalt; T (crystallization) = 1040°C (1900°F). Fe and Mg in garnet-cpx pairs at 10–20 kilobars (1.0–2.0 GPa); T = 700–1000°C (1290–1830°F). Garpyol at 37–43 kilobars (3.7–4.3 GPa; py = pyroxene); T = 930–1230°C (1700–2250°F). Assemblage ol-cpx-opx at 30 kilobars (3.0 GPa); T = 1575°C (2867°F). Assemblage opx-cpx in xenoliths in kimberlite at 32 kilobars (3.2 GPa), T = 920°C (1690°F); at 47 kilobars (4.7 GPa), T = 1315°C (2399°F). Fe and Mg in coexisting olivine pyroxene, up to 1440°C (2550°F).

(b) Feldspars have been much used in geologic thermometry. For example, the amount of the albite molecule in plagioclase and coexisting alkali feldspar has indicated temperature of extrusion of a quartz trachyte to be about 1000°C (1830°F); crystallization of quartz syenite, 100–700°C (1830–1290°F); and late granite crystallization, 700–650°C (1290–1200°F). Composition of plagioclase and coexisting magma (glass) gives temperatures that are close to those indicated by TiO_2 in coexisting iron oxide minerals, as does the two-feldspar geothermometer.

(c) Rare-earth-element partitioning among hydrous magma, amphibole, garnet, cpx, and opx has been used to estimate temperature of crystallization of a peridotite; also, the partitioning of the rare-earth elements and major elements in cpx has been used to estimate the temperature of crystallization of a spinel peridotite.

(d) Distribution of an ion between two inequivalent cation lattices can be used to indicate temperature of formation, if pressure can be estimated independently. With three mineral lattices, both temperature and pressure can be estimated. Epidote, binary pyroxene and amphibole solid solutions, aluminum silicate (Al_2SiO_5) polymorphs, and feldspars are among the minerals that should be most useful with this technique.

(e) Distribution of magnesium oxide (MgO) between phenocrysts of pyroxene and olivine in basalt, and glassy inclusions in the phenocrysts, has been used to estimate temperature of formation of the phenocrysts. Better results can be obtained if there are glassy inclusions in phenocrysts of feldspar that formed at the same time as the others.

Inclusions. When one mineral is included in another during crystallization and the association is cooled to room temperature, stresses develop in the mineral grains. By determining conditions under which relative compression between host and inclusion is eliminated, conditions of formation can be estimated.

Electrochemical methods. A cell is made using pairs of minerals that apparently formed in equilibrium and the temperature is varied until electromotive force (emf) = 0. This method has the advantage that it is not necessary to determine compositions or cell dimensions or to do other tedious experiments. *See* ELECTROMOTIVE FORCE (CELLS).

Fossil assemblages. By determining the temperatures of the water in which certain types of organisms grow, it is possible to infer the temperature of the water at the time that strata containing fossils of the same species, or perhaps closely similar species, were laid down. This method has been confined to differentiation between cold- and warm-water assemblages, but as more information is obtained about temperatures at which certain species lived, it should be possible to assign temperature ranges to the water in which the formations containing them were laid down.

Properties dissipated by heating. Properties such as thermoluminescence, radiation colors, and metamictization, which are exhibited by many minerals (and thermoluminescence by some rocks), are dissipated by heating. Most thermoluminescence is dissipated below 250°C (480°F), although a few materials have been reported which retain some capacity for thermoluminescence up to about 500°C (930°F). Likewise, most radiation colors in minerals are dissipated below 300°C (570°F), but in a few they persist up to 500°C (930°F) or higher.

Metamictization is the destruction of the regular internal structure of a mineral produced by emanations from contained radioactive elements. The metamictization of minerals can be dissipated and the original structure restored by heating them to about 450–900°C (840–1650°F), depending upon the mineral and rate and time of heating. *See* METAMICT STATE.

Fission tracks are another feature that can be annealed out by heating. This is important because of the effect it can have on fission track dating. For example, the fission track age of a rock from diopside

and zircon was 4.5×10^8 years, but from sphene, hornblende, and apatite, it was 7×10^6 years. Original solidification and cooling took place at 4.5×10^8 years. There was a later reheating, 7×10^6 years ago, to a temperature greater than 600°C (1110°F) but less than 700°C (1290°F) which erased tracks in sphene (600°C or 1110°F) and hornblende and apatite (450°C or 840°F) but not in diopside (825°C or 1520°F) or zircon (700°C or 1290°F). Some minerals and temperatures at which tracks are erased are as follows: silica glass (650°C or 1200°F), tektites and pigeonite (525°C or 980°F), enstatite and hypersthene (475°C or 890°F), muscovite (450°C or 840°F), and phlogopite (375°C or 710°F). *See* FISSION TRACK DATING.

Possession of these properties by minerals does not mean that they were formed at temperatures lower than those at which the properties are dissipated, but that they have not been heated to higher temperatures since the properties were acquired. In deducing the thermal history of such materials, therefore, the problem of when the property was acquired becomes important.

Crystallography. The generalization has been made that crystals grown at relatively low temperatures are likely to be simple in habit; those grown at high temperatures, complex. The composition and pH of the solution, presence of impurities, rate of growth, and other factors can also affect crystal habit.

Potassium feldspar is a good example of the change of crystal habit with temperature of formation. Phenocrysts of potassium feldspar in porphyrites (800°C or 1470°F±) are dominated by base, clinopinacoid, and orthodomes, giving crystals elongate parallel to the a axis. Crystals in pegmatites (500°C or 930°F±) are elongate parallel to the c axis and are dominated by (010), (001), and (110). In high-temperature veins such as those of the Alps (350°C or 660°F±), the (110) form becomes more prominent and the crystals are simpler. In very low-temperature veins the potassium feldspar is adularia (150°C or 300°F±), and only the two forms (110) and (101) remain, so the crystals are as simple as possible. Other examples are quartz, calcite, and fluorite. *See* CRYSTAL STRUCTURE.

Liquid inclusions. Minerals crystallizing from aqueous solutions commonly have imperfections that retain samples of the solution as liquid inclusions in the final crystals. When the crystal cools, the solution contracts, and a vapor bubble appears in each liquid inclusion. By heating plates of the mineral on a heating stage on a microscope and determining the temperature at which the solution just fills the cavities, it is possible to estimate temperature of formation if the pressure at formation was essentially the same as the vapor pressure of the solution. For significantly higher pressures, it is necessary to estimate what the pressure was and apply a correction.

The following necessary assumptions appear to be justified in most, if not all, cases. (1) The inclusion cavities were just filled with fluid under the temperature and pressure prevailing during crystallization. (2) Change of volume of the mineral itself is not significant. (3) Changes in volume and concentration brought about by deposition of material during cooling are such as not to affect the result. (4) Primary and secondary liquid inclusions can be distinguished under the microscope. (5) There has been no leakage from or into the inclusions. (6) The liquid is an aqueous solution containing no carbon dioxide or other gas in large concentration. (7) Pressure-volume-temperature relations are near enough to those of pure water or chloride solutions that have been studied so that no serious errors are introduced by using the available data.

Temperature ranges that have been estimated by this method for some common vein and pegmatite minerals are as follows: calcite, 40–362°C (100–684°F); sphalerite, 75–275°C (170–527°F); fluorite, 83–350°C (181–660°F); vein quartz, 100–440°C (212–824°F); pegmatite quartz, 200–530°C (390–990°F); and topaz, 275–500°C (527–930°F).

Behavior of organic material. Organic compounds are generally more sensitive to thermal changes than are inorganic ones. Therefore, when they are heated in nature, even to moderate temperatures, characteristic changes take place that can be used to estimate the temperatures to which various kinds of organic matter have been exposed. Persistence of organic compounds indicates that the temperature of the environment to which they have been exposed has not been above the temperature at which they decompose at a rate significant with respect to duration of their burial.

Degree of condensation, loss of volatiles, thermal stability of individual compounds, and other factors have been used as indications of temperatures reached. For example, temperatures of condensation to asphalt, bernalite, carbonite, and so on have been estimated to range from about 30 to 370°C (90–700°F). In general, mostly water is lost from coal up to about 100°C (212°F); CO_2 and hydrogen sulfide (H_2S) are given off to about 350°C (660°F); methane (CH_4), ammonia (NH_3), and some hydrogen (H_2) to 800°C (1470°F); above 800°C (1470°F), mostly hydrogen. "Degree of coalification" (concomitant with loss of volatiles) and reflectance of vitrinite have also been used to estimate temperatures of formation of coal. *See* COAL.

Colors developed in conodonts indicate the degree of heating that the sediments containing them have undergone, ranging from pale yellow through brown to black with heating to 350°C± (660°F±). Higher temperatures dissipate the dark colors, and a temperature of 950°C (1740°F) for several hours makes the conodonts almost clear.

Logs on the Colorado Plateau have been coalified at about 125°C (247°F). In coal formed at 4000 m (2.5 mi) depth (≃165°C or 329°F), vitrinite reflectance approaches 4%, and the coal retains about 4% of volatiles. At 5% reflectance, volatile content approaches zero.

Carboxyl groups of porphyrins remain in petroleum at temperatures lower than 200°C (390°F). Amino acids can be used in the same way in some fossils. Melting points of bitumens indicate maximum

temperatures for their formation. Fractionation series near dikes and veins give thermal gradients on a relative scale. Chemical processes such as hydrogenation, elimination of oxygen or nitrogen, carboxylation, and so forth, indicate exposure to increased temperature.

Most of the processes affecting organic materials are time-dependent as well as temperature-dependent, and some of them are pressure-dependent as well. Therefore, although relative temperatures can usually be indicated, it may be difficult or impossible to suggest absolute temperatures unless time or pressure or formation, or both, can be approximated. Earl Ingerson

Isotopic Methods

The stable isotopes of many light elements, including H, C, O, and S, are not uniformly distributed in natural substances. The $^{18}O/^{16}O$ ratio, for example, ranges from 0.00188 in Antarctic ice, through 0.00200 in ocean water, to 0.00206 in limestone. Relative to ocean water, the reference standard for oxygen, the limestone is 3% or 30‰ enriched in the heavy isotope. Such variations are caused by both equilibrium and kinetic processes. Equilibrium fractionations are temperature-dependent and can therefore be used as thermometers.

H. Urey suggested in 1947 that the partitioning of oxygen isotopes between calcite and water could be used to determine the temperature at which a marine carbonate fossil was precipitated in the geologic past. Largely as a result of this suggestion, several hundred laboratories throughout the world are now equipped to undertake research in stable isotope geochemistry. Spectacular achievements of this research have been made in the field of paleoclimatology, and understanding of igneous, metamorphic, sedimentary, and hydrothermal processes has also advanced enormously.

Calibration of isotopic thermometers. Isotopic fractionations among phases at equilibrium arise from the influence of isotopic mass on the vibrational frequencies of solids and on the vibrational-rotational frequencies of molecules. Fractionation of the isotopes among molecular species can be calculated from spectroscopic data, using methods of statistical mechanics, but the effects for crystals and liquids are difficult to calculate accurately. Experimentally determined fractionations, generally considered more reliable than those calculated from spectroscopic data, are obtained by analyzing the isotopic compositions of coexisting phases that have been thoroughly equilibrated at known temperatures in the laboratory. After this has been accomplished, the isotopic thermometer is said to be calibrated (**Figs. 1** and **2**). *See* STATISTICAL MECHANICS.

In most cases, isotopic fractionation between phases increases with decreasing temperature, and statistical mechanics predicts zero fractionation at infinite temperature. Pressure does not significantly affect the equilibrium distribution of isotopes among most substances, because the net volumetric effects of isotopic substitutions are typically negligible. Researchers have noticed that high pressures increase the rate of isotopic exchange in hydrothermal experiments and thereby facilitate the attainment of equilibrium between phases.

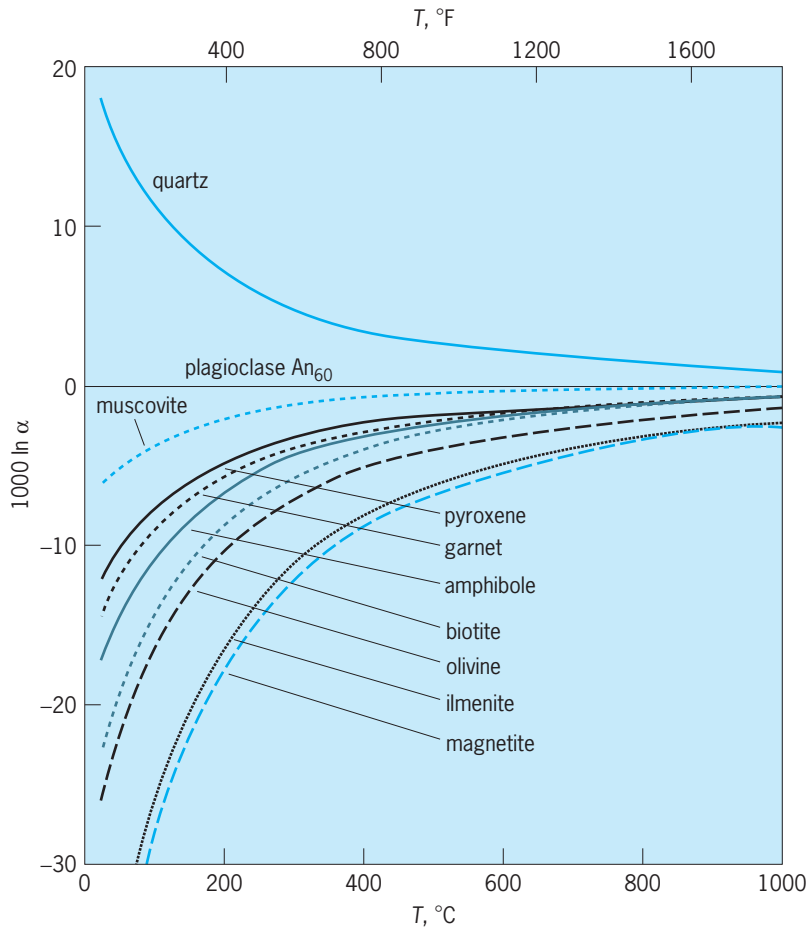

Fig. 1. Equilibrium oxygen isotope fractionations between several important rock-forming minerals and plagioclase feldspar (An_{60}). Minerals with the highest values have the strongest tendency to concentrate ^{18}O. The fractionations between different minerals are largest at low temperatures. (*Modified from I. Friedman and J. R. O'Neil, Compilation of stable isotope fractionation factors of geochemical interest, U.S. Geol. Surv. Prof. Pap. 440-KK, 1977*)

The requirements for a reliable temperature determination are that isotopic equilibrium was originally attained in the natural system at the time of crystallization or formation; that a suitable isotopic thermometer has been calibrated under known laboratory conditions; and that the original isotopic composition of the natural material has been preserved throughout its subsequent history. Unfortunately, the more readily equilibrium is attained in the laboratory, the less likely it is to be "frozen in" and subsequently preserved in natural materials. In fact, naturally occurring mineral assemblages probably could not retain their initial isotopic compositions over geologic time, were it not for their tendency to be coarser-grained and to be more dehydrated than samples equilibrated in hydrothermal laboratory experiments.

Terminology. The δ value is defined as the per mille (‰) difference in isotopic ratios between a sample

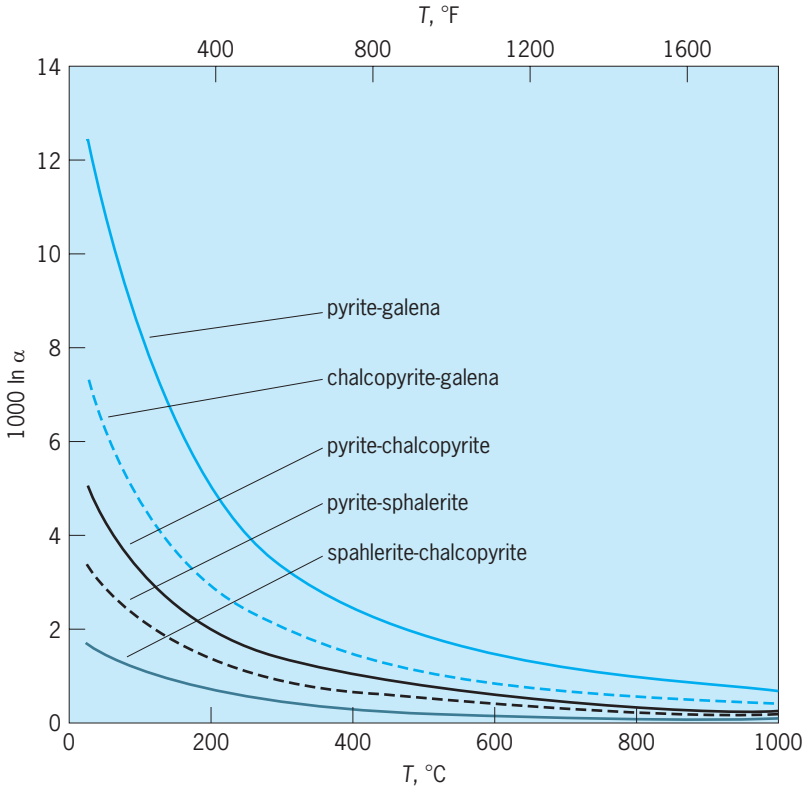

Fig. 2. Equilibrium sulfur isotope fractionations between different sulfide minerals. (*Modified from I. Friedman and J. R. O'Neil, Compilation of stable isotope fractionation factors of geochemical interest, U.S. Geol. Surv Prof. Pap. 440-KK, 1977*)

and a standard as shown in Eq. (1), where R_A is

$$\delta_A = \left(\frac{R_A - R_{\text{std}}}{R_{\text{std}}}\right) \times 10^3 \qquad (1)$$

the ratio of deuterium to hydrogen (D/H), carbon-13 to carbon-12 ($^{13}C/^{12}C$), oxygen-18 to oxygen-16 ($^{18}O/^{16}O$), or sulfur-34 to sulfur-32 ($^{34}S/^{32}S$) in substance A. The standard for hydrogen and oxygen isotopes is standard mean ocean water (SMOW). Calcite from a fossil belemnite known as PDB is used as a standard for both oxygen and carbon isotopes. Troilite from the Canyon Diablo meteorite is the standard for sulfur isotopes.

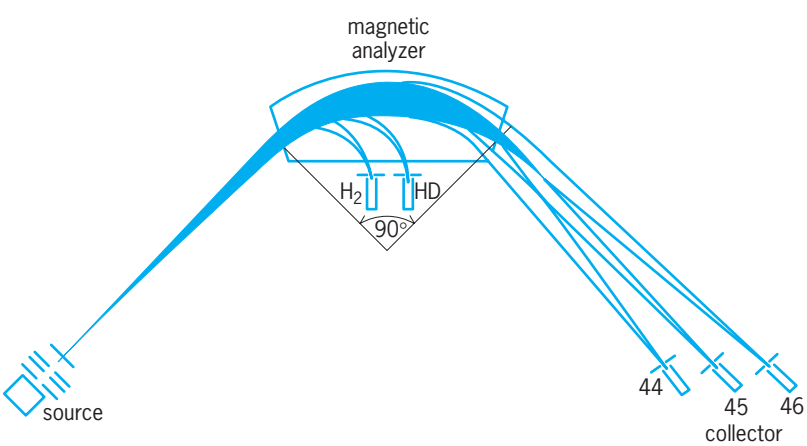

Fig. 3. A modern isotope-ratio mass spectrometer can analyze several different gases, and some can analyze samples as small as 100 micrograms or less. Modern mass spectrometers can be optimized for many different analytical purposes, but practically all rely on the traditional use of a magnetic field to effect the separation of ions having different charge to mass ratio. This MAT 252 spectrometer is set up to ionize and accelerate either hydrogen gas or carbon dioxide gas, separate the gas into distinct beams representing the different isotopic species (H_2 or HD; or masses 44, 45 and 46 for CO_2), and then precisely compare the intensities of the different mass beams. (*After R. E. Criss, Principles of Stable Isotope Distribution, Oxford University Press, 1999*)

The isotopic fractionation factor between substances A and B is defined as $\alpha_{AB} = R_A/R_B$, a quantity that is related to the equilibrium constant of an isotopic exchange reaction. An example is D/H exchange between hydrogen gas and water [reaction (2)].

$$HD + H_2O = H_2 + HDO \qquad (2)$$

The per mille fractionation, $10^3 \ln \alpha_{AB}$, is approximately equal to $\delta_A - \delta_B$ or Δ_{A-B}, which for reaction (2) is the difference between the δD values of the water and the hydrogen gas.

Analytical techniques. The element to be isotopically analyzed is quantitatively extracted from the sample and converted to a gas such as hydrogen (H_2), carbon dioxide (CO_2), or sulfur dioxide (SO_2). The gas is then introduced into a special mass spectrometer designed to simultaneously collect two or more ion beams differing in mass per unit charge, such as $[HD]^+$ and $[H_2]^+$, or $[^{12}C^{16}O^{18}O]^+$, $[^{13}C^{16}O^{16}O]^+$, and $[^{12}C^{16}O^{16}O]^+$. Electronic comparison of the ion-beam currents yields the desired isotopic ratio to a precision of ±0.5‰ for H, or ±0.05‰ for carbon (C), oxygen (O), and sulfur (S) [**Fig. 3**].

Many techniques can be used to obtain a gas from a sample for analysis by mass spectrometry. Examples are hydrogen extracted from water by reaction with hot zinc, chromium, or uranium metal; carbon dioxide obtained from carbonates by reaction with phosphoric acid (H_3PO_4) at a fixed temperature; organic compounds combusted to CO_2; oxygen liberated from silicates and oxides by reaction with hot fluorine (F_2) or bromine pentafluoride (BrF_5); and sulfides ground with copper(I) oxide (Cu_2O) and heated to produce SO_2. More recent developments include the ability to ablate and analyze microscopic samples using laser or ion microprobes, the analysis of tiny gas samples that are introduced as pulses into a stream of helium that continuously flows into the mass spectrometer, or using mass spectrometers that are directly interfaced with gas chromatographs or other analytical equipment to analyze isotopes of various compounds in complex mixtures. *See* MASS SPECTROMETRY.

Igneous rocks. Partitioning of oxygen isotopes among minerals at igneous temperatures can be studied (**Fig. 4**). The tendency of many minerals to concentrate the heavy isotope can be represented by B, the coefficient in Eq. (3), where α_{MP} is the equilibrium fractionation factor between the mineral and plagioclase (An_{60}), and T is absolute temperature.

$$1000 \ln \alpha_{MP} = \frac{B \times 10^6}{T^2} \qquad (3)$$

The temperatures at which igneous rocks crystallize can be estimated from experimental petrology

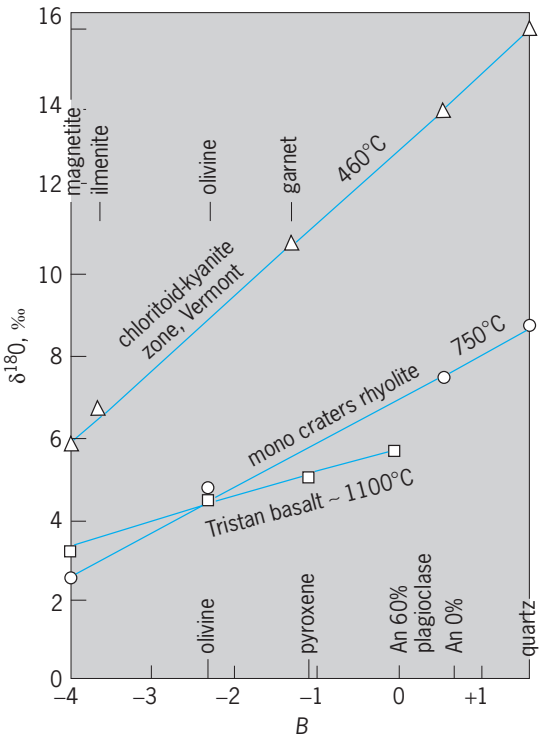

Fig. 4. Oxygen isotope compositions of minerals in a basalt, a rhyolite, and a metamorphic rock. The horizontal scale represents the tendency of each mineral to concentrate ^{18}O relative to plagioclase. The slopes decrease with increasing temperature of formation. °F = (°C × 1.8) + 32. (*After R. W. Fairbridge, ed., Encyclopedia of Geochemical and Environmental Sciences, Van Nostrand Reinhold, 1972*)

and direct observation, so isotopic thermometry is not especially valuable for this particular application. Crystal-melt fractionations are generally small and produce only small isotopic variations during igneous differentiation, unless crystallization is accompanied by assimilation of the wall rock.

Subsolidus exchange. Surprisingly, numerous departures from oxygen isotopic equilibrium have been noted in igneous rocks, and these provide far more interesting information about their genesis than does the isotopic thermometry of their crystallization. Plutonic rocks commonly show evidence of subsolidus isotopic exchange between coexisting minerals, produced during slow cooling. A good example is presented by the plagioclase-ilmenite fractionations of about 5‰ in rocks from the Labrieville anorthosite, indicating a temperature of 600°C (1110°F) that is much lower than the crystallization temperature of 1100°C (2010°F). In plutonic rocks with many minerals, slow cooling will typically result in disequilibrium fractionations that would indicate inconsistent isotopic "temperatures" for various mineral pairs, which are all significantly lower than the crystallization temperature.

Rocks that have been altered with infiltrating fluids typically display even more profound departures from isotopic equilibrium. In contrast with slow cooling, which merely causes the redistribution of oxygen isotopes within the rock, fluid infiltration conveys exotic oxygen into the rock and thereby can change the overall heavy isotope concentration. Thus, slow cooling can produce large, disequilibrium, fractionation factors as some minerals gain oxygen-18 at the expense of other minerals that lose oxygen-18, but fluid infiltration can cause all the minerals in the rock to become either richer or poorer in oxygen-18, mostly depending on the composition of the fluid. Proof that such fluid–rock interactions have occurred is therefore recorded in steep, positive slopes on plots where the delta values of one mineral in a rock are plotted directly against those of another mineral (**Fig. 5**).

Isotopic studies of several igneous provinces have revealed widespread depletion of ^{18}O and deuterium (D) in the upper crust due to interaction with circulating geothermal waters of meteoric origin. Such investigations have shown that the order of increasing resistance to hydrothermal ^{18}O exchange is feldspar-biotite-pyroxene-magnetite-quartz. This hierarchy helps explain the steep slopes of the altered rock suites in Fig. 5, and indicates that the frequently used quartz-magnetite isotopic thermometer is relatively resistant to hydrothermal alteration. *See* IGNEOUS ROCKS.

Large oxygen isotope variations discovered in mantle-derived eclogitic xenoliths in kimberlite pipes were originally attributed to crystal-melt fractionation at high pressure. However, it is now realized that the eclogites represent ancient oceanic crust that experienced isotopic exchange with convecting seawater at a spreading ridge. Oceanic crust has thus descended to depths of over 60 mi (100 km) and eventually has returned to the surface as fragments in volcanic eruptions. Some of the associated diamonds may represent organic carbon. *See* ECLOGITE.

Metamorphic rocks. Regularities in the oxygen isotope fractionations among coexisting quartz, plagioclase, pyroxene, garnet, rutile, ilmenite, and magnetite in regionally metamorphosed rocks suggest that these minerals commonly reached equilibrium at the maximum metamorphic temperatures attained, and preserved their isotopic compositions during subsequent retrograde cooling. Isotopic analyses of these minerals have yielded temperatures of metamorphism that are consistent with petrologic estimates. However, confidence in isotopic geothermometry is justified only when several mineral pairs provide concordant temperatures.

Isotopic thermometry indicates that lizardite-rich serpentinites formed at temperatures below 100°C (212°F), antigorite serpentinites at 200–400°C (370°–1020°F), glaucophane schists at 200–550°C (390–1020°F), garnet-grade schists at 450–500°C (390–750°F), and sillimanite schists and eclogites at 500–600°C (930–1110°F).

In addition, carbon and sulfur isotope fractionations have been applied to metamorphic thermometry. The concentration of ^{13}C into calcite relative to graphite is useful in marbles; the partitioning of ^{34}S among sulfide minerals is useful in metamorphosed ore deposits. *See* METAMORPHIC ROCKS.

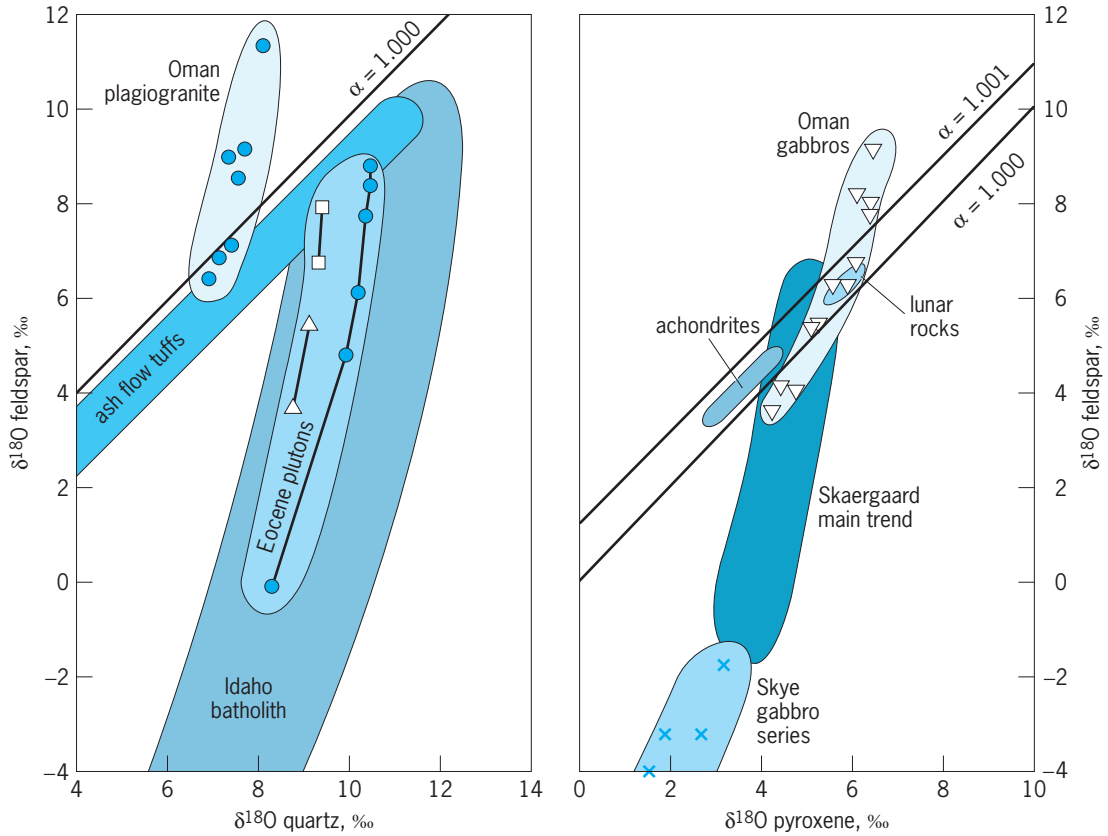

Fig. 5. Oxygen isotope compositions of coexisting quartz and feldspar, or pyroxene and plagioclase, in several igneous rock suites. Lunar rocks, achondritic meteorites, and many ash-flow tuffs cooled quickly and preserve small fractionations indicative of their magmatic crystallization temperatures; these plot along high-temperature isotherms which have unit slopes and small *y*-intercepts on this diagram. In contrast, many plutonic rock suites such as the Idaho batholith and the Skaergaard intrusion have been altered by fluids; such sample suites define steep positive slopes with anomalously large, variable, disequilibrium fractionation factors. The samples from Idaho, Skaergaard, and Syke interacted with meteoric waters which are low in ^{18}O, whereas those from Oman were infiltrated by sea water which is relatively high in ^{18}O. (*Modified from R. T. Gregory, R. E. Criss, and H. P. Taylor, Jr., Oxygen isotope exchange kinetics of mineral pairs in closed and open systems: Applications to problems of hydrothermal alteration of igneous rocks and Precambrian iron formations, Chem. Geol., 75:1–42, 1989*)

Geothermal waters. Oxygen isotope fractionations between water and dissolved sulfate from deep boreholes have yielded isotopic temperatures in reasonable agreement with the maximum measured temperatures in the geothermal reservoirs at Wairakei, New Zealand; Otake, Japan; and Larderello, Italy. The rate of isotope exchange between water and dissolved sulfate is sufficiently rapid to assure isotopic equilibrium in all reservoirs of significant size having temperatures above 140°C (280°F).

The rate of isotopic exchange is sufficiently slow, however, that no significant reequilibration is expected when geothermal waters ascend and escape to the surface via natural hot springs of significant flow rate. Unfortunately, the isotopic compositions of spring waters can be altered by boiling or dilution with near-surface waters. Surface meteoric waters usually contain less ^{18}O because they have not undergone the isotopic exchange with hot rocks experienced by deep geothermal waters. Boiling tends to increase both ^{18}O and D in the residual waters. Furthermore, oxidation of hydrogen sulfide can contribute sulfate to spring waters and thereby alter their sulfate compositions. Despite these difficulties, subsurface reservoir temperatures can be estimated from spring water isotopic analyses if appropriate corrections are made for boiling, dilution, and oxidation as indicated by chemical analyses. In this manner, the temperature of the deep reservoir at Yellowstone National Park (Wyoming, Montana, Idaho) has been estimated to be 360°C (680°F).

An additional isotopic thermometer that has been applied to geothermal systems is the partitioning of ^{13}C between carbon dioxide and methane.

Hydrothermal ore deposits. Although the majority of stable isotope studies of ore deposits are motivated by a desire to elucidate the origins of hydrothermal fluids, isotopic temperatures have also been obtained for several deposits. Sulfur isotope fractionations between coprecipitated sphalerite and galena in Providencia, Mexico, have yielded temperatures that are usually within 20°C (68°F) of the filling temperatures of associated fluid inclusions. A similar investigation of the Darwin lead-zinc replacement ores in California yielded reasonable isotopic temperatures ranging from 380°C (720°F) to 270°C (520°F). However, the wide range of sulfur isotope fractionations observed in low-temperature Mississippi Valley–type

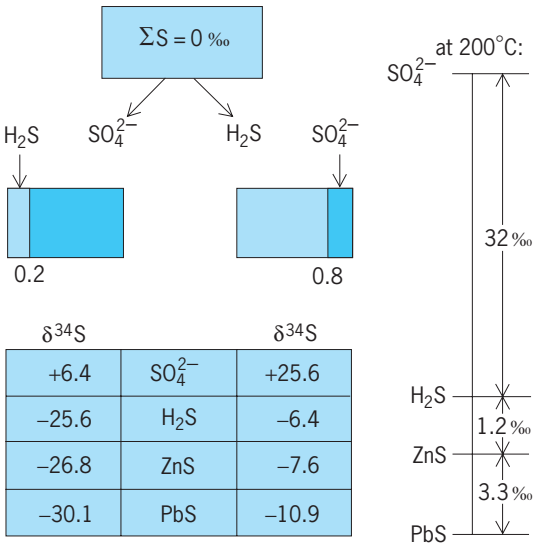

Fig. 6. Variation of $\delta^{34}S$ of sulfate, H_2S, and sulfide minerals with variation in H_2S/SO_4^{2-} of a hydrothermal solution at $T = 200°C$ (392°F), and $\delta^{34}S$ total = 0‰. (*After R. O. Rye and H. Ohmoto, Sulfur and carbon isotopes and ore genesis: A review, Econ. Geol., 69:826–842, 1974*)

deposits indicates that isotopic equilibrium may not occur in these deposits.

The effects of variations in H_2S/SO_4^{2-} ratio (or oxygen fugacity) on the isotopic compositions of hydrothermal minerals can be modeled (**Fig. 6**). It is clear that the oxidation of H_2S to SO_4^{2-} in a hydrothermal solution will have a profound influence on the sulfur isotope compositions of minerals forming from it. The consequent large isotopic variation observed in some deposits complicates the application of isotopic thermometry. *See* FUGACITY.

Sulfate-sulfide isotope fractionations in the metamorphosed Balmat-Edwards lead-zinc district, New York, indicate temperatures of about 620°C (1150°F) near the peak temperature of metamorphism. However, pyrite-sphalerite and sphalerite-galena fractionations indicate temperatures of 390°C (730°F) and 325°C (617°F), respectively. These are probably spurious temperatures caused by isotopic exchange during retrograde cooling.

Exceptionally large oxygen isotope differences have been observed between calcite and uraninite in hydrothermal veins. The δ values of calcite and uraninite in the Martin Lake mine, Saskatchewan, are +16.8 and −30‰, respectively. It is not clear if these values reflect a very large equilibrium fractionation, or deposition of the uraninite from isotopically light meteoric waters. Paleoclimate information has been gleaned from the isotopic composition of meteoric water trapped as fluid inclusions in epithermal fluorite from Idaho. *See* ORE AND MINERAL DEPOSITS.

Paleoclimates. The development of isotopic methods for determining the ancient temperatures at which fossil shells were grown has vastly increased understanding of the thermal history of the Earth's surface. The possibility of using that knowledge to predict the future course of climate is stimulating much research.

The most widely used isotopic thermometer is the fractionation of oxygen isotopes between the calcite shells of microscopic foraminifera and ocean water (**Fig. 7**). Isotopic variations exist for fossil planktonic foraminifera from Caribbean sediments during the past 730,000 years (**Fig. 8**). Seasonal variations are averaged out in such data because tens of individual shells are picked for each analysis from a section of sediment core representing hundreds of years of accumulation. If the isotopic composition of seawater were constant through time, the $^{18}O/^{16}O$ ratio of a fossil shell grown at equilibrium would depend only on its mean temperature of deposition. This is not the case because periodic accumulations of ^{18}O-depleted continental ice during the Pleistocene caused corresponding enrichments of ^{18}O within the oceans, thus enhancing the isotopic contrast between glacial and interglacial biogenic

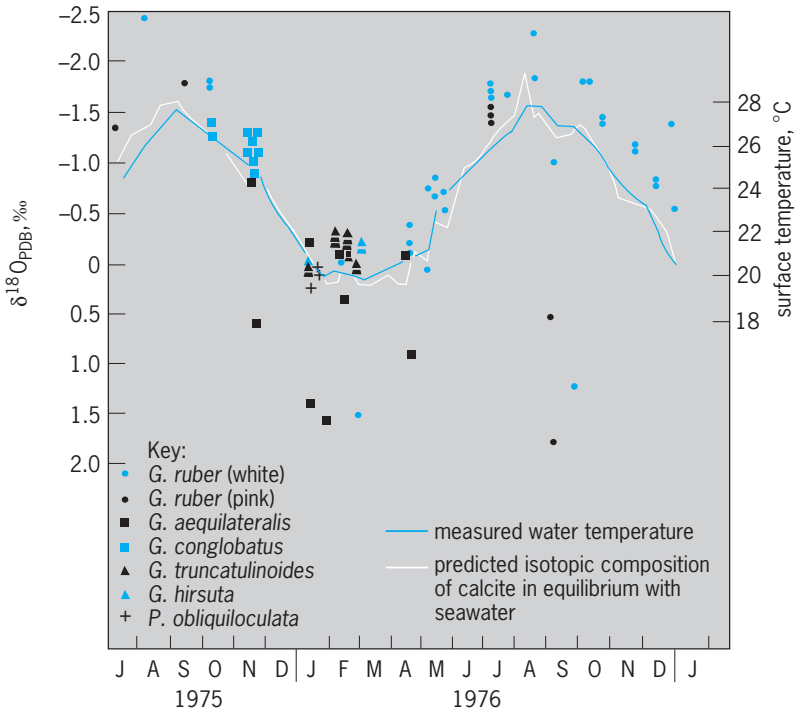

Fig. 7. Seasonal variation of $\delta^{18}O$ of living planktonic foraminifera (genera *Globigerina* and *Pachiderma*) off Bermuda. °F = (°C × 1.8) + 32. (*After D. F. Williams, A. W. H. Be, and R. G. Fairbanks, Seasonal oxygen isotopic variations in living planktonic foraminifera off Bermuda, Science, 206:447–449, 1979*)

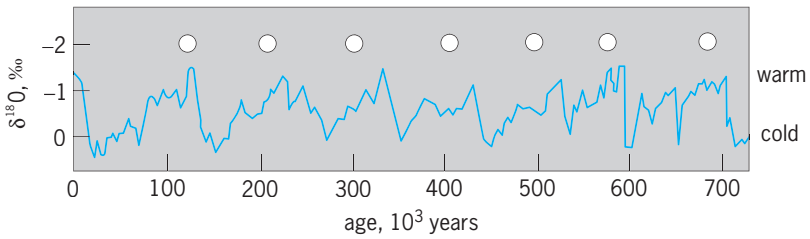

Fig. 8. Oxygen isotope variations in one species of planktonic foraminifera from Caribbean sediments during the past 730,000 years. The reference standard is PDB. The circles indicate times of maximum orbital eccentricity when northern summers occurred at perihelion and were relatively warm. (*After C. Emiliani, The cause of the ice ages, Earth Planet Sci. Lett., 37:349–352, 1978*)

calcium carbonate. This ice-volume influence during the past few glacial cycles can be estimated from analyses of benthic foraminifera that grew under consistently frigid temperatures. Seawater variation during the Pleistocene accounts for most of the 1.8‰ fluctuation in isotopic composition observed in planktonic foraminifera from tropical ocean sediments.

Isotopic analyses of benthic foraminifera from pre-Pleistocene sediments reveal variations in excess of possible ice-volume effects. Furthermore, the contrast between benthic and planktonic analyses is less in pre-Pleistocene sediments, presumably because deep water was warmer in preglacial times. Deep-sea temperatures were about 15°C (27°F) warmer 50 million years ago. Rapid, 1-part-per-million increases in the ^{18}O compositions of benthic foraminifera 35 and 15 million years ago probably indicate the initiation (temperature effect) and growth (ice-volume effect) of the Antarctic ice sheet, respectively. *See* CENOZOIC; PLEISTOCENE.

The recorded climatic cycles (Fig. 8) have led to several explanatory hypotheses. Statistical analysis has shown that a part of the variance in the climatic record is concentrated in three spectral peaks with periods of 100,000, 41,000, and 23,000 years. These closely match the periods of the Earth's orbital eccentricity, obliquity, and precession. This evidence supports the Milankovitch theory, which predicts extensive glaciation in the Northern Hemisphere a few thousand years in the future.

Orbital influences on solar radiation may not be the only cause of Pleistocene climate swings. For example, fluctuations in the carbon dioxide concentration in the atmosphere may also affect climate. Available data show an excellent correlation between the carbon dioxide trapped in Antarctic ice and the temperature variations derived from hydrogen isotope variations in the ice (**Fig. 9**). Heavy isotopes of hydrogen and oxygen are highly depleted in Antarctic snow because they preferentially "rain out" of air before it reaches the Antarctic interior. In addition, the abundances of these isotopes in precipitation correlate with air temperature. The mechanisms controlling atmospheric carbon dioxide concentration are complex and include ocean–atmosphere exchange, ocean mixing, biological productivity, volcanism, weathering, and fossil-fuel combustion.

In order to extend isotopic thermometry to Precambrian oceans, it is necessary to analyze materials that are less amenable than calcite to postdepositional isotopic exchange with circulating waters. Chert is considered the most suitable candidate for this role. The cherts that exhibit the highest $\delta^{18}O$ values within each age group are considered to be the most pristine; the other cherts have probably experienced isotopic exchange with adjacent rocks, with waters at elevated temperatures, or with isotopically light meteoric waters. The involvement of meteoric waters can be ascertained by means of D/H analyses of hydroxyl groups that are incorporated into the cherts.

There are three competing explanations of the

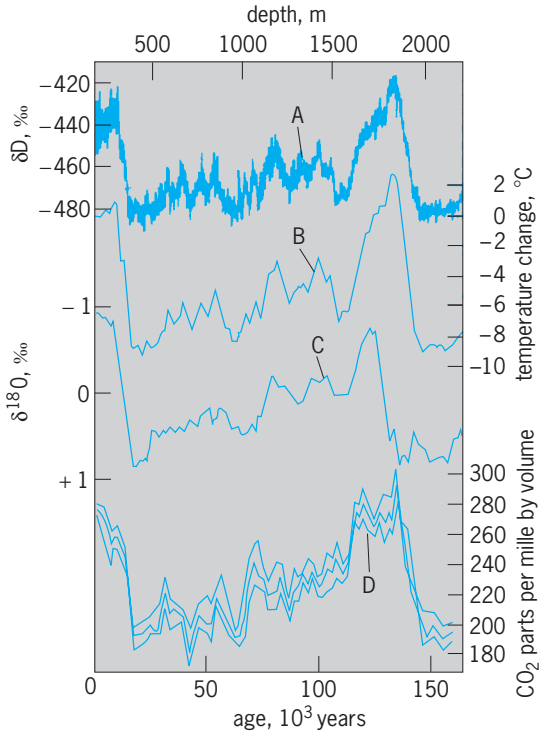

Fig. 9. Graphs showing correlation between carbon dioxide trapped in ice and inferred temperature variations. Curve A: hydrogen isotope profile in Vostok core, Antarctica [deuterium (D) per mille versus standard mean ocean water (SMOW)]. Curve B: temperature profile (relative to modern temperature) derived from curve A after correcting for isotopic changes in seawater derived from curve C. Curve C: oxygen (O) isotope changes in seawater derived from analyses for foraminifera. Curve D: carbon dioxide (CO_2) content of air trapped in ice (with uncertainty bands). (*After C. Lorius et al., Antarctic ice core: CO_2 and climatic change over the last climatic cycle, EOS, 69:681–684, 1988*)

trend toward higher $\delta^{18}O$ values with decreasing age. The first assumes fairly uniform temperatures and a progressive enrichment of ^{18}O in ocean water over geologic time. The second assumes that the ^{18}O content of ocean water is nearly constant over time because it is buffered by seawater–basalt interactions at 200–300°C (390–570°F) along axes of sea-floor spreading. The chert trend can then be interpreted in terms of progressive cooling of the Earth's surface temperature over time, with the climate being a steamy 70°C (160°F) about 3.4 billion years ago. The third possibility is that the ancient chert has not preserved its original isotopic composition. *See* CHERT; PALEOCLIMATOLOGY. G. Donald Garlick; Robert E. Criss

Bibliography. A. T. Anderson, Mineralogy of the Labrieville anorthosite, Quebec, *Amer. Mineral.*, 51:1671–1711, 1966; R. E. Criss, *Principles of Stable Isotope Distribution*, Oxford University Press, 1999; C. Emiliani, The cause of the ice ages, *Earth Planet. Sci. Lett.*, 37:349–352, 1978; R. W. Fairbridge (ed.), *Encyclopedia of Geochemical and Environmental Sciences*, Van Nostrand Reinhold, 1972; I. Friedman and J. R. O'Neil, *Compilation of Stable Isotope Fractionation Factors of Geochemical Interest*, U.S. Geol. Surv. Prof. Pap. 440-KK, 1977; R. T. Gregory, R. E. Criss, and H. P. Taylor H.P., Jr., Oxygen

isotope exchange kinetics of mineral pairs in closed and open systems: Applications to problems of hydrothermal alteration of igneous rocks and Precambrian iron formations, *Chem. Geol.*, 75:1–42, 1989; C. Lorius et al., Antarctic ice core: CO_2 and climatic change over the last climatic cycle, *EOS*, 69:681–684, 1988; H. C. Urey, The thermodynamic properties of isotopic substances, *J. Chem. Soc.*, 1947:562–581, 1947; D. F. Williams, A. W. H. Be, and R. G. Fairbanks, Seasonal oxygen isotopic variations in living planktonic foraminifera off Bermuda, *Science*, 206:447–449, 1979.

Geologic time scale

An ordered, internally consistent, internationally recognized sequence of time intervals, each distinct in its own history and record of life on Earth, including the assignment of absolute time in years to each geologic interval. The geologic time scale (see **table**) has two essential components: a relative scale, consisting of named intervals of geologic history arranged in chronologic sequence from oldest (bottom) to youngest (top); and a numerical (or absolute) time scale, providing estimates of absolute ages for the boundaries of these intervals.

Geologic time scale

EON ERA Period [system] Epoch [series]	Age at beginning of interval, in million years	Interval length, in million years
PHANEROZOIC		
CENOZOIC		65
Quaternary (Q)*		1.8
Recent	0.01	0.01
Pleistocene	1.8	1.79
Tertiary (T)	65	63.2
Pliocene (Tpl)	5.3	3.5
Miocene (Tm)	23.0	17.7
Oligocene (To)	33.9	10.9
Eocene (Te)	55.8	21.9
Paleocene (Tp)	65.5	9.7
MESOZOIC	251	185.5
Cretaceous (K)	145.5	80
Jurassic (J)	199.6	54.1
Triassic (Tr)	251	51.4
PALEOZOIC	542	291
Permian (P)	299	48
Carboniferous (M,P)	359.2	60.2
Devonian (D)	416	56.8
Silurian (S)	443.7	27.7
Ordovician (O)	488.3	44.6
Cambrian (C)	542	53.7
PRECAMBRIAN		
PROTEROZOIC	2500	1958
LATE (Z)† (Neoproterozoic)	1000	458
Ediacaran	630	88
Cryogenian	850	220
Tonian	1000	150
MIDDLE (Y) (Mesoproterozoic)	1600	600
EARLY (X) (Paleoproterozoic)	~2500	900
ARCHEAN	3600	
LATE (W)	2800	300
MIDDLE (U)	3200	400
EARLY (V)	>3600	>400

*In parentheses are the symbols for the periods and epochs used on geologic maps and figures in North America, as well as other parts of the world.
†Letter designations of Precambrian age intervals are used by the U.S. Geological Survey.

Relative time scale. The essential principles and concepts needed to develop a relative time scale for the Phanerozoic Eon (see table) had been recognized by the end of the eighteenth century. The principle of original horizontality implied that sediments, which eventually were lithified into sedimentary rocks, accumulated as essentially horizontal layers. The principle of superposition implied that in any undisturbed sequence of layered rocks the oldest layer was always on the bottom; hence, the bottom-to-top ordering of the geologic time scale. Uniformitarianism embodied the hypothesis that an understanding of modern geologic processes, such as weathering, erosion, and deposition of sediment, could be used to interpret the record of ancient processes, even though it is only partially complete. Igneous rocks, such as granites, were convincingly demonstrated to have been originally molten (magma) and were thus younger than any host rock into which the magma was injected. A conglomerate, very coarse detrital sedimentary rock, is younger than any of the pebbles or boulders contained within it. This is the principle of components. The application of these principles and concepts clearly demonstrated that the Earth had an unequivocally long history and facilitated interpretations about virtually all geologic materials formed as well as the determination of relative ages of different rocks, where continuously exposed. In addition, they were used to aid in correlating sequences of rocks by determining the comparable ages of rocks that were widely separated geographically. *See* CONGLOMERATE; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; EROSION; IGNEOUS ROCKS; WEATHERING PROCESSES.

Faunal succession. The problem of stratigraphic correlation was resolved through faunal succession, established by the British civil engineer William Smith in the early 1800s. Smith collected fossils across southern England and Wales. His observations showed that specific sequences of sedimentary rocks had a characteristic fossil content and that the fossils typical of a succession of layers were always found in a predictable order. If he found a fossiliferous layer in a new area, he could predict the fossils that would be found in younger or older layers. Using the principles of faunal succession and superposition, Smith was able to establish the sequence of sedimentary rocks over much of England and to produce the first regional geologic map, which showed the areal distribution of distinctive assemblages of rocks of different relative ages. The principle of faunal succession also permitted the correlation of British sedimentary rocks with sections in western Europe and ultimately with rocks worldwide. This was the

foundation of historical geology. *See* FOSSIL; GEOLOGY; INDEX FOSSIL.

Correlation. The recognition that the fossil record was uniquely ordered through time, and that distinctive fossil assemblages within this orderly sequence could be used to determine approximate contemporaneity (the principle of correlation) among sedimentary rocks from different areas, allowed the major components of the relative geologic time scale to be compiled. Sequences of rocks with particular faunal and lithologic characteristics were given names reflecting either distinctive features or the geographic areas where they were best expressed. Examples include the coal-bearing sequences of England (Carboniferous), the prominent chalk cliffs of Dover (Cretaceous) [*creta* is Latin for chalk], and the rocks and fossils that lay below the Carboniferous and were particularly well exposed in the Devon area of southwest England (Devonian).

Distinctive sequences of sedimentary rocks and associated faunal assemblages soon became formalized as systems, and the time interval they represented became periods. Because sequences often contained considerable thicknesses of rocks and there were sequential differences in the faunal successions within them, they were divided into smaller intervals characterized solely by their faunas. The primary divisions, often simply designated as lower, middle, and upper, were later given formal names as series. Series were further divided into stages, which were subdivided into zones.

Because the fundamental principles of historical geology developed in England and rapidly spread to western Europe, almost all names from the major rock sequences that make up the relative geologic time scale had their origin in Europe. Ultimately, the principle of faunal succession facilitated correlation of rocks of Carboniferous, Cretaceous, and Devonian ages, and for many of the series, stages, and zones within them, in many parts of the world. In many cases, different local sequences of series, stages, and zones had been developed and, at least tentatively, named. A proliferation of named stratigraphic sequences led to the need for some international standardization of nomenclature. This process started at the first International Geological Congress in the late 1800s, and the sequence of systems, although lacking approximate absolute ages for their boundaries, was established by the end of the nineteenth century. Stabilization of international nomenclature for intervals within systems, and more accurate definition of the ages of interval boundaries, is an ongoing process that is the responsibility of working groups and subcommissions of the International Commission on Stratigraphy. *See* SEDIMENTARY ROCKS; STRATIGRAPHIC NOMENCLATURE.

Absolute time scale. Absolute age determinations involve estimating the absolute age of formation of geologic materials or the absolute age of a specific process that affected the materials. A better understanding of the regional history and relative ages of igneous and metamorphic rocks, and how igneous or metamorphic events compared in time with the deposition of sedimentary sequences, was made possible with the discovery of radioactivity in the late nineteenth century and the recognition that several important rock-forming minerals contained elements with radioactive isotopes. Demonstrated in the early twentieth century, isotopic age dating of geologic materials assumes that in a chemically closed system (such as a mineral grain in an igneous or metamorphic rock that has not experienced further metamorphism, melting, chemical weathering, or mechanical fracturing after grain growth) the decay products of radioactive isotopes of elements (such as potassium, uranium, thorium, or rubidium) that are found in sufficient abundance in particular minerals will be retained in the mineral grain below a specific, narrow range of isotopic blocking temperatures. The isotopic blocking temperature range is dependent on mineral structure, mineral grain size, and rate of thermal cooling. Once the rate of decay of a particular radioactive isotope (parent) has been determined, the ratio of the remaining parent to its daughter decay products retained in a closed system will provide an estimate of the numerical (absolute) age, as well as associated error of the decay system. Decay systems of this kind may be applied to common (for example, biotite, potassium feldspar, and amphibole) and less common but still ubiquitous (for example, zircon, sphene, and badellyite) minerals in a wide range of igneous and metamorphic rocks. *See* BASEMENT ROCKS; METAMORPHIC ROCKS; RADIOACTIVITY.

The decay rate for each radioactive isotope is different, and therefore some isotopic decay schemes are ideal for very old rocks and some are better suited for young rocks or surface deposits. A slightly different system involves carbon-14 (^{14}C), which is a radioactive isotope produced by cosmic-ray interaction with nitrogen-14 (^{14}N) in the upper levels of the atmosphere. The production of ^{14}C is influenced by the intensity of the Earth's geomagnetic field. ^{14}C decays to ^{14}N on a time scale of several thousands of years. Over time, the ratio of unstable ^{14}C to the far more common carbon isotope ^{12}C has become stabilized in the atmosphere, and this ratio enters into the carbon incorporated by every living organism, where it is maintained as long as the organism is alive. When an organism dies, the ^{14}C gradually decays to ^{14}N, and the time at which the organism died can be estimated by how much ^{14}C is retained in the remains (for example, bone, shell, or plant fragment). Because of the relatively rapid rate of decay of ^{14}C, this radioactive isotopic decay scheme can be used only to date geologically young materials (that is, less than about 70,000 years). Isotopic age determinations of specific minerals and, sometimes, whole rocks constitutes geochronology. *See* DATING METHODS; GEOCHRONOMETRY; RADIOCARBON DATING; ROCK AGE DETERMINATION.

Geochronology did not become of geologic importance until the 1940s, when development of mass spectrometers permitted accurate determination of minute quantities of individual isotopes. Continual refinement of the instruments used for isotopic

analysis today permits high-precision age determinations of parts of individual mineral grains less than a few hundreds of micrometers in any dimension. The accuracy of refinements to the absolute geologic time scale continues to improve. *See* MASS SPECTROMETRY; MASS SPECTROSCOPE.

Parts of the relative time scale can be assigned absolute age estimates, and substantially refined, if a datable igneous rock such as a volcanic ash or lava flow is interlayered with fossiliferous strata. Few such fortuitous occurrences exist at or even near boundaries between specific intervals in the relative time scale. There are exceptions. In the 1990s, examination of strata deposited across the boundary between the Precambrian and the Cambrian periods (see table) in northern Siberia identified numerous horizons of volcanic origin, which contain grains of zircon that allow for high-precision uranium-lead (U-Pb) isotopic age determinations. Based on these data, the Precambrian-Phanerozoic (Cambrian) boundary was adjusted from about 570 million years ago to about 542 million years ago. Another example of a recent refinement is that of the age of the last full reversal of the Earth's magnetic field (from the Matuyama reverse polarity chron into the Brunhes normal polarity chron, in which we are now living). This is also taken as the base of the middle Pleistocene subseries. Early age determinations suggested that the Matuyama/Brunhes boundary was about 730,000 years ago. Recognition that young marine sedimentary sequences recorded astronomically tuned cycles led to the conclusion that the estimated age of the boundary could not be correct. Recent high-precision $^{40}Ar/^{39}Ar$ age determinations on lavas slightly older and younger than the boundary, as well as those that actually record the field polarity transition, have demonstrated that the actual age of the boundary is about 780,000 years. The boundary between the Cretaceous and the Tertiary (see table), the terminal extinction of the dinosaurs and many marine species, is now considered to be about 65 to 65.5 million years, depending on the time scale chosen, based on high-precision $^{40}Ar/^{39}Ar$ and U-Pb age determinations of meteorite impact–related materials from the Chixulub impact in the Yucatan peninsula, Mexico, and Deccan plateau basalts, which were erupted across the Cretaceous-Tertiary boundary, in India. In a general sense, geochronologists are making higher-precision estimates of the ages of specific geologic events to better define the geologic time scale. Intrusive igneous rocks that crosscut sedimentary rocks, igneous rock clasts in conglomerates, and very rare metamorphic rocks with preserved fossils can be used to provide some general definition of absolute ages of parts of the relative time scale. The geologic circumstances associated with each absolute age estimate dictate that many absolute age estimates for boundaries between intervals of the time scale must be interpolated between well-defined time points, and are thus constantly adjusted as better age data become available. The Phanerozoic geologic time scale (see table) is an evolving product of combining absolute and relative age components. All of the nomenclature has been long derived from the relative time scale, which is controlled by the fossil record, and absolute ages of specific boundaries are the best estimates available.

Based on isotopic age determinations, the age of events during the early, Precambrian part of Earth history, before there was a significant fossil record, has become far better understood. There is little international consensus for the nomenclature of the divisions of the Archean and Proterozoic eons. Geologists studying Precambrian rocks adhere to nomenclatures that are controlled solely by absolute ages, used to (loosely) define interval boundaries in the Precambrian (see table). *See* ARCHEAN; PRECAMBRIAN; PROTEROZOIC.

John W. Geissman; Allison R. Palmer

Bibliography. F. M. Gradstein, J. G. Ogg, and A. G. Smith (eds.), *A Geologic Time Scale*, 2004; W. B. Harland et al., *A Geologic Time Scale*, 1989; J. C. Reed et al. (eds.), *Precambrian: Conterminous United States*, vol. C-2 of *Geology of North America*, 1993; N. J. Snelling (ed.), *The Chronology of the Geologic Record*, Geol. Soc. (London) Mem. 10, 1985; S. Winchester, *The Map That Changed the World: William Smith and the Birth of Modern Geology*, 2001.

Geology

The science of the Earth. The study of the Earth's materials and of the processes that shape them is known as physical geology. Historical geology is the record of past events. *See* EARTH; EARTH SCIENCES.

Energy from two sources continually produces changes in the Earth. Radiant energy from the Sun causes ocean currents, winds, waves, rainfall, weathering, soil formation, and a myriad of other physical and chemical changes in the outermost rocky portion of the solid Earth (lithosphere), in the fluid envelopes of water (hydrosphere) and air (atmosphere), and in the totality of living matter (the biosphere). Heat energy inside the Earth causes slow convective movements deep in the Earth's interior. The internal motions break the rigid lithosphere into large fragments called tectonic plates, which move laterally at velocities up to around 5 in. (12 cm) a year. Collisions and other interactions between moving plates of lithosphere produce the Earth's gross topography—the ocean basins, mountain ranges, even the shapes of the continents themselves. *See* ATMOSPHERE; BIOSPHERE; HYDROSPHERE; INSOLATION.

Geologists examine rocks exposed at the Earth's surface and samples recovered from drilling. However, the radius of the Earth is 3982 mi (6371 km), and so the inner portions of the Earth must be studied remotely by means of the Earth's magnetic, electrical, gravitational, elastic, and other physical properties.

Following the spacecraft landings on the Moon, and space exploration of the other planets and their moons, the study methods of geology have been used in comparative planetology, in which the origin, development, and history of all solid bodies in the solar

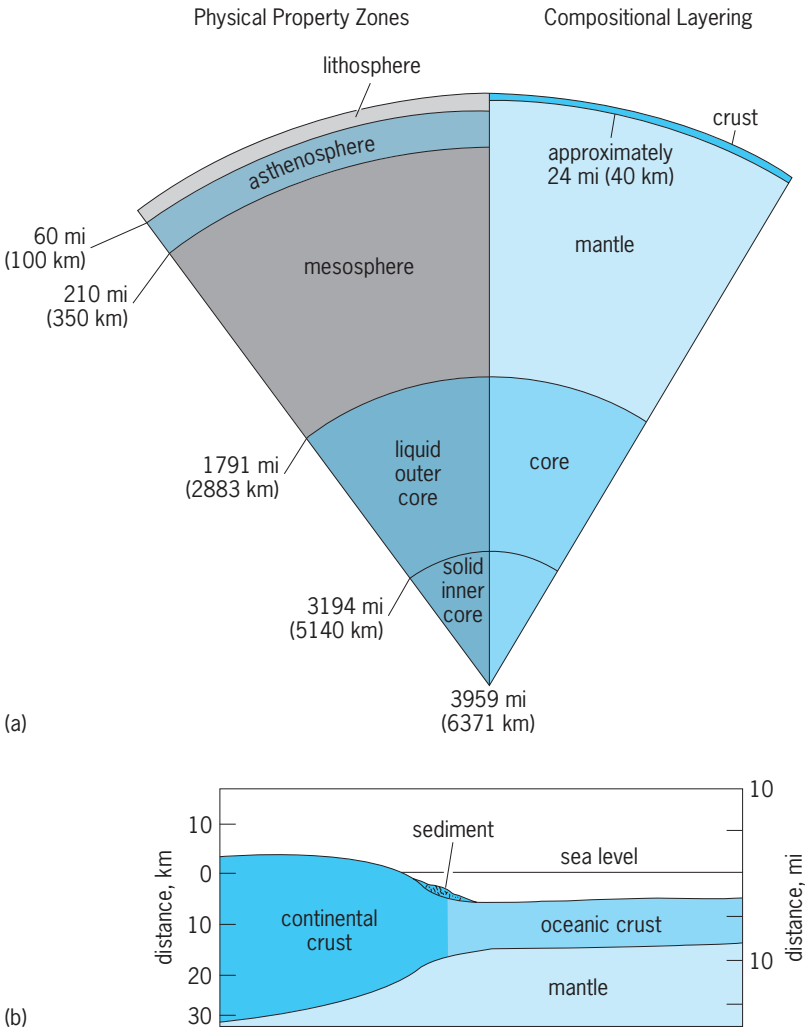

Fig. 1. Layers in the Earth. (*a*) Comparison of compositional layers and physical property changes in the Earth; temperature and pressure increase with depth. (*b*) Continental crust and oceanic crust.

system are compared. Geology became a universal science in the second half of the twentieth century, and an understanding of the geological evolution of the Moon, Mars, Venus, and other planetary bodies has provided a new perspective on the Earth's history. *See* PLANETARY PHYSICS.

Physical and historical geology. Geology is an interdisciplinary subject that overlaps and depends on other scientific disciplines. Physical geology is concerned primarily with the Earth's materials (minerals, rocks, soils, water, ice, and so forth) and the processes of their origin and alteration. Chemistry and physics are the two scientific disciplines most closely related—study of the chemistry of the Earth's materials is geochemistry, study of the physical properties of the Earth is geophysics. *See* GEOCHEMISTRY; GEOPHYSICS.

Historical geology is based on two complementary disciplines, stratigraphy and paleontology. Stratigraphy is the systematic study of stratified rocks through geologic time. The stratigraphic record reveals the sequence of events that have affected the Earth through eons of time. Absolute dates for the stratigraphic record are provided from geochemical studies of naturally occurring radioactive isotopes. Paleontology is the study of fossilized plants and animals with regard to their distribution in space and time. The fossil record is much more abundant in strata deposited during the present, or the Phanerozoic Eon, covering the past 545 million years. Paleontology is closely related to biology. The distinctions between physical and historical geology are more matters of convenience than substance, because it is increasingly clear, within the framework of plate tectonics, that all aspects of geology are interrelated. *See* PALEOBOTANY; PALEONTOLOGY; STRATIGRAPHY.

Internal structure and plate tectonics. The solid Earth is layered with respect to both composition and the physical properties (**Fig. 1**). The outermost compositional layer, the crust, is of two kinds: the oceanic crust, averaging about 6 mi (10) in thickness, everywhere underlies the ocean basins; the continental crust, averaging about 24 mi (40 km) in thickness, everywhere underlies the continents. Although the oceanic and continental crusts differ in composition, both are richer in silicon, aluminum, potassium, and sodium than the mantle which lies below. The mantle cannot be sampled directly, and its compositional heterogeneity is still uncertain. The innermost compositional layer is the core, a metallic mass composed largely of iron and nickel, with small amounts of silicon, oxygen, sulfur, and other chemical elements.

Layering of the Earth's physical properties arises from changes in pressure and temperature with depth. The outermost 60 mi (100 km) of the Earth, the lithosphere, is rigid and tough, and it can be fractured readily. From a depth of about 60–220 mi (100–350 km) is the asthenosphere, a region where temperatures are sufficiently high that rocks in the mantle are weak. Instead of fracturing, rocks in the asthenosphere flow and deform plastically. Beneath the asthenosphere is the mesosphere, a region where plastic properties decrease. Another distinct physical-property boundary occurs within the core—the outer core is molten, while the inner core, which has the same composition as the outer core, is solid. *See* ASTHENOSPHERE; EARTH INTERIOR; LITHOSPHERE.

Plate tectonics is one result of the Earth's layering of physical properties. Rigid plates of lithosphere float on the fluidlike asthenosphere. Continental crust and oceanic crust ride on top of lithospheric plates. Magma rising from deep in the mantle creates new oceanic crust where plates move apart. Where plates converge, the old oceanic crust is resorbed into the interior. Continental crust is characterized by a low density and cannot be dragged down into the mantle and resorbed in the same way as high-density oceanic crust. Instead, continental crust is continually moved around, colliding to form larger continents, fragmenting to form smaller units, in a fashion that endlessly changes the shapes of continents and the disposition of continents

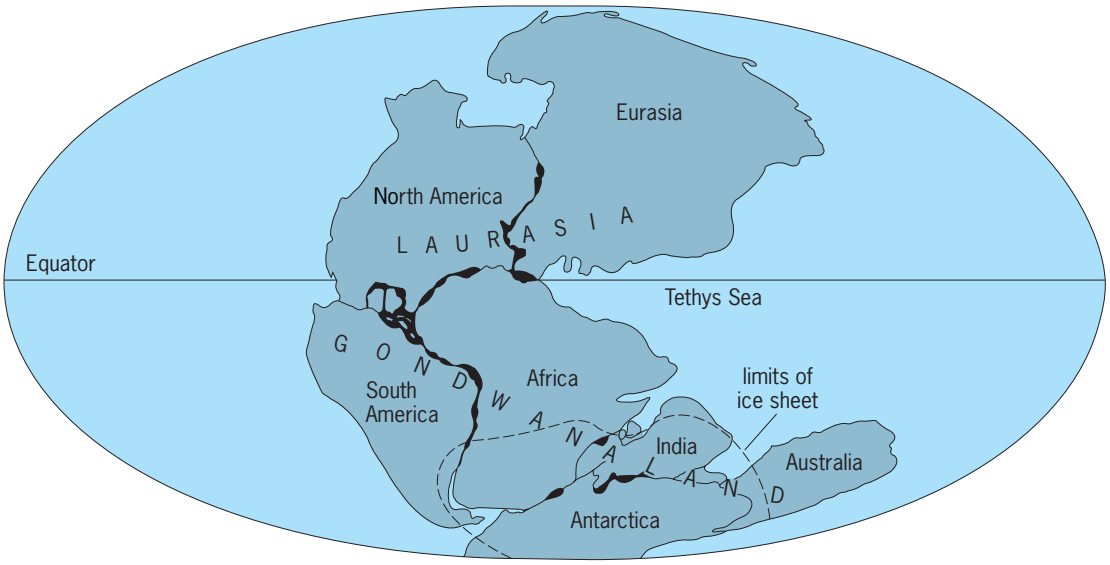

Fig. 2. Continents 250 million years ago, when they were assembled into a single large landmass called Pangaea. The northern portion of Pangaea is called Laurasia, the southern portion Gondwanaland. (*After B. J. Skinner and S. C. Porter, Physical Geology, Wiley, 1987*)

and ocean basins. Continental masses have collided and broken apart repeatedly throughout the Earth's history (**Fig. 2**). Geological activity is concentrated along plate boundaries and is most obviously manifested in earthquakes, volcanic eruptions, and the formation of mountain ranges such as the Himalaya and the Alps. *See* CONTINENT; CONTINENTAL DRIFT; CONTINENTS, EVOLUTION OF; EARTH, CONVECTION IN; EARTH CRUST; EARTHQUAKE; MAGMA; OROGENY; PLATE TECTONICS; VOLCANO.

Mineralogy. Minerals are the basic building blocks of rocks. About 3600 minerals have been identified, but fewer than 50 are common constituents in the types of rocks that are abundant in the Earth. The most common minerals in the crust are feldspars, quartz, micas, amphiboles, pyroxenes, olivine, and calcite. Each mineral is characterized by a distinct geometric packing of its constituent atoms, known as its crystal structure. Under ideal growth conditions the atomic packing is expressed as a crystal, which is a natural geometric solid bounded by plane faces. When a substance crystallizes in bulk, crowding of grains growing from neighboring centers prevents formation of recognizable crystals. An interlocking aggregate of irregularly shaped grains results. Most rocks have such an interlocking fabric of grains. Modern laboratories have effective devices for resolving the mineral content of rock materials; even the ultramicroscopic particles in clays are clearly defined under the electron microscope. *See* CRYSTAL STRUCTURE; ELECTRON MICROSCOPE; MINERAL; MINERALOGY.

Petrology. This is the study of rocks, their physical and chemical properties, and their modes of origin. Rocks are divisible into three primary families and three secondary ones. The primary families are igneous rocks, which have solidified from molten matter (magma); sedimentary rocks, made of fragments derived by weathering of preexisting rocks, of chemical precipitates from sea or lake water, and of organic remains; and metamorphic rocks derived from igneous or sedimentary rocks under conditions that brought about changes in mineral composition, texture, and internal structure (fabric). The secondary rock families are pyroclastic rocks, which are partly igneous and partly sedimentary rocks because they are composed largely or entirely of fragments of igneous matter erupted explosively from a volcano; diagenetic rocks are transitional between sedimentary and metamorphic rocks because their textures or compositions were affected by low-temperature, postsedimentation processes below conditions of metamorphism; migmatites are transitional between metamorphic and igneous rocks because they form when metamorphic rocks are raised to temperatures and pressures so that small localized fractions of the rock start to melt but the melting is insufficient for a large body of magma to develop. *See* PETROGRAPHY; PETROLOGY; ROCK.

Igneous rocks. Igneous rocks are formed as either extrusive or intrusive masses; that is, they are solidified at the Earth's surface or deep underground. Both kinds range widely in composition; silica, the most abundant ingredient, varies from about 40% to more than 75%. The most silica-rich varieties of igneous rock such as granite and granodiorite tend to be found in the continental crust. Basalt and gabbro, both containing about 50% silica, are the primary igneous rocks of the oceanic crust. *See* SILICA MINERALS.

Intrusive bodies (plutons) of granite and other silica-rich igneous rocks that formed at various depths are most numerous in mountain zones for two reasons: first, mountain belts have been much deformed, and abundant evidence indicates that crustal disturbance has favored igneous action;

Fig. 3. Nearly horizontal sedimentary rocks approximately 4000 ft (1200 m) thick along the Colorado River Canyon, Arizona. Some tilting of layers is discernible in the left section. (*Spence Air Photos*)

second, uplifts in mountain lands have permitted erosion to great depths so that plutonic masses are exposed. Evidence suggests that many large plutons solidified in magma reservoirs that formerly supplied volcanoes in action. *See* PLUTON.

Volcanic materials are erupted through two kinds of openings—central vents and long fissures. Central eruptions build up conical mountains; the materials are in part products of explosion, pyroclastic rocks, and partly interspersed lava flows. Lavas issuing from fissures do not build up volcanic cones but instead form vast fields of volcanic rock, chiefly basalt. The most extensive and most active volcanic fissures lie beneath the sea along the centers of the mid-ocean ridges. There, basaltic magma formed in the mantle rises and creates new oceanic crust. *See* BASALT; IGNEOUS ROCKS; LAVA; MID-OCEANIC RIDGE.

Fig. 4. Strata deformed by the compressive forces that created the Alps, near Val Malenco, northern Italy. (*Courtesy of Kurt Bucher*)

Sedimentary rocks. Bedrock exposed to the atmosphere and hydrosphere is broken into pieces, large and small, which are moved by running water and other agents to lower ground, and spread in sheets over river floodplains, lake bottoms, and sea floors. Dissolved matter is carried to seas and other bodies of water, and some of it is precipitated chemically and by action of organisms. The material deposited in various ways becomes compacted, and in time much of it is cemented into firm sedimentary rock. Generally, the deposition is not continuous but recurrent, and sheets of sediment representing separate events come to form distinct layers of sedimentary rock. The individual layers are beds or strata, and the rocks are said to be stratified (**Fig. 3**).

Large areas in every continent are underlain by sedimentary rocks that represent deposits during many periods of the Earth's history. In part, these bedded rocks are nearly horizontal, as they were originally; but in many places, particularly along the margins of present or past tectonic plates, elongate belts of bent and fractured strata can be observed in mountain belts (**Fig. 4**).

The principal kinds of sedimentary rock are conglomerate, sandstone, siltstone, shale, limestone, and dolostone. There are many other kinds, and some, though less important quantitatively, have large practical value; examples include common salt, gypsum, rock phosphate, iron ore, and coal. *See* SEDIMENTARY ROCKS.

Metamorphic rocks. These rocks have been developed from earlier igneous and sedimentary rocks by heat and pressure, especially in mountain zones found along plate margins, and near large masses of intrusive igneous rock. Thermal metamorphism adjacent to plutons results from rising temperatures, often with addition of new elements by circulating fluids, but without pronounced deformation of strata. Common effects are hardening and crystallizing of the affected rock, with changes in mineral composition.

Dynamic metamorphism involves elevated temperatures but also results from shearing stresses in rocks subjected to high pressure; effects are development of cleavage planes and growth of platy minerals in parallel arrangement. Dynamic metamorphism is most commonly observed in mountain belts where the crust is thickened and deformed due to plate collisions.

The common metamorphic rocks are in the two general classes, foliated (including slate, phyllite, schist, and gneiss) and nonfoliated (including marble and quartzite). *See* METAMORPHIC ROCKS; METAMORPHISM.

Regolith. Bedrock at and near the Earth's surface is subject to mechanical and chemical changes in a complex group of processes called weathering. Blocks and small chips that become detached due to mechanical fragmentation are especially vulnerable to chemical attack, which makes radical changes

in the mineral content. Some soluble products of alteration are removed by percolating water; and in the less soluble residue the most common constituents are quartz and clay, which are the basic minerals of soil. The effectiveness of weathering and the nature of its products are controlled by climate, topography, kinds of bedrock, and other variables. Organic processes play a major role in chemical weathering, which is most effective under conditions that favor development of bacteria, plants, and ground-dwelling animals. The blanket of weathered material that covers most of the Earth's surface is called the regolith.

Weathering prepares the way for removal of rock materials and reshaping of land surfaces by several agents of erosion. The most obvious of these agents is running water, which during a single rainstorm may cut deep gullies into plowed fields and sweep vast quantities of soil, sand, and coarser debris into brooks and eventually into channels of major streams. Abrasion by such moving loads deepens and widens stream channels in hard bedrock. Study of drainage systems brings conviction that even the largest and deepest valleys have been fashioned by the action of running water. *See* WEATHERING PROCESSES.

The pull of gravity causes the regolith to move downslope, from high points to low. Soil on slopes, even those covered with grass and other vegetation, creeps slowly downward. Blocks dislodged from cliffs build steep masses of sliderock which slowly migrate downslope. Frequently, in mountain lands great masses, including loose material and bedrock, rush down as landslides. All downslope movements caused by gravity are termed mass wasting. *See* LANDSLIDE; MASS WASTING; REGOLITH.

Water moving through underground openings dissolves and carries away great quantities of material. Caverns, large and small, are a conspicuous result. In high latitudes and in some mountain regions, glaciers are powerful eroding agents. In arid regions, quantities of sand and dust are moved by wind. Large-scale erosion along the coasts of seas and lakes is performed by waves and currents. *See* CAVE; DESERT EROSION FEATURES; GLACIATED TERRAIN.

Each major agent of erosion fashions characteristic features in landscapes; the subdiscipline of geology devoted to studies of these features is geomorphology. The net tendency of erosion is to reduce the height of landmasses. Various stages in the history of reduction are indicated by forms of valleys and slopes and by relations of land surfaces to the underlying bedrock. Some wide regions have been uplifted, and the streams have been rejuvenated after an advanced stage was reached in a cycle of erosion. Rejuvenation of an eroded land surface is believed to result, at least in part, from plate tectonics. *See* EROSION; GEOMORPHOLOGY.

Sediments and sedimentation. Sediment now deposited provides essential clues to understanding processes of past sedimentation and hence of the origin of sedimentary rock. Sediment and sedimentary rock are all important in geology. On the basis of depositional environment, sediments are assigned to three categories: terrestrial, those laid down on land; marine, those deposited on sea floors; and mixed terrestrial-marine, those laid down in transitional zones such as deltas, marine estuaries, and areas between high and low tide. In each major group the sediments are further described as clastic (consisting of rock fragments) and chemical (formed either as inorganic precipitates or partly through organic agencies). Study of modern sediments in the several environments takes account of physical peculiarities and the included remains of organisms.

Terrestrial sediments are extremely variable. They are laid down chiefly through the agencies of mass wasting, running water, glacier ice, and wind. The deposits are in large part temporary, as the tendency is for them to shift downslope and seaward in continued erosion of the lands. Marine sediments, deposited beneath oceans, comprise material derived from the land, from shells and skeletons of marine animals and plants, from the ocean by chemical precipitation, and from space (particles of meteorites). *See* MARINE SEDIMENTS; NEARSHORE PROCESSES.

Tectonics and structural geology. Through careful study of rock masses, it is possible to distinguish primary structures (such as bedding in sedimentary rocks, and flow structures in igneous rocks) that were acquired as a result of the genesis of the rock, from secondary structures that result from later deformation. Significant features in sedimentary rocks make them especially valuable for registering later changes in form.

Deformation may proceed rapidly or slowly. Rapid deformation usually involves large fractures (faults) with instantaneous displacements of several meters (**Fig. 5**). Rapid deformation is generally attended by strong earthquakes. Slow deformation leads to broad-scale warping, bending, and folding. The principal kinds of structural features that record past deformation are folds, joints, faults, cleavage, and unconformities. Tectonics is closely related to structural geology but deals with regional features,

Fig. 5. Fault breaking through to the surface, near Hilo, Hawaii. (*R. S. Fiske, U.S. Geological Suvey*)

such as mountain ranges. *See* FAULT AND FAULT STRUCTURES; FOLD AND FOLD SYSTEMS; GEODYNAMICS; STRUCTURAL GEOLOGY.

The most pronounced crustal deformation tends to be found in mountain belts, where erosion has exposed exceptionally thick sections of sedimentary rocks. These rocks record long histories of slow subsidence and sedimentation, interrupted by large-scale deformation and uplift. The deformation is due to lateral compressive forces arising from plate tectonic movements. *See* CORDILLERAN BELT; MOUNTAIN SYSTEMS.

Major relief features of the Earth reflect differences in density of the underlying rocks. Continental rocks have appreciably lower average density than the basaltic rocks of the oceanic crust. Continents therefore stand high, while ocean basins are low. Great mountain blocks, such as the Alps and the Himalaya, represent thickened parts of continental crust. The condition of approximate balance among diverse parts of the crust is called isostasy (equal standing). *See* ISOSTASY.

Stratigraphy. Studies of sedimentary rocks in comparison with the many kinds of modern sediments provide a basis for recognizing conditions under which the rocks were formed. Generally, a close interpretation is possible; bodies of rock are confidently classified as deposits on floodplains, at margins of glaciers, in large lakes, in shallow seas near shore, or in deep-sea troughs. Each distinctive type of deposit represents a facies. A rock type that is essentially uniform over a considerable area constitutes a lithofacies. An assemblage of fossils that is nearly uniform in a large unit of sedimentary deposits, indicating an environment suited to certain forms of life, is a biofacies. Deposits formed at the same time may differ in both lithofacies and biofacies, reflecting differences in topography, climate, and other items of environment. *See* FACIES (GEOLOGY).

A fundamental principle in stratigraphic studies, known as the law of superposition, is that in a normal sequence of strata any layer is older than the layer next above it. This elementary law is of importance in studies of many belts of highly deformed rocks where thick sections of strata have been overturned, even completely inverted, and can be resolved only through criteria that indicate original tops of beds. Close matching of the many kinds of modern sediments with materials in sedimentary rocks formed over an immense span of time has established the uniformitarian principle, which holds that processes now operating on the Earth have operated in similar fashion through the ages.

The history of the Earth's crust is encoded chiefly in sedimentary rocks, which record a sequence of events, changing physical environments, developments in plant and animal life, and effects of crustal movements. Additional records are supplied by volcanic rocks, which in many areas are interlayered with, and grade into, sedimentary strata; by relations of intrusive igneous bodies to older and younger rocks; and by erosion surfaces, some displayed in present landscapes, others revealed in exposures of unconformities. The long history includes radical changes in physical geography, featuring a contest between land and sea; the birth, rise, and wasting away of successive mountain systems; and the evolution of living forms in seas and on lands. *See* UNCONFORMITY.

Each continent (other than ice-buried Antarctica) displays a wide lowland or platform that was occupied repeatedly by seas and is now mantled with little-deformed marine strata. The total deposit on each platform represents a long span of time and ranges from a few tens to a few thousand meters thick. The join between continental crust and oceanic crust marks the geological boundary between an ocean basin and a continent. The boundary is beneath the sea, at the foot of the continental slope, and it is here that sediment eroded from the continent is accumulated. The sites of such accumulations of sediment classically were called geosynclines. When continental collisions occur, the wedge of sedimentary strata draped along and over the continental margin becomes deformed through folding and faulting, and elevated into a mountain range (Fig. 5). Examples of this process can be found in the Appalachian Mountains and the Alpine-Himalayan chain. One of the triumphs of plate tectonics is the illumination it has thrown on this long-puzzling aspect of mountain range development. *See* CONTINENTAL MARGIN; GEOSYNCLINE; PHYSICAL GEOGRAPHY.

Geologic column and geologic time scale. At any one locality a sequence of sedimentary beds, from older to younger, can be determined through physical evidence. Persistence of some peculiar units may establish approximate correlations through moderate distances, occasionally hundreds of miles. But the key to relative dating of stratigraphic units and to confident worldwide correlations is provided by fossils of animals and plants, which record progressive evolution in living forms from ancient to recent times. Through correlation a worldwide composite diagram has been developed that combines, in chronological order, the succession of known strata on the basis of fossil contents or other evidence. This composite diagram is the geologic column.

Many of the oldest known sedimentary rocks are now highly metamorphosed, and any fossils they may have held must have been obliterated. Some thick sections of old strata that are not appreciably altered have yielded only sparse indications of life, such as patterns of marine algae and burrows made by worms or other lowly forms. Successively younger groups of strata hold abundant fossils of marine invertebrates, marine fishes, land plants that progressed from primitive to more modern kinds, reptiles, birds, and small mammals, followed by diverse and generally more advanced kinds of mammals, primitive humans, and finally modern humans. Some forms evolved rapidly, and short-lived species that were equipped to become widely dispersed are of greatest value for correlation.

Geologic column and scale of time

Eon	Era	Period	Epoch	Dates (10^6 years BP)
Phanerozoic	Cenozoic	Quaternary	Holocene	0.01
			Pleistocene	1.8
		Tertiary	Pliocene	5.3
			Miocene	23
			Oligocene	33.9
			Eocene	58.8
			Paleocene	65.5
	Mesozoic	Cretaceous		145
		Jurassic		200
		Triassic		251
	Paleozoic	Permian		299
		Pennsylvanian		318
		Mississippian		359
		Devonian		416
		Silurian		444
		Ordovician		488
		Cambrian		542
Proterozoic*	No subdivisions in wide use			2500
Archean*	No subdivisions in wide use			3600
Hadean	No subdivisions			4560

*Proterozoic plus Archean also called Precambrian.

Absolute dates can be given to points on the geologic column, and a geologic time scale developed thereby, through the use of certain naturally occurring radioactive isotopes. Such dating is known as geochronology. The names given to units of the geologic column (see **table**) are derived largely from the paleontologic record—or its absence. *See* GEOCHRONOMETRY; GEOLOGIC TIME SCALE.

Geologic mapping. A geologic map represents the lithology and, so far as possible, the geologic age of every important geologic unit in a given area. Each distinctive unit that can be shown effectively to the scale of the map is a geologic formation. A good topographic base map is essential for representing relations of bedrock to land surface forms. Cooperation of workers with specialized qualifications—for example, in petrology, paleontology, or structural geology—required for accurate mapping and description of complex areas. Large organizations, such as federal geological surveys and some commercial firms, possess diversified personnel, laboratories equipped for varied analyses, and special field equipment. Aerial photographic surveys serve as a guide in field work and help in plotting accurate locations. Satellite imagery has provided vast quantities of data for mapping. *See* AERIAL PHOTOGRAPH; REMOTE SENSING; TOPOGRAPHIC SURVEYING AND MAPPING.

A completed geologic map should indicate important structural details, such as inclinations of strata, locations of faults, and axial traces of folds. Usually the map is supplemented by vertical sections on which structural features seen at the surface are projected to limited depths. Maps of small scale may represent sedimentary rocks only according to the systems to which they belong. With larger scale a given system may be represented by several formations, each recording an important episode in the history of the region. In some European countries, geologic mapping has been completed to fairly large scales; but in all continents great areas have been mapped only in reconnaissance fashion.

Economic geology. A general knowledge of geology has many practical applications, and large numbers of geologists receive special training for service in solving problems met in the mining of metals

and nonmetals, in discovering and producing petroleum and natural gas, and in engineering projects of many kinds. Human use of materials has become so great that waste materials are influencing natural geological processes. As a result, a new discipline, environmental geology, is starting to emerge. *See* ENGINEERING GEOLOGY; PETROLEUM GEOLOGY.

Brian J. Skinner

Bibliography. P. W. Birkeland, *Soils and Geomorphology*, 1999; H. Blatt, R. J. Tracy, and B. Owens, *Petrology: Igneous, Sedimentary, and Metamorphic*, 3d ed., 2005; R. J. Chorley, S. A. Schumm, and D. E. Sugden, *Geomorphology*, 1985; A. Cox and R. B. Hart, *Plate Tectonics: How It Works*, 1988; J. R. Craig, D. J. Vaughan, and B. J. Skinner, *Resources of the Earth*, 3d ed., 2001; G. H. Davis and S. J. Reynolds, *Structural Geology of Rocks and Regions*, 2d ed., 1996; G. M. Friedman, J. E. Sanders, and D. C. Kopaska, *Principles of Sedimentary Deposits: Stratigraphy and Sedimentation*, 1994; C. Klein, *Manual of Mineralogy*, revised 22nd ed., 2001; A. D. Miall, *The Geology of Stratigraphic Sequences*, 1996; A. D. Miall, *Principles of Sedimentary Basin Analysis*, 3d ed., 2000; B. J. Skinner and S. C. Porter, *The Blue Planet: An Introduction to Earth System Science*, 2d ed., 1999; B. J. Skinner and S. C. Porter, *The Dynamic Earth: An Introduction to Physical Geology*, 5th ed., 2003; S. M. Stanley, *Earth System History*, 2d ed., 2004; S. M. Stanley, *Exploring Earth and Life Through Time*, 1993.

Geomagnetic variations

Variations in the natural magnetic field measured at the Earth's surface and elsewhere in the Earth's magnetosphere (for example, at the geostationary orbit). These are field changes with periodicities from about 0.3 second to hundreds of years. (These boundaries are set to distinguish geomagnetic variations from the quasipermanent field and higher-frequency waves.) Many of these observed variations—from very short periods (seconds, minutes, hours) to daily, seasonal, semiannual, solar-cycle (11-year), and secular (~60–80 years) periods—arise from sources that either are external to the Earth (but superposed upon the larger, mainly dipolar field) or internal to the Earth (the magnetic-dipole and higher-harmonic trends and variations on the scales of hundreds and even thousands years). The daily and seasonal motions of the atmosphere at ionospheric altitudes cause field variations that are smooth in form and relatively predictable, given the time and location of the observation. During occasions of high solar–terrestrial disturbance activity that give rise to aurorae (northern and southern lights) at high latitudes, very large geomagnetic variations occur that even mask the quiet daily changes. These geomagnetic variations are so spectacular in size and global extent that they have been named geomagnetic storms and substorms, with the latter generally limited to the polar regions.

Solar quiet-time variations. The recurrent patterns of daily geomagnetic field changes (**Fig. 1**) arise in the upper atmosphere through dynamo current processes occurring at 100–120-km (60–75-mi) altitudes in the E ond lower F regions of the ionosphere. The charged particles of the daytime ionosphere are driven by thermotidal and wind forces through the main geomagnetic field to form a local current, extending primarily over the sunlit side of the Earth.

The ionization and the tidal and wind forces vary with time of day and season; the tidal and wind forces are dependent upon geographic location. However, the Earth's main field is offset from the geographic axis alignment and thus provides some geomagnetic coordinate organization to solar quiet-time (*Sq*) variations. Because the Earth's interior composition is electrically conducting, the solar quiet-time current in the ionosphere causes a secondary current to be induced within the Earth; the magnetic field

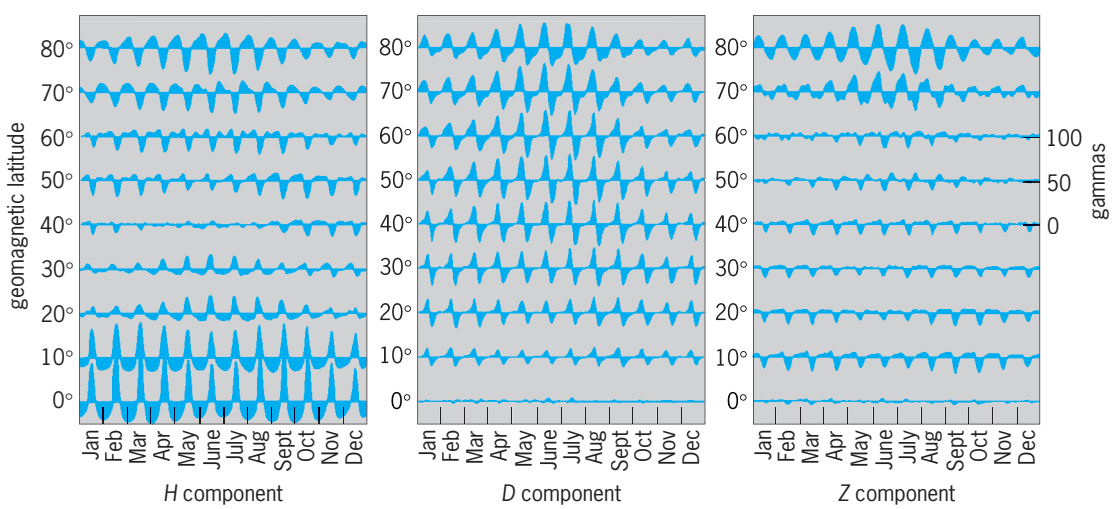

Fig. 1. Annual picture of the midmonth daily variations of quiet-time geomagnetic field, in local time, for the Northern Hemisphere from the Equator (0°) to 80° latitude displayed for *H*, *D*, and *Z* field components. The scale size between baselines is 50 nanotesla. (1 nT = 1 × 10^{-9} tesla = 1 gamma—a geomagnetic induction unit used in the past.)

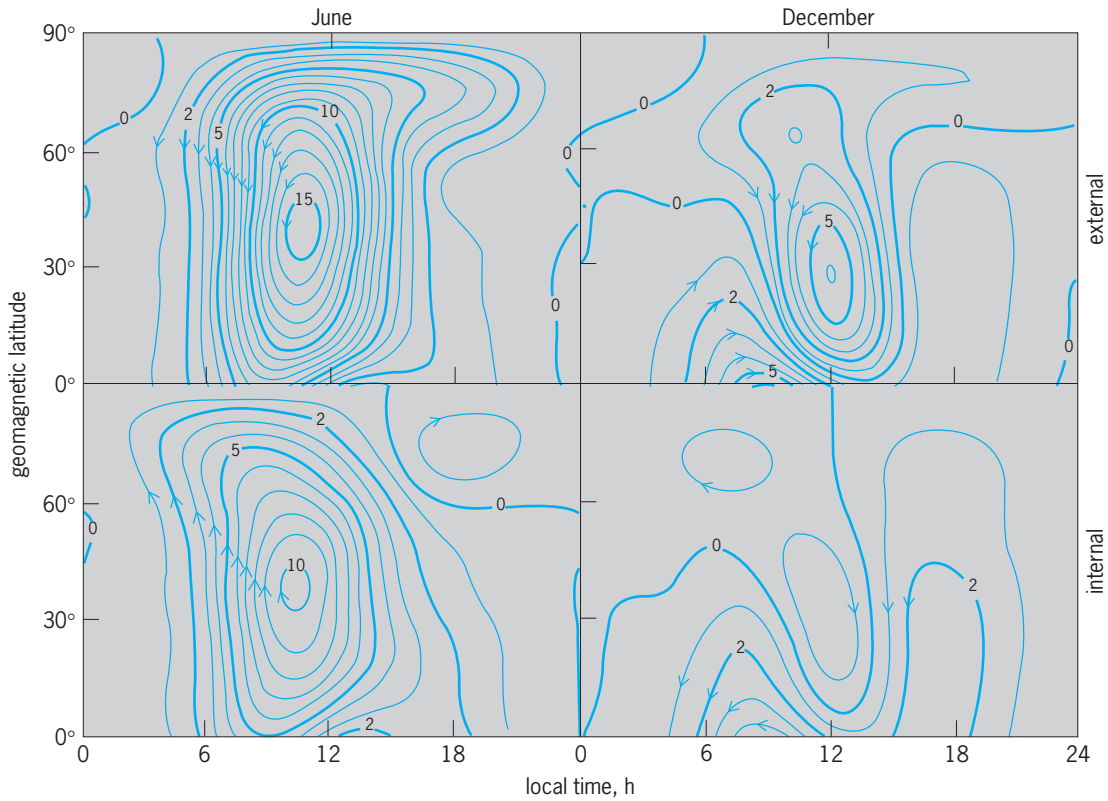

Fig. 2. Streamlines for the equivalent external ionospheric source currents (top) and internal induced currents (bottom) for solar quiet-time (*Sq*) variations of the geomagnetic field in the Northern Hemisphere for midmonth quiet conditions, in June and December. Each pattern, in local time versus geomagnetic (dipole) latitude coordinates, shows the equivalent current contours, with 10^{-4} ampere flowing in the region between contour lines. Arrows indicate the required flow direction. The midnight zero level that is assumed tor the display has no effect on the current pattern.

variations generated by these two currents are observed at the Earth's surface. In addition, a resolution of these external and internal parts of the observed fields can be used to calculate the remote deep-earth electrical conductivity profiles.

Figure 2 illustrates the computed "equivalent" external *Sq* current patterns in the ionosphere and those induced in the Earth (internal) for the summer and winter months of the Northern Hemisphere that could provide the observed fields. The denser the contours, the higher the surface field contribution.

The right-hand rule (with the right hand wrapped around the current vector and the thumb pointing in the current direction, the fingers point in the resulting magnetic field direction) can be applied showing the daily changes of field illustrated by Fig. 1. Near the Equator, the (almost) horizorital orientation of the main geomagnetic field lines causes an especially high ionospheric conductivity for the *Sq* system; as a result, there arises an intense eastward daytime current called the equatorial electrojet. In the polar cap region, there are quiet-time, Birkeland field–aligned currents flowing between the magnetosphere and ionosphere that add to the dynamo source of quiet-time geomagnetic field variations. Because magnetospheric behavior is sensitive to the direction of the interplanetary magnetic field (IMF) arriving with the solar wind particles blown out from the Sun, and because polar region field lines reach the outer parts of the magnetosphere, a signature of this interplanetary field is embedded in the polar cap *Sq* records. *See* ATMOSPHERIC TIDES; GEOMAGNETISM; INTERPLANETARY MAGNETIC FIELD; IONOSPHERE; MAGNETOSPHERE; SOLAR WIND.

Lunar variations. The seinidiurnal lunar tidal oscillations of the atmosphere drag the E-region ionization through the main field of the Earth and produce another dynamo current, *L*. The fields of this current are quite small, less than 10% of the *Sq* amplitude, and so special analytical methods are required to isolate their contribution to the observed geomagnetic records. The lunar tidal "day" is 50.5 minutes longer than the solar day, and the global *L*-amplitude patterns depend upon the twice-daily lunar tidal forces on the atmosphere, the E- and F-region ionization, and the direction and magnitude of the main magnetic field of the Earth.

Eclipse and solar flare effects. Temporary conductivity increases in the ionization due to direct x-ray radiation from solar flares, or decreases in ionization due to solar eclipses, can modify the *Sq* and *L* dynamo currents. Such variations appear as single half-cycle changes in the geomagnetic field that last from several minutes to about an hour, with maximum amplitudes rarely larger than 10 nanotesla. These small-amplitude variations are best observed on the extremely quiet *Sq*-condition days. *See* SUN.

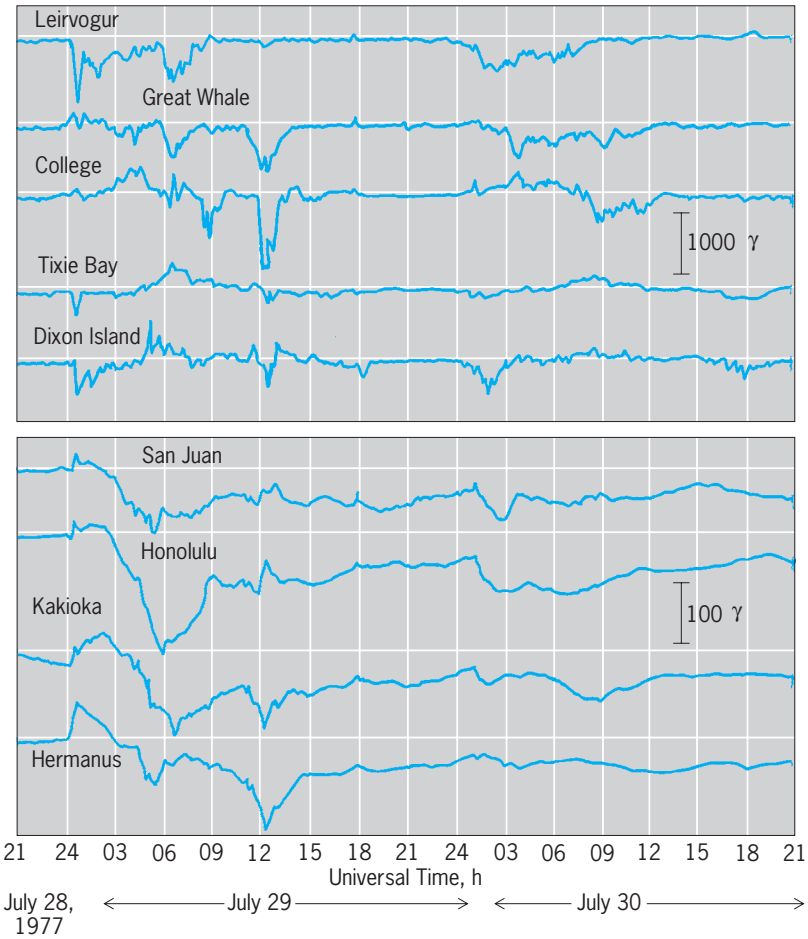

Fig. 3. Common scale magnetograms for the *H* component of a geomagnetic storm variation recorded at two groups of stations: the auroral zone (top) and the low-latitude regions (bottom). Gamma (γ) is the geomagnetic induction unit. (*Records prepared by National Geophysical Data Center, Boulder*)

Magnetic storms. Large geomagnetic disturbances (storms) are approximately one-hundredth of the Earth's main-field strength and are caused by shocks the magnetosphere experiences when significant solar wind disturbances (also known as coronal mass ejections, CME) arrive at the Earth's orbit. As a result, solar-wind-charged particles enter the magnetosphere and increase the Earth's ring current at 3.5 to 9 earth radii. The *H* (magnetic northward) component of the variation field typically shows the greatest arnplituldes (**Fig. 3**). *See* SOLAR CORONA.

Many storms have a similar appearance in the *H* component that is divisible into three parts: sudden commencement, or initial phase; main phase; and recovery phase. The sudden commencements are recorded simultaneously (within minutes) about the entire Earth. Usually there are a sudden onset and then the increase in the northward field strength that may continue for as long as several hours. During the main phase, the *H* component decreases, and this decrease often lasts longer than the initial phase and is several times larger in amplitude. The follow-on recovery to the quiet-time level takes longer than the other two phases and may even extend to several days. The more intense storms show an increase of both amplitude and duration. At some observatories, only the storm's main phase is observed.

Storm intensity and occurrence. The storms are most intense at the latitudes of the nightside auroral zone, where they can be about six to ten times larger than at middle latitudes. They show minimum amplitudes in the region of 20 to 30° geomagnetic latitude on the nightside of the Earth and have a secondary maximum near the dayside equatorial region. At middle latitudes, about 10 storms per year attain a magnitude of over 50 nT; about one or two storms per year attains over 250 nT. There is an increase in activity during the equinoxes, and the storm variation amplitudes are slightly larger in the winter hemisphere. The number and intensity of storms vary with changes in the sunspot number and lag behind the 11-year solar activity cycle by about a year or two. There are direct relationships of the storms with the solar outbursts and solar magnetic field orientation (which changes every 22 years), as well as the high-velocity solar wind and the interplanetary magnetic field direction.

The planetary-wide geomagnetic activity is measured at each magnetic observatory by the local *K* index, a quasilogarithmic scale indicating the range of most disturbed components of the geomagnetic field in a 3-hour interval. The *K* index is obtained from the range of the field changes about the estimated *Sq* variation during the same period; this value is normalized for the observatory location. An average of the indices from selected world observatories provides a "planetary" index, *Kp*. Storm magnitudes are arranged in size either by their highest *Kp* value or by the ring current index, *Dst*.

Polar substorms. A significant amount of energy can be delivered into the Earth's magnetosphere by a group of related physical processes called polar substorms. These phenomena occur when solar wind blows from the Sun at higher (500–800 km/s) than usual (350–400 km/s) velocities. If the incoming solar wind carries along a southward-directed magnetic field (that is, opposite to the Earth's field at its interface with the solar wind at about 10–12 earth radii distance from the Earth's center) for a prolonged time, a polar substorm can be triggered and then observed in the high-latitude magnetic and ionospheric parameter variations. There follows an explosive precipitation of particles into the midnight sector of the ionosphere, a spectacular increase in aurorae, a massive flow of field-aligned currents to and from the auroral region, and the dramatic development of geomagnetic westward and eastward electrojet currents in the auroral zone ionosphere (**Fig. 4**) and in the polar cap. These currents give rise to local heating of the high-altitude atmosphere and to the decrease and violent variations of the *H* magnetic field component seen on the magnetic storm records at auroral latitudes (top four records of Fig. 3). Very

short period pulsations (from seconds to minutes) of the geomagnetic field, known as ultralow-frequency (ULF) micropulsations, that are identified with corresponding fluctuations of the auroral luminosity are observed at such times.

The substorm disturbance often proceeds through three stages: (1) the growth phase (of several tens of minutes), representing the time of injection of energy from the nightside magnetosphere; (2) the expansive or explosive phase (several minutes or more), when the disturbance rises to its maximum in amplitude and effective area; and (3) the decay phase (up to an hour or two) as the event subsides. Consecutive substorms can blend (or follow in a sequence) for several hours during the main phase of a magnetic storm. A peak of the substorm activity is usually restricted to a region of less than 5° in latitude and 100° in longitude at the Earth's nightside ionosphere. The ratio of the high- to low-frequency components of the geomagnetic variations decreases rapidly with distance from the disturbance center. Evidence of the substorm is carried to lower latitudes by electric and magnetic fields from the closure of strong auroral-region westward electrojet currents within the ionosphere and from the Birkeland currents of the magnetosphere. The auroral electrojet index, AE, compiled from magnetic records obtained at observatories located in the auroral latitudes, is used as a measure of the substorm intensity. *See* AURORA.

Ring current. During the main phase of the magnetic storms and often during the growth phase of the substorms, energetic ions and electrons are fed into a ringlike region at about 3–9 earth radii distance in the equatorial plane of the Earth. There the complicated field and particle interactions generate a westward-flowing ring current due to a charge separation as the energetic protons and electron move toward the Earth from the magnetospheric tail. By using the right-hand rule, it can be seen that this current causes a worldwide southward field roughly parallel to the Earth's dipole axis. The resulting depression of the measured fields about the entire Earth, particularly apparent at the lower and equatorial latitudes, contributes to the main phase of the geomagnetic storm. With the decay phase of the substorm, the source of maintenance protons disappears; this allows a slow decay of the ring current. This process is seen as a recovery phase of the magnetic storm, and it lasts a few hours to several days, depending upon the intensity of the substorms that might occur during the storm's main phase. A ring current index, *Dst*, is derived from the storm-time (*st*) disturbance (*D*) field levels of low-latitude stations, with the *Sq* variation removed from the data and only the axially symmetric contributions to the field being considered. Actually, at the onset of the storm the positive values of *Dst* represent the compression of the magnetosphere by the solar wind rather than a ring current (which causes negative *Dst* values). These compression and ring current characteristics can be seen as the initial arid recovery phases of the magnetic storms on the low-latitude observatory traces (bottom four) of Fig. 3.

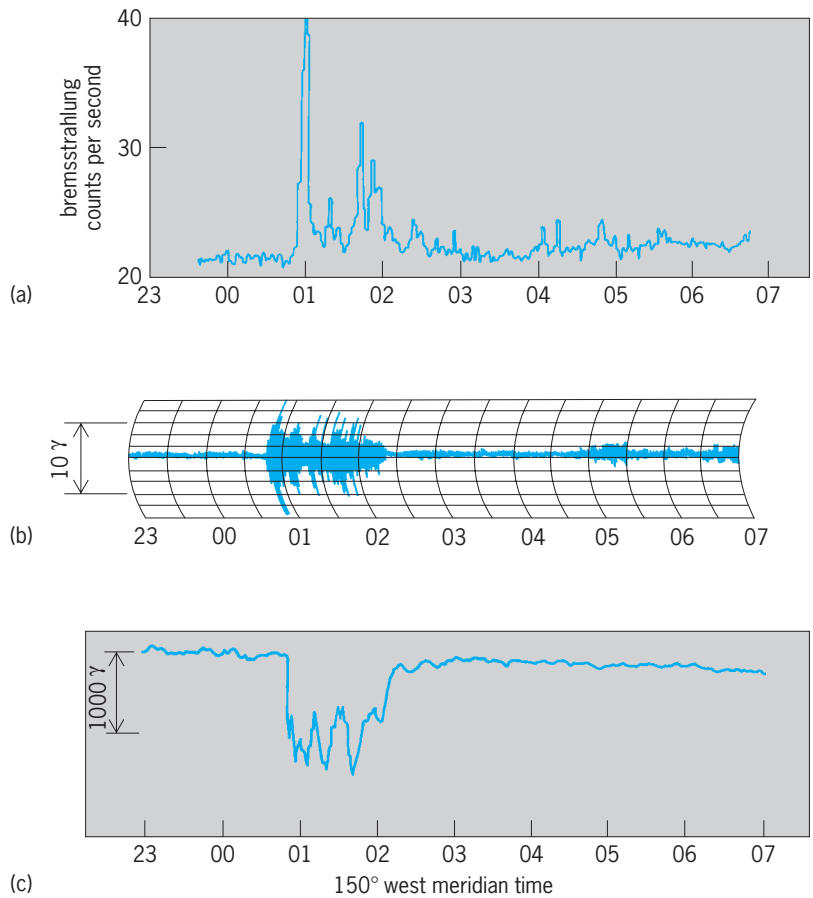

Fig. 4. Concurrent polar substorm phenomena observed at College, Alaska. (*a*) Low-energy electron precipitation into the ionosphere measured as the bremsstrahlung x-ray count rate. (*b*) Geomagnetic field micropulsations recorded by a north-south axis sensor. (*c*) *H* component of magnetic field. 1 gamma (γ) = 1 nT.

Rapid variations. The geomagnetic spectrum from several minutes to about a third of a second shows activity assolciated with solar–terrestrial disturbances. The irregularly shaped (on an amplitude–time trace) *Pi* pulsations are ideintified with the substorm onset (Fig. 4). The more smoothly varying (continuous) pulsations, *Pc*, also occur in association with the unsettled magnetospheric environment; they are recognized as having special period groups of several minutes (*Pc*4, *Pc*5), about 30 to 15 s (*Pc*3, *Pc*2), and about 0.5 to 5 s (*Pc*1) for which amplitudes near 10, 0.3, and 0.05 nT, respectively, are often reported.

These oscillations arise as hydromagnetic waves whose periods and amplitudes are governed by the charged-particle population and main-field configuration within the magnetosphere and by the transmission of energy from the magnetosphere into the ionosphere. All these variations except *Pc*1 are closely associated with auroral luminosity fluctuations. The *Pc*1 micropulsations (**Fig. 5**) have been shown to arrive at the high-latitude ionosphere as hydromagnetic waves that subsequently propagate in the F-region ionospheric duct

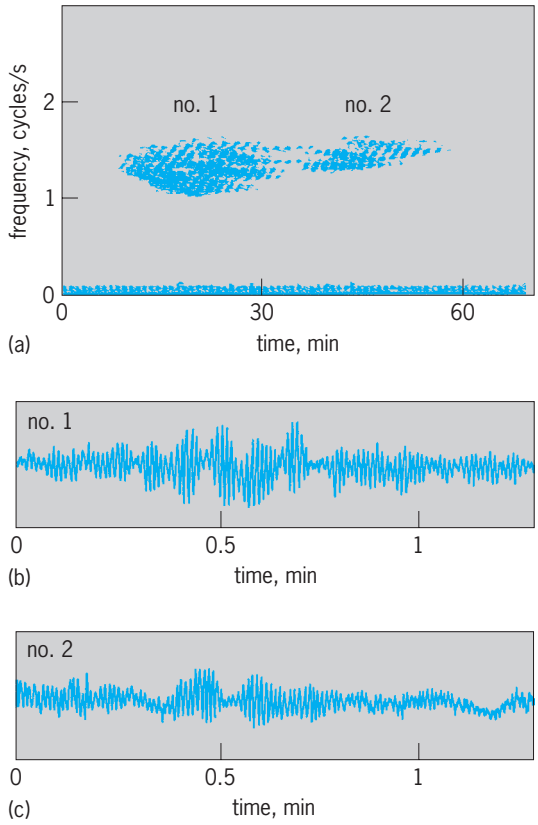

Fig. 5. Example of *Pc*1 geomaignetic pulsation event recorded at Boulder, Colorado. (*a*) Frequency-versus-time display of a two-part event. Note the unique rising frequency structure about a midfrequency near 1.5 cycles/s. (*b*) Amplitude-versus-time representation for no. 1 event and (*c*) no. 2 event. The beating appearance in the amplitudes is largely due to the overlapping elements of the rising frequency structure.

to the lower latitudes. The *Pc*1 pulsations occur most frequently during times of high magnetic activity in the week following major substorms.

Vladimir O. Papitashvili; Wallace H. Campbell

Bibliography. S.-I. Akasofu and S. Chapman, *Solar Terrestrial Physics*, Oxford Clarendon Press, 1972; W. H. Campbell, *Introduction to Geomagnetic Fields*, Cambridge University Press, 1997; S. Chapman and J. Bartels, *Geomagnetism*, vols. 1 and 2, Oxford University Press, 1940; J. Jacobs (ed.), *Geomagnetism*, vol. 3, Academic Press, New York, 1989; F. Lowes et al. (eds.), *Geomagnetism and Paleomagnetism*, 1988; A. Nishida, *Geomagnetic Diagnosis of the Magnetosphere*, Springer-Verlag, New York, 1978; T. A. Potemra (ed.), *Magnetospheric: Currents*, AGU Geophys, Monogr. 28, Washington, DC. 1984.

Geomagnetism

The magnetism of the Earth; also, the branch of science that deals with the Earth's magnetism. Formerly called terrestrial magnetism, geomagnetism involves any topic pertaining to the magnetic field observed near the Earth's surface, within the Earth, and extending upward to the magnetospheric boundary. Modern usage of the term is generally confined to historically recorded observations to distinguish it from the sciences of archeomagnetism and paleomagnetism, which deal with the ancient magnetic field frozen respectively in archeological artifacts and geologic structures. *See* PALEOMAGNETISM; ROCK MAGNETISM.

The primary component of the magnetic field observed at the Earth's surface is caused by electric currents flowing in its liquid core, and is called the main field. Vectorially added to this component are the crustal field of magnetized rocks, transient variations imposed from external sources, and the field from electric currents induced in the Earth from these variations.

Main geomagnetic field. The geomagnetic field is specified at any point by its vector **F**. Its direction is that of a magnetized needle perfectly balanced before it is magnetized, and freely pivoted about that point, when in equilibrium. The north pole of such a needle is the one that at most places on the Earth takes the more northerly position. Over most of the Northern Hemisphere, that pole dips downward (**Fig. 1**). The elements used to describe the vector **F** are H, the component of the vector projected onto a horizontal plane; its north and east components X and Y, respectively; Z the vertical component; F the magnitude of the vector **F**; the angles I, the dip of the field vector below the horizontal; and D the magnetic declination or deviation of the compass from geographic north. By convention, Z and I are positive downward, and D is positive eastward (or may be indicated as east or west of north). These elements can be related to each other by trigonometric equations. *See* MAGNETIC COMPASS.

However, when deriving and using accurate models of the field, a spherical coordinate system from

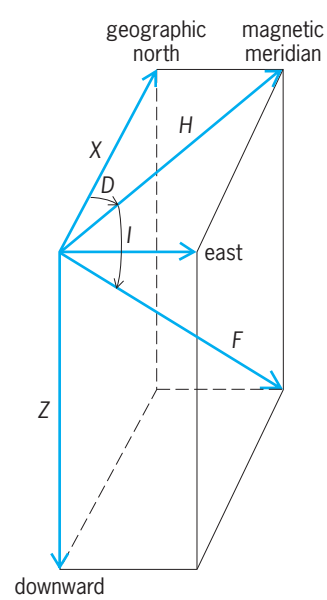

Fig. 1. Elements of the geomagnetic field. D = declination, I = inclination, H = horizontal intensity, X = north intensity, Y = east intensity, Z = vertical intensity, F = total intensity.

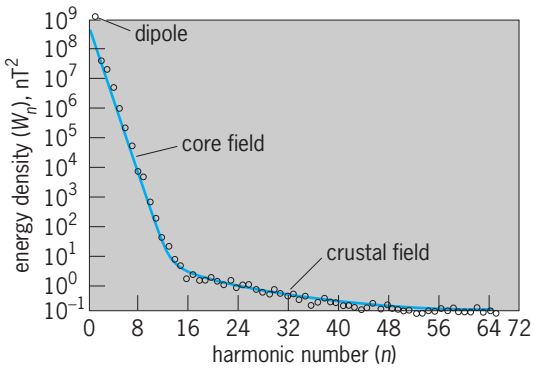

Fig. 2. Energy density spectrum of the geomagnetic field from an analysis of *Magsat* data at an average altitude of 260 mi (420 km).

the center of the Earth is used. The usual spherical coordinates are r, the distance from the Earth's center; θ, the geocentric colatitude; and ϕ, the east longitude. The components of field in this system are F_r, F_θ, and F_ϕ, which are related to those in the topocentric system by Eqs. (1)–(3), where δ is a positive angle

$$X = -F_\theta \cos\delta - F_r \sin\delta \tag{1}$$

$$Y = F_\phi \tag{2}$$

$$Z = -F_r \cos\delta + F_\theta \sin\delta \tag{3}$$

under 0.2°. It can be computed from trigonometric relations by using the accepted oblate spheroid for the Earth. *See* GEODESY.

It is possible to use a scalar potential function because very little current flows in the region between the Earth and the ionosphere, which begins at about 56 mi (90 km) altitude. However, currents above these heights do influence satellite observations, and their determination and removal from such data are the subject of current research. FORTRAN codes utilizing such coefficients to compute field are in general use. *See* MAGNETIC MONOPOLES; SPHERICAL HARMONICS.

A spectrum of the energy density of the field, as observed by the satellite *Magsat* at an average altitude of 260 mi (420 km) (**Fig. 2**), is computed by summing the squares of the coefficients for each n at a given value of r [Eq. (4)]. The spectrum reveals

$$W_n = (n+1)\left(\frac{a}{r}\right)^{2(n+2)} \sum_{m=0}^{n} \left[\left(g_n^m\right)^2 + \left(h_n^m\right)^2\right] \tag{4}$$

the presence of two different sources of the field: one in the core and one in the Earth's crust where the temperature is below the Curie point. This crustal magnetization can be completely due to the present field, or it can also have a remanent component of the ancient field frozen in the rocks. Near the Earth's surface the core field dominates up to about $n = 14$. Beyond $n = 14$ the crustal source is the more important. *See* CURIE TEMPERATURE.

Projection of the core component downward using the term

$$\left(\frac{a}{r}\right)^{2(n+2)}$$

shows a flat or so-called white spectrum just under the seismically determined core–mantle boundary, some 1800 mi (2900 km) below the surface. The crustal spectrum becomes flat no more than a few kilometers below the surface, as would be expected for sources at such a depth.

Magnetic poles. A magnetic pole is defined as a location where the field is vertically aligned, $H = 0$. Due to the presence of sometimes strong (for example, >1000 nT) magnetic anomalies at the Earth's surface, there are a number of locations where the field is locally vertical. However, those field components that extend to sufficient altitude to control charged particles can be accurately located by using the computations from a spherical harmonic expansion using degrees up to only about $n = 10$. Indeed, a pole can be defined by using only the main dipole ($n = 1$), or many terms. *See* AURORA.

The $n = 1$ poles are sometimes referred to as the geomagnetic poles, and those computed using higher terms as dip poles. The term geomagnetic could also refer to the eccentric geomagnetic pole, which can be computed from $n = 1$ and $n = 2$ harmonics so as to be the best representation of a dipole offset from the center of the Earth. The latter has been used as a simplified field model at distances of 3 or 4 earth radii. Due to the more rapid fall-off of the higher terms with distance from the Earth, the two principal poles approach those of the $n = 1$ term with increasing altitude, until the distortions due to external effects begin to predominate. *See* MAGNETOSPHERE.

Whereas for 1990–1995 the geomagnetic poles were located at 79°N and 71°W (and the corresponding southern antipodal points), the addition of higher-degree terms estimates the dip poles to be at 79°N, 105°W and 65°S, 138°E. Transient variations may change these positions up to several tens of kilometers over the course of a day, and more during times of magnetic disturbance.

A number of magnetic coordinate systems depend on the main magnetic field. The invariant latitude is based on the physics of trapped particles and used to organize data from the radiation belts. Corrected magnetic coordinates, based on the first few terms of the spherical harmonic expansion, are used for observations of conjugate geophysical phenomena such as aurora. Conjugate phenomena are those that occur at both ends of a magnetic field line extending from one hemisphere to another. An adjunct to such systems that becomes especially important in polar regions is a concept known as magnetic time, which is based on the $n = 1$ poles. Magnetic noon occurs when the point of observation lies between the $n = 1$ pole and the Sun, and magnetic midnight, when the pole lies between the observer and the Sun. More sophisticated transformations also use some of the lower degrees of n in defining magnetic time.

The distribution of the dip angle I over the Earth's surface can be indicated on a globe or map by contours called isoclines, along which I is constant. The isocline for which $I = 0$ (where a balanced magnetized needle rests horizontal) is called the dip

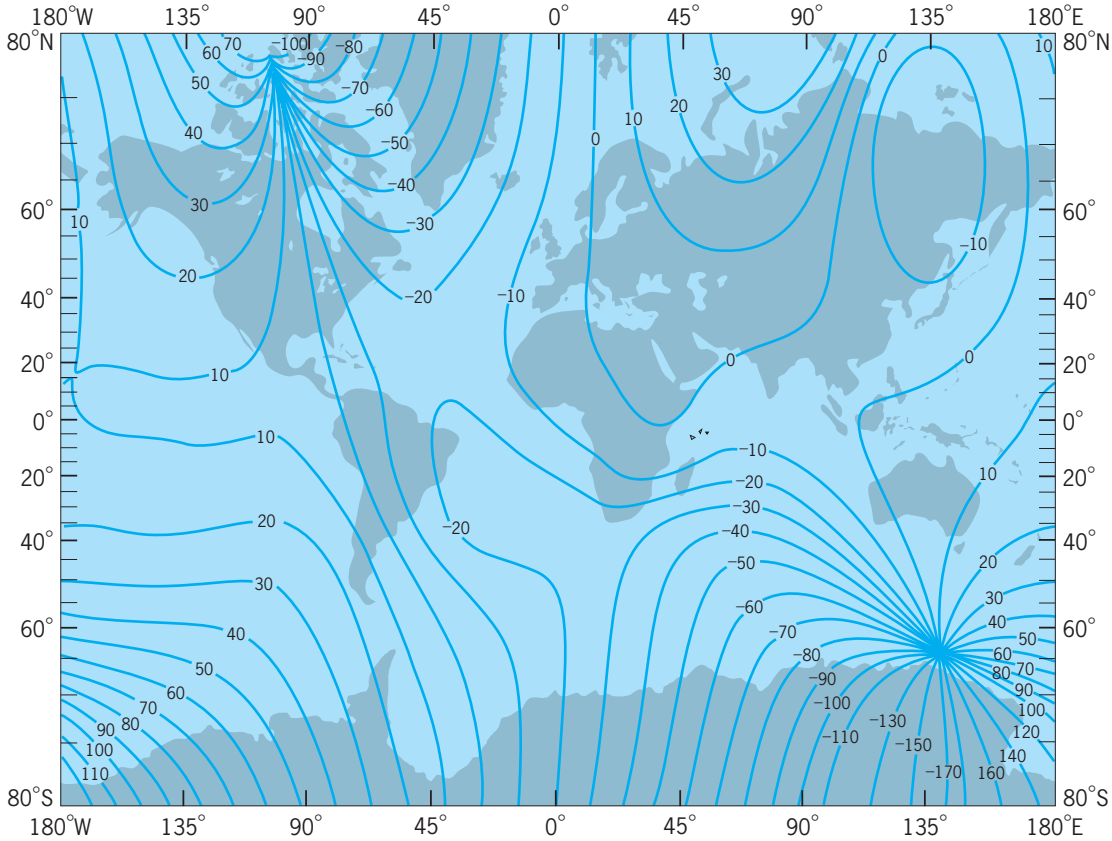

Fig. 3. Lines of equal geomagnetic declination in degrees for current epochs. (*U.S. Naval Oceanographic Office and British Geological Survey*)

equator. The dip equator is geophysically important because there is a region in the ionospheric E layer in which small electric fields can produce a large electric current called the equatorial electrojet. *See* GEOMAGNETIC VARIATIONS; IONOSPHERE.

Magnetic declination. A magnetized compass needle can be weighted to rest and move in a horizontal plane at the latitudes for which it is designed, thus measuring the declination D. The lines on the Earth's surface along which D is constant are called isogonic lines or isogones (**Fig. 3**). The compass points true geographic north on the agonic lines where $D = 0$. At nonpolar latitudes, D is a useful tool for marine and aircraft navigational reference. Indeed, isogones appear on navigation charts, electronic navigational aids are referenced to D, and airport runways are marked with D/10. A runway painted with the number 11 indicates that its direction has a compass heading of 110°. The compass needle becomes less reliable in polar regions because the horizontal component H becomes smaller as the magnetic poles are approached. *See* NAVIGATION.

Intensity patterns. The intensity of the field can also be represented by maps, and the lines of equal intensity are called isodynamic lines (**Fig. 4**). The dipole dominates the patterns of magnetic intensity on Earth in that the intensity is about double at the two poles compared to the value near the Equator. However, it can also be seen that the next terms of the spherical harmonic expansion also have a significant effect, in that there is a second maximum in Siberia, and an area near Brazil that is weaker than any other. This so-called Brazilian anomaly allows charged particles trapped in the magnetic field to reach a low altitude and be lost by collisions with atmospheric gases. The highest intensity of this smooth field is about 70 microteslas near the south magnetic pole in Antarctica, and the weakest is about 23 μT near the coast of Brazil.

Magnetic anomalies. The term anomaly has become clearer than it was previously because it is recognized that the geomagnetic field has a continuous spectrum (Fig. 2) but with two distinct contributors. Originally, the term meant a field pattern that was very local in extent; the modern definition is that portion of the field whose origin is the Earth's crust. The sizes of the strong and easily observable features are generally up to only a few tens of kilometers. Their intensity ranges typically from a few hundred nanoteslas up to several thousand, and they are highly variable depending on the geology of the region.

The former usage has continued in many practical applications where magnetic anomaly maps are constructed from a smooth contouring of aeromagnetic data, ignoring the larger scale but weaker anomalies. Such maps are derived from data taken on a number of flight lines after reduction, using an internationally agreed-upon reference model that is supposed to represent the noncrustal or core component. Such models have been termed IGRF for a

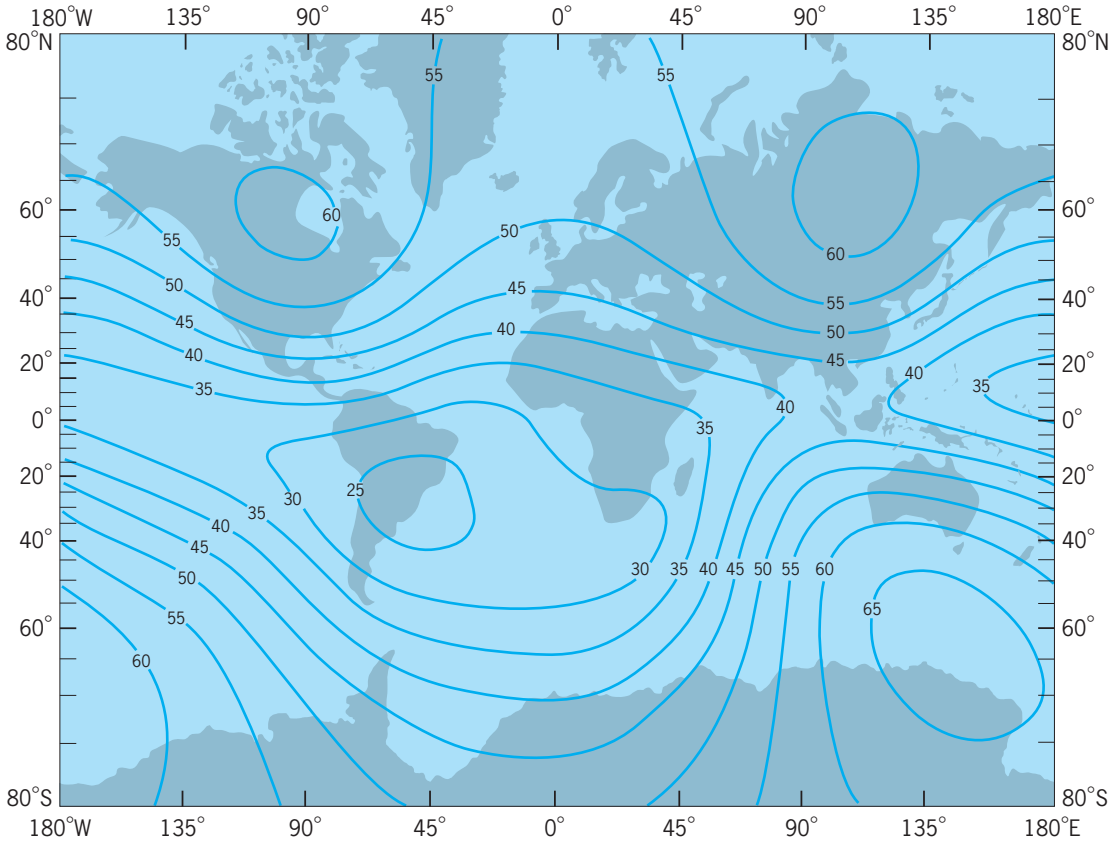

Fig. 4. Lines of equal magnetic intensity *F* in microteslas for recent epochs.

predicted International Geomagnetic Reference Field or DGRF where the attempt is made to correct past predictions and obtain a "definitive" set of spherical harmonic coefficients. Generally, the year or epoch of validity of such models accompanies the term. Thus, IGRF 1995 represents the best prediction of what the field will be at epoch 1995.0. To the present time, the IGRFs have been limited to $n = 8$ because of the difficulty of prediction of more detail with inadequate data, and the DGRFs to $n = 10$. The core field dominates that of the crust up to about $n = 14$ (Fig. 2), so that such models fall short of their goal. Neglect of the terms above $n = 10$ gives errors of the order of a few tens of nanoteslas in estimating the core contribution. Estimation of the core component is complicated by the presence in the lower-degree terms of contributions from crustal magnetization that are comparable in intensity to those seen above $n = 14$.

The best reduction of survey data that can be done is to remove the core component by subtracting a field computed by using the spherical harmonic models limited to $n = 14$ or 15. Formerly, it was thought that the regional or background field could be eliminated by treating data with a high-pass filter. Such filters can be devised to pass data with wavelengths shorter than some preset amount. However, such filtering also removes the long-wavelength components that exist even in the high-degree crustal field contribution. The background field removed with spherical harmonic of degree n has a scale size equal to $40{,}000/n$ (km).

With the advent of satellite surveys, the weaker, larger-size anomalies have become more visible and can be incorporated into spherical harmonic models. Those for Europe using harmonic degrees $n = 15$ through 50 were derived from the 1979–1980 survey by *Magsat* (**Fig. 5**). The strongest anomaly was found near Kursk, Soviet Union. Surface surveys of this region reveal two narrow strips of anomaly 36 mi (60 km) apart running from the northeast to the southwest. The most disturbed part of the major (northerly) strip is only 1.2 mi (2 km) wide, although the strip is 150 mi (250 km) long; Z is everywhere above normal and ranges up to 190 μT.

Exploration geophysicists work on a much smaller scale and are more concerned with removing the spatial gradients of the observations than with reduction to an absolute level. The practice is to first reduce the contoured residuals of their data; that is, the observed field reduced by that computed from the IGRF, followed by a local two-dimensional fit to further eliminate an average slope across the maps. Such reductions depict the very local anomalies resulting from near-surface mineral deposits, but not the larger-scale features. Maps so constructed generally have discontinuities at their borders when an attempt is made to combine them with others from adjoining regions.

An isolated anomaly such as that at Kursk is

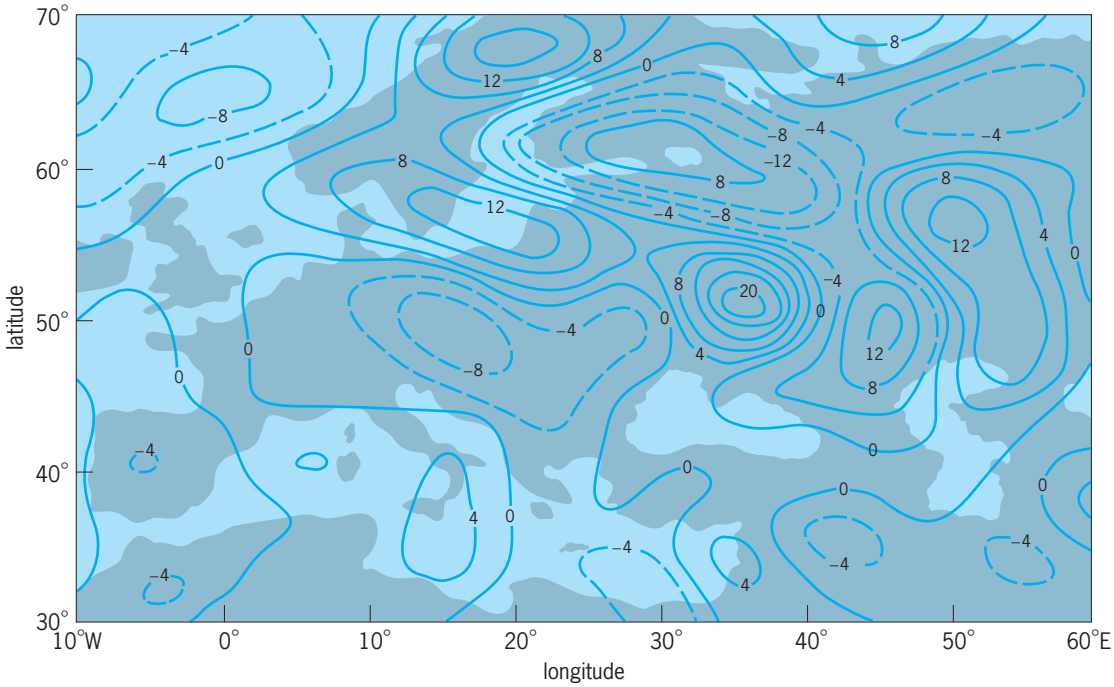

Fig. 5. Crustal anomaly field of the *Z* component as determined from a *Magsat* model ($n = 15$–50) over Europe for an altitude of 217 mi (350 km). Contours are in nanoteslas.

thought to be due to igneous intrusions. The other, broader patterns are assumed to be from differences of composition in the magnetic basement or crystalline rocks underlying sedimentary deposits, but still not so deep as to be hotter than the Curie point. The differences between continental and oceanic crusts are not clearly evident, nor do the patterns of anomalies about spreading centers show any obvious signatures at satellite altitudes. *See* DIPOLE; DIPOLE MOMENT; GEOPHYSICAL EXPLORATION; MAGNETITE; PLATE TECTONICS.

Secular magnetic variation. The main or core component of the geomagnetic field undergoes slow changes that necessitate continual adjustment of the model coefficients and redrawing of the isomagnetic maps. In any magnetic element at a particular place, the variation may be an increase or a decrease and is not constant in either magnitude or sign. This distribution of the rate for any element can be indicated on isoporic maps by lines (isopors) along which the rate is constant (**Fig. 6**). Typically, the pattern of isopors is more complex than that of the isomagnetic lines for the same element, partly because the spectrum of such change is not dominated by the dipole as is the case of the static field. There has been no accurate satellite monitoring since 1980. The greatest uncertainty in maps of secular magnetic variation are for the southern ocean areas where satellite data are inadequate to overcome the gaps in the distribution of surface magnetic observatories.

Studies indicate that the dipole component of the field 2000 years ago was about 50% stronger than the present. Its average decay rate has averaged about 0.05% per year (15 nanoteslas per year) since about 1840 when absolute measurements were first begun, but accelerated from 1970 to its 1994 value of 0.08% per year (24 nT/yr). However, there is also evidence that the decade of the 1940s showed a rate of only about 10 nT per year. A linear projection of the present rate would have the dipole decreasing to zero in less than 1500 years. Although archeomagnetic evidence indicates that the field has indeed decayed to near zero level within the last 50,000 years with a subsequent return to the present polarity, and paleomagnetic results show that the field has reversed its polarity many times since the Earth's formation (the last time, about a million years ago), there is no model that can predict the future course of field change.

One other smooth variation that has been detected in secular change is the apparent westward drift of some of the patterns of change. Isopors of the vertical field are seen to move preferentially to the west at about 0.2° per year. However, many changes are unpredictable. For example, what has been termed the magnetic jerk was first applied to a discontinuity noted in 1969 (**Fig. 7**). The term jerk is taken from the mechanical analog whereby the acceleration on a moving body is suddenly changed. Between the times of such changes, the secular variations appear to be smooth. Similar changes have probably gone unobserved because of the sparseness of the recorded data. The last irregularity occurred in 1983, though it has not yet been adequately analyzed.

Theory of core field and secular variation. Deriving a suitable model that explains the source of the Earth's magnetic field has been one of the most frustrating problems that theoreticians have faced. Starting with the physical laws that should govern the behavior of a highly conducting, rotating, spherical

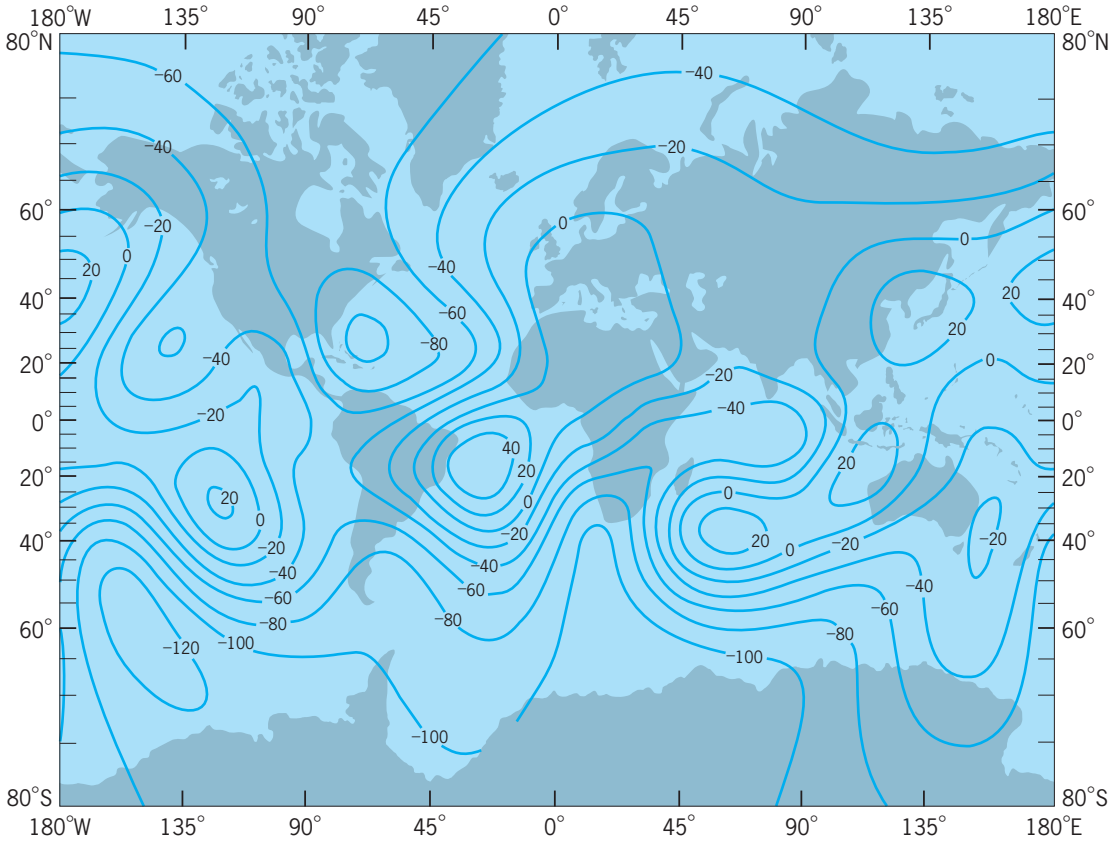

Fig. 6. Contours of equal annual rate of change (isopors) of the total intensity of the field in nanoteslas per year estimated for 1990. (*U.S. Naval Oceanographic Office and British Geological Survey*)

fluid and coming up with a model of the geomagnetic field is exceedingly difficult. The fact that the planets Mercury, Jupiter, Saturn, Uranus, and Neptune also have dipolelike fields indicates that it is a common process, most likely a hydromagnetic dynamo. Dynamo means that a current is generated as an electrical conductor is moved through a magnetic field, as in a dynamo supplying electrical power. *See* GEODYNAMO; JUPITER; MERCURY (PLANET); NEPTUNE; SATURN; URANUS.

Magnetic surveys and models. The main source of data for magnetic maps and models before the advent of satellites was fixed magnetic observatories. These stations, numbering about 140, provided the continuous record of changes in the magnetic field at their location. Their data are generally accurate and an excellent indicator of both secular change and the transient variations, but their global coverage is too sparse for a determination of the whole field. Spherical harmonic analyses based only on such data produce distorted results because of the large gaps in coverage, especially because of the sparseness of observing locations in southern oceanic regions.

An international effort known as the World Magnetic Survey (WMS) was initiated by the International Association of Geomagnetism and Aeronomy (IAGA) during the 1960s to intensify surface and aeromagnetic surveys. A panel was established under IAGA to provide standard magnetic models for public use, and to derive the first such model, IGRF 1965. This project also resulted in a bilateral agreement between the United States and the Soviet Union for the first sharing of satellite-derived data. The spacecraft that were part of this agreement were the POGO half of the OGO series (*Polar Orbiting Geophysical Observatory*; United States) and *Cosmos 49* (Soviet Union; see **table**).

Field models were initially constructed by using only the very accurate scalar data (errors less than 1 nT) from the OGO satellites series. However, it was found that such models were subject to a weakness known as the Backus ambiguity, whereby the vertical component (Z) of the field near the magnetic dip equator was poorly determined. It was thus

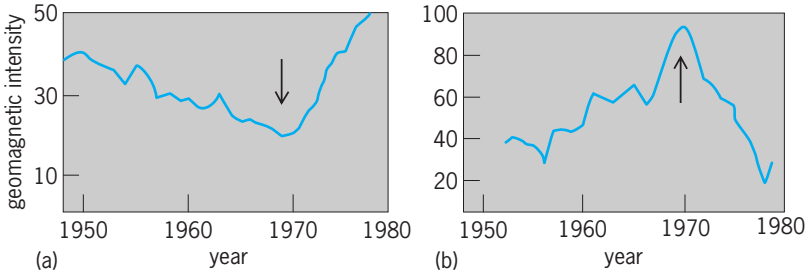

Fig. 7. Secular variation of the geomagnetic intensities (*a*) from the Hartland Magnetic Observatory in Great Britain and (*b*) from the Fredericksburg Observatory in the United States. The arrows point to the time of the jerk or abrupt chang of trend in late 1969 to early 1970. (*After V. Cortillot and J. L. LeMouël, Geomagnetic secular variations impulses, Nature, 311:709–716, 1984*)

Low-altitude satellite geomagnetic measurements, 1958–1993

Spacecraft	Orbit: Inclination, degrees	Orbit: Altitude range, mi (km)	Interval	Instruments	Coverage
Sputnik 3	65	260–360 (440–600)	May–June 1958	Flux gates	Soviet Union
Vanguard 3	33	310–2250 (510–3750)	Sept.–Dec. 1959	Proton precession	Near ground stations
Kosmos 26	49	160–2421 (270–403)	Mar. 1964	Proton precession	Whole orbit
Kosmos 49	50	157–293 (261–488)	Oct.–Nov. 1964	Proton precession	Whole orbit
1964–83C	90	620–653 (1040–1089)	Dec. 1964–June 1965	Rubidium vapor	Near ground stations
OGO 2	87	248–906 (413–1510)	Oct. 1965–Sept. 1967	Rubidium vapor	Whole orbit
OGO 4	86	247–545 (412–908)	July 1967–Jan. 1969	Rubidium vapor	Whole orbit
OGO 6	82	238–659 (397–1098)	June 1969–June 1971	Rubidium vapor	Whole orbit
Kosmos 321	71	162–402 (270–403)	Jan.–Mar. 1970	Cesium vapor	Whole orbit
Magsat	97	211–347 (352–578)*	Nov. 1979–June 1993	Cesium/flux gates	Whole orbit
Dynamic Explorer 2	90	190–630 (309–1012)	Aug. 1981–Feb. 1983	Flux gates	
POGS	90	Circular 470 (750)	Jan. 1991–Sept. 1993	Flux gates	
UARS	57	Circular 360 (585)	Sept. 1991–1994	Flux gates	
Freja	60	370–960 (600–1550)	Oct. 1992–1994	Flux gates	

*The altitude decayed throughout the interval. Some data were taken as low as 110 mi (190 km).

necessary to combine surface vector data with those from satellites. The coefficients of the models were expanded into series, such as $g(t) = g_o + \dot{g}(t - t_o)$, where the secular change terms $\dot{g}$ were determined primarily by the observatory data.

The *Magsat* survey (see table) was the first spacecraft attempt to obtain accurate vector data. The problem in such projects is to determine the spacecraft orientation, or attitude, to sufficient accuracy. Whereas optical pumping magnetometers (such as used on the OGO spacecraft) are accurate to better than 1 nT, this corresponds in a 50,000-nT field to an angular accuracy of 1/50,000 radians, or about 4 seconds of arc. *Magsat* achieved a nominal 20 seconds accuracy during its short data collection interval. The field models for 1980 were determined accurately up to about $n = 50$, beyond which the coefficients reached the noise level (Fig. 2).

Since 1980, there has not been a spacecraft designed to accurately follow secular change or to improve the crustal field definition pioneered by the Kosmos, OGO, and *Magsat* surveys. The satellites after 1980 (see table) make field observations with scalar errors greater than 20 nT and have little knowledge of the attitude. Field models that can be derived with post-1980 surface and satellite data can only bound the errors and provide models useful for applications where high accuracy is not required. *See* MAGNETISM.

Joseph Cain

Bibliography. G. Backus, R. Parker, and C. Constable, *Foundations of Geomagnetism*, 2005; W. H. Campbell, *Introduction to Geomagnetic Fields*, 2d ed., 2003; J. A. Jacobs (ed.), *Geomagnetism*, 4 vols., 1987–1991; J. A. Jacobs, *Reversals of the Earth's Magnetic Field*, 2d ed., 1993; F. J. Lowes et al., *Geomagnetism and Paleomagnetism*, 1988.

Geometric phase

A unifying mathematical concept that describes the relation between the history of internal states of a system and the system's resulting orientation in space. Under various aspects, this concept occurs in geometry, astronomy, classical mechanics, and quantum theory. In geometry it is known as holonomy. In quantum theory it is known as Berry's phase, after M. Berry, who isolated the concept (which was already known in special cases) and explained its wide-ranging significance.

Examples. A few examples of phenomena in which geometric phases appear are discussed below. The first concerns a man sitting on a platform, free to rotate about a perpendicular axis, that is, a soda-fountain stool. The problem is how can he rotate himself without touching anything external or shifting his position on the stool. The solution does not seem so difficult. He might, for example, extend his arm and rotate it to the right. Then he and the platform will rotate together by a smaller amount to the left. This procedure does not yet provide a fully satisfactory solution, however, because he remains with his arm sticking out. He would like a solution such that he returns to his initial posture. Simply to reverse the motion will not do, of course, because during the reverse motion the rotation of his body and the platform will also be canceled. The solution involves pulling the arm back close to his body before undoing its rotation. At the end of this cycle, he and the platform will have rotated together through a nonzero angle, and he will have returned to his initial posture.

A more elaborate form of the same problem concerns an astronaut initially at rest, that is, not rotating in space. The astronaut has the problem of reorienting herself. By simple maneuvers of the type just discussed, extending, rotating, and retracting the arm in different directions, the astronaut can indeed point herself in any desired direction, while returning to her original posture.

The same problem, essentially, must be solved by a cat dropped from rest in an arbitrary orientation. It must perform such contortions to reorient itself and land feet first.

A person's ability to control his orientation contrasts profoundly with his inability to perform a

displacement in space. It is a theorem of mechanics that the center of mass of an isolated body moves at a constant velocity, independently of whatever machinations the body might perform.

In the examples discussed so far, a surprising nonzero net rotation or reorientation is found after a cycle involving no net change in configuration, for a system basically at rest. A more complex situation, important for several of the more interesting applications of the geometric phase, occurs when a nonzero rate of rotation is expected, but this rate is manipulated by cyclic changes of configuration.

A classic realization of this situation concerns a spinning particle (henceforth termed simply a spin) interacting with a magnetic field whose magnitude is constant but whose direction can vary. If the direction of the magnetic field were fixed, then the spin would precess around the magnetic field direction, say at a rate ω. Thus after time t, the spin will have precessed through the angle $\theta_{\text{static}} = \omega t$. Now, if the direction of the field is slowly varied through a cycle before it is restored to its original direction, Berry showed that the total angle is instead $\theta_{\text{Berry}} = \omega t + \gamma_{\text{Berry}}$, where γ_{Berry} is half the solid angle swept out by the direction of the field. Thus, the value of γ_{Berry} is independent of almost all details of the cycle and has a purely geometric character. *See* PRECESSION; SPIN (QUANTUM MECHANICS).

Spins can be manipulated, and their directions monitored, by using the techniques of nuclear magnetic resonance. In this way, effects, such as the one just described and others more intricate, have been exhibited experimentally. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

General concept. With these examples, it is possible to gain some understanding of the general concept. A system is envisioned—stool sitter, skater, astronaut, cat, spin in magnetic field—whose possible states can be visualized as points in a suitable abstract space. At the same time, the system has some position or orientation in another space, which in all these examples (but not always) is just ordinary physical space. A history of internal states can be represented by a curve in the first space; the effect of this history on the disposition of the system, by a curve in the second space. The mapping between these two curves is described by the geometric phase. Especially interesting is the case when a closed curve (cycle) in the first space maps onto an open curve in the second, for then there is no net change in internal state, yet the disposition of the system with respect to the outside world is altered.

The power of the geometric phase ideas is that they make it possible, in complex dynamical problems, to find some simple universal regularities without having to solve the complete equations. Significant uses of these ideas include demonstrations of the fractional electric charge and quantum statistics of the quasiparticles in the quantum Hall effect, and of the occurrence of anomalies in quantum field theory. *See* ANYONS; HALL EFFECT; QUANTUM FIELD THEORY.

Frank Wilczek

Bibliography. A. Shapere and F. Wilczek (eds.), *Geometric Phases in Physics*, 1989.

Geometrical optics

The geometry of light rays and their images, through optical systems. Geometrical optics is by far the oldest model proposed for accounting for the behavior of light, going back to classical Greece. It was not until around the beginning of the nineteenth century that the wave nature of light was seriously considered, and in the modern view of the nature of light, geometrical optics as a fundamentally correct model is simply wrong. In spite of this geometrical optics is remarkably robust, remaining as a most practical tool in the solution of optical problems. It has been applied to analyzing laser resonators, to solving problems in interference and diffraction, and even to analyzing the behavior of waveguides, where at first glance it would seem to be totally inappropriate. These developments have been made possible by the generation of "fictitious" rays (all rays are fictitious) or by attributing to rays properties which cannot be accounted for in a strictly geometrical optical model. Nevertheless, the principal application of geometrical optics remains in the field of optical design, where it has been employed since the first optical instruments were developed in the early seventeenth century. This article concentrates on this application. *See* DIFFRACTION; INTERFERENCE OF WAVES; LASER; WAVEGUIDE.

Basic concepts. Light is a form of energy which flows from a source to a receiver. It consists of particles (corpuscles) called photons. All photons of a single pure color have the same energy per photon. Different colors have different energies.

For green light there are 2.5×10^{18} photons per joule of energy, or 2.5×10^{18} photons per second per watt of power. This factor increases for red light and diminishes for blue. *See* LIGHT; PHOTON.

Refractive index and dispersion. The speed with which the particles travel depends on the medium. In a vacuum this speed is $3 \times 10^8\ \text{m} \cdot \text{s}^{-1}$ ($1.86 \times 10^5\ \text{mi} \cdot \text{s}^{-1}$) for all colors. In a material medium, whether gas, liquid, or solid, light travels more slowly. Moreover, different colors travel at different rates. This change in speed is a property of the medium, and is measured by the ratio of the speed in a vacuum to the speed in the medium. This ratio is called the refractive index of the medium. The variation in refractive index with color is called dispersion. Generally speaking, blue light travels more slowly than red light, so the medium has a higher refractive index for blue than for red light. In order to deal more precisely and quantitatively with color, and therefore with refractive index and dispersion, it is customary to identify the different colors by their wavelengths, even though waves have no place in geometrical theory. *See* COLOR; DISPERSION (RADIATION); REFRACTION OF WAVES.

Rays. The paths that particles take in going from the source to the receiver are called rays. The

particles never interact with each other. The time that it takes a particle to travel from one point to another along a ray is determined by the speed at which the particle travels, which is determined by the refractive index of the medium. The product of the refractive index and the path length is called the optical path length along the ray. The optical path length is equal to the distance that the particle would have traveled in a vacuum in the same time interval.

Sources and wavefronts. A point source is an infinitesimal region of space which emits photons. An extended source is a dense array of point sources. Each point source emits photons along a family of rays associated with it. For each such family of rays there is also a family of surfaces each of which is a surface of constant transit time from the source for all the particles, or alternatively, a surface of constant optical path length from the source. These surfaces are called geometrical wavefronts (**Fig. 1**), although, again, waves have no place in geometrical theory. The reason for this name is that they are often good approximations to the wavefronts predicted by a wave theory. In isotropic media, rays are always normal to the geometrical wavefronts, and the optical path length from one geometrical wavefront to another is the same for all the rays in the family.

Ray paths. The ray path which any particle takes as it propagates is determined by Fermat's principle, which states that the ray path between any two points in space is that path along which the optical path length is stationary (usually a minimum) among all neighboring paths. In a homogeneous medium (one with a constant refractive index) the ray paths are straight lines. In a homogeneous medium containing a point source, the family of rays from the source all radiate outward and the associated geometrical wavefronts are concentric spherical shells.

In a system that consists of a sequence of separately homogeneous media with different refractive indices and with smooth boundaries between them

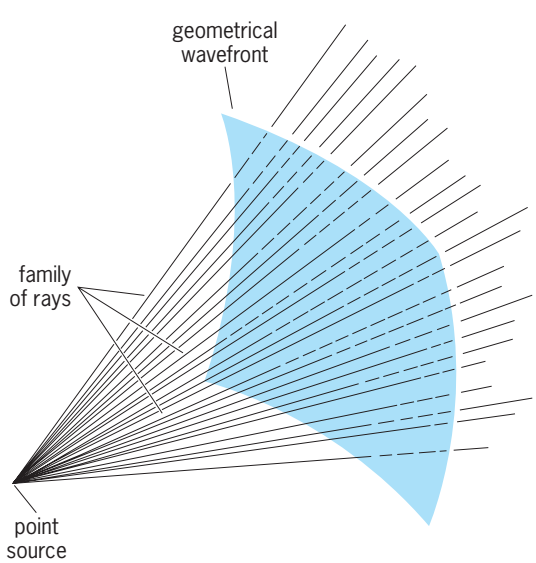

Fig. 1. Geometrical wavefront for a family of rays associated with a point source.

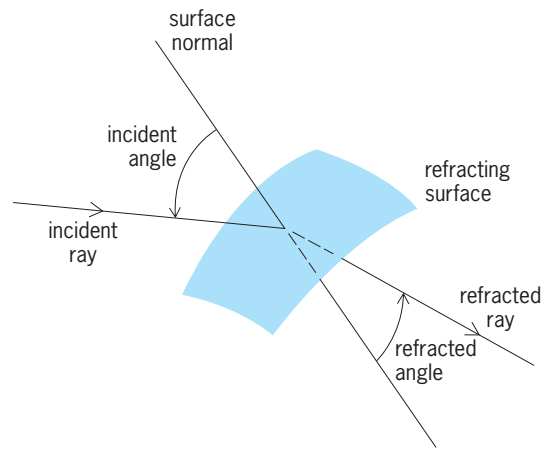

Fig. 2. Refraction of a ray at a smooth surface between two homogeneous media.

(such as a lens system), the ray paths are straight lines in each medium, but the directions of the ray paths will change in passing through a boundary surface. This change in direction is called refraction, and is governed by Snell's law, which states that the product of the refractive index and the sine of the angle between the normal to the surface and the ray is the same on both sides of a surface separating two media. The normal in question is at the point where the ray intersects the surface (**Fig. 2**).

Because this product is the same on both sides of the interface, the angle of the ray in the medium with the lower refractive index is always larger than the corresponding angle on the other side of the surface. When the sine of the larger angle becomes one, the largest value it can have, the angle itself is 90°, but the other angle is smaller. The latter angle is called the critical angle for the two media. If the light is traveling from the denser (higher-index) medium to the less dense medium, there is nothing to prevent the incidence angle from being larger than the critical angle, in which case the light is totally reflected by the surface and does not pass through to the other side. This phenomenon is called total internal reflection. *See* REFLECTION OF ELECTROMAGNETIC RADIATION.

Ideal image formation. The primary area of application of geometrical optics is in the analysis and design of image-forming systems. An optical image-forming system consists of one or more optical elements (lenses or mirrors) which when directed at a luminous (light-emitting) object will produce a spatial distribution of the light emerging from it which more or less resembles the object. The latter is called an image. *See* OPTICAL IMAGE.

In order to judge the performance of the system, it is first necessary to have a clear idea of what constitutes ideal behavior. Departures from this ideal behavior are called aberrations, and the purpose of optical design is to produce a system in which the aberrations are small enough to be tolerable. *See* ABERRATION (OPTICS).

First consider the requirements for ideal behavior. In an ideal optical system the rays from every point in the object space pass through the system so that they

converge to or diverge from a corresponding point in the image space. This corresponding point is the image of the object point, and the two are said to be conjugate to each other (object and image functions are interchangeable). Another way of expressing the same thing is that an ideal optical system converts every spherical geometrical wavefront in the object space into a spherical geometrical wavefront in the image space with the image point located at its center of curvature.

It is not enough, however, to have a system which forms a perfect point image for every point object. The image points must be in the proper geometrical relationship to constitute a good image.

Since there is a one-to-one correspondence between the conjugate object and image points, the geometry of the object and image spaces must be connected by some mapping transformation. The one generally used to represent ideal behavior is the collinear transformation. The intrinsic properties of the collinear transformation are as follows. If three object points lie on the same straight line, they are said to be collinear. If the corresponding three image points are also collinear, and if this relationship is true for all sets of three conjugate pairs of points, then the two spaces are connected by a collinear transformation. In this case, not only are points conjugate to points, but straight lines and planes are conjugate to corresponding straight lines and planes.

It is attractive to have an ideal behavior in which for every point, straight line, or plane in the object space there is one and only one corresponding point, straight line, or plane in the image space. This does not, however, guarantee that the three-dimensional mapping is distortion-free.

Another feature usually incorporated in the ideal behavior is the assumption that all refracting or reflecting surfaces in the system are figures of revolution about a common axis, and this axis of symmetry applies to the object-image mapping as well. With this axial symmetry, an object line which coincides with the axis has as its conjugate an image line also coinciding with the axis. Therefore every object plane containing the axis, called a meridional plane, has a conjugate which is a meridional plane coinciding with the object plane, and every line in the object-space meridional plane has its conjugate in the same plane. In addition, every object plane perpendicular to the axis must have a conjugate image plane which is also perpendicular to the axis, because of axial symmetry. In the discussion below, the terms object plane and image plane refer to planes perpendicular to the axis unless otherwise modified.

An object line parallel to the axis will have a conjugate line which either intersects the axis in image space or is parallel to it. The first case is called a focal system, and the second an afocal system.

Focal systems. The point of intersection of the image-space conjugate line of a focal system with the axis is called the rear focal point (**Fig. 3**). It is conjugate to an object point on axis at infinity. The image plane passing through the rear focal point is the rear focal plane, and it is conjugate to an object plane at infinity. Every other object plane has a conjugate located at a finite distance from the rear focal point, except for one which will have its image at infinity. This object plane is the front focal plane, and its intersection with the axis is the front focal point. Every object line passing through the front focal point will have its conjugate parallel to the axis in image space.

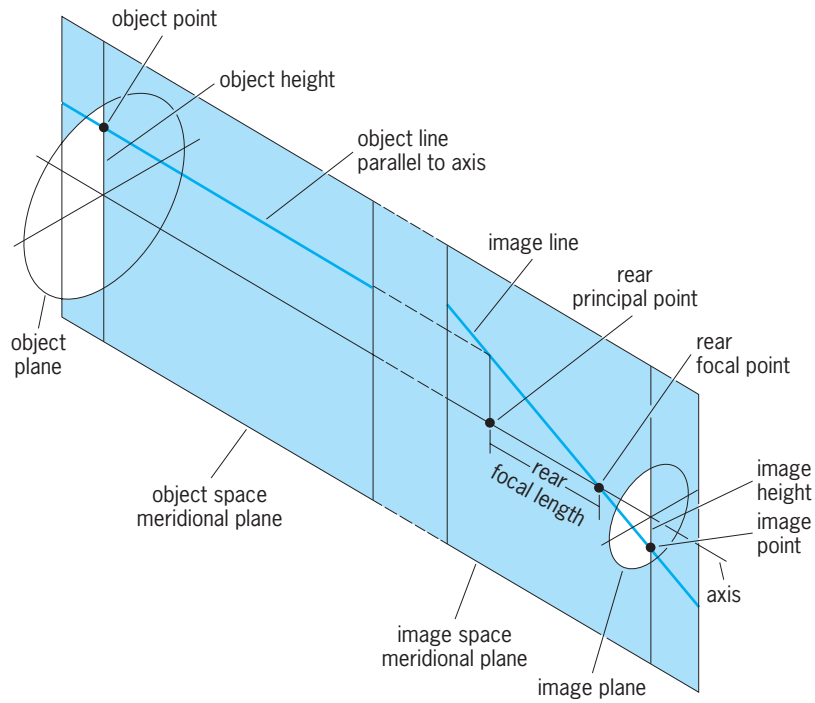

Fig. 3. Focal system.

Now take an arbitrary object plane and its conjugate image plane. Select a point off axis in the object plane and construct a line parallel to the axis passing through the off-axis point. The conjugate line in image space will intersect the axis at the rear focal point and the image plane at some off-axis point. The distance of the object point from the axis is called the object height, and the corresponding distance for the image point is the image height. The ratio of the image height to the object height is called the transverse magnification, and is positive or negative according to whether the image point is on the same or the opposite side of the axis relative to the object point.

Every pair of conjugate points has associated with it a unique transverse magnification, and a given transverse magnification specifies a unique pair of conjugate planes. The rear focal plane and its conjugate have a transverse magnification of zero, whereas the front focal plane and its conjugate have an infinite transverse magnification.

The conjugate pair which have a transverse magnification of $+1$ are called the (front and rear) principal planes. The intersections of the principal plane with the axis are called the principal points. The distance from the rear principal point to the rear focal point is called the rear focal length, and likewise for the front focal length. *See* FOCAL LENGTH.

The focal points and principal points are four of the six gaussian cardinal points. The remaining two are a conjugate pair, also on axis, called the nodal

points. They are distinguished by the fact that any conjugate pair of lines passing through them make equal angles with the axis. The function of the cardinal points and their associated planes is to simplify the mapping of the object space into the image space.

In addition to the transverse magnification, the concept of longitudinal magnification is useful. If two planes are separated axially, their conjugate planes are also separated axially. The longitudinal magnification is defined as the ratio of the image plane separation to the object plane separation, and is proportional to the product of the transverse magnifications of the two conjugate pairs. In the limit of the separation between the pairs approaching zero, the longitudinal magnification becomes proportional to the square of the transverse magnification.

Only in planes perpendicular to the axis is there an undistorted mapping, because only in such planes is the magnification a constant over the field. An object plane not perpendicular to the axis has a conjugate plane also inclined to the axis, and the magnification varies over the field, resulting in keystone distortion.

Afocal systems. In the case of afocal systems, any line parallel to the axis in object space has a conjugate which is also parallel to the axis (**Fig. 4**). Conjugate planes perpendicular to the axis are still uniquely related, but the transverse magnification is constant for the system, and the same value is applied to every pair of conjugate planes. The longitudinal magnification, also constant for the system, is proportional to the square of the transverse magnification regardless of the separation of the pair of conjugate planes. Cardinal points do not exist for afocal systems.

The most common use of an afocal system is as a telescope, where both the object and the image are at infinity. The apparent sizes of the object and image are determined by the angular subtense, which is defined, for finite object or image distance, as the ratio of the height to the distance from the observer, and, for infinite distance, by holding this ratio constant as the distance approaches infinity. The concept of angular magnification is therefore useful; this is defined as the ratio of the transverse magnification to the longitudinal magnification, and is inversely proportional to the transverse magnification. The power of a telescope or a pair of binoculars is the magnitude of the angular magnification. *See* MAGNIFICATION; TELESCOPE.

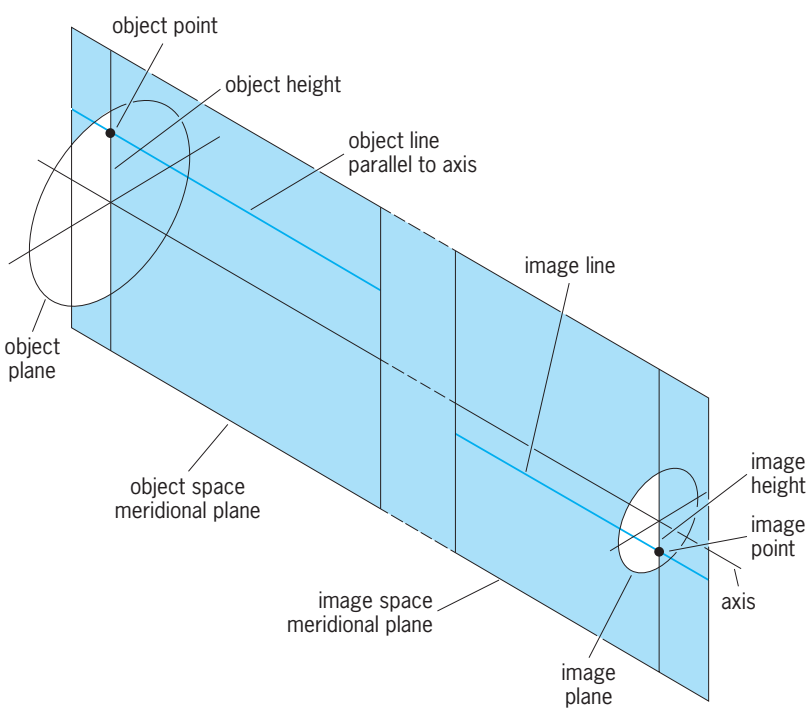

Fig. 4. Afocal system.

Paraxial optics. The above discussion of the properties of the ideal image-forming system did not consider the optical properties of the system where the ray paths are determined by Snell's law. Real optical systems do not, in fact cannot, obey the laws of the collinear transformation, and departures from this ideal behavior are identified as aberrations. However, if the system is examined in a region restricted to the neighborhood of the axis, the so-called paraxial region, where angles and their sines are indistinguishable from their tangents, a behavior is found which is exactly congruent with the collinear transformation. Paraxial ray tracing can therefore be used to determine the ideal collinear properties of the system. If a more extended region than the paraxial is considered, departures from the ideal collinear behavior are observed because additional terms in a series expansion for the trigonometric functions must be taken into account. Nevertheless, even in this extended region the paraxially determined quantities represent the first-order behavior of the system.

Ray tracing is usually done by using a cyclic algorithm, with the coordinate system being transferred from surface to surface through the system as the ray tracing progresses. Two operations are involved, a refraction and a transfer. In refraction, given the incident ray direction and position on the surface, Snell's law is used to determine the emergent direction of the ray. In transfer, the position of the ray on the next surface is determined.

In paraxial ray tracing, which uses the paraxial approximation to Snell's law, the operations of refraction and transfer are simple linear operations, and only two rays need be traced in order to locate the cardinal points of the system. Once the cardinal points have been determined, there are simple equations available to locate any pair of conjugate planes and their associated magnification.

Apertures. The above discussion does not take into account the fact that the sizes of the elements of the optical system are finite, and, in fact, for any optical system, the light that can get through the system to form the image is limited. To determine how much light can get through the system, consider the tracing of a sequence of rays from the axial object point meeting the first surface at increasingly greater heights. Eventually one of these rays and all others beyond it will be blocked by the edge of some aperture in the system. Assuming the aperture is circular and centered on the axis, all the object rays inside

the cone determined by the ray which just touches the edge of the aperture will get through the system and participate in the formation of the image. The aperture which limits the cone of rays admitted is called the aperture stop of the system.

An observer who looks into the front of the system from the axial object point sees not the aperture stop itself (unless it is in front of the system), but the image of it formed by the elements preceding it. This image of the aperture stop is called the entrance pupil, and is situated in the object space on the system. The image of the aperture stop formed by the rear elements is the exit pupil, and is situated in the image space of the system. The beam of light which would go from the object point to the image point on axis is spindle-shaped, with a circular cross section everywhere (**Fig. 5***a*).

Light from the edge of the object field will also be limited by the same aperture stop and, to first order, will have the same circular cross section at each surface as the axial beam but will be displaced laterally except at the aperture stop and the pupils. The beam will consist of a sequence of sections of skew circular cones (Fig. 5*b*). The connection between the entrance pupil and the object-space segments of the axial and the off-axis beams is shown in Fig. 5*c*.

It is also possible that a portion of the beam will be further blocked by the edge of some element which, although large enough to admit the axial beam, will not be large enough to admit all of the off-axis beam. This latter effect is called vignetting.

Marginal ray and chief ray. The characteristics of the axial beam are completely specified by a ray traced from the axial object point toward the edge of the entrance pupil. This ray will graze the edge of the aperture stop and emerge from the edge of the exit pupil proceeding to the axial image point. Such a ray is often called a marginal ray, although other names are also common.

The characteristics of the beam from the edge of the field are the same as those of the axial beam except for the displacement of the center of the beam at each surface. Thus the characteristics of this off-axis beam are described by the marginal ray, determining the cross-sectional radius, and a ray traced from the edge of the object field toward the center of the entrance pupil. This ray will pass through the center of the aperture stop and emerge from the center of the exit pupil proceeding to the image point at the edge of the image field. This ray is commonly called the chief ray.

These two rays, the marginal ray and the chief ray, are all that need be traced to determine the first-order properties of the image-forming system.

Lagrange invariant. At any surface in the system the marginal ray and the chief ray are determined by their heights at the surface and angles they make with respect to the axis. Moreover, they are connected by an interesting invariant relationship, often called the Lagrange invariant. This is obtained by multiplying the height of each ray by the product of the refractive index and the angle of the opposite ray, and taking the difference between the resulting two products.

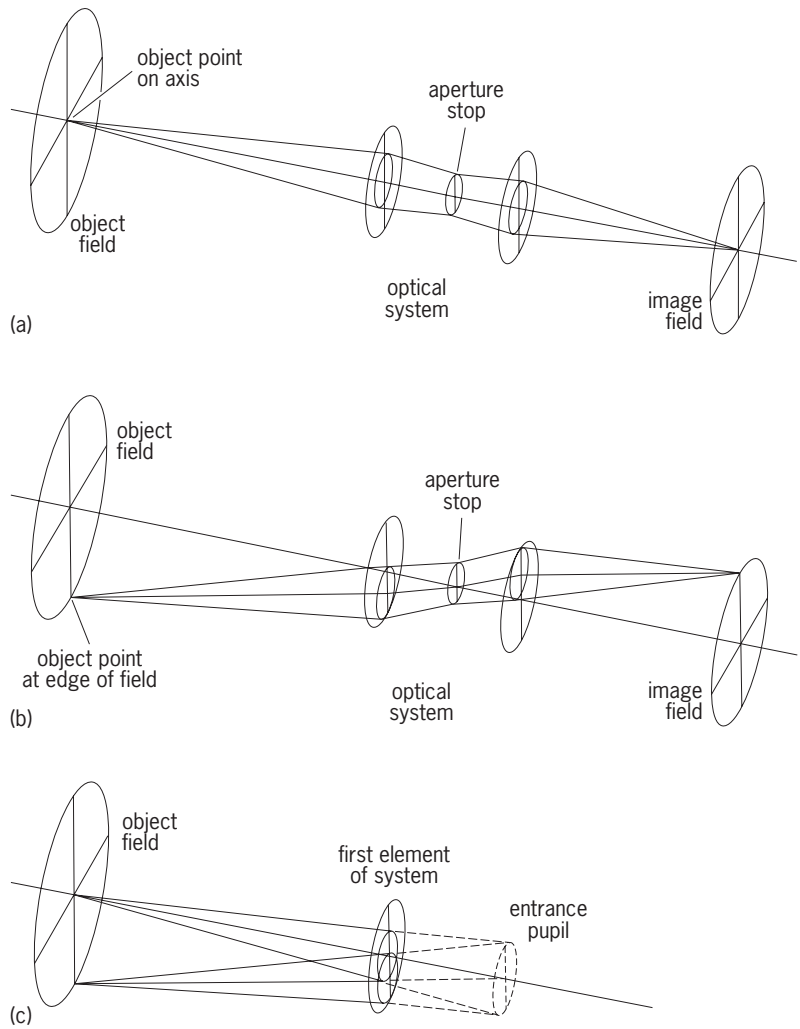

Fig. 5. Optical system with aperture stop. (*a*) Axial beam determined by the aperture stop. (*b*) Off-axis beam determined by the aperture stop. (*c*) Entrance pupil and its connection with the axial and off-axis beams in object space.

Because the Lagrange invariant holds at all surfaces throughout the optical system, in particular the object, the image, and the pupils, it is useful in solving many problems in the early stages of laying out an optical system, especially those involving a sequence of subsystems. It also plays a major role in the radiometry of image-forming systems; the total amount of light which can pass through a system is proportional to the square of the Lagrange invariant.

Although the determination of the first-order properties of an optical system is of major importance, this is only the beginning of the task of designing a good optical system. It is then necessary to trace real rays, determine the aberrations, and adjust the system parameters to improve the system's performance while maintaining its first-order behavior. *See* BINOCULARS; CAMERA; EYEPIECE; GUNSIGHTS; LENS (OPTICS); MIRROR OPTICS; OPTICAL MICROSCOPE; OPTICAL PRISM; OPTICAL PROJECTION SYSTEMS; OPTICAL SURFACES; OPTICAL TRACKING SYSTEMS; PERISCOPE; RANGEFINDER (OPTICS).

Roland V. Shack

Bibliography. G. L. James, *Geometrical Theory of Diffraction*, rev. ed., 1986; M. Katz, *Introduction to Geometrical Optics*, 1994; R. Kingslake, *Optical*

System Design, 1983; M. V. Kline and T. E. Furtak, *Optics,* 2d ed., 1986; W. Smith, *Modern Optical Engineering*, 3d ed., 2000.

Geometry

A branch of mathematics concerned with the properties of space, including points, lines, curves, planes and surfaces in space, and figures bounded by them.

Euclidean geometry. Geometry as a high school subject is largely based on the *Elements* of Euclid of Alexandria, a 13-volume work written about 300 B.C., of which the first 6 volumes present plane geometry, the next 4 are concerned with numbers and length, and the last 3 develop solid geometry. Euclid's *Elements* present in a logical order a sequence of over 400 mathematical propositions or theorems, proved on the basis of a set of 10 axioms or postulates which are assumed to be "self-evident" or true without proof—for example: "two points determine a line." Many theorems involve properties of triangles, circles, and other geometric figures, concerning lengths, angles, and congruence.

The most famous of Euclid's assumptions is the fifth, called the parallel postulate, which asserts that: given a line l and a point P not on l, there is one and only one line through P in the plane containing l and P that does not intersect l. Many unsuccessful attempts were made over the centuries to deduce Euclid's parallel postulate from the other axioms. Equivalent to this postulate is the assumption that the sum of the interior angles in a triangle is 180°. In a triangle bounded by arcs of three great circles on a sphere, the sum of the angles always exceeds 180°. *See* EUCLIDEAN GEOMETRY.

Noneuclidean geometries. K. F. Gauss is credited with discovering an "elliptic" noneuclidean geometry in which there are no parallels, but the other euclidean axioms are satisfied. One way to model it is to call each diameter of a sphere a "point," and call each plane through the center (intersecting the sphere in a great circle) a "line." Then each two points determine a unique line and each two lines determine one point.

Another type of noneuclidean geometry, called hyperbolic geometry, was discovered about 1830 by J. Bolyai and N. I. Lobachevski. Through a given point P not on a line l, there are many lines that do not meet l. This geometry may be modeled by defining as "lines" those circular arcs that meet a large circle C at right angles and restricting the term "points" to those within the absolute circle C. Then two points determine a unique line, but there are many lines through a point P not on a line l that do not intersect l. Since consistent geometries exist in which the parallel postulate does not hold, this postulate is independent of the others. *See* NONEUCLIDEAN GEOMETRY.

Projective geometry. Projective geometry studies geometric properties invariant under projections (as in photography) which may distort angles and preserve neither lengths nor ratios of lengths, whereas euclidean geometry is confined to properties of figures such as angles and ratios of lengths that are invariant under rigid motions, reflections, and similarity transformations. If A, B, C, and D are four collinear points, the so-called cross ratio $(BA/BC) \div (DA/DC)$ is preserved under projection. To each set of parallel lines in euclidean geometry is added a "point at infinity" or "vanishing point" in projective geometry, so that each two lines meet in just one point. Lines joining a point P not on l to four points A, B, C, and D on l have the same cross ratio as the four points. A comprehensive theory is derived from the invariance of cross ratio, including a complete theory of conic sections. Hyperbolic, elliptic, and euclidean geometries may be studied in terms of projective transformations that fix respectively a real nondegenerate conic, or an imaginary nondegenerate or degenerate line conic, the latter consisting of two imaginary points at infinity that lie on all circles, but on no other conics. *See* PROJECTIVE GEOMETRY.

Algebraic geometry. Algebraic geometry is a study of solutions of systems of polynomial equations $f_j(x) = 0$ in several variables x_i thought of as coordinates of a point $x = (x_1, x_2, \ldots, x_n)$ in n-dimensional space. Classical studies emphasize properties of an irreducible algebraic plane curve $f(x_1, x_2) = 0$ (f an irreducible polynomial), such as singular and multiple points and genus. The curve is rational (of genus 0) if x_1 and x_2 can be expressed as rational functions of a parameter t. Two curves $f(x_1, x_2) = 0$ and $g(u_1, u_2) = 0$ are birationally equivalent if both the coordinates of each curve can be expressed as rational functions of the coordinates of the other. In modern studies, the coefficients in the polynomials f_j are chosen from an arbitrary field k, and solutions x are sought in an algebraically closed field K containing k. The set X of all points x at which all the f_j vanish is called an algebraic closed set, or variety. If X is not empty, then the set of all polynomials $f(x)$ that vanish on X form an ideal with a finite basis in the ring of polynomials. Modern algebraic geometry studies these ideals and corresponding varieties. *See* POLYNOMIAL SYSTEMS OF EQUATIONS; RING THEORY.

Differential geometry. Differential geometry uses the tools of the calculus to study properties of curves and surfaces, usually in a euclidean space where the squared distance between nearby points is given in rectangular coordinates by the pythagorean formula $ds^2 = dx^2 + dy^2 + dz^2$. Arc length s is defined along "smooth" curves. Geodesics on a surface, like great circles on a sphere, are curves of shortest length between points not too far apart. Curvature is defined at each point of a smooth space curve, and torsion (twisting) at points not belonging to a straight portion of the curve; both curvature and torsion are constant on a helix (similar to a coiled spring). Curvatures of surfaces play an important role. *See* DIFFERENTIAL GEOMETRY.

Riemannian geometry. Riemannian geometry, named for G. F. B. Riemann, generalizes the concepts of differential geometry to noneuclidean spaces of any number n of dimensions in which there may or may not be any "straight" lines of infinite extent. Points P are specified by coordinates x^i like longitude and latitude on the Earth's surface.

Squared distance is expressed by a quadratic form $ds^2 = \int g_{ij} dx^i dx^j$, summed over i and j from 1 to n, where the functions g_{ij} are components of a "metric tensor" that is a scalar function of n vectors, varying in value from point to point. Under a differentiable change of coordinates (somewhat more general than the change from rectangular to spherical coordinates in euclidean analytic geometry), the components g_{ij} are changed. But geodesic paths and a certain curvature obtained from second derivatives of the g_{ij} are independent of the choice of coordinates x^i and describe intrinsic properties of the space. *See* RIEMANNIAN GEOMETRY.

Concepts of riemannian geometry are employed in Albert Einstein's theory of relativity. In his special theory (1905), the invariant interval between events in space-time is $ds^2 = dt^2 - (dx^2 + dy^2 + dz^2)/c^2$, where c is the velocity of light, about 3×10^8 m/s (1.86×10^5 mi/s). Thus neither the apparent time interval dt between events nor the space interval but a combination of the two is the same for all observers. In his general theory (1916), Einstein introduces a more general riemannian metric with g_{ij}'s representing the local gravitational potential, and explains the acceleration ascribed by Newton to a gravitational force as acceleration due to motion along curved world lines in space-time that are bent by the gravitational field. *See* RELATIVITY; SPACE-TIME. J. Sutherland Frame

Bibliography. E. Ballico (ed.), *Projective Geometry with Applications*, 1994; I. Chavel, *Riemannian Geometry: A Modern Introduction*, 1995; H. S. M. Coxeter, *Introduction to Geometry*, 2d ed., 1989; A. Helfer, *Introduction to Modern Differential Geometry*, 1991; W. V. Hodge and D. Pedoe, *Methods of Algebraic Geometry*, 3 vols., 1994; I. R. Shafarevich (ed.), *Algebraic Geometry I: Algebraic Curves, Algebraic Manifolds and Schemes*, 1994.

Geomorphology

The study of landforms, including the description, classification, origin, development, and history of planetary surface features. Emphasis is placed on the genetic interpretation of the erosional and depositional features of the Earth's surface. However, geomorphologists also study primary relief elements formed by movements of the Earth's crust, topography on the sea floor and on other planets, and applications of geomorphic information to problems in environmental engineering.

Geomorphologists analyze the landscape, a factor of immense importance to humankind. Their purview includes the structural framework of landscape, weathering and soils, mass movement and hillslopes, fluvial features, eolian features, glacial and periglacial phenomena, coastlines, and karst landscapes. Processes and landforms are analyzed for their adjustment through time, especially the most recent portions of Earth history.

History. Geomorphology emerged as a science in the early nineteenth century with the writings of James Hutton, John Playfair, and Charles Lyell. These men demonstrated that prolonged fluvial erosion is responsible for most of the Earth's valleys. Impetus was given to geomorphology by the exploratory surveys of the nineteenth century, especially those in the western United States. By the end of the nineteenth century, geomorphology had achieved its most important theoretical synthesis through the work of William Morris Davis. He conceived a marvelous deductive scheme of landscape development through the action of geomorphic processes acting on the structure of the bedrock to induce a progressive evolution of landscape stages.

Perhaps the premier geomorphologist was Grove Karl Gilbert. In 1877 he published his report "Geology of the Henry Mountains." This paper introduced the concept of equilibrium to organize tectonic and erosional process studies. Fluvial erosion was magnificently described according to the concept of energy. Gilbert's monograph "Lake Bonneville" was published in 1890 and described the Pleistocene history of the predecessor to the Great Salt Lake (**Fig. 1**). The monograph is a masterpiece of dynamic analysis. Concepts of force and resistance, equilibrium, and adjustment—these dominated in Gilbert's study of geomorphology. He later presented a

Fig. 1. The great bar of Pleistocene Lake Bonneville at Stockton, Utah. (*After G. K. Gilbert, U.S. Geol. Surv. Monogr. 1, 1890*)

Fig. 2. Surveying large transverse gravel bars created by flooding of the Medina River, Texas, in August 1978.

thorough analysis of fluvial sediment transport and the environmental effects of altered fluvial systems. He even made a perceptive study of the surface morphology of the Moon.

Despite Gilbert's example, geomorphologists in the early twentieth century largely worked on landscape classification and description according to Davis's theoretical framework. Toward the middle of the twentieth century, alternative theoretical approaches appeared. Especially in France and Germany, climatic geomorphology arose on the premise that distinctive landforms and processes are associated with certain climatic regions. Geomorphology since 1945 has become highly diversified, with many groups specializing in relatively narrow subfields, such as karst geomorphology, coastal processes, glacial and Quaternary geology, and fluvial processes.

Process geomorphology. Modern geomorphologists emphasize basic studies of processes presently active on the landscape (**Fig. 2**). This work has benefited from new field, laboratory, and analytical techniques, many of which are borrowed from other disciplines. Geomorphologists consider processes from the perspectives of pedology, soil mechanics, sedimentology, geochemistry, hydrology, fluid mechanics, remote sensing, and other sciences. The complexity of geomorphic processes has required this interdisciplinary approach, but it has also led to a theoretical vacuum in the science. At present many geomorphologists are organizing their studies through a form of systems analysis. The landscape is conceived of as a series of elements linked by flows of mass and energy. Process studies measure the inputs, outputs, and transfers for these systems. Although systems analysis is not a true theory, it is compatible with the powerful new tools of computer analysis and remote sensing. Systems analysis provides an organizational framework within which geomorphologists are developing models to predict selected phenomena.

The future. Geomorphology is increasing in importance because of the increased activity of humans as a geomorphic agent. As society evolves to more complexity, it increasingly affects and is threatened by such geomorphic processes as soil erosion, flooding, landsliding, coastal erosion, and sinkhole collapse. Geomorphology plays an essential role in environmental management, providing a broader perspective of landscape dynamics than can be given by standard engineering practice.

The phenomenal achievements of nineteenth-century geomorphology were stimulated by the new frontier of unexplored lands. The new frontier for geomorphology in the late twentieth century lies in the study of other planetary surfaces (**Fig. 3**). Each new planetary exploration has revealed a diversity of processes that stimulates new hypotheses

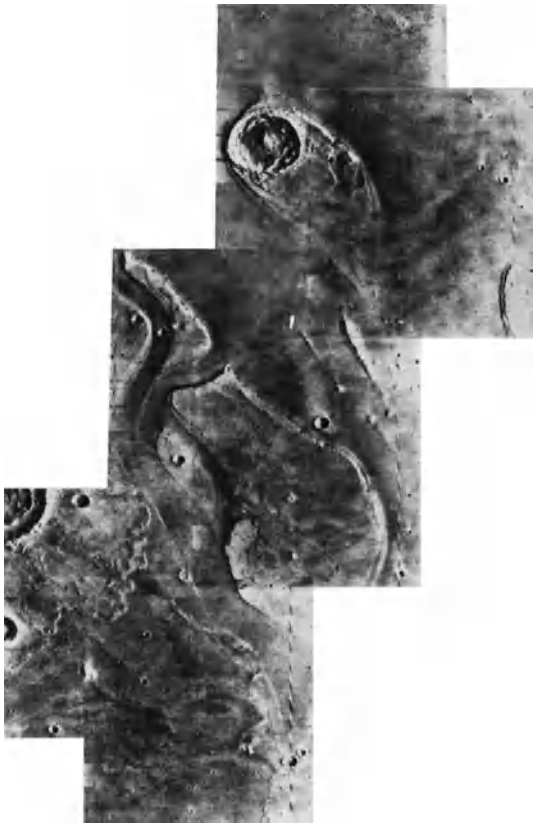

Fig. 3. Streamlined uplands and large sinuous channels in the Chryse Planitia region of Mars. (*National Aeronautics and Space Administration*)

for features on Earth. Geomorphology must now solve the mysteries of meteor craters on the Moon and Mercury, great landslides and flood channels on Mars, phenomenally active volcanism on Io, and ice tectonics on Ganymede. *See* COASTAL LANDFORMS; EROSION; GLACIATED TERRAIN; KARST TOPOGRAPHY.

Victor R. Baker

Bibliography. V. R. Baker and S. J. Pyne, G. K. Gilbert and modern geomorphology, *Amer. J. Sci.*, 278:97–123, 1978; A. L. Bloom, *Geomorphology: A Systematic Analysis of Late Cenozoic Landforms*, 3d ed., 1997; A. F. Pitty, *Geomorphology: Themes and Trends*, 1985; D. F. Ritter, *Process Geomorphology*, 3d ed., 1995.

Geophagia

Soil ingestion by animals. Grazing animals such as sheep and cattle ingest varying amounts of soil when they graze herbage contaminated with it.

Pastures become contaminated with soil when livestock walk across the herbage, particularly in wet conditions, as the treading action pushes the plants against the soil surface while soil also brushes off their hooves. An increase in the amount of soil ingested by the animals is associated with an increase in the ash content of the feces. Thus, the daily intake of soil can be determined from the daily fecal output and the fecal ash content.

The amount of soil ingested by sheep and cattle is influenced by soil type, stock density, earthworm activity, management practices, and various seasonal factors. Soils that are well drained and have a strong structure do not break up so readily and contaminate the herbage as is the case for poorly drained, weak-structured soils. When the density of stock grazing in a given area of herbage is increased, the amount of treading is increased, while the herbage is grazed more closely. The overall effect is that more soil is transferred to the herbage and ingested. Earthworm casts deposited at the soil surface are also ingested when the herbage is closely grazed.

Geophagia is subject to seasonal variations. The wetter and cooler conditions of autumn and winter result in muddier herbage and an increase in soil ingestion by grazing animals. During the spring and summer, the greater growth of the herbage and drier conditions result in cleaner herbage, and there is a marked decrease in intake of soil.

The ingestion of soil affects the teeth of sheep. For instance, under the conditions prevailing in intensive systems of sheep farming, the incisor teeth of the sheep slowly wear down, so that the older animals have only a small amount of tooth showing above the gum. The rate of teeth wear has been found to be associated with the amount of soil ingested, the causative agent being the abrasive action of the soil.

Soil can be a source of mineral nutrients. Since soils are higher than herbage in iron, manganese, zinc, copper, cobalt, selenium, and iodine, they may contribute to the mineral nutrition of the grazing animals. *See* AGRICULTURAL SCIENCE (ANIMAL); AGRICULTURE; SOIL CHEMISTRY.

Neville D. Grace

Geophysical exploration

Making, processing, and interpreting measurements of the physical properties of the Earth with the objective of practical application of the findings. Most exploration geophysics is conducted to find commercial accumulations of oil, gas, coal, or other minerals, but geophysical investigations are also employed with engineering objectives, in studies aimed at predicting the nature of the Earth for the foundations

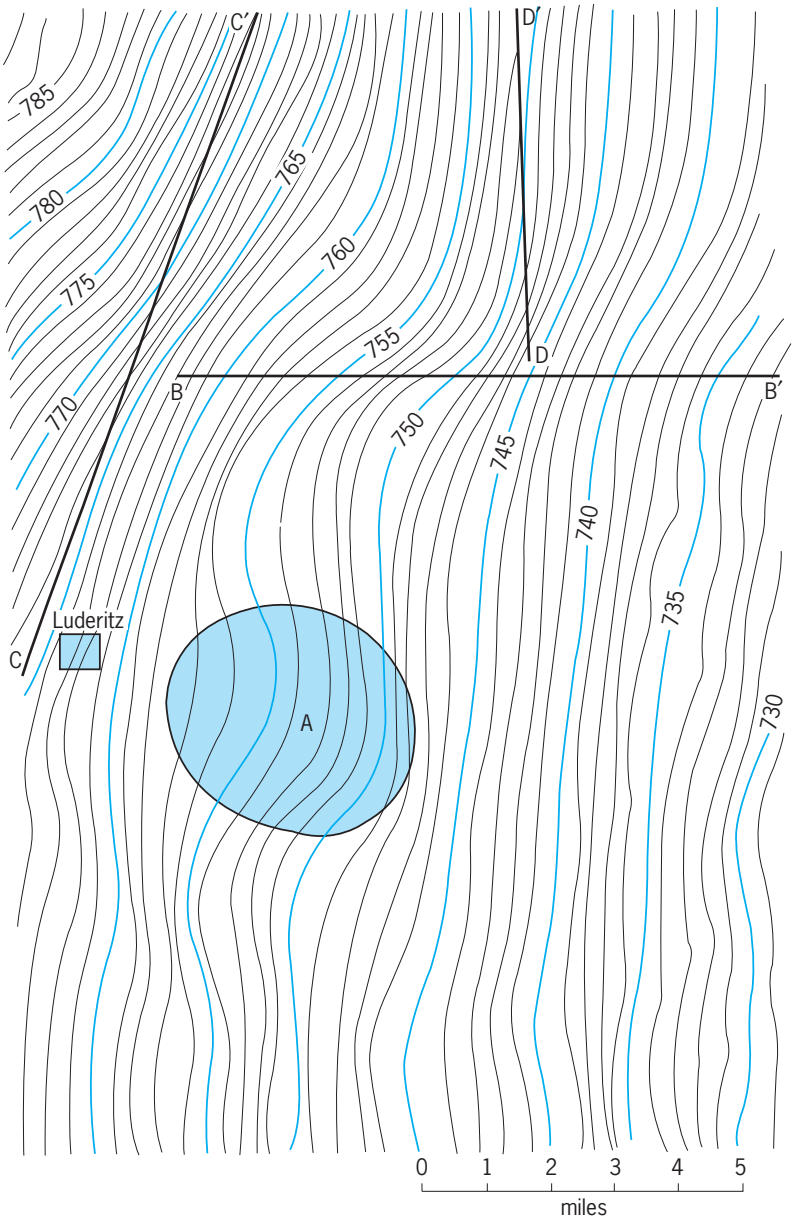

Fig. 1. Bouguer gravity of a portion of the Perth Basin, Western Australia. Contour interval is 1 mGal; datum is arbitrary. Departures from regularity are called anomalies. The bulge *A* results from an uplift area, the contour offset along *BB′* is an east-west trending fault, downthrown to the south. Other faults (*CC′* and *DD′*) are indicated by a closer spacing of the contours; both are downthrown to the east. Some of the variations in contour spacing result from measurement errors, some from interpolating between control points. 1 mi = 1.6 km. (*Western Australian Petroleum Pty. Ltd.*)

of roads, buildings, dams, tunnels, nuclear power plants, and other structures, and in the search for geothermal areas, water resources, pollution, archeological ruins, and so on.

Geophysical exploration, also known as applied geophysics or geophysical prospecting, is often divided into subsidiary fields according to the property being measured, such as magnetic, gravity, seismic, electrical, electromagnetic, thermal, or radioactive. *See* GEOPHYSICS.

General principles. A number of principles apply to most of the different types of geophysical exploration. Occasionally, prospective features can be mapped directly, such as iron deposits by their magnetic effects, but most features are studied indirectly by measuring the properties or the geometry of rocks that are commonly associated with certain mineral deposits.

Ordinarily, an anomaly is sought, that is, a departure from the uniform geologic characteristics of a portion of the Earth (**Fig. 1**). The primary objective of a survey is usually to determine the location of such departures. Sometimes, areas of anomalous data are obvious, but more often they are elusive because the anomaly magnitude is small compared to the background noise or because of the interference of the effects of different features. A variety of averaging and filtering techniques are used to accentuate the anomalous regions of change.

An anomaly usually seems smaller as the distance between the anomalous source and the location of a measurement increases (**Fig. 2**). Hence, a nearby source usually produces a sharp anomaly detectable only over a limited region, although possibly of large magnitude in this region. The detail of measurement required to locate anomalies must be compatible with the depth of the sources of interesting anomalies. If the source of an anomaly is deep in the Earth, then the anomaly is spread over a wide area, and its magnitude is small at any given location. As the depth of the anomaly increases, more sensitive instruments are needed because the effects become much smaller. Hence, the depth of the feature sought governs both the amount of detail and the precision required in measurements. Many of the differences in geophysical methods derive from the different depths of interest. Engineering, mineral, and ground-water objectives are usually shallow, often less than 100 ft (30 m), whereas petroleum and natural gas accumulations are usually quite deep: 0.6–5 mi (1–8 km).

Geophysical data usually are dominated by effects that are of no interest, and such effects must be either removed or ignored to detect and analyze the anomalous effects being sought. Noise caused by near-surface variations is especially apt to be large. The averaging of readings is the most common way of attenuating such noise.

The interpretation of geophysical data is almost always ambiguous. Since many different configurations of properties in the Earth can give rise to the same data, it is necessary to select from among many possible explanations those that are most probable, and (usually) to select from among the probable explanations the few that are most optimistic from the point of view of achieving set objectives. In engineering and pollution applications, the most pessimistic interpretation may be sought so that the worst situation can be studied.

Geologic features affect the various types of measurements differently; hence more can be learned from several types of measurements than from any one alone. Combinations of methods are particularly useful in mining exploration. In petroleum surveys,

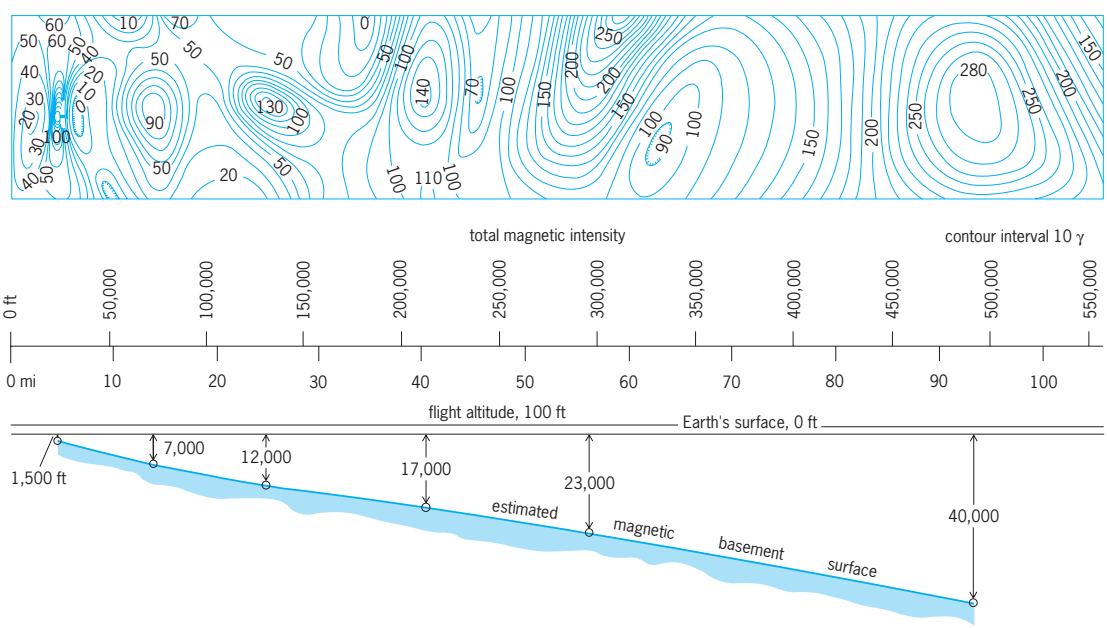

Fig. 2. Portion of a magnetic map (top) and interpretation (below). On the map, sharper features at the left indicate basement rocks are shallow, whereas broad anomalies to the right indicate deep basement. 1 ft = 0.3 m; 1 mi = 1.6 km. (*After S. L. Hammer, 4th World Petroleum Congress Proceedings, Rome, Sec. 1, 1955*)

magnetic, gravity, and seismic explorations are apt to be performed in that sequence, which is the order of relative cost. Less expensive methods are used first to narrow down the region to be explored by more expensive methods.

Usually the properties that geophysicists are able to measure are not directly related to objectives of interest; hence some association must be developed between the measured properties and features of interest. Geophysical measurements can determine something about the location, lateral extent, and depth of the source of an anomaly but not its specific cause. For example, an anomalous area located by ground-penetrating radar might be caused by a multitude of things, many of them of no interest. Interpretation utilizes other known information; for example, knowledge that potentially hazardous materials have been buried in the area might suggest this explanation for an anomaly. Reasoning is usually somewhat inferential; an ore often associated with a geological feature might be found by looking near geophysical indications of these features. The inference that the factors that produced a particular structural feature may have also affected sedimentation may lead to the discovery of a stratigraphic accumulation. *See* STRATIGRAPHY.

Magnetic surveying. Rocks and ores containing magnetic minerals become magnetized by induction in the Earth's magnetic field so that their induced field adds to the Earth field. Magnetic exploration involves mapping variations in the magnetic field to determine the location, size, and shape of such bodies. *See* GEOMAGNETISM.

The magnetic susceptibility of sedimentary rock is generally orders of magnitude less than that of igneous or metamorphic rock. Consequently, the major magnetic anomalies observed in surveys of sedimentary basins usually result from the underlying basement rocks. Determining the depths of the tops of magnetic bodies is thus a way of estimating the thickness of the sediments. Gradiometers are used to detect sedimentary structures such as faults. *See* ROCK MAGNETISM.

Except for magnetite and a very few other minerals, mineral ores are only slightly magnetic. However, they are often associated with bodies such as dikes that have magnetic expression so that magnetic anomalies may be associated with minerals empirically. For example, placer gold is often concentrated in stream channels where magnetite is also concentrated.

Instrumentation. Several types of instruments are used for measuring variations in the Earth's magnetic field. Because the magnetic field is a vector quantity, its magnitude and direction can be measured or, alternatively, components of the field in different directions. Usually, however, only the magnitude of the total field is measured.

Optically pumped, proton and fluxgate magnetometers are used extensively in magnetic exploration. Although it is fairly easy to achieve magnetometers with even greater sensitivity, those used in exploration typically are accurate to 0.1–10 nanoteslas (0.1–10 gammas), which is compatible with uncertainties in noise background. More sensitive magnetometers such as the cryogenic superconducting quantum interference device (SQUID) are sometimes used where more precise measurements are needed. *See* MAGNETOMETER; SQUID.

Field methods. Most magnetic surveys are made by aircraft, because a large area can be surveyed in a short time, and thus the cost per unit of area is kept very low. Aeromagnetic surveying is especially adapted to reconnaissance, for locating those portions of large, unknown areas that contain the best exploration prospects so that future efforts can be concentrated there. Other types of measurements (for example, electromagnetic, gamma-ray) are often made simultaneously.

The spacing of measurements must be finer than the size of the anomaly of interest. Petroleum exploration usually concentrates only on large anomalies, hence a survey for such objectives may involve flying a series of parallel lines spaced 0.6–2 mi (1–3 km) apart, with tie lines (perpendicular lines) every 6–9 mi (10–15 km) to assure that the data on adjacent lines can be related properly. The flight elevation is usually 900–3000 ft (300–1000 m). In mineral exploration, lines are usually located much closer—sometimes less than 300 ft (100 m) apart—and the flight elevation is as low as safety permits. Helicopters are sometimes used for mineral magnetic surveys.

Aircraft are usually equipped with Global Positioning System (GPS) and Doppler radar navigation and with both aneroid and radar altimeters so that the locations of measurements are known accurately. Aircraft also use radionavigation measurements and aerial photographs or other means to locate the aircraft. *See* AERIAL PHOTOGRAPHY; ALTIMETER; DOPPLER RADAR; SATELLITE NAVIGATION SYSTEMS.

The immediate product of aeromagnetic surveys is a graph of the magnetic field strength along lines of traverse. After adjustment, the data are usually compiled into maps on which magnetism is shown by contours (isogams) that connect points of equal magnetic field strength.

In ground mineral exploration, the magnetic field is measured at closely spaced stations. The effects of near-surface magnetic bodies is accentuated over measurements made in the air. Magnetic surveys are also carried out on the ground to delineate near-surface features such as buried drums and tanks or archeological artifacts. Magnetic gradient measurements [measurements made at nearby points (sometimes 10 ft or 3 m apart) so that differences give the magnetic gradients] are especially sensitive to near-surface features.

Magnetic surveying is often done in conjunction with other geophysical measurements, because it adds only a small increment to the cost and the added information often helps in resolving interpretational ambiguities.

Data reduction and interpretation. The reduction of magnetic data is usually simple. Often, measurement

conditions vary so little that the data can be interpreted directly, or else require only network adjustments to minimize differences at line intersections. The magnetic field depends on the elevation at which it is measured, but data can be continued to a different elevation; that is, the magnetic field at one elevation can be determined from knowledge of the field at a different elevation. Where different parts of an area have been surveyed at different elevations, continuation can be used to reconcile them. In surveys of large areas, the variations in the Earth's overall magnetic field may be removed (magnetic latitude correction). In exceptional cases, such as in land surveys made over very irregular terrain, as in bottoms of canyons where some of the magnetic sources may be located above the instrument, reduction of the data can become difficult.

The sharpness of a magnetic anomaly depends on the distance to the magnetic body responsible for the anomaly. Inasmuch as the depth of the magnetic body is often the information being sought, the shape of an anomaly is the most important aspect. Modeling is used to determine the magnetic field that would result from bodies of certain shapes and depths. The model anomalies are examined for a parameter of shape that is proportional to the depth (**Fig. 3**). The shape parameter is measured on real anomalies and scaled to indicate how deep the body responsible for the anomaly lies. Such estimates are typically accurate to 10–20%, sometimes better.

Iterative modeling techniques are used in more detailed studies. The field indicated by a model is subtracted from the observed field to give an error field. Then the model is changed to obtain a new error field. This process is repeated until the error field is made sufficiently small. The model then represents one possible explanation of the anomaly.

One commonly employed interpretation technique is to assume that the measured field results from near-vertical dikes or the edges of thin horizontal sheets. Seven or so successive equally spaced measurements can then be solved for the depth and other parameters of anomalies. Computers solve for each successive set of measurements, and then sophisticated filtering techniques eliminate solutions that are not consistent.

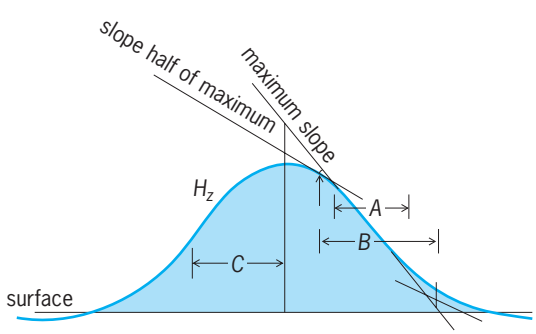

Fig. 3. Variation of shape measurements across an anomaly. Among the shape factors sometimes measured are *A*, the distance over which the slope is maximum; *B*, the distance between points where the slope is half of the maximum slope; *C*, half the width of the anomaly at half the peak magnitude. Such shape measurements multiplied by index factors give estimates of the depth of the body responsible for the anomaly.

Gravity exploration. Gravity exploration is based on the law of universal gravitation: the gravitational force between two bodies varies in direct proportion to the product of their masses and in inverse proportion to the square of the distance between them.

Because the Earth's density varies from one location to another, the force of gravity varies from place to place. Gravity exploration is concerned with measuring these variations to deduce something about rock masses in the immediate vicinity. *See* EARTH, GRAVITY FIELD OF.

Vertical density changes affect all stations equally and so do not produce easily measured effects. Gravity field variations are produced by lateral changes in density. Absolute density values are not involved, only horizontal changes in density. The product of the volume of a body and the difference between the density of the body and that of the horizontally adjacent rocks is called the anomalous mass.

Gravity surveys are used more extensively for petroleum exploration than for metallic mineral prospecting. The size of ore bodies is generally small; therefore, the gravity effects are quite small and local despite the fact that there may be large density differences between the ore and its surroundings. Hence, gravity surveys to detect ore bodies have to be very accurate and very detailed. In petroleum prospecting, on the other hand, the greater dimensions of the features more than offset the fact that density differences are usually smaller. *See* PETROLEUM; PROSPECTING.

Instrumentation. The most common gravity instrument in use is the gravity meter or gravimeter. The gravimeter basically consists of a mass suspended by springs comprising a balance scale. The gravimeter can be balanced at a given location, then moved to another location, and the minute changes in gravitational force required to rebalance the instrument can be measured. Hence the gravimeter measures differences in a gravity field from one location to another rather than the gravity field as a whole.

A gravimeter is essentially a very sensitive accelerometer, and extraneous accelerations affect the meter in the same way as the acceleration of gravity affects it. Typically, gravimeters read to an accuracy of 0.01 mGal, which amounts to 1/100,000,000 of the Earth's gravitational field. Anomalies of interest in petroleum exploration are often of the magnitude of 0.5 to 5 mGal. Extremely sensitive gravimeters measuring to microgal accuracy are used in boreholes to locate cavities such as caves and tunnels, and in archeological applications to search for burial chambers. *See* ACCELEROMETER; GRAVITY METER.

Field surveys. Almost all gravity measurements are relative measurements; differences between locations are measured although the absolute values remain unknown. Ordinarily, the distance between the stations should be smaller than one-half the depth of the structures being studied.

Gravity surveys on land usually involve measurements at discrete station locations. Such stations are

spaced as close as a few meters apart in some mining or archeological surveys, about 0.3 mi (0.5 km) for petroleum exploration, and 6–10 mi (10–20 km) for some regional geology studies. While it is desirable to have gravity values on a uniform grid, often this is not convenient, and so stations are located on traverses around loops. For petroleum exploration, the gravimeter might be read every 0.3 mi (0.5 km) around loops of about 4 by 6 mi (6 by 10 km). Helicopters are used for transport between stations in areas of difficult terrain. Location and elevation are then determined by an inertial navigation system, also carried by the helicopter.

The gravity field is very sensitive to elevation. An elevation difference of 9 ft (3 m) represents a difference in gravity of about 1 mGal. Hence, elevation has to be known very accurately, and the most critical part of a gravity survey often is determining elevations to sufficient accuracy.

Gravity measurements can be made by ships at sea. Usually the instrument is located on a gyrostabilized platform which holds the meter as nearly level as possible. The limiting factor in shipboard gravity data is usually the uncertain velocity of the meter, especially east to west, since the ship is moving. The velocity of a ship traveling east adds to the velocity because of the rotation of the Earth. Consequently, centrifugal force on the meter increases and the observed gravity value decreases (Eötvos effect).

Gravity measurements are also sometimes made by lowering a gravimeter to the ocean floor and balancing and reading the meter remotely. Gravity measurements have been made by aircraft using techniques like those used at sea, but are not sufficiently accurate to be useful for most exploration.

Specialized gravimeters are used to make measurements in boreholes. The main difference between gravity readings at two depths in a borehole is produced by the mass of the slab of earth between the two depths; this mass pulls downward on the meter at the upper level and upward at the lower level. Thus the difference in readings depends on the density of this slab. In sedimentary rocks, the borehole gravimeter is used primarily for measuring porosity. The density of most minerals in sedimentary rocks is about the same, but very different from water-filled pore spaces.

Variations in the Earth's gravity field affect sea level. Orbiting satellites can measure their elevation with respect to sea level with sufficient accuracy to map variations in the Earth's field over the oceans. Satellites can measure gravity anomalies at sea that are larger than about 5 mGal and 15 mi (25 km) width.

Data reduction. Gravity measurements have to be corrected for factors other than the distribution of the Earth's mass. Meters drift or change their reading gradually because of various reasons. The Sun and Moon pull on the meter in different directions during the course of a day. The gravitational force varies with the elevation of the gravimeter both because at greater elevations the distance from the Earth's center increases (free-air correction) and because mass exists between the meter and the reference elevation, which is usually mean sea level (Bouguer correction). Gravity varies with latitude because the Earth's equatorial radius exceeds its polar radius and because centrifugal force resulting from the Earth's rotation varies with latitude. Nearby terrain affects a gravimeter; mountains exert an upward pull, valleys cause a deficit of downward pull. Thus the effects of nearby elevation differences add, whether the differences are positive or negative. This is the most critical correction to be made to gravity data in areas of rough terrain.

Gravity measurements that have been corrected for all of these effects are called Bouguer anomalies, or free-air anomalies if the Bouguer correction has not been made. They therefore represent the effects of local masses within the Earth, that is, effects for which corrections have not been made. Most gravity maps display contours (isogals) of free-air or Bouguer anomaly values.

Data interpretation. The most important part of gravity interpretation is locating anomalies that can be attributed to mass concentrations being sought, isolating these from other effects (Fig. 1). Separating the main part of the gravitational field, which is not of interest (the regional), from the parts attributed to local masses, the residuals, is called residualizing.

Many techniques for gravity data analysis are similar to those used in analyzing magnetic data. Shape parameters are used to determine the depth of the mass's center. Another widely used technique is model fitting: a model of an assumed feature is made, its gravity effects are calculated, and the model is compared with field measurements.

Continuation is a process by which calculations are made from measurements of the gravity field over one surface to determine what values the field would have over another surface. A field can be continued if there is no anomalous mass between the surfaces. Continuing the field to a lower surface produces sharper anomalies as the anomalous mass is approached. However, if the process is carried too far, instability occurs when the anomalous mass is reached. The technique, however, is very sensitive to measurement uncertainties and often is not practical with real data.

Seismic exploration. The seismic method is the predominant geophysical method. Seismic waves are generated by one of several types of energy sources and detected by arrays of sensitive devices called geophones or hydrophones. The most common measurement made is of the travel times of seismic waves and the amplitude of the waves, with less attention being given to changes in their frequency content or wave shape.

The seismic method is divided into two major classes, refraction and reflection, and two types based on the objectives, exploration and reservoir studies. Method classification depends on whether the predominant portion of wave travel is horizontal or vertical, respectively.

Principles of seismic waves. A change in mechanical stress produces a strain wave that radiates outward

as a seismic wave, because of elastic relationships. The radiating seismic waves are like those that result from earthquakes, though much weaker. Most seismic work involves the analysis of P waves (compressional waves) in which particles move in the direction of wave travel, analogous to sound waves in air. S waves (shear waves) are occasionally studied, but most exploration sources do not generate very much shear energy. Surface waves, especially Rayleigh waves, are also generated, but these are mainly a nuisance because they do not penetrate far enough into the Earth to carry much useful information. Recording techniques are designed to discriminate against them. *See* SEISMOLOGY.

A seismic wave is a vector, involving both magnitude and direction. Historically, measurements have been made of only the vertical component of motion (with geophones), or only of the magnitude (with hydrophones). Attention is now being shifted to measuring all components of wave motion in order to study the conversion of P-waves to S-waves, and vice versa, and anisotropy.

The amplitude of a seismic wave reflected at an interface depends on the elastic properties, often expressed in terms of seismic velocity and density on either side of the interface. When the direction in which the wave is traveling is perpendicular to the interface, the ratio of the amplitudes of reflected and incident seismic waves is given by the normal reflection coefficient $R_\perp$ as shown in Eq. (1), where

$$R_\perp = \frac{\Delta(\rho V)}{2(\rho \bar{V})} \tag{1}$$

$\Delta(\rho V)$ is the change in the product of velocity and density and $(\rho \bar{V})$ is the average of the product of velocity and density on opposite sides of the interface. The relationships are much more complicated where wave travel is not perpendicular to interfaces. The variation of the reflection coefficient with angle of incidence is now routinely used to indicate the kind of fluid in the pore space and the lithology.

Seismic waves are bent when they pass through interfaces, and Snell's law holds, shown in Eq. (2),

$$\frac{\sin \sigma_1}{V_1} = \frac{\sin \sigma_2}{V_2} \tag{2}$$

where σ_i is the angle between a wavefront and the interface in the ith medium where the velocity is V_i (**Fig. 4**). Because velocity ordinarily increases with depth, seismic-ray paths become curved concave-upward (**Fig. 5**).

The resolving power (ability to separate features) with seismic waves depends inversely on their wavelength λ and is often thought of as of the order of $\lambda/4$. The wavelength is often expressed in terms of the wave's velocity and frequency f: $\lambda = V/f$. Most seismic work involves frequencies from 15 to 70 Hz, and most rocks have velocities from 4500 to 18,000 ft/s (1500 to 6000 m/s) so that wavelengths range from 90 to 900 ft (30 to 300 m). Usually, the frequency becomes lower and the velocity higher as depth in the Earth increases, so that wavelength increases

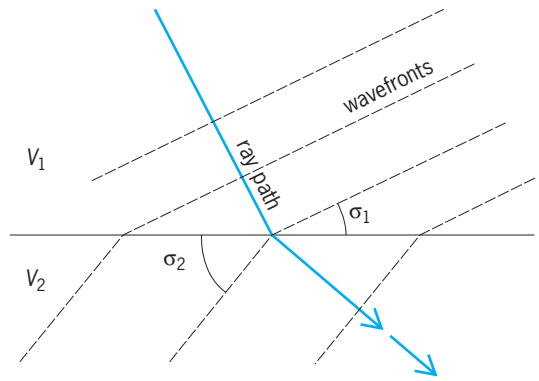

Fig. 4. Bending of seismic waves at interface; $V_2 > V_1$.

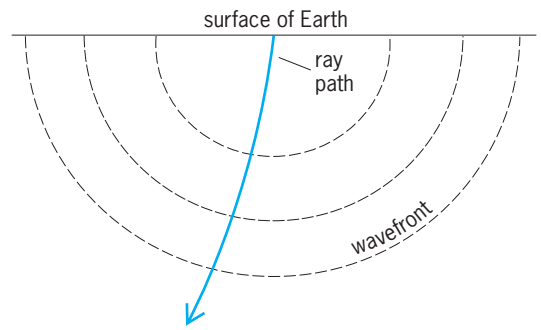

Fig. 5. Wavefronts become more widely spaced and ray paths become curved as velocity increases with depth.

and resolving power decreases. Very shallow, high-resolution work involves frequencies higher than those cited above, and long-distance refraction (and earthquakes) involve lower frequencies. *See* ECHO SOUNDER.

Reflection exploration. Seismic-wave energy partially reflects from interfaces where velocity or density changes. The measurement of the arrival times of reflected waves (**Fig. 6**) thus permits mapping the interfaces that form the boundaries between different kinds of rock. This, the predominant geophysical exploration method, can be thought of as similar to echo sounding. When a seismic source S generates seismic energy, it is received at detectors located at intervals, say from A to B. The distance to the reflector can be obtained from the arrival time of the reflection if the velocity is known. If the reflector dips towards A, the reflection will arrive sooner at B

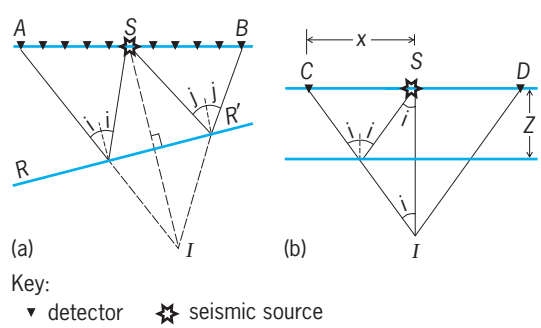

Fig. 6. Measuring reflected seismic wave energy. (*a*) Dipping reflector. (*b*) Flat reflector.

than at A; the difference in arrival times is a measure of the amount of dip.

For a flat reflector (Fig. 6*b*) and constant velocity V, the arrival time at detector C is t_c and the arrival time at a detector at the source S is $t_0 = 2Z/V$, where Z is the depth. From the pythagorean theorem (for the triangle CSI; Fig. 6*b*), Eqs. (3) and (4) are obtained. These give both the values of V and Z. Similar relationships can be used for nonflat reflectors or nonconstant velocity to yield velocity information.

$$(Vt_c)^2 = (Vt_0)^2 + x^2 = (2Z)^2 + x^2 \qquad (3)$$

$$V = x\left(t_c^2 - t_0^2\right)^{1/2} \qquad (4)$$

Usually a number of detector groups are used, and the arrival of reflected waves is characterized by coherency. Thus, if all of the detectors in a line move in a systematic way, a seismic wave probably passed. Multiple detectors make it possible to detect coherent waves in the presence of a high noise level and also to measure distinguishing features of the waves. *See* SEISMIC EXPLORATION FOR OIL AND GAS.

Refraction exploration. Refraction seismic exploration involves rocks characterized by high seismic velocity. Wavefronts are bent at interfaces (**Fig. 7**) so that appreciable energy travels in high-velocity members and arrives earlier at detectors distant from the source than energy that has traveled in overlying lower-velocity members. Differences in arrival time at different distances from the energy source yield information on the velocity and attitude (dip) of the high-velocity member.

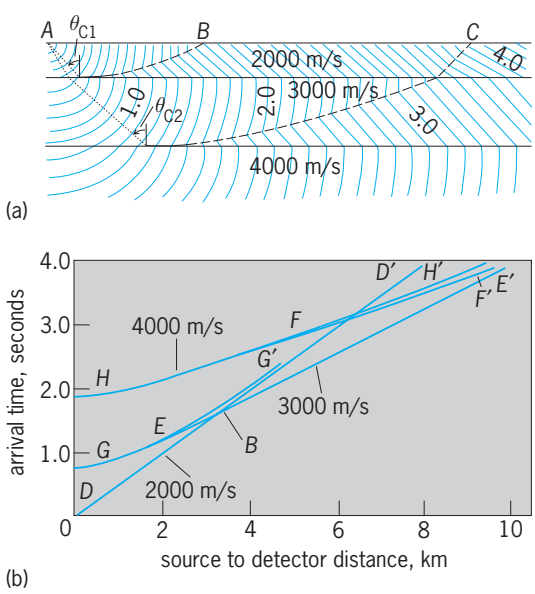

Fig. 7. Refractive seismic exploration data. (*a*) Section through model of layered earth. Curves (wavefronts) indicate location of seismic energy at successive times after a shot at *A*. Beyond *B*, energy traveling in the second layer arrives first; beyond *C*, that in the third layer arrives first. (*b*) Arrival time as a function of source-to-detector distance. *DD′*, direct wave; *EE′*, refracted wave (head wave) in the second layer; *FF′*, refracted wave in the third layer; *GG′*, reflection from interface between first and second layers; *HH′*, reflection from that between second and third layers. 1 km = 0.6 mi. 1 m/s = 3.3 ft/s.

A variant of refraction seismic exploration is the search for high-velocity masses in an otherwise low-velocity section, by looking for regions where seismic waves arrive earlier than expected. Such arrivals, called leads, were especially useful in locating salt domes in Louisiana, Texas, Mexico, and Germany in the late 1920s and 1930s.

Refraction seismic techniques are used in groundwater studies, engineering geophysics, and mining to map the water table and bedrock under unconsolidated overburden, with objectives such as foundation information or locating buried stream channels in which heavy minerals might be concentrated or where water might accumulate. Refraction techniques are also used in petroleum exploration and for crustal studies.

Channel-wave exploration. Seismic waves can become trapped once they are generated in low-velocity formations. The low-velocity formation constitutes a waveguide, and the waves are called channel, guided, or seam waves. Coal often satisfies waveguide requirements, and channel waves are used for determining the continuity of coal beds. The objective usually is to ascertain that the coal measures are not interrupted by faults or channels that would interfere with the operation of longwall mining machines. Sources and geophones are located in the coal bed in mining tunnels; and both reflection and transmission ray paths are used, the former where sources and geophones are in the same tunnel and the latter where they are in different tunnels. Channel waves are also sometimes studied in borehole-to-borehole measurements to ascertain the continuity of reservoirs.

Instrumentation. Detectors of seismic energy on land (geophones or seismometers) are predominantly electromechanical devices. A coil moving in a uniform magnetic field generates a voltage proportional to the velocity of the motion. Usually the coil has only one degree of freedom and is used so it will be sensitive to vertical motion only. Three mutually perpendicular elements in a three-component detector are coming into increased use to distinguish the type of waves (compressional, shear, Rayleigh, and such) or to determine the direction from which the waves come. *See* SEISMOGRAPHIC INSTRUMENTATION.

Detectors in water are usually piezoelectric. Pressure changes produced as a seismic wave passes distort a ceramic element and induce a voltage between its surfaces. Such detectors are not directionally sensitive. *See* HYDROPHONE.

Detectors are usually arranged in groups (arrays) spread over a distance and connected electrically so that, in effect, the entire group acts as a single large detector. Such an arrangement discriminates against seismic waves traveling in certain directions. Thus a wave traveling horizontally reaches different detectors in the group at different times, so that wave peaks and troughs tend to cancel; whereas a wave traveling vertically affects each detector at the same time, so that the effects add. The principles of seismic array design are similar to those in radio antenna design.

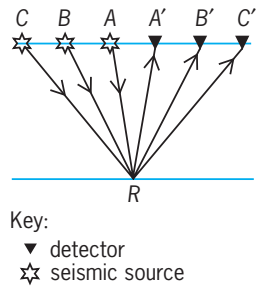

Fig. 8. Reflection of seismic energy obtained from reflection point *R* (common midpoint) by positioning a source at *A* and a detector at *A′*. The same point is involved with a source at *B* and a detector at *B′*, or *C* and *C′*. The redundancy in measuring reflections from the same point many times (often 48- to 96-fold) permits sorting out different kinds of waves.

The signal from the detectors is transmitted to recording equipment over a cable, streamer, or radio link, and then amplified and recorded. The output level from a geophone varies tremendously during a recording. Seismic recording systems are linear over ranges of 120 dB or more. Seismic amplifiers employ various schemes to compress the range of seismic signals without losing amplitude information. They also incorporate adjustable filters to permit discriminating on the basis of frequency.

Many sources are used to generate seismic energy. The classical land energy source is an explosion in a borehole drilled for the purpose, and solid explosives continue to be used extensively for work on land and in marshes. The explosion of a gas mixture in a closed chamber, a dropped weight, a hammer striking a steel plate, and other sources of impulsive energy are used in land work. An air gun, which introduces a pocket of high-pressure air into the water, is the most common energy source in marine work. Other marine energy sources involve the explosion of gases in a closed chamber, a pocket of high-pressure steam introduced into the water, the discharge of an electrical arc, and the sudden mechanical separation of two plates (imploder).

An oscillatory mechanical source (vibroseis) is the predominant source being used on land. Such a source introduces a long wave train so that individual reflection events cannot be resolved without subsequent processing (correlation with the input wave train), which, in effect, compresses the long wave train and produces essentially the same result as an impulsive source.

Field techniques. Most petroleum exploration seismic work has been carried out along lines of survey often run parallel to each other at right angles to the geological strike with occasional perpendicular tie lines, often run on a regular grid. Long lines many kilometers apart are sometimes run for regional information, but lines are often concentrated in regions in which anomalies have been detected by previous geophysical work. Most seismic work has the objective of mapping interfaces continuously along the seismic lines to map the geological structure.

Geophone groups are spaced 80–300 ft (25–100 m) apart with 48–120 adjacent groups of 6–24 geophones each being used for each recording. The source is sometimes located at the center of the active groups (split spread), sometimes at one end (end-on spread).

Following a recording, the layouts and sources are advanced down the line by some multiple of the group interval for redundant coverage (**Fig. 8**).

Small marine operations, often called profiling, may consist of an energy source and a short streamer containing a number of hydrophones and feeding a recorder. Larger marine operations (**Fig. 9**) involve ships 180 ft (60 m) or more in length towing a streamer 1 to 3 mi (2 to 8 km) long with 250 groups of hydrophones spaced along the streamer. An energy source is towed near the ship. Recordings are made as the ship is continuously under way at a speed of about 6 knots (3 m/s).

The foregoing methods acquire data along lines of traverse, but most seismic work today is designed to acquire data uniformly over an area. Such methods are known as three-dimensional, and they result in acquiring a volume rather than a cross section of data. A variety of geophone and source arrangements are used on land, most often with several parallel lines of geophones and perpendicular lines of sources. Often more than 1000 geophone groups will be recorded for each source location. Most marine three-dimensional data are acquired by ships towing two sources and up to 12 streamers pulled to the sides of the ship by paravanes, so that several closely spaced parallel lines of data are acquired on a single traverse. Except for requiring more data channels, instrumentation and methods are similar to those used for two-dimensional data acquisition. *See* OCEANOGRAPHY.

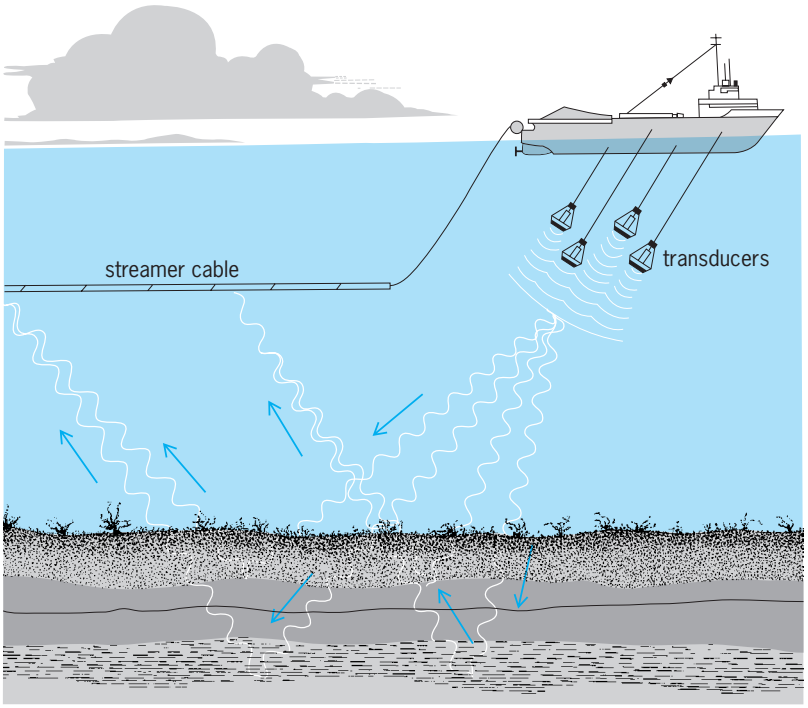

Fig. 9. Seismic surveying at sea is done by towing a long streamer containing detectors that sense seismic energy. Seismic energy is generated by several air guns that are also towed from the ship.

Three-dimensional methods generally require more precision in determining source and detector locations. The Global Positioning System and electronic distance instruments are used on land. Marine surveys use highly redundant location measurements of several different kinds. Global Positioning System instruments usually locate the ship and streamer tail buoys; acoustic transducers measure distances between the sources, streamers, and ship; and magnetic compasses within the streamers determine their orientations. A large volume of positioning data is acquired and reduced by computer in real time so that corrections can be made immediately.

Data reduction and processing. Seismic data are corrected for elevation and near-surface variations on the basis of survey data and observations of the travel time of the first energy from the source to reach the detectors, which usually involves travel either in a direct path or in shallow refractors.

Most seismic data processing is done either to reduce the noise so that structural and stratigraphic features can be seen more clearly, or to reposition features to display correctly their positions relative to each other so that they can be located by drilling a well. Seismic data processing is one of the larger users of giant computers.

Almost all data are processed by computers, with the first step often being editing, wherein data are merged with identifying data, rearranged, checked for being either dead or wild (with bad values sometimes replaced with interpolated values), time-shifted in accordance with elevation and near-surface corrections that have been determined in the field, scaled, and so on.

Following the editing, different processing sequences may follow, including (1) filtering (deconvolution) to remove undesired natural filter effects, trace-to-trace variations, variations in the strength or wave shape of the source, and so on; (2) grouping according to common midpoint (Fig. 8) or some other arrangement; (3) analyzing to see what velocity values will maximize coherency as a function of source-to-detector distance; (4) statistically analyzing to see what trace shifts will maximize coherency; (5) trace-shifting according to the results of steps 3 or 4; (6) stacking by adding together a number of individual traces; (7) migrating by rearranging and combining data elements in order to position reflection events more nearly under the surface locations where the appropriate reflecting surface is located; (8) another filtering; and (9) displaying of the data.

The techniques for processing two- and three-dimensional seismic data are nearly the same. When data are acquired only along lines of traverse, there is always ambiguity as to the directions from which the waves come, which produces ambiguity as to where features are located perpendicular to the line of acquisition. This ambiguity is removed with three-dimensional data sets, resulting in major improvements in the ability to resolve features. Three-dimensional data also permit many ways of displaying data that help in interpreting the significance of features.

Data interpretation. The travel times of seismic reflections are usually measured from record section displays (**Fig. 10**), which result from processing. Appropriate allowance (migration) must be made because reflections from dipping reflectors appear at locations downdip from the reflecting points. Allowance must also be made for variations in seismic velocity, both vertically and horizontally. Seismic events other than reflections must be identified and explained.

In petroleum exploration the objective is usually to find traps, places in which porous formations are high relative to their surroundings and in which the overlying formation is impermeable. If oil or gas,

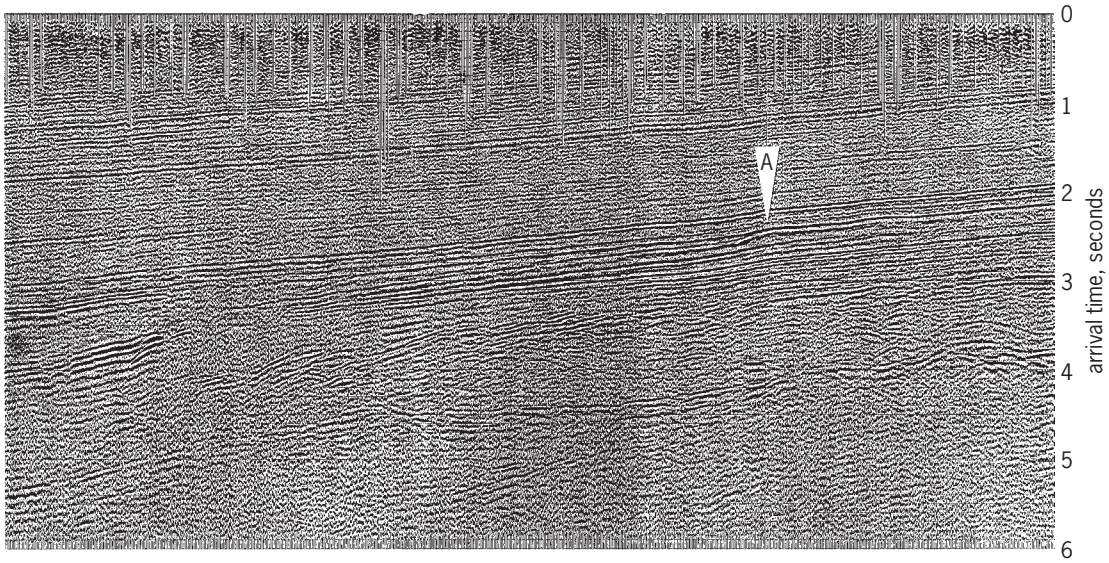

Fig. 10. Seismic record section in East Texas after processing. The rock bedding giving the various reflections was nearly horizontal when deposited, and the present dips result from tilting and other deformation subsequent to deposition. The reflection event at *A* is attributed to the Edwards Reef, which formed a barrier at the time the adjacent sediments were deposited. Downdip to the left of the reef formations the Woodbine sands can be seen thinning and pinching out in the updip direction. These are productive of oil and gas in this area. (*Grant Geophysical*)

which are lighter than water, are present, they float on top of the water and accumulate in the pores in the rock at the trap. Seismic exploration determines the geometry, hence where traps might be located. However, it usually is not possible to tell conclusively from seismic data alone whether oil or gas was ever generated, whether the rocks have porosity, whether overlying rocks are impermeable, or whether oil or gas might have escaped or been destroyed, even if they were present at one time.

In addition to mapping the structural patterns within the Earth, seismic data are analyzed for evidence that might identify the nature of the formations, the environment in which they were formed, and the nature of the fluid in the pore spaces. To reconstruct the geologic history, often several reflections at different depths are mapped, and attempts are made to reconstruct the position of the deeper reflectors at the time of deposition of shallower rocks.

After some experience has been developed in an area, patterns in the seismic data that distinguish certain reflectors or certain types of structure or stratigraphy often can be recognized. Seismic velocity measurements are helpful here. Seismic stratigraphy is considered an important aspect of sequence stratigraphy. *See* SEQUENCE STRATIGRAPHY.

In relatively unconsolidated sediments and in some other circumstances, gas and oil may lower the seismic velocity or rock density or both sufficiently to produce a distinctive reflection, usually evidenced by strong amplitude (a "bright spot") and other distinguishing features (**Fig. 11**). Coal and peat beds are also characterized by reflections of strong amplitude.

Much interpretation, including almost all three-dimensional interpretation, is done at computer workstations. The interpreter views data, base maps, and the progressing interpretation on computer screens. It is possible to call up any of the available data, display the data in various ways in various colors, pick events representing various horizons with the aid of an automatic picker, pick faults, carry out computations, and manipulate the data in various ways. In a three-dimensional interpretation, the interpreter can slice through the three-dimensional volume in many ways, including vertical zig-zag sections, slices at constant travel time, slices along picked horizons, slices parallel to faults, and combinations of horizontal and vertical displays. Workstations also make it easier to incorporate geologic, well-log, engineering, and production data into interpretation. The result is a much more complete and precise interpretation than conventional methods produce.

Seismic methods can also be applied to petroleum engineering. A number of techniques, including three-dimensional, delineate oil and gas fields in enough detail to permit increased recovery. Correlation of seismic amplitude with porosity, permeability, and other rock properties permits a more complete description of inhomogeneities, which affect fluid flow through reservoirs. Seismic methods, especially three-dimensional methods, had major impact on the discovery and production of oil and gas in the 1990s. This was especially important in revealing pools of hydrocarbons that had been missed or bypassed and thus increasing production from,

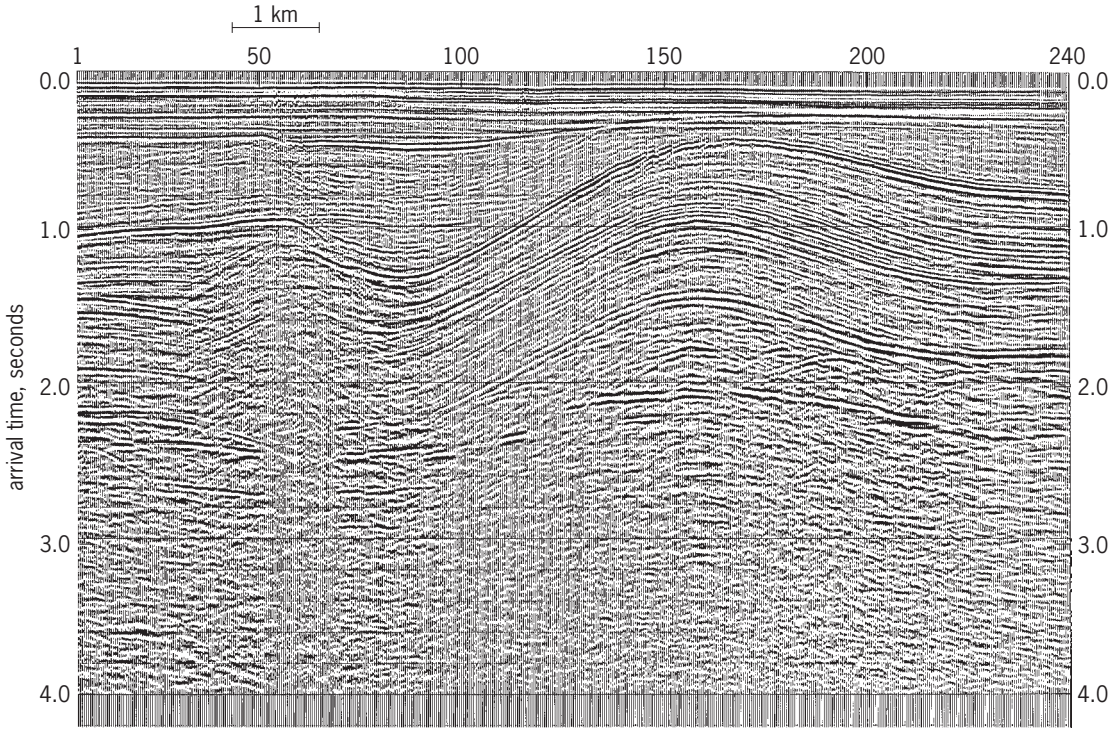

Fig. 11. Seismic record section from the North Sea. The layer from about 2 to 2.3 s at the right edge is composed primarily of salt that has undergone plastic flow. The folding of the rock formations is attributed to this salt movement. The upfolding (an anticline) has produced a trap containing gas at 1.1 s. The scale at the top refers to source points, based on a specified beginning point. 1 km = 0.6 mi. (*Grant Geophysical*)

and lengthening the life of, oil and gas fields. Seismic methods have the potential under some circumstances to monitor the flow of fluids as production progresses. Such information would permit improved hydrocarbon recovery.

Electrical and electromagnetic exploration. Variations in the conductivity or capacitance of rocks form the basis of a variety of electrical and electromagnetic exploration methods, which are used in metallic mineral prospecting and in engineering, ground-water, and other surveys with shallow objectives. Both natural and induced direct current and low-frequency alternating currents are measured in ground surveys, and ground and airborne electromagnetic surveys involving the lower radio frequencies are made.

Natural currents in the Earth, called telluric current, affect large areas. The current density of telluric currents varies with rock conductivity. Comparisons are made between readings observed simultaneously at various locations and at a reference location in order to determine resistivity differences between the locations.

Changes in electrical current flow give rise to associated magnetic fields, and the converse is also true, according to Maxwell's equations. Natural currents are somewhat periodic. Magnetotellurics involves the simultaneous measurement of natural electrical and magnetic variations from which the variation of conductivity with depth can be determined. *See* GEOELECTRICITY; ROCK, ELECTRICAL PROPERTIES OF.

Certain mineral ores store energy as a result of current flow and, after the current is stopped, transient electrical currents flow. This phenomenon is called induced polarization. Observations of the rate of decay of these transient currents are studied in time-domain methods.

Alternating currents tend to flow along the surface of conductors rather than in their interior. The thickness which contains most of the current is called the skin depth. The skin depth, in meters, is given by $(2/\sigma\mu\omega)^{1/2}$, where σ is the conductivity in mhos per meter, μ is permeability in henrys per meter, and ω is angular frequency in radians per second. Since the skin depth becomes greater as frequency becomes lower, measurements at different frequencies give information on the variation of conductivity with depth. Methods in which apparent resistivity is determined as a function of frequency are called frequency-domain methods.

Direct-current and low-frequency alternating-current ground surveys are carried out with a pair of current electrodes, by which electrical current is introduced into the ground, and a pair of potential electrodes across which the voltage is measured. The equipment is often simple, consisting essentially of a source of electrical power (battery or generator), electrodes and connecting wires, ammeter, and voltmeter. A key problem here is providing equipment that will generate enough electrical or electromagnetic energy in the ground but is reasonably portable. The depth of current penetration depends on the geometry of disposition of electrodes, on the frequency used, and on the conductivity distribution. There are two basic types of measurement: (1) electrical sounding, wherein apparent resistivity is measured as the electrode separation is increased—these measurements depend mainly on the variation of electrical properties with depth; (2) electrical profiling, in which variations are measured as the electrode array is moved from location to location.

Electromagnetic methods generally involve a transmitting coil, which is excited at a suitable frequency, and a receiving coil, which measures one or more elements of the electromagnetic field. The receiving coil is usually oriented in a way that minimizes its direct coupling to the transmitter, and the residual effects are then caused by the currents that have been induced in the ground. A multitude of configurations of transmitting and receiving antennas are used in electromagnetic methods, both in ground surface and airborne surveys. In airborne surveys, the transmitting and receiving coils and all associated gear are carried in an aircraft, which normally flies as close to the ground as is safe. Airborne surveys often include multisensors, which may record simultaneously electromagnetic, magnetic, and radioactivity data along with altitude and photographic data. Sometimes several types of electromagnetic configurations or frequencies are used.

The effective penetration of most of the electromagnetic methods into the Earth is not exceptionally great, but they are used extensively in searching for mineral ores within about 300 ft (100 m) of the surface. Electrical methods are effective in exploring for ground water and in mapping bedrock, as at dam sites. They are used also for detecting the position of buried pipelines and in land-mine detection and other military operations.

Ground-penetrating radar can provide images of shallow features in a manner analogous to seismic methods, and it is increasingly used in environmental studies to locate waste dumps and other features that may be pollution threats, in studying archeological sites, and for other applications. Ground-penetrating radar reflects because of changes in electrical resistivity, especially from metallic objects such as drums of waste. The principal feature limiting such radar penetration into the earth is the presence of ground water, so this radar is used more in arid regions than where the water table is near the surface.

Radioactivity exploration. Natural radiation from the Earth, especially of gamma rays, is measured both in land surveys and airborne surveys. Natural types of radiation are usually absorbed by a few feet of soil cover, so that the observation is often of diffuse equilibrium radiation. The principal radioactive elements are uranium, thorium, and the potassium isotope ^{40}K. Radioactive exploration has been used primarily in the search for uranium and other ores, such as niobium (columbium), which are often associated with them, and for potash deposits. The scintillation counter is generally used to detect and measure the radiation. *See* SCINTILLATION COUNTER.

Remote sensing. Measurements of natural and induced electromagnetic radiation made from

high-flying aircraft and Earth satellites are referred to collectively as remote sensing. This comprises both the observation of natural radiation in various spectral bands, including both visible and infrared radiation, and measurements of the reflectivity of infrared and radar radiation. *See* REMOTE SENSING.

Well logging. A variety of types of geophysical measurements are made in boreholes, including self-potential, electrical conductivity, velocity of seismic waves, natural and induced radioactivity, and temperature variations. Borehole logging is used extensively in petroleum exploration to determine the characteristics of the rocks that the borehole has penetrated, and to a lesser extent in mineral exploration.

Measurements in boreholes are sometimes used in combination with surface methods, as by putting some electrodes in the borehole and some on the surface in electrical exploration, or by putting a seismic detector in the borehole and the energy source on the surface. *See* WELL LOGGING. R. E. Sheriff

Bibliography. A. R. Brown, *Interpretation of Three-Dimensional Seismic Data*, 5th ed., 1999; M. B. Dobrin and C. Savit, *Introduction to Geophysical Prospecting*, 1988; R. E. Sheriff, *Encyclopedic Dictionary of Exploration Geophysics*, 3d ed., 1991; R. E. Sheriff, *Geophysical Methods*, 1989; R. E. Sheriff and L. P. Geldart, *Exploration Seismology*, 2d ed., 1995; W. M. Telford et al., *Applied Geophysics*, 1990; K. H. Waters, *Reflection Seismology: A Tool for Energy Resource Exploration*, 3d ed., 1992.

Geophysical fluid dynamics

The branch of physics that studies the dynamics of naturally occurring large-scale flows in both the atmosphere and oceans. Examples of such flows are weather patterns, atmospheric fronts, the Gulf Stream, coastal upwelling, and El Niño. The fluids are either air or water in a moderate range of temperatures and pressures. *See* EL NIÑO; GULF STREAM.

Because of their large scale (from tens of kilometers up to the size of the planet), geophysical flows are strongly influenced by the diurnal rotation of the Earth, which is manifested in the equations of motion as the Coriolis force. Another fundamental characteristic is stratification, that is, density heterogeneity within the fluid in the presence of the Earth's gravitational field, which is responsible for buoyancy forces. Thus, geophysical fluid dynamics may be considered to be the study of rotating and stratified fluids. It is the common denominator of dynamical meteorology and physical oceanography. *See* CORIOLIS ACCELERATION; EARTH; METEOROLOGY; OCEANOGRAPHY.

Historical development. Although geophysical fluid dynamics first became recognized as a scientific discipline in the late 1950s, its foundations can be traced to developments in fluid mechanics over a much longer period of time. During the nineteenth and early twentieth centuries, a number of mathematicians (for example, P. S. de Laplace), fluid dynamicists (Lord Kelvin, H. von Helmholtz, G. I. Taylor), meteorologists (L. F. Richardson, V. Bjerknes, C. G. Rossby), and oceanographers (A. Defant, V. W. Ekman) made significant contributions to the study of large-scale flows that later became incorporated in the discipline of geophysical fluid dynamics. *See* FLUID MECHANICS.

During World War II the demand for improved forecasts of the weather and sea state provided a major impetus and resulted in a new generation of dynamical meteorologists and physical oceanographers who continued to make contributions long after the war ended. The main catalyst, however, was the launching of the geophysical fluid dynamics Summer Program at the Woods Hole Oceanographic Institution, in Massachusetts, which began in 1959.

Dynamics of rotating flows. The first of the two distinguishing attributes of geophysical fluid dynamics is the effect of the Earth's rotation. Because geophysical flows are relatively slow and spread over long distances, the time taken by a fluid particle (be it a parcel of air in the atmosphere or water in the ocean) to traverse the region occupied by a certain flow structure is comparable to, and often longer than, a day. Thus, the Earth rotates significantly during the travel time of the fluid, and rotational effects enter the dynamics. Fluid flows viewed in a rotating framework of reference are subject to two additional types of forces, namely the centrifugal force and the Coriolis force. (Properly speaking, these originate not as actual forces but as acceleration terms to correct for the fact that viewing the flow from a rotating frame—the rotating Earth in the case of geophysical fluid dynamics—demands a special transformation of coordinates.) Contrary to intuition, the centrifugal force plays no role on fluid motion because it is statically compensated by the tilting of the gravitational force caused by the departure of the Earth's shape from sphericity. Thus, of the two, only the Coriolis force acts on fluid parcels. History reveals that Laplace wrote the correct equations of fluid flow in a rotating frame for a study of ocean tides some decades before G. G. de Coriolis, but somehow the name of the latter has remained attached to the terms representing the rotational effect in the equations of motion. *See* EARTH ROTATION AND ORBITAL MOTION.

The importance of the Coriolis force in the dynamics is measured by the Rossby number, Ro [Eq. (1)], where U is a typical velocity value within

$$\mathrm{Ro} = \frac{U}{fL} \tag{1}$$

the flow, $f = 2\Omega \sin$ (latitude) the Coriolis parameter [the Earth's angular rotation rate $\Omega = 7.29\ 10^{-5}$/s], and L a representative distance along the flow. The smaller this number, the stronger the effect of rotation on the flow.

At low speeds or long length scales, not atypical of geophysical flows, the Rossby number may fall far below unity. In such case, the rotating effects

become dominant, and the balance of horizontal forces is primarily an equilibrium between the pressure-gradient force and the Coriolis force [Eqs. (2) and (3)]. Here ρ is the density, f the Coriolis

$$-\rho f v = -\frac{\partial p}{\partial x} \tag{2}$$

$$+\rho f u = -\frac{\partial p}{\partial y} \tag{3}$$

parameter, u and v the eastward and northward velocity components respectively, p the pressure, and x and y the eastward and northward coordinates respectively. This balance is called geostrophy. *See* GEOSTROPHIC WIND.

One remarkable property of geostrophic flows is their vertical rigidity: All fluid parcels on the same vertical share the same velocity and therefore move as one rigid column. This is known as the Taylor-Proudman theorem. A corollary is another property: If there is any vertical velocity, fluid columns must go up and down without shrinking or stretching, occasionally forcing isolation of entire fluid columns above bottom bumps or dips, called Taylor columns. In the atmosphere and oceans, this property is manifested by the tendency of slow flows to follow topographic contours.

When the Rossby number is less than unity but not extremely small, the dynamics are less degenerate, and flow fields can exhibit a number of time-dependent behaviors, including inertia-gravity waves and vorticity waves. Inertia-gravity waves are the classical surface gravity waves with a modification due to the Coriolis effect. Their frequency-wavenumber relationship is shown in Eq. (4). Here g is the Earth's

$$\omega = \sqrt{gbk^2 + f^2} \tag{4}$$

gravitational acceleration, b the fluid's resting depth, f the Coriolis parameter, and k the wavenumber (2π divided by the wavelength). One property of these waves is that, unlike their purely gravitational counterparts, their frequency cannot fall below a minimum (f), lest they become evanescent (exponential instead of periodic behavior in one of the spatial directions). The Kelvin wave is the special case of a wave that is propagating along a boundary (shoreline) and evanescent away from it. This wave also has the property of having no velocity component transverse to the direction of propagation.

The other kind of waves is specific to rotating fluids. Instead of gravity, their restoring force is a vorticity gradient, which is created either by the Earth's curvature or by a bottom slope. There are thus two kinds of vorticity waves: planetary and topographic waves. Synoptic weather formations at midlatitudes share characteristics with planetary waves, whereas topographic waves tend to be associated with airflow over mountain ranges and ocean currents along the continental slope. A third source of vorticity gradient is a gradient in the horizontal shear of a horizontal flow. In this case, however, a portion of the kinetic energy stored in the flow is easily transferred to the wave, and the result is a growing wave that greatly distorts the original flow field. This is called barotropic instability. *See* WAVE MOTION; SURFACE WAVES.

In the vicinity of a top or bottom boundary, friction can become important, and the combination of rotation with friction leads to the development of Ekman layers. These layers differ from their cousins in classical fluid mechanics by having a well-defined thickness and a helicoidal flow field. An unexpected result is the existence of a significant angle between the frictional stress causing the boundary layer and the direction of the fluid transport within it. *See* FRICTION.

Dynamics of stratified flows. Variations of moisture in the atmosphere, of salinity in the ocean, and of temperature in either can modify the density of the fluid to such an extent that buoyancy forces become comparable to other existing forces. The fluid then has a strong tendency to arrange itself vertically so that the denser fluid sinks under the lighter fluid. The resulting arrangement is called stratification, the second distinguishing attribute of geophysical fluid dynamics. The greater the stratification in the fluid, the greater the resistance to vertical motions, and the more potential energy can affect the amount of kinetic energy available to the horizontal flow. A practical measure of stratification is the Brunt-Vaisala frequency, N, defined from its square according to Eq. (5), where $\rho(z)$ is the density function of the

$$N^2 = -\frac{g}{\rho}\frac{d\rho}{dz} \tag{5}$$

vertical z and decreasing upward ($d\rho/dz < 0$). In many ways, the Brunt-Vaisala frequency N is to stratified fluids what the Coriolis parameter f is to rotating fluids, and a close analogy can be developed between stratified and rotating fluids—namely, vertical properties of the latter are similar to horizontal properties of the former. The dimensionless number of stratified flows corresponding to the Rossby number of rotating flows is the internal Froude number, Fr [Eq. (6)], where H is the height of the fluid region

$$\mathrm{Fr} = \frac{U}{NH} \tag{6}$$

under consideration. *See* DENSITY.

Just as the water-air density discontinuity can support gravity waves, the gradual density variation of a stratified fluid can support gravitational waves, called internal waves. Like their surface counterparts, internal waves can be affected by the Earth's rotation, and unlike them, they can propagate not only horizontally but also vertically. Internal waves are ubiquitous in the ocean, where their frequency spectrum very often displays a standard structure suggesting saturation. In the atmosphere, mountain waves and cloud rolls are manifestations of internal waves.

On occasion, a source of energy is sufficiently intense to create vertical motions that overcome the

stratification, and the result is vertical mixing over a partial extent of the fluid system. A common situation is one where turbulence induced by a shear stress in the proximity of a boundary generates a mixed boundary layer. In the ocean, a mixed layer is created almost daily in the top few meters under the action of surface winds, and it also exists at the bottom of the ocean wherever near-bottom currents can generate enough turbulence to overcome the stratification. *See* TURBULENT FLOW.

Another mechanism capable of erasing an existing stratification is thermal convection. Whenever a fluid is heated from below or cooled from above, it becomes top heavy, and negative buoyancy forces generate velocities that rearrange the fluid toward a more stable state. During daytime, the ground is warmed by the Sun and effectively heats the atmosphere from below. The result is a state of convection that gradually erases the previous nocturnal stratification and creates a well-ventilated layer called the atmospheric boundary layer. The turbulence accompanying this state of convection greatly facilitates the dispersal of pollution. In lakes and oceans, heat loss at the surface during winter generates convection that can mix the water over several tens or hundreds of meters, depending on the intensity of the cooling and on the strength of the stratification created during the previous summer. *See* CONVECTION (HEAT).

Dynamics of rotating and stratified flows. A quantity central to the understanding of geophysical flows, which are simultaneously rotating and stratified, is the potential vorticity, q. This quantity incorporates both rotation and stratification [Eq. (7)] and has the

$$q = \left(f + \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)\frac{\partial \rho}{\partial z} + \left(\frac{\partial u}{\partial z}\frac{\partial \rho}{\partial y} - \frac{\partial v}{\partial z}\frac{\partial \rho}{\partial x}\right) \quad (7)$$

property of being nearly conserved by fluid parcels as they travel with the flow. One situation in which potential-vorticity conservation rules the fate of the fluid is geostrophic adjustment. This occurs whenever a heterogeneous fluid initially away from equilibrium seeks a lower level of energy; potential-vorticity conservation then limits the amount of vertical stretching or squeezing that every parcel can undergo, and the result is a final state with a residual motion in geostrophic balance. Many atmospheric and oceanic density fronts exhibit attributes of geostrophic adjustment.

The interplay between rotation and stratification effects creates a preferential length scale at which a large portion of the energy available in geophysical flows tends to be concentrated. This privileged length, L, called the radius of deformation, arises when the Rossby and Froude numbers are equal [Eq. (8)]. Numerous structures such as jets, vortices,

$$L = \frac{NH}{f} \quad (8)$$

and fronts have horizontal dimensions that are comparable to the radius of deformation. In the atmosphere, the radius of deformation sets the size of synoptic weather patterns (500–1000 km; 310–620 mi), whereas in the ocean it is usually smaller (10–50 km; 6–30 mi) and related to the width of currents such as the Gulf Stream.

Because they do not correspond to a state of absolute minimal energy, geostrophically adjusted states retain a certain amount of energy that can still be made available for subsequent motions. The process by which a geostrophic state relaxes toward a yet lower energy level is called baroclinic instability. The instability begins with a horizontal meandering of the flow that has a vertical phase shift creating a mutual reinforcement of the meanders at the different vertical levels of the fluid. Growing meanders eventually detach as vortices, and the ultimate situation is a complex pattern of a wavering flow embedded in an assortment of interacting vortices. The preferential length scale of the meanders and vortices is again the radius of deformation.

Geophysical flows are replete with vortices, resulting from baroclinic instability. Their interactions generate highly complex flows not unlike those commonly associated with turbulence. Unlike classical fluid turbulence, however, geophysical flows are wide and thin (with, furthermore, a high degree of vertical rigidity as a result of rotational effects), and their turbulence is nearly two-dimensional.

Applications. Geophysical fluid dynamics finds much of its use in dynamical meteorology and physical oceanography. In meteorology, geophysical fluid dynamics has been the key to understanding the essential properties of midlatitude weather systems, including the formation of cyclones and fronts. Geophysical fluid dynamics also explains the dynamical features of hurricanes and tornadoes, sea and land breezes, the seasonal formation and break-up of the polar vortex that is associated with high-latitude stratospheric ozone holes, and a host of other wind-related phenomena in the lower atmosphere. *See* CYCLONE; HURRICANE; TORNADO.

In oceanography, successes of geophysical fluid dynamics include the explanation of major oceanic currents, such as the Gulf Stream. Coastal river plumes, coastal upwelling, shelf-break fronts, and open-ocean variability on scales ranging from tens of kilometers to the size of the basin are among the many other marine applications. The El Niño phenomenon in the tropical Pacific is rooted in processes that fall under the scope of geophysical fluid dynamics. *See* EL NIÑO; GULF STREAM.

Highly successful applications of geophysical fluid dynamics have also been made to other planetary-scale fluids, such as the molten rock in the Earth's outer core, Jupiter's Great Red Spot, and convective gases in stars, including the Sun. *See* JUPITER; SUN.

Benoit Cushman-Roisin

Bibliography. B. Cushman-Roisin, *Introduction to Geophysical Fluid Dynamics*, Prentice Hall, 1994; A. E. Gill, *Atmosphere-Ocean Dynamics*, Academic Press, 1982; J. Pedlosky, *Geophysical Fluid Dynamics*, 2d ed., Springer-Verlag, 1987; H. M. Stommel and D. W. Moore, *An Introduction to the Coriolis Force*, Columbia University Press, 1989.

Geophysics

The study of the Earth and its relations to the rest of the solar system using the principles and practices of physics. Geophysics is considered by some to be a branch of geology, by others a branch of physics. It is distinguished from geology by its use of instruments to make direct and indirect measurements of the phenomena being studied in contrast to the more direct observations of geology, and by its concern with other members of the solar system. *See* EARTH; GEOLOGY; SOLAR SYSTEM.

Subdivisions of Geophysics

Geophysics is divided into a variety of subdisciplines. These can be grouped according to the portion of the Earth with which each is concerned, although there is much overlap. Solid-earth geophysics is concerned with the Earth's interior; oceanography and hydrology with the aqueous parts; meteorology with the lower atmosphere; and aeronomy with the upper atmosphere. Because students of the rest of the solar system use methods similar to those employed by geophysicists to study the Earth, geophysics has grown to encompass studies of the other planets, the Sun, and the space in which these bodies and the Earth move.

Solid-earth geophysics. This discipline is subdivided, according to the methods used to study the Earth, into the sciences of geodesy, geologic thermometry, seismology, and tectonophysics.

Geodesy. Geodesy is the science of the shape and size of the Earth. Estimates of its diameter were made at least as early as the second century B.C. A basic gridwork of carefully located points on the Earth's surface was started early in the seventeenth century, and today forms the basis of all surveys and maps of the Earth's surface. *See* GEODESY.

From the dimensions of the Earth and its gravity, its mass can be calculated. The distribution of mass in the interior is determined from the variation of gravity from place to place on and above the surface and from calculations involving the moment of inertia and the velocities of seismic waves in the Earth's interior. Precise measurements of the orbits of satellites map the gravitational field in the space surrounding the Earth. *See* EARTH, GRAVITY FIELD OF.

Geothermal studies. Heat is escaping continually from the Earth's interior by conduction and by volcanic processes. Measurements of the variations in the rate of upward heat flow provide evidence of the processes active in the Earth's interior. Studies of the thermal conductivity of rocks and of the rate of heat generation by disintegration of radioactive trace elements in the rocks make it possible to estimate the temperature distribution in the Earth's interior. The history of temperature variations in the interior can be modeled by using different assumptions of starting conditions and comparing the results with the limited knowledge of present conditions to test ideas on the evolution of the Earth. *See* EARTH, HEAT FLOW IN; GEOLOGIC THERMOMETRY.

Volcanic processes supplement the heat loss and are an important area of study because of their danger to human activities and because they provide information on how ore deposits form. Studies of volcanic processes on the ocean floor have revealed an environment for living creatures whose energy sources are not derived from sunlight, and may yield clues to the early evolution of life on Earth. *See* DEEP-SEA FAUNA; PREBIOTIC ORGANIC SYNTHESIS; VOLCANOLOGY.

Seismology. This is the science of earthquakes and other ground vibrations. Seismological studies may lead to the development of techniques for reliable earthquake prediction. In addition, earthquake studies may provide better understanding of the nature of the resultant motions, so that buildings and other structures can be designed to withstand their vibrations. By using earthquake vibrations, the structure of the Earth's interior and the patterns of deformation of the Earth can be mapped. *See* EARTHQUAKE; SEISMOLOGY.

Study of the times of passage of seismic waves through the Earth's interior is the principal source of information on the distribution of different types of rocks. Studies of the waves' rate of absorption and coefficients of reflection give information on the physical properties of different layers of the Earth. From the motions recorded at places on the surface, it is possible to locate where within the Earth each earthquake occurs, details of the mechanism of generation of the waves, and the amount of energy released. This provides clues as to the processes of change occurring in the Earth's interior which lead to short- and long-range deformation of the surface.

Because large explosions produce seismic waves which are recorded at great distances from the source, much effort has been given to seismic research as a means of monitoring compliance with a nuclear test ban treaty.

Tectonophysics. Also called geodynamics, this is the science of the deformation of rocks. It consists of tectonics, the study of the broader structural features of the Earth and their origins, as in mountain building; and rock mechanics, the measurement of the strength and related physical properties of rocks. *See* CONTINENTS, EVOLUTION OF; GEODYNAMICS; OROGENY; PLATE TECTONICS; ROCK MECHANICS.

Because the increase of temperature and pressure with depth in the Earth profoundly changes the physical behavior of rocks, a proper understanding of processes in the Earth's interior is possible only when based on the results of laboratory experiments carried out under extreme conditions. Pressures at depths greater than a few tens of miles can be duplicated only under very transient conditions, and usually not at the high temperatures which are prevalent. The atomic structure of minerals goes through phase changes at high pressures which result in the occurrence at great depths of mineral varieties different from those encountered at the surface, with resultant uncertainty as to composition of the Earth's deepest layers. The Earth's interior is likely to remain incompletely explored for a long time in spite of the

best efforts of geophysicists. *See* EARTH INTERIOR; HIGH-PRESSURE MINERAL SYNTHESIS.

Hydrospheric geophysics. Water is the compound that makes the Earth unique among large bodies in the universe. It occurs as a gas, a liquid, a solid, and as a component (often a trace element) in rocks. That part of the Earth's water which occurs as a liquid and as ice, largely free to move relatively easily from place to place, is called the hydrosphere. Specialized branches of geophysics have evolved to study water in this region. *See* HYDROSPHERE.

Oceanography. Scientific study of the oceans is concerned with the shape and structure of the ocean basins, the physical and chemical properties of seawater, ocean currents, waves and tides, thermodynamics of the ocean, and the relations of these to the organisms which live in the sea. Knowledge of the ocean is important because it is a source of food, much of the world's commerce travels across its surface, and its heat balance is a major factor affecting weather worldwide. *See* HEAT BALANCE, TERRESTRIAL ATMOSPHERIC; MARINE GEOLOGY; OCEANOGRAPHY.

Hydrology. Fresh water in lakes, in streams, and in the pores of near-surface rocks is the concern of hydrology. Every organism requires water to live. The distribution and purity of the water supply are consequently important. Surface water also carries away much of society's wastes, both sanitary and industrial. Hydrology is concerned with the chemical, physical, and biological processes by which these wastes are changed and removed as water moves through and over the ground toward the oceans. The constant exchange of water between the Earth's interior, the hydrosphere, and the atmosphere is an important aspect of hydrology. The availability of pure water for personal, agricultural, and industrial uses is so essential that pollution control is one of the most rapidly growing scientific and technological problems of modern society. *See* GROUND-WATER HYDROLOGY; HYDROLOGY; WATER POLLUTION.

Glaciology. An appreciable fraction of Earth's water occurs in the form of snow and ice. If all of the water frozen in glaciers were to melt, the ocean surface would rise several hundred feet, changing coastlines substantially. Glaciology, which is often considered a branch of hydrology, is the scientific study of the distribution and movement of this frozen water, how it accumulates and melts, and what it does to the underlying rocks as it flows over them. *See* GLACIOLOGY; TERRESTRIAL WATER.

Meteorology. Meteorology is the study of the composition and movements of the mass of air known as the atmosphere; of the interaction of the atmosphere with living organisms, the hydrosphere, and the solid earth; and of the flow of energy within the atmosphere and to and from the space beyond. An important goal of meteorology is to predict changes in atmospheric conditions, the weather and climate, from observations of current and past conditions and theoretical calculations based on models of how the atmosphere behaves. For this purpose, temperature, pressure, and the moisture conditions in the atmosphere are measured continuously or periodically at a large number of places at and above the Earth's surface, reported to data-processing centers, and used to make regular weather predictions. These are widely distributed by government agencies, private companies, and the mass media, and are used to guide individual and organizational activities. *See* AGRICULTURAL METEOROLOGY; CLIMATE MODIFICATION; INDUSTRIAL METEOROLOGY.

Short-term weather predictions attempt to describe conditions to be expected within a few hours or days; but attempts are made at predicting long-term trends for weeks, months, or even years ahead as well. Data gathered by satellites orbiting the Earth play an important role in making predictions and understanding atmospheric processes. It has become the responsibility of government to warn the public of upcoming extreme weather. *See* ATMOSPHERE; MESOMETEOROLOGY; METEOROLOGY; SATELLITE METEOROLOGY; WEATHER FORECASTING AND PREDICTION; WEATHER MODIFICATION.

Aeronomy. There is no distinct upper boundary to the atmosphere, which becomes progressively less dense with elevation, eventually merging with the Sun's extended atmosphere through which Earth moves. However, the properties of the atmosphere undergo a radical change at an elevation of about 60 mi (100 km), where the atmosphere becomes so highly ionized that its electrical properties become important in controlling its behavior. Aeronomy is the science of this upper part of the atmosphere. Solar radiations are mainly responsible for this ionization, although cosmic rays play a role. Molecules are dissociated into their component atoms, and selective diffusion results in an upward concentration of hydrogen, which may be lost continuously from the Earth. Free electrons as well as ions are present. The electrical particles move, forming currents which induce fluctuations of the Earth's magnetic field. *See* COSMIC RAYS; IONOSPHERE; UPPER-ATMOSPHERE DYNAMICS.

The impact of charged particles from the Sun can be so violent as to produce magnetic storms which disrupt radio communication. It also produces the aurora in high latitudes. Farther out, moving charged particles carried by the Earth in its motion around the Sun form a layer called the magnetosphere, whose shape is distorted by the flow of material radiated from the Sun. The charged layers of the atmosphere protect the Earth from the strongest of the Sun's radiations. Knowledge of the nature and variability of these radiations and their effects on the Earth and its inhabitants has been accumulating rapidly only in the last few decades, and much remains to be discovered. Exploration of the upper atmosphere and the outlying space became possible only with development of rocket probes and placing of artificial satellites in orbits around the Earth. As these researches into the outer fringes of the Earth's atmosphere have progressed, they have overlapped astronomical studies of the Sun to such a degree that solar science has become largely accepted as a part of geophysics. *See* AERONOMY; AURORA; MAGNETOSPHERE; SOLAR WIND; SUN.

Planetology. Until the exploration of space using rockets, the only information available about planets and natural satellites was obtained by use of astronomical telescopes. With the advent of rocket probes, the same diversity of measurements made on the Earth could be carried out on each of the other bodies. Because the methods of measurement used are similar to those used by geophysicists to examine the Earth, studies of the other planets and their moons have become a branch of geophysics known as planetology. *See* MOON; PLANET; PLANETARY PHYSICS.

Overlapping Fields

In addition to the regional subdivisions of geophysics defined above, there are other overlapping distinct specialties.

Geomagnetism. Geomagnetism is concerned with a detailed description of the Earth's magnetic field and its changes and with the magnetic properties of rocks. Because the Earth's magnetic field is used as a guide for navigation and to locate north in surveying, knowledge of its current patterns and changes both at the Earth's surface and in the surrounding space is important. The magnetism of rocks is also useful in delineating the history of changes in the Earth, providing the principal line of evidence showing how the present patterns of continents and oceans have evolved from very different arrangements in the past. Variations in magnetic field strength are also useful in mapping variations in the distribution of rocks of different compositions.

Changes in the magnetic field induce electric currents in the Earth's interior and in the atmosphere. Because magnetism and electricity are so closely linked, geoelectricity is generally considered a part of geomagnetism. Patterns of electric currents can be used to map the variations in the electrical conductivity of buried rocks, providing evidence of the Earth's internal structure. *See* GEOELECTRICITY; GEOMAGNETISM; ROCK MAGNETISM.

Geochronology. This field deals with the dating events in the Earth's history. The principal technique is based on radioactive disintegrations: the proportion of daughter to parent elements in a mineral or rock is a measure of the age of the material. Other methods depend on the redshift of the spectra of distant stars, the rate of recession of the Moon, annual variations in the growth rates of plants and animals including fossils, and the rates of erosion and sedimentation. *See* DATING METHODS; GEOCHRONOMETRY.

Geocosmogony. This is the study of the origin of the Earth. The many hypotheses fall into two groups: those which postulate that the Earth is primarily an aggregate of once smaller particles, and those which claim that it is essentially a fragment of a larger body. Current speculation favors the former theory. Geocosmogony is initimately linked with the origin of the solar system and the Milky Way Galaxy. Many lines of evidence suggest that the formation of the Earth was a typical minor event in the evolution of the Milky Way or of the universe as a whole, occurring 5–10 $\times$ 10^9 years ago. *See* COSMOCHEMISTRY; MILKY WAY GALAXY.

Exploration and prospecting. Geophysical techniques are widely used not only to study the general structure of the Earth but also to prospect for petroleum, mineral deposits, and ground water and to map the sites of highways, dams, and other structures. Seismic methods are the most widely used, but electrical, electromagnetic, gravity, magnetic, and radioactivity surveying methods are also well developed. Many types of geophysical surveys can be made by lowering measuring apparatus into boreholes. *See* GEOPHYSICAL EXPLORATION; PROSPECTING; WELL LOGGING.

B. F. Howell, Jr.

Bibliography. J. De Bremaecker, *Geophysics: The Earth's Interior,* 1985; C. M. R. Fowler, *The Solid Earth: An Introduction to Global Geophysics*, 2d ed., 2004; M. N. Hill et al., *The Sea*, vols. 1–8, 1962–1983; P. Kearey, M. Brooks, and I. Hill, *An Introduction to Geophysical Exploration*, 3d ed., 2002; F. K. Lutgens and E. J. Tarbuck, *The Atmosphere: An Introduction to Meteorology*, 10th ed., 2006; W. M. Kaula, *An Introduction to Planetary Physics*, 1976; P. V. Sharma, *Geophysical Methods in Geology*, 2d ed., 1986.

Geostrophic wind

A hypothetical wind based upon the assumption that a perfect balance exists between the horizontal components of the Coriolis force and the horizontal pressure gradient force per unit mass, with the implication that viscous forces and accelerations are negligible. Application of the geostrophic wind facilitates an approximation of the wind field from the pressure data over vast regions in which few wind observations are available.

Bases of the approximation. The geostrophic wind blows parallel to the isobars (lines of equal pressure) with lower pressure to the left of the direction of the wind in the Northern Hemisphere and to the right in the Southern Hemisphere. Its speed is given by the equation below, where ρ is the density of air, Ω is

$$V_{\text{geo}} = \frac{1}{2\rho\Omega \sin\phi}\frac{\partial p}{\partial n}$$

the angular speed of rotation of the Earth about its axis, p is the atmospheric pressure, ϕ is the latitude, and n is a coordinate normal to the isobars and directed toward higher pressure. The approximation now known as the geostrophic wind was first derived empirically by C. H. D. Buys-Ballot in 1857 and has been known as Buys-Ballot's law.

The geostrophic wind represents a good approximation to the actual wind at elevations greater than about 3000 ft (1 km), except in instances of strongly curved flow, large variations in wind speed, and in the vicinity of the Equator.

Thermal wind. This is a term denoting the net change in the geostrophic wind over some specific vertical distance. This change arises because the rate of change of pressure in the vertical is different in two

air columns of different air density, so that the horizontal component of the pressure gradient force per unit mass varies in the vertical. The thermal wind is directed approximately parallel to the isotherms of air temperature with cold air to the left and warm air to the right in the Northern Hemisphere, and vice versa in the Southern Hemisphere. Thus, for example, the increasing predominance of westerly winds aloft may be viewed as a consequence of the warmth of tropical latitudes and the coldness of polar regions. *See* CORIOLIS ACCELERATION; GRADIENT WIND; WIND; WIND STRESS.

Frederick Sanders; Howard B. Bluestein

Geosynthetic

Any synthetic material used in geotechnical engineering. *See* MANUFACTURED FIBER.

Geotextiles

Geotextiles are used with foundations, soils, rock, earth, or other geotechnical material as an integral part of a manufactured project, structure, or system. These textile products are made of synthetic fibers or yarns, constructed into woven or nonwoven fabrics that weigh from 3 to 30 oz/yd^2 (100 to 1000 g/m^2). Geotextiles are more commonly known by other names, for example, filter fabrics, civil engineering fabrics, support membranes, and erosion control cloth.

Permeable geotextiles perform three basic functions in earth structures: separation (the fabric provides a boundary that segregates materials); reinforcement (the fabric imparts tensile strength to the system, thereby increasing its structural stability); and filtration (the fabric retains soil particles while allowing water to pass through). Such geotextiles can thus be adapted to numerous applications in earthwork construction. The major end-use categories are stabilization (for roads, parking lots, embankments, and other structures built over soft ground); drainage (of subgrades, foundations, embankments, dams, or any earth structure requiring seepage control); erosion control (for shoreline, riverbanks, steep embankments, or other earth slopes to protect against the erosive force of moving water); and sedimentation control (for containment of sediment runoff from unvegetated earth slopes). No one geotextile is suited for all these applications. Each use dictates a specific fabric requirement to resist installation stresses and to perform its function once installed.

Stabilization applications. Soft or low-strength soils on a project site present costly problems in the construction and maintenance of haul roads, storage yards, railroads, and other areas which must support vehicular traffic. Poor soil conditions also cause rapid deterioration of paved and unpaved roads, city streets, and parking lots. Problems are caused by subgrade deformation and intermixing between subgrade soil and aggregate

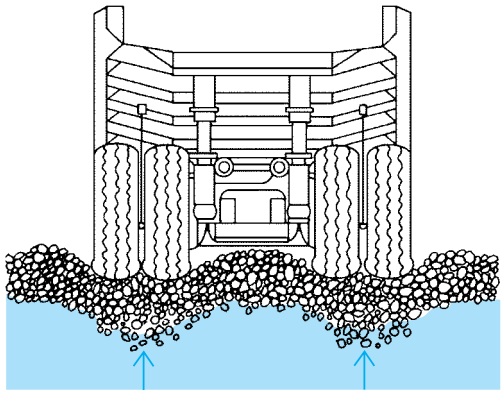

(a)

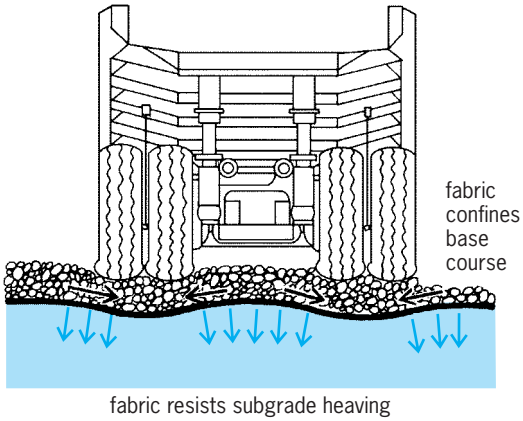

(b)

Fig. 1. Geotextiles are used to stabilize soils. (*a*) A rutted and unstable surface results from subgrade deformation and intermixing between subgrade soil and aggregate base. (*b*) Geotextiles resist rutting through separation and reinforcement (*from Mirafi Fabrics for the Mining Industry, Celanese Fibers Marketing Co.*). (*c*) A highway base course constructed with geotextile between subgrade and aggregate base (*from Mirafi Family of Construction Fabrics/MPB8, Celanese Fibers Marketing Co.*).

base, resulting in a rutted and unstable surface that impedes or even prohibits the traffic flow (**Fig. 1***a*).

Geotextiles can eliminate or reduce the effect of these soft-soil problems through separation and reinforcement. When placed over a soft soil, the geotextile provides a support membrane for placement and compaction of aggregate base. The fabric barrier prevents aggregate particles from intruding into the soft soil and prevents soil particles from pumping up into the aggregate layer. As a continuous membrane between soil and aggregate base, the geotextile helps confine the aggregate against lateral and vertical movement. This confinement maintains the density and hence the load-distributing characteristics of the aggregate. The fabric also resists the upward heaving of subgrade between wheel paths. If the subgrade is extremely soft and cannot support vehicle loads, the fabric will act as a reinforcing membrane to assist the subgrade in supporting loads (Fig. 1*b* and *c*).

Drainage applications. Soil moisture control through drainage is essential to maintain stability in pavements, foundations, cut slopes, and earth dams. Drainage is accomplished by providing a trench or blanket of porous rock for soil moisture to seep into. A perforated pipe is often installed within the porous rock to collect the moisture and transport it to an outlet. To ensure effective performance and long life, drain structures need a filter to retain soil particles that would clog the drain.

Graded aggregate filters are conventionally used. A properly designed graded aggregate will confine or retain the soil, thus preventing significant particle movement. If, however, the drainage aggregate is too coarse, the voids between rocks at the soil-aggregate interface will be too large for soil particles to bridge across. The resulting lack of soil-particle confinement will result in erosion when water seeps out of the soil, that is, soil piping (**Fig. 2***a*). *See* GROUND-WATER HYDROLOGY.

The pore structure and permeability of some geotextiles are similar to those of graded aggregate filters. These fabrics can provide the same particle retention at the soil-drain interface while permitting unrestricted flow of water from the adjacent soil. Geotextiles can thus eliminate the need for graded aggregate filters in drains. When drainage fabric is used, no special aggregate gradation is required, because the fabric prevents soil from washing into the drain (Fig. 2*b* and *c*). *See* DAM; FOUNDATIONS; PAVEMENT.

Erosion control applications. Embankments along coastal shorelines and inland waterways are subjected to wave and current action that can cause severe erosion, instability, and even destruction of the earth slope. To protect against erosion, earth slopes have often been covered with armor (riprap, concrete blocks, concrete slabs). Despite the armor covering, erosion can still occur, undermining the armor's foundation. To assure long-term stability and performance, the erosion-control structure must include a barrier that shields the soil surface from scouring. This barrier should be permeable so that any moisture seeping from the soil slope can escape without buildup of hydrostatic pressure. Traditionally, granular filters of specially graded sand, gravel, or stone have been used to prevent slope erosion beneath the armor. But granular filters are expensive, particularly when not locally available, and even when properly installed are subject to erosive forces that can wash them away.

Some geotextiles are ideal erosion control barriers. Erosion control fabrics will shield a soil slope from the erosive force of moving water, and they are permeable so that seepage from the earth slope can pass through freely. These fabrics will remain intact, covering the soil slope as long as the armor stone remains above it. *See* EROSION.

Sedimentation control. Severe erosion can occur during earthwork construction or mining operations when protective vegetation is removed and soil slopes are left temporarily unprotected. Resultant sediment runoff can create serious downstream damage, for example, contaminated waterways, clogged storm drains, or sediment-covered forests or

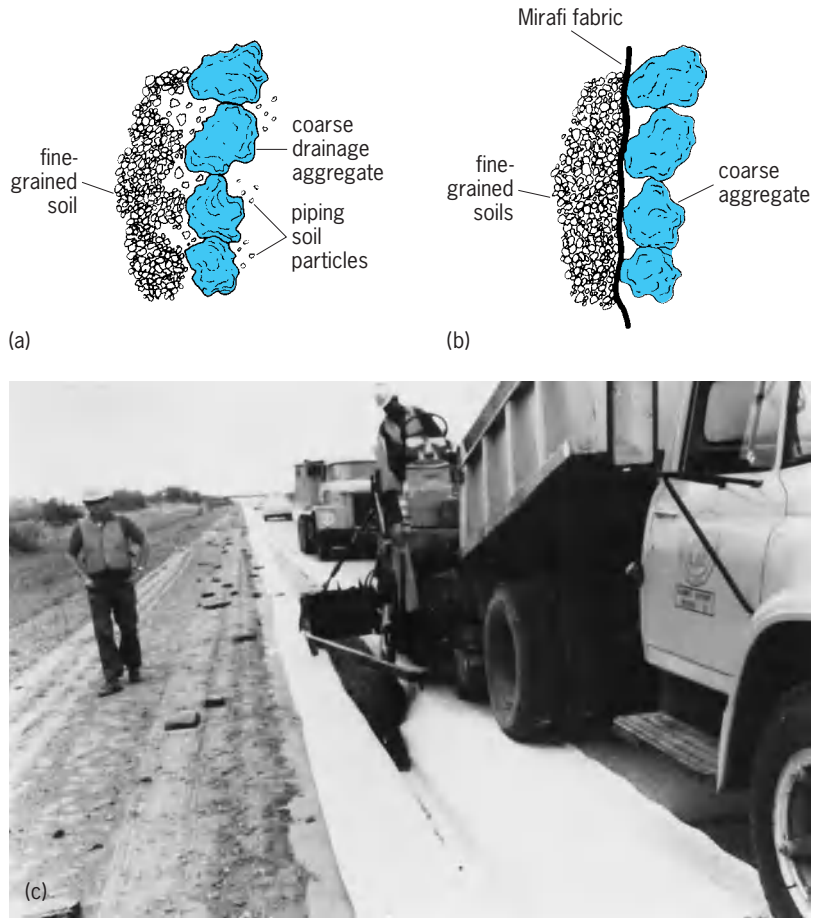

Fig. 2. Geotextiles find wide use for drainage. (*a*) Without a filter, soil particles may wash into and clog the drain. (*b*) Drainage fabric provides retention at the soil-drain interface, while allowing water to pass through (*from Mirafi Fabrics for the Mining Industry, Celanese Fibers Marketing Co.*). (*c*) Drain trench lined with a geotextile and backfilled with coarse aggregate.

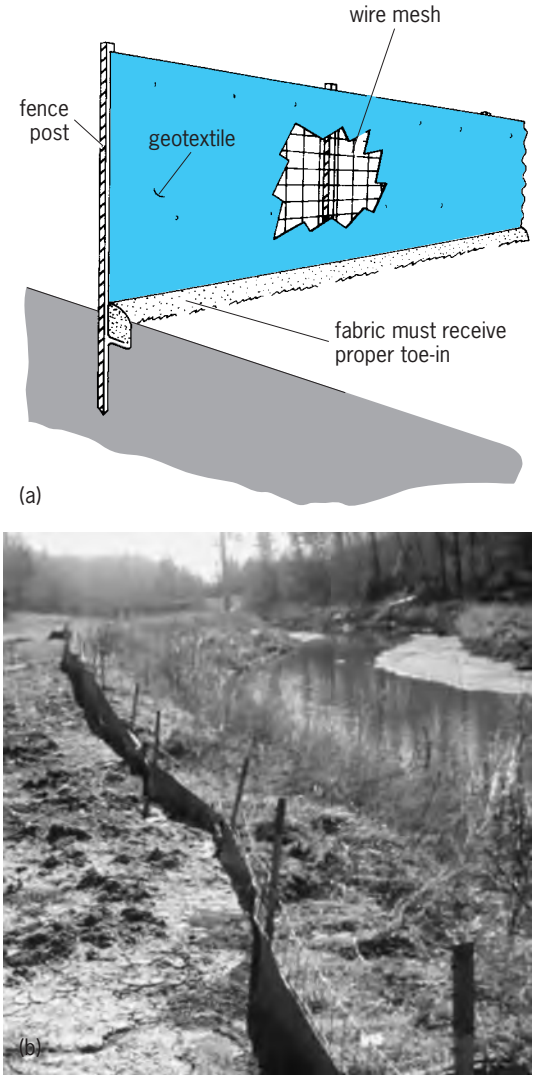

Fig. 3. A silt fence is a fabric-lined fence structure. When lined with a geotextile, it retains sediment runoff, preventing contamination of adjacent waterways. (*Mirafi Fabrics for the Mining Industry, Celanese Fibers Marketing Co.*)

pastures. Government agencies at the federal and state level have recognized the problems associated with sedimentation, and legislation has been passed requiring the control of sediment runoff from any disturbed land area. As a result, earthwork contractors and mine operators are faced with the responsibility and cost of controlling sediment runoff from work sites.

Silt fences constructed with geotextiles offer a cost-effective solution to sedimentation control. A silt fence is a fabric-lined structure installed on, or at the base of, an unvegetated slope. It acts as a water-permeable barrier that retains sediment runoff from the slope. The silt fence can be thought of as a temporary impoundment structure that forms a sediment pond aboveground (**Fig. 3**). When installed along the perimeter of a construction site, a silt fence can prevent sediment from leaving the disturbed area. By installing silt fences along stream banks, sediment can be kept from reaching the waterway.

Robert G. Carroll, Jr.

Geomembranes

A geomembrane is any impermeable membrane used with soils, rock, earth, or other geotechnical material in order to block the migration of fluids. These membranes are usually made of synthetic polymers in sheets ranging from 0.01 to 0.14 in. (0.25 to 3.5 mm) thick. Geomembranes are also known as flexible membrane liners, synthetic liners, liners, or polymeric membranes.

Early liners included clay, bentonite, cement-stabilized sand, and asphalt. Such liners, however, have measurable permeability or tend to crack and fissure when exposed to certain environments and chemicals. Traditional soil liners such as clay are regarded as good barriers if their permeability is 10^{-7} cm/s. Geomembranes, by contrast, have no true permeability if there are no leaks or holes in the membrane, and they possess inherent flexibility to accommodate geotechnical settlement and shifting. Migration through a geomembrane takes place at extremely low levels via chemical diffusion. Effective permeabilities for geomembranes can be calculated from measured diffusion rates, and vary from 10^{-11} to 10^{-13} cm/s, depending on the type of polymeric membrane. Modern geomembranes are commonly made of medium-density polyethylenes that are very nearly high-density polyethylenes (HDPE), several types of polyvinyl chloride (PVC), chlorosulfonated polyethylene (a synthetic rubber), ethylene propylene diene monomer (EPDM), and several other materials. Some geomembranes require reinforcement with an internal fabric scrim for added strength, or plasticization with low-molecular-weight additives for greater flexibility. *See* POLYMER.

Applications. Geomembranes are able to contain fluids, thus preventing migration of contaminants or valuable fluid constituents. Since they prevent the dispersal of materials into surrounding regions, geomembranes are often used in conjunction with soil liners, permeable geotextiles, fluid drainage media, and other geotechnical support materials. The major application of geomembranes has been containment of hazardous wastes and prevention of pollution in landfill and surface impoundment construction (**Fig. 4**). They are also used to a large extent in mining to contain chemical leaching solutions and the precious metals leached out of ore, in aquaculture ponds for improved health of aquatic life and improved harvesting procedures, in decorative pond construction, in water and chemical storage-tank repair and spill containment, in agriculture operations, in canal construction and repair, and in construction of floating covers for odor control, evaporation control, or wastewater treatment through anaerobic digestion. *See* HAZARDOUS WASTE.

Basin liners. When geomembranes are used as basin liners, they are often applied in direct contact and on top of soil liners. Such a composite lining system offers extra security for containment because the soil liner backs up the geomembrane. Rolls of geomembrane are deployed downslope, with seams parallel to the slope of the basin. Texture-surface

geomembranes have been developed to enhance the stability of steep slopes.

Final cover. Final caps form covers over waste impoundments or other containments. In final-cover applications, geomembranes prevent intrusion of precipitation into isolated and contained areas and, in addition, prevent the escape of gases from the containment. Often the surface over which the final cover is applied is unstable or subject to settlement. Seaming can take place on floating barges, if necessary, or floated into position in the construction of floating covers.

Final covers over landfills are often subject to deformation from differential settling of the subsoils. Geomembranes capable of extreme elongation and flexibility should be selected to accommodate the differential settlement, so that they can be stretched a great distance before tearing and can be bent (flexed) with ease. Textured-surface geomembranes are desirable for final covers because they impart long-term slope stabilities.

Vertical cut-off walls. Geomembranes are hung vertically in repair of tank linings and in construction of vertical barrier walls in soils. They can be used either alone or in conjunction with bentonite slurry trenches to cut off and isolated sections of ground water in cases of pollution remediation, salt-water intrusion, or dike construction. Geomembrane panels are inserted directly into these cut-off walls to depths of 100 ft (30 m) by using vibratory hammers and sheet pile drivers. The adjacent panels are not joined by heat seaming but are interlocked through prefabricated complementary joint sections; the locking mechanism is engaged as the panels are slid alongside one another into the cut-off wall.

Design and installation. A geomembrane is not intended to provide tensile support or load-carrying capacity. Many applications, in waste containment particularly, require maximum durability and longevity, with a lifetime of as much as hundreds of years. Therefore, in addition to selecting materials that are suitably resistant to chemicals and to wear, it is desirable to limit fatigue-inducing tension stresses through good design and installation practice.

Soil subgrades and cover soils must be free of sharp objects such as sticks and angular stones that would introduce puncture stresses on a linear used as a containment barrier. Heavy equipment used to place covering soils over the synthetic liner should be separated from the geomembrane by a suitable thickness of the soil. If necessary, geotextiles are placed above or below a geomembrane in order to increase puncture resistance.

Seaming. Joining adjacent panels of geomembrane is usually accomplished through the application of heat. Some heat seaming (welding) is done in factories during construction of prefabricated panels. However, for most panel installations, some if not all seams must be constructed in the field. Types of welding include extrusion, hot-wedge, and hot-air.

In extrusion welding, a strand of molten polymer is deposited at the edge of the overlapped geomembrane panels, bonding the sheets together.

Fig. 4. Calabasas solid-waste landfill in Los Angeles County, California, being lined with geomembrane.

In hot-wedge seaming, a wedge of hot steel is passed between the overlapped panel edges, followed directly by pressure rollers effecting the seam. The most advanced variation creates two welded tracks separated by an air gap. This allows testing by air pressure in the gap to detect potential leaks in the seam (**Fig. 5***a*).

In the hot-air welding method the sheets are melted by hot air blown between them, and then the molten sheets are pressed together. This welding system is not considered sufficiently consistent for final seaming operations; it is generally used only in conjunction with extrusion welding to keep the panels stationary before application of the extrusion weld.

Alternatively, solvent-adhesive seams can be effected for certain geomembranes. In solvent-adhesive welding operations the solvent-adhesive mixture is brushed onto panel edges, and various methods are used to press the panel edges together to form the seam. *See* ADHESIVE.

Testing. Geomembrane seams are tested very carefully to ensure that they are free of leaks and adequately joined. Nondestructive testing is used to check the entire seamed distance for leaks, and if any are found they are repaired.

Destructive testing is used to assess the quality of the bonding between panels. It requires that samples of the finished seam be removed and pulled apart in order to observe bonding. True welding is required, as opposed to surface attachments analogous to those in soldered metals. It is important to determine if welds are fully integrated because a mere surface attachment can be disrupted by absorbed chemicals. Samples are pulled apart by bending back the top panel from the overlapped section of the bottom panel in an attempt to peel the surface between

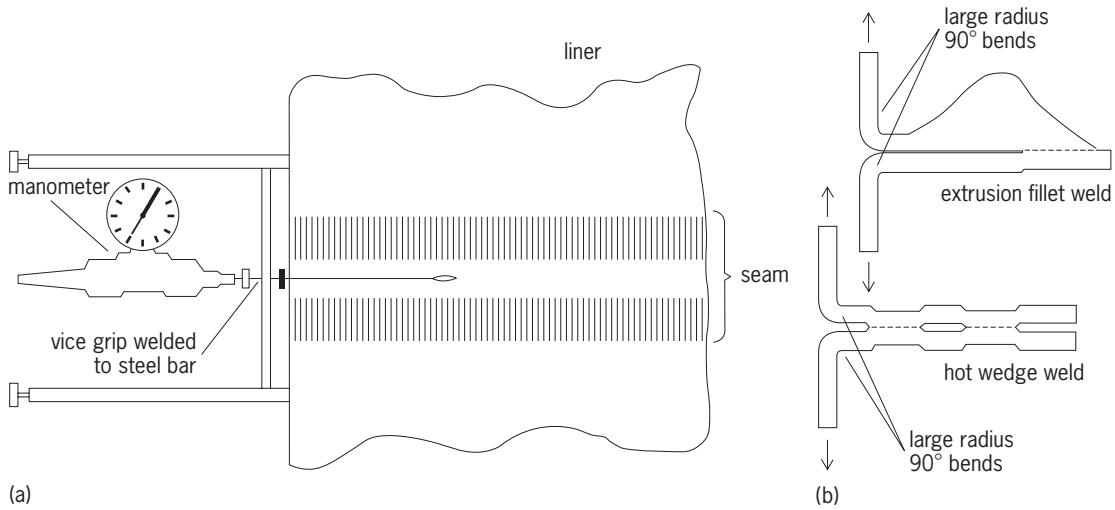

Fig. 5. Schematic diagrams of testing. (*a*) Nondestructive seam air-pressure test for leaks. (*b*) Destructive peel testing of both extrusion and dual-track hot-wedge seams to assess quality of welding, showing specimen configuration for peeling apart seams.

the seamed panels (Fig. 5*b*). If true welding has occurred, the resulting tear is through one of the panels as opposed to delamination (peeling). In contrast to nondestructive testing, destructive testing is limited to an interval basis. *See* ENGINEERING GEOLOGY; SOIL MECHANICS. Mark Cadwallader; Hal Pastner

Bibliography. J. E. Fluet, Jr. (ed.), *Geotextile Testing and the Design Engineer*, 1987; R. M. Koerner and R. F. Wilson-Fahmy (eds.), *Geosynthetic Liner Systems: Innovations, Concerns, and Designs*, 1996.

Geothermal power

Thermal or electrical power produced from the thermal energy contained in the Earth (geothermal energy). Use of geothermal energy is based thermodynamically on the temperature difference between a mass of subsurface rock and water and a mass of water or air at the Earth's surface. This temperature difference allows production of thermal energy that can be either used directly or converted to mechanical or electrical energy.

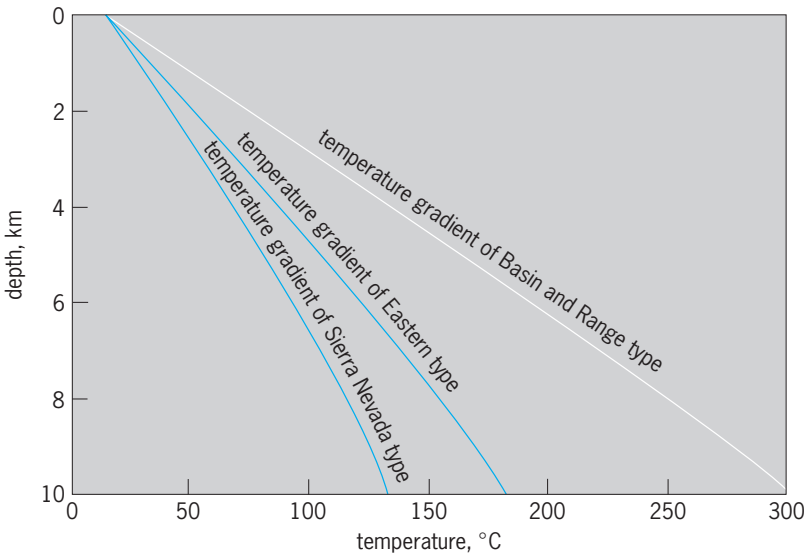

Fig. 1. Calculated average conductive temperature gradients in representative heat-flow provinces of the United States. 1 km = 0.6 mi. °F = (°C × 1.8) + 32. (*After D. E. White and D. L. Williams, eds., Assessment of Geothermal Resources of the United States–1975, USGS Circ. 726, 1975*)

Characteristics

Temperatures in the Earth in general increase with increasing depth, to 400–1800°F (200-1000°C) at the base of the Earth's crust and to perhaps 6300–8100°F (3500–4500°C) at the center of the Earth. Average conductive geothermal gradients to 6 mi (10 km; the depth of the deepest wells drilled to date) are shown in **Fig. 1** for representative heat-flow provinces of the United States. The heat that produces these gradients comes from two sources: flow of heat from the deep crust and mantle; and thermal energy generated in the upper crust by radioactive decay of isotopes of uranium, thorium, and potassium. The gradients of Fig. 1 represent regions of different conductive heat flow from the mantle or deep crust. Some granitic rocks in the upper crust, however, have abnormally high contents of uranium and thorium and thus produce anomalously great amounts of thermal energy and enhanced flow of heat toward the Earth's surface. Consequently thermal gradients at shallow levels above these granitic plutons can be somewhat greater than shown on Fig. 1. *See* EARTH, HEAT FLOW IN.

The thermal gradients of Fig. 1 are calculated under the assumption that heat moves toward the Earth's surface only by thermal conduction through solid rock. However, thermal energy is also transmitted toward the Earth's surface by movement of molten rock (magma) and by circulation of water through interconnected pores and fractures. These processes are superimposed on the regional conduction-dominated gradients of Fig. 1 and give rise to very high temperatures near the Earth's surface. Areas characterized by such high temperatures

are the primary targets for geothermal exploration and development.

Commercial exploration and development of geothermal energy to date have focused on natural geothermal reservoirs—volumes of rock at high temperatures (up to 662°F or 350°C) and with both high porosity (pore space, usually filled with water) and high permeability (ability to transmit fluid). The thermal energy is tapped by drilling wells into the reservoirs. The thermal energy in the rock is transferred by conduction to the fluid, which subsequently flows to the well and then to the Earth's surface.

Natural geothermal reservoirs, however, make up only a small fraction of the upper 6 mi (10 km) of the Earth's crust. The remainder is rock of relatively low permeability whose thermal energy cannot be produced without fracturing the rock artificially by means of explosives or hydrofracturing. Experiments involving artificial fracturing of hot rock have been performed, and extraction of energy by circulation of water through a network of these artificial fractures may someday prove economically feasible.

There are several types of natural geothermal reservoirs. All the reservoirs developed to date for electrical energy are termed hydrothermal convection systems and are characterized by circulation of meteoric (surface) water to depth. The driving force of the convection systems is gravity, effective because of the density difference between cold, downward-moving, recharge water and heated, upward-moving, thermal water. A hydrothermal convection system can be driven either by an underlying young igneous intrusion or by merely deep circulation of water along faults and fractures. Depending on the physical state of the pore fluid, there are two kinds of hydrothermal convection systems: liquid-dominated, in which all the pores and fractures are filled with liquid water that exists at temperatures well above boiling at atmospheric pressure, owing to the pressure of overlying water; and vapor-dominated, in which the larger pores and fractures are filled with steam. Liquid-dominated reservoirs produce either water or a mixture of water and steam, whereas vapor-dominated reservoirs produce only steam, in most cases superheated.

Natural geothermal reservoirs also occur as regional aquifers, such as the Dogger Limestone of the Paris Basin in France and the sandstones of the Pannonian series of central Hungary. In some rapidly subsiding young sedimentary basins such as the northern Gulf of Mexico Basin, porous reservoir sandstones are compartmentalized by growth faults into individual reservoirs that can have fluid pressures exceeding that of a column of water and approaching that of the overlying rock. The pore water is prevented from escaping by the impermeable shale that surrounds the compartmented sandstone. The energy in these geopressured reservoirs consists not only of thermal energy, but also of an equal amount of energy from methane dissolved in the waters plus a small amount of mechanical energy due to the high fluid pressures. *See* AQUIFER; GROUNDWATER HYDROLOGY.

Use of Geothermal Energy

Although geothermal energy is present everywhere beneath the Earth's surface, its use is possible only when certain conditions are met: (1) The energy must be accessible to drilling, usually at depths of less than 2 mi (3 km) but possibly at depths of 4 mi (6–7 km) in particularly favorable environments (such as in the northern Gulf of Mexico Basin of the United States). (2) Pending demonstration of the technology and economics for fracturing and producing energy from rock of low permeability, the reservoir porosity and permeability must be sufficiently high to allow production of large quantities of thermal water. (3) Since a major cost in geothermal development is drilling and since costs per meter increase with increasing depth, the shallower the concentration of geothermal energy the better. (4) Geothermal fluids can be transported economically by pipeline on the Earth's surface only a few tens of kilometers, and thus any generating or direct-use facility must be located at or near the geothermal anomaly.

Direct use. Equally important worldwide is the direct use of geothermal energy, often at reservoir temperatures less than 212°F (100°C). Geothermal energy is used directly in a number of ways: to heat buildings (individual houses, apartment complexes, and even whole communities); to cool buildings (using lithium bromide absorption units); to heat greenhouses and soil; and to provide hot or warm water for domestic use, for product processing (for example, the production of paper), for the culture of shellfish and fish, for swimming pools, and for therapeutic (healing) purposes. J. Patrick Muffler

Major localities where geothermal energy is directly used include Iceland (30% of net energy consumption, primarily as domestic heating), the Paris Basin of France (where 140–160°F or 60–70°C water is used in district heating systems for the communities of Melun, Creil, and Villeneuve la Garenne), and the Pannonian Basin of Hungary.

Electric power generation. The use of geothermal energy for electric power generation has become widespread because of several factors. Countries where geothermal resources are prevalent have desired to develop their own resources in contrast to importing fuel for power generation. In countries where many resource alternatives are available for power generation, including geothermal, geothermal has been a preferred resource because it cannot be transported for sale, and the use of geothermal energy enables fossil fuels to be used for higher and better purposes than power generation. Also, geothermal steam has become an attractive power generation alternative because of environmental benefits and because the unit sizes are small (normally less than 100 MW). Moreover, geothermal plants can be built much more rapidly than plants using fossil fuel and nuclear resources, which, for economic purposes, have to be very large in size. Electrical utility systems are also more reliable if their power sources are not concentrated in a small number of large units.

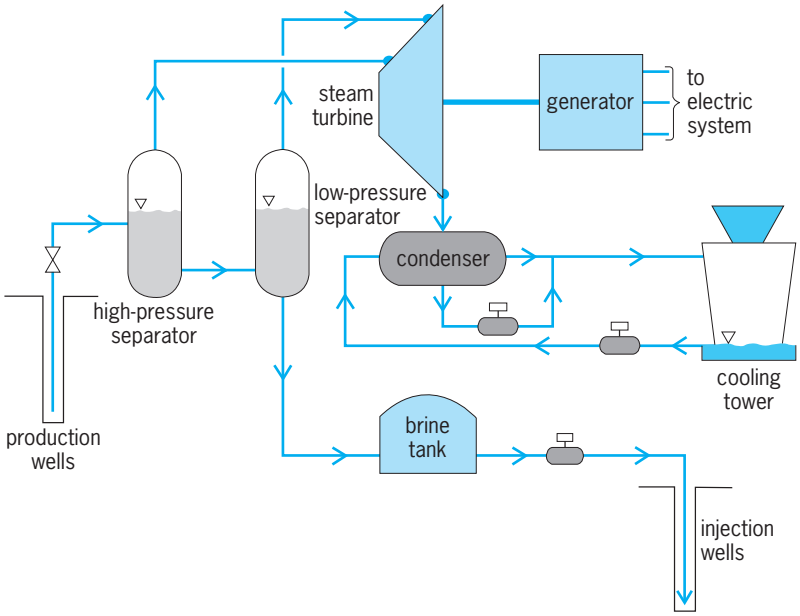

Fig. 2. Schematic diagram of the steam flash process.

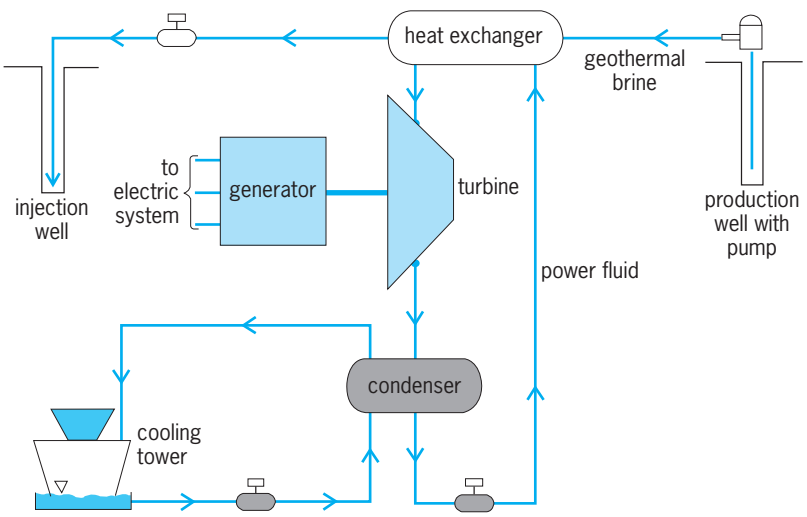

Fig. 3. Schematic diagram of the binary process.

In the United States a law was passed in 1978 that required the output from geothermal power generation projects (and others not based on fossil fuel resources, and cogeneration projects) to be purchased by electrical utilities at the cost that was avoided by the utility as a result of obtaining the power from a geothermal power plant. The legislation is called the Federal Public Utility Regulatory Policies Act (PURPA) and has created an incentive for the development of geothermal power projects.

The process used for generating power varies in accordance with the characteristics of the geothermal resource. The characteristics that affect the process are the temperature, the suspended and dissolved solids in the resource, and the level of noncondensable gases (primarily carbon dioxide) entrained in the geothermal brine, or steam. Almost all resources discovered to date are of the hydrothermal type (pressurized hot water) which can be produced from a well by two methods. If the temperature of a hydrothermal resource is below 400°F (204°C), a geothermal well can be produced with a pump, which maintains sufficient pressure on the geothermal brine to keep it as pressurized hot water. For hydrothermal resources over 400°F, the more suitable method of production is to flow the wells naturally, yielding a flashing mixture of brine and steam from the wells.

Steam flash process. The most common process is the steam flash process (**Fig. 2**), which incorporates steam separators to take the steam from a flashing geothermal well and passes the steam through a turbine that drives an electric generator. For the greatest efficiency in this process, a double-entry turbine is utilized which enables the most amount of steam available in the production from the geothermal well to be converted to electric power. If the resource has a high level of suspended and dissolved solids, it may be necessary to incorporate scaling control equipment in the steam flash vessel at the front of the plant and solids-settling equipment at the tail end of the plant. This will keep the process equipment from becoming plugged and allows a clean residual brine to be maintained for reinjection into the reservoir. If there are significant amounts of noncondensable gases, it may be necessary to install equipment to eject these gases out of the condenser to keep the back pressure on the system from rising and thereby cutting down on the efficiency of the process.

There are, at present, two resources in operation that have "dry steam," which is produced from the wells directly. These are very easy to convert to electric power and use the above described process without the necessity of the separation and brine injection equipment.

Binary process. A more efficient utilization of the resource can be obtained by using the binary process (**Fig. 3**) on resources with a temperature less than 360°F (180°C). This process is normally used when wells are pumped. The pressurized geothermal brine yields its heat energy to a second fluid in heat exchangers and is reinjected into the reservoir. The second fluid (commonly referred to as the power fluid) has a lower boiling temperature than the geothermal brine and therefore becomes a vapor on the exit of the heat exchangers. It is separately pumped as a liquid before going through the heat exchangers. The vaporized, high-pressure gas then passes through a turbine that drives an electric generator. The vapor exhaust from the turbine is then condensed in conventional condensers and is pumped back through the heat exchangers. There is a distinct environmental advantage to this process since both the geothermal brine and power fluid systems are closed from the atmosphere. Hydrocarbons, such as isobutane and propane, are common power fluids used in this process. *See* ELECTRIC POWER GENERATION.

Thomas C. Hinrichs

Production and Pollution Problems

The chief problems in producing geothermal power involve mineral deposition, changes in hydrological

conditions, and corrosion of equipment. Pollution problems arise in handling geothermal effluents, both water and steam.

Mineral deposition. In some water-dominated fields there may be mineral deposition from boiling geothermal fluid. Silica deposition in wells caused problems in the Salton Sea, California, field; more commonly, calcium carbonate scale formation in wells or in the country rock may limit field developments, for example, in Turkey and the Philippines. Fields with hot waters high in total carbonate are now regarded with suspicion for simple development. In the disposal of hot wastewaters at the surface, silica deposition in flumes and waterways can be troublesome.

Hydrological changes. Extensive production from wells changes the local hydrological conditions. Decreasing aquifer pressures may cause boiling water in the rocks (leading to changes in well fluid characteristics), encroachment of cool water from the outskirts of the field, or changes in water chemistry through lowered temperatures and gas concentrations. After an extensive withdrawal of hot water from rocks of low strength, localized ground subsidence may occur (up to several meters) and the original natural thermal activity may diminish in intensity. Some changes occur in all fields, and a good understanding of the geology and hydrology of a system is needed so that the well withdrawal rate can be matched to the well's long-term capacity to supply fluid.

Corrosion. Geothermal waters cause an accelerated corrosion of most metal alloys, but this is not a serious utilization problem except, very rarely, in areas where wells tap high-temperature acidic waters (for example, in active volcanic zones.) The usual deep geothermal water is of near-neutral pH. The principal metal corrosion effects to be avoided are sulfide and chloride stress corrosion of certain stainless and high-strength steels and the rapid corrosion of copper-based alloys. Hydrogen sulfide, or its oxidation products, also causes a more rapid degradation than normal of building materials, such as concrete, plastics, and paints. *See* CORROSION.

Pollution. A high noise level can arise from unsilenced discharging wells (up to 120 decibels adjusted), and well discharges may spray saline and silica-containing fluids on vegetation and buildings. Good engineering practice can reduce these effects to acceptable levels.

Because of the lower efficiency of geothermal power stations, they emit more water vapor per unit capacity than fossil-fuel stations. Steam from wellhead silencers and power station cooling towers may cause an increasing tendency for local fog and winter ice formation. Geothermal effluent waters liberated into waterways may cause a thermal pollution problem unless diluted by at least 100:1.

Geothermal power stations may have four major effluent streams. Large volumes of hot saline effluent water are produced in liquid-dominated fields. Impure water vapor rises from the station cooling towers, which also produce a condensate stream containing varying concentrations of ammonia, sulfide, carbonate, and boron. Waste gases flow from the gas extraction pump vent.

Pollutants in geothermal steam. Geothermal steam supplies differ widely in gas content (often 0.1–5%). The gas is predominantly carbon dioxide, hydrogen sulfide, methane, and ammonia. Venting of hydrogen sulfide gas may cause local objections if it is not adequately dispersed, and a major geothermal station near communities with a low tolerance to odor may require a sulfur recovery unit (such as the Stretford process unit). Sulfide dispersal effects on trees and plants appear to be small. The low radon concentrations in steam (3–200 nanocuries/kg or 0.1–7.4 kilobecquerels/kg), when dispersed, are unlikely to be of health significance. The mercury in geothermal stream (often 1–10 microgram/kg) is finally released into the atmosphere, but the concentrations created are unlikely to be hazardous. *See* AIR POLLUTION.

Geothermal waters. The compositions of geothermal waters vary widely. Those in recent volcanic areas are commonly dilute (<0.5%) saline solutions, but waters in sedimentary basins or active volcanic areas range upward to concentrated brines. In comparison with surface waters, most geothermal waters contain exceptional concentrations of boron, fluoride, ammonia, silica, hydrogen sulfide, and arsenic. In the common dilute geothermal waters, the concentrations of heavy metals such as iron, manganese, lead, zinc, cadmium, and thallium seldom exceed the levels permissible in drinking waters. However, the concentrated brines may contain appreciable levels of heavy metals (parts per million or greater).

Because of their composition, effluent geothermal waters or condensates may adversely affect potable or irrigation water supplies and aquatic life. Ammonia can increase weed growth in waterways and promote eutrophication, while the entry of boron to irrigation waters may affect sensitive plants such as citrus. Small quantities of metal sulfide precipitates from waters, containing arsenic, antimony, and mercury, can accumulate in stream sediments and cause fish to derive undesirably high (over 0.5 ppm) mercury concentrations. *See* WATER POLLUTION.

Reinjection. The problem of surface disposal may be avoided by reinjection of wastewaters or condensates back into the countryside through disposal wells. Steam condensate reinjection has few problems and is practiced in Italy and the United States. The much larger volumes of separated waste hot water (about 55 tons or 50 metric tons per megawatt-electric) from water-dominated fields present a more difficult reinjection situation. Silica and carbonate deposition may cause blockages in rock fissures if appropriate temperature, chemical, and hydrological regimes are not met at the disposal depth. In some cases, chemical processing of brines may be necessary before reinjection. Selective reinjection of water into the thermal system may help to retain aquifer pressures and to extract further heat from the rock. A successful water reinjection system has operated for several years at Ahuachapan, El Salvador.

A. J. Ellis

Bibliography. H. C. H. Armstead, *Geothermal Energy*, 2d ed., 1983; R. Bowen, *Geothermal Resources*, 2d ed., 1989; M. Economides and P. Ungemach (eds.), *Applied Geothermics*, 1987; L. Edwards et al. (eds.), *Handbook of Geothermal Energy*, 1982; J. Elder, *Geothermal Systems*, 1981; M. A. Grant et al., *Geothermal Reservoir Engineering*, 1983; K. Wohletz and G. Heiken, *Volcanology and Geothermal Energy*, 1992.

Geraniales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the superorder Rosidae of Eudicotyledon. The order consists of 6 families (Francoaceae, Geraniaceae, Greyiaceae, Ledocarpaceae, Melianthaceae, Vivianiaceae), 15 genera, and approximately 700 species. The Geraniaceae constitute the vast majority of the order and are temperate herbs or soft shrubs with deeply cleft or compound leaves (see **illus.**). The other families are mainly woody and are found in South America or Africa. Flowers typically have 5 sepals and petals, 10 stamens, and 5 fused carpels that separate from the central axis of the pistil when in fruit. Many of the species possess volatile compounds, as in *Pelargonium* (Geraniaceae) and *Melianthus* (Melianthaceae). *See* MAGNOLIOPHYTA; PLANT KINGDOM; ROSIDAE.

K. J. Sytsma

A common eastern United States species of geranium (*Geranium maculatum*), which is characteristic of the order Geraniales. (*A. W. Ambler, National Audubon Society*)

Gerbil

The common name for 88 species of small rodents comprising the subfamily Gerbillinae in the family Muridae (see **table**). They inhabit the desert regions, steppes, and sandy wastes of Asia, Africa, and southern Russia, often living many miles from water.

Genera and common names of gerbils

Genus	Numbers of species	Common name
Gerbillus	38	Northern pygmy gerbils
Microdillus	1	Somalian gerbil
Gerbillurus	4	Southern pygmy gerbils
Tatera	12	Large naked-soled gerbils
Taterillus	8	Small naked-soled gerbils
Desmodillus	1	Cape short-eared gerbil
Desmodilliscus	1	Brauer's gerbil
Pachyuromys	1	Fat-tailed gerbil
Ammodillus	1	Walo
Sekeetamys	1	Bushy-tailed jird
Meriones	16	Jirds
Brachiones	1	Przewalski's gerbil
Psammomys	2	Fat sand rats
Rhombomys	1	Great gerbil

Morphology. Most gerbils are about the size of a rat and are protectively colored to resemble the desert sands in which they live. They have a head and body length of 50–200 mm (2–8 in.), a tail length of 56–245 mm (2.2–9.8 in.), and a weight of 10–227 g (0.4–8 oz). They have large ears, large hindfeet, well-developed claws, and a long tail. Some have hair on the soles of the feet and bristles on the toes, presumably to prevent them from sinking into the loose soil. The smallest is the Brauer's gerbil with a head and body length of 41–74 mm (1.6–2.9 in.), a tail length of 33–49 mm (1.3–1.9 in.), and a weight of 6–14 g (0.2–0.5 oz). The fat-tailed gerbil (*Pachyuromys*) has a tail only half as long as its body. The tail is thickened and sausage-shaped, and serves as a fat-storage organ for this species, which does not store food in its burrows.

Many small mammals (gerbils, kangaroo rats, jerboas, golden hamsters, elephant shrews, etc.) living in hot arid conditions have relatively large auditory bullae (the bulbous bones enclosing their hearing apparatus). The arid air of the desert has poor sound-carrying qualities in comparison to cooler humid air. Thus, the external ears are large and very sensitive and are said to be suitable for receiving long-range low-frequency sounds. The inflated bullae increase the volume of the air-filled chambers surrounding the middle ear, which in turn reduces the resistance to the inward movement of the tympanic membrane. The malleus, part of which is greatly lengthened, is allowed to rotate more freely and has increased leverage. This transforms relatively weak vibrations of the greatly enlarged tympanic membrane into more powerful movements which are transmitted to the incus and the small footplate of the stapes, which rests against the oval window of the inner ear. *See* EAR (VERTEBRATE).

Gerbils display varying degrees of saltatory (jumping) locomotion (see **illustration**). They are often called antelope rats because of the way they move about. They hop rather than scamper or scurry in typical rat fashion. In this respect they remind one of jerboas, or jumping rodents; indeed, "gerbil" is just another form of the name "jerboa." *See* JERBOA.

Habitat and behavior. Gerbils are primarily nocturnal mammals. They are active year-round, although

The gerbil has a long tail that is used for balance when it hops.

their activity may be reduced during the winter in some areas. They feed on seeds, grasses, and roots, which for the most part contain a little moisture. Food may be stored for winter use. Some species may also include insects, birds, and bird eggs in their diet.

Fat sand rats (*Psammomys*) have the most efficient kidney known, but the necessity for efficiency is not a lack of water but an abundance of salt. Sand rats are desert rodents, but their habitat is restricted to the edges of saline and brackish waters. Sand rats feed on plants containing 80–90% water with about twice the salinity of seawater. The plentiful salt content makes it necessary for the kidney of the sand rat to be able to greatly concentrate the urine.

Gerbils live in burrows excavated in the sand. Burrows are variable in structure and may contain several levels with nest chambers, lengthy tunnels, and food storage chambers. Gerbils may be sociable and share community tunnels, or they may lead a solitary life. Males and females usually do not nest together.

The largest of the gerbils are the great gerbil and the large naked-soled gerbils. The great gerbil (*Rhombomys*) lives in central Asia from the Caspian Sea to southern Mongolia and north-central China, Iran, Afghanistan, and western Pakistan. Adults have a head and body length of 150–200 mm (5 $^3/_4$–7 $^3/_4$ in.) and a tail length of 130–160 mm (5–6 in.). The fur is thick and soft, and the tail is hairy, almost bushy. The soles of the feet are hairy, and the toes end in large claws. Females are polyestrous and produce two or three litters annually. Litter size may range from one to fourteen, and the young are born after a gestation of 23–32 days. Females may reach sexual maturity at 3–4 months. Longevity ranges 2–4 years. This species is considered a pest in Central Asia, where it damages crops, railway embankments, and irrigation channels. This species is also a reservoir for plague.

The large naked-sole gerbils (*Tatera*) have a head and body length of 90–200 mm (3 $^1/_2$–7 $^1/_2$ in.), a tail length of 115–245 mm (4$^1/_2$–9$^1/_2$ in.), and a weight of 30–227 g (1–8 oz). The body is heavy and ratlike. The dorsal coloration ranges from grayish to brownish; the underparts are white or whitish. The soles of the feet are naked. These gerbils usually walk on all four limbs, and when alarmed they escape by bounding as high as 1.5 m (4.9 ft). One species is said to be able to cover 3.5 m (11.5 ft) in one leap.

The Mongolian gerbil (*Meriones unguiculatus*) has become the most widely known species of gerbil since the 1950s because it has become a laboratory and pet animal and is now commercially traded almost worldwide. This species is gregarious and makes a good pet in captivity. *See* RODENTIA.

Donald W. Linzey

Bibliography. *Grzimek's Encyclopedia of Mammals*, vol. 3, McGraw-Hill, 1990; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Germ layers

The first cell layers that appear in the development of all animals, and from which the embryo body and extraembryonic membranes (when present) are constructed. The cells of the early embryo (the blastula) are organized into supracellular units, the germ layers. These are more or less distinct anatomically but do not necessarily have sharp boundaries of demarcation, in part because one layer gives rise to another. Germ layers are universal among animal embryos; they are an efficient method for establishing discontinuities of architectural importance and for setting aside what will become lineages of cells with different fates. *See* DEVELOPMENTAL BIOLOGY.

Types. Traditionally, three germ layers are recognizable: ectoderm, endoderm, and mesoderm as the outer, inner, and middle layers, respectively. Coelenterates and sponges possess only two of these layers—ectoderm and endoderm—and so are diploblastic. All other metazoans (multicellular animals) have the three germ layers and so are triploblastic. In recent years, a fourth germ layer—the neural crest—has been recognized as a distinctive layer restricted to vertebrates, which are therefore quadroblastic. The evolutionary innovations that followed the evolution of mesoderm (and that characterize all invertebrates other than sponges and coelenterates) are paralleled by evolutionary innovations that followed the evolution of the neural crest (which characterizes all vertebrates). *See* METAZOA; VERTEBRATA.

Endoderm and ectoderm are present in unfertilized eggs; they are set aside in the egg during oogenesis as a result of gene products deposited into the egg by the mother. Consequently, endoderm and ectoderm are regarded as primary germ layers—they arise first in individual development and were the first to appear phylogenetically. Mesoderm and neural crest arise later in development through inductive interactions: mesoderm from interactions between endoderm and ectoderm; neural crest following interaction between mesoderm and neurectoderm (neural ectoderm) and from interactions within the neurectoderm. Consequently, mesoderm and neural crest are regarded as secondary germ

layers—they arise secondarily in individual development and evolved subsequent to primary germ layers. The posterior region of many, if not all, vertebrate embryos does not form from germ layers but from a germinal zone of cells that does not separate into germ layers. As this occurs at neurulation (differentiation of nerve tissue and formation of the neural tube), we speak of primary neurulation (from germ layers) and secondary neurulation (from a germinal zone).

Three of the germ layers are named in accordance with their definitive positions in the spherical type of gastrula seen in sea urchins and amphibians, in which the ectoderm surrounds the embryo, the endoderm has invaginated as the precursor of the embryonic gut, and the mesoderm has delaminated to lie between ectoderm and endoderm. The neural crest is named from its topographical position in the crests of the neural folds at the onset of neurulation. In blastoderm types of blastulae, as in birds, a more or less distinct and complete lower endodermal layer (hypoblast) and an upper layer (epiblast) containing the prospective ectodermal and mesodermal layer-forming cells are present. In gastrulae of reptiles, birds, mammals, and a few invertebrate types, where the spherical form of the blastula has been modified into a two-layered blastodisc, the definitive positions of the layers remain essentially the same. The terms epiblast, mesoblast, and hypoblast are sometimes used as synonyms for ectoderm, mesoderm, and endoderm, respectively. However, these are geographical (topographical) terms; early in development the epiblast can contain future ectoderm and mesoderm or even all three future germ layers. Therefore, these terms should not be used as synonyms for germ layers. *See* BLASTULATION; GASTRULATION.

Also, the terms ectoderm and mesoderm should not be confused with the terms epithelia and mesenchyme; the former are germ layers, the latter subsequent types of cellular organization. Epithelia are sheets of polarized laterally connected cells (in simplest form, one cell layer thick) situated on a basement membrane secreted and deposited by the epithelial cells. Mesenchyme is a meshwork of unconnected cells that reside within extracellular matrix (ECM) which they synthesize and deposit. The ECM may be solid as in bone or fluid as in blood. Epithelia and mesenchyme can arise from all four germ layers, so the names for germ layers and for cellular organization should not be conflated. Sometimes, we use the terms mesenchyme for mesenchyme of mesodermal origin, and ectomesenchyme or endomesenchyme for mesenchyme of ectodermal or endodermal origin. *See* EMBRYONIC DIFFERENTIATION; EMBRYONIC INDUCTION.

Origin. The germ-layer structure of embryos has been known for almost 200 years. In 1817 Heinrich Christian Pander described the three-layered structure of the chick blastoderm, and within a decade (1828–1837) Karl Ernst von Baer recognized that the layer concept held true for many types of embryos, both vertebrate and invertebrate. During the latter half of the nineteenth century, the following concept of the origin of the germ layers gradually developed. The blastula was considered to be a single-layered hollow sphere which became converted into a two-layered gastrula by a process of invagination or delamination of cells from one wall of the blastula. The outer layer thus became the ectoderm; the inner layer, the endoderm. A third layer, the mesoderm, then arose from part of either the inner or outer layer, depending upon the group of animals; sometimes the terms endomesoderm and ectomesoderm are used to reflect the endodermal or ectodermal origins of mesoderm.

This concept of the origin of one germ layer from another during the process of gastrulation came to be accepted as the general and basic method of development of most metazoans. Today, we recognize that specific germ-layer-forming regions are present prior to the actual arrangement of the cells into distinct layers. We speak of these as future or presumptive germ layers, a state normally found in zygotes (for the two primary germ layers), in blastulae for mesoderm and in neurulae for neural crest. The fate of presumptive germ layers is not fixed; transplanting future ectoderm from a frog blastula into the future endoderm of another blastula results in the "ectodermal" cells forming endoderm. If the same transplantation is done using gastrula-stage ectoderm, we find that the cells no longer can change their fate—they form a piece of skin (an ectodermal derivative) within the gut (an endodermal derivative).

Fate of the layers. Once the germ layers have become established in their definitive anatomical positions, their further development consists, among other changes, in the segregation of subdivisions within them and in their morphological rearrangement. In vertebrates the ectodermal layer becomes subdivided into a skin-forming area and a neural-forming area. The latter area folds up to form the neural tube, from the crest of which will arise neural crest cells. The mesodermal layer in many invertebrates and vertebrates separates into two sublayers, one of which becomes applied to the overlying ectoderm, the other to the underlying endoderm. The space between the sublayers develops into the coelomic cavity. In vertebrates the portion of the mesoderm closest to the midline boxy axis (paraxial mesoderm) undergoes segregation into dorsal mesoderm to form a medial and dorsal, axial notochord, a rodlike structure around which the vertebral column is later formed. The notochord does not extend into the head; in the head, prechordal mesoderm and neural crest both form mesenchyme, with head muscles arising from mesoderm and the craniofacial skeleton arising from mesoderm and from neural crest. Lateral to the notochord, the mesoderm differentiates into blocks of mesoderm known as somites (one member of each pair on either side of the notochord and developing spinal cord), which later contribute to formation of skeletal muscles, vertebrae, and other structures. Initially epithelial, somites transform from an epithelial to a mesenchymal organization, providing a nice example of the

distinction between mesoderm and mesenchyme and of the origin of an epithelium from mesoderm. Still more laterally, the ventral mesoderm (the lateral plate) is further subdivided into an outer, or somatic, and inner, or splanchnic, layer, with the coelomic space between. An area of unsegmented mesoderm between the somites and the lateral plate, known as the intermediate mesoderm, contributes to the kidneys. This positioning and subdivision of primary and secondary germ layers constitute the basic or primitive body plan characteristic of all vertebrates. Further development of the germ layers consists of the formation of all the tissues and organs of the adult, except in some species of the germ cells, which arise from a separate germ plasm. The following list indicates the germ layer origin of some common organs and tissues of adult vertebrates.

Ectoderm
- Epidermis (hair, feathers, lens of eye)
- Central nervous system (brain, spinal cord, cranial and spinal nerves)
- Epithelia of sense organs, mouth, salivary glands
- Pituitary gland
- Enamel of teeth

Mesoderm
- Muscle
- Cartilage
- Bone
- Blood vessels
- Blood
- Kidney
- Gonads
- Liver (in part)
- Thyroid (in part)

Endoderm
- Epithelia of pharynx, thyroid, lungs
- Inner lining of digestive tract
- Bladder (in part)
- Urethra (in part)
- Liver (in part)

Neural crest
- Cartilage
- Bone
- Dentine of teeth
- Adrenal gland (part)
- Septa and valves of the heart
- Peripheral nervous system
- Enteric ganglia

Brian K. Hall; Nelson T. Spratt

Bibliography. S. F. Gilbert and S. R. Singer, *Developmental Biology*, 8th ed., 2006; B. K. Hall, *The Neural Crest in Development and Evolution*, 1999; J. M. W. Slack, *Essential Developmental Biology*, 2d ed., 2006; L. Wolpert et al., *Principles of Development*, 3d ed., 2006.

Germanium

A brittle, silvery-gray, metallic chemical element, Ge, atomic number 32, atomic weight 72.59, melting point 937.4°C (1719°F), and boiling point 2830°C

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

(5130°F), with properties between silicon and tin. Germanium is distributed widely in the Earth's crust in an abundance of 6.7 parts per million (ppm). Germanium is found as the sulfide or is associated with sulfide ores of other elements, particularly those of copper, zinc, lead, tin, and antimony. *See* PERIODIC TABLE.

Germanium has a metallic appearance but exhibits the physical and chemical properties of a metal only under special conditions since it is located in the periodic table where the transition from nonmetal to metal occurs. At room temperature there is little indication of plastic flow and consequently it behaves like a brittle material.

As it exists in compounds, germanium is either divalent or tetravalent. The divalent compounds (oxide, sulfide, and all four halides) are easily reduced or oxidized. The tetravalent compounds are more stable. Organogermanium compounds are many in number and, in this respect, germanium resembles silicon. Interest in organogermanium compounds has centered around their biological action. Germanium in its derivatives appears to have a lower mammalian toxicity than tin or lead compounds.

The properties of germanium are such that there are several important applications for this element, especially in the semiconductor industry. The first solid-state device, the transistor, was made of germanium. Single-crystal germanium is used as a substrate for vapor-phase growth of GaAs and GaAsP thin films in some light-emitting diodes. Germanium lenses and filters are used in instruments operating in the infrared region of the spectrum. Mercury-doped and copper-doped germanium are used as infrared detectors; synthetic garnets with magnetic properties may have applications for high-power microwave devices and magnetic bubble memories; and germanium additives increase usable ampere-hours in storage batteries.

Paul S. Gleim

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; R. Hull and J. C. Bean (eds.), *Germanium Silicon: Physics & Materials*, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; Z. Rappoport (ed.), *The Chemistry of Organic Germanium, Tin and Lead Compounds*, 2003.

GERT

A procedure for the formulation and evaluation of systems using a network approach. Problem solving with the GERT (graphical evaluation and review technique) procedure utilizes the following steps:

1. Convert a qualitative description of a system or problem to a generalized network similar to the critical path method—PERT type of network.
2. Collect the data necessary to describe the functions ascribed to the branches of a network.
3. Combine the branch functions (the network components) into an equivalent function or functions which describe the network.
4. Convert the equivalent function or functions into performance measures for studying the system or solving the problem for which the network was created. These might include either the average or variance of the time or cost to complete the network.
5. Make inferences based on the performance measures developed in step 4.

Both analytic and simulation approaches have been used to perform step 4 of the procedure. GERTE was developed to analytically evaluate network models of linear systems through an adaptation of signal flow-graph theory. For nonlinear systems, involving complex logic and queuing situations, *Q*-GERT was developed. In *Q*-GERT, a simulation of the network is performed in order to obtain statistical estimates of the performance measures of interest.

GERT networks have been designed, developed, and used to analyze production lines, quality control in manufacturing systems, capacity of air terminal cargo facilities, equipment allocation in construction planning, planning and control of marketing research, and system reliability. In recent time, GERT has become an obscure method which has seen little use. *See* DECISION THEORY; PERT; SIMULATION.

A. Alan B. Pritsker

Bibliography. U. Brandes and T. Erlebach (eds.), *Network Analysis: Methodological Foundations*, 2005; A. A. B. Pritsker, *GERT: Graphical Evaluation and Review Technique*, RM-4973-NASA, 1966; A. A. B. Pritsker, *Modeling and Analysis Using Q-GERT Networks*, 2d ed., 1980.

Gestation period

In mammals, the interval between fertilization and birth. It covers the total period of development of the offspring, which consists of a preimplantation phase (from fertilization to implantation in the mother's womb), an embryonic phase (from implantation to the formation of recognizable organs), and a fetal phase (from organ formation to birth).

In some mammals, the gestation period is extended by a special process known as delayed implantation or embryonic diapause: the fertilized egg develops in the normal way to form a hollow ball of cells (the blastocyst) and development is then arrested. Attachment of the blastocyst to the inner wall of the womb (implantation) is usually the next step in other mammals, but in cases of delayed implantation the blastocyst remains free in the womb and does not become attached until a later stage. In some cases, the period of delay before implantation is fixed. One example is the European roe deer, in which delayed implantation seems to ensure an optimal interval between mating in late summer and early fall, and birth in the spring. In other cases, the period of delayed implantation is variable. For instance, in many marsupials (such as the kangaroo) it is typical for a mother to have offspring in the pouch and one or more blastocysts free in the uterus. If the pouch offspring fail to survive, a free blastocyst can implant rapidly, thus permitting the mother to give birth again rapidly. For this reason, in these marsupials and in other mammals with an optional delay the gestation period can vary (see **table**).

Gestation periods of selected mammals

Mammal	Days	Mammal	Days
Baboon	173	Horse	337
Camel	402	Human	267
Cat	63	Lion	109
Chimpanzee	235	Macaque	165
Cheetah	93	Marmoset	148
Cow	284	Mole	35
Dog	63	Mouse	19
Elephant (Indian)	626	Opossum	13
Giraffe	456	Orangutan	250
Gorilla	259	Rat	21
Guinea pig	67	Sheep	148
Hamster	20	Wolf	63
Hedgehog	33	Zebra	360

There is widespread confusion over the duration of the gestation period in humans because of the way in which it is defined medically. The time of ovulation, and hence the time of fertilization, is difficult to determine in humans, so for purely practical reasons doctors measure the duration of pregnancy as the interval between the last menstrual period and birth, which is typically about 40 weeks or 280 days. For comparison with other mammals, however, the true gestation period between fertilization and birth in humans is about 267 days.

The length of the gestation period in mammals depends primarily on body size and the state of development of the offspring at birth. Large-bodied mothers have big offspring that take longer to develop, and development is also prolonged for offspring that are born at an advanced stage of development. There is, in fact, a major distinction between different groups of mammals according to the state of development of the offspring at birth. In some mammal groups (for example, marsupials, most insectivores, carnivores, and most rodents) the mother gives birth to a litter of poorly developed offspring (altricial offspring) that are typically kept in some kind of nest for some time after birth until they have grown enough to move around independently. Such altricial offspring are usually hairless at birth, and the eyes and ears are sealed off by membranes.

Marsupials are superaltricial, as their offspring are extremely poorly developed at birth and usually crawl into a pouch of some kind to continue their development before emerging into the outside world. In other mammal groups (for example, hoofed mammals, primates, whales, and dolphins), mothers typically give birth to a single, well-developed offspring, and a nest is not usually required. Such precocial offspring usually have a thick coat of fur at birth, and the eyes and ears are typically open. This distinction between mammal groups probably has a long evolutionary history. Fossil specimens from the Eocene Period (some 50 million years ago) show that relatives of the modern horse already had a single, well-developed offspring.

Because gestation periods generally increase in a regular fashion with the body size of the mother, this trend must also be taken into account when making comparisons between mammal species. For instance, the gestation period of the elephant (660 days) is about $2^1/_2$ times as long as the true human gestation period (267 days), but this difference can be attributed almost entirely to the enormous difference in body size. Once body size effects and special cases have been excluded, it emerges that marsupials have by far the shortest gestation periods. Among placental mammals, altricial mammals have markedly shorter periods than precocial mammals. Indeed, for any given body size, a precocial mammal has a gestation period three to four times longer than an altricial mammal. The human gestation period of 267 days is, for example, almost three times longer than the gestation period of 93 days for the cheetah, which has a similar body weight but gives birth to altricial offspring. Compared to all other mammals, human beings are found to have one of the longest gestation periods relative to body size.

One remarkable feature of mammalian gestation periods is that they show very little variability within a species. After excluding exceptional cases, departures from the average usually lie in a range of no more than ±4%. This is one of the smallest degrees of variability found in any biological dimension. *See* FERTILIZATION (ANIMAL); PREGNANCY; REPRODUCTIVE SYSTEM. R. D. Martin

Geyser

A natural spring or fountain which discharges a column of water or steam into the air at more or less regular intervals. It may be regarded as a special type of spring. Perhaps the best-known area of geysers is in Yellowstone Park, Wyoming, where there are more than 100 active geysers and more than 3000 noneruptive hot springs. Other outstanding geysers are found in New Zealand and Iceland. The most famous geyser is probably Old Faithful (see **illus.**) in Yellowstone Park, which erupts about once an hour. Then for about 5 min the water spouts to a height of 100–150 ft (30–45 m). Other geysers are less regular, but some intermittently discharge water and steam to heights of 250 ft (75 m) or more.

Old Faithful, Yellowstone Park, Wyoming. (*National Park Service, U.S. Department of the Interior*)

The eruptive action of geysers is believed to result from the existence of very hot rock, the relic of a body of magma, not far below the surface. The neck of the geyser is usually an irregularly shaped tube partly filled with water which has seeped in from the surrounding rock. Far down the pipe the water is at a temperature much above the boiling point at the surface, because of the pressure of the column of water above it. Its temperature is constantly increasing, because of the volcanic heat source below. Eventually the superheated water changes into steam, lifting the column of water out of the hole. The water may overflow gently at first but, as the column of water becomes lighter, a large quantity of hot water may flash into steam, suddenly blowing the rest of the column out of the hole in a violent eruption. *See* SPRING (HYDROLOGY). Albert N. Sayre; Ray K. Linsley

Ghost image (optics)

A feature or shape at the focal plane of a camera or other optical instrument that is not present in an actual scene, or an unfocused duplicate image that is overlaid upon a desired image. Ghost images, or ghosts, are caused by reflections from the surfaces of lenses or windows. Each glass surface divides incoming light into two parts: a refracted part that passes through the surface, and a reflected part that is turned back. If the reflected light is turned back again by reflection from another glass surface or a mirror, it may travel to the focal plane to form a ghost image. Ghost images may appear as an out-of-focus blur or smudge, a sharp circle or polygon with the

shape of the camera iris or other aperture, or a false image of an object within a scene. *See* REFLECTION OF ELECTROMAGNETIC RADIATION.

Antireflection coatings. The brightness of ghost images is greatly reduced by applying antireflection coatings to lenses and windows. Antireflection coatings are thin films of transparent material that reduce the reflectance of a glass surface. A common antireflection coating is a single layer of magnesium fluoride (MgF_2). Most glasses reflect a little more than 4% of the light from each surface. A single layer of magnesium fluoride with a thickness equal to one-quarter of the wavelength of light reduces the reflectance of glass to a little more than 1%. An even lower reflectance is obtained by applying two or more layers with properly chosen materials and thicknesses. Design of antireflection coatings is a specialized discipline in which an optimum mix of coating materials and layer thicknesses is developed with the help of computers and coating-design software. *See* INTERFERENCE FILTER.

Optical instrument design. Ghost images may be reduced or eliminated by properly designing the components of an optical instrument. The simplest way to eliminate ghosts is to use mirrors instead of lenses. When this is not possible, ghosts are controlled by dispersing them or moving them. A ghost is dispersed by adjusting the shape, location, or number of optical components so as to spread the ghost image over a broad area of the focal plane. As the size of the ghost increases, its brightness falls until it is no longer visible. A ghost image is moved by tilting the surfaces that cause the ghost. Tilting a surface changes the direction of the reflected light, which moves the ghost off the image plane. Tilting a single surface on a lens or window causes one edge of the part to be thicker than the other edge; this configuration is called a wedge. Ghosts caused by two reflections within a single lens element or window are not affected by tilting the entire component. Redirection of light from one tilted surface is undone by the tilt in the other surface.

Designing optical components that disperse and move ghosts is done with the help of computers and optical design software. Actions that remove ghost images often reduce image quality or increase the cost and complexity of an instrument. A computer calculates how changes to an instrument affect both the desired image of a scene and the brightness and shape of ghosts. A designer can then choose component shapes and materials that give the best performance.

Optical design software works by representing the propagation of light as lines, or rays, which lie along the path that light takes through an optical instrument; this technique is called ray tracing. At each lens surface, rays are split in two—one ray tracing the path of the transmitted light and the other ray tracing the path of reflected light. Numbers are assigned to each ray that represent the transmitted or reflected power. Rays that finally intersect the focal plane, along with their assigned power, are collected or binned together to calculate the size and brightness of each ghost. If the ghosts are too bright, changes that reduce them are made to the optical design. The effect of these changes on both image quality and ghost brightness is assessed, and adjustments are made in the design until an acceptable balance is achieved between the brightness of the ghost and other measures of the performance and cost of the instrument. Modern computers and software are capable of tracing many millions of rays, so accurate prediction of size, brightness, and cause of ghosts is a common part of instrument design. *See* GEOMETRICAL OPTICS.

Other sources. Reflections from lenses and windows are the most common causes of ghost images, but there are other sources. Optical components are often placed within a mechanical housing and then held in place by rings of metal. Bright sources, often outside the field of view of the instrument, illuminate the interior of the lens housing and retaining rings. Reflections from metal mounting surfaces may travel to the focal plane and form streaks or spots known as flares. Flares are also caused by reflection from the edges of lenses. Flares of this type are suppressed by applying black paint to the edge of an offending lens. Ideally, the index of refraction for the paint matches the index of refraction of the glass, so light that illuminates the edge of the lens passes easily into the paint and is absorbed. Diffuse scatter from housing walls often produces broad patches of illumination referred to as veiling glare. This kind of ghost is reduced by placing a thin metal annulus, called a vane, within the housing to block reflected and scattered light, or by machining fine grooves or threads in the walls of the housing to scatter light away from the focal plane.

Narcissus. Thermal infrared cameras work at wavelengths from 8 to 13 micrometers. At these long wavelengths, which cannot be seen by the eye, all warm objects produce infrared radiation. Infrared cameras that use cold detectors may exhibit a dark ghost around the center of their focal plane; this ghost is called narcissus. It is present if the detector sees a ghost image of itself reflected in one or more lenses of the camera. Because the detector is cold, it produces less infrared radiation than warm objects within the camera, so a ghost image of the detector is dark instead of bright. *See* INFRARED IMAGING DEVICES; INFRARED RADIATION. Gary L. Peterson

Bibliography. E. L. Dereniak and G. D. Boreman, *Infrared Detectors and Systems*, Wiley, New York, 1996; H. A. Macleod, *Thin-Film Optical Filters*, 3d ed., Institute of Physics Publishing, Bristol, U.K., and Philadelphia, 2001; W. J. Smith, *Modern Optical Engineering*, 3d ed., McGraw Hill, New York, 2000.

Giant nuclear resonances

Elementary modes of oscillation of the whole nucleus, closely related to the normal modes of oscillation of coupled mechanical systems. Giant nuclear resonances occur systematically in most, if not all, nuclei, with oscillation energies typically in the range

of 10–30 MeV. Among the best-known examples is the giant electric dipole (E1) resonance, in which all the protons and all the neutrons oscillate with opposite phase, producing a large time-varying electric dipole moment which acts as an effective antenna for radiating gamma rays. *See* GAMMA RAYS.

Giant resonances are usually classified in terms of three characteristic quantum numbers: L, S, and T, where L is the orbital angular momentum, S is the (intrinsic) spin, and T is the isospin carried by the resonance oscillation. The number L is also the multipole order, with possible values $L = 0$ (monopole), $L = 1$ (dipole), $L = 2$ (quadrupole), $L = 3$ (octupole), and so on. The spin quantum number S is either 0 or 1. The $S = 0$ resonances are often called electric, and the $S = 1$ ones magnetic (EL or ML, where L is the multipole order), stemming from the fact that these giant resonances have strong decay modes involving the emission of either electric (for EL resonances) or magnetic (for ML resonances) multipole photons of the same multipole order as the resonance. A giant resonance with $S = 0$ corresponds to a purely spatial oscillation of the nuclear mass (or charge density), while one with $S = 1$ corresponds to a spin oscillation. The isospin quantum number T, which is also either 0 or 1, determines the relative behavior of neutrons versus protons; in a $T = 0$ or isoscalar giant resonance, the neutrons and protons oscillate in phase, whereas in a $T = 1$ or isovector resonance the neutrons and protons oscillate with opposite phase. Examples of well-established giant resonances are the E0, E1, and M1 isovector modes, and the E0 and E2 isoscalar modes. Other types of giant resonances also exist, notably the giant Gamow-Teller resonance. *See* MULTIPOLE RADIATION; NUCLEAR MOMENTS.

These resonances are called giant because of their great strength, 50–100% of the theoretical limit, concentrated in a compact energy region. The oscillation energy is characteristic of the type of giant resonance and usually varies smoothly with mass. It is determined by the restoring force and the nuclear mass; the force is due to the nuclear attraction between nucleons, the most important part being the component of the same multipole order as the giant resonance.

Width. The width of a giant resonance generally contains contributions from three different sources: decay, damping, and fragmentation. If the resonance has sufficient energy, as most do, then the decay width stems primarily from the emission of nucleons and, in some cases, composite particles. Damping width is associated with mixing (damping) of the giant resonance into the sea of more complicated, underlying states, a process analogous to the frictional damping of a classical oscillator.

In the absence of decay and damping width, a giant resonance ideally would occur at a single sharp energy in a spherical nucleus. The degree to which the strength is fragmented among several different energies is accounted for by the fragmentation width.

Giant electric dipole (E1) resonance. This, the oldest and best known of the nuclear giant resonances, is the dominant feature in reactions initiated by gamma rays (**Fig. 1**). The absorption of a gamma ray induces the giant E1 oscillation, which breaks up, in this case, by emitting neutrons. This resonance is also the dominant feature in the reverse process, in which gamma rays are produced by proton and neutron bombardments of nuclei. The resonance, which is isovector ($L = 1$, $S = 0$, $T = 1$), occurs at an energy $E_x \cong 32A^{-1/3} + 20.6A^{-1/6}$ MeV (where A is the mass number of the nucleus) in medium and heavy nuclei (for example, 16.5 MeV at $A = 100$), and somewhat lower in light nuclei. Its strength is essentially equal to the theoretical limit. In deformed nuclei the resonance splits into several overlapping components (two if the shape is axially symmetric). In the most strongly deformed nuclei the resonance is partially resolved into two separate components.

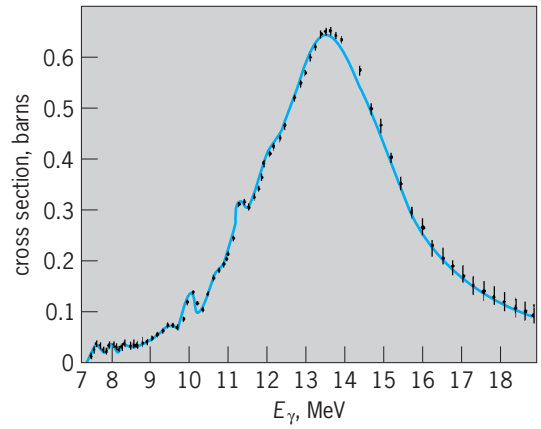

Fig. 1. Cross section of neutrons produced by gamma-ray bombardment as a function of the gamma-ray energy E_γ. The large peak is due to the giant electric dipole resonance. 1 barn = 10^{-28} m^2. (*After A. Veyssiere et al., Photoneutron cross sections of ^{208}Pb and ^{197}Au, Nucl. Phys., A159:561–576, 1970*)

Reaction selectivity. Because the different types of giant resonances often overlap in energy, particular resonances must be selectively excited in order to clearly delineate their properties. This is analogous to using the right driving force to excite particular normal modes of a coupled mechanical system. Inelastic alpha-particle scattering is very effective for studying isoscalar giant electric resonances. Both giant magnetic and electric resonances are studied by inelastic electron scattering, whereas charge changing reactions like the (p, n) reaction (neutron production by proton bombardment) are very useful for probing spin and isospin ($S = 1$; $T = 1$) modes.

Giant E0 and E2 resonances. The isoscalar giant E0 (electric monopole; $L = 0$, $S = 0$, $T = 0$) resonance lies at about $80A^{-1/3}$ MeV (17 MeV at $A = 100$), very close in energy to the giant E1 resonance, whereas the isoscalar giant E2 (electric quadrupole; $L = 2$, $S = 0$, $T = 0$) resonance lies somewhat lower, at about $65A^{-1/3}$ (14 MeV at $A = 100$). Both are strongly excited in forward-angle inelastic scattering of energetic alpha particles (**Fig. 2**). The angular dependence of the giant resonance excitation probability (cross section) is very different for the E0 and E2 resonances (Fig. 2), a feature which was essential in identifying the E0 mode.

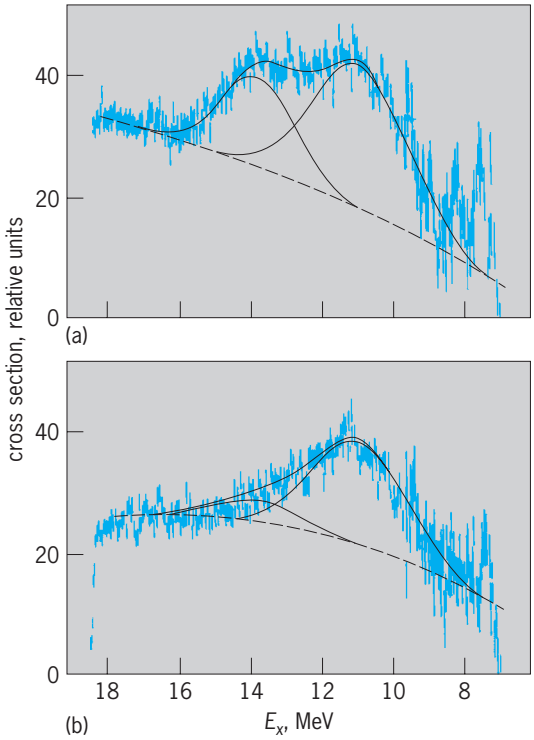

Fig. 2. Partial spectra of 120-MeV alpha particles scattered inelastically from lead-208 at angles of (*a*) 0° to 1.5° and (*b*) 1.5° to 3°. The broad peaks at excitation energies E_x near 11 and 14 MeV are the isoscalar giant quadrupole resonance and the isoscalar giant monopole resonance, respectively. (*After S. Brandenberg et al., Decay of the isoscalar giant monopole resonance in ^{208}Pb, Nucl. Phys., A466:29–69, 1987*)

The E0 resonances are spherically symmetric. The isoscalar E0 resonance is called the breathing mode, as the whole nucleus undergoes a purely radial oscillation, alternately expanding and contracting. The isoscalar E0 resonance energy is important in determining the nuclear compressibility.

The isovector E0 giant resonance ($L = 0$, $S = 0$, $T = 1$) is also a purely radial oscillation but one in which the protons oscillate against the neutrons. As a result, it lies at a higher energy than the isoscalar E0 resonance, at about $60A^{-1/6}$ MeV (28 MeV at $A = 100$). The isovector E2 giant resonance ($L = 2$, $S = 0$, $T = 1$) lies at around $130A^{-1/3}$ MeV (also 28 MeV at $A = 100$).

Giant Gamow-Teller resonance. In ordinary nuclear beta decay, a neutron inside a nucleus is transformed into a proton, and an electron and an antineutrino are produced. In one of the simplest types of beta decay, called Gamow-Teller decay, the transformed neutron is otherwise undisturbed, except that its spin may be reversed. As a result, the nucleus usually gains a small amount of energy. If beta decay involved a higher energy transfer to the nucleus, it would drive the giant Gamow-Teller resonance, which is a pure spin oscillation where the neutron spin and the proton spin oscillate out of phase ($L = 0$, $S = 1$, $T = 1$). A giant Gamow-Teller resonance is a strong feature in the (p, n) reaction in which neutrons emerge at 0° from nuclei struck by energetic protons (**Fig. 3**). This reaction substitutes a proton for a neutron in the nucleus via a spin-dependent interaction, in a manner analogous to beta decay but with a much larger energy transfer.

The properties of the giant Gamow-Teller resonance are important in certain problems in nuclear astrophysics. For example, the low-energy tail of the giant Gamow-Teller resonance in ^{37}Ar is important in determining the expected rate of the inverse beta decay reaction $\nu_e + {}^{37}\text{Cl} \rightarrow {}^{37}\text{Ar} + e^-$, which is used to detect neutrinos produced by nuclear reactions in the Sun. *See* RADIOACTIVITY; SOLAR NEUTRINOS.

Giant resonances in highly excited nuclei. Studies of the giant electric dipole resonance have been extended to highly excited hot nuclei. These studies provide unique information about the properties of such nuclei, in particular their shape. The shape sensitivity arises from the resonance splitting in a deformed nucleus, as described above. The size of the splitting gives the magnitude of the deformation, whereas the relative strength of the components determines the sense of the deformation: prolate (football-shaped) or oblate (doorknob-shaped).

The highly excited nuclei are formed by heavy-ion fusion, in which an energetic nucleus from an accelerator collides and fuses with a target nucleus. The large energy transferred to the excited nucleus is distributed between giant resonances and other, mostly random forms of energy analogous to heat, which is characterized by a nuclear temperature T.

These studies demonstrate the persistence of ground-state deformation in highly excited nuclei, with temperatures such that $kT \cong 1$ MeV (k = Boltzmann's constant). On the other hand, nuclei that are spherical at low temperature and spin undergo large thermal-shape fluctuations at temperatures $kT \geq 1$ MeV. Furthermore, hot nuclei that are spherical at low spin become deformed (most likely oblate-shaped) at higher spin. The same tendency is shown by a spherical liquid drop (or spherical water balloon) to become oblate when rotated.

Multiple giant resonances. The double giant E1 resonance is made up of two giant E1 resonances coexisting in the same nucleus. It is a so-called two-phonon giant resonance, or, in other words, a giant

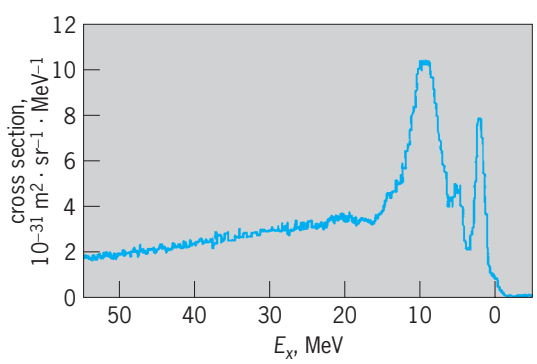

Fig. 3. Spectrum of neutrons emitted at 0° by bombarding zirconium-90 with 200-MeV protons. The broad peak at excitation energy E_x near 9 MeV is the giant Gamow-Teller resonance. (*After T. N. Taddeucci, Polarization transfer in (p, n) reactions, Can. Phys., 65:557–565, 1987*)

E1 resonance built upon a giant resonance. Double giant E1 resonances were first produced in (π^+, π^-) reactions, in which two neutrons inside the nucleus are transformed into protons and the resonance is excited. In subsequent studies, the double giant E1 resonance has been produced in distant collisions of heavy target nuclei with heavy projectiles moving at relativistic speeds. In such collisions, the repulsive electric force between the two positively charged nuclei has sufficiently high frequency Fourier components to induce a two-step process in which the normal (single) giant E1 resonance is first excited, and then, before the excited nucleus has a chance to decay, it is struck by a second so-called virtual photon, producing the double resonance. *See* FOURIER SERIES AND TRANSFORMS.

Giant resonances in supernovae explosions. Giant resonances play an important role in energetic nuclear reactions occurring in nature. Among the best examples are supernovae explosions. In a supernova, a massive star, having burned most of its nuclear fuel, undergoes gravitational collapse. During the collapse, nuclei in the core region produce neutrinos by a process called electron capture, which is a kind of inverse beta decay. The rate of these reactions, which cool the core and accelerate the collapse, depends on the properties of the giant Gamow-Teller resonance.

The stellar collapse continues until the stiffness of nuclear matter resists further compression, at which point a shock wave is created which rebounds out-ward and blows off the mantle of the star. The strength of the shock wave is directly related to the nuclear compressibility discussed above in the context of the giant isoscalar E0 resonance.

Left behind in the central region is a hot neutron star, which cools by emitting neutrinos. These higher-energy neutrinos travel outward and heat the nuclei in the mantle via inelastic scattering reactions which excite various giant resonances. This heating of the mantle, which can occur before the shock wave arrives, may contribute to the explosion. Certain elements found in nature may have been produced primarily as giant resonance decay products in these reactions. *See* GRAVITATIONAL COLLAPSE; NEUTRINO; NEUTRON STAR; NUCLEAR REACTION; NUCLEAR SPECTRA; NUCLEAR STRUCTURE; RESONANCE (QUANTUM MECHANICS); SUPERNOVA.

Kurt A. Snover

Bibliography. B. L. Berman and S. C. Fultz, Measurements of the giant dipole resonance with monoenergetic photons, *Rev. Mod. Phys.*, 47:713–761, 1975; Proceedings of the Gull Lake Conference of Giant Resonances, *Nucl. Phys.*, A569:1c–420c, 1994; Proceedings of the Workshop on Isovector Excitations in Nuclei, *Can. J. Phys.*, 65:535–698, 1987; K. A. Snover, Giant resonances in excited nuclei, *Annu. Rev. Nucl. Part. Sci.*, 36:545–603, 1986; J. Speth, *Electric and Magnetic Giant Resonances in Nuclei*, 1991; A. Van der Woude, Giant resonances, *Prog. Part. Nucl. Phys.*, 18:217–293, 1987; S. E. Woolsey et al., The neutrino process, *Astrophys. J.*, 356:272–301, 1990.

Giant star

An intermediate state in the evolution of a star in which it swells to enormous proportions before its death. During the longest and most stable phase of a star's life, the star, like the Sun, derives its energy from the thermonuclear fusion of hydrogen into helium deep in its dense, hot (10^7 K and up) core. It is then said to be on the main sequence. When the hydrogen fuel is gone and the central energy source is thereby exhausted, the core contracts and heats under the action of gravity, fresh hydrogen is ignited in a shell that surrounds the spent core, and the star becomes much more luminous, larger, and cooler at its surface. The lower surface temperature produces a redder color, hence the common term red giant. Stars like the Sun brighten by a factor of 1000 and grow in radius by a factor of 100 to about half the size of Earth's orbit (4.7×10^7 mi or 7.5×10^7 km).

There are actually three separate giant states. The first, described above, is terminated when the core temperature climbs so high (over 10^8 K) that the helium ignites and fuses into carbon. This event stabilizes the star; though the star is still a giant, it then contracts and dims as it quietly fuses its helium. When this helium is exhausted, the earlier behavior is repeated. The core contracts and is finally stabilized by electron degeneracy, becoming essentially a white dwarf. Helium then fuses to carbon in a shell around the core, and farther out hydrogen fuses to helium. The star then enters the asymptotic giant branch of the Hertzsprung-Russell diagram and swells to enormous proportions, perhaps two astronomical units (1.9×10^8 mi or 3×10^8 km), becoming even redder than before. It may pulsate and be seen as a long-period variable star, and loses much or most of its mass through a strong wind. *See* HERTZSPRUNG-RUSSELL DIAGRAM; STELLAR EVOLUTION.

James B. Kaler

Bibliography. I. Iben, Jr., Stellar evolution within and off the main sequence, *Annu. Rev. Astron. Astrophys.*, 5:571–626, 1967; I. Iben, Jr., and A. Renzinin, Asymptotic giant branch evolution and beyond, *Annu. Rev. Astron. Astrophys.*, 21:271–342, 1983; J. B. Kaler, *The Cambridge Encyclopedia of Stars*, 2006; J. B. Kaler, *Stars and Their Spectra: An Introduction to the Spectral Sequence*, 1997; J. M. Pasachoff and A. Filippenko, *The Cosmos: Astronomy in the New Millennium*, 3d ed., 2007.

Giardiasis

A disease caused by the protozoan parasite *Giardia lamblia*, characterized by chronic diarrhea that usually lasts 1 or more weeks. The diarrhea may be accompanied by one or more of the following: abdominal cramps, bloating, flatulence, fatigue, or weight loss. The stools are malodorous and have a pale greasy appearance. Infection without symptoms is also common. As with most other protozoa inhabiting the intestinal tract, the life cycle of *Giardia* involves two stages: trophozoite and cyst.

Trophozoites stay in the upper small-intestinal tract, where they actively feed and reproduce. When the trophozoites pass down the bowel, they change into the inactive cyst stage by rounding up and developing a thick exterior wall, which protects the parasite after it is passed in the feces. People become infected either directly by hand-to-mouth transfer of cysts from feces of an infected individual or indirectly by drinking feces-contaminated water. After the cyst is swallowed, the trophozoite is liberated through the action of digestive enzymes and stomach acids, and becomes established in the small intestine.

Epidemiology. Giardiasis occurs worldwide. Surveys conducted in the United States have demonstrated *Giardia* infection rates ranging from 1 to 20%, depending on the geographic location and age of persons studied. In community epidemics caused by contaminated drinking water, as many as 50 to 70% of the residents have become infected. Outbreaks also occur among backpackers and campers who drink untreated stream water. Both human and animal (beaver) fecal contamination of stream water has been implicated as the source of *Giardia* cysts in waterborne outbreaks. *Giardia* species in dogs and possibly other animals are also considered infectious for humans.

Epidemics resulting from person-to-person transmission occur in day-care centers for preschool-age children and institutions for the mentally retarded. Infants and toddlers in day-care centers are more commonly infected than older children who have been toilet-trained.

Why some people become ill when infected with *G. lamblia* and others do not has not been fully explained. Host immunity undoubtedly plays a role, but the exact immune mechanisms involved have not been identified. A number of other factors, such as the number of *Giardia* cysts swallowed (dose), varying virulence between *Giardia* strains, and origin of the parasite (human or animal), have been postulated, but not proved, as having an influence on the clinical course of infection. *See* EPIDEMIC; EPIDEMIOLOGY.

Diagnosis. The diagnosis of *Giardia* infection is most commonly made by identifying the causative agent, *G. lamblia*, in the feces. It is also possible to identify the parasite in digestive juices or biopsy material taken from the small intestine. In individuals with watery diarrhea, trophozoites are most commonly found in stools, but a few cysts may also be present. After the acute stage has passed, stools are more often semiformed or formed, and contain the more hardy cyst form of the parasite. Because *Giardia* cysts are passed in the feces on an intermittent basis, a minimum of three stool specimens (one every other day) should be obtained and examined to minimize the chance of missing an infection. The parasites may be stained in iodine or by more permanent staining methods for purposes of differentiating them from other bowel-inhabiting protozoa.

Treatment. Three drugs are available in the United States for the treatment of giardiasis: quinacrine, metronidazole, and furazolidone. Quinacrine is considered the drug of choice for adults and older children. Although quinacrine is effective in young children, the drug frequently causes vomiting in this age group. Metronidazole gives cure rates similar to quinacrine, and is generally well tolerated by both adults and children. Furazolidone is also an effective drug; it is the only anti-*Giardia* preparation that is supplied in pediatric suspension.

Prevention. Epidemic giardiasis most commonly results from ingestion of contaminated drinking water. The long-term solution to municipal waterborne outbreaks requires improvement in, and widespread use of, water filtration equipment in the water treatment process. Many cities in the United States rely solely on chlorination to disinfect drinking water; however, the amount of chlorine used does not kill *Giardia* cysts. Backpackers and campers should not drink water directly from streams or lakes, no matter how clean the water looks. If stream water must be used for drinking, it should be boiled for 1 min to kill *Giardia* as well as other infectious organisms that might be present. Chemical disinfectants such as laundry bleach or tincture of iodine may also be used to disinfect water of uncertain purity. These products work well against most bacterial and viral organisms, but are not considered as reliable as heat in killing *Giardia*. If water is cloudy, it should be strained through a clean cloth into a container to remove any sediment or floating matter. Then the water should be treated with chemicals.

Person-to-person transmission of *Giardia* can be prevented by practicing good personal hygiene and maintaining a sanitary environment. Hand washing should be stressed, especially after using the toilet or handling soiled diapers of infants. Quick and thorough cleanup of fecal accidents at home or in institutions also reduces the risk of spreading *Giardia* to others. *See* MEDICAL PARASITOLOGY.

Dennis D. Juranek

Bibliography. *Health Information for International Travel*, HEW Publ. (CDC) 1999; *Waterborne Transmission of Giardiasis*, U.S. Environmental Protection Agency, EPA-600/9-79-001, June 1979.

Gibberellin

Any of the members of a family of higher-plant hormones characterized by the *ent*-gibberellane skeleton shown below. Some of these compounds have

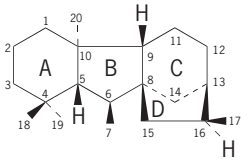

profound effects on many aspects of plant growth and development, which indicates an important regulatory role.

The discovery of gibberellins dates back to Japan in the 1920s, when plant pathologists were trying to understand, and ultimately control, the bakanae or

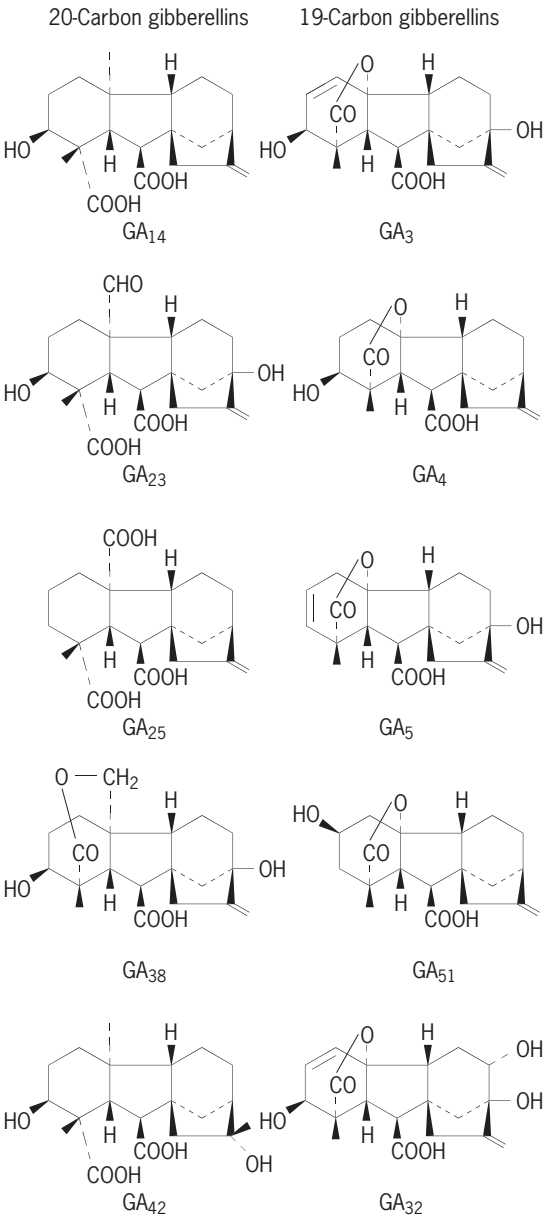

Fig. 1. Structural formulas of representative 20- and 19-carbon gibberellins.

"foolish seedling," disease of rice, which is caused by the fungus *Gibberella fujikuroi*. However, the chemical structure of the compound was not elucidated until 1954, when scientists reported the structure of gibberellic acid (GA_3) [**Fig. 1**]. *Gibberella fujikuroi* was subsequently shown to produce numerous other gibberellins, although GA_3 is by far the most abundant.

The discovery that gibberellic acid could restore normal growth in certain dwarf mutants of corn and peas led to the conclusion that gibberellic acid, or compounds structurally similar to it, is an endogenous growth regulator in higher plants. In 1958, the first higher-plant gibberellin, GA_1, was isolated from immature seeds of runner bean, and more than 70 gibberellins have subsequently been identified in all families of angiosperms and gymnosperms as well as several species of other fungi and ferns.

Nomenclature. All gibberellins are assigned a number and referred to as gibberellin A_1 . . . A_x, or simply as GA_1 . . . GA_x. In order for a compound to be given the GA designation, three criteria must be met: the structure and stereochemistry must be known; the compound must have the *ent*-gibberellane ring structure with a carboxyl group at carbon 7; and the compounds must occur naturally in plants or fungi. Compounds with unknown or incompletely known structures, such as those that have been detected in mosses, liverworts, algae, and bacteria, are referred to as gibberellin-like substances.

Chemistry. Gibberellins are tetracyclic diterpene acids with the *ent*-gibberellane skeleton. The two classes are based on the presence of the 19-carbon gibberellins and the 20-carbon gibberellins (Fig. 1). The 19-carbon gibberellins, which have lost carbon 20, contain a five-member lactone bridge that links carbons 4 and 10. The 19-carbon gibberellins, formed from 20-carbon gibberellins, are the biologically active forms. Gibberellins also vary according to the position and number of hydroxyl groups linked to the carbon atoms of the *ent*-gibberellane skeleton. Hydroxylation has a profound influence on biological activity. For example, compounds with a hydroxyl group on carbon 2 (such as GA_8, GA_{29}, and GA_{51}) are almost totally inactive and are generally thought to be the metabolic deactivation products of more active gibberellins. In contrast, gibberellins with a hydroxyl group on carbon 13 (GA_5 and GA_{20}) or carbon 3 (GA_4 and GA_7) possess much higher biological activity. The dihydroxylated gibberellins GA_1 and GA_3, which contain hydroxyl groups on carbons 3 and 13, are generally the most active.

Analytical methods. With the exceptions of *G. fujikuroi* and immature seeds of certain higher plant species, tissue levels of gibberellins are normally quite low, usually 1 to 100 nanograms per kilogram fresh weight. This complicates analysis of gibberellins, because extensive purification is often required before the actual analysis can be performed.

Analysis of gibberellin begins with extraction from the plant tissue using an organic solvent such as methanol or acetone. The organic extract is usually subjected to one or more procedures, such as solvent-solvent partitioning, that separate gibberellins from major plant constituents, including chlorophyll and other nonpolar pigments. The partially purified gibberellin extract is then subjected to chromatography, which affords additional purification and separates individual gibberellins from one another. In the past, this was achieved by paper or thin-layer chromatography; however, the use of high-performance liquid chromatography has largely replaced these two techniques. *See* CHROMATOGRAPHY; LIQUID CHROMATOGRAPHY.

The final step in gibberellin analysis is detection. Early work relied on gibberellin-specific bioassays that exploited the growth response of seedlings to applied gibberellins. Dwarf genotypes of rice, corn, and peas that have reduced levels of endogenous gibberellins have been most useful because of their sensitivity. In all these systems, the amount

of growth elicited is proportional to the amount of gibberellin applied. Another widely used bioassay is based on the promotion of starch hydrolysis in de-embryonated cereal seeds: the amount of starch hydrolysis is proportional to the concentrations of applied gibberellin. In practice, the response of the bioassay to several known concentrations of a gibberellin—which is usually GA_3, because of commercial availability—is determined and plotted graphically to construct a standard curve. The amount of gibberellin in a plant tissue extract can then be estimated from the standard curve by interpolation.

Bioassays provide a highly specific, simple, and inexpensive method for detection of gibberellins, but they have several inherent disadvantages. First, the degree of the response is a function of the logarithm of the concentration, so that a tenfold increase in the concentration of applied gibberellin only doubles the response, making it difficult to accurately measure small changes in GA levels. Second, bioassays give little structural information about the gibberellin, so that although the amount of GA-like substances present in an extract can be estimated, little can be deduced about exactly which gibberellin is being measured. Finally, and most important, not all gibberellins elicit the same response in a given bioassay. Some, such as GA_1 and GA_3, are highly potent in all bioassays, whereas gibberellins with hydroxyl groups on carbon 2 (for example, GA_8, GA_{29}, and GA_{51}) are almost totally inactive. *See* BIOASSAY.

A combination of gas chromatography and mass spectrometry overcomes many of the disadvantages of bioassays because such a technique provides definitive structural as well as quantitative information about the compounds being analyzed. A sample is injected into a gas chromatograph, which separates the components in the mixture. The gaseous component is fed into the mass spectrometer, where it is bombarded with a stream of electrons. As a result, the molecules break apart and form a series of charged fragments. Each compound has a unique pattern of fragment size and quantity, and so an unknown compound can be identified when its mass spectrum matches that of a reference compound. The mass spectrometer can also measure gibberellin levels, because the number of ions produced during fragmentation is proportional to the amount of the compound injected. *See* GAS CHROMATOGRAPHY; MASS SPECTROMETRY.

Biosynthesis. Gibberellins are a small subset of the terpene family of compounds. As with other terpenoids, gibberellin biosynthesis begins with mevalonic acid (**Fig. 2**). The *ent*-gibberellane skeleton is built by successive combinations of 5-carbon isoprene units. The 20-carbon diterpene geranylgeranyl pyrophosphate, with four isoprene units, is cyclized to form kaurene, which is a tetracyclic compound that resembles the *ent*-gibberellane skeleton except that ring B contains six carbons instead of five, as in gibberellins (Fig. 2). The conversion of kaurene to a gibberellin starts with the sequential oxidation of carbon 19, followed by hydroxylation at carbon 7 to form *ent*-7-α-hydroxykaurenoic acid (Fig. 2). Contraction of ring B by the extrusion of carbon 7 leads to the formation of GA_{12}-aldehyde. All gibberellins are subsequently derived from this compound by three major metabolic activities: (1) the oxidation of the carbon-7 aldehyde to a carboxyl group; (2) the formation of 19-carbon gibberellins by the sequential oxidation of carbon 20 to an aldehyde, followed by its removal and subsequent formation of a lactone bridge from carbon 4 to carbon 10; and (3) hydroxylation of one or more carbons contained in the *ent*-gibberellane skeleton. This process can occur in either 19- or 20-carbon gibberellins. In corn and peas, the formation of GA_{12}-aldehyde is followed by early hydroxylation at carbon 13 (GA_{53}); converson to GA_{20}, a 19-carbon gibberellin; and hydroxylation at carbon 3 (GA_1), and then at carbon 2 (GA_8). Alternatively, GA_{20} can be hydroxylated once at carbon 2 to form GA_{29} (**Fig. 3**). Hydroxylation at carbon 2 serves as a deactivation process. *See* TERPENE.

Glucose conjugates of gibberellins have been detected in many plant tissues, especially in seeds. Linkage to glucose takes place through either a carboxyl group (glucose ester) or an ether at a hydroxyl group (glucoside). Conjugation of gibberellins to glucose serves as another deactivation process in addition to hydroxylation at carbon 2. The conjugates may also serve as storage forms of gibberellin. *See* CONJUGATION AND HYPERCONJUGATION.

Gibberellin biosynthesis and metabolism takes place in three stages, based on characteristics of the component enzymes and properties of their substrates. First, soluble cytoplasmic enzymes convert mevalonic acid to kaurene. Most of the enzymes are common to the biosynthesis of other terpenes, and their substrates are highly soluble in water, because many contain a pyrophosphate group. The second stage begins with kaurene and ends with the formation of GA_{12}. Stage II oxidative enzymes are microsomal, and they require molecular oxygen (O_2) and the reduced form of nicotinamide adenine dinucleotide phosphate (NADPH); they are mixed-function oxidases involving the participation of cytochrome P-450. In addition, the substrates are highly nonpolar and not readily soluble in water. Stage III reactions include the conversion of 20-carbon gibberellins to 19-carbon gibberellins and all hydroxylations. These enzymes are soluble oxidases that contain iron and require NADPH and α-ketoglutarate as cofactors. *See* CYTOCHROME.

Of all gibberellins native to a given species, only one is believed responsible for biological activity, the others being either precursors to or deactivation products of that one gibberellin. It has been shown that GA_1 controls stem growth in both corn and peas and may also be important for biological action in most other species.

Determination of the gibberellin biosynthesis pathway has practical implications, such as the control of the stature of plants. Synthetic plant-growth regulators called growth retardants are used commercially to regulate plant growth and many of these compounds reduce endogenous gibberellin levels by

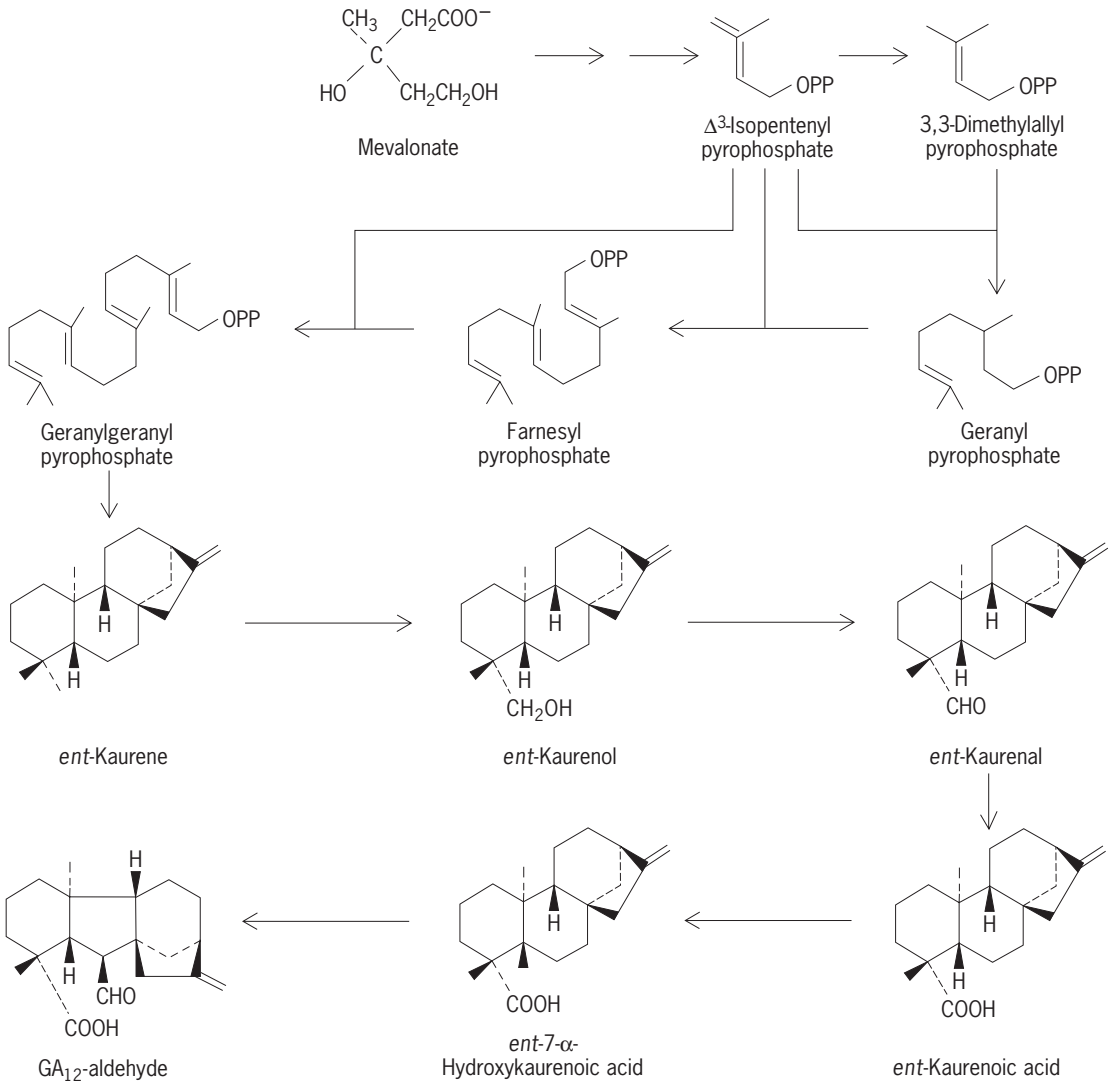

Fig. 2. Biosynthetic pathway of gibberellins from mevalonic acid in higher plants and the fungus *Gibberella fujikuroi*.

inhibiting gibberellin biosynthesis. Some growth retardants inhibit the cyclization of geranylgeranyl pyrophosphate to kaurene; others block the oxidation of kaurene to kaurenoic acid.

Control of growth and development. The involvement of gibberellins in specific aspects of plant growth and development can be inferred from several lines of experimentation.

Stem elongation. Probably the best-defined role for gibberellins in regulating the developmental processes in higher plants is stem growth. The absolute need for gibberellins in this process has been demonstrated by restoration of normal growth with gibberellin application to gibberellin-deficient dwarf mutants of corn, rice, peas, and other species. Rosette plants, which represent a special case of dwarfism, grow as dwarfs until they receive some inductive environmental stimulus such as a change in day length or temperature; gibberellins can often substitute for the stimulus. In the case of spinach, which requires long days to initiate stem growth, the inductive stimulus dramatically increases the level of certain gibberellins through higher activity of specific enzymes in the gibberellin biosynthetic pathway. The cellular basis for gibberellin-induced stem growth can be either an increase in the length of pith cells in the stem, as appears to be the case in lettuce hypocotyls, or primarily the production of a greater number of cells, as in many rosette plants following an inductive stimulus. *See* PHOTOPERIODISM.

Seed dormancy and germination. Freshly shed seeds from many species are often unable to germinate under ideal conditions; these seeds are said to be dormant. Depending on the species, overcoming dormancy requires a period of dry storage (afterripening) or an environmental stimulus such as light or low temperature. Applied gibberellins can often promote germination of dormant seeds, a capability suggesting that gibberellins are involved in the process of breaking dormancy.

Gibberellins are intimately involved in other aspects of seed germination as well. In the early stages of germination, the stored reserves that nourish the young seedling are mobilized until its photosynthetic

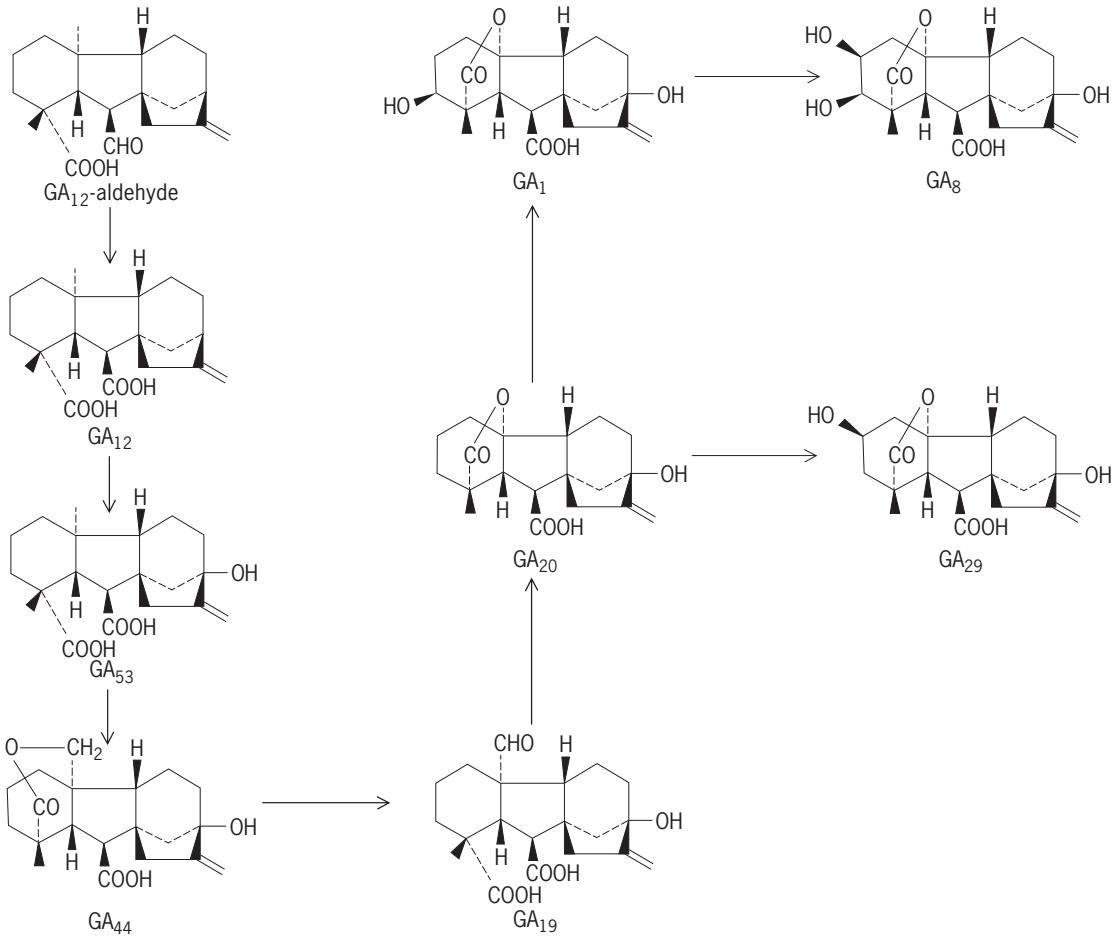

Fig. 3. Predominant pathway of gibberellin metabolism in corn and peas.

apparatus develops sufficiently. In the case of cereal grains, the breakdown of starch to glucose in the endosperm begins a few hours after the imbibition of water begins. Gibberellin in the seed embryo is believed to signal starch hydrolysis following action of the enzyme α-amylase, which is synthesized and released by aleurone cells that envelop the endosperm. Starch hydrolysis does not occur in de-embryonated seeds. Studies of molecular aspects of this process indicate that gibberellin causes increased transcription of the gene coding for the α-amylase enzyme, which results in higher levels of its messenger ribonucleic acid. *See* DORMANCY; RIBONUCLEIC ACID (RNA); SEED.

Flowering. Applied gibberellins promote or induce flowering in plants that require either cold or long days for flower induction. Gibberellin is probably not the flowering hormone or floral stimulus, because the floral stimulus appears to be identical or similar in all response types.

The sexuality of imperfect flowers (flowers with only male or female parts) is genetically determined, although environmental factors such as day length or temperature can be overriding factors. The application of gibberellins often modifies sex expression, usually causing an increase in the number of male flowers; in corn, however, feminization occurs. Gibberellin-deficient dwarf varieties of corn have male flowers on the ear (female inflorescence) indicating a natural role for gibberellins in sex expression. *See* FLOWER; PLANT GROWTH.

Commercial uses. Although gibberellins have limited use in agriculture compared with other agricultural chemicals such as herbicides, several important applications have been developed, including the production of seedless grapes. Application of gibberellin at bloom results in increased berry size and reduced berry rotting. Gibberellins are also used to increase barley malt yields for brewing and to reduce the time necessary for the malting process to reach completion. *See* MALT BEVERAGE.

Gibberellins have found significant applications in plant breeding. Many conifers do not flower until they are at least 10 years old, but the application of gibberellins can bring on cone production in juvenile plants. Thus, practical genetic improvement programs can be hastened by shortening the juvenile periods. In biennial vegetables such as carrots or brussels sprouts, seed can be produced in one season instead of the normal two following applications of gibberellin.

Other uses for gibberellin in agriculture include reduction of rind discoloration in citrus fruits, increased yield in sugarcane, stimulation of fruit set in fruit trees, and increased petiole growth in celery. *See* PLANT HORMONES.

James D. Metzger

Bibliography. A. Crozier (ed.), *The Biochemistry and Physiology of Gibberellins*, 1983; P. J. Davies (ed.), *Plant Hormones and Their Role in Plant Growth and Development*, 1987; J. E. Graebe, Gibberellin biosynthesis and control, *Annu. Rev. Plant Physiol.*, 38:419–465, 1987; D. S. Letham, P. B. Goodwin, and T. J. V. Higgins (eds.), *Phytohormones and Related Compounds: A Comprehensive Treatise*, 1978.

Ginger

An important spice or condiment; also the plant from which it is obtained, *Zingiber officinale*, of the ginger family (Zingiberaceae). The plant is a native of southeastern Asia. It is an erect perennial herb (see **illus.**) having thick, scaly, branched rhizomes which contain starch, gums, an oleoresin (gingerin) responsible for the pungent taste, and an essential oil which imparts the aroma. The rhizomes, dug up after the aerial parts have withered, are treated in different ways to produce green ginger or dried ginger. Ginger is used in medicine, in culinary preparations (soups, curries, puddings, pickles, gingerbread, and cookies), and for flavoring beverages such as ginger ale and ginger beer. The plant is grown in China, Japan, Sierra Leone, Jamaica, Queensland, and Indonesia. *See* SPICE AND FLAVORING; ZINGIBERALES.

Ginger (*Zingiber officinale*). (*USDA*)

Perry D. Strausbaugh

Ginkgoales

An order of gymnosperms in the class Ginkgoopsida (Pinophyta) with only one extant species, *Ginkgo biloba* (the maidenhair tree). Leaves identified as *Ginkgo* appeared first in the Upper Triassic, but *Ginkgo*-like leaves have been discovered as early as the Permian. During the Jurassic and Cretaceous periods, *Ginkgo* was a large taxon containing many species and having a circumpolar distribution in the Northern Hemisphere. The group diminished in size during the Early Cretaceous and into the Tertiary periods. It is probable that *Ginkgo*, considered sacred by the Chinese, would have become extinct had it not been cultivated in Chinese temple gardens, where the ginkgos became magnificent, old specimens several hundred feet tall.

Although *Ginkgo* is superficially different from other gymnosperms, details of seed morphology and reproduction clarify its nature as a gymnosperm. Young trees are excurrent; that is, they have a straight trunk from which lateral branches diverge (**Fig. 1**), but with age their excurrent habit often becomes obscured by the development of several major branches. The plant produces both long and short (spur) shoots. Long shoots, which result from extensive internodal growth, have alternately arranged leaves. Short shoots, characterized by almost no internodal growth, develop in the axils of leaves on long shoots and bear terminal clusters of leaves. The fan-shaped leaves, with dichotomous venation, are highly distinctive and may have a deep, apical notch (**Fig. 2***a*). Those borne on short shoots often lack the apical notch.

Ginkgo is dioecious (male cones and ovules are borne on different trees). Male cones, produced in the axils of leaves on short shoots, consist of an axis bearing helically arranged stamenlike structures, each of which terminates, usually, in a pair of microsporangia in which pollen is produced (Fig. 2*b*). Ovules commonly occur in terminal pairs on slender

Fig. 1. Two maidenhair trees (*Ginkgo biloba*).

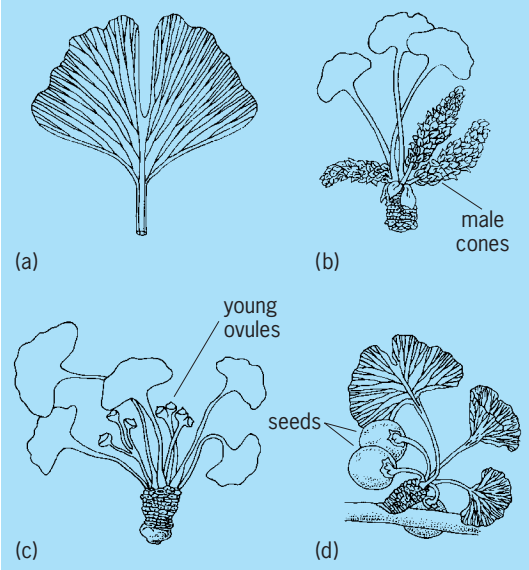

Fig. 2. Essential features of *Ginkgo biloba*. (*a*) Leaf showing deep apical notch and open, dichotomous venation of blade. (*b*) Short (spur) shoot with mature male cones and leaves lacking an apical notch. (*c*) Short shoot bearing young ovules with conspicuous basal collars. (*d*) Short shoot with mature seeds. (*After G. M. Smith et al., A Textbook of General Botany, 5th ed., Macmillan, 1953*)

stalks borne among the leaves of short shoots. Each ovule is enclosed basally by a collar (Fig. 2*c*). As the ovule matures, the outer integument becomes thick and fleshy, giving the seed the appearance of a fruit (Fig. 2*d*). Upon disintegration, the fleshy integument releases an unpleasant odor.

Ginkgo is used commonly as a shade tree in cities around the world because of its beauty and its resistance to disease and the effects of automotive pollution.

Several extinct genera are thought to be closely related to *Ginkgo*. Among these are *Trichopitys* and *Sphenobaiera*, which first appear in the early Permian, and *Baeira*, *Arctobaiera*, and *Eretmophyllum* from the Mesozoic. The reproductive structures of these ginkgophytes are unknown or not well understood. While their leaf form and stomatal characteristics resemble those of *Ginkgo*, it is not clear whether they should be included in Ginkgoales or assigned to one or more separate orders. *See* EMBRYOBIONTA; GINKGOOPSIDA; PINOPHYTA. Charles B. Beck

Bibliography. H. C. Bold, C. J. Alexopolous, and T. Delevoryas, *Morphology of Plants and Fungi*, 5th ed., 1990; R. F. Scagel, *Plants: An Evolutionary Survey*, 1984; W. N. Stewart and G. W. Rothinell, *Paleobotany and the Evolution of Plants*, 2d ed., 1993; T. N. Taylor and E. L. Taylor, *Biology and Evolution of Fossil Plants*, 1993.

Ginkgoopsida

A class of largely extinct gymnosperms (Pinophyta). Included orders are Calamopityales, Callistophytales, Arberiales, Peltaspermales, Ginkgoales, Leptostrobales, Caytoniales, Pentoxylales, and Ephedrales. The most ancient taxa, Calamopityales and Callistophytales, lived during the Carboniferous; the Arberiales, from late Carboniferous into the Triassic; the Peltaspermales, from Permian through Jurassic; the Leptostrobales and Caytoniales, from Triassic into the Cretaceous; and Ginkgoales, predominantly, from Triassic into the Cretaceous periods, with one species, *Ginkgo biloba*, persisting to the present. Ephedrales is the only largely extant group, but has a pollen fossil record beginning in the Upper Triassic. *See* CAYTONIALES; EPHEDRALES; GINKGOALES.

These taxa, divergent in many characteristics, are unified by the presence in all of platyspermic (bilaterally symmetrical) seeds lacking cupules. Presumed seeds of the most primitive order (Calamopityales) are of the *Lyrasperma* type, where the integument is fused with the nucellus over most of its length, and the nucellus, uncovered at its apex, terminates in a flasklike, pollen-catching organ called a salpinx. In all other orders, the seed is of the *Callospermarion* type, in which the integument, largely free from the nucellus, covers it except for a small opening at the apex, the micropyle. In some taxa, the seed may be secondarily platyspermic (derived from a radially symmetrical ancestor). In the most primitive, and some other, taxa of ginkgoopsids, the seeds and microsporangia are thought to have been borne on pinnately (featherlike) branched fertile structures. During the evolution of more advanced taxa, the individual microsporangia aggregated, fused to form synangia, and shifted onto leaves, causing additional changes to the microsporophylls and phyllosperms (seed-bearing leaves). In particular, the phyllosperms became greatly modified, often into peltate structures (as in advanced Peltaspermales) and stemlike structures (as in *Ginkgo*). *See* PINOPHYTA; PLANT KINGDOM. Charles B. Beck

Bibliography. C. B. Beck (ed.), *Origin and Evolution of Gymnosperms*, 1988; K. U. Kramer and P. S. Green (eds.), *The Families and Genera of Vascular Plants: Pteridophytes and Gymnosperms*, vol. 1, 1990.

Ginseng

The common name of the genus *Panax*, a group of perennial herbs of the aralia family (Araliaceae), native to the woodlands of the North Temperate Zone. *Panax schinseng* of Manchuria, extensively cultivated, was in such demand among the Chinese that the supply became insufficient. Then *P. quinquefolius* of eastern North America was discovered, and soon it was being exported to China in large quantities (see **illus.**). The price paid for the dried roots was so high that in a relatively short time the collectors nearly exterminated the plants. The Chinese used ginseng as a general panacea for many

***Panax quinquefolius*, ginseng, showing shoot and base.**

ills, but there is no evidence that the drug has therapeutic value. *See* APIALES.

Perry D. Strausbaugh; Earl L. Core

Giraffe

Member of the family Giraffidae represented by a single species, *Giraffa camelopardalis*. The only other living species in this family is the okapi (*Okapia johnstoni*); many fossil species are known. The giraffe occurs in the savanna regions of tropical Africa and the okapi ranges through the forested areas of the Congo. Both species are ruminants and belong to the mammalian order Artiodactyla (even-toed ungulates).

Giraffe (*Giraffa camelopardalis*). (*Photo by Arthur J. Emmrich; © 1999 California Academy of Sciences*)

The giraffe is the tallest of all mammals and may reach a height of 18 ft or 5.5 m (see **illustration**). The neck is long because of the extreme elongation of the neck vertebrae rather than an increase in their number. Giraffes are browsing animals, feeding mainly on acacia tree leaves. They are well adapted to this mode of feeding, having long lips and a prehensile tongue that can extend up to 20 in. (50 cm). These features allow them to pluck leaves from the trees and to avoid the thorns. There are two prominent horns on the forehead which are bony outgrowths covered by skin, and there is a short mane along the back of the neck. While these animals may weigh 1 ton (0.9 metric ton), they are agile and can travel at a good rate of speed. The senses of sight, hearing, and smell are well developed, and danger can be sensed at considerable distances. The giraffe lives in small herds with many females and usually one mature and several immature males. Old males are excluded and lead a life of isolation. Giraffes are social animals and may be seen with zebra, ostrich, and gnu. Gestation for the giraffe lasts about 15 months, and a single young is born. The young is about 6 ft (2 m) tall at the time of birth.

The okapi was not discovered until about 1900. It is not a common animal and lives in inaccessible areas of the eastern Congo. It is cryptically colored, having a hazel coat and striped hindquarters, and blends into its environment. The head shape, the lips, the tongue, and the horns of the male are the same as the giraffe's, but the neck is not elongate. The okapi closely resembles the extinct *Palaeotragus*, a ruminant that preceded the giraffe and occurred in Greece during the Miocene. The okapi is a nocturnal animal that lives singly or in pairs. It is a browser and quite dependent upon water. A single young is born after a gestation period of about 14 months. These animals breed readily in captivity, are rather shy, and become quite docile. *See* ARTIODACTYLA.

Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999.

Girvanella

A genus of fossil algae. *Girvanella* is characterized by flexuous, tubular filaments of uniform diameter, composed of thick, calcareous walls (see **illustration**). External diameters average between 10 and 30 micrometers, although specimens less than 10 μm and up to about 100 μm have been identified as *Girvanella.* Filaments may occur free (unattached), but usually occur in groups, twisted together to form nodules and encrusting masses on

Girvanella **in a thin section of Cambrian limestone. Tube diameter is about 20 m.**

various objects. The genus is intergrown with encrusting foraminifers in some Paleozoic limestones.

Girvanella is now generally placed in the blue-green algae (Cyanophyta), although in the past it has also been described variously as a foraminifer, sponge, and green algae. This genus is interpreted to be the calcified sheath of a variety of filamentous blue-green algae, similar to several living types.

Girvanella is a very common fossil, with a worldwide distribution. Occurring mainly in marine rocks, it has been reported from the Cambrian to Cretaceous. The apparent absence of *Girvanella* in rocks younger than Cretaceous age has not been satisfactorily explained. *See* ALGAE. John L. Wray

Bibliography. R. Riding, *Girvanella* and other algae as depth indicators, *Lethaia*, 8:173–179, 1975; J. L. Wray, *Calcareous Algae*, 1977.

Glacial geology and landforms

The scientific study of the processes and impacts of ice sheets, valley glaciers, and other ice masses on the Earth's surface, both on land and in ocean basins. The processes include understanding how ice masses move, erode, transport, and deposit sediment. The impacts on glaciated landscapes are enormous in terms of topographic change and floral and faunal modification. In those areas peripheral to glaciated areas, drainage patterns are altered, and climatic, vegetation, and soil conditions are severely changed. In addition, glacial geology involves studying the causes of glaciation, the chronology of glaciation in geologic time (the retreat and advance of ice masses at all scales), glacial stratigraphy, sea-level change, and how glaciations affect oceans, climate, flora, fauna, and human society globally. Closely allied to glacial geology are studies into the physics of ice masses (glaciology), global climatology (paleoclimatology), and paleoenvironmental reconstructions (paleoecology). *See* GLACIOLOGY; PALEOCLIMATOLOGY; PALEOECOLOGY.

Glacial sediments, both lithified and unlithified, are found on every continent, and the present oceanic basins of the Earth are covered in great thicknesses with glacial sediments. At the maximum of the past Pleistocene glaciation, sea level was at least 120 m (390 ft) below the present level. As a result, temporary land bridges appeared in the immediate postglacial period between Siberia and Alaska and between South Asia and the Indonesian Archipelago, and large parts of the North Sea in Europe were dry land. Such land bridges acted as corridors through which flora, fauna, and humans colonized North America, northwestern Europe, and parts of Australasia.

The impact of glacial geology on society is enormous, especially in the midlatitudes and poleward in both the Northern and Southern hemispheres. The soils, ground water, and construction are affected by glacial sediments and glaciated terrains. Cities such as Boston, Massachusetts, have been built on glacial landscapes such that urban planning and transportation systems must adapt to the topography. The siting of dams, the intricacies of ground-water pathways, and the location and discovery of minerals are tied to glaciations and problems in glaciated terrains. Of the total freshwater on Earth, 98% is held in ice sheets and other ice masses. (Freshwater is <2.5% of total water resources on Earth.) The impact of glaciation in the United States on its population is shown in **Fig. 1**, where considering land area in proportion to population, it can be seen that at least 60% of the population lives and works in areas once covered by the last glaciation (the Wisconsinan). In Canada, the comparative figure would be closer to 99% of the population, and in western Europe at least 50–60% of the population inhabits glaciated terrain.

Glacial History

Today, approximately 10% of the Earth's surface is covered by ice masses. The largest ice masses are the Greenland Ice Sheet and the West and East Antarctic Ice Sheets. The remaining ice masses are various mountain glaciers, small ice caps and fields, and ice shelves attached to the fringes of the Antarctic and the Canadian Arctic Archipelago (**Fig. 2***a*). Present ice-sheet thickness varies from around 2500 m (8200 ft) in central Greenland to over 4600 m (15,000 ft) in Antarctica. Valley glaciers and smaller ice masses are considerably thinner. During the maximum extent of the last glaciation in North America (Late Wisconsinan, Quaternary Period), the thickness of the Laurentide Ice Sheet in the region of Hudson Bay would have been similar to the Antarctic today. *See* ANTARCTICA; ARCTIC AND SUBARCTIC ISLANDS.

The extent of glaciation during the Quaternary appears to have covered at least 30% of the continents, but if the oceanic basins in the Arctic, north Pacific, north Atlantic, and Southern oceans are included, over 60% of the Earth's surface would have been icebound (Fig. 2*b*). It is likely that similar percentages of ice cover would have applied in earlier geological times.

Global glaciations can now be detected in all geological periods except the Jurassic. The Earth has alternated between greenhouse (warm, nonglacial periods, also know as interglacials) and icehouse

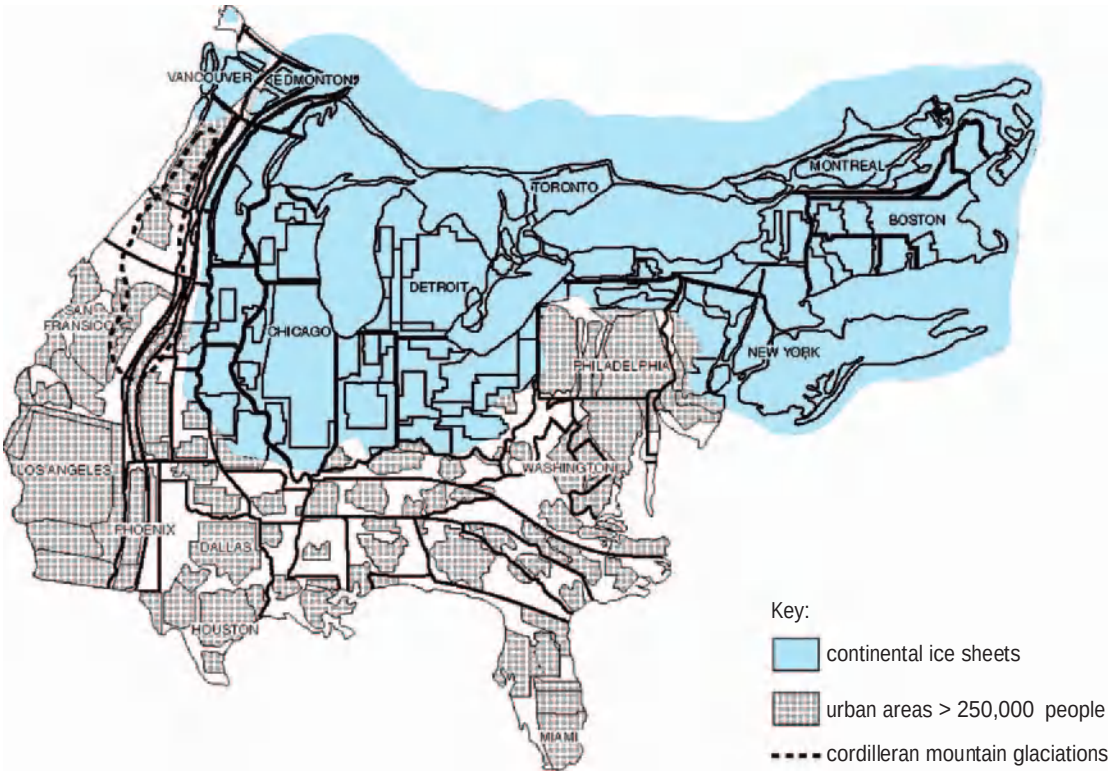

Fig. 1. Isodemographic map (1975/1976) overlain by the areal extent of the Laurentide and Cordilleran ice sheets at the last glacial maximum limits, 18,000 years ago. (*Modified from the Cart Department of Queens University, Kingston, Ontario, and R. C. Berg et al.*)

(cold, glacial periods) conditions in cycles ranging from 100,000 to 150,000 years. Recognition of the repetition of glaciation is a relatively new concept. In the nineteenth century, glacial lithified sediment (the Talchir Boulder Beds) was recognized in India, as well as the Permian glacial sediments of Shropshire, England, but a full recognition of repeated global glaciations became established only in the midtwentieth century. *See* GEOLOGIC TIME SCALE.

For a long time in Europe and North America, it was thought that a single glacial period [the Pleistocene (the ice age)] was to have occurred, giving it a unique position within the geological record. The Pleistocene is of immense importance since not only are the flora and fauna we see today largely a by-product of that glacial period, but also the emergence of human ancestors appears to stem from that period. In addition, the Pleistocene was a period of considerable upheaval in terms of plant speciation and the demise of several larger animals such as the wooly rhinoceros, giant elk, and mammoth.

When ice ages began to be recognized in Europe, a subdivision of four major glaciations was established. This fourfold division also was used in North America. Only with oceanic deep-drilling programs in the north Atlantic and Pacific oceans did it become apparent that, instead of four, 17 to 20 major glaciations had occurred during what has become a much longer and colder Pleistocene Period. Where in the past it was thought that glaciations worldwide were synchronous in their advances and retreats, it has become apparent that ice-mass margins varied considerably between continents and along different sections of the margin of the same ice sheet. For example, along the margins of the Laurentide Ice Sheet in North America, retreat and advance varied considerably between the east coast of New England and the Canadian Maritimes, as compared with the margins along the prairie regions of the Dakotas and the Canadian Prairie Provinces. This lack of synchroneity, certainly over periods less than was 2000 years, suggested a much more complex glacial geology than was previously considered. As dating methods advanced and the use of oxygen isotope ratios established warm and cold periods from data supplied from ocean drill cores, a more exact method of determining glacial and nonglacial periods has emerged using isotope stages. *See* PLEISTOCENE.

It appears we now live in an interglacial period that may be prolonged due to global warming and greenhouse gases. In the longer time span, it is possible that early accelerated global warming will lead to increased precipitation in the midlatitudes, which may have a negative feedback effect, thereby increasing the chance for a return to global glaciation. *See* GLACIAL HISTORY.

Causes of Glaciation

The causes of glaciation can be summarized by six possible mechanisms: cosmic, volcanic, tectonic, climatological, the "snowball earth," and astronomical. The repetitive nature of glaciation over geologic time must be of a paramount consideration. It has long

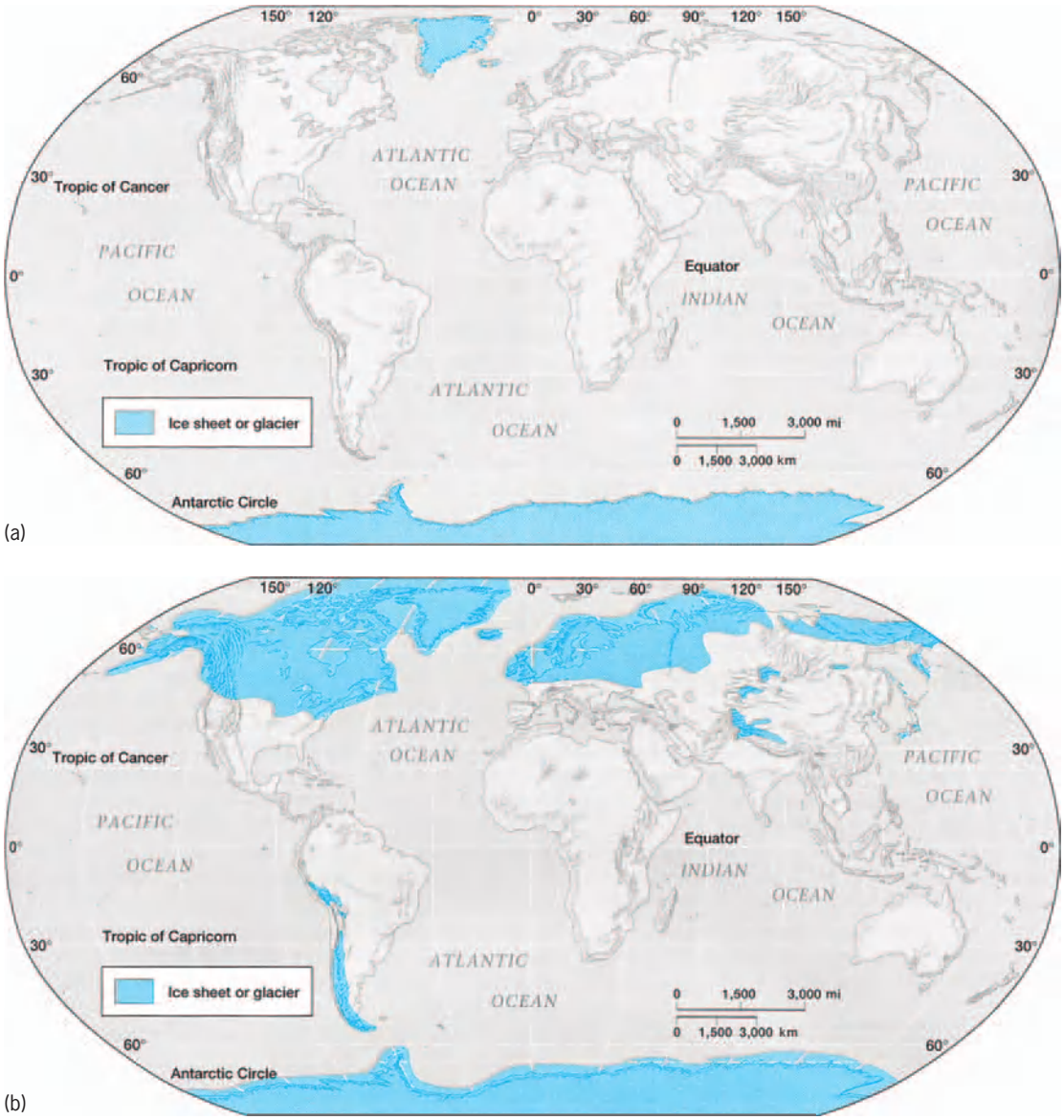

Fig. 2. Worldwide glaciation: (*a*) present distribution and (*b*) 18,000 years ago at the maximum of the Late Wisconsinan. (*Modified from T. L. McKnight and D. Hess, Physical Geography, 6th ed., 2000*)

been recognized that a reduction in solar energy striking the Earth's surface has the effect of lowering the mean annual surface temperatures by a few degrees (1–2°C), which is sufficient for snow banks and snowfields in high mountains to survive summer heat and gradually begin accumulating snow, leading to a sufficient thickness of snow that will transform to glacial ice. *See* INSOLATION; SNOWFIELD AND NÉVÉ.

Cosmic dust storms have been suggested as a mechanism of solar radiation reduction. They do occur and over the short term reduce overall solar radiant energy, but since they are of a nonrepetitive nature they cannot be a significant cause of glaciation.

Similar reductions in solar radiation occur with major volcanic eruptions when vast plumes of volcanic dust are shot high into the atmosphere and circle the Earth in the jet stream for considerable time. In 1991, the explosive volcanic eruptions from Mount Pinatubo in the Philippines caused a general global reduction in annual surface temperature of 0.3–1°C. However, this was a nonrepetitive event and not a major causative mechanism for global glaciation.

Over geologic time, continental landmasses have changed isostatically with respect to sea level, and landmasses such as the Tibetan Plateau have risen due to plate collision and orogenic forces. It has been suggested that as landmasses rise in elevation they may have been the sites of earlier glacial initiations. Again, even though much as this fact has been established, it is not repetitive and is therefore not a prime mechanism for glaciation. *See* PLATE TECTONICS.

That a climatological mechanism may trigger glaciation by changing patterns of seasons with much colder winters and cooler summers as was witnessed in medieval to midnineteenth-century Europe

cannot be denied, but such a sequence of events is unlikely to cause repeated global glaciation.

In recent years, the snowball theory, or runaway albedo effect, has been proposed in which the Earth became completely encased in a vast ice sheet, with even the oceans covered by thick sea ice. Evidence from pre-Quaternary sediments points to glaciations in the Neoproterozoic (1000–543 million years ago) that seem to have begun at low latitudes close to sea level. Such a glaciation would have required the ocean to cool to at least 30°N and S latitude, at which point a runaway albedo effect would have rapidly caused the Earth to lose energy by reflecting longwave radiation and becoming entirely ice covered. Alternatively, the Earth's tilt (at least 54°) then was greater than today (23.5°), and that may have caused great surface temperature anomalies. It is possible that the composition of the atmosphere was markedly different from today in terms of greenhouse gas effects which, when added to plate tectonic motion of continental landmasses, may have led to glaciation. Evidence from the Neoproterozoic points to the viability of the concept, but since that time glaciation has not occurred in the same globally cataclysmic manner. The snowball earth remains an interesting and controversial theory but does not seem to explain any of the global glaciations since the Neoproterozoic. *See* ALBEDO.

The most likely cause of glaciation on a global scale, but allied to global variations in climate and ocean currents, appears to be solar forcing. This astronomical theory has its beginnings in the midnineteenth century when James Croll and, later, Milan Milankovitch suggested that the movement of the Earth around the Sun (eccentricity) in association with the tilt of the Earth in relation to the Sun (tilt) and the Earth's wobble on its axis (precession), when in synchroneity, might be sufficient to reduce the Earth's surface temperature enough to initiate glaciation. This mechanism, coupled to other oceanic and climatological changes, appears to be the most likely mechanism to trigger repeated global glaciation. *See* EARTH ROTATION AND ORBITAL MOTION; PRECESSION OF EQUINOXES.

It is apparent that no single factor can be said to cause global glaciation. A combination of astronomical and climatological factors in relation to ocean currents and surface ocean temperatures, coupled with the plate tectonic movement of landmasses and the overall atmospheric quantities of greenhouse gases, appear to be necessary to trigger a global glaciation. Much remains to be understood about global climatic conditions and changes in relation to the many of factors discussed above. *See* CLIMATE HISTORY.

Landscape Development

In the Pleistocene, it seems that the vast ice sheets covering the Earth's surface began a relatively slow buildup to repeated maximum extensions into the midlatitudes (Fig. 2*b*). These ice sheets had an enormous impact in moving across terrain by eroding, transporting, and depositing vast quantities of sediments both on land and in ocean basins. The surface topography of the glaciated continents was totally altered with distinctive landscapes. Landforms were repeatedly overprinted until the terrains now seen in the northern United States and throughout Canada and northern Europe were finally exposed from beneath the ice some 10,000 years ago. Although the ice-sheet buildup was relatively slow, the melting of the ice sheets was remarkably fast, with sea level rising very quickly. Since the ice sheets isostatically had depressed the continental landmasses, in many instances sea level rose above present levels, drowning large coastal areas. Later, the continental landmasses adjusted to the loss of the ice-sheet load, and slowly raised shorelines and cliffs began to reappear above sea level. In Europe, these raised beaches and strandlines became the focus of the colonization pathways of early humans as they moved into northern Europe.

Glaciated landscapes are dominated by the effects of erosion, especially as seen on bedrock surfaces in the form of glacial scratches (striae) and chattermarks and the effects of high-pressure meltwater scour (**Fig. 3**). The evidence of glacial transport is in the boulder trains and isolated erratic boulders left strewn across many glaciated landscapes (**Fig. 4**). Evidence of glacial deposition occurs as thick glacial sediments, landforms, glacial lake sediments, and the immense thicknesses of glacial sediments within marine environments.

Glacial erosion. Glacial erosion is fundamental to the production of sediments for transport and deposition. Erosion processes occur in all glacial environments. Although limited, erosion in the form of abrasion and meltwater action occurs on the surfaces of ice masses, especially during the summer months when melted ice and snow move across glacier surfaces. In mountainous areas, considerable windblown debris and avalanched and mass-movement debris often end up on top of the ice mass (that is, the supraglacial environment), where debris may

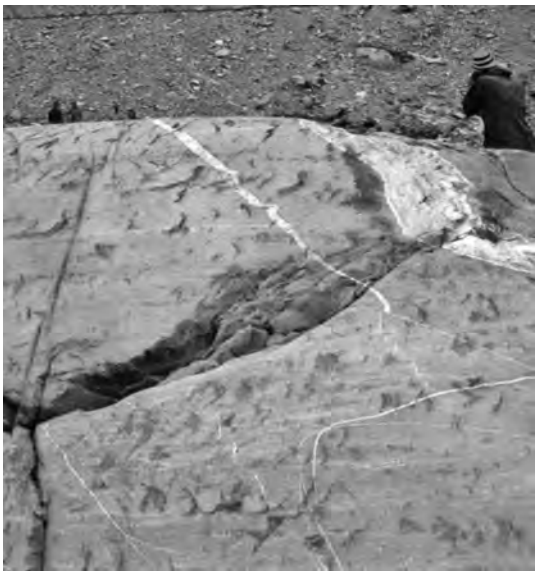

Fig. 3. Striations and chattermarks on bedrock in front of Omsbreen, Norway, ice moving from right to left. People are shown for scale.

Fig. 4. Erratic boulder, approximately 3 m (10 ft) high, sitting on the side of a field in southern Ontario, Canada.

suffer abrasion from percussion and from meltwater action. Even more limited is the amount of erosion that can occur within ice masses (that is, the englacial environment) in tunnels and galleries. Glacial debris is transported spasmodically, depending on meltwater activity within an ice mass, the concentration of debris within the ice mass, and the amount of debris that enters from the supraglacial environment via crevasses and meltwater moulins (cylindrical shafts), the connectivity of englacial tunnels, and meltwater discharge in the summer melt season. Significant volumes of debris are released from ice masses along lateral and frontal margins (that is, the proglacial environment) in valley glaciers or along the front margins of large ice sheets or into oceanic basins where ice margins are floating or have ice shelves attached. Debris exiting into the proglacial environment is subject to substantial erosion largely due to meltwater transport and mass-movement activity. The dominant location of glacial erosion occurs below the glacier (that is, the subglacial environment), where there is a high level of stress, considerable meltwater activity often under hydrostatic pressures, and vast sources of erodable material.

Processes. Glacial erosional processes can be subdivided as abrasion, plucking/quarrying, meltwater action, chemical action, and freeze–thaw processes. In general, all of these processes may operate on the same rock surfaces, such that examples of glacial erosion typically possess all their characteristic marks.

Abrasion is the wear or attrition of bedrock surfaces and rock fragment surfaces by the scouring processes of debris-laden ice and meltwater. The debris within the base of the ice or debris moving within high-pressure meltwater passes over the surface of bedrock or other rock surfaces, scratching and wearing down the surface. The evidence of abrasion is seen as minute scratches, or striae, on rock surfaces, or extremely smooth rock faces with tortuous geometries illustrative of rapid wear. These latter forms of erosion are called P-forms (**Fig. 5**). To be effective, such abrasion processes demand high basal-ice debris concentrations or high debris content in high-pressure meltwater streams, debris that is sharp, angular, and harder than the rock surfaces to be cut, sufficiently high basal-ice pressures, and an effective means of evacuation of the abraded debris. This last requirement is essential so that the abrasion process can continue, that is, not become clogged. The production of such immense volumes of abraded debris is apparent in the downstream "milky" nature of glacial streams, which exhibit a blue or greenish-blue color due to the high content of fine debris known as glacial milk or rock flour.

Plucking or quarrying is a set of processes that remove fractured, jointed, or disaggregated rock

Fig. 5. Large pothole (glacial P-form), Finland. (*Geological Survey of Finland*).

Fig. 6. Roche moutonée, approximately 25 m (80 ft) in length in front of the Nigardsbreen Glacier, Norway, with ice moving from right to left.

fragments from bedrock or the surface of rocks. Typically, rocks fracture under tensile or compressive stresses produced by the overlying moving ice mass or by percussive processes, whereby rock fragments crash against rock surfaces and produce flakes and shards of rock. Rock plucking can produce enormous boulders as well as minuscule rock flakes. Where large bedrock knobs have been overridden, roche moutonée may be produced as distinctive erosional landforms with smooth up-ice (stoss) sides and steep craggy down-ice (lee) sides (**Fig. 6**). Where percussive plucking occurs, tiny chattermarks and lunate fractures are developed on rock surfaces.

Meltwater, in association with high debris content and high-pressure discharge, creates other forms of abrasive wear on rock surfaces. A little understood process is chemical weathering beneath ice masses, where hydrostatically pressured meltwater, high stress levels, and fluctuating temperatures lead to carbonate, silicate, and iron solution and reprecipitation. Characteristically, such chemical processes can usually be observed as distinctive stains or "trails" on rock surfaces, often on the lee side of bedrock protrusions. In a glacial environment, freeze–thaw processes (seasonal and diurnal) are active, typically producing large volumes of frost-shattered rock. *See* WEATHERING PROCESSES.

Landforms. Glacial erosional landforms can be subdivided into areal and linear forms. The range of scale of these landforms can be immense, from centimeter-sized "rat-tails" and "flutes" on rock surfaces (**Fig. 7**) to roche moutonnée tens of meters in height (Fig. 6).

Fig. 7. Fluted and streamlined bedrock forms: (*a*) small-scale flutes, Wilton, Ontario; (*b*) P-forms in bedrock, Espanola, Ontario; (c) subglacial meltwater erosion marks; and (*d*) P-forms, Espanola, Ontario (scale card is 8.5 cm long).

Fig. 8. U-shaped valley, Muick, Scotland.

Vast areas of the Canadian and Fenno-Scandian shields best portray widespread areal glacial erosion. Linear forms occur in all glaciated regions, such as fiords, troughs (U-shaped valleys) [**Fig. 8**], finger lakes, tunnel valleys, and, at the much smaller scale, fluted bedrock knobs (P-forms), and roche moutonnée. All these landforms reflect the erosive power of glaciers, ice sheets, and associated meltwater action.

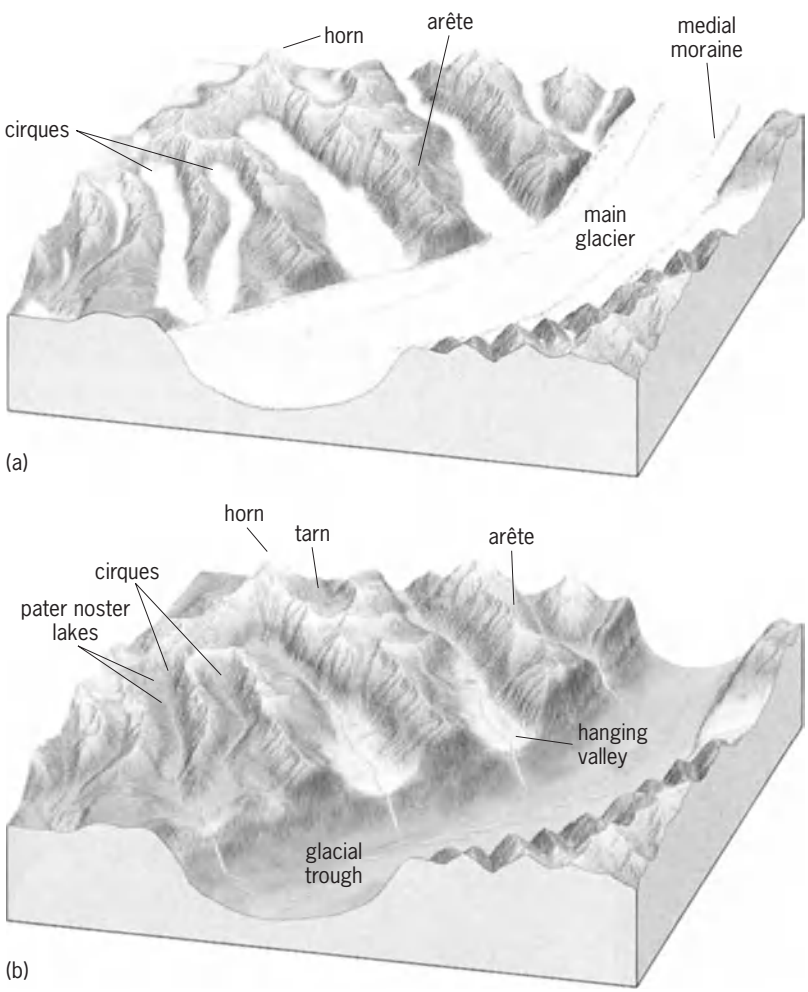

Fig. 9. Erosional landscape system for alpine glaciation. (*a*) Region during period of maximum glaciation. (*b*) Glaciated topography. (*Modified from E. J. Tarbuck and F. K. Lutgens, 2003*)

Landscape systems. Perhaps the best example of a glaciated landscape system is high mountain glaciation, where pyramidal peaks (horns), arêtes, cirques, tarn lakes, rock steps (riegels), and hanging valleys are found in close association (**Fig. 9**). No similar landscape system exists for ice-sheet erosive landscapes except for the areal scour, bedrock trough, and fluted-bedrock-knob landscapes of the Canadian Shield.

Glacial transport. In altering a landscape through erosion and subsequent deposition, glaciers and ice sheets transport vast quantities of sediment from fine-grained clays and silts to huge boulders. Sediment is transported on the surface of the ice as supraglacial debris, within the ice as englacial debris, beneath the ice as subglacial debris, and beyond the ice margins as proglacial sediment. Each environment affixes the transported sediment with a potentially distinctive signature that, although complex due to overprinting and retransportation, can be used to differentiate glacial sediment types.

Glacial deposition. The deposition of glacial sediment is largely a function of the means by which that sediment has been transported. Since ice masses scavenge sediments from their glaciated basins (in the case of ice sheets, this may be continental landmasses), the provenance of glacial sediment is typically immense, containing far-traveled sediment and boulders as well as local rock materials. The idea that glacial sediments are principally composed of far-traveled rocks is erroneous since most glacial sediments reflect a provenance only a few tens of kilometers up-ice from any point of deposition.

Glacial deposits can be subdivided into unsorted sediments with a wide range of particle sizes, and sorted sediments deposited by meltwater. Unsorted sediments are typically referred to as till, or glacial diamicton, and are exceedingly complex in their sedimentology and stratigraphy. They occur in all subenvironments of the glacier system and range from subglacial high-clay consolidated lodgement tills, to coarse supraglacial flow tills, to subglacial and submarginal areas of basal ice melting, leading to melt-out tills. In recent years, this classification with reference to subglacial tills is probably inaccurate. It would be more rigorous to call these tills tectomicts—products of the complex subglacial environment. *See* TILL.

Sorted, or fluvioglacial, sediments range from fine-grained clays deposited in glacial lakes (glaciolacustrine), to coarser sands and gravels deposited in front of ice masses as outwash fans (sandur), to rainout sediments deposited in marine settings (glaciomarine).

Processes. The processes of glacial deposition include direct smearing-on of subglacial debris as lodgement tills, mass movement of sediments from the frontal and lateral margins of glaciers as flow tills, and direct ablation of ice and melt-out tills formed in

front of ice masses in proglacial areas and in melt-out within subglacial cavities and caverns.

Landforms. Glacial landforms, or bedforms, range from those formed transverse to ice motion and parallel to ice motion, to unoriented nonlinear forms, ice-marginal forms, and fluvioglacial landforms. Although a complex number of glacial depositionary forms occur in most glaciated areas, the dominant landforms include Rogen moraines, drumlins and fluted moraines, hummocky moraines, end moraines, and eskers and kames.

Rogen moraines, or ribbed moraines, are a series of conspicuous ridges transverse to ice movement. These ridges typically rise 10–20 m (33–66 ft) in height, 50–100 m (164–328 ft) in width, and are spaced 100–300 m (328–984 ft) apart. They tend to occur in large numbers as fields in Quebec and Finland. They are composed of a range of subglacial sediments and are thought to be formed by the basal deformation of the underlying sediment, possibly at times in association with a floating ice margin (**Fig. 10**).

Fig. 10. Rogen moraine near Uthusslön, Sweden. (*Courtesy of Jan Lunqvist*)

Drumlins are one of the most known glacial landforms. A considerable literature exists as to their formation, and debate continues on their mechanism of origin. Drumlins are streamlined, roughly elliptical, or ovoid-shaped hills with a steep stoss side facing up-ice and a gentler lee side facing down-ice (**Fig. 11**). Drumlins range in height from a few meters to over 250 m (820 ft) and may be from 100 m (330 ft) to several kilometers in length. They tend to be found in vast swarms or fields of many thousands. In central New York State, over 70,000 occur. Likewise, in Finland, Poland, Scotland, and Canada, vast fields exist. Drumlins are composed of a wide range of dominantly subglacial sediments but may contain boulder cores, bedrock cores, sand dykes, and "rafted" nonglacial sediments. It seems likely that these landforms were developed below relatively fast-moving ice, beneath which a deforming layer of sediment developed inequalities and the preferentially stiffer units became nuclei around which sediment plasters and the characteristic shape evolved. Other hypotheses of formation are the streamlining by erosion of preexisting glacial sediments, changes in the dilatancy of glacial sediments leading to stiffer nuclei at the ice-bed interface, fluctuations in pore-water content and pressure within subglacial sediments again at the ice–bed interface, and the infilling of subglacial cavities by massive subglacial floods.

Fluted moraines are regarded as subglacial, streamlined bedforms akin to drumlins (**Fig. 12**). These landforms are typically linked in formation with drumlins and Rogen moraines. They tend to be much smaller in height and width as compared to drumlins but may stretch for tens of kilometers in length. They are composed of subglacial sediments. Fluted moraines have been found to develop in the lee of large boulders but may also be attenuated drumlin forms, the result of high basal-ice shear stress and high ice velocities. In central New York State, fluted moraines occur beside and among the large drumlin fields.

Hummocky moraine is a term used to denote an area of terrain in which a somewhat chaotic deposition pattern occurs, similar to a series of small sediment dumps. These moraines rarely rise above a few meters but may exist over considerable areas. Hummocky moraine often marks locations where massive downwasting of an ice mass may have occurred.

End moraines or terminal, retreat, or frontal moraines occur at the front margins of ice masses. Typically, these landforms, transverse to ice motion, may be a few meters to tens of meters in height and, if built up over several years, may be even higher and of considerable width (1–5 km or 0.6–3 mi). These moraines contain subglacial and supraglacial debris, and often have an arcuate shape closely mirroring the shape of an ice front. Since these moraines mark the edge of an ice mass at any given time, the longer the ice remains at that location, the higher and larger the moraines become. As an ice mass retreats with periodic stationary periods (stillstands), a series or sequence of moraines develop that mark the retreat stages (**Fig. 13**). Vast sequences of moraines can be observed in the midwestern states of Illinois, Ohio, and Michigan.

Eskers are products of subglacial meltwater streams in which the meltwater channel has become blocked by fluvioglacial sediments. Eskers are long ridges of sand and gravel that run across the landscape (**Fig. 14**). They range in height from a few meters to >50 m (160 ft) and may run for tens of kilometers. In Canada, some eskers cross the Canadian Shield for over 100 kilometers (60 mi). These landforms often have branching ridges and a dendritic morphology. Eskers are dominantly composed of fluvioglacial sand and gravel, with distinctive faulted strata along the edges where ice-tunnel walls melted

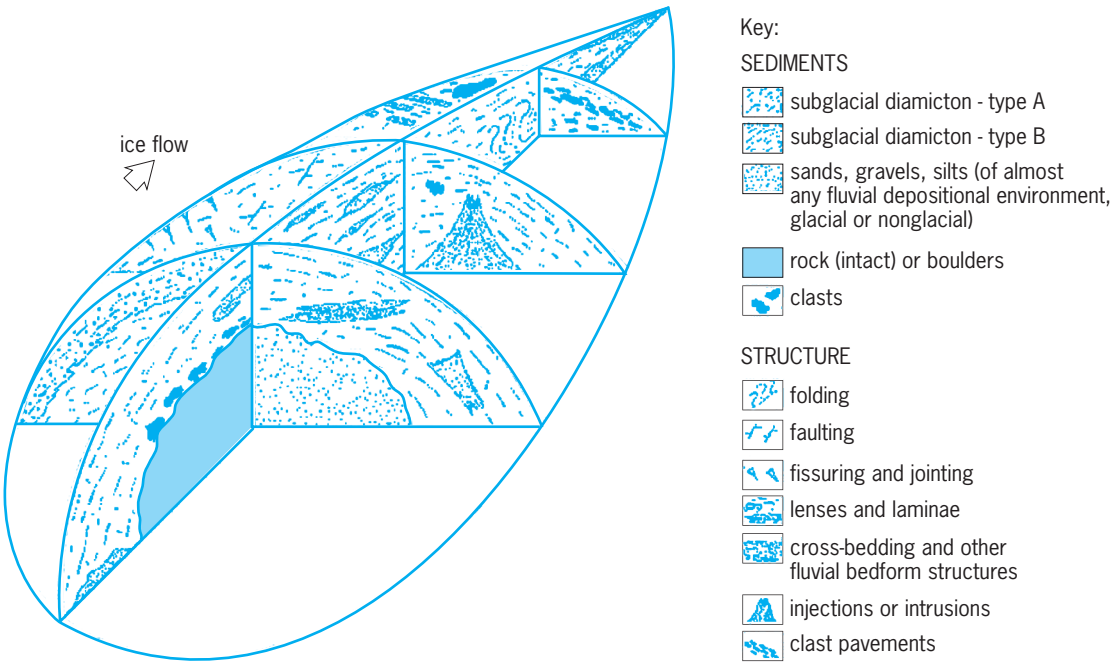

(a)

(b)

Fig. 11. Drumlin S. (*a*) General model, showing the variability of the internal composition. (*b*) Green Bay Lobe, Wisconsin. Ice direction from the upper right to lower left. (*Courtesy of D. Setz*)

and collapsed. In many instances, eskers are observed "running" uphill or obliquely crossing over drumlin—evidence of their formation within high-pressure subglacial meltwater tunnels.

In many glaciated areas during ice retreat, large, roughly circular dumps of fluvioglacial sand and gravel occur. These forms, called kames, appear to form when infilled crevasses or buried ice melts, leaving behind the sediment in a complex but chaotic series of mounds. The term kame or kame delta has been applied to fluvioglacial sediments that formed temporary deltas on entering now-gone glacial lakes. In Finland, a long line of such kame deltas formed into a moraine-like series of linear ridges (Salpausselkä) transverse to the ice-front retreat.

Fig. 12. Fluted moraine, Storbreen Glacier, Norway. Note boulder at he head of the flute; ice direction bottom right to middle left.

Fig. 14. Esker ridge, Bylot Island, Canada. (*Courtesy of C. Zdanowicz*)

Distinctive suites of glacial depositional landforms can be found in valley glacier and ice-sheet settings (**Fig. 15**). Typically, these sequences of landforms are often overprinted due to subsequent glaciations, but in the midwestern United States such suites of depositional landforms from the last glaciation (Late Wisconsinan) can be clearly discerned (**Fig. 16**). Distinctive landscape systems attributable to the marginal areas of ice sheets also carry a unique supraglacial suite of landforms, as can be observed in parts of the midwestern United States.

Periglacial effects. Areas beyond the ice limits are strongly influenced by glaciation due to deteriorating climatic conditions, the deposition of windblown dust (loess), the divergence of river systems where headwaters or partial drainage basins may be intersected by advancing ice, the impact of outburst floods from ice fronts (jökulhlaups), and major faunal and floral changes due to encroaching ice. In terms of the impact of ice masses on human life and

(a)

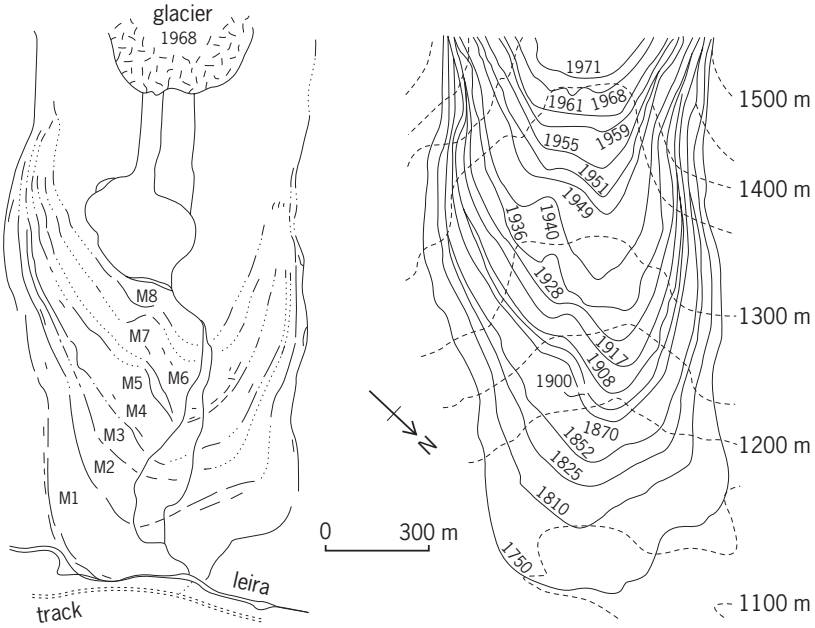

(b)

Fig. 13. End moraine. (*a*) Findelen Glacier, Switzerland (*Courtesy of J. Matthews*). (*b*) Sequence of retread of end moraines from Storbreen Glacier, Norway.

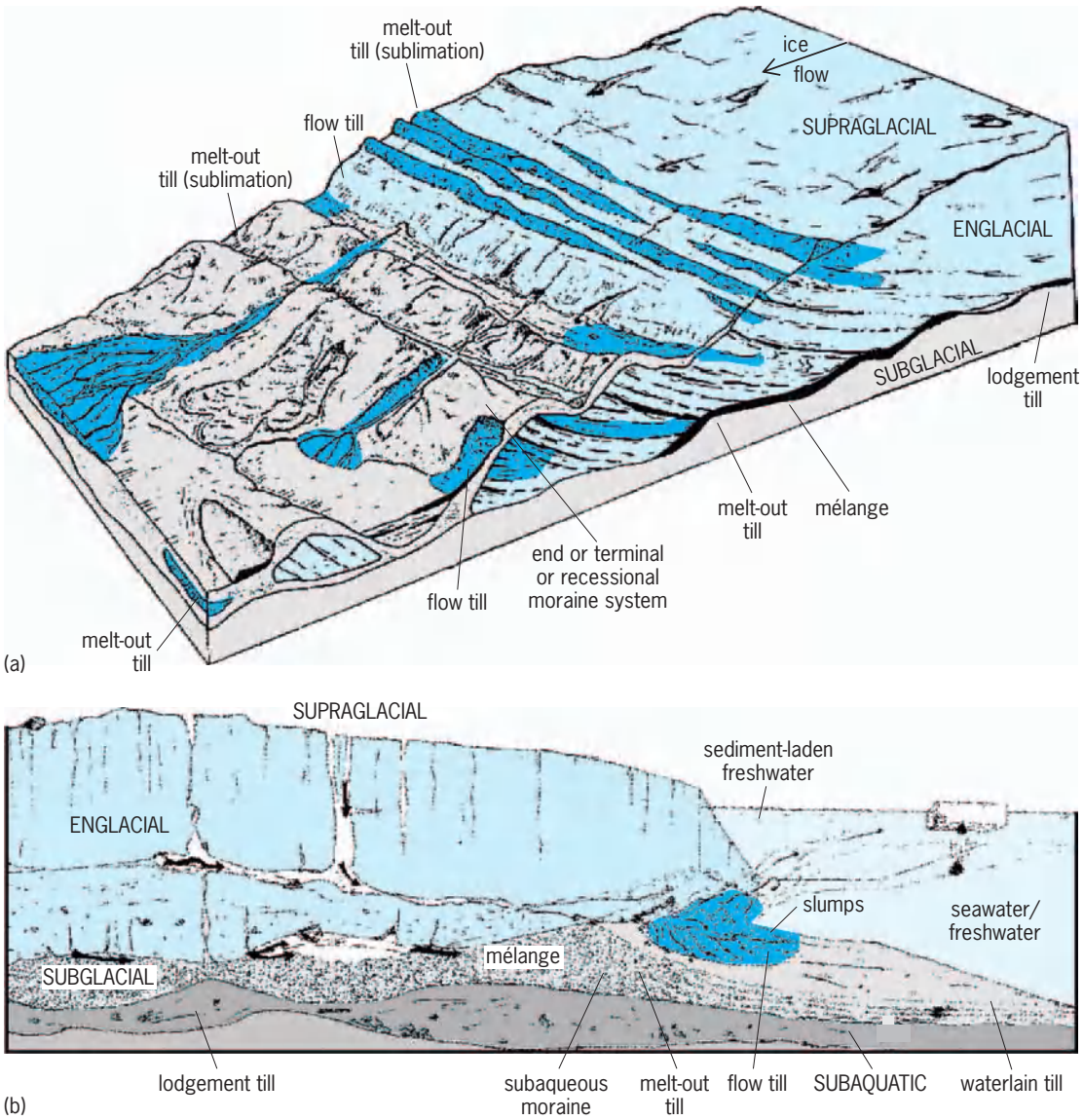

Fig. 15. Models of (*a*) land-based glacial depositional system and (*b*) depositional system of an ice mass with a marine moraine.

society, the only evidence is somewhat anecdotal, and at the close of the Late Wisconsinan may have acted to spur human migration, for example, through the short-lived ice-free corridor between the Cordilleran and Laurentian ice sheets in central Alberta, Canada, into the prairies to the south.

Within a few hundreds of kilometers of the vast ice sheets that covered North America and Europe, climatic conditions must have been severe, with strong katabatic cold winds descending from the ice. Associated with the poor climates, periglacial conditions must have prevailed in which the ground became permanently frozen to a considerable depth. The effects of periglacial activity produced frozen ground phenomena such as the movement of sediments down slopes, ice and sand wedges, localized ice lenses, and large ice-cored mound (pingo) formation. Dramatic changes in vegetation types and animal life occurred in the Northern Hemisphere, with the southern migration of many species. Unlike today, for example, central-southern Texas would have supported hardwood forests of the type now found in Ohio and Pennsylvania.

Due to the katabatic winds, fine sediments were picked up from the proglacial areas along the margins of the ice sheets, and the dust was transported away from the ice and deposited as thick, massive loess sediments (glacioaeolian sediments). Considerable thicknesses of these sediments occur in the midwestern United States, especially in Iowa and Kansas, as well as in central Europe, Hungry, and China. *See* LOESS.

Where ice advance cuts across drainage divides, rivers reduced and occasionally diverted around end moraines. In some instances, vast outpourings of meltwater (jökulhlaups) flooded from the ice fronts leading to distinctive heavily dissected terrains being formed (badlands). Such an occurrence is recorded in the Columbia River Badlands of Washington and Oregon states, where a huge lake

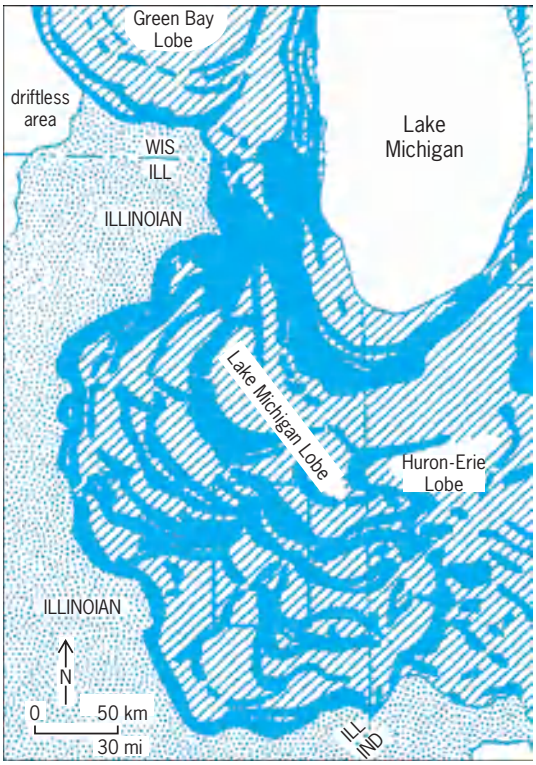

Fig. 16. Late Wisconsinan end moraines of the Green Bay Lobe, Lake Michigan Lobe, and Huron-Erie Lobe. (*Modified from J. C. Frye and H. B. Willman, 1973*)

formed and then dramatically drained (Lake Missoula floods). John Menzies

Bibliography. D. I. Benn and D. J. A. Evans, *Glaciers and Glaciation*, 1998; M. R. Bennett and N. F. Glasser, *Glacial Geology*, 1996; J. Menzies (ed.), *Modern and Past Glacial Environments*, rev. student ed., 2002; D. Mickelson and J. Attig (eds.), Glacial Processes Past and Present, Spec. Pap. 337, Geological Society of America, 1999.

Glacial history

The glacial history of the Earth is complex and extends back in geological time to the Proterozoic and possibly the Archean. Global glaciations have occurred during every geological period except the Jurassic. This persistence of global glaciation, as well as repeated and continuing glaciation in high mountainous areas, wherever altitudes exceed the local snowline, is so consistent that one might argue that the Earth is essentially a glacial planet. Evidence suggests that glacial epochs have repeatedly occurred almost every 100,000 to 150,000 years. *See* CLIMATE HISTORY; GEOLOGIC TIME SCALE; GLACIOLOGY.

Context. Since medieval times, there have been curious explanations of features that are now recognized as the result of glaciation, such as erratic boulders once regarded as the putting stones of giants, or meltwater potholes as the devil's punchbowls. The Norsemen in Iceland and Greenland, as reflected in the tenth- and eleventh-century sagas, recognized the power and impact of glacier advance on the surrounding terrain in terms of glacial erosion scouring bedrock surfaces, the advance and retreat of glaciers across pastureland, and the subsequent loss or reappearance of the land. In Switzerland, France, Germany, and Norway, the effects of glaciation were recognized by the 1700s. During the twelfth to eighteenth centuries, a cold climate period, known as the Little Ice Age, affected much of northern Europe, resulting in famines because of the drastic reductions in crop yields. During this period, glaciers expanded in the Swiss and French Alps and overran pastureland, and permanent snow beds in the highlands of Europe and in snowbound mountain passes were reported by travelers. This Little Ice Age also occurred in North America in what is referred to as the Neoglacial.

It was not realized until the 1840s that significant glaciations had affected Europe, when Louis Agassiz, professor at the University of Neuchâtel, Switzerland, and later at Harvard University, published *Études sur les Glaciares*. By the early part of the twentieth century, it was largely accepted that major glaciations had affected the Earth.

Although Agassiz's work is often hailed as the precursor to the realization that the Earth had been repeatedly glaciated during the Ice Age, it was apparent to several early geologists that ancient global glaciations could be detected from the examination of much older rocks. As early as 1855, it was suggested that the Permian rocks of Shropshire, England, were glacial in origin. Similarly, the Permo-Carboniferous Talchir boulder beds of India and others at Hallet's Cove in Australia testified to massive glaciations. In northern Ontario, the Northwest Territories of Canada, Wyoming, Virginia, Greenland, Spitsbergen, and Siberia, lithified sediments (tillites/diamictites) have been reported as evidence of ancient glaciations. Many of the locations where ancient glacial sediments have been found occur today in tropical areas such as the Amazon Basin, Namibia, and Mauritania. Long before the understanding of plate tectonics and continental drift, paradoxical questions arose as to how such glaciations could have occurred.

By the early 1900s, a fourfold sequence of glacial advance, followed by interglacial periods of warmer temperatures (akin to today's climate), had been derived for Europe based largely upon examination of the north alpine German river terraces by A. Penck and E. Brückner (see **table**). By the mid-1900s, the idea that these glaciations had spread to the other parts of Europe and North America was generally accepted, such that a fourfold set of glaciations, called the Pleistocene, became the accepted norm worldwide. *See* PLEISTOCENE.

With increasing marine exploration and the development of various dating techniques, especially oxygen isotope dating, it was increasingly clear that instead of four major global glaciations in the Pleistocene, there were more than 17 and possibly even more than 20 (**Fig. 1**). Since ocean basins are superb repositories of terrestrial sediment with little or no erosion, complete sequences of sediment

Major Pleistocene glacials and interglacials of North America and Europe following the classical system

European Alps	Northwest Europe	Britain	North America
Würm	**Weichsel**	**Devensian**	**Wisconsinan**
R/W	*Eem*	*Ipswichian*	*Sangamon*
Riss	**Warthe/ Saale/ Drenthe**	**Wolstonian**	**Illinoian**
M/R	*Holstein*	*Hoxnian*	*Yarmouth*
Mindel	**Elster**	**Anglian**	**Kansan**
G/M	*Cromerian*	*Cromerian*	*Aftonian*
Gunz			**Nebraskan**

*From bottom to top, the time sequence is from earlier to later Pleistocene. Interglacials in italics; glacials in bold.

layers, which reflect terrestrial processes and climatic conditions, are found in ocean deeps. Thus, repeated glaciation on land is well preserved and represented in the ocean sediments. *See* DATING METHODS; MARINE SEDIMENTS; PALEOCLIMATOLOGY.

Impact of glaciation. Apart from the obvious impact of glacier ice in eroding, transporting, and depositing vast volumes of sediments, glaciation has led to many indirect effects on the landscape and life. With the advance and retreat of vast ice sheets and the more localized fluctuations of mountain glaciers, considerable impact can be seen on local and regional fauna and flora. In some cases, especially at the close of the ice ages, extinction of species seems to have occurred, such as the woolly mammoth, the saber-toothed tiger, and the giant Irish elk. The migration of plants and animals occurred across the rapidly cooling and tundralike central plains of North America in advance of ice expansion from the Laurentide ice sheet and associated ice caps and fields. With the retreat of the ice, a northern migration ensued. Similar migrations can be seen in Europe, Asia, New Zealand, and South America. *See* BIOGEOGRAPHY; POSTGLACIAL VEGETATION AND CLIMATE.

In direct impacts upon the biogeography can be seen in increasing periods of aridity and later in periods of higher-than-normal precipitation (pluvials), the diversion of surface streams and rivers, and the huge fluctuations in some major river discharge regimes (for example, the Mississippi), causing significant changes thousands of kilometers downstream and into neighboring oceans. In the North Atlantic, vast outpourings of sediment and associated meltwater can be observed in the detection of Heinrich events that tell of vast changes in the subglacial glaciodynamics of the Laurentide ice sheet. As ice sheets grow and advance, crustal depression (isostatic depression) occurs. Following ice retreat, the reverse, or crustal rebound, occurs. Even today in northern United States and Canada, this rebound is ongoing. *See* ISOSTASY.

In close association with crustal depression and rebound is the fact that ice sheets act as enormous reservoirs of water evaporated from the ocean and deposited as snow on distant ice sheets. The ice sheets hold this water for thousands of years, resulting in a net decrease in ocean level (eustatic change). It is likely that sea level fell by as much as 120 m (390 ft) at the maximum extension of the last Pleistocene ice sheets. Unlike crustal rebound, which is

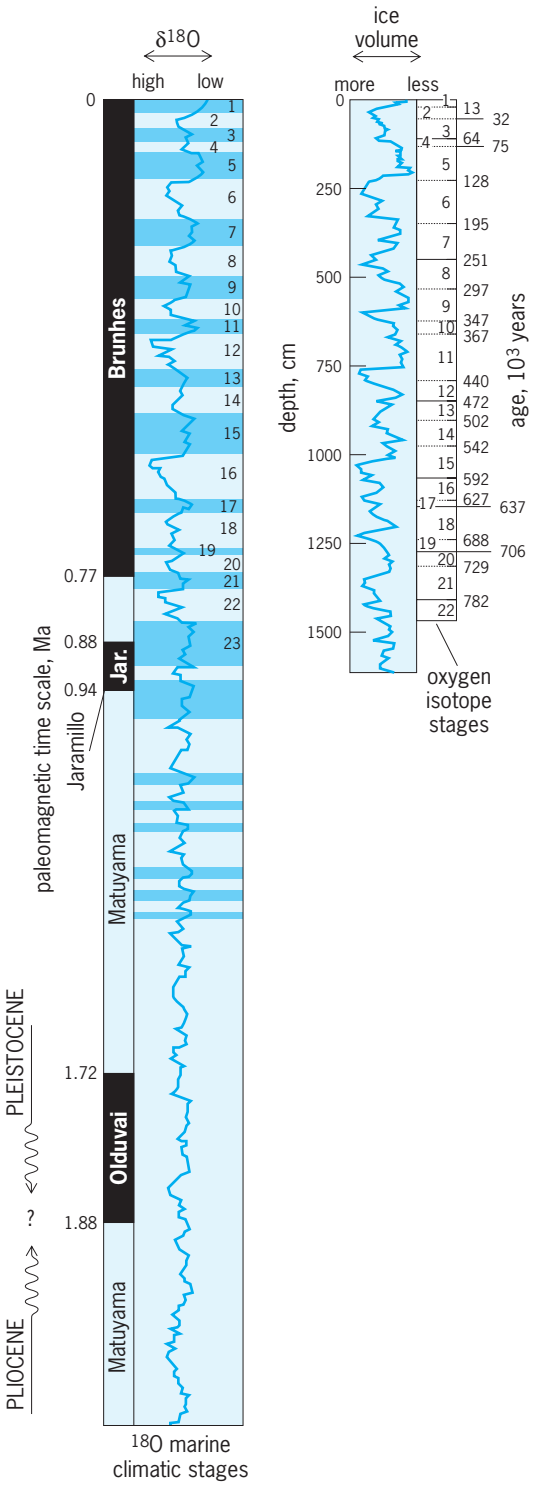

Fig. 1. Oxygen isotope records for past 1.88 million years, indicating global ice volume. In isotope stages 1 to 23, that indicates glaciations.

initially rapid but quickly slows down, sea level rises very rapidly following ice retreat and wastage. The impact of such sea-level rise is to drown coastlines, which reemerge from the sea once crustal rebound catches up as evidenced by raised beaches and abrasion platforms and exhumed spits and bars.

During and following the advance and retreat of global ice masses in the Pleistocene as well as dramatic shifts in biogeographic zones, early human evolution and later human migrations appear to have taken place. At the height of the late Pleistocene, migration of the first people into North America occurred across Beringia (dry land where the Bering Strait is now) and then southward, presumably following migrating caribou along the corridor formed between the relatively nearby margins of the western edge of the Laurentide ice sheet and eastern margin of the coalescent edges of the Cordilleran ice sheet in central Alberta and Montana. In Europe, human migration along the edges of the emerging coastlines of western Europe seems to have taken place close to the end of the Pleistocene.

Cause of global glaciation. The cause of global glaciation has long been debated. From the midnineteenth century to about 1970, a unique, nonrepetitive causation was looked for, such that ice ages were viewed as singular aberrations of unusual climatic conditions. It slowly became apparent that global climatic conditions were cyclic. Although other critical factors might intervene, the dominant causation of global glaciation must be climatic- and oceanic-based, tied directly to solar fluctuations in association with the Earth's solar orbit and tilt (**Fig. 2**). *See* EARTH ROTATION AND ORBITAL MOTION; INSOLATION; PRECESSION OF EQUINOXES.

As early as 1864, J. Croll had surmised that global cold periods leading to global glaciations were connected to the relationship between the Earth and the Sun. This idea, known as solar forcing, was expanded in the 1920s by M. Milankovitch. In the past several decades, these ideas have been refined and perfected to include the influence and impact of oceanic currents and global weather patterns in relation to Rossby waves and ocean surface temperatures (for example, El Niño), such that today a much sounder basis for understanding and potentially predicting global climate changes exists. However, the impact or onset of human-forced climatic global warming remains a central issue still debated, and its repercussions could be globally catastrophic.

Distribution of glaciers and ice sheets. The term global glaciation is perhaps imprecise. Although roughly synchronous major glaciation occurred in both the Northern and Southern hemispheres, ice never extended much beyond 40°N or 40°S latitude. The debate that surrounds the snowball Earth hypothesis suggests that ice, once north and south of 30°, developed a runaway feedback related to global albedo, causing a catastrophic glaciation that covered the whole Earth. Based upon existing geological evidence, it is likely that this hypothesis is inaccurate. *See* ALBEDO.

The distribution of ice cover during the Pleistocene is shown in **Fig. 3**, where approximately 30% of the continents are covered. If the ocean basin areas that were covered by ice are also included, the portion would rise to around 60%. Glacier buildup, leading to the development of vast ice sheets, requires high snow precipitation, low summer melt, and the survival of snow cover from winter through the following summer to the next winter. As snow accumulates, the intense reflection off the snow (its albedo) acts to maintain a low-temperature surface, reducing snowmelt and maintaining snow cover. Therefore, snow accumulates in those areas of the Earth's surface where low temperatures prevail, such as high mountain areas, polar areas, and mid- to upper-latitude continental areas, such as central Canada and central Russia. In the Suthern Hemisphere, the landmass that lies beneath the Antarctic acted as a location for gradual snow accumulation. In the

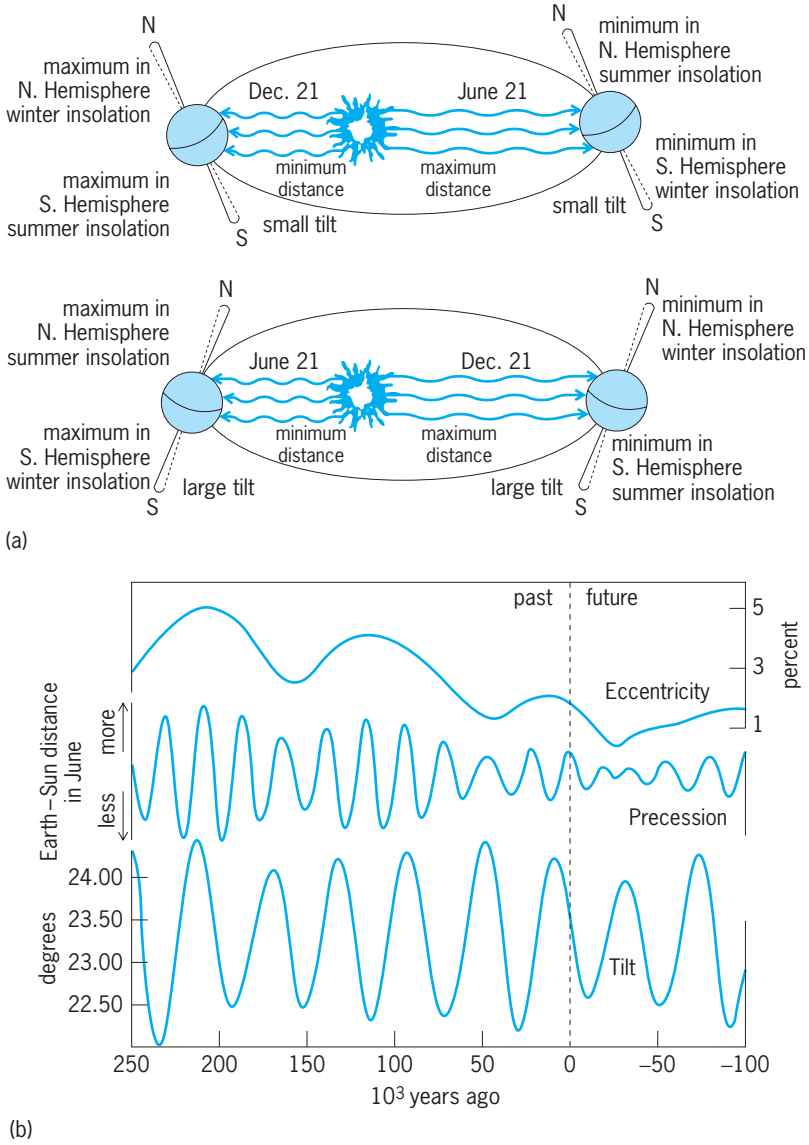

Fig. 2. Models of the Earth's orbital elements: eccentricity, tilt, and precession. (*a*) Relationship of orbital elements in a calendar year (*adapted from W. F. Ruddiman and H. E. Wright, eds., North America and Adjacent Oceans During the Last Deglaciation, 1987*). (*b*) Changes in orbital elements over the past 205,000 years and 100,000 years into the future (*adapted from J. Imbrie and K. Imbrie, Ice Ages: Solving the Mystery, 1979*)

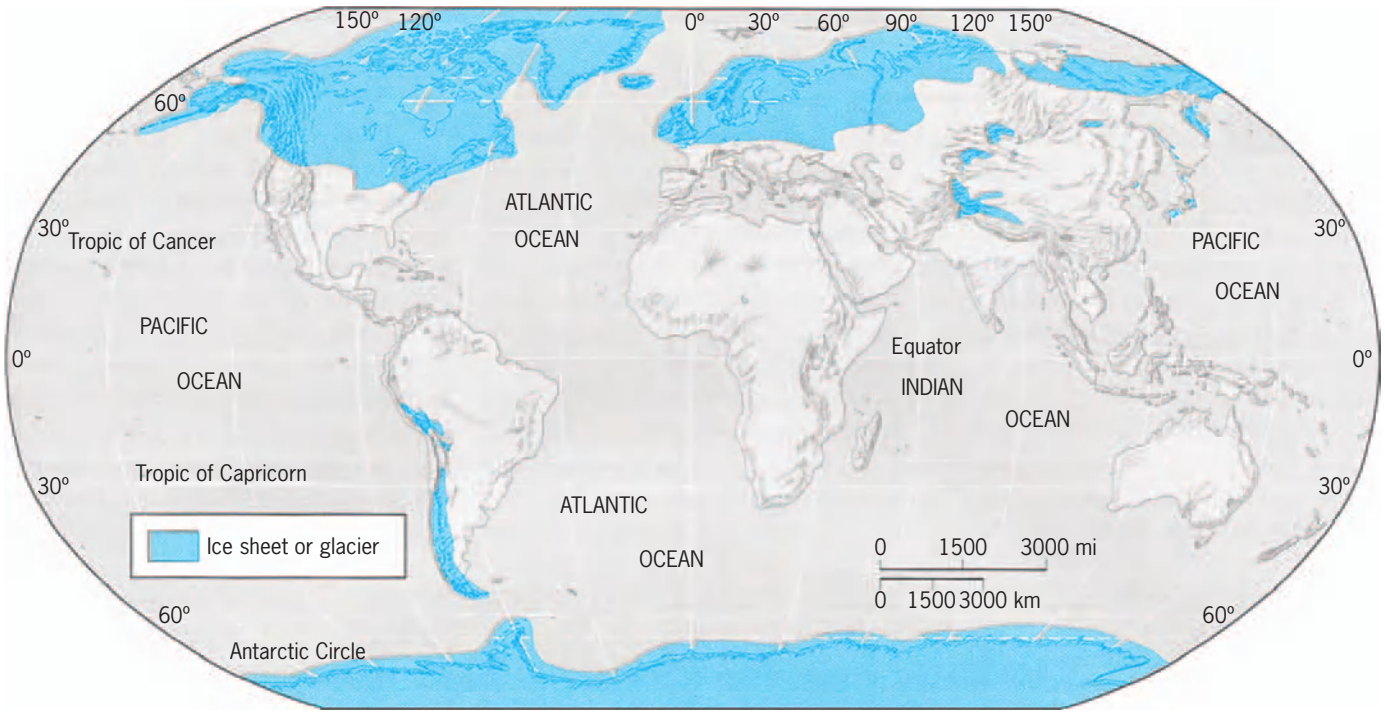

Fig. 3. Extent of worldwide glaciation approximately 18,000 years ago at the maximum of the late Wisconsinan. (*Modified from T. L. McKnight and D. Hess, Physical Geography, 6th ed., 2000*)

Northern Hemisphere with only the Arctic Ocean and no substantive landmass, snow accumulated on Greenland, the Canadian Arctic islands, central Canada, central Russia, and the Ural Mountains, as well as on high mountainous areas on all continents such as the Himalayas, the Pamirs and Tien Shan, the European Alps, the Carpathians, the Norwegian and Scottish Highlands, the Pyrenees, and the Rocky Mountains. In the Southern Hemisphere other than in Antarctica, glaciers developed in the Andes, the Southern Alps of New Zealand, the highlands of New Guinea, and the mountains of East Africa.

Evidence of glaciation. Terrestrial and marine evidence supports the glacier footprint in most parts of the world. Terrestrial clues are found in the following: specific sediment types indicative of glacial processes; the presence of erratic boulders far-traveled from a known source; distinctive landforms and landform suites (such as end moraine sequences, eskers, and drumlins); as well as the presence of buried organics, such as pollen grains, which can be traced to distinctive vegetation communities typical of arctic conditions. The presence of glacier ice on land or on fringing continental shelves is reflected in marine environments by specific sediment types with dropstones (carried by icebergs) and evidence of ice-rafted debris, as well as by the presence in stratigraphic sequences of suites of sediments containing distinctive organic remains indicative of varying terrestrial and sea temperature changes. *See* GLACIAL GEOLOGY AND LANDFORMS; MORAINE; PALYNOLOGY; POLLEN.

More recently, it has been possible to use sophisticated techniques, such as optical luminescence and oxygen isotope ratios, to date and climatically characterize specific sediment strata or horizons within sediment packages. There are instances where the specific origin of possible glacial sediment remains in doubt, even given the geological or stratigraphic context in which it is found. Considerable research remains to be done in differentiating glacial sediments both on land and in ocean basins. This problem is accentuated when attempting to characterize older lithified sediments that have suffered considerable geochemical and diagenetic changes. Glacial sediments, especially on land, often are geologically ephemeral, easily eroded or reworked, relatively thin and not widespread, and retain signatures that can easily be mistaken for sediments from other geological environments. *See* DIAGENESIS.

Pleistocene glaciations. It is now widely accepted that the Pleistocene began approximately 1.8 million years ago. It seems likely that the climate following the end of the Pliocene rapidly shifted to glacial conditions in both hemispheres. The onset of cold climatic conditions led to the buildup and the formation of the Laurentide and Cordilleran ice sheets covering most of the northern half of North America, and the European ice sheet expanded from the Alps into Germany and France. Meanwhile, the vast Fenno-Scandian ice sheet covered most of northern Europe, expanding into Poland, the Netherlands, and northern Germany. A vast ice sheet covered northern Russia, with a considerable extension across the Arctic Islands into the Arctic Ocean. In the Southern Hemisphere, the Antarctic ice sheet, which seems to have been in existence since the Oligocene, expanded into the Southern Ocean, while glaciers expanded and traveled to the coasts in South Island, New Zealand, and to the Antarctic and Pacific oceans

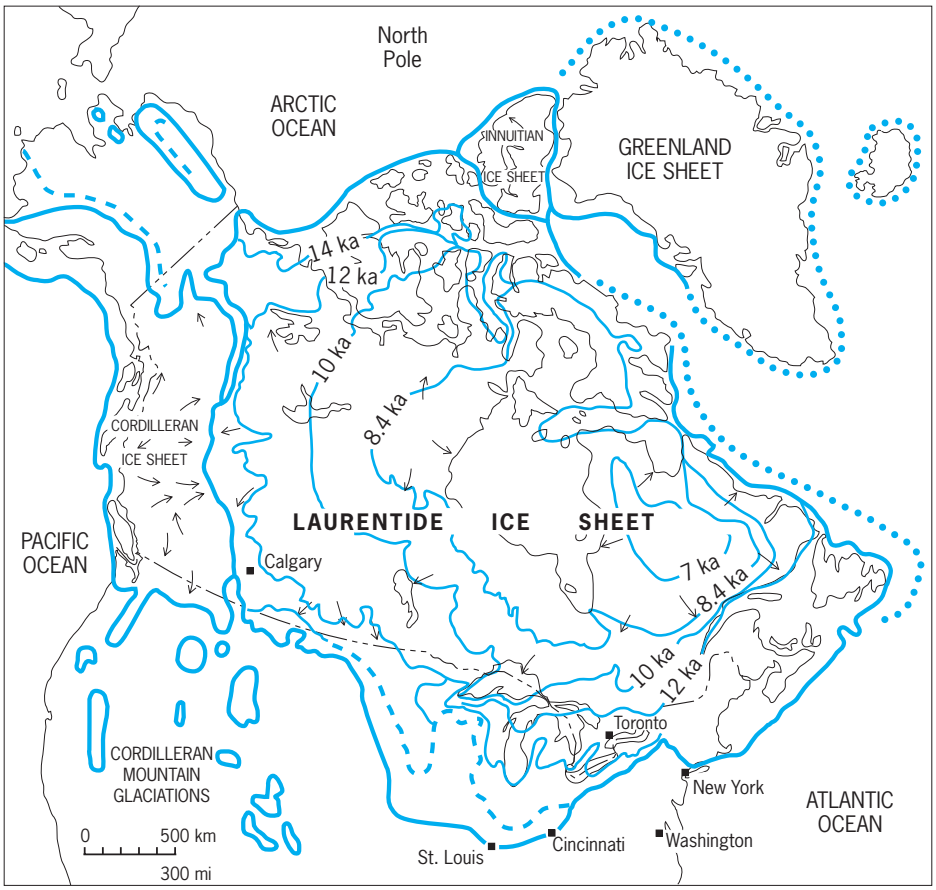

(a)

Key:

- area glaciated during Wisconsinan age
- additional area glaciated during earlier ages
- conspicuous end moraines of Wisconsinan age
- end moraines of earlier glacial ages

(glaciated area in Cordilleran region is not differentiated and is only approximate)

(b)

Fig. 4. Pleistocene glaciation. (*a*) Approximate maximum extent of glaciation and main ice sheets in North America in the late Wisconsinan 18,000 years ago and subsequent retreat margins to approximate position 7000 years ago; ka = 1000 years (*adapted from Fulton, 1989, and Clark, 1997*). (*b*) Extent of late Wisconsinan glaciation along the southern margin of the Laurentide and Cordilleran ice sheets 18,000 years ago (*modified from R. F. Flint, Glacial and Quaternary Geology, 1971*).

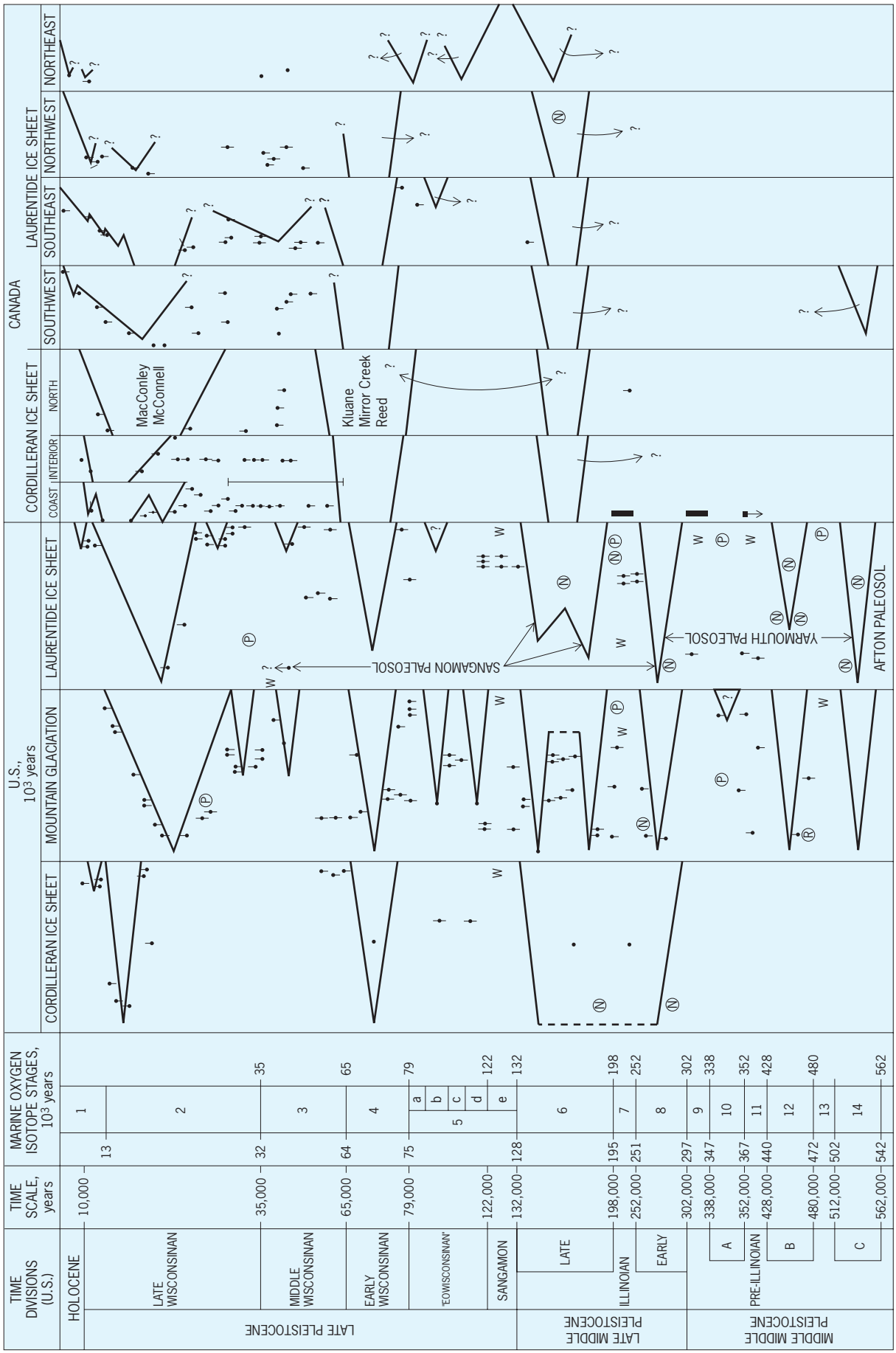

(*continued on next page*)

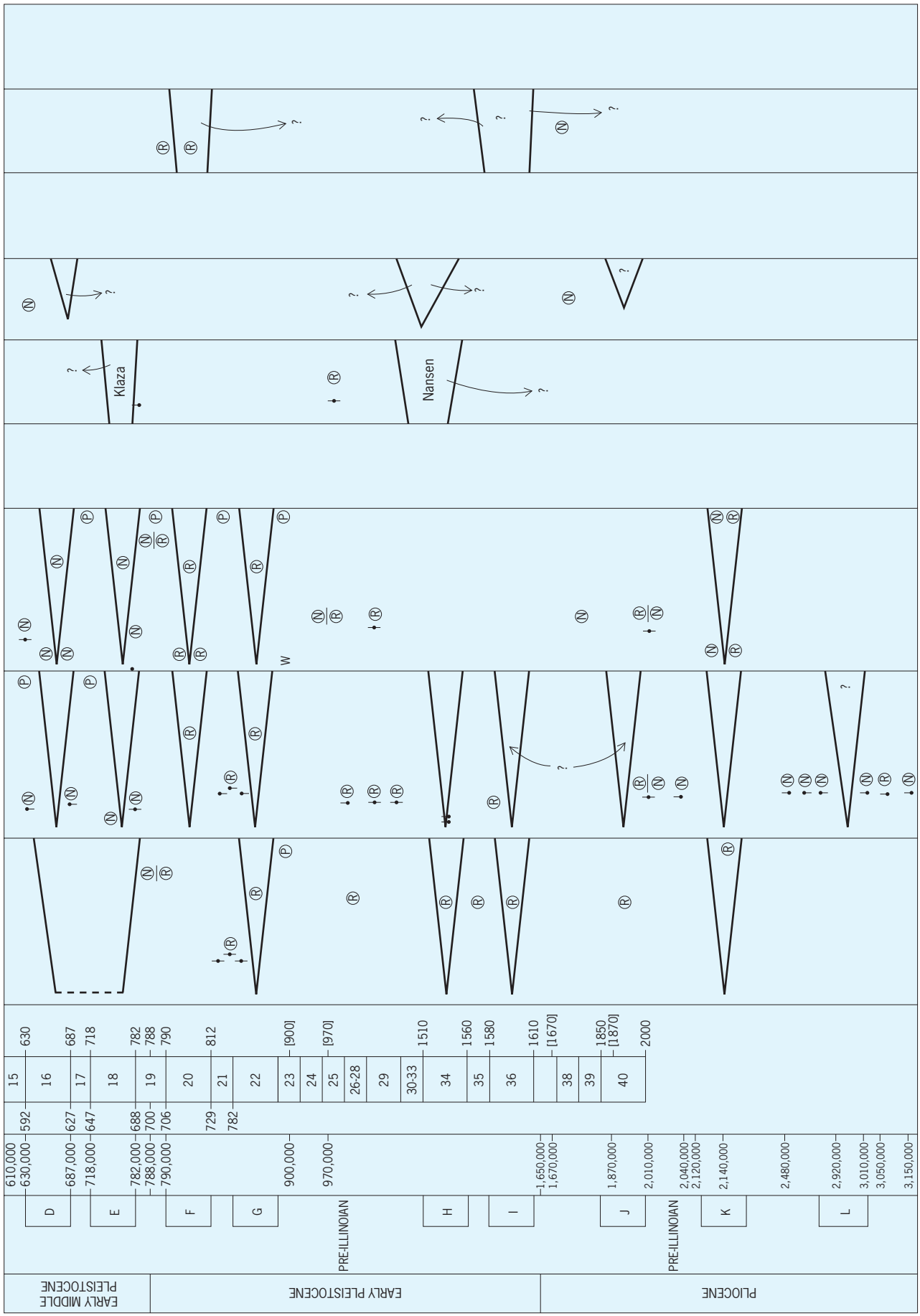

Fig. 5. Correlation chart of late Cenozoic (Pliocene and Pleistocene periods) glaciations in North America. (*Modified from Bowen et al., 1986*)

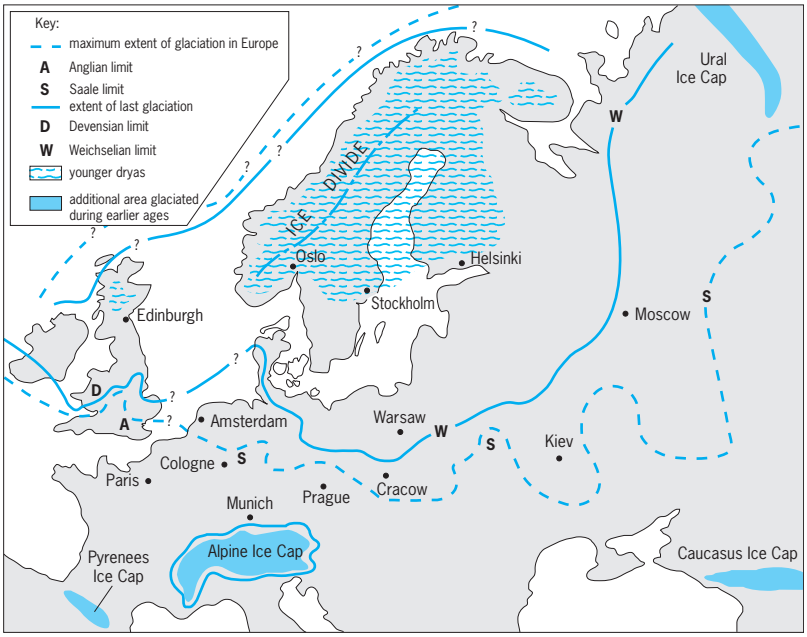

Fig. 6. Extent of late Pleistocene glaciations in Europe approximately 18,000 years ago. (*Modified from T. Nilsson, The Pleistocene: Geology and Life in the Quaternary Ice Age, 1983*)

in southern Patagonia in South America (Fig. 3). Elsewhere, smaller, independent ice sheets and ice fields covered much of the United Kingdom and Ireland, the highlands of New Guinea, the Sierra Nevadas in the western United States, other highlands in Central America, and the Andes from the tropics of Ecuador and Venezuela to the dry Andean Plateau of Bolivia and Peru. *See* OLIGOCENE; PLIOCENE.

The Pleistocene must be understood in the context of repeated phases of global glacier buildup, followed by major ice-sheet retreat and the resumption of warm interglacial times (100,000 years in length) in which climatic conditions returned to temperatures similar to today. At other times, ice-sheet retreat was much less dramatic, and interstadial conditions of cool, wet climates prevailed for short periods (10,000 years), followed quickly by ice-sheet advances (**Figs. 4** and **6**). In all the continents affected by the Pleistocene glaciations, similar advances/retreats and interglacial and interstadial phases generally occurred, but the specific details and timing of events differed not only among ice sheets but also along the various edges of the same ice sheet. At the time scale of thousands of years, there was a synchroneity. But at shorter periods, the idiosyncrasies of topographic variations, the local and regional fluctuations in ice discharge and mass balance, and local and regional basal glaciodynamic and regional weather patterns were such that synchroneity of glacial and nonglacial events, and of glacial pulses, retreats, surges, and rapid downwasting permutations caused huge variations in ice-front behavior.

The end of the Pleistocene occurred about 12,000 years ago, when the ice sheets began to melt, climate returned to conditions similar to today, sea level rapidly rose, and crustal rebound started. However, this approximate date needs considerable qualification since the ice sheet retreated from most of the northern United States somewhat earlier, following the maximum extension of the ice 18,000 years ago. In contrast, the resumption of warmer, nonglacial conditions did not begin in the northerly parts of Canada until 6000 years ago. Similar variations occurred in Europe and in the Southern Hemisphere.

With the end of the Pleistocene, the present interglacial period began. Considerable climatic fluctuations have continued to occur as witnessed by the Little Ice Age of the twelfth to eighteenth centuries, by the much warmer period of the Hypsithermal (climatic optimum) 9200–7300 years ago, and the possibly human-aided, global warming of the next few centuries. **Figure 7** shows the fluctuations in temperature and consequent ice-sheet advances for the past 150,000 years and the projections for the next 150,000 years, in which it is conjectured that following a period of global warming a sequence of glacial and interglacial periods will resume.

Pleistocene of North America. The major glaciations of North America can be divided into early, middle, and late Pleistocene. Within these time divisions, the major glacial periods of the Pre-Illinioan, Illinionan, and Wisconsinan can be subdivided according to those in the United States related to the Cordilleran ice sheet, mountain glaciation, and Laurentide ice sheet (Fig. 5). In Canada, a twofold subdivision of

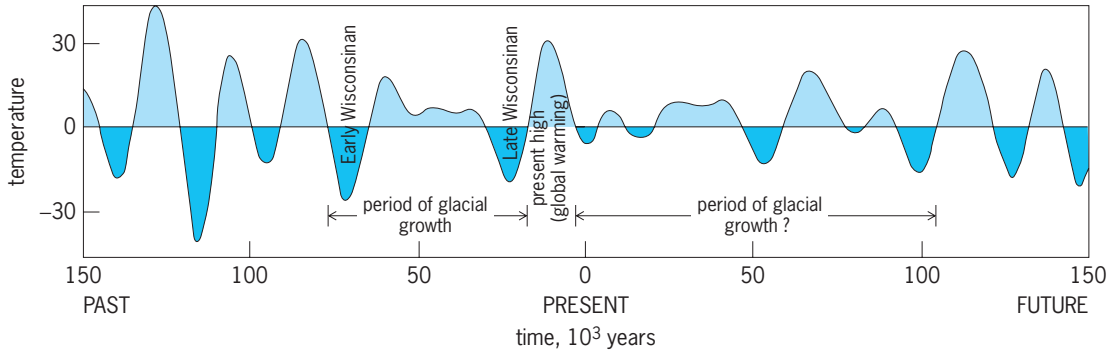

Fig. 7. Insolation curves for 45°N covering the past 150,000 years and the future 150,000 years. Past and future glaciations are shown below the zero degrees line. (*Adapted from W. S. Broecker and J. van Donk, J. Rev. Geophys. Space Phys., 8:167–198, 1970*)

Cordilleran and Laurentide ice-sheet glaciations can be distinguished.

Evidence based upon tills indicates that the Laurentide ice sheet had expanded into Iowa by the late Pliocene (glaciation K in Fig. 5). Glaciofluvial sediments in the western prairies of Canada are considered late Pliocene or early Pleistocene in age. From the early Pleistocene onward, a series of glacial advances and retreats can be traced throughout North America. The maximum extent of the last (late Wisconsinan) Laurentide ice sheet occurred along the southern margin between 24,000 and 14,000 years ago, while along the northern and eastern margins in Canada it occurred between 12,000 and 8000 years ago. The maximum extension of mountain glaciation seems to have occurred around 22,000 years ago, while the maximum extent of the Cordilleran ice sheet occurred 15,000 to 14,000 years ago. At the regional level in the Great Lakes basin, a complex sequence of events followed the retreat of the maximum ice-sheet cover approximately 18,000 years ago. Multiple lobes and ice streams formed and subdivided as the general northward retreat of the ice continued, with ice retreating from the Lake Ontario basin by around 11,000 years ago and from the north shore of Lake Superior by around 8000 years ago. *See* TILL.

The future. Predictions of future glacial periods can be made based on our past understanding of global climatic change and the tendency for the Earth's climate to swing between warmer and colder phases (Fig. 7). To date, the best climate models only suggest possible scenarios, typically based upon imperfect knowledge. All indications point to accelerated global warming, concomitant sea-level rise, and in the short time dramatic changes of climate and weather patterns, which some source; suggest are already occurring. Such changes over the next 50 to 100 years could lead to local and regional famine, coastal flooding, and socioeconomic and political upheaval at a scale never before witnessed. Glacial history is so wrapped up in the evolution and development of human civilization that it is crucial we understand past glacial history to predict potential global environmental change. John Menzies

Bibliography. R. M. Alley, *The Two-Mile Time Machine: Ice Cores, Abrupt Climate Change, and Our Future*, 2000; B. M. Fagan, *The Little Ice Age: How Climate Made History 1300-1850*, 2001; J. Grove, *The Little Ice Age*, 1988; J. Imbrie and K. Imbrie, *Ice Ages: Solving the Mystery*, 1979, reprint 2005; E. LeRoy LaDurie, *Times of Famine, Times of Feast: The History of Climate since the Year 1000*, 1971; P. A. Mayewski and F. White, *The Ice Chronicles: The Quest To Understand Global Climate Change*, 2003; J. Menzies (ed.), *Modern and Past Glacial Environments*, rev. student ed., 2002.

Glaciology

A broad field encompassing all aspects of the study of ice. While many glaciologists focus their attention on glaciers, the largest ice masses on Earth, glaciology also includes the study of ice that forms on rivers, lakes, and the sea; ice in the ground, including both permafrost and seasonal ice such as that which disrupts roads in the spring; and ice that crystallizes directly from the air on structures such as airplanes and antennas. All forms of snow research, including snow hydrology and avalanche forecasting, fall under the broad rubric of glaciology. Even planetary geologists are involved, as two of the moons of Jupiter, Ganymede and Callisto, are believed to be composed largely of ice. This article, however, will be restricted to discussion of glaciers.

Fig. 1. Storglaciären, a small valley glacier in northern Sweden.

Classification of glaciers. Glaciers are classified principally on the basis of size, shape, and temperature. Cirque glaciers occupy spectacular steep-walled, overdeepened basins a few square kilometers (1 $km^2 = 0.36$ mi^2) in area, called cirques. Most cirques are in high mountain areas that have been repeatedly inundated by ice. The cirques and the deep valleys leading away from them were, in fact, eroded by larger glaciers over the past 3 million years. *See* CIRQUE.

As a cirque glacier expands, it is usually constrained, at least initially, to move down such a valley. It then becomes a valley glacier (**Fig. 1**). Where such a valley ends in a deep fiord in the sea, the glacier is called a tidewater glacier. *See* FIORD.

In contrast, some glaciers are situated on relatively flat topography. Such glaciers can spread out in all directions from a central dome. When small, on the order of a few tens of kilometers across, these are called ice caps. Large ones, like those in Antarctica and Greenland, are ice sheets. *See* ANTARCTICA; ICE FIELD.

Thermally, glaciers are usually classified as either temperate or polar. In the simplest terms, a temperate glacier is one that is at the melting point throughout. The term melting point is used in this context rather than 0°C (32°F), because the temperature at which ice melts decreases as the pressure increases. Thus, the temperature at the base of a temperate glacier that is 500 m (1700 ft) thick will be about −0.4°C (31.2°F), but if heat energy is added to the

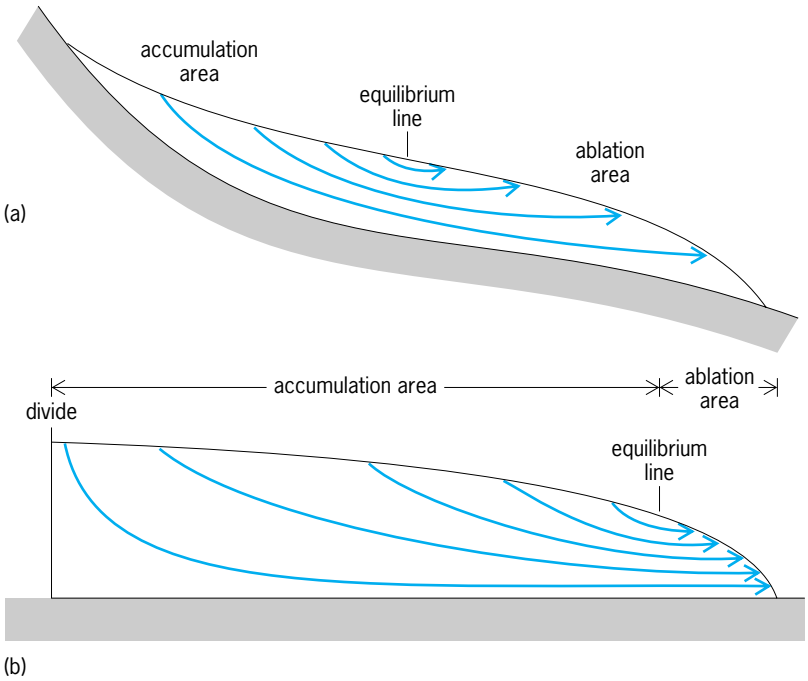

Fig. 2. Longitudinal profiles of (*a*) a valley glacier and (*b*) a continental ice sheet. The equilibrium line on the valley glacier is near the middle of the glacier, while that on the ice sheet is much nearer the margin.

ice, it can melt without an increase in temperature. Most valley glaciers are temperate.

In polar glaciers the temperature is below the melting point nearly everywhere. The temperature of a polar glacier increases with depth, however, because the deeper ice is warmed by heat escaping from within the Earth and by frictional heat generated by deformation of the ice. Thus, at its base, a polar glacier may be frozen to the substrate or may be at the melting point. Ice caps and ice sheets are normally polar, as are some valley glaciers in high latitudes.

As was the case with the classification based on size and shape, there is a continuum of thermal regimes in glaciers. The most common intermediate type has a surficial layer of cold ice, a few tens of meters (1 m = 3.28 ft) thick in its lower reaches, but is temperate elsewhere. Such glaciers are sometimes called subpolar or polythermal.

Glacier growth and mass balance. Glaciers exist because there are places where the climate is so cold that some or all of the winter snow does not melt during the following summer. The next winter's snow then buries that remaining from the previous winter, and over a period of years a thick snow pack or snowfield develops. Deep in such a snow pack the snow is compacted by the weight of the overlying snow. In addition, evaporation of water molecules from the tips of snowflakes and condensation of this water in intervening hollows results in rounding of grains. These processes of compaction and metamorphism gradually transform the deeper snow, normally known as firn, into ice. Melt water percolating downward into this firn may refreeze, accelerating the transformation. *See* SNOWFIELD AND NÉVÉ.

Ice is a mineral, like quartz or diamond. It has a well-defined crystal structure consisting of sheets of hexagonal rings stacked one upon the other so that individual hexagons lie one above another. These stacks form a tube. In the science of crystallography, the direction parallel to the axis of this tube is called the *c* axis of the mineral. *See* CRYSTALLOGRAPHY.

Crystal structures, natural or otherwise, are never perfect. Crystal lattices are distorted by imperfections called dislocations. When stressed, these dislocations move through a crystal by breaking and reforming one atomic bond at a time, resulting in deformation. This is why glaciers can flow. Were crystals perfect, many such bonds would have to break and reform simultaneously in order for deformation to occur, and this would require vastly higher stresses.

Ice does not deform significantly if the stresses are purely hydrostatic, as in a lake. However, the surfaces of snowfields are sloping, and this results in nonhydrostatic forces. These forces push the ice toward areas where the surface elevation is lower. Because the ice is weakened by dislocations, it responds to this force by flowing. When flow rates exceed a few meters per year, the ice mass is properly called a glacier.

Glaciers obviously flow downhill. Not so obvious, however, is the fact that here the term "hill" is defined by the slope of the glacier surface, not by the slope of the bed. Glaciers can in fact, flow up along a bed that is sloping in a direction opposite to that of the glacier surface.

Places where winter snowfall exceeds summer melt, which typically occurs either at high altitudes or at high latitudes, serve as accumulation areas for glaciers (**Fig. 2**). At high altitudes, the ice flows down valleys to lower elevations where temperatures are warmer (Fig. 2*a*). The part of a glacier lying in an area where summer temperatures are high enough to melt some of the ice flowing in from higher elevations in addition to all of the previous winter's snow is called the ablation area. In the case of ice caps or ice sheets at high latitudes, such as in Greenland, the ice either flows to the ocean where it breaks off readily and floats away (a process called calving), or it flows to either lower elevations or more southerly latitudes where temperatures are again high enough to melt both the winter snow and some of the inflowing ice (Fig. 2*b*). *See* ICEBERG.

The boundary between the ablation area and the accumulation area at the end of a melt season is called the equilibrium line (Fig. 2). The position of the equilibrium line for a given year can be determined by locating the upper limit of extensive bare ice just before the first major autumn snowfall. The altitude of the equilibrium line will be higher during warm years or years with little winter snowfall, and lower during cold years or years with a lot of snow fall. *See* SNOW LINE.

When, during a given year, the mass of snow added in the accumulation area of a glacier exceeds the mass of ice lost from the ablation area, the glacier is said to have had a positive mass balance. If such a situation persists for several years, the glacier will advance to lower elevations or more temperate latitudes, thus increasing the size of its ablation area

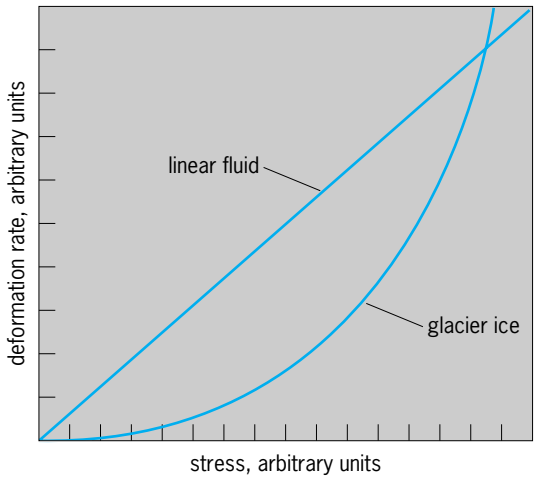

Fig. 3. Plots of deformation rate versus stress for a nonlinear substance such as glacier ice and for a linear fluid such as water.

and the mass loss. Conversely, persistent negative mass balances lead to retreat. Contrary to one implication of the word retreat, a retreating glacier does not flow backward. Rather, a glacier retreats when the ice flow toward the terminus is less than the melt rate at the terminus.

Because of time lags in a glacier flow system, valley glaciers may take decades, and continental ice sheets millennia, to respond to a change in mass balance. For example, studies of cores from the Greenland ice sheet suggest that the increase in temperature that precipitated the end of the last ice age occurred 11,600 years ago over a period of only a few years. However, the Laurentide ice sheet that covered the northern half of North America persisted until approximately 8000 years ago.

Flow of ice. Ice is a nonlinear material. Its deformation rate, $\dot{\epsilon}$, is proportional to the cube of the stress, τ, thus $\dot{\epsilon} = (\tau/B)^3$, where B is a measure of the viscosity of the ice. The cubic dependence means that at higher stresses the deformation rate increases very rapidly with any increase in stress. This behavior is distinct from that of linear fluids such as water (**Fig. 3**). *See* VISCOSITY.

Although this relation appears simple, it is not. In order to calculate τ, for example, one has to know all of the stresses acting at any given point in a glacier, and there are nine of them—three normal and six shear. (Fortunately, three of the shear stresses equal the other three, which simplifies the problem slightly.) The same applies to $\dot{\epsilon}$. In addition, B is not constant; it decreases as the temperature increases, so warm ice deforms more readily than cold ice. In ice near the melting point, viscosity also decreases with increasing interstitial water content. The water forms films along crystal boundaries, and this enables crystals to slip past one another. Finally, the viscosity is sensitive to the orientation of crystals in the ice. The reason for this is that ice deforms about 10 times more easily when the *c* axes are perpendicular to the plane of shear instead of parallel to it. Once the ice has undergone a small amount of deformation, crystals that are favorably oriented for deformation tend to grow at the expense of others. Thus, in ice that has been deformed appreciably, the viscosity tends to be lower than in undeformed ice.

Owing to these complications, sophisticated computer models are necessary for the most accurate calculations of glacier flow. Under suitable circumstances, however, approximations can be made that lead to simple but useful relations. One such result predicts that the flow rate of ice should be proportional to the fourth power of its thickness. Thus, a 10% decrease in thickness results in about a 35% decrease in flow rate. Such a decrease in flow rate would reduce the input of ice to the lower reaches of a glacier appreciably. Unless the ablation rate were to decrease by a similar amount, the glacier terminus would retreat significantly. This is the reason for the frequently observed evidence indicating that a relatively modest decrease in thickness of a glacier was associated with a substantial retreat of the snout. For example, Storgläcieren (Fig. 1) has retreated several hundred meters from its 1915 maximum position, but has only thinned a few tens of meters.

Sliding and till deformation. Both temperate glaciers and polar glaciers that are at the melting point at the base are able to move over the bed by sliding. In addition, a glacier resting on a bed of unconsolidated material often drags this material along. This material is commonly a mixture of clay, silt, sand, and gravel formed by glacier erosion and known as till. The thickness of the deforming till layer may range from centimeters to a few meters. *See* TILL.

Sliding can occur over a deforming substrate or over bedrock. Sliding speeds are normally lower over bedrock because the roughness of the surface is greater.

Two basic processes are involved in sliding: regelation and plastic flow. In regelation, ice melts on the upglacier sides of bumps, where the pressure is high and the melting point is depressed. The meltwater flows past the bump and refreezes in its lee, where the pressure is lower and the temperature is higher. The latent heat of freezing is conducted from the warmer lee side to the colder stoss or upglacier side where it melts more ice.

Plastic flow is simply the process whereby ice deforms around a bump. During plastic flow, ice may separate from the bed in the lee of a bump, leaving a cavity that is generally filled with water.

Regelation is most important on small bumps, because the path along which heat is conducted from the lee to the stoss side of the bump is shorter and the temperature gradient is higher. Plastic flow is most important on large bumps, because the stress scales with the size of the bump and the flow rate is proportional to the cube of the stress. The two processes are of roughly equal importance for bumps that are about 2–8 cm (1–3 in.) high and a meter or so in length.

Foliation. Glaciers commonly appear to be banded (Fig. 1). This appearance, a result of intercalation of layers of ice with differing bubble and dirt contents, is known as foliation. Variations in bubble content commonly arise from the vagaries of snow

accumulation and meltwater penetration, and are thus subparallel to the original layering of the snow in the accumulation area. In addition, in areas where stresses in the ice near the surface are tensile, fractures called crevasses may form. Crevasses often become filled with snow in such a way that the filling has either a higher or lower bubble content than the surrounding ice. Variations in dirt content are most common in ice near the base of a glacier, and result from episodic entrainment of the debris.

By the time ice reaches the snout of a glacier, these layers have been stretched to such an extent that they are commonly less than 1% of their original thickness, and even layers that were once nearly vertical, such as crevasse fillings, have been rotated so that they are subparallel to the bed or valley sides. This is particularly true of ice that has moved near the bed or sides. It is then usually impossible to determine the origins of variations in bubble content, but the variations themselves are still readily visible.

The planes separating layers of ice with differing bubble or dirt content have been referred to as shear planes, but this is not accurate. They may be subparallel to the plane of maximum shear strain rate, particularly near the bed, but they normally are not surfaces of localized intensive shear.

Water movement. In temperate ice there are veins along the intersections of each three neighboring ice crystals. Water can move through these veins and in so doing dissipate viscous or frictional heat. This heat melts ice, thus enlarging the veins.

Larger veins carry more water per unit of wall area, and they are thereby enlarged at the expense of smaller ones. At depths of several meters, tubes a few millimeters (25 mm = 1 in.) in diameter may be found which join downward to make successively larger conduits, forming an arborescent (treelike) network. Trunk conduits in the network eventually reach the bed. Thence, some of the water may percolate into the substrate, but most flows as streams along the bed, either in conduits melted upward into the ice or, more rarely, in channels cut downward into the bedrock.

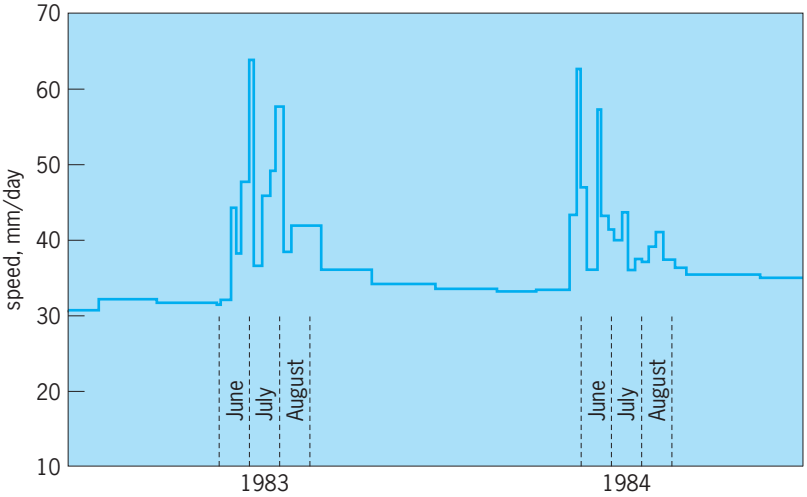

Fig. 4. Speed of a stake planted in Storglaciären (Fig. 1), over a 2-year time span.

Despite the vein system, ice near the surface of a glacier is not very permeable. However, crevasses can reach depths of a few tens of meters, and they thus give surface water access to the millimeter-scale and larger conduits at these depths. Although, initially, these conduits may not be large enough to carry all of the water tumbling into the crevasse, the viscous heat dissipated in the conduits rapidly enlarges them.

Crevasses move with the ice, and may eventually reach a place where stresses at the surface are compressive. The crevasse then closes, but the water easily maintains an opening, called a moulin, in the glacier surface.

The pressure at depth in a glacier is largely dependent on the ice thickness. However, the water pressure in conduits, although variable, is almost always lower than that in the ice. Conduits are thus constantly being squeezed closed by plastic flow of the ice. This closure is offset by melting of conduit walls by viscous heat. The equilibrium conduit size for any given water discharge is that at which closure and melting are balanced. If closure is too fast, the conduit becomes constricted and the pressure builds up in it, thus decreasing the rate of closure.

Because of the pressure distribution at the bed, subglacial streams do not, in general, follow topographic valleys in the bed. Rather, they flow along valley sides, flow uphill, and even cross ridges. Where melting of debris-laden ice forming the sides of subglacial conduits delivers to the streams more sediment than they can carry, sinuous ridges of gravel, called eskers, are deposited. Major esker systems in places such as Maine and Minnesota reveal the courses of vanished subglacial streams. *See* ESKER.

Water flow into a glacier varies on time scales of hours to months. It is highest in the summer, when snow and ice are melting and rainstorms provide additional flow. This is when the conduit system is largest. In the autumn and winter, when water inputs to the glacier decline, often becoming negligible, conduits close. Thus in the spring, when water fluxes increase again, conduits are small and the drainage system is deranged. Water pressures may then build up to high levels for a few days.

Water at a glacier bed normally has access to hundreds of cavities. When water pressures are high, each of these cavities becomes a hydraulic jack. The water pushes upglacier against the rock and down-glacier against the roof of the cavity. This additional downglacier force makes the glacier slide faster (**Fig. 4**). Such accelerations are likely to be most dramatic early in the melt season before the conduit system adjusts to renewed water input. However, they can also occur in the middle of the melt season when warm weather or rain increases the water input. Smaller accelerations may also occur on a diurnal time scale, particularly on sunny days, in response to the peak in melt that occurs in the early afternoon.

Glacier surges. Under certain rather rare circumstances, the high subglacial water pressures that develop in the spring do not dissipate quickly but

persist for weeks. This occurs under glaciers that have been thickening for several years or decades but have not advanced appreciably as a result of the thickening. On these occasions, the increase in sliding speed resulting from the increased water pressure inhibits development of an integrated subglacial conduit system, so water pressures remain high. The glacier then may advance at speeds of meters to tens of meters per day, in what is known as a surge.

During a surge, ice that has been accumulating on higher parts of the glacier is transferred to lower areas. The upper part of the glacier becomes thinner and the surface slope decreases. Eventually, a conduit system becomes established, the water pressure drops, and the surge ends. The cycle then starts over again, with years or decades of gradual thickening as the glacier builds toward another surge.

Ice streams. Flow as rapid as that in surging glaciers is also found in some parts of the Antarctic Ice Sheet. These high flow rates occur in streams of ice tens of kilometers wide and hundreds of kilometers long, and are sustained for centuries. These streams are bounded not by valley sides but by ice that is moving much more slowly. Ice Stream B, for example, which drains to the Ross Ice Shelf in West Antarctica, has a maximum speed of 825 m/yr, while ice on either side of it is moving at only 10–20 m/yr. Roughly parallel to Ice Stream B are four other ice streams, some of which are presently inactive. One of these, Ice Stream C, shut down abruptly about a century ago. The high speeds of these ice streams are attributed to slippery conditions at the bed, where high water pressures reduce friction between the ice and the bed. Changes in paths of water flow at the bed are believed to be responsible for the changes in ice stream activity. If rising sea level eventually destabilizes the Ross Ice Shelf, discharge of ice through these ice streams could lead to rapid drawdown of the West Antarctic Ice Sheet, thus accelerating the rise in sea level.

Jökulhlaups. Among the hazards associated with glaciers are jökulhlaups, or sudden releases of water from lakes dammed by glaciers. Jökulhlaup is an Icelandic word; it has entered the vocabulary of geology because such floods are common in Iceland where localized volcanic heat is responsible for the presence of deep lakes surrounded by ice. In other regions, the lakes are more commonly formed where a glacier in a trunk valley extends across the open mouth of a tributary valley.

The basic cause of jökulhlaups is easy to visualize. Ice floats in water, so when such a lake becomes sufficiently deep the water forces its way out beneath the ice dam. However, some jökulhlaups occur before the water actually becomes deep enough to float the ice.

Once water begins to escape under such an ice dam, the conduit may be enlarged rapidly, resulting in a catastrophic release and total draining of the lake. This is most likely where the surface area of the lake is large, as the lake level then remains high while the conduit is being enlarged. High lake levels result in high water pressures in the conduits, and this inhibits closure of the conduit by plastic flow. Conduits thus can become quite large. When the surface area of the lake is small, water pressures in the developing conduit may drop quickly as the water level falls, and the conduit may squeeze shut before the lake drains completely.

Among the largest floods in the geologic record are those resulting from repeated catastrophic draining of an ice-dammed lake in Montana, Lake Missoula, during the ice ages. About a third of eastern Washington was covered by these floods, which had discharges estimated to have exceeded 17×10^6 m^3/s (4.5×10^9 gal/s). Smaller jökulhlaups have occurred during recorded history in places such as Alaska, Greenland, Norway, the Alps, and the Himalayas, and in many cases they have been the cause of property damage and loss of life.

Glacier erosion. Despite the fact that ice is relatively soft, glaciers are able to erode rock as hard as granite. The two principal erosional processes are abrasion and quarrying. Abrasion occurs when rocks in the sole of a glacier are forced into contact with the bedrock over which the ice is moving, much as sandpaper may be forced into contact with a board. The contact force depends not on the thickness of the ice but on the rate of melting by geothermal and frictional heat at the bed. In the absence of such melting, a rock in the sole of the glacier would simply be forced upward into the ice. However, when melting occurs, ice flows around the rock toward the bed to replenish that lost. The drag thus exerted on the rock forces it into intense contact with the bed.

Bedrock surfaces that have been abraded are smooth on a scale of decimeters. The surface generally appears polished, but deep scratches, formed by larger stones, mar the polish. These scratches are called striations.

Quarrying is a more complicated process. It involves fracturing of the bedrock and entrainment of blocks of rock isolated by these fractures. At first it was questioned whether a glacier could break solid rock. However, it is now believed that repeated water-pressure fluctuations in subglacial cavities cause propagation of microcracks that are present in any apparently solid rock, and that these microcracks eventually join to isolate blocks. Propagation of the microcracks is a result of fatiguing, much as a wire bent repeatedly in the same place will break.

Entrainment is also facilitated by water-pressure fluctuations. Increases in water pressure in cavities transfer some of the weight of the glacier from bedrock bumps to the water. Pressure-release freezing may then occur on top of the bump. This increases the traction that the glacier exerts on the rock.

Water-pressure fluctuations are likely to be highest and most frequent in areas of crevassing, where water has ready access to the glacial drainage system. One such area of water input is a crevasse called the bergschrund that forms at the head of the glacier (**Fig. 5**). Another site is where convexities in the

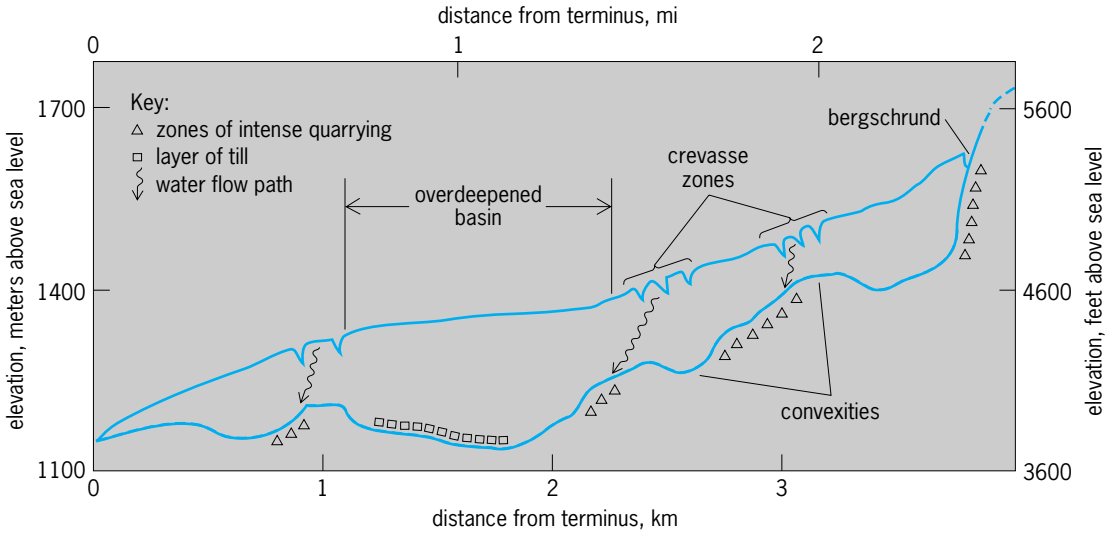

Fig. 5. Longitudinal profile of Storglaciären (Fig. 1) showing relation of crevasses to convexities in the glacier bed and to zones of intense quarrying. (*After R. LeB. Hooke, Positive feedbacks associated with the erosion of glacial cirques and overdeepenings, Geol. Soc. Amer. Bull., 103:1104–1108, 1991*)

glacier bed lead to tensile stresses and hence crevassing at the surface. Thus, the areas of most intense quarrying are the cirque headwall and the steep lee faces of these convexities. In the latter case, there is a positive feedback loop: the convexity causes crevassing that admits water to the lee face, where further quarrying accentuates the convexity.

Because it is localized on and thus accentuates steep faces, quarrying is responsible for some of the most dramatic attributes of glaciated valleys, and hence for the spectacular scenery that is associated with glaciated landscapes in mountainous areas (**Fig. 6**).

Water flowing along a glacier bed from areas of thick ice to areas of thinner ice must warm up in order to remain at the melting point. Some of the viscous energy dissipated by this water is consumed in warming it. When water flows up an adverse slope leading out of an overdeepened basin, the energy available may not be adequate to maintain the water at the melting point. Freezing then occurs, constricting the subglacial conduits and forcing the water to follow englacial (inside the glacier) paths. In the absence of subglacial water flow, the products of glacier erosion are not flushed out, but accumulate to form a protective layer that inhibits further erosion of the bed. For this reason, convexities are not scoured away, but become amplified by the erosional processes. *See* GLACIATED TERRAIN.

Fig. 6. Glaciated mountain landscape.

Glacier deposition. The load of rock debris that glaciers acquire by quarrying and abrasion is eventually deposited when the ice melts. Among the most important types of deposits are moraines. Moraines are ridges of debris deposited directly from the ice at the glacier margin (Fig. 1). They may be only a few meters in width and height or reach several kilometers in width and tens of meters in height. Because water has played little role in their deposition, they are typically heterogeneous, with abundant silt- and clay-sized material.

The time required for the formation of large moraines ranges from decades to centuries. They mark positions where the ice margin was stable for a long period of time. Where they are associated with organic material that can be dated by radiocarbon techniques, they are important in reconstructing the glacial history of an area; thus moraines are valuable in paleoclimatic studies. *See* DATING METHODS; MORAINE; PALEOCLIMATOLOGY.

Climatic record in glaciers. In inland areas of the Antarctic and Greenland ice sheets, it is so cold that there is no melting, even in summer. The snow that is deposited there, together with any impurities from the atmosphere, is compacted into ice with little or no chemical alteration. Accordingly, it preserves a record of conditions at the time of deposition. For this reason, much effort has been focused on obtaining continuous cores through these ice sheets.

One of the most important paleoclimatic signals in glacier ice is provided by the ratio of the heavy isotope of oxygen (^{18}O) to the normal isotope, (^{16}O). Snow deposited during warmer periods contains more ^{18}O. By making detailed measurements of the

ratio of these two isotopes, the difference between snow deposited in winter and that deposited in summer can be detected. Thus, much as trees can be dated by counting rings, scientists can date glacier ice that is up to several thousand years old. Using isotopes, it is easy to detect the onset and termination of ice ages (**Fig. 7**).

Recently, traces of the variations in bubble content and ice texture in annual layers have been found preserved in ice up to several thousand years old. This has enabled scientists to study the rate of climate change at the end of the last ice age. As the climate warmed, the accumulation rate on the Greenland Ice Sheet increased by a factor of nearly two, because the warm air held more moisture. Melting at the ice sheet margins, however, more than balanced the increase in accumulation and resulting increase in flow, so the ice sheet retreated.

This change in climate occurred over a period of about 3 years. An abrupt change of this magnitude today would wreak havoc with agriculture throughout the world.

Air bubbles that are trapped when snow is compacted into ice contain a record of the composition of the atmosphere. Among the constituents of such air bubbles is carbon dioxide (CO_2). By knowing the age of the ice from the $^{18}O/^{16}O$ ratio, a record of variations in atmospheric CO_2 can be reconstructed (Fig. 7). From this, it is clear that high atmospheric CO_2 concentrations are associated with warm periods in the past, and conversely. The postindustrial increase in CO_2 concentration in the atmosphere from about 280 ppm to about 360 ppm is as large as the change at the end of the last ice age, about 12,000 years ago. If this recent increase in CO_2 concentration leads to an increase in worldwide temperature comparable to that at the end of the ice age, glaciers will retreat, leading to a rise in sea level of 1–2 m (3–6 ft) in the twenty-first century. This, in conjunction with shifts in agricultural belts and other changes, could lead to major dislocations of human activities.

Another trace chemical present in ice is lead (Fig. 7). Variations in the concentration of lead in cores from Greenland indicate that lead is about 10 times more abundant in the atmosphere today than it was before humans learned how to use it. Such findings, in conjunction with medical research into health hazards posed by lead, have resulted in laws limiting its use in products such as paint and gasoline. *See* CLIMATE HISTORY.

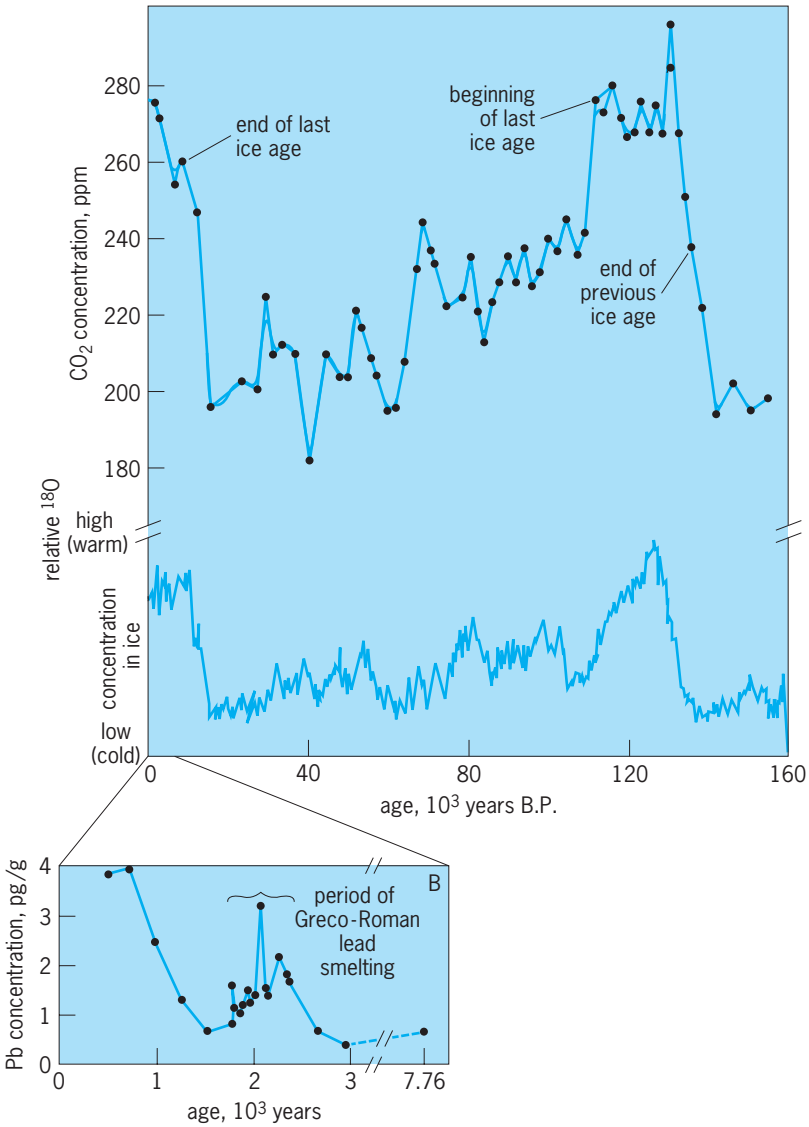

Fig. 7. Plot of $^{18}O/^{16}O$ ratio and of CO_2 and Pb variations with time in ice cores. The broad peak in lead concentration about 2000 years ago coincides with the period of maximum production of lead by the Roman Empire.

Glaciers of the ice ages. Glacier ice contains more ^{16}O than does seawater. When moisture from the oceans begins to accumulate as snow on the land, leading to growth of ice sheets, the $^{18}O/^{16}O$ ratio in the ocean increases. The $^{18}O/^{16}O$ ratios in shells of microscopic animals that live in the oceans reflect this change. Thus, through study of cores of sediment from the ocean floor, a record of the advance and retreat of continental ice sheets can be reconstructed.

This record reveals that over the past 3 million years situations like that at present, with relatively little ice outside the polar areas, are exceptional. Such comparatively ice-free periods tend to last about 10,000 years, the length of time since the end of the last ice age. Then, in response to well-known and predictable changes in the Earth's orbit that lead to cooler summers in the Northern Hemisphere, ice sheets begin to advance. The advance, punctuated by partial retreats, culminates after about 90,000 years. When roughly half of North America and Europe is again covered by ice, it will be difficult for the world of the year 90,000 to support a human population like the present one. *See* GLACIAL EPOCH; GLACIAL GEOLOGY.

Roger LeB. Hooke

Bibliography. D. I. Benn and D. J. A. Evans, *Glaciers and Glaciation*, 1997; M. R. Bennett and N. F. Glasser, *Glacial Geology: Ice Sheets and Landforms*, 1996; D. Drewry, *Glacial Geologic Processes*, Edward Arnold, London, 1986; M. Hambrey and J. Alean, *Glaciers*, 2d ed., 2004; W. S. B. Paterson, *The Physics of Glaciers*, 3d ed., 1999.

Gland

A structure which produces a substance or substances essential and vital to the existence of the organism and species. Glands are classified according to (1) the nature of the product; (2) the structure; (3) the manner by which the secretion is delivered to the area of use; and (4) the manner of cell activity in forming secretion. A commonly used scheme for the classification of glands follows.

I. Morphological criteria
   - *A.* Unicellular (mucous goblet cells)
   - *B.* Multicellular
     1. Sheets of gland cells (choroid plexus)
     2. Restricted nests of gland cells (urethral glands)
     3. Invaginations of varying degrees of complexity
        - *a.* Simple or branched tubular (intestinal and gastric glands)—no duct interposed between surface and glandular portion
        - *b.* Simple coiled (sweat gland)—duct interposed between glandular portion and surface
        - *c.* Simple, branched, acinous (sebaceous gland)—andular portion spherical or ovoid, connected to surface by duct
        - *d.* Compound, tubular glands (gastric cardia, renal tubules)—branched ducts between surface and glandular portion
        - *e.* Compound tubular-acinous glands (pancreas, parotid gland)—branched ducts, terminating in secretory portion which may be tubular or acinar

II. Mode of secretion
   - *A.* Exocrine—the secretion is passed directly or by ducts to the exterior surface (sweat glands) or to another surface which is continuous with the external surface (intestinal glands, liver, pancreas, submaxillary gland)
   - *B.* Endocrine—the secretion is passed into adjacent tissue or area and then into the bloodstream directly or by way of the lymphatics; these organs are usually circumscribed, highly vascularized, and usually have no connection to an external surface (adrenal, thyroid, parathyroid, islets of Langerhans, parts of the ovary and testis, anterior lobe of the hypophysis, intermediate lobe of the hypophysis, groups of nerve cells of the hypothalamus, and the neural portion of the hypophysis)
   - *C.* Mixed exocrine and endocrine glands (liver, testis, pancreas)
   - *D.* Cytocrine—passage of a secretion from one cell directly to another (melanin granules from melanocytes in the connective tissue of the skin to epithelial cells of the skin)

III. Nature of secretion
   - *A.* Cytogenous (testis, perhaps spleen, lymph node, and bone marrow)—gland "secretes" cells
   - *B.* Acellular (intestinal glands, pancreas, parotid gland)—gland secretes noncellular product

IV. Cytological changes of glandular portion during secretion
   - *A.* Merocrine (sweat glands, choroid plexus)—no loss of cytoplasm
   - *B.* Holocrine (sebaceous glands)—cells undergo dissolution and are entirely extruded, together with the secretory product
   - *C.* Apocrine (mammary gland, axillary sweat gland)—only part of the cytoplasm is extruded with the secretory product

V. Chemical nature of the product
   - *A.* Mucous goblet cells (submaxillary glands, urethral glands)—the secretion contains mucin
   - *B.* Serous (parotid gland, pancreas)—secretion does not contain mucin

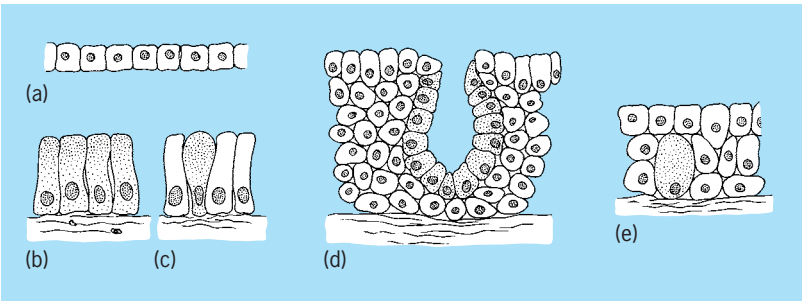

Fig. 1. Intraepithelial gland types. (*a*) Embryonic epithelium. (*b*) Epithelial sheet of glandular cells. (*c*) Unicellular goblet gland. (*d*) Multicellular gland. (*e*) Leydig's cell.

Development. Glands arise from the epithelium either by modification of cells in situ or as outgrowths. In situ glands are either unicellular and the result of cellular modifications, epithelial sheets of contiguous single cells, or intraepithelial groups or nests of cells (**Fig. 1**). Outgrowth glands are multicellular and arise as evaginations which grow into the surrounding tissue and may incorporate this tissue as part of the gland's structure. Outgrowths may be hollow or solid and of three main types: (1) those glands which detach themselves entirely from parental epithelium, such as the thyroid; (2) solid, club-shaped masses which become tubelike, and may be simple tubular, simple branched tubular, or compound tubular structures, of which the secreting portion of the gland is tubular, straight, or coiled (**Figs. 2** and **3**); and (3) bulbous outgrowths which form sac-shaped secreting units (**Fig. 4**). Each unit is called an acinus or alveolus. Such glands may be simple acinus, simple branched acinous, and compound acinous. Compound glands, either tubular (Fig. 3) or acinous (Fig. 4), are distinguished from simple glands by the presence of many ducts. The various ducts eventually join one or more common ducts which lead to the surface outlet. Compound glands incorporate surrounding tissues, including blood, lymph, and nerve tissues, into their substance as development proceeds. Such glands become large and have a tendency to subdivide into smaller divisions or lobes. Some compound glands are composed of tubular

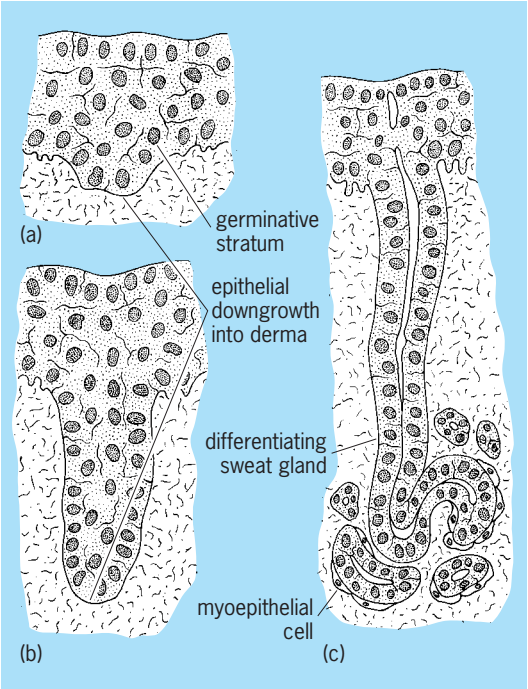

Fig. 2. Development of sweat gland as outgrowth from epithelial layer. (*a*, *b*) Epithelial downgrowth from germinative stratum. (*c*) Differentiating gland.

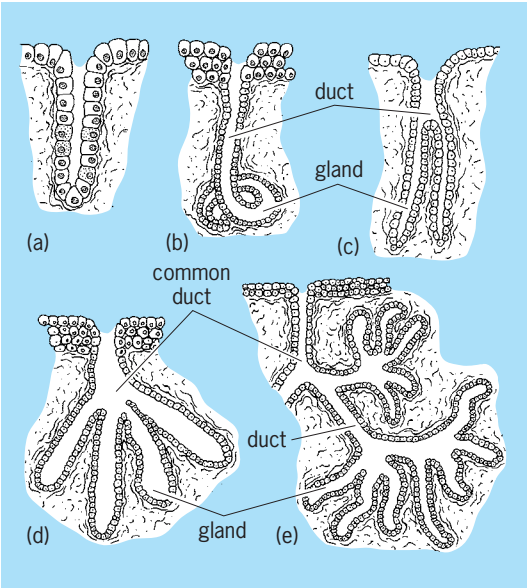

Fig. 3. Examples of tubular glands. (*a*) Simple (gland of Lieberkühn). (*b*) Coiled (sweat gland). (*c*) Simple branched (gastric gland). (*d*) Simple branched (Brunner's gland). (*e*) Compound (liver).

and acinous structures. They consist of a series of branched tubules which possess saclike, acinous outgrowths from their walls or distal ends.

Secretions. Glands may be classified on the basis of the kinds of secretions they produce. Sebaceous glands secrete oil or oily materials; serous glands, albuminous, watery material; mucous glands, a gelatinous substance; cytogenous glands liberate living cells; and granular glands produce and secrete granules.

Skin glands. All skin glands which include the in situ and epithelial-outgrowth glands are exocrine.

In situ glands are unicellular and hence represent modified single epithelial cells. They remain at the surface of, or within, the epithelial layer of the skin. Small unicellular glands together with larger single-cell glands known as club cells are abundant in the epidermis of fishes. They are mucus-secreting and odoriferous. The glands of Leydig are unicellular structures present in the epidermis of larval urodele amphibia such as salamanders and newts, and in adult urodeles, such as *Necturus*, which retain a quasi-larval condition in the adult. They are found also in gymnophionan amphibia. Leydig cells resemble the club cells of fishes. They secrete mucus and some may produce a poisonous substance. Unicellular hatching glands are present in the epidermis of the snout region of frog and toad larvae previous to hatching, and probably also in the larvae of most other amphibia. Their secretion digests the egg capsule and permits the larva to hatch.

Epithelial-outgrowth glands in their development present bulbous, epithelial outgrowths into the subepithelial tissues, and they may be classified into two groups. One group contains simple unbranched, or slightly branched acinous glands of fishes and amphibians. In most cases, they are mucus-secreting, although in some fishes and amphibians they are poison-secreting. The second group are simple unbranched, branched acinous, or tubuloacinous glands of reptiles, birds, and mammals. The secretion is thick and sebaceous. These glands are abundant in mammals and generally associated with the hair follicles. However, in some areas of the mammalian skin such as the nose, lips, nipples, upper eyelids, around the anus, and external genitals, these glands arise independently as invaginations of the germinative stratum of the epidermis. Compound

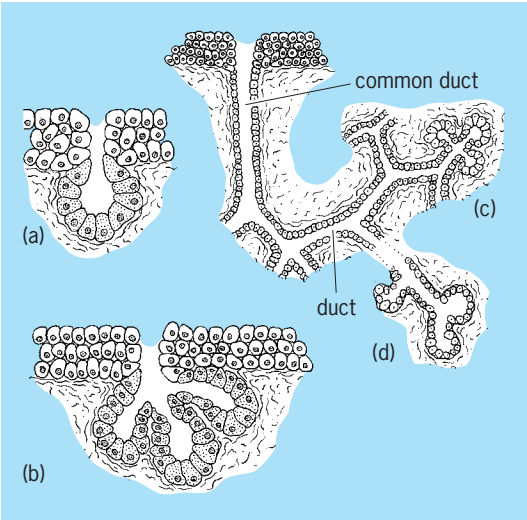

Fig. 4. Examples of typical acinous (alveolar) glands. (*a*) Simple. (*b*) Branched. (*c*) Compound. (*d*) Tubuloacinous, in which the distal secreting units are simple tubes with acinous side chambers.

tubular, nasal glands (salt glands) are situated near the eyes in marine birds and reptiles. These salt glands eliminate excess salt taken in by drinking seawater and with food. The fluid produced by salt glands is many times as salty as the bird's blood or body fluids; the kidneys produce more fluid, but with a much lower sodium chloride content. *See* ENDOCRINE SYSTEM (VERTEBRATE); LACRIMAL GLAND; POISON GLAND; SALT GLAND; SCENT GLAND; SEBACEOUS GLAND; SWEAT GLAND; UROPYGIAL GLAND.

Olin E. Nelsen

Glanders

A contagious zoonosis affecting primarily horses, mules, and donkeys and caused by the bacterium *Burkholderia* (*Pseudomonas*) *mallei*. Glanders primarily involves the respiratory systems, skin, and lymphatics. It is marked by a purulent inflammation of mucous membranes and an eruption of nodules on the skin, forming deep ulcers. Glanders was once common throughout the world but is now found only in the Mideast and parts of Africa, Russia, Asia, and South America. *Burkholderia mallei* is a gram-negative, non-acid-fast, nonsporulating, nonmotile, unencapsulated bacillus occasionally showing bipolar staining; it is obligately aerobic and oxidase-positive. It is closely related to *B. pseudomallei*, the cause of melioidosis, a disease of humans and animals in Southeast Asia and northern Australia. *Burkholderia mallei* is highly infectious for humans, who may acquire it by handling or treating glanderous animals or during laboratory investigations. Untreated acute disease in humans has a 95% mortality rate within 3 weeks. There is no vaccine against infection. *Burkholderia mallei* is considered a category B biological agent because it is moderately easy to disseminate, causes moderate morbidity with high mortality in untreated cases, and requires special enhanced laboratory diagnostic capability. *See* PSEUDOMONAS.

Epidemiology and transmission. Glanders is usually contracted by ingestion of contaminated food or water, by contact, or by inhalation of infectious droplets. The organism survives for months under most warm conditions. All equids are highly susceptible, particularly donkeys and mules. The disease is most frequent in endemic areas in the rainy season. Historically, glanders has been a scourage of military horses, in which the disease rapidly become endemic in wartime. During World War I it caused great loss on the eastern front in Russia. In World War II it was the subject of intense study as a biological weapon in Japan, Russia, and the United States, and caused heavy losses in horses and mules in the Far East.

Clinical findings and pathology. The disease is usually acute and often fatal in donkeys and mules, and chronic in horses, some of which may ultimately recover but continue to carry *B. mallei*. It is characterized by formation of nodules and ulcerations of the skin and respiratory membranes and by granulomatous nodules in the lungs, lymphatic channels, and lymph nodes. Acute disease is often rapidly fatal because of multiplication and dissemination of the organism in the bloodstream.

Humans develop a chronic or acute form with painful cutaneous ulcerating nodules, lymphangitis, and abscessation as in equids. Nodules progress to pyemia, pneumonia, and death within 3 weeks unless antibiotic treatment is provided.

Diagnosis and treatment. *Burkholderia mallei* may easily be cultured on agar medium supplemented with glycerol. The addition of bacitracin, polymixin B, and actidione is useful to suppress contaminants. The colonies are small, round, translucent, and yellowish and give off a slight fruity odor. The polymerase chain reaction (PCR) may be used for rapid, safe identification.

Although *B. mallei* is sensitive to sulfonamides and tetracyclines, affected horses are not usually treated since the destruction of cases is extremely effective for control and eradication. Essential components of diagnosis include clinical examinations at frequent intervals to detect the cutaneous and nasal forms, immunological tests to detect serum antibody, and skin and intradermopalpebral (within the skin of the eyelid) injection of mallein, a glycoprotein of *B. mallei*, to detect hypersensitivity. *See* AGGLUTINATION REACTION; COMPLEMENT-FIXATION TEST; HYPERSENSITIVITY.

Control and prevention. Early detection and elimination of infected animals is the key control measure. Detection and destruction of infected animals and regional quarantine measures have been very effective in control and eradication. Stables used by affected horses should be depopulated for several months before introducing susceptible stock.

John F. Timoney

Bibliography. A. Bauernfeind et al., Molecular procedure for the rapid detection of *Burkholderia mallei* and *Burkholderia pseudomallei*, *J. Clin. Microbiol.*, 36:2737–2741, 1998; Glanders, in *Manual of Standards for Diagnostic Tests and Vaccines*, pp. 576–581, World Organization for Animal Health, Paris, 2000; J. Lopez et al., Characterization of experimental equine glanders, *Microb. Infect.*, 5(12):1125–1131, 2003.

Glass

Materials made by cooling certain molten materials in such a manner that they do not crystallize but remain in an amorphous state, their viscosity increasing to such high values that, for all practical purposes, they are solid. Materials having this ability to cool without crystallizing are relatively rare, silica (SiO_2) being the most common example. Although glasses can be made without silica, most commercially important glasses are based on it. *See* AMORPHOUS SOLID.

Glass products are enormously varied, including windows, doors, bottles and vials, tableware, optical instruments, fiber optics, mirrors, colored filters, chemical apparatus, pipe, crucibles, thermal and

electrical insulation, cloth, and automobile, boat, and aircraft bodies. It is even possible to make a glass in which a photographic image can be developed.

Chemical properties. Chemically, most glasses are silicates. Silica by itself makes a good glass (fused silica), but its high melting point (3133°F or 1723°C) and its high viscosity in the liquid state make it difficult to melt and work. Fused silica products are expensive and are used only when their special properties (low thermal expansion, high softening point, light transmission characteristics, and corrosion resistance) are essential. *See* SILICA MINERALS.

To lower the melting temperature of silica to a more convenient level, soda (Na_2O) is added in the form of sodium carbonate (soda ash). This has the desired effect, but unfortunately the resulting glass has no chemical durability and is soluble even in water (water glass).

To increase durability, lime (CaO) in the form of calcium carbonate (limestone) is added to the glass to form the basic soda-lime-silica glass composition that is used for the bulk of common glass articles, such as bottles and window glass. Although these are the main ingredients, commercial glass contains other oxides (aluminum and magnesium oxides) and ingredients to help in melting, oxidizing, fining, or decolorizing the glass batch.

Special kinds of glass have other oxides as major ingredients. For example, boron oxide (B_2O_3) is added to silicate glass to make a low-thermal-expansion glass for chemical glassware that must withstand rapid temperature changes, for example, Pyrex glass. This type of glass is known as a borosilicate. Also, lead oxide (PbO) is used in optical glass and lead crystal because it gives a high index of refraction. Many other special types of glass have been developed for particular uses; some of these are discussed in this article.

Structure. Physically, glass is an arrangement of atoms quite like that in the liquid state; that is, it has no long-range order. In a crystal, the atoms are arranged in a regular, repeating pattern, but in a glass, although the arrangement in the neighborhood of a single type of atom may be the same throughout the glass (four oxygens around each silicon), the overall structure lacks periodicity of atomic arrangement (**Fig. 1**).

Melting. Production of glass articles begins with batch mixing of raw materials (sand, soda ash, and limestone) and their melting. For small production and special glass, melting may be done in pots or crucibles containing up to 1 ton of glass. Several pots may be heated by one central furnace into which the pots are inserted and later removed through large doors.

Larger batches are melted in large covered furnaces or tanks to which heat is supplied by a flame playing over the glass surface. Usually, these glass tanks are fired regeneratively; that is, the hot exhaust gases pass through an open brick lattice (checker) on one side of the furnace and heat it. After about 30 min, the flow is reversed, combustion air is brought in through the hot checkers and preheated, and the exhaust gases go out through checkers on the opposite side of the tank. Regeneration conserves energy and increases the flame temperature. Most glass tanks are fired by gas or oil; however, auxiliary heating with electricity (boosting) is common in the United States and all electric melting is widely used in Europe.

A glass tank may be either a batch (day tank) or a continuous type. The latter is divided into two sections joined by a narrow passage or throat. The raw materials and the scrap glass (cullet) to be remelted are charged continuously into one end of the first or melting section, where they are melted, mixed thoroughly, and refined (freed from bubbles) at high temperature (2910°F or 1600°C). The molten glass is then cooled as it passes through the throat to the second or conditioning section and exits the tank. It is cooled to the temperature (about 2012°F or 1100°C) for forming (shaping), and is taken by forehearths to the forming operation, which is different for each type of product. Continuous tanks range in capacity from 50 to 500 tons (45 to 450 metric tons) of glass per day.

Heat treatment. After forming, glass must be slowly cooled or annealed, usually in a long oven called a lehr. The purpose of annealing is to reduce the internal stresses, which can degrade the strength or optical properties of glass; for example, large internal stresses can cause the glass to crack during cooling and make the glass birefringent. These stresses

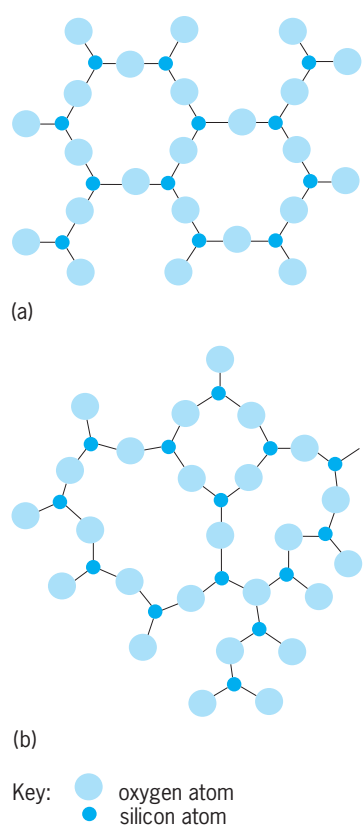

Fig. 1. Comparison of arrangements of atoms in crystals and glass. (*a*) Regular arrangement of atoms in crystal. (*b*) Disordered arrangement of atoms in glass.

are created when, because of temperature variations throughout the piece, different parts of the ware become rigid at different times. The opposite of annealing is tempering, in which the glass is rapidly chilled by a blast of air. This results in large compressive stresses in the surface of the glass. Because glass invariably fails by tension, this compressive stress must be overcome before the glass breaks. Thus tempered, glass is very strong and is used for doors and windows. Glass that is annealed usually breaks irregularly into fragments called shards. If tempered glass breaks, it shatters into many small pieces of roughly rectangular shape because of high internal stresses.

Products. The most common glass products are container ware and flat glass. The latter is principally used in transportation and architecture. Flat glass is produced by the float process, developed in the 1950s and considered one of the important technological achievements of the twentieth century (**Fig. 2**). Here a stream of viscous glass from a long melting furnace is poured continuously (floated) onto a shallow pool of molten tin at about 1920°F (1050°C). The glass spreads to a width of 130 in. (3.3 m), forms a sheet or ribbon with parallel surfaces, and attains a natural equilibrium thickness of about 0.25 in. (6.8 mm). As it continues down the tin bath, it cools and is sufficiently solid to be lifted off the bath exit at 1200°F (650°C). It then travels through an annealing lehr before being cut to size. Thinner glass may be made by stretching the ribbon with top rollers, and thicker glass may be made by partially damming the glass on the tin surface. The tin bath must be sealed to maintain an inert atmosphere and prevent oxidation of the tin.

Float glass has completely displaced two older flat-glass processes. Sheet glass with a wavy but brilliant fire-polished surface is made by drawing molten glass upward and solidifying it; and plate glass with parallel surfaces is made by drawing glass through rollers, cooling and annealing it, and then grinding and polishing both surfaces. Advantages of float glass are high optical quality, no waviness, fire-polished surface, and large thickness range [less than 0.04 in. (1 mm) to over 1 in. (25 mm)]; the process is continuous, and well suited to computer control and on-line coating application.

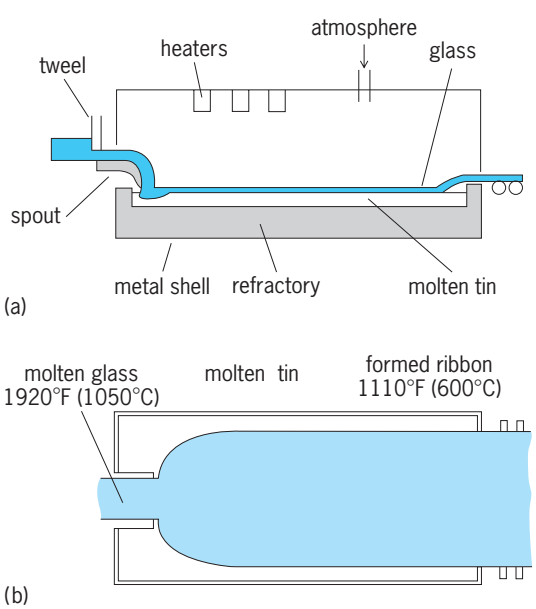

Fig. 2. Float process for manufacture of flat glass. (*a*) Float bath. (*b*) Glass sheet with equilibrium thickness. (*After W. C. Hynd, Flat glass manufacturing processes, in D. R. Uhlmann and N. J. Kreidl, eds., Glass: Science and Technology, vol. 2: Processing 1, Academic Press, 1984*)

Flat glass has come to be considered a commodity, and value-added products are made by bending, tempering, laminating, encapsulating, double-glazing, and coating the glass. Glass is bent for greater freedom of form, and it is tempered to make it stronger and safer. Laminated safety glass used in windshields is made by bonding two sheets of thin glass (0.08–0.12 in. or 2–3 mm) with a transparent organic material under heat and pressure. Thin film coatings on flat glass modify light and heat transmission; they may be electrically conductive, reflective, semireflective, antireflective, colored, low-emissivity, and so forth. Electrochromic coatings darken the glass when voltage is applied. *See* DIELECTRIC MATERIALS; ELECTROCHROMIC DEVICES; OPTICAL MATERIALS.

Container ware is largely made on automatic machines; some special pieces may be made by hand blowing into molds, and some very special ware may be made by the free or offhand blowing of a skilled glass blower. In the automatic process (**Fig. 3**), a stream of glass flows from the forehearth and is cut by shears into individual gobs, which are fed to a blank mold. Here the gob is formed into a rough blank, or parison, by either a plunger or compressed air; at this stage the opening of the bottle gets its final shape. The blank mold opens and the parison, supported by the bottle opening, is transferred to the final or blow mold, where it is blown to shape. The bottle is then transferred to a traveling belt, which carries it through the annealing lehr and on to automated inspection and packing.

Electric light bulb envelopes are molded on a special machine which converts a fast-moving ribbon of glass into over 10,000 bulbs per minute.

Glass fibers are either continuous for reinforcement or discontinuous for insulation. Continuous fibers are usually made by extruding molten glass through multiple (about 2000) nozzles in a platinum bushing built into the bottom of a forehearth; they may also be made by remelting and extruding glass marbles. After extrusion, fibers are drawn rapidly [up to 240 ft/s (80 m/s)] to attenuate their diameter, coated with bonding and protective chemicals (collectively called a size), and wound together on a spool or collet. Most continuous fiber is an E-glass composition, which contains the oxides of silicon, boron, calcium, and aluminum but no soda or other alkali. Continuous glass fiber is used to weave glass cloth or is chopped and used to reinforce plastics.

Glass wool is most often made by spinning, or forcing molten glass centrifugally out of small holes (around 20,000) in the periphery of a rapidly rotating steel spinner, and attenuating the fibers by

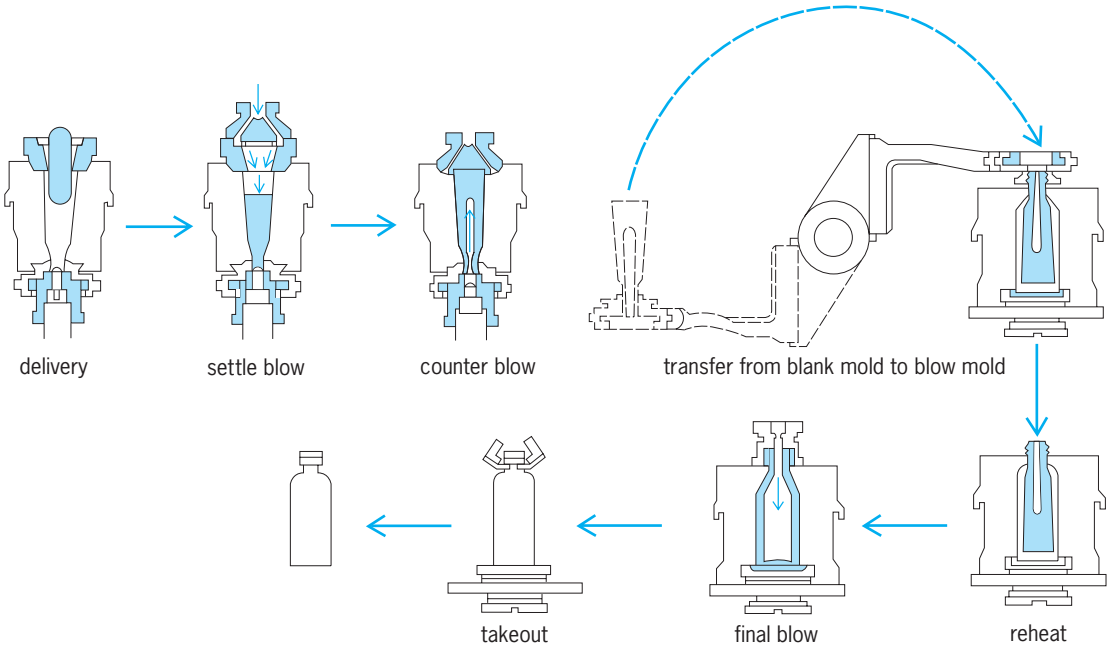

Fig. 3. Steps in automatic bottle manufacture. (*Glass Packer*)

entrainment with gas burners. As the fibers fall to a conveyor, they are sprayed with a binder that preserves their open structure. The resulting glass wool mat is mainly composed of still air, which accounts for its excellent insulation properties. Since these fibers are of extremely small diameter (about 5 micrometers) and thus have a large surface area, they are susceptible to chemical corrosion. Glass wool is similar in composition to window glass, but contains 5–10% boric oxide to increase the chemical durability. Many other processes exist for making discontinuous glass fiber. When made from impure natural materials or from slags, it is often called rock or mineral wool.

Optical fibers are hair-thin transparent glass fibers that transmit light around curves and over long distances by total internal reflection. Many fibers are combined into one fiber bundle or cable. A fiberscope allows image transmission for remote inspection or in medical procedures. Fiber optics for telecommunication carry digital pulses of light, which may be amplified by using a fiber laser amplifier. Power fibers carry high-power laser light used in welding or surgery. Fiber-optic sensors measure ambient conditions wherever a fiber optic can be inserted, such as pH and chemical composition in a reaction vessel, air quality, or blood gases and pressure. *See* FIBER-OPTIC SENSOR; FIBER-OPTICS IMAGING; OPTICAL FIBERS.

Foam glass, used where a self-supporting insulation material is needed, is formed by creating a large volume of small bubbles in a matrix of glass.

Optical glass is generally melted in platinum-lined tanks or pots, from which it is drawn or cast into blanks which can be ground to final shape. An alternate method is to cool the glass in the pot, break it apart, and select flaw-free pieces. Because an important requirement in optical glass is clarity, the purest raw materials are used.

Special glass. Besides the fairly common types of glass mentioned above, there are hundreds of special types; for example, the milk-white opal glass used for dinnerware and bottles, which owes its appearance to tiny crystals grown in the glass.

The process of crystallization in a glass, known as devitrification, is generally avoided; however, in some cases it is desired. An example is a type of glass ceramic that is shaped in the same manner as an ordinary glass, and then, by appropriate heat treatment, converted to a completely crystalline material. It is transparent because of the extremely small crystallite size (60 nanometers) and has outstanding thermal shock resistance since its thermal expansion is nearly zero (comparable to fused silica).

Another special glass is sensitive to light, which creates a latent image within the glass, the image being developed by heat treatment. This glass is used to make signs, and radio and clock faces. A variant of this glass is one in which the portion exposed to light is more soluble than the unexposed portions; it can be used to form, photographically, a fine intricate mesh or screen.

So-called solder glass has a relatively low softening point (below 930°F or 500°C). It is used to join two pieces of higher-melting glass, without softening and deforming them. The solder glass is applied to the joint as a powder and melted to produce a seal, as in a vacuum tube.

Vycor glass is a nearly pure silica glass formed without the production problems of fused silica. The starting glass, a soda borosilicate, is formed and, under heat treatment, separated into two phases, one of which can be leached out. The remaining porous material is nearly pure SiO_2 and can be converted

to a dense, clear glass by heating. During the latter treatment, it shrinks considerably.

Chemical vapor deposition produces very pure silica glass for fiber optics by hydrolysis of silicon tetrachloride in a flame and densification of the silica soot so produced. *See* VAPOR DEPOSITION.

Properties. The most important properties are viscosity; thermal expansion; strength; index of refraction; dispersion; light transmission (both total and as a function of wavelength); corrosion resistance; and electrical properties. These properties depend on glass composition, so glasses with specific values of these properties can often be tailor-made for specific applications.

The viscosity of glass is important mainly in its manufacture. The viscosity increases continuously as the temperature decreases, the logarithm of the viscosity being nearly proportional to the reciprocal of the temperature. Important points on the viscosity-temperature curve are the melting range (viscosity about 100 poises), the working range (10^3–10^6 poises), the softening point ($10^{7.6}$ poises), the annealing point (10^{13} poises), and the strain point ($10^{14.5}$ poises); the annealing range lies between the last two points. Traditionally the unit for viscosity is the poise (a relic cgs unit) but the pascal-second (Pa-s) the SI unit, is used increasingly in glass science; 10 poises = 1 Pa-s.

The coefficient of thermal expansion of various glass families ranges from close to zero to over 200 (units of $10^{-7}/^{\circ}C$). Silica, titania-silica ultralow-expansion glass, and specially designed glass ceramics have a coefficient of thermal expansion of nearly zero (from 0 to 5.5) and are used for telescope mirror supports, where good dimensional stability on temperature change is necessary. Glass ceramics are used for cook tops and ovenware, where outstanding thermal shock resistance is necessary. Common soda-lime glass has a coefficient of thermal expansion of 92, and it is subject to breakage when cooled rapidly. Alkali borosilicates have a coefficient of thermal expansion of 32 and are intermediate in thermal shock resistance. In making glass-to-metal seals and in enameling glass onto metal, it is important to match the coefficient of thermal expansion of the metal with that of the glass; this is generally possible because the thermal expansion may be fixed at any desired value by adjusting glass composition suitably. *See* THERMAL EXPANSION.

The theoretical strength of glass is extremely high, about 40 gigapascals (GPa; 6,000,000 lb/in.^{2}), but this is reduced markedly in practice because of the presence of many submicroscopic flaws on the glass surface. These act as stress concentrators and effectively control practical glass strength. The flaws may be mechanical in origin (handling the glass surface inevitably introduces flaws) or particulate (surface devitrification, dust, metal particles from forming equipment, and so forth). High-strength glass fibers may have strengths as high as 14 GPa (2,100,000 lb/in.^{2}); commercial glassware is 50–150 megapascals (MPa; 7000–22,000 lb/in.^{2}); and abraded glass may be as low as 15 MPa (2200 lb/in.^{2}). Hard oxide and lubricating coatings are applied to glass bottles to reduce friction, thereby reducing surface damage and maintaining bottle strength.

Glass articles may be strengthened by prestressing the glass to put the surface in compression. This may be done by several means: thermal tempering or toughening by quenching (automotive glass except windshields; doors; tableware), chemical strengthening (eyewear, basketball backboards), enameling, lamination of core and clad glasses with different thermal expansions, or controlled surface devitrification. Tempered glass is most common; it is five times as strong as annealed glass and is more resistant to thermal shock.

Glass breaks with no warning by brittle fracture. The low practical strength of glass and its brittle fracture are perhaps the most serious limitations on its use as an engineering material.

Index of refraction and its variation with wavelength (dispersion) depend on glass type and composition. Crown glasses (such as soda-lime glass) have low dispersion, while flint glasses (such as lead glass) are higher. Because of the dispersion of colors by glass, the image produced by white light through a simple lens is blurred, a phenomenon called chromatic aberration. This is reduced considerably by combining a crown glass with a flint glass in a compound lens. *See* REFRACTOMETRIC ANALYSIS.

Light transmission is important for materials such as colored glass and filters. However, it also determines the tint of ordinary window and bottle glass. The main coloring impurity, iron oxide, enters the glass via impure raw materials and imparts a greenish tinge, which may be seen by looking edgewise through "clear" float glass. In nominally clear float glass the iron concentration is 0.1% iron(III) oxide (Fe_2O_3). For optical glass, in which maximum transmission is needed, the iron oxide must be kept below 0.01% through beneficiation of raw materials, and its oxidation state must be controlled in order to minimize absorption. In fiber optics, with even lower absorption requirements (about 0.1 dB/km), impurities must be held to parts per million. If a clear glass is desired and total transmission is not important, it can be achieved by adding a so-called decolorizer such as selenium or cobalt, which colors the glass blue and balances the green iron absorption; the resulting glass has a light gray coloration that cannot be easily detected by the eye.

The electrical conductivity of glass, as for most insulators, increases with temperature, the logarithm of the conductivity varying linearly with the reciprocal of the absolute temperature. The surface of glass can be made conducting by a transparent tin oxide coating.

Dimensional stability is of importance in precision instruments, such as clinical thermometers, in which a delayed dimensional change in the glass can destroy the calibration of the instrument. If the glass is not carefully annealed and aged, it undergoes a compaction with time because the atoms slowly draw nearer to each other, a condition which is more stable at room temperature. This leads to changes in properties such as density, index of refraction, and strength.

J. F. McMahon; J. Wenzel

Colored and stained glass. Stained glass is colored by any of several means and assembled to produce a varicolored mosaic or representation. Stained glass windows, with the pieces arranged either to produce a pattern or a picture, have long been a decorative feature of buildings and especially old churches. Stained glass is also assembled into chandeliers and ornamental objects. The glass is colored by one of three means: fusing with metallic oxides, enameling, or painting. The addition of small percentages of metal compounds, usually oxides, to molten glass or pot metal produces throughout the glass a color characteristic of the compound.

The glass can also be coated with transparent metallic oxides and fired to bond the enameling to the glass base. Transparent pigments can be painted onto the glass in solid colors, or in several colors to render a picture; the pigments are then baked or burned onto the glass. With a hot iron, the glazier cuts the glass to shape, joins it into the design with channeled lead strips, and solders the joints. Either the tracery of the window or iron frames support the finished glaziery.

From the ancient art of making stained glass has evolved the modern technology of controlling the transmittance and color of glass. The process of adding metallic oxides to glass is widely used to produce colored glass for other than decorative purposes. Among these technological uses are the screening out of ultraviolet portions of solar radiation by display-window glass to decrease fading of objects on exhibition, the filtering out of infrared radiation especially in motion-picture and slide projectors by means of heat filters, and the reducing of visual contrast by means of a tinted auto windshield. Enameled glass is used for washable wall coverings and for luminous signs, and colored lenses in traffic lights, railroad and airport signals, and brake lights.

Melting highly colored glasses with flames poses problems because radiation from the heat of the flame cannot penetrate the glass melt, and the furnace bottom remains cold. Such glasses are often melted electrically. Forming is also affected to a lesser degree. J. Wenzel

Bibliography. N. P. Bansal and R. H. Doremus, *Handbook of Glass Properties*, 1996; C. A. Harper, *Handbook of Ceramics Glasses, and Diamonds*, 2001; G. McLellan and E. Shand (eds.), *Glass Engineering Handbook*, 1984; H. Rawson, *Glasses and Their Applications*, 1991; J. E. Shelby, *Introduction to Glass Science and Technology*, 2d ed., 2005; F. Tooley (ed.), *The Handbook of Glass Manufacture*, 2 vols., 1984; D. Uhlmann and N. Kreidl (eds.), *Optical Properties of Glass*, 1991; A. Varshneya, *Fundamentals of Inorganic Glasses*, 1994.

Glass switch

A glassy, solid-state device used to control the flow of electric current. Useful solid-state devices can be made from glassy as well as crystalline semiconductors. Crystals possess long-range order; that is, given the position of any particular atoms and the orientation of the neighboring atoms, the location of any other atom is known, no matter how far away from the atom under consideration. A glass is a special case of a noncrystalline class of materials, namely, amorphous solids. These do not exhibit long-range order, although they tend to have the same local structure (that is, short-range order) as the corresponding crystal. A glass is an amorphous solid that is formed by cooling rapidly from the liquid phase.

The first applications of glassy semiconductors were switches made from chalcogenide (that is, alloys containing tellurium, selenium, or sulfur) glasses. The two basic structures are known as the Ovonic Threshold Switch (OTS) and the Ovonic Memory Switch (OMS). They are active devices consisting simply of a thin film (about 10–100 nanometers thick) of glass between two metallic contacts. The device characteristics depend on the bulk properties of the semiconductor material rather than on the contacts. Consequently, the switches are symmetrical in that they respond identically to voltages and currents of both polarities. The OMS and OTS differ in terms of the composition of their amorphous semiconductor thin-film materials and their functional performance.

Amorphous semiconductors can be formulated so they can be doped by adding small amounts of impurities to change their electrical properties in the same way as crystalline semiconductors, or they can be designed to be insensitive to the effects of impurities. The glass switches typically use impurity-independent material compositions, and they are also highly resistant to the effects of radiation.

Device characteristics. Both the OTS and OMS show a rapid and reversible transition between a highly resistive and a conductive state effected by applied electric fields. The main difference between the two devices is that, after being brought from the highly resistive state to the conducting state, the OTS returns to its highly resistive state when the current falls below a holding current value. On the other hand, the OMS remains in the conducting state until a current pulse returns it to its highly resistive state. The OMS thereby remembers the last applied switching command, and it is from this property that the device receives its name.

Intermediate resistance states are also possible for OMS devices, which can be used in applications requiring a "gray scale." In all of these devices, the transitions between states are completely reversible.

The composition of the active material determines whether the device functions as an OTS or OMS, and also affects the values of certain device parameters. The device geometry, such as thickness and cross-sectional area of the active film, also affects the numerical values of the device parameters.

In the OTS (**Fig. 1**), conduction in the highly resistive state follows Ohm's law at fields below about 10^4 V/cm. At higher fields the dynamic resistance R_{dyn} decreases monotonically with increasing voltage. Typical values are R_{dyn} = 2–10 megohms at 1 V and R_{dyn} = 0.1–0.5 megohm just prior to breakdown.

When the applied voltage exceeds a threshold voltage V_T, the OTS switches along the load line to the conducting state. The transition time τ_t of this switching process has been shown to be less than 150 picoseconds. V_T is a function of both film thickness and active material composition and can be obtained in the range 2–300 V.

Current in the conducting state can be increased or decreased without significantly affecting the voltage drop across the device; the dynamic resistance is of the order of 1–10 ohms. Most of the voltage falls near the two contacts, due to barriers induced prior to switching; this accounts for about a 0.4-V drop. The field across the bulk is only about 1 kV/cm.

If the current is decreased below a critical value I_H, the OTS switches back to the original highly resistive state. I_H depends on circuit parameters and also can be varied; typical values are 0.1–1 mA.

The foregoing description of the static characteristics of the OTS holds for a slowly varying applied voltage. Upon application of a fast-rising pulse somewhat in excess of threshold voltage, the OTS ordinarily does not immediately switch to the conducting state but remains in the high-resistance state for a period of time τ_d, called the delay time. The magnitude of this delay is strongly dependent on the extent to which the threshold voltage is exceeded. For an applied voltage pulse slightly greater than V_T, τ_d can be several microseconds. However, it rapidly decreases with increasing voltage in excess of threshold, and essentially vanishes above a critical voltage that is proportional to film thickness. Above this point, the speed of switching is only circuit-limited, and total switching times less than 150 picoseconds have been observed.

Switching is an electronic effect, induced by the appearance of a critical electric field across a part or all of the film. The field induces a sharp transition which increases the free-carrier concentration by a factor of about 10^8. This results in a constant current density of approximately 10 kA/cm^2. At such current densities Joule heating effects are negligible.

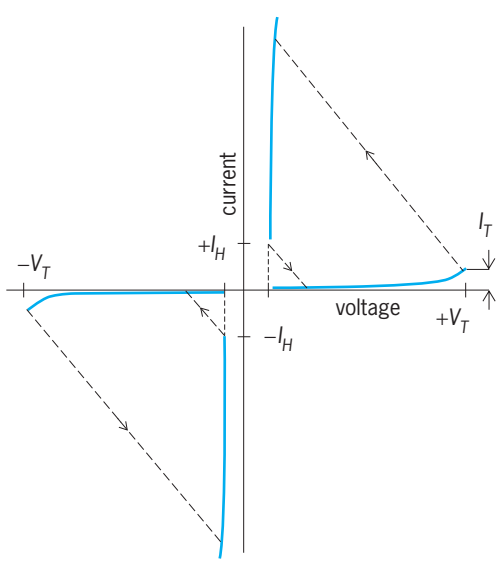

Fig. 1. OTS current-voltage characteristics.

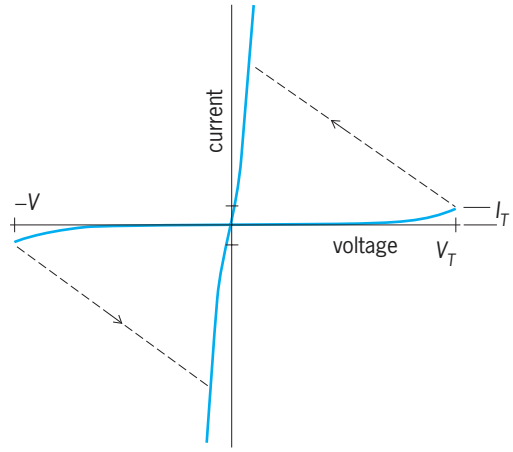

Fig. 2. OMS current-voltage characteristics.

The high conduction occurs primarily through a central filament whose area varies proportionally with the current.

In the OMS, the properties of the highly resistive state are essentially the same as for the OTS (**Fig. 2**). The OMS is switched to the conductive state when the threshold voltage is exceeded. After switching, however, the OMS is designed so that the amorphous semiconductor material within the conducting filament region can undergo a change in atomic structure. This change results in a change in electrical conductivity which remains after power is removed. The structural change is reversible by providing another pulse with a different current profile to the OMS so the device can be repeatedly programmed into a selected conductivity state.

Amorphous switch materials. Materials used in an OTS are specifically chosen to meet several important requirements. First, the film should have high electrical resistivity to avoid Joule heating effects. Second, the amorphous phase should be sufficiently stable to prevent crystallization or other structural changes. Third, the material should be chosen so that irreversible breakdown does not occur. All of these conditions can be fulfilled by the use of chalcogenide glasses. Chalcogenides are alloys one of whose major components comes from one of the elements in group 16 of the periodic table, ordinarily selenium or tellurium. These atoms ordinarily form polymeric-type chains in the solid state. Their electronic structure is such that they necessarily contain large but equal densities of positively and negatively charged traps in the amorphous phase. These traps keep the film resistivity high and retard irreversible breakdown. Structural stability is maintained by using alloys with relatively large densities or cross-linking atoms, ordinarily from groups 14 and 15 of the periodic table. *See* ELECTRICAL BREAKDOWN; TRAPS IN SOLIDS.

Memory material is chosen, in general, to contain relatively weaker bonds and a smaller density of cross-linking atoms, so that structural changes are more easily attained. A class of very rapidly crystallizing memory material alloys have been used to fabricate OMS devices that can be programmed at

voltages less than 3 V, with high speed (less than 1 nanosecond), extremely long cycle life (greater than 10^{13}), and the capability to be directly reprogrammed into a number of selected intermediate conductivity states. These materials can exist in the amorphous state, the fully crystalline state, and states of partial crystallinity, without compositional phase separation. Consequently, changes in the atomic structure of these materials can occur without atomic diffusion, allowing for rapid programming and very long reprogramming life.

Applications. Although Ovonic switches can be fabricated from bulk amorphous material, they are most conveniently produced as thin-film structures in which the active material and electrodes are vacuum-deposited or sputtered films, photolithographically defined. This economical process is compatible with transistor technology and with the methods used to produce passive components. It also lends itself to the fabrication of densely packed arrays.

Nonvolatile memories. Integrated arrays of OMSs can be used as electrically erasable programmable read-only memories (EEPROMs). Readout of these devices is extremely rapid, limited only by the readout circuit's characteristics; and, because of the capability of the OMS to be programmed into a number of selected conductivity states, more than one bit of data can be stored in each memory cell. This gives the OMS-based EEPROM an advantage over conventional EEPROM devices in terms of data-storage density, and consequently, cost per bit of stored information. Additionally, because of the high programming speed and very long cycle life of the device, an opportunity exists to develop unique semiconductor memory devices based on the OMS that can play the role in a computer of both nonvolatile archival memory and high-speed system memory. Availability of these devices can greatly simplify computer architecture and improve processing speed. *See* COMPUTER STORAGE TECHNOLOGY; SEMICONDUCTOR MEMORIES.

Neural networks. The OMS can be used as an electrically reconfigurable electrical interconnection in neural networks. In this application it provides a simple, high-density means to accomplish the many thousands or millions of programmable electrical interconnections required in practical neural-network circuits that can exhibit artificial intelligence. The thin-film nature of the OMS also allows for the design of multiple-layer, three-dimensional memory or neural-network circuits with data-storage or circuit-interconnect densities significantly greater than can be accomplished by using crystalline semiconductor technology. *See* NEURAL NETWORK.

Photographic films. Films have been developed in which the quality and amount of structural changes can be controlled by the amount of energy incident upon the film. These films are the only non-silver-based film with high resolution and amplification. They can also be used in updatable imaging products.

Transistors. A transistor, using an OTS as the emitter, has been developed. This can be used as a threshold amplifier, as a threshold latching amplifier, or as the basis for a computer using ternary logic. In the conducting state, OTSs can be used to provide a constant current density and to inject hot electrons into crystalline-semiconductor devices. Their threshold characteristics provide the opportunity for another important control function in these applications.

Other promising application areas include ac control where use is made of the inherent symmetry of these components, and microwave generation made possible by the inherently fast switching transition. It is anticipated that the exploration of new applications for amorphous switches will accelerate as knowledge of them becomes more widespread. *See* ALLOY STRUCTURES; AMORPHOUS SOLID; GLASS; SEMICONDUCTOR.

Stanford R. Ovshinsky; David Adler

Bibliography. D. Adler, *Sci. Amer.*, 236(5):36–48, May 1977; D. Adler et al., Threshold switching in chalcogenide-glass thin films, *J. Appl. Phys.*, 51:3289–3309, 1980; S. R. Ovshinsky, *Disordered Materials: Science and Technology*, 1991; S. R. Ovshinsky, Optically induced phase changes in amorphous materials, *J. Cryst. Solids*, 141:200–203, 1992.

Glass transition

The transition which occurs when a liquid is cooled to an amorphous or glassy solid. This can occur only if the cooling rate is fast enough to prevent crystallization which would otherwise occur if time had been sufficient for the sample to reach true equilibrium at each temperature. Since the crystal is invariably the thermodynamically stable low-temperature phase, the glass transition corresponds to a transition from a high-temperature liquid into a nonequilibrium metastable low-temperature solid. Experimentally, this transition occurs at a temperature T_g where the shear viscosity becomes immeasurably large, greater than 1×10^{15} poises (1×10^{14} pascal-seconds). Changes in the heat capacity and coefficient of thermal expansion are also generally observed near T_g. *See* AMORPHOUS SOLID; CRYSTAL STRUCTURE; VISCOSITY.

Basic phenomena. The liquid state of matter is characterized by zero shear modulus and a nonvanishing fluidity. (Fluidity is the reciprocal of viscosity.) When a liquid is slowly cooled, the system usually crystallizes through a discontinuous or first-order phase transition (that is, one with nonvanishing latent heat) at a melting temperature T_m. At T_m, the fluidity vanishes abruptly, and the system is in a new, stable-ordered, solid state, that is, a crystal. However, in many systems, it is possible to supercool the liquid to temperatures below T_m by rapid cooling. In those systems in which crystallization can be avoided, several interesting phenomena occur. The fluidity is observed to decrease continuously, reaching a limiting value so small that it is unobservable and the time required to relax a shear becomes very large, of the order of days or longer. When this has occurred, large-scale flow processes cease and the material appears solid on human time scales. This new state,

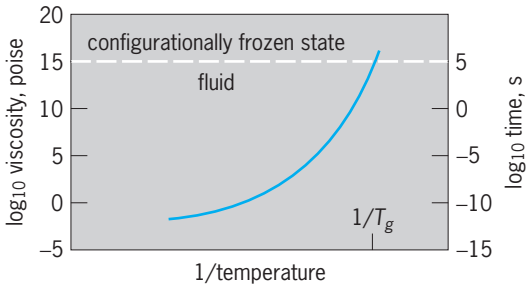

Fig. 1. Plot of the logarithms of viscosity and of the time constant for flow versus reciprocal temperature for a material cooled from the liquid to the glassy state. [$\log_{10}$ viscosity, Pa · s] = [$\log_{10}$ viscosity, poise] − 1.

which is referred to as a glass, differs from the crystal in that it is disordered and only metastable thermodynamically. Put simply, the glassy state is an extension of the liquid state, in which the viscosity (**Fig. 1**) increases above 1×10^{15} poises (1×10^{14} Pa · s). *See* PHASE TRANSITIONS.

Nonequilibrium character. All known systems are less stable in their glassy state than in some crystalline state. This is consistent with the observation that T_g is always lower (usually by a significant amount) than the melting temperature T_m. However, it is probably also true that the glass transition can occur in all liquids, provided crystallization can be bypassed. For many organic and polymeric systems, the difficulty of molecular packing and the steric hindrances are sufficient to prevent crystallization, and glass formation in these systems is relatively easy. In other systems, for example, metallic systems, rapid quench rates on the order of 1×10^6 K/s (2×10^{6}°F/s) may be necessary to avoid crystallization. For still other systems, particularly monatomic liquids with simple spherically symmetric interactions such as liquid argon, glass formation has never been observed in the laboratory. However, in computer simulations, where much faster cooling rates are achievable (1×10^{12} to 1×10^{14} K/s or 2×10^{12} to 2×10^{14}°F/s), such materials have formed glasses. This suggests that any system can be quenched from the liquid state to an amorphous glassy state assuming that the system can be cooled rapidly enough.

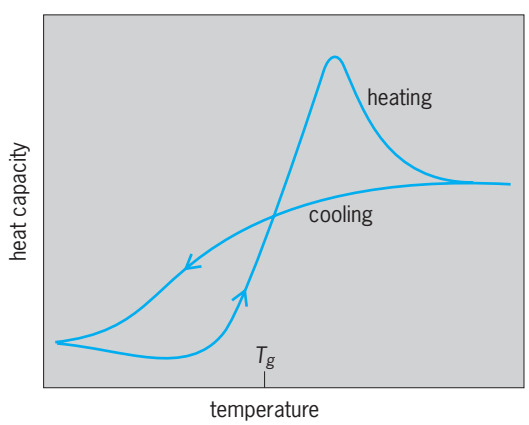

Fig. 2. Plot of heat capacity versus temperature for a system cooled and then reheated through the glass-transition region.

Variation of heat capacity. During the freezing of the supercooled liquid into a glass, the heat capacity and coefficient of thermal expansion change markedly within a narrow temperature range near T_g. This decrease in the heat capacity to a level near that of the crystalline system (**Fig. 2**) reflects a loss of configurational freedom which the sample undergoes at the glass transition. The magnitude of this drop in heat capacity in going from the equilibrium liquid to the glass is usually very large, approximately 50% for most glass formers, though it is barely detectable in the tetrahedrally coordinated network glasses silicon dioxide (SiO_2) and germanium dioxide (GeO_2). The exact value of T_g and the temperature dependence of the heat capacity depend on the heating and cooling rate of the measurement and on the thermal history of the sample. This behavior reflects that the observed changes in the heat capacity are not the result of a true equilibrium-phase transition but are a consequence of the system falling out of equilibrium when the time scale of the measurement becomes comparable to the relaxation time of the system. Below T_g, the relaxation rates are so slow that the system acts as a rigid solid, even though it is still flowing, although at a rate too slow to be measurable. *See* HEAT CAPACITY.

Related materials. The rapid increase in response times that characterize a dense supercooled liquid just above the glass transition has also been observed for a wide variety of disordered systems. These systems include a class of magnetic systems referred to as spin glasses, in which the interaction between the magnetic moments are "in conflict" with each other (for example, AuFe and CuMn); certain disordered insulators, in which highly localized electrons interact via the long-range Coulomb potential (the Coulomb glass); and orientational dipolar glasses (KBr:KCN), in which randomly located elastic-dipole defects interact via the strain field. These systems differ from ordinary liquids in that they do not have an ordered low-temperature phase, such as a crystal. Instead, the most stable state is disordered. However, as the temperature is lowered, these systems all reach a characteristic temperature in which their response becomes so slow that they essentially freeze, unable to reach their lowest free-energy state, very similar to the glass transition in ordinary liquids. *See* GLASS; METALLIC GLASSES; SPIN GLASS. Gary S. Grest

Bibliography. S. R. Elliott (ed.), *Physics of Amorphous Materials*, 2d ed., 1990; J. P. Hansen, D. Levesque, and J. Zinn-Justin (eds.), *Liquids, Freezing and the Glass Transition*, 1990.

Glaucoma

A disease of the eye in which damage is caused by elevated pressure within the eye. The incidence in persons over the age of 40 is about 0.5%, making glaucoma one of the most common and serious eye disorders, surpassed in the United States only by cataracts

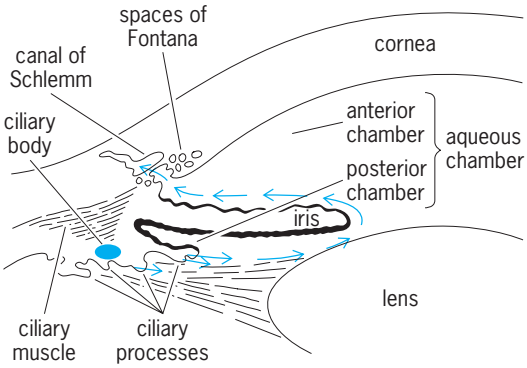

Pathway of circulation of aqueous humor in the eye.

as a cause of blindness. *See* CATARACT; VISUAL IMPAIRMENT.

The normal pressure within the eye measures 10–20 mmHg (1.3–2.7 kilopascals) and is maintained by a delicate balance between the inflow and outflow of fluid called aqueous humor. The aqueous humor is produced by the ciliary body and passes from the posterior chamber, that is, the area behind the iris, through the pupillary space into the anterior chamber, that is, the space in front of the iris (see **illus.**). It then drains through the trabecular meshwork into the canal of Schlemm, which leads to the exterior of the eye. *See* EYE (VERTEBRATE).

Classification. Four major types of glaucoma have been identified: (1) congenital or infantile glaucoma, which is evident at birth or shortly thereafter; (2) primary open-angle glaucoma, the most common kind, which is usually painless and is marked by a blockage in the outflow channels; (3) primary (acute) angle-closure glaucoma, in which the root of the iris interrupts drainage and causes a sudden painful blockage and acute rise in internal pressure; and (4) subacute or chronic angle-closure glaucoma, in which the root or base of the iris falls across the drain temporarily but repeatedly, resulting in transient increases in pressure but with scarring after each episode until the drain can no longer become unblocked. Secondary glaucoma results from inflammation, injury, surgery, or eye diseases such as swollen cataract.

Symptoms and effects. Infantile glaucoma often enlarges the eye because of a lack of rigidity in its white coat, known as the sclera; the condition is called buphthalmos. Angle-closure glaucoma is marked by pain, sudden visual loss, and a steamy or cloudy cornea; the semidilated and fixed pupil does not respond to changes in light intensity. In chronic angle-closure glaucoma, the periodic rises in pressure are accompanied by the above symptoms and also by halo vision, that is, the person sees haloes around a light source. In this type of glaucoma the pressure rises and falls as the angle reopens, and adhesions of scar tissue, called synechiae, form. The attacks continue until the drainage channel is sealed.

In the most common type of glaucoma, the asymptomatic open-angle form, the chief threat is a gradual, imperceptible loss of vision. The disease is bilateral, with progressive field loss mostly in the periphery where it escapes notice. For people over the age of 30 or 40, intraocular pressure should be measured every 2 or 3 years. Open-angle glaucoma is a heritable condition, but no racial or sexual patterns have been found.

Elevated intraocular pressure without apparent damage is a condition known as ocular hypertension. It should be considered a type of glaucoma. Frequent observation of the optic disk is recommended, with treatment as indicated.

Tests. A number of diagnostic tests are known. Eye pressure is measured by a tonometer, an instrument with weights and a scale. Applanation tonometry, which uses a slit lamp or special microscope, is more accurate. Gonioscopy, performed in conjunction with the slit lamp, is an examination of the drain or angle for susceptibility to sudden closure or gradual obstruction. Tonography provides a tracing, called a tonogram, which can be interpreted much like a cardiogram. It is usually obtained electronically by means of a special device attached to the eye to record the eye pressure over a 4-min period. These measurements make it possible to calculate the rate of elimination of aqueous humor from the eye. A low rate may be first sign of open-angle glaucoma.

So-called provocative tests are performed in open-angle or closed-angle glaucoma to elicit a pressure rise; they can help to provide a diagnosis when eye pressures are normal under ordinary conditions. The visual field is very important in diagnosing and following chronic glaucoma. It can be plotted and measured in various ways, from the simple tangent screen to the more sophisticated Goldmann perimeter.

Treatment. Blindness can usually be prevented by early treatment to maintain normal eye pressure. Eye drops or oral medication is usually effective, reducing the flow of aqueous humor into the eye and increasing its outflow. If medication fails, as in angle-closure glaucoma, or if sudden complete blockage occurs, surgery is indicated. Eye pressure is then lowered by creating an artificial outlet, usually with the argon or yttrium aluminum garnet (YAG) laser. The surgical procedure called peripheral iridectomy has largely been superseded by laser treatment. Congenital glaucoma, which occurs as a defect in development of the angle, must be treated surgically, however.

Open-angle glaucoma, or chronic painless glaucoma, can also be treated by trabeculoplasty with the argon laser, which is often successful. If it fails, a new outlet can be opened by a surgical procedure known as trabeculectomy. Lasers are preferred, if only because they carry less risk of infection than conventional surgery. *See* LASER. Jack Hartstein

Bibliography. G. K. Krieglstein, *Glaucoma*, 1993; M. B. Shields, *Textbook of Glaucoma*, 4th ed., 1997.

Glauconite

The term glauconite as currently used has a twofold meaning. It is used as both a mineralogic and morphologic term. The mineral glauconite is defined

as an illite type of clay mineral. It is dioctahedral with considerable replacement of aluminum by iron and magnesium. Structural substitutions result in a charge deficiency in both the tetrahedral and octahedral sheets, and interlayer cations seem to balance both of these charges. Calcium and sodium as well as potassium are the interlayer cations. A fundamental characteristic of glauconite is that the unit cell is composed of a single silicate layer rather than the double layer of most other dioctahedral micas. *See* CLAY MINERALS; ILLITE.

Glauconite is known to occur in flakes and as pigmentary materials. When used in the morphological sense, the term glauconite often refers to small, green, spherical, earthy pellets. Some of these pelletal varieties are composed solely of the mineral described above, others are a mixed-layer association of this mineral and other three-layer structures.

The mineral glauconite is formed during marine diagenesis. Because of the frequent association of glauconite with organic residues, it has been generally concluded that the presence of organic material is necessary for the formation of this mineral. Glauconite forms in a marine environment from a variety of raw materials. It is known to form as an alteration product of biotite mica, when the alteration takes place under reducing conditions. The deposition of sediment must be slow during this mineralogic transformation to allow complete alteration before burial. If burial is accomplished before alteration, the biotite mica persists. *See* AUTHIGENIC MINERALS; DIAGENESIS.

Relatively shallow water is another requisite for the formation of glauconite. It has been shown that glauconite forms in the shallow sea in agitated waters which are not highly oxygenated. A reducing environment is also probably necessary for the formation of glauconite.

The magnesium content of glauconite is very uniform, and the ratio of ferric to ferrous iron is rather constant. This suggests that a critical content of magnesium and a particular oxidation-reduction potential might be required for this mineral to form. It has been suggested that certain concentrations of sodium and potassium ions are also necessary for glauconite formation, since the ratio of these cations in the interlayer positions is rather distinctive.

Glauconite readily occurs in pellet form and has been identified in both recent and ancient sediments. It is a major component in some "greensand" deposits and has been used commercially for the extraction of potassium from such sources. *See* MARINE SEDIMENTS.

Floyd M. Wahl; Ralph E. Grim

Glaucophane

A monoclinic sodic amphibole with composition close to $Na_2(Mg_3Al_2)Si_8O_{22}(OH)_2$. This mineral exhibits a characteristic blue color with distinct pleochroism from colorless to lavender blue when viewed in thin section by plane-polarized light. Outcrops of glaucophane-rich metamorphic rocks are commonly blue and tend to have good foliation; these rocks are termed blueschists. *See* AMPHIBOLE; BLUESCHIST; PLEOCHROISM.

Composition. Classification of amphiboles by the International Mineralogical Association is based on the standard amphibole formula $AB_2{}^{VI}C_5{}^{IV}T_8O_{22}(OH)_2$. Sodic amphiboles are defined to include monoclinic amphiboles in which $Na_B > 1.50$. Analyses of naturally occurring glaucophanes show complex chemical substitution of octahedrally coordinated magnesium (Mg) by ferrous iron (Fe^{2+}) and of aluminum (Al) by ferric iron (Fe^{3+}). Glaucophanes have $Fe^{2+}/(Fe^{2+} + Mg)$ ratios less than 0.5 and ${}^{VI}Al > {}^{VI}Fe^{3+}$, whereas magnesioriebeckites have ${}^{VI}Al < {}^{VI}Fe^{3+}$. Sodic amphiboles with higher $Fe^{2+}/(Fe^{2+} + Mg)$ ratios include ferroglaucophanes (${}^{VI}Al > {}^{VI}Fe^{3+}$) and riebeckites (${}^{VI}Al < {}^{VI}Fe^{3+}$). Glaucophanes may also contain minor calcium (Ca) substitution of sodium (Na), with the Ca/(Na + Ca) ratio less than 0.25; this ratio is between 0.25 and 0.75 for winchite-ferrowinchite, and greater than 0.75 for tremolite-actinolite. These amphiboles have a vacant A site. Substitution of $Na_A + {}^{IV}Al = {}^{IV}Si$ results in nyböite and its ferric and ferro equivalents. Nyböites, similar to other sodic amphiboles, have been reported in subduction-zone metamorphic rocks such as the Nyböeclogite from Norway and the Donghai eclogite from eastern China. The increasing use of electron microprobe analyses of natural glaucophanes during the last three decades shows that sodic amphiboles exhibit complex compositional variations even in a single crystal. The difference in various substitutions of sodic amphibole depends on the pressure-temperature conditions and rock compositions. *See* ECLOGITE.

Stability. Stability relations of glaucophane and other sodic amphiboles have been extensively investigated. Experimental studies of synthetic iron-free glaucophane suggest a very high pressure stability field (>12 kilobars or 1200 megapascals), but natural glaucophane containing 9 wt % total iron as ferrous oxide (FeO) has been experimentally shown to decompose only at pressures below 4 kbar (400 MPa) and temperatures above 550°C. Many workers studying various blueschist terranes have noted that the ferric oxide (Fe_2O_3) content of glaucophane significantly extends its stability toward lower pressures.

End-member glaucophane has not been conclusively synthesized in the laboratory, and the pressure-temperature stability fields of glaucophane-magnesioriebeckite assemblages are highly dependent on bulk rock composition (hence, the mineral assemblage). Despite problems in laboratory studies and differences in experimental results, the available data suggest that formation of glaucophane requires pressures higher than 4 kbar (400 MPa; greater than about 7.5 mi or 12 km depth) and temperatures lower than 800°C (see **illus.**). This conclusion is consistent with investigations of natural occurrences in many blueschist and eclogite terranes within the Eurasian continent, Central America, and the circum-Pacific and Himalayan-Alpine orogenic belts. Within

these regions, glaucophane occurs together with lawsonite, jadeitic pyroxene, and, with increasing temperature, epidote and garnet in addition to chlorite, white mica, and stilpnomelane. These assemblages define the glaucophane-schist or blueschist facies of metamorphism. The aluminum oxide (Al_2O_3) content of the glaucophane-magnesioriebeckite assemblage coexisting with the epidote + actinolite + chlorite + albite + quartz assemblage decreases systematically with decreasing pressure, and has been used to estimate pressure (hence, depth) for blueschist formation in California, Japan, New Caledonia, and New Zealand.

Both experimental data and thermodynamic calculations show that glaucophane of the end-member composition is stable over a wide range of pressure-temperature conditions from 7 to 31 kbar at 400°C and 11 to 32 kbar at 600°C (see illus.). Reaction of glaucophane to form jadeite + talc defines the upper pressure limit of glaucophane; the equilibrium line was calculated at 30.5 kbar and 600°C with a gentle positive pressure-temperature slope. Similarly, the calculated maximum pressure of the reaction glaucophane + kyanite = jadeite + garnet + talc lies at 33 kbar at about 640°C and 30 kbar at about 750°C. Nyböite, like glaucophane, has not been synthesized from its own or other composition, and amphiboles with about 70 mol % nyböite are stable to pressures greater than 32 kbar at 600–900°C, but break down to albite + nepheline + sodian phlogopite at 15 kbar and 600–900°C. Apparently, both glaucophane and nyböite are stable within the stability field of coesite, a high-pressure polymorph of quartz. These sodic amphiboles have been identified in coesite-bearing ultrahigh-pressure metamorphic rocks from the Dora-Maira massif of the Italian Western Alps and the Dabie Mountains of central China, with the estimated minimum pressures of 27–29 kbar, equivalent to mantle depths of nearly 80–90 km. *See* MASSIF.

Tectonic significance. Glaucophane is an index mineral of blueschist, which is generated under unusually high pressures at low temperatures in a tectonic environment exclusively associated with a subducted lithospheric slab or related tectonic loading. The glaucophane-bearing assemblages occur in recrystallized graywackes and pelitic rocks and in metabasites and metacherts of oceanic affinity; they are typically found in subduction zone complexes at plate boundaries, a setting first recognized in the Jurassic and Cretaceous Franciscan Complex of northern California. Blueschists are most common and best developed in Mesozoic and Cenozoic terranes: some Paleozoic and even latest Precambrian blueschists have been described in Russia and China. Blueschists formed earlier in geologic time may have been eroded or been recrystallized under normal geothermal conditions. The preservation of glaucophane in blueschists of continental or island arc margins indicates either rapid uplift or maintenance of low geothermal gradients by steady-state subduction for tens of million years. *See* GRAYWACKE; SUBDUCTION ZONES.

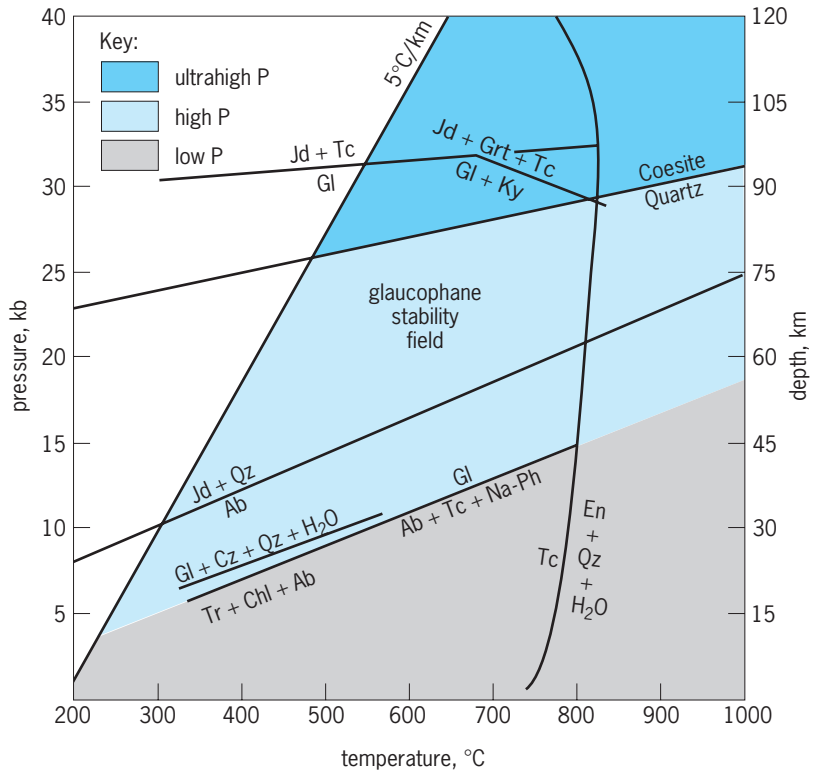

Pressure-temperature diagram showing the stabilities of glaucophane and glaucophane-bearing assemblages. Tc, talc; Jd, jadeite; Gl, glaucophane; Cz, clinozoisite; Qz, quartz; Tr, tremolite; Chl, chlorite; Grt, garnet; En, enstatite; Na-ph, sodium phologopite; Ab, albite; H_2O, water. 1 kilobar = 10^2 MPa, 1 km = 0.6 mi. The geothermal gradient with increasing temperature at 5°C per kilometer depth is indicated. (*Modified after J. H. Carman and M. C. Gilbert, 1983; S. Maruyama and others, 1986; R. Y. Zhang and J. G. Liou, 1994*)

Porphyroblastic sodic amphibole is also common as a primary phase in eclogites of many continent collisional orogens, including those in the western Alps, central China, and southern Urals. However, the development of glaucophane in some eclogites, including those tectonic blocks of the Franciscan mélange in California, has been explained as the beginning of a retrograde evolution from the stable conditions of the primary eclogite paragenesis. This reaction is accompanied by a variation in the distribution of iron and magnesium between the amphibole, garnet, and omphacite. In some pre-Alpine high-grade metamorphic rocks, amphiboles of the glaucophane-magnesioriebeckite series developed in the pre-eclogite and eclogite stages of the early Alpine orogeny; they persisted into the post-eclogite stage, before being replaced by blue-green Na-Ca amphibole during the post-eclogitic and late Alpine stages. Compositions of different-stage glaucophane crystals vary even in a single rock and reflect different prograde and retrograde pressure-temperature conditions. *See* OROGENY; PORPHYROBLAST.

J. G. Liou; R. Y. Zhang; S. Maruyama

Bibliography. J. H. Carman and M. C. Gilbert, Experimental studies on glaucophane stability, *Amer. J. Sci.*, 283:A414–A437, 1983; W. A. Deer, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, vol. 28: *Double-Chain Silicates*, 2d ed., 1997; B. E. Leake et al., Nomenclature of amphibole: Report of the

Subcommittee on Amphiboles of the International Mineralogical Association, Commission on New Minerals and Mineral Names, *Amer. Mineral.*, 82:1019–1037, 1997; J. G. Liou and S. Maruyama, Parageneses and compositions of amphiboles from Franciscan jadeite-glaucophane type facies series metabasites at Cazadero, California, *J. Metamor. Geol.*, 5:371–395, 1987; S. Maruyama, M. Cho, and J. G. Liou, Experimental investigations of blueschist-greenschist transition equilibria: Pressure dependence of Al_2O_3 contents in sodic amphiboles—a new geobarometer, *Geol. Soc. Amer. Mem.*, no. 164, pp.1–16, 1986; S. Maruyama, J. G. Liou, and M. Terabayashi, Blueschists/eclogites of the world and their exhumation, *Int. Geol. Rev.*, 38:490–596, 1996; R. Y. Zhang and J. G. Liou, Coesite-bearing eclogite in Hanah Province, central China: Detailed petrology, glaucophane stability and PT path, *Eur. J. Mineral.*, 6:217–233, 1994.

Glazing

The application of finely ground glass, or glass-forming materials, or a mixture of both, to a ceramic body and heating (firing) to a temperature where the material or materials melt, forming a coating of glass on the surface of the ware. Glazes are used to decorate the ware, to protect against moisture absorption, to give an easily cleaned sanitary surface, and to hide a poor body color.

Glazes are classified and described by the following characteristics: surface—glossy or matte; optical properties—transparent or opaque; method of preparation—fritted or raw; composition—such as lead, tin, or boron; maturing temperatures; and color. Opaque glazes contain small crystals embedded in the glass, but special glazes in which a few crystals grow to recognizable size are called crystalline glazes. A glaze may be applied during the firing; such a glaze is called salt glaze. Common salt, NaCl, or borax, $Na_2B_4O_7 \cdot 10H_2O$, or a mixture of both is introduced into the kiln at the finishing temperature. The salt evaporates and reacts with the hot ware to form the glaze. This type of glaze has been applied to sewer pipe and some fine stoneware.

The most important factor in compounding a glaze, after a suitable maturing temperature has been obtained, is the matching of the coefficient of thermal expansion of the glaze and the body on which it is applied. A slightly lower coefficient for the glaze will place it in compression (the desired condition) when the ware cools. The reverse state (with the glaze in tension because it has a higher coefficient) leads to the formation of fine hairline cracks, a condition known as crazing. *See* CERAMICS; GLASS; METAL COATINGS. J. F. McMahon

Bibliography. E. Cooper, *The Potter's Book of Glaze Recipes*, rev. ed., 2004; G. Daly, *Glazes and Glazing Techniques*, 1995; R. Hopper, *The Ceramic Spectrum: A Simplified Approach to Glaze and Color Development*, 2d ed., 2001.

Glide-path indicator

An aircraft landing instrument that provides the pilot with a set of vertical and horizontal cross pointers that indicate deviation from a radio-transmitted course to the threshold of the runway. A dual-frequency transmitter sends out one frequency to the right of the runway centerline and a second frequency to the left of the runway centerline. The reception of these signals in the aircraft biases the vertical needle to the left or right depending on the position of the aircraft relative to the transmitted 5° localizer path. Simultaneously, another dual-frequency transmitter causes a horizontal needle to indicate high or low as the aircraft descent path is compared to the transmitted 3° glide path. This indicator, when properly used with the other navigation instruments, allows approaches to be made to within 200 ft (60 m) of altitude and 0.5 mi (0.8 km) of the runway threshold, at which time the approach transfers to a visual type, or is aborted if no visual contact is made. *See* INSTRUMENT LANDING SYSTEM (ILS). James W. Angus

Bibliography. D. A. Lombardo, *Advanced Aircraft Systems*, 1993; B. L. Stevens and F. L. Lewis, *Aircraft Control and Simulation*, 1992.

Glider

An unpowered flying device that attempts to copy the flight of soaring birds as accurately as possible.

Early development. Otto Lilienthal made hundreds of flights with several designs of hang gliders before his fatal crash in 1896. His gliders were of rigid construction, generally without movable control surfaces, and were controlled by shifting the pilot's weight. These gliders had no landing gear other than the pilot's legs.

The next major advance in glider design was made by Wilbur and Orville Wright, who were inspired by Lilienthal. The Wrights' 1902 glider in its final version made hundreds of perfectly controlled glides and set a distance record of 622 ft (189 m) and a duration record of 26 s on the dunes of Kitty Hawk, North Carolina. This glider did not depend on weight shift for control, but had aerodynamic controls consisting of movable elevator, rudder, and wing tips. This system, in principle, has been used to the present time, even on very large and fast gliders and powered aircraft. The success with this glider led the Wrights to construct a slightly larger aircraft, with engine and propellers, that enabled them to take off and fly from level ground for the first time on December 17, 1903. The first successful powered vehicle, now called an airplane, was not a practical or useful aircraft, but with engineering development did reach that goal in a few years. *See* AIRPLANE; FLIGHT CONTROLS.

Methods of flight. In October 1911, Orville Wright made a gliding flight of nearly 10 min duration, and demonstrated that gliders could stay up for long periods in the rising air caused by the wind blowing against a sand dune or hill. This condition of flight,

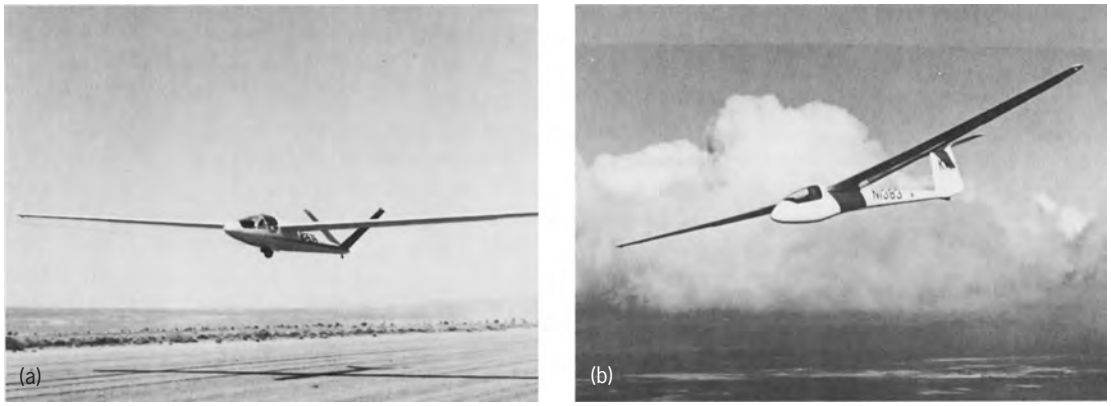

Fig. 1. Sailplanes. (*a*) In tow during launch. (*b*) In soaring flight. (*Photographs by George Uveges*)

called slope soaring, was the basic method of soaring flight until about 1930. Thermal soaring, the next step, was accomplished by flying in areas of rising convection currents, which are almost always present to some degree in the atmosphere. The development of thermal soaring and its practice is now principally due to the use of the variometer, calibrated as a sensitive rate-of-climb instrument, which enables the pilot to find the thermal and make the best use of it. By the use of thermal flight, the modern glider can fly almost anywhere in the world for extended time and distances over 500 mi (800 km) in one flight. Other methods of soaring make use of clouds and standing-wave phenomena in the atmosphere. These techniques have enabled gliders to achieve altitudes of 46,000 ft (14,000 m) under the proper conditions. High-performance gliders (sailplanes; **Fig. 1**) may be launched by towing behind powered aircraft to a release height of about 2000 ft (700 m), by winch-launching to about 800 ft (250 m), or by car towing, which is used to a lesser extent. Some gliders have been fitted with a small motor and propeller, which enables them to take off and climb to an altitude where rising air permits them to soar unpowered.

Francis M. Rogallo

Sailplanes. Sailplane construction traditionally has been of wood and plywood, although the use of aluminum alloy has become common. The greater strength and stiffness of modern aluminum alloy permits higher aspect ratios and improved performance. The use of fiber glass as primary structure has also come into prominence, since it is possible to produce the external shapes in accurate molds with greater precision, resulting in improved performance.

Sailplanes are equipped with dive brakes on the wings for emergency descent and for landing in small areas. Properly designed brakes hold the aircraft to its "never-exceed" dive speed in a vertical dive. The laminar-flow airfoil achieves its lowest drag in a certain range of lift coefficient and corresponding speed range. By the use of properly designed flaps, this range can be shifted to higher or lower speeds at the pilot's control, which provides improved performance at a wider range of speeds. *See* AIRFOIL.

As sailplane performance increases, each new drag item becomes important, and there has been a tendency to use reclining, and almost full reclining, positions for the pilots. This permits fuselages with overall heights as low as 30 in. (75 cm) but introduces control system and visibility problems. *See* FUSELAGE.

Ernest Schweizer

Flexible-wing gliders. In the late 1950s the National Aeronautics and Space Administration (NASA) investigated various methods of returning crewed spacecraft to Earth. Two kinds of glider were investigated: aircraft with rigid delta or lifting body shapes and very high landing speeds, which later evolved into the space shuttle; and flexible-wing craft which could be packed like parachutes and deployed for slow, controlled landings in almost any open field, a concept that had been proposed 10 years earlier. After NASA demonstrated flexible-wing capability, many segments of the Department of Defense and industry also became interested in flexible-wing gliders for a variety of applications. Crewed and radio-controlled flights were made with flexible-wing gliders with or without power or towed by cars or aircraft. Some gliders were completely flexible, and some were stiffened with springy battens, aluminum tubes, or fabric tubes either pressurized or ram-inflated. Although extensive military and space applications are still undeveloped, flexible wings have had an effect on sport flying devices such as kites, hang gliders, and deployable gliders or gliding canopies used by sky divers. The completely flexible, deployable gliders have maximum glide ratios of only 3 to 4, which are very low compared to those of sailplanes, but a substantial change from the zero glide of the traditional parachute or the glide ratio of less than unity obtained from modified parachutes. Flexible, deployable gliders are used by sky divers who have demonstrated great ability to maneuver, penetrate the wind, and land on a chosen spot.

Modern foot-launched hang gliders with aluminum tube frames are the result of many improvements by private individuals and small manufacturers. Hang gliders have flown over 100 mi (160 km) in straight-line distance, have reached about 20,000 ft (6000 m) altitude, and have remained aloft more than 15 h. But it is not so much their performance that makes hang gliders popular, as their low cost, their

Fig. 2. Hang glider in flight. (*Photography by J. Foster Scott*)

convenience of folding into a small package for transport or storage, and the fact that no license is required for glider or pilot. Hundreds of thousands of people have learned to fly hang gliders (**Fig. 2**).

Propellers and motors of about 10 hp (7.5 kW) have been attached to some hang gliders, enabling them to take off from level ground and climb in still air if necessary, while still able to soar in updrafts with the engine off. Some hang gliders have wheels that can be used optionally. In the summer of 1979 five powered hang gliders took off from California and, after many stops for rest and refueling, landed on the east coast. *See* FLIGHT. Francis M. Rogallo

Bibliography. D. Mondey (ed.), *International Encyclopedia of Aviation*, 1977.

Global climate change

The periodic fluctuations in global temperatures and precipitation, such as the glacial (cold) and interglacial (warm) cycles of the Pleistocene (a geological period from 1.8 million to 10,000 years ago). Presently, the increase in global temperatures since 1900 is of great interest. Many atmospheric scientists and meteorologists believe it is linked to human-produced carbon dioxide (CO_2) in the atmosphere.

Greenhouse effect. The greenhouse effect is a process by which certain gases (water vapor, carbon dioxide, methane, nitrous oxide) trap heat within the Earth's atmosphere and thereby produce warmer air temperatures. These gases act like the glass of a greenhouse: they allow short (ultraviolet; UV) energy waves from the Sun to penetrate into the atmosphere, but prevent the escape of long (infrared) energy waves that are emitted from the Earth's surface. *See* ATMOSPHERE; GREENHOUSE EFFECT.

Global warming. Human-induced changes in global climate caused by release of greenhouse gases into the atmosphere, largely from the burning of fossil fuels, have been correlated with global warming. Since 1900, the amount of two main greenhouse gases (carbon dioxide and methane) in the Earth's atmosphere has increased by 25%. Over the same period, mean global temperatures have increased by about 0.5°C (0.9°F).

Of all the greenhouse gases produced by humans, the most concern centers on carbon dioxide. Not only is carbon dioxide produced in much greater quantities than any other pollutant, but it remains stable in the atmosphere for over 100 years. Methane, produced in the low-oxygen conditions of rice fields and as a by-product of coal mining and natural gas use, is 100 times stronger than carbon dioxide in its greenhouse effects. Methane, however, is broken down within 10 years.

Chlorofluorocarbon (CFC) pollution, from aerosol propellants and coolant systems, affects the Earth's climate because CFCs act as greenhouse gases and they break down the protective ozone (O_3) layer. The ozone layer normally prevents most UV radiation originating in outer space from entering the Earth's atmosphere. A thinner ozone layer allows more UV radiation to penetrate the Earth's atmosphere. CFC production has declined since the late 1970s, with further reductions imminent, but the benefit of this decrease will be realized only slowly because CFCs remain in the atmosphere up to 100 years.

Other pollutants released into the atmosphere are also likely to influence global climate. Sulfur dioxide (SO_2) from car exhaust and industrial processes, such as electrical generation from coal, cool the Earth's surface air temperatures and counteract the effect of greenhouse gases. Nevertheless, there have been attempts in industrialized nations to reduce sulfur dioxide pollution because it also causes acid rain. Since sulfur dioxide remains in the atmosphere for only a week, reduction of sulfur dioxide emissions will immediately lessen its impact on global climate. *See* AIR POLLUTION; OZONE.

Predictions. A rise in mean global temperatures is expected to cause changes in global air and ocean circulation patterns, which in turn will alter climates in different regions. While many regions have already warmed [the United States and western Europe had temperature elevations of about 0.4°C (0.7°F) during the twentieth century], some areas may experience cooling trends. Changes in precipitation have already been detected. In the United States, total precipitation has increased, but it is being delivered in fewer, more extreme events, making floods (and possibly droughts) more likely. *See* OCEAN CIRCULATION.

Range shifts. One impact of global warming on wildlife has been changes in the distribution of a species throughout the world. By analyzing preserved remains of plants, insects, mammals, and other organisms which were deposited during the most recent glacial and interglacial cycles, scientists have been able to track where different species lived at times when global temperatures were either much warmer or much cooler than today's climate. The range of most species was several hundred kilometers closer to the Poles or several hundred meters higher in elevation during times when the Earth was 4–5°C (7–9°F) warmer than it is today. Likewise, during glacial periods, species lived closer to the Equator and at lower elevations than they do now. Several studies have documented poleward and upward shifts of many plant and insect species during the

current warming trend. In the western United States, the Edith's checkerspot butterfly lives, on average, 92 km (57 mi) farther north and 124 m (407 ft) farther up the mountains than it did in the early part of the twentieth century. During the same time period, many species of mountain plants have shifted to higher elevations in the Swiss Alps at rates up to 4 m (13 ft) per decade. The magnitude of these shifts in species' ranges northward and upward parallels the magnitude of warming that those regions have experienced. Climate is presumed to have been the driving force for the changes.

Phenological shifts. Changes in the timing of growth and breeding events in the life of an individual organism, called phenological shifts, have resulted from global warming. The beginning of spring is determined by length of the day and by local climatic conditions. People have long been interested in the events that mark the beginning of spring, such as blooming of the first spring flower, leafing out of trees, and nest building by birds. There have been changes in the timing of these events over the last few decades. In the Northern Hemisphere, trees such as oak, birch, and maple are leafing up to 20 days earlier. Almost one-third of British birds are nesting earlier (by 9 days) than they did 25 years ago; the other two-thirds have not changed. Five out of six species of British frog are laying eggs 2–3 weeks earlier.

Community reassembly. Community reassembly refers to changes in the species composition of communities. Communities are assemblages of interacting species living in the same area. Not all species have the same response time to environmental change For example, the Edith's checkerspot butterfly has moved almost exactly as much as predicted by the climatic change (see **illus.**). Alpine plants have lagged behind, moving at a rate of only 4 m (13 ft) per decade, when an immediate response to the warming trend should have resulted in shifts of 8–10 m (26–33 ft) per decade.

Further, not all species in a community will be equally limited by climate. Some will be more sensitive to small changes than others. Some will be primarily restricted by nonclimatic factors, such as soil type, or competitive displacement (the inability to successfully live in an area because a second species dominates local resources).

Euphydryas editha **(female), high elevation (10,000 ft or 3050 m) in the Sierra Nevada.**

These differences in response to large climatic changes can be seen in the fossil record. Because not all species were moving at the same rate, communities were broken apart, and sometimes new communities with no modern-day counterpart formed (nonanalog communities).

Extinctions. Extinction is the end of the existence of a species, but the term is also applied to the loss of a distinct subspecies or species within a given geographic area. To data, there have been no extinctions of species directly attributable to climate change. However, there is mounting evidence for drastic regional declines. The abundance of zooplankton (microscopic animals and immature stages of many species) has declined by 80% off the California coast. This decline has been related to gradual warming of sea surface temperatures. Zooplankton are a major food source for oceanic wildlife, and the decline of this food supply has been harmful to many birds, fish, and mammals. The sooty shearwater has declined by 90% since 1987. Populations of Cassin's auklet and rockfish have also decreased. A puzzling observation is that these species are not simply moving northward to colder waters where the zooplankton supply is still healthy, but seem unable to alter their behavior to respond to the changed environment. This very rigid, sedentary lifestyle does not bode well for their long-term survival, and may even be driving such species toward extinction. *See* CLIMATE HISTORY; CLIMATE MODIFICATION; CLIMATIC PREDICTION; EXTINCTION (BIOLOGY). Camille Parmesan

Bibliography. Intergovernmental Panel on Climate Change, *Climate Change 1995: 2d Assessment Report, Working Groups I and II*, Cambridge University Press, 1996; Intergovernmental Panel on Climate Change, *The Regional Impacts of Climate Change*, Cambridge University Press, 1998; P. M. Kareiva, J. G. Kingsolver, and R. B. Huey (eds.), *Biotic Interaction and Global Change*, Sinauer Associates, 1993; R. L. Peters and T. E. Lovejoy (eds.), *Global Warming and Biological Diversity*, Yale University Press, 1992.

Globe (Earth)

A sphere with a map of the world on its surface, the traditional model of the Earth. The use of a round object whose surface is equidistant at every point from the center ignores the Earth's true shape. Satellite measurements from space show that the Earth is slightly flattened at the Poles and bulges in the Southern Hemisphere. However, when scaled down to the size of a globe, these irregularities disappear. Compared with flat maps, all of which distort the Earth's surface patterns, the graphical representation produced by a spherical terrestrial globe eliminates serious inaccuracies and distortions of the relative size, direction, and shape of areas on Earth.

The world map may be drawn, engraved, or painted directly on the globe's surface, but it is more commonly a series of gores, or map segments in

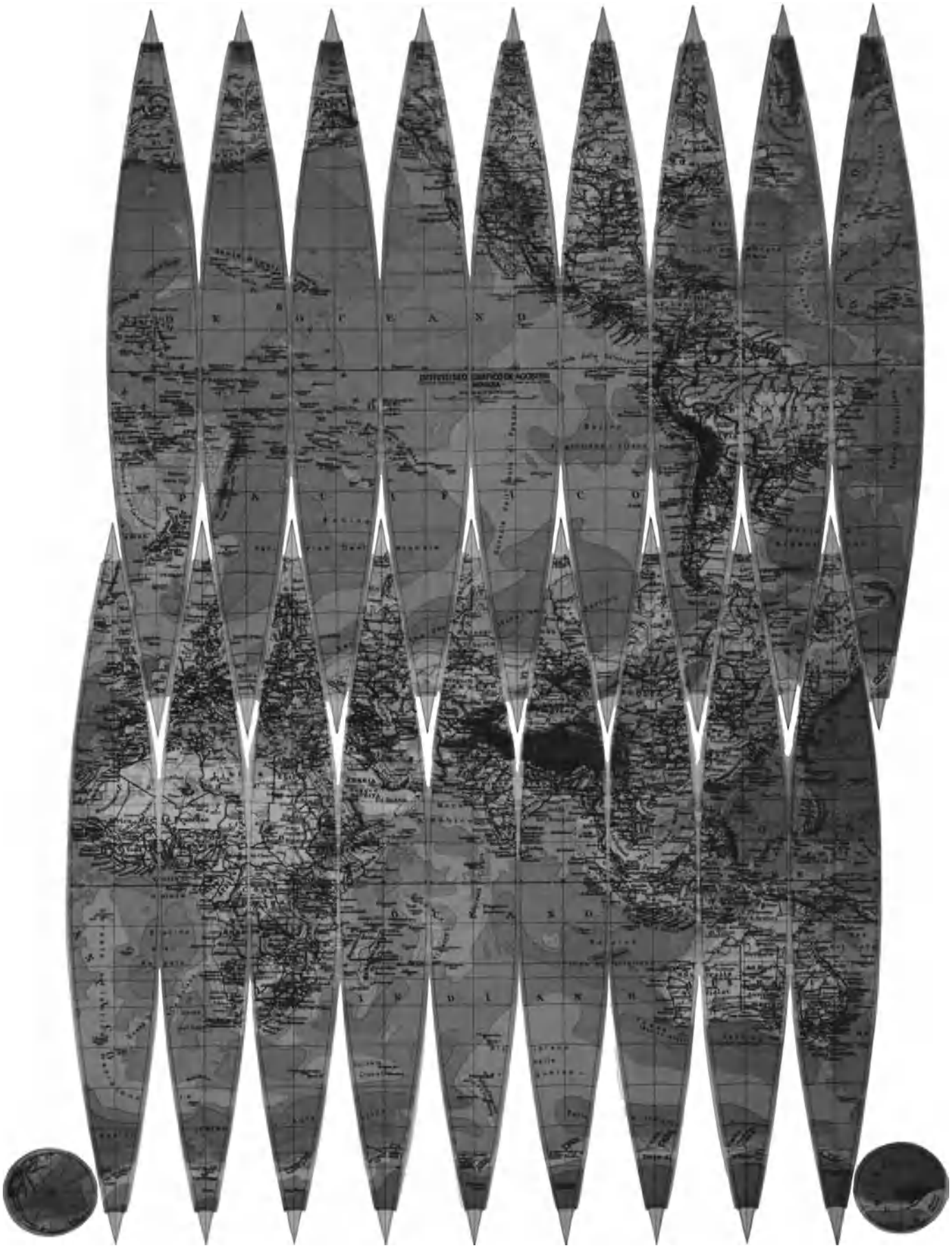

Fig. 1. Globe gores from collections of the Library of Congress. (*Istituto Geografico di Agostini, Novara, Italy*)

other designs, affixed to a cardboard, papier-maché, plastic, wooden, or metal ball (**Fig. 1**). Transparent globes with the map on the inner surface prevent soiling or wear. The surface of a raised relief globe has protrusions representing the Earth's landforms. To suggest mountains and basins, the vertical scale of these surface irregularities is exaggerated. With the development of sensors capable of measuring the ocean depths, recent globes depict relief on the ocean floor as well as on land. Globes constructed from a compilation of satellite images of Earth below depict continent shapes through the Earth's atmosphere.

Uses. Globes are artistically interesting and scientifically useful. Their value is in stimulating sound concepts of worldwide patterns, in rectifying errors induced by the limitations of flat maps, and in measuring distance and time. Long employed as aids in navigation, in teaching of earth sciences, and as library or parlor ornaments, globes are also used as toys, games, advertisements, exhibits, references for travel, and a means to illustrate missile and satellite paths. *See* CARTOGRAPHY.

Modern globes often have special attachments to enhance their utility. A meridian ring, extending from pole to pole, is calibrated in degrees to measure latitude. The longitude of points directly beneath that ring is indicated at the intersection of the ring with the equatorial scale. A horizon ring at right angles to the meridian ring may be calibrated in miles or meters, degrees, and hours to expedite distance and time measurement.

A hinged horizon ring may be lifted to serve as a meridian ring or placed in an oblique position to show great circle routes and distances. Some globes have a time disk loosely capped over the North Pole (**Fig. 2**). When set to the time of a point directly under the meridian ring, it shows the simultaneous time at other longitudes around Earth. The International Date Line is usually plotted in mid-Pacific with an analemma, shaped like a number 8, near it, which shows the latitude at which the Sun is directly overhead on each day of the year. Additional attachments convert globes into sundials, navigation instruments, and models for satellite orbits. As educational devices, globes are often lighted from the inside and may have transparent hemispheric covers on which to write. For interactive learning, a computerized globe was developed which speaks answers to questions about the location of any point on its surface that is touched. Handy inflatable plastic globes are available for the classroom. *See* INTERNATIONAL DATE LINE.

Fig. 2. Globe with hinged horizon ring. At the North Pole is a small disk which rotates independently to show simultaneous times over the Earth. (*Rand McNally and Co.*)

Fig. 3. Wilson globe, from collections of the Library of Congress. This globe is dated 1822. Note that the upper part of the analemma extends north of the horizon ring.

Sizes and mountings. Few globes are more than 18 in. (45 cm) in diameter, and most are 12 in. (30 cm) or smaller. A 52-ft (16-m) globe was exhibited at the Pan Pacific Exposition in 1915. At the Christian Science Publishing House in Boston, people enter a 30-ft (9-m) translucent ball to view the world map from the inside. On an outdoor motor-driven spindle, the 25-ft (8.5-m) Babson Globe was completed in 1955 in Wellesley, MA. A 75-in. (2-m) relief globe was built in 1957 by the Geo-Physical Map Co., now part of Rand McNally and Co. A 50-in. (1.3-m) globe, designed especially for White House use during World War II, had a free-rolling base. Globes in stands become decorative furniture. Others sit on tilted spindles or ride free in cradles or rings. Modern plastics and air compressors allow portable globes large enough for people to stand in to be set up for temporary display.

Famous globes. In the 2d century B.C., Greek geographer Crates of Mallus constructed the earliest known globe. The Germanic Museum in Nürnberg holds the oldest terrestrial globe in existence. This 20-in. (51-cm) original built in that city by Martin Behaim in 1492 shows more than 1100 place names,

is made of pasteboard and gypsum supported by a wooden frame, and is covered with parchment. Of almost equal age is an engraved and gilded copper ball, 7 in. (17 cm) in diameter, discovered in 1850 at Laon, France. During the late Renaissance, the Church of Rome foreswore its 100-year-old ban on the idea of a round Earth, and the forging of ornate metal globes became important work in the coppersmith's craft. Other sixteenth- and seventeenth-century globes were made by European cartographers such as Mercator, Hondius, Heyden, Blaeu, Delisle, and Senex. The most publicized were made by Coronelli for Louis XIV of France. His largest, made in 1680, was 15 ft (4.5 m) in diameter, covered with engraved gores, and equipped with a door to permit approximately 30 people to stand inside. Scores of famous globes reside in the headquarters of the Coronelli-Weltbund der Globusfreunde in Vienna. The New York Public Library holds a globe made in 1506, one of the earliest constructed after the discovery of America. James Wilson built the first globes made in the Americas at Bradford, VT, in 1810 (**Fig. 3**). Near-globes, such as the icosahedron of 20 spherical triangles or the dymaxion of six squares and eight triangles, are made today.

Arch C. Gerlach; Ute J. Dymon; Nancy L. Winter

Bibliography. J. Campbell, *Map Use and Analysis*, 3d ed., 1998; J. Goss, *The Mapmaker's Art: An Illustrated History of Cartography*, 1993; A. H. Robinson, *Elements of Cartography*, 6th ed., 1994.

Globule

A small, opaque nebula seen in silhouette against a rich star field or a bright nebula. Globules were first cataloged in the 1920s. In 1947, B. J. Bok called attention to their potential significance for star formation, and since then they have been commonly known as Bok globules. A globule is a region of the interstellar medium containing a high density of interstellar grains that obscure the more distant background stars and cause the region to appear as a dark nebula in optical photographs. Only relatively nearby globules can be identified, because if there are many stars in front of the nebula the contrast with the background is too weak. The distances to such nebulae have been estimated by counting the number of stars photographed in the line of sight and comparing this count with an analogous star count in a clearer comparison field. Comparative star counts have also been used to evaluate the degree of extinction produced by the grains in the nebula; if the optical properties of the grains are known, the total number of grains in a cut through the nebula can be estimated. *See* INTERSTELLAR EXTINCTION; NEBULA.

The material contained in interstellar grains represents only a small fraction (about 1%) of the total mass of a globule; most of its mass is in gaseous form. The grains play a key role in shielding the nebular gas from the surrounding starlight, thus creating an environment in which molecules can survive and interact. Radio astronomers have been able to detect carbon monoxide (CO) emission lines in dark nebulae, and in the densest cores of nebulae even heavier molecules have been identified. Dark nebulae are now commonly called molecular clouds: the most abundant molecule is molecular hydrogen, but it is difficult to detect directly. Molecular clouds are of great interest as the regions in which stars are formed. *See* COSMOCHEMISTRY; INFRARED ASTRONOMY.

Nearby globules have been surveyed from the radio emission lines of carbon monoxide, ammonia (NH_3), and other molecules. These data, when combined with optical estimates of distances, have led to fairly accurate determinations of the dimensions of globules. Careful studies of ammonia emissions indicate that a typical (molecular core) globule has a diameter of 0.1 parsec (1 parsec = 1.9×10^{13} mi or 3.1×10^{13} km) and mass four times that of the Sun; larger, more massive globules are also known. The kinetic temperature in a globule is low; it is estimated to be about 10 K (−442°F). *See* RADIO ASTRONOMY.

Calculations predict that in the absence of internal support a typical globule undergoes gravitational collapse in less than a million years to produce one or more protostars. These young objects radiate in the infrared, a region of the spectrum to which the nebular grains are transparent. The *Infrared Astronomical Satellite (IRAS)* has mapped the sky at several infrared wavelengths, and astronomers have identified many protostars in the dense cores of molecular clouds. Some but not all globules have given birth to protostars, and these probably represent the youngest stars of the Milky Way. *See* INFRARED ASTRONOMY; STELLAR EVOLUTION.

Beverly T. Lynds

Bibliography. E. Levy, J. I. Lunine, and M. S. Matthews (eds.), *Protostars and Planets*, vol. 3, 1993; B. Lynds (ed.), *Dark Nebulae, Globules, and Protostars*, 1971.

Globulin

A general name for any member of a heterogeneous group of serum proteins precipitated by 50% saturated ammonium sulfate, and thus differing from albumin, the protein present in greatest concentration in normal serum. Originally, the globulins were further subdivided on the basis of solubility (pseudoglobulin) or insolubility ("true globulins") in salt-free pure water. However, the sharp demarcations implied by these definitions do not exist since globulins show varying degrees of solubility in water and in solutions of various ionic strengths. *See* PROTEIN; SERUM.

The introduction of electrophoresis during the 1930s permitted separation of albumin from globulin on the basis of differing electric charge, and of subdivision of the globulins into alpha, beta, and gamma globulins on the basis of relative mobility at alkaline pH (8.6). However, each of these subgroups, though electrophoretically homogeneous, consists of a great variety of proteins with different biological

properties and markedly different sizes and chemical properties other than net charge. Thus the α_2-globulins, for example, as defined by moving boundary or paper electrophoresis, contain proteins ranging in molecular weight from approximately 50,000 to approximately 1,000,000 (α_2-macroglobulin), each with differing functions; for example, haptoglobin binds free serum hemoglobin, ceruloplasmin is involved in copper transport and metabolism, and the α_2-macroglobulin binds or stabilizes certain enzymes, such as trypsin and plasmin. *See* BLOOD; ELECTROPHORESIS.

The beta globulins are similarly heterogeneous, containing transferrin (the iron-binding protein), various complement components, beta lipoproteins, and other biologically active proteins. The globulins of slowest electrophoretic mobility, the gamma globulins, contain a family of at least five distinct proteins, each associated with antibody activity. *See* ANTIBODY; COMPLEMENT. H. Hugh Fudenberg

Bibliography. F. W. Putnam (ed.), *The Plasma Proteins*, vols. 1–2, 1975; N. J. Russell and G. M. Powell, *Blood Biochemistry*, 1983.

Glow discharge

A mode of electrical conduction in gases. Glow discharge commonly occurs under conditions of relatively low pressure and generally in the pressure range of 1–10 mm of mercury (100–1000 pascals). The discharge typically gives off light, so that the region of the discharge appears to glow with considerable intensity. This glow is quite diffuse as contrasted to a higher-pressure discharge, such as a high-pressure arc. Typical currents may be of the order of tens or hundreds of milliamperes, whereas the potential drop may be of the order of 100 V.

The most important application of the glow discharge is in the so-called voltage regulator or voltage reference tube. This device maintains a relatively constant difference of potential across itself as the current is varied over an appreciable range, and consequently is very useful in cases where a constant reference potential is required.

In terms of the potential-current characteristic, the glow discharge occurs after the potential has been increased so that the Townsend region has been passed. Thus the discharge is field-sustained. On the other hand, a continued increase in current leads first to the region referred to as the abnormal glow and beyond this to the arc discharge. The transition from the abnormal glow to the arc generally is almost discontinuous and is accompanied by a spark. For a discussion of this relationship *see* ARC DISCHARGE; ELECTRIC SPARK; ELECTRICAL CONDUCTION IN GASES; TOWNSEND DISCHARGE.

Regions of discharge. There are three main regions of interest in the glow discharge, similar to those in the arc. These are the cathode fall, the positive column, and the anode region. These will be discussed separately, but it is appropriate first to examine some of the general features of the mode (see **illus.**). The appearance is that of successive more or less well-defined luminous and dark regions. Starting from the cathode, there is a dark space which generally extends for a few millimeters, the Aston dark space. This is followed by a luminous region, also of limited extent, known as the cathode glow. This is succeeded by a somewhat longer dark space, designated the Crookes or Hittorf dark space. After this comes the negative glow region, the boundaries of which are rather poorly defined. Following this is the Faraday dark space, which is also more extensive and poorly defined. This changes gradually into the positive column which is luminous and of length determined by the pressure and distance between electrodes. This region may or may not contain striations, and if present they may be either stationary or moving. At the end of the positive column is a thin layer of greater luminosity, designated the anode glow. Between this and the anode is the anode dark space.

Cathode fall. The events occurring at the cathode are vital to the discharge. The current in the cathode circuit is primarily due to positive ions. However, it is necessary to produce enough electrons at the cathode to maintain the discharge. These electrons gain energy as they move in the electric field toward the anode, and produce excitation and ionization. It appears that these electrons are secondary electrons resulting from positive ion bombardment of the cathode surface. The drop in potential which occurs at the cathode depends on the kind of gas and the cathode material. Generally, this potential drop is a large fraction of the total potential drop across the

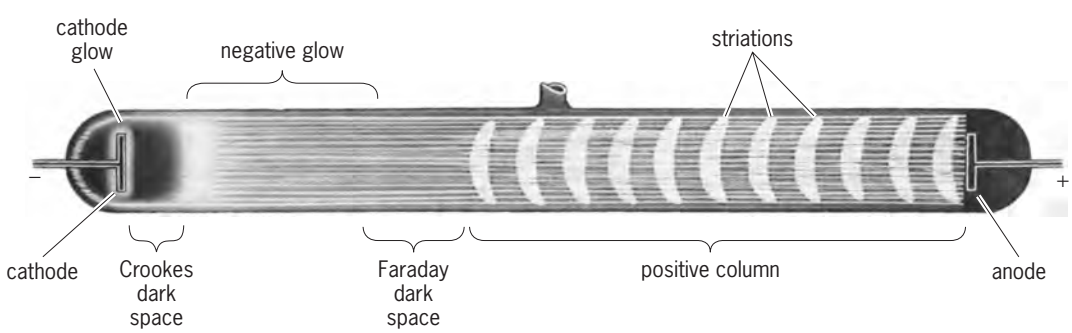

Glow discharge at approximately 0.1 mm (13 Pa) pressure, showing successive more or less well-defined luminous and dark regions. (*After J. B. Hoag and S. A. Korff, Electron and Nuclear Physics, 3d ed., Van Nostrand, 1948*)

discharge. The production of secondary electrons by this means is rather inefficient, which explains why the drop must be large.

Electrons starting at rest from the cathode must gain energy before they can produce excitation. This can be accomplished only by motion in the electric field, and hence there is a minimum distance which the electrons must move before they can produce excitation and consequent light emission. This explains the existence of the Aston dark space. It might be thought that the cathode glow could be explained by this also, but it is not likely that much of the light from this region is brought about by the secondary electrons. It appears that most of this light results from the positive ions that have struck the cathode and are returning to the ground state as neutral atoms. There are two facts of importance in this connection. First, the electron density is still rather low at such a short distance from the cathode. Second, the wavelengths present in the radiation indicate transitions involving states of a rather high degree of excitation. These high-energy states probably could not be produced by the electrons from the cathode at this point.

The Crookes dark space is actually a region of nearly uniform electric field. Most of the cathode drop occurs in this region, and here the positive ions gain most of their energy before striking the cathode. The electrons from the cathode gain enough energy here to produce cumulative ionization near the end of the space. In the negative glow, which follows, the potential is relatively constant. Here electrons, both from the cathode and from cumulative ionization, lose energy by inelastic collisions and produce a large amount of excitation. The boundary at the anode end of this space is poorly defined because of the broad distribution in electron energy. The slowing down of the electrons at the end of this region results in a negative space charge. Thus the electrons that move into the Faraday dark space gain energy.

Positive column. The beginning of the positive column is the result of excitation by these electrons. The situation in the positive column is the result of a balance between several processes. There is a nearly uniform potential drop which results in ionization throughout this entire region. On the other hand, there must be a loss of ions to make up for this production. This takes place primarily by diffusion to the walls, although there is also recombination. The electrons with their greater mobility diffuse to the walls, producing a slight negative potential relative to the center of the discharge. This negative potential both limits further electron diffusion and produces positive ion diffusion outward. This process is known as ambipolar diffusion. The positive column is not essential in maintaining the discharge. If the distance between electrodes is changed, with the pressure and current held constant, the extent of the positive column and the potential across the discharge change accordingly. The features of the anode and cathode regions remain unaltered under such a change up to the point where the positive column no longer exists.

A feature of this region is a succession of alternately luminous and dark regions, called striations, which usually occur when the discharge is operated at relatively high pressure; they may be stationary or moving. Their presence is related to the fact that in general the atomic species in the discharge are deexcited in times short compared to those required for them to diffuse through the positive column. Pure, inert gases do not show the effect, probably because they are excited into metastable, long-lived states.

Anode region. At the anode end of the positive column, the positive ions are repelled. This produces an increase in electric field, which causes the electrons to gain energy and excite more effectively. Thus the positive column ends in a region of increased luminosity, the anode glow.

Other aspects. There are many other aspects of the glow discharge that are interesting and important. One such phenomenon is cathode sputtering. Here the positive ions that are accelerated into the cathode knock out atoms or groups of atoms from the surface. Another aspect is that of abnormal glow. The voltage across the discharge remains nearly constant while the current is increased in the normal glow mode. This current increase is accompanied by an increase in the area of the cathode glow. When the glow has completely covered the cathode, a further current increase results in an increase in the cathode potential drop, and hence the potential drop across the discharge. This is the abnormal glow. It is characterized by more intense light emission and increased sputtering. *See* SPUTTERING.

It should be stated that many of the details of the discharge are uncertain. The processes are generally quite complicated. Reliable and accurate measurements are difficult at best, and most of the information is of a qualitative nature. Glenn H. Miller

Bibliography. B. Chapman, *Glow Discharge Processes: Sputtering and Plasma Etching*, 1980; M. N. Hirsh and H. J. Oskam (eds.), *Gaseous Electronics*, vol. 1: *Electrical Discharges*, 1978; R. K. Marcus (ed.), *Glow Discharge Spectroscopies*, 1993; Yu. P. Raizer, *Gas Discharge Physics*, 1991, reprint 1997.

Glucagon

The protein hormone secreted by the alpha cells of the pancreas which is known to influence a wide variety of metabolic reactions. Glucagon, along with insulin and other hormones, plays a role in the complex and dynamic process of maintaining adequate supplies of blood sugar (glucose). Glucagon has often been called the hyperglycemic-glycogenolytic factor because it causes the breakdown of liver glycogen to sugar (a process known as glycogenolysis) and thereby increases the concentration of sugar in the bloodstream (a condition known as hyperglycemia). Glucagon may also be involved in the regulation of protein and fat metabolism, gastric acid secretion and gut motility, excretion of electrolytes (such as sodium, potassium, and chloride) by the kidney, contractility of heart muscle, and release of

insulin from the pancreas. On the basis of present evidence, glucagon may be considered to act in a close and subtle cooperation with other hormones and with enzymes to help control the dynamic processes of life. Glucagon is used in human medicine chiefly in certain diabetic conditions when a dangerously low blood sugar must be rapidly raised. *See* CARBOHYDRATE METABOLISM; DIABETES; ENDOCRINE MECHANISMS; ENZYME; GLYCOGEN; INSULIN; LIVER; PANCREAS.

Discovery and isolation. The history of glucagon is closely interwoven with that of insulin. During 1922–1927 several investigators were surprised to observe that injections of crude preparations of insulin caused an unexpected, rapid hyperglycemia followed by the expected drop in blood sugar. Interest in the dramatic, lifesaving effects of insulin greatly overshadowed this hyperglycemic activity found in extracts of pancreas. However, J. R. Murlin and his associates investigated the hyperglycemic effect, became convinced that it was caused by another substance in the pancreas, and termed that substance glucagon, meaning "mobilizer of sugar." Interest in this "contaminant" of insulin lay partially dormant for about two decades until 1953, when A. Staub, L. Sinn, and O. K. Behrens succeeded in purifying and crystallizing glucagon. The availability of the pure hormone led to an intensive study of its chemical and biological properties.

Assay. All measurements of glucagon potency are related arbitrarily to the crystalline hormone; no international standard has been designated as yet. The specific biological potency is usually determined by injecting glucagon intravenously or subcutaneously into cats or rabbits and measuring the degree of hyperglycemia. As little as 0.2–0.3 microgram injected intravenously into a fasted adult cat causes an increase (within 15 min) of 30–50 milligrams of sugar per 100 milliliters of blood. Glucagon may also be measured by a means of a "radioimmune" assay, in which radioactively labeled glucagon and antibodies to glucagon are employed. *See* BIOASSAY.

Chemistry. Glucagon is a single-chain polypeptide composed of 29 amino acid residues (**Fig. 1**), and has a molecular weight of 3485. The complete chemical synthesis of glucagon was achieved late in 1967.

Despite the fact that glucagon and insulin are difficult to separate, showing a similarity in some chemical and physical properties, the actual chemical composition and amino acid sequences of the two hormones are quite different. Glucagon is only slightly soluble in water between pH 4 and pH 8. It forms threadlike polymers, called fibrils, when heated in dilute acids. The hormone is very rapidly cleaved by enzymes such as trypsin and chymotrypsin, leading to loss of hyperglycemic activity. This is the chief reason that glucagon must be injected rather than taken by mouth. William W. Bromer

Biological activity. The mechanism by which glucagon brings about an increase in blood sugar level by inducing the breakdown of liver glycogen has been the subject of extensive research over the past few decades. Glucagon molecules bind to

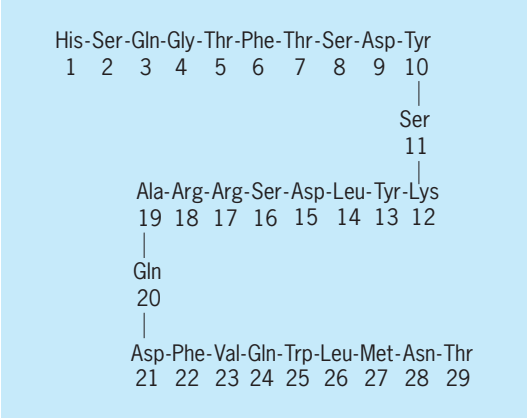

Fig. 1. Amino acid sequence in a molecule of glucagon. The abbreviations indicate: histidine (His), serine (Ser), glutamine (Gln), glycine (Gly), threonine (Thr), phenylalanine (Phe), aspartic acid (Asp), tyrosine (Tyr), lysine (Lys), leucine (Leu), arginine (Arg), alanine (Ala), valine (Val), tryptophan (Trp), methionine (Met), and asparagine (Asn).

glucagon-specific protein receptors in the membrane of target cells. When these receptors bind glucagon, they interact with other proteins in the cell membrane called G proteins and "activate" these protein switches. The activated G proteins then activate an enzyme on the inner surface of the cell membrane called adenylate cylase. Activated adenylate cyclase then produces a small molecule called 3′,5′-cyclic adenosine monophosphate (cAMP), commonly referred to as a second messenger molecule (glucagon being the first messenger in this case). It is cAMP that activates the glycogen breakdown to sugar in the liver cells. *See* SIGNAL TRANSDUCTION.

Cloning and analysis of the gene encoding glucagon has led to the discovery of two glucagon-like peptides called GLP-1 (glucagon-like peptide 1) and GLP-2 (glucagon-like peptide 2). In recent years, research has focused on production and biological activity of GLP-1 and GLP-2. Glucagon, GLP-1, and GLP-2 are all encoded by the same gene. This gene produces a protein called proglucagon in cells in the pancreas, gastrointestinal tract, and brain. In the pancreas cells, proglucagon is processed and converted into glucagon. In the gastrointestinal tract and brain cells, proglucagon is processed and converted into GLP-1 and GLP-2 (**Fig. 2**). *See* GENE ACTION.

GLP-1. GLP-1 has the opposite effect of glucagon. It lowers blood sugar levels by a variety of means: it stimulates the production of insulin, which enhances

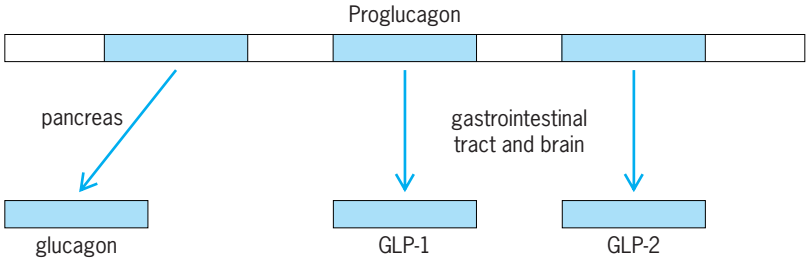

Fig. 2. Proglucagon protein is processed in the pancreas to yield the glucagon peptide and is processed in the gastrointestinal tract and brain to yield the GLP-1 and GLP-2 peptides.

the consumption of sugar by tissues; it inhibits gastric secretions, which decreases the absorption of sugars following meals; and it works within the central nervous system to reduce appetite, which also leads to low sugar levels. The ability of GLP-1 to stimulate insulin secretion has generated a great deal of interest in the use of GLP-1 in therapies for type II diabetes, in which insulin secretion is impaired. Studies have shown that injections of GLP-1 indeed stimulate insulin secretion, but the effect is short lived. Currently, extensive work is under way to find modifications of GLP-1 that will extend its half-life and make GLP-1 an effective therapy for type II diabetes.

GLP-2. Less is known regarding the role of GLP-2. It is an intestinal growth factor that stimulates the division of the cells of the gastrointestinal lining. Thus GLP-2 enhances nutrient absorption during digestion. Current research is aimed at using GLP-2 in clinical treatments for intestinal damage and other related problems. M. Todd Washington

Bibliography. T. M. Devlin (ed.), *Textbook of Biochemistry with Clinical Correlation*, 5th ed., Wiley, New York, 2001; D. Voet and J. G. Voet, *Biochemistry*, Wiley, New York, 1995.

Glucose

A simple sugar, or monosaccharide, with the chemical formula $C_6H_{12}O_6$; also called dextrose, glucopyranose, grape sugar, and corn sugar. *See* MONOSACCHARIDE.

Chemistry. Glucose is a carbohydrate with six carbon atoms and has the structure

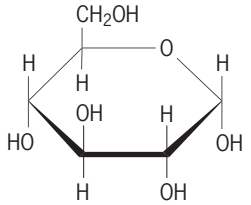

Glucose is synthesized in nature by photosynthetic organisms from carbon dioxide and water according to the general equation:

$$\underset{\text{Water}}{6H_2O} + \underset{\text{Carbon dioxide}}{6CO_2} + \underset{\text{Energy}}{\text{sunlight}} \longrightarrow \underset{\text{Glucose}}{C_6H_{12}O_6} + \underset{\text{Oxygen}}{6O_2}$$

It is stored as starch in plants and as glycogen in animals. Starch and glycogen are polymers of glucose, in which individual glucose molecules have been strung together in long chains that efficiently store the chemical energy contained in the individual glucose molecules. In addition, glucose molecules serve a variety of functions in the natural world. Many of these uses are structural. For example, cellulose, which forms the cell walls of plants, is a polymer of glucose. In addition, the exoskeleton of insects and the cell walls of fungi are formed from a chemical derivative of glucose that has been assembled into chitin, a lengthy polymer. In its various polymeric forms, glucose is likely the most abundant organic molecule in the biosphere. *See* CARBOHYDRATE; CELLULOSE; GLYCOGEN; STARCH.

Metabolism. Glucose is a major source of energy for humans and other animals. However, our bodies do not maintain a large reservoir of individual glucose molecules. Rather, animals store glucose in their bodies in the form of glycogen, a large polymer of glucose found primarily in liver and muscle tissue. When our bodies need energy, for example between meals or during periods of extended fasting, individual glucose molecules can be extracted from glycogen. In the form of glycogen, glucose is securely stored as a large and stable polymer, yet it is readily accessible when needed by the body. After being released from glycogen stores, glucose molecules travel in the bloodstream to the various tissues, where they are absorbed into individual cells. Once inside a cell, glucose undergoes a series of chemical reactions that allows the cell to tap the energy contained within its chemical bonds. This process is referred to as the catabolism, or breakdown, of glucose. The overall equation for the catabolism of glucose is

$$\underset{\text{Glucose}}{C_6H_{12}O_6} + \underset{\text{Oxygen}}{6O_2} \longrightarrow \underset{\text{Carbon dioxide}}{6CO_2} + \underset{\text{Water}}{6H_2O} + \underset{\text{Heat and ATP}}{\text{energy}}$$

Since the catabolism of glucose involves the consumption of oxygen, this process is also called cellular respiration. The energy in the above equation represents heat energy, which is dissipated and lost, and energy from the chemical bonds of glucose that has been captured in the chemical bonds of the organic molecule adenosine triphosphate (ATP). ATP is a convenient energy source that can be used to drive a variety of chemical reactions. For example, cells harness the energy contained in the chemical bonds of ATP to perform many kinds of cellular work, such as the contraction of muscles and the biosynthesis of organic molecules. *See* ADENOSINE TRIPHOSPHATE (ATP).

The overall equation for glucose catabolism shown above takes place in cells as a series of chemical reactions that are grouped into three distinct steps. In the first step, glucose undergoes glycolysis, that is, the splitting of glucose. In this process, one glucose molecule (which contains six carbon atoms) is split into two molecules of pyruvate (which contains three carbon atoms). Glycolysis allows the cell to transfer some of the energy contained within glucose to ATP and another organic molecule called NADH. In the second step, the pyruvate molecules enter another series of chemical reactions, called the Krebs cycle, which allows the cell to capture more of the energy of glucose in the form of NADH and yet another organic molecule called $FADH_2$. The NADH and $FADH_2$ molecules produced during glycolysis and the Krebs cycle are coupled to the synthesis of ATP by a process called oxidative phosphorylation, which is the third and final step of glucose catabolism. The complete catabolism of glucose generally yields about 36 molecules of ATP per

molecule of glucose. The catabolism of glucose is summarized schematically in the **illustration**. *See* CARBOHYDRATE METABOLISM; ENERGY METABOLISM; METABOLISM.

Disease. The brain and other organs depend on glucose as a source of energy. Humans and other animals do not generally consume large amounts of free glucose molecules from their diets. Rather, glucose is usually ingested as starch, which is a polymer of glucose, or as sucrose (table sugar), lactose (milk sugar), and maltose (malt sugar). The latter three sugars are disaccharides, in which a glucose molecule has been chemically linked to another carbohydrate molecule. Our bodies have the ability to convert the starch and disaccharide molecules into individual glucose molecules. Since glucose is shuttled between the tissues of our body by the cardiovascular system, the concentration of glucose in the blood is carefully regulated in order to meet the energy demands of the brain and other tissues. Indeed, humans and other mammals maintain remarkably stable blood glucose concentrations despite sporadic food intake. Between meals, blood glucose levels decline and the liver produces a compensatory amount of glucose from glycogen and other sources; after eating a meal, glucose is absorbed from the gut into the bloodstream. The hormone insulin causes this circulating glucose to be taken up by skeletal muscle and adipose (fat) tissue. This complicated three-way balance involving glucose absorption from the small intestine, output from the liver, and uptake by muscle and fat is normally carefully regulated such that the concentration of glucose in the blood is maintained in the range of 70–120 milligrams glucose per deciliter of blood (4–7 millimolar) in humans. Deviations from this narrow concentration range can be very detrimental to health. For example, low blood glucose levels can lead to seizures, coma, and even death due to in adequate supply of glucose to the brain. In contrast, a prolonged elevated blood glucose level, as occurs in the diabetic state, can result in many disease complications. These include kidney failure, blindness, nerve damage, and cardiovascular disease.

Of the several types of diabetes, Type 1 or insulin-dependent diabetes mellitus (IDDM) and Type 2 or non-insulin-dependent diabetes mellitus (NIDDM) are the most prevalent. Type 1 occurs when the body's own immune system mistakenly attacks and destroys a specific type of cell located in the pancreas called the beta cell. The beta cells produce insulin, so the loss of these cells means that the body can no longer make its own insulin. Type 1 diabetics must carefully monitor their own blood glucose levels and take insulin, usually by injection, prior to eating. Type 2 diabetes mellitus is a complicated disease that usually develops over time. Early in the disease progression, muscle and fat tissue no longer respond appropriately to insulin and fail to efficiently absorb glucose. To compensate, the beta cells increase their output of insulin in an effort to encourage muscle and fat cells to take up glucose. All too often, however, the extra demands placed upon the beta cells cause them to wear out. Indeed, over time the beta cells may become unable to produce enough insulin to coax muscle and fat cells to absorb glucose from the blood. This condition results in the buildup of glucose in the blood associated with diabetes and may lead to disease complications. *See* DIABETES; INSULIN.

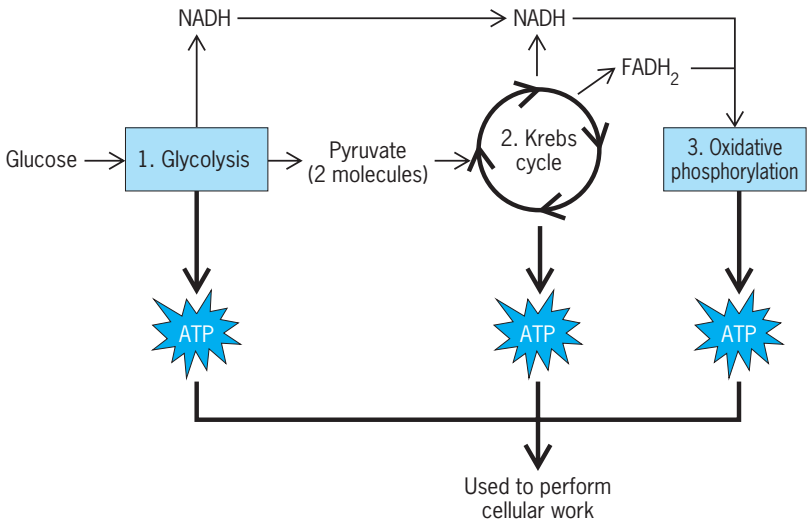

Breakdown (catabolism) of glucose. Glucose is catabolized in three steps: The final product is ATP (adenosine triphosphate), which is an organic molecule that cells use as a source of energy to perform cellular work.

Glucose transporter. Facilitated glucose transport is the movement of glucose across cell membranes that is driven by the glucose concentration gradient, but assisted (facilitated) by carrier proteins. It is energy-independent, and it is stereospecific in that only the D-glucose isomer is transported; the L-glucose isomer is excluded. This process occurs in all mammalian cells and is essential for the maintenance of whole-body glucose metabolism and energy balance. Currently, there are five established functional facilitative glucose transporters in mammalian cells, termed GLUT1, GLUT2, GLUT3, GLUT4, and GLUTx. Each of these transporters has distinct but overlapping tissue distributions, which underscore their specific physiologic function.

GLUT1 is generally expressed and is thought to be responsible for the basal (minimum) uptake of glucose. GLUT2 is predominantly expressed in the liver and pancreatic beta cells, where it functions as part of a sensor that mediates hepatic glucose output during states of fasting and insulin secretion in the postprandial (after a meal) absorption state. In contrast, neurons primarily express the relatively high-affinity GLUT3 isoform necessary to maintain high rates of glucose metabolism for energy production. GLUTx appears to provide important function during early embryogenesis, whereas GLUT4 is exclusively expressed in insulin-responsive tissues, adipose tissue, and striated (skeletal) muscle. These latter tissues provide the key functional elements responsible for the insulin stimulation of glucose uptake, and are key targets for disregulation in states of insulin resistance and diabetes.

Jeffrey Pessin; Robert Watson

Bibliography. G. I. Bell and K. S. Polonsky, Diabetes mellitus and genetically programmed defects in beta-cell function, *Nature*, 414:788–791, 2001; M. A. Lazar, How obesity causes diabetes: Not a tall tale, *Science*, 307:373–375, 2005; S. O'Rahilly, I. Barroso, and N. J. Wareham, Genetic factors in Type 2 diabetes: The end of the beginning?, *Science*, 307:370–373, 2005; C. J. Rhodes, Type 2 diabetes—a matter of β-cell life and death?, *Science*, 307:380–384, 2005; R. T. Watson, M. Kanzaki, and J. E. Pessin, Regulated membrane trafficking of the insulin-responsive glucose transporter 4 in adipocytes, *Endocr. Rev.*, 25:177–204, 2004; P. Zimmet, K. G. M. M. Alberti, and J. Shaw, Global and societal implications of the diabetes epidemic, *Nature*, 414:782–787, 2001.

Gluons

The hypothetical force particles believed to bind quarks into strongly interacting particles. Theoretical models in which the strong interactions of quarks are mediated by gluons have been successful in predicting, interpreting, and explaining many phenomena in particle physics, but free gluons remain undetected in experiments (as do free quarks). According to prevailing opinion, an individual gluon cannot be isolated.

Color. In 1961 M. Gell-Mann and Y. Ne'eman independently suggested that the strong (nuclear) interaction respected the unitary symmetry SU(3) and that the strongly interacting particles called hadrons could be classified according to the patterns prescribed by SU(3). The family groups, or supermultiplets, that emerged were confined to a few of the simplest possibilities permitted under SU(3) symmetry. Mesons, the hadrons with integral spin in units of $\hbar$ (Planck's constant h divided by 2π), occur only in families with 1 or 8 members. The baryons, which possess half-internal spin, fit into groups with 1, 8, or 10 members. Gell-Mann and G. Zweig separately showed in 1963 that this circumstance could be explained by the hypothesis that hadrons were composites of fundamental constituents that have come to be called quarks. In this quark model of hadrons, a meson is composed of one quark and one antiquark, and a baryon is composed of three quarks. All the hadrons then known could be built out of three different varieties (or flavors) of quarks, denoted up, down, and strange. Subsequent experiments have revealed the existence of three more flavors of heavy quarks: charm, bottom, and top. To account for the observed pattern of mesons and baryons, quarks must be spin-$^1/_2$ particles. *See* QUARKS; UNITARY SYMMETRY.

Although these rules reproduce the properties of the observed hadron states, they lead to a theoretical inconsistency. The characteristics of the unstable hadron resonance known as Δ^{++} (1232 MeV/c^2), which decays into a proton and a positively charged pi meson, require that it be composed of three quarks in a configuration that is symmetric under the interchange of any pair of quarks. However, according to the Pauli exclusion principle (which first emerged in the description of atomic structure), identical spin-$^1/_2$ particles cannot occupy the same quantum state. The quark model could be brought into agreement with the Pauli principle, without compromising any of its successes, if a new attribute were ascribed to the quarks that would make the three up quarks distinguishable. For fanciful reasons, this new attribute is now known as color, though it has no connection with the color of visible light. Quarks are said to come in three colors, most frequently given the arbitrary labels red, blue, and green. A Δ^{++} resonance composed of one red up quark, one blue up quark, and one green up quark will then have the observed properties and be consistent with the laws of quantum mechanics. In this picture the antiparticle of a red up quark is an anti-red anti-up quark, so that the mesons are described as colorless quark-antiquark pairs.

Support for the idea that each quark flavor comes in three distinguishable colors has come from many quarters including the rate at which strongly interacting particles are produced in electron-positron annihilations. Theoretical predictions for such observables are sensitive to the number of distinct quark species and thus to the number of colors.

The fundamental particles that do not experience strong interactions are the leptons, which like the quarks are spin-$^1/_2$ particles that are structureless at the current limits of resolution. The most familiar examples of leptons are the electron, the muon, and the neutrinos. Each lepton flavor comes in but a single species, which is to say that leptons are colorless. It is therefore appealing to regard color as the strong-interaction analog of the electric charge. Like electric charge, color cannot be created or destroyed in any of the known interactions; it is said to be conserved. *See* COLOR (QUANTUM MECHANICS); LEPTON.

Gauge symmetry. The existence of a conserved quantity is quite generally a consequence of a continuous group of symmetry transformations that leave the laws of physics unchanged in form. For example, the conservation of energy follows from the fact that physical laws depend upon the time interval between occurrences, and not upon an absolute time measured on some master clock. Translation in time (that is, the resetting of clocks) is a symmetry of the equations of physics. Symmetries relating to internal properties of particles, like electric charge, are known as internal symmetries. In the case of conservation of the color charge, a natural choice is the unitary group SU(3), now applied to color rather than flavor. A conservation law follows, by Noether's theorem, from invariance under a global (position-independent) continuous symmetry. If the equations of physics are required to be invariant in form under local symmetry transformations that may be different at every point in space and time, the interactions related to the symmetry are completely fixed. The manner in which this could be accomplished for any continuous symmetry was indicated by C. N. Yang and R. L. Mills in 1954. *See* SYMMETRY LAWS (PHYSICS).

Yang-Mills theory. It frequently happens that the symmetries respected by a phenomenon are recognized before a complete theory has been developed. The question thus arises as to whether a complete theory of nuclear forces could be deduced from a knowledge of the symmetry. Nuclear forces had long been known not to distinguish between the proton and neutron. From the point of view of nuclear forces, the designations proton and neutron are purely conventional. This symmetry among protons and neutrons is called isospin invariance. Yang and Mills investigated the consequences of the hypothesis that the nuclear force among protons and neutrons could be derived by imposing local isospin invariance (which is to say that the convention could be chosen independently at every point of space and time). In general, the requirement of local gauge invariance implies that the interaction must occur through the exchange of massless spin-1 bosons. One species of force particle corresponds to each conserved quantity. This made the Yang-Mills theory unacceptable as a description of nuclear forces. It predicted that nuclear forces were mediated by three massless "gauge bosons," whereas the short range (on the order of 10^{-15} m) over which nuclear forces are observed to act demands that the force particles be massive, as proposed by H. Yukawa. *See* GAUGE THEORY; I-SPIN; QUANTUM FIELD THEORY.

Quantum chromodynamics. Applying similar reasoning to the idea that a local color gauge symmetry should prescribe the strong interaction among quarks leads to the gauge theory of strong interactions that has been called quantum chromodynamics (QCD). The mediators of the strong interaction are eight massless vector bosons, which are named gluons because they make up the "glue" that binds quarks together. It is hoped that the infinite range of the forces mediated by the gluons may help to explain why free quarks have not been isolated. The gluons themselves carry color. Hence, strong interactions among gluons will also occur through the exchange of gluons. It is therefore believed that gluons, as well as quarks, may be permanently confined. According to this view, only colorless objects may exist in isolation. The notion of color confinement is supported by many calculations, notably those made in the discrete (lattice) version of QCD. A complete analytic proof has not yet been constructed. *See* QUANTUM CHROMODYNAMICS.

Experimental evidence. No evidence has been reported for isolated or free gluons. As indicated above, the current interpretation of quantum chromodynamics is that free gluons cannot exist. Therefore it is necessary to devise indirect means to test the idea that gluons exist with all the desired attributes.

Inelastic electron-proton scattering. Early support for the existence of an electrically neutral glue within the proton came from 1968 experiments on inelastic electron-proton scattering carried out at the Stanford Linear Accelerator Center (SLAC). These experiments indicated that the electrons were not scattered electromagnetically from the proton as a whole, but from individual pointlike charged objects subsequently identified with the quarks. They also showed that only about half of the energy of a rapidly moving proton is carried by its charged constituents. The remainder must then be borne by neutral constituents that do not interact electromagnetically. This role would naturally be played by the gluons.

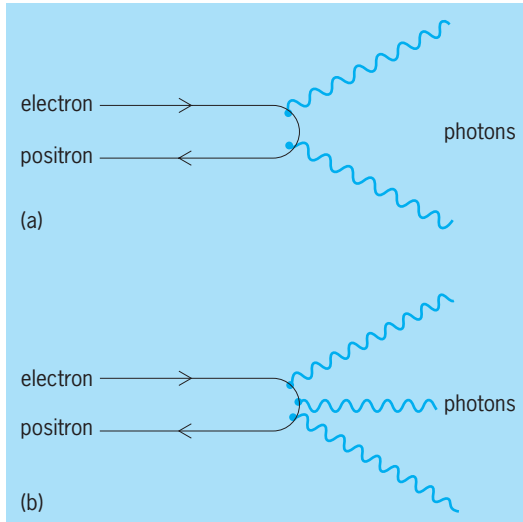

Fig. 1. Positronium decay. (*a*) Decay of parapositronium into two photons. (*b*) Decay of orthopositronium into three photons.

Charmonium lifetimes. Further evidence for the utility of the gluon concept was provided by the unusually long lifetime of the strongly decaying charmonium state J/ψ. In quantum electrodynamics the atom composed of an electron and an antielectron (positron) is known as positronium. Positronium occurs in two forms: orthopositronium, in which the electron and positron spins are aligned, and parapositronium, in which the spins are opposite. The electron and positron may annihilate into photons. The spinless parapositronium state may decay into two photons (**Fig. 1***a*), but the spin-1 orthopositronium state must decay into three photons (Fig. 1*b*). The difficulty of radiating an additional photon is reflected in the fact that orthopositronium lives

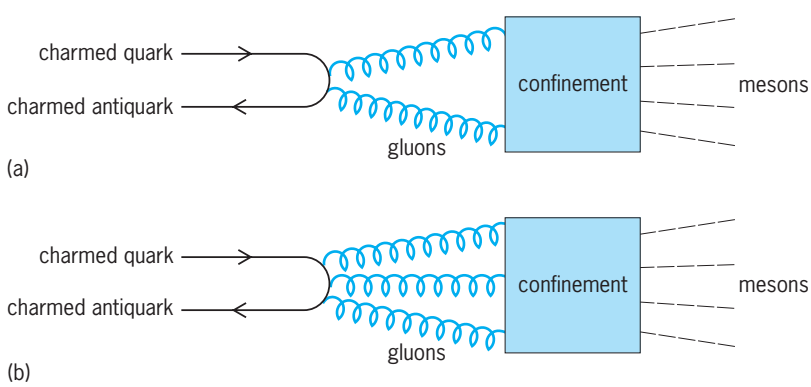

Fig. 2. Charmonium decay. (*a*) Decay of paracharmonium through a two-gluon semifinal state. (*b*) Decay of orthocharmonium through a three-gluon semifinal state.

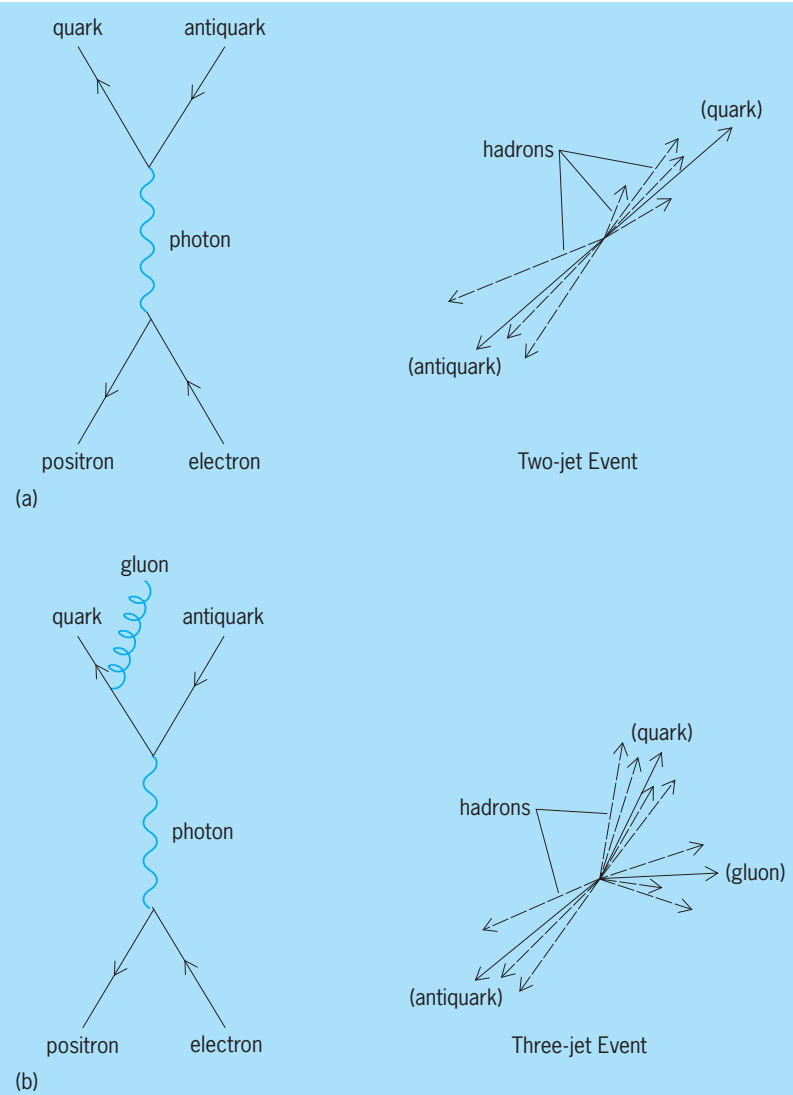

Fig. 3. Mechanisms for hadron production in electron-positron annihilation. (*a*) Two-jet event produced by the mechanism electron positron → quark + antiquark. (*b*) Three-jet event produced by the mechanism electron + positron → quark + antiquark + gluon. The observed hadrons are represented by broken lines.

1120 times longer than parapositronium. In similar fashion, charmonium, the strong-interaction "atom" composed of a charmed quark and a charmed antiquark, decays by the annihilation of the quark and antiquark into gluons. The gluons materialize into the observed hadrons through the action of the confinement mechanism, with unit probability. For the pseudoscalar paracharmonium level, designated η_c, the semifinal state is composed of two gluons (**Fig. 2***a*). The vector particle J/ψ, which corresponds to orthocharmonium, must decay into three gluons (Fig. 2*b*). The remarkably long lifetime of J/ψ and the large ratio of J/ψ to η_c lifetimes (approximately 500) argue for the aptness of the analogy. Decays of the still heavier quarkonium state upsilon also support this. *See* CHARM; J/PSI PARTICLE; POSITRONIUM.

Three-jet pattern. Compelling evidence for the existence of gluons was reported in 1979 by a number of experimental groups working at the high-energy electron-positron storage ring PETRA at the Deutsches Elektronen-Synchrotron (DESY) in Hamburg. It had earlier been established in experiments at SLAC and DESY that the dominant mechanism for hadron production in electron-positron annihilations is electron + positron → quark + antiquark, with the quarks materializing into hadrons. This interpretation explains the rate of particle production and the characteristic angular distribution of the sprays or jets of hadrons that emerge from the collisions (**Fig. 3***a*). If quantum chromodynamics is correct, one of the outgoing quarks may occasionally radiate an energetic gluon, just as a fast electron may radiate an energetic photon. When this happens, the hadrons may be expected to emerge in a three-jet pattern (Fig. 3*b*). Three-jet events are commonplace in electron-positron collisions at center-of-mass energies exceeding about 24 GeV. The characteristics of these events are consistent in every particular with the interpretation that the gluon radiation predicted by quantum chromodynamics is being observed indirectly.

High transverse-momentum jets. Because protons are composed of quarks and gluons, hard collisions of high-energy protons and antiprotons occur through scattering of gluons from gluons, or gluons from quarks or antiquarks, or quarks or antiquarks from quarks or antiquarks. Two-jet events have been measured in great detail in 630-GeV proton-antiproton collisions at the Superproton Synchrotron (SPS) Collider of the European Organization for Nuclear Research (CERN) near Geneva, Switzerland, and in 1.96-TeV proton-antiproton collisions at the Tevatron Collider at Fermilab in Batavia, Illinois. They are in complete accord with the predictions of quantum chromodynamics and cannot be understood without gluons.

Other implications. The existence of gluons with the properties implied by quantum chromodynamics has additional consequences. The interactions of quarks and gluons specified by quantum chromodynamics suggest the existence of a number of new species of hadrons. The most important of these from the gluon perspective are quarkless mesons, composed entirely of gluons. The simplest of these glueballs, as they are sometimes called, may be searched for in the radiative decay J/ψ → photon + 2 gluons, or in the decays of heavier quarkonium states. Several glueball candidates have been observed in this way, most notably the iota with mass 1440 MeV/c^2. Further experimental work will be required to determine whether the iota or other candidates must be identified as quarkless states. *See* MESON.

Further experimental tests of the properties that have been imputed to gluons now abound. The angular distribution of hadron jets in W-boson + 1-jet events probes the gluon spin. Multijet events produced in high-energy proton-antiproton collisions provide incisive tests of quantum chromodynamics and the nature of the gluon. Tests of the gluon spin are to be had from analysis of the pattern of scaling violations in inelastic lepton scattering, and from

comparisons of the hadronic decay rates of various quarkonium states.

The existence of gluons with the canonical properties vindicates the idea of color gauge symmetry and provides strong encouragement for quantum chromodynamics and, by derivation, for grand unification theories of the strong, weak, and electromagnetic interactions. *See* ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS; GRAND UNIFICATION THEORIES; STANDARD MODEL. C. Quigg

Bibliography. F. E. Close, Hadron spectroscopy (theory), *Int. J. Mod. Phys.*, A20:5156–5163, 2005; R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics*, Cambridge University Press, 1996; S. L. Glashow, Quarks with color and flavor, *Sci. Amer.*, 233(4):38–50, October 1975; G. Kane, *Modern Elementary Particle Physics,* rev. ed., 1993; T. B. W. Kirk and H. D. I. Abarbanel (eds.), *Proceedings of the 1979 International Symposium on Lepton and Photon Interactions at High Energies*, Fermilab, Batavia, Illinois, 1980; Y. Nambu, The confinement of quarks, *Sci. Amer.*, 235(5):48–60, November 1976; A. Watson, *The Quantum Quark*, Cambridge University Press, 2004.

Glutathione

A biological molecule that is a tripeptide comprising the three amino acids glutamate, cysteine, and glycine, and is involved in multiple metabolic processes. While the cysteine residue is linked to the glycine residue through a normal α-amide bond, the glutamate residue is linked to the cysteine residue through an unusual γ-amide bond. Glutathione is found in reduced and oxidized forms (**Fig. 1**). The reduced form (GSH) contains only a single glutathione unit with the side chain of cysteine in the sulfhydryl form. The oxidized form (glutathione disulfide, GSSG) contains two glutathione units covalently attached through a disulfide bridge between the cysteine residues. *See* AMINO ACIDS.

Synthesis. Glutathione is synthesized from free amino acids in two steps, each catalyzed by a different enzyme (**Fig. 2**): [1] The enzyme γ-glutamylcysteine synthase joins one molecule of glutamate and one molecule of cysteine to form γ-glutamyl cysteine. To fuel the formation of γ-glutamyl cysteine, this enzyme hydrolyzes one molecule of adenosine triphosphate (ATP) to adenosine diphosphate (ADP) and inorganic phosphate (P_i). [2] The enzyme glutathione synthase joins one molecule of γ-glutamyl cysteine and one molecule of glycine to form one molecule of glutathione. This enzyme also fuels this reaction by hydrolyzing one molecule of ATP to ADP and P_i. These two steps of glutathione synthesis are also part of the γ-glutamyl cycle. *See* ADENOSINE TRIPHOSPHATE (ATP); ENZYME.

Role in oxidative metabolism. Glutathione plays an important role in regulating the oxidative state inside the cell. Critical to this function is the enzyme glutathione reductase, which catalyzes the conversion of one molecule of GSSG to two molecules of GSH. This enzyme contains a flavin adenine dinucleotide (FAD) electron-transferring cofactor and catalyzes this reaction in two steps: [1] The oxidized form of the enzyme binds one molecule of NADH. Two electrons are transferred from the NADH molecule to the FAD cofactor of the enzyme. This results in the reduced form of the enzyme and NAD^+, which is released. [2] The reduced form of the enzyme binds one molecule of GSSG. The two electrons are transferred from the FAD cofactor of the enzyme to the glutathione disulfide, resulting in the formation of two molecules of GSH, which are released.

Glutathione reductase maintains the cellular levels of GSH in 100-fold excess over the cellular levels of GSSG. This allows GSH to act as an intracellular reducing agent that maintains the sulfhydryl groups on the side chains of cysteine residues in proteins in their normal states. Two molecules of GSH can react with a protein containing an undesired disulfide between two cysteine residues to form a protein with two cysteine residues with correct sulfhydryl side chains and one molecule of GSSG.

Reactions utilizing glutathione. Several other enzymes utilize glutathione for a variety of metabolic processes. One of these enzymes is glutathione peroxidase, which catalyzes the conversion of two molecules of GSH and one organic hydroperoxide to one molecule of GSSG, one molecule of the corresponding alcohol, and one water molecule. This reaction is important in red blood cells, where hydrogen peroxide reacts with lipids in the cell membrane to form lipid peroxides that can rupture the cell. The action of glutathione peroxidase prevents

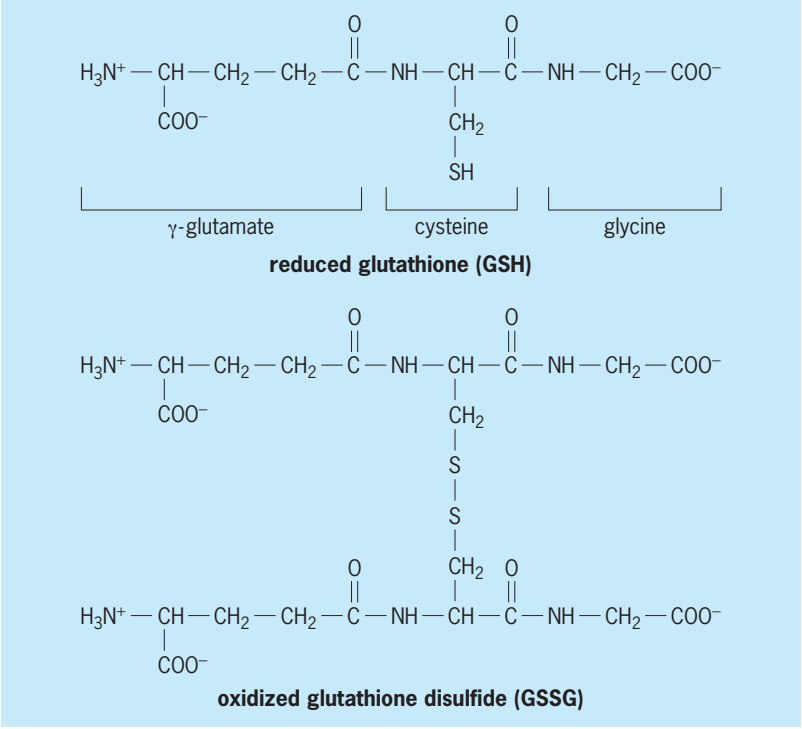

Fig. 1. Reduced and oxidized forms of glutathione.

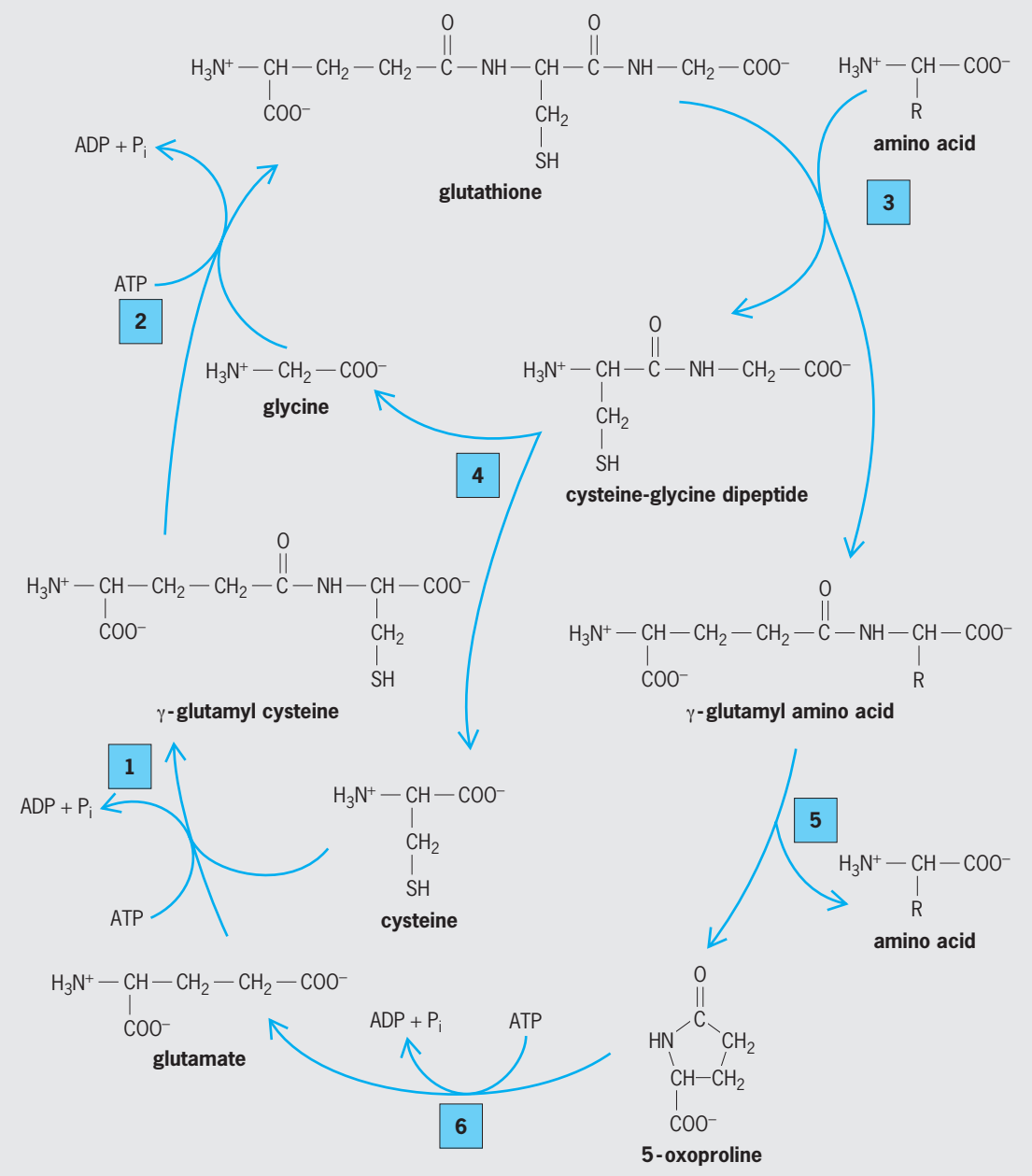

Fig. 2. γ-Glutamyl cycle. See text for description of steps.

cell rupturing by removing lipid peroxides from the cell membrane. *See* LIPID.

Another of these enzymes is glutathione-S-transferase, which catalyzes the covalent attachment of glutathione to a variety of normal metabolites or xenobiotics (chemicals not normally found within cells, such as toxins or drugs) through the sulfhydryl group of its cysteine residue. For example, this enzyme attaches glutathione to a compound called leukotriene A_4, a normal metabolite that is formed from arachidonic acid, to form leukotriene C4, a peptidoleukotriene that is a precursor of molecules involved in anaphylaxis, a severe and sometimes fatal allergic response. *See* ANAPHYLAXIS.

Glutathione also can attach itself nonenzymatically to proteins by forming a disulfide bond between the side chain of the cysteine in the glutathione and the side chain of a cysteine of the protein. Recent research shows that this "glutathionylation" is an important means of regulating the function of these target proteins. For example, glutathionylation has been observed with signal transduction components, such as *ras*, and transcription factors, such as *jun*, involved in promoting cell proliferation.

γ-Glutamyl cycle. Glutathione is also a key component of the γ-glutamyl cycle (Fig. 2), which is important for transporting amino acids from the outside of the cell to the interior of the cell. This cycle comprises six steps: [1, 2] As described above, glutathione is synthesized from glutamate, cysteine, and glycine. [3] The enzyme γ-glutamyl transpeptidase, which is found on the cell membrane,

binds an amino acid from the exterior of the cell and one molecule of glutathione from the interior of the cell. This enzyme transfers the γ-glutamyl group from the glutathione to the amino acid and then releases both the γ-glutamyl amino acid and the cysteine-glycine dipeptide to the cell's interior. [4] The cysteine-glycine dipeptide is broken down into cysteine and glycine by an intracellular protease. [5] The enzyme γ-glutamyl cyclotransferase converts the γ-glutamyl amino acid to the free amino acid and 5-oxoproline. [6] The enzyme 5-oxoprolinase converts the 5-oxoproline to glutamate and hydrolyzes one molecule of ATP to ADP and P_i in the process.

M. Todd Washington

Bibliography. D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 4th ed., 2004; D. Voet and J. G. Viet, *Biochemistry*, 3d ed., 2004.

Glycerol

The simplest trihydric alcohol, with the formula $CH_2OHCHOHCH_2OH$. The name glycerol is preferred for the pure chemical, but the commercial product is usually called glycerin. It is widely distributed in nature in the form of its esters, called glycerides. The glycerides are the principal constituents of the class of natural products known as fats and oils.

Properties. When pure, glycerin is a colorless, odorless, viscous liquid with a sweet taste. It is completely soluble in water and alcohol but is only slightly soluble in many common solvents, such as ether, ethyl acetate, and dioxane. Glycerin is insoluble in hydrocarbons. It boils at 290°C (554°F) at atmospheric pressure and melts at 17.9°C (64.2°F). Its specific gravity is 1.262 at 25°C (77°F) referred to water at 25°C (77°F), and its molecular weight is 92.09. It has a very low mammalian toxicity.

Production. Glycerin was first discovered in 1779 by Carl W. Scheele, who made it by heating olive oil with litharge. Until after World War II, nearly all the glycerin of commerce was produced as a by-product in the manufacture of soap or from the hydrolysis (splitting) of fats and oils. However, a substantial portion of the material made in the United States is now prepared synthetically from propylene.

In the process of soapmaking, called saponification, fat reacts with aqueous sodium hydroxide. The crude product is coagulated by the addition of salt. The acid portion of the fat combines with the sodium hydroxide to form solid soap, and the glycerin liberated in the reaction remains in the salt solution, which is called spent lye.

In the process involving the hydrolysis of fats, they react with water to give the component acids and glycerin. An aqueous solution of the latter is produced, called glycerin sweet water. From this liquid and from spent lye from soapmaking, glycerin is obtained in much the same way. After a preliminary treatment to remove impurities, the water of solution is evaporated under reduced pressure. The residual glycerin is filtered while hot to remove precipitated salts. For most applications, it is necessary to refine it further by fractional distillation under reduced pressure.

Three synthetic routes from propylene to glycerin are used on a large scale. In the first of these, propylene is converted to allyl chloride, which is then treated with aqueous chlorine to make glycerin dichlorohydrins which may be directly hydrolyzed to glycerin. Alternately, the dichlorohydrins are hydrolyzed to epichlorohydrin with a further hydrolysis to glycerin. In a second process, propylene is oxidized in the vapor phase to acrolein. The acrolein is converted to glycerin by successive reactions with hydrogen peroxide, water, and hydrogen. In the third route, propylene is reacted with aqueous chlorine to make propylene oxide. This is isomerized to allyl alcohol, which in turn is reacted with peracetic acid and then with water to make glycerin.

Several grades of glycerin are marketed, including high gravity, dynamite, yellow distilled, USP (U.S. Pharmacopoeia), and CP (chemically pure). USP is water-white and suitable for use in foods, pharmaceuticals, and cosmetics, or for any purpose where the product is designed for human consumption.

Uses. Glycerin is used in nearly every industry. With dibasic acids, such as phthalic acid, it reacts to make the important class of products known as alkyd resins, which are used as coatings and in paints. Because of its valuable emollient and demulcent properties, it is used in innumerable pharmaceutical and cosmetic preparations. It is an ingredient of many tinctures, elixirs, cough medicines, and anesthetics. It is a basic medium for toothpaste.

In foods, it is an important moistening agent for baked goods and is added to candies and icings to prevent crystallization. It is used as a solvent and carrier for extracts and flavoring agents and as a solvent for food colors.

Because of its humectant properties, it is sprayed on tobacco before it is processed to prevent crumbling and is added to adhesives and glues to keep them from drying too fast. Many specialized lubrication problems have been solved by using glycerin or glycerin mixtures.

Many millions of pounds of glycerin are used each year to plasticize various materials. As much as 15% is added to cellophane to render it pliable. It is included in meat casings and special types of paper for the same purpose. Sheets and gaskets made from ground cork are plasticized with glycerin.

Of its chemical derivatives, the esters of glycerin are the most important. Nitroglycerin (glyceryl trinitrate) is used in the manufacture of dynamites and propellants. The mono- and diesters of higher fatty acids can be formed by direct reaction of glycerin with the acid or by a transesterification reaction of a glyceride with glycerin. These esters are used as emulsifiers in foods and preparation of baked goods and for modification of alkyd resins. *See* ALCOHOL; FAT AND OIL; FAT AND OIL (FOOD); POLYOL.

Philip H. Cook

Bibliography. J. A. Kent (ed.), *Riegel's Handbook of Industrial Chemistry*, 9th ed., 1992; *Kirk-Othmer Encyclopedia of Chemical Technology*, 3d ed., vol. 11, 1998; C. S. Miner and N. N. Dalton (eds.), *Glycerol*, ACS Monogr. 117, 1953.

Glycogen

The primary reserve polysaccharide of the animal world. It is found in the muscles and livers of all higher animals, as well as in the cells of lower animals. Because of its close relationship to starch, it is often called animal starch, although glycogen is found in some lower plants, fungi, yeast, and bacteria. A polysaccharide similar to glycogen was isolated in one case from a higher plant, Golden Bantam sweet corn (*Zea mays*). *See* LIVER; STARCH.

Properties. Glycogen is a nonreducing, white, amorphous polysaccharide which dissolves readily in cold water, forming an opalescent, colloidal solution. It gives a reddish brown color with iodine, is precipitated by alcohol, and has a specific rotation $[\alpha]_D^{20}$ of approximately $+200°$. It is very resistant to the action of alkalies and may be prepared by boiling liver or muscle tissue in 30% potassium hydroxide to destroy the proteins, and precipitating the glycogen with ethyl alcohol. *See* OPTICAL ACTIVITY.

The molecular weight of glycogen is usually very high, and it varies with the source and the method of preparation; molecular weights of the order of $1–20 \times 10^6$ have been reported.

In its biochemical reactions, glycogen is similar to starch. It is attacked by the same plant amylases that attack starch, and like starch, it is degraded to maltose and dextrins. Both glycogen and starch are broken down by animal or plant phosphorylase enzyme in the presence of inorganic phosphate with the production of α-D-glucose-1-phosphate. *See* CARBOHYDRATE METABOLISM.

Molecular structure. Chemical studies, based on methylation and periodate oxidation procedures, show glycogen to possess a branched structure similar to the amylopectin starch fraction. The molecules of both these polysaccharides consist of chains of D-glucose residues joined by α-1,4 linkages, having similar chains attached through α-1,6 linkages at the branch points (see **illus**.). Depending on the source, the average chain length of a branch (which is the average length in glucose units of the outer and inner branches) in a glycogen molecule is 11–18 D-glucopyranose units. In amylopectin the average chain length of a branch is 22–27.

Synthesis. About 1940 C. F. Cori and G. T. Cori first showed that glycogen could be synthesized in the test tube and that two enzymes, muscle phosphorylase and a branching enzyme which is present in animal tissues, are required for the formation of this important polysaccharide. The muscle phosphorylase catalyzes a stepwise transfer of α-D-glucosyl units from α-D-glucose-1-phosphate to a nonreducing end of a "primer" (acceptor substrate). Synthesis of polysaccharide from α-D-glucose-1-phosphate

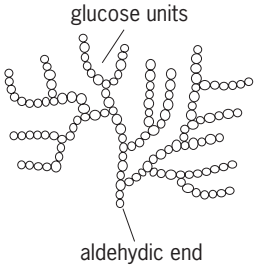

Branched structure of glycogen.

does not occur unless a small amount of starch, glycogen, or dextrin is present as a priming agent. In the presence of the primer, the enzyme adds D-glucose units to a preexisting polysaccharide chain, forming long linear chains joined through α-1,4 linkages. The reaction can be written as reaction (1).

$$x\ \alpha\text{-D-Glucose-1-phosphate} + \underset{\text{Acceptor}}{(\text{D-glucose})_n} \rightleftharpoons \underset{\text{Glycogen chain}}{(\alpha\text{-1,4-D-glucose})_{n+x}} + x\ \text{phosphate} \quad (1)$$

The second enzyme (branching factor) has the ability to unite such chains through α-1,6 linkages, resulting in the formation of a highly branched glycogen structure. It was believed that the combined action of the two enzymes was responsible for the synthesis of glycogen in the animal body.

In 1957 L. F. Leloir and his coworkers obtained an enzymic preparation from liver and other animal tissues that catalyzed the transfer of D-glucose from uridine diphosphate D-glucose to an acceptor (glycogen) according to reaction (2).

$$x\ \text{Uridine diphosphate-D-glucose} + \underset{\text{Acceptor}}{(\alpha\text{-1,4-D-glucose})_n} \longrightarrow \underset{\text{Glycogen chain}}{(\alpha\text{-1,4-D-glucose})_{n+x}} + x\ \text{uridine diphosphate} \quad (2)$$

This reaction resembles that of the phosphorylase enzyme, but its equilibrium starting with uridine diphosphate D-glucose is more favorable for synthesis of glycogen than that starting with α-D-glucose-1-phosphate. Since the product obtained is branched, the enzyme preparations must contain another enzyme (6-glucosyltransferase), which transfers α-1,4-glucosyl chains produced by the first enzyme (glycogen glucosyltransferase) to form a branched molecule. The existence of two mechanisms of glycogen formation raises the question of which of the two mechanisms operates in the animal body. Evidence indicates that synthesis of glycogen in the body occurs through a transfer of D-glucose units from uridine diphosphate D-glucose by an enzyme, glycogen glycosyltransferase, whose function is entirely synthetic. When the chains become about 10 units long, portions are transferred to other chains forming α-1,6-glycosyl linkages by a branching enzyme. The function of phosphorylase is considered to be concerned chiefly with degradation of glycogen.

Metabolic pathways. The metabolic formation of glycogen from glucose in the liver is frequently termed glycogenesis. In fasted animals, glycogen formation can be induced by the feeding, not only of materials that can be hydrolyzed to glucose and other monosaccharides, such as fructose, but also of various other materials. A number of L-amino acids, such as alanine, serine, and glutamic acid, upon deamination in the liver give rise to substances, such as pyruvic acid and α-ketoglutaric acid, that can be converted in the liver to glucose units which are subsequently converted to glycogen. Furthermore, substances such as glycerol derived from fats, dihydroxyacetone, or lactic acid can all be utilized for glycogen synthesis in the liver. Such noncarbohydrate precursors are termed glycogenic compounds. The process of glycogen formation from these precursors is known as glyconeogenesis. The term glycogenolysis is used to connote glycogen breakdown. *See* POLYSACCHARIDE. William Z. Hassid

Bibliography. E. D. Atkins, *Polysaccharides*, vol. 8, 1986; R. W. Stoddart (ed.), *The Biosynthesis of Polysaccharides*, 1984.

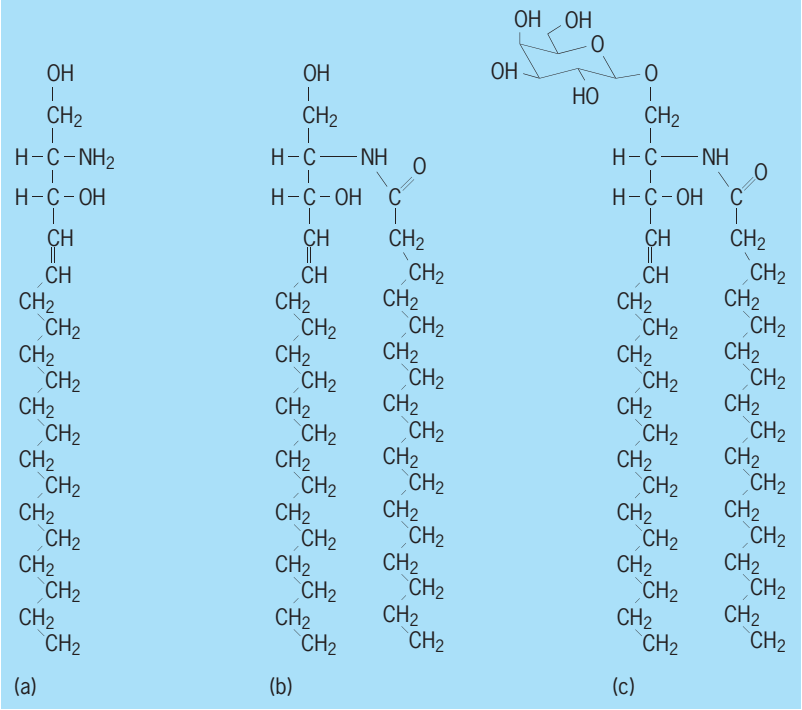

Components of glycosphingolipids. (*a*) Sphingosine. (*b*) Ceramide. (*c*) Galactocerebroside.

Glycolipid

One of a class of compounds having solubility properties of a lipid and containing one or more molecules of a covalently attached sugar.

Glycosphingolipids, the most abundant and structurally diverse type of glycolipids in animals, are glycosides of ceramide, a fatty acid amide of the amino alcohol sphingosine (see **illus.**). Cerebrosides are monosaccharide glycosides of ceramide, which generally contain either galactose or glucose. Galactosyl ceramide is enriched in brain tissue and is a major component of the myelin sheaths around nerves. Psychosine, *O*-sphingosyl galactoside, is the deacylated product of galactosyl ceramide. Glucosyl ceramide is present in the cell membranes of many cell types and is abundant in serum. Lipidoses (inherited lysosomal lipid storage diseases that lead to abnormal development and mental retardation) result from genetic defects in enzymes that degrade glycosphingolipids. For example, in Gaucher's disease, glucocerebrosidase, an enzyme that removes the glucose from glycosyl ceramide, is defective or missing.

Larger, neutral glycosphingolipids containing more than one sugar include lactosyl ceramide, abundant in leukocyte membranes; globosides; and other oligosaccharyl ceramides, some of which are important antigens defining blood groups. Gangliosides are oligosaccharyl ceramides, abundant in brain, spleen, erythrocytes, liver, and kidney, that contain glucose, galactose, *N*-acetylglucosamine, and sialic acids. Sialic acids are *N*-acyl derivatives of neuraminic acid, a nine-carbon sugar with a carboxylic acid group:

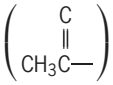

They occur at the terminus of many types of glycoconjugates, and display enormous structural diversity, with over 30 different structures known. Sulfatides are glycosphinglipids of various types that have sulfate esters, generally attached to galactose.

Glycosphingolipids carry blood group antigens and define tumor-specific or developmental antigens. In addition, they serve as receptors for many microorganisms and toxins, as modulators of cell surface receptors that mediate cell growth, and as mediators of cell adhesion. *See* ANTIGEN; CELL ADHESION.

Glycosyl phosphatidylinositols are a class of glycolipids that serve as membrane anchors for a multitude of proteins in organisms ranging from yeast to protozoa to humans. Glycosyl phosphatidylinositols consist of a molecule of phosphatidlyl inositol attached to a nonacetylated glucosamine and a trimannosyl core linked via an ethanolamine phosphodiester bridge to the C-termini of proteins. Furthermore, glycosyl phosphatidylinositol–core structures can have many different modifications, depending upon the protein and cell type. Lipophosphoglycans are glycosyl phosphatidylinositols attached to large polysaccharide structures that coat the surfaces of many parasitic protozoa, such as *Leishmania donovani*, the causative agent of visceral leishmaniasis (kala azar), afflicting millions of individuals each year. Lipophosphoglycans appear to protect these organisms from host defenses.

Mannosylphosphoryl dolichol, glucosylphophoryl dolichol, and oligosaccharyl phosphoryl dolichols are glycolipids with sugars attached to large polyisoprenoids (for example, C_{95} hydrocarbons) by phosphate esters. Dolichols are structurally related

to cholesterol. Saccharylphosphoryl dolichols serve as important biosynthetic intermediates in the assembly of both asparagine-linked glycoproteins and glycosyl phosphatidylinositols. *See* GLYCOPROTEIN.

Glycosyl glycerides are glycolipids that have a structure analogous to phospholipids. They are the major glycolipids of plants and microorganisms but are rare in animals.

Bacteria produce a wide variety of glycolipids not easily categorized. Examples include fatty acid esters of carbohydrates, such as cord factor, an ester of the disaccharide trehalose with two molecules of the complex fatty acid, mycolic acid. Cord factor is a toxic component of the waxy capsular material of virulent strains of *Mycobacterium tuberculosis*, the causative agent of tuberculosis. Mycosides, glycolipids that are also found in tubercle bacilli, comprise long-chain, highly branched, hydroxylated hydrocarbon terminated by a phenol group, with the sugar glycosidically attached to the phenolic hydroxyl. *See* LIPID; SPHINGOLIPID; TUBERCULOSIS. Gerald W. Hart

Glycoprotein

A compound in which carbohydrate (sugar) is covalently linked to protein. The carbohydrate may be in the form of monosaccharides, disaccharides, oligosaccharides, or polysaccharides, and is sometimes referred to as glycan. The sugar may be linked to sulfate or phosphate groups. One to two hundred glycan units may be present in different glycoproteins. Therefore, the carbohydrate content of these compounds varies markedly, from 1% (as in the collagens), to 60% (in certain mucins), to >99% (in glycogen). *See* COLLAGEN; GLYCOGEN.

Glycoproteins are ubiquitous in nature, although they are relatively rare in bacteria. They occur in cells, in both soluble and membrane-bound forms, as well as in the intercellular matrix and in extracellular fluids, and include numerous biologically active macromolecules. A particularly rich source of glycoproteins is human serum, in which, of the more than 60 proteins that have been identified, only two—albumin and prealbumin—do not contain sugar. Another rich source is hen egg white, where probably all proteins, apart from lysozyme, have carbohydrate attached to their molecules. A number of glycoproteins are produced industrially by genetic engineering techniques for use as drugs; among them are erythropoietin, interferons, colony stimulating factors, and blood-clotting factors. *See* GENETIC ENGINEERING.

Composition and structure. Of the numerous monosaccharides found in nature, only a small number are common constituents of glycoproteins. In most glycoproteins, the carbohydrate is linked to the polypeptide backbone by either N- or O-glycosidic bonds. A different kind of bond is found in glycoproteins that are anchored in cell membranes by a special carbohydrate-containing compound, glycosylphosphatidylinositol, which is attached to the C-terminal amino acid of the protein. A single glycoprotein may contain more than one type of carbohydrate-peptide linkage. N-linked units are typically found in plasma glycoproteins, in ovalbumin, in many enzymes (for example, the ribonucleases), and in immunoglobulins. O-linked units are found in mucins; collagens; and proteoglycans (typical constituents of connective tissues), including chondroitin sulfates, dermatan sulfate, and heparin. *See* ALBUMIN; ENZYME; IMMUNOGLOBULIN.

Microheterogeneity. Within any organism, all molecules of a particular protein are identical. In contrast, a variety of structurally distinct carbohydrate units are found not only at different attachment sites of a glycoprotein but even at each single attachment site—a phenomenon known as microheterogeneity. For instance, ovalbumin contains one glycosylated amino acid, but over a dozen different oligosaccharides have been identified at that site, even in a preparation isolated from a single egg of a purebred hen.

Biosynthesis. The protein moieties of all glycoproteins are synthesized, on polyribosomes, in the same manner as nonglycosylated proteins. Formation of the carbohydrate units is catalyzed by glycosyltransferases, enzymes that transfer sugars from suitable donors to particular acceptors, and is not directed by a genetically specified template. Over 100 distinct glycosyltransferases, located in the rough endoplasmic reticulum of the cell and in the Golgi apparatus, participate in the biosynthesis of glycoproteins, each forming an individual glycosidic bond. In addition, many other enzymes are required for completion of the carbohydrate units, such as the attachment of sulfate or phosphate. O-linked units are formed by sequential incorporation of one monosaccharide at a time from a sugar nucleotide, first to a hydroxyl of a particular amino acid residue of a polypeptide, and then to a hydroxyl of the protein-bound sugars. *See* ENDOPLASMIC RETICULUM; GOLGI APPARATUS.

N-linked oligosaccharides are synthesized in a different, more complicated sequence of reactions. First, a single precursor oligosaccharide is formed, linked by pyrophosphate to a long-chain lipid. The oligosaccharide is then transferred from the lipid to an amide of an asparagine residue of a growing polypeptide chain. This occurs as the polypeptide emerges from the ribosome and enters the lumen of the rough endoplasmic reticulum. Subsequently, the polypeptide-linked oligosaccharide undergoes a series of processing reactions which begin in the rough endoplasmic reticulum and continue and end in the Golgi apparatus, eventually yielding the N-linked oligomannose units.

Catabolism. One pathway of glycoprotein degradation in the body starts with the stepwise hydrolysis of the major part of the glycans by the sequential action of lysosomal glycosidases, followed by the enzymatic disassembly of the protein and cleavage of the carbohydrate-peptide linkages. An alternative pathway starts with proteolysis of the polypeptide backbone and requires the participation of

endoglycosidases, which release intact glycans from their linkages with amino acids. Individuals that lack one of the glycosidases or other enzymes (for example, sulfatases) needed for glycoprotein degradation suffer from serious, often fatal diseases. Examples are the mucopolysaccharidoses, such as the Hunter-Hurler syndrome. In these diseases, undegraded glycosaminoglycans accumulate in the tissues of the afflicted individuals, resulting in developmental abnormalities, skeletal deformations, and mental retardation. Another example is the I-cell disease, caused by lack of the first enzyme required for the formation of the mannose-6-phosphate on lysosomal enzymes, and thus the deficiency in the lysosomes of many of the hydrolases required for degradation of carbohydrate-containing compounds.

Functions. The carbohydrate may modify, in a variety of ways, the properties of the protein to which it is attached. Commonly, it changes the physicochemical properties of the protein, such as electrical charge, solubility, or viscosity, and increases its thermostability. Frequently, it protects the polypeptide backbone against proteolytic digestion. It is also required for the correct folding of proteins during their biosynthesis and for the stabilization of the three-dimensional structure of the mature glycoprotein.

Effects on biological activity. In a small number of glycoproteins the biological activity of the protein part is known to be affected by, or to depend on, the presence of the carbohydrate. A well-documented case is human chorionic gonadotropin, which loses its hormonal activity upon removal of its N-glycans. Another example is tissue plasminogen activator, a protease that induces clot lysis and is used clinically for this purpose. Two molecular species of the enzyme are found in tissues that are structurally identical except for the number of N-linked glycans (two or three); they differ markedly in their enzymatic activity. In general, however, the carbohydrate has no effect on the catalytic activity of glycoprotein enzymes, nor for example on the carbohydrate-binding properties or hemagglutinating activity of glycoprotein lectins.

Often, the carbohydrate affects the immunological properties of the protein to which it is attached, mainly by introducing novel antigenic determinants. This is illustrated by the A, B, and O blood group glycoproteins that are present on human erythrocytes, as well in other tissues, such as gastric mucosa and salivary glands. *See* BLOOD GROUPS.

Recognition determinants. Carbohydrates of glycoproteins can act as recognition determinants. Thus, infection by influenza virus is initiated by the attachment of the virus to sialic acid residues of the cell-surface glycoproteins. Many other systems in which carbohydrates play a similar role are known. For example, bacteria such as *Escherichia coli* adhere to the respiratory, intestinal, or urinary tract by binding to mannose residues on glycoproteins present on the epithelial cells lining these organs, a step prerequisite for the initiation of infection. Also, the binding of leukocytes to endothelial cells lining the blood vessels, crucial for the control of the movement of the leukocytes in the body, and for the initiation of inflammation, is mediated by complex oligosaccharides present on both the leukocytes and the endothelial cells. *See* CARBOHYDRATE; CELLULAR IMMUNOLOGY; GLYCOSIDE; MONOSACCHARIDE; OLIGOSACCHARIDE; POLYSACCHARIDE; PROTEIN. Nathan Sharon

Bibliography. R. C. Hughes, *Glycoproteins*, 1983; N. Sharon and H. Lis, Carbohydrates in cell recognition, *Sci. Amer.*, 268(1):74–82, 1993.

Glycoside

A large important class of sugar derivatives in which the sugar is combined with a nonsugar. In their cyclic forms, monosaccharides (simple sugars) possess one carbon (C) atom (the anomeric carbon) that is bonded to two oxygen (O) atoms; one oxygen atom forms a part of the ring, whereas the other is outside the ring (exocyclic) and is part of a hydroxyl (OH) group. If the oxygen atom of the anomeric hydroxyl group becomes bonded to a carbon atom, other than that of a carbonyl (C═O) group, the resulting compound is a glycoside. A glycoside thus consists of two parts (**Fig. 1***a*): the sugar (glycosyl) unit, which provides the anomeric carbon, and the moiety (the aglycon), which is the source of the exocyclic oxygen and carbon atoms of the glycosidic linkage. Such compounds frequently are referred to as O-glycosides to distinguish them from analogs having a sulfur (thio- or *S*-glycosides), nitrogen (amino- or N-glycosides), or carbon (anomalously called C-glycosides) as the exocyclic atom on the anomeric carbon. *See* HYDROXYL; MONOSACCHARIDE.

The formation of glycosides is the principal manner in which monosaccharides are incorporated into more complex molecules. For example, lactose (Fig. 1*b*), the most abundant disaccharide in mammalian milk, has a glycosidic bond involving

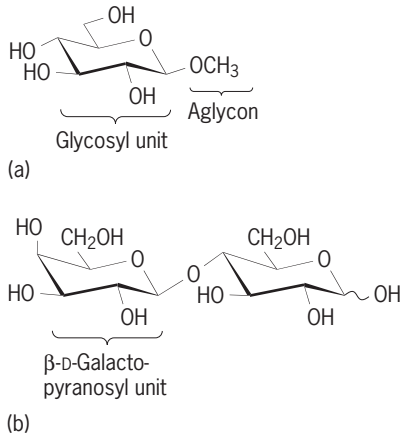

Fig. 1. Structural formulas of two glycosides. (*a*) Methyl *β*-D-glucopyranoside. (*b*) Lactose, 4-*O*-*β*-D-galactopyranosyl-D-glucopyranose; the wavy bond indicates that the group may have various orientations in space.

the anomeric carbon of D-galactose and the C-4 hydroxyl of D-glucose. The anomeric carbon atom can exist in either of two stereoisomeric configurations. In the D-series of carbohydrates a glycoside is designated as having the α configuration if the exocyclic oxygen is attached to the anomeric carbon by a bond that projects downward, or the β configuration if that bond projects upward, when the ring is oriented with the ring oxygen away from the viewer (Fig. 1*b*). In the L-series the orientation is reversed. The anomeric configuration is of immense importance to the chemistry and biochemistry of glycosides. For example, the principal structural difference between cellulose and amylose is that cellulose is β-glycosidically linked whereas amylose is α-linked. Humans are able to digest amylose but are unable to utilize cellulose for food. *See* CELLULOSE; LACTOSE; STEREOCHEMISTRY.

Chemical synthesis. The transformation of a monosaccharide into a glycoside involves the conversion of a hemiacetal (or hemiketal) into an acetal (or ketal). Such reactions normally are catalyzed by an acid and may involve a temporary loss of chirality (dissymmetry) at the anomeric carbon. Glycosides are susceptible to hydrolysis in aqueous acid, but they are stable under alkaline conditions. The classical (Fischer) method for synthesizing glycosides from unprotected (that is, underivatized) monosaccharides and low-molecular-weight alcohols involves heating the sugar and alcohol in the presence of an acid catalyst under anhydrous conditions. *See* ACETAL.

The simplicity of the method belies the complexity of the reactions; these give an equilibrium mixture of α- and β-pyranosides (having six-membered rings) and α- and β-furanosides (having five-membered rings) in a ratio that is dependent upon the monosaccharide, the alcohol, and the precise conditions of reaction. The furanosides are formed more rapidly, but pyranosides are more stable; thus, in general, the principal products at equilibrium are the pyranosides. The restrictions inherent in this method limit its synthetic usefulness.

A much more common method is the Königs-Knorr reaction, which utilizes a carbohydrate derivative in which the anomeric hydroxyl group has been replaced by a halogen atom—usually bromine (Br)—and all of the other hydroxyl groups of the sugar have been converted to esters or ethers, thereby making them unavailable for glycoside formation. This fully protected glycosyl halide is capable of forming a glycosidic bond with virtually any compound having a free hydroxyl group. A large number of variations of this method have been devised; most use a Lewis acid, such as a salt of silver or mercury, as catalyst. This method affords a relatively high yield, and predictability as to the anomeric configuration of the glycoside. When the glycopyranosyl halide is protected with (participating) ester groups, the phenomenon of anchimeric assistance (neighboring-group effect) becomes operative, and the major product is the so-called 1,2-*trans*-glycopyranoside in which the functional group on C-2 is trans to the aglycon [reaction (1), where R is an alkyl or aryl group].

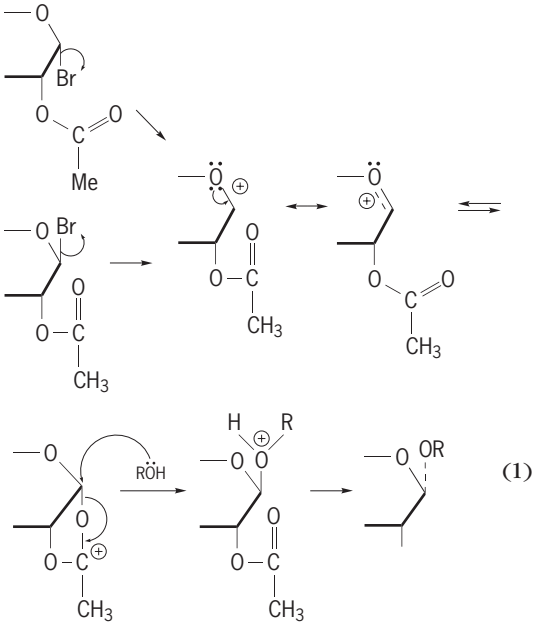

When the glycopyranosyl halide is protected with (nonparticipating) ether groups, the preferred product is the 1,2-*cis*-glycopyranoside. Thus, regardless of the glycosyl halide used, either an α- or a β-glycopyranoside can be formed as the major product by the appropriate selection of the protecting groups. A very powerful extension of this method employs *n*-pentenyl glycosides (specifically, 4-pentenyl), which are activated in the presence of halogen cations (halonium ions), usually supplied by iodonium collidine perchlorate (IDCP). A key feature is that if the *n*-pentenyl glycoside is protected by ester substituents it reacts much more slowly than one protected by ether substituents. The esterified glycoside is said to be disarmed and the etherified one armed with respect to their reactivity with halonium ions. Thus, when both are present in a reaction mixture the ether-protected (armed) 4-pentenyl glycosides react preferentially.

The sugars can be joined together by a glycosidic bond [reaction (2), where Ac is acetyl

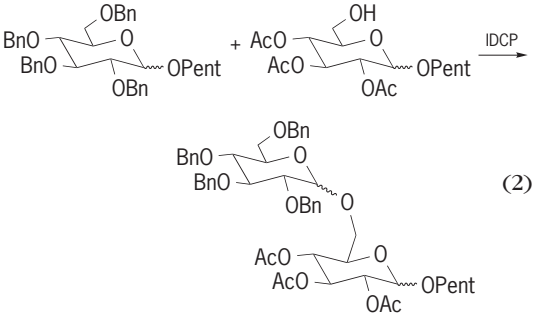

[$H_3C—C(=O)—$], Bn is benzyl ($—CH_2—C_6H_5$), and Pent is 4-pentenyl]. A 4-pentenyl glycoside, fully

substituted with ether groups, is reacted in the presence of IDCP with another 4-pentenyl glycoside having one free hydroxyl group and all the others esterified. The disarmed esterified glycoside reacts slowly with the free hydroxyl group, whereas the armed etherified glycoside reacts rapidly. Therefore, almost all of the reaction results in the replacement of the 4-pentenyloxy group of the armed glycoside by the free hydroxyl group of the disarmed glycoside.

Another ingenious method for glycosidic bond formation between sugar units uses an insoluble polymer as the site of reaction, and so is described as a solid-phase synthesis. Solid-phase syntheses of oligopeptides and oligonucleotides have been used for many years, but this method is the most practical application for oligosaccharides. These are formed by a sequence of couplings. The method begins by attaching a glycal (a monosaccharide having a double bond between carbons 1 and 2, and no hydroxyl group on either of those carbons) to the solid polymer; all of the free hydroxyl groups of the glycal are protected, and the double-bonded carbon atoms are incorporated into a three-membered cyclic ether (an epoxide). The use of 3,3-dimethyldioxirane effects this with high trans-stereoselectivity and yield [reaction (3), where R is benzyl ($—CH_2—C_6H_5$), R′ is

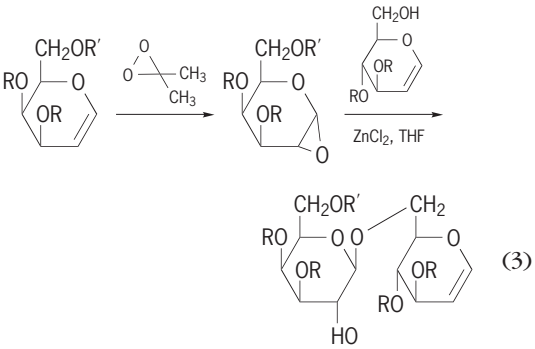

a polymer, and THF is tetrahydrofuran]. The epoxide is reacted in tetrahydrofuran, in the presence of zinc chloride ($ZnCl_2$), with another partially protected glycal having one free hydroxyl group. The epoxide acts as the glycosyl donor and is opened by a back-side attack by the free hydroxyl of the glycosyl acceptor, thereby forming a glycoside and concurrently generating a free hydroxyl group at C-2. This valuable outcome can afford an acceptor for branched-oligosaccharide synthesis, or the hydroxyl group can be protected by esterification or etherification and the glycal epoxidized and used for further glycosidation. *See* ESTER; ETHER; OLIGOSACCHARIDE; ORGANIC SYNTHESIS; REACTIVE INTERMEDIATES.

Naturally occurring glycosides. A very large number of glycosides exist in nature, many of which possess important biological functions. In many of these biologically important compounds the carbohydrate portion is essential for cell recognition, the terminal sugar units being able to interact with specific receptor sites on the cell surface.

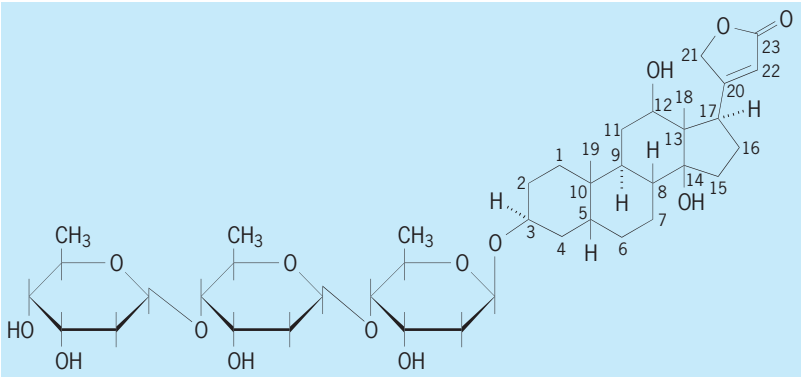

Fig. 2. Structural formula of digoxin, a cardiac glycoside.

One class of naturally occurring glycosides is called the cardiac glycosides because they exhibit the ability to strengthen the contraction of heart muscles. These cardiotonic agents are found in both plants and animals and contain complex aglycons, which are responsible for most of the drug action; however, the glycoside may modify the biological activity. The best-known cardiac glycosides come from digitalis and include the drug digoxin (**Fig. 2**). *See* DIGITALIS.

Glycosidic units frequently are found in antibiotics. For example, daunomycin and adriamycin, antibiotics used as antitumor agents, incorporate an aminodeoxy sugar unit glycosidically linked to a polycyclic aglycon (**Fig. 3***a*); and the important drug erythromycin A possesses two glycosidically linked sugar units (Fig. 3*b*). *See* ANTIBIOTIC.

Perhaps the most ubiquitous group of glycosides in nature is the glycoproteins; in many of them carbohydrates are linked to a protein by *O*-glycosidic bonds to L-serine, L-threonine, 5-hydroxy-L-lysine, or 4-hydroxy-L-proline. These glycoproteins include many enzymes (for example, the human intestinal disaccharidases), hormones, such antiviral compounds as interleukin-2, and the so-called antifreeze glycoproteins found in the sera of fish from very cold marine environments. *See* AMINO ACIDS; ANTIFREEZE (BIOLOGY); CARBOHYDRATE; ENZYME; GLYCOPROTEIN; HORMONE.

Glycolipids are a very large class of natural glycosides having a lipid aglycon. These complex glycosides are present in the cell membranes of microbes, plants, and animals. Animal glycolipids usually have ceramide derivatives as the aglycon. A ceramide is a general term for a component that consists of sphingosine (or one of its analogs) to which a long-chain fatty acid is connected by an amide linkage. If the glycosyl unit incorporates one or more *N*-acylneuraminic acid residues, the glycoside is called a ganglioside (Fig. 3*c*). Plant and microbial glycolipids are more diverse and include glycosides having *myo*-inositol derivatives, hydroxy fatty acids, or glycerol derivatives as the aglycon. Although their biological function is poorly understood, some are useful diagnostic antigens for certain types of lung

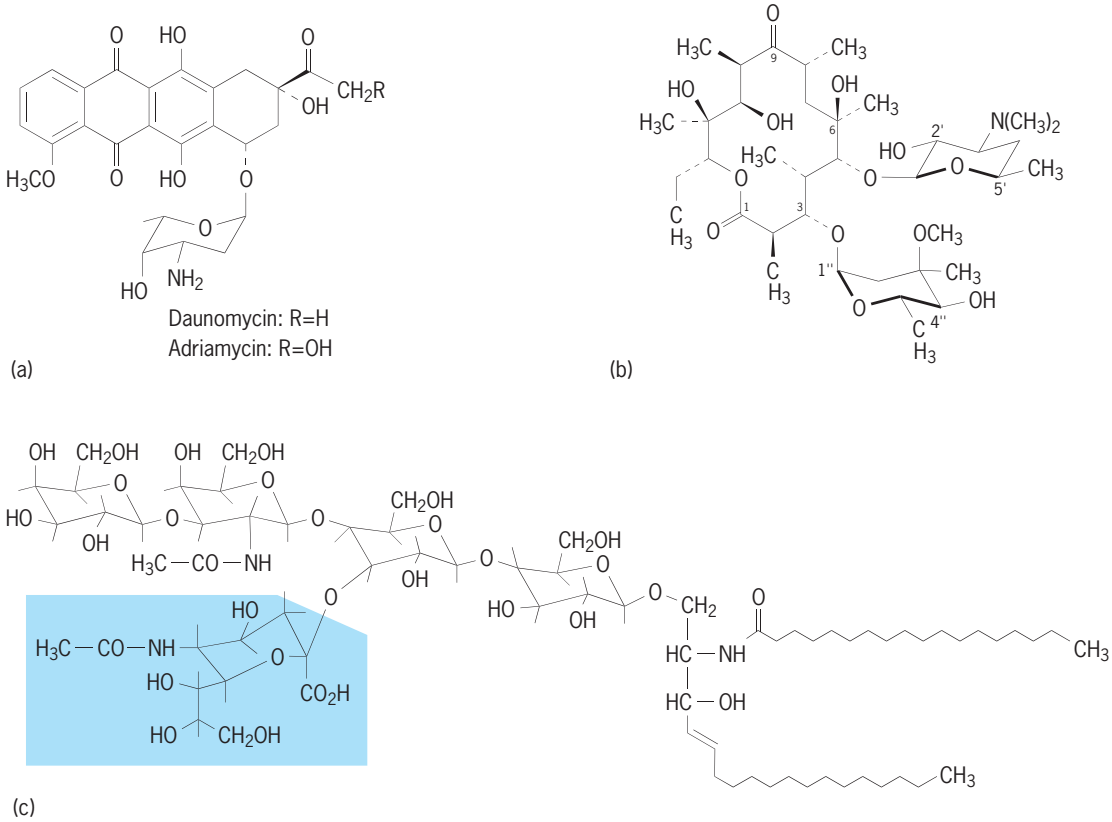

Fig. 3. Structural formulas of some naturally occurring glycosides. (*a*) Anthraquinonyl glucosaminosides (anthracycline antibiotics). (*b*) Erythromycin A, a macrolide (large-ring) glycoside antibiotic. [Me = methyl (—CH_3)]. (*c*) *N*-acetylneuraminic acid, a ganglioside; shaded area is the ceramide.

and colon carcinomas. *See* GLYCOLIPID; IMMUNOASSAY; LIPID. G. W. Hay

Bibliography. R. W. Binkley, *Modern Carbohydrate Chemistry*, 1988; A. E. Bochkov and G. E. Zaikov, *Chemistry of the O-Glycosidic Bond, Formation and Cleavage*, 1979; W. Pigman and D. Horton, *The Carbohydrates, Chemistry and Biochemistry*, 1972; L. G. Wade, Jr., *Organic Chemistry*, 4th ed., 1998.

Gnathostomata

A superclass of the subphylum Vertebrata (Craniata). Gnathostomes are animals with jaws, involving a vertical biting that developed from modified gill arches. They are further characterized by a notochord (an elongated dorsal cord of cells that is the primitive axial skeleton in all chordates) that is present in the ontogeny of all lineages but replaced by vertebral centra (the main bodies of vertebrae) in most taxa; early fossil forms with a bony exoskeleton that was lost in the higher lineages; limb girdles (secondarily lost in some forms) supporting paired appendages in all but the most primitive taxa; myelinated neurons; an adaptive immune system; intrinsic eye muscles; a sperm duct linked to the urinary system; and a distinct cerebellum. Some of the characters were carried over from the primitive superclass Agnatha, but jaws and paired appendages are unique to Gnathostomata. *See* ANIMAL EVOLUTION; JAWLESS VERTEBRATES; VERTEBRATA.

The gnathostomes are divided into three major grades, the hierarchical classification of which is shown below (the taxa are featured elsewhere in the Encyclopedia):

Superclass Gnathostomata
- Grade Placodermiomorphi
  - Class Placodermi
- Grade Chondrichthiomorphi
  - Class Chondrichthyes
- Grade Teleostomi
  - Class: Acanthodii
    - Actinopterygii
    - Sarcopterygii

In the past, gnathostomes were recognized as two superclasses, Pisces and Tetrapoda, which constitute roughly the aquatic and the terrestrial vertebrates, respectively. The term Pisces is no longer recognized in classification, but "jawed fishes" is a convenient term for all taxa from placoderms (extinct armored fishes) to coelacanths (lobefin fishes) and lungfishes (sarcoptergyian fishes), with about 27,900 species. The remainder of the sarcopterygians are tetrapods, with about 26,750 species. The earliest gnathostomes, the placoderms, appeared in the fossil record no later than the Middle Silurian, about 425 million

years before the present. *See* PISCES (ZOOLOGY); SARCOPTERYGII; TETRAPODA. Herbert Boschung

Bibliography. J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Gnathostomulida

A phylum of microscopic marine worms related to Rotifera and Micrognathozoa, mainly characterized by complex cuticular structures in the pharynx and a monociliated skin epithelium. Discovered in the 1920s on the German coast but not described until 1956, Gnathostomulida have since been reported from many sheltered sandy shores around the world. With fewer than 100 known species, this is one of the smallest animal phyla.

Morphology. Gnathostomulids range from 0.01 to 0.14 in. (0.3 to 3.5 mm) in length. They are worm-shaped, cylindrical, and semitransparent (or bright red), and sometimes have the external divisions of head and tail (see **illustration**). The skin is a one-layered epithelium that is completely monociliated; each of the polygonal epidermal cells bears only one cilium which can be up to 20 micrometers long. Parenchyma is poorly developed, and excretory organs may be present as three paired groups of protonephridia. The nervous system is largely basiepithelial, consisting of an unpaired buccal and unpaired frontal ganglion (brain) from which one to three pairs of longitudinal nerves originate. The sensory system usually consists of one or two pairs of simple bristles and three or four pairs of compound bristles (frontally and laterally) and a row of stiff cilia (dorsally on the head). In some genera, ciliary pits and spiral ciliary organs of unknown function are present in the anterior head region.

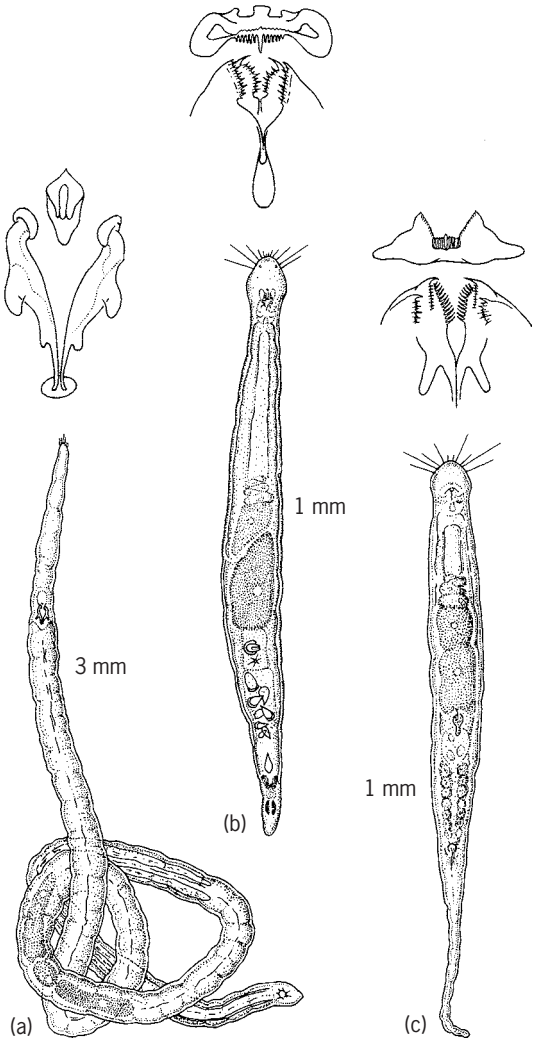

Representative Gnathostomulida showing details of basal plate and jaws. (*a*) *Haplognathia simplex*. (*b*) *Austrognathia riedli*. (c) *Gnathostomula mediterranea*.

Great diversity exists in the foregut area. The mouth, located ventrally in the head (in *Haplognathia* far from the anterior end), is equipped with complex cuticularized structures: sometimes with a jugum, a cartilaginous structure, in the upper lip, mostly with paired jaws and an unpaired basal plate in the lower lip area. The basal plate occurs in many different forms throughout the species but always bears lamellae or teeth in its center. The large mouth cavity is surrounded by a complex muscle apparatus (consisting of a dozen pairs of individual muscles) which becomes progressively more compact from *Haplognathia* to *Gnathostomula* types. The same is true for the jaws which vary from simple forceps types (*Haplognathia*) to complicated lamellar snap jaws with three rows of up to 60 teeth (see illustration). The midgut is simple; large cells surround a gut cavity; a permanent anus is lacking.

All gnathostomulids are hermaphrodites. The reproductive system consists of an unpaired dorsal ovarium, and paired (in most genera) or unpaired (in *Austrognathia*) caudolateral testes in the same specimen. In the majority of the genera, the male apparatus includes a simple tubular copulatory stylet, composed of a concentrically arranged bunch of cuticular rods and accompanied by glands. This type of male organ is always associated with a female organ consisting of a dorsal bursa (for sperm storage) with a lamellar wall on whose anterior extension a cuticularized opening ("mouthpiece") is found. *Haplognathia*, on the other hand, has no bursa at all, whereas *Austrognathia* often has a vagina and a bursa, the latter lacking cuticular structures. Sperm are very diverse, ranging from tiny aflagellar droplets (in the majority of genera) or large cones (in *Austrognathia*) to motile threads propelled by a single flagellum (as in *Haplognathia*).

Physiological activities. Fertilization is internal; after injection through the skin or the vagina and subsequent storage in the bursa or in the gut wall, the sperms one by one reach the mature egg. Oviposition (egg deposition or laying) is effected by rupture of the dorsal body epithelium, whereupon the egg becomes spherical and sticks to the substratum. Development is direct, with spiral cleavage and gastrulation as is typical for mesolecithal eggs (that is, eggs containing an intermediate amount of yolk). The hatching juvenile measures some 100 μm in length; of its pharynx apparatus, only the central part of the

basal plate is developed. Despite its impressive armature, the sophisticated pharyngeal apparatus simply serves to scrape the molecular film of organic material [together with cyanobacteria (blue-green algae), bacteria, and fungal hyphae] off the sand grains. Body fragmentation is a phenomenon that cannot be explained as yet, whereas the formation of a mucous cyst, as observed in some *Austrognathia*, might help to overcome deterioration of the environment.

Ecology and distribution. The late discovery of the phylum can be attributed to at least three facts: (1) Sand rich in organic detritus [which is the typical biotope (an area of uniform environmental conditions and biota)] did not receive much attention in classical sand-fauna research. (2) Many gnathostomulids prefer a deeper (and mostly anaerobic) sand layer, especially the boundary zone between the oxidized surface and the reduced "black" layer, where steep surface chemical gradients occur. (3) Methods of extraction, mostly based on the fact that in a sediment sample the animals migrate to the surface, were such that H_2S-producing samples were discarded prior to the rather late appearance of gnathostomulids. The aspect of the biotope suggests that at least part of the metabolism (and obviously also development) takes place under anaerobic conditions. Investigations have shown that whenever the typical gnathostomulid biotope is encountered, it regularly produces several genera (up to six) and species (up to 15 so far), often in hundreds of specimens per liter of sand. Biotopes range in depth from the upper tidal to 1300 ft (400 m). The distribution of gnathostomulids is worldwide, the majority of localities being known from European coasts and the east coast of North America, with a growing number of reports from the Pacific Ocean.

Relationships. The gnathostomulids belong to the general grouping of "lower worms." First considered to be aberrant Turbellaria, they are now thought to be a separate phylum related to Rotifera and Micrognathozoa. A relationship with the enigmatic fossil conodonts could not be substantiated. *See* CONODONT; ROTIFERA. W. E. Sterrer; Rupert J. Riedl

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; M. V. Sørensen and W. Sterrer, New characters in the gnathostomulid mouth parts revealed by scanning electron microscopy, *J. Morphol.*, 253: 310–334, 2002; M. V. Sørensen, W. Sterrer, and G. Giribet, Gnathostomulid phylogeny inferred from a combined approach of four molecular loci and phylogeny, *Cladistics*, 22:32–58, 2006; W. Sterrer and S. Tyler (comp.), Taxonomic database of the Gnathostomulida, Version 1.0 [http://devbio.umesci.maine.edu/styler/gnathostomulida], 2003.

Gneiss

An ancient geological term for coarse-grained, banded crystalline rock. The word, first applied by German miners to rocks widely exposed in the Harz Mountains, was adapted to scientific usage in the late eighteenth century.

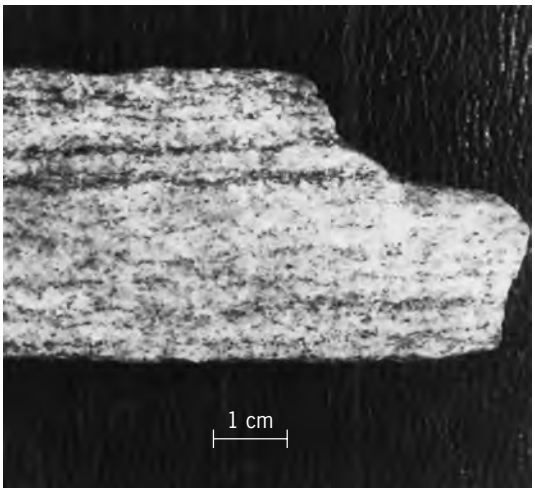

Gneiss formed by metamorphism of preexisting granite. Dark minerals are mica; light-colored minerals are quartz and feldspar. The streaky nature of banding is typical of gneisses. The sample is from the Great Smoky Mountains of North Carolina.

Gneiss is composed of mineral grains large enough to be seen with the naked eye (see **illus.**). Banding arises from segregation of the various minerals present, typically into dark- and light-colored layers. Individual bands are commonly 0.04 to 0.4 in. (1 mm to 1 cm) thick. Although individual mineral grains are often flattened parallel to banding, such shape orientation is not present in many gneisses. Sheetlike minerals such as micas may be present but form only a subordinate amount of the rock. Banded rock of coarse grain containing substantial amounts of such minerals is named schist. Crystalline rock which has flattened grains but lacks obvious banding is generally called leptite. *See* SCHIST.

Gneiss is defined by its texture, or arrangement of mineral grains, rather than by its mineral composition. However, the term gneiss is often taken to imply a mineral composition of granitic type, dominated by quartz and feldspar. Gneisses of other compositions are identified by qualifying terms such as compositional rock names, as in diorite gneiss and amphibolite gneiss, or a partial list of minerals present, as in biotite-plagioclase gneiss and hornblende-plagioclase gneiss. *See* FELDSPAR; QUARTZ.

Most gneisses are formed by recrystallization of preexisting rock during intense regional metamorphism. Shear stress present during such metamorphism causes formation of gneissic banding, although the exact mechanisms of this process are not well understood. Thus gneissic banding should never be mistaken for preexisting structures in the rock, such as sedimentary bedding. In most cases, recrystallization of minerals is dominant over shear, and individual minerals are not deformed. However, bent and fractured minerals may occur in mylonitic gneisses, in which shear was dominant over recrystallization. A variety of preexisting rock

types (igneous, sedimentary, or metamorphic) may be converted into gneiss. Some granitic gneisses may be of direct igneous origin, having formed by the shear arising from differential flow of viscous granitic magma. *See* MYLONITE.

Gneisses typically occupy large areas within the high-grade cores of regional metamorphic belts. Such terranes are often difficult to understand, because the processes which cause formation of gneissic texture are also sufficient to obscure preexisting rock structures. High temperature and shear are sufficient to cause plastic flow of gneissic rock on a gigantic scale. Major, map-size structures which have been identified in such terranes include mantled gneiss domes and nappes.

Such conditions of metamorphism are probably brought about by deep tectonic burial and major regional compression. Thus gneissic terranes may be expected to form in areas of convergent plate tectonics. *See* METAMORPHIC ROCKS; METAMORPHISM; METASOMATISM; PLATE TECTONICS. David W. Mohr

Bibliography. M. J. Hibbard, *Petrography to Petrogenesis*, 1994; A. Miyashiro, *Metamorphic Petrology*, 1994; A. Spry, *Metamorphic Textures*, 1969.

Gnetales

The only order of the class Gnetopsida in the division Gnetophyta. There are three living families, each with a single genus: Ephedraceae (*Ephedra*; 65 species in arid regions), Gnetaceae (*Gnetum*; 29 species in the tropics), and Welwitschiaceae (*Welwitschia*; 1 species in Namibia). Among the gymnosperms, Gnetales are usually considered the closest living relatives of the angiosperms. This relationship is supported by such morphological features as wood vessels, netted venation, and double fertilization, but contemporary molecular evidence is ambiguous, sometimes favoring this relationship and sometimes implying a closer affinity with the conifers. The fossil record is sparse but increasing, perhaps extending back to the Triassic Period. Each modern genus is quite distinctive. Species of *Ephedra* (Mormon tea) somewhat resemble the spore-bearing horsetails (*Equisetum*), with scale leaves on photosynthetic, jointed stems. Most species of *Gnetum* are lianas, woody vines closely resembling the Australian flowering plant family Austrobaileyaceae. The only species of *Welwitschia* is unlike any other living plant. From its stumplike trunk, just two leaves grow ever wider and longer [up to 7 m (23 ft)], tattering at the tips. All three genera are dioecious, with separate male and female individuals. The pollen and seed cones are compound, with flowerlike buds in the axils of bracts. *See* CYCADOPSIDA; PINOPHYTA; PLANT KINGDOM. James E. Eckenwalder

Bibliography. W. E. Friedman (ed.), Biology and evolution of the Gnetales, *Int. J. Plant Sci.*, 157(6,suppl.):S1–S125, 1996.

Gnotobiotics

The science involved with maintaining a microbiologically controlled environment, and with the knowledge necessary to obtain and use biological specimens in this environment. The roots of the word are *gnotos*, meaning well known, and *biota*, the combined flora and fauna of a region.

All exposed surfaces of an animal are teeming with microbes. The contents of the large intestine may contain 3 trillion microbes per ounce (100 billion per gram), belonging to several hundred species. This microbiota (bacteria, viruses, molds, yeasts, protozoa, and small parasites) is so complex, so incompletely characterized (some cannot yet be grown in pure culture), and so subject to changes in relative proportions that knowledge of what that microbiota is and what it does cannot be directly gained. Pathogens are more amenable to study under such ill-defined conditions because of their dramatic, overriding effects on the host.

Even if the animal itself is the primary interest of a researcher, there is no direct way to determine how many of an animal's normal characteristics are truly its own, and how many involve interaction with or reaction against resident microbiota. The only way to determine this is comparison with animals that have no microbiota. If differences are found, then the role of individual microbial species can be studied by inoculating pure cultures of these species into the animals without a microbiota.

Thus, gnotobiotics evolved initially to answer questions about what difference the resident microbiota makes, and which members of the microbiota make the difference. Answers become more and more essential in going beyond the effects of pathogenic microbes to the harmful or helpful long-term effects of environmental chemicals, compounds produced in the host's own metabolism, and therapeutic drugs being tested for efficacy, toxicity, or carcinogenicity. The activity of the microbiota is proportionately much greater in laboratory animals than it is in humans, and could have a decisive effect on such chemicals, especially since the chemicals are often received in small doses. For example, intestinal microbes turn a minor component of the cycad bean, a South Pacific foodstuff, into a carcinogen. One of the best drugs against parasitic schistosomes in humans is turned into a carcinogen by a single species of intestinal streptococcus. Gnotobiotic studies are designed to detect such possibilities.

Gnotobiotes. Gnotobiote is the term applied to an animal (or plant) with a defined microbiota. The most simple of gnotobiotes, and the ultimate source of all other gnotobiotes, is the animal with no microbiota. Invertebrate animals of this type are most frequently called axenic. Vertebrate animals may also be called axenic, but are more frequently called germfree. These terms suffice until defined microbial species are added to the germfree or axenic animal. Then the axenic animals becomes gnotoxenic. The germfree animal, which was already a germfree gnotobiote, becomes an associated gnotobiote. As

defined operationally by the Committee on Standards of the Institute of Laboratory Animal Resources, "A gnotobiote is one of an animal stock or strain derived by caesarean section or sterile hatching of eggs that is reared and continuously maintained with germfree technics under isolator conditions and in which the composition of any associated fauna and flora, if present, is fully defined, by accepted current microbiology."

Gnotobiology. This is a term sometimes used to designate studies involving gnotobiotes, although it tends to suggest that there is a unified body of knowledge which results from studying gnotobiotes. In fact, the gnotobiote is a more precisely defined laboratory animal which helps elucidate biological phenomena in immunology, nutrition, physiology, oncology, gastroenterology, microbial ecology, gerontology, pathogenic microbiology, parasitology, and so on. Nevertheless, some of these findings serve a dual purpose; besides their contribution to other disciplines, they prove to be necessary knowledge for obtaining, maintaining, or intelligently using gnotobiotic animals, and thus, by definition, belong in the science of gnotobiotics.

History and significance. Louis Pasteur, the originator of the concept (not the term) gnotobiotics, suspected a profound symbiosis and even mutualism between a host animal and its microbiota. He said that if he had the time he would like to grow animals free of all germs to test his assumption that they would not be able to continue life. In 1886, M. Nencki, the discoverer of ptomaines, countered with chemical reasons to support the opposite hypothesis of E. Metchnikoff, namely, that much of an animal's microbiota has a toxic, life-shortening effect and that the germfree animal would be better off by not having any. Early attempts at rearing germfree chickens and guinea pigs by both factions showed that animals could survive germfree, at least briefly, but poor health resulted; thus the experiments could be used to argue for or against either hypothesis.

It is now realized that both sides were right. Herbivorous animals consuming their natural diet are completely dependent on their microbiota to supply vitamin B_{12} and possibly other substances. G. Nuttall and H. Thierfelder, who in 1985, carried out the first experiment with germfree guinea pigs, discovered that they could rear them better with animal-derived feeds than with vegetable feeds; this was 55 years before identification of animal protein factor, which is actually a microbial product, vitamin B_{12}. However, the researchers were dissatisfied with the results of both diets and did not continue. Other one-shot experiments to test either hypothesis were carried out over the next three decades. They yielded only short-term results because of inadequate isolator design and materials, and because nutritional science was not yet equal to the task of providing diets which were adequate after sterilization and made up for any microbial contribution to digestion and to production of nutrients.

Furthermore, the very first germfree experiments with guinea pigs revealed an unexpected effect of microbes on the host. Germfree rodents and rabbits develop a conspicuous distention of the cecal pouch at the junction of the small and large intestines. The cecum is already relatively large in these species under open animal room conditions. The presence of this large amount of germfree cecal contents was subsequently found to decrease basal metabolism and cardiac output. The fact that surgically removing the cecum brought the germfree parameters toward those in germ-bearing animals indicated that the microbes' effect on the cecum and on metabolism was the result of their removing some materials which were generated by the host during digestion and accumulated in the cecum. It has taken years of experimentation with different kinds of diet to reduce cecal distention to minor proportions.

As a result of knowing more of what microbes do to benefit their host and learning how the experimenter can substitute for these microbial effects by dietary manipulation, Metchnikoff's hypothesis that the microbiota is life-diminishing can now be tested. In fact, germfree rats lived one-third longer than their open animal room counterparts (often referred to as conventional or classic animals). This is so not merely because they escape overt infections but because they show a much delayed development of nephrosis and of tumors and cancers.

In thus answering the question posed by Pasteur and Metchnikoff a century ago, the basic importance of gnotobiotic research is seen. The whole animal, integrating all the subtle, long-term effects of a microbiota, is the basis for a whole-life experiment. In an era when molecular biology, cell and tissue culture, and genetic engineering are used to find molecular explanations for biological phenomena, gnotobiotics is a reminder that it is the animal, at a higher level of organization, which points out the phenomena to be explained. *See* GENETIC ENGINEERING; MOLECULAR BIOLOGY; TISSUE CULTURE.

Building on the work of researchers during the 1930s and 1940s with animals in absolute isolation (**Fig. 1***a*), clinical surgeons and immunologists have learned that the use of antibiotics to eliminate major elements of the human microbiota, and the use of partial or complete isolation to prevent their reentry (**Fig. 2**), can greatly reduce infections in human patients with spontaneous immune deficiency or with iatrogenic immune deficiency resulting from procedures to cure cancer or prevent transplant rejection. Children with severe combined immune deficiency of genetic origin have lived as well as can be expected inside germfree isolators for periods up to 12 years. Knowledge and techniques developed with germfree animals are potentially relevant to such clinical applications. *See* IMMUNOLOGICAL DEFICIENCY.

Laboratory methods. Obtaining, maintaining, testing, and using gnotobiotes has become routine for the most common laboratory and domestic animals (rats, mice, pigs, guinea pigs, rabbits, chickens, quail, dogs, cats, primates, goats, cattle, and horses); the basic procedures are the same for all (Fig. 1*b*).

Obtaining gnotobiotes. Obtaining an animal germfree or with so few associates that they can all be specified is the initial step. Nature protects the immunodeficient fetus in the womb or the chicken in the egg by keeping it germfree; to keep it so throughout life it must be removed aseptically into a germfree environment. This is done by cesarean section in a surgical isolator, with the germ-laden mother underneath the plastic floor of the isolator. The young are transferred to a rearing isolator for hand feeding. Alternatively the whole uterus may be removed, clamped shut, and passed through a germicidal trap in the same way that eggs are passed through (Fig. 1*b*).

Experience has shown that nature is not completely reliable in prenatal isolation of the young from all microbial forms. While there are extensive rat colonies which have never shown any evidence of bacterial, viral, fungal, or parasitic associates, gnotobiotic mice of all strains carry a leukemogenic virus which is present in cells of the unborn fetus, and generally remains latent unless activated by repeated small doses of radiation. Exhaustive tests are therefore essential before a cesarean-derived line can be declared germfree or gnotobiotic.

Chickens are easy to obtain germfree. The egg surfaces are chemically sterilized in a lock or trap (Fig. 1*b*), and the chicks hatch inside the isolator where they are able to feed without assistance. Rearing of the first germfree small mammals, however, forced gnotobiotic researchers to pioneer development of sterilizable substitutes for the milk of each species, and of miniaturized feeding devices. Standard laboratory-animal solid diets proved inadequate for germfree animals until modified, primarily in vitamin and mineral content, to compensate for losses during sterilization and the absence of microbial vitamin synthesis. Once reproducing germfree colonies were established, they not only maintained themselves but provided foster mothers for cesarean-derived young of new strains.

As an alternative to cesarean derivation, treatment of older animals with antibiotic mixtures and germicidal baths has proved quite effective in eliminating bacteria, fungi, and protozoa, but not any viruses that may be present. Several months of treatment, with frequent transfers to new isolators, are required.

Maintenance of gnotobiotes. Maintenance for long periods demands absolute barriers (Fig. 1). The barrier must be complete and must consist of materials proven impervious to microbes. Gloves that are an integral part of the barrier permit manipulation of the isolator contents. A sterilizable lock, with inner and outer doors, provides for entry of supplies (**Fig. 3**) and exit of wastes, and for transfers between isolators through a sterilizable corridor connecting two locks, similar to the connecting sleeve shown in Fig. 3. The lock may be replaced by a trap of liquid germicide in which the liquid forms a barrier to airflow and sterilizes the surfaces of objects passed through it (Fig. 1*b*). Air is sterilized by passage through fine glass-wool filters. Diets and bedding are not suitable for surface sterilants, and must be sterilized by steam, radiation, or penetrating gas in containers that can then be surface-sterilized or connected to a lock.

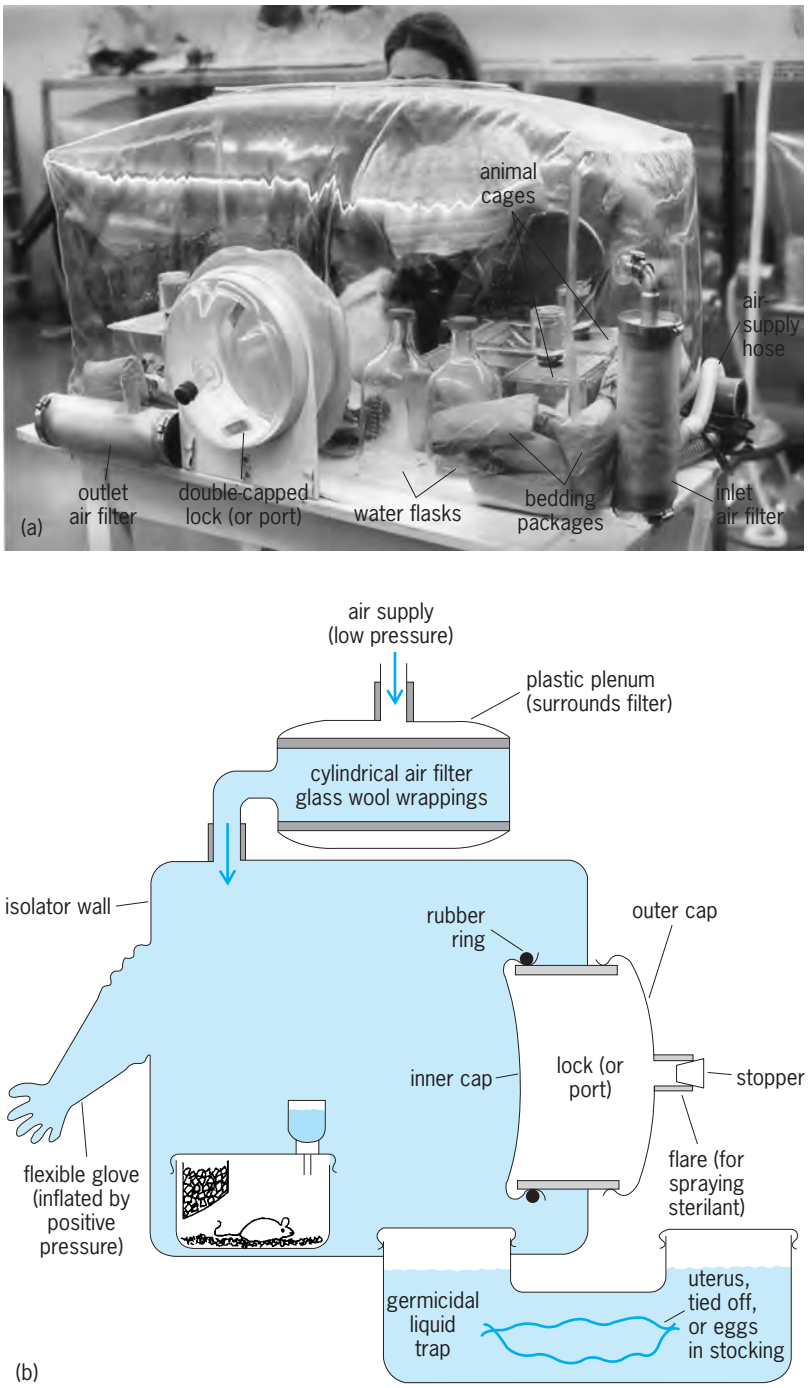

Fig. 1. Typical Trexler-type flexible film isolator. (*a*) Operating unit; operator is handling a bag of diet. (*b*) Diagrammatic cross section, including germicidal liquid trap to show one way of introducing new germfree animal lines.

Isolating systems have included heavy-metal isolators capable of holding steam under pressure for maximum sterilizing efficiency; lighter-weight metal isolators that could be put inside a large steam autoclave; sterile rooms entered by operators in diving suits; rigid plastic boxes; and flexible, transparent plastic isolators (Trexler-type isolators) which now represent the most common, simple, and cheap type

Fig. 2. Patient isolator used to protect immune-compromised patients from outside germs, and to protect hospital personnel from infection by patients with highly lethal tropical diseases. (*Vickers Americal Medical Corp.*)

of secure isolation for gnotobiotes (Fig. 1). This last type was made possible by the development of cold sterilants such as peracetic acid which can sterilize plastic and metal surfaces inside the isolator, the surfaces of objects passing through the lock, and the air initially within the isolator or lock.

Testing. Testing is required to verify both the germfree state of the initial animals and the successful maintenance of absolute barriers thereafter. The best methodologies must be applied: a variety of culture media and incubation conditions, light and electron microscopy, tissue culture, animal symptoms, and the animals' possession of antiviral antibodies. The maintenance of a reproducing colony gnotobiotically provides time for exhaustive testing of that line of animals since it was cesarean-derived. Thereafter, simpler testing of fecal matter for microscopically visible and cultivable bacteria usually suffices to detect breaks in the barrier, since the most ubiquitous microbes in the environment are relatively easy to detect, and a rare one is not likely to pass through a break alone.

Fig. 3. Transfer sleeve connecting an isolator to a cylinder containing food and bedding which has been steam-sterilized. A thin plastic film covering the cylinder end inside the sleeve will be punctured from inside the isolator after the sleeve has been spray-sterilized. The same sleeve is used for sterile transfers between isolators.

Characteristics of germfree vertebrates. Germfree vertebrates obtained by the above operations reveal characteristics which are valuable data for potential users, demonstrate the far-reaching influence of the microbiota on the host, and provide new insights in basic and applied disciplines. In the following capsule summaries, characteristics have been selected in which the germfree animal differs notably from the usual or conventional animal.

Anatomy. Hearts, lungs, and livers of germfree animals are as much as 20% smaller than those of conventional animals. Germfree rodents and rabbits (not other species) have ceca about five times the conventional size, yet their small intestines are generally thinner-walled and have less surface area and white cells.

Function. Germfree intestines show slower peristalsis, reduced sloughing of intestinal surface cells, and higher enzyme content in those cells. Germfree intestinal capillaries are 20 times less responsive to Adrenalin than those of conventional animals.

Effects of gnotobiotic association

Host	Associated microbe	Effect
Chicken	*Bacillus subtilis*	Heart lesions and death
	Streptococcus faecium	Growth depression
	Clostridium perfringens	Growth depression; local immune response
	Streptococcus faecalis	Humoral immune response
	Lactobacillus species	Improved starch digestion; elimination of *Escherichia coli*
	Eimeria tenella	No clinical symptoms in associated gnotobiote; death of conventional host
Turkey	*Histomanas meleagridis*	Much milder disease in associated gnotobiote than in conventional host unless *Escherichia coli* added
Guinea pig	*Escherichia coli*	Enteritis and death
	Streptococcus faecalis	Prevention of *Escherichia coli* pathogenicity
	Entamoeba histolytica	No growth parasite in associated gnotobiote; death of conventional host
	Bacillus subtilis	Host death in presence of *Entamoeba histolytica*
	Nippostrongylus muris	Growth of rat parasite only in associated gnotobiote guinea pig
	Nematospiroides dubius	Growth of mouse parasite only in associated gnotobiote guinea pig
Mouse	*Nematospiroides dubius*	Slower worm growth in associated gnotobite than in conventional host; incomplete life cycle in associated gnotobiote
	Trichinella spiralis	Milder disease in associated gnotobiote than in conventional host
	Escherichia coli	Elimination of *Candida albicans*
	Escherichia coli plus *Proteus mirabilis* and *Streptococcus faecalis*	Elimination of *Vibrio cholerae*
	Mycoplasma pneumonia	No infection in associated gnotobiote; pneumonia in conventional host
Mouse irradiated	*Escherichia coli*	Shorter survival in associated gnotobiote than in conventional host
	Lactobacillus acidophilus	Longer survival in associated gnotobiote than in conventional host
Rat	*Streptococcus mutans*	Rampant and typical tooth decay
	Actinomyces naeslundii	Periodontal (gum and bone) disease
	Bacillus anthracis	Much greater pathogenicity in associated gnotobiote than in conventional host
	Escherichia coli	Rapid elimination of *Shigella flexneri*
	Salmonella typhimurium	No mortality on one diet, 25% mortality on another
Pig	*Treponema hyodysenteriae*	No infection in associated gnotobiote; severe infection in conventional host
	Lactobacillus species or *Streptococcus* species	More rapid maturation of natural killer cells
Dog	Distemper virus	Much milder symptoms in associated gnotobiote than in conventional host
	Histoplasma capsulatum	Milder disease in associated gnotobiote than in conventional host

Germfree rats have a resting metabolic rate one-fourth less than conventional rats and one-third reductions in heart output and liver blood flow. Thyroid hormone levels do not decline with age in germfree rats and mice, as conventional levels do.

Growth rate. Growth rates are superior for germfree chickens, primates, pigs, and Japanese quail. Most other species show equal germfree and conventional rates, but germfree rabbits and guinea pigs have a lower rate than conventional ones.

Life-span. Germfree rats live one-third longer than conventional rats. Species like the rabbit and guinea pig in which cecal size is not well controlled by diet die earlier in germfree than in conventional status because the cecum may accidentally twist on itself.

Cancers. Spontaneous tumors occur later in germfree than in conventional rats and are usually endocrine-related. The longevity of germfree rats permits spontaneous appearance of new cancer types, such as prostate tumors, to provide animal models for human cancer. Not all types of chemical carcinogens are active in germfree animals.

Cholesterol metabolism. Germfree rats do not excrete cholesterol and bile acids in their feces as efficiently as they would microbially modified forms and therefore accumulate more cholesterol than conventional rats.

Radiation sickness. Germfree animals tolerate larger doses of radiation, live longer, and permit study of radiation effects per se without supravening infection.

Nutrition. Germfree animals need less vitamin A, protein, lysine, and vitamin C than conventional animals, but show poorer utilization of iodine and synthetic vitamin K. Germfree animals fed pure chemical diets have provided the first proof that all essential nutrients are now known.

Immune system. Germfree animals survive thymus removal at birth, permitting long-term study of the organ's role in disease resistance and aging. Germfree animals receiving bone marrow transplants show little reaction of the foreign white cells against their host. This fact has provided the basis for treatment of leukemia with transplantation preceded by antibiotics and isolation.

Germfree animals show conspicuous reductions in circulating antibodies, antibody-producing cells, size of lymph nodes and the number of their active centers, and blood clearance of bacteria by phagocytic cells. These differences can be accentuated by use of antigen-free, ultrafiltered water-soluble diets. Despite reduced ongoing function, germfree animals retain the ability to respond to challenges, and can be used to study natural antibody, natural killer cells, and the nature of a true primary immune response.

Associated gnotobiotes. Otherwise germfree animals associated with one or more defined microbial

species supply models for identifying the beneficial or harmful species among the resident microbiota, with the ultimate goal of controlling these in the host's favor. Since germfree animals are also gnotobiotes, several terms to describe associated gnotobiotes have been devised. Gnotophoric indicates the animal's carrying of known microbes, and gnotoxenic indicates the animal's association with known microbes. The **table** shows a selection from the more dramatic effects of associating germfree animals with one or several defined microbial species. Many experiments too complex for the table have shown that combinations of certain microbes from the resident microflora strongly resist the establishment of pathogenic microbes in the gut, a phenomenon called colonization resistance. Bacteria selected for this capacity through gnotobiotic research are now used to reconstitute the microflora of human patients after extensive antibiotic therapy.

Nearly all experiments with parasites (protozoa or worms) added to germfree animals have shown dramatic differences from the parasites' effects in conventional hosts. The parasites either need bacteria to establish themselves, or are dramatically antagonized by bacteria, or provoke less pathogenic effects or immune response when alone.

In terms of numbers of animals involved, the greatest use of gnotobiotes is to upgrade the quality of nearly all experimental rodents. Before the advent of gnotobiotes, all rat and mouse colonies carried ancestral pathogens, such as murine viral pneumonias, which were passed from parents to offspring after birth. These latent infections became active when the animals were stressed or kept for long periods, as occurred in many experiments. Gnotobiotes provided commercial animal breeders with pathogen-free parents or foster parents for new animal colonies which now supply the majority of the millions of rats and mice used annually in research. These are not gnotobiotic but owe their freedom from specific pathogens to the gnotobiotic colonies which commercial breeders maintain. Julian R. Pleasants

Bibliography. M. W. Balk and E. C. Melby, Jr., *The Importance of Laboratory Animal Genetics: Health and Environment in Biomedical Research*, 1984; J. Kawamata and E. C. Melby, Jr. (eds.), *Animal Models: Assessing the Scope of Their Use in Biomedical Research*, 1987; W. F. Loeb and F. W. Quimby (eds.), *The Clinical Chemistry of Laboratory Animals*, 2d ed., 1999; B. S. Wostmann et al., *Germfree Research: Microflora Control and Its Application to the Biomedical Sciences*, 1985.

Goat production

An agricultural business concerned with breeding goats, primarily for meat, milk, and fiber (called mohair and cashmere). The domesticated goat is closely related and similar in size to sheep, but has many anatomical and physiological differences. Domesticated goats of the major breeds have been developed in, and are mainly derived from, five geographical areas: Swiss mountains (for milk); Indian, Arabian, and northeast African drylands (for meat and milk); west African lowland (for meat and milk); south African prairie (for meat); Turkish highland (for mohair production). *See* CASHMERE; MOHAIR; SHEEP.

Accompanying colonizing immigrants and released from the live-food supplies of early merchant ships, goats became established in the Americas, and today make up a small (compared to cattle, swine, poultry, and sheep) but increasingly visible and important part of agriculture in the United States. Mohair production from Angora goats is of economic significance in Texas and ranks first in the world, followed by that of Turkey. Six different dairy goat breeds are recognized in the United States: Alpine, LaMancha, Nubian, Oberhasli, Saanen, and Toggenburg. *See* BEEF CATTLE PRODUCTION; DAIRY CATTLE PRODUCTION; SWINE PRODUCTION.

Breeds. Alpine goats are of Swiss origin. They have many colors, especially faded shades of white into black, a straight face, erect ears, and sickle-shaped horns, when present. Alpine does (mature females) may weigh at least 135 lb (60 kg), measure 30 in. (75 cm) at the withers (top of shoulder), and are (with Saanen and Nubian breeds) the biggest and heaviest dairy goats. Alpine goats are ranked second in milk-producing ability, behind the Saanen breed.

La Mancha goats are of Spanish origin but were developed in California. They have many colors, a straight face, but characteristically no external ears. They are smaller than Alpine goats and produce less milk, but their offspring are fleshier. They are very calm goats, and persistent milk producers in the winter when the other breeds often dry up.

Nubian goats are mostly of Indian origin, but were developed in England. They too have many colors, but characteristically arched faces and long pendulous ears. Nubian goats produce less milk than the Swiss breeds, but their milk-fat level is the highest. They are the most popular in the United States and are most heat-tolerant. They are also more fleshy. Horns in Nubian, as in LaMancha, goats are rarely seen. Although anatomically present, they are routinely and permanently removed in early age to facilitate animal handling. Horns in all Swiss breeds are of the upright sickle shape, while Indian, Spanish, and Turkish breeds have more sideways-growing horns with varying degrees of spiraling. Goats can be selected for the genetic trait of hornlessness, but this is only partially advisable since the trait is genetically linked to sterility, and a certain percentage of intersex or hermaphrodite offspring can be expected when both parents carry the hornless trait.

Oberhasli goats are of Swiss origin. They are related to Alpine goats and similar to them, except that they are often colored solid red or black. They are smaller and lower in milk production than Alpine goats, but very popular in Europe where they are well adapted to high mountain grazing and long-distance traveling. They have also been called

Fig. 1. A typical Saanen doe with good udder.

Chamoisee, Brienz, or Swiss Alpine goats in contrast to the French Alpine and British Alpine breeds which are today just called Alpine goats.

Saanen goats are of Swiss origin. They are only white and have straight faces and erect ears (**Fig. 1**). They are rated as the top milk producers around the world and have been used for the upgrading of many native breeds in underdeveloped countries. All dairy goat breeds in the United States are generally short-haired, although some Swiss strains may have partially long hair, which is of some advantage to them in alpine climates.

Toggenburg goats are also of Swiss origin. They are very popular in the United States and rank in milk-producing ability next to the Alpine and Saanen goats. They have only brown color with characteristic white facial markings and white ear and leg stripes. They are the smallest of the Swiss breeds, approximately 120 lb (54 kg) for the mature does and 26 in. (65 cm) height at withers. They have straight faces and erect ears.

In addition to the six American dairy breeds above (and the Angora), there are also two major goat populations in the United States for mixed purposes, mostly meat, brush clearance, pets, or laboratory use: the Spanish goat, mainly kept on open range land, and the West African Pygmy.

Milk. Milk is the main product of dairy goats in the United States (**Fig. 2**). There are two manufacturing plants for dry goat milk, one in California and one in Arkansas. Interest, mostly of limited local scale, is increasing in goat cheese production: soft curd (similar to cottage-ricotta), cheddar, and feta. Impetus comes from European and Near East traditions and imports, especially from France, which has a thriving, sophisticated Chèvres (soft goat cheese) industry. Liquid raw goat milk can be officially marketed in only 14 states in the United States, those with public health laws permitting such sales. In all other states, goat milk must be pasteurized. There is a strong interest nationwide in raw goat milk by people with malnourished children, by persons with cow milk allergies and gastric problems, and by those interested in natural foods. This situation has led to a wide dispersion of mostly small goat herds throughout the United States to satisfy needs on a family level.

Fig. 2. Milking a dairy goat on an elevated stand.

Goat milk is valued for raising orphaned foals and puppies and for the production of prime veal. Biochemically goat milk differs in several important aspects from cow milk. The major casein in cow milk (alpha-s-1) is found in neither goat nor human milk. The casein type in goat milk makes a softer curd and is more easily digested. The fat in goat milk consists of smaller globules than in cow milk and is more easily digested. The goat milk fat consists characteristically of much more short-chain fatty acids than cow milk.

There are also some consistent differences in vitamin, mineral, and enzyme contents between goat and cow milk; however, both are more similar in gross composition than either is to sheep milk, which has a much higher solids content and smaller daily yield. Daily milk production of good dairy goats in the United States is at least 1 gallon (3.8 liters) with an average content of 2.8% fat, 3.4% protein, 4.6% lactose, and 0.8% minerals, totaling 11.6% solids. *See* CHEESE; MILK.

Breeding. Breeding of dairy goats is similar to that of sheep. Goat kids, born in the spring, are ready for breeding in the fall of the same year and will have their next generation 150 days after conception, usually the next spring. This cycle coincides with the seasonal changes of vegetation; the spring with new pasture growth favors milk production and provides greater survival opportunity. Thus, the dairy goats (for example, the Swiss breeds) of the Northern Hemisphere with its distinct seasons evolved as seasonal breeders. The disadvantage is the consequent shortage of goat milk during winter months. Goats of Mediterranean, Near East, and Indian origin show less seasonality. Seasonal breeding is governed physiologically by day length; and by artificially reducing daylight in controlled housing, dairy goats can be brought into earlier estrus during the summer. This is a practical and successful method for obtaining an even goat milk supply throughout the year.

Breeding of dairy goats is still mostly by natural service. Bucks produce, in the scent glands on top of their head, odors attractive to does but obnoxious to people. Estrus lasts 1–2 days and may recur in 21 days if impregnation failed. Artificial insemination with frozen semen stored in liquid nitrogen tanks at -320°F (-196°C) is gaining popularity since it is

the best and most economical avenue of genetic improvement. Pregnancy testing is possible with ultrasonic scanning and radiography, but more practical by routine testing of milk samples for progesterone levels at 3 weeks after breeding.

Kidding in dairy goats is usually easier than in cows or sheep because of the generally lean condition and narrow bone structure of kids. Two or three kids per pregnancy, not necessarily identical, is the rule rather than the exception, including backward presentations. Rebreeding is not inhibited by nursing, as in some other mammalian species, but rather by the increasing day length of spring.

Modern developments are estrus synchronization of does with hormone implants or injections to overcome estrus detection problems and to facilitate artificial insemination, and embryo transfer in conjunction with estrus synchronization to enable wider use and greater number of offspring of genetically superior does. Embryo freezing is possible even now, and the sorting between different sex embryos as well as cloning from single superior embryos will be practical in the near future, thus opening exciting new methods of dairy goat improvements. Considering that in many parts of the world, where the majority of the 450 million goats are found, the productivity per animal usually averages only 10–20% of the average goat milk production in the United States, the modern breeding techniques in developed countries have contributed to the solution of malnutrition problems in the third world. *See* BREEDING (ANIMAL).

Feeding. Feeding of dairy goats is similar to that of cows and sheep, except that goats need and thrive on browse like deer, and some breeds may go without water for days. High-producing dairy goats need scientifically balanced rations as do high-producing dairy cows. *See* ANIMAL FEEDS.

Diseases. Facilities and equipment are not strongly developed for goats specifically; but as with housing, milking machines, and feed storage, medications against parasites, diseases, and metabolic disorders are adapted from, or patterned after, those for cattle. Unlike in dairy cows, the largest problem in dairy goats is internal parasites, which appear to adapt often to certain standard medications. Also, unlike in dairy cows, tuberculosis and brucellosis are not a problem in dairy goats in the United States, although annual testing continues to be advocated for safety reasons. On the other hand, there are viral diseases causing arthritis, bacterial infections producing abscesses, and intestinal bacterial disorders which are more serious in dairy goats than in dairy cows—particularly because veterinary research, education, and care in the United States are more focused on cattle than on goats. *See* AGRICULTURAL SCIENCE (ANIMAL). George F. W. Haenlein

Bibliography. J. L. Ayers and W. C. Foote, Proceedings of the 3d International Conference on Goat Production and Disease, *Dairy Goat J.*, 1982; G. F. W. Haenlein et al., Proceedings of the International Symposium on Dairy Goats, *J. Dairy Sci.*, 63(10):1591–1781, 1980; G. F. W. Haenlein and D. L. Ace, *Goat Extension Handbook*, University of Delaware, Department of Animal Science and Agricultural Biochemistry, 1983; P. Morand-Fehr et al., Nutrition and systems of goat feeding, *Proceedings of the International Symposium at Tours*, Institut National de Recherche Agronomique, Paris, 2 vols., 1981; National Research Council, *Nutrient Requirements of Goats: Angora, Dairy, and Meat Goats in Temperate and Tropical Countries*, 1981; S. N. Singh and O. P. S. Sengar, *Final Technical Report: Studies on the Combining Ability of Desirable Characters of Important Goat Breeds for Meat and Milk Separately and in Combination*, Department of Animal Husbandry and Dairying, Raja Balwant Singh College, Bichpuri (Agra), India, 1979.

Gobiesociformes

An order of bony fishes, also known as the Xenopterygii, or clingfishes, equipped with a thoracic sucking disk which serves for attachment to the substrate. This papillose disk involves the four rays of each pelvic fin, pelvic and pectoral girdles, and dermal flaps. There are single dorsal and anal fins that lack fin spines (see **illus.**). The body is scaleless, and

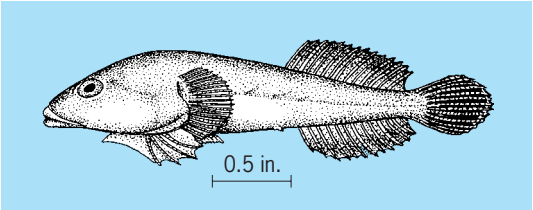

Northern clingfish (*Gobiesox maeandricus*). (*After D. S. Jordan and B. W. Evermann, The Fishes of North and Middle America, U.S. Nat. Mus. Bull. 47, 1900*)

there are no ribs and no swim bladder. The posttemporal is not forked. The order consists of a single family that is classified into 8 subfamilies, 33 genera, and nearly 100 Recent species. A Miocene genus is questionably assigned to the family. The Gobiesocidae are thought to be related to the Batrachoidiformes, or toadfishes. Clingfishes inhabit the intertidal zone of tropical to temperate shores of all continents, and a few occur in swift coastal streams of the American tropics. *See* ACTINOPTERYGII; BATRACHOIDIFORMES; SWIM BLADDER. Reeve M. Bailey

Bibliography. J. C. Briggs, *A Monograph of the Clingfishes (Order Xenopterygii)*, Stanford Ichthyol. Bull., vol. 6, 1955.

Gödel's theorem

The result, proved by K. Gödel in 1931, that any sufficiently advanced mathematical system must be incomplete in that there must always be a true sentence that is not provable in the system. Roughly speaking, Gödel showed how, for each such system, a

sentence could be constructed that asserted its own nonprovability in the system.

Provability. As an illustration, it is useful to consider the following sentence:

This sentence can never be proved.

If the above sentence is false, then what it says is not the case, which means that it can be proved, but false sentences cannot be proved. Thus it cannot be false; it must be true. Now, it has just been proved that the sentence is true. Since it is true, then what it says really holds, which means that it can never be proved. But the sentence was, in fact, just proved.

Thus, there is an apparent paradox. The fallacy in this reasoning is that the notion of proof is not well defined. No precise definition of proof has ever been given in any absolute sense; only provability within a given axiom system is discussed. It is now useful to consider an axiom system, which may be called system S, in which the notion of provability within the system is clearly defined, and to suppose also that the system is wholly correct in that every sentence provable in the system is really true. Then the following sentence may be considered:

This sentence is not provable in system S.

No paradox now arises, but instead the interesting fact that the above sentence must be true but not provable in system S. Indeed, if the sentence is false, then what it says is not so, which means that the sentence is, in fact, provable in system S, contrary to the given condition that the system proves only true sentences. And so the sentence must be true, hence not provable in system S (as the sentence says).

Self-reference via Gödel numbering. In attempting to construct a sentence that asserted its own nonprovability, Gödel considered mathematical systems that did not involve sentences or provability, but numbers, sets of numbers, and other purely mathematical entities. He avoided this problem by assigning to each sentence a positive whole number, now called the Gödel number of the sentence. Then he constructed an ingenious sentence G which asserted that a certain number g failed to belong to a certain set P of numbers. This set P was the set of Gödel numbers of all the provable sentences of the system and, more amazingly, the number g was the Gödel number of the very sentence G. Thus, G asserted that its own Gödel number g was not the Gödel number of a provable sentence. If G were false, then g would, in fact, be the Gödel number of a provable sentence, which would mean that G was provable, contrary to the fact that only true sentences are provable in the system. Therefore, G must be true, hence g really is not the Gödel number of a provable sentence, which means that G is not provable in the system. Thus, G is true but not provable in the system.

Gödel's second theorem. Gödel demonstrated a still more remarkable result that for these same systems, which include the most comprehensive mathematical systems known. He showed that, if these systems are consistent, then they cannot prove their own consistency. This is known as Gödel's second incompleteness theorem.

Unfortunately, there are widespread misconceptions about this result. For example, Gödel's second theorem is sometimes thought to imply the impossibility of knowing that these systems are consistent. In reality, the fact that a system cannot prove its own consistency does not constitute the slightest grounds for doubting its consistency. Indeed, if a system could, in fact, prove its own consistency, that, of course, would not be any guarantee that the system was consistent, since an inconsistent system can prove anything. To trust the consistency of a system just because it can prove its own consistency would be as foolish as trusting a person's veracity just because he or she claimed to always tell the truth. In reality, whether a system can or cannot prove its own consistency does not have the slightest bearing on whether or not the system is actually consistent. The consistency of the systems considered by Gödel is known rather by self-evident nature of the axioms and the obvious correctness of the rules of reasoning. Still, it is of interest that the systems, though obviously consistent, cannot prove their own consistency.

Impossibility of mechanizing mathematics. In the eighteenth century, G. Leibniz envisioned a universal calculating machine that could solve all mathematical problems. The impossibility of such a device has been conclusively demonstrated by further ramifications of Gödel's work developed by A. Church and A. Turing.

It is useful to consider a type of computing machine that is programmed to generate a set of positive whole numbers. For example, a generating machine might be given the following two instructions: (1) Print out 2. (2) If you print out n, then also print out $n + 2$. Then the machine will successively print out the numbers $2, 4, 6, 8, 10, \ldots$, in other words, the set of even numbers. A machine can be instructed to generate the set of odd numbers simply by changing the first instruction to "Print out 1." Now, a set S of numbers is called solvable (the more technical term is recursive) if there is a machine M to generate S and another machine M' to generate the set S' of all numbers that are not in S. (For example, the set of even numbers is solvable, since there is a machine that generates the even numbers and another that generates the odd ones.) Given the machines M and M', it is possible to mechanically decide of any given number x whether or not it is in S by simply setting both machines going at the same time and waiting to see whether M prints out x, in which case x is in S, or M' prints out x, in which case x is not in S but in S'. *See* RECURSIVE FUNCTION.

There are infinitely many of these possible generating machines, as many as there are positive whole numbers, and each number x is the label of one and only one machine, which may be denoted M_x. Now, it may or may not happen that the set of numbers

generated by M_x will contain the very label x of M_x; if this happens, x is called a special number, and if not, x is called an ordinary number. The interesting thing is that one of the machines C prints out all the special numbers but none of the ordinary ones. This machine C, so to speak, keeps track of the activity of all the machines, and whenever it observes that M_x prints out x, then C also prints out x. Now, the distinction between special and ordinary numbers is of considerable mathematical interest, because all formal mathematical problems can be reduced to determining which numbers are special and which are ordinary. And so if the set of special numbers were solvable, then there would exist a completely mechanical method of solving all formal mathematical problems, and Leibniz's universal calculating machine would be realized. Since there is the machine C that generates all the special numbers, the key question then is whether there is a machine C' that generates the set of ordinary numbers. In fact, there cannot be such a machine C' by the following argument:

If there were, indeed, such a machine, then it would have some label b. Thus M_b would print out all ordinary numbers but no special ones. The question then arises as to whether the number b is special or not. In either case, there is a contradiction. On the one hand, if b, indeed, is special, then M_b prints out b, contrary to the fact that M_b prints out only ordinary numbers. On the other hand, if b is ordinary, then, since M_b prints out all ordinary numbers, M_b must print out b, which makes b special. Thus, b cannot be either ordinary or special without contradiction, and hence there cannot exist a machine M_b that generates the set of ordinary numbers. Therefore, the set of special numbers is not solvable.

Thus, there is no purely mechanical method of determining which numbers are special and which are not, and so mathematics simply cannot be mechanized. Creativity and ingenuity will always be required. This is perhaps the most important consequence of Gödel's work. Another way of stating this consequence is to say that human beings can never eliminate the necessity of using their own intelligence, regardless of how cleverly they try. *See* LOGIC.

Raymond M. Smullyan

Bibliography. M. Davis, *The Undecidable*, 1973; N. Shankar, *Mathematics, Machines, and Gödel's Proof*, 1994; S. G. Shanker, *Gödel's Theorem in Focus*, 1988; R. Smullyan, *Gödel's Incompleteness Theorems*, 1992.

Gold

A chemical element, Au, atomic number 79 and atomic weight 196.967, a deep yellow, soft, and very dense metal. Gold is classed as a heavy metal and as a noble metal; commercially, it is the most familiar of the precious metals. Copper, silver, and gold are in the same group of the periodic table of elements. The Latin name for gold, *aurum* (glowing dawn), is the source of the chemical symbol Au. There is only

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

one stable isotope of gold, that of mass number 197. *See* PERIODIC TABLE.

Uses. Consumption of gold in jewelry accounts for about three-fourths of the world's production of gold. Industrial applications, especially electronic, consume another 10–15%. The remainder is divided among medical and dental uses, coinage, and bar stock for governmental and private holdings. Gold coins and most decorative gold objects are actually gold alloys, because the metal itself is too soft (2.5–3 on Mohs scale) to be useful with frequent handling. *See* GOLD ALLOYS.

Radioactive ^{198}Au is used in medical irradiation, in diagnosis, and in a number of industrial applications as a tracer. Another tracer use is in the study of movement of sediment on the ocean floor in and around harbors. The properties of gold toward radiant energy have led to development of efficient energy reflectors for infrared heaters and cookers and for focusing and retention of heat in industrial processes.

Occurrence. Gold occurs widely throughout the world, but usually very sparsely, so that it is quite a rare element. Sea water contains low concentrations of gold, on the order of 10 μg per ton (10 parts of gold per trillion parts of water). Somewhat higher concentrations accumulate on plankton or on the ocean bottom. At present, no economically feasible process is visualized for extracting gold from the sea. Native, or metallic, gold and various telluride minerals are the only forms of gold found on land. Native gold may occur in veins among rocks and ores of other metals, especially quartz or pyrite, or it may be scattered in sands and gravel (alluvial gold). *See* GOLD METALLURGY.

Properties. The density of gold is 19.3 times that of water at 20°C (68°F), so that 1 ft^3 of gold weighs about 1200 lb (1 m^3, about 19,000 kg). Masses of gold, like those of other precious metals, are measured on the troy scale, which counts 12 oz to the pound. Gold melts at 1064.43°C (1947.97°F) and boils at 2860°C (5180°F). It is somewhat volatile well below its boiling point. Gold is a good conductor of heat and electricity. It is the most malleable and ductile metal. It can easily be made into translucent sheets 0.0000039 in. (0.00001 mm) thick or drawn into wire weighing only 0.00005 oz/ft (0.5 mg/m). The quality of gold is expressed on the fineness scale

as parts of pure gold per thousand parts of total metal, or on the karat scale as parts of pure gold per 24 parts of total metal. Gold readily dissolves in mercury to form amalgams. Gold is one of the least active metals chemically. It does not tarnish or burn in air. It is inert to strong alkaline solutions and to all pure acids except selenic acid.

Compounds. Gold may be either unipositive or tripositive in its compounds. So strong is the tendency for gold to form complexes that all the compounds of the 3+ oxidation state are complex. The compounds of the 1+ oxidation state are not very stable and tend to be oxidized to the 3+ state or reduced to metallic gold. All compounds of either oxidation state are easy to reduce to the metal.

In its complex compounds gold forms bonds most readily and stably with halogens and sulfur, less stably with oxygen and phosphorus, and only weakly with nitrogen. Bonds between gold and carbon are fairly stable, as in the cyanide complexes and a variety of organogold compounds. William E. Cooley

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; J. R. Davis (ed.), *Metals Handbook: Desk Edition*, 2d ed., ASM International, 1998; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; J. Marsden and I House, *Chemistry of Gold Extraction*, 2d ed., 2006; H. Schmidbauer, *Gold: Chemistry, Biochemistry and Technology*, 1999.

Gold alloys

Combinations of gold and other metals. Gold is not oxidized in air, and for thousands of years was the only metal available that remained bright indefinitely. This, coupled with its pleasing yellow color, made it cherished for personal adornment and other decorative uses and valuable as a medium of exchange. The high specific gravity of gold (19.3) also gave some insurance against dilution with cheaper metals of lower density. Pure gold is soft. The addition of copper hardens the gold, and ultimately gold-copper alloys became standard for coinage. Gold coins in the United States contained 10% copper, the balance gold. Prior to 1934 the value of coinage had been based on the price of gold. In that year coinage of gold ceased in the United States, and gold was priced at a market value. The United States Mint sold gold to licensed users for industrial and ornamental uses from 1934 to March 17, 1968. From this date to January 1, 1975, some banks and one broker were licensed to sell gold, but only to licensed users for industrial and artistic purposes. Since 1975, there have been no restrictions on the buying and selling of gold.

Jewelry. Pure gold is weak, having a tensile strength of less than 20,000 lb/in.^{2} (138 megapascals) when annealed; however, by alloying with copper, sometimes in conjunction with silver or nickel, and often a little zinc, gold alloys with strengths of 60,000–100,000 lb/in.^{2} (414–690 MPa) may be made. Addition of these metals changes the color of gold so that red, yellow, greenish, and white golds result. The proportion of gold in solid gold jewelry is designated in karats (k); pure gold is 24 k, 18 k is 18/24 or 75% pure gold, and 14 k is 14/24 or 58.3% pure gold.

Very early it was found that reasonably pure gold could be hammered into extremely thin foil, about 0.000005 in. (125 nanometers) thick, and this was used for architectural and other decorations in pre-Biblical times. Such foil, also known as gold leaf, is still made and used in about the same way.

Because of the relatively high cost of gold it is often used in laminated form, a thin layer of karat gold being welded to a supporting metal, such as nickel or brass, which is rolled or drawn into complex forms without rupturing the gold-alloy surface. This material is called gold-filled or rolled-gold plate, and the ratio of the weight of the gold alloy to the total weight of the composite material is expressed as a fraction, along with the karat of the gold alloy; for example, $^1/_{\sqrt{10}}$ 12 k gold means that the article contains 5% of pure gold by weight. An exception to this method of designating quality occurs in watch cases, for which a definite thickness of gold alloy is stipulated.

Properties and uses. Gold is not oxidized in air, even on heating. It is insoluble in all single acids, although it can be dissolved in a mixture of nitric and hydrochloric acids (aqua regia) or in a solution of alkaline cyanides in the presence of air or an electric current. This solubility in cyanide is widely employed in extraction of finely divided gold from ores. The method can recover gold that has escaped amalgamation and many waste heaps from early mining operations have been reworked to extract additional amounts of gold. When an electric current is passed through a cyanide solution, gold can be deposited on other metals as an electroplate; this reaction is often used for protecting and decorating base metals.

Industrial uses of gold depend primarily upon the corrosion resistance and secondarily upon the strength that can be secured by alloying alone or by alloying and heat treating. Many alloys used in dentistry contain gold, silver, and copper, often with small amounts of platinum and palladium; these alloys can be heat-treated to develop strengths above 150,000 lb/in.^{2} (1.0 gigapascal). The latter have good spring properties. Alloys of this type find many electrical uses as contacts, particularly where rubbing is involved. Softer alloys are used for make-and-break electrical contacts, at currents less than 0.5 A, such as those employed in many instruments and in telephone equipment. For the latter application, gold alloys have now largely been replaced by palladium which is more economical and more resistant to electrical erosion. Gold electroplate, often thin, is employed on high-frequency conductors, such as those in radar equipment, because of the high electrical conductivity and tarnish resistance of gold. For the same reason gold is employed in the construction of many transistors, microcircuits, printed circuits,

and integrated circuits. In these applications ease of deposition by vacuum evaporation, cathodic sputtering, thermal decomposition, or electrode-position are important. Most such devices are so small that the cost of gold is relatively unimportant.

An excellent thermocouple has been developed that produces nearly the same electromotive force as the widely used Chromel-Alumel thermocouple; one leg is a palladium-platinum-gold alloy and the other leg is a gold-palladium alloy. Thin coatings of gold applied to metallic or glass surfaces may be employed for reflectors, particularly for infrared radiation.

Gold, which melts at a temperature intermediate between the melting points of silver and copper, finds some use, with its alloys, as a corrosion-resistant brazing material in chemical equipment and also in certain electronic vacuum devices. A gold coating is also sometimes applied to the grids of vacuum tubes to minimize electron emission from the grid.

Because gold does not oxidize when heated in air, appropriate gold compounds can be decomposed by heat to liberate the metal. Compounds of this type are used in the decoration of china and also for the production of printed electrical circuits on ceramics. These materials, known as liquid bright golds, are applied in the form of varnish, which is dried and then heated to redness, leaving a thin film of gold firmly attached to the underlying ceramic. This coating may be as thin as 0.00004 in. (1.0 micrometer) but may be made thicker. It is even possible to utilize this technique for applying thin gold coatings to some of the more stable plastics, which are being used for special printed circuits. Certain gold alloys with platinum and palladium can also be applied in this manner to produce stable electrical resistors. *See* GEM; GOLD; GOLD METALLURGY.

Charles R. Marsland; George Sistare

Bibliography. G. Gajda, *Gold Refining*, 2d rev. ed., 1980; K. Von Mueller, *Gold Refiner's Manual*, 1984.

Gold metallurgy

Extracting gold from ores, refining it, and preparing it for use. Total world resources of gold are estimated at about 83,000 tons (75,000 metric tons). South Africa has about half of these resources, and Brazil, Russia, and the United States have about 12% each. The United States produces about 330 tons (300 metric tons) per year. In the United States, gold is used for jewelry and arts (70%), industry and electronics (23%), and dentistry (7%). There are only a few dozen placer mines in the United States, nearly all in Alaska. In such mines, gold is processed with the modern equivalent of gold panning—sluicing, tabling, and jigging. In addition, by-product gold from copper mining is only about 10% (historically, this source used to be much larger). There are several hundred lode mines in the United States, where the ore is mined from solid rock. This gold is often difficult to recover because it is associated with sulfide or carbonaceous minerals. As technology has improved, possibilities for processing the more difficult-to-treat (refractory) ores have expanded. A particular ore is more or less refractory depending on its combination of chemical compounds and minerals. *See* GOLD; PLACER MINING.

Although each ore's treatment process is unique, a number of steps are common to the various methods.

Cyanide leaching. In this process, ore is first crushed dry in a gyratory crusher and ground wet in a semiautogenous grinding mill [a 30-ft-diameter (9-m) rotating cylinder containing steel balls]. During the wet grinding process, cyanide and lime are added. The ore leaves the grinding mill as a slurry of muddy water. At this point, granules of activated carbon are introduced, and the slurry flows into large stirred tanks. The gold gradually leaches out of each tiny ore particle (during its 20–30-h residence time) and dissolves into solution. As the gold dissolves, it is immediately picked up by the carbon, which is then screened from the muddy slurry, rinsed off, and pumped to a gold desorption tank, where gold is stripped from the carbon. Typically, a 20% ethanol stripping solution at 80°C (176°F) is used. The gold has now become concentrated in the stripping solution, and is pumped with the solution into electrowinning tanks containing steel wool as cathodes. There, it is gold-plated onto the steel wool. After several hours, the resulting "golden fleece'' is melted with flux. The flux removes the steel and impurities, and the molten gold is cast into shiny bars. The gold bars are usually sold for further refining. The carbon, now free of its gold, is heated without oxygen to 800°C (1470°F) to reactivate it and to restore its large surface area before it is returned to the slurry tanks to adsorb more gold. In the meantime, the slurry leaves the tanks and, after solid-liquid separation in a thickener, eventually ends up in the tailings pond.

Sometimes lower-grade ore (<0.05 oz/ton or 1.57 g/metric ton of gold) is simply crushed and placed in heaps where it is slowly leached with cyanide solutions. Even though heap leaching gives lower gold recovery (sometimes lower than 60%), the cost is lower. If too much clay is present in the crushed ore, it should be agglomerated by mechanically tumbling the ore with a small amount of water, lime, and portland cement. Agglomeration increases heap permeability and gold recovery and reduces the leaching time. Heap leaching is a form of solution mining that can also be applied to old tailings piles and mined overburden dumps. *See* SOLUTION MINING.

Refractory ores. The vat leaching configuration is known as carbon-in-leach and is one of the simplest ways to overcome preg-robbing—one major characteristic of refractory ores. Preg-robbing ores are difficult to treat because the gold, after first being dissolved into the cyanide solution (now "pregnant" with gold), adsorbs back onto the ore, thereby "robbing" the solution. However, in the carbon-in-leach process the freshly activated carbon is placed directly in the leaching tank. It attracts the gold more

strongly than preg-robbing constituents (such as, humic acids, carbonaceous compounds, and certain clay minerals). Furthermore, the carbon flows countercurrent to the slurry, so that the freshest carbon is mixed with the most gold-depleted slurry—an important engineering principle giving maximum efficiency.

If the ore is more refractory, stronger ore oxidation treatments are included in the process. These intense oxidation treatments overcome the other major characteristic of refractory ores, that is, pyrite encapsulation of the gold. The treatments also neutralize or partially neutralize the preg-robbing tendencies. In these treatments, the pyrite, or other encapsulating sulfide mineral, is oxidized. The oxidation of pyrite breaks down the mineral structure, exposes the gold, and makes the gold amenable to subsequent cyanide dissolution. These oxidizing treatments, in order of increasing power and risk, include air or oxygen oxidation of slurries in open tanks, chlorine oxidation of slurries in closed tanks, high-temperature roasting of dry powders in furnaces, high-pressure oxidation of slurries in autoclaves, bacterial oxidation of slurries in open tanks, and bacterial oxidation of coarse crushed ore in heaps.

If the ore is less refractory, the process might be modified to include less powerful alternatives to carbon-in-leach, such as carbon-in-pulp and countercurrent decantation. In the carbon-in-pulp process, the carbon is mixed with the slurry in downstream tanks after leaching. In countercurrent decantation, the carbon is packed in columns. The slurry leaves the leaching tanks and flows through a series of thickeners for solid-liquid separation and thorough washing. Clarified, gold-rich solution overflows from the end thickener and into the carbon columns, where gold is adsorbed as the solution percolates through. Carbon columns are also often used for clarified solutions from heap leaching. Carbon adsorption of silver is less efficient than gold. If silver values in the leach solutions are relatively high [greater than 1.5–3.0 oz/gal (10–20 g/liter)], a zinc cementation process replaces the carbon process. Here, both silver and gold are electrochemically displaced from solution after a countercurrent-decantation-type solid-liquid separation, filtering, vacuum deoxygenation, and the addition of zinc powder. Instead of replacing carbon entirely, another alternative is precipitation with silver sulfide either before or after the carbon adsorption step. Many other alternatives exist for processing a particular gold ore. *See* REFRACTORY.

Chemistry. Gold (Au) dissolution occurs electrochemically in a redox couple: at anodic sites on gold surfaces [reactions (1) and (2)] or at cathodic

$$\mathrm{Au(solid)} \longrightarrow \mathrm{Au^+(aqueous)} + e^- \quad (1)$$

$$\mathrm{Au^+(aqueous)} + 2\mathrm{CN^-(aqueous)} \longrightarrow \mathrm{Au(CN)_2^-(aqueous)} \quad (2)$$

sites on or near gold surfaces [reactions (3) and (4)].

$$\mathrm{O_2(gas)} + 2\mathrm{H^+(aqueous)} + 2e^- \longrightarrow \mathrm{H_2O_2} \quad (3)$$

$$\mathrm{H_2O_2} + 2\mathrm{H^+(aqueous)} + 2e^- \longrightarrow 2\mathrm{H_2O} \quad (4)$$

Dissolution rates depend on both cyanide (CN) and dissolved oxygen (O_2) concentration. If cyanide concentrations get low, the anodic reaction is rate limiting and the whole process becomes too slow. Depending on the ore, cyanide concentrations should be kept as low as possible without slowing rates. This helps prevent dissolution of impurities, such as mercury, that must be chemically removed downstream. If oxygen concentrations get too low, the cathode reaction is rate limiting, and the whole process likewise slows down. Sometimes normal stirring and flow is not enough, and compressed air or oxygen must be injected.

The sodium cyanide dissociates into sodium and cyanide ions [reaction (5)]; the cyanide ions react

$$\mathrm{NaCN} \longrightarrow \mathrm{Na^+} + \mathrm{CN^-} \quad (5)$$

with water (H_2O) to form cyanide gas [HCN; reaction (6)]. Protective alkalinity must be maintained

$$\mathrm{CN^-} + \mathrm{H_2O} \longrightarrow \mathrm{HCN(gas)} + \mathrm{OH^-} \quad (6)$$

(pH 10.5–11.5) to avoid deadly cyanide gas. The hydroxyl ion (OH^-) drives reaction (6) to the left if the pH is high enough, but too high pH causes destruction of CN^- and unnecessary expense.

Although cyanide is the dominant lixiviant, alternatives such as thiourea, thiosulfate, chlorine, bromine, and iodine have all been proven to work. Their advantages include lower toxicity, more favorable pH ranges, and stronger dissolving power for some refractory ores.

When pyrite is oxidized [one path is shown in reactions (7)–(9), any encapsulated gold is liberated

$$\mathrm{FeS_2} + {}^7/_2\mathrm{O_2} + \mathrm{H_2O} \longrightarrow \mathrm{FeSO_4} + \mathrm{H_2SO_4} \quad (7)$$

$$14\mathrm{FeSO_4} + 7\mathrm{H_2SO_4} + {}^7/_2\mathrm{O_2} \longrightarrow 7\mathrm{Fe(SO_4)_3} + 7\mathrm{H_2O} \quad (8)$$

$$7\mathrm{Fe_2(SO_4)_3} + \mathrm{FeS_2} + 8\mathrm{H_2O} \longrightarrow 15\mathrm{FeSO_4} + 8\mathrm{H_2SO_4} \quad (9)$$

and can be dissolved by cyanide. Bacteria (*Thiobacillus ferrooxidans*) in vats and heaps catalyze the conversion of ferrous (Fe^{2+}) to ferric (Fe^{3+}) ion in acidic sulfate [reaction (8)] solutions (pH < 6; the optimal growth rate is pH 2.0–2.5). Ferric ion is transported to the pyrite grain, which is oxidized [reaction (9)]. Ferric ion in the solution wetting the ore particles is the key to pyrite oxidation. *See* BIOLEACHING; ELECTROCHEMISTRY; METALLURGY; PH; PYRITE.

Environmental considerations. Government regulations and permit requirements are designed to protect the safety of individual workers, the public, and the environment. Mills must have spill basins. Any storage or tailings ponds must have fences and netting to protect migratory birds and other wildlife, and all accidental wildlife destruction must be reported. Cyanide cannot enter the environment; all cyanide must eventually be destroyed. Cyanide detoxification processes include the sulfur dioxide (SO_2)/air process, the ferrous sulfate process, the alkaline chlorination process, and the bacterial degradation process.

One problem is acid mine drainage. Bacterial action on even tiny amounts of pyrite or other sulfide minerals left in tailings generates acid [reactions (7)–(9)]. Drainage is further contaminated because the acid dissolves toxic metals. The whole process can continue long after the mine is closed. One answer is to continue operating water treatment plants, which typically use lime to neutralize the acid and precipitate metals. *See* LAND RECLAMATION.

Keith A. Prisbrey

Bibliography. R. W. Bartlett, *Solution Mining, Leaching, and Fluid Recovery of Materials*, 2d reprint ed., 1992; M. C. Fuerstenau and J. L. Hendrix (eds.), *Advances in Gold and Silver Processing*, 1989; P. Hayes, *Process Principles in Minerals and Materials Production*, 1993; J. C. Yannapoulos, *The Extractive Metallurgy of Gold*, 1991.

Goldhaber triangle

The phase space triangle, or Goldhaber triangle, corresponds to the kinematically allowed boundary for a high-energy reaction leading to four or more particles. In a high-energy reaction between two particles a and b yielding four particles 1, 2, 3, and 4 in the final state ($a + b \rightarrow 1 + 2 + 3 + 4$), it is convenient to consider the reaction in terms of the production of two intermediate-state quasiparticle composites x and y, which then decay into two particles each, as in expression (1).

$$\begin{array}{l} a + b \rightarrow x \qquad\qquad + y \\ \qquad\qquad\quad \hookrightarrow 1 + 2 \quad \hookrightarrow 3 + 4 \end{array} \tag{1}$$

Most high-energy interactions indeed proceed through such intermediate steps, in which, for specific values of the invariant masses $m_x = M_x^*$ and $m_y = M_y^*$, the quasiparticle composites may form resonances.

Kinematical limits. The kinematical limits in this representation are particularly simple, namely, they form a right-angle isosceles triangle. A Goldhaber triangle plot corresponds to plotting a point (m_x, m_y) for each event occurring in the above high-energy reaction. Because of the kinematical constraints, these points must all lie inside the triangle.

If one considers the general reaction given in Eq. (1), then the length of each of the two equal sides of the triangle is Q, defined in Eq. (2). Here W

$$Q = W - \sum_{i=1}^{4} m_i \tag{2}$$

is the total energy in the center of mass of particles a and b, and m_i, for $i = 1$ to 4, is the mass of the particles 1 to 4. Hence Q corresponds to the total kinetic energy available in the reaction. All quantities are in energy units of millions or billions of electronvolts.

The values of m_x and m_y run over the intervals $m_1 + m_2 \le m_x \le m_1 + m_2 + Q$ and $m_3 + m_4 \le m_y \le m_3 + m_4 + Q$, respectively. The effect of changing the incident momentum, and thus Q, is then simply to move the hypotenuse of the triangle, leaving the two sides, as well as the location of any resonances which may occur for certain mass values of the composites x and y, fixed.

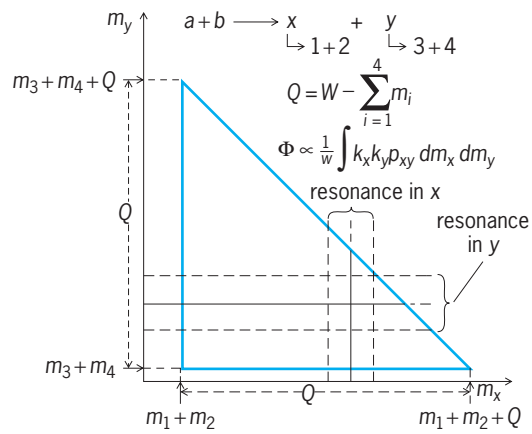

Definition of the kinematical boundary of the Goldhaber triangle for four particles.

In the triangle corresponding to the general reaction (see **illus.**), the vertical and horizontal bands indicate resonances at masses M_x^* and M_y^* with full width at half-maximum height or Γ_x and Γ_y, respectively. The bands shown of width 2Γ represent the regions usually chosen if the events corresponding to a given resonance are selected.

Phase space distribution. The phase space is given by $\Phi \propto 1/W \int k_x k_y p_{xy} dm_x dm_y$ where the integral extends over the triangle. Here k_x and k_y are the momenta in the center of mass of the composites x and y, respectively, and p_{xy} is the outgoing momentum of each of the composites in the overall center of mass of particles a and b. It is noteworthy that, along each of the three sides of the triangle, one of the factors in the integrand vanishes.

Comparison with Dalitz plot. Superficially there is a great similarity between the Dalitz plot and the Goldhaber triangle plot. There are, however, several important differences: (1) The Dalitz plot applies to three particles in the final state and corresponds to plotting m_y^2 versus m_x^2, however, in this case the two composites x and y have one particle in common. (2) The Dalitz plot has the advantage that the phase space distribution is uniform but the kinematical boundary is a more complicated function of the variables. In the Goldhaber triangle plot the phase space distribution is more complicated, but the kinematical boundary is very simple—a triangle. (3) If two resonances overlap on the plot, the interpretation is completely different. In the Dalitz plot the overlap corresponds to interference between the two resonances; in the Goldhaber triangle plot the overlap corresponds to double resonance formation, that is, both the x and y composites form essentially independent resonances at the same time. *See* DALITZ PLOT.

Gerson Goldhaber

Bibliography. R. L. Cool and R. E. Marshak (eds.), *Advances in Particle Physics*, vol. 2, 1968.

Golgi apparatus

An organelle found in all eukaryotic cells and absent from prokaryotes such as bacteria. The Golgi apparatus is named after the Italian histologist Camillo Golgi, who visualized an "internal reticular apparatus" by light microscopy in 1898. This organelle consists of flattened membrane-bounded compartments known as cisternae. In most cells, the Golgi cisternae are organized into stacks. Different cell types contain from one to several thousand Golgi stacks. The Golgi apparatus sorts newly synthesized proteins for delivery to various destinations, and modifies the oligosaccharide chains found on glycoproteins and glycolipids. *See* CELL ORGANIZATION.

Structure and location. The detailed morphology of the Golgi apparatus is variable depending upon the cell type. A Golgi stack may contain 4–30 or more cisternae. Usually the stacks are highly ordered (**Fig. 1**), but certain yeasts contain individual Golgi cisternae that are scattered throughout the cell. A typical cisterna measures about a micrometer in diameter. In vertebrate cells, tubules extend from the rims of the cisternae and connect the individual stacks to form a continuous ribbon. During vertebrate cell division, the Golgi ribbon breaks down and then reassembles. By contrast, plant cells contain isolated Golgi stacks that remain intact during cell division. A universal feature of the Golgi apparatus is the presence of multiple transport vesicles of 60–80 nanometers diameter surrounding the cisternae.

Golgi stacks are often located next to endoplasmic reticulum exit sites, which are specialized for protein export. In vertebrate cells, pre-Golgi intermediates are transported away from exit sites by movement along microtubules, thereby generating the Golgi ribbon near the nucleus. When microtubules in vertebrate cells are disrupted, the Golgi ribbon disintegrates to yield isolated Golgi stacks near exit sites.

Protein sorting and secretion. The Golgi apparatus acts at an intermediate stage in the secretory pathway. A subset of the proteins synthesized by the cell are inserted into the endoplasmic reticulum. Most such proteins are then delivered to the Golgi apparatus by means of coat protein II (COPII) transport vesicles, which form at endoplasmic reticulum exit sites (**Fig. 2**). Newly synthesized proteins traverse the Golgi stack until they reach the trans-most Golgi compartment, which is termed the trans-Golgi network to connote its extensive tubulation. The trans-Golgi network sorts the proteins into several types of vesicles. Clathrin-coated vesicles carry certain proteins to lysosomes. Other proteins are packaged into secretory vesicles for immediate delivery

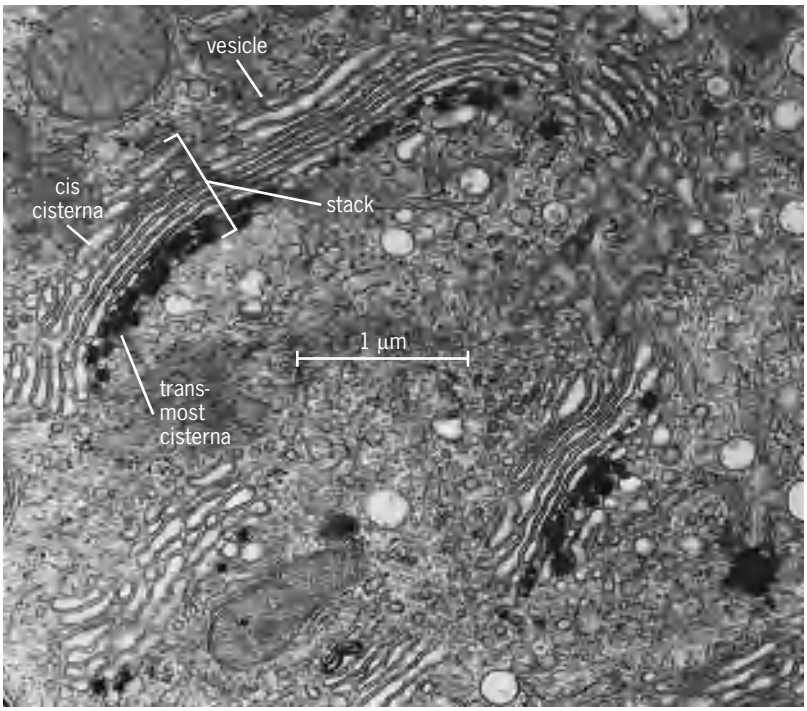

Fig. 1. Electron micrograph of the Golgi apparatus in a rat epididymis cell that was cytochemically reacted for thiamine pyrophosphatase using lead as a capturing agent. The trans-most cisternae are labeled with a black precipitate. (*Photo by Daniel S. Friend*)

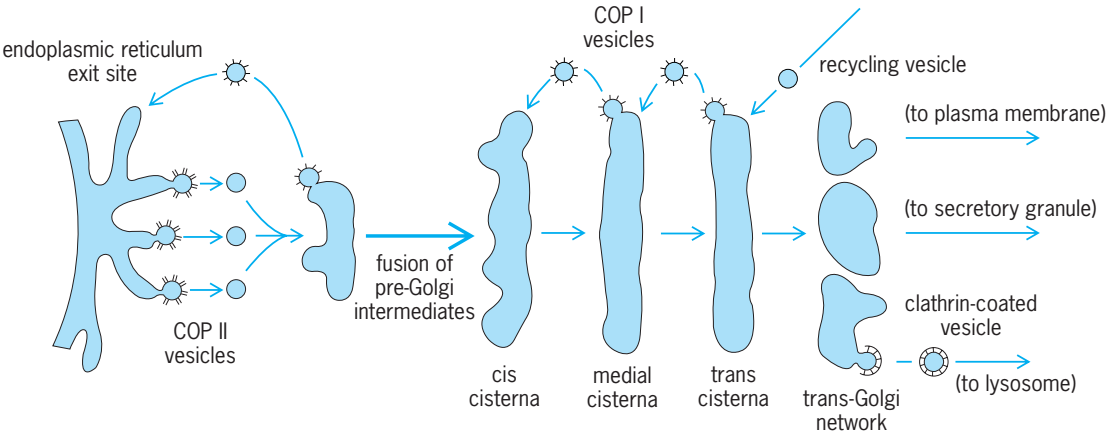

Fig. 2. Cisternal maturation model. At an endoplasmic reticulum exit site, COP II vesicles bud and then fuse with one another to generate pre-Golgi intermediates, which fuse in turn to generate a new cis cisterna. This cisterna then migrates through the stack to the trans face, maturing in the process. Maturation is driven by the recycling of resident Golgi enzymes in retrograde-directed COPI vesicles. At the trans-Golgi network stage, the cisterna is consumed by the formation of vesicles targeted to lysosomes, secretory granules, and the plasma membrane.

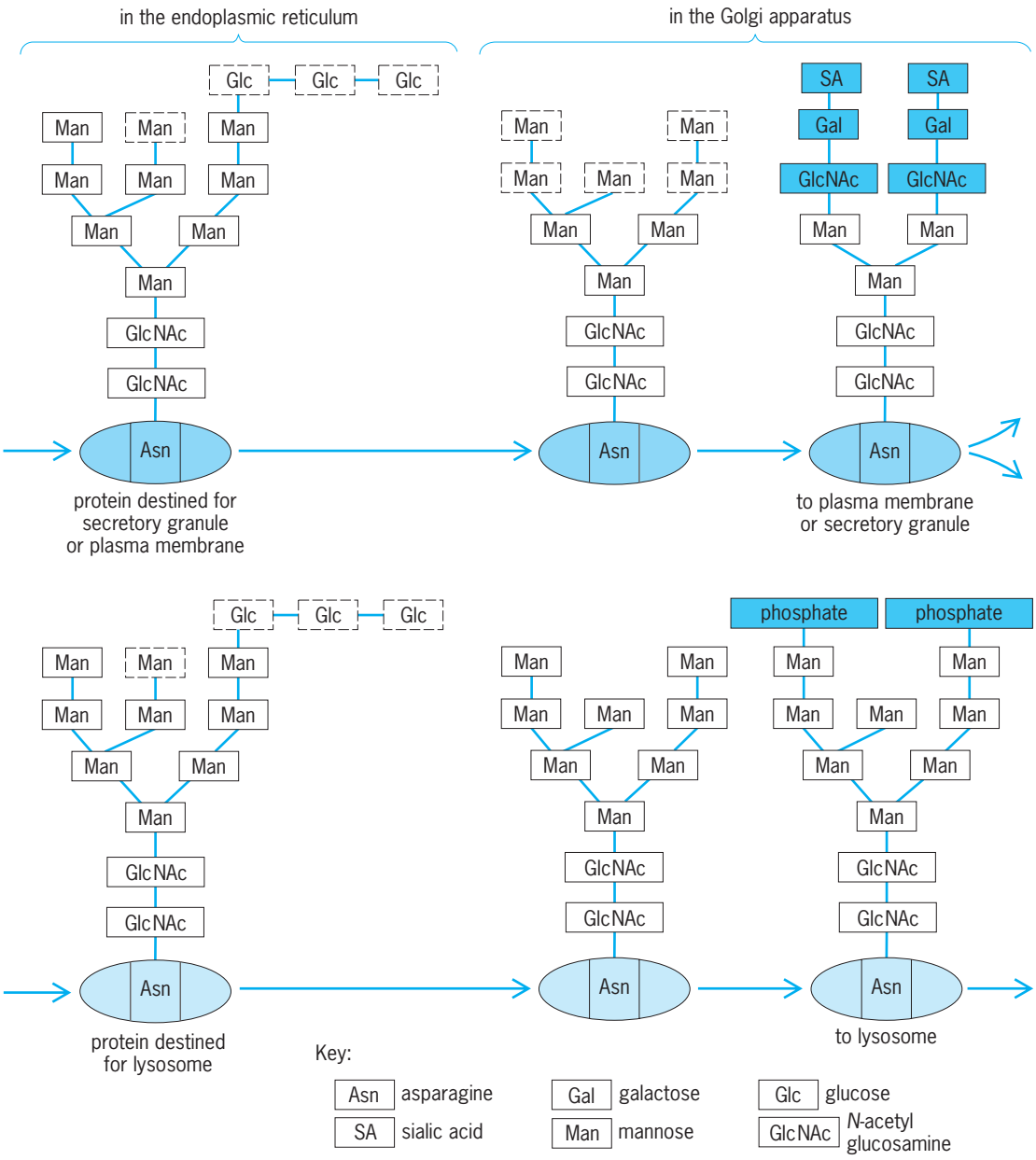

Fig. 3. Remodeling of asparagine-linked oligosaccharides by the vertebrate Golgi apparatus. Boxes with broken outlines represent sugar residues removed in the indicated organelle. (*After J. E. Rothman, Compartmental organization of the Golgi apparatus, Sci. Amer., 253:74–79, 1985*)

to the cell surface. Still other proteins are packaged into secretory granules, which undergo regulated secretion in response to specific signals. This sorting function of the Golgi apparatus allows the various organelles to grow while maintaining their distinct identities. *See* CELL MEMBRANES; ENDOPLASMIC RETICULUM; LYSOSOME.

Transport through the stack. Newly synthesized proteins enter the Golgi apparatus at the cis face and then travel through the medial and trans portions of the stack. The mechanism of protein transport through the Golgi stack is still being debated, but this process somehow involves coat protein I (COPI) vesicles. One theory holds that COPI vesicles carry newly synthesized proteins forward from one cisterna to the next. However, this model does not readily explain how proteins that are too large to fit inside a COPI vesicle can travel through the stack. Most researchers now believe that the cisternae themselves act as forward carriers. According to this cisternal maturation model (Fig. 2), a new cisterna forms at the cis face of the stack by the fusion of COPII vesicles. This cisterna then progresses through the stack to the trans face, where it breaks apart into various types of vesicles. Meanwhile, COPI vesicles move in the retrograde direction to recycle the resident Golgi enzymes. This recycling mechanism allows an individual Golgi enzyme to process a large number of substrate molecules. For example, a newly synthesized protein traverses the vertebrate

Golgi apparatus in about 20 min, whereas a typical Golgi enzyme remains in the organelle for many hours.

Compartmental organization. Each cisterna of the Golgi stack has a slightly different function. The major purpose of the trans-Golgi network is protein sorting. Meanwhile, the other cisternae function to process newly synthesized proteins and lipids. This processing is performed by enzymes that act in ordered pathways. The sequence of processing reactions is reflected in the distribution of Golgi enzymes across the stack. For example, enzymes that catalyze the initial processing steps are concentrated in the cis-most cisternae, whereas enzymes that catalyze the final processing steps are concentrated in the trans-most cisternae. As a result, the various cisternae of the Golgi stack are biochemically distinct (Fig. 1). According to the maturation model, as a cisterna progresses through the stack, it changes its composition by exporting one set of Golgi enzymes to younger cisternae while receiving a different set of Golgi enzymes from older cisternae.

Processing reactions. The best understood of the processing reactions carried out by the Golgi apparatus is the remodeling of oligosaccharides (chains of six-carbon sugars) that are attached to glycoproteins (**Fig. 3**). During insertion of a newly synthesized protein into the endoplasmic reticulum, one or more copies of a 14-sugar oligosaccharide may be attached to the amino acid asparagine at specific locations in the polypeptide chain. As the protein passes through the Golgi stack, the asparagine-linked oligosaccharides are modified to generate a diverse range of oligosaccharide structures. Additional oligosaccharides may become linked to the amino acids serine and threonine. Although the particular oligosaccharide modifications are quite different in animal, plant, and fungal cells, the Golgi apparatus always functions as a "carbohydrate factory." *See* OLIGOSACCHARIDE.

In vertebrate cells, olgiosaccharide remodeling can influence the fate of a newly synthesized protein (Fig. 3). If a protein is destined for either delivery to the plasma membrane or storage in secretory granules, several mannose units are removed from the oligosaccharide. Mannose removal paves the way for addition of variable numbers of *N*-acetylglucosamine, galactose, and sialic acid units. Mannose removal and *N*-acetylglucosamine addition occur in cis and medial cisternae, whereas galactose addition and sialic acid addition occur in trans cisternae. However, if a protein is destined for lysosomes, the mannose units are retained and are modified by the addition of phosphate. When a lysosomal precursor protein reaches the trans-Golgi network, the mannose-phosphate tags are recognized by a specific receptor, and the protein is packaged into a clathrin-coated vesicle for delivery to a lysosome.

The Golgi apparatus also carries out other processing events, including the addition of sulfate groups to the amino acid tyrosine in some proteins, the cleavage of protein precursors to yield mature hormones and neurotransmitters, and the synthesis of certain membrane lipids such as sphingomyelin and glycosphingolipids. *See* LIPID; PROTEIN.

Biochemical and genetic studies. Recent insights into Golgi apparatus function have come from analyzing the molecular components that drive protein transport through the secretory pathway. One experimental strategy was to develop a test tube system that reconstitutes transport through the Golgi stack. This biochemical approach identified COPI, as well as proteins that catalyze the fusion of vesicles with their target membranes. A complementary strategy was to generate yeast mutants with defects in secretion. This genetic approach identified COPII, as well as proteins that guide transport vesicles to their correct destinations. These molecular studies revealed that the principles underlying Golgi apparatus function have been conserved across eukaryotic evolution. *See* CELL (BIOLOGY); METABOLISM.

Benjamin S. Glick

Bibliography. E. G. Berger and J. Roth (eds.), *The Golgi Apparatus*, Birkhäuser Verlag, Basel, 1997; B. S. Glick and V. Malhotra, The curious status of the Golgi apparatus, *Cell*, 95:883–889, 1998; J. E. Rothman, Compartmental organization of the Golgi apparatus, *Sci. Amer.*, 253:74–89, 1985; J. E. Rothman and F. T. Wieland, Protein sorting by transport vesicles, *Science*, 272:227–234, 1996.

Gonorrhea

A common sexually transmitted disease caused by the bacterium *Neisseria gonorrhoeae*. Gonococci are gram-negative microorganisms that usually appear as pairs (diplococci). Humans are the only natural hosts for *N. gonorrhoeae*, which directly infects the epithelium of the mucous membranes of the human genital tract, pharynx, rectum, or conjunctiva. Local epithelial cell destruction usually occurs, but the organisms may spread to adjacent organs or disseminate via the bloodstream. In women, local complications include inflammation of the uterine lining (endometritis), inflammation of the fallopian tube (salpingitis), inflammation of the abdominal wall (peritonitis), and inflammation of Bartholin's glands (bartholinitis); in men, periurethral abscess and inflammation of a duct connected to the testes (epididymitis). Systemic manifestations such as arthritis or dermatitis may develop, and rarely endocarditis or meningitis.

Women are disproportionately affected by the complications of gonorrhea. Acute pelvic inflammatory disease and salpingitis, the most serious complications of gonorrhea, result in ectopic pregnancy and infertility. Gonococcal infection during pregnancy may also predispose women to premature rupture of membranes, delivery in less than full term, and postpartum endometritis. During childbirth, the gonococcus may infect the conjunctiva of the infant and result in the infection ophthalmia neonatorum. This infection is a serious complication that remains common in less developed countries and can lead to permanent damage to the eye and

blindness. *See* INFERTILITY; REPRODUCTIVE SYSTEM DISORDERS.

Epidemiology. Gonorrhea continues to be the most commonly reported communicable disease in the United States. The decline in the incidence of gonorrhea in the United States since 1984 has lagged behind that of many European countries such as Sweden where most cases are the result of importation. Data from developing countries are incomplete; however, gonorrhea rates are generally higher than those in western Europe or the United States.

Risk markers and risk factors for gonorrhea are similar to those for other sexually transmitted diseases. Risk markers for gonorrhea, that is, correlates of the probability of encountering an infected partner, include urban residence, young age, race, socioeconomic status, and marital status. Risk factors that may influence the probability of infection include number of sexual partners, lack of barrier contraceptives (for example, condoms), and young age. In the United States, over 71% of the cases reported in males and more than 82% of those in females occur between the ages of 15 and 29 years.

Gonorrhea is an infection spread by physical contact with the mucosal surfaces of an infected person, usually a sexual partner. There is no evidence that natural transmission occurs from toilet seats or similar objects. The risk of infection depends on the anatomic site, the amount of substance containing bacteria, and the number of exposures. Variations in host susceptibility have not been well defined. However, about one-third of men acquire the disease from a single exposure of vaginal intercourse with an infected female; after four exposures, the risk of infection increases to 60–80%. The transmission rate from an infected male to a female after a single exposure is thought to be 50–60%. In a small but significant proportion of infections, there are no symptoms. These individuals are important in the epidemiology of this disease because gonorrhea is usually spread by carriers who have no symptoms or have ignored symptoms.

Control and prevention. Control of gonorrhea depends on early diagnosis, effective treatment, and identification of asymptomatic individuals. The last has been accomplished, in part, through screening programs that have contributed to the detection and treatment of gonorrhea, particularly in women. However, complete control has not been possible because of the emergence and spread of strains that are resistant to less-expensive antimicrobial treatments such as penicillin and tetracycline.

There is no evidence that infected individuals develop long-lasting immunity to reinfection, and vaccination is not available. Thus, the prevention of gonorrhea relies on behavior modification and risk reduction, use of appropriate screening and diagnostic tests, routine use of highly effective antibiotics, early identification and treatment of sexual partners of individuals with gonorrhea, and the appropriate use of barrier methods such as condoms.

Clinical aspects. The usual incubation of gonococcal urethritis in the male is 2 to 7 days from the time of exposure; however, longer intervals are not uncommon. Acute gonococcal urethritis is characterized by discharge of pus, painful urination and swelling and localized redness of the urethral opening. A scant, almost nonpurulent discharge may develop in approximately one-fourth of men, and a few (5–10%) may never develop urethral discharge. Before antibiotic treatment became available, symptoms of urethritis persisted for an average of 8 weeks. Local complications of gonococcal urethritis, which include acute epididymitis, chronic inflammation of the prostate, inflammation of the lymph nodes, inflammation of the seminal vesicles, inflammation or abscess of Cowper's gland, and periurethral abscess, are uncommon occurrences in industrialized countries. Urethral stricture remains a common complication in developing countries.

In women, the glandular mucous membrane of the cervix (endocervix) is the primary site of urogenital infection. Symptomatic and asymptomatic urethral infection occurs in a substantial portion of women with endocervical infection. Although most women are asymptomatic, clinical signs usually develop within 10 days after infection and include increased vaginal discharge due to inflammation of the endocervix, painful urination, abnormal menstrual bleeding due to endometritis, and anorectal discomfort. Pelvic inflammatory disease, which occurs in approximately 15–20% of women with gonorrhea, is the most serious local complication. It typically results from the ascending spread of *N. gonorrhoeae* from the endocervix to the uterus and fallopian tubes. Symptoms may include fever, nausea, vomiting, lower abdominal pain, low back pain, and abnormal vaginal bleeding. Early treatment of acute pelvic inflammatory disease and salpingitis with antibiotics is important in order to prevent further complications. Chronic salpingitis may cause scarring and blockage of the fallopian tubes, which may result in infertility or tubal pregnancy.

Rectal infections, which tend to be asymptomatic, are present in as many as 50% of women with gonorrhea and are common in homosexual men. Infection of the pharynx may occur in approximately 20% of heterosexual women and homosexual men. Pharyngeal gonorrhea may produce an exudative tonsillitis, but characteristically there are no symptoms and the infection usually clears spontaneously over several weeks.

Gonococcal conjunctivitis is rare in adults and is probably due to self-inoculation; if untreated the infection may lead to ulceration of the cornea and blindness. Ophthalmia neonatorum can be prevented by prophylactic use of 1% silver nitrate eyedrops. Additionally, infants born to infected mothers may also develop infections of the pharynx or respiratory tract as well as disseminated infections due to *N. gonorrhoeae*.

Treatment. An increasing proportion of infections are due to antibiotic-resistant strains of

N. gonorrhoeae. Chromosomally mediated resistance to multiple antibiotics as well as plasmid-mediated resistance to beta-lactam antibiotics and tetracycline occurs in strains from both developed and developing countries. Nevertheless, infections can be effectively treated with third-generation cephalosporins (for example, ceftriaxone) or fluoroquinolones (for example, ciprofloxacin or ofloxacin). *See* SEXUALLY TRANSMITTED DISEASES.

Sandra A. Larsen; Stephen A. Morse

Bibliography. K. K. Holmes et al. (eds.), *Sexually Transmitted Diseases*, 3d ed., 1998; S. A. Morse et al. (eds.), *Atlas of Sexually Transmitted Diseases*, 1990; B. B. Wentworth et al. (eds.), *Laboratory Methods for the Diagnosis of Sexually Transmitted Diseases*, 2d ed., 1991.

Gonorynchiformes

A primitive order of ostariophysan fishes in which the first three vertebrae are specialized and associated with one or more cephalic ribs, thus representing a primitive Weberian apparatus (a series of bony ossicles that form a chain connecting the swim bladder with the inner ear). The order is further characterized by having the following features: an epibranchial (=suprabranchial) organ present, which consists of lateral pouches in the posterior part of the branchial chamber behind the fourth epibranchials; a small mouth and toothless jaws; no postcleithra and orbitosphenoid bones; and small parietal bones. The order comprises four families, treated below. *See* OSTARIOPHYSI; OSTEICHTHYES.

Chanidae (milkfishes). One extant species (see **illustration**) occurs in marine and brackish waters of the tropical and subtropical Indian and Pacific oceans and occasionally inhabits freshwater. The body is fusiform and compressed, and the mouth is terminal. The species feeds on benthic algae and invertebrates. It attains a maximum length of 1.7 m (5.6 ft) but usually is 1 m (3.3 ft). Milkfishes are valued food in Southeast Asia, where they are cultured extensively in ponds.

Gonorynchidae (beaked sandfishes). One genus and five species are found primarily in the Indo-Pacific region. These fishes possess an elongate body; an inferior mouth with a protrusible upper jaw; a sharply pointed snout with one barbel at the tip; and ctenoid scales on the head and body. Gonorynchidae is the most primitive teleost family with ctenoid scales. It attains a maximum length of 60 cm (24 in.). *See* TELEOSTEI.

Kneriidae (knerias). Four genera and 30 species inhabit the freshwaters of tropical Africa and the Nile River. The mouth is inferior or subterminal; the upper jaw is protrusible. Some species have cycloid scales and a lateral line, whereas in others (smaller species) the body is naked and lacks a lateral line. Maximum length is about 15 cm (6 in.).

Phractolaemidae (snake mudheads). These fishes comprise a monotypic (pertaining to a taxon that contains only one immediately subordinate taxon) family in the freshwaters of tropical Africa. They have an elongate body, small superior mouth, and a lunglike swim bladder that allows them to survive in poorly oxygenated waters. Maximum length is about 16 cm (6.3 in.). *See* SWIM BLADDER.

Herbert Boschung

Bibliography. R. Froese and D. Pauly (eds.), FishBase, World Wide Web electronic publication, <www.fishbase.org>, version 05/2005; T. Grande and F. Poyato-Ariza, Phylogenetic relationships of fossil and Recent Gonorynchiform fishes (Teleostei: Ostariophysi), *Zool. J. Linn. Soc.*, 125(2):197–238, 1999; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

***Chanos chanos*. (*Photo by Jack E. Randall with permission*)**

Gooseberry

A small fruit represented by about six species of the genus *Ribes* of the plant order Rosales. The gooseberry is a thorny, spreading bush which grows to a

Gooseberry branch bearing leaves and fruits. (*USDA*)

height of about 3 ft (0.9 m) and produces red, yellow, or green berries (see **illus.**). The most desirable hardier types in the United States are of American parentage, or are hybrids between American and European species. Commercial culture has declined and is limited to a few states, notably Oregon, Michigan, and Washington, but gooseberries are found in home gardens throughout most of the United States except the far South and Southwest.

The fruit is very acid and only a few European varieties, when fully ripe, are suitable for eating fresh. The fruit may be canned or frozen for use in pies or as preserves.

For diseases of gooseberry *see* CURRANT; ROSALES.

J. Harold Clarke

Gopher

The name for the North American rodents of the family Geomyidae. There are 39 species in 8 genera; 11 species are found in the United States. The distinctive name pocket gopher is applied to these animals, since they have large, fur-lined cheek pouches which open outward on the side of the face. Two common species are the northern pocket gopher (*Thomomys talpoides*), a small burrowing species, most numerous in the western United States and Mexico, and the prairie pocket gopher (*Geomys bursarius*).

Pocket gopher (*Thomomys talpoides*) of North America.

These are small to medium-sized animals with a broad, blunt head, small eyes and ears, and a short, thick tail which has very little hair (see **illus.**) and a tactile sensory function. The gopher's body is stout and its legs are short. The forelimbs of the animal have large feet, terminating in long, strong claws. Gophers have 20 teeth with the dental formula I 1/1 C 0/0 Pm 1/1 M 3/3. The teeth lack roots and grow continuously. One litter, which consists of 2–6 young, is raised each year.

These animals are vegetarians and, although they do not hibernate, hoard food in their burrows for the winter. Apparently they derive sufficient moisture from their food, because they have not been observed to drink. Gophers are pugnacious and defend themselves and their territory against their natural enemies, the hawks, owls, weasels, and the bull snake, and even against dogs and humans. *See* RODENTIA.

Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, The Johns Hopkins University Press, 1999; J. O. Whitaker, Jr. and W. J. Hamilton, Jr., *Mammals of the Eastern United States*, Cornell University Press, Ithaca, 1998; D. E. Wilson and S. Ruff, *The Smithsonian Book of North American Mammals*, 1999.

Gorgonacea

An order of the cnidarian subclass Alcyonaria. The Gorgonacea are the horny corals which often form fanlike or featherlike colonies with branches spread radially or oppositely in one plane. They attach to objects by somewhat enlarged bases of tufts of stolons. They are more widely distributed than the Alcyonacea and extend from the littoral zone to some great depth.

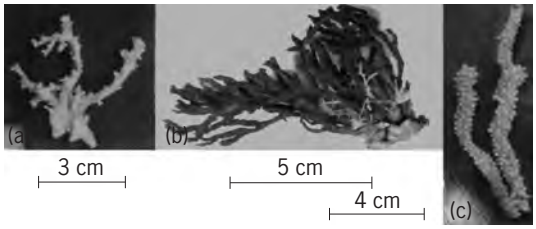

Fig. 1. Gorgonacea colonies. (*a*) Skeleton of *Corallium konojoi*. (*b*) *Melitodes* sp. (preserved specimen). (*c*) *Anthoplexaura dimorpha* (preserved specimen).

The gorgonians are characterized by the presence of an axial rod composed of horny matter or gorgonin. The order is divided into two suborders, the Scleraxonia and Holaxonia, according to the structure of the rod which has an outer cortex and an inner core, or medulla. In the Scleraxonia, which includes such genera as *Paragorgia, Corallium* (**Fig. 1***a*), and *Melitodes* (Fig. 1*b*),

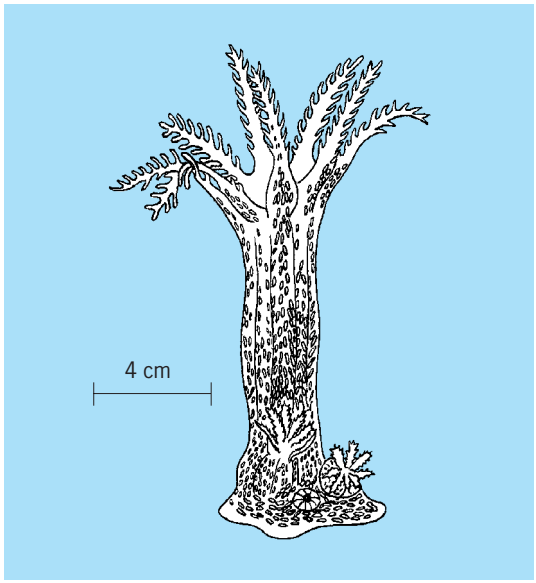

Fig. 2. Primary polyp of *Anthoplexaura dimorpha*.

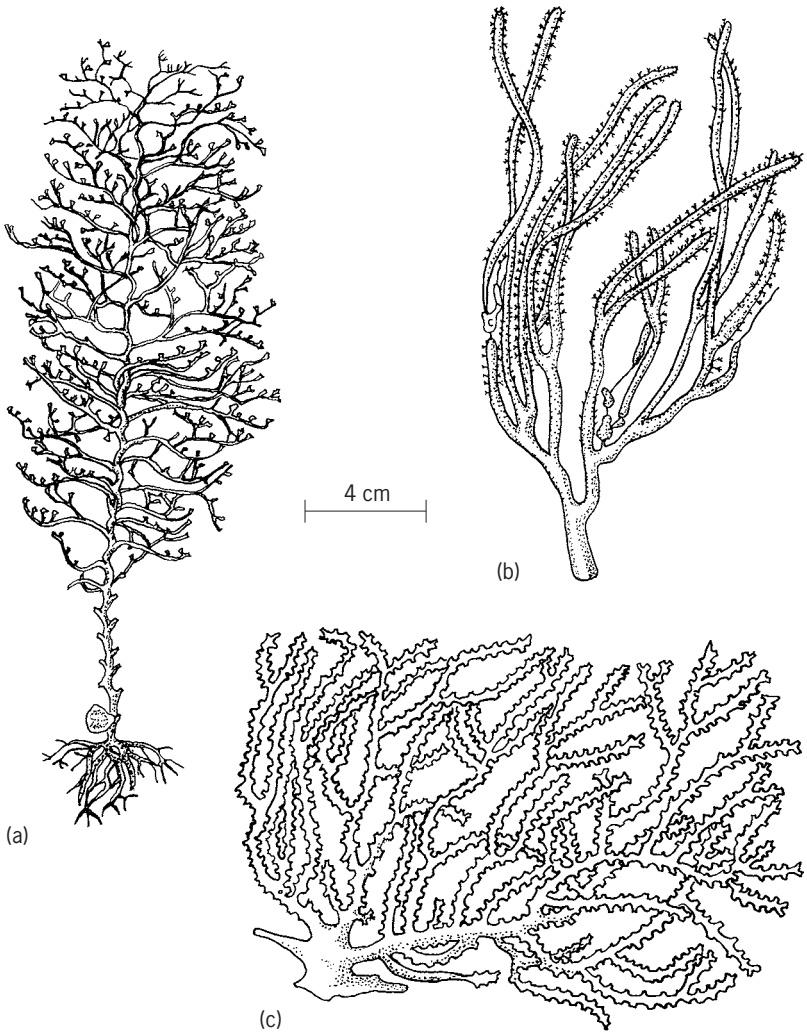

Fig. 3. Holaxonian gorgonians. (*a*) *Chrysogorgia flexilis* var. *africana*. (*b*) *Lophogorgia crista*. (*c*) *Paramuricea kükenthali*.

the axial skeleton contains calcareous spicules. In the Holaxonia, which are the typical gorgonians, a spiculeless skeleton is formed by the gorgonin and a small amount of calcareous matter. The latter group includes many genera: *Gorgonia* or the common sea fan, *Muricella, Anthoplexaura* (Fig. 1*c* and **Fig. 2**), *Chrysogorgia* (**Fig. 3***a*), *Lophogorgia* (Fig. 3*b*), *Paramuricea* (Fig. 3*c*), and *Keratoisis*. The last has a peculiar segmented axis with alternating calcareous and horny portions.

Most gorgonians are monomorphic. The autozooid (a feeding individual possessing fully developed organs and exoskeleton) resembles that of the Alcyonacea in many respects. Corallidae and *Paragorgia* are dimorphic; the siphonozooid has no tentacle. *See* OCTOCORALLIA (ALCYONARIA).

Kenji Atoda

Gout

A hereditary disease of ancient lineage that is caused by a derangement in purine metabolism, particularly uric acid synthesis. Uric acid is a normal breakdown product of purine metabolism and is derived from either ingested foods or body tissues. The disease is characterized by increased levels of blood uric acid (hyperuricemia), inflammatory arthritis, tophaceous deposits of urate crystals, renal insufficiency, and urolithiasis. Hyperuricemia is caused by increased production or decreased renal excretion of uric acid. *See* ARTHRITIS.

Gouty arthritis may be acute or chronic, but the basic pattern is one of acute arthritis in one joint followed by periods which are completely free of symptoms. These acute attacks are violent; and they frequently occur in conjunction with excessive intake of rich food and drink. The most frequent site of the initial attack is the big toe. Acute arthritis is characterized by the presence of birefringent monosodium urate crystals in synovial fluid leukocytes. Chronic gouty arthritis may be associated with extensive bone erosions and joint destruction.

Primary gout occurs most frequently in middle-aged males and is passed on as a familial or hereditary trait which for some unknown reason does not appear as often in females. Secondary gout refers to the disease when it is associated with some

underlying disorder causing an elevation in uric acid production, for example, myeloproliferative disorders.

The uric acid can be precipitated in soft tissues as localized deposits of chalky-white urate crystals called tophi. They occur in regions that are most susceptible to trauma, such as over bony prominences or adjacent to the small joints of hands and feet, in the cartilages of the ear, and in the kidney. Less commonly, they are found in eyelids, muscles, tendons, and heart valves. The mere presence of a tophus does not produce arthritis or pain, but there may be a surrounding inflammatory reaction.

Drugs that have been very helpful in treating gout include colchicine, which controls the acute arthritic attack; probenecid, which increases renal excretion of uric acid; and allopurinol, which inhibits xanthine oxidase, an enzyme required for uric acid formation. The clinical course is marked by long intervals of almost complete remission, but the disease flares up from time to time because of stress, infection, surgery, excessive alcohol intake, or unknown causes. Chronic gout causes functional impairment of the kidneys and greater incidence and earlier onset of atherosclerosis, but the general prognosis is good. *See* ARTERIOSCLEROSIS; PROTEIN METABOLISM; URIC ACID.

Robert Searles

Governor

A device used to control the speed of a prime mover. A governor protects the prime mover from overspeeding and keeps the prime mover speed at or near the desired revolutions per minute. When a prime mover drives an alternator supplying electrical power at a given frequency, such as 60 Hz, a governor must be used to hold the prime mover at a speed that will yield this frequency. An unloaded engine will fly to pieces unless its speed is under governor control. *See* PRIME MOVER.

Speed control. A governor regulates the speed of a prime mover by properly varying the flow of energy to or from it. In the case of gas and steam turbines and internal combustion engines, the fuel furnishes the energy to the prime mover. For such applications, the governor usually controls the speed of the unit by regulating the rate at which fuel, and hence energy, is furnished to the prime mover. The governor controls the fuel flow so that the speed of the prime mover remains constant regardless of load and other disturbances, or changes in accordance with such operating conditions as changes in speed setting.

For a diesel engine the governor is connected to the rack which controls the amount of fuel injected. A governor on a gas or gasoline engine is attached to the engine throttle. A steam turbine governor strokes the steam valve or valves which regulate the steam flow to the turbine. *See* DIESEL ENGINE; INTERNAL COMBUSTION ENGINE; STEAM TURBINE.

The output mechanism of a gas turbine governor is connected to the fuel valve with the stroke normally limited in each direction by the allowable combustion chamber temperature and other factors. If the fuel rate is too low, the fire will go out. Compressor stall must be avoided. To give rapid control, combustion chamber temperature is often computed in the governor by measuring other variables and applying the laws of thermodynamics. *See* GAS TURBINE.

A hydraulic turbine governor regulates the flow of water to the turbine by varying the openings of gates or other components. *See* HYDRAULIC TURBINE.

An aircraft propeller governor varies the pitch of the propeller to keep constant the speed of the engine attached to the propeller. This type of governor varies the load on the engine and thus controls the speed by regulating the energy flow from the prime mover. *See* PROPELLER (AIRCRAFT).

Ballhead governor. The speed of a prime mover is usually measured by a ballhead that contains flyweights driven at a speed proportional to the speed of the prime mover. The force from the flyweights is balanced, at least in part, by the force of compression of a speeder spring (**Fig. 1**). The upper end of this spring is positioned according to the speed setting of the governor.

The ballhead toes press against one end of a plunger. In the simplest version of a governor, the plunger position is a function of the engine speed as a result of the balance between the centrifugal and spring forces. The plunger is directly connected to the throttle or other energy controlling means of the prime mover. Because the governor output power is drawn directly from the speed measuring means, the output power and the precision of such a governor are severely limited.

Response. In automatic control theory, the input to the governor is taken to be the difference between the speed setting (reference) and the actual prime mover speed. This difference is the speed error. For the simplest governor, the position of the plunger is the output. The action of the governor is proportional since its output is proportional to its input. In equilibrium, where the engine is running in steady state, the speed of the prime mover depends on both the load and the speed setting. With the mechanical governor and a fixed speed setting, the equilibrium speed decreases as the load increases. A governor with this property is referred to as a droop governor,

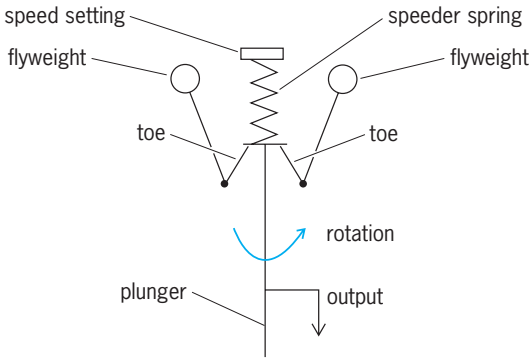

Fig. 1. Ballhead governor.

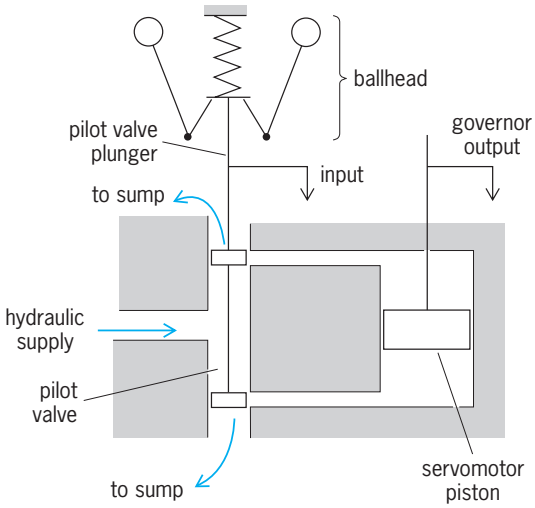

Fig. 2. Isochronous governor.

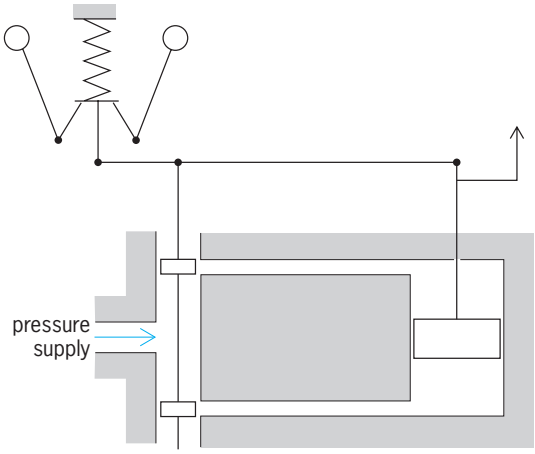

Fig. 3. Hydraulic governor.

or a governor operating on droop. The governor–prime mover unit is then said to be running on droop.

To increase the power output of a governor, a hydraulic amplifier is often employed. A governor that keeps the speed of a prime mover constant is said to be isochronous. In a simple isochronous governor, the ballhead senses the speed and strokes a pilot valve plunger that regulates the flow of fluid to a servomotor (**Fig. 2**). Normally, the fluid is oil. This governor is intended to bring the prime mover speed back to the speed setting after any change in load. *See* SERVOMECHANISM.

A hydraulic droop governor is obtained from the simple isochronous governor by introducing feedback from the governor output to the pilot valve (**Fig. 3**). This governor behaves like a simple mechanical governor except that a smaller ballhead is generally employed, greatly improving the precision of the governor by reducing hysteresis, dead band, and friction in the ballhead. The power output of the governor is much larger so that the effect of the load on governor performance incurred in moving the throttle or equivalent mechanism, is considerably diminished. A hydraulic governor may be sensitive to speed changes of as little as 1/1000 of 1%. For normal disturbances, prime mover speed error may be kept to 0.1% or better.

Use of dashpot. The performance of the simple isochronous governor is often greatly improved by the introduction of a dashpot in the feedback path from the output to the ballhead. If there is little damping in the prime mover, instability often occurs when the simple isochronous governor is used, whereas this instability is removed when the dashpot is incorporated. A system is stable when for each disturbance that dies out, the response of the system settles to an equilibrium condition. When instability occurs, unless some protection means is provided, the prime mover speed oscillates continuously or increases indefinitely until the unit breaks up.

When a dashpot is incorporated into a governor (**Fig. 4**), the governor output becomes a function of the speed error and the integral of this error. Such a governor action is referred to as a proportional plus integral controller. The velocity of the servomotor then depends on prime mover speed and prime mover acceleration. The time lag in the dashpot makes the governor sensitive to prime mover acceleration. As this lag is increased, the response of the governor to prime mover acceleration tends to increase. This increase in lag is accomplished by moving the dashpot needle valve, which controls the orifice area, toward the closed position.

Instead of mechanical feedback from the governor, force feedback is generally preferred (**Fig. 5**). An isochronous dashpot governor is turned into a droop governor by adding direct mechanical feedback from the servomotor to the ballhead.

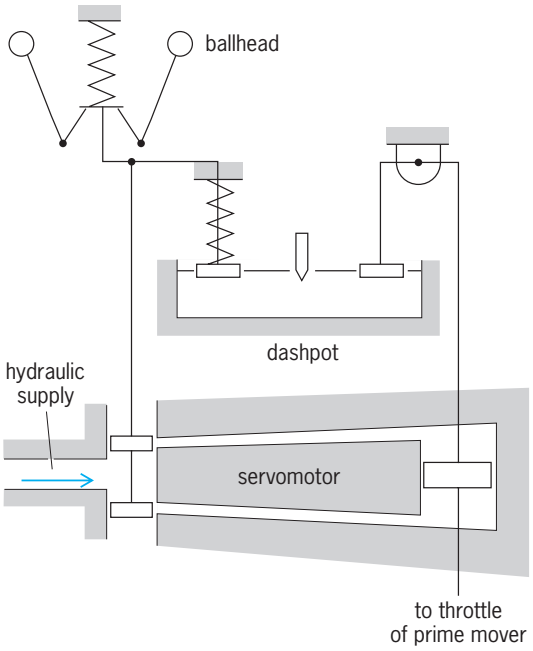

Fig. 4. Isochronous dashpot governor.

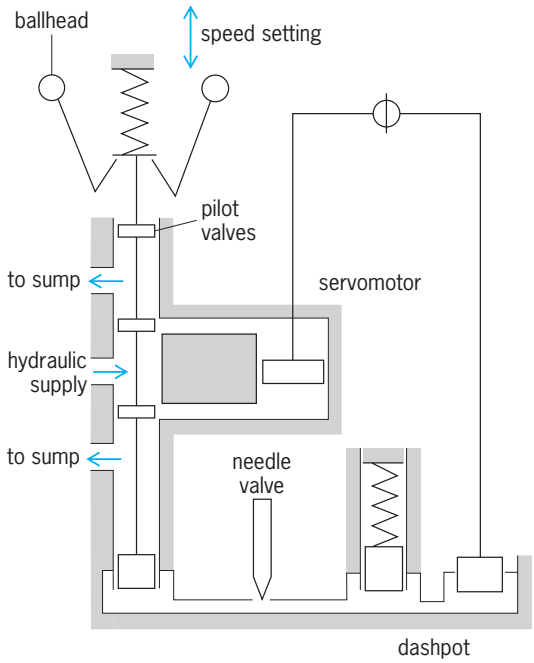

Fig. 5. Governor with force feedback.

Use of flywheel. Acceleration governors are sometimes used in place of governors with dashpots. In such governors a flywheel is employed instead of a dashpot. The prime mover drives the flywheel through a spring. This combination yields a motion proportional, except for a time lag, to the acceleration of the prime mover. *See* FLYWHEEL.

Applications. Governors for large hydraulic turbines require a second stage of hydraulic amplification. The governor servomotor piston is connected to the plunger of a relay valve that regulates the flow of oil to the turbine servomotors. The governor servomotor is then termed the controller. Turbine servomotors often require 10 hp (7500 W) or more.

The speed setting of a governor is often adjusted by an electric, hydraulic, or pneumatic motor as a function of auxiliary variables. Thus in an electrical power system the deviation of the frequency of the system from the desired value is used to position the governor speed setting so as to bring the frequency to the right value. Adjustment of governor speed settings is required to keep electric clocks on time. Oil and gas pipeline governors control pressure rather than speed. The pressure in the line is measured and the speed setting of the governor is adjusted accordingly. This affects the speed of the prime mover, which drives a compressor, raising or lowering the pressure.

Computer control. Traditional mechanical governor systems are being displaced by computer speed control systems. The desired controls are programmed directly into a computer, which reads a speed signal and provides for proportional/integral/differential (PID) control of speed by varying energy input to the prime mover. The computer-controlled governor can be programmed to provide functions similar to traditional mechanical governing systems but without the mechanical failures inherent in such systems. *See* COMPUTER; DIGITAL COMPUTER; DIGITAL CONTROL.

Paralleled prime movers. Prime movers may be paralleled to supply power to the same load. In an electrical power system, prime movers drive alternators electrically connected to the system. At most, one of the governors of paralleled prime movers can be isochronous; the rest must be on droop. The prime movers may all be on droop.

When two or more identical prime mover-generator units are paralleled, and one is controlled by an isochronous governor while the others are on droop, electrical coupling will force all units to run at the same steady-state speed. The alternators then supply electrical power to the electric grid at the same frequency.

Speed control by governor droop is an integral part of frequency control on an electric power system. Governors provide frequency control, and control how increases or decreases in load on the generators (generation pick-up) is shared among the remaining generators in the event of a loss of a generator (or how generation reduction is shared among generators in the event of a loss of load). System coordination of governor droop settings provides for equitable division of replacement generation until system automatic generation control computers assign a new generation dispatch. Plant control computers also participate in controlling the output of individual generators in a plant. In the North American power system, governor droop settings range between 4 and 7%. The droop setting means that for the stated percent frequency deviation the governor calls for the generator's maximum power. For example, on a 60-Hz system, if the droop setting is 5%, during a 5% (3-Hz) frequency deviation governors call for maximum power from the generators. Governors do not return the power system to 60 Hz; automatic generation control computers perform this function. *See* ELECTRIC POWER SYSTEMS.

Aircraft engines are synchronized and synchrophased by using one engine as the master, and adjusting the speed settings of the governors on the other engines to make them follow the master.

Design. Proper design of a governor involves making the governor characteristics fit those of the prime mover and load so as to give acceptable overall performance. Desirable and undesirable nonlinearities occur in governors to complicate design calculations. Thus centrifugal force varies as the square of speed. To compensate for this, nonlinear speeder springs are often employed. Bypass is often used in dashpots to limit the range of the dashpot output.

The output of a generator driven by the prime mover is sometimes used to obtain a voltage proportional to prime mover speed. This voltage, properly filtered, is fed to an electronic, magnetic, or other circuit, and eventually to a transducer to move the pilot valve. A gear with permanent magnet teeth

is often employed with a pickup coil, where the gear is driven at a speed proportional to that of the prime mover. The output pressure of a hydraulic pump is sometimes employed as a measure of prime mover speed. Other speed-measuring means are utilized.

Limit or topping governors are often employed to meet fail-safe specifications. The limit governor is usually a mechanical governor that takes over control from the main governor to shut the prime mover down if its speed reaches a fixed overspeed above its rated speed. The flyweight force increases with an increase in the distance of the flyweights from the axis of rotation of the ballhead; this flyweight force opposes the speeder spring force. The flyweight force increases as the square of prime mover speed. When it is sufficiently large, the governor gain becomes infinite, so that a small input to the governor causes an unlimited output. This causes a snap action of the governor at the overspeed for which the governor is set.

In the interests of economy the four-way valve servomotor (Fig. 2) is often replaced by a two-way valve-differential servomotor (**Fig. 6**). The area on one side of the servomotor is half the area on the other side. The supply pressure on the small area is sometimes replaced by a spring.

A load is termed isolated if the power supplying this load comes from one prime mover only. In the design of a governor for a prime mover supplying power to an isolated load, a prime mover differential equation is obtained relating the input to the prime mover, such as throttle position, to the speed of the prime mover, which is taken to be its output. This equation is obtained by mathematically balancing the torques on the prime mover shaft. The equation also involves a delay between a change in the input to the prime mover and the resulting change in the driving output torque. Part of the delay is a dead time. This dead time is normally 0.01–0.5 s. The rest of the delay is usually in the order of 0.1 s.

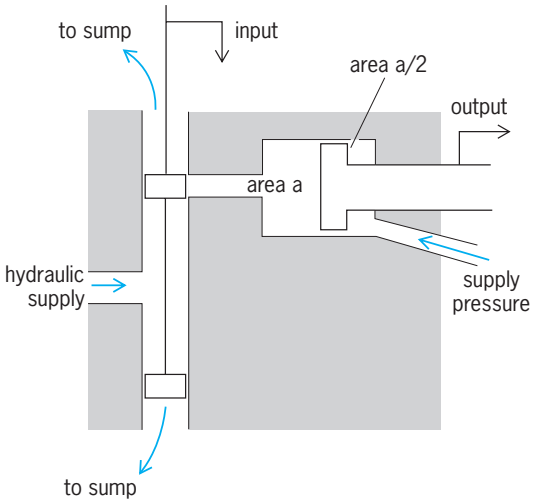

Fig. 6. Two-way valve-differential servomotor.

The nature of a prime mover may introduce other terms in the equation. For example, with hydraulic turbines the equation is complicated because of water hammer; a sudden closing of the gates to decrease the speed of the turbine causes the turbine to speed up initially, resulting in a correction opposite to that desired. As the gates are moved, pressure waves travel up and down the hydraulic conduit from the source of the water to the discharge from the turbine. These waves tend to destabilize the unit. *See* WATER HAMMER.

The equation of the prime mover is combined with the equation of the governor to yield an equation for overall system response from which the governor response is determined. When full load on a diesel engine is rejected, an overspeed of 3% is often considered acceptable with the response dying out in 1 s or less. For a hydraulic turbine this overspeed may be 30–40% and the response may endure as much as 10–20 s. Performance for other prime movers under governor control falls between these extremes. Rufus Oldenburger; Donald Davies

Bibliography. P. M. Anderson and A. A. Fouad, *Power System Control and Stability*, 2d ed., 2003; E. A. Avallone, T. Baumeister III, and A. Sadegh (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 11th ed., 2007; L. L. Grigsby (ed.), *The Electric Power Engineering Handbook*, 2001; H. A. Rothbart (ed.), *Mechanical Design and Systems Handbook*, 2d ed., 1985.

Graben

A block of the Earth's crust, generally with a length much greater than its width, that has been dropped relative to the blocks on either side (see **illus.**). The

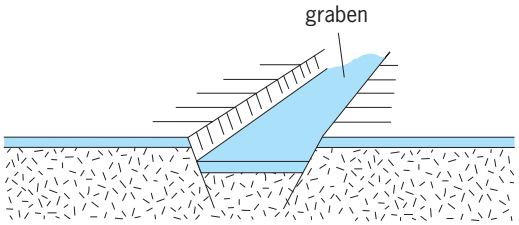

Diagram of simple graben. (*After A. K. Lobeck, Geomorphology, McGraw-Hill, 1939*)

size of a graben may vary; it may be only a few inches long or it may be hundreds of miles in length. The faults that separate a graben from the adjacent rocks are inclined from 50–70° toward the downthrown block and have displacements ranging from inches to thousands of feet. The direction of slip on these indicates that they are gravity faults. Graben are found in regions where the crust has undergone extension. They may form in the crests of anticlines or domes, or may be related to broad regional warpings. *See* EARTH CRUST; FAULT AND FAULT STRUCTURES; HORST; RIFT VALLEY. Philip H. Osberg

Gradient of a scalar

The result of the application of the distributive vector differential operator ∇, $\nabla = \mathbf{i}\,\partial/\partial x + \mathbf{j}\,\partial/\partial y + \mathbf{k}\,\partial/\partial z$, to a differentiable scalar function $S(x, y, z)$; thus $\nabla S = \mathbf{i}\,\partial S/\partial x + \mathbf{j}\,\partial S/\partial y + \mathbf{k}\,\partial S/\partial z$. The letters $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are symbols for the base vectors associated with x, y, z. The gradient of S is also denoted by grad S, while the symbol ∇ is usually called del and less frequently nabla.

The key properties of ∇S are implicit in its relation to the directional derivative of S, which is essentially the rate of change of S with respect to distance in a specified direction. To illustrate, if $S(x, y, z)$ gives the temperature at each point of a room, then from a given point of observation S may increase in certain directions (say toward a heat source) and decrease in others. Specifically, let P be a given point, C a curve given in terms of its arc length s by $\mathbf{r} = x(s)\mathbf{i} + y(s)\mathbf{j} + z(s)\mathbf{k}$ and further, let C pass through the point P and satisfy the requirement ($d\mathbf{r}/ds|$ at P) $= \boldsymbol{\tau}$ (a given prescribed unit vector). Along C, $x = x(s)$, $y = y(s)$, $z = z(s)$ and $S = S[x(s), y(s), z(s)]$ and if the functions are of class C'; then the equation below holds.

$$\left.\frac{dS}{ds}\right|_{P,C} = \frac{\partial S}{\partial x}\frac{dx}{ds} + \frac{\partial S}{\partial y}\frac{dy}{ds} + \frac{\partial S}{\partial z}\frac{dz}{ds}\bigg|_P$$
$$= \nabla S|_P \cdot \tau = |\nabla S|_P| \cos(\nabla S|_P, \tau)|$$

Here P, C indicates that the derivative is taken at the point P and for the curve C. This equation shows that $dS/ds|_{P,C}$ depends on S, P, and $\boldsymbol{\tau}$ only; hence, the alternative notation $dS/ds|_{P,\tau}$ is more appropriate. The quantity $dS/ds|_{P,\tau}$ is called the directional derivative.

Evidently, if (i) S and P are fixed, (ii) $\boldsymbol{\tau}$ is variable in direction, and (iii) $\nabla S|_P \neq 0$, then $dS/ds|_{P,\tau}$ can vary with the direction of $\boldsymbol{\tau}$ and will take on its maximum value when $\cos(\nabla S|_P, \boldsymbol{\tau}) = 1$. Thus, $\nabla S|_P$ points in the direction of the greatest rate of increase of S at P and $|\nabla S|_P|$ is this greatest rate of increase. Also, if $S(x, y, z) = S(a, b, c)$, (a, b, c fixed) defines a surface, then by confining C to the surface it may be shown that if $\nabla S|_P \neq 0$, then $\nabla S|_P$ is perpendicular to the surface. *See* CALCULUS OF VECTORS.

Homer V. Craig

Gradient wind

A hypothetical wind based upon the assumption that the sum of the horizontal components of the Coriolis acceleration and the atmospheric pressure gradient force per unit mass is equivalent to a wind acceleration which is normal to the direction of the wind itself (centripetal acceleration), with no viscous forces acting. The direction of the gradient wind is the same as that of the geostrophic wind. Its speed is determined by the equation below, where V_{geo} is

$$V_{\text{grad}} = \frac{V_{\text{geo}}}{\dfrac{1}{2} + \sqrt{\dfrac{1}{4} + \dfrac{V_{\text{geo}}}{2R\Omega \sin\phi}}}$$

the speed of the geostrophic wind, Ω is the angular speed of rotation of the Earth about its axis, ϕ is the latitude, and R is the radius of the curvature of the air trajectory, considered positive when the trajectory is curved in the cyclonic sense (center of curvature on low-pressure side) and negative when the curvature is anticyclonic. The gradient wind speed is less than the geostrophic speed when the air moves in a cyclonically curved path and greater when the air moves in an anticyclonically curved path. The gradient wind is a good approximation of the actual wind and is often superior to the geostrophic wind, particularly when the flow is strongly curved in the cyclonic sense. *See* CORIOLIS ACCELERATION; GEOSTROPHIC WIND; STORM; WIND.

Frederick Sanders

Grain boundaries

The internal interfaces that separate neighboring misoriented single crystals in a polycrystalline solid. Most solids such as metals, ceramics, and semiconductors have a crystalline structure, which means that they are made of atoms which are arranged in a three-dimensional periodic manner within the constituent crystals. Most engineering materials are polycrystalline in nature in that they are made of many small single crystals which are misoriented with respect to each other and meet at internal interfaces which are called grain boundaries. These interfaces, which are frequently planar, have a two-dimensionally periodic atomic structure. A polycrystalline cube 1 cm on edge, with grains 0.0001 cm in diameter, would contain 10^{12} crystals with a grain boundary area of several square meters. Thus, grain boundaries play an important role in controlling the electrical and mechanical properties of the polycrystalline solid. It is believed that the properties are influenced by the detailed atomic structure of the grain boundaries, as well as by the defects that are present, such as dislocations and ledges. Grain boundaries generally have very different atomic configurations and local atomic densities than those of the perfect crystal, and so they act as sinks for impurity atoms which tend to segregate to interfaces. *See* CRYSTAL DEFECTS; CRYSTAL STRUCTURE.

Structure. It was first believed that grain boundaries were made of an amorphous cement which held together the polycrystalline solid. With the advent of modern experimental probes of the atomic and defect structure of solids, such as electron microscopy and x-ray diffraction, it was determined that the grain boundary structure is frequently periodic in two dimensions. The geometry of a grain boundary is described by the rotation axis and angle, θ, that relate the orientations of the two crystals neighboring the interface, and the interface plane (or plane of contact) between the two crystals. Grain boundaries are typically divided into categories characterized by the magnitude of θ and the orientation of the rotation axis with respect to the interface plane. When θ is less than (arbitrarily) 15°, the boundary is called small-angle, and when θ is greater

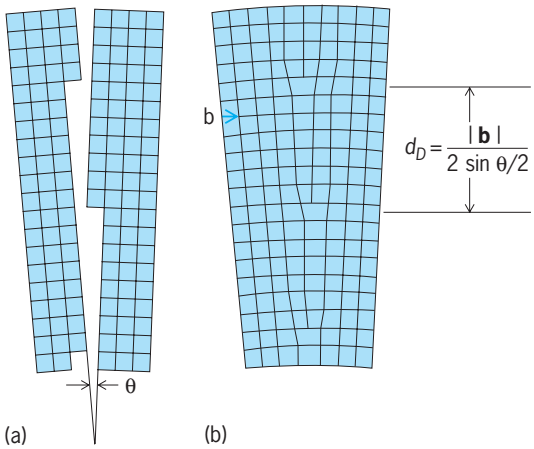

Fig. 1. The formation of a small-angle tilt grain boundary. (*a*) Unrelaxed configuration. (*b*) Relaxed configuration. (*After W. T. Read, Jr., Dislocations in crystals, McGraw-Hill, 1953*)

than 15°, the boundary is large-angle. *See* ELECTRON MICROSCOPE; X-RAY DIFFRACTION.

Small-angle boundaries. When the rotation axis is in the plane of the interface, the boundary is called a tilt boundary. **Figure 1** is a schematic diagram showing the formation of a small-angle tilt boundary in a simple cubic material. Simple cubic means that the atoms are located at the corners of a cube-shaped unit cell, which is translated in three dimensions to build up the crystal. The two crystals have been cut along a vertical plane, the interface plane. It is seen in Fig. 1*a* that the surface of the boundary-plane-to-be in each crystal is stepped in a periodic fashion on the atomic scale. When the two misoriented crystals are brought together and the atoms in the vicinity of the interface are allowed to move (relax) into lower energy configurations, then the result is as shown in Fig. 1*b*. It is seen that the boundary consists of a periodic array of edge dislocations with Burgers vector **b** and spacing d_D, which is given by Eq. (1). For small angles this equation becomes Eq. (2). Between the

$$d_D = \frac{|\mathbf{b}|}{2 \sin \theta/2} \tag{1}$$

$$d_D = \frac{|\mathbf{b}|}{\theta} \tag{2}$$

dislocations are regions of good cubic material. Because of the periodic nature of the boundary structure, the strain field associated with the boundary falls off exponentially (rapidly) along the direction normal to the interface. Thus the boundary region can be thought of as a thin layer with a thickness of $2d_D$ (or less), inside of which is disturbed material.

When the rotation axis is normal to the interface plane, then a twist boundary is present. **Figure 2** is a schematic diagram showing the formation of a small-angle twist boundary. In the unrelaxed configuration (Fig. 2*a*), it is possible to pick out regions of good cubic stacking (or good matching) and regions of bad cubic stacking (or bad matching). If the atoms at the interface are allowed to relax (Fig. 2*b*), then the regions of good match tend to expand, and all the mismatch is squeezed into narrow bands, which consist of a square array of screw dislocations with Burgers vector **b**, and spacing d_D given by the same expression obtained for tilt boundaries. **Figure 3** shows a transmission electron micrograph of a small-angle twist boundary in silicon, which contains a square array of screw dislocations, as predicted in Fig. 2*b*.

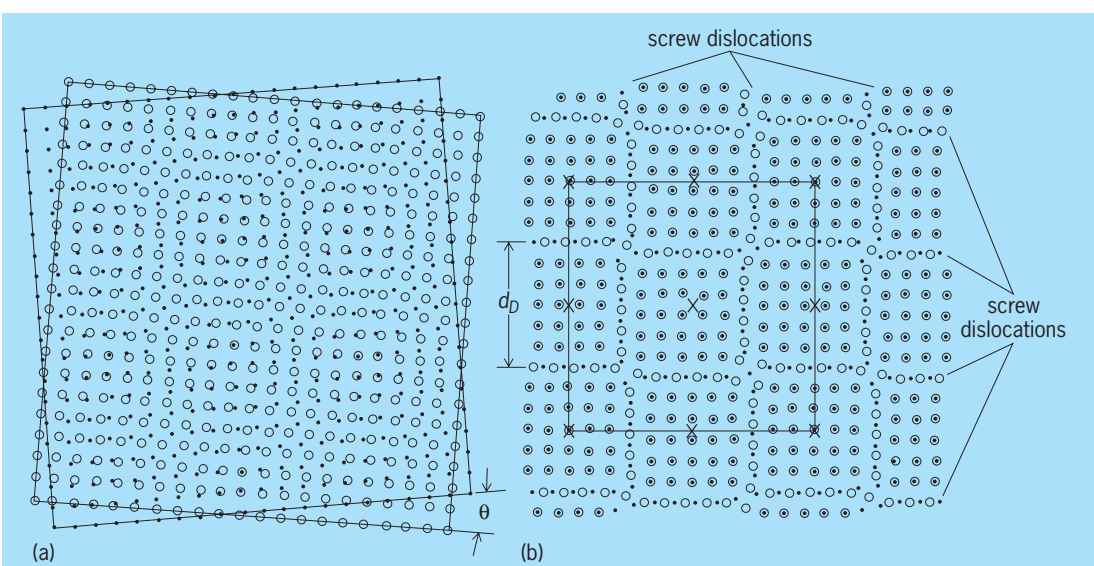

Fig. 2. Atomic configurations in the planes just above and below the boundary plane in a small-angle ⟨001⟩ twist grain boundary (*a*) Unrelaxed configuration. (*b*) Relaxed configuration, illustrating the formation of the screw dislocation network. The periodicity is that of the coincidence site lattice, the unit cell of which is indicated by the square. The crosses indicate the location of O-lattice points. (*After T. R. Schober and R. W. Balluffi, Quantitative observation of misfit dislocation arrays in low and high angle twist boundaries, Phil. Mag., 21:109–148, 1970*)

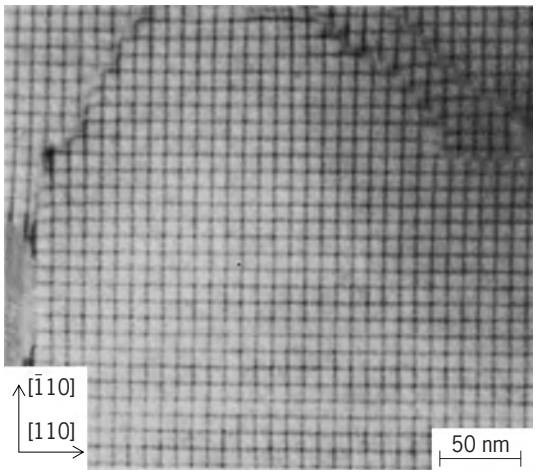

Fig. 3. Transmission electron micrograph, showing a small-angle ⟨001⟩ twist boundary in silicon. (*From H. Föll and D. Ast, TEM observations on grain boundaries in sintered silicon, Phil. Mag. A, 39:589, 1979*)

Large-angle boundaries. When the structures of the small-angle boundaries are extended to the large-angle regime, the expected dislocation spacing for a $\theta = 30^\circ$ boundary is $\approx 2\mathbf{b}$, which for most metals is 0.6 nanometer or less. It is unrealistic to postulate the existence of discrete dislocations at such small spacings, since their cores would overlap. A physically appealing approach to the structure of large-angle boundaries is to initially ignore their dislocation character and concentrate on their two-dimensionally periodic nature. This is done by using the concept of the coincidence site lattice (CSL).

The coincidence site lattice and its relation to grain boundary structure can be understood by using the structure of the small-angle twist boundary in Fig. 2*b*. At the corners of the outlined square, atoms in the two crystals are in exact match (or good coincidence); thus, these points define a simplified version of the coincidence site lattice. The plane lattice made up of the points determined by the coincident atoms in the two crystals gives the two-dimensional coincidence site lattice, and the number of such points expressed as a fraction of the lattice points in one plane is defined as $1/\Sigma$. The two-dimensional coincidence site lattice in the boundary plane shown in Fig. 2*b* describes the actual periodicity of the grain boundary structure, even when atoms move from their perfect crystal positions to take on a lower energy configuration.

There are additional locations, besides the points of the coincidence site lattice, where the two crystals are in good coincidence (or matching). These locations lie on a sublattice of the coincidence site lattice which is called the O-lattice, and it is easily seen that the two crystals can rotate with respect to each other at the O-lattice points (marked by crosses in Fig. 2*b*) to achieve good matching on a local scale. In the small-angle case, the O-lattice points are separated by bands of mismatch (with spacing d_D), which are screw dislocations possessing Burgers vectors of the perfect crystal (and hence are called primary dislocations).

For large misorientations and special angles θ_{CSL}, the dimensions of the unit cell of the coincidence site lattice can be relatively small. It is believed that boundaries with small unit cells and low values of Σ (see **table**) have low energy per unit area. **Figure 4***a* shows four unit cells of a $\Sigma = 13$ twist grain boundary in gold where the atom positions have been determined by x-ray diffraction techniques to an uncertainty of less than 0.01 nm. Careful examination of the structure shows that it can be modeled by using polyhedral arrangements of atoms, as illustrated for inclined octahedra in Fig. 4*a*. The three-dimensional atomic configuration determined by x-ray diffraction is shown in Fig. 4*b*. Arrays of octahedra, tetrahedra, and archimedean antiprisms are interwoven to build up the grain boundary structure. The atomic density across the actual interface plane is approximately 8% less than in the bulk material.

In the discussion of the structure of small-angle boundaries, it has been shown that local regions of good matching occur which are separated by dislocations possessing Burgers vectors of the perfect crystal. In a similar manner, it is expected that in boundaries with misorientations that deviate from the special values of θ_{CSL} there will be local rotations to achieve local regions of low-Σ structure, separated by grain boundary dislocations. These dislocations have Burgers vectors which are not present in the perfect crystal and, therefore, are called secondary dislocations, to distinguish them from the primary dislocations in Figs. 1*b*, 2*b*, and 3. The spacing d_x of the secondary dislocations with Burgers vectors $\mathbf{b_s}$ is given by Eq. (3). Figure 2*b* can also represent the

$$d_s = \frac{|\mathbf{b_s}|}{|\theta - \theta_{CSL}|} \tag{3}$$

structure of a large-angle twist boundary slightly deviated from θ_{CSL} with the regions of good matching replaced by regions of low Σ and primary dislocations replaced by secondary dislocations.

Boundaries with inclined rotation axes. The structures of tilt and twist boundaries are the easiest to visualize and thus have been the subject of most research. For the more general type of boundary, where the axis describing the misorientation between the two crystals is inclined to the interface plane, the boundary

Examples of special misorientation angles*

Misorientation angle (θ_{CSL})	Σ
22.62°	13
28.07°	17
36.87°	5

*For grain boundaries with a ⟨001⟩ misorientation axis, in the cubic system.

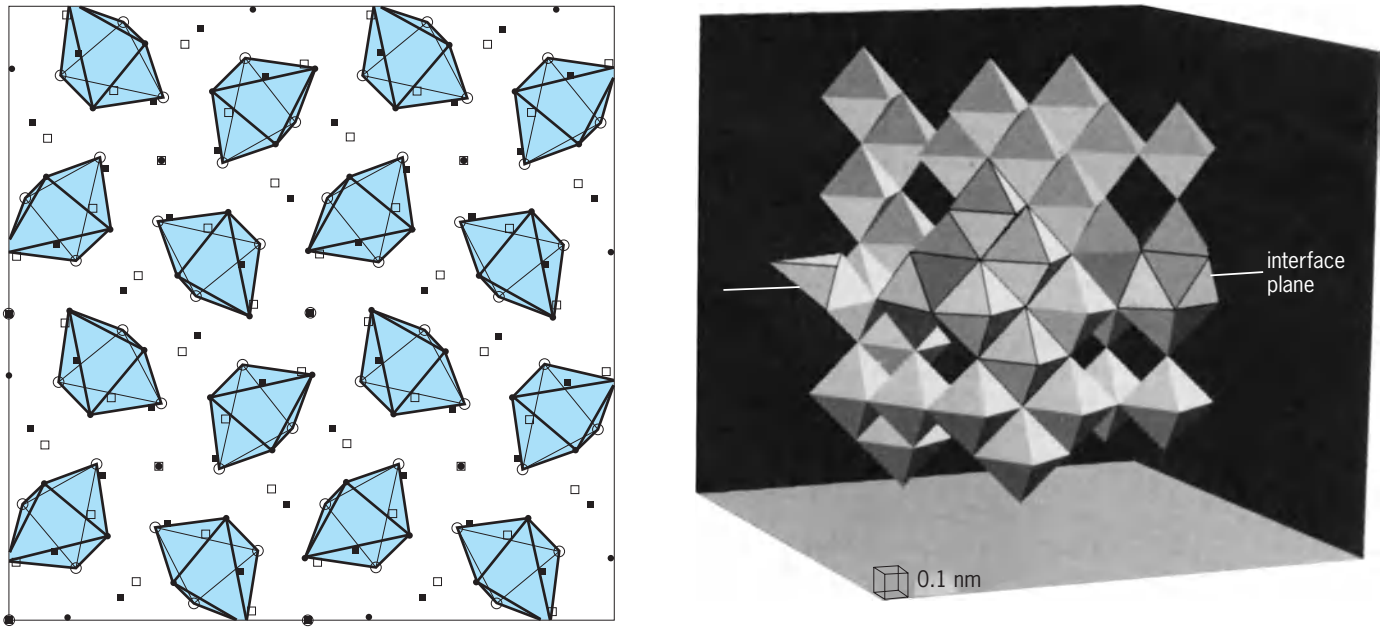

(a) (b)

Fig. 4. Structure of the $\Sigma = 13$ ($\theta = 22.62^\circ$) $\langle 001 \rangle$ twist boundary in gold. (*a*) Four unit cells, showing the polyhedral arrangements of atoms, in particular, inclined octahedra. (*b*) Three-dimensional perspective view of the different polyhedral arrangements of atoms that form the $\Sigma = 13$ boundary. (*After M. R. Fitzsimmons and S. L. Sass, The atomic structure of the $\Sigma = 13$ ($\theta = 22.6^\circ$) $\langle 001 \rangle$ twist boundary in gold determined using quantitative x-ray diffraction techniques, Acta Metallurg., 37:1009–1022, 1989*)

structure can be thought of as having tilt and twist components, and so, for example, in the small-angle case, will be made up of a mixture of edge and screw dislocations.

Influence on material properties. Because of the large differences in atomic structure and density between the grain boundary region and the bulk solid, the properties of the boundary are also quite different from those of the bulk, and have a strong influence on the bulk properties of the polycrystalline solid.

Mechanical properties. The mechanical behavior of a solid, that is, its response to an applied stress, often involves the movement of dislocations in the bulk, and the presence of boundaries impedes their motion since, in order for deformation to be transmitted from one crystal to its neighbor, the dislocations must transfer across the boundary and change direction. The detailed structure at the interface influences the ease or difficulty with which the dislocations accomplish this change in direction.

Since grain boundaries in engineering materials are not in a high-purity environment, the presence of impurities dissolved in the solid may have a strong influence on their behavior. The presence of one-half of a monolayer of impurity atoms, such as sulfur or antimony in iron, at the grain boundary, can have a drastic effect on mechanical properties, making iron, which is ductile in the high-purity state, extremely brittle, so that it fractures along grain boundaries. The segregation of the impurity atoms to the boundaries has been well documented by the use of Auger electron spectroscopy, and studies have led to the suggestion that the change in properties may be related to a change in the dislocation structure of the grain boundary induced by the presence of these impurities. *See* METAL, MECHANICAL PROPERTIES OF; PLASTIC DEFORMATION OF METAL.

Diffusion. The movement of atoms in and through solids is important for carrying out a variety of processes in the solid state. Among these processes are the motion of interfaces, which occurs during a variety of deformation and shaping treatments, and processes that occur during phase transformations, which are frequently used to enhance mechanical properties. Again, because the local atomic structure and density of the boundary region are different from that of the bulk, the movement of atoms will also be much different. Atoms move, or diffuse, much faster down grain boundaries than in the bulk. This enhanced motion of atoms down boundaries is called grain boundary diffusion, and is sometimes faster by as much as a factor of 10^6 over the bulk. Atoms also move through the crystal interior in a process called lattice diffusion. The amount of mass transport through a solid is the product of the velocity of individual atoms and the area available for the atoms to move through. At high temperatures, lattice diffusion tends to dominate since the cross-sectional area of the interior of the crystal is huge compared to the cross-sectional area of a grain boundary. At low temperatures, however, where lattice diffusion is slow, grain boundary diffusion tends to dominate. All these effects are of great importance to the electronics industry because, as part of the manufacture of electrical devices, thin films containing many grain

boundaries which serve as short-circuits for diffusion must be deposited. *See* DIFFUSION.

Electrical properties. Since modern electronic devices are fabricated from semiconductors, which may be polycrystalline, the presence of grain boundaries and their effect on electrical properties is of great technological interest. In a semiconductor such as silicon, the local change in structure at the interface gives rise to disruption of the normal crystal bonding, or sharing of valence electrons. One consequence can be the charging of the grain boundaries, which produces a barrier to current flowing across them and thus raises the overall resistance of the sample. This polycrystalline effect is exploited in devices such as zinc oxide varistors, which are used as voltage regulators and surge protectors. *See* SURGE SUPPRESSOR; VARISTOR.

Other devices, such as inexpensive solar cells, are also made of polycrystalline semiconductors. Grain boundaries can act as recombination centers in these cells, thus severely limiting their overall efficiency. The disruption of crystal bonding can encourage the segregation of impurities to the grain boundaries during crystal growth and subsequent processing. This effect (known as gettering) can be beneficial to solar cells because the crystalline areas are then left cleaner, so that the devices can be more efficient. *See* PHOTOVOLTAIC CELL; PHOTOVOLTAIC EFFECT; SOLAR CELL.

Stephen L. Sass

Bibliography. M. R. Fitzsimmons and S. L. Sass, The atomic structure of the $\Sigma = 13$ ($\theta = 22.6^\circ$) $\langle 001 \rangle$ twist boundary in gold determined using quantitative x-ray diffraction techniques, *Acta Metallurg.*, 37:1009–1020, 1989; P. E. J. Flewitt and R. K. Wild, *Grain Boundaries: Their Microstructure and Chemistry*, 2001; G. Gottstein and L. S. Shvindlerman, *Grain Boundary Migration in Metals: Thermodynamics, Kinetics, Applications*, 1998; A. Kelly, G. W. Groves, and P. Kidd, *Crystallography and Crystal Defects*, rev. ed., 2000; R. Raj and S. L. Sass (eds.), Interface Science and Engineering '87, *J. Phys.*, vol. 49, colloque C5, 1989; W. T. Read, Jr., *Dislocations in Crystals*, 1953; M. Ruhle et al. (eds.), International Conference on the Structure and Properties of Internal Interfaces, *J. Phys.*, vol. 46, colloque C4, 1985; T. Schober and R. W. Balluffi, Quantitative observation of misfit dislocation arrays in low and high angle twist boundaries, *Phil. Mag.*, 21:109–148, 1970; H. Zhang, *Grain Boundary Migration in Metals: Molecular Dynamics Simulations*, 2006.

Grain crops

Crop plants that belong to the grass family (Gramineae), generally grown for their edible starchy seeds. They also are referred to as cereal crops and include wheat, rice, maize (corn), barley, rye, oats, sorghum (jowar), and millet. The grain of all these cereals is used directly for human food and also for livestock, especially maize, barley, oats, and sorghum.

These large seeded grasses were among the first domesticated plants. Archeological studies indicate that some of these crops have been cultivated for 12,000–15,000 years. Cereal grains are the cheapest source of calories for human consumption and provide the most important energy source for three-fourths of the world population.

Cultivation. Wheat, barley, rye, and oats are cool season or temperate zone crops and are generally grown in low-rainfall areas (10–30 in. or 25–75 cm) because they are better adapted to cultivation under these conditions than other domesticated crops. However, they can produce large crops under higher rainfall, irrigation, and fertilizer applications. Rice is primarily a tropical or subtropical cereal, but Japanese plant breeders have developed types that grow at 45° latitude. Sorghum and some millets originally were tropical crops, but the area of adaptation has been greatly expanded by breeding new types. Maize also was a subtropical crop but is most productive in a temperate climate. Successful rice production is dependent on an abundance of water. Upland or dryland rice is available but limited in production. Some millets produce crops under dry, low-fertility conditions where other cereals often fail.

Storage. Another important attribute of these grain crops is the easy manner in which they can be stored. The grain often dries naturally before harvest to a safe moisture content (10–12%), or can easily be dried with modern equipment. Grain placed in adequate storage facilities can then be protected against insect infestations and maintained in sound condition for years. *See* CORN; BARLEY; MILLET; OATS; RICE; SORGHUM; WHEAT.

E. G. Heyne

Bibliography. J. E. Pratley, *Principles of Field Crop Production*, 3d ed., 1994.

Gram-equivalent weight

A quantity of a substance that contains the same number (known as the Avogadro number) of molecules as the number of atoms contained in exactly 12.000 g of carbon-12 (^{12}C). This convention stems from the concept that the central principle guiding chemical calculations is the relation of quantities of reacting substances to the numbers of molecules involved. *See* AVOGADRO'S NUMBER.

An added convenience in stoichiometric calculations is to incorporate the combining capacity (n), as well as the number of molecules, so that an equivalence of reacting substances is implicit without the need to examine balanced equations each time. Thus, the gram-equivalent weight of a substance is its gram-molecular weight divided by n. In acid-base reactions, n of the acid or base is given by the number of protons released or consumed in the reaction. For hydrochloric acid (HCl), ammonia (NH_3), acetic acid (CH_3COOH), and the acetate ion [CH_3COO; as in sodium acetate ($NaOOCCH_3$)], $n = 1$. For carbonic acid (H_2CO_3), sodium carbonate (Na_2CO_3), and ethylenediamine ($H_2NCH_2CH_2NH_2$), $n = 2$.

In precipitation and other metathetical reactions, the charge (valence) of the ion involved governs the value of n. Thus, $n = 3$ for ferric chloride ($FeCl_3$) and $n = 6$ for ferric sulfate [$Fe_2(SO_4)_3$] when precipitation yields either ferric hydroxide [$Fe(OH)_3$], barium sulfate, ($BaSO_4$), or silver chloride (AgCl). When oxidation-reduction is involved, the change of valence, rather than the valence itself, defines n. When the ferric ion (Fe^{3+}) acts as an oxidant [with the ferrous ion (Fe^{2+}) as product], $n = 1$ for ferric chloride ($FeCl_3$) and $n = 2$ for ferric sulfate. When potassium permanganate ($KMnO_4$) reacts in acid medium, $n = 5$, but in neutral or basic solution $n = 3$.

Further changes in the definition arise in dealing with metal complex formation, such as FeF_6^{3-}, $FeCl_4^-$, $Zn(EDTA)^{2+}$, and $Bi(EDTA)^-$; EDTA is the polydentate metal chelating agent ethylenediaminetetraacetate ion. The combining capacity of metals in complex formation depends on their coordination number, which, as these examples demonstrate, differs with different complexing agents, or ligands. *See* ELECTROCHEMICAL EQUIVALENT; EQUIVALENT WEIGHT; ETHYLENEDIAMINETETRAACETIC ACID; MOLE (CHEMISTRY); STOICHIOMETRY; VALENCE.

Henry Freiser

Gram-molecular weight

The molecular weight of an element or compound expressed in grams (g), that is, the molecular weight on a scale on which the atomic weight of the ^{12}C isotope of carbon is taken as 12 exactly. This replaces the earlier scale on which the atomic weight of oxygen was taken as 16.00 g. In the International System of Units, gram-molecular weight is replaced by the mole.

The ratio of the gram-molecular weights of any two elements or compounds must be identical with the ratio of the absolute weights of their individual molecules. Therefore, the gram-molecular weights of all elements or compounds contain the same number of molecules. This number, called the Avogadro number, N, is 6.022×10^{23}.

Since they contain the same number of molecules, the gram-molecular weights of all gases occupy the same volume at the same temperature and pressure. At 0°C and 1 atm (100 kilopascals) pressure this volume, called the gram-molecular volume, is 22.4 liters. *See* AVOGADRO'S NUMBER; GAS; MOLE (CHEMISTRY); MOLECULAR WEIGHT; RELATIVE MOLECULAR MASS.

Thomas C. Waddington

Grand unification theories

Attempts to unify three fundamental interactions—strong, electromagnetic, and weak—with a postulate that the three forces, with the exception of gravity, can be unified into one at some very high energy. The basic idea is motivated by the incompleteness of the electroweak theory of S. Weinberg, A. Salam, and S. Glashow, which has been extremely successful in the energy region presently accessible with the use of accelerators, and by the observation that the coupling constant for strong nuclear forces becomes smaller as energy increases whereas the fine-structure constant ($\alpha = 1/137$) for electromagnetic interactions is expected to increase with energy. The Weinberg-Salam-Glashow theory, which unifies electromagnetic and weak interactions, is incomplete in that strong interactions are not included; two coupling constants in the theory are unrelated; and in spite of many properties shared by them, quarks and leptons are unrelated, and their mass spectra and other properties remain unexplained. *See* GRAVITATION; STRONG NUCLEAR INTERACTIONS; WEAK NUCLEAR INTERACTIONS.

Models and successes. The simplest grand unification theory (GUT), proposed by H. Georgi and Glashow, is based on the assumption that the new symmetry that emerges when the three forces are unified is given by a special unitary group SU(5) of dimension 24. This symmetry is not observable in the low-energy region since it is badly broken. In this model, as in most GUTs, the coupling constants for the three interactions do actually merge into one at an energy of about 10^{14} GeV. Quarks (constituents of the proton, neutron, pion, and so forth) and leptons (the electron, its neutrino, the muon, and so forth) belong to the same multiplets, implying that distinctions between them disappear at the energy of 10^{14} GeV or above. In addition to the known 12 quanta of strong, electromagnetic, and weak interactions (namely 8 gluons, the photon, and the W^+, W^-, and Z^0 particles), there appear, in this model, 12 new quanta with the mass of 10^{14} GeV. These generate new but extremely weak interactions that violate baryon- and lepton-number conservation. The most spectacular prediction of GUTs is the instability of the proton, which is a consequence of baryon-number (and lepton-number) violation. In the simplest SU(5) model, the decay $P \rightarrow e^+ + \pi^0$ is a dominant proton-decay mode, with a predicted proton lifetime of $10^{29\pm2}$ years. The baryon-number violation can also induce neutron-antineutron oscillations. *See* GLUONS; INTERMEDIATE VECTOR BOSON; LEPTON; PHOTON; PROTON; QUARKS; SYMMETRY BREAKING; SYMMETRY LAWS (PHYSICS).

GUTs, in general, explain why the charge of the electron is precisely that of the proton with the opposite sign. In some GUTs with left-right symmetry, the lepton-number violation provides a novel way to generate the mass of neutrinos. Thus massive neutrinos are a distinct possibility in GUTs, and the smallness of their mass can also be understood. This leads to a possibility of neutrino oscillations which allow transformation of one neutrino species into another. Neutrinos with the appropriate mass may explain the missing mass in galaxies and clusters

and perhaps the formation of galaxies. The prediction of the SU(5) theory of the ratio of the two unrelated constants in the Weinberg-Salam-Glashow model agrees impressively with observation, and the ratios of certain lepton and quark masses can also be calculated with reasonable success. A characteristic of the SU(5) model is the existence of the so-called desert between 100 GeV, which is the mass scale of the Weinberg-Salam-Glashow model, and 10^{14} GeV. This desert may bloom in more general GUTs. *See* NEUTRINO.

When applied to cosmology, GUTs have many implications. According to the scenario based on the GUTs, the universe underwent a phase transition when its temperature cooled to 10^{27} K which corresponds to 10^{14} GeV in energy and to the first 10^{-35} s after the big bang. This was when, for example, the SU(5) symmetry would have been broken. The phase transition caused an exponential expansion (10^{30}-fold in 10^{-32} s) of the universe, which explains why the observed 3 K microwave background radiation is uniform (the horizon problem), and why the universe behaves as if space is practically flat (the flatness problem). Baryon-number–violating interactions that were efficient in the early universe can qualitatively account for the observed ratio of the baryon-number density n_B and the photon-number density n_γ in the present universe, $n_B/n_\gamma \simeq 10^{-10}$, which has been a long-standing puzzle in the standard hot big bang model of the universe. Another prediction common to many popular GUTs is the production of magnetic monopoles with a mass of 10^{16} GeV or less in the early universe. *See* BIG BANG THEORY; COSMOLOGY; PHASE TRANSITIONS; UNIVERSE.

Problems. In spite of its theoretical triumph and spectacular predictions, the simple SU(5) model is practically untested by experiment and appears to be incomplete or even incorrect. No experimental evidence of proton decay has been established. In fact, experiments have set a lower limit for the proton lifetime of 6×10^{31} years, which is somewhat larger than the SU(5) prediction. Although various modifications and generalizations of the simple SU(5) model can solve this problem, they do not, in general, provide significant improvements over the simple version. The problems which GUTS leave unsolved are numerous and include the following: (1) gravity is not included in the unification; (2) the number of unknown parameters turns out to be unsatisfactorily large because of the complexity of symmetry-breaking mechanisms; (3) the mass spectra of the leptons and quarks remain unexplained; and (4) it is difficult to understand why there exist at least two vastly different but stabilized energy scales, one at 100 GeV and the other at 10^{14} GeV.

Prospects. Perhaps a true unification theory including gravity will emerge when the nature of symmetry breakings, the possible substructure of the leptons and quarks, and implications of supersymmetry, which is a new kind of symmetry relating matter (quarks and leptons) and quanta (such as the photon and the W^+, W^-, and Z^0 particles), are explored and properly understood. Unification theories with supersymmetry, known as supergravity, have been favored as candidates for such a theory. *See* ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS; SUPERGRAVITY; SUPERSYMMETRY. Chung W. Kim

Bibliography. A. H. Guth and P. J. Steinhardt, The inflationary universe, *Sci. Amer.*, 250(5):116–128, May 1984; A. Linde, *Particle Physics and Inflationary Cosmology*, 1990; G. Ross, *Grand Unified Theories*, 1985; C. R. Storey, *Grand Unified Theory Made Easy*, 1993.

Granite

A crystalline igneous rock that consists largely of alkali feldspar (typically perthitic microcline or orthoclase), quartz, and plagioclase (commonly calcic albite or oligoclase). Its average grain size is 0.04–1.0 in. (1–25 mm); finer-grained rocks of this composition include rhyolite and aplite, and coarser-grained ones are granite pegmatite. *See* APLITE; PEGMATITE; RHYOLITE.

The revised nomenclature of the International Union of Geological Sciences (IUGS) subcommission defines granite as containing 80–100% by volume quartz, alkali feldspar, and plagioclase in specific proportions (see **illus.**), and 20–0% accessory minerals. The three essential minerals must include 20–60% quartz, and alkali feldspar must constitute 65–90% of the total feldspar. The variety alkali feldspar granite is similar except that alkali feldspar constitutes 90–100% of its total feldspar. The most common accessory minerals are biotite (typically forming 5–15% of the rock), iron-titanium (Fe-Ti) oxides (usually traces to 2% of magnetite or ilmenite), and traces

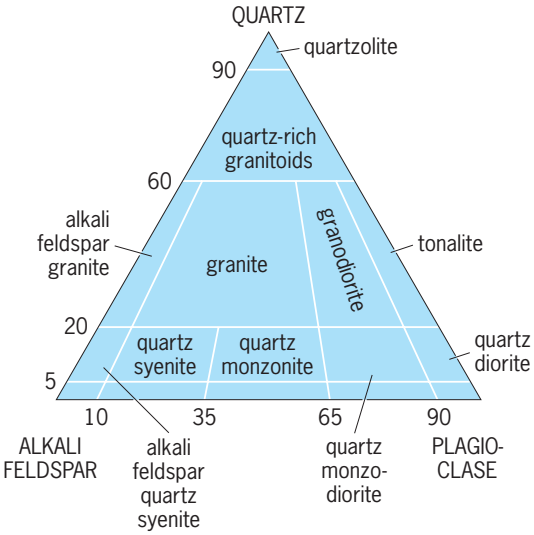

Classification of granitic rocks by the IUGS scheme. All proportions are by volume percentages.

of apatite and zircon. Hornblende or muscovite also are common accessories. The term granitic rocks includes granodiorite and tonalite as well as granite, and as used by some geologists may include quartz syenite to quartz diorite (see illus.). *See* BIOTITE; FELDSPAR; GRANODIORITE; HORNBLENDE; MUSCOVITE.

Texture. Granites range widely in texture. Some are massive or without structure, and all minerals are randomly oriented. Others show well-oriented tabular feldspars or flaky biotite grains, and are termed foliated or gneissic. Some show uniform grain size, some are seriate, and others exhibit phenocrysts (especially of microcline or orthoclase, rarely of quartz) one to several centimeters long set in a finer-grained matrix. *See* PHENOCRYST.

Occurrence. More than 99% of Earth's granites were emplaced in continental crust. The minor siliceous [silicon dioxide (SiO_2) $\geq$ 70 wt % of the rock] intrusive rocks of the ocean basins—as found in island arcs or mid-ocean ridges—typically are tonalite and contain little or no alkali feldspar. About 80% of all granitic rocks were formed in Precambrian times, and thus are older than 570 million years. Perhaps half of these older rocks are found in Archean (older than 2.5 billion years) greenstone-granite terranes (for example, in northern Ontario and Quebec). The remaining 20% are widely scattered through terranes of Phanerozoic (Paleozoic, Mesozoic, and Cenozoic) age.

Types. Granites may be divided into three major types: calc-alkaline, peraluminous, and alkaline. The peraluminous type is found with either of the other types in many regions. Each type has characteristic chemical and mineralogic features that are related to the geologic environment of its enclosing rocks and to its mode of origin.

Calc-alkaline. These granites, which include 70–80% of all granites, are the major type in orogenic belts of about 2 billion years to modern age, where oceanic crust is known to have been subducted at the margins of continental plates. They are also prominent in Archean crust, whose plate-tectonic mechanism of growth is controversial, but which may be like that of younger crust and occur by accretion of oceanic and other rocks at convergent margins.

Calc-alkaline granites typically are biotite or biotite-hornblende granites, some contain augite, and sphene is a common accessory. This type of granite tends to be more sodic [sodium oxide (Na_2O) = 3.3–5 wt %] than other types, has relatively low initial strontium-87/strontium-86 ($^{87}Sr/^{86}Sr$) isotopic ratios (0.7004–0.7008 for Phanerozoic ones and 0.7001–0.704 for Precambrian ones), and shows primary oxygen-18 ($\delta^{18}O$) values of 7–10 per mille. Because their parental rocks are largely of igneous type, calc-alkaline granites often are termed I-type granites. Those of the oceanic-continental convergent margins occur mostly in continental-margin batholiths, which consist of many tens to several hundreds of individual intrusive bodies of ovoid to irregular plan and of a few to several tens of kilometers in the maximum horizontal dimension. The individual intrusives of these batholiths commonly range from quartz diorite to tonalite, granodiorite, and granite.

Large continental-margin batholiths, like the Sierra Nevada batholith of the Sierra Nevada Range, California, may be several hundred kilometers long by 47–60 mi (75–100 km) broad. They form in the crust directly above subduction zones. Hot (2000–2300°F or 1100–1250°C) basaltic or andesitic magma that originates in the mantle at or above the subducted oceanic plate rises and pools in the lower or intermediate crust below the site of the batholith. The crust at such continental margins ranges from accreted oceanic rocks (basaltic, andesitic, and dacitic volcanic rocks, pelagic sediments, and ophiolites) to continental-margin sediments (especially turbidites) and continental crust (mafic to granitic igneous rocks, metamorphic rocks, and sedimentary rocks). The mantle-derived liquid directly melts and incorporates crustal rocks, while simultaneously precipitating minerals such as pyroxene, hornblende, or plagioclase. By this process the composition of the liquid changes; it becomes more siliceous and of dioritic to tonalitic composition. At the same time the mantle liquid in this lower-crustal chamber heats the enclosing rocks, which partially melt to granodioritic and granitic compositions. At a critical stage the primary magma body breaches its roof and rises as a discrete unit or diapir, and is emplaced in the upper crust. Afterward, the melted portions of the enclosing rocks accrete at the site of the former chamber and then rise up the conduit breached by the primary diapir. Thus magmas of a wide range of compositions are formed from both mantle and lower or intermediate crust. *See* CONTINENTAL MARGIN; DIAPIR; EARTH CRUST; OPHIOLITE; PYROXENE; TURBIDITE.

Calc-alkaline granites and granodiorites also form 25–50% of the granitic and gneissic rocks that contain the scattered Archean greenstone belts. These belts probably are accretionary collages of intra-oceanic island arcs; of submarine fans of graywacke and shale derived from nearby continental masses; of continental-shelf quartzite, carbonate rocks, volcanic rocks, graywacke, and other rocks; of fragments of older continental rocks; and of other rocks. These belts range from 3 to 9 mi (5 to 15 km) thick and cover 4000 to 12,000 mi^2 (10,000 to 30,000 km^2). During or after accretion the greenstone belts were folded and tectonically thickened, and they were intruded by basaltic magma from the underlying mantle. This hot magma was emplaced mostly in the lower crust, which it partially melted, and whose various rocks gave magmas that ranged in composition from tonalitic to granodioritic and granitic. These magmas rose gravitationally and were emplaced in the intermediate to upper crust. *See* GRAYWACKE; PLATE TECTONICS; SHALE; SUBMARINE.

Peraluminous. These granites contain aluminum in excess of that in feldspars. Biotite is the most

common accessory mineral. Highly peraluminous granites, known as S-type granites, also may contain muscovite (biotite-muscovite granites are called two-mica granites), garnet, cordierite, sillimanite, and other aluminous minerals. Peraluminous granites are of two general types: one is found in or above accreted masses of graywacke, pelite, and minor basalt, pelagic sediments, and other rocks (that is, accretionary prisms) and consists largely of biotite granite and granodiorite; the other is associated with metamorphosed shale or pelite and is of the highly peraluminous two-mica type. Biotite granite and granodiorite of the first type are common in Archean to modern crust, whereas the biotite-muscovite type is known from the Precambrian but mostly is found in Paleozoic, Mesozoic, and even early Cenozoic fold belts. They occur in relatively heterogeneous plutons of lenticular or elongate plan, of a few to as much as 60 mi (100 km) in maximum dimension, and oriented parallel to regional fold trends. *See* CORDIERITE; GARNET; PLUTON; SILLIMANITE.

Peraluminous granites typically show high ratios of K/Na and Rb/Sr, high abundances of U and Th, initial $^{87}Sr/^{86}Sr$ ratios higher than about 0.710, and $\delta^{18}O$ values greater than 10 per mille. Their mode of origin remains a matter of some debate: one school holds that deep folding and thickening of graywacke or pelite heats them so that 10–50% melting occurs. The magma produced is a mixture of this melt and of pelitic xenoliths and xenocrysts that may be carried in suspension during ascent and emplacement of the magma. In some fold belts other stratified rocks, such as graywacke, arkose, or rhyolitic volcanic rocks, may be interlayered with pelite. Partial melting of such rocks tends to give calc-alkaline liquids, so that magma formed from such a mixed source will give granite intermediate in character between peraluminous and calc-alkaline types. Another school contends that as an oceanic plate is subducted at a continental margin, its covering of turbidites, other sediments, and minor igneous rocks is offscraped to give an accretionary prism. This prism, which may have a maximum thickness of 20–30 km, thus consists largely of graywacke and pelite. Its composition, position just above the Moho, and large mass make it a ready source for the generation of granite. Percolation of very hot basaltic liquid into such a prism from a leaky, underlying and subducting plate may account for many peraluminous granites.

Alkaline. Subalkaline to alkaline granites are characterized by iron-rich mafic minerals and relatively sodic alkali feldspar [molecular Na/(Na + K) from 0.3 to 0.5; K = potassium]. The subalkaline type typically contains ferruginous biotite or hornblende, or both, but varieties containing ferrohedenbergite or fayalite are not uncommon. Allanite and zircon are common accessories. The alkaline type contains the Na-Fe minerals aegirine or riebeckite-arfvedsonite, or both, and may also contain ferruginous biotite or even astrophyllite, eudialyte, or other rare minerals. Both subalkaline and alkaline types show 5–6 wt % K_2O, 3–4% Na_2O, low ratios of Fe^{3+}/Fe^{2+}, magnesium oxide (MgO) less than 0.1%, high abundances of F and light rare-earth elements [lanthanum (La), cerium (Ce), neodymium (Ne), samarium (Sm)], variable initial $^{87}Sr/^{86}Sr$ ratios, and low contents of water. These granites are found in continental interiors, in most instances along rift zones at moderate distances (60–500 mi or 100–800 km) from the continental margin. Alternatively, some of these granites also may be found along traces of hot spots, or ascending plumes of basaltic magma derived from the deep mantle (for example, the linearly disposed Mesozoic granitic intrusives of Nigeria). Minor dikes or plutons of diabase or gabbro, nepheline syenite, quartz syenite, and anorthosite are associated in many instances, and with the granites form a suite of genetically related rocks. The origin of this suite involves source material from both mantle and crust. During rifting or plume formation, basaltic magma is produced deep in the mantle. Some of this liquid may be extruded or emplaced in the upper crust as dikes, but a major part of it pools in the lower or intermediate crust. Here, as in the model of origin of continental-margin batholiths given above, this very hot liquid reacts with and partially melts the crustal rocks—many of which in this continental crust are granitic—and assimilates a low-melting or granitic component. Liquidus minerals, especially clinopyroxene and plagioclase, and refractory minerals of the crustal rocks are precipitated at the same time. In this process the liquid becomes dioritic, syenitic, and ultimately granitic, and precipitation gives pyroxenite, anorthosite, and mixed rocks. The syenitic and granitic magmas formed rise buoyantly, and are emplaced in the upper crust. In many cases a part of the granitic magma is extruded as rhyolitic volcanic rocks. *See* DOLERITE; GABBRO; IGNEOUS ROCKS; MAGMA; METAMORPHIC ROCKS; METAMORPHISM; NEPHELINE; PETROLOGY; PORPHYRY; QUARTZ.

Fred Barker

Bibliography. M. G. Best, *Igneous and Metamorphic Petrology*, 1982; A. Miyashiro, *Metamorphic Petrology*, 1994; L. A. Raymond, *Petrology: The Study of Igneous, Sedimentary, and Metamorphic Rocks*, 1994; H. J. Stein and J. L. Hannah (eds.), *Ore-Bearing Granite Systems: Petrogenesis and Mineralizing Processes*, 1990.

Granodiorite

A phaneritic (visibly crystalline) plutonic rock composed chiefly of sodic plagioclase (oligoclase or andesine), alkali feldspar (microcline or orthoclase, usually perthitic), quartz, and subordinate dark-colored (mafic) minerals (biotite, amphibole, or pyroxene). Granodiorite is intermediate between granite and quartz diorite (tonalite). Alkali feldspar is dominant over plagioclase in granite but is subordinate to plagioclase in granodiorite. Quartz diorite carries little or no alkali feldspar. For convenience granite and granodiorite are commonly grouped and referred to as granite. *See* GRANITE; IGNEOUS ROCKS.

Carleton A. Chapman

Granulite

An important class of metamorphic rocks exposed at the surface of the Earth's crust, and inferred to make up a large portion of the deeper crust. Granulites are known from experimental petrologic studies to have formed at higher temperatures, and in many cases, higher pressures, than most other crustal rock assemblages. Thus, they are believed to have formed at considerable depths in the crust. The transport of whole granulite terrains, sometimes hundreds of kilometers in lateral dimension, from depths of 12–18 mi (20–30 km), poses a major problem.

Granulites may be of many different bulk compositions, inherited from precursor sedimentary, igneous, or lower-grade metamorphic rocks. The high temperatures of crystallization have resulted in very low water content, reflected in nearly anhydrous mineralogy. Characteristic minerals of granulite metabasalts are plagioclase, orthopyroxene, clinopyroxene, hornblende, and garnet. These minerals are also characteristic of granulites of intermediate to granitic composition, together with progressively greater amounts of quartz and potassium feldspar. The association of potassium feldspar with orthopyroxene is definitive for charnockite, a granulite of approximately granitic composition characteristic of ancient high-grade terrains.

Almost all granulites are Precambrian. The most extensive granulite terrains are late Archean, 2.5–2.9 $\times$ 10^9 years old, as in Fennoscandia, southwestern Greenland, southern India, western Australia, and Antarctica, although some extensive granulite areas are younger Precambrian, as in the Grenville Province of southeastern Canada and northern New York. In several well-studied terrains, such as southwestern Greenland, southern Norway, and southern India, there are gradual regional transitions from rocks of lower metamorphic grade, rich in micas and other hydrous minerals, to granulites. The **illustration** shows one such terrain, in southern India and Sri Lanka, in which gradations from lower-grade rocks to granulites take place over distances of tens to a few hundreds of kilometers. These gradations presumably represent crustal gradients of increasing temperature or pressure, related to increasing crustal depth during the Archean metamorphic episode. How the high-temperature and high-pressure mineralogy was "frozen in" without substantial back reaction to lower-grade conditions during slow uplift to the surface after the metamorphic period is poorly understood; granulites share this problem with all other metamorphic rocks. Granulites characteristically contain CO_2-rich fluid inclusions in the mineral grains, in contrast to the more aqueous fluid inclusions of other kinds of rock. This suggests that action of volatiles low in H_2O and rich in CO_2, probably of subcrustal origin, were important in crustal metamorphism early in the Earth's history. *See* BASALT; GRANITE; METAMORPHIC ROCKS; METAMORPHISM. R. C. Newton

Bibliography. R. A. Howie, Charnockites, *Sci. Progr.*, 52:628–644, 1964; P. R. A. Wells, Chemical and thermal evolution of Archaean sialic crust, southern West Greenland, *J. Petrol.*, 20:187–226, 1979.

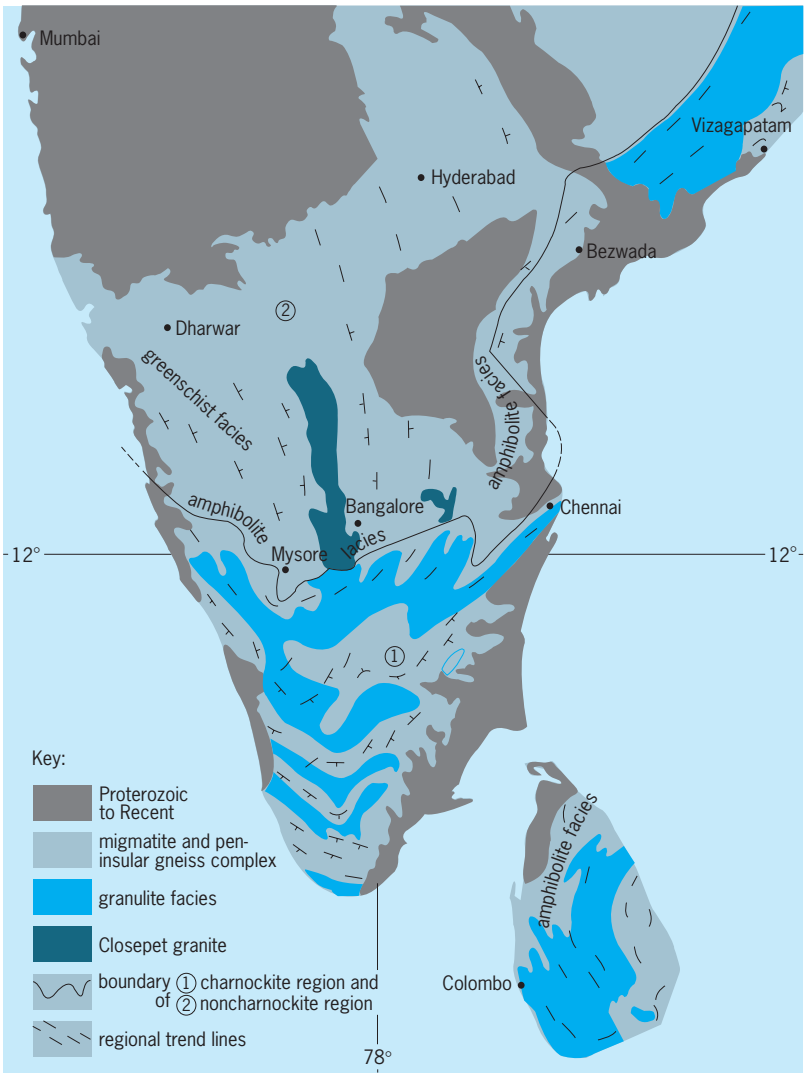

Southern part of peninsular India and Sri Lanka, showing broad tectonic and metamorphic features. (*After A. P. Subramaniam, Charnockites and granulites of southern India: A review, Dan. Geol. Foren., 17:473–493, 1967*)

Granuloma inguinale

A mildly infectious, chronic, granulomatous disease principally affecting skin and subcutaneous tissues of the genital and rectal areas. Although rare in the United States, the disease is very common in New Guinea, the Caribbean, and other tropical and subtropical areas.

The causative organisms, *Calymmatobacterium granulomatis* (*Donovania granulomatis*), are gram-negative encapsulated bacteria isolated only with difficulty from lesion biopsy material by using chick chorioallantoic membrane or coagulated egg yolk slants. An organism similar to *C. granulomatis* is isolatable from feces, and such organisms are not

infrequently detected in vaginal specimens.

The usual diagnostic procedure is the microscopic examination of smears of lesion scrapings that have been stained with Wright-Giemsa stain. There is no uniform serological procedure for either detecting antibodies against *C. granulomatis* or identifying the organism.

Although the method of transmission of granuloma inguinale is controversial, there is a definite correlation with sexual activity and a frequent association with homosexual behavior. Improved personal hygiene and individual awareness could correspondingly reduce morbidity and tissue destruction. *See* SEXUALLY TRANSMITTED DISEASES.

Tetracyclines are the drugs of choice, with streptomycin an effective alternative. Resistance to these drugs is countered by changing to either chloramphenicol or gentamicin. Resolution of the lesions begins in 1 week and is complete within 3–5 weeks, unless treatment is discontinued. *See* ANTIBIOTIC; MEDICAL BACTERIOLOGY. Douglas S. Kellogg

Bibliography. R. F. Dodson et al., Donovanosis: A morphologic study, *J. Invest. Derm.*, 62:611, 1974; G. Hart, Psychological and social aspects of venereal disease in Papua New Guinea, *Brit. J. Vener. Dis.*, 50:453, 1974; S. Lal and C. Nicholas, Epidemiological and clinical features in 165 cases of granuloma inguinale, *Brit. J. Vener. Dis.*, 46:461, 1970; E. H. Lennette et al. (eds.), *Manual of Clinical Microbiology*, 1974.

Grape

The two genera of grapes are *Vitis* and *Muscadinia*. *Vitis vinifera* has intermittent forked tendrils, bark that sheds, a diaphragm at the node, and elongated clusters with berries that adhere to the pedicels at maturity (**Fig. 1**). This species also has thin, smooth, shiny leaves with three, five, or seven lobes. Berries may be round or oval and have edible skins that adhere to the flesh. In the American species, skins slip from the pulp. Many American species have a characteristic musky or foxy odor and taste. *Muscadinia* can be easily distinguished from *Vitis* by bark that does not shed and simple tendrils that do not fork. *See* FRUIT.

Viticulture is the science of grape production. In a broad sense, viticulture includes studies of grape varieties; methods of culture such as trellising, pruning, and training; insect and disease control; propagation; and raisin production.

Cultivation. Countries with the largest acreages are Spain, Italy, France, Russia, and Turkey. California ranks ninth in acreage, and it produces 3% of the world's grapes. Worldwide, about 8000 varieties of grapes have been named and described. Of this number, perhaps about 100 are of major importance.

In the United States, *V. vinifera* is grown on the west coast, and most of the grapes cultivated east of the Rocky Mountains have been derived from American native species such as *V. labrusca* and *V. aestivalis*, or from crosses between them and *V. vinifera*. There is also a native Caribbean species and several Asiatic species. There are three main species of *Muscadinia* that are found mostly in the southeast portion of the United States.

Fig. 1. Seyval grapes. (*Bountiful Ridge Nurseries*)

Grape stems are rather flexible, and so a trellis is usually needed to support the vine. In California the three main systems are the head-, cordon-, and cane-pruned vines (**Fig. 2**). For mechanical harvesting purposes, the vines are usually trained to a vertical, two- or three-wire trellis. In New York and the eastern United States the umbrella Kniffin, the four-cane Kniffin, and the Geneva double curtain are frequently used. The different kinds of training found around the world are virtually innumerable.

Uses. Table grapes are utilized for food and decorative purposes. They must have an attractive appearance, be palatable, have good shipping and storage qualities, and be resistant to injury in handling. Large berries of uniform size that have a firm pulp, a tough skin, a sturdy rachis, and a strong adherence of berries to cap stems are desirable. In the United States there is a marked preference for seedless grapes. Some of the leading table grapes in California are Emperor, Tokay, Thompson Seedless, Cardinal, and Perlette. Some of the principal commercial American varieties are Concord, Catawba, Delaware, and Niagara. Some of the important varieties of *M. rotundifolia*, the Muscadine grape, are Scuppernong, Thomas, and Hunt. The Republic of South Africa and Australia are also important table grape–growing countries.

Important wine grapes in California include Cabernet Sauvignon, Carignane, Chardonnay, Grenache,

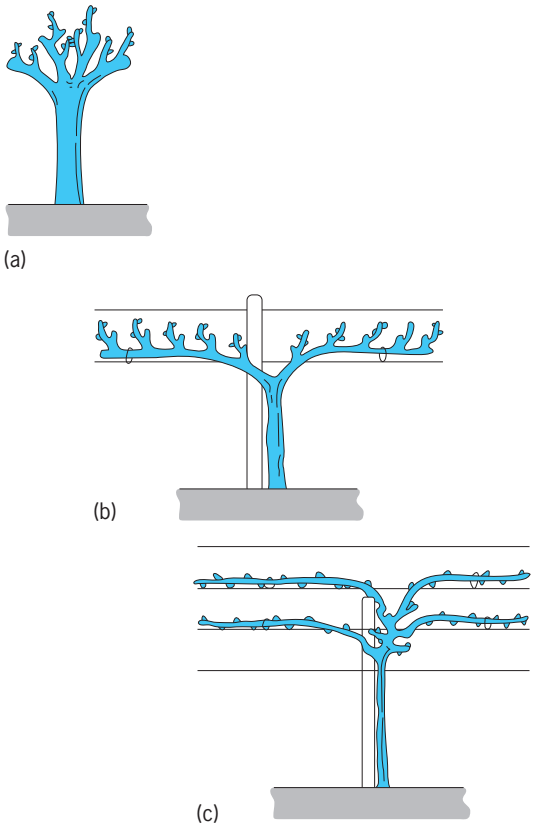

Fig. 2. Three main training-pruning systems in California. (*a*) Head-trained–spur-pruned. (*b*) Cordon-trained–spur-pruned (*c*) Head-trained–cane-pruned. (*After R. J. Weaver, Grape Growing, John Wiley and Sons, 1976*)

French Colombard, and Zinfandel. Many North American and *rotundifolia* species used for eating purposes are also used for wine.

A good dried raisin must be soft, textured, not sticky, early to ripen, and preferably seedless. The leading raisin grapes are Black Corinth, Thompson Seedless, and Muscat of Alexandria. The last variety has seeds that can be removed by machine. California produces 40% of the world's raisins, and there most are dried in the sun between the rows on paper trays. Other important raisin producers are Turkey, Greece, Spain, Iran, and Australia. In Middle Eastern countries the raisins are often dried on flat areas of soil or cement, and in Australia the raisins are dried in the shade on racks. It is important to dry raisin grapes before the rains begin. In rainy areas, facilities for rapidly covering the drying grapes must be available. Raisin grapes can also be dried in dehydrators.

Harvesting, processing, and marketing. Many wine grapes are now harvested mechanically. The slapper-type harvester (**Fig. 3**) has double banks of flexible horizontal rods that shake the vines and remove the clusters. The pulsator type of harvester strikes the posts rapidly so that the berries fall off. Fruit is immediately taken to the winery, where it is crushed and made into wine. Table grapes and most raisin grapes are still picked by hand. Soon after picking, table grapes start to deteriorate. If they are not to be eaten right away, they should be cooled as soon as possible. Most table grapes grown on the west coast of the United States are shipped east by train for marketing. Robert J. Weaver

Diseases. Grapes are susceptible to many diseases, which under certain climatic conditions may limit the kinds of grapes that can be grown. Disease control is particularly difficult in areas with high humidity, frequent summer rains, and high temperatures. Some disease control measures are required in all grape-growing areas.

Fungi. Powdery mildew (caused by *Uncinula necator*), anthracnose (caused by *Elsinoë ampelina*), black rot (caused by *Guignardia bidwellii*), and downy mildew (caused by *Plasmopara viticola*), are important grape diseases worldwide. Unlike the other three, powdery mildew occurs in dry climates. Symptoms include grayish-white powdery growth on the surface of green parts of the vine and dropping, discoloration, or splitting of the fruit (**Fig. 4**).

Anthracnose symptoms include lesions or cankers on all green parts of the vine; fruit lesions are circular, sunken, and ashy gray, and are surrounded by a dark margin (**Fig. 5**). Black rot affects young tissue of all parts of the vine, but damage to the immature fruit is most serious. Symptoms are reddish-brown, circular spots on leaves and, on bunch grapes, a rot which within 7–10 days transforms berries into black, hard, shriveled mummies that remain attached to the bunch. Downy mildew symptoms include light-yellow translucent spots on the upper leaf surfaces, patches of white mildew on the lower surfaces of leaves, and malformation of shoots, tendrils, or berries early in the season. Other fungal diseases include *Eutypa* dieback (caused by *E. armeniacae*), *Botrytis* rot (caused by *B. cinerea*), and bitter rot (caused by *Melanconium fuligineum*).

Vineyard sanitation, proper pruning, and planting on sites with good air drainage help reduce the severity of fungal diseases. Chemical spraying is usually necessary. Sulfur dusts or sprays control powdery

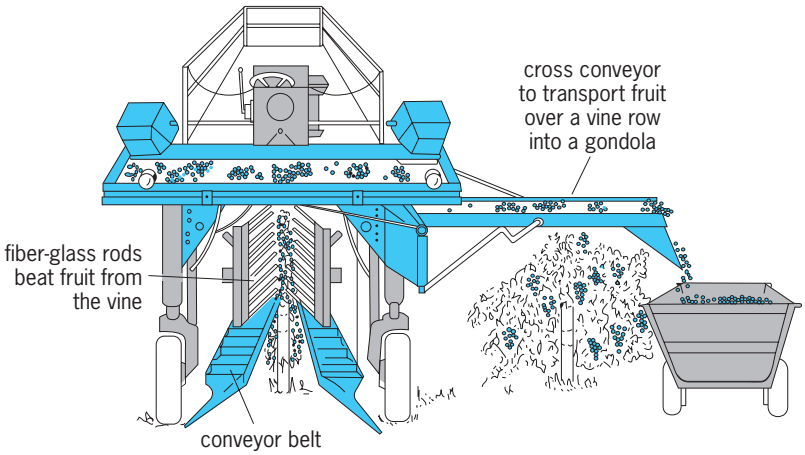

Fig. 3. Slapper-type mechanical harvester. (*After R. J. Weaver, Grape Growing, John Wiley and Sons, 1976*)

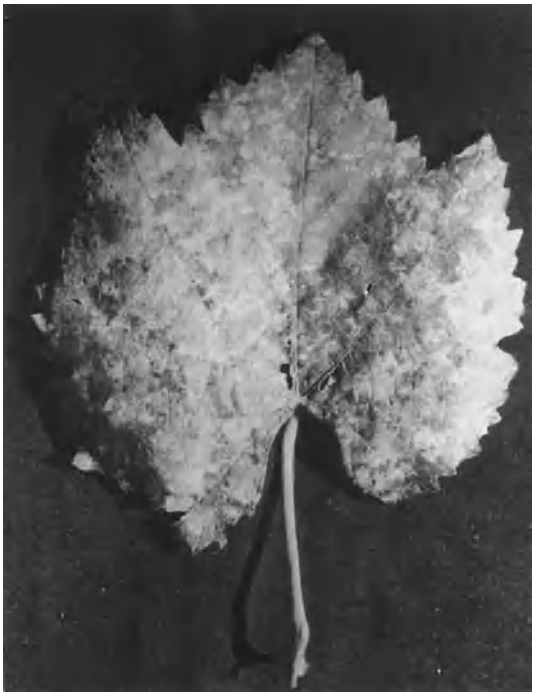

Fig. 4. Powdery mildew on a grape leaf.

mildew but injure the vines at high temperatures. For anthracnose control, a dormant spray of liquid lime-sulfur is beneficial. Bordeaux mixture is effective against most diseases, but injures young grape tissue. Several organic fungicides are effective. These include ferbam, captan, folpet, maneb plus zinc sulfate, zineb, and benomyl. Benomyl and ferbam are not effective against downy mildew.

Fig. 5. Anthracnose on the grape fruit, commonly called bird's-eye rot.

Viruses. Grape leafroll, fanleaf, yellow mosaic, veinbanding, and corky bark are important viral diseases worldwide. Many other grape viruses are serious in certain locations. Since the grape viruses either have no vector or are nematode-transmitted, the best control is the use of virus-free planting stocks. Fumigation may also be used to reduce reinfections by the nematode-transmitted viruses such as fanleaf, yellow mosaic, and veinbanding.

Bacteria. Bacterial blight (caused by *Erwinia vitivora*) and crown gall (caused by *Agrobacterium tumefaciens*) are minor problems, but Pierce's disease is devastating to grapes and limits the production of bunch grapes in the southeastern United States. Pierce's disease symptoms include marginal necrosis of leaves, wilting and drying of fruit, decline of vigor, and usually death of the grapevine. The only control for Pierce's disease is the use of resistant or tolerant cultivars. Plants may be freed of the causal agent by hot water treatment (25 min at 50°C), and this is valuable in preventing the movement of the pathogen in propagating wood. *See* PLANT PATHOLOGY.

D. L. Hopkins

Bibliography. D. Flaherty (ed.), *Grape Pest Management*, rev. ed., 1992; R. C. Pearson and A. C. Goheen (eds.), *Compendium of Grape Diseases*, 1988.

Grapefruit

A citrus fruit, *Citrus paradisi*. It apparently arose as a hybrid of shaddock or pummelo and sweet orange in the West Indies. Its first recorded mention was in Barbados in 1750, and the first use of the term grapefruit occurred in Jamaica in 1814. It was thought to have been introduced into Florida by Count Odelle Phillipe near Safety Harbor on Tampa Bay around 1823. The term grapefruit was derived from the tree's tendency to produce large clusters of fruit, as grape vines do.

Plant structure. The tree is a large evergreen, spreading in habit and becoming larger than most other edible citrus species. Grapefruit wood is slightly less cold-tolerant than sweet orange but more resistant to heat. The individual leaves are relatively large compared to those of sweet oranges, and their broadly winged petioles slightly overlap the leaf bases. Flowers are white-petaled and perfect with pollination occurring naturally without the aid of bees. Fruit is relatively large and the peel thick compared to sweet oranges. Fruit shape is oblate or flattened at each end unless grown from off-bloom (not regular bloom) or under growing conditions promoting excessive vigor, in which case the fruit is often pear-shaped or sheep-nosed. The yellow peel color is not related to cool temperature as in the case of sweet oranges, but fruit picked early in the season must be degreened with ethylene to develop a satisfactory peel color. *See* ETHYLENE; FRUIT.

Hedging (shearing back the sides) and topping (shearing off the tops) of trees with large mechanical equipment are used to maintain trees within

allotted spaces, providing for better light penetration and more efficient production and harvesting operations. Topping of larger trees reduces the top–to–root system ratio and results in larger fruit of better quality.

Cultivars. The original grapefruit were white-fleshed and extremely seedy; however, current important commercial cultivars are seedless or contain few seeds. Cultivars include the commercially seedless white Marsh, pink Marsh or Thompson, Redblush or Ruby Red, and Star Ruby and the seedy Duncan. Cultivars resembling grapefruit, but of no current commercial interest, include the extremely seedy but tasty Triumph and Royal. Grapefruit cultivars are not divided into early, midseason, and late as are sweet oranges, as there are only slight differences in maturity dates of these cultivars grown commercially.

All important grapefruit cultivars have sterile pollen, and the seedless ones are highly parthenocarpic and largely ovule-sterile. Most grapefruit cultivars have arisen from bud mutations in commercial groves, with their subsequent recognition and propagation by growers rather than as a result of traditional breeding programs. The Star Ruby was developed from seed of the seedy, pink-fleshed Hudson cultivar that was irradiated. A recent product of a breeding program at the University of California is the Oroblanca, a triploid developed from the hybridization of a grapefruit and pummelo.

Composition. Grapefruit ripens slowly over an extended period, storing well on the tree after reaching edible quality, with fruit of a given cultivar harvested from early fall to midsummer. Composition is, therefore, important not only for indicating nutritive values but also for determining proper time of harvest. The composition of fruit varies not only with the degree of maturity, but also with climate, cultivar, cultural practices, and other factors. Grapefruit may frequently be spot-picked for large size with the remainder left on the tree for later harvest.

The fresh weight of grapefruit consists of 35–50% juice, with the remainder made up of peel, pulp, and seeds. Fresh juice consists of 88–93% water, 7–12% soluble solids with sugars (sucrose and reducing) and acids (chiefly citric) constituting 85–90% of these solids. The edible quality of grapefruit depends in large measure upon the ratio of sugars to acids in juice. The nutritive value of juice is in part related to its vitamin C content. The juice also contains a number of other vitamins and mineral elements required in a well-balanced human diet. The principle giving grapefruit its distinctive bitter flavor is naringin, a glucoside not found in its progenitor the pummelo or in other commercial citrus. *See* ASCORBIC ACID; CITRIC ACID.

Uses. Profits from grapefruit production traditionally come from that portion of the crop marketed for fresh fruit. Thus, those factors influencing the percentage of the crop meeting fresh grade (pack-out) are very important. Excessive fertilization which results in poorly shaped fruit and blemishes from wind scarring and pests reduce market grade.

Fig. 1. Greasy spot lesions on lower side of grapefruit leaf.

The United States produces about 75% of the world's grapefruit, and Florida accounts for 75% of this total. The remainder is produced in Texas, California, and Arizona. Florida's grapefruit production has averaged just over 50 million 85-lb (38.5-kg) boxes annually, while combined production of the other United States citrus-producing areas has averaged just over 17 million boxes. Typically, 36% of Florida's crop is marketed as fresh fruit with about 50% of the crop in California-Arizona and Texas utilized in that form. The remainder of the crop in Florida and the other production areas is utilized for

Fig. 2. Rootrot lesion on susceptible grapefruit scion.

processed products, primarily juices. Fresh grapefruit exports from Florida account for a large part of the fresh market shipments. In 1981–1982, for example, one-third of the total shipments were exported. Japan is the major export market for Florida grapefruit, accounting for about 70% of the export shipments. Most of the grapefruit juice production is sold in the United States market. Grapefruit juice is sold as either canned juice, ready-to-serve product, or frozen concentrate. Canned juice sales account for more than 50% of the volume sold through the grocery stores.

Diseases. Fungus diseases of foliage and fruit which must be controlled by spraying include greasy spot, caused by *Mycospharella citri* (**Fig. 1**); melanose, caused by *Diaporthe citri*; and scab caused by *Elsinoë fawcetti*. Rootrot, caused by the fungus *Phytophthora parasitica*, can result in a high percentage of young grapefruit tree loss. However, its incidence may be reduced by the use of tolerant rootstocks, high budding, and fungicides. High budding reduces the proximity of the susceptible grapefruit scion to the soil surface, the source of the fungal inoculum (**Fig. 2**). Grapefruit trees are susceptible to the important citrus viruses, including exocortis, psorosis, xyloporosis, and tristeza. However, the economic impact of these viruses is now minimized with the use of virus-free propagation material and tolerant rootstocks. Rust mites, the cause of rind russetting, must also be controlled if fruit is to be marketed fresh. *See* PLANT PATHOLOGY.

David P. H. Tucker

Bibliography. W. Reuther et al., *The Citrus Industry*, 5 vols., 1967–1989; R. W. Ward and R. L. Kilmer, *The Economics of the United States Citrus Industry: A Domestic and International Perspective*, 1988.

Graph theory

A branch of mathematics that belongs partly to combinatorial analysis and partly to topology. Its applications occur (sometimes under other names) in electrical network theory, operations research, organic chemistry, theoretical physics, and statistical mechanics, and in sociological and behavioral research. Both in pure mathematical inquiry and in applications, a graph is customarily depicted as a topological configuration of points and lines, but usually is studied with combinatorial methods. *See* COMBINATORIAL THEORY; TOPOLOGY.

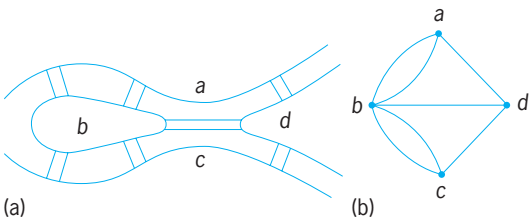

Fig. 1. Königsberg bridge problem. (*a*) The seven bridges of Königsberg. (*b*) Corresponding graph.

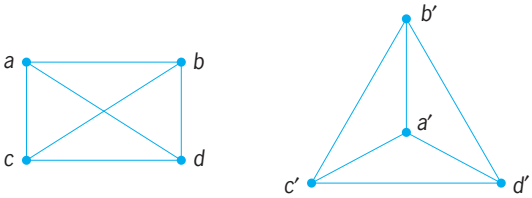

Fig. 2. Two isomorphic graphs.

Origin of graph theory. Graph theory and topology are said to have started simultaneously in 1736 when L. Euler settled the celebrated Königsberg bridge problem. In Königsberg, there were two islands linked to each other and to the banks of the Pregel River by seven bridges. **Figure 1** illustrates both this setting and its topological abstraction as a graph. The points a, b, c, d correspond to land areas, and the connecting lines to bridges. The problem is to start at one of the land areas and to cross all seven bridges without ever recrossing a bridge. Euler proved that there is no solution, and he established a rule that applies to any connected graph: such a traversal is possible if and only if at most two points are odd, that is, each is the terminus for an odd number of lines. Euler also proved that the number of odd points in a graph is always an even number. Thus, a complete traversal without recrossing any lines is possible if the number of odd points is zero or two. If zero, the complete traversal ends at the starting point.

In geometry a graph might arise as the set of vertices and edges of a convex, three-dimensional polyhedron, such as a pyramid or a prism. Euler derived an important property of all such polyhedra. Let V, E, and F be the numbers of vertices, edges, and faces of such a polyhedron. Euler proved that $V - E + F = 2$, which is now called the Euler equation. For instance, a cube has $V = 8$, $E = 12$, and $F = 6$, so that $8 - 12 + 6 = 2$. Euler's observations have been extended to a theorem about embeddings of graphs in surfaces and to the Euler-Poincaré characteristic for cell complexes in combinatorial topology.

Definitions. A graph consists of a set of points, a set of lines, and an incidence relation that designates the end points of each line. In many applications no line starts and ends at the same point. (Such a line would be called a loop.) Also, no two lines have the same pair of end points. A graph whose lines satisfy these conditions is called simplicial. The valence of a point is the number of lines incident on it, calculated so that a loop is twice incident on its only end point. Two graphs are isomorphic if there is one-to-one correspondence from the point set and line set of one onto the point set and line set, respectively, of the other that preserves the incidence relation. The point correspondence $a \to a'$, $b \to b'$, $c \to c'$, $d \to d'$ indicates an ismorphism between the two graphs of **Fig. 2**.

An automorphism of a graph is an isomorphism of a graph with itself. For instance, a plane rotation of 120° would yield an automorphism of either of the two graphs in **Fig. 3**. A plane reflection through

Fig. 3. Two graphs for which either a 120° rotation or a vertical reflection is an automorphism.

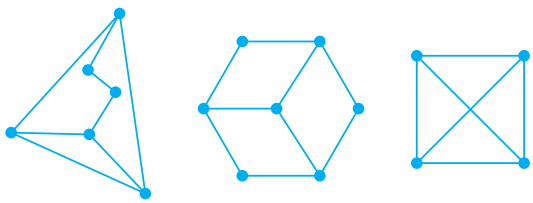

Fig. 4. Homeomorphic but nonisomorphic graphs.

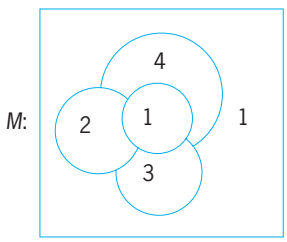

Fig. 5. Plane map requiring four colors.

a vertical axis would also yield an automorphism of either of them. The set of all automorphisms of a graph G forms the automorphism group of G. R. Frucht proved in 1938 that every finite group is the automorphism group of some graph. Two graphs are homeomorphic (**Fig. 4**) if, after smoothing over all points of valence 2, the resulting graphs are isomorphic. *See* GROUP THEORY.

Map coloring problems. Drawing a graph on a surface decomposes the surface into regions. One colors the regions so that no two adjacent regions have the same color, rather like a political map of the world. It is a remarkable fact that for a given surface, there is a single number of colors that will always be enough no matter how many regions occur in a decomposition of the surface. The smallest such number is called the chromatic number of that surface. It is easy to draw a plane map, as in **Fig. 5**, that requires four colors. In 1976 K. Appel and W. Haken settled a problem dating back to about 1850, by showing that four colors are always enough for plane maps.

Some maps on more complicated surfaces require more than four colors. For instance, **Fig. 6** illustrates a map on a torus (the surface of a doughnut) that needs seven. To obtain the toroidal map from the rectangular drawing, first match the top to the bottom to get a cylindrical tube. Then match the left end of the cylinder to the right end to complete the torus. Whereas, before this matching, region 7 meets only regions 1 and 6, after the matching it also meets region 2 along FG, region 3 along GH, region 4 along AB, and region 5 along BC. In fact, after the matching, each of the seven regions borders every other region. It follows that seven colors are necessary. No map on the torus needs more than seven colors, as P. J. Heawood proved in 1890. G. Ringel and J. W. T. Youngs completed a calculation in 1968 of the chromatic numbers of all the surfaces except the plane or sphere.

Planarity. A graph is planar if it can be drawn in the plane so that none of its lines cross each other. Neither of the two graphs in **Fig. 7** can be drawn in the plane. K. Kuratowski proved in 1930 that a graph is planar if and only if it contains no subgraph homeomorphic to either of those two graphs. Testing all the subgraphs might be a very tedious process, even on a fast computer. In 1974 J. Hopcroft and R. Tarjan obtained an extremely fast alternative planarity test. The time it takes a computer to perform the Hopcroft-Tarjan test is linearly proportional to the time it takes to read its point set into the computer.

There are methods to decide for any graph and any surface whether the graph can be drawn on the surface without edge crossings. The time to execute such methods is unfeasibly large for most graphs and most surfaces except the plane or the sphere. Ringel has constructed many important special drawings on higher-genus surfaces.

Variations. In a directed graph, or digraph, each line ab is directed from one end point a to the other end point b. There is at most one line from a to b. The adjacency matrix $M = (m_{ij})$ of a digraph D with points $b_1, b_2, \ldots, b_n$ has the entry $m_{ij} = 1$ if the line b_ib_j occurs in D; otherwise $m_{ij} = 0$ (**Fig. 8**).

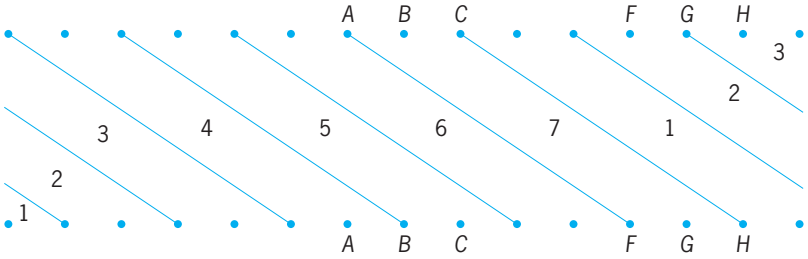

Fig. 6. Map on a torus (doughnut) that requires seven colors. To form the torus, paste opposite sides of the rectangle together.

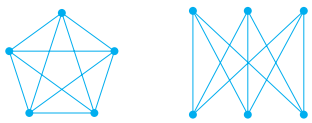

Fig. 7. Prototypes of all nonplanar graphs.

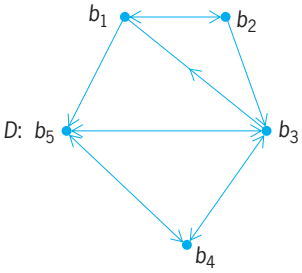

$$M = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Fig. 8. A digraph and its adjacency matrix.

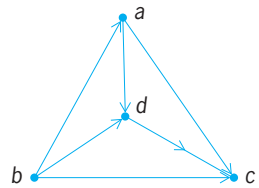

Fig. 9. A tournament.

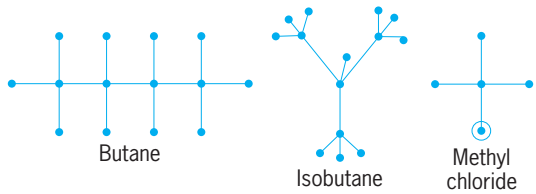

Fig. 10. The two isomers of a saturated hydrocarbon.

An oriented graph is obtained from an ordinary graph by assigning a unique direction to every line. If there is one line between ever pair of points and no loops, an ordinary graph is called complete. An oriented complete graph is called a tournament (**Fig. 9**).

Applications. A. Cayley reformulated the problem of counting the number of isomers of saturated hydrocarbons (C_nH_{2n+2}) in graphical language (**Fig. 10**). Each isomer is a tree all of whose vertices have valence 1 for hydrogen, or 4 for carbon. G. Polya devised a general theorem in 1937 for enumeration to provide a solution to such problems. F. Harary and others have solved many related problems by applying and extending Polya's theorem. Extremely effective use of Polya's theorem occurs in theoretical physics, where G. Ford and G. Uhlenbeck solved several graphical enumeration problems arising in statistical mechanics. *See* MOLECULAR ISOMERISM.

Suppose that some of the points of a graph correspond to workers $x_1, \ldots, x_m$, that the rest of the points correspond to jobs $y_1, \ldots, y_m$, and the presence of a line between x_i and y_j means that worker x_i is capable of performing job y_j. The personnel assignment problem is to find m lines so that each worker x_i is matched to exactly one possible job. In the optimal assignment problem, labels on the lines tell how well a worker can do a particular job. An algorithm due to H. Kuhn and J. Munkres solves the optimal assignment problem.

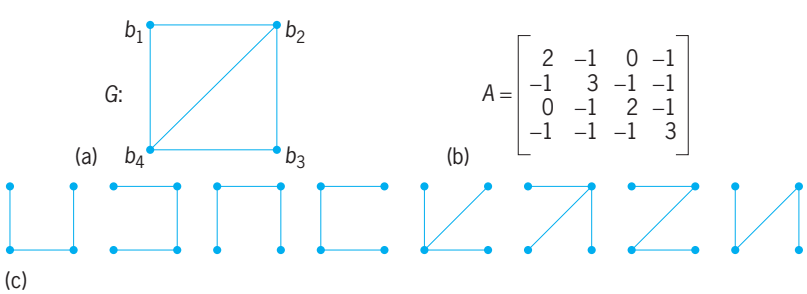

Fig. 11. Matrix-tree theorem. (*a*) Graph *G*. (*b*) Corresponding matrix *A*. (*c*) Spanning trees of *G*.

If the points of a graph represent cities and the lines between them are labeled with distances, one might want to find the shortest path from one point to another. An efficient method to determine a shortest path was developed by E. Dijkstra in 1959. K. Menger proved in 1927 that if A and B are disjoint sets of a connected graph G, then the minimum number of points whose deletion separates A from B equals the maximum number of disjoint paths between A and B. L. Ford and D. Fulkerson generalized Menger's theorem into a method for solving network flow problems. *See* LINEAR PROGRAMMING.

According to the physical laws of G. Kirchhoff and G. Ohm, any set of voltages applied to the input nodes of an electrical network determines the voltages at all other nodes and the currents on every branch. Kirchhoff also proved the result known as the matrix-tree theorem: Let G be a connected graph with points $b_1, \ldots, b_n$ and let $A = (a_{ij})$ be the matrix such that a_{ij} is the valence of b_i and, for $i \neq j$, $a_{ij} = -1$ if b_i is adjacent to b_j or 0 otherwise. Then the cofactor of each entry a_{ij} equals the number of spanning trees of G, that is, the number of trees in G that includes every point of G (**Fig. 11**). *See* NETWORK THEORY.

Numerous applications of graph theory to social and behavioral science have been developed, many by Harary and his coauthors. If points represent persons and lines represent such interrelationships as communication, linking, or power, then a graph may depict various aspects of social organization. Anthropologists use graphs to describe kinship, and management scientists use them to display corporate hierarchy.

Graph theory is presently in a phase of rapid growth. Two of the major branches not described here are extremal graph theory, founded by P. Turan and developed by P. Erdös, and hypergraph theory, developed by C. Berge. Matroid theory, originated by H. Whitney and expanded by W. Tutte, is closely related to graph theory. Tutte is also responsible for important results in many other areas of graph theory and combinatorial research, including connectivity, decomposition, and chromatic numbers.

Jonathan L. Gross

Bibliography. N. L. Biggs, E. K. Lloyd, and R. J. Wilson, *Graph Theory 1736-1936*, 1986, reprint 1999; B. Bollobas, *Graph Theory: An Introductory Course*, 1979; J. L. Gross and J. Yellen, *Graph Theory and its Applications*, 1999; J. G. Michaels and K. H. Rosen (eds.), *Applications of Discrete Mathematics*, 1991; D. B. West, *Introduction to Graph Theory*, 2nd ed., 2000; R. J. Wilson, *Introduction to Graph Theory*, 1996.

Graphic methods

Procedures and techniques for visually representing on paper or a screen information pertaining to data analysis and decision making or relationships between variables by means of diagrams or charts.

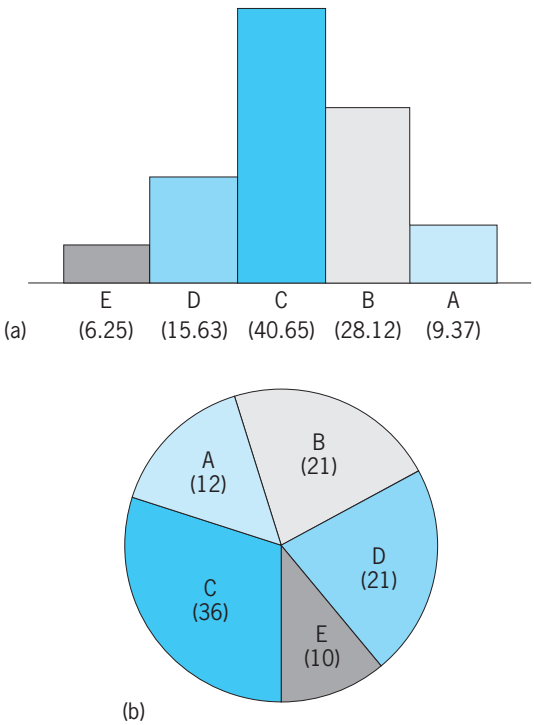

Fig. 1. Percentages of distribution of grades in a class, depicted by (*a*) bar graph and (*b*) circle graph.

Structural information is usually depicted by a bar graph, also known as a histogram (**Fig. 1***a*), or a circle graph, also known as a sectorgram or a pie chart (Fig. 1*b*). Such charts are commonly used to display data from industry, business, international trade, and government. The field of descriptive statistics relies heavily on graphs to simplify the presentation of data.

A graph representing the relation between an independent and a dependent variable is a two-dimensional line, whereas a graph representing the relation between two independent variables and a dependent variable is a three-dimensional surface. **Figure 2**, for example, is a line graph representing the functional relation $y = 5 - x$ between the nonnegative independent variable x and the dependent variable y.

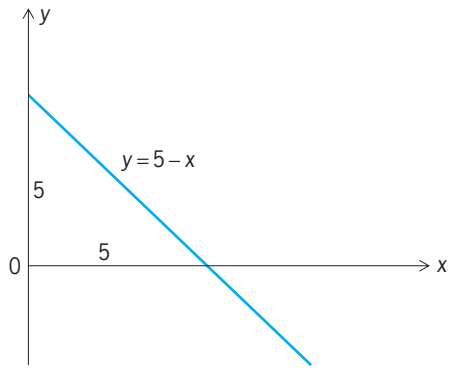

Fig. 2. Line graph of the functional relation $y = 5 - x$.

Graphical methods are frequently used to solve problems of curve fitting, correlation and regression analysis, nomographs and alignment charts, numerical integration, areas under curves, and interpolation and extrapolation. The bell-shaped curve representing the normal probability distribution is a line graph which is extensively used in problems involving sampling theory, estimation, and hypothesis testing in statistical analysis.

Line graphs are by far the most important type because they can be used not only to represent functional relationships but also to solve problems. For example, the roots of an equation in one variable can be found with the help of a graph. If $f(x)$ is a function which becomes 0 for a certain value of x, that value is a root of the equation $f(x) = 0$. The equation is solved by finding the x coordinates of points where the graph of $f(x)$ crosses the x axis. Two further examples of the solution of problems by line graphs will be briefly discussed: determination of equilibrium price and linear programming problems. *See* ROOT (MATHEMATICS).

Equilibrium price. If the price for a commodity is considered as the independent variable and the demand for the commodity as the dependent variable, the demand curve is a line graph illustrating the fact that as the price falls the demand increases. Likewise, the supply curve illustrates the fact that as price goes up the supply of the commodity is increased. If demand and supply curves are shown together, the point of intersection defines the equilibrium price under the law of supply and demand.

Linear programming. The line graphs of first-degree equations in two variables x and y of the type $ax + by = c$ (where a, b, and c are constants) are straight lines. A line divides the plane into two halves, one of which is represented by the inequality $ax + by \leq c$. The region enclosed by a finite number of half-planes and the two coordinate axes is a convex polygon called the feasible region, assumed to be nonempty. The problem of standard linear programming is that of finding a point in the feasible region at which the value of the objective function (a first-degree function of x and y) is a maximum. Even though there is an infinite number of points in the polygon, there is a vertex of the polygon at which the objective function attains its maximum value. Thus a graphical method can be used to solve this linear programming problem in two variables by drawing the polygon and computing the coordinates of all its vertices (a finite number) from which an optimal vertex can be easily chosen. *See* COORDINATE SYSTEMS; CURVE FITTING; EXTRAPOLATION; INTERPOLATION; LEAST-SQUARES METHOD; LINEAR PROGRAMMING; NOMOGRAPH; NUMERICAL ANALYSIS; STATISTICS.

V. K. Balakrishnan

Bibliography. R. W. Bowen, *Graph It!: How to Make, Read and Interpret Graphs*, 1991; R. D. Gustafson and P. D. Frisk, *Functions and Graphs*, 2d ed., 1991; S. M. Kosslyn, *Elements of Graph Design*, 1993; E. M. Mikhail, *Observations and Least Squares*, 1982; J. Molnar, *Nomographs: What They Are and How to Use Them*, 1981.

Graphic recording instruments

Instruments that make a graphic record of one or more quantities as a function of another variable, usually time. Signals representing information, such as the shape of time-varying electronic waveforms or the movements of a machine, are presented in graphical form by these devices. *See* TRANSDUCER.

The name of an instrument with a recording device incorporated often includes "-graph," such as a barograph for recording barometric pressure data.

Although there are still many graphic recording instruments in use, the trend now is to interface the sensor providing the signal to a computer. The signal can be processed immediately in any desired way and presented as a graph on the computer display or by a printer attached to the computer. This practice is much more versatile and often less expensive than the use of graphic recording instruments.

Classification. Recorders are of two forms: those that plot an input variable with respect to time (denoted x-t) and those that plot two different variables (denoted x-y). Graphic recorders can also be classified as being either directly or indirectly driven by the input signal. In addition, they can be classified by their exhibiting means, recording means, number of marking devices or channels, and marking means.

Direct- and indirect-acting recorders. Direct-acting units (**Fig. 1**) are suitable when the primary variable exerts sufficient and appropriate force to overcome the frictional loads of the bearings and marking means. Direct-drive recorders lack general-purpose usefulness and are, therefore, more costly to produce than indirect alternatives. In some circumstances, however, they are essential.

In the indirect form an intermediate stage is used to convert the input signal into the equivalent mechanical movements needed. Indirect-acting recorders generally accept voltage or current input signals, converting these into equivalent mechanical positions on a suitable exhibiting medium. Other input variables, such as pressure and temperature, are first transformed into electrical signals. Such signals may range from microvolts to megavolts or from picoamperes to kiloamperes. Recorders are often supplied with adjustable input stages so that the input signal can be matched to the scale size needed for the graph.

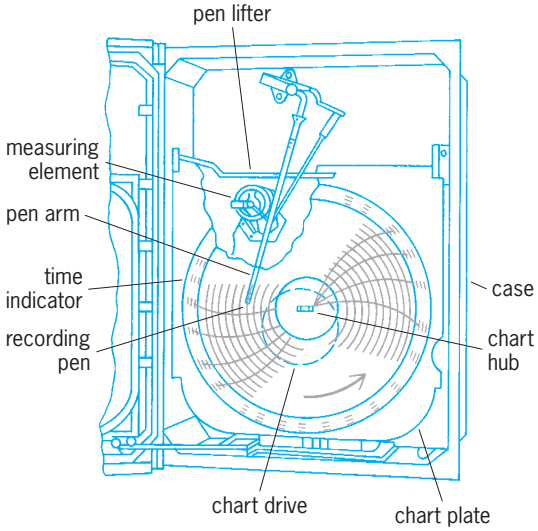

Fig. 1. Single-pen, direct-acting, circular-chart pressure recorder. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

Exhibiting means. Exhibiting means can be of either circular form or linear form, the latter being continuous strip or a single sheet. Circular charts vary from 4 to 12 in. (100 to 300 mm) in diameter. Strip-chart widths range from 2 to 12 in. (50 to 300 mm), with flat-bed sheets being the commonly used paper sheet sizes. Roll-chart lengths are available from 65 ft (20 m) to more than 330 ft (100 m). The paper is often fan-folded for convenience. The medium used can be either plain or specially treated paper or film. The recording means may be continuous, intermittent (when the marking device is retracted from the chart between measurements), or sequential (when more than one variable is intermittently recorded).

Multichannel recording. Multichannel recorders can record several signals simultaneously. This can be done by using several recording mechanisms combined in a single unit (**Fig. 2**). Alternatively a single marking mechanism can sequentially record values of each channel. One popular method, the dotting recorder, makes dots on the paper of a different color for each channel. Being closely placed, the dots produce a set of continuous, differently colored traces.

Marking methods. Many different methods are used to produce the permanent trace on the recording medium. They include the use of fluid inks and fiber pens that mark in direct contact with the paper; pressure-forced ink jetting; marking with a heated stylus on heat-sensitive paper; pressure- or voltage-sensitive paper on which the stylus moves in contact; pressure of an inked ribbon onto the paper after the marking head has been positioned; exposure of heat-sensitive or photographically sensitive paper or film; and electrostatically charged images in a laser printer.

Drive systems. Motive power is required to move the marking device relative to the exhibiting medium. In the x-t recorder the marker is driven across an axis mounted perpendicular to the medium that is moved under the medium at a steady rate with time (Fig. 2). In the x-y version the marker can be driven in two perpendicular directions, the medium being fixed (**Fig. 3**). Some systems move a single sheet backward and forward under the markers.

Two basic kinds of drive are used in recorders. The first is a speed-controlled drive that moves a strip of exhibiting medium under the marker; the other positions a marker along the axis at a distance that is proportional to the input magnitude.

Spring-motor chart drives are sometimes used, but the majority of x-t recorders are driven by electric motors of various kinds. Stepping motors have become prolific. Gearing is used to obtain the required chart speeds which can vary from several feet per second to a fraction of an inch per day. Speed is regulated by the use of servo velocity-controlled drives

or synchronous motors when suitable alternating-current (ac) supplies are available. *See* MOTOR; SERVOMECHANISM; STEPPING MOTOR; SYNCHRONOUS MOTOR.

Marker positioning systems make use of open- or closed-loop positioning systems. For these, rotary or linear drive motors are used. *See* CONTROL SYSTEMS.

A simple, extensively used way to move a marker to a position proportional to an input electrical signal is to place it on the end of a needle driven by an electromagnetic meter movement. The disadvantage of this method is that the friction of the contacting pen and the circular sweep of the pen can introduce error. Ink pen systems are also not suitable when very rapid following of the response is required. In the ultraviolet recorder and film oscillographs the meter movement (galvanometer) is used to rotate a small mirror that deflects a beam of radiation onto a sensitive exhibiting medium. These instruments can record kilohertz waveforms. *See* AMMETER; ELECTRICAL MEASUREMENTS; GALVANOMETER; VOLTMETERS.

For more robust operation, feedback control systems are used to control the speed of chart drives and the position of marker units. Each uses a similar principle of control.

As the marker unit of a typical position-control feedback system (**Fig. 4**) moves across the chart, it also moves the connection point made on a resistance potentiometer unit that is fed with electric current from a biasing dc voltage source. The input voltage (e) and the potentiometer output voltage (e_s) are fed to the two inputs of a differential electronic amplifier. If they are not equal, the amplifier produces a difference signal that is applied to the drive motor (M). Depending on the polarity of this difference, the motor moves the marker (as a smooth and continuous process) to a point where the difference is kept close to zero. At that point the marker is at a position on the chart that is proportional to the input signal magnitude. Similar means are used to control speed, but there a velocity sensor is used instead of the potentiometer. *See* DIFFERENTIAL AMPLIFIER.

Analog and digital forms. Graphic recorders (commonly referred to as plotters) are designed for use with either analog or digital electrical input signals. In the analog variation the input circuitry accepts signals that have the information to be displayed carried in terms of the variable amplitude of a voltage or current. A variable-gain input stage allows input signal magnitudes to be matched, by amplification or attenuation, to the set sensitivity of the positioning system.

In the digital plotter the digital signal (having only two states of existence for each of the bits forming the digital word equivalent to the analog alternative) can be of either serial format or parallel format. In the former a time sequence of serially occurring binary bits represents the signal amplitude. In the latter it is the state of binary signals simultaneously existing on several parallel lines that represents the amplitude. The signal lines carrying these digital signals, for purposes of interconnection, are called the communication interface.

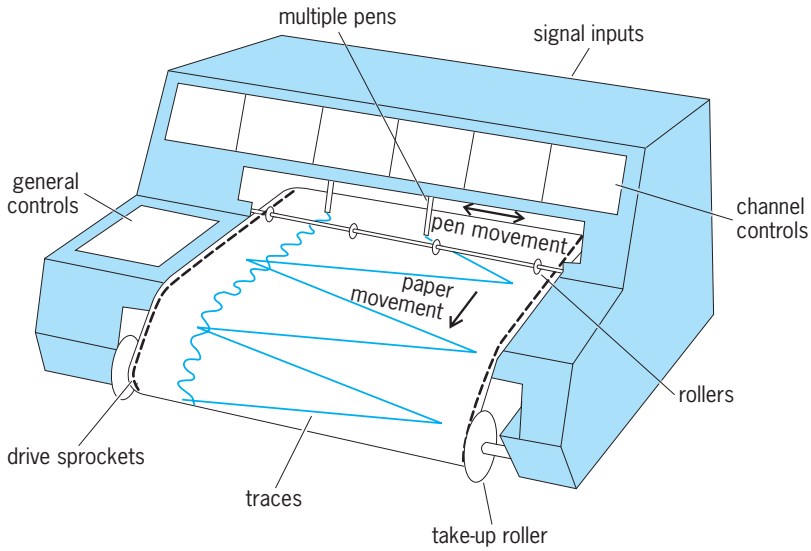

Fig. 2. Multichannel, x-t recorder with several recording mechanisms combined in one unit. Signal inputs feed channels at rear, and each channel has one set of controls. Multiple pens are staggered to miss each other as they traverse exhibiting medium, and produce traces which may be in different colors. Paper is held down by rollers and moved forward by sprocket-driven perforations.

As the digital interface is capable of transmitting multiple channels of information to the plotter, and because digital plotters contain a microprocessor computer that provides processing power and a measure of intelligence, such plotters offer more facilities than analog models. For example, in addition to plotting more than one input signal, they can be instructed (using a command language) to generate common geometric shapes, select axis sensitivities, change pens to provide different colors or line widths, draw text of required size, provide annotated axes to charts, add data point markers, lift the pen, set the drawing speed, and set the plot to a chosen paper or plotted size. Storage buffers are often provided that store data of the signal to be recorded.

Dynamic performance. Inadequate dynamic response from a recorder will give traces that have less

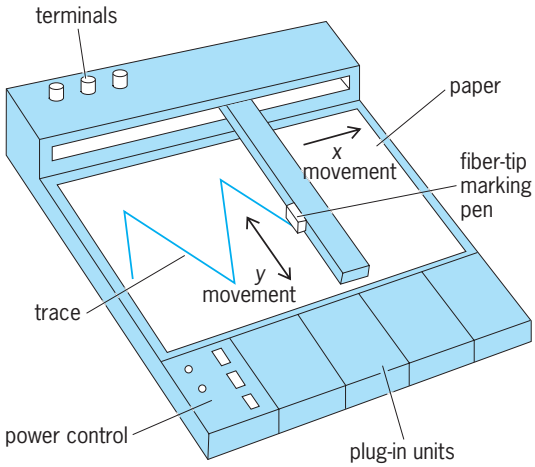

Fig. 3. One form of flat-bed, x-y chart recorder. Plug-in units can be used to configure recorder range, filter signals, and drive the x movement from time.

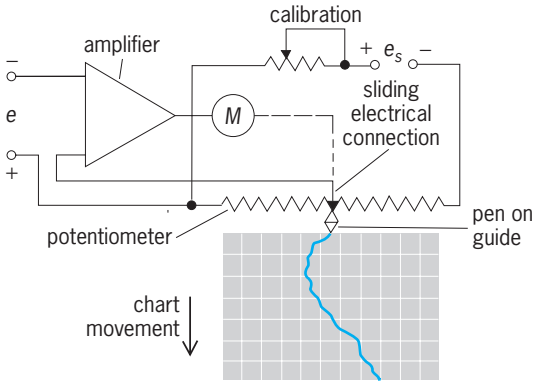

Fig. 4. System for potentiometric feedback positioning of a marker.

amplitude and possibly a different time shape than the actual signal. For multiple-trace systems, it can also result in incorrect phase information between the channels. Dynamic response is described by the time taken to trace a stated amplitude on the exhibiting medium. A high-speed oscillograph can trace a full-scale waveform in a millisecond, whereas a typical ink pen recorder will need 250 ms to cover the width. The buffered digital plotter largely avoids dynamic plotter response problems by recording the data in a high-speed electronic buffer, plotting output at a speed suited to the observer.

Transient recorders. Recording details of fast-changing signals requires considerable chart length. If the events desired only rarely arise, a transient recorder can be used. Here incoming data are continuously captured by using digital electronic storage. As new data are entered, the first part of earlier recorded data is destroyed. If the set of data held at any time contains a recognizable event of interest, the output is fed to a conventional recorder. Transient recorders are used to capture transients on electrical power lines and to record large amplitudes of seismic vibrations.

Application. The extensive acceptance of computer-based data logging and the general availability of hard-copy printers have contributed to a shift in usage from dedicated chart recorders to digital printers that provide permanent displays in the form of computer printouts. The plotting mechanisms incorporated within printers are similar to those of conventional graphic recording instruments, the majority being driven with stepping motors and using thermal- or laser-based printing. Dedicated graphic recording devices are still in use because it is often not economical to replace them with modern equivalents. They also find application where records need longer paper lengths than a computer printer can provide. *See* COMPUTER PERIPHERAL DEVICES; STEPPING MOTOR.

Peter H. Sydenham

Bibliography. G. K. McMillan (ed.), *Process/Industrial Instruments and Controls Handbook*, 5th ed., 1999; L. Michalski et al., *Temperature Measurement*, 2d ed., 2001; R. L. Moore, *Process Analyzers and Recorders*, vol. 2 of *Basic Instrumentation Lecture Notes and Study Guide*, ISA, rev. ed., 1989; B. E. Noltingk, *Instrumentation Reference Book*, 2d ed. 1995; J. G. Webster (ed.), *The Measurement, Instrumentation, and Sensors Handbook*, 1999.

Graphite

A low-pressure polymorph of carbon, the common high-pressure polymorph being diamond. Several other rare polymorphs have been synthesized or discovered in meteorites. The contrast in physical properties between these two polymorphs is remarkable: Graphite is metallic in appearance and very soft, whereas diamond is transparent and one of the hardest substances known.

Graphite is hexagonal, space group P 6_3/m mc, $a = 0.248$ nanometer, $c = 0.680$ nm, with 4C in the unit cell. Its atomic arrangement consists of sheets of carbon atoms at the vertices of a planar network of hexagons (**Fig. 1**). Thus, each carbon atom has three nearest-neighbor carbon atoms. The layer distance between the sheets is $c/2 = 0.340$ nm. Diamond, on the other hand, is a three-dimensional framework structure with the carbon atoms in tetrahedral (fourfold) coordination. Crystals of graphite are infrequently encountered since the mineral usually occurs as earthy, foliated, or columnar aggregates often mixed with iron oxide, quartz, and other minerals (**Fig. 2**). *See* CRYSTAL STRUCTURE; DIAMOND.

Properties. The sheetlike character of the graphite atomic arrangement results in distinctive physical properties. The mineral is very soft, with hardness $^1/_2$; it soils the fingers and leaves a black streak on paper, hence its use in pencils. The specific gravity is 2.23, often less because of the presence of pore spaces and impurities. The color is black in earthy material to steel-gray in plates, and thin flakes are deep blue in transmitted light. One perfect cleavage is parallel to the hexagonal sheets, allowing the mineral to be split into thin flexible but nonelastic folia.

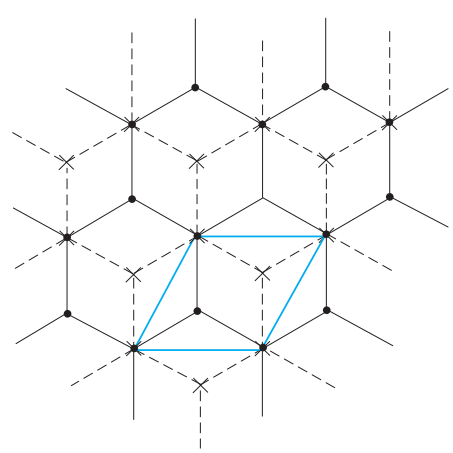

Fig. 1. The graphite atomic arrangement down the *c* axis. There are two distinct layers of carbon atoms, related by symmetry. The first level consists of carbon atoms (black circles) at the vertices of a network of hexagons. The second level, also a hexagonal network with carbon atoms at the vertices (crosses), is shifted relative to the first level.

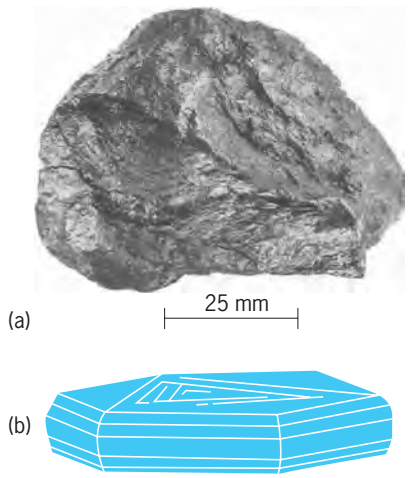

Fig. 2. Graphite. (*a*) Earthy aggregate (*American Museum of Natural History specimen*). (*b*) Hexagonal crystal of graphite with triangular markings on face (*after C. Palache, H. Berman, and C. Frondel, Dana's System of Mineralogy, vol. 1, 7th ed., John Wiley and Sons, 1944*)

Graphite is a conductor of electricity, distinguishing it from amorphous carbon (lampblack).

Graphite closely resembles molybdenite, MoS_2, a mineral with similar crystal structure, but the two can be distinguished by the greenish streak and much higher (4.70) specific gravity of the latter mineral. Early confusion with the brittle gray lead sulfide, galena, resulted in the synonymous trade names plumbago and black lead for graphite. In a similar manner, the misleading term "lead pencil" has persisted and is still in common usage.

Occurrence. Graphite arises from the thermal and regional metamorphism of rocks such as sandstones, shales, coals, and limestones which contained organic products not exposed to an oxidizing environment. It also can form in a strongly reducing environment, such as in serpentinites and limestones where hydrogen gas may reduce carbon dioxide. Platy graphite showing crude crystal surfaces often occurs speckled in coarsely crystallized marbles. The major sources of graphite are in gneisses and schists, where the mineral occurs in foliated masses mixed with quartz, mica, and so on. Noteworthy localities include the Adirondack region of New York, Korea, and Sri Lanka. In Sonora, Mexico, graphite occurs as a product of metamorphosed coal beds. Graphite is also observed in meteorites, where the mineral was formed under strongly reducing conditions, usually in association with metallic iron. Paul B. Moore

Synthetic Graphite

Graphite has a highly developed crystalline structure, and its softness, high thermal and electrical conductivity, and self-lubricating qualities differentiate it from other forms of carbon.

Carbon in graphitic form has both metallic and nonmetallic properties. Commercially produced synthetic graphite is a mixture of crystalline graphite and cross-linking intercrystalline carbon. Its physical properties are the result of contributions from both sources. Thus, among engineering materials, synthetic graphite is unusual because a wide variation in measurable properties can occur without significant change in chemical composition.

At room temperature the thermal conductivity of synthetic graphite is comparable to that of aluminum or brass. An unusual property of graphite is its increased strength at high temperature. The crushing strength is about 20% higher at 1600°C (2900°F) and the tensile strength is 50–100% higher at 2500°C (4500°F) than at room temperature.

Graphite is resistant to thermal shock because of its high thermal conductivity and low elastic modulus. It is one of the most inert materials with respect to chemical reaction with other elements and compounds. It is subject only to oxidation, reaction with and solution in some metals, and formation of lamellar compounds with certain alkali metals and metal halides.

Uses. Graphite has many uses in the electrical, chemical, metallurgical, nuclear, and rocket fields: electrodes in electric furnaces producing carbon steel, alloy steel, and ferroalloys (**Fig. 3**); anodes for the electrolytic production of chlorine, chlorates, magnesium, and sodium; motor and generator brushes; sleeve-type bearings and seal rings; rocket motor nozzles; missile nose cones; metallurgical molds and crucibles; linings for chemical reaction vessels; and, in a resin-impregnated impervious form, for heat exchangers, pumps, pipings, valves, and other process equipment.

Preparation. Synthetic graphite can be made from almost any organic material that leaves a high carbon residue on heating to 2500–3200°C (4500–5800°F). In commercial operations, raw materials are carefully selected because not all substances with high carbon content undergo a suitably complete transformation to graphite at these temperatures. Petroleum coke is raw material for the most commonly used production process. After calcining and sizing, the coke is mixed with coal tar pitch, heated to about 165°C (329°F), and formed by extrusion or molding to so-called green shapes. Baking to 750–1400°C (1400–2600°F) in gas- or oil-fired kilns follows the forming operation. Graphite is produced by heating the baked shapes to 2600–3000°C (4700–5400°F) by passing electricity amounting to 1.6–3.0 kW per pound of graphite through the bed of a furnace

Fig. 3. Graphite reflector for a test reactor.

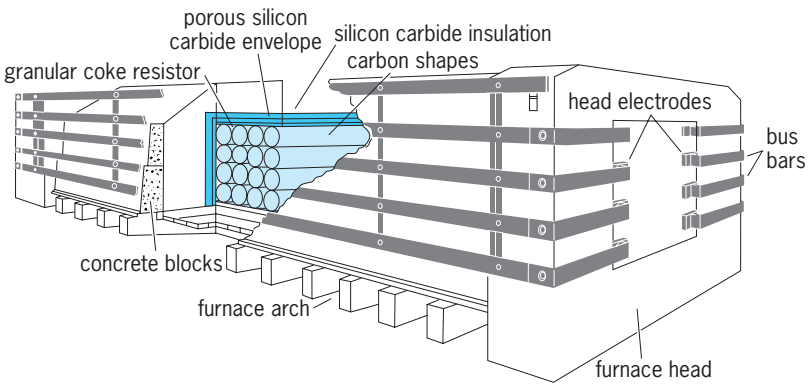

Fig. 4. Graphitizing furnace, for making synthetic graphite.

made of the shapes laid in granular coke (**Fig. 4**). The whole bed is covered by an insulating blanket of silicon carbide, coke, and sand. Higher-density synthetic graphite can be obtained by impregnating the baked carbon with pitch prior to graphitization. Graphite with total ash content less than 20 parts per million is needed for a number of nuclear and electrolytic uses, and is obtained by heating the graphite shapes electrically to about 2500°C (4500°F) while bathing them in a purifying gas. *See* CARBON; COKING (PETROLEUM); ELECTROCHEMICAL PROCESS.

Highly ordered crystalline graphite can be produced up to about $^1/_4$-in. (0.6-cm) thickness by pyrolyzing organic gases under controlled conditions at 1400–2000°C (2600–3600°F). Pyrolytic graphite exhibits a high degree of anisotropy (varying properties in different directions). Parallel to the thickness the thermal conductivity is comparable to copper, but perpendicular to the thickness conductivity is about 1/200 the conductivity of copper. Tensile strength and thermal expansion also vary greatly with orientation. The room temperature density of pyrolytic graphite reaches up to 2.22, about 98% of the 2.26 density of the graphite single crystal, whereas the density of graphite electrodes ranges from 1.5 to 1.7. Pyrolytic graphite has been formed into rocket nose cones.

Approaching the density limit, very highly oriented graphite can be prepared by stress-annealing pyrolytic carbons. Though polycrystalline, the highly oriented pyrolytic graphite exhibits most of the properties of the ideal graphite crystal and, therefore, has been very useful in studies of the properties of graphite and its intercalation compounds. Because of its high reflectivity, the highly oriented pyrolytic graphite has important applications as monochromators for x-rays and thermal neutrons.

Graphite Fibers

Carbon fibers are filamentary forms of carbon, with a fiber diameter normally in the 6–10-micrometer range. The product is offered in the form of yarns or tows containing from 1000 to 500,000 filaments per strand. The fibers offer a unique combination of properties. They are flexible, lightweight, thermally and to a large extent chemically inert, and are good thermal and electrical conductors. In their high-performance varieties, carbon fibers are very strong and can be extremely stiff. The fibers are made by carbonizing (charring) an organic polymer yarn. Many precursor materials have been proposed, but only three, rayon, polyacrylonitrile (PAN), and pitch, are being used in commercial production. The descriptive terms carbon and graphite are sometimes used to indicate that the material has been carbonized at a temperature below approximately 2500°C (4500°F, carbon) or above (graphite). Quite frequently, however, they are used synonymously. With the exception of some very-high-modulus (over 80,000,000 lb/in.^{2} or 552 gigapascals) pitch-based fibers, the commercial products do not have the three-dimensional order characteristic of graphite, and should properly be called carbon.

Preferred orientation. The most important parameter to characterize carbon fibers is the preferred orientation, that is, the extent to which carbon (or graphite) microcrystallites are oriented parallel to the fiber axis. **Figure 5** shows the Young's modulus E_c, corrected for fiber porosity, as a function of the orientation q. Fibers from all three raw materials fall on the same curve, in excellent agreement with theory. The thermal and electrical conductivity of carbon fibers have a similar increase with greater orientation. The longitudinal (in the direction of the fiber axis) coefficient of thermal expansion in the −100 to +100°C (−148 to +212°F) temperature range becomes slightly negative when the fiber modulus approaches 30,000,000 lb/in.^{2} (207 GPa), and decreases further for fibers with higher preferred orientation. This property is important because it permits the design of composite structures, such as large space antennas and mirrors, having a geometry that is not affected by temperature changes.

The intrinsic tensile strength of carbon fibers also increases with increased preferred orientation. For some fibers, this value probably exceeds 1,000,000 lb/in.^{2} (6.9 GPa). However, the practically

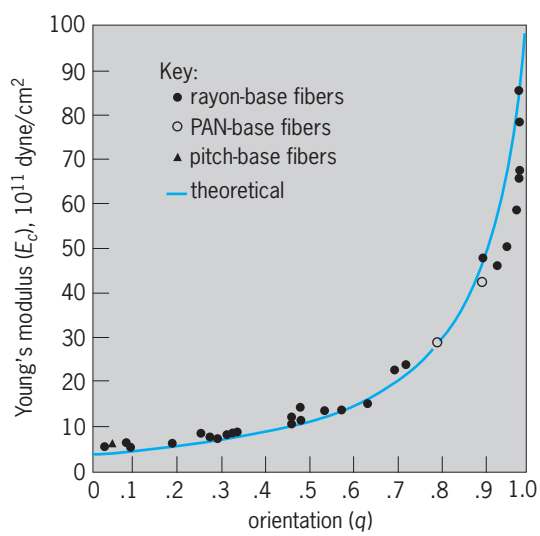

Fig. 5. Young's modulus (E_c) shown as a function of the orientation q; 1 dyne/cm^2 = 0.1 pascal. (*After R. Bacon, Carbon fibers from rayon precursors, Chem. Phys. Carbon, 9:1–102, 1973*)

useful fiber strength depends on the frequency and the severity of strength-limiting flaws such as surface irregularities, holes, or foreign particle inclusions in the fibers, and rarely exceeds 500,000 lb/in.^{2} (3.4 GPa).

Categories. Carbon fibers fall into two categories: low-modulus (under 20,000,000 lb/in.^{2} or 138 GPa) and high-performance (moduli above 25,000,000 lb/in.^{2} or 172 GPa).

Low-modulus fibers. These products are intended primarily for nonstructural applications. They are made from rayon or pitch and are offered as yarns, cloth, felt, and mats. The yarns are impregnated with fluoropolymer resin or graphite powder, braided, and used as packing materials for pumps and valves. Formerly limited to some high-temperature uses where the performance of asbestos packings was inadequate, they are expanding because of environmental concerns about asbestos usage. Rayon-based cloth products are used as reinforcement in resin composites where short-time high-temperature strength and good ablative characteristics are desirable. Cloth carbonized to temperatures above 2100°C (3800°F) is used in rocket nozzle throats and ablation chambers. Low-temperature (under 1800°C or 3300°F) cloth is used where good strength and low thermal conductivity are required, such as in reentry vehicle heat shields, rocket nozzle entrance sections, and exit cones of spacecraft. Rayon-based felts are also used in phenolic matrices as heat shield materials. Other applications are as thermal insulations for high-temperature vacuum furnaces and as the electrical conductor for the sulfur side of sodium/sulfur batteries. Carbon fiber mats are made predominantly from a pitch precursor. The carbon is used in the form of a thin veil mat (50–100 g/m^2 or 0.16–0.32 oz/ft^2) as a surface layer in automotive sheet-molded components (for example, hoods), where it provides sufficient electrical conductivity for electrostatic shielding and electrostatic painting without a conductive primer. Other applications are as thermal insulation similar to carbon felt, and, in a chopped form, as reinforcement for injection moldings to provide high stiffness, good electrical conductivity, and resistance to abrasion. Automotive brake formulations containing short mat fibers as a partial replacement for asbestos are being evaluated.

High-performance fibers. These were originally also made from rayon. To obtain the high modulus and high tensile strength, it was necessary to stretch these fibers to several times their original length at carbonizing temperatures above 2800°C (5100°F). The high cost of this stretching process has made these rayon-based fibers obsolete. High-performance carbon fibers are made from either a PAN or pitch precursor. Polyacrylonitrile is wet- or dry-spun and stretched 500–2000% at approximately 100°C (212°F) to increase the polymer orientation. Ordinary pitch is isotropic and forms a fiber with little or no preferred orientation. By first converting the pitch to a mesophase or liquid crystal state, it is possible to spin very highly oriented pitch fibers whose orientation is maintained and further enhanced during carbonization. After spinning, the PAN and pitch fibers are still thermoplastic and are converted into thermosets by oxidative cross-linking. During this operation, PAN fibers must be kept under tension to maintain the preferred orientation; mesophase pitch fibers can be oxidized without tension at significant savings in process costs. After cross-linking, both types of fibers are carbonized in an inert (nitrogen) atmosphere. The chemical carbon yield from PAN is about 50%, and from mesophase pitch approximately 80%. The carbonized fibers normally receive a surface treatment to enhance the resin-fiber bonding and, hence, the shear strength and transverse tensile

Typical properties of high–modulus carbon yarns and their composites*

Yarns	Units	Thornel† 300	Thornel† 50 (PAN)	Thornel‡ P55S	Thornel‡ 75S
		PAN and pitch			
Tensile strength (10 in. gage length)	10^3 lb/in.^{2}	500	350	300	300
Tensile modulus	10^6 lb/in.^{2}	33.5	55	55	75
Density	g/cm^3, Mg/m^3	1.75	1.81	2.02	2.06
Electrical resistance	μohm-m	18	9.5	7.5	4.6
Thermal conductivity	Cal/s-cm K	.02	.16	.26	.44
		Unidirectional laminates (60% fiber volume in high-performance epoxy)			
Tensile strength (coupon)	10^3 lb/in.^{2}	250	195	150	115
Tensile modulus	10^6 lb/in.^{2}	20	33	34	44
Flexural strength	10^3 lb/in.^{2}	260	190	120	100
Flexural modulus	10^6 lb/in.^{2}	19	29	28	38
Compressive strength	10^3 lb/in.^{2}	230	120	70	60
Compressive modulus	10^6 lb/in.^{2}	20	30	28	35
Short beam shear	10^3 lb/in.^{2}	17	10	5	9
Coefficient of thermal expansion	$\Delta l/l^\circ$ F $\times 10^{-6}$				
Longitudinal $\times 10^{-6}$		−.3	−.4	−.5	−.7
Transverse $\times 10^{-6}$		15	15	1515	15

*1 in. = 2.54 cm; 10^3 lb/in.^{2} = 6.895 MPa; 10^6 lb/in.^{2} = 6.895 GPa; $\Delta l/l^\circ$ C $\times 10^{-6}$ = 1.8$\Delta l/l^\circ$ F $\times 10^{-6}$.
†Typical of intermediate and high-modulus PAN-based fibers. Thornel is a registered trademark of Union Carbide Corporation.
‡Pitch-based fibers.

strength of the composite. Most, if not all, carbon fibers also receive a thin coating of a polymer finish or sizing for the primary purpose of protecting the brittle fibers from self-abrasion during shipping and handling.

The principal use of high-performance carbon fibers is as the reinforcing component in structural composites. The **table** lists the representative properties of some PAN and pitch-based carbon fibers and of their epoxy matrix composites. Due to initially high cost, the original applications were almost exclusively for lightweight, high-stiffness, and high-strength composites for the aerospace industry.

The second major usage of high-performance carbon fibers is in sporting goods, such as golf club shafts, tennis rackets, fishing rods, and sailboat structures. The major matrix material for both aerospace and sporting goods applications is epoxy; polyimides are used where higher temperature requirements (350–650°F or 177–343°C) exist. Other resins, including polyesters and thermoplastics such as nylon and polyphenylenesulfide, are increasingly used to lower fabrication costs for industrial and automotive applications. Metal matrices, primarily aluminum and magnesium, are used in development programs for space structures such as antennas. These metal matrix composites require the use of extremely high-modulus (over 75,000,000 lb/in.^{2} or 517 GPa) carbon fibers. *See* COMPOSITE MATERIAL; MANUFACTURED FIBER; METAL MATRIX COMPOSITE.

H. F. Volk

Bibliography. D. B. Chung, *Carbon Fiber Composites*, 1994; J. Delmonte, *Technology of Carbon and Graphite Fiber Composites*, 1981, reprint 1987; M. S. Dresselhaus et al. (eds.), *Graphite Fibers and Filaments*, 1988; L. H. Peebles, Jr., *Carbon Fibers: Information, Structure, and Properties*, 1994.

Graptolithina

A group of marine organisms that were common in the early Paleozoic (the Late Cambrian to Early Devonian periods). Graptolites became extinct in the late Paleozoic Era. They were minute animals that built communal skeletons. Each graptolite colony contained from two to many hundreds of separate graptolite animals (referred to individually as zooids). All of the zooids within a single colony (a rhabdosome) were formed by asexual budding from the founder zooid, which probably grew from a fertilized egg. Thus, each colony started with a sexually produced animal and then enlarged by budding new zooids from one another. The result was a spreading rhabdosome composed of a few to many hundreds of minute tubes (thecae), each containing its own zooid (**Fig. 1**). Graptolite thecae exhibit a range of diameters (30 μm–2 mm), but are most commonly about 0.5–1.0 mm in diameter and several times this length. The thecae consist of the protein collagen, which is organized into two main building blocks. The most prominent are the fuselli, which have the form of narrow half-rings stacked one upon the other like a set of semicircular bricks. Collectively, the fuselli formed the main substance of the thecal tube. The second building block was deposited on the outside of the theca in the form of bandagelike strips or thin continuous sheets (**Fig. 2**). Also built of collagen fibers, this second layer (the cortex) is much thinner, but probably added greatly to the strength of the theca.

With the exception of a small group of living organisms (for example, *Rhabdopleura* of the class Pterobranchia; Fig. 1) that may be relatives of the graptolites, little is known for sure of the detailed anatomy and ecology of these enigmatic organisms. Although graptolites are common as fossils in dark shales that were deposited under conditions of low oxygen concentration, the original soft body parts are not preserved—only the empty tubes remain. Thus, many basic features (such as exactly how they built their colonies) are still subjects of debate among specialists. Most specialists think that graptolites and pterobranchs are closely related to one another, and that graptolites employed a specialized organ similar to that used by *Rhabdopleura* zooids to secrete the fuselli and cortex from their body (externally) in a mortaring fashion (**Fig. 3**). A smaller group of specialists interpret the evidence very differently. They suggest that the thecae were formed internally beneath an enveloping tissue layer by zooids unlike those of *Rhabdopleura*, to which graptolites, therefore, would not be closely related.

Based on the shape and patterns of occurrence of their colonies, graptolites were probably suspension feeders that extracted food particles such as bacteria, algae, and other organic matter from the surrounding ocean water. Many graptolite species were very widely distributed throughout the world's oceans during the Ordovician, Silurian, and Early Devonian periods (an interval of about 110 million years in total). Graptolites evolved very rapidly. Most species occupied a relatively short segment of this portion of the Earth's history. The average duration for individual graptoloid graptolite species was approximately a million years. As a result of this combination of wide occurrence and short duration, graptolite species make very useful index fossils. For example, finding specimens of a particular species in both China and Australia indicates that the rocks containing these two fossils were formed during the same short time of the Earth's history when this particular species lived. Using relations like these, paleontologists and geologists have untangled the complex history of the Earth. *See* DEVONIAN; FACIES (GEOLOGY); INDEX FOSSIL; ORDOVICIAN; SILURIAN.

Types. Graptolites are mainly of two sorts: bottom-living (benthic) and free-floating (planktic). Benthic graptolites generally either formed horizontal, sheet-like colonies attached to the sea floor and other hard objects like shells, or grew as upright, bushy or cone-shaped colonies. Some of these erect benthic colonies became relatively large, longer than 10–15 cm (4–6 in.) from base to tip. Many of these benthic graptolites are grouped by graptolite taxonomists as the Dendroidea. Benthic graptolites also

include several other graptolite orders (such as the Tuboidea); however, these are very small and very rare fossils. Dendroids are characterized by long, tubular thecae of two distinctly different sizes that are arranged in pairs within their branches. The larger thecae are known as autothecae, and the smaller as bithecae. Thecae within the dendroid colony were interconnected through a special chord called the stolon, which probably permitted coordinated activity and the sharing of nutrients among zooids. Benthic graptolites are less common than planktic graptolites and appear to be of limited value as index fossils. Thus, they are less well studied and less well understood than the planktic graptolites. For instance, it is unclear why the thecae have a regular division into autothecae and bithecae. It may reflect some division of labor among different-sized zooids, or sexual dimorphism, or possibly some other factor. Dendroids first appeared in the Late Cambrian Period (around the same time as most other marine invertebrate animals, including arthropods, brachiopods, mollusks, and echinoderms) and became extinct in the Carboniferous Period.

In contrast to the benthic graptolites, species of the Graptoloidea were planktic. The founding individual of the colony in these graptolites did not settle to the ocean floor to complete its maturation as did the founders of dendroid rhabdosomes (Fig. 1*a*). Instead, a graptoloid larva matured while in the ocean currents, and there it initiated formation of its rhabdosome. This change in habit was accompanied by a corresponding change in the form of the tube constructed by this zooid (the sicula). The sicula of graptoloid graptolites is conical with a fine threadlike extension (nema) on the closed end (Fig. 1*b*). The earliest planktic graptolites retained many of the features of their dendroid ancestors, including the many-branched colony form and differentiation of the thecae into autothecae and bithecae. Thus, these early graptoloids are not easily distinguished from some dendroid species except by the form of the sicula and their planktic habit. The evolutionary appearance of planktic graptolites is one of the key events that marks the beginning of the Ordovician Period.

Early Ordovician graptoloids evolved simpler colonies over the next several tens of millions of years. Their rhabdosomes came to possess fewer branches and to consist of only large thecae comparable to the autothecae of dendroids. Whatever purpose bithecae served in benthic graptolites, they evidently were not beneficial to planktic graptolites. These changes, which resulted in lighter, more streamlined, and possibly more mobile rhabdosomes, are presumed to be adaptations to life in the plankton, but once again the exact nature of the advance afforded by these changes is the subject of continued study.

During the Ordovician the evolutionarily advanced (but simpler) graptoloids diversified into a great many different species with a wide range of colony forms and sizes (**Fig. 4**). At the same time that the overall rhabdosome form became simpler, the form of the individual zooid tubes (thecae) became much more varied and complex. Some graptoloids exhibit hooked, sharply flexed, hooded, or spiny forms. In some species the rhabdosome was reduced to a scaffoldinglike set of rods without any continuous skeletal material between them. The colony form of graptoloids shows numerous evolutionary trends, but overall quite different forms dominate different intervals during the Ordovician to Early Devonian. Graptoloids became extinct in the Early Devonian.

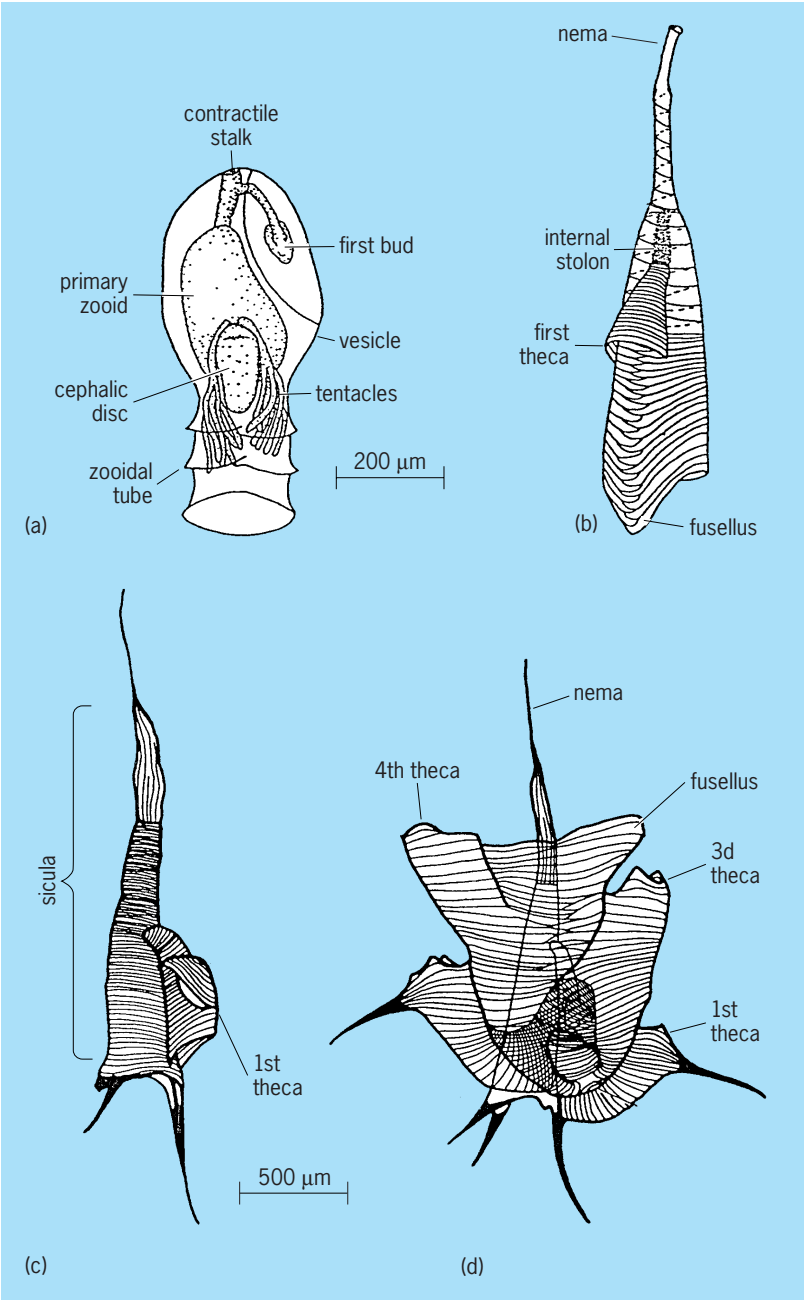

Fig. 1. Sketches of portions of pterobranch and graptolite colonies showing the dwelling tubes and the constructional units of the graptolite skeleton. (*a*) Young *Rhabdopleura compacta* colony with founder zooid and first asexual bud (Recent). Benthic graptolites may have been similar to this pterobranch. (*b*) Comparable graptoloid graptolite sicula (Early Ordovician). (*c, d*) Later growth stages of the more derived graptoloid *Hustedograptus uplandicus* (Middle Ordovician).

Fig. 2. Scanning electron micrograph of a graptolite (*Pseudoclimacograptus scharengergi*) showing the fuselli (horizontal ridges) that compose graptolite thecae, and the cortical bandages (oblique strips) that were deposited on top of the fuselli (Middle Ordovician); specimen is approximately 1 mm in width.

Preservation. Graptoloids lived mostly in oxygen-minimum zone waters at shelf margins. Thus, they are most common in rocks deposited in outer-shelf and slope environments. Most graptoloid graptolites are preserved in gray to black shales as black or silvery films flattened on the surface of the rock. Rough handling or careless washing of the rock may completely remove these delicate fossils. Occasionally, the mineral pyrite may form within the hollow graptolite rhabdosomes prior to compression of the original mud into shale, and an exquisite three-dimensional fossil may result. Rarely, graptolites may be preserved in limestones. Often graptolites in these rocks are preserved uncompressed, and with the use of hydrochloric or acetic acid, they can sometimes be liberated from their limestone matrix (Fig. 2). Most understanding of graptolites' detailed structure comes from material of this kind.

Life history and extinction. The evolutionary history of planktic graptolites was complex, and studies of their patterns of morphological change and fluctuations in species diversity have provided important examples for evolutionary paleobiology about the nature of adaptation to this unusual lifestyle and to the causes of changes in organic diversity in the world's oceans. For instance, the biosphere experienced a catastrophic decline in species number among a wide array of marine invertebrates, but especially among trilobites, brachiopods, conodonts, corals, and graptolites near the end of the Ordovician Period. This mass extinction has been estimated to have been the second or third most severe extinction event in the history of marine metazoans. During this event, planktic graptolites almost became entirely extinct.

Many planktic graptolite species became extinct during the Late Ordovician due to environmental changes that have been linked to global climate changes. Environmental changes in the oceans included a sea-level fall of as much as 100 m (330 ft) and the loss of many oceanic upwelling sites located near continental shelf margins. These changes in ocean conditions developed as a consequence of global climate change from an ice-free, "greenhouse" world climate to an "icehouse" global climate lasting approximately 1 million years during which ice covered the South Pole. Lowered sea level during glaciation drained most marine environments that were spread broadly across shelves of the time. When the ice melted, global climates and marine environments returned to preglaciation conditions.

The majority of preglaciation Late Ordovician graptolites lived in hypoxic (oxygen-deficient) waters, most of which developed in oxygen-minimum zones under ocean upwelling waters at shelf margins. Some planktic graptolites lived in hypoxic environments in basins within shelves. Locally endemic species occur in a number of these shelf basins. A few

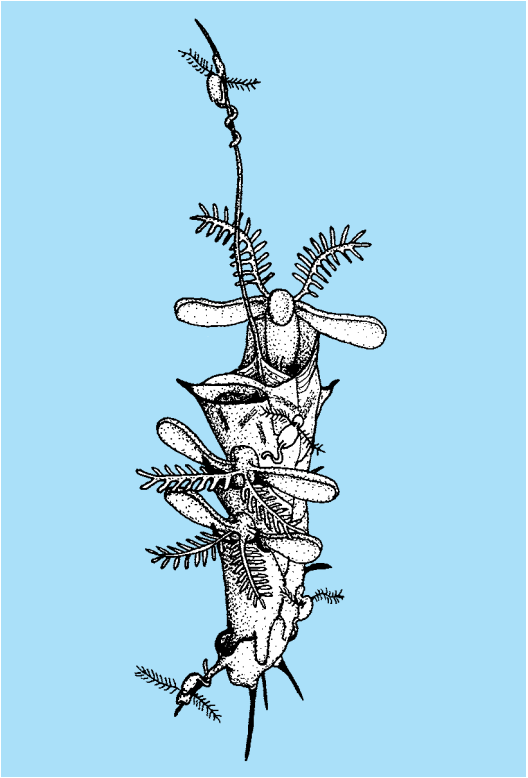

Fig. 3. Hypothetical form of zooids that inhabited colony tubes, based on a planktic graptolite. Large zooids are shown with possible oarlike swimming organs used to move the colony through the ocean water and tentacles for feeding. Also shown are two possible modes of graptolite colony construction: large zooids linked together built the whole colony, or smaller free-roaming zooids may have helped the larger zooids by adding to the cortex and spines. A third hypothesis (not shown) involves an entirely different zooid reconstruction (see text).

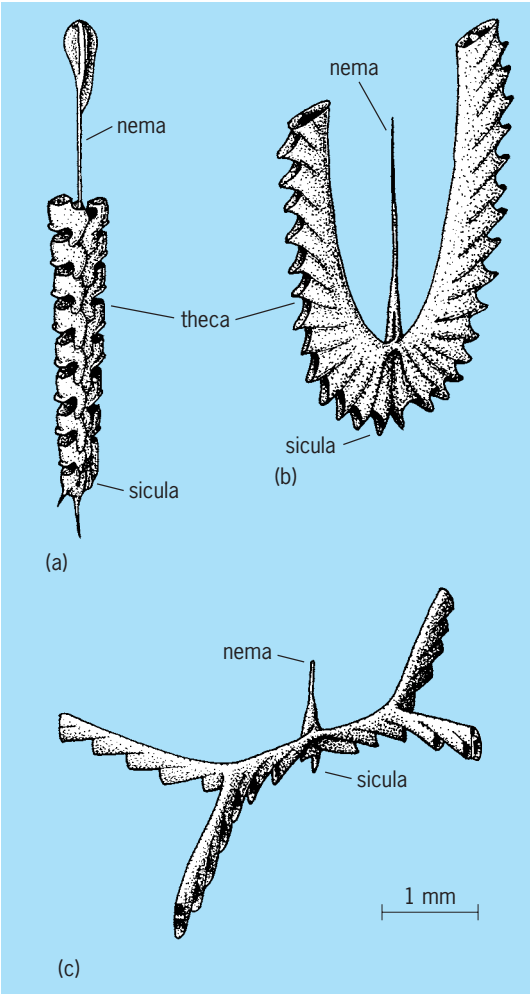

Fig. 4. Examples of some planktic graptolites, reconstructed to show three-dimensional form.

Late Ordovician graptolites (primarily those called normalograptids) lived in relatively oxic (oxygen-containing) waters near or at the ocean surface. Some of these species occur in both shelf basin and shelf marginal rock sequences.

Hypoxic environments preferred by many graptolites were lost during glaciation as a consequence of the draining of shelf basins and the loss of shelf margin oxygen-minimum zones. Most planktic graptolites living in hypoxic environments became extinct during the glacial episode. In contrast, those graptolites living in oxic waters thrived during glaciation. They even radiated modestly with a number of new species appearing in certain areas.

A Late Ordovician–Early Silurian planktic graptolite refugium (an area that has escaped the great changes that occurred in the region as a whole, often providing conditions in which relic colonies can survive) developed in a shelf basin (Yangtze Basin) in modern south China. Freshwater runoff from surrounding land areas flowed into that basin. Because freshwater is less dense than the deeper basinal waters, they most likely served as a seal that enabled the deep, hypoxic basin waters to be density-stratified. Preglaciation Late Ordovician planktic graptolites living in hypoxic waters beneath the freshwater seal of that basin survived through the interval during which other graptolites living in sites where hypoxic waters were diminished or had disappeared became extinct. Turnover among species and rates of extinction of species living in this basin remained relatively similar to preclimate change conditions throughout the climate change interval.

Deglaciation and resultant sea-level rise gradually led to a return of the preglaciation ocean conditions. Many new graptolite species appeared during the early phases of sea-level rise and redevelopment of ocean upwelling sites with oxygen-minimum zones and shallow shelf basins in which hypoxic conditions redeveloped.

Viruses, bacteria, and other microorganisms as well as floating larvae of many organisms are abundant in many modern hypoxic environments. Molecular biologists have discovered an extensive system of gene exchanges mediated by plasmids, episomes, and viruses among bacteria. Viruses in hypoxic environments may invade bacteria and obtain a portion of the host genetic material. When the host wall ruptures, viruses containing new genetic information may spread out to infect another host and transfer genetic information from the old to new host.

Late Ordovician graptolites living in hypoxic to anoxic (oxygen-absent) waters likely became crowded together in small enclaves as their environments diminished during the glaciation-related sea-level fall and the loss or near-loss of the hypoxic environments in which they lived. Thus, viral transfer of genetic information could have taken place between graptolite larvae in Late Ordovician diminished hypoxic environments. Lateral transfer of genetic information by viruses, if it did take place, could have been one mechanism for the relatively rapid species originations seen in strata deposited during post-glacial sea-level rise and redevelopment of relatively broader hypoxic environments.

Although planktic graptolites radiated relatively rapidly to a number of lineages during the early part of the Silurian, they appear to have undergone as many as eight significant reductions in species number during the Silurian. During one event that took place in the mid-Silurian (Late Wenlock), planktic graptolites almost became extinct. Nearly all of the survivors of this near-extinction lived in oxic waters. Some of the other severe reductions in graptolite taxa that took place during the Silurian have been linked to climate-induced sea-level changes and related environmental changes. The majority of pre-Late Wenlock Silurian planktic graptolites appear to have lived in or near hypoxic waters. The marked reductions in planktic graptolite diversity at intervals within the pre-Late Wenlock Silurian appear to coincide with diminished hypoxic environments in a pattern similar to that of the Late Ordovician extinction.

The majority of post-Late Wenlock graptolites probably lived in oxic waters where they no longer had the protection from predators afforded by life in hypoxic environments. Many taxa appear to have lived in waters over the shelves of the time.

Graptolite species diversity diminished significantly during the Late Silurian into the Early Devonian. Finally, planktic graptolites became extinct late in the early part of the Devonian. *See* EXTINCTION (BIOLOGY). William B. N. Berry; Charles E. Mitchell

Bibliography. D. E. B. Bates, Graptolites: Strange plankton of the past, *Endeavor, New Series*, 13:54–62, 1989; W. B. N. Berry and H. Hartman, Graptolite parallel evolution and lateral gene transfer, pp. 397–404 in M. Syvanen and C. L. Kado (eds.), *Horizontal Gene Transfer*, 2d ed., Academic Press, San Diego, 2002; W. B. N. Berry, Late Ordovician climate change: Impact on graptolite biodiversity. *Earth Systems Processes 2, Abstracts with Programs 55*, 2005; O. M. B. Bullman, in R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, Pt. V: *Graptolithina*, Kansas University Press, 1970; P. R. Crowther, The Fine Structure of Graptolite Periderm, *Spec. Pap. Palaeont.*, no. 26, 1981; S. C. Finney and W. B. N. Berry, New perspectives on graptolite distributions and their use as indicators of platform margin dynamics, *Geology*, 25:919–922, 1997; A. Hallam and P. B. Wingnall, *Mass Extinctions and Their Aftermath*, Oxford University Press, 1997; C. E. Mitchell et al., Was the Yangtze Platform a refugium for graptolites during the Hirnantian (Late Ordovician) mass extinction?, pp. 523–526 in G. L. Albanese, M. S. Beresi, and S. H. Peralta (eds.), *Ordovician from the Andes*, Instituto Superior de Correlacion Geologica (INSUEGO) Serie Correlacion Geologica 17, 2003; R. B. Rickards, Palaeoecology of the Graptolithina, an extinct class of the phylum Hemichordata, *Biol. Rev.*, 50:397–436, 1975; P. Storch, Biotic crises and post-crisis recoveries recorded by Silurian planktonic graptolite faunas of the Barrandian area (Czech Republic), *Geolines*, 3:59–70, 1995.

Grass crops

Members of the family Gramineae cultivated as forage and grain for consumption. The grasses are the most useful of all the plants that cover the Earth. The cereal grasses (rice, wheat, maize, rye, barley, oats, sorghum, and the millets) supply directly three-fourths of the energy and over half of the protein in food consumed by humans. Indirectly, these cereals together with the forage grasses supply most of the food for the domestic animals that provide milk, meat, eggs, and much of the draft power required to grow crops. Deer, antelope, rabbits, and many other wild game depend on grasses for much of their sustenance. Sugarcane produces more than half of the world supply of sugar. Starch, and most of the alcohol for beverage and industrial uses, comes from the cereal grasses. *See* BARLEY; CEREAL; CORN; MILLET; OATS; RICE; RYE; SORGHUM; SUGARCANE; WHEAT.

The bamboos are of vast importance in the Indo-Malay region, where they are used in building houses, bridges, furniture, rafts, water pipes, vessels for holding water, and so forth. Slender bamboo stems make excellent fishing poles, whereas the giant bamboos, measuring up to 12 in. (30 cm) in diameter and over 100 ft (30 m) high, are important substitutes for wood. *See* BAMBOO.

The grasses protect soil from erosion and help conserve water resources. More than any other family of plants, the sod-forming grasses blanket golf courses, athletic fields, lawns, parks, and cemeteries with a protective covering that beautifies and enhances the environment. No other family of plants in the vast plant kingdom is so useful to humans. *See* EROSION; LAWN AND TURF GRASSES; SOIL CONSERVATION.

Structure. Grass stems have solid joints (nodes) and leaves arranged in two rows, with one leaf at each joint (**Fig. 1**). The leaves consist of the sheath, which fits around the stem like a split tube, and the blade, which is commonly long and narrow. Seed heads are made up of minute flowers on tiny branchlets, often several crowded together, but always two-ranked like the leaves. The flowers are generally wind-pollinated. The seeds are enclosed between two bracts, or glumes, which remain on the seed when ripe.

Growth characteristics. The 600 genera grouped into 14 tribes that make up the grass family may be annual or perennial. In the tropics, elephant grass (*Pennisetum purpureum*) can produce up to 37 tons of dry matter per acre (75 tons per hectare per year) whereas Tifdwarf Bermuda grass (*Cynodon* spp.) on a golf green will produce only a few pounds of matter in the same period. Annuals and some perennial grasses are bunch grasses which spread only by seeding. Others, mostly perennials, also spread by creeping stems called stolons when above ground and rhizomes when below the soil surface. Stoloniferous and rhizomatous species usually are most tolerant of continuous defoliation by animals. The creeping grasses form the best sods and surpass others for soil conservation; they are also the best turf grasses. All grasses have fine fibrous root systems that permeate the soil extensively to depths ranging from much less than 3 ft to more than 10 ft (1 to 3 m). The roots are short-lived, are continually being replaced, and in the process increase the organic matter content of the soil.

Distribution. Grasses are distributed throughout the world. A few profuse seeders with a very short life-cycle such as annual bluegrass (*Poa annua*) can be found growing from the Equator to the Arctic Circle. Annual species predominate in the adverse environments found in the deserts and the arctic areas. Temperature is the principal factor determining the distribution of perennial grasses. Most tropical grasses are killed when temperatures drop much below 32°F (0°C), whereas temperate perennial grasses can tolerate much lower temperatures. However, there is a great deal of overlapping of tropical and temperate grasses that may differ in temperature hardiness among varieties within some species. *See* BERMUDA GRASS.

Perennial grasses are frequently classed as cool- or warm-season grasses depending upon the season in which they make their best growth. Generally, cool-season grasses such as bluegrass (*Poa pratense*) are also temperate grasses and fail to survive the long hot

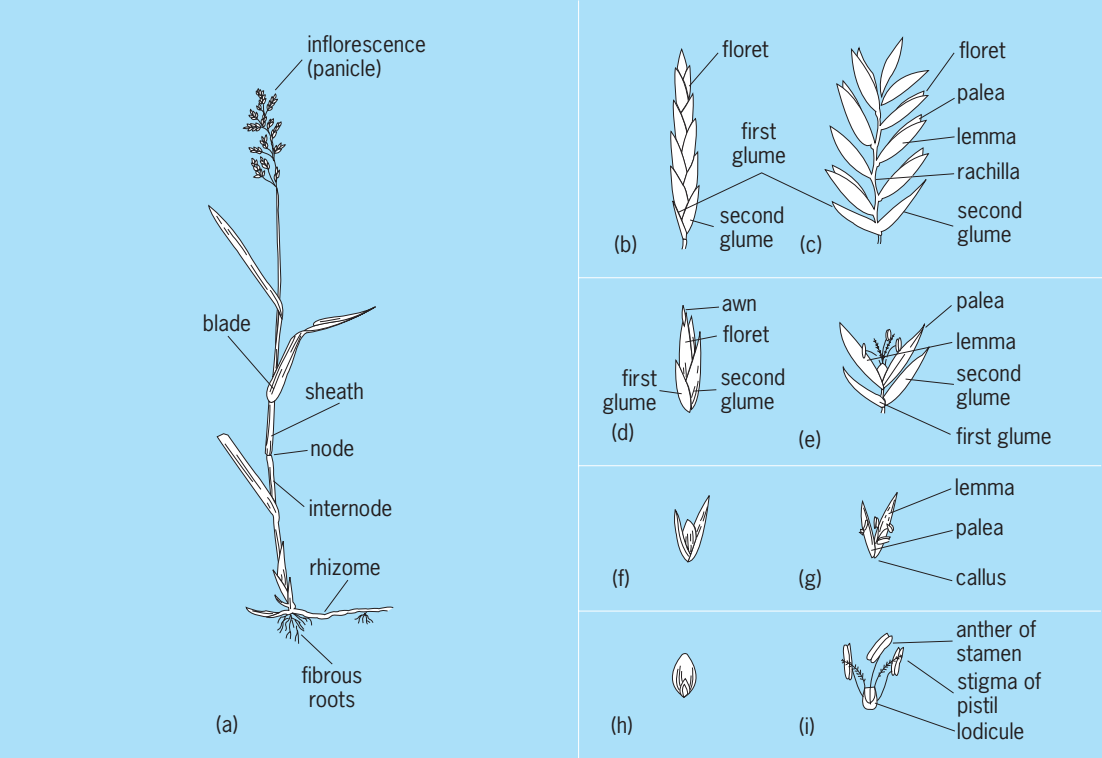

Fig. 1. Parts of a typical grass plant. (*a*) Complete plant. (*b*) Many-flowered spikelet. (*c*) Generalized spikelet. (*d*) One-flowered spikelet. (*e*) Spikelet at flowering. (*f*) Floret in fruit. (*g*) Floret in flower. (*h*) A fruit. (*i*) A flower. (*After P. D. Strausbaugh and E. L. Core, Flora of West Virginia, West Va. Univ. Bull., ser. 52, no. 12–2, pt. 1, p. 67, 1952*)

summers in the tropics. Usually diseases contribute to their death. Warm-season grasses taken into the temperate zone usually die during winter. *See* BLUEGRASS.

Some 10% of the 1400 or more species of grasses found in the United States originated in other countries. Many of the weedy ones were introduced accidentally with seed or feed brought in with livestock. Johnson grass, one of the South's worst weeds, was introduced from the Middle East as a forage grass for cattle. Pensacola bahia grass (*Paspalum notatum*) came to Pensacola, Florida, in the digestive tracts of cattle shipped from Santa Fe, Argentina, in the early 1900s. Today Pensacola bahia grass is widely grown in pastures and roadsides in Florida and the southern half of the states bordering the Gulf of Mexico. Most of the cool-season grasses planted in the United States for forage and turf originated in Europe. Among these, the bluegrasses (*Poa* spp.), fescues (*Festuca* spp.), bents (*Agrostic* spp.), cocksfoots (*Dactylis* spp.), and timothy (*Phleum pratense*) are the most common. Africa and South America are the original homes of many of the warm-season grasses such as the Bermuda (*Cynodon* spp.), the bahia and dallis grasses (*Paspalum* spp.), and Pangola grass (*Digitaria* spp.) that are planted in the southern United States. *See* BERMUDA GRASS; TIMOTHY.

Reproduction and propagation. Annuals and most perennial grasses reproduce sexually and are propagated by seed. Many of the grasses are cross-pollinated largely by the wind. In some species, cross-pollination is facilitated by self-incompatibility that occurs at variable frequencies. *See* POLLINATION.

Most grasses produce perfect flowers containing both male and female organs. The male organs (anthers) must be carefully removed before they shed pollen to make controlled hybrids. In maize, the male (in the tassel) and the female (in the ear) organs are separated, a characteristic that greatly facilitates hybrid production. Controlled hybridization without emasculation in the *Pennisetum* spp. is made possible by their characteristic exsertion of the female pollen-receptive stigmas at least 24 h before their pollen-shedding anthers. Cytoplasmic male sterility has been discovered in a few species and is used to produce commercial F_1 hybrid seed of maize, sorghum, and pearl millet. *See* FLOWER.

A few species, largely tropical perennials, reproduce by apomixis, simply defined as vegetative reproduction through the seed. Apomictic seeds produce the same genotype as the plant that produced them. An obligate apomict such as tetraploid bahia grass rarely produces a sexual seed, whereas facultative apomicts such as Kentucky bluegrass produce both sexual and apomictic seeds. Apomixis permits the increase and use of superior genotypes with the added advantage of seed propagation. Some perennial bunch grasses, such as elephant grass and sugarcane, are important enough to warrant the added cost of vegetative propagation from stem cuttings. A few superior genotypes of grasses that spread rapidly by stolons or rhizomes are propagated vegetatively. Usually these grasses are sterile or produce too few

seeds to permit seed propagation. They are also heterozygous F_1 hybrids that would lose most of their superiority if they could be propagated by seed. Examples are Pangola grass and Coastal and Midland Bermuda grasses that have been planted on several million acres. The sterile interspecific Bermuda grass hybrids developed for turf and widely grown across the southern United States are planted vegetatively. Machines designed to harvest and plant vegetative parts of these grasses greatly facilitate vegetative planting. *See* PLANT PROPAGATION; REPRODUCTION (PLANT).

Nutritive value. Cereal grasses are excellent sources of carbohydrates, but tend to be low in protein and deficient in lysine when used as a sole source of food for humans and animals. Protein and lysine content in cereal grain has been improved by breeding for these traits. The quality of the young foliage of forage grasses is usually adequate for ruminants. As the leaves and stems age, digestibility and protein content decreases, indigestible lignin increases, and the performance of animals consuming them declines. The protein and mineral content of grasses growing on most soils can be increased by fertilizing the soil with nitrogen and minerals. The nutritive value of forages can be improved by breeding. *See* BREEDING (PLANT); GRAIN CROPS; GRASSLAND ECOSYSTEM. Glenn W. Burton

Diseases. Although grasses are the oldest agricultural crop, it has only been in recent years that the diagnosis and control of their diseases have received major attention. Of the 95 diseases known to occur on the various cultivated grass species, not all affect the same grass, nor do they all occur in the same region of the world. Kentucky bluegrass (*Poa pratensis*), for example, is susceptible to 52 known diseases, while tall fescue (*Festuca arundinacea*) is subject to only 6. With each of the cultivated grass species, however, there is at least one known disease for each season of the year.

Pathogens. Fungi are the most common of the biotic agents that cause diseases of grasses. The types of pathogenic fungi range from those that infect and remain primarily inside the plant tissue and are visible only with the aid of a microscope, to those that can be seen with the naked eye as they grow freely over the surface of the leaves, to the highly conspicuous mushrooms and toadstools that cause fairy rings. Fungi that parasitize the leaves of grasses cause the most conspicuous symptoms. Depending on the disease, these symptoms vary from distinct leaf spots to indiscriminate blotches to complete blighting of the entire leaf (**Figs. 2** and **3**). Many of these fungi can also cause destructive rots of the root systems of the same plants. Several of the fungal diseases are very severe, and can cause significant reductions in yields and nutritive value of harvested forage, the longevity of the productive life of the stand of forage or pasture grass, or the landscape or recreational value of turfgrass.

Next in importance as grass pathogens are nematodes. While some species attack the floral tissue of grass plants, the most important parasitic nematodes are those that feed on the roots. When nematode populations are high, their feeding activity destroys large amounts of root tissue. As a result, the plants recover more slowly from mowing or grazing. Also, the overall yield of forage is reduced, and plant resistance to various environmental and use stresses is lowered.

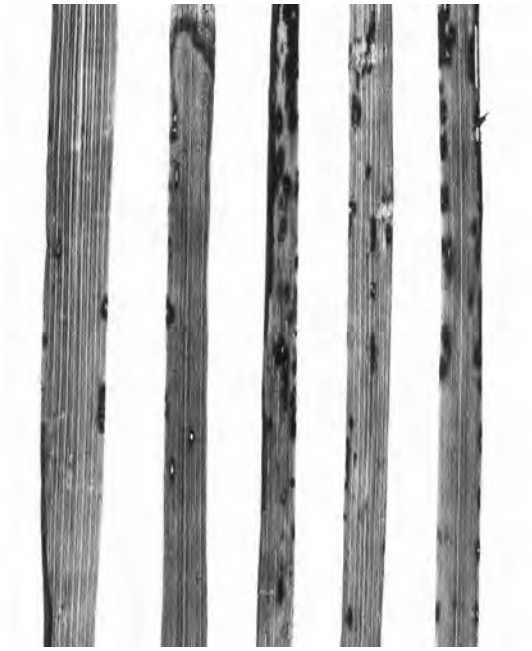

Fig. 2. Cerospora leaf spot of tall fescue, an example of a leaf spot-producing disease. (*From H. B. Couch, Diseases of Turfgrasses, 2d ed., Robert Krieger Publishing Co., 1973*)

Bacteria, mycoplasmas, spiroplasmas, and viruses also cause diseases of grasses. In some instances these diseases are of major importance, but collectively the total impact on forage yield or turfgrass quality is not as great as those of fungi and nematodes.

Control measures. Control of grass diseases is achieved by the use of disease-resistant varieties

Fig. 3. Fusarium patch of Kentucky bluegrass, a winter disease of grasses and an example of a leaf-blighting disease. (*From H. B. Couch, Diseases of Turfgrasses, 2d ed., Robert Krieger Publishing Co., 1973*)

(especially with grasses grown for pasture and harvested forage) and by the application of pesticides. Many grass varieties have both multiple resistance to several pathogens and high growth potential. In general, it is easier to develop varieties resistant to foliar diseases than to obtain resistance to diseases of the root system. Diseases that are caused by fungi with a high capacity for hybridization, for example, the rusts and powdery mildews, are also difficult to control by the development of resistant varieties. In these instances it is not uncommon for new races of the pathogen to appear that are highly pathogenic to a variety that was formerly resistant to the disease. Application of fungicides and nematicides for the control of grass diseases is generally considered practical only in the cultivation of landscape and recreation turf. Commercially available pre- and postplant nematicides are very effective in reducing the populations of soil-inhabiting nematodes that parasitize grass roots. Also, there are several commercial fungicides that provide high levels of control of parasitic fungi on the leaves of grasses. Both nematicides and fungicides are formulated as granules for application with a spreader, and as powders and liquids for use with sprayers. *See* FUNGISTAT AND FUNGICIDE; PESTICIDE; PLANT PATHOLOGY. H. B. Couch

Bibliography. H. B. Couch, *Diseases of Turfgrasses*, 3d ed., 1995; A. S. Hitchcock, *Manual of Grasses of the United States*, 2d ed., 1983; M. B. Jones and A. Lazenby, *The Grass Crop: The Physiological Basis of Production*, 1988; H. B. Sprague (ed.), *Grasslands of the United States: Their Economic and Ecological Significance*, 1974.

Grassland ecosystem

A Biological community that contains few trees or shrubs, is characterized by mixed herbaceous (nonwoody) vegetation cover, and is dominated by grasses or grasslike plants. Grassland ecosystems range from the dense bamboo of the Amazonian tropics to the cool northern steppes of Russia, and from dry plains in the western United States to Canadian arctic grasslands. Mixtures of trees and grasslands occur as savannas at transition zones with forests, as in the east-central United States, or where rainfall is marginal for trees, such as in south-central Africa and Australia. About 1.2×10^8 mi^2 (4.6×10^7 km^2) of the Earth's surface is covered with grasslands, which make up about 32% of the plant cover of the world. The proportion of original grasslands varies widely among continents with about 44% remaining in Europe and less than 10% in Australia, although the latter has vast expanses of savannas. In North America, grasslands include the Great Plains, which extend from southern Texas into Canada. The European meadows cross the subcontinent, and the Eurasian steppe ranges from Hungary eastward through Russia to Mongolia; the pampas cover much of the interior of Argentina and Uruguay. Vast and varied savannas and velds can be found in central and southern Africa and throughout much of Australia. *See* SAVANNA.

Most civilizations have developed in grassland and savanna regions, and were it not for the abundance and widespread distribution of grasses, the human population of the world would not have attained its present level. Significant portions of the world's grasslands have been modified by grazing or tillage or have been converted to other uses. The most fertile and productive soils in the world have developed under grassland, and in many cases the natural species have been replaced by cultivated grasses (cereals). *See* CEREAL; GRASS CROPS.

Grasslands occur in regions that are too dry for forests but that have sufficient soil water to support a closed herbaceous plant canopy that is lacking in deserts. Thus, temperate grasslands usually develop in areas with 10–40 in. (25–100 cm) of annual precipitation, although tropical grasslands may receive up to 60 in. (150 cm). Grasslands are found primarily on plains or rolling topography in the interiors of great land masses, and from sea level to elevations of nearly 16,400 ft (5000 m) in the Andes. Because of their continental location they experience large differences in seasonal climate and wide ranges in diurnal conditions. In general, there is at least one dry season during the year, and drought conditions occur periodically. *See* DROUGHT; PLANT-WATER RELATIONS.

Classification. Different kinds of grasslands develop within continents, and their classification is based on similarity of dominant vegetation, presence or absence of specific dominant species, or prevailing climate conditions (see **illus.**). In North America, the tallgrass prairie lies between the eastern deciduous forest and the Central Plains. Annual rainfall is 30–40 in. (75–100 cm), and under optimum conditions the dominant bluestem (*Andropogon*) grass species may exceed a height of 6 ft (2 m). The mixed-grass prairie contains species of the tallgrass prairie and the shortgrass prairie, which dominates farther west. Annual precipitation amounts range from 20 to 30 in. (50 to 75 cm), and the height of the dominant species has a similar range. Shortgrass prairive precipitation amounts of 10–20 in. (25–50 cm) and are dominated by grasses adapted to dry conditions, such as grama grasses (*Bouteloua*) and buffalo grass (*Buchloe dactyloides*). Desert grasslands have annual rainfalls of 10–18 in. (25–45 cm), which usually fall in summer (July–August) and winter (December–January). Between those wet periods, the desert grasslands are subjected to extreme drought.

Other North American grasslands include the annual grassland in the central valley of California. Originally, these grasslands were dominated by perennial species, but with grazing pressure most of them have been replaced by native and exotic annual species that are more resistant to grazing. Mountain grasslands occupy higher elevations and are frequently mixed in a parklike appearance with trees up to elevations of 8200 ft (2500 m). Shrub steppe grassland vegetation can be found in the northwestern

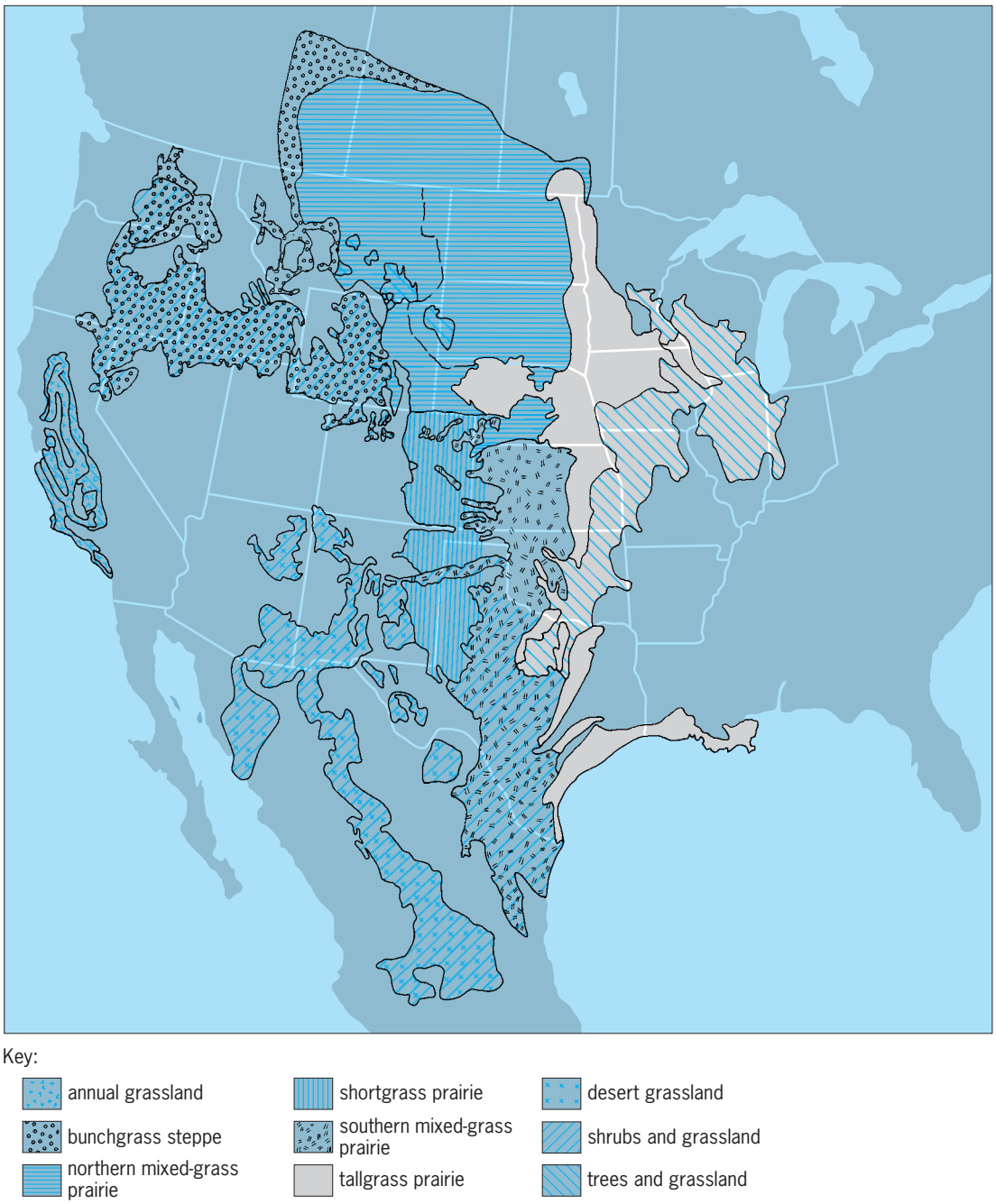

Map of grassland types in central and western North America.

United States and includes many herbaceous grassland species in combination with woody species such as sagebrush (*Artemisia*).

Climate. The climate of grasslands is one of daily and seasonal extremes. Deep winter cold does not preclude grasslands since they occur in some of the coldest regions of the world. However, the success of grasslands in the Mediterranean climate shows that marked summer drought is not prohibitive either. In North America, the rainfall gradient decreases from an annual precipitation of about 40 in. (100 cm) along the eastern border of the tallgrass prairie at the deciduous forest to only about 8 in. (20 cm) in the shortgrass prairies at the foothills of the Rocky Mountains. A similar pattern exists in Europe, with the taller grasses along the northwestern coast and decreasing plant height and rainfall toward the south and east into the plains of Hungary. Growing-season length is determined by temperature in the north latitudes and by available soil moisture in many regions, especially those adjacent to deserts. Plants are frequently subjected to hot and dry weather conditions, which are often exacerbated by windy conditions that increase transpirational water loss from the plant leaves.

Soils. Soils of mesic temperate grasslands are usually deep, about 3 ft (1 m), are neutral to basic, have high amounts of organic matter, contain large

amounts of exchangeable bases, and are highly fertile, with well-developed profiles. The soils are rich because rainfall is inadequate for excessive leaching of minerals and because plant roots produce large amounts of organic material. Humus, partially decomposed organic material that may constitute up to 10% of the soil, expands its capability to retain water by as much as 20% and binds soil particles into clumps, increasing the effectiveness of the soil to make nutrients and water available to the plants. With less rainfall, grassland soils are shallow, contain less organic matter, frequently are lighter colored, and may be more basic. Tropical and subtropical soils are highly leached, have lower amounts of organic material because of rapid decomposition and more leaching from higher rainfall, and are frequently red to yellow. Almost all true grasslands soils are classified as Mollisols and Aridisols, although a few from drier climates are Alfisols and some coastal grasslands are Vertisols. *See* HUMUS.

Grassland soils are dry throughout the profile for a portion of the year. If the soil profile is relatively shallow over a rock subsurface or an impervious layer, the total water stored may be too small to support trees even in geographical areas of relatively high rainfall, which accounts for the presence of grasslands in regions whose climate would support forests. Sandy soils reduce loss of water from runoff and increase water storage in the soil profile because of increased percolation. As a result, sandy soils support more vigorous grassland vegetation, or permit tree growth in a relatively arid region, when compared with soils containing higher amounts of silt and clay particles. Because of their dense fibrous root system in the upper layers of the soil, grasses are better adapted than trees to make use of light rainfall showers during the growing season. When compared with forest soils, grassland soils are generally subjected to higher temperatures, greater evaporation, periodic drought, and more transpiration per unit of total plant biomass. *See* BIOMASS; SOIL.

Vegetation. Worldwide, there are approximately 600 genera and 7500 species of grasses. Temperate North American dominant grass genera include the bluestems (*Andropogon*, *Schizachyrium*), grama grasses (*Bouteloua*), switchgrasses (*Panicum*), wheatgrasses (*Agrypyron*), wire grasses (*Artistida*), fescues (*Festuca*), and bluegrasses (*Poa*). However, many other herbaceous grassland species are broad-leaved forbs, which are plants other than grasses, and like most of the grasses, are perennial. Throughout the year, flowering plants bloom in the grasslands with moderate precipitation, and flowers bloom after rainfall in the drier grasslands. The number of plant species in any one grassland is relatively small, usually 50 to 350 species. Dominant species that flower late in the growing season tend to be taller than the early, cool-season species. In general, the dominant species persists from year to year with relatively small changes in biomass production or changes in importance. The subdominant species demonstrate much greater fluctuations from year to year, depending upon the weather conditions and especially in relation to periods of seed germination and growth.

Just as the above-ground canopy is layered, so are the roots below ground. In mesic grasslands, the roots of perennial grasses have rooting depths greater than 5 ft (1.5 m); many forbs have even deeper roots. Usually, however, 75% or more of the root biomass occurs in the top 10 in. (25 cm) of the soil profile.

With increasing aridity and temperature, grasslands tend to become less diverse in the number of species; they support more warm-season species (which mature late in the growing season); the complexity of the vegetation decreases; the total above-ground and below-ground production decreases; but the ratio of above-ground to below-ground biomass becomes smaller.

Many grassland plant species demonstrate adaptations to minimize damage from grazing. Unlike forbs and woody species, the growing tissues in grasses are at the base of the leaves near the soil surface, so when the leaves are grazed, the meristematic region is protected and can still produce new leaves. Some forbs also have antigrazing mechanisms, such as spines or tough structural material, chemicals that discourage grazing because of taste, supine growth forms so that the leaves stay near the soil surface, or growth flushes of plentiful forage under advantageous environmental conditions so that some of the plants are likely to be ungrazed.

Animal life. There are many more invertebrate species than any other taxonomic group in the grassland ecosystem. For example, more than 1600 different species of insects have been identified from a shortgrass prairie in Colorado. In North America, each prairie state has at least 100 species of grasshopper; Kansas has nearly 300 species. Invertebrate species occur both above and below the soil surface, and depending upon their life cycle, some are found in both locations. Invertebrates play several roles in the ecosystem. For example, many are herbivorous, such as grasshoppers, and eat leaves and stems, whereas others, like cutworms, feed on the roots of plants. Earthworms process organic matter into small fragments that decompose rapidly, scarab beetles process animal dung on the soil surface, flies feed on plants and are pests to cattle, and many species of invertebrates are predaceous and feed on other invertebrates. Soil nematodes, small nonarthropod invertebrates, include forms that are herbivorous, predaceous, or saprophagous, feeding on decaying organic matter. In a South Dakota mixed-grass prairie, $2–6 \times 10^6$ herbivorous nematode forms were found within a soil sample 3 ft (1 m) square and 8 in. (20 cm) deep. *See* SOIL ECOLOGY.

Most of the reptiles and amphibians in grassland ecosystems are predators. For example, lizards and box turtles prey on insects, and snakes prey on rodents and small invertebrates.

Relatively few bird species inhabit the grassland ecosystem, although many more species are found in the flooding pampas of Argentina than in the dry

grasslands of the western United States. In North American prairies, perhaps only about a dozen bird species are restricted to the grasslands, and another two dozen are characteristic of but not limited to the grasslands. The small number of species is related to the uniform habitat. Typical groups of birds include hawks, grouse, meadowlarks, longspurs, and sparrows. Their role in the grassland ecosystem involves consumption of seeds, invertebrates, and vertebrates; seed dispersal; and scavenging of dead animals.

Small mammals of the North American grassland include moles, shrews, gophers, ground squirrels, and various species of mice. Among intermediate-size animals are the opossum, fox, coyote, badger, skunk, rabbit, and prairie dog; large animals include the mule deer (*Odocoileus hemionus*), white-tailed deer (*O. virginianus*), pronghorn (*Antilocapra americana*), and elk (*Cervus canadensis*). The most characteristic large mammal species of the North American grassland is the bison (*Bison bison*), but the 60–75 $\times$ 10^6 of these animals were largely eliminated in the late 1800s and are now mostly confined to reserves. Mammals include both ruminant (pronghorns) and nonruminant (prairie dogs) herbivores, omnivores (opossum), and predators (wolves).

Except for large mammals and birds, the animals found in the grassland ecosystem undergo relatively large population variations from year to year. These variations, some of which are cyclical and others more episodic, are not entirely understood and may extend over several years. Many depend upon predator–prey relationships, parasite or disease dynamics, or weather conditions that influence the organisms themselves or the availability of food, water, and shelter. *See* POPULATION ECOLOGY.

Microorganisms. Within the grassland ecosystem are enormous numbers of very small organisms, including bacteria, fungi, algae, and viruses. One study found 9 $\times$ 10^6 bacteria in a gram of tallgrass prairie soil; another study found about 6560 ft (2000 m) of fungal hyphae (strands of fungus) in 0.04 oz (1 g) of short-grass prairie soil. From a systems perspective, the hundreds of species of bacteria and fungi are particularly important because they decompose organic material, releasing carbon dioxide and other gases into the atmosphere and making nutrients available for recycling. Bacteria and some algae also capture nitrogen from the atmosphere and fix it into forms available to plants. *See* NITROGEN FIXATION; SOIL MICROBIOLOGY.

System functions. The grassland ecosystem consists of several components. Producers are plants that use the Sun's energy to capture carbon as carbon dioxide from the atmosphere and, with available nutrients from the soils, produce more plant material. Consumers consist of animals and microorganisms that feed upon plant parts (herbivores) and other animals (carnivores). Decomposer microorganisms and invertebrates convert dead organic matter to released carbon dioxide and available nutrients in the soil. The carbon cycle involves the transfer of carbon from the atmosphere into plants, through various animals and microorganisms, and back into the atmosphere. Energy is captured first from the Sun and then cycled through the system as organic material until this organic material is eventually decomposed to carbon dioxide. The nutrients are then released again to the soil. Only about 1% of the total solar radiation is captured by the vegetation. *See* BIOGEOCHEMISTRY.

The amount of vegetation produced depends on the type of grassland and the level of water and nutrients in the soil. In the tallgrass prairie, annual aboveground plant production is 1.6–2.0 oz/ft^2 (500–600 g/m^2), about 1.1 oz/ft^2 (350 g/m^2) in the mixed-grass prairie, 0.5–0.6 oz/ft^2 (150–200 g/m^2) in the shortgrass prairie, and as much as 13.1 oz/ft^2 (4000 g/m^2) in tropical grasslands. Total annual primary production is two or more times those amounts, because much of the energy is translocated below ground. *See* BIOLOGICAL PRODUCTIVITY; ECOLOGICAL ENERGETICS.

Grasslands have been used for grazing for the last 1.5–2.0 $\times$ 10^7 years, probably beginning in the Miocene. For all practical purposes, it can be assumed that grasslands have evolved under the influence of herbivores. In North America, the bison roamed across the landscape, heavily grazing the grassland they traversed. As the bison moved on, however, the rangelands recovered. Where the grassland is grazed heavily, as in prairie dog towns, around waterholes, and in fenced-in livestock pastures, grazing exerts enormous pressure on the vegetation. Because many of the native species are palatable, they are selectively grazed, and weedy, less palatable species remain. Moreover, those influences may extend to other components of the grassland ecosystem and increase the vulnerability of the soil to erosion or the susceptibility of the rangeland to insect herbivory or diseases. *See* ECOLOGICAL SUCCESSION.

Much of the grassland ecosystem has been burned naturally, probably from fires sparked by lightning. Human inhabitants have also routinely started fires intentionally to remove predators and undesirable insects, to improve the condition of the rangeland, and to reduce cover for predators and enemies; or unintentionally. Thus, grasslands have evolved under the influences of grazing and periodic burning, and the species have adapted to withstand these conditions. If burning or grazing is coupled with drought, however, the grassland will sustain damage that may require long periods of time for recovery by successional processes. Because of the prehistoric and historic burning of prairies, some rangeland management strategies have included periodic fires to remove woody species and old or dead herbaceous vegetation. *See* AGROECOSYSTEM; ECOSYSTEM.

Paul G. Risser

Bibliography. G. P. Chapman, *Desertified Grasslands: Their Biology and Management*, 1992; A. Langley, *Grasslands*, 1993; S. P. Long, M. B. Jones, and M. J. Roberts (eds.), *Primary Productivity of Grass Ecosystems*, 1991.

Gravel

An unconsolidated sedimentary aggregate containing more than 50% by weight of gravel-sized particles (mean diameter greater than 0.08 in. or 2 mm). The gravel-sized particles are termed the framework; those less than 0.08 in. (2 mm) in diameter are the matrix. There is an important distinction between framework-supported and matrix-supported gravels. The latter may possess a muddy matrix, in which case they are termed diamictons. Typically, diamictons are unstratified internally, contain subangular framework clasts, and are deposited by mass-flow processes such as debris flow or glacial-ice transport. Water-laid gravels are typically stratified or cross-stratified and are framework-supported, with subangular to rounded clasts in a sand matrix. Less commonly, they may be sand-matrix–supported, or they may lack matrix and then are termed openwork gravels. Water-worn gravel clasts tend to conform to the shape of triaxial ellipsoids and develop preferred orientation; with long axes normal to stream flow and intermediate axes dipping gently upstream.

The consolidated equivalents of gravels are conglomerates and breccias, the latter including only angular particles. Paleoenvironmental indicators for conglomerates include stratification, size grading, particle roundness, particle orientation, and matrix: framework relations. *See* BRECCIA; CONGLOMERATE; SEDIMENTARY ROCKS. Brian R. Rust

Bibliography. W. C. Krumbein and F. J. Pettijohn, *Manual of Sedimentary Petrography*, 1988; F. J. Pettijohn and P. E. Potter, *Sand and Sandstone,* 2d ed., 1987.

Gravimetric analysis

That branch of quantitative chemical analysis in which a desired constituent is converted (usually by precipitation) to a pure compound or element of definite, known composition, and is weighed. In a few cases, a compound or element is formed which does not contain the constituent but bears a definite mathematical relationship to it. In either case, the amount of desired constituent can be determined from the weight and composition of the precipitate. The following are the essential steps in a conventional gravimetric analysis.

Dissolution. A sample is weighed and dissolved in a suitable solvent. Water and dilute mineral acids dissolve most inorganic substances, but occasionally concentrated acids or the specific effect of hydrofluoric acid are required. Some refractories require fusions to convert them to acid-soluble products. *See* BALANCE; SOLUBILIZING OF SAMPLES.

Precipitation. After removal of any interfering substances, the desired constituent is precipitated by the addition of the appropriate reactant to the properly prepared solution. Conditions are regulated so that coprecipitation of foreign substances is minimized. In most cases, favorable conditions are attained by precipitating the constituent from a well-stirred, highly diluted, hot solution by the drop-by-drop addition of a dilute reagent in only slight excess over the amount theoretically required. In order to obtain a more nearly pure product, it is usually desirable to dissolve the precipitate in a suitable solvent and to form it a second time. *See* PRECIPITATION (CHEMISTRY).

Digestion. In order to be readily filterable, a suspension of a precipitate is ordinarily allowed to stand at high temperature for the period of time necessary to permit amorphous particles to clot, or crystalline particles to increase in size. This digestion is usually carried out on an electric hot plate so regulated as to maintain a temperature just below the boiling point of the liquid. The addition of paper pulp is helpful in aiding subsequent filtration, but may be used only in cases where the precipitate is to be ignited at a temperature that causes the pulp to burn off.

Filtration. Filtration is accomplished by pouring the suspension of a precipitate through a suitable filtering medium. Whatever the medium, as much of the supernatant liquid as possible is decanted through the filter, and transference of the precipitate is delayed as long as possible. Common filters are as follows.

Filter paper for gravimetric use is specially prepared and has undergone a treatment with hydrofluoric acid and with other acids so that on ignition it gives an ash of known, and usually negligible, weight. Papers of different degrees of porosity are available, and in a given filtration that grade is chosen which gives as rapid a filtration as possible and yet retains the precipitate completely. In general, gelatinous precipitates require a coarse-mesh paper; fine crystalline precipitates, a fine-mesh paper. Suction is almost never used in paper filtration.

A Gooch crucible is a perforated-base porcelain crucible, and the filtering medium is a pad of asbestos produced by pouring into the crucible a suspension of asbestos fibers which are matted by applying suction. A perforated plate on top of the pad holds it in place. Fiber-glass disks may be used instead of asbestos. Suction is applied during filtration; the crucible is dried to constant weight and is weighed before and after the filtration. Gooch crucibles are usually used for precipitates that attain a definite composition at the moderate temperature of a drying oven.

A Munroe crucible differs from a Gooch crucible in that it is made of platinum and uses a platinum sponge (produced by igniting alcohol-moistened ammonium chloroplatinate) as the filtering medium. It retains even fine precipitates and can be heated to a high temperature. It is too expensive for routine use.

An Alundum crucible is made from aluminum oxide and is porous throughout. It therefore needs no additional filtering medium and can be heated to a high temperature.

A Selas crucible has glazed porcelain sides and an unglazed porcelain base which serves as the filtering medium.

A glass filtering crucible has a sintered glass base for the filtering medium and is favored in cases where the precipitate needs to be heated only to the moderate temperatures of a drying oven.

Washing. Precipitates are washed (usually with hot water) until essentially free from soluble foreign matter. Occasionally an aqueous solution of an electrolyte is used to prevent peptization of the precipitate, with resulting conversion to a colloidal solution. Ammonium nitrate is favored for this purpose since it is removed by volatilization when the precipitate is subsequently ignited. Washing is more efficient if the wash water is added in several portions with intermediate drainage.

Drying and ignition. Some substances can be dried to constant weight by heating to relatively low temperatures (110–275°C or 230–530°F) in drying ovens. High-temperature ignition is usually carried out on a precipitate that has been filtered on paper. First the paper is smoked off at a low temperature, and the residue is ignited either in an electric muffle furnace or over a free flame. In the latter case, one of the following types of burners is used: a bunsen burner, a simple tube with an opening at the base to permit air to be mixed with the illuminating gas used; a tirrill burner, a modification of the bunsen burner which allows greater flexibility in the adjustment of the air-gas mixture; a meker burner, which is larger in diameter than the tirrill burner and has a grid at the top to give broad, hot flame; a blast lamp, which supplies air (or oxygen) under pressure to the illuminating gas used. The temperature delivered to the contents of a platinum crucible by a bunsen or tirrill burner is about 850°C (1600°F); that by a grid-top burner is about 1000°C (1800°F); that by a gas-air blast lamp about 1150°C (2100°F).

Cooling. A dried or ignited precipitate is cooled in a desiccator, which is a jarlike receptacle containing a desiccant, and which, in analytical chemistry, is used principally to allow a heated crucible and its contents to come to room temperature without taking on moisture from the air. Anhydrous calcium chloride, although not a very efficient desiccant, is commonly used for this purpose because of its low cost. Other desiccants in the order of increasing effectiveness are anhydrous barium perchlorate, sodium hydroxide sticks, silica gel, aluminum oxide, anhydrous calcium sulfate, calcium oxide, anhydrous magnesium perchlorate, barium oxide, and phosphorus pentoxide.

Calculations. At least two weighings are required for each analysis—the original sample, and the dried or ignited residue. From these weights, the percentage or proportion of the desired constitutent may be calculated from the equation below.

$$\%A = \frac{\text{wt of residue} \times \text{factor} \times 100}{\text{wt of sample}}$$

The factor is determined from a knowledge of the chemical relationships between the weight of substance A contained in, or equivalent to, a fixed weight of residue of known composition. *See* ANALYTICAL CHEMISTRY; QUANTITATIVE CHEMICAL ANALYSIS; STOICHIOMETRY.

Stephen G. Simpson

Bibliography. G. D. Christian, *Analytical Chemistry,* 5th ed., 1993; A. I. Vogel et al., *Vogel's Textbook of Quantitative Chemical Analysis*, 5th ed., 1989.

Gravitation

The mutual attraction between all masses and particles of matter in the universe. In a sense this is one of the best-known physical phenomena. During the eighteenth and nineteenth centuries gravitational astronomy, based on Newton's laws, attracted many of the leading mathematicians and was brought to such a pitch that it seemed that only extra numerical refinements would be needed in order to account in detail for the motions of all celestial bodies. In the twentieth century, however, Albert Einstein with his general theory of relativity and the concurrent development of quantum mechanics shattered this complacency. The subject is currently in a healthy state of flux.

Until the seventeenth century, the sole recognized evidence of this phenomenon was the gravitational attraction at the surface of the Earth. Only vague speculation existed that some force emanating from the Sun kept the planets in their orbits. Such a view was expressed by Johannes Kepler, the author of the laws of planetary motion, to justify his empirical laws. But a proper formulation for such a force had to wait until Isaac Newton founded Newtonian mechanics, with his three laws of motion, and discovered, in calculus, the necessary mathematical tool. *See* CELESTIAL MECHANICS; NEWTON'S LAWS OF MOTION; PLANET.

Newton's law of gravitation. Newton's law of universal gravitation states that every two particles of matter in the universe attract each other with a force that acts in the line joining them, the intensity of which varies as the product of their masses and inversely as the square of the distance between them. Or, the gravitational force F exerted between two particles with masses m_1 and m_2 separated by a distance d is given by Eq. (1), where G is called the constant of gravitation.

$$F = \frac{Gm_1m_2}{d^2} \tag{1}$$

A force varying with the inverse-square power of the distance from the Sun had been already suggested—notably by R. Hooke but also by other contemporaries of Newton, such as E. Halley and C. Wren—but this had been applied only to circular planetary motion. The credit for accounting for, and partially correcting, Kepler's laws and for setting gravitational astronomy on a proper mathematical basis is wholly Newton's.

Newton's theory was first published in the *Principia* in 1686. According to Newton, it was formulated in principle in 1666 when the problem of elliptic motion in the inverse-square force field was solved. But publication was delayed in part because

of the difficulty of proceeding from the "particles" of the law to extended bodies such as the Earth. This difficulty was overcome when Newton established that, under his law, bodies having spherically symmetrical distribution of mass attract each other as if all their mass were concentrated at their respective centers.

Newton verified that the gravitational force between the Earth and the Moon, necessary to maintain the Moon in its orbit, and the gravitational attraction at the surface of the Earth were related by an inverse-square law of force. Let E be the mass of the Earth, assumed to be spherically symmetrical with radius R. Then the force exerted by the Earth on a small mass m near the Earth's surface is given by Eq. (2), and the acceleration of gravity on the Earth's surface, g, by Eq. (3).

$$F = \frac{GEm}{R^2} \tag{2}$$

$$g = \frac{GE}{R^2} \tag{3}$$

Let a be the mean distance of the Moon from the Earth, M the Moon's mass, and P the Moon's sidereal period of revolution around the Earth. If the motions in the Earth-Moon system are considered to be unaffected by external forces (principally those caused by the Sun's attraction), Kepler's third law applied to this system is given by Eq. (4).

$$\frac{4\pi^2 a^3}{P^2} = G(E + M) \tag{4}$$

Equations (3) and (4), on elimination of G, give Eq. (5).

$$g = 4\pi^2 \frac{E}{E + M} \frac{a^2}{R^2} \frac{a}{P^2} \tag{5}$$

Now the Moon's mean distance from the Earth is $a = 60.27R = 3.84 \times 10^8$ m (2.39×10^5 mi), and the sidereal period of revolution is $P = 27.32$ days $= 2.361 \times 10^6$ s. These data give, with $E/M = 81.35$, $g = 9.77$ m/s^2 (32.1 ft/s^2), which is close to the observed value.

This calculation corresponds in essence to that made by Newton in 1666. At that time the ratio a/R was known to be about 60, but the Moon's distance in miles was not well known because the Earth's radius R was erroneously taken to correspond to 60 mi (97 km) per degree of latitude instead of 69 mi (111 km). As a consequence, the first test was unsatisfactory. But the discordance was removed in 1671 when the measurement of an arc of meridian in France provided a reliable value for the Earth's radius.

Gravitational constant. Equation (3) shows that the measurement of the acceleration due to gravity at the surface of the Earth is equivalent to finding the product G and the mass of the Earth. Determining the gravitational constant by a suitable experiment is therefore equivalent to "weighing the Earth."

In 1774, G was determined by measuring the deflection of the vertical by the attraction of a mountain. This method is much inferior to the laboratory method in which the gravitational force between known masses is measured. In the torsion balance two small spheres, each of mass m, are connected by a light rod, suspended in the middle by a thin wire. The deflection caused by bringing two large spheres each of mass M near the small ones on opposite sides of the rod is measured, and the force is evaluated by observing the period of oscillation of the rod under the influence of the torsion of the wire (see **illus.**). This is known as the Cavendish experiment, in honor of H. Cavendish, who achieved the first reliable results by this method in 1797–1798. More recent determinations using various refinements yield the results: constant of gravitation $G = 6.67 \times 10^{-11}$ SI (mks) units; mass of Earth $= 5.98 \times 10^{24}$ kg. The result of the best available laboratory measurements, announced in 2000, is $G = (6.674215 \pm 0.000092) \times 10^{-11}$ in SI (mks) units. As a result, the mass of the Earth is $(5.972254 \pm 0.000082) \times 10^{24}$ kg and the Sun's mass is $(1.988435 \pm 0.000027) \times 10^{30}$ kg.

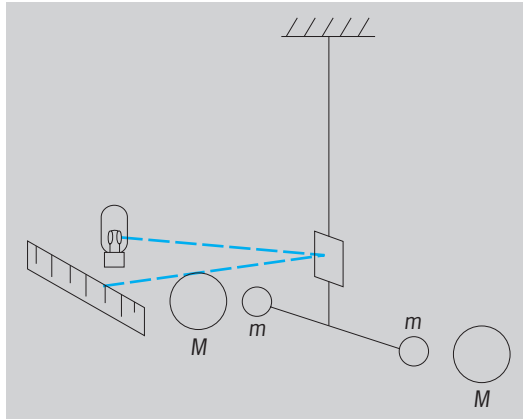

Diagram of the torsion balance.

In a more basic sense, G is a conversion factor between the units used for mass and distance and the units used for force. This definition must give the same force unit as the electromagnetic methods. These other methods are much more accurate than our current knowledge of G. Hopefully, improvements in the value of G will give the same size for the unit of force. Thus, knowing G more accurately will check the consistency of our definition of force.

In Newtonian gravitation, G is an absolute constant, independent of time, place, and the chemical composition of the masses involved. Partial confirmation of this was provided before Newton's time by the experiment attributed to Galileo in which different weights released simultaneously from the top of the Tower of Pisa reached the ground at the same time. Newton found further confirmation, experimenting with pendulums made out of different materials. Early in this century, R. Eötvös found that different materials fall with the same acceleration to within 1 part in 10^7. The accuracy of this figure has been extended to 1 part in 10^{11}, using aluminum and gold, and to 0.9×10^{-12} with a confidence of 95%, using aluminum and platinum.

With the discovery of antimatter, there was

speculation that matter and antimatter would exert a mutual gravitational repulsion. But experimental results indicate that they attract one another according to the same laws as apply to matter of the same kind. *See* ANTIMATTER.

A cosmology with changing physical "constants" was first proposed in 1937 by P. A. M. Dirac. Field theories applying this principle have since been proposed by P. Jordan and D. W. Sciama and, in 1961, by C. Brans and R. H. Dicke. In these theories G is diminishing; for instance, Brans and Dicke suggest a change of about 2×10^{-11} per year. This would have profound effects on phenomena ranging from the evolution of the universe to the evolution of the Earth. There is no firm evidence at present to support a time variation of G. For instance, detailed analyses of the motion of the Moon using the lunar laser ranging data, the analysis of solar system data, and especially the ranging data using Viking landers on Mars, and the timing analysis of the binary pulsar PSR B1913 + 16 have all resulted in only upper limits on a possible temporal variation of G. The present upper limit on the relative rate of time variation of G is a few parts in 10^{14} per year. This result is theory-independent; that is, it is based purely on observations without relying on any particular theory in which G is presumed to vary with time. *See* PULSAR.

Mass and weight. In the equations of motion of Newtonian mechanics, the mass of a body appears as inertial mass, a measure of resistance to acceleration, and as gravitational mass in the expression of the gravitational force. The equality of these masses is confirmed by the Eötvös experiment. It justifies the assumption that the motion of a particle in a gravitational field does not depend on its physical composition. In Newton's theory the equality can be said to be a coincidence, but not in Einstein's theory, where this equivalence becomes a cornerstone of relativistic gravitation.

While mass in Newtonian mechanics is an intrinsic property of a body, its weight depends on certain forces acting on it. For example, the weight of a body on the Earth depends on the gravitational attraction of the Earth on the body and also on the centrifugal forces due to the Earth's rotation. The body would have lower weight on the Moon, even though its mass would remain the same. *See* CENTRIFUGAL FORCE.

Gravity. This should not be confused with the term gravitation. Gravity is the older term, meaning the quality of having weight, and so came to be applied to the tendency of downward motion on the Earth. Gravity or the force of gravity is today used to describe the intensity of gravitational forces, usually on the surface of the Earth or another celestial body. So gravitation refers to a universal phenomenon, while gravity refers to its local manifestation.

A rotating planet is oblate (or flattened at the poles) to a degree depending on the ratio of the centrifugal to the gravitational forces on its surface and on the distribution of mass in its interior. The variation of gravity on the surface of the Earth depends on these factors and is further complicated by irregular features such as oceans, continents, and mountains. It is investigated by gravity surveys and also through the analysis of the motion of artificial satellites. Because of the irregularities, no mathematical formula has been found that satisfactorily represents the gravitational field of the Earth, even though formulas involving hundreds of terms are used. The problem of representing the gravitational field of the Moon is even harder because the surface irregularities are proportionately much larger. *See* EARTH, GRAVITY FIELD OF.

In describing gravity on the surface of the Earth, a smoothed-out theoretical model is used, to which are added gravity anomalies, produced in the main by the surface irregularities.

Gravity waves are waves in the oceans or atmosphere of the Earth whose motion is dynamically governed by the Earth's gravitational field. They should not be confused with gravitational waves, which are discussed below.

Gravitational potential energy. This describes the energy that a body has by virtue of its position in a gravitational field. If two particles with masses m_1 and m_2 are a distance r apart and if this distance is slightly increased to $r + \Delta r$, then the work done against the gravitational attraction is $Gm_1m_2\Delta r/r^2$. If the distance is increased by a finite amount, say from r_1 to r_2, the work done is given by Eq. (6). If $r_2 \to \infty$, Eq. (7) holds.

$$W_{r_1,r_2} = Gm_1m_2 \int_{r_1}^{r_2} \frac{dr}{r^2} = Gm_1m_2\left(\frac{1}{r_1} - \frac{1}{r_2}\right) \tag{6}$$

$$W_{r_1,\infty} = \frac{Gm_1m_2}{r_1} \tag{7}$$

If one particle is kept fixed and the other brought to a distance r from a very great distance (infinity), then the work done is given by Eq. (8). This is called

$$-U = \frac{-Gm_1m_2}{r} \tag{8}$$

the gravitational potential energy; it is (arbitrarily) put to zero for infinite separation between the particles. Similarly, for a system of n particles with masses $m_1, m_2, \ldots, m_n$ and mutual distance r_{ij} between m_i and m_j, the gravitational potential energy $-U$ is the work done to assemble the system from infinite separation (or the negative of the work done to bring about an infinite separation), as shown in Eq. (9).

$$-U = -G\sum_{i<j} \frac{m_im_j}{r_{ij}} \tag{9}$$

A closely related quantity is gravitational potential. The gravitational potential of a particle of mass m is given by Eq. (10), where r is distance measured from

$$V = \frac{-Gm}{r} \tag{10}$$

the mass. The gravitational force exerted on another mass M is M times the gradient of V. If the first body

is extended or irregular, the formula for V may be extremely complicated, but the latter relation still applies. *See* POTENTIALS.

A good illustration of gravitational potential energy occurs in the motion of an artificial satellite in a nearly circular orbit around the Earth which is affected by atmospheric drag. Because of the frictional drag the total energy of the satellite in its orbit is reduced, but the satellite actually moves faster. The explanation for this is that it moves closer to the Earth and loses more in gravitational potential energy than it gains in kinetic energy. *See* ENERGY.

Similarly, in its early evolution a star contracts, with the gravitational potential energy being transformed partly into radiation, so that it shines, and partly into kinetic energy of the atoms, so that the star heats up until it is hot enough for thermonuclear reactions to start. *See* STELLAR EVOLUTION.

Another related phenomenon is that of speed of escape. A projectile launched from the surface of the Earth with speed less than the speed of escape will return to the surface of the Earth; but it will not return if its initial speed is greater (atmospheric drag is neglected). For a spherical body with mass M and radius R, the speed of escape from its surface is given by Eq. (11). For the Earth, V_e is 11.2 km/s

$$V_e = \left(\frac{2MG}{R}\right)^{1/2} \tag{11}$$

(7.0 mi/s); for the Moon, it is 2.4 km/s (1.5 mi/s), which explains why the Moon cannot retain an atmosphere such as the Earth's. By analogy, a black hole can be considered a body for which the speed of escape from the surface is greater than the speed of light, so that light cannot escape; however, the analogy is not really exact since Newtonian mechanics is not valid on the scale of the universe. *See* BLACK HOLE; ESCAPE VELOCITY.

Application of Newton's law. In modern times, Eq. (5)—in a modified form with appropriate refinements to allow for the Earth's oblateness and for external forces acting on the Earth-Moon system—has been used to compute the distance to the Moon. The results have been superseded in accuracy only by radar measurements and observations of corner reflectors placed on the lunar surface.

Newton's theory passed a much more stringent test than the one described above when he was able to account for the principal departures from Kepler's laws in the motion of the Moon. Such departures are called perturbations. A notable triumph of the theory occurred when the observed perturbations in the motion of the planet Uranus enabled J. C. Adams in 1845 and U. J. Leverrier in 1846 independently to predict the existence and calculate the position of a hitherto-unobserved planet, later called Neptune. When yet another planet, Pluto, was discovered in 1930, its position and orbit were strikingly similar to predictions based on the method used to discover Neptune. But the discovery of Pluto must be ascribed to the perseverance of the observing astronomers; it is not massive enough to have revealed itself through the perturbations of Uranus and Neptune. *See* PERTURBATION (ASTRONOMY).

F. W. Bessel observed nonuniform proper motions of Sirius and Procyon and inferred that each was gravitationally deflected by an unseen companion. It was only after his death that these bodies were telescopically observed, and they both later proved to be white dwarfs. More recently evidence has been accumulated for the existence of some planetary masses around stars. The discovery of black holes (which will never be directly observed) hinges in part on a visible star showing evidence for having a companion of sufficiently high mass (so that its gravitational collapse can never be arrested) and on observations of the motions of stars near galactic centers. *See* BINARY STAR.

Newton's theory supplies the link between the observed motion of celestial bodies and certain physical properties, such as mass and sometimes shape. Knowledge of stellar masses depends basically on the application of the theory to binary-star systems. Analysis of the motions of artificial satellites placed in orbit around the Earth has revealed refined information about the gravitational field of the Earth and of the Earth's atmosphere. Similarly, satellites placed in orbit around the Moon have yielded information about its gravitational field, and other space vehicles have yielded the best information to date on the masses and gravitational fields of other planets. *See* SATELLITE (SPACECRAFT); SPACE PROBE.

Newtonian gravitation has been applied to the motion of stars in galaxies as well as the motion of galaxies in clusters of galaxies. Careful observations of the motions of stars and gas (using the 21-cm line of neutral hydrogen) in the outer reaches of spiral galaxies have shown that the speed of rotation far from the center of a galaxy does not diminish with distance as would be expected on the basis of the gravitational influence of the rest of the stars as implied by the light distribution in the galaxy, but stays essentially constant. This circumstance leads to a rotation curve (the plot of the rotational speed versus the radius away from the center of a disk-shaped spiral galaxy) that is flat at large distances from the nucleus of the galaxy. The phenomenon associated with flat rotation curves for the outer regions of spiral galaxies has been well established, and is now generally interpreted in terms of the existence of the so-called dark matter, that is, excess matter that is invisible. Of course, the discrepancy with the prediction of Newtonian gravitation could be due to the breakdown of the theory. Newtonian gravitation and its relativistic generalization (general relativity) have been well tested in the solar system, but significant deviations over larger scales cannot be ruled out in principle. Though phenomenological modifications of the theory have been proposed to explain the discrepancy, no alternative theory that is based on fundamental physical principles exists. Therefore, the discrepancy is generally attributed to the existence of dark matter halos in which the visible galaxies are presumed to be embedded. Dark halo models

together with rotation curves are used to calculate the density of dark matter in galaxies. It is estimated that the visible mass of a galaxy is only about 10% of its total mass. Furthermore, the discrepancy between the luminous and dynamical masses is thought to be larger for clusters of galaxies. Even more dark matter is needed if certain models of cosmology, such as the inflationary scenario, are taken into account. The result is that more than 90% of the mass of the universe could be in the form of dark matter. Proposals for the nature of dark matter range from exotic subatomic particles to black holes. The determination of the nature of dark matter constitutes a fundamental problem in astrophysics and cosmology. *See* COSMOLOGY; DARK MATTER; GALAXY, EXTERNAL; INFLATIONARY UNIVERSE COSMOLOGY; MILKY WAY GALAXY; UNIVERSE.

Accuracy of Newtonian gravitation. Newton was the first to doubt the accuracy of his law when he was unable to account fully for the motion of the perigee in the motion of the Moon. In this case he eventually found that the discrepancy was largely removed if the solution of the equations was more accurately developed. Further difficulties to do with the motion of the Moon were noted in the nineteenth century, but these were eventually resolved when it was found that there were appreciable fluctuations in the rate of rotation of the Earth, so that it was the system of timekeeping and not the gravitational theory that was at fault. *See* EARTH ROTATION AND ORBITAL MOTION.

A more serious discrepancy was discovered by Leverrier in the orbit of Mercury. Because of the action of the other planets, the perihelion of Mercury's orbit advances. But allowance for all known gravitational effects still left an observed motion of about 43 seconds of arc per century unaccounted for by Newton's theory. Attempts to account for this by adding an unknown planet or by drag with an interplanetary medium were unsatisfactory, and a very small change was suggested in the exponent of the inverse square of force. This particular discordance was accounted for by Einstein's general theory of relativity in 1916.

Newtonian mechanics considered gravity as a direct action-at-a-distance and in the early days did not properly account for the time delay between the action of gravity experienced by a body and the motion of the source. General relativity, on the other hand, considers gravitational effects as being geometric and time-related. *See* RELATIVITY.

Gravitational lens. From a relativistic point of view, light is expected to be deflected when it passes through a gravitational field. An analogy can be made to the refraction of light passing through a lens. For example, a galaxy situated between an observer and a more distant source will have a focusing effect, and this accounts for some of the observed properties of quasistellar objects. The multiple images of the quasar (Q0957 + 561 A,B) are almost certainly caused by the light from a single body passing through a gravitational lens. While this is the best-studied gravitational lens, many other examples of this phenomenon have been discovered. *See* GRAVITATIONAL LENS; QUASAR.

Testing of gravitational theories. One of the greatest difficulties in investigating gravitational theories is the weakness of the gravitational coupling of matter. For instance, the gravitational interaction between a proton and an electron is weaker by a factor of about 5×10^{-40} than the electrostatic interaction. (If gravitation alone bound the hydrogen atom, then the radius of the first Bohr orbit would be 10^{13} light-years, or about 1000 times the Hubble radius!) Of the four basic interactions (that is, strong, electromagnetic, weak, and gravitational) that are known at present, gravitation is by far the weakest force in nature. [A unification of the electromagnetic and weak interactions into the so-called electroweak interaction has been carried out, and further unification of this interaction with the strong (nuclear) force exists. So far, no sucessful realistic unification with gravitational interactions has been accepted.] *See* ELECTROWEAK INTERACTION; FUNDAMENTAL INTERACTIONS.

Many attempts have been made to discover Yukawa-type forces that would couple to matter with strengths close to that of gravity. These forces are expected to have finite ranges, however, in contrast to the infinite range of universal gravitation. Thus, deviations from Newton's inverse-square law of gravitation could reveal the presence of such interactions. These forces also appear in certain theories that attempt to unify gravity with the other fundamental interactions. Compelling evidence for the existence of any anomalous macroscopic force has not been found.

Results of experiments to search for new spin-dependent interactions place strict upper limits on anomalous spin-dependent couplings. Furthermore, torsion-balance experiments using ordinary matter have been interpreted as indicating the universality of free fall for matter as well as antimatter.

There is no firm evidence at present for any deviation from the Newtonian law of gravitation in the nonrelativistic regime.

Relativistic theories. Before Newton, detailed descriptions were available of the motions of celestial bodies—not just Kepler's laws but also empirical formulas capable of representing with fair accuracy, for their times, the motion of the Moon. Newton replaced description by theory, but in spite of his success and the absence of a reasonable alternative, the theory was heavily criticized, not least with regard to its requirement of "action at a distance" (that is, through a vacuum). Newton himself considered this to be "an absurdity," and he recognized the weaknesses in postulating in his system of mechanics the existence of preferred reference systems (that is, inertial reference systems) and an absolute time. Newton's theory is a superb mathematical one that represents the observed phenomena with remarkable accuracy.

The investigation of electric and magnetic phenomena culminated in the second half of the nineteenth century in the complete formulation of the laws of electromagnetism by J. C. Maxwell. Maxwell

based his theory on M. Faraday's field concept. The electromagnetic field propagates with the speed of light; in fact, Maxwell's theory unified the science of optics with electricity and magnetism. Maxwell's theory of the electromagnetic field was extended and strengthened with the subsequent observations of electromagnetic waves by H. Hertz and the successes of the theory of electrons developed by H. A. Lorentz. *See* ELECTROMAGNETIC RADIATION; MAXWELL'S EQUATIONS.

The theory of relativity grew from attempts to describe electromagnetic phenomena in moving systems. No physical effect can propagate with a speed exceeding that of light in vacuum; therefore, Newton's theory must be the limiting case of a field theory in which the speed of propagation approaches infinity. Einstein's field theory of gravitation (general relativity) is based on the identification of the gravitational field with the curvature of space-time. The geometry of space-time is affected by the presence of matter and radiation. The relationship between mass-energy and the space-time curvature is therefore a relativistic generalization of the Newtonian law of gravitation. The relativistic theory is mathematically far more complicated than Newton's. Instead of the single Newtonian potential described above, Einstein worked with 10 quantities that form a tensor. *See* TENSOR ANALYSIS.

Principle of equivalence. An important step in Einstein's reasoning is his principle of equivalence, saying that a uniformly accelerated reference system imitates completely the behavior of a uniform gravitational field. Imagine, for instance, a scientist in a space capsule infinitely far out in empty space so that the gravitational force on the capsule is negligible. Everything would be weightless; bodies would not fall; and a pendulum clock would not work. But now imagine the capsule to be accelerated by some agency at the uniform rate of 981 cm/s^2 (32.2 ft/s). Everything in the capsule would then behave as if the capsule were stationary on its launching pad on the surface of the Earth and therefore subject to the Earth's gravitational field. But after its original launching, when the capsule is in free flight under the action of gravitational forces exerted by the various bodies in the solar system, its contents will behave as if it were in the complete isolation suggested above. This principle requires that all bodies fall in a gravitational field with precisely the same acceleration, a result that is confirmed by the Eötvös experiment mentioned earlier. Also, if matter and antimatter were to repel one another, it would be a violation of the principle. *See* FREE FALL.

Einstein's theory requires that experiments should have the same results irrespective of the location or time. This has been said to amount to the "strong" principle of equivalence.

Classical tests. The ordinary differential equations of motion of Newtonian gravitation are replaced in general relativity by a nonlinear system of partial differential equations for which general solutions are not known. Apart from a few special cases, knowledge of solutions comes from methods of approximation. For instance, in the solar system, speeds are low so that the quantity v/c (v is the orbital speed and c is the speed of light) will be small (about 10^{-4} for the Earth). The equations and solutions are expanded in powers of this quantity; for instance, the relativistic correction for the motion of the perihelion of Mercury's orbit is adequately found by considering no terms smaller than $(v/c)^2$. This is called the post-newtonian approximation. (Another approach is the weak-field approximation.)

Einstein's theory has appeared to pass three famous tests. First, it accounted for the full motion of the perihelion of the orbit of Mercury. (Mercury is the most suitable planet, because it is the fastest-moving of the major planets and has a high eccentricity, so that its perihelion is relatively easily studied.) Second, the prediction that light passing a massive body would be deflected has been confirmed with an accuracy of about 5%. Third, Einstein's theory predicted that clocks would run more slowly in strong gravitational fields compared to weak ones; interpreting atoms as clocks, spectral lines would be shifted to the red in a gravitational field. This, again, has been confirmed with moderate accuracy. Locations on the Earth's surface can be determined with great accuracy using the artificial satellites that make up the Global Positioning System (GPS). The clocks in that system must make relativistic corrections for both their motions relative to the Earth (special theory) and their gravitational potential relative to the Earth's surface (general theory). *See* SATELLITE NAVIGATION SYSTEMS.

Predictions of the theory have been confirmed in an experiment in which radar waves were bounced off Mercury; the theory predicts a delay of about 2×10^{-4} s in the arrival time of a radar echo when Mercury is on the far side of the Sun and close to the solar limb. Tests, similar in principle, have been conducted using observations of the Mariner space vehicles, the accuracy of confirmation being in the region of 4%. A greater level of accuracy has been achieved, by better than an order of magnitude, using data from transponders on Viking orbiters and landers on Mars. Furthermore, the deflection of microwave radiation passing close to the Sun has been observed using radio interferometry with a baseline of 22 mi (35 km). The amount of bending that has been found is 1.015 ± 0.011 times the amount predicted by general relativity. In another test, the precession of a gyroscope in orbit around the Earth is to be studied for evidence of the so-called geodetic precession as well as the precession due to the gravitomagnetic field of the rotating Earth. The lunar laser-ranging data have been used to measure the de Sitter precession of the Moon's orbit. This is due to the geodetic precession of the Earth-Moon orbital angular momentum in the gravitational field of the Sun. It amounts to an advance in the lunar node and perigee by about 2 seconds of arc per century, a prediction first made by W. de Sitter in 1916 soon after the advent of general relativity. Such small secular effects are suitable for study since they accumulate in time. Other periodic (noncumulative) orbital

effects have until recently been too small to observe. But the current revolution in observational techniques and accuracy has changed the situation; postnewtonian terms are now routinely included in many calculations of the orbits of planets and space vehicles, and comparison with observations will furnish tests of the theory.

Mach's principle. One of the most penetrating critiques of mechanics is due to E. Mach, toward the end of the nineteenth century. Some of his ideas can be traced back to Bishop G. Berkeley early in the eighteenth century. Out of Mach's work there has arisen Mach's principle; this is philosophical in nature and cannot be stated in precise terms. The idea is that the motion of a particle is meaningful only when referred to the rest of the matter in the universe. Geometrical and inertial properties are meaningless for an empty space, and the motion of a particle in such space is devoid of physical significance. Thus the behavior of a test particle should be determined by the total matter distribution in the universe and should not appear as an intrinsic property of an absolute space. Mach's principle suggests that gravitation and inertia are equivalent. This idea strongly influenced the development of general relativity; however, general relativity, having an absolute concept of rotation, does not fully satisfy Mach's principle.

Brans-Dicke theory. This is a classical field theory of gravitation that was developed in 1961 by Brans and Dicke on the basis of an interpretation of Mach's principle. In this theory the gravitational field is described by a tensor and a scalar, the equations of motion being the same as those in general relativity. The addition of a scalar field leads to the appearance of an arbitrary constant, whose value is not known exactly. The Brans-Dicke theory predicts that the relativistic motion of the perihelion of Mercury's orbit is reduced compared with Einstein's value, and also that the light deflection should be less. With regard to the orbit of Mercury, Dicke pointed out that if the Sun were oblate, this might account for some of the motion of the perihelion. In 1967 he announced that measurements showed a solar oblateness of about 5 parts in 100,000 (or a difference in the polar and equatorial radii of about 21 mi or 34 km). His observations and discussion are still subject to some controversy. The difference between the theory and that of general relativity can be parametrized by the number ω, where $\gamma = (1 + \omega)/(2 + \omega)$; for general relativity, $\gamma = 1$. Dicke has proposed $\omega \approx 7.5$; but the results of measurement of deflection of radiation by the Sun indicate a value of ω greater than 23, for which the predictions of the two theories would not be greatly different. Subsequent data from the Viking spacecraft imply that this constant must be greater than about 500, thus rendering the Brans-Dicke theory almost indistinguishable from general relativity.

There are, of course, many other theories not mentioned here.

Supergravity. This is the term applied to highly mathematical theories of gravitation attempting to form a part of a unified field theory in which all types of forces are included. String, loop, and M-brane theories are under investigation. *See* FUNDAMENTAL INTERACTIONS; QUANTUM GRAVITATION; SUPERGRAVITY; SUPERSTRING THEORY.

Gravitational waves. The existence of gravitational waves, or gravitational "radiation," was predicted by Einstein shortly after he formulated his general theory of relativity. They are now a feature of any relativity theory. Gravitational waves are "ripples in the curvature of space-time." In other words, they are propagating gravitational fields, or propagating patterns of strain, traveling at the speed of light. They carry energy and can exert forces on matter in their path, producing, for instance, very small vibrations in elastic bodies. The gravitational wave is produced by change in the distribution of some matter. It is not produced by a rotating sphere, but would result from a rotating body, or pairs of bodies, not having symmetry about their axis of rotation or from pulsars and supernova explosions. In spite of the relatively weak interaction between gravitational radiation and matter, the measurement of this radiation is now technically possible. *See* SUPERNOVA.

There are currently (2006) five experiments attempting to detect gravitational radiation by interferrometric methods: the two LIGO (Laser Interferometer Gravitational-Wave Observatory) detectors located in the United States (near Livingston, Louisiana, and at the Washington Hanford Nuclear Research Area), VIRGO (France and Italy), GEO 600 (Germany), and TAMA (Japan). These instruments can detect a shift of the order of 1 part in 10^{21}. It is possible, if these instruments meet their design standards, that they may be sensitive enough to detect nearby wave sources. LIGO 2, which is under design, should increase sensitivity by a factor of 10. *See* LIGO (LASER INTERFEROMETER GRAVITATIONAL-WAVE OBSERVATORY.

The National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA) are planning a space-based interferometer, the Laser Interferometer Space Antenna (LISA). It will consist of three spacecraft in an equilateral triangle fromation with arms 5×10^6 km (3×10^6 mi) long. It will be located in the Earth's orbital path about 20° behind the Earth. Communication lasers will be located in each spacecraft. This design eliminates Earth noise and avoids the problems of folded light paths. *See* GRAVITATIONAL RADIATION.

J. M. A. Danby; Bahram Mashhoon; John L. Safko

Bibliography. H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; J. B. Hartle, *Gravity: An Introduction to Einstein's General Relativity*, 2003; W. Kopczynski and A. Trautman, *Spacetime and Gravitation*, 1992; C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, 1973; H. C. Ohanian and R. Ruffini, *Gravitation and Spacetime*, 2d ed., 1994; K. S. Thorne, *Black Holes and Time Warps: Einstein's Outrageous Legacy*, 1994; C. M. Will, *Theory and Experiment in Gravitational Physics*, rev. ed., 1993.

Gravitational collapse

The stage in the evolution of a star in which the pressure in the star is insufficient to maintain the star at a stable size. The material in the star or in the core of the star then falls inward under its own gravitational attraction. Depending on the mass, composition, and spin of the star, the collapse may proceed to the formation of a neutron star or black hole, possibly accompanied by a supernova explosion. *See* STELLAR EVOLUTION.

Stability of stars. Stars similar to the Sun maintain a stable size by continually burning nuclear fuel. In most stars, as for the Sun, this burning involves the conversion of hydrogen to helium with a release of energy as the nuclear reactions take place. This energy supplies heat to the interior of the star, which in turn keeps the core of the star hot enough so that nuclear reactions can continue. Heat is also continually transferred from the interior to the exterior of the star, where it is lost, primarily in the form of radiation.

Just as the temperature decreases from the interior to the exterior of the star, so does the pressure. The change in the pressure over an interval of distance is called a pressure gradient. A pressure gradient is required to keep the star in equilibrium, as can be seen by the following argument. First, any chunk of gas in a stable star is in equilibrium, on the average, if the star is stable. Second, there is a net gravitational force acting on any chunk of gas as a result of the attraction to the rest of the matter in the star, and this force, which is normally called the weight of the gas, must be directed toward the center of the star. Finally, the pressure gradient supplies a force to balance the gravitational force, since the pressure on the bottom of the chunk of gas (bottom meaning closer to the center of the star) is larger than the pressure on the top of the chunk. In equilibrium, the variation of pressure in the star is just that for which all chunks of gas are in equilibrium.

Contraction and the onset of collapse. Suppose now that the nuclear burning is turned off, as is the case if the hydrogen in the core is used up and only helium remains. Then, no more heat is supplied to the core of the star, and its temperature drops as remaining heat is transported to the surface of the star and is lost. Likewise, since the pressure in the gas depends on the temperature, the pressure in the core also drops, and more importantly, the pressure gradient decreases throughout the star. The various chunks of gas are no longer in equilibrium, and they move inward. This compression of the gas then causes a rise in temperature and the temporary reestablishment of the required temperature gradient. However, as heat is transported outward, the temperature again drops, the gas is further compressed, and so on. In this stage, the star is undergoing contraction, the heating coming from the gravitational potential energy of the star rather than nuclear reactions.

Eventually, the star condenses, and the temperature rises to a sufficiently high value that helium burning can take place. Once again, the star is in stable equilibrium until the reservoir of fuel in the core of the star is used up, after which contraction takes place until conditions are right for nuclear burning of higher elements. The contraction discussed here refers to what is going on in the core of the star; the atmosphere may actually be expanding in the process.

When the star has a core which is composed of iron, no further nuclear burning can take place, since iron is the most stable element. Continued contraction must then take place. In fact, theoretical calculations indicate that under normal conditions the temperature does not rise high enough for nuclear burning to proceed beyond carbon, at which point the star effectively runs out of fuel, and continued contraction also takes place. If, after ejection of material by the star, the mass of the star is less than about 1.2–1.4 solar masses, the limiting value known as the Chandrasekhar limit, the contraction takes place to a white dwarf. Theoretical studies indicate that normal stars smaller than about 4 solar masses will eject enough matter in the process of evolution to end up as white dwarfs. Observations suggest that stars of up to 8 solar masses can eject sufficient matter to become white dwarfs. A white dwarf is stable without any nuclear burning taking place. The pressure gradient is produced by the same kind of quantum interactions among the electrons as those which make atoms stable. The white dwarf then can only cool off with time, since it has no source of heat. The cooling period is very long, however, up to 1 billion years. As a result, many white dwarfs are observed in the sky, since the radiation continues long after the star, in some sense, has died. *See* WHITE DWARF STAR.

The mass loss by stars in the range of around 4 to 8 solar masses is not well understood. However, theoretical studies indicate that stars of more than about 8 solar masses will not lose enough mass in their evolution to become white dwarfs. What happens then is that the contracting star at some point becomes unstable; that is, the heating as a result of contraction is insufficient to produce the required pressure gradient to support the gas against gravity. This instability occurs first in the core of the star, and the core of the star starts to fall inward on itself. This is the stage of gravitational collapse. During collapse, the released gravitational potential energy goes into kinetic energy of motion of the core rather than into heating the core, since a near-equilibrium situation is required for efficient transfer of gravitational energy into heat energy. In essence, the core is then freely falling inward under its own gravitational attraction.

Collapse to a neutron star. There are several possible end points to the gravitational collapse of a star. White dwarfs have already been mentioned as a possible end point of stellar evolution, but a collapsing core will be too condensed to form a white dwarf, even if its mass is less than the maximum allowed mass. The only other known possibilities are collapse to a neutron star and collapse to a black hole. A neutron star is a highly condensed star in which the predominant constituent of matter is in

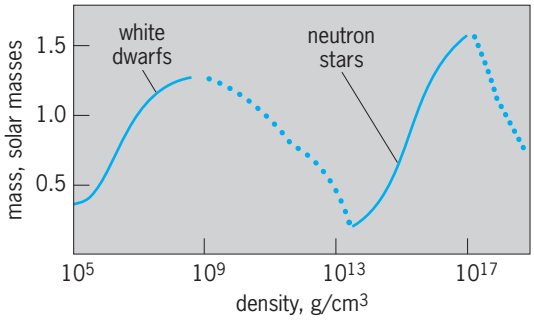

Fig. 1. Masses of white dwarfs and neutron stars as a function of central density. The solid lines indicate possible stable configurations.

the form of neutrons. The stability of such stars is a result of the quantum-mechanical interaction of neutrons, the same kind of interaction which, for electrons, leads to stability of white dwarfs. *See* NEUTRON STAR.

A neutron star is so condensed that the density of matter in the neutron star is comparable with the density of matter in the nucleus of an atom. For example, a neutron star of 1 solar mass has a radius of about 6 mi (10 km). As with white dwarfs, there is a maximum mass for stable neutron stars, in the range of 1.4 to 3 solar masses, the actual value being somewhat uncertain because nuclear forces are important at such a high density. The upper limit of 3 solar masses is fairly certain, however, since beyond that limit gravity will make the neutron star unstable, independent of the nature of the nuclear forces. **Figure 1** shows how the mass of white dwarfs and neutron stars depends on the central density of the star. The actual curve for white dwarfs depends on the composition of the star, and the actual curve for neutron stars is uncertain because of lack of knowledge of nuclear forces at high density.

Theoretical studies of collapsing stars indicate that the star develops a collapsing iron core as result of nuclear reactions induced in the compressed matter. When the core is sufficiently compressed, protons in the nuclei of the ions in the core are converted to neutrons, with the emission of neutrinos in the process. This takes place because electrons, due to quantum-mechanical interactions, are being forced into higher and higher energy states as the core of the star gets smaller and smaller. Protons can then absorb electrons to become neutrons. This is just the inverse of the process by which a free neutron decays to a proton plus an electron (and an antineutrino), a process which is prohibited in the interior of a neutron star because the electron does not have sufficient energy to be emitted. *See* RADIOACTIVITY.

Once the core reaches nuclear densities, repulsive nuclear forces become important and, if the mass of the core is small enough, the collapse is stopped and the core "bounces." The current thinking is that the bounce produces a shock wave which blows off the outer core and envelope of the star while a neutron star may be formed at the center. The large kinetic energy of the collapsing core is then converted to heat which ends up in radiation and in the kinetic energy of the material blown off the star. This process is thought to be the origin of some supernovae, although current models of collapsing stars do not follow this scenario in full three-dimensional detail. The problems with the models are thought to arise from simplifying assumptions, for example, no rotation or magnetic fields, which are needed to make the calculations tractable. *See* SHOCK WAVE; SUPERNOVA.

Collapse to a black hole. Some collapsing stars are expected to end up with a core that has a mass larger than the maximum mass of a neutron star. In addition, neutron stars (and white dwarfs as well) that had been formed earlier could accrete enough matter to become more massive than the maximum mass of a neutron star. In these cases, the only known alternative is for such stars to collapse to a black hole. A black hole is a region in space in which gravity is so strong that even light cannot escape from its surface. Although black holes had been conjectured earlier, Karl Schwarzschild found the first black-hole solution of general relativity in 1916, although the significance of the solution as a black hole was not realized at the time. After suggestions that black holes might be an end point of stellar evolution, J. R. Oppenheimer and H. Snyder showed in 1939 that a black hole must result from spherically symmetric gravitational collapse if the mass of the collapsing body is large enough. At present, a black hole is believed to be the only result of a collapsing star that cannot lose enough matter to become a white dwarf or neutron star. Black holes are even more condensed than neutron stars. For example, a 1-solar-mass black hole would have a radius of about 2 mi (3 km). *See* BLACK HOLE.

The collapse of a star to a black hole could take place without a supernova occurring, in which case the disappearance of the star might not be seen as a visible event. However, one potentially observable characteristic of such a collapse would be the emission of gravitational waves (discussed further below) as the black hole settled down to its equilibrium configuration. Gravitational waves interact with matter even more weakly than neutrinos, but detectors being planned and built would be sensitive enough to measure the gravitational radiation coming from collapse of a nearby star to a black hole. *See* GRAVITATION; GRAVITATIONAL RADIATION.

Because black holes emit no light, they are intrinsically dark and directly unobservable. However, they can still have an effect on nearby matter because their gravitational field is still present. In fact, there is strong evidence that at least one condensed body in a binary star system, Cygnus X-1, is a black hole. Its gravitational field results in matter heating up sufficiently to emit x-rays, which can only be done by a highly condensed body, and its mass is determined from the orbital motion to be larger than 5 solar masses. From theory, the only known body which has these properties is a black hole. Black holes have also been identified by similar evidence in

several other x-ray binaries, including LMC X-3 and A0620-00. There is no proof that such black holes came from the gravitational collapse of a star, but that is the most reasonable hypothesis. *See* ASTROPHYSICS, HIGH-ENERGY; BINARY STAR; X-RAY ASTRONOMY.

One might think that the range of masses of black holes formed from collapse is limited, since a normal star can burn nuclear fuel under stable conditions only if its mass is less than about 60 solar masses. Of course, black holes of larger than 60 solar masses could exist if sufficient additional matter fell into existing black holes or if many black holes coalesced to form a larger black hole. However, it has been conjectured that supermassive stars could exist, with masses between 10^3 and 10^7 solar masses, in which the pressure gradient arises from radiation pressure, and the release of gravitational energy during contraction provides the major or, for the more massive stars, the only energy source. Theoretical studies have shown that the more massive stars would become, at some stage in their evolution, unstable to gravitational collapse, ending up as supermassive black holes. There is strong evidence for supermassive black holes at the centers of some galaxies, including the Milky Way Galaxy, and supermassive black holes are at the centers of quasars. Even if the existence of such large black holes is confirmed, it would not prove that the black holes necessarily came from the collapse of a supermassive star. *See* GALAXY, EXTERNAL; MILKY WAY GALAXY; QUASAR; SUPERMASSIVE STARS.

General relativistic collapse. All scenarios of gravitational collapse involve Einstein's theory of gravitation, general relativity, in an intimate way. In some cases, general relativity is responsible for the instability in the star which starts the collapse. In other cases, general relativity prevents the formation of any final state of matter other than a black hole. While these effects are important, they do not illustrate the dramatic way in which general relativity changes perception of the geometry of space and time. This is best illustrated by imagining what gravitational collapse to a black hole would look like, both to an observer looking at the event from outside the system and to an observer riding down on the surface of the collapsing star. In order to lessen the effect of the gravitational tidal forces, the same kind of forces responsible for the tides on the Earth, it is assumed that the collapse is of a supermassive star.

The outside observer sees the star initially falling inward, not unlike what one would expect from newtonian gravity. However, as the radius of the star decreases to a value close to the Schwarzschild radius (the radius of the black hole that will be formed), the rate of falling inward slows. In fact, the star never quite visibly reaches the Schwarzschild radius, even after an infinite time, as far as the outside observer is concerned. Another important effect comes about because of the gravitational redshift, the shifting of light toward the red part of the spectrum when the emitting atoms are in a region of large gravitational potential. The light from the star is shifted more and more toward longer wavelengths as the star approaches the Schwarzschild radius. To the outside observer, any light emitted by the star becomes shifted quickly outside the visible range of the spectrum and the star becomes dark. Finally, since the redshifting of the light just reflects the apparent slowing down of the atomic clocks in the atoms in the star, other clocks, including the life processes of the observer riding down on the surface of the star, will be similarly slowed down. As the images of the star and the observer on the surface fade out of view, there will be left a picture of the star and the observer at some definite time, as when a motion picture slows to a halt. *See* GRAVITATIONAL REDSHIFT.

The observer riding down on the surface has a quite different view of the process. Nothing particularly unusual is observed as the star collapses past its Schwarzschild radius, which occurs at the same time that the outside observer sees the star frozen in place as it fades out of view. However, crossing that radius does limit the options of the observer on the star; no matter what the observer does, escape past the Schwarzschild radius into the outside universe is impossible. Moreover, no observer, riding on the surface of a collapsing star, can escape the fate that awaits him or her. Initially, as the star gets smaller, unusual optical effects are noticed—the surface of the star appears to rise around the observer and the sky becomes smaller. More ominous is the increase in the tidal force on the observer. At the same time, this force stretches the observer from head to toe and squeezes the observer from the sides. After a short time this force rises to an infinite value, crushing both the observer and the star to infinite density. At this point the star is said to have reached a singularity. Once inside the Schwarzschild radius, the observer has no way of avoiding the singularity. Even if the observer were able to move outward at the speed of light, he or she would be crushed by the singularity in a short time. *See* RELATIVITY. Philip C. Peters

Observation of collapse. Although gravitational collapse does not necessarily imply a violent astronomical event, the existence of supernovae in the Milky Way Galaxy and in other galaxies is taken to be evidence for the collapse of the cores of stars. Moreover, in the Milky Way Galaxy the Crab and Vela supernovae now show the remnants of a dramatic explosion. Both remnants also have pulsars approximately at their center, pulsars being identified as rapidly rotating neutron stars. The Crab supernova was noted when it occurred in 1054, so there is no question that the remnant seen today came from an explosive event. Prior to 1987, these examples provided the best evidence that at least some stars undergo gravitational collapse to a neutron star, in the process generating a supernova explosion. However, many questions could not be answered, in particular the nature of the star before collapse and the details of the collapse itself. *See* PULSAR.

On February 24, 1987, a supernova was observed in the Large Magellanic Cloud, a small galaxy associated with the Milky Way Galaxy, and therefore

relatively close to Earth as far as supernovae are concerned. This supernova, named 1987A, must be considered one of the most important astronomical events of the twentieth century. For the first time a supernova occurred in a star that could be identified and studied on existing plates, and the record of the evolution of the supernova has been obtained with remarkable detail, using all parts of the spectrum from radio waves to x-rays and gamma rays. Moreover, prior to the appearance of the supernova optically, neutrinos from the supernova were observed in two widely separated detectors, confirming for the first time the predicted production of neutrinos in supernova explosions. Information from the light curve yields information on the radioactive elements produced in the explosion. The gravitational collapse of the core of the star was expected to end up as a neutron star. However, not all the details corresponded to existing models. For example, the progenitor star was a blue supergiant, not a red supergiant, which is the kind of star, based on models, thought most likely to explode. Perhaps connected with this is the fact that the supernova was initially much dimmer than a normal supernova would have been at that distance. Substantial mass loss of the progenitor stars played an important role in its evolution. The effects of this mass loss have been seen in the complex cocoon-like structure that developed around 1987 as the rapidly expanding stellar ejecta from the explosion interacted with the more slowly expanding ejecta from the progenitor star. *See* SUPERNOVA.

Events nowadays widely believed to be connected with the core collapse of stars are gamma-ray bursts. These bursts have durations from a few milliseconds to a few minutes and emit gamma rays, the most energetic photons in the universe. Gamma-ray bursts shine hundreds of times brighter than typical supernova and are observed roughly once a day from random positions on the sky. It is believed that gamma-ray bursts are created in extremely energetic supernovae called hypernovae. Only about one out of 100,000 supernovae is a hypernova, and the cores of stars collapsing in hypernovae are believed to be massive enough to produce a black hole, connecting gamma-ray bursts with the formation of black holes. While the core collapses to form a black hole, a blast wave is created which moves through the star at a speed close to the speed of light. Once the blast wave collides with the stellar material inside the star, the energy of the blast wave is partially converted into the production of gamma rays, which leave the star's surface with the speed of light. Roughly 2700 gamma-ray bursts have been observed with the *Compton Gamma Ray Observatory's* Burst And Transient Source Experiment (BATSE). *See* GAMMA-RAY BURSTS.

Future measurements of gravitational collapse will focus on measuring gravitational wave signals. In the case of a nonspherical symmetric collapse, the general theory of relatively perdicts the emission of gravitational waves. These waves are ripples in space-time, moving through it with the speed of light and perturbing it (**Fig. 2**). The basic idea is to measure the perturbation to space-time that the gravitational wave induces. The best way to do so is in interferometer experiments. An interference pattern is generated by a laser beam, which is split into two beams initially in phase with each other and moving perpendicularly away from each other. These beams are reflected several times between a pair of mirrors. When a gravitational wave passes the experiment, the space between the mirrors in each perpendicular arm of the experiment will be changed differently, resulting in a change in the interference pattern that can be measured. Several experiments—ground-based such as the Laser Interferometer Gravitational-Wave Observatory (LIGO), and space-based such as the Laser Interferometer Space Antenna (LISA)—will be trying to measure these signatures. *See* LIGO (LASER INTERFEROMETER GRAVITATIONAL-WAVE OBSERVATORY).

Fig. 2. Artist's conception of gravitational waves. (*LIGO*)

Sadegh Khochfar; Philip C. Peters; Joseph Silk

Bibliography. D. Arnett, *Supernovae and Nucleosynthesis*, 1996; H. A. Bethe and G. Brown, How a supernova explodes, *Sci. Amer.*, 252(5):60–68, 1985; T. Piran, S. Weinberg, and J. C. Wheeler (eds.), *Supernovae*, 1991; S. L. Shapiro and S. A. Teukolsky, *Black Holes, White Dwarfs, and Neutron Stars*, 1983; J. C. Wheeler, *Cosmic Catastrophes: Supernovae, Gamma-ray Bursts, and Adventures in Hyperspace*, 2000.

Gravitational lens

A massive body producing distorted, magnified, or multiple images of more distant objects when its gravitational fields bend the paths of light rays. Lenses have been observed when the light from very distant quasars is affected by intervening galaxies and clusters of galaxies, producing several different images of the same quasar. A. Einstein predicted the occurrence of this phenomenon in 1936, but the discovery of real gravitational lenses did not occur until 1979. Gravitational lenses, in addition to being intrinsically interesting, can reveal the intrinsic properties of galaxies, active galaxies, and quasars, and provide information on the universe and its contents, including dark matter.

Action of gravity. The lens phenomenon exists because gravity bends the paths of light rays, which is

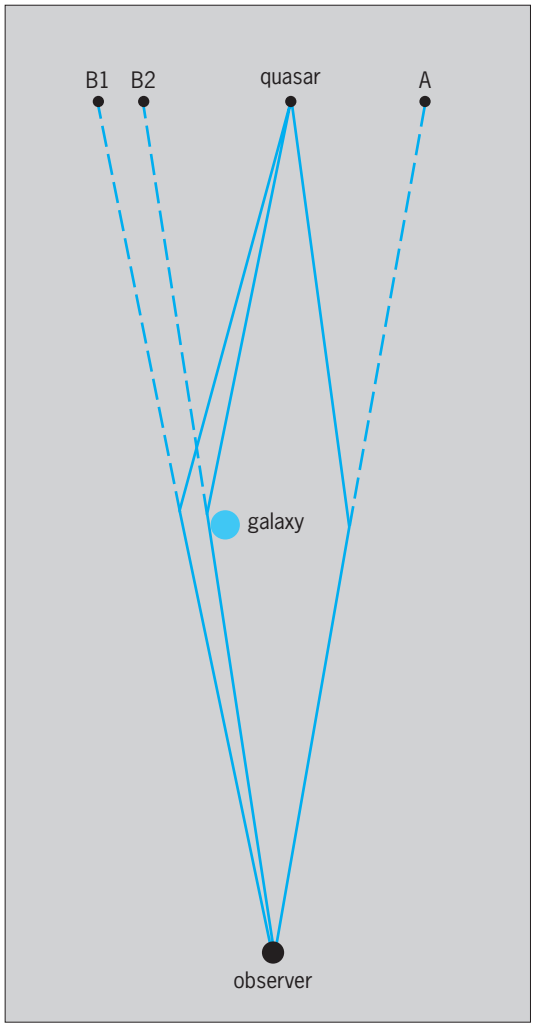

Fig. 1. Schematic illustration of a gravitational lens. The angles are exaggerated for clarity. Here, the lens action produces three images of a quasar (A, B1, and B2), since light from the image can travel along three different curved paths and still reach the observer.

predicted by Einstein's general theory of relativity. Since photons, the carriers of light energy, have no mass, Newton's theory of gravity indicates that light would always travel in a straight line even if there were heavy, massive objects between the source and the observer. (Even if photons are given mass in Newton's theory, the predicted bending of light is different from the result in general relativity.) But in general relativity, gravity acts by producing curvature in space-time, and the paths of all objects, whether or not they have mass, are also curved if they pass near a massive body. *See* GRAVITATION; RELATIVITY.

Numerous eclipse observations have confirmed Einstein's prediction with modest accuracies of 20–30%. Radio astronomers have measured changes in the positions of quasars that occur when the Sun passes near them in the sky. The precision of these experiments, which fit Einstein's predictions, is now at the level of tenths of a percent.

A massive object acts as a gravitational lens when light rays from a distant quasar are bent around or through it and are focused to form an image, which can be seen or photographed by an astronomer on Earth (**Fig. 1**).

Sometimes a lens can amplify the total intensity of light in a quasar image, making it considerably brighter than the quasar would appear to be in the absence of a lens. If the galaxy and quasar are sufficiently well aligned, several images of the same quasar will appear, since light can travel on many different paths and still arrive at the detecting telescope. In some cases the lensed image is part of a ring. It has been shown that if there are multiple images of the same quasar, there must be an odd number of them, as long as the galaxy is big enough so that it does not act as a point mass.

Discovery of lenses. Astronomers treated gravitational lenses as curiosities for a long period of time. Space is so empty that the probability of two stars being aligned accurately enough is extremely small. But the discovery of quasars, hyperactive galaxies which are bright enough to be visible even though they are nearly 10^{10} light-years (1 light-year is equal to 5.88×10^{12} mi or 9.46×10^{12} km) away, made it possible to probe a much larger volume of space. Now, there was a reasonable chance that a galaxy might lie in the path of light traveling from a quasar to the Earth. *See* QUASAR.

A survey of faint blue objects in the northern sky by Richard Green and Maarten Schmidt, completed in 1977, was one of the keys to establishing gravitational lenses as actual (as opposed to hypothetical) astrophysical objects. This Palomar-Green survey isolated several pairs of quasars, where two very similar faint blue objects were very close to each other in the sky. Subsequent investigation demonstrated that these two objects were not only very similar; they were two distinct images of the same object that were being produced when light from the object passed around a galaxy that was between the quasar and the Earth. In some cases, there were three images; **Fig. 2** shows one of the early multiple

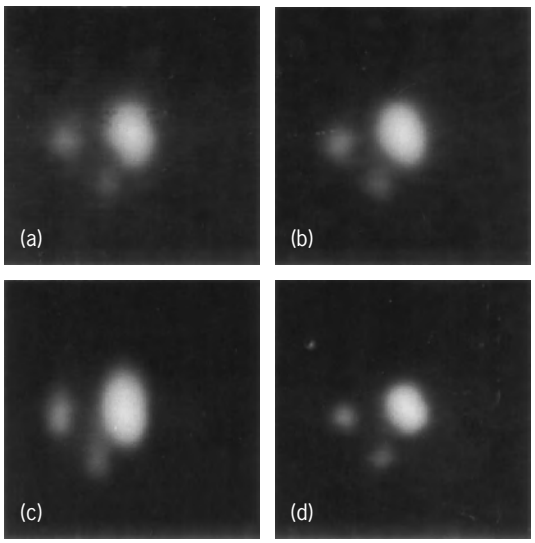

Fig. 2. Images of the triple quasar PG 1115 + 08. (*a*) Blue filter. (*b*) *V* band (yellow) filter. (*c*) Unfiltered image. (*d*) Red filter. (*From E. K. Hege et al., Morphology of the triple QSO PG1115 + 08, Nature, 287:416–417, 1980*)

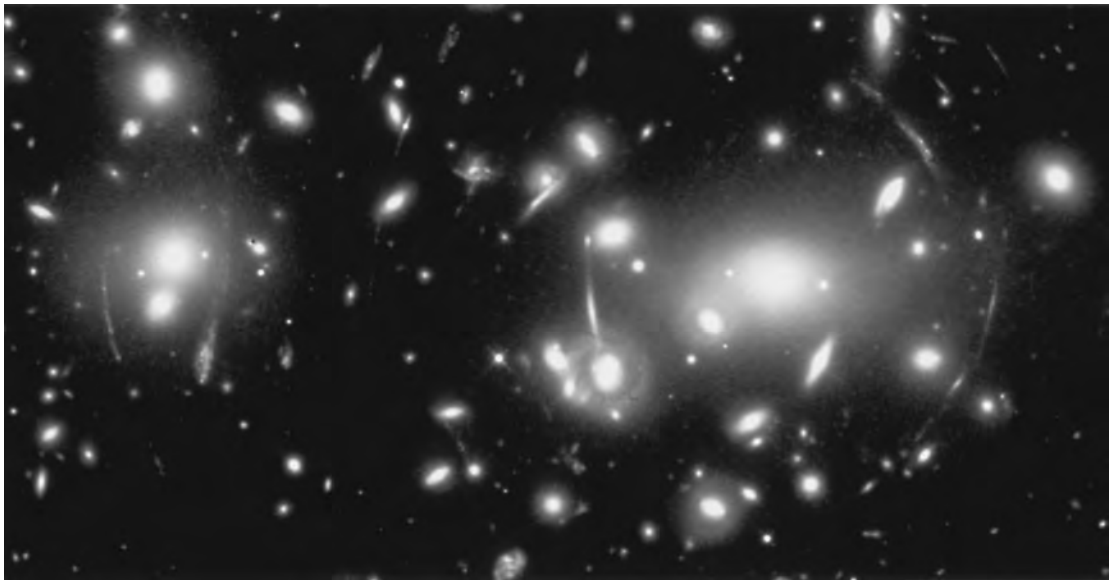

Fig. 3. Image of the galaxy cluster Abell 2218 obtained from the *Hubble Space Telescope*. The ring of light is an illusion, light from a very distant object that is bent by the gravity from the galaxy cluster. (*NASA*)

quasars, PG1115 + 08. (PG designates the Palomar-Green survey, and the numbers provide an approximate position in the sky in terms of the astronomical coordinates of right ascension and declination.)

More striking are some ringlike arcs produced when the distant object, the lensing galaxy or galaxy cluster, and the observer are exactly in a straight line. **Figure 3** is a picture, taken with the *Hubble Space Telescope*, of one of the most impressive arcs. When light passes through the cluster, it is bent. The distant object could be either a galaxy or a quasar, but it is definitely much farther away than the cluster. The cluster acts like a little telescope of its own, focusing the light from the distant object toward the Earth and making it brighter than it otherwise would be. *See* SATELLITE (ASTRONOMY).

Implications. The discovery of gravitational lenses affects astronomers' understanding of the universe on the very largest scales. The very existence of this phenomenon indicates that nearly a dozen quasars—the ones that are being lensed—are more distant than the galaxies that are focusing their light. When quasars were first discovered, some astrophysicists argued that their redshifts were produced by exotic new physics and the quasars were just beyond the boundary of the Milky Way Galaxy. This controversy has largely subsided but has not been completely resolved. The lens phenomenon shows that at least some quasars are billions of light-years away, well beyond the edge of the Milky Way. If some quasars are billions of light-years away and others look like them, it is reasonable to conclude that all quasars are billions of light-years away at the edge of the observable universe.

Gravitational lenses can be used to determine the distance scale of the universe. Most quasars change the amount of light that they produce. In the case of a multiply imaged quasar like PG 1115+080 (Fig. 2), observers on Earth could see that change occur at different times because light travels on different paths to get here. The image where light travels on a more direct path would brighten first, and the one taking a more roundabout route would brighten later. The differences in the two path lengths can be used to deduce the distance to the quasar and the lensing object. Astronomers can then measure the redshifts of these distant objects and use the lens as another way to determine how fast the universe is expanding.

This seemingly easy idea is hard to implement in practice. Only one of 500 quasars is lined up in exactly the right way, and the mass distribution in the lensing object must be fully understood in order to interpret the data correctly. Sharp infrared pictures from the *Hubble Space Telescope* severely constrained possible models of the lens, making the interpretation much more secure. The data indicate a Hubble constant of 70 kilometers per second per megaparsec (the conventional units for measuring the Hubble constant), meaning that, if the universe has a very low density, its age is 14 billion years. *See* COSMOLOGY; HUBBLE CONSTANT.

Dark matter. Gravitational lenses also enable the discovery of invisible objects. The speed with which stars move in galaxies and galaxies move in galaxy clusters indicates that many galaxies may be surrounded by massive dark halos. Since the matter that composes these halos cannot be seen, the name "dark matter" has been used to describe it. The dark matter could be brown dwarfs (objects not massive enough to be stars), dead stars, Jupiter-sized objects, or subnuclear particles. The more massive forms of dark matter are termed MACHOs (massive compact halo objects). *See* BROWN DWARF.

If a MACHO passed directly between a distant star and the Earth, the light from the star could be temporarily brightened as the MACHO focused the starlight toward the Earth. Precise calculations of this event indicate that the brightening should last about

a week. Several teams of astronomers have made repeated observations of a nearby galaxy, the Large Magellanic Cloud, in search of this phenomenon, and seven such events have been detected. These events indicate that MACHOs, which are probably low-mass white dwarf stars, make up a sizable fraction of the mass of the halo of the Milky Way Galaxy, probably at least 20% of the dark matter and possibly as much as 100%. *See* MAGELLANIC CLOUDS; MILKY WAY GALAXY.

Harry L. Shipman

Bibliography. N. Cohen, *Gravity's Lens: Views of the New Cosmology*, 1988; J. Glanz, Is the dark matter mystery solved?, *Science*, 271:595–596, 1996; R. Schild, Gravity is my telescope, *Sky Telesc.*, 81(4):375–379, April 1991; G. Taubes, OGLEing, MACHOs, and the search for dark matter, *Science*, 260:492–493, 1993; K. Thorne, *Black Holes and Time Warps: Einstein's Outrageous Legacy*, 1994.

Gravitational radiation

A wave that is generated by the acceleration of mass and travels at the speed of light. Gravitational waves are an implicit outcome of the special theory of relativity, and were explicitly put forward in Albert Einstein's theory of general relativity in 1916. Einstein showed that the acceleration of masses generates time-dependent gravitational fields which can carry energy away from their source at the speed of light. *See* RELATIVITY.

Properties. The essence of general relativity is that mass and energy produce a curvature of four-dimensional space-time and that matter moves in response to this curvature. The Einstein field equations proscribe the interaction between mass and space-time curvature, much as Maxwell's equations proscribe the relationship between electric charge and electromagnetic fields. One solution to the field equations is a weak, oscillating perturbation to space-time curvature, that is, a gravitational wave, just as electromagnetic waves, in the form of light, microwaves, or x-rays, are solutions to Maxwell's equations. Gravitational waves can be thought of equivalently in several ways: as an oscillating perturbation to a flat, or Minkowski, space-time metric; as an oscillating strain in space-time; as an oscillating tidal force between free test masses. *See* ELECTROMAGNETIC RADIATION; MAXWELL'S EQUATIONS; SPACE-TIME.

As with electromagnetic waves, gravitational waves travel at the speed of light and are transverse in character; that is, the strain oscillations occur in directions orthogonal to the direction in which the wave is propagating. Whereas electromagnetic waves are dipolar in nature, gravitational waves are quadrupolar: the strain pattern contracts space along one transverse dimension, while expanding it along the orthogonal direction in the transverse plane (see **illustration**). This characteristic arises from the fact that gravitational radiation is produced by oscillating multipole moments of the mass distribution of a system, just as electromagnetic waves are produced by oscillating moments of the charge distribution. The principle of mass conservation rules out monopole radiation, just as charge conservation does so in electromagnetism, and the principles of linear and angular momentum conservation rule out gravitational dipole radiation. Quadrupole radiation is the lowest allowed form and is thus usually the dominant form. In this case, the gravitational-wave field strength is proportional to the second time derivative of the quadrupole moment of the source, and if the gravitational wave is to carry away energy, the field strength must fall off in amplitude inversely with distance from the source. The tensor character of gravity—the hypothetical graviton is a spin-2 particle—means that the transverse strain field comes in two orthogonal polarizations, rotated with respect to each other in the transverse plane by 45°. *See* ANGULAR MOMENTUM; CONSERVATION LAWS (PHYSICS); GRAVITON; LINEAR

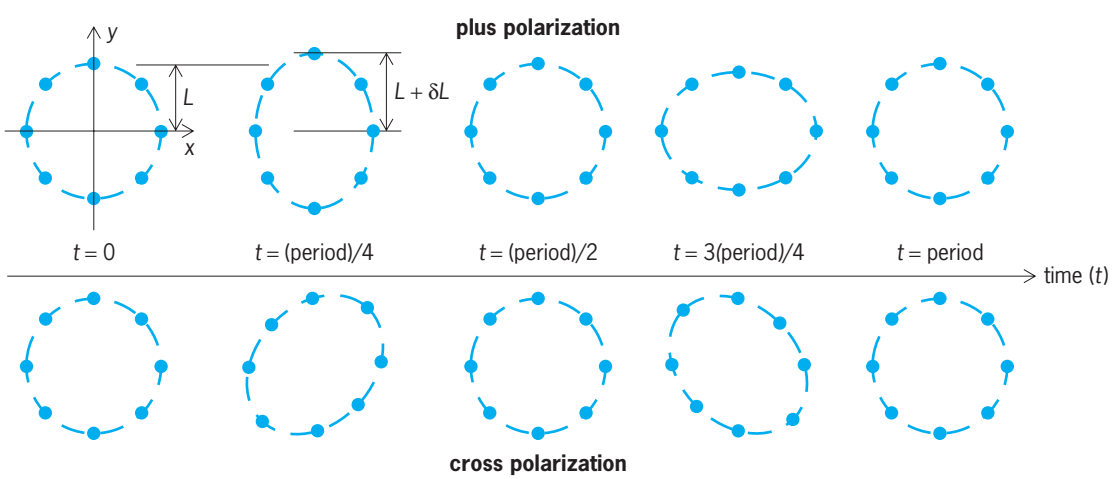

Effect of a gravitational wave, passing in a direction perpendicular to the plane of the paper, on a ring of test particles. The ring is distorted into an ellipse, elongated in one direction in one half-cycle of the wave, and elongated in the orthogonal direction in the next half-cycle. The waves come in two independent polarization states, denoted plus polarization and cross polarization; the strain patterns for the two states are rotated 45° from each other.

MOMENTUM; MOMENTUM; MULTIPOLE RADIATION; POLARIZATION OF WAVES.

Gravitational waves differ from electromagnetic waves in that they propagate essentially unperturbed through space, as they interact only very weakly with matter. Furthermore, gravitational waves are intrinsically nonlinear, because the wave-energy density itself generates additional curvature of space-time. However, this phenomenon is significant only very close to strong sources of waves, where the wave amplitude is relatively large. More usually, gravitational waves distinguish themselves from electromagnetic waves by the fact that they are very weak. One cannot hope to detect any waves of terrestrial origin, whether naturally occurring or artificial; instead one must look to very massive astrophysical objects moving at very high velocities. For example, strong sources of gravitational waves that may exist in our galaxy or nearby galaxies are expected to produce wave strengths on Earth that do not exceed strain levels of one part in 10^{21}.

Sources. The prototypical example of a gravitational-wave source is a pair of compact astrophysical objects such as stars in mutual orbit about their center of mass, otherwise known as a binary star system. The pair could consist of two neutron stars, two black holes, or one of each type of object, both of which are remnants of once-giant, massive stars. Both types of objects are extremely dense, which is key to generating strong gravitational waves. A binary system will radiate waves with a period equal to half the orbital period. This period, and the physical separation between the objects, will decrease over time as gravitational waves carry energy away, a process known as an inspiral. Eventually, the objects will come close enough to each other that they merge into a single object, during which gravitational waves will continue to be emitted. This coalescence of two neutron stars or a neutron star and a black hole is now thought to be the source of the so-called short gamma-ray bursts that have been detected by various gamma-ray observatories. Finally, the merged objects may form a black hole, which itself would oscillate in a normal mode, emitting further gravitational energy in heavily damped sinusoidal waves. *See* BINARY STAR; BLACK HOLE; GAMMA-RAY BURSTS; NEUTRON STAR.

Another example of a source is an isolated, spinning neutron star, which deviates in shape from a perfect sphere (so that it has a quadrupole moment). Such a star would emit nearly monochromatic gravitational waves, with the wave frequency changing only very slowly as energy is lost. On the other hand, short bursts of gravitational waves could be emitted by supernovae and other stellar collapses or by the infall of a massive object into a black hole. Finally, there may be a stochastic background of gravitational waves, analogous to the cosmic microwave background, that originates early in the evolution of the universe. Indeed, the inflationary theory of cosmology predicts that such a background should exist, though it would be very weak in amplitude. Other posited cosmological sources of a stochastic background include a network of "cosmic strings," and rapidly expanding temperature bubbles, formed by a phase transition that may have taken place as the universe cooled and expanded. *See* COSMIC BACKGROUND RADIATION; COSMIC STRING; COSMOLOGY; GRAVITATIONAL COLLAPSE; INFLATIONARY UNIVERSE COSMOLOGY; PHASE TRANSITIONS; SUPERNOVA; UNIVERSE.

Detection. Because gravitational waves are inherently so weak—very much weaker than electromagnetic waves—Einstein concluded that they would never be detected. Nevertheless, efforts to directly detect gravitational waves have been made since the 1960s, though as yet with no detection. Even indirect evidence of their existence did not come until many decades after their prediction, with the 1974 discovery and subsequent observations of a binary pulsar by Russell Hulse and Joseph Taylor. By precisely monitoring the orbital period of this binary star system, thought to be made up of two neutron stars, they were able to confirm that the orbit was speeding up at just the rate predicted by the general-relativistic emission of gravitational waves. *See* PULSAR.

Resonant-mass detectors. The effort to detect gravitational waves directly started in the early 1960s with Joseph Weber's experiments with resonant acoustic bar antennas. This type of detector consists of a large mass equipped with highly sensitive vibration detectors. The mass is typically a cylinder of aluminum weighing several tons and having a resonant vibrational mode around 1000 Hz. A passing gravitational-wave burst would resonantly excite the cylinder's mode, and the mode amplitude could be measured using some type of electromechanical transducer that is tuned to the mode. This line of research expanded considerably after 1969 in response to Weber's apparently claimed detection of gravitational waves (though Weber himself later said that he had not actually claimed a detection), as over a dozen groups tried to construct similar bar detectors to see if Weber's results could be reproduced. None were able to confirm the detection. The sensitivity of bar detectors has increased significantly since the initial experiments, with the detectors being cooled to cryogenic temperatures to reduce thermal noise and isolated from terrestrial vibrations by a series of mechanical filters. Present acoustic bar detectors have achieved a strain sensitivity to short bursts of several parts in 10^{20}.

Free-mass detectors. The second, newer method of detecting gravitational waves on Earth uses "free masses" separated from each other by long baselines whose lengths are precisely monitored using a laser interferometer. The basic detection idea was first suggested by Mikhail E. Gertsenshtein and V. I. Pustovoit in 1962 and was independently conceived of by Rainier Weiss in 1972. Robert Forward built the first prototype gravitational-wave interferometer in the 1970s, and this early work was quickly followed by more prototypes at numerous research laboratories. As with the bar detectors, the sensitivities of the prototype interferometers improved by several orders of

magnitude over the first detectors. But it was clear from the start that interferometers would need to be greatly scaled up in size before they could hope to achieve a sensitivity sufficient to detect gravitational waves. The laboratory prototype research eventually led to the design and construction of several long-baseline interferometric detectors around the world: the LIGO (Laser Interferometer Gravitational-Wave Observatory) detectors at two locations in the United States, constructed with 4-km-long (2.5-mi) arms; the Virgo detector in Cascina, Italy, with 3-km-long (1.9-mi) arms; the GEO detector outside Hannover, Germany, with 600-m-long (2000-ft) arms; and the TAMA detector near Tokyo, Japan, with 300-m (1000-ft) arms. Presently the LIGO detectors are the most sensitive and are able to detect strains as small as one part in 10^{21}. *See* LIGO (LASER INTERFEROMETER GRAVITATIONAL-WAVE OBSERVATORY).

All of these detectors are L-shaped Michelson interferometers, a configuration that takes advantage of the quadrupolar field pattern of a gravitational wave. The test-mass mirrors of the interferometer are suspended as low-frequency pendula, so that they are free to be moved by a passing gravitational wave. The interferometer is illuminated with a powerful and extremely well-stabilized laser, and the arms and main components are contained in a high-vacuum enclosure to reduce fluctuations from air molecules. When a gravitational wave passes through the interferometer, the light travel time down and back one arm will become shorter or longer relative to the round-trip time in the other arm; this will result in a light signal at the interferometer output, proportional to the wave amplitude. The optical beams are usually folded many times within the arms, typically using resonant optical cavities, to increase the effective length of the arms and thus the change in light travel time. This type of detector is more broadband than the resonant-bar type, and is sensitive to gravitational waves in a frequency band starting around 10 Hz and extending to nearly 10 kHz. At frequencies below 10 Hz, terrestrial vibrations and forces simply become too large, and at the upper end of the band the interferometers lose sensitivity as the gravitational wavelength becomes shorter than the effective arm length. *See* INTERFEROMETRY.

Detectors in space. Searching for gravitational waves at frequencies below 10 Hz requires detectors in space, several of which have been conceived. For gravitational-wave periods between a few minutes and a few hours, the best detection method at present is the Doppler tracking of spacecraft. In this method, a highly coherent microwave signal is generated on Earth and transmitted to the spacecraft. The microwaves are received by the spacecraft and transponded back to Earth. The frequencies of the transmitted and received signals are compared, and from this the relative velocity between Earth and the spacecraft is determined. A gravitational wave would produce fluctuations in the measured Doppler shift, proportional to the wave amplitude. Interestingly, each feature in the gravitational waveform shows up three times in the Doppler shift, a fact that can be used to help distinguish gravitational waves from noise in the data. In December 2001, the *Cassini* space probe was tracked in this way as it cruised toward Mars. An upper bound was set on the amplitude of gravitational waves at frequencies around a milliHertz, at a level of 2 parts in 10^{15}.

Monitoring pulsar signals. At even lower frequencies, below 10^{-5} Hz, very good limits on gravitational waves have come from observations that monitor the timing signals from pulsars. The basic idea here is that the rotation period of the pulsar's neutron star is extremely stable. By comparing the timing of the received pulsar signal with the most stable atomic clocks on Earth, one can look for fluctuations in the pulsar timing that would occur if gravitational waves pass through the intervening space. These measurements have set tight bounds on the strength of gravitational waves having periods of around 1 year.

LISA project. The proposed LISA project—the Laser Interferometer Space Antenna—would bridge some of the frequency gap between the terrestrial interferometers and spacecraft tracking. The LISA concept is to construct an interferometric detector in space, where it would be free of terrestrial background noise and the measurement baseline can be made extremely long. The plan calls for a trio of satellites to be launched into orbits around the Sun, separated from each other by 5×10^6 km (3×10^6 mi). With such a long baseline, the instrument would be sensitive to gravitational waves in the frequency band from 1 mHz to 0.1 Hz. The technology for this mission is under development.

Cosmic background measurements. Finally, there is hope that future measurements of the cosmic microwave background will show signs of gravitational radiation arising from inflation, the hypothesized epoch when the universe underwent nearly exponential expansion. Gravitational waves produced during inflation would leave their imprint on the polarization of the microwave background, which may still be detectable today.

Gravitational-wave astronomy. Physicists and astrophysicists are hopeful that these detectors will soon provide direct observations of gravitational waves, opening up the field of gravitational-wave astronomy. Our present knowledge of the universe, what it is made of and its history, has been acquired largely through observations of electromagnetic radiation. As technology progressed and each new region in the electromagnetic spectrum was opened up for observation, unexpected and important information was revealed. But for many phenomena, gravitational radiation carries information that cannot be found in the electromagnetic spectrum. Probing the universe with gravitational radiation is thus likely to bring exciting surprises and unanticipated astrophysics. *See* GRAVITATION.

Peter K. Fritschel

Bibliography. B. Barish and R. Weiss, LIGO and the detection of gravitational waves, *Phys. Today*, 52(10):44–50, October 1999; M. Bartusiak, Catch a gravity wave, *Astronomy*, 28(10):54–59, October 2000; M. Bartusiak, *Einstein's Unfinished Symphony: Listening to the Sounds of Space-Time*,

Joseph Henry Press, Washington, DC, 2000; K. S. Thorne, Gravitational radiation, in S. Hawking and W. Isrrael (eds.), *3000 Years of Gravitation*, chap. 9, pp. 330–458, Cambridge University Press, 1987.

Gravitational redshift

A shift toward longer wavelengths of spectral lines emitted by atoms in strong gravitational fields. It is also known as the Einstein shift. One of three famous predictions of the general theory of relativity, this shift results from the slowing down of all periodic processes in a gravitational field. The amount of the shift is proportional to the difference in gravitational potential between the source and the receiver. For starlight received at the Earth the shift is proportional to the mass of the star divided by its radius. In the solar spectrum the shift amounts to about 0.001 nanometer at a wavelength of 500 nanometers. In the spectra of white dwarfs, whose ratio of mass to radius is about 30 times that of the Sun, the shift is about 0.03 nm, which can easily be measured if it can be separated from the Doppler effect. This was first done by W. S. Adams for the companion of Sirius, a white dwarf whose true velocity relative to the Earth can be deduced from the observed Doppler effect in the spectrum of Sirius. The measured shift agreed with the prediction based on Einstein's theory and on independent determinations of the mass and radius of Sirius B. A more accurate measurement was carried out in 1954 by D. M. Popper, who measured the gravitational redshift in the spectrum of the white dwarf 40 Eridani B. Similar measurements, all confirming Einstein's theory, have since been carried out for other white dwarfs. Attempts to demonstrate the gravitational redshift in the solar spectrum have thus far proved inconclusive, because it is difficult to distinguish the gravitational redshift from so-called pressure shifts resulting from perturbations of the emitting atoms by neighboring atoms. *See* WHITE DWARF STAR.

In 1960 R. V. Pound and G. A. Rebka, Jr., succeeded in measuring the gravitational redshift in a laboratory experiment. ^{57}Fe nuclei bound in a solid emit and absorb gamma rays in an exceedingly narrow frequency range ($\Delta\nu/\nu = 3 \times 10^{-13}$) centered on a sharply defined frequency near 14.4 keV. A source and an absorber of this accurately monochromatic radiation were separated by a vertical distance h of 74 ft (22.5 m), so that the radiation incident on the absorber was shifted in frequency by an amount given by the equation below, according to Einstein's theory, where g is the acceleration of gravity and c is the speed of light. The rate of absorption was measurably diminished relative to the rate that would have been observed in the absence of the gravitational redshift. A suitable Doppler shift introduced by a small, accurately measurable relative motion of the source and the absorber restored the measured absorption rate to its normal value. The Doppler shift needed to accomplish this restoration agreed with the value predicted by Einstein's theory to within 10%. Subsequently, the error has been reduced to 1%. *See* MÖSSBAUER EFFECT.

$$\frac{\Delta\nu}{\nu} = \frac{gh}{c^2} = 2.5 \times 10^{-15}$$

In 1976 R. F. C. Vessot and collaborators used a rocket-borne hydrogen maser with a frequency stability of 1 part in 10^{14} to measure a gravitational redshift $\Delta\nu/\nu = 4 \times 10^{-10}$, corresponding to the difference in gravitational potential between the ground and the apogee of the rocket. This experiment confirmed the predicted redshift to better than 1 part in 10^4. *See* MASER.

Attempts have also been made to deduce stellar masses from measurements of the gravitational redshift, but the difficulty of allowing properly for the Doppler effect and for pressure shifts renders these determinations very uncertain.

Gravitational redshift has also been detected as temperature fluctuations in the observed cosmic microwave background. In 1967 R. K. Sachs and A. M. Wolfe showed that large-scale fluctuations in cosmic density should lead to different gravitational redshifts of observed microwave photons; the measurement of angular temperature variations therefore gives direct information about the mass distribution. This effect was observed in 1992 by the *Cosmic Background Explorer (COBE)* satellite, and has become a key tool in measuring the initial density perturbations from which cosmological structures are believed to have formed by gravitational instability. *See* COSMIC BACKGROUND RADIATION; COSMOLOGY; RELATIVITY.

David Layzer; Anthony Aguirre

Bibliography. R. d'Inverno, *Introducing Einstein's Relativity*, Clarendon Press, Oxford, 1996; C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, W. H. Freeman, San Francisco, 1973.

Graviton

A theoretically deduced particle postulated as the quantum of the gravitational field. According to Einstein's theory of general relativity, accelerated masses (or other distributions of energy) should emit gravitational waves, just as accelerated charges emit electromagnetic waves. And according to quantum field theory, such a radiation field should be quantized; that is, its energy should appear in discrete quanta, called gravitons, just as the energy of light appears in discrete quanta, namely photons. *See* PHOTON; QUANTUM FIELD THEORY; RELATIVITY.

The properties of the graviton follow from the properties of the classical gravitational field. That is, its rest mass and charge are zero (like the photon); it has spin 2 in units of $h/2\pi$, where h is Planck's constant, and is therefore a boson, which is a particle that obeys Bose-Einstein statistics. Because of its vanishing rest mass, its spin is restricted to be parallel to its motion, so that a graviton has only two independent spin states (again like a photon). *See* BOSE-EINSTEIN STATISTICS.

Observation of gravitational radiation are difficult, because matter is very weakly coupled to the gravitational field (the gravitational force between an electron and a proton is only 10^{-39} times the electrical force between them), so that the rate of emission and absorption of gravitational radiation is very small. The observation of quanta of the gravitational field, gravitons, is practically impossible according to present knowledge. As far as is known, the only physical situations in which the quantization of gravitation plays any appreciable quantitative role are the evaporation of black holes (light ones particularly), and the very high-temperature phase early in the history of the universe, namely within 10^{-33} s after the big bang. *See* BIG BANG THEORY; BLACK HOLE; COSMOLOGY; ELEMENTARY PARTICLE; GRAVITATION; GRAVITATIONAL RADIATION; UNIVERSE. Charles J. Goebel

Bibliography. G. Kane, *Modern Elementary Particle Physics*, updated ed., 1993; M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, 1995.

Gravity

The gravitational attraction at the surface of a planet or other celestial body. The quantity g is often referred to simply as "gravity" or "the force of gravity" of Earth, both of which are incorrect. The force of gravity means the force with which a celestial body attracts an object, that is, the weight of the object. The letter g represents the acceleration caused by the gravitational force and, of course, has the dimensions of acceleration. *See* EARTH, GRAVITY FIELD OF; GRAVITATION. Dirk Brouwer; G. M. Clemence

Gravity meter

A device that measures local acceleration due to the Earth's gravity; it is also called a gravimeter. Such instruments fall into two categories: relative gravity meters, which are used to determine gravity differences among a number of geographic locations or changes in gravity that occur at a single location over time; and absolute gravity meters, which can measure the true value of the acceleration due to gravity at a given location and time.

The local value of gravity is the acceleration undergone by a freely falling mass upon which gravity is the only force acting. Because the value of gravity at any particular position depends on the distribution of mass throughout the Earth (and also slightly on the Earth's rotation), measurements of the gravity field can yield information on the density of underlying rock. Thus, gravity meters are used for geologic studies and for oil and mineral exploration. Local gravity also depends on the shape of the Earth; the observation of gravity over time, then, provides a measure of deformations in the Earth that can be caused by a wide variety of phenomena, including tides, tectonic activity, and volcanism.

Nowhere on the surface of the Earth does the value of gravity differ from the nominal value of 980 Gal by more than about 0.5%. (The SI unit for gravity is the meter/second^2; the more commonly used unit is the Gal, defined as 1 cm/s^2, or the milliGal, which equals 0.001 Gal.) Values of gravity predicted with a latitude-and-height-dependent Earth model usually agree with observed values to within about 30 mGal. The gravitational acceleration produced by the mass of a 1-m-thick (3-ft) sheet of water (having infinite lateral extent) is 0.043 mGal.

A number of different instruments are available. One of two methods is used in all gravity meters. The first, employed by absolute gravity meters, is the direct determination of the acceleration of a test mass falling inside a vacuum chamber by using optical interferometry. The second is the observation of variations in the position of a mass supported by a mechanical or magnetic spring. This method, applied in relative gravity meters and shipboard gravity meters, is usually used in conjunction with an additional applied force (nulling force) that maintains the mass at a null position. The small nulling force is a relative measure of gravity.

Absolute gravity meters. In this type of gravity meter a mechanical system inside a high-vacuum chamber effects the release of a test mass (**Fig. 1**). A corner-cube prism mounted to the test mass

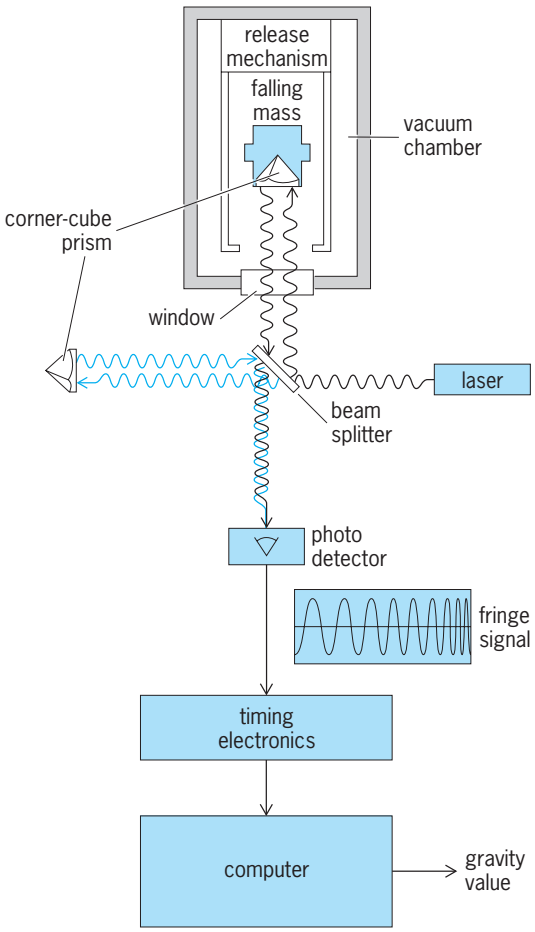

Fig. 1. Operation of an absolute gravity meter.

reflects laser light entering the chamber through a window in the bottom. A beamsplitter (or half-silvered mirror) combines this reflected light with that similarly reflected from a fixed corner-cube. The two reflections interfere with one another, producing successive light and dark fringes. A photodetector senses the fringes and generates an electrical signal from which the local value of gravity (or the rate at which the falling mass accelerates) can be calculated through appropriate timing analysis. Typically, the mass falls a distance of 10–20 cm (4–8 in.). Because the measurement is based on calibrated length and time standards—the wavelength of the laser light and the frequency of the timing oscillator—the measurement is absolute. *See* INTERFEROMETRY.

The mechanical systems vary. In some, the test mass's fall is initiated by a small elevator inside the vacuum chamber, which also surrounds the mass as it falls and acts as a shield against minute nongravitational forces (Fig. 1). In others, the mechanical system resides at the bottom of the vacuum chamber and launches the test mass upward, providing a mass in free flight, the acceleration of which can be measured during both its rise and fall. Another important variation is the use of a spring suspension system to isolate the reference reflector from vibrations.

The best instruments achieve an accuracy approaching 0.002 mGal. A complete gravity determination at a single location requires 1–2 days, the time required to make a thousand measurements (drops), the average of which is required to smooth out imperfectly corrected environmental and tidal effects. These instruments weigh several hundred pounds but are fairly portable.

A very recent technique for making an absolute measurement of gravity involves measuring the acceleration of individual atoms in free fall. While largely a laboratory exercise (but with promise for a portable instrument eventually), this technique relies on the laser-cooling and trapping of atoms. The accelerations of falling cesium atoms are determined with an atom interferometer; accuracy is similiar to that of the macroscopic free-fall method.

Relative gravity meters. These are more common than absolute gravity meters. In a relative gravity meter, a mass is suspended by a spring and its position is monitored electronically with respect to some reference position (**Fig. 2***a*). As the force of gravity on the mass changes (because, for example, the meter has been moved to a different location), the length of the spring changes slightly in order to maintain equilibrium. An additional force applied to the mass to restore its position to the original reference point is a measure of the difference in gravity between the two observation sites. In some cases, the mechanical spring is replaced by a magnetic suspension system.

Spring gravity meters. For a given change in gravity, the change in length of the spring in a relative gravity meter is proportional to the square of the period of the spring mass system. Maximizing the period enhances the meter's sensitivity; this, however, involves softening the spring and increasing the length of the suspension system. In one gravity meter of this type, a hinged beam connects the test mass to the instrument frame and the spring is oriented diagonally (Fig. 2*b*). This suspension, while requiring only a few centimeters of length, achieves a natural period in the spring-mass system of about 15 s (a vertically oriented spring having this period would stretch to a length of 55 m or 180 ft). The meter is read by noting the amount of rotation in a spring-adjusting screw (nulling screw) that is required to restore the beam to its original equilibrium position. The instrument can sense gravity differences as small as 0.003 mGal. This type of instrument is the most common type of gravity meter. Because the instrument weighs only a few pounds, can be easily transported, and requires only battery power, it is used extensively in geophysical surveying and oil exploration. *See* GEOPHYSICAL EXPLORATION.

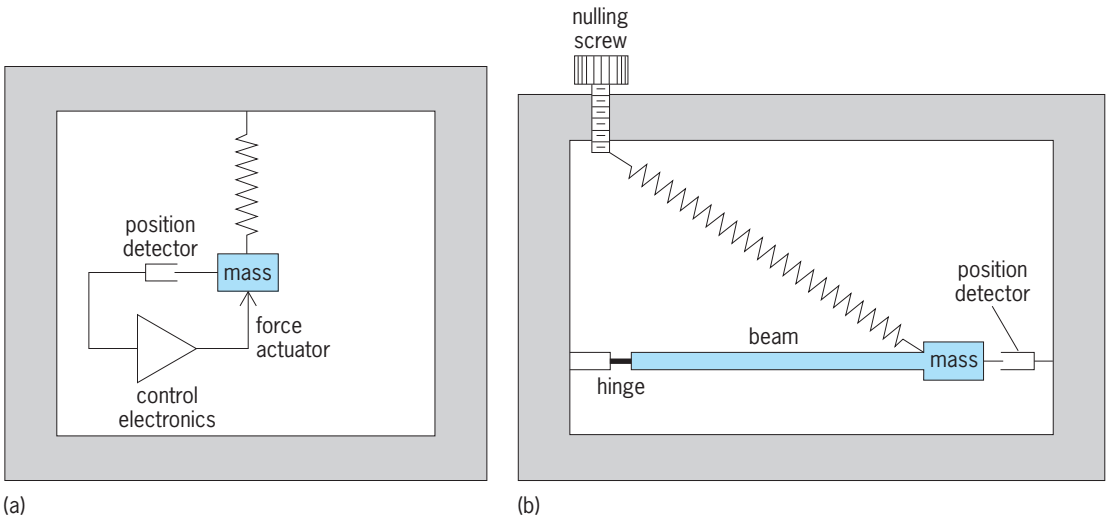

Fig. 2. Relative gravity meters. (*a*) Mass suspended by a spring whose position is monitored electronically. (*b*) Mass connected to the instrument frame by a hinged beam, with the spring oriented diagonally. (*LaCoste & Romberg*)

Superconducting gravity meters. Another type of relative gravity instrument is the superconducting gravity meter. In this instrument a 2.5-cm-diameter (1-in.) superconducting sphere is levitated by persistent currents in a pair of superconducting coils immersed in a liquid helium bath. The sphere's position in sensed capacitively by electrodes, which are also employed to apply an electrostatic force on the sphere to maintain it at a constant position. The voltage required to maintain the null position of the sphere changes as the local value of gravity changes. Superconducting gravity meters are the most sensitive gravity recording devices available, capable of detecting changes in gravitational acceleration as small as 0.0001 mGal. Because these meters must be operated inside liquid helium dewars, they are not easily moved; generally, they are used to monitor changes in gravity that occur over time at fixed locations. *See* DEWAR FLASK; LIQUID HELIUM; SUPERCONDUCTING DEVICES.

Shipboard gravity meters. Gravity variations are frequently mapped from ships or airplanes. The accelerations encountered by a moving ship are significant compared to the minute variations in gravitational acceleration produced by the suboceanic density structures of interest. By averaging the gravity field over several minutes, however, changes of 1 mGal can be detected. The gravity sensor must be mounted on a stabilized platform to keep it aligned with the gravity vector (local vertical). This is accomplished by means of a double-gimbal motorized platform, the position of which is controlled electronically by using gyroscopes and accelerometers. *See* ACCELEROMETER; GYROSCOPE.

Two types of shipboard gravity systems are in wide use. One is similar to the land meter that uses a diagonally oriented spring (Fig. 2*b*), except that the motion of the beam-mass pair is heavily overdamped. In another commonly used shipboard gravity meter, a wire coil is wound around a pivoted mass held between the poles of a permanent magnet. The current in the coil required to maintain the mass at a null position is proportional to gravity. *See* GRAVITY.

Mark A. Zumberge

Bibliography. O. Francis et al., Calibration of a superconducting gravimeter by comparison with an absolute gravimeter FG5 in Boulder, *Geophys. Res. Lett.*, 25(7):1075–1078, 1998; I. Marson and J. E. Faller, *g*, the acceleration of gravity: Its measurement and its importance, *J. Phys. E: Sci. Instrum.*, 19:22–32, 1986; A. Peters, K. Y. Chung, and S. Chu, Measurement of gravitational acceleration by dropping atoms, *Nature*, 400(6747):849–852, 1999; W. M. Telford, L. P. Geldart, and R. E. Sheriff, *Applied Geophysics*, 2d ed., 1990; W. Torge, *Gravimetry*, 1989.

Graybody

An energy radiator which has a blackbody energy distribution, reduced by a constant factor, throughout the radiation spectrum or within a certain wavelength interval. The designation "gray" has no relation to the visual appearance of a body but only to its similarity in energy distribution to a blackbody. Most metals, for example, have a constant emissivity within the visible region of the spectrum and thus are graybodies in that region. The graybody concept allows the calculation of the total radiation intensity of certain substances by multiplying the total radiated energy (as given by the Stefan-Boltzmann law) by the emissivity. The concept is also quite useful in determining the true temperatures of bodies by measuring the color temperature. For a discussion of the Stefan-Boltzmann law and color temperature *see* HEAT RADIATION; BLACKBODY.

Heinz G. Sell; Peter J. Walsh

Graywacke

A well-indurated dark gray sandstone that is characterized by abundant dark-colored detrital rock fragments and more than 15% clay matrix minerals between sand grains. Graywacke sands were deposited chiefly in marine basins near the edge of continental margins where plate subduction was taking place. Subsequent compressional deformation and uplift of rocks in the sedimentary basins results in the occurrence of most graywackes in Alpine-type (compressional) mountain ranges. *See* CLAY MINERALS; CONTINENTAL MARGIN.

The term graywacke evolved from initially informal usage for sandstones examined as hand specimens in the field to later formal usage for sandstones defined rigorously on microscopic criteria. Graywacke was applied initially to sandstones of Late Devonian to Early Carboniferous age in the Harz Mountains of Germany that are distinctive in that they are nonporous, strongly indurated, clayey, and dark gray. Geologists who had seen the Harz Mountain exposures later applied the term graywacke to sandstones of similar appearance elsewhere in the world. It was not until the 1950s and 1960s that descriptions in English were made of the microscopic texture and mineral composition of graywackes. P. D. Krynine, F. J. Pettijohn, and R. L. Folk provided important descriptions of graywacke and incorporated the term into formal sandstone classifications, but their definitions all differed. Krynine defined graywacke as being composed of more than 25% detrital micas and micaceous rock fragments; Folk defined it as being composed of more than 25% detrital grains derived from (exclusive of micas) metamorphic rocks; and Pettijohn defined it as dark gray, well-indurated sandstone that contained more than 15% detrital clay matrix. All three experts studied the same type of sandstone, but they based their definitions on different criteria. Several additional definitions have been proposed since these studies were completed. Because of the confusion that developed from its use, the term graywacke has been omitted from most formal classification schemes. It is used, however, as a field term referring to sandstone with the characteristic detrital rock fragments and clay matrix minerals.

Mineral composition. Graywackes have a wide range in mineral composition, which reflects the varied source rocks from which the detritus in them was derived. They tend to be quartz-poor (10–50%), to be rich in both feldspar and unstable rock fragments, and to contain several percent of unstable accessory minerals such as micas, pyroxenes, and amphiboles. Feldspathic graywackes (those in which feldspar exceeds rock fragments) are derived chiefly from plutonic cores of denuded island arcs. Lithic graywackes (those in which rock fragments exceed feldspar) are derived either from volcanic island arcs or from sedimentary rocks in adjacent basins that were deformed and uplifted. Volcanic rock fragments characterize the former type of lithic graywackes, whereas sandstone, shale, and their weakly metamorphosed equivalents characterize the latter type.

Matrix origin. The term matrix is applied to mineral particles which are less than 0.001 in. (0.03 mm) in diameter and which are interstitial to coarser, sand-size grains. In graywacke, such matrix minerals are chiefly illite and chlorite. In the 1950s it was believed that matrix was composed of detrital particles deposited simultaneously with the sand grains. The poor sorting reported for many graywackes was interpreted as the result of deposition of the sediment from muddy marine turbidity currents. Subsequent work demonstrated that "matrix clays" are either (1) sand grains of claystone and shale that were mashed between more resistant quartz grains so that evidence of this detrital origin is obscured or obliterated; (2) chemically precipitated cements (such as chlorite, smectite, or other clay species) that formed when water in pores in the sand reacted with detrital grains, chiefly volcanic rock fragments; or (3) clay minerals of origins 1 and 2 that underwent recrystallization upon encountering higher temperatures and pressure upon deep burial.

Depositional origin. Most graywackes were deposited in submarine fans and adjacent basin-plain environments by turbidity currents. They commonly display graded bedding, Bouma sequences, and current-formed and biogenic sole marks. The term Bouma sequence refers to five divisions of a single, ideal turbidity current deposit. Graywackes are interbedded with shale beds that were deposited by dilute turbidity currents and other marine processes. Thicknesses of several miles of interbedded turbidite graywacke and shale accumulated in many basins. Burial and subsequent compressional deformation of these sequences resulted in the generation of clay matrix, loss of porosity, and strong induration. The gray color of the sandstone is derived from rock fragments and organic-stained clay minerals. *See* ARKOSE; SANDSTONE; SEDIMENTARY ROCKS; SHALE; TURBIDITE; TURBIDITY CURRENT. Earle F. McBride

Bibliography. R. L. Folk, *Petrology of Sedimentary Rocks*, 1982; P. D. Krynine, *The Megascopic Study and Field Classification of Sedimentary Rocks*, 1948; F. J. Pettijohn and P. E. Potter, *Sand and Sandstone*, 2d ed., 1987.

Great Basin

A semiarid region of the western United States characterized by internal drainage and a distinctive topography of narrow, elongate, predominantly north-trending mountain ranges with intervening deep valleys. The Great Basin is about 400,000 km^2 (144,000 mi^2) in area and includes most of Nevada and western Utah, as well as parts of eastern California, southeastern Oregon, and southern Idaho (**Fig. 1**). The region is bounded by the Sierra Nevada mountains to the west, the Columbia Plateau to the north, the Wasatch Range and Colorado Plateau to the east, and the Mojave Desert to the south. The province was named by John Fremont, who explored the area in 1843–1845 and recognized that the region is internally drained, with no rivers having outlets to the sea.

Landscape features. The Great Basin forms the northern part of the Basin and Range physiographic province, named for its characteristic topography of narrow fault-block mountain ranges separated by broad flat valleys, many of which form closed basins (**Fig. 2**). Extensional block faulting has resulted in uplift of the mountains relative to the valley floors along range-bounding normal faults. Generally, east-west extension and stretching of the continental crust began about 20 million years ago and is still active in much of the Great Basin.

Faulting and localized volcanism has blocked drainages and contributed to the formation of the internal drainage system that defines the region today. Within the Great Basin, rivers have no outlet to the sea but terminate in local depressions, where they form playas or saline lakes. The longest river is the Humboldt, which flows west and then south across Nevada, eventually terminating in the Humboldt Sink. Major lakes include the Great Salt Lake,

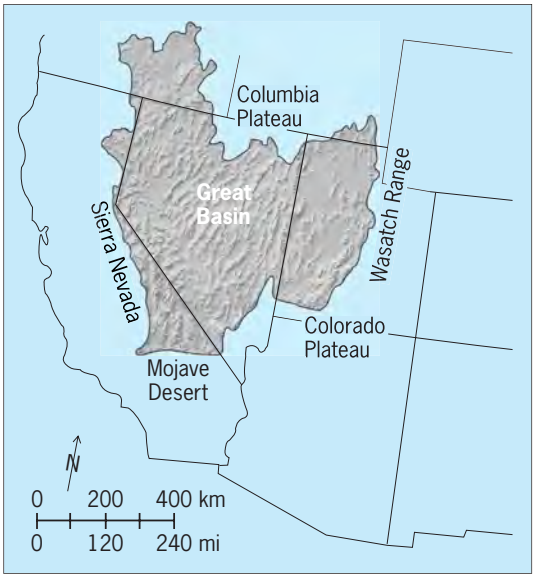

Fig. 1. Location of the Great Basin and bounding topographic features.

Fig. 2. Railroad Valley and the Grant Range, a typical basin and range in eastern Nevada.

fed by runoff from the Wasatch Mountains, and Mono and Pyramid lakes, fed by runoff from the Sierra Nevada. *See* BASIN; DESERT; DESERT EROSION FEATURES; FAULT AND FAULT STRUCTURES; PLATE TECTONICS; PLAYA; VOLCANO.

Climate and vegetation. The Great Basin is generally arid, as rainshadow effects caused by the bounding mountain ranges block the flow of moisture into the region from both the Pacific Ocean to the west and the Gulf of Mexico to the southeast. Precipitation averages between 100 and 300 mm (4 and 12 in.) per year, but is highly variable from location to location and from year to year. Rainfall increases with elevation, and the mountain ranges receive significantly more precipitation than adjacent valleys. The Great Basin is considered a "cold" desert, with mean winter temperatures below 0°C (32°F) and most winter precipitation falling as snow.

Currently, vegetation in the Great Basin is dominated by sagebrush (*Artemisia* spp.), with few trees except for some pinion-juniper woodlands at higher elevations and cottonwoods along perennial streams. The ecology has been significantly altered by the spread of introduced invasive species, including cheatgrass, Russian thistle, and tamarisk that thrive in the disturbed areas created by fires and overgrazing.

Human impact. The Great Basin remains one of the least populated regions of the continental United States. Rapid population growth and urbanization have centered on the cities of Las Vegas, Reno, and Salt Lake City, all located on the borders of the region. Outside of these urban areas, cattle grazing and mining remain the major economic activities. Human environmental impact on the region includes overgrazing and resulting vegetation changes; land disturbance and waste rock from mining operations; lowered water tables and loss of surface stream flows due to heavy use of water resources for irrigation; and the interbasin transfer of water to supply rapidly growing populations in cities such as Las Vegas and Los Angeles.

David C. Greene

Bibliography. J. Hudson, *Across This land: A Regional Geography of the United States and Canada*, Johns Hopkins University Press, 2002; J. Laity, Desert environments, in A. Orme (ed.), *The Physical Geography of North America*, pp. 380–401, Oxford University Press, 2002; S. Trimble, *The Sagebrush Ocean: A Natural History of the Great Basin*, University of Nevada Press, 1989.

Great circle, terrestrial

A circle or near-circle representing a trace on the Earth's surface of a plane that passes through the center of the Earth and divides it into equal halves (**Fig. 1**). The Equator is a great circle, the trace of the plane that bisects and is perpendicular to the Earth's axis. Planes through the Poles cut the Earth along meridians. All meridians are great circles; actually, they are not quite circular because of the slightly flattened Earth. The equatorial diameter is 1.0034 times the size of the polar diameter. All parallels other than the Equator are called small circles, being smaller than a great circle.

The shortest distance between two points on the Earth's surface follows a great circle arc. It is therefore advantageous for commercial airlines to plan long-distance, nonstop flights along great circle routes called orthodromes. Also important in navigation are rhumb lines or loxodromes, which are lines of constant compass directions. Except where they coincide with meridians or the Equator, loxodromes are not great circle arcs.

By using two map projections—the gnomonic and the Mercator—one can navigate by compass and follow great circle routes. The gnomonic is the only map projection on which any straight line is a great circle arc. The Mercator projection shows rhumb lines as straight lines. Navigators plot the origin and destination of trips on a gnomonic projection to obtain shortest routes. They then select from the gnomonic projection points along a route for transfer to a Mercator projection. A series of line segments or rhumb lines connecting points on the Mercator projection thus approximate the shortest route. Using the charted route on the Mercator projection, navigators can follow the compass bearing except when changing it is required at intermediate points.

Although sophisticated instruments have replaced the compass as a navigational tool, intercontinental air routes are still laid out as nearly as possible to true great circles. For example, a trip from Seattle to Tokyo passes near Anchorage, Alaska, and the Aleutian Islands, and a direct flight from Seattle to Amsterdam crosses Canada's Northwest Territories, Iceland, and Scotland. Both are great circle routes.

Despite its desirable property of showing great circle arcs as straight lines, the gnomonic projection cannot show an entire hemisphere and greatly exaggerates shapes and sizes away from its center. Moreover, direct distance measures on the projection are unreliable.

Two common methods can be used to calculate the distance of a great circle arc. One method uses trigonometric functions: $\cos D = \sin a \sin b + \cos a \cos b \cos c$. Here D is the arc distance between points A and B in degrees, a is the latitude of A, b is the latitude of B, and c is the difference in longitude between A and B. After D is calculated, it can be converted to a linear distance measure by multiplying D by the length of one degree of the Equator, which is 111.32 km or 69.17 mi.

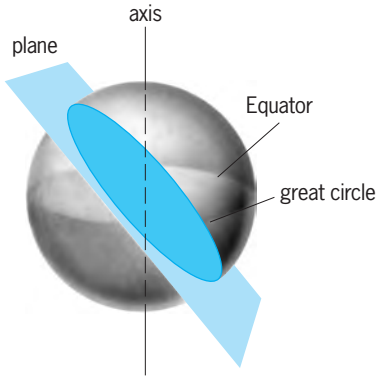

Fig. 1. Diagram of a great circle described by a plane through the center of the Earth.

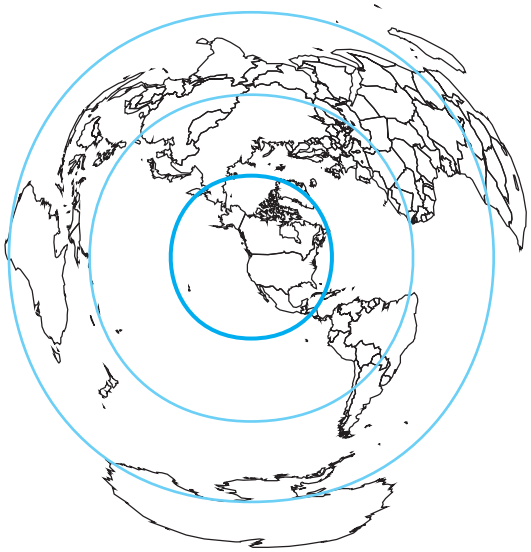

Fig. 2. Azimuthal equidistant projection centered at Seattle, Washington. Circles away from Seattle have an interval of 5000 km (3120 mi).

The other method uses the azimuthal equidistant projection (**Fig. 2**). Unlike the gnomonic projection, the azimuthal equidistant projection can be centered at any point on the Earth's surface and can show the entire sphere. More importantly, a straight line from the center of the projection to any other point is a great circle route and the distances are at a comparable (consistent) scale between the two points. The azimuthal equidistant projection is therefore useful in showing any movement directed toward or away from a center, such as seismic waves, radio transmissions, missiles, and aircraft flights. Kang-tsung Chang

Bibliography. B. D. Dent, *Cartography: Thematic Map Design*, 5th ed., William C. Brown Publishers, 1998; A. H. Robinson et al., *Elements of Cartography*, 6th ed., John Wiley, 1995.

Green chemistry

The discovery, development, and application of sustainable chemical reactions that create no waste or generate only benign wastes. The idea that toxic chemicals, once created, can be contained or prevented from entering the environment is unrealistic. Since some portion of the toxins will always enter the environment, whether through use, processing, disposal, or unforeseen circumstances such as accidents, terrorism, or natural disasters, green chemistry solutions are applied at the molecular level to avoid creating toxins in the first place. Using green chemistry to prevent pollution provides a more efficient approach to preserving human health and the environment than waste processing or remediation.

Developing practical approaches for sustainable manufacturing is one of the major goals of green

chemistry. Green chemistry faces many challenges, including social, business, political, and technological barriers. Even within the field of chemistry, a major challenge exists to shift the focus of mainstream chemical research away from just developing compounds and processes to solving short-term (1–5 years) problems of practical and economic importance. Green chemistry is acting now to ameliorate the long-term environmental impact (50–200 years) of chemical technologies through molecular-level design, acting in concert with green principles and a sustainable vision of the future.

Roots of green chemistry. The 12 principles of green chemistry put forth by P. T. Anastas and J. C. Warner provide an effective framework for developing greener methods. In summary, the principles are (1) prevent waste before it is created; (2) be atom-economical; (3) use benign substances; (4) make benign products; (5) use less solvents; (6) use less energy; (7) use renewables; (8) avoid protecting groups; (9) use catalysts; (10) make things biodegradable; (11) analyze in real time; (12) be safe.

When every atom in the reactants appears in the products, a maximum of atom economy is reached. A similar concept is the efficiency or E factor, which is equal to the mass of inputs for a reaction process (including solvents and purification steps) divided by the mass of useful products obtained from a reaction. The lowest possible E factor is 1. Most industrial processes have E factors ranging from about 3–5 for high-volume chemicals such as acetic acid or monomers/polymers and 20–100 for high-value chemicals such as pharmaceuticals or semiconductors. Improving atom economy or lowering the E factor simultaneously saves money and protects the environment, unless doing so substantially increases the energy requirements for the reaction.

Examples. Two representative examples of green chemistry are oxidative catalysis and alternative solvents.

Oxidative catalysis. Few oxidants are simultaneously inexpensive and reactive enough to be commercially viable. **Figure 1** shows the main commercial oxidants and ranks them qualitatively in terms of oxidizing power versus greenness. In order to replace environmentally undesirable oxidizing systems such as heavy-metal ions or active chlorine, switching to oxygen-based oxidants is not enough. Additional activation is necessary, preferably catalytic activation with a benign metal ion via an enzymatic or synthetic catalyst that satisfies green principles.

T. J. Collins and coworkers have developed a viable green hydrogen peroxide (H_2O_2)-activating catalyst using an iron (Fe) complex of an oxidation-resistant tetra amide macrocyclic ligand (TAML). FeTAML complexes are constructed from benign components and activate H_2O_2 in water over a wide pH range. Long-lived under oxidizing conditions, FeTAML complexes efficiently bleach lignins, dyes, or chlorophenols at very low catalyst concentrations. The main commercial application of FeTAML complexes is for totally chlorine-free (TCF)

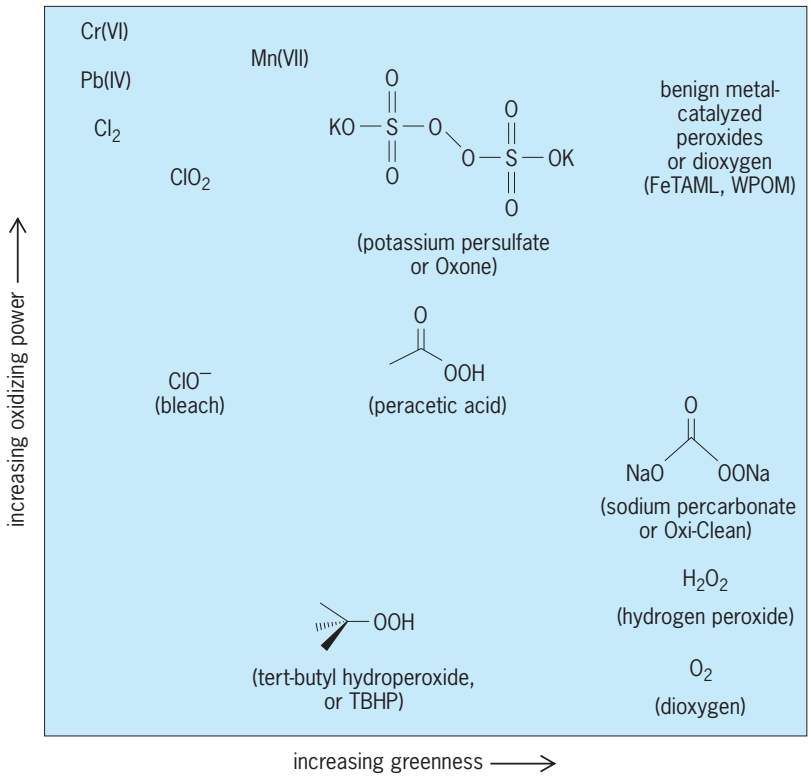

Fig. 1. Comparison of some common oxidants in terms of oxidizing power versus greenness.

paper bleaching. TCF paper bleaching avoids the formation of dioxins and absorbable organic halogens, unlike paper bleaching with chlorine dioxide (ClO_2).

N. Mizuno and coworkers have developed a tungsten polyoxometallate (WPOM) catalyst that activates hydrogen peroxide to perform selective olefin (double-bond) oxidation reactions. The catalyst uses hydrogen peroxide efficiently and exhibits very high selectivities, useful features for eventual commercial development as an alternative to the use of lead-, chromium-, manganese-, or organic peroxide-based oxidants. Molybdenum polyoxometallate (MoPOM) complexes, developed by C. Hill and others, activate oxygen (O_2) for oxidizing a variety of different substrates, including lignin bleaching for TCF paper manufacture and in-situ detoxification of chemical warfare agents. *See* CATALYSIS; HYDROGEN PEROXIDE; OXIDATION PROCESS.

VOC alternatives. Most solvents used today are volatile organic compounds (VOCs). VOCs readily escape to the atmosphere when used, causing a substantial fraction of all air pollution. Eliminating VOCs is environmentally desirable but requires that practical and economical VOC solvent alternatives be developed.

Supercritical water and supercritical carbon dioxide (CO_2) continue to provide successful green approaches for replacing VOCs in chemical processes such as decaffeinating coffee, dry cleaning, and demanding chemical reactions.

R. Rogers and others have developed a new class of solvents called room-temperature ionic liquids (RTILs). Many RTILs are based on chloroaluminate

anions or alkyl imidazolium cations. Most RTILs exhibit a low melting point, a high boiling point, and a high viscosity. As solvents, RTILs have extremely low vapor pressures, an important green feature for replacing VOCs and decreasing atmospheric pollution. Many of the chemical properties of ionic liquids such as tunable polarity, good dissolving power for organic molecules, and easy drying are identical to those of VOCs. The major hurdles for commercialization of RTILs are cost, limited toxicological data, and difficulties associated with removing solutes from the RTIL to allow the ionic liquid to be recycled or reused.

C. J. Li and coworkers have developed an arsenal of useful synthetic organic reactions that take place in water instead of VOCs. Controlling chemical reactions in water is an important green objective; however, care must be taken that water-soluble reagents do not leak out of the chemical process and become widely dispersed in the environment.

IBM has pioneered a technique that uses soap and water to degrease silicon wafers, thereby avoiding the use of chlorinated fluorocarbons (CFCs, or Freons) during semiconductor manufacturing. Another CFC-free wafer-cleaning technology has been developed by J. DeSimone that uses surfactants and supercritical CO_2 as the solvent. *See* IONIC LIQUIDS; SOLVENT; SUPERCRITICAL FLUID.

New approaches. Exciting developments in green chemistry are coming from new ideas and research projects that use the green framework as a starting point instead of trying to retrofit green solutions onto existing technologies.

Supramolecular synthesis. J. C. Warner and coworkers have used supramolecular chemistry (which aims to generate complex chemical systems from components bound together by noncovalent intermolecular forces) to challenge the fundamental idea of forming covalent bonds as a means of controlling chemical reaction pathways. When an undesirable covalent bond is formed, it is difficult to cleave and often leads to the production of chemical waste or to the use of high-energy conditions to try to control the process of covalent bond formation. Supramolecular interactions are much weaker than covalent bonds and therefore more reversible. Warner has used weak supramolecular interactions to direct the formation of titanium dioxide (TiO_2) at room temperature, thereby avoiding the use of temperatures as high as 1000°C (1832°F), to synthesize the next generation of dye-sensitized TiO_2 solar photovoltaic cells. In contrast, current methods for manufacturing solar cells are very energy intensive, thus rendering solar technology expensive and compromising most, if not all, of the benefit derived from obtaining energy from the sun. Currently, most energy used in manufacturing processes is derived from combustion, including the synthesis of solar cells. *See* SOLAR CELL; SUPRAMOLECULAR CHEMISTRY; TITANIUM OXIDES.

Copper complexes. In the United States, most exterior wood is pressure-treated with chromated copper arsenate (CCA). Pressure-treated wood is green in color because of the inclusion of copper, but its antifungal and anti-insect properties are typically attributed to the use of chromium and arsenic. CCA-treated wood is used for playground lumber, exposing children to hexavalent chromium or arsenic. In addition, as the CCA leaches out of the treated wood into waterways, background levels of chromium and arsenic in drinking water are increased. Quaternary alkylammonium copper complexes have been developed by Chemical Specialties Inc. for pressure-treating wood that exhibit lifetime, antifungal, and anti-insect properties comparable to CCA-treated wood, eliminating the need for chromium or arsenic. This technology provides a tremendous step forward using some relatively simple green chemistry. *See* COPPER.

Poly(lactic acid). P. Gruber at Cargill-Dow has applied a green vision to establish a large-scale poly(lactic acid) plant that makes plastic from corn. The process depends on a renewable feedstock (corn) instead of an exhaustible one (petroleum). Variation in the chemical and physical properties of the resulting plastics is achieved by controlling the separation and chemical modification steps used on the corn feedstock. The process steps have been designed to be as green as possible, including incentives for the corn suppliers to use sustainable agricultural practices.

Education. Education plays a key role in green chemistry. How people learn chemistry affects how they will solve real-world problems. If chemical education fails to incorporate green principles, the next generation of chemists will continue with waste-generating practices. **Figure 2** compares a conventional chemical laboratory experiment with one designed to incorporate green principles. The new, greener procedure preserves all of the pedagogical elements of the older laboratory, but also teaches the powerful lesson that green chemistry

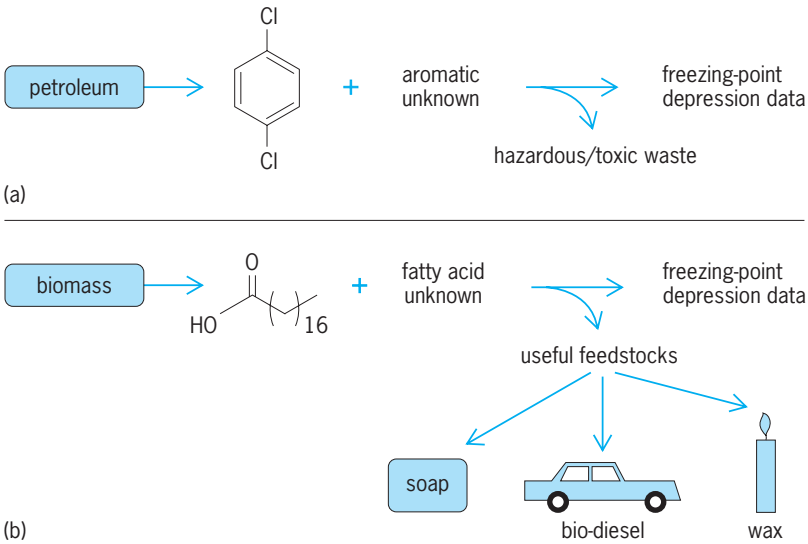

Fig. 2. Comparison of laboratory experiments for measuring colligative properties (which depend on the number of atoms or molecules but not their nature). (*a*) Old laboratory approach uses an exhaustible feedstock, uses hazardous reagents, creates toxic waste, and has an E factor of ∞. (*b*) New greener laboratory approach uses a renewable feedstock, uses benign reagents, creates almost zero waste, and approaches an E factor of 1.

can create knowledge and compounds without harming the natural world or creating toxic waste.

S. W. Gordon-Wylie

Bibliography. P. T. Anastas and J. C. Warner, *Green Chemistry: Theory and Practice*, Oxford University Press, 1998; M. C. Cann and M. E. Connelly, *Real World Cases in Green Chemistry*, American Chemical Society, 2000; S. S. Gupta et al., Rapid total destruction of chlorophenol pollutants by activated hydrogen peroxide, *Science*, 296:326–328, 2002; K. Kamata et al., Efficient epoxidation of olefins with ≥99% selectivity and use of hydrogen peroxide, *Science*, 300:964–966, 2003; R. J. Lempert et al., *Next Generation Environmental Technologies: Benefits and Barriers*, RAND Science and Technology Policy Institute, 2003; W. M. Nelson, *Green Solvents for Chemistry: Perspectives and Practice*, Oxford University Press, 2003; R. D. Rogers and K. R. Seddon, *Ionic Liquids: Industrial Applications for Green Chemistry*, American Chemical Society, 2002; B. M. Trost, The atom economy: A search for synthetic efficiency, *Science*, 254:1471–1477, 1991.

Green fluorescent protein

A protein produced by the bioluminescent jellyfish *Aequorea victoria* that fluoresces in the lower green portion of the visible spectrum. The gene for green fluorescent protein (GFP) has been isolated, and the protein has been extensively characterized biochemically and found useful in biological research. After synthesis of GFP, three of its amino acids undergo a self-catalyzed reaction that produces an internal fluorophore (fluorescent molecule). The protein does not require a specific cellular environment to become fluorescent and can be artificially expressed in cell types as diverse as bacterial, plant, and animal cells. It can be attached to other proteins of interest using recombinant DNA techniques, making it possible to easily trace the synthesis, location, and movement of these proteins in single living cells using conventional fluorescence microscopy. *See* BIOLUMINESCENCE; FLUORESCENCE; FLUORESCENCE MICROSCOPY; HYDROIDA.

Biological role. The natural function of the GFP in the jellyfish is to convert the blue light emitted by a calcium-regulated protein called aequorin into green light. When the animal is disturbed (for example, when it is touched lightly), calcium stores inside some of its cells are released, causing excitation of the aequorin. The excitation energy is then transferred to GFP (probably by a quantum-mechanical process) and released as a brief green glow. The biological significance of this process is still not known. When expressed alone, the protein (in the absence of aequorin) is fluorescent and emits green light when excited with ultraviolet light, and to some extent when excited with blue light.

Tracking and quantitation. Green fluorescent protein technology, by allowing observations of processes in living cells over time, has begun to provide new biological insights. GFP can be attached to a protein marker for specific organelles, such as mitochondria or the Golgi apparatus, allowing the organelle's location and morphology to be followed over time. Changes in organelle structure in response to drug treatments, or natural processes such as differentiation or mitosis can be studied.

Green fluorescent protein tagging also makes it possible to follow the trafficking of proteins between different parts of the cell or between different organelles (via video microscopy). Moreover, the exact levels of the tagged protein in different organelles can be quantitated, since the intensity of light emission from GFP is directly proportional to the number of molecules. It is possible to quantitate the absolute number of GFP molecules in an organelle of interest. *See* VIDEO MICROSCOPY.

A laser scanning confocal microscope, which scans a laser over the sample, can be used for imaging GFP within cells or within small organisms, and unlike conventional microscopes it can map the location of GFP in three dimensions. GFP, like other fluorophores, can be photobleached (rendered nonfluorescent) when exposed to intense light. The laser in a confocal microscope provides an ideal tool for selective bleaching of GFP in a particular region of an organelle or of an entire organelle. Recovery by diffusion of GFP into the bleached area or by transport from other organelles can be monitored, providing important information about the mobility of the protein not previously available. This technique, known as fluorescence recovery after photobleaching (FRAP), has been utilized to show that resident proteins of the Golgi apparatus are freely mobile (rather than anchored to fixed positions), and to track movement of proteins between organelles. *See* CONFOCAL MICROSCOPY; CYTOCHEMISTRY.

Other uses. Green fluorescent protein can also be placed under the control of promoters (regulatory DNA sequences) that allow its selective expression only in specific cell types, for example, neurons. This has proved highly useful in studies of developmental biology for examining the differentiation and migration of a wide variety of cell types in development, particularly in small, relatively transparent organisms such as *Drosophila* embryos, nematodes, and zebrafish. It is also possible to do genetic studies of the promoter sequences themselves (for example, by examining the effects on GFP expression of changes of the DNA sequence at different regions of the promoters) to better determine which specific DNA sequences within the promoter contribute to its capacity to regulate expression of a linked gene. *See* DEOXYRIBONUCLEIC ACID (DNA).

Cells expressing GFP-tagged proteins can be separated from nonexpressing cells or cells expressing different levels of GFP using devices called cell sorters. Cells are sent one at a time past a laser beam set at a frequency that excites fluorescent molecules. The cells can then be scored for fluorescence and deposited into distinct pools. One application is identification of transfected cells after foreign DNA coding for a gene of interest and also for GFP is introduced. The GFP identifies the transfected cells, and

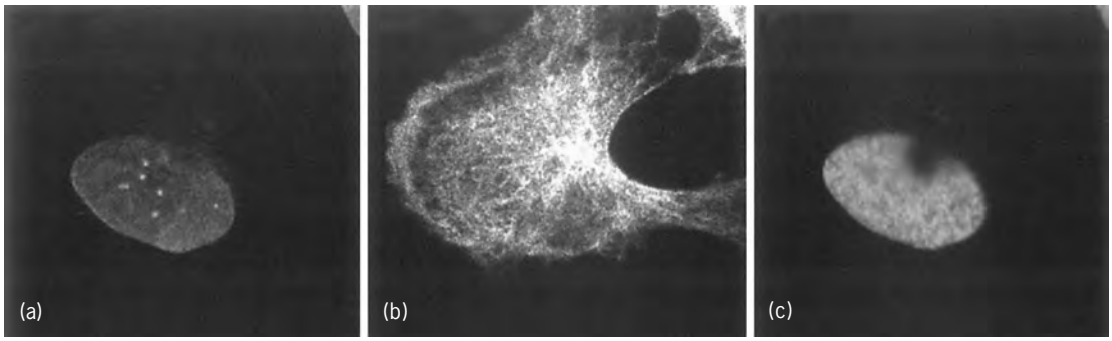

Labeling of multiple organelles in the same cell with spectrally distinguishable green fluorescent protein variants. (*a*) Lamin-binding receptor (a protein found in the nuclear envelope) tagged with CFP. (*b*) Microtubules containing tubulin tagged with YFP. (*c*) Histone 2b (a protein that binds to DNA in the nucleus) tagged with RFP. (*Jan Ellenberg, European Molecular Biology Laboratory*)

the cell sorter can then provide a pure population of them.

Spectral variants. New improved variants of the protein have been devised that are brighter than the original and show a preference for blue light (which is nontoxic to cells) rather than ultraviolet (which can rapidly kill cells being observed). Also, variants have been engineered which excite and emit at different wavelengths, allowing tagging and following two or more different proteins in the same cell. Two fluorescent proteins, a cyan-emitting variety (CFP) and a yellow-green-emitting variety (YFP), have proved particularly useful in double-label experiments since they can be visualized separately in a fluorescence microscope. Recently, a related red fluorescent protein (RFP) has been identified from other coelenterates, so it is now possible to tag three different proteins in the same cell or in a cell population (see **illustration**). John Presley

Bibliography. M. Chalfie and S. R. Kain (eds.), *Green Fluorescent Protein: Properties, Applications and Protocols*, 2d ed., Wiley, New York, 2005; J. Ellenberg, J. Lippincott-Schwartz, and J. F. Presley, Dual-color imaging with GFP variants, *Trends Cell Biol.*, 9:52–56, 1999; B. W. Hicks, *Green Fluorescent Protein: Applications & Protocols (Methods in Molecular Biology)*, Humana Press, Totowa, NJ, 2002; J. R. Lakowicz and C. D. Geddes, *Green Fluorescent Protein (Topics in Fluorescence Spectroscopy)*, Springer, New York, 2006; J. F. Presley et al., ER-to-Golgi transport visualized in living cells, *Nature*, 389:81–85, 1997; K. F. Sullivan and S. A. Kay, *Green Fluorescent Proteins*, Academic Press, New York, 1999.

Greenhouse effect

The ability of a planetary atmosphere to inhibit heat loss from the planet's surface, thereby enhancing the surface warming that is produced by the absorption of solar radiation. For the greenhouse effect to work efficiently, the planet's atmosphere must be relatively transparent to sunlight at visible wavelengths so that significant amounts of solar radiation can penetrate to the ground. Also, the atmosphere must be opaque at thermal wavelengths to prevent thermal radiation emitted by the ground from escaping directly to space. The principle is similar to a thermal blanket, which also limits heat loss by conduction and convection. In recent decades the term has also become associated with the issues of global warming and climate change induced by human activity. *See* ATMOSPHERE; SOLAR RADIATION.

Of the terrestrial planets, Venus has by far the strongest greenhouse effect. Though only about 1% of the incident solar radiation penetrates to the ground, the thermal opacity of Venus's atmosphere, which is 100 times more massive than the Earth's and composed mostly of carbon dioxide, is exceedingly large. As a result, the trapped solar radiation on Venus generates a surface temperature of nearly 460°C (860°F), which is hot enough to melt lead and vaporize mercury. This is 500°C (900°F) hotter than the surface would be if it were simply in thermal equilibrium with the global mean solar energy absorbed by Venus (−40°C, −40°F), without benefit of the greenhouse effect. For the Earth, the strength of the greenhouse warming effect is a more modest 33°C (60°F), due primarily to the thermal opacity of atmospheric water vapor, clouds, and carbon dioxide. The trapped sunlight makes the global mean surface temperature of the Earth a relatively comfortable 15°C (59°F) instead of an otherwise frigid −18°C (−1°F). On Mars, the very tenuous carbon dioxide atmosphere can muster only a few degrees (≈3°C, 5°F) of greenhouse warming. Meanwhile, on Mercury, which has no tangible atmosphere, there is no greenhouse effect at all. Mercury's surface temperature is determined solely by local thermal equilibrium with the absorbed sunlight. *See* MARS; MERCURY (PLANET); VENUS.

Mechanism. Basic understanding of the greenhouse effect dates back to the 1820s, when the French mathematician and physicist Joseph Fourier performed experiments on atmospheric heat flow and pondered the question of how the Earth stays warm enough for plant and animal life to thrive; and to the 1860s, when the Irish physicist John Tyndall demonstrated by means of quantitative spectroscopy that common atmospheric trace gases, such as water vapor, ozone, and carbon dioxide,

are strong absorbers and emitters of thermal radiant energy but are transparent to visible sunlight. It was clear to Tyndall that water vapor was the strongest absorber of thermal radiation and, therefore, the most influential atmospheric gas controlling the Earth's surface temperature. The principal components of air, nitrogen and oxygen, were found to be radiatively inactive, serving instead as the atmospheric framework where water vapor and carbon dioxide can exert their influence. Based on his understanding of the radiative properties of absorptive gases in the atmosphere, Tyndall speculated in 1861 that a reduction in atmospheric carbon dioxide could induce an ice age climate.

Relying on the work of Tyndall and on careful measurements of heat transmission through the atmosphere compiled by the American astronomer Samuel Langley, the Swedish chemist and physicist Svante Arrhenius was the first to develop a quantitative mathematical model of the Earth's greenhouse effect. In 1896 he published his heat-balance calculations of the Earth's sensitivity to carbon dioxide change. Given that Arrhenius's basic interest was to explain the likely causes of ice age climate, his greenhouse model was successful. He showed that reducing atmospheric carbon dioxide by a third would cool the global surface temperatures by −3°C (−5.5°F), and that doubling carbon dioxide would cause the tropical latitudes to warm by 5°C (9°F), with somewhat larger warming in polar regions. These results are in remarkably close agreement with current expectations for global climate change in response to carbon dioxide forcing. *See* SPECTROSCOPY.

Large seasonal and epochal (ice age) variability is evident in the terrestrial climate record. This implies that the greenhouse effect on Earth operates rather differently from the static, or at least slowly evolving, carbon dioxide greenhouses on Venus and Mars. Partly, this is the result of changes in the atmospheric concentration of carbon dioxide over geological time scales. More importantly, it is due to the dominant strength of water vapor as the largest contributor to the terrestrial greenhouse effect, and the fact that the amount of atmospheric water vapor is strongly dependent on temperature. Because of this, and also because condensing water vapor produces clouds which typically cool the Earth, there is strong coupling and interaction between the amount of atmospheric water vapor and temperature.

This aspect of water vapor behavior was noted by the American geologist Thomas Chamberlin who, in 1905, described the greenhouse contribution by water vapor as a positive feedback mechanism. Surface heating due to another agent, such as carbon dioxide or solar radiation, raises the surface temperature and evaporates more water vapor which, in turn, produces additional heating and further evaporation. When the heat source is taken away, excess water vapor precipitates from the atmosphere, reducing its contribution to the greenhouse effect to produce further cooling. This feedback interaction converges and, in the process, achieves a significantly larger temperature change than would be the case if the amount of atmospheric water vapor had remained constant. The net result is that carbon dioxide becomes the controlling factor of long-term change in the terrestrial greenhouse effect, but the resulting change in temperature is magnified by the positive feedback action of water vapor.

It is believed that in the early stages of the solar system, when the Sun had only about 75% of its present luminosity, Venus may also have had an ocean and an atmosphere much like that of the Earth. However, in its early days Venus was apparently unable to sequester its store of carbon in the form of limestone, as did the Earth. As a result, carbon dioxide emitted from volcanic eruptions continued to accumulate in the Venusian atmosphere, increasing its greenhouse effect and evaporating more and more water vapor into the atmosphere to further magnify the growing strength of its greenhouse. Eventually all ocean water was vaporized, and the water vapor was carried high into the atmosphere where most of it was photolyzed and lost to space. Perhaps the climate disaster on Venus can be attributed to the lack of appropriate life forms that might have limited the buildup of carbon dioxide. In any case, Venus stands as an example of a runaway greenhouse effect that has rendered the planet inhospitable to life.

Besides water vapor, many other feedback mechanisms operate in the Earth's climate system and impact the sensitivity of the climate response to an applied radiative forcing. Determining the relative strengths of feedback interactions between clouds, aerosols, snow, ice, and vegetation, including the effects of energy exchange between the atmosphere and ocean, is an actively pursued research topic in current climate modeling. Current best estimates from climate modeling results show that without feedback effects, a doubling of atmospheric carbon dioxide will raise the global surface temperature by 1.2–1.3°C (2.2–2.4°F), but that the net feedback magnification due to water vapor, snow/ice melting, and cloud changes magnifies the no-feedback result by approximately a factor of 3. Interestingly, carbon dioxide accounts for approximately 7°C (13°F) of the total 33°C (60°F) terrestrial greenhouse effect, and the other nonvolatile greenhouse gases such as ozone, methane, nitrous oxide, and anthropogenic chlorofluorocarbons add another 3–4°C (5–7°F) of greenhouse strength. In this context, the nonvolatile greenhouse gases can be thought of as providing the core forcing for the Earth's greenhouse effect. The larger greenhouse contribution by water vapor and clouds may then be considered to represent the feedback response that magnifies the core forcing by the atmospheric feedback factor of 3. This distinction helps to identify more clearly the causes and effects in an otherwise highly interactive climate system. *See* CLIMATE MODIFICATION.

The modern approach to studying Earth's climate and the human impact on climate change began in

the 1930s with the work of Guy Callendar, a British engineer who systematically documented the time trend of anthropogenic fossil fuel use and linked it to the corresponding increase in atmospheric carbon dioxide. Having had available to him the precise spectroscopic measurements of the absorption characteristics of water vapor, carbon dioxide, and other heat-absorbing gases, Callendar could link his estimates of carbon dioxide increase with the 0.4°C (0.7°F) temperature increase recorded in the 50 years prior to 1950. In 1958 Charles Keeling, a research chemist at Scripps Institute in California, began making high-precision measurements of carbon dioxide accumulation in the atmosphere. This ushered in a new era of precise measurement and documentation of the atmospheric increases of carbon dioxide and other greenhouse gases such as methane and nitrous oxide. The new measurement techniques were subsequently applied to measuring the precise gaseous composition of air bubbles trapped in glacial ice, thereby extending knowledge of atmospheric composition over time scales going back to the ice ages. The measurements show that atmospheric concentrations of carbon dioxide increased from a preindustrial 280 parts per million in 1850 to the present value of 370 ppm. In geological context, during the last ice age when the global temperature was 5°C (9°F) colder, atmospheric carbon dioxide was at levels near 150 ppm.

Given accurate knowledge of the atmospheric concentration of greenhouse gases, their impact on the strength of the atmospheric greenhouse effect can be accurately calculated using the extensive spectroscopic databases that are available now. Inasmuch as the term "greenhouse effect" refers to determining the radiative efficiency of trapping solar heat at the ground surface by the absorption and emission of thermal radiation, a moderate computational effort can provide the answer. Insofar as human activity induces global warming and climate change, it is necessary to incorporate the uncertainties that are associated with atmospheric feedback processes and with atmosphere/ocean dynamical interactions that contribute to greenhouse efficiency. At present, it is necessary to rely on general circulation climate models to determine the feedback contribution (due mostly to clouds and water vapor) to the total atmospheric greenhouse effect, since remote sensing measurements capable of measuring feedback-related water vapor and cloud changes are not currently available. *See* ATMOSPHERIC GENERAL CIRCULATION; CLIMATE MODELING; REMOTE SENSING.

Human impact. Both Arrhenius and Callendar considered the projected global temperature increases due to anthropogenic carbon dioxide buildup as likely to be beneficial in warming the cold regions of the Earth. But human-induced contributions to climate change are not necessarily benign. For example, discovery of an ozone hole in the Antarctic in the early 1980s pushed the world community to the 1987 Montreal agreement to phase out production of chlorofluorocarbons to reverse ozone destruction in the stratosphere caused by these compounds. As a by-product, this also helped to reduce the rate of increase of the anthropogenic greenhouse forcing. In late 1997 the world community met in Kyoto to consider what could be done to help stabilize and curtail the steady increase in fossil fuel use and the resulting increase in the strength of greenhouse forcing. There are good reasons to be concerned: (1) The projected global warming is much larger than anything previously experienced in human history. (2) The climate response to a globally uniform forcing is not necessarily uniform and may include large regional fluctuations. (3) Extreme events, both droughts and floods, are more likely with the stronger hydrological cycle of a warmer Earth. (4) The inevitable rise in sea level will put low-lying coastal areas at increased risk. (5) Unanticipated changes in global climate could occur if the climate system were to exceed some critical threshold that is beyond the modeling capability of current climate models. *See* STRATOSPHERE.

Andrew A. Lacis

Bibliography. G. E. Christianson, *Greenhouse*, 1999; J. R. Fleming, *Historical Perspectives on Climate Change*, 1998; J. T. Houghton et al., *IPCC WGI 1995: Climate Change 1995*.

Greenhouse technology

Low tunnels, high tunnels, and greenhouses are structures used to grow plants under protected conditions. The progression of terms shows the level (low to high) of technical sophistication in the plant-growing systems. Low tunnels, also called row covers, primarily advance the growing season for outdoor crops (for example, tomatoes, melons, strawberries, and sweet corn). Low tunnels are created using long, narrow strips of transparent plastic material (often polyethylene) buried in the ground along their outer edges to cover one or several adjacent rows of plants grown directly in the soil. French cloches and traditional hot caps are variations of this technology. High tunnels are large versions of low tunnels, raised sufficiently above the ground that people can walk within them.

Types of greenhouses. Greenhouses (or glasshouses) are relatively permanent structures (glass or plastic with aluminum or steel frames) equipped with several means of environmental modification. Free-standing greenhouses are the most basic structural type; they are typically 7–10 m (23–33 ft) wide and 5–10 times as long as wide. Cross-sectional shapes can be classified as arch, hoop, or gable (see **illus.**). The hoop shape is common for high tunnels. Multispan greenhouses are typically connected by a series of roof gutters to create a single air space. Large multispan greenhouses can cover several hectares under one roof, and they are the design of choice for larger commercial greenhouse operators. Eave height continues to increase, and modern structures 5 m (16 ft) tall at the eaves are not uncommon. The greater height improves environmental control, particularly ventilation.

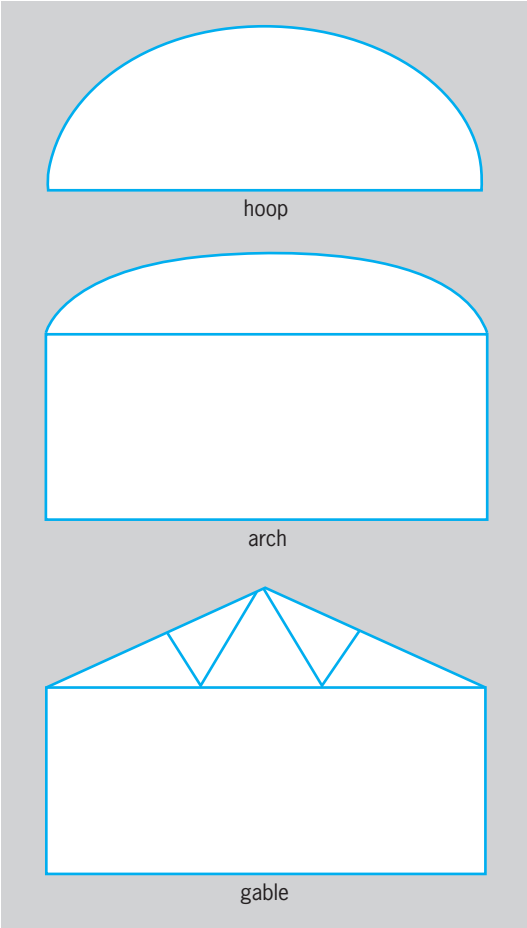

Common commercial greenhouse shapes.

Construction. Aluminum, steel, wood, and concrete have been used to build greenhouse frames, but steel and aluminum are most common. The most prevalent shape has a straight sidewall and a gable roof, although arched roof designs are frequently seen (see illus.). The greenhouse can be a single free-standing unit, or multiple units can be joined at the eaves to cover several hectares. Floors are frequently made of concrete, although gravel floors with concrete walkways may be used to reduce cost.

Light transmittance is important when selecting a covering material. Glass provides the most light to the plants and retains its light transmittance; however, various rigid and film plastic glazing materials are used because of initial lower costs. Plastic film glazing provides a greater ratio of diffused to direct light, which may be of some benefit for plant growth. *See* GLASS.

Recommended greenhouse orientation depends on latitude. If the latitude is greater than 35°, an east-west orientation of the ridge line is preferred to admit more natural light in winter when the Sun does not rise far above the horizon, even at midday. If the latitude is less than 35°, the preferred orientation is north-south to reduce midday solar heating and prevent shade lines caused by long structural elements oriented east-west from casting shadows that remain in one place for a long time, which retards plant growth in the shadow areas.

Environment control. Environment control typically encompasses air temperature, supplemental light, air movement (circulation and mixing), and carbon dioxide concentration. Some degree of relative humidity control may also be included. Natural ventilation may be used to provide less stringent control, but mechanical ventilation is used when more precise temperature control is needed and evaporative cooling is required during hot weather.

Integrated control by computer provides the flexibility of zoned control of each environmental parameter without conflicting control signals (for example, ventilating while supplementing carbon dioxide). Most modern greenhouse facilities have installed environment control computers. Additionally, computerized control usually provides data summaries useful for management decisions. Continuing developments in sensor technology will undoubtedly lead to increasingly sophisticated control of aerial and root environments and fault detection capabilities.

Heating. The function of a greenhouse is to grow plants; structural insulation opportunities are minimal and heat requirements are high in comparison to most other types of buildings. Efforts to conserve heat, such as insulating the north wall and part of the north roof, have often shown negative benefits by reducing natural light and degrading plant growth and quality. Movable, horizontal, indoor curtain systems (which also double as movable shade systems) can save approximately one-quarter of the yearly heat in a cold climate.

Greenhouses can be heated by oil or natural gas. Electric heat is uneconomical. Large greenhouses may be heated using heavier industrial heating oils; propane and wood have been used in specialized situations. Heat delivery is either hydronic (hot water) or by steam. In some cases, particularly in smaller greenhouses, hot air systems can be used if care is taken not to direct the hot air directly onto the plants. No matter what the heat source, care is required to prevent incompletely burned flue gases from entering the greenhouse and causing plant damage (phytotoxicity). *See* HEAT.

Cooling. Solar loads in greenhouses are so great that mechanical cooling, as by air conditioning, is prohibitively expensive. Options for greenhouse cooling are thus limited. If natural light is generally high, simple shading (preferably a movable shade, but it can also be semipermanent) can cool a greenhouse several degrees. The most typical cooling mechanism is ventilation, either natural or mechanical. Natural ventilation is not controlled, but mechanical ventilation is generally designed to keep the temperature rise of the ventilation air to no more than 4–5°C (7–9°F). This may require an air exchange rate as high as one total greenhouse volume each minute. The next step of cooling is to use evaporative means. Evaporative cooling can be used in all but the most hot and humid climates to reduce midday temperature extremes. Cooling can be obtained by spraying a fine mist into the ventilation air or by pulling outside air through matrices (such as aspen fiber pads or

manufactured units having large surface areas but limited airflow restriction) or structures that are wetted to cool the air flowing past them. A less common evaporative cooling means is to spray the outside of the greenhouse with water, or perhaps to wet a shade cloth covering the greenhouse.

Lighting. Greenhouse lighting may be used for photoperiodic reasons, or for enhancing growth. Photoperiodic lighting is a very low intensity light during the night to break the darkness period and induce plant responses representative of summer (short nights and long days). Plants use light within the wavelength band of 400–700 nanometers for growth, and grow proportionally to the number of photons within that energy band. This is in contrast to the human eye which has highly variable sensitivity over the visible light band. The photon flux is measured in moles of photons (one mole of photons is equal to Avogadro's constant, approximately 6.02×10^{23}). The average daily solar photon flux over the Earth is approximately 26 mol/m^2, but greenhouse lighting generally provides less due to the cost of electricity. Greenhouse supplemental lighting is usually provided by high-pressure sodium (HPS) lights because of their relatively high energy efficiency. *See* LIGHT; PHOTOPERIODISM; PHOTOSYNTHESIS.

Carbon dioxide. Ambient carbon dioxide concentrations range from 360 parts per million in rural areas to nearly 400 ppm in urbanized areas. Plants, however, are able to use carbon dioxide efficiently at concentrations up to 2000 ppm. Optimum concentrations are often in the range 800–1000 ppm, which can lead to 25% greater growth provided that other inputs are not limited. Carbon dioxide can be added through carefully controlled flue gases from the greenhouse heating system, or from tanks of liquid carbon dioxide.

Solar applications. Greenhouses are, by their nature, solar collectors, but commercial greenhouses generally do not include specialized or enhanced solar features. Home (and hobby) greenhouses, in contrast, frequently include additional thermal mass and other features to enhance their passive solar collection potential.

Mechanization and automation. Many greenhouse operations that were formerly done by hand are now mechanized or automated. Root medium is mixed, fertilized, and placed directly into flats and pots by machine. Seeding can be totally by machine. Transplanting seedlings into larger containers can be by machine. Plant watering and fertilizing (termed "fertigation" when combined) can now be automated using overhead watering booms, drip irrigation emitters, ebb-and-flow benches, or flooded (and then drained) floors. Automatic material movement at harvest, coordinated by a computer, is no longer unusual. Although the human worker is unlikely to be totally replaced, mechanization and automation have reduced labor requirements to five workers per hectare, or fewer, in the most modern greenhouse operations.

Plant nutrition management. Sixteen elements are believed to be essential for plant growth, including carbon, oxygen, and hydrogen. Minerals required for plant nutrition have traditionally been grouped into two categories: macronutrients and micronutrients. The difference is based on the quantity of the nutrient required for good plant growth.

Plant fertilizers are composed of a mix of salts that are electrically conducting when dissolved into water. This characteristic leads to the use of electrical conductivity as a measure of fertilizer concentration. Computer programs have been developed that are suitable for balancing a nutrient mix to achieve close approximations to the desired molar ratios of elements. *See* PLANT MINERAL NUTRITION.

Hydroponics. Hydroponics is defined as growing plants without using soil. However, a root medium such as sand, gravel, or rockwool may be used. Two common hydroponics systems that use no root medium are the nutrient film technique (NFT) and deep flow troughs (DFT). *See* HYDROPONICS.

The nutrient film technique is a closed system for growing plants so their roots remain in a shallow stream of recirculating nutrient solution. There is no root medium other than the nutrient solution. However, seedlings are typically started in a small cube of root medium, which is then transferred to the nutrient film technique system when roots begin to emerge from the cube. All required nutritional elements are dissolved in the nutrient solution. The system typically uses shallow troughs or channels to support the plant roots and contain the flowing nutrient solution. The defining characteristic of true nutrient film technique is that the nutrient stream is very shallow. Plant roots rest partly within the stream and partly within the air above. However, nutrient solution moves by capillary action around the roots that extend into the air to bathe them in water and nutrients, while permitting free access by the roots to oxygen. The nutrient stream may flow continuously or may be intermittent with the roots alternately submerged and exposed to air (on a rapid cycle of numerous times each hour so even the smallest roots do not desiccate).

The deep flow troughs system for plant production in closed systems is also identified as ponds, deep-flow hydroponics, and raft systems. The plants float on a raft (typically a sheet of foam plastic such as polystyrene) in a shallow tank less than 0.3 m (1 ft) deep. The tank, or pond, is filled with nutrient solution, and the plant roots hang down into the solution. The nutrient solution is monitored, oxygenated, replenished, and recirculated as required. The mass of water in the troughs provides several distinct advantages. First, it provides significantly greater buffering than do other hydroponic techniques such as the nutrient film technique. Changes of nutrient levels, pH, temperature, and dissolved oxygen (DO) concentration (as examples) are significantly damped. Additionally, the ponds can be used for material movement with a minimum of mechanization or physical effort.

Economics. Growing plants in greenhouses is another form of farming and is subject to many of the risks inherent in agriculture. Modern greenhouse

technologies have mirrored developments in most of agriculture in that increased labor efficiency, larger sizes of greenhouse operations, and mass production of a few crops, or even a single crop, have become the rule to be profitable. Large buyers of greenhouse products (such as discount chains) can deal best with large producers, which has led to considerable industry consolidation. There remain, of course, many niche opportunities for small grower operations. The current dynamic in the greenhouse industry in the United States is characterized by the entry of many growers in small, specialized operations, and consolidations and mergers of large operations. *See* FLORICULTURE; PLANT GROWTH; PLANT-WATER RELATIONS.

Louis D. Albright

Bibliography. R. A. Aldrich and J. W. Bartok, Jr., *Greenhouse Engineering*, Natural Resource, Agriculture and Engineering Service (NRAES), Cornell University, Ithaca, NY, 1994; E. Goto et al. (eds.), *Plant Production in Closed Ecosystems*, Kluwer Academic, Dordrecht, The Netherlands, 1997; J. J. Hanan, *Greenhouses: Advanced Technology for Protected Horticulture*, CRC Press, Boca Raton, FL, 1998; R. W. Langhans, *Greenhouse Management*, Halcyon Press, Ithaca, NY, 1990; H. M. Resh, *Hydroponic Food Production*, Woodbridge Press, Santa Barbara, CA, 1995.

Greenland

The world's largest island. It is a major land and ice mass in the North Atlantic Ocean between North America and Europe, bounded by the Arctic Ocean to the north. The total area of Greenland is 2,166,086 km^2 (836,330 mi^2), or a little more than three times the size of Texas. Its most distinctive feature is an ice sheet that covers about 80% of the island like a dome, with elevations toward the center of the island exceeding 3000 m (9800 ft). The Greenland Ice Sheet is the largest ice mass in the Northern Hemisphere and is the second largest ice sheet in the world. (The Antarctic Ice Sheet is about ten times larger.) The Greenland Ice Sheet affects weather systems in the North Atlantic Ocean and northern Europe, and has the potential to affect the environment by contributing to global sea-level rise. Greenland has a total ice volume of 2,600,000 km^3 (624,000 mi^3) which, if melted, would cause global sea level to rise about 6 m (20 ft). Significant melting would also introduce a large volume of freshwater into the North Atlantic Ocean, which could disrupt the Gulf Stream current. Recent studies, using ice thickness measurements from airborne ice-penetrating radar and ice surface movement from satellite measurements, show accelerated ice discharge into the ocean. An example of an ice thickness profile along a flight line from the west to the east coast is shown in the **illustration**. *See* ARCTIC AND SUBARCTIC ISLANDS; ARCTIC OCEAN; GLACIOLOGY; GULF STREAM.

Greenland is also a country that has a population of about 56,000, most of whom live in small villages along the coast, with about 25% of the population living in the capital Godthab (Nuuk). Emblematic of its polar location, the summer solstice is a national holiday. Greenland is a self-governing parliamentary democracy but is part of the Kingdom of Denmark and uses the Danish krone as its currency. Denmark handles its foreign affairs and defense responsibilities, and provides financial support that is equivalent to about half the Greenland government revenues.

Greenland's climate is classified (Köppen system) as polar tundra along the coastal margins, with exception of the northern regions bordering the Arctic

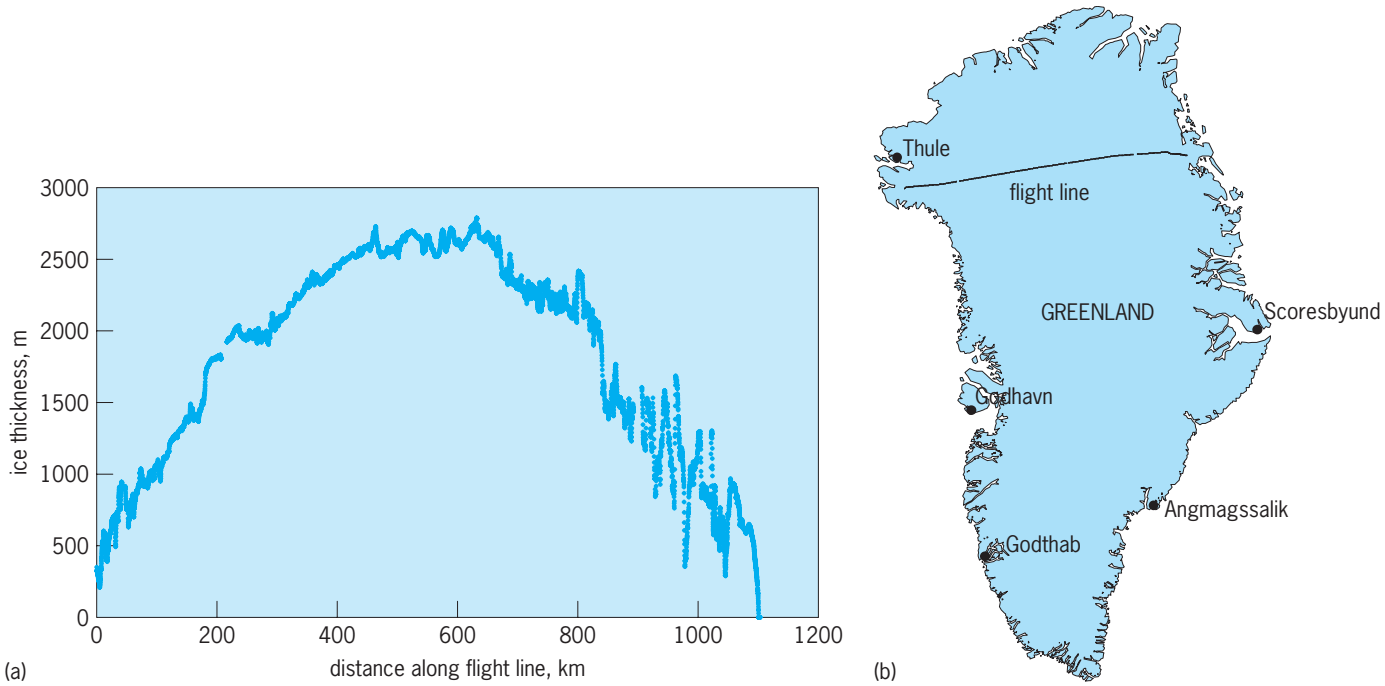

Aerial survey using ice-penetrating radar, showing (*a*) the ice thickness profile along (*b*) a flight line. (*University of Kansas*)

Ocean. For this region and the interior region of Greenland beginning at the ice-sheet edge, the climate is classified as polar ice cap. Along the coastal margins, the average temperature and precipitation generally decrease from south to north. There are profound climatic differences between the east and the west coasts of Greenland because of differences in the ocean current direction. Along the east coast, the East Greenland Current transports cold arctic water southward, reducing the air temperature. Along the west coast, a northward-flowing current warmed by a branch of the Gulf Stream helps to moderate the air temperature. *See* KÖPPEN CLIMATE CLASSIFICATION SYSTEM; POLAR METEOROLOGY; TUNDRA.

The eastern part of Greenland is dominated by the East Greenland Mountains, much of which is buried under the Greenland Ice Sheet. Exposed mountain peaks near the ice-sheet margin range from 1800 to 3700 m (5900 to 12,000 ft). The terrain along the west coast ranges from rolling hills to mountainous, with peaks of 700 to 1000 m (2300 to 3300 ft). The East Greenland Mountains strongly influence the ice flow of the Greenland Ice Sheet, with the main ice flow toward the west. David A. Braaten

Bibliography. Central Intelligence Agency, *The World Factbook*, 2006; S. Gogineni et al., Coherent radar ice thickness measurements over the Greenland Ice Sheet, *J. Geophys. Res.*, 106(D24):33,761–33,772, 2001; E. Rignot and P. Kanaratnam, Changes in the velocity structure of the Greenland Ice Sheet, *Science*, 311:5763, 986–990, 2006; A. Weidick, *Greenland: Satellite Image Atlas of Glaciers of the World*, U.S. Geol. Surv. Prof. Pap. 1386-C, 1995.

Greenockite

A mineral having composition CdS (cadmium sulfide) and crystallizing in the hexagonal system (dihexagonal pyramidal class). Pyramidal crystals are rare, and greenockite usually occurs as earthy coatings with resinous luster and yellow-to-orange color. There is good prismatic cleavage; the hardness is 3 (Mohs scale) and specific gravity is 4.9. Greenockite and wurtzite, ZnS, are isostructural, and a complete solid-solution series exists between the two minerals. Although greenockite is the most common cadmium mineral, no deposits of it are sufficiently large to warrant mining it solely as a source of cadmium. It is commonly associated with sphalerite, and thus the supply of cadmium comes as a by-product from the treating of zinc ores. *See* CADMIUM; SPHALERITE.

Cornelius S. Hurlbut, Jr.

Green's function

A solution of a partial differential equation for the case of a point source of unit strength within the region under examination. The Green's function is an important mathematical tool that has application in many areas of theoretical physics including mechanics, electromagnetism, acoustics, solid-state physics, thermal physics, and the theory of elementary particles. The underlying physics in each of these areas is generally described by some linear partial differential equation which relates the physical variable of interest (electrostatic potential or pressure amplitude in a sound wave, for example) to a source function. For present purposes the source may be regarded as an independent entity, although in some applications (for example, particle physics) this view masks an inherent nonlinearity. The source may be physically located within the region of interest, it may be simulated by certain boundary conditions on the surface of that region, or it may consist of both possibilities. A Green's function is a solution to the relevant partial differential equation for the particular case of a point source of unit strength in the interior of the region and some designated boundary condition on the surface of the region. Solutions to the partial differential equation for a general source function and appropriate boundary condition can then be written in terms of certain volume and surface integrals of the Green's function.

Helmholtz equation. As a primary example, consider the inhomogeneous Helmholtz equation with parameter k, Eq. (1), where ρ_k is the source, ∇^2 is

$$(\nabla^2 + k^2)\,\phi_k(\vec{r}) = -4\pi\rho_k(\vec{r}) \qquad (1)$$

the laplacian operator, and ϕ_k is the function of the independent variable $\vec{r}$ whose solution is sought. This is an important partial differential equation whose solutions can be related through Fourier transformation to solutions of the inhomogeneous wave equation (that is, the wave equation with a source). *See* CALCULUS OF VECTORS; FOURIER SERIES AND TRANSFORMS; LAPLACIAN; WAVE EQUATION.

The special case of $k = 0$ gives Poisson's equation which describes both electrostatics and newtonian gravitation. The Green's function for the Helmholtz equation obeys Eq. (1) with the source replaced by a Dirac delta function $\delta^3(\vec{r} - \vec{r}')$ as appropriate for a point source. The Dirac delta function is zero except at $\vec{r} = \vec{r}'$ where its value is infinite. This infinity is such that the volume integral of $\delta^3(\vec{r} - \vec{r}')$ over $\vec{r}$ (for fixed $\vec{r}'$) is one. While it is convenient to think of $\vec{r}'$ as the location of the source and $\vec{r}$ as the location at which the strength of the Green's function (due to the source) is determined, the inherent symmetry property of Eq. (2) implies that source and field

$$G_k(\vec{r}, \vec{r}') = G_k(\vec{r}', \vec{r}) \qquad (2)$$

points can be interchanged. A unit source at $\vec{r}'$ produces the strength $G_k(\vec{r},\vec{r}')$ at $\vec{r}$, and a unit source at $\vec{n}$ produces the strength $G_k(\vec{r},\vec{r}')$ at $\vec{r}'$. Green's theorem enables the general solution to be determined from the Green's function, using Eq. (3), where $\vec{n}'$ is the

$$\begin{aligned}\phi_k(\vec{r}) = &\int_V dV' G_k(\vec{r}, \vec{r}')\rho_k(\vec{r}) \\ &+ \oint_s dS'\; G(\vec{r}, \vec{r}')\vec{n}' \cdot \vec{\nabla}'\phi_k(\vec{r}') \\ &- \oint_s dS'\vec{n}' \cdot \vec{\nabla}' G(\vec{r}, \vec{r}')\phi_k(\vec{r}') \qquad (3)\end{aligned}$$

outward normal to the surface S which bounds the region V, and $\vec{\nabla}'$ is the gradient with respect to $\vec{r}'$. If $\phi_k(\vec{r})$ is specified on the boundary (the Dirichlet condition), it is required that $G_k(\vec{r},\vec{r}') = 0$ for $\vec{r}'$ on the boundary. If $\vec{n} \cdot \vec{\pi}_k(\vec{r})$ is specified on the boundary (the Neumann condition), then it is required that $\vec{n}' \cdot \vec{\nabla}' G(\vec{r},\vec{r}') = 0$ for $\vec{r}'$ on the boundary. In principle, a more complicated boundary condition can be imposed. The physical content of Eq. (3) is that the solution $\phi_k(\vec{r})$ arises from the sources within V (the volume integral) and the effective sources on the surface S (the surface integrals). *See* GREEN'S THEOREM; POTENTIALS.

Boundary effects. If the boundary recedes to infinity, $G_k(\vec{r},\vec{r}')$ approaches 0 as $|\vec{r} - \vec{r}'|$ approaches infinity, and the Green's function for the Helmholtz equation is given in closed form by Eq. (4) for gen-

$$G_k(\vec{r}, \vec{r}') = \frac{e^{\pm ik|\vec{r}-\vec{r}'|}}{|\vec{r} - \vec{r}'|} \tag{4}$$

erally complex values of k the top sign is taken for the imaginary part of k (Im k) greater than 0 and the bottom is taken for Im k less than 0. The need for complex values of k comes about when the Green's function for thewave equation, Eq. (5), is found by

$$\left(\nabla^2 - \frac{1}{S^2}\frac{\partial^2}{\partial t^2}\right)\phi(\vec{r}, t) = -4\pi\rho(\vec{r}, t) \tag{5}$$

Fourier transformation of $G_k(\vec{r},\vec{r}')$. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

For finite domains the determination of $G_k(\vec{r},\vec{r}')$ is difficult and usually analytically impossible, but for a number of boundary shapes the Green's function can be expressed as the infinite series in Eq. (6), where $u_n(\vec{r})$ is an eigenfunction of the ho-

$$G_k(\vec{r}, \vec{r}') = \sum_{n=1}^{\infty} \frac{u_n(\vec{r})u_n^*(\vec{r}')}{k^2 - k_n^2} \tag{6}$$

mogeneous Hemholtz equation (for the specified boundary condition) and k_n^2 is its associated eigenvalue. In principle, it is possible to explicitly carry out some of the summation implicit in Eq. (6), and in practice this procedure is often very useful. For parallelepipeds, right circular cylinders, and spheres, the eigenfunctions and eigenvalues can be determined in terms of functions such as sines and cosines, Bessel functions and Legendre functions. In some cases the Green's function can be found in closed form. For the half-space $z > 0$ the Green's function is given by Eq. (7), where R_- and R_+ are given by Eqs. (8).

$$G_k(\vec{r}, \vec{r}') = \frac{e^{ikR_-}}{R_-} \mp \frac{e^{ikR_+}}{R_+} \qquad (\text{Im } k > 0) \tag{7}$$

$$\begin{aligned} R_- &= [(x - x')^2 + (y - y')^2 + (z - z')^2]^{1/2} \\ rR_+ &= [(x - x')^2 + (y - y')^2 + (z + z')^2]^{1/2} \end{aligned} \tag{8}$$

In Eq. (7) the minus sign between the two terms is for the Dirichlet condition on the surface $z = 0$ and the plus sign is for the Neumann condition. The case Im $k < 0$ can be treated as in Eq. (4). For Poisson's equation ($k = 0$), closed-form expressions can also be found for several other geometries by the technique referred to as the method of images. An image source is one placed outside the region V in such a way that the boundary condition is satisfied while leaving the source function unchanged in V. Equation (7) is an example of the method of images for the Helmholtz equation with the image source placed at the point $(x', y', -z')$. The second term in Eq. (7) is the contribution from the image source and the first term is the contribution from the source. *See* BESSEL FUNCTIONS; EIGENFUNCTION; LEGENDRE FUNCTIONS; SPECIAL FUNCTIONS; TRIGONOMETRY.

Spectral property. An important property of the Helmholtz Green's function is that it provides the eigenvalue spectrum of the Helmholtz equation. It is evident from the eigenfunction expansion [Eq. (6)] that $G_k(\vec{r},\vec{r}')$ has a pole in the complex k plane at $k^2 = k_n^2$ with residue related to the eigenfunction u_k. This property is particularly useful if it is possible to find the Green's function by some means other than the eigenfunction expansion. The property is also of great importance in applications to complicated many-particle systems.

Quantum-mechanical applications. The time-independent Schröodinger equation for a single particle in an external potential can be cast in the form of Eq. (1), but the source function depends linearly on the wave function $\phi_r(\vec{r})$. In this case Eq. (1) becomes a linear integral equation for the wave function whose kernel is the Helmholtz Green's function. Alternatively, a Green's function can be introduced directly for the Schröodinger equation which has an eigenfunction expansion similar to Eq. (6) in which $u_n(\vec{r})$ is an eigenfunction of the Schröodinger equation and the k^2 is its energy eigenvalue (to within a scale factor). *See* EIGENVALUE (QUANTUM MECHANICS); INTEGRAL EQUATION; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

It is also possible to describe many-particle quantum-mechanical systems in terms of a Green's function formalism. Such systems occur in condensed matter physics (notably, solid-state physics), nuclear physics, and elementary particle physics. Unfortunately, the equations obeyed by these Green's functions are generally nonlinear and approximation methods are required. But the spectral property of Eq. (6) remains an important feature, and it is sometimes possible to detect effective single-particle–like behavior in complicated many-particle motion. These quasi-particles appear as (approximate) poles in the Green's function for the many-particle system.

If the interactions between particles are relatively unimportant, it is possible to develop an approximation scheme in which the Green's function for an individual free particle plays an essential role. These Green's functions are called propagators and form the basis for many calculations in modern relativistic quantum field theory. *See* DIFFERENTIAL EQUATION; PROPAGATOR (FIELD THEORY); QUANTUM CHROMODYNAMICS; QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY.

Peter Shaw

Bibliography. G. B. Arfken and H.-J. Weber (eds.), *Mathematical Methods for Physicists*, 5th ed., 2000; G. Barton, *Elements of Green's Functions and Propagation*, 1989; E. N. Economou, *Green's Functions in Quantum Physics*, 1990; G. D. Mahan, *Many-Particle Physics*, 3d ed., 2000; R. D. Mattuck, *A Guide to Feynman Diagrams in the Many-Body Problem*, 2d ed., 1976, reprint 1992.

Green's theorem

A term used variously in mathematical literature to denote either (1) the Gauss divergence theorem (shown below) or some one of several forms or

$$\iiint \nabla \cdot \mathfrak{I}\, dV = \iint \mathfrak{I} \cdot \boldsymbol{\nu}\, dS$$

immediate consequences of their theorem, or (2) the plane case of Stokes' theorem. *See* GAUSS' THEOREM; STOKES' THEOREM.

The variation of the Gauss divergence theorem in which $\mathfrak{I}$ is the product $u\ \nabla v$ of a scalar function and the gradient of a scalar function is sometimes called Green's first identity. In this case Eq. (1) holds, and Gauss' theorem assumes the form of Eq. (2), where

$$\nabla \cdot \mathfrak{I} = \nabla \cdot (u\ \nabla v) = u\ \nabla^2 v + \nabla u \cdot \nabla v \qquad (1)$$

$$\iiint u\ \nabla^2 v\, dV + \iiint \nabla u \cdot \nabla v\, dV = \iint u \frac{\partial v}{\partial n}\, dS \qquad (2)$$

$\partial v/\partial n = \nabla\ v \cdot \boldsymbol{\nu}$ and is therefore the directional derivative of v in the direction of $\boldsymbol{\nu}$ the unit normal to the surface; thus n may be taken to be arc length along any curve of class C' which is tangent to $\boldsymbol{\nu}$ at the point in question. The derivative $\partial v/\partial n$ is a function of the surface coordinates.

By interchanging u and v in Eq. (2) with subsequent subtraction of the new form from the original, there results the relationship shown as Eq. (3),

$$\iiint u\ \nabla^2 v\, dV - \iiint v\ \nabla^2 u\, dV = \iint \left(u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} \right) dS \qquad (3)$$

which is known as Green's second identity as well as Green's theorem. Two additional equalities associated with Green may be obtained by taking $u = v$ and $u = 1$ in Eq. (2); thus Eqs. (4) and (5) result.

$$\iiint u\ \nabla^2 u\, dV + \iiint \nabla u \cdot \nabla u\, dV = \iint u \frac{\partial u}{\partial n}\, dS \qquad (4)$$

$$\iiint \nabla^2 v\, dV = \iint \frac{\partial v}{\partial n}\, dS \qquad (5)$$

The formulas (2)–(5) provide powerful devices for the investigation of differential equations involving the laplacian operator ∇^2, such as Poisson's equation and the wave equations. *See* CALCULUS OF VECTORS.

Homer V. Craig

Gregarinia

A subclass of the class Telosporea. These protozoans occur principally as extracellular parasites in the digestive tracts and body cavities of invertebrates. Their spores are formed directly by the zygote. There are three orders: the Archigregarinida, whose life cycle embraces both sexual and asexual phases; the Eugregarinida, which increases only by sporogony; and the Neogregarinida, whose life cycle involves schizogony and gamont formation. The most familiar gregarines belong to the Eugregarinida and are represented by two types: cephaline, whose trophozites (sporadins) are divided into an anterior protomerite and a larger posterior deutomerite by a transverse septum; and acephaline, which lack the septum. *See* ARCHIGREGARINIDA; NEOGREGARINIDA; PROTOZOA; SPOROZOA; TELOSPOREA.

Elery R. Becker; Norman Levine

Greisen

A type of hydrothermal wall-rock alteration and a class of tin-tungsten deposits (so-called greisen deposits). Hydrothermal wall-rock alteration is the process whereby rocks on the margins of hydrothermal flow channels are changed from an original assemblage of minerals to a different one. This change occurs because of heat and mass exchange between water and rock. The term greisen was originally used by miners in Saxony with reference to relatively coarse-grained aggregates of quartz and muscovite found on the borders of tin veins in granite of the Erzegebirge.

Granitic rocks altered to greisen are known as apogranites. They are composed mainly of quartz, topaz (fluor-aluminosilicate), and muscovite (white mica), accompanied by accessory minerals such as tourmaline (complex boro-aluminosilicate), fluorite (calcium fluoride), and zinnwaldite (iron-lithium mica). Abundant veins of quartz-topaz are characteristic of intensely greisenized zones. Skarn and limestone on the margins of apogranites may also be altered to greisen (aposkarn greisen and apocarbonate greisen, respectively) with abundant fluorite. Apogranite greisen commonly is accompanied by other types of hydrothermal wall-rock alteration, including early feldspathic and late sericitic and lesser argillic.

Tin-tungsten-(beryllium-molybdenum) deposits in peraluminous granites commonly are accompanied by greisen. Ore minerals may include cassiterite, wolframite, scheelite, molybdenite, bismuth, and bismuthinite, accompanied in some deposits by

pyrrhotite and sphalerite, in addition to chalcopyrite and other sulfides. *See* GRANITE.

Tin greisens represent the dominant world source of lode tin, with examples in Southeast Asia (Malaya, Indonesia, Burma, Thailand); southeast China; Tasmania, Australia; Zinnwald and Altenberg (Erzegebirge), Germany; and Cornwall-Devon, southwest England. Greisenized skarn and apocarbonate greisen, also mostly tin deposits (but including beryllium and tungsten), are found in western Tasmania, Australia; Seward Peninsula, Alaska; and Yunnan (tungsten), China. Many of the tungsten deposits of southeast China, the richest tungsten province in the world, occur in greisenized granite. *See* TIN; TUNGSTEN.

Although a continuum likely exists between greisen and lower-temperature sericitic and argillic alteration, a distinction is made between these different alteration types on the basis of fluoride-, boron-, and lithium-enrichment in greisen. Additional mass changes relative to fresh granite include loss of sodium and calcium, gain or loss of potassium and silica, and gain of volatile substances (especially fluorine and sulfur). The chemical components added during greisenization originate in a magmatic aqueous fluid given off during the final stages of crystallization of granitic magmas. Greisenization and ore deposition results from the interaction of these fluorine-rich fluids with granite and its wall rocks as temperatures decline below 600°C (1110°F). *See* METASOMATISM; ORE AND MINERAL DEPOSITS; PNEUMATOLYSIS.

Marco T. Einaudi

Bibliography. R. G. Taylor, *Geology of Tin Deposits*, 1979.

Grignard reaction

A reaction between an alkyl or aryl halide and magnesium metal in a suitable solvent, usually absolute ether. This is represented by reaction (1), where R

$$RX + Mg \xrightarrow{\text{ether}} RMgX \qquad (1)$$

stands for alkyl or aryl, and X for chlorine, bromine, or iodine. The organomagnesium halides produced by this reaction are known as Grignard reagents and are useful in many chemical syntheses. They are named after Victor Grignard, who discovered them and developed their use as synthetic reagents, for which he received a Nobel prize in 1912.

The scope of the Grignard reaction is extremely broad, and Grignard reagents have been prepared from many kinds of alkyl and aryl halides. In general, alkyl chlorides, bromides, and iodides and aryl bromides and iodides react readily. A few halides, such as aryl chlorides, react very sluggishly, and require specially activated magnesium, modified reaction techniques, the use of high-boiling solvents, and long reaction periods. Vinylmagnesium halides can be prepared by using tetrahydrofuran as the solvent.

The structure of a Grignard reagent is usually written RMgX, where X represents a halogen. However, the actual structure in ether solution is more complex. Equilibrium is rapidly established between the organomagnesium halide (RMgX) and the corresponding dialkylmagnesium (RMgR), as in reaction (2). These two species (RMgX and R_2Mg) are

$$2RMgX \rightleftharpoons R_2Mg + MgX_2 \qquad (2)$$

reactive, and are solvated in ether solvents by coordination of the ether oxygen to magnesium. In solution they further associate as dimers or higher polymers. Thus, while oversimplified, a Grignard reagent can be considered as RMgX.

Probably the most common synthetic use of the Grignard reaction involves the reaction of Grignard reagents with carbonyl compounds, followed by hydrolysis to produce a variety of products containing new carbon-carbon bonds. For example, in reaction (3), a Grignard reagent with carbon dioxide at

$$RMgX + CO_2 \rightarrow RCOOMgX \xrightarrow{H_2O} RCOOH + MgXOH \qquad (3)$$

low temperatures produces salts of carboxylic acids, from which the acids can be readily isolated. Additon of Grignard reagents to the carbonyl group of many aldehydes and ketones produces primary, secondary, or tertiary alcohols. The production of these alcohols is shown in reactions (4)–(6).

$$RMgX + CH_2{=}O \longrightarrow RCH_2OMgX \xrightarrow{H_2O} \underset{\text{Primary alcohol}}{RCH_2OH} + MgXOH \qquad (4)$$

$$RMgX + R'CH{=}O \longrightarrow RR'CHOMgX \xrightarrow{H_2O} \underset{\text{Secondary alcohol}}{RR'CHOH} + MgXOH \qquad (5)$$

$$RMgX + R'R''C{=}O \longrightarrow RR'R''COMgX \xrightarrow{H_2O} \underset{\text{Tertiary alcohol}}{RR'R''COH} + MgXOH \qquad (6)$$

The reaction of Grignard reagents with carbon dioxide, aldehydes, and ketones serves as a convenient means of lengthening the carbon chain. Further, a hydrocarbon chain can be increased by two carbon atoms by the reaction of a Grignard reagent with ethylene oxide, as in reaction (7). The result-

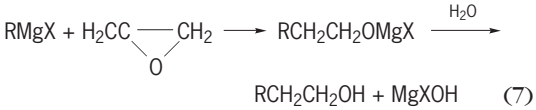

(7)

ing alcohol can be converted to the corresponding halide, and the process repeated.

Grignard reagents react with esters to produce ketones initially. The reaction as a rule does not stop at this stage, however, since additional Grignard

reagent reacts with the ketone to yield a tertiary alcohol.

Acid chlorides and anhydrides react in an analogous manner to produce ketones, which react further to give tertiary alcohols. A valuable modification of this procedure involves prior reaction of the Grignard reagent with cadmium or zinc chloride. The organocadmium or organozinc compounds are less reactive than those of magnesium and react so slowly with ketones that the conversion of the ketone to the tertiary alcohol can usually be avoided.

When aliphatic Grignard reagents are exposed to air or oxygen, they are oxidized to alkoxides, which on hydrolysis produce alcohols. Aromatic Grignard reagents react poorly in this way, and only unsatisfactory yields of phenols can be obtained.

The reaction of Grignard reagents with alkyl halides which contain reactive halogen atoms produces hydrocarbons by a process of alkylation, as in reaction (8). In a similar manner, the addition

$$\mathrm{RMgX} + \mathrm{XR} \longrightarrow \mathrm{R{-}R} + \mathrm{MgX_2} \qquad (8)$$

of a variety of metal halides, such as silver bromide or cobaltous chloride, tends to couple Grignard reagents, as in reaction (9).

$$2\mathrm{RMgX} + 2\mathrm{AgBr} \longrightarrow \mathrm{R{-}R} + 2\mathrm{Ag} + 2\mathrm{MgXBr} \qquad (9)$$

Compounds containing active hydrogen atoms react readily with Grignard reagents. Reaction with water brings about immediate decomposition and formation of the corresponding hydrocarbon, as in reaction (10). In certain instances, hydrogen at-

$$\mathrm{RMgX} + \mathrm{H_2O} \longrightarrow \mathrm{RH} + \mathrm{MgXOH} \qquad (10)$$

tached to carbon is sufficiently acidic to react with Grignard reagents. The reactions of ethyl-magnesium bromide with acetylene, reaction (11), and cyclopentadiene, reaction (12), are examples. These reactions

$$2\mathrm{C_2H_5MgBr} + \mathrm{HC{\equiv}CH} \longrightarrow 2\mathrm{C_2H_6} + \mathrm{BrMgC{\equiv}CMgBr} \qquad (11)$$

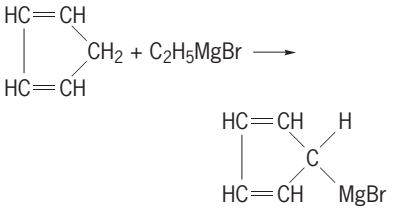

$$+ \mathrm{C_2H_6} \qquad (12)$$

enable formation of additional Grignard reagents which could otherwise not be prepared by conventional methods.

Grignard reagents react with a variety of nitrogen- and sulfur-containing compounds, such as imines, nitriles, and sulfoxides. Furthermore, the reaction of Grignard reagents with halides of boron, phosphorous, silicon, and tin is a very convenient procedure for producing organometallic compounds of these elements. An example, reaction (13), is the

$$4\mathrm{C_6H_5MgBr} + \mathrm{SiCl_4} \rightarrow (\mathrm{C_6H_5})_4\mathrm{Si} + 4\mathrm{MgBrCl} \qquad (13)$$

preparation of tetraphenylsilane from phenylmagnesium bromide and silicon tetrachloride. *See* ORGANOMETALLIC COMPOUND. Paul E. Fanta

Bibliography. L. Kurti and B. Czako, *Strategic Applications of Named Reactions in Organic Synthesis*, 2005; J. J. Li, *Name Reactions: A Collection of Detailed Reaction Mechanisms*, 2d ed., 2003; M. B. Smith and J. March, *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*, 5th ed., 2001; T. W. G. Solomons, *Organic Chemistry*, 8th ed., 2003.

Grimmiales

An order of the true mosses (subclass Bryidae) which consist of two families and about five genera and generally grow in dry, exposed places, especially on rock. They are often dark, even blackish, with erect-ascending and simple or forked stems, or prostrate and freely branched stems. The leaves, in many rows, are often hair-pointed (see **illus.**) with a single and well-developed costa. The cells are generally smooth but commonly have side walls that are nodulose-thickened. The capsules are immersed to exserted (sometimes on curved setae), and are usually erect, symmetric, and sometimes ribbed. The operculum is differentiated, and the single peristome (rarely lacking) is composed of 16 teeth which are commonly perforated or variously cleft. The calyptrae are cucullate or mitrate. The chromosome numbers are generally 12 and 13.

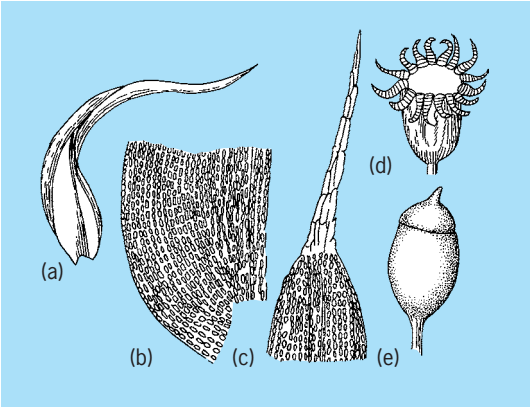

***Grimmia apocarpa*; grows generally on rock. (*a*) Leaf. (*b*) Portion of base of leaf. (*c*) Apex of leaf. (*d*) Urn and peristome. (*e*) Capsule. (*After W. H. Welch, Mosses of Indiana, Indiana Department of Conservation, 1957*)**

Grimmia has many resemblances to the genus *Orthotrichum*, which belongs to a double-peristome series. Both Hypnales and Orthotrichales seem intermediate between the acrocarpous and pleurocarpous groups of mosses and grow in similar, dry habitats. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA; ORTHOTRICHALES. Howard Crum

Grinding mill

A machine that reduces the size of particles of raw material fed into it. The size reduction may be to facilitate removal of valuable constituents from an ore or to prepare the material for industrial use, as in preparing clay for pottery making or coal for furnace firing. Coarse material is first crushed. The moderate-sized crushings may be reduced further by grinding or pulverizing.

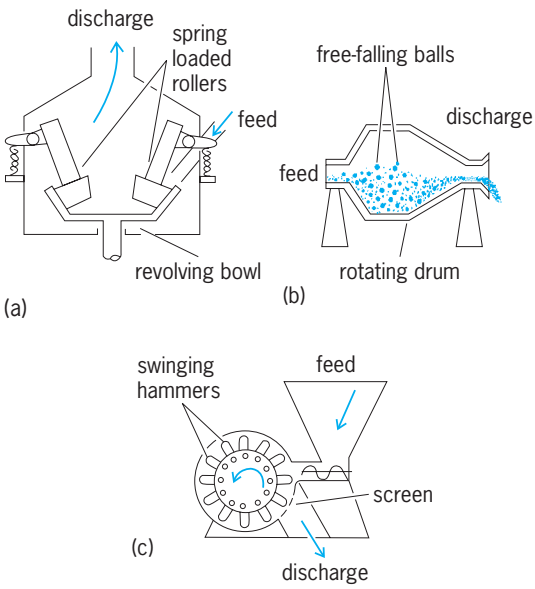

Basic grinding mills. (*a*) Ring-roller mill. (*b*) Tumbling mill. (*c*) Hammer mill.

Grinding mills are of three principal types, as shown in the **illustration**. In ring-roller pulverizers, the material is fed past spring-loaded rollers. The rolling surfaces apply a slow large force to the material as the bowl or other container revolves. The fine particles may be swept by an airstream up out of the mill. In tumbling mills the material is fed into a shell or drum that rotates about its horizontal axis. The attrition or abrasion between particles grinds the material. The grinding bodies may be flint pebbles, steel balls, metal rods lying parallel to the axis of the drum, or simply larger pieces of the material itself. In hammer mills, driven swinging hammers reduce the material by sudden impacts. *See* BALL-AND-RACE-TYPE PULVERIZER; PEBBLE MILL; TUMBLING MILL.

Depending on the required fineness and uniformity of the finished particles, the discharge may or may not be classified by size. Oversized particles may be returned in a closed circuit to the grinding mill for further reduction. Material may be ground dry or wet, in batches or continuously. *See* CRUSHING AND PULVERIZING; MECHANICAL CLASSIFICATION.

Ralph M. Hardgrove

Bibliography. F. T. Farago, *Abrasive Methods Engineering*, vols. 1–2, 1976–1980; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997.

Gromiida

An order of the subclass Filosia. The test of these protozoa is mostly chitinous in some species, rather thin and a bit flexible in others. Siliceous particles of endogenous origin, or sometimes minute platelets, may be embedded in the chitinous layer; some species typically reinforce the test with sand grains or other extraneous particles. In addition to *Gromia*, the genera *Chlamydophrys*, *Clypeolina*, *Plagiophrys* (**illus.** *a*), and *Pseudodifflugia*

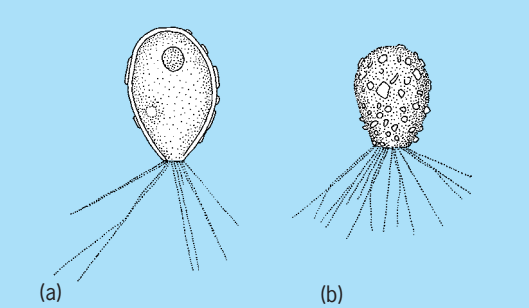

Representative Gromiida. (*a*) *Plagiophrys parvipunctata*. (*b*) *Pseudodifflugia fulva*. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)

(illus. *b*), among others, have been assigned to the order. A few species are marine, the rest fresh-water types. Little is known about most Gromiida, although mitosis and encystment have been described in *Chlamydophrys*. *See* PROTOZOA; RHIZOPODEA; SARCODINA; SARCOMASTIGOPHORA.

Richard P. Hall

Ground-penetrating radar

A nondestructive technique using electromagnetic waves to locate objects or interfaces buried beneath the Earth's surface or located within a visually opaque structure; also termed ground-probing, surface-penetrating, or subsurface radar.

Principle of operation. Ground-penetrating radar (GPR) transmits a regular sequence of low power bursts of electromagnetic energy into the ground and detects the weak reflected signal from the buried target. The energy is in the form of either a very short duration impulse or a sweep over a range of frequencies. Most GPR systems, which are regulated and licensed by the countries in which they are used to comply with radio spectrum requirements, operate within the range of frequencies from 10 MHz to 10 GHz and can have a bandwidth of several gigahertz. The average radiated power is in the order of a thousandth of a watt. The receiver is highly sensitive and can detect reflected signals of less than 10^{-12} W. *See* MICROWAVE; RADAR.

The buried target can be a conductor of electricity, a nonconducting dielectric, or combinations of both; the surrounding host material can be soil, earth materials, rocks, ice, fresh water, or human-made materials such as concrete or brick. A typical GPR achieves

Fig. 1. Ground-penetrating radar equipment. (*a*) Portable equipment with 500-MHz antenna and controller that displays scrolling vertical slices of the soil in real time (*USRADAR*). (*b*) Portable equipment with 400-MHz antenna being used in archeological survey at Franklin D. Roosevelt's Hyde Park Estate in New York (*Geological Survey Systems, Inc.*).

a range of up to a few meters in depth, but some special systems can penetrate up to hundreds of meters. A few GPR systems have been operated from aircraft and even from satellites. The range of GPR in the ground is limited because of the absorption the signal undergoes while it travels on its two-way path through the ground. GPR works well through materials such as granite, dry sand, snow, ice, and fresh water, but it will not penetrate clay or salt water because of the high absorption of electromagnetic energy by these materials. In air the GPR signal travels at the speed of light, but it is slowed down in materials in the ground by their permittivities (dielectric constants). Hence true range requires calibration for each material. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; PERMITTIVITY.

GPR can be used on vehicles for rapid survey by means of an array of antennas. Other GPR systems are designed to be inserted into boreholes to provide tomographic images of the intervening rock. Most GPR systems use separate, portable, transmit-and-receive antennas which are placed on the surface of the ground and moved linearly to provide an image of the cross section of the ground traversed (**Fig. 1**). By systematically surveying the area in a regular grid pattern, a radar image of the ground can be built up. GPR images are displayed either as two-dimensional representations, using horizontal (x or y) and depth (z) axes or a horizontal plane representation (x,y) at a given depth (z), or as a three-dimensional reconstruction.

The image of a buried target generated by a GPR does not correspond to its geometrical equivalent and is generally of much lower resolution. Unprocessed GPR images often show bright spots caused by multiple internal reflection, as well as a distortion of the aspect ratio of the image of the target caused by variations in the velocity of propagation. Symmetrical targets, such as spheres or pipes, cause migration of the reflected energy to a hyperbolic pattern. GPR images can be processed to compensate for these effects. This is usually done off-line. GPR can be

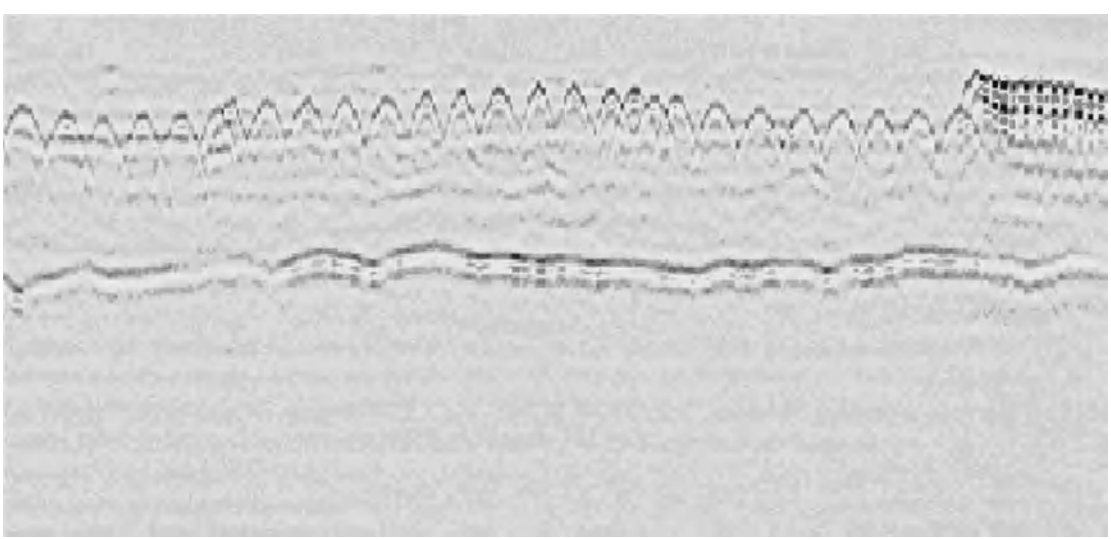

Fig. 2. Radar cross-sectional image of reinforced concrete road. (*ERA Technology UK*)

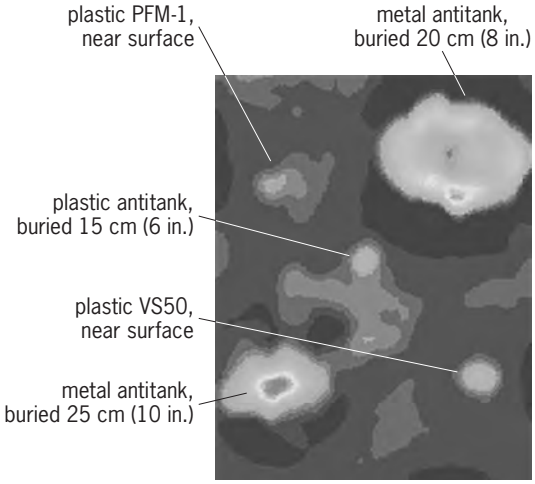

Fig. 3. Radar plan image of mines, area 0.9 by 1 m (3.0 by 3.3 ft). (*ERA Technology UK*)

designed to detect specific targets such as interfaces in roads, pipes, and cables; and localized objects such as cubes, spheres, and cylinders. It is capable of detecting features many hundreds of years old; hence a prospective site should remain unexcavated prior to survey, to preserve its information.

Applications. GPR technology has many applications, some well established and some still undergoing research and development. A typical radar cross-sectional image is shown in **Fig. 2**; it represents a survey line 4.5 m (14.9 ft) long over a reinforced concrete road. The positions of the steel reinforcing mesh rods can be seen at the top of the image, while the interface between the concrete and the road subbase can be seen at a depth of 0.5 m (1.7 ft) half way down the image. Plastic pipes and cables can also be detected by GPR, and this is one of the routine applications of modern GPR equipment.

Typical applications for GPR are as follows:

Archeological investigations
Bridge deck analysis
Borehole inspection
Building condition assessment
Contaminated land investigation
Detection of buried mines (antipersonnel and antitank)
Evaluation of reinforced concrete
Forensic investigations
Geophysical investigations
Medical imaging
Pipes and cable detection
Planetary exploration
Rail track and bed inspection
Remote sensing from aircraft and satellites
Road condition survey
Security applications
Snow, ice, and glacier
Timber condition
Tunnel linings
Wall condition

GPR has been used very successfully in forensic investigations. The most notorious cases occurred in the United Kingdom in 1994 when the grave sites, under concrete, of the victims of a serial murderer were pinpointed. Similar use of GPR was made in Belgium in 1996. GPR in the hands of an expert allows rapid, nondestructive investigation of suspected sites and saves much fruitless excavation.

Abandoned antipersonnel land mines and unexploded ordnance are a major hindrance to the recovery of many countries from war. Their effect on the civilian population is catastrophic, and major efforts are being made by the international community to eliminate the problem. Mines are located mostly with metal detectors, which respond to the large amount of metallic debris in abandoned battlefield areas and hence have difficulty in detecting minimum metal or plastic mines. GPR technology is being applied to this problem (**Fig. 3**).

Archeological applications of GPR have been numerous, ranging from attempts to detect Noah's Ark to the exploration of Egyptian and Indian sites as well as castles and monasteries in Europe. The quality of the radar image can be exceptionally good, although correct understanding normally requires

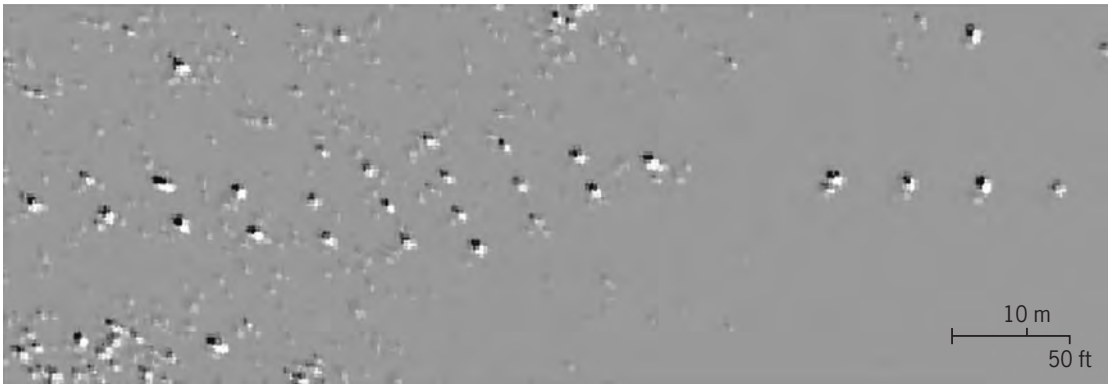

Fig. 4. Synthetic aperture radar (SAR) image of buried metal mines taken from an altitude of 400 m (1300 ft) above the Yuma desert. (*Courtesy of R. Vickers, SRI, USA*)

joint interpretation by the archeologists and radar specialists. GPR has been used for many different types of geological surveys ranging from exploration of the Arctic and Antarctic icecaps and the permafrost regions of North America, to mapping of granite, limestone, marble, and other hard rocks as well as geophysical strata. While most GPR systems are used close to the ground, airborne systems have been able to map ice formations and glaciers, and penetrate through forest canopy. Airborne GPR, processed using synthetic aperture techniques, has been used to detect buried metallic mines from a height of several hundred meters (**Fig. 4**). *See* ARCHEOLOGY; REMOTE SENSING. David J. Daniels

Bibliography. C. E. Baum (ed.), *Detection and Identification of Visually Obscured Targets*, Taylor and Francis, 1998; S. Cloude, *An Introduction to Electromagnetic Wave Propagation and Antennas*, UCL Press, 1995; D. J. Daniels (ed.), *Ground Penetrating Radar*, 2d ed., IET Publishing, 2004; J. D. Taylor (ed.), *Introduction to Ultra-wideband Radar Systems*, CRC Press, 1995.

Ground proximity warning system

A system carried on many aircraft to warn the pilot that the aircraft may be in danger of inadvertent contact with the ground. It is intended to reduce the occurrence of controlled-flight-into-terrain (CFIT) accidents, in which aircraft with no apparent mechanical difficulty or defect strike the ground while under the direct or indirect control of the pilot. These accidents usually occur in conditions of poor visibility (due to atmospheric obscuration such as fog or rain, or darkness of night) where external visual cues are not available to provide added assurance of terrain clearance. Since 1975, federal aviation regulations have required installation of the system on large turbine-powered aircraft in commercial service.

Components. The heart of the ground proximity warning system (GPWS) is a computer which receives inputs from several sensors on the aircraft and issues warnings to the pilot through visual and aural alerting devices. The primary sensors providing inputs to the computer are (1) the radio altimeter, which measures height of the aircraft above the terrain; (2) the barometric altimeter, which measures altitude with respect to an atmospheric pressure datum; (3) the electronic glideslope of the Instrument Landing System (ILS), which provides the pilot with vertical descent guidance during the final portions of an ILS approach and landing; and (4) sensors which indicate aircraft configuration such as the position of flaps and landing gear. When a hazardous condition is detected by the system, the pilot is warned by illumination of caution lights on the instrument panel, accompanied by an aural alarm and a synthetic voice message indicating the type of hazard. *See* ALTIMETER; INSTRUMENT LANDING SYSTEM (ILS).

Operation. The system is designed to detect and warn the pilot of the following conditions: excessive descent rate near the ground; excessive terrain closure rate; approaching the ground with landing gear or flaps not in the landing configuration; and descending significantly below the ILS electronic glideslope when on approach to landing. Also, during takeoff and immediately after initiating a missed-approach go-around, the system warns the pilot if the aircraft is descending when it should normally be climbing. The Radio Technical Commission for Aeronautics (RTCA), in its *Minimum Performance Standards for Airborne Ground Proximity Warning Equipment*, prescribes the parameters that must be monitored and the thresholds at which the system should declare a hazardous condition. There are provisions for the pilot to test the operation of the system, and also to inhibit certain warnings in unusual or emergency conditions (if, for instance, an approach and touchdown is planned with the landing gear or flaps retracted due to mechanical failure). However, the operation of the system generally requires no action on the part of the pilot.

Limitations. While GPWS has been credited with significantly reducing the incidence of CFIT accidents, it does have limitations. Its sensors cannot detect hazards which may be ahead of the aircraft such as steeply rising terrain or artificial obstacles. In addition, the effectiveness of GPWS is largely dependent on the pilot's prompt reaction to the system's warnings. However, these limitations have been surmounted in some respects with the availability of enhanced ground proximity warning systems (EGPWS) since 1996. The enhancement adds a detailed terrain database that can be interrogated to identify hazards along the current course in time for the pilot to take evasive action. Consequences of the tendency of pilots to delay taking evasive action are also mitigated by the additional time to react that is provided by an EGPWS. *See* AIR NAVIGATION. George W. Flathers II

Bibliography. R. Hurst and L. Hurst (eds.), *Pilot Error: The Human Factors*, 1982; *Minimum Performance Standards: Airborne Ground Proximity Warning Equipment*, Radio Technical Commission for Aeronautics, DO-161A, May 1976.

Ground state

In quantum mechanics, the stationary state of lowest energy of a particle or a system of particles. The ground state may be bound or unbound; when bound, its energy generally is a finite amount less than the energy of the next higher or first excited state. In the typical circumstances that the potential energy is zero at infinite separation, the magnitude of the negative ground-state energy is the binding energy, that is, the energy required to separate all the particles infinitely. *See* ENERGY LEVEL (QUANTUM MECHANICS); EXCITED STATE; NONRELATIVISTIC QUANTUM THEORY; NUCLEAR BINDING ENERGY. Edward Gerjuoy

Ground-water hydrology

The occurrence, circulation, distribution, and properties of any liquid water residing beneath the surface of the earth. This article excludes any water molecules which constitute mineral species such as clays. Generally, ground water is that fraction of precipitation which infiltrates the land surface and subsequently moves, in response to various hydrodynamic forces, to reappear once again as seeps or in a more obvious fashion as springs. Most of ground-water discharge is not evident because it occurs through the bottoms of surface water bodies; in fact, large fresh-water springs are relatively common on the ocean bottom off the eastern coast of Florida.

Ground water can be found, at least in theory, in any geological horizon containing interconnected pore space. Thus a ground-water reservoir (an analogy to an oil reservoir) can be a classical porous medium, such as sand or sandstone; a fractured, relatively impermeable rock, such as granite; or a cavernous geologic horizon, such as certain limestone beds. Ground-water reservoirs which readily yield water to wells are known as aquifers; in contrast, aquitards are formations which do not normally provide adequate water supplies, and aquicludes are considered, for all practical purposes, to be impermeable. These terms are, of course, subjective descriptions; the flow of water which constitutes an economically viable supply depends upon the intended use and the availability of alternative sources. *See* AQUIFER.

Physics of flow. Ground water and its dissolved constituents move according to mathematical relationships derived from the physics of flow through porous media. These relationships are simplest in the case of the saturated flow of a homogeneous fluid in a granular formation. There are a number of governing equations for this family of problems, and many solutions are available for problems of practical importance. There are also equations describing the behavior of dissolved constituents in flowing ground water. This class of problems has gained prominence because of the recognition of existing and potential instances of ground-water contamination. However, these equations are more difficult to solve, and consequently the simulation of most field situations involving contaminant transport requires the use of large digital computers. When the unsaturated zone must be considered, the governing equations are still available, but there is a paucity of solutions. The major research advances in unsaturated flow can be attributed to soil physicists, who are primarily concerned about water movement through the root zone during the irrigation of crops.

Regional flow. Ground-water flow in a regional sense is concerned with movement attributable to gravitational forces. Ground water moves from points of high potential energy to points of low potential energy. Water in transit from regional topographic highs to regional topographic lows dissipates energy through friction. The flow path may be relatively short or extend for hundreds of miles, depending on the physical system. The generally observed correspondence between changes in the elevation of the surface of the saturated zone and changes in the land topography is maintained because of a continual source of water via precipitation. When precipitation patterns change, the ground-water surface responds accordingly, rising in wet years and falling in dry years. The topographic lowlands often correspond to areas of ground-water discharge. They may be characterized by springs, wetlands, or surface water bodies such as lakes, rivers, and oceans. *See* WETLANDS.

Artesian systems. The water level in a drilled well is a measure of the potential energy of the ground water it encounters. In a recharge area, where water moves downward under gravitational force, the fluid potential energy, and therefore the water level in wells, decreases with depth. On the other hand, in a discharge area where ground water moves toward the surface, the fluid potential increases with depth. This latter situation results in a well-water elevation which is higher than that observed for the upper surface of the saturated zone in the vicinity of the well. This phenomenon has been designated an artesian head. In some instances the well-water-level elevation may actually be higher than the land surface elevation. In this situation the ground water flows naturally from the well, and no pump is required. Such wells are called flowing artesian wells. It has often been observed that the discharge of flowing wells decreases with time. This is a natural consequence of providing a ground-water flow path to the surface which is less energy-dissipating than the original porous medium flow. *See* ARTESIAN SYSTEMS.

Ground-water mining. It is important to recognize that when ground water is used consumptively, that is, when it is removed permanently from the hydrologic system, the ground-water reservoir must be affected. The most immediate impact is a lowering of the saturated zone elevation. As the influence of the pumping reaches surface water bodies, water is generally induced from them. The rate of loss from the surface water body will increase until an equilibrium is established between water withdrawn from the well and that entering the aquifer. If no surface water bodies are available, the ground-water surface will fall as long as ground water is utilized: this is called ground-water mining. Practically speaking, the process is self-limiting, because the cost of pumping ground water increases as the water levels in the wells decline.

Chemical and physical characteristics. As ground water moves through a regional flow system, its physical and chemical characteristics are modified by the environments encountered. It dissolves soluble ions, and thereby changes its chemical composition. Thus the chemistry of ground water can sometimes be used to decipher its complex flow history. During transit through the subsurface, when ground water encounters elevated rock temperatures, the water is

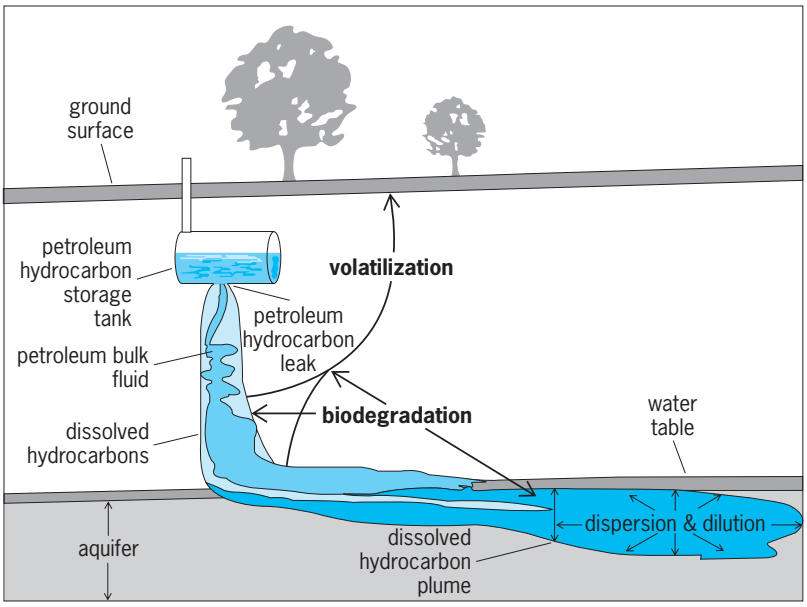

Fig. 1. Movement of a lighter-than-water petroleum hydrocarbon liquid in a flowing ground-water system. (*After D. F. Pope and J. N. Jones, Monitored Natural Attenuation of Petroleum Hydrocarbons, EPA/600/F-98/021, May 1999*)

heated. Under certain circumstances, the water may reach the surface as hot springs or geysers. When surface manifestations are absent, hot water and occasionally steam may be tapped by wells to provide a source of energy. High-temperature water can be used to produce electricity; lower-temperature water may heat homes, or may be employed effectively in selected industries. *See* GEOTHERMAL POWER.

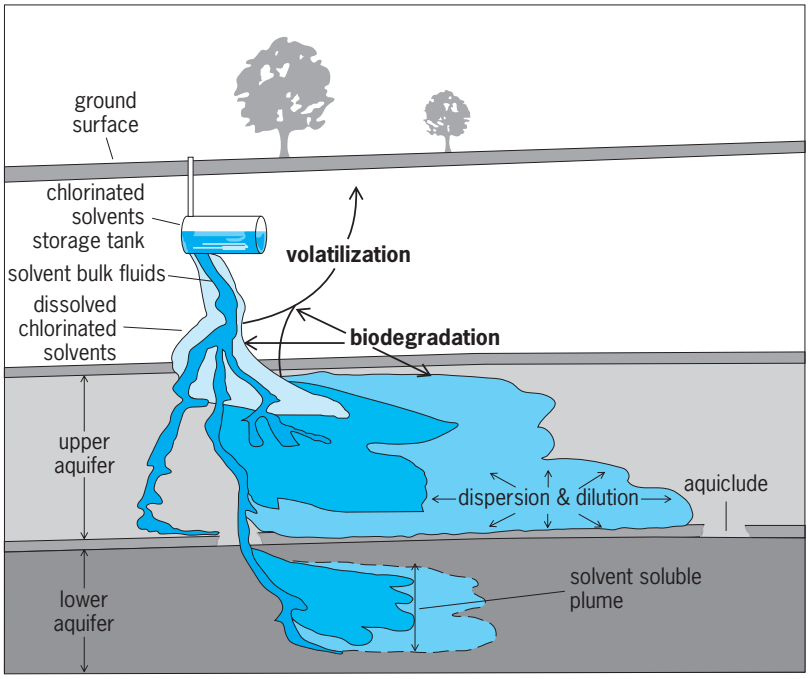

Fig. 2. Movement of a heavier-than-water chlorinated solvent in a flowing ground-water system. (*After D. F. Pope and J. N. Jones, Monitored Natural Attenuation of Chlorinated Solvents, EPA/600/F-98/022, May 1999*)

Reservoir behavior. To effectively utilize ground water as a natural resource, it is necessary to be able to forecast the impact of exploitation on water availability. When ground water is used for water supply, a concern is the potential energy in the aquifer as reflected in the water level in the producing well or neighboring wells. When a ground-water reservoir that does not readily transmit water is tapped, the energy loss associated with flow to the well can be such that the well must be drilled to prohibitively great depths to provide adequate supplies. On the other hand, in a formation able to transmit fluid easily, water levels may drop because the reservoir is being depleted of water. This is generally encountered in reservoirs of limited areal extent or those in which natural infiltration has been reduced either naturally or through human activities.

To forecast reservoir behavior, some type of model of the ground-water system must be employed. Such models may be statistical or deterministic. When problems involving a change in pumping demand are encountered, deterministic models founded on physical principles are generally employed. They can be physical (scale models of the prototype), electrical (based on the analogy between porous media flow and electric current flow), or mathematical (involving analytical or numerical solutions of the governing physical equations).

Early models focused on forecasting well performance. From the point of view of water supply, it was important to forecast how the water level in the pumping well and neighboring wells would respond to pumping. Analytical solutions to the appropriate governing equations were developed for a number of field situations. The solution for the long-term equilibrium case is known as the Thiem equation. The Theis equation is similar, but expresses transient behavior. These, and related expressions, can be readily modified to accommodate a number of possible hydrologic complications. For example, when a well is located in the vicinity of a stream, the stream will, unless hydraulically disconnected from the aquifer, eventually contribute some of its flow to satisfying the well discharge. On the other extreme, when a well is located near the aquifer boundary, the well water levels will be lower than would be encountered had the well been located away from any such impermeable barrier.

These analytical solutions have, in fact, a dual purpose. It is evident that, given appropriate aquifer parameters, these equations can be used to forecast ground-water reservoir response to pumping. A less evident application involves the determination of the aquifer parameters. This normally requires a carefully controlled field experiment called a pumping test, which normally involves a central pumping well and a series of nearby observation wells. Analysis of the response of the ground-water system after the cessation of pumping is known as the recovery test. Both the pumping and recovery tests are analyzed by using the analytical formulas discussed above.

Ground-water quality. Problems involving ground-water quantity were once the primary concern of

hydrologists; interest is now focused on ground-water quality. Ground-water contamination is a serious problem, particularly in the highly urbanized areas of the United States. Legislation restricting the design and construction of waste-disposal facilities has accelerated and expanded ground-water quality studies. Predictive models which describe the convective and dispersive transport of contaminants have been developed and applied to field situations. Such models can be used to investigate the effectiveness of remedial measures designed to contain or remove contaminated ground water. *See* WATER POLLUTION.

Nonaqueous-phase contamination. The recent discovery of ground-water contamination by toxic liquids that are only slightly soluble in water has focused attention on the characteristics and behavior of non-aqueous-phase liquids in the subsurface. A lighter-than-water non-aqueous-phase liquid reaches the water table, whereupon it begins to spread and extend its area (**Fig. 1**). A denser-than-water intruding liquid moves downward through the water table and continues vertically until it encounters an impeding geological horizon (**Fig. 2**). It then begins to move in the direction of the slope of the top of this horizon. Due to the fact that many non-aqueous-phase liquids have low solubility, they present a significant threat to ground-water quality over long periods of time.

Applications. Although this discussion has considered the ground-water phenomenon in a rather narrow sense, the same fundamental principles are applicable to a number of related but distinctly different disciplines. Soil mechanics, the study of foundation engineering, relies heavily upon porous flow physics for its governing equations. This is also true of excavation dewatering in construction and mining, drainage problems in agricultural engineering, and reservoir simulation problems in the oil, gas, and geothermal industries. Ground-water hydrology has played a historically important role in the exploitation of water resources in the United States, and continues to contribute to their protection and effective utilization. *See* HYDROLOGY; SOIL MECHANICS.

George F. Pinder

Access to ground-water resources by wells has enabled the growth of agriculture, population, and commerce throughout the world. The types and uses of wells vary substantially, reflecting the widespread dependence on water supply as well as the efforts to assess and protect the quality and quantity of water for future uses. *See* WELL.

Water supply wells. These serve a variety of water needs ranging from those of an individual home or farmstead to those of urban population centers and industries. Water supply wells are usually designed, constructed, and operated in order to maximize the yield of water. The design factors should include both durable materials and pumping strategies that will provide long-term sustained yield from the aquifer (**Fig. 3**). The diameter of these wells may range from 4 to greater than 24 in. (10 to > 60 cm) depending on the aquifer properties and the pump required for lifting the amount of water needed. The well screen interval is normally placed in the most transmissive (or high-water-yielding) geologic formation available. Above the screen, a well casing made of iron, steel, or a thermoplastic material completes the subsurface portion of the well. Surface completion methods vary; however, the casings are carefully sealed with clay and cement grouting materials to ensure the surface water does not enter the well. In this way, the infiltration of microbial organisms or runoff chemicals from the land surface are prevented from entering the produced water. A pump and water discharge line are contained within the well and are fitted with appropriate valves at the wellhead.

The hydraulics of well operation (Fig. 1) relate to the response of the static water level in the aquifer to pumping. Pumping causes a decline in the water level (drawdown; Fig. 1) to the pumping water level. In certain hydrogeologic settings, the water level

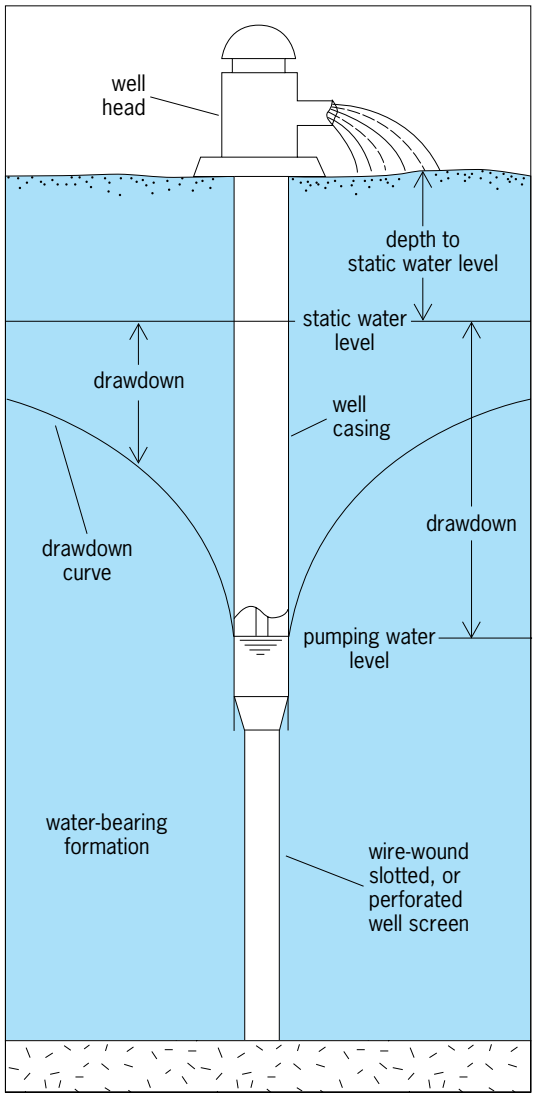

Fig. 3. Diagram of a production well, showing aspects of design that relate to well hydraulic performance. (*After F. G. Driscoll, Groundwater and Wells, 2d ed., Signal Environmental Systems, Johnson Division, 1986*)

in the aquifer may be above the land surface (because the water in the geologic formation is confined or is at higher pressure than overlying formations), and the well may yield water without pumping. Such wells are termed artesian. Once completed, the well would provide water without pumping as long as the pressure differential between the land surface and the screened formation was maintained.

The long-term yield and hydraulic performance of production wells are usually evaluated on the basis of a pumping test, which provides information on the hydraulic properties of the aquifer under pumping conditions. In this type of test, an array of water-level-observation wells are located within the estimated zone of drawdown of the production well. In a typical configuration, there are several small-diameter observation wells screened in the same formation as the production well (**Fig. 4**). One observation well (no. 4 in Fig. 4) would be used to determine if the clay layer shown between 46 and 50 ft (14 and 15 m) below land surface acts as a confining layer between the upper water-table aquifer and the lower aquifer. Water levels would be collected before pumping of the production well was begun, and then at successive time intervals over several days to weeks as the pumping water levels declined from static (nonpumping) levels. The amount of water in storage in the aquifer and the transmissive properties of the geologic formation can then be determined in order to establish a pumping strategy that would provide long-term yield of water without causing excessive drawdown (so-called mining) of the resource.

Monitoring or observation wells. These wells, also known as piezometers, are often used to permit the observation of water levels and to collect samples of ground water for microbiological or chemical analyses. They have many similarities to water supply wells in design and construction. The principal difference between monitoring and production wells is that monitoring wells are most often screened in discrete formations to provide specific hydrogeologic and chemical information rather than to produce large volumes of water. In general, these wells have diameters from 1 to 4 in. (2.5 to 10 cm), and the screened intervals are of the order 2 to 10 ft (0.6 to 3 m). The construction materials may be similar to those used for water supply wells, but designs often call for the use of stainless steel or chemical-resistant thermoplastics for durability under potentially contaminated subsurface conditions (**Fig. 5**). The precautions to seal the monitoring well bore from surface infiltration with bentonite clay and cement may be far more involved than in the construction of a production well. The need to achieve a proper seal becomes all the more important when surface soils or overlying formations may be contaminated. Percolation of contaminated water down the well bore would otherwise bias the analytical results from the well.

Since monitoring or observation wells are designed to provide depth-discrete information, they are frequently built in nested installations. In this situation, two or more wells are finished at different depth ranges to provide information on vertical hydraulic-head (that is, pressure) differences or concentration distributions of chemical or microbial parameters.

Water sample collection is conducted in monitoring wells by small-diameter submersible pumps, surface pumps (if the static water level is within approximately 18 ft or 5.4 m of land surface), or grab samplers. Water-level measurements are made during sampling events to record water levels relative to a surveyed benchmark for analysis of ground-water flow velocity and direction.

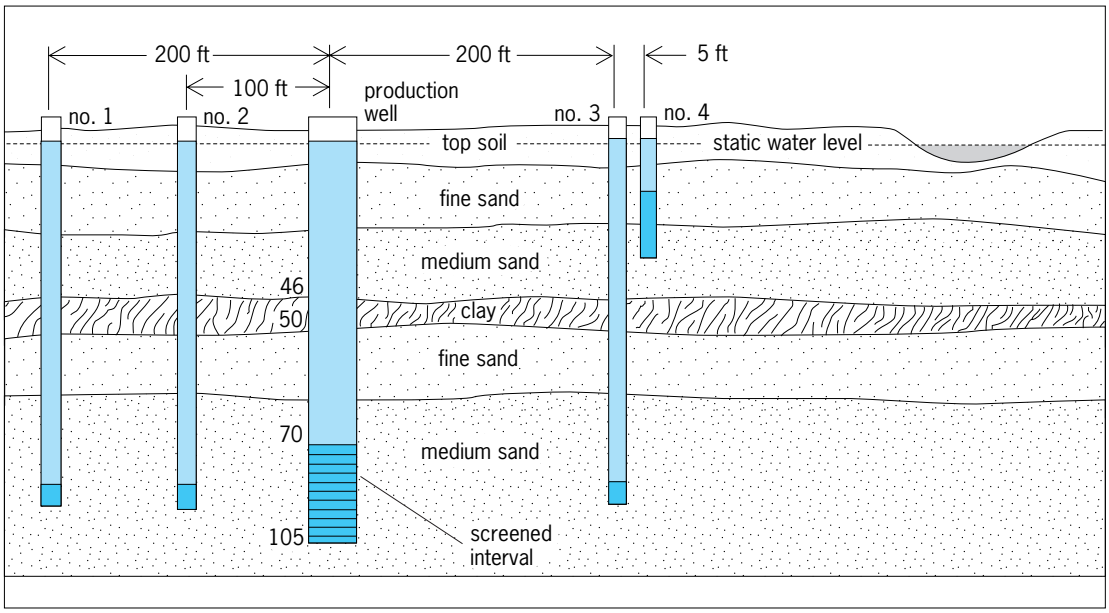

Fig. 4. Geologic section on a line through a test well and observation wells. Numbers along the production well indicate depth in feet. 1 ft = 0.3 m. (*After F. G. Driscoll, Groundwater and Wells, Signal Environmental Systems, Johnson Division, 2d ed., 1986*)

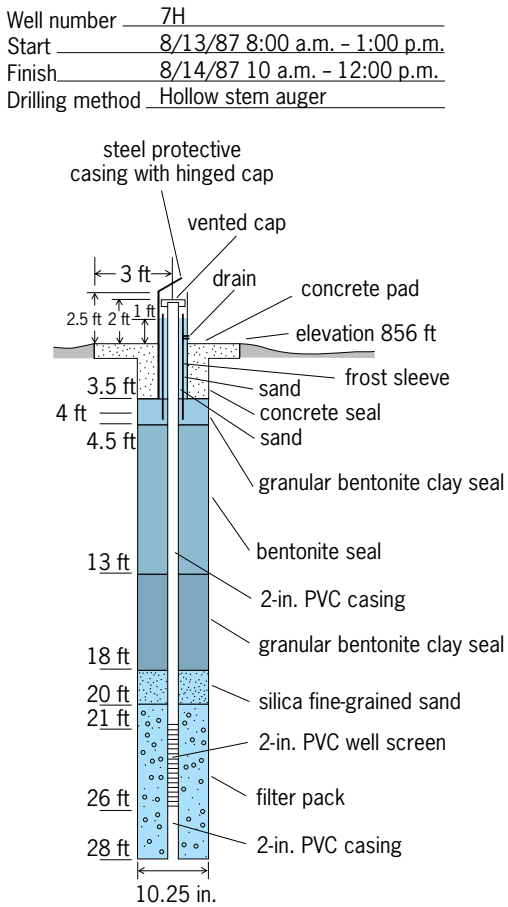

Fig. 5. Design detail for a typical monitoring well. (*After L. Aller, Handbook of Suggested Practices for the Design and Installation of Ground-Water Monitoring Wells, National Water Well Association, 1989*)

Alternative designs for nested monitoring installations include multilevel sampling arrays that employ discrete sampling ports linked to the wellhead by tubing. In these cases, mechanical samplers or surface pumps are used to retrieve water samples. These types of installations can provide significantly more vertical detail in chemical distributions than nested, conventionally screened monitoring wells. However, their use has been linked mainly to shallow depths (less than 100 ft or 30 m below land surface) because of constraints on the size of the sampling mechanism and the lift capabilities. *See* WATER SUPPLY ENGINEERING.

Michael J. Barcelona

Bibliography. M. J. Barcelona et al. (eds.), *Handbook of Groundwater Protection*, 1988; J. Bear, *Hydraulics of Groundwater*, 1980; P. B. Bedient, H. S. Rifai, and C. J. Newell, *Groundwater Contamination*, 1999; R. J. Charbeneau, *Groundwater Hydraulics and Pollutant Transport*, 1999; S. N. Davis and R. J. M. De Wiest, *Hydrogeology*, 1966, reprint 1991; L. Dingman, *Physical Hydrogeology*, 1993; F. G. Driscoll, *Groundwater and Wells*, 2d ed., 1986; C. W. Fetter, Jr., *Applied Hydrogeology*, 1980; R. A. Freeze and J. A. Cherry, *Groundwater*, 1979; M. E. Harr, *Groundwater and Seepage*, 1962, reprint 1992; D. M. Nielsen (ed.), *Practical Handbook of Ground Water Monitoring*, 1991; M. E. Renz, *Practical Groundwater Hydrology*, 1995; U.S. Environmental Protection Agency, *Ground Water Handbook*, vol. 1: *Ground Water and Contamination*, EPA 625/6-90/016a, and vol. 2: *Methodology*, EPA 625/6-90/016b, 1991; F. van der Leeden, F. L. Troise, and D. K. Todd, *The Water Encyclopedia*, 2d ed., 1990.

Grounding

Intentional electric connections to a reference conducting plane, which is typically earth (hence the use of the term "ground," of "earth" in British usage), but which more generally consists of a specific array of interconnected electrical conductors, referred to as the grounding electrode conductor. The symbol that denotes a connection to the grounding conductor is three parallel horizontal lines, the lower two being shorter than the one above (**Fig. 1**). The electric system of an airplane, spacecraft, satellite, or ship observes specific grounding practices with prescribed points of grounding, but no connection to earth is involved in this scheme. A connection to such a reference grounding conductor which is independent of earth is denoted by the symbol shown in **Fig. 2**.

The subject of grounding may be divided into two categories: system grounding and equipment

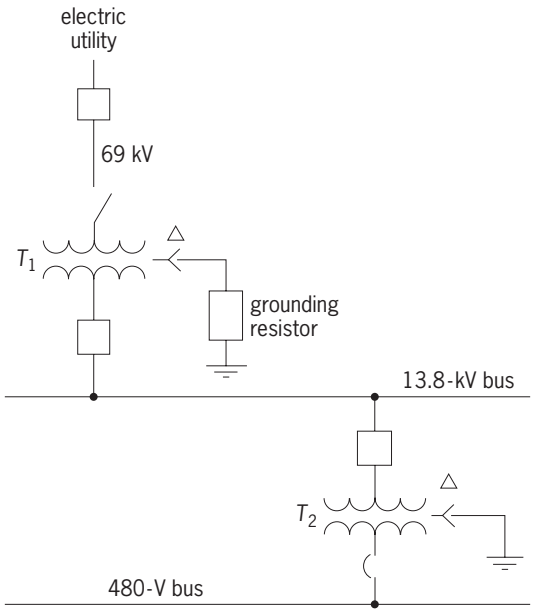

Fig. 1. Each conductively isolated portion of a distribution system requires its ground.

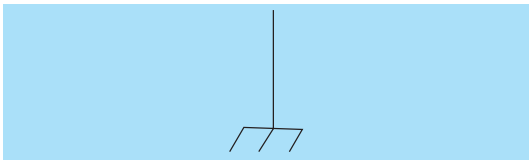

Fig. 2. Symbol to denote connection to a reference ground that is independent of earth.

or "safety" grounding. System grounding refers to an intentional connection from the electric power system conductors to earth for the purpose of securing superior performance qualities in the electrical system. The term "equipment or safety grounding" relates to a grounding connection from the various electric-machine frames, equipment housing, metal raceways containing energized electrical conductors, and closely adjacent conducting structures judged to be vulnerable to contact by an energized conductor. The purpose of such equipment (safety) grounding is to avoid environment hazards such as electric shock to area occupants, fire ignition hazard to the building or contents, and sparking or arcing between building interior metallic members which may be in loose contact with one another. Such ground connections provide a low-impedance path for the return of fault current, which will cause the overcurrent protective device on the affected circuit to "open," thereby removing the hazard. The design of outdoor open-type installations presents special grounding problems.

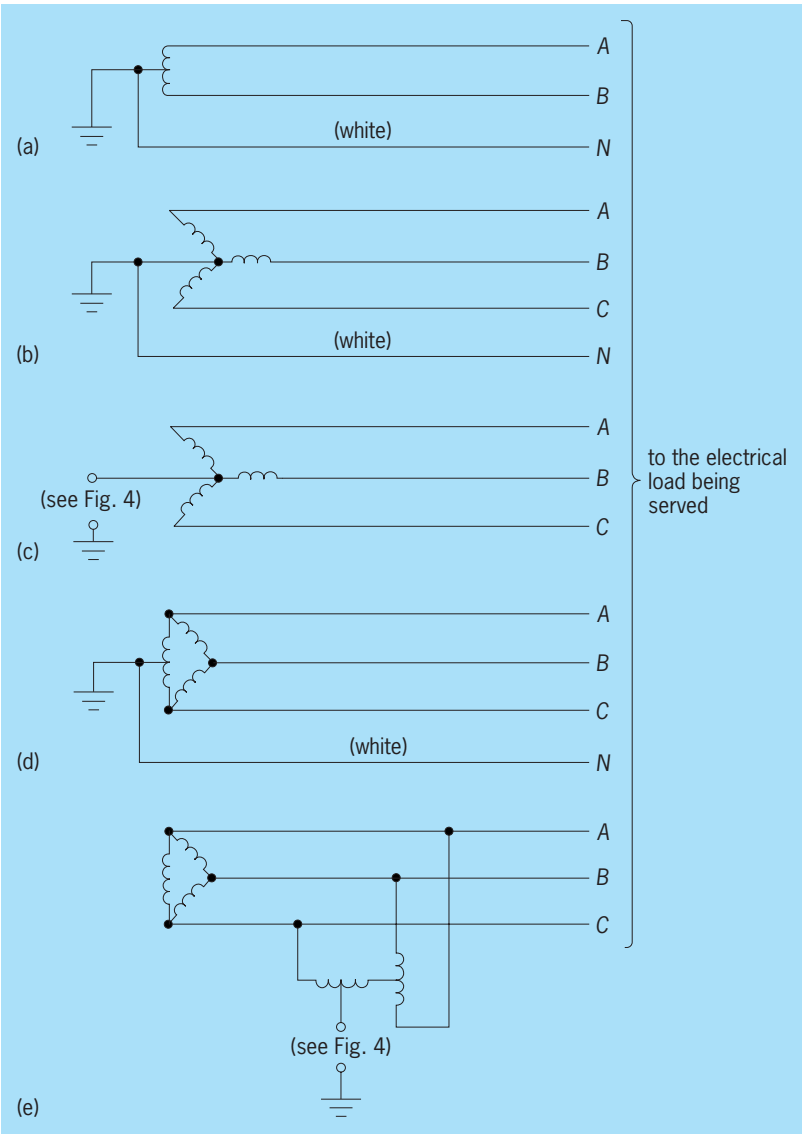

Fig. 3. Commonly used patterns of grounding. (*a*) Single-phase three-wire 240/120-V service. (*b*) Three-phase four-wire 208/120- or 480/277-V service. (*c*) Three-phase three-wire pattern. (*d*) Three-phase four-wire Δ 240-V line-to-line pattern. (*e*) Three-phase Δ with derived neutral.

System grounding. Appropriately applied, system grounding can (1) avoid excess voltage stress on electrical insulation within the system, leading to longer apparatus life and less frequent breakdown; (2) improve substantially the operating quality of the overcurrent protection system; (3) greatly diminish the magnitude of arc fault heat energy released at an insulation breakdown point, lessening arc burning damage and fire ignition possibility; and (4) provide for lightning discharge and minimize damage should high voltages be accidentally impressed on lower-voltage system conductors and equipment.

Sectionalization. Each voltage transformation point employing an insulating transformer interrupts the continuity of the system grounding circuit, and is defined as a "separately derived system" by the National Electrical Code (NFPA-70-2005) [Fig. 1]. System grounding connections made on the 69-kV electric service companies' lines extend their influence only to the 69-kV winding of transformer T_1. The grounding connections established at the 13.8-kV winding of transformer T_1 extend its influence only to the 13.8-kV winding of transformer T_2. The grounding connections established at the 480-V secondary winding of transformer T_2 apply to the 480-V conductor system only. Two distinct advantages result from this. First, the system grounding arrangement of each voltage-level electric system is independent of all others. Second, the type and pattern of system grounding to be used with any individual voltage-level electric system can be tailored to optimize the performance of that particular electric-system section. *See* ELECTRIC POWER SYSTEMS.

It is mandatory to locate the grounding connection at the source-point electrical neutral of the particular voltage-level separately derived system, and mandatory to do so at the service entrance point if the point of origin is outside the local building.

Common patterns. The great majority of system grounding patterns fall into one of the varieties that are shown in **Fig. 3**. The most used varieties of system grounding impedance are illustrated in **Fig. 4**.

The use of solid grounding exclusively for grounding patterns of Fig. 3*a*, *b*, and *d* are influenced by two considerations: (1) Overcurrent protection is present in only the phase conductor of single-phase, one-side grounded circuits. (2) The National Electrical Code (NEC) requires any electric system that can be solidly grounded to limit phase-to-ground potential to not more than 150 V.

Solid grounding is also used almost exclusively in the case of operating voltages of 69 kV and higher, as in Fig. 3*c*, to achieve the most rigid control of overvoltage stress. Such control allows reduced-rating lightning arresters, which in turn permits the successful use of reduced insulation level on station apparatus and equipment.

The high, unrestricted magnitude of short-circuit current created by a line-to-ground (L-G) fault on

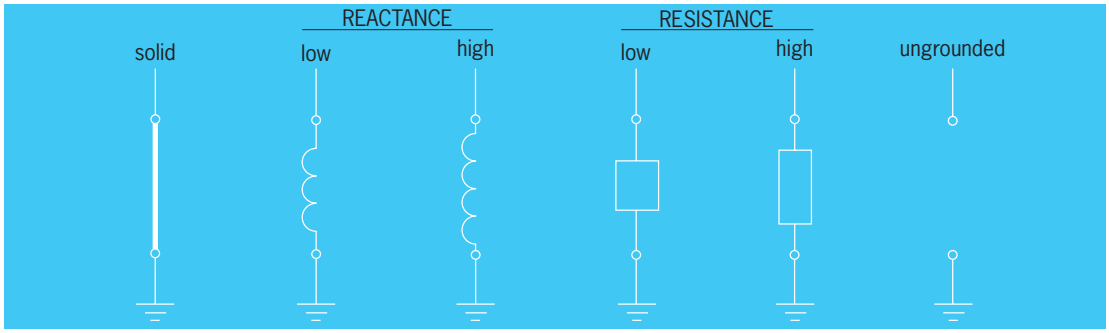

Fig. 4. Varieties of system grounding impedances in common use.

a solidly grounded system can pose severe design problems with costly solutions. The desire to artificially reduce the magnitude of L-G fault current is the chief reason for the use of other grounding impedances.

A low-reactance (inductive) grounding connection impedance can be used to cause a moderate reduction in the L-G short-circuit current, particularly for the purpose of avoiding excess short-circuit current flow in a phase winding of a rotating machine or of accomplishing a desired distribution of neutral unbalanced load current among source machines. Using reactance to achieve the reduction in L-G fault current to below about 40% of the three-phase value enters the high-reactance region, which is subject to the generation of damaging transient overvoltages as a result of a ground fault condition, unless appropriate resistive damping circuits are added. Only one special case of high-reactance grounding is free of overvoltage trouble: the ground-fault neutralizer case, in which the reactance magnitude is carefully matched with the electric system L-G capacitance to create a natural frequency of oscillation almost exactly equal to power system operating frequency. It is critical, however, because a small deviation from the resonant value will destroy the overvoltage immunity. The need for impedance grounding, in general, has been somewhat reduced with the introduction of overcurrent protective devices with ratings of up to 300,000 AIR (ampere interrupting rating).

Resistance grounding. By substituting a resistive grounding impedance (a totally dissipative impedance), much greater reductions in the L-G fault current can be intentionally created without danger of transitory overvoltages.

The low-resistance region is characterized by an established level of available ground-fault current well below the three-phase fault value yet ample to properly operate protective devices responsive to ground-fault current flow. Typical current values in use range downward from a few thousand amperes. Present-day protective practices allow the current value to be set at 400 A for general purpose medium-voltage electric systems widely used in industry. The far more critical electric-shock hazards incident to electric power supply to portable excavating machinery have led to the selection of a much lower level of available ground-fault current, typically in the 25–50-A region.

Most electrical breakdowns occur as between line and ground, and many remain so throughout the interval of detection and isolation. A summary of the operating advantages achieved by intentional reduction of available ground-fault current is given below and illustrated in **Fig. 5**. (1) Low heat-energy release at the fault location because of the low current magnitude. (2) No noticeable dip in the system line-to-line voltages, which means no disturbance to the operation of all healthy load circuits. (3) Diminished interrupting duty on the circuit interrupter (low current and high power factor), which contributes to infrequent maintenance requirements. (4) Diminished duty imposed on the equipment grounding conductor network, which allows superior performance achievement at lower cost.

High-resistance grounding relates to a mode of operation in which the fault location and subsequent corrective action are undertaken manually by skilled maintenance personnel. It is used principally in electric systems that serve critical continuous-process machines.

The available ground-fault current is reduced to the same order as the electric system charging current to ground (generally less than 5 A). This resistive component of fault current is sufficient to arrest the generation of transient overvoltages by L-G

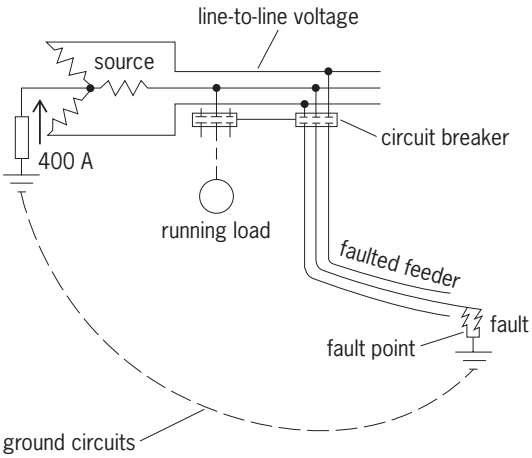

Fig. 5. Resistance grounding. (*General Electric Co.*)

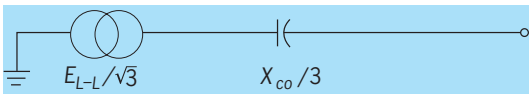

Fig. 6. Equivalent supply circuit to one phase conductor (relative to ground) of an "ungrounded" electric power system. E_{L-L} = line-to-line voltage; $X_{CO}/3$ = capacitive reactance coupling to ground.

fault disturbances and provides a positive signal for identifying the presence of an L-G fault somewhere on the system. Successful results depend upon the presence of a skilled maintenance crew who can respond quickly to the ground-fault alarm, promptly locate the faulty circuit element, and take effective action to remove this faulty circuit from the system before a second insulation failure is induced.

Ungrounded system. Although a system may have no intentional grounding connection, and hence is named an ungrounded system, it is in fact unavoidably capacitively grounded, which is the reason that a fatal shock can be received when a person contacts an equipment enclosure and a phase conductor simultaneously. The layer of insulation surrounding every energized conductor metal constitutes the dielectric film of a minute distributed capacitor between power conductors and ground. In the aggregate this capacitance can amount to a substantial fraction of a microfarad for the complete metallically connected system. Surge voltage suppression filters typically include line-to-ground-connected capacitors which add to the distributed capacitors inherent in the system proper. Unless modified by the stabilizing qualities of grounding connections previously described, L-G fault disturbances can create dangerous overvoltage transients (impressed on system insulation) in about the same ways possible had the grounding connection impedance of Fig. 4 been a capacitor. The equivalent circuit of one phase conductor of an ungrounded three-phase power supply (relative to ground) takes the form illustrated in **Fig. 6**.

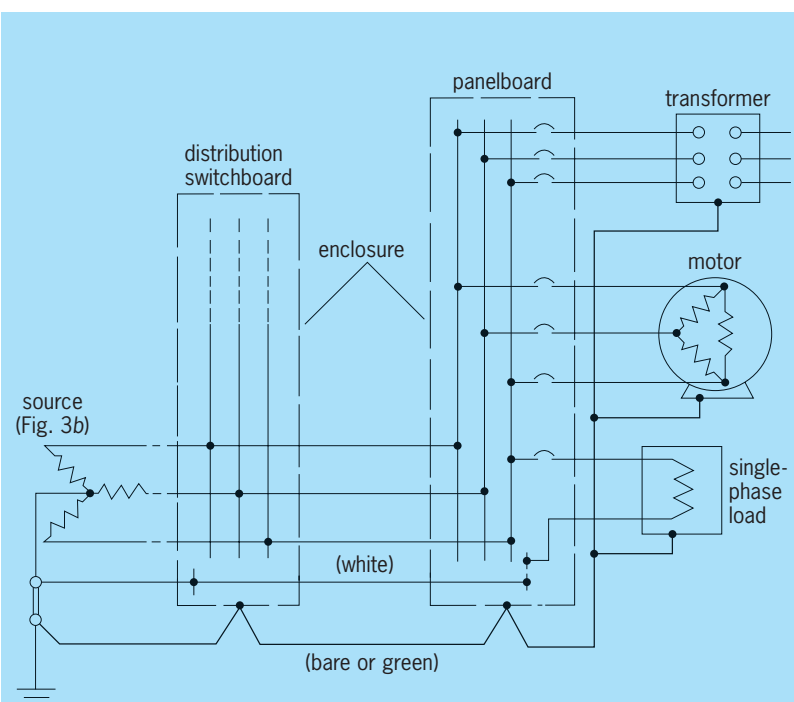

Fig. 7. Heavy line shows typical equipment grounding conductor.

Equipment (safety) grounding. At each unit of electrical equipment, safety grounding serves to establish a near-zero potential reference plane (even during L-G fault conditions). This reference extends to the outer reaches of the particular voltage-level electric system to which a solid equipment (safety) grounding connection can be made from the metal frames of served electric machines, the metal housings that contain switching equipment or other electric-system apparatus, and the metal enclosures containing energized power conductors; these enclosures may be metal cable sheaths or metal raceways.

The purposes of this interconnected mesh conducting network, drawn as the heavy lines in **Fig. 7**, are listed below.

1. To avoid electric shock hazard to any occupant of the area who may be making bodily contact with a metallic structure containing energized conductors, one of which has made an electric fault connection to the mentioned structure.
2. To provide an adequately low impedance to the return path of L-G fault current so as to cause automatic operation of the affected circuit's overcurrent protective device and not to interfere with the operation of system overcurrent protectors of unaffected circuits.
3. To provide ample conductivity (cross-sectional area) to carry the possible magnitude of ground-fault current for the duration controlled by the overcurrent protectors in the electric system.
4. To avoid, by installation, with appropriate geometric spacing (relative to phase conductors) dangerous amounts of ground-fault current diverted into paralleling conductive paths.

The fast-growing use of low-signal-level input high-speed electronic information technology systems places added emphasis on the need to minimize the transmission of stray electrical noise from the electric power circuits to the surrounding space in which these critical equipments are located. The presence of fast-acting solid-state switching devices among the electric-system switching components can aggravate the problem by intensifying the amount of high-frequency disturbance present with the electric-system voltage carried by the power conductors.

The complete metal enclosure of the electric power system, as shown in **Fig. 8**, contributes immensely to the elimination of electrical noise by reducing radiated interference. The NEC requires that all such metal enclosures be interconnected to form an adequate continuous electrical circuit, and also that they be grounded (with some exceptions). The effect of this construction is to enclose the entire electric power system conductor array within a continuous shell of grounded metal that functions

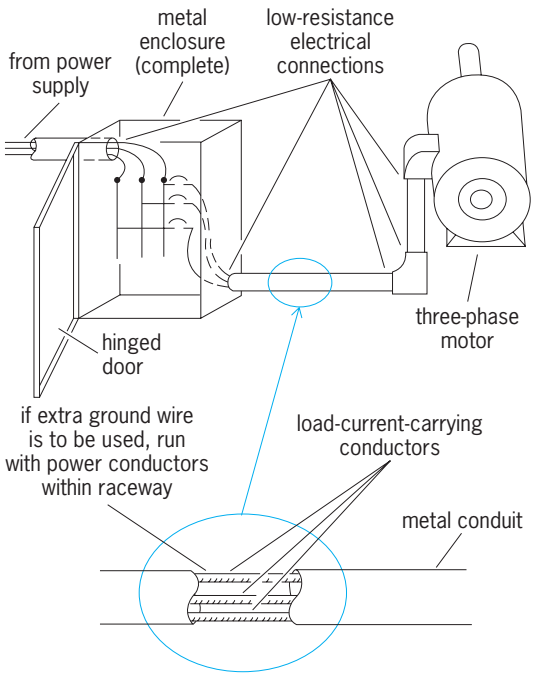

Fig. 8. Complete enclosure of power conductors within a continuous grounded steel shell confines electric and magnetic effects of power currents.

as a Faraday shield to confine the electrostatic and electromagnetic fields associated with the power conductors to the space within the metallic shell. The contribution of electrical noise external to the enclosures is reduced to almost zero. *See* ELECTROMAGNETIC COMPATIBILITY.

To avoid a by-pass circuit by which power-conductor noise voltages might be conductively transmitted to the outside of the metallic enclosure, a careful check of the integrity of the insulation of the grounded power conductor (white wire) throughout the building interior is warranted. The NEC prescribes that the power-system grounded conductor be connected to the grounding conductor at the point of service entrance to the building, and at no other point within the building beyond the service equipment, with some exceptions (such as a separately derived system). Only if a supply feeder extends from one building into another is it permissible to reground the white wire, and then only at the point of entrance to the second building. *See* ELECTRICAL CODES; WIRING.

Protection considerations. The importance of limited magnitude L-G fault current in easing the problem of electric shock exposure control will be evident from the discussion below. The NEC contains a mandatory requirement for installation of automatic ground-fault-responsive tripping of the power supply at the service equipment for all solidly grounded wye-connected electric services of more than 150 V L-G but not those exceeding 600 V line to line (L-L), for each service, main building, or feeder disconnecting means rated 1000 A or more. In the interest of assuring superior electric shock protection, Section 210.8 of the NEC mandates the installation of ultrasensitive ground-fault-responsive tripping features for personal protection (type-GFCI personal protectors) on certain 120-V grounded circuits. Protection is accomplished by deenergizing the supply power to the receptacle if more than 5 milliamperes of the circuit current becomes diverted to a return path other than the grounded power conductor of the circuit.

Grounding of outdoor stations. Installations where earth is used as a reference ground plane present special problems. It is difficult to design an earth "floor surface" for an outdoor open-type substation that will be free of dangerous electric shock voltage exposure to persons around the station.

Personal electric shock danger. A beginning step in dealing with the shock hazard is to establish the magnitude of electric shock exposure which can be accepted by a person without ill effect. The evaluation of dangers from electric shock can be aided by the simple "person model" (**Fig. 9**). The organs which can be vitally affected are the heart and lungs. Lung muscles can be propelled to a tight closed spasm, shutting off the respiratory action. However, the lungs resume normal action once the severe shock voltage is removed. On the other hand, a single short-time excess-value electric shock incident may throw the heart muscles into a state of fibrillation from which they are unable to automatically recover normal action, resulting in death within a few minutes unless defibrillation is accomplished.

Electric shock is communicated to the body via electrical contact to the body extremities (Fig. 9), hand to hand, hand to foot, or foot to foot, each of which creates nearly the same shock intensity at the central chest area. The hand-to-foot case applies to a person standing on the substation floor and touching nearby conducting parts of the station with the extended arms; this applied shock hazard is named the E_{touch} exposure. The foot-to-foot case applies to a person walking across the substation floor during which activity the person's two feet, with considerable separation distance between, contact the substation surface; this shock hazard is named the E_{step} exposure. The effect of a surface layer of higher-resistance material over the active area of the station may be represented by an added resistance element R_{fc} (Fig. 9).

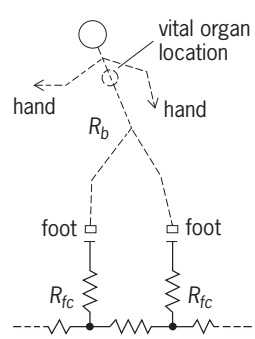

Fig. 9. A simple "person model" to be used in the evaluation of electric shock voltage-exposure intensity.

The basic limits of tolerable electric shock exposure are expressed by Eq. (1), where i is the current

$$i = \frac{0.116 \text{ ampere}}{\sqrt{t}} \tag{1}$$

magnitude through the body and t the elapsed time in seconds, or, in similar form, by Eq. (2).

$$i^2 t = 0.0134 \text{ ampere}^2 \text{ seconds} \tag{2}$$

The sources of potential electric shock hazard are commonly expressed in terms of volts (rather than amperes), but the limiting shock exposure tolerance can be converted to terms of voltage and time if the body resistance R_b (Fig. 9) is known. The value of R_b may go up into the megohm region in the absence of excess moisture or perspiration, but drops drastically when these substances are present on the skin. Assuming the lower limit of the body resistance to be 1000 ohms, the maximum allowable impressed body voltage V_s is given by Eq. (3).

$$V_s = i(R_b) = i(1000) \text{ volts} = \frac{116 \text{ volts}}{\sqrt{t}} \tag{3}$$

Use of ground rods. The voltage exposure values of both E_{touch} and E_{step} depend heavily on the magnitude of voltage gradients along the surface of the outdoor station floor. An outdoor station might receive incoming power at 69 kV and an L-G circuit fault at this terminal would result in the injection of 3000 A into the grounding conductor at the station. During this current flow the station grounding conductor system could rise above mean earth potential by as much as 2500 V or more. The resulting potential distribution pattern across the station floor must be evaluated and tailored to achieve the limiting values of E_{touch} and E_{step}.

As a first step, one may consider driving a standard ground rod (8 ft or 2.4 m long, 0.75 in. or 19 mm in diameter) in the center of the substation floor area. In a typical soil makeup this might be found to establish a 25-ohm connection to earth. It takes only 100 A injected into such a connection to produce a voltage drop of 2500 V, about half of which takes place within a radius of about 2.5 ft (0.8 m). As a voltage probe is moved away from the rod on the station surface, the potential drops rapidly in a conical pyramidal fashion.

The earth-connection properties of a single driven ground rod are thus controlled, almost totally, by a small cylinder of earth, about 10 ft (3 m) in diameter, immediately surrounding it. The resistance value is usually substantial (on the order of 25 ohms) and is reduced only slightly by a second ground rod within the influence zone of the first. A low-value-resistance connection to earth therefore requires a multiplicity of distributed ground rods, spaced to be nearly independent of the influence fields of each other.

To meet the electric shock safety requirements in open-type outdoor stations usually requires an elaborate array of metallic grounding conductors buried in the surface soil of the station area. The array is commonly composed of parallel horizontal conductors with perhaps 8 ft (2.4 m) horizontal spacing at a depth of some 1 to 2 ft (0.3 to 0.6 m) below the surface (**Fig. 10***a*). The geometry of this conductor system usually matches that of the station structure and makes use of the below-grade concrete reinforcing steel involved in the construction of the station. The ground-surface voltage, within one specific bay, relative to the station grounding conductor potential and above the mean earth potential, resembles that shown in Fig. 10*b*. An artificially controlled level of available ground-fault current at 400 A is common in industrial electric power system design, in contrast to a value of 3000 A, not uncommon elsewhere, representing a 7-to-1 advantage.

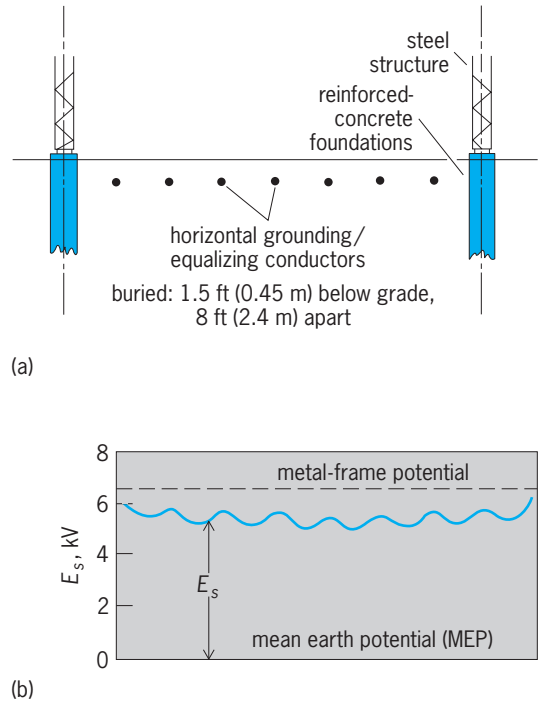

Fig. 10. Surface-potential control in an earth-floor electrical station obtained by a mesh grid of buried grounding equalizing conductors. (*a*) Physical pattern of conductors in one bay. (*b*) Station surface potential E_s above mean earth potential during a local L-G fault.

Hazards near station boundary. There is the possibility that unacceptable levels of electric shock voltage exposure E_{step} and E_{touch} may be found in some places external to the outdoor station area. To ensure the absence of electric shock danger to persons who may frequent these areas adjoining the substation, it is important to check out suspect spots and institute corrective measures as necessary.

R. H. Kaufmann; Brian J. M. Partland

Bibliography. *IEEE Guide for Safety in AC Substation Grounding*, IEEE Stand. 80–2000, 2000; *IEEE Recommended Practice for Electric Power Systems in Commercial Buildings*, IEEE Stand. 241–1990, 1990; *IEEE Recommended Practice for Grounding of Industrial and Commercial Power Systems*, IEEE Stand. 142–1991, 1991; National Fire Protection Association, *National Electrical Code*, 2005; R. P. O'Riley, *Electrical Grounding*, 6th ed., 2002.

Group theory

Any set of elements which is equipped with an operation (called multiplication) satisfying the requirements (1), (2), (3), and (4) is called a group. Group theory is the branch of mathematics devoted to the properties of groups. Group theory has applications in many branches of physical sciences, and in some branches of algebra and in analytic function theory.

Four requirements on group operation. The operation of multiplication is supposed to satisfy the following requirements.

Closure. The product of two elements, for example, g_1 and g_2, in the group G gives another element g_3 (called their product) in the group. That is, requirement (1) is satisfied, where $g \in G$ means that g is an element of G.

$$\text{For any } g_1\, g_2 \in G, \qquad g_1 g_2 = g_3, g_3 \in G \tag{1}$$

Associative law. If g_1, g_2, and g_3 are any elements of the group, then Eq. (2) is satisfied; that is, in forming a

$$(g_1 g_2) g_3 = g_1 (g_2 g_3) \tag{2}$$

product of three elements the same result is obtained by first multiplying g_1 and g_2 and then multiplying the result and g_3, as first multiplying g_2 and g_3 and then multiplying g_1 and the result.

Existence of an identity element. There is in the group an element e (called the identity) which satisfies Eq. (3) for every element g of the group.

$$eg = ge = g \tag{3}$$

Existence of inverses. For each element g of the group there is an element g^{-1} (called the inverse of g) which satisfies Eq. (4).

$$g^{-1} g = g g^{-1} = e \tag{4}$$

Examples. An example of a group is the set of positive real numbers equipped with the operation of ordinary multiplication. The identity element is then the number 1 and the inverse of an element is its reciprocal. Another example of a group is the set of all real numbers equipped with the operation of ordinary addition. The identity element is then the number 0, and the inverse of an element is its negative.

Both these examples are commutative groups (also called abelian groups after the mathematician Niels Abel) because their multiplication law satisfies Eq. (5) for every g and h in the group.

$$gh = hg \tag{5}$$

An example of a noncommutative group (also called nonabelian) is the group of rotations of a three-dimensional rigid body around a point. Here an element of the group is a rotation, that is, the act of rotating the body around a certain axis in space by a certain angle. The group operation means merely successive transformation; $g_1 g_2$ stands for the act of rotation obtained by carrying out the acts of rotation g_2 and g_1 successively in that order. The identity element is the act of rotating through zero angle or making no rotation. The inverse of an element is the rotation around the same axis by the same angle but in the opposite sense. It is easy to check that if g_1 is a rotation around the vertical by 90° and g_2 a rotation around some horizontal axis by 90°, that $g_1 g_2 \neq g_2 g_1$, so the group is noncommutative.

Groups may have a finite or infinite number of elements. Furthermore, the elements may be discrete, or form a continuum in the topological sense described below. A simple example of a discrete finite group is the group with two elements, the identity e and another e_1 satisfying $e_1 e_1 = e$. All other examples given above are of groups with an infinite number of elements, that is, infinite groups; the first two are examples of discrete infinite groups, while the third (the group of three-dimensional rotations) is an example of a continuous group.

Topological groups. For infinite groups, it often occurs that a group has a natural geometry. For example, the group of positive real numbers described above has the geometry of the real numbers. This geometry is compatible with the group operations in the sense that the product gh is a continuous function of g and h, and the inverse g^{-1} is a continuous function of g. Generalizing this scheme to arbitrary groups leads to the notion of a topological group (or less frequently a continuous group) which is a set of elements equipped not only with a group operation but also with a topology, the two notions being required to be compatible in the above sense. A topology can be specified in a number of ways; the net effect of any of them is to give meaning to the phrase: Two group elements are close to one another. For example, in the case of the rotation group discussed above, a topology can be defined by specifying that two rotations are close to one another when their axes have nearly the same direction and their angles of rotation differ by very little. *See* TOPOLOGY.

There is a particular class of topological groups which is of great importance. These are the so-called Lie groups (named after the mathematician Sophus Lie). These are topological groups in which it is possible to label the group elements by a finite number of coordinates in such a way that the coordinates of gh and g^{-1} are analytic functions of the coordinates of g and h and of g, respectively. That means that they should be representable as convergent power series in those coordinates. Most of the topological groups occurring in applications are Lie groups. One important example of a group which is not a Lie group (because its elements cannot even be labeled by a finite number of coordinates in such a way as to make the group operation continuous) is the group of all permissible coordinate transformations in the general theory of relativity. The significant transformations in the special theory of relativity, the so-called Lorentz transformations, do form a Lie group. *See* LIE GROUP; RELATIVITY.

Arthur S. Wightman

Applications to physical sciences. Group theory, the mathematical theory of symmetry, has many

applications in physics and allied areas of chemistry and materials science. Some major ways in which group theory is used in physics will be discussed in conceptual terms. For a more systematic survey of the physically important symmetry groups *see* SYMMETRY LAWS (PHYSICS)

Classification. A symmetry may be so powerful that its possible realizations can be classified. To accomplish such classifications is the task of the theory of transformation groups. An outstanding example is the classification of possible forms of crystals. In the nineteenth century N. Fedorov demonstrated that there are exactly 230 essentially different possible transformation groups of crystal lattices. The two-dimensional version of this problem (which is also of physical interest) can be interpreted as the problem of classifying the essentially different periodic patterns of planar ornaments, and there are exactly 17 such patterns. *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Reduction of variables. If a physical system is invariant under some group of transformations, then its behavior generally depends upon fewer variables than might appear in general. For instance, if the potential energy due to the interaction between two particles depends on their positions, then in general it depends upon six variables, namely three coordinates for the position of each particle. However, if this energy does not depend upon the absolute position of these particles in space, but only on their relative position—that is, it is invariant under the group of spatial translations—then this energy can depend only on three parameters, namely the three coordinates of the relative position or, equivalently, the three components of the vector connecting the particles. Further, this energy does not depend on the absolute orientation of the particles—that is, it is invariant under the group of spatial rotations—then it can depend on only one parameter, namely the distance between the particles. Thus, exploiting symmetry can drastically simplify the description of the problem.

In this simple example, the precise reduction in complexity entailed by the symmetries can be inferred rather easily. In more complex cases, it becomes all but essential to use systematic mathematical procedures to carry out similar reductions. The general idea is that physical behavior can depend only on quantities that are invariant under the symmetry transformations; and the relevant branch of group theory, which provides systematic procedures for identifying such quantities, is called invariant theory.

Spectroscopy. A major problem in many branches of quantum mechanics is to find the energy levels of the system under study. Group theory is invaluable in predicting such energy levels. For if the fundamental laws describing a physical system have a certain symmetry, then states related by symmetry operations will all have the same energy; and if the fundamental laws are only approximately symmetric, or if some external condition breaks the symmetry, then these states will have nearly but not quite the same energy. The transitions between a cluster of states having nearly but not quite the same energy, and some other given state, will then give rise to a cluster of closely spaced spectral lines. The laws of group theory make it possible to anticipate the patterns of the closely spaced clusters of spectral lines that result. *See* ENERGY LEVEL (QUANTUM MECHANICS); NONRELATIVISTIC QUANTUM THEORY.

For example, the fundamental laws governing a hydrogen atom are very nearly invariant under transformations that rotate the electron's spatial position, while leaving the direction of its spin unchanged. However, a more accurate treatment shows that there are a number of relatively small effects which are not invariant under this transformation. Spectral features that at low resolution appear to be single lines will therefore, when examined more closely, exhibit fine structure, whose details depend in very direct and characteristic fashion upon the symmetries of the states involved. *See* ATOMIC STRUCTURE AND SPECTRA; FINE STRUCTURE (SPECTRAL LINES).

Another important situation that frequently occurs is that a transition between two specific energy levels is forbidden because they differ in symmetry. Such selection rules were very important historically in establishing the laws of quantum mechanics and testing their validity. In modern applications the procedure is often reversed, and the symmetry of the states is inferred from the observed selection rules. This information, in turn, is invaluable in forming models of the structure and dynamics of the system under consideration. *See* SELECTION RULES (PHYSICS).

Important examples of this procedure abound. In elementary particle physics, the systems of interest were, by a complicated set of indirect inferences, proposed to be composites of three quarks (for baryons) or a quark and an antiquark (for mesons). For these systems, states of definite energy appear as short-lived particles, or resonances, in high-energy collisions. The quark hypothesis, augmented with group-theoretic arguments, enabled deductions of the kind described above for the number of closely spaced resonances and the selection rules for transitions among them. Experiments that confirmed these deductions gave great support to the quark idea. *See* BARYON; ELEMENTARY PARTICLE; MESON; QUARKS.

Similarly, a definite model of the spatial structure of a molecule allows predictions of the possible energy levels of electrons in that molecule, which can be compared with observed spectra. Again, the shell model allows prediction of the possible energies of systems consisting of a given number of protons and neutrons, that is, an atomic nucleus. *See* MOLECULAR STRUCTURE AND SPECTRA; NUCLEAR STRUCTURE.

In applications of group theory to spectroscopy (in the broad sense used above), the mathematical theory of unitary representations plays a central role. *See* SPECTROSCOPY; UNITARY SYMMETRY.

Guide to new laws. In modern physics, symmetry, or group theory, is increasingly used as a guide to

formulating new laws. For example, the discovery that the weak interaction is not invariant under spatial inversion or parity (*P*) was an important clue for formulating the proper theory of this interaction. The operation of combined parity (*CP*), a transformation involving both changing particles into their antiparticles and spatial inversion, is an approximate symmetry. However, some tiny violations of *CP* have also been observed, whose inner meaning has not been completely elucidated. *See* PARITY (QUANTUM MECHANICS); WEAK NUCLEAR INTERACTIONS.

A more abstract and wide-ranging symmetry postulate, gauge invariance, plays a central role in the modern theory of fundamental interactions. Roughly speaking, gauge invariance postulates not only an overall symmetry among different entities, but that the symmetry transformations can be made independently at every point of space-time. When applied to the color symmetry of quarks, this powerful hypothesis leads directly to quantum chromodynamics, which is presently accepted as the fundamental theory of the strong interaction. *See* GAUGE THEORY; QUANTUM CHROMODYNAMICS.

A new form of group, a supergroup, allows more kinds of symmetry transformations than were previously considered. Specifically, supersymmetry makes it possible to transform particles of different spin into one another. Demanding supersymmetry leads to models of elementary particle interactions that are especially attractive in several ways. At present there is no direct evidence that these models are correct, but supersymmetry provides a striking example of how group theory can be used as a guide to suggesting possible new physical laws. *See* SUPERSYMMETRY.

Frank Wilczek

Applications within mathematics. The applications of group theory within mathematics itself are numerous and important. The part of mathematics in which the notion of group was first clearly isolated was the theory of algebraic equations.

Algebra. Here, the basic idea is to associate a group (the so-called Galois group, named after the mathematician E. Galois) with algebraic equation (6) in

$$a_0 + a_1x + \cdots + a_nx^n = 0 \qquad (6)$$

such a way that the structure of the group reflects the type of operations necessary to compute the roots from the coefficients $a_0, \ldots, a_n$. For example, this method is used to obtain a simple proof that for n greater than 4, no solution in roots of the general equation exists. *See* ALGEBRA; EQUATIONS, THEORY OF.

Topology. In topology, groups are used to classify the structure of various geometric objects. The homology and homotopy groups serve this purpose for a manifold. As an example consider the homotopy group. To define it, a point P of the manifold is chosen and all continuous closed paths starting and ending at P are considered. A notion of product of such paths is defined: If x_1 and x_2 are two paths, the product x_1x_2 is the closed path which is obtained by tracing out x_2 and then x_1. The inverse of a path is then the same curve traversed in the opposite direction and the identity path consists of the point P alone. Paths are divided into equivalence classes, two paths lying in the same equivalence class if one can be deformed continuously into the other. The notion of product then extends to equivalence classes, the product of two equivalence classes being obtained as the equivalence class of the product of any two representative paths. With this definition of product, the set of equivalence classes of continuous closed paths starting at P is a group, the (one-dimensional) homotopy group (also called fundamental group). Provided the manifold under discussion is connected (that is, provided any two points of the manifold can be connected by a continuous curve lying in the manifold), the fundamental group is essentially the same, independent of which point is taken as the starting point for the paths. If the fundamental group of the manifold consists of one element only, the manifold is said to be simply connected. *See* MANIFOLD (MATHEMATICS).

Analytic functions. Group theory plays a fundamental role in the theory of analytic functions. This is true not only for the thoroughly studied case of one complex variable, but also for the case of several complex variables, which is less well understood. Only the former will be discussed here. The starting point is the fact that each analytic function has a natural domain of definition, its Riemann surface. There is another Riemann surface associated with it, its universal covering surface, which is simply connected. The universal covering surface may have several sheets for each sheet of its underlying Riemann surface, each point of the latter being replaced by several "lying above" it. Any function which is analytic on the underlying Riemann surface can be regarded as analytic on the universal covering surface; it will take the same values at those points which lie above a given point. The analytic function regarded as defined on the universal covering surface is invariant under transformations which carry the points lying above any point of the Riemann surface into themselves. These transformations form a group. Now a second basic fact is that any simply connected Riemann surface can be mapped in a one-to-one and analytic fashion onto one of three regions: the interior of the unit circle, the complex plane, or the complex plane closed by adding a point at infinity. Thus, an analytic function on an arbitrary Riemann surface can be regarded as an analytic function defined in one of these three regions and invariant under a certain group of transformations. Such functions are called automorphic functions. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES; RIEMANN SURFACE.

Structure of groups. If G is a group and a subset G' of its elements is a group under the same law of multiplication, then G' is called a subgroup of G. Especially significant are the invariant subgroups which have the property that if g is any element of G' and h any element of G, then hgh^{-1} is an element of G'. Let G and H be two groups and suppose that f is a mapping of G into H [that is, a function which for each element, g, of G yields an element $f(g)$ in H]

with the property $f(gh) = f(g)\,f(h)$. Then f is called a homomorphism of G into H and G is homomorphic to H. The subset of G consisting of those elements which f maps onto the identity in H form an invariant subgroup, the kernel of the homomorphism f. If $H = G$ so that f maps G into itself, it is called an endomorphism. If f maps G one-to-one onto H, then it is called an isomorphism. Finally, if f is an isomorphism of G onto G, it is called an automorphism of G. The idea of isomorphism gives precise meaning to the vague notion that two groups have the same structure.

An important concept, also used in applications of group theory to the physical sciences, is that of a representation of a group. This is a correspondence between group elements g and linear operators $M(g)$ satisfying Eq. (7). Evidently, a representation of a

$$M(g_1)M(g_2) = M(g_1g_2) \tag{7}$$

group is a homomorphism of the group into the group of linear transformations of a vector space.

For topological groups it is natural to require of a homomorphism that it preserve not only the group structure but also the topological structure; for a topological group, a homomorphism is defined as a continuous mapping, f, of G into H satisfying $f(gh) = f(g)f(h)$. The group structure and the topological structure of a topological group impose severe restrictions on each other. Evidence for this statement is the solution of Hilbert's fifth problem (posed in 1900 and solved about 50 years later), which may be stated roughly: Every topological group whose elements can be labeled by a finite number of coordinates so that the group operations are continuous is isomorphic to a Lie group.

For some topological groups it is possible to define a notion of integration on the group which has invariance properties under group multiplication. Namely, if F is a complex-valued function defined on the group and $\int F(g)\,d\mu(g)$ is the integral of F over the group with respect to a measure $d\mu(g)$, then $d\mu(g)$ is said to be left-invariant if Eq. (8) holds for all in-

$$\int F(g)\,d\mu(g) = \int F(hg)\,d\mu(g) \tag{8}$$

tegrable F and all h in the group. Analogously, right-invariance is defined by Eq. (9). When a left-invariant

$$\int F(g)\,d\mu(g) = \int F(gh)\,d\mu(g) \tag{9}$$

measure exists so does a right-invariant, and they are unique up to a multiplicative factor. For a compact group, that is, one for which $\int d\mu(g) < \infty$, left- and right-invariant measures coincide. *See* MEASURE THEORY.

Invariant integration is a powerful tool in the study of representations of groups. Consider, for example, two irreducible representations of a compact group, G, by unitary matrices. (A representation is irreducible if there is no proper subspace of vectors carried into itself by all matrices of the representation. It is unitary if all the matrices are unitary.) Then Eq. (10) holds, where the first alternative holds if

$$\int M^{(1)}(g)_{\kappa\lambda}\overline{M^{(2)}(g)}_{\mu\nu}\,d\mu(g) = \begin{Bmatrix} 1 \\ 0 \end{Bmatrix} \frac{\delta_{\kappa\mu}\delta_{\lambda\nu}\int d\mu(g)}{\dim M^{(1)}(g)} \tag{10}$$

$M^{(1)} = M^{(2)}$ and the second if they are not equivalent. $M^{(1)}$ and $M^{(2)}$ are equivalent if there exists a unitary operator U such that $M^{(1)}(g) = UM^{(2)}(g)U^{-1}$ for all g in G; dim $M^{(1)}(g)$ stands for the number of rows (or columns) in $M^{(1)}(g)$. Angular momentum selection rules of quantum-mechanical systems are consequences of Eq. (10) in the case that G is the three-dimensional rotation group. *See* ANGULAR MOMENTUM; MATRIX THEORY.

For Lie groups there is an important method of analysis, the so-called infinitesimal method. Consider, for example, a Lie group of matrices with matrix multiplication as group multiplication. (Not all Lie groups are isomorphic to such groups, but many are.) A one-parameter subgroup is a subset of elements, $g(t)$, labeled continuously by a real parameter, t, and satisfying Eq. (11) for all real t_1 and t_2. Such a subgroup can be written in the form of Eq. (12).

$$g(t_1)g(t_2) = g(t_1 + t_2) \tag{11}$$

$$g(t) = \exp tx \tag{12}$$

The matrix x is called the infinitesimal element of the one-parameter subgroup. If the group elements are labeled by n parameters, n distinct one-parameter subgroups can be chosen with linearly independent infinitesimal elements $x_1, \ldots, x_n$. These matrices will then satisfy commutation relations (13).

$$x_jx_k - x_kx_j = \sum_{l=1}^{n} C^l_{jk}x_l \tag{13}$$

The constants C^l_{kj} are called the structure constants of the Lie group, and they largely determine the structure of the group. *See* GRAPH THEORY; RING THEORY; SET THEORY.

Arthur S. Wightman

Bibliography. D. Bailin and A. Love, *Supersymmetric Gauge Field Theory and String Theory*, 1994; T. P. Cheng and L.-F. Li, *Gauge Theory of Elementary Particle Physics*, 1988; D. Marris and M. Bertolucci, *Symmetry and Spectroscopy*, 1989; D. J. Robinson, *A Course in the Theory of Groups*, 2d ed., 1996; J. Rotman, *An Introduction to the Theory of Groups*, 4th ed., 1995; D. H. Sattinger and O. L. Weaver, *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics*, 1993; H. Weyl, *Symmetry*, 1952.

Group velocity

The velocity of propagation of a group of waves forming a wave packet; also, the velocity of energy flow in a traveling wave or wave packet. The pure sine waves used to define phase velocity v_p do not ever really exist, for they would require infinite extent. What do exist are groups of waves, wave packets, which are combined disturbances of a group of

sine waves having a range of frequencies and wavelengths. Good approximations to pure sine waves exist, provided the extent of the media is very large in comparison with the wavelength of the sine wave. In nondispersive media, pure sine waves of different frequencies all travel at the same speed υ_f, and any wave packet retains its shape as it propagates. In this case, the group velocity υ_g is the same as υ_p. But if there is dispersion, the wave packet changes shape as it moves, because each different frequency which makes up the packet moves with a different phase velocity. If υ_p is frequency-dependent, then υ_g is not equal to υ_p. *See* PHASE VELOCITY; SINE WAVE; WAVE MOTION.

The relationship between υ_g and υ_p is very easily derived in one dimension. Consider a wave packet made of two waves: the first has wavelength λ and phase velocity υ_p; the second has wavelength $\lambda + \Delta\lambda$ and phase velocity $\upsilon_p + \Delta\upsilon_p$. The wave packet is the combined disturbance which is just the sum of the two waves taken over their infinite extent. The "position" of the wave packet may be taken as the position of maximum disturbance which occurs at some point x_0 at time $t = 0$, as shown in **illus**. *a* at the position of the crests labeled 2. For clarity, the waves are drawn separated. After some time t_0, crest 2 of the second wave will have moved ahead of crest 2 of the first wave, and the crests labeled 1 will be aligned as shown in illus. *b*. The position of the wave packet has moved forward a distance D in time t_0, so the group velocity is $\upsilon_g = D/t_0$. But $D = \upsilon_p t_0 - \lambda$ is evident from the figure, as is $\lambda + \Delta\lambda - \lambda = \Delta\lambda = (\upsilon_p + \Delta\upsilon_p)t_0 - \upsilon_p t_0 = \Delta\upsilon_p t_0$. Therefore, $t_0 = \Delta\lambda_p/\Delta\upsilon_p$, and upon combining these one has $\upsilon_g = \upsilon_p - \lambda(\Delta\upsilon/\Delta\lambda)$. A calculus derivation would yield Eq. (1), which is usually given as the basic relationship. A more convenient form utilizes the equation $\upsilon_p = \lambda f$, where f is the frequency associated with wavelength λ. By the chain rule of differentiation, $d\upsilon_p/d\lambda = f + \lambda(df/d\lambda)$ so $\upsilon_g = \upsilon_p - \lambda f - \lambda^2(df/d\lambda) = -\lambda^2(df/d\lambda)$. But then one uses the wave number $k = 2\pi/\lambda$ to finally write Eq. (2), where $\omega = 2\pi f$ is the angular frequency and $\omega/k = \upsilon_p$. This last form is readily generalized to more than one dimension.

$$v_g = v_p - \lambda \frac{dv_p}{d\lambda} \tag{1}$$

$$v_g = \frac{d\omega}{dk} \tag{2}$$

For waves in deep water, $\upsilon_p = ck^{-1/2}$, where c is a constant. Thus $\omega = ck^{1/2}$ and $\upsilon_g = d\omega/dk = {}^1/_2\,\upsilon_p$. These waves are very strongly dispersive, and the individual wave crests slide through the group at twice the group speed. *See* LIGHT; SOUND; WAVE EQUATION; SURFACE WAVES. S. A. Williams

Bibliography. A. P. French, *Vibrations and Waves*, 1977; K. U. Ingard, *Fundamentals of Waves and Oscillations*, 1988; H. J. Pain, *Physics of Vibrations and Waves*, 5th ed., 1999.

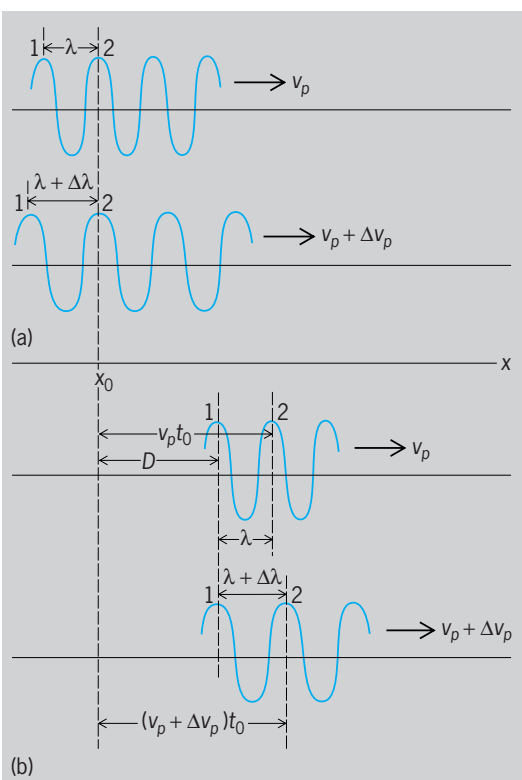

One-dimensional wave packet made of two waves. (*a*) Maximum disturbance of wave packet at time t_0 is located at x_0, where crests labeled 2 coincide. (*b*) After time t_0, wave packet has moved distance *D* to a position at which crests labeled 1 now coincide.

Grout

A binding or structural agent used in construction and engineering applications. Grout is typically a mixture of hydraulic cement and water, with or without fine aggregate; however, chemical grouts are also produced.

The type most commonly specified in construction and engineering is cementitious grout, which is used where its more conventional sister material, concrete, is less suited because of placing limitations or restrictions on coarse-aggregate contents.

Grout can be formulated from a variety of cements and minerals and proportioned for specific applications. Neat cement grout refers to formulations without aggregate, containing only hydraulic cement, water, and possibly admixtures. Sanded grout is any mix containing fine aggregate and it is formulated much like masonry mortar. Whether neat or sanded, cementitious grouts derive their strength and other properties from the same calcium silicate-based binding chemistry as concrete.

Through manual application or pumping, cementitious grouts are used to fill voids and cracks in pavements, building and dam foundations, and brick and concrete masonry wall assemblies; to construct floor toppings or provide flooring underlayment; to place ceramic tile, and to bind preplaced-aggregate concrete. Pressure grouting is employed in raising or leveling uneven slabs, foundations, or structures.

Grouting practice also includes design and placement of flowable fill, a product used for trenches and backfilling. Grouts mixed with special expansive cements are employed for demolition as alternatives to techniques relying on explosive devices or heavy equipment.

Cementitious grouts are either batched in hoppers on site or delivered in truck mixers. Specialty grouts for small-scale applications are prepackaged. Grouts are formulated from either ASTM type I portland cement or other hydraulic cements, depending upon project conditions. Mineral and chemical admixtures, which impart special properties to concrete mixes, are similarly specified for grouts. Mixes are designed for requirements or factors such as minimum compressive strength, placing conditions and consistency, workability and working time, and setting and hardening. Physical and structural properties of the hardened product are predicated on water-cement ratio, cement particle hydration, and the subsequent bonding with fine aggregate or surfaces in contact with grout. *See* CEMENT; CONCRETE.

Don Marsh

Bibliography. S. H. Kosmatka, *Cementitious Grouts and Grouting*, Portland Cement Association, 1990.

Growth factor

Any of a group of biologically active polypeptides which function as hormonelike regulatory signals, controlling the growth and differentiation of responsive cells. Indeed, the distinction between growth factors and hormones is frequently arbitrary and stems more from the manner of their discovery than from a clear difference in function. *See* CELL DIFFERENTIATION; HORMONE.

The discovery of growth factors has been primarily a direct result of the development of techniques for animal-cell culture. Modern tissue-culture techniques permit the removal of a variety of cell types from animals and the growing of the cells for indefinite periods in artificial growth media. However, supplying cells with the nutritional essentials (such as glucose, vitamins, minerals, and amino acids) is not sufficient for growth; cells also require a number of hormones and growth factors. The particular hormone–growth factor combination that is most favorable differs for various cell types. *See* TISSUE CULTURE.

Growth factors are first identified merely as growth-promoting agents; subsequently, they are identified as highly purified discrete molecular species. Generally, growth factors discovered in this way have been polypeptides. Like polypeptide hormones, the action of growth factors on cells depends on binding to specific cell-surface receptors (proteins).

The physiological functions of growth factors can be inferred from their activities in tissue cultures and probably involve the regulation of cell proliferation and differentiation during embryonal development and during wound healing. However, in many cases there has been no direct demonstration of the physiological importance of particular growth factors. Some of the most fully characterized growth factors are described below.

Families. The sequence of amino acids has been determined for several growth-factor polypeptides (see **illus.**). This information permits a number of growth factors to be placed into families, members of which have related amino acid sequences, suggesting that they evolved from a single ancestral protein. The insulin family comprises somatemedins A and C, insulin, insulinlike growth factor (IGF), and multiplication-stimulating factor (MSF). A second family consists of sarcoma growth factor (SGF), transforming growth factors (TGFs), and epidermal growth factor (EGF). In addition, there are growth factors, such as nerve growth factor (NGF), fibroblast growth factor (FGF), and platelet-derived growth factor (PDGF), for which structural homologs have not been identified. *See* INSULIN; PROTEIN.

Nerve growth factor. Nerve growth factor is produced in large amounts in mouse submaxillary gland and in the prostate gland of several mammals; it is probably produced in smaller amounts in other tissues. Nerve growth factor appears to function in promoting survival of, and extension of, axons from nerve cells of the peripheral nervous system. For example, injection of nerve growth factor into newborn mice causes excessive axon growth from sympathetic and sensory nerve cells, while injection of antiserum to nerve growth factor causes these nerve cells to die. The mouse-gland protein which has been most fully characterized exists as a 136,000-molecular-weight aggregate of α, β, and γ subunits. The β subunit, which consists of two noncovalently associated 13,000-molecular-weight peptide chains, contains the full nerve growth factor activity, while the γ subunit is a proteolytic enzyme which is closely related to glandular kallikrein. *See* NERVOUS SYSTEM (VERTEBRATE).

Epidermal and transforming growth factors. Epidermal growth factor, like nerve growth factor, is produced in large amounts in the submaxillary gland of male mice (a fact of uncertain physiological consequence) but is apparently produced by other tissues also. Submaxillary-gland epidermal growth factor, like nerve growth factor, is found in a high-molecular-weight aggregate. The 60,000-molecular-weight epidermal growth factor aggregate contains the 6000-molecular-weight biologically active subunit, as well as a kallikrein-like proteolytic enzyme which is very similar to, but distinct from, the γ subunit of nerve growth factor. Both epidermal growth factor and nerve growth factor are produced by proteolytic cleavage of large precursor proteins, and it appears that the kallikrein-like enzymes perform this function. Epidermal growth factor injected into newborn mice causes premature opening of eyelids and eruption of incisors, suggesting that epidermal growth factor normally functions in promoting growth of epithelial tissue. Transforming growth factors are a family of factors with similar biological activity. Several transforming growth factors have been shown

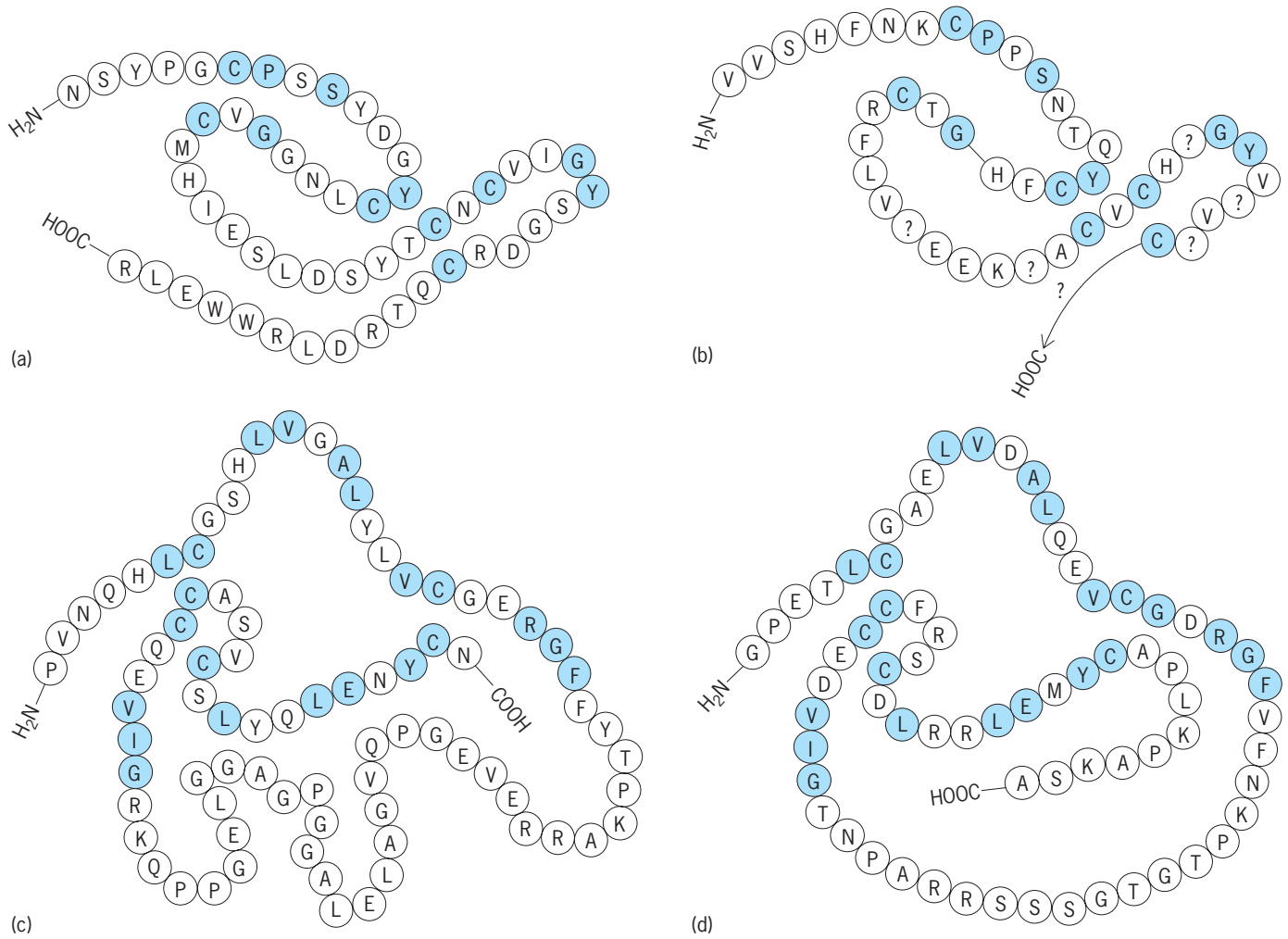

Comparison of the amino acid sequences of (*a*) mouse epidermal growth factor, (*b*) rat transforming growth factor, (*c*) bovine proinsulin, and (*d*) human insulinlike growth factor I. Shaded circles depict positions of structural homology between epidermal growth factor and transforming growth factor and between proinsulin and insulinlike growth factor. Amino acid abbreviations are: A, alanine; C, cysteine; D, aspartate; E, glutamate; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, trytophan; Y, tyrosine.

to be structurally related to epidermal growth factor (illus. *a* and *b*). Transforming growth factors are secreted by a variety of cultured cells, particularly tumor cells, and have the property of causing normal fibroblasts to behave like tumor cells. Consistent with their structural similarity to epidermal growth factor, some transforming growth factors will interact with the epidermal growth factor cell-surface receptor.

Somatemedin growth factors. Insulinlike growth factors I and II are distinct but similar, differing only in a few amino acids. Somatemedins A and C and multiplication-stimulating factor are themselves quite similar to the insulinlike growth factors. They share a number of biological activities, including stimulation of sulfation of cartilage by chondrocytes, stimulation of proliferation of fibroblasts, and insulinlike stimulation of glucose metabolism in fat cells. Insulin is a more distantly related member of this family (illus. *c*). Nevertheless, the relationship is close enough to permit insulin to bind to somatemedin receptors and, at high concentrations, to stimulate proliferation of fibroblasts.

Platelet-derived growth factor. This growth factor is stored in blood platelets and is released during blood clotting. Since platelet-derived growth factor stimulates fibroblast cell proliferation tissue cultures, it is likely that release of platelet-derived growth factor at wound sites may be important for stimulation of wound healing. Fibroblast growth factor has not been so fully characterized as platelet-derived growth factor, but while they share many similarities, it appears to be a distinct protein.

A broader family. While there are no readily apparent structural similarities among epidermal growth factor, platelet-derived growth factor, and insulin, the cell-surface receptor proteins for these growth factors do have remarkable similarities, which suggests that epidermal growth factor, insulin, and platelet-derived growth factor families may themselves be related. The similarity of the receptors lies in the fact that each is a protein kinase (an enzyme which

phosphorylates proteins), and the preferred substrate of the receptor kinase activity is the receptor itself. The receptor protein kinase activities are unusual in preferring tyrosine residues in proteins as substrates (most known kinases prefer serine or threonine). The only other tyrosine-specific protein kinases known are tumor-causing proteins produced by ribonucleic acid (RNA) tumor viruses. *See* TUMOR VIRUSES.

Growth factors and cancer. The stimulation of cell proliferation by several growth factors is similar in some ways to the rapid cell proliferation characteristic of tumor cells. Furthermore, the growth factor receptors are similar to the tumor-causing proteins produced by several RNA tumor viruses. It has been demonstrated that platelet-derived growth factor is virtually identical to the tumor-causing protein of the RNA tumor virus, simian sarcoma virus. Thus, what has long been a subject of speculation can no longer be doubted: some forms of cancer involve improper function of growth factors. *See* CANCER (MEDICINE); ONCOLOGY. Mark Bothwell

Bibliography. H. N. Antoniades and A. J. Owen, Growth factors and regulation of cell growth, *Annu. Rev. Med.*, 33:445–463, 1982; G. Carpenter and S. Cohen, Epidermal growth factor, *Annu. Rev. Biochem.*, 48:193–216, 1979; D. Gospodarowicz and J. S. Moran, Growth factors in mammalian cell culture, *Annu. Rev. Biochem.*, 45:531–558, 1976; L. A. Greene and E. M. Shooter, The nerve growth factor, *Annu. Rev. Neurosci.*, 3:353–402, 1980.

Gruiformes

A highly heterogeneous, worldwide order of field, marsh, and aquatic birds that may be closely related to the shorebirds and their allies (see **illustration**). Diagnosis of this order is difficult because of its great diversity; some taxonomists divide the gruiforms into a number of separate and possibly unrelated orders. Some families, such as the Turnicidae and the Pedionomidae, are considered by some workers to be members of the Charadriiformes. Many of the families contain a small number of species and live in restricted areas. *See* CHARADRIIFORMES.

American coot (*Fulica americana*). (*Photo by Antonio J. Ferreira, © California Academy of Sciences*)

Classification. The order Gruiformes is arranged into 10 suborders and 20 families, as listed below.

Order Gruiformes
- Suborder Mesoenatides
  - Family Mesoenatidae (mesites, 3 species; Madagascar)
- Suborder Turnices
  - Family: Turnicidae (button quail, 14 species; Old World from Europe to Australia)
    - Pedionomidae (plains wanderer; 1 species; Australia)
- Suborder Grues
  - Family: Geranoididae (fossil; Eocene of North America)
    - Eogruidae (fossil; Eocent to Oligocene of Mongolia)
    - Ergilornithidae (fossil; Oligocene to Miocene of Asia)
    - Gruidae (canes; 15 species; worldwide, except South America)
    - Aramidae (limpkin; 1 species; New World tropics)
    - Psophiidae (trumpeters; 3 species; South America)
- Suborder Ralli
  - Family Rallidae (rails, gallinules, coots; 142 species; worldwide, including numerous oceanic islands)
- Suborder Heliornithes
  - Family Heliornithidae (sun grebes; 3 species; pantropics)
- Suborder Apterornithes
  - Family Apterornithidae (fossil; Quaternary of New Zealand)
- Suborder Rhtnocheti
  - Family Rhynochetidae (kagu; 1 species; New Caledonia)
- Suborder Eurypygae
  - Family Eurypygidae (sun bittern; 1 species; New World tropics)
- Suborder Cariamae
  - Family Cunampaiidae (fossil; Eocene of Argentina)
    - Cariamidae (seriemas; 2 species; South America)
    - Bathornithidae (fossil; Oligocene of the United States)
    - Idiornithidae (fossil; Eocene and Oligocene of France)
    - Phorusrhacidae (fossil; Eocene to late Pliocene of Europe to South America, with the greatest radiation from the Oligocene to the late Pliocene of South America)
- Suborder Otides
  - Family Otididae (bustards; 24 species: Old World, including Australia)

Fossil record. The fossil record of some groups within the Gruiformes is well known from the early

Eocene, including the cranes, rails, and seriemas. Other living families have either no fossil record or a poor one. The Apterornithidae of New Zealand are large, flightless birds only distantly related to any other gruiform group. Rails are common as recent fossils on many oceanic islands; many are flightless, as are some living island species. Cranes are well represented as fossils from the early Eocene of the Northern Hemisphere. Some of the Ergilornithidae possess only two toes, which led a few workers to conclude that this family is ancestral to the ostriches, but that idea has been almost universally rejected. *See* STRUTHIONIFORMES.

The most fascinating group of gruiform fossils is found in the Cariamae and includes the large radiation of the giant, flightless, predatory Phorusrhacidae and their relatives. It was centered in South America from early Oligocene to late Pliocene times, but with a late Pliocene form found in Florida and an Eo-Oligocene species discovered in France. The North American Bathornithidae and the European Idiornithidae, known from the late Eocene to the early Miocene, represent a separate northern radiation of the Cariamae. All that remains of these large, diverse radiations are the two living species of seriemas of Brazil and Argentina.

Characteristics. The Gruiformes range from small to very large. Some have short legs and others long, and the bill varies from short and straight to long and decurved. Some have long wings and are strong fliers, whereas others are flightless. Wading, terrestrial, or aquatic, they are usually found in open country, but some inhabit dense forests. The sexes in most species are similar in plumage, and most species are monogamous. The downy young usually leave the nest shortly after hatching and are cared for by both parents. Northern species are migratory; the migratory concentrations and flights of cranes are spectacular. Some forms, including the kagu and a number of rail species, are flightless. Many gruiforms are secretive and fly only when other forms of escape are impossible. Although rails are reluctant to fly, many species undertake long migrations and have colonized many isolated oceanic islands, where they tend to evolve flightlessness.

Despite the fact that cranes are revered as symbols of good fortune and luck by many cultures, most species have been seriously reduced in numbers, mainly by habitat destruction but also by hunting. Most of their 15 species are seriously endangered and are barely maintaining their low numbers despite rigorous conservation efforts. Several species of larger bustards are equally endangered because of habitat loss and high-power lines which are hazardous to the flying birds. Many island rails have become extinct primarily as a result of the introduction of small mammalian predators. *See* AVES; ENDANGERED SPECIES.

Walter J. Bock

Bibliography. J. del Hoyo et al. (eds), Order Gruiformes, *Handbook of the Birds of the World*, vol. 3, pp. 34–273, Lynx Edicions, 1996; P. A. Johnsgard, *Bustards, Hemipodes and Sandgrouse: Birds of Dry Places*, Oxford University Press, 1992; P. A. Johnsgard, *The Cranes of the World*, Indiana University Press, 1983; B. Taylor, *Rails: A Guide to the Rails, Crakes, Gallinules and Coots of the World*, Yale University Press, 1998.

Grylloblattodea

A relatively new order of flightless insects first described from Sulphur Mountain near Banff, Canada, in 1914 by entomologist E. M. Walker. There are now 26 described species of living Grylloblattodea in five genera. Called ice crawlers, they are typically found at high elevations in East Asia, Japan, Siberia, and western North America. *See* INSECTA; ORTHOPTERA.

Description. Ice crawlers are found under rocks, in leaf litter, and occasionally in caves. Some Asian cave dwellers are blind. *Grylloblatta* species were one of the first groups of insects to colonize Mount St. Helens (Washington) within a few years after the 1980 eruption.

Some species appear to require a narrow temperature range: a *Grylloblatta* living on Mount Rainier (Washington) undergoes heat convulsions above about 14°C (57°F), and has a metabolism that functions best in a thermal window between −6 and 12°C (between 21 and 54°F). These insects appear to survive by occupying moderate habitats such

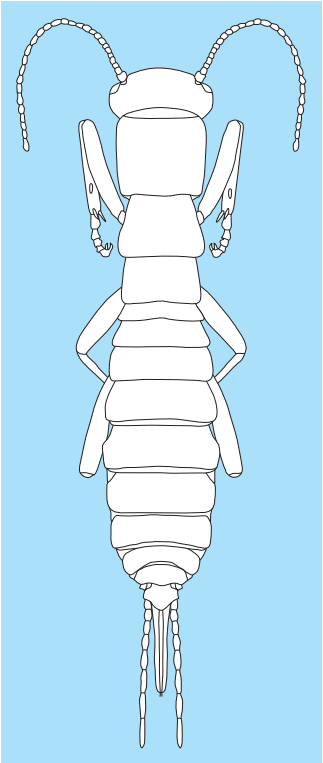

Dorsal view of the first ice crawler species to be described: *Grylloblatta campodeiformis* as. (*Redrawn from the original paper by E. M. Walker, A new species of Orthoptera, forming a new genus and family, Can. Entomol., 46(3):93–99, 1914*)

as those with temperatures moderated by snow cover.

Living Grylloblattodea are defined by a number of unique characters, including an eversible sac on the underside of the first abdominal segment, a spine on the underside of the meta (third) thoracic segment, and distinct genitalic and ovipositor structures (see **illustration**).

Phylogeny. There is some debate about the relationships of Grylloblattodea with both living and extinct insects. With regard to living species, some analyses of molecular (DNA) characters and analyses of a combination of molecular and structural characters show Grylloblattodea as sharing a common ancestry with earwigs (order Dermaptera). More recent molecular analysis pairs Grylloblattodea with the flightless African Mantophasmatodea, named in 2002 and displacing Grylloblattodea as the newest described insect order. One analysis places stick insects (Phasmatodea) in the same group as Grylloblattodea and Mantophasmatodea. This grouping of two to three orders shares a common ancestor with the lineage (Dictyoptera) comprising cockroaches [order Blattaria (Blattodea)] and praying mantids (order Mantodea). *See* BLATTARIA (BLATTODEA); DERMAPTERA.

Likely Jurassic and Permian ancestors of modern Grylloblattodea include fossil insects (some with fully formed wings) that share characters such as ovipositor structure with living species. However, some investigators argue that the order Grylloblattodea includes a much broader diversity of fossil insects representing a group that was once very widespread and common in the Paleozoic and early Mesozoic and whose habitats included much warmer temperate and tropical climes than occupied by modern ice crawlers. This broad view of Grylloblattodea suggests that they were ancestral to orthopteroid orders as well as Dermaptera and Plecoptera (stoneflies). These extinct species included within Grylloblattodea comprised about half of all insects from some Permian sites. Many were consumers of pollen and spores. Other authorities argue that this broad view of the past diversity of Grylloblattodea is based on an inappropriate grouping of unrelated extinct families. *See* PLECOPTERA.

Darryl T. Gwynne; Glenn K. Morris

Bibliography. D. Grimaldi and M. S. Engel, *Evolution of the Insects*, Cambridge University Press, 2005; V. R. Vickery and D. K. M. Kevan, *The Grasshoppers, Crickets and Related Insects of Canada and Adjacent Regions*, Canadian Government Services, Ottawa, 1985.

Guava

A plant, *Psidium guajava*, of tropical America that has long been in cultivation. It is a shrub or low tree which belongs to the myrtle family (Myrtaceae). The fruit (see **illus.**) is a berry, yellow when ripe, and quite variable in size depending on variety and growing conditions, the average being about $2^1/_2$ in. (7.2 cm) long. The guava is quite aromatic, sweet, and juicy. It is used mostly for jellies and preserves, but also as a fresh fruit. *See* MYRTALES.

***Psidium guajava*, showing a branch with leaves and two berries and a berry cut in half.**

Perry D. Strausbaugh; Earl L. Core

Guayule

A desert plant, *Parthenium argentatum*, of the composite family (Compositae), which produces rubber. It is a native perennial shrub growing in the Chihuahuan Desert of north-central Mexico and southwestern Texas. The plant is bushy with dense branches, thick clusters of silverlike leaves, a strong taproot, and a thick crown (see **illus.**). At maturity

Guayule plant, maximum height 40 in. (102 cm).

the plant varies in height from 30 to 40 in. (76 to 102 cm). *See* CAMPANULALES.

Early history. The presence of rubber in guayule was known by the Aztecs and certain North American Indians for several centuries. However, the first published account of guayule was not until 1852, when J. M. Bigelow, a medical doctor, discovered the plant while working with the Mexican Boundary Survey Team. His specimens were sent to the Gray Herbarium of Harvard University, where guayule was first described taxonomically.

Guayule was widely exposed to the American public for the first time in 1876, when the Mexican government sent an exhibit to the Centennial Exposition at Philadelphia. During the next 25 years several companies researched guayule rubber. In 1904 W. A. Lawrence developed the pebble mill extraction method. The Mexicans were very enthusiastic about guayule, and within 5 years over a dozen extraction factories were built to harvest and process guayule from wild stands. By 1910 one-half of the United States supply of natural rubber came from guayule in Mexico.

Breeding and cultivation. The wild stands of guayule were quickly depleted, and with the exhaustion of the natural supply, the Continental-Mexican and intercontinental rubber companies undertook studies to domestically cultivate guayule. Extensive commercial development began in Salinas, California, in 1925, and between 1931 and 1941, 3,068,630 lb (1,391,907 kg) of rubber were milled. During World War II the domestication of guayule became a national issue when the supply of *Hevea* rubber from the Far East was cut off. With the development of synthetic rubber and the end of the war, economic and political incentives for guayule production disappeared, and all projects were terminated by 1959.

Guayule is a tetraploid of 72 chromosomes and reproduces by apomixis. Diploids of 36 chromosomes are also found, and they reproduce sexually. Breeding programs in guayule seek improved varieties with larger rubber yields, increased disease resistance, and greater cold tolerance. There are several closely related species of *Parthenium*, some treelike and some very weedy. Crossing these species with guayule gives wide variation for selection and breeding.

In 1976 research was reinitiated in Arizona and California. The rising price of oil and the accompanying political difficulties with oil-producing countries rekindled interest in guayule as a source of natural rubber. Also, guayule could be very important for agricultural economy in certain parts of Texas, Arizona, and California, where the high summer temperatures, low rain fall, and absence of freezing winters are favorable for guayule cultivation. As water becomes a diminishing resource, a drought-tolerant crop such as guayule could be very important for farmers. It seems possible that if certain problems in cultivation and processing are successfully resolved, guayule could become a major cash crop. *See* RUBBER; RUBBER TREE. David D. Rubis; Linda J. Cassens

Guidance systems

The algorithms and computers utilized to steer a vehicle along a path. The types of vehicles include airplanes, rockets, missiles, ships, torpedoes, drones, and material transport vehicles within factories and so forth. The means of steering depend on the vehicle and can be the rudder, elevators, and other control surfaces on an airplane, the rudder on a ship, the control surfaces on a missile or on a torpedo, the gimbal angle of the motor on a rocket, and others. In every case the guidance system utilizes knowledge of the difference between where the vehicle should be and where it is. The difference between these two vectors is processed by the guidance algorithm. The output is a steering command intended to reduce the error between the desired and the actual paths. *See* DRONE; ELEVATOR (AIRCRAFT); FLIGHT CONTROLS; SHIP POWERING, MANEUVERING, AND SEAKEEPING.

The configuration shown in **illus.** *a* applies where a path or trajectory for the vehicle has been computed and stored in the computer. As time passes, the corresponding values for the trajectory, desired position, and velocity are made available at the summation device. Meanwhile, sensors measure the actual position and velocity of the vehicle. This set of signals is fed back to the summation device and is subtracted from the reference trajectory. The output of the summation device is the difference between these two signals, that is, the error signal. This configuration would be utilized, for example, to boost a satellite into orbit. Inertial navigation utilizing gyroscopes could provide the sensing operation. The role of the guidance system is to process this error signal and use the results to steer the vehicle along the reference trajectory. *See* AUTOPILOT; CONTROL SYSTEMS.

The configuration shown in illus.*b* applies when the vehicle is attempting to track a target which may or may not be moving. The absolute position of the target is not sensed, but rather the difference between the vector from the vehicle to the target and the vector describing the vehicle orientation. Thus the sensor is different from that used in the configuration of illus. *a*. A common example would be a radar rigidly mounted to the nose of a vehicle. Here the sensed target bearing would automatically be measured relative to the attitude or orientation of the vehicle. The role of the guidance system is to process this error signal and use the result to steer the vehicle toward the target. *See* HOMING.

Several important performance attributes contribute to the effectiveness of the system. These attributes are governed by the guidance system and by the other system components, including the vehicle itself and its dynamic behavior.

A primary concern is accuracy, which is certainly a function of the guidance system but under the best of circumstances can be no better than the quality of the sensor. Whether the goal is to insert a satellite into synchronous orbit or to try to intercept an

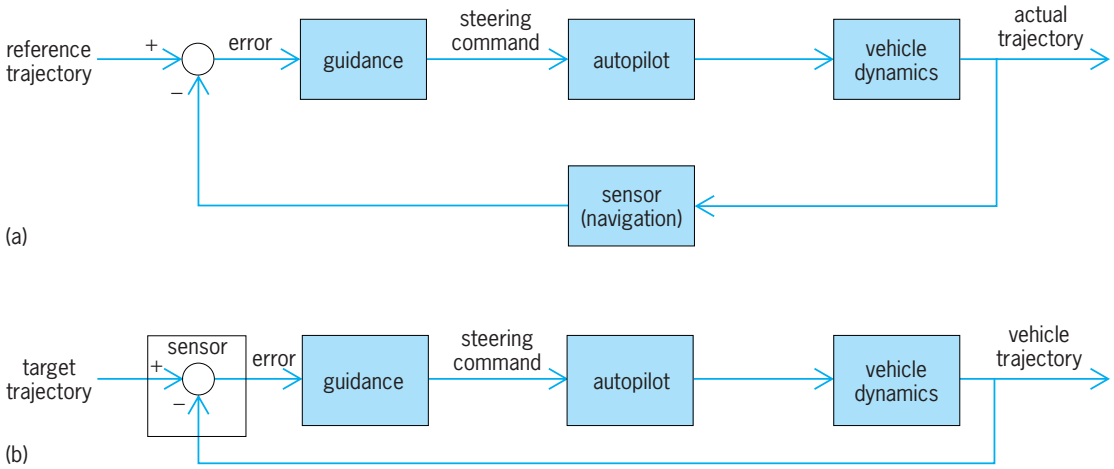

Guidance system for (*a*) following a specified trajectory and (*b*) homing on a target.

enemy aircraft with an air-to-air missile, the accuracy of the sensor and the properties of the guidance system are the principal factors in overall accuracy.

Another concern is speed of response. Here the dynamics of the vehicle itself can be a limiting factor. The guidance system must compensate to the extent possible in providing a fast, responsive system. The system should be able to recover from errors as quickly as possible and return to the desired path. In the case of homing on a target, this is crucial if the target can maneuver. Coupled with the need for a quick response is the simultaneous need for a stable response. The system must not respond so fast as to overshoot its proper heading excessively, causing possible growing oscillations. Consideration of stability places certain limitations on the guidance system and on the speed of response of the overall systems.

Another important feature of the system is its robustness. The guidance system design is based on a mathematical model of the vehicle, the autopilot, and the sensor. In fact, none of these behaves exactly like its model. The guidance system must provide good overall performance despite these modeling errors. Environmental effects (winds, ocean currents, and so forth) may also be different than what was expected. A robust guidance system provides good performance despite these uncertainties.

Reliability is also important. Efforts must be made to accommodate failures of various kinds. In many cases, backup components are provided for redundancy. This is frequently the case for the digital computer of the guidance system, especially for crewed space flight. *See* RELIABILITY, AVAILABILITY, AND MAINTAINABILITY.

Gerald Cook

Bibliography. R. H. Battin, *An Introduction to the Mathematics and Methods of Astrodynamics*, 1987, revised, 1999; M. Kayton, Navigation: Ships to space, *IEEE Trans. Aerosp. Electr.*, 24(5):474–519, 1988; C.-F. Lin, *Modern Navigation, Guidance, and Control*, 1991; G. W. Stimson, *Introduction to Airborne Radar*, 2d ed., 1998.

Guided missile

An uncrewed, controlled-flight vehicle that is guided to a target by guidance and control equipment. This equipment may be carried in the missile vehicle itself, or guidance may be directed from the launch site. The term is generally reserved for aerodynamic maneuverable missiles that may be guided to predetermined targets for military purposes.

Classification. Guided missiles are classified by launch or target mode such as air to air (AAM), air to surface (ASM), surface to air (SAM), surface to surface (SSM), and other possible modes such as subsurface launch to air, surface, or subsurface targets, and surface or air launch to subsurface targets. Missiles may be classified by range (short, medium, long) or by techniques related to tracking and guidance (radar, infrared heat seeker, optical or television, laser, radio, wire control command, fiber optics, inertial, acoustical). Some missiles make use of terrain following, which permits the missile to look at the terrain, compare it with a predetermined mapped route, and in effect fly a course as if by following a road map. Particularly for long-range flight, missiles may also make use of celestial navigation (star tracking) or may use electronic guidance assistance provided by satellites or midcourse surface stations. With optical missiles, a two-way data link may also be incorporated that provides the launcher with a continuous display from the missile video after launch and egress, and permits target updating or even reselection from a safer standoff distance (**Fig. 1**). *See* ELECTRONIC NAVIGATION SYSTEMS; GUIDANCE SYSTEMS; SATELLITE NAVIGATION SYSTEMS; STAR TRACKER.

Propulsion. Guided missiles are generally self-propelled, and may use rocket motors (liquid or solid), air-breathing turbojet engines, ramjets, or various types of combined-cycle engines such as air-augmented rockets, ducted rockets, turboramjets, or hybrid rockets (for example, solid fuel and liquid oxidizer). Future use may be made of nuclear rocket engines or various types of electrical rockets, such as charged particles or ion rockets. Advances

Fig. 1. GBU-15 with data link system, mounted on F-4 Phantom II fighter. (*Hughes Aircraft Co.*)

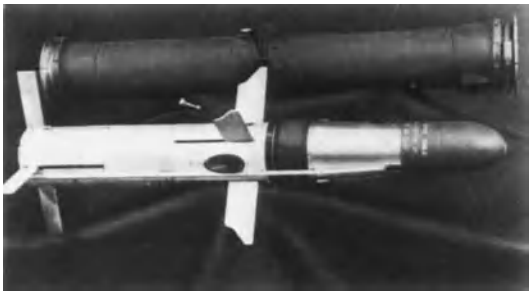

Fig. 2. TOW (tube-launched, optically tracked, wire-guided) antitank missile and launch tube. Wings and tails fold to permit tube mounting. (*Hughes Aircraft Co.*)

in the development of scramjet engines could provide the propulsion for a hypersonic cruise missile in the Mach number speed range from about 4 up to about 7. For some missions, particularly air-to-surface missions, unpowered, gliding, guided missiles may be used. *See* ION PROPULSION; ROCKET PROPULSION; TURBINE ENGINE SUBSYSTEMS; SCRAMJET.

Warhead. The kill mechanism for a missile consists of some form of explosive warhead and a system for detonation (fusing and arming). Warheads are typically either a high explosive or nuclear, and vary according to the type of fragment that is ejected and the spray pattern of the fragment. Warheads may be exploded upon contact with the target, by command from an external source, by a proximity fuse that senses the target, by preset timers, and so on.

Capabilities and constraints. Guided missiles are intended to be a type of munition extension with capabilities beyond that of commonly used, pure ballistic, free-flying projectiles. The capability extension may be in accuracy, warhead size, mobility, speed, range, altitude, endurance, or target selection. By the nature of the increased capability, missiles can be complicated and costly. Functions required of a missile system require the combined efforts of specialists in fields of aerodynamics, propulsion, electronics, materials, structures, warheads, sensors, fuses, reliability, safety, packaging, carriage, launching, and testing and evaluation. A severe constraint is imposed by the fact that a missile is a "throwaway" item, and its cost versus effectiveness must be carefully weighed. It is necessary that the required mission be achievable, but at an affordable cost. That is to say, a missile cost that exceeds the target cost would be undesirable; hence an effective missile with simplicity and low cost is a system to be sought.

Early development. During and shortly after World War I, the British tested radio-controlled airplanes and rockets, while the Americans tested an airplane controlled by on-board, preset pneumatic and electrical controls. The Germans became interested in missiles in the mid-1930s and undertook experiments with large guided rockets. This led to the initiation of guided missile warfare in 1943, with the use of radio-controlled, rocket-powered glide bombs against ships, and to the launching of several thousand V-1 and V-2 surface-to-surface missiles during 1944 and 1945. At the close of World War II, the Germans were developing a surface-to-air guided missile designed to carry out complex pursuit and evasion

Fig. 3. Launching of a field-mobile, short-range, all-weather surface-to-air Roland missile. The German–French Roland has been licensed for production in the United States for use as a U.S. Army air defense weapon. (*U.S. Army*)

Fig. 4. F-15 Eagle fighter aircraft equipped with a mixture of short-range, infrared Sidewinder and medium-range, radar Sparrow air-to-air missiles. (*McDonnell-Douglas*)

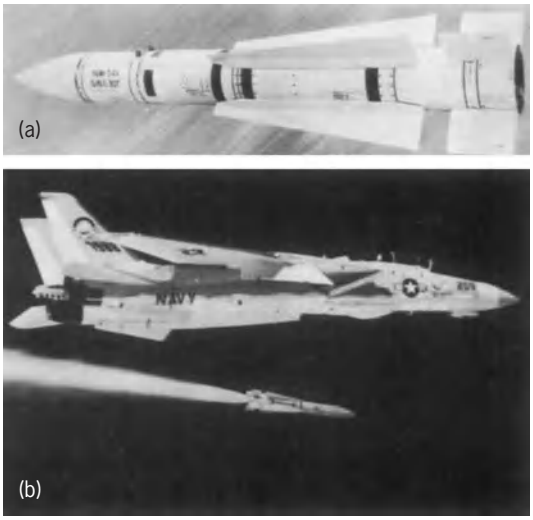

Fig. 5. Phoenix long-range air-to-air supersonic missile. (*a*) Side view (*Hughes Aircraft Co.*). (*b*) Launch from F-14 fighter aircraft (*U.S. Navy*).

maneuvers, and research based on this system continued in the United States.

Use. Since World War II, guided missiles have experienced only limited use in combat, primarily in Southeast Asia, in the Middle East wars, and in the Falklands; thus a measure of combat effectiveness is difficult to assess. However, missiles have changed the nature of warfare in many ways. Ground troops have an increased firepower through the use of human-portable antiair and antitank missiles (**Fig. 2**). Ground troop protection is increased through the use of mobile air defense missiles (**Fig. 3**). Ground warfare has also been changed through the use of helicopter-launched missiles for both ground and air target attack.

Navy vessels have an increased firepower through the use of antishipping-type missiles, as well as increased protection with air defense missiles. Some unique missile deployment systems have appeared in the Soviet navy. The Kirov cruiser, for example, carries a variety of missiles suggestive of both defensive and offensive capability, including antisubmarine and two antiair types, as well as long-range cruise surface-to-surface antishipping or land-attack missiles. The Slava cruiser, as another example, is equipped with long-range surface-to-surface cruise missiles and two types of surface-to-air missiles. Newer ships, such as these, make use of vertical launch systems which provide certain advantages when launching from a moving platform in that launcher elevation and azimuth settings are not critical to the missile trajectory, and target acquisition is performed by the missile itself after launch.

Fig. 6. Falcon air-to-air missile family, including missiles with infrared and radar guidance, and high-explosive and nuclear warheads. (*Hughes Aircraft Co.*)

Fig. 7. Tomahawk long-range, subsonic cruise missile. (*a*) Side view. (*b*) Submerged launch. Missile may also be launched from surface or air. (*General Dynamics*)

Aircraft have an increased effectiveness through the use of a variety of air-to-air combat missiles (**Figs. 4, 5,** and **6**) and air-to-surface missiles ranging from battlefield ground support to long-range cruise stand-off missiles.

The cruise missile (**Figs. 7** and **8**), either subsurface-, surface-, or air-launched, provides increased flexibility for tactical and strategic missions from stand-off distances that provide improved safety for the launch platform. Various degrees of stand-off distances are achieved with ballistic missiles by using a variety of types of midcourse and terminal guidance, types of warheads, propulsion schemes, and launch techniques. *See* AIR ARMAMENT; ANTISUBMARINE WARFARE; ARMY ARMAMENT; MILITARY AIRCRAFT; NAVAL ARMAMENT.

Guided missile systems are essential elements of the arsenals of many nations. Some of the world missile systems are listed in the **table**.

Some omissions and uncertainties are probably due to incomplete sources of information. The listing

World missile systems

United States

AAM

Falcon
Sparrow
Oriole
Meteor
Ding-Dong
Sidewinder
Eagle
Genie
Diamondback
Phoenix
Agile
Brazo
CLAW
SHAG
Seek Bat
AMRAAM
ASAT (ALMV)

SAM

Lark
Terrior
Nike Ajax
Bomarc
Loki
Hawk
Talos
Tartar
Nike Hercules
Mauler
Nike Zeus
Redeye
Davy Crockett
Bull Goose
Typhoon
Nike X/Sprint
Chaparral
SAM-D
Standard
Spartan
Stinger
Patriot
RAM
ADATS

ASM

Rascal
Bullpup
Bulldog
Dove
Gorgon
Bold Orion
Corvus
Hound Dog
Skybolt
Wagtail
Quail
Crossbow
Green Quail
Shrike
Condor
SRAM
WASP
Standard ARM
Teton
Cobra
LASRM
Maverick
Hornet
Walleye
Viper
ZAP
SCAD
Harpoon
ALCM
HARM
Hellfire
HOBOS
Paveway
ASALM
SLAM
GBU-15

SSM

Matador
Corporal
Regulus I
Regulus II
Honest John

SSM (*cont.*)

Snark
Navaho
Redstone
Lacross
Hermas
Rigel
Atlas
Dart
Little John
Sergeant
Jupiter
Polaris
Thor
Titan
Triton
Clam
Slam
Cobra
Lobber
Mace
MAW
Minuteman
Pershing
Shillelagh
Law
Taurus
TOW
Lance
Poseidon
Dragon
Trident
Tomahawk
Viper
Petrel
Able
Asroc
Subroc
Astor
Rat
Able-Alfa
MX
Copperhead

Russia

AAM

AA-1
AA-2
AA-3
AA-4
AA-5
AA-6
AA-7
AA-8
AA-9
AA-XP-1
AA-XP-2

SAM

SA-1
SA-2
SA-2/3
SA-2/4
SA-3
SA-4
SA-5
SA-6
SA-7
SA-8
SA-9
SA-10
SA-11
SA-12
SA-13
SA-N-3
SA-N-4
SA-N-5
SA-N-6
SA-N-7
SA-N-8
Galosh
ABM-X-3
SH-4
SH-8

ASM

AS-1
AS-2
AS-3
AS-4
AS-5
AS-6
AS-7
AS-8
AS-9
AS-X-10
AS-11
AS-X-15
BL-10

SSM

AT-1
AT-2
AT-3
AT-4
AT-5
AT-6
Salish
Samlet
Scrubber
Shaddock
Styx
Sark
Serb
Scud
Shyster
Sandal
Skean
Sapwood
Saddler
Sa-sin
Scarp
Scrag
Sego
Scaleboard
Savage
Scapegoat
Scrooge

SS-16
SS-17
SS-18
SS-19
SS-20
SS-21
SS-22
SS-23
SS-X-24
SS-X-25
SS-N-6
SS-N-7
SS-N-8
SS-N-9
SS-N-10
SS-N-11
SS-N-12
SS-NX-13
SS-N-14
SS-N-15
SS-N-16
SS-N-17
SS-N-18
SS-NX-19
SS-N-20
SS-NX-21
SS-NX-22
SS-NX-23
SSC-X-4

Argentina

Martin Pescador
Mathogo

Australia

Ikara

Brazil

Carcara
MSS-1

China

CSS-2
CSS-3
CSS-X-4

France

ACRA
Entac
SS-11 (Harpoon)
SS-12
Pluton
Magic R550
R-511
R-530
AS-12
AS-20
AS-30
Martel
Crotale R-440
Super 530
Masurca
Exocet
Malafon
HOT
Milan
Otomat
Mercure
Hirondelle
MICA
Fakir-H
ANS
AS-15
ASMP
ARMAT
SATCP
AS-11
SSBS S-3
MSBS M-20
MSBS M-4
Sica/Shahine

Germany

Cobra
Roland
Viper
Kormoran
Armburst
Mamba
Jumbo
Tiralleur

Israel

Shafrir
Gabriel
Picket
Python
PDM

Italy

Airtos
Aspide
Indigo
Sea Killer
Marte
Mosquito
Sparviero

Japan

KAM-3D
KAM-9
ASM-1
AAM-1
AAM-2
Tan-SAM

Norway

Penguin
Teme

Republic of South Africa

Kukri

Sweden

Bantam RB53
RB04
RB05
RB08
RB70
RB56
RB15

Switzerland

Micon
Mosquito

Taiwan

Kun Wu

United Kingdom

Blue Steel
Swingfire
Vigilant
Blowpipe
SLAM
Rapier
Firestreak
Red Top
SRAAM
Bloodhound
Thunderbird
Tiger Cat
Hellcat
Seacat
Sea Dart
Seaslug
Seawolf
Sea Skua
SAMM10
Hawkswing
Beeswing
Skyflash
ASRAAM
ALARM
Javelin
Sea Eagle

Fig. 8. Harpoon antiship cruise missile. (*a*) Ship-launched. (*b*) Air launch version on A-6 intruder attack aircraft. (*McDonnell-Douglas*)

also includes many experimental and defunct systems. M. Leroy Spearman

Bibliography. Air Force Magazine, *Air Force Almanac*, annually; Aviation Week and Space Technology, *Aerospace Forecast and Inventory*, annually; J. R. Chambers, *Partners in Freedom: Contributions of the Langley Research Center to U. S. Military Aircraft of the 1990s*, National Aeronautics and Space Administration, Washington, DC, 2000; E. Fleeman, *Tactical Missile Design*, 2d ed., American Institute of Aeronautics & Astronautics, Reston, VA, 2006; Flight International, *World Missile Directory*, annually; General Dynamics—Pomona Division, *The World's Missile Systems*, revised periodically; B. Gunston, *The Illustrated Encyclopedia of the World's Rockets and Missiles*, Smithmark Pub, 1987; M. J. Hemsch (ed.), *Tactical Missile Aerodynamics: General Topics*, American Institute of Aeronautics & Astronautics, Reston, VA, 1992; M. L. Spearman, *Historical Development of World Wide Guided Missiles*, NASA TM 85658, 1983; M. L. Spearman, *A Historical Review of Missile Airframe Designs*, AIAA Pap. 97–2277, 1997.

Guild

A group of species that utilize the same kinds of resources, such as food, nesting sites, or places to live, in a similar manner. Emphasis is on ecologically associated groups that are most likely to compete because of similarity in ecological niches, even though species can be taxonomically unrelated. The term was derived from the guild in human society composed of people engaged in an activity or trade held in common.

Richard B. Root first used the term guild in 1967, specifying the "foliage-gleaning guild" as a group of birds that habitually harvest arthropods (mainly insects and spiders) from tree leaves. In 1973 he distinguished three guilds based on feeding methods of herbivorous invertebrate animals in a community utilizing one plant species: the sap-feeding guild, the strip-feeding guild, and the pit-feeding guild. Members of the sap-feeding guild sucked juices from the plant with tubular mouthparts and included plant lice (Aphidae), frog hoppers (Cercopidae), leaf hoppers (Cicadellidae), and plant bugs (Miridae). Members of the strip-feeding guild chewed off strips of leaves with cutting mandibles or similar mechanical shears, and included caterpillars (Lepidoptera), crickets and grasshoppers (Orthoptera), and snails and slugs (Mollusca). Members of the pit-feeding guild used chewing or sucking mouthparts but concentrated feeding in small areas, resulting in many pits of consumed or damaged tissue; they included leaf beetles (Chrysomelidae), weevils (Curculionidae), and thrips (Thripidae). This was the first time that unrelated organisms were grouped into ecologically meaningful units smaller than the whole community; these groups represent the major functional roles in the community. *See* ECOLOGICAL COMMUNITIES.

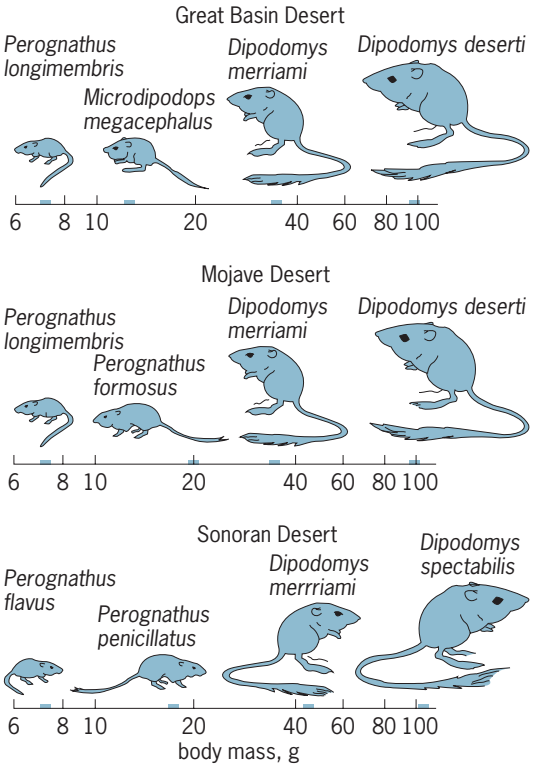

Guilds of granivorous rodents in three deserts in the southwest United States. Mean weight (in grams) of each species is plotted on the body mass axis. The body mass scale is logarithmic, so that equal spacing shows equal ratios. (*After M. A. Bowers and J. H. Brown, Body size and coexistence in desert rodents: Chance or community structure?, Ecology, 63:391–400, 1982*)

The guild concept focuses attention on the ways in which ecologically related species differ enough to permit coexistence, or avoid competitive displacement. For example, new places to live for some plants are provided by badger mounds in dense tall-grass prairie vegetation. Many species were associated with these disturbances, and since mounds were relatively small and rare, the potential for competition in this guild was high. One important difference between guild members was dispersal of seeds, so that some colonized new mounds quickly and others slowly, the result being a temporal divergence in utilization. The other major difference to which plants responded was the soil moisture in mounds. The combination of these characteristics permitted many species of plants to coexist on a population of badger mounds.

Seeds provide the major source of food for many desert rodents. Studies on such granivorous rodent guilds have shown that they are highly structured in relation to body sizes of member species. In one study it was found that although species assemblages differ between the Great Basin, Mojave, and Sonoran deserts, four species per guild coexist, with distinct body sizes at about 100, 40, 20, and 7 g (3.5, 1.4, 0.7, and 0.2 oz, respectively; see **illus.**). This uniformity in guild composition suggests the strong influence of competition organizing the guild. *See* POPULATION ECOLOGY.

Not all guilds are as clearly structured as the granivorous rodent guild. Often the hypothesis of random assemblage of species cannot be refuted in many cases based on size differences. Much research interest is focused by ecologists on the role of competition in structuring guilds and communities.

The guild is also commonly used as the smallest unit in an ecosystem in studies relating to environmental impact, wildlife management, and habitat classification. A representative species of a guild may be selected for study involving the uncertain assumption that environmental impact will influence this species in the same way as other guild members. *See* ECOSYSTEM. Peter W. Price

Bibliography. M. A. Bowers and J. H. Brown, Body size and coexistence in desert rodents: Chance or community structure?, *Ecology*, 63:391–400, 1982; P. W. Price, *Insect Ecology*, 3d ed., 1997; R. B. Root, Organization of a plant-arthropod association in simple and diverse habitats: The fauna of collards (*Brassica oleracea*), *Ecol. Monogr.*, 43:95–124, 1973; D. Simberloff and W. Boecklen, Santa Rosalia reconsidered: Size ratios and competition, *Evolution*, 35:1206–1228, 1981.

Gulf of California

A young, elongate ocean basin on the west coast of Mexico. It is flanked on the west by the narrow mountainous peninsula and continental shelf of Baja California, while the eastern margin has a wide continental shelf and coastal plain. The floor of the gulf consists of a series of basins 3300–12,000 ft (1000–3600 m) deep, whereas the northern gulf is dominated by a broad shelf which is the result of deltaic deposition from the Colorado River. The structural depression of the gulf continues northward into the Imperial Valley of California, which is cut off from the ocean by the delta of the Colorado River. *See* CONTINENTAL MARGIN.

The gulf formed along the plate edge between the Pacific and the North American lithospheric plates (see **illus.**). The plate edge enters the gulf at the north as the San Andreas Fault System and passes through a series of short segments of transform faults and spreading axes to where it becomes the East Pacific Rise entering the mouth of the gulf at the south. Opening of the gulf is believed to have been initiated between 5 and 6 million years before present, contemporaneous with an eastward jump of the trend of the San Andreas system in southern California, and the gulf was created by approximately 180 mi (300 km) of northwestward displacement of the peninsula of Baja California as a part of the Pacific plate. Thus, the floor of the gulf consists of young oceanic crust beneath the basins, adjacent to thin subsided continental crust along the continental margins. *See* FAULT AND FAULT STRUCTURES; PLATE TECTONICS.

Most of the gulf lies within an arid climate, with 4–6 in. (10–15 cm) of annual rainfall over Baja California and ranging on the eastern side from 4 in. (10 cm) in the north to about 34 in. (85 cm) in the southeast. No year-round streams enter the gulf on the west; a series of intermediate-size rivers flow in on the east side; and the major source of freshwater sediment came from the Colorado River at the

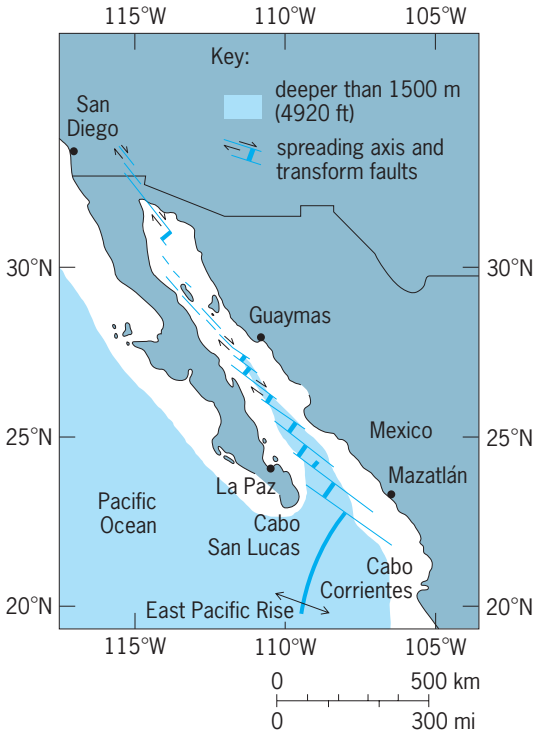

Simplified tectonics of the Gulf of California. (*After J. R. Curray et al., Tectonics and geological history of the passive continental margin at the tip of Baja California, in Initial Reports of the Deep Sea Drilling Project, vol. 64, pt. 2, U.S. Government Printing Office, 1982*)

north prior to damming it upstream in the United States.

Water circulation is driven by seasonal wind patterns. Surface water is blown into the gulf in the summer by the southwesterly wind regime. In the winter, surface water is driven out of the gulf by the northwesterly wind regime, and upwelling occurs along the eastern margin, resulting in high organic productivity. Bottom sediments of the gulf range from deltaic sediments of the Colorado River at the north and coalesced deltas of the intermediate-size rivers on the east. A strong oxygen minimum occurs between 990 and 3000 ft (300 and 900 m) water depth, where seasonal influx of terrigenous sediments and blooms of diatoms due to upwelling produce varved sediments consisting of alternating diatom-rich and clay-rich layers. Rates of sediment accumulation are high, and total sediment fill beneath the Colorado River delta at the north may attain thicknesses of greater than 6 mi (10 km), even though the structural depression and the underlying crust are geologically young. *See* BACILLARIOPHYCEAE; DELTA; MARINE SEDIMENTS; OCEAN CIRCULATION; UPWELLING; VARVE. Joseph R. Curray

Bibliography. J. K. Crouch and S. B. Bachman (eds.), *Tectonics and Sedimentation Along the California Margin*, Pacific Section, Society of Economic Paleontologists and Mineralogists, 1984; J. R. Curray et al., *Initial Reports of the Deep Sea Drilling Project*, vol. 64, pts. 1 and 2, 1982; R. A. Schwartzlose and J. R. Hendrickson, *Bibliography of the Gulf of California: Marine Sciences (Through 1981)*, Inst. Cienc. Mar Limnol., Univ. Nal. Auton. Mex., vol. 7, 1983; Tj. H. van Andel and G. G. Shor, Jr. (eds.), *Marine Geology of the Gulf of California: A Symposium*, Amer. Ass. Petrol. Geol. Mem. 3, 1964.

Gulf of Mexico

A subtropical semienclosed sea bordering the western North Atlantic Ocean. It connects to the Caribbean Sea on the south through the Yucatán Channel and with the Atlantic on the east through the Straits of Florida. To the north, it is bounded by North America, to the west and south by Mexico and Central America, and on the east and southeast by Florida and Cuba respectively. The Gulf of Mexico Basin lies wholly within the boundaries of the North American tectonic plate. *See* BASIN; INTRA-AMERICAS SEA; PLATE TECTONICS.

Marine geology. The continental shelves surrounding the gulf are very broad along the eastern (Florida), northern (Texas, Louisiana, Mississippi, Alabama), and southern (Campeche) area, averaging 125–186 mi (200–300 km) wide (see **illus.**). The continental shelves along the western and southwestern (Mexico) and southeastern (Cuba) boundaries of the gulf are narrow, often being less than 12 mi (20 km) wide. Between the continental shelves and the Sigsbee Abyssal Plain are three steep continental slopes: the Florida Escarpment off west Florida, the Campeche Escarpment off Yucatán, and the Sigsbee Escarpment south of Texas and Louisiana. Two major submarine canyons crease the gulf's shelf areas: the De Soto Canyon near the Florida-Alabama border, and the Campeche Canyon west of the Yucatán Peninsula. *See* CONTINENTAL MARGIN; ESCARPMENT; HOLOCENE; MARINE GEOLOGY.

Meteorology. Compared with the North American rivers, the Mexican rivers are short, but they still provide approximately 20% of the fresh-water input to the gulf because of extensive orographic rainfall from the trade winds that dominate the southern flank of the basin. Meteorologically, the Gulf of Mexico is a transition zone between the tropical wind system (easterlies) and the westerly frontal-passage-dominated weather (in winter particularly) to the north, punctuated with intense tropical storms in summer/autumn called the West Indian Hurricane. Much of the atmospheric moisture supplied to the North American heartland during spring and summer has its origin over the gulf, and thus it is a vital element in the so-called North American Monsoon. *See* HURRICANE; MONSOON METEOROLOGY; STORM SURGE; TROPICAL METEOROLOGY.

Circulation. The Gulf Stream System dominates the oceanic circulation in the Gulf of Mexico. The Yucatán Current, flowing northward into the eastern Gulf of Mexico, is the first recognizable western boundary current in the Gulf Stream System. North of the Yucatán Peninsula, the flow penetrates into the eastern gulf (where it is called the Gulf Loop Current) at varying distances with a distinctive chronology, loops around clockwise, and finally exits through the Straits of Florida, where it is called the Florida Current. This intense current reaches to more than 3300 ft (1000 m) depth, and transports 1.1×10^9 ft^3/s (3×10^7 m^3/s) of water, an amount 1800 times that of the Mississippi River. *See* GULF STREAM; MEDITERRANEAN SEA; OCEAN CIRCULATION.

Over a nearly annual cycle, the Gulf Loop Current flows between Mexico and Cuba, moving northward to the Mississippi Delta, then southward along the Florida Escarpment, and finally eastward into the Straits of Florida between Florida and Cuba. After reaching its maximum northerly position, an anticyclonic eddy, approximately 186 mi (300 km) in diameter, separates. The main flow then reforms between the Yucatan Channel and the Florida Keys, and the process starts again. *See* OCEAN CIRCULATION; SEAWATER.

Biogeography. Surrounding the Gulf of Mexico are many population centers that exploit the numerous estuaries, lagoons, and oil and gas fields. Coral reefs off Yucatán, Cuba, and Florida provide important fishery and recreational activities. There are extensive wetlands along most coastal boundaries with ecological connections to many seagrass beds nearshore and coastal mangrove forests of Mexico, Cuba, and Florida. This biogeographic confluence creates one of the most productive marine areas on Earth, providing the food web for commercially important species such as lobster, demersal (bottom-dwelling) fish, and shrimp; this same ecology supports large populations of sea turtles and marine mammals. The coastal and nearshore waters also support large phytoplankton populations.

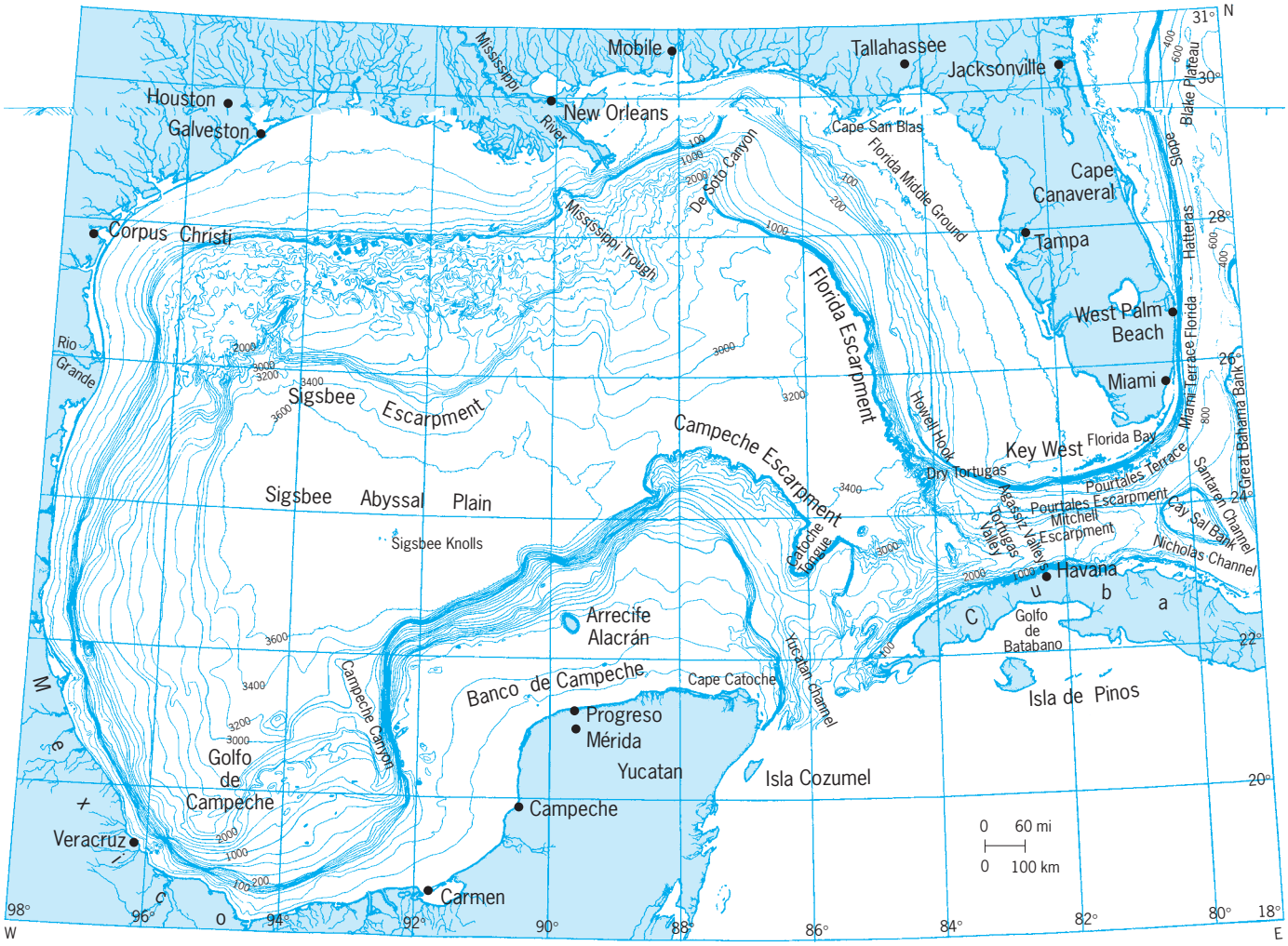

Bottom topography and physiographic provinces in the Gulf of Mexico. Contours in meters.

The juxtaposition of these enormous marine resources and human activities has led to a distinctive anthropogenic impact on the health of the marine ecosystem. *See* BIOGEOGRAPHY; ESTUARINE OCEANOGRAPHY; FOOD WEB; MANGROVE; MARINE ECOLOGY; REEF; WETLANDS. George A. Maul

Bibliography. L. R. A. Capurro and J. L. Reid (eds.), *Contributions on the Physical Oceanography of the Gulf of Mexico*, 1972; G. A. Maul and F. M. Vukovich, The relationship between variations in the Gulf of Mexico Loop Current and Straits of Florida volume transport, *J. Phys. Oceanog.*, 23(5):785–796, 1993; H. Stommel, *The Gulf Stream*, 2d ed., 1965.

Gulf Stream

A great ocean current transporting about 7×10^7 tons (6.3×10^7 metric tons) of water per second (1000 times the discharge of the Mississippi River) northward from the latitude of Florida to the Grand Banks off Newfoundland. Before the days of scientific oceanography, it was supposed that the origin of the water in the Gulf Stream was the Gulf of Mexico. The origin has since been traced farther upstream, through the Gulf of Mexico and Caribbean, to the Great North and South Equatorial currents of the Atlantic. The Gulf Stream is thought of now as a portion of a great horizontal circulation in the ocean, where particles of water execute closed circuits, sometimes moving slowly in mid-ocean regions and other times rapidly in strong currents like the Gulf Stream. Thus the beginning and end of the Stream have arbitrary geographical limits. *See* ATLANTIC OCEAN; GULF OF MEXICO.

The Gulf Stream is a narrow (60 mi or 100 km) and swift (up to 5 knots or 250 cm/s) eastward-flowing current jet which is surrounded by intense eddies. As it leaves the coast at Cape Hatteras, the Stream becomes unstable and meanders from side to side like a river. The meanders usually move downstream with a speed of 0.2 knot (10 cm/s), becoming progressively larger; they have a particularly large amplitude near the New England seamounts near 60°W. South of New England, meanders are 60–120 mi (100–200 km) in length and width, and change position with periods of several days to several weeks. Meanders pinch off from the Stream 10–15 times per year, forming energetic eddies, typically 120 mi (200 km) in diameter (**Fig. 1**). These eddies are called rings because they consist of a ring of Gulf Stream water encircling water in the center of the eddy. Rings frequently drift westward and eventually coalesce with the Stream after a lifetime of months to years.

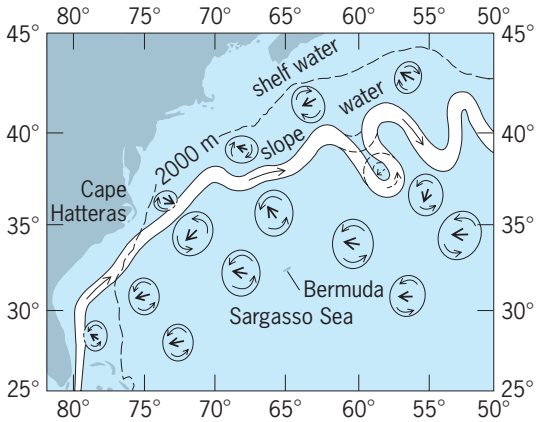

Fig. 1. The path of the Gulf Stream and the distribution and number of rings. Warm rings are formed to its north and cold rings to its south. 2000 m = 6600 ft. (*After P. L. Richardson, Gulf Stream rings, Oceanus, 19(3):65–68, 1976*)

The meanders and rings are thought to force counterrotating gyres that recirculate water from the Gulf Stream on both sides of it (**Fig. 2**).

The near-surface Gulf Stream transports warm water from southern latitudes eastward to the Grand Banks, where the flow becomes broader and weaker, separating into several branches and eddies. About half the near-surface flow continues eastward across the Mid-Atlantic Ridge, and half recirculates southwestward.

The subsurface Gulf Stream coincides with a strong horizontal temperature gradient marking the juncture of cold slope water on the north and warm Sargasso water on the south. Although the thermal gradient extends deeper than 6600 ft (2000 m) under the Stream, the deep current structure is complex and is not a simple downward extension of the near-surface flow. The mean velocity of the Gulf Stream decreases with depth and extends to the sea floor even in depths of 4000 m. The deep Gulf Stream gyre including most of the recirculation is located west of the Mid-Atlantic Ridge and north of Bermuda. *See* MID-OCEANIC RIDGE.

The volume transport of the Gulf Stream increases from 30 sverdrups (1 Sv = 10^6 m^3/s or 3.5×10^7 ft^3/s) off Florida to 150 Sv north of Bermuda, and then decreases eastward (Fig. 3). A large part of the increase in transport is caused by eastward components of the recirculating gyre which extend deeply to the sea floor.

The Gulf Stream is driven predominantly by the large-scale wind pattern, the westerlies in the north and the trades in the south. The winds exert a torque on the ocean that, because of the shape and rotation of the Earth, causes a large western-intensified gyre that transports 30 Sv. Part of the Atlantic circulation is driven by the large-scale density differences caused mainly by heating in low latitudes and cooling in high latitudes. Cold, deep water is formed in northern seas and flows southward as a deep western boundary current transporting roughly 15 Sv; warm upper layer water flows northward and replaces it. This flow is called the meridional overturning circulation. The deep western boundary current which augments the westward-flowing portion of the northern recirculating gyre encounters the Gulf Stream off Cape Hatteras, North Carolina. Some of the deeper water in the deep boundary current continues southward along the boundary passing under the mean axis of the Stream. The rest of the deeper water is entrained into the Gulf Stream, some to eventually cross the mean axis of the Stream, recirculate westward toward the western boundary, and continue southward. Energetic fluctuation of the deep currents and the Gulf Stream make it much more complex than shown in schematics which attempt to portray the mean flow. *See* OCEAN CIRCULATION; WIND STRESS.

Almost half of the Gulf Stream transport in the Straits of Florida originates in the South Atlantic. The rest comes from the large-scale recirculation or flow in the subtropical gyre of the North Atlantic. Near the Grand Banks of Newfoundland, the Gulf Stream divides into several branches. The first branch flows northward as a western boundary current east of Newfoundland, turns eastward near 50°N, and flows across the Mid-Atlantic Ridge as the North Atlantic Current. The second branch flows eastward across the Mid-Atlantic Ridge south of the Azores, and is concentrated into the relatively swift eastward-flowing Azores Current centered near 34°N. The third branch recirculates water south of the Gulf Stream.

Philip Richardson

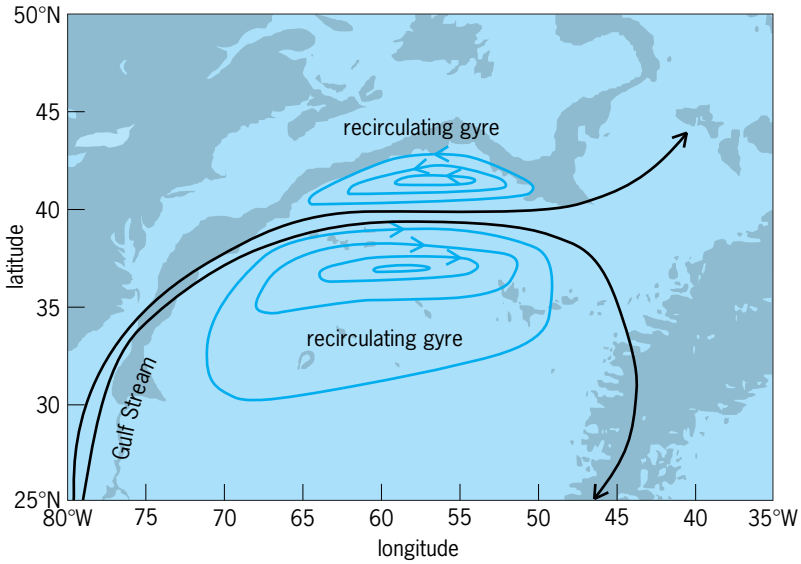

Fig. 2. Average transport of the Gulf Stream and recirculating gyres. Each line represents approximately 15 Sv of transport. (*After N. G. Hogg, On the transport of the Gulf Stream between Cape Hatteras and the Grand Banks, Deep-Sea Res., 39:1231–1246, 1992*)

Bibliography. W. E. Johns et al., Gulf Stream structure, transport, and recirculation near 68°W, *J. Geophys. Res.*, 1995; A. R. Robinson (ed.), *Eddies in Marine Science*, 1983; W. J. Schmitz, Jr., and M. S. McCartney, On the North Atlantic circulation, *Rev. Geophys.*, 1993; H. Stommel, *The Gulf Stream: A Physical and Dynamical Description*, 2d ed., 1965; B. A. Warren and C. Wunsch (eds.), *Evolution of Physical Oceanography*, 1981; L. V. Worthington, *On the North Atlantic Circulation*, Johns Hopkins Oceanographic Series, vol. 6, 1976.

Gum

A class of high-molecular-weight molecules, usually with colloidal properties, which in an appropriate solvent or swelling agent are able to produce gels (highly viscous suspensions or solutions) at low dry-substance content. The molecules are either hydrophilic or hydrophobic; that is, they do or do not have an affinity for water. The term gum is applied to a wide variety of substances of gummy characteristics, and therefore cannot be precisely defined.

Various rubbers are considered to be gums, as are many synthetic polymers, high-molecular-weight hydrocarbons, or other petroleum products. Chicle for chewing gum is an example of a hydrophobic polymer which is termed a gum but is not frequently classified among the gums. Quite often listed among the gums are the hydrophobic resinous saps that often exude from plants and are commercially tapped in balsam (gum balsam) and other evergreen trees (gum resin). Incense gums such as myrrh and frankincense are likewise fragrant plant exudates.

Usually, however, the term gum, as technically employed in industry, refers to plant polysaccharides or their derivatives. These are dispersible in either cold or hot water to produce viscous mixtures. Modern usage of the term includes water-soluble derivatives of cellulose and derivatives and modifications of other polysaccharides which in the natural form are insoluble. Usage, therefore, also includes with gums the ill-defined group of plant slimes called mucilages. *See* CELLULOSE; COLLOID; POLYSACCHARIDE.

Viscosity. Gums are important in that they impart viscosity to aqueous solutions. Their physical properties are manifestations of their chemical structure, the kind and amount of solvent, and the kind and concentration of ions and other substances dissolved in the solvent. Because gums are commonly composed of several different kinds of monomer units with many possible variations in regard to degree of branching, length of branches, and types of linkages, an almost infinite number of structures is possible. Forces act between molecules, between different parts of the same molecule, and between polymer and solvent. These forces include hydrogen bonding, ionic charges, dipole and induced dipole interactions, and van der Waals forces.

All of these forces affect such properties as gel-forming tendency, viscosity, and adhesiveness. The types of linkage due to their effects on chain flexibility are important also in determining physical properties. For example, it is known that linear molecules make more viscous solutions than do long-branch molecules of similar molecular weights, but they have a tendency to precipitate because of association of the chains. If this association is prevented, stability can be achieved without much sacrifice of viscosity. This can be done by introducing groups with ionic charges that repel one another, or by attaching many very short side branches to prevent

Sources and sugars of important gums

Gum	Source	Sugars present and linkages
Seaweeds		
Agar	Red algae (*Gelidium* sp.)	D-Galactose β-(1 $\rightarrow$ 4), 3,6-anhydro-L-galactose α-(1 $\rightarrow$ 3) + sulfate acid ester groups
Algin	Brown algae (*Macrocystis pyrifera*)	D-Mannur onic acid β-(1 $\rightarrow$ 4), L-gulur onic acid β-(1 $\rightarrow$ 4), Na salt
Carrageenan	Red algae (*Chondrus crispus, Gigartina stellata*)	D-Galactose, 3,6-anhydro-D-galactose + sulfate acid ester groups
Fucoidan	Brown algae (*Fucus* sp., *Laminaria* sp.)	L-Fucose + sulfate acid ester groups
Laminaran	Brown algae (*Laminaria* sp.)	D-Glucose, D-mannitol, β-(1 $\rightarrow$ 3) chain and β-(1 $\rightarrow$ 6) branches
Plant exudates		
Gum arabic	*Acacia* sp.	L-Arabinose, D-galactose, L-rhamnose, D-glucur onic acid
Ghatti	*Anogeissus latifolia*	L-Arabinose, D-xylose, D-galactose, D-mannose, D-glucur onic acid
Karaya	*Sterculia urens*	D-Galactose, D-rhamnose, D-galactur onic acid
Tragacanth	*Astragalus* sp.	D-Galactose, D-xylose, D-glucur onic acid
Plant extracts		
Pectin	Cell walls and intracellular spaces of all plants (commercial source, citrus waste)	D-Galactur onic acid α-(1 $\rightarrow$ 4) partially esterified, L-arabinose α-(1 $\rightarrow$ 5) and α-(1 $\rightarrow$ 3) branches, D-galactose β-(1 $\rightarrow$ 4)
Larch arabino-galactan	Western larch	D-Galactose, L-arabinose
Ti	Tubers of *Cordyline terminalis*	D-Fructose, D-glucose
Plant seeds		
Corn-hull gum	Corn seed coat	D-Xylose, L-arabinose, D-galactose, L-galactose, D-glucur onic acid
Guar	*Camposia teragonolobus* endosperm	D-Mannose β-(1 $\rightarrow$ 4), D-galactose α-(1 $\rightarrow$ 6) branches
Locust bean	Carob tree (*Ceratonia siliqua*) endosperm	D-Mannose β-(1 $\rightarrow$ 4), D-galactose α-(1 $\rightarrow$ 6) branches
Quince seed	*Cydonia vulgaris*	L-Arabinose, D-xylose, hexuronic acid, monomethyl hexuronic acid
Psyllium seed	*Plantago* sp.	D-Xylose, L-arabinose, D-galactur onic acid, L-rhamnose, D-galactose
Flax seed	*Linum usitatissimum*	D-Galactur onic acid, D-xylose, L-rhamnose, L-arabinose, L-galactose, D-glucose
Tamarind	*Tamarindus indica*	D-Glucose, D-galactose, D-xylose
Wheat gum	Wheat	D-Xylose β-(1 $\rightarrow$ 4), L-arabinose branches
Miscellaneous		
Cellulose derivatives	Plant cell walls, wood pulp, and cotton	D-Glucose β-(1 $\rightarrow$ 4)
Starch	Cereal grains and tubers	D-Glucose α-(1 $\rightarrow$ 4), D-glucose α-(1 $\rightarrow$ 6) at branch points
Amylose		
Amylopectin		
Dextran	Bacterial action on sucrose	D-Glucose α-(1 $\rightarrow$ 6) and α-(1 $\rightarrow$ 3)
Chitin	Exoskeleton of animals of the phylum Arthropoda	*N*-Acetyl-D-glucosamine β-(1 $\rightarrow$ 4)

close approach of the chains. Much work remains to be done in this area. While it is known that solutions of some gums are slimy or mucilaginous, whereas those of other gums are tacky, the reasons for these differences are unknown. Rheological properties of the different gum solutions also differ. *See* HYDROGEN BOND; POLARIZATION OF DIELECTRICS; RHEOLOGY.

Source and structure. Many plant gums originally used by humans still are important items of commerce. Most of these are dried exudates from trees and shrubs, produced with or without artificial stimulation by injury. These gums are collected by hand, usually in the hot, semiarid regions of the world, often from wild plants, but sometimes from cultivated plants. For the most part, these gums have highly branched structures containing two to five different sugar units. They frequently are acidic in that they contain carboxyl groups which are portions of glycuronic acids. Commercial gums often are mixtures of two or more different polysaccharides.

Seed gums, which also have been used for many centuries, likewise are important items of commerce today. The more ancient gums were extracted from quince, psyllium, or flax seeds. The ground endosperm of locust trees grown in the Mediterranean area also has been an item of commerce for many years. It is still important today, along with a similar gum obtained from the endosperm of the guar plant which is native to India and Pakistan but is grown to a small extent in the southwestern United States.

Seaweed gums produced by hot water or alkaline extraction of seaweed are other important industrial raw materials.

Industrially important or potentially useful gums are listed in the **table**, together with the plant source, the sugar units, and the most prevalent glycosidic linkage when known.

Use. Gums are used in foods as stabilizers and thickeners. They form viscous solutions which prevent aggregation of the small particles of the dispersed phase. In this way they aid in keeping solids dispersed in chocolate milk, air in whipping cream, and fats in salad dressings. Gum solutions also retard crystal growth in ice cream (ice crystals) and in confections (sugar crystals). Their thickening and stabilizing properties make them useful in water-base paints, printing inks, and drilling muds. Because of these properties they also are used in cosmetics and pharmaceuticals as emulsifiers or bases for ointments, greaseless creams, toothpastes, lotions, demulcents, and emollients. Gums are used to modify texture and increase moisture-holding capacity, for example, as gelling agents in canned meats or fish, marshmallows, jellied candies, and fruit jellies. Their adhesive properties make them useful in the production of cardboard, postage stamps, gummed envelopes, and as pill binders. Other applications include the production of dental impression molds, fibers (alginate rayon), soluble surgery films and gauze, blood anticoagulants, plasma extenders, beverage-clarifying agents, bacteriological culture media, half-cell bridges, and tungsten-wire-drawing lubricants. *See* FOOD MANUFACTURING.

Roy L. Whistler

Bibliography. A. A. Lawrence, *Natural Gums for Edible Purposes*, 1977; C. L. Mantell, *The Water Soluble Gums*, 1947, reprint 1965; W. Pigman and M. L. Wolfrom (eds.), *Advances in Carbohydrate Chemistry*, vol. 13, 1958; K. R. Stauffer, *Handbook of Edible Gums,* 1988; R. L. Whistler and J. N. BeMiller (eds.), *Industrial Gums*: *Polysaccharides and Their Derivatives*, 3d ed., 1993; R. L. Whistler and M. L. Wolfrom (eds.), *Methods in Carbohydrate Chemistry*, vol. 5: *General Polysaccharides*, 1965.

Gunsights

Optical instruments which establish an optical line or axis for the purpose of aiming a weapon. The axis includes the observer's eye, a suitable mark in the instrument, and the target. Most gunsights employ as their basis a telescope or a partially reflective mirror.

Rifle sights. A typical rifle sight (**Fig. 1**) consists of a terrestrial telescopic system having an objective, an eyepiece, an erector lens, and a reticle. Sometimes a field lens is employed to ensure uniform illumination. The magnifying powers customarily range from unity to 12 diameters, magnifications of $2^1/_2$ to 9 being most common. Instruments are available in which the power may be varied optically over a wide range to suit existing conditions. Large eyepieces are used to provide a wide field of view and a substantial eye relief, the latter so that the eyepiece will not strike the user's eye in recoil. In order to adjust for elevation or windage, either the reticle may be moved, usually by means of screws, or the entire sight may be tilted relative to the gun barrel. Machine-gun sights are also telescopic in nature and customarily employ a prism for erecting the image. Magnifying powers range between $1^1/_2$ and 3, and the reticle may be illuminated by a small lamp to permit night use. *See* EYEPIECE; TELESCOPE.

Aircraft gunsights. These are usually of the reflector type (also known as reflex sights) and employ in their simplest form a lamp, a reticle, a collimating lens, and a glass plate or partially reflecting mirror (**Fig. 2**). The collimator images the reticle pattern at infinity, and the mirror superimposes this image over the target area. The collimator's pupil is so large that the eye can be positioned at any point within a large cylindrical volume in back of the sight. The collimating system may employ a simple spherical mirror, a Schmidt lens, or a Mangin mirror. A Mangin mirror consists of a negative meniscus lens, the shallower

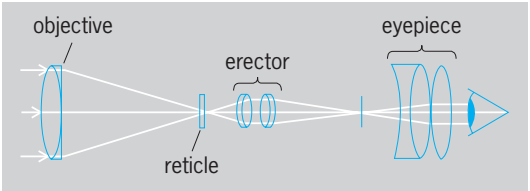

Fig. 1. Rifle sight.

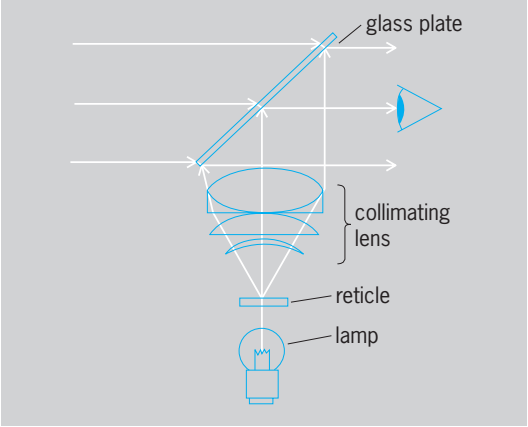

Fig. 2. Reflex sight.

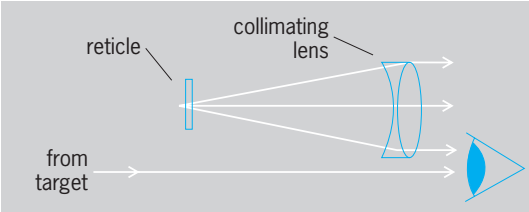

Fig. 3. Collimator sight.

surface of which is silvered to act as a spherical mirror, while the other surface corrects for the spherical aberration of the reflecting surface. The magnification of this type of sight is unity. *See* MIRROR OPTICS.

Artillery sights. These can assume various forms, the simplest of which is the collimator sight (**Fig. 3**), consisting of an objective having a reticle at its focus. When the eye is so placed as to receive light simultaneously from the target and the reticle, the latter appears superimposed on the former and a line of sight is established. Other forms of artillery sights include the elbow telescope, which is a conventional terrestrial telescope containing a prism for producing a right-angle bend, and the panoramic sight. For a discussion of panoramic sights *see* LENS (OPTICS); PERISCOPE. Edward K. Kaprelian

Bibliography. R. Kingslake, *Optical System Design*, 1983; Optical Society of America, *Handbook of Optics*, 2d ed., 1994; W. J. Smith, *Modern Optical Engineering*, 3d ed., 2000; B. H. Walker, *Optical Engineering Fundamentals*, 1995.

Gymnolaemata

A class of predominantly marine, colonial animals belonging to the phylum Bryozoa. Gymnolaemate zooids (individual colony members) are characterized by their short, wide, vaselike or boxlike zooecia (skeletal outer structures) and circular lophophores (food-gathering organs).

Classification and history. The gymnolaemates include several thousand species—mostly marine, some brackish, and a very few freshwater—belonging to the orders Ctenostomata and Cheilostomata. Appearing early in the Early Ordovician, when they included the probable stem group of the Ectoprocta, gymnolaemates remained quite inconspicuous until the Cretaceous, when they rose to the position of dominance among bryozoans which they still maintain. *See* BRYOZOA; CHEILOSTOMATA; CTENOSTOMATA.

Morphology. Highly varied in size and shape, most gymnolaemate colonies are small and delicate, but a few are large, conspicuous growths. Characteristics of these colonies remain essentially uniform throughout their extent (that is, they are not divisible into distinct endozone and exozone regions). The individual zooids may be relatively isolated, loosely grouped side by side, or tightly packed together.

The boxlike gymnolaemate zooecia lack internal calcareous cross-partitions like diaphragms. The zooecial wall may be a thin chitinous membrane (rarely coated externally with gelatinous material), a moderately thick and firm chitinous wall, or a thin to thick calcareous wall (covered externally by a thin chitinous cuticle). Just inside the zooecial wall is an epidermis, and immediately interior to that is a peritoneum, both composed of flat, thin cells; however, the body wall lacks a muscle layer. The zooecial aperture may be as wide as, but is often narrower than, the zooecium below.

An epistome (flap covering the mouth) is not present. The visceral cavities of adjacent zooids are not directly interconnected. The lophophore bears few to moderate numbers of tentacles (8–34, where counted). Individual gymnolaemate zooids are usually hermaphroditic, sometimes male or sometimes female. Individual colonies are always hermaphroditic, usually because they bear hermaphroditic zooids but sometimes because they bear both male and female zooids.

Life cycle. Asexual budding, which produces a new zooid within a gymnolaemate colony, begins by separating off a small sac- or boxlike portion of the parent zooid's visceral or body cavity. Cells along the inside of the body wall (cystids) then develop into a new set of internal soft-part organs (polypides) suspended in the newly separated sac or box. Periodically, the polypides of gymnolaemate zooids degenerate into brown bodies; the new succeeding polypides are regenerated from the original zooids' body walls. Only the very few freshwater gymnolaemates produce resistant resting bodies, known as hibernacula. Roger J. Cuffey

Bibliography. A. H. Cheetham and P. L. Cook, General features of the class Gymnolaemata, pp. 138–207 in R. A. Robison (ed.), *Treatise on Invertebrate Paleontology*, part G, *Bryozoa*, revised, vol. 1, Geological Society of America & University of Kansas, Boulder & Lawrence, 1983; R. J. Cuffey and J. E. Utgaard, *Bryozoans*, pp. 204–216, vol. 1, in R. Singer (ed.), *Encyclopedia of Paleontology*, Fitzroy Dearborn, Chicago, 1999; F. K. McKinney and J. B. C. Jackson, *Bryozoan Evolution*, Unwin Hyman, Boston, 1989; J. S. Ryland, *Bryozoans*, Hutchinson University Library, London, esp. pp. 27, 29, 43–85, 1970; P. D. Taylor, *Bryozoa*, pp. 465–488 in M. J.

Benton (ed.), *The Fossil Record 2*, Chapman & Hall, London, 1993.

Gymnostomatida

An order of the Holotrichia which contains a large, widely distributed group of what are believed to be the most primitive ciliate protozoa. These organisms occur abundantly in sands of intertidal zones, as well as in the more usual fresh- and salt-water habitats. The body size is frequently large; ciliation is simple and plentiful; and the oral area lacks buccal ciliature, as the name of the order implies. The cytopharynx, however, is reinforced by fibrillar rods known as trichites, or nemadesmata. Members of two families exist as harmless commensals in such herbivores as horses and camels and in a few other mammals. *Prorodon*, *Holophrya*, *Didinium*, and *Dileptus* are examples of carnivorous gymnostomes. *Chilodonella* and *Nassula* are herbivorous and anatomically more complex. *Ichthyophthirius*, a fish parasite, is often classified erroneously as a gymnostome, but belongs in the order Hymenostomatida. It is found on fish in home aquaria and causes the disease known as the ich. *See* CILIOPHORA; HOLOTRICHIA; HYMENOSTOMATIDA; PROTOZOA.

John O. Corliss

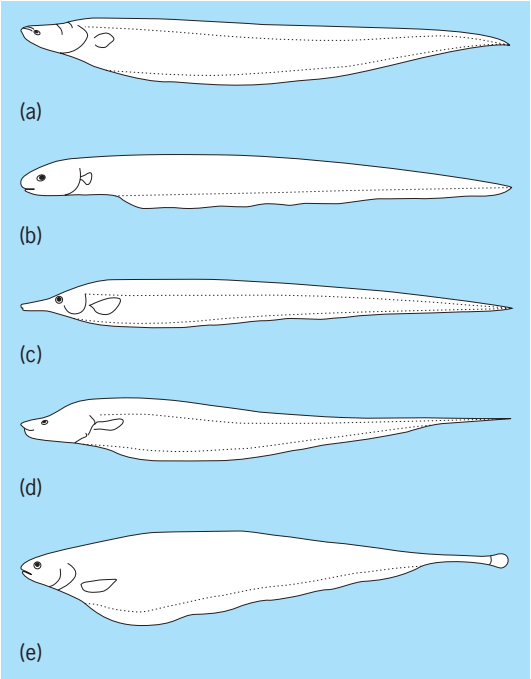

Gymnotiform fishes. (*a*) Nakedback knifefish. (*b*) Electric knifefish. (*c*) Sand knifefish. (*d*) Glass knifefish. (*e*) Ghost knifefish. (*From J. S. Nelson, Fishes of the World, 4th ed., Wiley, New York, 2006*)

Gymnotiformes

An order of teleost fishes in the superorder Ostariophysi, series Otophysi, characterized by a combination of the following characters: Weberian ossicles (a series of small bones that form a chain connecting the swim bladder with the inner ear); an eel-like body, usually compressed but sometimes rounded; absence of a dorsal fin and absence of a pelvic girdle and pelvic fins; a very long anal fin; a caudal fin that is absent or very small; and electric organs. The anal rays are articulated in a way that allows for circular motion; thus undulation of the anal fin can move the gymnotiform forward or backward. The order Gymnotiformes, commonly known as American knifefishes, comprise five families (see **illustration**) described below. *See* ELECTRIC ORGAN (BIOLOGY); OSTARIOPHYSI; SWIM BLADDER.

Gymnotidae (nakedback knifefishes). Nakedback knifefishes are represented by two genera. The principal genus, *Gymnotus*, with about 32 species, ranges in freshwater through southern Mexico, Central America, South America, to Argentina, as well as Trinidad. It has a superior mouth, small scales, and a weak electrical discharge, and attains a total length of 60 cm (24 in.).

The other genus, *Electrophorus*, with one species, *E. electricus* (electric knifefish, commonly known as the electric eel), was formerly in the family Electrophoridae. It occurs primarily in the Orinoco and Amazon river basins, and differs from *Gymnotus* in having a subterminal mouth, no scales, and a strong electrical discharge (up to 600 volts). It attains a total length of 230 cm (90 in.). Gymnotids are nocturnal fishes, usually occurring in quiet waters from deep rivers to swamps. *See* EELS.

Rhamphichthyidae (sand knifefishes). Sand knifefishes, with three genera and 12 species, are known only from freshwaters of South America. The following combination of characters distinguishes them from other gymnotids: an elongated compressed body, no caudal fin, a long tubular snout, nostrils close together, no teeth on lower jaws, and myogenic (muscular) electric organs. They are nocturnal fishes that burrow in the sand during daylight hours. The maximum recorded length is 100 cm (39 in.).

Hypopomidae (bluntnose knifefishes). Hypopomids range from the Río de la Plata of Argentina (35°S) to the Río Tuira of Panama (8°N) and are known from all South American countries except Chile. They have their greatest diversity in the Amazon River basin. The family, consisting of about 35 species in six genera, is distinguished from other gymnotiforms by the absence of teeth from both jaws, a moderate to short snout length, well-separated nostrils, small eyes, and no caudal fin. The smallest species attains a total length of only 8 cm (3 in.), and the largest a moderate length of 35 cm (14 in.).

Sternopygidae (glass knifefishes). The range of glass knifefishes is essentially the same as that of the bluntnose knifefishes. The family comprises six generia and 30 valid species, with 11 more species assigned manuscript names and awaiting descriptions. The family possesses the following unique combination of characters among gymnotiforms: multiple rows of small villiform (brushlike) teeth in both jaws, large eyes [diameter equal to or greater than the distance between the nares (nostrils)], large infraorbital

bones with expanded bony arches, anterior nares located outside the gape, no urogenital papilla, and no caudal fin. Some sternopygid species are specialized for life in main river channels, and many species are highly compressed laterally and translucent. Sternopygid species are medium- to large-sized, varying in total length from 12 to 140 cm (4.7 to 55 in.).

Apteronotidae (ghost knifefishes). The range of ghost knifefishes is essentially the same as that of the bluntnose knifefishes and glass knifefishes. Apteronotids are the only gymnotiform fishes with a caudal fin and a dorsal organ (a longitudinal strip of fleshy tissue firmly attached to the posterodorsal midline). Apteronotid species are further characterized by one or two rows of conical teeth in both jaws, infraorbital bones ossified as slender tubes, anterior nares located outside the gape, and no urogenital papilla. Apteronotids also possess a high-frequency tone-type electric organ discharge (more than 750 Hz at maturity). There are 52 valid species in 13 genera, and 12 additional species assigned manuscript names and awaiting descriptions. Apteronotid species range in total length from 16 to 130 cm (6 to 51 in.). Herbert Borschung

Bibliography. J. S. Albert, *Species Diversity and Phylogenetic Systematics of American Knifefish (Gymnotiformes, Teleostei)*, Misc. Publ. Mus. Zool. Univ. Mich. 190, 2001; J. S. Albert, Sternopygidae (Glass knifefishes, Rattail knifefishes), pp. 487–491, and Hypopomidae (Bluntnose knifefishes), pp. 494–496 in R. E. Reis, S. O. Kullander, and C. J. Ferraris, Jr. (eds.), *Checklist of the Freshwater Fishes of South and Central America*, Edipucrs (Editoria Universitaria da PUCRS), Porto Alegre, Brasil, 2003; J. S. Albert and R. Campos-da-Paz, Phylogenetic systematics of Gymnotiformes with diagnoses of 58 clades: A review of the available data, pp. 419–446 in L. R. Malabarba et al. (eds.), *Phylogeny and Classification of Neotropical Fishes*, Edipucrs, 1998; J. S. Albert and W. G. R. Crampton, Family Hypopomidae (bluntnose knifefishes), pp. 494–496 in R. E. Reis et al., (eds.), *Checklist of the Freshwater Fishes of South and Central America*, Edipucrs, 2003; C. J. Ferraris, Jr., Family Rhamphichthyidae (sand knifefishes), pp. 492–493 in R. E. Reis et al. (eds.), *Checklist of the Freshwater Fishes of South and Central America*, Edipucrs, 2003; R. Froese and D. Pauly (eds.), FishBase, World Wide Web electronic publication, www.fishbase.org, version 05/2005; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Gypsum

The most common sulfate mineral, characterized by the chemical formula $CaSO_4 \cdot 2H_2O$; it shows little variation from this composition.

Gypsum is one of the several evaporite minerals. This mineral group includes chlorides, carbonates, borates, nitrates, and sulfates. These minerals precipitate in seas, lakes, caves, and salt flats due to concentration of ions by evaporation. When heated or subjected to solutions with very large salinities, gypsum converts to bassanite ($CaSO_4 \cdot H_2O$) or anhydrite ($CaSO_4$). Under equilibrium conditions, this conversion to anhydrite is direct. The conversion occurs above 108°F (42°C) in pure water. The presence of halite (NaCl) or other sulfates in the solution lowers this temperature, although metastable gypsum exists at higher temperatures. *See* ANHYDRITE; HALITE.

Crystals of gypsum are commonly tabular, diamond-shaped, or lenticular; swallow-tailed twins are also common. The mineral is monoclinic with symmetry $2/m$. The common colors displayed are white, gray, brown, yellow, and clear. Cleavage surfaces show a pearly to vitreous luster. Gypsum is the index mineral chosen for hardness 2 on Mohs scale with a specific gravity of 2.32. In addition to free crystals, the common forms of gypsum are satin spar (fibrous), alabaster (finely crystalline), and selenite (massive crystalline).

Gypsum is used for a variety of purposes, but chiefly in the manufacture of plaster of paris, in the production of wallboard, in agriculture to loosen clay-rich soils, and in the manufacture of fertilizer. Plaster of paris is made by heating gypsum to 392°F (200°C) in air. A hemihydrate is formed as part of the water of crystallization is driven off. Later, when water is added, rehydration occurs. The interlocking, finely crystalline texture that results forms a uniform hardened mass. The slightly increased volume of the set plaster serves to fill the mold into which it has been poured. *See* PLASTER OF PARIS.

Gypsum deposits play an important role in the petroleum industry. The organic material commonly associated with its formation is considered a source of hydrocarbon (oil and gas) generation. In addition, these deposits act as a seal for many petroleum reservoirs, preventing the escape of gas and oil. *See* PETROLEUM.

Because most gypsum deposits form by evaporation from solution, they are indicators of former arid and semiarid climates. The worldwide distribution of gypsum is an indicator of climatic change through geologic time. Large-scale gypsum deposits are found in strata of nearly all ages, the oldest recognized being 3.4×10^9 years old (North Pole, Pilbara Block, Australia). The largest known deposits formed in Silurian (400×10^6 years ago), Pennsylvanian-Permian (295–245×10^6 years), Jurassic (200×10^6 years), and Miocene (5×10^6 years) time.

Gypsum deposits are mined throughout the world, with the United States being a world leader in gypsum production. The majority of United States gypsum is mined in Michigan, Iowa, Texas, California, and Oklahoma. Canada is the world's second largest producer. Most Canadian production is in the province of Nova Scotia. Among the other leading producers are France, Japan, Iran, Russia, Italy, Spain, and the United Kingdom.

Marc L. Helman; Charlotte Schreiber

Bibliography. American Society for Testing and Materials, *Annual Book of ASTM Standards*, vol. 04.01: *Cement; Lime; Gypsum*, 1993; Bureau

of Mines–U.S. Department of the Interior, *Minerals Yearbook*, centennial ed., vol. 1: *Metals and Minerals*, 1982; L. L. Y. Chang, R.A. Howie and J. Zussman, *Rock-Forming Minerals*, vol. 5B: *Nonsilicates: Sulphates, Carbonates, Halides*, 2d ed., 1996; C. J. Dixon, *Atlas of Economic Mineral Deposits*, 1979; C. Klein and C. S. Hurlbut, Jr., *Manual of Mineralogy,* 21st ed., 1993, revised 1998; C. C. Plummer and D. McGeary, *Physical Geology,* 8th ed., 1999.

Gyrator

A linear, passive, two-port electric circuit element whose transmission properties are such that it is effectively a half wavelength longer for one direction of transmission than for the other direction of transmission. Thus a gyrator is a device that causes a reversal of signal polarity for one direction of propagation but not for the other. (A two-port element has a pair of input terminals and a pair of output terminals.) This device is novel, since it violates the theorem of reciprocity. *See* RECIPROCITY PRINCIPLE.

Until the early 1950s, all known linear passive electrical networks obeyed the theorem of reciprocity. However, several different types of nonreciprocal networks are now widely applied, principally at microwave frequencies. These devices are used to control the direction of signal flow and to protect or isolate components from undesired signals. One common application of a three-port nonreciprocal network, called a circulator, is to permit connection of a transmitter and a receiver to the same antenna. This is accomplished with minimum interference and virtually no power loss of either transmitted or received signal. *See* CONTINUOUS-WAVE RADAR; NETWORK THEORY.

Reciprocity. The theorem of reciprocity can be stated in many equivalent forms, and perhaps the simplest to consider and understand is the particular form it takes when it is expressed especially for a two-port microwave network, as shown in **Fig. 1***a*. Here d_1 represents a complex signal wave propagating toward the network input port, r_1 represents the wave propagating away from the network on the same side, and d_2 and r_2 represent the same quantities on the other side. In a linear network the relationship between the waves can be expressed as shown in Eqs. (1). Thus r_1 consists of a part which

$$\begin{aligned} r_1 &= S_{11}d_1 + S_{12}d_2 \\ r_2 &= S_{21}d_1 + S_{22}d_2 \end{aligned} \tag{1}$$

arises from the partial reflection of d_1 at the input port and a part contributed by the fraction of d_2 which is transmitted through the network. The reflection and transmission properties of the network are described by the complex coefficients S_{ij}, which are termed the scattering parameters of the network.

The network can be characterized by selectively setting one of the input waves to zero and measuring the resulting response. Thus when d_2 is zero, Eq. (2)

$$S_{21} = \frac{r_2}{d_1} \tag{2}$$

holds. The quantity S_{21} gives the amplitude and phase of the wave emerging from the right-hand side of the network when a unit wave, with zero phase angle, is incident on the left-hand side. The significance of the other scattering parameters can be deduced by similar arguments. If there is no reflection of the incident waves, then S_{11} and S_{22} are zero, and the network is said to be matched.

The theorem of reciprocity is stated by Eq. (3).

$$S_{12} = S_{21} \tag{3}$$

That is, the network has the same transfer characteristics for one direction of propagation as it has for the other. Thus, if a matched network causes a particular insertion loss and phase shift for signals transmitted from left to right, signals transmitted from right to left must suffer the same loss and phase shift. An ideal gyrator is one which has the scattering matrix given in Eq. (4), and hence violates reciprocity.

$$[S] = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} 0 & -S_{21} \\ S_{21} & 0 \end{bmatrix} \tag{4}$$

The terminal currents i_1, and i_2 and voltages of V_1, and V_2 of the network (Fig. 1*b*) are related by impedances Z_{ij} as shown in Eqs. (5). The theorem of

$$\begin{aligned} V_1 &= Z_{11}i_1 + Z_{12}i_2 \\ V_2 &= Z_{21}i_1 + Z_{22}i_2 \end{aligned} \tag{5}$$

reciprocity now requires the condition of Eq. (6).

$$Z_{12} = Z_{21} \tag{6}$$

Theoretical gyrators. The first comprehensive treatment of nonreciprocal two-port networks was given in 1948 by B. D. H. Tellegen, who applied the word gyrator to describe such networks. Tellegen restricted his analysis to an ideal gyrator whose impedance matrix has the form given in Eq. (7). It

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} = \begin{bmatrix} 0 & -R \\ R & 0 \end{bmatrix} \tag{7}$$

may be easily shown that for such a gyrator to be

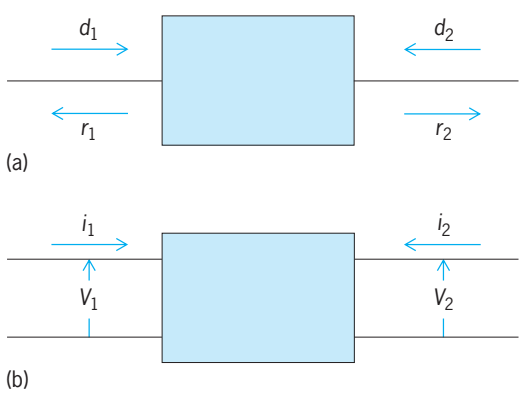

Fig. 1. Two-port microwave network. (*a*) Scattering matrix representation. (*b*) Currents and voltages.

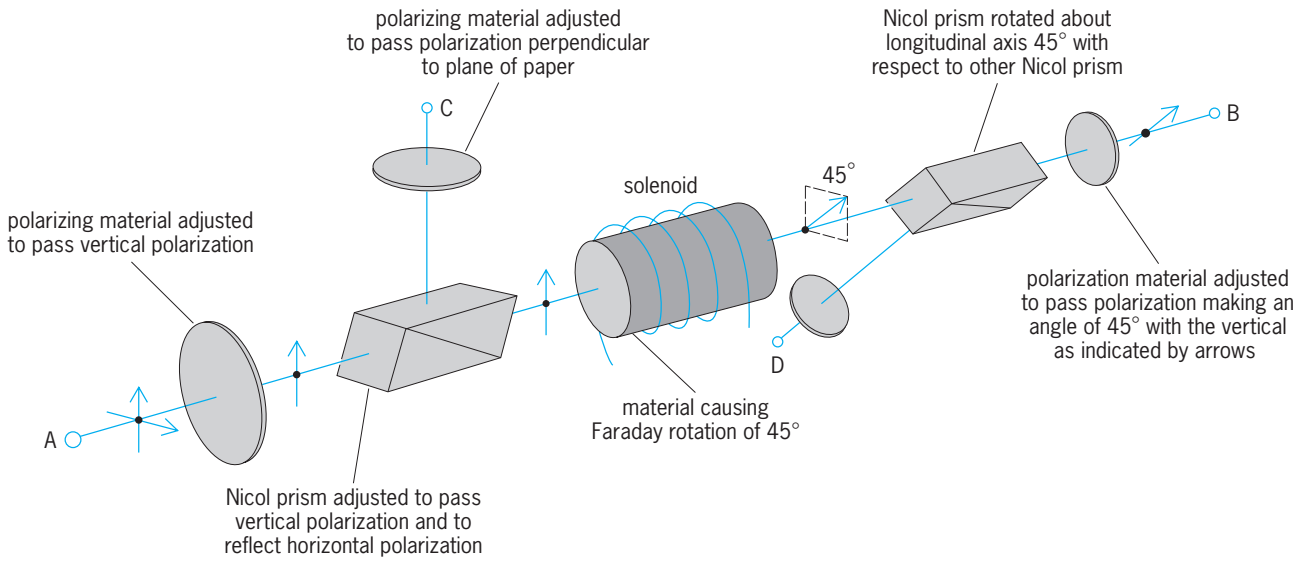

Fig. 2. Passive nonreciprocal optical device based on Faraday rotation.

nondissipative, the Z_{ij} must be real. However, it is not necessary that the diagonal terms, Z_{11} and Z_{22}, be zero. Furthermore, this restriction on the diagonal terms is not needed in order to realize two of the most important properties of the gyrator, namely, the construction of one-way transmission systems and the phase inversion of signal polarity for one direction of transmission.

Tellegen's ideal gyrator has the additional property of impedance inversion. If the gyrator is terminated by an impedance Z_L, Eq. (8), the input impedance

$$Z_L = \frac{V_2}{i_2} \tag{8}$$

Z_{in} of the gyrator for Eq. (5) is given by Eq. (9). This property of impedance inversion is not, how-

$$Z_{\text{in}} = \frac{V_1}{i_1} = \frac{R^2}{Z_L} \tag{9}$$

ever, unique to nonreciprocal networks. Indeed, if one makes the same requirement on a nondissipative reciprocal two-port, that $Z_{11} = Z_{22} = 0$. the input impedance Z_{in} of the terminated network is that given by Eq. (10), where X_{12} is the transfer reac-

$$Z_{\text{in}} = \frac{V_1}{i_1} = \frac{X_{12}^2}{Z_L} \tag{10}$$

tance of a nondissipative network, $Z_{12} = jX_{12}$. Such a reciprocal impedance-inverting network can be easily realized.

Practical gyrators. The theorem of reciprocity has in the past been considered so universally valid that present-day textbooks still make the statement that if the condition stated in Eq. (6) is valid, the two-port is passive, and that if the condition does not hold for a particular network, it cannot be a passive one.

Although reciprocity is as universally valid in mechanical, acoustical, or optical systems as it is in electrical ones, there are passive systems that can be constructed in each of these areas which are nonreciprocal. For example, it has been long known that a mechanical system which contains a gyroscopic coupler does not obey the theorem of reciprocity.

Perhaps the first passive nonreciprocal system was an optical one proposed by Lord Rayleigh, making use of the rotation of the plane of polarization of light when it passed through a transparent material in the presence of a magnetic field. This phenomenon is called Faraday rotation. If polarized light is propagated through a transparent medium along the direction of the magnetic field, the plane of polarization of the light is rotated through some angle θ per unit length, which is determined by the properties of the medium and the strength of the magnetic field. Faraday rotation is unusual in that it is nonreciprocal. Thus the sense (clockwise or counterclockwise) of the rotation is the same whether the light travels parallel to the applied magnetic field direction or contraparallel to it. Hence, if the plane of polarization is rotated through an angle θ in traversing the Faraday cell and the ray is reflected back through the cell toward its source, it will again be rotated through an angle θ, so that when it arrives back at the source, the plane of polarization will have been rotated through a total angle 2θ.

Lord Rayleigh's one-way system, shown in **Fig. 2**, consisted of two polarizing Nicol prisms oriented so that their planes of acceptance made an angle of 45° with each other. The material causing the Faraday rotation was placed between them. If the Faraday rotator is adjusted to cause a 45° rotation, light which is passed by the first crystal is passed by the second also. In the reverse direction, however, the 45° rotation added by the Faraday cell produces light rays polarized horizontally which are reflected by the Canada balsam cement in the first Nicol prism and directed toward point C. Thus light admitted to the device at point A is transmitted to point B; light admitted at point B is transmitted to point C; light admitted at point C is transmitted to point D; and light admitted at point D is transmitted to point A.

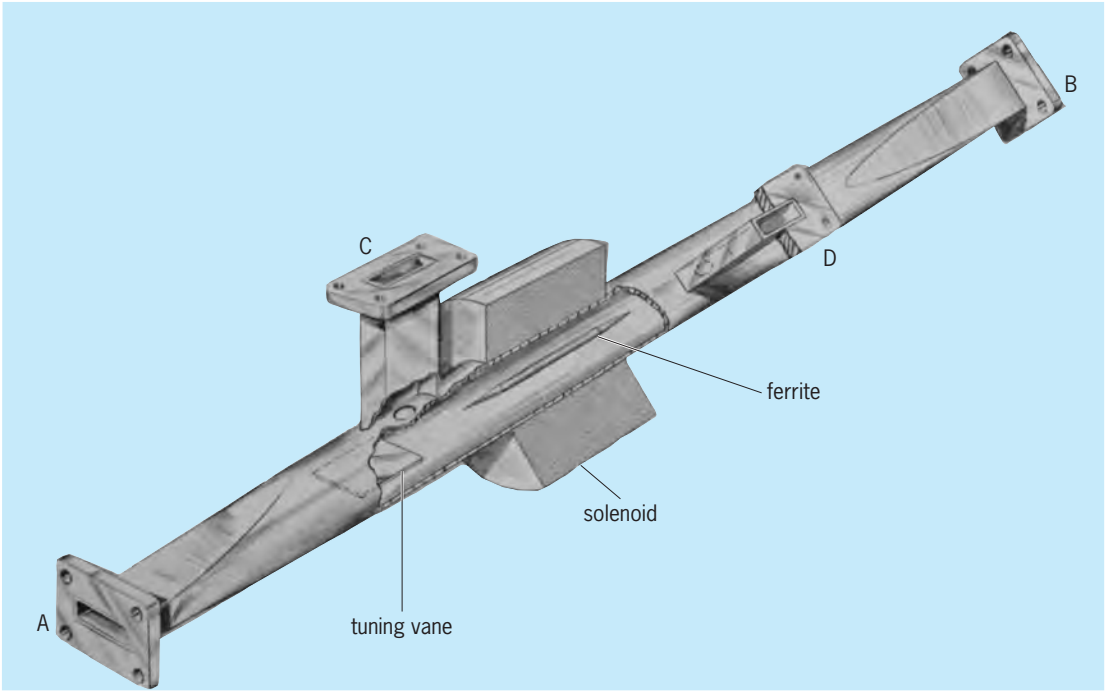

Fig. 3. Microwave analog of nonreciprocal optical device.

See FARADAY EFFECT; POLARIZED LIGHT.

The microwave analogy of Lord Rayleigh's device was proposed by C. L. Hogan and is shown in **Fig. 3**. Since it circulates microwave power from waveguide A to B, from B to C, from C to D, and from D to A, it has been called a (four-port) circulator. The nonreciprocal medium used here is a ferrimagnetic material called ferrite. In such a material, infinitesimal magnetic dipole moments which arise from the electronic structure of the material act gyroscopically when a steady magnetic field is applied, as shown in **Fig. 4**. They precess about the applied field direction in a counterclockwise sense, thus permitting strong coupling to the component of a microwave-frequency magnetic field which is circularly polarized in the same sense. The component with the opposite sense of polarization is weakly coupled. Thus energy exchange between the magnetic dipoles and the microwave field is polarization-sensitive. *See* FERRIMAGNETISM; FERRITE; FERRITE DEVICES.

Present-day circulators utilize the properties of electromagnetic fields in ferrite loaded microwave circuits. Consider the three-port circulator shown in **Fig. 5**, which comprises a circular disk resonator filled with ferrite connected to three transmission lines. When microwave energy is transmitted to the resonator along one of the transmission lines, an electromagnetic field is established in the resonator which is stationary in space, as shown in **Fig. 6**. The application of a dc magnetic field perpendicular to the plane of the disk resonator will rotate this stationary pattern through an angle dependent on the strength of the applied field. The field pattern is dipolar in nature, and hence has a region where the microwave magnetic field intensity is low. If the pattern is oriented to position this low-intensity region at one of the output transmission line ports, very little microwave power will leave the resonator via this port: it is isolated from the input. From the symmetry of the junction, this stationary pattern will advance 30°

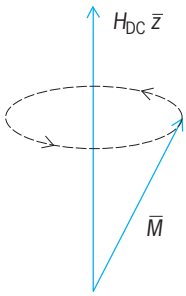

Fig. 4. Precessional motion of the ferrite internal magnetization vector, $\bar{M}$, with a static magnetic field H_{DC} applied in the $\bar{z}$ direction.

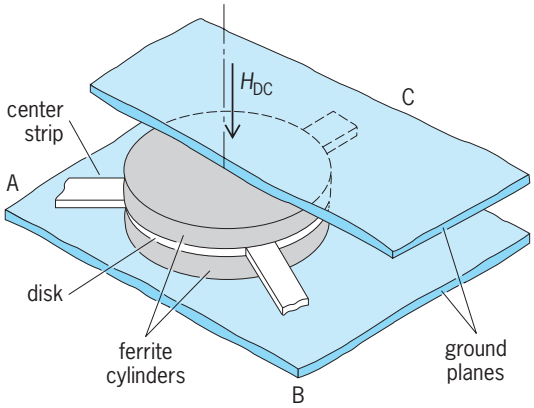

Fig. 5. Stripline Y-junction circulator. (*After C. E. Fay and R. L. Comstock, Operation of ferrite junction circulator, IEEE Trans. Microwave Theory Tech., MTT-13:1–13, January 1965*)

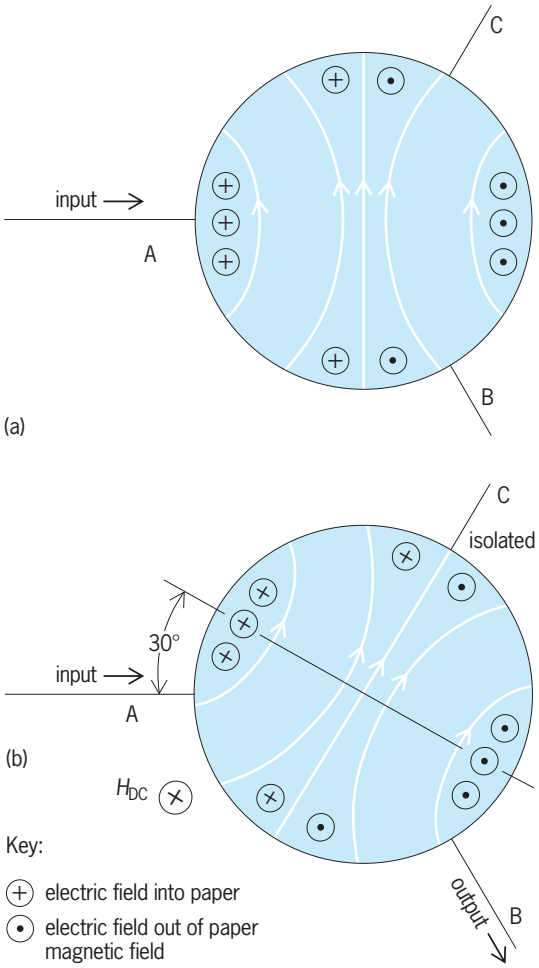

Fig. 6. Stationary electromagnetic field pattern in disk resonator of stripline circulator. (*a*) Junction is not magnetized, resulting in symmetric coupling to ports B and C. (*b*) Junction is magnetized to rotate pattern 30°, for circulation. Port C is isolated. (*After C. E. Fay and R. L. Comstock, Operation of ferrite junction circulator, IEEE Trans. Microwave Theory Tech., MTT-13:1–13, January 1965*)

if the ports are excited sequentially. Thus energy at port A will be transmitted to port B, energy incident on port B will be transmitted to port C, and energy incident on port C will be transmitted to port A. The ideal three-port scattering matrix of this junction circulator is given in Eq. (11).

$$[S] = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{11}$$

Devices based on this pattern rotation principle have been realized in a variety of transmission line geometries, including rectangular waveguide, stripline, and microstrip. High performance can be obtained over wide microwave-frequency bandwidths (with ratio of upper to lower frequency limit less than or equal to 4:1) with isolation in the order of 0.01 of the incident power and with about 95% of the power transmitted to the desired port. These devices are found in most microwave systems, where they are employed to control the flow of microwave signals. For example, their use permits a radar transmitter and a radar receiver to share the same antenna.

In another use, a transmission line–matched termination connected to one of the ports of a three-port circulator produces a one-way transmission device, known as an insolator. (Any signal incident on this termination is absorbed.) Such two-port devices are widely used to prevent unwanted reflections from altering the properties of oscillators or amplifiers and to eliminate frequency-dependent transmission behavior caused by reflections in networks. Isolators realized from terminated circulators are compact, operate over wide bandwidths, and can be made to dissipate high microwave power levels through the use of an external high-power termination. Commercial low-power devices are in "drop-in" form compatible with microwave integrated (printed) circuit technology. *See* INTEGRATED CIRCUITS.

Fred J. Rosenbaum

Bibliography. R. E. Collin, *Fundamentals for Microwave Engineering*, 1992; J. Helszajn, *Microwave Engineering: Passive, Active, and Non-Reciprocal Circuits*, 1992; S. Y. Liao, *Microwave Devices and Circuits*, 3d ed., 1990.

Gyrocompass

A north-seeking form of gyroscope used as a directional reference in navigation. The first practical gyrocompasses were developed by H. Anschütz (Germany) in 1908, E. A. Sperry (United States) in 1911, and S. G. Brown (England) in 1916. Modern gyrocompasses are so reliable and so much more accurate than magnetic compasses that they are now used as the prime navigational instrument on nearly every ship and on major aircraft and missiles. *See* MAGNETIC COMPASS.

Basic operation. A gyrocompass combines the action of two devices, a pendulum and a gyroscope, to produce alignment with the Earth's spin axis. The principle is demonstrated with the model of **Fig. 1**, which consists of a rapidly spinning, heavy gyro

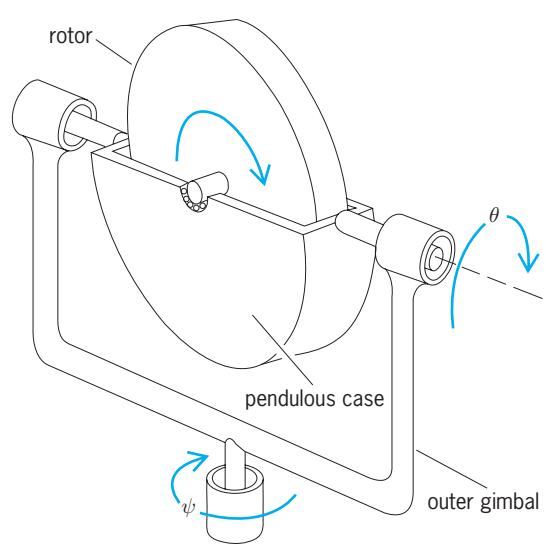

Fig. 1. Gyrocompass model.

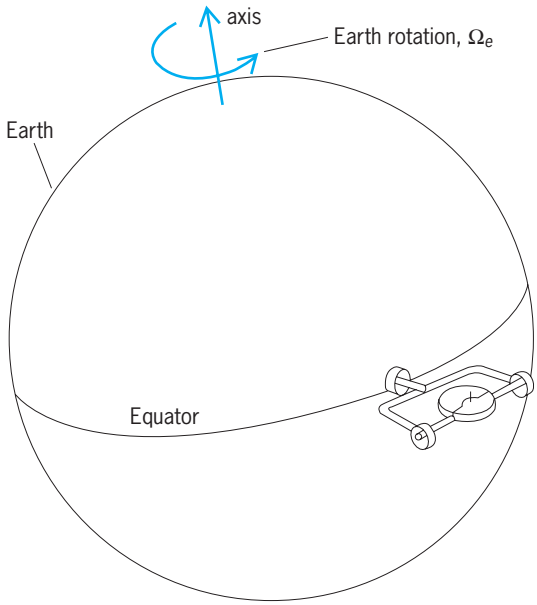

Fig. 2. Precession due to Earth's rate of rotation.

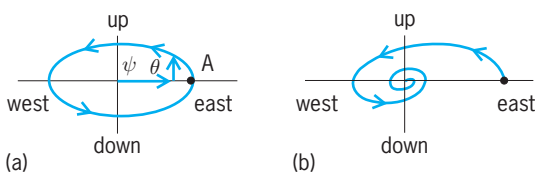

Fig. 3. Path of rotor axle of gyrocompass. (*a*) Typical undamped pattern. (*b*) Pattern with damping added.

rotor, a pendulous case which permits the rotor axle to nod up and down (angle θ), and an outer gimbal which permits the axle to rotate in azimuth (angle ψ).

In **Fig. 2** the model is positioned at the Equator of the Earth. As the Earth rotates, the gimbal moves with it. So long as the rotor's spin axis is aligned with the Earth's axis, the gyro experiences no torque from Earth rotation. If there is misalignment, however, a sequence of restoring torques is initiated.

The restoring action can be seen by imagining a beam of light to be projected by the gyro rotor axle onto a vertical piece of paper held just at the north end of the axle. **Figure 3*a*** shows a typical pattern traced on the paper by the light spot. At point A the rotor axle has been displaced east by $\psi = 5°$. The spin axis of the gyro is no longer parallel to that of the Earth. Therefore, as the Earth rotates, the north end of the gyro rotor tilts upward. The ensuing dip angle results in a pendulum torque, down at the south end and up at the north end of the rotor axis. This torque causes the north end of the gyro axis to precess westward. As soon as the gyro axis passes through the north-south meridian, the action reverses. The north end of the rotor tilts down and the pendulum torque causes the gyro to precess eastward. Thus, the axis traces the elliptic path in Fig. 3*a*. In practical instruments damping is added and the path converges to true north in a spiral motion, as in Fig. 3*b*. *See* GYROSCOPE.

Shipboard installation. In a shipboard installation the system must be mounted in a complete set of gimbals to isolate it from rolling, pitching, and yawing motions of the ship. Friction must be minimized. Moreover, Schuler tuning is employed to keep horizontal accelerations of the ship from producing false torques on the pendulum; the unique combination of gyro spin speed and pendulosity is chosen so that no acceleration of the instrument can disturb its vertical reference. Then the light spot in Fig. 3 encircles the true north point once every 84 min. *See* SCHULER PENDULUM.

If the ship travels north or south, the ship's velocity V over the Earth creates an apparent Earth rate of $\Omega_s = V/R$ (R is the Earth's radius) about an axis perpendicular to the Earth's spin axis. The gyrocompass aligns to the vector sum of Ω_e and Ω_s, which results in a north steaming error angle whose tangent is Ω_s/Ω_e. This error is corrected by using V from the ship's pitometer to precess the gyro in the opposite direction at an equal rate.

Practical gyrocompass design is shown by the typical system depicted in **Fig. 4**. The rotor and case of Fig. 4 replace the rotor and case of Fig. 1, except that in Fig. 4 this assembly is not itself pendulous. Instead the case carries a coupling pin which rides in a slot on the ballistic ring, which is pendulous by virtue of the two mercury bottles and connecting tube known as a mercury ballistic. Gyroscope damping is introduced by making the coupling pin slightly eccentric.

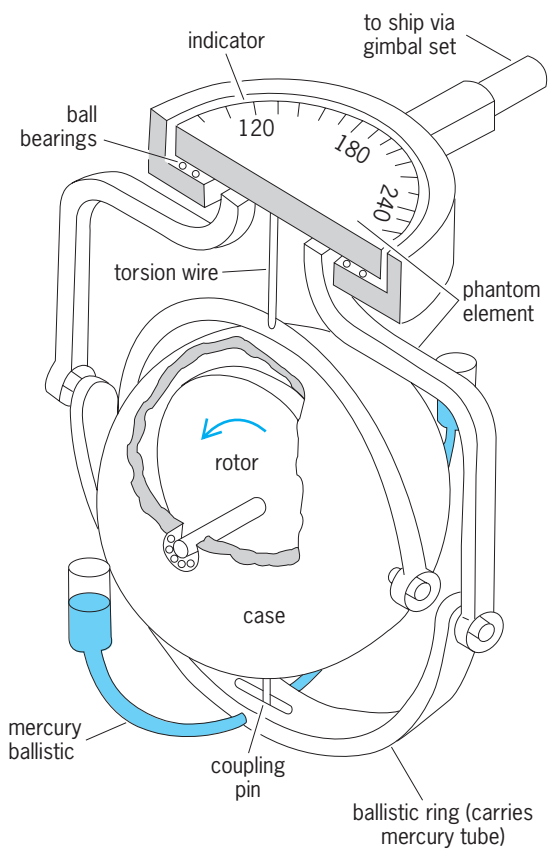

Fig. 4. Typical gyrocompass design.

The ballistic ring is supported by bearings from the phantom element, which replaces the outer gimbal of Fig. 1. The gyro assembly is also supported from the phantom element by a torsion wire which provides frictionless support. A servomotor is arranged to drive the phantom element whenever there is any twist in the wire. Thus the phantom element always stays aligned with the gyro. The phantom element carries an indicator from which ship's heading is read. Robert H. Cannon, Jr.

Aircraft application. For many years the use of gyrocompasses in aircraft was impractical because of their high speed and large, rapid changes in attitude. As mentioned above, the north-south component of vehicle velocity produces an error which depends on the velocity magnitude. Aircraft applications of gyrocompasses therefore use a modified version of the marine gyrocompass. The gyroscopes are mounted on a platform that is stabilized by signals from the gyroscopes. The platform is aligned to the local vertical and to north prior to takeoff by using essentially the same technique as for a marine gyrocompass. The 84-min Schuler period is shortened by amplifying signals from tilt sensors or accelerometers on the platform to rapidly remove platform tilt and align to north. Alignment times range from 5 to 30 min, depending upon the desired accuracy. The heading and vertical, once established, are "remembered" by the gyroscopes during flight. Vehicle velocity can be computed from the accelerometer data and used to correct for vehicle velocity and for dead-reckoning navigation. The need for preflight heading and vertical alignment can be eliminated by "gyrocompassing" this system in-flight by using an independent velocity sensor such as a Doppler radar or the Global Positioning System. *See* DOPPLER RADAR; INERTIAL GUIDANCE SYSTEM; SATELLITE NAVIGATION SYSTEMS.

The combination of the stable inertial platform and the radar allows accurate determination of local vertical and velocity of the aircraft. These data are put into the platform azimuth gyro servo loop to allow it to perform the gyrocompassing function correctly during flight. The Doppler radar must be capable of very high accuracy. To achieve 0.1° compass accuracy, the aircraft velocity must be determined to about 1 knot (0.5 m/s) at a latitude of 45°.

The complex gimbals that support the platform have been eliminated in newer gyrocompasses by the use of body-mounted or strapdown gyroscopes and accelerometers. The development of the ring-laser gyroscope has resulted in small lightweight gyrocompasses of comparable accuracy but lower cost than earlier gimbaled systems. Doppler radar velocity data can be used for in-flight alignment. Ring-laser gyroscopes are very reliable and retain their accuracy over very long periods of time. They are applied extensively in both commercial and military applications.

A gyrocompass must not be confused with a directional gyro. The latter is a gyroscope which is slaved to a magnetic compass and is not itself a north-seeker. Most conventional aircraft are equipped with north-slaved directional gyros. Heinz Buell

Bibliography. A. Frost, *Marine Gyro Compasses for Ships' Officers*, 1982; E. H. Pallett, *Automatic Flight Control*, 4th ed., 1994.

Gyrocotylidea

An order of flatworms, 2–3 cm (1 in.) long, in the intestine of Holocephali (the chimaeras, a subclass of the Chondrichthyes fishes). There are at least 10 species in two genera, *Gyrocotyle* and *Gyrocotyloides*. (The Cestodaria, in which they were formerly included along with the Amphilinidea, is probably not a valid taxon.) They lack proglottids (the segments found in tapeworms) and an intestine, and have a posterior attachment organ, the rosette (modified to form a cuplike sucker at the end of a caudal stalk in *Gyrocotyloides*) [see **illustration**]. Their body margins undulate. Most testes follicles are located in groups near the anterior end. The follicular vitellarium (the part of the female reproductive system that produces nutritive cells filled with yolk) is lateral, from the anterior to posterior body end. The ovary is located in the posterior third of the body. The uterus is sac-shaped or (in one species) branched. A ciliated lycophora larva about 0.3–0.5 mm (0.012–0.02 in.) long, with two separate anterior nephridiopores (external excretory structure openings) and five pairs of posterior hooks, hatches from the egg. The life cycles are unknown. Young hosts harbor many worms; large ones harbor usually no more than two. No pathological effects are known except in high-intensity infections. *See* AMPHILINIDEA; CESTODA; PLATYHELMINTHES. Klaus Rohde

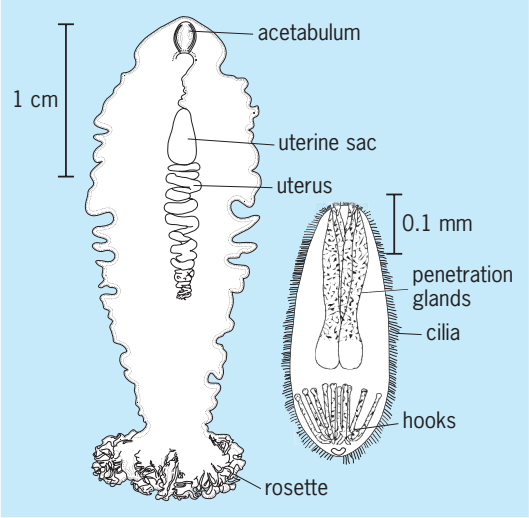

Gyrocotylidean adult and larva.

Bibliography. K. Rohde, The minor groups of parasitic Platyhelminthes, *Adv. Parasitol.*, 33:145–234, 1994; W. Xylander, Gyrocotylidea (unsegmented tapeworms), in K. Rohde (ed.), *Marine Parasitology*, pp. 89–92, CSIRO Publishing, Melbourne, and CABI Publishing, Wallingford, England, 2005.

Gyromagnetic effect

An effect arising from the relation between the angular momentum and the magnetization of a magnetic substance. It is the effect which is exploited in the measurement of the gyromagnetic ratio of magnetic materials. The gyromagnetic effect is demonstrated by a simple experiment in which a freely suspended magnetic substance is subjected to a magnetic field. Upon a change in direction of the magnetic field, the magnetization of the substance must change. In order for this to happen, the atoms must change their angular momentum. Since there are no external torques acting on the system, the total angular momentum must remain constant. Thus the sample must acquire a mass rotation which may be measured. In this way, the gyromagnetic ratio may be determined. Two common methods of determination are the Einstein-de Haas method and the Barnett method. *See* GYROMAGNETIC RATIO.

Einstein-de Haas method. This is usually used to determine the gyromagnetic ratio of ferromagnetic materials. Imagine a ferromagnetic substance in the shape of a cylinder suspended on one end by a torsion fiber, forming thereby a torsional pendulum. A magnetic field is applied parallel to the axis of the cylinder, for example, by means of a coil surrounding the cylinder. The ferromagnetic cylinder is now magnetized in the direction of the magnetic field. If the field is suddenly reversed, the magnetization will reverse, and the accompanying change in angular momentum of the atoms will be balanced by a mass rotation of the cylinder which can be measured by noting the change in amplitude of displacement of the torsional pendulum. A measurement is made of the change in magnetization by means of a magnetometer. The ratio of magnetization change to angular momentum change is the magnetomechanical ratio, the reciprocal of the gyromagnetic ratio.

Barnett method. This is an alternative to the Einstein-de Haas technique. In this experiment, a ferromagnetic bar with a coil surrounding it is spun rapidly about its axis. The atoms, therefore, acquire an angular momentum in the direction of rotation. The magnetization thereby produced is measured by stopping the rotation abruptly and measuring the voltage induced in the surrounding coil which is part of a fluxmeter circuit. Now the same increase in magnetization in the direction of the rotation could be achieved by applying a magnetic field in that direction instead of rotating the bar. In the first case, if an angular momentum L makes an angle θ with the axis of rotation, it will experience a torque $L\omega \sin\theta$, where ω is the angular velocity of rotation, tending to turn it into the axis of rotation. In the second case, the magnetic field H would produce a torque on L in the same direction of $(1/\gamma)\, LH \sin\theta$ where $1/\gamma$ is the magnetomechanical ratio (L/γ = magnetic moment). If the magnetizations in the two cases are the same, then the torques are the same also, and Eq. (1)

$$L\omega \sin\theta = \frac{1}{\gamma} LH \sin\theta \qquad (1)$$

holds. Thus Eq. (2) is obtained.

$$\gamma = \frac{H}{\omega} \qquad (2)$$

In this way the magnetomechanical ratio can be measured. In an experiment similar to the one described, S. J. Barnett showed in 1914 that the magnetomechanical ratio for a ferromagnet was close to e/mc, the ratio of charge to mass of an electron.

Elihu Abrahams; Frederic Keffer

Bibliography. C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996.

Gyromagnetic ratio

The ratio of angular momentum to magnetic moment for atomic systems. This ratio is usually expressed in terms of the magnetomechanical factor g', as in Eq. (1). The ratio is written here in electromagnetic

$$\frac{\text{Angular momentum}}{\text{Magnetic moment}} = \frac{2mc}{g'e} \qquad (1)$$

units; thus, e/c and m are the charge and mass of the electron. The factor g' is sometimes loosely called the gyromagnetic ratio.

Magnetomechanical ratio. This quantity is the inverse of the gyromagnetic ratio. It is usually denoted by γ and is equal to $g'e/2\,mc$.

The magnetomechanical ratio of a substance identifies the origin of the magnetic moment. For example, for electron spin the angular momentum is $^1/_2\,\hbar$, where $\hbar$ is Planck's constant divided by 2π. The magnetic moment is the Bohr magneton $e\hbar/2mc$. Thus, the magnetomechanical ratio is given by Eq. (2).

$$\gamma = \frac{e\hbar/2mc}{\hbar/2} = \frac{e}{mc} \qquad (2)$$

Since $\gamma = g'e/2mc$, for electron spin $g' = 2$. For orbital angular momentum, $\gamma = e/mc$ and $g' = 1$. The experimental values of g' for most ferromagnetic materials are in the neighborhood of 2, showing that the major contribution to the magnetization comes from the electron spin. Deviations of g' from 2 show the extent to which the orbital motion contributes to the magnetization. In superconductors, the fact that $g' = 1$ shows that the diamagnetic currents which cause the Meissner effect are caused by electrons. For discussion of the measurement of the magnetomechanical ratio *see* ELECTRON SPIN; GYROMAGNETIC EFFECT; MEISSNER EFFECT; SUPERCONDUCTIVITY.

Spectroscopic splitting factor. This is the measure of the energy-level splittings of an atomic system in a magnetic field. For a free electron spinning in a magnetic field H, the energy levels are split according to $\Delta E = g\mu_B H$ where μ_B is the Bohr magneton and g is the spectroscopic splitting factor, or g factor, and is equal to 2.00. For electron spins in paramagnetic salts, there are complicated energy-level schemes because of the spin-orbit coupling and crystalline field interactions; the g factor is different from case to case and may depend upon the orientation of the magnetic field with respect to the crystal axes. Under

these circumstances, the g factor is defined quantum mechanically. For ferromagnetic materials, the spectroscopic splitting factor is defined similarly; it is the factor in the ferromagnetic resonance condition which gives the splitting of the energy levels in the magnetic field. It is not generally equal to the magnetomechanical factor g', as the two are affected differently by the effects of orbital angular momentum. For free atoms, the spectroscopic splitting factor is identical to the Landé g factor. For further discussion of g factors. *See* MAGNETIC RESONANCE.

Elihu Abrahams; Frederic Keffer

Gyroscope

A device that is used to define a fixed direction in space or to determine the change in angle or the angular rate of its carrying vehicle with respect to a reference frame. Gyroscopes (also called gyros) respond to vehicle angular rates, that is, rates of change of angles between vehicle axes and reference axes, from which these angles can be computed. Gyros are used for guidance, navigation, and stabilization, for example, to measure the angular deviation of a guided missile from its desired flight trajectory; to determine the heading of a vessel for steering; to determine the heading of an automobile as it turns through city streets; to indicate the heading and orientation of aircraft during and after a series of maneuvers; and to stabilize and point radar dishes and satellites. *See* AIR NAVIGATION; GYROCOMPASS; INERTIAL GUIDANCE SYSTEM; MARINE NAVIGATION; NAVIGATION.

Gyros can be utilized either mounted on a stabilized platform, whose orientation in a moving vehicle remains fixed in space by means of two, three, or four gimbals, or directly attached to the vehicle's body, so that the gyro experiences the same maneuvering as the vehicle, an operating mode referred to as strapdown. Strapdown operation is desirable because it enables a much less expensive system; it became feasible only in the 1970s and 1980s, when the very high digital computing speed required to mechanize the strapdown algorithms became available.

Gyros can be operated closed-loop or open-loop. Closed-loop means that feedback from the gyro output introduces a restoring mechanism either inside the gyro (for example, torquing in mechanical gyros) or counterrotating platform motions to maintain the gyro at its null (initial) setting. In open-loop operation, the gyro is allowed to operate off its null position as it responds to the input angular rates. *See* CONTROL SYSTEMS; SERVOMECHANISM.

Gyro performance is usually characterized by the drift-rate parameter, which is routinely reported in degrees per hour. A drift rate of 0.015°/h corresponds to an accumulated position error of 1 nautical mile (1.85 km) per hour on the Earth's surface at the Equator. Another important performance parameter is angle random walk, which is reported in degrees per square root hour. This is an error which results from the integration of white noise into the angular rate measurement. The most precise navigation or guidance systems require performance to better than 0.001°/h; aircraft navigation requires 0.01°/h; tactical missiles and automobile navigation require 0.1–10°/h; autopilots, seeker heads, and so forth, require 10–100°/h; and automobile chassis control and virtual-reality games require only a few thousand degrees-per-hour performance. Incorporating Global Position System (GPS) receivers in an integrated navigation system can reduce the requirements on gyro accuracy by an order of magnitude. Many gyro error sources such as bias (that is, the output even when there is no input rate), sensitivity to temperature, and acceleration can be calibrated or modeled and compensated for, although not completely, in the navigation or control-system computer. *See* AUTOPILOT; SATELLITE NAVIGATION SYSTEMS.

Gyroscopes use different physical phenomena to respond to input angular rates; for example, spinning-mass gyros sense changes in angular momentum from Coriolis acceleration; resonator gyroscopes sense deflections from Coriolis acceleration; and optical gyros sense phase shifts (the Sagnac effect) between counterpropagating beams of light. Instruments that do not have spinning masses are not technically gyros but angular rate sensors. However, the term "gyro" is commonly used for all rate-sensing devices. *See* CORIOLIS ACCELERATION.

Gyroscopic precession. The classical spinning-mass gyroscope is based upon the phenomenon that the spin axis of a spinning mass points in a fixed direction in space unless acted upon by an external influence. This phenomenon can be observed with the toy gyroscope, whose orientation remains generally fixed (that is, defines a fixed reference) no matter how its base is moved. (Such a body is called a free gyro.) However, the spin axis can be made to rotate if a torque (or rotational rate) is applied at right angles to the spin vector. The spin vector then begins to rotate (precess) about a third axis, perpendicular to the spin axis and the applied torque; that is, the spin axis tries to align itself with the applied torque. This is the law of gyroscopic precession; measurement of precession is what makes the spinning-mass gyro useful for knowing the changes in direction of the carrying vehicle.

A mass, of inertia I, spinning at rate ω_s, defines an angular momentum vector $\mathbf{H}$ along the axis of rotation or spin axis (**Fig. 1**). When a torque $\mathbf{M}$ is applied about a perpendicular axis, a mass particle at a point A on the rim is forced downward and reaches its maximum downward velocity when it has rotated through 90° to a point A'. After rotation through another 90° to a point A'', the mass particle has moved back up to zero displacement; however, the net force pattern was down. On the other side, a similar particle is experiencing an upward force pattern. Thus, the spinning mass is twisted (precesses) about the third orthogonal axis, called the output axis. The precession rate Ω is given by Eq. (1).

$$\Omega = \frac{M}{H} = \frac{M}{I\omega_s} \tag{1}$$

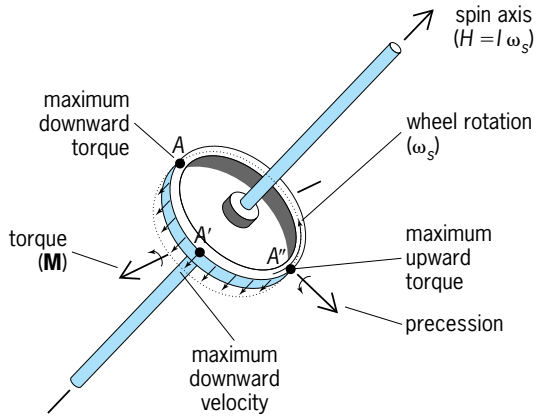

Fig. 1. Gyroscopic precession.

See ANGULAR MOMENTUM; MOMENT OF INERTIA; TORQUE.

A simple experiment, requiring a bicycle wheel with handles on the hub, can be used to demonstrate precession. The wheel is spun up at fairly high speed and supported vertically by a string tied to one of the handles. The wheel will not fall but will remain vertical, precessing slowly around the string because of the torque applied by the wheel's weight. If a freely rotatable base is available, a person standing on the base can cause self-precession by holding the wheel with spin axis horizontal and twisting the handles clockwise or counterclockwise.

Conversely, if a spinning mass experiences an input angular rate Ω about an axis orthogonal to the spin axis, an output torque **M** appears about the third orthogonal axis. For gyroscopic instruments, since vehicle rotations are to be measured, the torque produced by an applied rate is the quantity of interest. *See* PRECESSION; RIGID-BODY DYNAMICS.

Spinning-mass gyroscopes. The free gyroscope's spinning mass is isolated from the rotations of the case or the carrying vehicle so that it remains fixed in space. The relative position of the spinning mass to the gyro case is proportional to the vehicle's angle of rotation. Low-performance versions are used for artificial horizons, turn indicators, and directional (heading) indicators.

Rate gyro. The rate gyro's spinning mass is forced to rotate with the vehicle rotation rate Ω about one particular axis (for example, pitch or roll). The output torque **M** causing the spinning mass to turn (precess) is opposed by an elastic restraint of spring constant K. The angle θ, measured by pickoffs, that the spinning mass turns through is proportional to the vehicle rotation rate Ω, as given in Eq. (2).

$$\Omega = \frac{M}{H} = \frac{K\theta}{H} \tag{2}$$

See AUTOMATIC HORIZON.

Rate gyros are used in applications in which errors greater than 10°/h can be tolerated, such as autopilots and stability-augmentation flight control. For measuring rapid rate changes or when higher accuracy is required, a precision gyroscope may be required. *See* STABILITY AUGMENTATION.

Floated gyro. The floated, single-degree-of-freedom, rate-integrating gyro, or floated gyro as it is commonly known, is basically a rate gyro in which the spring restraint is replaced by viscous damping. The spinning mass is contained inside a "float" which is isolated from the case by means of the viscous flotation fluid, low-restraint electromagnetic suspensions, pickoffs, and torque generators (**Fig. 2**). Under an applied vehicle rotation rate Ω, the spinning mass (float) begins to precess (rotate) at rate given by Eq. (3), where C is the damping coefficient. Integrating this equation with respect to time gives Eq. (4) for the change in the carrying vehicle's angular direction θ_v. The integration is done mechanically inside the gyro via the fluid, hence the term "integrating;" pickoffs (usually electromagnetic) measure θ directly. *See* VISCOSITY.

$$\Omega = \frac{M}{H} = \frac{C\dot{\theta}}{H} \tag{3}$$

$$\theta_v = \frac{C\theta}{H} \tag{4}$$

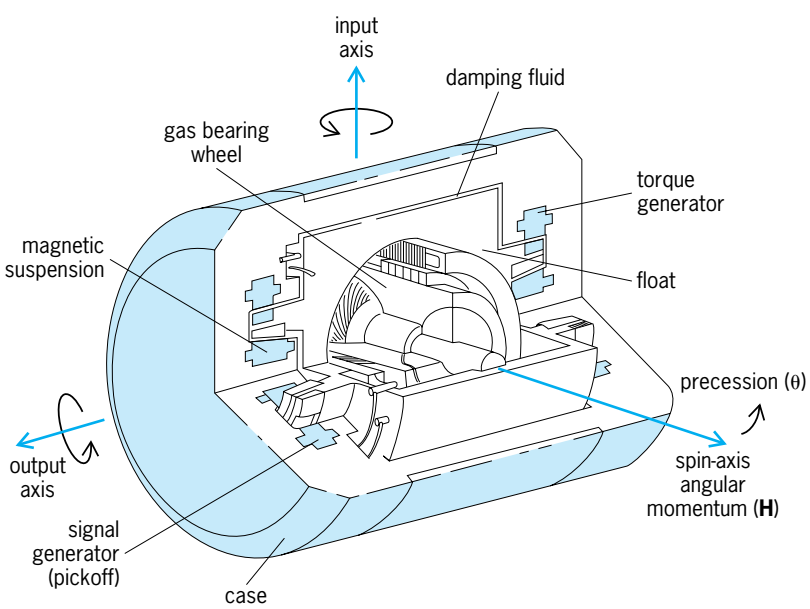

Fig. 2. Single-degree-of-freedom, rate-integrating floated gyroscope. (*After L. Finkelstein and I. V. Grattan, eds., Concise Encyclopedia of Measurement and Instrumentation, Pergamon Press, 1994*)

Floated, integrating gyros went from revolutionizing military aircraft navigation in the 1950s to enabling strategic missile guidance, submarine navigation, space flight (for example, the Apollo spacecraft), and satellite stabilization (for example, the *Hubble Space Telescope*). However, floated gyros have relatively high costs, are labor-intensive with respect to their overall life cycle, and have long warmup times—factors which motivated the development of other spinning-mass and non-spinning-mass gyroscopes.

DTG. The dynamically (or dry-) tuned gyroscope (DTG) was invented in the early 1960s (**Fig. 3**). It is a two-degrees-of-freedom gyro whose spinning mass is suspended by a universal, or Hooke's, joint. By

selecting, or tuning, the spin speed, the inertia of the gimbal portion (typically 15–40 mm or 0.6–1.6 in. in diameter) of the universal joint cancels the elastic restraint of the joint flexures; hence, the rotor is decoupled from the drive shaft, and the spinning mass operates as a free gyro, but with limited angular displacement. By symmetry, the DTG senses two axes of angular rate simultaneously, while avoiding the use of flotation fluids (hence the term "dry") and electromagnetic suspensions. Torquing or platform motions are used to keep the rotor aligned with the case. DTGs have been used for aircraft and missile navigation, north finders, heading and reference systems, seeker heads, and spacecraft stabilization. *See* UNIVERSAL JOINT.

Suspended gyros. The electrostatically suspended gyroscope (ESG) was developed in the 1960s and 1970s. Its spherical rotor (10–40 mm or 0.4–1.6 in. in diameter) is suspended in a spherical chamber in vacuum by electrostatic forces. External motor windings spin the rotor to the desired speed and are then turned off. The rotor will spin for days before requiring motor excitation. Optical or electrical techniques are used to pick off rotor position. The ESG is basically a free gyroscope and does not require a torquer to keep the rotor and case aligned. ESGs are very accurate (0.0001°/h) and are used for long-term navigation of submarines and aircraft and for land surveying. Cryogenic ESGs, with accuracies three orders better than the best floated gyros, have been developed for measurements testing relativity theory in space. A similar concept is the magnetically suspended gyro (MSG), which uses a magnetic field to suspend the rotor. The performance of MSGs is not quite as good as that of ESGs, but they are more rugged. *See* ELECTROSTATICS; MAGNETIC FIELDS; RELATIVITY.

Sagnac effect. Optical gyros use the Sagnac effect to detect rotation. The Sagnac effect pertains to the postulate of the theory of relativity that the speed of light c is constant, and independent of the motion of the source. If two identical light waves (wavelength λ) circulate in opposite directions along a closed path undergoing a rotation Ω, then the light beam traveling in the same direction as the rotation takes longer to travel around the path than the other beam, resulting in a changed interference pattern. Thus, the beams become phase-shifted with respect to each other by an angle ϕ given by Eq. (5), where A is the

$$\phi = \left(\frac{8\pi AN}{\lambda c}\right)\Omega \tag{5}$$

area enclosed by the path, and N is the number of times the light travels around the path. The quantity in parentheses, called the scale factor, relates the gyro's output to its input. It is apparent that the larger the enclosed area A (that is, the bigger the device), the better the performance.

Optical gyroscopes. These include the ring laser gyro (RLG) and fiber-optic gyros (FOGs).

Ring laser gyro. This gyro was invented in the 1960s and is now widely used in tactical and navigation

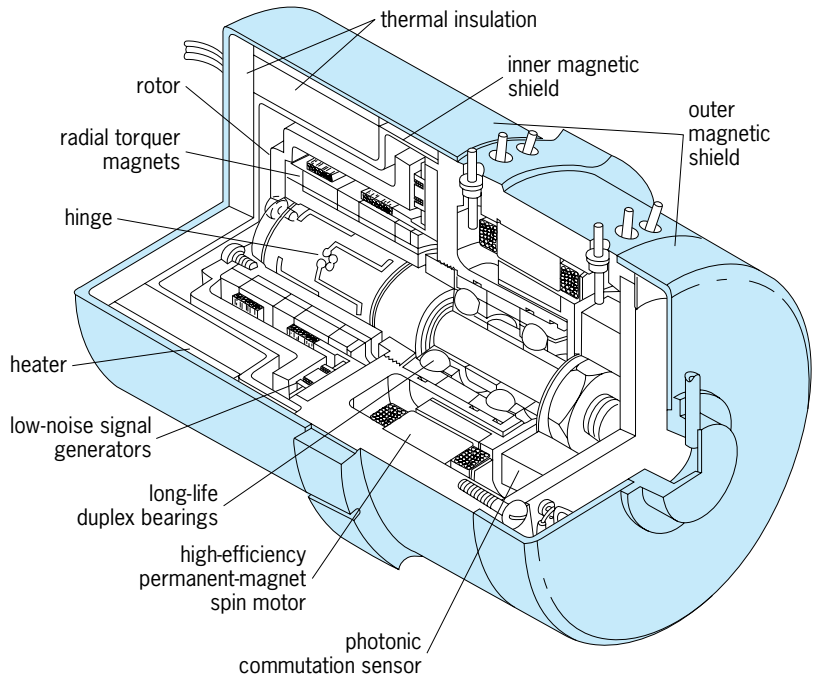

Fig. 3. Dynamically tuned gyroscope.

systems (**Fig. 4**). It comprises a closed optical cavity (usually a three- or four-sided block of low-expansion-coefficient material, such as modified quartz or lithium aluminum silicate), whose light path is defined by mirrors mounted at the corners. The light travels through holes bored in the block containing a low-pressure gas, usually a helium-neon mixture which lases when the anode and cathode are excited. Thus, the RLG is itself actually a laser (that is, it does not require an external light source), and is thus said to be an active device. *See* LASER.

The lased light propagates clockwise and counterclockwise so that there are two optical beams resonating in opposite directions in the cavity (typically with path lengths of 80–400 mm or 3–16 in.). Each optical beam is maintained in resonance (that is, each beam contains an integral number of wavelengths). Under a rotation rate about the gyro input

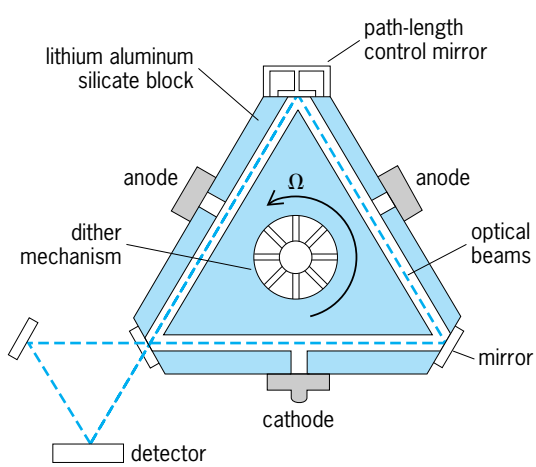

Fig. 4. Ring laser gyroscope. Input axis is perpendicular to plane of page.

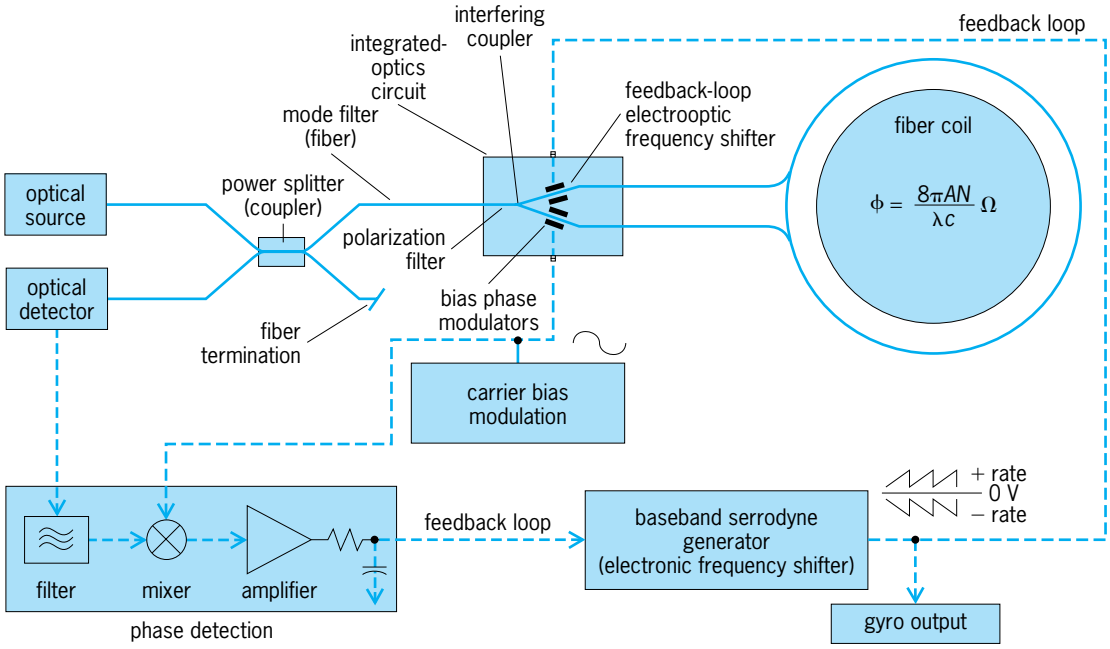

Fig. 5. Interferometric fiber-optic gyroscope. Solid lines show optical path; broken lines show electrical path.

axis, the resonant frequencies of the clockwise and counterclockwise beams change. Some light from both beams is transmitted through one of the mirrors to impinge on a detector. Because the beams have different frequencies, a changing interference pattern (fringes) appears. The detector has two sensitive elements which allow measurement of the fringe pattern changes across the detector in both the positive and negative direction, thereby determining the external rotation rate and direction. Each fringe is equivalent to an angle input, so that counting the number of fringes produces a measurement of the angle. Therefore, the RLG (together with its detector) is an open-loop integrating gyro.

The key element in the RLG is the mirrors, which must be very precise to minimize backscatter between the two beams. Backscatter causes the two light beams to lock frequencies so that low input rates cannot be detected. This phenomenon, known as lock-in, can be avoided by dithering the RLG about the input axis at variable frequencies by a piezoelectrically driven dither motor. Another way to avoid lock-in is to use four light beams in a slightly bent optical path and either Faraday or Zeeman rotators. These rotators alter the propagation constant of the light beams by rotating the polarization, resulting in a phase change, and thus separating the beams to prevent lock-in. The Faraday rotator performs this operation as the light beams travel through a magnetically active crystal, whereas the Zeeman rotator applies a magnetic field bias directly to the gas medium. High-performance RLGs also need a path-length control device integral with one of the mirrors to make the gyro insensitive to changes in the optical cavity length. *See* FARADAY EFFECT; OPTICAL ISOLATOR; ZEEMAN EFFECT.

Since 1980, RLGs have advanced rapidly into many applications, taking over from mechanical instruments. The RLG is an exceptional device for strapdown applications because it has excellent scale-factor stability and linearity, and almost negligible sensitivity to acceleration. It also has digital output, fast turn-on, excellent dormancy, and no moving parts, and it requires much less maintenance than mechanical devices.

Fiber-optic gyros. These gyros use optical fibers, in place of a lasing block, to define the optical path. The light source is external, and its light is split by a beam splitter or optical coupler to produce clockwise and counterclockwise light beams in the fiber-optic coil. FOGs are called passive devices because the optical source (a laser) is external. FOG development began in the 1970s. There are two principal types: interferometric and resonant.

The interferometric fiber-optic gyro (IFOG) [**Fig. 5**] has up to 1 km (0.6 mi) of optical fiber wound into a coil with both ends brought into a coupler. Light from the optical source passes through a power splitter (coupler) and a mode filter (fiber), and into the integrated optics circuit containing an interfering coupler (Y-junction). The interfering coupler splits the light into counterpropagating beams, and then recombines them after they have propagated through the fiber coil. The recombined beam then retraces the path through the mode filter, and is guided into an optical detector via the power splitter, thereby ensuring that both beams have traveled identical paths. The integrated optics circuit is an electrooptic crystal (lithium niobate) comprising optical waveguides fabricated by a proton-exchange process or a titanium-diffused process. It is used for polarization filtering, phase and frequency modulation, and precision interfering (optical phase–to–optical intensity conversion) of the recombined beams. Bias modulation ensures operation of the IFOG in the regime most sensitive to input rate. Backscatter errors are

eliminated by using a wide-band laser source. *See* ELECTROOPTICS; INTEGRATED OPTICS.

The basic open-loop IFOG is not an integrating gyro like the RLG, nor is its output (scale factor) linear with input rate except for a limited range. However, the addition of a feedback loop, including a frequency shifter via serrodyne (sawtooth-wave) phase modulation, can make the IFOG closed-loop and linear with rate. As the input rotation rate changes and is detected at the detector, the feedback loop provides a signal to the frequency shifter which changes the frequency in the opposite direction so that the detector reads at the null position. The sawtooth frequency is proportional to rate. Counting the cycles in the sawtooth provides angle, so that the gyro is now an integrator. Diameters of IFOGs vary from 180 mm (7 in.) for high-performance gyros down to 25 mm (1 in.) for gyros devoted to tactical applications.

The resonant fiber-optic gyro (RFOG) maintains the counterpropagating light beams in resonance, recirculating them in a short fiber-optic coil. Light from an external source enters into the coil by a coupler. In the RFOG, unlike the IFOG, the light continues to travel around the coil while constantly being balanced by the source. The coupler allows light to leave the coil after the beams become phase-shifted under a rotation of the coil. The clockwise beam and the counterclockwise beam are diverted onto separate detectors that measure the frequency shifts; the difference in clockwise and counterclockwise frequencies is proportional to rate. The frequency of the light entering the coil is shifted to maintain both beams in resonance and to provide closed-loop operation. A very narrow band laser is required to resolve coil resonances; therefore, RFOGs are susceptible to backscatter. RFOGs are lower-performing devices but offer size advantages over IFOGs. *See* FIBER-OPTIC SENSOR.

Vibrating gyroscopes. These gyros use an oscillating mass in place of a spinning mass to sense rates. The mass oscillates (sinusoidally) back and forth through a fixed angle; the amplitude of the oscillation is restrained by the elastic (spring) stiffness of the vibrating structure. Nearly all such gyros oscillate at the resonant frequency of the mass-spring system since the gyro output is maximized at this frequency; hence these gyroscopes are also called resonator gyros. *See* RESONANCE (ACOUSTICS AND MECHANICS); VIBRATION.

In the tuning-fork resonator gyro (**Fig. 6***a*), the two tines oscillate from side to side ($\dot{x}$) at a frequency ω, 180° out of phase with each other (that is, as one moves to the right, the other moves to the left). If an angular rate Ω is applied about the base, the Coriolis force causes the tines to move up and down, also at the frequency ω and also 180° out of phase with each other (Fig. 6*b*). The overall motion is thus an ellipse, and the up-and-down component of the vibration ($\dot{x}x$) can be used to sense Ω. *See* TUNING FORK.

There are many configurations of resonator gyros, many of which are made using micromachining techniques developed for the solid-state electronics industry, such as photolithography, masking, and etching, to carve mechanical shapes in quartz and silicon. Quartz devices use piezoelectric effects, via metallized electrodes deposited on the quartz, to drive the oscillation as well as to sense the response. Silicon devices use electrostatic forces to drive the oscillating mass and capacitive plates to sense the response. These devices can be made very small (300 micrometers × 600 micrometers). Micromechanical resonators offer the very inexpensive, high-production capability required for many commercial markets such as toys and automobiles. Because they are small, they also have the ability to withstand incredibly high acceleration forces (over 100,000 *g*), such as those experienced by artillery shells.

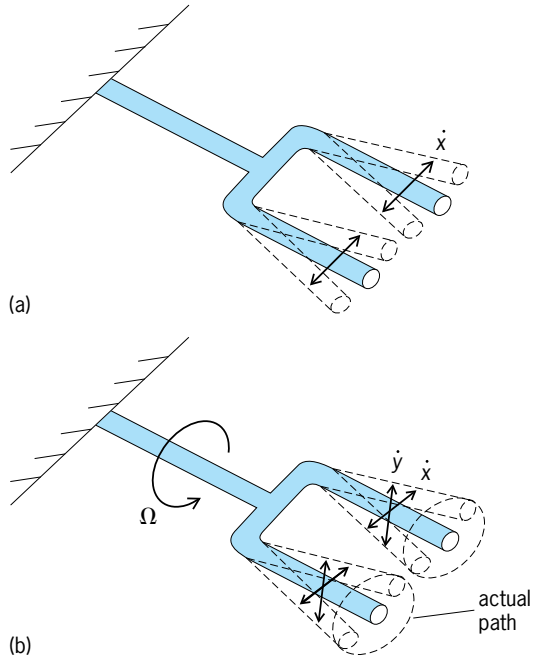

Fig. 6. Coriolis effect on a vibrating (resonator) structure. (*a*) Drive vibration ($\dot{x}$). (*b*) Vibration in presence of rotation rate Ω, which includes output (sense) vibration ($\dot{y}$).

Another resonator gyro offering high performance is the hemispherical resonator gyro (HRG), which consists of a hemispherical quartz shell (20–50 mm or 0.8–2 in. in diameter), electrostatically forced into resonance so that the rim of the shell becomes elliptical. As the gyro is rotated about the stem, the node shape of the vibrating rim changes, and this change is sensed capacitively by sensors adjacent to the rim. This type of gyro has been used in commercial aircraft navigation systems and for satellite attitude control.

Other types. While the most widely used gyroscope techniques have been described, there are several other ways of sensing rate. Gyros that have been developed are the fluidic type and the magnetohydrodynamic angular motion sensor. In the fluidic (or laminar jet) type, a fluid jet is deflected under a rate to impinge on temperature-sensitive elements; the power change due to differential cooling leads to a measure of rate. The magnetohydrodynamic angular motion sensor measures angular acceleration

by magnetohydrodynamic effects in a mercury-filled ring, and the angular acceleration is then integrated to obtain the angular rate. *See* MAGNETOHYDRODYNAMICS.

Gyros under development are the nuclear magnetic resonance (NMR) type, which uses the angular motion of atomic nuclei to respond to rate, and the Josephson junction gyro (JJG), which measures rotation by measuring a phase-induced change in the current in a superconducting ring. Because electrons have a much higher phase sensitivity than photons, the JJG is theoretically capable of being two orders of magnitude smaller than an RLG for equivalent performance. A promising technology which is in its infancy stages is inertial sensing based upon atom interferometry (sometimes known as cold atom sensors). A typical atom de Broglie wavelength is 30,000 times smaller than an optical wavelength, and because atoms have mass and internal structure, atom interferometers are extremely sensitive. Accelerations, rotations, electromagnetic fields, and interactions with other atoms change the atom interferometric fringes. In theory, this means that atom interferometers could make the most accurate gyroscopes, accelerometers, gravity gradiometers, and precision clocks, by orders of magnitude. *See* DE BROGLIE WAVELENGTH; INTERFEROMETRY; JOSEPHSON EFFECT; NUCLEAR MAGNETIC RESONANCE (NMR).

Neil Barbour

Bibliography. A. Lawrence, *Modern Inertial Technology: Navigation, Guidance, and Control*, 2d ed., 1998; H. Lefèvre, *The Fiber-Optic Gyroscope*, 1993; *Optical Fiber Sensors*, vol. 4: *Applications, Analysis, and Future Trends*, ed. by J. Dakin and B. Culshaw, Artech House, 1997; W. Wrigley, W. Hollister, and W. Denhard, *Gyroscopic Theory, Design, and Instrumentation*, 1969.

Gyrotron

One of a family of microwave generators, also called cyclotron resonance masers, in which cyclotron resonance coupling between microwave fields and an electron beam in vacuum is the basis of operation. This type of coupling has the advantage that both the electron beam and the associated microwave structures can have dimensions which are large compared with a wavelength. Thus, cyclotron resonance masers are potentially greatly superior to conventional microwave tubes with respect to power capability at short wavelengths.

The development of these power sources is particularly significant for magnetically confined plasma fusion experiments. Microwave heating is considered an attractive method of supplying the energy needed to bring a reactor to ignition temperature, and gyrotrons provide a potential means of producing sufficient microwave power at the very short wavelength required. Gyrotrons also have potential application in millimeter-wave radar and communications systems. *See* NUCLEAR FUSION.

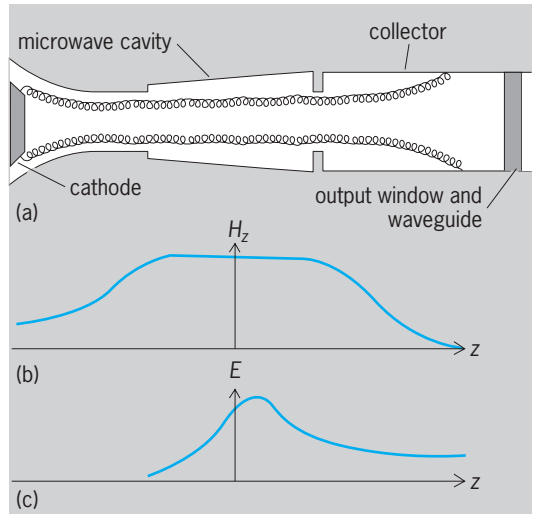

Fig. 1. Gyrotron. (*a*) Schematic diagram, showing elements. (*b*) Plot of typical dc magnetic field H_z as function of distance *z* along axis. (*c*) Plot of representative microwave electric field intensity $|E|$.

Basic characteristics. The basic cyclotron resonance condition is given by Eq. (1), where ω is the

$$\omega = n\omega_c \tag{1}$$

operating frequency, n is an integer, and ω_c is the cyclotron frequency or angular velocity of the electron given by Eq. (2). Here, B is the dc magnetic

$$\omega_c = \frac{eB}{\gamma m_0} \tag{2}$$

field, e is the electron charge, m_o is the rest mass, and γ is the relativistic mass factor. The fundamental cyclotron resonance occurs when $n = 1$. This is the strongest and most useful interaction. The resonance condition requires that very high magnetic fields be used for high-frequency devices. For example, a frequency of 120 GHz requires a magnetic field of about 45 teslas. Generally, the very high frequency gyrotrons have used superconducting magnets.

Larger values of n allow corresponding reductions in the required dc magnetic field. Practical devices have generally been limited to values of n no larger than 2 (second-harmonic operation).

The most important microwave field component in the gyrotron is the electric field tangential to the orbit of the electron. With the fundamental cyclotron resonance interaction, any spatial variation of the microwave fields is of little importance. It is this property that allows the gyrotron to use cross-section areas which are large compared with a wavelength.

Electron bunching in the gyrotron occurs by virtue of the relativistic mass effect included in Eq. (2). The transverse microwave electric field introduces a sinusoidal modulation of γ depending on the angular position of the electron in its orbit relative to the direction of the electric field. The modulation of γ results in a modulation of angular velocity as given by Eq. (2). As the beam drifts, this converts to

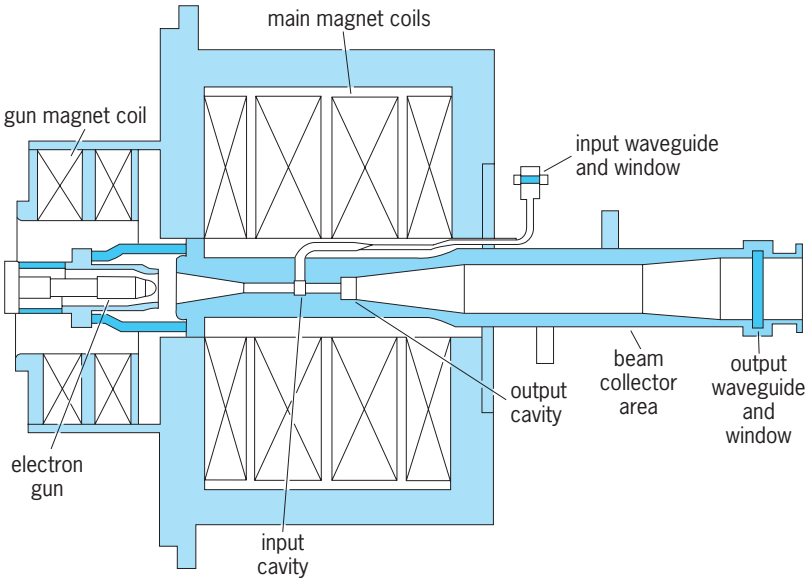

Fig. 2. Simplified cross section of pulsed gyroklystron amplifier.

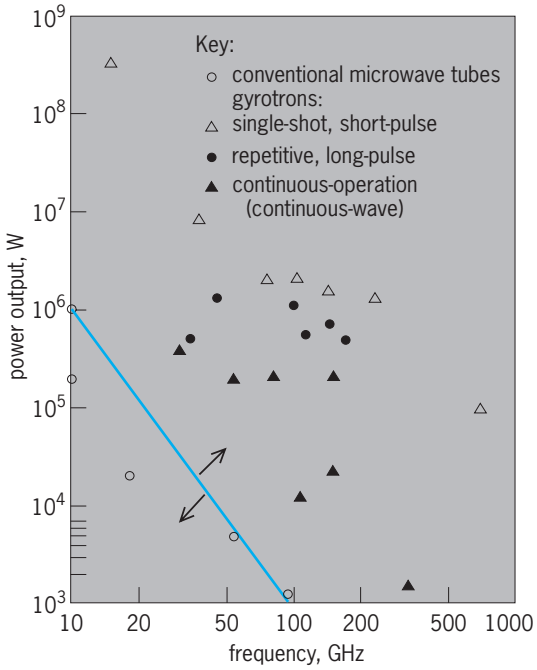

Fig. 3. Power output versus frequency typical of gyrotrons compared with that of conventional microwave tubes. For pulsed devices, peak power output is shown.

angular bunching in the coordinate system centered on each electron orbit. By proper adjustment of phase conditions, the bunched beam can give up most of its energy to microwave energy.

A number of tube configurations are possible using the cyclotron resonance interaction. The simplest form, and that used for most practical gyrotrons to date, is an oscillator using a single resonant cavity (**Fig. 1**). Here, the electron beam is a hollow beam with all electrons having helical motion. For efficient operation, all electrons must have a large fraction of their total energy contained in motion transverse to the device axis. A gyroklystron amplifier employing two or more cavities is another alternative (**Fig. 2**); and in a third variation a traveling-wave circuit, in analogy to a traveling-wave tube, is used.

Capabilities. Gyrotrons can produce higher power at high frequency than conventional klystrons and traveling-wave tubes (**Fig. 3**). Single-shot, short-pulse devices use intense relativistic beams which are not suitable for repetitive pulsing. *See* KLYSTRON; MICROWAVE TUBE; TRAVELING-WAVE TUBE. Howard R. Jory

Bibliography. J. Benford and J. Swegle, *High-Power Microwaves*, 1991; C. Edgecomb (ed.), *Gyrotron Oscillators*, 1993; R. J. Temkin (ed.), *17th International Conference on Infrared and Millimeter Waves*, 1992.

H

Hackberry — Hysteriales (lichenized)

Hackberry

A medium-sized to large tree, *Celtis occidentalis*, occasionally growing to 120 ft (36 m). It occurs in the eastern half of the United States, except the extreme south, and has corky or warty bark; alternate, long-pointed serrate leaves unequal at the base; and a small drupaceous fruit with thin, sweet, edible flesh. The pith of the twigs is chambered. The wood is used for furniture, boxes, and baskets. It is a shade tree and is also used for shelterbelts. Sugarberry (*C. laevigata*) is similar to hackberry. It grows in the southeast United States and has narrower leaves with entire margins and smaller fruit. It is used for furniture, boxes, and baskets; shelterbelts; and shade. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves/Kenneth P. Davis

Hadean

The portion, referred to as an eon, of geological time that extends for several hundred million years from the end of the accretion of the Earth to the formation of the oldest recognized rocks. According to current models, the inner planets formed by the accretion of planetesimals in an environment where gas and volatiles had been swept away by early intense solar activity. The accretion of the Earth appears to have been completed between 50 and 100 million years (m.y.) after the beginning of the solar system (T_0) as recorded in the oldest refractory inclusions in the Allende Meteorite, whose age of 4566 ± 2 m.y., ascertained by lead isotope dating, is taken as T_0. Core formation on the Earth appears to have been coeval with accretion and so preceded the Hadean. Any primitive atmosphere was removed by early collisional events, and the present atmosphere has arisen by a combination of degassing and additions from comets. *See* EARTH, AGE OF; LEAD ISOTOPES (GEOCHEMISTRY).

The Acasta Gneiss in the Northwest Territories of Canada, dated at 3960 m.y., is often regarded as the oldest rock. However, that date refers to relict zircon crystals in the rock rather than the age of formation of the rock itself. The oldest definitely dated rocks are at Isua, Greenland, with an age of 3650–3700 m.y. Thus the Hadean Eon begins around 4500–4450 m.y. ago and extends to between 3960 and 3650 m.y. ago depending on the age assigned to the oldest rock.

Lunar record. There is almost no evidence on Earth of the Hadean Eon, but the Nectarian (3850–?4200 m.y.) and Pre-Nectarian Systems (?4200–?4450 m.y.) on the Moon cover this missing period of time in the terrestrial rock record. The formation of the Moon is conventionally ascribed to a giant collision between an impacting body about 0.20 earth mass with the Earth about 50 m.y. after T_0, predating the Hadean. Part of the rocky mantle of the impactor (but not that of the Earth) was spun out into orbit, melted, and crystallized to form a feldspathic crust. The formation of this lunar highland crust is dated at 4440 ± 20 m.y. and represents the oldest recorded event in the Earth-Moon system. Accretion of the Moon was essentially complete at that time, but the Moon was subjected to a continuing bombardment that lasted until 3850 m.y. ago. Over 40 ringed basins over 300 km in diameter and 10,000 craters with diameters of 30–300 km were formed during this period. A similar but more intense bombardment must have struck the Earth during this period, causing chaotic interruptions to geological processes. A consensus is emerging that this bombardment was not continuous but mostly due to a spike or cataclysm around 3850–4000 m.y. Probably several spikes in the cratering flux were involved. The presence of a population of large impactors for 500 m.y. after the formation of the inner planets may be explained by the collisional disruption of a large body in the asteroid belt that provided a swarm of impacting bodies.

Magma oceans. Apart from the impacts of other major planetesimals during the formation of the Earth, the Moon-forming collision was sufficiently energetic to melt the Earth. However, there is little evidence of the sort of mineralogical or geochemical zoning in the upper mantle that typified the crystallization of the Moon and terrestrial layered intrusions and that might have been expected to result from the crystallization of a terrestrial magma ocean. Since it seems inevitable that the Earth was melted, probably the rapid crystallization of a magma ocean frustrated crystal settling and the formation of a differentiated mantle. Mantle temperatures during the Hadean were probably significantly higher due both to thermal inputs from impacts of large bodies and to a much higher heat flow. Over half the heat produced by the decay of ^{235}U to ^{207}Pb (isotopes of uranium and lead) in the Earth was released during Hadean time alone, adding as much as several hundred degrees to the Earth's internal temperature. *See* EARTH, HEAT FLOW IN; RADIOACTIVE MINERALS.

Early crusts. Observations from Mars, the Moon, Venus, and Vesta suggest that the crust of the Hadean Earth was basaltic. Enigmatic evidence suggests that there was some extraction of the relatively large elements (such as potassium, rubidium, thorium, and uranium) from the mantle at this time, but the scale of this extraction was probably minor as the evidence is restricted to very small areas of outcrop of early Archean rocks. Relict zircon crystals up to 4200 m.y. in age are the only surviving remnants of surface rocks that have been identified. The zircons appear to be derived from felsic igneous rocks. If so, small areas of such rocks presumably formed from remelting of basaltic crust that sank back into the mantle. *See* EARTH CRUST.

It is often thought that, by analogy with the Moon, the Earth formed an early plagioclase-rich (anorthositic) crust. Several reasons make this unlikely. First, the composition of the Moon is richer in calcium and aluminum than that of the terrestrial mantle, leading to the early appearance of plagioclase during crystallization of the lunar magma ocean. Second, plagioclase is unstable at shallow depths (40 km) in the Earth and will transform to garnet, thus locking up calcium and aluminum in a dense phase. In contrast, plagioclase will be stable in the Moon to depths of several hundred kilometers. Plagioclase will sink in a wet terrestrial basaltic magma. Finally, there is no sign of an ancient geochemical reservoir of europium or of primitive $^{87}Sr/^{86}Sr$ (strontium isotopes) that would have resided in an early strontium-rich rubidium-poor anorthostic crust.

Was there an early world-encircling crust of granite on the Hadean Earth? A primitive "sialic" crust is refuted, among other evidence, by the virtual absence of an old zircon population in younger sediments. The significant feature about the Earth, in contrast to the other terrestrial planets, appears to be the presence of liquid water at the surface that enables recycling of subducted basaltic crust through the mantle. It is this process that permits the slow production of the continental crust throughout geologic time. The conditions for the production of massive granitic crusts are probably unique to the Earth and require three or more stages of derivation from a primitive mantle composition. In other planets, the absence of subduction leads to the persistence of barren basaltic plains. Thus no good evidence exists for that enduring geological myth of a primordial crust of "sial" or granite. Such models originated through false analogies between the production of a silicic residuum during crystallization of basaltic magmas and conditions in an early molten Earth, and a lack of appreciation of the difficulties of producing granite. *See* SILICATE MINERALS; ZIRCONIUM.

Different picture of Earth. Conditions on the Hadean Earth bore little resemblance to more recent times. A picture dimly appears of a hot young Earth with a thick basaltic crust, covered by an ocean. Dry land was rare. Plate tectonics had not yet begun. A few remnant zircon crystals indicate the formation of an occasional felsic rock, produced by remelting of the basalt. Sporadic disruption of the surface was caused by the collisions of basin-forming impactors that probably culminated in a spike or cataclysm around 3850–4000 m.y. ago. Such events must have frustrated the origin and development of life, which emerged in post-Hadean time. Stuart Ross Taylor

Bibliography. C. J. Allègre and S. H. Schneider, The evolution of the Earth, *Sci. Amer.*, 271:44–51, 1994; P. Cloud, The Hadean Earth, in *Oasis in Space: Earth History from the Beginning*, chap. 2, Norton, New York, 1988; H. E. Newsom and J. H. Jones, *Origin of the Earth*, Oxford, 1990; P. J. Thomas, C. F. Chyba, and C. P. McKay, *Comets and the Origin of Life*, Springer, 1997.

Hadromerida

An order of sponges of the class Demospongiae with monactinal megascleres that usually have a terminal knob at one end. Microscleres in the form of streptasters, asters, sigmas, or small spined diactinals may occur. Megascleres tend to be arranged in tracts radiating from the center of the sponge; spongin is usually sparse in occurrence.

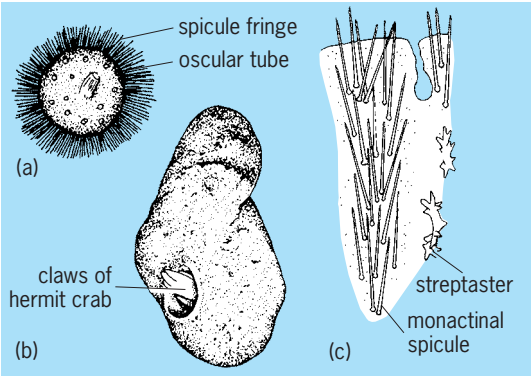

Hadromerine sponges. (*a*) *Radiella sol*, deep-sea species. (*b*) *Suberites ficus* living on shell occupied by hermit crab. (*c*) Spicule arrangement of *Spirastrella*.

In shape, hadromeridan sponges include radially symmetrical forms (see **illus.**) as well as encrusting, massive, or branching types. They occur in tidal and shallow waters of all seas and extend down to depths of at least 18,000 ft (5500 m). Included in the group are the limestone-excavating clionids. *See* BIOERODING SPONGES; DEMOSPONGIAE. Willard D. Hartman

Hadron

The generic name of a class of particles which interact strongly with one another. Examples of hadrons are protons, neutrons, the π, K, and D mesons, and their antiparticles. Protons and neutrons, which are the constituents of ordinary nuclei, are members of a hadronic subclass called baryons, as are strange and charmed baryons, for example, Λ^0_s and Λ^0_c. Baryons have half-integral spin, obey Fermi-Dirac statistics, and are known as fermions. Mesons, the other subclass of hadrons, have zero or integral spin, obey Bose-Einstein statistics, and are known as bosons. The electric charges of baryons and mesons are either zero or ± 1 times the charge on the electron. Masses of the known mesons and baryons cover a wide range, extending from the pi meson, with a mass approximately one-seventh that of the proton, to values of the order of 10 times the proton mass. The spectrum of meson and baryon masses is not understood. *See* BARYON; BOSE-EINSTEIN STATISTICS; FERMI-DIRAC STATISTICS; MESON; NEUTRON; PROTON.

It is believed that the net number of baryons in the universe is a conserved quantity. In making this count, baryons and antibaryons are arbitrarily assigned baryon numbers $+1$ and -1, respectively, so that production or annihilation of a baryon-antibaryon pair in a given reaction has no effect on the net baryon number. (Mesons are assigned baryon number zero.) In this scheme, the least massive baryon, the proton, is stable, and all other baryons unstable. However, it has been suggested, as a result of some theoretical attempts to unify the strong, weak, and electromagnetic interactions, that protons may be unstable with a lifetime many orders of magnitude greater than the age of the universe. *See* FUNDAMENTAL INTERACTIONS; GRAND UNIFICATION THEORIES; SYMMETRY LAWS (PHYSICS).

In addition to baryon number, hadrons and their antiparticles are assigned quantum numbers to represent other properties of hadronic matter. Among these are *i*-spin (related to electric charge), strangeness, charm, and two others. Strong interactions conserve *i*-spin, and hence hadrons may form *i*-spin multiplets whose members differ in mass by only a small fraction of the proton mass, for example, the proton-neutron doublet. Hadrons with nonzero values of the strangeness and charm quantum numbers, as well as hadrons without a decay mode that conserves *i*-spin, are quasistable; their decays take place either through the weak interaction, which does not conserve strangeness or charm, or through the electromagnetic interaction, which does not conserve *i*-spin. *See* CHARM; I-SPIN; STRANGE PARTICLES; WEAK NUCLEAR INTERACTIONS.

Based on an enormous body of data, hadrons, are now thought of as consisting of elementary, fermion constituents known as quarks which have electric charges of $+{}^2/_3|e|$ and ${}^1/_3|e|$, where $|e|$ is the absolute value of the electron charge. A quark-antiquark pair makes up a meson, while three quarks constitute a baryon. To satisfy a well-tested principle concerning the assignment of quantum numbers in a closed fermion system, the quarks are assigned a quantum number (color) with three possible values, in addition to the quantum numbers signifying other quark properties such as strangeness and charm. Massless, spin-1 gluons are considered to be the propagators of the force between quarks. A quantitative quantum theory based on these ideas, called quantum chromodynamics (QCD), has been developed. *See* ELEMENTARY PARTICLE; GLUONS; QUANTUM CHROMODYNAMICS; QUARKS. A. K. Mann

Bibliography. L. B. Okun, *Particle Physics: The Quest for the Substance of Substance*, 1985; D. H. Perkins, *Introduction to High Energy Physics*, 4th ed., 2000; C. Quigg, Elementary particles and forces, *Sci. Amer.*, 252(4):84–95, April 1985.

Hadronic atom

A hydrogenlike system that consists of a strongly interacting particle (hadron) bound in the Coulomb field and in orbit around any ordinary nucleus. The kinds of hadronic atoms that have been made and the years in which they were first identified include pionic (1952), kaonic (1966), Σ^- hyperonic (1968), and antiprotonic (1970). They were made by stopping beams of negatively charged hadrons in suitable targets of various elements, for example, potassium, zinc, or lead. The lifetime of these atoms is of the order of 10^{-12} s, but this is long enough to identify them and study their characteristics by means of their x-ray spectra. They are available for study only in the beams of particle accelerators. Pionic atoms can be made by synchrocyclotrons and linear accelerators in the 500-MeV range. The others can be generated only at accelerators where the energies are greater than about 6 GeV. *See* ELEMENTARY PARTICLE; HADRON; PARTICLE ACCELERATOR.

The hadronic atoms are smaller in size than their electronic counterparts by the ratio of electron to hadron mass. For example, in pionic calcium, atomic number $Z = 20$, the Bohr radius of the ground state is about 10 fermis (1 fermi $= 10^{-15}$ m), and in ordinary calcium it is about 2500 fermis. Thus the atomic electrons are practically not involved in the hadronic atoms, and the equations of the hydrogen atom are applicable. The close approach of the hadrons to their host nuclei suggests that hadron-nucleon and hadron-nucleus forces will be in evidence, and this is one of the motivations for studying these relatively new types of atoms.

X-ray emissions. Negative hadrons are captured into orbits of large principal quantum number n by

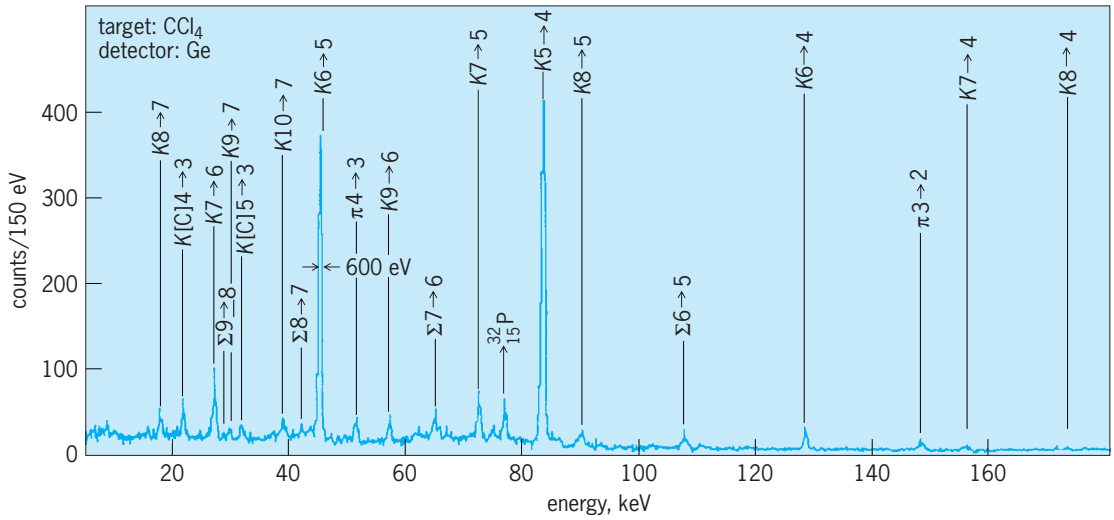

X-ray spectrum resulting from kaons stopped in carbon tetrachloride. Lines from Σ^- hyperonic atoms are seen along the kaonic x-rays. The pionic lines came from decay products. Nuclear gamma rays of the first excited state of phosphorus-32 are seen at 78 keV. (*Lawrence Berkeley Laboratory*)

the attraction of the positively charged nuclei. As the hadrons fall through successively smaller Bohr orbits (cascade), electrons are ejected from the cloud of atomic electrons (Auger effect). When a hadron has reached about the same radius as that of the electronic atom ground state, x-ray emission becomes the dominant method for the system to shed its excitation energy. X-rays, whose energy increases with each successive jump, are emitted until the hadron reaches the ground state ($n = 1$) of the hadronic atom or is absorbed by the nucleus in a strong interaction. *See* ATOMIC STRUCTURE AND SPECTRA; AUGER EFFECT; X-RAYS.

The precise measurement of the energies and shapes of the x-ray transitions not perturbed by the strong interaction is an excellent method to determine masses and magnetic moments of the hadrons. This is the case for the masses of negative pions, kaons, antiprotons, and Σ^-, as well as the magnetic moments of the last two.

The lines of the spectra of special interest are due to transitions between the lowest quantum levels because the hadrons are then closest to the nuclei and the nuclear forces perturb the orbits. The effects expected and observed are that some of the lines are slightly broadened (energy indefinite) and the average transition energy is different from that predicted solely on the basis of Coulomb effects.

The series of x-ray lines generally cuts off rather abruptly at $n > 1$. However, in light pionic atoms the ground state is reached. Kaons, Σ^-, and antiprotons ($\bar{p}$) can probably reach the ground state only when the nucleus is singly charged. In kaonic chlorine atoms, for example, the series ends at $n = 3$, and in kaonic lead the series ends at $n = 7$.

Experimentally, the hadrons are generated by a beam of protons incident on a metallic target. A secondary beam is used to transport the particles to the target in which the atoms are to be made and studied. The arrival of a hadron is signaled by a set of scintillation counters as it is slowed down by passage through a moderator of carbon or beryllium. The thickness of the moderator is adjusted so that a maximum number of hadrons stops in the target under investigation. The x-ray detectors are semiconductors of silicon or germanium. Efficiencies for detecting an x-ray that comes from within the target are around 5×10^{-3}. The energy resolution of the detectors is of paramount importance. The **illustration** is an example of a kaonic x-ray spectrum of chlorine obtained by an ultrapure germanium detector. The lines are labeled according to their hadronic transitions. The intensity of the lines average about 0.3 x-ray per stopped kaon for the principal lines ($\Delta n = -1$). About 5×10^6 kaons were stopped to obtain the spectrum. In special cases where the resolution is of vital importance, as was the case in determination of the pionic mass, a crystal spectrometer is used. *See* PARTICLE DETECTOR.

Pionic atoms. The interpretation of the x-ray spectra of pionic atoms is complicated by the necessity for the pion to react with two nucleons rather than one nucleon alone, as in reaction (1), where N stands

$$\pi + N + N \rightarrow N + N + \text{kinetic energy} \qquad (1)$$

for either a proton or a neutron. The two-nucleon final state is required for the reaction to conserve momentum.

Through the use of the line width and energy shift data of all the measurements on many elements throughout the periodic table, a calculation was made to determine the parameters of the low-energy pion-nucleus interaction. An important ingredient of this calculation is the short-range nucleon-nucleon correlation in nuclei, which determines the probability for reaction (1). The experimental data on the pion absorption in the nuclei can be fairly well accounted for in this phenomenological approach.

Kaonic atoms. Kaonic atoms were expected to yield valuable information concerning the surface of nuclei because a kaon reacts very strongly with

either a single neutron or a single proton, as in reaction (2).

$$K^- + N \rightarrow \pi + \text{hyperon} \qquad (2)$$

On the basis of theory it had been predicted that more neutrons than protons would be found on the surfaces of nuclei. One of the first interpretations of the behavior of the series of x-ray lines of various elements ranging from $Z = 3$ through 92 suggested that the neutron dominance of nuclear surfaces was verified. However, it was later pointed out that a resonance between K^- and proton, $Y^*(0)1405$, would probably enhance the affinity of kaons for protons on the nuclear outskirts, and this could account for the increased capture rate observed experimentally. The data on the kaonic atoms could not be analyzed uniquely in terms of the difference of the proton and neutron distributions at the nuclear surface. The K^- nucleon resonances at the threshold complicate the parametrization of elementary interaction (2) so much that the present quality of the data is not sufficient to determine the elementary interaction on the one hand and the nuclear matter distribution on the other. *See* NUCLEAR STRUCTURE.

Σ^- hyperonic atoms. When kaons are absorbed by nucleons, about 20% of the hyperons produced are Σ^- particles. In the light elements most of the Σ^- hyperons are ejected from the nucleus in which they are generated. Some of them are captured by target nuclei; Σ^- hyperonic atoms are formed and emit characteristic x-rays, just as do the kaonic atoms. Weak x-ray lines due to the hyperonic atoms are found along with the kaonic x-ray lines (see illus.). The hyperonic lines are of special interest because they are actually doublets due to the magnetic moment of the Σ^-. X-ray lines of Σ^- atoms in some heavy elements have been measured, and the Σ^- magnetic moment was determined to be $\mu = -(1.097 \pm 0.044)$ nuclear magnetons with an additional systematic error of ± 0.04 nuclear magneton. Perhaps Σ^- will turn out to be a suitable probe of the nucleus.

Antiproton atoms. Antiproton atoms are the latest in the series of hadronic atoms to be observed. Their x-ray lines are doublets due to the magnetic moment of $\bar{p}$. The splitting has been measured and μ found to be -2.795 ± 0.019 nuclear magnetons, which agrees very well with the proton magnetic moment but has a sign opposite to that expected for antiparticles.

The main research effort involving antiproton atoms has been dedicated to the investigation of the x-ray spectra of the antiprotonic hydrogen. The transitions to the ground state depend directly on the elementary antiproton-proton interaction at the threshold. If this interaction turns out to be simple enough, the antiprotonic atoms will be a future tool for measuring the matter distribution of the nuclear surface. Another source of low-energy antiprotons—the Low Energy Antiproton Ring (LEAR), which makes precision measurements on antiprotonic atoms feasible—was put into operation at CERN near Geneva, Switzerland.

There are two additional hadrons with lifetimes that are long enough to be candidates for hadronic atom formation: the negative xi (Ξ^-) and the negative omega (Ω^-), but even at the largest accelerators, these particles are too scarce for their atoms to be detected. Clyde E. Wiegand; Bogdan Povh

Bibliography. C. J. Batty, Exotic atoms, *Sov. J. Part. Nuc.*, 13:71–96, 1982; J. B. Warren (ed.), *Nuclear and Particle Physics at Intermediate Energies*, 1976.

Haemophilus

A genus of gram-negative, pleomorphic bacteria that are facultative anaerobes and are nonmotile and non-spore-forming. They have fastidious growth requirements, including either factor X (provided by hemin or other porphyrins) or factor V [provided by nicotinamide adenine dinucleotide (NAD), nicotinamide adenine dinucleotide phosphate (NADP), or nicotinamide nucleoside] or both, depending on the species. Erythrocytes represent an exogenous source of these factors.

Classification. *Haemophilus influenzae* was the first of the species to be isolated and is considered the type species. It was originally recovered during the influenza pandemic of 1889 and for a time was believed to be the causative agent of influenza; thus it was called the influenza bacillus. However, when this fallacy became apparent, the organism was renamed, still reflecting the historical association with influenza.

Haemophilus species are distinguished by a number of criteria, including the requirement for factor X or factor V or both (see **table**), the pattern of sugar fermentation, the ability to lyse horse erythrocytes, and the presence of catalase. Strains of *H. influenzae* and *H. parainfluenzae* can be subdivided into biotypes according to tests for the production of indole and the presence of urease and ornithine decarboxylase. Strains of *H. influenzae* can be separated into encapsulated and nonencapsulated forms. Encapsulated strains express one of six biochemically and antigenically distinct capsular polysaccharides that are designated serotypes a through f. Nonencapsulated strains are defined by their failure to agglutinate with antisera against the known capsular types and are referred to as nontypable. *See* INFLUENZA; MENINGITIS.

Human infections. *Haemophilus influenzae* is a human-specific pathogen that inhabits the upper respiratory tract and is acquired by exposure to airborne droplets or contact with respiratory secretions. Nontypable strains can be isolated from the nasopharynx of up to 80% of normal children and adults at any given time, usually in association with asymptomatic colonization. In the setting of compromised mucociliary clearance (due to viral respiratory infection, exposure to cigarette smoke, or underlying cystic fibrosis, chronic bronchitis, or allergic disease), nontypable *H. influenzae* can spread contiguously to produce localized disease in the lungs (bronchitis,

***Haemophilus* species associated with human disease**

Species	Growth factor X	Growth factor V	Primary clinical manifestations
H. influenzae nontypable	+	+	Exacerbation of chronic bronchitis, otitis media, sinusitis
serotype b			Bacteremia, meningitis, epiglottitis, pneumonia, septic arthritis
serotype a, c-f			Meningitis, pneumonia, septicemia
biotype IV			Neonatal septicemia, amnionitis, maternal bacteremia, acute salpingitis, tubo-ovarian abscess
biogroup aegyptius			Purulent conjunctivitis, Brazilian purpuric fever
H. parainfluenzae	−	+	Meningitis, endocarditis, brain abscess, genitourinary tract infection
H. aphrophilus	+*	−	Endocarditis, brain abscess
H. paraphrophilus	−	+	Endocarditis, brain abscess
H. haemolyticus	+	+	Pharyngitis
H. parahaemolyticus	−	+	Endocarditis, pharyngitis
H. segnis	−	+	Endocarditis
H. ducreyi	+	−	Chancroid (genital ulcers)

*Although *H. aphrophilus* was originally reported to require factor X for growth, recent studies suggest that it can grow independent of this factor.

pneumonia), middle ears (otitis media), or sinuses (sinusitis). Overall, these organisms are the leading cause of exacerbations of chronic bronchitis, and the second most common etiology of acute otitis media and sinusitis (see table).

On occasion, nontypable *H. influenzae* causes invasive disease such as meningitis, septicemia, endocarditis, epiglottitis, or septic arthritis. Invasive disease occurs most often in neonates and in patients with underlying immunodeficiency, especially when abnormalities in humoral immunity are present. The majority of individuals who develop nontypable *H. influenzae* meningitis have a passage between the respiratory tract and the subarachnoid space. Typical symptoms include fever, headache, and stiff neck. Since the 1970s, nontypable *H. influenzae* has emerged as an important cause of neonatal septicemia, particularly in premature infants. In most infants, symptoms develop within the first few hours of life, with manifestations of respiratory distress dominating the presentation. A large proportion of isolates from these patients are biotype IV, a subgroup of nontypable *H. influenzae* that is uncommon at other sites of infection. In 1984 a fulminant septicemic illness due to *H. influenzae* biogroup aegyptius was recognized in children in São Paolo and was called Brazilian purpuric fever. *Haemophilus influenzae* biogroup aegyptius is a subgroup of *H. influenzae* biotype III that was first described in the late 1800s and has a predilection for causing purulent conjunctivitis. In most patients, Brazilian purpuric fever begins with purulent conjunctivitis, which is followed 3–15 days later by fever, abdominal pain, vomiting, and vascular collapse. Although conjunctivitis due to *H. influenzae* biogroup aegyptius occurs in the United States and other countries, cases of Brazilian purpuric fever have been reported only in Brazil and Australia, apparently because only select clones are endowed with the necessary virulence determinants.

Encapsulated strains of *H. influenzae* are present in the nasopharynx of only 2–5% of children and an even smaller percentage of adults. Historically, *H. influenzae* type b strains were the primary cause of childhood bacterial meningitis and a majority of other bacteremic diseases in children. However, in recent years the incidence of disease due to *H. influenzae* type b has plummeted in the United States and other developed countries, reflecting the routine use of *H. influenzae* conjugate vaccines. The existing *Haemophilus* vaccines contain the type b capsular polysaccharide (a polymer of ribose and ribitol-5-phosphate) conjugated to an immunogenic carrier protein. These vaccines provide effective protection against disease due to *H. influenzae* type b but fail to protect against non-type b strains.

Haemophilus aphrophilus, *H. haemolyticus*, *H. parahaemolyticus*, *H. parainfluenzae* and *H. segnis* are members of the normal flora in the human oral cavity and oropharynx and have low pathogenic potential. *Haemophilus aphrophilus*, *H. haemolyticus*, and *H. segnis* are often present in dental plaque and gingival scrapings. *Haemophilus parainfluenzae* is occasionally isolated from the urethra and vagina. Among these species, *H. parainfluenzae* is the most common pathogen and has been reported in association with a variety of diseases, including meningitis, endocarditis, brain abscess, epidural abscess, epiglottitis, pneumonia, empyema, septicemia, and urinary and genital tract infections. *Haemophilus aphrophilus* and *H. paraphrophilus* are responsible for a similar spectrum of diseases (see table). *Haemophilus haemolyticus* and *H. parahaemolyticus* rarely produce illness in humans.

Haemophilus ducreyi is distantly related taxonomically to other *Haemophilus* species and is the causative agent of chancroid, a sexually transmitted disease characterized by genital ulceration and inguinal adenitis. Transmission occurs during sexual contact when the organism penetrates a break in the epithelium. Following an incubation period of 4–7 days, a papule develops, usually surrounded by a rim of redness. Over the next 2–3 days, the papule evolves to a pustule, which ruptures to form a sharply circumscribed ulcer. Occasionally multiple ulcers coalesce to form giant ulcers. Other diseases

associated with a similar presentation include syphilis, lymphogranuloma venereum, donovanosis, and genital herpes. Chancroid is most common in developing countries in Africa, Asia, and Latin America. Although relatively uncommon in the United States, a number of large outbreaks have been identified since 1981, usually revolving around prostitution and illicit drug use. Evidence suggests that chancroid may be an important cofactor in heterosexual spread of human immunodeficiency virus (HIV). *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); SEXUALLY TRANSMITTED DISEASES.

Therapy. Strains of *H. influenzae* are increasingly resistant to a wide variety of antibiotics, including penicillins, erythromycin, trimethoprim-sulfamethoxazole, and occasionally chloramphenicol. In most cases, penicillin resistance is mediated by a plasmid-encoded beta-lactamase, which has been detected in up to 60% of isolates in some geographic areas. Accordingly, an extended-spectrum cephalosporin is generally recommended for empiric treatment of serious disease. For localized respiratory tract infection, options include a combination of amoxicillin and clavulanate (a beta-lactamase inhibitor), an oral second- or third-generation cephalosporin, trimethoprim-sulfamethoxazole, azithromycin, or clarithromycin. Among isolates of *H. ducreyi*, plasmid-mediated resistance to ampicillin, chloramphenicol, tetracyclines, and sulfonamides is well described. Effective agents for the treatment of chancroid include ceftriaxone, azithromycin, erythromycin, ciprofloxacin, and amoxacillin-clavulanate. Traditionally, other *Haemophilus* species have been treated with ampicillin. However, resistance to ampicillin, usually related to betalactamase production, is recognized with increasing frequency. As a result, a second- or third-generation cephalosporin should be used until antibiotic susceptibility results are available. *See* ANTIBIOTIC; DRUG RESISTANCE.

Veterinary diseases. *Haemophilus agni* is the etiologic agent of acute septicemia in lambs. This illness typically presents as a fulminant process with death in less than 12 hours.

Haemophilus paragallinarum commonly colonizes the nasal passages of birds and is the causative agent of fowl coryza, which manifests as purulent nasal discharge.

Haemophilus parasuis is a normal resident in the oropharynx of swine and is the etiologic agent of Glasser's disease. Infection typically occurs on serosal surfaces in a wide range of locations, including the meninges (brain), pleura (lung), pericardium (heart), peritoneum (abdominal cavity), and synovia (joints). This disease occurs in young swine and presents with fever, respiratory distress, convulsions, and paresis (focal weakness).

Haemophilus parahaemolyticus typically infects young swine via the respiratory route, leading to pneumonia, which is sometimes accompanied by meningitis or arthritis.

Haemophilus somnus can be found in the nasopharynx and female genital tract of healthy cattle and causes infectious thromboembolic meningoencephalitis. This disease typically presents with abrupt onset of fever, weakness, and somnolence. Other neurological signs include ataxia, blindness, and paralysis. Occasionally pneumonia and arthritis are present. Alternatively, *H. somnus* can cause isolated arthritis, endometritis, or pneumonia, especially in young cattle. *See* MEDICAL BACTERIOLOGY.

Graham P. Krasan; Joseph W. St. Geme

Bibliography. S. J. Barenkamp, Other *Haemophilus* species, in R. D. Feigen and J. D. Cherry (eds.), *Textbook of Pediatric Infectious Disease*, 4th ed., W. B. Saunders, Philadelphia, 1998; J. G. Holt et al. (eds.), *Bergey's Manual of Determinative Bacteriology*, 9th ed., Williams & Wilkins, Baltimore, 1994; M. Killian, *Haemophilus*, in A. Balows et al. (eds.), *Manual of Clinical Microbiology*, 5th ed., American Society for Microbiology Press, Washington, DC, 1991; E. R. Moxon, *Haemophilus influenzae*, in G. L. Mandell, J. E. Bennett, and R. Dolin (eds.), *Principles and Practice of Infectious Diseases*, 4th ed., Churchill Livingstone, New York, 1995; J. I. Ward and K. M. Zangwill, *Haemophilus influenzae*, in R. D. Feigen and J. D. Cherry (eds.), *Textbook of Pediatric Infectious Disease*, 4th ed., W. B. Saunders, Philadelphia, 1998.

Haemosporina

A relatively small and generally rather compact group of protozoa in the subphylum Sporozoa. Authorities differ as to the group's taxonomic status; that assigned it by the Committee on Taxonomy and Taxonomic Problems of the Society of Protozoologists is followed here: a suborder of the order Eucoccida, subclass Coccidia, class Telosporea, subphylum Sporozoa. The Haemosporina are common protozoan parasites of vertebrates, and some of them are important as causes of illness and death. The best known of the group are the four species of malarial parasites of humans. Not so well known are at least 60 other species of malarial parasites, with a wide host distribution among terrestrial vertebrates, as well as numerous species of *Haemoproteus* and *Leucocytozoon*, and some species of *Hepatocystis*. All three genera are closely related to the genus *Plasmodium*, in which all the true malarial parasites are placed (see **table**). *See* MALARIA.

Transmission of these parasites is probably always effected in nature by the bite of some blood-sucking invertebrate. In the vertebrate host they reproduce asexually; sometimes this occurs in the tissues of certain internal organs, such as the lungs, liver, spleen, and brain; sometimes in the red blood cells (erythrocytes), or even in other types of blood and blood-forming cells; and often in both tissues and blood cells. The immature sex cells, gametocytes, always occur in erythrocytes or leukocytes (white blood cells). Gametocytes mature into gametes after ingestion by an intermediate host (arthropod). Fertilization ensues, with a subsequent period of development culminating in the production of numerous

Some of the better-known species of malarial parasites

Plasmodium species	Host	Vector
Human		
P. falciparum	Humans	*Anopheles* sp.
P. malariae	Humans and anthropoid apes	*Anopheles* sp.
P. ovale	Humans	*Anopheles* sp.
P. vivax	Humans	*Anopheles* sp.
Simian		
P. brasilianum	South American monkeys of various species	?
P. cynomolgi (includes several subspecies)	Monkeys, especially macaques (occasionally humans)	*Anopheles* sp.
P. inui	Monkeys, especially macaques; very widely distributed	*A. hackeri* and *A. leucosphyrus*
P. knowlesi	Monkeys, especially *Macaca irus* (occasionally humans)	*A. hackeri* and probably others
P. reichenowi	Anthropoid apes	?
Rodent		*A. dureni*
P. berghel	Congo tree rat; laboratory rats and mice	
Avian		
P. cathemerium	Numerous avian species (mostly passerines)	*Culex* sp. (possibly some anophelines)
P. circumflexum	Numerous avian species (mostly passerines)	*Culex* sp. and *Theobaldia* sp.
P. elongatum	Numerous species of passerines and others	*Culex* sp.
P. gallinaceum	Natural hosts Asiatic wildfowl; also occurs in domestic fowl	*Aedes* sp.
P. hexamerium	Numerous species of passerines	?
P. juxtanucleare	Domestic fowl and partridges (found originally in Brazil, but probably from Far East)	*C. pallens* and *C. sitiens*
P. nucleophilum	Catbirds and others	?
P. pinottii	Toucan of Brazil (various laboratory hosts, especially pigeon)	?
P. polare	Cliff swallow in North America; in other birds elsewhere	?
P. relictum	Numerous species of passerine and other birds	*Culex* sp., *Aedes* sp., *Theobaldia* sp., and probably even some anophelines
P. rouxi	English sparrows in Near East; elsewhere in other species (occurrence questionable in New World)	*C. pipiens*
P. vaughani	Mainly in robin in North America; in a number of other passerines elsewhere	?
Reptilian		
P. mexicanum	Lizards of a number of species	?

sporozoites. These tiny filamentous forms are infective for the vertebrate host. Since they can develop no further in the arthropod, infection in this host is self-limited, in the sense that no further buildup is possible; the insect is seldom harmed by the parasite. However, sporozoites may remain infective for a long time in the invertebrate host, perhaps as long as the insect lives.

Taxonomy. Protozoologists disagree about the number of families that should comprise the suborder Haemosporina. Some include two families, the Haemoproteidae and Plasmodiidae; others combine these two as a single family. Still others recognize another family, the Babesiidae, as belonging in the suborder. However, the Babesiidae, containing parasites of the red blood cells and usually referred to as babesias or piroplasms, differ in certain important respects, one of them being that transmission is by ticks and another that, although they inhabit red blood cells, they produce no pigment (hemozoin) from the homoglobin.

Subgenera. With the steady increase in the number of species of *Plasmodium* regarded as valid, proposals have been made to place all species infecting birds into four subgenera: *Haemamoeba,* to include parasites that produce round gametocytes; *Huffia,* to include those that produce elongate gametocytes, if all types of blood and blood-forming cells are parasitized; *Giovannolaia*, to include only those that produce relatively large erythrocytic schizonts; and *Novyella*, to include those that produce small erythrocytic schizonts. Similarly, it has been proposed to put mammalian species of *Plasmodium* into three subgenera: *Plasmodium, Laverania*, and *Vinckeia*. Differentiaton would be based on a variety of characters, but chiefly on shape of the gametocytes (crescentic for *Laverania* and spherical for *Plasmodium* and *Vinckeia*), and exoerythrocytic stages (primary and secondary for *Plasmodium*, primary only in the other two). Species infecting primates other than lemurs would fall into the first two of these three genera. Reptilian species of *Plasmodium* are less easily grouped, but three subgenera, *Sauramoeba*, *Carinamoeba*, and *Ophidiella*, have been created. Those species with larger schizonts go into the first, and species with smaller schizonts into the second; and the third genus is reserved for the single species known from snakes.

Plasmodiidae. A family of Haemosporina inhabiting the erythrocytes of the vertebrate host (and, in a few cases, other blood and blood-forming cells), in which both asexual reproduction and the formation of gametocytes occur. Hemoglobin metabolism results in the formation of a by-product, hemozoin, typical of these parasites. The invertebrate host is generally thought always to be a

mosquito, but for many species it is still unknown.

The Plasmodiidae are the true malarial parasites. Although there can be many types of host, the majority of known species are found in birds, mammals, and reptiles (especially lizards). Though, according to P.C.C. Garnham, the avian malarial parasites are "probably of the greatest antiquity," those of reptiles may be even older. Many of the reptilian and avian species of *Plasmodium* parallel one another very closely; this similarity in itself suggests that parasitism in this group of Haemosporina may have occurred in birds and reptiles much earlier than in mammals. Although primates seem to be the preferred hosts among the mammals, malaria also occurs in a number of other mammals, such as rats, squirrels, bats, antelope, and water buffalo. Some mammalian blood parasites formerly regarded as species of *Plasmodium* have been transferred to the genus *Hepatocystis* because they reproduce only in the tissues (in this case, the liver), as does *Haemoproteus*. It is thought that the malarial parasites of mammals may have evolved from *Hepatocystis;* the host distribution of both is much the same, though the latter is more restricted.

Haemoproteidae. Only the gametocytes of this family of Haemosporina occur in blood cells, the asexual stages being confined to the internal organs. If occurring in the erythrocytes, the parasites produce hemozoin as in malaria, except in the case of *Leucocytozoon*. This parasite was long thought to invade only leucocytes (usually assumed to be lymphocytes), in which it caused changes such as to make the type of host cell unrecognizable, but these cells have been shown to be erythrocytes, at least in some species. This finding is additional evidence of the close relationship of the Haemoproteidae to the true malarial parasites. The invertebrate host, where known, is a blood-sucking fly, such as *Lynchia* or *Simulium*, or a midge of the genus *Culicoides*. *Lynchia* may transmit *Haemoproteus columbae* of the pigeon, *Simulium* (blackfly) may transmit *Leucocytozoon* of ducks, and *Culicoides* may transmit both *Hepatocystis* and certain species of *Haemoproteus*.

Included in the family Haemoproteidae are the three genera *Haemoproteus, Leucocytozoon,* and *Hepatocystis*. It is believed (though without much proof) that parasites of the first two genera are highly host-specific. Whether this is also true of *Hepatocystis* is unknown. If anything like strict host specificity exists, there are indeed many species in all three genera. *Haemoproteus* and *Leucocytozoon* occur very widely in birds throughout the world, but for some reason *Hepatocystis* is limited to the Old World, and is found most commonly in rodents, bats, and monkeys. *Haemoproteus* may sometimes occur in reptiles.

Vector-host, parasite relations. The vectors for many species of *Haemoproteus, Leucocytozoon,* and *Hepatocysti* remain undiscovered. It is believed that species with mammalian hosts are usually transmitted by mosquitoes of the genus *Anopheles*, and those of birds by *Aedes* or *Culex*. This aspect of the life cycle of the reptilian malarias is almost wholly unknown, though development to the oocyst stage has been demonstrated in three species of *Culex* which had fed on lizards infected with *Plasmodium floridense*. Since well over 2700 species of mosquitoes are known, with almost equal variety in host preferences and biting habits, there is much room for research. The vector for only one species of *Hepatocystis*, *H. kochi* of African monkeys, has been discovered; it is a midge, *Culicoides adersi*.

It is noteworthy that many individuals are often malaria-resistant, even when bitten by a mosquito species known to be a very effective malaria vector. Some degree of racial immunity also exists. Blacks are less susceptible to vivax malaria than whites, and often more resistant to the consequences of falciparum malaria. In the latter case it is usually because they are heterozygotes for the red blood cell defect known as sickling; homozygotes are likely to be so handicapped that they fail to reach maturity.

A somewhat similar situation exists among mosquitoes, at least in the species *Culex pipiens*. C. Huff showed many years ago that resistance to the avian malarial parasite (*Plasmodium cathemerium*) behaved as a simple mendelian recessive.

Pathogenicity. The Haemosporina vary greatly in pathogenicity, as with other large groups of parasites, and while they are not on the whole very harmful to their hosts, there are some notable exceptions. Human malaria, though on the wane in many parts of the world, was and is a very important disease. Though there are no statistics for the number of cases that occur annually, it is likely that there are still 200,000,000; the majority occur in Africa and southeastern Asia.

With the discovery that several species of simian malaria may be transmitted to humans by mosquitoes (especially of the so-called leucosphyrus group of anophelines, which have a strong preference for primate blood), these malarias have become potentially important. Fortunately, humans are completely susceptible to only 4 of the 23 species and subspecies of simian *Plasmodium* generally recognized as valid; humans are partially susceptible to at least 4 more.

It is also possible that the numerous species of *Plasmodium* occurring in the lower vertebrates may play a significant part in the maintenance of biological balances, particularly since such infection is often acquired early in the life of the host, when susceptibility is likely to be greatest.

Certain species of *Haemoproteus* and *Leucocytozoon* are known to be quite pathogenic to their avian hosts. Notorious examples are *H. lophortyx* in various species of quail, *L. simondi* (*anatis*) in ducks, and *L. smithii* in turkeys. The incidence of infection with these three species is often high, and the latter two have been the causes of disastrous and fulminating epizootics. Young birds are especially susceptible.

Control of haemosporidian infections always depends primarily on control of the vectors. This usually involves thorough screening and the intelligent use of insecticides: DDT is still one of the most valuable, despite the tendency of insects to develop

resistance to it. (That it is also highly poisonous to humans, pets, and useful insect species must never be forgotten.)

Life cycles. The life cycles of the various Haemosporina seem to be very similar, although many are still incompletely known, and are exemplified by the malaria parasites of humans. Such life cycles also closely resemble those of the coccidia, a group of sporozoans to which the Haemosporina are closely related, and from which they probably evolved.

Sporozoites (the infective forms) introduced by the vector develop into intracellular asexually reproducing stages (exoerythrocytic forms) in the tissues. In malaria, after one or several generations of this stage, some of the young parasites (merozoites) escape into the bloodstream and infect red blood cells. Reproduction continues in these cells (and the exoerythrocytic cycle may also persist, it is believed) sometimes for years. Some parasites go on to further asexual reproduction, while others under unknown stimuli become gametocytes. *Haemoproteus* and *Leucocytozoon* undergo similar life cycles, except that reproduction is limited to the internal organs. So does *Hepatocystis*, though its behavior in *Culicoides* (a vector) presents interesting variations. The sexual phase of the cycle occurs only in the vectors and involves gametogenesis, fertilization, growth of the oocyst, and eventual development of large numbers of sporozoites. Some of these migrate to the salivary glands, where they await an opportunity to reach another vertebrate host.

One of the interesting aspects of the haemosporidian life cycle is its frequent correlation with the physiological activities of the host. The periodicity of asexual reproduction is known to be conditioned by both the genetic constitution of the parasite and diurnal variations in host physiology, at least in some species. Relapses in *Leucocytozoon* infections, and perhaps in malaria, are particularly common in the spring because of seasonal changes in reproductive activity.

Biological aspects. Much of what is known about the basic biology of malaria is the result of research on the malarias of the lower animals. Earlier investigations were done chiefly on malaria-infected birds, especially canaries, and later on ducks and chickens. *Plasmodium relictum* (*praecox*), *P. cathemerium*, *P. lophurae*, and *P. gallinaceum* were the species most used, and *P. fallax* has been added to the list.

Plasmodium berghei, a species occurring naturally in the African tree rat, has been found to be a very convenient tool for research on malaria in rats and mice. Their use has provided much knowledge about the life cycle of the malaria parasites, the mechanism of relapse, and chemotherapy.

Use of the electron microscope has supplied information about the ultrastructure of the malaria parasites. An especially interesting finding was the demonstration of an organelle interpreted as a cytostome in the erythrocytic stages of three avian and two simian species of *Plasmodium*. Though essentially similar in all five species, the organelle was about twice the size in the avian parasites (140–190 micrometers) as in the simian parasites (50–80 μm). By the use of this organelle, minute bits of host cell cytoplasm are said to be ingested. Food vacuoles form around these particles and, as digestion proceeds, malaria pigment appears.

A similar structure has also been observed in exoerythrocytic merozoites, but in them it does not seem to function in the ingestion of food. Its real use, and the uses of other organelles, remains obscure.

Sporozoites also have a cytostome, or micropyle, as it is called by Garnham and associates. They thought it might function as an avenue through which the contents of the parasite escape into the host liver or other cells to initiate the exoerythrocytic cycle. Not the least interesting of Garnham's findings is the resemblance of sporozoite structure to that of *Toxoplasma*. Garnham's studies included three of the species causing human malaria.

One of the most important unsolved problems in malariology is a simple and reliable method of culturing the parasites. Nevertheless, through the use of existing cultural methods, much has been learned about parasite physiology, which is known to be remarkably like that of the host cell. Glucose is the chief fuel of the erythrocytic stages, and both the Krebs and Embden-Meyerhof cycles are concerned with energy production. Little is known about the physiology of exoerythrocytic parasites, or of those in mosquitoes.

Protein requirements of the erythrocytic forms are largely met by the consumption of hemoglobin, though this must be supplemented by methionine indirectly derived from blood plasma. Also needed are *para*-aminobenzoic acid and the vitamin ascorbic acid. *See* PARA-AMINOBENZOIC ACID; ASCORBIC ACID.

The chief metabolic residue is malarial pigment, or hemozoin. Formerly thought to be the same as hematin, it is now believed to be a porphyrin-denatured protein complex. Since hemoglobins of different species (and even sometimes of different individuals) are not identical, it probably differs somewhat in different parasite and host species. *See* CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; FERTILIZATION (ANIMAL); GAMETOGENESIS; REPRODUCTION (ANIMAL); SPOROZOA; TOXOPLASMEA.

Reginald D. Manwell

Bibliography. L. R. Ash et al., *Atlas of Human Parasitology*, 4th ed., 1997; P. C. Beaver and R. C. Jung (eds.), *Animal Agents and Vectors in Human Disease*, 5th ed., 1985; R. Goldsmith, *Tropical Medicine and Medical Parasitology*, 1988; M. Tiru and W. Hennessen (eds.), *Diagnostics and Vaccines for Parasitic Diseases*, 1986.

Hafnium

A metallic element, symbol Hf, atomic number 72, and atomic weight 178.49. There are five naturally occurring isotopes. It is one of the less abundant elements in the Earth's crust. *See* PERIODIC TABLE.

Hafnium is a lustrous, silvery metal that melts at about 2222°C (4032°F). Reported values of the boiling point vary greatly, from about 2500 to about

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

5100°C (4530 to 9200°F). There are virtually no uses of the metal other than in control rods for nuclear reactors.

The chemistry of hafnium is almost identical with that of zirconium. The similarity of hafnium to zirconium is a consequence of the lanthanide contraction, which brings the ionic radii to very nearly identical values. Before (and since) the discovery of hafnium, this element was extracted with zirconium from its ores and passed with zirconium into all derivatives. Since the chemical properties are so similar, there has been no incentive to separate the hafnium except for making nuclear studies and components of nuclear reactors. *See* ZIRCONIUM. Warren B. Blumenthal

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; J. Hala (ed.), *Halides, Oxyhalides and Salts of Halogen Complexes of Titanium, Zirconium, Hafnium, Vanadium, Niobium and Tantalum*, 1989; M. F. Lappert, *Comprehensive Organometallic Chemistry II: Scandium, Yttrium, Lanthanides and Actinides, and Titanium, Zirconium, and Hafnium*, vol. 4, 1995.

Hail

Precipitation composed of lumps of ice formed in strong updrafts in cumulonimbus clouds. Individual lumps are called hailstones. Most hailstones are spherical or oblong, some are conical, and some are bumpy and irregular. By definition, hail must have a diameter of at least 0.2 in. (5 mm), and has been known to reach over 6 in. (150 mm). Thus, the largest stones are grapefruit or softball size, and the smallest are pea size.

Hailstone structure. Very often hailstones are observed to be made of alternating rings of clear and white ice (see **illus.**). These rings indicate the growth processes of the hail. The milky or white portion of the growth occurs when small cloud droplets are collected by the hailstone and freeze almost instantaneously, trapping bubbles of air between the droplets. The clear portion is formed when many droplets are collected so rapidly that a film of water spreads over the stone and freezes gradually, giving time for any trapped air bubbles to escape from the liquid.

Conditions for formation. The most favorable conditions for hail formation occur in the mountainous, high plains regions of the world, such as in the western United States, the Russian Federation, northern India, South Africa, and Argentina. The high mountainous region of Kenya also has a very high incidence of hail.

The growth of hail normally occurs in a two-stage process. First, hailstone embryos form as graupel, or frozen rain particles. Then they grow to hailstone size as the updrafts strengthen. Hailstone embryos can also form in weaker cloud cells feeding into the main storm. Hailstorms generally have relatively high, cool cloud bases and very strong updrafts within the clouds to carry the hail embryos and hailstones into the cooler regions of the cloud, where maximum growth occurs. Both small ice particles (the embryos) and supercooled liquid water (at temperatures below 32°F or 0°C) are needed to form hailstones. A hailstone larger than 0.5–1 in. (1–2.5 cm) will change size very little during its fall to earth; one smaller than 0.4 in. (1 cm) will probably melt completely by the time it reaches the ground. The largest hailstones fall at speeds approaching 88 mi/h (40 $m \cdot s^{-1}$), while the more numerous, smaller stones fall at 22–44 mi/h (10–20 $m \cdot s^{-1}$).

Hailstorm modification. Much damage to property and crops is caused by hailstorms. The damage to crops depends on the time of year; much hail occurs in the late summer in the Dakotas, but earlier in states such as Illinois, where less (but still significant) damage is done. Property damage due to hail can be as great as that done to crops. In May 1995 a hailstorm caused $1 billion worth of damage in the Dallas–Fort Worth area in Texas, an amount comparable

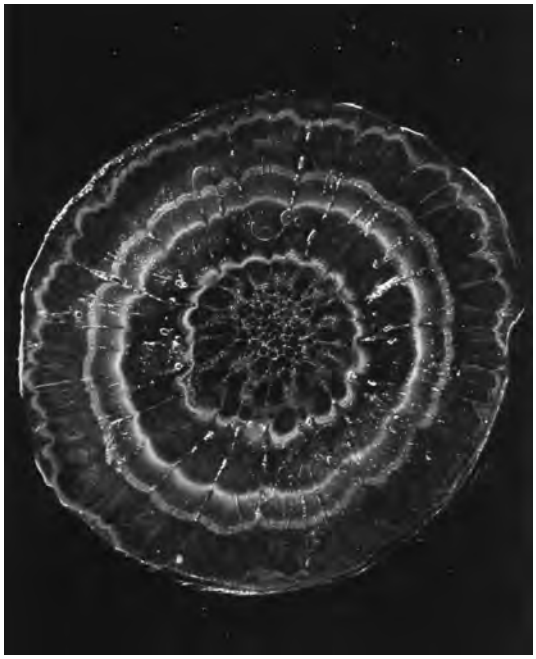

Cross section of a large hailstone showing the structure of alternating rings of clear and white ice. (*Alberta Research Council, Edmonton*).

to the annual crop damage by hail in the entire United States. Consequently, considerable effort has been made to modify hailstorms, with limited success. The experiments and operations use silver iodide to form many small ice particles (more than nature provides) which compete for the available supercooled liquid water to become hailstones and to cause more rapid transformation from supercooled liquid water to ice. The more numerous but smaller hailstones of the modified storm melt as they fall to earth and produce rain. Controversy exists as to whether this is the way the cloud seeding works in every situation, and whether it might cause more hail and less rain in certain cases. The tremendous complexity of a hailstorm makes the experiments to test the theories very expensive.

Evidence from the late 1980s is encouraging. Target and control analyses of hail suppression in southwestern North Dakota, and in southwestern France, where a 30-year-long operational project is in place, have indicated 40–45% decreases in hail in the target areas. A randomized crossover-designed hail suppression project in Greece has indicated similar hail reduction effects. Small rain increases have been detected, but in nonrandomized tests. Experiments in the field and with hailstorm models run on supercomputers continue to reveal hailstorm processes and how they might be changed. *See* CLOUD PHYSICS; PRECIPITATION (METEOROLOGY); WEATHER MODIFICATION.

Harold D. Orville

Bibliography. R. D. Farley, Numerical modeling of hailstorms and hailstone growth, pt. III: Simulation of an Alberta hailstorm—natural and seeded cases, *J. Climatol. Appl. Meteorol.*, 26:789–812, 1987; G. Foote and C. Knight (eds.), *Hail: A Review of Hail Science and Hail Suppression*, American Meteorological Society, December 1977; C. A. Knight and P. Squires (eds.), *Hailstorms of the Central High Plains*, vols. 1 and 2, National Hail Research Experiment, National Center for Atmospheric Research, 1982; P. L. Smith et al., An exploratory analysis of crop hail insurance data for evidence of cloud seeding effects in North Dakota, *J. Appl. Meteorol.*, 36:463–473, 1997.

Hair

Nonliving, specialized epidermal derivatives characteristic only of modern mammals. However, it is now thought that hair was present in at least some therapsid reptiles. It consists of keratinized cells, tightly cemented together, which arise from the matrix at the base of a follicle. A follicle is a tubular epidermal downgrowth that penetrates into the dermis and widens into a bulb (the hair root) at its deep end (**Figs. 1** and **2**). The follicle, together with a lateral outgrowth called the sebaceous gland, forms the pilosebaceous system. Rapid cell production in the matrix, and differentiation in the regions immediately above, produces a hair shaft which protrudes from the follicle mouth at the skin surface. *See* SEBACEOUS GLAND.

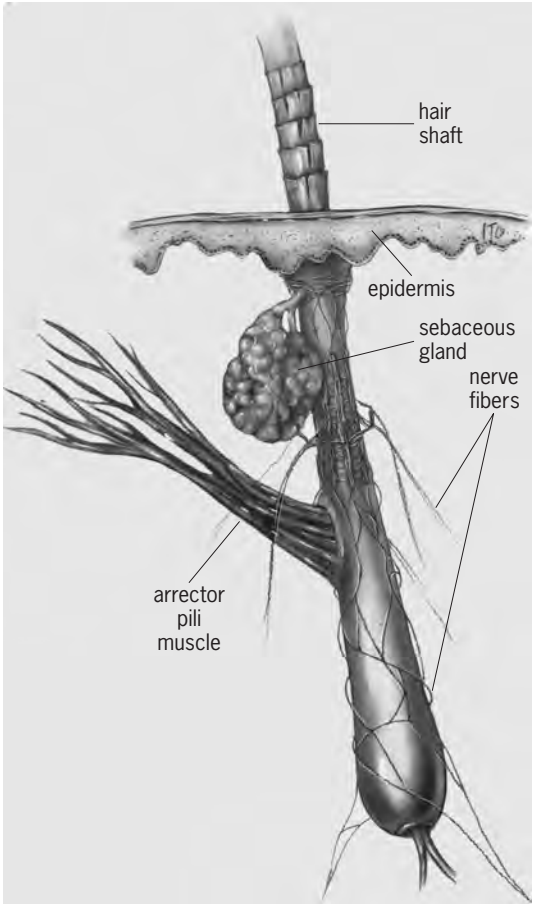

Fig. 1. Human hair follicle and associated structures.

Structure. The shaft consists of an outer cuticle and an innermost medulla, with a cortex in between (**Fig. 3**). The superficial cuticular cells overlap one another, with the free edges directed toward the hair tip. The cortex, which makes up the bulk of each hair, consists of longitudinally arranged spindle-shaped cells which are firmly attached to each other. The color of hair is due to the effect of the pigment in the cortical cells on light reflected from the central medulla. The medulla is less dense than the cortex, and its finer cells are more loosely attached to one another; it may therefore be continuous, discontinuous, or fragmented.

Growth and replacement. Hair follicles do not produce hairs continuously; growing phases alternate with periods of rest. In association with this alternation, the follicle and shaft pass through a complex series of morphological change which is referred to as the hair cycle (**Fig. 4**). During the growth phase, called the anagen phase (Fig. 4*a*), the follicle penetrates deep into the dermis; sections through its deepest portions show complex maturation patterns of cells as the various regions of the shaft are formed. Cytocrine activity of follicle melanocytes deposits pigment in the maturing cells. As the shaft attains its specific length, the follicle shortens (Fig. 4*b*), and eventually, during the quiescent or telogen phase (Fig. 4*c*), the dead "club" hair protrudes from

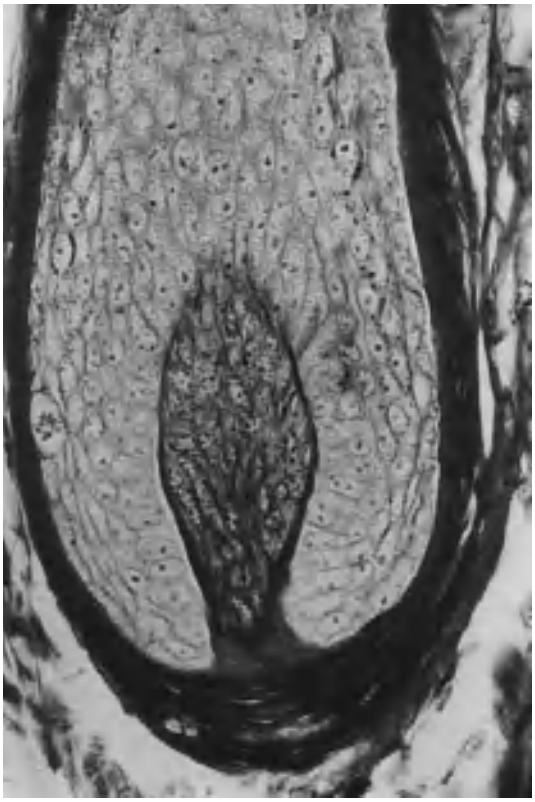

Fig. 2. Photomicrograph of longitudinal section of the lower half of a hair follicle, the root.

a follicle wherein no proliferative or differentiative activity can be seen. The growing period may be as long as 3 years or more, as in the human scalp, or as short as 12 days, as on the belly of a rat. Hairs cannot "turn gray" or "fall out" overnight (traditionally, shock is believed to have such effects), since they are dead structures joined proximally at the follicle base. However, stress may alter both the strength and the color of hairs during the growth phase, probably through a disruption of the endocrine balance during this critical period.

Hairs are not permanent structures but are continually replaced throughout the life of a mammal. In some species, for example, the rat, hamster, mouse, chinchilla, and rabbit, the replacement pattern is undulant, and waves of follicular activity can be traced across the body. In other species, for example, humans, cats, and guinea pigs, each follicle appears to cycle independently of others in the immediate area. Control of a cyclical follicular activity is imperfectly understood, but it certainly involves hormones. However, a systemic stimulant such as a hormone, if acting alone, would affect all follicles over the body, since it is carried by the bloodstream. It would be disastrous for a mammal if all its follicles were to grow in synchrony, since inevitably all functional club hairs would be shed simultaneously, thus depriving the animal of its thermal insulation. It is remarkable that both mammalian hairs and avian feathers, which function as thermal insulators, are developed and shed in various asynchronous patterns, so that the animals are never "naked." Since both hairs and feathers are epidermal appendages controlled and maintained by dermal papillae, it seems probable that part of the secret of the subtle cyclic control mechanisms must lie in these mesenchymal cells. *See* FEATHER.

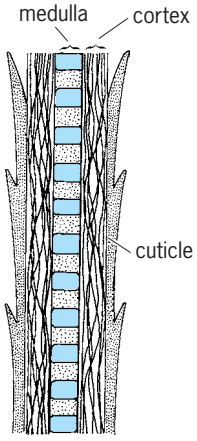

Fig. 3. Longitudinal section of a hair shaft.

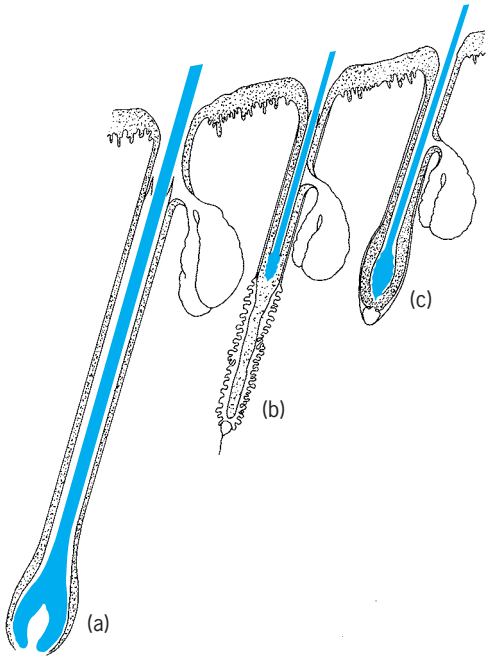

Fig. 4. Stages of the hair growth cycle. (*a*) Growing (anagen). (*b*) Intermediate. (*c*) Quiescent (telogen).

Apart from the sexual or ecological significance of annual pelage changes in many mammals, wild species could not survive if their follicles grew continuously. Merino sheep, whose follicles have a continuous (or, at least, extremely long) growth phase, are prized for the length and texture of their hair shaft, which makes the finest wool. However, strays which are not sheared regularly become entangled in shrubs and bushes and die.

Evolution and adaptive radiation. The major function of hairs in modern mammals is that of thermal insulation. However, individual hairs constitute the pelage, and, strictly, it is the pelage which insulates, since single hairs alone do not have an insulatory

function. A major evolutionary problem has been to determine how such complex individual units, constituting such a complex system, could have been produced by natural selection. Mammalian hairs were probably "protoadapted" to form a pelage; that is, they performed another function before they became units in an insulatory system. This ancestral function was initially the subject of selection; and eventually, as a by-product, the units became sufficiently dense to have an insulatory function which was later improved by selection.

It seems most likely that the primary, ancestral function was that of mechanoreception, and this purpose is suggested by the existence of so-called tactile or sinus hairs in most mammals, also known as vibrissae or whiskers. Such structures, which are characterized by a blood-filled sinus in the dermis surrounding the follicle, which acts as an erectile tissue, are associated with a complex network of nerve fibers. Although all hairs seem to have some degree of sensory function, most of the body is covered by coat hairs which lack the complex, innervated sinus. They may occur as coarse, tough guard hairs, as in pig bristles or porcupine quills, especially, or as short, fine underhairs grouped around the guard hairs. Although guard hairs serve predominantly for physical protection, the dense underhairs fulfill the insulatory function.

Humans appear to lack vibrissal-type hairs, but show a diversity of coat hair at different stages of life. Infants have fine, unpigmented lanugo hairs, which are replaced by coarser terminal hairs in the adult. Vellus hairs are intermediate in texture and in time of appearance. The time sequence of replacement of lanugo hair by terminal hairs varies according to the region of the body. *See* WOOL. P. F. A. Maderson

Bibliography. C. E. Orfanos et al. (eds.), *Hair Research: Status and Future Aspects*, 1981; R. I. Spearman and P. A. Riley (eds.), *The Skin of Vertebrates*, 1981.

Half-life

The time required for one-half of a given material to undergo chemical reactions; also, the average time interval required for one-half of any quantity of identical radioactive atoms to undergo radioactive decay.

Chemical reactions. The concept of the time required for all of the material to react is meaningless, because the reaction goes very slowly when only a small amount of the reacting material is left and theoretically an infinite time would be required. The time for half completion of the reaction is a definite and useful way of describing the rate of a reaction.

The specific rate constant k provides another way of describing the rate of a chemical reaction. This is shown in a first-order reaction, Eq. (1), where c_0 is

$$k = \frac{2.303}{t} \log \frac{c_0}{c} \qquad (1)$$

the initial concentration and c is the concentration at time t. The relation between specific rate constant and period of half-life, $t_{1/2}$, in a first-order reaction is given by Eq. (2). In a first-order reaction, the period

$$t_{1/2} = \frac{2.303}{k} \log \frac{1}{1/2} = \frac{0.693}{k} \qquad (2)$$

of half-life is independent of the initial concentration, but in a second-order reaction it does depend on the initial concentration according to Eq. (3).

$$t_{1/2} = \frac{1}{kc_0} \qquad (3)$$

See CHEMICAL DYNAMICS. Farrington Daniels

Radioactive decay. The activity of a source of any single radioactive substance decreases to one-half in 1 half-period, because the activity is always proportional to the number of radioactive atoms present. For example, the half-period of ^{60}Co (cobalt-60) is $t_{1/2} = 5.3$ years. Then a ^{60}Co source whose initial activity was 100 curies will decrease to 50 curies in 5.3 years. The activity of any radioactive source decreases exponentially with time t, in proportion to $\exp -0.693t/t_{1/2}$. After 1 half-period (when $t = t_{1/2}$) the activity will be reduced by the factor $e^{-0.693} = {}^1/_2$. In 1 additional half-period this activity will be further reduced by the factor $^1/_2$. Thus, the fraction of the initial activity which remains is $^1/_2$ after 1 half-period, $^1/_4$ after 2 half-periods, $^1/_8$ after 3 half-periods, $^1/_6$ after 4 half-periods, and so on.

The half-period is sometimes also called the half-value time or, with less justification, but frequently, the half-life. The half-period is 0.693 times the mean life or average life of a group of identical radioactive atoms. The probability is exactly $^1/_2$ that the actual life-span of one individual radioactive atom will exceed its half-period. *See* RADIOACTIVITY.

Robley D. Evans

Halichondrida

A small order of sponges of the class Demospongiae, subclass Ceractinomorpha, with a skeleton of diactinal or monactinal siliceous megascleres or both. Some megascleres may be arranged in loose tracts, but most are distributed irregularly in the flesh. Spongin is present in small amounts; microscleres are absent. A skinlike dermis is present and is often reinforced with tangentially placed spicules.

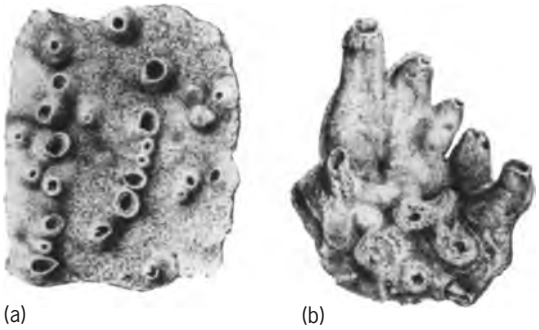

***Halichondria panicea*, shallow-water sponge. (*a*) Encrusting form. (*b*) Fistular form.**

Halichondrid sponges are encrusting, massive, lobate, or branching in shape. Common shallow-water species, such as *Halichondria panicea*, exhibit extensive intraspecific variations in shape associated with environmental conditions (see **illus.**). Halichondrids inhabit all seas, occurring chiefly in tidal areas and shallow waters of the continental shelf. Some species occur down to depths of at least 4900 ft (1500 m). Fossil species are unknown. *See* DEMOSPONGIAE. Willard D. Hartman

Halide

A compound containing one of the halogens [fluorine (F), chlorine (Cl), bromine (Br), iodine (I)] and another element or organic group. Halides have the general formula M_xX_y, where M is a metal or organic group and X is a halogen. Halides are composed of almost every element in the periodic table, and they are referred to as fluorides, chlorides, bromides, or iodides. *See* HALOGEN ELEMENTS; PERIODIC TABLE.

The halides are divided into classes that reflect the nature of bonding between the halogen and metal or organic species. The bonding of halides ranges from purely ionic to essentially covalent. The classes include ionic halides, molecular halides, halides and halogens that behave as ligands in coordination complexes, and organic halides.

Ionic halides such as sodium chloride (NaCl) and potassium chloride (KCl) are prepared from the vigorous reaction of the alkali and alkaline-earth metals with the halogens. These compounds possess high melting and boiling points and are soluble in very polar solvents. Ionic halides are extremely important to the chemical industry, where they are used to produce commodity chemicals such as sodium hydroxide (NaOH), hydrochloric acid (HCl, a hydrogen halide), and potassium nitrate (KNO_3).

The organic halides are divided into the alkyl halides (haloalkanes) and the aryl halides. The alkyl halides have the general formula RX, where R is any alkyl group and X is one of the halogens; for example, 1-chlorobutane ($CH_3CH_2CH_2CH_2Cl$). Halides are good leaving groups in nucleophilic substitution reactions and are good nucleophiles. The aryl halides are compounds where the halogen is attached directly to an aromatic ring and have the general formula ArX, where Ar is an aromatic group. *See* COORDINATION CHEMISTRY; COORDINATION COMPLEXES; ELECTROPHILIC AND NUCLEOPHILIC REAGENTS; HALOGENATED HYDROCARBON; HALOGENATION. Thomas J. Meade

Halimeda

A genus of marine, benthic, green algae (Chlorophyta) belonging to the family Codiaceae. Plants are attached by a holdfast, generally several centimeters high, and consist of calcareous segments separated by flexible, little-calcified nodes. Most species have a distinct, erect habit (see **illus.**), but some deep-water species have a vinelike growth form.

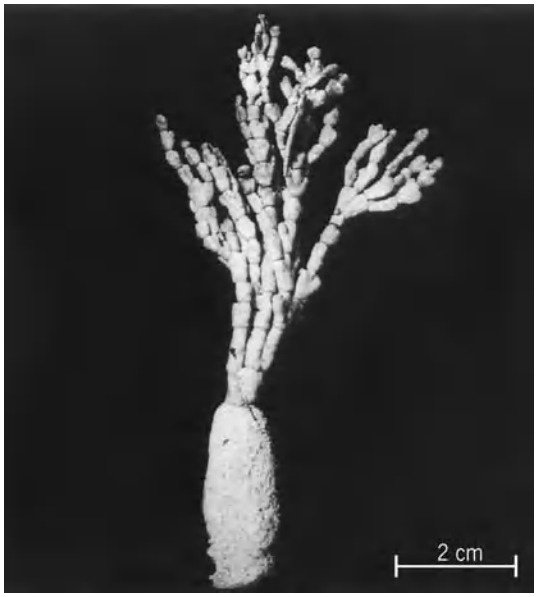

Halimeda **from southern Florida.**

Internally, *Halimeda* segments are made up of longitudinal filaments that develop lateral branches terminating in the surface layer. Calcification, consisting of the mineral aragonite, is variable within segments; also, older segments in a plant are more thoroughly calcified than younger ones.

Halimeda is an exclusively marine alga restricted to tropical waters, except for one or two species known from subtropical regions. These algae colonize sand and mud substrates, where rhizoids of the plant penetrate the soft bottom to develop holdfasts. Light requirements and other factors permit these plants to range in depth from just below low tide to about 330 ft (100 m). They are most common at shallow depths of a few meters, especially in tropical marine shelf and lagoonal environments.

Halimeda has been a prodigious sediment producer throughout its history, which extends from the Mesozoic. The first occurrence of *Halimeda* is not certain because of taxonomic problems in distinguishing morphologically similar genera; however, Middle Jurassic time is probably the earliest appearance of this modern calcareous green alga. *See* ALGAE; ATOLL. John L. Wray

Bibliography. J. L. Wray, *Calcareous Algae*, 1977.

Halite

One of the group of minerals referred to as evaporites, halite is commonly known as salt. Halite is one of many substances that are essential for human life. Evaporite minerals form when ions are concentrated to their saturation point by the progressive evaporation of seawater or saline lake water. Halite precipitates after calcium sulfate, but before the highly soluble salts of potassium and magnesium.

Halite (chemical formula NaCl) is composed of sodium cations and chlorine anions in equal

proportion. It is the most common chloride mineral in natural sequences which proceed beyond the precipitation of sulfates. Even in sequences which contain a high percentage of potassium and magnesium salts, halite is often the most common chloride present.

Crystals of halite are generally cubic or hopper-shaped (skeletal). The mineral is isometric with symmetry 4/m $\overline{3}$ 2/m. Although the mineral is colorless generally, impurities can color it gray, red, orange, or brown. Blue or violet halite results from exposure to radioactivity, which produces dislocations and defects in the crystal structure. Halite is characterized by a hardness of 2.5 on Mohs scale and a specific gravity of 2.16.

The U.S. Geological Survey ranks halite as a basic raw material essential to modern industry along with coal, limestone, iron, and sulfur. Chlorine gas and hydrochloric acid are two of the major products for which halite is a raw material. The agricultural industry uses halite for a variety of purposes, for example, in cattle feed stocks, in fertilizers, and as a weed killer. In many areas with severe winter weather, halite is used regularly as a road deicer. Salt is used for seasoning food and as a preservative both in the home and in the food processing industry. *See* SALT (FOOD).

The deformation of bedded halite deposits is of importance to the petroleum industry. Salt rises, in part as a result of density contrasts, to form dome-like structures. Hydrocarbons (oil and gas) are commonly associated with salt domes. Exploration for these structures by geophysical techniques often results in major discoveries by the petroleum industry. *See* GEOPHYSICAL EXPLORATION; PETROLEUM GEOLOGY; SALT DOME.

Salt resources are found throughout the world and are considered virtually unlimited for any future need. The United States is the world's largest producer, with over 50% of the production coming from Texas and Louisiana and most of the rest from New York, Ohio, and California. Other major producers include China, Russia, Germany, Indonesia, Canada, Mexico, France, Italy, Brazil, Spain, and Romania. *See* HALOGEN MINERALS; SALINE EVAPORITES.

Marc L. Helman; B. Charlotte Schreiber

Bibliography. Bureau of Mines, U.S. Department of the Interior, *Minerals Yearbook*, 1988, vol. 1: *Metals and Minerals*, 1990; Bureau of Mines Staff, *Chemical Industry Applications of Industrial Minerals and Metals*, 1993; W. A. Deer, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, vol. 5: *Non-Silicates*, 1962; C. J. Dixon, *Atlas of Economic Mineral Deposits*, 1979; C. S. Hurlbut, Jr., and C. Klein, *Manual of Mineralogy*, 21st ed., 1998; C. C. Plummer and D. M. McGeary, *Physical Geology*, 8th ed., 1999.

Hall effect

An effect whereby a conductor carrying an electric current perpendicular to an applied magnetic field develops a voltage gradient which is transverse to both the current and the magnetic field. It was discovered by E. H. Hall in 1879. Important information about the nature of the conduction process in semiconductors and metals may be obtained through analysis of this effect.

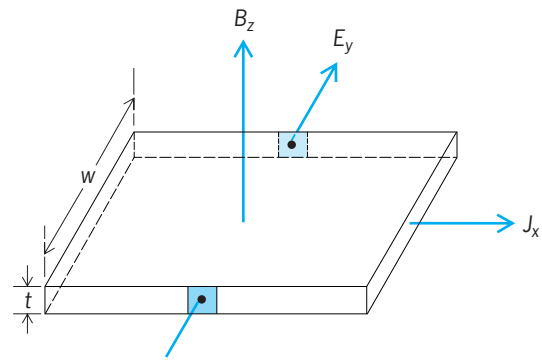

Fig. 1. Configuration of fields and currents in the Hall effect experiment.

Theory. A simple model which accounts for the phenomenon is the following. For a magnetic field of strength B in the z direction (**Fig. 1**), particles flowing with speed v in the x direction suffer a Lorentz force F_L in the y direction given by Eq. (1), where

$$F_L = -qvB \tag{1}$$

q is the charge of the particles. This force deflects the particles so that a charge imbalance develops between opposite sides of the conductor. Deflection continues until the electric field E_y resulting from this charge imbalance produces a force $F_y = qE_y$ which cancels the Lorentz force. In practice, the equilibrium condition $F_L + F_y = 0$ is achieved almost instantaneously, giving a steady-state Hall field as in Eq. (2). The current density is $J_x = nqv$, where n

$$E_y = vB \tag{2}$$

is the carrier density. The Hall resistivity, defined by Eq. (3), is thus given by Eq. (4). For a sample of thick-

$$\rho_{yx} = \frac{E_y}{J_x} \tag{3}$$

$$\rho_{yx} = \frac{B}{nq} \tag{4}$$

ness t and width w, the Hall voltage is $V_H = -E_y w$ and the total current is $I = J_x tw$ so that the Hall resistance, defined by Eq. (5) [which is the experimentally measured quantity], obeys Eq. (6) and is independent of

$$R_H = \frac{V_H}{I} \tag{5}$$

$$R_H = \frac{-B}{nqt} \tag{6}$$

the width but inversely proportional to the sample thickness. The Hall coefficient, defined by Eq. (7),

$$R_0 = \frac{\rho_{yx}}{B} \tag{7}$$

satisfies Eq. (8) and thus R_0 provides a measure of

$$R_0 = \frac{1}{nq} \tag{8}$$

the sign and magnitude of the mobile charge density in a conductor. Within the free-electron theory of simple metals, q is expected to be the electron charge $-e$, and n is taken to be $n = Zn_A$, where Z is the valence of the metal and n_A is the density of the atoms. This yields Eq. (9).

$$R_0 = \frac{-1}{n_A Ze} \tag{9}$$

See FREE-ELECTRON THEORY OF METALS.

Equation (9) is approximately valid in simple monovalent metals but fails drastically for other materials, often even giving the wrong sign. The explanation of the failures of Eq. (9) was one of the great early triumphs of the quantum theory of solids. The theory of band structure shows how collisions with the periodic array of atoms in a crystal can cause the current carriers to be holes which have an effective positive charge which changes the sign of the Hall coefficient. Band structure theory also accounts for the observed dependence of R_0 on the orientation of the current and the magnetic field relative to the crystal axes, an effect which is very useful for studying the topology of the Fermi surface. *See* BAND THEORY OF SOLIDS; FERMI SURFACE; HOLE STATES IN SOLIDS.

Effect in semiconductors. Because of band structure effects, semiconductors generally contain two types of carriers and their number density is highly temperature-dependent. The presence of two types of carriers causes the Hall coefficient to vary with the strength of the magnetic field. This effect can be used to deduce the relaxation times of each carrier species. *See* SEMICONDUCTOR.

Effect in magnetic materials. In a magnetic material the Hall resistivity has the form $\rho_{yx} = R_0B + R_SM$, where M is the magnetization and R_S is the anomalous (or extraordinary) Hall coefficient. The contribution of the anomalous term due to the magnetization may be as large as 100 to 1000 times that of the ordinary term due to the Lorentz force. The origin of this intrinsically quantum-mechanical effect is the asymmetry of the scattering of the electrons from the atoms when there is a net magnetization. The electrons scatter more frequently to the right (say) than the left, producing an extra transverse force distinct from the Lorentz force. For example, a plot of the Hall resistivity of amorphous ferromagnetic cobalt as a function of applied field displays an initial sharp rise, which is due to the increase in the normal component of magnetization of the sample and the resulting anomalous Hall effect. The magnetization reaches its saturation value M_S at an applied field of about 1.8 teslas. A subsequent linear decrease in the Hall resistivity is due to the ordinary Hall effect, which happens to have the opposite sign.

Quantization of Hall resistance. In certain special field-effect transistors, it is possible to create an electron gas which is effectively two-dimensional. The Hall resistance for an idealized system in two dimensions is given by Eq. (10), where n_S is the density of

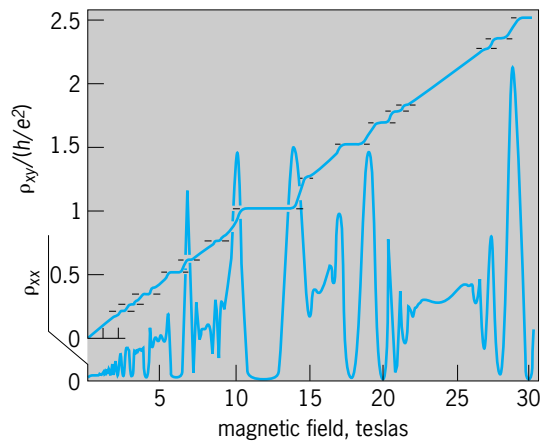

Fig. 2. Hall resistivity (ρ_{xy}, upper trace) and dissipative resistivity (ρ_{xx}, lower trace) as functions of magnetic field in a GaAs-AlGaAs heterostructure at a temperature close to absolute zero (0.085 K). (*After R. Willett et al., Observation of an even-denominator quantum number in the fractional quantum Hall effect, Phys. Rev. Lett., 59:1776–1779, 1987*)

$$\rho_{xy} = -\rho_{yx} = \frac{B}{n_S e} \tag{10}$$

electrons per unit area (rather than volume). However, if the measured value of ρ_{xy} for a high-quality (low-disorder) device is plotted as a function of B (**Fig. 2**), the linear behavior predicted by Eq. (10) is observed only at low fields. At high fields the Hall resistance exhibits plateau regions in which it is a constant independent of B. Furthermore, the values of ρ_{xy} on these plateaus are given quite accurately by the universal relation of Eq. (11), where h is Planck's

$$\rho_{xy} = \frac{h}{e^2 \nu} \tag{11}$$

constant and ν is an integer or simple rational fraction (Fig. 2). The precision and reproducibility of measurements of the quantized values of ρ_{xy} have reached a few parts in 10^8. The absolute accuracy with which Eq. (11) has been verified is better than 1 part in 10^6.

This extremely accurate quantization of ρ_{xy} allows the realization of a new standard of resistance based solely on fundamental constants of nature. In addition, the quantum unit of Hall resistance, $h/e^2 \simeq$ 25,812.80 ohms, determines the fine-structure constant. *See* ELECTRICAL UNITS AND STANDARDS; FUNDAMENTAL CONSTANTS.

The explanation of this remarkable phenomenon involves several subtle quantum-mechanical effects. In the quantum regime (small ν), ρ_{xx}, which is the dissipative (longitudinal) resistivity, approaches zero on the Hall plateaus. The quantization of the Hall resistance is intimately connected with this fact. It is speculated that at zero temperature the dissipation is zero and that Eq. (11) is then obeyed exactly. *See* QUANTUM MECHANICS.

The nearly complete lack of dissipation in the quantum Hall regime is reminiscent of superconductivity. In both effects the ability of the current to flow without dissipation has its origin in the existence of

a quantum-mechanical excitation gap, that is, a minimum threshold energy needed to disturb the special microscopic order in the system. In an ordinary conductor, flow of charge past the (inevitably present) impurities and random crystalline defects induces turbulence and hence dissipation, much as occurs in a rocky stream. Here, however, quantum mechanics causes the flow to be highly ordered, and production of turbulence costs a small but finite amount of energy which is not available to the system at temperatures close to absolute zero. Hence, just as in a superconductor, the current flow is nearly ideal and carries essentially zero entropy. *See* ENTROPY; SUPERCONDUCTIVITY.

In the integer quantum Hall effect [where ν in Eq. (11) in an integer], this excitation gap is a single-particle effect associated with the quantization by the strong magnetic field of the kinetic energy of the individual electrons into discrete states called Landau levels. In the fractional effect, the gap is associated with the highly collective, many-body ordering of the electrons into a quantum state which minimizes the strong Coulomb repulsion and hence lowers the overall energy. Thus, while the integer and fractional quantum Hall effects look superficially similar on a plot of resistivities versus magnetic field (Fig. 2), their physical origins are actually quite different. *See* DE HAAS-VAN ALPHEN EFFECT; GALVANOMAGNETIC EFFECTS.

Steven M. Girvin

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; J. P. Eisenstein and H. L. Störmer, The fractional quantum Hall effect, *Science*, 248:1510–1516, 1990; B. I. Halperin, The quantized Hall effect, *Sci. Amer.*, 254(4):52–60, April 1986; R. E. Prange and S. M. Girvin (eds)., *The Quantum Hall Effect*, 2d ed., 1990; C. T. Van Degrift, M. E. Cage, and S. M. Girvin, Resource letter QHE-1: The integral and fractional quantum Hall effects, *Amer. J. Phys.*, 58:109–123, 1990.

Halley's Comet

The most famous of comets, associated with many important events in history. Records of Halley's Comet appear at least as far back as 240 B.C., and they are found in the Bayeux Tapestry (the apparition of the comet in 1066) and in the *Nuremberg Chronicle* (the apparition of 1456 and probably those of 684 and 1301). The comet's size, activity, and favorably placed orbit, with the perihelion roughly halfway between the Sun and the Earth's orbit, ensure its visibility to the naked eye at each apparition.

Despite major observing campaigns on recent bright comets, specifically comets Hyakutake and Hale-Bopp, Halley's Comet is the basis for many ideas about comets in general. No other comet has had so organized a campaign as that of Halley's in 1985–1986, during which several space missions were launched that made on-site measurements and produced several images of the nucleus. Although detailed knowledge of one comet is invaluable, comets are likely to be highly individualistic, and it should not be assumed that they are all similar to Halley's Comet.

History. This comet was the first to have its return predicted, a feat accomplished by Edmond Halley in 1705. He computed the orbits of several comets with Isaac Newton's new gravitational theory. The orbits of comets observed in 1531, 1607, and 1682 were remarkably similar. Halley assumed that the sightings were of a single comet and predicted its return in 1758–1759. The prediction was verified, and the comet was named in his honor. Halley's Comet was observed in 1835 by F. W. Bessel and by numerous astronomers in 1910 and 1986. Halley's Comet displays the gamut of known cometary phenomena, including a long tail when sufficiently close to the Sun (**Fig. 1**). Because of its predictable orbit, brightness, and extensive activity, it was the prime target of the six spacecraft making up the Halley Armada in March 1986. *See* COMET.

Orbital properties. The comet's orbit is a very elongated ellipse, with an eccentricity of 0.967, which has a perihelion of 0.59 astronomical unit (1 AU = the average Earth–Sun distance = 9.30×10^7 mi = 1.496×10^8 km) and an aphelion of 35 AU, between the orbits of Neptune and Pluto. Halley's Comet was farthest from the Sun in 1948, and has been moving away from the Sun since its perihelion on February 9, 1986. Halley's Comet will return to perihelion in 2061. The average period of revolution is 76 years, and the comet's motion is retrograde, that is, opposite the planets' motion. The present relative positions of the comet's orbit and the Earth's orbit mean that the comet can approach Earth as close as 0.15 AU at the descending node. *See* CELESTIAL MECHANICS.

1986 apparition. Halley's Comet was detected for the first time on its most recent approach to the Sun by D. C. Jewitt and G. E. Danielson of the California Institute of Technology. The comet was recovered on October 16, 1982, by using the 200-in. (5-m) telescope on Palomar Mountain and an advanced electronic detector originally designed for the *Hubble Space Telescope*.

Casual viewers, if they were far enough south and well away from city lights, saw the comet in March and April 1986. However, Halley's Comet was not the spectacular object for public viewing that it was in 1910. The excitement for the 1985–1986 apparition lay primarily in the massive cooperative effort that scientists launched to observe the comet.

In March 1986, six uncrewed spacecraft successfully encountered Halley's Comet on the sunward side and made measurements in its vicinity (see **table**). These missions produced data that have greatly enhanced the understanding of comets. Observations of Halley's Comet were also made by spacecraft in orbit around the Earth and the planet Venus, and by ground-based instruments. The ground-based observations were coordinated by the International Halley Watch, composed of networks of astronomers and institutions worldwide formed to coordinate the total observing effort and to archive results.

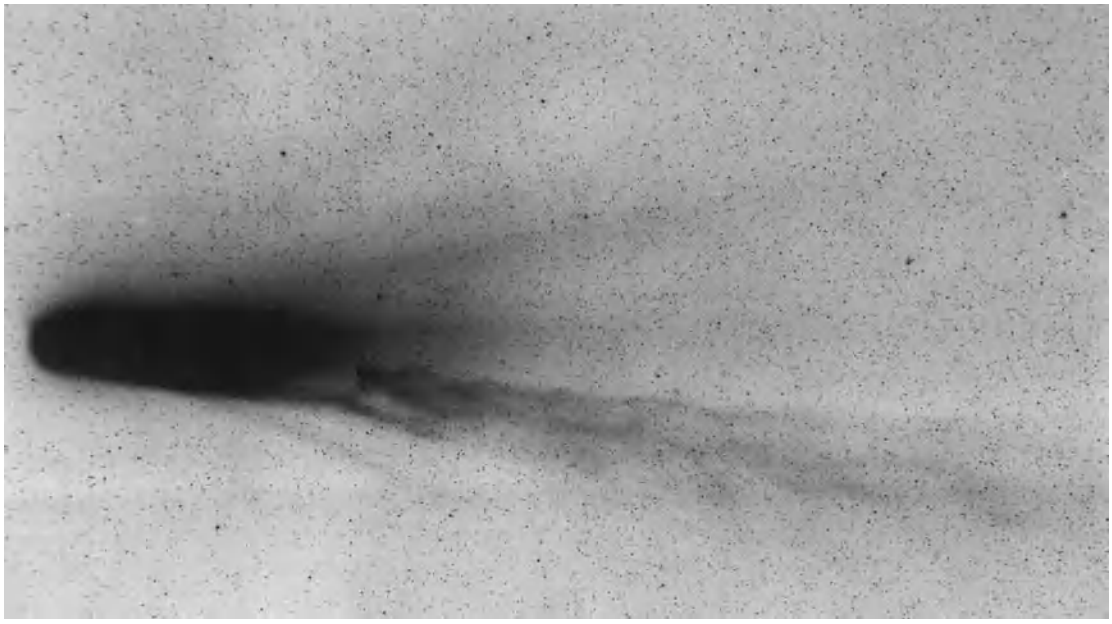

Fig. 1. Halley's Comet as photographed by the United Kingdom Schmidt telescope in Australia on March 9, 1986. Dust-tail structures are visible (above), and the plasma tail (below) also shows a completely detached portion called a disconnection event. (*Copyright © by Photolabs, Royal Observatory, Edinburgh*)

Results from both space and ground-based observations focus on three general areas. The first concerns the interaction of the comet with the solar wind. H. Alfvén's basic picture has been confirmed: the comet's plasma tail is indeed formed by the comet–solar wind interaction. Molecular ions from the comet are trapped onto solar-wind magnetic field lines, causing the magnetic field lines to drape around the comet to form the plasma tail. This process has been confirmed in detail by the spacecraft observations, including the missions to Halley's Comet and the *International Cometary Explorer*, which also encountered Comet Giacobini-Zinner at a distance of 7800 km (4845 mi) on September 11, 1985, and passed directly through the plasma tail. Somewhat surprising is the immense distance over which the interaction takes place, up to approximately 1×10^7 km (6×10^6 mi) or more from the comet, as measured by the spacecraft. Plasma processes in comets can produce spectacular results, such as disconnection events (Fig. 1).

The second area is chemical composition. The composition of the gas as measured in the inner coma is approximately 80% water (H_2O), roughly 10% carbon monoxide (CO) as determined from a rocket observation, approximately 3.5% carbon dioxide (CO_2), a few percent in complex organic compounds such as polymerized formaldehyde $(H_2CO)_n$, and the remainder in trace elements. The deuterium-to-hydrogen ratio (D/H) was initially found to be close to the value for terrestrial ocean water, although the uncertainty in the initial value was high. These results supported the view that comets may have supplied an important fraction of the volatile elements to the terrestrial planets, possibly including prebiotic molecules to Earth. However, a refined analysis gives a D/H value well above the value for ocean water. Thus, if the D/H value for Halley's Comet is typical, comets cannot have been the sole source of ocean water. Note that the composition values refer to a comet that has passed through the inner solar system many times. The values for the deep interior of Halley's Comet or for other comets may be different and would probably have a higher value of carbon dioxide.

The dust composition was found to be as follows. Some particles are composed primarily of the light atoms hydrogen (H), carbon (C), nitrogen (N), and

Space missions to Halley's Comet

Spacecraft	Sponsor	Approx. closest approach	Day in 1986
*Vega 1**	Soviet Union	8890 km (5520 mi)	March 6
Suisei	Japan	151,000 km (94,000 mi)	March 8
*Vega 2**	Soviet Union	8030 km (4990 mi)	March 9
Sakigake	Japan	7,000,000 km (4,350,000 mi)	March 11
*Giotto**	European Space Agency	605 km (375 mi)	March 14
International Cometary Explorer	NASA, United States	28,000,000 km (17,400,000 mi)	March 25

*Produced imaging of comet nucleus.

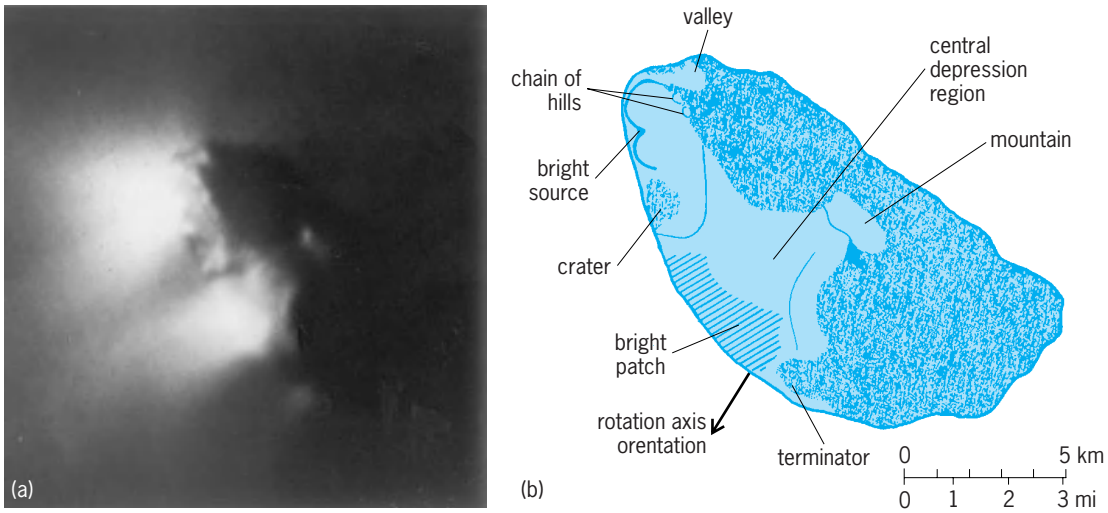

Fig. 2. Nucleus of Halley's Comet from *Giotto* spacecraft. (*a*) Composite image formed from 60 individual images. The resolution varies from about 800 m (2600 ft) at the lower right to about 80 m (260 ft) at the upper left. The material in the bright jets streams toward the Sun, leftward. (*b*) Matching drawing labeling the features on the nucleus. (*Harold Reitsema, Ball Aerospace; copyright © 1986 by Max Planck Institut für Aeronomie*)

oxygen (O), and are called CHON particles. Another kind has a silicate composition similar to the rocks that make up the crusts of Earth, Moon, and Mars, and of most meteorites. Most dust particles resemble a mixture of these two types, that is, they resemble carbonaceous chrondrites enriched in the light elements (hydrogen, carbon, nitrogen, and oxygen). These should resemble the Brownlee particles collected in the Earth's upper atmosphere.

The third general area on which the results focus is the cometary nucleus. Researchers have long based their understanding of the nucleus on F. Whipple's dirty, water–ice snowball model. According to this model, the Sun's heating of the nuclear surface produces sublimation of the ices, causing gas and dust to be released to form the cometary atmosphere and tails. All available evidence from the Halley observations, including images of the nucleus obtained by *Giotto* and the *Vega*s, confirms Whipple's model. The spacecraft observations also show that the nuclear body, shaped like a potato or peanut, is roughly 15 km (9 mi) long and 8 km (5 mi) across, somewhat larger than expected (**Fig. 2**). The surface of the comet, found to be darker than expected, has a reflectivity comparable to black velvet or coal. The nuclear surface is likely a crust of dust covering the sublimating ices. Solar energy heats the surface, and energy is conducted downward to the ices, where sublimation occurs. Some of the gases produced by the sublimation escape through thin areas or openings in the crust, forming the jets that were observed originating from the surface of Halley's nucleus.

Studies of the rotation of Halley's nucleus have provided important insights into the rotational state and interior structure of nuclei. Evidence for the rotation period of the nucleus has been presented to support values of 2.2 days and 7.4 days. In retrospect, the rotation of the highly asymmetrical nucleus is likely to be complex. Indeed, the final answer to this important question involves a model with five dominant active regions and a nucleus that rotates and spins and also precesses. This model satisfies the constraints placed on the motion by the images from the *Vega* and *Giotto* spacecraft, the observations of jets and shells, and the observed variations in brightness. The model implies an approximately constant density throughout the interior of the nucleus.

Some insight into the active comet's dusty environment can be inferred from the dust's effect on the spacecraft. The three spacecraft that passed closest to Halley—*Giotto* and the two *Vega*s—survived but were seriously damaged because of the density of the dust and the high speed of the spacecraft. The *Giotto* spacecraft was retargeted to intercept Comet Grigg-Skjellerup in July 1992.

1991 outburst. On February 12, 1991, when it was 14.4 AU from the Sun, Halley's Comet displayed a major outburst which lasted for months. The dust cloud was about 300,000 km (185,000 mi) across. Major activity this far from the Sun is unprecedented and nicely illustrates the range of physical processes in cometary science. Crystallization of amorphous ice, widely believed to be the form of water ice in the cometary interior, could be the cause of the outburst. The amorphous to crystalline ice transition releases energy that produces pockets of gas under pressure. These gas pockets could punch through the interior material and flow to the surface. The circumstances of the outburst are consistent with this model.

Associated meteors. Halley's Comet is associated with two meteor showers, the η-Aquarids and the Orionids. The origin of these showers undoubtedly involves meteoroids from the comet's nucleus that were gravitationally perturbed into their current orbits, which intersect the Earth's orbit. *See* METEOR.

John C. Brandt

Bibliography. J. C. Brandt, *Comets: A Scientific American Reader*, 1981; J. C. Brandt and R. D. Chapman, *Introduction to Comets*, 2d ed., 2004; J. Crovisier and T. Encrenaz, *Comet Science*, 2000;

M. Grewing, F. Praderie, and R. Reinhard (eds.), *Exploration of Halley's Comet*, 1988; W. F. Huebner (ed.), *Physics and Chemistry of Comets*, 1990.

Halloysite

A clay mineral similar in structure to kaolinite, having a 1:1 structure in which a silica tetrahedral sheet is joined to an alumina octahedral sheet. Unlike kaolinite, however, the structure is disordered in both the *a* and the *b* axis directions in successive layers, and it frequently contains water between the layers. *See* KAOLINITE.

Two principal modifications exist: a less hydrous form with a composition and structure near to that of kaolinite, $Al_2Si_2O_5(OH)_4$; and a hydrous form with the composition $Al_2Si_2O_5(OH)_4 \cdot 2H_2O$. The less hydrous form has a *c*-dimension of about 0.72 nanometer, whereas the hydrous form has a *c*-dimension of about 1.01 nm, the difference between them being roughly the thickness of a single sheet of water molecules. The hydrated form converts spontaneously and irreversibly into the less hydrous form when dried. The terminology for these two forms is confused. The 0.72-nm variety has been called metahalloysite, dehydrated halloysite, and halloysite. The 1.01-nm variety has been called halloysite, hydrated halloysite (or hydrohalloysite), and endellite. The recommended terminology is halloysite (0.7 nm) and halloysite (1.0 nm).

Electron microscopy reveals that the morphology of halloysite is usually tubular. Because the 1:1 layers in halloysite generally are separated from each other by water, halloysite has a larger cation exchange capacity, surface area, and catalytic activity than does kaolinite.

Halloysite is formed in nature from the weathering of feldspar under intense leaching conditions, and may also form in low-temperature hydrothermal systems. It has not been synthesized in the laboratory beyond doubt, although products resembling halloysite have been obtained by the artificial weathering of feldspar, and by the intercalation of kaolinite. Halloysite may precede kaolinite as a weathering product, and the transformation of halloysite into kaolinite may explain why halloysite is not common in sediments. *See* FELDSPAR; WEATHERING PROCESSES.

Halloysite is used as a catalyst and in the manufacture of ceramic products. *See* CLAY MINERALS.

Dennis Eberl

Bibliography. G. W. Brindley and G. Brown (eds.), *Crystal Structures of Clay Minerals and Their X-ray Identification*, 1980.

Halo

Either of two large circles of light surrounding the Sun or Moon that result from the refraction of sunlight by small, hexagonal ice crystals falling slowly through the air. Light passing through the side faces

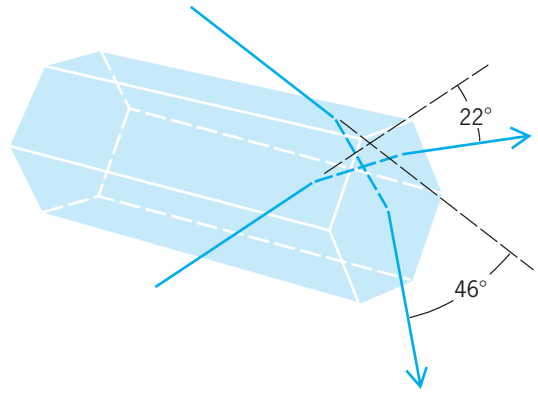

Fig. 1. Light rays passing through a hexagonal crystal to produce the 22° halo and the 46° halo.

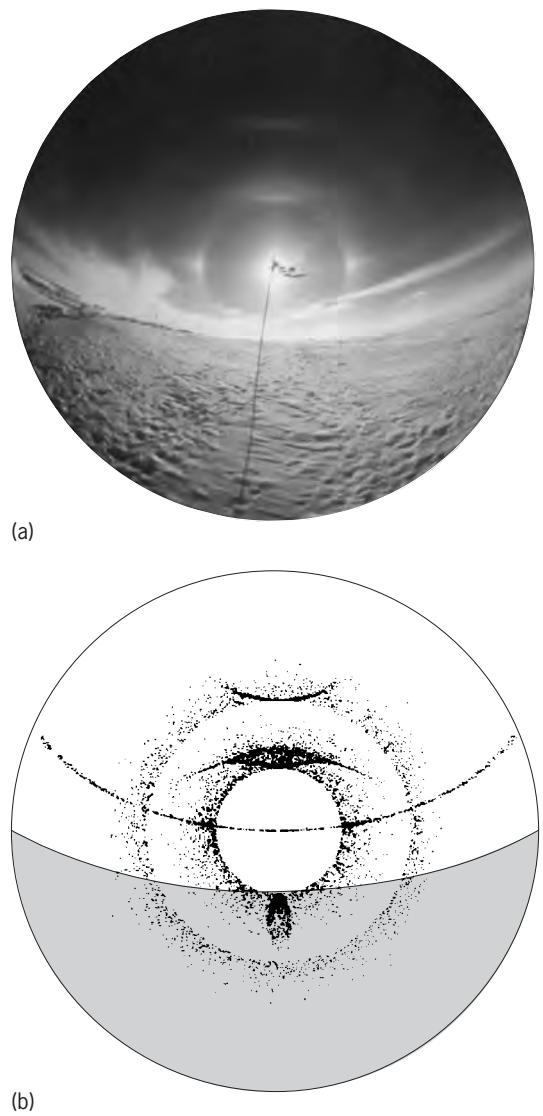

Fig. 2. Halo effects. (*a*) Photograph taken at the South Pole. (*b*) Matching computer simulations including the 22° and 46° halos, a sundog on either side of the Sun at the 22° halo position, a parhelic circle parallel to the horizon passing through the Sun, a combination of the upper tangent arc and the rarer Parry arc at the top of the 22° halo, and a circumzenithal arc at the top of the 46° halo. (*From R. Greenler, Rainbows, Halos, and Glories, Cambridge University Press, 1980*)

of a hexagonal prism is refracted by an amount that depends on the orientation of the crystal; but a collection of many crystals refracts light passing through two side faces by an average angle of about 22° (**Fig. 1**). If such crystals tumble randomly as they fall, they will produce the 22° halo, a circle around the Sun with an angular radius of 22°. Rays that pass through a side face and an end face of the prism similarly produce the larger and fainter 46° halo. The halos sometimes have a red inner edge and otherwise appear nearly white.

Many similar effects result from rays passing through ice crystals that assume special orientations as they fall, and from rays undergoing combination of reflection and refraction in an ice crystal. Usually, all of these effects are referred to as halo effects; the wide-angle photograph in **Fig. 2***a* shows several of these. The corresponding dot diagram (Fig. 2*b*) shows the simulation of the effects by a computer tracing of the rays through the crystal, made in order to identify the ray paths that cause each effect. *See* METEOROLOGICAL OPTICS; SUN DOG. Robert Greenler

Halocyprida

An order of the subclass Myodocopa, class Ostracoda, characterized by biramous antennae with the endopod and exopod of similar size, reduction or absence of the seventh pair of appendages, an unpaired male copulatory organ, and the absence of a median eye. The taxon is subdivided into two suborders, Halocypridina and Cladocopina; each has both Recent and fossil representatives. Extant halocyprids (see **illus.**) are small ostracodes, measuring less than 0.4 in. (10 mm). The shell, which may be smooth or ornamented, may be strongly calcified or soft and may or may not bear a rostrum at its anterior end. The two orders exhibit major differences. The Halocypridina possess a uniramous maxillule, whereas that of the Cladocopina is distinctly biramous. The seventh pair of appendages is reduced to one or two segments and a pair of bristles in the Halocypridina, but the sixth pair is elongate; in the Cladocopina both the sixth and seventh pairs are absent. A heart may or may not be present in the halocypridinids, but it is totally absent in cladocopinids. Differences are also evident in the mandible and armature of the caudal furca.

With the exception of one brackish-water representative, halocyprids are marine, cosmopolitan in all oceans, and found from surface waters to abyssal depths. Most are planktonic species, although a few are epibenthic or benthic, and they include filter feeders, carnivores, and possibly detritus feeders.

The stem protostracode has been hypothesized as having several characters in common with certain halocyprids, such as the structure of the sixth and seventh appendages and the pair of furca. Several primitive characters are recognized in the family Polycopacea, and the extinct Entomozoacea could be ancestral to that family. Although classified in the suborder Halocypridina, and presumably ancestral to both of its recent major taxa, the extinct Entomoconchacea are also thought to exhibit a distant connection with the Cladocopina. *See* CRUSTACEA.

Patsy A. McLaughlin

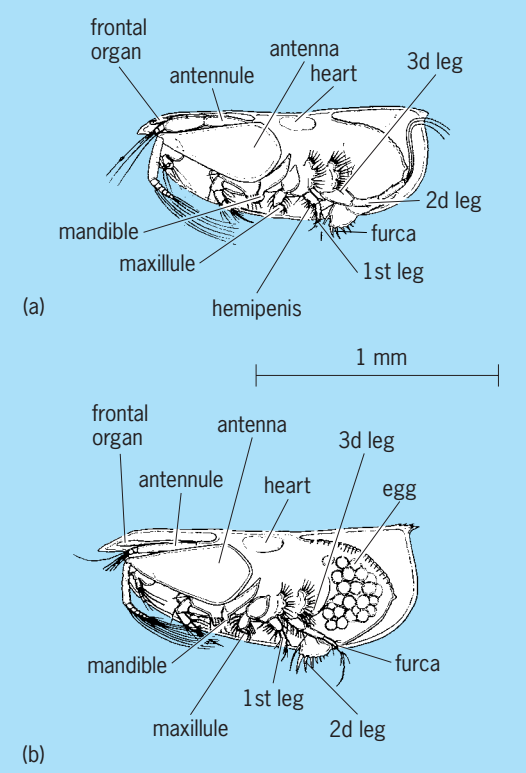

Conchoecia elegans **(superorder Halocyprida). (*a*) Male with left valve removed; long frontal organ extends into the rostrum; endopod of the antenna terminates in a hook-shaped clasper; second leg shows dimorphic enlargements. (*b*) Female with left valve removed; antennule lacks the sensory enlargement of one seta seen in the male; both first and second legs bear branchial plates. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. Q: Arthropoda 3, Geological Society of America, 1961, and D. L. McGregor and R. V. Kesling, Contrib. Mus. Paleontol. Univ. Mich., 1969*)**

Bibliography. L. G. Abele (ed.), *The Biology of Crustacea*, vol. 1: *Systematics, the Fossil Record, and Biogeography*, 1982; L. S. Kornicker and I. G. Sohn, Phylogeny, ontogeny, and morphology of living and fossil Thaumatocypridacea (Myodocopa: Ostracoda), *Smith. Contrib. Zool.*, vol. 219, 1976; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Halogen elements

The halogen family consists of the elements fluorine, F; chlorine, Cl; bromine, Br; iodine, I; and astatine, At. Estimates of the relative abundances of these elements in seawater in parts per million, and in the lithosphere (Earth's crust to a depth of 10 mi or 16 km) are given in **Table 1**. From direct measurements it is also known that all the halogen elements except astatine exist in the Earth's atmosphere.

The halogens are the best-defined family of elements. They have an almost perfect gradation of

TABLE 1. Relative abundances of the halogens, in ppm

Element	Seawater	Lithosphere
Fluorine	1.4	770
Chlorine	18,980	550
Bromine	65	1.6
Iodine	0.05	0.3

physical properties. The increase in atomic weight from fluorine through iodine is paralleled by increases in density, melting and boiling points, critical temperature and pressure, heats of fusion and vaporization, and even in progressively deeper color (fluorine is pale yellow; chlorine, yellow-green; bromine, dark red; and iodine, deep violet).

Chemically, fluorine is the most powerful oxidizing agent known. The heavier halogens have progressively less oxidizing ability. Each forms an acid with hydrogen, and salts with metals. The properties of these acids and salts show as consistent a relationship as the elements themselves. Organic halogen compounds generally show progressively increased stability in the order iodine, bromine, chlorine, fluorine.

Although all halogens generally undergo the same types of reactions, the extent and ease with which these reactions occur vary markedly. Fluorine in particular has the usual tendency of the lightest member of a family of elements to exhibit reactions not comparable to the other members. Each halogen must be considered individually, both in its preparation and in its reaction. *See* ASTATINE; BROMINE; CHLORINE; FLUORINE; HALIDE; HALOGENATION; IODINE; PERIODIC TABLE.

Despite the chemical similarities of the halogen elements, a rich variety of chemical forms and processes are involved in the injection of halogen-containing material into the atmosphere, the chemical and physical transformations that occur in the air, and the eventual removal of the halogen atoms from the atmosphere. Although their concentrations are known to vary with time and location, the values listed in **Table 2** are typical.

Particulates. In airborne particles whose diameters range from 0.1 micrometer to over 10 micrometers in the marine atmosphere, there is typically 5000 nanograms of chlorine per standard cubic meter (scm) of air, and lesser amounts of F, Br, and I as shown in Table 2. In continental air there is normally less chlorine, bromine, and iodine in airborne particles (aerosols) than near the sea, while the opposite is true for fluorine. In polluted areas there are often much higher levels of halogens in particles; specific examples include fluoride from aluminum and steel mills and phosphate-fertilizer plants and bromide from the burning of gasoline additives. Generally, the halogen form in particles and in precipitation is the halide: fluoride (F^-), chloride (Cl^-), bromide (Br^-), and iodide (I^-), although iodate (IO_3^-) may also occur. An interesting anomaly is that the amount of iodine in the marine aerosol is much larger compared to that for Cl, Br, and Na than that expected from seawater. Part of the explanation for the iodine richness of marine aerosols could be a correspondingly high iodine content of organic surface films on seawater that enter the air on bursting bubbles and sea spray.

Albert A. Gunkler

Gases. In gaseous form the halogens exist in the atmosphere in both organic and inorganic molecules. While there is at least one atmospherically significant naturally occurring organohalogen, methyl chloride (CH_3Cl), much of the atmospheric burden is anthropogenic. Table 2 shows that a typical concentration of organic F compounds is 1000 parts per trillion (ppt) by volume. Corresponding values for organochlorine, -bromine, and -iodine compounds are 2500 ppt, 20 ppt, and 3 ppt. The dominant individual species in this class of compounds are anthropogenic. They include the chlorofluorocarbons CF_2Cl_2, $CFCl_3$, $C_2F_4Cl_2$, $C_2F_3Cl_3$; the chlorocarbons CH_3CCl_3 and CCl_4 (which might also have a natural source), CH_2Cl_2, and C_2Cl_4; and perfluoromethane, CF_4. Generally, the environmental stability or inertness of these organohalogens increases with the number of halogen atoms substituted for hydrogen atoms and from the heaviest to the lightest halogen elements. The unusual stability of chlorofluorocarbons in sea level air allows accumulation in the air; their levels are increasing each year.

Dissolved halides. Most of the measurements of halides in precipitation are of the halide ion dissolved in rainwater, but some studies have been performed on melted snow. Table 2 shows typical concentrations measured in micrograms per liter of liquid precipitation. For fluoride the lowest levels, 5 micrograms/ liter, are seen far from areas influenced by continental dust or industrial pollution. For chloride the lowest values (100) are measured over continents, away from the influence of chloride-rich sea-salt aerosol. However, in polluted urban areas hydrochloric acid, HCl, is often a major component in acidic rain. The low values in Table 2 for Br^- and I^- in rainfall represent marine precipitation.

Chlorofluorocarbons. One class of halogen compounds, the anthropogenic chlorofluorocarbons mentioned above, despite being very useful as aerosol-spray-can propellants and as refrigerants,

TABLE 2. Atmospheric halogen concentrations

Source	F	Cl	Br	I
Particles (ng/scm) in marine air	0.2	5000	7	2
Organic gases (ppt)	1000	2500	20	3
Inorganic gases (ppt)	?	1000	1–10	1–5
Precipitation (μg/liter)	5–150	100–10,000	10	5

causes serious side effects. The same property of chemical inertness that makes CF_2Cl_2 and $CFCl_3$ useful also allows their concentrations to grow. Once in the air, there appears to be only one means of destroying these tightly constructed artificial molecules: upward air motions eventually carry these molecules into the upper atmosphere (stratosphere, 8–35 mi or 13–56 km above the surface) where they are attacked and decomposed by harsh ultraviolet (UV) light and energetic oxygen atoms. Chlorine atoms are released as the chlorofluorocarbons decompose, and the chain reaction below ensues:

$$\begin{array}{rcl} Cl + O_3 & \rightarrow & ClO + O_2 \\ ClO + O & \rightarrow & Cl + O_2 \\ \hline \text{net: } O + O_3 & \rightarrow & 2O_2 \end{array}$$

Ozone (O_3) and its precursor O atoms are thus destroyed in the high stratosphere. Stratospheric ozone is very important as a natural shield against biologically damaging ultraviolet rays from the Sun. Also, the absorption of solar energy by stratosphere O_3 is a major source of heat for the upper atmosphere. Natural wind systems are largely generated in this way. In addition to the chain reaction above, the chlorine species participate in many other photochemical reactions in the stratosphere. It is clear that human usage of CF_2Cl_2, $CFCl_3$, and similar chemicals is having a large impact on the chemistry of the high-altitude air, but it is not clear how much the atmospheric O_3 is being depleted by these pollutants. Continued usage of CF_2Cl_2 and $CFCl_3$ is being predicted to be capable of causing a 12% loss of atmospheric O_3 and to cause a large redistribution of the ozone with altitude. The latter distubance could affect global wind systems and climate. The facts that these chlorofluorocarbons are present in air everywhere on the Earth and have reached high altitudes are well established by reliable measurements. A second, equally important side effect of the chlorofluorocarbons is that their increasing concentrations are causing more of the Earth's infrared energy to be trapped. An appreciable global warming through a greenhouse effect is indicated. *See* GREENHOUSE EFFECT; PHOTOCHEMISTRY; STRATOSPHERIC OZONE.

The projected stratospheric consequences of chlorofluorocarbon use are largely due to the fact that each chlorofluorocarbon molecule can reside in the atmosphere for 50–100 years, thus allowing accumulation to occur. A great deal of elemental chlorine, Cl_2, is used in water treatment and in bleaching with no apparent long-lasting or global environmental effects. Chlorinated solvents are also used widely, but most of them are similarly short-lived in the open air. Potential problems due to accidental release of radioactive iodine (the 129 and 131 isotopes) from uranium fission include the possibility of accumulation in the human thyroid. *See* AIR POLLUTION; ATMOSPHERIC CHEMISTRY; FLUOROCARBON.

Ralph J. Cicerone

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; R. T. Morrison and R. N. Boyd, *Organic Chemistry*, 7th ed., 2000; D. F. Shriver and P. W. Atkins, *Inorganic Chemistry*, 3d ed., 1999.

Halogen minerals

Naturally occurring compounds containing a halogen as the sole or principal anionic constituent. There are over 70 such minerals, but only a few are common and can be grouped according to the following methods of formation. *See* HALOGEN ELEMENTS.

1. *Saline deposition by evaporation of seawater or salt lakes.* Halite (rock salt), NaCl, is the most important of this type and is found in beds covering many hundreds of square miles and ranging in thickness from a few feet to over 1000 ft (300 m). Of the other minerals associated with halite, sylvite, KCl, and carnallite, $KMgCl_3 \cdot 6H_2O$, are the most important. *See* HALITE; SALINE EVAPORITES.

2. *Hydrothermal deposition.* Fluorite, CaF_2, is the chief representative of this type and occurs in veins by itself or associated with metallic ores. Cryolite, Na_3AlF_6, may be of primary deposition or may result from the action of fluorine-bearing solutions on preexisting silicates. *See* CRYOLITE; FLUORITE.

3. *Secondary alteration.* Chlorides, iodides, or bromides of silver, copper, lead, or mercury may form as surface alterations of ore bodies carrying these metals. The most common are cerargyrite, AgCl, and atacamite, $Cu_2(OH)_3Cl$. *See* CERARGYRITE.

4. *Deposition by sublimation.* Halides formed as sublimation products about volcanic fumaroles include sal ammoniac, NH_4Cl; malysite, $FeCl_3$; and cotunnite, $PbCl_2$. At Mount Vesuvius, Italy, is the most noted occurrence of such minerals. *See* HALIDE.

5. *Meteorites.* Lawrencite, $FeCl_2$, has been found in iron meteorites. *See* METEORITE.

Cornelius S. Hurlbut, Jr.

Bibliography. L. H. Ahrens, *Origin and Distribution of the Elements*, 1979; J. D. Dana, *System of Mineralogy*, vol. 2: *Halides*, *Nitrates*, *Borates*, *Carbonates*, *Sulfates*, *Phosphates*, *Arsenates*, *Tungstates*, *Molybdites*, 7th ed., 1951.

Halogenated hydrocarbon

An aliphatic or aromatic hydrocarbon in which one or more hydrogen atoms are substituted by halogen. *See* HALOGEN ELEMENTS; HALOGENATION.

Alkyl halides. These are compounds in which one hydrogen of an alkane has been replaced by halogen [fluorine (F), chlorine (Cl), bromine (Br), or iodine (I)], for example, bromoethane (ethyl bromide; CH_3CH_2Br). Many alkyl halides have been prepared; the chlorides and bromides are most useful and most common. Since halogen atoms are much more electronegative than carbon, alkyl halides are polar molecules [dipole moment (μ) = 1.94 for chloromethane, CH_3Cl]. This property and the higher molecular weight lead to much higher boiling points and

densities compared to the parent alkanes; for example, ethane has a boiling point of −89°C (−128°F) and a density at this temperature of 0.54 g/cm^3 (0.31 oz/in.^{3}), while ethyl bromide has a boiling point of 38°C (100°F) and a density of 1.46 g/cm^3 (0.84 oz/in.^{3}). *See* DIPOLE; ELECTRONEGATIVITY; HALIDE; POLAR MOLECULE.

Preparation. Alkanes readily undergo chlorination by a radical substitution process [reactions (1), where

$$\begin{aligned} Cl_2 &\longrightarrow Cl\cdot + RH \longrightarrow R\cdot + HCl \\ R\cdot + Cl_2 &\longrightarrow RCl + Cl\cdot \end{aligned} \qquad (1)$$

RH represents an alkane, and R· is the free radical formed from the alkane]. Methyl chloride (CH_3Cl) and ethyl chloride (C_2H_5Cl) can be prepared in this way. However, the reaction is not selective for higher alkanes, and isomer mixtures are usually obtained. Free-radical chlorination is an important process in the industrial production of polychloro compounds. *See* ALKANE; FREE RADICAL.

A very general method for obtaining any alkyl chloride, bromide, or iodide is the reaction of the alcohol (ROH) with the corresponding hydrohalide (HX) or phosphorus halide (PX_3) [reactions (2) and (3), where RX represents the alkyl halide].

$$ROH + HX \longrightarrow RX + H_2O \qquad (2)$$

$$3ROH + PBr_3 \longrightarrow 3RX + H_3PO_3 \qquad (3)$$

Reactions. Alkyl halides are important starting materials for the preparation of many other functionally substituted compounds. A general reaction for chlorides, bromides, and iodides is nucleophilic substitution, in which an ion or molecule with an available electron pair (a nucleophile; Nuc:) displaces a halide ion [reaction (4)]. For example, *n*-propyl bromide

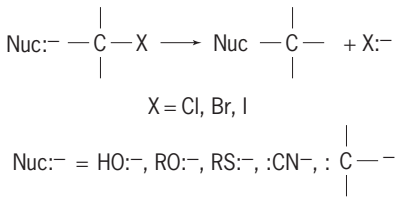

(4)

(**1**) with sodium cyanide (**2**) gives the alkyl cyanide [**3**; reaction (5)]; with a neutral nucleophile such as

$$\underset{(1)}{CH_3CH_2CH_2Br} + \underset{(2)}{NaCN} \longrightarrow \underset{(3)}{CH_3CH_2CH_2CN} \qquad (5)$$

a phosphine (**4**), a phosphonium salt (**5**) is the product [reaction (6)]. Alkyl iodides are more reactive

$$\underset{(4)}{CH_3CH_2CH_2Br + (C_6H_5)_3P} \longrightarrow \underset{(5)}{CH_3CH_2CH_2P^+(C_6H_5)_3Br^-} \qquad (6)$$

than the chlorides or bromides, but they are more expensive and are useful only for small-scale laboratory purposes. Alkyl fluorides do not undergo displacement. *See* REACTIVE INTERMEDIATES; SUBSTITUTION REACTION.

A related reaction, which competes with substitution, is elimination. In this case, the nucleophile acts as a base to remove a proton, and the product is an alkene [reaction (7)].

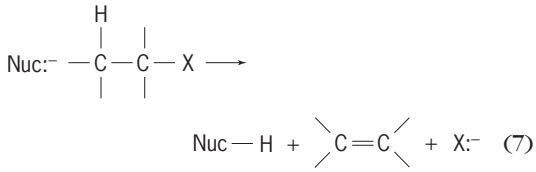

(7)

See ALKENE.

The occurrence of elimination as opposed to substitution depends on several factors. Substitution is most favorable with primary halides [RCH_2X; reaction (8)], and it is generally impractical with

$$\underset{\text{1-Chlorobutane (primary)}}{CH_3CH_2CH_2CH_2Cl} + NaOCH_3 \longrightarrow \underset{\text{Substitution product}}{CH_3CH_2CH_2CH_2OCH_3} + NaCl \qquad (8)$$

tertiary or other sterically encumbered halides, which undergo elimination instead [reaction (9)].

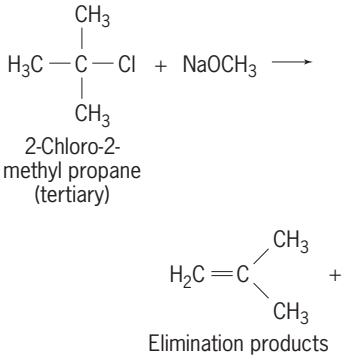

(9)

Strongly basic reagents and higher reaction temperatures promote elimination versus substitution. Alkyl fluorides are particularly prone to give the elimination product.

Another very important role of halides is in the formation of organometallic compounds. The reaction can be considered as insertion of a metal atom into the carbon-halogen bond. The organic group in the resulting organometallic compound acquires nucleophilic character, and these compounds have wide application in organic synthesis [reaction (10)].

$$R{-}Br + Mg \xrightarrow{\text{ether}} R{-}MgBr \rightleftharpoons R_2Mg + MgBr_2 \qquad (10)$$

The first organometallic compounds were alkylmagnesium halides (for example, R-MgBr, a Grignard reagent) which are readily available from any alkyl chloride, bromide, or iodide and magnesium. A major industrial use of this insertion reaction is the manufacture of dichlorodimethylsilane (**6**), which is the precursor to silicone polymers (**7**). About 600,000 tons (540,000 metric tons) per year

of methyl chloride is consumed in this process [reaction (11)].

$$2CH_3Cl + Si \longrightarrow (CH_3)_2SiCl_2 \xrightarrow{H_2O} \left(\begin{array}{c} CH_3 \\ | \\ OSi \\ | \\ CH_3 \end{array} \right)_n \quad (11)$$

(6) (7)

See GRIGNARD REACTION; ORGANIC SYNTHESIS; ORGANOMETALLIC COMPOUND; SILICONE RESINS.

Unsaturated halides. When a halogen atom is located on a doubly bonded carbon or a carbon adjacent to a double bond, the halides are termed vinylic and allylic, respectively.

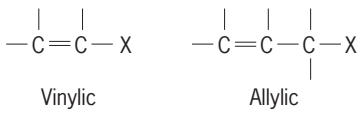

In vinylic halides, the C-X bond is shorter and stronger and undergoes substitution less readily than in a saturated halide. Conversely, halogen in an allylic position is activated toward nucleophilic substitution, and compounds with this structure are versatile intermediates in a number of synthetic applications. *See* BOND ANGLE AND DISTANCE; CHEMICAL BONDING.

The parent compound vinyl chloride (chloroethene) is one of the most important industrial organic chemicals. In the major manufacturing process, ethylene chloride is obtained by addition of chlorine to ethylene, and then it is converted to vinyl chloride (**8**) by elimination of hydrogen chloride (HCl). The chlorine used in the first step is generated by the oxidation of HCl with air. The overall process is called oxychlorination [reaction (12)]. Most of the output goes into the manufacture of poly(vinyl chloride) [PVC] and related copolymers. Vinyl chloride is a highly toxic and potentially carcinogenic compound, and stringent exposure limits are required in its manufacture and use. *See* COPOLYMER; POLYVINYL RESINS.

$$\begin{array}{l} H_2C{=}CH_2 + Cl_2 \longrightarrow ClCH_2CH_2Cl \longrightarrow H_2C{=}CHCl + HCl \\ 2HCl + O_2 \xrightarrow{CuCl_2} Cl_2 + H_2O \\ \hline H_2C{=}CH_2 + HCl + {}^1/_2O_2 \longrightarrow H_2C{=}CHCl + H_2O \end{array} \quad (12)$$

(**8**)

$$H_2C{=}CHCl + HF \longrightarrow CH_3CHClF \longrightarrow H_2C{=}CHF + HCl \quad (13)$$

(**9**)

$$H_2C{=}CHCH_3 \xrightarrow{Cl_2} H_2C{=}CHCH_2Cl \xrightarrow[NaOH]{HOCl}$$

(**10**)

$$ClCH_2CHOHCH_2Cl \xrightarrow{NaOH} H_2C\overset{O}{\diagup\diagdown}CHCH_2Cl \quad (14)$$

(**11**) (**12**)

Vinyl fluoride (**9**), also used in polymers, is obtained by addition of hydrogen fluoride (HF) to vinyl chloride followed by elimination of hydrogen chloride [HCl; reaction (13)].

Allyl chloride [**10**; reaction (14)] is another important industrial chemical, prepared by free-radical chlorination of propene. Addition of hypochlorous acid (HOCl) gives the dichloroalcohol (**11**), which is then converted to the oxide, epichlorohydrin (**12**), by reaction with sodium hydroxide (NaOH). The oxide is an intermediate in the manufacture of glycerol and in epoxy resins.

Halogen on a carbon adjacent to a benzene or other aromatic ring is termed benzylic; a compound with this structure has properties like those of an allylic halide. Benzyl chloride or bromide is obtained by free-radical halogenation of toluene under mild conditions. Further chlorination gives the dichloro- and trichloro-compounds.

Substitution reactions of benzyl chloride occur readily; benzyl alcohol is obtained by hydrolysis, and phenylacetic acid is prepared via the nitrile. Substitution with a long-chain tertiary amine leads to quaternary ammonium salts, which are widely used as bacteriostatic cationic detergents [reaction scheme (15)]. *See* QUATERNARY AMMONIUM SALTS.

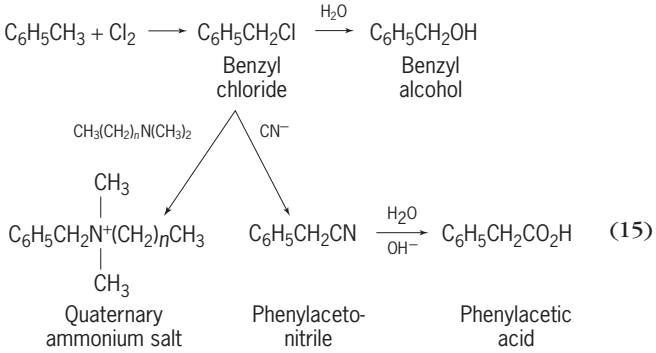

(15)

Aromatic halides. Compounds with halogen bonded directly to a benzene (**13**) or other aromatic ring are called aryl halides. Halogen is introduced by electrophilic substitution, with a Lewis acid catalyst such as $FeCl_3$ or $FeBr_3$ to enhance the positive character of the halogen [as illustrated for chlorobenzene (**14**) in reaction (16)].

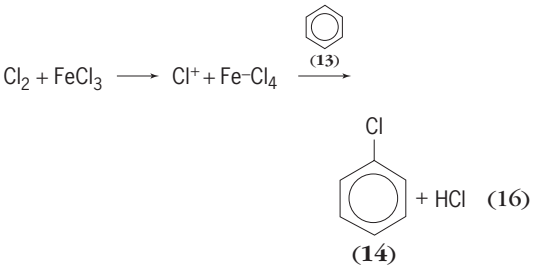

(16)

(**14**)

See ELECTROPHILIC AND NUCLEOPHILIC REAGENTS.

Halogen is an ortho-para directing, slightly deactivating group. Further substitution by electrophiles leads to a mixture of ortho (*o*) and para (*p*) isomers, with the latter predominating. Thus nitration gives

o- and *p*-chloronitrobenzene in a 1:2 ratio; both of these products are important commercial intermediates [reaction (17)].

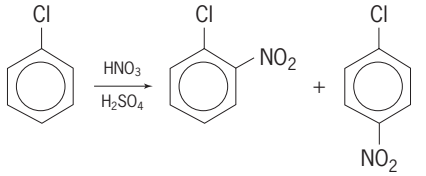

(17)

31% ortho 65% para

See NITRATION.

Chlorobenzene (**14**) is unreactive under conditions used for nucleophilic substitution in alkyl halides. At high temperature, chlorine is substituted by hydroxide ion to give phenol, but the process actually occurs by initial elimination to give the highly reactive intermediate benzyne (**15**) [reaction (18)].

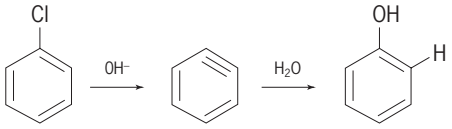

(18)

(**14**) (**15**) (**16**)

This reaction was for some years the basis of a major industrial process for phenol (**16**). *See* ORGANIC REACTION MECHANISM.

A few aryl halides that were at one time produced on a large scale are now proscribed for use in the United States. The insecticide DDT is the condensation product of chlorobenzene and chloral. Another product, a mixture of polychlorobiphenyls, was at one time widely used as a dielectric fluid in transformers. *See* AROMATIC HYDROCARBON; POLYCHLORINATED BIPHENYLS.

Fluorocarbons. In these compounds, every hydrogen atom is replaced by fluorine. Fluorocarbons can be named simply by using the prefix perfluoro- with the parent name. Because of the small atomic radius and high electronegativity of fluorine, these compounds are chemically inert and have properties quite unlike those of other halogenated organic compounds. *See* FLUORINE; FLUOROCARBON.

Hydrofluorocarbons. These organic compounds contain combinations of fluorine and hydrogen to satisfy the valency requirement of carbon. Hydrofluorocarbons are also known as HFCs. The development of specific molecules for particular applications, previously satisfied by chlorofluorocarbons, has been international in scope. Because they contain hydrogen, hydrofluorocarbons are more likely to be degraded in the lower regions of the atmosphere. Since they do not contain chlorine, these compounds do not contribute to ozone depletion.

Hydrofluorocarbons are prepared by methods similar to that used for chlorofluorocarbons and hydrochlorofluorocarbons. Process control, however, is much more critical. Asymmetrical tetrafluoroethane [HFC-134a (CF_3CH_2F)] can be prepared by addition of HF to the double bond of trichloroethylene ($Cl_2C{=}CClH$), followed by replacement of chlorine by fluorine [reaction (19)].

$$Cl_2C{=}CClH + 4HF \longrightarrow CF_3CH_2F + 3HCl \qquad (19)$$

Another useful method is the replacement of a C-Cl bond by a C-H bond by reaction with hydrogen. This conversion can begin with a suitable chlorofluorocarbon or hydrochlorofluorocarbon, as in reactions (20) and (21).

$$CF_3CCl_2F + 2H_2 \longrightarrow CF_3CH_2F + 2HCl \qquad (20)$$

$$CF_3CHCl_2 + 2H_2 \longrightarrow CF_3CH_3 + 2HCl \qquad (21)$$

Perfluorocarbons. These saturated compounds contain only carbon and fluorine. They are named by using the prefix perfluoro- along with the name of the equivalent hydrocarbon. Perfluorocarbons are chemically very inert and also have excellent thermal stability. This inertness and the resulting long atmospheric lifetime is reflected in higher global warming potentials compared to hydrofluorocarbons, hydrochlorofluorocarbons, and chlorofluorocarbons. This is due, partly, to the high strength of the C-F bond. The electronegativity of fluorine shields the carbon backbone from chemical attack. Under normal conditions, perfluorocarbons are unaffected by strong acids or bases and by oxidizing or reducing agents. *See* ELECTRONEGATIVITY.

Perfluorocarbons are dense, colorless substances. They possess unique physical and chemical properties when compared with their hydrocarbon analogs. Low boiling points, critical temperatures, and vapor pressures are the norm. Perfluorocarbons wet almost any solid because of their low surface tensions. They have unusual solubility characteristics, and their electrical properties make them outstanding insulators. These compounds do not mix with polar liquids such as water or nonpolar liquids such as hydrocarbons.

There are several methods available for the synthesis of perfluorocarbons. The first member of the series, carbon tetrafluoride (CF_4), can be prepared in a pure state by the direct reaction of fluorine and carbon. Higher homologs can also be made by this method. Another useful procedure is to react a hydrocarbon with a high-valency metal fluoride, such as cobalt trifluoride. Process conditions are adjusted for maximum fluorination. Electrochemical fluorination is also used. Here, a hydrocarbon, dissolved in anhydrous liquid HF, is subjected to an electric current. Fluorinated product and hydrogen is the end result. The replacement of a carbon-chlorine bond with a carbon-fluorine bond is also possible. This is done with HF as the fluorine source.

Polyhaloalkanes. A number of compounds with two or more halogen atoms are of special importance. Methane (CH_4) can be substituted with as many as four halogen atoms to give compounds such as CH_2Cl_2, CHI_3, and CF_3Br. Several polyhalomethane, -ethane, and -ethylene derivatives have major uses, and they are industrial chemicals produced in large quantities.

The accumulation of halogen atoms on a saturated chain reaches a limit at three carbons, except for the

important case of fluorine. The completely chlorinated (perchloro) propane C_3Cl_8 is known; but with longer chains, the perchloro compounds are alkenes and dienes such as hexachlorobutadiene. Similarly, exhaustive chlorination of a variety of C-5 alkanes or alkenes leads to further chlorination, elimination, and cyclization; the final product is hexachlorocyclopentadiene (**17**). Diels-Alder adducts of this diene [reaction (22)] have been used in the past

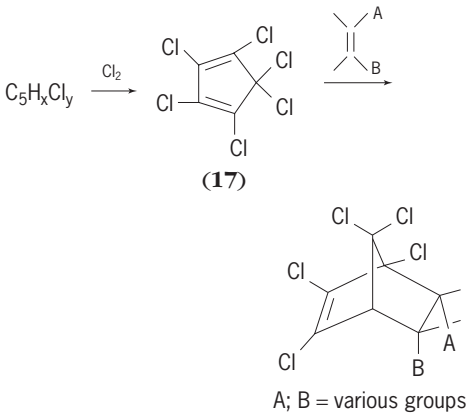

as insecticides but they have become greatly restricted in the United States. *See* DIELS-ALDER REACTION.

Chlorinated C_1 and C_2 compounds. Chlorination of methane leads to mixtures of mono-, di-, tri-, and tetrachloro products. The relative amounts can be controlled by adjusting the ratio of starting materials [reaction (23); with equimolar amounts of methane and chlorine].

$$CH_4 + Cl_2 \longrightarrow \underset{37\%}{CH_3Cl} + \underset{41\%}{CH2Cl_2} + \underset{19\%}{CHCl_3} + \underset{3\%}{CCl_4} \qquad (23)$$

Methyl chloride is manufactured by this chlorination process and also by reaction of methanol and HCl. Methylene chloride, CH_2Cl_2, is the major product from methane, and is utilized primarily as a cleaning solvent or as a blowing agent for plastic foam. Methylene chloride is more volatile and much less toxic than $CHCl_3$ or CCl_4.

The trihalomethanes are known as haloforms. Chloroform is produced by methane chlorination. Its major use is in the manufacture of chlorofluorocarbons. Chloroform found some early use as an inhalation anesthetic; while the margin between anesthetic and lethal concentrations was narrow, chloroform possessed the virtue of nonflammability.

Chloroform, bromoform, and iodoform (general formula CHX_3) can all be obtained by the haloform reaction, which was at one time the manufacturing process for chloroform. In this process a methyl ketone (**18**) is treated with halogen in the presence of base [reaction (24)]. The three α-hydrogen atoms

$$\underset{(18)}{R-\overset{O}{\overset{\|}{C}}CH_3} \xrightarrow[OH^-]{3X_2} \underset{(19)}{-\overset{O}{\overset{\|}{C}}CH_3} \xrightarrow{OH^-} R-CO^- + CHX_3 \qquad (24)$$

are substituted via the enol by Cl, Br, or I, and the trihaloketone (**19**) is then susceptible to cleavage by a base to give the haloform and a carboxylate anion. Before the availability of spectroscopic methods, this reaction was used as a diagnostic test for the —$COCH_3$ group in an unknown compound. Reaction of the substance in question with I_2 and a base gave a yellow precipitate of iodoform if this group was present.

The ease of formation of haloforms in aqueous solution has led to the problem that these compounds, referred to collectively as THM (trihalomethane), have become ubiquitous environmental contaminants. A major source is the halogenation process used in municipal water treatment. In the United States a maximum contaminant level (MCL) of 0.1 mg/liter for THM has been set by the Environmental Protection Agency. *See* WATER TREATMENT.

The main process for production of carbon tetrachloride has become the high-temperature chlorinolysis of mixed, partially chlorinated two- and three-carbon alkenes and alkanes that are by-products from other processes. Tetrachlorethylene ($Cl_2C{=}CCl_2$) is also obtained, and the relative amounts of CCl_4 and $Cl_2C{=}CCl_2$ can be varied by controlling the temperature and amount of chlorine.

Three chlorinated C_2 compounds are produced in very large volume, primarily for solvent use. 1,1,1-Trichloroethane (Cl_3CCH_3) is one of the least toxic of the high-chlorine compounds and is used extensively as a solvent in the manufacture of electronic components. Trichloroethylene ($HClC{=}CCl_2$) has been found to contribute significantly to photochemical smog, and its use is being curtailed. Tetrachloroethylene ($Cl_2C{=}CCl_2$) is the principal solvent in textile dry cleaning. *See* BROMINE; CHLORINE; IODINE.

Chlorofluorocarbons. These substances are methane and ethane derivatives with all hydrogen atoms replaced by combinations of chlorine and fluorine. Chlorofluorocarbons are known collectively as CFCs. The first of these compounds, dichlorofluoromethane (CCl_2F_2), was introduced in the 1930s as a nontoxic, nonflammable working fluid in refrigeration equipment to replace ammonia and sulfur dioxide. Other compounds were developed to meet the requirements of specific uses such as air-conditioning equipment in buildings and vehicles, and propellants for aerosols.

The principal method of preparation is replacement of chlorine by fluorine with HF in the presence of a metal fluoride. The reaction can be controlled to give the desired fluorine content [reaction (25)].

$$2CCl_4 + 3HF \longrightarrow CCl_3F + CCl_2F_2 + 3HCl \qquad (25)$$

Hydrochlorofluorocarbons. These simple organic compounds contain combinations of hydrogen, chlorine, fluorine, and carbon to satisfy the valency requirement of carbon. Also known as HCFs, the hydrochlorofluorocarbons have been developed as interim substitutes for chlorofluorocarbons. Since they have at least one hydrogen atom in the molecule,

they are more likely to be degraded in the troposphere by reaction with hydroxyl (OH) radicals. Thus, the potential that hydrochlorofluorocarbons have to deplete ozone by migration to the stratosphere is reduced.

The hydrochlorofluorocarbons with lower molecular weights are dense, colorless gases or water-immiscible liquids. Depending on the number of hydrogen atoms present in the molecule compared to fluorine and chlorine, they may be flammable or nonflammable. Nonflammability is required for most commercial applications.

Hydrochlorofluorocarbons are used as working fluids in refrigeration and air-conditioning equipment, as foam-blowing agents, and as solvents. Both rigid and flexible polyurethane, polystyrene, or polyethylene foam are used extensively in insulation and in packaging applications. The foaming agent reduces foam density and increases softness. Previously, trichlorofluoromethane ($CFCl_3$), designated as CFC-11, was used widely for this purpose. Hydrochlorofluorocarbons are also used as solvents to remove materials such as grease in the electronic and aerospace industries.

Simple hydrochlorofluorocarbons are prepared by methods similar to those used for chlorofluorocarbons. Because of the presence of hydrogen and the reduced stability, the process conditions are much more complex.

Ozone depletion potential. The relative potential of chlorofluorocarbons and other compounds to deplete ozone (O_3) in the stratosphere is known as the ozone depletion potential.

The Earth's atmosphere is a complicated and dynamic system. Simplistically, there are several regions of the atmosphere, defined by their temperature and structure. In the lowest region, the troposphere, temperature decreases with increasing altitude, and mixing of the air is rapid. The next region is called the stratosphere. Here, the temperature increases with altitude, and mixing is slower. These two regions are separated by the tropopause. The tropopause is roughly 5–6.8 mi (8–11 km) above the Earth's surface; and the stratosphere, which lies above it, extends to about 30 mi (50 km) above the Earth's surface. The stratosphere contains ozone, which shields the Earth's surface from high-energy solar radiation, ultraviolet-B. *See* ATMOSPHERE; ULTRAVIOLET RADIATION.

The concentration of ozone in the stratosphere is the result of a dynamic balance of a series of reactions that continually produce and destroy ozone, and winds that move and mix the air. Stable chlorofluorocarbons are transported by winds from the troposphere into the stratosphere. Once there, they are subjected to solar ultraviolet radiation of very high energy. Chlorine atoms are produced in this process, and they react with the ozone. A simplified ozone destruction cycle can be written as shown in reactions (26)–(30). This chain reaction continues

$$CCl_3F + \text{ultraviolet} \longrightarrow Cl + CCl_2F \quad (26)$$

$$Cl + O_3 \longrightarrow ClO + O_2 \quad (27)$$

$$O_3 + \text{light} \longrightarrow O + O_2 \quad (28)$$

$$ClO + O \longrightarrow Cl + O_2 \quad (29)$$

$$O_3 + O_3 \longrightarrow 3O_2 \quad (30)$$

until stopped by other mechanisms. The net result is shown in reaction (27). *See* CHAIN REACTION (CHEMISTRY).

Because of concern about ozone depletion, the use of chlorofluorocarbons in aerosol products in the United States was banned in 1978. Extensive studies linking seasonal ozone depletion over the Antarctic continent to chlorofluorocarbons and other compounds led to calls for action. An international agreement, known as the Montreal Protocol, required a phase-out of production of chlorofluorocarbons in developed countries by the year 1996. This accelerated the search for suitable alternatives in applications hitherto fulfilled by chlorofluorocarbons.

The ozone depletion potential depends on several factors. These include the number of chlorine atoms in the molecule and the ability of the compound to degrade in the troposphere before significant amounts can reach the stratosphere. In order to reduce the impact on the ozone layer, substitutes for chlorofluorocarbons have been designed to contain hydrogen to decrease their atmospheric stability. *See* AEROSOL; ATMOSPHERIC CHEMISTRY; STRATOSPHERIC OZONE.

Greenhouse effect. The thermal and radiation balance of the Earth is determined by a balance of heating by incoming solar energy and cooling by outgoing infrared radiation from the Earth's surface. Many compounds block the release of infrared radiation from lower regions of the atmosphere back into space. Chlorofluorocarbons, carbon dioxide (CO_2), and methane are a few examples. Gases, such as those mentioned, block release of infrared energy by absorbing energy frequencies that are characteristic of their structure. This retained energy results in a warming of the troposphere and the surface of the Earth known as the greenhouse effect. Carbon-halogen (especially fluorine) bonds are very efficient in absorbing certain frequencies of this infrared energy. The ability of compounds to trap radiation depends on such factors as stability in the atmosphere, extent of halogen substitution, and molecular structure. *See* GREENHOUSE EFFECT; HEAT BALANCE, TERRESTRIAL ATMOSPHERIC.

The global warming potential of halocarbons can be reduced by decreasing their stability in the atmosphere. This can be done by including hydrogen as part of the molecule. James A. Moore; V. N. M. Rao

Bibliography. R. E. Banks, B. E. Smart, and J. C. Tatlow (eds.), *Organofluorine Chemistry, Principles and Commercial Applications*, 1994; D. A. Fisher et al., Model calculations of the relative effects of CFCs and their replacements on stratospheric ozone, *Nature*, 344:508–512, 1990; M. Hudlicky, *Chemistry of Organic Fluorine Compounds*, 2d ed., 1992; L. E. Manzer and V. N. M. Rao, Catalytic synthesis of chlorofluorocarbon (CFC) alternatives, *Adv. Catal.*,

39:329–350, 1993; D. A. O'Sullivan, International gathering plans ways to safeguard atmospheric ozone, *Chem. Eng. News*, 67(26):33–36, 1989; F. Rowland, Stratospheric ozone in the 21st century: The chlorofluorocarbon problem, *Environ. Sci. Technol.*, 25:622–628, 1991; A. Streitwieser, C. Heathcock, and E. M. Kosower, *Introduction to Organic Chemistry*, 4th ed., 1992, revised 1998.

Halogenation

A chemical reaction or process which results in the formation of a chemical bond between a halogen atom and another atom. Reactions resulting in the formation of halogen-carbon bonds are especially important, and the emphasis in this article is primarily on such reactions. The halogenated compounds produced are employed in many ways, for example, as solvents, intermediates for numerous chemicals, plastic and polymer intermediates, insecticides, fumigants, sterilants, refrigerants, additives for gasoline, and materials used in fire extinguishers. *See* HALOGEN ELEMENTS.

Numerous compounds are halogenated by a wide variety of commercial and laboratory processes. Halogenation reactions can be subdivided in several ways, for example, according to the type of halogen (fluorine, chlorine, bromine, or iodine), type of material to be halogenated (paraffin, olefin, aromatic, hydrogen, and so on), and operating conditions and methods of catalyzing or initiating the reaction.

Halogenation reactions with elemental chlorine, bromine, and iodine are of considerable importance. Because of high exothermocities, fluorinations with elemental fluorine tend to have high levels of side reactions. Consequently, elemental fluorine is generally not suitable for direct fluorination. Two types of reactions are possible with these halogen elements, substitution and addition.

Substitution halogenation. This type of halogenation is characterized by the substitution of a halogen atom for another atom (often a hydrogen atom) or group of atoms (or functional group) on paraffinic, olefinic, aromatic, and other hydrocarbons. A chlorination reaction of importance that involves substitution is that between methane and chlorine. The overall reaction between methane and chlorine is shown in reaction (1).

$$CH_4 + Cl_2 \longrightarrow CH_3Cl + HCl \qquad (1)$$

Substitution reactions with paraffins involve several free-radical reaction steps: only the most important steps are shown below. In the initiating step, chlorine free radicals are formed as follows: Cl_2 (elemental chlorine) → 2 Cl· (chlorine free radical). Breaking elemental chlorine into free radicals requires energy, and can be accomplished thermally at high temperatures (about 480°F or 250°C, or higher) or at lower temperatures (including ambient temperatures) by means of radiation (gamma radiation, sunlight, and so on).

Heats of reaction in radical halogenation*

Reaction	F_2	Cl_2	Br_2
RX + X· → R· + HX	−33	−3	+13
R· + X_2 → RX + X·	−71	−24	−21

*In kilocalories per mole (1 kcal = 4184 joules).

The chlorine free radical starts a chain mechanism with methane, as in reactions (2) and (3).

$$CH_4 + Cl\cdot \longrightarrow CH_3\cdot + HCl \qquad (2)$$

$$CH_3\cdot + Cl_2 \longrightarrow CH_3Cl + Cl\cdot \qquad (3)$$

Similar reaction steps also occur with other paraffins or hydrogen. The relative rates of halogenation vary according to the type of carbon-hydrogen bond, as follows: tertiary > secondary > primary. Free-radical reactions such as those shown above are sometimes affected to a considerable extent by the surface of the reactor employed. The kinetics of chlorination can vary by factors of at least four or five, depending on the material of construction of the reactor or the past history of the reactor.

The **table** gives approximate heats of reaction for the two steps in the chain mechanism, such as those shown in reactions (2) and (3) for substitution halogenation with fluorine, chlorine, and bromine. Halogenation reactions are exothermic; the level of exothermicity is $F_2 > Cl_2 > Br_2 > I_2$. *See* SUBSTITUTION REACTION.

Addition halogenation. Chlorine, bromine, and iodine react readily with most olefins; the reaction between ethylene and chlorine, to form 1,2-dichloroethane, reaction (4), is one of considerable

$$CH_2{=}CH_2 + Cl_2 \longrightarrow CH_2Cl{-}CH_2Cl \qquad (4)$$

commercial importance, since it is used in the manufacture of vinyl chloride. This reaction occurs readily in the gas phase at temperatures in the range of 175 to 265°F (80 to 130°C) when the inner surface of the reactor is coated with calcium or lead chlorides.

Reaction (4) also occurs readily in the liquid phase by means of ionic reactions at temperatures from 85 to 158°F (30 to 70°C). Ferric chloride, often formed by reaction of chlorine with iron in the walls of the reactor, is an excelent catalyst.

Addition reactions with bromine or iodine are frequently used to measure quantitatively the number of —CH=CH— (or ethylenic-type) bonds on organic compounds. Bromine numbers or iodine values are measures of the degree of unsaturation of the hydrocarbons.

Halogenation of aromatics. Substitution halogenation on the aromatic ring can be made to occur by means of ionic reactions. The chlorination reactions with elemental chlorine occur under conditions similar to those used for addition chlorination of olefins; temperatures of 68–122°F (20–50°C) are suitable, and ferric chloride is an effective catalyst. Polychlorination of benzene results in various

isomers of dichloro-, trichloro-, and tetrachlorobenzenes; even higher levels of chlorination sometimes occur.

When liquid benzene is contacted with chlorine at essentially ambient temperature in the presence of radiation (light radiation, gamma radiation, and so on), but in the absence of catalysts, the chlorine adds to the benzene by means of a free-radical chain mechanism to form benzene hexachloride, an effective insecticide. The overall reaction is shown as reaction (5).

$$C_6H_6 + 3Cl \xrightarrow[\text{radiation}]{\text{86–104°F or 30–40°C}} C_6H_6Cl_6 \qquad (5)$$

Bromine and benzene also react in the presence of sunlight; benzene hexabromide is formed, however, with difficulty.

When alkyl benzenes are halogenated, the halogen can react either on the aromatic nucleus or on the alkyl group. Free-radical (and gas-phase) chlorinations generally result in substitution, primarily on the alkyl groups; such reactions occur readily at 248–266°F (120–130°C). Ionic, low-temperature reactions, however, favor substitution on the nucleus.

At 750–930°F (400–500°C), chlorobenzenes are produced from benzene by free-radical chlorination steps. The resulting products have an isomer distribution quite different from that obtained by means of ionic chlorinations at essentially ambient temperatures.

Halogenations with hydrogen halides. Numerous methods are available for halogenating with hydrogen halides. Oxychlorination reactions can be used with olefinic, paraffinic, and aromatic hydrocarbons. The reaction with ethylene (6) is of major industrial

$$CH_2{=}CH_2 + 2HCl + {}^1/_2O_2 \xrightarrow[\text{180–750°F or 250–400°C}]{CuCl_2} CH_2Cl{-}CH_2Cl + H_2O \qquad (6)$$

importance for the production of 1,2-dichloroethane in modern vinyl chloride plants.

Hydrogen halides can also be made to react with many ethylenic and acetylenic compounds. Examples are shown in reactions (7) and (8).

$$CH_2{=}CH_2 + HCl \xrightarrow[\text{212°F or 100°C}]{ZnCl_2} CH_3CH_2Cl \qquad (7)$$

$$HC{\equiv}CH + HCl \xrightarrow[\text{94–284°F or 90–140°C}]{HgCl_2} H_2C{=}CHCl \qquad (8)$$

Before 1965 reaction (8) was the main method used in producing vinyl chloride. Fluorocarbons widely used as refrigerants are produced from carbon tetrachloride; the production of trichlorofluoromethane is shown in reaction (9). Alcohols react

$$CCl_4 + HF \xrightarrow[\text{212°F or 100°C}]{SbF_5} CCl_3F + HCl \qquad (9)$$

with hydrogen halides to form alkyl halides, as shown in reactions (10) and (11).

$$CH_3OH + HBr \rightarrow CH_3Br + H_2O \qquad (10)$$

$$CH_3CH_2OH + HCl \xrightarrow[\text{230–284°F or 110–140°C}]{ZnCl_2} C_2H_5Cl + H_2O \qquad (11)$$

Electrochemical techniques are employed commercially for substituting fluorine atoms for hydrogen on the alkyl group of organic acids, amines, nitriles, and alcohols. In an electrolytic cell operated at low voltages hydrogen is also a product.

Miscellaneous halogenating agents. Many compounds have been employed as halogenating agents. Several examples are sodium hypochlorite, phosgene ($COCl_2$), thionyl chloride, sulfuryl chloride, ferric chloride, antimony chlorides, and phosphorus chlorides, which can be used for certain chlorination reactions; hypobromites, *N*-bromosuccinimide, and *N*-bromoacetamide, which have proved to be brominating agents; antimony pentafluoride, silver difluoride, and lead tetrafluoride, metal fluorides which are used as fluorinating agents; and iodine monochloride and alkali hypoiodites, which are iodination agents. *See* HALOGENATED HYDROCARBON. Lyle F. Albright

Bibliography. P. B. De la Mare and R. Bolton, *Electrophilic Additions to Unsaturated Systems*, 2d ed., 1982; R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; V. Gutmann, *Main Group Elements: Groups 6–7*, 1975.

Halophilism (microbiology)

The requirement of high salt (NaCl) concentrations for growth of microorganisms. Microorganisms (mainly bacteria) can be classified by their physiological tolerance to salt. Most normal eubacteria, such as *Escherichia coli* and *Pseudomonas fluorescens*, and most fresh-water microorganisms, are nonhalophiles (best growth at less than 1.2% NaCl). Slight halophiles (1.2–3% NaCl) include many marine microorganisms. Moderate halophiles (3–15% NaCl) include *Vibrio costicola*, *Paracoccus halodenitrificans*, and many others. Borderline extreme halophiles (9–25% NaCl) include the photosynthetic bacterium *Ectothiorhodospira halophila*, the actinomycete *Actinopolyspora halophila*, and the halophylic archaebacteria *Halobacterium volcanii* and *H. mediterranei*. Extreme halophiles (require at least 10% NaCl; optima 15–30% NaCl) are *Halobacterium salinarium* and *Halococcus morrhuae*. *See* METHANOGENESIS (BACTERIA).

The halophilic aerobic archaebacteria give a striking red color to hypersaline waters. They are found in the Dead Sea, the Great Salt Lake, Lake Magadi in Kenya, and other alkaline salt lakes, and in solar salterns where salt is prepared by evaporating seawater. Their red color is due to carotenoid pigments (bacterioruberins), which seem to protect them from strong sunlight in their natural environments. *See* CAROTENOID.

Some genera fit into more than one halophilic

classification. Thus, some *Ectothiorhodospira* species can grow in lower salt concentrations than *E. halophila*. Some species of *Dunaliella*, a flagellated alga found in salt lakes, behave as moderate halophiles, and others as borderline extreme halophiles.

The range of salt concentration at which any microorganism grows may be affected by nutritional conditions and by temperature. Generally, at lower temperatures cells grow better at lower salt concentrations, and they can grow at a wider range of salt concentrations in richer media.

Many species of salt-tolerant microorganisms also exist which do not require salt but can grow in high salt concentrations. These include lactic acid bacteria, which can grow in 8% NaCl or higher, and some bacilli and micrococci that can grow in 25% NaCl. Yeasts (for example, *Saccharomyces rouxii*) and filamentous fungi (for example, *Xeromyces* spp.) can grow in very high concentrations of salt or sugars, such as sucrose; usually, they do not require such concentrations for growth. Such microorganisms are often used in the processing of salted foods, such as sauerkraut or soy sauce; but they will spoil jams and jellies.

It is not always possible to distinguish between the direct action of salts and sugars on cells and their indirect action, through reduction of available water in the cell's environment. True halophiles, however, have a specific requirement for NaCl whose ions can be replaced only partially, if at all, by K^+, NO_3^-, or other ions.

Adaptation to salt; compatible solutes. Microorganisms that live in high concentrations of salt or other solutes do not exclude solutes from the interior of the cell. However, the internal solute composition is quite different from the outside composition. *Dunaliella* species have internal glycerol concentrations corresponding to the external concentration of NaCl. Other salt-tolerant algae and yeasts also have high internal concentrations of glycerol, or other polyols. Solutes which maintain osmotic equilibrium between inside and outside of the cell without interfering with the cell's physiological processes are called compatible solutes. *See* OSMOREGULATORY MECHANISMS.

In halophilic eubacteria (and in some nonhalophiles exposed to moderate salt concentrations), compatible solutes include sugars and amino acids, especially glutamate and proline. However, for most halophilic organisms, the concentrations of known compatible solutes do not add up to the osmotic equivalent of external salts.

Red halophiles. Mechanisms of adaptation to a highly saline environment have been best characterized in aerobic halophilic bacteria whose enzymes are able to function in high salt concentrations; indeed, most of them require such salt concentrations for activity, stability, or both. For a number of enzyme systems, KCl rather than NaCl is required, confirming the role of KCl as a compatible solute in these bacteria.

Other parts of the cells of these bacteria also require high salt concentrations for function of stability. Halobacteria lyse and their cell walls may completely dissolve unless salt concentrations are high. NaCl is specifically required for active transport of ions and nutrients in all halophilic bacteria.

Transport and sensory pigments. When halobacteria are grown in conditions of limiting oxygen supply, up to 50% of their cytoplasmic membrane becomes covered by organized patches of purple material. The color is due to bacteriorhodopsin, a combination of a protein, bacterioopsin, and retinal, the same visual pigment found in animal eyes. When membranes containing bactriorhodopsin are exposed to visible light, protons (H^+ ions) are transferred from inside to outside of the cell. The resulting proton gradient can be used as a source of cellular energy for adenosine triphosphate (ATP) production, for active transport, and for anaerobic growth. The purple-membrane system has probably provided the best direct evidence for the chemiosmotic hypothesis since in this system a proton gradient can be established by light alone. *See* ION TRANSPORT.

Purple membranes also exist in nature in blooms of halobacteria in hypersaline waters. They are thought to provide an extra source of energy to these microorganisms under conditions of low oxygen availability, which might occur often, since oxygen is much less soluble in concentrated salt solutions than in pure water.

Molecular biology and genetics. Halophilic archaebacteria, which are often easier to grow than methanogens or extreme thermophiles, provide much information regarding central metabolism, chromosome structure, transcription, translation, and the control of synthesis of ribosomal proteins and other proteins in archaebacteria. Some mutations of *Halobacteria* species take place at a very high rate because of the presence of insertion sequences. In *Haloferax* species, partly because of the relatively simple nutritional requirements of the latter halophiles, a number of genetic markers have been found. A good part of the genome of *Haloferax volcanii* has been mapped. Virtually nothing is known of the genetics of halophilic eubacteria, and few mutations have been characterized. *See* BACTERIAL GENETICS; MUTATION.

Phages. A number of phages of halophilic archaebacteria (especially *Halobacterium salinarium* and closely related species) have been isolated from salt ponds, salted fish, and apparently pure bacterial cultures. Some phages may exist in either lytic or carrier (lysogenic) states in their hosts, depending on environmental conditions. The phages of halophilic archaebacteria have the same salt requirement for stability as their hosts, and are inactivated within minutes in dilute solutions. In contrast, phages of halophilic eubacteria are relatively stable for days in the absence of salt. *See* ARCHAEA; BACTERIAL PHYSIOLOGY AND METABOLISM.

D. J. Kushner

Bibliography. A. D. Brown, *Microbial Water Stress Physiology: Principles and Perspectives*, 1990; M. Kates, D. J. Kushner, and A. T. Matheson (eds.), *The Biochemistry of Archaea (Archaebacteria)*,

1993; T. Kauri et al., A bacteriophage of a moderately halophilic bacterium, *Arch. Microbiol.*, 156:435–438, 1991; D. J. Kushner (ed.), *Microbial Life in Extreme Environments*, 1978; F. Rodriguez-Valera (ed.), *Halophilic Bacteria*, vols. 1 and 2, 1988; R. H. Vreeland and L. I. Hochstein (eds.), *The Biology of Halophilic Bacteria*, 1992.

Haloragales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Rosidae of the class Magnoliopsida (dicotyledons). The order consists of 2 families, with about 150 species in all. The Haloragales are herbs with perfect or often unisexual, more or less reduced flowers. The perianth is minute or vestigial and was probably originally tetramerous. The ovary is inferior, with only one ovule in each locule, and the seeds have a more or less abundant endosperm. Many of the species are aquatic. Entomophily (pollination by insects) has been largely abandoned in the group, and the pollen is commonly distributed by wind or water. The aquarium plant called parrot's feather (*Myriophyllum*, family Haloragaceae) and the very large-leaved plant *Gunnera* (family Gunneraceae) are well-known members of the Haloragales. *See* MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM; ROSIDAE.

Arthur Cronquist; T. M. Barkley

Hamamelidae

A small subclass of the class Magnoliopsida (dicotyledons) in the division Magnoliophyta (Angiospermae), the flowering plants, consisting of 11 orders (Trochodendrales, Hamamelidales, Daphniphyllales, Didymelales, Eucommiales, Urticales, Leitneriales, Juglandales, Myricales, Fagales, and Casuarinales), 24 families, and about 3400 species. They have strongly reduced, often unisexual flowers with poorly developed or no perianth. In the more advanced types, the flowers of one or both sexes are borne in catkins, and the mature fruit is unilocular and indehiscent, with only a single seed. With the notable exception of some of the Urticales, they are all woody plants. Pollination is usually by wind. Many of the families formerly grouped under the Amentiferae belong to the Hamamelidae. See articles on each order. *See* MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Hamamelidales

A small order of flowering plants, division Magnoliophyta (Angiospermae), which gives its name to the subclass Hamamelidae in the class Magnoliopsida (dicotyledons). The family Hamamelidaceae contains about 100 species and the Platanaceae about 6 species; the other 3 families have only 2 species each. Within its subclass, the order is more advanced than the Trochodendrales in having vessels in the wood, but less advanced than the other orders in that the gynoecium consists either of separate carpels or of united carpels that open at maturity to release the seeds. Witch hazel (*Hamamelis*; see **illus.**), sweet gum (*Liquidambar*, family Hamamelidaceae), and the plane tree or sycamore (*Platanus*) are familiar members of the Hamamelidales. *See* HAMAMELIDAE; MAGNOLIOPHYTA; PLANT KINGDOM; TROCHODENDRALES.

Arthur Cronquist; T. M. Barkley

American witch hazel (*Hamamelis virginiana*), a characteristic member of the family Hamamelidaceae and order Hamamelidales. The flowers of this species open in the autumn, when the leaves are already beginning to deteriorate. (*John H. Gerard, National Audubon Society*)

Hamilton-Jacobi theory

A theory that uses canonical transformations to construct a general procedure for solving mechanical problems. Either of two methods can be used.

1. If the Hamiltonian is conserved, one can transform to a set of canonical coordinates such that only the momenta appear. That is, the coordinates do not appear in the transformed Hamiltonian. Such coordinates are termed cyclic, or ignorable. This procedure generates equations with trivial solutions.

2. One can find a cononical transformation from the coordinates and momenta (q,p) at a time t to a new set of constant quantities. Usually these quantities are chosen to be the values of the coordinates and momenta (q_0,p_0) at $t = 0$.

With such a transformation, the equations of transformation relating old and new canonical variables are the desired solution to the problem. The new coordinates depend only upon the initial values and time. This latter method is more general and can also be used when the Hamiltonian is time-dependent.

Either of these procedures comprises the Hamiltonian-Jacobi theory. Both require a canonical transformation of the coordinates so that the transformed Hamiltonian is identically zero. The

transformation eliminates the original momenta and results in a set of $(f + 1)$ differential equations in the coordinates q_i and t, where f is the number of degrees of freedom of the system. These equations can, in principle, be solved. *See* CANONICAL COORDINATES AND TRANSFORMATIONS; DEGREE OF FREEDOM (MECHANICS); DIFFERENTIAL EQUATION; HAMILTON'S EQUATIONS OF MOTION.

Schrödinger equation. The resulting Hamilton-Jacobi equation bears a close formal resemblance—in fact, much more than formal—to the Schrödinger wave equation. The Hamilton-Jacobi equation determines the canonical transformation from a set of initial coordinates and momenta to those at time t. Likewise, in quantum mechanics the state vector at t is found from the state vector at an earlier time t_0 by a unitary transformation, as in Eq. (1). The

$$\psi(t) = U(t,t_0)\psi(t_0) \tag{1}$$

unitary operator U satisfies the Schrodinger equation (2), which resembles the Hamilton-Jacobi equa-

$$H\left(q, \frac{ih}{2\pi}\frac{\partial}{\partial q}, t\right) U = \frac{ih}{2\pi}\frac{\partial U}{\partial t} \tag{2}$$

tion for the canonical transformation. *See* NONRELATIVISTIC QUANTUM THEORY; SCHRÖDINGER'S WAVE EQUATION.

Action-angle variables. A system is said to be separable and multiply periodic if, for each coordinate pair q_j, p_j, there is a separate equation of motion such that each pair is periodic in the time, as in oscillatory motion, or p_j is a periodic function of q_j, as in rotational motion. In such systems it is useful to introduce the action variables J_k, also termed the action-angle variables, defined by Eq. (3). In this case

$$J_k = \oint p_k dq_k \tag{3}$$

the energy is expressible as a function of the action variables, as in Eq. (4).

$$E = H(J_1, \ldots, J_f) \tag{4}$$

See ACTION; QUANTUM MECHANICS.

Each action-angle variable J_k corresponds to a frequency ν_k at which the corresponding coordinate returns to its previous value. To calculate the frequencies ν_k of a separable, multiply periodic system, it is sufficient to find the energy as a function of the action variables. The derivative of this function with respect to any one action variable is the frequency of the corresponding motion. This relation provides a basis for perturbation theory.

John L. Safko; Philip Stehle

Bibliography. V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 2d ed., 1989; M. Born, *Mechanics of the Atom*, 1960; T. L. Chow, *Classical Mechanics*, 1995; H. C. Corben and P. Stehle, *Classical Mechanics*, 2d ed., Dover reprint, 1994; H. Goldstein, C. P. Poole, Jr. and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; I. Percival and D. Richards, *Introduction to Dynamics*, 1982.

Hamilton's equations of motion

A set of first-order ordinary differential equations that may be used to describe the motion of a mechanical system. Because of their remarkably symmetrical form [which appears in Eqs. (4), below], they are often referred to as the canonical equation of motion (where "canonical" is used in the sense of designating a simple general set of standard equations). The Lagrangian formulation of a system of f degrees of freedom generates f differential equations of second order in the time derivatives of the variables. Hamilton's equations, which are equivalent to Lagrange's equations, consist of $2f$ first-order and highly symmetrical equations. These properties make Hamilton's equations very useful for general discussions of the motion of systems. *See* DEGREE OF FREEDOM (MECHANICS); DIFFERENTIAL EQUATION; LAGRANGE'S EQUATIONS.

Definitions. Hamilton's equations follow from Lagrange's equations in a straightforward manner. The generalized coordinates of a system with f degrees of freedom can be labeled q_j $(j = 1, 2, \ldots, f)$, and the dynamical description of the system can be given by the Lagrangian $L(q,\dot{q},t)$, where q denotes all the coordinates and a dot denotes the total time derivative. Lagrange's equations are then given by Eq. (1).

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0 \tag{1}$$

The momentum p_j canonically conjugate to q_j is defined by Eq. (2).

$$p_j = \frac{\partial L}{\partial \dot{q}_j} \tag{2}$$

It is assumed that Eq. (2) is soluble for the velocities $\dot{q}_j$ in terms of the coordinates and momenta. *See* LAGRANGIAN FUNCTION.

The Hamiltonian $H(q,p,t)$ is defined by Eq. (3). Dif-

$$H(q, p, t) = \sum_{j=1}^{f} p_j\dot{q}_j - L(q, \dot{q}, t) \tag{3}$$

ferentiating this equation and then using Lagrange's equations leads to Eqs. (4), which are Hamilton's

$$\dot{q}_j = \frac{\partial H}{\partial p_j} \qquad p_j = -\frac{\partial H}{\partial q_j} \tag{4}$$

equations. It is essential that H be written as a function of the coordinates and momenta without the velocities appearing.

Phase space. Hamilton's equations are most easily interpreted by introducing the phase space of the system. This is a space of $2f$ dimensions in which the canonical coordinates and momenta serve as the Cartesian coordinates of a point, the system's phase point, that represents the state of motion of the system. Hamilton's equations give the velocity of this phase point as a function of its position in phase space and the time. The important Liouville theorem

follows directly from Hamilton's equations. Consider many points in a neighborhood of phase space at a given time. These points represent different possible states of motion of the system, or the motions of many distinct identical systems. As time proceeds, these points move according to Hamilton's equations. If these points are pictured as particles of a fluid flowing through phase space, this flow is that of an incompressible fluid because the divergence of the velocity field in the fluid vanishes. For additional information on Liouville's theorem and phase space *see* STATISTICAL MECHANICS

Applications. As they stand, Hamilton's equations (4) are seldom easier to integrate directly than Lagrange's. Elimination of the p's from Hamilton's equations leads to Lagrange's equations. Hamilton's equations are of great advantage in more general discussions, and they permit the making of canonical transformations that can lead to simplifications. They also lend themselves to numerical integration in some cases where Liouville's theorem is important, as in ion and electron optics. *See* CANONICAL COORDINATES AND TRANSFORMATIONS.

The Hamiltonian function H of classical mechanics is used in forming the quantum-mechanical Hamiltonian operator for systems having a classical analog. *See* NONRELATIVISTIC QUANTUM THEORY.

John L. Safko; Philip Stehle

Bibliography. V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 2d ed., 1989; H. C. Corben and P. Stehle, *Classical Mechanics*, 2d ed., 1960, reprint, Dover, New York, 1994; H. Goldstein, C. P. Poole, Jr., and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; I. Percival and D. Richards, *Introduction to Dynamics*, 1982.

Hamilton's principle

Hamilton's principle obtains Lagrange's equations by variations about possible motions of the system between two times (that is, it is an integral principle) rather than changes in the instantaneous state (which characteristic of a differential principle). *See* LAGRANGE'S EQUATIONS.

The meaning that is attached to the phrase "motion of the system between two times" (t_1 and t_2) needs explanation. The problem of determining how a system moves may be formulated in the following way: If the configuration of the system at time t_1 is specified by the generalized coordinates $q_1(t_1), \ldots, q_f(t_1)$ and at the time t_2 by $q_1(t_2), \ldots, q_f(t_2)$ [where f is the number of degrees of freedom of the system], then it is required to find the trajectory along which the system travels from the initial to the final configuration. Hamilton's principle addresses this problem similarly to the way that a geometer addresses the problem of finding the shortest path lying in a curved surface between two given points on the surface. The geometer specifies the distance ds between any two close-lying points in terms of the coordinates q_i of the two points and their differences, the coordinate differentials dq_i. Hamilton defined a characteristic function Φ,

$$\Phi = \int_{t_1}^{t_2} L(q, \dot{q}, t)\, dt$$

using the Lagrangian function $L(q, \dot{q}, t)$, called the Lagrangian of the system, in a way analogous to the geometer's metric. Hamilton's principle states that the system follows the trajectory that makes the integral in the equation above have a minimum value, provided the time interval between times t_1 and t_2 is not too great. This principle produces Lagrange's equations of motion for the system. Either variational principle (Lagrange's or Hamilton's) produces the same Lagrange's equations. *See* DEGREE OF FREEDOM (MECHANICS); DIFFERENTIAL GEOMETRY; LAGRANGIAN FUNCTION.

Hamilton's principle should not be confused with the Hamiltonian or Hamilton's equations of motion, discussed elsewhere. *See* CANONICAL COORDINATES AND TRANSFORMATIONS; HAMILTON-JACOBI THEORY; HAMILTON'S EQUATIONS OF MOTION.

Philip Stehle; John L. Safko

Bibliography. T. L. Chow, *Classical Mechanics*, 1995; H. C. Corben and P. Stehle, *Classical Mechanics*, 2d ed., 1960, reprint, Dover, New York, 1994; H. Goldstein, C. P. Poole, Jr. and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; D. Hestenes, *New Foundations for Classical Mechanics*, 2d ed., Kluwer, Dordrecht, 1999; W. Yourgrau and S. Mandelstam, *Variational Principles in Dynamics and Quantum Theory*, 3d ed., 1968, reprint 1979.

Hamster

The common name for any of 14 species of rodents in the family Cricetidae. The natural range of most of these species is Asia but a few, such as the common hamster (*Cricetus cricetus*), are found in Europe.

The common hamster is a solitary, aggressive animal with interesting burrowing and hoarding habits (see **illus.**). The burrows are not deep, rarely more than 2 ft (0.6 m), and consist of a large central chamber with radiating side chambers for special purposes, such as for hoarding food, for living quarters, and for excretion. Hamsters are clean animals and avoid soiling their living quarters. A so-called summer chamber is used for breeding. A single animal may hoard as much as 200 lb (90 kg) of roots, seeds, nuts, and various tubers, and each type of food is

The hamster, a pugnacious and solitary rodent.

stored in a separate chamber. The hamster goes into its burrow in the autumn, closes off the entrance, and goes to sleep. It does not hibernate deeply and wakes up from time to time to eat. *See* HIBERNATION AND ESTIVATION.

The hamster uses its incisors to carry large food materials, such as carrots, while small foods, such as corn or nuts, are carried in the large cheek pouches. Foraging is always nocturnal. This animal also hunts lizards, birds, mice, insects, and snakes. The hamster has 16 teeth and the dental formula is I 1/1 C 0/0 Pm 0/0 M 3/3.

This rodent has scent glands on each flank, and by rubbing any structure with its flanks, it marks a territory as a warning to other hamsters. When mating, the male invades the territory of the female after first mixing his scent with hers. After mating, the female drives the male away and raises the young herself.

The golden hamster (*Mesocricetus auratus*) is a closely related species, first known from one specimen found in 1839; it was not seen again until 1930, when a family of 13 animals was dug up in Syria. Since that time, it has not been found alive in the wild state. This species makes a good pet and is extensively used as a laboratory animal for experimental purposes, especially in studies of the physiological aspects of hibernation. Prior to hibernation, changes occur in many animals in which the body fats change chemically and form what is termed brown fat. It has been postulated that brown fat is related to the rapid production of heat during arousal. The gestation period is only 15 days and adulthood is attained in about 11 weeks. Other species of hamsters are *Calomyscus bailwardi*, the mouselike hamster found in Asia, and nine species of *Cricetelus*, the Eurasian hamsters found in Europe and western Asia. *See* RODENTIA; SCENT GLAND; TERRITORIALITY.

Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999.

Haplopoda

An order of carnivorous branchiopod crustaceans formerly included in the order Cladocera. Only one species, the fresh-water *Leptodora kindti*, is included.

The body is about 9 mm (0.4 in.) long. *Leptodora* is among the most transparent of multicellular freshwater animals. It swims slowly by means of enormous antennae and seizes its prey with its six pairs of segmented, grasping, thoracic limbs. The mandibles are styliform and of a type unique within the Branchiopoda. There is a single, median sessile eye. The carapace is reduced to a dorsal brood pouch and does not protect the body.

Parthenogenetic eggs and young are carried in the brood pouch in summer. In autumn, fertilized resting eggs are carried there which are shed freely, and the eggs overwinter. They hatch in spring as nauplii, a stage eliminated from the parthenogenetic phase of the life cycle.

Leptodora is widely distributed in the plankton of larger lakes of the Holarctic region. *See* BRANCHIOPODA.

Geoffrey Fryer

Haplosclerida

An order of sponges of the class Demospongiae, including species with a skeleton made up of a single category of siliceous megascleres embedded in spongin fibers or joined together in a network by a spongin cement. The megascleres are usually diactinal, sometimes monactinal. Microscleres (never asters) are present or absent. A modified dermal skeleton

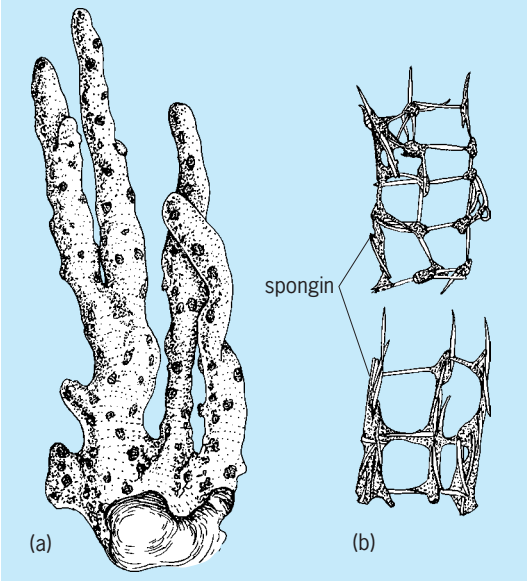

Fig. 1. *Haliclona* morphology. (*a*) Whole sponge. (*b*) Portions of the skeleton of two specimens of *H. oculata* showing varying amounts of spongin joining the spicules.

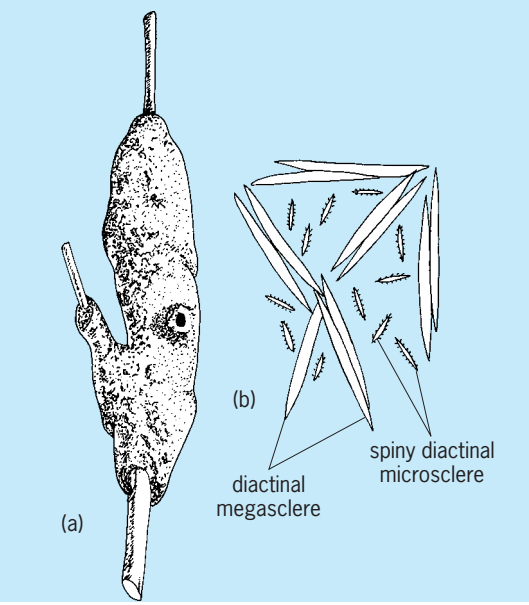

Fig. 2. Spongillid morphology. (*a*) An encrusting sponge growing on a twig. (*b*) Spiculation of *Spongilla lacustris*.

is absent. Several genera lack spicules and have a skeleton of spongin fibers only. Species of this order are encrusting, massive, or lobate, and many species form large upright branching colonies (**Fig. 1**), the branches often being hollow.

Haplosclerid sponges inhabit all seas and are especially abundant in tidal and shallow waters of the continental shelf. Some species occur down to depths of 6600 ft (2000 m).

The family Spongillidae is restricted to fresh water except for a few species which have secondarily invaded brackish water. Spongillids (**Fig. 2**) are encrusting, massive, or branching in form and typically inhabit the edges of streams, ponds, and lakes where light shade is present. They are chiefly gray, brown, or white in color. One species is bright yellow and many are green from the occurrence of zoochlorellae in the cells.

Fossil spongillids are known from the Jurassic Period. Spongillid spicules commonly occur in Pleistocene lake sediments. *See* DEMOSPONGIAE.

Willard D. Hartman

Haplosporea

A class of protozoa, often known as Acnidosporidia, in the subphylum Sporozoa. Haplosporea are distinguished from other similar groups by the production of spores lacking polar filaments. The spores are enclosed in a membrane, and each spore contains a single sporozoite.

Two orders make up the class according to the scheme proposed by the Committee on Taxonomy and Taxonomic Problems of the Society of Protozoologists. These are the Haplosporida and Sarcosporida. Haplosporida are parasites of invertebrates and primitive chordates (Ascidia). Certain species found in fish have also been assigned to this order by some investigators. Sarcosporida are muscle parasites of vertebrates, particularly warm-blooded ones. In many respects Sarcosporida resemble fungi, which some investigators have thought them to be. *See* HAPLOSPORIDA; PROTOZOA; SARCOSPORIDA; SPOROZOA.

Reginald D. Manwell

Haplosporida

A group of Protozoa sometimes regarded as an order within the class Haplosporea. The chief distinguishing characteristic of the Haplosporida, and the one from which their name is derived, is the production of uninucleate spores that lack polar capsules and polar filaments. A spore of *Urosporidium fuliginosum,* a haplosporidan parasitizing a polychaete annelid, is shown in **Fig. 1**. The spores of some species have lids.

Taxonomy. The number of species within the order is highly uncertain, partly because some species originally assigned to the Haplosporida have had to be reassigned to other groups after more intensive study, and partly because organisms which have no evident

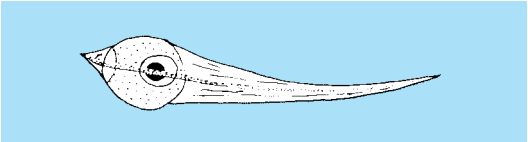

Fig. 1. Spore of *Urosporidium fuliginosum*.

importance often receive little study. M. Caullery, who with F. Mesnil proposed the order in 1899, recognized only 5 genera, which include fewer than 20 valid species. However, he also published a table listing 25 other genera which have been assigned to the order by various authors. Some of them may remain there, but the later research mentioned above resulted in the transfer of certain genera to other groups, for example, the fungi. There has also been disagreement among authorities as to whether existing knowledge justifies the recognition of certain genera as Haplosporida; R. R. Kudo includes several that Caullery rejected or regarded as uncertain.

Life cycle. Very little is known of the life cycles of most Haplosporida. The life cycle of *Icthyosporidium giganteum*, a fish parasite, was worked out many years ago (**Fig. 2**); however, this organism is possibly not a true haplosporidan. It is not in Caullery's list of valid genera, but is included among the Haplosporida by Kudo. The cycle begins with a uninucleate spore that lacks polar capsules and polar filaments. These spores are often of a bizarre shape and represent the infective stage. The spore develops into a multinucleate plasmodium in the coelom or tissues of the host. The multinucleate plasmodium gives rise to uninucleate forms known as sporoblasts, which in turn produce the spores.

Parasitic species. Haplosporida are mainly parasites of invertebrates, such as rotifers, annelids, crustacea, insects, and mollusks, but they also occur in

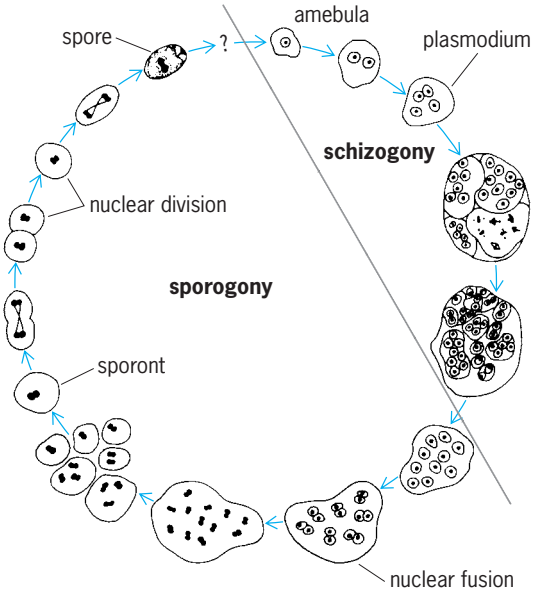

Fig. 2. Stages in the development of *Icthyosporidium giganteum*.

the bodies of Ascidia (tunicates), which are primitive chordates. If the genus *Icthyosporidium* is properly included in the order, Haplosporida also parasitize fish. One species causes "neon" fish disease. A species of *Rhinosporidium* was described from a human, but this genus has since been assigned to the fungi.

It is of considerable biological interest that some species of Haplosporida are hyperparasites, that is, parasitic on other parasites. They have been found in gregarines parasitizing annelids, and also in flukes.

Whether Haplosporida are of any real importance in nature remains unknown. Doubtless, they play some part in the maintenance of biological balances, but evidence suggests that few of them are pathogenic. *See* HAPLOSPOREA; PROTOZOA; SPOROZOA.

Reginald D. Manwell

Bibliography. A. R. Anderson, *Comparative Protozoology*, 1988; M. C. Meyer et al., *Essentials of Parasitology*, 4th ed., 1988; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Harbors and ports

A harbor is a geographic location with a body of water of sufficient depth for vessels to enter to seek rest, obtain supplies, and find shelter from storms or other natural phenomena. Harbors typically are naturally occurring but may be constructed. The modern harbor is a place where vessels are built, launched, anchored, and repaired, and usually contains terminals for incoming and outgoing vessels to transport cargo and people. Harbors typically are used for commercial, naval, and fishery vessels, as well as refuge for small craft. Most harbors that accommodate large vessels are situated at the mouth of a river or at some point where it is easy to transfer cargoes inland by river barges, railroads, or trucks.

Harbors may be classified as three types: landlocked with access to the sea usually by an inlet or channel of limited size; unprotected with significant exposure to the hazards of changing tides and weather-induced phenomena (such as waves, fog, and ice); and artificial. The artificial harbors are fashioned by dredging ship channels and by constructing jetties, breakwaters, and other features to protect and accommodate vessels as necessary.

Ports. The term harbor sometimes is confused with the term port. A port is located usually within a harbor and contains facilities dedicated to the movement of cargo and passengers to another destination. A modern port must be deep enough for large vessels to transit and anchor safely, and must provide a direct channel to open water. A modern port also must have enough room for docks, loading and unloading machinery, warehouses or open spaces, and efficient access and distribution. Many ports are situated within cities with large populations. Consequently, the modern harbor must help ensure that the working port does not deny the population access to the waterfront for recreation.

With the exception of some artificial harbors constructed during wartime, few harbors are entirely human-made. The ideal port is deep enough to accommodate the largest cargo freighters, container ships, tankers, and cruise ships in use. It must have enough room for large ships to maneuver around each other. The bed of the harbor should not be so rocky, sandy, or muddy that anchors cannot be safely dropped and secured. In latitudes where sea ice is present, harbors should be situated so that shipping operations are curtailed as little as possible during the ice season.

Harbors are developed according to prevailing natural conditions (**Fig. 1**). A natural coastal harbor is sheltered from the wind and sea by virtue of its location within a natural coastal indentation or in the protective lee of an island, cape, reef, or other natural barrier. When needed, breakwaters are constructed to provide shelter or supplement inadequate shelter already provided by natural sources. In some coastal harbors, locks or other mechanical devices may be used to provide sufficient water to float vessels at all stages of the tide. Facilities may consist of quays or wharves parallel to the banks of the channel, or piers or jetties that extend into the channel. In a natural river harbor, the waters are not restrained by any artificial means. In some river harbors, slips or basins have been excavated in the banks, obliquely or at right angles to the axis of the stream. As with coastal harbors, river harbors may be equipped with locks or other mechanical devices to provide sufficient water to float vessels at all stages of the tide. A canal or lake harbor is located in the interior portion of a canal or lake that is connected with the sea by a navigable waterway. An open roadstead is a port that has no natural or artificial barrier to provide shelter from the wind, sea, and swell.

Port construction. Among the most important structures at any modern port are breakwaters. These barriers are designed to intercept or deflect the force of moving water. Their main function is to protect the harbor area from severe wave action. Thus, breakwaters provide an artificial means of forming suitable harbors. They also are used to deflect currents and to control silting of channels. They must be strong enough to withstand the impact of waves and should be located to avoid detrimental current or wave deflections. Fine examples of ports protected by breakwaters are Dover, England, and Madras, India. *See* NEARSHORE PROCESSES.

Dredging operations are done at many ports, especially those located at the mouths of rivers to keep the channels free of silt and other obstructions that hamper navigation. The viability of some ports cannot be maintained without dredging. For example, the Port of New York and New Jersey, the third largest port in the United States and one of the largest ports in the world, naturally is about 6 m (19 ft) deep. Virtually all of its channels must be dredged regularly to maintain required depths, ranging from

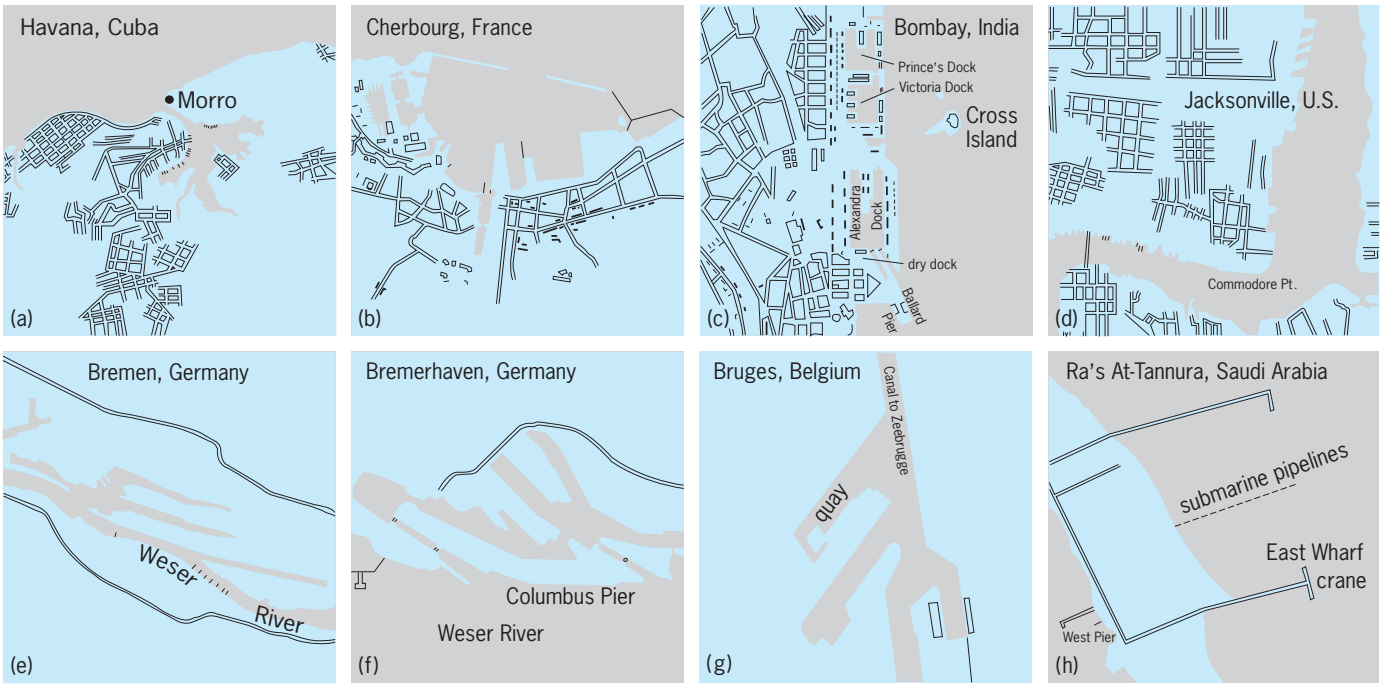

Fig. 1. Examples of different types of harbors. (*a*) Coastal [natural]. (*b*) Coastal [breakwater]. (*c*) Coastal [tide gates]. (*d*) River [natural]. (*e*) River [basins]. (*f*) River [tide gates]. (*g*) Canal or lake. (*h*) Open roadstead. (*U.S. Navy Hydrographic Office*)

10 to 15 m (35 to 50 ft). The large volume of material that is removed from the bottoms of the channels is used for beneficial purposes, ranging from beach nourishment to the upland capping and remediation of brownfield sites (properties where pollutants and contaminants complicate use). *See* RIVER ENGINEERING.

Port efficiency. The amount of cargo moving around the world via ships increases steadily each year. Constructing new ports to accommodate this cargo is neither practical nor inexpensive when such a need must compete with other, more desirable land uses. Therefore, existing ports must improve their efficiency, especially when ports compete with each other for business. The efficiency of a port depends predominantly on the size of the ships that will use it. The largest ports load and unload cargo mainly from container ships and tankers, and must dedicate large areas of land and buildings to store cargo temporarily as it awaits transit to end users. Electric cranes, automatic hoists, fast belt loaders, pneumatic tubes for grains, and rapid pumping facilities for oil, water, and other liquids have been developed to speed loading and unloading. Electronically controlled equipment regulates these operations where previously longshoremen accomplished the same tasks at considerable cost and time. To move cargo out of the port, a port must have both an efficient system of roads and railways. Provision must also be made for customs inspections, banking, and other commercial enterprises that handle the paper work involved in marine insurance, finance, and the preparation of shipping documents. *See* MATERIALS-HANDLING EQUIPMENT.

Cargo security and navigation safety also are prime considerations. The millions of enclosed containers that carry cargo around the world are a major focus of security. Security initiatives undertaken by ports include the hardening of the physical infrastructure, closed-circuit television, early warning systems, security awareness and training, evacuation and emergency planning, stakeholder coordination, cargo and container inspections, and coordination with various levels of government. To help ensure the safe navigation of ships, ports use lighthouses, buoys, beacons, radio direction finders, radar, sonar, geographical positioning systems (GPS), and tugboats with competent harbor pilots to facilitate vessel transit. *See* BUOY; COASTAL ENGINEERING; ELECTRONIC NAVIGATION SYSTEMS; LIGHTHOUSE; PILOTING.

Cargo facilities. General cargo facilities handle nonspecialized cargo such as fruit, coffee, fish, metals, cement, coal, meat, rubber, and wood. Cargo of this nature may be packaged in bags, on pallets, or loosely in large storage areas within the confines of the ship. General cargo is unloaded by a ship's gear (such as derricks, winches, electric cranes, or pneumatic processes) at varying locations along the length of the ship, or with equipment located on the pier. Unloaded cargo is transported from shipside by truck or train to landside temporary storage. While awaiting transport to its final destination, cargo may be stored in open areas or in sheds, depending upon the nature and volume of the cargo as well as the time interval before final transport.

Container terminal facilities. Container terminal facilities move cargo between land and ship in

preloaded standard-sized metal boxes [typically, 12 m (40 ft) long by 2.4 m (8 ft) wide by 2.6 m (8.5 ft) high]. Container terminals require open, paved storage areas adjacent to a ship berth to store containers temporarily while they await loading onto a ship or final shipment to their destination inland. Containerization began in the 1950s with the goal of shipping cargo more rapidly and using fewer personnel. Containers are brought to and from upland holding areas (container yards) by truck or train. They are brought close to the ship, onto which they are individually loaded. A special crane lifts a container from the truck or train vertically and then moves it horizontally to the appropriate storage location aboard ship. Containers are placed first in the hold (interior storage area) and then stacked on the deck of a ship. This process is reversed to unload a container from a ship. The first container ships were able to transport approximately 2000 containers. Today, one ship can move as many as 10,000 containers, and container ships continue to grow in size. *See* MERCHANT SHIP.

A container yard must be organized and managed carefully to ensure that all containers (which can number in the thousands or hundreds of thousands at any given time) can be accessed in the minimum amount of time. Containers also move refrigerated cargo.

Equipment for moving containers within the yard includes straddle carriers and side loaders (**Fig. 2**). Straddle carriers are versatile and provide for the efficient movement and repositioning of containers that will be loaded onto the ship, train, or truck by other equipment. Side loaders move containers primarily from the yard on and off trucks or trains, but are being used increasingly to help optimize temporary yard storage. Ports are seeking to improve the efficiency of handling containers within the yard and are assessing the feasibility of automated systems. *See* MARINE CONTAINERS.

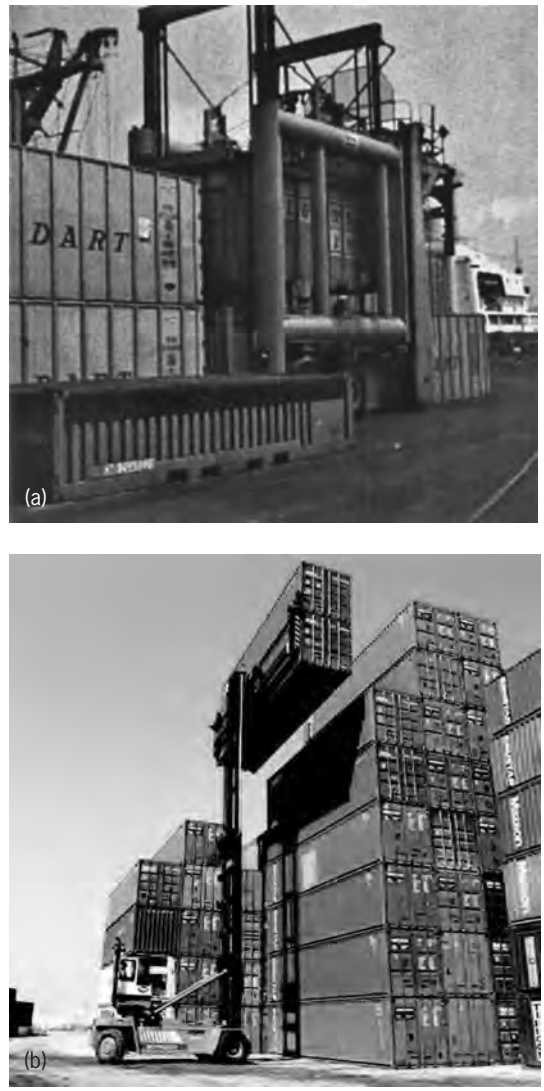

Fig. 2. Examples of container-moving equipment (*a*) Straddle carrier. (*b*) Side loader.

Other types of terminal facilities. General cargo-handling facilities include roll-on/roll-off terminals that accommodate ships that carry cargo in the form of wheeled vehicles. Ferry terminals are designed to provide the efficient transfer of people or vehicles between ferries and high-volume roads, highways, or mass-transit facilities in large urban environments.

Bulk-liquid terminals accommodate vessels transporting liquid cargoes in large volumes, such as crude oil. Bulk liquid is transferred between vessels and storage tanks on land by means of pipelines, with the cargo normally discharged by the ship's pumps. The terminal supplies the energy for pumping the bulk liquid from storage onto the vessel. Since many bulk liquids are flammable or toxic, bulk-liquid terminals are located away from population centers and are equipped with special fire and clean-up installations that are maintained continuously. *See* PIPELINE; PUMP.

Dry bulk terminals are designed to handle cargo such as minerals (for example, iron ore), coal, foodstuffs (for example, grain), and other dry commodities (for example, cement). Depending on the type of commodity, storage is of three basic types: open storage for commodities that are not subject to serious degradation from the elements, sheds for cargo that may be degraded in the short term by the elements, and silos for commodities such as grain or cement where protection from the elements is essential.

Thomas F. Costanzo

Bibliography. H. Agershore, *Planning and Design of Ports and Marine Terminals*, 2004; R. W. Gastil, *Beyond the Edge: New York's New Waterfront*, 2002; J. Gaywaithe, *Design of Marine Facilities for the Berthing, Mooring and Repair of Vessels*, 2004; H. Stevens, *The Institutional Position of Seaports: An International Comparison*, 1999; G. P. Tsinker, *Port Engineering: Planning, Construction, Maintenance and Security*, 2004.

Hardness scales

Arbitrarily defined measures of the resistance of a material to indentation under static or dynamic load, to scratch, abrasion, or wear, or to cutting or drilling. Standardized tests compare similar materials according to the particular aspect of hardness measured by the test. Widely used tests for metals are Brinell, Rockwell, and Scleroscope tests, with modifications depending upon the size or condition of the material. Indentation tests compare species of wood or flooring materials, and abrasion tests serve as an index of performance of stones and paving materials.

Hardness tests are important in research and are widely used for grading, acceptance, and quality control of manufactured articles. The hardness designation or scale is associated with the test method or instrument used.

Scratch hardness. Resistance to scratching is defined by comparison with 10 selected minerals, which are numbered in the order of increasing hardness. This mineralogical scale, called Mohs scale, is 1, talc; 2, gypsum; 3, calcite; 4, fluorite; 5, apatite; 6, orthoclase; 7, quartz; 8, topaz; 9, corundum; and 10, diamond. Minerals lower in the scale are scratched by those with higher numbers. The scale is extended to provide finer distinction of harder materials by additional minerals: 7, vitreous pure silica; 8, quartz; 9, topaz; 10, garnet; 11, fused zirconia; 12, fused alumina; 13, silicon carbide; 14, boron carbide; and 15, diamond.

File-test hardness. Materials are differentiated qualitatively according to resistance to scratching or cutting by files especially selected for the purpose. Whether or not a visible scratch is produced on the material indicates its hardness in comparison with a sample of desired hardness. The method is used for routine inspection of hardened surfaces in production.

Brinell hardness. Resistance to indentation by a hardened steel or tungsten carbide ball under specified load is the basis for Brinell hardness. Standard procedure uses a 10-mm ball with loads of 3000 kg for hard material, 1500 kg for intermediate, and 500 kg or less for soft materials. Various machines apply and control the specified load. The diameter of the impression is measured with a micrometer microscope. Brinell hardness number (Bhn), expressed in kilograms per square millimeter, is obtained by dividing the load by the spherical surface area of the impression. Different-size balls may be used according to size, thickness, and hardness of the specimen, and give the same hardness number provided the loads are proportional to the square of the ball diameter. Carboloy balls are used for very hard material. Time of load application, minimum thickness, and size of specimen are are standardized. A close relation exists between Bhn and ultimate tensile strength.

Vickers hardness. Indentation of a square-based diamond pyramid penetrator with an angle between opposite faces of 136° measures Vickers hardness. Applied load may be varied from 5 to 120 kg in increments of 5 kg according to size of test piece. Vickers hardness number, also called diamond pyramid hardness, is equal to the load divided by the lateral area of the pyramidal impression. The area is computed from measurements of the diagonals of the square impression. Vickers hardness is the most reliable measure for very hard material and is applicable to thin sheets and hardened surfaces.

Rockwell hardness. Depth of indentation of either a steel ball or a 120° conical diamond with rounded point, called a brale, under prescribed load is the basis for Rockwell hardness. The ball is normally $1/\sqrt{16}$ in. in diameter, but $1/8$-, $1/4$-, or $1/2$-in. balls are used for soft materials. A specially designed machine applies loads of 60, 100, or 150 kg. The depth of impression, referred to the position under an initial minor load, is indicated on a dial whose graduations represent the hardness number. Hardness is designated by a number with a standard system of prefix letters to indicate type of penetrator and load used.

Superficial Rockwell hardness is measured by a special machine differing from the standard Rockwell tester in that it applies lighter loads with a more sensitive depth-measuring dial. It produces a shallow impression and is suitable for thin sheet material and where surface hardness is of limited depth.

Monotron hardness. The pressure in kilograms per square millimeter required to embed a 0.75-mm hemispherical diamond penetrator to a depth of 0.0018 in., producing an impression 0.36 mm in diameter, is the measure of Monotron hardness. The depth is controlled by a separate dial graduated to 1 kg/unit area. The method is applicable to the entire range of hardness and is suited to thin sheet and case-hardened surfaces.

Shore Scleroscope hardness. Height of rebound of a diamond-tipped weight or hammer falling within a glass tube from a height of 10 in. and striking the specimen surface measures Shore Scleroscope hardness. The standard hammer is $1/4$ in. in diameter, $3/4$ in. long, and weighs $1/\sqrt{12}$ oz. The hardness number is the height of rebound referred to an arbitrary scale graduated to 140 divisions within the glass tube. The method is a dynamic load test, and the rebound reflects the size of indentation produced, which determines the energy absorbed by deformation and hence that available for rebound. A recording instrument with a dial indicates the rebound hardness directly. Both instruments are portable and permit rapid determinations.

Herbert pendulum hardness. Resistance to cold working is measured as Herbert pendulum hardness. The apparatus consists of a rocking device, called a pendulum, supported on a 1-mm steel ball in contact with the specimen. A curved level bubble measures amplitude of oscillation. The time hardness number is the time in seconds for 10 complete swings of the pendulum through a small arc. The work-hardening capacity is measured by the maximum time hardness after previously repeated single swings of the pendulum. The scale hardness number is the angular oscillation of a half swing after tilting the

pendulum through a definite angle before release. Scale hardness is taken as a measure of resistance to flow as in rolling, drawing, and stamping. The method is applicable to studies of machining and forming of metals.

Microhardness. Resistance to indentation over very small areas (as on small parts, the constituents of metal alloys, or for exploration of hardness variations) is called microhardness. One tester employs the Vickers square-based pyramidal diamond penetrator attached to the end of a vertically guided shaft having a weight of 25 g. An arrangement of microscopes permits centering and measurement of the diagonals of the impression. The hardness number is the pressure intensity in kilograms per square millimeter.

Another procedure employs a Tukon tester applying loads of 25–3600 g using a Knoop indenter, which is a diamond ground to produce a diamond-shaped impression with ratio of diagonal lengths of 7:1. The location of the indenter and measurement of the diagonals of the impression are accomplished with microscopes. The hardness number is the ratio of the applied load to the projected area of the impression.

Hardness of wood. The load required to embed a 0.444-in.-diameter steel ball to half its diameter expresses the hardness of wood. Used as a means of comparison, the values vary with species and grain characteristics. Hardness values of poplar and Douglas-fir are approximately 400 and 900 lb, respectively.

Hardness of paving. Wear or abrasion hardness applies primarily to natural stones, paving or flooring materials. It is measured by a specified test providing an index of service performance. Hardness of stone is reflected by weight loss of a cylindrical core rubbed on a sand bed. Deval abrasion test determines weight loss of a charge of broken stone tumbled in a cylinder. Similarly, the Los Angeles rattler and the standard rattler test for paving block tumble a charge including steel balls in a drum to determine percentage loss of weight as an index of wear. Special wear tests are applied to floor surfaces.

William J. Krefeld/Waldo G. Bowman

Bibliography. H. E. Davis, G. E. Troxell, and F. W. Hauck, *The Testing of Engineering Materials,* 4th ed., 1982; A. A. Ivan'ko, *Handbook of Hardness Data*, 1971; V. K. Sarin (ed.), *Science of Hard Materials*, vol. 3, 1989.

Hardy-Weinberg formula

A basic mathematical relation used in population genetics. It gives the proportion of the various genotypes in a randomly mating population in terms of the frequencies of the genes. The formula is useful for genetic analysis of populations, such as human populations or plants and animals in nature where experimental matings are not possible. It was discovered independently in 1908 by G. H. Hardy, a British mathematician, and W. Weinberg, a German physician who made a number of important contributions to the methodology of human genetics. *See* HUMAN GENETICS; POPULATION GENETICS.

In its simplest form the Hardy-Weinberg formula may be stated thus: If p is the proportion of gene A in the population and $q\ (= 1 - p)$ is the proportion of gene a, then after one generation of random mating the three genotypes AA, Aa, and aa will occur in the proportions p^2, $2pq$, and q^2. In other words the genotypes are given by the appropriate terms in the expansion of the binomial $(p + q)^2$. The extension to multiple alleles is direct.

As a numerical example, the recessive gene a for phenylketonuria, which is a metabolic deficiency resulting in mental retardation, has a frequency in the United States of about 0.01. Therefore the proportions of the three genotypes are:

Homozygous normal	AA	$(0.99)^2$	D 0.9801
Heterozygous normal	Aa	$2(0.99)(0.01)$	D 0.0198
Feeble-minded	aa	$(0.01)^2$	D 0.0001

There are about 200 times as many persons who are heterozygous carriers (Aa) of the gene as there are persons homozygous (aa) for it. This is characteristic of rare recessive factors: almost all of the affected children have normal (heterozygous) parents.

The formula holds only for an infinite population and assumes random mating in the absence of significant mutation pressure or gene transfer between populations. However, it is an accurate approximation in many populations. Random mating in this context means that matings occur without regard to the characters determined by the genes in question or the degree of relationship of the mates, and it is possible for a population to be mating at random with respect to some genes and not for others at the same time. For example, it is appropriate to regard the human population as mating at random for blood group genes, and the data actually show excellent agreement with Hardy-Weinberg predictions. However, the formula would not be expected to hold for genes that determine such characters as skin color or intelligence, which strongly influence the choice of mates.

Gene frequency analysis is a technique for testing genetic hypotheses in randomly mating populations. The Hardy-Weinberg formula or an extension of it is used to predict the frequency of certain types in the population or in the progeny of certain parental combinations, and these are compared with the frequencies actually observed. *See* GENETICS.

James F. Crow

Bibliography. L. M. Cook, *Case Studies in Population Biology*, 1988; D. S. Falconer, *Introduction to Quantitative Genetics*, 4th ed., 1996; E. L. Green, *Genetics and Probability*, 1979; J. Roughgarden, *Theory of Population Genetics and Evolutionary Ecology: An Introduction,* rev. ed., 1987.

Harmonic (periodic phenomena)

A sinusoidal quantity having a frequency that is an integral multiple of the frequency of a periodic quantity to which it is related. *See* MODE OF VIBRATION.

A harmonic series of sounds is one in which the basic frequency of each sound is an integral multiple of some fundamental frequency. The name exists for historical reasons, even though according to the usual mathematical definition such frequencies form an arithmetic series. An ideal string (or air column) can vibrate as a whole or in a number of equal parts, and the respective periods of vibration are proportional to the lengths. These increasingly shorter lengths or periods form a harmonic series. The name came from the harmonious relation of such sounds, and the science of musical acoustics was once called harmonics. Nowadays, it is customary to deal with ratios of frequency rather than ratios of length and, because frequency is the reciprocal of period, the definition of harmonic in acoustics becomes that given here. *See* MUSICAL ACOUSTICS. Robert W. Young

Harmonic motion

A periodic motion that is a sinusoidal function of time. It is often called simple harmonic motion. It is the simplest possible type of vibratory motion. The motion is symmetric about its midpoint, at which the velocity is greatest and the acceleration is zero. At the extreme displacements or turning points, the velocity is zero, and the acceleration is a maximum. The motion is characterized by a unique frequency (without overtones).

Harmonic motion may be present in very simple mechanisms. For example, if a wheel is rotating at constant speed about a fixed axis, the projection on any fixed line of the motion of a point on the wheel is simple harmonic. Harmonic motion may also result from the response of a vibrating system to a periodic—in particular a sinusoidal—force. Harmonic motion is the typical motion of most simple systems that have been displaced from a position of stable equilibrium and then released, provided that the damping is negligible. The motion of a pendulum is approximately simple harmonic for small amplitudes. *See* PENDULUM.

If x represents the displacement measured from the midposition and t the time, then harmonic motion can be described by either of the forms in Eqs. (1). The constants A, B, C, and δ are not all

$$\begin{aligned} x &= A\cos(\omega t) + B\sin(\omega t) \\ x &= C\sin(\omega t - \delta) \end{aligned} \tag{1}$$

independent, but have the relations $A = -C\sin\delta$, $B = C\cos\delta$. The amplitude C represents maximum displacement in one direction from the center (half the total motion between extreme positions); δ is a phase angle whose value depends on the precise instant at which the oscillation was started, or alternatively on the phase of the motion when $t = 0$. These quantities are illustrated in **Fig. 1**.

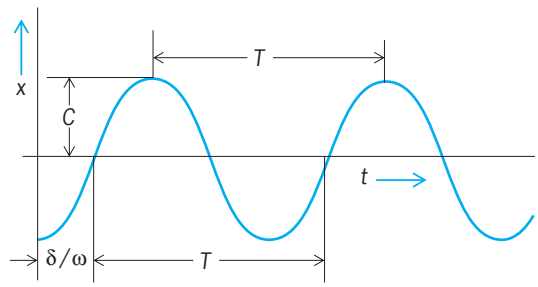

Fig. 1. Representation of simple harmonic motion.

The remaining constant, ω, is known as the angular frequency. Dimensionally, ω is the reciprocal of time. Thus the product ωt is a pure number, to be interpreted as an angle measured in radians. When ωt increases by 2π, the motion repeats. Thus the angular frequency ω is related to ordinary frequency f (number of complete oscillations in unit time) and period T (duration of one complete oscillation) by Eq. (2). The velocity v and acceleration a, obtained

$$\omega = 2\pi f = \frac{2\pi}{T} \tag{2}$$

by differentiating Eq. (1), are given by Eqs. (3) and (4). Because the net force acting on a body is equal to

$$v = \frac{dx}{dt} = \omega C\cos(\omega t - \delta) = \omega\sqrt{C^2 - x^2} \tag{3}$$

$$a = \frac{dv}{dt} = -\omega^2 C\sin(\omega t - \delta) = -\omega^2 x \tag{4}$$

the mass of the body multiplied by its acceleration, Eq. (5) shows that in simple harmonic motion the

$$F = ma = -m\omega^2 x \tag{5}$$

force must be proportional to the displacement.

Conversely, simple harmonic motion occurs whenever, for a body that is displaced from an equilibrium position, there is a net restoring force proportional to the displacement.

Energy considerations. The importance of harmonic motion lies in the simplicity of its time dependence, as given by Eqs. (1), and in the frequent occurrence of linear restoring forces. For sufficiently small displacements of almost any mechanical system from equilibrium, the restoring force or torque is always approximately proportional to the displacement.

This can be explained by some considerations about potential energy. At a point of stable equilibrium, potential energy is necessarily a minimum. If potential energy is a well-behaved function of position, it may be expanded as a series of powers of the distance from any point. (An important exception is electrostatic potential, which varies inversely as the distance from the charge and so is not well-behaved in the immediate neighborhood of the charge.) In the expansion of potential energy about a point of stable equilibrium, the special minimum property guarantees that the linear term must vanish. The next term,

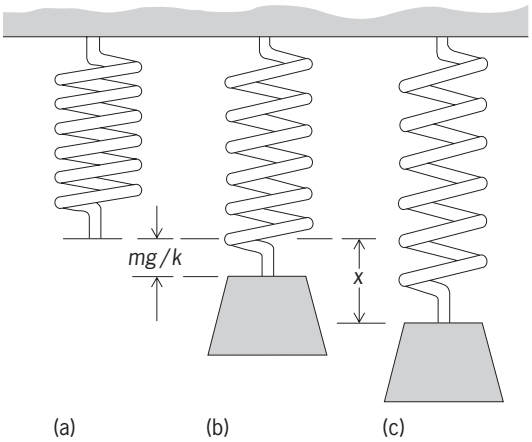

Fig. 2. Weight on an elastic spring. (*a*) Unloaded. (*b*) Statically loaded. (*c*) Weight and spring oscillating.

quadratic in the distance from the equilibrium point, does not ordinarily vanish and provides a good approximation of the potential energy for sufficiently small displacements. A quadratic, or parabolic, variation of potential with distance is equivalent to a force that varies linearly.

In a system oscillating freely about an equilibrium position, the energy changes from potential to kinetic and back again. For simple harmonic motion the potential energy, proportional to the square of the displacement, is greatest at the extremities of the motion. Conversely, the kinetic energy $^1/_2 mv^2$ is zero at the turning points and maximum when the body is going past its equilibrium position. The total energy (kinetic plus potential) is constant. In simple harmonic motion it is proportional to the square of the amplitude of the motion, and the average potential energy just equals the average kinetic energy.

The frequency of a freely oscillating system is determined by the stiffness and inertia of the system. If the motion is harmonic, it is also isochronous, which means that the frequency is independent of the amplitude of the motion.

Weight on elastic spring. If an elastic spring is stretched a distance x beyond its natural length, or compressed a distance x short of its natural length, it exerts a restoring force equal to $-kx$. The stiffness of the spring is measured by the spring constant k. Consider a spring which is suspended vertically and set into vertical oscillations, with a mass m attached to its free end (**Fig. 2**). If m is large compared to the mass of the spring, the latter can be neglected.

The equation of motion for the mass m is Eq. (6),

$$mg - kx = ma = m\frac{d^2x}{dt^2} \tag{6}$$

where g is the acceleration of gravity. The solution of this equation can be written in the form of Eq. (7).

$$x = \frac{mg}{k} + C\sin(\omega t - \delta) \qquad \omega^2 = \frac{k}{m} \tag{7}$$

The term mg/k is simply the extension of the spring due to the static weight, and marks the loaded equilibrium position. Displacement from this equilibrium position calls forth a linear restoring force, and the resulting motion is simple harmonic.

Potential energy (PE) for this example is the sum of an elastic energy $^1/_2 kx^2$ and a gravitational energy $-mgx$. Then Eq. (8) may be written, showing the

$$\begin{aligned} \text{PE} &= {}^1/_2 kx^2 - mgx \\ &= {}^1/_2 k\left(x - \frac{mg}{k}\right)^2 - \frac{m^2g^2}{2k} \end{aligned} \tag{8}$$

quadratic dependence on displacement from equilibrium. The constant term $-m^2g^2/2k$ is of no importance to the motion; it can always be removed by a new choice for the zero of potential energy.

Rotational harmonic motion. Rotational as well as translational simple harmonic motion may occur. In this case the angular displacement is a sinusoidal function of time. A free angular vibration will be simple harmonic if the angular displacement from an equilibrium orientation produces a restoring torque proportional to the displacement. An example is the torsional pendulum (**Fig. 3**). One end of a torsionally flexible elastic rod is held fixed. To the other end is fastened a disk, or another body of large moment of inertia. If the rod is twisted and then released, the rod and disk will undergo angular simple harmonic motion, provided that the torque in the rod is proportional to the angle of twist.

Atomic vibrations. The realization that atoms are continually vibrating in motions that are nearly harmonic is essential for understanding many properties of matter, including molecular spectra, heat capacity, and heat conduction.

In a diatomic molecule, the distance between the atoms is not precisely fixed. There is an equilibrium

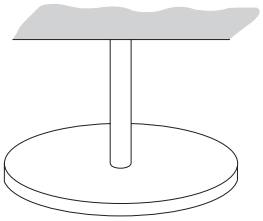

Fig. 3. Torsional pendulum.

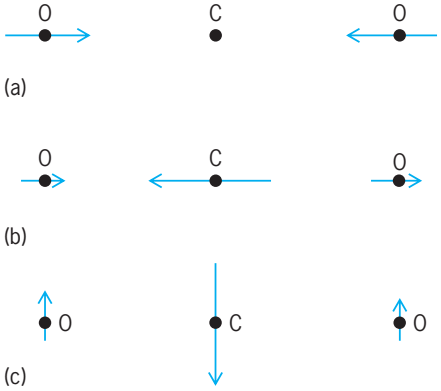

Fig. 4. Three normal modes of the CO_2 molecule. (*a*) Lower-frequency in-line mode. (*b*) Higher-frequency in-line mode. (*c*) Bending mode.

separation corresponding to a minimum in potential energy. The actual magnitude of the interatomic distance oscillates about the equilibrium distance with a motion that is approximately simple harmonic, if the energy of oscillation is small.

In a polyatomic molecule or in a crystalline solid, the atoms also vibrate about equilibrium positions. Because of the large number of degrees of freedom, however, the situation is more complicated. For example, in the carbon dioxide (CO_2) molecule neither oxygen atom by itself moves harmonically, or even periodically. On the other hand, the motion of the CO_2 molecule can be analyzed into a number of independent motions, called normal modes, each of which is by itself simple harmonic. In one such motion, for example, the two oxygen atoms move in phase toward or away from the carbon atom (**Fig. 4***a*). The actual motion of the atoms in a molecule is a superposition of the various normal modes and a rotation of the molecule as a whole. *See* DAMPING; FORCED OSCILLATION; HARMONIC OSCILLATOR; LATTICE VIBRATIONS; MOLECULAR STRUCTURE AND SPECTRA; PERIODIC MOTION; VIBRATION; WAVE MOTION.

Joseph M. Keller

Bibliography. R. Baierlain, *Newtonian Dynamics*, 1983; V. Barger and M. G. Olsson, *Classical Mechanics*, 2d ed., 1995; K. R. Symon, *Mechanics*, 3d ed., 1971.

Harmonic oscillator

Any physical system that is bound to a position of stable equilibrium by a restoring force or torque proportional to the linear or angular displacement from this position. If such a body is disturbed from its equilibrium position and released, and if damping can be neglected, the resulting vibration will be simple harmonic motion, with no overtones. The frequency of vibration is the natural frequency of the oscillator, determined by its inertia (mass) and the stiffness of its restoring force.

The harmonic oscillator is not restricted to a mechanical system, but might, for example, be electric. Typical electronic oscillators, however, are only approximately harmonic.

If a harmonic oscillator, instead of vibrating freely, is driven by a periodic force, it will vibrate harmonically with the period of the force; initially the natural frequency will also be present, but any damping will eventually remove the natural motion. The response of a harmonic oscillator driven by a general force $f(t)$, an arbitrary function of time, is described by a linear differential equation. If $x(t)$ represents the displacement as a function of time t, then the equation of motion is Eq. (1). The left side of Eq. (1) represents the

$$m\frac{d^2x}{dt^2} = -c\frac{dx}{dt} - kx + f(t) \qquad (1)$$

mass m multiplied by the acceleration. The force on the right side of the equation includes, in addition to the restoring force $-kx$ and the driving force, a "viscous damping" force $-c\ dx/dt$ that is proportional to the velocity. (This damping may vanish.) The explicit solution of Eq. (1) is Eq. (2), where $b = c/2m$,

$$x = e^{-bt}\left\{ A\sin(\omega t - \delta) + \frac{e^{-bt}}{m\omega}\int_{-\infty}^{t} e^{bT}\sin[\omega(t-T)]f(T)\,dT \right\} \qquad (2)$$

$\omega^2 = (k/m) - (c/2m)^2$, and A and δ are arbitrary constants which can be set equal to zero if the oscillator displacement and velocity were zero before application of the force $f(t)$. The variable of integration T is the time at which the force $f(t)$ is considered to act. *See* DAMPING; FORCED OSCILLATION; HARMONIC MOTION.

In both quantum mechanics and classical mechanics, the harmonic oscillator is an important problem. It is one of the few rigorously soluble problems of quantum mechanics. The quantum mechanical description of electromagnetic, electronic, mesonic, and other fields is usually carried out in terms of a (time) Fourier analysis. The individual Fourier components of noninteracting fields are independent harmonic oscillators.

The hamiltonian for a harmonic oscillator is Eqs. (3) where ω is the angular frequency character-

$$H = \frac{p^2}{2m} + \frac{m\omega^2 q^2}{2} \qquad (3a)$$

$$= \frac{P^2}{2} + \frac{1}{2}\omega^2 Q^2 \qquad (3b)$$

istic of the oscillator. The mass m in Eq. (3*a*) is made to disappear by replacing the coordinate q and the conjugate momentum p by $Q = m^{1/2}q$ and $P = m^{-1/2}p$ in Eq. (3*b*). The corresponding Schrödinger equation for the wave function ψ is then Eq. (4), where E is

$$-\frac{\hbar^2}{2}\frac{d^2\psi}{dQ^2} + \frac{\omega^2 Q^2}{2}\psi = E\psi \qquad (4)$$

the energy, and $\hbar$ is Planck's constant divided by 2π. This equation possesses quadratically integrable solutions only for the characteristic energy values as shown by Eq. (5), where n is any positive integer.

$$E = (n + {}^1/_2)\hbar\omega \qquad (5)$$

For these cases, ψ can be expressed in terms of the Hermite polynomial $H_n(y)$ of degree n, as in Eq. (6) where $y = (\omega/\hbar)^{1/2}Q$ is dimensionless, and Eq. (7) applies.

$$\psi_n = C_n e^{(-y^2/2)}H_n(y) \qquad (6)$$

$$C_n^2 = \left(\frac{\omega}{\pi\hbar}\right)^{1/2}\frac{1}{2^n n!} \qquad (7)$$

See ANHARMONIC OSCILLATOR; LAGRANGE'S EQUATIONS; NONRELATIVISTIC QUANTUM THEORY.

Joseph M. Keller

Bibliography. R. Baierlein, *Newtonian Dynamics,*1983; R. M. Eisberg, *Fundamentals of Modern Physics*, 1961; R. M. Eisberg and R. Resnick, *Quantum Physics*, 2d ed., 1985; H. Goldstein, C. P. Poole,

and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; C. Kittel, W. D. Knight, and M. A. Ruderman, *Mechanics*, Berkeley Physics Course, vol. 1, 2d ed., 1973; J. W. S. Rayleigh, *The Theory of Sound*, 2d ed., 1894, reprint 1945.

Harmonic speed changer

A mechanical-drive system used to transmit rotary, linear, or angular motion at high ratios and with positive motion. In the rotary version (**Fig. 1**), the drive consists of a rigid circular spline, an input wave generator, and a flexible spline. Any one of these can be fixed, used as the input, or used as the output. Any combination (fixed, driver, or driven) may be used. *See* SPLINES.

In Fig. 1 the fixed member is a rigid circular spline with 132 internal teeth. The driven part is a flexible spline, a ring gear with 130 external teeth of same size as on the rigid spline. The driving member is shown as a two-lobed member generating a traveling circular wave on the flexible spline.

The flexible spline meshes with the rigid circular spline at two diametrically opposite regions of the spline, which is shown flexed as an ellipse with its major axis nearly vertical. The teeth on the splines clear each other along the nearly horizontal minor axis.

As the wave generator rotates, the axes of the flexible spline correspondingly rotate along with the regions of contact and clearance between the teeth on the splines. Because the number of teeth on the flexible spline is less than the number of teeth on the fixed and rigid circular spline, when the wave generator rotates one turn, the flexible spline will rotate $^{2}/_{\sqrt{132}}$ parts of a turn in a direction opposite that of the driving wave generator. The speed reduction is therefore 2:132 (or 1:66). By increasing the number of teeth on the two splines, the flexible spline always having two less (with a two-lobed wave generator) than the number on the circular spline, the speed reduction can be increased correspondingly.

In actual construction (**Fig. 2**), the wave generator is composed of a ball bearing with an elliptical inner race. In rotating, the inner race flexes the outer race, engaging its external teeth along its major axis with the internal teeth of the rigid spline.

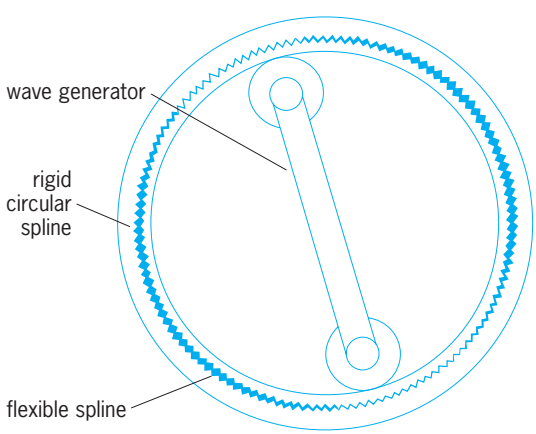

Fig. 1. Rotary-to-rotary harmonic speed changer.

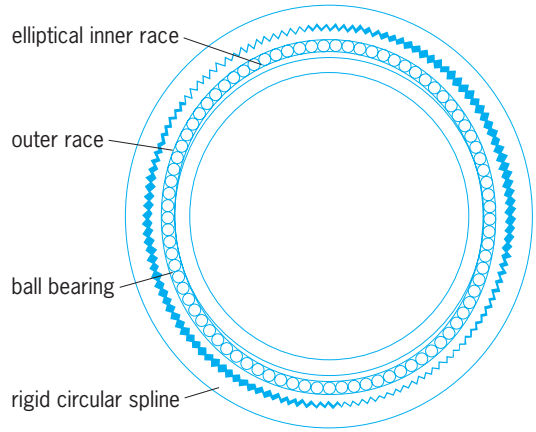

Fig. 2. Diagram of actual construction of a rotary-to-rotary harmonic speed changer.

The advantages of this drive are (1) high-ratio gearing in small space, (2) speed reduction (or increase) in fixed ratio up to 1000:1 in one unit, (3) negligible wear of teeth, (4) balanced bearing loads, (5) negligible backlash, (6) efficiency of approximately 80% in a gear ratio of 400:1, and (7) adaptability to rotary-to-rotary, rotary-to-linear, and linear-to-linear drives.

Paul H. Black

Bibliography. E. A. Dijksman, *Motion Geometry of Mechanism*, 1976; H. H. Mabie and F. W. Ocvirk, *Mechanisms and Dynamics of Machinery*, 4th ed., 1987.

Harpacticoida

An order of the Copepoda derived from the genus *Harpacticus* Milne-Edwards 1840, with about 1700 species known. This group of crustaceans is referred to as Podoplea Copepoda. Their form is variable but is generally linear and more or less cylindrical (see **illus.**). These animals are minute and vary in length from 0.016 to 0.12 in. (0.4 to 3 mm). As a rule, the first thoracic segment is incorporated with the cephalothorax, and the last thoracic segment is included in the abdomen. First antennae in the males are prehensile, while second antennae in both sexes are usually biramous.

The mouthparts of free-living forms are well developed, while in parasitic forms they are reduced or transformed. Maxillipeds as a rule are also prehensile. The first pair of swimming legs generally is modified for grasping. The second to fourth pairs of legs are natatory, with exopodites almost always three-jointed, and endopodites three- to one-jointed or absent. Sexual dimorphism can occur in all legs mentioned. The last pair of legs is reduced and dissimilar in the two sexes. The ovisac may be single or double, but is always ventral. Ova of a few species are laid free. *See* SEXUAL DIMORPHISM.

The nervous system consists of a supra- and a subesophageal ganglion, a circumesophageal ring, and a ventral chain with one pair of ganglia in each

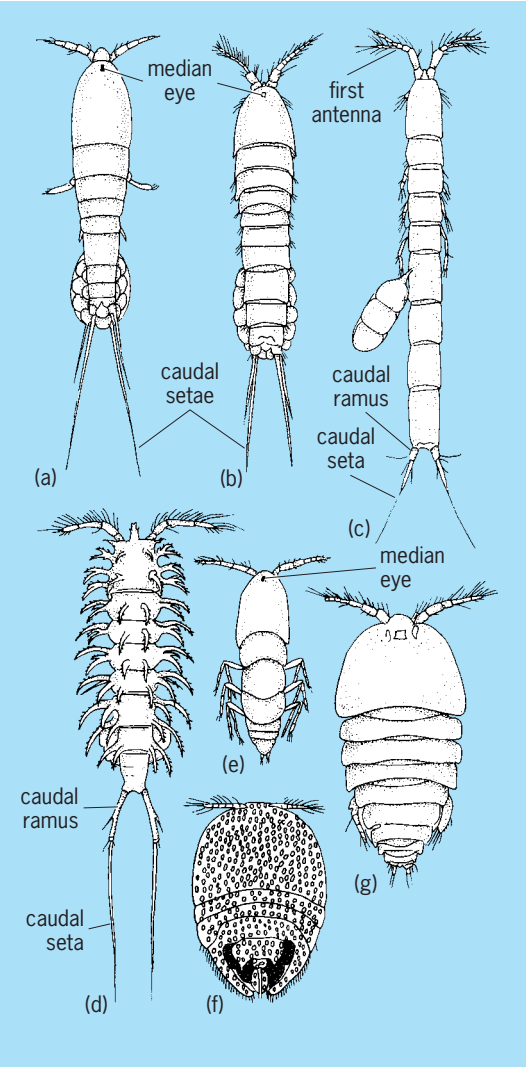

Some typical harpacticoids. (*a*) *Harpacticus chelifer*. (*b*) *Ameira longipes*. (*c*) *Cylindropsyllus laevis*. (*d*) *Anchorabolus mirabilis*. (*e*) *Parategastes spaericus*. (*f*) *Porcellidium ravanae*. (*g*) *Alteutha interrupta*.

segment. The supraesophageal ganglion innervates the eye, the first antennae, and if they are present, the frontal organs. The circumesophageal ring innervates the dorsal musculature of the cephalothorax, the second antennae, and the labrum, while the subesophageal ganglion innervates the rest of the mouthparts. The last-mentioned ganglion is provided with two transverse commissures.

As a rule, a tripartite median eye, the nauplius eye, is present. The parts of the eye, however, may be independent, or the eye may have disappeared. One pair of frontal organs is sometimes found. Antennulae have an olfactory organ, the esthete, on the fourth joint or its equivalent; males often have esthetes on other joints also.

The gonads are single. Their ducts open ventromedially on the first true abdominal segment. Two oviducts open into a seminal receptacle. The vas deferentia are single or paired, and sperm are contained in spermatophores.

The alimentary canal is similar to that in most other copepods. The excretory organ consists of a maxillary gland which is always present, and a rudimentary antennal organ occurs in some forms. Development follows the general plan as it occurs in other copepods. *Elaphoidella bidens* exhibits parthenogenesis, while heterogony is found in *Epactophanes*.

Harpacticoids have a worldwide distribution. They occur in all kinds of aquatic habitats, especially marine, but also among moss and leaves. In the sea they range from the shore to abyssal depths. In general, they are free-living and benthonic; some species are pelagic, parasitic, or commensal. The species may be mono- or polycyclic. A few fresh-water species of Harpacticoida produce resting eggs, or they estivate within cysts. *See* COPEPODA. Karl Lang

Bibliography. K. Lang, *Monographie der Harpacticiden*, 1948; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; G. O. Sars, *Crustacea of Norway*, vol. 5, 1911, vol. 7, 1921.

Hassium

A chemical element, symbol Hs, atomic number 108. It was synthesized and identified in 1984 by using the Universal Linear Accelerator (UNILAC) at Darmstadt, West Germany, by the same team (led by P. Armbruster and G. Müzenberg) which first identified bohrium and meitnerium. The isotope ^{265}Hs was produced in a fusion reaction by bombarding a ^{208}Pb target with a beam of ^{58}Fe projectiles. *See* PERIODIC TABLE.

1																	18
1												13	14	15	16	17	2
H	2																**He**
3	4											5	6	7	8	9	10
Li	**Be**											**B**	**C**	**N**	**O**	**F**	**Ne**
11	12											13	14	15	16	17	18
Na	**Mg**	3	4	5	6	7	8	9	10	11	12	**Al**	**Si**	**P**	**S**	**Cl**	**Ar**
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
K	**Ca**	**Sc**	**Ti**	**V**	**Cr**	**Mn**	**Fe**	**Co**	**Ni**	**Cu**	**Zn**	**Ga**	**Ge**	**As**	**Se**	**Br**	**Kr**
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
Rb	**Sr**	**Y**	**Zr**	**Nb**	**Mo**	**Tc**	**Ru**	**Rh**	**Pd**	**Ag**	**Cd**	**In**	**Sn**	**Sb**	**Te**	**I**	**Xe**
55	56	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86
Cs	**Ba**	**Lu**	**Hf**	**Ta**	**W**	**Re**	**Os**	**Ir**	**Pt**	**Au**	**Hg**	**Tl**	**Pb**	**Bi**	**Po**	**At**	**Rn**
87	88	103	104	105	106	107	108	109	110	111	112	113					
Fr	**Ra**	**Lr**	**Rf**	**Db**	**Sg**	**Bh**	**Hs**	**Mt**	**Ds**	**Rg**							

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

The discovery of bohrium and meitnerium was made by detection of isotopes with odd proton and neutron numbers. In this region, odd-odd nuclei show the highest stability against fission. Elements with an even atomic number are intrinsically less stable against spontaneous fission. The isotopes of hassium were expected to decay by spontaneous fission—which explains why meitnerium was synthesized before hassium.

As in the case of bohrium and meitnerium, the isotope was produced by fusion in a one-neutron de-excitation channel; in this case the compound system was ^{266}Hs. The reaction mechanism again was

cold fusion. The isotope ^{265}Hs has a half-life of about 2 ms and decays by emission of an alpha particle of 10.36 MeV.

The elements bohrium, hassium, and meitnerium are stabilized by shell effects against spontaneous fission; this special stability may occur because these nuclei may prefer a sausage shape which is predicted to be energetically most favorable for them. *See* BOHRIUM; MEITNERIUM; TRANSURANIUM ELEMENTS.

Peter Armbruster

Bibliography. S. Cwiok et al., Fission barriers of transfermium elements, *Nucl. Phys.* A410:254–270, 1983; A. G. Demin et al., On the properties of the element 106 isotopes produced in the reaction Pb + 54Cr, *Z. Phys.*, A315:197–200, 1984; S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; G. Muzenberg et al., The identification of elements 108, *Z. Phys.*, A317:235–236, 1984; G. Muzenberg et al., Observation of the isotopes 264108 and 265108, *Z. Phys.*, A328:49–59, 1987; Y. T. Oganessian et al., On the stability of the nuclei of element 108 with A = 263–265, *Z. Phys.*, A319:215–217, 1984.

Hatchettite

Mountain tallow, a yellow-white to yellow-green hydrocarbon occurring in Belgian coal seams. The material, also called hatchettine, frequently is found in geodes and also in fractures in associated rocks. *See* GEODE.

Hatchettite is translucent but darkens on exposure to air. It is soft, has no odor, is greasy to the touch, and consists of 85.5% carbon and 14.5% hydrogen leading to the empirical formula C_nH_{2n}. Its index of refraction is 1.47–1.50, it melts at 46–47°C (115–117°F), is sparingly soluble in alcohol or ether, decomposes in concentrated sulfuric acid, and has a specific gravity of 0.89–0.98.

Irving A. Breger

Hazardous waste

Any solid, liquid, or gaseous waste materials that, if improperly managed or disposed of, may pose substantial hazards to human health and the environment. Hazardous waste can be generated from many types of processes (**Table 1**). It generally has been considered a subset of solid waste and has been distinguished from municipal wastes and nonhazardous industrial wastes. This characterization is not based on science and engineering but on the regulatory history of hazardous wastes. For example, in the United States in the mid-1970s, statutes existed to control air and water pollution, as well as the design of landfills and other facilities to address solid wastes; hazardous waste, however, was not being addressed. As a result, the U.S. Congress enacted the Resource Conservation and Recovery Act (RCRA) in 1976. The law's primary goals are to protect human health and the environment from the potential hazards of waste disposal, to conserve energy and natural resources, to reduce the amount of waste generated, and to ensure that wastes are managed in an environmentally sound manner. The RCRA regulations are found in the Code of Federal Regulations at Title 40, Parts 260 through 280. In fact, RCRA is an amendment to the Solid Waste Disposal Act of 1956. The 1984 amendments to RCRA are known as the Hazardous and Solid Waste Amendments (HSWA). Subtitle C (hazardous waste) and Subtitle D (solid, primarily nonhazardous, waste) provide the structure for comprehensive waste management programs. In addition, RCRA regulates underground storage tanks under Subtitle I and medical waste under Subtitle J. Many other countries have similar programs to control hazardous wastes. Of the 13 billion tons of industrial, agricultural, commercial, and household wastes generated annually in the United States, 2% (more than 279 million tons) are hazardous as defined by RCRA regulations.

Properties. Hazardous wastes may be of any phase—solid, liquid, gas, or mixture. The hazard of the waste can result from its inherent properties, such as its physicochemical properties, including its likelihood to ignite, explode, react, cause irritations, or elicit toxic effects. The inherent properties of hazardous wastes can also be biological, such as that of infectious medical wastes.

Since wastes are transported by fluids (liquids and gases), especially in air and water, the likelihood of

TABLE 1. Typical hazardous wastes generated by selected industries

Waste generator	Types of wastes produced
Chemical manufacturers	Strong acids and bases Reactive wastes Ignitable wastes Discarded commercial chemical products
Vehicle maintenance shops	Paint wastes Ignitable wastes Spent solvents Acids and bases
Printing industry	Photography waste with heavy metals Heavy-metal solutions Waste inks Spent solvents
Paper industry	Ignitable wastes Corrosive wastes Ink wastes, including solvents and metals
Construction industry	Ignitable wastes Paint wastes Spent solvents Strong acids and bases
Cleaning agents and cosmetic manufacturing	Heavy-metal dusts and sludges Ignitable wastes Solvents Strong acids and bases
Furniture and wood manufacturing and refinishing	Ignitable wastes Spent solvents Paint wastes
Metal manufacturing	Paint wastes containing heavy metals Strong acids and bases Cyanide wastes Sludges containing heavy metals

SOURCE: U.S. Environmental Protection Agency, *RCRA: Reducing Risks from Wastes*, Rep. EPA530-K-97-004, September 1997.

TABLE 2. Densities of some important environmental fluids

Fluid	Density (kg m^{-3}) at 20°C unless otherwise noted
Air at standard temperature and pressure (STP) = 0°C and 101.3 N m^{-2}	1.29
Air at 21°C	1.20
Ammonia	602
Diethyl ether	740
Ethanol	790
Acetone	791
Gasoline	700
Kerosene	820
Turpentine	870
Benzene	879
Pure water	1000
Seawater	1025
Carbon disulfide	1274
Chloroform	1489
Tetrachloromethane (carbon tetrachloride)	1595
Lead	11,340
Mercury	13,600

SOURCE: D. A. Vallero, *Environmental Contaminants: Assessment and Control*, Elsevier Academic Press, Burlington, MA, 2004.

contamination strongly depends on physical properties, especially solubility, density, and vapor pressure. If the substances composing the hazardous waste are quite soluble in water, they are hydrophilic. If a substance is not easily dissolved in water, it is hydrophobic. Since many contaminants are organic (consisting of molecules containing carbon-to-carbon bonds and/or carbon-to-hydrogen bonds), the solubility can be further differentiated as to whether the substance is easily dissolved in organic solvents. Such substance are said to be lipophilic. A substance that is not easily dissolved in organic solvents is said to be lipophobic. *See* SOLVENT; VAPOR PRESSURE.

Density is a very important fluid property for environmental situations. For example, in responding to a leak or release of a hazardous substance, the density of the substance must be known. If a substance is burning, whether it is of greater or lesser density than water will be one of the factors on how to extinguish the fire. If the substance is less dense than water, the water will likely settle below the substance layer, making water a poor choice for fighting the fire. So, any flammable substance with a density less than water (**Table 2**), such as benzene or acetone, will require other fire-extinguishing substances. For substances heavier than water, such as carbon disulfide, water may be the proper fire-fighting material. *See* DENSITY.

Ground-water contamination. One of the chief concerns from hazardous wastes is the potential to contaminate ground water, which can be affected by any type of hazardous waste from numerous sources. When a fluid that is dense and miscible (that is, able to be mixed in any concentration without separation of physical phases) seeps underground through the vadose zone, below the water table and into the zone of saturation, the dense contaminants move downward (**Fig. 1**). When these contaminants reach the bottom of the aquifer, the bottom shape dictates their continued movement, along with the slope of the underlying bedrock or other relatively impervious layer. Solution and dispersion near the

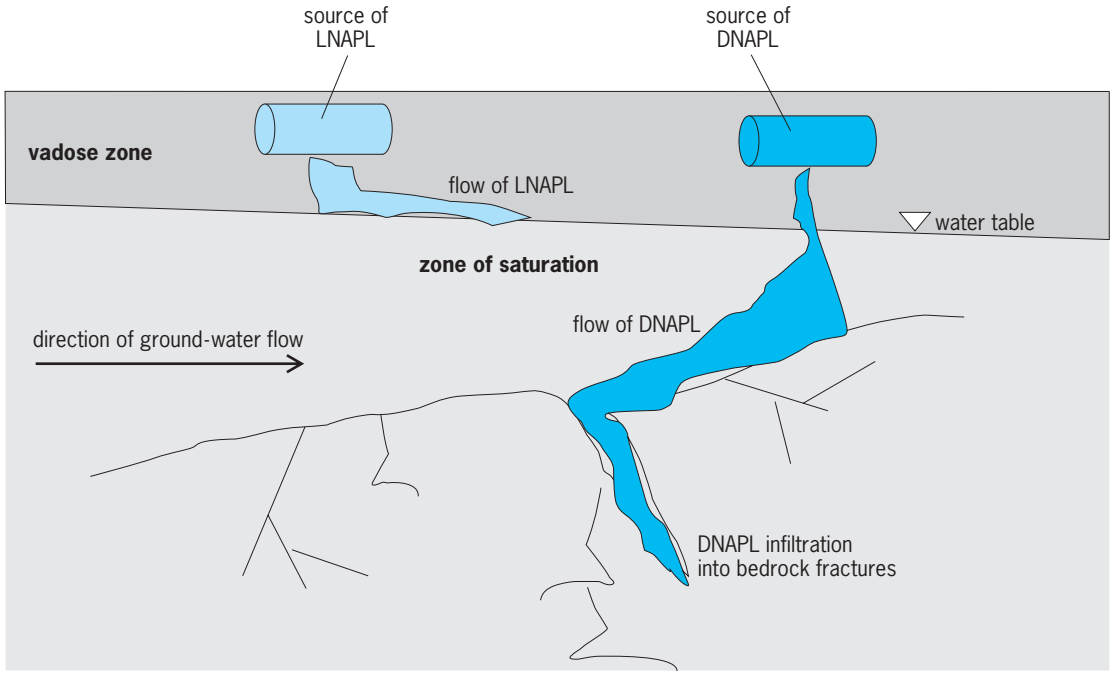

Fig. 1. Importance of density of fluids in ground-water contamination scenarios. The dense nonaqueous phase liquids (DNAPLs) can penetrate more deeply into the aquifer than do the light nonaqueous phase liquids (LNAPLs). The density reference for whether a compound is a DNAPL or an LNAPL is whether it is denser or lighter than water, respectively. The DNAPL may even be against the general flow of the ground water. (*After D. A. Vallero, Environmental Contaminants: Assessment and Control, Elsevier Academic Press, 2004; H. Hemond and E. Fechner-Levy, Chemical Fate and Transport in the Environment, Academic Press, 2000*)

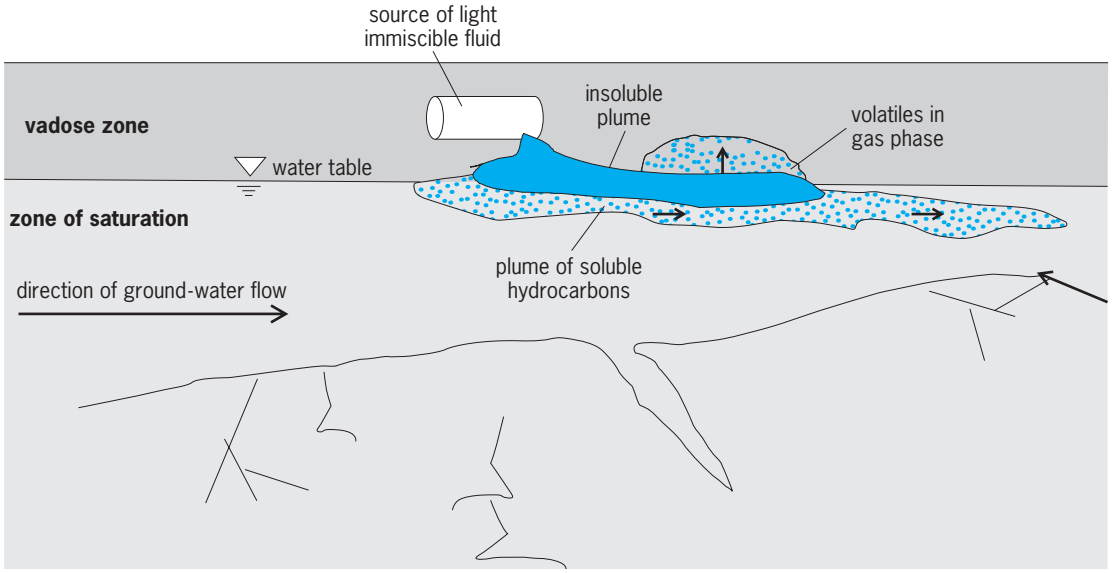

Fig. 2. Hypothetical plume of hydrophobic fluid. (*After D. A. Vallero, Environmental Contaminants: Assessment and Control, Elsevier Academic Press, 2004; and M. N. Sara, Groundwater Monitoring System Design, in Practical Handbook of Ground-Water Monitoring, D. M. ed. by Nielsen, Lewis Publishers, 1991*)

boundaries of the plume will generate a secondary plume that will generally follow the general direction of ground-water flow (**Fig. 2**). Organics more dense than water (DNAPLs) will penetrate more deeply, while the lighter organics (light nonaqueous-phase liquids, or LNAPLs) will float near the top of the zone of saturation.

The physics of this system determines the direction of the contaminant plume's movement. For example, the movement is greatly affected by the solubility and density of the contaminants, but other factors also influence the rate of transport, such as sorption, presence of other solvent besides water, and the amount of organic matter in the soil. In

TABLE 3. Key legal definitions related to hazardous wastes in the United States

Term	Definition
Hazardous waste	A solid waste, or combination of solid wastes, which because of its quantity, concentration, or physical, chemical, or infectious characteristics may (a) cause, or significantly contribute to an increase in mortality or an increase in serious irreversible, or incapacitating reversible, illness; or (b) pose a substantial present or potential hazard to human health or the environment when improperly treated, stored, transported, or disposed of, or otherwise managed.
Medical waste	Any solid waste which is generated in the diagnosis, treatment, or immunization of human beings or animals, in research pertaining thereto, or in the production or testing of biologicals. Such term does not include any hazardous waste identified or listed under subchapter III of this chapter or any household waste as defined in regulations under subchapter III of this chapter. (For "infectious waste" definition, see CFR Part 243 below.)
Mixed waste	Waste that contains both hazardous waste and source, special nuclear, or by-product material subject to the Atomic Energy Act of 1954 (42 U.S.C. 2011 et seq.).
Solid waste	Any garbage, refuse, sludge from a waste treatment plant, water supply treatment plant, or air pollution control facility and other discarded material, including solid, liquid, semisolid, or contained gaseous material resulting from industrial, commercial, mining, and agricultural operations, and from community activities, but does not include solid or dissolved material in domestic sewage, or solid or dissolved materials in irrigation return flows or industrial discharges which are point sources subject to permits under section 1342 of title 33, or source, special nuclear, or by-product material as defined by the Atomic Energy Act of 1954, as amended (68 Stat. 923) (42 U.S.C. 2011 et seq.). (For another definition, see CFR Part 243 & 257 below.)
Transuranic waste	Material contaminated with elements that have an atomic number greater than 92, including neptunium, plutonium, americium, and curium, and that are in concentrations greater than 10 nanocuries per gram, or in such other concentrations as the Nuclear Regulatory Commission may prescribe to protect the public health and safety.
High-level radioactive waste	(A) The highly radioactive material resulting from the reprocessing of spent nuclear fuel, including liquid waste produced directly in reprocessing and any solid material derived from such liquid waste that contains fission products in sufficient concentrations; and (B) other highly radioactive material that the Commission, consistent with existing law, determines by rule requires permanent isolation.
Low-level radioactive waste	Radioactive material that—(A) is not high-level radioactive waste, spent nuclear fuel, transuranic waste, or by-product material as defined in section 2014(e)(2) of this title; and (B) the Commission, consistent with existing law, classifies as low-level radioactive waste.

SOURCE: *U.S. Code, Title 42—The Public Health and Welfare*, Chapter 82, Solid Waste Disposal Subchapter I, General Provisions, §6903; Chapter 23, Development and Control of Atomic Energy, Division A, Atomic Energy Subchapter I, General Provisions, §2014; and Chapter 108, Nuclear Waste Policy, §10101.

Fig. 2, note the importance of vapor pressure. The volatile contaminants (that is, those with relatively high vapor pressure) can move upward from the plume, often reaching the atmosphere. Thus, hazardous wastes can contaminate the air, soil, aquifers, and surface waters. *See* GROUND-WATER HYDROLOGY; WATER POLLUTION.

Hazardous-waste characterization. The concentration and quantity determine whether a waste is hazardous. In most countries, these factors are codified to encourage proper handling of wastes (**Table 3**). In the United States, the amount of hazardous waste generated by a source determines the stringency of the rules. Any source that generates more than 1000 kg (2200 lb) of hazardous wastes per month is considered a large-quantity generator. A large-quantity generator is also one that has produced in any single month or has accumulated at any time 1 kg (2.2 lb) of acutely hazardous waste, or that generates or has accumulated at any time more than 100 kg (220 lb) of spill cleanup material contaminated with acutely hazardous waste. Depending on the hazard of the waste, a source that generates less than 100 kg of waste per month may be conditionally exempted. Other hazardous waste operations that are regulated include those that transport, as well as those that treat, store, and dispose of hazardous wastes.

Hazardous wastes often have been improperly disposed of and managed in the past, creating a need for response and remedial operations to clean up contaminated sites and operations. This has led to extensive environmental and public health threats with associated substantial legal and financial costs to the responsible parties. Cleanup continues to be a challenge to environmental engineers and scientists, but often the most prudent course of action is to eliminate the production of hazardous wastes in the design phase. Thus, pollution prevention, waste minimization, life-cycle analysis, and other "green engineering" programs are increasing. *See* HAZARDOUS-WASTE ENGINEERING. Daniel Vallero

Bibliography. H. M. Freeman (ed.), *Hazardous Waste Minimization*, 1990; H. M. Freeman (ed.), *Standard Handbook of Hazardous Waste Treatment and Disposal*, 2d ed., 1997; H. Hemond and E. Fechner-Levy, *Chemical Fate and Transport in the Environment*, 2000; D. M. Nielsen, *Practical Handbook of Ground-Water Monitoring*, 1991; D. A. Vallero, *Environmental Contaminants: Assessment and Control*, 2004.

Hazardous waste engineering

The control and management of hazardous wastes are truly among the most important challenges of our times. Environmental engineers play crucial roles in reducing the amount of hazardous substances produced, treating hazardous wastes to reduce their toxicity, and applying sound engineering controls to reduce or eliminate exposures to these wastes. Engineers design the facilities that generate the chemicals that, under the wrong circumstances, become hazards. The engineers also design and operate the containment and treatment facilities to deal with the wastes after they have been released, and are frequently called upon to address these wastes once they are in the environment.

A hazard is expressed as the potential of unacceptable outcome (**Table 1**). For chemicals, the most important hazard is the potential for disease or death, which is measured by epidemiologists as morbidity and mortality, respectively. The hazards to human health are referred to collectively in the medical and environmental sciences as toxicity. Toxicology is the study of these health outcomes and their potential causes. *See* EPIDEMIOLOGY; TOXICOLOGY.

In the United States, as in many countries, the term hazardous waste is defined by the federal government. Section 1004(5) of the Resource Conservation and Recovery Act (RCRA) defines a hazardous waste as a solid waste that may pose a substantial present or potential threat to human health and the environment when improperly treated, stored, transported, or otherwise managed.

Toxicity testing. The U.S. Environmental Protection Agency (EPA) has developed standard approaches and has set criteria to determine if substances exhibit hazardous characteristics. Because RCRA defines as hazardous a waste that presents a threat to human health and the environment when it is improperly managed, the government identified a set of assumptions that would allow for a waste to be disposed of if it is not subject to the controls mandated by Subtitle C of RCRA. (*See* HAZARDOUS WASTE.) This mismanagement scenario was designed to simulate a plausible worst case. Under a worst-case scenario, a potentially hazardous waste is assumed to be disposed along with municipal solid waste in a landfill with actively decomposing substances overlying an aquifer. When the government developed the mismanagement scenario, it recognized that not all wastes would be managed in this manner but that a dependable set of assumptions would be needed to ensure that the hazardous waste definition was implemented. The Hazardous and Solid Waste Amendments of 1984 (HSWA) established the Toxicity Characteristic Leaching Procedure (TCLP) to provide replicable results for contaminants commonly found in hazardous waste sites (**Table 2**). The specific medium used in the test is dictated by the alkalinity of the waste. The liquid extracted from the waste is analyzed for the 39 listed toxic constituents, and the concentration of each contaminant is compared to the TCLP standards specific to each contaminant.

Other hazards besides toxicity are important in hazardous waste engineering. The outcome may relate to environmental quality, such as an ecosystem stress, loss of important habitats, and decreases in the size of the population of sensitive species. Outcomes related to public and personal safety are also important. These may include a substance's potential to ignite, its corrosiveness, flammability, or explosiveness. A substance may be a public welfare hazard

TABLE 1. Properties of hazardous wastes

Designation	Description
H1	"Explosive": subtances and preparations which may explode under the effect of flame or which are more sensitive to shocks or friction than dinitrobenzene.
H2	"Oxidising": substances and preparations which exhibit highly exothemic reactions when in contact with other substances, particularly flammable substances.
H3A	"Highly flammable" liquid substances and preparations having a flashpoint of below 21°C (including extremely flammable liquids), or substances and preparations which may become hot and finally catch fire in contact with air at ambient temperature without any application of energy, or solid substances and preparations which may readily catch fire after brief contact with a source of ignition and which continue to burn or to be consumed after removal of the source of ignition, or gaseous substances and preparations which are flammable in air at normal pressure, or substances and preparations which, in contact with water or darnp air, evolve highly flammable gases in dangerous quantities.
H3B	"Flammable": liquid substances and preparations having a flashpoint equal to or greater than 21°C and less than or equal to 55°C
H4	"Irritant": noncorrosive substances and preparations which, through immediate, prolonged, or repeated contact with the skin or mucous membrane, can cause inflammation.
H5	"Harmful": substances and preparations which, if they are inhaled or ingested or if they penetrate the skin, may involve limited health risks.
H6	"Toxic": substances and preparations (including very toxic substances and preparations) which, if they are inhaled or ingested or if they penetrate the skin, may involve serious, acute, or chronic health risks and even death.
H7	"Carcinogenic": substances and preparations which, if they are inhaled or ingested or if they penetrate the skin, may induce cancer or increase its incidence.
H8	"Corrosive": substances and preparations which may destroy living tissue on contact.
H9	"Infectious": substances containing viable microorganisms or their toxins which are known or reliably believed to cause disease in humans or other living organisms.
H10[2]	"Toxic for reproduction": substances and preparations which, if they are inhaled or ingested or if they penetrate the skin, may produce or increase the incidence of nonheritable adverse effects in the progeny and/or of male or female reproductive functions or capacity.
H11	"Mutagenic": substances and preparations which, if they are inhaled or ingested or if they penetrate the skin, may induce hereditary genetic defects or increase their incidence.
H12	Substances and preparations which release toxic or very toxic gases, in contact with water, air, or an acid.
H13	Substances and preparations capable by any means, after disposal, of yielding another substance, e.g., a leachate, which possesses any of the characteristics listed above.
H14	"Ecotoxic": substances and preparations which present or may present immediate or delayed risks for one or more sectors of the environment.

SOURCE: United Kingdom Environment Agency, *Interpretation of the Definition and Classification of Hazardous Waste*, Technical Guidance WM2, Hazardous Waste Directive Annex III, 2003.

that damages property values or physical materials—for example, due to its corrosiveness or acidity. The hazard may be inherent to the substance. But more than likely, the hazard depends upon the situation and conditions where the exposure may occur. The substance is most hazardous when a number of conditions exist simultaneously, such as the hazard to firefighters using water in the presence of oxidizers.

Waste management hierarchy. There is growing acceptance throughout the world to use waste management hierarchies for solving hazardous waste problems. A typical sequence involves source reduction, recycling, treatment, and disposal. The waste generator is encouraged to begin at the top of the hierarchy and proceed down the hierarchy only if necessary. In the United States, the term pollution prevention is often used instead of the term source reduction, and waste minimization is used as an umbrella term for strategies and technologies that use source reduction and recycling techniques (**Fig. 1**).

Source reduction. Source reduction measures include process modifications, feedstock substitutions, improvements in feedstock purity, changes in housekeeping and management practice, increases in the efficiency of equipment, and recycling within a process. In the United States, there is a requirement in the federal environmental regulations that all large hazardous waste generators must have a program at their facility to encourage source reduction and recycling. Other countries, such as Austria, Germany, and Denmark, have initiated direct subsidies to encourage preferable waste management options.

Recycling. This is the use or reuse of hazardous waste as an effective substitute for a commercial product or as an ingredient or feedstock in an industrial process. It includes the reclamation of useful constituent fractions within a waste material or the removal of contaminants from a waste to allow it to be reused. *See* RECYCLING TECHNOLOGY.

Treatment. This refers to any method, technique, or process that changes the physical, chemical, or biological character of any hazardous waste to neutralize it; to recover energy or material resources from the waste; or to render the waste nonhazardous, less hazardous, safer to manage, amenable for recovery, amenable for storage, or reduced in volume.

Disposal. This is the discharge, deposit, injection, dumping, spilling, leaking, or placing of hazardous waste into or on any land or body of water so that the waste or any constituents may enter the air or be discharged into any waters, including ground water.

Treatment technologies. There are various alternative waste treatment technologies, including physical treatment, chemical treatment, biological treatment, incineration, and solidification or stabilization treatment. These processes are used to recycle and reuse waste materials, reduce the volume and toxicity of a waste stream, or produce a final

TABLE 2. Toxicity Characteristic Chemical Constituent Regulatory Levels for 39 Hazardous Chemicals [pursuant to Resource Conservation and Recovery Act of 1976 (RCRA) regulations, 40 CFR 261 and 262.11]

Contaminant	Regulatory level, mg L^{-1}
Arsenic	5.0
Barium	100.0
Cadmium	1.0
Chromium	5.0
Lead	5.0
Mercury	0.2
Selenium	1.0
Silver	5.0
Endrin	0.02
Lindane	0.4
Methoxychlor	10.0
Toxaphene	0.5
2,4-D	10.0
2,4,5 TP (Silvex)	1.0
Benzene	0.5
Carbon tetrachloride	0.5
Chlordane	0.03
Chlorobenzene	100.0
Chloroform	6.0
o-Cresol	200.0
m-Cresol	200.0
p-Cresol	200.0
Cresol	200.0
1,4-Dichlorobenzene	7.5
1,2-Dichloroethane	0.5
1,1-Dichloroethylene	0.7
2,4-Dinitrotoluene	0.13
Heptachlor (and its hydroxide)	0.008
Hexachloroethane	3.0
Hexachlorobutadiene	0.5
Hexachloroethane	3.0
Methyl ethyl ketone	200.0
Nitrobenzene	2.0
Pentachlorophenol	100.0
Pyridine	5.0
Tetrachloroethylene	0.7
Trichloroethylene	0.5
2,4,5-Trichlorophenol	400.0
2,4,6-Trichlorophenol	2.0
Vinyl chloride	0.2

residual material that is suitable for disposal. The selection of the most effective technology depends upon the wastes being treated.

The characteristics of the media in need of treatment determine the performance of any contaminant treatment or control. For example, sediment, sludge, slurries, and soil characteristics that will influence the efficacy of treatment technologies include particle size, solids content, and high contaminant concentration (**Table 3**).

Particle size is an important limiting characteristic for applying treatment technologies to sediments. Most treatment technologies work well on sandy soils and sediments. The presence of fine-grained material reduces the effectiveness of treatment system emission controls because it increases particulate generation during thermal drying, it is more difficult to dewater, and it has greater attraction to the contaminants (especially fine-grained clays). Clayey sediments that are cohesive also present materials-handling problems in most processing systems.

Solids content generally ranges from high (30–60% solids by weight) to low (10–30% solids by weight). Treatment of slurries is better at lower solids contents, but this can be achieved for high solids contents by adding water at the time of processing. It is more difficult to change a lower to a higher solids content, but evaporative and dewatering approaches, such as those used for municipal sludges, may be used. Also, thermal and dehalogenation processes are decreasingly efficient as the solids content is reduced. More water means increased chemical costs and an increased need for wastewater treatment.

Elevated levels of organic compounds or heavy metals in high concentrations must also be considered. Higher total organic carbon (TOC) content favors incineration and oxidation processes. The TOC can be the contaminant of concern or any organic, since they are combustibles with caloric value. Conversely, higher metal concentrations may make a technology less favorable by increasing contaminant mobility of certain metal species following processing.

A number of other factors may affect the selection of a treatment technology other than its effectiveness for treatment (**Table 4**). For example, vitrification and supercritical water oxidation have been used only for relatively small projects and have not been proven for use in full-scale sediment projects. Regulatory compliance and community perception are always a part of decisions regarding an incineration system. Land use considerations, such as the amount of acreage needed, are commonly confronted in solidification and solid-phase bioremediation projects, as well as in sludge farming and land application. Disposing of ash and other residues following treatment must be part of any process. Treating water effluent and air emissions must be part of the decontamination decision-making process.

Physical treatment. This includes processes that separate components of a waste stream or change the physical form of the waste without altering the chemical structure of the constituent materials. Physical treatment techniques are often used to separate the materials within the waste stream so that they can be reused or detoxified by chemical or biological treatment or destroyed by high-temperature incineration. These processes are very useful for separating hazardous materials from an otherwise nonhazardous waste stream so that the materials may be treated in a more concentrated form, separating various hazardous components for different treatment processes, and preparing a waste stream for ultimate destruction in a biological or thermal treatment process.

Physical treatment processes are important to most integrated waste treatment systems regardless of the nature of the waste materials or the ultimate technologies used for treatment or destruction. The physical processes that are commonly used in waste treatment operations include screening, sedimentation, flotation, filtration, centrifugation, dialysis, membrane separations, ultrafiltration, distillation, solvent extraction, evaporation, and adsorption. *See* ACTIVATED CARBON; ADSORPTION; CENTRIFUGATION; DIALYSIS; DISTILLATION; EVAPORATION; FILTRATION; FLOTATION; MEMBRANE SEPARATIONS;

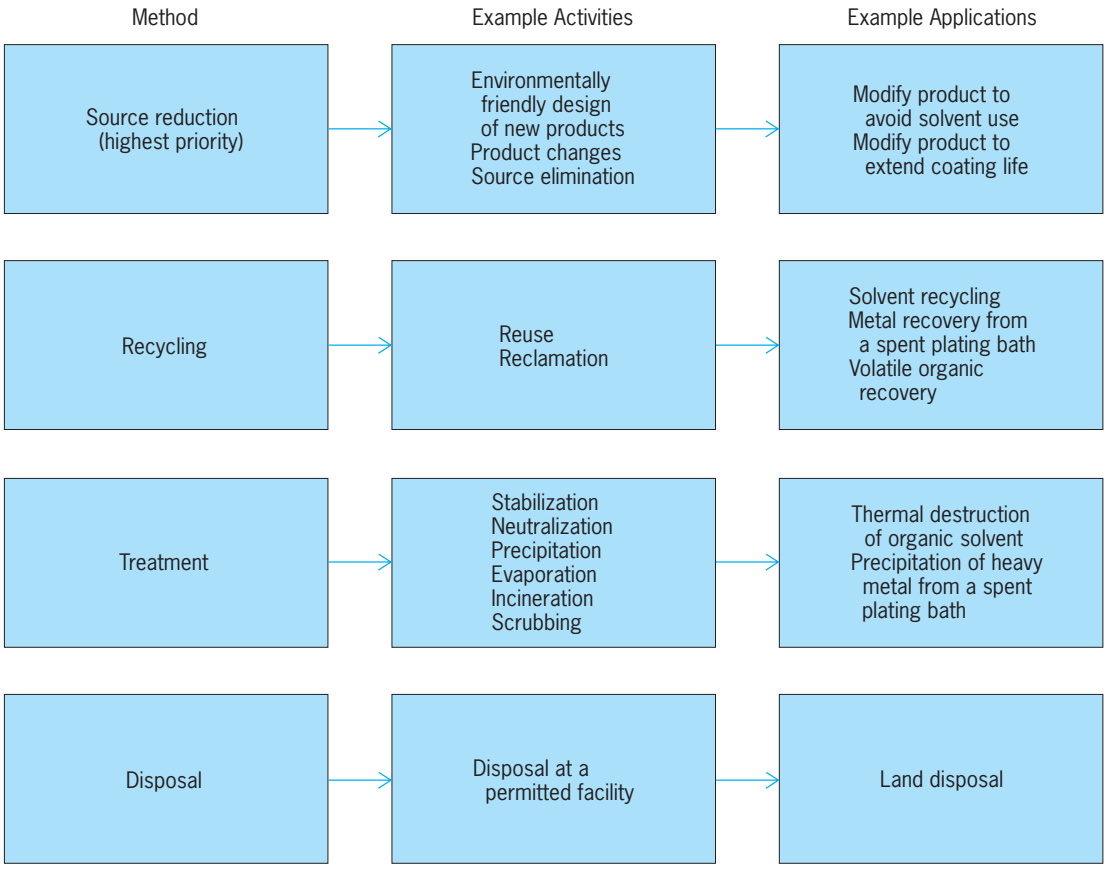

Fig. 1. Flow chart for management of waste: reduction from manufacture, use, treatment, and disposal. (*After U.S Environmental Protection Agency, Facility Pollution Prevention Guide, EPA/600/R-92/088, 1992*)

SCREENING; SEDIMENTATION (INDUSTRY); SOLVENT EXTRACTION; ULTRAFILTRATION.

Chemical treatment. Chemical treatment processes alter the chemical structure of the constituents of the waste to produce either an innocuous or a less hazardous material by taking advantage of chemical reactions. This can be said of many biological processes as well, so chemical treatment processes are understood to be abiotic (that is, chemical reactions anywhere outside of a living organism). Four

TABLE 3. Effect of particle size, solids content, and extent of contamination on decontamination efficiencies*

	Predominant particle size			Solids content		High contaminant concentration	
Treatment technology	Sand	Silt	Clay	High (slurry)	Low (in situ)	Organic compounds	Metals
Conventional incineration	N	X	X	F	X	F	X
Innovative incineration	N	X	X	F	X	F	F
Pyrolysis	N	N	N	F	X	F	F
Vitrification	F	X	X	F	X	F	F
Supercritical water oxidation	X	F	F	X	F	F	X
Wet air oxidation	X	F	F	X	F	F	X
Thermal desorption	F	X	X	F	X	F	N
Immobilization	F	X	X	F	X	X	N
Solvent extraction	F	F	X	F	X	X	N
Soil washing	F	F	X	N	F	N	N
Dechlorination	U	U	U	F	X	X	N
Oxidation	F	X	X	N	F	X	X
Bioslurry process	N	F	N	N	F	X	X
Compositing	F	N	X	F	X	F	X
Contained treatment facility	F	N	X	F	X	X	X

*F = sediment characteristic is favorable to the effectiveness of the process.
N = sediment characteristic has no significant effect on process performance.
U = effect of sediment characteristic on process is unknown.
X = sediment characteristic may impede process performance or increase cost.
SOURCE: U.S. Environmental Protection Agency, *Remediation Guidance Document*, EPA-905-B94-003, Chap. 7, 2003.

TABLE 4. Factors for selecting decontamination and treatment approaches*

Treatment technology	Implementability at full scale	Regulatory compliance	Community acceptance	Land requirements	Residuals disposal	Wastewater treatment	Air emission control
Conventional incineration		•	•				•
Innovative incineration		•	•				•
Pyrolysis		•					•
Vitrification	•	•					•
Supercritical water oxidation	•						
Wet air oxidation							
Thermal desorption					•	•	•
Immobilization				•			•
Solvent extraction					•	•	
Soil washing					•	•	
Dechlorination							•
Oxidation	•						
Bioslurry process	•						•
Composting				•			•
Contained treatment facility				•		•	•

*• indicates that the factor is critical in the evaluation of the technology.
SOURCE: U.S. Environmental Protection Agency, *Remediation Guidance Document*, EPA-905-B94-003, Chap. 7, 2003.

categories of chemical reactions are available to treat hazardous wastes: synthesis or combination, decomposition, single replacement, and double replacement.

Synthesis or combination. In combination reactions (1),

$$A + B \rightarrow AB \tag{1}$$

two or more substances react to form a single substance. Two types of combination reactions are important in environmental systems: formation and hydration. Formation reactions are those where elements combine to form a compound. An example is the formation reaction (2) of ferric oxide.

$$4Fe\,(s) + 3O_2\,(g) \rightarrow 2Fe_2O_3\,(s) \tag{2}$$

Hydration reactions involve the addition of water to synthesize a new compound, for example, in reaction (3) when calcium oxide is hydrated to form calcium hydroxide, and reaction (4) when phosphate is hydrated to form phosphoric acid.

$$CaO\,(s) + H_2O\,(l) \rightarrow Ca(OH)_2\,(s) \tag{3}$$

$$P_2O_5\,(s) + 3H_2O\,(l) \rightarrow 2H_3PO_4\,(aq) \tag{4}$$

Decomposition. Decomposition is often referred to as degradation when discussing organic compounds in toxicology, environmental sciences, and engineering. In decomposition reactions (5),

$$AB \rightarrow A + B \tag{5}$$

one substance breaks down into two or more new substances. For example, in reaction (6) calcium carbonate breaks down into calcium oxide and carbon dioxide.

$$CaCO_3\,(s) \rightarrow CaO\,(s) + CO_2\,(g) \tag{6}$$

Single replacement (single displacement). This reaction (7)

$$A + BC \rightarrow AC + B \tag{7}$$

commonly occurs when one metal ion in a compound is replaced with another metal ion, such as in reaction (8) when trivalent chromium replaces monovalent silver.

$$3AgNO_3\,(aq) + Cr\,(s) \rightarrow Cr(NO_3)_3\,(aq) + 3Ag\,(s) \tag{8}$$

Double replacement (metathesis or double displacement). In metathesis reactions (9), cations and anions trade

$$AB + CD \rightarrow AD + CB \tag{9}$$

places. They are commonly encountered in metal precipitation reactions, such as when lead is precipitated, as in reaction (10) of a lead salt with an acid

$$Pb(ClO_3)_2\,(aq) + 2KCl\,(aq) \rightarrow PbCl_2\,(s) + 2KClO_3\,(aq) \tag{10}$$

such as potassium chloride.

Chemical processes are attractive because they produce minimal air emissions, they can often be carried out on the site of the waste generator, and some

processes can be designed and constructed as mobile units. A few commonly applied chemical treatment processes are neutralization, precipitation (for example, waste steams containing heavy metals), ion exchange, dehalogenation (especially dechlorination and debromination), and oxidation-reduction for detoxifying toxic wastes. *See* ACID AND BASE; ION EXCHANGE; OXIDATION-REDUCTION; PRECIPITATION (CHEMISTRY).

Biological treatment. Biological waste-treatment processes are those that use microorganisms to decompose organic wastes into water, carbon dioxide, and simple inorganic substances, or into simpler organic substances such as aldehydes and acids. Contaminants, if completely organic in structure, are in theory completely destructible using microorganisms, with the engineering inputs and outputs summarized in reaction (11).

$$\text{Hydrocarbons} + O_2 + \text{microorganisms (+ energy)} \rightarrow CO_2 + H_2O + \text{microorganisms (+energy?)} \quad (11)$$

Contaminant wastes are mixed with oxygen and aerobic microorganisms, sometimes in the presence of an external energy source in the form of added nutrition for the microorganisms. In time, the by-products of gaseous carbon dioxide and water are produced and exit the top of the reaction vessel, while a solid mass of microorganisms is produced to exit the bottom.

If the waste contains other chemical constituents, in particular chlorine and/or heavy metals, and if the microorganisms are able to withstand and flourish in such an environment, the simple input and output relationship is modified to reaction (12).

$$\begin{aligned}&\text{Hydrocarbons} + O_2 + \text{microorganisms (+energy?)}\\&\quad + \text{Cl or heavy metal(s)} + H_2O + \text{inorganic salts}\\&\quad + \text{nitrogen compounds} + \text{sulfur compounds}\\&\quad + \text{phosphorus compounds} \rightarrow CO_2 + H_2O\ \text{(+energy?)}\\&\quad + \text{chlorinated hydrocarbons or heavy-metal(s)}\\&\qquad \text{inorganic salts} + \text{nitrogen compounds}\\&\quad + \text{sulfur compounds} + \text{phosphorus compounds}\end{aligned} \quad (12)$$

If the microorganisms do survive in this complicated environment, the potential exists for the transformation to a potentially more toxic molecule that contains chlorinated hydrocarbons, higher heavy-metal concentrations, as well as more mobile or more toxic chemical species of heavy metals.

All bioreactor systems share common attributes. All rely on a population of microorganisms to metabolize organic contaminants, ideally into the harmless by-products of $CO_2 + H_2O$ (+energy?). In all the systems, the microorganisms either must be initially cultured in the laboratory to be able to metabolize a specific organic waste, or must be given sufficient time to evolve to be able to digest the contaminant. The engineer must plan for and undertake extensive and continual monitoring and fine-tuning of each microbiological processing system during its operation. *See* BIOCHEMICAL ENGINEERING.

The advantages of biological treatment systems include (1) the potential for energy recovery; (2) volume reduction of the hazardous waste; (3) detoxification as selected molecules are reformulated; (4) basic scientific principles, engineering designs, and technologies that are well understood from a wide range of other applications, including municipal wastewater treatment at facilities; (5) application to most organic contaminants which as a group compose a large percentage of the total hazardous waste generated; and (6) possibility to scale the technologies to handle a single gallon/pound (liter/kilogram) or millions of gallons/pounds (liters/kilograms) of waste per day.

The disadvantages of the biotreatment systems include (1) operation of the equipment requires very skilled operators and is more costly as input contaminant characteristics change over time and correctional controls become necessary; and (2) ultimate disposal of the waste microorganisms is necessary and particularly troublesome and costly if heavy metals and/or chlorinated compounds are found during the monitoring activities. In general, hazardous waste containing heavy metals should not be bioprocessed without specific additional steps.

Variations on three different types of bioprocessors are available to treat hazardous wastes: trickling filter, activated sludge, and aeration lagoons.

Trickling filter. The classical design of a trickling filter system includes a bed of fist-sized rocks enclosed in a rectangular or cylindrical structure through which is passed the waste (**Fig. 2**). Biofilms are selected from laboratory studies and encouraged to grow on the rocks. As the liquid waste moves downward through the bed, the microorganisms come in contact with the organic contaminant/food source and ideally metabolize the waste into $CO_2 + H_2O +$ microorganisms (+energy?). Oxygen is supplied by blowers from the bottom of the reactor and passed upward through the bed. The treated waste moves downward through the bed and subsequently enters a quiescent tank where the microorganisms that are sloughed off the rocks are settled, collected, and disposed. Trickling filters actually are considered mixed treatment systems because aerobic bacteria grow in the upper, higher-oxygen layers of the

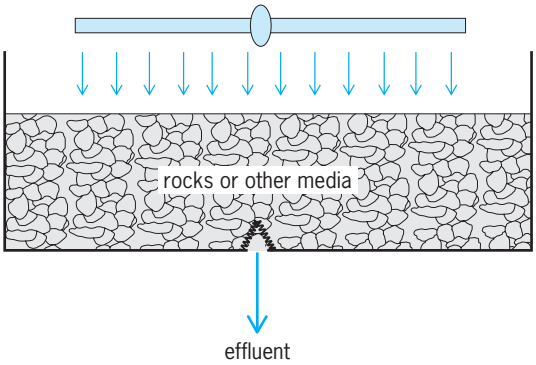

Fig. 2. Trickling filter treatment system. (*After D. Vallero, Engineering the Risks of Hazardous Wastes, Butterworth-Heinemann, 2003*)

media, while anaerobes grow in the regions lower in the system. *See* BIOFILM; INDUSTRIAL WASTEWATER TREATMENT.

Activated sludge. The key to the activated sludge system is that the microorganisms are recycled within the system (**Fig. 3**). This reuse enables the microorganisms to evolve over time and adapt to the changing characteristics of the hazardous waste.

In an activated-sludge system, a tank of liquid hazardous waste is injected with a mass of microorganisms. Oxygen is supplied in the aeration basin as the microorganisms come in contact, sorb, and metabolize the waste, ideally into $CO_2 + H_2O$ + microorganisms (+energy?). The heavy (fed) microorganisms then flow into a quiescent tank, where they settle by gravity and are collected and disposed. Depending on the operating conditions of the facility, some or many of the settled (now hungry) and active microorganisms are returned to the aeration basin, where they feed again. Liquid effluent from the activated-sludge system may require additional microbiological and/or chemical processing prior to release into a receiving stream or sewer system.

Aeration lagoons. These ponds treat liquid and dissolved contaminants for long terms, from months to years (**Fig. 4**). Persistent organic molecules, those not readily degraded in trickling filter or activated sludge systems, are potentially broken down by certain microbes into $CO_2 + H_2O$ + microorganisms (+energy?) if given enough time. The ponds are open to the weather, and ideally oxygen is supplied directly to the microorganisms from the atmosphere.

Since biological treatment systems do not alter or destroy inorganic substances, and high concentrations of such materials can severely inhibit decomposition activity, chemical or physical treatment may be required to extract inorganic materials from a waste stream prior to biological treatment. *See* SEWAGE TREATMENT.

Thermal treatment. If completely organic in structure, hazardous wastes are completely destructible using principles based in thermodynamics with the engineering inputs and outputs as in reactions (13) and (14).

$$\text{Hydrocarbons} + O_2\ (\text{+energy?}) \rightarrow CO_2 + H_2O\ (\text{+ energy?}) \quad (13)$$

$$(CH)_x + O_2 \rightarrow CO_2 + H_2O \quad (14)$$

Complete combustion may also result in the production of molecular nitrogen (N_2) when nitrogen-containing organics are burned, such as in the combustion of methylamine. Incomplete combustion can produce a variety of compounds. Some are more toxic than the original compounds being oxidized, such as polycyclic aromatic hydrocarbons (PAHs), dioxins, furans, and carbon monoxide. *See* COMBUSTION.

Wastes are mixed with oxygen, sometimes in the presence of an external energy source, and within several seconds the by-products of gaseous carbon dioxide and water are produced and exit the top of the reaction vessel, while a solid ash is produced that exits the bottom of the reaction vessel. Energy may also be produced during the reaction, and the heat may be recovered.

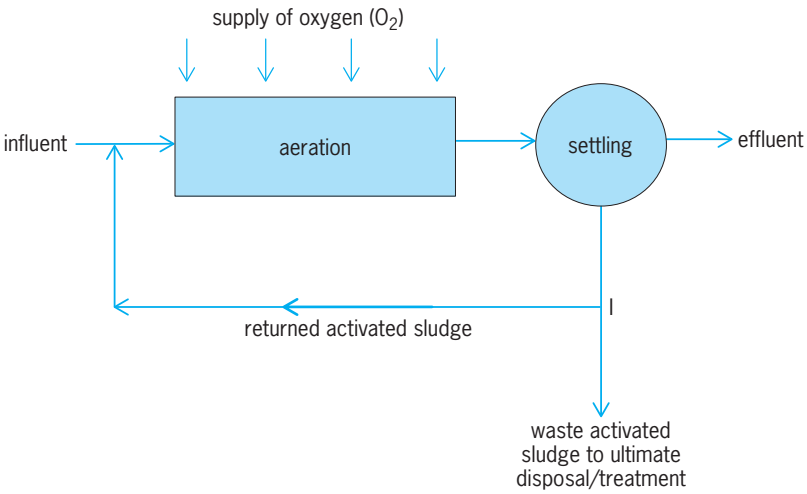

Fig. 3. Activated sludge treatment system. (*After D. Vallero, Engineering the Risks of Hazardous Wastes, Butterworth-Heinemann, 2003*)

If the wastes contain other chemical constituents, particularly chlorine and/or heavy metals, the situation becomes complex. For such wastes, the potential exists for destroying the initial contaminant as well as exacerbating the problem by generating hazardous off-gases containing chlorinated hydrocarbons or ashes containing heavy metals. For example, the improper incineration of certain chlorinated hydrocarbons can lead to the formation of the highly toxic chlorinated dioxins, furans, and hexachlorobenzene.

The advantages of thermal systems include (1) the potential for energy recovery; (2) volume reduction of the contaminant; (3) detoxification as selected molecules are reformulated; (4) basic scientific principles, engineering designs, and technologies that are well understood from a wide range of other applications, including electric generation and municipal solid-waste incineration; (5) application to most organic contaminants which compose a large percentage of the total contaminants generated worldwide; (6) possibility to scale the technologies to handle a single gallon per pound (liter per kilogram) of waste or millions of gallons per pound (liter per kilogram) of waste: and (7) small land area requirements relative to other hazardous waste management facilities such as landfills.

In general, the same reaction applies to most thermal processes of gasification, pyrolysis, hydrolysis, and combustion. The actual reactions from test burns

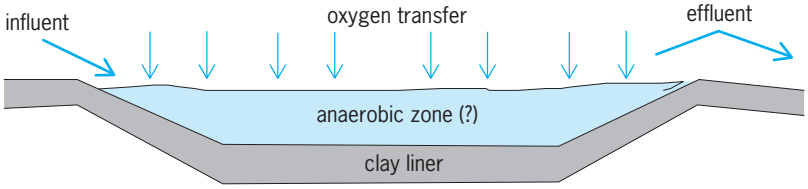

Fig. 4. Aeration pond. (*After D. Vallero, Engineering the Risks of Hazardous Wastes, Butterworth-Heinemann, 2003*)

for commonly incinerated compounds are provided in **Table 5**.

Of all of the thermal processes, incineration is the most common process for destroying organic contaminants in industrial wastes. Incineration simply is heating wastes in the presence of oxygen to oxidize organic compounds. There are a number of stages in the incineration process where new compounds may need to be addressed. Ash and other residues may contain high levels of metals, at least higher than the original feed. The flue gases are likely to include both organic and inorganic compounds resulting from temperature-induced volatilization or newly transformed products of incomplete combustion with higher vapor pressures than the original contaminants.

The disadvantages of hazardous waste incinerators include: (1) the equipment is capital-intensive, particularly the refractory material lining the inside walls of each combustion chamber which must be replaced as cracks form whenever a combustion system is cooled and heated; (2) the equipment requires very skilled operators and is more costly to operate when fuel must be added to the system; (3) the disposal of the ash is necessary and particularly troublesome and costly if heavy metals and/or chlorinated compounds are present; and, (4) the air emissions may be hazardous and thus must be monitored for chemical constituents and must be controlled, including the release of heavy metals and their compounds.

Pyrolysis is the process of chemical decomposition induced in organic materials by heat in the absence of oxygen. It is impossible to achieve a completely oxygen-free atmosphere, so pyrolytic systems run with less than stoichiometric quantities of oxygen. During pyrolysis, organic compounds are converted to gaseous components (such as CO, H_2, CH_4, and other hydrocarbons), along with some liquids, coke (the solid residue of fixed carbon), and ash. Pyrolysis generally takes place well above atmospheric pressure at temperatures exceeding 430°C (800°F). The secondary gases need their own treatment, such as by a secondary combustion chamber. Particulates must be removed by additional air pollution controls, such as fabric filters or cyclones.

The target contaminant groups for pyrolysis include semivolatile organic compounds, including pesticides, polychlorinated biphenyls (PCBs), dioxins, and PAHs. Pyrolysis allows for separating organic contaminants from various wastes, including those from refineries; coal tar; wood-preservative, creosote-contaminated, and hydrocarbon-contaminated soils; mixed radioactive and hazardous wastes; synthetic rubber processing; and paint and coating processes. Pyrolysis systems may be used to treat a variety of organic contaminants that chemically decompose when heated (cracking). Pyrolysis is not effective in either destroying or physically separating inorganic compounds that coexist with the organics in the contaminated medium. Volatile metals may be removed and transformed, but their mass balance will not change. *See* PYROLYSIS.

TABLE 5. Balanced combustion reactions for selected organic compounds

Compound	Combustion reaction
Chlorobenzene	$C_6H_5Cl + 7O_2 \rightarrow 6CO_2 + HCl + 2H_2O$
TCE	$C_2Cl_4 + O_2 + 2H_2O \rightarrow 2CO_2 + HCl$
HCE	$C_2Cl_6 + {}^1/_2O_2 + 3H_2O \rightarrow 2CO_2 + 6HCl$
CPVC	$C_4H_5Cl_3 + 4{}^1/_2O_2 \rightarrow 4CO_2 + 3HCl + H_2O$
Natural gas fuel (methane)	$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$
PTFE Teflon	$C_2F_4 + O_2 \rightarrow CO_2 + 4HF$
Butyl rubber	$C_9H_{16} + 13O_2 \rightarrow 9CO_2 + 8H_2O$
Polyethylene	$C_2H_4 + 3O_2 \rightarrow 2CO_2 + 2H_2O$
Wood*	$C + O_2 \rightarrow CO_2$
	$H + 0.25O_2 \rightarrow 0.5H_2O$

*Wood is considered to have the composition $C_{6.9}H_{10.6}O_{3.5}$. Therefore the combustion reactions are simple carbon and hydrogen combustion.

SOURCE: U.S. Environmental Protection Agency.

Emerging thermal technologies. New thermal processes include high-pressure oxidation and vitrification. High-pressure oxidation combines two related technologies—wet air oxidation and supercritical water oxidation—which combine high temperature and pressure to destroy organics. Wet air oxidation has generally been limited to conditioning of municipal wastewater sludges, but can degrade hydrocarbons (including PAHs), certain pesticides, phenolic compounds, cyanides, and other organic compounds.

Vitrification uses electricity to heat and destroy organic compounds and immobilize inert contaminants. The product is a solid, glasslike material that is very resistant to leaching.

Solidification and stabilization. Solidification and stabilization are treatment systems designed to improve handling and the physical characteristic of the waste, decrease the surface area across which transfer or loss of contained pollutants can occur, and limit the solubility of, or detoxify, any hazardous constituents contained in the wastes. In solidification, these results are obtained primarily by producing a monolithic block of treated waste with high structural integrity. Stabilization techniques limit the solubility or detoxify waste contaminants even though the physical characteristics of the waste may not be changed. Stabilization usually involves the addition of materials that ensure that the hazardous constituents are maintained in their least soluble or least toxic form.

Disposal. After treatment is completed, an inorganic valueless residue remains that must be disposed of safely. There are five options for disposing of hazardous waste: (1) Underground injection wells are steel- and concrete-encased shafts placed deep below the surface of the earth into which hazardous wastes are deposited by force and under pressure. Some liquid waste streams are commonly disposed of in underground injection wells. (2) Surface impoundment involves natural or engineered depressions or diked areas that can be used to treat, store, or dispose of hazardous waste. Surface impoundments are often referred to as pits, ponds, lagoons, and basins. (3) Landfills are disposal facilities where

hazardous waste is placed in or on land. Properly designed and operated landfills are lined to prevent leakage, and contain systems to collect potentially contaminated surface water runoff. Most landfills isolate wastes in discrete cells or trenches, thereby preventing potential contact of incompatible wastes. (4) Land treatment is a disposal process in which hazardous waste is applied onto or incorporated into the soil surface. Natural microbes in the soil break down or immobilize the hazardous constituents. Land treatment facilities are also known as land-application or land-farming facilities. (5) Waste piles are noncontainerized accumulations of solid, nonflowing hazardous waste. While some are used for final disposal, many waste piles are used for temporary storage until the waste is transferred to its final disposal site.

Of the hazardous waste disposed of on land in the United States, nearly 60% is disposed of in underground injection wells, approximately 35% in surface impoundments, 5% in landfills, and less than 1% in waste piles or by land application.

Comprehensive treatment facilities. In the United States, Canada, and many western European countries, there are large facilities capable of treating and disposing of many types of hazardous wastes. These facilities are able to realize economies of scale by incorporating many treatment processes that might not be economical for individual generators. Comprehensive facilities are also able to exploit the synergistic opportunities made possible by having many different types of waste present at a single site. These include using waste acids and alkalies to neutralize one another, waste oxidants to treat cyanides and organic contaminants in water, salts and acids to salt out organic compounds from wastewater, onsite incinerators to dispose of organic vapors generated by other onsite processes, ash and calcium and magnesium oxides to aid in stabilization processes, and combustible solids and liquids to produce blended liquid fuels. Although these technical advantages are compelling, the siting of these facilities is often resisted by neighbors and is often opposed politically. This opposition is known as NIMBY, "not in my back yard."

Abandoned disposal sites. There are sites in many countries where hazardous waste has been disposed of improperly in the past and where cleanup operations are needed to restore the sites to their original state. In the United States, about 20,000 of these sites have been identified, and approximately 2000 sites require immediate action. Denmark, the Netherlands, and Sweden have also identified sites requiring immediate attention. Environmental regulations lay out explicit steps for cleaning up contamination (**Fig. 5**). The first step of a contaminant cleanup is a preliminary assessment (PA). During the PA of a site, readily available information about a site and its surrounding area are collected to determine the threat to human health and the environment. Any possible emergency response actions may also be identified. If information beyond the preliminary data in the PA is needed, a site inspection is performed.

Cleaning up abandoned disposal sites involves isolating and containing contaminated material, removing and redepositing contaminated sediments, and in-situ and direct treatment of the hazardous wastes involved.

Ex-situ and in-situ treatment. Contaminated soil and sediment must often be removed and treated

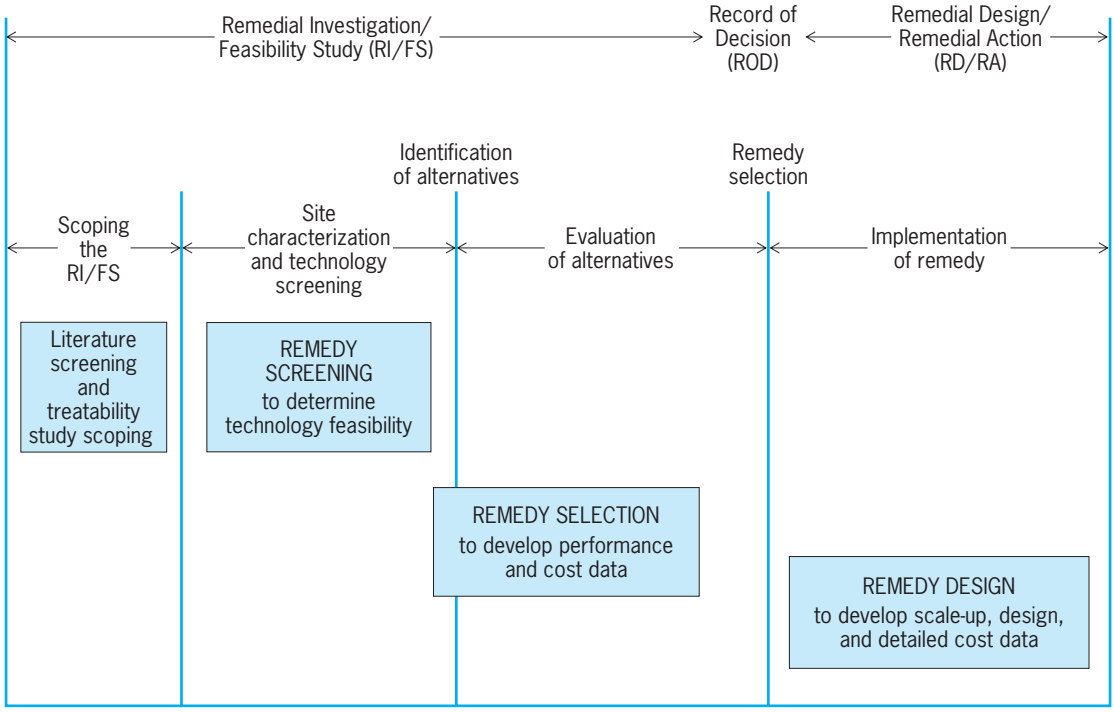

Fig. 5. Steps needed in the cleanup of an abandoned hazardous waste site. (*U.S. Environmental Protection Agency, Guide for Conducting Treatability Studies under CERCLA: Thermal Desorption, EPA/540/R-92/074 B, 1992*)

off-site (ex situ). Contaminated soil may be excavated and transported to kilns or other high-temperature operations, where it is mixed with combustible material. High-sand-content soils may be used as part of an asphalt mix. Contaminated soil may also be distributed onto an impermeable surface, allowing the more volatile compounds to evaporate. Microbial biodegradation can be accelerated and enhanced by adding nutrients and moisture to the soil. A faster process, thermal desorption, entails heating the soil to evaporate the contaminants and capture the compounds, and then burning them in a vapor treatment device.

Ground water generally is treated by drilling recovery wells to pump contaminated water to the surface. Commonly used ground-water treatment approaches include air stripping, filtering with granulated activated carbon (GAC), and air sparging. Air stripping transfers volatile compounds from water to air. Ground water is allowed to drip downward in a tower filled with a permeable material through which a stream of air flows upward. Another method bubbles pressurized air through contaminated water in a tank. Filtering ground water with GAC entails pumping the water through the GAC to trap the contaminants. In air sparging, air is pumped into the ground water to aerate the water. Most often, a soil venting system is combined with an air sparging system for vapor extraction.

Hazardous wastes treated where they are found, without first removing them, is known as in-situ remediation. Bioremediation makes use of living microorganisms to break down toxic chemicals or to render them less hazardous. This is often done by using bacteria, and in some instances algae and fungi, that are already living in the soil, sediment, or water. These microbes are exposed to increasing amounts of the chemical, so that they adapt to using the chemical as an energy (food) source. This process is known as acclimation. The acclimated microbes can then be taken from the laboratory and applied to the waste in the field. The most passive form of bioremediation is natural attenuation, where no direct engineering intervention is used, and the contaminants are allowed to be degraded by resident microbes over time. The only role for the engineer is to monitor the soil and ground water to measure the rate at which the chemicals are degrading. Natural attenuation can work well for compounds that are found in the laboratory to break down under the conditions found at the site. For example, if a compound is degraded under reduced, low-pH conditions in the laboratory, it may also degrade readily in soils with these same conditions (for example, in deeper soil layers where bacteria have adapted to these conditions naturally). Natural attenuation can sometimes be enhanced, for example, when soil moisture levels are increased.

Plant life may also be used to reduce the amount of hazardous wastes. In phytoremediation, contaminated areas are seeded, and plant roots extract the chemicals from the soil. The harvested plants are either treated on-site or transferred to a treatment facility. In reality, both microbial and macrophytic processes occur simultaneously. Fast-growing trees (for example, poplars), can help to treat areas contaminated with agricultural chemicals. Plants, such as grasses and field crops, have even been used to treat the very persistent polychlorinated biphenyls (PCBs), wood preservatives, and petroleum. Plants have also been used to extract heavy metals and radioactive substances from contaminated soil. Bioremediation has been used successfully to treat numerous other organic and inorganic compounds.

Various methods are available for treating substances after they have been released into the environment. However, eliminating the wastes before they are released is the best means of reducing hazards and the associated risks, especially by applying the principles of green chemistry and sustainable design. *See* GREEN CHEMISTRY. Daniel Vallero

Bibliography. Federal Remediation Technologies Roundtable, *Remediation Technologies Screening Matrix and Reference Guide*, 4th ed., 2002; H. M. Freeman (ed.), *Hazardous Waste Minimization*, 1990; H. M. Freeman (ed.), *Standard Handbook of Hazardous Waste Treatment and Disposal*, 2d ed., 1997; A. Gaudy and E. Gaudy, *Elements of Bioenvironmental Engineering*, 1988; C. N. Haas and R. J. Ramos, *Hazardous and Industrial Waste Treatment*, 1995; Metcalf and Eddy, Inc., by G. Tchobanoglous, F. L. Burton, and H. D. Stensel, *Wastewater Engineering*, 4th ed., 2002; J. J. Peirce, R. Weiner, and P. A. Vesilind, *Environmental Pollution and Control*, 4th ed., 1998; M. N. Sara, Groundwater monitoring system design, in D. M. Nielsen (ed.), *Practical Handbook of Ground-Water Monitoring*, Lewis Publishers, Chelsea, MI, 1991; H. Tammenagi, *The Waste Crisis: Landfills, Incinerators, and the Search for a Sustainable Future*, 1999; U.S. Environmental Protection Agency, *Remediation Guidance Document*, EPA-905-B94-003, Chap. 7, 2003; U.S. Environmental Protection Agency, *Test Methods for Evaluating Solid Waste*, vols. I and II (SW-846), 3d ed, 1986; D. A. Vallero, *Engineering the Risks of Hazardous Wastes*, 2003; D. A. Vallero, *Environmental Contaminants: Assessment and Control*, 2004; C. Wentz, *Hazardous Waste Management*, 1989; D. G. Wilson, History of Solid Waste Management, in *Handbook of Solid Waste Management*, 1997.

Head

The region of the body consisting of the skull, its contents, and related structures. Its two principal parts are the cranium, or braincase, and the face.

The skin, hair, and subcutaneous tissues over the top of the skull are collectively known as the scalp. The regions of the cranium take their names from the underlying bones, for example, the temporal, parietal, frontal, and occipital regions.

The intracranial contents include the brain and uppermost portion of the spinal cord with their coverings (meninges), blood vessels, and the important cranial nerves, as well as the cerebrospinal fluid system. Many openings, or foramina, afford means of

passage from within the skull for nerves and blood vessels. *See* BRAIN.

The face contains the mouth and jaws, the nose and eyes (the ears are in the cranial region), and many related structures. Muscles acting on the lower jaw and on the skin (facial muscles, or muscles of expression) cover much of the face and part of the cranium. Thomas S. Parsons

Headache

Pain within the head. It is probably the most common complaint for which people seek a physician's help. Headaches can be grouped into three primary categories: vascular, muscle-contraction, and organic.

Vascular headaches. Vascular headaches include classic and common migraine as well as cluster, toxic, and hypertensive headaches. All are caused by dilation of cerebral blood vessels. Constriction of the blood vessels may also occur in any part of the cerebral vasculature and cause the neurologic symptoms associated with some forms of vascular headache.

Migraine. Migraine is a French word derived from the Greek *hemicrania*, which means half a head. An estimated 14 to 16 million people in the United States suffer from migraine.

Migraine affects one side of the head but may be bilateral. This may occur several times each month, although some individuals experience attacks only once or twice yearly. Neurologic symptoms, especially visual disturbances, are common. The associated symptoms of migraine include nausea, vomiting, loss of appetite, and irritability. During the migraine attack, some patients complain of sensitivity to light and sound. In classic migraine, the attacks may be preceded for several minutes to an hour by visual disturbances such as blind spots, flashing lights, and decreased visual field, or other neurologic symptoms such as numbness, speech disorders, or perception of strange odors.

Migraine tends to be a familial disorder. Migraine usually begins in young adulthood but can appear early in childhood; it rarely first occurs after age 40. An estimated 60% of migraine sufferers are women, and often the first migraine coincides with menarche. The onset of an acute migraine attack has been linked to such factors as diet, stress, fatigue, menstruation, hormonal changes, hypoglycemia, and oversleeping. A small number of persons can be treated by the elimination of certain foods, such as chocolate, aged cheese, fermented foods, and alcoholic beverages from the diet. A regular schedule of meal times, adequate sleep, and a consistent time for awakening have also been shown to be beneficial. *See* ALLERGY.

After an attack has begun, the drugs of choice for treatment are the ergotamines. If used early in an attack, they may prevent its progression. For those in whom ergotamine is contraindicated or not tolerated, isometheptene mucate has been effective in the abortive treatment of migraine, as have the nonsteroidal anti-inflammatory agents such as naproxen sodium. *See* ERGOT AND ERGOTISM.

Prophylactic therapy should be considered when attacks occur more often than twice a month. The agent of choice in migraine prophylaxis is the beta blocker, propranolol. Other beta blockers have been used successfully in migraine treatment, but only propranolol has been approved for this therapeutic indication by the U.S. Food and Drug Administration.

Antidepressants and the nonsteroidal anti-inflammatory drugs have been used successfully in treating migraine, and calcium channel blockers are another potentially effective remedy. Serotonin, a substance found in blood platelets, has been implicated in blood vessel changes, and serotonin blockers may offer additional treatment options. *See* SEROTONIN.

Many physicians recommend nonpharmaceutical treatment such as biofeedback, a method by which the patient learns to control a physiological function that was previously considered involuntary. By using self-hypnotic phrases and progressive relaxation exercises, as well as instrumentation, the patient may learn to relax the muscles in the upper part of the body and to increase finger temperature, which have been shown to abort migraine attacks. Biofeedback training has been especially successful with children. *See* HYPNOSIS.

Cluster headache. Cluster headache is the occurrence of migraines in groups or series. Men are more prone to cluster headaches. The cluster headache is characterized by its one-sided, excruciating attack that is usually localized around one eye. Typically, the attack is brief, lasting from several minutes to 1 h, and it often awakens the person from sleep. The associated manifestations include facial flushing, nasal congestion or draining, tearing, and on occasion drooping of the eyelid and constriction of the pupil. The series of headaches lasts from a few weeks to several months and then disappears for several months or years. The series appear most frequently in spring and fall. Although the cause remains unknown, any alcohol consumption during a series of cluster headaches is known to precipitate an acute attack.

The drugs of choice for treating cluster headaches are the ergotamines, but they must be used as early as possible in the attack to be effective. The inhalation of pure oxygen has also proved beneficial. Because cluster headaches are self-limiting, therapy is short-term. Methysergide, a potent serotonin receptor antagonist, is the agent of choice for prophylaxis. Corticosteroids, including prednisone, can prevent the onset of attacks. For those with chronic cluster headaches, lithium carbonate has provided some relief.

Other forms of vascular headache may be caused by systemic infection or fever, which causes dilation of the blood vessels. The ingestion of alcohol, poisons, or some medications used to treat hypertension or cardiac disease may produce adverse effects, including vascular headaches. *See* HYPERTENSION.

Muscle-contraction headaches. The most common form of form of headache is the muscle-contraction or tension headache. It is characterized by dull, constricting pain that can either occur intermittently

or continue for days, months, or years. Intermittent muscle-contraction headaches usually occur once or twice a week or less frequently and show no definite pattern. Over-the-counter analgesics, including aspirin, acetaminophen, and the ibuprofen compounds, provide relief in most cases.

Muscle-contraction headaches usually affect both sides of the head and may be described as having a hat-band distribution of pain. The chronic form occurs on a daily or near-daily basis, with a worsening of pain in the morning and late evening. If the headache is associated with a sleep disturbance, it may be symptomatic of depression, and antidepressant therapy or psychological counseling may be beneficial.

A daily headache pattern carries the accompanying risk of habituation to prescription and over-the-counter analgesics. Hospitalization in a specialized unit dedicated to the problem of headaches may yield alternative means of obtaining pain relief.

Organic headaches. Very few headaches have an organic cause, such as brain tumor or aneurysm. Headache is not a prominent symptom of brain tumor: if present, headache will become progressively worse and constant, and it may not appear until late in the course of the tumor development. Other, more prominent neurologic symptoms usually arise before the onset of headache.

The headache associated with an aneurysm is usually mild until the aneurysm is at the point of rupture. If a patient complains of an exceptionally severe headache, organic disease, such as aneurysm, must be ruled out. Rupture of an aneurysm may have morbid consequences and must be treated immediately. *See* ANEURYSM.

Sinus headache is a common misnomer: chronic sinus disease does not cause headache. Acute sinus headache is characterized by nasal congestion and fever. The headache is minimal in the morning and increases in severity through the day. Treatment consists of decongestants as well as antibiotic therapy.

Temporamandibular joint (TMJ) disease involves a faulty bite or misalignment of the teeth and can cause a headache, but many people whose headaches have been incorrectly attributed to TMJ disease have not found relief from pain through dental appliances or surgery.

Eye conditions may also cause headache. The increased intraocular pressure of glaucoma, for example, may cause a headache, and so complaints of a recent onset of headache, particularly in the elderly, should prompt a screening for glaucoma. *See* GLAUCOMA; PAIN.

Seymour Diamond

Bibliography. S. Diamond and D. J. Dalessio, *The Practicing Physician's Approach to Headache*, 6th ed., 1999; H. G. Wolff, *Headache and Other Head Pain*, 5th ed., 1987.

Health physics

A branch of the environmental and occupational safety health sciences and professions concerned with the protection of people and the environment from possibly harmful effects of radiation, while providing for the utilization of radiation for the benefit of society. Health physicists are interdisciplinary radiation protection and safety specialists whose expertise draws from environmental science, mathematics, medicine, radiological health, radiation biology, chemistry, and physics. The subject requires understanding of the generation, measurement, and characteristics of radiation; environmental transport of radionuclides; dosimetry; effects of radiation in biological systems; and the regulations and recommendations governing the use of radiation. The field has expanded to include nonionizing as well as ionizing sources of radiation.

Although radiation and radioactivity were discovered in 1895, radiation protection activities began almost immediately, around 1900. The field received a major stimulus, expansion, and formalization during World War II as a part of the atomic bomb (Manhattan) project. Contemporary health physics includes basic and applied research on fundamental biological and physical radiation interaction mechanisms; on dosimetry, shielding, and instrumentation; on quantification of radiation effects; and on radioecology. Industrial hygiene, epidemiology, environmental engineering, and electronics and instrument development also include fundamental research in health physics.

Health physics specialists are involved in nuclear power radiation monitoring, control, and regulation required for the safe operation of facilities. In medicine, health physicists help design and survey radiation therapy facilities, and develop techniques to minimize patient doses so as to derive the maximum benefit from diagnostic radiology and nuclear medicine. Industrial applications employing health physics include industrial radiography and x-ray diffraction, gages, product and food sterilization by radiation, and radioactive tracer techniques. Health physics is employed in many aspects of nuclear waste management and remedial actions. Environmental health physics covers measuring and possibly reducing exposure to naturally occurring radioactivity, especially radon; monitoring nuclear facilities and waste sites; consumer-product radiation safety; and radiation accident monitoring, management, and mitigation. Health physics includes all aspects of radiation protection, from initial measuring and planning, through monitoring and implementation, to routine control and emergency management. *See* DECONTAMINATION OF RADIOACTIVE MATERIALS; DOSIMETER; ENVIRONMENTAL RADIOACTIVITY; NUCLEAR MEDICINE; RADIATION INJURY (BIOLOGY); RADIATION SHIELDING; RADIOACTIVE TRACER; RADIOACTIVITY; RADIOACTIVITY AND RADIATION APPLICATIONS; RADIOACTIVITY STANDARDS; RADIOECOLOGY; RADIOGRAPHY; RADIOLOGY; RADON.

Marvin Goldman

Bibliography. H. Cember, *Introduction to Health Physics*, 3d ed., 2000; National Council on Radiation Protection and Measurements, *Limitation of Exposure to Ionizing Radiation*, Rep. 116, 1993; B. Shleien (ed.), *The Health Physics and Radiological Health Handbook*, 3d ed., 1997.

Hearing (human)

The general perceptual behavior and the specific responses that are made in relation to sound stimuli. The auditory system consists of the ear and the auditory nervous system. The ear comprises outer, middle, and inner ear. The outer ear, visible on the surface of the body, directs sounds to the middle ear, which converts sounds into vibrations of the fluid that fills the inner ear. The inner ear contains the vestibular and the auditory sensory organs. The vestibular organ's function is the key to balance: this organ converts movements of the head into nerve signals. *See* EAR (VERTEBRATE).

The auditory part of the inner ear, known as the cochlea because of its snaillike shape, analyzes sound in a way that resembles spectral analysis. It contains the sensory cells that convert sounds into nerve signals to be conducted through the auditory portion of the eighth cranial nerve to higher brain centers. The neural code in the auditory nerve is transformed as the information travels through a complex system of nuclei connected by fiber tracts, known as the ascending auditory pathways (**Fig. 1**). They carry auditory information to the auditory cortex, which is the part of the sensory cortex where perception and interpretation of sounds are believed to take place. The auditory cortex is linked to other parts of the brain, but these connections are usually not considered part of the auditory system. Interaction between the neural pathways of the two ears makes it possible for a person to determine the direction of a sound's source. *See* BINAURAL SOUND SYSTEM; BRAIN.

Outer ear. The pinna collects sound, but because it is small in relation to the wavelengths of sound that are important for human hearing, the pinna plays only a minor role. The ear canal acts as a resonator: it increases the sound pressure at the tympanic membrane in the frequency range between 1500 and 5000 Hz; at 3500 Hz the increase is greatest, about 10 decibels, but together with the gain from the diffraction of the head, the sound pressure may be as much as 17 dB higher than outside the head. Because of diffraction by the head, the sound that reaches the ears depends on the angle to a sound source. The difference between the arrival time of a sound at each of the two ears and the difference in the intensity of the sound that reaches each ear are direct functions of the angle between a line through the two ears and the sound source. This difference is used by the auditory nervous system to determine the location of the sound source. That sound-localizing ability is most highly developed in the horizontal plane.

Middle ear. Sound that reaches the tympanic membrane causes the membrane to vibrate, and these vibrations set in motion the three small bones of the middle ear: the malleus, the incus, and the stapes. The footplate of the stapes is located in an opening of the cochlear bone—the oval window. Moving in a pistonlike fashion, the stapes sets the cochlear fluid into motion and thereby converts sound (pressure fluctuations in the air) into motion of the cochlear fluid. Since motion of the fluid in the cochlea begins the neural process known as hearing, the transformation of sound into motion of the cochlear fluid must occur as efficiently as possible. Sound is a weak force compared to the force required to set the cochlear fluid into motion; in other words, the air in which sound is transmitted has a low impedance compared to the impedance of the fluid that the sound must move. Because the most efficient transfer of energy occurs when the impedances of the energy source and the energy receiver are equal, direct transfer of sound to a fluid would be inefficient. In fact, 99.9% of the sound energy would be lost by reflection from the air–fluid interface. To improve efficiency,

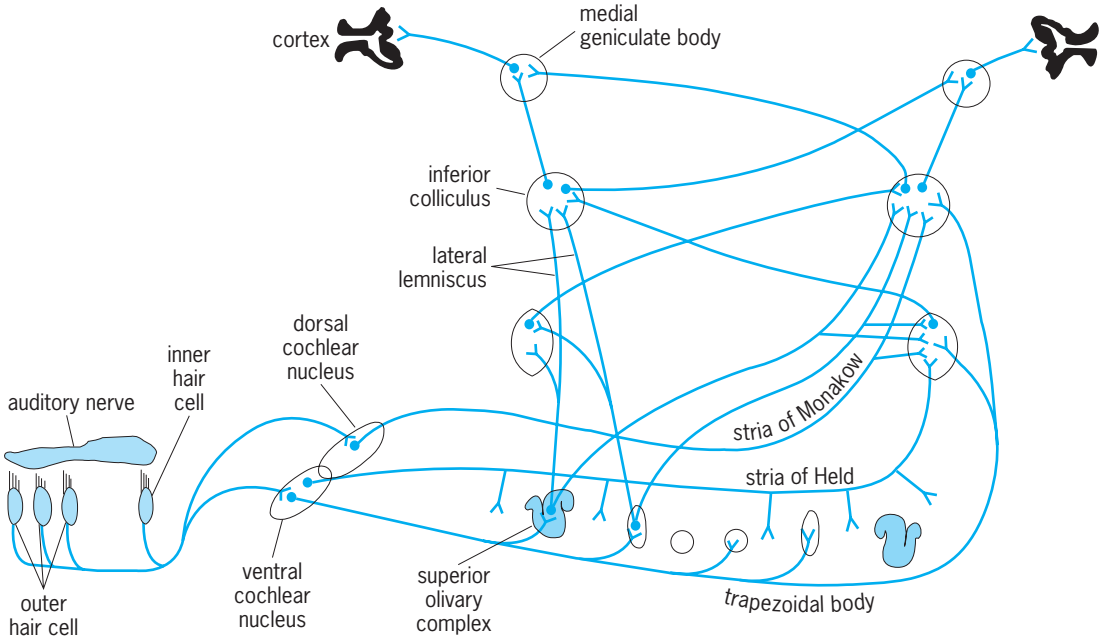

Fig. 1. Ascending auditory pathway. (*After A.R. Møller, Auditory Physiology, Academic Press, 1983*)

the middle ear in humans acts as an impedance transformer: it matches the high impedance of the cochlear fluid to the low impedance of air, thereby facilitating transmission of sound to the cochlear fluid. This transformer action mainly depends upon the ratio between the surface areas of the tympanic membrane and the stapes footplate. Optimum transfer of sound to the inner ear would occur if this ratio was equal to the square root of the ratio between the impedances of air and the cochlear fluid, giving a perfect impedance match. The properties of the human middle ear almost fit that criterion, reducing to a minimum the loss of sound energy from reflection.

The middle-ear bones are enclosed in the air-filled middle-ear cavity; it is important that the air pressure in this cavity be closely maintained at the ambient pressure. The Eustachian tube, which connects the middle-ear cavity to the nasopharynx, maintains that balance. This tube is usually closed, but it opens briefly during swallowing and yawning, thereby equalizing the air pressure in the middle-ear cavity.

There are two small muscles in the middle ear: the tensor tympani and the stapedius muscles. The former pulls the manubrium of the malleus inward, while the latter is attached to the stapes and pulls the stapes in a direction that is perpendicular to its pistonlike motion. The stapedius muscle is the smallest striated muscle in the body, and it contracts in response to an intense sound (of about 85 dB above the threshold of hearing). This is known as the acoustic middle-ear reflex. The muscle's contraction reduces sound transmission through the middle ear and thus acts as a regulator of input to the cochlea. Perhaps a more important function of the stapedius muscle is that it contracts immediately before and during a person's own vocalization, reducing the sensitivity of the speaker's ears to his or her own voice and possibly reducing the masking effect of an individual's own voice. The role of the tensor tympani muscle is less well understood, but it is thought that contraction of the tensor tympani muscle facilitates proper ventilation of the middle-ear cavity.

These two muscles are innervated by different cranial nerves: the facial (VIIth) nerve for the stapedius and the trigeminal (Vth) nerve for the tensor tympani. The acoustic stapedius reflex plays an important role in the clinical diagnosis of disorders affecting the middle ear, the cochlea, and the auditory nerve. A contraction of the stapedius muscle can easily be recorded by using a change in the ear's acoustic impedance as an indicator, because a contraction of the stapedius muscle makes the middle ear more stiff.

Cochlea. In the 1930s, G. von Békésy proved that the basilar membrane of the cochlea has frequency-selective properties. Using human cadaver ears, he showed that vibrations in the cochlear fluid set up a traveling wave on the basilar membrane and that this wave always progresses from the base of the cochlea toward the apex. The amplitude of the wave increases until it reaches a certain point, and then rapidly decreases. The location of the point of maximal deflection is a direct function of the frequency of the sound. When tones are used to set the cochlear fluid into vibration, one specific point on the basilar membrane will vibrate with a higher amplitude than any other. Therefore, a frequency scale can be laid out along the basilar membrane, with low frequencies near the apex and high frequencies near the base of the cochlea (**Fig. 2**). The propagation velocity of this traveling wave is much lower than the propagation velocity of the sound wave in the fluid, and most of the energy is transferred to the basilar membrane at its basal part. As the traveling wave approaches the point of maximal deflection, the propagation velocity decreases markedly. These properties of the basilar membrane are largely a result of the change in compliance along the membrane, which is greater than 1:100.

The sensory cells that convert the motion of the basilar membrane into a neural code in individual auditory nerve fibers are located along the basilar membrane (**Fig. 3**). They are also known as hair cells, because they have hairlike structures on their surfaces. These "hairs" are actually stereocilia, and they are deflected when the basilar membrane vibrates. It is this deflection of the hairs that controls the discharges in individual fibers of the auditory nerve.

Sensory transduction in hair cells. The sensory cells in the mammalian cochlea are similar to the cells that are found in, for instance, the airways, where they transport mucus by rhythmic movements (beats) of their hairs (cilia). Cells in the airways have two types of cilia: stereocilia and kinocilia. The hair cells of the mammalian cochlea lose their kinocilia around the time of birth, however, and only

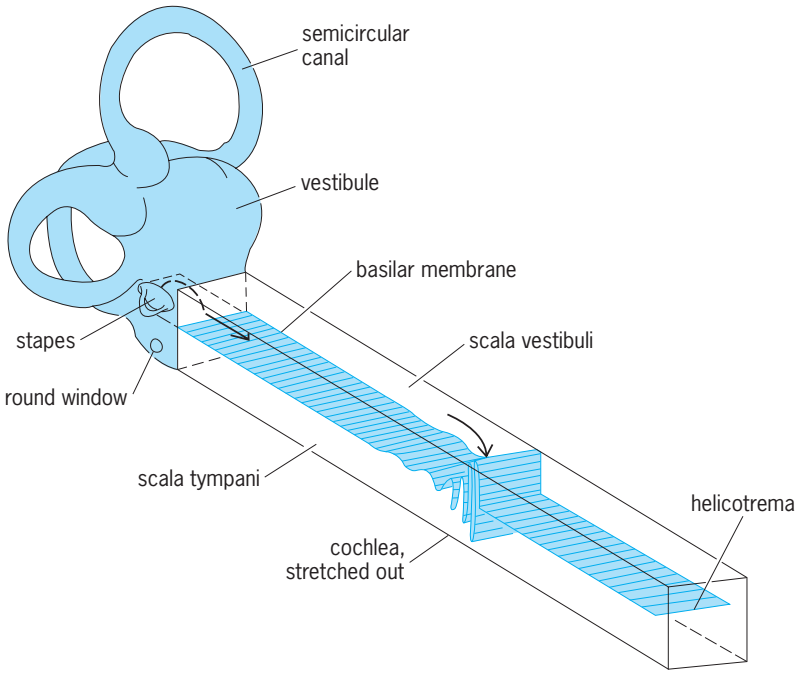

Fig. 2. Traveling wave along the basilar membrane in response to a pure tone that produces a maximum displacement of the membrane in its middle portion. (*After G. Zweig, R. Lipes, and J. R. Pierce, The cochlear compromise, J. Acoust. Soc. Amer., 29:1312–1317, 1957*)

stereocilia remain in the hair cells of the adult mammalian cochlea.

The hair cells in the mammalian cochlea function as mechanoreceptors: motion of the basilar membrane causes deflection of the hairs, starting a process that eventually results in a change in the discharge rate of the nerve fiber connected to each hair cell. This process includes the release of a chemical transmitter substance at the base of the hair cells that controls the discharge rate of the nerve fiber (**Fig. 4**). The hair cells also move when subjected to an electrical field, a property that may be the basis for the positive feedback in the cochlea.

Frequency tuning in auditory nerve fibers. Because of the frequency selectivity of the basilar membrane, individual auditory nerve fibers respond in accordance with the frequency of a sound; therefore, the frequency selectivity of the basilar membrane can be assessed by recording the neural impulses from single nerve fibers (**Fig. 5**). Tuning curves can be obtained by using a standard neurophysiological technique far less complex than that required for measuring the vibration of the basilar membrane.

The frequency selectivity of the basilar membrane had long appeared to be much less sharp than the tuning of single auditory nerve fibers, and so it was assumed that the neural transduction process in the cochlea included some unknown mechanism that sharpened the frequency selectivity of the basilar membrane. With the development of more refined methods for quantifying small vibration amplitudes, measurements could be made in living, anesthetized animals. Such measurements revealed that the frequency selectivity of the basilar membrane is in fact as sharp as neural tuning. Such studies have also shown that the frequency analysis that takes place in the cochlea is far more complex than was believed earlier, and some kind of positive feedback probably improves the frequency selectivity of the basilar membrane. This positive feedback in the cochlea may be compared with that used in early radio receivers to increase their frequency selectivity and so improve differentiation between two radio stations. The positive feedback in the cochlea requires biological energy and is therefore active only in the living, intact cochlea, which explains why von Békésy could not observe it in his studies. Because the effect

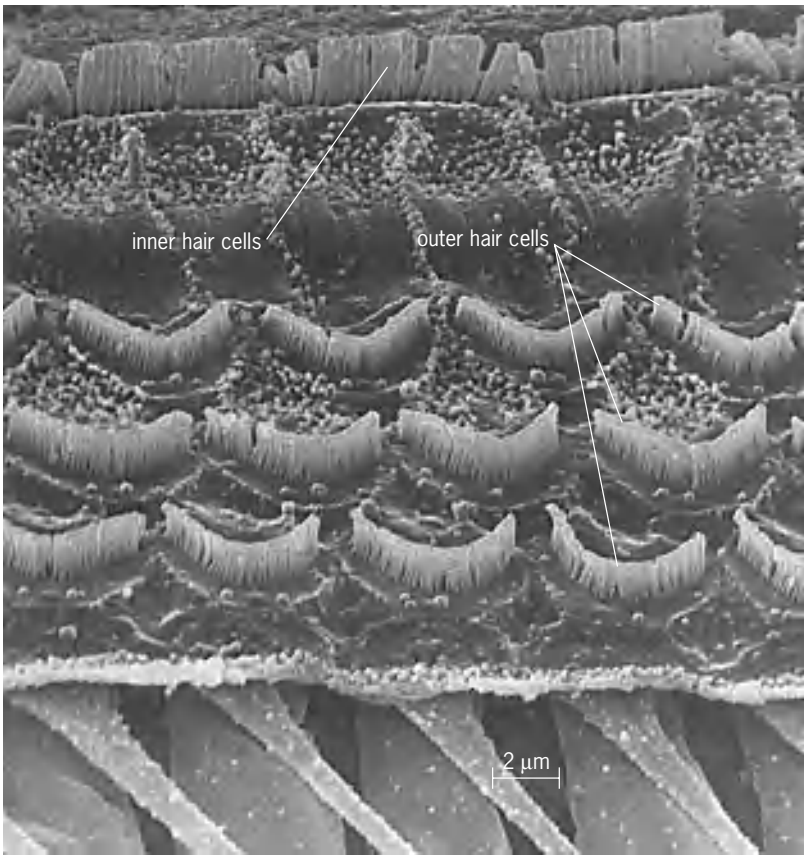

Fig. 3. Electron micrograph of hair cells along the basilar membrane. (*From A. R. Møller, Auditory Physiology, Academic Press, 1983*)

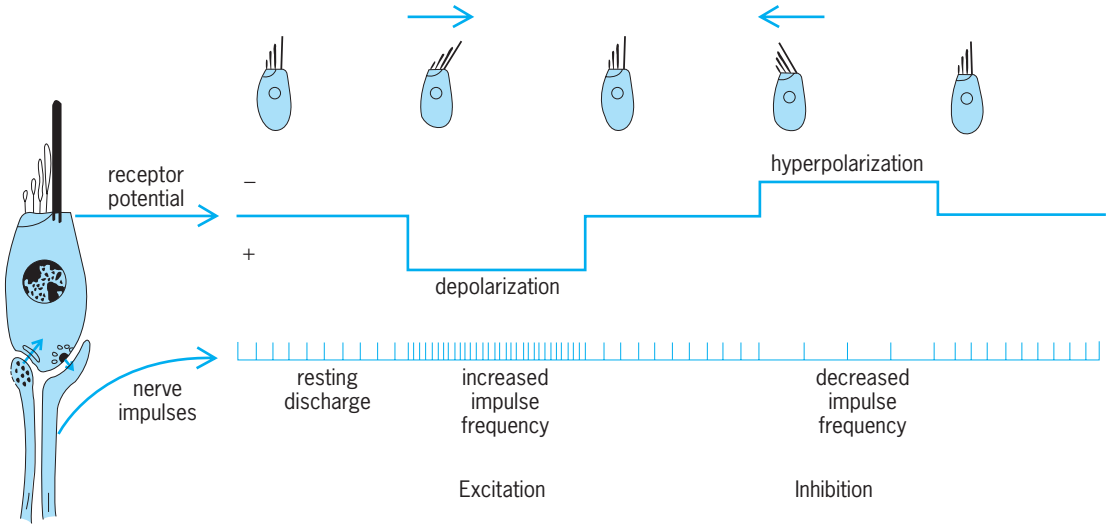

Fig. 4. Schematic illustration of the excitation in hair cells. A deflection of hairs in one direction results in an increase in the neural discharge rate in the nerve fiber associated with the hair cell, whereas a deflection in the opposite direction causes inhibition, or slowing, of this discharge rate. (*After A. Fiock, Transducing mechanisms in lateral line canal organ receptors, Proceedings of the Cold Spring Harbor Symposia on Quantitative Biology, 30:133–146, 1965*)

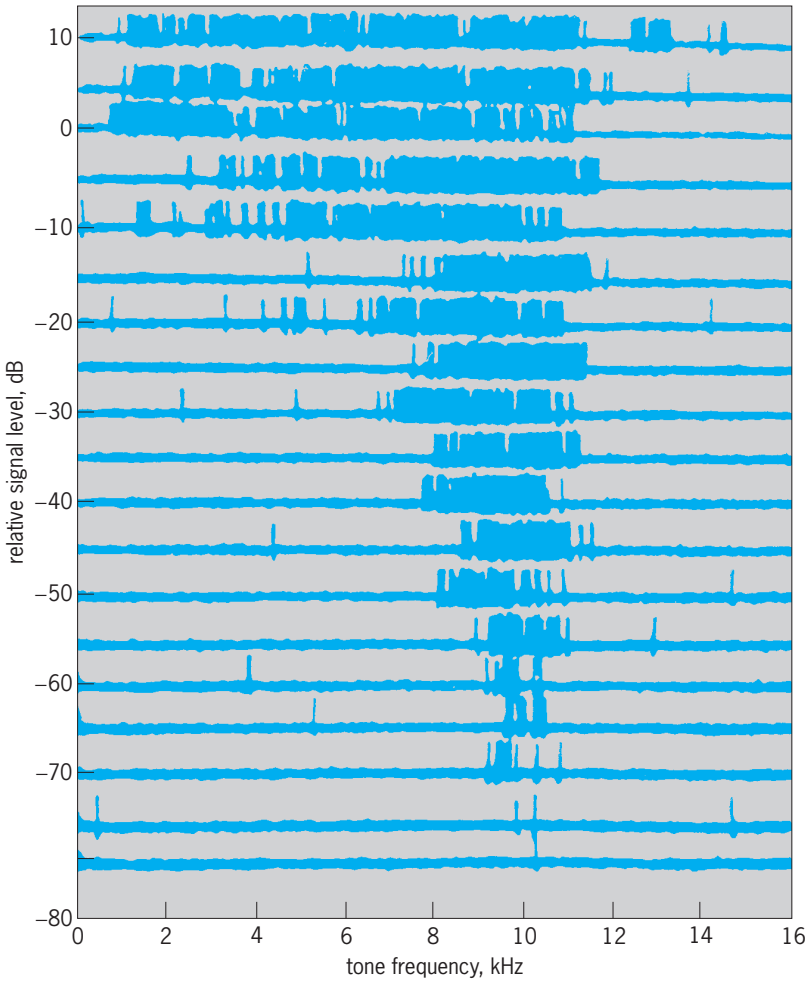

Fig. 5. Neural discharges evoked in a single auditory nerve fiber of an experimental animal exposed to tones of different frequencies. When a continuous tone was presented at different intensities and at variable frequencies, the neural discharges were recorded by using a microelectrode placed close to a single auditory nerve fiber. When the sound intensity is low, the fiber responds only within a narrow range of sound frequencies; as the sound intensity is raised, the nerve fiber responds over a larger and larger range. The contour that surrounds the area of increased neural activity is known as a tuning curve. (*After E. F. Evans, The frequency response and other properties of single fibers in the guinea pig cochlear nerve, J. Physiol. (London), 226:263–287, 1972*)

of positive feedback probably decreases as sound intensity increases, the cochlea has less frequency selectivity for intense sounds than for weak sounds. The basilar membrane therefore acts as a nonlinear filter—unlike most manufactured frequency filters, such as those used in telecommunications, that are usually linear. The nonlinearity of the basilar membrane compresses the sound intensities that reach the ear, thereby extending the range of sound intensities in which the hearing sense can function adequately.

Neural frequency code. The frequency selectivity of the basilar membrane provides the central nervous system with information about the frequency or spectrum of a sound, because each auditory nerve fiber is "tuned" to a specific frequency. However, the frequency of a sound is also represented in the time pattern of the neural code, at least for frequencies up to 5 kHz, because the probability of the occurrence of a nerve impulse is highest during a certain phase of motion of the basilar membrane. This means that the time at which activity occurs in auditory nerve fibers provides information to the central nervous system about the frequency of a sound. Thus, the frequency or spectrum of a sound can be coded for place and time in the neural activity in the auditory nervous system. Both ways of coding the frequency or spectrum of a sound require neural processing in order to be interpreted, and although the precise mechanism remains unclear, psychoacoustic experiments reveal that the auditory system is capable of such an analysis. *See* AUDIOMETRY; PITCH.

Auditory nervous system. The ascending auditory nervous system consists of a complex chain of clusters of nerve cells (nuclei), connected by nerve fibers (nerve tracts). The chain of nuclei relays and transforms auditory information from the periphery of the auditory system, the ear, to the central structures, or auditory cortex, which is believed to be associated with the ability to interpret different sounds.

The four main nuclei in the ascending auditory pathway are not connected in a simple series. Rather, each nucleus relays information not only to the next higher nucleus but also to other nuclei in the chain. The complexity of the auditory nervous system is evidenced by the first relay station of the ascending auditory nervous system, the cochlear nucleus: each auditory nerve fiber connects with nerve cells in at least three divisions of the cochlear nucleus. Such parallel processing, a prominent feature of the auditory nervous system, is not confined to the right or left side: there are connections between the nuclei of the two sides. A descending auditory nervous system, which connects nuclei above with nuclei at lower levels, is also present.

When pure tones are used as stimuli to study the responses from individual neurons in the various nuclei of the ascending auditory nervous system, it is evident that individual neurons respond selectively to tones with frequencies in a discrete range. Such frequency tuning of nerve cells is similar to the frequency specificity of single auditory nerve fibers that is described above. However, the shapes of the tuning curves of nerve cells in the nuclei of the ascending auditory pathway differ from those of single auditory nerve fibers, with the difference being less marked for more peripherally located nuclei such as the cochlear nucleus.

Neurons in the entire auditory nervous system are, in general, organized anatomically according to the frequency of a tone to which they respond best, which suggests a tonotopical organization in the auditory nervous system and underscores the importance of representations of frequency in that system. However, when more complex sounds were used to study the auditory system, qualities of sounds other than frequency or spectrum were found to be represented differently in different neurons in the ascending auditory pathway, with more complex representation in the more centrally located nuclei. Thus, the response patterns of the cells in each division of the cochlear nucleus are different, which indicates that extensive signal processing is taking place. Although the details of that processing remain to be

determined, the cells appear to sort the information and then relay different aspects of it through different channels to more centrally located parts of the ascending auditory pathway. As a result, some neurons seem to respond only if more than one sound is presented at the same time, others respond best if the frequency or intensity of a sound changes rapidly, and so on.

Another important feature of the ascending auditory pathway is the ability of particular neurons to signal the direction of sound origination, which is based on the physical differences in the sound reaching the two ears. Certain centers in the ascending auditory pathway seem to have the ability to compute the direction to the sound source on the basis of such differences in the sounds that reach the ears.

Knowledge of the descending auditory pathway is limited to the fact that the most peripheral portion can control the sensitivity of the hair cells. *See* SIGNAL DETECTION THEORY.

Cortical representation of sounds. The human auditory cortex, unlike that of many mammals, is buried deep in the temporal lobe and is, therefore, not easily accessible for study. As a result, most knowledge about cortical representation of sounds has been gained through experiments that use animals. Early in the history of neurophysiological research, it became evident that the surface of the cortex responds to simple sounds such as tones because it is tonotopically organized. However, with more complex sounds, complex maps better describe the cortical representation of sounds because they delineate the response not only to the frequency of a single tone but to many other qualities that may be related to behavior. Perhaps the best-known of such maps are those generated by the navigational sounds of the flying bats, which use their auditory systems much like radar for navigational purposes. The bat emits a sound and then listens for the echo, using the delay between emission of the sound and detection of the echo to determine the distance to an object. Actually, each animal emits a complex sound that contains components with rapidly changing frequencies and a series of harmonics. Analysis of the harmonic content of the echo helps to identify an animal's unique echo among those of other bats flying nearby. *See* ECHOLOCATION; PSYCHOACOUSTICS.

Neuronal computation of the bat's echoes has been shown to lead to specific cortical maps in which features relevant to navigation are represented on the surface of the cortex. These features are derived from the physical characteristics of a sound, which in this case are the echoes of the sounds emitted by the animal itself. Such features as the distance to obstacles or to a target (prey) and pure navigational data such as velocity and direction to a target were all represented in maps on the animal's auditory cortex. Maps that show how behavioral features appear for sounds that have meaning for other animals, such as humans, have yet to be devised.

Acoustic reflexes. Contraction of the middle-ear muscles in response to sound is involuntary and is called the acoustic middle-ear reflex. It is mediated through connections from the lower levels of the ascending auditory pathway to the motonuclei of the facial nerve (stapedius muscle) and trigeminal motonucleus (tensor tympani muscle). Other, more diffuse acoustic reflexes are the pinna reflex, prominent in some animals, and the startle reflex, which causes contractions of many skeletal muscles. *See* PHYSIOLOGICAL ACOUSTICS.

Auditory evoked potentials. Most knowledge about auditory nervous system function comes from studies of the discharge patterns of single nerve fibers or single nerve cells, which are recorded by placing microelectrodes near a single nerve fiber or nerve cell to record the electrical activity of that particular fiber or cell. Such recorded action potentials are identical waveforms, so that only the time of occurrence of a discharge is significant. Electrical activity of the nervous system can be studied, as well, by recording evoked potentials or compound action potentials, so named because they represent electrical activity in many hundreds or thousands of nerve cells or nerve fibers. By placing an electrode on the surface of those structures, electrical activity generated in specific parts of the nervous system can be recorded. Evoked potentials can also be recorded at a distance from these neural generators, that is, auditory evoked potentials can be recorded from electrodes on the scalp. Although evoked potentials provide

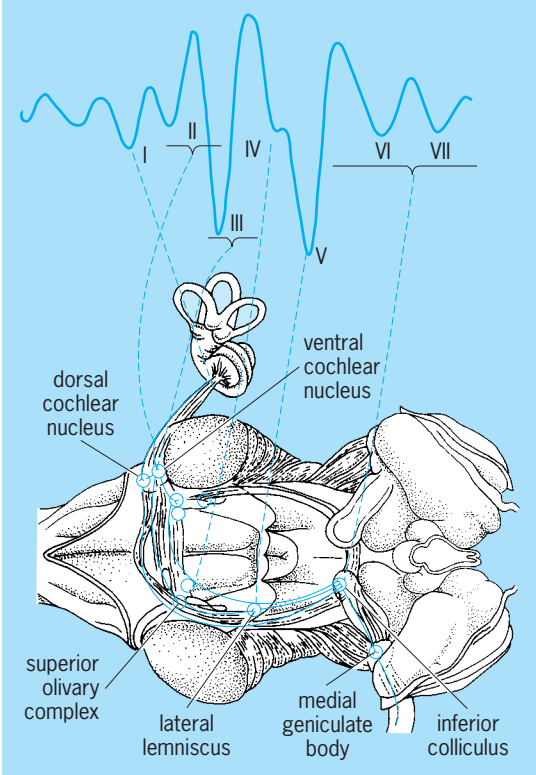

Fig. 6. Main generators of the five earlier peaks in the brainstem auditory evoked potential. (*After A. R. Møller and P. J. Janetta, Neural generators of the auditory brainstem response, in J. T. Jacobson, ed., The Auditory Brainstem Response, College Hill Press, 1984*)

only indirect information about the function of the auditory nervous system, they can be recorded noninvasively. Auditory evoked potentials recorded from the scalp, known as far-field potentials, have very small amplitudes, but repeating the same stimulus many times and adding the responses makes it possible to separate those potentials from background noise. The exact origin of the different components of the far-field potentials from the ascending auditory pathway, known as brainstem auditory evoked potentials or BAEP, is unclear, but sufficient information about their origin has been gathered (**Fig. 6**) to make them important tools in the diagnosis of disorders that affect the ear and the auditory nervous system. *See* HEARING IMPAIRMENT; LOUDNESS; MASKING OF SOUND; PHONORECEPTION; SOUND; SPEECH PERCEPTION.

Aage R. Møller

Bibliography. R. V. Harrison, *The Biology of Hearing and Deafness*, 1988; R. D. Luce, *Sound and Hearing*, 1992; B. C. Moore, *Frequency Selectivity in Hearing*, 1986; J. O. Pickels, *An Introduction to the Physiology of Hearing*, 2d ed., 1988; W. A. Yost, *Fundamentals of Hearing: An Introduction*, 4th ed., 2000.

Hearing (vertebrate)

The ability to perceive sound arriving from distant vibrating sources through the environmental medium (such as air, water, or ground). The primary function of hearing is to detect the presence, identity, location, and activity of distant sound sources. Sound detection is accomplished using structures that collect sound from the environment (outer ears), transmit sound efficiently to the inner ears (via middle ears), transform mechanical motion to electrical and chemical processes in the inner ears (hair cells), and then transmit the coded information to various specialized areas within the brain. These processes lead to perception and other behaviors appropriate to sound sources, and probably arose early in vertebrate evolution.

Sound. Sound is gathered from the environment by structures that are variable among species. In many fishes, sound pressure reaching the swim bladder or another gas-filled chamber in the abdomen or head causes fluctuations in volume that reach the inner ears as movements. In addition, the vibration of water particles that normally accompany underwater sound reaches the inner ears to cause direct, inertial stimulation. In land animals, sound causes motion of the tympanic membrane (eardrum). In amphibians, reptiles, and birds, a single bone (the columella) transmits tympanic membrane motion to the inner ears. In mammals, there are three interlinked bones (malleus, incus, and stapes). Mammals that live underground may detect ground-borne sound via bone conduction. In whales and other sea mammals, sound reaches the inner ears via tissue and bone conduction.

Auditory anatomy. The inner ears of all vertebrates contain hair-cell mechanoreceptors that transform motion of their cilia to electrochemical events resulting in action potentials in cells of the eighth cranial nerve. Patterns of action potentials reaching the brain represent sound wave features in all vertebrates.

In fishes the auditory portions of the ears are otolith organs containing a group of hair cells and a solid calcium carbonate otolith which is in contact with the hair-cell cilia. The sacculus is the major auditory organ in most fishes, but other organs (utricle and lagena) may function in hearing as well. In mammals these otolith organs serve the sense of balance.

Most amphibians have two auditory organs: the amphibian and basilar papillae that respond to lower and higher frequencies, respectively. The auditory organ of reptiles and birds is the basilar papilla situated on the basilar membrane.

The mammalian cochlea contains hair cells in the organ of Corti on the basilar membrane. The cochlea varies in length from 10 mm (in some rodents) to 60 mm (elephant). The inner ears of mammals and birds perform a mechanical sorting of frequencies through variations of basilar membrane stiffness from base to tip. In some amphibians, reptiles, and birds, frequency selectivity may also depend on the electrochemical mechanisms of hair-cell membranes. Eighth nerve cells are selective for different sound frequencies. All vertebrates have an analogous set of auditory brain centers. *See* EAR (VERTEBRATE); PHYSIOLOGICAL ACOUSTICS.

Hearing abilities. Experiments show that vertebrates have more commonalities than differences in their sense of hearing. The major difference between species is in the frequency range of hearing, from below 1 Hz to over 100,000 Hz. In other fundamental hearing functions (such as best sensitivity, sound intensity and frequency discrimination acuity, time and frequency analysis, and source localization), vertebrates have much in common. The hearing abilities of over 180 species from all vertebrate classes have been investigated (see **illus.**). All detect sound within a restricted frequency range. Audiograms graphically display the lowest sound intensities that an animal can hear (threshold) for the range of audible frequencies. All species are able to detect sounds in the presence of interfering sounds (noise), discriminate between different sound features, and locate the sources of sound with varying degrees of accuracy.

Fishes. Fishes with greatest sensitivity and frequency range are the "hearing specialists." These species have a mechanical link between the gas bladder and the ear (for example, goldfish and catfish). Fishes without a gas bladder (such as flatfish) detect only acoustic particle motion. The best sensitivity falls between 100 and 1000 Hz. The most sensitive fishes have the best sensitivity in the range between −20 and 0 dB.

Amphibians and reptiles. There are very few audiograms for amphibians and reptiles because it is difficult to train these animals. The two lowest curves in illus. *b* were obtained for two frogs using conditioning methods. The audiogram for a turtle (broken

line) shows poor sensitivity. For amphibians, best sensitivity approaches 0 dB.

Birds. The audiograms for 20 bird species are similar in sensitivity and frequency range, with the best sensitivity between −15 and 25 dB and the best frequency between 1000 and 6000 Hz. The barn owl is exceptional with best hearing between 2000 and 9000 Hz.

Mammals. The audiograms for 57 mammals show great variation in frequency range. The low-frequency limit varies from 20 to 10,000 Hz. The high-frequency limit varies from 11,000 to 150,000 Hz. No single species hears within this entire frequency range. Some marine mammals and bats have the best high-frequency hearing. Some rodents, elephants, and primates hear well at lower frequencies. Most mammals are similarly sensitive at their best frequency of hearing.

The sensitivity range is similar among all groups, with some species in all groups having a best sensitivity in the region of −20 to 0 dB. Fishes, amphibians, reptiles, and birds hear best between 100 and 5000 Hz. Only mammals hear at frequencies above 10,000 Hz. Humans and elephants have the poorest high-frequency hearing.

Sound intensity discrimination. All vertebrates have the ability to discriminate between sounds that differ only in sound level or intensity, but the acuity of this discrimination varies among species (1 to 3 dB differences can be discriminated). Level discrimination thresholds are approximately the same at all frequencies and intensities.

Temporal pattern. Most natural sounds vary over time in level or intensity. These patterns are important for the detection, identification, and classification of sources. The threshold for detecting a brief sound increases with sound duration in all vertebrates. At durations greater than 100 to 400 milliseconds, sensitivity reaches a maximum value. Features with durations as brief as 1 to 4 ms can be resolved by birds and mammals. Animals trained to discriminate between sounds differing only in duration can discriminate differences of 5–10%.

Sound source localization. All vertebrates are able to determine the direction of sound sources in the horizontal and vertical directions with some degree of accuracy. In land animals the information used for localization in the horizontal plane comprises differences in the times and intensities at which sound features reach the two ears.

Fishes localize sound sources with an accuracy of 10–20°, but the mechanisms for localization are not well understood. Time and intensity differences at the two ears are too small to be useful. However, the angle of acoustic particle motion is represented in eighth nerve cells due to the directional selectivity of the otolith organs and their hair cells.

Many amphibians, reptiles, and birds have close-set ears, so that time and intensity differences reaching the two ears are reduced. These animals localize sound sources in the horizontal and vertical planes with an accuracy of 10–20°. In some species, there are specialized mechanisms that exaggerate small time and intensity differences reaching the two ears.

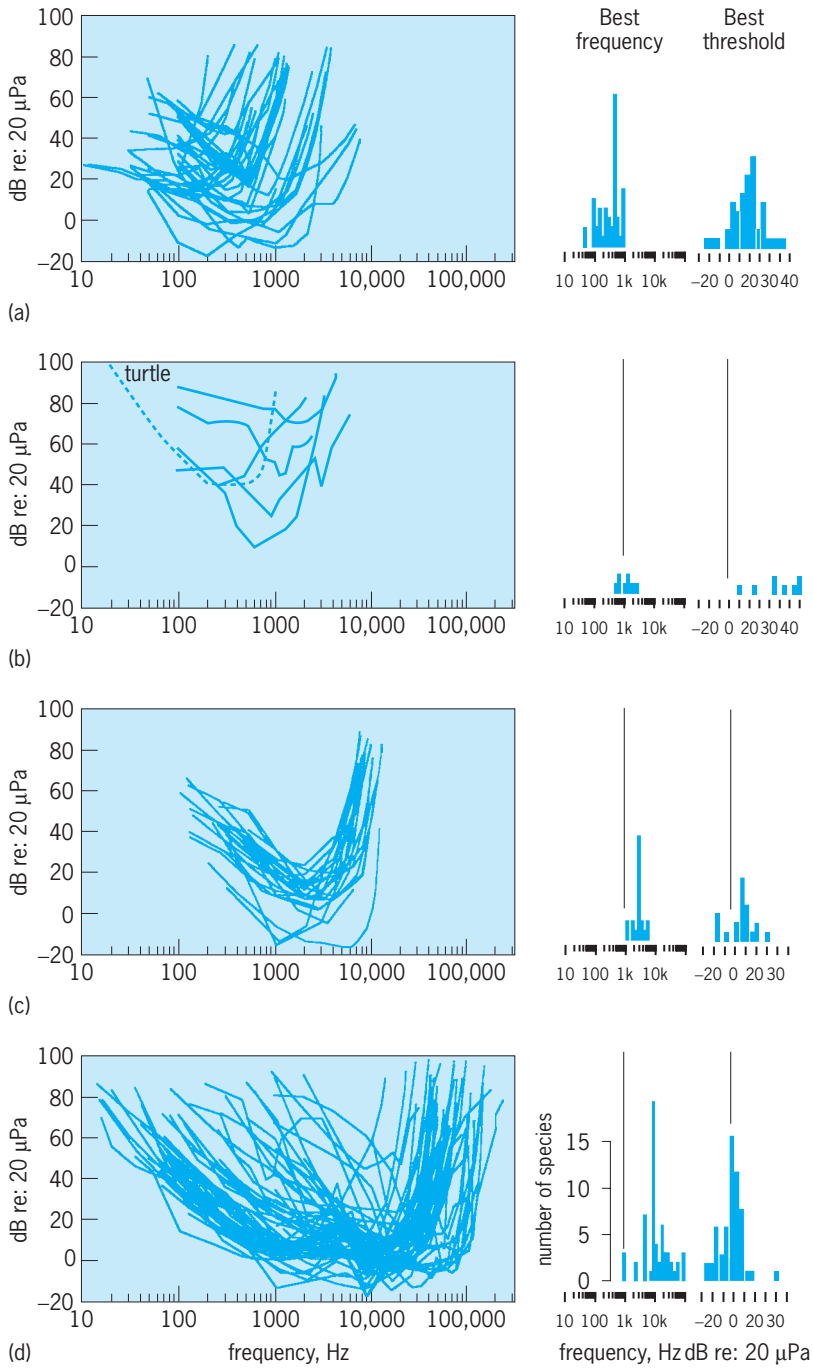

Behavioral audiograms for fishes ($n = 49$), amphibians ($n = 8$), a reptile ($n = 1$), birds ($n = 20$), and mammals ($n = 57$). Thresholds are expressed in decibels with respect to 20 micropascals. Thresholds for aquatic animals tested underwater are the sound pressure levels in air corresponding to the sound intensity levels underwater calculated from underwater sound pressure thresholds assuming far-field conditions. On the right are shown frequency distributions of best frequency and best, or lowest, threshold obtained from the audiograms. [*Modified after R. R. Fay (1988), with permission, where the references and tabled data for all audiograms can be found*]

Mammals with small heads also have reduced time and intensity difference cues for low frequencies. These animals rely on intensity difference cues for high-frequency sounds. Mammals with large heads and good sensitivity at low frequencies use time difference cues for horizontal localization. Acuity varies among mammals between 1 and 20°. Humans and

porpoises have the most acute localization abilities.

Frequency processing. All vertebrates tend to break down complex sounds into their constituent frequency components. This process enhances sound detection and is important for sound source identification and localization. Acuity in frequency selectivity includes the minimum discriminable difference between pure-tone frequencies, and measures of the frequency range within which one sound may interfere with, or mask, the detection of another sound.

Minimum discriminable frequency differences decline as sound level is raised, but grow larger in proportion to the frequency at which the discrimination is measured. Detecting the presence of a pure tone in the presence of masking noise bands of different center frequency and frequency range indicates that masking interference occurs only in a restricted frequency range surrounding the signal (the critical band).

For all vertebrates, the estimates of frequency resolution tend to be increasing, parallel functions of frequency. The greatest frequency acuity occurs in animals having poor high-frequency hearing and long cochleas (such as the human and elephant). These hearing functions likely have had a long and stable evolutionary history. Thus, all vertebrates have solved common problems of sound perception. *See* AUDIOMETRY; LOUDNESS; MASKING OF SOUND; PHONORECEPTION; SOUND. Richard R. Fay

Bibliography. R. R. Fay, *Hearing in Vertebrates: A Psychophysics Databook*, Hill-Fay Associates, Winnetka, IL, 1988; R. R. Fay, Structure and function in sound discrimination among vertebrates, in D. Webster, R. Fay, and A. Popper (eds.), *The Evolutionary Biology of Hearing*, Springer-Verlag, New York, 1992; R. S. Heffner and H. E. Heffner, Evolution of sound localization in mammals, in D. Webster, R. Fay, and A. Popper (eds.), *The Evolutionary Biology of Hearing*, Springer-Verlag, New York, 1992; W. C. Stebbins and M. A. Berkeley (eds.), *Comparative Perception: Complex Signals*, Wiley, New York, 1990.

Hearing aid

A device that amplifies sound for someone with a hearing loss. A typical aid consists of a microphone, an electronic amplifier, an earphone (called a receiver), and a battery. The sound is coupled to the ear with an earmold. The great majority of earmolds are cast from impressions of the individual patient's ear canal, but over-the-counter and mail-order hearing aids use prefabricated earmolds, commonly offered in several sizes. In the case of in-the-ear hearing aids, all components are small enough to be contained in an earmold shell that fits within the ear canal and concha, or even within the canal alone (**Fig. 1**). *See* AMPLIFIER; EAR (VERTEBRATE); EARPHONES; MICROPHONE.

Development. Pre-electric ear trumpets provided acoustic gain as early as the late eighteenth century. By the early 1920s, hearing aids containing miniature vacuum tubes could provide an amount of amplification limited only by the quality of the seal between the earmold and the ear (to prevent whistling caused by feedback). By 1962 the transistor plus subminiature microphones and earphones had made the components small enough to produce all-in-the-ear hearing aids. *See* TRANSISTOR.

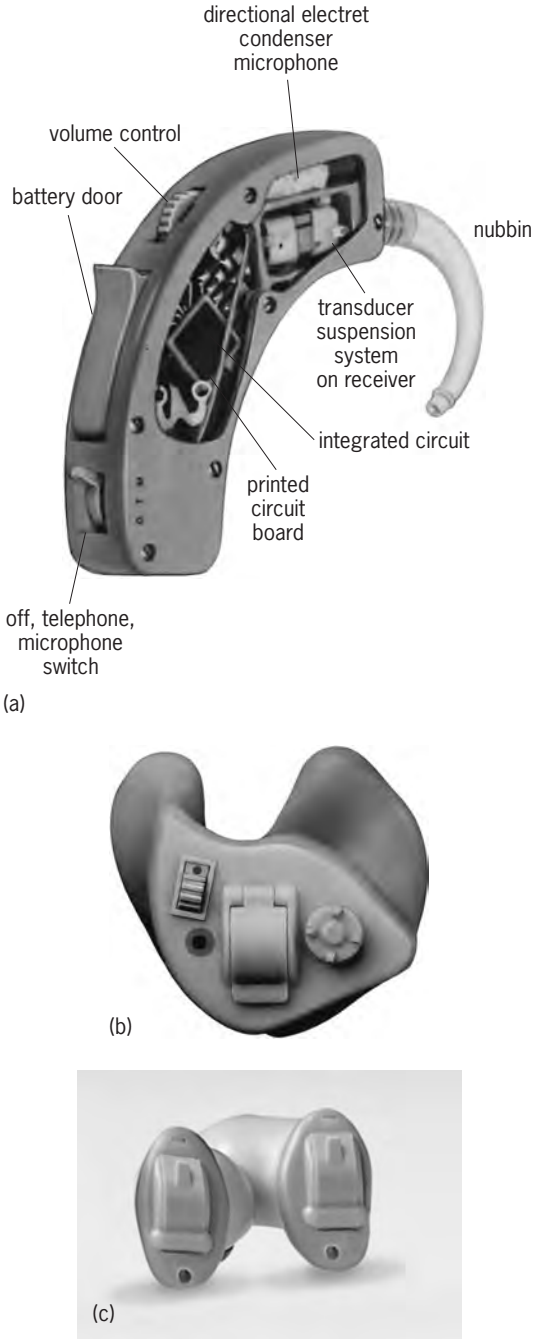

Fig. 1. Hearing aids. (*a*) Over-the-ear aid with case open to show components, and (*b*) in-the-ear aid (*Beltone USA, Glenview, IL*). (*c*) In-the-canal aid (*Siemens Hearing Instruments, Piscataway, NJ*).

By 1989, hearing-aid designers had come to understand that they needed to take the transmission characteristics of the ear into account to create normal frequency response at the eardrum, and that

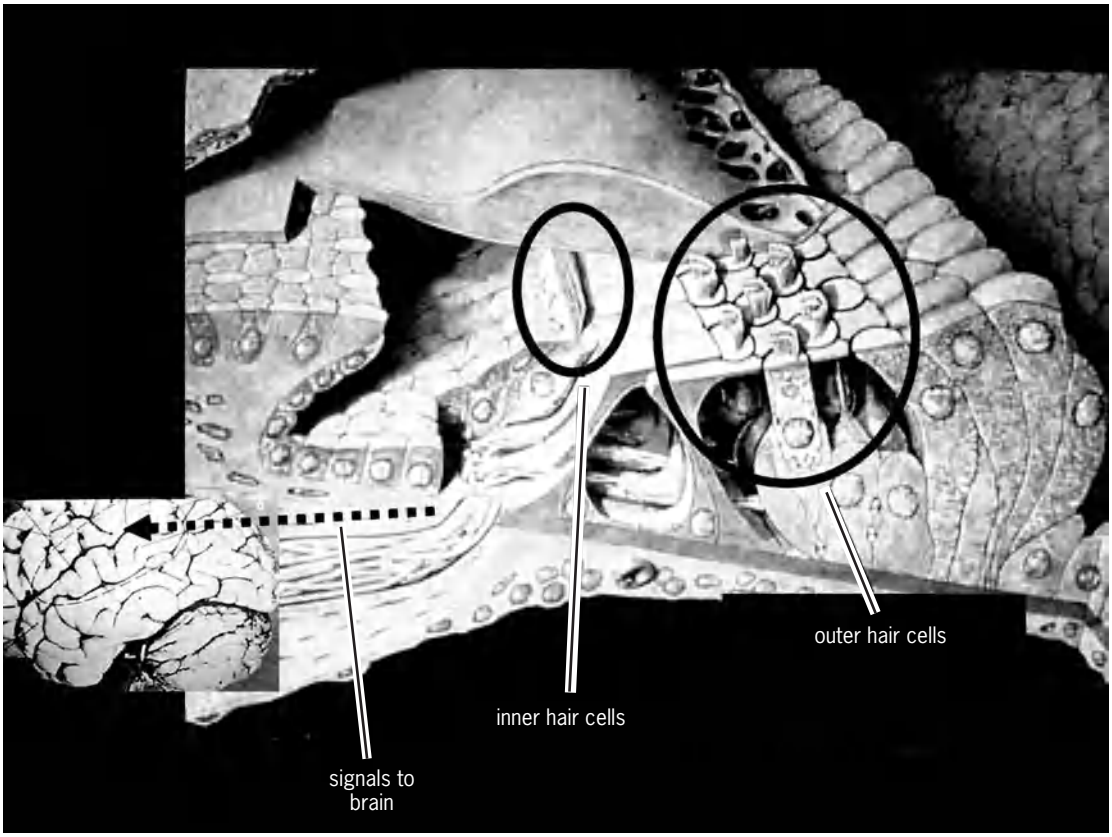

Fig. 2. Inner and outer hair cells found on the basilar membrane of the inner ear. Tiny hairs, or stereocilia, are visible on the top of each cell. Inner hair cells provide all the signals to the brain. Outer hair cells amplify quiet sounds.

hearing loss is typically accompanied by hearing distortions that need to be compensated by electronic processing of the signal. This understanding, combined with integrated circuits, high-quality miniature microphones and receivers, zinc-air batteries, and high-efficiency class D amplifiers, made it practical to produce hearing aids with a fidelity (exclusive of the processing) equal to the best home stereo systems. *See* BATTERY; INTEGRATED CIRCUITS.

Types of hearing loss. Hearing loss caused by abnormalities in the mechanical parts of the ear is called conductive; that created by loss of hair cells of the cochlea is called cochlear or sensorineural; and that created by pathology of the eighth nerve (much less common) is called neural. Conductive hearing loss is effectively overcome with surgery or hearing aids, while the latter two types of loss present a greater challenge for hearing aids.

In the normal inner ear, approximately 15,000 outer hair cells provide some 30–50 dB of mechanical amplification for quiet sounds, while approximately 5000 inner hair cells convert sound-induced mechanical energy into nerve impulses that go to the brain (**Fig. 2**). A major cause of damage to these hair cells is excessive exposure to loud noise and music. A loud ringing in the ear after such an exposure is an indication of likely death of some of the cells. The loss of hair cells causes two symptoms: a loss of hearing sensitivity to quiet sounds, and a loss of ability to understand speech in noise. The two symptoms can occur in various ratios. When loss of sensitivity is the dominant symptom, hearing aids can restore nearly normal hearing by amplifying these quiet sounds without overamplifying louder sounds, and by providing extra gain for those high-frequency sounds made inaudible by the typical high-frequency hearing loss. A signal-processing system called wide-dynamic-range compression (WDRC) does all of that. Although this system was widely criticized when it was introduced, it is now used in almost all analog and digital hearing aids. *See* SIGNAL PROCESSING.

Unfortunately, a loss of speech cues available to the brain often accompanies a loss of inner (as opposed to outer) hair cells, and the loss of redundant cues makes it difficult for the listener to understand speech in noise. This difficulty can be helped by increasing the signal-to-noise ratio in several ways: using directional-microphone hearing aids, leaning closer to the talker, using a microphone held close to the talker, or using a system of microphones worn by each of several talkers (a development that can improve the signal-to-noise ratio by a factor of 10). *See* HEARING (HUMAN); HEARING IMPAIRMENT.

User satisfaction. Surveys show that a nearly constant 60% of hearing-aid users are satisfied overall, with no increase of satisfaction from analog to digital aids. The dissatisfied 40% include those whose hearing aids are not properly fitted or are inappropriate to the loss; those whose earmold fit does not

prevent feedback or prevent their own voices from sounding hollow (the latter effect is called occlusion); those— a small percentage—for whom wearing a hearing aid is disturbing under any circumstances; and most important, those who are unable, even with the hearing aids, to understand speech in the presence of noise or competing speech. For the last group, both analog and digital aids may fail equally by providing audibility but not the ability to separate speech from interference. The surveys show no difference in overall satisfaction between inexpensive analog hearing aids and expensive digital hearing aids, although both analog and digital aids with directional microphones produce higher satisfaction. (As noted above, an even greater improvement in signal-to-noise ratio can be obtained using a microphone held close to the talker or using a system of microphones worn by each of several talkers.)

While many of the claims of the superiority of digital over analog hearing aids have been exaggerated, digital circuits do have potential advantages deriving from their flexibility, and the better ones provide sophisticated feedback reduction that does not simultaneously reduce audibility or intelligibility. From the standpoint of convenience, a digital circuit can analyze a listening environment and adjust the hearing aid to the processing the manufacturer has chosen for that environment (directional versus omni mode, for example), so that the user does not have to readjust the controls. Although many digital hearing aids have not achieved the sound quality of the best analog aids, digital aids have neither an inherent advantage or disadvantage in sound quality compared to analog aids.

Binaural hearing aids. Some years ago the use of binaural hearing aids (an aid in each ear) was looked on with some suspicion, as a way to get people to buy a second aid that offered no advantage. The Federal Trade Commission strictly limited advertising claims for the advantages of binaural hearing aids. It is now recognized that an aid in each ear offers three real advantages: (1) Binaural release from masking: because of the directional cues created by binaural listening, the ability to separate speech from interference is improved by several decibels. (2) Binaural loudness summation: a sound heard by two ears is louder than the same sound heard by one ear. This means that less amplification is required of each hearing aid of a binaural pair, and for the same loudness the physical level of sound going into each ear is less. The danger of feedback, and the danger of distortion created by overload of the mechanical parts of the ear, are thereby reduced. (3) The loss of intelligibility in the unaided ear that often occurs over time from the auditory deprivation is avoided. *See* MASKING OF SOUND.

Mead C. Killion; Edgar Villchur

Bibliography. K. W. Berger, *The Hearing Aid: Its Operation and Development*, National Hearing Aid Society, 1984; M. C. Killion, Guest editorial: Hearing aids, past, present, future—moving toward normal conversations in noise, *Brit. J. Audiol.*, 31:141–148, 1997; M. C. Killion, High fidelity and hearing aids, *Audio*, 75(1):42–44, 1990; S. Kochkin, On the issue of value: Hearing aid benefit, price, satisfaction, and brand repurchase rate, *Hear. Rev.*, 10(2):12–26, February 2003; S. Kochkin, 10-year customer satisfaction trends in the U.S. hearing instrument market, *Hear. Rev.*, 9(10):14–25, 46, October 2002; S. Silman, S. A. Gelfand, and C. A. Silverman, Late-onset auditory deprivation: Effects of monaural versus binaural hearing aids, *J. Acous. Soc. Amer.*, 76:1357–1362, 1984; E. Villchur, A different approach to the noise problem of the hearing impaired, *Amer. J. Audiol.*, 2(2):47–51, July 1993.

Hearing impairment

Any alteration of hearing capacity. Hearing impairment can be of various degrees, including mild, moderate, severe, profound, or total. The degree of impairment typically is categorized by the loss of hearing sensitivity, that is, how loud sounds must be for a listener to hear them. The degree of impairment can refer either to the loss of hearing sensitivity for individual pitches of sounds for each ear separately or to an overall loss of hearing sensitivity for both ears. Hearing impairment is further categorized as unilateral if present in only one ear and as bilateral if present in both ears.

Hearing impairment may be present at birth or acquired later in life. Because hearing serves many functions, from the simple detection of warning sounds to the more complex functions necessary for the perception of speech, the effects of a hearing impairment are quite varied. The most prominent effects depend on a variety of factors, including the degree of the impairment, the pitch range of the impairment, and the stage of language and speech development when the impairment occurred. Congenital hearing loss greatly interferes with normal language and speech development if it is bilateral and of severe or greater magnitude over the pitch range that covers speech sounds. Acquired hearing loss can occur gradually or suddenly at any time of life and therefore can also be defined in relation to the development of language and speech. Hearing impairment is often termed prelingual, perilingual, or postlingual if the hearing loss occurred prior to, during, or after the development of language and speech, respectively.

The term deafness has two meanings. If it refers only to a total lack of hearing function, the term is lowercase. If it refers to an individual with bilateral hearing loss who does not use oral language and speech, the term has an initial capital. Thus, an individual may be deaf in one ear and not Deaf; another individual may be Deaf, yet have some residual hearing ability.

The auditory system (see **illustration**) consists of the outer ear (1), the external ear canal (2), the eardrum (3), the system of bones in the middle ear called the ossicular chain (4) that conducts sound from the eardrum to the inner ear, which consists

of the hearing organ called the cochlea (5), and the vestibular system, which controls balance. The cochlea senses and analyzes sounds and sends corresponding impulses along the auditory nerve (6) to the brain. Hearing loss can be classified by the part of the auditory system that is defective. Conductive hearing loss results from abnormalities or diseases of the outer or middle ear; sensorineural hearing loss results from abnormalities or diseases of the inner ear or auditory nerve; and central hearing impairment results from abnormalities or diseases of the auditory portions of the central nervous system. Combined conductive and sensorineural hearing loss is referred to as mixed hearing loss. *See* AUDIOMETRY; EAR (VERTEBRATE).

Causes. Hearing loss results from a broad range of causes, some of which are reversible, and so hearing loss may be either temporary or permanent.

Conductive abnormalities. Many conditions exist whereby the transmission of sound through the conductive mechanism is impeded. These conductive abnormalities result in a reduction of sound reaching the inner ear, so that sounds must be made louder to be heard. The degree of reduction in sound transmission caused by the abnormality determines the degree of the hearing loss.

External canal abnormalities may be due to ear wax (cerumen), which can become hardened or impacted in the external ear canal. Bacteria, fungi, or other microorganisms may infect the skin of the external ear canal and cause sufficient swelling of the walls to block sound waves. Congenital malformations may also occlude or block the canal.

Middle-ear abnormalities may be associated with a sequence of conditions that often results in otitis media, an infection in the middle-ear cavity that is a common cause of conductive hearing loss, especially in children. This sequence may begin with a common cold, allergy, or upper respiratory tract infection. Nasal secretions can pass to the middle ear through the eustachian tube; if the tube becomes involved and cannot open properly, the surrounding tissues absorb the air in the middle-ear cavity. The higher air pressure in the external ear forces the drum membrane inward, restricting movement of the ossicular chain and inhibiting sound conduction. Watery secretions then form in the middle-ear space and may be thickened by the infection, resulting in suppurative (pus-forming/discharging) otitis media. If the infection does not resolve, the secretions may expand enough to perforate the eardrum and further impede the transmission of sound.

Air travel can be hazardous to the ears for some people. Aero-otitis media begins during flight when the difference in barometric pressure on the two sides of the eardrum is not corrected by swallowing. High-speed movement in elevators and deep-water diving may produce a similar effect.

Another middle-ear abnormality is otosclerosis, a hereditary disease. Portions of the bony capsule surrounding the inner ear, normally the hardest bone in the body, become decalcified and are replaced by a soft spongy bone that enlarges and hardens, often fixing the ossicular chain at the point where the middle-ear conductive mechanism joins the inner ear. The volume of the sound reaching the inner ear becomes greatly reduced. Otosclerosis is a progressive disease that usually begins in youth. The hearing loss is first apparent in adolescence or during the early twenties. It is more common among whites and East Indians than other racial groups, with a slight predominance in females.

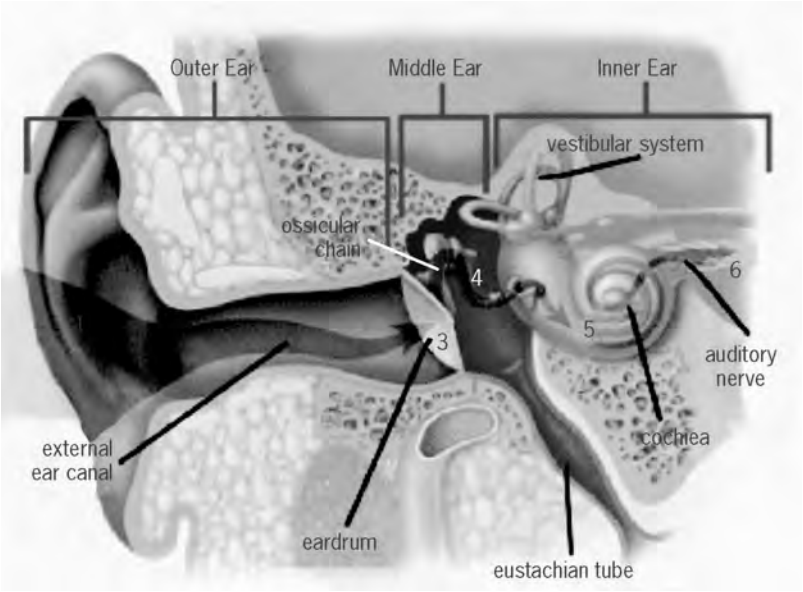

Major portions of the human auditory system.

Sensorineural abnormalities. Many conditions affect the generation of electrical potentials in the inner ear and their transmission by the auditory nerve. These abnormalities result in two different effects on hearing. When hearing sensitivity becomes impaired, sounds must be amplified to be heard. In addition, because the sensory and neural structures are involved (the structures that analyze sound into its components), the sound that is heard becomes distorted.

Inner-ear abnormalities may be due to congenital malformations that may arise independently or in association with other syndromes such as Waardenburg syndrome (which causes hearing loss and pigmentary anomalies) or Usher's syndrome (which causes hearing loss and progressive vision loss). Many drugs and chemicals can be damaging to the inner ear (that is, ototoxic). The effect may be temporary or permanent and depends on the dosage and how they are applied. The types of ototoxic drugs (with examples) include salicylates (aspirin), nonsteroidal anti-inflammatory drugs (naproxen), antibiotics (gentamycin), diuretics (ethacrynic acid), chemotherapeutic agents (cisplatin), and quinine. Inner-ear damage can be caused by bacterial infections such as meningitis and viral infections such as measles and mumps. Loud noise is an insidious environmental cause of inner-ear damage; it may be in the form of either a sudden explosive sound

such as a blast, or a prolonged exposure. Initially, noise exposure results in a temporary loss of hearing for the high pitches, known as temporary threshold shift. The condition may become permanent with habitual exposure. Damage is cumulative and usually occurs without pain. *See* ACOUSTIC NOISE; LOUDNESS.

Aging is one of the most common causes of inner-ear abnormalities, but it also affects other parts of the auditory system. Hearing loss associated with the aging process is called presbycusis, and frequently leads to complaints by the older person that other people do not speak distinctly. This condition results from the loss of sensory cells of the inner ear, particularly those associated with the hearing for high pitches that are characteristic of certain speech sounds. Auditory nerve abnormalities are occasionally caused by a tumor within the skull which can press against the auditory nerve and disturb its function.

Central nervous system abnormalities. Even though the hearing mechanism is sending the correct signals to the brain, the auditory pathways in the central nervous system can experience abnormalities that result in hearing impairments. Some causes of central nervous system deficits that may affect hearing include degenerative diseases such as multiple sclerosis, cerebral vascular accidents from hemorrhage or clotting, birth deformities, and brain trauma. *See* NERVOUS SYSTEM (VERTEBRATE).

Nature of hearing impairment. Conductive hearing loss impairs sensitivity to sound; if the sound is amplified, the impairment can usually be overcome. In noisy places where speaking more loudly is an automatic response, a person with mild conductive hearing loss, who hears less background noise and has learned to listen carefully for quiet speech, may actually have an advantage over a person with normal hearing. Most of these conditions can be reversed with medical or surgical intervention.

In sensorineural hearing loss, the abnormality can affect different portions of the inner ear so that often hearing sensitivity may be normal or nearly so for low-pitched tones but falls off sharply for higher tones. In addition, what is heard may be distorted. Persons with mild-to-moderate sensorineural hearing loss may hear the louder low-pitched portions of speech such as vowels but may not hear, or hear with distortion, quieter, high-pitched speech such as the voiceless consonants (that is, consonants produced without sound from the vocal cords). The person thus has the disquieting sensation of hearing speech but not understanding it. When hearing loss occurs gradually, the adult learns to accommodate with supplemental speech reading (lip reading) and filling in from the context of the conversation. The speech of the hearing-impaired person may deteriorate over time because the high-pitched speech sounds cannot be heard. A child born with sensorineural hearing loss that is undiagnosed or untreated may fail to acquire intelligible speech and may wrongly be thought inattentive, developmentally delayed, or ill behaved. *See* SPEECH DISORDERS; SPEECH PERCEPTION.

In most cases of sensorineural or mixed hearing loss, both sensitivity and clarity are impaired, so the perceived sound is weak and distorted. A hearing aid that amplifies sound fails to correct the distortion, and so the difficulty of understanding speech is only partially remedied. With severe or greater loss, the listener may hear only the cadence and the intensity variations of speech without being able to discriminate all the speech sounds. Speech deteriorates for adults who suffer this type of hearing loss after speech has developed normally. For children born with severe or profound bilateral sensorineural hearing loss, special education is required for the acquisition of speech and language. *See* HEARING AID; PSYCHOACOUSTICS.

Other symptoms characterize sensorineural hearing loss. Tinnitus head noise or ringing in the ears that is heard with no related acoustic stimulus may occur continuously or intermittently, and is described as a rushing or roaring noise. Vertigo and nausea may accompany the hearing deficit if the abnormality affects the vestibular system, as in Ménière's disease (a disorder of the inner ear causing fluctuating hearing loss, vertigo, tinnitus, and a sensation of pressure in the affected ear). In double hearing, or diplacusis, a single tone is heard at a different pitch in each ear, or simple tones may sound fuzzy or rough. Loudness recruitment, an abnormal increase in perceived loudness as a sound is intensified, may be associated with a low tolerance for loud noises.

Treatment and management. Prevention is the best approach to possible loss of hearing. Universal neonatal hearing screening programs are often mandated and have proven very successful. Many newborns have their hearing screened before being discharged from the hospital nursery. Though such programs do not prevent hearing loss, they do allow for early intervention that then can lead to normal language and speech development. Desirable preventive measures include immunization for viral and bacterial diseases, early medical intervention for upper respiratory infections or earaches, keeping both nostrils open when blowing the nose, control of allergies, avoidance of ototoxic drugs, and reduced exposure to loud sounds. Routine hearing tests throughout life are also advisable.

Antibiotic control of ear infections is a common treatment for conductive hearing loss, and surgical procedures are also available. Myringotomy is a simple surgical perforation of the tympanic membrane to drain fluids from the middle ear. Tympanoplasty, a more extensive surgical procedure, can eradicate middle-ear infection. A stapedectomy is the reconstruction of the ossicular chain that has been fixed in place in the oval window of the inner ear by otosclerosis. Under local anesthetic, the footplate of the third bone in the ossicular chain, the stapes, is removed and replaced with a graft of connective tissue and a pistonlike prosthesis that connects the flexible

oval-window graft to the remainder of the ossicular chain. Many of these procedures for conductive hearing loss allow hearing to return to normal or near-normal levels.

Sensorineural hearing impairment is generally permanent and not amenable to medical intervention. For sensorineural hearing losses that have not yet become permanent, or that fluctuate or that are the result of processes that may be reversible, the measures to correct the hearing loss are varied. Tranquilizing and antivertiginous drugs, vasodilative agents, and low-salt diets are among the treatments prescribed for the symptomatic management of Ménière's disease, but they have met with only moderate success. Various medications are sometimes used with certain sudden hearing loss, again with only moderate success. The lack of widely accepted and effective pharmaceutical approaches to sensorineural hearing loss has placed an emphasis on the use of a variety of devices to correct for hearing impairment.

Cochlear implants. For permanent severe or profound bilateral sensorineural hearing loss, the cochlear implant seeks to restore hearing by direct electrical stimulation of the auditory nerve, thereby producing an auditory sensation that represents environmental sounds and speech sounds. The cochlear implant consists of a very small electrode array that is inserted into the cochlea adjacent to the auditory nerve fibers. This electrode array is in turn connected to a small module located under the skin behind the ear. An external processor is then added that picks up acoustic signals in the environment and transmits them across the skin where they are picked up by the internal module and converted to electrical signals on the electrode that then activate any remaining auditory nerve fibers. Over 60,000 patients already have been implanted, and this number is growing at a very rapid rate. For the adult who experiences a permanent, severe or profound sudden bilateral sensorineural hearing loss, the cochlear implant often restores sufficient auditory function to allow normal or near-normal communication ability, including face-to-face conversation and telephone use. The device also helps substantially in the educational process for teaching oral language and speech to prelingual hearing-impaired children. Universal neonatal hearing screening programs are now a part of routine neonatal health care in many countries throughout the world. Many deaf children are detected at birth by such programs, then receive a cochlear implant at 1 year of age, attend special and intensive education programs centered on language and speech development, and by age 5 or 6 are able to enter regular schools with their hearing peers and with normal language and normal or near-normal speech.

Hearing aids. A conventional hearing aid is a battery-operated device that amplifies sound. Hearing aids come in a variety of configurations from small modules that fit behind the outer ear to tiny modules that fit completely in the ear canal. Over 90% of all hearing aids are now digital, which allows many additional features. These include adaptive directional microphones to help users screen out background or extraneous sounds; automatic reduction of the extraneous whistling associated with older hearing aids; more settings to fine-tune performance in different environments, such as crowds; automatic and coordinated adjustment of settings in both ears to account for changing conditions in the listening environment; controls linked to other devices, such as the user's wristwatch, which allow the owner to change settings more easily and discreetly; and improved integration with cell phones and other personal listening devices. Current hearing aids that are adjusted properly and worn for at least a short period often provide substantial benefit to many hearing-impaired individuals. For others, additional improvement can be obtained with additional aural rehabilitation which includes special auditory training to use this amplification and instruction in speech reading (lip reading).

Because the spontaneous development of language and speech is highly dependent on an infant's ability to hear speech, the child with severe to profound congenital hearing loss requires either a hearing aid or a cochlear implant (or both) and intensive special education in order to talk. Hearing-impaired individuals with this type of training are known as oral deaf. Hearing-impaired individuals who do not develop oral language and speech but instead are taught sign language belong to the Deaf community, comprising individuals who have sign language as their primary language. Gerald R. Popelka

Bibliography. H. Cooper and L. Craddock, *Cochlear Implants: A Practical Guide*, 2d ed., 2006; J. Katz (ed.), *Handbook of Clinical Audiology*, 5th ed., 2001; H. A. Newby and G. R. Popelka, *Audiology*, 6th ed., 1992; J. L. Northern and M. P. Downs, *Hearing in Children*, 5th ed., 2001; M. Valente (ed.), *Strategies for Selecting and Verifying Hearing Aid Fittings*, 2002.

Heart (invertebrate)

Hearts of invertebrates can be categorized according to the source of the electrical rhythmicity that underlies their beat. Rhythmic electrical activity can arise in the muscle itself (myogenic hearts) or in neurons that drive the heart muscle (neurogenic hearts). Most mollusks and some insects appear to have purely myogenic hearts; these hearts beat normally when isolated from neural inputs. Conversely, the hearts of the higher crustaceans and the xiphosuran *Limulus* are usually considered to be purely neurogenic: motor neurons impose their rhythmic electrical activity on heart muscle fibers by means of direct excitatory synapses. Without neural input, the heart ceases to beat. Other invertebrates, including gnathobdellid leeches and some insects, have hearts that can produce a myogenic beat but

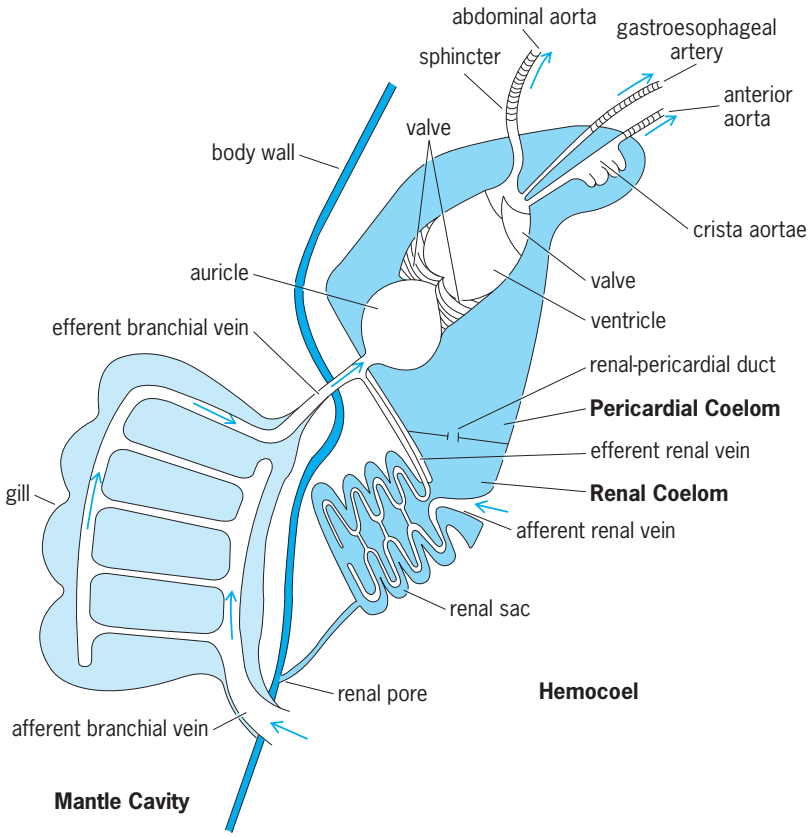

Fig. 1. ***Aplysia*** **circulatory system. Arrows indicate the normal direction of blood flow. (*After E. Mayeri et al., Neural control of circulation in Aplysia, I. Motoneurons, J. Neurophysiol., 37:458–475, 1974*)**

require rhythmic neural input to coordinate that beat and maintain the proper rate.

Myogenic hearts. In the marine snail *Aplysia*, a muscular heart consisting of an auricle and a ventricle is located in a dorsal pericardial cavity (**Fig. 1**). The rhythmic contractions of the auricle fill the ventricle with hemolymph, which is then pumped through the open circulatory system by the rhythmic contractions of the ventricle. The normal heartbeat period lasts about 3 s.

A pair of semilunar valves prevents backflow of hemolymph into the auricle during ventricular contraction. Three arteries issue from the ventricle toward the anterior, and a single semilunar valve prevents backflow from them during ventricular expansion. The arteries carry the hemolymph to the various body organs, where they end in tissue spaces. The hemolymph then collects in the hemocoel and returns to the heart by two parallel veins, one through the kidney and one through the gill.

Although the heart is innervated, its normal beat persists after denervation, and so the heart is myogenic. There is no discernible cardiac pacemaker muscle, but contraction of the ventricle usually begins at its base near the valves that separate it from the auricle.

The cardiovascular motor cells which innervate the heart are located in the abdominal ganglion. Two heart inhibitors which slow the heartbeat and a heart excitor which speeds the heartbeat have been identified. In addition, an excitatory motor neuron of unknown function innervates the heart. Three excitatory cardiovascular motor neurons which innervate certain main arteries have also been identified. The activity of these neurons causes the arteries to constrict and thus allows them to regulate blood pressure.

The cardiovascular motor cells are controlled by an elaborate network of heart interneurons within the abdominal ganglion. The interneuron network coordinates the activity of the motor cells so that appropriate cardiovascular adjustments can be made. For example, under anoxic conditions the snail displays a behavior called respiratory pumping, which is designed to maximize aeration of the gill. During this response, cardiac output decreases in an orderly way. The heart interneurons excite the heart inhibitors while inhibiting both the heart excitor and the vasoconstrictor motor neurons. This

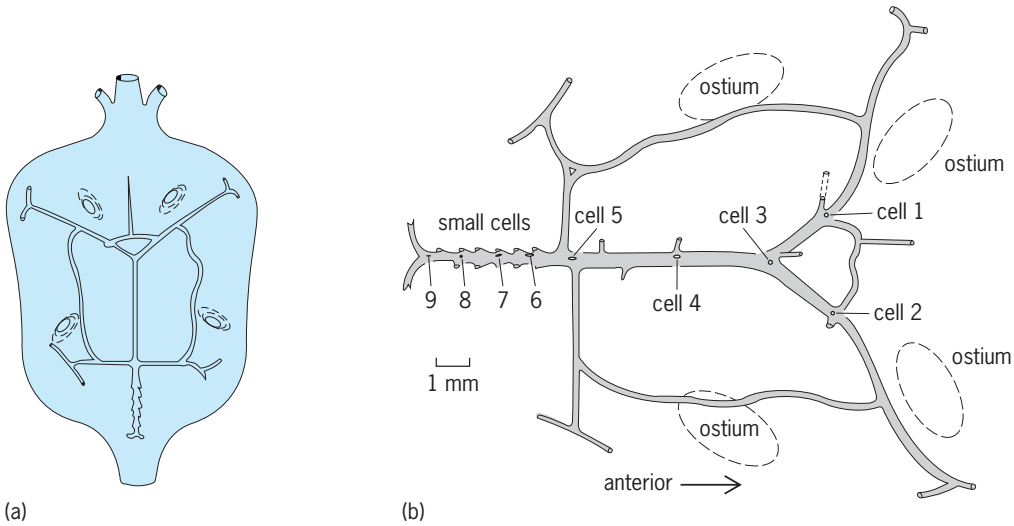

Fig. 2. Cardiac ganglion of *Homarus*. (*a*) Nerves on the inner dorsal surface of the heart. (*b*) Cell body position of all nine intrinsic neurons. (*After D. K. Hartline, Integrative neurophysiology of the lobster cardiac ganglion, Amer. Zool., 19:53–65, 1979*)

interneuronal action slows the heartbeat, lowers blood pressure, and prevents conflicting cardiovascular responses.

The heartbeat rhythm of *Aplysia* derives directly from the myogenic properties of the cardiac muscle fibers. However, a central neural network of cardiovascular motor cells and heart interneurons regulates the heartbeat rate and blood pressure. This network ensures that the cardiovascular system will respond appropriately to the environmental demands confronting the animal. The precise neural control of the heart and the existence of the functionally undefined heart motor neuron call into question the purely myogenic nature of the heart.

Neurogenic hearts. A muscular heart pumps hemolymph through the open circulatory system in lobsters (**Fig. 2***a*). This heart is located dorsally along the thoracic midline and is suspended within a pericardial cavity by ligaments. The heartbeat period lasts about 2 s. Large anterior- and posterior-going arteries, which branch extensively to supply various body organs, issue from the heart. Semilunar valves, located at the juncture of each artery with the heart, prevent backflow of blood into the heart when it relaxes. Hemolymph enters the heart from the pericardial sinus through six ostia, which have valves to ensure unidirectional flow.

The rhythmic discharge of motor neurons innervating the heart by way of excitatory chemical synapses produces the heartbeat. These motor neurons are located in the cardiac ganglion on the inner dorsal surface of the heart (Fig. 2*a* and *b*). The cardiac ganglion contains only nine neurons, which generate a simple two-phased rhythm of electrical activity. The four posterior small cells (cells 6–9) are interneurons, and the five anterior large cells (cells 1–5) are the motor neurons.

The neurons of the cardiac ganglion receive direct synaptic input from the central nervous system through one pair of inhibitory axons and two pairs of excitatory axons. Cardiac ganglion cells are excited by heart stretch, presumably by the direct activation of neuron collaterals which enter the heart muscle. A neurohormone released by the pericardial organ also directly excites cardiac ganglion cells.

All nine neurons of the ganglion produce nearly concurrent rhythmic impulse bursts of about 0.5 s in duration, separated by quiescent periods of 1.5 s. Impulse burst production begins with the interneurons, followed closely by the motor neurons which discharge at 50–100 Hz. These high-frequency impulse bursts in the motor neurons drive the heartbeat. In the absence of this rhythmic input, the heart ceases to beat; thus the heart is truly neurogenic.

The cardiac ganglion neurons generate normal rhythmic impulse bursts when totally isolated from the heart, pericardium, and central nervous system. Thus, the heartbeat rhythm is generated solely by the intrinsic neurons of the cardiac ganglion. Sensory feedback from heart stretch, input from the central nervous system, and neurohormones modulate only the rate and intensity of the heartbeat rhythm generated by the ganglion. This modulation adjusts the heartbeat rate to the lobster's internal state and to its environmental needs. For example, heartbeat rate is affected by activity, illumination, and feeding.

The interneurons of the cardiac ganglion form a positive feedback network by means of mutually excitatory chemical synapses. The interneurons also make excitatory chemical synapses on the motor neurons. All the neurons of the ganglion appear to be capable of generating an endogenous electrical rhythm which underlies impulse burst production. The motor neurons, however, have their activity cycle imposed upon them by the interneurons through strong excitatory synapses. Primary control of the overall impulse burst rhythm of the ganglion is exerted by a pacemaker interneuron, cell 9, which fires first in the ganglionic impulse burst. The excitatory synaptic interactions among the interneurons initiate and coordinate the production of impulse bursts in the other interneurons, once the pacemaker has begun to fire. Thus the interneurons produce a coordinated impulse burst, which is synaptically imposed on the motor neurons. The motor neurons then produce a synchronous impulse burst, which is synaptically imposed upon the heart muscle and causes it to contract.

The lobster cardiac ganglion can be thought of as a semiautonomous nervous system comprising two tiers of neurons (interneurons and motor neurons). This nervous system produces rhythmic impulse bursts in the motor neurons that drive the heartbeat. Sensory feedback, input from the central nervous system, and neurohormones modulate the frequency and intensity of the impulse burst rhythm produced by the ganglion cells.

In *Limulus*, which has a neurogenic heart with an organization similar to that of the lobster, the neuropeptide proctolin induces myogenic properties in the heart muscle. Although similar action of proctolin has not been reported in lobsters, whose nervous system contains the neuropeptide, this observation calls into question whether the lobster heart is purely neurogenic.

Hearts with combined properties. The medicinal leech (*Hirudo medicinalis*) is a segmented worm, and its segmentation is reflected in the anatomy and neural control of its circulatory system. The closed circulatory system comprises four longitudinal vessels which run the length of the animal (**Fig. 3**). These vessels are in fact coelomic sinuses. The sinuses communicate with one another at each end of the animal and in each segment by a series of transverse vessels. Smaller vessels and capillaries branch off the segmental transverse vessels and invade the body tissues. Spiral muscle cells located in the walls of the two lateral sinuses, or heart tubes, contract rhythmically to produce the rhythmic constrictions which move blood through the circulatory system. These constrictions of the lateral sinus constitute the heartbeat.

Normally the constriction cycle of each segmental heart tube section is coordinated with its neighbors

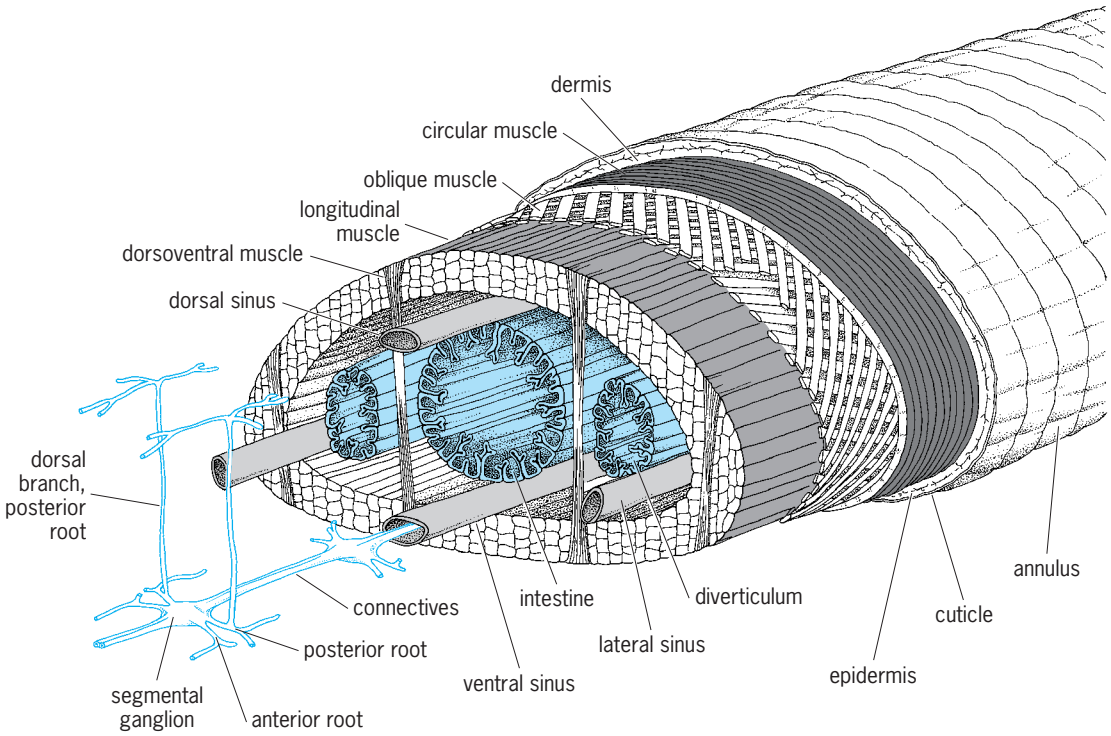

Fig. 3. Cross section of the leech *Hirudo medicinalis*, showing its major organs and muscles, including its ventral nerve cord and the four longitudinal coelomic sinuses which compose the circulatory system. The nerve cord is enclosed in the ventral sinus. The lateral sinuses constrict rhythmically and are called heart tubes. (*After J. G. Nicholls and D. van Essen, The nervous system of the leech, Sci. Amer., 230(1):38–48, 1974*)

so that on one side they constrict in a rear-to-front progression (peristalsis), while on the other they constrict nearly in synchrony. Reciprocal, right-left transitions between peristaltic and synchronous coordination states occur spontaneously every 10–20 heartbeat cycles. The heartbeat period varies from 10–30 s depending upon temperature and other undefined influences. The muscle of each segmental heart tube section, from the third through eighteenth segment, is innervated from each corresponding segmental ganglion of the ventral nerve cord by a rhythmically active segmental HE motor neuron. Each heart is also innervated intersegmentally by two cells, the HA modulatory neurons, that reside in the fifth and sixth segmental ganglia of the nerve cord.

Denervation of the hearts does not cause cessation of their rhythmic beat. However, the segmental heart tube sections lose all trace of intersegmental coordination, and their beat cycle period is prolonged. Thus, rhythmic neural activity is necessary to establish the complex coordination of the segmental heart tube sections and to maintain a normal heartbeat period. In contrast, the rhythmic activity of the HE cells in completely isolated nerve cords faithfully mimics the heartbeat pattern with all its subtle nuances. This observation shows that the complex heartbeat pattern, which normally paces the heart, is generated within the central nervous system.

The myogenic activity of the heart muscle is entrained by the rhythmic activity of the HE motor neurons through excitatory cholinergic synapses and is regulated by neuropeptides contained in these neurons and the HA neurons. The HA neurons contain the peptide FMRFamide, while the HE neurons contain that peptide or a closely related one. The HA modulatory neurons appear to regulate mainly beat strength through peptide release, while HE motor neurons appear to ensure that the heart will be rhythmically active (myogenic) through peptide release.

A set of segmental interneurons, HN cells, controls the activity of the HE cells by means of direct inhibitory synapses. Bilateral pairs of HN cells have been identified in the first seven segmental ganglia. Not all HN cells connect with HE cells, but those that do connect with all ipsilateral HE cells in more posterior ganglia.

HN cells produce rhythmic impulse bursts through combination of endogenous membrane properties and reciprocal inhibitory synapses. Their inhibitory synaptic interactions also coordinate the activity of the entire ensemble into a precise pattern. The HE cells are tonically active in the absence of inhibitory input from the HN cells. Normally they are driven into rhythmic impulse bursts by periodic inhibition from the HN cells. The intersegmental nature of HN cell to HE cell synaptic contacts ensures the proper intersegmental coordination of the HE cells. The rhythmic activity of the HE cells is then imposed upon the hearts themselves by the excitatory cholinergic synapses that they make on heart muscle. The establishment of the two different heartbeat coordination states and the reciprocal changes between

them are mediated by the HN cell pair in the fifth segmental ganglion.

Although the hearts of the medicinal leech are capable of disjointed myogenic contractions and their myogenic properties are regulated by modulatory neural inputs, an elaborate central neural network of heart interneurons and motor neurons also exists. This central network imposes its activity upon the hearts by means of direct excitatory synapses and causes them to beat with the proper period and intersegmental coordination. The central network also mediates the reciprocal changes in coordination state between the hearts. *See* NERVOUS SYSTEM (INVERTEBRATE).

Ronald L. Calabrese; Christine S. Cozzens

Bibliography. R. B. Hill and F. Lang (eds.), Symposium on the Comparative Physiology of Invertebrate Hearts, *Amer. Zool.*, 19:3–175, 1979; A. I. Selverston (ed.), *Model Neural Networks and Behavior*, 1985; G. S. Stent, W. J. Thompson, and R. L. Calabrese, Neural control of heartbeat in the leech and some other invertebrates, *Physiol. Rev.*, 59:101–136, 1979.

Heart (vertebrate)

The muscular pumping organ of the cardiovascular system.

Anatomy. The heart typically lies ventrally, near the anterior end of the trunk; it is ventral and medial to the gills in fish and at the base of the neck or in the chest region of tetrapods. In humans it is located behind the breastbone and ribs between the third and fifth costal cartilages. Its anterior portion or base is directed to the right and dorsally and is the area where the great vessels enter and leave the heart. The lower muscular portion ends in a blunt apex which lies behind the fifth costal cartilage on the left.

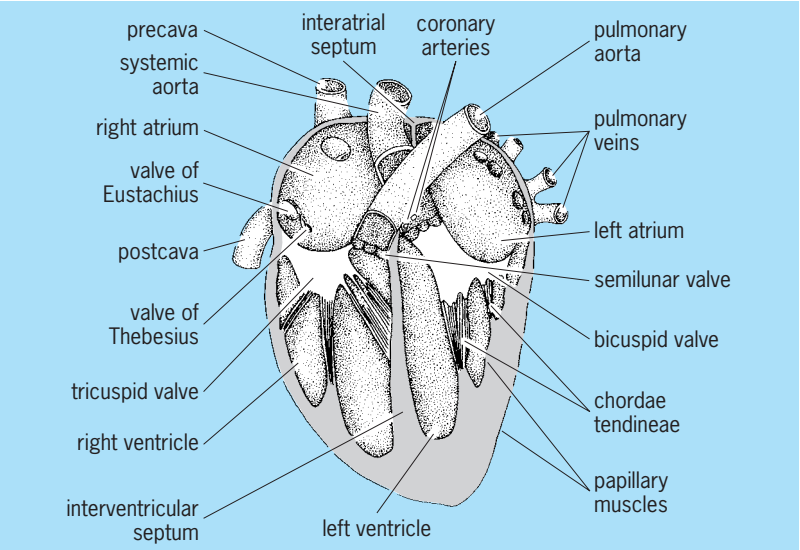

Fig. 2. Internal structure of four-chambered mammalian heart, ventral view. (*After C. K. Weichert, Anatomy of the Chordates, 2d ed., McGraw-Hill, 1958*)

The muscular wall of the heart, the myocardium, is lined by an inner endocardium and is covered externally by the membranous visceral pericardium. There are coronary arteries and veins to and from the heart, which has a specialized neuromuscular conducting system and autonomic nerve supply.

Water breathers. In fishes the heart is basically a simple tube (**Fig. 1**) which becomes subdivided into four successive chambers, the sinus venosus, atrium, ventricle, and conus arteriosus. Blood from the body enters the sinus and leaves the conus to go to the gills to be oxygenated. The ventricle supplies the main pumping force.

In fishes and all other animals with backbones the heart becomes bent to fit into the available space so that the atrium, originally posterior to the ventricle, comes to lie dorsal and finally anterior to it; in Fig. 1 the heart is shown as if it were still a simple straight tube.

Air breathers. When lungs are introduced into the system in lungfish and tetrapods, the mixing of oxygenated and nonoxygenated blood becomes a problem. In brief, the sinus venosus and conus arteriosus disappear, becoming incorporated into the other chambers or the bases of the great vessels. At the same time the atrium and later the ventricle become divided into right and left chambers by a median septum. Various stages in this development are seen in different amphibians and reptiles.

In birds and mammals including humans (**Fig. 2**) the medial fibromuscular septum divides the heart into two lateral halves, each consisting of a thin-walled receiving chamber or atrium (auricle, often used as a synonym of atrium, properly refers to a small, earlike projection of the atrium) and a thicker, muscular pumping chamber or ventricle. Blood enters the right atrium from the superior and inferior

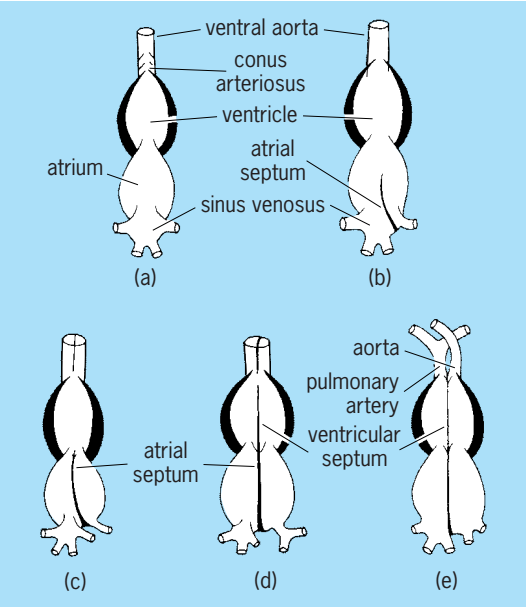

Fig. 1. Vertebrate hearts in ventral view, drawn in the primitive "stretched out" position to show the different chambers. (*a*) Dogfish. (*b*) Crossopterygian. (*c*) Amphibian. (*d*) Reptile. (*e*) Mammal.

venae cavae which drain most of the body. It passes through the tricuspid valve to the right ventricle and is pumped to the lungs during systole, or contraction of the heart. Blood returns from the lungs by way of the pulmonary veins to the left atrium, passes into the left ventricle through the mitral valve, and during contraction is pumped out into the aorta. *See* CARDIOVASCULAR SYSTEM. Thomas S. Parsons

Heart disorders

Pathologies of the heart and its blood vessels. Almost all cardiovascular disorders eventually progress to serious debilitating stages characterized by heart failure (reduced pumping function), dysrhythmias (abnormal electrical rhythms), and sudden death. Coronary atherosclerosis, a disease causing obstruction of the arteries that supply nutrient blood to heart muscle, is the leading cause of cardiovascular mortality. Hypertension (high blood pressure) is another important cause. Valvular heart disease, cardiomyopathies (disease of heart muscle), and congenital heart disease are less common.

Coronary disease. The coronary arteries carry oxygen and nutrients to the heart muscle. Obstructive disease of these arteries is almost always caused by atherosclerosis, a pathologic process that begins in late adolescence and early adulthood. In its initial stages, the disease produces a fatty deposition and thickening of the inner surface of the artery. This thickening gradually progresses to the atherosclerotic narrowing, eventually restricting the flow of blood through the artery. The major clinical manifestations of this disease are not usually seen until the sixth decade of life; men are affected earlier than women. *See* ARTERIOSCLEROSIS.

Because the size or cross-sectional area of the arterial lumen is progressively reduced by growth of the fatty plaque, the capacity to deliver blood to the heart muscle is increasingly limited. During exercise or other hemodynamic stress, the increased oxygen demand of the heart muscle cannot be met by an increase in coronary blood flow. Therefore, heart muscle function declines, and the individual experiences characteristic discomfort in the chest called angina pectoris. Angina pectoris is relieved by rest or medications that improve the balance between myocardial oxygen demands and coronary blood (and oxygen) supply. Electrocardiographic changes, which usually accompany these events, provide a basis for the exercise-stress test that can be used to diagnose coronary artery disease.

Myocardial infarction. Coronary artery disease can be complicated by the abrupt formation of a blood clot (thrombus) at the site of an atherosclerotic plaque. This formation usually occurs as a result of cracks or fissures on or adjacent to the plaque. Such rupture or disruption of plaque results in an irregularity along the inner surface of the artery that promotes formation of a thrombus with a reduction in blood flow through the artery. When the artery is incompletely obstructed, unstable angina or angina at rest may develop. This is a dynamic condition that is complicated by spasms of the involved artery. When the artery becomes completely occluded, a myocardial infarction or heart attack results, causing irreversible injury to the heart muscle (myocardium) supplied by the occluded artery. Complications and deaths are directly related to the size of the infarction. *See* INFARCTION; THROMBOSIS.

The diagnosis of a myocardial infarction is based on a clinical history of prolonged chest pain, serial electrocardiographic changes, and abnormal blood levels of enzymes that are released from damaged heart muscle. Treatment consists of prompt administration of a clot-dissolving drug so as to limit the extent of myocardial injury. The agents are administered with heparin and aspirin, which provide anticoagulant functions. On occasion, mechanical procedures, such as coronary angioplasty or bypass surgery, are used to reestablish flow in the occluded vessel.

During the hours and days after a myocardial infarction, the heart can lose its pumping ability, and consequently heart failure may develop. Other serious complications include destruction of a muscle that supports the mitral valve within the heart or rupture of the weakened wall of the heart. The injured myocardium also becomes susceptible to disorders of electrical rhythm, the most serious disorder being ventricular fibrillation. If this condition is not treated promptly, the dysrhythmia can be fatal. Most other disorders of cardiac rhythm, including atrial fibrillation, are much less dangerous.

Risk factors. Several characteristics or risk factors are associated with a high likelihood that coronary disease will develop. They include cigarette smoking, elevated levels of blood cholesterol, hypertension, and a family history of the disease. Other risk factors include diabetes, a lack of exercise, and obesity. By modifying or eliminating these factors, the risk of developing coronary disease can be reduced.

Cholesterol. Total blood cholesterol exceeding 240 mg per deciliter is associated with an increased risk of developing coronary disease. Values below 200 mg are desirable. This value represents the sum of all cholesterol carried by a number of distinct lipoproteins. The lipoproteins are combinations of cholesterol, other fats, and proteins. High-density lipoproteins (HDL) appear to provide some protection against coronary disease, while the low-density lipoproteins (LDL) are strongly associated with the development of coronary disease and its complications. However, this low-density fraction of the total cholesterol can be reduced by limiting the dietary intake of saturated fats and cholesterol. Drugs that are effective in reducing low-density lipoproteins include bile sequestrants, nicotine acid, and the inhibitors of cholesterol synthesis. When an elevated blood level of cholesterol is treated, the risk of the development of coronary disease is reduced. *See* CHOLESTEROL; LIPID METABOLISM.

Hypertension. Hypertension or high blood pressure is defined as a pressure exceeding 160/100 mmHg;

values below 140/90 mmHg are normal. A variety of vascular and neuroendocrine abnormalities contribute to the complications of hypertension, including stroke and renal failure. After a prolonged asymptomatic period, heart failure, coronary events, and stroke lead to disability or death. High blood pressure can be reduced by limiting dietary sodium and by weight loss. Several drugs, including diuretics, beta blockers, calcium channel blockers, and angiotensin converting enzyme inhibitors, are effective antihypertension agents. Lowering the blood pressure of hypertensive individuals reduces the incidence of heart failure and stroke and the probability of premature death. *See* HYPERTENSION.

Diagnosis. The diagnosis of coronary disease is often based on stress tests that use electrocardiographic monitoring during exercise. In addition, some diagnostic laboratories use an intravenous injection of radioactive tracers to study the distribution of myocardial blood flow during exercise. Other laboratories use ultrasonography (an echocardiogram) to detect changes in myocardial function during the hemodynamic stress of exercise. Unfortunately, the accuracy of such tests is imperfect, and further evaluation is necessary in some individuals, especially those with atypical symptoms. The most reliable procedure for detecting coronary disease and evaluating its severity is coronary arteriography. This procedure involves injection of a small amount of radiopaque dye into the coronary arteries (the dye is delivered through a small catheter that is inserted into an artery in the groin). A radiographic motion picture is made as the dye circulates through the artery, and atherosclerotic narrowings are identified and their severity is evaluated. *See* ELECTRODIAGNOSIS.

Treatment. The treatment of coronary disease can be considered in three stages. First, the well-known risk factors should be eliminated or modified: smoking must be discontinued, elevated blood cholesterol levels should be reduced, and high blood pressure should be treated. Second, attempts should be made to improve the balance between myocardial oxygen demand and supply with medications; nitroglycerin, beta blockers, and calcium channel blockers are effective therapeutic agents. Third, invasive procedures, such as angioplasty or surgery, may be necessary in individuals whose disease is too serious for medical therapy.

Coronary angioplasty is a procedure that reestablishes flow through a diseased segment of a coronary artery. In a manner similar to that used in coronary angiography, a balloon-tipped catheter is inserted at the groin and is directed to the diseased coronary artery. Under fluoroscopic visualization, the balloon is inflated, the atherosclerotic obstruction is stretched, and the lumen of the vessel is opened (**Fig. 1**). In most cases this procedure is effective in relieving symptoms, but on occasion the procedure damages the artery and urgent bypass surgery becomes necessary. Symptoms recur in as many as one-third of the individuals, and many of them require repeat angioplasty.

Coronary artery bypass surgery involves the placement of a vein (or artery) graft to bypass the diseased segment of a coronary artery (**Fig. 2**). This is a major thoracic surgical procedure that involves the use of a heart-lung machine. The risk of death or serious complications with this procedure is about 2–3%. Most individuals experience complete relief of symptoms; moreover, the procedure can prolong life in

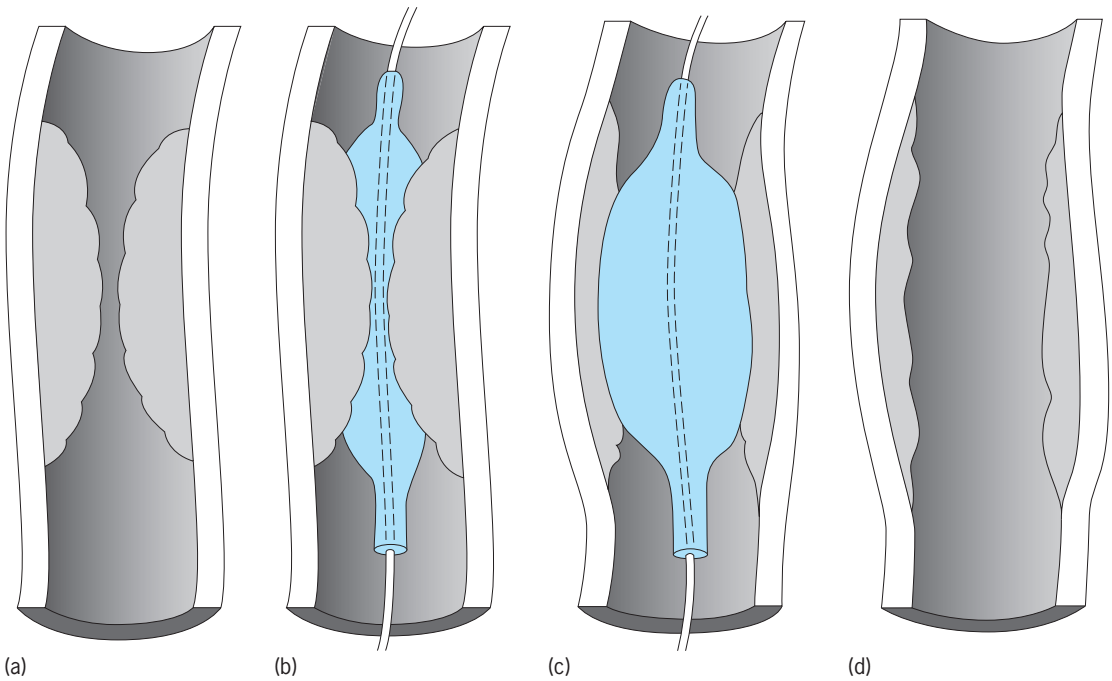

Fig. 1. Percutaneous transluminal coronary angioplasty. (*a*) Cholesterol deposits in the wall of the artery cause a partial obstruction. (*b*) A balloon-tipped catheter is positioned in the narrowed section of the diseased artery. (*c*) Inflation of the balloon compresses the atherosclerotic obstruction. (*d*) After the lumen is restored, the balloon is deflated and removed.

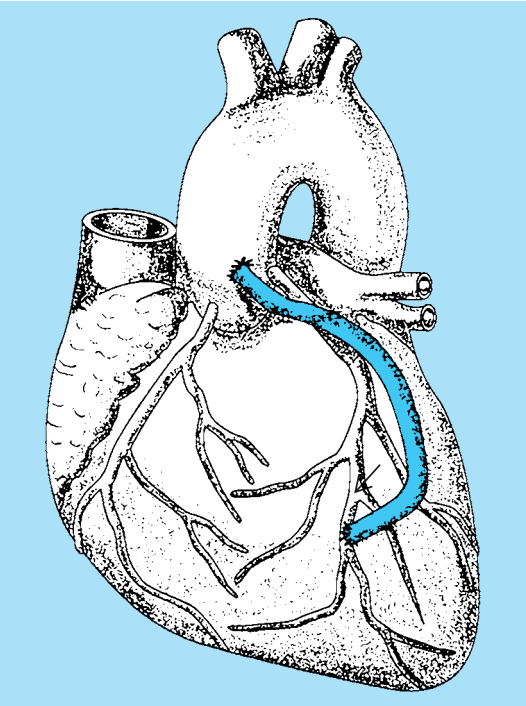

Fig. 2. Coronary artery bypass graft (shaded vessel). This graft provides blood flow to a region that is underperfused by a diseased or obstructed artery (arrow).

selected cases with severe disease. Unfortunately, the atherosclerotic process is not arrested, and therefore risk-factor modification remains important after operation. If symptoms recur, medical treatment and invasive procedures may be repeated.

Hypertensive heart disease. High blood pressure results in an excessive cardiac work load. The heart responds to the high pressure through a growth process called hypertrophy. As a result, the mass of the heart muscle is increased, and the work requirement of each unit of heart muscle returns toward normal. Unfortunately, this compensatory increase in cardiac mass eventually results in heart failure and other morbid cardiovascular events. Antihypertensive drugs are effective in the treatment of hypertension and its consequences.

Valvular heart disease. Heart valves ensure an efficient forward blood flow through the heart. Disease of these valves can be caused by rheumatic fever or degenerative noninflammatory processes. Mitral valve prolapse is a common cause of mitral regurgitation. In this condition, a portion of the blood pumped by the left ventricle leaks backward into the left atrium and causes increased pressure in the atrium and pulmonary veins and shortness of breath. Rheumatic mitral stenosis produces similar symptoms, but the cause is a narrowed valve that obstructs blood flow. A balloon valvuloplasty procedure is used to open a stenotic mitral valve, but surgical repair or replacement with an artificial valve is necessary in mitral regurgitation.

When the aortic valve is incompetent, a portion of the ejected blood leaks backward from the aorta to the left ventricle and causes enlargement of the chamber. By contrast, in aortic stenosis a narrowed valve obstructs blood flow through the valve. Thus, the ventricle is exposed to pressure overload, and it hypertrophies in a fashion similar to that in hypertensive heart disease. Surgical replacement of the valve not only relieves symptoms but can be life-saving. Unfortunately, artificial valves can malfunction or form blood clots, and on occasion additional surgery is necessary.

Congenital heart disease. About 1% of all newborns have a structural abnormality of the heart or adjacent blood vessels. Such congenital heart disease can be caused by maternal rubella, or it may be inherited, as in Down syndrome. However, the cause of most congenital heart disease is unknown.

Defects in the walls between the atria and the ventricles are the most common congenital malformations. In an atrial septal defect, a hole is present in the wall between the right and left atria. In a ventricular septal defect, a hole connects the right and left ventricles. Both conditions cause a shunting of blood from the left to the right heart chambers; thus, oxygenated blood is recirculated through the right heart and lungs. A blood shunt between the pulmonary artery and the aorta also causes a recirculation of oxygenated blood through the lungs; this defect occurs when there is an abnormal persistence of the fetal connection between the two vessels. These and a variety of other more complex abnormalities that occur alone or in combination can be cured or reduced by modern surgical techniques. *See* CONGENITAL ANOMALIES; DOWN SYNDROME.

Cardiomyopathies. Heart muscle disease that is caused by excessive intake of alcohol, some drugs, or infections may cause depression of myocardial function and cardiac enlargement. The resulting dilated cardiomyopathy causes shortness of breath and fatigue. Less commonly, inappropriate growth of the myocardium has occurred in the absence of enlargement of the chamber. This hypertrophic cardiomyopathy causes anginalike chest pain, shortness of breath, and fainting spells. If medical treatment is unsuccessful, surgical removal of a small portion of the myocardium can benefit selected individuals.

Heart failure. Congestive heart failure is a clinical syndrome that consists of shortness of breath, fatigue, and retention of fluid. It is caused by failure of the heart as a pump; thus, heart failure can be caused by almost any form of heart disease. Medical treatment includes digitalis to strengthen the heart muscle, diuretics to reduce fluid retention, and vasodilators to reduce the work load of the heart.

Cardiac transplantation can be performed in individuals whose heart failure does not respond to medical therapy. A normal donor heart (obtained from a victim of irreversible brain injury) is used to replace the diseased heart. After the operation, medications are given to prevent rejection of the transplanted heart. Unfortunately, a shortage of donor organs remains a major problem, and the use of mechanical hearts is limited by the formation of blood clots, infection, and a variety of engineering problems. *See* TRANSPLANTATION BIOLOGY.

Sudden cardiac death. Most cases of sudden death result from inadequate pumping and low cardiac output during a rapid cardiac dysrhythmia, such as ventricular tachycardia or ventricular fibrillation. Fortunately, antiarrhythmic drugs, implantable cardiac defibrillators, or surgical procedures can control the abnormal rhythms in selected individuals. *See* HEART (VERTEBRATE).

William H. Gaasch; Ferdinand J. Venditti, Jr.

Bibliography. E. Braunwald (ed.), *Heart Disease: A Textbook of Cardiovascular Medicine*, 5th ed., 1997; R. C. Schlant and R. W. Alexander (eds.), *Hurst's The Heart*, 9th ed., 1998.

Heartwater disease

A rickettsial disease, also known as cowdriosis, which is caused by the microorganism *Cowdria ruminantium* and is transmitted by ticks of the genus *Amblyomma*. The disease occurs in wild and domestic ruminants, primarily cattle, sheep, and goats, in sub-Saharan Africa and some Caribbean islands (for example, Guadeloupe, Antigua, and Marie-Galante); it is a major obstacle to improvement of livestock production in Africa.

Heartwater disease is characterized by fluid in the pericardium of the heart, high fever, lung edema, and nervous symptoms that range from mild incoordination and exaggerated reflexes to convulsions seen in acute infections. The course of acute heartwater disease is 2–6 days, and recovery is rare. However, young animals have a high rate of natural resistance.

Macrophages and the endothelial cells that line blood vessels, predominantly those in the brain, become infected with the rickettsia. The organisms develop within a membrane-bound vacuole, forming large inclusions. Within the inclusions, *C. ruminantium* divides by binary fission and subsequently develops into denser elementary bodies that are infective for other cells. In the tick vector, transmission of the rickettsia occurs while nymphal and adult ticks feed. The developmental cycle in ticks is similar to the one in endothelial cells, and involves replication of *C. ruminantium* within inclusions in gut and salivary gland cells.

Diagnosis of heartwater is by clinical symptoms and by demonstration of inclusions of *C. ruminantium* in brain tissue. Treatment is difficult because the course of the disease is rapid and must be given on the basis of clinical signs without waiting for laboratory confirmation. The organism is susceptible to tetracycline antibiotics. However, once marked nervous symptoms have developed, recovery usually does not occur. *See* ANTIBIOTIC.

Control and prevention of heartwater is achieved by tick control or immunization. Ticks are controlled by using acaracide dips or pour-ons to provide tick-free stock. An alternative program is to allow young animals to be exposed naturally to *C. ruminantium* by ticks. However, this approach requires enzootic stability of heartwater in which sufficient numbers of naturally infected ticks and animals are present in the field for continual exposure. A heartwater vaccine (available only in the Republic of South Africa) involves controlled exposure and infection of calves with frozen sheep blood infected with *C. ruminantium*. If a clinical reaction or temperature occurs, vaccinated cattle are treated with tetracyclines.

Katherine M. Kocan

Bibliography. J. L. Howard, *Current Veterinary Therapy: Food Animal Practice*, 4th ed., 1999; Z. Woldehiwet and M. Ristic (eds.), *Rickettsial and Chlamydial Diseases of Domestic Animals*, 1993.

Heartworms

Heartworm (*Dirofilaria immitus*) is a nematode parasite that resides within the host's large pulmonary arteries and right heart chambers. It primarily infests dogs but may also infest foxes, wolves, coyotes, ferrets, sea lions, horses, and cats.

A dog can be infested with one to several hundred adult heartworms, which can grow to 12 in. (30 cm). During their 3–5-year life-span, heartworms can cause serious and often life-threatening damage to the heart and lungs.

Initially a problem in southeastern coastal areas, heartworm infestation has spread rapidly throughout the United States in the past two decades. Endemic areas require a reservoir of infected animals (usually dogs) and the presence of mosquitoes, the intermediate host, which transmit the larval stages to a new host.

Life cycle. Heartworms are transmitted from dog to dog by several species of mosquitoes. Adult female heartworms release their microscopic offspring called microfilariae into the bloodstream. A mosquito becomes infested with these circulating microfilariae while taking a blood meal from the dog. The microfilariae develop into mature larvae within the mosquito during the next 10–14 days. As the mosquito feeds again, the mature larvae are injected into the new host. Once in the dog, it takes approximately 6 months for these larvae to complete the cycle by migrating to the large arteries of the lung and right chambers of the heart.

Pathology. Adult heartworms stimulate a progressive proliferation of the artery lining (endarteritis) that gradually restricts the blood flow to the lungs. The resulting increase in the pulmonary artery blood pressure (pulmonary hypertension) causes the right ventricle to pump harder, eventually leading to right heart failure. In advanced cases the lung fibrosis and heart changes may be permanent.

When adult worms die either naturally or with treatment, their fragments become lodged distally in the smaller pulmonary arteries causing an exaggerated proliferation of the vessel lining, the formation of blood clots, and an intense local inflammatory reaction. Blood flow is severely restricted or totally blocked, resulting in severe coughing, coughing up blood, and difficulty in breathing (dyspnea).

Symptoms of infestation. The clinical symptoms of heartworm disease (dirofilariasis) are proportional

to the number of infecting heartworms, the duration of the infection, and the host's response. Dogs with only a few adult worms or early in the course of the disease usually show no outward symptoms. Initial signs of disease include coughing, exercise intolerance, and weight loss. As the disease progresses, these symptoms become more pronounced. With advanced disease, dogs begin to exhibit progressive signs of pulmonary disease and associated heart failure, including fainting spells, collapse, difficulty in breathing, coughing up blood, and fluid accumulation around the lungs (hydrothorax) or within the abdominal cavity (ascites).

Rarely, a rapidly fatal condition called vena caval syndrome may be observed in young dogs with massive heartworm infestation. A large number of heartworms block the right heart valves and occlude the large vein returning to the heart (vena cava), severely compromising cardiac, liver, and kidney function. A dog with vena caval syndrome will suddenly stop eating, become listless, very weak, and pale, and may become jaundiced (icteric). An enlarged jugular vein is common. The urine is dark brown (hemoglobinuria). Without prompt surgical removal of the worm burden, most of these dogs will die within 24–72 h.

Diagnosis. Heartworm infection can usually be determined by examining a blood sample for the presence of circulating microfilariae. Unfortunately, circulating microfilariae are found in only 30–60% of infected dogs and in less than 10% of infected cats. Those infestations in which the adults produce no circulating microfilariae are termed occult infections. An occult heartworm infestation can be accurately diagnosed by identifying specific circulating immunologic substances (uterine antigens) released into the blood by adult females.

Treatment. Managed properly, all but the most advanced cases of heartworm disease can usually be treated successfully. There are six steps in heartworm treatment.

1. The severity of the infestation is assessed, and the degree of pathologic disease present is staged. A thorough medical history, a thorough physical examination, laboratory evaluation of the liver and kidneys, urine analysis, chest radiographs, and an electrocardiogram complete the evaluation. Dogs with advanced chronic heartworm disease, especially those with heart failure, are at greatest risk of complications with the adulticide therapy. Stabilizing the lung and cardiac function is strongly advised prior to initiating therapy to eliminate the adult worms. In these cases, the prospect of complete recovery is poor despite successfully killing all the adults.

2. The adult worms are killed with an adulticide drug. The treatment involves a series of injections with an arsenical compound. Serious side effects are possible; therefore the dog should be hospitalized and closely monitored for potential adverse reactions.

3. Exercise is restricted for 4–6 weeks following administration of the adulticide to minimize the potential for all the dead worms becoming lodged in the smaller pulmonary arteries at one time.

4. Medication is administered to eliminate the microfilariae, if present, once the threat of embolic pneumonia has passed. The microfilariae can cause damage to the dog's kidney (for example, glomerular nephritis). They are sources of infestation for other dogs in the area as well.

5. Once the microfilariae have been eliminated, daily or monthly preventive medication is initiated.

6. The antigen-based (occult) adult heartworm test is repeated to determine the success of the adulticide therapy.

Prevention. Daily or monthly heartworm preventive medication is strongly recommended, especially during the mosquito season, in infested areas. The objective is killing the infectious larval stages before they develop into adults. Only dogs that test negative for heartworms should be placed on preventive medication because of the risk of serious reactions.

William D. Fortney

Bibliography. P. A. Aiello and S. E. Aiello (eds.), *The Merck Veterinary Manual*, Merck & Company, 1998; American Heartworm Society, *Heartworm Disease in Dogs*, 1995; S. J. Ettinger and E. C. Feldman (eds.), *Textbook of Internal Medicine: Disease of the Dog and Cat*, W. B. Saunders, 1995.

Heat

For the purposes of thermodynamics, it is convenient to define all energy while in transit, but unassociated with matter, as either heat or work. Heat is that form of energy in transit due to a temperature difference between the source from which the energy is coming and the sink toward which the energy is going. The energy is not called heat before it starts to flow or after it has ceased to flow. A hot object does contain energy, but calling this energy heat as it resides in the hot object can lead to widespread confusion. *See* ENERGY; INTERNAL ENERGY.

Heat flow is a result of a potential difference between the source and sink which is called temperature. Work is energy in transit as a result of a difference in any other potential such as height. Work may be thought of as that which can be completely used for lifting weights. Heat differs from work, the other type of energy in transit, in that its conversion to work is limited by the fundamental second law of thermodynamics, or Carnot efficiency. This natural law is that the fraction of the heat Q convertible to work is determined by the relation $dW = Q(dT/T)$ for processes where the source and sink are differentially different in temperature, or by the relation $dW = dQ(T_1 - T_2)/T_1$ where the source (at T_1) and the sink (at T_2) differ by a finite temperature interval. *See* WORK.

For the above relations to be valid, temperature must be expressed on a thermodynamic temperature scale. Conversely, any temperature scale for which the above relations are valid, irrespective of the substance or material under

investigation, is a thermodynamic temperature scale. The perfect gas law defines a scale in which the temperature is proportional to the thermodynamic temperature. In order to make the two scales be identical, the triple point of water (temperature and pressure at which ice, water, and vapor are in equilibrium) is defined to be at 273.16 kelvins on both the ideal-gas and the thermodynamic scales. *See* TEMPERATURE; THERMODYNAMIC PRINCIPLES.

Harold C. Weber; William A. Steele

Bibliography. C. O. Bennett and J. E. Myers, *Momentum, Heat, and Mass Transfer*, 3d ed., 1982; K. S. Pitzer, *Thermodynamics*, 3d ed., 1995.

Heat balance

An application of the first law of thermodynamics to a process in which any work terms are negligible.

Closed systems. For a closed system, one that always consists of the same material, the first law is $Q + W = \Delta E$, where Q is the heat supplied to the system, W is the work done on the system, and ΔE is the increase in energy of the material forming the system. It is convenient to treat ΔE as the sum of changes in mechanical energy, such as kinetic energy and potential energy in a gravitational field, and of internal energy ΔU that depends on changes in the thermodynamic state of the material. (Scientific disciplines do not all use the same sign convention for work, so care is needed when consulting text books.)

The net heat input Q may have positive and negative components. For example, a satellite in space has no work interaction with its surroundings. It receives heat by radiation from the Sun on some surfaces (positive Q) and loses heat by radiation to space from others (negative Q). If the net heat input is positive, the internal energy of the satellite increases and its temperature rises; if the heat inputs and outputs are in balance, the internal energy of the satellite does not change and its temperature remains constant. Because the rates at which any changes occur are usually of interest, heat balances are often written in terms of heat flow rates (heat per unit time), sometimes denoted by a dot over the symbol, $\dot{Q}$, so that for a process with negligible work, kinetic energy and potential energy terms $\dot{Q} = \dot{Q}_{\mathrm{IN}} - \dot{Q}_{\mathrm{OUT}} = dU/dt$, the rate of change of internal energy with time.

The thermodynamic distinctions in terminology between heat, work, and energy are not always applied rigorously. For example, in a heat balance on an electric light bulb, the electrical power is strictly a rate of doing work on the system but might be treated as an electrical heat input, to be balanced against the rate of heat loss by radiation and convective cooling of the bulb by the surrounding air.

Open systems. Often it is more convenient to apply the first law or a heat balance to an open system, a fixed region or control volume across the boundaries of which materials may travel and inside which they may accumulate, such as a building, an aircraft engine, or a section of a chemical process plant. Then the first law is expressed by the equation below,

$$\dot{Q} + \dot{W}_S = \sum \dot{m}\left(h + c^2/2 + gz\right)_{\mathrm{OUT}} - \sum \dot{m}\left(h + c^2/2 + gz\right)_{\mathrm{IN}} + dE/dt$$

where $\dot{W}_S$ is the rate of doing shaft work on the system; $\dot{m}$ is the mass flow rate of any stream entering or leaving the control volume; h is the enthalpy per unit mass; c is the velocity; gz is the gravitational potential for each stream at the point of crossing the boundary of the control volume; and E is the energy of all material inside the control volume. When conditions inside the control volume do not change with time, although they need not be spatially uniform, $dE/dt = 0$, and the balance equation is known as the steady-flow energy equation. There is also a mass balance equation: the total incoming and outgoing mass fluxes must be equal if material cannot accumulate inside the control volume.

Enthalpy is a thermodynamic property defined by $h = u + pv$, where u is the specific internal energy (enthalpy per unit mass), p the pressure, and v the specific volume. It is used, along with shaft work $\dot{W}_S$, because the derivation of the first-law equation for a control volume from the more fundamental equation for a closed system involves work terms pv that are not available for use outside the control volume. Changes in enthalpy occur because of changes in temperature, pressure, physical state (for example, from liquid to vapor), and changes in chemical state.

For example, for a control volume around a domestic heating boiler, the entering material streams are the fuel and the air needed for combustion and the incoming flow of cold water; the outgoing streams are the combustion products and the hot water or steam. The only heat flow would be a small term for heat losses from the hot casing of the boiler to the surroundings. There might be a small electrical work input to drive a circulating pump for the water. The main terms in this heat balance would be the changes in enthalpy. The enthalpy change from fuel to combustion products is negative and is related to, but not quite equal to, the calorific value of the fuel. *See* ENTHALPY; THERMODYNAMIC PRINCIPLES; THERMODYNAMIC PROCESSES.

D. B. R. Kenning

Bibliography. Y. A. Cengel and M. A. Boles, *Thermodynamics: An Engineering Approach*, 3d ed., McGraw-Hill, 1998; M. C. Potter and C. W. Somerton, *Engineering Thermodynamics*, McGraw-Hill, 1995; K. Wark, Jr., *Advanced Thermodynamics for Engineers*, McGraw-Hill, 1994; P. B. Whalley, *Basic Engineering Thermodynamics*, Oxford University Press, 1992.

Heat balance, terrestrial atmospheric

The balance of various types of energy in the atmosphere and at the Earth's surface. Essentially all the energy that the Earth–atmosphere system receives

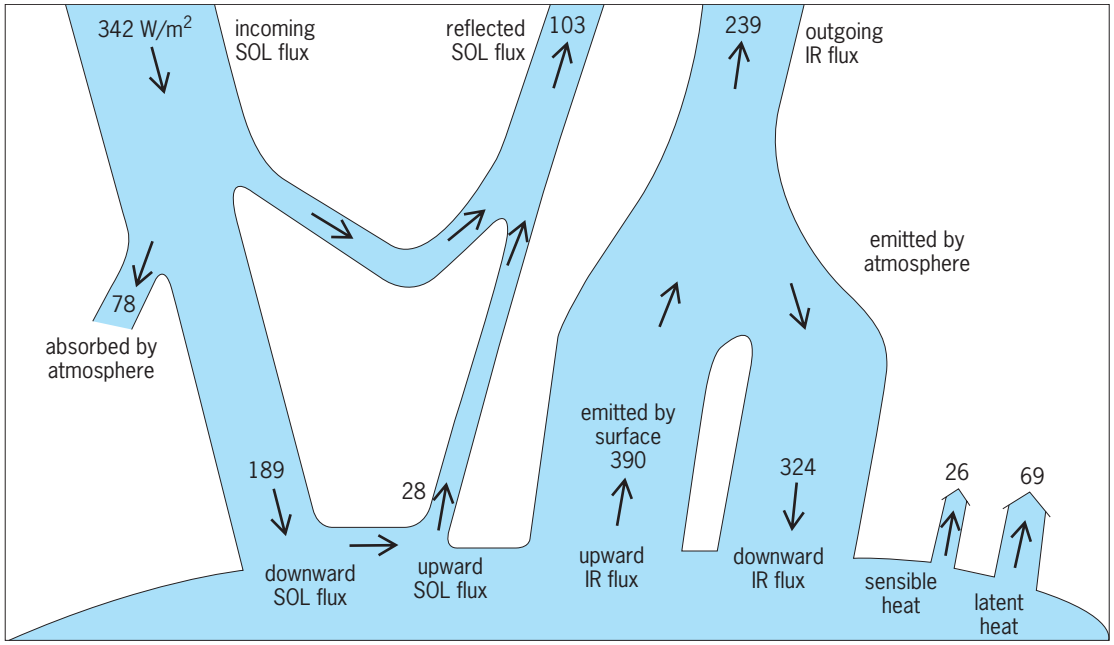

Fig. 1. Heat balance of the Earth and the atmosphere system. The solar (SOL) constant used in the calculations is 1366 W/m^2 so that the effective incoming solar flux is 342 W/m^2, while the global albedo used is 30%. The atmosphere given in the figure contains molecules, aerosols, and clouds. The atmospheric thermal infrared (IR) flux is emitted both upward and downward. The surface albedo and surface temperature used are 15% and 288 K, respectively. The width of the shaded area with an arrow is approximately proportional to the flux value. (*Data taken from K. N. Liou, 2002*)

comes from the Sun. This energy is conventionally referred to as the solar constant, and is defined as the flux of solar energy (energy per time) available on 1 square meter facing the Sun at the top of the atmosphere when the Earth is at its mean distance from the Sun. On the basis of recent satellite observations, a value of about 1336 watts per square meter (W/m^2) has been suggested. Because the area of the spherical Earth is four times that of its cross section facing the parallel solar beam, the top of the Earth's atmosphere receives an average of about 342 W/m^2. Based on analysis of the observed data from satellite radiation budget experiments in the last 40 years, about 30% of this is reflected back to space and is referred to as the global albedo. The reflecting power of the Earth–atmosphere system includes the scattering of molecules, aerosols, and clouds, as well as reflection of different types of surfaces. As a consequence of this global albedo, only about 70% of the incoming solar flux (that is, about 239 W/m^2) is available on average to warm the Earth–atmosphere system. For this system to be in thermodynamic equilibrium or balance so that an equilibrium temperature can be defined, it must radiate the same amount of energy (239 W/m^2) back to space. The emitted (or outgoing) terrestrial radiation from the Earth and the atmosphere having an equilibrium temperature of about 254 K (−2.5°F) is in the infrared portion of the electromagnetic spectrum and is called the thermal infrared radiation or longwave radiation. This is differentiated from the solar radiation or shortwave radiation from the Sun, which has an effective temperature of about 5800 K (10,000°F). *See* ALBEDO; ATMOSPHERE; HEAT BALANCE; SOLAR CONSTANT; SOLAR ENERGY; TERRESTRIAL RADIATION.

On global average and over a climatological period of say 40 years, the incoming solar flux of 342 W/m^2 must be balanced by the sum of the reflected solar flux of 103 W/m^2 and the emitted infrared flux of 239 W/m^2 at the top of the atmosphere, as displayed in **Fig. 1**. This is referred to as the heat balance of the Earth–atmosphere system.

Determination of the radiative heat-balance components in the atmosphere and at the surface requires theoretical calculation and analysis. Climatological profiles of temperature, cloud, aerosol, water vapor, ozone, molecular make up, and pertinent absorbing gases are needed in radiative transfer calculations. Also required are the globally averaged surface albedo, estimated to be about 15%, and the climatological surface temperature of 288 K (59°F). Using these parameters, the solar flux reaching the surface is found to be about 189 W/m^2, while the reflected component is 28 W/m^2. Thus, the absorbed solar flux at the surface is 161 W/m^2, which amounts to about 47% of the incoming solar flux. As a consequence, the atmosphere, including clouds, aerosols, ozone, air molecules, and other minor gases, absorbs 78 W/m^2, or about 23% of the incoming solar flux.

On the basis of theoretical calculations, the emitted downward infrared flux from the atmosphere reaching the surface is 324 W/m^2. The climatological surface temperature of 288 K emits 390 W/m^2 of infrared flux, leading to a net loss of the thermal infrared flux by the Earth's surface of 66 W/m^2. Consequently, the radiative heat balance at the Earth's surface is the sum of the gain due to the absorption

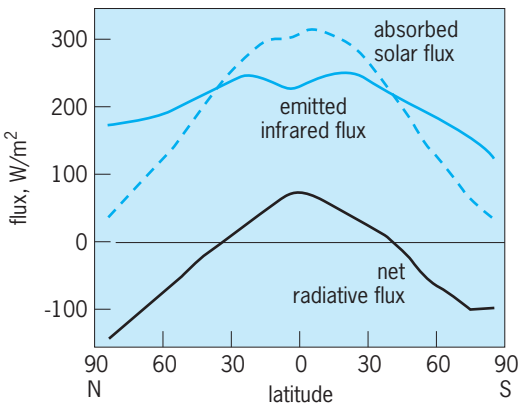

Fig. 2. Components of the annual mean absorbed solar flux, emitted infrared flux, and net radiative flux as functions of latitude at the top of the atmosphere derived from satellite broadband radiation measurements. (*Data taken from K. N. Liou, 2002*)

of solar flux and the net emission loss associated with thermal infrared flux, resulting in a positive gain of 95 W/m^2. This gain is countered by the transport of fluxes of sensible heat (related to potential temperature) and latent heat (related to water substances) out of the surface in order to maintain an overall heat balance. These fluxes are 26 W/m^2 and 69 W/m^2, respectively, based on the so-called Bowen ratio, which has a global value of 0.27. The surface heat balance is illustrated in Fig. 1.

On global average, the incoming solar flux absorbed by the Earth-atmosphere system is balanced by the emitted infrared flux, as presented above. However, substantial local variability occurs. **Figure 2** illustrates the latitudinal distribution of solar, infrared, and net radiative flux at the top of the atmosphere. The net radiative flux pattern depicts a maximum gain of radiative heat at the Equator associated with a minimum of emitted infrared flux due to the colder towering cumulus clouds frequently occurring in this region and the large amount of solar flux absorbed there. In the polar regions, we see large negative net fluxes associated with the high albedo of snow and ice as well as the sharp decrease in solar insolation during the winter season. This configuration clearly demonstrates the gain and loss of radiative energy in the tropical and subtropical areas and the polar regions, respectively, a pattern that establishes a temperature gradient from the Equator to the Poles. Consequently, thermally driven circulations are produced either directly or indirectly by horizontally heating and cooling gradients in the atmosphere. Thus, there is a substantial transport of sensible and latent heat by large-scale atmospheric circulations in addition to their upward transport from the surface to the atmosphere by convective activities, as shown in Fig. 1. *See* ATMOSPHERIC GENERAL CIRCULATION; INSOLATION.

The atmospheric latent heat component has maxima in the subtropics between about 20° and 30° latitude in both hemispheres, where evaporation rates are highest. Minimum patterns are evident between about 10°N and 10°S because of excess precipitation in the tropical regions, and between about 40° and 60° in both hemispheres because of surplus precipitation associated with storm activity. The atmospheric sensible heat component shows a large maximum in the tropics associated with higher temperatures and two small maxima located at about 40°N and 40°S related to the transport of heat by circulations. On an annual basis and over a climatological time period, about 20 W/m^2 of sensible and latent heat is lost between 40°S and 40°N, while about 50 to 70 W/m^2 is gained poleward of 60°. In addition to the transport of sensible and latent heat fluxes by atmospheric circulations, the ocean currents can transport heat. Best estimates show that about 40 W/m^2 is moved out of the tropics and about 25 W/m^2 is transported into latitudes poleward of 40° by ocean currents, but this transport is terminated at about 70°S where there are no oceans. The sum of the heat flux transport by the atmosphere and the ocean must be balanced by the net radiative flux at the top of the atmosphere on an annual basis. K. N. Liou

Bibliography. K. N. Liou, *An Introduction to Atmospheric Radiation*, 2d ed., Academic Press, 2002.

Heat capacity

The quantity of heat required to raise a unit mass of homogeneous material one unit in temperature along a specified path, provided that during the process no phase or chemical changes occur is known as the heat capacity of the material in question. The unit mass may be 1 g, 1 lb, or 1 gram-molecular weight (1 mole). Moreover, the path is so restricted that the only work effects are those necessarily done on the surroundings to cause the change to conform to the specified path. The path, except as noted later, is at either constant pressure or constant volume. This definition conforms to an average heat capacity for the chosen unit change in temperature.

Instantaneous heat capacity at a particular temperature is defined as the rate of heat addition relative to the temperature change at the temperature in question; that is, on a plot of heat addition Q as a function of temperature T, instantaneous heat capacity is given by the slope of the curve at the temperature in question. Units of heat capacity are energy units per unit mass of material per unit change in temperature.

In accordance with the first law of thermodynamics, heat capacity at constant pressure C_p is equal to the rate of change of enthalpy with temperature at constant pressure $(\partial H/\partial T)_p$. Heat capacity at constant volume C_v is the rate of change of internal energy with temperature at constant volume $(\partial U/\partial T)_v$. Moreover, for any material, the first law yields the relation in Eq. (1).

$$C_p - C_v = \left[P + \left(\frac{\partial U}{\partial V}\right)\right]_T \left(\frac{\partial U}{\partial T}\right)_P \qquad (1)$$

See ENTHALPY; INTERNAL ENERGY.

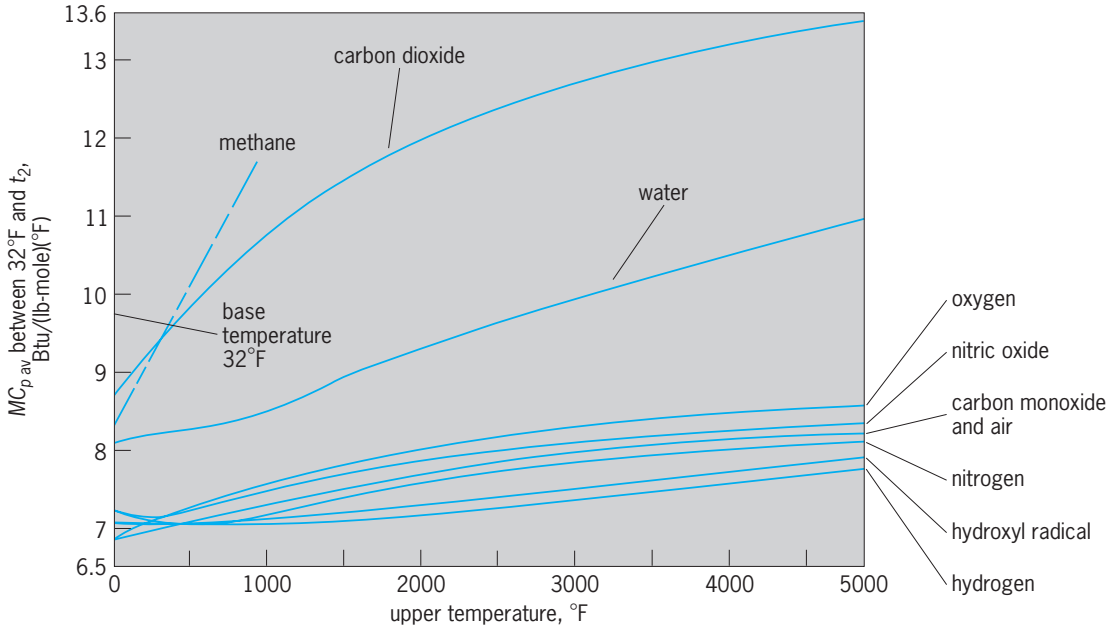

Fig. 1. Variation of average molar heat capacity with temperature for several common gases. t^2 = a given upper temperature. °C = (°F − 32)/1.8.

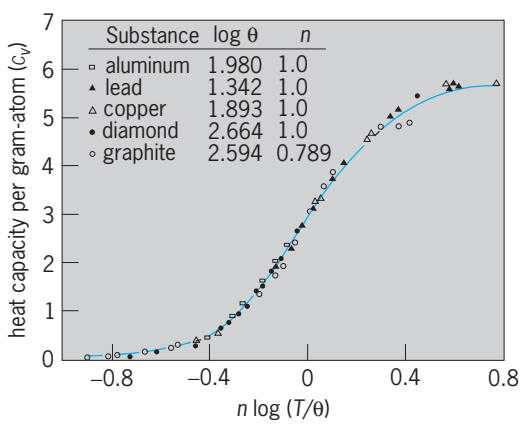

Fig. 2. Atomic heat capacity as a function of temperature (K).

Gases. For one mole of a perfect gas, the preceding relation becomes $MC_p - MC_v = R$, where M is the molecular weight of the gas under discussion and R is the perfect gas law constant. For gases the ratio C_p/C_v is usually designated by the symbol K.

For monatomic gases at moderate pressures, MC_v is about 1.67, and the heat capacity changes but little with temperature. For diatomic gases, MC_v is approximately 5 at 68°F (20°C) and moderate pressures. Change of heat capacity with temperature is usually small. The value of K is between 1.40 and 1.42. For triatomic gases at moderate pressures, MC_v varies from 6 to 7 and changes rapidly with temperature. The value of K varies but is always smaller than that for the less complex molecules at the same conditions of pressure and temperature.

For gases with more than three atoms per molecule, no generalizations are reliable. However, as molecular complexity increases, heat capacity increases, the influence of temperature on heat capacity increases, and K decreases. **Figure 1** shows average MC_p for several common gases. Up to pressures of a few atmospheres, the effect of pressure on heat capacity of gases is small and is usually neglected.

Solids. For solids, the atomic heat capacity (heat capacity when the unit mass under discussion is 1 at. wt) may be closely approximated by equation of type (2), where $n = 1$ for elements of simple crystalline

$$C_v = J\left(\frac{T}{\theta}\right)^n \tag{2}$$

form, but has a smaller value for those of more complex structures; θ is characteristic of each element; J is a function that is the same for all substances; T is absolute temperature. **Figure 2** compares measured with calculated values.

For all solid elements at room temperature, C_v is about 6.4 calories per gram atom per degree Celsius. This approximation may be used when no experimental data are available, but errors may be considerable, particularly for elements with atomic weights less than 39. Kopp's law states that for solids the molal heat capacity of a compound at room temperature and pressure approximately equals the sum of heat

Heat capacities of some elements

Element	Heat capacity
All heavy elements	6.4
Boron	2.7
Carbon	1.8
Fluorine	5.0
Hydrogen	2.3
Oxygen	4.0
Phosphorus and sulfur	5.4
Silicon	3.5

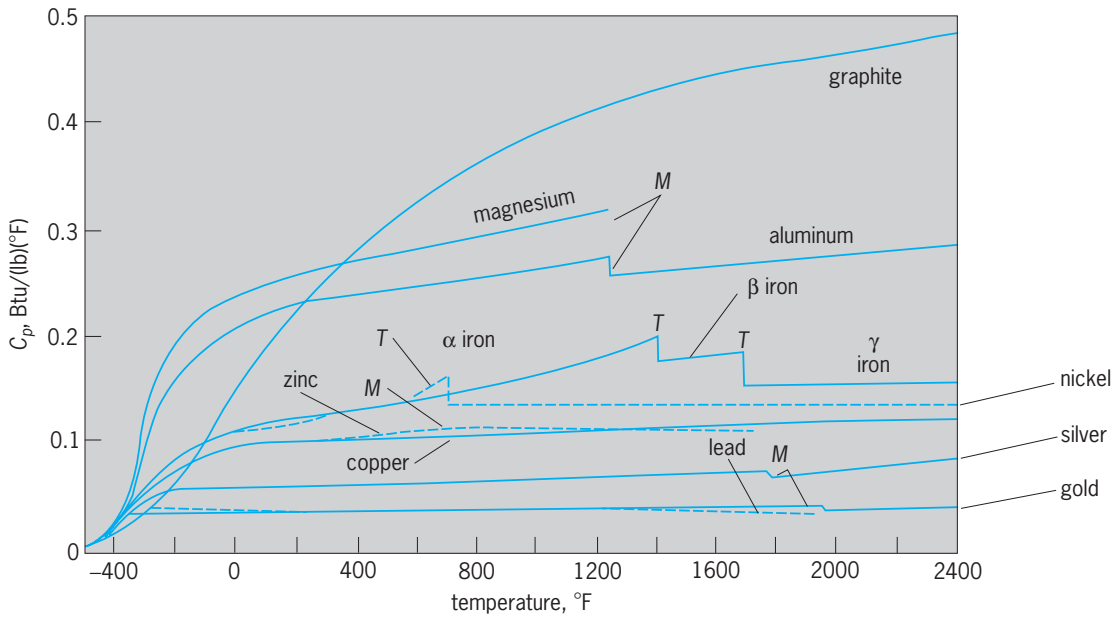

Fig. 3. Change in heat capacity of some industrially important solids with temperature. *M* = melting point; *T* = transition temperature. °C = (°F − 32)/1.8.

capacities of the elements in the compound. Errors are considerable but may be reduced by judicious choice of atomic heat capacities for the lighter elements. Recommended values for some of these are given in the **table** of constants for Kopp's law. Use of Kopp's law is justified only when no experimental data are available.

Figures 3 and **4** give instantaneous heat capacities for some industrially important solids.

Liquids. For liquids and solutions no useful generally applicable approximations are available. For aqueous solutions of inorganic salts the approximate heat capacity of the solution may be estimated by assuming the dissolved salt to have negligible heat capacity. Thus, in a 20% by weight solution of any salt in water 0.8 would be the estimated heat capacity.

Effect of pressure on heat capacities at any temperature may be calculated by the relations in Eqs. (3*a*) and (3*b*).

$$\left(\frac{\partial MC_p}{\partial P}\right)_T = -T\left(\frac{\partial^2 V}{\partial T^2}\right)_P \qquad (3a)$$

$$\left(\frac{\partial MC_v}{\partial V}\right)_T = T\left(\frac{\partial^2 P}{\partial T^2}\right)_V \qquad (3b)$$

Constant temperature. Not so familiar as C_p and C_v are the heat necessary to cause unit change

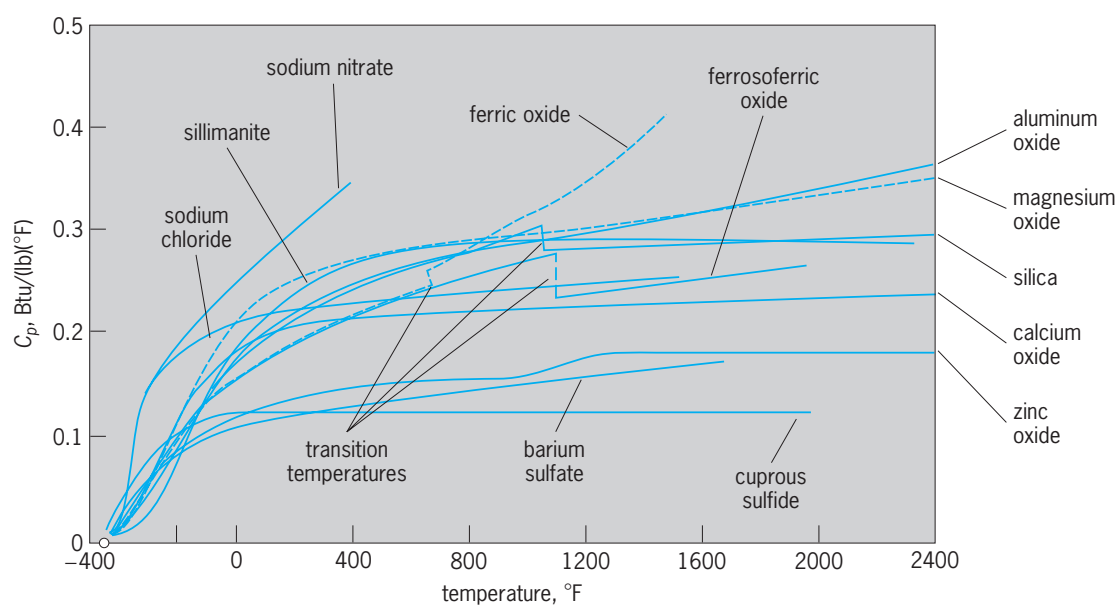

Fig. 4. Change in heat capacity of compounds with temperature. °C = (°F − 32)/1.8.

in pressure in a unit mass of material at constant temperature and the heat required to cause unit change in volume at constant temperature. These are designated $\partial Q_P/\partial P$ and $\partial Q_T/\partial V$. Similarly, $\partial Q_V/\partial P$ and $\partial Q_p/\partial V$ may be called heat capacities. *See* SPECIFIC HEAT; THERMODYNAMIC PRINCIPLES.

Harold C. Weber

Bibliography. I. M. Klotz and R. M. Rosenberg, *Classical Thermodynamics*, 5th ed., 1994; K. C. Rolle, *Thermodynamics and Heat Power*, 5th ed., 1998.

Heat exchanger

A device used to transfer heat from a fluid flowing on one side of a barrier to another fluid (or fluids) flowing on the other side of the barrier.

When used to accomplish simultaneous heat transfer and mass transfer, heat exchangers become special equipment types, often known by other names. When fired directly by a combustion process, they become furnaces, boilers, heaters, tube-still heaters, and engines. If there is a change in phase in one of the flowing fluids—condensation of steam to water, for example—the equipment may be called a chiller, evaporator, sublimator, distillation-column reboiler, still, condenser, or cooler-condenser.

Heat exchangers may be so designed that chemical reactions or energy-generation processes can be done within them. The exchanger then becomes an integral part of the reaction system and may be known as a nuclear reactor or catalytic reactor.

Heat exchangers are normally used only for the transfer and useful elimination or recovery of heat without an accompanying phase change. The fluids on either side of the barrier are usually liquids, but they may also be gases such as steam, air, or hydrocarbon vapors; or they may be liquid metals such as sodium or mercury. Fused salts are also used as heat-exchanger fluids in some applications.

With the development and commercial adoption of large, air-cooled heat exchangers, the simplest example of a heat exchanger would now be a tube within which a hot fluid flows and outside of which air is made to flow for cooling. By similar reasoning, any container of a fluid immersed in any fluid could serve as a heat exchanger if the flow paths were properly connected, or any container of a fluid exposed to air becomes a heat exchanger when a temperature differential exists.

Most often, the barrier between the fluids is a metal wall such as that of a tube or pipe. However, it can be fabricated from flat metal plate or from graphite, plastic, or other corrosion-resistant materials of construction. If the barrier wall is that of a seamless or welded tube, several tubes may be tied together into a tube bundle (see **illus.**) through which one of the fluids flows distributed within the tubes. The other fluid (or fluids) is directed in its flow in the space outside the tubes through various arrangements of passes. This fluid is contained by the heat-exchanger shell. Discharge from the tube bundle is to the head (heads) and channel of the exchanger. Separation of tube-side and shell-side fluids is accomplished by using a tube sheet (tube sheets).

Applications. Heat exchangers find wide application in the chemical process industries, including petroleum refining and petrochemical processing; in the food industry, for example, for pasteurization of milk and canning of processed foods; in the generation of steam for production of power and electricity; in nuclear reaction systems; in aircraft and space vehicles; and in the field of cryogenics for the low-temperature separation of gases. Heat exchangers are the workhorses of the entire field of heating, ventilating, air-conditioning, and refrigeration.

Classifications. The exchanger type described in general terms above and illustrated by the diagram is the well-known shell-and-tube heat exchanger. Shell-and-tube exchangers are the most numerous, but constitute only one of many types. Exchangers in use range from the simple pipe within a pipe—with a few square feet of heat-transfer surface—up to the

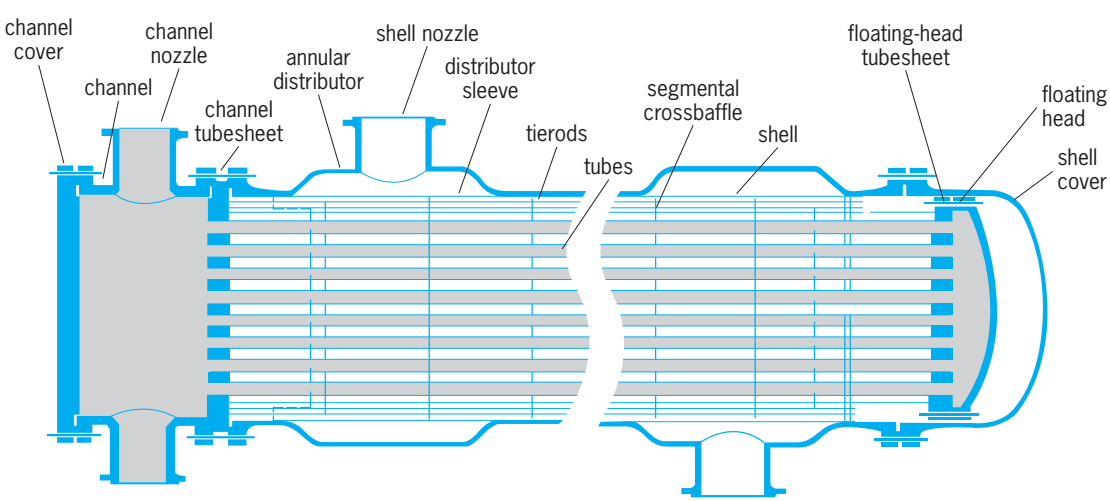

Schematic diagram of heat exchanger.

complex-surface exchangers that provide thousands of square feet of heat-transfer area.

In between these extremes is a broad field of shell-and-tube exchangers often specifically named by distinguishing design features; for example, U tube, fin tube, fixed tube sheet, floating head, lantern-ring packed floating head, socket-and-gland packed floating head, split-ring internal floating head, pull-through floating head, nonremovable bundle with floating head or U-tube construction, and bayonet type.

Also, varying pass arrangements and baffle-and-shell alignments add to the multiplicity of available designs. Either the shell-side or tube-side fluids, or both, may be designed to pass through the exchanger several times in concurrent, countercurrent, or cross flow to the other fluids.

The concentric pipe within a pipe (double pipe) serves as a simple but efficient heat exchanger. One fluid flows inside the smaller-diameter pipe, and the other flows, either concurrently or countercurrently, in the annular space between the two pipes, with the wall of the larger-diameter pipe serving as the shell of the exchanger.

To solve new processing problems and to find more economical ways of solving old ones, new types of heat exchangers are being developed. There has been much emphasis on cramming more heat-transfer surface into less and less volume. Extended-surface exchangers, such as those built with fin tubes, are finding wide application.

Water shortage has added a new dimension to heat-exchanger design and has led to use of air-cooled exchangers.

Plate-type heat exchangers, long used in the milk industry for pasteurization and skimming, have moved into the chemical and petroleum industries. Coiled tubular exchangers and coiled-plate heat exchangers are winning new assignments. Spiral exchangers offer short cylindrical shells with flat heads, carrying inlets and outlets leading to internal spiral passages. These passages may be made with spiral plates or with spiral banks of tubes. Exchangers with mechanically scraped surfaces are finding favor for use with very viscous and pastelike materials.

A somewhat unusual type of plate heat exchanger is one in which sheets of 16-gage metal, seam- and spot-welded together, are embossed to form transverse internal channels which carry the heat-transfer medium. This type is often used for immersion heating in electroplating and pickling.

Materials of construction. Every metal seems to be a possible candidate as a material of construction in fabrication of heat exchangers. Most often, carbon steels and alloy steels are used because of the strength they offer, especially when the exchanger is to be operated as a pressure vessel. Because of excellent heat conductance, brass and copper find wide use in exchanger manufacture.

Corrosion plays a key role in the selection of exchanger construction materials. Often, a high-priced material will be selected to contain a corrosive tube-side fluid, with a cheaper material being used on the less corrosive shell side.

For special corrosion problems, exchangers are built from graphite, ceramics, glass, bimetallic tubes, tantalum, aluminum, nickel, bronze, silver, and gold. *See* CORROSION.

Problems of use. Each of the fluids and the barrier walls between them offers a resistance to heat transfer. However, another major resistance that must be considered in design is the formation of dirt and scale deposited on either side of the barrier wall. It may become so great that the exchanger will have to be removed from service periodically for cleaning.

Chemical and mechanical methods may be used to remove the dirt and scale. For mechanical cleaning, the exchanger is removed from service and opened up. Perhaps the entire tube bundle is pulled from the exchanger shell if the plant layout has provided space for this to be done. If the deposit is on the inside of straight tubes, cleaning may be accomplished merely by forcing a long worm or wire brush through each tube.

More labor is required to remove deposits on the shell side. After removal of the tube bundle, special cleaning methods such as sandblasting may be necessary.

Much engineering effort has gone into the design of heat exchangers to allow for fouling. However, it has been suggested that methods are available to design heat exchangers that, by accommodating a certain amount of dirt in a thermal design, will allow heat exchangers to run forever without shutdown for cleaning. Commercial units designed in this fashion are in operation.

Another operating problem is allowance for differential thermal expansion of metallic parts. Most operating difficulties arise during the startup or shutdown of equipment. Therefore, the following general rules have been suggested: (1) Startup. Always introduce the cooler fluid first. Add the hotter fluid slowly until the unit is up to operating conditions. Be sure the entire unit is filled with fluid and there are no pockets or trapped inert gases. Use a bleed valve to remove trapped gases. (2) Shutdown. Shut off the hot fluid first, but do not allow the unit to cool too rapidly. Drain any materials which might freeze or solidify as the exchanger cools. (3) Steam condensate. Always drain any steam condensate from heat exchangers when starting up or shutting down. This reduces the possibility of water hammer caused by steam forcing the trapped water through the lines at high velocities. *See* CONDUCTION (HEAT); CONVECTION (HEAT); COOLING TOWER; DISTILLATION; EVAPORATOR; FURNACE CONSTRUCTION; HEAT RADIATION; HEAT TRANSFER; VAPOR CONDENSER. Raymond F. Fremed

Bibliography. N. Afgan et al. (eds.), *New Developments in Heat Exchangers*, 1996; S. Kakaç and H. Liu, *Heat Exchangers: Selection, Rating and Thermal Design*, 2d ed., 2002; T. Kuppan, *Heat Exchanger Design Handbook*, 2000; R. K. Shah and D. P. Sekulic, *Fundamentals of Heat Exchanger Design*, 2002.

Heat insulation

Materials whose principal purpose is to retard the flow of heat. Thermal- or heat-insulation materials may be divided into two classes, bulk insulations and reflective insulations. The class and the material within a class to be used for a given application depend upon such factors as temperature of operation, ambient conditions, mechanical strength requirements, and economics.

Examples of bulk insulation include mineral wool, vegetable fibers and organic papers, foamed plastics, calcium silicates with asbestos, expanded vermiculite, expanded perlite, cellular glass, silica aerogel, and diatomite and insulating firebrick. They retard the flow of heat, breaking up the heat-flow path by the interposition of many air spaces and in most cases by their opacity to radiant heat.

Reflective insulations are usually aluminum foil or sheets, although occasionally a coated steel sheet, an aluminumized paper, or even gold or silver surfaces are used. Refractory metals, such as tantalum, may be used at higher temperatures. Their effectiveness is due to their low emissivity (high reflectivity) of heat radiation. *See* EMISSIVITY.

Thermal insulations are regularly used at temperatures ranging from a few degrees above absolute zero, as in the storage of liquid hydrogen and helium, to above 3000°F (1650°C) in high-temperature furnaces. Temperatures of 4000–5000°F (2200–2760°C) are encountered in the hotter portions of missiles, rockets, and aerospace vehicles. To withstand these temperatures during exposures lasting seconds or minutes, insulation systems are designed that employ radiative, ablative, or absorptive methods of heat dissipation.

Heat flow. The distinguishing property of bulk thermal insulation is low thermal conductivity. Under conditions of steady-state heat flow the empirical equation that describes the heat flow through a material is Eq. (1), where q = time rate of heat flow,

$$\frac{q}{A} = -k\frac{\theta_2 - \theta_1}{l} \tag{1}$$

A = area, θ_1 = temperature of warmer side, θ_2 = temperature of colder side, l = thickness or lenght of heat-flow path, and k = thermal conductivity, representative values being listed in the **table**. For a given thickness of material exposed to a given temperature difference, the rate of heat flow per unit area is directly proportional to the thermal conductivity of the material. *See* CONDUCTION (HEAT).

In the unsteady state, or transient heat flow, the density and specific heat of a material have a strong influence upon the rate of heat flow. In such cases, thermal diffusivity $\alpha = k/\rho\ C_p$ is the important property. Here ρ = density and C_p = specific heat at constant pressure. In the simple case of one-dimensional heat flow through a homogeneous material, the governing equation is Eq. (2), where t =

$$\frac{d\theta}{dt} = \alpha \left.\frac{d^2\theta}{dx^2}\right|_0^l \tag{2}$$

time and x is measured along the heat-flow path from 0 to l.

Thermal conductivity. In general, thermal conductivity is not a constant for the material but varies with temperature. Generally, for metals and other crystalline materials, conductivity decreases with increasing temperature; for glasses and other amorphous materials, conductivity increases with temperature. Bulk insulation materials in general behave like amorphous materials and have a positive temperature coefficient of conductivity.

Thermal conductivity of bulk insulation depends upon the nature of the gas in the pores. The conductivities of two insulations, identical except for the gases filling the pore spaces, will differ by an amount approximately proportional to the difference in the conductivities of the two gases.

Thermal conductivities of selected solids*

Material	Density, lb/ft³ (kg/m³)	Temperatures, °F (°C)	Conductivity (k),† Btu/(h)(ft²)(°F/in.) [W/(m²)(°C/m)]
Asbestos cement board	120 (1920)	75 (24)	4 [0.6]
Cotton fiber	0.8–2.0 (15–32)	75 (24)	0.26 [0.037]
Mineral wool, fibrous rock, slag, or glass	1.5–4.0 (24–64)	75 (24)	0.27 [0.039]
Insulating board, wood, or cane fiber	15 (240)	75 (24)	0.35 [0.050]
Foamed plastics	1.6 (26)	75 (24)	0.29 [0.042]
Glass			3.6–7.32 [0.52–1.056]
Hardwoods, typical	45 (720)	75 (24)	1.10 [0.159]
Softwoods, typical	32 (510)	75 (24)	0.80 [0.115]
Cellular glass	9 (150)	75 (24)	0.40 [0.058]
Fine sand (4% moisture content)	100 (1600)	40 (4)	4.5 [0.65]
Silty clay loam (20% moisture content)	100 (1600)	40 (4)	9.5 [1.37]
Gypsum or plaster board	50 (800)	75 (24)	1.1 [0.16]

*From American Society of Heating, Refrigerating, and Air Conditioning Engineers, *Heating, Ventilating and Air Conditioning Guide*, 1959.
†Typical: suitable for engineering calculations.

Increasing the pressure of the gas in the pores of a bulk insulation has little effect on the conductivity even with pressures of several atmospheres. Decreasing the pressure has little effect until the mean free path of the gas is in the order of magnitude of the dimensions of the pores. Below this pressure the conductivity decreases rapidly until it reaches a value determined by radiation and solid conduction. A few materials have such fine pores that at atmospheric pressure their dimensions are smaller than the mean free path of air. Such insulations may have conductivities less than that of still air. *See* HEAT RADIATION; HEAT TRANSFER. Harry F. Remde

Bibliography. D. Brandreth (ed.), *Improved Thermal Insulation*, 1991; D. E. Croy and D. A. Dougherty, *Handbook of Thermal Insulation Applications*, 1984; R. M. Diamant, *Thermal and Acoustic Insulation*, 1986; F. J. Powell and S. L. Matthews (eds.), *Thermal Insulation: Materials and Systems*, 1987.

Heat pipe

A device for transferring heat efficiently between two locations by using the evaporation and condensation of a fluid contained therein. Originally developed for refrigeration and spacecraft thermionic generator applications, heat pipes (and their close relative, the thermosiphon) have become routine components in systems ranging from missiles to domestic tuner-amplifiers and furnaces. They have many applications in areas where reliable performance and low cost are of prime importance—for example, in electronics and heat exchangers.

Operating principle. The heat pipe, the idea of which was first suggested in 1942, is similar in many respects to the thermosiphon. A large proportion of applications for heat pipes do not use heat pipes as strictly defined below, but employ thermosiphons (**illus.** *a*), sometimes known as gravity-assisted heat pipes. A small quantity of liquid is placed in a tube from which the air is then evacuated, and the tube sealed. The lower end of the tube is heated, causing liquid to vaporize and the vapor to move to the cooler end of the tube, where it condenses. The condensate is returned to the evaporator section by gravity. Since the latent heat of evaporation is generally high, considerable quantities of heat can be transported with a very small temperature difference between the two ends. Thus the structure has a high effective thermal conductance. The thermosiphon, also known as the Perkins tube, has been used for many years. A wide variety of working fluids have been employed, ranging from helium to liquid metals.

One limitation of the basic thermosiphon is that in order for the condensate to be returned by gravitational force to the evaporator region, the latter must be situated at the lowest point. The heat pipe is similar in construction to the thermosiphon, but in this case provision is made for returning the condensate against a gravity head. Commonly a wick,

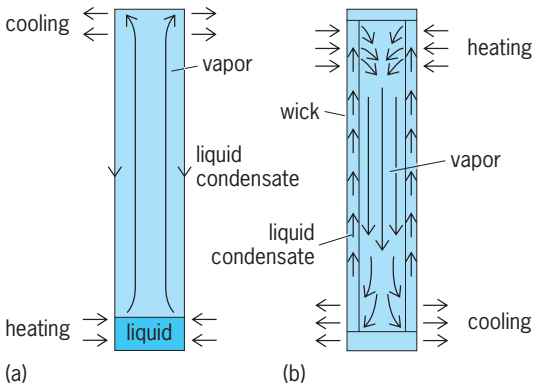

Heat transfer devices. (a) Thermosiphon, (b) Heat pipe. The heat pipe can be in any position, not just vertical as shown.

for example a few layers of fine gauze, is used. This is fixed to the inside surface of the tube, and capillary forces return the condensate to the evaporator (illus. *b*). Since the evaporator position is not restricted, the heat pipe may be used in any orientation. If, of course, the heat pipe evaporator happens to be in the lowest position, gravitational forces will assist the capillary force. Alternative techniques, including centripetal forces, osmosis, and others, may be used for returning the condensate to the evaporator (see **table**).

Capillary forces are by far the most common form of condensate return employed, but a number of rotating heat pipes are used for cooling of electric motors and other rotating machinery. In some applications a mechanical pump is used to return condensate in two-phase run-around coil heat recovery systems. While this may be regarded as a retrograde step, it is a much more effective method for condensate return than reliance on capillary forces.

Applications. The applications for which heat pipes are developed are related to five principal functions of the heat pipe: separation of heat source and sink, temperature flattening, heat flux transformation, temperature control, and acting as a thermal diode or switch. The two major applications, cooling of electronic components and heat exchangers, can involve all of these features. In the case of electronics cooling and temperature control, all features can be important. In heat exchangers employing heat pipes,

Methods of condensate return in heat pipes

Method	Device
Gravity	Standard thermosiphon
Capillary forces	Standard heat pipe
Centripetal forces	Rotating heat pipe
	Rotary heat pipe
Electrostatic volume forces	Electrohydrodynamic heat pipe
Magnetic volume forces	Magnetohydrodynamic heat pipe
Osmotic forces	Osmotic heat pipe
Vapor bubble pump	Inverse thermosiphon
Mechanical pump	Two-phase run-around coil

the separation of heat source and sink, and the action as a thermal diode or switch, are most significant.

Cooling of electronic components. The use of heat pipes for cooling or temperature control of electronic components is the largest field of application, and a wide variety of heat pipe geometries may also be employed. The basic tubular heat pipe remains the most common type, but flat heat pipes, direct-contact systems, and flexible heat pipes all play an increasingly important role. The most widely applied heat pipes in electronics cooling are units of conventional form that were developed for cooling of semiconductor devices and are manufactured by using a low-cost production method for spirally grooved wicks. By using water as the working fluid in a copper container, heat transport capabilities in excess of 1 kW are achieved. The heat pipes, which normally operate horizontally, typically have a diameter of 0.6 in. (16 mm), and give a 30% improvement in heat dissipation and a 50% weight reduction when compared with a conventional heat sink.

A wide range of audio amplifiers employ heat-pipe heat sink systems of this type. As well as improving heat dissipation, the use of heat pipes also yields electrical benefits, including improvements in the distortion factor. The heat pipes are used to transfer heat to the periphery of the amplifier, where a finned heat sink can dissipate it to atmosphere. This prevents local overheating within the electronics enclosure.

A second major application is in power amplifiers. The heat pipe radiator is used as a heat sink for high-power transistors and thyristors. It contributes to low package sizes and general costbenefits.

Heat exchangers. The use of bundles of multiple heat pipes as heat exchangers is second only to electronics thermal control in terms of number of units in operation. While heat-pipe heat exchangers in the form of air-air or air-liquid heat recovery systems are used in domestic and commercial buildings for conserving energy in heating, ventilating, and air conditioning, it is in the industrial field that the heat-pipe heat exchanger has the most applications. Gas-gas heat-pipe heat exchangers are used in a wide range of industrial processes, normally to recover heat from hot or humid exhaust streams to preheat make-up air. Heat may also be recovered for space heating.

The conventional heat-pipe heat exchanger is available in sizes covering heat transport duties ranging from a few tens of watts to several megawatts. The ability to separate source and sink conveniently has led to many derivatives which show the versatility of heat pipe and thermosiphon heat exchangers. These include systems for recovering heat from wastewater, condensers, and exchangers for improving the performance of domestic heat pumps; heat recovery in flue gas desulfurization; and heat transfer in spent nuclear fuel stores. Heat-pipe waste heat boilers (which use hot gases to generate steam for processes or space heating) have been used in situations where reactive fluids have to be separated, and the ability of heat-pipe heat exchangers (in common with the run-around coil, another heat recovery system) to accommodate different external surfaces on the evaporators and condensers benefits many energy-intensive processes. *See* HEAT EXCHANGER; HEAT TRANSFER.

David A. Reay

Bibliography. P. D. Dunn and D. A. Reay, *Heat Pipes*, 4th ed., 1994; D. A. Reay, *Advances in Heat Pipe Technology*, 1982; M. Terpstra and J. G. Van Veen, *Heat Pipes: Construction and Application*, 1987; L. L. Vasiliev, Low temperature heat pipes, *J. Heat Recovery Sys.*, 5:203–216, 1985.

Heat pump

The thermodynamic counterpart of the heat engine. A heat pump raises the temperature level of heat by means of work input. In its usual form a compressor takes refrigerant vapor from a low-pressure, low-temperature evaporator and delivers it at high pressure and temperature to a condenser (**Fig. 1**). The pump cycle is identical with the customary vapor-compression refrigeration system. *See* REFRIGERATION CYCLE.

Application to comfort control. For air-conditioning in the comfort heating and cooling of space, a heat pump uses the same equipment to cool the conditioned space in summer and to heat it in winter, maintaining a comfortable temperature at all times (**Fig. 2**). *See* AIR CONDITIONING; COMFORT HEATING.

This dual purpose is accomplished, in effect, by placing the low-temperature evaporator in the conditioned space during the summer and the high-temperature condenser in the same space during the winter (**Fig. 3**). Thus, if 70°F (21°C) is to be maintained in the conditioned space regardless of the season, this would be the theoretical temperature of the evaporating coil in summer and of the condensing coil in winter. The actual temperatures on the refrigerant side of these coils would need to be below 70°F in summer and above 70°F in winter to permit the necessary transfer of heat through the coil surfaces.

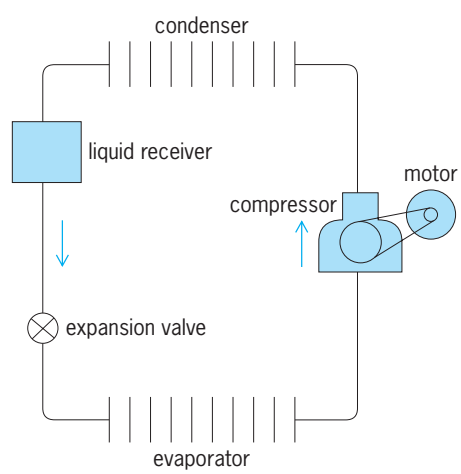

Fig. 1. Basic flow diagram of heat pump with motor-driven compressor. For summer cooling, condenser is outdoors and evaporator indoors; for winter heating, condenser is indoors and evaporator outdoors.

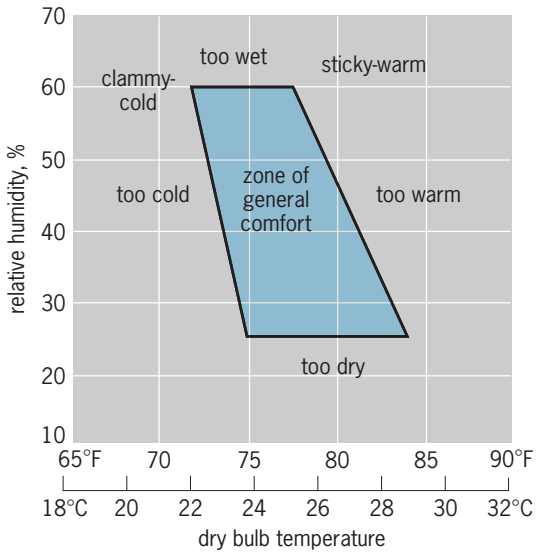

Fig. 2. Indoor climatic conditions acceptable to most people when doing desk work; continuous air motion with 5–8 air changes per hour.

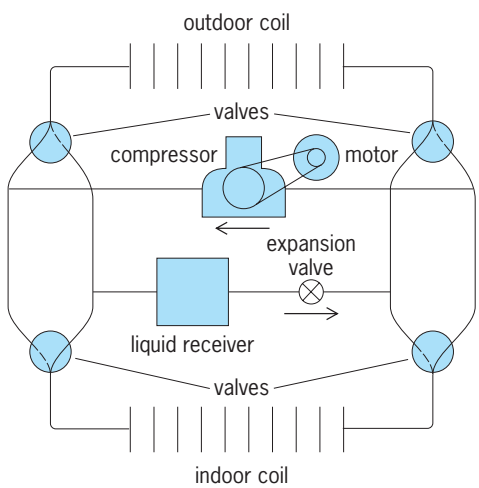

Fig. 3. Air-to-air heat pump installation; fixed air circuit with valves in the summer positions (the broken lines show the winter positions).

If the average outside temperatures are 100°F (38°C) in summer and 40°F (4°C) in winter, the heat pump serves to raise or lower the temperature 30°F (17°C) and to deliver the heat or cold as required. The ultimate ideal cycle for estimating performance is the same Carnot cycle as that for heat engines. The coefficient of performance COP_c as cooling machine is given in Eq. (1), and the coefficient COP_w

$$COP_c = \frac{\text{refrigeration}}{\text{work}} = \frac{T_c}{T_h - T_c} \qquad (1)$$

as a warming machine is given in Eq. (2), where T is

$$COP_w = \frac{\text{heat delivered}}{\text{work}} = \frac{T_h}{T_h - T_c} \qquad (2)$$

temperature in degrees absolute (Rankine) and the subscripts c and h refer to the cold and hot temperatures, respectively.

For the data cited, the theoretical coefficients of performance are as in Eqs. (3). The significance of

$$\begin{aligned} COP_c &= \frac{460 + 70}{(460 + 100) - (460 + 70)} = 17.7 \\ COP_w &= \frac{460 + 70}{(460 + 70) - (460 + 40)} = 17.7 \end{aligned} \qquad (3)$$

these coefficients is that ideally for 1 kilowatt-hour (kWh) of electric energy input to the compressor there will be delivered 3413 × 17.7 = 60,000 Btu/h as refrigeration or heating effect as required. This is a great improvement over the alternative use of resistance heating, typically, where 1 kWh of electric energy would deliver only 3413 Btu. The heat pump uses the second law of thermodynamics to give a much more substantial return for each kilowatt-hour of electric energy input, since the electric energy serves to move heat which is already present to a desired location. *See* CARNOT CYCLE; THERMODYNAMIC PRINCIPLES.

Effect of seasonal loads. When summer comfort cooling of space is desired, it is entirely possible and practical to use the same compressor equipment and coils for winter heating and for summer cooling and to dispense with the need for direct-fired apparatus using oil, gas, coal, or wood fuel.

For an economical installation, equipment must be of correct size for both the summer cooling and winter heating loads. Climatic conditions have a significant influence and can lead to imbalance on sizing. If the heating and cooling loads are equal, the equipment can be selected with minimum investment. However, generally the loads are not balanced; in the temperate zone the heating load is usually greater than the cooling load. This necessitates (1) a large, high-horsepower compressor fitted to the heating demand, (2) a supplementary heating system

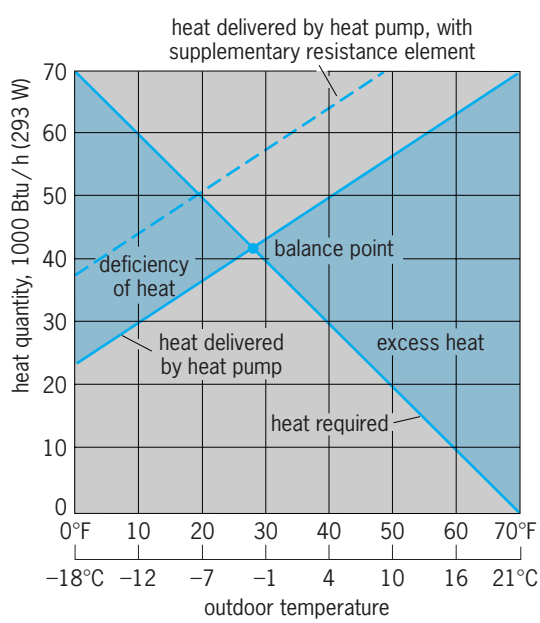

Fig. 4. Performance of air-to-air, self-contained, domestic heat pump on the heating cycle.

(electrical resistance, fuel, or solar energy), or (3) a heat-storage system. *See* COMPRESSOR; RESISTANCE HEATING; SOLAR ENERGY.

If well water or the ground serves as the heat source, the imbalance is less severe than when atmospheric air is the source. However, the uncertain heat transfer rates with ground coil, the impurities, quality, quantity, and disposal of water, and the corrosion problems mitigate the use of these sources.

Atmospheric air as a heat source is preferable, particularly with smaller domestic units. A self-contained, packaged unit of this type offers maximum dependability and minimal total investment. The performance of such a unit for the heating cycle is illustrated in **Fig. 4**. Curves of heat required and heat available show the limitations on capacity. The heat delivered by the pump is less than the heat required at low temperatures, so that there is a deficiency of heat when the outside temperature goes below, in this case, about 28°F (2°C). The intersection of the two curves is the balance point. There is an area of excess heat to the right and of deficiency of heat to the left of the balance point. Many devices and methods are offered to correct this situation, such as storage systems, supplementary heaters, and compressors operating alternatively in series or in parallel.

In temperate regions heat-pump installations achieve coefficients of performance on the order of 3 on heating loads when all requirements for power, including auxiliary pumps, fans, resistance heaters, defrosters, and controls are taken into account. Automatic defrosting systems, when air is the heat source, are essential for best performance, with the defrost cycle occurring twice a day.

Heat pumps are uneconomical if used for the sole purpose of comfort heating. The direct firing of fuels is generally more attractive from an overall financial viewpoint. The investment in heat-pump equipment is higher than that for the conventional heating system. Unless the price of electric energy is sufficiently low or the price of fuels very high, the heat pump cannot be justified solely as a heating device. But if there is need for comfort cooling of the same space in summer, the heat pump, to do both the cooling and heating, becomes attractive.

The heat pump is also used for a wide assortment of industrial and process applications such as low-temperature heating, evaporation, concentration, and distillation. Theodore Baumeister

Solar energy-assisted system. For the heat pump to be more attractive in colder regions, where ambient temperatures remain about (or below) 0°F (−18°C) for long periods, the source temperature for the heat pump should be increased. One possible method for achieving this is solar energy.

Solar collectors. For space heating utilizing solar energy, flat-plate collectors are generally used. A typical flat-plate collector contains a metallic plate painted black, with one or two glass covers and the sides and bottom of the collector well insulated. Solar energy is transmitted through the glass, and a significant part of that reaching the metallic plate—about 90%—is absorbed by the plate, increasing its temperature. The energy absorbed by the plate is, in turn, transferred to a working fluid—usually air or water. If air is the working fluid, it passes over the metallic plate; if water, it flows through tubes attached to the metallic plate.

Because of the heat transfer from the collector to the surroundings, not all the radiant energy reaching the collector is transferred to the working fluid. The efficiency of a collector, defined as the ratio of the heat transfer to the working fluid to the radiant energy reaching the collector surface, decreases with an increase in the absorber temperature. The higher the temperature at which the working fluid is operated, the higher is the temperature of the absorber surface and the lower the efficiency.

Working fluid and size. A heat pump system utilizing solar energy for space heating is schematically illustrated in **Fig. 5**. Air is the working fluid for the solar collectors, and the system uses an air-to-air heat pump. Energy storage is necessary as, during some periods, solar energy may not be available at all, such as during the night, or may be available in very limited quantities, such as on cloudy days. The system can operate in three different modes (Fig. 5). In mode 1, the solar energy stored is not at a sufficiently high temperature for direct use but at a temperature higher than ambient temperature—conditions likely to be met on cold, sunny days. The heat pump uses

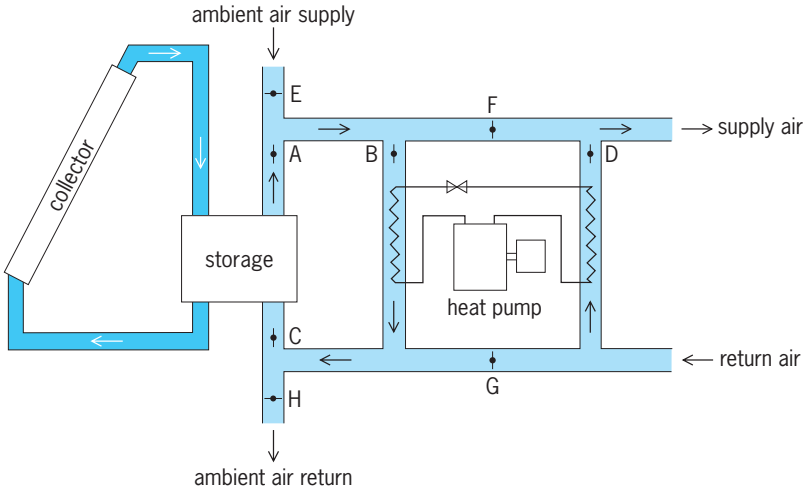

	open	close	mode
1	A,B,C,D	E,F,G,H	temperature of stored energy too low for direct use but higher than ambient air
2	A,C,F,G	E,B,D,H	stored solar energy temperature sufficiently high; direct utilization of solar energy (bypass heat pump)
3	B,D,E,H	A,C,F,G	stored solar energy temperature lower than ambient air temperature; heat pump source is ambient air

Fig. 5. Schematic arrangement of solar-assisted heat-pump system.

the stored solar energy in this mode. In mode 2, the stored solar energy is at a sufficiently high temperature for direct utilization—probable conditions during sunny days in the fall and spring. The heat pump is then bypassed. In mode 3—during extended periods of cloudy days or very cold weather—the stored solar energy is depleted and outside air is the source for heat pumps. *See* SOLAR HEATING AND COOLING.

N. V. Suryanaryana

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; L. Miles, *Heat Pumps: Theory and Service*, 1993; M. Taylor and J. Roberts, *Heat Pump Systems*, 1992.

Heat radiation

The energy radiated by solids, liquids, and gases as a result of their temperature. Such radiant energy is in the form of electromagnetic waves and covers the entire electromagnetic spectrum, extending from the radio-wave portion of the spectrum through the infrared, visible, ultraviolet, x-ray, and γ-ray portions. From most hot bodies on Earth this radiant energy lies largely in the infrared region. *See* ELECTROMAGNETIC RADIATION; INFRARED RADIATION.

Radiation is one of the three basic methods of heat transfer, the other two methods being conduction and convection. *See* CONDUCTION (HEAT); CONVECTION (HEAT); HEAT TRANSFER.

A hot plate at 400 K (260°F) may show no visible glow; but a hand which is held over it senses the warming rays emitted by the plate. A temperature of more than 1000 K (1300°F) is required to produce a perceptible amount of visible light. At this temperature a hot plate glows red and the sensation of warmth increases considerably, demonstrating that the higher the temperature of the hot plate the greater the amount of radiated energy. Part of this energy is visible radiation, and the amount of this visible radiation increases with increasing temperature. A steel furnace at 1800 K (2800°F) shows a strong yellow glow. If a tungsten wire (used as the filament in incandescent lamps) is raised by resistance heating to a temperature of 2800 K (4600°F), it emits a bright white light. As the temperature of a substance increases, additional colors of the visible portion of the spectrum appear, the sequence being first red, then yellow, green, blue, and finally violet. The violet radiation is of shorter wavelength than the red radiation, and it is also of higher quantum energy.

In order to produce strong violet radiation, a temperature of almost 3000 K (5000°F) is required. Ultraviolet radiation necessitates even higher temperatures, and there is no solid on Earth which can withstand such temperatures without melting. The Sun emits considerable ultraviolet radiation, as evidenced by the sunburn it produces. The spectral distribution of the Sun's radiation has been measured, and the temperature of the Sun's surface has been determined from Wien's displacement law and corresponds to about 6000 K or 10,000°F (Wien's law is discussed later). Such temperatures have been produced on Earth in gases ionized by electrical discharges. The vapor lamps used on highways, the fluorescent lamp used in offices, and the xenon compact-arc lamp used in searchlights are good examples of such gas discharges. They emit large amounts of ultraviolet radiation. Temperatures up to 20,000 K (35,000°F), however, are still much too low to produce x-rays or γ-radiation. Approaches to the utilization of nuclear energy have made use of the fusion of deuterons in magnetically constricted arcs at extremely high currents and with multiple, intense laser beams bombarding small spheres. By these means, temperatures above 1×10^6 K (2×10^6 °F) have been obtained for small fractions of a second. These devices require enormous amounts of energy to produce such high temperatures. Matter maintained at such temperatures emits x-rays and γ-rays. *See* NUCLEAR FUSION; SUN; ULTRAVIOLET RADIATION.

Heat radiators emit energy over a wide range of angles and wavelengths at a single temperature. Lasers, in a sense, are completely opposite to heat radiators, emitting a very narrow range of wavelengths within a narrow beam. In addition, a laser will not operate if the laser medium has only a single temperature. *See* LASER.

Theory. The emission of radiation is explained in terms of excited atoms and nuclei. For example, electrons in an atom can be ejected from their normal orbits around the atom into those farther from the nucleus. When this happens the atom is said to be in an excited state. This occurs when energy, supplied from outside a substance, is converted into thermal motion and finally into excitation. A short time after excitation, the electrons return to their normal orbits and give off their excess energy ΔE in the form of radiation of a particular frequency ν. This wavelength may be determined by the relation $\Delta E = h\nu$, where h is Planck's constant. *See* ATOMIC STRUCTURE AND SPECTRA; PLANCK'S CONSTANT.

In a gas, the thermal motion consists of substantially unhindered movement of the individual particles with different velocities. In a solid, on the other hand, the thermal motion is an oscillating movement of the particles, with varying displacements, about their fixed positions. The extent of the thermal motion depends upon the temperature. The hotter the substance, the greater the thermal motion and the higher the intensity and energy of the radiation. An energy distribution of the radiation intensity results, for example, from the distribution of velocities of the particles in a gas or from the distribution of displacements of the particles about their positions in a solid.

Further, the maximum available energy (excitation energy) depends upon temperature, and this explains why the energy of emitted radiation shifts to shorter wavelength (that is, higher energy) as the temperature is increased. For instance, a temperature of 1000 K (1300°F) produces just enough excitation energy for the dark red glow which contains

the longest wavelengths within the visible portion of the spectrum. As explained before, higher temperatures or greater excitation energies are necessary to excite measurable quantities of the shorter wavelength regions. It is obvious that with decreasing temperatures, less excitation energy is available, and the amount of heat radiation decreases until finally at absolute zero of temperature (0 K or −459.67°F) substances radiate no energy because all atomic motion has ceased. However, for a definition of zero point energy *see* QUANTUM MECHANICS

The radiated energy per second is commonly expressed in terms of joules per second, or watts. Other units often used are ergs per second or calories per second. These are related to each other as follows: 1 watt = 1 joule/s = 10^7 ergs/s = 0.239 cal/s. For instance, the Sun radiates onto 1 cm^2 of the Earth's surface 2 cal/min or $^1/_{30}$ cal/s or about $^1/_7$ W. The total energy radiated from 1 cm^2 (0.155 $in.^2$) of a tungsten wire in an incandescent lamp at 2800 K (4600°F) is 112 W. The same wire at room temperature emits only 0.0015 W. *See* SOLAR CONSTANT.

Energy distribution curves. To evaluate the usefulness of a heat radiator, energy distribution curves are used. These are graphs of relative or absolute radiated energy versus the wavelength of radiation (expressed in micrometers, nanometers, or angstroms) or the frequency (velocity of light/wavelength) expressed in hertz (cycles per second).

Such graphs show how the energy radiated from a substance at a certain temperature is distributed over the various portions of the spectrum (**Fig. 1**). The usefulness of these graphs lies in the fact that they provide information, for example, on the effectiveness of a radiator as a light source or as a heating element. Furthermore, the area under the energy distribution curve is equivalent to the total radiated energy.

The energy distribution of various substances differs because of their internal properties and their surface condition. As a common rule, which holds well above 3 micrometers, substances with good electrical conductivity, especially metals, are poor emitters of radiation and are good reflectors (for example, silver or aluminum). Insulators radiate strongly in the infrared region of the spectrum and have gaps of low radiation intensity near the visible portion of the spectrum, as shown in Fig. 1*c*. These gaps are due to the electronic band structure of insulators. Roughening the surface of all radiators increases the emitted energy. This is true because tiny holes in the surface act as cavity radiators, radiating almost blackbody energy. (Cavity radiators and blackbody radiation are discussed later.) *See* ELECTRIC INSULATOR.

Energy distribution curves are obtained by passing white (heterochromatic) light through a monochromator (quartz prism, grating, and the like) and measuring the spectral intensity of radiation with a phototube or a thermopile. The measured intensities at the various wavelengths are then plotted either as percent of the maximum intensity (relative energy)

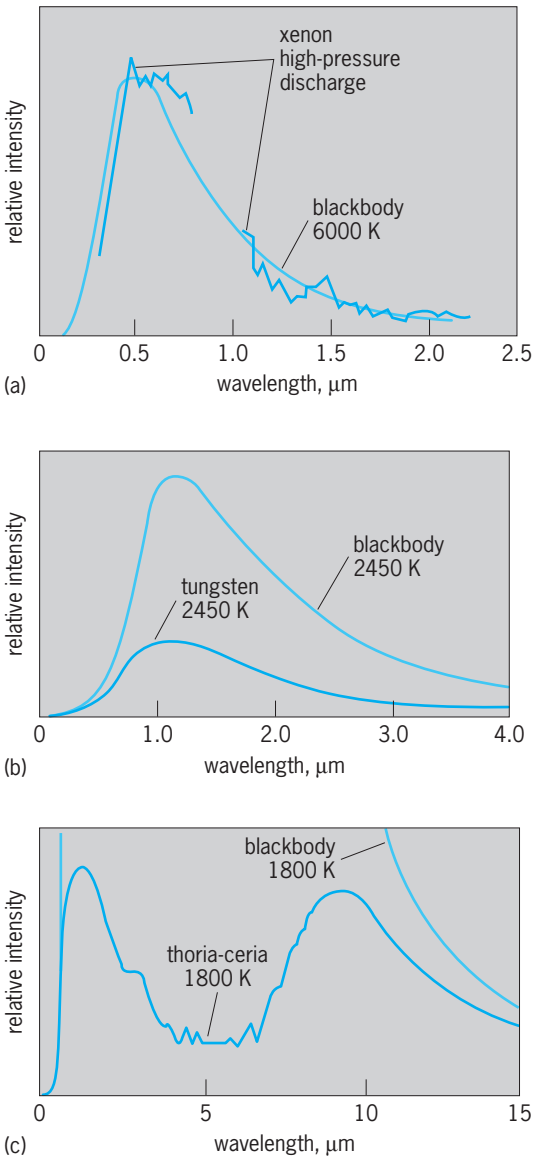

Fig. 1. Energy distribution curves for (*a*) xenon high-pressure electrical gas discharge (*after W. Meyer, ed., Technischwissen-schaftliche Abhandlungen aus dem Osramkonzern, vol. 6, Springer, 1953*); (*b*) tungsten (*after American Institute of Mining and Metallurgical Engineers, Pyrometry: The Papers and Discussions of a Symposium on Pyrometry, 1920*); (*c*) thoria plus 1% ceria, a typical ceramic (*after R. W. Pohl, Einführung in die Optik, Springer, 1948*). °F = (K × 1.8) − 460.

or as absolute intensity (absolute energy) versus the wavelength. *See* RADIOMETRY.

A radiator used in heating rooms should produce much infrared radiation (heat) and no light, whereas much visible light and little heat is desired from a light source. Unfortunately, an energy distribution curve gives a true picture of the radiator for one particular temperature only. If more information is needed, a set of such curves would have to be provided for the temperature range of interest.

Blackbody radiation. Because of the tedious experimental work involved in determining such curves, a different and more fruitful approach is generally taken. Two quantities characterize a heat radiator completely: the total emissivity and the spectral

emissivity, which are designated by ϵ and ϵ_λ, respectively, where the subscript λ designates wavelength. Both emissivities, in conjunction with the radiation properties of a blackbody, describe fully the behavior of a real heat radiator. The radiation properties of a blackbody are completely stated by Planck's radiation law.

Planck's law and the concept of blackbody radiation are of utmost importance for the understanding of heat radiation. The blackbody signifies in the domain of heat radiation what any other standard, such as the standard meter, signifies in its own domain. A blackbody is defined as a body which emits the maximum amount of heat radiation. Although there exists no perfect blackbody radiator in nature, it is possible to construct one on the principle of cavity radiation.

A cavity radiator is usually understood to be a heated enclosure with a small opening which allows some radiation to escape or enter. The escaping radiation from such a cavity has the same characteristics as blackbody radiation. Radiation energy which enters the cavity is almost completely absorbed because of the multiple reflections it encounters (**Fig. 2**). This

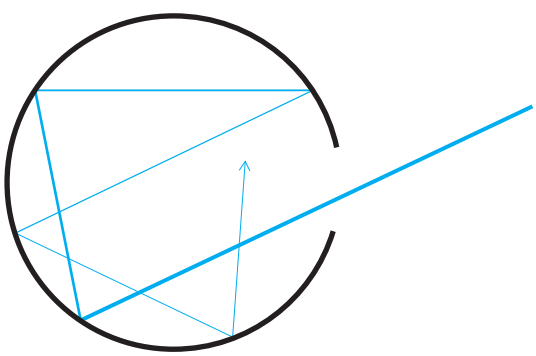

Fig. 2. Diagram of a cavity radiator.

follows because at each reflecting point some of the energy is absorbed by the walls. The absorptivity of the cavity hole is essentially unity, independent of the wall material. As a consequence of Kirchhoff's law, the emissivity of the cavity is also unity, and this fulfills the definition of a blackbody radiator. In practice, the cavity is approximated by a small hole or even a wedge cut into a surface. *See* BLACKBODY.

Kirchhoff's law. This law correlates mathematically the heat radiation properties of materials at thermal equilibrium. It is often called the second law of thermodynamics for radiating systems.

Kirchhoff's law can be expressed as follows: The ratio of the emissivity of a heat radiator to the absorptivity of the same radiator is a function of frequency and temperature alone. This function is the same for all bodies, and it is equal to the emissivity of a blackbody. When ϵ is the emissivity of a real radiator, α its absorptivity, and $E = 1$ the emissivity of a blackbody, Kirchhoff's law takes the form of Eq. (1).

$$\frac{\epsilon}{\alpha} = E = 1 \tag{1}$$

A substance, when brought without contact into an evacuated enclosure the walls of which are at a constant but higher temperature than the body, will assume the wall temperature after some time. However, it will not exceed it. Under these conditions, the exchange of energy can take place only by radiation. As the test body receives radiation from the walls it will absorb some of it, transforming it into motion of its elementary particles, and thereby raising its own temperature. Thermal equilibrium is obtained when the temperature of the walls and the test body is the same; in this case the test body must emit as much energy as it receives. If it absorbs all the impinging radiation, it is a blackbody. If it absorbs only a fraction of the impinging radiation, the other part must be reflected in order to maintain the equilibrium. These statements require that the absorptivity be equal to the emissivity. This is the form in which Kirchhoff's law is often stated. For opaque bodies, absorptivity plus reflectivity must be equal to unity, and therefore the emissivity and the absorptivity respectively must be unity minus the reflectivity. A consequence of Kirchhoff's law is the postulate that a blackbody has an emissivity which is greater than that of any other body.

Planck's radiation law. This celebrated law represents mathematically the energy distribution of the heat radiation from 1 cm^2 of surface area of a blackbody at any temperature. It is the only heat radiation law which is accurate throughout the entire spectrum. The basis of Planck's radiation law was experimental data obtained from measurements on cavity radiators.

Planck's radiation law has great importance. Formulated by Max Planck early in the twentieth century, it laid the foundation for the advance of modern physics and the advent of quantum theory. In determining the heat radiation of hot bodies Planck's radiation law is a basic tool in research and development, both in science and industry. The radiation law can be used to predict light output of incandescent lamps, the cooling time of molten steel, heat dissipation of nuclear reactors, the energy radiated from the Sun, the temperature of the stars, and many other important applications.

Although Planck's law can be derived on theoretical grounds alone, it was deduced from experiment. Prior attempts to calculate the heat radiation of a blackbody had described the radiation as consisting of electromagnetic waves whose energy content could vary continuously. Those attempts did not match the experimental results. Planck replaced the concept of continuous energy with the idea that the energy existed in bundles; that is, the energy was quantized. *See* QUANTUM MECHANICS.

This concept was a drastic innovation at that time. However, upon calculating the radiation, Planck found that the expression in Eq. (2) described the

$$R_\lambda = \frac{c_1}{\lambda^5[e^{c_2/(\lambda T)} - 1]} \tag{2}$$

experimental results completely. This is the mathematical expression of Planck's radiation law, where

R_λ is the total energy radiated from the body measured in watts per unit area per unit wavelength, at the wavelength λ. The quantity T is the temperature in kelvins, and e is the base of the natural logarithms.

Planck found that $c_1 = 2\pi hc^2$ and $c_2 = hc/k$ where h, c, and k are Planck's constant, the velocity of light, and Boltzmann's constant, respectively. In SI units, $c_1 = 3.7418 \times 10^{-16}$ watt meter^2, $c_2 = 1.4388 \times 10^{-2}$ meter kelvin, and R_λ is in watts per square meter per meter of wavelength. More practical units use area in square centimeters and wavelength in micrometers. Then $c_1 = 37{,}418$ and $c_2 = 14{,}388$, while R_λ is in watts per square centimeter per micrometer of wavelength. Planck's law can be graphed for various temperatures (**Fig. 3**). However, while various substances attain these temperatures, these substances will not radiate as predicted by Planck's law since they are not blackbodies themselves.

The radiation increases at every point of the energy spectrum as the temperature is increased (Fig. 3). At all temperatures, the energy radiated at the extremes of the energy spectrum approaches zero and has a maximum at some place in between. The total area under any of the energy distribution curves (Fig. 1) measures the total energy radiated by the body at the temperature represented by the curve.

Three important aspects of Planck's radiation law can be examined. First, the behavior of the law at the extremes of the energy spectrum leads to a discussion of Wien's radiation law, the Rayleigh-Jeans radiation law, and the so-called ultraviolet catastrophe. Second, the shift of the wavelength at which the maximum energy is radiated can be studied as the temperature is changed. This leads to Wien's displacement law. Finally, the total amount of energy radiated at any temperature can be investigated. This leads to the Stefan-Boltzmann law. The four laws mentioned were well known prior to the formulation of Planck's law. It is the fact that the Planck law so neatly sums up the four earlier laws and introduces the implication of energy quantization which made it of such importance in the development of modern physics during the twentieth century.

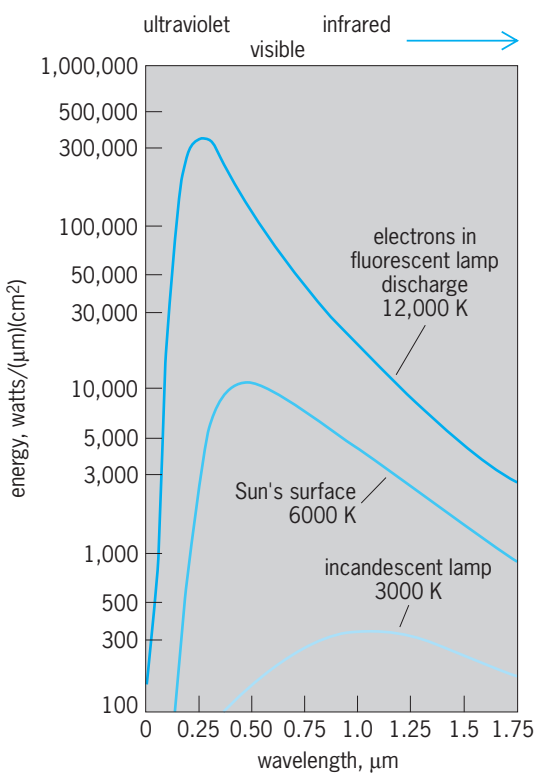

Fig. 3. Curves of Planck's law for various temperatures. Substances which attain these temperatures are also shown, although these substances will not radiate as predicted by Planck's law. °F = (K × 1.8) − 460.

Rayleigh-Jeans law. The heat radiation from a blackbody at long wavelengths is adequately described by the Rayleigh-Jeans radiation law. For larger values of λT, Planck's law simplifies to the Rayleigh-Jeans law, as shown in Eq. (3), where the numerical form

$$R_\lambda = \frac{c_1 T}{c_2 \lambda^4} = 2.6007 \frac{T}{\lambda^4} \tag{3}$$

of the equation is based on the practical units specified above. This law states that the energy radiated at any temperature increases without limit as the wavelength decreases. This law can be accurate only for wavelengths much larger than that at which the maximum occurs (Figs. 1 and 3). For wavelengths shorter than this maximum, the energy radiated from a blackbody actually decreases again. If a blackbody acted as predicted by the Rayleigh-Jeans law, then the energy radiated at very short wavelengths, in the ultraviolet region, would become extremely large and the total energy radiated would be infinite. This is known as the ultraviolet catastrophe, and would be valid at any temperature, no matter how low.

Wien's radiation law. This law is valid at short wavelengths and is obtained from Planck's law by taking λT as very small. Planck's formula then becomes Eq. (4). This law is accurate in the visible region of

$$R_\lambda = \frac{c_1}{\lambda^5} e^{-c_2/(\lambda T)} = \frac{37{,}418}{\lambda^5} e^{-14{,}388/(\lambda T)} \tag{4}$$

the spectrum below 3000 K (5000°F).

Wien's displacement law. This law is obtained from Planck's law by the process of differentiation. It describes the shift with temperature of the wavelength at which the maximum amount of radiation occurs by Eq. (5). Thus, the product of the temperature of

$$\lambda_{\max} T = 2898 \qquad \text{(micrometer-kelvins)} \tag{5}$$

a blackbody and the wavelength at which the maximum amount of radiation occurs is a constant. Wien's law has wider significance than this, however. Dividing Planck's law by T^5 results in Eq. (6). On the right-

$$\frac{R_\lambda}{T^5} = \frac{c_1}{(\lambda T)^5 [e^{c_2/(\lambda T)} - 1]} = \frac{37{,}418}{(\lambda T)^5 [e^{14{,}388/(\lambda T)} - 1]} \tag{6}$$

hand side of Eq. (6) the wavelength and temperature always appear multiplying each other. This means

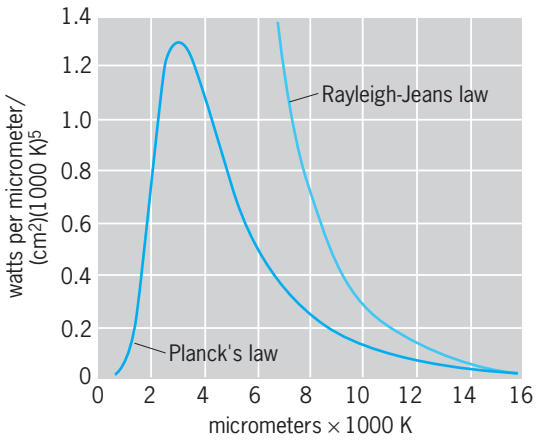

Fig. 4. Planck's law expressed as Wien's displacement law. Rayleigh-Jeans law is shown for comparison.

that only one curve is needed to express Planck's law, for all temperatures, if a graph is used in which the radiation energy divided by the fifth power of the temperature is plotted versus the product of the wavelength and temperature (**Fig. 4**). It is helpful here to measure temperature in thousands of degrees. For comparison, the Rayleigh-Jeans law can also be illustrated.

Wien's displacement law is helpful in determining the temperature of hot bodies. If λ_{max} can be measured, the temperature is immediately obtained from the displacement law. This is how an astronomer measures the temperature of a star. As an example, the radiation from the Sun's surface has a maximum in the green region of the energy spectrum in the vicinity of $\lambda = 0.5\ \mu$m (Fig. 3). From Wien's law the surface temperature of the Sun must then be about 6000 K (10,000°F).

Stefan-Boltzmann law. This law states that the total energy radiated from a hot body increases with the fourth power of the temperature of the body. This law can be derived from Planck's law by the process of integration and is expressed mathematically as Eq. (7), where R_T is the total amount of energy

$$R_T = 5.670 \times 10^{-12} T^4 \qquad (7)$$

radiated from the blackbody in watts per square centimeter. When R_T is multiplied by the total emissivity, the total energy radiated from a real heat radiator is obtained.

The rapid increase in heat radiation with temperature is quite evident from the Stefan-Boltzmann law. If the absolute temperature is doubled, say from 273 to 546 K (32 to 524°F), then the energy radiated increases 16-fold. Thus, the attainment of very high temperatures requires large amounts of energy to overcome the loss of energy by heat radiation. Temperatures greater than 3×10^7 K (5×10^7 °F) are encountered in a hydrogen bomb explosion. Such temperatures are 100,000 times higher than room temperature. Therefore, the energy radiated by 1 cm^2 (0.155 in.^{2}) of a substance at this high temperature will be 1×10^{20} times as much as that radiated at room temperature by the same substance. This energy, if radiated by a blackbody, would boil 2×10^7 tons (2×10^{10} kg) of ice water in 1 s.

Temperature determination. The apparent temperature of a real heat radiator can be determined by comparison with a blackbody whose temperature is known. This is done in any one of three customary ways, based upon the radiation laws described. (1) The radiation temperature of a surface is the temperature of the blackbody which radiates the same total energy per unit area in 1 s as does the surface. This temperature is based upon the Stefan-Boltzmann law. (2) The brightness temperature of a surface is the temperature of the blackbody which has the same brightness at a certain wavelength as has the surface. The wavelength is normally taken as 0.655 μm, and the brightness temperature thus measured can be used in conjunction with the spectral emissivity $\epsilon_{0.655}$ and Wien's radiation law to calculate the true temperature of the body. (3) The color temperature of a surface is the temperature of a blackbody whose radiation has the same or approximately the same energy distribution as the surface. This is illustrated in Fig. 1*a*, where it is seen that the xenon high-pressure discharge has a color temperature approximating 6000 K (10,000°F). For a graybody, that is, a body whose emissivity does not vary throughout the spectrum, the color temperature and true temperature are the same. *See* INCANDESCENCE.

Heinz G. Sell; Peter J. Walsh

Bibliography. M. Born, *Atomic Physics*, 8th ed., 1969, reprint 1989; R. P. Feynman, *QED: The Strange Theory of Light and Matter*, 1985; S. Fluegge (ed.), *Handbuch der Physik*, vol. 26, 1958; M. F. Modest, *Radiative Heat Transfer*, 1993; R. Siegel and J. R. Howell, *Thermal Radiation Heat Transfer*, 3d ed., 1992; M. R. Wehr et al., *Physics of the Atom*, 4th ed., 1984.

Heat storage systems

Systems that buffer temperature changes and reclaim energy that would otherwise be lost to the surroundings. Reclaiming energy increases the efficiency of systems in both industrial applications and smaller-scale applications such as household use of solar energy. In addition, heat storage materials and systems can be used in applications where the temperature needs to be kept within a certain range, such as prevention of damage to biological materials or electronic devices. The most important temperature range for heat storage materials for both industrial and consumer use is 0–100°C (32-212°F).

Heat storage materials. Heat storage systems rely on the materials within them to store thermal energy. These materials can be classified as sensible heat storage materials or phase-change materials (PCM).

Sensible heat storage. Every material stores energy within it as it is heated, and in this way is a "sensible heat storage material." The energy stored can be quantified in terms of the heat capacity C, the temperature change ΔT (= final temperature − initial

temperature), and the amount of additional heat stored ΔQ, such that $\Delta Q = C\Delta T$. The units for Q, C, and T are joules (J), J K^{-1}, and K, respectively (or Btu, Btu/°F, and °F). Clearly, other factors being equal, the higher the heat capacity (C) of a material, the greater will be the energy stored (ΔQ) for a given temperature rise (ΔT). Typical materials in a sensible heat storage device include rock, sand, wood, and water. An example of sensible heat storage is a night storage heater, which stores energy in bricks that are heated using nighttime electrical energy (available at reduced rates), and radiates this heat to the surroundings during the following day. *See* HEAT; HEAT CAPACITY; SPECIFIC HEAT.

The energy density (defined as $\Delta Q/V$, where V is the volume of the system) of a sensible heat storage system depends on both the heat capacity and the density. For example, balsa wood has a specific heat capacity (heat capacity per unit mass) of 2.9 J K^{-1} g^{-1}, which is rather high compared with that of marble (0.88 J K^{-1} g^{-1}). However, balsa has a rather low density (0.16 g cm^{-3}) compared with marble (2.6 g cm^{-3}), which means that 1 cm^3 of marble can store 2.3 J for a 1°C (or, equivalently, 1 K) temperature rise, whereas 1 cm^3 of balsa can store only 0.46 J over the same temperature increment. Ideally, it would be best to have as small a system as possible to store heat energy. One way to improve the energy density of the heat storage system is to use phase-change materials.

Phase-change materials. Although all materials increase their heat content Q as the temperature is increased, a very large increase in Q occurs when materials change phase. For example, the heat content of water increases considerably as it goes through the phase change from ice to liquid; this is the familiar melting process. The step in Q at the phase change is the latent heat associated with the transition, usually represented as $\Delta_{trs}H$ (see **illus.**). The step in Q at the transition is in addition to the sensible heat storage capacity of the material. *See* MELTING POINT.

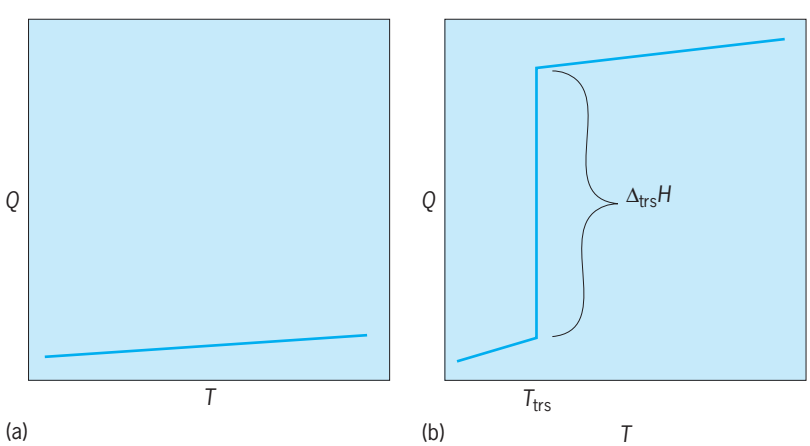

Heat content *Q* as a function of temperature *T*. (*a*) *Q* increases with increasing temperature, even if there is no phase transition, as in a sensible heat storage material. (*b*) When the material undergoes a phase transition at temperature T_{trs}, a dramatic increase in *Q* occurs; its jump corresponds to the value of the latent heat of the transition $\Delta_{trs}H$ as indicated on the diagram. This large increase in *Q* can be used to advantage in phase-change materials for heat storage.

A phase change can lead to a much larger quantity of energy stored, compared with sensible storage alone. The comparison for water is quite useful. Pure water has a heat capacity of 4.2 J K^{-1} g^{-1}, so for a 1°C temperature rise, 1 g of water can store 4.2 J. However, the latent heat associated with melting of ice is 330 J g^{-1}. So taking 1 g of ice from just below its melting point to just above (with a total temperature difference of 1°C) absorbs 334 J (latent heat plus 4.2 J from sensible heat storage), about 80 times as much as the sensible heat storage capacity alone.

Although phase-change materials can be used to provide more energy storage, they have their limitations. The most critical limitation is that, in order to be most useful, they must have their phase transition in the temperature range at which the system will operate. For example, at ordinary pressures water has only two phase transitions, one associated with melting at 0°C (32°F) and one with boiling at 100°C (212°F). If a heat storage system were required to operate in the temperature range of 20–40°C (68–104°F), water could operate only as a sensible heat storage material, not as a phase-change material.

Glauber salts are commonly used phase-change materials. They are inorganic salt hydrates that undergo melting-phase transitions that absorb heat. While these materials are highly efficient at storing heat, their maximum storage capacity, again, is over a limited range (transition temperature from 30 to 48°C or 86 to 118°F). In addition, there are problems associated with phase separation on melting, and there is a need for supercooling in order to recover all the stored heat. While these materials are inexpensive, their corrosive nature can lead to high storage container costs. However, the energy storage capacity of Glauber salts and related materials is high, and they are used in commercial products.

Paraffin waxes are also used as phase-change materials. When a paraffin wax melts, it absorbs heat and stores energy. However, there are drawbacks associated with using paraffin waxes as heat storage materials such as oxidation, volatility, flammability, and the large volume change on melting (about 20% increase). Once again, containment costs can be high. Paraffins are often microencapsulated in commercial phase-change materials applications. *See* PARAFFIN.

Pentaerythritol [$C(CH_2OH)_4$] and its derivatives absorb energy on going through phase transformations, but its phases (before and after the transition) are both solid. The attractive feature of using a solid-solid phase transition is that the mechanical properties of the material are much the same before and after the transition. In particular, the flow problems associated with melting transitions are overcome. The major problem with pentaerythritol is its relatively high transition temperature (188°C or 370°F), which severely limits its commercial use. The transition temperature can be lowered through chemical modification, but this also lowers the quantity

of heat absorbed in the transition. Nevertheless, development of new solid-solid phase transition phase-change materials offer considerable promise for heat storage systems.

Applications. Broadly speaking, applications of heat storage systems fall into two categories: scavenging of heat, and thermal control. Both rely on the principle that as energy is put into the system its temperature rises. Later, the energy is released as the temperature falls. In both cases, the goal is to achieve the maximum energy stored (and later released) for a given temperature change.

Scavenging of heat. These uses range from district heating or cooling to solar energy applications.

In district heating, heat storage systems are used to scavenge energy from what might otherwise be wasted sources, such as smokestacks associated with heavy industry. By appropriate heat exchange between the waste energy source and a fluid, this thermal energy can be moved to another location such as a steam heating system for housing and office buildings.

District cooling makes use of heat storage systems in a somewhat different way. For example, in Minato Mirai 21, a major commercial and residential development in Yokohama, Japan, an underground storage system makes use of off-peak nighttime electricity to cool phase-change materials to their low-temperature state. In the daytime, heat-exchange fluids carry the "cold" to the buildings in this district to provide air conditioning. When the fluids return to the phase-change materials system, their heat warms the material. The high energy density of the phase-change materials allows a large cooling capacity, while leveling out the diurnal electricity consumption.

In solar energy applications, the radiant energy of the Sun is used to increase the energy content of the heat storage system in the daytime. This energy is released from the system as the temperature decreases during the night, either through direct radiation or through exchange of energy with circulating fluids. As with the other heat-scavenging applications, the energy density is greatest if phase-change materials, optimized to the temperature cycle of the system, are employed. *See* SOLAR ENERGY; SOLAR HEATING AND COOLING.

Another example of a scavenging heat storage system is found in some automobiles. The heat storage system uses the engine's waste heat to store thermal energy that is later used to rapidly heat the air for the passenger compartment after a cold start.

Thermal control. Because a heat storage system aims to have a large energy storage for a small temperature change, these systems also have applications in thermal control.

There are now commercially available ski boots that make use of heat storage systems. Phase- change materials in the insole prevent major temperature drops. Any loss of energy because of the cold conditions is tempered by the presence of the material, which requires a large loss of energy (latent heat) before it cools into its low-temperature phase and then cools further. In a similar way, phase-change materials in insulated storage bags are used to keep pizza hot during delivery.

Preventing temperature extremes can be very important in applications involving biological materials or electronic devices. Here again, heat storage systems can be used to "buffer" against major temperature changes in the environment. If necessary, two phase-change materials can be used, with one providing a lower-temperature buffer and the other an upper-temperature buffer. If the temperature of the surroundings is too extreme, or out of range for too long, the storage capacity of the system can be overcome and the temperature can go outside the desired range.

Outlook. The convenience or actual savings associated with a heat storage system must outweigh its capital cost. One way to improve this situation is by the use of more efficient systems. Efficiency has two components: the energy recovered (which should be as high as possible) and the quality of the energy transferred (the fraction of the energy that is available for conversion to useful work, also called the exergy). New materials, such as high-efficiency phase-change materials with a wide or tunable range of phase-transition temperatures, would be very useful. New designs of the heat storage systems can further optimize efficiency. With further realization of the limitations of energy supplies, heat storage materials are likely to play a larger role in consumer and industrial energy distribution systems in the near future. *See* ENERGY STORAGE.

Mary Anne White

Bibliography. S. M. Hasnain, Review on sustainable thermal energy storage techniques, Part II: Cool thermal storage, *Energy Convers. Manag.*, 39:1139–1153, 1998; F. R. Kalhammer, Energy-storage systems, *Sci. Amer.*, 241(6):56–65, 1979; C. J. Weinberg and R. H. Williams, Energy from the Sun, *Sci. Amer.*, 263(3):146–155, 1990; M. A. White, *Properties of Materials*, Oxford University Press, New York, 1999.

Heat transfer

Heat, a form of kinetic energy, is transferred in three ways: conduction, convection, and radiation. Heat can be transferred only if a temperature difference exists, and then only in the direction of decreasing temperature. Beyond this, the mechanisms and laws governing each of these ways are quite different. This article gives introductory information on the three types of heat transfer (also called thermal transfer) and on important industrial devices called heat exchangers.

Conduction. Heat conduction involves the transfer of heat from one molecule to an adjacent one as an inelastic impact in the case of fluids, as oscillations in solid nonconductors of electricity, and as motions of electrons in conducting solids such as metals. Heat flows by conduction from the soldering iron to the work, through the brick wall of a furnace, through the wall of a house, or through the wall of a cooking

utensil. Conduction is the only mechanism for the transfer of heat through an opaque solid. Some heat may be transferred through transparent solids, such as glass, quartz, and certain plastics, by radiation. In fluids, the conduction is supplemented by convection, and if the fluid is transparent, by radiation.

The conductivities of materials vary widely, being greatest for metals, less for nonmetals, still less for liquids, and least for gases. Any material which has a low conductivity may be considered to be an insulator. Solids which have a large conductivity may be used as insulators if they are distributed in the form of granules or powder, as fibers, or as a foam. This increases the length of path for heat flow and at the same time reduces the effective cross-sectional area, both of which decrease the heat flow. Mineral wool, glass fiber, diatomaceous earth, glass foam, Styrofoam, corkboard, Celotex, and magnesia are all examples of such materials. *See* CONDUCTION (HEAT).

Convection. Heat convection involves the transfer of heat by the mixing of molecules of a fluid with the body of the fluid after they have either gained or lost heat by intimate contact with a hot or cold surface. The transfer of heat at the hot or cold surface is by conduction. For this reason, heat transfer by convection cannot occur without conduction. The motion of the fluid to bring about mixing may be entirely due to differences in density resulting from temperature differences, as in natural convection, or it may be brought about by mechanical means, as in forced convection.

Most of the heat supplied to a room from a steam or hot-water radiator is transferred by convection. In fact, the heat from the fire in the furnace heating the hot water or steam is transferred to the boiler wall by convection, and the hot water or steam transfers heat from the boiler to the radiator by convection. Iced tea is cooled and soup heated by convection. *See* CONVECTION (HEAT).

Radiation. Solid material, regardless of temperature, emits radiations in all directions. These radiations may be, to varying degrees, absorbed, reflected, or transmitted. The net energy that is transferred by radiation is equal to the difference between the radiations emitted and those absorbed.

The radiations from solids form a continuous spectrum of considerable width, increasing in intensity from a minimum at a short wavelength through a maximum and then decreasing to a minimum at a long wavelength. As the temperature of the object is increased, the entire emitted spectrum decreases in wavelength. As the temperature of an iron bar, for example, is raised to about 1000°F (800 K), the radiations become visible as a dark red glow. As the temperature is increased further, the intensity of the radiation increases and the color becomes more blue. This process is quite apparent in the filament of a light bulb. When the bulb is operated at less than normal voltage, the light appears quite red. As the voltage is increased, the filament temperature increases and the light progressively appears more blue.

Liquids and gases only partially absorb or emit these radiations, and do so in a selective fashion. Many liquids, especially organic liquids, have selective absorption bands in the infrared and ultraviolet regions. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION.

Transfer of energy by radiation is unique in that no conducting substance is necessary, as with conduction and convection. It is this unique property that makes possible the transfer of large amounts of energy from the Sun to the Earth, or the transfer of heat from a radiant heater in the home. It is the ready transfer of heat by radiation from a California orange grove to outer space on a clear night that sometimes results in a frost. The presence of a shield of clouds will tend to prevent this loss of heat and often prevent the frost. By means of heat lamps and gold-plated reflectors, heat may be transferred deep into the layer of enamel on a car body, with resultant hardening of the enamel from the inside out. It is also the transfer over great distance of quantities of radiant energy that makes the atomic bomb so destructive. *See* HEAT RADIATION.

Design considerations. By utilizing a knoweledge of the principles governing the three methods of heat transfer and by a proper selection and fabrication of materials, the designer attempts to obtain the heat flow required for his purposes. This may involve the flow of large amounts of heat to some point in a process or the reduction in flow in others. It is possible to employ all three methods of heat transfer in one process. In fact, all three methods operate in processes that are commonplace. In summer, the roof on a house becomes quite hot because of radiation from the Sun, even though the wind is carrying some of the heat away by convection. Conduction carries the heat through the roof where it is distributed to the attic by convection. The prudent householder attempts to reduce the heat that enters the rooms beneath by reducing the heat that is absorbed in the roof by painting the roof white. He may apply insulation to the underside of the roof to reduce the flow of heat through the roof. Further, heated air in the attic may be vented through louvers in the roof.

Heat transferred by convection may be transferred as heat of the convecting fluid or, if a phase change is involved, as latent heat of vaporization, solidification, sublimation, or crystallization. The human body can be cooled to less than ambient temperature by evaporation of sweat from the skin. Dry ice absorbs heat by sublimating the carbon dioxide. Heat extracted from the products of combustion in the boiler flows through the gas film and the metal tube wall and converts the water inside the tube to steam, all without greatly changing the temperature of the water.

Heat exchangers. In industry it is generally desired to extract heat from one fluid stream and add it to another. Devices used for this purpose have passages for each of the two streams separated by a heat-exchange surface in the form of plates or tubes and are known as heat exchangers. Needless to say, the

automobile radiator, the hot-water heater, the steam or hot-water radiator in a house, the steam boiler, the condenser and evaporator on either the household refrigerator or air conditioner, and even the ordinary cooking utensils in everyday use are all heat exchangers. In power plants, oil refineries, and chemical plants, two commonly used heat exchangers are the tube-and-shell and the double-pipe exchangers. The first consists of a bundle of tubes inside a cylindrical shell. One fluid flows inside the tubes and the other between the tubes and the shell. The double-pipe type consists of one tube inside another, one fluid flowing inside the inner tube and the other flowing in the annular space between tubes. In both cases, the tube walls serve as the heat-exchange surface. Heat exchangers consisting of spaced flat plates with the hot and cold fluids flowing between alternate plates are also in use. Each of these exchangers essentially depends upon convection heat flow through a film on each side of the heat-exchange surface and conduction through the surface. Countless special modifications, often also utilizing radiation for heat transfer, are in use in industry.

In these exchangers, the fluid streams may flow parallel concurrently or in mixed flow. In most cases, the temperatures of the various streams remain essentially constant at a given point, and the process is said to be a steady-state process. As the streams move through the exchangers, unless there is a phase change, the fluids are continuously changing in temperature, and the temperature gradient from one stream to the other may be continuously varying. To determine the amount of surface needed for a given process, the designer must evaluate the effective temperature gradient for the particular condition and exchanger.

With extremely high temperatures, or with gas streams carrying suspended solids, the use of conventional heat exchangers becomes impractical. Under these conditions, the transfer of heat from one stream to another becomes more economical by the alternate heating and cooling of refractory solids or by checkerwork as in the blast-furnace hot stove, in the glass-furnace regenerator, or in the Royster stove. At lower temperatures, metal packing is frequently employed, as in the Ljungstrom preheaters or in regenerators for liquid-air production. In petroleum refining and in the metallurgical industry, exchangers are being employed in which one or more of the streams are fluidized beds of solids, the large area of the solids tending to produce very high rates of heat exchange. In some of these devices and also in nuclear power reactors, large quantities of heat are being generated in the exchangers. Here one of the principal problems involves the rapid removal of this heat before the temperature rises to the point where the equipment is damaged or destroyed.

Often the heating or cooling of a body is desired. In this case, the body representing the second stream does not remain at constant temperature, the heat being transferred representing a change in the heat content of the body. Such a process is known as an unsteady-state process. The heating or cooling of food and canned products in utensils, refrigerators, and sterilizers; the heating of steel billets in metallurgical furnaces; the burning of brick in a kiln; and the calcination of gypsum are examples of this type of process. *See* HEAT; HEAT EXCHANGER.

Ralph H. Luebbers

Bibliography. Y. Bayazitoglu and M. N. Ozisik, *Elements of Heat Transfer,* 1988; M. Becker, *Heat Transfer: A Modern Approach,* 1986; A. Bejan and J. S. Jones, *Modern Heat Transfer*, 1993; A. J. Chapman, *Fundamentals of Heat Transfer,* 1987; J. P. Holman, *Heat Transfer*, 8th ed., 1997; F. Kreith and M. Bohn, *Principles of Heat Transfer*, 6th ed., 2000.

Heat treatment (metallurgy)

A procedure of heating and cooling a material without melting. The heating and cooling sequence may involve temperatures above, below, and at the ambient. Controlled heating and cooling rates, and a variety of furnace atmospheres and heating media may be used. Plastic deformation may be included in the sequence of heating and cooling steps, thus defining a thermomechanical treatment. Typical objectives of heat treatments are hardening, strengthening, softening, improved formability, improved machinability, stress relief, and improved dimensional stability. Heat treatments are often categorized with special names, such as annealing, normalizing, stress relief anneals, process anneals, hardening, tempering, austempering, martempering, intercritical annealing, carburizing, nitriding, solution anneal, aging, precipitation hardening, and thermomechanical treatment.

Steels and Other Ferrous Alloys

All metals and alloys in common use are heat-treated at some stage during processing. Iron alloys, however, respond to heat treatments in a unique way because of the multitude of phase changes which can be induced, and it is thus convenient to discuss heat treatments for ferrous and nonferrous metals separately. *See* IRON; STEEL.

General principles. At room temperature the equilibrium crystal structure of pure iron is body-centered cubic α-iron (**Fig. 1***a*), also known as

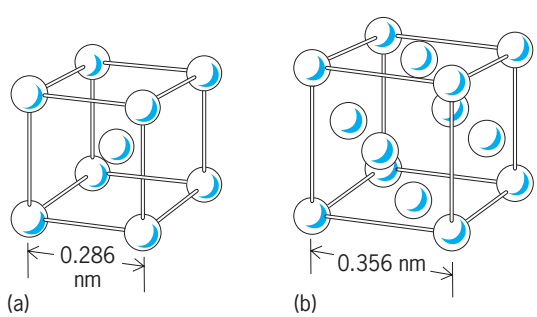

Fig. 1. Crystal structures of iron. (*a*) Body-centered cubic α-iron. (*b*) Face-centered cubic γ-iron.

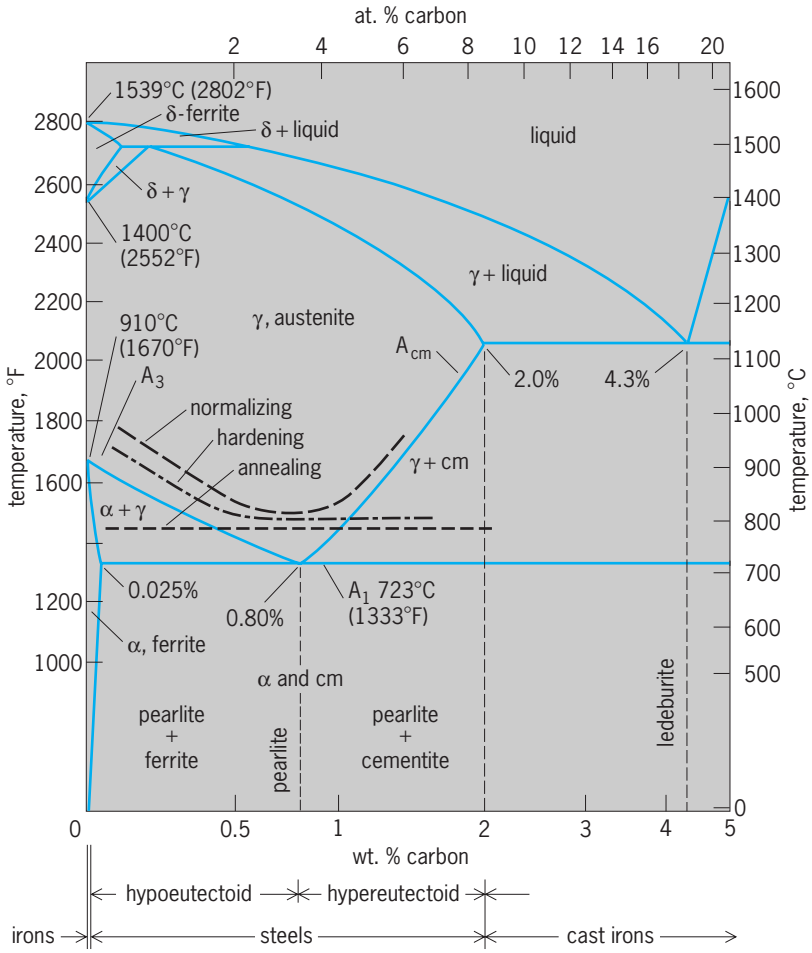

Fig. 2. Iron-carbon phase diagram.

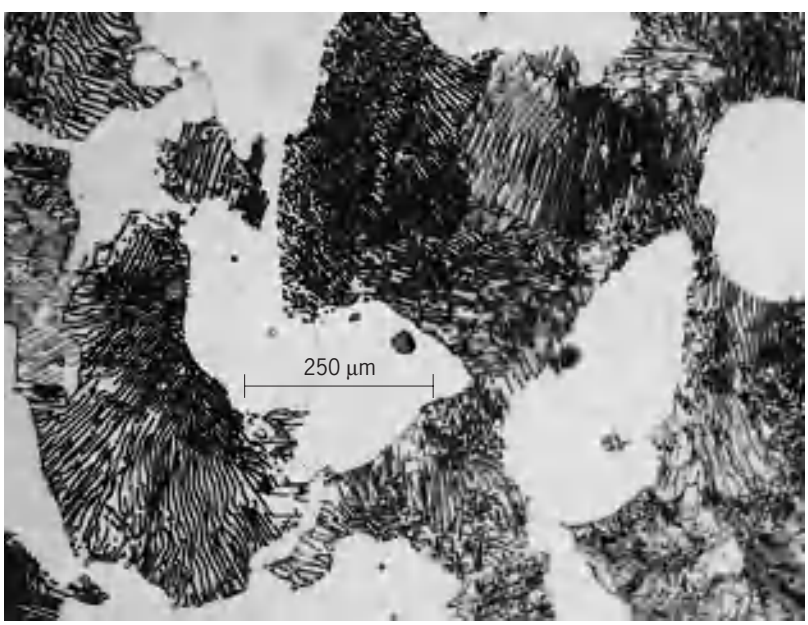

Fig. 3. Microstructure of annealed 0.40% carbon steel, after polishing and etching. White regions are ferrite; lamellar regions, pearlite.

ferrite. On heating above 1670°F (910°C), α-iron is transformed to face-centered cubic γ-iron (Fig. 1*b*), also called austenite. At 2552°F (1400°C) γ-iron transforms to δ-ferrite, which is also body-centered cubic and structurally similar to α-iron, but δ-ferrite is seldom involved in heat-treating procedures. The addition of carbon to iron influences the transformation from one form of iron to another, and the resultant structures are summarized in the iron-carbon phase diagram shown in **Fig. 2**. Practical steels and irons contain other elements such as manganese, silicon, sulfur, phosphorus, aluminum, silicon, chromium, molybdenum, and nickel. These alloy elements influence the shape of the iron-carbon diagram, but if the total alloy content is less than 2%, the phase diagram is not affected substantially.

The general principles of the heat treatment of plain-carbon and low-alloy steels may be understood from the basic iron-carbon diagram. The diagram indicates the microconstituents, or phases, which are observed for a given carbon content at each temperature. In addition to ferrite and austenite, the other principal phase is the intermetallic compound, Fe_3C, or cementite. Cementite is, however, not the most stable form of carbon in iron, and at true equilibrium, graphite is formed after prolonged heating. Graphite is a constituent of the common grades of cast iron, but these irons usually contain 2–4% silicon to promote graphitization. From a practical standpoint, cementite is, however, stable in most alloys with less than 2% carbon. Approximate ranges of the carbon contents for irons, steels, and cast iron are shown in Fig. 2.

Steels which contain 0.80% carbon are called eutectoid steels, those with less carbon are hypoeutectoid, and those with more are hypereutectoid. The changes which occur on slowly heating a typical hypoeutectoid steel containing 0.40% carbon are the following. At room temperature such a steel contains ferrite and cementite. On heating above the A_1 critical temperature, 1333°F (723°C), the cementite dissolves, austenite is formed, and the structure consists of austenite and ferrite. Above the A_3 temperature, which is 1472°F (800°C), for the alloy, the structure is entirely austenitic. Melting is observed at 2640°F (1450°C), and in practice, heat-treating and hot-forming temperatures are below 2200°F (1200°C). On cooling, the sequence of phase changes is reversed, but the shapes and distribution of the phases depend on the cooling rate. If the cooling is very slow, as in furnace cooling, the carbides will tend to have a spheroidal shape. If the cooling rate is more rapid, as in air cooling, the cementite will form a duplex lamellar structure with the ferrite, called pearlite, and the microstructure will consist of 50% ferrite grains and 50% pearlite colonies as shown in **Fig. 3**. A eutectoid steel, containing 0.80% carbon, will consist entirely of pearlite on air-cooling from above A_1. Hypereutectoid steels will consist of pearlite colonies and excess cementite. It should be emphasized that pearlite is not a phase, and it is shown on the phase diagram only for convenience.

Annealing heat treatments. Annealing heat treatments are used to soften the steel, to improve the machinability, to relieve internal stresses, to impart dimensional stability, and to refine the grain size. Several typical annealing treatments are discussed below.

Anneal or grain-refining anneal. Hypo- and hypereutectoid steels are heated just above A_1, and cooled moderately slowly to room temperature. Recrystallization of the ferrite grains will occur during the heating, and the material will be softened by the stress relief and the removal of crystallographic imperfections at temperature and during the slow cooling. Anneals of this type are used to improve the machinability of hypereutectoid steels since the blocky pearlite which forms (Fig. 3) aids in the breakup of the chips during machining.

Spheroidizing anneal. In this treatment, steels are heated just above A_1 and cooled very slowly through the critical temperature, using a programmed reduction of the temperature over a period of 10–15 h to 1200°F (650°C), and then cooling more rapidly to room temperature. In other procedures, the temperature is cycled slowly from just above to just below the critical, or the steel may be cooled to just below the critical and held for a long period. The latter are called isothermal anneals. The objective of all of these treatments is to produce spheroidized carbides. Such structures provide superior cold formability in hypoeutectoid steels and good machinability in hypereutectoid steels.

Stress-relief anneal. This treatment consists of heating below A_1 and cooling slowly. The objective is to relieve residual stresses, such as those introduced by cold forming and machining. Phase transformations are avoided in order to minimize distortion, and this treatment is used frequently to improve the dimensional stability of a complex part.

Normalizing. This treatment involves heating above A_3 or A_{cm} (cm indicates cementite) and air-cooling to room temperature. It is used for hypereutectoid steels to remove carbide networks which may have formed in previous processing (such as carburizing), and for hypoeutectoid steels to produce a material of intermediate strength with high ductility and low residual stresses. Typically such a treatment is specified for large shafts, made of hypoeutectoid steels, after forging. Normalizing is used to improve the machinability of hypoeutectoid steels, because the resultant microstructure contains pearlite, although the pearlite colonies will not be as well developed as those in a steel which has been given a full anneal.

In-process anneal. This is an annealing treatment introduced at the steel mill during the processing of steel bar, wire, sheet, tube, or other product. These may be spheroidizing, softening, normalizing, or stress relief anneals, depending on the ultimate use.

Hardening treatments. Steels are hardened by heating to a temperature at which austenite is formed and then cooling with sufficient rapidity to make the transformation to pearlite and/or ferrite unfavorable.

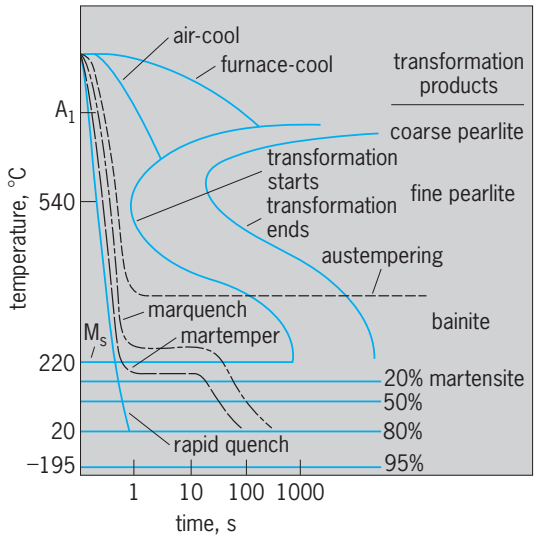

Fig. 4. Schematic isothermal transformation diagram for a eutectoid steel. °F = (°C × 1.8) + 32.

The phase diagram is useful in determining suitable austenitizing, or hardening, temperatures, but the isothermal transformation curve, also called a TTT (time, temperature, transformation) curve, is more helpful in assessing the hardenability of a steel. A typical isothermal transformation diagram for a eutectoid steel austenitized at 1500°F (815°C) is shown in **Fig. 4**. This diagram is obtained experimentally for each combination of steel and austenitizing temperature by austenitizing small pieces, cooling rapidly to a holding bath, typically a lead or salt bath, and quenching into water after various times at the holding temperature. The constituents which form isothermally are determined by metallographic methods and are indicated on the diagram.

The superposition of typical cooling rates used in heat treatments onto the isothermal diagrams will indicate the constituents which will form initially on cooling. A continuous cooling transformation diagram is used to record all of the phases which form at various cooling rates, but such diagrams are difficult to determine and the isothermal curves are frequently used for this purpose. For example, furnace and air-cooling curves intersect the transformation curve above the nose of the C curve (Fig. 4), and the resultant microstructure will be pearlite. A rapid quench, in water or in some oils, will miss the pearlite nose, and martensite will form.

Martensite has a body-centered tetragonal crystal structure (**Fig. 5**) with the *c* and *a* dimensions dependent on the carbon content. Martensite can be formed only from austenite and contains all of the carbon and alloy elements which are dissolved in the austenite. The hardening of steel occurs by the formation of martensite, and the hardness of the martensite increases with the carbon content, reaching a plateau at 0.80% C.

Martensite begins to form at the M_s temperature, which is 520°F (270°C) in the example shown in Fig. 4. The percentage of martensite is shown as a series of horizontal lines on the transformation diagram

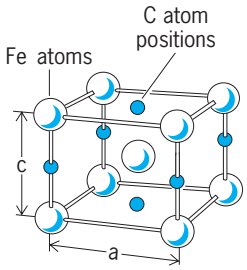

weight percent C	c (nm)	a (nm)
0	0.286	0.286
0.20	0.288	0.2858
0.40	0.291	0.2856
0.80	0.295	0.285

Fig. 5. Crystal structure of martensite.

since very little martensite formation occurs isothermally. At room temperature, such a steel will contain about 80% martensite by volume, with the remainder retained austenite. Some of the retained austenite may be transformed by cooling below room temperature, but the austenite cannot be eliminated solely by refrigeration, even to liquid helium temperatures. The M_s temperature is lowered, and the retained austenite contents increased, by increasing alloy and carbon contents of the austenite.

The position of the nose of the C-curve is dependent on the alloy and carbon contents, and this greatly influences the cooling rates required to harden a steel. In plain-carbon steels the nose is far to the left, and water or rapid oil quenches must be used. For medium-alloy steels (for example AISI 4340, which contains 0.40 C, 1.8 Ni, 0.8 Cr, 0.25 Mo) moderately thick sections may be hardened by oil quenching. For high-alloy steels (for example, type A-2 tool steel, which contains 1.0 C, 5.0 Cr, 1.0 Mo, 0.25 V) air cooling may be sufficient. The hardenability of steel is greater the farther to the right the position of the nose of the C-curve. Hardness should be distinguished from hardenability. Hardness refers to the resistance to indentation. Hardenability refers to the severity of cooling which must be used to produce martensite in a steel. If the steel can be hardened by using a slow cooling rate, such as air cooling, it has high hardenability. If a very rapid cooling rate, such as water quenching, must be used, it has low hardenability.

Freshly quenched martensite containing more than 0.10% C cannot be deformed easily, and a subsequent heating, or tempering, is used to increase the ductility. Microstructural changes occur on tempering, and the sequence of structures in low- and medium-alloy steels follows three stages. In the first, ϵ-carbide, which is hexagonal with the approximate composition Fe_2C, precipitates at temperatures up to 450°F (220°C). In the second stage, 410–590°F (210–310°C), the retained austenite decomposes to bainite, and in the third stage, 590–1000°F (210–540°C), ferrite and cementite are formed. Typically, cutting tools and bearings made of low-alloy steels are tempered through the first and second stages, and high-strength structural steels, such as AISI 4340, are tempered into the third stage in order to obtain sufficient ductility. The tempering sequence is somewhat different from high-alloy tool steels, such as type A-2 or the high-speed tool steels. In these steels, high-alloy carbides precipitate on tempering in the range 900–1000°F (490–550°C), lowering the M_s temperature of the retained austenite. The retained austenite transforms to martensite on cooling, and these steels thus undergo a secondary hardening on tempering in this range.

Austempering. Several variations in the quenching procedure are used to achieve special properties. Austempering involves quenching rapidly to a temperature below the pearlite nose, but above the M_s, as shown in Fig. 4. The objective is the formation of bainite, which is similar in form and has about the hardness of martensite tempered at the same temperature. Bainite, however, has better impact toughness and lower residual stresses than the corresponding martensite.

Marquenching and martempering. Marquenching involves quenching to a temperature just above the M_s and then air-cooling to room temperature. In martempering, the procedure is similar, but the interrupting temperature is just below M_s. The steels are then tempered in the normal fashion. These treatments are used to minimize the distortion which occurs in the direct quench to room temperature.

HSLA and dual-phase steels. High-strength low-alloy (HSLA) steels usually have a low carbon content (typically 0.02–0.10%) and sufficient content of alloy elements such as manganese, silicon, molybdenum, titanium, columbium, chromium, and vanadium to provide enough hardenability for the formation of low-carbon martensite in thin sheets. Such steels may harden on air-cooling from the austenitizing temperatures, although similar results may be obtained by water-quenching thin sheets of plain-carbon steel. Dual-phase (or duplex) steels are HSLA steels for which a particularly high degree of formability has been achieved by a special heat treatment. These steels are heat-treated by austenitizing in the $\alpha + \gamma$ phase region (this is often called an intercritical anneal) and cooling rapidly to room temperature. The temperature is chosen to result in a microstructure with approximately 80% ferrite and 20% martensite-plus-retained-austenite and this provides a combination of good formability and high strength in the finished part.

Surface treatments. Some heat treatments are used to alter the chemistry at the surface of a steel, usually to achieve preferential hardening of a surface layer. *See* SURFACE HARDENING OF STEEL.

Carburizing. Most carburizing is carried out in an atmosphere of partially combusted natural gas which is further enriched in carbon by additions of natural gas, or some other carbonaceous gas. The starting material is a low-carbon steel, and carbon diffuses into the outer layers, thus producing a duplex structure with a hypereutectoid steel in the case and a hypoeutectoid steel in the core. A hardening treatment may then be used to produce a hard case and a soft core.

Nitriding. Steels, usually containing aluminum, molybdenum, titanium, or chromium, are heated at 1000°F (540°C) in an atmosphere containing undissociated ammonia. Nitrogen diffuses into the

surface, forming nitrides and causing a substantial increase in hardness of the surface layer. Additional hardening treatments are not required after the nitriding.

Chromizing. Chromium may be added to the surface by diffusion from a chromium-rich material packed around the steel or dissolved in molten lead. This process is quite slow and must be carried out at an elevated temperature. It is used to improve the corrosion resistance of the surface.

Nonferrous Metals and Alloys

Many nonferrous metals do not exhibit phase transformations, and it is not possible to harden them by means of simple heating and quenching treatments as in steel. Unlike steels, it is impossible to achieve grain refinement by heat treatment alone, but it is possible to reduce the grain size by a combination of cold-working and annealing treatments. Some nonferrous alloys can be hardened, but the mechanism is one by which a fine precipitate is formed, and the reaction is fundamentally different from the martensitic hardening reaction in steel. There are also certain ferrous alloys that can be precipitation hardened. However this hardening technique is used much more widely in nonferrous than in ferrous alloys. In titanium alloys, the β phase can transform in a martensitic reaction on rapid cooling, and the hardening of these alloys is achieved by methods which are similar to those used for steels.

Annealing treatments. This process involves heating the metal or alloy to an elevated temperature, particularly after cold-working operations such as wire drawing, deep drawing, cold forming, cold extrusion, or heavy machining, and cooling slowly to room temperature. The range of annealing temperatures for copper-base alloys is 660–1560°F (350–850°C), for aluminum alloys 560–800°F (290–430°C), and for nickel-base alloys 1470–2190°F (820–1200°C). The metals are softened by a recrystallization mechanism, which involves the nucleation of a new set of strain-free grains, and the subsequent growth of these strain-free grains at the expense of the cold-worked grains. The temperature at which recrystallization starts is not precisely defined and depends primarily on the amount of cold work, that is, deformation, the material experienced. Recovery from cold work can also be achieved without the formation of new grains by prolonged heating below the recrystallization temperature. Recovery treatments are used in instances where stress relieving is required for dimensional stability but full softening is undesirable because of the accompanying reduction in strength.

Precipitation and age hardening. Certain nonferrous alloys can be hardened by a precipitation-hardening treatment, sometimes called age hardening. In order to be amenable to such a treatment, the solubility of a phase or a compound in the alloy must be such as to be completely dissolved at a high temperature and supersaturated at a lower temperature. For example, one of the well-known precipitation-hardening alloys is an aluminum alloy containing 4% copper, 0.5% manganese, and 0.5% magnesium. It is solution-treated by heating at about 950°F (510°C) and quenching into water. At the solution temperature all of the copper is in solid solution, and the solid solution is retained on quenching to room temperature. The alloy is soft after this treatment (or solution anneal) and may be readily machined or deformed. If the alloy is subsequently aged at room temperature, the strength and hardness gradually increase, reaching a maximum in 5 or 6 days. The aging may be accomplished more rapidly by heating for several hours at 400°F (175°C).

During the age-hardening period, copper-rich regions form in the aluminum lattice and a precipitate which is coherent with the aluminum matrix is gradually developed. The coherency refers to a matching of the lattice spacings in the copper-rich cluster with those in the aluminum, and the resultant local strains produce the hardening of the alloy. If the alloy is overaged, by heating above the precipitation-hardening temperature, the equilibrium precipitate, $CuAl_2$, which is not coherent, will be formed and the hardness will be reduced. Overaging is utilized where dimensional stability is favored over achieving the maximum strength.

Some precipitation-hardening alloys achieve maximum strength only after a cold-working operation is introduced following the solution treatment and prior to the precipitation treatment. These treatments are usually carried out at the mill, and the material is purchased in the hardened condition.

A variety of heat treatments and thermomechanical treatments are used for aluminum alloys, and these are designated by letters and digits following the alloy designation. Some common designations are the following:

F = fabricated
O = annealed
H = strain-hardened
W = solution heat-treated
T1 = naturally aged only
T2 = annealed (case products only)
T3 = solution heat-treated, cold-worked, and naturally aged
T4 = solution heat-treated and naturally aged
T5 = artificially aged only}
T6 = solution heat-treated and artificially aged
T7 = solution heat-treated and stabilized (over aged)
T8 = solution heat-treated, cold-worked, and artificially aged
T9 = solution heat-treated, artificially aged, and cold-worked
T10 = artificially aged and cold-worked

See ALLOY; METAL, MECHANICAL PROPERTIES OF.

B. L. Averbach

Bibliography. American Society for Metals, *Heat Treating Source Book*, 1986; American Society for Metals, *Metals Handbook*, 9th ed., vol. 4: *Heat Treating*, 1981; American Society for Metals, *Steels: Heat Treatment and Processing Principles*, 1989; O. F. Devereux, *Topics in Metallurgical Thermodynam-*

ics, 1989; G. Krauss, *Principles of Heat Treatment of Steel*, 1980, reprint 2000; P. H. Morton (ed.), *Surface Engineering and Heat Treatment*, 1992.

Heating system

An apparatus consisting of an energy source, a method of converting that energy to heat, and a transport system to convey the energy and heat to the point of use. Most heating systems include some manual or automatic method of controlling the heat output and delivery.

Energy sources. There are many sources of energy for use in heating. The earliest source, and still most common in underdeveloped countries, comprises wood and wood products, such as paper, wood chips, and sawdust; peat is used in some cultures. Solar use for heating and electrical generation has become widespread, although economics discourages more general use. The generation of electrical energy requires the use of fossil fuels, water power, geothermal energy, or nuclear energy. Fossil fuels are used directly in furnaces and boilers. Waste heat from various processes is often used to serve other parts of a common facility. For example, the methane gas generated by sewage disposal plants is sometimes used in special boilers to speed up the sewage reaction process. The heat pump is an adaptation of the refrigeration cycle in which the rejected heat is put to beneficial use. *See* ELECTRIC POWER SYSTEMS; ENERGY SOURCES; HEAT PUMP; SOLAR ENERGY.

Energy conversion. The energy source is converted into heat by various means. Wood and fossil fuels are converted by burning, that is, the combustion process. All combustion processes generate substances that are regarded as pollutants, and are more or less inefficient. Thus, there are frequent attempts to use waste heat and maximize combustion efficiencies. Modern standards and codes require that pollutants be removed to the greatest extent possible. This requirement adds to costs and sometimes has an adverse effect on efficiency. Electrical energy can be converted directly into heat by means of resistance heaters. Heat pump systems typically use electrical energy to drive the refrigeration compressor and fans; part of the useful heat is waste heat from the refrigeration process, but most of the heat comes from the air or water source. Solar energy requires collectors, which convert it into heat or electricity.

Transport systems. Heating systems use many methods to deliver heat to the point of use. Radiation systems take several forms. Cast-iron column radiators, using steam or hot water, have largely been superseded by convector radiators using steam, hot water, or electricity. Panel-type radiators are also used in ceilings or in floors. All require a piping or electrical distribution system. Forced-air warm-air heating, using electric motor–driven circulating fans, is very common. Residential warm-air heating systems use warm air with heat supplied by a central furnace burning fossil fuels or by a heat pump. Commercial, institutional, and industrial systems usually use central boilers for heat generation, with distribution through piping systems to heat exchangers at forced-air handling units, which deliver heated air to the point of use. In most cases the air systems also provide cooling for year-round air conditioning. In some cities, district heating systems supply heat through piping systems from a central plant to many users. *See* COMFORT HEATING; DISTRICT HEATING; HOT-WATER HEATING SYSTEM; PANEL HEATING AND COOLING; RADIANT HEATING; SOLAR HEATING AND COOLING; STEAM HEATING; WARM-AIR HEATING SYSTEM.

Roger W. Haines

Bibliography. American Society of Heating, Refrigerating and Air Conditioning Engineers, *ASHRAE Handbooks: HVAC Systems and Equipment*, 1992, *Fundamentals*, 1993.

Heavy minerals

Minerals with a density greater than 2.9 g/cm^3. The term is most commonly used to denote high-density components of siliciclastic sediments. Most heavy mineral studies are undertaken to determine sediment provenance, because heavy mineral suites provide important information on the mineralogical composition of source areas (see **table**). Since heavy minerals rarely constitute more than 1% of sandstones, their study normally requires them to be concentrated. This is achieved by disaggregation of the sandstone, followed by mineral separation using dense liquids such as bromoform, tetrabromoethane, or the more recently developed nontoxic polytungstate liquids. *See* DENSITY; PROVENANCE (GEOLOGY); SANDSTONE.

Geographic and stratigraphic variations in heavy mineral suites within a sedimentary basin can be used to infer differences in sediment provenance. Such differences result either from the interplay between a number of sediment transport systems draining different source regions, or from erosional unroofing within a single source area. Heavy mineral data therefore play an important role in the understanding of depositional history and paleogeography. In some cases, sophisticated mathematical and statistical treatment of heavy mineral data may be required to elucidate the interplay between multiple sediment transport systems. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; PALEOGEOGRAPHY; SEDIMENTOLOGY; STRATIGRAPHY.

The recognition of changes in provenance within a sedimentary sequence provides a basis for correlation of strata that is independent of more traditional biostratigraphic methods. Heavy mineral assemblages have proven to be especially useful for correlating sandstones that lack age-diagnostic fossils, especially nonmarine sequences. *See* SEQUENCE STRATIGRAPHY.

The identification of provenance on the basis of heavy mineral data requires careful consideration of all factors influencing the composition of the assemblages. A number of processes may alter the original

Provenance of the most commonly recorded nonopaque detrital heavy mineral species*

Mineral	Density, g/cm^3	Hardness (Mohs scale)	Stability in acidic weathering	Stability in burial diagenesis	Most common provenance
Andalusite	3.13–3.16	$6\frac{1}{2}$–$7\frac{1}{2}$	High	Low	Metapelites
Amphibole	3.02–3.50	5–6	Low	Low	Various igneous and metamorphic rocks
Apatite	3.10–3.35	5	Low	High	Various igneous and metamorphic rocks
Chloritoid	3.51–3.80	$6\frac{1}{2}$	Moderate	Moderate	Metapelites
Chrome spinel	4.43–5.09	$7\frac{1}{2}$–8	High	High	Ultramafic igneous rocks
Clinopyroxene	2.96–3.52	5–$6\frac{1}{2}$	Low	Low	Basic igneous and metamorphic rocks
Epidote group	3.12–3.52	6–$6\frac{1}{2}$	Low	Low	Low-grade metamorphic rocks
Garnet	3.59–4.32	6–$7\frac{1}{2}$	Moderate	Moderate-high	Metasediments
Kyanite	3.53–3.65	$5\frac{1}{2}$–7	High	Low-moderate	Metapelites
Monazite	5.00–5.30	5	High	High	Metamorphic rocks; granites
Rutile	4.23–5.50	6–$6\frac{1}{2}$	High	High	High-grade metamorphic rocks
Sillimanite	3.23–3.27	$6\frac{1}{2}$–$7\frac{1}{2}$	High	Low	Metapelites
Staurolite	3.74–3.83	$7\frac{1}{2}$	High	Moderate	Metapelites
Titanite	3.45–3.55	5	Moderate	Low-moderate	Various igneous and metamorphic rocks
Tourmaline	3.03–3.10	7	High	High	Metasediments: granites
Zircon	4.60–4.70	$7\frac{1}{2}$	High	High	Granites and other acidic igneous rocks

*High-stability minerals (such as zircon, rutile, and tourmaline) are commonly recycled.

provenance signal at various points in the sedimentation cycle. These include (1) weathering in the source region, which may affect the mineralogy prior to incorporation in the transport system; (2) abrasion during transport, which may cause destruction of mechanically unstable minerals; (3) weathering during periods of alluvial storage on the floodplain, which may cause depletion of chemically unstable components; (4) hydrodynamics during transport and at the final depositional site, which may affect the relativ e abundance of minerals with different hydraulic behavior; and (5) diagenesis, which may cause depletion of minerals that are unstable in the subsurface. The effects of diagenesis are particularly pervasive, with several case studies (for example, North Sea, United States Gulf Coast, Bay of Bengal) showing that mineral suites become progressively less diverse as burial depth increases. This is the result of dissolution of unstable minerals by circulating pore waters, a process known as intrastratal solution. *See* DIAGENESIS; WEATHERING PROCESSES.

Single-grain heavy mineral geochemical analysis (for example, by electron microprobe) gives additional sophistication to provenance determination. For instance, tourmaline compositions can be used to distinguish granitic and metasedimentary sources, and constraints on plate-tectonic settings of sedimentary basins have been inferred from geochemical analysis of detrital augite. Garnet compositions are particularly useful in provenance studies in view of their response to differences in metamorphic grade and protolith composition. Some heavy minerals, notably zircon and monazite, can be dated radiometrically, and the combined mineralogical and isotopic characterization of sediment source areas is a very powerful tool for provenance studies. *See* AUGITE; ELECTRON MICROSCOPE; GARNET; GRANITE; TOURMALINE.

Heavy minerals have important economic applications. Their use in paleogeographic reconstructions, especially in elucidating sediment transport pathways, is of particular value in hydrocarbon exploration, and their use in correlation has important applications in hydrocarbon reservoir evaluation and production. Recent advances have made it possible to utilize the technique on a real-time basis at the well site, where it is used to help steer high-angle wells within the most productive reservoir hori zo ns. Heavy minerals may become concentrated naturally by hydrodynamic sorting, usually in shallow marine or fluvial depositional settings. Naturally occurring concentrates of economically valuable minerals are known as placers, and such deposits have considerable commercial significance. Cassiterite, gold, diamonds, chromite, monazite, and rutile are among the minerals that are widely exploited from placer deposits. *See* DATING METHODS; MONAZITE; PLACER MINING; WELL; ZIRCON. Andrew Morton

Bibliography. M. A. Mange and H. F. W. Maurer, *Heavy Minerals in Colour*, Chapman and Hall, London, 1992; A. C. Morton and C. R. Hallsworth, Processes controlling the composition of heavy mineral assemblages in sandstones, *Sed. Geol.*, 124:3–29, 1999.

Heavy water

A form of water in which the hydrogen atoms of mass 1 (^{1}H) ordinarily present in water are replaced by deuterium (symbol D or ^{2}H), the heavy stable isotope of hydrogen of mass 2. The molecular formula of heavy water is D_2O (or ^2H_2O).

Properties. Because the mass difference between ^{1}H and ^{2}H is the largest for any pair of stable (nonradioactive) isotopes in the periodic table, many of

TABLE 1. Physical properties of ordinary and heavy water

Property	^1H_2O	^2H_2O (D_2O)
Molecular weight, ^{12}C scale	18.015	20.028
Melting point, °C	0.00	3.81
Normal boiling point, °C	100.00	101.42
Temperature of maximum density, °C	3.98	11.23
Density at 25°C, g/cm^3	0.99701	1.1044
Critical constants		
Temperature, °C	374.1	371.1
Pressure, mPa	22.12	21.88
Volume, cm^3/mol	55.3	55.0
Viscosity at 55°C, mPa · s	0.8903	1.107
Refractive index, n^{20}_D	1.3330	1.3283

the physical and chemical properties of the pure isotopic species and their respective compounds differ to a significant extent. Selected physical properties of ^1H_2O and ^2H_2O are compared in **Table 1**.

Heavy water, judging from its higher melting and boiling points, its higher viscosity, and its surprisingly high temperature of maximum density, is a distinctly more structured liquid than is ordinary water. Heavy water is more extensively hydrogen-bonded, and the hydrogen bonds formed by ^2H are somewhat stronger than are those of ^1H. Nonpolar solutes induce structure in ^2H_2O to a greater extent than in ^1H_2O, and structure-breaking salts are more disruptive in ^2H_2O, largely because there is more structure to break. With increasing temperature, structure is broken down more rapidly in ^2H_2O than in ^1H_2O, and as a result ^2H_2O may be more structured or less structured than is ^1H_2O, depending on the temperature. Many ionic solutes are distinctly less soluble in D_2O than in H_2O.

The nuclear properties of ^1H and ^2H are very different. The capture cross section (Σ_a) of deuterium for low energy of neutrons is much smaller than that of ordinary hydrogen. For ^1H_2O, Σ_a is 2.2×10^{-2} cm^{-1} as compared with the value of 8.5×10^{-6} cm^{-1} for ^2H_2O. The nuclear spin of ^1H is $^1/_2$, but that of ^2H is 1. Thus, both isotopes can be used in nuclear magnetic resonance spectroscopy. The resonance frequency for ^2H is much lower than that of ^1H at a given magnetic field, and the relative sensitivity for detection is much higher for ^1H than for ^2H.

The principal difference in chemical behavior between ^1H and ^2H derives from the generally greater stability of chemical bonds formed by ^2H. The most important factor contributing to the difference in bond energy is the lower zero-point vibrational energy (of the order of 5.021–5.275 kilojoules/mol) for chemical bonds formed by ^2H. Both kinetic and static isotope effects result. The ion product constant of ^2H_2O is 1.11×10^{-15} at 25°C (77°F) compared to 10^{-14} for ^1H_2O. The difference by a factor of 10 in ionization constant causes values of pH and pD for solutions of identical composition in the two media to differ significantly. The nonidentity of pH and pD for solutions of a given composition can be critical for phenomena sensitive to hydrogen ion activity, and can make for serious problems in the interpretation of experiments in many biological systems. In the pH range 2–9, it has been established that pD = pH + 0.41 (molar scale; 0.45 molal scale) as measured by an ordinary glass electrode.

Replacement of more than one-third of the ^1H by ^2H in the body fluids of animals by administration of heavy water, or two-thirds of the hydrogen in higher plants, is lethal. However, numerous green and blue-green algae have been successfully cultured in >99.5% ^2H_2O on carbon dioxide and inorganic nutrients. Fully deuterated organisms can, in turn, be used to start a food chain for the growth of nutritionally more demanding bacteria, molds, and even protozoa in fully deuterated form. These organisms contain >99.5% D, and cultures of these organisms of unnatural isotopic composition have been successfully maintained for years in a fully deuterated form.

Preparation. Deuterium is present in ordinary water to the extent of 0.0145%. Water is the only practical source of deuterium, and very large amounts of water must be processed to produce relatively small amounts of heavy water. Deuterium can be extracted from water (as highly enriched ^2H_2O) by electrolysis, by distillation, or by chemical exchange. Electrolysis of ordinary water is very costly in energy, and has been used only for small-scale production or to enrich partially concentrated material. Distillation of water has been successfully used to separate highly enriched D_2O from ordinary water, but again the energy costs are high per unit of separative work performed. More efficient is the distillation of liquid hydrogen at cryogenic temperatures. Excellent fractionation factors can be achieved in the distillation of liquid hydrogen. The power requirements are modest, but the requirements for very large hydrogen feeds of very high purity (traces of oxygen, nitrogen, carbon monoxide, and so on, are solids at liquid hydrogen temperatures and may clog the apparatus) have mitigated against large-scale use. As the production of hydrogen for synthetic ammonia or coal liquefaction expands, however, by-product extraction of deuterium from the hydrogen stream by distillation of liquid hydrogen may well become important.

Various chemical exchange reactions have been considered for concentrating deuterium from natural water (**Table 2**). Isotope exchange reactions between hydrogen gas and water or ammonia, or between hydrogen sulfide and water, provide the best point of departure for the large-scale manufacture of heavy water. The H_2O/H_2S process is the chemical exchange process of principal commercial interest. Because the equilibrium constant for the distribution of deuterium between H_2S gas and liquid H_2O is temperature-dependent, the process can be carried out in the form of a dual-temperature exchange process. In the dual-temperature H_2O/H_2S process, exchange of ^2H between H_2S (gas) and H_2O (liquid) is carried out at elevated pressure (~2 megapascals or 20 atm) at 120–140°C (250–280°F); ^2H displaces ^1H in the H_2S and becomes concentrated in the gas. At 30°C (86°F) the equilibrium is shifted and reversed,

TABLE 2. Chemical exchange reactions for concentrating deuterium from water

Reaction	Equilibrium constant: Low temp. (°C)	Equilibrium constant: High temp. (°C)
$H_2O(l) + HDS(g) \rightleftharpoons NDO(l) + H_2S(g)$	2.18 (20°)	1.83 (130°)
$NH_3(l) + HD(g) \overset{catalyst}{\rightleftharpoons} NH_2D(l) + H_2(g)$	6.60 (−50°)	4.42 (0°)
$H_2O(g) + HD(g) \overset{catalyst}{\rightleftharpoons} HDO(g) + H_2(g)$	3.62 (25°)	2.43 (125°)

and ^2H is stripped from the gas into the liquid water. The dual-temperature exchange process is carried out in a pair of gas-liquid contacting towers; the cold tower operates at 30°C (186°F), the hot tower at 120–140°C (250–280°F). Water fed to the system flows downward through the cold tower and then through the hot tower countercurrent to a stream of hydrogen sulfide. The water is progressively enriched in deuterium as it passes through the cold tower and progressively depleted in transit through the hot tower. Deuterium-enriched water is drawn from the bottom of the cold tower, and H_2S enriched in deuterium is removed from the top of the hot tower. Further enrichment then occurs in a second stage. To produce 1 metric ton of D_2O, the plant must process 41,000 tons (37,000 metric tons) of water and must cycle 135,000 tons (120,000 metric tons) of H_2S. The product from the exchange plant is enriched to around 15% deuterium, and enrichment to 99.5% is accomplished by vacuum distillation. The dual-temperature GS H_2S/H_2O exchange process is used by all heavy-water plants producing more than 20 metric tons of D_2O per year.

Analysis. The principal methods for determining the deuterium content of heavy water are density determinations, infrared spectroscopy in the ~3-micrometer region, and absorption spectrophotometry in the near-infrared region of the spectrum between 1 and 2 μm. Mass spectroscopy, interferometry, falling-drop methods, and nuclear magnetic resonance spectroscopy have also been used for determining the $^1H/^2H$ ratio in heavy water. Absorption spectrophotometry in the near infrared is a particularly useful procedure because it can be carried out in conventional recording spectrophotometers.

Uses. The only large-scale use of heavy water in industry is as a moderator in nuclear reactors. The Canadian heavy-water-moderated natural-uranium-fueled CANDU reactor uses 0.94 ton (0.85 metric ton) of heavy water per electrical megawatt of installed capacity. The heavy water is not consumed by reactor operations, and the demand for heavy water is determined by the rate at which new nuclear power stations are built. *See* NUCLEAR REACTOR.

Deuterium may become important in energy production by nuclear fusion by the reaction $^2_1H \rightarrow {}^3_1H + {}^1_1H + 4.0$ MeV. It has been estimated that the entire annual energy requirements for the United States in the year 2020 could be supplied by the D-D nuclear fusion reaction by the amount of deuterium contained in 5500 tons (5000 metric tons) of D_2O. *See* NUCLEAR FUSION.

Small amounts of heavy water are used to grow fully deuterated organisms, which serve as a source of fully deuterated compounds of biological importance. These are used in research techniques such as small-angle neutron scattering, in high-resolution nuclear magnetic resonance spectroscopy of immobilized samples, and in the study of isotope effects. *See* DEUTERIUM; ISOTOPE SEPARATION; LIQUID; TRITIUM; WATER.

Joseph J. Katz

Bibliography. J. J. Katz and H. L. Crespi, in *Isotope Effects in Chemical Reactions*, ACS Monogr. 167, 1971; I. Kirshenbaum, *Physical Properties and Analysis of Heavy Water*, 1951; G. M. Murphy (ed.), *Production of Heavy Water*, 1955; H. K. Rae (ed.), *Separation of Hydrogen Isotopes*, ACS Symp. Ser. 68, 1978.

Hederellida

A marine bryozoan order in the class Stenolaemata. It is recorded from scattered localities in the Ordovician through Permian systems of the Paleozoic Era but is found prominently in Devonian strata. The colonies comprise tubular zooecia that commonly encrust brachiopod shells, echinoderm plates and columnals, and the outer surfaces of corals. Although usually encrusting, the colonies sometimes may be erect structures. The initial zooecium is a bulbous ancestrula, which gives rise to a tubular zooecium. Succeeding tubular zooecia bud, one at a time, from the lateral wall of the preceding zooecium. The tubular zooecia have perforated walls, and the distal opening of the zooecium is commonly sealed by a plate (possibly perforate). The zooecial wall apparently consisted of two layers, but it lacks the distinctive laminate structures of the order Trepostomata (Stenolaemata) and shows no additional thickening in the outer parts of the colony.

The simple, perforate tubular zooecia, in which the commonly oval zooecial opening is the same diameter as the zooecial tube, and the lack of accessory tubes or structures separate the hederellids from the Paleozoic cyclostomes. There are six recognized genera, including *Hederella, Reptaria, Hederopsis*, and *Clonopora*.

June R. P. Ross

Bibliography. R. S. Bassler, The Hederelloidea, a suborder of Paleozoic cyclostomatous Bryozoa, *Proc. U.S. Nat. Mus.*, 87:25–91, 1939; R. A. Robison (ed.) *Treatise on Invertebrate Paleontology*, revised, 1983.

Hedgehog

Members of the family Erinaceidae (order Insectivora), which includes 7 genera and 19 species of medium- to large-size animals generally

Hedgehog (*Erinaceus europaeus*).

characterized by spines on their back and sides. This family, however, is subdivided into two subfamilies: the spiny hedgehog group which includes the European hedgehog (*Erinaceus europaeus*); and the hairy hedgehogs or gymnuras, long-tailed species which lack spines. The typical spiny hedgehog has well-developed eyes and a rudimentary to moderately long tail. In most species, there are five toes on each foot, although in some species these are reduced to four on the hindfeet. The spines can be erected by strong muscles, and serve as a protection for the naked or hairy belly, head, and limbs when the animal rolls itself into a ball.

The European hedgehog (see **illus.**) grows to a length of about 9 in. (23 cm). It is a nocturnal animal and rests in thickets or undergrowth, or in crevices during the day. It has 36 sharp teeth suited to its omnivorous diet, which includes worms, insects, reptiles, small rodents, and mollusks. The dental formula is I 3/2 C 1/1 Pm 3/2 M 3/3. The adult is relatively immune to the poison of venomous snakes; thus hedgehogs are valuable in controlling the viper population. The gestation period is 30–50 days and there may be two litters each year of from two to seven young, which are born blind and with soft spines. The young are weaned after a month, and reach sexual maturity at 10 months. *See* DENTITION.

The hairy hedgehogs such as the moon rat (*Echinosorex gymnurus*) are found in southern Asia, Borneo, and Sumatra. This is one of the four species of this little-known group of ratlike animals. The body is over a foot long, which makes it one of the largest living species of insectivore. These animals have a characteristic odor due to the presence of scent glands. It is thought that the diet consists of insects and fruits. *See* INSECTIVORA; SCENT GLAND. Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999.

Helicity (quantum mechanics)

A quantum-mechanical variable that specifies the component of spin-angular momentum of a particle along its direction of motion. The helicity of a particle with rest mass depends on the reference frame, because such a particle has velocity less than the velocity of light c in vacuum. Hence one can make a Lorentz transformation along the direction of motion to a frame in which the particle moves in the opposite direction. For a massless particle, reversal in direction of motion by Lorentz transformation is impossible, and helicity is a Lorentz-invariant quantity. For example, a beam of circularly polarized light consists of photons (zero-mass quanta of the electromagnetic field) with helicity $h = +1$ for right circular polarization, or $h = -1$ for left circular polarization (see **illustration**). *See* LORENTZ TRANSFORMATIONS; PHOTON; POLARIZATION OF WAVES; POLARIZED LIGHT; RELATIVITY; SPIN (QUANTUM MECHANICS).

According to quantum mechanics, the spin angular momentum of a spin-$^1/_2$ particle (for example, the electron or muon) can have only two possible values relative to any given axis: parallel or antiparallel. Thus the helicity of a spin-$^1/_2$ particle can always be thought of as a superposition of two components: a "left-handed" or "left-chirality" portion (spin opposed to direction of motion) and a right-chirality portion (spin parallel to direction of motion). [the word chirality means handedness]. *See* NONRELATIVISTIC QUANTUM MECHANICS.

Neutrinos (ν) and antineutrinos ($\bar{\nu}$) are spin-$^1/_2$ particles of very small rest mass that are produced and absorbed only in weak interactions, and then only in their left-chirality and right-chirality states, respectively. If ν and $\bar{\nu}$ were zero-mass particles (and most physicists believed this to be possible until 1998), their helicities and chiralities would be identical, and we would have $h(\nu) = -1$, $h(\bar{\nu}) = +1$. However, since 1998 neutrino-oscillation experiments have given strong evidence that at least two of three neutrino and antineutrino species have distinct rest masses greater than zero. Thus, although the weak coupling of a neutrino (antineutrino) occurs only in its left (or right) chirality state, the particle, once generated, propagates in a superposition of left- and right-chirality states, yielding a helicity that depends on the reference frame. For example, it is useful to consider the weak interaction (1) in

$$n \rightarrow p + e^- + \bar{\nu}_e \qquad (1)$$

which a neutron decays to a proton, an electron, and

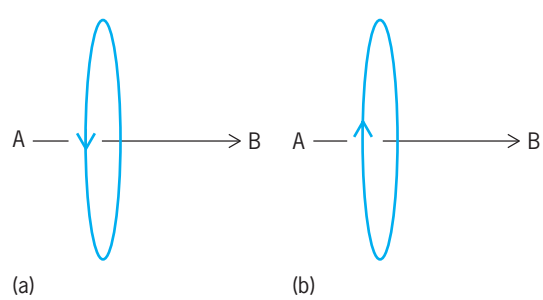

Schematic diagrams of circularly polarized light. (*a*) Right circular polarization. The arrow AB indicates the direction of light propagation. The arrow on the circle indicates the "rotation" associated with the angular momentum of the photon. This angular momentum is parallel to AB. In the description given by classical optics, the arrow on the circle indicates the sense of rotation of the electric and magnetic vectors in the electromagnetic field. (*b*) Left circular polarization. The symbols have the same meaning, but the angular momentum of the photon is in the opposite direction.

an electron-antineutrino. According to the standard theory of weak interactions, modified to account for finite neutrino mass, the average value of helicity for an ensemble of antineutrinos with velocity $v(\bar{\nu})$ is given by Eq. (2), where $<\ldots>$ means average. Since

$$\langle h(\bar{\nu})\rangle = +\frac{v(\bar{\nu})}{c} \qquad (2)$$

in the neutron rest frame virtually all emitted neutrinos in neutron beta decay have kinetic energy much greater than their rest energy, v is very nearly equal to c for these antineutrinos, and thus $\langle h(\bar{\nu})\rangle \cong +1$. In other words, in the neutron rest frame the helicity is nearly, but not exactly, the same as it was in general when a mass $m(\bar{\nu}) = 0$ was assumed. *See* NEUTRINO; WEAK NUCLEAR INTERACTIONS.

Chirality can be an important concept even in circumstances where helicity is irrelevant. For example, in chemistry one encounters examples of two distinct molecules (enantiomorphs) which have the same chemical formula but are mirror images of one another. The spatial structure of one is left-handed (has left chirality), while the other has right chirality. *See* AMINO ACIDS; STEREOCHEMISTRY.

Eugene D. Commins

Bibliography. A. Bertin, R. A. Ricci, and A. Vitale, *Fifty Years of Weak Interaction Physics*, Societa Italiana di Fisica, Bologna, 1984; E. D. Commins and P. H. Bucksbaum, *Weak Interactions of Leptons and Quarks*, Cambridge University Press, 1983; F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, Wiley, New York, 1984; J. J. Sakurai, *Advanced Quantum Mechanics*, Addison-Wesley, Reading, MA, 1978.

Helicobacter

A genus of gram-negative bacilli whose members are spiral shaped, showing corkscrewlike motility generated by multiple, usually polar flagella. *Helicobacter* require low concentrations of oxygen for maximum growth and produce the enzymes oxidase, catalase, and urease.

Different species of *Helicobacter* are found in the stomachs of different animals: *H. felis* in cats and dogs, *H. mustelae* in the domestic ferret, *H. nemestrinae* in the pigtailed macaque, and *H. acinonyx* in captive cheetahs with gastritis. The species of *Helicobacter* found in the human stomach, *H. pylori*, is extremely common. In the United States and similarly developed countries, its prevalence increases at about 1% per year of age so that the majority of adults above age 50 are infected. In less developed countries, infection rates are dramatically higher, with up to 80% of children infected.

Helicobacter pylori is present in virtually all cases of chronic gastritis, which progresses slowly (years or decades) from asymptomatic to atrophic gastritis with impaired acid secretion. Virtually all individuals with duodenal ulcers are infected with *H. pylori*, which colonize sites in the duodenum. The termination of treatment for duodenal ulcers leads to a high rate of recurrence of the ulcers, but ulcer treatment plus eradication of *H. pylori* from the stomach usually leads to a permanent cure. A significant proportion of individuals with *H. pylori*–associated atrophic gastritis develop intestinal-cell metaplasia in the stomach, a condition which is known to represent a precancerous state. Epidemiological studies have confirmed that 75–80% of individuals with gastric cancer have *H. pylori* gastritis. However, the low incidence of duodenal ulcer and gastric or duodenal cancer in *H. pylori*–infected individuals indicates that *H. pylori* infection, although important, is only one of many risk factors leading to these conditions; other risk factors may include diet and variations in the efficiency of the immune system. *See* CANCER (MEDICINE); ULCER.

Helicobacter pylori virulence factors include its shape and ability to rapidly move into and through the gastric mucous coating, which protects the organism from stomach acid, its surface-associated urease enzyme which neutralizes stomach acid near the organism, a cytotoxin, and a fibrillar adhesin which binds the organism to the surface of gastric epithelial cells. *Helicobacter pylori* also has the ability to survive in large numbers in spite of antibodies which are secreted into the stomach and the host immune cell response (inflammatory response) which is characteristic of gastritis.

Detection of *H. pylori* infection is not difficult. Procedures include detection of circulating anti–*H. pylori* antibody in the blood; histological examination of stomach biopsies in which *H. pylori* may be seen residing in gastric mucus or attached to gastric epithelial cells; cultivation of *H. pylori* from a stomach biopsy; and detection of the potent urease enzyme of *H. pylori* by direct testing of biopsy or indirectly by breath test.

Eradication of *H. pylori* from the stomach is difficult. The most successful therapies are combinations of bismuth compounds plus two or three different antibiotics. *See* MEDICAL BACTERIOLOGY; TOXIN.

Doyle J. Evans, Jr.

Bibliography. S. Baron, *Medical Microbiology*, 4th ed., 1996.

Helicoplacoidea

A small class of spindle-shaped, spirally pleated, primitive echinoderms in the subphylum Echinozoa, from the Early Cambrian (*Nevadella* Zone) in eastern California, western Nevada, and eastern British Columbia. Since their discovery in the early 1960s, three genera and six species have been described based on nearly 600 complete specimens, making helicoplacoids the most diverse and abundant echinoderm class known from the Early Cambrian. Most helicoplacoids have a spindle-shaped theca or body with diagonally spiraled pleats, each made up of three rows of flexibly sutured plates forming an interambulacral ridge. Smaller-plated ambulacra spiral around the theca every 7–13 pleats; the ambulacral

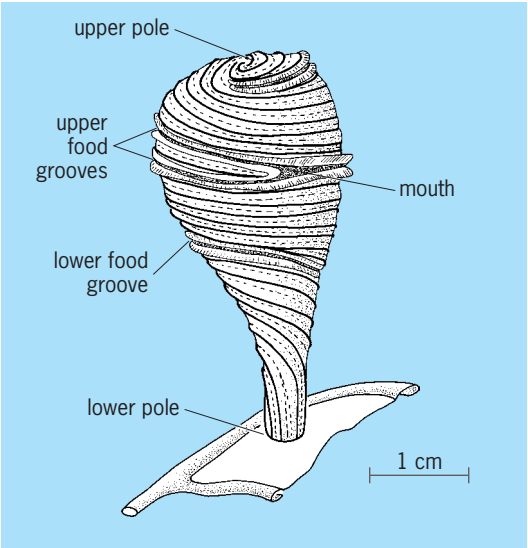

***Helicoplacus gilberti*, side view, reconstructed in living position. The mouth is inferred to be where the ambulacral branches join, and the cover plates over the food grooves are in an open (feeding) position. The lower pointed pole of the helicoplacoid is attached to a trilobite fragment lying on the sea floor. (*After K. Derstler*)**

plates cover a central food groove that has pores for tube feet on each side. The original authors inferred that the mouth was at the more rounded pole of the theca and that a single long ambulacrum branched once about one-third of the way down the theca. However, it has been argued that the mouth was located where the ambulacra split (see **illus.**), so that three ambulacral branches lead away from it, one spiraling partway down the theca and two spiraling up the theca a few pleats apart. No anal opening or pyramid has been found on any of the known specimens.

Helicoplacoids have been reconstructed either as lying recumbent on the sea floor or as standing erect and attached to objects on the sea floor by using the more pointed pole of the theca (see illus.). They were apparently low- to medium-level suspension feeders using their ambulacral tube feet to capture small food particles drifting past the theca. The covered ambulacra bearing pores for tube feet are similar to those of other echinozoans such as edrioasteroids and camptostromatoids. The unusual plating and spiral symmetry of helicoplacoids separate them from all other classes of echinoderms. *See* ECHINODERMATA; ECHINOZOA.

James Sprinkle

Bibliography. K. Derstler, *Studies on the Morphological Evolution of Echinoderms*, Ph.D. dissertation, University of California, Davis, 1985; J. W. Durham and K. E. Caster, *Helicoplacoidea*: A new class of echinoderms, *Science*, 140:820–822, 1963.

Helicopter

An aircraft characterized by its large-diameter, powered, rotating blades, attached to a substantially vertical axis. The helicopter can lift itself vertically by the reactive force generated as the rotating blades accelerate air downward. It can both lift and propel itself by accelerating air downward at an angle to the vertical. The helicopter is the most successful vertical takeoff and landing (VTOL) aircraft developed, by virtue of its relatively high efficiency in performing hovering and low-speed flight missions.

It was not until the 1930s that helicopters began to demonstrate practical capabilities. In 1937 Heinrich Focke built in Germany a twin-rotor helicopter, which ultimately demonstrated meaningful performance capabilities. In 1939 Igor Sikorsky built a relatively simple and controllable single-rotor helicopter, which evolved into the modern standard configuration. In 1946 a single-rotor machine developed by Lawrence Bell received the world's first commercial helicopter license.

Dynamics. The key to understanding the operation and control of a helicopter lies in a knowledge of the forces and resultant motion of each rotor blade as momentum is imparted to the air. Unlike a fixed-wing aircraft, which derives its lift from the translational motion of the fuselage and airfoil-shaped wing relative to the air, the helicopter rotates its wings (or rotor blades) about a vertical shaft and thus is able to generate lift while the fuselage remains stationary.

The rotational motion of the rotor blades creates additional forces that act on the blades as they revolve about the shaft. The principal additional force is the centrifugal force that arises from the circular motion of the blade mass. The centrifugal force creates an effective stiffening of the rotor blade so that a relatively limber blade structure can carry the aerodynamic forces necessary to lift the weight of the helicopter, similar to the phenomenon of supporting a rock attached to the end of a string by whirling it in the horizontal plane.

The major forces on an element of a rotor blade are the lift, drag, weight, and centrifugal force (**Fig. 1**). The forces of greatest interest are the lift, drag, and centrifugal forces, since the blade weight is small relative to the other forces. In order for the rotor to lift the helicopter in hovering flight, the average distributed lift from all blades must equal the weight of the helicopter. With these large lift forces distributed along a blade, it will cone upward until there is a balance of the moments created by the lift and centrifugal forces about the blade attachment point. In hovering flight, both the lift and drag forces are steady, and the blade cones to a constant, equilibrium position. In forward flight, these forces vary as the blade rotates; consequently, the blade's angular position relative to the hub, called the flapping angle (Fig. 1), is a function of these oscillatory aerodynamic forces. In essence, the rotor-blade flapping motion may be considered as a dynamics problem analogous to the forced response of a simple spring-mass-damper system, wherein the centrifugal force is equivalent to the restoring spring and the aerodynamic force provides both the damping and the forcing function for the blade mass.

A unique characteristic of the blade dynamic system just described is that the flapping natural

frequency of the blade is approximately equal to the rotational frequency of the rotor. Since the aerodynamic forcing function is also proportional to the rotational frequency, the rotor blade represents one of the few dynamic systems that is forced at its natural frequency or in resonance. This can be a catastrophic occurrence for mechanical systems without damping due to the large amplitudes of motion that may be generated. In the case of a rotor blade, however, the aerodynamic damping generated by the blade flapping motion is so high that it limits the forced response to acceptable amplitudes.

A major consequence of operating at flapping resonance is the resulting relationship between the maximum force on the blade and its maximum displacement. When a dynamic system is forced at its natural frequency, there is a 90° phase lag between the phase angle of the maximum force and the phase angle at which the maximum amplitude of the response occurs. Thus, when a rotor blade experiences a change in the aerodynamic loading, the blade responds correspondingly one-quarter of a revolution (that is, 90°) later. This behavior is important for controlling the flight motion of the helicopter.

For example, to achieve forward flight, the rotor must be tilted in the direction of the desired flight path. This is achieved by decreasing the lift on a blade rotating toward the nose (advancing blade) and increasing the lift on a blade rotating toward the tail of the helicopter (retreating blade). This difference in lift on the advancing and retreating blades, which is maximum when the blades are arrayed laterally, will cause the advancing blade to flap to a lower amplitude when it arrives over the nose while the retreating blade with more lift flaps to a higher amplitude when it arrives over the tail, thus creating the forward tilt of the rotor.

The method for achieving the desired control of the rotor tilt is to vary the angle of the blade (pitch) as the rotor rotates. This variation is achieved by a control device called a swash plate, which, when tilted, will cycle the blade pitch angle in a sinusoidal manner so that blade lift is alternately increased and decreased once every rotor revolution. As discussed below, this cyclic control is also required to accommodate the lift variations encountered when the rotor is in forward flight. *See* AERODYNAMIC FORCE.

Aerodynamics. The basic physics of rotor aerodynamics, particularly the aerodynamic relationships for a rotor in hover or vertical climb, closely parallels the physics of a propeller. In a climb, for example, the air flows through the rotor perpendicular to the rotor disk plane in the same manner that air flows into a propeller disk as it translates horizontally. Thus, the rotor or propeller blades experience the same aerodynamic environment as they rotate. *See* PROPELLER (AIRCRAFT).

As the helicopter translates into forward flight, significant differences occur. Because the rotor blades are mounted on a vertical axis and rotate in a horizontal plane, the translational velocity of the helicopter significantly alters the relative airspeed of a rotor blade element as the rotor blade revolves (**Fig. 2**). The blade moving in the direction of flight (advancing blade) encounters a relative airflow velocity which is equal to the vector sum of the helicopter's flight velocity V and the blade's rotational velocity. The rotational velocity is proportional to the radial distance along the blade, and at the tip is equal to the tip speed ΩR, where Ω is the angular velocity of the rotor and R is the radial distance to the blade tip. On the retreating side of the disk, the blade experiences a relative airflow velocity equal to the difference between the two velocities. When the blades are positioned fore and aft, the flight velocity does not contribute substantially to the relative airflow velocity across the blade.

This changing airflow velocity as the blades rotate produces a sinusoidal lift variation if the blade pitch angle is held constant. Since lift is proportional to the velocity squared, an advancing blade with its higher velocity will generate more lift than a retreating blade. Recalling the dynamic response characteristics of the blades discussed above, the blade with

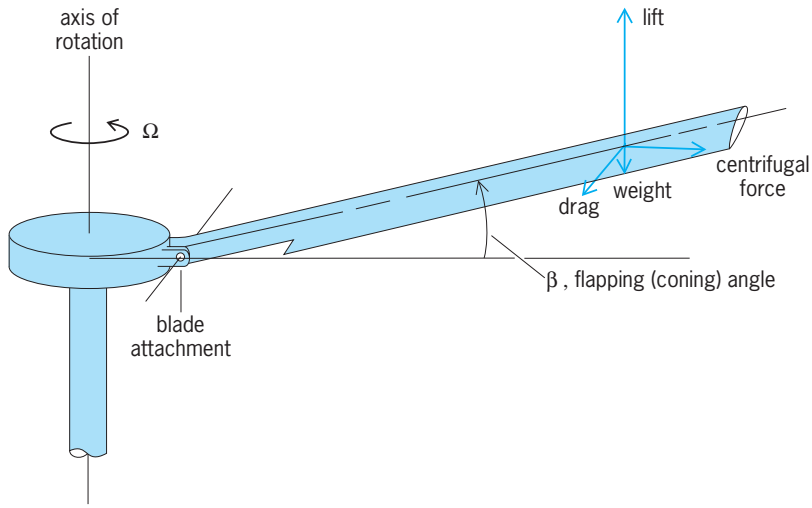

Fig. 1. Forces acting on a helicopter rotor blade.

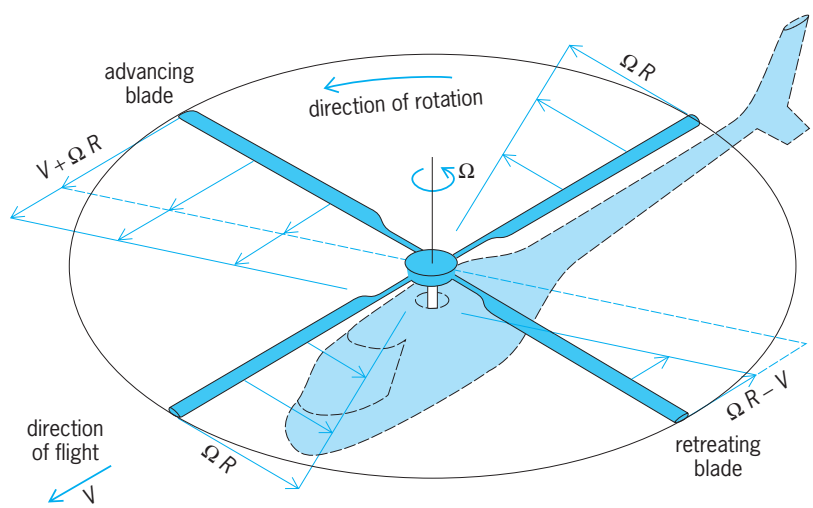

Fig. 2. Variations of relative airflow velocity of rotor blades in forward flight.

greater lift flaps to a larger angle than the blade with less lift, causing the rotor to tilt in a direction opposite to the direction of flight. To counter the difference in lift between the advancing and retreating blades, cyclic pitch control is again introduced to lower the pitch, and consequently the lift of the advancing blades, while increasing the pitch and lift of the retreating blades, thus maintaining the desired tilt of the rotor.

This difference in velocity between advancing and retreating blades is the principal reason that helicopters have a lower speed capability than a fixed-wing aircraft. Since helicopters generally operate with a tip speed ΩR in the range of 600–800 ft/s (180–240 m/s), it does not require much forward speed before the advancing blade is moving relative to the air at a speed approaching that of sound (1117 ft/s or 340 m/s). At these speeds, the airfoil sections of the blade experience compressible flow effects which drastically increase the blade drag and hence the power required. On the retreating blade, the situation is just the opposite. As the helicopter goes faster, the retreating blade experiences a relative airflow that is diminished by the forward velocity. Under these conditions, the blade tip would have no movement relative to the air if the forward velocity were equal to the tip speed of the rotor. Even at forward velocities much less than the tip speed of the rotor, there are regions where the forward velocity equals or exceeds the local velocity due to rotation. The diminishing velocity on the retreating blades is compensated for by increasing the blade pitch and thus the angle of attack of the blades' airfoil. However, an airfoil has a limiting angle that, if exceeded, will result in flow separation (stall). When stall occurs, there is an increase in drag and decrease in lift capability of the airfoil, which again causes an increase in the power required. These conditions make the design of a helicopter a compromise between compressibility effects on the advancing blades and retreating blade stall. *See* AERODYNAMICS; AIRFOIL.

Performance. Helicopter performance is determined by the power required to achieve the desired flight condition. There are four primary functions associated with helicopter flight that require power: the power required to turn the rotor (profile power), the power required to lift the helicopter (induced power), the power required to propel the helicopter in forward flight (parasite power), and the power required to climb or descend (climb power). Each of these elements contributes in a different manner to the total power required as the flight speed varies.

Profile power. The profile power has a moderate value in hover that is a function of the area of the blades being turned, the drag characteristics of the airfoil sections of the blades, and the cube of the effective speed of the rotor. In hovering flight, the effective speed of the rotor is only a function of the tip speed at which the rotor operates. As the helicopter moves into forward flight, the effective speed of the rotor blades increases with a corresponding increase in profile power. For the example illustrated, the drag characteristics of the airfoil section are assumed to be constant. If the rotor encounters stall or compressibility effects, the drag characteristics will increase and, correspondingly, the profile power required will increase more dramatically with airspeed.

Induced power. The induced power has just the opposite trend, as shown by the decrease in induced power as airspeed is increased. Since induced power is associated with the generation of lift, increasing or decreasing the weight of a given helicopter will cause a corresponding change in induced power.

The size of the rotor also plays a major role in determining the induced power. A rotor generates lift by accelerating air downward or, in effect, increasing the momentum of the downward flow. The larger the rotor diameter, the greater the air mass it will act on, and the lower the velocity change required to produce a given lift. Induced power is directly proportional to this induced velocity, which, in turn, is related to the disk loading of the rotor, that is, the ratio of the lift generated to the area of the planar disk described by the blade radius. Helicopters, with their large-diameter rotors, are low disk-loading aircraft, which accounts for their relatively good efficiency in hovering flight compared to other means of achieving powered vertical lift. Even so, the power required is still sizable, which explains the difficulties encountered by the pioneers in providing an engine light enough and powerful enough to achieve vertical flight.

In forward flight, the mass flow of air through the rotor is increased; consequently, the blades do not have to induce as much velocity change to this higher mass flow in order to produce a given lift. This reduction in induced velocity decreases the induced power required as speed is increased.

Parasite power. The power required to propel the helicopter in forward flight (parasite power) is predominantly a function of the drag characteristics of the fuselage and rotor hub and the cube of the flight velocity. Therefore, the parasite power is zero in hovering flight but rises rapidly as airspeed is increased. At high speeds, the parasite power is the predominant cause of power expenditure, illustrating the need for fuselages and rotor hubs that possess very low drag characteristics.

Climb power. The climb power is proportional to the weight of the helicopter and its velocity of climb or descent. When a helicopter is climbing, more power is required; however, when it is descending, the helicopter requires less power than it does in level flight. If the rate of descent is high enough, the helicopter can achieve a condition called autorotation wherein no engine power is required. Autorotation for a helicopter is similar to the glide capability of fixed-wing aircraft. The helicopter, however, has a distinct advantage over an airplane in this regard since a helicopter can autorotate at much lower airspeeds than the glide speed of an airplane. By using the energy

derived from the descent, the helicopter can produce the energy needed to generate the necessary lift and to maintain the desired rotational speed of the rotor. As the helicopter nears the ground, the pilot executes a maneuver, called a flare, which arrests any forward velocity, and the aircraft can be landed in a very small landing site.

Rotor configurations. Many different rotor arrangements have been used (**Fig. 3**), and most of the early attempts at vertical flight were made with machines having multiple or coaxial counterrotating rotors. Most modern helicopters employ the single rotor or the tandem rotor configurations.

In addition to the selection of the number and location of the lifting rotors, designers have developed varied methods for attaching the blades to the rotor hub. Very early experiments conducted with the blades rigidly attached to the hub were unsatisfactory because of the excessive moments applied to the rotor mast. The first satisfactory solution to this problem was applied by Juan de la Cierva in his development of the autogiro. Cierva incorporated a hinged blade attachment to allow the blades to flap freely and to relieve the undesired moments.

Teetering rotors. Based on the success achieved by the introduction of hinged attachments for the rotor blades, several configurations have been successfully manufactured (**Fig. 4**). The teetering rotor used on two-bladed configurations has one central hinge that allows the blades to move in unison (one up, one down) like a seesaw (Fig. 4*a*). Each blade also has pitch-change bearings to allow the blade pitch angle to be varied as required. The gimbaled rotor (Fig. 4*b*) is essentially equivalent to the teetering rotor and has been used on rotors with three or more blades. It allows the rotor disk to flap as a unit. In both the teetering and gimbaled arrangements, moments are reacted between the blades and are not transferred into the shaft.

Articulated rotor. The articulated rotor (Fig. 4*c*) has each blade attached to the hub by its own flapping hinge. In addition, a hinge is introduced to allow in-plane or lead-lag motion of the blade in order to relieve in-plane bending moments at the root end of the blade. The articulated rotor differs somewhat from the two preceding configurations in that the flapping hinge for each blade is offset from the center of rotation. This offset allows moments to be applied to the shaft and, by selection of a specific hinge offset, the magnitude of the moment is controlled, and is an aid in the control of the helicopter.

Hingeless rotor. The bearingless or hingeless rotor (Fig. 4*d*) is receiving renewed attention in research and development efforts. These efforts represent a closing of the loop in rotor development in that the earliest rotors were basically hingeless designs. There is, however, a major difference in the newer designs which makes them feasible. The early rotors were designed with relatively rigid blade attachments and, given the construction materials of the time, the only way to successfully deal with the large moments was to introduce hinges. With the availability of improved materials, current designs tailor the structural stiffness to allow flexing to take place without having a hinge. This concept provides a lower-weight rotor system and retains the ability to develop hub moments for control.

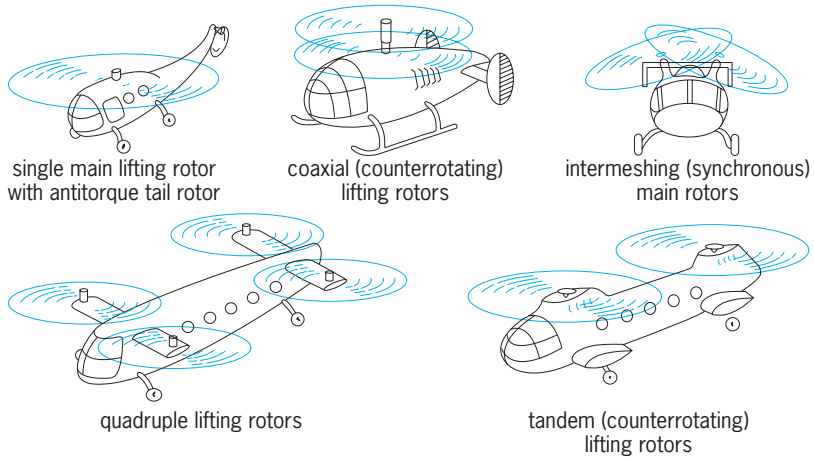

Fig. 3. Principal helicopter (rotor) configurations.

Flight control. In order to utilize fully the capabilities of a helicopter in hover, vertical, sideward, rearward, and forward flight, the flight controls are of necessity complex (**Fig. 5**).

Cyclic pitch control. The primary control, called cyclic pitch control, is introduced by a control stick located between the pilot's knees. Through a series of control linkages the motion of the stick is transferred to tilt the swash plate either longitudinally or laterally so that blade pitch is cyclically varied in a

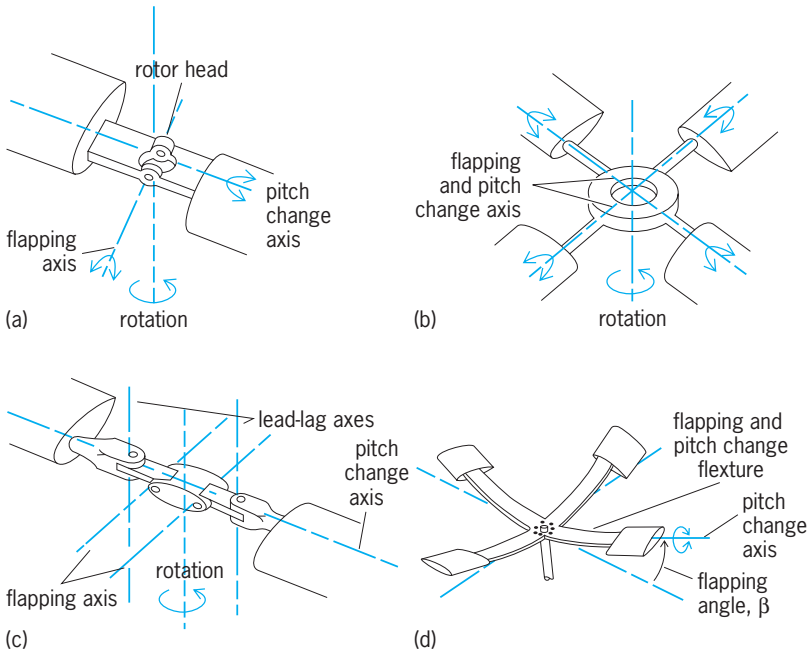

Fig. 4. Principal types of rotor hubs. (*a*) Teetering (semirigid) rotor head. (*b*) Gimbaled rigid rotor. (*c*) Fully articulated rotor head. (*d*) Hingeless rotor.

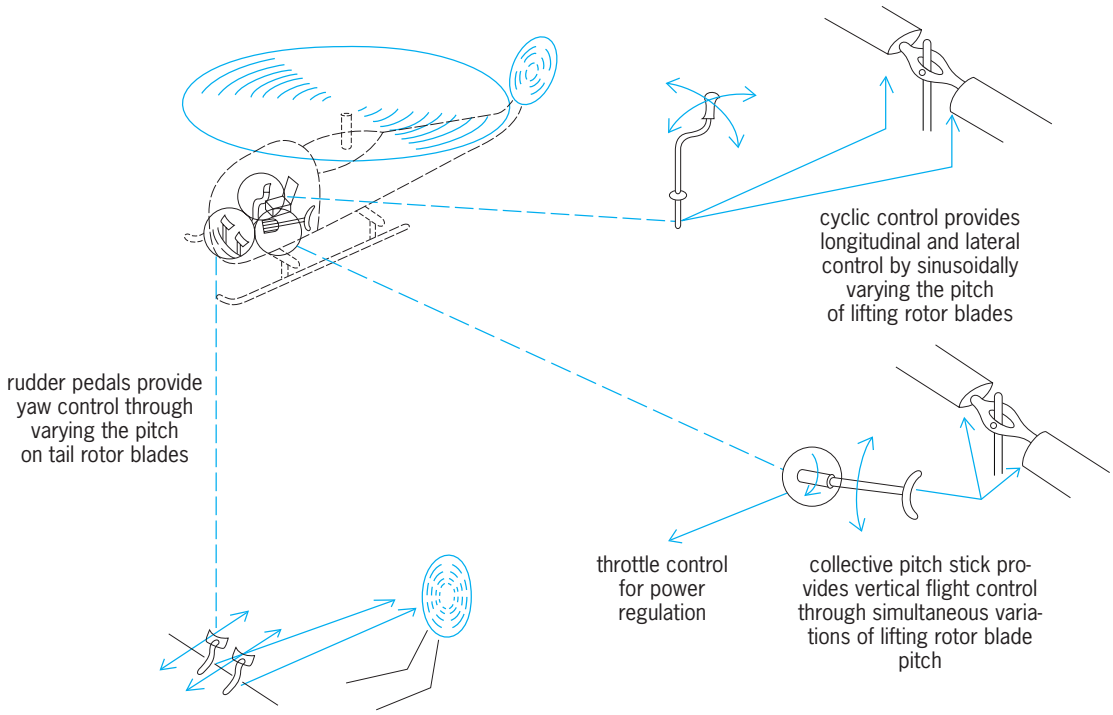

Fig. 5. Functions of helicopter flight controls.

prescribed manner. The cyclic pitch control directs the helicopter in the same direction in which it is moved: if the stick is moved forward, the helicopter moves forward; if it is moved to the right, the helicopter moves to the right, and so forth.

Collective pitch control. The vertical direction of movement is controlled by the collective pitch, which moves the aircraft up if the collective lever is raised and down if the stick is lowered. The collective lever, like the cyclic control, changes the pitch angle of the blades in order to increase or decrease the blade lift. However, a collective input to the blades will change their pitch by an equal amount regardless of their azimuthal position in the plane of rotation. The collective lever is located by the pilot's left hand and, although it may be locked temporarily with a friction lock, most of the time it must be held in the pilot's hand.

Yaw control. The yaw or directional control of the fuselage is provided by pedals similar to those employed by fixed-wing aircraft. On a single-rotor helicopter, the pedals are used to control the pitch of a small tail rotor. The thrust provided by the tail rotor generates the antitorque required to counteract the main-rotor driving torque or to change the directional heading of the helicopter.

Engine fuel control. Because of the varying power requirements of a helicopter, the engine throttle is constantly monitored. Normally, it is controlled by a motorcycle-type twist grip mounted on the collective pitch lever. On early reciprocating engine helicopters, the pilot had to coordinate the throttle position with the power demands. However, on modern turbine-powered helicopters, the engine fuel control is coupled to the engine's power turbine speed by a governor so that rotor rotational speed is maintained automatically. *See* FLIGHT CONTROLS.

Julian L. Jenkins

Higher harmonic control. This is a demonstrated active-control-system concept for helicopters that promises many improvements in such important areas as helicopter vibrations, noise, and even performance. Up to a 90% reduction in helicopter vibration levels has been achieved in flight tests, showing that it is feasible to reduce vibrations to levels comparable to those of a fixed-wing aircraft. Low vibration levels are important for crew and passenger comfort and also to reduce structural fatigue of the airframe, rotor system, drive train, and dynamic components. Efforts in acoustics have been largely directed at reducing or eliminating the annoying blade-slap noise generated by the interaction of a rotor blade with the trailing tip vortex shed by the preceding blade.

Higher harmonic control is an active control concept, in contrast to the conventional passive means of vibration control, such as vibration absorbers, vibration isolators, bifilar or Frahm absorbers, or nodal beam suspension. A passive vibration control device treats the vibratory loads after they have been generated, whereas an active vibration control concept alters or reduces the vibratory excitation at its source. In the case of higher harmonic control, the source of the vibrations is the vibratory aerodynamic forces which act on the rotor blades in flight. The rotor system transmits these unwanted vibrations to the airframe. Higher harmonic control senses these vibrations and applies high-frequency (20–35 Hz) pitch motion to the rotor blades at very

small angles (less than 1°) to suppress this aerodynamic excitation. The primary elements of the active vibration suppression system are acceleration transducers that sense the vibratory response of the fuselage; a higher-harmonic blade-pitch actuator system; a flight-worthy microcomputer, which incorporates the algorithm or mathematical model for reducing vibrations; and a signal-conditioning system (electronic control unit), which interfaces between the sensors, the microcomputer, and the higher-harmonic-control actuators. *See* ACTIVE SOUND CONTROL.

E. Roberts Wood

Applications. The growth of the helicopter industry in the United States is founded on the uses made by the armed forces. From the small two-place helicopter built by Sikorsky Aircraft during World War II to the modern air cavalry concept, the military has been a major proponent of the helicopter. The unique capabilities of the helicopter to operate in confined areas and to take off and land in unprepared sites have made it an indispensable tool in the military environment. Large-scale military orders for helicopters during the Korean campaign provided the impetus to the fledgling industry. The growth in the number of helicopters was paralleled by an equally impressive growth in applications. During the Vietnam conflict of the 1960s, the helicopter experienced a greatly expanded role in the military's tactics. From the outset, the Vietnam conflict was dominated by the helicopter. The absence of usable roads, coupled with the need for extreme mobility, required the flexibility of helicopter transportation to deploy troops and equipment rapidly. Procurement and use of helicopters soared in all the services for such varied tasks as observation, troop carriers, gun ships, rescue, and cargo resupply (**Fig. 6**). *See* MILITARY AIRCRAFT.

The technology that evolved to meet the needs of the military provided the base for an impressive growth in commercial applications. Initially, commercial operators were dependent entirely on derivatives of military aircraft for the equipment used. However, as commercial operations expanded into more areas, a market was established which would support the development of distinctly commercial aircraft.

With such diverse operations as crop spraying, logging, construction, police and ambulance service, and passenger and corporate transportation, the industry has responded with a variety of commercial helicopters (**Fig. 7**). One of the more dramatic areas of growth has been in energy exploration, particularly in the offshore oil operations in regions such as the North Sea and the Gulf Coast of the United States. Transportation of workers and equipment to and from the oil rigs is a role easily filled by the helicopter. However, as the oil rigs are moved farther from shore, the longer mission ranges present

Fig. 6. Military helicopter applications. (*a*) Observation (*Bell Helicopter Textron*). (*b*) Cargo resupply (*U.S. Army*). (*c*) Tactical transportation (*Sikorsky*). (*d*) Attack (*U.S. Army*).

Fig. 7. Commercial helicopter applications: (*a*) Spraying; (*b*) construction; (c) corporate transportation (*Bell Helicopter Textron*). (*d*) Oil exploration (*Sikorsky*).

a significant technical challenge to the industry. *See* OIL AND GAS, OFFSHORE.

Hybrid aircraft. The expanding applications of rotary winged aircraft have created a demand for improved capabilities, including higher speeds, longer range, and greater efficiency. In order to maintain the hovering efficiency of the helicopter while achieving the higher speeds of fixed-wing aircraft, hybrid or convertible aircraft have evolved that maintain the large-diameter rotor efficiency of the helicopter and then convert or change the configuration to achieve high-speed performance. Two configurations that use this approach are the tilt-rotor and the X-wing or stopped-rotor configurations.

Tilt-rotor configuration. The tilt-rotor aircraft (**Fig. 8**) uses large-diameter rotors located on pylons at each end of the wing. For hover and low-speed flight, the tilt-rotor flies in much the same manner as a conventional helicopter. That is, the aircraft is controlled by altering the magnitude and direction of the lift vectors on each of the two side-by-side rotors. For wingborne cruise flight, the pylons are tilted forward so that the rotors function as large propellers. This conversion accomplishes two significant things. First, by rotating the rotor and hub into an axial alignment with the free-stream flow, the parasite drag is considerably reduced, as is the power required for high-speed flight. Second, with the rotor operating in axial flow as a propeller, the sinusoidal flow conditions that exist in the helicopter mode are eliminated, and the rotor can produce propulsive force at much higher speeds.

Full conversion can be accomplished in seconds and it is also possible for the tilt-rotor to fly at intermediate conversion angles over a wide range of airspeeds. Full conversion to the airplane mode can be accomplished only at airspeeds above the stall speed of the wing.

When in the airplane mode, a tilt-rotor is controlled with conventional ailerons, elevators, and rudders in the same manner as any fixed-wing aircraft. In the helicopter mode, the rotors provide the flight control by using collective and cyclic pitch and pedals. Control is accomplished somewhat differently than the single-rotor control described above.

Fig. 8. Concept of advanced tilt-rotor aircraft. (*Bell Helicopter Textron*)

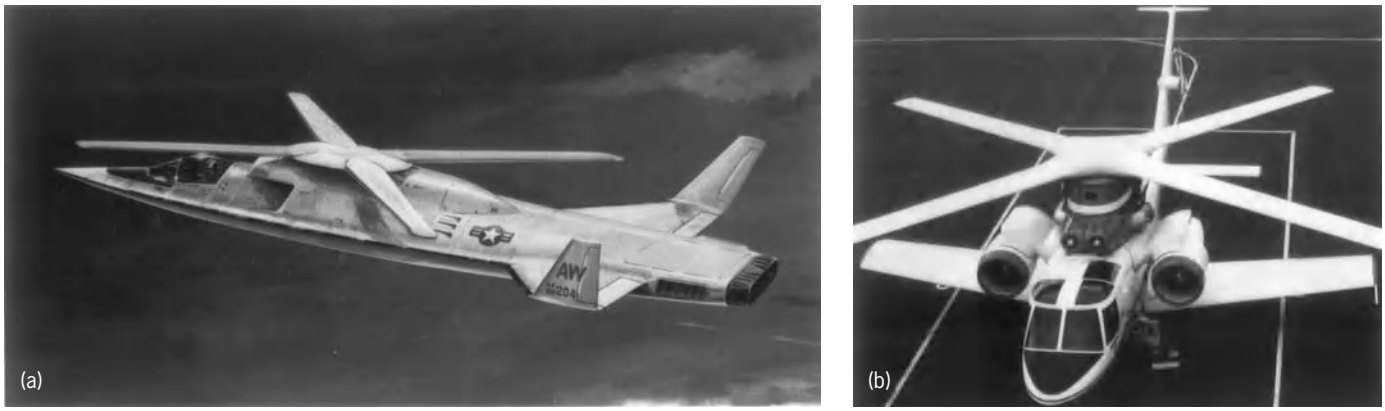

Fig. 9. X-wing aircraft. (*a*) Concept. (*b*) Experimental X-wing, developed by Defense Advanced Research Projects Agency (DARPA), installed on NASA/U.S. Army Rotor Systems Research Aircraft. (*Sikorsky Aircraft Division, United Technologies Corp.*)

Since there are two rotors, roll and yaw control are implemented by introducing differential forces between the rotors. Increasing the collective pitch on one rotor while decreasing it on the other introduces a rolling moment. While differential collective pitch is used for roll control, differential longitudinal cyclic pitch is used for yaw control. Since cyclic pitch tilts the rotor disk in the desired direction of flight, the aircraft executes a yaw turn when one rotor tilts forward and the other back.

The feasibility of the tilt-rotor concept was demonstrated by experimental research aircraft, first flown in 1977. Its success paved the way for the development of the V-22 Osprey tilt-rotor aircraft, flown in March 1989. *See* VERTICAL TAKEOFF AND LANDING (VTOL).

X-wing configuration. In the experimental X-wing or stopped-rotor approach to achieving high-speed flight, rather than conversion of the low-disk-loading rotor into a propeller, the rotor is slowed and stopped in an X position relative to the fuselage (**Fig. 9**). In order to accomplish this conversion, several unique design features must be provided. First, the rotor blades must be extremely rigid since the centrifugal stiffening of a rotating blade is lost when the rotor stops. Second, the ability to generate lift and to control the aircraft while the rotor slows to a stop creates an unusual requirement for both the airfoil section of the rotor/wing and the method of controlling the system. Finally, the power plant must also convert from a system that provides shaft power to the rotor for helicopter flight into a system that provides propulsive force when the rotor is stopped.

To achieve the desired rigidity of the rotor/wing, the X-wing aircraft uses an extremely rigid cantilevered rotor blade. In order to achieve the desired stiffness with a reasonable weight, graphite/epoxy composite materials are used to fabricate an I-beam type of structural spar. The beam is tailored to provide high bending stiffness in order to carry the lifting loads when the rotor is stopped. It is also tailored in torsional bending to allow the beam to be moderately twisted for changes in collective pitch. *See* COMPOSITE MATERIAL.

Unlike a conventional helicopter rotor, however, the X-wing does not require physical movement of the rotor blade as the primary method of achieving collective and cyclic pitch control. Instead, it uses a circulation control airfoil, which relies on the Coanda principle to alter the sectional lift generated. When a thin, low-pressure, high-speed stream of air is blown tangentially to a curved surface, the flow stream tends to follow the surface and thereby alter the velocity field of the airfoil. This change in the velocity field alters the lifting pressure on the airfoil. By modulating the internal air supplied to the slotted airfoil, the lift can be increased or decreased just as if the airfoil were changing its angle of attack. A pneumodynamic control system forces compressed air through ducts in the leading and trailing edges of the rotor/wing blades to provide circulation control lift in rotary and fixed-wing flight and during conversion between the two modes.

The leading edge of a rotor blade in helicopter flight becomes the trailing edge on the left wing when the rotor is stopped for the cruise condition. This unique conversion requires an elliptical airfoil with provisions for blowing air over both the leading and trailing edges of the airfoils during certain azimuth positions in order to maintain the desired aircraft lift and rolling moment trim as the conversion takes place. A convertible engine and declutching arrangement is required to provide an alternative propulsive system once the rotor system is stopped. The convertible engine has two operating modes. In one mode it provides shaft power through a power turbine to turn the main rotor blades when the X-wing operates as a helicopter. In the second mode, the power turbine output shaft is declutched from the rotor and the engine then operates as a turbojet to provide thrust for forward propulsion.

Julian L. Jenkins

Bibliography. J.-P. Harrison, *Principles of Helicopter Flight*, 1993; W. Johnson, *Helicopter Theory,* 1980, reprint 1994; R. W. Prouty, *Helicopter*

Performance, Stability, and Control, 1986, revised 1995; J. Seddon, *Basic Helicopter Aerodynamics*, 1990; W. Z. Stepniewski and C. N. Keys, *Rotary-Wing Aerodynamics*, 1984.

Helicosporida

An order of protozoa in the class Myxosporidea (subphylum Cnidospora). It is characterized by production of spores with a relatively thick, single intrasporal filament and three uninucleate sporoplasms. The spore membrane occurs as a single piece. *Helicosporidium parasiticum* is the only species in the order. The parasite infests the body cavity, fat bodies, and ganglia of mites (*Hericia hericia*) and the body cavity of fly larvae (*Dasyhelea obscura* and *Mycetobia pallipes*) found in the sap of horse chestnut and elm trees. *See* INSECT PATHOLOGY.

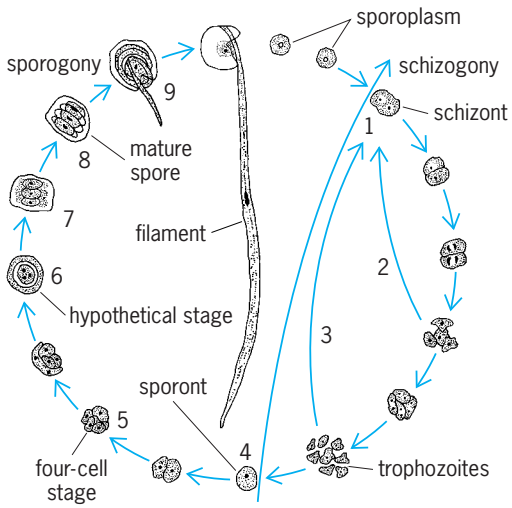

Probable development of a helicosporidan, diagrammatic: 1–3, schizogony, with formation of the schizont; 4–9, sporogony, with development of mature spore; 4–7, sporont arising from trophozoites and development into immature spore; 7, mature spore; 8, opening of mature spore; 9, escape of sporoplasms

Although the complete life cycle of helicosporidans is not known, stages in the development and division of the trophozoite and in the development of the spore have been described (see **illus.**). The sporont or sporoblast gives rise to four cells: One forms the spore membrane and three remain as sporoplasms. The origin of the filament is not known. *See* CNIDOSPORA; DIPTERA; MYXOSPORIDEA; PROTOZOA.

Ross F. Nigrelli

Helimagnetism

A property possessed by some metals, alloys, and salts of transition elements in which the atomic magnetic moments, at sufficiently low temperatures, are arranged in a spiral or helix. It may be seen from the **illustration** that simple antiferromagnets and ferromagnets can be considered as nonconical helimagnets with helical angles ϕ of 180 and 0°, respectively. In the same way, nonconical helimagnets may be considered as conical helimagnets with cone angle θ of 0°. Some typical helimagnets are listed in the **table**. The magnetic structures have been detected by neutron diffraction.

Helimagnetism arises when the exchange coupling parameters J_{ij} between the spins $\mathbf{S}_i$ and $\mathbf{S}_j$ of the magnetic atoms lie within a certain range of values. The exchange energy is given by Eq. (1), summed

$$E_{ij} = -2J_{ij}\mathbf{S}_i \cdot \mathbf{S}_j \tag{1}$$

over all pairs of atoms (i, j). *See* FERROMAGNETISM.

Consider an axial structure with strong ferromagnetic coupling (positive J_{ij}) between atoms which are in the same plane, and weaker ferromagnetic or antiferromagnetic coupling between atoms which are in the same plane, and weaker ferromagnetic or antiferromagnetic coupling between atoms which are in neighbor planes. Let this latter coupling between nearest neighbor planes be J_1, between next neighbor planes be J_2, and so on. Let the angle between the ferromagnetically coupled spins of the nth plane and those of the first plane be ϕ_n. The exchange energy is then given by expression (2), where J is the intraplanar exchange.

$$E \propto -\Sigma_n[J + 2J_1 \cos\left(\phi_{n+1} - \phi_n\right) + 2J_2 \cos\left(\phi_{n+2} - \phi_n\right) + \cdots] \tag{2}$$

If one assumes a helical array, then Eq. (3) holds,

$$\phi_n = nka + \text{constant} \tag{3}$$

where ka, which is the ϕ of array (c) of the figure, is the phase angle between neighbor planes. With a equal to the spacing between planes, k becomes the wave number describing the pitch of the helix.

On substituting Eq. (3) into Eq. (2), one obtains Eq. (4) where J_k is the Fourier transform of the exchange parameter J_{ij}, when the latter is considered a function of distance between planes only.

$$E \propto -[J + 2J_1 \cos ka + 2J_2 \cos 2ka + \cdots] = -J_k \tag{4}$$

It is seen that the exchange energy is minimized by that value of k which maximizes J_k. In this way the pitch of the helix is determined.

Ferromagnets have maximum J_k at $k = 0$. This cannot be achieved if the interplanar J_n are mostly negative, that is, antiferromagnetic. In the latter case what happens physically is that the planes try to order antiferromagnetically with respect to one another. It is impossible for them all to do this, and a helical compromise results.

Conical helimagnets have maximum J_k at $k \neq 0$, and also anisotropic coupling which favors spin alignment along the crystal axis.

Helimagnetism is found in most of the rare-earth metals and their alloys and in a few salts of the iron group. Some helimagnets have quite complicated patterns, such as manganese sulfate which, below

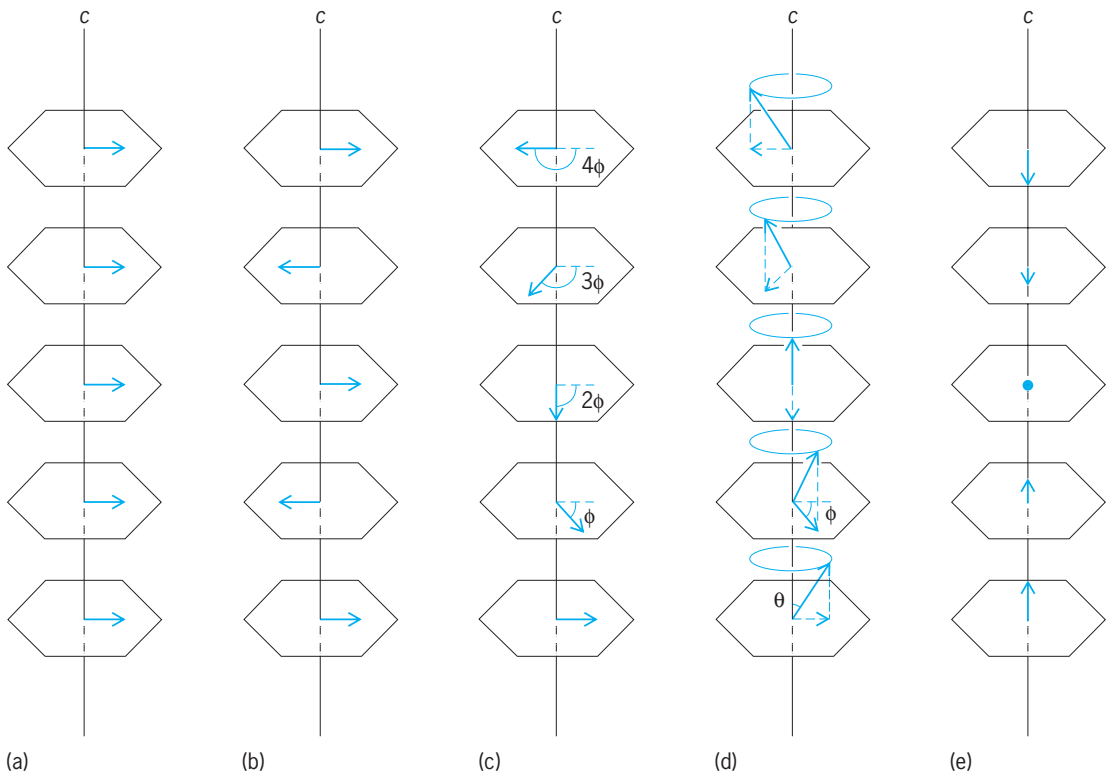

Some magnetically ordered arrays. Shown are directions taken by magnetic moments of consecutive planes along the *c* axes of hexagonal structures. (*a*) Simple ferromagnet; (*b*) simple antiferromagnet; (*c*) nonconical helimagnet of helical angle ϕ; (*d*) ferromagnetic conical helix of cone angle θ and helical angle ϕ; (*e*) sinusoidal antiferromagnet. In *a*, *b*, and *c* the moments all lie in the planes. In *d* there is a net ferromagnetic component along the *c* axis and in *e* there is an oscillating moment along the *c* axis.

Some representative helimagnets

Substance	Magnetic structure	Temperature, K (°F)
MnO	Nonconical helix	$0 < T < 84$ ($-460 < t_F < -308$)
$MnAu_2$	Nonconical helix	$0 < T < 363$ ($-460 < t_F < 194$)
Dy	Nonconical helix	$85 < T < 179$ ($-307 < t_F < -137$)
	Ferromagnet	$0 < T < 85$ ($-460 < t_F < -307$)
$MnCr_2O_4$	Simple ferrimagnet	$18 < T < 43$ ($-427 < t_F < -382$)
	Complex conical helix	$0 < T < 18$ ($-460 < t_F < -427$)
Er	Conical helix	$0 < T < 20$ ($-460 < t_F < -424$)
	Complex oscillation	$20 < T < 53$ ($-424 < t_F < -364$)
	Sinusoidal antiferromagnet	$53 < T < 85$ ($-364 < t_F < -307$)

10 K (−442°F), orders into two antiferromagnetically coupled conical helices. Frederic Keffer

Bibliography. B. Barbara, D. Gignoux, and C. Vettier, *Lectures on Modern Magnetism*, 1988; D. Craik, *Magnetism*, 1995; D. C. Jiles, *Introduction to Magnetism and Magnetic Materials*, 2d ed., 1998; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996.

Helioporacea

An order of the cnidarian subclass Alcyonaria (Octocorallia). The Coenothecalia have no spicules but form colonies with a massive skeleton composed of fibrocrystalline argonite fused into lamellae. The skeleton is perforated by both numerous wide cylindrical cavities occupied by the polyps, and narrow ones containing the solenial systems. In the calyx of the polyp, septalike structures of stony coral, or pseudosepta, are formed. The order includes a few genera, of which *Heliopora*, or the blue coral, is often found on coral reefs. *See* OCTOCORALLIA (ALCYONARIA); REEF. Kenji Atoda

Helioseismology

A technique for probing the interior of the Sun, using methods akin to terrestrial seismology. The Sun, although the nearest star by far, is a typical star, so what can be learned of its interior through helioseismology is of broad importance to the stars in general.

Solar waves. Like terrestrial seismology, helioseismology entails the analysis of many "seismic" wave

modes to determine the structure of the interior. However, although terrestrial seismic waves are initiated by a singular event such as an earthquake, waves within the Sun are continuously excited, probably by the turbulent convective motions in its outer layers. Thus the solar waves are always present at all points within the Sun and on its surface. The Sun is "ringing" like a bell, but not like one struck by a clapper; it vibrates more like a bell suspended in a sandstorm, continuously struck by tiny grains of sand. *See* SEISMOLOGY.

The solar waves are seen at the surface as up-and-down motions of the gases with a speed of about 0.3 mi/s (0.5 km/s) and a vertical displacement of about 30 mi (50 km). These waves are detected through the Doppler shift of the wavelength of absorption lines in the solar spectrum. They have periods clustering near 5 min (that is, with a frequency of one cycle in 5 min or about 0.003 cycle per second). As a result, the solar surface undulates up and down in a so-called five-minute oscillation (**Fig. 1**). The oscillation is actually the superposition of as many as 10^7 individual modes of oscillation of the Sun as a whole, where each mode has its own characteristic frequency (near, but not exactly at, 0.003 cycle per second) and spatial pattern on the solar surface.

Precise observations of the solar oscillations are difficult. The individual oscillation modes have velocities at the solar surface of only 12 in./s (30 cm/s) or less, and only extremely sensitive and stable spectrographs can measure such small motions. In addition, a nearly continuous stream of data extending over days is needed to separate the many individual modes with nearly identical oscillation frequencies. Ground-based observations are hampered because the day-night cycle allows observations of only 10 h or so at most sites. This restriction has been overcome by making observations from near the South Pole during the austral summer, and through networks of similar telescopes spaced at several longitudes around the globe (**Fig. 2**). Another very successful method is to obtain observations from a spacecraft located in an orbit experiencing continuous sunlight, such as the *Solar and Heliospheric Observatory* (*SOHO*) spacecraft, which orbits about the stable L1 Lagrange point, lying between the Earth and the Sun about 9×10^5 mi (1.5×10^6 km) from the Earth.

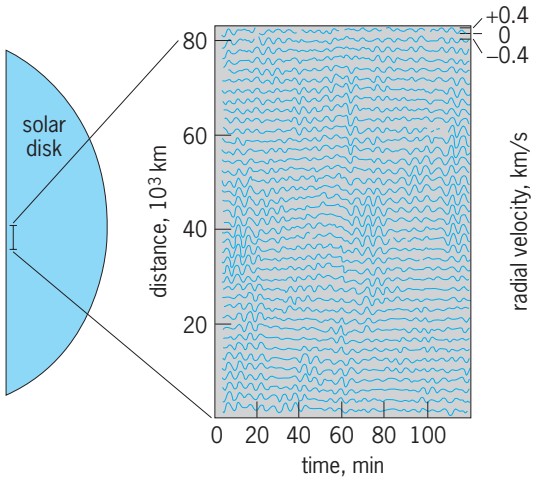

Fig. 1. Variation of radial velocity with time along an 80,000-km (50,000-mi) line on the solar surface. The 5-min oscillations, visible as wave packets localized in space and time, are actually the superposition of millions of discrete oscillation modes. 1 km = 0.6 mi. (*Adapted from S. Musman and D. Rust, Vertical velocities and horizontal wave propagation in the solar photosphere, Solar Phys., 13:261–286, 1970*)

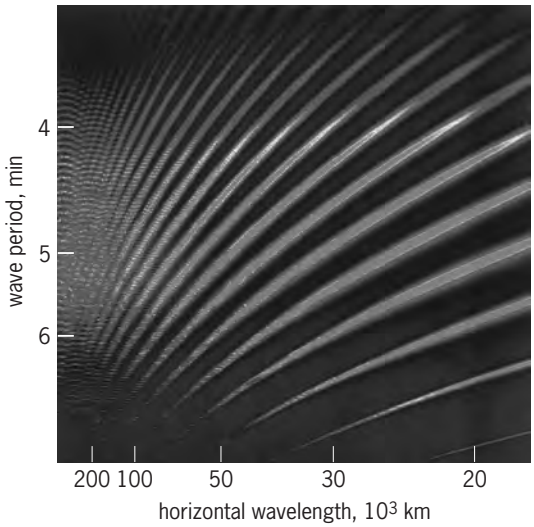

Fig. 2. Space-time spectrum of solar oscillations, from continuous observations obtained by the Global Oscillations Network Group (GONG) network of six solar telescopes spaced in longitude around the Earth. Ridges of power show how oscillation periods vary with horizontal wavelength; from lower right to upper left, each ridge is a higher overtone. 1 km = 0.6 mi. (*National Optical Astronomy Observatories*)

Each solar oscillation mode is produced within a specified region of the interior. The top of this resonant cavity is near the solar surface, which reflects the waves downward when they strike the surface from below. The bottom occurs at the depth where upward refraction reverses the propagation direction of descending waves. Waves with long horizontal wavelength, which travel nearly straight downward into the Sun, penetrate deeply, whereas waves with short horizontal wavelength remain near the surface. Within each cavity there can be both a lowest mode of oscillation and many higher-frequency overtones, or harmonics. Thus, a two-dimensional spectrum, displaying the oscillation power as a function of frequency and horizontal wavelength, shows a number of "ridges" of power, each corresponding to a separate overtone number (Fig. 2). *See* HARMONIC (PERIODIC PHENOMENA); MODE OF VIBRATION; VIBRATION.

Structure of solar interior. It is possible to predict the frequencies of oscillation modes in the Sun, given a model of the interior, that is, a numerical tabulation of temperature, density, pressure, and composition as a function of distance from the surface. When the oscillation frequencies predicted by the standard model in general use before the advent of helioseismology were compared with the observed

frequencies, small but significant discrepancies were found.

The discrepancies were lessened when models allowed for a higher initial abundance of helium in the core than first thought, as well as a decrease over the Sun's lifetime in the helium abundance near the surface due to gravitational sinking of helium relative to the lighter hydrogen atoms. The models pinpointed the bottom of the convection zone, at a depth of about 125,000 mi (200,000 km), or about 30% of the solar radius below the surface. They also suggested that the opacity of the solar material just below the bottom of the convection zone was larger than predicted; this was subsequently verified by revised opacity calculations. New calculations of the equation of state relating pressure to temperature and density throughout the interior led to still better agreement between models and helioseismology data. The most significant remaining difficulty is that the models predict a flux of solar neutrinos from the core which is at least twice that observed. It is impossible to adjust the structure so as to yield the observed neutrino flux without producing serious disagreement with helioseismology data, using standard physical descriptions of neutrinos. A possible resolution of this problem is that models of neutrinos as massless particles may be incorrect, and that instead they have a slight amount of mass, thereby causing neutrinos produced in the core to be changed into nondetectable form before reaching detectors at the Earth. *See* NEUTRINO; SOLAR NEUTRINOS.

Rotation in solar interior. Helioseismology offers insight into the structure of the solar interior and also into its rotation. Waves propagating with or against the direction of rotation are carried by it, and their effective propagation speed and frequency are increased or decreased. The frequency shift for any mode depends on the average rotation rate within the resonant cavity for that mode, and comparison of the shift for many modes with different cavities makes it possible to determine how the rotation varies with depth.

The surface of the Sun has long been known to rotate differentially with latitude; that is, at the Equator the surface rotation period is about 25 days while near the Poles it is about 34 days. Helioseismology provides a description of how the rotation varies with depth, down to and below the bottom of the convection zone which constitutes the outer 30% of the solar radius (**Fig. 3**). Roughly speaking, the increase of rotation period from Equator to Pole persists throughout the convection zone. However, at all latitudes the rotation period decreases slightly over the outer 10% of the solar radius, and then increases again to approximately its surface value at the bottom of the convection zone. At the bottom of the convection zone there is an abrupt transition to a deeper interior, which seems to rotate nearly uniformly and at the same speed as surface latitudes of about 35°. That is, the rotation period increases rather sharply across this transition zone near the Equator, and decreases sharply near the Poles. This zone of radial shear, called the tachocline, may be the

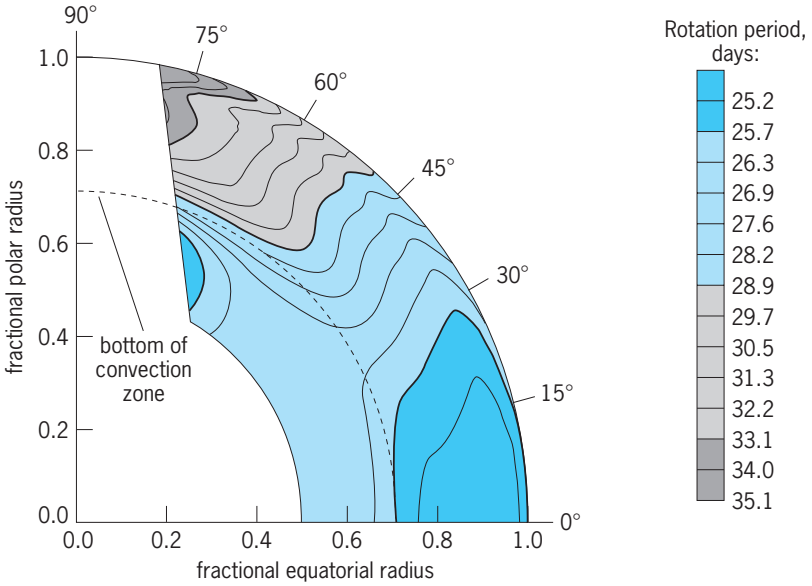

Fig. 3. Contour plot showing lines of constant rotation period within the outer 50% of the Sun, based on data from the Solar Oscillations Investigation (SOI) on the *Solar and Heliospheric Observatory (SOHO)* spacecraft. Tints refer to periods in days given by the key. Light color below the convection zone (broken line) indicates almost uniform rotation of the interior. Solar latitude is indicated from the Equator (0°, horizontal axis) to the Pole (90°, vertical axis). White areas represent regions inaccessible to current helioseismology measurements. (*S. Korzennik, SOI team; NASA*)

location of the solar magnetic dynamo, which creates the 22-year cycle of sunspot activity. *See* STELLAR ROTATION; SUN.

Robert W. Noyes

Bibliography. J. W. Leibacher et al., Helioseismology, *Sci. Amer.*, 253(3):48–57, September 1985; Special section on helioseismology, *Science*, 272:1281–1309, 1996.

Heliozoia

A subclass of the Actinopodea. Unlike Radiolaria, these protozoans have no central capsule. Most species live in fresh water. Pseudopodia may be either slender with an axial filament surrounded by cytoplasm (axopodia) or filamentous (filopodia). Axopodial filaments can be extended or retracted rapidly by mechanisms not yet explained. Certain floating species can roll along on the tips of their axopodia and also swim by moving their axopodia. Some species are naked; others have skeletal elements ranging from siliceous scales or spicules embedded in a gelatinous capsule to a reticulate chitinous skeleton often impregnated with silica. A centroplast may or may not be present. The subclass has three orders. Actinophryida lack skeletons and centroplasts and include both uninucleate and multinucleate genera (**illus.** *a*). In the Centrohelida a centroplast is present or else is assumed to be because the nucleus is eccentric (illus. *b*). The organisms may be covered with a gelatinous sheath or with a skeleton composed of discrete siliceous elements embedded in such a layer. The Desmothoracida have a perforate continuous skeleton which is composed of chitin, and sometimes impregnated with minerals

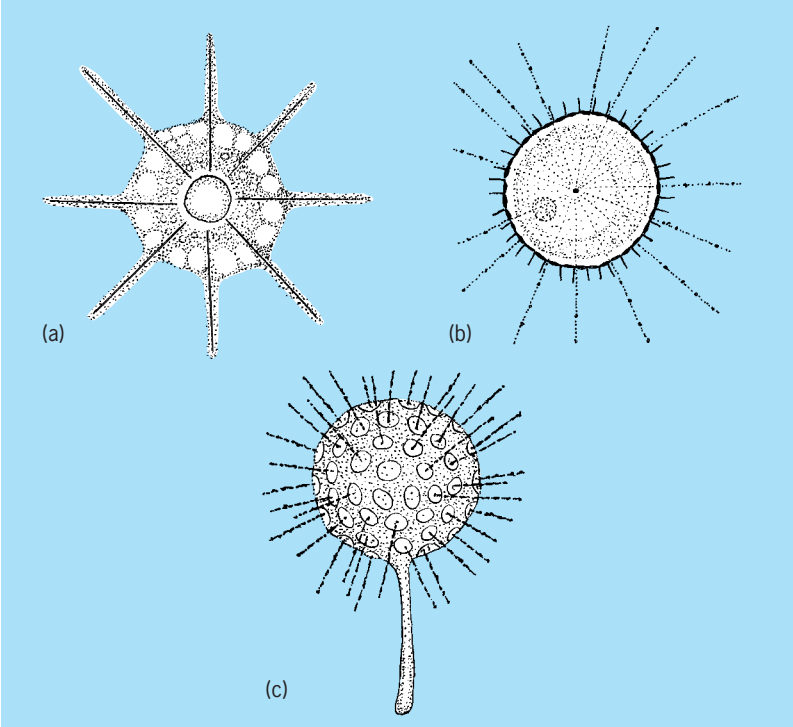

Organization of Heliozoia. (*a*) Actinophryida type. (*b*) Centrohelida type. (*c*) Desmothoracida type. (*After R. P. Hall*, *Protozoology*, *Prentice-Hall*, *1953*)

(illus. *c*). *See* ACTINOPHRYIDA; ACTINOPODEA; CENTROHELIDA; DESMOTHORACIDA; PROTOZOA; SARCODINA; SARCOMASTIGOPHORA. Richard P. Hall

Helium

A gaseous chemical element, He, atomic number 2 and atomic weight 4.0026. Helium is one of the noble gases in group 18 of the periodic table. It is the second lightest element. The world's chief source of helium is a group of natural gas fields in the United States. *See* INERT GASES; PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Helium is a colorless, odorless, and tasteless gas. It has the lowest solubility in water of any known gas. It is the least reactive element and forms essentially no chemical compounds. The density and the viscosity of helium vapor is very low. Thermal conductivity and heat content are exceptionally high. Helium can be liquefied, but its condensation temperature is the lowest of any known substance. The properties of helium are given in the **table**. *See* LIQUID HELIUM.

Helium was first used as a lifting gas in balloons and dirigibles. This use continues for high-altitude research and for weather balloons. The principal use of helium is in inert gas–shielded arc welding. The greatest potential for helium use continues to emerge from extreme-low-temperature applications. Helium is the only refrigerant capable of reaching temperatures below 14 K (−434°F). The chief value of ultralow temperature is the development of the state of superconductivity, in which there is virtually zero resistance to the flow of electricity. Other helium applications include use as a pressurizing gas in liquid-fueled rockets, in helium-oxygen breathing mixtures for divers, as a working fluid in gas-cooled nuclear reactors, and as a carrier gas for chemical analysis by gas chromatography.

Terrestrial helium is believed to be formed in natural radioactive decay of heavy elements. Most of this helium migrates to the surface and enters the atmosphere. The atmospheric concentration of helium (5.25 parts per million at sea level) could be expected to be higher. However, its low molecular weight permits helium to escape into space from the upper atmosphere at a rate roughly equal to its formation. Natural gases contain helium at concentrations higher than in the atmosphere. Arthur W. Francis

Helium is an element with a closed electronic shell, a large ionization potential, and a low polarizability, which makes it a very unlikely candidate to form chemical bonds. However, solid helium compounds have been found to form at high pressure, one with nitrogen [$He(N_2)_{11}$] and one with neon [$Ne(He)_2$]. These compounds belong to a class known as van der Waals compounds. *See* INTERMOLECULAR FORCES.

Other helium compounds have also been observed in a clathrate hydrate, $He(H_2O)_{6+\delta}$, and helium has been detected inside the carbon molecule buckminsterfullerene (C_{60}), forming HeC_{60}. Mixtures of helium and other components prevail under conditions of high pressure in the outer planets of the solar system and their satellites. Therefore, it is believed that helium compounds are important in the modeling of the interiors of such celestial bodies. The formation of helium compounds at high pressures illustrates that under such conditions different chemical behavior occurs compared to that observed under ambient conditions. *See* CHEMICAL BONDING; CLATHRATE COMPOUNDS; FULLERENE. Willem L. Vos

Helium-3 is a rare stable isotope of helium was discovered by L. W. Alvarez and R. Cornog in 1939. Its concentration in nature is so low, approximately one part per hundred million in well helium, that it was 1951 before sufficient quantities of pure gas became available for experimentation. The gas was then, and continues to be, obtained as a by-product from the decay of tritium, the heavy isotope of hydrogen. Tritium is produced in a nuclear reactor from the reaction between lithium and a neutron.

Properties of helium

Property	Value
Atomic number	2
Atomic weight	4.0026
Melting point* at 25.2 atm pressure	−272.1°C (1.1 K)
Triple point (solid, helium I, helium II)	−271.37°C (1.78 K)
Triple point = λ-point (helium gas, helium I, helium II)	−270.96°C (2.19 K)
Boiling point at 1 atm pressure	−268.94°C (4.22 K)
Gas density at 0°C and 1 atm pressure, g/liter	0.17847
Liquid density at its boiling point, g/ml	0.1249
Solubility in water at 20°C, ml helium (STP)/1000 g water at 1 atm partial pressure of helium	8.61

*The melting point varies with the pressure.

The ^{3}He nucleus is composed of two protons and one neutron, one fewer than for ^{4}He; as a consequence, ^{3}He is a fermion whereas ^{4}He is a boson. The two isotopes are the exemplars of Fermi-Dirac and Bose-Einstein systems, respectively. It is principally for this reason that helium, an apparently featureless chemical element, has been studied intensively. *See* BOSE-EINSTEIN STATISTICS; FERMI-DIRAC STATISTICS; TRITIUM.

Bernard M. Abraham

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., 2004; M. Ozima and F. A. Podosek, *Noble Gas Geochemistry*, 2001; J. Wilks and D. S. Betts, *An Introduction to Liquid Helium*, 2d ed., 1987.

Helix

Any nonplanar curve all of whose tangents make the same angle with a fixed line. Other characteristic properties are that all principal normals are parallel to a plane and that the ratio of torsion to curvature is constant. If a helix has constant curvature (and hence constant torsion), it is a circular helix; it lies on a circular cylinder whose elements it cuts at a constant angle (see **illus.**). Parametric equations for a circular helix (the curve of an untapered screw) are $x = a \cos t$, $y = a \sin t$, $z = kt$, $a > 0$, $k \neq 0$, and $-\infty < t < \infty$. *See* ANALYTIC GEOMETRY; DIFFERENTIAL GEOMETRY; PARAMETRIC EQUATION.

Leonard M. Blumenthal

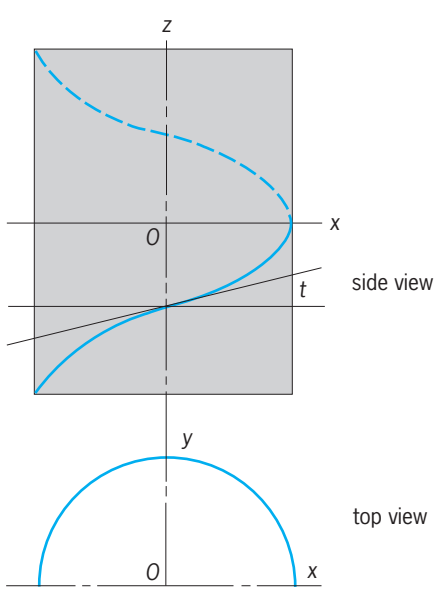

Diagrams of a circular helix.

Helmholtz coils

A pair of flat circular coils of small cross section, with equal numbers of turns and equal diameters, arranged with a common axis, and connected in series to have a common current (**Fig. 1**). The purpose of the arrangement is to obtain a magnetic field H_{AB} that is more nearly uniform than that of a single coil (**Fig. 2**) without the use of a long solenoid. *See* SOLENOID (ELECTRICITY).

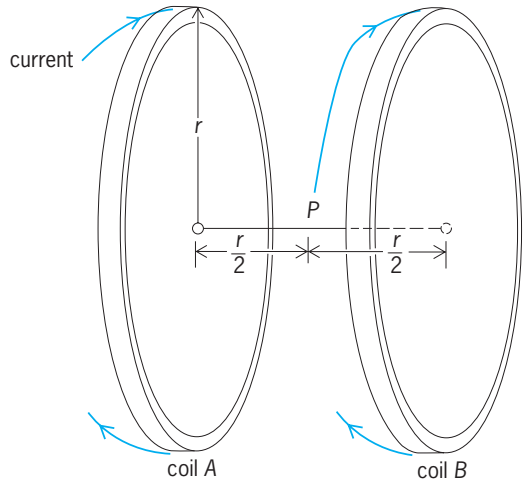

Fig. 1. Helmholtz coils, arranged with common axis and connected in series. (*After L. B. Loeb, Fundamentals of Electricity and Magnetism, 3d ed., Wiley, 1947*)

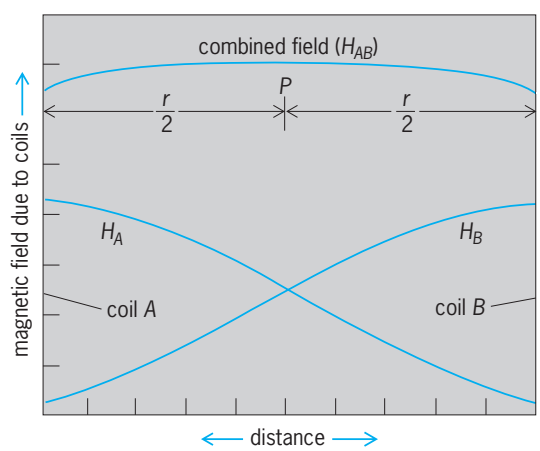

Fig. 2. Separate and combined fields of Helmholtz coils. Note the region of constant resultant field at point *P*. (*After L. B. Loeb, Fundamentals of Electricity and Magnetism, 3d ed., Wiley, 1947*)

The optimum arrangement is that in which the distance between the two coils is equal to the radius of one of the coils (Fig. 1). For this arrangement, the variation of the field strength near the center of the apparatus is a minimum, and the field is nearly uniform near the center. The field is the sum of the fields produced there by the individual coils. More complicated systems having more coils of finite cross section can be computed to give equal or better field uniformity and greater efficiency. Kenneth V. Manning

Bibliography. F. J. Bueche, *Introduction to Physics for Scientists and Engineers*, 4th ed., 1986; D. Halliday, R. Resnick, and K. Krane, *Physics*, 5th ed., 2002; W. H. Hayt and J. A. Buck, *Engineering Electromagnetics*, 7th ed., 2005; W. T. Scott, *The Physics of Electricity and Magnetism*, 2d ed., 1966, reprint 1977.

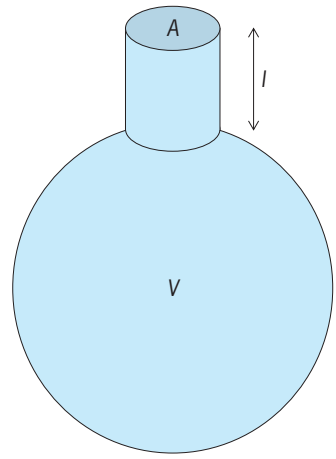

Helmholtz resonator. Symbols are explained in text.

Helmholtz resonator

An acoustic device consisting of a rigid cavity with one or more small openings. When exposed to sound or vibration, a Helmholtz resonator responds most strongly near a characteristic resonance frequency governed by the dimensions of its cavity and openings. A common example is a bottle or jug that produces sound near its resonance frequency when air is blown over the neck's opening. The Helmholtz resonator is named after the physicist and physiologist Hermann von Helmholtz, who constructed spherical glass resonators with small openings at opposite ends to analyze the frequency content of sounds, including musical tones.

Characteristics. The classical description of the Helmholtz resonator, as derived by Lord Rayleigh, assumes that all dimensions of the resonator are small compared to one-quarter wavelength of its resonance frequency, and that all sound amplitudes are small enough to allow linear acoustic analysis. In this case, the resonator's state can be well described by a uniform cavity pressure and a volume flow rate of air through the cavity openings.

At resonance, the system's oscillating potential energy, associated with compression and expansion of air in the cavity, is exactly balanced by the oscillating kinetic energy of air flowing through the openings. The resulting oscillations can be large compared to the externally applied sound or vibration, and can continue for some time after the external force is removed. This phenomenon occurs at the resonance frequency, also called the natural frequency, which is given for a Helmholtz resonator with a single opening (see **illustration**) by Eq. (1), where c is the speed

$$f_0 = \frac{c}{2\pi}\sqrt{\frac{A}{l' V}} \tag{1}$$

of sound for air (or other fluid employed), V is the cavity volume, A is the surface area of the opening, and $l' = l + \delta l$ is an effective neck length that includes both the true neck length l and an "end correction" describing an additional volume of fluid constrained to flow into or out of the openings. For an opening small compared to the resonator size, the end correction is approximated by $\delta l = 0.96\sqrt{A}$ for an outer opening within a large plane, and $\delta l = 0.48\sqrt{A}$ for an opening at the end of a long neck in free space.

Energy loss in Helmholtz resonators is associated both with friction of the fluid flowing through the resonator opening and with sound radiation into the surrounding air. These energy loss mechanisms affect the quality factor Q of the resonator, which is defined by Eq. (2), where f_1 and f_2 are the frequencies where the resonator response (as measured by its radiation or absorption of sound pressure) is one-half that at the resonance frequency f_0. Low damping results in a high Q value, corresponding to high-amplitude resonator response within a narrow frequency range around resonance. *See* Q (ELECTRICITY).

$$Q = \frac{f_0}{|f_1 - f_2|} \tag{2}$$

The behavior of a Helmholtz resonator is analogous to a lumped-element harmonic oscillator with a mass element $M = \rho A l'$, a spring element $K = \rho c^2 A^2/V$, and a damping or dashpot element R, where ρ is the mass density of the air or other fluid. The vibration response of this analog system is greatest near its resonance frequency, given by Eq. (3). Similarly,

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{K}{M}} \tag{3}$$

the Helmholtz resonator is analogous to a series inductor-resistor-capacitor (LRC) circuit with the natural frequency given by Eq. (4). Modified lumped-element models can incorporate deviations from the standard Helmholtz resonator configuration, such as flexible cavity walls in the case of intracranial aneurysms. *See* HARMONIC MOTION; HARMONIC OSCILLATOR; RESONANCE (ACOUSTICS AND MECHANICS); RESONANCE (ALTERNATING-CURRENT CIRCUITS).

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{1}{LC}} \tag{4}$$

If the resonator dimensions are not small compared to one-quarter of the wavelength $\lambda_0 = c/f_0$, the resonator can no longer be accurately regarded as a lumped-element system. In such cases, more detailed analysis, incorporating the spatial distribution of fluid pressure and velocity, is required to fully describe the resonator behavior. For example, a pipe open on one end, such as an organ flue pipe, has additional resonances at frequencies $f_0 = nc/(4l)$, where l is the effective pipe length and $n = 1, 3, 5, \ldots$, is any positive odd integer. *See* SOUND; VIBRATION.

Applications. Helmholtz resonance describes the lowest-frequency oscillation mode of many acoustic systems, including vehicles or dwellings with open windows, low-frequency loudspeakers, and the bodies of stringed musical instruments. Resonance may be induced by an external sound field, mechanical vibration, or an external flow field interacting with the cavity opening. Depending on the application, this resonance may be either a desired feature to be exploited or a nuisance to be controlled.

Vases acting as Helmholtz resonators were used by ancient Greek and Roman architects to modify the acoustic response of amphitheaters, as mentioned in Vitruvius' *De Architectura* (ca. 27 B.C.), and are also found in medieval European churches. In modern architecture, as well as acoustic filter and muffler design, Helmholtz resonators limit the transmission or reverberation of low-frequency sound that can be difficult to control by other means. To absorb sound over a wide frequency band, a collection of resonators with appropriately chosen resonance frequencies and quality factors can be employed. *See* MUFFLER.

T. Douglas Mast

Bibliography. H. L. F. Helmholtz, *On the Sensations of Tone*, 2d English ed., transl. by A. J. Ellis, Dover, New York, 1954; U. Ingard, On the theory and design of acoustic resonators, *J. Acous. Soc. Amer.*, 25:1037–1061, 1953; M. Long, *Architectural Acoustics*, Elsevier Academic, Burlington, MA, 2006; J. W. Strutt, Lord Rayleigh, *The Theory of Sound*, Dover, New York, 1945.

Hematologic disorders

Those disorders marked by aberrations in structure or function of the blood cells or the blood-clotting mechanism. Although many other diseases may be reflected by the blood and its constituents, the abnormalities of red cells, white cells, platelets, and clotting factors are considered to be primary hematologic disorders.

Erythrocytic abnormalities. Red-cell abnormalities are principally represented by the anemias and polycythemias. The anemias are marked by a decrease in the hemoglobin concentration, and may be due to blood loss or decreased production or excessive destruction of red cells. They are considered separately. *See* ANEMIA.

Polycythemias are disorders characterized by an increase in the numbers of circulating red cells and usually by a concomitant increase in hemoglobin. Relative polycythemia is really the result of hemoconcentration in which there is fluid loss from the blood, with corresponding increases in proportions of cellular elements. It is seen in excessive vomiting, diarrhea, dehydration, and similar conditions. *See* HEMOGLOBIN.

Secondary polycythemias result from a compensatory increase in the formation of red cells following hypoxia of the bone marrow. This condition is most often associated with chronic conditions, such as heart or lung diseases and prolonged exposure to high altitudes. The blood volume may be considerably expanded, and the individual may present a typically dusky red appearance. There are rare familial types due to synthesis of an abnormal hemoglobin with high oxygen affinity; this leads to decreased release of oxygen to the tissues, which in turn stimulates an increased production of red cells.

Primary polycythemia or polycythemia vera is a chronic and ultimately fatal disease of middle and old age in which there is a gradual increase in the number of red cells and usually an increase in the number of platelets and leukocytes. The viscosity of the blood is increased, which tends to slow the rate of blood flow; this factor, together with the high platelet count, predisposes to the thromboses commonly seen in the condition. Paradoxically there is also a hemorrhagic tendency, and the subject may bleed from the nose, stomach, and elsewhere. The cause of the disease is unknown, and although more benign, it has been likened to leukemia. The mainstay of treatment is repeated bloodlettings, supplemented when necessary by bone-marrow-suppressing agents such as radioactive phosphorus or hydroxyurea. *See* LEUKEMIA.

Leukocytic abnormalities. In a wide variety of conditions the many forms of white cell present in the circulation, bone marrow, and lymphoid tissues of the body may be altered in form or number.

The suffixes -philia and -penia denote increases and decreases, respectively, for the cells named. Leukopenia, neutrophilia, eosinophilia, and pancytopenia are examples of the wide range of possibilities. The absolute or relative increases or decreases of one or more types of leukocytes are often characteristic of certain disease states; for example, in infections neutrophils are increased both in absolute numbers and in relation to the other white cells. There may also be changes in the proportions of the different cells, and immature or atypical forms may be present.

The leukemias represent a special kind of malignancy in which there is usually an uncontrolled proliferation of one or more types of leukocytes, often reflected by great increases in the white cell count of the peripheral blood. The leukemias are generally classified, on clinical grounds, as acute, subacute, or chronic. They can also be subclassified further on the basis of cell-marker, karyotypes, or other studies. In addition, they are named to indicate the particular cell which shows neoplastic characteristics. The most common forms are lymphocytic and

myelogenous leukemia. In the former, large numbers of abnormal lymphocytes may be present in the blood; in the latter, the cells of the granulocyte series are involved. Other varieties exist, such as monocytic, eosinophilic, and plasma-cell leukemia, but they are much less frequently encountered.

Many similarities exist in the clinical course and pattern of lesions seen in the leukemias. In the acute forms, the onset is abrupt, with the appearance of fever, malaise, and weakness. The course in the absence of treatment is usually rapid, with susceptibility to infections and a tendency toward hemorrhage accounting for many complications or leading to death. The outlook for acute lymphocytic leukemia, which is the most common type seen in children, has improved dramatically: long-term remissions and probable cures following treatment are now seen in more than 50% of children and nearly 20% of adults.

Chronic leukemias may be nearly acute to almost completely benign. In general, the granulocytic forms tend to be more severe than the lymphocytic forms. The first evidence of chronic lymphatic leukemia may be discovered more or less by accident, by periodic physical examination, or by routine blood counts. Although more severe forms are seen, the course often extends over many years or even decades, and symptoms attributable to the leukemia may be very late in developing. On the other hand, chronic myelogenous leukemia (chronic granulocytic leukemia) is rarely discovered accidentally and, when so discovered, signs and symptoms attributable to the leukemia soon develop. It may progress into an acute form. Symptoms of the chronic leukemias include fatigue, lack of exercise tolerance, pallor, and weight loss.

The treatment of the leukemias is based on the type, the stage at which it is diagnosed, and the age of the patient. The most common type of acute lymphocytic leukemia in children responds to complex and novel treatment methods that include the use of a combination of two or more chemotherapeutic agents together with steroids. More than half the children will be free of symptoms 10 years after therapy ends. The same treatment has been used for the other types of childhood acute lymphocytic leukemia and all the adult forms, but relatively few respond well. High-dose chemotherapy followed by marrow transplantation is recommended for patients under 50 who have an HLA-compatible sibling. Although this form of therapy is not without great risk, cures have resulted. The prognosis for chronic lymphocytic leukemia that is discovered before signs and symptoms are present is excellent, and the mean survival time is approximately 7 years; survival 20 years or longer is not at all unusual. When symptoms or complications are present, the disease is usually treated with prednisone and an alkylating agent such as chlorambucil. Most patients with chronic myelogenous leukemia respond well initially to therapy with an alkylating agent such as busulfan or hydroxyurea, but relapses are typical and the mean survival time of approximately 3 years is only slightly increased by treatment. Most forms of the acute nonlymphocyctic leukemias are treated with chemotherapeutic agents, and while some show a good initial response, the prognosis is very poor. Bone marrow transplantation as described above for the treatment of acute lymphocytic leukemia has been recommended for the treatment of nonlymphocytic leukemia and also for chronic myelogenous leukemia.

Hemorrhagic disorders. The hemorrhagic disorders result from a large number of known and unknown causes or contributing factors, often of a diverse nature. The initial or primary events that halt bleeding from a very small wound are the formation of a platelet plug which seals the hole in the vessel wall and arteriolar vasoconstriction. The secondary event is the subsequent fortification of the plug by fibrin. The stimulus that causes the platelets to stick together and form the plug is believed to be exposure to substances released from the damaged blood vessels which act in concert with certain plasmatic factors, such as adenosine diphosphate (ADP) released from damaged red cells and the von Willebrand factor. A decrease or abnormality of the von Willebrand factor, a reduction in the number of platelets, a qualitative defect in platelets, or a defect in the vascular wall can all result in failure of the primary hemostatic mechanism, with spontaneous bleeding; this is referred to as purpura. *See* HEMORRHAGE.

Blood vessel damage may occur as a result of direct or indirect damage by microorganisms during infections, as the result of vitamin C deficiency (scurvy), and following hereditary defects in blood vessel development. A number of comparatively rare disorders, usually with a hypersensitivity component, also fall into this category. There are several forms of thrombocytopenia, all of which are characterized by a decrease in thrombocytes or platelets in the circulation. Qualitative changes in the platelets can result in impairment of their function, and may be inherited or acquired, as is seen following aspirin ingestion. Since the platelets are involved in primary plug formation, such deficiencies or abnormalities lead to purpura. Treatment is dependent on the cause and usually is restricted to those occasions when bleeding is severe or life-threatening or when surgery is about to be performed. Thus if the platelet count is very low or the platelets are defective, transfusion of platelets may be indicated. Most forms of von Willebrand's disease are now treated with desmopressin; however, in severe cases cryoprecipitate may be indicated.

Fibrin or clot formation is the result of the sequential or stepwise interaction of platelets and several blood-clotting factors which are proteins; tissue factor is also involved. A defect in fibrin formation gives rise to the type of bleeding seen in hemophilia, in which the primary hemostatic plug is formed normally but breaks down several hours or days later owing to lack of adequate fortification by fibrin. *See* HEMOPHILIA.

Defects in the clotting process that result in

excessive bleeding may be due to an acquired or hereditary deficiency or abnormality of a clotting factor, as in hemophilia A or B. The bleeding that may be associated with hereditary conditions is usually controlled by the intravenous administration of concentrates of the deficient factor, which are obtained from plasma. For example, factor VIII concentrates are used in hemophilia A, factor IX concentrates in hemophilia B, and cryoprecipitate in fibrinogen deficiencies. Fresh-frozen plasma is usually used to treat those conditions for which concentrates are not available. Acquired deficiencies of clotting factors are seen in many conditions, including liver disease, vitamin K deficiency, and disseminated intravascular clotting, but they may be caused by the spontaneous development of antibodies directed against one of the clotting factors. The bleeding associated with liver disease may be treated with fresh-frozen plasma and platelet transfusions; the treatment of disseminated intravascular clotting is directed at the primary cause of the condition. The bleeding that results from an antibody is treated with infusions of the clotting factor against which the antibody is targeted, immunosuppressives, and steroids. Sometimes attempts are made to bypass the action of the antibody by intravenous administration of an activated clotting factor. *See* BLOOD. Cecil Hougie

Bibliography. D. L. Bernard, *Clinical Hematology*, 1989; B. Brown, *Hematology: Principles and Practice*, 1993; C. Kjeldsberg et al., *Practical Diagnosis of Hematologic Disorders*, 3d ed., 2000; G. R. Lee et al., *Wintrobe's Clinical Hematology*, 10th ed., 1999; W. J. Williams et al., *Hematology*, 4th ed., 1990.

Hematopoiesis

The process by which the cellular elements of the blood are formed. Blood contains many free-floating cells which are moved throughout the body within the blood vessels. The three main types of cells are the red cells (erythrocytes), which serve to carry oxygen, the white cells (leukocytes), which function in the prevention of and recovery from disease, and the thrombocytes, which function in blood clotting. In humans there is only one white cell in the blood for every 700 red cells.

The formation of these cells is one of the most active and important processes in the body. Most of the circulating cells live only for a short time and must be replaced in order to maintain life. For instance, in the human adult a red blood cell has a life of 120 days; 250 billion new red cells have to be produced daily to replace those that are destroyed. Serious consequences result if the balance of replacement and destruction of blood cells is impaired by disease, by genetic defects (as in hemolytic anemia), or by irradiation with x-rays or other high-energy radiations. The lymphocytes, a class of leukocytes, are more sensitive to ionizing radiation than any other cell type in the body, and changes in the number of lymphocytes may be used as a good index of radiation damage. *See* HEMATOLOGIC DISORDERS; RADIATION INJURY (BIOLOGY).

Blood Cell Production in Adults

Blood cells originate in the reticuloendothelial tissue, which is a loose, fibrous, highly vascularized mesh of fibers, endothelial cells, and macrophages. Within the spaces of the tissue are found the precursor (blast) cells of the definitive adult types. For the sake of convenience, the reticuloendothelial tissue is divided into two general but imprecise types: lymphoid and myeloid tissue. This classification certainly does not apply to the situation present in the embryo and in certain blood disorders.

Lymphoid tissue. This tissue is primarily localized in the lymph nodes of the lymphatic system and is also in the spleen, thymus, and bone marrow. Several classes of white cells are produced, including the lymphocytes, macrophages, and monocytes. These cells function in producing antibodies and in removing detritus formed during infectious diseases; organisms in which the lymphoid tissue is reduced or destroyed (such as after sustaining irradiation damage) usually succumb quickly to disease. There is evidence that many of the early blast cells of the lymphoid tissue arise in the embryonic thymus; they then migrate to their final location late in embryonic development. In juveniles and young adults the thymus may also produce a humoral factor which is important for the production of some differentiated cell types and antibodies by the lymphoid tissues. *See* CELLULAR IMMUNOLOGY; LYMPHATIC SYSTEM; THYMUS GLAND.

Myeloid tissue. This tissue is normally limited in humans to the red bone marrow of the ribs, sternum, vertebrae, and proximal ends of the long bones of the body. It is concerned with the production of the erythrocytes and certain types of leukocytes. The latter are the granular leukocytes (called eosinophiles, basophiles, and neutrophiles on the basis of the affinity of granules in their cytoplasm for certain dyes) and megakaryocytes. Fragments of megakaryocytes form the blood platelets (thrombocytes), which are necessary for blood clotting.

Blast differentiation. There are several theories concerning the origin of the blood cells, and they differ mainly in the number of types of blast cells thought to be involved in the process. For instance, the currently favored view holds that there is only one type of blast cell and that it is capable of giving rise to all other blood cells. This pluripotential blast cell is morphologically indistinguishable from a circulating lymphocyte, and may be called a hemocytoblast. It is found in both lymphoid and myeloid tissues. The hemocytoblast may give rise to lymphocytes, macrophages, megakaryocytes, monocytes, granular leukocytes, and erythrocytes, that is, to all the definitive circulating blood cell types. The hemocytoblasts probably are ultimately derived from the mesenchyme of the early embryo. Well-defined stages of the differentiation of the various cell types can be recognized by the hematologist using various stains.

Though the identification and characterization of the progenitor cells that give rise to the red and white cells of the blood are difficult, incontrovertible proof of the existence of pluripotential blast cells has been furnished. If a suspension of bone-marrow cells is injected into the circulation of a lethally irradiated host mouse, some of the injected cells colonize the host's spleen and give rise to localized nodules of either erythrocytes, granulocytes, megakaryocytes, or lymphocytes. These nodules are macroscopically visible and appear before the host dies. It has been shown that the cells that compose a nodule are a clone, derived from a single cell. It has become possible to culture cells from the marrow in very elaborate tissue culture media and to demonstrate cells that have many of the properties expected of precursor blast cells.

Erythropoiesis. More is known about the transformation of a blast cell into an erythrocyte than about any of the other types of blood cells. A hormone, erythropoietin, regulates the production of differentiated cells from the blast cell. Erythropoietin is thought to be produced primarily in the kidney and is essential, at least in higher vertebrates, for the formation of red cells. In physiological or pathological states where a surfeit of red cells exists, less erythropoietin is produced. When a great demand for red cells exists (as in certain anemias), the level of erythropoietin may be high. Colonies of red cells do not form in the spleens of lethally irradiated recipients transfused with bone-marrow cells from untreated donors when the hormone is lacking. The hormone is thought to regulate the production of erythroblasts, the earliest recognizable cells of the erythrocytic series. *See* SPLEEN.

The stages of the transformation of a blast into an erythrocyte are morphologically and functionally distinct. The erythroblasts are spherical cells with a basophilic cytoplasm. They continue to divide and give rise to various types of polychromatic erythroblasts; it is in these cells that hemoglobin is actively synthesized, giving a reddish tinctorial quality to stained cells. Furthermore, the basophilia of the cytoplasm continually decreases because cytoplasmic ribosomes are lost. Toward the end of the polychromatic phase, the nucleus condenses and no further synthesis of nuclear ribonucleic or deoxyribonucleic acid can be detected. In mammals the nucleus is subsequently extruded, but in all other vertebrates the nucleus is retained, although it may be functionally inactive. After the nucleus is extruded some synthesis of hemoglobin still goes on, and the cells are called reticulocytes because of the reticulum of basophilic material present in the hemoglobin-filled cytoplasm. Some reticulocytes may be released to the circulation. Finally all synthesis of hemoglobin in the bone marrow ceases, and the cells are released to the circulation as the definitive erythrocytes. In some anemias the differentiative process may occur very much faster, or some stages may be skipped, and nucleated (immature) cells may be released to the circulation. *See* HEMOGLOBIN.

Embryonic Origins of Blood Cells

Circulating blood cells can be detected very early in the development of vertebrates; they are present as soon as there is a circulatory system to transport them. In the embryo there is a succession of changing sites of primary production. The details may vary among the different vertebrates, but generally one can detect three sites of origin. The yolk sac produces the first blood cells; subsequently certain other organs (spleen, liver, venous sinusoids) become productive. Finally, the normal adult sites (primarily, the bone marrow and lymph nodes) assume their important role near the time of birth. The sequence of different hematopoietic centers in humans is shown in the **illustration**; this sequence comprises most of the features found in all vertebrates.

Yolk sac. Definite areas of the embryo are allocated very early to the production of the first blood cells. These areas, called blood islands, are located in the mesoderm of the yolk sac. Groups of the primordial precursors of blood cells coalesce to form stellate, interconnected groups of highly basophilic cells (hemocytoblasts). The central portions of the blood islands undergo erythropoiesis (although it may differ in detail from the sequence previously described in adult bone marrow), while the cells on the periphery of the interconnected islands flatten and become the endothelial lining of the blood vessels of the yolk sac. Only erythrocytes form during this first wave of hematopoiesis, and the hemoglobins contained in these red cells differ considerably in their amino acid sequence from the hemoglobins of the adult. *See* YOLK SAC.

Formation of definitive blood cells. The yolk sac usually forms red cells for only a relatively short time. As its production wanes, hematopoiesis becomes detectable, and then intense, in a variety of extramedullary sites. In humans the liver next assumes important hematopoietic functions, and the process here is similar to that in the bone marrow of the adult. Nonnucleated red cells are produced, which may (in mice) or may not (in humans) contain hemoglobin identical to the adult. Granular leukocytes and macrophages also appear. Eventually hematopoietic activity in the liver ceases, and other sites, such as the spleen, become active. Later

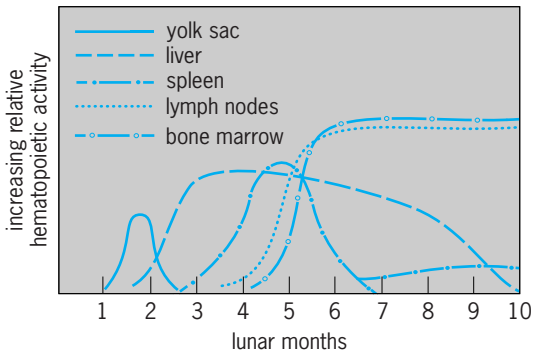

Hematopoietic activity in humans during the normal gestation period of 10 lunar months of 28 days each.

in development the lymph nodes form and become important centers of origin of lymphocytes. Near the time of birth, the bone marrow becomes active in producing red cells, granular leukocytes, and megakaryocytes.

There are many differences in the hematopoietic activity of the yolk sac compared with subsequent sites. For instance, in the human yolk sac, red cells are nucleated, large, and basophilic and contain an embryonic hemoglobin composed of two α-polypeptide chains (or special variants designated zeta) and two ϵ-polypeptide chains. The liver produces red cells that seem indistinguishable from those of the adult, but they contain a fetal hemoglobin composed of two α- and two γ-polypeptide chains. Eventually the bone marrow produces red cells containing the usual human adult hemoglobin (two α- and two β-polypeptide chains).

The regulation of the changing sites of hematopoiesis and their relation to the nature of the blood cells produced are still mysteries. *See* BLOOD; LIVER. Fred Wilt

Bibliography. D. W. Golde (ed.), *Hematopoiesis*, 1984.

Hemiascomycetes

A class of the phylum Ascomycota that includes the yeasts and yeastlike fungi. These are morphologically simple fungi; no ascoma is formed, and the asci are produced free on the host or substrate. Asexual reproduction occurs by the formation of blastospores (budding) or, less frequently, by fission arthrospores. Two main orders are recognized, the Saccharomycetales and the Taphrinales. *See* YEAST.

The vegetative body (thallus) of the Saccharomycetales may be either unicellular (true yeasts) or mycelial. In unicellular species, asci form when two vegetative cells fuse, and then the fused cell undergoes meiosis to form ascospores. In mycelial species, the hyphae are not very extensive. Sexual reproduction occurs when adjacent cells extend short lateral branches that fuse to form the asci. Variations on these modes of ascus formation, however, are common among the yeasts.

The Saccharomycetales are common on substrates high in sugars, such as plant exudates, ripe fruits, and flower parts. Because they are microscopic, they are recognized mainly from cultures that have a homogeneous appearance and a characteristic odor. The most important genus is *Saccharomyces; S. cerevisiae* is the common bakery and brewery yeast, and *S. ellipsoideus* is used in winemaking. An important mycelial species is *Nematospora coryli*, which causes yeast spot disease of various crops.

The order Taphrinales includes the leaf curl disease fungi. The most widely recognized species are *Taphrina deformans*, cause of leaf curl of peach and almond trees, and *T. caerulescens*, cause of leaf blister of oaks. These fungi produce a well-developed mycelium in the host tissue but, when grown in culture, form only a yeastlike colony of single cells—a phenomenon known as dimorphism. Asci are produced when special binucleate hyphal cells beneath the host cuticle undergo nuclear fusion and the resulting diploid cell elongates to form an ascus on the leaf surface. The nucleus undergoes meiosis, and ascospores are formed. *See* ASCOMYCOTA; EUMYCOTA; FUNGI.

On the basis of molecular data, some workers now propose separating the Taphrinales and the fission yeast, *Schizosaccharomyces*, into a new class, Archiascomycetes. These fungi are considered to be more primitive and phylogenetically basal to the rest of the ascomycetes. Richard T. Hanlin

Bibliography. C. J. Alexopoulos, C. W. Mims, and M. Blackwell, *Introductory Mycology*, 4th ed., John Wiley, New York, 1996; J. A. Barnett, R. W. Payne, and D. Yarrow, *Yeasts: Characteristics and Identification*, 1990; D. R. Berry, *Biology of Yeast*, 1982; N. J. W. Kreger-van Rij (ed.), *The Yeasts: A Taxonomic Study*, 3d ed., 1984; H. Nishida and J. Sigiyama, Archiascomycetes: Detection of a major new lineage within the Ascomycota, *Mycoscience*, 35:361–366, 1994.

Hemicellulose

The term adopted by E. Schulze in 1891 to designate the plant cell components which are made soluble by dilute alkali or which go into solution quite readily in hot dilute mineral acids with the formation of simple sugars. Hemicelluloses constitute about one-fourth of perennial plants and about one-third of annual plants. The term is usually applied to those polysaccharides in the cell wall of land plants which are extractable by dilute alkaline solutions. The term has also been used to include all the polysaccharide components of the cell wall other than cellulose. *See* CELL WALLS (PLANT).

How the hemicelluloses, which are often water soluble after extraction from plant tissue, are anchored in the cell wall so as to be unextractable by water is still a controversial question. As an explanation, hemicellulose-lignin or hemicellulose-cellulose covalent bonds have been postulated. Hemicelluloses extracted from different plant sources, although they often have many common characteristics, are rarely identical. In fact, many different hemicelluloses usually occur intermixed with each molecular type representing different degrees of polymerization. Because of this heterogeneity, few hemicelluloses have been isolated in a homogeneous state. Therefore, relatively little is known of the structure of these compounds that compose almost one-third of the carbohydrates in woody tissue.

Preparation. Hemicelluloses are generally prepared by alkaline extraction of dried, defatted plant tissue or a special delignified plant tissue which is called holocellulose. The latter is usually prepared by treatment of the defatted plant with chlorous acid solution, which solubilizes the lignin but leaves the polysaccharide behind in the same morphological relation as in the original plant tissue. Hemicelluloses

are more readily extractable from holocellulose than from dried, defatted plant tissue, and can be divided into two solubility classes, hemicellulose A and hemicellulose B.

Hemicellulose A is insoluble in slightly acidic solution, and therefore precipitates from the alkaline extract upon acidification. This class is composed mainly of high-molecular-weight linear glycans that contain small amounts of uronic acid. The degree of polymerization (DP) of these polysaccharides usually is 70–200. In comparison, the DP of cellulose is 2500–3500. For most plants, the A fraction is the largest because within this fraction occurs the abundant and widely distributed polysaccharide arabinoglucoxylan, which hydrolyzes to L-arabinose, D-glucuronic acid, and D-xylose. Sometimes the D-glucuronic acid unit has a methyl ether group at position C4 of the pyranose ring.

Hemicellulose B, which is precipitated from the neutralized alkaline extract by the addition of excess methanol or ethanol, consists of branched-chain molecules of a lower degree of polymerization than hemicellulose A. Its uronic acid content is higher than that of hemicellulose A. These hemicellulose molecules may also have D-galactopyranose and D-glucopyranose units attached at positions yet undefined. Further fractionation procedures, such as the formation of a complex between the hemicellulose molecules and copper salts and fractional precipitations with organic solvents, must be used to isolate a homogeneous polymer.

D-Xylose is the dominant building unit of the hemicelluloses of most woods and annual plants. The D-xylose content of hemicellulose from hardwood is generally higher (20–25%) than that from softwood (7–12%). D-Mannose is also very abundant in hemicelluloses; the mannose content of softwoods is usually higher than that of hardwood. Often it occurs as a polymer, mannan, or in combination with D-glucose or D-galactose as a glucomannan, galactomannan, or galactoglucomannan.

Uses. Hemicelluloses are important to the paper industry. In chemical wood pulps, hemicellulose is needed for satisfactory pulp quality. Its presence aids the swelling of the pulp, the bonding of the fibers, the bursting strength, tensile strength, tear resistance, folding endurance, opacity, and specific surface of the pulp sheet. The optimum hemicellulose content of a pulp will vary, depending on the type of wood and the conditions of the pulping operations. Excessive amounts of hemicellulose will produce a brittle or glassine paper. *See* PAPER.

On the other hand, if pulps are to be used for purposes other than papermaking, the presence of hemicellulose may be quite disadvantageous. For example, in the production of pulps which are to be used in the manufacture of rayon, cellophane, and cellulose esters and ethers, it is often essential to obtain an α-cellulose very low in hemicellulose. α-Cellulose is insoluble in 17.5% sodium hydroxide solution. In cellulose acetate manufacturing, small amounts of mannan or xylan cause hazy acetate solutions and clogged filter presses.

Hemicelluloses also serve as nutrients for yeasts, and they can be used for raw material in the production of furfural and ethyl alcohol. *See* CELLULOSE.

Roy L. Whistler

Bibliography. M. R. Brown, Jr. (ed.), *Cellulose and Other Natural Polymer Systems: Biogenesis, Structure, and Degradation*, 1982; M. F. Chaplin and J. F. Kennedy (eds.), *Carbohydrate Analysis,* 2d ed., 1994; M. P. Couglan and G. P. Hazelwood (eds.), *Hemicellulose and Hemicellulases*, 1993; R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 5th ed., 1994; F. Loewus and W. Tanner (eds.), *Plant Carbohydrates I: Intracellular Carbohydrates*, 1982.

Hemichordata

A group of deuterostome animals that includes the classes Enteropneusta, Pterobranchia, and Planctosphaeroidea. The last, a monospecific class, is represented by *Planctosphaera pelagica*, a planktonic larva that occupies low depths and resembles the larval tornaria of Enteropneusta. *See* ENTEROPNEUSTA; PTEROBRANCHIA.

The hemichordates have a slender tubular diverticulum that projects forward into the protosome from the roof of the buccal cavity. That organ, also called a stomochord or buccal diverticulum, is a supporting axial rod of the protosome and resembles the notochord of Chordata, but it does not have the same position, structure, origin, or function, and so they are not homologous. *See* CHORDATA.

The hemichordates are not plentiful animals. All are marine species that live in a wide range of habitats and depths, from intertidal to abyssal, and show a worldwide distribution. They vary in size from a fraction of an inch, or a few millimeters, such as *Rhabdopleura*, to 7 ft (2 m) or more, as with *Balanoglossus gigas*.

The members of the phylum differ widely. The enteropneusts are vermiform and solitary, whereas the pterobranchs are sacciform and colonial. The adult hemichordate is characterized by a division of both the body and the coelom into three parts: the anterior region, containing the protocoel (protosome); the middle region, containing two mesocoels (mesosome); and the posterior region, containing two metacoels (metasome). Those divisions differ in length and structure. The protosome contains a contractile heartlike vesicle, a blood sinus, and a glomerulus that rest on the buccal diverticulum. The mesosome contains a dorsal thickening of the nervous system, called a neurochord, which is isolated from the epidermis and which shows giant nerve cells and fibers that facilitate fast responses. The mesosome may also bear one or more pairs of dorsal tentaculated arms that resemble a lophophore. The mouth opens anteroventrally in the mesosome. The metasome is the longest region; anteriorly it possesses numerous paired gill slits that extend from the dorsal pharynx wall to the body surface, with one pair or none being found in

pterobranchs. The gill slits are considered to be homologous with the chordate gill slits. The gill apparatus itself may function to separate food from water, which is expelled through the gill slits. The gonads are also in the metasome and have a retroperitoneal origin. They are numerous in enteropneusts but limited to one or two in pterobranchs, and they open by separate dorsal pores. Sexes are separate, and mating organs are lacking.

The nervous system is primitive, consisting of an intraepidermal plexus. Two thickenings, middorsal and midventral, form the dorsal and ventral nerve cords. The neurochord perhaps represents a nervous center, but it lacks any special concentration of nervous elements. Scattered monociliated sensory cells respond to tactile and chemical stimuli. The circulatory system is open, with two main longitudinal contractile vessels. The colorless blood moves forward in the dorsal vessel, passes into a blood sinus on the buccal diverticulum, and is then pumped into the glomerulus by a pulsating heartlike vesicle before running backward to the ventral vessel. The glomerulus is formed from an infolding of the hind wall of the protocoel and possibly has excretory and osmoregulatory functions.

The gut is straight in enteropneusts, with a terminal anus at the end of the metasome; in pterobranchs, it is U-shaped and the anus is near the mouth. All hemichordates are filter feeders, taking in microscopic organisms by means of ciliary streams. The enteropneusts trap the food with the protosome, the pterobranchs use their tentacles, and *Planctosphaera* manipulates a ciliary band.

Spawning is synchronized, fertilization is external, and cleavage is holoblastic, radial, and nearly equal. Some species have large eggs that develop directly, without a larval stage; others have a typical larva, the tornaria, which resembles the bipinnaria or auricularia of echinoderms; the pterobranchs have a planulalike larva. In this latter group, asexual reproduction by budding occurs.

The hemichordates are a primitive group, having a tripartite body and coelom; their embryonic development resembles the echinoderms, with which they also share a primitive nervous system. The Hemichordata, therefore, may be a group at a low level of evolution, between echinoderms and chordates. Jesús Benito

Hemicidaroida

An extinct paraphyletic order of regular sea urchins (Echinoidea), identified by having a plain, unsculptured test, tubercles that are perforate and crenulate, and no suranal plate intercalated into the apical disc. They almost certainly include ancestors of later groups that developed imperforate and noncrenulate tuberculation and are thus paraphyletic.

Two families are generally placed within the order, Hemicidaridae and Pseudodiadematidae. Hemicidarids have large primary interambulacral tubercles and much smaller ambulacral tubercles, the ambulacra typically being narrow and sinuous and composed of simple plates adapically. Pseudodiadematids have equal-sized ambulacral and interambulacral tubercles at the ambitus, wider ambulacra, and trigeminate or polygeminate plate compounding throughout.

The oldest hemicidaroid is Late Triassic (Upper Norian) and the youngest comes from the Upper Cretaceous (Campanian). They appear to have been epifaunal grazers. *See* ECHINOIDEA. Andrew Smith

Hemidiscosa

An order of the subclass Amphidiscophora in the class Hexactinellida. These sponges are distinguished from the order Amphidiscosa in that the birotulates are hemidiscs with asymmetrical ends. *Hemidiscella* from the Cretaceous is an example. *See* AMPHIDISCOSA; HEXACTINELLIDA. Willard D. Hartman

Hemimorphite

A mineral sorosilicate having the composition $Zn_4Si_2O_7(OH)_2 \cdot H_2O$; an ore of zinc. It crystallizes in the orthorhombic system, pyramidal class, and thus the prismatic crystals have different forms at top and bottom (see **illus.**). Also, as a result of the

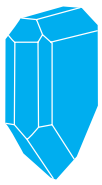

Hemimorphite crystal habit. (*After C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*)

symmetry, the vertical axis is polar and crystals display the properties of pyroelectricity and piezoelectricity. There is perfect prismatic cleavage. Botryoidal and staclactic aggregates, showing a crystalline surface and frequently impure, are common. Crystals are usually colorless and the aggregates white, but in some cases there are faint shades of green, yellow, and blue. The mineral has a vitreous luster, a hardness of $4^1/_2$ to 5 on Mohs scale, and a specific gravity of 3.45. Hemimorphite frequently resembles the zinc carbonate, or smithsonite, but can be distinguished by the reaction with hydrochloric acid. Hemimorphite is slowly soluble but smithsonite effervesces in cold acid. *See* PIEZOELECTRICITY; PYROELECTRICITY; SMITHSONITE.

Hemimorphite is a secondary mineral found in the oxidized portion of zinc deposits associated with smithsonite, sphalerite, cerussite, anglesite, galena, and more rarely the oxidized minerals of copper. Widely distributed, it has been mined in Belgium, Germany, Romania, England, Algeria, and Mexico. In the United States it is found at Sterling Hill, New Jersey; Friedensville, Pennsylvania; and Elkhorn Mountains, Montana. *See* SILICATE MINERALS; ZINC. Cornelius S. Hurlbut, Jr.

Hemiptera

An order of the class Insecta sometimes referred to as the Heteroptera. Both names refer to the forewings, which are differentiated into a thickened basal area and a membranous apical region. These are the true bugs. The related order, Homoptera, has forewings of uniform texture. Included in Hemiptera are such common insects as the bedbugs, stink bugs, plant bugs, lace bugs, and backswimmers. *See* HOMOPTERA.

The true bugs number about 25,000 species. They are known from all continents except Antarctica and occur on most islands. One group, the marine water striders (*Halobates*), inhabits the open ocean. Hemiptera range in size from small aquatic and ground-inhabiting forms approximately 0.04–0.08 in. (1–2 mm) in length to the giant water bugs that are 4 in. (100 mm) or more.

In habits, the true bugs range from strictly phytophagous types attached to a single host plant to general predators on other insects and even to specialized ectoparasites of bats. Many species are of economic importance as plant pests or vectors of disease. They occur in vegetation, on the ground, in and on the water, and in the nests of termites. One family, the Aradidae, lives on fungi on the bark of dead trees, and certain Reduviidae and Cimicidae live in caves. Among aquatic forms, members of the genus *Rhagovelia* swim on riffles of swift-flowing streams, and *Aphelocheirus* remains permanently submerged in rivers. Most water bugs depend on surface air held to the body by air spaces and hairs on the abdomen. As oxygen is depleted in the air bubble, it is replaced from the surrounding medium by diffusion.

A few Hemiptera, such as *Dacerla*, *Coquillettia*, and *Pilophorus*, mimic ants. Others, including cotton stainers (*Dysdercus*) and certain reduviid predators (*Phonoctonus*), resemble distasteful cantharids and other insects of orange or yellow and black colors, which is a protective Müllerian type of mimicry. Still other bugs resemble their surroundings and hence are concealed from birds, lizards, and other enemies. One of the most striking of these is the flat, barklike *Phloea* of South America. Stridulation, or sound production, is less conspicuous in Hemiptera than in the cicadas of Homoptera and katydids of Orthoptera but is developed in several groups. Stink bugs, flat bugs, the predacious reduviid bugs, and several aquatic groups are examples. The sounds are produced either by rubbing filelike portions of the legs against knife edges on the body wall, by rubbing parts of the wings against the body, or by rubbing the beak along a prosternal cross-striated stridulatory furrow, as occurs in the Reduviidae (**Fig. 1**). *See* PROTECTIVE COLORATION.

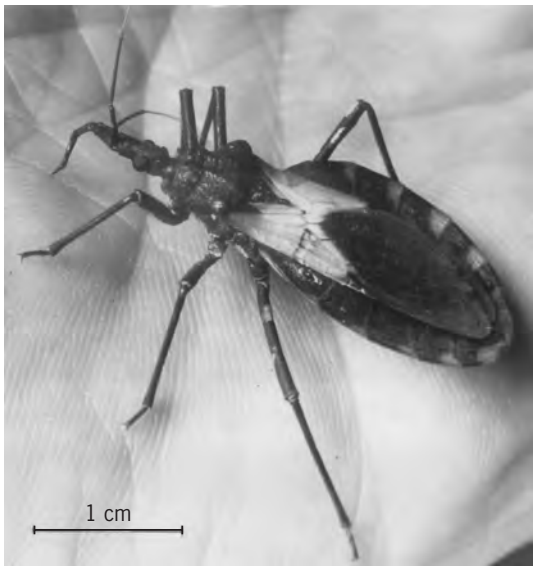

Fig. 1. ***Triatoma pallidipennis*, of the Reduviidae. (*Photograph by E. S. Ross*)**

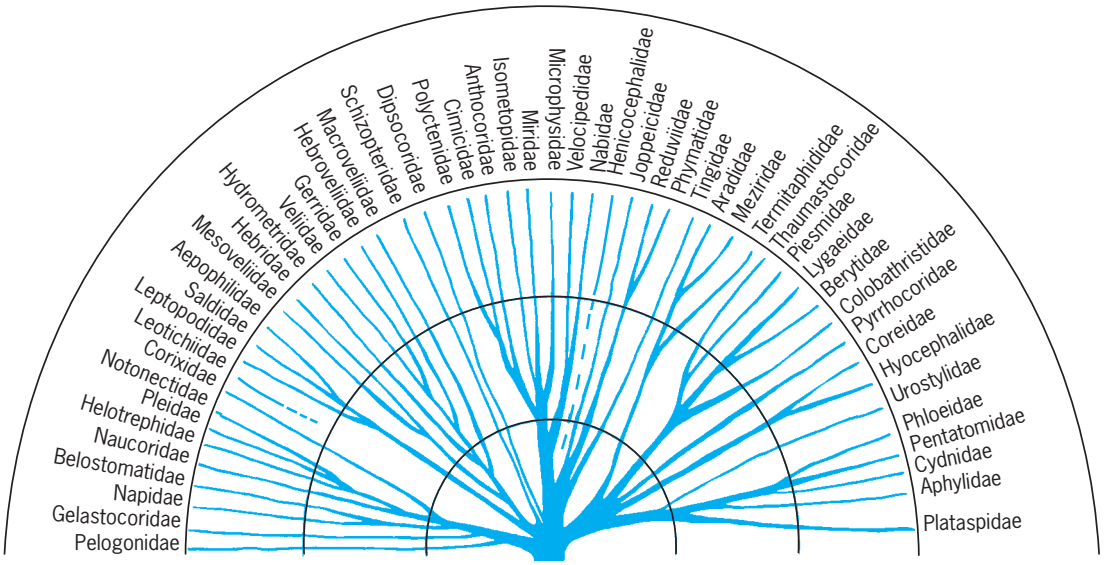

Fig. 2. Relationships of families of Hemiptera.

TABLE 1. Families of Hemiptera

Family	Common name	Distribution	No. of species
Subdivision Hydrocorisae			
Corixidae	Water boatmen	General	300
Nepidae	Water scorpions	General	170
Belostomatidae	Giant water bugs	General	140
Notonectidae	Backswimmers	General	170
Pleidae	None	General	20
Helotrephidae	None	Tropical	20
Naucoridae	Creeping water bugs	General	200
Gelastocoridae	Toad bugs	Tropical and subtropical	80
Ochteridae	Velvety shore bugs	Tropical and subtropical	20
Subdivision Amphibicorisae			
Gerridae	Water striders	General	300
Veliidae	Smaller water striders	General	200
Hydrometridae	Marsh treaders	General	50
Mesoveliidae	Water treaders	General	20
Hebridae	Velvety water bugs	General	40
Subdivision Geocorisae			
Superfamily Leptopodoidea			
Saldidae	Shore bugs	General	200
Leptopodidae	None	Tropical and subtropical	20
Leotichidae	None	Oriental	2
Superfamily Dipsocoroidea			
Dipsocoridae	None	General	
Schizopteridae	Jumping ground bugs	Tropical and subtropical	
Superfamily Cimicimorpha			
Cimicidae	Bat, bed, bird bugs	General	80
Anthocoridae	Flower bugs	General	300
Polyctenidae	Bat bugs	Tropical and subtropical	20
Miridae	Plant bugs	General	5000
Microphysidae	None	Palearctic	30
Plokiophilidae	None	Tropical	20
Nabidae	Damsel bugs	General	250
Tingidae	Lace bugs	General	700
Vianaidae	None	Neotropical	2
Thaumastocoridae	Palm bugs	Tropical	11
Superfamily Enicocephaloidea			
Enicocephalidae	Gnat bugs	General	300
Superfamily Reduvioidea			
Reduviidae	Assassin bugs	General	3500
Superfamily Aradoidea			
Aradidae	Flat bugs	General	800
Termitaphididae	Termite bugs	Tropical	10
Superfamily Pentatomorpha			
Idiostolidae	None	Chilean	2
Lygaeidae	Lygaeid bugs	General	2000
Thaumastellidae	None	Ethiopian	1
Colobathristidae	None	Tropical	70
Berytidae	Stilt bugs	General	100
Malcidae	None	Ethiopian, Oriental	30
Piesmatidae	Ash-gray leaf bugs	General but discontinuous	20
Pyrrhocoridae	Pyrrhocorid bugs	General	300
Largidae	None	General	100
Coreidae	Coreid bugs	General	2000
Rhopalidae	None	General	300
Stenocephalidae	None	Old World, Neotropical	20
Hyocephalidae	None	Australia	1
Pentatomidae	Stink bugs	General	2500
Phloeidae	Bark bugs	Neotropical	5
Plataspidae	None	Old World	400
Lestoniidae	None	Australia	1
Cydnidae	Ground or burrower bugs	General	600
Urostylidae	None	Oriental and Australian	50
Aphylidae (not placed)	None	Australian	2
Joppeicidae	None	Mediterranean	1

Most Hemiptera are bisexual and oviparous, but parthenogenesis is known and a pseudoplacental organ results in a special type of viviparity in Polyctenidae. Mating usually takes place on vegetation or on the ground, the pairing being end-to-end in stink bugs, squash bugs, chinch bugs, and similar species, and with the male above the female in most others. Bedbugs (Cimicidae) have a special organ of Ribaga (or organ of Berlese) through which spermatozoa are introduced to the hemocoel of the female where they penetrate the ovarioles and fertilize the ova. *See* INSECT PHYSIOLOGY; INVERTEBRATE EMBRYOLOGY.

Classification. The Hemiptera were first divided by P. Latreille, in 1825, into Hydrocorisae (water bugs) and Geocorisae (land bugs). L. Dufour, in 1833, proposed a third group, Amphibicorisae (surface water bugs). F. Fieber, in 1851, substituted the Cryptocerata (hidden antennae) for Hydrocorisae and Gymnocerata (exposed antennae) for Geocorisae. Both systems are in use today, but the Latreille-Dufour system conforms best to recent studies. In 1954, the Geocorisae were divided into Cimicomorpha and Pentatomomorpha by D. Leston, J. Pendergrast, and T. Southwood. R. H. Cobben, in 1968, recognized nine more or less equivalent major subdivisions in the Hemiptera. **Table 1** represents a balance between extremes and includes each distinctive group, down to the family level. The cladogram shown in **Fig. 2** summarizes the last published survey (1956) of the possible phylogeny of the families of the Hemiptera.

Habitats. The Hemiptera occupy various terrestrial and aquatic habitats (**Table 2**). Some species are intimately associated with plants and animals.

Morphology and Embryology

In Hemiptera and Homoptera the mouthparts are elongate and slender, forming a sucking mechanism with sheathlike labium and needlelike mandibulary and maxillary stylets (**Fig. 3**). In Homoptera, as in leafhoppers and aphids, the beak arises from the posterior part of the head, whereas in Hemiptera the position is anterior and the head is commonly directed forward or downward rather than backward.

The stylets are grooved and slide up and down during the act of feeding, the mandibles being toothed to grip the plant or other tissue that is being penetrated. Two tubes are formed by the stylets, one for sucking up the fluid food and the other for injecting salivary or anticoagulant fluids. The labium is jointed with three or four segments.

Hemiptera are further characterized (**Fig. 4**) by antennae, usually of four or five segments, a pair of compound eyes, and often two ocelli. The thorax consists of a prominent pronotum, a triangular mesothoracic scutellum, and a broad metathorax which is partly fused with the first abdominal segment. The mesothoracic wings, or hemelytra, overlap at their membranous apices when at rest, leaving the scutellum exposed. They consist of a leathery corium at the base, an inner clavus along the scutellum, and an apical membrane. A marginal fracture separates the apex of the corium into a cuneus in some groups. The venation is seldom clear but has been interpreted from the tracheae of nymphal wing pads as consisting of the subcosta at the anterior edge of the wing, the radius and media more or less fused behind, the cubitus as the posterior vein of the corium, and one or more anal veins in the clavus. The hindwings are hitched to the forewings in flight by grooves and pegs. In metathoracic wings, the subcosta, radius, and media are variously fused but usually form an elongate cell with or without a backward-directed spur vein or hamus. The anal and jugal areas are more or less developed with ill-defined veins and folds.

Wings are sometimes reduced to short pads and may be lacking in certain groups or even in members of a single species. Wing polymorphism, especially in water striders, appears to be determined by environmental factors as well as by genetic factors.

The front legs are frequently enlarged and sometimes chelate in predacious forms, and the middle and hindlegs are adapted for swimming in some groups. The foretibiae have a small plate or comb at the inner apex and sometimes a spongy pad for clinging to smooth surfaces or holding prey. Tarsi usually are of three segments, but some groups have only two segments and front tarsi are lacking in a few specialized types. Two claws are commonly present with or without leaflike or bristlelike arolia between.

The abdomen is of 10–11 segments, but only seven are commonly seen, the first two being fused and the

TABLE 2. Habitats of Hemiptera

Habitat	Families
Terrestrial	
On surface beneath stones and wood	Nabidae, Reduviidae, Lygaeidae, Enicocephalidae, Leptopodidae
In rotting plant material	Schizopteridae, Lygaeidae
Beneath surface of ground	Vianaidae
Aquatic	
True aquatics	Corixidae, Nepidae, Belostomatidae, Notonectidae, Pleidae, Helotrephidae, Naucoridae
Surface bugs	Gerridae, Veliidae, Hydrometridae, Mesoveliidae
Shore bugs	Gelastocoridae, Ochteridae, Hebridae, Mesoveliidae, Saldidae, Leptopodidae, Dipsocoridae
Marine intertidal	Saldidae
Other associations	
In bird nests	Anthocoridae, Cimicidae, Reduviidae, Lygaeidae
Ectoparasites	Polyctenidae
In termite colonies	Termitaphididae, Aradidae, Reduviidae
In spider webs	Plokiophilidae, Nabidae, Reduviidae
In fungi	Aradidae, Plataspidae
On seeds	Lygaeidae
On stems, foliage, flowers	Anthocoridae, Miridae, Nabidae, Reduviidae, Tingidae, Lygaeidae, Bertyidae, Piesmidae, Pyrrhocoridae, Coreidae, Pentatomidae, Cydnidae, Urostylidae, Phloeidae, Plataspidae

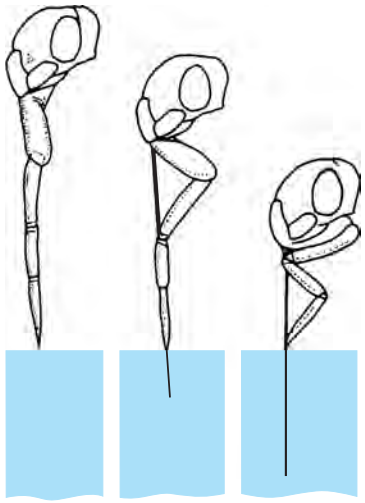

Fig. 3. Mouthparts and method of feeding of a plant bug.

last being partly concealed by the genitalia or withdrawn into the abdomen. Reproductive organs are shown for both sexes in copulation (**Fig. 5**). Males have the ninth segment variously developed as a genital capsule. A pair of parameres is commonly associated with the genital capsule, and the aedeagus, or penis, arises from its floor. The aedeagus is hinged to basal plates. An ejaculatory duct traverses the entire structure, terminating in the gonopore. Clasping organs may be present on either side of the thickened phallosoma, and various lobes and processes are developed on the often membranous endosoma. The latter is evaginated during mating, with the vesica variously inflated and inserted in the female genitalia. The female abdomen is variously modified but commonly has an ovipositor, with bent valves, arising from the eighth and ninth segments. A spermatheca is present in many groups and is distinctive in shape and form in some. Spiracles occur laterally, either in a ventral position or dorsally and most often in or near lateral connexival plates. Long sensory hairs, or trichobothria, occur on the ventrites of stink bugs, squash bugs, chinch bugs, and some others.

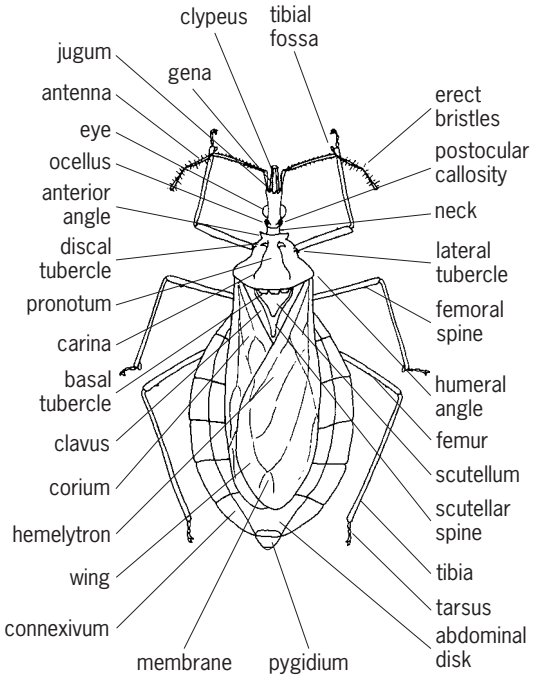

Fig. 4. External anatomy of a bug.

Internal anatomy. Internally the Hemiptera resemble other orders of insects, but several special features are noteworthy. Parts of the midgut have fluted ceca in stink bugs, squash bugs, and some others. These contain symbiotic bacteria which are specific and are transmitted from one generation to the next, in or on the ova. In stink bugs the eggs are smeared at the time of laying, and the first-stage nymphs, when hatched, cluster around the egg shells and suck up bacterial material. Cimicidae have no gastric ceca but possess a pair of mycetomes in about the third abdominal segment. The function of symbiotic bacteria is not clear, but in the Reduviidae it has been shown that vitamins of the B group are produced.

Salivary glands are of two parts, the bilobed principal glands and the tubular accessory glands. They empty into a common duct and thence to a pump which forces the fluid through the stylet tube. The Malpighian tubules are usually four in number.

Respiration is accomplished by spiracles, of which there are 10 pairs or less. In terrestrial forms a closing apparatus is developed to conserve moisture, whereas in water bugs the spiracles are open. Breathing tubes are present in Nepidae, enabling these water scorpions to use surface air while submerged. *Buenoa* and related backswimmers have hemoglobin in clusters of cells adjacent to the spiracles.

Endocrine glands include the corpus cardiacum and corpus allatum located behind the brain. These control molting and metamorphosis. The presence of hormone from the corpus allatum (juvenile, or inhibitory, hormone) prevents metamorphosis in young nymphs. Later, after four or five nymphal instars, the inhibitory hormone is no longer present and the growth and differentiation hormone acts to produce metamorphosis. *See* ENDOCRINE SYSTEM (INVERTEBRATE).

Scent glands are a characteristic feature of bugs, causing the well-known "buggy odor." Most adult bugs have a pair of metathoracic glands with openings on the metapleura, as in Pentatomidae and others, or on the hind coxae, as in Reduviidae.

A few have additional glands in the first abdominal segment. Nymphs of all but a few, such as some Reduviidae and the Hydrometridae, have dorsal abdominal scent glands. The arrangement varies from a full complement of three, opening at hind margins of third, fourth, and fifth tergites; two, on the third and fourth or the fourth and fifth tergites; one, on the fourth tergite; to none. The openings may be single (Miridae) or double and widely separated (Naucoridae). In adults of the water striders (Gerridae), there is a single opening (omphalium) at the middle of the metasternum.

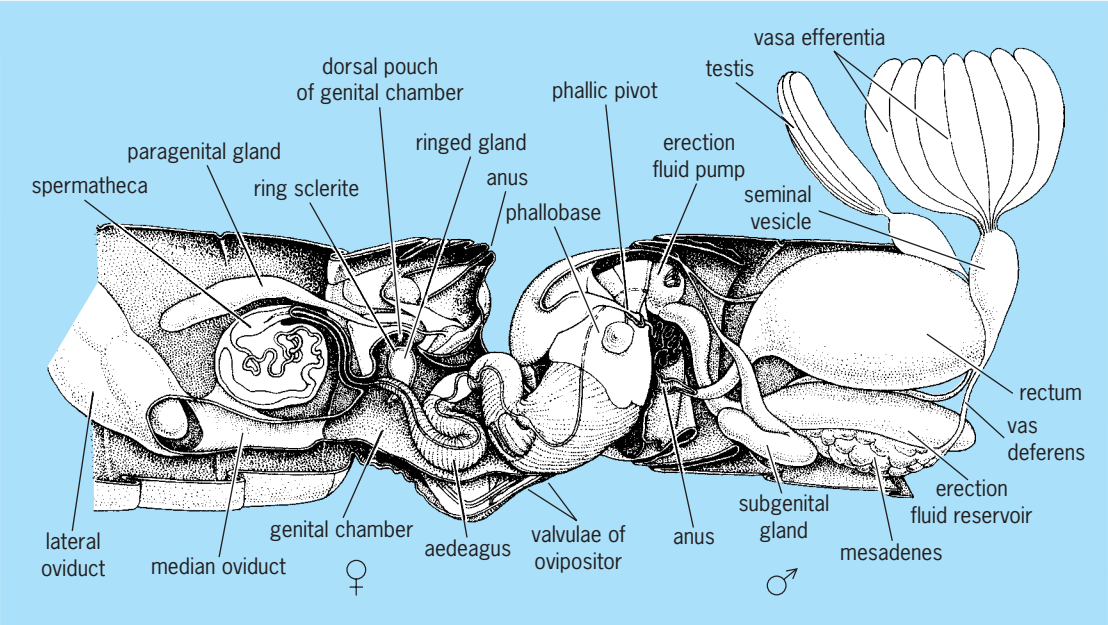

Fig. 5. Reproductive organs of female and male *Oncopeltus fasciatus* in copulation.

Internal reproductive organs include, in females, a pair of acrotrophic ovaries with ovarioles numbering 2–8, with 7 as the commonest number. A spermatheca is well developed and of characteristic shape for each of the main groups of stink bugs, squash bugs, and others, but is absent in plant bugs (Miridae) and their allies. In males, the paired testes comprise 1–7 follicles. The seminal vesicles are usually large and are formed by a glandular reservoir. Accessory glands known as mesadenia empty into the reservoir, and ectadenia join the ejaculatory bulb in some groups.

Chromosomes are best seen in late instar nymphs or during early adult life in males. The $2n$ (diploid) number may be as low as 6, in *Rhytidolomia*, or as high as 43, in *Ranatra*. Commonly, there are 12 autosomes and 2 X chromosomes or an X and a Y. In *Cimex* and some others there are supernumerary X chromosomes. *See* GENETICS.

Development and metamorphosis. Hemiptera undergo simple or incomplete metamorphosis including egg, nymph, and adult stages (**Fig. 6**). Eggs are characterized by a chorion of several layers, usually with hexagonal reticulations on the surface. A lid or cap is present in the plant bugs, bedbugs, and allies, but the anterior end of the egg is marked only by micropylar processes in other groups. An egg burster is well developed and T-shaped in Pentatomidae and less developed but visible in some other groups. Eggs are glued to surfaces by *Lethocerus*, inserted in plant tissues by Miridae, or laid free by some *Triatoma*. In *Belostoma* and *Abedus*, the eggs are laid on the backs of males, presumably for protection, and a corixid (*Ramphocorixa*) regularly lays its eggs on crayfish in pasture ponds. The eggs may be stalked, as in Corixidae; spindle-shaped, as in Hydrometridae; or oval- or barrel-shaped, as in Pentatomidae, and may be laid in clusters, as by *Zelus*, or singly. The female "broods" or "guards" the eggs in *Nerthra* and some Scutelleridae.

Embryonic development is endoblastic and is not very different from that of other insects. Glandular organs, the pleuropodia, occur on the first abdominal segment and produce the embryonic cuticle. In Polyctenidae the pleuropodia function as a pseudoplacental organ. The incubation period varies from a few days to several months, and development is arrested in species that overwinter in the egg stage, as do some Miridae. As development proceeds, red eyespots become visible through the chorion. The eggs of many Miridae are embedded in woody tissue and swell because of growth and absorption of moisture, forcing out a yolk plug (*Notostira*).

Eclosion, or hatching, occurs either by splitting or tearing the chorion or by raising the lid. A postembryonic molt occurs during eclosion, and commonly the exuviae are attached to the egg shell.

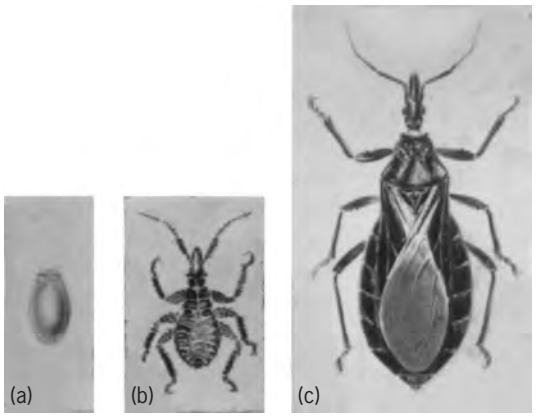

Fig. 6. *Triatoma protracta*. (*a*) Egg. (*b*) Nymphal instar. (*c*) Adult.

First instar nymphs are active or, in Pentatomidae, sedentary and gregarious. They take little or no food, the yolk remaining from the egg stage being sufficient, in most cases, to carry them through to the second instar. Bacteria are taken up from crypts amid the egg shells or from the egg surface in Pentatomidae, thus infecting the gastric ceca.

Wing pads begin to show on the hind margins of meso- and metanota at the third stage, the third nymphal instar. Nymphs, also called larvae or neanides, commonly undergo five instars, but rare cases have been reported with four in some Veliidae and Cimicidae. One case, *Dindymus sanguineus*, of the Pyrrhocoridae, has been reported with nine instars. The final molt and metamorphosis to the adult stage results in an increase in tarsal segmentation, from one to two or from two to three, an increase in number of antennal segments in some cases (four to five in Pentatomidae), and development of the wings and external genitalia. Also, the ocelli are fully developed at this time, and numerous internal changes occur.

Subdivisions

Important families of the Hemiptera and their distinguishing characteristics are treated under the respective subdivisions. Many of the economically important species are given in **Table 3**.

Hydrocorisae. This subdivision contains nine families of water bugs with concealed antennae and without a bulbus ejaculatorius in the male. Many species are predaceous.

Corixoidea. This is a superfamily which contains the single family Corixidae, or water boatmen. Corixidae lack ocelli, have three dorsal abdominal scent-gland openings as nymphs, and have a unique type of mouthpart. The rostrum is broad, with a pseudosegmentation, and the food consists of algae, as well as small animal forms gathered by the comblike front tarsi, or palae. Respiration is through an air bubble obtained at intervals by touching the surface with the pronotum. Corixids swim with the dorsal side uppermost, using the oarlike middle and hind legs. Eggs are stalked and there are five nymphal instars. Adults sometimes fly to lights in great numbers. In Mexico, the eggs are harvested from lakes to make a bread (ahuatle) and the adults (mosco) are dried and sold in the United States as food for pet birds and turtles.

Nepoidea. The Nepidae, or water scorpions, have a long breathing tube at the tip of the abdomen, through which they obtain air directly from the surface. The front legs are chelate, and the beak is short and stout to suck the juices of other insects on which they prey. Two common types of water scorpion are the long, slender *Ranatra*, of worldwide distribution, and the broad *Nepa*, of the eastern United States and Europe. The eggs are embedded in plant tissue and have two flaplike appendages.

Belostomatidae, or electric-light bugs, are related to the Nepidae but have short, straplike respiratory appendages at the tip of the abdomen. Giant water bugs of the genus *Lethocerus*, reaching 4 in. (10 cm) in length, fly to lights at night. They are pests in fish ponds where they attack fry. They can inflict a painful bite when handled carelessly by humans. In China, the adults, called kwai fa shim, are boiled in oil and sold in markets as food for humans. Eggs are laid on cattails and other plants at the water's edge. Control in fish ponds is by draining or by rearing fry and fingerlings in screened pools. Smaller, 1–2-in. (2.5–5-cm) belostomatids (*Belostoma*, *Abedus*) lay eggs on the backs of the males, possibly for protection.

Notonectoidea. The Notonectidae or backswimmers, as the name implies, swim ventral side uppermost, the hindlegs serving as oars. Breathing is facilitated by an air bubble, obtained by touching the tip of the abdomen to the surface. *Notonecta* is worldwide in distribution. It is stout in body form and swims with jerky movements or clings to underwater supports. *Buenoa* in the Western Hemisphere and *Anisops* in the Old World have hemoglobin around the spiracles and maintain hydrostatic balance in the water, thus being truly limnetic forms.

Pleoidea. These minute bugs of the families Pleidae and Helotrephidae are suboval, with legs not fitted for rowing. Unlike Notonectidae, the nymphs have a single scent gland opening on the abdominal tergum.

Naucoroidea. The creeping water bugs are sometimes separated into the Naucoridae and Aphelocheiridae. They are suboval in body form, with chelate front legs and with two nymphal scent gland openings widely separated at the hind margin of the third abdominal tergite. Respiration is either by an air bubble or by means of a plastron. In the Old World genus *Aphelocheirus*, the plastron consists of an ultramicroscopic hair pile (2,000,000 hairs per square millimeter) which acts as a physical gill; oxygen diffuses through it from the water to the spiracles of the bug. *Aphelocheirus* is thus able to remain submerged permanently, inhabiting the bottoms of rivers with swift-flowing, well-oxygenated water.

Gelastocoroidea and Ochteroidea. The Gelastocoridae, or toad bugs, and the Ochteridae (or Pelogonidae) are shore-line or mud inhabitants. Both have ocelli. The former are cryptically colored, resembling the sand or mud background. Their eyes are large and their antennae concealed. Ochterids are black with a silky sheen and the antennae, unlike those of other Cryptocerata, are visible from above, though short and partly concealed beneath the eyes. Eggs are laid in sand and mud.

Amphibicorisae. This subdivision contains surface water bugs with antennae exposed and without a bulbus ejactulatorius in the male. Spermatheca have a fecundation canal.

Only the single, diverse superfamily Gerroidea has been proposed for the surface water bugs. All have conspicuous antennae, and the body is clothed with hydrofuge hairs. Three pairs of sensory hairs, the trichobothria, are inserted between the eyes. All Gerroidea are predacious.

Gerridae are the large water striders with long middle and hind legs and a median scent gland opening, the omphalium, on the metasternum. The claws are

TABLE 3. Economically important Hemiptera

Name	Damage
Cimicidae	
Human bedbug (*Cimex lectularius*)	Sucks blood in human dwellings; eggs with small caps, glued in cracks
Tropical bedbug (*Cimex hemipterus*)	Sucks blood in human dwellings; lives in cracks and sleeping mats
Swallow bug (*Oeciacus vicarius*)	In mud nests of swallows; reported on humans in houses with swallows
Mexican chicken bug (*Haematosiphon inodora*)	Sucks blood in chicken roosts; eggs laid in cracks
Reduviidae	
Bloodsucking conenose bugs (*Triatoma* sp., *Rhodnius* sp., *Panstrongylus* sp.)	Several species live in nests of wood rats (*Neotoma*) and suck blood; they transmit Chagas' disease in Latin America
Miridae	
Tarnished plant bug (*Lygus lineolaris*)	Eggs inserted in plant tissue; overwinters as adult; feeding deforms leaves and stems; punctures cause cat-facing on fruits
Lygus bugs (*Lygus hesperus*)	Several species stunt alfalfa, also attack beans, cotton, and other plants; they overwinter in adult stage and lay eggs in plant tissue
Apple red bug (*Lygidea mendax*)	Feeding punctures cause pitting of fruits; winters in egg stage, with eggs inserted in bark
Ash plant bug (*Neoborus amoenus*)	Feeding stunts new growth in spring; eggs laid in stems and bark scars in summer–overwinter in egg stage
Suckfly (*Cyrtopeltis notatus*)	Feeding and egg laying in ring around stems of tomato and tobacco; causes tomato fruits to drop
Cotton flea hopper (*Psallus seriatus*)	Sucks cotton squares causing shedding and whiplike growth of plant; overwinters in egg stage
Garden flea hopper (*Halticus bracteatus*)	Feeding punctures spot leaves of many garden plants; overwinters as adult
Tingidae	
Azalea lace bug (*Stephanitis pyrioides*)	Introduced from Japan; overwinters in egg stage; three broods a year in New Jersey
Rhododendron lace bug (*Stephanitis rhododendri*)	Overwinters in egg stage; two broods in New Jersey; feeding punctures and fecal spots on undersides of leaves
Basswood lace bug (*Gargaphia tiliae*)	Overwinters in adult stage; eggs laid in spring, partially inserted on undersides of leaves
Eggplant lace bug (*Gargaphia solani*)	Hibernates in adult stage; leaves spotted and yellowed; up to six generations a year in Virginia
Chrysanthemum lace bug (*Corythucha marmorata*)	Overwinters as adult; eggs inserted deep in veins under leaves; feeding punctures and fecal spots
Sycamore lace bug (*Corythucha ciliata*)	Adults hibernate under bark; feeding punctures and fecal spots on under surface of leaves
Lantana lace bug (*Teleonemia scrupulosa*)	Feeds on leaves of lantana, causing leafdrop; introduced into Hawaii from Mexico to control lantana
Ash lace bug (*Leptotypha minor*)	Overwinters in adult stage; eggs, feeding punctures, fecal spots, and cast skins on undersides of leaves
Lygaeidae	
Chinch bug (*Blissus leucopterus*)	Attacks corn and small grains in the Middle Western states, causing wilting and drying of plants; overwinters in adult stage with 2±3 generations a year
Hairy chinch bug (*Blissus hirtus*)	Attacks lawns in eastern U.S., killing grass in spots; life cycle similar to *B. leucopterus*
False chinch bug (*Nysius ericae*)	Several species of the genus *Nysius* breed on weeds and migrate in large numbers to cultivated crops, causing wilting
Coreidae	
Boxelder bug (*Leptocoris trivitatus*)	Overwinters as adult; attacks seeds of boxelder; two generations a year
Squash bug (*Anasa tristis*)	Attacks most cucurbits, wilting leaves or killing small plants; overwinters as adult; the small brown eggs are laid in groups on the plants in the spring and early summer
Pyrrhocoridae	
Cotton stainers (*Dysdercus suturellus and Dysdercus minulus*)	In Florida and Arizona, puncture bolls and seeds and stain the lint
Pentatomidae	
Brown stink bug (*Euschistus servus*)	Breeds on various weeds and attacks most cultivated crops; feeding stains cotton, pits or causes cat-facing of pears, peaches, tomatoes, and other plants; overwinters in the adult stage
Dusky stink bug (*Euschistus tristigmus*)	
One-spot stink bug (*Euschistus variolarius*)	
Conchuela (*Chlorochroa ligata*)	
Say stink bug (*Chlorochroa sayi*)	
Green stink bug (*Acrosternum hilare*)	
Southern green stink bug (*Nezara viridula*)	
Harlequin bug (*Murgantia histrionica*)	Feeds on foliage, causing wilt and death of the plants; attacks cabbage and many other garden plant; breeding is more or less continuous, with adults and nymphs hibernating in cold weather

inserted before the tips of the tarsi. Most species of *Gerris* are pond or slow-moving-stream inhabitants, but *Metrobates* lives on large rivers and *Halobates* inhabits lagoons in tropical seas and even the open ocean. The marine forms are always wingless but all others are polymorphic, with fully winged forms occurring together with short-winged or apterous types. *Halobates* feeds on small anemones and other floating organisms and lays its eggs on feathers, pumice, and presumably anything else that floats. One species, *H. micans*, occurs in all tropical seas but the remaining 40 or more species are Indo-Pacific.

Veliidae are small water striders which have shorter legs and a longitudinal groove between the eyes. The claws are preapical. Like the Gerridae, these are pond inhabitants (*Microvelia*), stream-riffle bugs (*Rhagovelia*), and marine types (*Halovelia*), but the last are found only near shore in tropical reefs.

Hydrometridae are long, slender marsh treaders in which the head is longer than the thorax. The claws are apical. These have been reported as predacious on *Anopheles* mosquito larvae, and attempts have been made to transport them for purposes of biological control.

Mesoveliidae and Hebridae are two small families which differ from others in having the well-developed ocelli and the single dorsal abdominal scent gland openings of the nymphs. *Mesovelia mulsanti* is a swift water strider, but some other species are shoreline inhabitants, as are the Hebridae or velvet water bugs.

Geocorisae. This subdivision contains the land bugs with conspicuous antennae and an ejaculatory bulb in the male reproductive system. This subdivision can be divided into seven groups, each of which may be equivalent in rank to the Hydrocorisae and Amphibicorisae. The two largest groups are the Cimicimorpha and the Pentatomorpha, with the included families having certain common characteristics. The Cimicimorpha includes the Cimicoidea and Tingoidea, which have distinctly operculate eggs and are without a median spermatheca; the accessory salivary glands are vesicular, the abdomen is without trichobothria, and claws are without discrete lateral arolia. The Thaumastocoroidea also belong here. The Pentatomorpha contains the Pentatomoidea, Coreoidea, Pyrrhocoroidea, and Lygaeoidea. In these superfamilies, the eggs are nonoperculate, a median spermatheca is present, accessory salivary glands are tubular, the abdomen has trichobothria, and the claws have well-developed lateral arolia. Superfamilies that do not fit the above groupings are the Reduvioidea, formerly included in the Cimicimorpha; the Saldoidea, with three pairs of head trichobothria like Amphibicorisae; the Arodoidea, which are like the Pentatomorpha but without trichobothria; and the Enicocephaloidea and Dipsocoroidea, of isolated position in the system.

Saldoidea. The superfamily Saldoidea, of the group Leptopodoidea, comprise the shore bugs which have three pairs of trichobothria on the vertex and a single nymphal scent gland opening. Saldids lay their eggs amid moss or stones at the edges of streams and ponds. One genus, *Chiloxanthus*, lives on the tundra and overwinters for 9 months with hundreds of feet of permafrost beneath the winter ice above. An intertidal genus, *Aepophilus*, occurs along the coasts of Britain and France. Leptopodidae differ in the spiny body and appendages. They live under drier conditions, sometimes remote from the water. Leotichidae is a rare group of doubtful affinities; specimens have been found in caves.

Dipsocoroidea. This is a group of minute ground inhabitants, of which the Schizopteridae live in leaf mold and the Dipsocoridae are predators on small insects under bark or in rotten wood, such as *Ceratocombus*, or amid stones at the edge of streams, such as *Cryptostemma*.

Cimicimorpha. This is the largest group of the Geocorisae and is divided into three superfamilies, the Cimicoidea, Tingoidea, and Thaumastocoroidea.

1. Cimicoidea. Anthocoridae, or the flower bugs, are small predators on thrips, mites, and similar species in vegetation (*Orius*), in stored products (*Lyctocoris*), and in the nests of birds. The ocelli are distinct, and the front wings have a marginal fracture, the cuneus. Fertilization of the female is hemocoelic through special organs. The male right genital clasper is large and bladelike; the left is lacking.

Cimicidae contains the bedbugs, which probably were derived from the Anthocoridae but differ in habits, in the absence of ocelli, and in having the short wings reduced to pads. Fertilization is usually through an organ of Ribaga on the second, third, or fourth ventrites or the third, fourth, or fifth tergites. The food consists of blood of birds and mammals. The common bedbugs of humans are *Cimex lectularius*, in temperate regions and large cities, and *Cimex hemipterus*, in the tropics. *Oeciacus* lives on swallows in Europe and North America; *Paracimex*, on cave swiftlets in eastern Asia; *Ornithocoris* and *Haematosiphon*, on chickens in the Western Hemisphere; and *Cimex columbarius*, on pigeons in Europe. Most Old World species, including those of the genus *Cimex*, live on bats, and a few bat parasites are found in the New World. Bedbugs breed continuously in houses and thus are able to complete several generations in a year. Bedbugs and chicken bugs are controlled by residual sprays of 5% DDT or by fumigation. Resistance has been reported to standard dosages of DDT.

The Polyctenidae are bat ectoparasites which resemble the Cimicidae but lack eyes and have ctenidia and strong claws. Reproduction is by hemocoelic fertilization, and the embryo is nourished by a pseudoplacental organ, resulting in true viviparity.

The Miridae family contains a majority of the species of Hemiptera. Included are plant bugs of both herbivorous and predacious type. Ocelli are lacking, a cuneus is present, the antennae are four-segmented, and there is only one dorsal abdominal scent gland opening in nymphs. The eggs are embedded in plant tissue. Of the predators, *Cyrtorhinus mundulus* was introduced into the Hawaiian

Islands and is credited with controlling the sugarcane leafhopper by sucking its eggs. Plant pests include the tarnished plant bug, *Lygus lineolaris*, and other Lygus bugs on alfalfa, cotton, and beans. Control may be obtained with 5% DDT dust applied not later than 30 days before harvest. Lygus bugs overwinter in the adult stage and may undergo two or more generations in a year. Most plant bugs, however, overwinter in the egg stage and are univoltine.

Plant bugs and also some stink bugs and others leave their preferred weed hosts as the growing season advances to suck developing fruit such as apples and pears, thus causing deformity or "cat-facing." Control is by clean cultivation or by DDT dust or spray.

Two small families, the Plokiophilidae and Microphysidae, are close relatives of Miridae. Both are predacious. The Plokiophilids live in the webs of spiders and embiids.

Nabidae are the long "damsel" bugs which are slender predators on other insects. The ocelli are well developed, and the rostrum is four-segmented. The eggs are embedded in plant tissue. Nymphs have three abdominal scent gland openings. *Nabis ferus* is a cosmopolitan predator in grasses and other vegetation. The tropical Velocipedidae are related but differ in the broader form and darker coloration.

2. Tingoidea. This superfamily of the group Cimicimorpha contains the lace bugs of the family Tingidae which have wings with many lacelike areolae. They lack ocelli, have four segments in the antennae and beak, and commonly have one or more bulbous or hoodlike elevations on the thorax. Pests are known on ornamental plants such as Christmas berry (*Corythucha*), ash (*Leptoypha*), and other ornamental plants, and some species attack agricultural crops, such as *Gargaphia solani* on eggplant. The latter is controlled with Malathion as a 5% dust. Lace bug injury is characteristic. The leaves of infested plants have white areas due to feeding, and the undersurfaces of leaves have black fecal spots and usually cast skins adhering to the leaf hairs. Lace bugs commonly overwinter as adults and pass through more than one generation each year. *Teleonemia scrupulosa* was introduced into the Hawaiian Islands from Mexico in an effort to control the lantana plant. The rare South American Vianaidae are also included in this group.

3. Thaumastocoroidea. The only family, the Thaumastocoridae, occurs in Australia and the New World tropics; it includes the royal palm bug (*Xylastodorus*) of Florida.

Enicocephaloidea. This is a unique group in which the head is bilobed, the pronotum trilobed, and the wings completely membranous. There is no prosternal stridulatory groove, the eggs are without an operculum, and the males swarm like gnats. The life cycle is not fully known, but they live under stones, beneath bark, and in leaf mold and are predators. Some species are wingless; others can cast off their wings.

Reduvioidea. This group includes only the assassin bugs or conenose bugs of the family Reduviidae. Nearly all have a stridulatory furrow on the prosternum, which is scraped by the rostrum to produce a squeaking sound. Ocelli are generally present, the beak is three-segmented, and lateral scent gland openings usually lie beneath the hind coxae rather than on the metapleura. The corsair (*Rasahus*) is black with a spot on the membrane of the wings. It lives beneath stones and feeds on insects but flies to light, where it encounters and often bites man. Kissing bugs, the Triatominae, have a longer, more slender rostrum and lack scent glands on the nymphal abdominal tergites. They are exclusively bloodsuckers, living mostly in the nests of wood rats of the genus *Neotoma* in the United States. Several species regularly feed on humans in the American tropics: *Panstrongylus megistus*, *Rhodnius prolixus*, *Triatoma infestans*, and others. Up to 30% of the bugs are commonly infected with *Trypanosoma cruzi* in some places. The trypanosome which causes the disease American trypanosomiasis, or Chagas' disease, is picked up by the bug while sucking blood but is transmitted through the feces. Fecal contamination may be at the place of the bite or in the conjunctiva of the eye—the bites are often about the face (hence the name kissing bug) and the feces are rubbed into the eye. The trypanosomes affect muscles of the heart and other organs. Kissing bugs may be controlled with residual sprays, such as DDT, by eliminating nesting rodents, or by screening bedrooms. *Rhodnius prolixus* is maintained as a laboratory animal for studies of insect physiology. Other reduviids include the thread-legged Emesinae that live in spider webs and elsewhere, the large wheel bug (*Arilus*) of the southern United States, and many other predators of injurious insects. Phymatines with sometimes concealed antennae lie in wait in flowers and attack bees and other insects.

Aradoidea. Flat bugs of the family Aradidae and their specialized relatives, Termitaphididae, have the mandibular and maxillary setae coiled in the clypeus. They lack ocelli and have four-segmented antennae. Most, if not all, are fungus feeders. The Termitaphididae have no wings and the margins of the body have numerous small, setigerous lobes. They are known from termite nests in the Old and New World tropics. Aradidae are nearly cosmopolitan and fully winged, brachypterous, stenopterous, or apterous. Some species are polymorphic, the males having fully developed or stenopterous wings and females being short-winged.

Pentatomorpha. This group includes the superfamilies Lygaeoidea, Pyrrhocoroidea, Coreoidea, and Pentatomoidea.

1. Lygaeoidea. This is the first of the superfamilies with ventral trichobothria on the abdomen. These erect bristles are in groups of three on either side near the middle of the third and fourth segments and lateral on the fifth, sixth, and seventh segments. The antennae are four-segmented, as is the beak. Ocelli are present. There are only four simple veins in the membrane. The Lygaeidae include the chinch bug, *Blissus leucopterus*, the false chinch bugs, *Nysius* spp., and the milkweed bugs,

Oncopeltus. The last is reared as a laboratory animal for experimental purposes, using milkweed seeds as food. The chinch bug, which is a pest of corn and wheat in the Middle West, has two generations per year. Adults overwinter in grass and trash. Control is by spraying with Dieldrin or Toxaphene, applying to bases of plants where the bugs congregate. Barrier strips of Dieldrin may also be used in strips 4 rods wide.

1. Small families related to the lygaeids are the thread-legged Neididae (Berytidae), the small Piesmidae, formerly included in Tingidae, and the tropical Colobathristidae. Some of the last have a unique method of stridulation with a filelike arch on the sides of the head.

2. Pyrrhocoroidea. This group includes the cotton stainers (*Dysdercus*), which attack cotton bolls in the southwestern United States and over most of the tropics, and stout, dark bugs with a reddish border (*Largus*). Pyrrhocorids resemble lygaeids and coreids but lack ocelli. *Dysdercus* may be controlled by a 5% DDT spray.

3. Coreoidea. The squash bugs and their relatives include the Coreidae, *Anasa tristis*; Rhopalidae, *Rhopalus* and *Leptocoris*; Alydidae, *Leptocorisa*; and the Hyocephalidae. They have four-segmented antennae, a beak, distinct ocelli, and many veins in the membrane. Coreids resemble Lygaeids in arrangement of the trichobothria. The squash bug overwinters as an adult. Eggs are laid on squashes and other cucurbits. There are five nymphal instars and usually only a single drawn-out generation per year. Control of these pests is either by DDT 5% dust applied to bugs on soil or under trap boards or by nicotine sulfate in 40% solution.

4. Pentatomoidea. This large group has marginal trichobothria. The antennae are usually five-segmented and the beak is four-segmented.

Scutelleridae, or shield bugs, are not injurious in the United States, but *Eurygaster* is a pest of grains in Russia, and some metallic green species are plant pests in the tropics.

Cydnidae include the ground-burrowing bugs *Amnestus*), which attack strawberries in sandy soil, and the negro bugs (*Corimelaena*).

Smaller families of little or no economic importance in the United States are the large tropical Tessaratomidae, the Acanthosomatidae, the Dinidoridae, the Aphylidae, the Urostylidae, the barklike Phloeidae, the shining, oval Plataspidae, and the Lestoniidae.

The true stink bugs, Pentatomidae, include the black-and-red harlequin cabbage bug, *Murgantia histrionica* (Hahn), and several green stink bugs, *Thyanta*, *Euschistus*, and *Chlorochroa*, on cotton and other crops. Control is by DDT as a dust or spray on the foliage or, on cotton, BHC or Dieldrin dust. Stink bugs overwinter as adults; they may produce two generations in a season. Feeding punctures on pears, which are attacked as the cover crop dries, cause cat-facing. One group, the Asopinae, includes predators (*Podisus*) on caterpillars. *See* ENTOMOLOGY, ECONOMIC; INSECTA. Pedro W. Wygodzinsky

Hemispheric laterality

The human brain is a bilaterally symmetrical structure which is for the most part richly interconnected by two main bridges of neurons called the corpus callosum and anterior commissure. These structures can be surgically sectioned in humans in an effort to control the spread of epileptic seizures. Although there is no apparent change in everyday behavior of these patients, dramatic differences in cognitive function can be demonstrated under specialized testing conditions. Because of these studies it now can be said that in normal humans these cerebral commissures are largely responsible for behavioral unity; the neural mechanism keeps the left side of the body up to date with the activities of the right side, and vice versa.

Results of tests involving speech. Changes in behavioral responses of persons whose cerebral commissures have been sectioned are almost undetectable. The person walks, talks, and behaves in a normal fashion. Dramatic effects are observed only under testing conditions which utilize stimuli that are lateralized exclusively to one hemisphere or the other. When visual stimuli are flashed to the left of fixation of the eye, all information is exclusively projected to the right hemisphere, and vice versa. Likewise, when the left hand manipulates objects that are held out of sight, all relevant sensory information projects to the opposite, right hemisphere. Similarly, information from objects held in the right hand is relayed exclusively to the left hemisphere. This is the normal state and permits the separate input-output testing of each half-brain.

Using these sensory avenues, remarkable "deconnection" effects are seen in split-brain conditions. For example, if a picture of an apple is flashed in the right visual field, the person describes the object normally. However, if the same picture is flashed in the left visual field, in the early days of postoperative testing the person denies that the stimulus was presented at all. Gradually, after many test sessions the person may have the impression that something was flashed, but is unable to say what. In brief, this disparity of recognition in the two sides of the visual field occurs because the information is projected to the right hemisphere, which is incapable of speech. Because the right hemisphere is now disconnected from the left, information arriving in the right hemisphere cannot be communicated by means of speech.

Results of tests not involving speech. When tests are used which do not require a spoken response, numerous mental abilities are observed to be present in the "disconnected" right hemisphere. If, for example, instead of being asked for a verbal response, the person is required with the left hand to retrieve from a series of objects the one most closely related to the flashed visual stimulus, good performance is seen. Even though the person is unable to describe a picture of an orange flashed to the left field, when the left hand searches through a field of objects placed out of view, it correctly retrieves the orange. If asked what the object is, the person would

be unable to identify it. Here again the left hemisphere controls speech but cannot solve the problem. The right hemisphere solves the problem but cannot elicit speech.

Right hemisphere. Through the use of this kind of test procedure, a variety of other mental abilities have been observed to go on in the separated right hemisphere. It can read simple nouns but not verbs. It can recognize some adjectives. It can recognize negative constructions but cannot differentiate constructions in the active or passive voice. It cannot pluralize nouns. In brief, it can manage some aspects of language only.

Left hemisphere. Despite its linguistic superiority, the left hemisphere does not excel the right in all tasks. Tests have demonstrated that in some specialized functions, such as arranging blocks to match a pictured design or drawing a cube in three dimensions, the right hemisphere is decidedly superior to the left.

Emotional reactions. In the area of emotional reactions there appears to be equal reactivity in the two hemispheres. Although a brain-bisected person cannot describe an emotional stimulus presented to the right hemisphere, an emotional response such as laughter is nonetheless elicited. In one instance, when a picture of a nude was flashed to the right hemisphere, the person said she did not see anything and then chuckled. When she was asked what the laughing was about, she said that it was "a funny machine." Here the right hemisphere clearly had enough ability to elicit laughter but not speech. The left hemisphere simply observed the changed emotional state and said that something must be funny, but did not know what.

Localization of the speech processor. Experimental psychological research developed a procedure commonly called dichotic listening. It is a process where two auditory messages are played at the same time—one going to the left ear and one going to the right ear. Previous research showed that under these test conditions normal subjects tend to repeat only the message presented to the right ear and to suppress the message presented to the left. It turns out this phenomenon is exceedingly dramatic in the split-brain person. In one study, when two phonemes (one to each ear) were presented to the split-brain persons, they were totally unable to name phonemes presented to the left ear. This was true despite the fact that the left ear sends projections to the left hemisphere that are not affected by split-brain surgery. Indeed, when a phoneme or any auditory signal is presented to the left ear alone, the split-brain person experiences no difficulty whatsoever in describine auditory event in detail. The suppression comes about only when both ears are simultaneously stimulated, and reflects some kind of dynamic process involving the two cerebral hemispheres that is not well understood.

A more surprising aspect of this study, however, was that the right hemisphere using its own manual response system could not respond at all to phonemes, the basic unit of speech perception. This finding is unexpected in light of the above described language capacities of the right hemisphere. Can the right hemisphere understand simple spoken and written words without being able to understand one of the basic building-block units that make it up? Perhaps the manner in which the right hemisphere comprehends spoken language is drastically different than the way the left processes its information.

Specificity of cortical circuitry. Tests have been conducted in which not all of the subjects have had their entire cerebral commissure sectioned. This is because it is now believed total commissure section is not necessary to stop the interhemispheric spread of some kinds of seizure activity. In neuropsychological testing, these partially sectioned persons showed dramatic breakdown in interhemispheric transfer. When the posterior part of the callosum is sectioned, visual aspects of the syndrome appear. When it is spared and more interior regions are cut, however, tactile and auditory communications are blocked, but not visual ones. It also appears that no fundamental reorganization of the interhemispheric transfer system takes place, since years after surgery these same deficits are present and are not compensated for in any way.

Individual variation and age of surgery. Another fact is that there would appear to be a large variation in the lateralized talents of each half-brain. Thus, while the right hemisphere frequently appears to have some language talent, it is clear that not all split-brain persons have language skills in the right hemisphere. Similarly, visual spatial skills, which are usually present exclusively in the right hemisphere, are frequently bilaterally represented and sometimes represented only in the left speech hemisphere. There is even some evidence that the commissure system itself varies in what is transferred where. In some cases an intact anterior commissure in those persons with a full callosal section will allow for the complete interhemispheric transfer of visual information. Yet, in other cases visual transfer is blocked with a section of only the posterior callosal area. One possible important factor is the age of the surgical section and the extent of early brain damage which was experienced by the person. *See* BRAIN; PSYCHOLOGY.

Michael S. Gazzaniga

Bibliography. M. S. Gazzaniga, *Nature's Mind*, 1994; J. B. Hellige, *Hemispheric Asymmetry: What's Right and What's Left*, 1993; M. Kosslyn, *The Wet Mind*, 1992, reprint 1995.

Hemlock

The genus *Tsuga* of the pine family, characterized by flattened needles with two white lines beneath the needlelike leaves, which have distinct short stalks. The cones are small and pendent. Eastern hemlock (*T. canadensis*) grows to a height of about 90 ft (27 m) and occurs in eastern Canada, the Great Lakes states, and the Appalachians. Minutely toothed leaves, some of the smaller ones growing upside

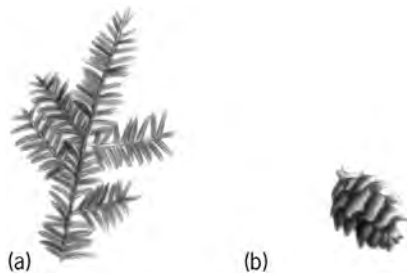

Eastern hemlock (*Tsuga canadensis*). (*a*) Branch. (*b*) Cone.

down, are characteristic of this species (see **illus.**). The wood is hard and strong, and is used for construction, boxes, crates, and paper pulp. The bark is one of the principal domestic sources of tannin. The eastern hemlock is a common ornamental tree. *See* PINALES.

Carolina hemlock (*T. caroliniana*), a species found in the southern Appalachians, has entire needles and is sometimes grown as an ornamental.

The western hemlock (*T. heterophylla*), which attains a height of 200 ft (60 m) or more, grows in the extreme Northwest and in Alaska. Its needles resemble those of the eastern hemlock, but the white lines beneath them are not so distinct. The greater part of the annual lumber cut comes from Washington. Much of this lumber cut goes into pulpwood for which it is a desired species. It is an important lumber tree, with uses similar to those of the eastern species. Another western species, *T. mertensiana*, is of lesser importance commercially. *See* FOREST AND FORESTRY; TREE. Arthur H. Graves; Kenneth P. Davis

Hemoglobin

The oxygen-carrying molecule of the red blood cells of vertebrates. This protein represents more than 95% of the solid constituents of the red cell. It is responsible for the transport of oxygen from the lungs to the other tissues of the body and participates in the transport of carbon dioxide in the reverse direction. More is known about the chemical structure and function of hemoglobin than any other protein.

Structure. Each molecule of hemoglobin comprises four smaller subunits, called polypeptide chains. These are the protein or globin parts of hemoglobin. A heme group, which is an iron-protoporphyrin complex, is associated with each polypeptide subunit and is responsible for the reversible binding of one molecule of oxygen. There are two different kinds of polypeptide subunits in the hemoglobin of the normal adult human (abbreviated Hb A). One is called the alpha chain and contains 141 amino acid residues. The other is called the beta chain and contains 146 amino acid residues. Two of each kind of polypeptide chain are arranged in the form of a truncated tetrahedron which has the overall shape of an ellipsoid of dimensions 5.5 × 5.5 × 7.0 nanometers (**Fig. 1**). The molecular formula of Hb A is $\alpha_2\beta_2$, and its molecular weight is 66,000. *See* AMINO ACIDS; PEPTIDE; PROTEIN.

Synthesis and metabolism. The polypeptide chains and the heme are synthesized and combine together in nucleated red cells of the bone marrow. As those cells mature, the nuclei fragment and the cells, now called reticulocytes, begin to circulate in the blood. After sufficient hemoglobin has been formed in the reticulocyte, all nuclear material disappears and the cell is then called an erythrocyte, or red blood cell. Each hemoglobin molecule lasts as long as the red cell, which has an average life of 120 days.

When the old red cell is destroyed, the heme is metabolized to iron, which is reused, and to porphyrin, which is degraded to bile pigments and excreted by the liver. The polypeptide chains are broken down to amino acids, which may then be reused in other metabolic processes, including the synthesis of new proteins. *See* BILIRUBIN; HEMATOPOIESIS; PORPHYRIN.

Heterogeneity and genetic control. Several different kinds of normal hemoglobins occur in humans. Embryonic hemoglobins are present in the first 2 months of intrauterine life. Normal fetal hemoglobin (Hb F) predominates during the rest of fetal life; it has a molecular formula of $\alpha_2\gamma_2$. The alpha chains are identical to those in Hb A, and the gamma chains are similar but not identical to the beta chains. Two types of gamma chains are normally produced. They differ from one another by only a glycine or alanine in position 136 of the gamma polypeptide chain. The gamma chains are not normally synthesized after birth, and Hb F disappears from circulation within the first 3–4 months of life. Another normal hemoglobin, Hb A_2, which has a molecular formula of $\alpha_2\delta_2$, makes up about 2–3% of the total hemoglobin in normal humans. Each type of polypeptide chain is determined by a separate genetic locus (**Fig. 2**). These genes determine the exact sequence of amino acid groups in the polypeptide chains. In addition to these genetically different hemoglobins, other minor components of hemoglobin are formed as the result of chemical modifications that occur after protein synthesis. One of the most interesting is Hb A_{Ic}, which results from the glycosylation of Hb A by glucose from the blood. The relative amount of Hb A_{Ic} is influenced by the concentration of blood sugar and the age of the red cell. *See* GENETIC CODE.

Mutations of these genes result in changes in the amino acid structure of the polypeptide chains. The abnormal gene responsible for sickle-cell anemia causes the normal glutamyl amino acid residue at position six of the beta chain of Hb A to be substituted by a valyl group. This single change out of the total 146 amino acid residues of the beta chain causes a severe, lifelong anemia in individuals who have two mutant genes. Almost 400 other abnormal human hemoglobins are known, but only about one-third are associated with anemia or other diseases. *See* HEMATOLOGIC DISORDERS; MUTATION.

The hemoglobins of different species are similar in molecular size and chemical properties. The corresponding polypeptide chains of closely related animals are very similar in their sequences of amino

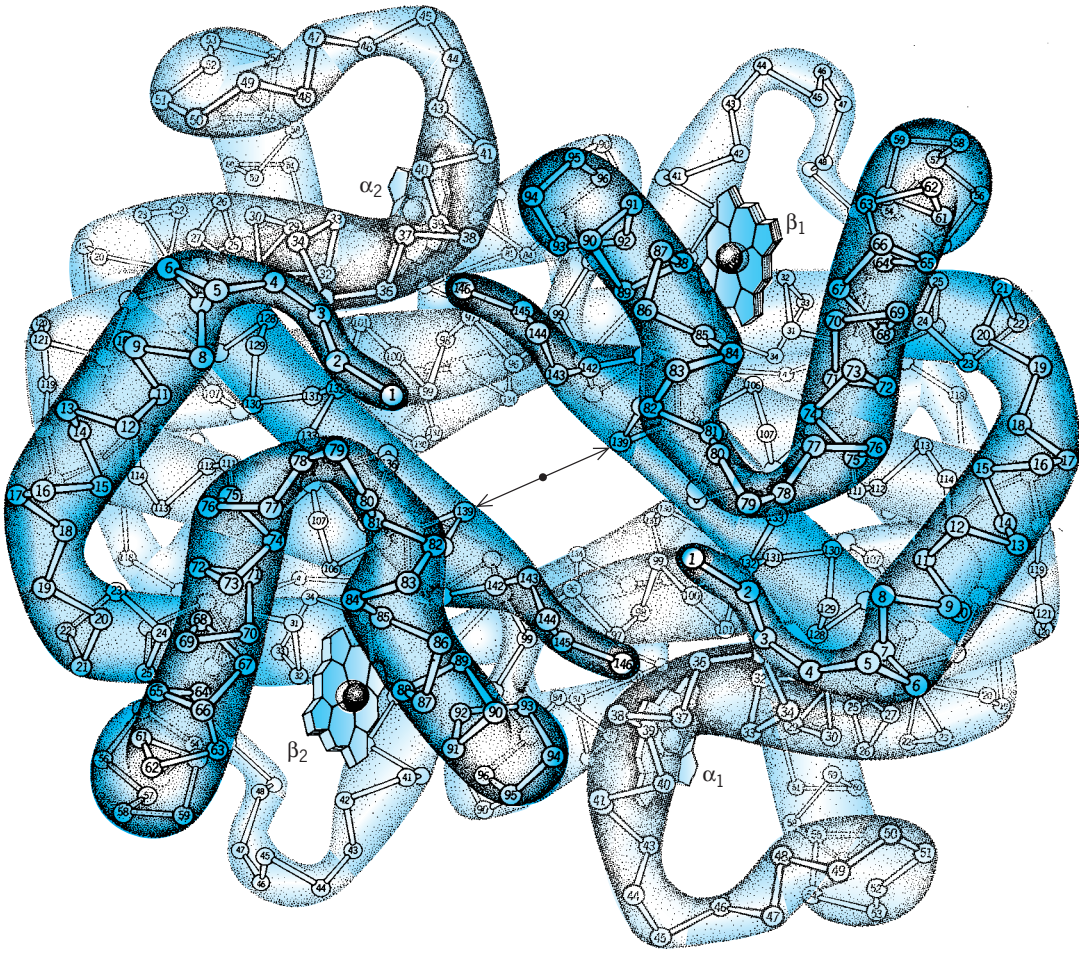

(a)

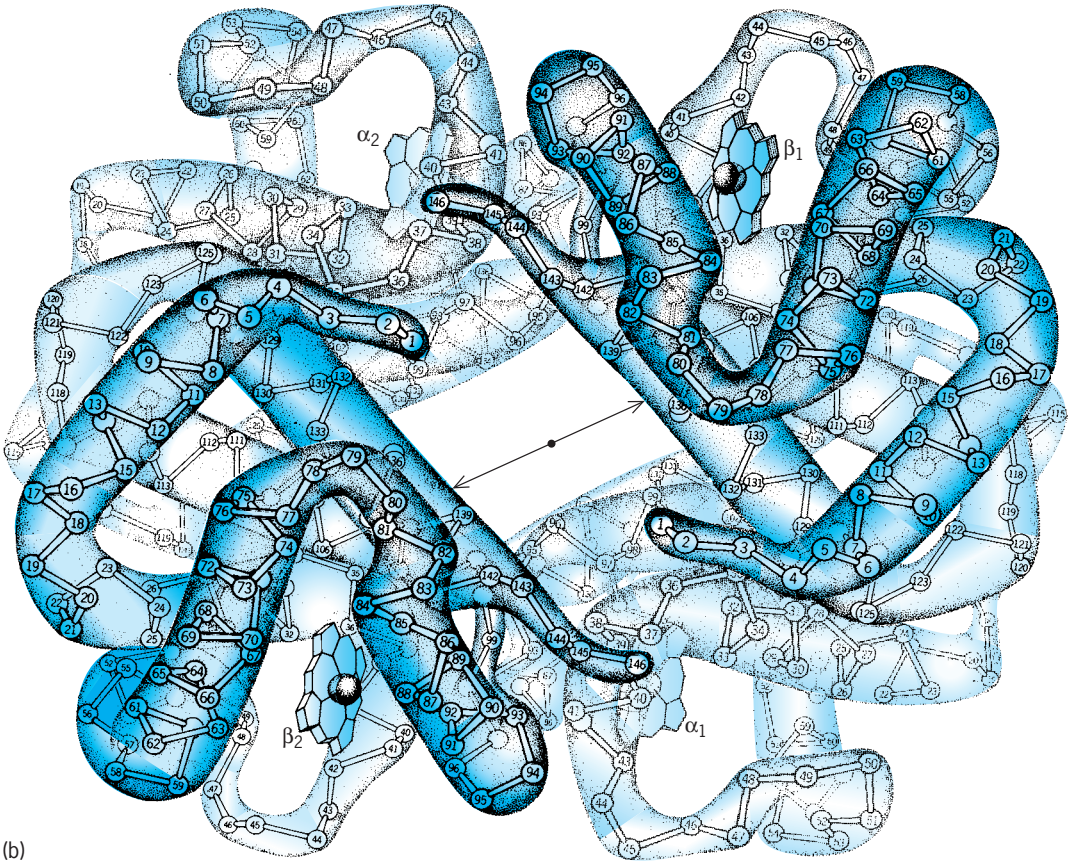

(b)

acid residues. Greater differences occur between hemoglobin of less related animals. Evolutionary differences between animals can be estimated by a comparison of the structure of their hemoglobins. *See* GENETICS.

Normal values and function. Normal adult males and females have about 16 and 14 g, respectively, of hemoglobin per 100 ml of blood; each red cell contains about 29×10^{-12} g of hemoglobin. Red cells normally comprise 40–45% of the volume of whole blood. Determination of these quantities is useful in diagnosis of hematologic disorders. Significant deviations from the values occur under altered conditions of red-cell production and destruction. *See* CLINICAL PATHOLOGY.

The reversible combination of hemoglobin and oxygen can be presented by the reaction shown below. The equilibrium constants for each step are

$$Hb + 4O_2 \overset{K_1}{\rightleftharpoons} HbO_2 + 3O_2 \overset{K_2}{\rightleftharpoons}$$

$$Hb(O_2)_2 + 2O_2 \overset{K_3}{\rightleftharpoons} Hb(O_2)_3 + O_2 \overset{K_4}{\rightleftharpoons} Hb(O_2)_4$$

not the same because an oxygen molecule on one heme group changes the affinity of the other hemes for additional oxygen molecules. This alteration in binding affinity during oxygenation is called heme-heme interaction and is due to small changes in the three-dimensional structure of the molecule. Some of the differences between the structure of oxygenated and deoxygenated hemoglobin are shown in Fig. 1. The principal change illustrated is a shift of the $\alpha_1\beta_2$ subunits relative to the $\alpha_2\beta_1$ subunits, resulting in an increase in the space between the beta chains in deoxyhemoglobin. These two structural conformations have different affinities for oxygen. The relationship between the pressure of oxygen and the oxygen content of a solution of hemoglobin can be graphed. The sigmoid curve obtained is a result of heme-heme interaction. Similar measurements of myoglobin, the oxygen storage molecule of muscle, reveal a rectangular hyperbola which indicates no heme-heme interaction (**Fig. 3**). Myoglobin is one-quarter the size and weight of hemoglobin and has only one heme per molecule.

Oxidation states. Hemoglobin combines reversibly with carbon monoxide about 210 times more strongly than with oxygen. This strong affinity for carbon monoxide accounts for the poisoning effects of this gas. The iron of each heme is normally in the ferrous state. It can be oxidized to the ferri state, in which case the hemoglobin, called ferrihemoglobin or methemoglobin, will no longer combine with either oxygen or carbon monoxide. Ferrihemoglobin will combine strongly with ions such as fluoride, cyanide, and hydroxide. The two oxidation states of hemoglobin and various chemical combinations can be determined by spectrophotometric measurements. The dark-bluish color of venous blood is due to the predominance of deoxygenated

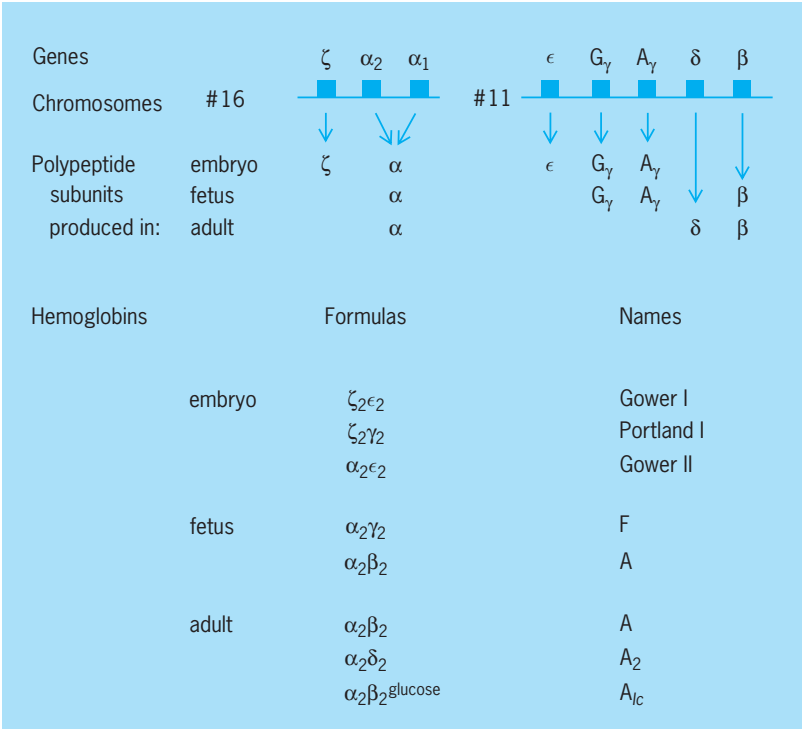

Fig. 2. The chemistry and genetics of hemoglobin. The ζ and two different α-globin genes are located on chromosome 16 (in humans). The other non-α-globin chain genes are located on chromosome 11. The two α genes produce the same α-globin polypeptide chain. The two γ genes make two slightly different polypeptide chains. The $\beta^{glucose}$ chain results from the glycosylation of the beta chains in Hb A, and occurs continuously at a slow rate throughout the life-span of the red blood cell.

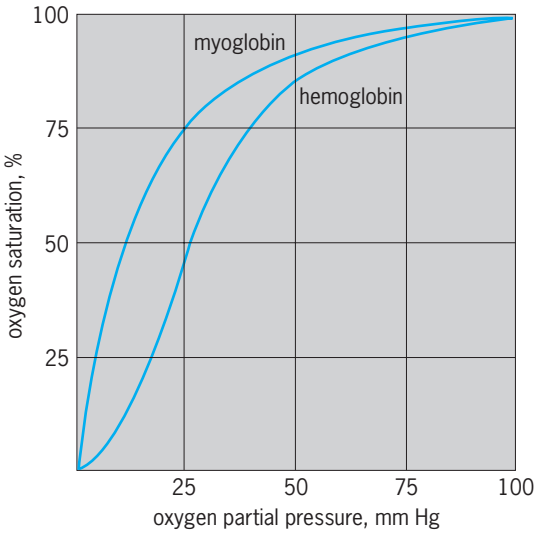

Fig. 3. Myoglobin and hemoglobin oxygen saturation curves. The sigmoid-shaped curve of oxygenation of hemoglobin results from several small changes in the structure of the hemoglobin molecule that occur as O_2 reacts with each heme group. 1 mmHg = 133.3 Pa.

Fig. 1. Three-dimensional structure of (*a*) oxyhemoglobin and (*b*) deoxyhemoglobin. The four polypeptide subunits are associated as two identical half-molecules, each containing one alpha and one beta chain. A single iron-containing heme group is associated with each globin chain subunit. The positions of the two alpha-beta half-molecules change relative to each other when oxygen combines with deoxyhemoglobin to form oxyhemoglobin. (*From R. E. Dickerson and I. Geis, Hemoglobin: Structure, Function, Evolution, and Pathology, Benjamin/Cummings, 1983; copyright by I. Geis*)

hemoglobin. The bright-red color of arterial blood is due to oxyhemoglobin. A cherry-red color results from the presence of carbonmonoxyhemoglobin.

Blood chemistry. Hemoglobin binds carbon dioxide by means of free amino groups of the protein but not by the heme group. The reversible combination with carbon dioxide provides part of the normal blood transport of this gas. Hemoglobin serves also as a buffer by reversible reactions with hydrogen ions. The acidic property of oxyhemoglobin is greater than deoxygenated hemoglobin. The extra binding of hydrogen ion by deoxyhemoglobin promotes the conversion of tissue carbon dioxide into bicarbonate ion and increases the amount of total carbon dioxide that can be transported by blood. *See* BLOOD; RESPIRATION. Richard T. Jones

Bibliography. H. F. Bunn and B. G. Forget, *Hemoglobin: Molecular, Genetic and Clinical Aspects*, 1986; A. D. Stephens (ed.), *Haemoglobin: Research and Applications*, 1987.

Hemophilia

A rare, hereditary blood disorder marked by a tendency toward excessive bleeding. It is almost entirely restricted to males, and is transmitted as a sex-linked mendelian recessive trait passing from an affected male through an unaffected or very mildly affected daughter to appear again in a grandson. Queen Victoria was a carrier, and several of her male descendants were affected. *See* HUMAN GENETICS; SEX-LINKED INHERITANCE.

Classical hemophilia (hemophilia A) is due to a deficiency of the antihemophilic factor or factor VIII, a clotting factor which is normally present in the blood in trace amounts and is essential for normal fibrin formation. Hemophilia B is a similar sex-linked bleeding disorder affecting males but characterized by a deficiency of another blood clotting factor, factor IX. It can only be distinguished from hemophilia A, which it closely resembles, by laboratory tests. Hereditary deficiencies of other clotting factors may give rise to bleeding disorders similar to hemophilia, but they are inherited as autosomal dominant or recessive characteristics. Most of the hereditary hemorrhagic disorders, including both types of hemophilia, have been observed in animals. Most of the clinical manifestations follow trauma. Even a slight blow may produce severe hemorrhage into the affected body tissues. Bleeding into the joints is seen in the most severe cases, and may result in crippling deformities. Bleeding into muscles and subcutaneous tissue is common, while bleeding from the mucous membranes occurs from time to time. There are large numbers of subtypes of each disorder, which vary markedly in clinical severity. Thus the clinical spectrum ranges from individuals who bleed excessively only following major hemostatic challenges and may live a normal life to those who have crippling deformities. Potent concentrates of factor VIII or IX prepared from human plasma are available, and are very effective in the management of hemophilia A and B, respectively. They must be given intravenously. When given in adequate amounts, they can restore the hemostatic potential of the blood to normal for many hours. *See* BLOOD; HEMORRHAGE. Cecil Hougie

Bibliography. S. Chein et al. (eds.), *Clinical Hemorheology*, 1987; W. J. Williams et al., *Hematology*, 3d ed., 1983.

Hemorrhage

The escape of blood from within the vascular system. Hemorrhage may result from either trauma or disease of the vessel wall.

Causes. The escape of blood following rupture of a vessel wall as a result of trauma is obvious and needs no further explanation. The causes other than trauma can be divided into three main groups.

The first group consists of these conditions in which there is a chronic disease process affecting the vessel wall, such as atherosclerosis or aneurysm formation. Either of these conditions, in association with an elevated blood pressure, can result in a break in the wall and subsequent hemorrhage. An infarct, or tissue death from any cause, may also result in hemorrhage. *See* ANEURYSM; ARTERIOSCLEROSIS; INFARCTION.

The second group consists of those causes in which there is an acute process affecting the vessel wall, such as septicemia, bacterial toxins, or anoxia. *See* HYPOXIA.

The third group consists of those hemorrhagic conditions which result from some defect in the blood itself. Under this heading are leukemia, thrombocytopenia, and the clotting disorders. *See* LEUKEMIA.

Petechiae are hemorrhages no larger than the head of a pin. Hemorrhages of greater size are termed ecchymoses. A localized mass of blood in tissue is a hematoma. Spontaneous hemorrhaging into the skin and mucosal surfaces is termed purpura. This usually denotes a disease of the vascular system or of the blood itself, such as a deficiency of blood platelets. Platelets are thought to be necessary to maintain capillary integrity.

Cerebrovascular accident. Cerebrovascular accident, or stroke, is an acute vascular lesion of the brain. This may be the result of hemorrhage from, thrombosis in, or embolism to a cerebral vessel. Cerebral hemorrhage can result from a rupture of the vessel wall, as from a defect in the wall such as a congenital aneurysm, or it may be secondary to death of cerebral tissue as occurs with an infarct. Intracerebral hemorrhages may be small petechial or perivascular hemorrhages, or they may be massive. Small hemorrhages can result from minimal trauma, poisoning such as from carbon monoxide, or from blood diseases such as leukemia. Massive cerebral hemorrhage is frequently the result of trauma or vascular disease. With cerebral hemorrhage or death of cerebral tissue there is loss of function of the regions involved. One such complication is paralysis of the muscles of the extremities on the side opposite the

lesion. Since the brain is enclosed in a nondistensible box, the skull, the increase in size due to the addition of the blood raises intracranial pressure. *See* EMBOLISM; THROMBOSIS.

Compensatory mechanisms. Following the sudden loss of a large quantity of blood, the body reacts to compensate for this loss and the blood clotting mechanism is activated. As the blood volume is decreased, the carotid sinus reflex comes into play. This results in a generalized vasoconstriction which affects the arteries and precapillary and capillary vessels. With a lowering of capillary pressure, there is a shift of fluid from the tissue spaces into the plasma. This tends to restore the blood volume but lowers the hematocrit (the relative percentage of erythrocytes, or red cells, in the blood). Following a loss of 500 ml of blood, a healthy person can restore blood volume in 24 h. The protein content can be restored in a few hours, but the replacement of the cellular elements takes days or even weeks. Failure of these mechanisms to compensate adequately can result in shock with dire consequences. *See* CARDIOVASCULAR SYSTEM.

Morphologic changes of hemorrhages. Small hemorrhages can be completely absorbed leaving no tissue alterations; larger volumes of blood act as a foreign substance which is destroyed and removed by the phagocytes and may result in scarring. The red cells are broken down, liberating hemoglobin. This released hemoglobin is divided into two moieties, the iron-free hematoidin and the iron-containing hemosiderin.

In animals dying of severe or fatal hemorrhage a state of severe ischemia of all the organs is found. The mucous and serous surfaces are pale; the organs are drier than normal and reduced in weight. *See* CIRCULATION DISORDERS; HEMATOLOGIC DISORDERS.

Romeo A. Vidone; Irwin Nash

Bibliography. J. M. Kissane (ed.), *Anderson's Pathology*, 9th ed., 1989; R. S. Cotran et al., *Robbins' Pathologic Basis of Disease*, 6th ed., 1999; W. J. Williams et al., *Hematology*, 3d ed., 1983.

Hemp

The fiber and the plant *Cannabis sativa*. It should not be confused with Manila hemp, which is not related to true hemp. *See* ABACA.

Hemp fiber, which for many years was the major raw material used in the manufacture of rope, now is used mostly in the production of small twines, linenlike fabrics and canvases, and, to some extent, in making special types of paper. *See* NATURAL FIBER.

The plant is about 6–8 ft (2–2.5 m) tall at maturity and has a stem which at midpoint is roughly the diameter of a lead pencil.

Hemp is an annual crop, most of which is produced in Eastern Europe and mainland China, with some production in South Korea, Turkey, and Italy. World production has declined over the years. In the United States, hemp fiber production gradually declined to none by the mid-1950s. Hemp seed production in the United States was centered in Kentucky and fiber production in the upper Midwest.

Hemp is planted about the same time as corn and requires about 4 months of frost-free weather to produce a crop of fiber. Hemp needs a fertile soil and good drainage. The fiber is separated by crushing the retted (partially rotted) stems and beating out the nonfibrous material.

Hemp contains the drug marijuana, and a permit from the Drug Enforcement Administration of the Department of Justice is required to grow or possess any part of this plant in the United States. *See* MARIJUANA.

Elton G. Nelson

Henequen

A fiber obtained from the leaves of *Agave fourcroydes*. It is produced only in Mexico, Cuba, and El Salvador. Henequen is sometimes incorrectly called sisal, which is a closely related plant grown in Brazil and Africa. *See* SISAL.

A well-developed henequen plant has a short stem about 12–18 in. (30–45 cm) in diameter and 3–5 ft (1–1.5 m) in height at maturity. Leaves are 3–6 ft (1–2 m) in length and 4–6 in. (10–15 cm) wide near the base. They are gray-green, thick, succulent, and smooth, with many small, curved spines on the edges and a sharp terminal spine 0.4–1.2 in. (1–3 cm) long. When the plant is 15 to 25 years old, a flower stalk grows up through the top to a height of 13–26 ft (4– 8 m) and develops terminal branches which bear flowers near their tips. The flowers are followed by seed or by bulbils or sometimes by both.

Cultivation. The henequen plant requires a dry, tropical climate and a well-drained soil, preferably of limestone formation. The average annual rainfall in Yucatan, the henequen-growing area of Mexico, is about 30 in. (750 mm) and the temperature is seldom below 59°F (15°C). Many of the plantations there are rocky, and land preparation consists of removal of trees, shrubs, and some of the native vegetation before setting out the plants. Where conditions permit, land for planting should be plowed and cultivated.

Henequen plants are propagated from bulbils or more generally from suckers. Bulbils must be cultivated in a nursery for 1 or 2 years before they are set out in the field. Suckers are dug when 16–24 in. (40–60 cm) in height and, after some of the top is cut back, are set out directly in the field. Plantations are usually in double rows about 3 ft (1 m) apart and spaced about 4 ft (1.25 m) apart in the row, with about 11 to 13 ft (3.5–4 m) between each pair of rows. Some planting is in single rows.

Diseases. Henequen is affected by fewer diseases than most plants. Disease symptoms are frequently caused by nutrient deficiencies. The fungus *Colletotrichum agaves* may attack the leaves and cause leaf spots, especially if they have been punctured by insects. Infestation by the sisal weevil, *Scyphophorus interstitialis,* which bores into the bud, can be

reduced by field sanitation and by spraying with aldrin or dieldrin.

Harvesting and processing. The first cutting of henequen leaves for fiber is in the sixth or seventh year. Successive harvests are at 6- to 12-month intervals for periods of 10 to 20 years. About 15 to 20 leaves at an angle of 40° or more from the vertical are cut each time. The terminal and marginal spines are trimmed off, and the leaves are tied in bundles and taken to a defibering machine where they are beaten and scraped. The fiber obtained is usually dried in the sun, but sometimes dryers are used.

Uses. The greatest quantity of henequen fiber goes into farm twine, followed by industrial tying twine and then light-duty rope. Padding for innerspring mattresses is made from the lowest grades of fiber and from flume tow, the short, tangled fiber that can be recovered from the cleaning operation. Henequen is exported as manufactured twine, rope, or padding, not as raw fiber. *See* NATURAL FIBER. Elton G. Nelson

Bibliography. *Kirk-Othmer Encyclopedia of Chemical Technology*, vol. 10, 3d ed., 1980.

Heparin

A highly sulfated mucopolysaccharide isolated from mammalian (chiefly beef) tissues, with blood anticoagulant activity. Heparin was first found in abundance in the liver, hence the name, but it is present in substantial amounts in the spleen, muscle, and lung as well. Chiefly composed of D-glucosamine and D-glucuronic acid residues, it usually has a molecular weight from 6000 to 20,000. It normally occurs as the sodium or potassium salts and undergoes some degree of desulfation when the free acid is liberated. Full desulfation of heparin yields heparamine, which on treatment with carboxylic acid anhydrides furnishes heparides, that is, amides of the polysaccharide, free of anticoagulant properties but still valuable for treating hyperlipidemia. In the blood of most mammals, heparin is an antagonist to thrombin, prothrombin, and thromboplastin. It lessens the tendency of platelets to agglutinate. When it is injected, its action is almost instantaneous, but it begins to decline rapidly in 2–3 h. Heparin is only effective when injected, and the site of the injection often becomes swollen or inflamed and causes pain. It is used in the treatment of venous thrombosis, embolism, myocardial infarction, and certain types of cerebral thrombosis. Heparin is the single most important anticoagulant used in medicine. *See* POLYSACCHARIDE. Frank Wagner

Bibliography. R. A. Bradshaw and S. Wessler (eds.), *Heparin-Structure: The Function and Clinical Implications*, 1975; W. D. Comper, *Heparin (and Related Polysaccharides): Structural and Functional Properties*, 1981; N. M. McDuffie (ed.), *Heparin: Structure, Cellular Functions, and Clinical Applications*, 1979; I. Witt (ed.), *Heparin: New Biochemical and Medical Aspects*, Proceedings of the Symposium of the Deustche Gesellschaft fur Klinishe Chemie, 1983.

Hepaticopsida

A class of lower green plants called liverworts that belong to the division Bryophyta. The class is divided into approximately 225 genera and 8500 species. Although there is a great diversity of external form, most of the gametophytes (gamete-producing plants) are dorsoventrally differentiated. These plants are considered among the most primitive of the existing land plants. Fossil remains of liverworts have been found in both the Devonian and Carboniferous. Since the fossils found do not differ significantly from modern liverworts, they are of little value in ascertaining phylogenetic relationships. Liverworts are widely distributed over the world, but have their greatest diversity in the tropics of the Americas and East Indies.

Except when the plants occur in masses, they are quite inconspicuous and are usually confused with mosses, which they resemble somewhat in their external appearance. In the presence of adequate moisture, they grow on soil, rocks, and tree trunks. Usually the plant body is a thin, prostrate thallus, sometimes having a short central axis with leaflike appendages. On the lower surface are rhizoids (rootlike structures) which function in anchorage and absorption. *See* GEOLOGY.

Reproduction. The sex organs, antheridium (male) and archegonium (female), may be produced on the same plant or on different plants. The sperms, produced in the antheridia, are flagellate and motile. The sporophyte (spore-producing generation), which develops as a result of fertilization, remains on the female gametophyte, and for a time is parasitic. The

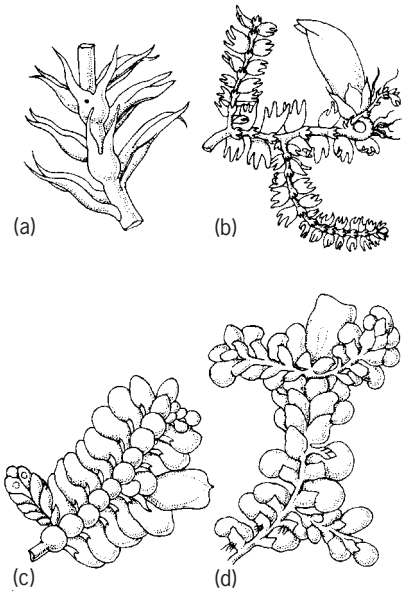

Fig. 1. Jungermanniidae, the leafy liverworts. (*a*) *Herberta*, showing three ranks of equal bifid leaves. (*b*) *Lepidozia*, ventral aspect, showing reduced ventral leaves. (*c*) *Leucole jeunea*, ventral aspect, showing reduced ventral leaves and complicate dorsal leaves. (*d*) *Radula*, ventral aspect, showing absence of ventral leaves but complicate dorsal leaves. (*After E. W. Sinnott and K. S. Wilson, Botany: Principles and Problems, 5th ed., McGraw-Hill, 1955*)

spores are developed within a sporangium, or capsule, which lacks a sterile region or columella internal to the spore-producing tissue as occurs in the mosses. After discharge from the sporangium the spores germinate, forming short protonemas from which arise new liverwort plants (gametophytes). *See* BRYOPSIDA.

Classification. The Hepaticopsida are divided into two subclasses, the Jungermanniidae and the Marchantiidae.

The Jungermanniidae, with about 190 genera and 8000 species, are called the leafy liverworts because of their chlorophyll-containing ribbonlike or leaflike bodies (**Fig. 1**). The thallus is never differentiated, and it lacks the air spaces and external pores characteristic of the Marchantiidae. Only unicellular, smooth-walled rhizoids are present. The sporophyte often produces a long stalk (sporangiophore) bearing the sporangium, or capsule, on the upper end. When mature the sporangium dehisces (opens) by four valves. The Jungermanniidae are composed of four orders: Takakiales, Calobryales, Jungermanniales, and Metzgeriales.

The Marchantiidae with about 35 genera and 450 species, are called thallose liverworts, the best-known example being *Marchantia polymorpha*. The flat plant body is composed of several distinct layers of tissue, having the chlorophyll-containing cells of the uppermost layer separated by numerous air chambers which are exposed to the external atmosphere through pores. Rhizoids are of two kinds, smooth-walled and tuberculate-walled. The male and female sex organs, each borne on separate plants, are generally found grouped in the tops of special stalked structures known respectively as antheridiophores and archegoniophores. The Marchantiidae are composed of three orders: Sphaerocarpales, Monocleales, and Marchantiales.

Alternation of generations. The life cycle of liverworts is well illustrated by *Marchantia* (**Fig. 2**). The gametophyte, or gamete-producing generation,

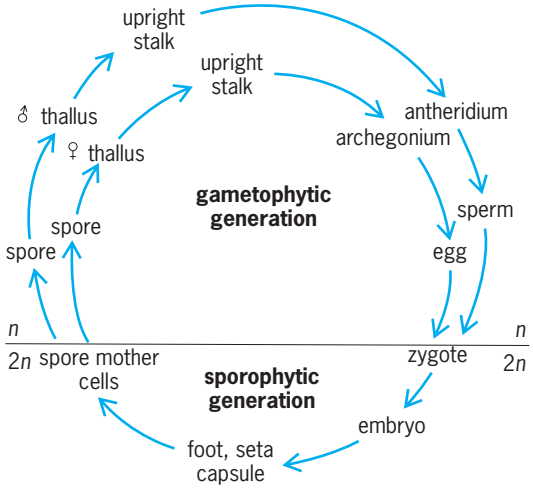

Fig. 2. Diagram of life cycle of the liverwort *Marchantia*, indicating stages at which chromosome changes occur. (*After H. J. Fuller, The Plant World, rev. ed., Holt, 1951*)

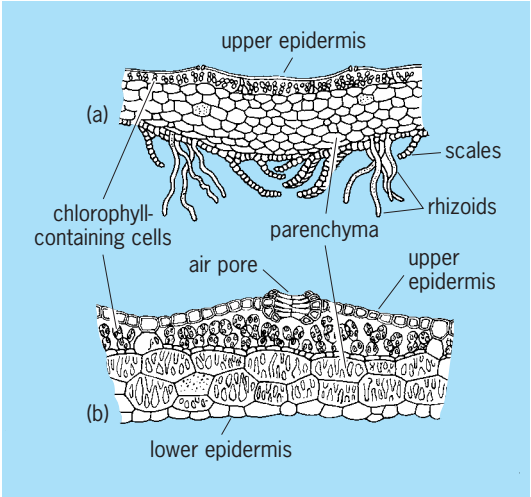

Fig. 3. Transverse sections through the thallus of *Marchantia*. (*a*) Middle portion with scales and rhizoids on the underside. (*b*) Margin of the thallus more highly magnified. (*After H. Kraemer, Applied and Economic Botany, 2d ed., Wiley, 1916*)

is the dominant phase. In *Marchantia*, the gametophyte is a flat, dichotomously branched (forked) thalloid structure produced by apical growth of the thallus. As development proceeds, the thallus produces numerous rhizoids from the ventral (under) surface. The internal anatomy shows considerable differentiation of tissues (**Fig. 3**). Starting with the upper surface, there is a single epidermal layer, subtended by an area of air chambers which are connected to the external atmosphere through regularly spaced pores or openings. The pores are surrounded by four rows of four cells each, the upper and lower circles being somewhat smaller, giving the whole structure a barrel-shaped appearance. Below these openings, the vertical pillars of cells and the multibranched filaments at the base of the chambers have numerous oval chloroplasts which give the rich, dark-green color so characteristic of the plant. The remainder of the thallus is made up of densely arranged parenchymatous cells. The lower parenchyma tissue has few chloroplasts, but it contains numerous large mucilage cells which probably function as storage tissue. *See* EPIDERMIS (PLANT); PARENCHYMA.

The plants are dioecious; that is, they have male and female sex organs on different plants. The antheridia, or male sex organs, are produced on stalked, disk-headed branches called antheridiophores which grow upward from the apical notch of the thallus or gametophyte (**Fig. 4**). The antheridia are embedded in cavities of the disk, each cavity having only one antheridium which, when mature, discharges the numerous biflagellated motile sperms to the exterior through a surface pore. Maturation of antheridia proceeds from the center of the antheridial disk outward and is often indicated by the cell walls becoming deep purple in color.

The archegonia are found on similar stalks called archegoniophores (**Fig. 5**). At the apex of each

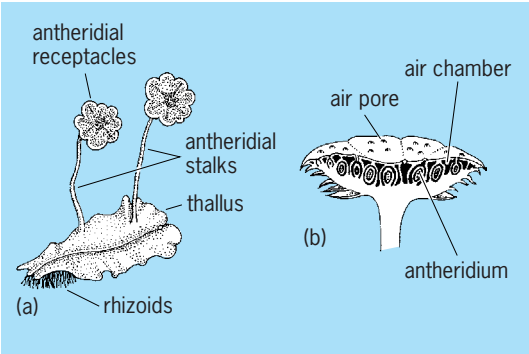

Fig. 4. ***Marchantia*. (*a*) Example of male gametophyte (antheridial plant) (*after H. J. Fuller and O. Tippo, College Botany, Holt, 1949*). (*b*) Section through an antheridial receptacle (*after W. J. Robbins and H. W. Rickett, Botany, 3d ed., Van Nostrand, 1939*).**

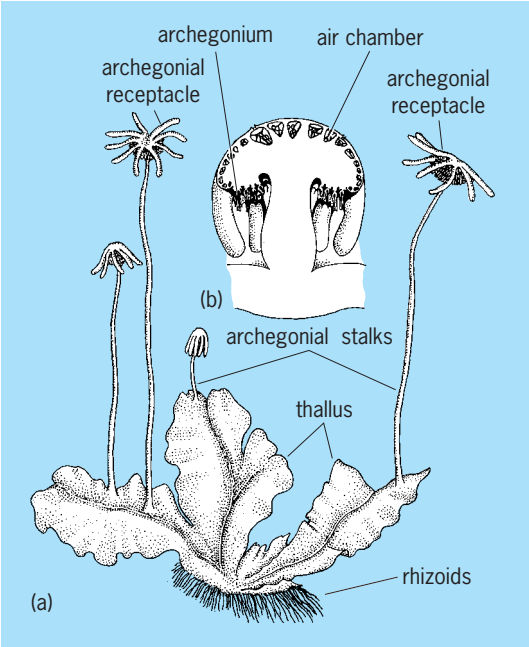

Fig. 5. ***Marchantia*. (*a*) Example of female gametophyte (archegonial plant) (*after H. J. Fuller and O. Tippo, College Botany, Holt, 1949*). (*b*) Section through an archegonial receptacle (*after W. J. Robbins and H. W. Rickett, Botany, 3d ed., Van Nostrand, 1939*).**

archegoniophore about nine finger-shaped structures project outward like the ribs of an umbrella. The archegonia are formed in groups near the bases of these arching projections. Each projection may produce 12–16 archegonia in succession from its tip inward.

There has long been some question as to how the sperms reach the egg cells in the archegonia. It has been observed that raindrops may splash the sperm as much as 2 ft (60 cm). Mites have been observed on the archegoniophore with live sperm on their bodies. Following a rain, or with heavy dew, the water film is probably adequate to allow the sperms to swim from the antheridia to the archegonia.

As a result of fertilization a zygote is formed which develops into a sporophyte (**Fig. 6**). This is accomplished by a series of divisions of the zygote (fertilized egg) which finally results in a structure having a foot embeded in the gametophyte; an elongated seta, or stalk, which shows its most rapid development at the time of spore maturation; and a sporangium, or capsule, within which the spores are developed following a meiotic division of the spore mother cells.

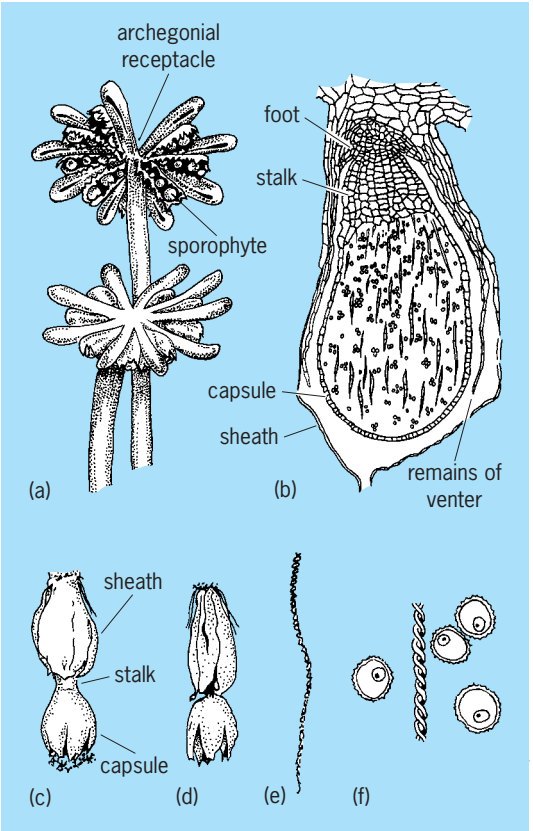

Fig. 6. ***Marchantia* sporophyte morphology. (*a*) Mature archegonial receptacles of sporophytes (*after W. J. Robbins and H. W. Rickett, Botany, 3d ed., Van Nostrand, 1939*). (*b*) *Marchantia polymorpha*, longitudinal section showing the internal structure of a nearly mature sporophyte; (*c*, *d*) surface views of sporophytes during and after shedding of spores and elaters; (*e*) portion of an elater; (*f*) spores and portion of an elater (*after G. M. Smith, Cryptogamic Botany, vol. 2, 2d ed., McGraw-Hill, 1955*).**

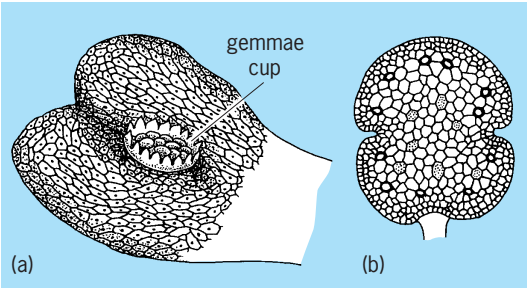

Fig. 7. Structures involved in vegetative reproduction of *Marchantia*. (*a*) Portion of surface of the thallus. Note gemma cup with gemmae (*after H. J. Fuller and O. Tippo, College Botany, Holt, 1949*). (*b*) Notched, lens-shaped gemma of *M. polymorpha* (*after G. M. Smith, Cryptogamic Botany, vol. 2, 2d ed., McGraw-Hill, 1955*).

Special pointed elongated cells, known as elaters, aid in the dispersal of spores by being sensitive to internal water tensions which cause them to coil and uncoil. Spores falling upon a suitable substrate germinate immediately and produce new gametophytes. *See* MEIOSIS.

In *Marchantia* there are also two methods of vegetative reproduction. The first results from the growth habit of the plant. As it grows and branches, the older portions of the plant decay, thus separating the branches into individual plants. The second method is in the formation of gemmae cups on the dorsal surface (**Fig. 7**). Within each cup several small, notched, lens-shaped bodies are produced which, when detached, grow from both ends simultaneously and develop into new plants. This provides for rapid dispersal during the growing season. *See* BRYOPHYTA; CALOBRYALES; JUNGERMANNIALES; JUNGERMANNIIDAE; MARCHANTIALES; MARCHANTIIDAE; METZGERIALES; MONOCLEALES; SPHAEROCARPALES.

Paul A. Vestal

Hepatitis

An inflammation of the liver caused by a number of etiologic agents, including viruses, bacteria, fungi, parasites, drugs, and chemicals. All types of hepatitis are characterized by distortion of the normal hepatic lobular architecture due to varying degrees of necrosis of individual liver cells or groups of liver cells, acute and chronic inflammation, and Kupffer cell enlargement and proliferation. There is usually some degree of disruption of normal bile flow, which causes jaundice. The severity of the disease is highly variable and often unpredictable. *See* LIVER.

Viral hepatitis. The most common infectious hepatitis is of viral etiology. Four types of viral hepatitis are recognized: type A (infectious) hepatitis; type B (serum) hepatitis; non-A, non-B (NANB) hepatitis; and delta hepatitis. The clinical course of viral hepatitis is highly variable, ranging from nearly asymptomatic disease to a severe prolonged illness capable of causing death. Signs and symptoms of viral hepatitis are nonspecific and consist of headache, fatigue, malaise, loss of appetite, nausea, jaundice, dark urine, and pain in the upper abdomen. A fever as high as 104°F (40°C) can occur in many with the disease.

Type A viral hepatitis is caused by a small ribonucleic acid (RNA) virus belonging to the picornavirus class. The disease is usually spread by the oral-fecal route, is associated with overcrowding and poor hygiene, and has an incubation period of 15–45 days. It occurs primarily in children and young adults, and is a usually mild disease leading to no residual effects. The virus is usually not detectable in the blood, but an antibody against it is usually detectable by serologic methods and can confirm the diagnosis.

Type B viral hepatitis is caused by a deoxyribonucleic acid (DNA) virus that belongs to a class referred to as hepadna viruses. In 1963 B. Blumberg discovered in an Australian aborigine an antigen which was initially called Australia antigen and, later, hepatitis-associated antigen, and which was subsequently shown to be part of the virus which causes type B hepatitis. The intact virion is called the Dane particle, and has been identified with the electron microscope. Hepatitis B is usually parenterally transmitted, has an incubation period of 50 to 160 days, occurs at any age, and is common in drug addicts. The infectious particle of hepatitis B (the Dane particle) is found in a variety of body fluids of those infected, including saliva, tears, seminal fluid, cerebrospinal fluid, ascites, breast milk, gastric juice, urine, and feces. Hepatitis B virus can be transmitted sexually and through the placenta during the birth process. Type B viral hepatitis has a mortality rate of up to 15%. In Asia and Africa, hepatitis B appears to be related to the development of liver cancer. It may proceed to chronic hepatitis in 10% of the cases. In addition, viral immune complexes of hepatitis B antigens and antibodies have been implicated as a causative agent in several relatively rare conditions—polyarteritis nodosa, membranous glomerulonephritis, and mixed cryoglobulinemia. There is an effective hepatitis B virus vaccine that is in widespread use in medical personnel and others who are at an increased risk. *See* ANTIBODY; ANTIGEN; IMMUNE COMPLEX DISEASE.

Non-A, non-B hepatitis. Liver inflammation referred to as non-A, non-B hepatitis is most common among people who have received transfused blood and whose serologic tests show no evidence of hepatitis A, hepatitis B, or other types of virus such as Epstein-Barr. The causative virus can also be found in clotting factor concentrates and is detected with relative frequency in hemophiliacs who receive large amounts of blood-clotting concentrates. *See* ANIMAL VIRUS; EPSTEIN-BARR VIRUS; HEMOPHILIA.

Delta hepatitis. Delta hepatitis, the most recently recognized hepatitis agent, is caused by the delta agent hepatitis D virus (HDV), a defective ribonucleic acid (RNA) virus that requires the helper function of hepatitis B virus for its replication and expression. Delta hepatitis has a worldwide distribution, but it occurs primarily in northern Africa, southern Europe, and the Middle East.

Alcoholic hepatitis. Another frequently occurring form of hepatitis is caused by excessive ethyl alcohol intake and is referred to as alcoholic hepatitis. It usually occurs in chronic alcoholics and is characterized by fever, high white blood cell count, and jaundice. Microscopically, the hepatic lobular architecture is disrupted by fatty change, acute inflammation with necrosis, and peculiar intracytoplasmic aggregates of keratin proteins called alcoholic hyaline. This form of hepatitis is reversible if the etiologic agent, ethyl alcohol, is withdrawn. Persons who develop alcoholic hepatitis are often misdiagnosed, with obstructed bile ducts being a common assumption, and some undergo exploratory surgery for that condition. *See* ALCOHOLISM.

Drug-induced hepatitis. Some drugs are capable of damaging the liver and can occasionally cause enough damage to produce clinical signs and

symptoms. The injury can be so severe that acute liver failure results. Among those drugs known to damage the liver are tetracycline, methotrexate, anabolic and contraceptive steroids, phenacetin, halothane, chlorpromazine, and phenylbutazone. Usually the damage caused by those drugs is relatively mild, and complete reversal follows cessation of medication. Many patients present with jaundice, but only after a careful history is taken will a drug-induced cause be found or considered. *See* ALLERGY; ANALGESIC; ANESTHESIA; ANTIBIOTIC; CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; STEROID.

Chronic hepatitis. Chronic hepatitis is a condition defined clinically by evidence of liver disease for at least 6 consecutive months. It may be caused by a prolonged bout of hepatitis B or non-A, non-B hepatitis. Persons with chronic hepatitis frequently develop nonspecific signs and symptoms such as mild fatigue and abnormal liver function tests. The disease is categorized pathologically into chronic aggressive and chronic persistent forms. The distinction is important, because the chronic aggressive form may progress to cirrhosis and require treatment with corticosteroids. *See* CIRRHOSIS.

Signs and treatment. Clinical features of hepatitis include malaise, fever, jaundice, and serum chemical tests revealing evidence of abnormal liver function. In most mild cases of hepatitis, treatment consists of bedrest and analgesic drugs. In those individuals who develop a great deal of liver cell necrosis and subsequently progress into a condition known as hepatic encephalopathy, exchange blood transfusions are often used. This is done with the hope of removing or diluting the toxic chemicals thought to be the cause of this condition. *See* LIVER DISORDERS.

Samuel P. Hammar

Bibliography. R. J. Gerety, *Hepatitis B*, 1985; G. Gitnick (ed.), *Modern Concepts of Acute and Chronic Hepatitis*, 1988; I. D. Gust and S. M. Feinstone (eds.), *Hepatitis A*, 1988; Y. F. Liaw (ed.), *Chronic Hepatitis*, 1986; R. E. Sampliner, *Preventing Viral Hepatitis*, 1988; T. Shikata, R. H. Purcell, and T. Uchida, *Viral Hepatitis C, D, and E*, 1991; A. J. Zuckerman, *Viral Hepatitis and Liver Disease*, 1988.

Herbarium

A collection of pressed and dried plant specimens, and a description of when, where, and by whom they were collected, arranged in a systematic manner, and serving as a permanent physical record of the occurrence of an individual plant at a specific place and time. Herbaria may contain specimens from the full range of organisms that have classically been considered plants: fungi, lichens, algae, bryophytes, ferns and their allies, gymnosperms, and angiosperms. Many herbaria also accumulate and manage special collections such as liquid-preserved parts for anatomical studies, wood, seeds, or specially preserved material suitable for extraction of deoxyribonucleic acid (DNA) or other chemical constituents. Many groups of plants, especially those with succulent or fleshy parts, are not suitable for preservation as dry, flat specimens because they lose many of their important features in the drying process. Consequently, these plants are often preserved in liquid.

Purpose and uses. A herbarium is basically a library of physical specimens, readily accessible for reference or detailed study. The information embodied in each specimen and its label forms an important resource which can be used to answer many kinds of questions. Much of what is known about the majority of the world's species of plants is based on the specimens in herbaria. A large herbarium contains the results of the accumulated effort and information from thousands of people over hundreds of years. Herbarium specimens represent a sampling of the individual variation exhibited by plants and provide the primary data source for developing concepts of species, naming them, and communicating about them. Examining a large number of individuals allows for a more accurate assessment of the nature of a species and its ecological and geographical distribution. Specimens are used in taxonomic and ecological research, such as morphological studies, and for comparative identification and verification of unknown specimens derived from multidisciplinary scientific projects, governmental or private entities, and interested individuals. Herbaria are also major educational resource tools for training new scientists and professionals who deal with plant materials, and for providing basic information on plants (for example, morphology or phytogeography) for many different education levels.

Herbarium collections may include specimens of both wild and cultivated origin, and can give insights into the evolution of domesticated species or point to wild species of potential interest for new genetic traits. They can also provide information on the historical distribution of rare, threatened, or endangered species. *See* PLANT GEOGRAPHY; PLANT TAXONOMY.

Specimen preparation. A high-quality herbarium specimen should be carefully selected and be representative of the individual from which it is made. For many herbaceous plants one or more entire individuals can be included in a specimen, but for larger individuals the parts to be preserved must be chosen by the collector and should exhibit the full range of morphological features of the plant. A specimen is prepared to fit within the dimensions of a standard mounting sheet (11.5 × 16.5 in. or 29 × 42 cm) and must be arranged before drying so that all pertinent parts will be visible. Two or more mounting sheets may be needed to accommodate large specimens.

Each specimen is placed in a single folded sheet of newspaper and pressed between blotters and corrugated cardboard, forming a sandwich. Alternating units of specimen, blotter, and corrugate are then stacked to form a press with a wooden frame at each

end and are bound with two web straps or ropes. The press is then placed over a source of low heat with the corrugations running vertically. The warmed air rises through the corrugations and draws moisture from the plants, while the pressure of the straps ensures that the plants dry flat.

The dried specimen is mounted on a sheet of stiff paper by using glue or gummed tape or by sewing with heavy thread. A standard mounting sheet should be of archival quality and made from 100% rag fiber. The mounting sheet provides support for the specimen and makes it relatively easy to manipulate without damage. Each mounting sheet also has a label glued in the lower right corner with the scientific name of the plant and information on where and when it was collected and by whom. A good label provides detailed information on features of the plant that are not apparent from a direct examination of the dried specimen, such as flower color, height or growth habit, and ecological factors such as habitat, soil type, and associated species.

Structures that are too large to be pressed, such as cones, large fruits, or palm inflorescences, are dried and stored in boxes or bags and are cross-referenced to the conventional, flat specimen to which they belong.

Organization. A herbarium may be arranged in any convenient way as long as the specimens and their information can be easily retrieved. Herbaria are usually arranged by family, genus, species, and infraspecific name, often combined with some geographical scheme. In many collections, specimens are arranged alphabetically by genus and species within a family, or the arrangement may follow an accepted phylogenetic scheme. Specimens are filed in manila folders, often called genus covers, with one or more individual sheets per folder, depending on the thickness of the specimens. In some collections, added protection is provided for each specimen by placing it in a fold of thin, high-quality paper before filing in the genus cover.

Storage units. In the earliest herbaria, specimens were glued to pages, that were bound into volumes. As the number of specimens in collections increased, this method was abandoned, primarily because it was impossible to add to or change the arrangement of the specimens in the bound volumes. Most herbaria today store specimens in steel or wooden cases fitted with two columns of compartments measuring 17 in. (43 cm) deep, 13 in. (33 cm) wide, and 6 in. (15 cm) high (see **illus.**). Because standard cases require permanent aisles wide enough to open the doors, many collections have converted their specimen storage space to mobile modules (compactors), where the units, driven by manual assists or electric motors, move along rails and require no permanent aisles. This is the most compact and efficient means for storing herbarium specimens in limited space.

Protection. Once a plant specimen is prepared and dried, it will essentially last indefinitely if it is properly protected from environmental factors, such as insect pests and high relative humidity. There are specimens from the sixteenth and seventeenth centuries that are as well preserved and useful today as when they were first gathered. As with other items containing cellulose, such as paper, frequent changes in temperature and relative humidity can have an adverse effect and hasten deterioration. The most significant problem, however, for any herbarium is damage from the larvae of various stored-product pests. The most widespread and destructive is the tobacco beetle (*Lasioderma serricorne*). There are no known treatment methods for specimens which will permanently protect them from insect attack. Therefore, a collection must be regularly monitored and any infestations eliminated to prevent damage. An integrated pest management program that combines monitoring, collection isolation and freezing, and judicious use of repellents or pesticides is effective, taking into account the constraints imposed by the physical facility.

Standard herbarium case. (*Harvard University Herbaria; photo by Stephen Jennings*)

Distribution. The *Index Herbariorum* provides information on the majority of the world's herbaria (2639 herbaria in 147 countries holding nearly 273 million specimens). Within the United States there are 628 listed herbaria, with more than 60 million specimens. The oldest herbaria date from the middle of the sixteenth century, but the majority were formed during the twentieth century. Globally, the average herbarium contains about 100,000 specimens, but nearly 65% have less than 25,000 specimens.

James C. Solomon

Bibliography. A. Cronquist, *An Integrated System of Classification of Flowering Plants*, 1992; L. Forman and D. Bridson (eds.), *The Herbarium Handbook*, 3d ed., 1998; P. K. Holmgren, N. H. Holmgren, and L. C. Barnett, *Index Herbariorum*, 8th ed., 1990.

Herbicide

Any chemical used to destroy or inhibit plant growth, especially of weeds or other undesirable vegetation. The concepts of modern herbicide technology began to develop about 1900 and accelerated rapidly with the discovery of dichlorophenoxyacetic acid (2,4-D) as a growth-regulator-type herbicide in 1944–1945. A few other notable events should be mentioned. During 1896–1908, metal salts and mineral acids were introduced as selective sprays for controlling broad-leafed weeds in cereals; during 1915–1925 acid arsenical spray, sodium chlorate, and other chemicals were recognized as herbicides; and in 1933–1934, sodium dinitrocresylate became the first organic selective herbicide to be used in culture of cereals, flax, and peas. Since the introduction of 2,4-D, a wide variety of organic herbicides have been developed and have received wide usage in agriculture, forestry, and other industries. Today, the development of highly specific herbicides that are intended to control specific weed types continues. Modern usage often combines two or more herbicides to provide the desired weed control. Worldwide usage of herbicides continues to increase, making their manufacture and sale a major industry.

The control of weeds by means of herbicides has provided many benefits. Freeing agricultural crops from weed competition results in higher food production, reduced harvesting costs, improved food quality, and lowered processing costs, contributing to an abundant supply of low-cost, high-quality food. Not only are billions of dollars saved through increased production and improved quality, but costs of labor and machinery energy for weed control are reduced, livestock is saved from the effects of poisonous weeds, irrigation costs are reduced, and insect and disease control costs are decreased through the removal of host weeds for the undesirable organisms.

Additional benefits due to appropriate herbicide use result as millions of people are relieved of the suffering caused by allergies to pollens and exposure to poisonous plants. Recreational areas, roadsides, forests, and parks have been freed of noxious weeds, and home lawns have been beautified. Herbicides have reduced storage and labor costs and fire hazards for industrial storage yards and warehouse areas. Modern herbicides even benefit the construction industry, where chemicals applied under asphalt prolong pavement life by preventing weed penetration of the surface.

Classification. There are well over a hundred chemicals in common usage as herbicides. Many of these are available in several formulations or under several trade names. The many materials are conveniently classified according to the properties of the active ingredient as either selective or nonselective. Further subclassification is by method of application, such as preemergence (soil-applied before plant emergence) or postemergence (applied to plant foliage). Additional terminology sometimes applied to describe the mobility of postemergence herbicides in the treated plant is contact (nonmobile) or translocated (mobile—that is, killing plants by systemic action). Thus, glyphosate is a nonselective, postemergence, translocated herbicide.

Selective herbicides are those that kill some members of a plant population with little or no injury to others. An example is alachlor, which can be used to kill grassy and some broad-leafed weeds in corn, soybeans, and other crops.

Nonselective herbicides are those that kill all vegetation to which they are applied. Examples are bromacil, paraquat, or glyphosate, which can be used to keep roadsides, ditch banks, and rights-of-way open and weed-free. An important use for such chemicals is the destruction of vegetation before seeding in the practice of reduced tillage or no tillage. Some are also used to kill annual grasses in preparation for seeding perennial grasses in pastures. Additional uses are in fire prevention, elimination of highway hazards, destruction of plants that are hosts for insects and plant diseases, and killing of poisonous or allergen-bearing plants.

Preemergence or postemergence application methods derive naturally from the properties of the herbicidal chemical. Some, such as trifluralin, are effective only when applied to the soil and absorbed into the germinating seedling, and therefore are used as preemergence herbicides. Others, such as diquat, exert their herbicidal effect only on contact with plant foliage and are strongly inactivated when placed on contact with soil. These can be applied only as a postemergence herbicide. However, the distinction between pre- and postemergence is not always clear-cut. For example, atrazine can exert its herbicidal action either following root absorption from a pre- emergence application or after leaf absorption from a postemergence treatment, and thus it can be used with either application method. This may be an advantage in high-rainfall areas where a postemergence treatment can be washed off the leaf onto the soil and nevertheless can provide effective weed control.

Herbicidal action. Many factors influence herbicide performance. A few are discussed below.

Soil type and organic matter content. Soils vary widely in composition and in chemical and physical characteristics. The capacity of a soil to fix or adsorb a preemergence herbicide determines how much will be available to seedling plants. For example, a sandy soil normally is not strongly adsorptive. Lower quantities of most herbicides are needed on a highly adsorptive clay soil. A similar response occurs with the organic portion of soil; higher organic matter

content usually indicates that more herbicide is bound and a higher treatment rate necessary. For example, cyanazine is used for weed control on corn at the rate of 1.5 lb/acre (1.7 kg/ha) on sandy loam soil of less than 1% organic matter, but 5 lb/acre (5.6 kg/ha) is required on a clay soil of 4% organic matter.

Leaching. This refers to the downward movement of herbicides into the soil. Some preemergence herbicides must be leached into the soil into the immediate vicinity of weed seeds to exert toxic action. Excessive rainfall may leach these chemicals too deeply into the soil, thereby allowing weeds to germinate and grow close to the surface.

Volatilization. Several herbicides in use today will volatilize from the soil surface. To be effective, these herbicides must be mixed with the soil to a depth of 2 to 4 in. (5 to 10 cm). This process is termed soil incorporation. Once the herbicide is in contact with soil particles, volatilization loss is minimized. This procedure is commonly employed with thiocarbamate herbicides such as EPTC and dinitroaniline herbicides such as trifluralin.

Leaf properties. Leaf surfaces are highly variable. Some are much more waxy than others; many are corrugated or ridged; and some are covered with small hairs. These variations cause differences in the retention of postemergence spray droplets and thus influence herbicidal effect. Most grass leaves stand in a relatively vertical position, whereas the broad-leafed plants usually have their leaves arranged in a more horizontal position. This causes broad-leaf plants to intercept a larger quantity of a herbicide spray than grass plants.

Location of growing points. Growing points and buds of most cereal plants are located in a crown, at or below the soil surface. Furthermore, they are wrapped within the mature bases of the older leaves. Hence they may be protected from herbicides applied as sprays. Buds of many broad-leafed weeds are located at the tips of the shoots and in axils of leaves, and are therefore more exposed to herbicide sprays.

Growth habits. Some perennial crops, such as alfalfa, vines, and trees, have a dormant period in winter. At that time a general-contact weed killer may be safely used to get rid of weeds that later would compete with the crop for water and plant nutrients.

Application methods. By arranging spray nozzles to spray low-growing weeds but not the leaves of a taller crop plant, it is possible to provide weed control with a herbicide normally phytotoxic to the crop. This directed spray technique is used to kill young grass in cotton with MSMA or broad-leafed weeds in soybeans with chloroxuron in an oil emulsion. Another example of selective application technology is the use of wiper applicators. This apparatus commonly employs absorbent ropes which draw herbicide solution by capillary action from a reservoir. The equipment is arranged so that the ropes can pass above crop plants and transfer the herbicide solution to weeds taller than the crop by a wiping action. Use of a nonselective translocated herbicide in this recirculating sprayer system will then selectively remove the weeds from the crop. Glyphosate is being used commercially in such systems to remove johnson grass from cotton and soybeans.

Protoplasmic selectivity. Just as some people are immune to the effects of certain diseases while others succumb, so some weed species resist the toxic effects of herbicides whereas others are injured or killed. This results from inherent properties of the protoplasm of the respective species. One example is the use of 2,4-DB or MCPB (the butyric acid analogs of 2,4-D and MCPA) on weeds having β-oxidizing enzymes that are growing in crops (certain legumes) which lack such enzymes. The weeds are killed because the butyric acid compounds are broken down to 2,4-D or MCPA. Another example is the control of a wide variety of weeds in corn by atrazine. Corn contains a compound that conjugates with the atrazine atom, rendering it nontoxic; most weeds lack this compound. A third important example of protoplasmic selectivity is shown by trifluralin, planavin, and a number of other herbicides applied through the soil which inhibit secondary root growth. Used in large seeded crops having vigorous taproots, they kill shallow-rooted weed seedlings; the roots of the crops extend below the shallow layer of topsoil containing the herbicides and the seedlings survive and grow to produce a crop.

Properties. Several factors of the commercial herbicide influence selectivity to crops.

Molecular configuration. Subtle changes in the chemical structure can cause dramatic shifts in herbicide performance. For example, trifluralin and benefin are very similar dinitroaniline herbicides differing only in the location of one methylene group. However, this small difference allows benefin to be used for grass control in lettuce, whereas trifluralin will severely injure lettuce at the rates that are required for weed control.

Herbicide concentration. The action of herbicides on plants is rate-responsive. That is, small quantities of a herbicide applied to a plant may cause no toxicity, or even a slight growth stimulation, whereas larger amounts may result in the death of the plants. It has long been known that 2,4-D applied at low rates causes an increase in respiration rate and cell division, resulting in an apparent growth stimulation. At high application rates, 2,4-D causes more severe changes and the eventual death of the plant.

Formulation. The active herbicidal chemical itself is seldom applied directly to the soil or plants. Because of the nature of the chemicals, it is usually necessary that the commercial product be formulated to facilitate handling and dilution to the appropriate concentration. Two common formulations are emulsifiable concentrates and wettable powders. Emulsifiable concentrates are solutions of the chemical in an organic solvent with emulsifiers added which

permit mixing and spraying with water. Wettable powders are a mixture of a finely divided powder, active chemical, and emulsifiers which allow the powder to suspend in water and to be sprayed. An additional ingredient called an adjuvant is sometimes added to the formulation or to the spray tank when the spray solution is mixed. These adjuvants are normally surface-active agents (surfactants) which improve the uniformity of spray coverage on plants and the plant penetration of the herbicide. An example is the addition of surfactant to paraquat spray solutions to improve its nonselective postemergence action on weeds. Another material sometimes added to spray solutions is a nonphytotoxic oil. These oils may be used to improve postemergence action of herbicides such as diuron which normally have limited foliar absorption. *See* PESTICIDE; SURFACTANT.

Rodney O. Radke

Bibliography. F. M. Ashton and T. J. Monaco, *Weed Science: Principles and Practices*, 3d ed., 1991; J. R. Crister, Jr., *Herbicides*, 1978; *Herbicide Handbook of the Weed Science Society of America*, 7th ed., 1994; P. C. Kearney and D. Kaufman (eds.), *Herbicides: Chemistry, Degradation, and Mode of Action*, 1975–1976; R. D. King, *Farmers Weed Control Handbook*, 1985; H. A. Roberts (ed.), *Weed Control Handbook: Principles*, 8th ed., 1990; W. T. Thomas, *Agricultural Chemicals: Herbicides*, 1993, Book II, 1993.

Herbig-Haro objects

Compact nebulae with peculiar spectra that were first studied in the 1950s by George Herbig and, independently, by Guillermo Haro. Herbig-Haro (HH) objects are typically located within or adjacent to star-forming regions and appear as small patches of nebulosity with highly variable luminosity. They exhibit spectroscopic emission lines that are characteristic of highly ionized gas that has been shocked upon interaction with the surrounding interstellar medium. Doppler shifts of the spectral lines reveal that the gas is moving at supersonic velocities of 100 km/s (60 mi/s) to more than 1000 km/s (600 mi/s). Typical masses of an individual clump of gas are 1–20 times that of Earth with temperatures of 8000–12000 K. They are composed almost entirely of hydrogen and helium with less than 1% consisting of heavier elements. Over 400 individual Herbig-Haro objects are known. Many of the closest ones are observed to emanate from a very young stellar source with proper motions that are discernible in *Hubble Space Telescope* images taken over periods of several years. *See* ASTRONOMICAL SPECTROSCOPY; DOPPLER SHIFT; HUBBLE SPACE TELESCOPE.

The discovery of bipolar outflows from young stars in the late 1970s and early 1980s revealed Herbig-Haro objects to be associated with jets of ionized gas that emanate from young "exciting" stars surrounded by circumstellar accretion disks (**Fig. 1**). When identified, the source stars are pre-main-sequence T Tauri stars or young protostars still hidden at optical wavelengths by a collapsing envelope of dust and gas. The ionized jets are typically highly collimated and oppositely directed along the rotational axis of a disk (**Fig. 2**). Millimeter-wave observations of carbon dioxide emission often reveal less collimated molecular outflows moving in the same general direction but at lower velocities. These are usually found to contain Herbig-Haro objects when studied in sufficient detail. Both ionized jets and molecular outflows may extend for several parsecs from their exciting sources, where they are distributed over a much greater range of angles than nearer to the source. *See* PROTOSTAR; T TAURI STAR.

It is now understood that bipolar jets and outflows responsible for Herbig-Haro objects probably play an important role in triggering continued star formation. A large number of irradiated Herbig-Haro objects are found in clustered star forming regions such as the Orion Nebula, where they deliver enough energy to substantially perturb the medium and could have been responsible for initiating gravitational collapse in the densest clumps of gas and dust. In addition, Herbig-Haro objects reveal centuries-old histories of jet emission from star-disk systems. Consequently, they are important laboratories for

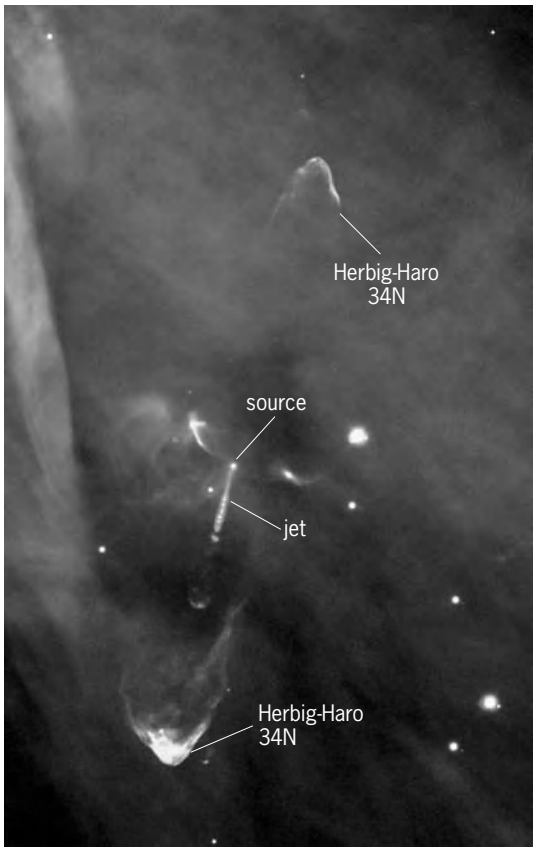

Fig. 1. Jet and bow-shock emission from Herbig-Haro 34. Two opposite jets run into the surrounding interstellar medium, where they terminate in Herbig-Haro objects with a characteristic bow-shock structure. The jets are composed of "bullets" of dense gas ejected at high velocities. (*European Southern Observatory*)

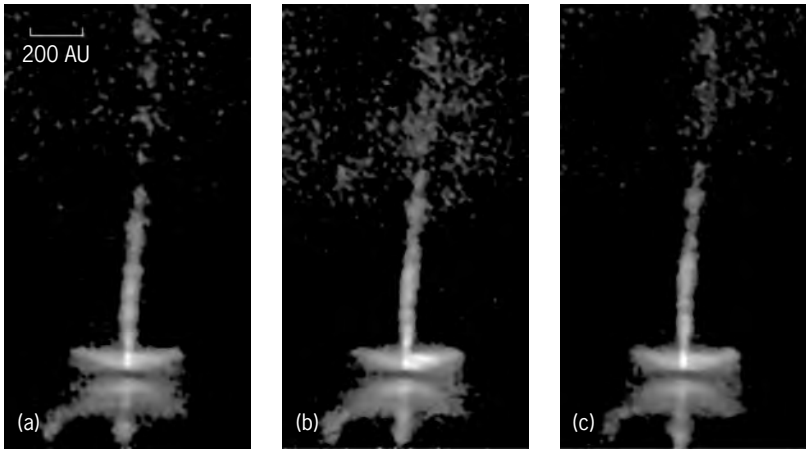

Fig. 2. Disk and jets associated with Herbig-Haro 30, showing changes between images taken in (*a*) 1995, (*b*) 1998, and (*c*) 2000 with the Wide Field and Planetary Camera 2 aboard the *Hubble Space Telescope*. A slightly flared disk of dust and gas appears edge-on at the bottom of the images. The jet and associated Herbig-Haro objects are believed to arise from complex interactions between the magnetic fields associated with accreting disk material and a rapidly rotating young star hidden from view in the surrounding dust and gas. (*NASA; Alan Watson, Universidad Nacional Autonoma de Mexico; Karl Stapelfeldt, Jet Propulsion Laboratory; John Krist, Space Telescope Science Institute; Chris Burrows, European Space Agency/Space Telescope Science Institute*)

star-formation studies, including the conditions prior to collapse and the evolution of outflow systems. *See* ORION NEBULA; STELLAR EVOLUTION. David Koerner

Bibliography. T. P. Ray, Fountains of Youth: Early days in the life of a star, *Sci. Amer.*, 283(4):42–47, August 2000; B. Reipurth and J. Bally, Herbig-Haro flows: Probes of early stellar evolution, *Annu. Rev. Astron. Astrophys.*, 39:403–455, 2001; B. Reipurth and S. Heathcote, Herbig-Haro Objects and the birth of stars, *Sky Telesc.*, 90(4):38–40, October 1995.

Herbivory

The consumption of living plant tissue by animals. Plants and the animals that consume them constitute roughly one-half of the scientifically described species. Herbivorous species occur in most of the major taxonomic groups of animals, including the vertebrates (such as grazing fish, tortoises, geese, and hoofed mammals), echinoderms (sea urchins), mollusks (snails and slugs), nematodes (roundworms), and arthropods (including crabs, lobsters, copepods, amphipods, and isopods; mites; and especially insects, such as beetles, caterpillars of moths and butterflies, larvae of many flies, sawflies, grasshoppers, aphids and their relatives, thrips, and stick insects). Herbivorous insects alone may account for one-quarter of all species. The fraction of all biomass produced by plants that is eaten by herbivores varies widely among plants and ecosystems, ranging from less than 1% to nearly 90%. Thus, in terms of both the number of species involved and the role that herbivory plays in the flow of energy and nutrients in ecosystems, herbivory is a key ecological interaction between species. *See* MARINE ECOLOGY.

Plant adaptations. Either because the herbivore is much smaller than the plant, as when caterpillars feed upon trees, or because part of the plant is inaccessible to the herbivore, herbivory usually does not kill the plant outright, although there are striking exceptions (such as bark beetle outbreaks that decimate conifer trees over thousands of square kilometers). Nevertheless, chronic attack by herbivores can have dramatic cumulative effects on the size, longevity, or reproductive output of individual plants. As a consequence, plants have evolved several means to reduce the level of damage from herbivores and to ameliorate the impact of damage.

Many plants possess physical defenses that interfere mechanically with herbivore feeding on or attachment to the plant. Spines and thorns deter feeding by large herbivores, smaller hairs on the leaf surface can jam insect mouthparts (and in some cases, sharp hooks or sticky resin on these hairs can entrap insects), and the waxy surface layer of plant leaves makes it difficult for insect herbivores to move around the plant without falling off. Mature oak leaves cannot easily be penetrated by the mandibles of newly hatched winter moth caterpillars, which in part explains the small size and low survival of caterpillars fed mature leaves. In addition to these physical defenses, plant tissues may contain chemical compounds that render them less digestible or even toxic to herbivores. For example, tannins and quinones bind to proteins in the food, reducing the nutritional quality of the consumed plant tissue. Many plant compounds (such as heart poisons in milkweed leaves and delphinine in larkspur foliage) can cause death if consumed by unadapted herbivores. While natural selection imposed by herbivores was the likely force driving the elaboration of these plant chemicals, humans have subsequently found many uses for the chemicals as active components of spices (capsacin, which makes cayenne pepper hot), stimulants (caffeine), relaxants (nicotine), hallucinogens (mescaline),

poisons (strychnine), and drugs (ephedrine).

An exciting recent finding is that some plants possess induced resistance, elevated levels of physical or chemical defenses that are brought on by herbivore damage and confer enhanced resistance to further damage. Acacia trees grow branches with longer and denser thorns after they are browsed by giraffes. Damaged leaves of wild parsnip (*Pastinaca sativa*) contain higher concentrations of antiherbivore compounds, and caterpillars of the cabbage looper (*Trichoplusia ni*) grow far less rapidly when fed previously damaged leaves. Two reasons why plants may produce certain defensive chemicals only after damage (rather than always producing them) are that those chemicals may be partially toxic to the plant and that their production may involve a physiological cost. For example, wild parsnip leaves after damage have higher rates of respiration that can be accounted for entirely by the production of energy required to synthesize damage-induced antiherbivore compounds. Beetle-damaged soybean plants are initially less palatable to beetles than are undamaged plants, but after about 20 days the damaged plants actually become more palatable, a possible delayed cost of induced resistance. If damage is unpredictable, plants may save energy or scarce nutrients for other uses by defending heavily only those tissues that are highly likely to be attacked by herbivores, and producing defensive chemicals in other tissues only after they are damaged. Wild parsnip fruits, which are consistently attacked by herbivores, always contain high levels of defensive chemicals, whereas root tissues, which receive herbivore damage only rarely, produce high concentrations of chemical defenses only after being damaged.

Induced resistance may eventually be exploited for agricultural purposes. For example, populations of spider mites on grape vines that had received prior damage by another species of mite were substantially lower than populations on previously undamaged plants. Thus, rather than using pesticides, farmers may be able to "vaccinate" their crops with relatively benign herbivores to protect them from later-feeding herbivores that cause more damage. Induced resistance may also be a factor contributing to periodic outbreaks of herbivore populations, through the following sequence: A high herbivore population induces resistance and consequently declines, leading to reduced damage, a decline in the induced resistance as the plants regrow, and finally the return of rapid herbivore population growth to restart the cycle.

When defenses fail to prevent herbivore damage, plants can utilize several mechanisms to compensate for the effects of damage. Fireweed plants (*Epilobium latifolium*) whose growing shoot tips are destroyed by moth larvae grow longer side branches to replace the leaf tissue that they were prevented from producing on the damaged main stem. In one location, scarlet gilia plants whose growing tips were eaten by grazing mammals actually grew larger and produced more seeds than undamaged plants, but such beneficial effects of herbivory are unlikely to occur commonly, even for scarlet gilia across its entire range. Tissues that plants regrow after herbivore damage commonly have higher photosynthetic rates than undamaged tissues, which may hasten the process of compensating for the lost tissue. In at least two species of nonagricultural plants, families of related individuals have been shown to vary in their ability to maintain high seed production after receiving herbivore damage. Tolerance of herbivory in plants thus appears to have a genetic basis, and can evolve in response to herbivore pressure.

Herbivore adaptations. Herbivores can either avoid or counteract plant defenses. Many herbivores avoid consuming the plant tissues that contain the highest concentrations of toxic or antinutritive chemicals. Insects that feed on plants whose leaves contain secretory canals through which toxic substances are transported to sites of herbivore damage often sever the canals before feeding. Herbivores have also evolved an elaborate array of enzymes to detoxify otherwise lethal plant chemicals. Because few herbivores have the ability to detoxify the chemical compounds produced by all the plant species they encounter, many herbivores have restricted diets; the larvae of more than half of all species of butterflies and moths include only a single genus of plants in their diets. Some insect species that have evolved the means to tolerate toxic plant chemicals have also evolved ways to use them in their own defense. Larvae of willow beetles store plant compounds in glands along their back. When the larvae are disturbed, the glands exude droplets of the foul-smelling compounds, which deter many potential predators.

If a plant evolved the ability to produce a novel chemical compound that its herbivores could not detoxify, the plant and its descendants would be freed for a time from the negative effects of herbivory. A herbivore that then evolved the means to detoxify the new compound would enjoy an abundance of food and would increase until the level of herbivory on the plant was once again high, favoring plants that acquire yet another novel antiherbivore compound. These repeated rounds of evolution of plant defenses and herbivore countermeasures have been termed coevolution. Coevolution over long periods of time helps to explain similar patterns of evolutionary relatedness between groups of plant species and the herbivorous insect species that feed on them.

Ecological context. Plants and their herbivores seldom occur in isolation, and other species can influence the interaction between plants and herbivores. Plant defenses that slow the growth of insect larvae increase the insects' vulnerability to their predators. Many parasites use characteristic plant odors to help them find herbivorous host insects with restricted plant diets. Certain fungi that grow within the leaves of grasses are rather harmless to the plants but produce chemicals that deter feeding by insects and mammals, causing infected plants to outperform uninfected plants in

the presence of herbivores. Mammalian herbivores often rely upon gut microorganisms to digest cellulose in the plant material they consume. Thus, herbivory occurs against a backdrop of multiple interactions involving the plants, the herbivores, and other species in the ecological community. *See* GRASS CROPS; POPULATION ECOLOGY.

William F. Morris

Bibliography. R. S. Fritz and E. L. Simms, *Plant Resistance to Herbivores and Pathogens: Ecology, Evolution, and Genetics*, 1992; R. Karban and I. T. Baldwin, *Induced Responses to Herbivory*, 1997; G. A. Rosenthal and M. R. Berenbaum, *Herbivores: Their Interactions with Secondary Plant Metabolites*, 1992; L. M. Schoonhoven, T. Jermy, and J. J. A. van Loon, *Insect-Plant Biology*, 1998.

Hermaphroditism

A condition in which components of both testes and ovaries are present in the same individual. Although true hermaphroditism is common among lower forms of animals such as annelids and mollusks, it is rare in humans. A more common condition in humans is pseudohermaphroditism, which simulates hermaphroditism. In female pseudohermaphroditism, or gynandry, the external sexual characteristics are in part or wholly of the male aspect, but internal female genitalia are present. In male pseudohermaphroditism, or androgyny, the individual has external sexual characteristics of female aspect, but has testes (usually undescended).

The four general forms of hermaphroditism are bilateral, the presence of an ovary and testis on each side; lateral, the presence of an ovary on one side and a testis on the other; ovatesticular, the presence of an ovatestis on one or both sides; and unilateral, the combination of an ovatestis on one side with an ovary or testis on the other.

During the fourth to sixth week of embryonic life the genital ridge differentiates into cortical and medullary components. If the embryo is a genotypic male (XY), its destiny is to be a phenotypic male; the cortical component atrophies while the medullary component develops into the seminiferous tubules. If the embryo is a genotypic female (XX), destined to be a phenotypic female, the medullary component atrophies and the cortical component persists to become the ovary. Any failure of the tissue components to develop or function, or any aberration of the chromosomes during embryogenesis, results in some degree of abnormal sexual characteristics; sexual alterations may also occur at puberty. Thus, variable phenotypic patterns of masculinization and feminization occur in accordance with genotypic determinants, and with the elaboration of hormones. *See* OVARY; REPRODUCTIVE SYSTEM DISORDERS; TESTIS.

Sybil P. Parker

Hernia

The protrusion of an organ, part of an organ, or other structure through the wall of the body cavity normally containing it. Various organs may be involved, including the bladder, brain, esophagus, intestine, ovary, and rectum. The most common location for a hernial bulge to appear is the abdominal wall, particularly the groin.

Among the most infrequent but life-threatening hernias is a cerebral hernia in which part of the brain protrudes through an opening in the skull.

A diaphragmatic hernia, which occurs when a defect is present in the muscular diaphragm separating chest from abdomen, may be present at birth or result from an injury later in life. Abdominal organs, such as the liver, spleen, stomach, and intestine, can pass through the diaphragmatic defect and lodge in the chest cavity, so that the lungs become compressed and breathing is impaired. Hiatal or esophageal hernia results when a portion of the stomach slides into the chest cavity through the normal diaphragmatic opening for the esophagus. Hiatal hernias occur principally in the elderly and primarily produce symptoms of indigestion or heartburn. Both diaphragmatic and hiatal hernias can be cured surgically.

There are various abdominal-wall hernias. Among the most common is an incisional or ventral hernia which occurs when an abdominal incision does not heal correctly or separates because of abnormal strain. The defect allows a protrusion of the intestine, necessitating surgical repair. Umbilical or navel hernias are usually congenital, representing a defect left by the passage of umbilical cord structures during fetal life. Unlike all other hernias, childhood umbilical hernias tend to close spontaneously. In adults, umbilical hernias may be caused by a number of predisposing factors, including pregnancy and abdominal distension due to an abnormal collection of fluid or obesity. Epigastric hernias are located strictly in the midline of the abdomen and above the navel. Both adult umbilical and epigastric hernias require little more than a simple surgical procedure for permanent cure.

A Spigelian hernia is located on the lateral abdominal wall, while a lumbar or dorsal hernia occurs through the posterior abdominal wall. Both types are quite rare and difficult to diagnose. They are associated with vague complaints of discomfort and an associated ill-defined mass. Operative repair is accomplished by the simple reapproximation of separated muscle and connective tissue.

Groin hernias consist of two major types, inguinal and femoral. Inguinal hernias account for 75% of all hernias of the body, and are divided into two anatomic variants, indirect and direct. In large inguinal hernias, the groin protrusion may stretch into the scrotum. Indirect inguinal hernias are caused by a weakness in the abdominal wall that corresponds to an area where the testis descended into the scrotum during embryological development. With direct inguinal hernias, the defect results mainly from strain

on the abdominal muscles which have been weakened by age—hence the term "older man's" hernia. Inguinal hernias are 10 times as common in men as in women. Femoral hernias are more common in women, but are infrequent. The weakness in a femoral hernia originates in the area where the major veins, arteries, and nerves pass from the abdomen into the lower extremities. A femoral hernia bulge is always located in the upper inner part of the thigh, just below the groin crease.

Clinical manifestations. A groin hernia develops when the outer layers of the lower abdominal wall weaken or actually separate. Through this defect, the inside lining or peritoneal membrane of the abdominal cavity gradually protrudes and forms a sac. The hernia sac contains displaced intestine or other structures. When combined with the outer layers of stretched-out muscle, connective tissue, and skin, a hernial bulge forms.

Development of a groin hernia usually results from chronic strain on the lower abdominal muscles which in turn causes general loss of tissue tone. Age and heredity also play important roles. Other risk factors include a chronic cough, constipation, generalized obesity, and pregnancy. Activities such as lifting a heavy object or sudden twists or pulls have long been regarded as instant causes of groin hernias, but do not reflect the complicated pathophysiology of groin herniation.

Most inguinal hernias cause little acute pain. They are usually associated with a variety of milder symptoms, including weakness or pressure in the groin, an aching discomfort in the region, a burning feeling in the area of the bulge, or the presence of intestinal gurgling.

A groin hernia is termed reducible if the protruding sac and its intestinal contents can be pushed back into place inside the abdominal cavity. However, upon standing or especially with exercise the bulge quickly reappears. If the sac cannot be manually reduced, the condition is termed an incarcerated hernia. In rare instances, the blood supply to the incarcerated intestine becomes pinched off. This complication is termed a strangulated hernia and is a surgical emergency. Rapidly worsening pain, occasional vomiting, and a firm tender bulge are signals that the herniated tissue has become strangulated and is in the process of turning gangrenous and dying. Within a few hours, this condition becomes life-threatening and requires immediate surgical attention.

Diagnosis and treatment. Most groin hernias are self-diagnosed. However, the presence of a visible bulge in the groin or scrotum is enough to warrant the attention of a surgeon. It is important that other disease entities, including accumulation of fluid in the testis (hydrocele), tumor of the lymph gland (lymphoma), and swelling of spermatic cord (varicocele), be ruled out.

With rare exceptions, all hernias should be corrected surgically to prevent the possibilities of incarceration, intestinal obstruction, and strangulation. Groin herniorrhaphy is typically an elective operation, performed on an ambulatory basis and over 98% successful. Surgical repair involves strengthening the area of weakness through one of many surgical techniques.

Groin hernia repair methods utilize tension-free techniques in which a small incision is made at the site of the hernia and a piece of plastic mesh is inserted to cover the area of the abdominal-wall defect. There is no tension because the surrounding tissues are not sewn together. For example, a conical mesh, termed a hernia plug, can be applied so that it adheres to the confines of the hernia defect, much like placing a cork in a bottle. The mesh is safe and well tolerated by the body's natural tissues. Such modern herniorrhaphies take approximately 20 min to perform, recovery is swift, and the likelihood of the hernia recurring is less than 2%. Also, since these methods of groin hernia repair cause little postoperative discomfort, individuals are able to resume most normal activities in 2–3 days. Ira M. Rutkow

Bibliography. I. M. Rutkow, *Hernia Surgery*, 1993; G. E. Wantz, *Atlas of Hernia Surgery*, 1991.

Herpes

Any virus of the herpesvirus group, which comprises a family of 70 species, 5 of which are pathogenic to humans; the term also refers to any infection caused by these viruses. Since these pathogens are ubiquitous in nature, most individuals of all populations are exposed to and thus immunized to these viruses. The five pathogenic groups include herpes simplex I and II, varicella-zoster, cytomegalovirus, and the Epstein-Barr virus.

In nonimmunized hosts, the vast majority of all herpes infections are subclinical; that is, they do not cause specific symptoms or signs attributable to herpes infections, but present symptoms of nonspecific viral illnesses which resolve spontaneously. However, the infections that cause clinical disease in fact may cause serious morbidity and mortality in afflicted individuals. A whole range of problems accompanies infections, including the mental anguish of the victims of the venereal disease caused by herpes simplex II (genitalis), and the great concern of public health officials over the epidemic proportions of this disorder. Reactivation of herpes infection, characteristic of the immunocompromised host, is an important cause of mortality in the treatment of patients with advanced cancer, and is a dreaded potential complication of an otherwise possibly curable systemic disease. The outlook for the treatment of patients afflicted with serious herpesvirus infections becomes brighter as potential treatments for several herpesvirus disorders are identified.

Characteristics of the virus. All known viruses are composed of a core of nucleic acids and a protein barrier, the capsid, which protects the gene structure. Herpesviruses have a deoxyribonucleic acid (DNA) core and are 150 to 200 nanometers in size with icosahedral symmetry, and are coated by an envelope derived from the infected host cells. The surface of the virions in general contains protein-carbohydrate

structures which allow cellular attachment and thus cellular penetration. All viruses require living cells for their replication, since they carry no metabolic machinery themselves, but rely upon the host-cell metabolic processes. The virus may replicate and destroy the cell, or replicate and allow cell survival, or incorporate its viral gene structure into the host gene structure.

This incorporation phenomenon is designated as latency, and, for example, herpes simplex virus exhibits the phenomenon of latency within nerve cells in the area of previous infection. The Epstein-Barr virus characteristically causes latent infection in lymphocytes (white blood cells in the circulating blood), and the cytomegalic virus also causes latent infection within lymphocytes and possibly within nerve cells. Once the viral genome is incorporated into the host cell, antiviral drugs are of no use, since therapeutic agents cannot selectively destroy or inhibit the viral genome. Viral biologic avenues of infection are still not completely understood. Factors which are possibly involved in the reactivation of latent virus generally revolve around some depression of the host immune response system, and this may be mediated in part through damage by irradiation, chemical agents, mechanical injuries, hormones, emotions, or superimposed infections. Viral genome incorporation into host cells is of great interest as several herpesvirus types are implicated in the development of cancer (herpes simplex II with uterine cervical cancer, Epstein-Barr virus with Burkitt's lymphoma and nasopharyngeal carcinoma). *See* VIRUS INFECTION, LATENT, PERSISTENT, SLOW.

The fact that viruses require host-cell machinery for replication implies that metabolites can be used to compete with or inhibit the assembly of viral DNA. This is the foundation of therapeutic intervention for all herpesviruses, which involves a series of chemicals with structures similar to the base pairs which compose the DNA structure. The base analogs compete with or inhibit viral enzymes necessary for the assembly of viral DNA. The agents are useful in selected clinical situations for given age groups and virus types, and have been only for investigational use in the prophylaxis of certain herpes-type infections. *See* VIRUS.

Herpes simplex I and II. These infections are spread by intimate contact of mucocutaneous surfaces during the period of virus shedding from active lesions. Clinical infections involving herpes type II usually affect the genitalia, but may affect the oral mucosa, causing exquisitely painful ulcerations which crust and heal. Upon healing, the virus resides in latent form within local nerve cells. Factors related to viral reactivation are poorly understood, but may relate in part to the host immune system. The type II virus has been linked to the development of uterine cervical carcinoma, since women with primary genital herpes display roughly a 12-fold greater risk than women without primary genital herpes. Also, women with uterine cervical cancer have greater blood antibody levels against herpesvirus type II than controls. Uterine cervical cancer often is preceded by an abnormal growth of cervical cells known as dysplasia, and a similar correlation between incidence of dysplasia and the presence of herpes antigens in these precancerous cells is established. Other factors which predispose to the development of cervical cancer include the beginning of sexual intercourse at an early age and multiple sexual partners. Thus, in part, cervical cancer is a venereal disease, and this lends support to the notion that a transmissible agent is involved. However, the precise role of herpes type II remains a question, as the virus may act as a cocarcinogen along with some other agent in the cervix, or in fact may be an incidental phenomenon which tends to occur with greater frequency in incipient cancers of the cervix. *See* SEXUALLY TRANSMITTED DISEASES.

Active herpes lesions in prepartum women are an indication for cesarean section, since neonatal herpes may be acquired through the birth canal and may result in serious neonatal morbidity and mortality.

Herpes simplex virus I (cold sores, fever blisters) afflicts 20–40% of the population in the United States and usually affects the oropharynx, causing pharyngitis, tonsillitis, gingivostomatitis, or keratitis (eye inflammation) as primary infections. Inflammation of the mouth, eye, or brain may occur as a secondary infection.

Acyclovir is recommended for the treatment of primary genital herpes, but no therapy is recommended for recurrent herpes genitalis. Also, no therapy is recommended for primary labial herpes, but Acyclovir may be used in high-risk groups, and topical triflurothymidine may be used for keratitis.

Herpes varicella-zoster. Primary infection (airborne) due to varicella-zoster usually affects preschool children, causing chickenpox, a systemic cutaneous vesicular skin eruption. Rare complications include encephalitis, hepatitis, and pneumonia, with the complications usually affecting the immunocompromised host. Secondary infection usually afflicts the elderly when latent viral reactivation occurs, presumably due to an immune imbalance in the host, and involves the spread of virus along the skin in the anatomic distribution of nerve (this disorder is known as shingles). Infection in the first trimester of pregnancy may cause fetal anomalies, and infection of the mother near time of delivery may cause neonatal herpes, which can be treated with hyperimmune serum containing antibodies to herpes varicella-zoster. Vidarabine is recommended as therapy in the compromised host or high-risk patient. *See* CHICKENPOX AND SHINGLES.

Cytomegalovirus. Like all herpes, cytomegalovirus is ubiquitous, with the majority of infections remaining subclinical. Adult syndromes which are well recognized include a mononucleosislike syndrome and hepatitis, both of which are self-limited diseases in the normal host. However, reactivation of latent infection is a major source of morbidity and especially mortality in the compromised host, for example, the patient being treated with

chemoradiotherapy for advanced malignant disease. Disseminated cytomegalic disease in these patients is often fatal. *See* HEPATITIS.

Since rubella infections in pregnancy have been controlled, cytomegalovirus has become one of the most important serious pregnancy infections as serious fetal sequela are common, for example, prematurity, malformations, mental retardation, or death. There is no recommendation for drug therapy, although investigational drugs are in use. *See* CYTOMEGALOVIRUS INFECTION.

Epstein-Barr virus. The characteristic clinical syndrome caused by Epstein-Barr virus infection includes generalized lymphadenopathy, hepatosplenomegaly, pharyngitis, tonsillitis, and general fatigue and fever. This disorder affects individuals of all ages, but predominantly adolescents. The majority of children are subclinically infected. This mononucleosis syndrome is usually a self-limited disorder, and investigational drugs in use for prophylaxis of high-risk individuals include interferons and acyclovir. Epstein-Barr virus is suspected to be of etiologic importance in Burkitt's African lymphoma, in which the genomes of Epstein-Barr virus are routinely displayed in tumor cells of lymphocytes. However, the Epstein-Barr genome is not found in the non-African cases of Burkitt's lymphoma. Virological and immunological data support an etiological link of Epstein-Barr virus and a cancer of the nasopharynx. *See* ANIMAL VIRUS; EPSTEIN-BARR VIRUS; VIRUS CHEMOPROPHYLAXIS.

David J. Dabbs

Bibliography. D. L. Ablashi (ed.), *Epstein-Barr Virus and Human Diseases*, 1987; D. A. Baker, *Acyclovir Therapy for Herpes Virus Infection*, 1990; E. Lycke and E. Norrby (eds.), *Textbook of Medical Virology*, 1983; S. Robbins and R. Cotran, *Pathologic Basis of Disease*, 6th ed., 1999; W. C. Russell and J. W. Almond (eds.), *Molecular Basis of Virus Disease*, 1987; D. O. White and F. Fenner, *Medical Virology*, 4th ed., 1994.

Herrings

The common name for a family (Clupeidae) of about 70 genera of fishes in the order Clupeiformes. Used extensively as food all over the world, they occur in all seas except the Arctic and Antarctic. Many live in vast shoals and are caught in great numbers.

These fishes are the most primitive of the higher bony fishes. The fins have no supporting spines and are soft-rayed. There are usually four gill clefts, with the pectoral fins behind the gill openings. Scales are present on the body but absent on the head, and the swim bladder and lateral line may be missing. *See* SWIM BLADDER.

Some species live permanently in fresh water, while many species are anadromous, spending most of their life in the sea and spawning in fresh water. When spawning, the female extrudes the eggs in estuaries while rubbing against seaweeds, rocks, or other objects to which the eggs adhere. The male swims closely behind the female and fertilizes the eggs. The herrings along the North American shores breed once a year, while those in the seas along the Scandinavian coast breed twice yearly, in the spring and autumn. Herrings mature at the age of 4 years, but are not suitable for spawning until about 5–8 years. After spawning, mature herrings disperse to their feeding grounds and congregate in large schools. Mature fish return to their specific spawning grounds during the breeding season. Some herrings live 20 years.

A shad, typical representative of the herring family.

The herring *Clupea harengus* has a circumpolar distribution. About eight other species of this genus are recognized, including the sprat or brisling (*C. sprattus*), which occurs in the Mediterranean and seas of western Europe, and the gizzard shad (*Dorosoma cepedianum*), which is a common species in the Potomac River. In Europe the herring is either salted, pickled, or smoked and cured as kippers. In Canada and the United States young herring are canned as "sardines."

Sardine. This herring, *Sardina pilchardus*, is known commonly as the pilchard and is found along the European coasts in the Atlantic. *Sardina sagax* occurs along the coasts of Japan, Chile, California, and South Africa. Pilchards occur in large schools, feed on pelagic organisms, and spawn in the summer. The entire fish may be processed and preserved in oil, since the bones are soft and all parts are edible.

Shad. These herrings, members of the genus *Alosa*, occur in northern waters on both sides of the Atlantic (see **illus.**). They spawn in fresh water but spend most of their life in the sea. They breed in rivers from spring to early summer, and the fry remain in fresh water for a year or so. The allis shad (*A. alosa*) occurs along the coasts of Europe and may reach a length of 30 in. (75 cm) and weigh 8 lb (3.6 kg). The American shad (*A. sapidissima*) is valued for food.

Anchovies. A family of herringlike fish, the Engraudidae together with the Clupeidae belong to the suborder Clupoidea. They are found in the Mediterranean and range along the European coast as far north as Norway. Spawning occurs in the summer; the female lays eggs which are sausage-shaped and transparent, and float near the surface. Within 3 days the minute, transparent larvae hatch. *See* CLUPEIFORMES.

Charles B. Curtin

Hertzsprung-Russell diagram

A two-dimensional diagram used extensively in astronomy, developed independently by Ejnar Hertzsprung in 1911 and Henry Norris Russell in 1913. In its original form, the Hertzsprung-Russell (H-R) diagram was a plot of absolute visual magnitude versus spectral type (O, B, A, and so on). Variants are now commonly used, avoiding requirements of and uncertainties due to spectral classification. The vertical axis of the diagram is some suitable measure of the power output of the star, while the horizontal axis indicates the temperature (or color) of the star's visible surface, or the corresponding spectral type. Each point in the plot represents a nearby star of known distance. In any of its forms, the diagram reveals the most fundamental correlation among observed stellar properties discovered to date. *See* MAGNITUDE (ASTRONOMY); SPECTRAL TYPE.

Observational and theoretical forms. Different quantities and units may be used for the two axes of the Hertzsprung-Russell diagram depending on whether it relates observational or theoretical stellar properties. Observationally, a star's power output is estimated from measurements made in a specific photometric (wavelength) band, with resulting absolute magnitude determined from the measured apparent magnitude using a knowledge of the star's distance. In its observational form, also referred to as a color-magnitude diagram, absolute magnitude is used as ordinate, although apparent magnitude may be used for a collection of stars at a common distance. Brighter stars (that is, those with higher luminosities, and smaller numerical values of the magnitude) appear at the top of the diagram. The color scale is usually a color index constructed as the difference between the magnitudes measured in two chosen spectral bands. For historical reasons, the bluest color index (corresponding to the highest temperatures) appears at the left. *See* COLOR INDEX.

In its theoretical form, the Hertzsprung-Russell diagram relates more physically fundamental stellar quantities, which can be predicted or modeled from theories of stellar structure and evolution. In this case the ordinate gives the total bolometric magnitude, or the bolometric luminosity, which measure the star's total power output, usually on a logarithmic scale normalized to the luminosity of the Sun. An effective temperature is used as the theoretical counterpart to the color index. Closely related to the temperature of the stellar surface, the effective temperature is tied to the luminosity and stellar radius through the Stefan-Boltzmann law. Consequently, lines of constant stellar radius run diagonally across the Hertzsprung-Russell diagram, from top left to bottom right, with stellar radius increasing from bottom left to top right. *See* HEAT RADIATION.

Theoretical modeling describes stellar structure and evolution in terms of fundamental properties such as mass, luminosity, surface gravity, and effective temperature. These properties may in principle be transformed into observable quantities such as magnitudes and color indices. It is the comparison between observed and theoretical quantities which allows stellar structure and evolutionary theories to be tested and refined. As such, the Hertzsprung-Russell diagram provides a very powerful framework for the verification and development of theories of star formation and evolution, ultimately leading to a rigorous foundation for exploring and explaining the evolutionary history of the Milky Way Galaxy as a whole.

There are numerous complications in transforming between theoretical and observed quantities. A major problem in deriving the power output is that a star's apparent magnitude measured at a telescope must be converted into an absolute magnitude, requiring an accurate knowledge of the star's distance. Direct distance determinations in astronomy are based on trigonometric parallax

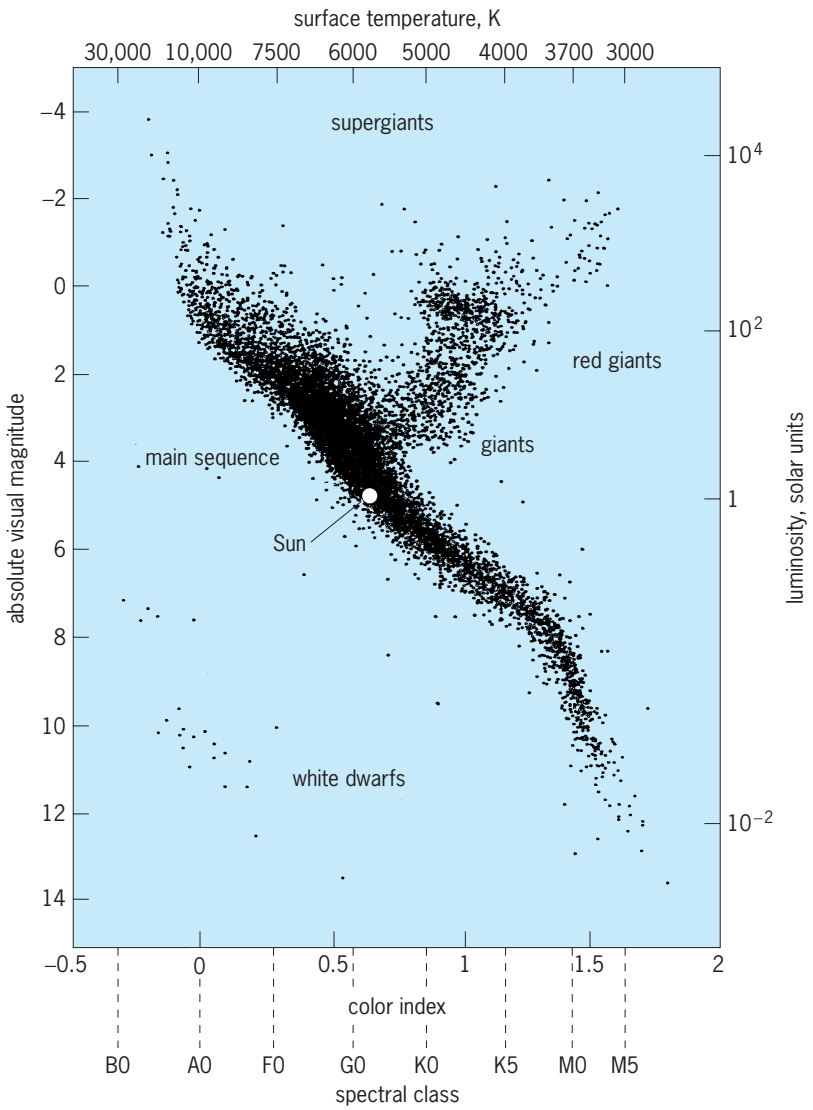

Fig. 1. Hertzsprung-Russell diagram for about 15,000 stars within a sphere of radius 100 parsecs, taken from the *Hipparcos Catalogue*. The color index and absolute visual magnitude scales are directly measured. The spectral class, surface temperature, and luminosity (in terms of solar luminosity) are approximate relationships appropriate for the main sequence.

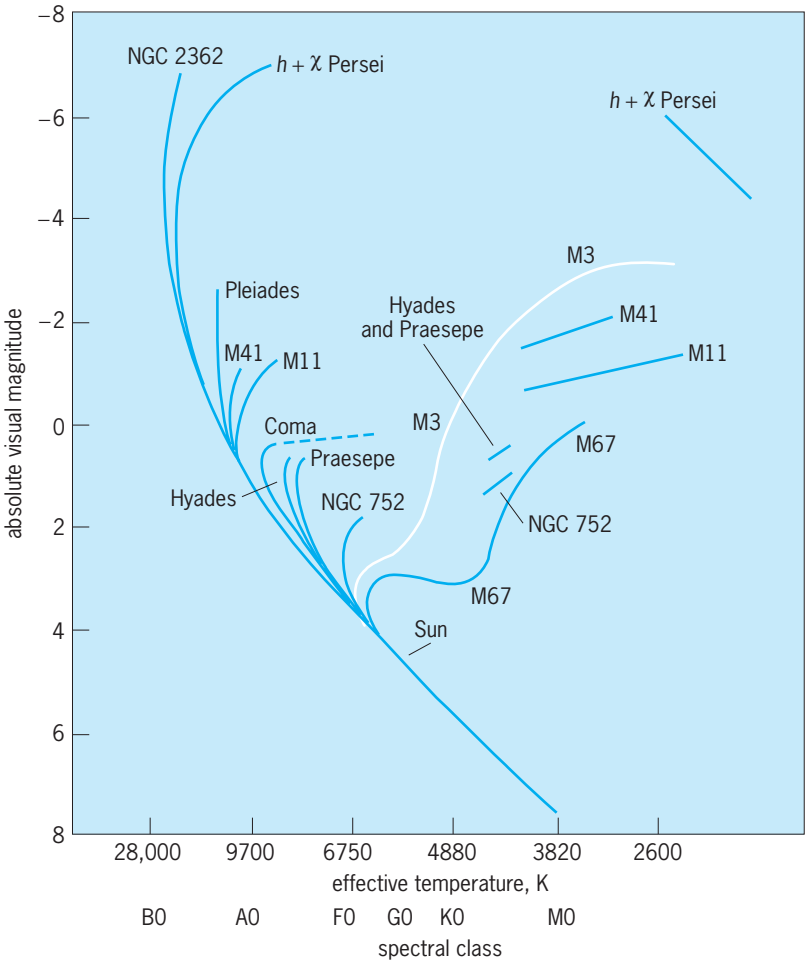

Fig. 2. Composite Hertzsprung-Russell diagram for several representative galactic clusters. NGC 2362 is the youngest and M67 the oldest cluster shown in the diagram, with ages of about 2×10^6 and 5×10^9 years, respectively. (*After A. Sandage, Observational approach to evolution, I. Luminosity functions, Astrophys. J., 125:435–444, 1957*)

measurements, which employ the motion of the Earth around the Sun to provide an estimate of the star's distance by triangulation. Ground-based parallax determinations are complicated by atmospheric motions and other disturbances, and such distance estimates have provided accuracies better than about 20% for only a few hundred nearby stars. The *Hipparcos* satellite, which was devoted to stellar distance measurements from space, has allowed accurate determination of individual stellar distances for many tens of thousands of stars out to distances of a few hundred light-years from the Sun, providing a significant clarification in the structure of the observational Hertzsprung-Russell diagram. *See* PARALLAX (ASTRONOMY).

General features. Figure 1 shows the Hertzsprung-Russell diagram for about 15,000 single stars from the compilation of nearly 120,000 stellar distances measured by the *Hipparcos* satellite. The sample has been limited to those with relative distance errors smaller than 10%, and lying within a sphere of radius 100 parsecs (326 light-years, 3.1×10^{15} km, or 1.9×10^{15} mi), which is around 1% of the distance from the Sun to the center of the Galaxy. The figure illustrates the basic features of the classical Hertzsprung-Russell diagram. The absolute visual magnitude scale runs from −5 to 15, corresponding to a range of 10^8 in star luminosity. The color index scale corresponds to effective temperatures ranging from around 100,000 K (180,000°F) at the left to about 2500 K (4000°F).

From the upper left (blue, high-luminosity stars) to the lower right (red, low-luminosity stars) a prominent concentration of objects defines the main sequence. Stars located on the main sequence are also called dwarfs. They include stars such as Sirius, and are assigned luminosity class V in the MK stellar classification system. (In this system, two parameters, spectral type and luminosity class, categorize each star.) Along the main sequence, the luminosity of a star and its surface temperature are tightly correlated. Stellar structure theory successfully models this relationship. The main-sequence stars are at the early phases of their lives, and are powered by the fusing of hydrogen to helium in their centers. Masses of the main-sequence stars increase going toward the upper left of the diagram (reaching almost 100 times the Sun's mass) and decrease going to the lower right (to about one-tenth of the Sun's mass). Due to their higher central temperatures and pressures, the more massive stars are burning hydrogen more rapidly and are therefore brighter. The main sequence has a certain width, which arises from numerous contributions: remaining uncertainties in the distance and magnitude measurements, stellar variability, the presence of undetected binary systems, evolutionary changes, and variations in the chemical abundances of the stars. *See* BINARY STAR; DWARF STAR; VARIABLE STAR.

Extending from the main sequence in the direction toward lower temperatures, and at roughly constant luminosity, are the luminosity class III giant stars (such as Vega) and the clump of more luminous red giants. Even more luminous supergiants, of luminosity class I (such as Arcturus and Procyon), are sparsely represented but occupy a broad range of color index at the very highest luminosities. They reach absolute magnitudes of less than −5, corresponding to luminosities some 10^4 times brighter than the Sun, and with radii around 1000 times that of the Sun. The lower left part of the diagram is not entirely empty and contains the white dwarfs: hotter than the Sun, but much less luminous (typically 10^4 times fainter) and of much smaller radius (about 1% of the Sun's radius). *See* GIANT STAR; SUPERGIANT STAR; WHITE DWARF STAR.

The distribution of stars in the Hertzsprung-Russell diagram based on observations such as those in Fig. 1 gives a somewhat incomplete impression of the occurrence of stars of different spectral type and luminosity class for two reasons, related to the luminosity and to the space density of different types of stars. First, such measurements generally extend only down to some observational limit in apparent magnitude, around a visual magnitude of 12 in the case of the *Hipparcos* satellite observations. In any given volume of space, say one extending to a

radius of 100 parsecs, the intrinsically low-luminosity stars, such as the white dwarfs and the fainter red main-sequence stars, are actually too faint to be observable farther out, even though their space density may be very high. Second, when the Hertzsprung-Russell diagram is constructed using well-defined limits on the accuracy of the distance measurements, the very luminous supergiants, which are relatively rare but can nevertheless be seen in considerable numbers out to very large distances, are poorly represented. For the same reason, other relatively uncommon stars, such as the hot but subluminous nuclei of planetary nebulae, which would appear between the main sequence and the white dwarf sequence, are not represented in Fig. 1. To determine distances (and hence luminosities) for rarer types of stars farther out in the Milky Way Galaxy, a range of alternative, indirect methods must be adopted. Their corresponding locations in the Hertzsprung-Russell diagram, the interpretation of their physical properties and, ultimately, a complete and unified picture of stellar and galactic evolution rest heavily on the confidence that can be placed in these distance estimates.

Stellar evolution. The Hertzsprung-Russell diagram says nothing, at least directly, about the mass, chemical composition, or age and state of evolution of a star. However, comparisons between observations (such as Fig. 1) and the predictions of stellar evolution theory allow stringent constraints to be placed on models of the structure, chemical composition, and evolution of stars. In the pre-main-sequence evolutionary phase, protostars form from collapsing material contained in interstellar clouds, liberating gravitational potential energy, and increasing in mass and temperature to the point where thermonuclear fusion of hydrogen into helium becomes an important energy source. At this point, the star settles onto the main sequence, its precise position depending primarily on its mass but also on its chemical composition. The subsequent evolution, in which hydrogen is slowly consumed in the stellar core, has such a long duration that most visible stars are found in this phase. The Sun, for example, is roughly halfway through its estimated main-sequence lifetime of about 10^{10} years. When the supply of hydrogen in the central region is depleted, structural adjustments occur, making the star expand and decreasing the temperature of its outer layers. The star becomes a red giant, and the point representing it on the Hertzsprung-Russell diagram moves up and to the right of the main sequence. Other changes occur at successive evolutionary stages and depend upon the mass, chemical composition, and other detailed properties. In particular, high-mass stars evolve at a much faster rate than low-mass stars and exhaust their nuclear fuel over a much shorter interval of time. *See* PROTOSTAR.

The effects of evolution are particularly evident in open clusters, physical aggregations of typically several hundred stars, in which all objects are essentially at the same distance from the Sun. Due to the common distance of the objects, the detailed features of the diagram are preserved when working in terms of apparent magnitude, even though the zero point of the absolute magnitude scale may be undetermined by the observations. Clusters are important laboratories for testing theories of star evolution, since all stars in a cluster are believed to have been formed at the same time. In a cluster of a given age, all stars brighter than a given limit will have evolved off the main sequence (**Fig. 2**). Thus, while the youngest clusters, such as NGC 2362, still contain hot O stars, progressively older clusters contain no stars bluer than, or to the left of, a certain turn-off point. As the cluster ages, the turn-off point moves down the main sequence to stars of progressively lower mass. In this way, the age of a star cluster can be inferred from its appearance of its Hertzsprung-Russell diagram. *See* STAR CLUSTERS.

Interpretation is further complicated by the effects of the star's initial chemical composition. Stars with the same mass but of different chemical composition do not lie at the same location in the Hertzsprung-Russell diagram. Stars categorized by astronomers as population II (for example, globular cluster stars) are old objects, having estimated ages nearly equal to that of the universe itself. These objects formed early in the history of the Milky Way Galaxy, with low abundances of elements heavier than hydrogen and helium, and a chemical composition nearly reflecting that of the primordial material of the universe. In contrast, open clusters are population I objects, having chemical abundances similar to that of the Sun. Their constituent stars were formed from material already enriched by the products of earlier phases of star formation, and have a wide range of ages, dating anywhere from the initial formation of the Milky Way Galaxy up to the present time. *See* MILKY WAY GALAXY; STAR; STELLAR EVOLUTION; STELLAR POPULATION.

Michael A. C. Perryman

Bibliography. S. P. Maran, *The Astronomy and Astrophysics Encyclopedia*, John Wiley, 1997; S. Mitton, *The Cambridge Encyclopedia of Astronomy*, Cambridge University Press, 1984; F. H. Shu, *The Physical Universe: An Introduction to Astronomy*, University Science Books, 1982.

Hesperornithiformes

An extinct group of Cretaceous toothed birds. All members were foot-propelled, seagoing birds that could reach the size of a modern Emperor penguin. The largest known hesperornithiform reached a maximum length of over 5 ft (1.5 m). All of the known fossils are from marine rocks that are Cretaceous in age (about 100–65 million years old); they are most abundant from North America, especially Kansas. Although these birds are thought to be closely related to living birds (the group or clade Neornithes)—along with Ichthyornithiformes—the evolutionary interrelationships among Hesperornithiformes are not well understood by paleontologists. *See* AVES; CRETACEOUS; ICHTHYORNITHIFORMES; NEORNITHES.

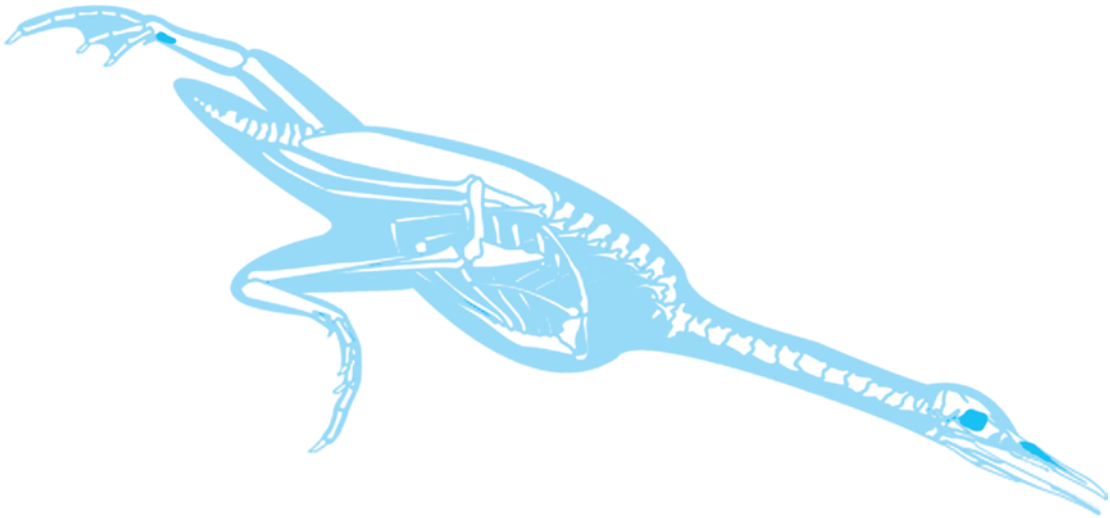

***Hesperornis*, a marine hesperornithiform.**

Fossil record. Hesperornithiforms are known from all over the Northern Hemisphere—North America, England, Sweden, Russia, Kazakhstan, and Mongolia—and are preserved in rocks that record shallow, epicontinental seas. As these birds had only rudimentary wings, it is known that they were flightless; the miniscule forelimbs may have been used for steering when swimming. Soft-tissue preservation has revealed that the feathers of Hesperornithiformes were filamentlike, much like many flightless birds today.

Description. Hesperornithiforms had highly reduced, nonfunctional wings. They probably spent almost all their time in the water, except presumably in the breeding and egg-laying season. Like modern-day cormorants, hesperornithiforms appear to have led an aquatic lifestyle, probably diving to catch fish.

One of the best-known hesperornithiforms is *Hesperornis* (see **illustration**). When this bird was first described from the Late Cretaceous rocks of Kansas in the 1870s, it caused a great deal of excitement around the world. *Hesperornis* was touted as a modern bird, living at sea, and some 50 million years younger than its counterpart (and the only other fossil bird known at the time), *Archaeopteryx* ("ancient-wing") from the Jurassic of Germany (about 150 million years old). *Hesperornis* caused a stir because, unlike *Archaeopteryx*, it had lost its ability to fly, but like *Archaeopteryx*, still possessed teeth. The lower jaw (dentary) and the back portion of the upper jaw (maxilla) of *Hesperonis* had many sharp teeth. It is likely that they swam and fed much like modern penguins, preying on small fish. *See* ARCHAEORNITHES.

Extinction. Hesperornithiformes are one of the several lineages of nonmodern birds (along with, for example, the Ichthyornithiformes) that disappear from the fossil record at the end of the Cretaceous. No records of these birds are from the succeeding geological period, the Tertiary, although this was the time of the massive modern avian evolutionary radiation. *See* EXTINCTION (BIOLOGY). Gareth Dyke

Bibliography. S. Chatterjee, *The Rise of Birds*, Johns Hopkins Press, Baltimore, 1997; L. M. Chiappe and L. M. Witmer, *Mesozoic Birds: Above the Heads of Dinosaurs*, University of California Press, Berkeley, 2002; A. Feduccia, *The Origin and Evolution of Birds*, 2d ed., Yale University Press, New Haven, 1999; S. L. Olson, *The Fossil Record of Birds*, in *Avian Biology*, vol. 8, Academic Press, New York, 1985.

Heterochrony

An evolutionary phenomenon that involves changes in the rate and timing of development. As animals and plants grow from their earliest embryonic stages to the adult, they undergo changes in shape and size. This life history of an individual organism is known as its ontogeny. The amount of growth that an organism experiences during its ontogeny can be more or less than its ancestor. This can apply to the organism as a whole or to specific parts.

Evolution can be viewed as a branching tree of modified ontogenies. Heterochrony that produces these changes in size and shape may be the link between genetics at one extreme and natural selection at the other.

Processes. If a character of one species in an evolutionary sequence undergoes less growth than its ancestor, the process is known as pedomorphosis. If it undergoes more growth, the process is known as peramorphosis. Each state can be achieved in three basic ways, by varying the timing of onset, offset, or rate of development.

Pedomorphosis. If development is stopped at an earlier growth stage in the descendant than in the ancestor (for example, by earlier onset of sexual maturity), ancestral juvenile features will be retained by the descendant adult (progenesis). If the onset of development of a particular structure is delayed in a descendant, the structure will develop less than in the ancestor (postdisplacement). The third process

that produces pedomorphosis is neoteny, whereby the rate of growth is reduced. Few, if any, organisms grow isometrically. Different morphological features grow with varying allometries; some are positive, some negative. Changing growth rates change the allometric coefficients and produce a structure in the descendant that is different in shape from that in the ancestor.

Peramorphosis. The three ways that this type of growth can be achieved are just the opposite at pedomorphosis. Development can start earlier in the descendant than in the ancestor (predisplacement); or the rate of development can be increased, thus increasing the allometric coefficient (acceleration); or development can be extended by a delay in the onset of sexual maturity (hypermorphosis). In such cases, ancestral allometries are extended, producing descendant adults that are morphologically quite different from their ancestors.

Heterochrony can affect both the differentiation of meristic characters (differentiative heterochrony), such as, the timing of induction of limb bud growth in vertebrates or the formation of spines in a sea urchin, and also affect the structure once it has been formed (growth heterochrony). The complex interplay between these two aspects of heterochrony may have played a crucial role in the evolution of a number of major evolutionary novelties.

Cellular level. As an organism grows, the numbers of cells that it produces increases. Ultimately, changes to rate and timing of growth are reflections of changes to the timing of onset and rate of cell development, and the balance between cell growth and cell death. Morphogens and growth hormones play a major role in controlling development, in terms of initiation, rate of division, and migration. Therefore, changes to the timing of their expression affect the shape and size of the final adult structure. Inception of hormonal activity is under the control of genes that regulate the timing of its production.

Variations between individuals or species in the timing of induction of cellular development are manifested in differentiative heterochronic changes. Changes to the rates of cell division are reflected in growth heterochronies, affecting the shape and size of specific morphological structures. Any slight perturbations to cellular differentiation early in embryonic development can have major effects on the resultant adult form. Growth heterochronies occurring later in ontogeny produce relatively minor morphological differences. In addition, cell size can be a factor in determining heterochronic processes. Pedomorphic salamanders possess larger cells than nonpedomorphic forms. Likewise, trends toward the evolution of pedomorphic species of lungfishes over the last 350 million years are accompanied by an increase in cell size.

Variation within species. Much intraspecific variation within animals or plants involves slight differences in the shape or size of morphological features, or overall body size, and is likely to have arisen by heterochrony. The importance of change in growth proportions during ontogeny is seen by comparing the growth of a dog skull with that of a cat skull. In terms of overall body shape and proportions, breeds of dogs exhibit a far greater range of morphologies than do breeds of domestic cats. This greater range of morphologies arises from the much lower proportionate growth rate of the cat skull.

A number of species of insects show polymorphism, whereby a number of morphologically quite distinct forms occur within a single species. Many of these polymorphisms are due to heterochrony. For example, parthenogenetic thrips are always males and morphologically more advanced than nonparthenogenetic forms. This condition is accompanied by progenesis, that is, precocious onset of sexual maturity. Whether a female produces such young is determined by environmental factors, in this case photosensitivity that results in the production of a neurosecretory brain hormone, virginoparin. Aphids that lack wings are also an example of pedomorphosis. This condition is also attained by progenesis, and such forms experience rapid reproduction, produce more offspring, and have shorter generation times.

Sexual dimorphism in animals often arises from heterochrony. This frequently involves variations in body size, arising from either differences in maturation times or different growth rates. Thus, in *Homo sapiens* the generally smaller adult female body size arises from an earlier onset of sexual maturity, juvenile growth trajectories prior to maturation being similar in the two sexes.

Variations between species. Heterochrony is a crucial factor in speciation, providing the raw material upon which natural selection operates. Many examples of speciation, both from the fossil record and from living examples, are the product of heterochrony. Species are usually a product of a mosaic of heterochronic characters. Some features demonstrate a higher level of development than in the ancestor by peramorphic processes; others developed less by pedomorphic processes. *See* MOSAICISM; SPECIATION.

A classic example of evolution is the numerous species of finches first described from the Galápagos Islands by Charles Darwin. These finches possess a wide variety of bill shapes that have allowed a wide range of food types to be utilized. Some bills are used for crushing; others are for probing or biting. Such adaptations are generally regarded as having evolved in order for the species to feed on a particular object. However, heterochrony played a crucial role in the evolution of these bills. Because of small changes to the rates of growth of upper and lower bills, some reducing in size and others growing more and becoming much more robust, a wide range of species has evolved. If suitable food sources were available that could be exploited by a particular bill shape, this shape would be selected for and a new species would evolve.

Thus the evolution of species arises from the natural selection of particular morphologies that have arisen by heterochronic changes to the rates and timing of growth, and that confer an adaptive

advantage to the organism. Because such adaptive changes often result in changes to life history strategies, resultant behavioral changes can in some instance be considered to have arisen ultimately from heterochrony.

Evolutionary trends. Many evolutionary trends documented from the fossil record were fueled by heterochrony. Through a succession of species, certain traits that evolved are shown to be either increasingly pedomorphic or increasingly peramorphic. Such trends are called, respectively, pedomorphoclines and peramorphoclines and have evolved along a particular environmental gradient. Many examples have been described from the fossil record in trilobites, echinoids, ammonoids, brachiopods, and bivalves.

A pedomorphocline develops when an ancestral species that passes through a number of ontogenetic stages (say, A to M) gives rise to a descendant species that passes through fewer (say, stages A to K). As the adult state K occurred in the ancestral juvenile, the descendant species is pedomorphic. If this species subsequently gives rise to a pedomorphic species that passes through only stages A to I, a pedomorphocline has been established. A peramorphocline exists when successive younger species pass through successively more morphological stages during their ontogeny. Often a mosaic of heterochronic characters occurs in such trends, with some traits forming a pedomorphocline, and others forming a peramorphocline, in the same sequence of species.

Major evolutionary novelties. While groups such as flightless birds, amphibians, and humans are thought to have arisen by pedomorphosis, it is more likely that a critical combination of pedomorphic and peramorphic characters (mosaic heterochrony) was more important in causing such drastic evolutionary changes. For instance, the evolution of horses through the Cenozoic in North America shows trends of increased body size (by hypermorphosis), while changes to the structure of the leg arose by a pedomorphic reduction in digit number, with a peramorphic increase in growth of the one remaining digit to produce a hoof.

Human development has for a long time been ascribed to the pedomorphic retention of many primate juvenile characters. However, while a few features, such as dental development, may be pedomorphic, the most distinctive features of the human species are relatively long life-span, long juvenile growth period, large brain size, and large body size. These are peramorphic features produced by a sequential hypermorphosis, wherein the time spent in successive preadult periods is extended, relative to the same time periods in other primates. *See* ANIMAL EVOLUTION.

Kenneth J. McNamara

Bibliography. S. J. Gould, *Ontogeny and Phylogeny*, 1977; B. K. Hall, *Evolutionary Developmental Biology*, 1998; M. L. McKinney (ed.), *Heterochrony in Evolution: A Multidisciplinary Approach*, 1988; M. L. McKinney and K. J. McNamara, *Heterochrony: The Evolution of Ontogeny*, 1991; K. J. McNamara (ed.), *Evolutionary Trends*, 1990.

Heterocorallia

A small, extinct late Paleozoic order of fossil corals known from Europe, North Africa, Asia, North America, and Australia but limited to the Middle Devonian and to the Late Mississippian. They are found in calcareous shales and in limestones. The preserved material, called a corallite, is a calcareous substrate secreted by a scleractinian-like polyp. The corallites attained elongated shapes (**Fig. 1***a*) and show a bifurcational pattern of septal insertion during growth in the diameter. Nevertheless, the known materials show little change in diameter from one end to the other. There are rare indications of branching, which can be interpreted as budding of new polyps in the soft tissues of the original polyp. *See* DEVONIAN; MISSISSIPPIAN.

Each corallite is a framework of septa (radial longitudinal elements) and tabulae (horizontal skeletal elements). Each such element consists of fibrous calcium carbonate; the fibers are perpendicular to the growth lamellation, which is weakly apparent in this section (Fig. 1*b–d*). An external sheath, or epitheca, is absent or very rare.

The septa and tabulae of heterocorals justify their classification in the subclass Zoantharia of the class Anthozoa. However, the pattern of septal insertion indicates that they are a separate group (that is, a separate evolutionary lineage) aside from the Tabulata, Rugosa, and Scleractinia.

As interpreted by O. H. Schindewolf, there are four original septa. Each of these may split into two near the periphery; new septa are inserted only in the four spaces formed between the two split portions of each original septum. Of these four insertion spaces, two opposing ones regarded as lateral show

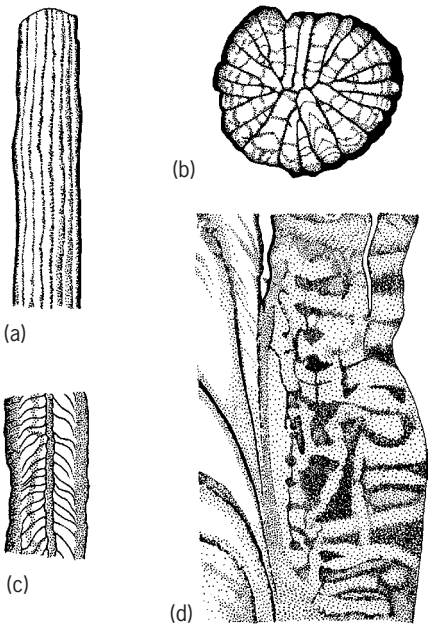

Fig. 1. Morphology of the heterocorallian skeleton. (*a*) External view. (*b*) Transverse thin section. (*c*) Median longitudinal thin section. (*d*) Enlarged part of c. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. F, Geological Society of America, University of Kansas Press, 1956*)

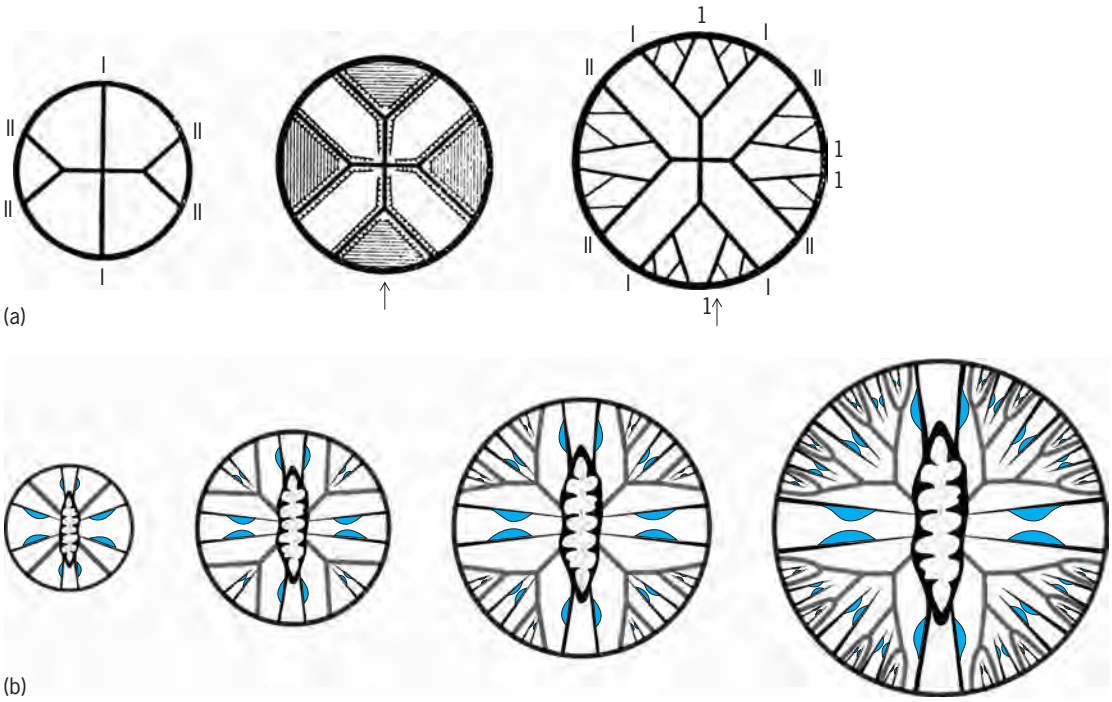

Fig. 2. Details of the heterocorallian body. (*a*) Generalized pattern of septal insertion (*from O. Schindewolf, Grundfragen der Paläontologie, Schweizerbart, Stuttgart, 1950*). (*b*) Reconstruction of the mesenterial in the soft body (*from M. Gudo, Konstruktion, Evolution und riffbildendes Potential rugoser Korallen, Courier Forschungsinstitut Senckenberg, 228:1–153, 2001*).

far more newly formed septa than the remaining two. This pattern of septal insertion, and the knowledge of the general functional design of cnidarian polyps, allows the reconstruction of the soft body of the heterocorals (**Fig. 2**). The soft body of the polyp must have had six original twins of mesenteries and four growth zones in which new twins of mesenteries were inserted in an exponential manner. As a result, the septa which mold the arrangement of mesenteries were formed in a bifurcational pattern.

The order contains only one family, the Heterophylliidae, with two genera, *Hexaphyllia* and *Heterophyllia*. *See* CNIDARIA. Michael Gudo; Dorothy Hill

Bibliography. B. Berkowski, Famennian Rugosa and Heterocorallia from southern Poland, *Palaeontologia Polonica*, 61:3–88, 2002; M. Gudo, Konstruktion, Evolution und riffbildendes Potential rugoser Korallen, *Courier Forschungsinstitut Senckenberg*, 228:1–153, 2001; M. Montenari, U. Leppig, and D. Weyer, Heterocorallia from the Early Carboniferous of the Moldanubian Southern Vosges Mountains (Alsace, France), *Neues Jahrbuch für Geologie und Paläontologie, Abhandlungen*, 224(2):223–254, 2002; O. Schindewolf, *Grundfragen der Paläontologie*, Schweizerbart, Stuttgart, 1950; C. Teichert (ed.), *Treatise on Invertebrate Paleontology*, pt. F, suppl. 1, 1986.

Heterocyclic compounds

Cyclic compounds in which the rings include at least one atom of an element different from the rest. Homocyclic compounds are cyclic compounds in which all the ring atoms are the same. In organic homocyclic compounds the annular atoms are all carbons. If the molecule contains carbon atoms, then it is organic (most types of heterocyclic compounds studied to date are organic compounds). An example of an organic heterocyclic compound is oxazoline (**1**), which is composed of five annular atoms, two

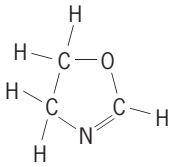

(**1**)

of which are not carbon. An example of an inorganic heterocyclic compound is phosphonitrilic chloride (**2**); the six-membered ring is composed of alternat-

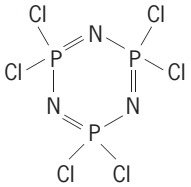

(**2**)

ing atoms of nitrogen and phosphorus.

The smallest possible ring is three-membered, for example, ethylene oxide (**3**), but very large rings are

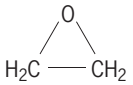

(**3**)

possible, as in the crown ethers, for example, 18-crown-6 (**4**). Thus, considerable diversity in the ring

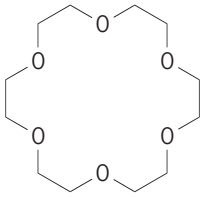

(**4**)

system is possible since not only may the size of the ring vary, but also the nature and number of heteroatoms in the ring can vary. The cycle may contain only single bonds and is thus saturated; it may include one or more double bonds; or it may possess aromatic unsaturation characteristics of benzene, that is, it is heteroaromatic. As with the homocyclic compounds, heterocyclic compounds can contain more than one ring, either heterocyclic or homocyclic.

The more familiar heteroelements in an organic compound are oxygen, nitrogen, and sulfur, but many elements have been or could be incorporated into a heterocyclic system. Compounds containing a ring formed by internal hydrogen bonding are not classed as heterocyclic. There is no limitation as to what kind of heteroatom may participate in ring formation, provided it has the appropriate bond angles and geometry.

Properties. The chemical properties of saturated heterocyclic compounds are usually very similar to those of appropriate open-chain analogs, provided account is taken of conformational differences. For example, the chemistry of *N*-methylpiperidine (**5**)

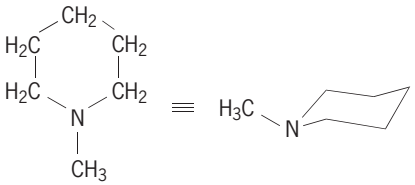

(**5**)

is not very different from that of ethylmethylpropylamine (**6**). Differences in properties originate in

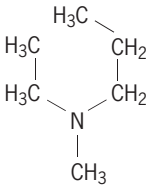

(**6**)

stereochemical factors. One such factor concerns four- and three-membered heterocycles in which the ring bonds are distorted with respect to bond angle and bond length, just as are those in cyclobutane and cyclopropane. These strained heterocyclic systems show enhanced activity in processes involving ring opening. Thus ethylene oxide (**3**), although formally an ether, is far more reactive than its open-chain analog, dimethylether (**7**).

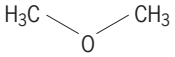

(**7**)

On the other hand, some polyunsaturated heterocyclic systems exhibit chemical properties implied in the term aromatic. For example, pyridine (**8**),

(**8**)

the parent of six-membered heteroaromatic nitrogen compounds, is a thermally stable material resisting oxidation, undergoing aromatic substitution (particularly nucleophilic substitution—the ring nitrogen atom makes electrophilic substitution difficult and the electrophile usually attacks the nonbonded pair of electrons on the ring nitrogen atom), and possessing appreciable resonance energy. Thiophene (**9**) is

(**9**)

a five-membered sulfur heterocycle, but it has aromatic properties similar to those of benzene. Such aromatic character may be expected in planar, unsaturated rings in which six electrons derived from ring unsaturation and from available electron pairs of the heteroatoms are delocalized over the whole ring system.

Nomenclature. Heterocyclic compounds may be named systematically. Many heterocycles, however, have nonsystematic names that are usually preferred by practicing chemists over the systematic ones. In the systematic approach to nomenclature the ring size is denoted by the appropriate stem. For example, three-membered saturated rings without nitrogen would have a name ending in -irane. The nature of the heteroatom is denoted by such prefixes as oxa-, thia-, or aza-, for oxygen, sulfur, or nitrogen, respectively. Thus, ethylene oxide (**3**) becomes oxirane. A five-membered unsaturated ring would have a name ending in -ole. Thiophene (**9**) thus becomes thiole. A six-membered unsaturated ring containing nitrogen would have a name ending in -ine according to this scheme. Pyridine (**8**) then becomes azine. Actually, the trivial names for these three systems are commonly accepted, and the systematic names are not often used. The degree of unsaturation is specified in the suffix as indicated in the **table**. It is important to note that the suffix is slightly modified when nitrogen is absent from the heterocyclic ring. The numbering of the ring begins with the heteroatom of highest priority and proceeds around the ring so as to give other heteroatoms or substituents the lowest number possible. The **illustration** shows how the systematic naming rules are applied and also how the ring atoms are numbered. An apex at which no

Suffixes for degree of unsaturation

	Number of double bonds in ring							
	With nitrogen				Without nitrogen			
Ring size	Max.	2	1	0	Max.	2	1	0
3			irine	iridine			irine	irane
4		ete	etine	etidine		ete	etene	etane
5	ole		oline	olidine	ole		olene	olane
6	ine	*	*	*	in	*	*	inane or ane
7	epine	*	*	*	epin	*	*	epane

* Add prefix, such as dihydro and tetrahydro, to the name of the ring with maximum unsaturation.

specific atom is shown [for example, the equivalent structure (**10**) for isoxazole] represents ═CH—.

Natural and synthetic compounds. Heterocyclic compounds are encountered in a very large number of groups of organic compounds. Inorganic heterocyclic compounds are also very common, but have not been studied systematically in terms of their heterocyclic characteristics in most cases. Naturally occurring heterocyclic compounds are extremely common as, for example, most alkaloids, sugars, vitamins, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), enzymic cofactors, plant pigments, many of the components of coal tar, many natural pigments (such as indigo, chlorophyll, hemoglobin, and the anthocyanins), antibiotics (such as penicillin and streptomycin), and some of the essential amino acids

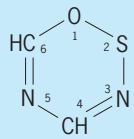

1,2,3,5-Oxathiadiazine

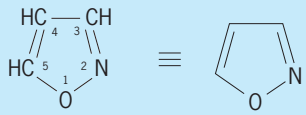

1,2-Oxazole, or preferably, isoxazole

(**10**)

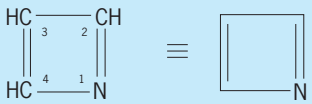

Azete or azacyclobutadiene

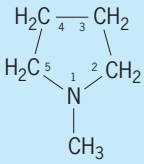

1-Methylazolidine, or preferably, 1-methylpyrrolidine

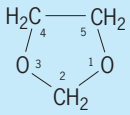

1,3-Dioxolane

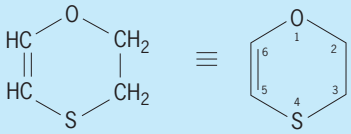

2,3-Dihydro-1,4-oxathine

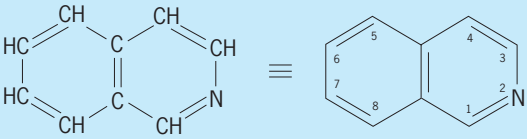

2-Azanaphthalene, or preferably, isoquinoline

Structural formulas of heterocyclic compounds. The ring positions are numbered.

(for example, tryptophan), and many of the peptides (such as oxytocin). Some of the most important naturally occurring high polymers are heterocyclic, including starch and cellulose. The major groups of natural products that are not mainly heterocyclic are the fats and most of the terpenes, steroids, and essential α-amino acids, though exceptions do exist.

Important synthetic dyes, such as phthalocyanins and phthaleins, as well as a large number of drugs, poisons, and medicinal products (both natural and synthetic) are heterocyclic. These include sulfothiazole, pyrethrin, strychnine, most of the antihistamines, the ergot alkaloids, morphine, the barbiturates, the tranquilizers Librium and Valium, and dihydrocannabinol (the main active constituent in marijuana). Many synthetic polymers (such as polyvinylpyridine, polyvinylpyrrolidone, and melamineformaldehyde), a number of industrial solvents (such as pyridine, dioxane, and tetrahydrofuran), and certain bulk chemicals (ethylene oxide, propylene oxide, and several aromatic nitrogen heterocycles obtained by synthesis from petroleum hydrocarbons or from coal tar) are heterocyclic.

For details about the specific heterocyclic systems *see* FURAN; INDOLE; PYRIDINE; PYRIMIDINE; PYRROLE; THIOPHENE.

R. A. Abramovitch

Bibliography. D. J. Coffey (ed.), *Rodd's Chemistry of Carbon Compounds*, vol. 4, pt. G: *Heterocyclic Compounds*, 1977; G. P. Ellis, *Synthesis of Fused Heterocycles*, 1987; T. L. Gilchrist, *Heterocyclic Chemistry*, 3d ed., 1997; R. R. Gupta, *Physical Methods in Heterocyclic Chemistry*, 1983; A. R. Katritzky, *Advances in Heterocyclic Chemistry*, vols. 1–32, 1963–1982; A. R. Katritzky and C. W. Rees (eds.), *Comprehensive Heterocyclic Chemistry: The Structure, Reactions, Synthesis, and Uses of Heterocyclic Compounds*, 8 vols., 1984.

Heterocyclic polymer

Essentially, linear high polymers comprising heterocyclic rings, or groups of rings, linked together by one or more covalent bonds. As the search has continued for polymeric materials having useful properties at high temperatures (500°C or 930°F or higher), much attention has been given to heterocyclic polymers. The possibility of forming rigid molecules that can be ordered into anisotropic arrays having exceptional stiffness may become of even greater interest. As a group such polymers are often both mechanically rigid and inherently resistant to thermal degradation.

Some of these polymers form molecules in which the rings are fused together, as shown symbolically in the **illustration** (ladder polymers), and some form molecules in which fused rings are joined by single bonds (stepladder polymers). Similar considerations hold for simple aromatic systems (for example, linear polymers of benzene), but the heterocyclic systems have been more useful in application.

In practice, the heterocyclic resins have rather high molecular weights and high glass transition temperatures. Some cross-linking may also be introduced during curing. Because of the insolubility and infusibility of the unmodified polymers, processing and fabrication was originally accomplished in stages. First, soluble prepolymers were prepared and fabricated into the final form desired (film, molding, coating, impregnated glass cloth, and so on). In this stage, the heterocyclic rings are not yet closed. Closure of the rings by condensation reactions was then effected by heating, and volatile by-products were eliminated. In contrast, poly(amide-imide) and the unmodified resins can be processed by more conventional techniques.

Major applications for these polymers are as metal-to-metal adhesives and as laminating resins for fibrous composites for structural applications in the aerospace industry. Other applications requiring both strength and resistance to oxidation at elevated temperatures have developed, including valve seats, bearings, and turbine blades. *See* POLYMERIC COMPOSITE.

Polyimide resins. The basic synthetic reaction to form the prepolymer is the condensation of an aromatic dianhydride with an aromatic diamine. Thus, pyromellitic dianhydride may be added to 4,4′-diaminodiphenyl ether in an anhydrous medium to give a high-molecular-weight poly(amic acid), as shown in reaction (1). Solvents such as *N*,*N*-dimethylformamide are suitable media. The ingredients must be extremely pure, and must be present in equal molar amounts. Other dianhydrides and diamines may be used.

The prepolymer solution may then be used to impregnate fiber-glass cloth or other reinforcement, or applied to other substrates. Then, the solvent must be driven off and the rings closed by condensation, usually in stages [reaction (2)]. Frequently, most of the solvent is removed in a precuring stage and the major condensation effected later, at temperatures in the range 180–380°C (360–720°F). A subsequent postcuring step, also at elevated temperatures, may be used. The structure shown is of the stepladder type; ladder structures may be obtained by use of phenyl or fused-ring diamines.

Properties depend on the structures of the ingredients and on the reaction and curing conditions. Fiberglass composites that are made by using

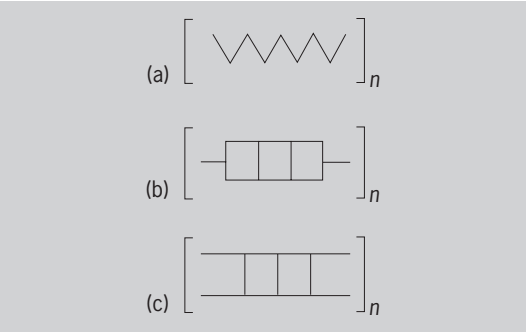

Structural units in linear polymers. (*a*) Simple linear polymer. (*b*) Stepladder polymer. (*c*) Ladder polymer.

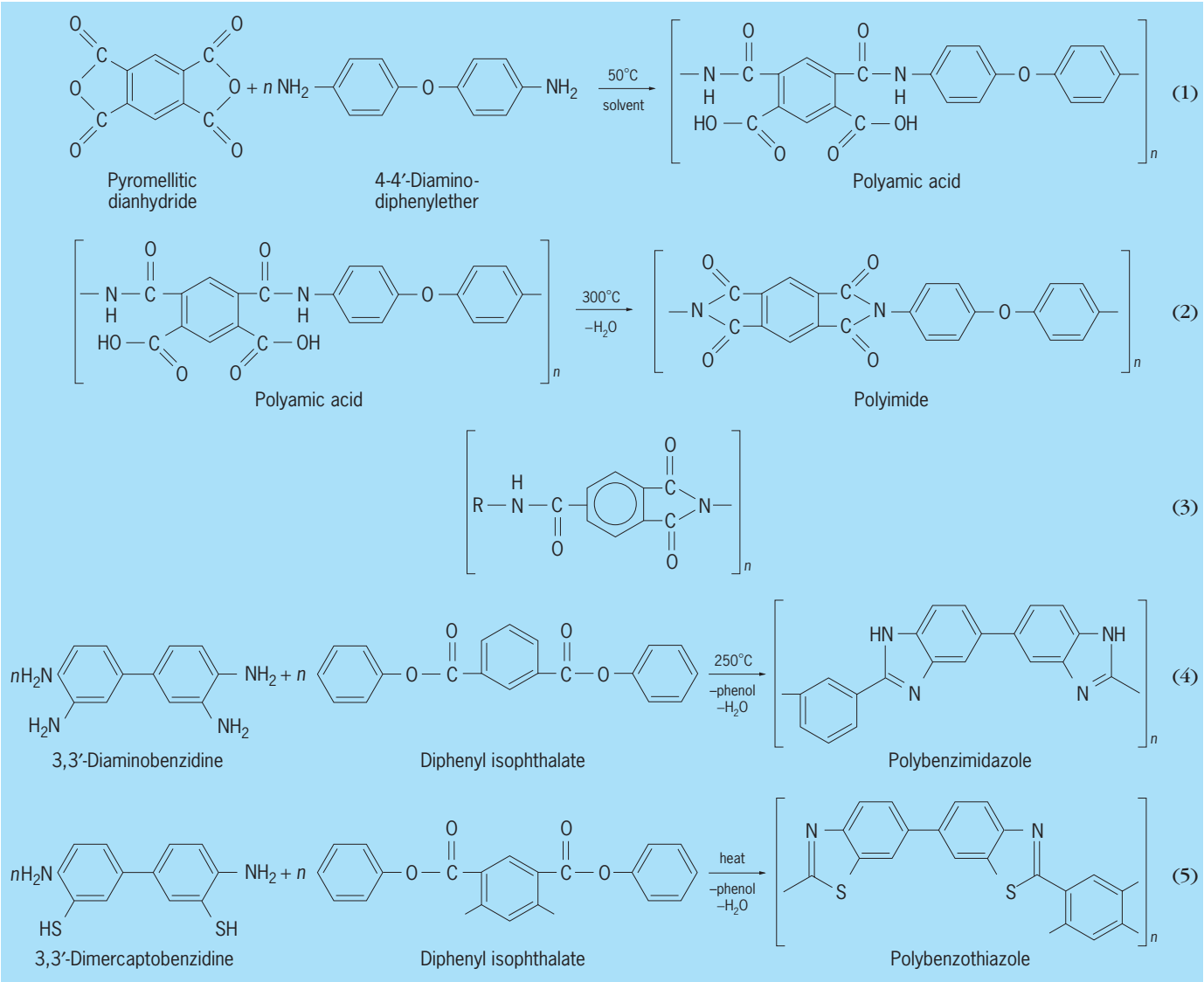

typical resins retain considerable strength after they are aged for 100 h at 500–600°C (930–1100°F). Adhesive bonds also show good resistance to aging at high temperatures. Certain polyimides are available as films and as molding powders, and some can be spun into fibers.

By introducing amide groups into a polyimide, some thermal stability is sacrificed, but improved processibility is gained. A typical structure is given in notation (3). Applications have developed in high-performance electrical connectors, engine parts, pumps, valves, and turbines.

Polybenzimidazoles. The basic reaction for the synthesis of aromatic polybenzimidazoles is the reaction of an aromatic tetramine with an aromatic diacid or diester. Although the details of the intermediate steps are not completely understood, the overall reaction for a typical example, the condensation of 3,3′-diaminobenzidine and the diphenylester of isophthalic acid, is as shown in reaction (4).

Polymerizations may be conducted in the melt, and to varying degrees of reaction. Partially polymerized resins may be dissolved in solvents such as *N,N′*-dimethylformamide and applied as a solution for laminating and adhesive applications. As with other heterocyclic polymers, final curing is effected at elevated temperatures (up to about 400°C or 750°F) under pressure.

Properties depend on the structures of the ingredients, and on reaction and curing conditions. In general, polybenzimidazoles are somewhat less stable in air at high temperatures than the polyimides. Applications are mainly as laminating and adhesive resins for composites and metals.

Polybenzothiazoles. Polybenzothiazoles are typically prepared by the reaction of a dimercaptobenzidine with an aromatic diacid, diester, or diacyl chloride, as shown in reaction (5).

A mainly soluble prepolymer is prepared by carrying the reaction forward to only a limited extent. As with the other resins in this family, the solvent is removed and the reaction is completed by heating after application to the substrate desired. Polybenzothiazoles find the same

applications as other heterocyclic polymers and are intermediate in stability between polyimides and polybenzimidazoles. *See* POLYESTER RESINS; POLYETHER RESINS; POLYMER; POLY-P-XYLYLENE RESINS; POLYSULFONE RESINS. John A. Manson

Bibliography. H. F. Mark, *Encyclopedia of Polymer Science and Technology*, 12 vols., 3d ed., 2004; K. L. Mittal (ed.), *Polyimides and Other High Temperature Polymers: Synthesis, Characterization, and Applications*, 2005; G. Odian, *Principles of Polymerization*, 4th ed., 2004.

Heterodontiformes

An order of galeomorph sharks that are commonly known as the bullhead sharks. These sharks are distinguished by the following combination of characters: a typically sharklike (subcylindrical) trunk; a head that is elevated above the eyes, but with a blunt and flattened snout; a small mouth; teeth that are similar in both jaws; spiracles present; eyes that lack a nictitating fold or membrane (which is at the inner angle of the eye or below the eyelid and is capable of extending over the eyeball); two dorsal fins, each with a well-developed spine; anal fin present; oviparous development; and an egg case that is horny with spiral flanges but no tendrils. They comprise one family (Heterodontidae), one genus (*Heterodontus*), and nine species, which vary in maximum total length from 57 to 165 cm (22 to 65 in.) [see **illustration**].

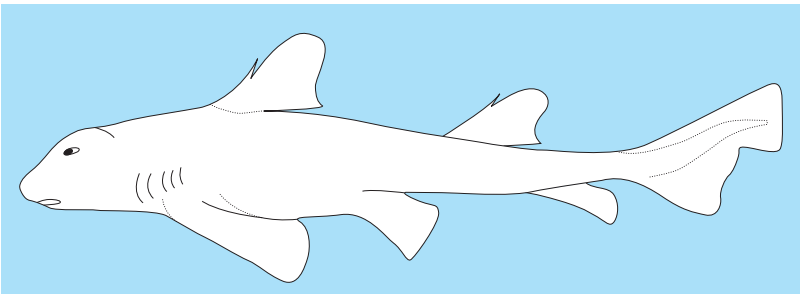

***Heterodontus* sp. (*J. S. Nelson, Fishes of the World, 4th ed., Wiley, New York, 2006*)**

Bullhead sharks occur in warm-temperate and tropical continental waters of the western Indian Ocean and western and eastern Pacific Ocean, but are absent from the Atlantic Ocean and from oceanic insular waters. They live on the bottom, often occupying crevices and caves by day and feeding at night on sea urchins, starfishes, crabs, mollusks, and small fishes. *See* ELASMOBRANCHII; SELACHII. Herbert Boschung

Bibliography. L. J. V. Compagno, Checklist of Chondrichthyes, pp. 503–547 in W. C. Hamlett (ed.), *Reproductive Biology and Phylogeny of Chondrichthyes: Sharks, Batoids, and Chimaeras*, Science Publishers, Enfield, NH, 2005; L. J. V. Compagno, Sharks of the world: An annotated and illustrated catalogue of shark species known to date, *FAO Species Catalogue for Fishery Purposes*, no. 1, vol. 2: *Bullhead, Mackerel, and Carpet Sharks* (*Heterodontiformes, Lamniformes, and Orectolobiformes*), FAO, Rome, 2001; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Heterodyne principle

The principle that multiple frequencies applied to a nonlinear device produce new frequencies that are sums and differences of the applied frequencies and their harmonics. Heterodyning, or mixing, is the process of multiplying a weak signal by a strong sinusoidal carrier, sometimes called the local oscillator, to shift the frequency of the signal in such a way that the information carried by the signal is preserved. *See* CARRIER (COMMUNICATIONS).

Applications. Signal frequencies are changed in a wide variety of electrical and electronic equipment, including radio, television, communications, radar, and measuring instruments. In modulation, low-frequency signals, such as audio or video, are translated to higher frequencies to permit the multiplexing of signals in a common communication system, transmission by a waveguide incapable of handling low frequencies, efficient radiation by an antenna of reasonable size, or generation of frequencies appropriate for physical measurements by a particular device. In demodulation, low-frequency audio or video information is removed from a high-frequency carrier at a receiver. In a superheterodyne receiver a low-power incoming signal is translated to a fixed lower-frequency band, eliminating the need for tuning an intermediate-frequency amplifier. Input signals of different frequencies are accommodated by merely changing the frequency of the local oscillator. Parametric amplifiers provide power gain to a small signal at the same frequencies or at changed frequencies. *See* DEMODULATOR; MODULATION; MODULATOR; PARAMETRIC AMPLIFIER; RADIO RECEIVER; TELEVISION RECEIVER.

Nonlinear devices. New frequencies cannot be generated in linear elements; nonlinearity is required. The form of nonlinearity varies with technology, ranging from iron-cored devices at power frequencies to minute semiconductor devices at microwave or millimeter-wave frequencies and nonlinear optical materials in laser systems. These nonlinear devices may often be idealized as nonlinear resistors, capacitors, or inductors. *See* INDUCTOR; NONLINEAR OPTICAL DEVICES; RESISTOR.

The analysis of nonlinear devices is difficult and has usually been restricted to special cases. The Manley-Rowe relations are an exception, providing general relations between the powers at the different frequencies in nonlinear inductors and capacitors. These results are remarkable in that they are independent of the shape of the nonlinear characteristic, the magnitudes of the signals at the different frequencies, and the external circuit connected to the nonlinear elements. They delineate what is possible and what is not. For example, they show which devices are stable and which are not; and

they indicate that up-converters, used to produce a high frequency from a low frequency in transmitters, can have gain, whereas down-converters, used to reduce the input frequency to a lower frequency in receivers, have loss. They have been applied to the Raman effect at optical frequencies and to a wide variety of physical systems, including rotating electrical machinery, nonlinear inductors and capacitors and equivalent semiconductor devices, and microwave ferrites. *See* ELECTRIC ROTATING MACHINERY; ELECTRICAL INSTABILITY; FERRITE DEVICES; GAIN; RAMAN EFFECT.

Harrison E. Rowe

Bibliography. J. M. Manley and H. E. Rowe, Some general properties of nonlinear elements, I. General energy relations, *Proc. IRE*, 44:904–913, 1956; M. Schwartz, *Information Transmission, Modulation, and Noise*, 4th ed., 1990; H. Taub and D. L. Schilling, *Principles of Communication Systems,* 1986.

Heterogeneous catalysis

A chemical process in which the catalyst is present in a separate phase. In the usual case, the catalyst is a solid, and the reactants and product are in gaseous or liquid phases. *See* CATALYSIS.

Heterogeneous catalysis proceeds by the formation and subsequent reaction of chemisorbed complexes which can be considered to be surface chemical compounds. In the simple case where A → B is slow in the absence of catalyst, reaction (1) might

$$\left.\begin{array}{l} * + A \rightarrow A* \\ A* \rightarrow B* \\ B* \rightarrow B + * \end{array}\right\} \text{Chain propagating steps} \qquad (1)$$
$$A \rightarrow B \qquad \text{Overall reaction}$$

occur. might occur. Reaction A → B is fast if the three preceding steps are fast. Here, * represents a catalytic site on the surface of the catalyst, A* → B* is called a surface reaction, * + A → *A represents the chemisorption of A, and B* → B + * represents desorption of B. *See* ADSORPTION.

With most sets of reactants, more than one reaction will be thermodynamically possible. The degree to which a given catalyst favors one reaction compared with other possible reactions is called the selectivity of the catalyst for reaction (1). Two aspects of a catalyst are of particular importance: its selectivity and its activity, which can be taken as the rate of conversion of reactants by a given amount of catalyst under specified conditions. Ideally, the rate will be proportional to the amount of catalyst.

History. Heterogeneous catalytic processes were first recognized by Humphry Davy in 1817 (the oxidation of hydrocarbon gases on platinum) and by L. J. Thénard in 1818 (the decomposition of hydrogen peroxide on various solids). Industrial applications started about 50 years later with the Deacon process for oxidizing hydrogen chloride (HCl) to chlorine (Cl_2), reaction (2), with a copper chloride catalyst, and the contact process for making sulfuric acid (H_2SO_4), reactions (3), with a

$$O_2 + 4HCl \rightarrow 2Cl_2 + 2H_2O \qquad (2)$$

$$\begin{array}{l} 2SO_2 + O_2 \rightarrow 2SO_3 \\ SO_3 + H_2O \rightarrow H_2SO_4 \end{array} \qquad (3)$$

platinum catalyst. The original platinum catalyst in (3) has been replaced by vanadium pentoxide on silica gel.

The rate of development of industrial applications of heterogeneous catalysis accelerated after about 1900. Manufacture of nitric acid (HNO_3) by the oxidation of ammonia (NH_3), reaction (4), began in

$$NH_3 \xrightarrow{O_2} NO \xrightarrow{O_2} NO_2 \xrightarrow{H_2O,O_2} HNO_3 \qquad (4)$$

1906, although in 1838 C. F. Kuhlmann had discovered that platinum catalyzed the first step of the process. In modern usage, the catalyst is platinum-rhodium gauze at 900°C (1652°F).

By the 1950s, heterogeneous catalytic processes had come to dominate the petroleum, petrochemical, and chemical industries. Heterogeneous catalysis is a critical feature in energy conservation and interconversion, is a key feature in the production of synthetic fuels from coal and oil shale, and is the primary strategy for control of automobile air pollution. *See* ALCOHOL FUEL; COAL GASIFICATION; FISCHER-TROPSCH PROCESS.

Reactors. Both industrially and in the laboratory, catalytic reactions are usually effected in one of three kinds of reactors: batch, tube-flow, and gradientless flow. Examples are shown in the **illustration**. In the flow reactors, there is usually a large change in concentrations of products and reactants between the entrance and exit of the bed. In the gradientless reactor, the changes are kept very small, either by recirculation of 99% of the exit gases from the catalyst bed in *b* back to the entrance to the bed, or by use of a stirred-flow reactor, of which the fluidized reactor in *c* is one form.

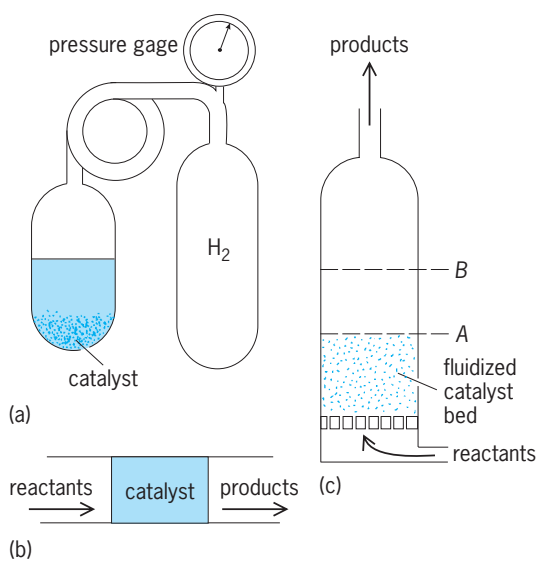

Catalytic reactors. (*a*) Batch. (*b*) Continuous fixed-bed. (*c*) Continuous fluidized reactor.

Some typical heterogeneous catalytic reactions

Catalyst	Reaction	T, °C*
	Hydrogenation	
Pt, Pd, Rh, Ni, as powders or supported on SiO_2, Al_2O_3, or C	$H_2 + C_2H_6 \rightarrow C_2H_6$	−100
	$H_2 + D_2 \rightarrow 2HD$	−180
	$C_3H_8 + D_2 \rightarrow C_3H_7D, C_3H_7D{-}C_3D_8 + HD$	50
	Cyclopropane $+ H_2 \rightarrow C_3H_8$	40
	$C_2H_6 + H_2 \rightarrow 2CH_4$	250
Cr_2O_3 activated to generate coordinatively unsaturated Cr^{3+}	$H_2 + C_2H_6 \rightarrow C_2H_6$	−100
	$H_2 + D_2 \rightarrow 2HD$	−180
	$D_2 + C_3H_8 \rightarrow C_3H_7D + HD$	200
Fe†	$3H_2 + N_2 \rightarrow 2NH_3$	400‡
Cu/ZnO	$2H_2 + CO \rightarrow CH_3OH$	350‡
Pt, Cr_2O_3	Heptane → toluene $+ 4H_2$	450
Pt, Cu	Acetone $+ H_2 \rightarrow$ 2-propanol	75
	Polymerization	
Cr^{2+}/SiO_2	$C_2H_4 \rightarrow$ linear polyethylene	50
	Olefin metathesis	
Mo^{4+}/Al_2O_3	$2C_3H_6 \rightarrow C_2H_4 + CH_3CH{=}CHCH_3$	50
	Oxidation	
Pt, many oxides of transition metals	$2H_2 + O_2 \rightarrow 2H_2O$	0−200
	$2CO + O_2 \rightarrow 2CO_2$	50−200
	$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$	200−350
Ag	$2C_2H_4 + O_2 \rightarrow$ ethylene oxide	200
Bismuth molybdate	$C_3H_6 + NH_3 + 3/2O_2 \rightarrow CH_2{=}CHCN + 3H_2O$	450
V_2O_5/SiO_2	Naphthalene $+ O_2 \rightarrow$ phthalic anhydride	350
Fe_3O_4	$H_2O + CO \rightarrow H_2 + CO_2$	450

*The lowest temperature at which significant yields are obtained in a flow reactor. °F = (°C × 1.8) + 32.
†In commercial practice, a few percent of a promoter such as potassium oxide plus aluminum oxide ($K_2O + Al_2O_3$) is added to the iron (Fe) to improve the performance and extend the life of the catalyst.
‡Because of the small value of the equilibrium constants at the operating temperatures, these reactions are run at about 100 atm (10 megapascals) in order to get adequate conversions.

Catalysts. A selection of heterogeneous catalytic processes of scientific or industrial interest is shown in the **table**.

Oxidation reactions of organic compounds may be divided into two categories, complete oxidation (or combustion) and partial (or selective) oxidation. The oxidation of methane (CH_4) shown in the table is an example of complete oxidation. One of the most widely applied uses of complete oxidation is the catalytic converter in motor vehicles. Here CH_4, other hydrocarbons, carbon monoxide (CO) in the engine exhaust, and oxygen are converted to carbon dioxide (CO_2) and water (H_2O) by passage over a supported platinum-palladium-rhodium catalyst.

Reaction (5) represents a particular catalytic partial oxidation reaction of commercial interest. The catalyst of choice is Ag/Al_2O_3 with various promoters, for example, Cs^+ and Cl^-. Reaction is carried out in a packed-bed flow reactor in the range of 200–300°C (392–572°F), and selectivites of the order of 80% are achieved.

$$2CH_2{=}CH_2 + O_2 \longrightarrow \underset{\text{Ethylene oxide}}{2CH_2{-}CH_2 \atop \diagdown\ \diagup \atop O} \qquad (5)$$

Since catalytic activity will ordinarily be proportional to surface area, most catalysts are used in forms with large specific areas, a_s. The low-area catalyst of reaction (4) is a rather unusual case. Higher-area metal powders are often used for liquid-phase reactions in batch reactors like that shown in *a*. For example, finely divided nickel, $a_s = 25–40\ m^2\ g^{-1}$, is used for the hydrogenation of unsaturated glycerides in the manufacture of margarine from vegetable oils. *See* HYDROGENATION.

Supported catalysts are widely used. In these, the catalytic ingredient is dispersed in the internal porosity of such supports as silica gel, γ-alumina, and charcoals. These supports have large areas in the internal porosity, 100–1000 $m^2\ g^{-1}$, and their average pore diameters are 2–20 nanometers. Tiny crystallites of such metals as platinum, palladium, rhodium, and nickel can be formed in the pore structure. Supported oxides and sulfides of transition metals are also used. Supported catalysts have the advantage that the area of the catalytic ingredient can be very large. Since, however, the support granules can have diameters in the 1–5 mm range, large flows of gas produce only moderate pressure drops across a catalyst bed. Supported catalysts are much more resistant to coalescence of the catalytic ingredient than are powders. Further, deposition of carbonaceous residues accompanies many catalytic reactions. In

some cases, the catalyst can be regenerated by burning off the residues. Such regeneration would result in drastic losses of area with metal powders.

Supported catalysts are particularly prone to problems with diffusion. Reaction $^* + A \rightarrow {}^*A$ must be preceded by the diffusion of the reactant through the pore structure of the support to a catalytic site. Reaction $B^* \rightarrow B + {}^*$ must be followed by the diffusion of the product out of the support. Heat must also flow in and out of the granule of support. Such matters receive particular attention in the chemical engineering aspects of catalysis. *See* DIFFUSION.

One important type of catalyst exposes strongly acidic sites in its internal porosity. Such catalysts are used to crack larger molecules of hydrocarbon into smaller ones in petroleum refining. Other catalysts, called dual-functional catalysts, have a hydrogenating catalytic ingredient on an acidic support. These are also of major importance in processing petroleum. The acidic function for catalytic cracking is commonly provided by a zeolite, a crystalline porous material with strongly acidic sites. *See* CRACKING; HYDROCRACKING; ZEOLITE.

Another type of catalyst consists of an organometallic complex deposited on such supports as silica or alumina. These catalysts have been called heterogenized homogeneous catalysts, and they accompany a development in which the nature of surface sites on heterogeneous catalysts has been interpreted in terms of coordination chemistry and homogeneous catalysis. *See* COORDINATION CHEMISTRY.

Most catalysts exhibit coordinatively unsaturated surface sites (*cus*) which are capable of reacting with molecules in the gas or liquid phases to form chemisorbed intermediates, as in reaction $^* + A \rightarrow {}^*A$. For example, the atoms at the surface of a crystallite of platinum must be coordinatively unsaturated. Considerable progress has been made in applying the studies of surface chemistry and physics on particular crystal surfaces—(111), (110), or (100)—of such metals as platinum to the interpretation of catalytic reactions and chemisorption on metals. *See* ELECTRON DIFFRACTION; SURFACE PHYSICS.

High catalytic activity for a given reaction requires adsorption of the reactants at *cus* sites to be of intermediate strength. If the adsorption is too weak, the reactants are unlikely to be activated; if too strong, A^* will be unreactive. With good catalysts, then, it is likely that there will be other compounds which adsorb more strongly than the reactants, with partial or complete blockage of the catalytic reaction. Such compounds are said to be poisons. For example, transition metals are poisoned by soft bases like R_2S, R_3P, and CO, where R represents a functional group. Poisoning is a problem in all types of catalysis.

Mechanism. The kinetics of heterogeneous catalytic reactions are often rather complicated. In general, rate = (amount of catalyst)$f(T, P_i$, or $C_i)$, where T is the temperature of the catalyst, P_i is the partial pressure of component i, and C_i is the concentration of component i. The index i covers reactants and products. A reaction having a rate-limiting process $A^* + B^* \rightarrow C^*$ will have a rate proportional to $\theta_A\theta_B$, (where θ_A is the fraction of surface sites represented by * which have been converted to A^* and θ_B) that converted to B^*. For $A^* + B^* \rightarrow C^*$ to be a rate-limiting process, A^* must be in equilibrium with A(g) and B^* with B(g), where g indicates the gaseous phase. Thus, θ_A and θ_B can be expressed in terms of P_A and P_B through the appropriate adsorption isotherms. If θ_A is small, it is apt to be proportional to P_A or $P_A^{0.5}$. If $\theta_A \approx 1$, it will be very nearly independent of P_A. *See* ADSORPTION.

Considerable work has been aimed at the determination of the chemical mechanism of particular catalytic reactions. Such work has depended upon kinetics, isotopic tracers, stereochemistry, and structure variation in reactants. A simple example is that which is commonly accepted for the hydrogenations of olefins, the Horiuti-Polanyi mechanism as illustrated for ethylene adsorption in reactions (6). It is

$$H_2C{=}CH_2 + 2{*} \longrightarrow H_2C{-}CH_2 \quad (\text{with } {*} \text{ and } {*})$$

or (6)

$$H_2C{=}CH_2 + {*} \longrightarrow H_2C{-}CH_2 \quad (\text{with } {*})$$

as yet unclear which formulation for the chemisorption of ethylene is most appropriate. The geometries of the two forms would be nearly the same. The adsorption of ethylene is accompanied by the dissociative chemisorption of hydrogen, reaction (7), followed by reactions (8) and (9).Much still remains to

$$H_2 + 2{*} \rightarrow 2H{*} \quad (7)$$

$${*}H_2C{-}CH_2{*} + H{*} \rightarrow {*}CH_2CH_3 + 2{*} \quad (8)$$

$${*}CH_2CH_3 + H{*} \rightarrow CH_3CH_3(g) + 2{*} \quad (9)$$

be learned about the details of catalytic mechanisms. *See* HOMOGENEOUS CATALYSIS.

Robert L. Burwell, Jr.; Gary L. Haller

Bibliography. J. R. Anderson and M. Boudart (eds.), *Catalysis: Science and Technology*, vols. 1–11, 1981–1996; J. A. Moulijin, P. Van Leeuwen, and R. A. Van Santen (eds.), *Catalysis: An Integrated Approach to Homogeneous, Heterogeneous, and Industrial Catalysis*, 1993; J. M. Thomas and W. J. Thomas, *Principles and Practice of Catalysis*, in G. Ertl, H. Knözinger, and J. Weitkamp (eds.), *Handbook of Heterogeneous Catalysis*, vols. 1–5, 1997.

Heteronemertini

An order of the class Anopla in the phylum Rhynchocoela, with an unarmed proboscis, a thick partly fibrous dermis, and a three-layered body musculature composed of outer longitudinal, median circular, and inner longitudinal strata. Cerebral organs, cephalic

grooves and slits, and eyes are generally present. The alimentary system consists of a mouth, foregut, intestine with regular lateral diverticula but no cecum, and anus. The principal family, Lineidae, contains some of the most common and best-known rhynchocoelan genera of temperate seashores, such as *Lineus* ("bootlace worms"), *Micrura*, and *Cerebratulus* ("ribbon worms"); the last one widely used in studies on regeneration, embryology, and nutrition. Heteronemertini are mainly marine littoral in habit. *See* ANOPLA; ENOPLA; PALAEONEMERTINI; NEMERTEA.

J. B. Jennings

Heterophile antigen

The serologic reactions of the tissue and blood-cell antigens of most animals are normally characteristic of the species. Significant serologic cross reactions usually occur only with antisera to the corresponding antigens of closely related species. The numerous groups of heterophile antigens—of which the Forssman antigens are the best studied—constitute significant exceptions. Heterophile antigens link the species hog-ox-human (blood group A), cat-horse, and dog-hog-cat-human, while several heterophile groups link otherwise diverse microorganisms. Links between pneumococcus type XIV and the human blood groups are also known, as well as similarities between antigens in mammalian hearts and the cell walls of the group A hemolytic streptococcus—a bacterium that is a common cause of rheumatic fever in humans. The cross reactions between the *Proteus* bacillus and the *Rickettsiae* are important in the diagnosis of typhus fever.

In 1911 J. Forssman reported that the injection of tissues from the guinea pig into rabbits stimulated the formation of antibodies which, together with complement, lysed sheep red blood cells. The organs of the horse, dog, cat, mouse, fowl, and tortoise also stimulate the production of antibodies with affinities for sheep cells, while the organs of humans, the rabbit, ox, rat, goose, eel, and frog do not. The Forssman antigen is also widely distributed among microorganisms, for example, pneumococci, anthrax, and dysentery bacilli. Cross reactions between the Forssman antigen in human tissues and human blood-group-A substance are said to be due to the sharing of a common disaccharide, α-*N*-acetyl-galactosaminoyl-(1$\rightarrow$3)-D-galactose. "Purified" Forssman antigen derived from pooled sheep red cells was found to contain as the principal ingredients galactose, galactosamine, lingoceric acid, and sphingosine. Some human sera react strongly with the erythrocytes of animals (sheep, horse, ox) that carry heterophile antigens. These reactions are commonly used for the diagnosis of infectious mononucleosis. *See* ANTHRAX; ANTIBODY; ANTIGEN; BACILLARY DYSENTERY; INFECTIOUS MONONUCLEOSIS; PNEUMOCOCCUS; RICKETTSIOSES; SEROLOGY.

Margaret J. Polley

Bibliography. R. M. Aloisi, *Principles of Immunology and Immunodiagnostics*, 1988; L. G. Schook and J. G. Tew, *Antigen Presenting Cells: Diversity, Differentiation, and Regulation*, 1987.

Heterosis

Hybrid vigor or increase in size, yield, and performance found in hybrids, especially if the parents have previously been inbred. The application of heterosis has been one of the most important contributions of genetics to scientific agriculture in providing hybrid corn, and vigorous, high-yielding hybrids in other plants and in livestock. *See* BREEDING (ANIMAL); BREEDING (PLANT); GENETICS.

It has been known for several centuries that hybrids between varieties and even between species are frequently of unusual size and vigor. The proverbial hardiness of the mule is often given as an example in animals. Charles Darwin devoted many years to a study of the effects of inbreeding and hybridization in several kinds of domestic plants. He reported, and his findings have since been abundantly confirmed, that inbreeding usually leads to decreased size and weakness, and to an increased frequency of abnormalities. The changes are cumulative so that with successive generations of inbreeding the deleterious effects are more and more pronounced. However, all the effects of inbreeding are immediately eliminated by outcrossing. Darwin also noted that many plants have elaborate mechanisms that prevent self-fertilization, thus avoiding the deleterious effects of close inbreeding.

A satisfactory explanation of heterosis awaited the development of mendelian genetics. It is now known that the effect of inbreeding in a population is to make homozygous many genes that had previously been heterozygous. Thus the decline in vigor and size must have its explanation in the increased homozygosity that accompanies inbreeding.

Likewise, heterosis may be explained by the reverse effect of increased heterozygosity following outcrossing, since a hybrid will in general be more heterozygous than its parents. *See* MENDELISM.

Hypotheses. There are two principal hypotheses to account for the association of size and vigor with heterozygosity.

Dominance. The dominance hypothesis notes that any noninbred population carries a number of recessive genes that are harmful to a greater or lesser extent, but which are rendered ineffective by their dominant alleles. As they become homozygous through inbreeding, they exert their harmful effect. With hybridization, some of the detrimental recessives contributed to the hybrid by one parent are masked by dominant alleles from the other, and an increase in vigor is the result. This hypothesis is supported by numerous experimental studies showing the large number of harmful recessives carried in actual populations, a number that is quantitatively adequate to account for the observed decline in vigor with inbreeding and recovery on crossing. *See* DOMINANCE.

Overdominance. The alternative hypothesis is that there are loci at which the heterozygote is superior in vigor to either homozygote. This, the overdominance hypothesis, also has the consequence that vigor is proportional to heterozygosity. It is argued that studies of individual genes have revealed only a very small number with this property, but on the other hand such genes would be expected to persist in a population and to make a disproportionate contribution to its variability. The dominance hypothesis has been more widely accepted, but the two are very difficult to distinguish experimentally, and it is likely that overdominant loci are playing an appreciable role in heterosis, particularly in determining why one hybrid is better than another.

In the fruit fly (*Drosophila melanogaster*) the decreased viability with inbreeding is about two-thirds caused by genes with drastic effects—most of them being lethal when homozygous—and about one-third by a much larger number of very mildly deleterious genes that act cumulatively. The evidence that these genes are numerous comes from two sources. One is the high rate of mutation; either each gene mutates very often or there are many of them, and the latter is much more likely. The second line of evidence is the discovery of a great amount of heterogeneity for proteins. Since these proteins are direct gene products, they reflect the genetic variability. *Drosophila* studies have shown that the average fly is heterozygous for about 12% of the gene loci. The values in vertebrates are smaller, being about 4-5%. Whether all this variability is maintained in the population by overdominance or by recurrent mutation is not yet clear. However, these are presumably the kind of genes that are responsible for one-third of the inbreeding effect, as well as for the differences between one hybrid and another.

Use. The most important example of the systematic use of heterosis has been the development of high-yielding hybrid corn. Almost all the corn grown in the north-central corn-producing states of the United States is now hybrid. About 1910 two American geneticists, E. M. East and G. H. Shull, discovered that by inbreeding corn, selecting the inbred strains, and then making hybrids between the selected inbred strains, they could obtain excellent plants. The hybrids not only had a high yield but also were extremely uniform, and by choosing the right inbred strains for hybridization, the breeder could incorporate into the hybrids other desirable traits such as disease resistance, straight stalks, and well-shaped ears.

The main difficulty in this scheme is that the seeds from which the hybrid plants develop are grown on ears of inbred plants. Hence the seed producer has a low yield, and the seed becomes expensive. This difficulty is circumvented by using the double cross, or four-way-cross, originated by D. F. Jones. The breeder starts with four inbred strains and makes two different crosses yielding two hybrids. The two hybrids are then crossed to produce the commercial double-cross seed. In this way the commercial seed is produced on high-yielding hybrid plants. The double-cross hybrids have about the same yield as the single-cross hybrids, though they are less uniform. It has become possible to produce higher-yielding inbred lines, so almost all hybrid corn is now single cross, with an improvement in uniformity.

Following the pattern set for hybrid corn, breeders have produced many other agricultural plants as hybrids. There has also been an increasing tendency to use heterosis in animal breeding, both by employing mating schemes that maximize heterosis and by hybridization between inbred lines, the latter being widely practiced especially in poultry breeding.

James F. Crow

Bibliography. G. W. Burns and P. J. Bottino, *The Science of Genetics*, 6th ed., 1988; J. F. Crow, The rise and fall of overdominance, *Plant Breed. Rev.*, 17:225-257, 2000; V. Loeschcke (ed.), *Genetic Constraints on Adaptive Evolution*, 1987.

Heterostraci

Extinct, armored, jawless vertebrates with a single pair of external gill openings. Heterostraci first appeared in the Ordovician, according to some authorities, and became common in the Silurian and Devonian. Their bony dermal armor consists mainly of aspidine (bone with apparent growth layers, lacking cavities for bone cells) and is composed of three structural layers: a basal layer of lamellar bone, a middle layer usually with a honeycomb arrangement of cavities and partitions, and a superficial layer of dentine tubercles or ridges. *See* JAWLESS VERTEBRATES; OSTRACODERM.

Morphology. The head region was usually enclosed in spoon-shaped upper and lower bony plates or

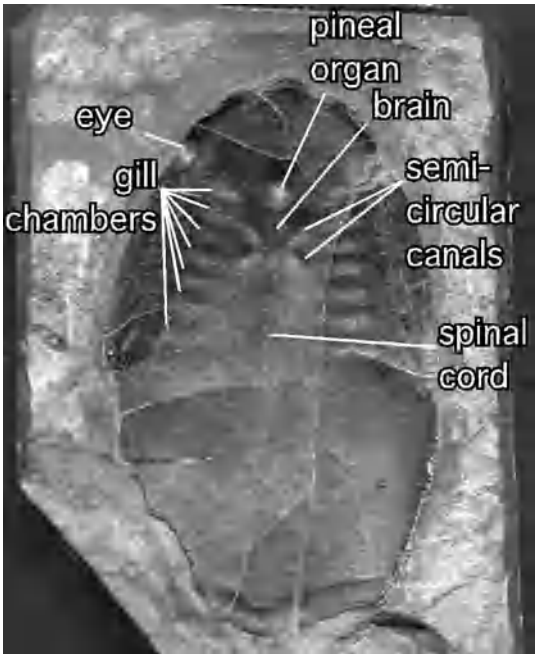

Fig. 1. Internal surface of the dorsal head shield of a cyathaspidiform heterostracan from northern Canada. Locations of organs are indicated by impressions in the surface of the shield. Length of shield, 3.5 cm (1.4 in.).

Fig. 2. ***Dinaspidella elizabethae*, a typical cyathaspiform heterostracan from the Early Devonian of northern Canada. Anterior is to the right; total length, 11 cm (4.3 in.).**

shields, joined to each other by smaller plates near the eyes, mouth, and gill chambers. The inner surface of the head armor often bore impressions of the brain, pineal organ, two pairs of semicircular canals, and the gill chambers (**Fig. 1**). The mouth lacked jaws, but small movable plates formed its rear margin. The eyes were widely separated on the sides of the head. The seven, paired, internal gill chambers emptied through a single pair of external gill openings. The body was covered in thick scales, each shaped like a diamond or parallelogram. The tail had a strong lower lobe into which the axis of the body continued, but an almost equally strong upper lobe was often present, with smaller scale-covered lobes between them (**Fig. 2**).

Phylogeny. Ordovician genera such as *Arandaspis* and *Astraspis* are often considered to be close relatives to or members of the Heterostraci, though they lack some of the group's unifying features. In the Silurian, primitive heterostracans such as *Athenaegis* and *Tolypelepis* occur. In Late Silurian and Early Devonian rocks, typical heterostracans in the Cyathaspidiformes (such as *Dinaspidella*, *Anglaspis*, *Poraspis*, *Ctenaspis*) are very diverse and among the most common fossil vertebrates, while other forms such as *Lepidaspis* and *Aserotaspis* are more poorly known. In the Late Silurian and during the Devonian, more advanced forms including Pteraspidiformes (such as *Errivaspis*, *Pteraspis*, *Phialaspis*) and Psammosteiformes (such as *Drepanaspis*, *Psammolepis*) diversified and became dominant. The Heterostraci declined in diversity and abundance toward the end of the Devonian and were extinct before the end of that period. *See* DEVONIAN; ORDOVICIAN; SILURIAN.

Fossil record. Heterostraci were common in North America and in Europe but are not known from China or the Southern Hemisphere. Their fossils are useful to geologists as indicators of the ages of the rocks in which they are found. Silurian representatives lived in marine habitats, while later forms occupied both marine and freshwater environments. The latest surviving Heterostraci in the Psammosteiformes were confined to fresh or brackish waters and reached body sizes up to 1.5 m (5 ft). *See* FOSSIL.

Mark V. H. Wilson

Bibliography. P. Janvier, *Early Vertebrates*, Clarendon Press, Oxford, 1996; J. A. Long, *The Rise of Fishes: 500 Million Years of Evolution*, Johns Hopkins University Press, Baltimore, 1995; J. A. Moy-Thomas and R. S. Miles, *Palaeozoic Fishes*, 2d ed., W. B. Saunders, Philadelphia, 1971.

Heterotardigrada

An order of the tardigrades, the majority of whose genera have widely varied structure. Cephalic appendages having a sensorial function are present, as well as cirrus lateralis and clava. Pharyngeal pockets are strengthened by uninterrupted ridges, except in *Pseudechiniscus islandicus*, which has interrupted ones. Toes or claws are uniform in structure and completely separated from one another. A preanal gonopore is present; excretory glands are absent. This order of tardigrades is divided into two suborders, Arthrotardigrada and Echiniscoidea.

Members of the suborder Arthrotardigrada have toelike terminations of the legs (**Fig. 1**). The tubular middle part of the leg telescopes into the broad proximal part. These animals are marine organisms found in sand or on algae. One species, *Tetrakentron*, is found on the buccal tentacles of a sea cucumber, *Leptosynapta*. Species of *Batillipes* were recorded from North Carolina and Florida.

In the suborder Echiniscoidea the legs terminate with claws (**Fig. 2**). The middle part of the leg is

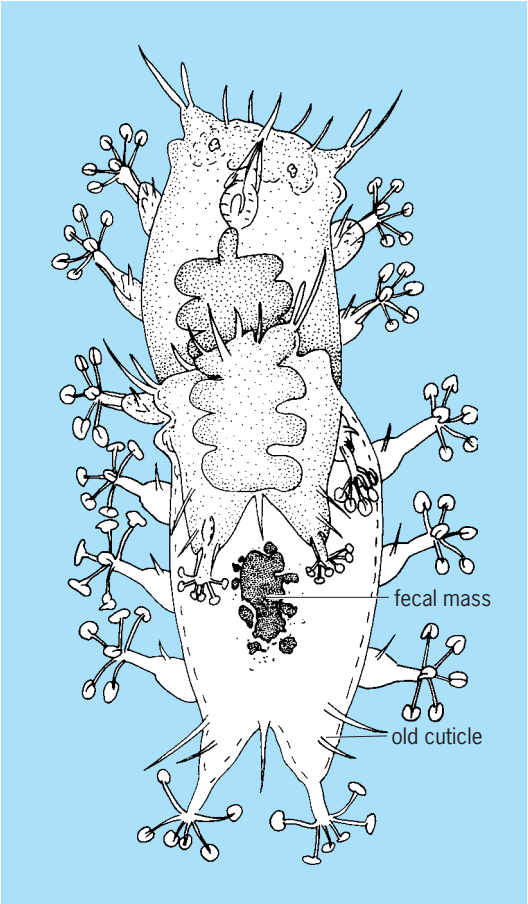

Fig. 1. ***Batillipes*; defecation during molt.**

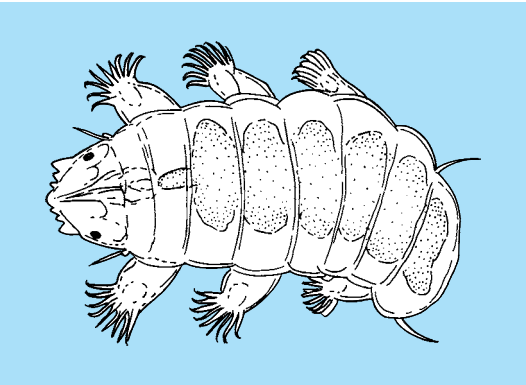

Fig. 2. ***Echiniscoides sigismundi*, a representative of the suborder Echiniscoidea.**

partially retractable into the proximal part. At least the fourth pair of legs has a distinct fold. There are two families: Nudechiniscidae, with a uniform cuticle; and Scutechiniscidae, with segmental and intersegmental thickenings (plates) of cuticle. Frequently these animals are red because of the presence of carotenoid pigments. Many species are active at relatively low humidity. *See* CAROTENOID; TARDIGRADA.

Eveline Marcus

Heterotrichida

A large order of the Spirotrichia containing many well-known, sizable species. The buccal ciliature, both membranes and the adoral zone of membranelles, is well developed, although in a number of families the somatic, or body, ciliature is really holotrichous in nature. Heterotrichs have become adapted to all sorts of habitats, including the digestive tracts of a variety of invertebrate and a few vertebrate hosts. Man is not involved. Some species contain pigments in their cytoplasm giving them such coloration as blue, green, pink, brown, or black.

Because of their size and amazing regenerative powers, a number of heterotrichs have been widely used in experimental research paralleling the studies by embryologists in the fields of morphogenesis and differentiation. *Stentor* (most popular in morphogenetic investigations), *Blepharisma*, and *Folliculina* are all genera with colored species. *Condylostoma*, *Climacostomum* (see **illus.**), and *Spirostomum*, which may reach a length of $^1/_8$ in. (3 mm), are common species. *Nyctotherus* is found in the digestive tract of amphibians and many invertebrates. *Balantidium*, long considered a heterotrich and important as the only ciliate parasitizing humans, is actually a trichostome. *See* CILIOPHORA; PROTOZOA; SPIROTRICHIA.

John O. Corliss

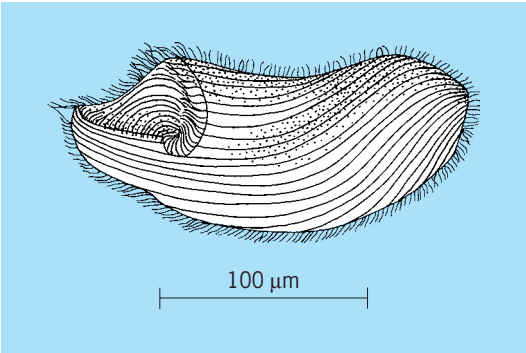

***Climacostomum*, an example of a heterotrich.**

Heulandite

A mineral belonging to the zeolite family of silicates and crystallizing in the monoclinic system. It usually occurs in crystals with prominent side pinacoid, often having a diamond shape. There is perfect side pinacoid cleavage on which the luster is pearly; elsewhere the luster is vitreous. The crystals often have undulating faces, and are made up of subindividuals in nearly parallel position (see **illus.**). In polarized light they show optical anomalies of a sectoral nature. The hardness is $3^1/_2$ to 4 on Mohs scale; specific gravity is 2.18–2.20. The mineral is usually white or colorless but may be yellow or red.

Heulandite crystal habit. (*After C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*)

Heulandite is essentially a hydrous calcium aluminum silicate, $Ca(Al_2Si_7O_{18}) \cdot 6H_2O$. Small amounts of sodium and potassium usually substitute for calcium. Heulandite is a secondary mineral found in cavities in basalts associated with other zeolites and calcite. Notable localities are in the Faeroe Islands, India, Nova Scotia, and West Paterson, New Jersey. *See* SILICATE MINERALS; ZEOLITE.

Clifford Frondel; Cornelius S. Hurlbut, Jr.

Hexacorallia

A subclass of cnidarian class Anthozoa that includes stony and black corals and sea anemones; also known as Zoantharia. With few exceptions, these animals are benthic as adults, living attached to firm substrata or burrowed into soft sediments; they occur at all latitudes, and from the high intertidal zone to the deepest parts of the oceans. As in all anthozoans, the adult form is a polyp—a cylindrical organism that has, at its free end, the single body opening, the mouth, which is surrounded by retractile, rarely branched, tubular tentacles **(Fig. 1)**. The tentacles number from six to several hundred, but typically the number is approximately a multiple of six, which gives the subclass its name, Hexacorallia. The opposite end of the polyp is attached to or burrowed into the substratum, or it emerges from the tissue that unites the members of the colony. *See* ANTHOZOA; SEA ANEMONE.

Fig. 1. Anthozoan anemone polyp. (*Photo by Eugene Weber, © California Academy of Sciences*)

Diversity. Orders Scleractinia (stony corals) and Zoanthidea (mat anemones) have some members that are colonial, some clonal, and some solitary; all members of Antipatharia (black corals) are colonial, and all members of Ceriantharia (tube anemones) are solitary; orders Actiniaria (sea anemones in the strict sense) and Corallimorpharia (mushroom anemones) contain some members that are solitary and some that are clonal (in at least one anemone species, clone mates may not separate completely, but whether this should be considered coloniality is debatable). Actiniaria and Scleractinia have the greatest number of species, each with somewhat more than a thousand; the other orders include a few tens to a hundred or so species. *See* ACTINIARIA; ANTIPATHARIA; CERIANTHARIA; CORALLIMORPHARIA; SCLERACTINIA; ZOANTHIDEA.

Previously, a seventh order of Hexacorallia was recognized, Ptychodactiaria. Comprising only three species, each in its own genus, this group is now considered a member of Actiniaria (where the two species known at the time had been placed until the middle of the twentieth century). The order Corallimorpharia, which is morphologically intermediate between Scleractinia (which it resembles in internal anatomy and the type of nematocysts it possesses) and Actiniaria (which it resembles in lacking a calcareous skeleton and lacking colonial members), may actually belong to Scleractinia. *See* PTYCHODACTIARIA.

Morphology. As in all anthozoans, the internal body space (the coelenteron, or gastrovascular cavity) is divided by longitudinally oriented sheets of tissue, the mesenteries **(Fig. 2)**. Along the free edge of most mesenteries is a filament that contains cilia, nematocysts, and gland cells; typically a longitudinal retractor muscle runs along one side of each mesentery. Commonly, two mesenteries—the directives—attach to each siphonoglyph, a histologically specialized channel running the length of the actinopharynx. A typical member of Actiniaria and Antipatharia has two siphonoglyphs diametrically opposite one another; a member of Cerianthiaria and Zoanthidea has a single one, conventionally considered to be on opposite sides of the animal in the two orders, the position in the former termed asulcal and in the latter, sulcal. As far as is known, scleractinians lack siphonoglyphs, which are typically absent or poorly developed in corallimorpharians.

In many hexacorals, the number of mesenteries increases as the diameter of the polyp does. An antipatharian polyp does not grow much after it is formed; in some species, a polyp never has more than the six initial mesenteries; in others, four or six additional mesenteries develop. Mesenteries are added only in certain regions and bilaterally—they form in couples, the halves arising simultaneously on each side of the animal. In Actiniaria, Corallimorpharia, Scleractinia, and Zoanthidea, two mesenteries (which constitute a pair) arise side by side simultaneously. In Zoanthidea, one member of most pairs is complete and one incomplete; in the other orders, the members of a pair are typically the same size. Mesenteries are added around the perimeter of an actiniarian, corallimorpharian, and scleractinian polyp, with the pairs forming diametrically opposite one another and constituting a couple. In a zoanthidean polyp as well, a couple constitutes two pairs of mesenteries, but addition is only on each side of the directive mesenteries that connect to the siphonoglyph. In a cerianthid polyp, mesenteries are also added in a single growth zone, but on the side of the polyp opposite the siphonoglyph; and a couple constitutes two single (not paired) mesenteries. Thus, the oldest mesenteries flank the siphonoglyph and are sequentially younger toward the opposite side of the animal, whereas the youngest mesenteries in a zoanthidean flank the siphonoglyph and are sequentially older toward the opposite side of the animal.

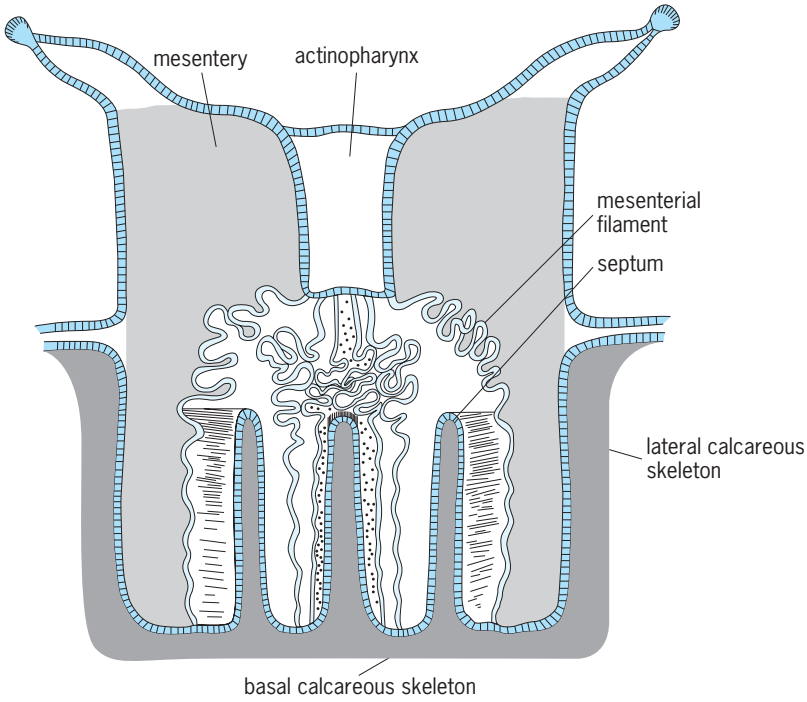

Fig. 2. Diagrammatic longitudinal section through a scleractinian polyp.

Characteristics. As is true of all cnidarians, hexacorals are carnivores. Many shallow-water species, particularly those in warm waters, have intracellular dinoflagellate symbionts (zooxanthellae) that contribute fixed carbon to the animal and remove wastes from it. Gametes form and mature in the mesenteries, between the filaments and the retractor muscles, which rupture to release them when they are ripe; some kinds of hexacorals are hermaphroditic and some are gonochoric (with separate sexes). Typically gametes are spawned freely into the sea, where egg and sperm meet and the planula (ciliated, free-swimming larva) develops; in some hexacorals, however, embryos are brooded internally or on the surface of the parent, so a free-swimming stage of the life cycle is lacking.

Polymorphism of the sort that is common in octocorals is virtually absent in hexacorals. In both clonal and colonial species, some individuals may have distinctive morphological attributes, but in contrast to octocorals these differences appear to be situational, so an individual can develop or lose the attributes. For example, in some species of scleractinians, the polyps at the edge of a colony that encounters cnidarians of a different species can develop elongate tentacles with nematocysts specialized for aggression. In actiniarians of a few clonal species, individuals at the periphery of a clone have highly developed acrorhagi (marginal bulges with dense nematocysts that are used in aggressive encounters) and form few or no gametes, whereas individuals at the core of the clone have small acrorhagi and their gametogenic tissue is richly developed.

Reproduction and propagation. Sexual reproduction typically results in a planula larva, which eventually metamorphoses into a polyp. In clonal and colonial species, asexual (vegetative) reproduction occurs in several ways, typically in only one way per species. Budding involves an outgrowth of a new polyp from an existing one or from the tissue connecting members of the colony; other hexacorals divide into two, either longitudinally or transversely; some sea anemones regenerate from detached tentacles or from bits of tissue that tear from the edge of the base; and apparently some anemones and corals can develop from bits of tissue shed from the inside of the animal, although the precise mechanism for this mode of propagation is unclear.

Skeletal features. A skeleton is absent in Actiniaria, Ceriantharia, and Corallimorpharia; zoanthideans do not form skeletons either, but in many taxa a polyp can incorporate into the mesoglea of its column particles such as sponge spicules or sand grains to form a sort of adventitious skeleton. In life, an antipatharian skeleton is hidden under the thin layer of polyps that secrete it; after their death, the polyps rot away, leaving the bushy or whiplike black skeleton that gives the taxon its common name, black corals. Composed of organic material so it is somewhat flexible, it has traditionally been formed into bracelets thought to ward off illness (hence the name Antipatharia) and is made into other jewelry as well. All scleractinian polyps secrete around themselves a cuplike calcareous skeleton in the crystal form aragonite; in some taxa the animals are solitary; in some the colonial skeleton is massive, with the "cups" of adjacent polyps sharing walls. Each polyp also deposits radial calcareous septa within the cup in the space between members of each larger mesenterial pair (Fig. 2). This skeleton is the major structural component of tropical and subtropical coral reefs and of the structures known as deep-sea coral reefs. *See* REEF.

Fossil record. Several orders of Hexacorallia have excellent fossil records, but fossils of groups that are exclusively soft-bodied are rare and almost never show taxonomically informative features. The earliest-known fossil corals are of Middle Cambrian age, but their sporadic occurrence in the Cambrian and earliest Ordovician suggests that early hexacorals lacked skeletons and that skeletalization evolved several times, possibly not in the same lineage each time. Members of the two major orders of Paleozoic corals first appeared in the Ordovician—Tabulata in the Early Ordovician and Rugosa in the Middle Ordovician. Middle Cambrian corals seem to represent two or more "experiments" in skeletalization, and Tabulata a later one. Rugosa seems to be polyphyletic: it can be logically derived from Tabulata at least twice. A minor order commonly thought to have separated from Tabulata in the Middle Ordovician became common and even a reef-builder in the Silurian and Devonian but was extinct before the end of the Middle Devonian. Another minor order separated from the Rugosa during the Late Devonian but became extinct during the early Carboniferous.

Lineage. Whether the lineages of these corals survived beyond the Paleozoic is debated. Zoanthideans bear considerable anatomical resemblance to rugosans. It has also been suggested that Scleractinia descended from Rugosa, possibly polyphyletically, with various scleractinian groups arising from different rugosan families. However, because of differences in ontogeny (mode of septal insertion), skeletal composition (the calcareous rugosan skeleton has the crystal form calcite), and the time gap (there are no known Early Triassic corals), it is more likely that Scleractinia evolved from one or more groups of early Mesozoic soft-bodied hexacorals. These ancestral animals are often referred to as sea anemones, but there is no independent evidence that order Actiniaria is so ancient. Because of similarities between scleractinians and actiniarians, it is also possible that sea anemones are derived from a lineage of ancestral scleractinians through loss of the calcareous skeleton and other morphological modifications; preliminary molecular evidence favors this direction of evolution. *See* RUGOSA.

Further complicating hexacorallian phylogeny is the strong evidence that Scleractinia is polyphyletic. All hexacorals with paired, coupled mesenteries and a calcareous skeleton are considered scleractinians, but details of skeletal structure and of animal anatomy and biology vary greatly, and the groups

with similar skeletons do not necessarily share other attributes, so it may be that Scleractinia is a grade, not a clade. Moreover, the skeleton is inferred to have been lost and regained in some lineages. The greatest difference between Scleractinia and Corallimorpharia is that a polyp of the former has a calcareous skeleton whereas one of the latter is askeletal; the relative ephemerality of the scleractinian skeleton has led to the suggestion that corallimorpharians are simply naked scleractinians. Daphne G. Fautin

Bibliography. D. F. Dunn, Cnidaria, pp. 669–706 in S. P. Parker (editor-in-chief), *Synopsis and Classification of Living Organisms*, vol. 1, McGraw-Hill, New York, 1982; H. Erhardt and D. Knop, Order Zoantharia, pp. 269–278 in *Corals: Indo-Pacific Field Guide*, Ikan, Frankfurt, 2005; D. G. Fautin and G. R. Allen, *Field Guide to Anemonefishes and Their Host Sea Anemones*, Western Australian Museum, 1992; L. H. Hyman, Class Anthozoa: Subclass Zoantharia, pp. 566–632 in *The Invertebrates: Protozoa through Ctenophora*, vol. 1, McGraw-Hill, New York, 1940.

Hexactinellida

A class of sponges whose skeletons are made of siliceous hexactine spicules. These exclusively marine sponges are widely distributed in modern oceans. Their fossil record extends from the late Precambrian to the Recent. The basic spicule type of the class is a triaxial hexactine, in which the three pairs of opposed rays are at right angles to each other and lie along one of the three axes of a cube. Proximal ray ends and axial filaments meet at the center of the cube. These principal spicules and variants of that form make up skeletons of the sponges.

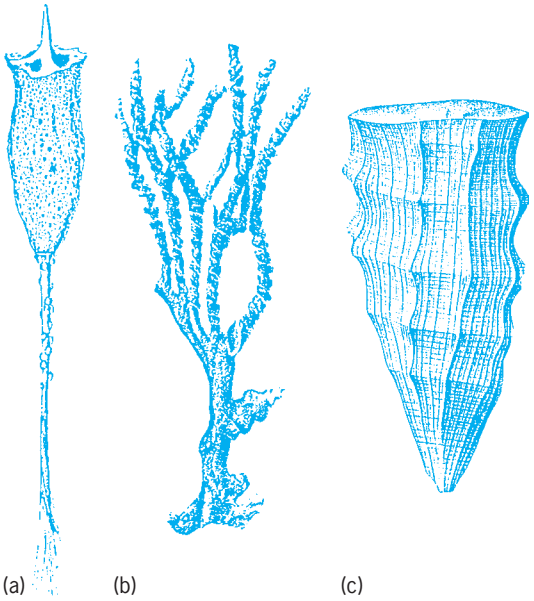

Fig. 1. Living hexactinellids (*a*) *Hyalonema* (*Cyliconema*) *thomsoni* Marshall, order Amphidiscosida (*after F. E. Schulze, 1887*) and (*b*) *Regadrella phoenix* Schmidt, order Lyssacinosida (*after E. A. Minchin, 1900*). Fossil hexactinellid (*c*) *Hydnoceras tuberosum* Conrad, order Reticulosa, Devonian (*after J. Hall and J. M. Clarke, 1900*).

Recent hexactinellid sponges are chiefly upper bathyal marine animals and are most common in depths of 200–2000 m (660–6560 ft), although many species are known to inhabit lower bathyal depths. Only a few species are known to range into hadal depths, below 6000 m (19,685 ft), but there they may form dense aggregations. Confirmed records of the shallowest living hexactinellids are at 25–30 m (80–100 ft) in the Pacific Ocean near Victoria, British Columbia. The most diverse bathyal hexactinellid faunas are those in Antarctic seas, Indo-west Pacific, and North Pacific regions.

Morphology and physiology. Living hexactinellid sponges are commonly goblet- or vase-shaped, although branched, massive, tubular, or ropy-appearing sponges also occur in the class (**Fig. 1**). Many have root tufts of long spicules that anchor them in place and support them above the sea floor.

The body wall of hexactinellid sponges is relatively open and cavernous. Soft parts are suspended on the mineralized skeleton, and consist of inner and outer layers of syncytial filaments or trabeculae and an internal layer of variously spaced, thimble-shaped flagellated chambers, which are suspended in the reticulate networks of the other layers (**Fig. 2**). These syncytial filaments are composed of protoplasm that contains scattered nuclei, but they are not subdivided into distinct cells. The trabecular network covers all surfaces that come into contact with water flowing through a sponge, and may be a continuous cytoplasmic network throughout a sponge.

Trabeculae have no regular direction in the inner and outer networks, but form bounding dermal and gastral membranes on the external and internal margins of the wall. The dermal membrane is perforated by pores, which allow water to enter into and circulate through the sponge. Sponges may have a flagellate chamber system only or may possess a canal system for water circulation. Small-diameter inhalant canals may extend in from the dermal pores to flagellate chambers, and moderately distinct, larger-diameter exhalant canals may lead from those chambers in the gastral part of the wall. Those canals commonly allow water to exit into a spongocoel or cloaca and out one or more large exhalant oscula. Such canals are seen as gaps in the trabecular structure or as interruptions in the mineralized skeleton.

Skeleton and spicules. Spicules of the skeleton are composed of hydrated amorphous silica, close to opal in composition, but may include some organic matter as well. The silica is deposited around an organic axial filament, which has a square cross section (**Fig. 3***a*). Such filaments meet at the center of the spicule and form an axial cross. The silica may show a layered structure, presumably a result of discontinuous secretion. Spicules may be formed in all parts of the trabecular network by multinucleate scleroblast syncytia, at least where origins are known.

Spicules are separable into two types: larger main elements termed megascleres, and usually smaller

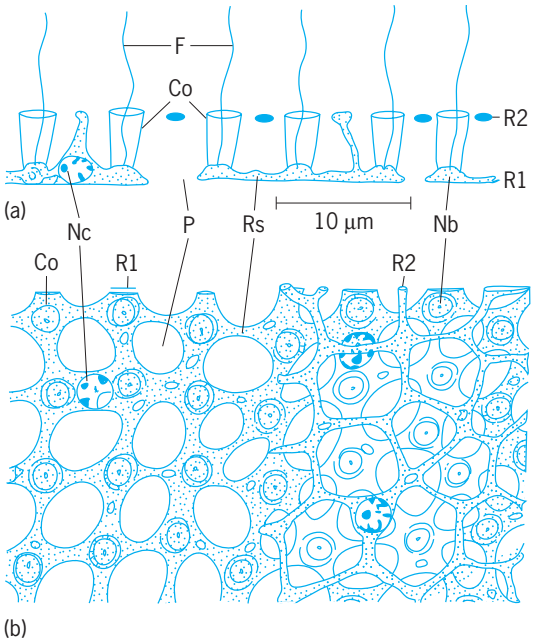

Fig. 2. Diagrammatic representation of sections through the choanosyncytium of *Aphrocallistes vastus* (*a*) in vertical section and (*b*) in tangential section from the interior of the chamber; the secondary reticulum is omitted for clarity. Co, collar; F, flagellum; Nb, nodal body; Nc, choanosyncytium nucleus; P, prosopyle; RS, connecting strands of main reticulum; R1, main reticulum of chamber wall; R2, secondary reticulum of chamber wall. (*Adapted from H. M. Reiswig, 1979*)

secondary elements termed microscleres (Figs. 3 and 4). Many variants of the basic triaxial hexactine megasclere (Fig. 3*c*) are known. For example, pentactines (pentacts) may result where one ray is reduced or lost (Fig. 3*d*), stauractines (stauracts) where two opposed rays are lost (Fig. 3*e*), or rhabdodiactines where all but two opposed rays are lost (Fig. 3*f*). Some such spicules still show a rudimentary axial cross, even in fossil examples, so their relationships to the basic orthotriaxial hexactine are clearly demonstrated. Other modifications may take place as well, such as in spicules where one ray becomes greatly elongated or enlarged, or where some rays are bent near their origins and curve away from the cubic axes. Spicules also may be modified by addition of solid spines that lack axial filaments (Fig. 3*b*, *g*, *i*). Such spines, for example, may be laterally directed on the distal ray of a pentactine to produce a conifer-appearing pinnule, or they may be thornlike to produce an uncinate of a rhabdodiactine.

Megascleres may be united in several ways to form rigid skeletal frameworks in some Hexactinellida (Fig. 3*j*–*l*). In the simplest skeletal framework, described as lyssacine, adjacent megascleres may be fused where rays cross one another or lie side by side. Commonly these rays are still recognizable as distinct elements. Where rays are close to one another but not in contact, they may be fused by siliceous rodlike synapticulae that bridge between them. Branched or interbraided siliceous filaments also may bridge between rays of associated spicules and fuse them together. Lyssacine frameworks may have loose, irregularly shaped meshes. In a dictyonine framework, overlapping rays or rays that meet tip to tip may be enveloped in a siliceous coating to form beams, in which individual rays are not identifiable except by their separated parallel axial filaments. This framework is usually a regular three-dimensional rectangular structure. The more complex lychniscose framework appears similar to a regular dictyonine framework, but has 12 small diagonal bracing struts developed at each axial megasclere node (Fig. 3*l*) to form an even more rigid skeletal framework.

Paleozoic hexactinellid sponges commonly have unfused megasclere spicules, but many still have their quadrate graph-paper-like skeletal structure preserved. Such fossilization may have been a result of their burial in relatively quiet environments, or it may suggest that their skeletons were held together by nonmineralized trabecular material or perhaps by unpreserved spongin. The skeleton of regularly arranged and ranked spicules like that in the Cambrian *Protospongia* is characteristic of early Paleozoic hexactinellids. Bundled unfused hexactines and pentactines, along with rhabdodiactines, produced the rectangularly three-dimensional skeletons of middle and late Paleozoic dictyosponges

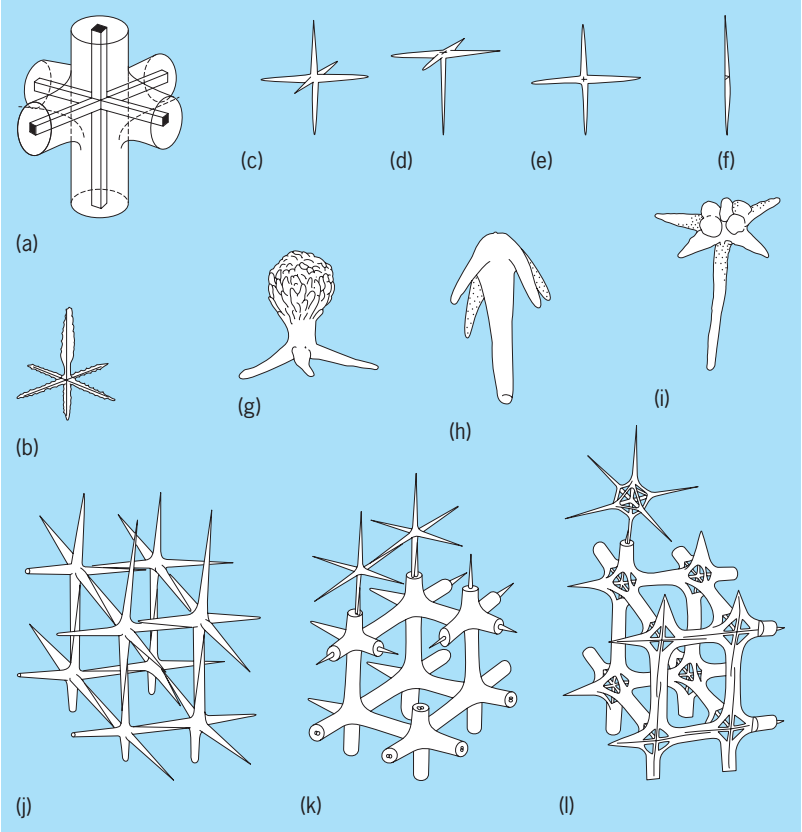

Fig. 3. Kinds of megasclere spicules and skeletal nets in the Hexactinellida. (*a*, *b*) Pattern of junction of coring axial filaments, with square cross sections, and spicule rays in pinnulated hexactine of *Sympagella nux* (*after H. M. Reiswig, 1971*). (*c*) Hexactine. (*d*) Pentactine. (*e*) Stauractine. (*f*) Rhabdodiactine. (*g*) Pinnular pentactine. (*h*) Anchorate root tuft pentactine. (*i*) Hexactine with nodes. (*j*) Lyssacinosid skeletal net with unfused hexactines. (*k*) Hexactinosid or dictyonine skeletal net with overlapping rays fused in siliceous coating and with simple spicule centers. (*l*) Lychniscosid skeletal net with fused rays and bracing struts at spicule centers. (*Parts j, k, l from J. K. Rigby, 1987*)

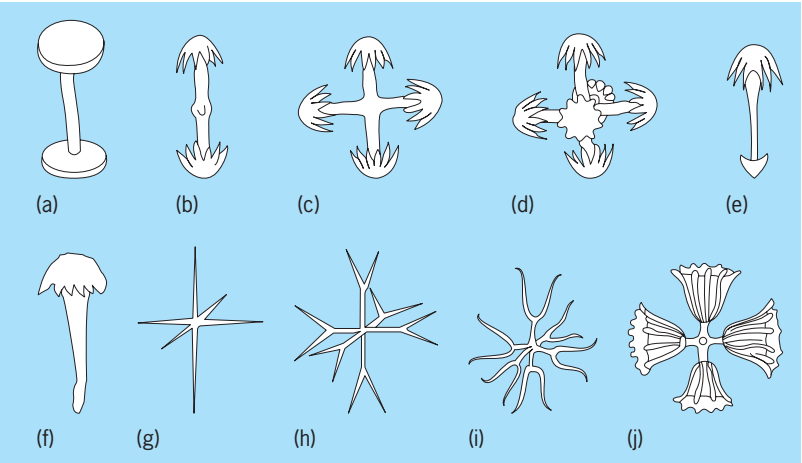

Fig. 4. Representative microsclere spicules found in the Hexactinellida. (*a–e*) Birotulates have umbrellalike terminations; (*f*) paraclavules have only one such termination; (*h*, *i*) hexasters are small hexactinal spicules. (*a*) Birotulate. (*b*) Amphidisc. (*c*) Staurodisc. (*d*) Hexadisc. (*e*) Hemidisc. (*f*) Paraclavule. (*g*) Microhexactine. (*h*) Hexaster. (*i*) Hexaster with curved terminations. (*j*) Floricome, with rays toward and away from the viewer omitted.

like *Hydnoceras* (Fig. 1*c*). Mesozoic and Cenozoic hexactinellid sponges, like Recent ones of the class, commonly have fused rigid skeletons.

Microsclere spicules are small (**Fig. 4**), commonly occur free within soft parts of the sponges, and are not united as part of the solid skeletal structure of the sponge. Microscleres typically range 10–100 micrometers long and 1 μm or less in diameter, only one-tenth the size of most megascleres. They may be small versions of megascleres, or they may be of two other basic types, hexasters or birotulates. Birotulates (Fig. 4*a–e*) are the small spicules with umbrellalike terminal expansions. Hexasters (Fig. 4*g–j*) are small hexactinal spicules that usually have branched ends, which may appear delicately flowerlike. Because microscleres are small and occur loose in the sponge, they are not as commonly preserved in fossils as are the megascleres. Taxonomically, microscleres are very useful spicules. The subclass Amphidiscophora is distinguished by having microscleres that are birotulates and never hexasters; whereas the subclass Hexasterophora has microscleres that are the small six-rayed hexasters, and lacks birotulates. Within the Amphidiscophora, sponges in the order Amphidiscosa have microscleres that are amphidiscs, and in some taxa also staurodiscs or hexadiscs, whereas sponges in the order Hemidiscosa have microscleres that are hemidiscs (Fig. 4*e*). Orders within the subclass Hexasterophora are differentiated on the basis of the firm rigid skeletal frameworks produced by the megascleres.

Taxonomy. The following classification is a combination of ones used in living and fossil sponges.

Class Hexactinellida
- Subclass Hexasterophora
  - Order: Lyssacinosida
    - Hexactinosida
    - Lychniscosida
- Subclass Amphidiscophora
  - Order: Reticulosa
    - Amphidiscosa
    - Hemidiscosa

The class Hexactinellida are exclusively marine sponges whose skeletons are made of siliceous hexactinal megasclere spicules with axial fibers or canals that have square cross sections, and microscleres that include hexaster or amphidisc microscleres; Precambrian–Recent.

The subclass Hexasterophora are hexactinellids in which microscleres are hexasters, and megascleres may be free and unconnected or, more commonly, fused to form a rigid skeleton; Ordovician–Recent. *See* HEXASTEROPHORA.

The order Lyssacinosida are hexasterophorans in which the megascleres are usually distinct and free, although some may be fused by supplementary siliceous elements; Ordovician-Recent.

The order Hexactinosida are hexasterophorans in which rows of hexactinal megascleres are fused by a supplementary coating of silica to form a rigid reticulate skeleton; Triassic-Recent. *See* HEXACTINOSA.

The order Lychniscosida are hexasterophorans in which the rigid framework of fused megascleres has each spicule node, termed a lychnisc, braced by 12 struts; Triassic-Recent.

The subclass Amphidiscophora are hexactinellids in which microscleres are birotules but never hexasters; megascleres are free and not fused to form a rigid skeleton; Precambrian-Recent.

The order Reticulosa are amphidiscophorans in which a dermal skeleton of parallel stauractines, pentactines, or hexactines forms a major part of the skeleton, and in which the microscleres include paraclavules; Precambrian–Permian. *See* RETICULOSA.

The order Amphidiscosa are amphidiscophorans in which the microscleres are equal-ended amphidiscs and not paraclavules; Ordovician-Recent. *See* AMPHIDISCOSA.

The order Hemidiscosa are amphidiscophorans whose principal microscleres are hemidiscs; Upper Pennsylvanian, Cretaceous-Recent. *See* HEMIDISCOSA.

There is general agreement on placement of most major taxa. However, the order Reticulosa is included in the subclass Amphidiscophora by paleontologists, based on the mode of occurrence of paraclavule microscleres (Fig. 4*f*) that indicates they functioned like amphidiscs in several Mississippian fossil sponges. The Reticulosa are considered by paleontologists to be an exclusively Paleozoic group of extinct sponges. However, some zoologists include the peculiar, branched *Sclerothamnus* as the only living and post-Paleozoic representative of the Reticulosa, and include the order in the subclass Hexasterophora.

Fossil record. Fossil hexactinellid sponges are known as impressions in the Precambrian Ediacaran beds of Australia and as isolated spicules from approximately equivalent beds in the Yangtze Gorge of China; but the earliest known body fossils with

clearly preserved intact spicule nets are from lowermost Cambrian rocks in Hunan Province, China. The best example of an early sponge fauna is that of the Middle Cambrian Burgess Shale of western Canada, which includes not only well-preserved hexactinellid sponges but other types as well. *See* BURGESS SHALE.

Ordovician hexactinellid sponges document the beginning of thick-walled dictyosponges, which became major elements in Devonian and Mississippian faunas of North America when one major lineage developed a three-dimensional bundled skeletal structure. Essentially contemporary hexactinellids with disordered skeletons also developed but were not as diverse or as abundant as the dictyosponges. Both lineages persisted until the end of the Paleozoic.

The Permian-Triassic boundary marks a major break in the history of many groups of fossils, including the hexactinellid sponges. Reticulose amphidiscophoran sponges became extinct at the end of the Permian and, of the subclass, only the minor hemidiscosan sponges continued to the Recent. Hexasterophoran sponges, however, became abundant and diverse during the Mesozoic, although those with lyssacine skeletons have left only a poor record in Jurassic and younger rocks. However, dictyonine hexactinosid sponges, which have a fused reticulate skeleton, produced an abundant and diverse Jurassic and Cretaceous record, particularly in Europe, and are represented in modern seas by several surviving families. Lychniscosid hexasterophoran sponges also experienced great diversity and abundance in the Jurassic and Cretaceous of Europe, but only two of those families survived into the Cenozoic, and only one is known in modern seas. *See* PORIFERA. J. Keith Rigby

Bibliography. P. R. Bergquist, *Sponges*, University of California Press, Berkeley, 1978; J. Hall and J. M. Clarke, *A Memoir of the Paleozoic Reticulate Sponges Constituting the Family Dictyospongidae*, N.Y. State Mus. Mem. 2, 1888, 1900; J. N. A. Hooper and R. W. M. Van Soest (eds.), *Systema Porifera, A Guide to the Classification of Sponges*, vol. 2, Kluwer Academic/Plenum, New York, 2002; R. L. Kaesler (ed.), *Treatise on Invertebrate Paleontology*, pt. E (rev.), vols. 2 and 3, 2003, 2004; E. A. Minchin, Sponges: Phylum Porifera, pp. 1–78 in E. R. Lancaster (ed.), *A Treatise on Zoology*, pt. 2, chap. 3: The Porifera and Coelenterata, Adam and Charles Black, London, 1900; H. M. Reiswig, The axial symmetry of sponge spicules and its phylogenetic significance, *Cah. Biol. Mar.*, 12:505–514, 1971; H. M. Reiswig, Histology of Hexactinellida (Porifera), pp. 173–180 in C. Lévi and N. Boury-Esnault (eds.), *Biologie des Spongiaires—Sponge Biology*, Colloques Internationaux du Centre National de la Recherche Scientifique, 2901, Centre National de la Recherche Scientifique, Paris, 1979; J. K. Rigby, Phylum Porifera, pp. 116–139 in R. S. Boardman, H. Cheetham, and A. J. Rowell (eds.), *Fossil Invertebrates*, Blackwell Scientific, Palo Alto, CA, 1987; F. E. Schulze, Report on the Hexactinellida collected by H.M.S. "Challenger," *Challenger Report: Zoology*, 1887.

Hexactinosa

An order of sponges of the subclass Hexasterophora in the class Hexactinellida. The parenchymal megascleres in this order are united to form a rigid framework and consist wholly of simple hexactins which are arranged in parallel linear series. The members of each series are united one to another by a secondary envelope of silica. Examples include *Hexactinella*, *Aphrocallistes*, *Eurete*, and *Farrea*. *See* HEXACTINELLIDA; HEXASTEROPHORA. Willard D. Hartman

Hexanchiformes

An order of Squalomorpha that are known as the six-gill sharks. These sharks are distinguished by a combination of the following characters: six or seven gill slits, all anterior to the pectoral fins; a single dorsal fin without a spine; anal fin present (the only order of squalomorphs with an anal fin); the dorsal, anal, and pelvic fins placed far posteriorly; eyes lacking a nictitating fold or membrane (at the inner angle of the eye or below the eyelid that is capable of extending over the eyeball); spiracles present, but small and well behind the eyes; and ovoviviparous development. They comprise two extant families (chlamydoselachidae and Hexanchidae), four genera, and five species. *See* ELASMOBRANCHII; SELACHII.

Chlamydoselachidae (frill sharks). This family, comprising only one species (*Chlamydoselachus anguindus*), is characterized by six gill openings, with the margin of the first gill opening continuous across the throat; very elongate body; subcylindrical trunk; terminal mouth; and teeth with three elongate cusps that are similar in the upper and lower jaws. Maximum length is about 196 cm (77 in.) [**illus.** *a*].

The frill shark occurs on continental and insular slopes [usually at depths between 120 and 1280 m (393 and 4200 ft)] of the western North Atlantic, eastern Atlantic from Norway to South Africa, southwestern Indian, western Pacific, and eastern Pacific from California to Chile. It feeds on squids, bony fishes, and other sharks.

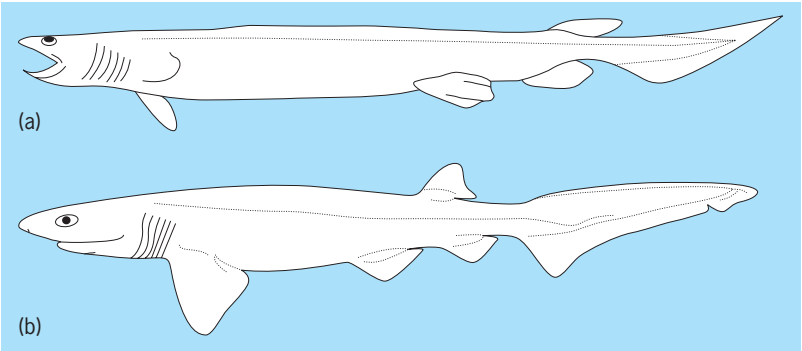

Examples of Hexanchiformes. (*a*) Frill shark. (*b*) Cow shark. (*J. S. Nelson, Fishes of the World, 4th ed., Wiley, New York, 2006*)

Hexanchidae (cow sharks). The cow sharks, comprising three genera and four species, are identified by six gill openings in the genus *Hexanchus*; seven gill openings in the genera *Heptranchias* and *Notorynchus* (each sometimes placed in its own family, Heptranchiidae and Notorynchidae, respectively); the margin of the first gill opening not continuous across the throat; inferior mouth (overhung by a long snout); upper and lower teeth different. Maximum length is about 470 cm (15 ft) [illus. *b*].

These sharks are circumglobal in temperate and tropical waters of continental and insular shelves and slopes. They feed on relatively large marine organisms, including other sharks, rays, bony fishes, crustaceans, and carrion. Herbert Boschung

Bibliography. M. R. de Carvalho, Higher-level elasmobranch phylogeny, basal squaleans, and paraphyly, pp. 35–62 in M. L. J. Stiassny, L. R. Parenti, and G. D. Johnson (eds.), *Interrelationships of Fishes*, Academic Press, San Diego, 1996; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Hexasterophora

A subclass of sponges of the class Hexactinellida, in which the parenchymal microscleres are typically hexasters (small, six-rayed spicules, often with branched ends). This is a diverse assemblage of sponges commonly firmly fixed to the substratum by the base, less commonly anchored by means of basal spicule tufts or mats. The spicules of the body are sometimes free and unconnected, but the parenchymal megascleres are often fused to form a rigidly connected skeleton (see **illus.**) The following orders are recognized: Hexactinosa, Lychniscosa, Lyssacinosa, and Reticulosa. *See* HEXACTINELLIDA; HEXACTINOSA; RETICULOSA. Willard D. Hartman

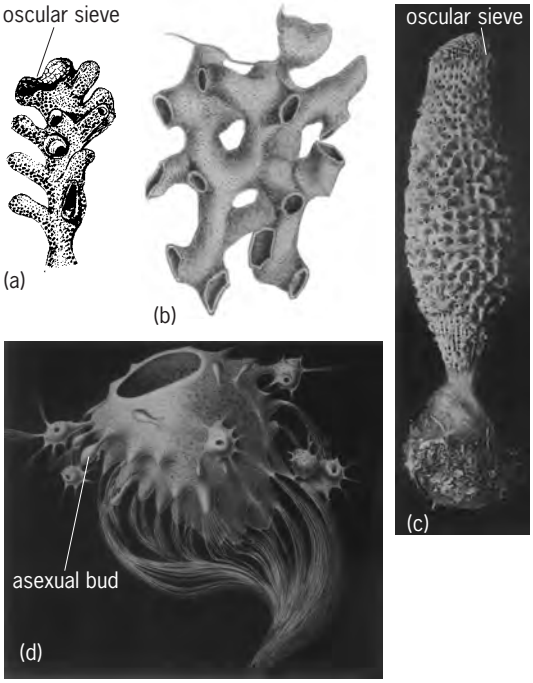

Representative hexasterophorans. (*a*) *Aphrocallistes*. (*b*) *Eurete*. (*c*) *Polylophus*, showing asexual buds. (*d*) *Euplectella oweni*.

Hibernation and estivation

A general term applied to a condition of dormancy and torpor found in cold-blooded (poikilotherm) vertebrates and invertebrates. This rather universal phenomenon can be readily seen when body temperatures of poikilotherm animals drop in a parallel relation to ambient environmental temperatures. In the strict sense of the word, hibernation is a term which physiologists apply to relatively few species of mammals and birds (warm-blooded vertebrates).

Poikilotherm Animals

Hibernation occurs with exposure to low temperatures and, under normal conditions, occurs principally during winter seasons when there are lengthy periods of low environmental temperatures. Animals that hibernate naturally can be induced to do so under controlled laboratory conditions; consequently, hibernating animals are becoming useful as experimental subjects. The potential use of hibernators and animals in similary depressed metabolic states has attracted the attention of scientists in a variety of biological disciplines, including space biology and medical research. It has been found that hibernating animals resist lethal levels of radiation, that they show resistance to infection and parasitism, and that antibody and immune responses are also greatly altered. These examples represent only a few areas of active modern research utilizing hibernators.

A related form of dormancy is known as estivation. Many animals estivate when they are exposed to prolonged periods of drought or during hot, dry summers. For all practical purposes, hibernation and estivation in animals are indistinguishable, except for the nature of the stimulus, which is either cold or an arid environment.

There is no complete list of animals that hibernate; however, many examples can be found among the poikilotherms, both vertebrate and invertebrate. The poikilotherms are sometimes referred to as ectothermic, because their body temperatures are not internally regulated but follow the rise and fall of environmental temperatures. During hibernation and winter torpor, body temperatures reflect the environmental temperature, often to within a fraction of a degree. Among the classic examples of hibernators or estivators are reptiles, amphibians, and fishes among the vertebrates, and insects, mollusks, and many other invertebrates.

Reptiles and amphibians. Many terrestrial reptiles, such as lizards, snakes, and turtles, become dormant and hibernate by burrowing in crevices under rocks, logs, and in the ground below the frost line. Terraqueous turtles also become cold-torpid and may

often be found completely submerged in mud and in ponds under ice.

Since the hibernating reptile is subject to the caprices of duration of seasonal low temperature, there is no well-defined period of dormancy. The period of hibernation may often be related to latitudinal positions, as evidenced by the turtle family Emydae. These are aquatic and semiaquatic species, some of which range in the United States from northern latitudes southward through the Mississippi Valley. Species that inhabit the northern climes will hibernate longer than their southern relatives, thus showing hibernation periods which are proportional to the length of the winter period. The physiological characteristics of these hibernating reptiles show many striking alterations of body chemistry. Hibernating reptiles show a loss of appetite and discontinue the ingestion of food. Although the metabolic rate is reduced as much as 95% in hibernating turtles, there is some utilization of stored food products. There are two principal types of reserve food: (1) lipids, which are often found in solid fatty masses in the viscera, under the skin, and distributed in tissues such as muscle, kidney, and particularly liver; and (2) glycogen, the animal starch, which is less stable and more rapidly used than fats. Glycogen is generally localized in tissues such as liver and muscle. There is evidence that these reserve foods are selectively utilized. In hibernating turtles, the tissue glycogen is used during the initial days and weeks of hibernation; later, the lipids are utilized.

The Crocodylia, such as caimans, crocodiles, and alligators, are least able to hibernate. These live in tropical regions where the winter temperatures are relatively uniform. The American alligator (*Alligator mississippiensis*) becomes lethargic and stops eating for several months during winter. This animal undergoes other seasonal physiological modifications as is evidenced by a winter hypoglycemia, that is, a reduction in blood sugar level. The blood sugar level of animals is a physiological index of metabolism and hormonal balance. A major hazard to hibernating poikilotherms is death from freezing; ice crystals form in free protoplasmic water and ultimately destroy the cells and tissues, causing the death of the animal. Frogs, salamanders, and turtles are often surrounded by ice while in hibernation. These animals are able to survive, despite the reduction in body temperatures to about 32 to 31°F (0 to −1°C). As winter approaches, the water content of the tissues becomes reduced and the blood more concentrated. Experiments with turtles have shown that, during winter, there is a lowering of the freezing point of blood. Insects are also credited with the ability to lose body water during winter. Therefore, any physical adjustments which these animals are able to make to rid themselves of water will aid their survival. Hibernating frogs die if, during a severe winter, the frost penetrates their mud hibernacula. Frequently, the uppermost layers of mud freeze and so do its inhabitants, while the animals buried further down in the unfrozen mud remain safe. Frogs, as well as fishes, often freeze to death in shallow artificial pools made of concrete and lacking a mud substrate.

In consideration of the geographical distribution of the amphibians, the limitations imposed by permafrost and the frozen tundra of the northern latitudes can readily be realized. Amphibians are not ordinarily found in permafrost climates.

Fishes. Hibernation in fishes does not occur. Many fishes do, however, spend much of the winter in a state of quiescence while partially frozen in mud and ice. Shallow ponds in temperate, subarctic, and Arctic latitudes often freeze solid in winter. Fishes such as the carp may become encased in the ice or partially buried in the bottom mud. Thus they remain torpid for lengthy periods. The Alaskan black fish (*Dallia pectoralis*) has received some fame, regarding its ability to "freeze solid." Actually, *Dallia*, or any other fish, does not freeze, but becomes encased in ice. If the fish were to freeze solid, with the production of ice crystals in its tissues, it would die. Fishes exposed to cold have a much-reduced metabolism; therefore their oxygen requirements are readily adjusted to the oxygen content of the cold waters. The situation is similar to that which occurs in hibernating aquatic turtles.

The phenomenon of estivation is best known in the dipnoans, that is, the lungfishes. These fishes (*Neoceratodus* in Australia, *Protopterus* in Equatorial Africa, and *Lepidosiren* in South America) are restricted to tropical regions marked by repeated seasons of drought. They survive the dry seasons by becoming dormant and torpid. The lungfishes and their relatives were abundant during the Paleozoic Era. They are, therefore, among the more primitive air-breathing animals possessing a lung which utilizes atmospheric oxygen. This lung becomes the primary organ of respiration during the torpidity of estivation. In general, the lungfishes follow a similar behavioral pattern as the dry seasons approach. *Protopterus*, for example, burrows in the bottom mud as the water begins to diminish during the dry season. Once in its burrow, the fish assumes the coiled posture of seasonal dormancy (**Fig. 1**). A lifeline of air is provided by the tunnel from the burrow to the surface. In preparation for estivation, *Protopterus* secretes a slimy mucus around itself which hardens in a tight cocoonlike chamber. This chamber forms an effective hygroscopic structure which prevents the desiccation of the fish. There is but one opening, formed around the mouth. Thus the air from the tunnel enters the mouth and passes to the lung apparatus. At the termination of the dry season, water slowly enters the burrow, softens the contents, and awakens the lungfish. The metabolism of the lungfish is at a low ebb during estivation, with the energy for its modest life processes provided by the utilization of tissue protein. This is evidenced by the large increase in the formation of urea, a product of protein metabolism. The urea is promptly excreted in large quantities when the fish returns to water. *Lepidosiren*, the South American member of the family of lungfishes, sometimes called the mud siren, also

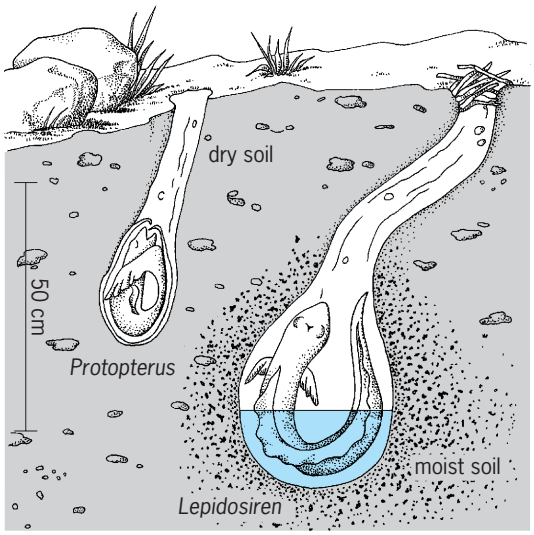

Fig. 1. Lungfishes in estivation.

burrows. The burrow entrance is loosely plugged with surface clay, perforated with air holes, each about 0.5 in. (1.25 cm) in diameter. Mud sirens burrow in the deepest and most concave bottoms, seemingly to ensure that they will receive the inflowing water when it first returns to the pond. The fishes' tunnellike burrows are about 2–3 in. (5–7.5 cm) in diameter and extend downward about 2–3 ft (0.6–0.9 m). The mud siren estivates in a gourd- shaped, gelatinous-coated burrow which may be partially filled with water. These fishes, therefore, differ from their African relatives in that they maintain a soft, moist medium. Lungfishes in estivation are readily handled and transported, indicating the extremely quiescent nature of their torpor.

Invertebrates. This is often more than a period of seasonal dormancy, because in some snails estivation may be extended for years at a time, and among the insects and spiders the period of hibernation becomes intimately associated with a phase in the life cycle. During the winter months and during a hot dry summer, the soil contains a remarkable variety of torpid invertebrates, for example, earthworms, snails and slugs, nematodes, insects and spiders, grubs, larvae, and pupae of many insects, egg cases, and cocoons. The hibernacula for dormant invertebrates are also provided by rotting logs and vegetation, the undersurfaces of bark, and crevices in rock and building foundations.

Insecta. Insects overwinter, for the most part, in the egg or larval stage of metamorphosis (egg→larva→pupa→adult). Hibernation frequently becomes integrated with the diapause, or arrested development, of the egg or larva which occurs during the winter. The egg of the silkworm (*Bombyx mori*) undergoes diapause in conjunction with hibernation. The eggs are laid in autumn and must be exposed to about 32°F (0°C) for several months to ensure hatching in the spring. The Syrphidae, commonly referred to as the hover flies, are ordinary garden inhabitants. They live in close association with plants, as commensals and parasites, and are also known to infest animals, producing myiasis. The larvae of the syrphids are particularly susceptible to hibernation and estivation. They become markedly desiccated but are quickly reactivated by exposure to water. The Lepidoptera, butterflies and moths, offer some of the most common examples of hibernating larvae and pupae. The familiar cocoon of the butterfly is the hibernaculum of the larva and pupa. The disappearance of mosquitoes in the Arctic, subarctic, and temperate zones in autumn and winter is also a reflection of insect hibernation. The adult hibernates, after its last autumnal blood meal, and presumably subsists on its stored food reserves. Provision for overwintering in *Culex* and *Anopheles* mosquitoes lies in their development of a reserve food material in the form of a fat body. *See* INSECT PHYSIOLOGY; MYIASIS.

Mollusca. Land mollusks, such as snails and slugs, are noted for their ability to hibernate and estivate. Land snails are particularly cold-hardy and have frequently been found encased in ice. Among the more cold-resistant land snails are species of *Physa*, which normally live in northern Siberia, 73°30′N latitude, where the mean annual temperature is below −4°F (−20°C). In the subarctic and temperate zones most fresh-water mollusks can live in water under the ice and are often enclosed in solid ice. The common garden snail (*Helix aspersa*) and the edible snail (*H. pomatia*) exhibit the chief characteristics of hibernation common to the Gastropoda. In autumn, these snails cease ingesting food and burrow into the ground under logs, leaves, and other forms of natural debris, where they remain in hibernation for about 6 months. The animal makes this burrow with its muscular foot and settles down to a winter sleep. The hibernating position is such that the shell aperture is turned upward. The hibernating snail may be found singly or in clusters, with numerous individuals in contact with each other. The animal seals its aperture by secreting a membrane, the epiphragma, chiefly composed of calcium phosphate. In the epiphragma, there remains a minute respiratory opening. The respiratory movements are drastically reduced and, as in other poikilotherms, the rate of heart beat is related to temperature; at 86°F (30°C) the heart rate is 50–60 beats per minute, whereas in hibernation the rate is reduced to about 4–12 per minute. The energy for the lowered metabolism during hibernation appears to be derived from the utilization of stored fats, with a subsequent loss of body weight.

The slugs, referred to as garden snails without shells, are also hibernants. Their hibernaculum is similar to that of the shelled forms. They hibernate in the ground under leaves, tree bark, and other forms of decomposing vegetation. Estivation of land mollusks occurs chiefly in tropical climates, where they bury themselves deeply in the ground or under rocks. They also form a type of epiphragma as a defense against evaporation and desiccation. Tropical snails are able to estivate for lengthy periods, months and possibly years at a time. A speciman of *Helix desertorum* is reputed to have awakened after 5 years of dormancy. *See* MOLLUSCA.

The two marine mollusks periwinkles (*Littorina*) and mussels (*Mytilus* and *Modiolus*) cannot be regarded as hibernators, yet they are extremely cold-hardy These animals are inhabitants of the intertidal zone. In such latitudes as Woods Hole, Massachusetts, and Hebron, Labrador, they may be exposed repeatedly to temperatures of −4°F (−20°C) in winter. At such temperatures, ice forms inside the shelled animals. These seashore mollusks can live through the winter season with as much as 70–75% of the body cavity being recurrently filled with ice. At such low temperatures the animals are being dehydrated, since a great deal of the body water is being converted into unusable ice and the remaining fluid becomes a more concentrated salt solution. In these animals, the body water which freezes is not tissue or cellular protoplasmic fluid, but is chiefly extracellular water.

Nemata. Nematodes are cosmopolitan and have been found in practically all ecological settings. However, the terrestrial and fresh-water forms are most apt to respond to winter exposure by becoming quiescent. Pasture puddles, which often freeze solid in winter, harbor a particularly rich nematode fauna. Literally millions per acre are found in the top few inches of soil. Among the "cushion" plants such as mosses and lichens one finds frequent nematode "parasitism." In addition to a rich fauna of nematodes, these plants also harbor tardigrades, rotifers, and even protozoans. These animals are characterized by their capacity to endure the extremes of drought, heat, desiccation, and cold. Under such forms of environmental stress they go into a dormant state, and revive when environmental conditions become less severe and again conducive to active life. The particular nematodes such as *Mononchus, Dorylaimus, Plectus, Monhystera, Cephalobus, Trilobus, Trypyla, Tylenchus,* and *Aphelenchus* are considerably alike throughout the world, and in spite of wide longitudinal and latitudinal separation their overwintering and drought behavior are similar.

Their resistance notwithstanding, nematodes do die and decline in numbers during the winter. *See* NEMATA (NEMATODA); ROTIFERA; TARDIGRADA.

Bryozoa. The Ctenostomes are the only Bryozoa with fresh-water representatives. These animals bear a resemblance to moss and seaweeds and are collectively called moss animals. The colonial fresh-water *Paludicella* and *Victorella* form hibernacula. These are specialized external buds which persist through the winter while the rest of the colony breaks down. At the termination of winter, the shell of the hibernaculum splits into two halves and a young colony is brought forth. *See* BRYOZOA.

Protozoa. The phenomenon of encystment is commonplace in the Protozoa, or single-celled animals. Encystment is remarkably similar to estivation and hibernation, and an encysted protozoon is extremely quiescent and almost nonmetabolizing. It is generally accepted that encystment is a protective measure against cold and desiccation, and yet it may coincide with some general physiological, protoplasmic reorganization or dedifferentiation in the species. For example, during encystment of the ameba *Pelomyxa carolinensis*, there is a reduction in number of nuclei, a concentration of protoplasm, and the formation of a cyst membrane and shell. The ciliate *Colpoda cucullus* during its encystment undergoes nuclear reorganization, and multiplication of the organism takes place. The ability of protozoa to encyst is at least one of the reasons for their worldwide distribution. The Arctic tundra, which remains frozen for 6–8 months of the year, teems with protozoa as well as other invertebrate life. These microscopic animals could not live through these winter rigors unless they had this capacity for self-preservation and protection.

The free-living protozoa offer many examples of encystment due to winter low temperatures and to evaporation due to drought. The cyst may be regarded as the hibernaculum, and it is usually found in mud or ice and soil. Cysts of the single-celled organisms can be recovered from ice and mud samples taken in midwinter and also from dried soil collected during the most arid summers. The protozoan cyst is often, as with many invertebrates, a function of reproduction and therefore a stage in the life cycle of the species. The ciliates *Colpoda* and *Didinium* are noted for their cyst formation. Encysted *Colpoda* can resist experimental freezing for as long as 2 months, and can be activated after being in a dried state for as long as 5 years. *See* PROTOZOA.

Hibernacula. The hibernacula of poikilotherm vertebrates and invertebrates are as varied as the animals themselves (**Fig. 2**). The minute cysts in protozoa, the cocoon and egg case of insects and spiders, the burrows and crevices of reptiles, and the dried mucous case of the lungfish, in all instances, protect the animal from evaporation or desiccation and freezing.

X. J. Musacchia

Warm-Blooded Vertebrates

Many mammals and some birds spend at least part of the winter in hiding, but remain no more drowsy than in normal sleep. On the other hand, some mammals undergo a profound decrease in metabolic rate and physiological function during the winter, with a body temperature near 32°F (0°C). This condition, sometimes known as deep hibernation, is the only state in which the warm-blooded vertebrate, with its complex mechanisms for temperature control, abandons its warm-blooded state and chills to the temperature of the environment. Between the drowsy condition and deep hibernation are gradations about which little is known. The bear, skunk, raccoon, and badger are animals which become drowsy in winter. Although usually considered the typical hibernator, the bear's body temperature does not drop more than a few degrees. Body temperature, measured rectally, was 96°F (35.5°C) at an air temperature of 40°F (4.4°C). At 25°F (−3.5°C), body temperature was 88°F (31.2°C).

Deep hibernators. The deep hibernators are confined to five orders of mammals: the marsupials, the Chiroptera or bats, the insectivores, the rodents, and, probably, the primates. Some of the South American

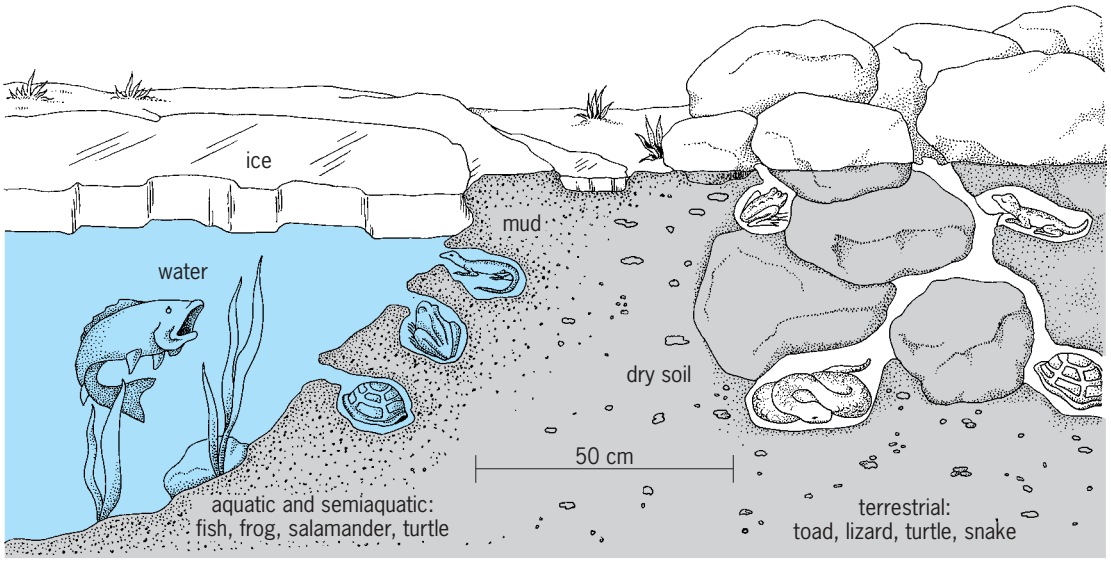

Fig. 2. Hibernacula of various cold-blooded vertebrates.

opossums are reported to hibernate. Most, if not all, of the insect-eating bats of temperate climates not only hibernate in the winter, but also drop their body temperature when they roost and sleep. The advantage of this for a small mammal with a disproportionately large heat-losing surface is obvious when conservation of energy is considered. Some monotremes and marsupials hibernate. Among the insectivores, the European hedgehog is a deep hibernator, and the Madagascan tenrec (*Centetes*) probably hibernates or estivates during the summer. Many rodents are deep hibernators, including ground squirrels, woodchucks, dormice, and hamster. The fat-tailed and mouse lemurs are primates that hibernate or estivate. Among birds, the poorwill (*Phalaenoptilus*) and some hummingbirds and swifts undergo a lowering of body temperature and metabolic rate in cold periods.

With all deep hibernators, except the bats, hibernation is seasonal, usually occurring during the cold winter months. In all cases, it occurs in animals which would face extremely difficult conditions if they had to remain active and search for food. A few desert-living species disappear during the driest, hottest months, and it is probable that they enter into a stage of deep hibernation in their cool burrows, though no studies have been made on this, or on the estivation of animals during the tropical rainy season. Hibernators are restricted typically to more temperate zones, but the Alaskan ground squirrel lives near the Arctic Circle.

Physiology. During a preparation period for hibernation, the animals either become fat, like the woodchuck, or store food in their winter quarters, like the chipmunk and hamster. Prior to hibernation, there is a general involution of the endocrine glands, but at least part of this occurs soon after the breeding season and is not directly concerned with hibernation. Animals such as ground squirrels become more torpid during the fall, even when kept in a warm environment, indicating a profound metabolic change which may be controlled by the endocrine glands. On the other hand, ablation of any of the glands has not brought on hibernation. Lack of food will cause hibernation in the spiny pocket mouse (*Perognathus*), but in most hibernators this has little if any effect, and the stimulus for hibernation is not known. It has been reported that an extract from the blood of an animal in hibernation will induce hibernation when infused into an active potential hibernator, indicating that the factor which produces hibernation may be bloodborne.

Neural mechanism. Hibernation in mammals is not caused by an inability to remain warm when exposed to cold, for hibernators are capable of very high metabolic rates and sometimes do not enter hibernation if exposed to cold for months at a time. When the animal is entering hibernation, heart rate and oxygen consumption decline before body temperature, indicating that the animal is actively damping its heat-generating mechanisms. The autonomic nervous system is involved in this process, for the heart is slowed to a lesser extent in animals entering hibernation when their parasympathetic system is blocked by atropine. As normal hibernation deepens, the heart rate, blood pressure, metabolic rate, and body temperature slowly drop, but in some animals periodic bouts of shivering and increased oxygen consumption occur, elevating the body temperature temporarily and causing a stepwise entrance into hibernation. Presumably this is caused by bursts of activity of the sympathetic nervous system, which otherwise remains in abeyance during entrance into hibernation. The preoptic-hypothalamic area of the brain is the center for temperature control in mammals. Ground squirrels with lesions in this area will enter into a state resembling hibernation, but detailed physiological comparisons of lesioned and normal hibernators have not been made. *See* AUTONOMIC NERVOUS SYSTEM.

Homeostasis. In deep hibernation at a steady state the body temperature is 33–35.5°F (0.5–2°C) above

that of the environment, and it is a peculiarity of hibernators that the vital processes can function at lower temperatures than those of nonhibernators. The heart rate varies between 3 and 15 beats per minute, and blocking or stimulating the vagus nerve has no effect on the heart rate, indicating that the parasympathetic nervous system cannot be regulating the changes in heart rate. The metabolic rate is less than one-thirtieth of the warm-blooded rate at rest, and the main source of energy is fat, since the respiratory quotient (RQ) is about 0.7. An RQ above this figure would indicate that proteins or carbohydrates were being used. The torpid black bear, with its relatively high body temperature, also uses fat for energy, whereas the starved, active bear uses fat and protein.

In spite of its low body temperature, the hibernating animal retains a remarkably rigid control of its internal environment. Blood sugar levels are low in some hibernating animals, but there is no known consistent variation in any of the blood constituents which can be reasonably considered as a cause of the hibernating state. The hibernating animal senses changes in its environment and responds to increases in atmospheric CO_2 by increasing its respiratory rate. This sensitivity to the environment varies with time, even when body temperature remains the same, but the causes of this variation are not known. In its most sensitive state an animal in hibernation will respond with muscular activity to the movement of one of its hairs or to a light breeze playing over its back.

Temperature control. If the environmental temperature drops to 32°F (0°C), the hibernating animal may respond either by increasing its metabolic rate and remaining in hibernation or by a complete arousal from the hibernating state. Chilling the preoptic-anterior hypothalamic area of the brain in hibernating golden-mantled ground squirrels causes an increase in metabolic rate, which varies linearly with the decrease in the temperature of the brain, indicating that the hibernating animal is truly regulating its body temperature. Unlike animals in the warm-blooded state, hibernating hamsters and ground squirrels do not sense the increasing cold with peripheral receptors, for heart and metabolic rate do not increase as chilling proceeds if the brain temperature does not decline. However, dormice, which hibernate in nests exposed to rapid temperature variations, increase heart and metabolic rate when the periphery is chilled and brain temperature remains unchanged. Hibernating marmots also sense peripheral cold. In this case the temperature sensors in the brain are insulated from the cold by the curled position of a relatively large animal, and the marmot may need peripheral input to avoid freezing. Thus, some animals at the low temperatures of hibernation appear to have the same reactions to a cold environment as mammals in the warm-blooded state. Between an environmental temperature of about 39 and 59°F (4 and 15°C) the hibernating animal makes no attempt to regulate body temperature at a preferred level, and in this aspect it differs from the warm-blooded animal, which regulates its temperature between a low and a high point. If hibernation does involve lowering the "set point" in order to control body temperature, the method by which this is accomplished is unknown.

Temperature and pH effects. A hibernating mammal reduces its metabolic rate by nearly 30-fold and shifts from glycogen to lipid (that is, fat stores) as the major fuel source for metabolism. The magnitude of metabolic rate reduction is far in excess of what would be expected solely as a result of a hibernator's lowered body temperature (T_b, down from 98.6 to 44.6°F, or 37 to 7°C, in some species). Moreover, suppression of glycogen metabolism during hibernation must be poised for regular and rapid relaxation during periods of arousal (which are fueled by glycolysis) as well as at the end of the hibernation period.

Mechanisms controlling these aspects of hibernation metabolism appear to be the relative acidification of the intracellular fluids of the hibernator. This is a consequence of the hibernator's tendency to continuously regulate its blood pH (at about pH 7.4, termed pH stat) while its T_b drops, and of the adoption of a modified breathing pattern that, although variable among species, is typified by periods of apnea lasting up to 2 h that are interspersed between 3–30 min intervals of rapid ventilation.

Just as the neutral pH of water changes with temperature due to shifts in its dissociation constant, varying the T_b of an ectotherm causes shifts in the pH of its blood and intracellular fluids, at an average of −0.015 pH unit/°C. At 98.6°F (37°C) T_b, the blood pH of most ectotherms would be about 7.4, and about 7.85 at 44.6°F (7°C) T_b. Due to both the physical nature of intracellular fluids and the ectotherm's homeostatic and metabolic compensation for these thermally caused pH shifts (that is, mainly through the regulation of total CO_2), the histidine imidazole buffer groups of many of its regulatory proteins (enzymes) gain or release protons in order to maintain a constant dissociation ratio. By preserving the dissociation state (termed alpha stat regulation) of a key enzyme, its activity and in turn the metabolic process it catalyzes could remain near optimal at the specific combinations of intracellular temperature and pH imposed by the range of T_b experienced by the ectotherm. *See* PH REGULATION (BIOLOGY).

Since hibernating mammals tend to maintain their blood pH near 7.4, blood and most intracellular pH values (except those in the heart and liver which may have to function normally to sustain life at such low rates) are acidotic relative to the temperature-pH relationships known for ectotherms. Alpha stat regulation does not occur, and increased protonation of the imadazole buffers leads to changes in the kinetic and structural properties of enzymes, depending upon their pH and temperature sensitivity. Phosphofructokinase (PFK), a key enzyme in the control of glycolysis, becomes completely inactivated at the intracellular temperature (43°F or 6°C) and pH (7.0) conditions prevailing in the muscle of a hibernating ground squirrel *(Spermophilus)*. Glycogen metabolism is similarly reduced in hibernating

bats *(Eptesicus)*, even though their rate of fat oxidation is unaffected by this state.

Inactivation of PFK by imidazole protonation conserves glycogen stores, and may account for the large decrease in the metabolic rate of hibernators. Also, since it occurs rapidly and is reversible, this deactivation would permit arousal metabolism. In addition to the temperature-pH effect, PFK is also inhibited by low amounts of urea. Also, apnea (through CO_2 retention) causes acidosis, which apart from a large drop in T_b could inactivate PFK. This suggests a glycogen-energy–serving strategy used by mammals that hibernate or enter daily torpor (for example, tropical bats) in warm habitats (where T_b would never drop to 44.6°F or 7°C). Through apnea or rebreathing of hibernacula air containing high CO_2, PFK inactivation might be induced. Finally, the different effects of the pH stat and alpha stat strategies on energy metabolism and tissue function have important biomedical implications for human hypothermic surgery. Jeffrey B. Graham

Arousal. The hibernator is capable of waking at any time, using self-generated heat, and this characteristic clearly separates the hibernating state from any condition of induced hypothermia. During the total period of hibernation, the hibernator spontaneously wakes from time to time, usually at least once a week. In the period of wakefulness the stored food is evidently eaten, but animals which do not store food rely on their fat for the extra energy during the whole winter. The cause of the periodic arousals has not been definitely determined, but it is theorized that the arousal is due to the effect of the accumulation of a metabolite or other substance which can be neutralized only in the warm-blooded state.

Arousal is a complex, coordinated event which results in warming in the shortest time, usually about 3 h when in an environmental temperature of 41°F (5°C). As soon as the animal starts to rouse, the heart speeds and oxygen consumption rises, followed shortly by an increase in body temperature. The anterior parts of the body warm rapidly, but because of vasoconstriction, the posterior parts remain cold until late in the warming process. Shivering can be detected early, using the electromyogram, and it becomes more obvious as the animal warms. The energy for warming, at least in rodents, is obtained from glycogen stored in the liver and other tissues. It has been calculated that one waking uses as much energy as does 10 days of hibernation, so that waking must use most of the winter stores, whether fat or fodder. Charles P. Lyman

Anaerobiosis

For many ectothermic vertebrates (fishes, amphibians, and reptiles) the ability to avoid seasonal and periodic environmental rigors by entering a state of metabolic inactivity is a crucial element in their survival. Specifically, winter dormancy and summer estivation—the usual context in which these terms are applied to ectotherms—permit these animals to survive and flourish, first, by reducing the impact of seasonal extremes (such as cold, heat, dryness, and ice) and, second, by significantly lowering the ectotherm's energetic costs during times that would not be favorable for activity (that is, when food is available).

Regulating factors. As in hibernating endotherms (birds and mammals), a key factor regulating seasonal torpor in ectotherms is the continuous internal monitoring of environmental cues (such as day length) which in turn triggers temporally precise seasonally adaptive changes in systemic function, metabolism, and behavior. Thus, both the desert-inhabiting spadefoot toad, *Scaphiopus couchi*, and the horned lizard, *Phrynosoma*, are obligatory hibernators; they enter winter dormancy at a fixed time each year, even if local conditions might favor their activity for a slightly longer period. A second important factor is the presence in ectotherms of a bioenergetic metabolic system that, when compared to mammals and birds, operates at a much lower intensity and has less absolute dependence on molecular oxygen. The metabolic energy adaptations for seasonal torpor in ectothermic vertebrates are to a large extent similar to those required by vigorious activity or prolonged diving, and thus involve the processing or storage of intermediate metabolitics such as lactic acid, the regulation of intra- and extracellular pH, and enduring periods without access to oxygen.

Energy production. Both endotherms and ectotherms acquire energy through the metabolic breakdown of stored food reserves (carbohydrates, fats, and protein) and the transfer of the released chemical energy into molecules of adenosine triphosphate (ATP). Both anaerobic (glycolysis) and aerobic (oxidative phosphorylation) metabolic pathways are used in ATP production, and in most cells these two processes are coupled. The product of glycolysis, acetyl coenzyme A (AcCoA), flows directly into the aerobic pathway, where it is fully oxidized to water and carbon dioxide. Glycolysis has a much lower rate of ATP yield per volume of substrate than does oxidative phosphorylation; but it will proceed in the absence of oxygen, whereas oxidative phosphorylation cannot. Cells metabolizing without oxygen convert AcCoA into lactic acid, the anaerobic storage product, until oxygen becomes available. For example, during vigorous activity, the short-term energy utilization by an animal's muscle far exceeds oxygen availability and, following activity, the animal repays the oxygen debt accrued by this tissue in the oxidation of the lactic acid. In ectotherms, glycolytic energy production contributes from as little as 50% to as much as 98% of the total energy utilized during activity, which is a much higher proportion than in endotherms. Even in the presence of oxygen, however, aerobic glycolysis occurs, and the AcCoA produced is not always shunted into oxidative phosphorylation. Except that constant low levels of lactic acid are present in the tissues, little is known about resting levels of aerobic glycolysis in animals. However, prolonged confinement in burrows or in a submerged hibernaculum may require an ectotherm to

suspend oxidative metabolism and rely solely upon anaerobic metabolism. *See* ENERGY METABOLISM; METABOLISM.

Terrestrial ectotherms. Subterranean refugia permit escape from extreme surface temperatures, and cool subsurface temperatures reduce metabolism. During winter dormancy, spadefoot toads buried 20 in. (50 cm) in sand lower aerobic metabolism to 20% of the resting level. Some toads and frogs envelop themselves in cocoons made from many layers of unshed skin. In summer the yellow mud turtle *Kinosternon* leaves drying ponds to estivate in a dry burrow for as long as 3 months. During estivation, oxygen consumption declines by 75% and the animal becomes anaerobic. Upon emergence, the mud turtle quickly restores its considerable water losses, unloads a large volume of carbon dioxide, and then resumes a normal rate of oxygen consumption.

Aquatic ectotherms. Seasonal inactivity in aquatic ectotherms imposes an additional set of hardships in which specializations related to diving come into play. Winter-dormant turtles (*Pseudemys*) and bullfrogs (*Rana*) have subsurface burrows that penetrate deeply into anoxic mud; thus they are without access to oxygen. The tropical marine turtle *Chelonia* enters a temperature induced torpidity and remains submerged, sometimes in anoxic areas, for considerable periods. *Pseudemys* can survive long periods without oxygen.

If lactic acid is stored in the body, it must be neutralized or the tissue pH will be affected. This is usually done by bicarbonate and chloride, but in turtles (*Chrysemys*) kept submerged without oxygen for 3 months lactate levels become so high that calcium and magnesium must be mobilized from its bone and shell to mitigate the effect. In fact, anoxia stimulates a compensatory lowering of the amounts of cytochrome oxidase, an important aerobic enzyme, in its brain, heart, and muscle. A key glycolytic enzyme, pyruvate kinase, undergoes an increase in the turtle's muscles, but not in its heart. A great many fishes endure hypoxic and anoxic conditions under winter ice, in stagnant swamps that may also become dry, and in deep lakes and ocean basins. *Rasbora*, a cyprinid, can survive over 100 days without oxygen. Goldfish and carp (*Carassius*) can overwinter under ice without oxygen for from 2 to 5 months. Goldfish are seasonally adapted to these conditions. In anoxia this fish dramatically rearranges its anaerobic metabolism, and lactic acid does not concentrate. Rather, it is remetabolized to AcCoA and carbon dioxide. AcCoA is in turn broken down to ethyl alcohol which, along with carbon dioxide, is excreted by the anoxic goldfish. Even some air-breathing fish endure anoxia in their drying mud burrows. The aerobic metabolism of an estivating African lungfish (*Protopterus*) is reduced to 20% of resting. The South American lungfish (*Lepidosiren*) and the swamp eel (*Synbranchus*), both competent air breathers, have very high levels of glycogen and glycolytic enzymes in their hearts to serve them when they become anaerobic during prolonged burrow dives. Jeffrey B. Graham

Bibliography. A. F. Bennett and J. A. Ruben, Endothermy and activity in vertebrates, *Science,* 206:649–654, 1979; S. F. Hand and G. N. Somero, Phosphofructokinase of the hibernator *Citellus beecheyi*: Temperature and pH regulation of activity via influences on the tetramer-dimer equilibrium, *Physiol. Zool.*, 56:380–388, 1983; M. Harless and H. Morlock (eds.), *Turtles*: *Perspectives and Research*, 1989; C. Heller and G. W. Colliver, *Amer. J. Physiol.*, 227:538–589, 1974; P. W. Hochachka, *Living Without Oxygen*, 1980; J. Hudson, Torpidity in mammals, *Comparative Physiological Thermoregulation*, vol. 3, pp. 97–165, 1973; L. L. McClanahan, R. Ruibal, and V. H. Shoemaker, Rate of cocoon formation and its physiological correlates in a ceratophryd frog, *Physiol. Zool.*, 56:430–435, 1983; A. Malan, Respiration and acid-base state in hibernation, *Hibernation and Torpor in Mammals and Birds*, pp. 237–282, 1982; F. N. White, A comparative physiological approach to hypothermia, *J. Thorac. Cardiovasc. Surg.*, 82:821–831, 1981; F. N. White and G. N. Somero, Acid-base regulation and phospholipid adaptations to temperature: Time courses and physiological significance of modifying the milieu for protein function, *Physiol. Rev.*, 62:40–90, 1982; M. E. Yacoe, Adjustments of metabolic pathways in the pectoralis muscles of the bat, *Eptesicus fuscus*, related to carbohydrate sparing during hibernation, *Physiol. Zool.*, 56:648–658, 1983.

Hickory

Any species of the genus *Carya*, formerly known botanically as *Hicoria*. Hickories are mostly tall forest trees characterized by strong, terminal, scaly winter buds, pinnately compound leaves (see **illus.**), solid pith (not chambered), and fruit with an outer husk or exocarp which splits more or less readily into four parts, revealing a nut with a hard shell or endocarp. *See* FAGALES.

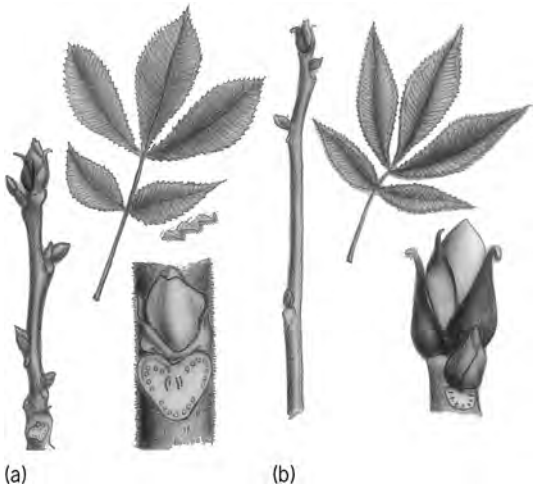

Twigs, buds, and leaves of hickories. (*a*) Shagbark (*Carya ovata*). (*b*) Pignut (*C. glabra*).

The shagbark hickory (*C. ovata*) grows to a height of about 120 ft (36 m) and is found in the eastern half of the United States and adjacent Canada. It is the most important species because of the commercial value of its nuts, the hickory nuts of commerce, and of its wood. It is easily recognized by the bark, which in older trees exfoliates from the trunk in long, curving, irregular plates, and by the leaves, which usually have five leaflets, of which the terminal one is larger.

The pecan (*C. illinoensis*) is also a valuable species because of its commercially popular, thin-shelled, sweet nuts. Although its native range is limited to the Mississippi Valley region and Mexico, many varieties are cultivated in the southern United States.

Other species are the mockernut and pignut hickories, widely distributed through the eastern half of the United States, and the shellbark hickory found in the east-central United States. The remarkably tough and strong wood of all species makes it the world's best wood for tool handles. It is also used for parts of vehicles, furniture, flooring, boxes, and crates, and for smoking meats. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Hidden variables

Additional variables or parameters that would supplement quantum mechanics so as to make it like classical mechanics. Hidden variables would make it possible to unambiguously predict (as in classical mechanics) the result of a specific measurement on a single microscopic system. In contrast, quantum mechanics can give only probabilities for the various possible results of that measurement. Hidden variables would thus provide deeper insights into the quantum-mechanical probabilities. In this sense the relationship between quantum mechanics and hidden variables could be analogous to the relationship between thermodynamics (for example, temperature) and statistical mechanics (the motions of the individual molecules). *See* STATISTICAL MECHANICS.

As an example, in Young's double-slit experiment the pattern observed on a distant screen consists of alternating bright and dark regions. This pattern is the average of the results for a large number of photons passing through the double-slit arrangement, and represents the quantum-mechanical probability that a photon will be detected in any specific region of the screen. Hidden variables would, in principle, make it possible to precisely predict where each photon would be detected. *See* INTERFERENCE OF WAVES.

The concept of hidden variables appeared during the early development of quantum theory, apparently the result of an instinctive response engendered by the deterministic nature of classical mechanics. Specifically, it was felt that if the same measurement was made on each member of a group of identically prepared systems, the result of each measurement should be the same. If several different results were observed, then those identically prepared systems must not have been identical; that is, it was believed that the systems must have been in different microstates that were unknown at the time of preparation. The different microstates would be specified by the hidden variables. The variables were hidden only in the sense that they had not been observed. There were no restrictions as to whether or not they would, in principle, be possible to observe.

F. J. Belinfante formulated a three-section classification scheme for hidden variable theories—zeroth kind, first kind, and second kind. Most interest, both theoretically and experimentally, has been focused on hidden variable theories of the second kind, also known as local hidden variable theories.

Hidden variables of zeroth kind. In 1932 J. von Neumann provided an axiomatic basis for the mathematical methods of quantum mechanics. As a sidelight to this work, he rigorously proved from the axioms that any hidden variable theory was inconsistent with quantum mechanics. This was the most famous of a number of proofs, appearing as recently as 1980 and purporting to show the impossibility of any hidden variable theory. In 1966 J. S. Bell pinpointed the difficulty with von Neumann's proof—one of his axioms was fine for a pure quantum theory which makes statistical predictions, but the axiom was inherently incompatible with any hidden variable theory. The other impossibility proofs have also been found to be based on self-contradictory theories. Such theories are called hidden variable theories of the zeroth kind.

An impossibility proof originally based on a theorem by A. M. Gleason is, however, instructive since it does rigorously rule out one form of hidden variable theories whose precise predictions for a single system do not depend on a specific measurement procedure. This is also consistent with N. Bohr's dictum that a measurement is a joint result of both the system being measured and the measuring apparatus. It is easy to forget this point and be seduced by the false expectation that the result of a quantum-mechanical measurement is a preexisting property having nothing to do with the measuring apparatus.

Hidden variables of the first kind. Hidden variable theories of the first kind are constructed so as to be self-consistent and to reproduce all the statistical predictions of quantum mechanics when the hidden variables are in an "equilibrium" distribution.

D. Bohm introduced the first important theory of this type in 1952. It was based on ideas that had already been introduced in 1927 by L. de Broglie in his description of the "pilot wave." In Bohm's theory the hidden variable for a particle is identified with the position, x, of the particle, and the hidden variable equilibrium distribution is set equal to the particle's position probability given by quantum mechanics. Basically, this theory is a reinterpretation of nonrelativistic quantum mechanics in terms of hidden variables. However, it is inherently nonlocal (as below), and no satisfactory relativistic formulation has been given. As yet no prescriptions have

been found for experiments that would distinguish the theory from conventional quantum mechanics. Thus, as a description of nature, it is an unnecessary complication of quantum mechanics, and the principle of Occam's razor may be applied.

Two first-kind theories that can be tested experimentally are those of N. Wiener and A. Siegal, and of Bohm and J. Bub. These postulate an equilibrium distribution of hidden variables that is gaussian. Deviations from the statistical predictions of quantum mechanics can be observed only when this equilibrium distribution is disturbed. Consequently, these theories can be tested by performing an experiment that disturbs this equilibrium distribution and then performing a second experiment before the hidden variables can relax. C. Papaliolios did such a test by observing the passage of photons through two closely spaced linear polarizers. The results were in agreement with the statistical predictions of quantum mechanics. Hence, the hidden variables described by these theories would have to relax to their equilibrium distributions in this experiment in a time less than about 2×10^{-14} s. From a heuristic point of view, Bohm and Bub estimate a relaxation time of about 10^{-13} s. Thus, this experiment decreases the credibility of these theories but does not entirely rule them out. *See* DISTRIBUTION (PROBABILITY).

Hidden variables of the second kind. These theories predict deviations from the statistical predictions of quantum mechanics, even for the "equilibrium" situations for which theories of the first kind agree with quantum mechanics. They are generally called local hidden variable theories because they are required to satisfy a locality condition. Intuitively, this seems to be a very natural condition. Locality requires that an apparatus at one location should operate independently of any settings or actions of a second apparatus at a spatially separated location. In the strict Einstein sense of locality, the two apparatus must be independent during any time interval less than the time required for a light signal to travel from one apparatus to the other.

Einstein-Podolsky-Rosen thought experiment. The focus for much of the discussion of local hidden variable theories is provided by a famous thought experiment (a hypothetical, idealized experiment in which the experimental results are deduced) that was introduced in 1935 by A. Einstein, B. Podolsky, and N. Rosen. Their thought experiment involves an examination of the correlation between measurements on two parts of a single system after the parts have become spatially separated. They used this thought experiment to argue that quantum mechanics was not a complete theory. Although they did not refer to hidden variables as such, these would presumably provide the desired completeness. The Einstein-Podolsky-Rosen (EPR) experiment led to a long-standing philosophical controversy; it has also provided the framework for a great deal of research on hidden variable theories.

A version of the Einstein-Podolsky-Rosen experiment is based on the single system consisting of the two photons emitted in an atomic cascade. A state for two such photons moving in opposite directions along the z axis is given by the equation below, where $|x\rangle_i$ and $|y\rangle_i$ are linear polarization states for

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\{|x\rangle_1|x\rangle_2 - |y\rangle_1|y\rangle_2\}$$

photon i in the x and y directions, respectively; x and y are arbitrary orthogonal directions in a plane perpendicular to the z axis. Quantum mechanics assumes this state remains the same regardless of the distance between the two photons, and it predicts very strong correlations between the measurements of the polarizations for the two spatially separated photons. In fact, if the measurement of polarization for photon 1 gives a result along the x (or y) axis, measurement of the polarization for photon 2 can be predicted with absolute certainty to also be along the same x (or y) axis. This is an obvious result in view of the conservation of angular momentum for the entire, spatially separated, two-photon system; but in view of the spatial separation, it is an extraordinary result that already suggests a hidden variable explanation (that is, each photon would carry with it a prescription that predetermines the result of any polarization measurement). Furthermore, this same result holds for any different set of axes, say x', y'. Thus, by simply measuring polarization for photon 1, the result of measuring polarization for photon 2 can be predicted with absolute certainty in either the primed or unprimed axes (or, in fact, any such set of axes) without disturbing it. Einstein, Podolsky, and Rosen argued that, for any axis, the polarization of photon 2 must therefore somehow be real (predetermined); and since quantum mechanics does not encompass this, it must be an incomplete theory. N. Bohr has supplied the most famous response to this argument by questioning the Einstein-Podolsky-Rosen definitions of real and complete, and emphasizing the meaning of measurement as a function of both the microsystem and the entire measuring apparatus.

The nonlocality aspect of the Einstein-Podolsky-Rosen experiment is especially important. In the above example, if photon 1 passes a polarizer in any orientation, photon 2 absolutely must also pass another polarizer in the same orientation. Similarly if photon 1 does not pass, photon 2 absolutely must not pass. This kind of behavior is easily understood if the photons carry prescriptions (hidden variables) telling them what to do in every measurement case. But suppose such prescriptions do not exist. How then does photon 2, together with the second polarizer, "know" whether or not photon 2 should pass if locality is enforced and it cannot "know" the polarizer orientation for the measurement being made on photon 1?

Bell's result. The Einstein-Podolsky-Rosen experiment appeared in 1935; nevertheless, the von Neumann "impossibility" proof was dominant for many years. However, new efforts were stimulated in 1952 when Bohm did the "impossible" by designing the

hidden variable theory of the first kind described above. Bohm's theory was explicitly nonlocal, and this fact led Bell to reexamine the Einstein-Podolsky-Rosen experiment. He came to the remarkable conclusion that any hidden variable theory that satisfies the condition of locality cannot possibly reproduce all the statistical predictions of quantum mechanics.

Specifically, in Einstein-Podolsky-Rosen type experiments, quantum mechanics predicts a very strong correlation between measurements on the spatially separated parts. Bell showed that there is an upper limit on the strength of these correlations in the statistical prediction of any local hidden variable theory. Bell's result can be put in the form of inequalities which must be satisfied by any local hidden variable theory but which may be violated by the statistical predictions of quantum mechanics under appropriate experimental conditions.

Local hidden variable experiments. Experiments performed under conditions in which the statistical predictions of quantum mechanics violate Bell's inequalities can test the entire class of local hidden variable theories. It should be emphasized that the quantum-mechanical predictions resulting from most conceivable systematic errors tend to yield a weaker correlation that does not violate the inequalities. However, all existing experiments have required supplementary assumptions regarding detector efficiencies. The experiments considered to be most definitive involve polarization correlation between two photons from either an atomic cascade or from parametric down conversion in a nonlinear crystal. Some experimental highlights in the history of those two-photon correlation experiments are (1) the 1972 experiment by S. Freedman and J. Clauser, which was the first experimental test; (2) the 1976 experiment by E. Fry and R. Thompson, which provided a dramatic improvement in the signal-to-noise ratio by using laser excitation, and also employed an anisotropic initial-state density matrix; (3) the 1982 experiment by A. Aspect, P. Grangier, and G. Roger, which introduced two-channel polarizers; (4) the subsequent 1982 experiment by Aspect, J. Dalibard, and Roger, which used time-varying polarizers and was the first effort to enforce Einstein locality; (5) the experiements in 1987 by Y. H. Shih and C. A. Alley and in 1988 by Z. Y. Ou and L. Mandel, which were the first to use parametric down conversion to generate the photon pair; (6) the 1995 experiment by the groups of A. Zeilinger and Y. Shih, which introduced a new type II down conversion source and used all four possible Bell states for experimental tests; (7) the 1998 experiment by the group of A. Zeilinger, which was the first to rigorously enforce Einstein locality; and (8) the 1998 experiment by the group of N. Gisin, which required the lifetime of the entangled state (for example, $|\Psi\rangle$ above) to be greater than 30 μs, as compared to a few nanoseconds in previous experiments.

Nevertheless, due to the supplementary assumptions, small loopholes still remain, and experiments have been proposed to eliminate them. One, an exact experimental realization of Bohm's version of the Einstein-Podolsky-Rosen thought experiment, has been undertaken by the group of E. Fry and Th. Walther. It involves massive, spin-$^1/_2$ nuclei (fermions) of mercury-199 rather than massless photons (bosons), and requires the entangled-state lifetime to exceed several milliseconds.

In conclusion, the overwhelming experimental evidence is against any theory that would supplement quantum mechanics with hidden variables and still retain the locality condition; that is, any hidden variable theory that reproduces all the statistical predictions of quantum mechanics must be nonlocal. The remarkable Einstein-Podolsky-Rosen correlations have defied any reasonable classical kind of explanation. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS. Edward S. Fry

Bibliography. F. J. Belinfante, *A Survey of Hidden-Variables Theories*, 1973; J. F. Clauser and A. Shimony, Bell's theorem: Experimental tests and implications, *Rep. Prog. Phys.*, 41:1881–1927, 1978; B. D'Espagnat, The quantum theory and reality, *Sci. Amer.*, 241:158–181, 1979; E. S. Fry and Th. Walther, Fundamental tests of quantum mechanics, *Adv. Atom. Mol. Phy.*, 42:1–27, 2000; E. S. Fry, Th. Walther, and S. Li, Proposal for a loophole free test of the Bell inequalities, *Phys. Rev.*, 52:4381–4395, 1995; N. D. Mermin, Is the Moon there when nobody looks? Reality and the quantum theory, *Phys. Today*, 38(4):38–47, April 1985; J. A. Wheeler and W. H. Zurek (eds.), *Quantum Theory and Measurement*, 1983.

Higgs boson

A hypothetical massive scalar elementary particle, the avatar (embodiment) of electroweak symmetry breaking in the Glashow–Weinberg–Salam theory. Interactions with the Higgs boson endow the quarks, leptons, and weak gauge bosons with mass. Although experiments provide indirect evidence for the influence of the Higgs boson, no direct observation has yet been made (as of 2006).

Since the mid-1990s, experiments in laboratories around the world have elevated the electroweak theory to a law of nature that holds over a remarkable range of distances, from the subnuclear (10^{-19} m) to the galactic (10^{20} m). The electroweak theory offers a new conception of two of nature's fundamental interactions, ascribing them to a common underlying symmetry principle. It joins electromagnetism with the weak interactions—which govern radioactivity and the energy output of the Sun—in a single quantum field theory. Dozens of measurements have tested and confirmed the agreement between theory and experiment at the level of one part in a thousand. *See* ELECTROWEAK INTERACTION; FUNDAMENTAL INTERACTIONS; STANDARD MODEL.

Weak interactions and electromagnetism. Weak interactions and electromagnetism are linked through symmetry, but their manifestations in the everyday world are very different. The influence of

electromagnetism extends to unlimited distances, while the influence of weak interactions is confined to dimensions smaller than an atomic nucleus, less than about 10^{-17} m. The range of an interaction in quantum theory is inversely proportional to the mass of the force particle, or messenger. Accordingly, the photon, the force carrier of electromagnetism, is massless, whereas the W and Z particles that carry the weak forces are heavyweights, with masses nearly a hundred times that of the proton, the nucleus of hydrogen. *See* ELECTROMAGNETISM; INTERMEDIATE VECTOR BOSON; PHOTON; QUANTUM ELECTRODYNAMICS; WEAK NUCLEAR INTERACTIONS.

The many successes of the electroweak theory and its expansive range of applicability are most impressive, but the present understanding of the theory is incomplete. We have not yet learned what differentiates electromagnetism from the weak interactions—what endows the W and Z with great masses while leaving the photon massless. It is commonplace in physics to find that the symmetries observed in the laws of nature are not manifest in the consequences of those laws. A liquid is a disordered collection of atoms or molecules held together by electromagnetism that looks the same from every vantage point, reflecting the fact that the laws of electromagnetism are indifferent to direction. A crystal is an ordered collection of the same atoms or molecules, held together by the same electromagnetism, but it does not look the same from every vantage point. Instead, a crystal displays ranks, files, and columns that single out preferred directions. The rotational symmetry of electromagnetism is hidden in the regular structure of a crystal. *See* CRYSTAL STRUCTURE; SYMMETRY BREAKING; SYMMETRY LAWS (PHYSICS).

Higgs condensation. The central challenge in particle physics today is to understand what hides the symmetry between the weak and electromagnetic interactions. The simplest guess goes back to theoretical work by Peter Higgs and others in the 1960s. According to this picture, the diversity of the everyday world is the consequence of a vacuum state that prefers not a particular direction in space but a particular particle composition. The vacuum of quantum field theory—and of the world—is a chaotic confusion of many kinds of virtual particles winking in and out of existence. The laws of physics do not change if we interchange the identities of different particles. At high temperatures, as in the disordered liquid, the symmetry of the laws of physics is manifest in the egalitarian throng of particles. But at low temperatures, as in the ordered crystal, one kind of particle is preferred, and condenses out in great numbers, hiding the symmetry. *See* QUANTUM FIELD THEORY.

The condensate of Higgs particles can be pictured as a viscous medium that selectively resists the motion of other particles through it. In the electroweak theory, the drag on the W and Z particles caused by their interactions with the Higgs condensate gives masses to the weak-force particles. Interactions with the condensate could also give rise to the masses of the constituent particles—quarks and leptons—that compose ordinary matter. Today's version of the electroweak theory shows how this could come about, but does not predict what the mass of the electron, the top quark, or other particles should be. *See* LEPTON; QUARKS.

Experimental searches. If it were possible to heat up the vacuum enough, we could see the symmetry restored: all particles would become massless and interchangeable. For the present, it is beyond human means to heat even a small volume of space to the energy of 1 TeV (10^{12} electronvolts)—the temperature of 10^{16} kelvins. However, there is hope of succeeding on a smaller scale: it is possible to excite the Higgs condensate and see how it responds. The minimal quantum-world response is the Higgs boson: an electrically neutral particle with zero spin.

Although the Higgs boson has not been observed, experiments have begun to offer some evidence for its presence, beyond the shadowy role as giver of mass for which it was conceived. Experiments carried out over the past decade, using the Large Electron-Positron Collider (LEP) at CERN (European Laboratory for Particle Physics) and other instruments, are sensitive not just to the structure of the electroweak theory but also to quantum corrections. Long before the top quark was discovered at Fermilab in 1995, precision measurements detected its virtual quantum effects and anticipated its extraordinarily great mass. The influence of the Higgs particle is subtle, but increasingly incisive experiments strongly hint that a Higgs boson with mass less than about 200 GeV (1 GeV = 10^9 electronvolts) is required by precision measurements. *See* PARTICLE ACCELERATOR.

The electroweak theory predicts that the Higgs boson will have a mass, but it does not predict what that mass should be. (Consistency arguments require that it weigh less than 1 TeV.) Nevertheless, enough is known about the Higgs boson's properties—about how it could be produced and how it would transform into lighter particles or decay (as do all elementary particles except stable particles like the electron)—to guide the search.

LEP searches. The most telling searches have been carried out in experiments investigating electron-positron annihilations at energies approaching 210 GeV at LEP. The quarry of the LEP experimenters was a Higgs boson produced in association with the Z. A Higgs boson accessible at LEP would decay into a b (bottom) quark and a b antiquark, and have a total width (inverse lifetime) smaller than 10 MeV. The **illustration** shows how the decay pattern of a standard-model Higgs boson depends on its mass.

In the final year of LEP experimentation (before the machine was shut down in November 2000 to make way for construction of the Large Hadron [LHC]), accelerator scientists and experimenters made heroic efforts to stretch their discovery reach to the highest Higgs mass possible. At the highest energies explored, a few tantalizing four-jet events showed the earmarks of Higgs + Z production. A statistical analysis shows a slight preference for a

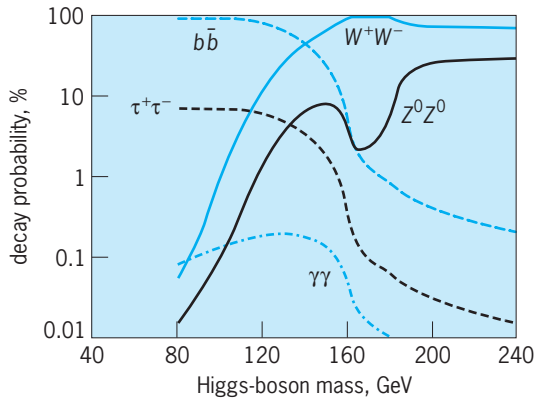

Dependence of the decay probabilities of a standard-model Higgs boson on its mass.

Higgs-boson mass near 116 GeV, but not enough to establish a discovery. Time will tell whether the LEP events were Higgs bosons or merely background events. The LEP observations do place an experimental lower bound on the Higgs-boson mass: the standard-model Higgs particle must weigh more than about 114 GeV.

Prospects. In 2006, experiments at Fermilab's Tevatron, where 1-TeV protons collide with 1-TeV antiprotons, were poised to extend the search, looking for Higgs + W or Higgs + Z production. The best hope for the discovery of the Higgs boson lies in experiments at the LHC at CERN, where 7-TeV beams of protons are designed to collide head-on. The LHC is scheduled to be commissioned in 2007, and it should make possible the study of collisions among quarks at energies approaching 1 TeV. A thorough exploration of the 1-TeV energy scale should determine the mechanism by which the electroweak symmetry is hidden and reveal what makes the W and Z particles massive.

Further questions. Once a particle that seems to be the Higgs boson is discovered, a host of new questions will come into play. Is there one Higgs boson or several? Is the Higgs boson the giver of mass not only to the weak W and Z bosons but also to the quarks and leptons? How does the Higgs boson interact with itself? What determines the mass of the Higgs boson? To explore the new land of the Higgs boson in more ways than the LHC can do alone, physicists are planning a TeV linear electron-positron collider.

Origins of mass. While the Higgs boson may explain the masses of the W and Z, and perhaps the masses of the quarks and leptons, it is not accurate to say that the Higgs boson is responsible for all mass. Visible matter is made up largely of protons and neutrons, and their masses reflect the energy stored up in the strong force that binds three quarks together in a small space. The origins of mass suffuse everything in the world around us, for mass determines the range of forces and sets the scale of all the structures we see in nature.

Extensions of electroweak theory. The present inability to predict the mass of the Higgs boson is one of the reasons many physicists believe that the standard electroweak theory needs to be extended. The search for the Higgs boson is also a search for extensions that make the electroweak theory more coherent and more predictive. Supersymmetry, which entails several Higgs bosons, associates new particles with all the known quarks, leptons, and force particles. Dynamical symmetry breaking interprets the Higgs boson as a composite particle whose properties we may hope to compute once its constituents and their interactions are understood. These ideas and more will be put to the test as experiments explore the 1-TeV scale. *See* ELEMENTARY PARTICLE; SUPERSYMMETRY.

Chris Quigg

Bibliography. F. Close, *Lucifer's Legacy: The Meaning of Asymmetry*, Oxford University Press, 2000; M. Kado and C. Tully, The searches for Higgs bosons at LEP, *Annu. Rev. Nucl. Part. Sci.*, 52:65–113, 2002; G.'t Hooft, Nobel Lecture: A confrontation with infinity, *Rev. Mod. Phys.*, 72:333–339, 2000; M. J. G. Veltman, Nobel Lecture: From weak interactions to gravitation, *Rev. Mod. Phys.*, 72:341–349, 2000; F. Wilczek, Masses and molasses, *New Scientist*, 162(2181):32–37, April 10, 1999.

High magnetic fields

Magnetic fields that are large enough to significantly alter the properties of objects that are placed in them. Valuable research is conducted at high magnetic fields. *See* MAGNETISM.

High-field magnets. Interest in the production of high magnetic fields, driven primarily by opportunities for new fundamental and applied science, extends back more than a hundred years. Until the late 1940s and early 1950s, magnets were confined to fields below 2 tesla due to the saturation of the iron cores used in laboratory magnets and the problems in cooling the conventional wire-wound coils. The attainable magnetic fields were extended more than 20 times by (1) the development of superconducting wire, fabricated from type II superconductors, which is capable of carrying high currents, and (2) the development of a new resistive magnet design by Francis Bitter using coils made of copper plates with the cooling water flowing through the plates themselves. Superconducting magnets have made fields up to about 15 T accessible to many laboratories, while higher-field magnets are restricted to a small number of facilities of 20 T. *See* SUPERCONDUCTING DEVICES.

Research and development efforts in magnets and magnet materials have led to gradual increases in the fields available for scientific research to fields near 20 T from superconducting magnets, 33 T in copper-core (resistive) magnets, and 45 T for hybrid magnets, which combine a superconducting outer core and a copper inner core. Superconducting magnets, which are used not only for research but also for magnetic resonance imaging (MRI), have the advantage that they use no electrical power once the field is established and the temperature is maintained at liquid-helium temperatures of 4.2 K (−452°F) or below.

The disadvantage is that there is a critical magnetic field, H_{c2}, determined by the type of conductor, that limits the attainable field to about 22 T in superconducting materials currently available. It is anticipated that new high-temperature superconductors may exceed this limit by 50% or more. Resistive magnets have no such limit, but are limited by the strength of the copper composite plates and the ability to cool them. The disadvantage is that they consume enormous amounts of power, in the 12–30-MW range. This limitation, plus the fact that they are very expensive to build and operate, has confined them to a few central facilities worldwide. *See* MEDICAL IMAGING; SUPERCONDUCTIVITY.

Another class of resistive magnets is pulsed magnets, in which the field is maintained for only brief periods of time. Advanced pulsed magnets that are not self-destructing provide fields beyond 70 T for about 0.1 s. Pulsed magnets using explosive magnetic flux compression have achieved fields above 500 T for periods of 10 microseconds. While the experiments that can be performed in the restricted times of the pulse are limited, much important research has been carried out with these magnets. *See* MAGNET.

Materials research. Research at very high magnetic fields spans a wide spectrum of experimental techniques for studies of materials. These techniques include nuclear magnetic resonance (NMR) in biological molecules utilizing the highest-field superconducting magnets, while the resistive magnet research is primarily in the investigation of semiconducting, magnetic, superconducting, and low-dimensional conducting materials. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

Semiconductors. Much of the progress in semiconductor physics and technology has come from high-field studies. For example, standard techniques for mapping the allowed electronic states (the Fermi surface) of semiconductors and metals are to measure the resistance (in the Shubnikov–de Haas effect) or magnetic susceptibility (in the de Haas–van Alphen effect) as a function of magnetic field and to observe the oscillatory behavior arising from the Landau levels of the electron orbits. Measurements at low fields are limited to low impurity concentrations since the orbits are large and impurity scattering wipes out the oscillations. At high fields of 20–200 T, the orbits are smaller, and higher impurity concentrations (higher carrier concentrations) have been studied. *See* DE HAAS-VAN ALPHEN EFFECT; FERMI SURFACE; SEMICONDUCTOR.

In two-dimensional semiconductor sandwiches (quantum wells), quantum Hall effects are observed at very low temperature (a few millikelvin) and high field. In such samples the Hall voltage shows quantized steps as the field is increased. The integer quantum Hall effect was first observed by Klaus von Klitzing. Subsequently, fractional quantum Hall steps were observed by Daniel C. Tsui and Horst L. Störmer, and explained by Robert B. Laughlin. The fractional Hall effect studies in fields up to 30 T have led to the observation of a new quasiparticle, the composite fermion, consisting of an electron and two flux quanta bound together. Measurements at fields of 45 T and beyond are predicted to provide evidence for the existence of a quasiparticle consisting of an odd number of flux quanta bound to an electron, a composite boson, and to lead to connections between the quantum Hall effect and superconductivity. In addition to the new physics, the quantum Hall resistance is now the international resistance standard, and there are many potential applications in microelectronics. *See* ELECTRICAL UNITS AND STANDARDS; HALL EFFECT; RESISTANCE MEASUREMENT.

High-temperature superconductors. Another area in which very high magnetic fields have an important role is in high-temperature superconductors, which have great potential for high-field applications, from magnetic resonance imaging, to magnetically levitated trains, to basic science. In type II superconductors, the practical superconductors, nonsuperconducting quantized vortices penetrate the superconductor when a magnetic field is applied. The superconductor can be thought of as resembling Swiss cheese, with the vortices as the holes in the superconductor. The number of vortices increases linearly with increasing field until the vortices overlap and the superconductivity is destroyed throughout the superconductor at the critical magnetic fields (H_{c2}). Motion of vortices driven by magnetic field or current is an energy-loss process leading to a voltage across the superconductor. It is thus essential to understand the vortex dynamics and pinning mechanisms of superconductors for practical applications. Low-temperature superconductors, such as niobium-tin (Nb_3Sn), the most commonly used, have values of H_{c2} slightly above 20 T; in high-temperature superconducting materials, H_{c2} is several times greater. High-temperature superconductors also have the complexity of a vortex liquid phase in addition to the solid phase. The phase boundary is dependent on both field and temperature. Transport, specific heat, magnetization, and nuclear magnetic resonance studies up to 30 T showed that the phase boundary is independent of field above 9 T, an unexpected result. Further work at higher fields is necessary to clarify the behavior.

Magnetic materials. Magnetism is one of the most challenging subjects of condensed-matter physics, and studies at high magnetic fields have played an important role in advancing understanding of magnetic materials. For example, in many organic conductors the conduction electrons (or holes) are confined to one or two dimensions, leading to very rich magnetic phase diagrams. High-field phases above 20 T include spin-density waves, a modulation of the electron magnetic moments that can propagate through the crystal, modifying the conduction and magnetic properties. Magnetic-field-driven metal insulator transitions are another area of great interest. Studies of unusual metallic compounds of rare-earth and actinide elements with both localized electron magnetic moments and conduction electrons,

at fields at or above the interaction energy of the electrons and moments, have greatly enhanced the understanding of this interaction. *See* ACTINIDE ELEMENTS; ORGANIC CONDUCTOR; PHASE TRANSITIONS; RARE-EARTH ELEMENTS.

Magnetic levitation. Another area of interest is the magnetic levitation of diamagnetic materials (the most common materials). For a diamagnetic material, an upward force is generated that is proportional to the product of the magnetic field and the magnetic field gradient. When the levitation condition is satisfied at high magnetic fields, zero gravity is simulated. Uses include growth of higher-purity crystals (aligned materials, including biological molecules, for detailed structural studies, much cheaper to produce in high magnetic fields than in satellite environments), studies of the effects of gravity and magnetic field on plant growth, and studies of the complex assembly processes of small particles under zero gravity and high magnetic field conditions. *See* SPACE BIOLOGY; SPACE PROCESSING; WEIGHTLESSNESS. William G. Moulton

Bibliography. J. S. Brooks, J. E. Crow, and W. G. Moulton, Science opportunities at high magnetic fields, *J. Chem. Phys. Sol.*, 59:569–590, 1998; R.G. Clark (ed.), *Proceedings of the 5th International Symposium on Research at High Magnetic Fields*, North-Holland Publishers, 1997; Z. Fisk et al. (eds.), *Physical Phenomena at High Magnetic Fields II*, World Scientific Press, 1995; E. Manousakis et al. (eds.), *Physical Phenomena at High Magnetic Fields I*, Addison-Wesley, 1991; *National High Magnetic Field Laboratory*, H. Schneider-Muntau (ed.), *High Magnetic Fields: Applications, Generation, Materials*, World Scientific, 1997; M. Springford, L. J. Challis, and H. U. Karow (eds.), *The Scientific Case for a European Laboratory for 100 Tesla Science*, European Science Foundation, 1998.

High-pressure chemistry

Chemistry at very high pressures, arbitrarily chosen to be above 10,000 bars (10^9 pascals), and mainly concerned with solid and liquid states. A bar is 10^6 dynes/cm^2, or 1.0197 kg/cm^2, or 0.9869 atm, or 10^5 Pa. Multiples of the bar are the kilobar (1 kilobar = 10^3 bars = 10^8 Pa) and the megabar (1 megabar = 10^6 bars = 10^{11} Pa). At 25°C (77°F) and 10 kilobars (10^9 Pa), nearly all ordinary gases are liquid or solid, and only a few liquids are not frozen; thus most high-pressure chemistry involves either higher temperatures, at which chemical reactions can occur at appreciable rates, or studies of internal arrangements in solids.

Figure 1 illustrates the range of high pressures which exist in nature as well as those which can be attained in the laboratory. Three broad ranges of pressures are also shown. In the lowest range, from 1 bar (10^5 Pa) to about 10^5 bars (10^{10} Pa), normal low-pressure chemical behavior prevails, and only minor departures from the usual valence and coordination rules are found. However, as discussed later, many interesting changes in materials can be effected in this pressure range as atoms are forced into new bonding arrangements. In the second range, from 10^5 to 10^9 bars (10^{10} to 10^{14} Pa), the energy added by compression becomes comparable with chemical bond energies, so that outer-shell electronic orbits are distorted and atoms and molecules change in character. A general tendency toward more metallic behavior is observed as the electrons become less strongly fixed to particular atoms, and chemical bonds may be broken. In the third region, upward of about 10^9 bars (10^{14} Pa), the delocalization of electrons is extensive, and the material consists of a mixture of ions and electrons, so that chemical bonds are of little importance. The boundaries on these three pressure ranges are, of course, only approximate, and show some variation according to the temperature and the atoms involved.

Equipment. High-pressure chemical phenomena can be studied in a wide variety of types of equipment, depending on the pressure and temperature range and the object of the study. The highest laboratory pressures, several megabars, are achieved, albeit but for a few microseconds, by the accelerative forces generated by high explosives or high-velocity impacts in the so-called shock-wave techniques. *See* SHOCK WAVE.

Static pressures which may be exerted for minutes or hours have reached a maximum of about 400 kilobars (4×10^{10} Pa) between diamond faces on specially supported anvils of the type shown in **Fig. 2***a*, but the specimens are quite thin, with typical diameters of 1 mm, and the temperature range is limited. Larger specimens, 1 cm^3 (0.6 in.^{3}) or more in size, may be studied up to about 100 kilobars (10^{10} Pa) and 2000°C (3600°F) or higher in cylindrical, tetrahedral, or cubical apparatus of the types that are indicated in Fig. 2*a*, *c*, and *d*.

There is no theoretical upper limit to the static pressures that may be achieved, but practical limits are imposed by the magnitude of the forces required, the materials of construction available, and the stress gradients in them. The changes occurring in the compressed material may be monitored in place, for example, by optical, x-ray, or electrical techniques, for in many cases the material reverts to its original state upon release of pressure. Pressure apparatus is conveniently calibrated by observing definite changes in certain substances, for example, resistance changes in bismuth at about 25 kilobars (2.5×10^9 Pa) or in lead at 130 kilobars (1.3×10^{10} Pa). *See* HIGH-PRESSURE MINERAL SYNTHESIS.

High-pressure effects. The simplest effect of high pressure is the closer compression of atoms. The noble gases and alkali metals are quite compressible (potassium shrinks to half its original volume under a pressure of 100 kilobars or 10^{10} Pa), whereas most oxides and the stronger metals are considerably stiffer. However, at a pressure exceeding about 100 kilobars (10^{10} Pa), most of the easily compressed electronic clouds are tightened up, and the compressibilities of most substances

approach each other. Usually, only minor amounts of energy compared with chemical bond energies can be added by compression to 100 kilobars. Nevertheless, the atoms of the substance can thereby be forced much closer together than they would be by cooling to low temperatures. An immediate consequence is an increase in melting temperature with pressure for most substances, since nearly all solids, with the exception of unusual ones like ice and bismuth, expand when they melt, and the process of melting absorbs heat as the entropy of the substance increases. The thermodynamic relationship shown in Eq. (1) expresses the

$$\frac{dT}{dP} = \frac{T\,\Delta V}{\Delta H} \qquad (1)$$

change in melting temperature T with pressure P as a function of the volume increase ΔV and the heat absorbed in melting ΔH for a given quantity of the substance. Thus, pressures of 50–100 kilobars ($5–10 \times 10^9$ Pa) increase melting points about 100°C (212°F) for substances such as iron, whose ΔV of melting is small, to several hundred degrees for substances such as sodium chloride (NaCl), whose ΔV of melting is large. At 1 bar (10^5 Pa), NaCl melts in an iron crucible, but at 100 kilobars (10^{10} Pa) iron can be melted in NaCl crucible.

Substances that consist of large molecules are easily stiffened or frozen by high pressures. The mobility of the molecules is sharply decreased by a sort of interlocking and tangling effect; thus for the substance to be sheared, chemical bonds must be broken, a process which requires considerable energy. For example, ordinary oils become so stiff at a few dozen kilobars that they are useless for transmitting pressure, and a droplet of pressure-frozen oil is capable of denting a steel plate. This stiffening phenomenon limits the study of most reactions of organic molecules to low pressures because they are rather large and "freeze" easily, but are usually not stable enough to withstand the temperatures necessary for liquefaction or intermolecular reactions.

The thermodynamic relationship shown in Eq. (1) applies not only to melting but also to phase changes or internal structure changes in solids. In general, the higher-density forms are favored at high pressures, but since ΔH can be positive or negative, the effects of pressure and temperature may oppose or reinforce each other in determining the stability of a particular high-pressure phase. Phase changes between solids rarely run freely to follow the theoretical equilibrium line, but instead tend to be sluggish or exhibit a region of indifference. The stronger or more refractory the solid, or the greater the atomic displacements involved in the change, the broader the region of indifference. An outstanding example is diamond, which at room conditions persists at nearly 20 kilobars (2×10^9 Pa) out of its stability field. The existence of a region of indifference can hamper the study of some high-pressure phenomena, but on the other hand, it may permit the recovery of high-pressure phases for more detailed analysis at room conditions.

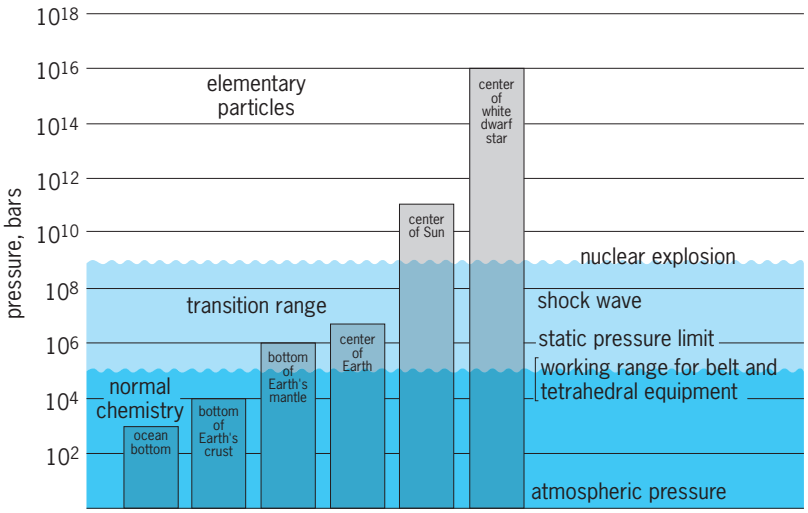

Fig. 1. Range of existing natural high pressures. 1 bar = 10^2 kPa.

The region of indifference may shrink drastically in the presence of solvents or catalysts that can promote the phase change. One of the more effective solvents, especially for systems containing oxides, is hot water. For chemically reducing systems such as carbon, the diamond-graphite transformations are assisted by the group VIII metals, iron, nickel, platinum, and so on. Other carbon solvents, such as silver chloride (AgCl) or cadmium oxide (CdO), exist but do not permit diamond to form, apparently because they do not favor carbon cations. For transformations in the boron nitride (BN) system, nitrides, either molten or in metallic solution, are effective catalysts. The rule "like dissolves like" is a useful, though not infallible, guide in these matters. *See* DIAMOND.

Some interesting chemical effects of pressure are related to shifts in chemical equilibria. The free-energy change ΔG determines chemical equilibria at temperature T through an expression of the

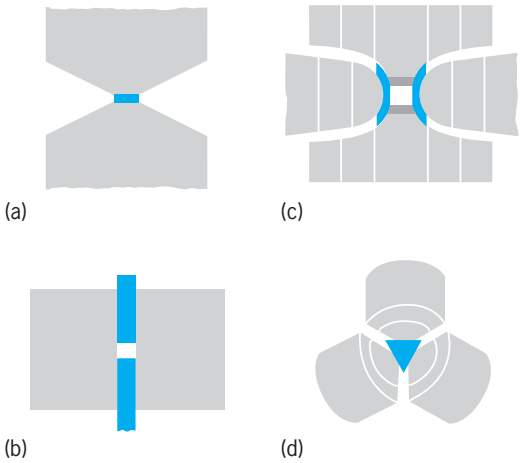

Fig. 2. Basic types of static high-pressure equipment. (*a*) Bridgman anvil. (*b*) Piston and cylinder. (*c*) Belt design. (*d*) Tetrahedral design.

form of Eq. (2), where R is the gas constant, and

$$\Delta G = RT \ \ln\left(\frac{a_m a_n \cdots}{a_x a_y \cdots}\right) \qquad (2)$$

$a_m, a_n, \ldots,$ are terms related to the concentrations of reactants and products in a reaction of the type $m + n + \cdots = x + y + \cdots$. For most reactions, ΔG ranges about 50 kcal (200 kilojoules) each side of zero. At 43 kilobars (4.3×10^9 Pa), a change in volume of 1 cm^3 (0.06 $in.^3$) corresponds to a change in energy of 10 kcal (40 kJ), and a change of 1 cm^3 (0.06 $in.^3$) per gram mole of reacting substance is not large, so that it is easy to find large shifts in chemical equilibria produced by pressure. Phase changes can also be regarded as special kinds of intramolecular reactions whose equilibria are shifted by pressure when interatomic bonds are broken and reformed.

Several rules exist regarding phase changes or electronic bonding changes in solids produced by pressure. The first rule recognizes that open structures stabilized by relatively weak ionic or van der Waals forces can easily collapse under pressure to denser structures. For example, eight crystalline forms of water ice have been identified. Potassium chloride (KCl) collapses from the NaCl structure to the more closely packed cesium chloride (CsCl) structure at about 25 kilobars (2.5×10^9 Pa). White phosphorus, made up of separate rings of four P atoms, is permanently pushed into the denser, more stable, semiconducting black form at pressures of a few tens of kilobars.

The second rule results from the compression of the outermost electronic shells and the delocalization of the electrons so that they are not as firmly fixed to particular atoms, and so that greater numbers of atoms can cluster around a given atom as the strongly directed valence forces are weakened. This is the same behavior observed as atomic number increases: Most of the lighter elements, with a few valence bonds per atom, are nonmetallic, whereas most of the heavier elements are metallic with higher valence numbers. Thus, pressure favors not only more metallic behavior but also behavior more like that of elements of higher atomic number. For example, upon compression to 100–130 kilobars ($1–1.3 \times 10^{10}$ Pa), silicon and germanium, normally semiconductors with the relatively open diamond structure, collapse to the white tin structure and become good metallic electrical conductors, equivalent to aluminum in that respect. As another example, silica, with four oxygen atoms about each silicon atom, changes at pressures of about 100 kilobars (10^{10} Pa) and moderate temperature in the presence of water to stishovite, which has the rutile structure of titanium oxide (TiO_2), with six oxygen atoms about each silicon. Many other examples of similar behavior in silicate systems are known wherein pressures of 100 kilobars or so force common crustal minerals into crystalline forms embodying higher coordination numbers, similar to those found in oxide compounds of heavier atoms. These effects are important in considerations of the nature of the deeper layers of the Earth's crust. *See* STISHOVITE.

The approach to the metallic state with increasing pressure may take different forms. Substances such as phosphorus, iodine, and selenium become substantially metallic at 100–150 kilobars ($1–1.5 \times 10^{10}$ Pa). On the other hand, when normally insulating organic compounds such as pentacene or hexacene, which consist of five or six aromatic rings fused together, are compressed to 150–200 kilobars ($1.5–2 \times 10^{10}$ Pa), they become semiconductors, since electrons are able to travel between the large molecules as the electronic clouds overlap. Certain complex compounds containing linear chains of metal atoms, such as Magnus's green salt, $Pt_2Cl_4(NH_4)_4$, increase in electrical conductivity up to 160 kilobars (1.6×10^{10} Pa) and then show a decrease; evidently, conductivity is favored only in particular ranges of interatomic spacing. Mössbauer studies have shown that ferric ion is reversibly reduced to ferrous ion in many compounds at pressures in the 100–200 kilobars ($1–2 \times 10^{10}$ Pa) range. Theoretical calculations indicate that even hydrogen could become metallic at pressures variously estimated to be 2–18 megabars ($2–18 \times 10^{11}$ Pa), but such pressures would be extremely difficult to generate and use, even with shock-wave techniques.

By changing the environment of atoms, high pressures can have strong effects on cooperative phenomena such as magnetism and superconductivity. For example, the high-pressure form of iron found above about 110 kilobars (1.1×10^{10} Pa) at 25°C (77°F) is not ferromagnetic. Superconducting transition temperatures are lowered by the application of high pressures to the metallic superconductors. Pressure studies guided the attainment of the ideal compositions of the high-temperature superconductors based on copper oxides reported in 1989, although pressure usually affects these superconductors adversely.

Many chemical reactions proceed through an intermediate state whose volume may be larger or smaller than that of the reactants. When the volume is smaller, the rate of reaction may be increased by pressure. Usually the intermediate state is more voluminous, especially in solids, and the rate of reaction is reduced by pressure. Here, as with chemical equilibria, the energy change due to pressure in the 30–100 kilobars ($3–10 \times 10^9$ Pa) range can be comparable with ordinary chemical bond or reaction-activation energies, and large pressure effects are possible. Certain reaction pathways may be effectively blocked so that new paths are followed. For example, at 130 kilobars (1.3×10^{10} Pa), where diamond is stable at temperatures up to 3000°C (5400°F) or more, the pyrolysis of some organic compounds, especially those consisting mainly of aromatic rings, produces graphite as the initial product whereas paraffin or polyethylene loses hydrogen to form waxy, dense solids of increasing microcrystalline diamond content as the pyrolysis temperature increases. A hexagonal form

of diamond can be prepared by subjecting highly crystalline graphite to pressure of 130 kilobars (1.3×10^{10} Pa) and 1500°C (2700°F). The reaction proceeds to some extent at 25°C (77°F) but proceeds much further on heating. (Poorly crystalline or partly amorphous graphite yields ordinary cubic diamond.) The hexagonal diamond can be recovered at room pressure and temperature if it has been heated at high pressure. Hexagonal graphitic BN can be converted to a wurtzite form by the application of 130 kilobars (1.3×10^{10} Pa) at 30°C (86°F), but at high temperatures, or in the presence of a molten catalyst-solvent, the cubic form of BN, which is slightly more stable than the wurtzite form, is obtained. Robert H. Wentorf, Jr.

Bibliography. W. Holzapfel and N. Isaacs (eds.), *High Pressure Techniques in Chemistry and Physics: A Practical Approach*, 1997; N. Isaacs, *Liquid Phase High Pressure Chemistry*, 1981; W. F. Sherman and A. A. Stadtmuller, *Experimental Techniques in High Pressure Research*, 1988; R. van Eldik and F.-G. Klärner (eds.), *High Pressure Chemistry: Synthetic, Mechanistic, and Supercritical Applications*, 2002.

High-pressure mineral synthesis

A laboratory technique for studying the behavior of minerals under high-pressure conditions.

The structure and nature of minerals vary with temperature and pressure. An example is ice, an important "mineral" in the polar regions of the Earth and in the interior of some planets: the solid form of water (H_2O), ice changes into a liquid and then into a gas with increasing temperature. Pressure dramatically alters the structure and nature of this material as well. When liquid H_2O is compressed at 300 K (80°F), it freezes at approximately 1 gigapascal (10 kilobars). This high-pressure form of ice (known as ice-VI) is denser than the coexisting liquid; thus ice-VI sinks into the bottom of water, while the density of ice at atmospheric pressure is lower than that of liquid water (**Fig. 1**). In ordinary ice, the H_2O molecule possesses hexagonal symmetry, which is manifested in the striking and beautiful shapes of snow crystals. In ice-VI, however, the H_2O molecules are arranged in a different manner, and the crystal shape is also quite different. With further compression of ice-VI to a pressure of 2 GPa (20 kilobars), another high-pressure phase forms, known as ice-VII, which has cubic symmetry. All of these forms of ice are colorless, but it is believed that the ice becomes opaque and "metallic" at ultimate high pressure.

The nature of minerals as they exist at atmospheric pressure represents only a very limited part of their real nature. The range of pressure and temperature prevailing at the surface of the Earth is very limited compared to the ranges that exist in the other planets of the solar system. The bottom of the ocean, which is at the highest pressure that can be observed directly, is only 0.1 GPa (1 kilobar), while

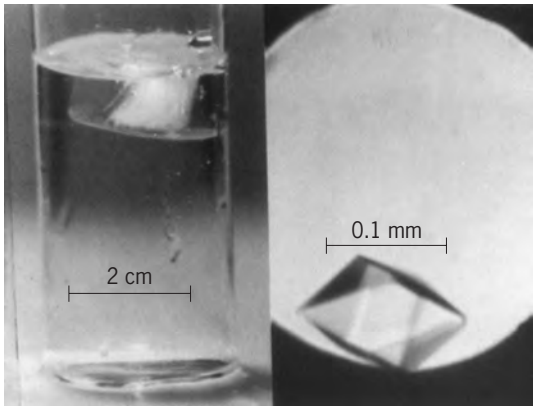

Fig. 1. Differing ice density. Normal ice floats on water while high-pressure ice formed at 1 GPa (10 kilobars) sinks to the bottom. (*From N. Miura, T. Yagi, and H. Ishimoto, The world of extreme, Kotai Butsuri (Japanese), 23:424–430, 1988*)

the pressure at the center of the Earth is 390 GPa (3900 kilobars). Pressures at the centers of large planets such as Saturn and Jupiter exceed 1000 GPa (10,000 kilobars). Therefore, to study the formation and structure of the Earth and other planets, it is essential to study the behavior of minerals under high pressure. It has become clear through high-pressure experiments that the minerals constituting the Earth's lower mantle (which extends from 650 to 2900 km or 400 to 1800 mi from the surface and occupies more than 50% of the entire volume of the Earth) are mostly so-called silicate perovskites that can never be formed on the surface of the Earth. *See* EARTH INTERIOR; JUPITER; SATURN.

Experiments and apparatus. Pressure is defined as a force per unit area; therefore, in order to apply a high pressure, it is necessary to concentrate a large force in a small area. Because of the limited strength of materials used to produce sample chambers, many different techniques are required, depending on the pressure range (**Fig. 2**).

Static compression. There is a specific type of equipment used to produce static high pressures. Mineral synthesis under conditions similar to those at the surface of the Earth is carried out by using a hydrothermal bomb apparatus. The starting materials are sealed in a high-pressure vessel (made of a hardened steel) together with a pressure-transmitting medium such as water or inert gas, and the whole vessel is heated in a large furnace. In this apparatus, the sample chamber is large (on the order of cubic centimeters), and the temperature throughout the sample is uniform, but the pressure is limited to about 1 GPa (10 kilobars).

The piston-cylinder arrangement is the high-pressure apparatus most commonly used to study the minerals in the Earth's crust. In this apparatus, the piston is subjected to the same pressure as that of the sample; it is difficult to achieve pressures higher than 3 GPa (30 kilobars), even when the piston and cylinder are fashioned from tungsten carbide, an exceptionally hard alloy. The multianvil apparatus was developed to overcome this difficulty.

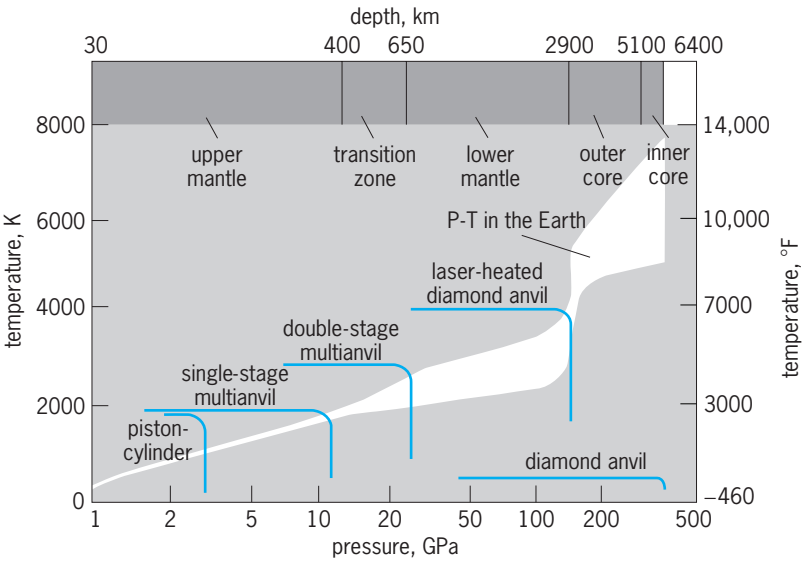

Fig. 2. Diagram of pressure (P) and temperature (T) within the Earth, showing capabilities of various types of high-pressure apparatus. 1 GPa = 10 kilobars. 1 km = 0.6 mi.

Very high pressure achieved on the top of the anvil is supported by a much larger area at the bottom. The upper portion of the anvil is also supported laterally, which increases the strength of the anvil. Consequently, pressures achieved by the apparatus are much higher than is possible with the simple piston-cylinder apparatus. However, when tungsten carbide is used for anvil material, 25 GPa (250 kilobars) is the practical upper limit of the pressure. Higher pressures have been achieved by using sintered diamond as the anvil material. In this multianvil apparatus, a sample and a small furnace are embedded in a solid pressure-transmitting medium, and the entire assembly is compressed. Typically, the sample chamber has a volume of several cubic millimeters, and it can hold a sample weighing several milligrams.

Although the amount of sample used is usually very small (typically a few micrograms), the diamond-anvil apparatus is often chosen for studying the behavior of minerals under a very wide range of pressures and temperatures. The principle of operation of the diamond anvil is very simple (**Fig. 3**). The sample and the pressure-transmitting medium are placed in a small hole (typically 0.3 mm or 0.012 in. in diameter and 0.1 mm or 0.004 in. deep) of a metal gasket and then squeezed between two flat faces of gem-quality diamonds. The metal gasket is deformed and high pressure is applied to the sample. Diamond is the hardest material known, and the highest pressure that has been achieved by this apparatus is more than 400 GPa (4000 kilobars), which is more than that at the center of the Earth. Although the size of the sample is very small, the development of many sophisticated experimental techniques, such as electron-probe microanalysis, very intense x-rays from synchrotron radiation, and Raman microanalysis, has made it possible to study the structure and the nature of minerals synthesized in such small quantities. The pressure in the sample chamber can be measured by using the wavelength shift with pressure of the ruby fluorescence line. Small chips of ruby are placed in a sample chamber and are irradiated by blue light from a laser; they emit red fluorescent light, and the pressure can be calculated precisely by measuring the wavelength of this light. *See* DIAMOND; ELECTRON SPECTROSCOPY; RAMAN EFFECT; SYNCHROTRON RADIATION.

The sample is heated by using strong laser light. Since diamond is transparent to visible and near-infrared light, the light from the laser can reach the sample through the diamond without much loss of energy. The energy of the light is absorbed and changed into heat either by the sample itself or by a dark material, such as graphite or platinum powder, that is mixed with the sample. Temperatures as high as 4000 K (6740°F) can be achieved, and formation of molten diamond has been reported. With this technique, it has become possible to produce in the laboratory conditions very similar to those believed to exist at the center of the Earth.

A disadvantage of the diamond-anvil apparatus is that while the heated area is very small (only 20 micrometers or less), a large temperature gradient occurs within the sample. In order to heat the whole area, it is necessary to scan the hot spot; thus, the thermal history of the sample becomes complicated. By contrast, the heating in multianvil apparatus is much more uniform, and a much larger sample can be treated. Therefore, the multianvil apparatus is a very powerful tool for studies of multicomponent systems or for synthesis of single crystals of high-pressure minerals.

Shock compression. In nature, shock compression caused by the impact of meteorites plays an

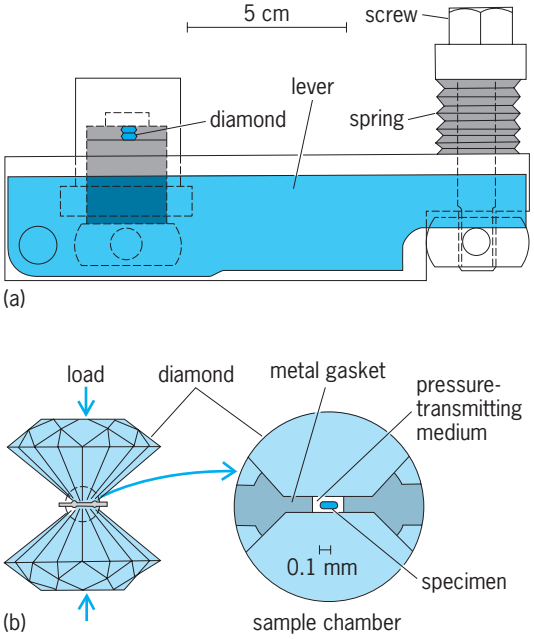

Fig. 3. Lever and spring type of diamond-anvil apparatus. (*a*) Cross section. (*b*) Expanded view, showing placement of the specimen. (*After T. Yagi, Diamond anvil and high-pressure experiments, Ouyou Butsuri (Japanese), 50:1337–1341, 1981*)

important role in the formation of high-pressure minerals. Shock compression experiments are carried out in the laboratory by using either explosions or collisions of projectiles into sample material. An example of this type of equipment is a single-stage gas gun that is used to study equations of state (pressure-volume-temperature relationships) or to synthesize high-pressure minerals. The projectile is accelerated to 4–5 km/s (2.4–3 mi/s) and collides with the sample chamber; a very high pressure is generated for a few microseconds. Usually, the sample is broken into small pieces after the impact; thus, proper design of the sample chamber is required to facilitate recovery of the sample. *See* METEORITE.

In-place observations. It has been found that in many minerals the phases formed under high pressure and temperature can be recovered at a pressure of 1 atm (10^5 pascals) if the temperature is reduced rapidly while the sample is still maintained under pressure; this is known as quenching the high-pressure phase. However, some high-pressure phases are known to undergo retrogressive transformation when the pressure is released. Many calcium-bearing silicates and iron oxides are in this category. To study these high-pressure minerals, it is necessary to carry out the observations while the sample is still under high pressure. Such in-place observations at high pressures become more and more important when the pressure range is expanded, because more minerals become unquenchable under these extreme conditions. X-ray and optical techniques provide the means for such in-place observations. By using very strong x-rays from synchrotron radiation or strong light from a laser, it has become possible to perform high-quality observations even in a pressure range of 100 GPa (1000 kilobars). *See* LASER.

Effects on structure. Minerals are composed of anions [such as oxygen (O) and fluorine (F)] and cations [such as silicon (Si), magnesium (Mg), and iron (Fe)]. The basic building unit of a mineral is known as a coordination polyhedron, in which a cation is surrounded by several equidistant anions. The number of surrounding anions is known as the coordination number. In silicate, the fundamental SiO_4 tetrahedron is usually the basic unit of the structure, and silicates are classified into several different groups depending on how these SiO_4 tetrahedra are connected with each other in the structure. The effects of pressure on mineral structure also depend on how the polyhedra are connected. *See* CRYSTAL STRUCTURE; SILICATE MINERALS.

When high pressure is applied to a silicate in which the SiO_4 tetrahedra are isolated, vacancies among the polyhedra shrink first, because the polyhedra are much "harder" than the vacancies. When the polyhedra are connected to each other, by sharing corners or edges, the crystal will tend to reduce the total volume by rotating the polyhedra. When rotation is no longer possible, the polyhedra become compressed. When the compression of the crystal exceeds a certain value, rearrangement of the atoms occurs, and the crystal structure is transformed; this is accompanied by a sharp decrease in volume. Sometimes, the coordination number increases in these pressure-induced phase transformations. The reason for this is that a large anion is usually more compressible than a small cation; and by increasing the pressure, the ratio of the radius of cation to that of the anion changes. For example, when silicon dioxide (SiO_2) and tin dioxide (SnO_2) transform from coesite into stishovite (the rutile-type structure) and from the rutile-type structure into the fluorite-type structure, the coordination numbers of Si and Sn increase from 4 to 6 and from 6 to 8, respectively. *See* COESITE; FLUORITE; RUTILE; STISHOVITE.

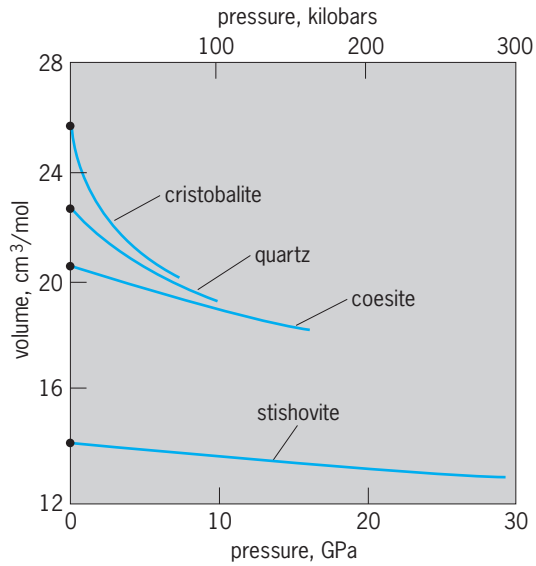

Fig. 4. Compression curves of various polymorphs of SiO_2. 1 cm^3 = 0.06 $in.^3$

Volume compression curves vary for various polymorphs of SiO_2 (**Fig. 4**). Cristobalite is a high-temperature polymorph of quartz. Silicate tetrahedra are connected by sharing oxygen at the corners, and the structure possesses many vacancies. The volume can be reduced by rotating tetrahedra; and consequently the compressibility of this structure is almost three times larger than that of quartz. When quartz transforms into coesite at about 3 GPa (30 kilobars), the density increases approximately 6%, but the coordination number of silicon remains unchanged. At about 10 GPa (100 kilobars), coesite transforms into stishovite with a density increase of 28%, and the coordination number of silicon is increased to six. This is a very close-packed structure, and the compressibility is very small. *See* QUARTZ.

A large number of phase transformations has been found in minerals under high pressure, but most of these structures have already been observed in other minerals existing under atmospheric pressure. For example, rutile-type SiO_2 (stishovite) is formed only above 10 GPa (100 kilobars), but the same structure is obtained at atmospheric pressure when the Si ion is replaced by the larger germanium (Ge) ion. This implies that crystal structure is determined mainly by the ratio of the cation radius to that of the anion.

When the very dense structure is compressed further, the bond length becomes shorter and shorter, and the orbitals of the electrons around the ions begin to overlap. This means that the orbital electrons can move freely in the material, which changes into a so-called metallic state. This metallic transition is believed to occur in all materials when they are subjected to high enough pressure. Even hydrogen, helium, and ice are believed to exist in the metallic state in the interiors of Jupiter and Saturn. In the laboratory, however, this transformation into the metallic state under pressure has been confirmed in only a limited number of materials such as Si, Ge, and gallium arsenide (GaAs). *See* BOND ANGLE AND DISTANCE; FREE-ELECTRON THEORY OF METALS.

The effects of shock pressure are very similar to those of static pressure. In the laboratory, however, because of the small size of the projectile, the duration of shock pressure is limited to the order of microseconds. When the synthesis of a high-pressure mineral requires rearrangement of atoms, the high-pressure state sometimes does not last long enough to accomplish phase transformation. In a naturally occurring impact of a meteorite, the size of the colliding material is much larger, and consequently the shock compression state lasts for a much longer time. This difference in time duration may explain the fact that many high-pressure minerals such as coesite and stishovite are found in natural craters while it is difficult to synthesize these minerals in the laboratory. Another important factor is that in shock compression it is almost impossible to control the temperature. When materials are compressed by shock waves, the temperature rises quickly and higher than that of the adiabatic compression curve. The cooling process is usually much slower, causing the retrogressive transition. Although shock compression is a very interesting and powerful means of synthesizing high-pressure minerals, the process is much more complicated than static compression. *See* METEORITE; SHOCK WAVE.

Transformations in mantle minerals. Several thousand minerals have been found to occur on the surface of the Earth, but the number of important minerals occurring in its deep interior is believed to be relatively small. The three major minerals that are believed to exist in the upper part of the mantle are olivine, pyroxene, and garnet. Many phase transformations are found to occur in these minerals up to 30 GPa (300 kilobars).

Olivine. Olivine is probably the most abundant mineral in the mantle, and the behavior of this material has been studied in detail. Based on the observations of the isomorphous structures of magnesium germanium oxide (Mg_2GeO_4), J. D. Bernal first pointed out the possibility that Mg_2SiO_4 olivine may transform into spinel structure in the Earth's mantle, and that the discontinuous density increase associated with this transition might be the cause of the seismic discontinuity in the mantle. The prediction of the olivine-spinel transformation in silicate was proved to be correct more than 30 years later by an experimental study on nickel silicate (Ni_2SiO_4). The coordination numbers of both silicon and nickel remain unchanged during this olivine-spinel transition; but because of the better packing of polyhedra the spinel structure is approximately 10% denser than olivine.

Natural olivine is a solid solution, and its composition is close to $(Mg_{0.9}Fe_{0.1})_2SiO_4$. Phase diagrams in the system Mg_2SiO_4-Fe_2SiO_4 have been studied in detail. With increasing magnesium content, the pressure from the olivine-to-spinel transition increases rapidly. When the magnesium content increases to more than 80%, a new phase appears that gives an x-ray diffraction pattern similar to that of spinel phase. The structure of this phase was elucidated by using an isostructural single-crystal cobalt silicate of (Co_2SiO_4) synthesized under high pressure; it was characterized as possessing a modified spinel structure. It was a unique structure never found in other minerals formed at atmospheric pressure, and it does not satisfy the Pauling's rule for the stability of minerals. Because of these observations, it was at first believed that this phase might not be thermodynamically stable but instead might be formed by the retrogressive transition from an unquenchable phase. This possibility was contradicted by the high-pressure in-place observation. Because of the existence of this phase, a sequence of the phase transformations expected to occur in the Earth became complicated. The properties of density and elasticity of this modified spinel structure are similar to those of spinel structure, and the major discontinuity at 400 km (250 mi) in the upper mantle can be explained by the transition from olivine to modified spinel. *See* MOHO (MOHOROVIČIĆ DISCONTINUITY); SOLID SOLUTION.

Many interpretations were advanced concerning the nature of postspinel transition; they included the transition to a potassium nickel fluoride (K_2NiF_4) structure and a strontium lead oxide (Sr_2PbO_4) structure, and the decomposition into MgO plus SiO_2. Experimental study confirmed that Mg_2SiO_4 spinel decomposes into a mixture of perovskite-type $MgSiO_3$ plus MgO at about 25 GPa (250 kilobars). In so-called ABX_3-type compounds, the most common structure at atmospheric pressure is that of perovskite. However, this structure is formed only when the larger cation (A) is similar in size to the X anion and the smaller cation (B) is large enough so that it can be sixfold-coordinated by X anions (**Fig. 5**) Therefore, under ambient conditions, this structure can be formed when cation A is very large, such as calcium (Ca), strontium, and barium (Ba), while smaller cation B is relatively large such as titanium (Ti), vanadium (V), and Fe. At atmospheric pressure, Mg and Si are far too small to satisfy these conditions. These silicate perovskites are stabilized only under very high pressure, where the oxygen anion is compressed much more than the cations and the ratio of the ionic radii has changed. An interesting observation is that when the large cation A is a transition-metal ion such as Fe or Co, this silicate perovskite is not formed but is decomposed into a mixture of a rocksalt-type AO compound plus a rutile-type BO_2 compound.

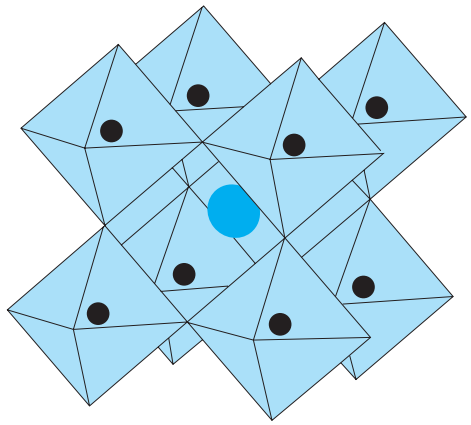

Fig. 5. Perovskite structure; the oxygen atoms (black circles) are on the corner of each octahedron, and a large cation (color circle) is surrounded by eight octahedra.

This is probably because the site for the cation A in perovskite structure has an unfavorable energy for such transition-metal ions. *See* OLIVINE; PEROVSKITE; SPINEL; TRANSITION ELEMENTS.

Pyroxene. Pyroxene is the second most abundant mineral in the Earth. The phase boundary between ortho- and clinopyroxene in $MgSiO_3$ is somewhat ambiguous; this is probably because the transition is extremely sensitive to the uniaxial component of the pressure. Above 18 GPa (180 kilobars) pyroxene decomposes into a modified spinel-type structure, Mg_2SiO_4 plus SiO_2. This modified spinel transforms into spinel structure at about 20 GPa (200 kilobars). Further compression forms a single phase of ilmenite structure, and this phase transforms into perovskite structure. At very high temperatures, a garnet phase is found to exist. *See* ILMENITE; PYROXENE.

Garnet. The first silicate perovskite was found in pyrope when this material was subjected to a pressure above 30 GPa (300 kilobars). The structure of this perovskite is believed to be complicated because of the existence of aluminum (Al) ion, and the structure has not been analyzed in detail. *See* GARNET.

Major minerals within the Earth. It has become clear that many of the major phases of silicate transform into the perovskite structure above 25 GPa (250 kilobars). Therefore it is believed that silicate perovskite is the most abundant mineral within the Earth, although it is an exotic mineral on the surface. In order to clarify the nature of the high-pressure minerals believed to be present in the interior of the Earth (**Fig. 6**) many studies have been made using various techniques. For this type of study, it is important to obtain a single crystal. The multianvil apparatus has been widely used for such experiments, and various single crystals of high-pressure minerals such as silicate perovskite, spinel, and stishovite have been synthesized.

Phase transformations in gases. Gases such as hydrogen, helium, and methane are important constituents in the outer planets. Since these gases are very compressible, it was difficult to study their behavior at high pressure in the laboratory. However, by using diamond-anvil apparatus it has become possible to liquefy these gases either by cooling or by applying pressures of a few thousand atmospheres and to load them into high-pressure sample chambers. All of these gases are found to solidify below 10 GPa (100 kilobars) at room temperature; and, as is found in other solids, many polymorphs are formed under very high pressure. In some materials, experiments are carried out up to 100 GPa (1000 kilobars), and the results of these studies provide important information for understanding the nature of the interiors of the planets.

Significance. A relatively narrow range for atmospheric pressure is an absolute requirement for the existence of most life forms on Earth, but this range is just a very small portion of the pressure range found in the universe. Therefore, high-pressure mineral synthesis is an indispensable tool for understanding the real nature of minerals and to study the interiors of planets. Development of experimental techniques has made it possible to perform such studies on various materials under practically the entire pressure range existing in the Earth. Many efforts are in progress to extend the range of pressures and temperatures for such experiments.

High-pressure synthesis is a powerful method not only for use in earth and planetary sciences but also for the creation of new materials. Many industrial diamonds are synthesized by using high-pressure techniques, and some new high-pressure materials, such as cubic boron nitride, have found wide application. Pressure is one of the most fundamental parameters that can alter the state of materials, and research in this field is also expected to expand in the future. *See* HIGH-PRESSURE CHEMISTRY; HIGH-PRESSURE PHYSICS; SILICATE PHASE EQUILIBRIA;

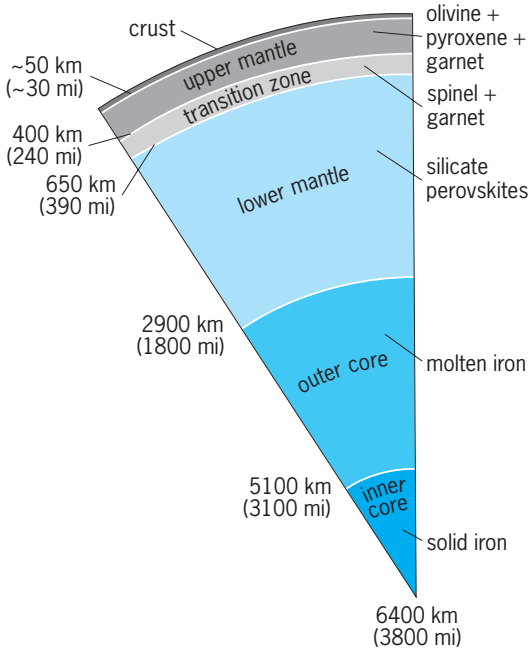

Fig. 6. Schematic representation of a section through the Earth, showing major minerals associated with each layer.

SOLID-STATE CHEMISTRY; SOLID-STATE PHYSICS; THERMODYNAMIC PROCESSES. Takehiko Yagi

Bibliography. H. Aoki, Y. Syono, and R. J. Hemley (eds.), *Physics Meets Mineralogy: Condensed Matter Physics in the Geosciences*, 2000; P. W. Bridgman, *The Physics of High Pressure*, 1949; L. G. Liu and W. A. Bassett, *Elements, Oxides, Silicates, High-Pressure Phases with Implications for the Earth's Interior*, 1986; M. H. Manghnani and T. Yagi, *Property of Earth and Planetary Materials at High Pressure and Temperature*, 1998; D. C. Rubie, T. S. Duffy, and E. Ohtani (eds.), *New Developments in High-Pressure Minerals Physics and Applications to the Earth's Interior*, 2005.

High-pressure physics

High-pressure physics is concerned with the effects of high pressure on the properties of matter. Since most properties of matter are modified by pressure, the field of high-pressure physics encompasses virtually all branches of physics. The justification for classifying high-pressure physics as a separate field is that rather specialized and often ingenious techniques are needed, both to produce high pressures and to make measurements of changes of physical properties of a material at high pressure. This field is therefore analogous to the fields of low- and high-temperature physics. Indeed, high-pressure experiments have been performed at temperatures approaching absolute zero and at temperatures as high as 9000°F (5000°C). *See* LOW-TEMPERATURE PHYSICS.

Pressure is defined as force per unit area; commonly used units of pressure include pounds per square inch (psi), kilograms per square centimeter (kg/cm^2), atmospheres (atm), bars, and pascals (Pa). The bar, equal to 10^6 dynes/cm^2 or about 14.5 lb/in.^{2} (that is, slightly less than 1 atm), is the basic unit that has been used by most researchers, although the approved unit of pressure in the International System (SI) of Units is now the pascal, which is equal to 10^{-5} bar. *See* PRESSURE.

The "high" of high-pressure physics connotes experimental difficulty. At liquid-helium temperatures, pressures of several hundred bars are considered high. In general, however, the high-pressure range may be arbitrarily regarded as extending from about 1 kbar (100 MPa or 14,500 lb/in.^{2}) upward to the present experimental limit. Prolonged static pressures in excess of 1 megabar (100 gigapascals or 1.45×10^7 lb/in.^{2}) can be achieved in very small samples weighing about 1 microgram. For perspective, a 100-micrometer-diameter (0.004-in.) grain of sand weighs about 1 μg. If simultaneous prolonged high temperatures in the range from 1800 to 3600°F (1000 to 2000°C) are required, the limit of static high pressures attainable is reduced to about 200 kilobars (20 GPa or 2.9×10^6 lb/in.^{2}). Additional experimental requirements, such as the need for nonmagnetic pressure vessels in magnetic resonance studies, further limit maximum attainable pressures.

Transient pressures as high as about 10^7 bars (1000 GPa or 1.45×10^8 lb/in.^{2}) have been attained in shock waves produced by high explosives or by projectile impact. This article is primarily concerned with static high-pressure experimentation, but it should be noted that the results of shock-wave experiments generally complement those of prolonged static high-pressure experiments. *See* SHOCK WAVE.

High-pressure effects. The major effects of high pressure on matter include diminution of volume, phase transitions, changes in electrical, optical, magnetic, and chemical properties, increases in viscosity of liquids, and increases in the strength of most solids. The magnitudes of these pressure effects may be illustrated by some examples, both general and specific. If a gas, initially at low pressure and a temperature below critical, is compressed, the first effect of pressure is to reduce the free space separating the atoms or molecules. At pressures on the order of several hundred bars, the volume will be about one-thousandth of the initial volume and a change of phase from gas to liquid will occur. Further compression of the liquid, by pressures up to about 50 kbar (5 GPa or 7.3×10^5 lb/in.^{2}), results in an additional volume decrement of only 20–50%. For liquids the greatest effect of pressure may be the increase in viscosity, which can be as high as a millionfold for a pressure increase of 10 kbar (1 GPa or 1.45×10^5 lb/in.^{2}).

For liquids that freeze to a solid phase of greater density than the liquid, the effect of pressure is to raise the freezing temperature. In the case of substances, such as water, having a solid phase of lower density than the liquid, the freezing point is lowered by pressure. The freezing point of water is lowered to −4°F (−20°C) by a pressure of about 2 kbar (0.2 GPa or 2.9×10^4 lb/in.^{2}). Above pressures of 2 kbar, however, water crystallizes in new solid forms which are denser than the liquid, and the freezing point is raised by further increases of pressure. At 45 kbar (4.5 GPa or 6.5×10^5 lb/in.^{2}) the freezing point of water is 374°F (190°C).

In general solids are less compressible than liquids, and the compressibility of both solids and liquids decreases with increasing pressure. However, solids and liquids show wide individual variations in compressibility. At a pressure of 200 kbar (20 GPa or 2.9×10^6 lb/in.^{2}) solid sodium is reduced to about half its initial volume, but diamond is reduced in volume by only a few percent at the same pressure. The electrical conductivity of metals is generally increased by pressure, but there are numerous exceptions to this rule. Changes in the electrical conductivity of metals are typically on the order of 10% for a 10-kbar (1-GPa or 1.45×10^5 lb/in.^{2}) pressure change. Phase transitions in metals generally result in discontinuities in the pressure-versus-resistance curves. These resistance discontinuities in various metals, for example, bismuth, iron, and lead, may be used as calibration points for high-pressure apparatus. *See* ELECTRICAL CONDUCTIVITY OF METALS.

At high pressure many solids exhibit polymorphic phase changes, that is, a rearrangement of the atoms or molecules in the solid. There are no universally applicable rules governing the number of phase changes or the kind of phase change to be expected at high pressure, but there is a thermodynamic requirement that the phase that is stable at high pressure must have a smaller volume than the phase that is stable at low pressure. Camphor has 11 probable solid phases, and seven solid phases of water are known to exist in the pressure range extending to 45 kbar (4.5 GPa or 6.5×10^5 lb/in.^{2}). On the other hand, many elements and compounds are known to exist in only one solid phase up to pressures of over 1000 kbar (100 GPa or 1.45×10^7 lb/in.^{2}). *See* THERMODYNAMIC PRINCIPLES.

Frequently, dramatic changes in physical properties result from phase changes. Ferromagnetic iron transforms to a paramagnetic form at pressures somewhat above 100 kbar (10 GPa or 1.45×10^6 lb/in.^{2}). In the same pressure range, the semiconducting element germanium transforms into a metallic phase that has an electrical conductivity greater than a million times that of the semiconductor. Similar semiconductor-to-metal transitions at high pressure have been observed in the cases of silicon, indium arsenide, gallium antimonide, indium phosphide, aluminum antimonide, and gallium arsenide. *See* SEMICONDUCTOR.

Many phases that form at high pressure transform back to low-pressure phases as the pressure is released. However, some high-pressure phases may be retained in a metastable condition at low pressures, and some low-pressure phases can persist metastably at high pressure. Diamond, the high-pressure form of carbon, is thermodynamically unstable at room temperature and pressures below about 12 kbar (1.2 GPa or 1.74×10^5 lb/in.^{2}). Nonetheless diamond persists indefinitely as a metastable phase at low temperatures; it transforms to the stable form, graphite, only when heated to temperature in excess of 1800°F (1000°C) at low pressure. Graphite can be exposed to pressures in excess of 200 kbar (20 GPa or 2.9×10^6 lb/in.^{2}) without transforming to diamond. Simultaneous high temperatures are required before diamond can form. Through use of a molten metal catalyst to lower the activation energy for the transformation, diamond is commercially synthesized at pressures of about 40 kbar (4 GPa or 5.8×10^5 lb/in.^{2}) and temperatures of about 2700°F (1500°C). Without the catalyst, pressures in excess of 100 kbar (10 GPa 1.45×10^6 lb/in.^{2}) and temperatures above 3600°F (2000°C) are required for diamond synthesis. *See* METASTABLE STATE.

An important area of research is the study of electronic transitions at high pressure. Mössbauer spectral measurements of iron compounds have demonstrated the reversible reduction of Fe^{3+} (ferric) ions to Fe^{2+} (ferrous) ions at high pressures. Similar reductions of Mn^{3+} to Mn^{2+} and of Cu^{2+} to Cu^{+} have been observed by optical spectroscopy at high pressures. This research is relevant to the problem of estimating the possible chemical composition of the interior of the Earth. Some researchers now hypothesize that only a negligible amount of ferric iron (Fe^{3+}) is present at depths below 600 mi (1000 km).

Experimental methods and problems. The simplest type of high-pressure apparatus is a thick-walled, hollow cylinder closed by sliding pistons to form a cavity. The sample is contained in the cylinder and force is applied to the pistons. Both the force and the change in volume can be measured by ordinary instruments. However, large corrections must be made for the pressure-induced distortion of the sample cavity and for the friction between pistons and cylinder. For liquids, the stresses will be hydrostatic, and the corrections for sample-cavity distortion may be calculated from elasticity theory. However, additional corrections, for example, for friction and for compression of the seals used to prevent leakage, must be determined empirically. Electrical leads may be introduced into the pressure cavity; the liquid may then serve as a pressure-transmitting fluid during measurements of the electrical properties of various materials. The electrical leads can also be connected to a coil of wire made from a metal or alloy having a well-known pressure variation of resistivity. The pressure in the liquid may then be determined independently by measuring the resistance change in the wire. Measurement of the compressibility of solid samples can best be made if the sample is immersed in a fluid of known compressibility.

At pressures above about 25 kbar (2.5 GPa or 3.6×10^5 lb/in.^{2}) at room temperature, however, most liquids have solidified, and solid pressure-transmitting media must be used. At these higher pressures, an important part of the experimental problem is to ensure that the nonhydrostatic, or shearing, components of stress are as small as possible in comparison with the hydrostatic component. The pressure-transmitting medium must be a soft solid, chosen empirically as one which does not become prohibitively stiff at high pressure.

The strength of materials presently available limits the simple piston-and-cylinder apparatus to a maximum pressure of about 50 kbar (5 GPa or 7.2×10^5 lb/in.^{2}). Apparatus designed for higher pressures takes advantage of the fact that the strength of many materials is greatly enhanced by pressure. Multistage apparatus has been built in which the sample is contained in a pressure vessel, which is in turn contained in a larger pressure vessel. Pressures up to about 90 kbar (9 GPa or 1.3×10^6 lb/in.^{2}) have been reached in two-stage apparatus, at the cost of greatly increased complexity of the apparatus and marked reduction in the accuracy with which measurements can be made. Through ingenious use of geometrical factors, the benefits of multistaging can be realized without having to construct concentric pressure vessels. However, the problem of measurement is enormously complicated, because only a small and imprecisely known fraction of the total force applied to the system actually serves to compress the sample. In some types of apparatus, over 98% of the total applied force is used to provide support for the highly stressed regions of the apparatus.

Temperature may be varied in high-pressure experiments by externally heating or cooling the entire pressure vessel. The temperature limit for externally heated pressure vessels is about 1800°F (1000°C), a limit determined by the reduction in strength of the vessel material at high temperatures. Higher temperatures may be obtained in internally heated pressure chambers in which heat is generated by passage of current through a resistance element; alternatively, the heat of a chemical reaction may be used. Temperature is measured by thermocouples within the pressure cavity, and suitable corrections must be made for the effect of pressure on thermocouple output. The effect of pressure on thermocouple output varies greatly according to the combination of metals or alloys chosen for the thermocouple junction. If two different thermocouple junctions are subjected to the same temperature and pressure, their outputs may be used to determine both temperature and pressure. *See* THERMOCOUPLE.

Sustained pressures in the megabar (100 GPa or 1.45×10^7 lb/in.^{2}) range are achieved in the diamond anvil apparatus. The sample cavity consists of a hole in a metal sheet compressed between two diamond surfaces. A typical sample cavity, having a diameter of 100 μm (0.002 in.) and a height of 50 μm (0.002 in.), can contain several ruby chips and a pressure-transmitting medium in addition to the material under investigation. The diamond anvils also serve as windows permitting direct optical observation of the sample and ruby chips. When ruby is illuminated by ultraviolet light, it fluoresces (emits light) of a characteristic wavelength that depends on pressure. This relationship between pressure and wavelength is often called the ruby pressure scale, and it is commonly used for determination of pressure in the diamond anvil apparatus. Because of their inertness and low shear strengths, solidified rare gases such as helium and neon are frequently used as pressure-transmitting media.

X-ray diffraction has become an important tool for studying volume changes and phase changes at high pressure. A portion of the pressure chamber must be constructed of a material of low atomic number, such as boron, beryllium, or diamond, to provide a "window" that is relatively transparent to x-rays. If an x-ray beam of suitable wavelength is passed through the sample, a portion of this beam is diffracted by the sample. The x-ray diffraction pattern, which may be recorded on film, is characteristic of the crystal structure and the atomic spacings of the material under investigation. If a calibrant (that is, a material that behaves in a known manner under pressure) is mixed with the specimen, the diffraction patterns of calibrant and specimen may be simultaneously recorded at high pressure. The diffraction pattern of the calibrant then serves as an internal standard of pressure. *See* X-RAY DIFFRACTION.

Previously, long exposure times (about a day) were required to record an x-ray diffraction pattern at a single pressure. Now many researchers transport their diamond anvil apparatus to very intense synchrotron x-ray sources where they can accomplish as much work in a week as would have taken a year or more with conventional x-ray sources. *See* SYNCHROTRON RADIATION.

Other ingenious techniques have been used for high-pressure studies of the Mössbauer effect, nuclear magnetic resonance, diffusion in solids, and many other effects of current interest in solid-state physics research. The primary problem of high-pressure physics continues to be accuracy of measurement.

No article on high-pressure research would be complete without mention of the late Percy W. Bridgman, who won the 1946 Nobel Prize in physics. For 40 years, beginning in 1908, Bridgman's work completely dominated the field of high-pressure physics. Virtually all high-pressure research depends heavily on apparatus designed by Bridgman and on his pioneering measurements of changes in physical properties at high pressure. *See* HIGH-PRESSURE MINERAL SYNTHESIS; HIGH-PRESSURE PROCESSES.

R. K. Linde; P. S. DeCarli

Bibliography. J. R. Asay and M. Shahinpoor (eds.), *High Pressure Shock Compression of Solids*, 1993; R. A. Graham, *Solids Under High Pressure Shock Compression*, 1992; R. M. Hazen, *The New Alchemists: Breaking Through the Barriers of High Pressure*, 1993; H. D. Hochheimer and R. D. Etters (eds.), *Frontiers of High Pressure Research*, 1992; C. Homan et al., *High Pressure in Science and Technology*, 3 pts., 1984; W. F. Sherman and A. Stadtmuller, *Experimental Techniques in High Pressure Research*, 1987.

High-pressure processes

Changes in the chemical or physical state of matter subjected to high pressure. The earliest high-pressure chemical process of commercial importance was the Haber synthesis of ammonia from hydrogen and nitrogen developed in Germany prior to World War I. The synthesis of diamonds from graphite developed in the early 1950s is a high-pressure physical process. Raising the pressure on a system may result in several kinds of change. It causes a gas or vapor to become a liquid, a liquid to become a solid, a solid to change from one molecular arrangement to another, and a gas to dissolve to a greater extent in a liquid or solid. These are physical changes. A chemical reaction under pressure may proceed in such a fashion that at equilibrium more of the product forms than at atmospheric pressure; it may also take place more rapidly under pressure; and it may proceed selectively, forming more of the desired product among multiple possible products.

Pressures higher than that of the atmosphere are expressed in bars and kilobars as well as in other units. A bar is 10^5 pascals, or 10^5 newtons per square meter, which are the units for pressure in the International System of units. These units are too small for convenient use in high-pressure processes, hence the bar is used. The bar equals 0.9869 standard atmosphere, 760 mmHg.

Physical processes. Increasing the pressure on a gas or vapor compresses it to a higher density and so to a smaller volume. If the pressure exceeds the vapor pressure, the vapor will condense to a liquid which occupies a still smaller volume. A vapor may be condensed at a higher temperature when it is under pressure; this permits the use of cooling water to remove the latent heat instead of more costly refrigeration.

Solids also change from a less dense phase to a more dense phase under the influence of increases in pressure. The density of diamond is about 1.6 times greater than that of graphite because of a change in the spatial arrangement of the carbon atoms. The temperatures and pressures used in the commercial synthesis of diamond range up to 5000°F (3000 K) and 100,000 atm. A molten metal is required as a catalyst to permit the atomic rearrangement to take place at economical rates of conversion. Metals such as tantalum, chromium, and iron form a film between graphite and diamond.

The highest static pressure attained in the laboratory is about 600,000 atm. At still higher pressures, the electrons are stripped from the atomic nuclei and matter loses its identity as recognizable atoms. This situation apparently exists in the centers of the white dwarf stars, where pressures of the order of 10^{16} atm prevail.

Chemical processes. In a manner similar to its effect during a physical change in which the volume of a system decreases, pressure also favors a chemical change where the volume of the products is less than the volume of the reactants. This is Le Chatelier's principle, which applies to systems in equilibrium. This general principle may be derived more precisely by thermodynamic reasoning, and thermodynamics is used to predict the effect of pressure on physical and chemical changes which lead to an equilibrium state.

Ammonia production. Ammonia is formed according to the reaction shown below. The ammonia content at

$$N_2 + 3H_2 \rightleftharpoons 2NH_3$$

equilibrium in a mixture which contains initially a ratio of 3 moles of hydrogen to 1 mole of nitrogen is shown in **Fig. 1**. At 1 atm only a fraction of 1% ammonia is formed. The ammonia content increases greatly when the pressure is raised. At 100 atm and 392°F (200°C) there would be about 80% ammonia at equilibrium. However, a very long time is required to form ammonia under these conditions, and consequently commercial processes operate at higher temperatures and pressures and use a catalyst to obtain higher rates of reaction. Many combinations of pressure and temperature have been used. The largest number of plants now operate in the region of 300 atm and 840–930°F (450–500°C). A higher pressure process is carried out at about 1000 atm and 930–1200°F (500–650°C).

A catalyst must be used or the reaction is too slow. Thus the process does not operate at 100 atm and 390°F (200°C) in spite of the favorable conversion

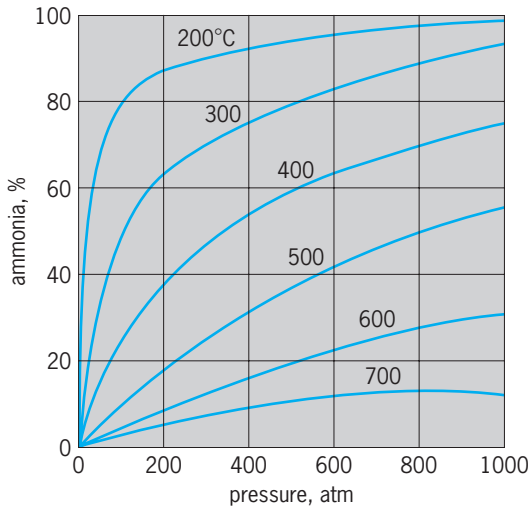

Fig. 1. Effect of combined temperature and pressure on the formation of ammonia. °F = (°C × 1.8) + 32; 1 atm = 1.01 bar. (*After E. W. Comings, High Pressure Technology, McGraw-Hill, 1956*)

at equilibrium because no catalyst has been found to provide an adequate rate of reaction under these conditions. Iron is the catalyst used at higher temperatures and pressures. The physical form of the catalyst is very important. Iron shavings or lumps are not effective. A suitable catalyst has been made from a fused mixture containing 66% Fe_2O_3, 31% FeO, 1.0% K_2O, and 1.8% Al_2O_3. The last two components are promoters and significantly increase the activity of the catalyst, but they alone are not catalysts. The mixture must not be contaminated with other impurities or the catalyst may be poisoned. Sulfur, phosphorus, and arsenic poison the catalyst permanently, whereas gases such as oxygen, water vapor, and carbon oxides reduce the catalyst's activity only while thay are in contact with it and are temporary poisons. Before the catalyst is used, the iron oxides are first reduced to iron by the hydrogen of the reacting mixture. The catalyst is activated by this treatment and must not be exposed to gases, which may contain poisons. Its life may be from several months to several years. Increasing the pressure increases the rate of reaction of the mixture in contact with a suitable catalyst.

A flow sheet for an ammonia process operating at 1000 atm is shown in **Fig. 2**. Nitrogen is obtained from the air, and hydrogen comes from natural gas. The process involves preparation of the nitrogen-hydrogen mixture, purification to exclude catalyst poisons, compression to high pressure, and circulation of the synthesis gas through a closed loop which contains two synthesis converters charged with catalyst. Because of the incomplete conversion in one converter, the unreacted gas is passed on to the next converter after ammonia has been condensed from the partially reacted gas mixture.

The temperature of the reaction is carefully controlled. Higher temperatures shorten the life of the catalyst and reduce the degree of conversion at equilibrium, whereas lower temperatures reduce the rate

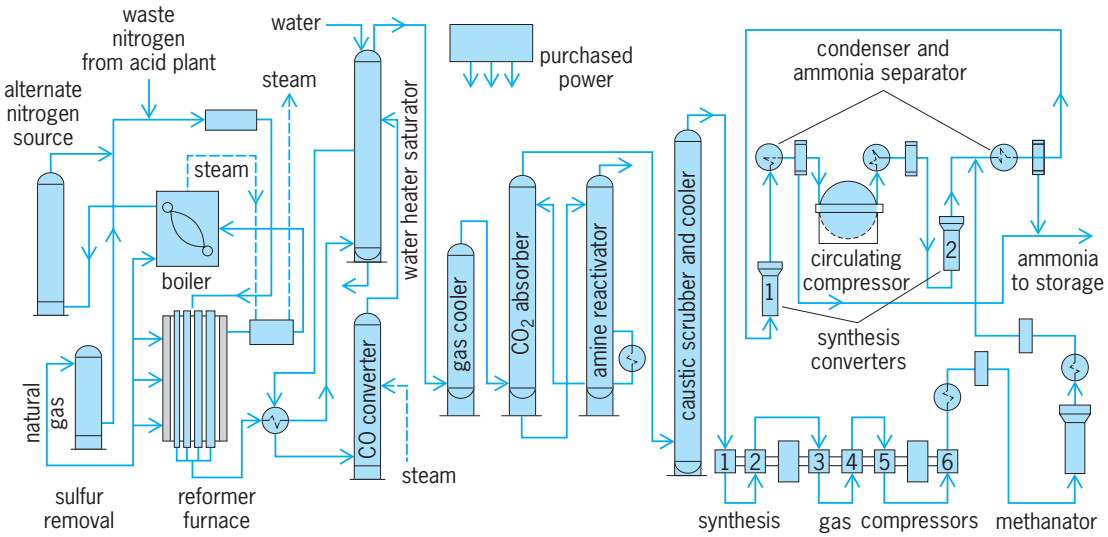

Fig. 2. Ammonia process at 1000 atm. 1 atm = 1.01 bar. (*After W. H. Shearon and H. L. Thompson, Ind. Eng. Chem., 44(2):254–264, 1953*)

of reaction. The reaction mixture entering the converter is relatively cool. It is heated to the reaction temperature inside the converter by heat exchange with the reacting gas and with the gas that has passed through the catalyst. The reaction is exothermic, and the converter is designed to remove some of the heat of reaction while the gas is passing through the catalyst.

The converters are large pressure vessels containing a heat exchanger and a catalyst basket. The converter for the 1000-atm process contains 16.5 ft^3 (0.5 m^3) of catalyst and produces 250–400 lb (110–180 kg) of ammonia per hour in each cubic foot of catalyst. This is comparable to 144 ft^3 (4 m^3) of catalyst in a converter for a 300-atm process, which produces 50 lb (22.5 kg) of ammonia per hour per cubic foot of catalyst. Molecular sieves (refractory compounds with controlled porosity) are coming into use as a base for catalysts to reduce the volume and increase the life of the catalysts. The high reaction temperature is confined to the catalyst basket inside the vessel, and the thick walls of the vessel are held at a lower temperature.

Methanol production. Methanol is synthesized from hydrogen and carbon monoxide at 200 atm and 600°F (315°C) in a similar manner. The catalyst contains aluminum oxide, zinc oxide, chromium oxide, and copper. Higher alcohols are produced at pressures of 200–1000 atm and temperatures up to 1000°F (538°C) with a similar catalyst to which potassium carbonate or chromate has been added.

Polyethylene production. Polyethylene has been produced at pressures in the ranges 3–4, 20–30, 40–60, and 1000–3000 atm. The latter is probably the highest pressure yet used in the commercial synthesis of an organic chemical product. The ethylene is polymerized in a stainless steel tubular reactor at 375°F (191°C) with small amounts of oxygen as a catalyst.

Phenol production. Phenol can be formed from chlorobenzene mixed with 18% sodium hydroxide solution at a pressure of 330 atm. Pressure is employed in this instance to maintain the mixture in the liquid phase at a temperature high enough for the hydrolysis reaction to proceed at an acceptable rate. A tubular reactor installed in a furnace provides for heating the mixture to 680°F (360°C).

Hydrogenation. Hydrocracking and hydrodesulfurization in the refining of gasoline and fuel oils are carried out at pressures up to 200 atm and temperatures of 800°F (427°C) and higher. The large volumes processed require very large reaction vessels, 8 ft (2.4 m) in diameter by 80 ft (24 m) long with 8-in.-thick (20-cm) walls, for example.

Other chemical reactions are carried out at elevated pressure. A reaction which has been observed to take place in the laboratory under high pressure may later be found to take place at lower pressure with a suitable catalyst. *See* HYDROCRACKING; HYDROGENATION.

Apparatus. High-pressure apparatus is carefully designed to provide for the measurement of pressure and temperature, a vessel with sufficient strength, a closure to prevent leakage, easy access to the interior, for the compression of gases and liquids at high pressures, and for the safety of people and equipment in case of ruptures or explosions. The deadweight gage is the principal primary gage for calibrating other pressure-measuring instruments. The deadweight gage consists of a free piston that is balanced between the force of gravity acting on weights at one end and the pressure of oil contained in the system at the other end.

P. W. Bridgman received the Nobel Prize in physics in 1946 for his pioneering work in the physics of high pressure. He was the originator of a self-sealing closure based on the unsupported area principle. This principle is illustrated in **Fig. 3**. The sealing gasket is initially under a low or moderate pressure. As pressure rises in the vessel, the pressure in the gasket is automatically maintained at a level higher than in

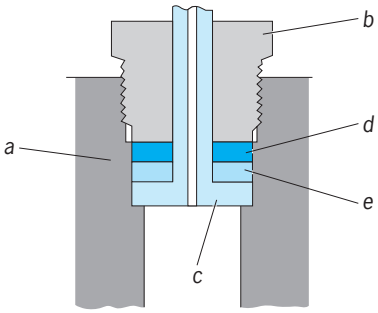

Fig. 3. Bridgman unsupported area gasket. Refer to the text for *a*–e. (*After E. W. Comings, High Pressure Technology, McGraw-Hill, 1956*)

the vessel. This ensures against leakage of the contents. Pressure, contained in the vessel *a*, acts upward on the steel disk *c*, which in turn acts on the gasket *e*. The gasket is supported by the steel ring *d*, which has a smaller area than *c* because of the unsupported area in the center. A balance of forces on the gasket requires that the forces acting downward balance those acting upward. The upward force is the product of the pressure in the vessel and the area of the lower face of *c*. This force is balanced by an equal force downward, which is the product of the pressure in the gasket and the area of the gasket. Because the area of the gasket is smaller than the area of the lower face of *c*, the pressure in the gasket must always be higher than the pressure in the vessel. *See* CHEMICAL EQUILIBRIUM; HIGH-PRESSURE CHEMISTRY; HIGH-PRESSURE PHYSICS; PHASE EQUILIBRIUM. Edward W. Comings

Bibliography. P. W. Bridgman, *The Physics of High Pressure*, reprint 1949; C. Homan et al., *High Pressure in Science and Technology*, 1984; K. Ishizaki, J. K. Tien, and E. Hodge (eds.), *Design Effects on Materials Processing and Design*, 1992; G. N. Peggs (ed.), *High Pressure Measurement Techniques*, 1983; I. L. Spain and J. Paauwe, *High Pressure Technology: Applications and Processes*, vol. 2, 1977; R. Van Eldik, *Inorganic High Pressure Chemistry: Kinetics and Mechanisms*, 1987; R. Van Eldik and J. Jonas (eds.), *High Pressure Chemistry and Biochemistry*, 1987.

High-temperature chemistry

The study of chemical phenomena occurring above 500 K (440°F). High temperatures represent one of the important variables available to scientists for increasing the variety of possible chemical reactions over that expected for classical ground-state atoms and molecules. One can enhance the relative population of excited rotational, vibrational, and electronic states by increasing the temperature and thus effectively create new species and new mechanisms for reaction. The potentialities of this approach are well illustrated by the three laws of high-temperature chemistry: (1) At high temperatures everything reacts with everything. (2) The higher the temperature, the faster the reaction. (3) The products may be anything. With an infinity of species available at high temperatures, the "golden age" of chemical synthesis is probably still in the future.

High temperatures also provide a common tie among the various options for energy production, conversion, or storage. For maximum thermodynamic efficiency, an energy production cycle should operate with a working fluid at as high a temperature as possible, and exhaust the spent fluid at as low a temperature as possible. Thus, in the combustion of coal to produce electric power or in the combustion of gasoline or diesel fuel to propel a car or an airplane, there is a need for materials of construction which allow operation of such devices at high temperatures. In the evaluation of new fuels or propellants, higher flame temperatures are among the desirable properties often sought.

It is convenient to discuss temperatures in terms of energy and to note that 11,500 K (20,200°F) corresponds to 1 electronvolt. In this sense, the particles emitted by radioactive nuclei or accelerated in cyclotrons and synchrotrons, which have energies in the keV, MeV, and BeV ranges, are effectively at temperatures of $\sim 10^7$ K, $\sim 10^{10}$ K, and $\sim 10^{13}$ K, respectively, and "high-energy physics" is synonymous with "ultrahigh temperature chemistry."

Traditional high-temperature chemistry has been mainly concerned with phenomena in the range of 500–3000 K (440–4900°F), although exotic flames can produce temperatures up to ~6000 K (10,300°F), shock waves can generate temperatures up to ~25,000 K (45,000°F), electric arcs can be operated in constricted modes to produce temperatures of ~50,000 K (90,000°F), and nuclear processes begin to occur at temperatures in the millions of degrees range. Laser excitation of selected energy states can produce species with effective temperatures in the range of 10^8 K. The major goals of high-temperature scientists include (1) the characterization of all important gaseous molecules, ions, and condensed phases—molecular formulas and structures, energy levels, thermodynamic properties, and the details of chemical bonding; (2) the establishment of reaction rate parameters and correlations with molecular properties; (3) the development of unique approaches to chemical syntheses and the preparation of new materials; and (4) the development of new techniques for generating or containing or utilizing high-temperature fluids in connection with energy production, conversion, and storage.

Gaseous species. Studies of thermal decomposition, vaporization, and/or sublimation of inorganic compounds have shown that complex gaseous species are much more common than usually realized. For example, condensed alkali metal sulfates vaporize appreciably as M_2SO_4 (gas); alkali metal carbonates vaporize as M_2CO_3 (gas); and even NH_4X (solid) yields NH_4X (gas). At high temperatures the relative fraction of complex species in the vapor actually increases with the increasing temperature for many systems. Vapors at high pressures and high temperatures can thus be complex. Rapid mass

spectrometric sampling techniques have demonstrated unequivocally the existence of polymers such as $(Ar)_m$, $(CO_2)_n$, and $(H_2O)_q$ where m, n, and q are 2,3,4, . . . , 8,9,10, ... As a further complication, various oxidation states can exist in high-temperature systems and thereby a still greater variety of possible molecular species. Typical systems which have been characterized by a combination of mass spectrometric, optical, infrared, nuclear magnetic resonance, and electron spin resonance spectroscopy and chemical studies are the difluorides of group 14 (CF_2, SiF_2, GeF_2, SnF_2, PbF_2) and their polymers.

In particular, the technique of generating high-temperature gaseous species and then reacting them with low-temperature molecules on a surface which is maintained at low temperatures (∼−190°C or −310°F) has been very productive. Practical approaches to a great variety of organometallic syntheses have been developed through the use of molecules such as SiF_2, C_2, C_3, BF, SiO, and SiS, and various atoms—such as C, Si, Li, Mg, Fe, Cu, Ni, Pd, and Pt. Matrix-isolation spectroscopy at 4–100 K (−452 to −279°F) provides the opportunity for gaining detailed information about atomic/molecular parameters—energy levels, frequencies, bond angles—which are necessary for calculation of thermodynamic functions and the prediction of chemical reactions.

Condensed systems. Condensed systems at high temperatures provide additional versatility in the chemical world since one fails to observe almost uniquely in high-temperature systems that atoms do combine in the ratio of small whole numbers. Such observations on gases led to the formulation of the law of definite proportions and other basic rules of early chemistry, but these rules no longer hold exactly for condensed systems at temperatures in the 1000–3000 K (1300–4900°F) range. Experiments have established melting points of refractory solids (binary carbides, borides, silicides, and so forth) at 1 atm (10^2 kPa) pressure in the 2000–4000 K (3100–6700°F) range, and unless one goes to higher pressures than 1 atm there really is no sold-state chemistry at temperatures greater than ∼4000 K (6700°F). Obviously, there is a chemistry of liquid systems which might be of some technological importance, and ingenious techniques for generating and maintaining reactive high-temperature liquids have been described.

Phase diagrams have demonstrated conclusively that one can prepare crystalline solids in an almost infinite number of compositions. For example, there are "pure" compounds fitting classical valences, and then one can substitute, as in $MgAl_2O_4$, where Al^{3+} ions can occupy some of the Mg^{2+} sites or Mg^{2+} ions can occupy some of the Al^{3+} sites so that there is an almost continual variation in properties. One can prepare Ta-O-C phases in which the gross stoichiometry is $Ta_1O_xC_{1-x}$ and, further, find it extremely difficult to detect the fact that oxygen is even present since the sizes of the oxygen and carbon atoms are so nearly the same in this lattice. Discovery of oxynitride phases and the establishment of essentially continuous solid solutions over wide ranges of compositions for many oxide, sulfide, carbide, boride, and other systems have been reported. The variations in compositions of alloys are similarly extremely complex; the variety possible is only beginning to be appreciated. Techniques for presenting phase diagrams have been described in which an attempt is made to present a single diagram which summarizes the behavior of whole classes of alloy systems rather than that of individual binary or ternary combinations. *See* NONSTOICHIOMETRIC COMPOUNDS.

Properties of high-temperature solids are, of course, not reliably interpreted until the exact stoichiometric and structural data have been obtained. The status of high-temperature thermodynamic properties is acceptably characterized by saying that no material is described thermodynamically to better than ±0.5% up to 1500 K (2200°F), or better than to ±1–2% from 1500–2000 K (2200–3100°F). There are practically no reliable data at temperatures greater than 2000 K (3100°F) except for a handful of basic materials—tungsten, molybdenum, tantalum, and aluminum oxide.

There is a great need for experimental techniques by which thermodynamic measurements can be reliably extended into this high-temperature region. There is already experimental evidence which is not explainable by current theoretical viewpoints on solids and liquids at high temperatures. Vaporization, sublimation, and other heterogeneous equilibrium studies have been made for many systems, and these are, of course, sensitive to the exact nature of the surfaces, to variations in stoichiometry as gases are evolved from a condensed phase, and to other factors which may accompany deviations from equilibrium.

Electrochemical techniques. One of the unique ways in which high-temperature thermodynamic data have been obtained involves the use of electrochemical techniques of the same sort as those which were widely applied during the 1920–1940 period for the establishment of reliable thermodynamic reference data for aqueous systems and for simple solids at temperatures near 25°C (77°F). High-precision free energies of formation for nonstoichiometric phases can be derived over fairly wide ranges of temperature. However, eventually at 1500 K (2200°F) and higher, one begins to have difficulty with vaporization, melting, and high diffusion rates.

Calorimetry. Calorimetry at temperatures up to 3000 K (4900°F) became routine through the use of electron bombardment heating and levitation heating. In electron bombardment heating, one boils electrons out of a thermionic emitter, such as thoriated tungsten, and accelerates them across a potential drop of a few thousand volts to strike the conducting sample and raise its temperature by means of the kinetic energy transferred and also by the heating caused by the electric resistance of the material. Levitation heating of conducting samples is accomplished with standard radio-frequency induction heaters and a pair of oppositely wound (left-handed and right-handed) coils. Samples weighing

up to 1000 g (35 oz) can be levitated, melted, and cast in a containerless, controlled atmosphere. Copper, gold, platinum, tantalum, graphite, and many other materials have been levitated and, in all cases except graphite, raised to temperatures above their melting points (3000 K or 4900°F and higher). Levitation calorimetry has provided heats of fusion and heat capacities for many metals, alloys, and conducting compounds in both solid and liquid states.

Chemical kinetics. From the viewpoint of heterogeneous kinetics, the high-temperature area has been studied extensively, with the practical concerns involving (1) the rates of interaction between various corrosive gases (oxygen, nitrogen, sulfides, halogens, and so on) and surfaces of various pure metals, alloys, and ceramic structural materials, and (2) the catalytic effects of solids of various stoichiometries on various gas reactions as in hydrocarbon refining, in the preparation of SO_3 or NH_3, or in the catalytic conversion of $CO + H_2$ to methanol, of CO to CO_2, or of NO_x to N_2 and O_2.

Homogeneous gas kinetics in the region above 1000 K (1300°F) has been mainly concerned with reactions of neutral atomic and molecular species which were stable at high temperatures, as in typical flames. The roles of H, O, OH, and of intermediate combustion products like HO_2, MO, and CHO have been established by direct mass-spectrometer probing of flames. The importance of electrons and atomic or molecular ions in flame kinetics has also been recognized. This knowledge is of interest in identifying the most efficient combustion systems for energy generation and in elucidating the chemical parameters which are crucial to the development of flame-retardant materials.

Another fertile area for high-temperature kinetics research has been the study of the reaction rates of atoms or molecules which have been produced by either photochemical, laser-pulse, or electric-arc excitation or by pulsed thermal dissociation processes and then allowed to react either in systems at relatively higher pressures by random collisions or in systems at low pressures by molecular-beam techniques. Thus, a hydrogen atom created from the H_2 molecule in an arc or by thermal dissociation requires 2.24 eV/atom, and therefore is a chemical species of the sort that might be expected in a very high-temperature system. The thermodynamic potential of such atoms for reaction is much greater than that for most molecular species, and reaction rates are usually measurable although relatively fast. Selective excitation to specific energy levels is possible with tunable dye lasers. *See* ULTRAFAST MOLECULAR PROCESSES.

Several types of monitoring techniques, rapidscan infrared spectroscopy, electron spin resonance spectroscopy, and mass spectrometry, are typically used in these studies. Gas reaction rates in mass spectrometers also have been widely explored by scientists interested in ion-molecule reactions. Of course, since ions are created endothermally by the impartation of several electronvolts of energy, they are high-temperature species. Many new types of high-temperature reactants can be created. It is possible that a synthetic chemistry making use of ion-molecule reactions may someday be of economic significance. Ion sputtering is a widely used technique for the preparation of electronic circuitry. *See* CHEMICAL DYNAMICS.

John L. Margrave

Bibliography. L. Eyring (ed.), *Advances in High Temperature Chemistry*, vols. 1–4, 1967–1975; J. W. Hastie, *High Temperature Vapors*, 1975; J. L. Margrave (ed.), *Modern High Temperature Science*, 1984; E. T. Turkdogan, *Physical Chemistry of High Temperature Technology*, 1980.

High-temperature materials

Materials that serve above about 1000°F (540°C). In the broad sense, high-temperature materials can be identified by the following classes of construction solids: stainless steel (limited), austenitic superalloys, refractory metals, ceramics and ceramic composites, metal-matrix composites, and graphitic composites. The first three classes are well proven in industrial use, although stainless steels serve but slightly above 1000°F (540°C) and refractory metals are usually limited to nonoxidizing atmospheric conditions. The other classes are under extensive worldwide research to establish whether they can be utilized to replace and extend the capabilities of austenitic superalloys, which are the mainstay of high-temperature service.

The most demanding applications for high-temperature materials are found in the aircraft jet engines, industrial gas turbines, and nuclear reactors. However, many furnaces, ductings, and electronic and lighting devices operate at such high temperatures. In order to perform successfully and economically at high temperatures, a material must have two essential characteristics: it must be strong, since increasing temperature tends to reduce strength; and it must have resistance to its environment, since oxidation and corrosion also increase with temperature.

High-temperature materials have acquired their importance because of the pressing need to provide society with energy and transportation. Machinery that produces electricity or some other form of power from a heat source operates according to a series of thermodynamic cycles, including the basic Carnot cycle and the Brayton cycle, where the efficiency of the device depends on the difference between its highest operating temperature and its lowest temperature. Thus, the greater this difference, the more efficient is the device—a result giving great impetus to create materials that operate at very high temperatures. Following is a discussion of the characteristics of some of the more significant alloys, a brief review of strengthening mechanisms, and a description of several developments. *See* CARNOT CYCLE; EFFICIENCY; MECHANICAL ENGINEERING.

Metallic materials designs. Alloys used at high temperatures in heat engines are composed of several elements. High-temperature metallurgists, using both

theoretical knowledge and application of empirical experimental techniques, depend upon three principal methods for developing and maintaining strength. The alloys are composed of grains of regular crystalline arrays of atoms and show usable ductility and toughness when one plane of atoms slips controllably over the next (dislocation movement); excessive dislocation movement leads to weak alloys. Thus, in terms of achieving mechanical strength, alloy design attempts to inhibit but not completely block dislocation movement.

The most common technique is solid solution strengthening. Foreign atoms of a size different from that of the parent group cause the crystal lattice to strain. This distortion impedes the tendency of the lattice to slip and thus increases strength. *See* ALLOY; CRYSTAL STRUCTURE.

A second significant mechanism is dispersion strengthening. This is the introduction into the alloy lattice of extremely hard and fine foreign phases such as carbide particles or oxide particles. Carbide dispersions can usually be created by a solid-state chemical reaction within the alloy to precipitate the particles, while oxide particles are best added mechanically, as described later. These hard particles impede slippage of the metal lattice simply by intercepting and locking the dislocations in place.

Still another strengthening technique is coherent-phase precipitation; a foreign phase component of similar but modestly differing crystalline structure is introduced into the alloy—always by a chemical solid-state precipitation mechanism. The phase develops significant binding strength with the mother alloy in which is resides (it is "coherent") and, like the other mechanisms, then impedes and controls dislocation flow through the metal lattice. Coherent phase precipitation is a special case within precipitation hardening.

However, certain advances in processing have had significant effects on these classical strengthening approaches in recent years, and they are covered separately below. *See* HEAT TREATMENT (METALLURGY); SOLID-STATE CHEMISTRY.

Metallic materials systems. These include superalloys, stainless steel, eutectic alloys, and intermetallic compounds.

Superalloys. High-temperature alloys also must be resistant to chemical attack by the atmosphere in which they exist, more often than not a high-speed high-pressure gas of highly oxidizing nature. Protection from the atmosphere is achieved by causing the material to develop a tough diffusion-resistant oxide film from various elements in the alloy, or by applying a protective coating to the material.

From all this, a high degree of success has been achieved with the superalloy class of materials. Superalloys serve under tough conditions of mechanical stressing and atmospheric attack to over 2000°F (1100°C), which often is within a few hundred degrees of their melting point. These alloys are capable of supporting modern aircraft jet engines (**Fig. 1**).

Undoubtedly, the most complex and sophisticated group of high-temperature alloys is the superalloys. A superalloy is defined as an alloy developed for elevated temperature service, where relatively severe mechanical stressing is encountered and high surface stability is frequently required. Superalloys, classically, are those utilized in the hottest parts of aircraft jet engines and industrial turbines; in fact, the demand of these technically sophisticated applications created the need for superalloys.

Superalloys are strengthened by all of the methods described above, as well as by some even more subtle ones, such as control of the boundaries between the grains of the crystalline metal by minor additives and other little-understood solid-state chemical reactions. **Figure 2** illustrates how these mechanisms combine to make nickel-base superalloys very strong. The superalloy René 77 (a nickel-base casting alloy used for turbine blades) is shown in two magnifications as a basis for illustrating some of these strengthening mechanisms. The continuous matrix of the alloy is the face-centered-cubic-phase austenite, referred to as γ. The most effective

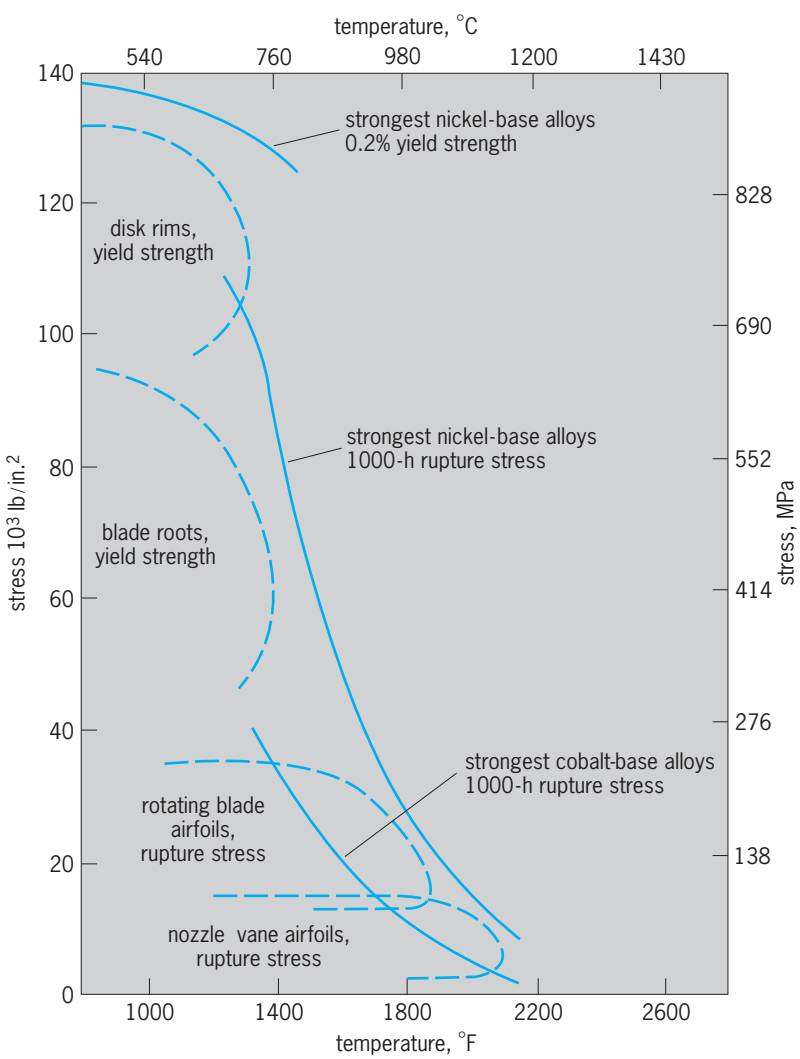

Fig. 1. Capabilities (broken lines) of superalloys utilized in critical parts of aircraft jet engines. (*After C. T. Sims, N. S. Stoloff, and W. C. Hagel, eds., Superalloys II, John Wiley and Sons, 1987*)

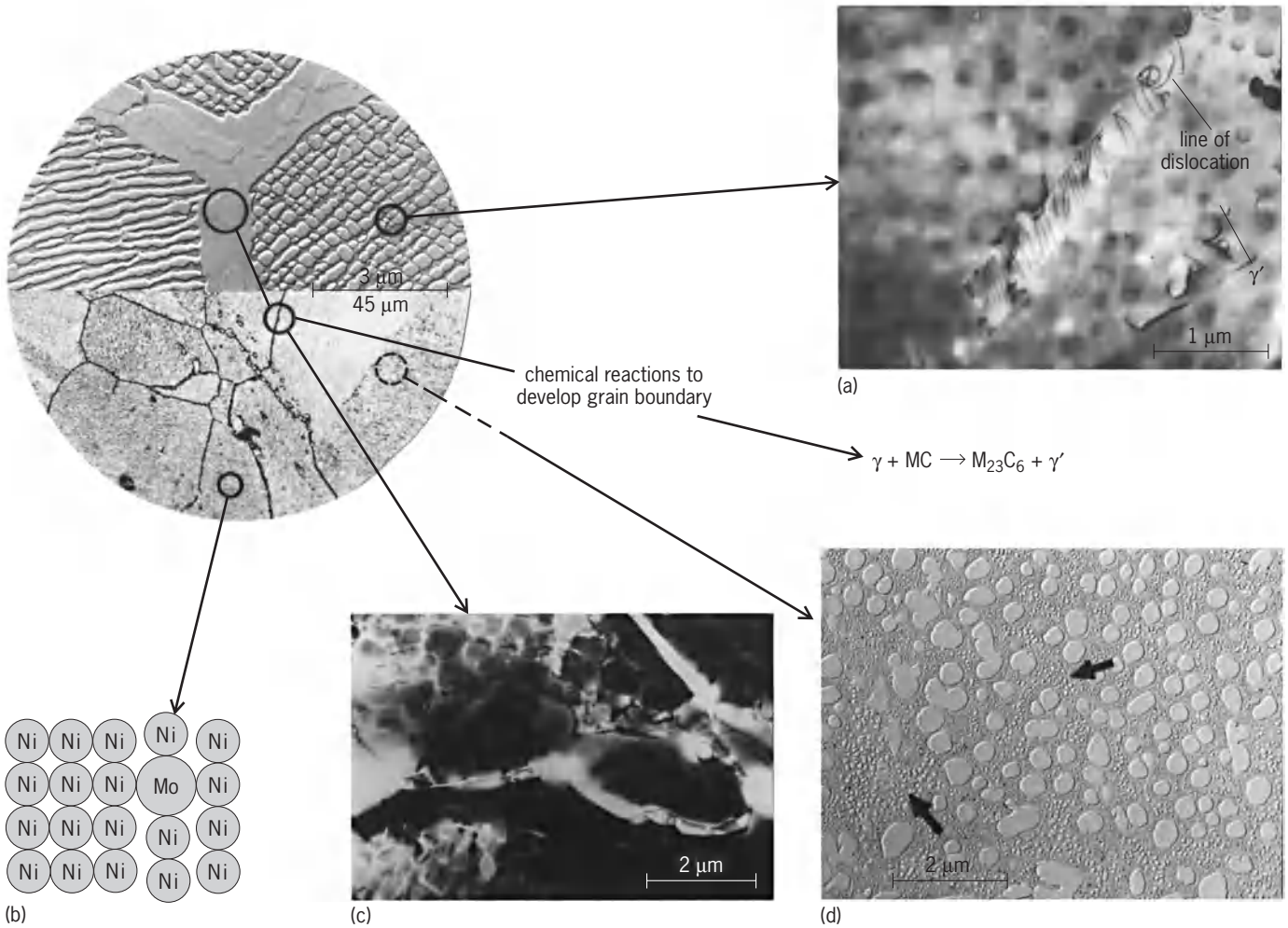

Fig. 2. Major features for achieving useful strength in nickel-base superalloys. (*a*) Dislocation interaction with γ', a coherent crystalline phase, shows dislocation movement. (*b*) Lattice after solid solution strengthening with molybdenum. (*c*) Carbide strengthening at grain boundaries. (*d*) Fine γ (arrows) for low-temperature strengthening; the big blobs are large γ'. (*From C. T. Sims and W. C. Hagel, eds., The Superalloys, Wiley-Interscience, 1973*)

strengthening effect comes from precipitation of γ', an intermetallic compound of nickel (Ni) and aluminum (Al), basic composition Ni_3Al. It is coherent and readily controlled by heat treatment. Chemical reactions between the γ and MC (a simple metallic carbide often present in superalloys) create $M_{23}C_6$ (a complex metallic carbide). These carbides, precipitated at boundaries between the grains, lock the grain boundaries in place and generate the dislocations. While these alloys were originally processed by melting in air, the large amounts and complex control of reactive metals needed eventually demanded processing in vacuum to maintain alloy cleanliness; these alloys contain as many as 15 different alloying additions.

Since the mid-1960s, advanced processing applied to superalloys has provided much additional capability. It was found that directional solidification could be utilized to reduce thermal fatigue failures and to create single-crystal components with still greater advantages. Dispersion of fine oxide particles in superalloys by powder metallurgy techniques also produced advantages.

Stainless steel. Stainless steels are strengthened principally by carbide precipitations and solid solution strengthening, since they cannot form the γ' coherent precipitate, their use is limited to about 1200°F (650°C), except where strength may not be needed. *See* STAINLESS STEEL.

Refractory alloys. Refractory metal alloys are based on elements that have extremely high melting points—greater than 3000°F (1650°C). These elements—tungsten, tantalum, molybdenum, and niobium (columbium)—are strengthened principally by a combination of solid solution strengthening, carbide precipitation, and unique metalworking. However, they cannot be used in modern heat engines because no commercially successful method of preventing their extensive reaction with oxidizing environments for broad application has been found. *See* REFRACTORY; SOLID SOLUTION.

Eutectic alloys. These metal-based systems have been under study since the late 1960s, but service applications have not been developed. They are usually similar to or based on superalloys, but they are processed by directional solidification so that a particularly

strong stable phase forms as a needlelike structure, giving unusual strength. However, they are expensive to process, and industrial use is questionable.

Intermetallic compounds. These are metal-based systems centered on the fixed atomic compositions occurring in metallic systems of aluminum with nickel, titanium (Ti) and niobium (Nb), such as Ni_3Al, Ti_3Al, TiAl, and Nb_3Al. Intermetallic compounds are of interest because they often possess a lattice arrangement of atoms that leads to higher melting points and less ease of deformation. Some have shown ductility and toughness potential, and alloying to optimize properties is under significant study. For instance, Ti_3Al is fabricable and ductile, particularly when alloyed with niobium. None are in gas turbine service, but there is a possibility that Ti_3Al will become acceptable if the danger of titanium combustion and hydrogen embrittlement can be avoided. *See* INTERMETALLIC COMPOUNDS.

Oxidation and corrosion. In addition to possessing strength, high-temperature alloys must resist chemical attack from the environment in which they serve. Most commonly this attack is characterized by a simple oxidation of the surface. However, in machines which utilize crude or residual oils or coal or its products for fuel, natural or acquired contaminants can cause severe and complex chemical attack. Involved reactions with sulfur, sodium, potassium, vanadium, and other elements that appear in these fuels can destroy high-temperature metals rapidly—sometimes in a matter of a few hours. This is known as hot corrosion, and it is a major problem facing otherwise well-suited alloys for service in turbines operating in the combustion products of coal. *See* CORROSION.

Some nuclear reactors operate above 1000°F (540°C), and the working fluid is often not an oxidizing gas. For instance, high-temperature gas-cooled reactors in the development stages utilize high-temperature, high-pressure helium as the working fluid. However, the helium inevitably contains very low levels of impurities—oxygen, carbon, hydrogen, and others—and it does not contain enough oxygen to form a protective oxide film on the metals which contain it. Impurities enter these alloys and reduce strength by precipitating excessive amounts of oxide and carbide phases and by other deleterious effects. Refractory metals, nominally of high interest here, have shown long-term degradation by carbide-related mechanisms, and so metallurgists have been working to develop resistant superalloys. *See* HELIUM.

In other reactor applications such as the liquid metal fast-breeder reactor, construction metals must resist the high-temperature liquid metal used to transfer heat from hot uranium-plutonium fuel. The liquid metal is usually sodium or potassium. All classes of high-temperature alloys—stainless steels, superalloys, and refractory metals—have been under evaluation for service. As in helium gas, small amounts of impurities are the critical item. They can react with the metal at one temperature and transfer it to a component operating at a lower temperature. Stainless steels remain the prime construction materials for these breeder reactors. *See* NUCLEAR REACTOR.

The most significant problem, however, remains that of resistance to oxidation and high-temperature corrosion in present heat engines. Superalloys contain small-to-moderate amounts of highly reactive elements such as chromium and aluminum which react easily with oxygen to form a thick tenacious semiplastic oxide surface film. This prevents further reaction of the aggressive environment with the underlying metal. This is the reason why stainless steels are "stainless" all contain a minimum of 10% chromium. Superalloys follow approximately the same rule, but often also contain about 5% aluminum, which further enhances oxidation resistance. It has been found that small (<1.0%) additions of rare-earth elements (such as yttrium) further improve oxidation resistance by reducing oxide spalling during the inevitable thermal cycles which gas turbines experience. *See* RARE-EARTH ELEMENTS.

The natural protective system works well when oxidation is the only or primary type of attack. However, when the contaminants—vanadium and others—are present, they react in myriad ways in the 1000–2000°F (540–1100°C) temperature range to destroy the protective oxide and eventually the alloy. Coating is a method used to combat this problem.

Superalloy processes. Metallurgists have been seeking ways to increase the capability of superalloys through development of new processes.

Oxide dispersion strengthening. One technique to generate improved strength is known as oxide dispersion strengthening (**Fig. 3**). The objective is to distribute very fine, uniformly dispersed nonreactive oxide particles throughout the alloy. The processing usually involves starting with very fine particles of the metal itself, to which the oxide is added by a chemical or mechanical process step. The alloy is then consolidated by a mechanical pressing operation and forged into a final useful shape. Oxide dispersion strengthening materials are characterized by unusual creep resistance at very high temperatures; however, intermediate-temperature creep and all tensile properties are mediocre.

Success has been obtained in combining some of the classic strengthening factors described previously with oxide dispersion strengthening. This gives a balance of property enhancement which can be particularly useful to turbine metallurgists—good high-temperature creep strength from oxide dispersion strengthening and good intermediate-temperature strength from the other mechanisms. In this process, prealloyed superalloy powder (carbide- and γ'-strengthened) is hammer-milled in the presence of a very reactive metal, such as yttrium (Y), with oxygen (O) present. The result is a superalloy containing finely dispersed Y_2O_3 particles. *See* CREEP (MATERIALS).

Directional solidification. A technique adaptable to investment casting is that of directional solidification, which is particularly suitable to the complex shapes of airfoil parts used in gas turbines. It has been found

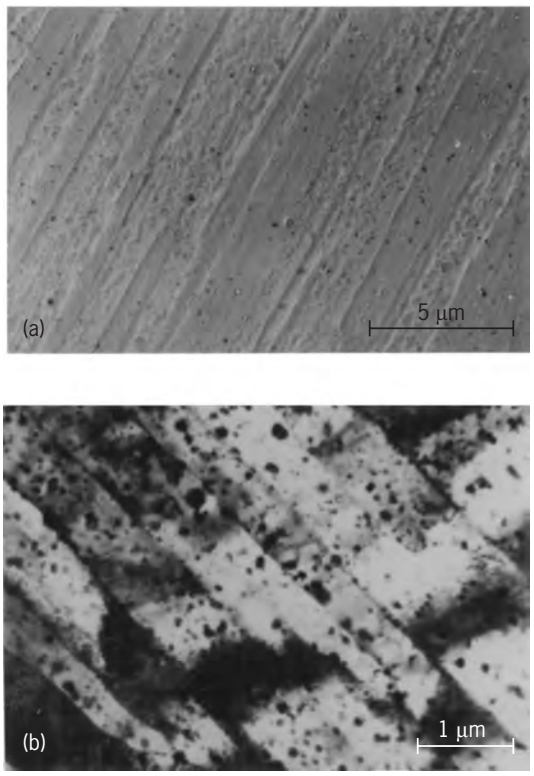

Fig. 3. Structure of thoria-strengthened nickel showing thoria particles. (*a*) Electron micrograph. (*b*) Transmission electron micrograph. (*From C. T. Sims and W. C. Hagel, eds., The Superalloys, Wiley-Interscience, 1973*)

that by commencing the freezing of molten superalloys (held in a ceramic mold of the shape desired) at the bottom, then allowing the freezing process slowly to proceed up through the shape to be cast, the grains of the structure acquire a long slender shape in the direction of freezing. This significantly increases the ability of the structure to withstand mechanical load in the freezing direction, and particularly increases resistance of the superalloy to a complex phenomenon involving stress and temperature cycling or temperature gradients, known as thermal fatigue. Thermal fatigue failures commence at transverse grain boundaries by this process.

Another approach to advanced airfoil materials is to eliminate grain boundaries entirely. In directional solidification, a number of grains nucleate in the first solid to freeze, and the few with the most preferred orientation for growth choke off solidification of the others and grow along the length of the airfoil. By reducing the cross section of the first solid to form, fewer grains are nucleated. If the ceramic mold also has a short length of spiral cavity before the airfoil shape, solidification of metal through this spiral "pigtail" will allow only a single grain to grow into the airfoil section. If a particular crystal orientation with respect to the airfoil is required, a seed crystal is used instead of a pigtail. The seed placed at the bottom of the mold is long enough that the very bottom portion is not melted in the directional solidification furnace. As the mold-containing molten alloy is withdrawn, solidification occurs on the seed, with essentially the same crystal orientation as the seed. The seed can be cut from the solidified airfoil and used to seed additional airfoils. These single crystals, also known as monocrystals, allow elimination of grain-boundary strengtheners such as zirconium, carbon, boron, or hafnium. This may allow more flexibility in the remainder of the composition, and higher heat treatment temperatures to optimize alloy mechanical properties and microstructure. Further, freezing occurs in the 001 crystallographic direction, preferred because it is in the low-elastic-modulus direction for the face-centered cubic alloys.

Powder metallurgy. An old process, powder metallurgy, has found increased use for high-temperature superalloys, and offers the possibility of added flexibility in alloy chemistry. Alloys originally developed for cast and wrought processing have been modified slightly for use as powder metallurgy materials. For large structures such as turbine disks, this has resulted in much more homogeneous chemistry and microstructure and achievement of outstanding thermal fatigue resistance. Metal flow resistance in isothermal forging is reduced because of nearly superplastic grain-boundary sliding. *See* POWDER METALLURGY; SUPERPLASTICITY.

Coatings. Superalloys and stainless steels must possess a balance of properties for high-temperature service. The most oxidation- and corrosion-resistant alloys do not have acceptable strength for most structural applications. Therefore, a viable solution is to utilize strong alloys for airfoil structures but then coat them to create environmental stability. This is done by adding elements such as chromium and aluminum into the surface of the alloy; these elements react with the aggressive environment to form highly protective oxides. The coating is applied by chemically reacting the turbine part with an atmosphere containing chromium or aluminum halides at very high temperature so that the active elements diffuse into the surface of the alloy to form the protective layer. Thus, ultimately, the coating layer is composed mainly of nickel from the alloy, and aluminum or chromium or both added in the coating process. Other methods develop overlayer coatings, which form not by reaction with the nickel substrate but by deposition from sources of the desired overlayer chemistry that are made by processes such as physical vapor deposition from an electron-beam–heated source, or by low-pressure plasma spraying of powders of desired coating compositions. Coatings formed by any of these methods are bonded to the substrate in the high-temperature heat treatment to cause the necessary interdiffusion. Such aluminum- and chromium-rich coatings can triple the life of industrial gas-turbine parts at temperatures such as 1600°F (870°C) in oxidizing atmospheres. *See* METAL COATINGS; SPUTTERING.

Eventually, however, the coatings fail by oxidizing away or by further interdiffusion with the superalloy underneath. If a very thin layer of platinum is used, the concentration of aluminum in the surface appears to be enhanced, creating an even greater measure of protection. A metallographic picture of a

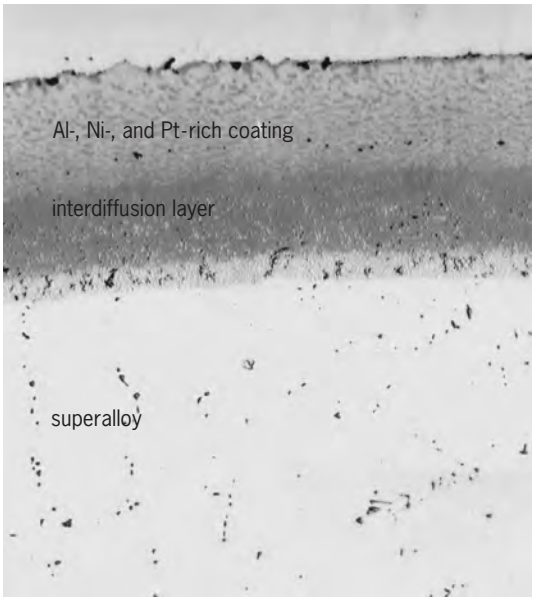

Fig. 4. Aluminide-platinide coating on a nickel-base superalloy. (*E. Buchanan, General Electric Co.*)

superalloy with a coating made by the electron-beam process, with aluminum increasing in concentration toward the surface, is shown in **Fig. 4**.

Nonmetallic materials. Considerable activity involving attempts to adapt ceramics to high-temperature applications has occurred since the mid-1960s. These are ceramics of the covalent-bonded type, such as silicon carbide and silicon nitride. Oxide ceramics, such as aluminum oxide (Al_2O_3) and zirconium oxide (ZrO_2), possess ionically bonded structures, and so tend not to possess usable high-temperature creep resistance. Turbine designers and materials engineers are struggling with the problems of utilizing the covalent ceramics, since they possess great strength. It has also become apparent that these ceramics possess great oxidation and corrosion resistance—features that would be useful in turbine equipment which must handle high-temperature products of combustion from coal. Data have shown that silicon carbide (SiC) and silicon nitride (Si_3N_4) are attacked to only 2 or 3 mils (0.05–0.08 mm) in depth after 6000–8000 h exposure in corrosive atmospheres; most high-temperature metals or alloys cannot meet this performance level. However, their complete lack of ductility means that new design techniques are required to prevent early and catastrophic failure, and only very small parts have been shown to be useful so far. It is not certain that such ceramics will be usable in heat engines.

To bypass the brittleness problem, the concept of ceramic composites has emerged. In these materials, a ceramic matrix is filled with ceramic fibers, which strengthen the whole body. When subjected to a high mechanical load, the fibers pull against the matrix and slip slightly, giving the material a certain level of tolerance. This is known as fiber pull-out. However, the effect is not elastic, so that it does not approach the deformation tolerance normally seen by metallic alloys that is essential for safe structures in tension.

The use of graphite and graphite/graphite compositions has also been explored. Graphite has a strikingly unique combination of very low density and high elastic modulus, and it demonstrates increasing mechanical strength with increasing temperature. It is capable of mechanical service at tempertures of 2200°C (4000°F) or higher. However, since graphite is a form of carbon, it has virtually no oxidation resistance. Therefore, attempts to utilize graphite must involve truly extraordinary success in protecting it from oxidation. So far, protective coatings have had but limited success, and it appears that use may well be limited to rather low temperatures. *See* CERAMICS; COMPOSITE MATERIAL; GRAPHITE; METAL, MECHANICAL PROPERTIES OF.

Chester T. Sims

Bibliography. American Society for Metals, *Metals Handbook, Desk Edition*, 2d ed., 1999; American Society for Metals, *Sourcebook on Materials for Elevated-Temperature Applications,* ed. by E. F. Bradley, 1979; W. Betteridge and J. Heslop (eds.), *The Nimonic Alloys,* 2d ed., 1974; C. T. Sims, N. S. Stoloff, and W. C. Hagel (eds.), *Superalloys II,* 1987.

Highway bridge

A structure that crosses over a body of water, traffic, or other obstruction, permitting the smooth and safe passage of vehicles. In highway transportation systems, the term "bridge" is usually reserved for structures over bodies of water. However, many other structures are generally considered highway bridges. An overhead is a structure carrying a highway over a railroad, and an underpass is a structure providing passage of a highway under a railroad. An overcrossing is a structure carrying a county road or a city street over a state highway, and an undercrossing is a structure providing passage of a county road or a city street under a state highway. A separation is a structure separating into two state highways. A connector ramp is a structure connecting intersecting highways and roads. An interchange is the group of ramps and structures providing connections for traffic between intersecting highways (**Fig. 1**). A viaduct is an elevated structure carrying a highway over streets, railroads, or other features. *See* BRIDGE; HIGHWAY ENGINEERING.

Highway bridges can be made of steel (**Fig. 2**), concrete (**Fig. 3**), timber, stone, metal alloys, or advanced composite materials, and may have different structural systems such as girder (beam), truss, arch, cable-stayed, and suspension. *See* STRUCTURAL MATERIALS.

Design. Bridge design is a combination of art and science. Conceptual design is usually the first step. Before any theoretical analysis and detailed design proceeds, the designers visualize the bridge in order to determine its function and performance. The conceptual design process includes selection of bridge systems, materials, proportions, dimensions,

foundations, esthetics, and consideration of the surrounding landscape and environment. A bridge may be straight or horizontally curved, or have skewed supports. The width of a highway bridge is determined by the number and width of the traffic lanes and the shoulder or sidewalk width, and is typically the same dimension as the approaching highway.

The selection of bridge type is influenced by many factors such as span length, site geology and foundation requirements, design loads, surrounding geographical features, width requirements, clearance requirement below the bridge, transportation of construction materials, erection procedures, and construction cost and duration. The **table** shows the span lengths appropriate for various bridge types. A bridge is required to fulfill its function as a thoroughfare while blending and harmonizing with its surroundings.

The final design process involves structural analysis, member and detail design, and preparation of construction drawings and specifications. Structural analysis commonly involves computer models, which use appropriate material properties, member discretization, boundary conditions, and loads. Members and connection joints are proportioned to carry all possible loads (permanent loads, vehicular live loads, wind loads, and earthquake loads), combined and factored in accordance with the requirements of applicable design standards and codes. Two standards are the American Association of State Highway and Transportation Officials Load and Resistance Factor Design (AASHTO-LRFD) Bridge Design Specifications (2004) and AASHTO Standard Specifications for Highway Bridges (2002) in the United States. **Figure 4** shows the AASHTO "design truck." The variable axle spacing between the 145-kilonewton loads is adjusted to create a critical condition for the design of each location in the structure. Well-prepared construction documents are biddable, constructable, and cost-effective. *See* STRUCTURAL ANALYSIS; STRUCTURAL DESIGN.

Construction. Bridge construction involves managing the fabrication and erection operations needed for safe and efficient building of the structures in accordance with the construction documents.

A fundamental part of construction engineering is project management (planning, scheduling, and controlling). Planning starts with analyses of the type and scope of work to be accomplished, as well as the identification of construction methods and equipment, and type and size of the work force. Scheduling establishes a sequence of operations and accounts for the interrelation of operations at the job site, as well as allocation of the work force and equipment. Controlling consists of procuring materials in accordance with construction staging, dictating overall construction procedures, engineering inspection, and maintaining complete records of daily operations, payments, and expenditures. Proper planning, scheduling, and controlling prevents costly delays and cost overruns. *See* CONSTRUCTION ENGINEERING.

Fig. 1. Aerial view of I-105/605 Interchange in California. (*California Department of Transportation*)

Maintenance. The primary objective of bridge maintenance is to ensure public safety. Proper maintenance also reduces life cycle costs and earns public confidence. Before 1970, components of the United States' highway network exhibited rapid aging. Since then, bridge maintenance programs have evolved into a sophisticated bridge management system.

The Federal Highway Administration (FHWA) requires all state agencies to submit highway bridge

Fig. 2. Steel girder bridge.

Types of bridges and applicable span length

Bridge type	Span range	Leading bridge and span length
Prestressed concrete girder	10–300 m (33–984 ft)	Stolmasundet, Norway; 301 m (988 ft)
Steel I/box girder	15–376 m (49–1234 ft)	Sfalassa Bridge, Italy; 376 m (1234 ft)
Steel truss	40–550 m (131–1804 ft)	Quebec, Canada; 549 m (1801 ft)
Steel arch	50–550 m (164–1804 ft)	Shanghai Lupu, China; 550 m (1804 ft)
Concrete arch	40–425 m (131–1394 ft)	Wanxian, China; 425 m (1394 ft) [steel-tube-filled concrete]
Cable-stayed	110–1100 m (361–3610 ft)	Sutong, China; 1088 m (3570 ft)
Suspension	150–2000 m (492–656 ft)	Akaski-Kaikyo, Japan; 1991 m (6532 ft)

data on a Structure Inventory and Appraisal (SI&A) sheet. Bridge data include description of the structure, structural data and history, traffic information, load rating, condition and appraisal ratings, and inspection findings. The information on SI&A sheets is a valuable aid to establish maintenance and replacement priorities, as well as to project the maintenance cost of the highway bridges.

FHWA requires that each public bridge be inspected at a regular interval not exceeding 2 years. Underwater bridge components that cannot be visually evaluated during periods of low flow by tactile means are required to be inspected at an interval not exceeding 5 years. The purpose of the bridge inspection is to obtain the information necessary to properly evaluate the bridge capacity and the adequacy of the bridge.

The major tasks of the bridge inspection are to (1) identify minor problems that can be corrected before they develop into major repairs; (2) locate bridge components that require repairs in order to avoid replacement; (3) identify unsafe conditions, prepare accurate condition reports, and recommend corrective actions; (4) investigate more serious damage and its effect on the structure; and (5) generate accurate bridge records.

Bridge ratings. The evaluation or rating of existing bridges is a continuous activity of the bridge owners to ensure the public safety. The evaluation provides criteria to repair and rehabilitate superstructures, to issue special permits, and to post, close, or replace the existing bridge.

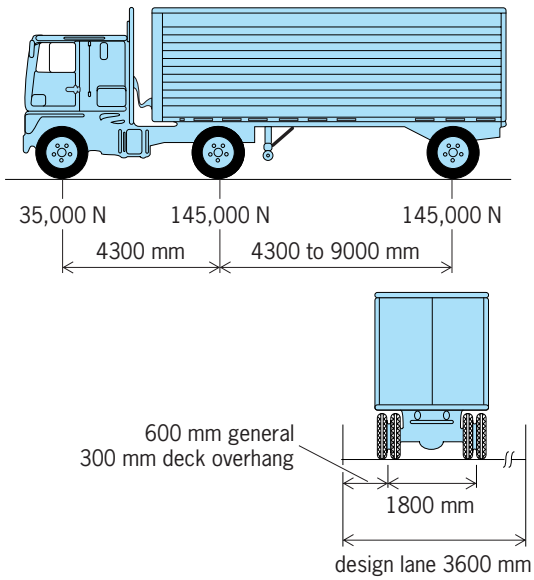

35,000 N = 7868 lbf
145,000 N = 32,600 lbf
4300 mm = 169 in.
9000 mm = 354 in.
600 mm = 23.6 in.
300 mm = 11.8 in.
1800 mm = 71 in.
3600 mm = 142 in.

Fig. 4. AASHTO design truck (load).

It is generally believed that when a bridge is designed for the AASHTO design vehicles, and is constructed and maintained in accordance with the contact documents, the bridge will have adequate capacity to handle the actual present traffic. However, changes in details during construction, failure to meet minimum material strength requirements, unexpected settlements of foundation after construction, and unforeseen damage to a member could affect the capacity of the bridge. In addition, old bridges might be designed for a lighter vehicle than the present one, or a different design code.

Sometimes, an industry needs to transport their heavy machinery from one location to another location. These vehicles could weigh much more than the design vehicles, and thus the bridge owner may need to determine the current live-load-carrying

Fig. 3. Concrete girder bridge.

capacity, or rating, of the bridge to issue special permits. There are two levels of rating for bridges: inventory and operating. The inventory rating determines the load that can be safely carried by a bridge for an indefinite period. The operating rating reflects the absolute maximum permissible load that can be safely carried by the bridge.

When a bridge is found to have inadequate capacity for design vehicles, engineers look at several alternatives prior to closing the bridge. Possible alternatives include limiting allowable vehicle weight and speed, reducing vehicular traffic volume, restricting vehicles to certain lanes, and recommending repairs to alleviate the problem. In addition, when evaluations show the structure is marginally inadequate, frequent inspections may be recommended to monitor the physical condition of the bridge in the later stages of its life cycle. Lian Duan

Bibliography. AASHTO, *AASHTO LRFD Bridge Design Specifications*, 3d ed., American Association of State Highway and Transportation Officials, Washington, DC, 2004; AASHTO, *Standard Specifications for Highway Bridges*, 17 ed., American Association of State Highway and Transportation Officials, Washington, DC, 2002; R. M. Barker and J. A. Puckett, *Design of Highway Bridges*, Wiley, New York, 1997; W. F. Chen and L. Duan (eds.), *Bridge Engineering Handbook*, CRC Press, Boca Raton, FL, 2000.

Highway engineering

A branch of civil engineering that includes planning, design, construction, operation, and maintenance of roads, bridges, and related infrastructure to ensure effective and efficient movement of people, goods, and services that use this mode of transportation. Highway engineering has a profound impact on residential, commercial, recreational, and industrial locations and operations, and therefore has a far-reaching influence on the cultural and economic activities of a region. *See* CIVIL ENGINEERING.

Planning and project development. At the planning stage, the need for the highway development is established. Depending on the scale of the proposed development, highway planning may occur at any level of the administrative hierarchy such as the state, regional, or local levels, and public input and involvement is critical. The need for highway development may arise out of new construction, major structural repair or replacement, existing or projected future travel demand in excess of available capacity, and need for increased mobility through capacity enhancement along the planned or existing highway corridors. Other needs arise from safety concerns that can be addressed only through improvements in the geometric design of roads, such as curves and grades. The cornerstone of highway planning is the estimation of current and future traffic volumes on the road network and comparison of the expected volume and the capacity of each individual highway section to yield a measure of the expected level of performance or service at that section.

Traffic volume. This may be measured in terms of annual average daily traffic (AADT) and computed as the total yearly traffic count divided by the number of days in a year. AADTs are obtained by conducting 48-hour counts and adjusting these values by a factor derived from continuous counts. The 48-hour counts are traditionally done using mechanical devices such as pneumatic tubes laid across the road. Continuous counts are done using automatic traffic recorders permanently embedded in the pavement. Continuous counts provide factors by which values from 48-hour counts can be adjusted to reflect seasonal or temporal variations in traffic volumes.

For urban highway systems, traffic volumes are generally predicted using a four-step process. This begins with trip generation, where the study area is divided into zones and the number of trips generated by each zone is computed on the basis of socioeconomic characteristics of that zone. This is followed by a distribution of trips from and to each zone. The third step, modal split, predicts the proportion of trips that use the major available modes of transport. Finally, the trips are assigned on the various links of the network.

For purposes of design, traffic volumes are needed for a representative period of traffic flow. However, because traffic flow is not uniform, it is inappropriate to use traffic volumes per average hour or traffic volumes during the hour of highest flow, as these would result in underestimation and overestimation, respectively, of traffic flow. The actual traffic flow rate used for design is often a compromise value that falls within these two extremes. Traffic flow rate is the number of vehicles that pass a given spot in a specified time interval, such as 1 hour. In order to incorporate different vehicle classes in the traffic stream, vehicle flows are expressed in terms of passenger-car equivalents.

Capacity. This is the maximum theoretical traffic flow rate that a highway section can accommodate under a given set of environmental, highway, and traffic conditions. The capacity of a highway section is influenced by factors such as the number of lanes, lane width, presence of traffic control systems and their effectiveness, frequency and duration of traffic incidents, and efficiency of collection and dissemination of highway traveler information. To increase the capacity of an existing highway section, engineers traditionally widen existing lanes, provide additional lanes, or make horizontal curves and vertical grades smoother and gentler. Where physical, institutional, or environmental constraints preclude such geometric improvements, capacity-enhancing technologies that are collectively termed intelligent transportation systems (ITS) can be pursued. ITS increase highway capacity in a variety of ways, including enhanced traffic management, reduced probability of crashes, quick detection and clearance of incidents, and effective delivery of traveler information. *See* TRAFFIC-CONTROL SYSTEMS.

Level of service. Traffic conditions arising from the interplay of traffic volume and section capacity are perceived by road users in a way that is termed level of service (LOS). LOS ranges from an index of A (very good) to F (very poor). When traffic volume approaches highway capacity, vehicular flow becomes congested and speeds fall below the design speed. Under such conditions, queues are formed and operations within the queue are characterized by stop-and-go waves. Such a situation is indicative of a poor level of service (LOS F). When traffic volumes are much lower than highway capacity, vehicles travel at a freely flowing speed and are virtually unaffected by the presence of other vehicles in the traffic stream, indicating LOS A or B. In most cases, highway engineers design for an LOS of C. An important part of highway planning is to specify the desired minimum level of service.

Public participation. Public participation in highway project planning and development should not be limited to conventional public meetings where completed design alternatives are presented for public input, but should occur early in the project development process. A consensus is often reached among identified stakeholders regarding the existence of a problem, the need for the highway project, and to some extent, how it should be fixed.

Environmental considerations. Highway facilities often have adverse effects on the environment, such as noise pollution, air pollution, water pollution, and degradation of habitats and the flora and fauna they support. As such, the prediction and mitigation of imminent environmental impacts of a highway are a key aspect of transportation planning. Traffic noise arises from sources such as vehicle/air interaction, tire/pavement interaction, vehicle exhausts, and engines, and may be addressed at source, in the transmitting medium, or at the receptor. Engine insulation and mufflers are used to reduce engine and exhaust noise, respectively; noise absorbing pavement materials or longitudinal tining (texture) reduce tire/pavement noise; and traffic noise barriers protect highway neighbors from traffic noise. Highway air pollution is caused by the dispersion of gas or liquid droplets and solid particles emitted from vehicle tailpipes, such as carbon monoxide, nitrogen oxides, volatile hydrocarbons, sulfur oxides, and particulate matter. Pollutants are also released during refueling and from wear of tires and brake linings. Air pollution is generally associated with respiratory damage to humans and damage to structures and vegetation. In addition, highway system construction, operation, and maintenance degrade water quality by contributing a variety of pollutants to runoff, such as solids, heavy metals, oil and grease, and pesticides. In many countries, transportation-related legislation has been passed to protect the environment from adverse effects of highway usage. *See* AIR POLLUTION.

Geometric design. Highway engineering design involves selecting the appropriate location, alignment, and cross section of a planned highway by considering three major factors—human, vehicle, and roadway/environment—and how these factors interact to provide a safe highway. Human factors for design include reaction time for braking and steering, visual acuity for traffic signs and signals, and car-following behavior. Vehicle considerations include vehicle size and dynamics, which are essential for determining lane width and maximum slopes as well as for selecting the design vehicles, which range from passing cars to tractor trailers. The physical dimensions of vehicles affect several design elements, including road corner radii and clearance of highway overpasses. Highway engineers design road geometry to ensure that vehicles negotiating curves and grades are stable, and to provide adequate sight distances for vehicles that undertake passing maneuvers along curves on two-lane two-way roads.

Location planning. This involves fitting the road efficiently onto the surrounding terrain and environment. Given a set of specific terminal points, a number of locations are usually evaluated, and the best alternative is chosen according to criteria such as travel time, vehicle operation cost, section length, environmental impacts, total cost, ease of maintenance, accessibility/connectivity, and landscape aesthetics.

Horizontal alignment. This can be represented in an aerial view of the highway (**Fig. 1***a*). It consists of straight lines and curves. Curves are fitted to provide a smooth transition between straight highway sections. A transition curve may be used to connect the circular and straight sections. Elements of curve design include the selection of an appropriate radius and the degree of superelevation in the curve (raising one lateral end of the highway above the other). To ensure vehicle stability at high speeds, large curve radii and maximum superelevation are necessary. However, for curves at low-speed sections, such as local streets, flat curves of small radii may be acceptable. Intersection and interchange designs are part of horizontal alignment.

Intersections and interchanges. Where two or more highways cross each other at the same level, various vehicle maneuvers (turning, crossing, and through movements) occur within a limited area, and as the volumes of these movements increase there is increased likelihood of traffic conflicts and crashes. Channelization—the provision of islands or raised curbs—reduces such danger by limiting each stream to a unique path. In areas where traffic volumes are high, the movement of traffic streams can be separated in time using multiphase traffic signals that permit movement of only nonconflicting streams in each phase. Where highways cross each other at different levels, grade separation (overhead bridges with or without access) is used.

Vertical alignment. The vertical alignment of a highway is represented by its longitudinal profile, which illustrates the elevation of all points along the length of the highway. The purpose of vertical alignment design is to determine the level of the highway at each point in an overall bid to ensure adequate safety and drainage along the entire length of the section. Vertical alignment consists of grade tangents connected

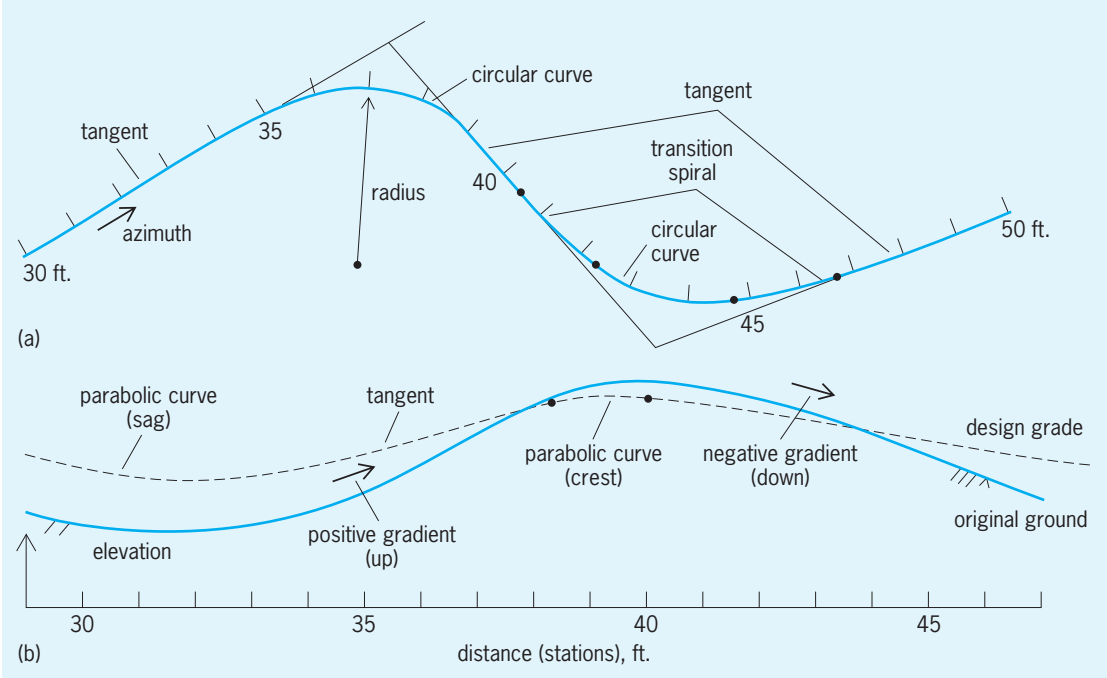

Fig. 1. Typical (*a*) plan view and (*b*) profile view of a highway section.

with parabolic curves. The desirable maximum slope of the tangent and rate of slope change depends on the road classification and characteristics of the design vehicle. The vertical alignment is generally kept as close as possible to the natural terrain to minimize earthwork volumes. However, challenging highway geometry, such as steep slopes, may be unavoidable at locations of difficult topography. Also, the combined effect of the longitudinal grade and the length over which it occurs is an important consideration. A vertical curve could be a crest or a sag curve (Fig. 1*b*). Construction of vertical curves is typically expensive because of the volume of earthwork involved. In the design of vertical alignment, it is desirable to balance cuts and fills, so that suitable material excavated in cuts can be used to fill low-lying sections. A challenge often encountered in highway engineering design is to minimize earthwork volumes while providing acceptable levels of safety. An alignment with acceptable level of safety is one that provides drivers with adequate sight distance to enable them to stop their vehicles in response to a perceived obstruction. Braking distance, the theoretical minimum distance that a vehicle travels while braking to a stop, depends on factors such as vehicle speed and friction between the vehicle tire and pavement. The distance traveled by the driver between the time of perception and braking is added to the braking distance to give the total required stopping sight distance. For sag curves, the critical concern is the headlight sight distance (length of road illuminated by vehicle headlights) for nighttime driving.

Highway cross section. This refers to the road profile perpendicular to the direction of travel and within the limits of the highway right-of-way. Highway cross section consists of driving lanes and may include other elements such as bicycle or pedestrian lanes, shoulders, medians, barriers, curbs, cross slope for drainage, and superelevation. **Figure 2** shows a typical cross section for a rural highway. The elements

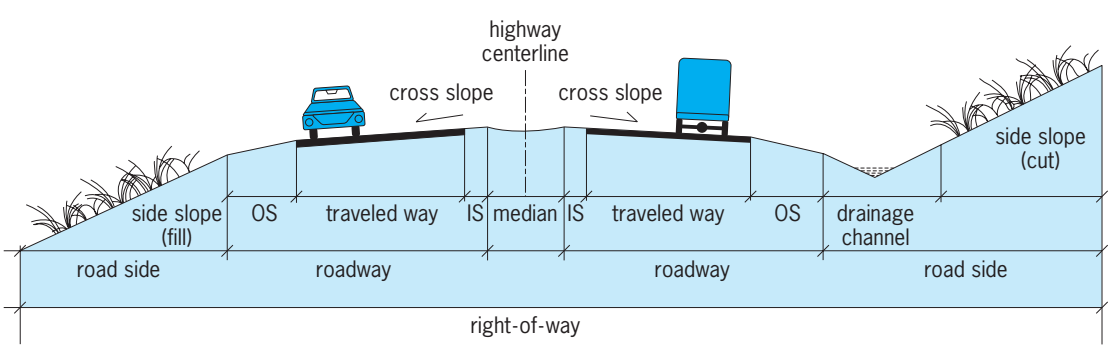

Fig. 2. Typical highway cross section. IS, inside shoulder; OS, outside shoulder.

and dimensions of highway cross section at a given location depend on road classification and land use (urban or rural).

Pavement design. This is the process of selecting appropriate types and thicknesses of pavement layers so that the pavement structure can withstand expected loads in a cost- effective manner. The original ground, or subgrade, typically lacks adequate bearing capacity and may need replacement or overlay using more competent material to reduce the stresses imposed by the vehicle loads to tolerable levels. This way, the natural ground can withstand loads without deforming.

Pavement materials. Pavement materials include surface and subsurface layers that consist mainly of mineral aggregate. Aggregate may be unbound (in the case of gravel roads) or bound using cementitious material such as ordinary portland cement (for rigid pavements) or asphaltic cement (for flexible pavements). Pavement materials have to meet specified standards before they are used in construction. A wide range of physical and chemical tests are carried out on ingredients as well as pavement mixes to ensure that they are sufficiently strong and durable.

Ingredients. Commonly used mineral aggregates include crushed rock and natural gravel and sand. An aggregate mixture whose particles are of a wide range of sizes, with the smaller particles filling the voids between the bigger ones, is described as a well-graded mix. A well-graded aggregate mix provides the maximum degree of aggregate interlock and generally yields competent material. In some cases, physical or chemical stabilizers such as lime are added to improve the properties of available material. In recent years, efforts have been made to recycle waste materials as pavement materials, including crushed glass, ground-up tires, and recycled bituminous mixes scraped from old pavements.

Mixes. For rigid pavements, ordinary portland cement is mixed with water and aggregates to produce a viscous concrete mix that is poured into prepared forms and vibrated. The mix hardens, or cures, as the cement hydrates, and is usually watered to enhance the curing process. Special additives are usually added to the concrete mix to retard or to accelerate the curing process or to improve performance. For flexible pavements, asphaltic cement (bitumen) is liquefied by heating, solvent addition, or emulsification, and is mixed with mineral aggregates. The asphaltic concrete mix is then laid on the prepared base course in compacted layers. The laid material is cured by cooling or by evaporation of the solvent/emulsifying agent. *See* CEMENT.

Pavement types. There are generally three types of pavements: gravel, flexible, and rigid. The choice of any type depends on a variety of factors, including traffic, construction and maintenance costs, and availability of materials, labor, equipment, or funding. *See* PAVEMENT.

1. *Gravel pavement.* This is the simplest type of pavement and is often designed for lightly traveled roads. Material used for such pavements often consists of well-graded, naturally occurring silty or clayey gravel. This material is spread, watered, and compacted in layers to provide a crowned shape that is good for drainage purposes. Increased traffic may result in accelerated pavement deterioration and may require more frequent maintenance. *See* GRAVEL.

2. *Flexible pavement.* This is a multilayered structure that includes a subbase, a base, and an asphaltic wearing course (**Fig. 3***a*). It is described as flexible because the main structural layer is based on asphalt, a viscous material that causes inability of asphaltic concrete layers to span a gap without sagging into it. *See* ASPHALT AND ASPHALTITE.

3. *Rigid pavement.* This pavement type consists of a plain or steel-reinforced portland cement concrete slab laid on a prepared surface such as crushed-stone base course (Fig. 3*b*). It is described as rigid because the slab is competent enough to bridge any minor gaps in the underlying layers without undue failure. *See* REINFORCED CONCRETE.

Highway construction. Highway construction usually follows planning and design, and involves the implementing facility designs to yield new or reconstructed facilities such as pavements, bridges, drainage structures, and traffic control devices. Road construction is often preceded by detailed stakeout surveys and preparation of the subgrade. *See* CONSTRUCTION ENGINEERING; CONSTRUCTION EQUIPMENT.

Surveys and subgrade preparation. The purpose of stakeout surveys is to lay out the designed position of all the road elements on the field as a guide for the contractor during construction. Stakeout surveys are carried out continuously during the construction period to monitor the work. After completion of the

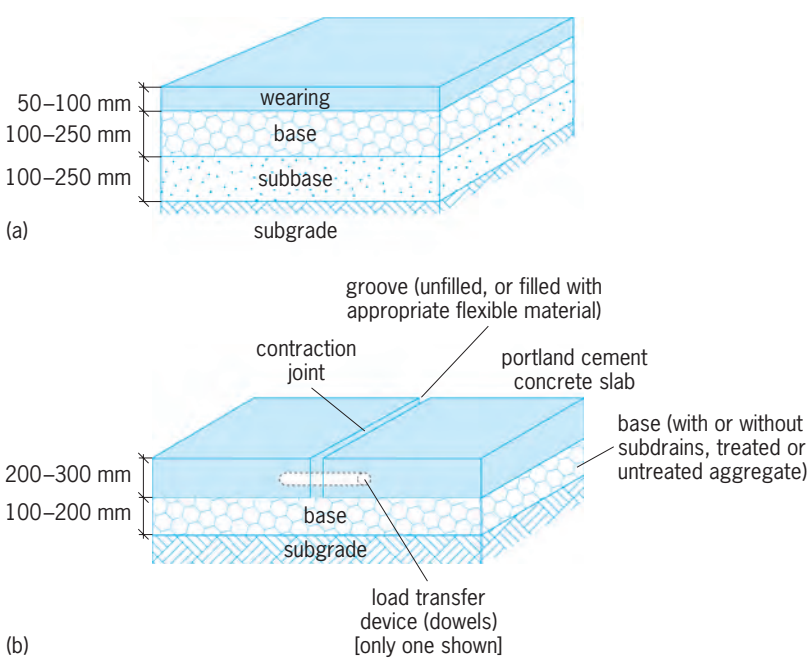

Fig. 3. Cross sections of (*a*) typical flexible pavement and (*b*) typical rigid pavement. Some designs include a subbase layer, between the base and subgrade.

surveys, clearing and topsoil removal may be necessary where the road alignment passes through vegetated areas. Where the alignment passes through built-up areas, clearing includes demolition and relocation of existing structures and utilities within the road right-of-way. If the formation level (the level of foundation on which pavement will sit) is below the existing ground level, excavation is needed to lower the existing ground, thus forming a cut section. If the formation level is above existing ground, filling is carried out to raise the road and an embankment is subsequently formed. Embankments are constructed one layer at a time, and each layer is spread uniformly and compacted using an appropriate combination of rolling, watering, vibrating, and kneading, depending on the type of fill material.

Pavement. When the subgrade elevation (formation level) is reached either in a cut or in a fill section, it is compacted thoroughly to cross-fall shape, and the subbase and base courses are laid. These courses may consist of natural gravel, crushed rock, or stabilized material, depending on pavement design specifications. The type of wearing surface ranges from gravel, sand-seal, or chip-seal for low-volume roads, to asphaltic or portland cement concrete for high-volume roads.

Traffic operations and control. Traffic signals are important traffic control devices for smooth operation of highways. The typical traffic signal at an intersection displays a sequence of green, amber, and red. One complete signal sequence is called a cycle. Traffic signals are either pretimed or demand-actuated. Pretimed signals repeat a predetermined unchanging cycle. For demand-actuated signals, the lengths of display time for each color vary depending on the presence of vehicles or pedestrians. Flow-concentration controllers are capable of sensing detailed demand information and respond appropriately by revising the signal cycle length and phasing patterns. Advanced traffic control systems usually employ a central computer to monitor and control traffic flows on large highway networks. Special types of demand-actuated signals are able to recognize and give priority to particular types of vehicles such as buses, light-rail vehicles, and emergency vehicles.

Highway infrastructure management. The practice of highway engineering has evolved to include the effective and efficient management of existing infrastructure (such as pavement, bridges, and traffic equipment) and operations (such as safety and congestion) throughout the life cycle of the asset. In all functional areas of highway infrastructure management, the key is to arrive at optimal facility-level and network-level scoping and timing decisions so that the highest levels of service can be accrued at least cost to the agency and user over the life cycle. The levels of service are typically expressed in terms of a variety of performance measures that include facility condition, safety, and protection from accidental or natural disasters.

Pavement management. Pavement condition is monitored over a period of time using a condition index or serviceability rating. Most highway agencies have implemented pavement management systems for cost-effective rehabilitation and maintenance decision-making. The type of intervention for a particular section of highway depends on the severity and frequency of pavement distress as well as the road class. Severe and widespread occurrences of potholes on high-volume roads, for instance, may warrant extensive rehabilitation in the form of newly added pavement layers, with or without removal of the existing layers. However, scattered and light occurrences of cracking on a local road may require only routine maintenance in the form of crack sealing. For flexible pavements, maintenance is often needed to address common surface problems, including potholes, cracks, and ruts. Potholes and cracking are caused by a variety of factors, including the freezing/thawing of moisture beneath the pavement and traffic loads, and are exacerbated by lack of adequate lateral restraint to the pavement layers. Rutting in the wheel tracks may be due to excessive wheel loads, poor asphaltic concrete mixes, poor subgrade material, or subgrades weakened by the ingress of moisture. For rigid pavements, maintenance addresses problems such as flexural cracking caused by stresses induced by loss of subgrade or base support. Other problems on rigid pavements include spalling or faulting at the joints, surface deterioration, and corrosion of steel reinforcement, which is accelerated by the deicing salts used during winter maintenance.

Bridge management. Highway agencies responsible for bridge maintenance and repair are faced with the challenge of monitoring the condition of bridges and deciding what types of repairs need to be applied and when. Through the development and implementation of bridge management systems, many agencies establish decision support tools that analyze and summarize data, use mathematical models to make predictions of bridge conditions, and facilitate rational evaluation of alternative policies and programs, preferably on the basis of multiple performance measures.

Congestion management. In highway congestion management, a system-wide plan is developed for implementing measures to mitigate the magnitude and duration of traffic congestion within a given analysis period and with a given budget. The cost and effectiveness (benefits) of each mitigation measure are key inputs to the management process. Mitigation measures include a variety of approaches involving travel demand management as well as technological innovations such as intelligent transportation systems. The purpose of these approaches is to reduce demand, encourage off-peak travel, increase capacity, and enhance traffic flow. Provisions may include improved signalization, use of frontage roads, arterial access control measures, provision of special lanes for high-occupancy vehicles, carpooling, improved public transportation services, electronic toll collection, and incident response services.

Safety management. This is a systematic process to ensure that highway safety resources are appropriately allocated to reduce the frequency and severity

of all types of highway crashes. Safety management involves identifying, considering, implementing as appropriate, and evaluating all safety-enhancing opportunities in all phases of highway development. As such, highway engineers seek to minimize the frequency and severity of crashes by improving highway planning, design, construction, maintenance, and operation. Highway safety management generally consists of several components, including analyzing and predicting crash trends, identifying hazardous links and intersections in highway networks, identifying crash causes, and evaluating effectiveness and cost-effectiveness of various countermeasures. Specific highway-safety countermeasures include the provision of traffic signalization, channelization, luminous median markers, road signs, and crash barriers. Well-maintained pavements with adequate surface friction can also contribute to improved highway safety.

Kumares C. Sinha; Samuel Labi

Bibliography. N. J. Garber and L. A. Hoel, *Traffic and Highway Engineering*, 3d ed., Thomson-Engineering, Toronto, 2001; F. L. Mannering, W. P. Kilareski, and S. S. Washburn, *Principles of Highway Engineering and Traffic Analysis*, 3d ed., Wiley, New York, 2004; C. H. Oglesby and R. G. Hicks, *Highway Engineering*, 4th ed., Wiley, New York, 1982; *A Policy on Geometric Design of Highways and Streets*, 5th ed., American Association of State Highway and Transportation Officials, Washington, DC, 2004.

Hilbert space

An abstract notion of great power and beauty, which has been central to the development of mathematical analysis and forms the backdrop for many applications of analysis to science and engineering. Its essence lies in the fact that the objects of primary interest in analysis (namely, functions) enjoy geometrical properties which are in important ways analogous to the geometry of physical space. Thus the highly developed human visual and spatial intuition can lead to significant truths about functions.

The prototype of Hilbert space is familiar three-dimensional euclidean space, but the general notion requires "space" to be understood in an abstract way. With the development of the axiomatic method, mathematicians have come to use "space" for any collection of mathematical entities possessing in their ensemble some structure analogous to physical space. The individual entities are then called points of this space, even though they may be elaborate constructs such as functions or infinite sequences. This device of focusing on the relationship between mathematical objects and temporarily ignoring their natures is one of the most fruitful constituents of modern analysis, allowing a dramatic extension of the range of application of geometrical ideas. *See* EUCLIDEAN GEOMETRY.

Generalizations of euclidean space. Vectors (or directed line segments) in euclidean space have a rich structure. It makes sense to multiply them by real numbers or to add two of them together. It is possible to speak of their lengths (or magnitudes) and the angles between them. Vectors may be perpendicular, and Pythagoras' theorem holds. *See* LINEAR ALGEBRA.

Suitable classes of functions have some, but not usually all, of these properties. Functions can be added, and there are various ways (appropriate in different contexts) of defining the magnitude of a function and even the angle between two functions. If certain of the desirable properties of euclidean space are isolated and adopted as postulates, a class of spaces is defined that may include spaces of functions that are of concern in analysis. During the early decades of the twentieth century it was gradually realized that it is possible to select a small number of properties in such a way that the resulting spaces possess virtually all the desirable features of euclidean space except those which are closely linked to finite dimensionality. Furthermore, many examples of such spaces arise naturally in analysis, and many classical results, processes, and problems can be formulated and analyzed within the framework of these spaces. They are named after David Hilbert, who took a decisive step toward their introduction in 1906, when he proved what is known as the spectral theorem. From a contemporary viewpoint his discovery is best described as a decoupling result for certain transformations of Hilbert space, but Hilbert himself did not use the term Hilbert space, and the geometric interpretation was developed by other mathematicians, notably E. Schmidt, F. Riesz, and J. von Neumann.

Role of Hilbert space. Hilbert's innovation occurred while he was investigating integral equations which arose from mathematical physics. These equations are continuous analogs of the systems of simultaneous linear equations which are encountered in elementary algebra, but the unknown entity, instead of being a finite set of numbers, is a function. Typical examples of functions are physical quantities which depend on position and time, such as temperature and pressure. Physical laws impose constraints on these functions and so give rise to equations. Just as graphical methods give insight into the solution of a pair of simultaneous linear equations in two unknowns, so the development of Hilbert-space geometry had great consequences for the understanding of integral equations. The geometry is infinite-dimensional in general; that is, elements of the space cannot be specified by a finite set of coordinates. In consequence, ordinary geometric imagination is a shaky guide, and it must be complemented by rigorous definition and proof. By the 1940s the appropriate concepts and basic theorems were well established, and it was clear that Hilbert space is a truly fundamental mathematical structure which appears in widely disparate branches of pure and applied mathematics. A striking instance is quantum mechanics, where observable quantities are modeled by linear transformations of Hilbert space. *See* INTEGRAL EQUATION; LINEAR SYSTEMS OF EQUATIONS;

NONRELATIVISTIC QUANTUM THEORY; OPERATOR THEORY.

Definition. The definition of Hilbert space requires use of the concept of a vector space over the real or complex numbers. This is any collection of objects (to be called vectors) for which are defined two operations satisfying certain natural laws. The operations are to add any pair of vectors and to multiply any vector by a real or complex number. The laws include commutative, associative, and distributive laws. Vector spaces are important in their own right, but they are too general to be of much use in analysis, and it is necessary to postulate further structure having to do with lengths and angles.

A Hilbert space is defined to be a vector space H over the real or complex numbers that possesses an inner product and is complete with respect to this inner product. Here an inner product means a scalar (that is, real or complex number) valued function (x, y) of the pair of vectors x, y in H which satisfies conditions (1), (2), and (3).

$$(x, y) \text{ is linear in } x \tag{1}$$

$$(y, x) = \overline{(x, y)} \tag{2}$$

$$(x, x) > 0 \text{ unless } x = 0 \tag{3}$$

The overbar in Eq. (2) denotes complex conjugation; it can be omitted for real Hilbert space. Completeness is a technical condition which is automatically satisfied in finite-dimensional spaces but is needed in infinite-dimensional ones to ensure the validity of limiting procedures. The inner product corresponds roughly to the orthogonal projection of one vector on another. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES. N. J. Young

Bibliography. S. K. Berberian, *Introduction to Hilbert Space*, 2d ed., 1999; J. Dieudonné, *A History of Functional Analysis*, 1983; P. Halmos, *A Hilbert Space Problem Book*, rev. ed., 1994; F. Riesz and B. Szökefalvi-Nagy, *Functional Analysis*, 1955, reprint 1990; W. Rudin, *Functional Analysis*, 2d ed., 1991; N. J. Young, *An Introduction to Hilbert Space*, 1988.

Hill and mountain terrain

Land surfaces characterized by roughness and strong relief. The distinction between hills and mountains is usually one of relative size or height, but the terms are loosely and inconsistently used. Because of the prevalence of steep slopes, hill and mountain lands offer many difficulties to human occupancy. Cultivable land is scarce and patchy, and transportation routes are often difficult to construct and maintain. However, many rough lands, especially those readily accessible to centers of population, attract tourists because of their scenic quality and the opportunities they may afford for outdoor recreation.

High mountain ranges set up major disturbances in the broad pattern of atmospheric circulation, and thus affect climates over extensive areas. More locally, by inducing turbulence or forcing moist air to rise in crossing them, rough lands commonly induce condensation and precipitation that makes them moister than surrounding lowlands. Within the rough country, local differences in elevation and in exposure to sun and wind produce complex patterns of local climatic contrasts. These variations, in turn, are often accompanied by unusual local variety in vegetation and animal life.

Development of rough terrain. Uplift of the Earth's crust is necessary to give mountain and hill lands their distinctive elevation and relief, but most of their characteristic features—peaks, ridges, valleys, and so on—have been carved out of the uplifted masses by streams and glaciers. The upraised portions of the crust may have been formed as broad, warped swells, smaller arched folds or domes, upthrust or tilted blocks, or folded and broken masses of extreme complexity. Some limited areas owe their elevation and certain of their local features to the outpouring of thick sheets of lava or the construction of volcanic cones. Hill lands, with their lesser relief, indicate only lesser uplift, not a fundamentally different course of development.

In some rough lands, for example the Appalachians or the Scottish Highlands, the most intense crustal deformation is known to have occurred hundreds of millions of years ago, while in others, such as the Alps, the Himalayas, or the California Coast Ranges, intense deformation has occurred quite recently. Since mountains can be erosionally destroyed in relatively short spans of geologic time, however, the very existence of mountains and hills indicates that structural deformation, at least simple uplift, has continued in those areas until recent times. This conclusion is reinforced by the fact that the major cordilleran belts are currently foci of earthquake activity and volcanism.

Distribution pattern of rough land. Hill and mountain terrain occupies about 36% of the Earth's land area. The greater portion of that amount is concentrated in the great cordilleran belts that surround the Pacific Ocean, the Indian Ocean, and the Mediterranean Sea. Additional rough terrain, generally low mountains and hills, occurs outside the cordilleran systems in eastern North and South America, northwestern Europe, Africa, and western Australia. Eurasia is the roughest continent, more than half of its total area and most of its eastern portion being hilly or mountainous. Africa and Australia lack true cordilleran belts; their rough lands occur in patches and interrupted bands that rarely show marked complexity of geologic structure. *See* CORDILLERAN BELT.

Relationship to plate tectonics. The broad-scale pattern of crustal disturbance, and hence of rough lands, is unquestionably but complexly related to plate tectonics, that is, the relative movements of a worldwide set of immense crustal plates. Strong crustal deformation tends to be concentrated at or near plate margins, where adjacent plates are being jammed together, rifted apart, or moved laterally along a common border. Comparison of a world map of plate convergence zones with maps of earthquake

frequency and high mountains indicates that plate convergence zones are generally both earthquake-prone and mountainous, suggesting probable causal relationships. However, large areas of hill and mountain lands lie far from plate boundaries and hence cannot be explained by current plate convergence alone.

Where crustal plates converge, dense, flexible oceanic crust turns downward into the underlying mantle and usually undergoes partial melting at depth, often giving rise to volcanic activity above the sinking slab. However, low-density, more rigid continental crust resists sinking and tends to become intensely folded, thrust-faulted, and thickened in the convergence (sometimes in these circumstances called the collision) zone. The extensive cordilleran systems of southern Eurasia have been carved from complex structures formed by collision of the Eurasian plate with the African-Arabian and Australian plates, the crust involved being largely continental. In contrast, the island arcs of the Aleutians, the Kuriles, the Marianas, the eastern West Indies, and southern Indonesia are largely volcanic, having developed where convergence has involved only oceanic crust. Japan, the Andes, and the Cascade Range of the northwestern United States represent zones of convergence where oceanic crust from one plate has underrun continental crust from the other, producing structures combining complex deformation with prominent volcanic features. *See* EARTH CRUST.

At present, most plate divergence or rifting is associated with the extensive system of mid-oceanic ridges and is of little direct concern to continental terrain character. In a few areas, however, notably in the Red Sea–Gulf of Aden–East Africa rift zone, rifting extends into the continental blocks. Here the characteristic downfaulted troughs are in part occupied by arms of the sea or by lakes and are flanked by dissected bands of high ground, in places marked by volcanic flows and cones. A less extensive example is the Gulf of California trough and adjacent highlands in northwestern Mexico. *See* MID-OCEANIC RIDGE.

The many areas of hill or mountain terrain far from current plate margins are less obviously related to movements of the existing plate system. In some of these areas the geologic structures are similar to those now found at plate margins, but are ancient and have clearly been produced by convergence in a much older and very different plate system. The Appalachian–New England–Laurentian highland of North America and its structural continuations in northern Britain and Scandinavia are a good example. The Urals, eastern Australia, many of the mountainous areas of northern and eastern China, and parts of the uplands of eastern Brazil owe their existence to some form of late reuplift of the ancient structures. Other examples, notably the Rocky Mountains of the western interior of the United States and southern Canada, are even more difficult to explain by plate tectonics, for although their structures date from times since the present plate system came into being, they have developed far from plate margins. The broad, gentle upwarps of ancient shields or sedimentary platforms in plate interiors, such as the Ozark, Lake Superior, and Great Plains uplands of the United States, although simpler in form, are similarly perplexing.

Among the factors suggested to account for such intraplate structures and uplifts, sometimes acting in combination, are (1) isostatic adjustment resulting from continued erosional stripping of the thick, low-density crust of ancient structures; (2) unusually strong compressional and torsional stresses transmitted into the interior from the plate margins; and (3) movement of the crustal plates across areas of upwelling currents (plumes or hot spots) in the underlying mantle, with accompanying uplift, more or less volcanic activity, and differential movements of small sections of the crust. *See* EARTH, CONVECTION IN; EARTH INTERIOR; HOT SPOTS (GEOLOGY); PLATE TECTONICS.

Predominant surface character. The feature of hill and mountain lands are chiefly valleys and divides produced by sculpturing agents, especially running water and glacier ice. Local peculiarities in the form and pattern of these features reflect the arrangement and character of the rock materials within the upraised crustal mass that is being dissected.

Stream-eroded features. The principal features of most hill and mountain landscapes have been formed by stream erosion together with landslides and slower forms of gravity-induced movements (mass wasting) on the valley sides. The major differences between one rough land and another are in the size, shape, spacing and pattern of the stream valleys and the divides between them. These differences reflect variations in the original form and structure of the uplifted mass and in the ways in which erosion has been produced. *See* LANDSLIDE; MASS WASTING.

In consequence of uplift, streams have a large range of elevation through which they can cut, and as a result usually possess steep gradients early in their course of development. At this stage they are swift, have great erosive power, and are marked by many rapids and falls. Because of the rapid downcutting, the valleys are characteristically canyonlike with steep walls and narrow floors. At the same early stage, however, divides are likely to be high and continuous, and broad ridge crests common. In hill lands such ridges often provide easier routes of travel than do the valleys. In the Ozark Hills of Missouri, most of the highways and railroads follow such broad ridge crests. In high mountain country the combination of gorgelike valleys and continuous high divides makes crossing unusually difficult. In the Himalayas and the central Andes, and to a lesser degree in the Rocky Mountains of Colorado, the narrow canyons are very difficult of passage, and the divides are so continuously high and steep as to provide no easy pass routes across the ranges.

As erosional development continues, the major streams achieve gentle gradients and their valleys continue to widen. Divides become narrower and deep notches develop at valley heads, with well-defined peaks remaining between them. At this stage, which is represented by much of the Alps and by the Cascade Mountains of Washington, the principal valleys and the relatively low passes at their heads furnish feasible routes across the mountain belts. If other conditions are favorable, the wide valleys may afford significant amounts of cultivable or grazing land, as is true in the Alps. Still further erosion continues the widening of valleys and reduction of divides until the landscape becomes an erosional plain upon which stand only small ranges and groups of mountains or hills. The mountains of New England and many of the mountains and hill groups of the Sahara and of western Africa represent late stages in the erosional sequence.

Glaciated rough terrain. Glacial features of mountain and hill lands may be produced either by the work of local glaciers in the mountain valleys or by overriding glaciers of continental size. Continental glaciation has the general effect of clearing away crags, smoothing summits and spurs, and depositing debris in the valleys. The resulting terrain is less angular than is usual for stream-eroded hills, and the characteristic glacial trademark is the numerous lakes, most of them debris-dammed, a few occupying shallow, eroded basins. Examples of rough lands overridden by ice sheets are the mountains of New England, the Adirondacks, the hills of western New York, the Laurentian Upland of eastern Canada, and the Scandinavian Peninsula.

In contrast, mountains affected by local glaciers are made rougher by the glacial action. The long tongues of ice that move slowly down the valleys are excellent transporters of debris and are able to erode actively on shattered or weathered rock material. Valleys formerly occupied by glaciers are characteristically steep-walled and relatively free of projecting spurs and crags, with numerous broad cliffs, knobs, and shoulders of scoured bedrock. At their heads they generally end in cliff-walled amphitheaters called cirques (**Fig. 1**). The valley bottoms are commonly steplike, with stretches of gentle gradient alternating with abrupt rock-faced risers. Especially in the lower parts of the glaciated sections are abundant deposits of rocky debris dropped by the ice. Sometimes these form well-defined ridges (moraines) that run lengthwise along the valley sides or swing in arcs across the valley floor. Lakes are strung along the streams like beads, most of them dammed by moraines but some occupying shallow eroded basins. *See* MORAINE.

Because of rapid erosion and steepening of valley walls, either that effected directly by the ice or that attendant upon the exposure of the rock surface by continual removal of the products of weathering, glaciated mountains are likely to be unusually rugged and spectacular. This is true not only of such great systems as the Himalayas, the Alps, the Alaskan Range, and the high Andes, but also of such lesser ranges as those of Labrador, the English Lake District, and the Scottish Highlands. *See* EROSION; WEATHERING PROCESSES.

Most of the higher mountains of the world still bear valley glaciers, though these are not as large or as widespread as formerly (**Fig. 2**). Dryness and long, warm summers limit glaciers in the United States to a few large groups on the

Fig. 1. Head of a glaciated mountain valley ending in a cliff-walled amphitheater, called a cirque, showing a steplike valley bottom with lakes and waterfalls. (*Photograph by Hileman, Glacier National Park*)

Fig. 2. Upper reaches of Susitna Glacier, Alaska, showing cirques and snowfields. The long tongues of glacier ice carry bands of debris scoured from valley walls. (*Photograph by Bradford Washburn*)

Fig. 3. Hills in the western Appalachians, West Virginia. (*Photograph by John L. Rich, Geographical Review*)

higher peaks of the Pacific Northwest and numerous small ones in the northern Rockies. *See* GLACIATED TERRAIN.

Effects of geologic structure. Form and extent of the elevated areas, pattern of erosional valleys and divides, and, to some extent, sculptured details of slope and crest reflect geologic structure.

Some areas, such as the Ozarks, the western Appalachians, and the coast ranges of Oregon, are simply upwarped areas of roughly homogeneous rocks that have been carved by irregularly branching streams into extensive groups of hills or mountains (**Fig. 3**). Others, like the Black Hills or the ranges of the Wyoming Rockies, are domes or arched folds, deeply eroded to reveal ancient granitic rocks in their cores and upturned younger stratified rocks around their edges. The Sierra Nevada of California is a massive block of the crust that has been uplifted and tilted toward the west so that it now displays a high abrupt eastern face and a long canyon-grooved western slope. The central belt of the Appalachians displays long parallel ridges and valleys that have been hewn by erosion out of a very old structure of parallel wrinkles in the crust. The upturned edges of resistant strata form the ridges; the weaker rocks between have been etched out to form the valleys. The Alps and the Himalayas are eroded from folded and broken structures of incredible complexity involving almost all varieties of rock materials.

Most volcanic mountains, like the Cascades of the northwestern United States or the western Andes of Peru and Bolivia, are actually erosional mountains sculptured from thick accumulations of lava and ash. In these areas however, modern eruptive vents give rise to volcanic cones that range from small cinder heaps to tremendous isolated mountains. The greater cones, such as Fuji, Ararat, Mauna Loa, Rainier, or Shasta, are among the most magnificent features of the Earth's surface. *See* MOUNTAIN; MOUNTAIN SYSTEMS; TERRAIN AREAS; VOLCANO.

Edwin H. Hammond

Bibliography. K. W. Butzer, *Geomorphology from the Earth*, 1977; A. Doerr and J. F. Coling, *Fundamentals of Physical Geography*, 1993; G. H. Dury, *An Introduction to Environmental Systems*, 1981; P. Keary and F. J. Vine, *Global Tectonics*, 1990; D. P. McKenzie, The Earth's mantle, J. Francheteau, The Oceanic crust, and B. C. Burchfiel, The continental crust, *Sci. Amer.*, pp. 66–145, September 1983; G. T. Trewartha et al., *Fundamentals of Physical Geography*, 3d ed., 1977.

Hiodontiformes

An order formerly in Osteoglossiformes but now independent, with the two orders comprising the extant Osteoglossomorpha, a primitive group of teleost fishes. The Hiodontiformes consist of one family, one genus, and two extant species, *Hiodon alosoides* (goldeye) and *H. tergisus* (mooneye). Hiodontids superficially resemble clupeids and for many years were classified as clupeiforms. The body is deep and laterally compressed and has a ventral keel, but the keel is not serrated as in most clupeids; the eyes are far forward and large, with the diameter greater than the length of the snout; the scales are large and cycloid and either bright silvery or golden; the anal fin base is much longer than the dorsal fin base; and the caudal fin is forked. The goldeye differs from the mooneye in having 9 or 10 principal dorsal fin rays versus 11 or 12, and a ventral keel extending anteriorly past the pelvic fins versus not extending anteriorly past the pelvic fins. *See* CLUPEIFORMES; OSTEICHTHYES; OSTEOGLOSSIFORMES.

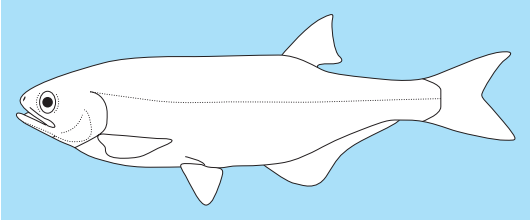

Example of a hiodontid. (*From J. S. Nelson, Fishes of the World, 4th ed., Wiley, New York, 2006*)

Both species are limited to the freshwaters of North America between the Continental Divide and the Appalachians. *Hiodon alosoides* has the greatest north-south distribution of North American freshwater fishes, ranging from 69° to 31°N latitude and including the MacKenzie River delta and Great Plains of northwestern Canada, isolated populations in the Hudson River drainage in eastern Canada, and populations southward in large rivers of the Mississippi Basin to Louisiana and Alabama. The species is absent from the Great Lakes basin. *Hiodon tergisus* ranges between latitudes 60° and 30°N, including the Mississippi River basin east of the Great Plains and west of the Appalachians to the Gulf Coast, parts of the Great Lakes, the St. Lawrence River, and tributaries to Hudson Bay northward.

Both species are opportunistic feeders, foraging in late evening and at night, at the surface, primarily on

aquatic insects, crustaceans, mollusks, small fishes, and frogs. *Hiodon alosoides* attains a total length of 52 cm (20.5 in.), and *H. tergisus* a total length of 47 cm (18.5 in.). Herbert Boschung

Bibliography. E. J. Hilton, Comparative osteology and phylogenetic systematics of fossil and living bony tongue fishes (Actinopterygii, Teleostei, Osteoglossomorpha), *Zool. J. Linn. Soc.*, 137:1–100, 2003; G.-Q. Li and M. V. H. Wilson, Phylogeny of Osteoglossomorpha, pp. 163–174 in M. L. J. Stiassny, L. R. Parenti, and G. D. Johnson (eds.), *Interrelationships of Fishes*, Academic Press, San Diego, 1996; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Hippopotamus

The name for two species of even-toed (artiodactylid) ungulates which form the family Hippopotamidae. Both species occur in Africa; the great African hippopotamus (*Hippopotamus amphibius*) inhabits the rivers of tropical Africa, and the pygmy hippopotamus (*Choeropsis liberiensis*) lives near the rivers of western Africa but is more terrestrial.

The great African hippopotamus, the largest living artiodactylid, reaches a length of 18 ft (5.5 m), stands about 5 ft (1.5 m) at the shoulder, and may weigh up to 4 tons (3.6 metric tons; see **illustration**). The ears are small and flexible, and the nostrils and eyes protrude so that they are out of the water as the animal floats. The skin may be 2 in. (5 cm) thick in places and almost devoid of hair. The tail is quite short and has some hair, as does the muzzle and inside of the ears. The feet end in four toes enclosed in round hoofs. The lateral toes are almost as well developed as the median ones, and all four toes touch the ground—a primitive condition. Both canines and incisors continue to grow during the life of these animals, a condition similar to that of rodents. These teeth are few in number and are long only in the lower jaw. The canines, which are curved and resemble tusks, may reach a length of 2 ft (0.6 m) and weigh 5–6 lb (2.3–2.7 kg). The other teeth are hidden behind the lips, but there are numerous powerful molars. The dental formula is I 2/2 C 1/1 Pm 4/4 M 3/3 for a total of 40 teeth. These animals have a gestation period of about 34 weeks, and a single calf is born.

The pygmy hippopotamus is about the size of a large pig and may weigh up to 400 lb (180 kg). It is a more solitary species and does not live in large herds, as does the great African species. The dental formula differs from that of *H. amphibius*, being I 2/1 C 1/1 Pm 4/4 M 3/3, for 38 teeth. *See* DENTITION.

These animals migrate regularly, following the river upstream during the rainy season and downstream during the dry season to new pastures. They come onto land, especially at night, to feed on vegetation. The males of the great African hippopotamus occupy and maintain territories within which are small herds of females and juveniles.

Hippopotamus (*Hippopotamus amphibius*). (*Photo by Gerald and Buff Corsi; © California Academy of Sciences*)

The closest living relatives are the pigs. *See* ARTIODACTYLA. Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999.

Hirudinea

A subclass of the annelid worms commonly known as leeches. These organisms are parasitic or predatory and have terminal suckers for attachment and locomotion. Most inhabit inland waters, but some are marine and a few live on land in damp places. The majority feed by sucking the blood of other animals, including humans.

Morphology. Leeches differ from other annelids in having the number of segments in the body fixed at 34, chaetae or bristles lacking, and the coelomic space between the gut and the body wall filled with packing (botryoidal) tissue. In a typical leech the first six segments of the body are modified to form a head, bearing eyes, and a sucker, and the last seven segments are incorporated into a posterior sucker (see **illus.**). Each segment is divided externally into 2–16 rings or annuli. The number per segment is constant for a particular species in the midbody region, but decreases toward the extremities.

The mouth of a leech opens within the anterior sucker, and there are two main methods of piercing the skin of the host to obtain blood. The rhynchobdellid leeches have the anterior part of the gut modified to form an eversible proboscis. The anterior sucker is planted firmly in position on the skin of the host, and the proboscis is forced out of the leech's mouth and into the host tissue. The arhynchobdellid leeches lack the proboscis, and in its place have three jaws, each shaped like half a circular saw, placed just inside the mouth. When the anterior sucker is placed in position on the skin of the host, the three jaws are pushed foward and rocked in such a way

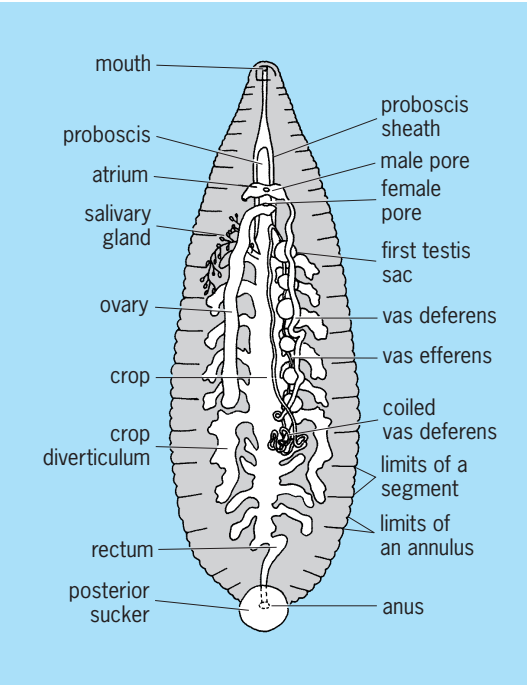

General structure of a leech. Male reproductive system is shown on the right, the female on the left. (*After K. H. Mann, A key to the British freshwater leeches, Freshwater Biol. Ass. Sci. Publ., 14:3–21, 1954*)

as to cause a Y-shaped incision. Salivary glands are present in both orders of the leeches, and their secretion helps to prevent the clotting of blood. The jawed leeches are able to penetrate the skin of humans and other mammals, but a proboscis is less efficient, and rhynchobdellid leeches usually have to seek out softer tissues such as the lining of the nostril of a bird, or the gills of a fish. Some arhynchobdellid leeches have given up bloodsucking and feed instead on such animals as earthworms and insects. In these leeches, the jaws are often reduced to muscular ridges that are used merely to grasp the prey, which is normally swallowed whole. *See* ARHYNCHOBDELLAE; RHYNCHOBDELLAE.

The bloodsucking leeches have a large region of the gut, the crop, modified for the storage of blood. It consists of a central chamber with a number of paired diverticula, or side branches. When a meal of blood is taken, the crop becomes considerably distended, so that some leeches can store up to 10 times their own weight in blood.

The process of digestion is very slow, and a meal may last a leech for 9 months. The carnivorous forms have lost most or all of their gut diverticula and resemble earthworms in having a straight, tubular gut.

Life cycle. Leeches are hermaphroditic, having a single pair of ovaries and several pairs of testes. The latter are arranged on either side of the gut, and the sperms pass forward in paired ducts to a single midventral opening placed about one-third of the distance from the anterior to the posterior sucker. In many leeches there is an eversible penis, and sperms are transferred directly to the female pore of another leech. However, in others the penis is lacking, and the sperms are made up into packets, or spermatophores, which are attached to the body surface of another leech. From here the sperms migrate through the tissues to fertilize the eggs. The female pore is in the midventral line, a short distance behind the male pore, and at the time of egg laying a thickened, glandular region of the body wall, the clitellum, secretes a barrel-shaped cocoon around the body, over the genital pores. The fertilized eggs are passed into this cocoon, after which the leech works the cocoon over its head and then closes the ends. Some leeches loosely deposit the cocoons in a damp place, others cement them to stones or vegetation under water, while still others place them under their bodies and brood over them. In the last case the young, on hatching, attach themselves to the undersurface of the parent and are carried about for some time. The method of reproduction is very like that of the earthworm, and leeches may be regarded as earthworms which have become modified for the parasitic mode of life. *Acanthobdella* is regarded as a transitional form. It has only 30 segments in the body, of which 5 bear bristles, or chaetae. There is no anterior sucker, and the coelomic space is not entirely filled by packing tissue. *See* OLIGOCHAETA.

Economic importance. The importance of leeches as a means of making incisions for the letting of blood or the relief of inflammation is declining, and in civilized countries the bloodsucking parasites of mammals are declining, because of lack of opportunity for contact with the hosts. In other countries they are still serious pests. *Limnatis nilotica* of the Middle East enters the mouth or nostrils of animals or humans drinking at streams, and causes bleeding, nausea, and vomiting, often followed by death from anemia or suffocation. The leech may sometimes be removed, with difficulty, by the use of irritants such as salt water, chloroform water, or vinegar. *Hirudo* is widespread in Europe and Asia. Its attacks on humans and domestic animals are usually not serious, since they usually occur on the body or the limbs, but occasionally serious damage is caused by attacks on delicate tissues such as the eyes. In South America the attacks of *Haemadipsa chiliani* on horses and cattle can be fatal. *Theromyzon*, which enters the throat and nostrils of water birds, may cause serious losses of domestic ducks and geese when crowded conditions lead to an increase of the leech population. The fish parasite *Piscicola* may be a serious pest in hatcheries or heavily stocked ponds.

On the other hand, many of the leeches which are carnivorous in habit form an important contribution to the diet of fresh-water fishes. Records from the stomach contents of fish do not often show this, because of the rapid and complete digestion of the leeches. *See* ANNELIDA.

Kenneth H. Mann

Bibliography. K. J. Muller et al. (eds.), *Neurobiology of the Leech*, 1981; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; R. T. Sawyer, *Leech Biology and Behavior*, 3 vols., 1986.

Histamine

A biologically active amine that is formed by the decarboxylation of the amino acid histidine. It is widely distributed in nature and is found in plant and animal tissues as well as in insect venoms. In humans, histamine is a mediator of inflammatory reactions, and it functions as a stimulant of hydrochloric acid secretion in the stomach.

Most tissue histamine is found stored in mast cells, where it can be released by a variety of stimuli. Once released, it can cause many effects, including constriction of bronchiolar, gastrointestinal, uterine smooth muscle, and lowering of blood pressure. If histamine is released in the skin, itching, a flare (area of redness) due to vasodilation, and a wheal due to leaking of fluid into the tissue are observed. The increase in vascular permeability that permits this leakage is due to an action on the endothelial cells of postcapillary venules.

All of these actions of histamine are mediated by the activation of histamine receptors, designated either H-1 or H-2. Antihistamine drugs exert their effects by blocking the combination of histamine with these receptors. *See* ANTIHISTAMINE.

Histamine release can be caused by tissue injury, by physical stimuli such as cold or pressure, by drugs such as heroin, and most importantly by immunologic mechanisms. Mast cells in the skin, the lung, the nasal passages, or other sites may become sensitized to antigens such as ragweed or other pollens, and then release histamine and other biologically active substances upon exposure to them. The released histamine may then cause wheals, bronchoconstriction, itching, runny nose, and other effects commonly associated with allergic responses. If the allergic reaction becomes generalized and severe, life-threatening anaphylactic shock may ensue. The prompt administration of epinephrine, which exerts effects opposite to those of histamine, can be life-saving in such cases. *See* ALLERGY; ANTIGEN; EPINEPHRINE; IMMUNOLOGY. Alan Burkhalter

Bibliography. M. Garcia-Caballero, L. J. Brandes, and S. Hosada (eds.), *Histamine in Normal and Cancer Cell Proliferation*, 1993; J. C. Schwartz and H. Haas, *The Histamine Receptor*, 1992.

Histocompatibility

A term used to describe the genes that influence acceptance or rejection of grafts. When grafts of tissue are exchanged between genetically dissimilar individuals, profound immunological rejection generally takes place. In contrast, grafts between genetically similar individuals, such as identical twins, are normally tolerated; they are histocompatible. Most known examples of histocompatibility (or H) genes encode polymorphic (that is, tending to differ between individuals) cell-surface proteins. The major histocompatibility complex (MHC) contains a set of histocompatibility genes, termed major because mismatching at these genes invokes rapid rejection. Other, minor antigens are revealed after precise matching of the major antigens.

The main function of MHC genes involves distinguishing self from nonself in the immune system, as part of preventing the spread of infectious disease. The body employs special mechanisms to avoid rejection of the fetus, which is effectively an allograft, that is, a graft from a donor to a genetically dissimilar recipient of the same species; in this case, the mechanisms include a diminution of MHC gene expression.

Genetics. Histocompatibility antigens are encoded by genes that are inherited in a codominant, Mendelian fashion. The MHC, the locus to which the main barriers to tissue rejection map, contains a spectrum of genes, many of which influence processing and presentation of antigens to the immune system. In mice, the MHC is designated the H-2 complex; in humans, it is referred to as the HLA complex (for human leukocyte A system). Mice and other mammals seem to have a similar arrangement of genes in their MHCs. Other vertebrate MHCs have been identified, such as fish, chicken, and amphibian, and these are beginning to show some differences in arrangement of their genes. *See* MENDELISM.

The human MHC is normally divided into three regions, class I, II, and III. Located at the middle of the chromosome, the class II region contains several sets of related genes encoding the class II antigens (see **illus.**). The class II region also contains a group of genes that seem to be involved in processing antigens that are destined to be presented on class I molecules.

In the direction of the tip of the chromosome arm, the class III region consists of a mixture of over 40 genes, many of which bear no relationship to the immune system. Exceptions include genes for cytokines such as tumor necrosis factor. The class I region contains the *HLA*, *B*, and *C* genes as well as over 16 related loci. The class I genes are interspersed with a large number of genes of nonimmune function. *See* ANTIGEN; BLOOD GROUPS; CELLULAR IMMUNOLOGY; CHROMOSOME; CYTOKINE.

Polymorphism. Some of the MHC class I and class II genes have the remarkable property of being highly variable between individuals. Thus at each of the HLA loci there are multiple variant genes (alleles) in the population. Because humans are diploid, each individual can have two alleles at each locus. Each allele encodes one HLA antigen. Thus, an individual expresses two different HLA antigens from each series. The HLA system exhibits extreme polymorphism because there are so many different alleles (and corresponding antigens) in the population. When the alleles at each locus are considered in combination, millions of different HLA phenotypes can be generated, and it is rare to find unrelated individuals who share all the same antigens encoded at the HLA loci; hence the difficulties in tissue matching for transplantation. Siblings inherit one set of MHC genes (together called a haplotype) from each parent. Thus, their chances of having identical

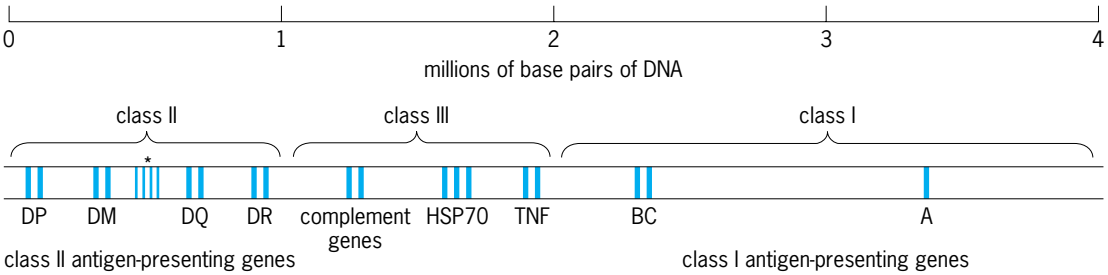

Arrangement of genes in the human major histocompatibility complex on the short arm of chromosome 6.

haplotypes is greater than for the general population. Transplant of a kidney from an HLA-identical (and ABO-identical) sibling has a high survival potential because there are no HLA incompatibilities.

Precise typing of HLA antigens or genes in tissue-typing laboratories aims to maximize matching organs from unrelated donors. Precise matching is critical in bone marrow transplantation, where there is a high load of lymphoid tissue and graft-versus-host reactions. It is progressively less important in transplantation of small bowel, kidney, heart, lung, and liver. However, in many cases, the use of powerful immunosuppressant drugs circumvents this problem. *See* IMMUNOSUPPRESSION; MUTATION; POLYMORPHISM (GENETICS).

Function. The MHC class I and class II products present fragments of pathogens to the immune system, in particular, to T cells.

In general, class I molecules bind peptides in the groove at the top end of the molecule. These peptides are derived from proteins that have been synthesized in the cytoplasm of the cell. The T cell contains a cell-surface molecule called a T-cell receptor which can bind to the class I molecule. Each T cell expresses a different receptor, that has been previously selected in the thymus to recognize a self MHC molecule only if it contains a nonself peptide. If the particular T-cell receptor binds to the MHC-peptide combination, the presenting cell is then killed.

An additional set of genes within the class II region, the *LMP* and *TAP* genes, are thought to be involved in processing antigen. The *LMP* gene products form part of a complex termed the proteasome, which cleaves proteins into peptides. These peptides then have to be transported across an internal membrane into the correct compartment where they encounter class I molecules. *TAP* gene products are responsible for this part of the process.

Class II molecules also bind peptides, usually those derived from proteins outside the cell. The foreign protein may be bound to the cell surface by means of an immunoglobulin molecule, expressed on a B cell. The foreign protein is then taken to a cellular compartment where it is broken into peptides capable of binding into the groove of the class II molecule. The class II molecule together with the foreign peptide, if recognized by an appropriate T cell, stimulates the T cell to produce cytokines which will cause the B cell to proliferate and produce immunoglobulin to neutralize the foreign protein.

The structure of class I and class II molecules clarifies their polymorphic quality. The amino acids that are most variable between alleles occupy those positions on the molecule that point into the peptide binding groove. This enables different individuals to bind different sets of peptides from diverse organisms. The inherent flexibility of the system is presumably important to impede the spread of pathogens since agents such as influenza virus or the human immunodeficiency virus (HIV) throw up altered protein sequences from time to time. Proteins from these escape mutants may not provide optimum peptides for binding into the grooves of certain MHC molecules, but their spread is checked by at least some of the many variant structures in the population that do elicit binding. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); IMMUNOLOGICAL ONTOGENY; INFLUENZA.

HLA and disease. Alleles of the various MHC genes are associated with a large number of diseases. These fall into three categories: autoimmune diseases, infectious diseases, and genetically linked diseases. Most common are the autoimmune diseases such as insulin-dependent diabetes mellitus, rheumatoid arthritis, and multiple sclerosis; these tend to occur more often with certain alleles. For example, multiple sclerosis tends to occur more frequently in individuals who possess the HLA-DR2 antigen. Ankylosing spondylitis, a disorder of the spine, is strongly associated with HLA-B27 individuals. Any explanation for these associations likely reflects the known biological function of HLA molecules in presenting antigens to the immune system. The second category of MHC-associated diseases is the infectious type; malaria is an example, but compelling evidence for others is lacking in humans. The third category consists of diseases for which the responsible gene is genetically linked to the MHC but is not thought to be involved with the immune system. There is a gene near the MHC alteration which disturbs iron uptake in the gut, causing hemochromatosis (that is, excessive accumulation of iron in the liver), for example. Another disease association is narcolepsy, in which affected individuals fall asleep involuntarily. It is almost exclusively associated with the HLA-DR2 antigen, although the reasons behind this remain obscure.

The associations between HLA and disease are important because of the insight that they provide as to the disease etiology. However, the HLA region

provides just one of the factors influencing the course of these diseases, and important roles are also played by other unlinked genes located on other chromosomes, as well as by poorly understood environmental factors such as past and present microbial infection. *See* AUTOIMMUNITY.

Geography. HLA haplotypes are important demographic markers effective in tracing the spread of populations. For example, it is possible to monitor certain HLA alleles of ancient populations throughout the Americas, some of which are found in Asian populations. Their distribution is consistent with theories of migration across the Bering Straits from northeast Asia. In some isolated South American Indian populations, new *HLA-B* alleles seem to have arisen by recombination, possibly as a direct result of high pathogen loads during periods of their history. It is possible to trace so-called Celtic haplotypes through Europe, and to the west and north of the British Isles. By polymerase chain reaction, it is possible to analyze cadaveric samples, from bones or mummies, which may provide further clues to population movement. *See* GENETIC MAPPING.

Xenotransplantation. There is much interest in transplanting organs between species, using other mammals such as pigs as organ donors for humans. MHC incompatibility may not be as marked in this situation as in allotransplantation (that is, mismatching in the same species). This may be partly due not only to poor class I and class II recognition between species but also to incompatible surface adhesion molecules necessary for T-cell stimulation. There are a number of other problems, though, such as naturally occurring antibody to carbohydrate antigens, which may differ markedly between species. These antibodies can cause hyperacute rejection within minutes. It has been estimated that humans and Old World monkeys commit up to 1% of their antibody repertoire to a response against a carbohydrate present on other species studied for transplantation. *See* ANTIBODY; GENETICS; HUMAN GENETICS; TRANSPLANTATION BIOLOGY. John Trowsdale

Bibliography. J. Klein and U. Klein (eds.), *Molecular Evolution of the Major Histocompatibility Complex*, 1991; R. Srivastava, G. P. Ram, and P. Tyle (eds.), *Immunogenetics of the Major Histocompatibility Complex*, 1991; A. R. Williamson and M. W. Turner, *Essential Immunogenetics*, 1987.

Histogenesis

The developmental processes by which the definite cells and tissues which make up the body of an organism arise from embryonic cells. Among animals, the ectoderm, endoderm, and mesoderm, also known as the primary germ layers, provide the stem cells which gradually transform into distinctive kinds of cells and tissues. In the higher plants, meristematic cells, which occur wherever extensive growth takes place, provide the basis for tissue formation. Parenchyma, a simple permanent tissue, may dedifferentiate and become meristematic. *See* APICAL MERISTEM; EMBRYOLOGY; GERM LAYERS; HISTOLOGY; LATERAL MERISTEM; PARENCHYMA; PLANT TISSUE SYSTEMS. Charles B. Curtin

Histology

The study of the structure and chemical composition of tissues of animals and plants as related to their function. The primary aim is to understand how tissues are organized at all structural levels, including the molecular and macromolecular, the entire cell and intercellular substances, and the tissues and organs.

The four tissues of the animal body include cells and intercellular substances. They are (1) epithelium, in which the cells are generally closely applied to each other and separated by very little intercellular substance; (2) connective tissue, in which the cells are usually separated by greater amounts of intercellular substance, which may indeed form the great bulk of the tissue; (3) muscular tissue, whose cells are primarily concerned with contractility; and (4) nervous tissue, whose components are concerned primarily with rapid conduction of impulses.

The tissues are combined in various ways to form integrated patterns which are characteristic of every organ of the body. In this way, imperceptibly, the study of the minute structure and relations of organs merges into gross anatomy.

The structures studied in histology extend from those which are just too small to be seen with the hand lens to those which are beyond the resolution of the electron microscope, for example, those which measure about 2 nanometers. For this purpose, a wide range of instruments has been used. These include x-ray diffraction units, x-ray absorption microscope, electron microscope, polarization microscope, dark-field microscope, ultraviolet microscope, visible light microscope, phase contrast and interference microscope, and the dissecting microscope. *See* MICROSCOPE.

In earlier days, fresh material to be examined by the histologist was either teased or sliced freehand, scraped, or smeared so that it would be thin enough to be viewed with a microscope by transmitted light. Later, microtomes were introduced which could cut sections as thin as 1 micrometer; now they are capable of cutting sections 7 nm in thickness. The period of use of the microtome corresponds roughly with the use of fixatives, which were introduced to preserve and retain structure. The structural rearrangements, distortions, and losses caused by the use of fixatives and the use of irrelevantly esthetic stains eventually caused a reaction among histologists. This took the form of a return to the study of fresh material under strictly controlled conditions, the development of tissue culture and micromanipulation, microcinematography, intravital staining (of living cells), and supra vital staining (of surviving cells). Another practical development was fixation by freezing and drying. In this method of

preservation specimens are frozen very rapidly by immersion in fluids cooled to −292°F (−180°C) or less, dried in a vacuum at a temperature of −58°F (−50°C) or less, and later infiltrated in paraffin. *See* MICROTECHNIQUE; TISSUE CULTURE.

The major fields of histological studies are (1) morphological descriptions; (2) developmental studies; (3) histo- and cytophysiology; (4) histo- and cytochemistry; and (5) fine (or submicroscopic) structure. Histo- and cytophysiology deal with correlations between morphological changes and functional activity. Histo- and cytochemistry deal with the chemical composition of morphological structures. The study of fine structure deals with the arrangements of structures which are below the resolution of the light microscope (about 0.2 μm). *See* CONNECTIVE TISSUE; EPITHELIUM; MUSCULAR SYSTEM; NERVOUS SYSTEM (VERTEBRATE). Isidore Gersh

Bibliography. D. H. Cormack, *Ham's Histology*, 9th ed., 1987; D. Fawcett, *Bloom and Fawcett: A Textbook of Histology*, 12th ed., 2001; L. C. Junqueira et al. (eds.), *Basic Histology,* 9th ed., 1998; T. S. Leeson, C. R. Leeson, and A. A. Paparo, *Text/Atlas of Histology*, 1988; I. R. Telford and C. F. Bridgman, *Introduction to Functional Histology*, 2d ed., 1994; L. Weiss (ed.), *Cell and Tissue Biology: A Textbook of Histology*, 6th ed., 1988.

Historadiography

The technique for taking x-ray pictures of cells, tissues, or sometimes the whole animal or plant, if it is a small one. Soft x-rays, those with low penetrating power and relatively long wavelengths, are required for this type of picture. The best pictures are obtained when the tissues contain deposits of metallic elements which have a high absorption capacity for x-rays. *See* X-RAYS.

In applying the technique to tissues, a relatively thin section is placed against an x-ray film and irradiated with a beam of x-rays. When the film is developed, a picture of the object or section of tissue shows on the film. Another method attempts to focus the x-rays after they pass through the specimen. In this arrangement, a lens is placed between the specimen and the film. X-rays are very difficult to focus. The lenses must be the reflecting type of mirror surfaces. A simple two-mirror system has two cylindrically curved surfaces set at right angles to each other. The glancing angle must be small to get reflection of the x-rays. The lenses also have certain aberrations that limit the resolution to about 0.2 micrometer, which is no better than the resolution of the best light microscope images. However, x-rays with their short wavelengths are potentially capable of much higher resolution if they could be focused without much aberration. The advantage of the x-ray picture is that it may resolve detail not visible with light in calcified bone tissues or in tissues injected with an x-ray- absorbing material.

An interesting use of historadiography in biology is the determination of the relative mass of cellular structures. The tissue is quick-frozen and dried at low temperatures. Then sections are prepared and photographed by the method described above. A relation exists between the mass of the various parts of the specimen and the contrast of the film, that is, the number of silver grains or the amount of silver deposited. Quantitative determinations can be made by scanning the photograph with a densitometer. *See* NUCLEAR RADIATION (BIOLOGY). J. Herbert Taylor

Hodgkin's disease

A malignant lymphoid neoplasm, usually arising in lymph nodes characterized by morphological heterogeneity and bizarre giant tumor cells referred to as Reed-Sternberg cells. The etiology of Hodgkin's disease is unknown, although current epidemiological data suggest an infectious etiology. The disease has been recognized in all or several members of "close-knit" groups, usually occurring in these persons several years after they have separated from one another. As in non-Hodgkin lymphomas, there is morphological and immunological evidence to suggest a viral etiology. *See* ANIMAL VIRUS.

Reed-Sternberg cells. Hodgkin's disease is diagnosed by finding Reed-Sternberg cells in histologic sections of involved lymph nodes. Classically, Reed-Sternberg cells are greater than 40 micrometers in diameter and have large bilobed nuclei and very prominent eosinophilic nucleoli surrounding by a halo. Several variants of Reed-Sternberg cells have been identified.

Classification. Until 1966 Hodgkin's disease was classified morphologically as paragranuloma, granuloma, and sarcoma. In 1966 R. Lukes and coworkers reclassified Hodgkin's disease into six categories using the following morphologic criteria: (1) the number of lymphocytes or histiocytes, or both, present in the involved lymph node; (2) the type, number, and distribution of the Reed-Sternberg cells; and (3) the amount and distribution of the associated connective tissue in the diseased tissue. The disease was then reclassified into four major groups: lymphocyte predominant; nodular sclerosis; mixed cellular; and lymphocyte depleted. Each of the four groups has a distinct microscopic appearance and different clinical course, Progression from one subtype to another is not uncommon.

Clinical features. Persons with Hodgkin's disease usually seek medical advice because of enlarged painless lymph nodes. They may also have fever, weight loss, anorexia, pruritus, and anemia. The clinical extent of the disease is determined by a process of staging based on physical examination, various biopsies, and usually a laparotomy for examination of the spleen and liver. The disease is divided into the following stages: stage 1—disease limited to one anatomic region or two contiguous anatomic regions on the same side of the diaphragm; stage 2—disease in more than two anatomic regions or in two noncontiguous regions on the same side of the diaphragm; stage 3—disease on both sides of

the diaphragm but not extending beyond the involvement of lymph nodes, spleen, or Waldeyer's ring; stage 4—involvement of the bone marrow, lung parenchyma, pleura, liver, bone, skin, kidneys, gastrointestinal tract, or any organ in addition to lymph nodes, spleen, or Waldeyer's ring. Each stage is also classified as A or B depending on whether the patient has systemic symptoms of disease.

Susceptibility. Hodgkin's disease more commonly affects males than females, except for the nodular sclerosing variety, which occurs with equal frequency in both sexes. It generally is a disease of persons between the ages of 20 and 40, but it may affect the very young and the very old. Characteristically, persons with Hodgkin's disease exhibit a loss of cell-mediated immunity and become susceptible hosts for infection with a variety of microorganisms such as tubercle bacilli.

Cellular origin. The cellular origin of Hodgkin's disease remains uncertain. One hypothesis suggests that it is caused by a defective T-lymphocyte immune system which permits the development of neoplastic Reed-Sternberg cells. *See* CELLULAR IMMUNOLOGY.

Treatment. The treatment of Hodgkin's disease is dependent on its clinical stage and microscopic appearance. A combination of chemotherapy (anticancer drugs) and radiation therapy is commonly used. The best prognosis is in those persons with stage 1 disease who have the lymphocyte predominant or nodular sclerosing variety. Very encouraging therapeutic results have occurred, with many apparent cures. *See* CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; LYMPHATIC SYSTEM; RADIOLOGY.

Samuel P. Hammar

Bibliography. E. Braunwald et al. (eds.), *Harrison's Principles of Internal Medicine*, 15th ed., 2001; F. Cavalli, G. Bonodonna, and M. Rozencweig (eds.), *Malignant Lymphomas and Hodgkin's Disease: Experimental and Therapeutic Advances*, 1986; B. W. Dana (ed.), *Malignant Lymphomas, Including Hodgkin's Disease: Diagnosis, Management, and Special Problems*, 1993.

Hoisting machines

Mechanisms for raising and lowering material with intermittent motion while holding the material freely suspended. Hoisting machines are capable of picking up loads at one location and depositing them at another anywhere within a limited area. In contrast, elevating machines move their loads only in a fixed vertical path, and monorails operate on a fixed horizontal path rather than over a limited area. *See* ELEVATING MACHINES; MONORAIL.

The principal components of hoisting machines are sheaves and pulleys, for the hoisting mechanism; winches and hoists, for the power units; and derricks and cranes, for the structural elements.

Block and tackle. Sheaves and pulleys or blocks are a means of applying power through a rope, wire, cable, or chain. Sheaves are wheels with a grooved periphery, in appropriate mountings, that change the direction or the point of application of a force transmitted by means of a rope or cable. Pulleys are made up of one or more sheaves mounted in a frame, usually with an attaching swivel hook, eye, or similar device at one or both ends. Pulley systems are a combination of blocks.

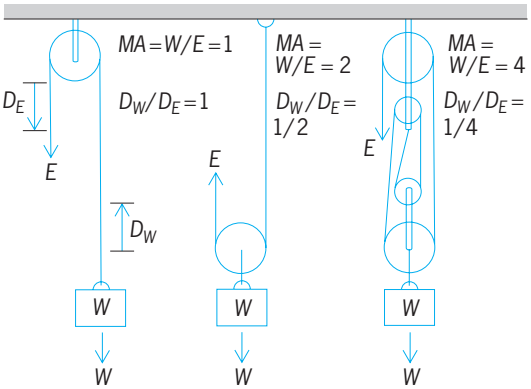

Fig. 1. Mechanical advantage of pulley systems.

The mechanical advantage MA of a pulley system is the ratio of weight lifted W to input effort exerted E. The mechanical advantage, neglecting friction, for any of various arrangements can be determined readily because it equals the number of strands that support the load (**Fig. 1**). The ratio of distance D_W through which the weight is lifted to distance D_E through which the effort is expended is inversely proportional to the mechanical advantage.

Sometimes used alone, sheaves and pulleys find their most usual application as the hoisting tackle of derricks and cranes.

Winches and hoists. Normally, winches are designed for stationary service, while hoists are mounted so that they can be moved about, for example, on wheel trolleys in connection with overhead crane operations.

A winch is basically a drum or cylinder around which cordage is coiled for hoisting or hauling. The drum may be operated either manually or by power, using a worm gear and worm wheel, or a spur gear arrangement. A ratchet and pawl prevent the load from slipping; large winches are equipped with brakes, usually of the external band type. Industrial applications of winches include use as the power element for derricks and as the elevating mechanism with stackers (**Fig. 2**).

Floor- and wall-mounted electric hoists are used for many hoisting and hauling jobs from fixed locations in industrial plants and warehouses. Heavy-duty types are standard equipment for powering ship's gear in cargo handling. They are also mounted on over-the-road carriers to facilitate the moving of heavy bulky loads, and serve as the power units of power cranes and shovels. A railroad car puller employs the same principle; however, the drum is mounted vertically and is used for spotting railroad cars in freight yards.

Hoists are designed to lift from a position directly above their loads and thus require mobile mountings.

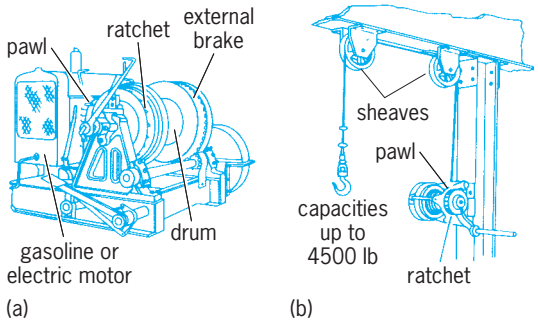

Fig. 2. Powered and hand winches. (*a*) Heavy-duty single-drum winch. (*b*) Wire-rope hand winch. 4500 lb = 2042 kg.

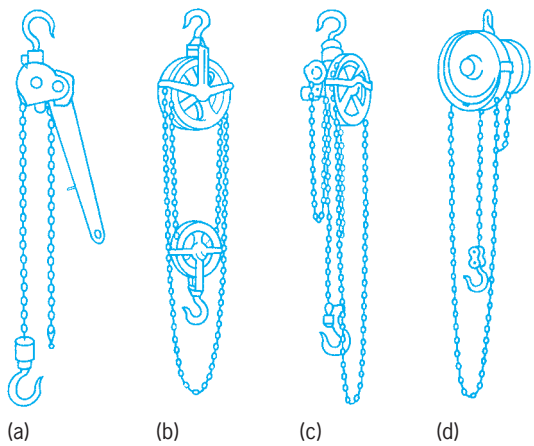

Fig. 3. Chain hand hoists. (*a*) Lever (ratchet). (*b*) Differential. (*c*) Worm gear. (*d*) Spur gear.

Hoists are classified by their power source, such as hand, electric, or pneumatic.

Hand hoists are chain operated. There are four types: spur geared, worm geared, differential, and pull-lift or lever (**Fig. 3**). The last, with its lever for operation and its ratchet for holding, is the simplest and most economical. However, since the operating lever is located on the anchor end of the hoist, it is not as convenient for vertical lifting as it is for horizontal pulling.

The spur-gear type costs the most but is the most economical to operate, with an efficiency as high as 85%. Where hoists are to be used frequently, it is the type most recommended.

Worm-gear hoists transform about one-third to one-half the input energy to useful work. Offsetting this low efficiency is the locking characteristic of the worm drive: The load cannot turn the mechanism; consequently the load is at all times restrained from running away. In contrast, with a ratchet the load is locked only at positions where the pawl engages a step of the ratchet. Differential hoists use only about one-third the energy input; they too prevent loads from running away during lowering. Spur-gear hoists are more efficient, but require a brake to restrain loads during lowering or holding.

Electric hoists lift their loads by either cable or chain (**Fig. 4**). The cable type has a drum around which a wire cable is coiled and uncoiled for hoisting and lowering. Chain models have either a roller chain and sprocket or a link chain and pocketed wheel for hoisting and lowering.

There are innumerable below-the-hook attachments, such as slings, hooks, grabs, and highly specialized devices to facilitate practically any handling requirement. Many of these devices are designed to pick up and release their loads automatically. All chain hoists are designed with the lower hooks as the weakest parts; not being interchangeable with the anchor hook, therefore, if the hoist is overloaded, the lower hook spreads or opens up. If the inside contour of the hook is not a true arc of a circle, this is an indication that the hook has been overloaded.

Pneumatic or air hoists are constructed with cylinders and pistons for reciprocating motion and air motors for rotary motion. Compressed air is the actuating medium in both. Various arrangements admit air to and discharge it from cylinders mounted to operate vertically or horizontally. Pneumatically operated hoists provide smooth action and sensitive response to control; these characteristics account for their wide use in handling fragile materials, such as molds in foundries. In addition, freedom from sparking makes them useful in locations where the presence of explosive mixtures make electrical equipment hazardous.

Derricks and cranes. A derrick is distinguished by a mast in the form of a slanting boom pivoted at its lower end and carrying load-supporting tackle at its outer end. In contrast, jib cranes always have horizontal booms.

Derrick masts are supported by guy lines or stiff legs; some are arranged to rotate 360°. Winches, hand or powered, usually in conjunction with pulleys, do the lifting. Derricks are standard equipment on construction jobs; they are also used on freighters for loading and unloading cargo, and on barges for dredging operations.

Jib cranes, when carried on self-supporting masts, are called pillar cranes; those mounted on walls are called wall-bracket cranes. Cranes with jiblike booms are frequently used in shops, mounted on columns

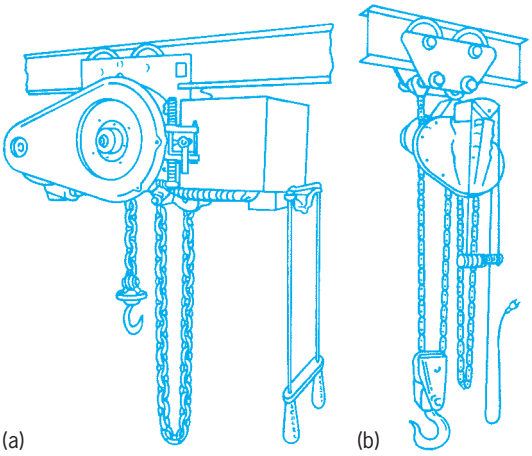

Fig. 4. Hoists. (*a*) Chain. (*b*) Roller chain.

or walls, but have limited coverage. They may have their own running gear or be mounted on trucks (**Fig. 5**). Mobile types for heavier service are called yard cranes or crane trucks. They may or may not be able to rotate their booms. More powerful machines of this type belong to the power crane and shovel group. *See* BULK-HANDLING MACHINES.

Overhead-traveling and gantry cranes. Hoisting machines with a bridgelike structure spanning the area over which they operate are overhead-traveling or gantry cranes. In the overhead-traveling type, the bridge is carried by, and moves along, overhead trackage which is usually fixed to the building structure itself. The gantry crane is normally supported by fixed structures or arranged for running along tracks on ground level. Gantry cranes are standard equipment in shipside operations. Basic arrangements of overhead-traveling cranes are top running and underhung. In the former, the bridge's end trucks ride on top of the runway rails; in the latter, the end trucks carry the bridge suspended below the rails (**Fig. 6**). Types for relatively light duty can be made of elements used in the construction of overhead track. *See* MONORAIL.

Both overhead-traveling and gantry cranes are called bridge cranes. These cranes span a rather large area and differ among themselves primarily only in the construction of the bridge portion of the crane, and in the method of suspension of the bridge. Where smaller areas are to be spanned, standard beams are used for the bridge structure; however, built-up girders or trusslike bridge structure are used for larger spans. A full gantry crane has both its supporting elements erected on the ground, usually riding on tracks. A half-gantry crane has a supporting structure at one end of the bridge that reaches to the ground, and the other end is carried directly on overhead tracks. Selection of either type depends primarily on building design and the areas in which the crane is to be used.

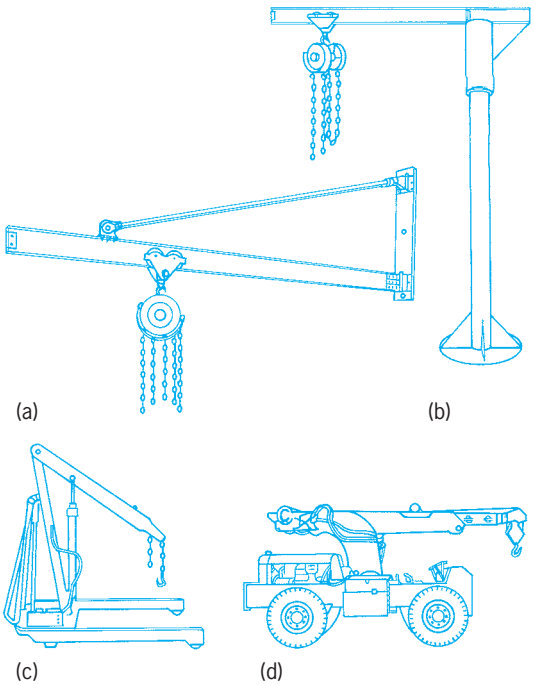

Fig. 5. Four types of jib cranes. (*a*) Wall crane. (*b*) Pillar crane. (*c*) Movable hydraulic crane. (*d*) Crane truck.

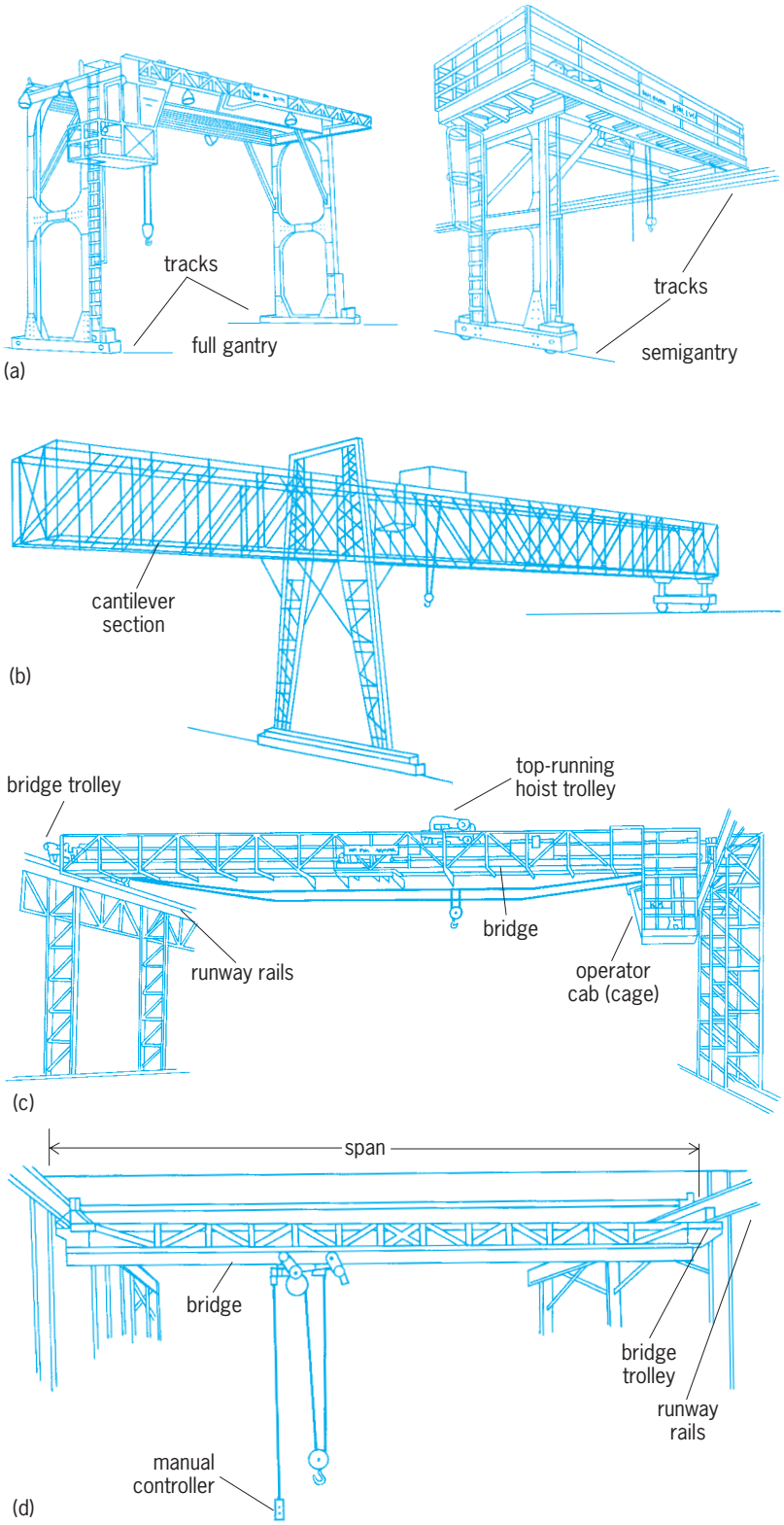

Fig. 6. Basic types of overhead-traveling cranes. (*a*) Traveling gantry cranes. (*b*) Cantilever crane. (*c*) Top-running crane. (*d*) Underhung crane.

Any one or several of the hoists described earlier may be attached to the bridge, usually suspended from a trolley attached to an I-beam track. The combination of a hoist on a track and the bridge crane itself moving on tracks provides for usable movement of equipment within a rectangular area governed only by the length of the bridge and the total horizontal movement of the bridge or gantry crane.

In simpler units the bridges and trolley hoists may be hand-propelled. In heavy-duty units hand operation is not practical and separate electric motors drive each motion. Controls for the motors vary from pendant-type push buttons operated from the ground to remote or automatic control. Pendant controls are satisfactory when a crane has intermittent use during the work day; in larger units, where the crane is in constant use and heavy loads are the rule, an operator may be stationed in a cab mounted to the bridge structure. A more sophisticated extension of this is the operatorless crane, which is worked by means of an electronic control.

Another type of lifting mechanism, used on a modified type bridge crane, is a fork-lift attachment suspended from the overhead truck or bridge; it is referred to as a stacker crane. This unit is especially used for such locations as stock rooms, die racks, or finished-goods storage because it has the advantage of being able to handle unit loads in narrow aisles. Arthur M. Perrin

Bibliography. J. M. Apple, *Plant Layout and Materials Handling*, 3d ed., 1977, reprint 1991; Electric Overhead Crane Institute, *Specifications and General Information for Standard Industrial Service: Electric Overhead Traveling Cranes*; W. A. Rossnagel, *Handbook of Rigging: For Construction and Industrial Operations*, 4th ed., 1988; H. I. Shapiro, *Cranes and Derricks*, 3d ed., 1999.

Holasteroida

An order of irregular echinoids (sea urchins) of the superorder Atelostomata, with a strongly bilaterally symmetrical test. They have a small oval mouth lacking buccal notches that lies close to the anterior on the lower surface: a lantern is never present. Ambulacral plating is simple. The anterior ambulacrum is often differentiated, and petals are developed adapically (see **illus.**). The anus lies on the posterior face. The elongate apical disc, in which ocular plates II and IV abut along the midline and separate the anterior and posterior pairs of genital plates, is the most distinctive characteristic. In a few genera, the apical disc is split, anterior and posterior parts being separated by intervening interambulacral plates. Phyllodes of penicillate feeding tube feet with their associated characteristically large pores are typically developed around the mouth.

The 105 Recent species are divided into 13 genera. All are deep-sea forms, living at depths between 325 and 24,600 ft (100 and 7500 m), but are most commonly found between 4900 and 14,800 ft (1500 and 4000 m). Their fossil record indicates that they were much more common and diverse in the past, with 61 genera divided into seven families, most of these being shallow-water forms. The group first appeared in the Early Cretaceous, diversified greatly during the middle and Late Cretaceous, and then declined during the Tertiary to their present low levels. They are all deposit feeders, living either epifaunally or shallowly buried in unconsolidated substrata. *See* ECHINODERMATA. Andrew Smith

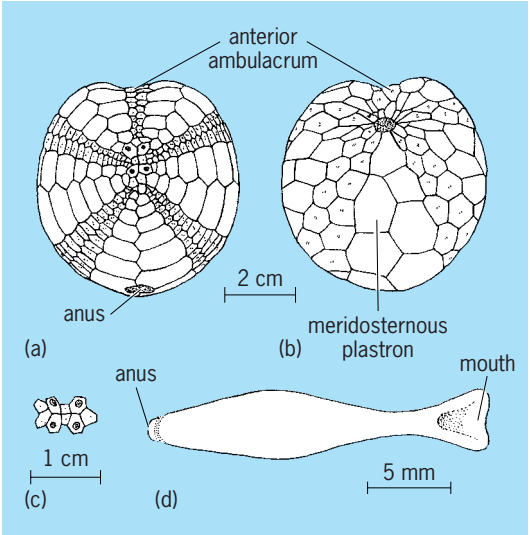

Diagnostic features of holasteriods. (*a*) Aboral aspect. (*b*) Adoral aspect. (*c*) Apical system. (*d*) *Echinosigra paradoxa*, a deep-sea species of Pourtalesiidae, adoral aspect.

Hole states in solids

Vacant electron energy states near the top of an energy band in a solid are called holes. A full band cannot carry electric current; a band nearly full with only a few unoccupied states near its maximum energy can carry current, but the current behaves as though the charge carriers are positively charged. The situation can be understood in terms of the definition of the effective mass: if the energy band is specified by a function $E(k)$, where k is the magnitude of the wave vector $\mathbf{k}$, the effective mass for a spherical band is given by the equation below, where $\hbar$ is Planck's constant

$$m* = \hbar^2 \left(\frac{\partial^2 E}{\partial k^2} \right)^{-1}$$

divided by 2π. Near a maximum of the band, the second derivative of the energy is negative, so the effective mass is negative. States for which the effective mass is negative are defined as hole states. Carriers in such states behave under the influence of an external electromagnetic field as though they carry positive charge. *See* BAND THEORY OF SOLIDS.

The process of conduction in such a system may be visualized in the following way. An electron moves against an applied electric field by jumping into a vacant state. This transfers the position of the vacant

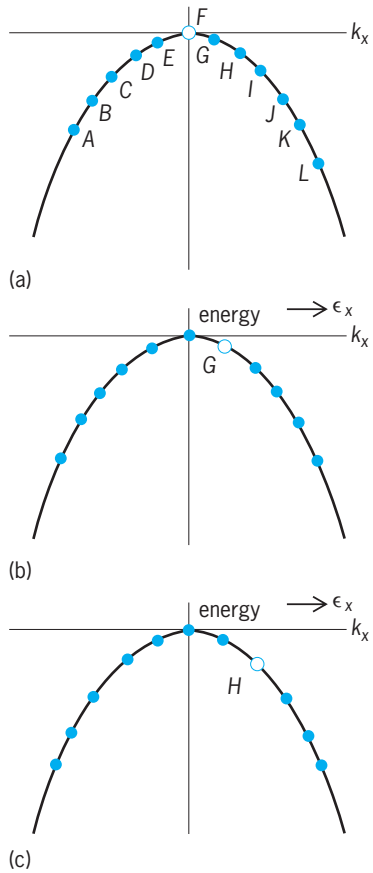

Process of hole conduction. (*a*) At time $t = 0$, energy states *A* through *L* are filled except *F*. (*b*) An electric field ϵ_x is applied in the $+x$ direction. The force on the electrons is in the $-k_x$ direction, and all electrons make transitions in the $-k_x$ direction, moving the hole to *G*. (*c*) After a further interval, the electrons move farther along, and the hole is now at *H*. (*After C. Kittel, Introduction to Solid State Physics, 3d ed., Wiley, 1956*)

state, or propagates the hole, in the direction of the field, as shown in the **illustration**. Whether conduction occurs by electrons or holes is determined experimentally from the sign of the Hall emf. If a current is carried in the presence of a magnetic field perpendicular to the current, an emf is developed perpendicular to the current and to the field. The sign of this emf depends on the charge on the carriers. *See* HALL EFFECT.

Hole conduction is important in many semiconductors, notably germanium and silicon. The occurrence of hole conduction in semiconductors can be favored by alloying with a material of lower valence than the "host." Semiconductors in which the conduction is primarily due to holes are called *p* type. Hole conduction is also observed in some metals, including iron and chromium. In other metals, including aluminum and bismuth, both holes and electrons may be present in equilibrium. *See* ELECTRICAL CONDUCTIVITY OF METALS; SEMICONDUCTOR.

Joseph Callaway

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; A. Sutton, *Electronic Structure of Materials*, 1993.

Holectypoida

An extinct order of primitive irregular sea urchins of the class (Echinoidea), which retain a functioning lantern tern throughout life. Although their periproct opens on the oral surface in the posterior interambulacrum, the test still shows considerable radial symmetry. The mouth is large, circular, and centrally positioned and is indented by sharp buccal notches. The lantern has stout teeth that are wedge-shaped in cross section. Ambulacral plating is trigeminate adorally but usually simple adapically; the tube feet and their pore pairs are undifferentiated. Tubercles are arranged radially on the oral surface, and tuberculation is distinctly finer adapically. The holectypoids are distinguished from other primitive irregular echinoid groups by the presence of a fifth genital plate in the apical disc that may be perforated by a gonopore.

The approximately 130 nominal species are divided into nine genera and two families, Holectypidae and Discoididae. The families are differentiated on the presence or absence of internal buttressing. Holectypoids first appeared in the late Lower Jurassic and survived to the end of the Cretaceous (Maastrichtian), when they became extinct. They were mostly infaunal deposit feeders living in relatively coarse, permeable substrata. *See* ECHINODERMATA; ECHINOIDEA.

Andrew Smith

Holly

The American species of holly (*Ilex opaca*) attains a maximum height of 40–50 ft (12–15 m) and has evergreen leaves. It grows naturally in the eastern and southeastern United States close to the Atlantic and Gulf coasts, in the Mississippi Valley, and westward to Oklahoma and Missouri. It is best known for its bright red berries, which make a pleasing contrast with the deep-green spiny leaves, and for this reason it is valued for decorations at the Christmas season (see **illus.**). The wood is hard, tough, and close-grained. The heartwood is ivory white when first cut, but becomes brownish with age or on exposure, and takes a high polish. It is used for cabinet work and musical instruments; because it resembles ivory, it is sometimes used for keys for pianos and organs. Its fine grain makes it valuable for wood-engraving work.

The English holly (*I. aquifolium*) is cultivated extensively in the extreme northwestern United States, but is not hardy in the northeastern states. Its spiny

The American holly (*Ilex opaca*).

leaves are glossier than those of the American holly and have wavier margins. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Holmium

A chemical element, Ho, atomic number 67, atomic weight 164.93, a metallic element belonging to the rare-earth group. The stable isotope ^{165}Ho makes up 100% of the naturally occurring element. The

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

metal is paramagnetic, but as the temperature is lowered, it changes to antiferromagnetic and then to the ferromagnetic system. *See* ANTIFERROMAGNETISM; PERIODIC TABLE; RARE-EARTH ELEMENTS.

Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner, Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

Holocene

That portion of geologic time that postdates the latest episode of continental glaciation. The Holocene Epoch is synonymous with the Recent or Postglacial interval of Earth's geologic history and extends from 10,000 years ago to the present day. It was preceded by the Pleistocene Epoch and is part of the Quaternary Period, a time characterized by dramatic climatic oscillations from warm (interglacial) to cold (glacial) conditions that began about 1.6 million years ago. The term Holocene is also applied to the sediments, processes, events, and environments of the epoch.

In 1833 British geologist Charles Lyell applied the term Recent to the interval of time of human existence. Now it is known that humans have been around much longer than the last 10,000 years, but that fact does not discount the need for a term for the part of Earth's history closest to us. In 1846 Forbes proposed Postglacial as an alternative term for Recent. However, Postglacial is inappropriate in higher latitudes where the postglacial interval did not begin until several thousand years after it did in the middle latitudes.

In 1867 the term Holocene (meaning wholly recent) was proposed for the interval of time following the Pleistocene, and in 1885 it was formally submitted to the International Geological Congress. The U.S. Geological Survey officially adopted the term Holocene in 1967, and the U.S. Commission on Stratigraphic Nomenclature endorsed it in 1969.

The Holocene represents the current interglacial episode. The Holocene interval is treated differently in geologic time charts than are earlier interglacial episodes, however. It is designated as a separate epoch of the Quaternary Period. Although there is little reason to suspect that the present interglacial interval differs substantially from earlier ones and there is little doubt that another glacial interval will follow, a separate designation is appropriate for the Holocene. It holds a unique place in geologic history—it is the time of rapid development of human culture and the rise of humankind as a geologic agent whose activities have substantially modified Earth's environment.

As the interval of time closest to us, the Holocene Epoch is very convenient to study. Holocene sediments cover virtually every part of the Earth's surface and represent almost every environment of deposition. With the development of ^{14}C dating (a method of age determination based on the measurement of radioactive carbon decay), Holocene sediments are relatively easy to date. From a scientific standpoint, the Holocene Epoch is of great interest because it provides a recent analog for past environments and processes. Its sediments and landforms provide important clues to changes that occurred as a result of the last shift from the glacial to the nonglacial climatic mode. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; RADIOCARBON DATING.

The Pleistocene/Holocene transition was a time of dramatic environmental change. The huge ice sheets that had developed over the northern and western parts of North America (Laurentide and Cordilleran, respectively) and most of Scandinavia were at their maximum geographic extent about 18,000 ^{14}C years B.P. (before present, where present is defined as the

CENOZOIC	QUATERNARY	Holocene
		Pleistocene
	TERTIARY	Pilocene
		Milocene
		Oligocene
		Eocene
		Paleocene
MESOZOIC	CRETACEOUS	

year 1950) and in full retreat by 14,000 ^{14}C years B.P. By 10,000 ^{14}C years B.P., the Laurentide ice sheet had withdrawn from the Great Lakes. The ice sheets survived in the northern latitudes for another 3000 ^{14}C years or so. The progress of deglaciation was complex, because the overall glacial meltback was interrupted by intervals of glacier readvance. It remains unclear whether these readvances were synchronous on a hemispheric or global scale and what role ice sheet/oceanic interactions played in the deglaciation. *See* PLEISTOCENE.

Boundaries. Because the deglaciation began earlier in the middle latitudes than it did in the high latitudes, the boundary between sediments of the Pleistocene and Holocene epochs is time-transgressive (varies in age from place to place). A precise location for the boundary between sediments corresponding to the two epochs has never been agreed upon, although an arbitrary age of 10,000 years ^{14}C years B.P. (defined by the Holocene Commission for the International Quaternary Association) is commonly used as the age for the boundary.

The time-transgressive nature of the boundary between Pleistocene and Holocene sediments is verified by ^{14}C dating, and some geologists have argued that it is more realistic and practical to represent boundaries between late Quaternary events and sediments as diachronous (beginning and ending at different times in different places) than to use geologic time classifications that assume time-synchronous (same age everywhere) boundaries for events and sediments. Since 1983, when the North American Stratigraphic Nomenclature Committee introduced diachronic units, informal and formal diachronic classifications for late Quaternary time and event units have been proposed as alternatives for geologic time units with time-synchronous boundaries. An example of a diachronic unit is the Hudson Episode. It is based on postglacial sediments that occur stratigraphically above deposits of the Wisconsin (last glacial) Episode. The Hudson Episode, which began several thousand years earlier in Illinois than it did at Hudson Bay, is an alternative designation to the Holocene Epoch, a geologic time unit that began everywhere 10,000 years ago.

Study and subdivision. Most studies of the Holocene apply an interdisciplinary approach. Such an approach is necessary to address complex interactions among the various natural systems, which respond with different sensitivities and at different rates to change. Traditionally, climate changes in the Holocene have been studied chiefly through the technique of palynology (pollen analysis), but increasingly paleontology (for example, fossil plant remains, beetles, ostracods, diatoms), dendochronology (tree ring analysis), marine and ice core records, and fluctuations of alpine glaciers have been used to reconstruct past climates. *See* PALEOCLIMATOLOGY.

Although commonly viewed as a relatively stable climatic interval compared to the Pleistocene Epoch, the Holocene Epoch has experienced small, yet significant climatic changes. It is subdivided informally into an early phase (from about 10,000 to 8000 ^{14}C years B.P.) influenced by the waning effects of receding ice sheets, a middle phase (from about 8000 to 4000 ^{14}C years B.P.) characterized by the warmest interval of the present interglacial episode, and a late phase (the last 4000 ^{14}C years) that included climatic reversal at the onset of Neoglaciation (a time of renewed glacial activity that some assume to be the first phase of the next glacial episode).

Early Holocene. The early phase of the Holocene was geologically the most eventful. The periglacial (near the edge of the ice) landscape was unstable and very dynamic. As the Pleistocene ice sheets melted, enormous volumes of water, stored as glacier ice for many thousands of years, returned to the oceans via meltwater streams or by way of ice streams that flowed directly to the ocean.

As the ice sheets shrank, sea level rose an average of 130 m, drowning the continental margins and closing many land bridges, including the land bridge across the Bering Strait between Asia and North America that had enabled humans to migrate to the Americas. In parts of Canada and Scandinavia, temporary marine invasions occurred when the ice melted from low areas where the Earth's crust had been depressed by the weight of the ice sheets.

In natural catchment basins along the retreating Cordilleran and Laurentide ice margins, huge lakes formed. Some were on the order of several hundred thousand square kilometers in areal extent. Many of these lakes were ephemeral and flashy (there one year and gone the next over several thousands of years). When ice or sediment dams gave way, catastrophic floods scoured the landscape and the lakes drained quickly, only to refill and repeat the process. Examples include glacial Lakes Missoula and Columbia, whose torrents repeatedly swept across Washington state to the Pacific Ocean via the Columbia River valley, and glacial Lakes McConnell and Agassiz, whose floodwaters alternately spilled to the Arctic Ocean, the Gulf of Mexico, and the Atlantic Ocean. Some lakes shrank more gradually or dried up completely when they were no longer fed by glacial meltwaters. Other lakes, such as the Great Lakes, found new, lower outlets, which sometimes altered the drainage course of vast watersheds.

As the ice sheets waned, the Earth's crust rose, rebounding from the release of the weight of thousands of meters of glacier ice and creating uplifted shoreline features and sediments. Parts of Hudson Bay and Scandinavia were uplifted several hundred meters. Maximum uplift occurred in the early Holocene, but uplift continues even today although at much slower rates.

With the retreat of the Northern Hemisphere ice sheets came a wave of migration that eventually resulted in the present distribution of plant communities. Pollen studies of cores from lakes and bogs in glaciated areas of western Europe and North America indicate that tundra areas bordering the ice initially gave way to spruce forests. Later, the spruce were

replaced by pine and birch, followed still later by broad-leaf deciduous trees such as oak and elm.

Animal migrations followed the changes in the distribution of plant communities. A dramatic phenomenon in the late Quaternary was the extinction of a large number of animals, especially large mammals, at the close of the last glacial episode. Most of the extinctions in North America took place before 10,000 ^{14}C years B.P., and almost all had occurred by 8000 ^{14}C years B.P. The cause of these extinctions is not known, but they coincide with the Pleistocene/Holocene transition, the time of maximum climate and vegetational change, the opening of an ice-free corridor between the Cordilleran and Laurentide ice sheets in Canada, and massive migrations of human hunters south to the margins of the ice sheets.

As the plant and animal communities adjusted to the changing climate of the early Holocene, other environmental adjustments also occurred. These adjustments were highly dependent on the geographic area and the direct or indirect impact of glaciation on the landscape. In the recently deglaciated areas, soils had begun to form, the effects of periglacial conditions on the landscapes had diminished, and slopes were becoming more stable by the close of the early Holocene interval. In the tropics and the equatorial belt, precipitation increased and lakes were filled to high levels, indicating a climate that was much wetter than that of today.

Middle Holocene. The middle phase of the Holocene has been called the hypsithermal, a name for the warmest interval of the present interglacial episode (see **illus.**). It has also been referred to as the climatic optimum, a term which is more appropriately applied to the peak warmth of the hypsithermal phase. At the climatic optimum, world temperature was probably 2 or 3°C higher than today. Like other late Quaternary phases, the hypsithermal reached the lower latitudes earlier than it did the middle and higher latitudes. In the tropics the climatic optimum may have corresponded more closely to the interval of maximum insolation (exposure to the Sun), about 10,000 ^{14}C years B.P., but in the middle latitudes of the midwestern United States it occurred about 7000 ^{14}C years B.P., whereas in Labrador it occurred about 4000 ^{14}C years B.P. Fossil finds reveal that the geographic ranges of numerous species of plants and animals were several hundred kilometers farther north during the hypsithermal than they are today. For example, broad-leaved trees reached farther north than they now do in northern Europe and North America. The descendants of fossil beetles found in central Ontario do not extend north of southern Ohio today. In the higher latitudes, permafrost (ground that remains frozen even during the summer) likely extended to a greater depth below the surface and was much less widespread than today. The climate was warm enough to melt much of the sea ice in the Arctic Ocean, as indicated by the occurrence of fossil driftwood (dated at 4000–6000 ^{14}C years B.P.) on uplifted beaches.

In mountainous areas of the middle latitudes, the treeline and other vegetational boundaries shifted during the middle Holocene to several hundred meters higher than their present elevations. In the mountains of the western United States, many alpine glaciers disappeared. In the midwestern United States, the middle Holocene was a time not only of maximum warmth but of relatively dry conditions. Due to summer drought, the prairie/forest border extended more than 100 km east of its present position. Today relict prairie soils can be found in areas that became forested during the late Holocene. The Great Plains became more arid, which is believed to have caused a major reduction of human and animal populations in that area during the hypsithermal.

Late Holocene. After the climatic optimum, the Earth experienced climatic cooling. The shift to a cooler, moister climate began about 5000–4000 ^{14}C years B.P. in the midcontinent. In western North America at about 5000 ^{14}C years B.P., the mountain glaciers began to expand again. This renewed glacial activity is called Neoglaciation (see illus.). At least three intervals of glacial expansion have occurred in the late Holocene. The glacial advances are cyclic. In the mountains of the western United States, the three advances have been dated at about 5000, 2800, and 300 ^{14}C years B.P. The most impressive of the three glacial intervals is the last, called the Little Ice Age. It is well documented because it occurred in historic time. Between the intervals of glacier expansion were times of climatic warming. One, called the Little Climatic Optimum to differentiate it from the hypsithermal of the middle Holocene, peaked about 1800 ^{14}C years B.P.

Many Rocky Mountain glaciers advanced as much as 1–2 km during the Little Ice Age. Trees that were overrun by advancing glaciers now are being exposed as the ice recedes due to a climatic warming that began over 100 years ago. Little Ice Age moraines are well defined in the Rocky Mountains and in Europe. The moraines commonly occur only

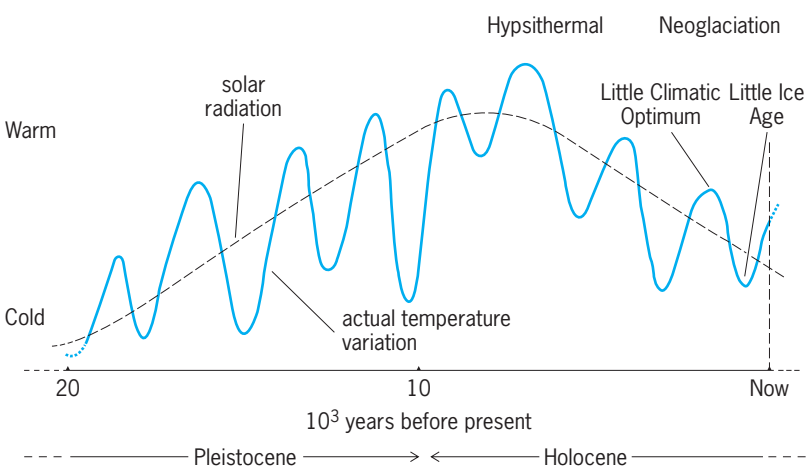

Climatic variations during the last 20,000 years. The broken line is solar radiation reaching the Earth at 65°N latitude. The solid line is a speculative reconstruction of actual temperature variation. (*From E. C. Pielou, After the Ice Age: The Return of Life to Glaciated North America, University of Chicago Press, 1991*)

a few hundred meters downslope from the present glacier margin. Their surfaces are bare of vegetation, unlike older Holocene moraines. *See* MORAINE.

The late Holocene phase saw the spread of peatland (also called muskeg) across much of Canada and the northern Great Plains, particularly in areas of former lake plains. Peatland development is attributed to the return to a wetter, cooler climatic regime. Treelines that had migrated northward and to higher elevations during the hypsithermal interval returned southward and to lower elevations.

During the late Holocene, human populations expanded and human culture developed into the complex agricultural, industrial, and technological society of today. The result is that humans have become significant factors in altering the Earth's surface environment, including, most believe, Holocene climate. A number of scientists attribute recent global warming (nearly 0.5°C in the past 150 years) to the buildup of carbon dioxide and other greenhouse gases in the atmosphere as a result of human activity. Some scientists warn of a much greater increase to come in the next 50 years. Only geologic time will tell the impact of humankind on the Earth's environment, including the natural climatic cycles. *See* GEOLOGIC TIME SCALE; GREENHOUSE EFFECT; QUATERNARY.

Ardith K. Hansel

Bibliography. C. L. Matsch, *North America and the Great Ice Age*, McGraw-Hill, New York, 1976; E. C. Pielou, *After the Ice Age: The Return of Life to Glaciated North America*, University of Chicago Press, 1991; H. E. Wright (ed.), *Late Quaternary Environments of the United States*, vol. 2: *The Holocene*, University of Minnesota Press, Minneapolis, 1983.

Holography

A technique for recording, and later reconstructing, the amplitude and phase distributions of a coherent wave disturbance. Invented by Dennis Gabor in 1948, the process was originally envisioned as a possible method for improving the resolution of electron microscopes. While this original application has not proved feasible, the technique is widely used as a method for optical image formation, and in addition has been successfully used with acoustical and radio waves. This article discusses holography with electromagnetic waves in the optical and microwave regions of the electromagnetic spectrum, and its potential use with x-rays. For holography with sound waves. *See* ACOUSTICAL HOLOGRAPHY.

Optical Holography

Optical holography makes use of a highly coherent beam of light, such as supplied by a laser source. *See* LASER.

Fundamentals of the technique. The technique is accomplished by recording the pattern of interference between the unknown object wave of interest and a known reference wave (**Fig. 1**). In general, the object wave is generated by illuminating the (possibly three-dimensional) subject of concern with the coherent light beam. The waves reflected from the object strike a light-sensitive recording medium, such as photographic film or plate. Simultaneously a portion of the light is allowed to bypass the object, and is sent directly to the recording plane, typically by means of a mirror placed next to the object. Thus incident on the recording medium is the sum of the light from the object and a mutually coherent reference wave.

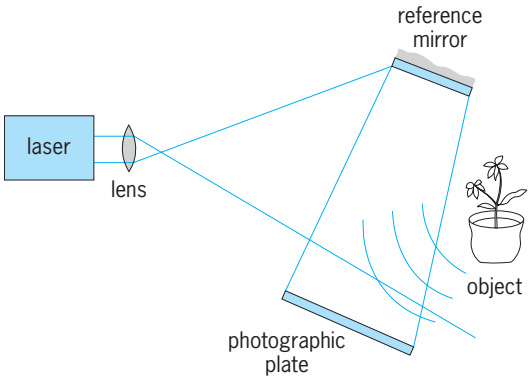

Fig. 1. Recording a hologram.

While all light-sensitive recording media respond only to light intensity (that is, power), nonetheless in the pattern of interference between reference and object waves there is preserved a complete record of both the amplitude and the phase distributions of the object wave. Amplitude information is preserved as a modulation of the depth of the interference fringes, while phase information is preserved as variations of the position of the fringes. *See* INTERFERENCE OF WAVES.

The photographic recording obtained is known as a hologram (meaning a total recording); this record generally bears no resemblance to the original object, but rather is a collection of many fine fringes which appear in rather irregular patterns (**Fig. 2**).

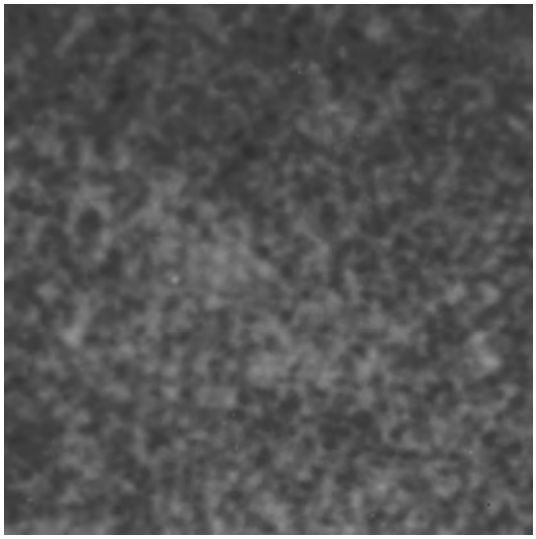

Fig. 2. Typical appearance of a hologram (under magnification).

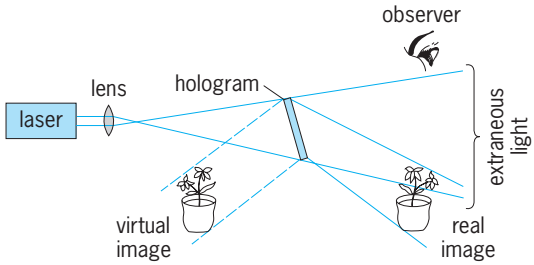

Fig. 3. Obtaining images from a hologram.

Nonetheless, when this photographic transparency is illuminated by coherent light, one of the transmitted wave components is an exact duplication of the original object wave (**Fig. 3**). This wave component therefore appears to originate from the object (although the object has long since been removed) and accordingly generates a virtual image of it, which appears to an observer to exist in three-dimensional space behind the transparency. The image is truly three-dimensional in the sense that the observer's eyes must refocus to examine foreground and background, and indeed can "look behind" objects in the foreground simply by moving the head laterally.

Also generated are several other wave components, some of which are extraneous, but one of which focuses of its own accord to form a real image in space between the observer and the transparency. This image is generally of less utility than the virtual image because its parallax relations are opposite to those of the original object.

Applications. The holographic technique has a number of unique properties which make it of great value as a scientific tool.

Microscopy. Historically, microscopy is the potential application of holography that has motivated much of the early work, including the original work of Gabor. The use of holography for optical microscopy has been amply demonstrated, but these techniques are not serious competitors with more conventional microscopes in ordinary microscopy.

Nonetheless, there is one area in which holography offers a unique potential for optical microscopy. This area might be called high-resolution volume imagery. In conventional microscopy, high transverse resolution is achieved only at the price of a very limited depth of focus; that is, only a limited portion of the object volume can be brought into focus at one time. It is possible, of course, to explore a large volume in sequence by continuously refocusing to examine new regions of the object volume, but such an approach is often unsatisfactory, particularly if the object is a dynamic one, continuously in motion. A solution to this problem is to record a hologram of the object by using a pulsed laser. The dynamic object is then "frozen" in time, but the recording contains all information necessary to explore the full object volume with an auxiliary optical system. Sequential observation is acceptable because the object (that is, the holographic image) is no longer dynamic. This approach has been fruitfully applied to the microscopy of three-dimensional volumes of living biological specimens and to the measurement of particle-size distributions in aerosols.

Interferometry. Holography has been demonstrated to offer the capability of several unique kinds of interferometry. This capability is a consequence of the fact that holographic images are coherent; that is, they have well-defined amplitude and phase distributions. Any use of holography to achieve the superposition of two coherent images will result in a potential method of interferometry.

The most powerful holographic interferometry techniques are based on the following property: When a photographic emulsion is multiply exposed to form several superimposed holograms, upon reconstruction the several corresponding virtual images are formed simultaneously and therefore interfere. Likewise the various real images interfere.

The most dramatic demonstrations of this type of interferometry use a pulsed ruby laser. Two laser pulses are used to record two separate holograms on the same transparency. Any changes of the object between pulses result in well-defined fringes of the interference in the reconstructed image (**Fig. 4**). The technique is particularly well suited for performing interferometry through imperfect optical elements (for example, windows of poor quality), thus making possible certain kinds of interferometry that could not be achieved by any classical means. *See* INTERFEROMETRY.

Memories. Optical memories for storing large volumes of binary data in the form of holograms have been intensively studied. Such a memory consists of an array of small holograms, each capable of reconstructing a different "page" of binary data. When one of these holograms is illuminated by coherent light, it generates a real image consisting of an array of bright or dark spots, each spot representing a binary digit. This image falls on a detector array, with one detector element for each binary digit. Thus to read a single binary digit at a specific location in the memory, a beam deflector causes light to illuminate the appropriate hologram page, and the output of the proper detector element is interrogated to

Fig. 4. Image taken by the technique of holographic interferometry, showing the compressional waves generated by a high-speed rifle bullet. (*Courtesy of R. E. Brooks, L. O. Heflinger, and R. F. Wuerker*)

determine whether a bright spot of light exists at that particular location in the image. *See* COMPUTER STORAGE TECHNOLOGY.

Display. There has been interest in the use of holography for purposes of display of three-dimensional images. Applications have been found in the field of advertising, and there is increased use of holography as a medium for artistic expression. A significant technical development in this area has been the perfection of a type of recording known as a multiplex hologram. Such a recording typically consists of a large number of separate holograms, all in the form of thin, contiguous, vertical strips on a single piece of film. Each of these holograms produces a virtual image of a different ordinary photograph of the subject of interest. In turn, each such photograph was originally taken from a slightly different angle. Thus when the observer examines the virtual image produced by the entire set of holograms, each eye looks through a different hologram and sees the subject from a different angle. The resulting stereo effect produces a nearly perfect illusion of three-dimensionality. Furthermore, as an observer moves the head horizontally, or as the collection of holograms is rotated, the observer's two eyes continuously see a changing pair of images. If the original set of photographs is properly chosen, the image can be made to move or dance about in nearly any desired fashion. Very dramatic three-dimensional displays of animated subjects can thus be constructed from a series of ordinary photographs. Such displays do not require a laser for viewing, but rather can be utilized with white-light sources.

Holographic optical elements. A hologram consisting of the interference of a plane reference wave and a diverging spherical wave, upon illumination by a reconstruction plane wave, will generate a diverging spherical wave (the virtual image) and a converging spherical wave (the real image), each traveling in a different angular direction. Thus such a hologram behaves as an optical focusing element, with properties similar to those of a lens (or, more accurately, a pair of lenses). More complex holograms can generate a multitude of foci, in virtually any pattern desired. Alternatively, by varying the periodicity of the gratinglike structure of the hologram, a small laser beam can be deflected through an angle that is controlled by the local period of the structure. Holograms which are used to control transmitted light beams, rather than to display images, are called holographic optical elements. Interest in such elements has grown substantially, and commercial applications have been found. Most notable is the use of holograms in supermarket scanners at checkout stands. Light from a helium-neon laser falls on a small region of a holographic optical element, which was recorded on a disk and is rotating continuously. As the hologram rotates, different portions of the hologram containing different grating periods are illuminated, and the angle of deflection of the laser beam sweeps through a pattern that was predetermined when the hologram was recorded. In this way the laser beam is caused to follow a complicated scan pattern, which ultimately allows the reading of information from the bar-code patterns recorded on each product. *See* CHARACTER RECOGNITION; GEOMETRICAL OPTICS; OPTICAL IMAGE.

Security applications. Holograms are widely used in various applications where authenticity is of the utmost importance, for example, in credit cards and in bank notes to prevent forgery. A holographic postage stamp is in use in the United States to prevent counterfeiting. Security applications have had the greatest commercial impact of any applications of holography.

Other applications. A variety of other applications of holography has been proposed and demonstrated, including the analysis of modes of vibration of complicated objects, measurement of strain of objects under stress, generation of very precise depth contours on three-dimensional objects, and high-resolution imagery through aberrating media. These and other applications of holography will be useful in future scientific and engineering problems. Joseph W. Goodman

Thick holograms. Thick, white-light reflection holograms were introduced by Y. Denisyuk in 1962. By illuminating a photosensitive plate so that the object and reference waves come from opposite sides of the plate, fringes can be formed that run parallel to the surfaces of the plate. These fringes act in concert as a multilayer mirror whose layers are warped. The warping is such that when the hologram is illuminated a reconstruction of the original object results. In addition, in polychromatic illumination the layers act as a thin-film stack to select the original wavelength. Multiple fringe patterns can be recorded to reproduce a color reconstruction of the object.

Rainbow hologram. A rainbow hologram, invented by S. A. Benton in 1969, is composed of vertically aligned narrow strips. Each strip reconstructs a single perspective of a three-dimensional (3-D) object. A strip is selected for viewing by locating the pupil of the eye on a line connecting the strip to a point in the light source. Because the light path to the left and right eyes will pass through different strips, the viewer has a stereoscopic impression of the object. Moving the head then gives a full 3-D impression. Object motion can also be incorporated by capturing the object at different times in different strips. For example, a 3-D object may appear to move as the viewer walks around a cylindrical hologram.

Embossed holograms. The embossed hologram is stamped into the medium. For this application, a mask is generated in a material such as photoresist. The mask is coated with a hard material to become a master for generating stamps. The stamps are then used on softer but durable materials to create the holographic emblems that are widely seen today. The technique of rainbow holography can also be applied to create pseudo-3-D images from white-light illumination.

Digital holography. In digital holography the recording is done by the combination of a digital sensor such as a charged-coupled device (CCD) and computer memory. The fringe pattern that is formed between

the optical object and reference waves is stored digitally. The computer rather than wave propagation generates the reconstruction. *See* CHARGE-COUPLED DEVICES; DIGITAL COMPUTER.

Computer-Generated Holography

Information in a hologram is contained in patterns called fringes (**Fig. 5***a*). For optical holography, the fringe patterns are formed by the interference of two beams of light. In contrast, the fringes in a computer-generated hologram (CGH) are computer-calculated and then drawn, plotted, or printed on a transparent or reflective material. The computer-generated hologram was invented by A. W. Lohmann in 1966. It drew on the recent introduction of the fast Fourier transform, the recent performance gains of computers of that era, and the availability of relatively precise plotters. Insights into the relation of communications theory to holography developed by Lohmann around 1956 also played a crucial role in the development of computer-generated holography. Precursors to the computer-generated hologram can be found in the artificial holograms of W. L. Bragg (1939) and of G. L. Rogers (1952). A variation of the computer-generated hologram, the Kinoform, was described by L. B. Lesem in 1969 and has found important applications.

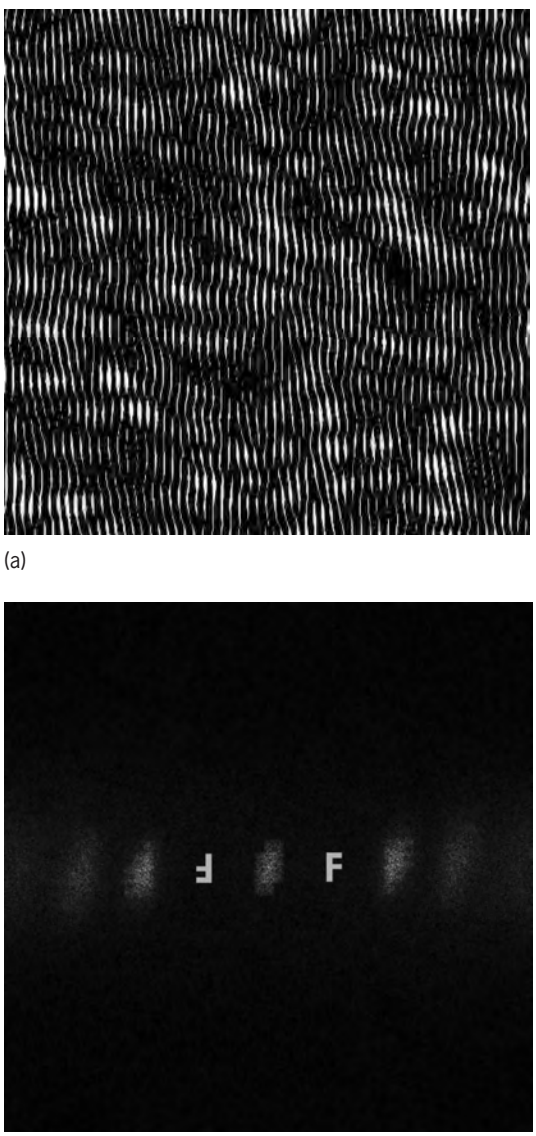

Fig. 5. Computer-generated hologram (CGH). (*a*) Mask. (*b*) Reconstruction from the hologram.

The computer-generated hologram, like the optical hologram, is an entire recording. Information is stored at each point on the computer-generated hologram's surface about both the energy of the light and the direction in which the light is moving. Figure 5*a* is a computer-generated hologram mask; the mask is photographically reduced onto transparent material such as a slide. If Fig. 5*a* is reduced to a size of about 5 mm × 5 mm onto slide film, the resulting slide will be a computer-generated hologram. If a point of red light, for example, a laser-point spot on a distant wall, is observed though the computer-generated hologram, the observer sees not a point but the letter F as illustrated in Fig. 5*b*. The letter appears to hover at the location of the point source of light. The computer-generated hologram has converted the simple optical wave emanating from the point into the more complicated wave that would have come from the F. The object in this illustration is flat, but a three-dimensional object could have been chosen just as easily. Because the hologram is artificially generated, there is almost unlimited flexibility with regard to what information is put into it.

In addition to the desired image, the F, there are other structures visible in Fig. 5*b*. The technical term for the desired image is the true image. An inverted image called the twin image is also seen, which comes from the fact that this type of computer-generated hologram is purely absorbing and does not change the phase of the light. There is a single point in the center of the reconstruction called the D.C. spike (D.C. is from the electrical term direct current), which is again an effect of the absorptive nature of this computer-generated hologram. The fuzzy clouds are known as high-diffraction-order noise and come from the fact that this type of computer-generated hologram is binary (that is, it contains only black and white) rather than graytone (containing shades of gray).

Constructing a computer-generated hologram. There are many ways of designing a computer-generated hologram. The three original Lohmann computer-generated hologram types seen in **Fig. 6** provide examples. A virtual grid is imposed on the surface that will become the computer-generated hologram. A rectangular aperture is inserted into each cell of the grid. For the type III computer-generated hologram, that aperture is centered vertically and its width is half the width of the cell. The height of the aperture is proportional to the amplitude A. The horizontal displacement is proportional to the phase Φ.

Applications. Computer-generated holograms can form images. The images can be located at any distance from the computer-generated hologram. The image location can be either in front of the computer-generated hologram or behind it. The images can be

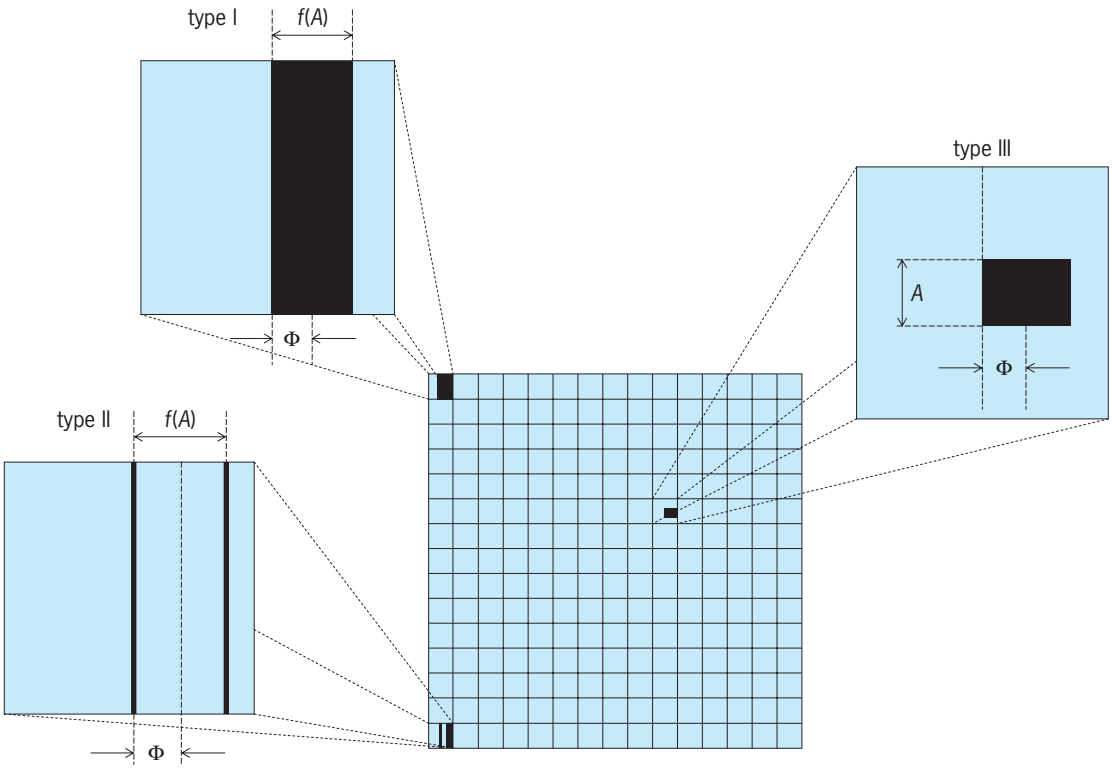

Fig. 6. The three original types of binary computer-generated holograms introduced by A. W. Lohmann.

three-dimensional and can be in color. Computer-generated hologram systems have even been built that produce 3-D television. They are presently limited to showing small images by the large amount of information needed to produce the images. The advantage of the computer-generated hologram over optical holograms is that the hologram may be of an object that has never existed. All that is needed is a mathematical description of the object.

Optical testing. In optical testing, an element under test, the test-piece, is compared to an element of known quality, the reference-piece. One way to perform such a test is to form overlapping images of the two pieces in laser light. The interference patterns formed show the difference between the pieces to a very high accuracy. The use of computer-generated holograms in optical testing was pioneered by A. J. MacGovern and J. C. Wyant in 1971. The basic idea in using computer-generated holograms for optical testing is to replace, or augment, the conventional reference-piece with a computer-generated hologram. Computer-generated holograms can be used in either transmission or in reflection. The primary advantage of a computer-generated hologram is its flexibility. Much more complicated reference pieces can be realized as computer-generated holograms than are practical to fabricate by traditional shaping and polishing of optical glass. *See* INTERFEROMETRY.

Diffractive optical elements. Just as holograms can be fabricated that serve as holographic optical elements, a computer-generated hologram can be created that acts as a diffractive optical element (DOE). The diffractive optical element mimics the optical properties of a glass element, such as a prism or lens. The advantages of such a diffractive optical element are that it is thin and weighs very little. The Kinoform is a type of computer-generated hologram with a structure that is similar to the glass door on a shower, though the dimples in the glass are very thin, about a micrometer, and perfectly controlled rather than random. Kinoforms have been widely applied in imaging systems as lenses. They are gaining acceptance in illumination systems, especially those that are based on light-emitting diodes. A stepped version of the Kinoform introduced by H. Dammann in 1970 has been widely used because of its relatively low cost. *See* LIGHT-EMITTING DIODE.

Diffractive optical elements have inherent dispersion; they act differently on different colors of light. Dispersion can be a direct advantage in devices such spectrometers that rely on it to function. It can be a disadvantage in lenses used for imaging. *See* DISPERSION (RADIATION).

Diffractive optical elements have combined with conventional glass lenses in camera objectives. The dispersion of the two types of elements cancel out, allowing commercial production of camera objectives that are significantly lighter and smaller than their all-glass counterparts. *See* CAMERA.

Optical vortices. Computer-generated holograms can be used to produce waves that rotate about the direction of propagation. These so-called helical waves can be used as propeller beams, to apply torque to components (cranks) of micromechanical devices; as optical vortices to trap and rotate microscopic particles; and in fiber-optic power transmission to maximize the transmission capabilities by minimizing the concentration of light at the center of the fiber. Helical waves are characterized by the

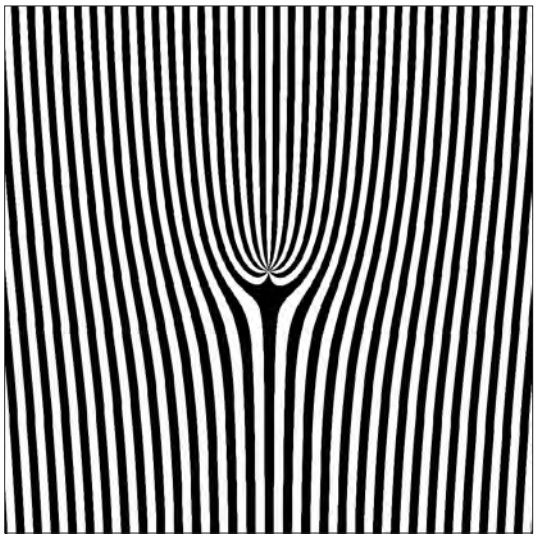

Fig. 7. Computer-generated hologram for topological charge of nine.

number of rotations per wavelength of propagation. This number is called the topological charge or the circular-harmonic order. **Figure 7** shows a computer-generated hologram that can be used to produce a wave with a topological charge of nine.

William J. Dallas

Microwave Holography

Microwave holography is microwave imaging by means of coherent continuous-wave electromagnetic radiation in the wavelength range from 1 mm to 1 m. As a long-wavelength imaging modality, it differs from techniques which employ echo timing (for example, conventional radar) by its requirement for phase information. In this respect it resembles optical holography, from which it has departed significantly. The technique usually involves small-scale systems, that is, systems in which the effective data acquisition aperture is of the order of tens or hundreds of wavelengths. Microwave holographic imaging is characterized by high lateral-resolution capability in comparison with images obtained from echo timing. The natural image format of the data it presents to the human observer enhances its diagnostic potential. In particular, it conveniently produces phase imagery which increases further its diagnostic capability. *See* MICROWAVE; RADAR.

Microwave holographic imaging originated from the two-stage optical process consisting of data recording in the form of an interferogram, and image reconstitution (from a transparency reduced in size because of the larger wavelength) by optical diffraction. The first microwave (and acoustic) holograms were recorded in 1951 before the availability of lasers. The first publicized demonstration of small-scale microwave imaging occurred in 1965 (**Fig. 8**). The object was a metal letter A with a height of approximately 7 ft (2 m), that is, 70 wavelengths, illuminated by X-band microwave radiation with a wavelength of approximately 30 mm. The hologram (approximately 10 × 10 ft or 3 × 3 m) was mapped by recording the field intensity and converting it to a small transparency for image construction by laser light. Subsequently the methods of data recording and the replacement of the optical diffraction process by digital computation transformed microwave holography into a diagnostic imaging technique in its own right.

Data recording. The replacement of the optical diffraction process by computer processing using a fast Fourier transform algorithm has important implications for the data-recording stage. Instead of obtaining the microwave interferogram analogously to the optical process, the microwave field scattered by the object is recorded directly in amplitude and phase by using a microwave receiver which compares the measured field at any point in space with a reference value. For the forward-scatter case (**Fig. 9***a*), the object (which may be semitransparent to microwaves) is illuminated from a microwave source and transmitting antenna T_x. A receiving antenna scans through known coordinates in the surface S and feeds the field values at each point to the receiver. Since a portion of the source energy is fed directly to the receiver by a separate reference channel (either a free space path or a waveguide), the receiver can generate the complex field values (phase and amplitude) at each point. An alternative recording geometry, the backscatter mode (Fig. 9*b*) is the usual configuration for radar systems. The transmitting and receiving antennas are either the same antenna or two antennas close together, as shown. The antennas scan as a unit over the desired surface S, and the complex field values are recorded. Because the illumination from the transmitting antenna also scans the object in this case, the resolution of the system is doubled in comparison with the forward-scatter case.

Fig. 8. Optically reconstructed image using laser light (wavelength equal to 0.6328 μm) of metal letter A 70 wavelengths high at the X band (wavelength of approximately 30 mm). (*From R. P. Dooley, X-band holography, Proc. IEEE., 53:1733–1755, 1965*)

Computer processing. The sampled field values recorded over the surface *S* may be expressed as an array of complex numbers and are therefore suitable for computer processing. The computer algorithm is designed to reconstitute an image from the particular scan geometry used. The process can be thought of as effectively inverting the propagation process that brought the scattered waves to the surface *S*. The inversion process usually incorporates a version of the fast Fourier transform algorithm to convert the data recorded on *S* (not strictly a hologram) into the reconstructed object. The computer transfers its output to a memory and then a television monitor display. The important advantages of this digital microwave holographic process are (1) the availability of numerical field values with high accuracy and low noise; (2) the separate operations on the phase and amplitude values, and the separate display of these values; (3) the possibility of computer-observer interaction at any stage of the processing; and (4) the options of monochrome or false color display format. *See* HARMONIC ANALYZER.

Imaging applications. The efficacy of microwave holography as an imaging modality independent of optical holography is evidenced by a comparison of the microwave image of an object (**Fig. 10***a*) with the optical photograph (Fig. 10*b*). Considerable resolution of detail can be observed in an object that is only 20 wavelengths long at the microwave frequency. However, the role of microwave holography is not to mimic optical holography. Until 1979, perhaps the most useful diagnostic application of microwave holographic imaging was the metrology of large reflector antennas. The data acquisition procedure is a variant of the general forward-scatter case (Fig. 9*a*) since the object itself is scanned in both azimuth and elevation to synthesize the holographic aperture. In this arrangement (**Fig. 11***a*), the test antenna itself feeds the complex field values to the microwave receiver, and so functions simultaneously as the receiving antenna and the object. The transmitting antenna is located either on the ground, in the near field of the test antenna, or on board a synchronous satellite. The image, that is, the conventional notion of an image (Fig. 11*b*), is obtained by quantifying the amplitude distribution over the reflector, and also shows the support legs and the focal region "laboratory." More important is the phase image (Fig. 11*c*), which corresponds to the errors in the reflector profile, that is, deviations from the ideal paraboloidal shape. Other important diagnostic information can be derived, for example, the astigmatism due to gravitational distortion which is apparent in Fig. 11*d*.

Microwave holography is also useful in applications where images of concealed structure are required. Microwave radiation penetrates a variety of dielectric media to a depth depending on the attenuation of a given wavelength in a particular medium. One such application is the mapping of subsurface pipes and cables. A scanning arrangement for this purpose (**Fig. 12***a*) uses the backscatter mode (Fig. 9*b*). The detection of the backscatter from buried pipes, which is very weak after suffering attenuation in the soil, is assisted by the polarization discrimination of the receiving antenna. The data acquisition and computer processing follow the normal procedure with compensation for microwave propagation in the soil. Plastic pipes as well as metal pipes can be imaged (Fig. 12*b*). Hence this noninvasive microwave technique has a diagnostic power greater than the normal metal detectors. *See* NONDESTRUCTIVE EVALUATION.

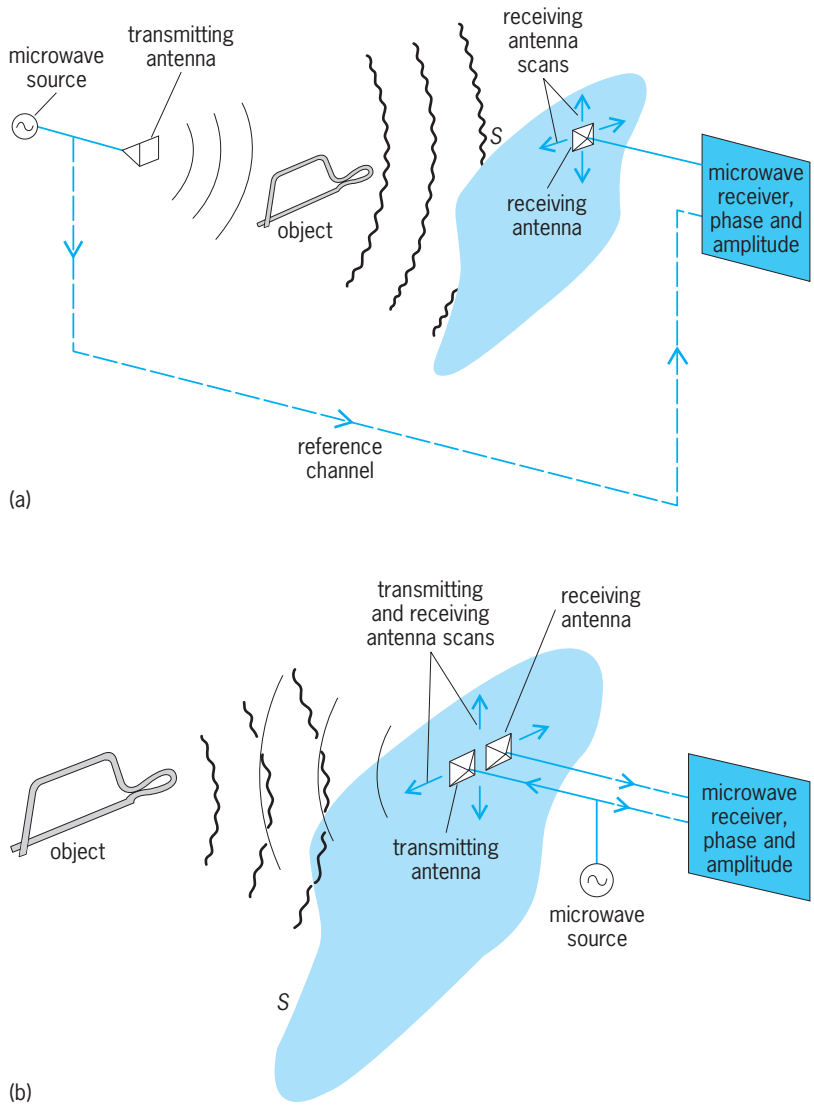

Fig. 9. Typical scan geometries for recording complex field data. (*a*) Forward-scatter mode. (*b*) Backscatter mode.

Microwave tomography. The major limitation of the microwave holographic techniques discussed above is that the images produced are essentially two-dimensional. This may seem surprising, given the fact that optical holography is a three-dimensional image construction process. The reason is that the microwave wavelength is so long (10^4–10^6 times that of light) that the depth of focus of the microwave hologram is prohibitive. This disadvantage is overcome by employing a tomographic mode of imaging which exploits the ability of microwaves to penetrate many materials and thereby characterize their three-dimensional structure more accurately. This development is analogous to the technique of computer-aided tomography used in x-ray scanning

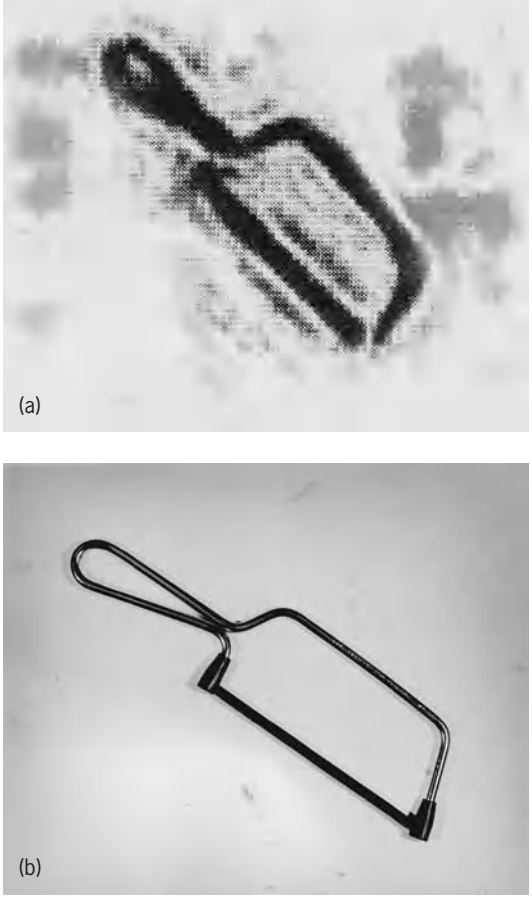

Fig. 10. Comparison of (*a*) digitally reconstructed microwave image of object 20 wavelengths long at the Q band (wavelength of 9 mm), showing reflections from surroundings, with (*b*) optical photograph of object.

systems. Microwave holographic tomography requires holograms to be recorded from different views of the object and synthesized. Again, the availability of phase imagery increases its diagnostic potential. *See* COMPUTERIZED TOMOGRAPHY. Alan P. Anderson

X-Ray Holography

Physicists and life scientists have been engaged in research that will ultimately allow three-dimensional imaging of living organisms with resolution and contrast far beyond the reach of optical microscopes. The impetus for this activity is the imminent availability of high-intensity coherent sources of electromagnetic radiation with wavelengths between 0.1 and 10 nanometers. Much of the study is concentrated on holographic imaging because it can eliminate the need for focusing elements which are difficult to fabricate with enough precision to achieve diffraction-limited resolution in the soft x-ray regime. Furthermore, several of these new sources promise extremely high intensity and subnanosecond pulses, and can circumvent the problem of killing and altering the specimen with the x-ray exposure by extracting an image from the specimen before it is obliterated. *See* X-RAY OPTICS; X-RAYS.

X-ray sources. To be suitable for holography, the x-radiation must be monochromatic and have a relatively high degree of coherence. Synchrotrons using magnetic undulators can generate a narrow band of intense radiation. Use of monochrometers and pinhole apertures can improve the coherence of this radiation at the sacrifice of intensity but with retention of sufficient intensity to image biological specimens on time scales from a few seconds to a few hours. *See* SYNCHROTRON RADIATION.

There are several promising sources. Nonlinear optical frequency multiplication techniques produce intense picosecond pulses of tunable coherent radiation, and have reached wavelengths as short as 40 nm. Similarly, multiphoton excitation can pump atoms to higher energy levels that have lasing transitions at wavelengths much shorter than the excitor laser. X-ray lasers driven by nuclear explosives and by more conventional laboratory sources are under development. X-ray and gamma-ray lasers will be inherently short-pulse, high-intensity devices because they will probably not have resonant cavities, so the radiation being amplified can make only a single passage through the active medium; and the creation and maintenance of a high density of excited atomic states of short lifetime and high quantum energy require enormous power, which

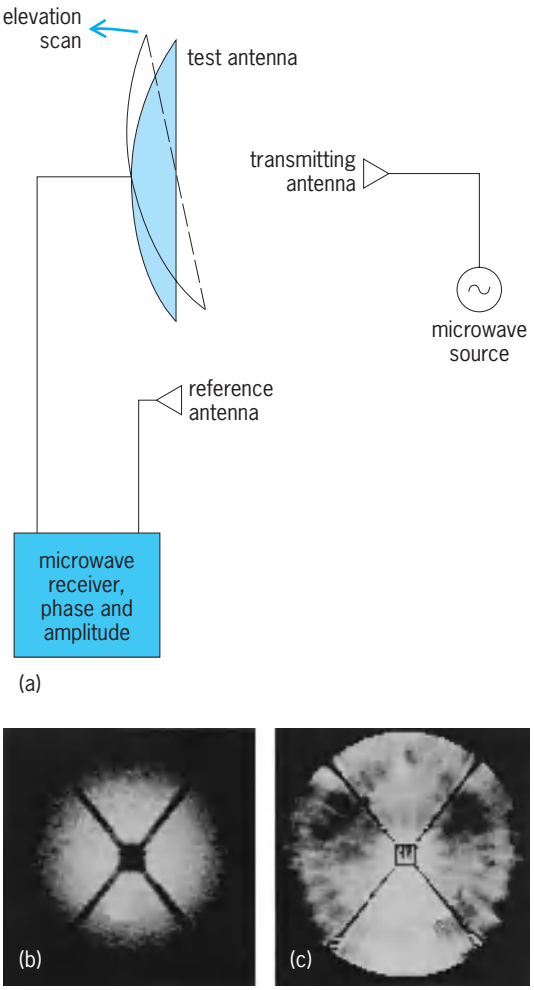

Fig. 11. Applications of microwave holographic imaging to metrology of an 82-ft-diameter (25-m) paraboloidal reflector antenna structure used in satellite communications. (*a*) Scan geometry used for data recording. (*b*) Amplitude image showing aperture illumination distribution. (*c*) Phase image showing deviations from the ideal paraboloid.

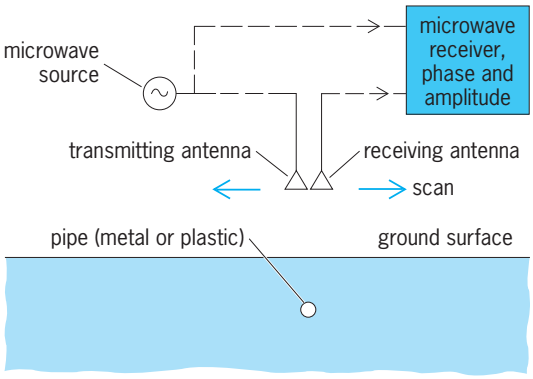

(a)

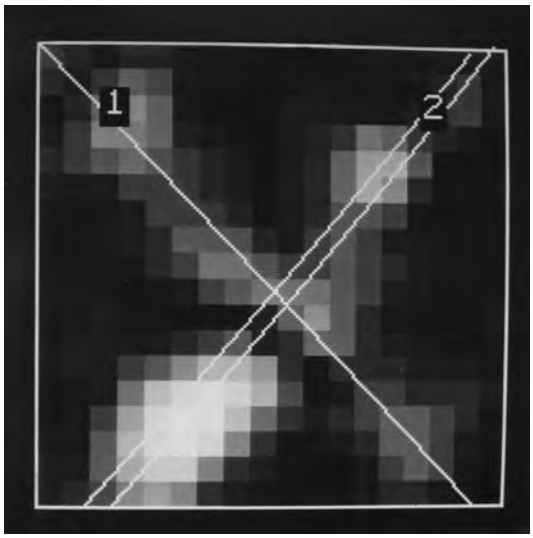

(b)

Fig. 12. Subsurface imaging. (*a*) Scan system. (*b*) Microwave amplitude image, at wavelength of 0.6 m (2 ft), of two crossed pipes whose positions are indicated by solid lines. Pipe 1 is plastic; pipe 2 is metal. Field of view is 7 × 7 ft (2 × 2 m).

terrestrial sources can supply only in the form of pulses. *See* NONLINEAR OPTICS.

Geometries. There are three principal geometries for holography (**Fig. 13**). The Fresnel transform techniques use planar reference waves and have resolution limited by the grain size of the recording medium. The on-axis (Gabor) form is inherently simple but suffers from overlap of the real and virtual images. The off-axis (Leith-Upatnieks) modification reduces the image overlap problem but requires a mirror and a broadened beam for system illumination; both forms may be difficult at x-ray wavelengths. The Fourier transform (Stroke) geometries, using curved wavefronts, achieve large fringe spacings and are therefore less sensitive to grain size.

Coherence is characterized by the effective finite length of a photon wave train in the transverse direction (spatial). Both coherence length and geometry limit the holographable volume of a specimen. For most specimens of biological interest, spatial and temporal coherence lengths of 10 micrometers to 1 nm are adequate.

Interactions of x-radiation. The interaction of x-radiation with matter is quite different from the interaction of visible light with matter. Whereas the extinction of a visible beam traversing matter is mainly due to scattering, the extinction of an x-ray beam is mainly due to absorption. X-rays can also be scattered, but usually the cross section for coherent scattering is very much smaller than for absorption. In the visible regime, holographic images are primarily formed by refraction or reflection, whereas in the x-ray regime they are dominated by diffraction. The greatest contrast in x-ray absorption between water, which composes most of the cytoplasm, and protein (or the nucleic acids) occurs between the K edges of oxygen and nitrogen.

Snapshot x-ray holography. Existing x-ray sources, in particular, synchrotron radiation sources, have been used to make holograms. However, they require long exposures, limiting their usefulness for research on living specimens. More coherent sources may also be developed, but those of low intensity will be similarly limited, since ionization will have decomposed molecules, modified compositions, and altered biological functions before enough radiation can be received to form a useful hologram. Snapshots are essential for x-ray holography of living specimens. Fortunately, it is likely that x-ray sources producing brief intense bursts will be developed.

With an intense pulsed coherent source (such as an x-ray laser), hydrodynamic expansion, initiated by sudden heating, rather than normal biological activity, chemical change, or thermal agitation, will limit the time during which recording of the hologram must be accomplished. Analytical expressions for the explosion of a semiopaque feature (such as a protein

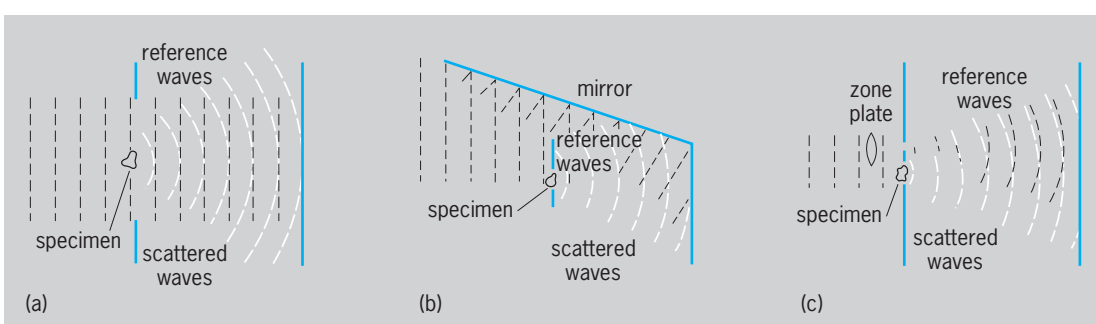

Fig. 13. Geometric configurations of x-ray holographic techniques. (*a*) On-axis Fresnel transform (Gabor holography). (*b*) Off-axis Fresnel transform (Leith-Upatnieks holography). (*c*) Planar Fourier transform (Stroke holography).

globule) are useful for estimating the radiation requirements for typical cases. They are based on the criterion that, to achieve a linear resolution δ, a specified minimum number of photons must have been coherently scattered in a volume δ^3 and that, during the exposure time Δt, no dimension of the specimen should have increased by more than δ. For most biological specimens, intensities on the order of 10^{12} W · cm^{-2} with pulse lengths on the order of 10^{-11} s will be required to obtain an x-ray hologram with resolution of 10 nm.

Recording. An x-ray hologram can be registered by radiation-induced prompt or latent chemical change, or by photoelectron emission. Photographic emulsion is unsatisfactory for Fresnel transform x-ray holography because the resolution is limited by grain size. If an electron microscope could be used to image the points of electron emission from a photocathode reference surface, time-gated holography might be possible. However, the continuous distributions in energy and in angles of emission of electrons from a photocathode preclude the formation of sharp electron-optical images, imposing a tradeoff between quantum efficiency and resolution, unless image-deblurring analysis can be applied. Photoresists (materials that lose resistance to chemical etching at points exposed to radiation) have grain sizes that approach 5 nm, which is entirely adequate for x-ray holograms with resolutions of 10 nm. To reconstruct a photoresist hologram, a transmission electron micrograph could be formed and viewed with visible laser illumination, or a transmission electron microscope could be used to scan and digitize the photoresist for analysis by computation, which can also mitigate nonlinearities that may be troublesome in optical reconstruction. *See* ELECTRON MICROSCOPE; PHOTOEMISSION; PHOTOGRAPHIC MATERIALS.

By using Fourier transform x-ray holography, it is possible to arbitrarily adjust the fringe spacing at the sacrifice of intensity and thereby record with common photographic emulsions. However, when compared with photoresists on the basis of number of quanta required to produce a developable speck, it is not clear that the greater sensitivity of photographic emulsions offers any advantage, and consequently there is no clear advantage of Fourier over Fresnel methods.

Practical considerations. The realization of x-ray holography as a practical research tool still awaits the solution of some challenging technical problems:

1. Development of sources that can generate intense coherent radiation at the precise wavelengths to optimize contrast among specimen constituents. Perhaps nonlinear mixing with tunable visible radiation will be necessary.

2. Termination of exposure within a sufficiently brief time interval and with sufficient intensity to achieve the desired resolution, as discussed above. Frequency multiplication techniques and multiphoton excitation lasers can achieve these short pulses because the optical laser driving them can be modelocked. Corresponding schemes are difficult to envisage for x-ray or gamma-ray lasers, and their pulse lengths are likely to be much longer. A shutter or gate, somewhere in the system, that operates when full intensity is reached will be essential. *See* OPTICAL PULSES.

3. In principle, photoelectric recording could be time-gated. However, complexity and precision required of the electronics, and blurring associated with initial electron velocity distribution make this approach unattractive. On the other hand, exposure control in photoresist recording is not likely to be managed by a gate; therefore exposure control must be provided elsewhere in the system.

4. Leith-Upatnieks holography may be necessary to avoid image overlap obscuration, and this requires an x-ray mirror. A synthetic Bragg crystal may suffice, and thermal expansion, if sufficiently uniform, can provide automatically time-gated reflection.

Johndale C. Solem

Bibliography. M. F. Adams and A. P. Anderson, Synthetic aperture tomographic imaging (SAT) for microwave diagnostics, *Proc. IEEE*, 129:83–88, 1982; A. P. Anderson, Microwave holography, *Proc. IEEE*, 124:946–962, 1977; S. A. Benton, Hologram reconstructions with extended incoherent sources, *J. Opt. Soc. Amer.*, 59:1545–1546A, 1969; B. R. Brown and A. W. Lohmann, Complex spatial filtering with binary masks, *Appl. Opt.*, 5:967–969, 1966; H. Dammann, Blazed synthetic phase-only holograms, *Optik*, 31:95–104, 1970; Y. Denisyuk, Photographic reconstruction of the optical properties of an object in its own scattered field, *Sov. Phys. Doklady*, 7:543–545, 1962; J. W. Goodman, *Introduction to Fourier Optics*, 3d ed., 2005; P. Hariharan, *Basics of Holography*, 2002; J. E. Kasper and S. A. Feller, *The Complete Book of Holograms: How They Work and How To Make Them*, 1987, reprint 2001; T. Kreis, *Handbook of Holographic Interferometry*, 2005; S. H. Lee (ed.), *Selected Papers on Computer-Generated Holography and Diffractive Optics*, 1992; E. N. Leith and J. Upatnieks, Photography by laser, *Sci. Amer.*, 212(6):24–35, 1965; L. B. Lesem, P. M. Hirsch, and J. A. Jordan, Jr., The Kinoform: A new wavefront reconstruction device, *IBM J Res. Dev.*, 13:150–155, 1969; A. W. Lohmann, Optische Einseitenbanduebertragung angewandt auf das Gabor-Mikroskop (Optical single-sideband transmission applied to the Gabor microscope), *Opt. Acta*, 3:97–99, 1956; J. Ludman, H. J. Caufield, and J. Riccobono (eds.), *Holography for the New Millennium*, 2002; A. J. MacGovern and J. C. Wyant, Computer generated holograms for testing optical elements, *Appl. Opt.*, 10:619–624, 1971; G. Saxby, *Practical Holography*, 3d ed., 2003; U. Schnars and W. Jüptner, *Digital Holography*, 2005; J. Solem and G. Baldwin, Microholography of living organisms, *Science*, 218:229–235, 1982.

Holostei

An unranked group of the fish subclass Neopterygii consisting of several fossil orders and the extant orders Lepisosteiformes and Amiiformes. The holosteans are descended from the older

Chondrostei and in turn are ancestral to the great mass of modern bony fishes, the Teleostei. It is not certain whether holosteans evolved from a single stock or multiple stocks of the Chondrostei. *See* TELEOSTEI.

Phylogeny. Holosteans made their first appearance in the Upper Permian as the order Semionotiformes; three additional orders arose in the Triassic Period, and the fifth and last order, the Aspidorhynchiformes, evolved in the Middle Jurassic. In the Jurassic and Lower Cretaceous, holosteans dominated actinopterygian fish life, but by the Late Cretaceous they had been largely replaced by teleosts. The specialized Pycnodontiformes persisted until the Eocene, but the spindle-shaped, predacious Aspidorhynchiformes and the Pholidophoriformes, which are likely ancestors of the Teleostei, died out in the Cretaceous. Fragmentary remnants of two large orders persist to the present time as survivors of the Mesozoic: Lepisosteiformes, as the North American and Middle American gars (family Lepisosteidae), and Amiiformes, as a single species, the bowfin (family Amiidae) of eastern North America. Understandably, both groups have been intensively studied by biologists in search of clues to the life of the past.

Morphology. Holosteans, although highly varied in body form, were structurally as well as temporally intermediate between chondrosteans and teleosts, to which group they passed on substantial advances. Mouthparts were improved by horizontal suspension of the hyomandibular from the skull, a more forward positioning of the angle of the gape, and the development of a strong coronoid process on the mandible. The maxilla was freed posteriorly, and the entire feeding mechanism became more mobile and was strengthened, thus permitting diversification in food habits, though most holosteans were predacious. The caudal fin is typically abbreviate heterocercal, with the posterior vertebrae upturned but not forming a long upper lobe. Typically the anterior upturned centra each support a single, slender hypural. Dorsal and anal fin rays are strengthened and reduced in number to approximate serial equivalence with the internal supports. In early forms scales are often thick and rhomboidal, as in chondrosteans, but in certain advanced types are thin and rounded; they retain an enamellike outer layer (ganoine) that is lost in all but the earliest teleosts. In living holosteans the swim bladder is highly vascularized, and auxiliary aerial respiration is possible, a sometimes essential faculty in oxygen-poor waters of swamps. *See* ACTINOPTERYGII; AMIIFORMES; ASPIDORHYNCHIFORMES; PYCNODONTIFORMES; SEMIONOTIFORMES.

Reeve M. Bailey

Holothuroidea

A class of Echinozoa characterized by a cylindrical body and smooth leathery skin, and known as sea cucumbers. There are no arms, but a ring of five or more tentacles may surround the mouth, which is usually at one end of the body. There are no pedicellariae. Tube feet may be present or lacking. There are no ambulacral grooves, although they are represented by internal epineural canals overlying the radial nerves. E. Deichmann (1957) regards them as the most aberrant group of extant echinoderms. *See* ECHINOZOA.

Holothurians resemble worms because the pentamerous symmetry is largely concealed by a secondary bilateral symmetry, and the general absence of external spines distinguishes them from the other extant echinoderms. They tend to rest on one side, so that the axis of radial symmetry becomes horizontal. This habit leads to the differentiation of an upper (dorsal) surface and a lower (ventral) one. The dorsal side corresponds to the interradius which contains the madreporite, and therefore the ventral side is one of the radii. Each radius runs from the anterior to the posterior end. If tube feet are developed, their disposition indicates the radii (**Fig. 1**).

The 1100 living species have been grouped in 170 genera arranged in six orders: the Dendrochirotida, Dactylochirotida, Aspidochirotida, Elasipodida, Molpadida, and Apodida. Species range in size from 1.2-in. (3-cm) body length up to 5 ft (1.5 m); the largest types are tropical Synaptidae. Colors vary widely; the most brilliant colors are found among the Synaptidae. Yellow, red, violet, and fawn tints occur, but many species are somber shades or black. *See* ECHINODERMATA.

Relation to humans. Some holothurians are esteemed as food, in particular about 20 species of the genera *Stichopus* and *Holothuria*. These are fished mainly in the Indian and Pacific oceans. The animals are boiled for 20 min in seawater, smoked, and dried. The product is then called chekin, trepang, or bêchede-mer. The annual export trade to the Far East and Mediterranean countries exceeds 10,000 tons (9072 metric tons). In Italy also, fishermen take local species, but these are not esteemed.

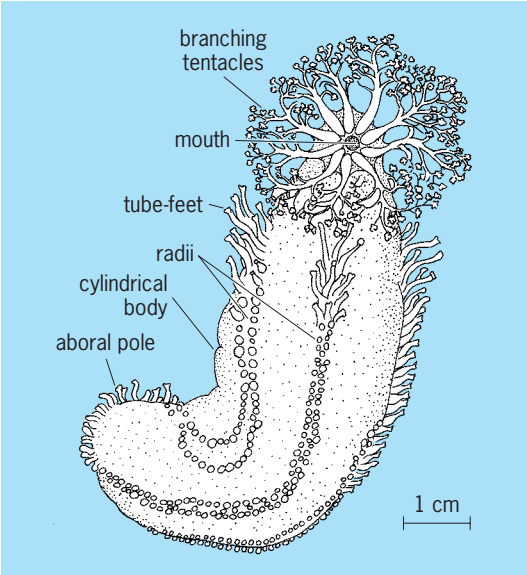

Fig. 1. ***Cucumaria*, a representative holothurian.**

A few holothurians are poisonous, especially to fish. In the Pacific some species of *Actinopyga* and *Holothuria* yield a secretion used by the islanders as an aid in fishing rock pools, in the same way as rotenone.

Ecology. The Dendrochirotida and Apodida have more restricted ranges than the other orders, which are believed to be global. Holothurians occur in all seas, from low-tide level down to the greatest depths explored. The Vitiaz expedition in 1957 took a holothurian from the Tonga Trench at a depth of $6^1/_2$ mi (10,415 m). At depths below $5^1/_2$ mi (8850 m) holothurians comprise 90% of the total mass of living matter, the rest being mainly starfishes. Two pelagic genera are known.

Parasites include protozoa, flatworms, nematodes, and annelids. Some crabs, such as *Pinnotheres*, inhabit the cloaca or respiratory trees, and another, *Lissocarcinus*, lurks between the tentacles. Gastropods such as *Entoconcha* and *Entocolax* bore into the skin, body cavity, or foregut. A fish, *Fierasfer*, inhabits the cloaca of large Aspidochirotida.

Skeleton. This structure usually comprises no more than a ring of 10 (or 5) calcareous plates around the esophagus, and numerous small platelets or spicules scattered in the skin. In Psolidae the skin plates are large and overlap like scales, but they do not form radial and interradial series. The skin platelets may assume distinctive forms, such as anchors, wheels, or other shapes, and may be useful as taxonomic characters.

Muscular system. The muscular system is well developed, and the body can assume a variety of shapes. The chief muscles are five radial longitudinal bands in the body wall; outside these are transverse fibers. Some species have pharyngeal retractor muscles which invert the anterior part of the body. This feature is used as an aid to classification.

Alimentary system. The viscera lie in the coelom, which contains a fluid. The mouth leads into a short esophagus, which connects to the long and usually looped intestine. Its front part may be differentiated as a stomach, and the posterior loop serves as a rectum, opening at the cloaca (**Fig. 2**).

Most holothurians are mud swallowers, living on the bacteria and other organic material present in the substrate. The Dendrochirotida, however, feed on plankton, which is caught on their sticky tentacles and transferred to the mouth. Some Aspidochirotida have paired cloacal glands, the Cuvierian organs, which extrude a mass of sticky threads to trap small animals and to repel predators.

Respiration. Respiration occurs partly in the skin and tube feet and sometimes in tubular branches of the gut, the respiratory trees. Rhythmic contractions of the cloaca cause seawater to flow to and fro in them.

Hydrocoele. Although poorly developed in the Molpadida, and vestigial in the Apodida, the water-vascular system is typical in the other three orders. The one or more madreporites lie in the dorsal interradius, internal in nearly all cases, and open into the coelom. The single stone canal which arises from them runs downward to the ring vessel around the esophagus. A single polian vesicle lies in one interradius (Fig. 2). The radial vessels run forward to the base of the tentacle ring and then backward in the body wall, following the radii. The oral tentacles represent specialized tube feet, and may have ampullae.

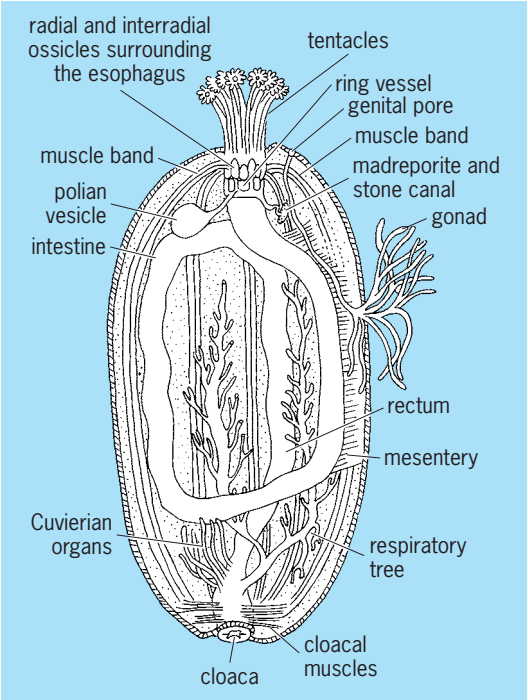

Fig. 2. Holothurian anatomy.

Nervous system. This system is typical for the phylum, following the pattern of the water-vascular system, lying between the canals and the body wall. The whole surface of the body is sensitive to light, but some Synaptidae have in addition an eyespot at the base of the tentacles. These are analogous to the eyespots of starfishes and, like them, do not form an image, but are able to detect shadow movements and thus indicate the presence of potential enemies or prey.

Life history. Holothurians seem to attain sexual maturity after about 3 years and continue to grow for several more years. The life span may therefore be longer than in sea stars and sea urchins, perhaps 10 years.

Reproductive system. The reproductive system comprises either one or two gonads, with their ducts. The gonopore is usually anterodorsal (Fig. 2). The sexes are usually distinct. Hermaphroditism is common in the Apoda, but individuals are either male or female at any given time, so self-fertilization does not occur. The life history may include a free-swimming larva (auricularia). Many cold-water species incubate their young in various ways. *Leptosynapta minuta* carries the developing embryos in the coelom until they escape by way of the cloaca. Some species of *Psolus* carry the young in dorsal pockets in the skin, covered by calcareous plates. Others shelter the eggs and young by lying on them. A few species, such as *Cucumaria planci*,

habitually reproduce asexually by spontaneous transverse fission. Most species can regenerate lost organs. Autoevisceration is a peculiar protective habit which is often displayed in response to interference. The whole gut, the gonads, and respiratory trees are ejected through the cloaca and cast off. Regeneration of the lost structures occurs from the torn mesenteries, and requires several months. During regeneration the animal cannot feed; instead, the muscles are slowly digested in place. Howard B. Fell

Holotrichia

A major subclass of the class Ciliatea. These protozoa have a fairly uniform body ciliation, as the name implies. Separate articles appear on the groups listed in the following classification:

Subclass Holotrichia
- Order: Gymnostomatida
  - Trichostomatida
  - Chonotrichida
  - Apostomatida
  - Astomatida
  - Hymenostomatida
  - Thigmotrichida

The cilia are typically arranged in longitudinal rows over the body, although scattered exceptions exist. Cirri are absent, and buccal membranelles are absent or inconspicuous, except in some members of two groups. A mouth is often, although not always, present. Ciliary organelles associated with it are generally not conspicuous. Trichocysts are present in species belonging to a number of families.

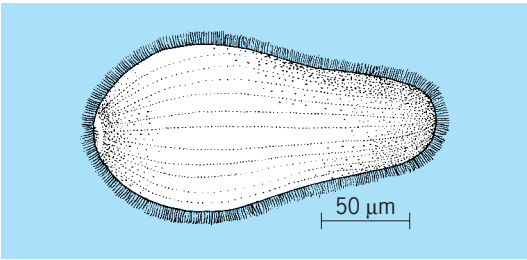

***Prorodon*, a primitive holotrich.**

The seven orders contain several thousand species and include groups of, presumably, the most primitive ciliates living today. The prototype of the Holotrichia is exemplified by the form portrayed in the **illustration**. *See* CILIATEA. John O. Corliss

Homalozoa

A subphylum of echinoderms, made up of members having a flattened theca or body lacking pentameral symmetry. Homalozoans (also called carpoids) include four extinct classes of relatively uncommon primitive echinoderms ranging in age from the Early or Middle Cambrian to the Late Carboniferous. Homalozoans have a flattened, asymmetrical to bilaterally symmetrical theca often composed of a marginal frame of large elongate plates surrounding top and bottom central areas that had numerous smaller plates and were probably flexible. One homalozoan class (Ctenocystoidea) has no attached appendages; two classes (Stylophora and Homostelea) have a single plated armlike or taillike appendage used for feeding or swimming; and one class (Homoiostelea) has two plated appendages at opposite ends of the theca, one used for feeding, the other for swimming. The mouth is located either on one edge of the theca or at the base of the armlike appendage, while the anal pyramid is usually at the opposite margin or corner of the theca. All homalozoans were apparently mobile, benthic, detritus or suspension feeders that had adopted a flatfish way of life. Because they have a typical echinoderm skeleton with single-crystal, multiporous, sutured, calcite plates, most researchers consider homalozoans to be true echinoderms. However, they seem only distantly related to other echinoderms that have well-developed pentameral symmetry. It has been argued that one class of homalozoans (stylophorans or calcichordates) is intermediate between echinoderms and several groups of chordates, thus making the class directly ancestral to the vertebrates. However, this proposal has not been widely accepted by other echinoderm or vertebrate paleontologists. *See* CARPOIDS; ECHINODERMATA. James Sprinkle

Bibliography. R. P. S. Jefferies, *The Ancestry of the Vertebrates*, 1986; R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, pt. S, 1968.

Homeostasis

Either the relatively constant conditions within organisms or the physiological processes by which such conditions are maintained in the face of external variation.

Similar homeostatic controls are used to keep factors such as temperature and blood pressure nearly constant despite changes in an organism's activity level or surroundings. Such servosystems, which are of wide technological use as well, operate by detecting changes in the variable that the system is designed to hold constant and initiating some action that offsets any change. All incorporate a sensor within the system that responds when the actual condition differs from the desired one, a device to ensure that any action taken will reduce the difference between actual and desired, and an effector to take the needed action as directed. The crucial aspect is that information is fed back from effector to sensor and action is taken to reduce any imbalance—hence the term negative feedback.

Biological mechanisms. Blood pressure is, at least on a moment-to-moment basis, regulated by a system for which the sensors are stretch-sensitive cells located in the neck arteries that carry blood from heart to brain. An increase in blood pressure triggers

sensor activity; their signal passes to the brain; and, in turn, the nerve supplying the heart (the vagus) is stimulated to release a chemical (acetylcholine) that causes the heart to beat more slowly—which decreases blood pressure.

The volume of the blood is subject to similar regulation. Fluid (mainly plasma) moves between the capillaries and the intercellular fluid in response to changes in pressure in the capillaries. A decrease in blood volume is detected by sensors at the base of the brain; the brain stimulates secretion of substances that cause contraction of tiny muscles surrounding the blood vessels that lead into the capillaries. The resulting arteriolar constriction reduces the flow of blood to, and the pressure within, the capillaries, so fluid moves from intercellular space into capillaries, thus restoring overall blood volume.

Body temperature in mammals is regulated by a sensor that consists of cells within the hypothalamus of the brain. If that area is experimentally heated, the rest of the body cools off. Several effectors are involved, which vary among animals. These include increasing heat production through nonspecific muscle activity such as shivering; increasing heat loss through sweating, panting, and opening more blood vessels in the skin (vasodilation); and decreasing heat loss through thickening of fur (piloerection) and curling up. Humans (but not pigs, cats, or dogs) sweat, but they retain only a vestige of piloerection ("goose flesh"). *See* THERMOREGULATION.

Regulation using negative-feedback servosystems occurs in more localized neuromuscular systems as well. When stimulated by nerves, muscles pull harder—how much shorter the muscle becomes depends on the load it is pulling against. But one often makes a muscle shorten by a fixed and load-independent amount. It is done with sensors within the muscle that effectively monitor its length. Extra load (for example, a book laid on an outstretched arm) stretches muscles such as the biceps; that stretch is detected and stimulates the brain to send more frequent nerve impulses to the muscle, restoring its original length—but at the higher tension needed for its greater load. Similar machinery permits a person standing with slightly bent knees to lift one leg off the ground without collapsing—the tensions of the extensor muscles of the opposite leg are doubled to offset the doubling of their load. *See* MUSCLE.

While the homeostatic mechanisms first described involved the neural and endocrine systems of mammals, it is clear that such arrangements pervade systems from genes to biological communities, and that they are used by the simplest and the most complex organisms. If a bacterium such as *Escherichia coli* is grown in a culture medium that lacks the amino acid tryptophan, it reacts to that absence by turning on the appropriate genes and making the enzymes that it needs to synthesize tryptophan. If the same bacterium is transferred to a medium in which lactose rather than glucose is the sole energy source, it begins to make the enzymes necessary to live on lactose. In either case the bacterium does what is necessary to maintain the level of some substance within itself.

Other mechanisms. Servosystems of the kind described above are limited in their response speed by the necessity for the external change to affect the internal sensor. In practice, organisms commonly overlay such systems with anticipatory devices. The initial heat-conserving response to a sudden chill is triggered by information from cold-sensitive receptors in the skin. Since alteration in blood flow to the skin is part of the response to cold and since skin temperature is not closely regulated, these receptors cannot function as the primary source of the feedback signal—but they do initiate a rapid response. Similarly, when given access to water, a mammal deprived of water drinks rapidly. The animal stops drinking when sufficient water has been consumed to restore its normal content. But that happens well before the main detectors for water content (detectors that monitor blood composition) have been affected.

Organisms of every kind develop, mature, and even shift physiological states periodically—between day and night, with seasons, or as internal rhythms. Thus organisms cannot be considered constant except over short periods. However, all such changes appear to involve the same basic sensing of the results of the past activity of the system and the adjusting of future activity in response to that information. Development of an organism from a fertilized egg is far from a direct implementation of a genetic program; probably no program could anticipate all the variation in the external context in which an organism must somehow successfully develop. *See* BIOLOGICAL CLOCKS; ENDOCRINE MECHANISMS; NERVOUS SYSTEM (VERTEBRATE); SERVOMECHANISM.

Steven Vogel

Bibliography. C. A. R. Boyd and D. Noble (eds.), *The Logic of Life: The Challenge of Integrative Physiology*, 1993; D. D. Chiras, *Human Biology: Health, Homeostasis and the Environment*, 3d ed., 1999; S. Vogel, *Vital Circuits*, 1993.

Homeotic (Hox) genes

The formation of a normal plant or animal body structure or organ in place of another at an abnormal site. Examples of homeosis (also called homeotic transformation) are most obvious in insect appendages, where an appendage that is characteristic of one segment, for example the antennae on an insect head segment, are transformed into insect legs that normally develop only on trunk segments (see **illus**.). Similar examples of homeotic transformations can also occasionally be found in vertebrates where lumbar vertebrae are transformed into thoracic vertebrae which then extend into rib processes, or in floral organs where petals are transformed into sepals. Homeotic transformations rarely occur in nature in living organisms, and are due to genetic defects in a class of proteins called homeotic proteins, the products of homeotic genes. In addition, homeotic

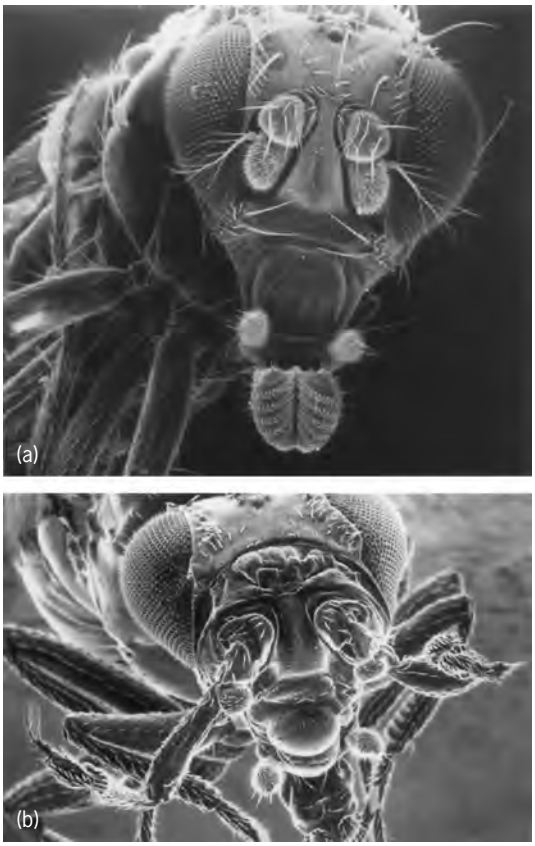

Scanning electron micrographs of the head of *Drosophila melanogaster*. (*a*) Normal individual; the antennae are the small, paired, bulbous structures between the large, faceted eyes. (*b*) An *antennapedia* mutant individual; leglike appendages replace the normally small, bulbous antennae.

transformations are due to the accidental or deliberate manipulation of homeotic gene expression so that homeotic proteins are produced in the wrong place or at the wrong time in developing plants and animals.

Animals. Homeotic transformation is best understood in animals, particularly in the fruit fly *Drosophila*. During early embryogenesis, the anterior-posterior axis of most animals is divided into segmental units or metameres. These metameres may be thought of as the basic building blocks of body form, usually incorporating large numbers of individual cells. During normal development, metameric building blocks of cells require additional instructions to indicate whether they are fated to develop as head, thoracic, or abdominal regions of the body. Thus, the cells of different metameres express unique combinations of the different homeotic proteins that act as developmental switches that assign different fates to different members of this serial array of embryonic developmental fields. For example, if a combination of homeotic proteins that usually assigns abdominal development is produced throughout the entire *Drosophila* embryo, the result is an animal that consists of reiterated series of abdominal segments (and the animal rapidly dies for obvious reasons) in place of normal head and thoracic segments. *See* FATE MAPS (EMBRYOLOGY).

There are eight *Drosophila* homeotic genes, which are specialized to assign head, thoracic, or abdominal development. These genes map in clusters on chromosome 3 of the fly in an array where the order of the genes in the cluster mimicks the order of the metameres in the body they control, with the head genes on the left, the thoracic genes in the middle, and the abdominal genes on the right. Each of these genes is similar at the molecular level in that they all contain a homeobox DNA sequence, which specifies a similar but distinct homeodomain of 60 amino acids in the different homeotic proteins. The homeodomain allows homeotic proteins to bind to the regulatory sequences of many other genes, and the regulation of these other genes presumably explains the global effects that homeotic proteins can induce during development.

Most animals use a similar genetic system to control which structures will develop in which positions on the head-tail axis of the body. Mice and humans have four clusters of homeobox-containing genes (called Hox, for homeobox genes) that are very similar to the *Drosophila* homeotic gene clusters. The order of the genes in the fly versus mammal is conserved, and the mammalian genes express homeodomain proteins that are produced in unique combinations on the developing head-tail axis of the embryos.

The common ancestor to present-day fruit flies and humans (probably a wormlike creature that existed 600 million to 1 billion years ago) probably used a cluster of proto-homeotic genes to assign head, trunk, and tail regional identities to embryonic cells of the body plan. Apparently, after this system evolved, it proved so useful (or so difficult to eliminate) for the animals with it that the system is still in use in most or all present-day animals to assign embryonic axial positional identities. In fact, at an early stage of embryogenesis (but not the earliest stages), invertebrate and vertebrate embryos are quite similar in at least these developmental control molecules.

Plants. A flower consists of four organ types arranged in whorls which can be thought of as developmentally similar to, though not evolutionarily homologous to, animal metameres. From innermost to outermost, these whorls are carpels, stamens, petals, and sepals. Plant homeotic genes that are involved in flower development are not structurally similar to animal homeotic genes. Many of the plant homeotic genes belong to the MADS-box gene family, which shares some important characteristics with the animal homeotic genes. For example, different floral MADS-box homeotic proteins are produced early in flower development only in a subset of the whorls of the floral primordia, and the combination of floral homeotic proteins determines whether cells in a certain whorl will develop as petal or carpel. Another shared characteristic is that the MADS-box proteins are also deoxyribonucleic acid-binding proteins that regulate the expression of other genes, and so produce their effects on the development of the flower by modulating which of a number of other genes are active or inactive in various subregions of the

developing flower. *See* CELL DIFFERENTIATION; DEVELOPMENTAL BIOLOGY; DEVELOPMENTAL GENETICS; MUTATION. William McGinnis

Bibliography. W. Bateson, *Materials for the Study of Variation*, 1894; T. Jack, L. L. Brockman, and E. M. Meyerowitz, The homeotic gene APETALA3 of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens, *Cell*, 68:693–697, 1992; W. McGinnis and R. Krumlauf, Homeobox genes and axial patterning, *Cell*, 68:283–302, 1992.

Homing

A process of navigation by which a destination is approached by keeping some navigation parameter constant. In its early uses, the most commonly chosen parameter was the relative bearing from the vehicle to destination as determined from a signal emitted at or near the destination point. The vehicle then steers to travel in the direction of its destination. The signal can be of many forms, ranging from a visual image to a radio wave or even an odor. This simple form of homing requires minimal on-board equipment, but the path taken by the vehicle over the Earth's surface is influenced by vehicle drift due to winds, currents, or other causes. *See* DIRECTION-FINDING EQUIPMENT.

Use of navigation aids. A higher level of homing is available through certain radio navigation aids (navaids), such as the very high frequency omnidirectional range (VOR) and the aircraft instrument landing system (ILS). While the signals from these aids define a path in space that can be followed by the user, the signals themselves do not indicate the courses that should be selected to remain on the path. In the case of VOR, the signals define the azimuth to or from the station, with the user being required to adjust the vehicle heading to acquire the desired path and to compensate for any drift in order to remain on the path. In common usage, the user specifies the azimuthal radial to be followed, and then the user equipment presents indications of any deviations from the desired path. When treated properly, the ILS signals define a precise fixed path in space, which the user equipment processes to yield deviation indications in the lateral and vertical directions. These aids require only limited on-board databases for effective use. *See* ELECTRONIC NAVIGATION SYSTEMS; INSTRUMENT LANDING SYSTEM (ILS).

Direct-to paths. The advent of affordable on-board digital processing accompanied by extensive geographical databases has resulted in supporting direct-to paths from present position to destinations that need not radiate a signal. These paths can be defined to the precision supportable by the navigation aids in use, and modern equipment generally presents the user with guidance indications to acquire and remain on the desired path. The user interface is essentially indistinguishable from that presented by classic homing devices. Almost any geographically referenced navigation aid can be used to support this technique, but any errors in the database are reflected directly into destination or path errors. There is some disagreement as to whether traveling a direct path to a destination that does not radiate distinctive signals is a true homing technique.

The navigation aids used most for establishing direct-to paths include VOR/DME (distance-measuring equipment), DME/DME, Loran-C, and the Global Positioning System (GPS). GPS has the advantage that it is the only global, all-weather precision navigation aid that can be used for all phases of navigation. *See* DISTANCE-MEASURING EQUIPMENT; LORAN; SATELLITE NAVIGATION SYSTEMS.

This form of navigation requires that the geographical location of the destination be known to the user. The requirement rules out use of the technique to home on vehicles in motion unless the location in time of the vehicle is also known to the user.

Vehicle-to-vehicle homing. One of the most useful navigation features developed after the advent of radionavigation is vehicle-to-vehicle homing, where a vehicle homes on signals radiated by another vehicle. This is especially valuable in rescue operations. For example, when a ship responds to an emergency distress signal from a foundering ship, it uses its radio direction finder to determine the direction of the transmission from the ship in distress, and then can proceed to the location of the ship. Additionally, lifeboats are equipped with special rescue beacons to furnish a transmission on which the rescue vessel can home.

In the United States, aircraft are required to carry emergency locator transmitters (ELTs). These small transmitters, operating at 121.5 MHz for civil and 243 MHz for military aircraft, automatically turn on when the aircraft crashes. Search-and-rescue teams, both airborne and ground based, can then home on the crash site by using direction-finding techniques. For ships, there is a comparable location transmitter, the emergency position indicating radio beacon (EPIRB). These transmitters are augmented by a satellite-based search-and-rescue (SAR) network named SARSAT, which operates at 121.5, 243, and 406 MHz. Retransmission of the original ELT or EPIRB transmissions allows ground stations to process the data from these transmissions to estimate approximate locations of these transmitters. This can allow direction of search vehicles to the area where they can home in on the transmitters directly, thus reducing search times greatly. There is a Russian counterpart to SARSAT called COSPAS. SARSAT involves placing transponders in National Oceanic and Atmospheric Administration (NOAA) polar-orbiting satellites and Geostationary Operational Environmental Satellites (GOES). At present, signals intercepted by the GOES satellites are used for alerting and identification of 406 MHz signals only; they cannot be used for position fixing.

Military missions. The most serious military homing mode is that which enables a missile to home on any source of radiation. Typical sources are radio transmitters of all kinds, including radar and navigation aids, and the infrared exhausts of jet engines,

tanks, and ships. Countermeasures against such activity include the use of short bursts of transmission, frequency hopping, high physical mobility, and decoys and chaff to attempt to attract the missile away from its intended target. *See* ELECTRONIC WARFARE; GUIDED MISSILE; MISSILE. Eugene O. Frye

Bibliography. *Air Navigation: Homing*, U.S. Air Force Manual 51–40, 1954; N. Bowditch, *The American Practical Navigator*, 2 vols., reprint 2002; M. Kayton and W. Fried, *Avionics Navigation Systems*, 2d ed., 1997; U.S. Department of Transportation, *Federal Radionavigation Plan*, biennially.

Homocysteine

A sulfur-containing amino acid that is structurally and metabolically related to the essential amino acid methionine (see structures). Discovered by Vincent du

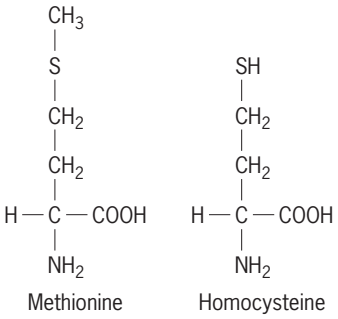

Vigneaud in 1932, homocysteine is found in both prokaryotic and eukaryotic organisms. It is a sensitive marker of nutritional status, vitamin deficiency, and kidney function.

It has been known since the 1960s that elevated levels of homocysteine and homocystine (the disulfide form of homocysteine) are found in the blood and urine of children with certain hereditary diseases of methionine metabolism. Clinical features of these rare diseases include skeletal deformities, abnormalities of the lens of the eye, mental retardation, abnormal blood clots, and premature death. Recent interest in homocysteine has been stimulated by emerging evidence that elevated blood levels of homocysteine are linked to coronary heart disease, stroke, Alzheimer's disease, and birth defects.

Homocysteine metabolism. Unlike most amino acids, homocysteine is not incorporated into polypeptide chains during protein synthesis but functions as a key intermediate in the metabolism of methionine (see **illustration**). Methionine is derived from dietary protein and is the direct precursor of the ubiquitous methyl donor, *S*-adenosylmethionine (SAM). By donating one-carbon methyl groups to hormones, neurotransmitters, nucleic acids, phospholipids, proteins, and other substrates, SAM serves an essential function in mammalian homeostasis. Homocysteine is produced as a by-product of SAM-dependent methyl transfer reactions.

Homocysteine occupies a critical branch point in the methionine metabolic cycle. Depending on the metabolic need for methionine and SAM, homocysteine either can serve as a substrate for the regeneration of methionine or can be diverted from the methionine cycle to produce cystathionine and, ultimately, cysteine (see illustration). The major

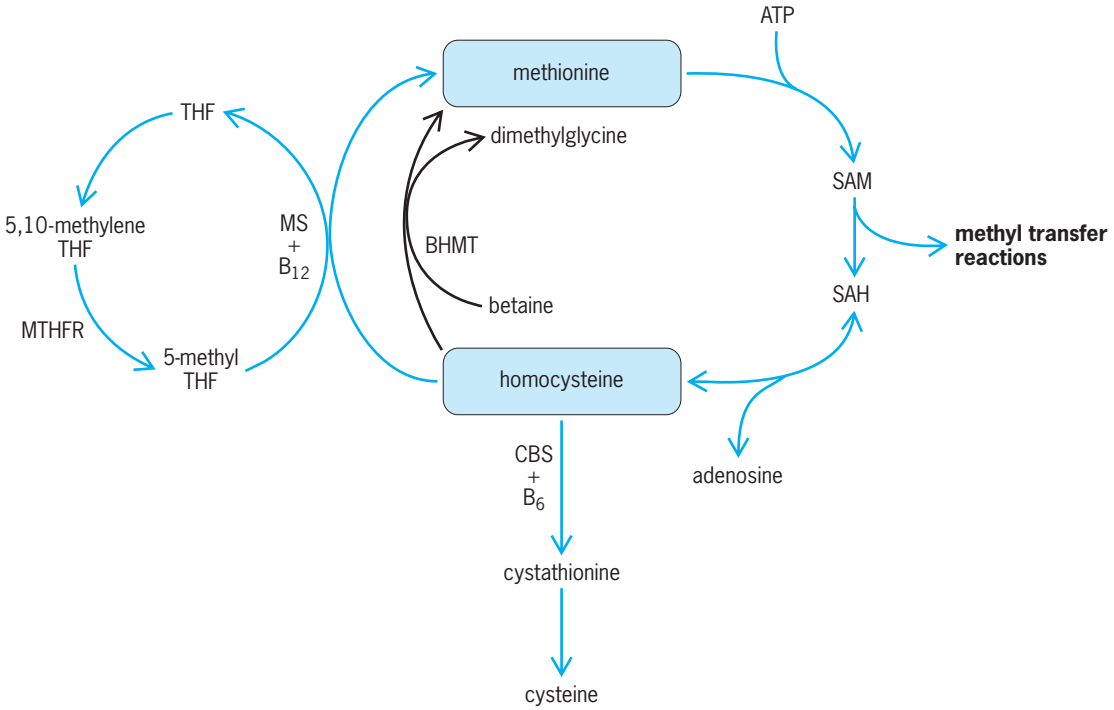

Homocysteine metabolism. CBS, cystathionine ß-synthase; MTHFR, 5,10-methylene tetrahydrofolate reductase; MS, methionine synthase; BHMT, betaine:homocysteine methyltransferase; SAM, *S*-adenosylmethionine; SAH, *S*-adenosylhomocysteine; 5-methyl THF, 5-methyltetrahydrofolate; 5,10-methylene THF, 5,10-methylene tetrahydrofolate; B_6, vitamin B_6; B_{12}, vitamin B_{12}. Black arrows indicate an alternate pathway that does not require B vitamins.

pathway for the regeneration of methionine from homocysteine in most tissues is catalyzed by the enzyme methionine synthase (MS). The MS reaction requires two B vitamins, methylcobalamin (a form of vitamin B_{12}) and 5-methyl tetrahydrofolate (a form of folate). 5-Methyl tetrahydrofolate, in turn, is produced by the enzyme 5,10-methylene tetrahydrofolate reductase (MTHFR). An alternative pathway for the regeneration of methionine from homocysteine is catalyzed by the enzyme betaine:homocysteine methyltransferase (BHMT). This enzyme is found mainly in liver and kidney and does not require B vitamins.

The rate-limiting reaction in the conversion of homocysteine to cysteine is catalyzed by the enzyme cystathionine β-synthase (CBS), which requires pyridoxal phosphate (a form of vitamin B_6). Thus, the level of homocysteine in a given tissue is determined by the amount of methionine in the diet, the bioavailability of B vitamins, and the relative activities of the enzymes MS, MTHFR, BHMT, and CBS. *See* AMINO ACIDS.

Hyperhomocysteinemia. Homocysteine contains a free thiol (SH) group, but only a small fraction (<2%) of the total homocysteine in blood is found in this (aminothiol) form. The remainder is a mixture of disulfide derivatives, including homocystine, homocysteine-cysteine mixed disulfide, and protein-bound disulfides. These various forms of homocysteine in blood are all derived from homocysteine that has been exported from the liver and other tissues. Sensitive and reliable methods for measurement of total homocysteine in blood have become widely available during the past decade. The normal concentration of total homocysteine in blood plasma is 5 to 15 μmol/L. Levels of total homocysteine tend to increase with age and are higher in men than in women. The term "hyperhomocysteinemia" refers to elevation of the blood level of total homocysteine above 15 μmol/L.

Severe hyperhomocysteinemia. Severe hyperhomocysteinemia, which is defined as a total homocysteine concentration greater than 100 μmol/L, occurs classically in patients with a homozygous defect of the gene that encodes CBS. Because these patients also have elevated levels of total homocysteine in their urine, they are said to have homocystinuria. Severe hyperhomocysteinemia also can be caused by hereditary defects in vitamin B_{12} metabolism or deficiency of vitamin B_{12} due to pernicious anemia.

Moderate hyperhomocysteinemia. Moderate hyperhomocysteinemia is defined as a blood level of total homocysteine between 15 and 100 μmol/L. This moderate degree of elevation of total homocysteine can be caused by genetic defects in CBS or MTHFR, kidney failure, or dietary deficiencies of folate or vitamin B_{12}. Moderate hyperhomocysteinemia also can be produced by drugs that interfere with the bioavailability of B vitamins, including certain medications used for the treatment of cancer, epilepsy, infections, asthma, and Parkinson's disease.

Hereditary homocystinuria. Hereditary homocystinuria (a cause of severe hyperhomocysteinemia) is quite rare, occurring in approximately one in 100,000 live births. The classic form of this disease, homozygous CBS deficiency, produces defective transsulfuration of homocysteine to cysteine, leading to accumulation of high blood levels of both homocysteine and methionine. When untreated, patients with homozygous CBS deficiency may develop skeletal deformities, dislocation of the lens of the eye, mental retardation, and premature death due to blood clots in the arteries or veins. Treatment with pharmacological doses of vitamin B_6, vitamin B_{12}, and folate, in conjunction with methionine restriction and supplemental betaine, lowers the plasma concentration of total homocysteine in most patients and markedly decreases the risk of vascular complications.

Hereditary homocystinuria also can be caused by genetic defects in the transport or metabolism of vitamin B_{12} or folate. Like homozygous CBS deficiency, these disorders cause markedly elevated levels of total homocysteine in the blood and urine, but levels of methionine are usually lower than normal. The clinical features of these disorders are very similar to those of homozygous CBS deficiency and include a high rate of cardiovascular events. It was the observation that vascular disease is a characteristic feature of homocystinuria caused by distinct metabolic defects that led the pathologist Kilmer McCully to postulate in 1969 that homocysteine may be a causative agent in atherosclerosis. *See* VITAMIN; VITAMIN B_6; VITAMIN B_{12}.

Birth defects and pregnancy complications. Moderate hyperhomocysteinemia is associated with an increased risk of birth defects resulting from incomplete closure of the neural tube during early embryogenesis. Although the overall prevalence is low (approximately one per 1000 live births), over 400,000 children are born each year with neural tube defects worldwide, and many of the cases are disabling.

The risk of a neural tube defect can be decreased by oral supplementation with folic acid during pregnancy, but only if the supplement is started before pregnancy or near the time of conception. To help prevent neural tube defects, enriched cereal grain products in the United States have been fortified with folic acid since 1998.

Moderate hyperhomocysteinemia also has been implicated as a possible risk factor for other complications of pregnancy, including intrauterine growth retardation, severe preeclampsia, birth defects other than neural tube defects, and premature separation of the placenta from the wall of the uterus. It is not yet known, however, whether treatment with folic acid or other homocysteine-lowering therapies will protect pregnant women from these complications. *See* CONGENITAL ANOMALIES.

Vascular disease. The hypothesis that homocysteine may cause cardiovascular disease was proposed over 30 years ago by Kilmer McCully, a pathologist who observed advanced vascular lesions in

children with hereditary homocystinuria. McCully's pioneering observations were confirmed in 1985 by an analysis of over 600 patients with severe hyperhomocysteinemia due to CBS deficiency. By the age of 30, approximately half of these patients had suffered from a cardiovascular event (stroke, myocardial infarction, or abnormal blood clots). It is now known that the risk of cardiovascular events in patients with hereditary homocystinuria can be markedly reduced by homocysteine-lowering therapy.

The mechanisms by which elevated levels of homocysteine predispose to vascular pathology are incompletely understood. Homocysteine may damage the endothelial cells that line the surface of blood vessels, interfering with their ability to regulate blood flow and prevent blood clots and inflammation. Homocysteine also may produce an oxidative stress that leads to atherosclerosis through the generation of oxidized lipoproteins.

Although severe hyperhomocysteinemia is a proven risk factor for adverse vascular events, the importance of moderate hyperhomocysteinemia as a cardiovascular risk factor is still undefined. Since the 1980s, a large number of epidemiological studies have suggested that moderate hyperhomocysteinemia may be a very prevalent risk factor, occurring in approximately 30% of patients with stroke, myocardial infarction, poor circulation in the arteries, or blood clots in the veins. In fact, some recent epidemiological studies suggest that blood levels of total homocysteine within the high-normal range (10–15 μmol/L) may confer an increased risk of cardiovascular disease.

An association between moderate hyperhomocysteinemia and future cardiovascular events also has been observed prospectively, although a few prospective studies have failed to demonstrate this association. It is still uncertain, therefore, whether moderate hyperhomocysteinemia is an independent risk factor for cardiovascular disease or simply a marker for (that is, associated with) another factor. Nevertheless, because total homocysteine in blood can be lowered by oral administration of folic acid or combinations of B vitamins, there is growing enthusiasm for treatment of moderate hyperhomocysteinemia as a strategy for prevention of cardiovascular disease and its complications.

Guidelines from the American Heart Association emphasize that it is not yet known whether reduction of blood levels of total homocysteine through increased intake of folic acid or other B vitamins will decrease cardiovascular risk in patients with moderate hyperhomocysteinemia. Several large clinical trials to test this hypothesis are ongoing. *See* VASCULAR DISORDERS.

Neurological disease. Regulation of homocysteine metabolism appears to be especially important in the central nervous system, presumably because of the critical role of methyl transfer reactions in the production of neurotransmitters and other methylated products. Abnormal accumulation of homocysteine in the brain may lead to increased levels of *S*-adenosylhomocysteine (see illustration), which can inhibit methyl transferase reactions. Another metabolite of homocysteine, homocysteic acid, may directly damage the brain by activating a specific receptor on neurons.

It has been known for decades that mental retardation is a feature of severe hyperhomocysteinemia due to hereditary homocystinuria. It also is well known that impaired cognitive function can result from pernicious anemia, which causes hyperhomocysteinemia due to deficiency of vitamin B_{12}. Hyperhomocysteinemia also may be linked to several other neurological disorders, including depression, schizophrenia, multiple sclerosis, and dementia (including both Alzheimer's disease and other types of dementia). *See* ALZHEIMER'S DISEASE; METABOLIC DISORDERS; NERVOUS SYSTEM DISORDERS.

Steven R. Lentz

Bibliography. R. Carmel and D. W. Jacobsen (eds.), *Homocysteine in Health and Disease*, Cambridge University Press, UK, 2001; R. Clarke, Homocysteine-lowering trials for prevention of heart disease and stroke, *Semin. Vasc. Med.*, 5:215–222, 2005; S. R. Lentz, Mechanisms of homocysteine-induced atherothrombosis, *J. Thromb. Haemost.*, 3:1646–1654, 2005; J. Selhub, Homocysteine metabolism, *Annu. Rev. Med.*, 19:217–246, 1999; S. Seshadri et al., Plasma homocysteine as a risk factor for dementia and Alzheimer's disease, *N. Engl. J. Med.*, 346:476–483, 2002; D. S. Wald, M. Law, and J. K. Morris, Homocysteine and cardiovascular disease: Evidence on causality from a meta-analysis, *BMJ*, 325:1202, 2002.

Homogeneous catalysis

A process in which a catalyst is in the same phase as the reactant. A homogeneous catalyst is molecularly dispersed (dissolved) in the reactants, which are most commonly in the liquid state. Catalysis of the transformation of organic molecules by acids or bases represents one of the most widespread types of homogeneous catalysis. In addition, the catalysis of organic reactions by metal complexes in solution has grown rapidly in both scientific and industrial importance. *See* CATALYSIS.

It is worth noting that sulfuric acid manufacture began in the eighteenth century using the lead chamber process in which oxides of nitrogen are used as homogeneous catalysts for the oxidation of sulfur dioxide. The basic chemistry involved is shown in reactions (1) and (2).

$$2NO + O_2 \longrightarrow 2NO_2 \qquad (1)$$

$$NO_2 + SO_2 + H_2O \longrightarrow H_2SO_4 + NO \qquad (2)$$

See SULFUR.

Acid-base catalysis. The two principal areas are specific acid (or base) catalysis and general acid (or base) catalysis. Specific acid catalysis refers to reactions in which only the oxonium ion (H_3O^+) can act as the catalyst. A common example is the

hydrolysis of simple acetals, reaction (3). Specific

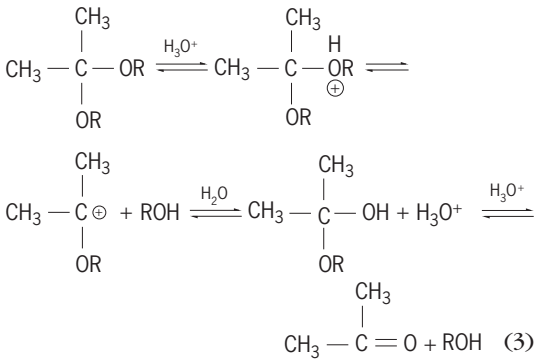

(3)

acid catalysis is found to be characteristic of reactions in which there is rapid, reversible protonation of the substrate before the slow, rate-limiting step.

Reactions that are catalyzed by proton donors in general are considered to be subject to general acid catalysis. General acid catalysis often becomes important only at higher acidity levels. The proton is a convenient and powerful agent for the distortion of the electronic configuration of a substrate in order to facilitate reaction. The mechanism by which this occurs has many variants. For example, a covalent bond may be more easily broken after protonation of one of the bonded atoms; the reaction, $ROH_2^{\oplus} \rightarrow R^{\oplus} + H_2O$ is easier than $ROH \rightarrow R^{\oplus} + OH^{\ominus}$.

Exactly the same distinction can be made in catalysis by bases as was made above for acids. Thus, in specific base catalysis the reaction rate is proportional to the concentration of $OH^{\ominus}$.

Metal complexes. In homogeneous catalysis by coordination compounds of transition metals, the catalyst is usually deployed in solution and most commonly exists in a molecularly dispersed form. Thus, all sites are potentially active for catalysis, and in many cases catalysis is observed under much milder reaction conditions than found with heterogeneous catalysis by metals and metal oxides.

The catalysis of the incorporation of carbon monoxide into organic substrates by transition metal complexes is technologically important. The hydroformylation or oxo reaction (4) in which an olefin

$$RHC{=}CH_2 + CO + H_2 \longrightarrow RHC_2{-}CH_2CHO + RCH(CHO){-}CH_3 \quad (4)$$

is reacted with carbon monoxide and hydrogen to generate a mixture of linear and branched aldehydes, was discovered in 1938. The first catalyst found was dicobalt octacarbonyl, $Co_2(CO)_8$, and this is still used extensively today in commercial operations. The steps involved in this reaction are summarized in reactions (5).

$$Co_2(CO)_8 + H_2 \rightleftharpoons 2HCo(CO)_4 \quad (5a)$$

$$HCo(CO)_4 + \text{olefin} \rightleftharpoons HCo(CO)_3(\text{olefin}) + CO \quad (5b)$$

$$HCo(CO)_3(\text{olefin}) \rightleftharpoons RCo(CO)_3 \quad (5c)$$

$$RCo(CO)_3 + CO \rightleftharpoons RCo(CO)_4 \quad (5d)$$

$$RCo(CO)_4 + CO \rightleftharpoons (R{-}CO)Co(CO)_4 \quad (5e)$$

$$(RCO)(CO)_4 + HCo(CO)_4 \longrightarrow RCHO + Co_2(CO)_8 \quad (5f)$$

Some of these steps represent transformations that are common to many sequences found in homogeneous catalysis. Thus, step (5*c*), the insertion of a coordinated olefin into the metal-hydride bond to generate a metal-alkyl bond, is a frequently encountered method of metal-carbon bond formation. Step (5*e*), the formation of a metal-acyl bond by alkyl migration to a coordinated carbon monoxide, is a key step in most catalytic (and stoichiometric) syntheses involving the incorporation of carbon monoxide into organic molecules.

While the reaction steps above can be conducted in a stoichiometric manner under very mild reaction conditions, in order for the system to function catalytically at rates which are desirable for industrial processes, the reaction temperature is maintained at greater than 248°F (120°C) and the reaction pressures are usually in excess of 200 atm (20 megapascals).

Other catalyst systems have been discovered which can perform hydroformylation reactions under much milder reaction conditions than cobalt. In particular, rhodium complexes containing triarylphosphine ligands can catalyze hydroformylation reactions at very rapid rates at ~212°F (~100°C) and 30 atm (3 MPa) pressure of synthesis gas ($CO + H_2$). Another important difference between the rhodium and cobalt catalysts is that the rhodium system can generate a much higher proportion of linear aldehyde product [reaction (4)]. This effect is related to the greater steric crowding around the metal when triarylphosphines are present in the coordination sphere. This is an example of the influence of stereochemistry around the metal on the stereochemical course of the catalytic reaction, and this phenomenon is an important feature of many homogeneously catalyzed reactions. *See* STERIC EFFECT (CHEMISTRY).

Another reaction involving the catalysis of the incorporation of carbon monoxide which has assumed considerable commercial importance is the synthesis of acetic acid from methanol, reaction (6).

$$CH_3OH + CO \longrightarrow CH_3CO_2H \quad (6)$$

The reaction is catalyzed by both cobalt and rhodium complexes in the presence of an iodide cocatalyst or promoter. The mechanism of the rhodium-catalyzed reaction is reasonably well understood, as shown in reactions (7). The reaction which

$$CH_3OH + HI \rightleftharpoons CH_3I + H_2O \quad (7a)$$

$$[Rh(CO)_2I_2]^- + CH_3I \rightleftharpoons [Rh(CO)_2(CH_3)I_3]^- \quad (7b)$$

$$[Rh(CO)_2(CH_3)I_3]^- \longrightarrow [Rh(CO)(COCH_3)I_3]^- \quad (7c)$$

$$[Rh(CO)(COCH_3)I_3]^- + CO \rightleftharpoons [Rh(CO)_2(COCH_3)I_3]^- \quad (7d)$$

$$[Rh(CO)_2(COCH_3)I_3] \longrightarrow [Rh(CO)_2I_2]^- + CH_3COI \quad (7e)$$

$$CH_3COI + H_2O \longrightarrow CH_3CO_2H + HI \qquad (7f)$$

generates the metal carbon bond, that is, step (7*b*), is rate-determining in the catalytic cycle.

The commercial reactors utilizing rhodium catalysts are operated at temperatures in the range of 300–390°F (150–200°C) and pressures of less than 40 atm (4 MPa). The rate of the reaction is sufficiently rapid that the amount of the very expensive rhodium catalyst required is very small.

A wide range of olefin transformation reactions are catalyzed by transition-metal complexes. Some of the more important reactions are isomerization, dimerization, polymerization, and metathesis. Olefin polymerization and oligomerization reactions involve a variety of homogeneously catalyzed processes and represent extremely large industrial applications of such systems (for example, polyethylene and polypropylene). Several different types of catalysts can be employed, including those based on free radicals, acids, carbanions, and transition metals. The properties of the polymer can be markedly influenced by the choice of catalyst. For example, free-radical polymerization of ethylene generates low-density polyethylene, whereas transition-metal catalysts give so-called high-density polyethylene. Examples of the transition-metal catalysts are the Ziegler-Natta systems produced by reacting titanium chloride ($TiCl_4$) with alkyl-aluminum compounds, and they are usually heterogeneous. Such transition-metal catalysts frequently introduce stereoregularity into the polymer, most likely because the reaction occurs at a crystal face. While detailed mechanistic information about these transition-metal catalyst systems is scarce, it is likely that the reaction occurs after formation of a metal-hydride either in solution or on a surface, followed by multiple olefin insertions to give polymer, as shown in reaction (8).

$$\text{Ti—H} \xrightarrow{C_2H_4} \text{Ti—CH}_2\text{—CH}_3 \xrightarrow{nC_2H_4} \text{Ti(CH}_2\text{CH}_2)_n\text{CH}_3 \qquad (8)$$

See POLYMERIZATION.

Oxidation. Transition-metal complexes act as homogeneous catalysts in many different types of oxidation process. Two main categories of reaction can be recognized, involving either one-electron or two-electron processes.

Autoxidation. The involvement of transition-metal complexes in one-electron, radical processes is most evident in the so-called autoxidation reactions whereby hydrocarbons are oxidized to various oxygen-containing compounds by radical chain processes. The general scheme is shown in reactions (9). While metal species can enhance the rate of several of the above steps, the most common pathway for catalysis of liquid-phase autoxidations involves the metal-catalyzed decomposition of alkyl hydroperoxides, of which reactions (10) and (11) are

$$\text{Initiation:} \quad \text{Initiator} + RH \longrightarrow inH + R\cdot \qquad (9a)$$

$$\text{Propagation:} \quad R\cdot + O_2 \longrightarrow RO_2\cdot \qquad (9b)$$

$$RO_2\cdot + RH \longrightarrow RO_2H + R\cdot \qquad (9c)$$

$$\text{Termination:} \quad R\cdot + RO_2\cdot \longrightarrow RO_2R \qquad (9d)$$

$$2RO_2\cdot \longrightarrow RO_4R \qquad (9e)$$

$$RO_4R \longrightarrow \text{nonradical products} + O_2 \qquad (9f)$$

$$RO_2H + Co^{II} \longrightarrow RO\cdot + Co^{III}(OH) \qquad (10)$$

$$(RO)_2Co^{III} \rightleftharpoons RO_2\cdot + Co^{II} \qquad (11)$$

examples. Cobalt and manganese salts are particularly effective in promoting autoxidation processes. The oxidation of *p*-xylene to terephthalic acid (the key monomer involved in the manufacture of polyester) is carried out on a very large scale using a cobalt-bromide catalyst.

Indirect oxidation. Transition-metal complexes find utility in the catalysis of various types of indirect, two-electron oxidations. Examples of these indirect processes are the so-called Wacker reaction, in which olefins are oxidized to aldehydes or ketones by palladium(II) compounds, with concomitant reduction of the palladium. The palladium is then reoxidized in a separate reaction by a combination of a copper salt and oxygen. The best-known example of the Wacker reaction is the oxidation of ethylene to actetaldehyde, reaction (12). The reaction is conducted in an

$$HC_2{=}CH_2 + {}^1/_2\,O_2 \longrightarrow CH_3CHO \qquad (12)$$

aqueous medium in the presence of palladium and copper chlorides as the catalyst system. The generally accepted mechanism is shown in reactions (13).

$$PdCl_4^{2-} + C_2H_4 \rightleftharpoons [PdCl_3(C_2H_4)]^- + Cl^- \qquad (13a)$$

$$[PdCl_3(C_2H_4)]^- + H_2O \rightleftharpoons [PdCl_2(H_2O)(C_2H_4)] + Cl^- \qquad (13b)$$

$$[PdCl_2(H_2O)(C_2H_4)] + H_2O \rightleftharpoons [PdCl_2(OH)(C_2H_4)]^- + H_3O^+ \qquad (13c)$$

$$[PdCl_2(OH)(C_2H_4)]^- + H_2O \rightleftharpoons [HOCH_2CH_2PdCl(H_2O)] + Cl^- \qquad (13d)$$

$$[HOCH_2CH_2PdCl(H_2O)] \longrightarrow CH_3CHO + Pd + HCl + H_2O \qquad (13e)$$

$$Pd + 2CuCl_2 \longrightarrow PdCl_2 + 2CuCl \qquad (13f)$$

$$CuCl + {}^1/_2O_2 + HCl \longrightarrow CuCl_2 + {}^1/_2H_2O \qquad (13g)$$

Another important indirect oxidation process is the metal-catalyzed epoxidation of olefins with alkyl hydroperoxides, reaction (14). Various molybdenum, vanadium, and chromium complexes act as catalysts for this reaction, by pathways that are still rather poorly understood.

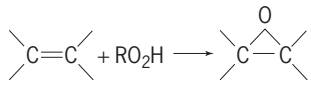

(14)

Adiponitrile, $NC(CH_2)_4CN$, is produced as a precursor of hexamethylenediamine, one of the building blocks of Nylon 66. The selective addition of two

moles of hydrogen cyanide to butadiene has been developed into a valuable new synthesis of adiponitrile. In one variant of this process, a zero-valent nickel catalyst, $Ni[P(OAryl)_3]_4$, can be used to bring about a series of reactions including HCN addition to butadiene, isomerization of cyanolefins, and hydrocyanation of 4-pentene-nitrile, reactions (15)–(17). A key step in reactions (15) and (17) appears to

$$HC_2{=}CH{-}CH{=}CH_2 + HCN \longrightarrow HC_3{-}CH{=}CH{-}CH_2{-}CN \quad (15)$$

$$HC_3{-}CH{=}CH{-}CH_2{-}CN \rightleftharpoons HC_2{=}CH{-}CH_2{-}CN \quad (16)$$

$$HC_2{=}CH{-}CH_2{-}HC_2CN + HCN \longrightarrow NC(CH_2)_4CN \quad (17)$$

be generation of a nickel hydride species capable of reacting with an olefin. An outline of the mechanism is shown in reactions (18)–(22), where L represents

$$NiL_4 + HCN \rightleftharpoons HNi(CN)L_3 + L \quad (18)$$

$$HNi(CN)L_3 + RCH{=}CH_2 \longrightarrow HNi(CN)(olefin)L_2 + L \quad (19)$$

$$HNi(CN)(olefin)L_2 \longrightarrow RCH_2CH_2Ni(CN)L_2 \quad (20)$$

$$RCH_2CH_2Ni(CN)L_2 \longrightarrow NiL_2 + RCH_2CH_2CN \quad (21)$$

$$NiL_2 + 2L \longrightarrow NiL_4 \quad (22)$$

phosphite or phosphine.

Perhaps the most elegant illustration of the selectivity achievable with homogeneous catalysts is found in the asymmetric hydrogenation of unsymmetrical olefins in the presence of rhodium complexes containing optically active phosphine ligands. Through this process, it is possible to prepare a number of optically active α-amino acids from the corresponding unsaturated precursors, reaction (23).

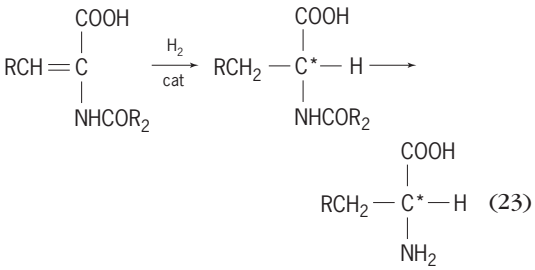

(23)

(The carbon marked with an asterisk is an asymmetric center.) The energy difference between the optical enantiomers is very small, but nevertheless, with suitable optically active phosphine ligands, one isomer can be produced with greater than 90% selectivity. This approach is used in the synthesis of L-dopa, the drug that is used in the treatment of Parkinson's disease. *See* ASYMMETRIC SYNTHESIS; OPTICAL ACTIVITY; STEREOSPECIFIC CATALYST.

Enzymes are naturally occurring catalysts that are responsible for the myriad biological transformations involved in the synthesis and breakdown of biomaterials. The enzymes, which are usually macromolecular proteins, frequently operate homogeneously. They use many types of homogeneous catalysis, such as acid, base, free-radical, and oxidation and reduction catalysis. *See* ENZYME; ORGANIC SYNTHESIS.

Denis Forster

Bibliography. I. N. Levine, *Physical Chemistry,* 5th ed., 2000; J. A. Moulijin, P. Van Leeuwen, and R. A. Van Santen (eds.), *Catalysis: An Integrated Approach to Homogeneous, Heterogeneous, and Industrial Catalysis*, 1993; G. W. Parshall and S. D. Ittel, *Homogeneous Catalysis: The Applications and Chemistry of Catalysis by Soluble Transitic Metal Complexes*, 1992.

Homoptera

An order of the class Insecta related to the order Hemiptera. This is a major group of sucking insects, with more than 30,000 species, even though in Asia and Africa the number of undiscovered species probably still exceeds the discovered ones. Common examples are the cicadas, aphids, and leafhoppers. The group is difficult to characterize because of the large number and diverse forms of the species it contains.

The head of these insects is hypognathous or opisthognathous, the beak appearing to arise from the ventral posterior margin of the head or even from the prosternum. The gula is membranous or absent. As in the Heteroptera, the beak consists of two pairs of stylets, formed by the maxillae and the mandibles, ensheathed in the labium. The maxillary stylets fit together to form a double tube, one channel serving for the passage of food and the other for saliva.

Most winged species have four wings, but male scale insects have only two. In most forms both pairs of wings are membranous and transparent, but in some, the forewings are somewhat thickened and may then be either coriaceous and translucent, or opaque, and with or without an apical membranous area. When the insects are at rest, the forewings are usually held, rooflike, over the dorsum, with the apex of one of them slightly overlapping the apex of its complement (**Fig. 1**).

The digestive tract in a vast majority of species is peculiarly complex in that it forms a filter chamber, a structure consisting essentially of a close association

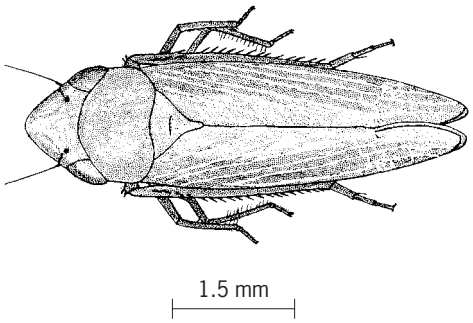

Fig. 1. A cicadellid in dorsal view.

of the posterior end of the midgut with the posterior end of the foregut. The approximated portions of the loop thus formed are believed to permit certain elements of the food to bypass most of the digestive portion of the gut. This anatomical feature, coupled with the prodigal feeding habits of most of the species, has led to the theory that certain factors in plant sap are present in such small quantities that large amounts of the sap must be imbibed, and the unneeded component shunted across the loop of the filter chamber, in order to obtain a sufficient quantity of the factors needed to sustain the insect. All species for which the food habits are known are phytophagous, deriving their nourishment from plant sap, except a few which are believed to be mycetophagous as immatures.

In most species metamorphosis is gradual, but in a few it is practically holometabolous. The adults and nymphs of most species are terrestrial, but a few species are subterranean in all stages and others are subterranean only in the immature stages. A number of species are vectors of virus diseases of plants.

Series Coleorrhyncha. This group is characterized by the origin of the beak, formed at the anteroventral extremity of the face, and by the fact that the propleura form a sheath for the base of the beak. The hindwings are absent, and the forewings are held flat over the abdomen in repose. The flight function has been lost. The prothorax is provided with dorsolateral expansions, the paranota, absent in other modern insects, but similar to structures found in Permocarboniferous fossil insects. There are additional anatomical features which have been cited as evidence of primitiveness. These serve as a basis for placing the Peloridiidae, the only known family, in the lowest position among the Homoptera. The species lack a filter chamber in the digestive tract, and are rare. They occur in Tasmania, New Zealand, and South America.

Series Auchenorrhyncha. This series and the Sternorrhyncha are the major groups of the Homoptera. In the Auchenorrhyncha the beak arises at the anteroventral extremity of the face and is not sheathed by the propleura. The labium arises well in front of the anterior legs. The antennae usually have one to three basal segments surmounted by a seta. In repose, the forewings are usually placed in a rooflike manner over the abdomen. They are more heavily sclerotized than the hindwings in some species, and may have a terminal membrane, thus somewhat resembling the forewings of the Heteroptera. Many species are good flyers. The hindwings nearly always have a jugum. The tarsi are nearly always three-segmented in the adults. Hindlegs are often adapted for jumping.

A filter chamber is apparently a constant feature of the digestive tract, although it may be greatly reduced in some species. There are usually four Malpighian tubules, originating in the anterior portion of the hindgut in the filter chamber. These frequently have an enlarged portion to which a glandular function has been attributed. Some species are cryptonephridial. Sound production by the males has been demonstrated to be of widespread occurrence in the Auchenorrhyncha and probably occurs throughout the group. A sound-producing apparatus has also been demonstrated in the females of some species.

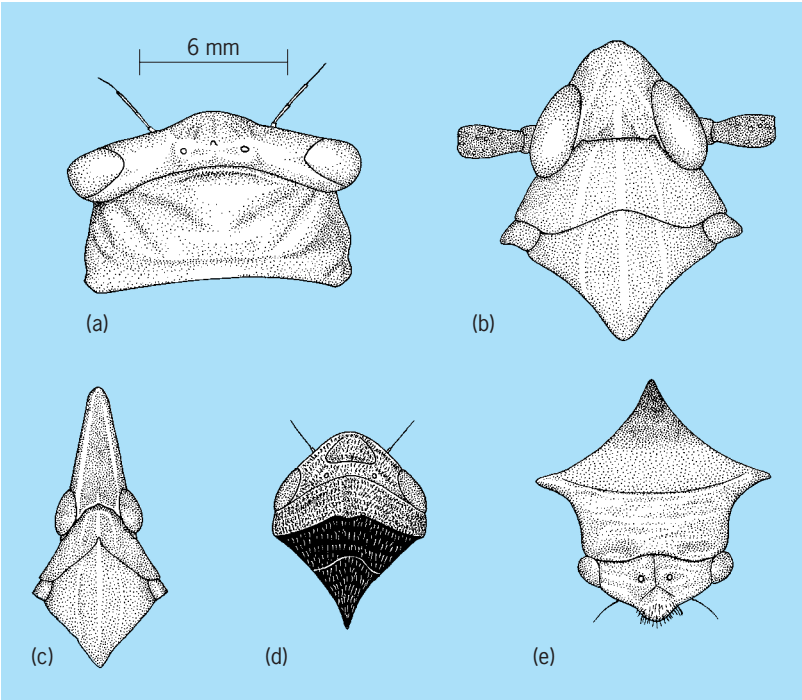

Fig. 2. Anterior dorsum of various homopterans: (*a*) Cicada, (*b*) fulgoroid (apical part of antenna omitted), (*c*) another species of fulgoroid, (*d*) cercopid. (*e*) A membracid, anterior view of face and pronotum.

The Auchenorrhyncha includes a large number of species. A number of classifications have been proposed, most of them differing chiefly in the rank assigned to the higher categories. The classification adopted here is a common one. It divides the series into the superfamily Fulgoroidea and the families Cicadidae, Cercopidae, Membracidae, and Cicadellidae. These families are not subordinate to the superfamily Fulgoroidea.

Superfamily Fulgoroidea. These are insects commonly known as lantern flies. They are distinguished from other Auchernorrhyncha by the following characteristics (**Fig. 2***b* and *c*). The middle coxae are the same length as the anterior coxae and are joined to the body at some distance from the median line. Tegulae, small, scalelike sclerites, are usually present at the base of the forewings. The hindwings usually lack a submarginal vein parallel to the wing margins.

The antennae, situated beneath the eyes, have two well-developed basal segments, of which the second is enlarged and often provided with numerous sensillae. Longitudinal carinae are often found on the face. There may be two or three ocelli, also located on the face. The forewings may be tectiform, vertical, or horizontal with their apices overlapping in repose. The anal, or claval, veins are usually fused to form a Y-shaped pattern. Contradictory statements have been made regarding the presence of a filter chamber in the digestive tract.

The popular name, lantern flies, resulted from old reports of bioluminescence in the head of a large South American species. This phenomenon has not been observed since, but the common name has survived.

This group is subdivided into 20 families and includes many species which are important because of the economic damage they do while feeding or because they carry virus diseases of plants.

Family Cicadidae. Included in this group are the cicadas, harvest flies, and jar flies. The insects in this family are probably better known to the layman than any other homopterous family because of their large size and the strident songs of the males. The following combination of characteristics will separate them from other families of Auchenorrhyncha. The short middle coxae differ from the anterior coxae and are joined to the body near the medial line. Tegulae are absent. The hindwings have a submarginal vein parallel to the wing margin. There are three ocelli arranged in a small triangle (Fig. 2*a*). The immature forms have digging legs.

The head of the adult is large, with protuberant eyes (Fig. 2*a*). The upper median portion (postclypeus) of the face is swollen, and this area is bounded throughout its length by the lora, or mandibular plates. The anterior and posterior tentorial arms are connected in the head. The forewings are membranous, tectiform in repose, having a single anal vein or two more or less parallel anal veins. The anterior legs have dilated femora which are spiny beneath. The posterior legs are not saltatorial. In the digestive tract, there is a long filter chamber in which the two parts of the gut are spirally interwound, with the hindgut issuing from its anterior end. There are four Malpighian tubules.

Adults are usually found on trees or shrubs. At least in some species, the songs of the males assemble local populations. After mating, the female lays eggs in the shoots of plants, making a slitlike incision for them by means of an ovipositor which is provided with sawlike blades or valves. These incisions frequently result in the death of the shoots. After hatching, the young cicadas fall to the ground and take up a subterranean existence which may last 1–17 years. During this period they feed on plant roots. Several species of North American periodical cicadas, with life cycles of 13 or 17 years, attract much attention because of occasional appearance in large numbers. These are often referred to as 13- or 17-year "locusts" in popular accounts, possibly as a result of association by early North Americans with the biblical plagues of locusts. There is no basis for such an association other than large numbers of individuals, and the use of the term locusts to refer to cicadas is inadvisable. *See* POPULATION ECOLOGY.

Family Cercopidae. Spittle bugs and froghoppers are common examples of this group. Insects in this group most often attract attention in the immature stages, during which they surround themselves with a mass of froth or spittle. The family has the following distinguishing characters. The middle coxae, the tegulae, and the hindwings are as described for the Cicadidae. There are usually two ocelli, never three, occasionally none. When present they are located on the crown of the head, of which the median apical portion is usually distinctly delimited from the remainder by sulci (Fig. 2*d*).

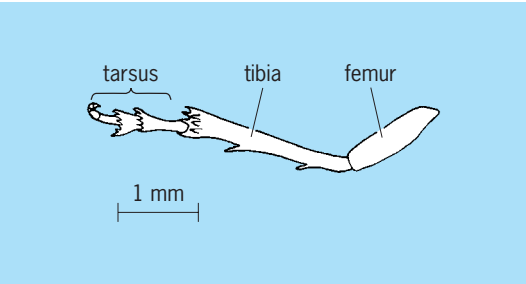

Fig. 3. Hindleg of a cercopid.

The head of the adult is proportionately smaller than in the Cicadidae, but the structure of the tentorium is similar. The pronotum, which does not extend posteriorly, is horizontal or sloped, but not vertical. The forewings, tectiform in repose, are usually somewhat sclerotized and different in texture from the membranous hindwings. The hind legs are saltatorial, with the coxae short and conical, not transversely dilated. The tibiae have one or a few stout spurs and a cluster of small outgrowths at the apex (**Fig. 3**). In the digestive tract, there is a complicated filter chamber in which a number of convolutions are found. The hindgut issues from its posterior end. There are four cryptonephridial Malpighian tubules arising from two basal stalks.

The froth, or spittle, which covers the immature insects consists of extruded anal fluid with which air bubbles are mixed. The frothy mass does not evaporate readily, a fact which seems correlated with the reported inability of the young cercopids to survive in a dry atmosphere. There are many more species in the tropics than in temperate climates. One species, *Philaenus spumarius*, the meadow spittle bug, is very common in the temperate portion of the Northern Hemisphere, and its masses of spittle are familiar sights. Some species of Australia and the East Indies live within a calcium carbonate tube attached to stems or leaves.

Family Membracidae. The treehoppers are small to medium in size and seldom attract attention. Most of them feed on woody plants and are found on the stems in sunny locations. Frequently a number of specimens are arranged in a vertical row on the stem, all with their heads downward. Membracidae may be recognized by the following combination of characters. The middle coxae and the tegulae are as described above for the Cicadidae. There are two ocelli, or none. The pronotum extends backward over the abdomen, sometimes almost completely covering the wings (**Fig. 4***a*). The upper portion of the head is vertical (Fig. 2*e*).

In the adult head, the anterior and posterior arms of the tentorium are not connected. The antennae consist of two larger basal segments and a

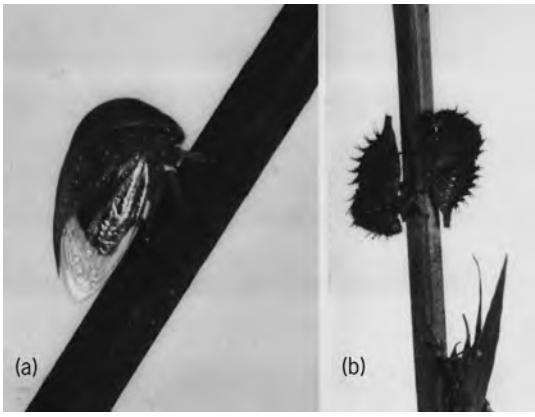

Fig. 4. Membracids on stems. (*a*) Adult. (*b*) Nymphs. (*Courtesy of C. H. Hanson*)

terminal seta with a large number of subsegments. The forewings are usually membranous but may be partly sclerotized, and are tectiform in repose. The radial and medial veins are not fused basally. There is often a submarginal vein in the hindwings. The anterior and middle tibiae are often dilated. The hind legs are saltatorial and spiny. The digestive tract has a filter chamber with fewer coils than found in many Homoptera. There may be two or four Malpighian tubules which are cryptonephridial, at least in some species.

The greatest number of species occurs in the warmer regions of the world. In many species the enlarged pronotum has adornments, excrescences, and processes which are astonishing in appearance, some of them nearly as large as the remainder of the insect. Most species are not active, leaping only to take flight, seldom flying, and then only for short distances. In North America, *Stictocephala bubalus*, the buffalo treehopper, causes economic damage to fruit and other trees by slitting the twigs during oviposition. The nymphs (Fig. 4*b*) leave the trees and feed on herbaceous plants, often occurring in great numbers in pastures. The adults return to woody plants before oviposition.

Family Cicadellidae. The leafhoppers are included in this large family. These usually small insects are known to many people by sight but not by name, because of their common occurrence in great numbers at night near lights. The species may be distinguished from other families by the following combination of characters. The coxae and tegulae are as described for Cicadidae. The pronotum does not extend backward over the abdomen and does not have a median ridge. The upper portion of the head is never vertical. There is usually a submarginal vein parallel to the wing margin in the hindwings (**Fig. 5**). There are two ocelli, or none. When present, they may occur on the face, on the crown of the head, or on the margin between the face and crown. The hind tibiae have many spines arranged in rows (Fig. 1).

On the face, the lora border the postclypeus for only a short distance. The anterior and posterior arms of the tentorium are not connected in the head. The forewings may be membranous or heavily sclerotized and, in the latter case, often have a membranous apical portion. They are usually tectiform in repose. Radial and medial veins arise from a common stalk. The hind legs are saltatorial. The digestive tract has a filter chamber, but reports indicate a considerable degree of variation in its detailed structure. There are four cryptonephridial Malpighian tubules.

Probably the greatest number of species occurs in tropical areas, but the majority of these have not been described. Temperate North America also has a large number of species. Leafhoppers occasionally bite man, but apparently they have never been seen taking blood. Several species have been found to be vectors of virus diseases of plants. Some of the more important of these virus diseases are phloem necrosis of elms, curly top virus of sugar beets, aster yellows virus, rice dwarf virus, and papaya bunchy top virus, in addition to virus diseases of potatoes, clover, alfalfa, grapes, peaches, sugarcane, eggplant, corn, maize, wheat, and cranberries. In a few cases, the virus is transmitted transovarially. *See* PLANT PATHOLOGY.

Series Sternorrhyncha. In this group of homopterous families, the beak appears to arise either between the fore coxae or behind them. The antennae are usually long, filamentous, and have no well-differentiated terminal setae. Wingless forms are common. The wings, when present, are usually membranous, with reduced venation, and usually without closed cells. The hindwings have no jugum. The tarsi of the adults are 1- or 2-segmented. A filter chamber is usually present in the digestive tract.

The Sternorrhyncha includes a large number of species, many of them of great economic importance. The winged forms are not strong flyers, but they are so light that they may be borne considerable distances by air currents. There have been several classifications proposed for the Sternorrhyncha. The following subdivision into superfamilies and families appears to be the most widely used classification.

Family Psyllidae. The Psyllidae are known as jumping plant lice. This family has also been known as the Chermidae. Its representatives resemble cicadas in appearance, but are much smaller. About 1000 species are known. Adult insects have a

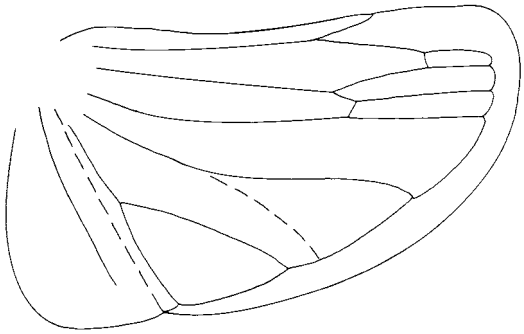

Fig. 5. Cicadellid wing. The more basal broken line represents the jugal fold.

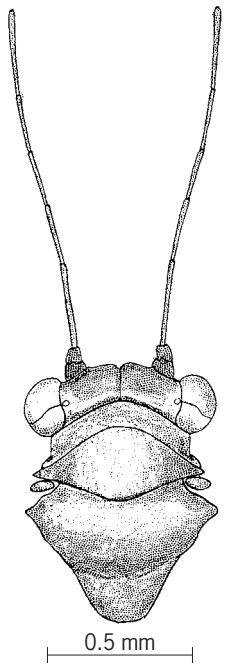

Fig. 6. Anterior dorsum of a psyllid.

transverse head, usually emarginate anteriorly, with protuberant eyes (**Fig. 6**). The face is often in the form of two conspicuous cones which are directed ventrally. Three ocelli occur. The antennae are 6- to 10-segmented, with the latter condition being common. They consist of two wider basal joints and a terminal portion which bears sensillae. The apical antennal segment bears two fine setae. The winged forms have four wings which are usually membranous, but the forewings are thicker than the hindwings. The wings are tectiform in repose. The venation is reduced, but conspicuous, and cross veins are seldom present. In the forewings, the radius, media, and cubitus arise from a common stalk. The tarsi are 2-segmented with a pair of apical claws. The hind legs of the adult are saltatorial, with the coxae greatly dilated.

The hindgut and the posterior portion of the esophagus are looped around each other. Possibly this is a functional filter chamber, although a specialized clearly delimited filter chamber, similar to that found in some other groups of Homoptera, is apparently absent. There are four Malpighian tubules opening separately into the midgut.

The species of psyllids may be monophagous or polyphagous. There may be an alternation of host plants, but the nymphs and adults occur on the same plant. Hibernation may occur in the egg, nymphal, or adult stages. The nymphs have flattened bodies and do not have saltatorial hind legs. They are frequently covered by a secretion which may be wooly, waxy, or viscous.

Some psyllids produce severe damage to their food plants. This damage may result from the mere feeding by tremendous numbers of individuals, from the resulting yellowing or rolling of leaves, or from galls produced on the leaves. Indirect damage may result from the growth of fungi on leaves which have become coated with the sugary excrement, the honeydew, of the psyllids. *Psylla pyricola*, the pear psylla, is a species of considerable economic importance in North America. Its damage includes premature falling of leaves and fruit, reduced quality of fruit, and blackening of the leaves from fungus.

Family Aleyrodidae. The whiteflies are 0.3 in. (7 mm) or less in length and usually lightly covered with a white, powdery, waxy material which has led to their common name. The head varies in shape, and there are two ocelli. The antennae are usually 7-segmented, with the two basal segments thicker than the other segments, which may have sensoria and may terminate in a bristle. The eyes may be reniform, or divided, in which case they appear as four eyes. Adults have four wings which are opaque, being covered with a whitish powder. The wings are tectiform in repose. The venation is greatly reduced and there are no cross veins. The tarsi are 2-segmented, with a pair of terminal claws between which is a median structure which varies in form from a blade to a bristle. The thorax is separated from the abdomen by a constriction. There is a filter chamber in the gut, but the first ventriculus of the gut is not greatly dilated where the filter chamber occurs. There are only two Malpighian tubules. The stylets of the mouthparts, when not in use, are held in the crumena, a pouch within the thorax.

Several species of Aleyrodidae are polyphagous. The eggs are attached to the leaves of plants. The newly hatched insects are ambulatory, but they soon choose a feeding site and remain virtually sessile until reaching the adult stage. The sessile forms, with their appendages greatly reduced, resemble scale insects, and the last larval instar is anchored to the plant by a waxy secretion. Both bisexual reproduction and parthenogenesis occur within the group.

Whiteflies directly damage plants by their feeding. Indirectly, damage results from spotting at the feeding site, growth of fungus on the excreted honeydew, or from increased susceptibility of leaves to winter damage. Two species are important pests in the United States. *Dialeurodes citri*, the citrus whitefly, an Asian species, is a pest of citrus and other plants in California, Florida, and elsewhere. *Trialeurodes vaporarium*, the greenhouse whitefly, is a pest in greenhouses and on several cultivated crops. Other species carry virus diseases to cassava, cotton, cucumber, sunflower, and tobacco.

Superfamily Aphidoidea. Members of this large superfamily of small insects have four wings, or none. The wings, usually membranous or whitish and opaque, are held tectiform over the abdomen in repose. The forewings are much larger than the hindwings. The tarsi are 2-segmented and usually have two claws. The hind legs ordinarily are not saltatorial. The intersegmental lines are distinct in the abdomen. Both parthenogenetic females and sexual females occur. The life history may be extremely complicated, involving several host plants as well as several forms of

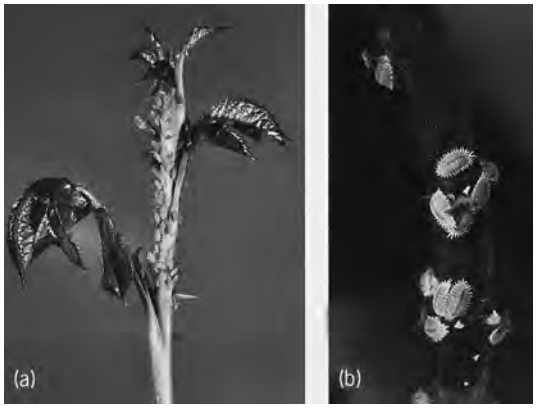

Fig. 7. Plant pests. (*a*) A colony of aphids on a rose shoot. (*b*) The citrus mealybug (*Ohio Agricultural Experiment Station*).

the insects. Honeydew is usually produced. This superfamily includes the families Aphididae and Chermidae.

Family Aphididae. The true aphids (**Fig. 7***a*) are members of this family in which the sexual females are oviparous and the parthenogenetic forms are viviparous. The sexual females, and usually the males, have a functional beak and a continuous digestive tract. Cornicles are usually present.

Compound eyes are present in the adults and three ocelli occur in the winged forms. Usually the antennae are 5- or 6-segmented with the apical segment ending in a tapering terminal process. Sensoria are present on some antennal segments. The beak has five segments. The wings extend caudad beyond the apex of the abdomen. In the forewing, veins Rs (radial sector) separates from the stigma and reaches the wing apex. The last abdominal segment is extended to form a cauda, with the anus opening beneath it. A poorly developed filter chamber is said to occur in the gut of some aphids, but a number of aphids lack it. Malpighian tubules are absent throughout the family (**Fig. 8**).

Many species of aphids have both an alternation of generations and an alternation of hosts, one of which is usually obligatory. In temperate climates, the eggs are laid most commonly on a woody host in autumn where they overwinter. The females which hatch from the eggs produce, by viviparous parthenogenesis, a variable number of generations on the primary host. Winged forms are eventually produced which migrate to a herbaceous alternate host. Here a number of generations are produced by viviparous parthenogenesis, some of the individuals having wings, but most of them being wingless. In autumn, a generation of winged females occurs which migrate to the primary host. Here they produce parthenogenetically the winged sexual males and females. After mating, the females lay eggs. There are many departures from this generalized life history. Some species are monophagous. In others, alternation of hosts appear optional, the aphid species being able to reproduce on the primary host, secondary host, or both.

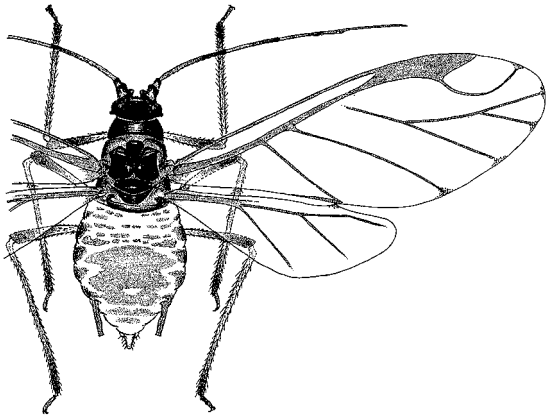

Fig. 8. An aphid, dorsal aspect.

Many species of aphids that produce honeydew are very attractive to ants, and colonies of very small aphid species which otherwise might escape notice can often be located by observing the attending ants. In a few species, this relationship has progressed to the point where ants are necessary for the survival of the aphid species, as in the corn root aphid, *Anuraphis maidiradicis*. The eggs of this species are cared for by the corn field ant, and the young aphids are transferred to corn roots early in the growing season. *See* SOCIAL INSECTS.

Many aphid species are important because of damage done in feeding. The pea aphid, *Acyrtosiphum pisum*, is a pest to peas and other crops in the United States. The cotton aphid, *Aphis gossypii*, is a pest of cotton, melons, and other crops in the same region. Other species are important because of their ability to transmit virus diseases to such plants as potatoes, sandalwood, squash, beets, cauliflower, celery, clover, cucumber, groundnut, alfalfa, onion, pea, sugarcane, tobacco, turnip, and citrus.

Family Chermidae. This is a small family of minute insects, the adelgids and phylloxeriids, in which both the sexual and parthenogenetic females are oviparous. Cornicles are absent. The antennae are 3- to 5-segmented and bear sensilae. Both winged and wingless forms occur. The following are subfamilies of Chermidae.

In the subfamily Cherminae all forms have a beak, and the digestive tract is not closed; the wings are tectiform in repose, and there is no branched vein extending toward the posterior margin of the forewing. There is frequently a waxy, flocculent secretion. The life cycle is extremely complicated and may involve an alternation of plant species and an alternation of generations in the insects, with several morphological forms occurring in the life history of one species. The primary host is always spruce. Secondary hosts are other conifers.

In the subfamily Phylloxerinae, the sexual forms lack mouthparts. The parthenogenetic females have a beak but the digestive system is closed (not continuous), and there is no honeydew. In repose, the membranous wings lie flat over the abdomen. There is a branched vein extending toward the posterior margin of the forewing. Waxy secretions, rarely present,

are not flocculent. Although winged migrants occur commonly, there is no secondary host.

The grape phylloxera, *Phylloxera vitifoliae*, has been a severe pest of cultivated grapes and once threatened the entire wine industry of France. It was discovered that native North American grape roots were not damaged greatly by the phylloxera, and consequently these were used as grafting stock in Europe to reduce damage.

Superfamily Coccoidea. The scale insects and mealy bugs, which are members of this large and important superfamily, are usually small. More than 4000 species have been described. In the males, the hindwings are reduced to clublike halteres. The wings are usually held flat over the back in repose, and the venation is greatly reduced. The females are wingless. When legs are present the tarsi are usually 1-segmented, but in some males the tarsi are 2-segmented. There is a single tarsal claw. The hind legs are not saltatorial. In some species, the abdominal segmentation is much modified or obliterated.

The males are usually very small, even in species where the female is much larger. They have no beak, do not eat, and have a nonfunctional digestive tract. Young males resemble corresponding female stages, but they molt more often, eventually passing through a stage so much like the pupal stage of holometabolous insects that they have presented an obstacle to systematists of higher categories of insects who used the type of metamorphosis as a fundamental criterion in classification. Adult males normally have long antennae consisting of 10 or fewer segments, although written accounts have stated, as a result of erroneous observations, that as many as 25 segments were present. The eyes of the adult male are compound in some species, but they usually consist of a series of isolated facets. The legs are well developed. The caudal end of the abdomen often bears an elongate process. Males are unknown in some species, and apparently of rare occurrence in some others. A few species have wingless males.

Females are usually sedentary, with the legs absent or reduced and nonfunctional. The body may be soft, or gall-like, or covered with powdery or tufted wax, or with scales or other hardened secretions. Some species form cysts which resemble small pearls. Others inhabit plant galls. In some groups, the females superficially resemble aphids, but cornicles are never present. The females usually have a 1- to 3-jointed beak which arises behind the bases of the anterior legs and varies considerably in shape and size. Its stylets are coiled in a ventrally located internal crumena when not in use. Antennae are present in adult females and may be reduced to flat disks, or may be elongate, with as many as 11 joints. The eyes are very simple and somewhat resemble ocelli. There are usually a number of glands, pores, and ducts in the integument and these features are useful in taxonomy. There is no pupalike stage, as occurs in males. The gut may be either continuous, and with a complex filter chamber, or discontinuous. There are two, three, or four Malpighian tubules.

In the first nymphal stadium, both sexes are mobile and known as crawlers. At this period in their development, they spread over the plant and to other plants.

Scale insects injure the host plant by their feeding on leaves, stems, or roots, and a number of species are very important economically. Most species produce large quantities of honeydew. A few species have been shown to be vectors of virus diseases of cucumbers, tobacco, and cacao. The sedentary habits and small size of the females have contributed to their wide distribution on economically important plants. Monophagous and polyphagous species occur.

Although specialists in scale insects generally agree that the included groups of insects form a superfamily, there seems to be little agreement on the constitution of the several included families. For this reason the taxa discussed here can be treated most conveniently, at present, as subfamilies. As many as 20 taxa in the family group have been accepted by some authors.

In the Margarodinae, abdominal spiracles are present in all stages of development. The adult males have compound eyes, and usually 10-jointed antennae. Adult females have at least some segmentation of the legs, and usually two bristles on the tarsal claw. The antennae of the adult female are not contiguous basally, and the anal tube is apical if well developed. No stages have a flat anal ring which bears pores and setae. Some *Margarodes* and related species live underground and have become lawn pests in some parts of the southern United States, usually in areas where the soil is sandy. The females form glassy cysts in which they are able to live quiescent for a considerable time without food. The cysts, or "ground pearls," are lustrous and are used as jewelry by natives in some parts of the world.

In the Ortheziinae, abdominal spiracles are present in all stages. Immature forms of both sexes and all females have a flat anal ring which bears pores and setae. Adult males have 9-jointed antennae with an apical seta on the terminal segment. Adult males have a strongly bivalved penis sheath. The body of the female is covered with hard white waxy plates. *Orthezia insignis* has a widespread distribution in greenhouses.

The Monophlebinae posses the characters listed above for the Margarodinae but the anus is dorsal. This subfamily includes the cottony-cushion scale, *Icerya purchase*, a species introduced to California from Australia and which at one time threatened the citrus industry in California. It was brought under control by an introduced Australian coccinellid, *Rodolia cardinalis*, in one of the earliest known instances of biological control.

In the Aspidiotinae, the adult male does not have compound eyes. Adult females have greatly reduced antennae and are without legs. There are no abdominal spiracles in any of the stages. There are no setae at the anal opening. In the females and nymphs, the apical abdominal segments are fused to form a

compound pygidium. The scales are shieldlike and nearly circular. The San Jose scale, *Aspidiotus perniciosus*, is one of the most important members of the subfamily and is thought to be a native of China. It received its common name from the California locality where it was first found in the United States. It attacks many kinds of fruit trees and shrubs, as well as woody ornamentals. It occurs throughout the United States and Canada and in other parts of the world in temperate regions.

The species of Lepidosaphinae have dark-colored scales and features similar to those of the Aspidiotinae, except that the scale is not circular. Three species of *Lepidosaphes* are of economic importance, two on citrus, and one, the oystershell scale, a cosmopolitan species which infests many kinds of woody plants.

In the Lacciferinae, the adult male does not have compound eyes. The abdomen is without spiracles in all stages, and in the nymphs and females its apical segments are not coalesced to form a pygidium. The abdominal apex has a tubular projection. The body is not covered by a scale, but the insects are enclosed in a mass of resin. The Oriental species, *Laccifer lacca*, or lac insect, found on a number of host plants, as well as related species, secretes the resinous material from which shellac is produced. *See* SHELLAC.

The Eriococcinae have many characteristics like those of the Lacciferinae, but the abdominal apex is without the long tubular projection. The adult females and late instar nymphs have an anal ring. The beneficial cochineal insects, *Dactylopius coccus*, belong to this group. The harmful mealybugs, species of *Pseudococcus* (Fig. 7*b*), and other genera also belong here. These soft-bodied insects are covered with flocculent wax secretion which suggested their common name. A number of species are harmful to cultivated plants, and some are serious pests of greenhouse and house plants. *See* ENTOMOLOGY, ECONOMIC; HEMIPTERA; INSECT PHYSIOLOGY; INSECTA. David A. Young

Homosclerophorida

An order of primitive sponges of the class Demospongiae, subclass Tetractinomorpha, with a skeleton consisting of equirayed, tetraxonid, siliceous spicules and their derivatives formed through reduction in number of rays (see **illus.**). In some species the ends of the rays of the tetraxons branch many times to form spicules known as candelabra. Sponges of the genus *Oscarella*, considered to be primitive, lack spicules but are reinforced by a mesenchymal collagen bearing elastin fibers. The genus *Corticium* and related forms have a cartilagelike dermal region.

Homosclerophorid sponges are mostly small in size and encrusting to massive in shape. They occur in tidal and shallow waters, down to depths of at least 1640 ft (500 m). Fossil sponges with spicules suggesting homosclerophorid affinities are scattered through the fossil record from Carboniferous strata upward. *See* DEMOSPONGIAE. Willard D. Hartman

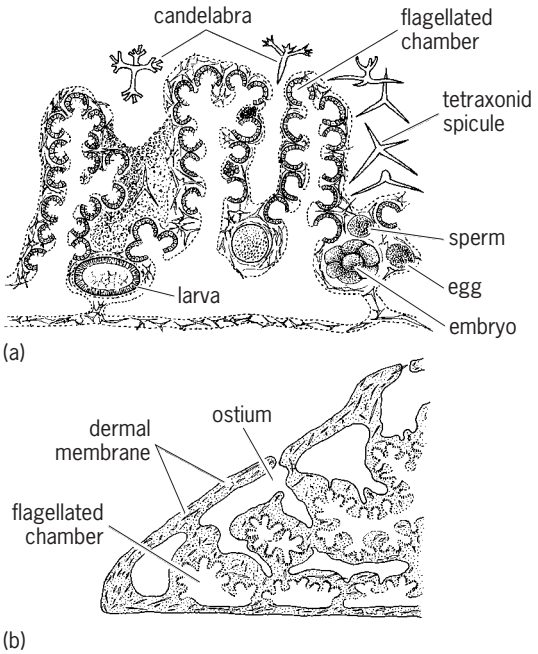

Homosclerophorida morphology. (*a*) Section through *Plakina*. (*b*) Section through *Plakortis* with cortex.

Honey Dew melon

A long-keeping cultivar of muskmelon, *Cucumis melo inodorus*, of the gourd family, Cucurbitaceae (see **illus.**). *See* VIOLALES.

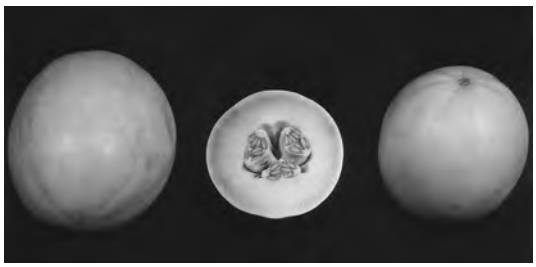

Honey Dew melon, *Cucumis melo inodorus*.

Description. Vines are vigorous and prolific and have large leaves and stems. The vines usually bear andromonoecious flowers which are pollinated by bees. The fruits are large, 5–7 lb (2.3–3.2 kg), slightly oval (length 7–8 in. or 18–20 cm, diameter 6–7 in. or 15–18 cm), smooth, creamy yellow to ivory when ripe, and with little or no net. They usually remain attached to the stem at harvest; some later introductions develop an abscission layer (slips) at harvest maturity. The flesh is thick, light green, tender, juicy, and very sweet with mild aroma and flavor; it contains 10% or more sugar when ripe and is rich in potassium and vitamin C, but not as rich in vitamin A as the orange-fleshed cantaloupe.

Quality. Honey Dew melons can be stored for 2–3 weeks at temperatures of 45–50°F (7–10°C). Melons do not ripen satisfactorily at temperatures below 50°F (10°C). Commercially, Honey Dew melons after harvest are often treated with ethylene gas (500–1000 parts per million) to induce and hasten ripening. As the fruits mature, the skin changes in color from greenish to creamy white, skin texture changes from a fuzzy feel to a smoothness, and the flesh increases in sugar, decreases in acidity, and releases the characteristic aroma. Sugars do not increase after the fruit is removed from the vine. Flavor and texture improve for a few days after harvesting and attain their highest quality in fruits harvested at their maximum sugar content. In California, state laws require that the edible portion contain at least 10% soluble solids before Honey Dews can be marketed.

Cultivation. Honey Dew is of African origin, has been grown in France for many years as the cultivar White Antibe, and was introduced into the United States in 1911. It was named Honey Dew in 1915. It requires a long frost-free season of 110–125 days and grows well at average temperatures of 70°F (21°C) or higher. Most of the crop is harvested in the summer in California; almost all of the remaining crop is harvested in the spring in Texas, with less than 5% harvested in the fall in Arizona. The chief growing areas are the Sacramento and San Joaquin valleys of California. In these areas high sunlight, low humidity, and the absence of rain tend to prevent fungus diseases that often defoliate the plants in humid areas. Honey Dew is propagated by seeds planted directly in the soil or by transplants. Plants are spaced about 12 in. (30 cm) in rows 6–8 ft (2–2.5 m) apart.

Diseases. Diseases include powdery mildew, which occurs under high humidity and is controlled by sulfur dusts, and Fusarium and Verticillium wilts, which inhabit the soil and cause wilting of the foliage and early death. Virus diseases, including watermelon mosaic, cucumber mosaic, and squash mosaic, are transmitted by insects. Major insect pests include wireworms, cutworms, leaf miners, cucumber beetles, aphids, and spider mites. Nematodes attack the roots of plants grown in infested soils. *See* MUSKMELON; PLANT PATHOLOGY. Oscar A. Lorenz

Bibliography. California Melon Research Board, *Informational Bulletin*, quarterly; N. B. Childers, *Modern Fruit Science*, 9th ed., 1995; R. A. Seelig, *Honeydews*, Fruit and Vegetable Facts and Pointers series, 1973.

Hookeriales

An order of the true mosses (subclass Bryidae), found mostly in the tropics (see **illus.**). The Hookeriales are no doubt heterogeneous, but most have well-developed double costae. The plants are usually shiny and often yellowish. The stems are creeping to spreading or ascending, and sparsely to freely and irregularly branched, or frondose. The leaves are often asymmetric and complanate, and commonly bordered and toothed. The cells are often large and lax, and the alar cells are not differentiated. The sporophytes are lateral, with setae short to elongate and often rough, even bristly. The capsules may be erect to inclined, but are essentially symmetric. The double peristome consists of 16 teeth which are long and slender, gradually tapered, papillose throughout or cross-striolate below, bordered, and often furrowed, with an endostome consisting of 16 keeled segments from a basal membrane. The calyptrae may be cucullate or mitrate and often ciliate-fringed.

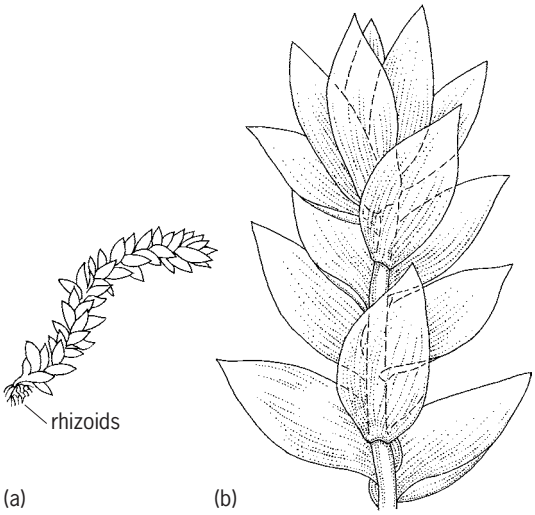

***Hookeria acutifolia*. (*a*) A portion of an entire plant. (*b*) Apical portion enlarged. (*After W. H. Welch, Mosses of Indiana, Indiana Department of Conservation, 1957*)**

As presently constituted, the order consists of two or three families and about 40 genera. The strong double costa, well-developed double peristome, with teeth slender and often furrowed, and mitrate calyptrae, which are often somewhat hairy and sometimes fringed at the base, allow some clustering of genera around a natural center. However, the relationships of the Daltoniaceae and Ephemeropsidaceae (with papillose exostomes) to the Hookeriaceae, (with cross-striate exostomes) are unclear. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA. Howard Crum

Hooke's law

A generalization applicable to all solid materials, stating that stress is directly proportional to strain and expressed as

$$\frac{\text{Stress}}{\text{Strain}} = \frac{S}{\epsilon} = \text{constant} = E$$

where E is the modulus of elasticity, or Young's modulus, in pounds per square inch. The constant relationship between stress and strain applies only to stress below the proportional limit. For materials having a nonlinear stress-strain diagram, the law is an approximation applicable to low stress values. *See* STRESS AND STRAIN; YOUNG'S MODULUS. W. J. Krefeld; W. G. Bowman

Bibliography. R. G. Budynas, *Advanced Strength and Applied Stress Analysis*, 2d ed., 1998; H. E.

Davis, G. E. Troxell, and F. W. Hauck, *The Testing of Engineering Materials*, 4th ed., 1982.

Hop

A plant (*Humulus lupulus*) belonging to the family Cannabaceae, also including nettles and hemp. The plant is a rough-stemmed, tall-twining dioecious perennial herb. It is propagated by cuttings of underground stems. The male flowers are borne on loose stalks. The anthers burst to produce large quantities of pollen, which is windborne. The female flowers occur in a cluster consisting of a compressed central axis that holds many flowers. As the clusters mature, the central axis elongates and papery leaves replace each flower, growing into a conelike structure, which is the hop of commerce (see **illus.**).

Mature hop cones.

The commercial value of the hop is in the golden-yellow lupulin glands on the bracteoles. The glands contain the resins, which impart bitterness, and essential oils, which impart hop flavors, to beer and ale. Dried hops are also used as soporifics. *See* MALT BEVERAGE.

Hop is grown primarily in the temperate areas of the world. In the United States, commercial production is in Washington, Oregon, and Idaho. Hop is planted on a 6.5–8-ft (2–2.5-m) spacing. The vine is trained by hand to strings fastened to a trellis 16.5–18 ft (5–5.5 m) in height. The female inflorescences are picked mechanically and dried in forced-air, direct-fired kilns at temperatures of 115–149°F (46–65°C). Hops are marketed in 190–200-lb (85–90-kg) bales. Other major hop-producing countries include Germany, the Czech Republic, Ukraine, China, United Kingdom, and Slovenia.

The principal pests are the two-spotted spider mite (*Tetranychus urticae*) and the hop aphid (*Phorodon humuli*). Major fungal diseases are downy mildew (caused by *Pseudoperonospora humuli*), powdery mildew (caused by *Sphaerotheca humuli*), Verticillium wilt (caused by *Verticillium albo-atrum*), and Phytophthora crown rot (caused by *Phytophthora citricola*). The fungal diseases are controlled by either using fungicides or growing resistant varieties. The severity of the fungal diseases varies between growing regions because of differences in environment. *See* FUNGISTAT AND FUNGICIDE.

Viruses causing diseases in hop include the *Prunus* necrotic ringspot virus and apple mosaic virus (both ilarviruses); and hop latent virus, hop mosaic virus, and American hop latent virus (all carlaviruses). Viruses are controlled by planting virus-free rootstock or resistant varieties. Two viroids that infect hop are hop stunt viroid and hop latent viroid. *See* PLANT PATHOLOGY; PLANT VIRUSES AND VIROIDS.

C. B. Skotland; Stephen T. Kenny

Bibliography. R. A. Neve, *Hops*, 1991; V. Rybacek (ed.), *Hop Production*, 1991.

Hophornbeam

The genus *Ostrya* of the birch family, represented in North America by two species. *Ostrya virginiana*, a small tree which may reach a height of 60 ft (18 m), is widely distributed in the eastern half of the United States and in the highlands of southern Mexico and Guatemala. It can be recognized by its fruit, which closely resembles that of the hop vine, and by its very scaly bark. The scales usually occur in narrow, more or less parallel, vertical strips. The winter buds are usually tinged with green showing about six striate scales, and the leaves are sharply and doubly serrate (see **illus.**). This is one of several trees known as ironwood because of its hard, strong wood, and like the hornbeam, it is used for fence posts, tool handles, mallets, and other articles requiring hardness and strength. *See* FAGALES; TREE.

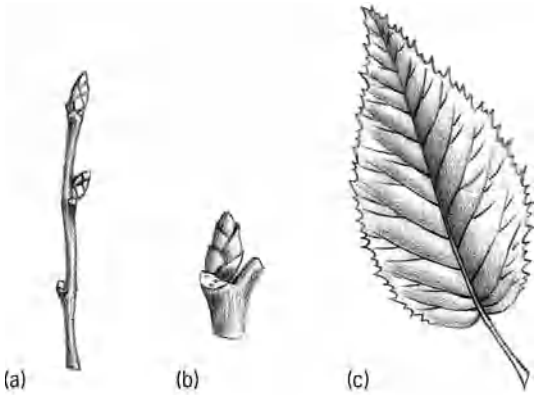

American hophornbeam (*Ostrya virginiana*). (*a*) Twig. (*b*) Terminal bud. (*c*) Leaf.

Arthur H. Graves; Kenneth P. Davis

Hoplocarida

A subclass of Crustacea, with a single extant order, Stomatopoda, commonly known as mantis shrimps. The Haplocarida was formerly included as a taxon

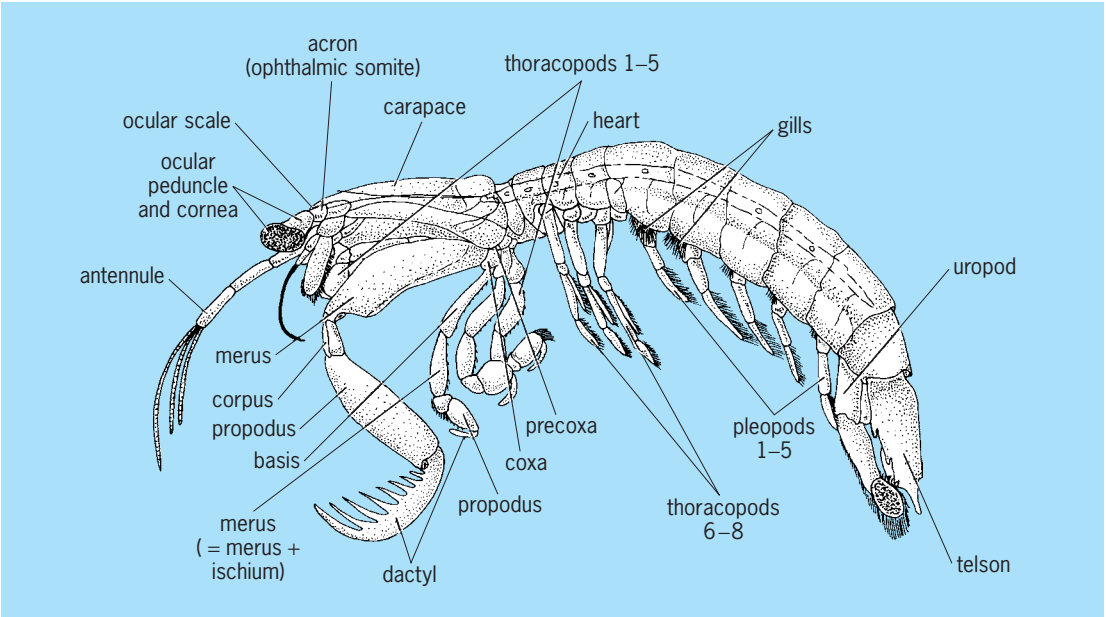

Haploid morphotype. (*After P. A. McLaughlin, Comparative Morphology of Recent Crustacea, W. H. Freeman, 1980*)

within the Eumalacostraca, and disagreement regarding its independent origin still persists. The controversy is centered on the question of whether those elements of the eumalacostracan caridoid facies observed in haplocarids represent examples of homology or convergence. Investigations of fossil (Paleostomatopoda) and Recent Stomatopoda suggest that a distinct set of morphological features, sometimes referred to as a hoploid morphotype, clearly delineate the Hoplocarida from the Eumalacostraca (see **illus.**). Those characters include cephalic kinesis, triflagellate antennules, thoracopods with trisegmented protopods, five pairs of subchelate thoracic appendages ("maxillipeds"), abdominal gills, and an elongate tubular heart that extends the length of the body and has segmentally arranged ostia and arteries. Morphological aspects of feeding observed in modern-day species, as well as features of the digestive system and abdominal musculature, tend to support the hypothesis of a distinct origin. *See* CRUSTACEA; STOMATOPODA. Patsy A. McLaughlin

Bibliography. F. R. Schram (ed.), *Crustacean Phylogeny*, vol. 1 of *Crustacean Issues*, 1983; F. R. Schram, Polyphyly in the Eumalacostraca?, *Crustaceana,* 16:243–250, 1969.

Hoplonemertini

An order of the class Enopola of the phylum Rhynchocoela, characterized by possession of an elaborate armed proboscis consisting of an anterior thick-walled tube, a median portion armed with stylets, and a posterior blind tube. The alimentary system possesses a cecum, and in some species (such as *Amphiporus lactifloreus*) the foregut (stomach) can be everted into the prey to achieve extracorporeal digestion prior to ingestion. There are two suborders, the Monostylifera and Polystylifera, separated on the number of stylets in the proboscis. Monostylifera include fresh-water, terrestrial, and symbiotic species, as well as the more common marine littoral forms. *See* ANOPLA; BDELLONEMERTINI; ENOPLA; NEMERTEA. J. B. Jennings

Horizon

Traditionally, the apparent boundary between the sky and the Earth or the sea. Without complicating factors, it would be a line located 90° from the zenith or overhead point; because of trees, hills, or buildings, the visible horizon differs from the ideal one. *See* ASTRONOMICAL COORDINATE SYSTEMS; CELESTIAL SPHERE.

Black holes. The term horizon also refers to the imaginary boundary of a black hole. The boundary, often referred to as the event horizon, is a spherical surface, with a radius equal to 1.8 mi (3 km) times the number of solar masses in the mass of the hole. A black hole forms when an object becomes so small that its gravity prevents nearby objects from escaping from its gravitational attraction. Objects close enough to a black hole, within its horizon, would have to move faster than light in order to escape its gravitational field. Because this is impossible according to the theory of relativity, escape cannot occur. *See* BLACK HOLE.

Cosmology. In cosmology the term horizon takes on still another meaning. According to the widely accepted big bang theory, cosmic evolution began from an explosion 10 to 20 billion years ago. Because nothing can travel faster than light, there is a limit to how far an observer can see or influence the surroundings; this limit is referred to as the horizon or horizon length. For example, when the universe

is a million years old, the horizon length is a million light-years or about 3×10^5 parsecs. In classical cosmology the horizon length is the distance that light can travel since the beginning of time.

The uniformity of the universe is difficult to understand, considering the small size of the horizon when the universe was young. How is it possible that twos of the universe which were separated by several million light-years or several horizon lengths when the universe was a million years old were at the same temperature at that time? Neither light nor heat could be transmitted from one region to the other in the short time that had passed since the big bang, and yet temperature uniformity is shown by observation. The so-called inflationary scenario presents a slightly modified version of the early stages of the big bang model and provides an explanation. *See* BIG BANG THEORY; COSMOLOGY; INFLATIONARY UNIVERSE COSMOLOGY; UNIVERSE.

Harry L. Shipman

Hormone

One of the chemical messengers produced by endocrine glands, whose secretions are liberated directly into the bloodstream and transported to a distant part or parts of the body, where they exert a specific effect for the benefit of the body as a whole. The endocrine glands involved in the maintenance of normal body conditions are pituitary, thyroid, parathyroid, adrenal, pancreas, ovary, and testis (see **illustration**). However, these organs are not the only tissues concerned in the hormonal regulation of body processes. For example, the duodenal mucosa, which is not organized as an endocrine gland, elaborates a substance called secretin which stimulates the pancreas to produce its digestive juices. The placenta is also a very important hormone-producing tissue.

Classification. The hormones obtained from extracts of the endocrine glands may be classified into four groups according to their chemical constitution: (1) amino acid derivatives, such as epinephrine, norepinephrine, thyroxine, and triiodothyronine; (2) proteins, such as many of the anterior pituitary hormones; (3) peptides, such as insulin, glucagon, ACTH, vasopressin, oxytocin, and secretin; and (4) steroids, such as estrogens, androgens, progesterone, and corticoids (see **table**). Hormones, with a few exceptions like pituitary growth hormone and insulin, may also be classified as either tropic hormones or target-organ hormones. The former work indirectly through the organs or glands which they stimulate, whereas the latter exert a direct effect on peripheral tissues. Many of the target-organ hormones are steroids whose production and secretion are controlled by tropic hormones, which themselves are either proteins or peptides; for example, the target-organ steroid cortisol, from the adrenal gland, is released by the action of the tropic peptide hormone pituitary ACTH. *See* PEPTIDE; PROTEIN; STEROID.

Production and circulation. Hormones are responsible for the normal functioning of the body; thus, the organism's recurring needs call for a continuing cycle of hormonal production and circulation. The various hormones are secreted chiefly under the control of the anterior lobe of the pituitary gland, known as the master gland, and in terms of their physiological role they are all either metabolic or gonad-stimulating; that is, they are involved either in the regulation of body chemistry or in the sex cycle. A key to the physiological importance of hormones is that these substances can manifest an effect even when present in very minute or trace quantities. As little as 0.05 mg of the estrogenic hormone of the ovary will induce uterine bleeding in a woman.

Regulation. This maintenance of the normal state of body with respect to the hormonal functioning of the various endocrine glands depends upon a feedback mechanism which tends to preserve an internal balance. For example, the anterior pituitary gland and the adrenal cortex are parts of such a system. The pituitary hormone ACTH stimulates the adrenal cortex to secrete its steroid hormones. If one adrenal gland is removed, the ACTH stimulation causes the cortex of the remaining adrenal to increase in size. If, on the other hand, the anterior pituitary is removed, the consequent lack of ACTH results in considerable atrophy of the cortices of both adrenal glands. If ACTH is then injected, the adrenals regain their usual size. Under normal conditions, the concentration of adrenal cortical hormones in the bloodstream controls the secretion of ACTH by the pituitary; when this concentration is high, less ACTH is released; when the concentration is low, the output of ACTH increases. The same sort of push-and-pull mechanism is believed to operate in the secretion of sex hormones by the ovaries and testes.

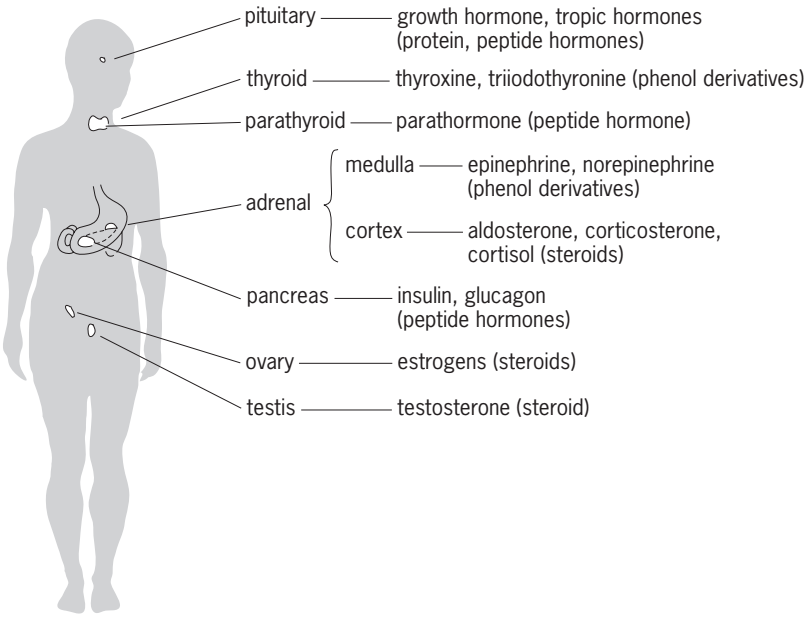

Major endocrine glands and the hormones they secrete.

Classification of hormones released from principal endocrine glands*

Endocrine gland and hormone	Target tissue	Principal actions	Chemical nature
Posterior lobe of pituitary			
Antidiuretic hormone (ADH)	Kidneys	Stimulates reabsorption of water; conserves water	Peptide (9 amino acids)
Oxytocin	Uterus	Stimulates contraction	Peptide (9 amino acids)
	Mammary glands	Stimulates milk ejection	
Anterior lobe of pituitary			
Growth hormone (GH)	Many organs	Stimulates growth by promoting protein synthesis and fat breakdown	Protein
Adrenocorticotropic hormone (ACTH)	Adrenal cortex	Stimulates secretion of adrenal cortical hormones such as cortisol	Peptide (39 amino acids)
Thyroid-stimulating hormone (TSH)	Thyroid gland	Stimulates thyroxine secretion	Glycoprotein
Luteinizing hormone (LH)	Gonads	Stimulates ovulation and corpus luteum formation in females; stimulates secretion of testosterone in males	Glycoprotein
Follicle-stimulating hormone (FSH)	Gonads	Stimulates spermatogenesis in males; stimulates development of ovarian follicles in females	Glycoprotein
Prolactin (PRL)	Mammary glands	Stimulates milk production	Protein
Melanocyte-stimulating hormone (MSH)	Skin	Stimulates color change in reptiles and amphibians; unknown function in mammals	Peptide (two forms; 13 and 22 amino acids)
Thyroid gland			
Thyroxine (thyroid hormone)	Most cells	Stimulates metabolic rate; essential to normal growth and development	Iodinated amino acid
Calcitonin	Bone	Lowers blood calcium level by inhibiting loss of calcium from bone	Peptide (32 amino acids)
Parathyroid glands			
Parathyroid hormone	Bone, kidneys, digestive tract	Raises blood calcium level by stimulating bone breakdown, stimulates calcium reabsorption in kidneys; activates vitamin D	Peptide (34 amino acids)
Adrenal medulla			
Epinephrine (adrenaline) and norepinephrine (noradrenaline)	Smooth muscle, cardiac muscle, blood vessels	Initiate stress responses; raise heart rate, blood pressure, metabolic rate; dilate blood vessels; mobilize fat; raise blood glucose level	Amino acid derivatives
Adrenal cortex			
Aldosterone	Kidney tubules	Maintains proper balance of Na^+ and K^+ ions	Steroid
Cortisol	Many organs	Adaptation to long-term stress; raises blood glucose level; mobilizes fat	Steroid
Pancreas			
Insulin	Liver, skeletal muscles, adipose tissue	Lowers blood glucose level; stimulates storage of glycogen in liver	Peptide (51 amino acids)
Glucagon	Liver, adipose tissue	Raises blood glucose level; stimulates breakdown of glycogen in liver	Peptide (29 amino acids)
Ovary			
Estradiol	General	Stimulates development of secondary sex characteristics in females	Steroid
	Female reproductive structures	Stimulates growth of sex organs at puberty and monthly preparation of uterus for pregnancy	
Progesterone	Uterus	Completes preparation for pregnancy	Steroid
	Mammary glands	Stimulates development	
Testis			
Testosterone	Many organs	Stimulates development of secondary sex characteristics in males and growth spurt at puberty	Steroid
	Male reproductive structures	Stimulates development of sex organs; stimulates spermatogenesis	
Pineal gland			
Melatonin	Gonads, pigment cells	Function not well understood; influences pigmentation in some vertebrates; may control biorhythms in some animals; may influence onset of puberty in humans	Amino acid derivative

*Reproduced with permission from P. H. Raven and G. B. Johnson, *Biology*, 6th ed., McGraw-Hill, New York, 2002.

Molecular mechanisms of action. Hormones exert their effects on peripheral tissues by regulating the behavior of cells within tissues. They do this by interacting with discrete structures in cells called hormone receptors. These hormone receptors are also generally proteins, and they bind to a select hormone with a high degree of specificity. Hormone receptors are generally of two types: cell surface receptors or intracellular receptors. For a cell surface receptor, binding of its specific hormone generates a cascade of intracellular signaling messages that regulate cell behavior. This is how the insulin receptor works and how most of the tropic hormones work. For an intracellular receptor, the specific hormone must first enter the cell and bind to the intracellular receptor. After that, the hormone and the receptor travel together to the nucleus of the cell to bind to deoxyribonucleic acid (DNA) and, thereby, regulate the expression of specific genes. Thyroid hormone and the steroid hormones generally work in this manner. *See* ADENOHYPOPHYSIS HORMONE; ALDOSTERONE; ENDOCRINE MECHANISMS; EPINEPHRINE;

INSULIN; GENE; NEUROHYPOPHYSIS HORMONE; SIGNAL TRANSDUCTION. Choh Hao Li; Theodore Mazzone

Bibliography. F. F. Bolander, *Molecular Endocrinology*, Elsevier Science & Technology, 2004; F. F. Greenspan and G. J. Strewler (eds.), *Basic & Clinical Endocrinology*, 5th ed., Appleton & Lange, Stamford, CT, 1997; A. W. Norman and G. Litwack, *Hormones*, Academic Press, San Diego, 1997; C. C. Quirk and J. H. Nilson, Hormones and gene expression—Basic principles, in L. J. DeGrool and J. L. Jameson (eds.), *Endocrinology*, pp. 3–14, Saunders, New York, 2001.

Hornbeam

The genus *Carpinus* of the birch family, represented in the United States by *C. caroliniana*, the American hornbeam or blue beech. Hornbeam is a small tree sometimes attaining a height of 35 ft (10.7 m), and it has a smooth, steel-gray, fluted bark. It grows throughout the eastern half of the United States, especially in moist soil along banks of streams; it is sometimes called water beech. When mature, it is easily recognized by its peculiar bark, by the doubly serrate leaves resembling those of sweet birch, and by the small, pointed, angular winter buds with scales in four rows (see **illus.**). The fruit is a small nutlet subtended by a three-lobed serrate bract. The wood is very hard, giving rise to the name ironwood. The bole of this tree is rarely used for sawn products, because it is short and usually crooked. Leaves turn scarlet or orange in fall.

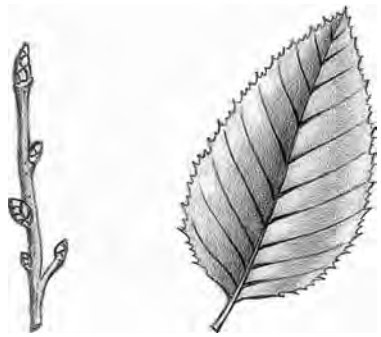

American hornbeam *Carpinus caroliniana*.

The European hornbeam (*C. betulus*) is often cultivated in parks and estates. It can be distinguished by its larger size, larger winter buds, and larger three-lobed, almost entire fruiting bracts. *See* FAGALES; FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Hornblende

The name that was traditionally assigned to common calcic amphiboles of metamorphic and igneous rocks. However, a nomenclature scheme for amphiboles was introduced in 1997 in which the names now carry strict compositional restrictions. Magnesiohornblende (contains magnesium) and ferrohornblende (contains iron) are monoclinic amphiboles with end-member compositions $Ca_2(Mg_4Al)(Si_7Al)O_{22}(OH)_2$ and $Ca_2(Fe_4^{2+}Al)(Si_7Al)O_{22}(OH)_2$, respectively (Ca = calcium, Mg = magnesium, Al = aluminum, Si = silicon, O = oxygen, Fe = iron, OH = hydroxyl). Most natural compositions differ significantly from these ideal end members. In particular, they show solid solution toward pargasite and ferropargasite with ideal end-member compositions $NaCa_2(Mg_4Al)(Si_6Al_2)O_{22}(OH)_2$ and $NaCa_2(Fe_4^{2+}Al)(Si_6Al_2)O_{22}(OH)_2$ [Na = sodium]. Significant deviations from these compositions are denoted by the addition and replacement of prefixes and adjectival modifiers characteristic of the compositions involved. Thus fluorohornblende (contains fluorine, F) has the end-member composition $Ca_2(Mg_4Al)(Si_7Al)O_{22}F_2$, in which all of the OH in hornblende has been replaced by F. When used to denote an amphibole of known chemical composition, the term hornblende is never used without a prefix or adjectival modifier. The unmodified term hornblende specifically refers to a calcic amphibole identified by physical or optical properties without characterization of the chemical composition. *See* AMPHIBOLE.

Physical and optical properties. In hand specimen, hornblende varies in color from dark green to brown to black. Crystals are normally anhedral to subhedral and prismatic, with prominent {110} faces meeting at 56° and 124°; end sections show {110} cleavage, and prism faces show longitudinal cleavage traces. In thin section, hornblende is strongly pleochroic (changes color on rotation in plane-polarized light) in various shades of green and brown. Electronic absorption spectroscopy shows that this is due primarily to intervalence charge-transfer reactions among the octahedrally coordinated transition-metal cations in the structure. The complexity of chemical substitutions possible in hornblende precludes characterization of chemical variation by optical methods. However, the textural information from this technique is important in the interpretation of hornblende paragenesis. *See* COORDINATION CHEMISTRY; CRYSTALLOGRAPHY.

Occurrence. Hornblende is a common rock-forming mineral in medium- and high-grade metamorphic rocks, particularly those of mafic and ultramafic composition. In mafic rocks, it first appears in the upper part of the low grade by a chemical reaction involving the disappearance of actinolite, a nonaluminous calcic amphibole. This change is extremely noticeable in thin sections, very pale-green actinolite giving way to blue-green hornblende. With prograde metamorphism, the composition of the hornblende gradually changes in a highly complex manner that is a function of temperature, pressure, oxygen fugacity (a measure of the activity of oxygen), and the chemical composition of the rock. This causes a gradual color change from blue-green through various shades of green to olive green and brown. At the middle of the high grade, hornblende becomes unstable and breaks down to form pyroxene (plus other minerals). The prominence of

hornblende in medium-grade metabasic rocks has led to these rocks being called amphibolites. *See* PYROXENE.

Hornblende is commonly found as a minor phase in a wide variety of igneous rocks. Magnesium-rich hornblendes do occur as primary phases in basic and ultrabasic rocks, but this is not common. Igneous amphiboles are most abundant in calcic-alkaline diorites, granodiorites, and granites, becoming more iron-rich with increasing acidity of the host rock. This compositional trend is also characterized by a progressive increase in the alkali content of the amphibole, and hornblende grades into hastingsite, riebeckite, and arfvedsonite in granitic rocks. *See* GRANITE; GRANODIORITE; IGNEOUS ROCKS.

Due to its complex structure and chemistry, hornblende contains much information on its formation. Its behavior is understood reasonably well, and hornblende is of considerable use in interpreting the geological history of the rocks in which it occurs. *See* AMPHIBOLITE; METAMORPHISM.

Exsolution. For temperatures at which the amphiboles form, amphiboles of intermediate composition occur. These are unstable at lower temperatures and unmix, forming spectacular exsolution textures that are very useful in deciphering details of the thermal history of the host rock. Hornblende shows partial to complete miscibility with other amphibole subgroups, in particular the ferromagnesian and the alkali amphiboles. In medium-grade metamorphic rocks, coexisting hornblende and cummingtonite-grunerite is very common, and hornblende is characterized by prominent exsolution lamellae of cummingtonite-grunerite. Details of the orientation of these exsolution lamellae can give considerable information about the cooling history of the rock. *See* CUMMINGTONITE.

Frank C. Hawthorne

Bibliography. F. C. Hawthorne, The crystal chemistry of the amphiboles, *Can. Mineral.*, 21:173–480, 1983; F. C. Hawthorne, The crystal chemistry of the amphiboles, *Rev. Mineral.*, 9A:1–102, 1982; B. E. Leake et al., Nomenclature of amphiboles, *Can. Mineral.*, 35:219–246, 1997; P. Robinson et al., Phase relations of metamorphic amphiboles: Natural occurrence and theory, *Rev. Mineral.*, 9B:1–227, 1982.

Hornfels

A metamorphic rock that has been subjected to heating during contact metamorphism around intrusive igneous rocks. Hornfels is typically fine-grained, although where it is subjected to high temperatures, large crystals called porphyroblasts can form. In outcrop, hornfels is notoriously tough and can be difficult to sample. Mineral grains in hornfels are randomly oriented, with no preferred alignment of crystals to form foliation or cleavage planes. This texture indicates that the hornfels was not subjected to significant stresses during contact metamorphism.

Hornfels generally originates from sediments that undergo mineralogical changes, the nature of which depend on the magnitude of heating. The types of minerals that form are strongly dependent on the bulk composition. Minerals in hornfels formed from metamorphism of limestones, which are rich in calcium oxide (CaO), carbon dioxide (CO_2), and various amounts of magnesium oxide (MgO), iron oxide (FeO), and aluminum oxide (Al_2O_3), include (from high to low temperature) fosterite, diopside, tremolite, talc, and brucite. Other minerals that may be present include wollastonite, vesuvianite, anorthite, and grossular garnet, depending on the bulk composition of the rock. These minerals are also common in skarn-type deposits commonly found within contact metamorphic environments. *See* LIMESTONE; SKARN.

Pelitic sediments are rich in chemical constituents such as silicon dioxide (SiO_2), Al_2O_3, MgO, FeO, potassium oxide (K_2O), and water (H_2O), with relatively minor amounts of CaO, sodium oxide (Na_2O), manganese oxide (MnO), and titanium dioxide (TiO_2). Metamorphism of these sediments to form hornfels results in formation of minerals such as chlorite, muscovite, biotite, andalusite, sillimanite, cordierite, garnet, staurolite, and K-feldspar. At extremely high temperatures (>800°C or 1470°F) aluminum-rich minerals such as sapphirine, spinel, and corundum form. Deposits of emery, utilized for abrasives, are aluminum-rich hornfels that are products of high-temperature contact metamorphism. Chemical study of emeries indicates a general lack of alkali elements (K, Na, and Ca), which has been used to argue that they form as a result of extraction of a melt phase during high-temperature contact metamorphism. Examples of emery deposits include the Cortlandt Complex in New York; the Martinsville Intrusive Complex in Virginia; and localities in Turkey, Siberia, and Australia. *See* EMERY.

During contact metamorphism, hornfels typically forms in the highest-temperature part of aureoles adjacent to the pluton. Further away from the pluton, metamorphism of sediments results in development of schists and phyllites. For example, the Ballachulish Igneous Complex in Scotland is a composite pluton consisting of granite and granodiorite. Around the pluton, low-grade chlorite-bearing slates are progressively metamorphosed, resulting in the systematic appearance from low to higher temperature of cordierite + biotite + muscovite phyllite to cordierite + K-feldspar + biotite hornfels. In hornfels of a slightly different composition, muscovite is preserved, resulting in a hornfels with the composition andalusite + K-feldspar + cordierite + biotite + muscovite. Adjacent to the contact with the pluton, these muscovite-bearing hornfels undergo partial melting, resulting in the segregation of K-feldspar + plagioclase + quartz from the metamorphosed sediment as a result of partial melting. *See* PLUTON; SLATE.

Metamorphic studies of hornfels provide an important avenue to documenting the temperature and, in particular, the pressure (that is, the depth)

during emplacement of the intrusive igneous rock that provides the heat. For example, for hornfels metamorphosed at low pressures [<2 kilobars (200 megapascals) or ~7 km (4 mi)] the predominant porphyroblasts present in hornfels are predicted to be cordierite + K-feldspar + andalusite or sillimanite. The presence of andalusite versus sillimanite is a function of temperature, with sillimanite favored at higher temperatures. At pressures greater than 2 kbar (200 MPa) but less than 3.5 kbar (350 MPa), staurolite is found in addition to cordierite + K-feldspar + andalusite or sillimanite. A pressure zone of 3.5–4 kbar (350–400 MPa) is distinctive because of the presence of garnet, and at still higher pressures (>4 kbar or 400 MPa) is marked by the lack of andalusite and the predominance of kyanite or sillimanite along with garnet and staurolite. The utility of this pressure scheme is that the depth at which sedimentary rocks were subjected to contact metamorphism can be evaluated by determining the metamorphic minerals in hornfels. This is a routine endeavor completed during field mapping and petrographic examination of hornfels. One difficulty in this technique is that the mineralogy used to distinguish between pressure zones is also a function of bulk rock chemistry, which means that the growth of a particular metamorphic mineral may vary within a sequence of metamorphosed sedimentary rocks. *See* METAMORPHIC ROCKS; METAMORPHISM; MINERALOGY.

Matthew W. Nyman

Bibliography. D. M. Carmichael, Metamorphic bathozones and bathograds: A measure of the depth and postmetamorphic uplift and erosion on the regional scale, *Amer. J. Sci.*, 278:769–797, 1978; D. M. Kerrick (ed.), *Contact Metamorphism*, Mineralogical Society of America, 1991.

Horology

Measurement of the time dimension. In practice, horology is the search for a steady or repetitive action, and the design of an instrument to perform that action and to indicate (read out) a measure of the action. Until early in the twentieth century, horology dealt with mechanical instruments, with effort distributed between improving accuracy and decreasing size of timepieces. Increasingly, however, electronic instruments have provided means for meeting these objectives. *See* ATOMIC CLOCK; ATOMIC TIME; CHRONOMETER; CLOCK (MECHANICAL); DYNAMICAL TIME; EARTH ROTATION AND ORBITAL MOTION; ESCAPEMENT; MASER; OSCILLATOR; PENDULUM; PULSAR; QUARTZ CLOCK; SATELLITE NAVIGATION SYSTEMS; SUNDIAL; TIME; WATCH.

Frank H. Rockett; Donald C. Backer

Bibliography. J. E. Burnett, *Time's Pendulum*, 1998; J. T. Fraser et al. (eds.), *The Study of Time,* vols. 1–10, 1972–2000; J. L. Jesperson and D. W. Hanson, Special issue on time and frequency, *Proc. IEEE*, 79(7):894–1079, 1991; J. L. Jesperson and J. Fitz-Randolph, *From Sundials to Atomic Clocks*: *Understanding Time and Frequency*, 2d ed., 2000; P. Kartaschoff, *Frequency and Time*, 1978; J. D. Kraus, *Radio Astronomy*, 2d. ed., 1986; A. R. Thompson, J. M. Moran, and G. W. Swenson, *Interferometry and Synthesis in Radio Astronomy*, 2d ed., 2001; M. E. Whitney, *The Ship's Chronometer*, 1984.

Horse production

The science of breeding, raising, and caring for the horse. Successful horse production depends upon the use of the most recent advances in breeding methods such as artificial insemination and embryo transfer, feeding, physical conditioning, genetics, health care, marketing, and general management.

Breeds and types. Horses are grouped in four classes—ponies, light horses, draft horses, and coach horses—based upon their height and size. Height is measured at the horse's shoulder by the standard unit of the hand, which is 4 in. (10 cm). Ponies are less than 14.2 hands, and most weigh less than 900 lb (408 kg). Horses are over 14.2 hands and are classified further as light or draft based upon their size: light horses usually weigh 900–1400 lb (408–635 kg), and draft horses weigh over 1400 lb (635 kg). Light horses are also classified according to their type and use, including the polo, racing, stock, hunter, jumper and gaited horses. Race horses include harness horses, such as Standardbred trotters and pacers; flat racers, such as Thoroughbreds and Quarter Horses; steeplechasers; and endurance racers. Driving horses include the heavy harness, fine harness, and roadster types, which are all used to hitch to some vehicle.

Within each of the classifications there are usually several breeds, each of which has certain distinguishing characteristics such as conformation, color, and sometimes function.

Pony breeds. There are four common pony breeds in the United States—Welsh, Shetland, Connemara, and Pony of America. The Welsh or Welsh Mountain Pony originated in Wales and was mainly used for work in coal mines prior to importation (**Fig. 1**). The Shetland Pony was developed in the Shetland Islands and is of small size (9–11 hands). Connemaras are the largest pony breed and were developed in Ireland. The Pony of America was developed in the United States by crossing a Shetland stallion with an Appaloosa mare; it must retain the Appaloosa color pattern to be eligible for registration.

Light horses. The Arabian horses, all of which descend from five foundation mares known as the Al-Khamesh, are the foundation for most of the light horse breeds. Arabians are distinguished from other breeds by a thin dished face and prominent forehead, and have gained recognition for their endurance.

Thoroughbred horses have been bred for their racing ability for several hundred years (**Fig. 2**). All Thoroughbreds trace back to three stallions—Byerly Turk, Darley Arabian, and Godolphin Barb—which were imported to England in the late 1600s and early

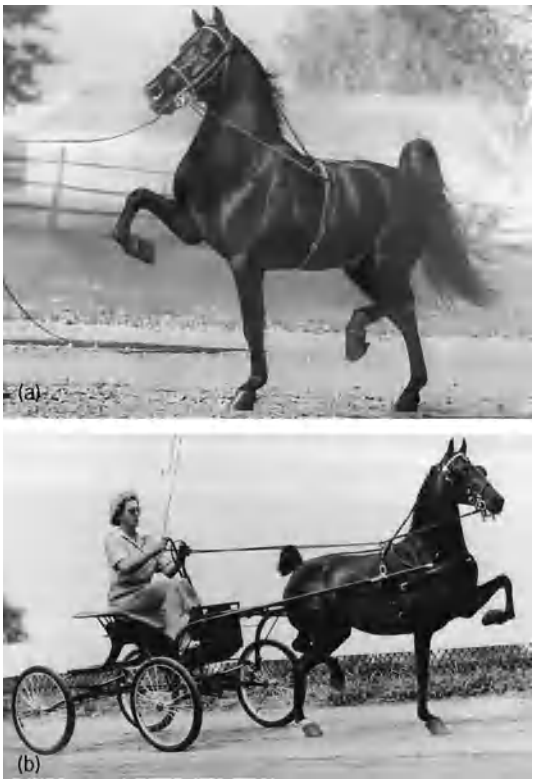

Fig. 1. Two popular pony breeds. (*a*) Welsh pony stallion. (*b*) Hackney pony mare. (*Photographs by McClasky*)

1700s, and are responsible for the Thoroughbred's refinement and speed. The Thoroughbred is now recognized for its versatility and is used for many purposes.

The American Quarter Horse is the most popular horse breed in America (**Fig. 3**), as proved by the registration of over 2,000,000 head since the establishment of the breed association in 1941. The history of the breed traces to the colonial era when these horses were used to run short races, but later they became the most popular horse for the cowboys working cattle on the western ranches. Within the breed there are several types, including working cow horses, racers, hunters, and jumpers.

Standardbred horses were developed for their ability to race as trotters or pacers, and their name was derived from the original registration requirement of racing a mile in a standard time (**Fig. 4**). Morgan horses, known for their versatility, trace their development to one sire, Justin Morgan, and became popular on New England farms.

Various other breeds were developed for their ease of riding (**Fig. 5**). The American Saddlebreds excel as show horses and are shown as three- and five-gaited and as fine harness horses. (Five-gaited horses are shown at the rack and slow gait in addition to the walk, trot, and canter required of the three-gaited horse.) Tennessee Walking horses were developed for their ability to perform the running walk, an easy gliding gait. Breed associations whose only requirement is a specific color pattern include the White-Horse Registry, Cream Horse Registry, America Buckskin Registry Association, International Buckskin Horse Association, and Appaloosa, Pinto, Paint, and Morocco Spotted Horse associations.

Draft horses. Draft horses were developed for their pulling power, and are characterized by their size (1500–2000 lb or 680–907 kg), heavy muscling and bone, and deep, wide body (**Fig. 6**). The Percheron originated in France and has more refined characteristics than some of the other draft breeds. Suffolks were developed in England and were used exclusively for farm work; they are the only draft breed that breeds true for color—chestnut. Clydesdales are noted for their long hair on the back of their legs. Belgians, because of their action, endurance, and strength, were the primary horses used by the Roman cavalry. Shires, which also have long hair on the backs of their legs, were the least popular of the breeds used in the United States for farming.

Coach horses. Heavy harness or coach horses were used to pull coaches, and included breeds such as Hackney, Cleveland Bay, French Coach, Russian Orloff, Yorkshire, and American Carriage Horse.

Breeding practices. The horse breeding industry has become very specialized and technical. Computerized records of performance events and pedigrees have allowed the breeder to have instant access to necessary information. This information and the individuals conformation are used by breeders for selection or culling of breeding stock.

The natural breeding season of horses is during the spring and early summer. Most mares do not have estrous cycles during the winter. Fillies reach puberty at about 1 year of age and at that time start to have estrous cycles, but they are usually not bred until they are at least 3 years of age. Each estrous cycle lasts about 21–23 days. The period of estrus when a mare will accept a stallion is 5–6 days; ovulation

Fig. 2. The Thoroughbred. (*Photograph by The Thoroughbred Record*)

occurs about 2 days before the end of the estrous period. If the mare is bred naturally, she is bred every other day during estrus, but if she is bred by artificial insemination, she is bred only once, just prior to ovulation. Mares are bred an average of two estrous cycles per conception. *See* ESTRUS.

Colts reach sexual maturity when they are about 15 months old. They may be used to breed 6–8 mares when they are 2 years old, but most are not used as breeding stallions until they are 4–5 years old. During a breeding season, stallions can successfully breed about 60 mares, unless artificial insemination is utilized in which case 200 or more mares may be bred by a stallion. Stallions produce 50–75 ml of semen, which contains about 15 billion spermatozoa, per ejaculation. When a mare is inseminated, 500 million actively motile spermatozoa are used.

Embryo transfer is sometimes used to obtain a foal from some mares that are infertile, or from mares that need to remain in competitive training. By this procedure the mare is bred and the embryo is flushed from the uterus when it is 6–8 days old. The embryo, which sometimes is stored frozen, is then transferred to a host mare that carries the fetus to term. Identical twins can be produced by dividing the embryo shortly after conception and placing each half in a host mare.

Colts that do not have the desired conformation or fail to meet performance expectations are castrated, after which they are known as geldings. Most colts are castrated between 6 and 12 months of age but can be castrated at any age. Very few stallions are used for pleasure or recreational purposes because of their unpredictable behavior; geldings, however, are usually docile and easily handled.

Feeding. Even though many scientific advances have been made in the understanding of the nutrient requirements of horses, the "eye of the master" determines the success of feeding a horse properly. Since there is a wide variation in the nutrients required to maintain proper physical condition of the same types of horses kept under similar circumstances, most horses are fed as individuals. Horses evolved on the open grassy plains and can obtain their nutrient requirements (as well as good exercise) from good pastures. Today, however, most horses are kept in paddocks and are fed hay for the roughage portion of their diet. *See* ANIMAL FEEDS.

Mature horses receiving a moderate amount of exercise can be maintained exclusively on hay. When exercised heavily, however, their energy requirement increases, so they are fed less hay and grain is added to their diet. The ratio of hay to grain depends on the amount of exercise a horse gets; more active horses are fed more grain. The standard hays fed to horses are alfalfa, oat, timothy, and various grasses, while the most commonly used grains are oats, corn, milo, and barley.

Young growing horses, lactating mares, and pregnant mares (in their last 90 days of pregnancy) have increased requirements for protein, minerals, and vitamins. Higher protein needs are met by feeding legume hays or protein supplements such as linseed or soybean meals, and mineral supplements such as dicalcium phosphate provide the extra calcium and phosphorus that is usually needed. If adequate vitamins are not present in the natural feeds, artificial sources are added. Salt should always be available. The average horse drinks 8–10 gallons (30–38 liters) of water per day, a requirement which increases during hot weather and with intensive exercise.

J. Warren Evans

Fig. 3. American quarter horse. (*Photograph by J. A. Stryker*)

Disease and parasite control. Control measures frequently require the services of a professional veterinarian. However, good management practices can

Fig. 4. Standardbred trotter. (*Photograph by The Harness Horse*)

Fig. 5. Saddle horse breeds. (*a*) Five-gaited American saddle horse (*photograph by McClasky*). (*b*) Tennessee walking horse (*USDA photograph*). (*c*) Morgan mare. (*d*) Arabian stallion (*USDA photograph*). (*e*) Palomino gelding (*photograph by J. L. A. duPont*). (*f*) Appaloosa stallion (*photograph by H. H. Sheldon*).

prevent some troubles. Daily inspection may permit early detection of disease and treatment before the condition becomes serious.

The breathing of a horse should be free, soft, and noiseless. Normal respiration rate for a horse at rest is 8–16 per min. Normal pulse rate for a horse at rest is 36–40 per min. The pulse may be taken by palpating an artery that crosses the jawbone just in front of the large cheek muscle. Normal body temperature (rectal) for a horse at rest is 100°F (38°C). A horse's coat should be smooth and glossy, and its skin loose and supple. The weight should be borne equally on both front feet. The hindlegs, however, are rested alternately.

Common preliminary symptoms of disease include high temperature; fast, irregular, or noisy respiration; rapid or weak pulse; loss of appetite; sweating without known cause; stiffness; lameness; coughing; inflamed mucous membranes; discharges from nose, eyes, or genitals; diarrhea; constipation; dejection; restlessness; rolling; groaning; and heat or swelling in any part.

Internal parasites of horses must be controlled by treatment. Effective prevention of infestation is impossible. Conditions that favor good pasture also favor the completion of the life cycle of horse parasites and the likelihood of infestation.

Strongyles, or bloodworms, are the most damaging of the internal parasites of horses. These intestinal worms range from $^1/_3$ to $2^1/_2$ in. (8 to 64 mm) in length. When present in large numbers they cause anemia, weakness, and emaciation. Veterinarians often prescribe thiabendazole or cambendazole to eliminate strongyles.

Ascarids, or large roundworms, are particularly harmful to foals and young growing horses. Yellowish-white in color and up to 1 ft (30 cm) in length, ascarids stay in the small intestine. Toxins produced by these worms may kill a foal. Carbon disulfide given by means of a stomach tube eliminates roundworms. However, a piperazine–carbon disulfide complex or combendazole are more commonly used.

Bots, the larvae of the botfly, attach to the wall of the horse's stomach and interfere with the normal work of the stomach. Treatment is the same as for ascarids. *See* MYIASIS.

Lice are the most common of the external parasites of horses. Lice can be eliminated by several insecticides, of which coal tar dips and rotenone are the safest. *See* ANOPLURA; INSECTICIDE.

Sales. Selling horses profitably is the end and aim of most horse breeders. Young stock must be well bred, well fed, and well managed to command high prices. In addition, older horses must give evidence of good training and ability to perform well. Although most horses are sold privately, horse auctions are common throughout the United States. The most noted of these auctions are those for Thoroughbred yearlings held each summer at the Keeneland racetrack, Lexington, Kentucky, and at Saratoga, New York. Only a very select group of yearlings from

Fig. 6. Breeds of work horse. (*a*) Percheron mare; (*b*) Belgian stallion (*J. F. Abernathy Live Stock Photography Co.*). (*c*) Clydesdale stallion; (*d*) Shire stallion; (*e*) Suffolk stallion (*USDA photographs*).

the best studs is handled at these sales. Standardbred yearling sales patronized by the leading market breeders are held each fall at Lexington, Kentucky, and at Harrisburg, Pennsylvania. John M. Kays

Horseradish

A hardy perennial crucifer, *Armoracia rusticana*, of eastern European origin belonging to the plant order Capparales. Horseradish is grown for its pungent roots, which are generally grated, mixed with vinegar and salt, and used as a condiment or relish. Propagation is by root cuttings, and the crop is grown like an annual. The individual roots or sets are uncovered by hand usually twice during the summer and stripped of all side roots (see **illus.**). Maliner Kren is a common variety. Harvesting of the roots occurs in the fall, usually $3^1/_2$ to 4 months after planting. Production in the United States is limited to northern areas; Illinois, Wisconsin, and Missouri are important producing states. *See* CAPPARALES. H. John Carew

Roots of horseradish stripped of side roots.

Horst

A segment of the Earth's crust, generally long as compared to its width, that has been upthrown relative to the adjacent rocks (see **illus.**). Horsts range in size from those that have lengths and upward displacement of a few inches to those that are tens of miles long with upward displacements of thousands of feet. The faults bounding a horst on either side commonly have inclinations of 50–70° toward the downthrown blocks, and the direction of

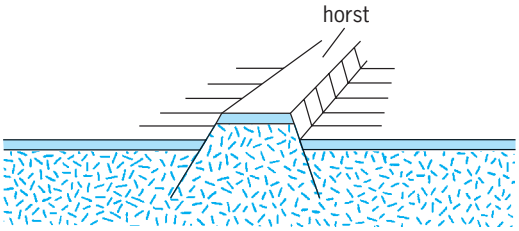

Simple horst with associated faults. (*After A. K. Lobeck, Geomorphology, McGraw-Hill, 1939*)

movement on these displacements indicates that they are gravity faults. These relationships suggest that horsts develop in regions where the crust has undergone extension. They may form in the crests of anticlines or domes, or may be related to broad regional warpings. *See* EARTH CRUST; FAULT AND FAULT STRUCTURES; GRABEN. Philip H. Osberg

Horticultural crops

Intensively managed plants cultivated for food or for esthetic purposes. Plant agriculture is divided traditionally into the fields of agronomy (herbaceous field crops, mainly grains, forages, oilseeds, and fiber crops), forestry (forest trees and products), and horticulture (garden crops, particularly fruits, vegetables, spices and herbs, and all plants grown for ornamental use). Most horticultural plants are utilized in the living state, with water essential to quality; thus most horticultural plants and products are highly perishable. *See* AGRICULTURAL SCIENCE (PLANT); AGRONOMY; FLORICULTURE; FOREST AND FORESTRY.

Classification. Custom has defined the classification of many crops. Thus field corn (maize, usually consumed as an animal feed grain) is agronomic; sweet corn (maize with sweet immature kernels consumed by humans) is horticultural. Pines grown for timber or naval stores are a forest crop; pines grown as landscape trees are a horticultural crop. Some crops that could be considered horticultural are not treated as such; tobacco is an example. Finally, many tropical species which are sometimes treated in a special category as plantation or estate crops are included as horticultural crops. These include coffee, tea, and cacao (beverage crops), plants grown for drugs and medicines (medicinal crops), tree crops grown for oils as oil palm and coconut (oil seed crops), and condiments and flavorings (spice crops).

Horticultural crops are usually classified as edibles or ornamentals. Edible crops which are used for direct human consumption are commonly subdivided into fruits or vegetables, but this classification is traditional and difficult to define precisely.

Fruit crops. Fruit crops in the horticultural sense are cultivated for tissues associated with the botanical fruit, that is, seed-bearing structures derived from the flower, which are usually pulpy and tasteful.

Fruit plants are typically woody perennials, but there are exceptions (strawberry, banana). Temperate fruit crops are deciduous; subtropical and tropical fruits are usually evergreen. Fruits borne on low-growing shrubs, vines, or herbaceous plants are known as small fruits. Trees or shrubs bearing nuts, characterized by a hard shell separated by a firm inner kernel (the seed), are often treated as a special category of fruit crops.

A traditional classification of fruit crops (with some examples) is presented below.

- Temperate fruits
  - Small fruits
    - Berries (blueberry, cranberry, strawberry)
    - Brambles (blackberry, raspberry)
    - Vine fruits (grape, kiwifruit)
  - Tree fruits
    - Pome fruits (apple, pear, quince)
    - Stone fruits (apricot, cherry, peach, plum)
    - Nuts (almond, chestnut, filbert, pecan, walnut, pistachio)
- Subtropical and tropical fruits
  - Small fruits
    - Vine fruits (passion fruit)
    - Herbaceous fruits (banana, papaya, pineapple)
  - Tree fruits
    - Citrus fruits (grapefruit, lemon, lime, orange, pummelo, tangerine)
    - Other fruits (avocado, date, fig, mango, mangosteen)
    - Nuts (Brazil nut, cashew, macadamia)

See FRUIT; FRUIT, TREE; NUT CROP CULTURE.

Vegetable crops. Vegetable crops in the horticulture sense are commonly herbaceous plants grown as annuals or biennials and occasionally as perennials that have edible parts (including, confusingly, the botanical fruit). Examples of edible parts include the root (sweet potato), tuber (potato), young shoot (asparagus), leaf (spinach), flower buds (cauliflower), fruit (tomato), and seed (pea). Temperate-tropical distinctions are not as important in vegetables as in fruit crops. Many tropical perennials (for example, tomato) are grown as annuals in temperate areas. A classification of vegetables based on use and botanical affinity is listed below.

- Cole crops (broccoli, cabbage, cauliflower)
- Legumes (pulse) crops (bean, lentil, pea)
- Solanaceous fruit (chili pepper, eggplant, tomato)
- Vine crops or cucurbits (cucumber, muskmelon, squash, watermelon)
- Greens (chard, spinach)
- Salad crops (celery, lettuce, parsley)
- Root crops (beet, carrot, sweet potato)
- Tuber crops (Jerusalem artichoke, potato)
- Bulb and corms (garlic, onion, shallot)
- Herbs (rosemary, sage)
- Mushrooms [*Agaricus, Lentinus* (shiitake)]

Ornamental crops. Plants grown for ornamental use, such as cut flowers, bedding plants, interior foliage plants, or landscape plants, represent an enormous group and include thousands of species. They may be grouped as follows.

- Flowers and bedding plants:
  - True annuals (marigold, petunia, zinnia)
  - Biennials (English daisy, foxglove)
  - Perennials (daylily, delphinium, iris, peony, rose)
- Bulbs and corms (crocus, gladiolus, narcissus, tulip)
- Foliage (interior) plants (philodendron, sansevieria)
- Ground covers and vines (English ivy, Japanese spurge, myrtle)

Lawn (turf) plants (Bermuda grass, bluegrass, fescue, perennial ryegrass)
Evergreen shrubs and trees:
Broadleaf (holly, rhododendron)
Narrowleaf (fir, juniper, yew)
Deciduous shrubs (dogwood, forsythia, lilac, viburnum)
Deciduous trees (ash, crabapple, magnolia, maple)

See ORNAMENTAL PLANTS. Jules Janick

Hospital infections

Infections acquired during a hospital stay; also known as nosocomial infections. They may be recognized during or after hospitalization. They usually appear during hospitalization, but as many as 25% of infections related to surgery occur after discharge. Data suggest that in the United States 4.5% of all hospitalized patients can expect at least one such infection, which translates into almost 2 million infections annually. Thus this major public health problem accounts for 299 million patient-days in United States hospitals each year.

The mere presence of microorganisms in or on a patient does not indicate an infection and may only represent colonization. Colonization implies establishment, growth, and multiplication of organisms (such as normal bacteria in the mouth), but without clinical symptoms. Although viruses, bacteria, and fungi are always present in the human body, infection occurs as a result of some alteration in the normal balance between an infectious organism and a susceptible person. Factors that may influence the development of an infection include the individual's defense mechanisms and the site of entry of the microorganism. Infectious agents may enter the body through breaks in the skin and mucous membranes or by way of the respiratory, gastrointestinal, or urinary tract. Resistance to infection may be influenced by age (very young or very old), chronic disease (such as diabetes and sickle-cell disease), or cancer. Genetic, hormonal, nutritional, and hygienic factors also play a role in modifying the body's defense mechanisms. Most nosocomial infections (93%) are caused by bacteria; fungi account for about 6%; and viruses, protozoa, and parasites account for the remaining 1%.

Nosocomial infections occur most frequently in the urinary tract, in surgical wounds, as a complication of pneumonia, and in association with bacteremia.

Urinary tract infection. Urinary tract infection is an inflammatory process occurring in the kidney, ureter, bladder, or adjacent structures. Microorganisms usually enter through the urethra. The infection increases in severity as it moves through the urinary tract; at any point it may lead to kidney damage, disseminated infection, bacteremia, and associated morbidity and mortality. Urinary tract infection is of major importance, not only because it accounts for almost half of all nosocomial infections, but because of the potential for serious complications. Urine is normally sterile, and so the presence of bacteria in a properly collected urine specimen is usually evidence of infection.

The most common cause of nosocomial urinary tract infections is an indwelling urinary catheter. Bacteria are believed to enter at the time of catheter insertion or later by migration along the catheter wall into the bladder. *Escherichia coli*, an aerobic gram-negative bacillus, is responsible for about 80% of these infections. They can be prevented by maintaining asepsis during insertion and removal of urinary catheters and removing them at the earliest opportunity. *See* URINARY TRACT DISORDERS.

Surgical wound infections. Surgery is associated with more than 40% of all hospital admissions but with more than 70% of all nosocomial infections, which are not necessarily confined to the surgical wound itself. Almost 75% of all pneumonias, 56% of all urinary tract infections, and 54% of all cases of bacteremia occur in surgical patients.

Approximately 5% of all incisions become infected, but the incidence varies widely with the type of operation and the health of the patient. The risk of a surgical wound infection increases with increasing age of the patient, length of the operation, and duration of hospital stay before surgery. Infection rates are lowest in incisions made through clean tissue, that is not inflamed, infected, or contaminated, and highest when the tissue is already infected or contaminated, as in cases of trauma or ruptured appendix.

Staphylococcus aureus is the most common bacterium found in surgical wound infections, but *E. coli* is also common in such wounds. Other bacilli, such as *Pseudomonas aeruginosa* and *Serratia marcescens*, are seen in debilitated patients. The principal source of infection is the patient's own flora, and these microorganisms probably gain entrance during surgery. Some microorganisms are, however, transmitted from operating-room personnel, possibly through microscopic holes in surgical gloves. Airborne infections occur rarely and only under unusual circumstances. *See* STAPHYLOCOCCUS; SURGERY.

Respiratory infections. Pneumonia accounts for about 10% of all hospital-acquired infections, in part because the normal respiratory defense mechanisms are often weakened in critical illnesses. Anesthesia, alcohol intoxication, or convulsions may suppress the normal cough reflex and permit material to be aspirated into the lungs. Fluid accumulations in the lungs from congestive heart failure, chest injuries, or viral infection may encourage bacterial growth and lead to pneumonia.

Shortly after admission to a hospital, the patient's upper respiratory tract becomes colonized with microorganisms unique to the hospital environment and, in large measure, resistant to antibiotics commonly used there; that obviously complicates treatment. Furthermore, ventilatory support

equipment may harbor bacteria that increase the risk to the patient. Such respiratory infections are commonly caused by *Klebsiella pneumoniae* and *Pseudomonas aeruginosa*. *See* PNEUMONIA.

Bacteremia. Bacteremia, the invasion of the normally sterile bloodstream by bacteria, is a serious condition that may rapidly lead to shock and death. Invasion from an infection elsewhere in the body is termed secondary bacteremia. If no such source can be found, the bacteremia is classified as primary. Although primary bacteremia accounts for no more than 5% of all nosocomial infections, it is important because of its seriousness and its preventability. Many if not most cases are related to intravenous fluid therapy, which is administered to more than half of all patients at some time during their hospital stay.

Some outbreaks of bacteremia in the United States have been caused by contamination of commercially prepared intravenous fluids, but the principal source of infection is the insertion site of the needle or catheter. Small, short-bore steel needles are safer than the longer and larger plastic cannulae, but in either case the risk is correlated with the duration of use. Sterile insertion and proper maintenance of intravenous needles and their sites are of the utmost importance.

Other causes and consequences. Some antibiotics have direct toxic effects. Some drugs and (rarely) antibiotics injure the bone marrow, reducing the production of white blood cells needed to fight infection.

About 25% of hospital patients are given antibiotics, all of which cause major alterations in the microbial populations of the skin and the gastrointestinal and respiratory tracts. Infections with antibiotic resistant organisms may be more difficult to treat. *See* ANTIBIOTIC.

Acquired immune deficiency syndrome. Viral infections, although sometimes difficult to recognize, are becoming an increasing concern in hospitals, particularly because of the possibility of transmitting the human immunodeficiency virus (HIV), the causative organism for acquired immune deficiency syndrome (AIDS). The virus may be transmitted within the hospital through transfusion of infected blood or blood products, although the screening of all donated blood for HIV antibodies makes it extremely unlikely. Hospital workers are at risk for acquiring the virus from needle pricks only if the needle has contacted HIV-infected blood. Since the virus is not transmitted by the respiratory route or by casual contact, no one is endangered simply by being in the same hospital with HIV-infected patients. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); INFECTION; MEDICAL BACTERIOLOGY.

Bruce B. Dan

Hot spots (geology)

The surface manifestations of plumes, that is, columns of hot material, that rise from deep in the Earth's mantle. Hot spots are widely distributed around the Earth (**Fig. 1**). One of their characteristics is an abundance of volcanic activity which persists for long time periods (greater than 1 million years). When the lithosphere (the rigid outer layer of the Earth) moves over a plume, a chain of volcanoes is left behind that progressively increases in age along its length. Hot spots are believed to be fixed with respect to each other and the deep mantle so that the age and orientation of these chains provide information on the absolute motions of the tectonic plates. *See* LITHOSPHERE; PLATE TECTONICS.

Volcanic expression. The Hawaiian-Emperor seamount chain in the central Pacific Ocean is a good example of a volcanic chain that was generated at a hot spot. The 3400-mi-long (5700-km) chain is made up mainly of tholeiitic lavas and ash tuff and pumice deposits. The lavas may have evolved from an initial submarine shield-building stage, through an explosive stage as they build up to sea level and finally to a subaerial posterosional stage. The largest volumes of lava formed during the shield-building stage, when extrusive processes built massive pillow and sheet lava flows, up to 3 mi (5 km) in height and 60 mi (100 km) in width, above the mean sea-floor depth. *See* LAVA; SEAMOUNT AND GUYOT.

Not all hot-spot volcanism is expressed in terms of highly lineated, multistage, volcanic chains. The Réunion hot spot, for example generated the Deccan Traps, which extend laterally for thousands of kilometers over the Indian subcontinent. Large basalt provinces have also been mapped at sea: the seaward-dipping reflector sequences of the Norwegian and Greenland passive margins and the deep-water volcanic sill complexes of the eastern Atlantic and Pacific oceans are examples. Aseismic ridges that extend up to or close to the axes of mid-oceanic ridges are another example of hot-spot volcanism. When a hot spot (for example, Iceland) is centered on the axis, pairs of ridges such as the Iceland-Faeroes Rise and the Greenland Rise are formed. Sometimes the plate (for example, Africa) has migrated off the hot spot (such as Tristan da Cunha), leaving behind ridge systems that no longer extend to the ridge axis (such as Rio Grande Rise and Western Walvis). *See* MID-OCEANIC RIDGE; VOLCANO.

Topographic swells. Another characteristic of hot spots is their association with broad swells in the Earth's topography. At Hawaii, for example, the swell reaches widths of 900 mi (1500 km) and heights of up to 0.7 mi (1.2 km) above normal depths of the sea floor (**Fig. 2**). The Hawaiian hot-spot swell is believed to have been formed in response to either thermal or dynamic effects in an underlying mantle plume. The crustal and upper-mantle structure, which is constrained by seismic refraction data, shows that the oceanic crust is of uniform thickness beneath the swell. The long-wavelength correlation that is observed between the gravity anomaly and the topography (about 37 mGal mi^{-1} or 22 mGal km^{-1}; Fig. 2) indicates that the mass excess of the swell is compensated by a low-density, high-temperature region below the crust. Swells of similar dimensions are found around Bermuda, Cape Verde, and Réunion.

Fig. 1. Global distribution of hot spots and hot-spot traces. The broken lines indicate gaps in the traces, usually because of movement of a ridge over a hot spot. Line AB indicates the approximate location of the profile shown in Fig. 2. (*After R. S. Detrick and S. T. Crough, Island subsidence, hot spots, and lithospheric thinning, J. Geophys. Res., 83: 1236–1244, 1978*)

Some swells (for example, French Polynesia) are very large, reaching more than 1200 mi (2000 km) across, and may involve more than one hot spot (for example, Tubai, Society, Pitcairn, and possibly the Marquesas). The lack of a suitable reference complicates the identification of swells in continental regions. Swells similar in size to Hawaii have, however, been described around the volcanic centers of Dafur, Hoggar, and Tibesti in northern Africa. Others have been described from elsewhere in Africa and North America, where they may have influenced the development of sedimentary basins.

The uplift of hot-spot swells is believed to result from thermal perturbations in the underlying plume. The excess heights of swells suggest, on isostatic grounds, that temperature differences of about 450°F (250°C) occur between the plume and the surrounding mantle. Theoretical studies suggest that plumes are relatively narrow (about 180 mi or 300 km) when they begin their ascent through the mantle, but they widen to about 600–1200 mi (1000–2000 km) as they spread out beneath the lithosphere (**Fig. 3**). Hot ascending plumes may raise the temperature of the overlying lithosphere, thereby thinning it. The mechanism of thinning is not well understood, however, and it may involve either conductive heating of preexisting lithosphere or the removal of lithospheric material by some form of mantle flow.

Models. Two classes of models have been proposed to explain hot-spot swells: the reheating and dynamic models. In the reheating model swell, uplift is produced by thermal expansion that is confined to the conducting portion of the lithosphere (the thermal boundary layer). In the dynamic model, however, there is a contribution to the uplift that is produced by vertical normal stresses exerted to the seismically defined base of the lithosphere (the mechanical boundary layer) by convection. Unfortunately, it is difficult to use surface observations to distinguish between these models. Geoid anomaly data suggest, for example, that the excess mass of swells is compensated by low-density, high-temperature material at depth. However, these data indicate only the depth of the greatest temperature differences. They are therefore indeterminate as to whether the temperature differences are confined to within the lithosphere as predicted by the reheating model, or whether they extend below as implied by the dynamic model.

The main distinguishing feature between the different uplift models is that the reheating model predicts a higher heat flow than the dynamic model. Discrimination between these models therefore depends on how the subsidence history, heat flow, and long-term strength (which is controlled mainly by the temperature) differ from those for unperturbed lithosphere of the same age. Many seamounts and oceanic islands show a subsidence history that follows closely what would be expected for the age of the lithosphere on which they were emplaced.

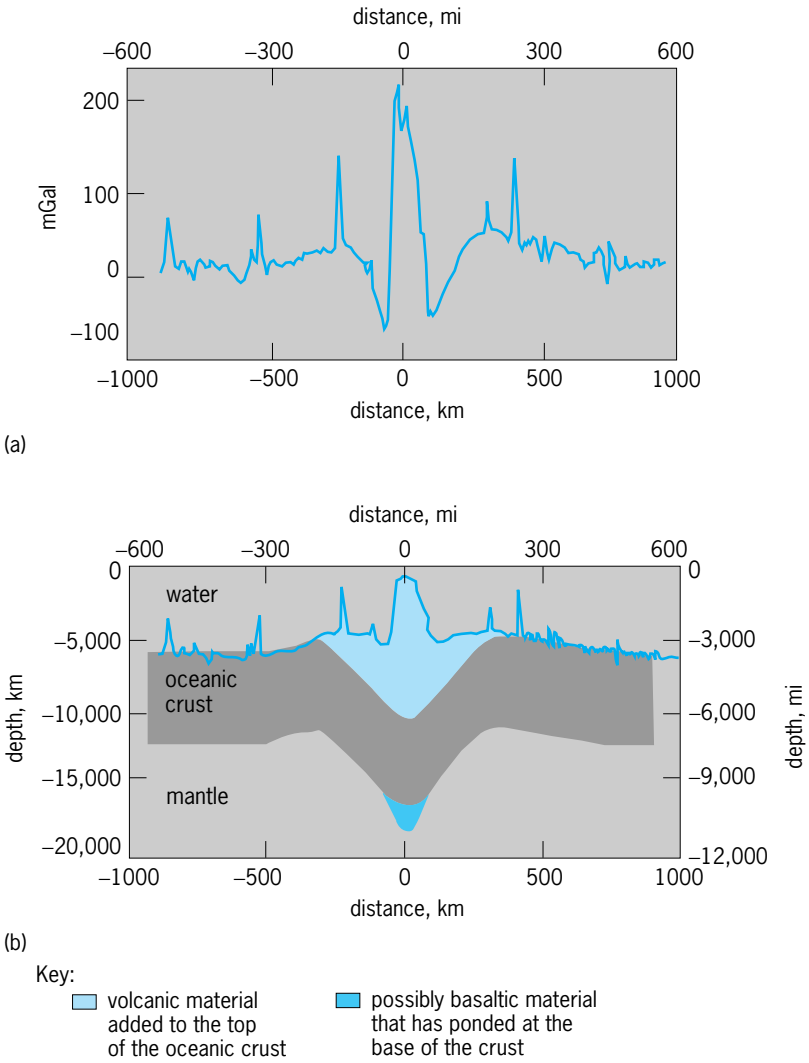

Fig. 2. Profile of the Hawaiian hot spot that crosses the Hawaiian-Emperor seamount chain in the region of Oahu. (*a*) Free-air gravity anomaly. (*b*) Crustal and upper-mantle model.

Others, however, show a much greater subsidence. The subsidence history of Eniwetok and Bikini atolls in the Western Pacific, for example, is more typical of 25-million-year-old oceanic lithosphere than the 90-million-year-old lithosphere on which they were emplaced. This is supportive of a reheating model.

Heat-flow and elastic-thickness data, however, are not so easily explained by the reheating model. At Hawaii, for example, the measured heat flow is not significantly higher over the swell than the surrounding sea floor. Furthermore, the elastic thickness [a measure of the long-term (greater than 1-million years) strength of the lithosphere] is close to what would be expected on the basis of the tectonic-plate age. One explanation is that the perturbed heat has been stored in the lithosphere but has moved away with it as the Pacific plate (which is moving at rates greater than 100 mm yr^{-1}) migrated over the Hawaiian plume. If this is so, the effects of reheating should be more apparent on slow-moving plates such as Africa, since there should been enough time for the plate to have been heated. The Cape Verdes, for example, have a relatively low elastic thickness (about 6 mi or 10 km less than expected), and high heat flow (about 20 mW m^2 more than expected). These data can be explained if the 125-million-year-old oceanic lithosphere that underlies the Cape Verdes has been thermally reset by a hot spot to a thermal age of about 60–80 million years. The corresponding change in the depth of the sea floor, however, amounts only to about 1650–2700 ft (500–820 m), which is significantly less than the observed swell height of 6900 ft (2100 m).

The reheating model is therefore unable to explain the full height of the Cape Verdes swell, suggesting that other factors, such as dynamic effects, may be involved in its support. Finally, some hot-spot swells (for example, French Polynesia) show both normal and low elastic-thickness values. These observations are difficult to explain by either of the uplift models, suggesting instead a high degree of volcano individuality and the possibility of local rather than regional controls on some of the geophysical parameters that are used to characterize hot spots.

Chemical characteristics of volcanoes. In contrast to the chemically depleted basalts of the mid-oceanic ridges, hot-spot volcanoes have a higher proportion of those elements that tend to enter the liquid phase first when the solid mantle is heated, as compared to mid-oceanic ridge magmas that have experienced the same degree of crystal fractionation. This observation suggests that mid-oceanic basalts are derived from a mantle source that has previously undergone a melt extraction event or that the hot-spot magma source has been enriched in these so-called incompatible elements by some other mechanism. A higher concentration of radiogenic elements in hot-spot volcanoes also points to a magma source that is chemically distinct from that of mid-oceanic ridges. The chemical variation within hot-spot volcanoes, which reflects the heterogeneous nature of mantle sources and the interaction of chemically enriched melts with depleted mantle and crustal-derived melts, has been described in terms of a number of end-member types, which include depleted mantle, enriched mantle, and material with a component possessing a high ratio of uranium-238 to lead-204 ($^{238}U/^{204}Pb$) known as HIMU. One of the best known of the enriched mantle components, known as DUPAL, occurs in the French Polynesia superswell region, suggesting a link between them. HIMU components are also found within the superswell region. Possible sources for these components include subcontinental lithosphere (DUPAL) and subducted oceanic lithosphere (HIMU) now residing at the lower-upper mantle or core-mantle boundary that has become entrained in the convecting upper mantle before being involved in melting. *See* EARTH INTERIOR.

Material additions. Although the temperature differences in plumes are relatively small and imply small amounts of melting, the fact that mantle material is continually being fed through the melting region of the plume means that substantial volumes of magma can be generated by hot spots. Seismic

refraction data, which indicate the amount of volcanic material that has been added to the top and bottom of the crust, indicate large variations in the production rates between hot-spot volcanoes. For example, rates that vary from 0.005 mi^3 (0.02 km^3) yr^{-1} for the Canary Islands to about 0.07–0.1 mi^3 (0.3–0.5 km^3) yr^{-1} for Réunion. The Canary Islands and Réunion were formed on old and young lithosphere, respectively, suggesting that the plume flux may depend on the thermal age of the overlying lithosphere: thick lithosphere inhibiting melt production, thin lithosphere encouraging it. One explanation of this observation is that plumes may be sites of decompression melting. If the lithosphere is thick, little melt is produced because insufficient decompression melting can occur. When the lithosphere is thin, a plume may develop beneath young oceanic lithosphere (that is, at a ridge crest), and then large quantities of melt can be produced. *See* MAGMA.

There is a close association between magmatism and extension in the Earth's crust, such as continental margins that formed as a result of continental break-up. Some magmatism is to be expected at continental margins because of decompression melting. However, some margins (for example, Norway, Greenland, and Britain) are associated with such large thicknesses of volcanic material (up to 6 mi or 10 km) that they require high temperatures and high upwelling rates in order to explain them. Because of their association with Iceland, it has been proposed that the volcanic material at these so-called volcanic margins was generated at a hot spot. *See* CONTINENTAL MARGIN.

An important problem is the relative timing of initial rifting and the volcanism. The earliest volcanism may occur as the plume is rising, and before rifting begins; but the volume of such volcanism is likely to be small. According to the decompression hypothesis, a bulk of the melting will follow initial rifting when the continental crust is thinned. Seafloor spreading would therefore be expected to follow volcanism. Indeed, the separation of the Seychelles and Saya da Malha Banks from India followed the formation of the Deccan Traps at the Réunion hot spot 60–65 million years ago, and the opening of the North Atlantic followed the formation of the Brito-Arctic volcanic province at the Iceland hot spot 55–65 million years ago. The impingement of rising plumes on the lithosphere did not, however, lead to continental rifting in all cases (for example, Siberian Traps). Also, in some (for example, Karoo) there has been a long delay before rifting begins, suggesting that other controls may, in some cases, be more important.

Role of mantle plumes. Ever since W. J. Morgan introduced the concept of hot spots, there has been considerable discussion on the role of mantle plumes in the driving mechanism of plate tectonics. Early discussions were concerned with how ascending plumes interacted with convective motions in the mantle, especially the large-scale flow between mid-oceanic ridges and subduction zones. They now focus on whether hot spots play an active or passive role in the generation of plate motions. The uplift of plume-generated midplate swells may be sufficient, for example, to contribute to the forces that drive plate motions. However, plumes are passive, as is seen by their tendency to become entrained at mid-oceanic ridges (for example, Iceland) and large-offset fracture zones (for example, Louisville). *See* SUBDUCTION ZONES.

A. B. Watts

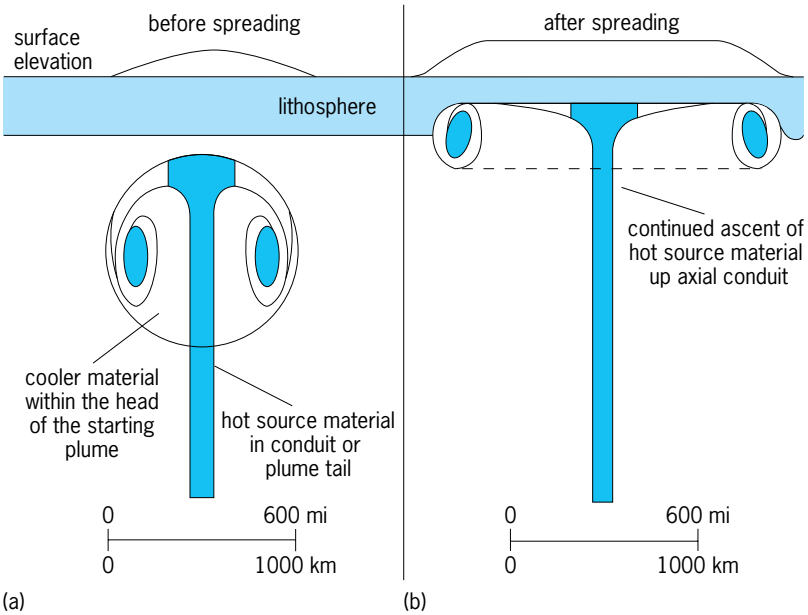

Fig. 3. Ascent of a plume (*a*) before and (*b*) after near-surface spreading. Darker areas indicate higher temperatures. The upper panels indicate the horizontal scale over which significant uplift would be expected to occur above such a plume. (*After R. I. Hill, Starting plumes and continental break-up, Earth Planet. Sci. Lett., 104: 398–416, 1993*)

Bibliography. S. T. Crough, Hotspot swells, *Annu. Rev. Earth Planet. Sci.,* 11:165–193, 1983; R. S. Detrick and S. T. Crough, Island subsidence, hot spots, and lithospheric thinning, *J. Geophys. Res.*, 83:1236–1244, 1978; R. I. Hill, Starting plumes and continental break-up, *Earth Planet, Sci. Lett.*, 104:398–416; N. H. Sleep, Hotspot volcanism and mantle plumes, *Annu. Rev. Earth Planet. Sci.*, 20:19–43, 1992; R. S. White and D. P. McKenzie, Magmatism at rift zones: The generation of volcanic continental margins and flood basalts, *J. Geophys. Ref.*, 94:7685–7729, 1989.

Hot-water heating system

A heating system for a building in which the heat-conveying medium is hot water. Heat transfer in British thermal units (Btu) equals pounds of water circulated times drop in temperature of water. For other liquids, the equation should be modified by specific heats. The system may be modified to provide cooling.

A hot-water heating system consists essentially of water-heating or -cooling means and of heat-emitting means such as radiators, convectors, baseboard radiators, or panel coils. A piping system connects the heat source to the various heat-emitting units and includes a method of establishing circulation of

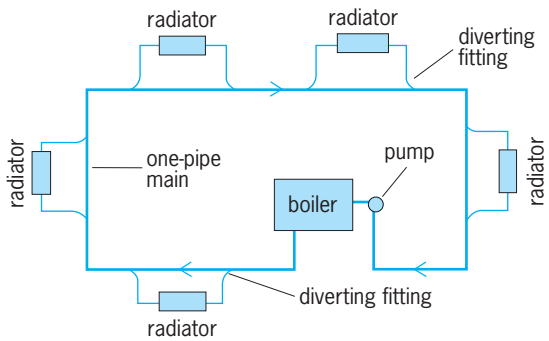

Fig. 1. One-pipe hot-water heating system.

the water or other medium and an expansion tank to hold the excess volume as it is heated and expands. Radiators and convectors have such different responses that they should not be used in the same system.

Types. In a one-pipe system (**Fig. 1**), radiation units are bypassed around a one-pipe loop. This type of system should only be used in small installations.

In a two-pipe system (**Fig. 2**), radiation units are connected to separate flow and return mains, which may run in parallel or preferably on a reverse return loop, with no limit on the size of the system.

In either type of system, circulation may be provided by gravity or pump. In gravity circulation each radiating unit establishes a feeble gravity circulation; hence such a system is slow to start, is unpredictable, and is not suitable for convectors, baseboard radiation, or panel coils because circulating head cannot be established, and circulation cannot be supplied to units below the mains. The pipes must be large. For these reasons gravity systems are no longer used.

In forced circulation a pump is used for motivation. Circulation is positive and units may be above or below the heat source. Smaller pipes are used.

Operation. For perfect operation it is imperative that the friction head from the heat source through each unit of radiation and back to the heat source be the same. This usually requires careful balancing after installation and during operation.

Expansion tanks are open or closed. Open tanks are vented to the atmosphere and are used where the water temperature does not exceed 220°F or 123°C (at sea level). They provide the safest operation, almost free from explosion hazards. Closed tanks, used for higher water temperatures, are provided with safety devices to avoid possible explosions.

One outstanding advantage of hot-water systems is the ability to vary the water temperature according to requirements imposed by outdoor weather conditions, with consequent savings in fuel. Radiation units may be above or below water heaters, and piping may run in any direction as long as air is eliminated. The system is practically indestructible. Flue-gas temperatures are low, resulting in fuel savings. The absence of myriads of special steam fittings, which are costly to purchase and to maintain, is also an important advantage. Hot water is admirably adapted to extensive central heating where high temperatures and high pressures are used and also to low-temperature panel-heating and -cooling systems.

Circulating hot-water pumps must be carefully specified and selected. On medium-size installations, it is recommended that two identical pumps be used, each capable of handling the entire load. The pumps operate alternately but never together in parallel. On large installations three or more pumps may be used in parallel, provided they are identical and produced by the same manufacturer. The casings and runners must be cast from the same molds, and the metals and other features that affect their temperature characteristics must be identical. All machined finishes must be identical, and the pumps must be thoroughly shop-tested to operate with identical characteristics.

When the system is in operation, the pump can be disconnected from the boiler by throttling down the valve at the boiler return inlet; it should not be closed completely. This permits the water in all boilers to be at the same temperature so that when a boiler is thrown back into service, the flue gases do not impinge on any cold surfaces, thus producing soot and smoke to further contaminate the outdoor atmosphere. *See* COMFORT HEATING; DISTRICT HEATING; OIL BURNER.

Erwin L. Weber

Bibliography. American Society of Heating, Refrigerating, and Air Conditioning Engineers, *Handbook*: *Applications*, 1993, *Equipment*, 1983, *Systems*, 1992, *Fundamentals*, 1994; J. E. Traister, *Residential Hot Water Systems: Repair and Maintenance,*1986.

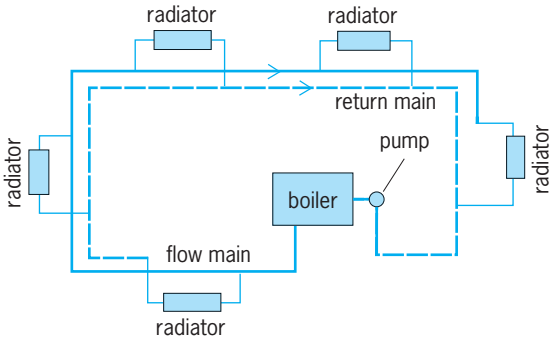

Fig. 2. Two-pipe reverse return system.

Hubble constant

A number that characterizes the expansion rate of the universe and is required to determine its age. In the standard big bang model, the local universe expands uniformly according to the Hubble law, $v = H_0 d$, where v is the velocity of a galaxy at a distance d, and H_0 is the Hubble constant. The wavelength of radiation is stretched due to the expansion of space so that the spectra of objects become progressively redder at greater distances. (For nearby objects, the observed redshift can be described as a Doppler effect.) The constant is named after Edwin P. Hubble, who discovered that the velocity of

recession of a galaxy is proportional to its distance. A reliable and accurate measurement of the Hubble constant, an independent estimate of the ages of the oldest objects in the universe, and a further measurement of the average mass-energy density in the universe are all separately required in order to test and ultimately provide strong constraints on cosmological models. Measuring an accurate value of H_0 was one of the motivating reasons for building the *Hubble Space Telescope*. *See* DOPPLER EFFECT; HUBBLE SPACE TELESCOPE; REDSHIFT.

Although measurement of the Hubble constant is extremely simple in principle, it is much more difficult in practice. For decades, it remained an outstanding problem in cosmology following Hubble's original discovery in the 1920s. There were two main reasons for the difficulty: First, measuring distances turned out to be immensely challenging. Second, while the velocities can be measured very simply and accurately (from measurements of the positions of spectral lines in galaxies), galaxies are known to interact gravitationally with their neighbors. In so doing, their velocities are perturbed (inducing so-called peculiar motions superimposed on the general expansion). Hence, while an accurate measurement of the Hubble constant requires that an accurate extragalactic distance scale be established, this already-difficult task must be done at distances great enough that peculiar motions of galaxies are small compared with the overall cosmic expansion velocity, the Hubble flow. Measurements from two major efforts, the Hubble Space Telescope Key Project and the *Wilkinson Microwave Anisotropy Probe* (*WMAP*), have led to a convergence on the value of this important parameter. *See* ASTRONOMICAL SPECTROSCOPY.

Distances to galaxies and measurement. Astronomy is unique among the physical sciences in that most length scales cannot be measured directly. The size scales, especially in a cosmological context, are too vast. In general, the basis for estimating distances in astronomy is the inverse-square radiation law. If objects can be identified whose luminosities are either constant (standard candles), or perhaps related to a quantity that is independent of distance (for example, period of oscillation, rotation rate, velocity dispersion, or color), then, given an absolute calibration, their distances can be gauged. The standard candles must be independently calibrated (to absolute physical units) so that true distances (in meters or megaparsecs; 1 Mpc $= 3.08 \times 10^{22}$ m $= 3.26 \times 10^6$ light-years) can be determined.

Cepheid variables. Primary among the distance indicators are the Cepheid variables, stars whose outer atmospheres pulsate regularly with periods ranging from 2 to about 100 days. Empirically it has been established that the period of pulsation (a quantity independent of distance) is very well correlated with the intrinsic luminosity of the star. High resolution is the key to discovering Cepheids in other galaxies. The Cepheids must be identified against the background of fainter, resolved and unresolved stars that contribute to the light of the galaxy. From the Earth, turbulence in the atmosphere degrades the image resolution, smearing the light from the stars of interest and decreasing their contrast against the background. The resolution of the *Hubble Space Telescope* is about 10 times better than can be generally obtained through the Earth's atmosphere, and, moreover, it is stable. As a result, the volume of space made accessible by the *Hubble* increased by a factor of 1000. *See* CEPHEIDS.

Other distance indicators. The reach of Cepheid variables as distance indicators is limited, however, even with the *Hubble Space Telescope*. For distances beyond 20 Mpc or so, brighter objects than ordinary stars are required, for example, luminous supernovae or the luminosities of entire galaxies. The principle for measuring all of these distances is the same. Whether it is the rotation speed or velocity dispersion of a galaxy, or the rate at which a supernova fades, all of these distance-independent quantities are indicators of intrinsic luminosity. Given the apparent luminosity (corrected for intervening extinction by dust), a simple application of the inverse-square law then yields a distance. The absolute calibration for all of these methods is presently established using the Cepheid distance scale.

One of the most promising cosmological distance indicators is the luminous supernovae classified as type Ia. These objects have luminosities comparable to entire galaxies of moderate luminosity, and hence can be observed to distances of hundreds of megaparsecs. Unfortunately, the exact mechanism for the ignition of the explosion has not yet been theoretically or observationally established, nor are the progenitors known with any certainty. The basis for most of the relative distance indicators remains empirical. Ultimately, the confidence in these empirically based methods will be strengthened as the theoretical basis is more firmly established. *See* SUPERNOVA.

For spiral galaxies, the total luminosity shows an excellent, empirical correlation with the maximum rotation velocity of the galaxy (the Tully-Fisher relation). Independent of distance, galaxy rotation rates can be measured spectroscopically (from Doppler shifts of spectral features of hydrogen at radio or optical wavelengths). This relation has been measured for hundreds of galaxies within clusters, and in the general field.

For elliptical galaxies, a correlation exists between the stellar velocity dispersion and the intrinsic luminosity, analogous to the relation between rotation velocity and luminosity for spirals. Moreover, the surface brightness of an elliptical galaxy inside a given radius is tightly correlated with the velocity dispersion of the galaxy. Another method useful for measuring distances to elliptical galaxies makes use of the fact that the resolution of stars within galaxies is distance-dependent. *See* GALAXY, EXTERNAL.

Key project. The examples described above provide a means of measuring relative distances to galaxies, while the absolute calibration for these methods relies on the Cepheid distance scale. A key project of the *Hubble Space Telescope* has provided Cepheid distances to a sample of 18 galaxies useful for setting

the absolute distance scale for galaxies, in addition to other methods. All of these methods can be applied at distances where the peculiar motions of galaxies contribute less than 10% of the overall cosmic expansion velocity.

A controversy existed for many decades over the value of the Hubble constant, with published values disagreeing at times by a factor of 2. However, the Cepheid distances from the Hubble Space Telescope Key Project have provided a means of calibrating and comparing a number of relative distance methods; and for the first time, to within an uncertainty of ±10%, all of these methods are consistent with a value of the Hubble constant of 72 kilometers per second per megaparsec.

Microwave anisotropy measurements. The detection in 1965 by Arno Penzias and Robert Wilson of the cosmic background radiation provides one of the strongest foundations for big bang cosmology. In 2002, the *WMAP* satellite measured small fluctuations in the microwave sky to a few parts in 100,000. Encoded within these fluctuations is information about the geometry, the Hubble constant, as well as the matter-plus-energy density of the universe. By themselves, these data provide only a weak limit on the Hubble constant. But combined with measurements of the large-scale distribution of galaxies, the *WMAP* results yield a value of the Hubble constant of 71 kilometers per second per megaparsec, with an uncertainty of only 5%. This value is in excellent agreement with that from the Hubble Key Project, and is a completely independent determination based on very different underlying physics. *See* COSMIC BACKGROUND RADIATION.

Age of the universe. The Hubble constant sets the expansion time scale for the universe. In order to measure the time since the big bang, it is necessary to determine the expansion rate and the matter-plus-energy density of the universe. Increasing evidence suggests that the total matter density of the universe is only about 30% of the total mass-energy density of the universe. The remaining 70% appears to be in a mysterious form with a repulsive pressure, which causes an acceleration of the universe (and results in an older universe). Measurements of type Ia supernovae provide evidence for an acceleration of the universe, consistent with a dark-energy component. This phenomenon is also consistent with the *WMAP* observations. *See* DARK ENERGY.

With a Hubble constant of 72 kilometers per second per megaparsec, and a universe composed of one-third matter density and two-thirds dark energy, the expansion age is calculated to be 13.5 billion years. The expansion age of the universe can be compared to the ages of the oldest stars within the Milky Way Galaxy, those in globular clusters. Results from the *Hipparcos* satellite, applied to models of stellar evolution, yield ages of about 12 billion years with an uncertainty of about ±20%. Given the uncertainties in the two independent measurements of the time scale (one based on the theory of the evolution of stars and the other based on the dynamics of the big bang), there is a remarkable agreement. *See* ASTROMETRY; BIG BANG THEORY; COSMOLOGY; STAR CLUSTERS; STELLAR EVOLUTION; UNIVERSE.

Wendy L. Freedman

Bibliography. J. D. Barrow, *The Origin of the Universe*, Basic Books, 1994; K. Ferguson, *Measuring the Universe*, Walker, New York, 1999; W. Freedman, The expansion rate and size of the universe, *Sci. Amer. Quart.*, 1:92–97, Spring 1998; W. Freedman and M. Turner, Cosmology in the new millennium, *Sky Telesc.*, 106(4):31–41, October 2003.

Hubble Space Telescope

The *Hubble Space Telescope* is the largest visible-light observatory ever placed into space. *Hubble*'s orbit, some 612 km (380 mi) above Earth's surface (**Fig. 1**), keeps it above almost all of Earth's atmosphere, at a location where its view of the heavens is much clearer than that of ground-based telescopes. The superior view afforded by *Hubble*'s orbit has made the telescope a unique resource for astronomers worldwide and has led to fundamental discoveries about the size and age of the universe, the birth and death of stars, and the development of galaxies.

The spacecraft. The *Hubble Space Telescope* is 13.2 m (43.5 ft) long, is 4.2 m (14 ft) in diameter at its widest, weighs 11,110 kg (24,500 lb), and orbits Earth once every 96 minutes. At its core is a reflecting telescope with a primary mirror 2.4 m (94.5 in.) in diameter (**Fig. 2**). The primary mirror directs light from astronomical objects to a 30-cm (12-in.) secondary mirror, which then bounces it back through a hole at the center of the primary mirror to the scientific instruments. The optical design is a Ritchey-Chrétien variant of a Cassegrain telescope.

Hubble's optical system enables the telescope to record astronomical images with unprecedented precision in the optical, ultraviolet, and infrared spectral bands. In order to take full advantage of the clearer view above Earth's atmosphere, *Hubble*'s mirrors had to be polished until they were extremely smooth: The largest bumps on *Hubble*'s primary mirror are analogous to the height of a baseball on a surface as wide as the continental United States. A flaw in the overall shape of the primary mirror hampered observations for several years after launch. However, because the primary mirror was so smooth, corrective optics installed in 1993 were able to realize *Hubble*'s expected performance (**Fig. 3**). Its angular resolution of 0.05 arcsecond at optical wavelengths is equivalent to being able to distinguish two fireflies 1 m (3 ft) apart at a distance of 5000 km (3000 mi). *See* INFRARED ASTRONOMY; TELESCOPE; ULTRAVIOLET ASTRONOMY.

Recording sharp images during exposure times that can approach 1 hour requires very precise and stable pointing of the telescope. Guide stars are used to keep *Hubble* locked on target. Each observation typically relies on two guide stars that *Hubble*'s guiding system keeps fixed at preselected locations in

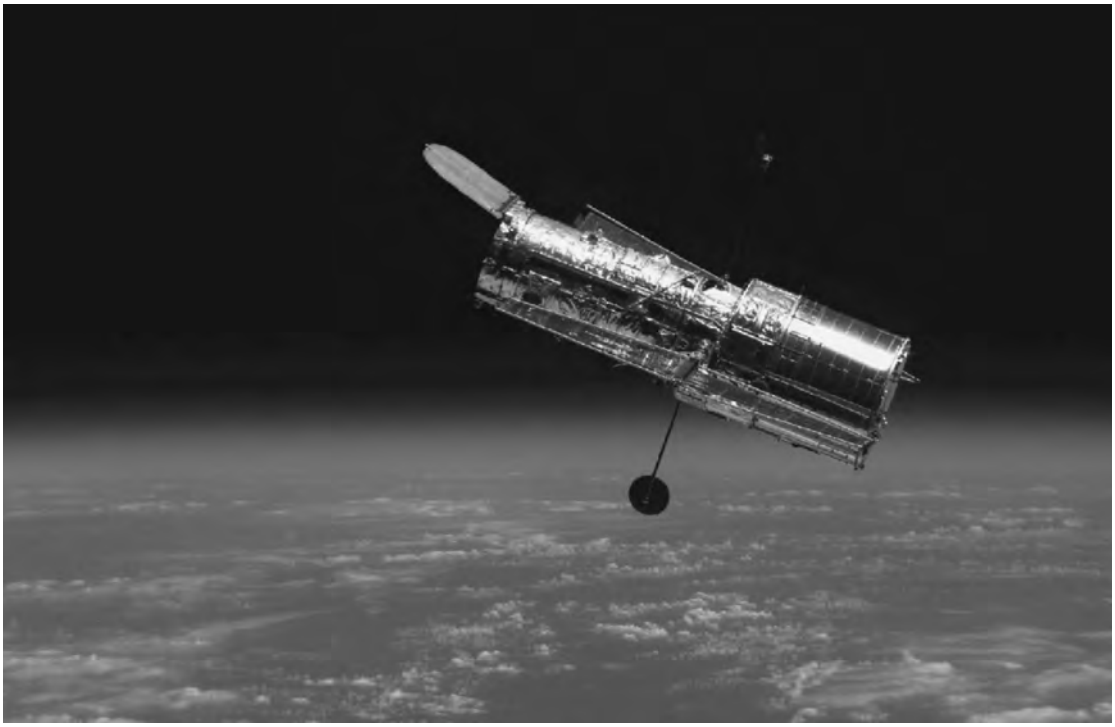

Fig. 1. ***Hubble Space Telescope*** **in orbit around Earth.**

two of its three Fine Guidance Sensors. Whenever *Hubble* drifts slightly off target, the guiding system returns these stars to their preselected locations in the telescope's field of view, placing the telescope back on target. This system keeps *Hubble* pointed to within 0.007 arcsecond, equivalent to keeping a laser in New York locked on a small coin in Washington, DC.

Hubble's complement of scientific instruments handles a wide range of observational tasks. Its cameras have recorded images of astronomical objects at wavelengths ranging from 115 to 2500 nanometers. *Hubble*'s spectrographs have analyzed the spectra of these objects between wavelengths of 115 and 1030 nm. Because these instruments can be removed and replaced, they have been upgraded several times during *Hubble*'s stay in orbit (see **table**). A diverse assortment of support equipment surrounds the telescope and its scientific instruments. To supply power, the spacecraft has two solar arrays and a set of storage batteries that keep the spacecraft operating when it passes through Earth's shadow. To rotate the telescope toward different spots on the sky, *Hubble* has four reaction wheels; increasing their spin rate in one direction causes the spacecraft to rotate in the opposite direction. To store and communicate the data gathered, *Hubble* has a set of solid-state data recorders and communications antennas that link to NASA's Tracking and Data Relay Satellite System (TDRSS). Governing the whole spacecraft is an onboard computer that interprets and executes the instructions relayed from the ground.

History. Two disadvantages of ground-based observing were well known to early-twentieth- century astronomers: (1) Because of the turbulent motions of Earth's atmospheric gases, the paths of light rays passing through the atmosphere are constantly shifting, distorting our view of astronomical objects. To human eyes these distortions are rather subtle; for example, they are responsible for the twinkling of stars. However, this effect limits the sharpness of most images taken with ground-based telescopes to a resolution of no better than 1 arcsecond (Fig. 3). (2) The Earth's atmosphere is quite transparent to visible light but blocks much of the infrared and ultraviolet light from the cosmos. Both of these wavelengths are scientifically important. *See* TWINKLING STARS.

In the 1920s, the decade in which the American scientist Robert Goddard launched the first liquid-fueled rockets, the German scientist Hermann Oberth published the first serious papers describing the advantages of a space-based telescope. Development of the V2 rocket in Germany during World War II made Oberth's speculations seem much more realistic. After the war, in 1946 the American astronomer Lyman Spitzer produced a detailed report for the RAND Corporation on what a large space-based telescope might accomplish. This report is widely regarded as the birth of the Hubble Space Telescope.

Even though the rationale for a space-based telescope was clear in 1946, the technology was not yet ready. Launch vehicles capable of placing a large telescope into orbit were not developed until the 1960s, when the planning of the *Hubble Space Telescope* began in earnest. Several smaller precursors, such as the *Orbiting Astronomical Observatory* (*OAO*) and the *International Ultraviolet Explorer* (*IUE*), preceded *Hubble* into space, demonstrating both

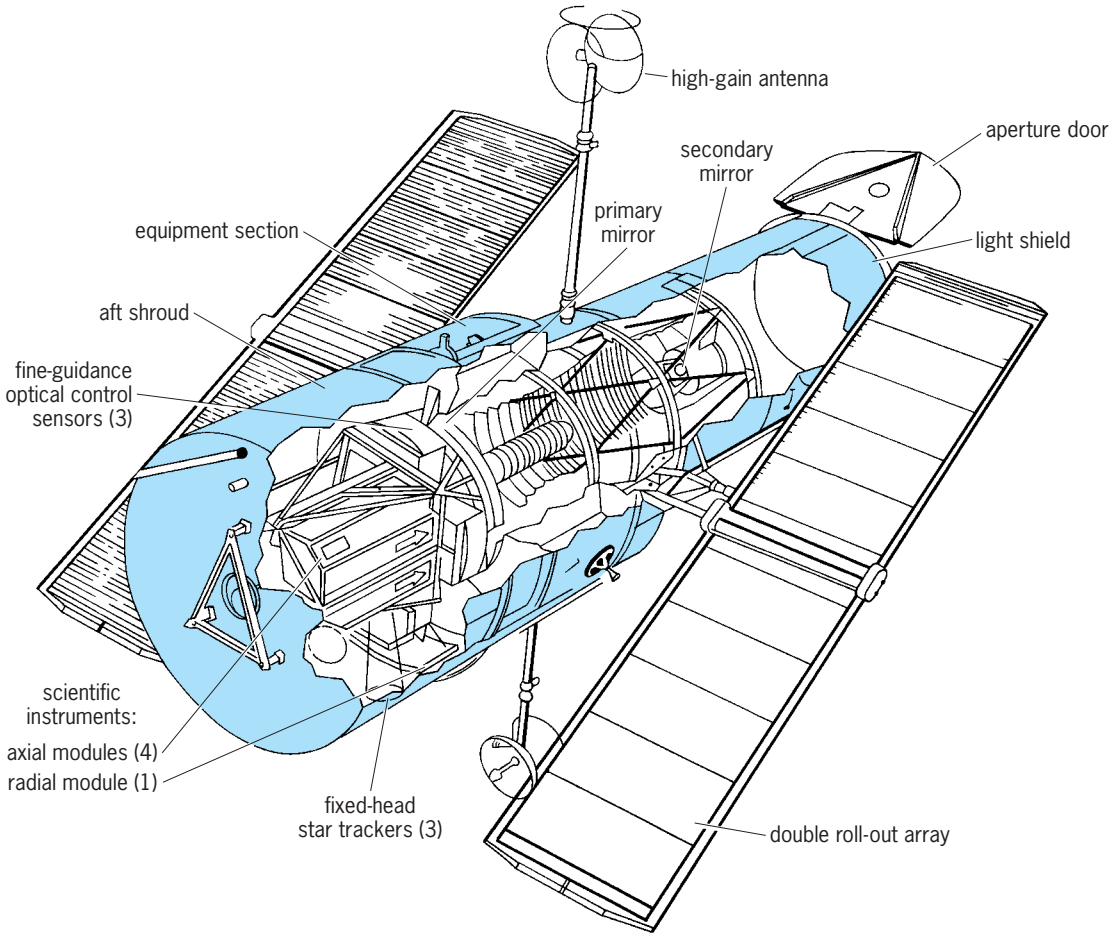

Fig. 2. Diagram of the *Hubble Space Telescope*.

the feasibility and the rewards of astronomical observing from above Earth's atmosphere. Meanwhile, NASA and the astronomy community worked to build political support for a much larger, much more expensive telescope in space. *See* SATELLITE (ASTRONOMY).

Congress approved the so-called Space Telescope in 1977. Lockheed Corporation was chosen to build the spacecraft, and Perkin-Elmer Corporation to grind and polish the primary mirror. The spacecraft, renamed the *Hubble Space Telescope* in 1983 in honor of American astronomer Edwin Powell *Hubble*, was ready for launch in 1986, but the tragedy of the *Challenger* space shuttle explosion delayed *Hubble*'s launch until 1990.

The space shuttle *Discovery* finally carried *Hubble* aloft on April 24, 1990, forty-four years after Spitzer's seminal report. The shuttle's robotic arm released the spacecraft into orbit the following day. However, the ultrasharp pictures expected from *Hubble* did not begin to arrive until $3^1/_2$ years later.

Shortly after launch, astronomers realized that *Hubble* could not be properly focused. The source of the problem was a flaw in the primary mirror which

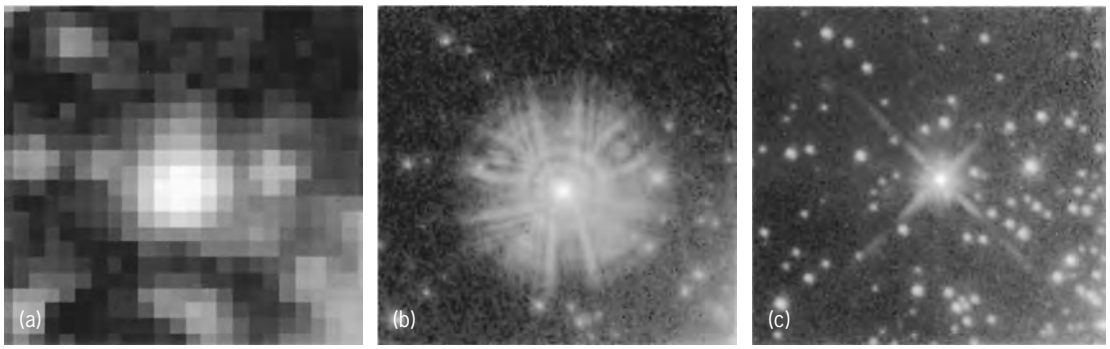

Fig. 3. Comparison of images of a field of stars in the 30 Doradus Nebula. (*a*) Ground-based image obtained under conditions of good seeing. (*b*) Same star field at the same scale taken with the *Hubble Space Telescope* before the first servicing mission. (*c*) *Hubble* image after the first servicing mission when the spherical aberration was corrected. (*NASA*)

History of *Hubble*'s instruments*

	Instruments	Dates
Original instruments	Wide Field/Planetary Camera (WF/PC), *Hubble*'s original visible/ultraviolet light camera	1990–1993
	Faint Object Camera (FOC), *Hubble*'s highest-resolution camera	1990–2002
	Faint Object Spectrograph (FOS), for analyzing light from faint objects	1990–1997
	Goddard High Resolution Spectrograph (GHRS), designed to perform detailed analyses of spectra	1990–1997
	High Speed Photometer (HSP), for measuring rapid variations in the brightness of astronomical objects	1990–1993
Instruments from first servicing mission	Wide Field Planetary Camera 2 (WFPC2), a visible/ultraviolet-light camera with corrective optics to compensate for the flaw in the primary mirror	1993–present
	Corrective Optics Space Telescope Axial Replacement (COSTAR), a device that placed corrective optics over *Hubble*'s original instruments to compensate for the flawed primary mirror	1993–present (no longer in use; current instruments designed with corrected optics)
Instruments from second servicing mission	Near Infrared Camera and Multi-Object Spectrometer (NICMOS), *Hubble*'s primary camera for observing infrared light	1997–present (inoperable 1999–2002)
	Space Telescope Imaging Spectrograph (STIS), a much more efficient optical/UV spectrograph than the FOS or GHRS; can also be used as an optical/UV camera	1997–present (inoperable since 2004)
Instrument from fourth servicing mission	Advanced Camera for Surveys (ACS), an optical-light camera much more sensitive than WFPC2 and with a wider field of view	2002–present
Planned instruments	Cosmic Origins Spectrograph (COS), an extremely sensitive UV spectrograph	Awaiting installation
	Wide Field Camera 3 (WFC3), an advanced optical-infrared camera	Awaiting installation

*As of 2006.

caused an effect known as spherical aberration. *Hubble*'s mirror was exquisitely polished, but its overall shape was incorrect.

Fortunately, NASA had designed *Hubble* for periodic servicing and upgrades (see table). During the first servicing mission in December 1993, astronauts were able to install a new camera (WFPC2) with optics that corrected for the flaw in *Hubble*'s mirror, as well as a device (COSTAR) that placed corrective optics in front of *Hubble*'s original instruments. That mission also installed new solar arrays on *Hubble*, eliminating some troublesome vibrations of the spacecraft produced by the original arrays. With these improvements, *Hubble*'s performance surpassed the original specifications, enabling *Hubble* to fulfill its scientific promise (Fig. 3).

Subsequent servicing missions performed additional upgrades. The second servicing mission, in February 1997, installed two new instruments: STIS, a vastly improved spectrograph, and NICMOS, *Hubble*'s primary infrared camera. The third mission, in December 1999, replaced four of *Hubble*'s guiding gyroscopes and fixed some of *Hubble*'s multilayer insulation. The fourth mission, in March 2002, installed a state-of-the-art visible-light camera, the Advanced Camera for Surveys (ACS), a new set of solar arrays, and a cooling system that revived the NICMOS camera, which had been inoperable since 1999.

A fifth servicing mission, originally scheduled for 2005, has been delayed by the loss of the shuttle *Columbia* and ongoing problems with the Space Shuttle program. The mission will install a more sensitive spectrograph (COS) and an upgraded optical/infrared camera (WFC3), along with new batteries and gyroscopes that are essential for extending *Hubble*'s useful lifetime. Each month that passes without an upgrade now places the *Hubble* spacecraft in greater danger of an equipment failure that would prematurely end its mission.

Highlights of Hubble science. The *Hubble Space Telescope*'s contributions to astronomy are numerous and wide-ranging. The following are a few of the most notable.

Measuring the size and age of the universe. The primary method for measuring distances to other galaxies is to measure the brightness of stars whose intrinsic light outputs are known by other means. A class of stars known as Cepheid variables is particularly useful because the period over which they vary in brightness is directly related to their total light output. By comparing the apparent brightness of these stars to the total light output inferred from their brightness variations, astronomers can calculate their distance. *Hubble*'s high resolution has enabled the study of individual Cepheid variable stars in distant galaxies and the measurement of distances to galaxies up to 100 million light-years away. From the distances to these galaxies and the speeds at which they are moving away from each other, astronomers can calculate how long it took for galaxies to reach their current positions. Current estimates put this amount of time—the age of the universe—about 13.5 billion years. *See* CEPHEIDS; COSMOLOGY; HUBBLE CONSTANT.

Observations of young galaxies. *Hubble* has supplied the first clear pictures of what the universe was like when it was only 1 or 2 billion years old. Images

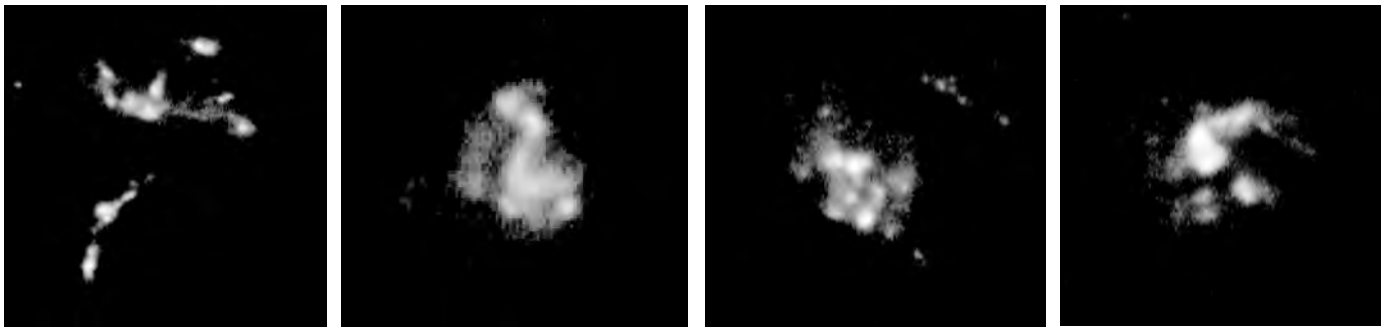

Fig. 4. Distant galaxies from the Hubble Ultra Deep Field. The distorted appearances of these galaxies suggest that they collided with other galaxies shortly before their light began the long journey to Earth.

such as the Hubble Deep Field and Hubble Ultra Deep Field reveal a multitude of galaxies at a wide range of distances, some of them over 10 billion light-years away. Light collected from these most distant galaxies took over 10 billion years to reach Earth. Thus, the images of these galaxies show how they looked over 10 billion years ago. Many of them appear quite strange and distorted in comparison to present-day galaxies, leading astronomers to believe that disruptive collisions between galaxies were much more common early in time than they are now (**Fig. 4**). *See* GALAXY, EXTERNAL.

Disks around young stars. The motions of planets in the solar system, which generally follow roughly circular paths sharing the same orbital plane, have long led astronomers to speculate that the solar system formed from a disk-shaped collection of matter with the Sun at its center. *Hubble* has helped to verify this suspicion through its observations of young stars in the process of formation. All stars, including the Sun, were born in clouds of interstellar hydrogen gas. Observations of clouds in which stars are currently forming show that brand-new stars are generally surrounded by disks of gas and tiny, solid particles called dust grains. The gas and dust in at least some of these disks are eventually expected to clump into planets similar to those that orbit the Sun. *See* SOLAR SYSTEM; STAR; STELLAR EVOLUTION. Mark Voit

Bibliography. E. J. Chaisson, *The Hubble Wars*, 1994; C. C. Petersen and J. C. Brandt, *Hubble Vision: Further Adventures with the Hubble Space Telescope*, 1998; R. W. Smith, *The Space Telescope*, 1989; R. W. Smith and D. DeVorkin, *Hubble Space Telescope: Imaging the Universe*, 2004; M. Voit, *Hubble Space Telescope: New Views of the Universe*, 2000.

Hudson Bay

A horseshoe-shaped bay, approximately 1050 km (650 mi) wide and 1370 km (850 mi) long, stretching from the Canadian provinces of Quebec and Ontario in the south to the Northwest Territories. With a surface area of approximately 827,000 km^2 (320,000 mi^2), Hudson Bay ranks twelfth in surface area among the Earth's seas and oceans. The bay is connected to the Arctic Ocean by the Foxe Channel and Roes Welcome Sound and to the Atlantic Ocean by the Hudson Strait (see **illus.**). Several islands are located within the bay, the largest of which is Southampton (41,212 km^2; 15,912 mi^2). *See* ATLANTIC OCEAN; NORTH AMERICA.

Early exploration of Hudson Bay was prompted by interest in finding an Arctic shortcut through North America to Asia. The bay was discovered by English navigator Henry Hudson. In 1610, Hudson directed his ship, the *Hopewell*, on a 3-month investigation of the bay's eastern shore and islands. Another early explorer, Thomas Button, reached the site of present-day Churchill on the bay's western shore in 1612. The following summer, Button passed Southampton Island on his way to the Hudson Strait. Others who explored the bay include William Baffin in 1615, and Luke Fox and Thomas James in 1631.

Geologic history. Hudson Bay is located within a depression of the Canadian Shield and is underlain with Precambrian rocks more than 500 million years old. The shield covers 4.8 million square kilometers (1.8 million square miles) in the United States and Canada and encompasses portions of Wisconsin, Minnesota, New York, Michigan, the Northwestern Territories, Saskatchewan, Ontario, Quebec, Baffin Island, and Labrador. The shield is the oldest region within the North American crustal plate, and is considered a craton—an area of the Earth's crust that has been stable for millions of years. Bordering the bay are gently dipping Paleozoic limestones, sandstones, and dolomite rocks. *See* CRATON; PALEOZOIC; PRECAMBRIAN.

The bay was formed 2.4 million years ago during the Pleistocene geologic epoch. Centered on what is now Hudson Bay, the Laurentide ice sheet stripped away soil and deposited rocks and sediments as it moved across the landscape. The enormous weight of the accumulated ice caused areas of the Earth's crust to compress and sink. Melting and ice retreat began about 13,000 years ago during the Wisconsin glacial period. As the ice melted, relatively warm seawater invaded the Hudson Bay through the Hudson Strait, causing additional shrinking of the ice mass. Eventually, the ice sheet separated into the Keewatin and Labrador ice centers, which disappeared

completely between 6500 and 5000 years ago. As the ice sheet melted, the crust began to slowly rebound due to isostatic adjustment. This process is not yet complete, as the area surrounding the bay continues to rise about 0.6 m (2 ft) each century. Uplifting of the land surface is most obvious along portions of the coast where lines representing former beaches run parallel to the shore. A rapid rate of uplift suggests that the Hudson Bay will become shallower over time and may disappear completely when isostatic equilibrium is reached. *See* GLACIAL EPOCH; ISOSTACY; PLEISTOCENE.

Physical geography. Hudson Bay's underwater physiography is represented by broad contours that are mostly concentric with its periphery. The floor of the bay is predominantly smooth but has been incised in some places. Its average depth is 128 m (420 ft) with a deepest known point of 183 m (600 ft). The basin containing the bay has a north-south running elongation, a reflection of its bedrock structure. Marine currents are the most important agent of sediment transport in the bay. Snow and ice, covering the bay for several months of the year, inhibits sediment discharge from streams. Sediment, coarse gravel, and even large boulders are moved away from shore in a process known as ice rafting. *See* BASIN.

Shoreline areas within the northern reaches of the bay are low in elevation, rocky, and indented with numerous small islands and inlets. Some areas surrounding the bay are characterized by broad, flat plains. To the north of Southampton Island and Quebec's Ungava Peninsula are areas with elevations exceeding 305 m (1000 ft). Hudson Bay's drainage basin is nearly 4 million square kilometers (1.5 million square miles). On the west and southwest shorelines are extensive areas of drowned swampland. Deltas, estuaries, and tidal flats are common through the area. The tundra surrounding much of the bay is a cold desert (moisture is scarce). Tundra plant cover includes lichens, grasses, and scattered shrubs. Plants must complete their annual cycles during brief summers in soils that are poorly drained as a result of the underlying permafrost. Soils on land areas surrounding the bay include infertile entisols, inceptisols, histosols, and spodosols. *See* SOIL.

Water properties and biology. The waters found within Hudson Bay are uniform in temperature, averaging near freezing with temperatures slightly cooler near the center of the bay in deeper water. The general pattern of water circulation is counterclockwise. River outflow inhibits Atlantic waters from entering the Hudson Strait, lowering the salinity of the bay water, especially in the spring and summer months. Tides range from just less than 1 m (3 ft) to greater than 4 m (13 ft) in western portions of the bay. Pack ice covers the open water from October to June.

Hudson Bay is connected to the Atlantic Ocean by the Hudson Strait. Compared to the bay, the strait is predominantly deep, in some places extending to 500 m (1640 ft). The strait forms a zone where water from the West Greenland Current mixes with water from the bay, creating an environment that supports plankton and diverse fish species.

Map of the Hudson Bay.

Climate. The climate of Hudson Bay is controlled by cold, dry, and stable continental polar and Arctic air masses. During much of the year, the bay is covered by snow and ice. On the land area surrounding the bay in January, the temperatures range from −30 to −20°C (−22 to −4°F), while in July the temperatures range from 10 to 20°C (50 to 68°F). The coldest land areas adjacent to the bay are located on its northwest side, extending from Manitoba's border with the Northwest Territories to Southampton Island. Rainfall is highest during summer months.

Thomas A. Wikle

Bibliography. A. W. Bally and A. R. Palmer (eds.), *The Geology of North America: An Overview*, Geological Society of America, Boulder, 1989; J. F. Lewry and M. R. Stauffer, *The Early Proterozoic Trans-Hudson Orogen of North America, Newfoundland*, Geological Society of Canada, 1990; B. MacLean (ed.), *Marine Geology of Hudson Strait and Ungava Bay, Eastern Arctic Canada: Late Quaternary Sediments, Depositional Environments, and Late Glacial–Deglacial History Derived from Marine and Terrestrial Studies*, Geol. Surv. Canada Bull. 566, 2001.

Huebnerite

A mineral with the chemical composition $MnWO_4$. Huebnerite is the manganese member of the wolframite solid-solution series. It commonly contains small amounts of iron. It occurs in monoclinic, short, prismatic crystals. Fracture is uneven. Luster varies from adamantine to resinous (see **illus.**). Hardness is 4 on Mohs scale and specific gravity is 7.2. Huebnerite is transparent and yellowish to reddish-brown in color; streak is brown. It is fusible with difficulty. *See* WOLFRAMITE.

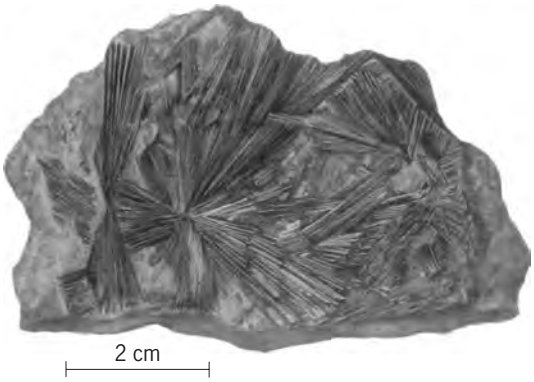

Radiating groups of huebnerite crystals in quartz vein, Silverton, Colorado. (*Specimen from Department of Geology, Bryn Mawr College*)

Edward C. T. Chao

Human biological variation

Anthropologists study human biology to better understand the extent of human biological variability, to explain the mechanisms that create and pattern this variability, to relate variability to health and disease, and to understand the sociocultural factors that interact with, and are influenced by, our biology. The major areas of study of human biological variation are growth and development, genetic variation, variation related to climate, infectious and noninfectious diseases, and demography. *See* ANTHROPOLOGY.

Biological versus biocultural variation. For a species with worldwide distribution, humans are remarkably similar in biology, behavior, emotions, and cognition. All humans are bipedal, have a relatively narrow range of body size and physical features (compared with other primate species), require the same essential nutrients, can eat a wide range of foods, depend on complex technology, organize their social lives via formal kinship relations and social networks, use symbolic language, and have religion. To account for the fundamental unity of the human species, anthropologists adopt an evolutionary perspective to explain variation through time and across space and a biocultural perspective that considers the interactive effects between human culture and biology.

Current evidence from the fossil record and genetic linkages among living people indicate that modern humans, *Homo sapiens sapiens*, had their origin in Africa about 125,000 years ago. Some of those African people migrated to the Middle East, Asia, Europe, Australia, and eventually the Americas. Modern humans carried with them a sophisticated array of technology, social organization, and ideological beliefs and practices. In other words, they had culture, and they used their cultural skills, as well as their biology and technology, to adapt to and exploit the environments of new lands as they migrated. The original human pattern of biology and culture was naturally selected because it conferred greater reproductive success on our ancestors. Biocultural variations that developed in the past 125,000 years allowed for continued adjustments to local climate, food, and lifestyle and continued reproductive success. *See* FOSSIL HUMANS.

History of human variation study. Until the mid-twentieth century, human variation and its biocultural nature were scarcely appreciated. From the time of Plato and Aristotle, living people and their cultures were considered to be imperfect copies, more or less, of an ideal type of human being and culture. In 1758, Carl von Linné (Linnaeus) formalized a system of human types in the 10th edition of his *Systema Naturae*. This book provided European science with the system of nomenclature for living things still used today. Linnaeus divided the human species into four types based on geography, skin color, temperament, clothing style, and other traits of material culture. In essence these four groups were Americans, Asians, Africans, and Europeans. A sense of superiority of Europeans to other "types" was inherent in this grouping and was elaborated into the "race science" of the nineteenth and early twentieth centuries. The typological approach of Linnaeus and the "race science" that followed ignored the true nature of human variation. "Races," it was believed, were fixed types in terms of size, shape, and other physical and behavioral traits. Moreover, these types were believed to be immutable, that is, not responsive to changes in the environment. *See* ANIMAL SYSTEMATICS; SPECIES CONCEPT.

A countercurrent of thought was evident in the late nineteenth century as migration of people from Europe to the United States was shown to alter the physical features of the migrants. At the forefront of the antiracial movement was Franz Boas, the founder of American anthropology. In a series of studies based on large samples of migrants to New York City, Boas demonstrated that the migrants changed their body size and shape, becoming taller and longer-legged, with narrower skulls, in response to the new environment. He referred to this process of change as plasticity in human biology. Boas noted that the earlier in life migration took place, the greater the plasticity of response. Those born in the United States showed the greatest change from their parents' "type" and often looked more like long-term residents than their own parents. Boas correctly concluded that such plasticity was due to the adjustment of growth and development of the body to the new environment,

especially in relation to improvements in nutrition and health.

Full appreciation of Boas' work waited until after the genocide committed during World War II, in the name of "racial purification," and new discoveries in genetics after 1950. Anthropologists then began to reject the typological approach and "race" in favor of a population approach to the study of human variation and adaptation. By the 1960s, a series of research programs, coordinated by the International Biological Program, were implemented to study human adaptability and variability, and similar research continues today.

Growth and development. An interest in human growth variation is natural for anthropologists because the way a human being grows is the product of an interaction among the biology of our species, the physical environment in which we live, and the social, economic, and political environment that every human culture creates. The basic pattern of human growth and development is shared by all people and is the outcome of the evolutionary history of our fossil ancestors and modern people. That genetically encoded pattern includes approximately three years of infancy, four years of childhood, several juvenile years, and eight adolescent years prior to reaching reproductive and social maturation. Given the relatively long period of growth, the variety of human environments, and the plasticity of human biology, it is not surprising that people achieve significant variation in size, shape, and body composition (amount of muscle, fat, and bone). *See* ANIMAL GROWTH.

The shortest population is the Efe pygmies of central Africa; men average 145 cm (57 in.) and women 136 cm (54 in.). The tallest population are the Dutch of the Netherlands; men average 184 cm (72 in.) and women 172 cm (68 in.). There is much variation between populations, as some Efe are taller than some Dutch. Within populations there is also much variation—at least 60 cm (24 in.) of difference between the shortest and tallest "normal" people. Genetic factors are involved in some of this variation; however, inadequate nutrition and high rates of disease are the two most important environmental factors that lead to poor growth. Recent immigrants from Latin America and Asia (like Boas' migrants) become both taller and longer-legged due to improved nutrition and health in the United States. Over the past 200 years there have been general improvements in health in the wealthier nations due to better public hygiene, safer drinking water, and better diet. Along with these changes there has been a general increase in average stature, 1–2 cm (0.5 in.) per decade. There has also been a faster rate of maturation, as the average age of menarche (first menstruation) declined by 3–4 months per decade. Archeological studies show that the growth trend has returned us to the stature of our Paleolithic, hunting and gathering ancestors. It seems that cultural "advances" such as agriculture and urbanization were harmful to human health and growth in the past 12,000 years. Even today, urban dwellers are less muscular and have less dense bones than our Paleolithic ancestors. This is probably due to the sedentary nature of modern urban life compared with the hunting and gathering exercises of our ancestors. Because the human body changes so rapidly in response to the biocultural environment, many researchers use growth and development of populations as a "mirror" reflecting the material and moral conditions of the society. *See* PHYSICAL ANTHROPOLOGY.

Genetics, geography, and human variation. Genetics is concerned with the structure of deoxyribonucleic acid (DNA), how DNA is organized, how it operates, and how it is transferred across generations. Anthropological genetics focuses on the spatial distribution of genes within breeding populations, that is, groups of people who are likely to mate with each other. Geography, language, religion, nationality, and other sociocultural factors usually delineate human breeding populations. Anthropologists also study how changes in the DNA of breeding populations, brought about by mutation, migration, nonrandom mating, and random mortality in small populations, may be preserved or eliminated by natural selection, which may lead to new genetic variation and adaptation to local environments.

The well-known case of the sickle-cell hemoglobin allele (form of a gene) and its protection against malaria in the heterozygous individual is an example of genetic adaptation to the environment in the human species. Briefly stated, the introduction of farming in tropical Africa created excellent breeding environments for malaria-carrying mosquitoes. Malaria was then, and is still today, a major cause of death in children and teens in Africa. People carrying a single copy of the mutant sickle-cell allele paired with a normal red blood cell allele (heterozygotes) have resistance to malaria. Natural selection, then, increased the frequency of the mutant gene, even though it is lethal to those who inherit a double dose.

Richard Lewontin made the most important discovery of anthropological genetics when in 1972 he reported that more than 90% of human genetic variation is found within populations and less than 10% is found between populations. This is true for large continental populations, such as Africans compared with Europeans and Asians, and also for small isolated human groups, such as South American Indians. Many other geneticists have repeated Lewontin's finding, and together these studies totally disprove the notion of geographic "races." More recently, geneticists found that African populations express the greatest absolute range of genetic variation, while populations in the rest of the world show narrower ranges of variation. The non-African population variation appears to be a genetic subset of Africa. These observations support the African-origin hypothesis of modern human evolution. *See* EARLY MODERN HUMANS; HUMAN GENETICS; MOLECULAR ANTHROPOLOGY.

Human adaptation to climate. Human beings evolved in Africa, which has a range of climates from tropical to moderately cold, from sea level to high altitude, and a high to moderate exposure to ultraviolet (UV) radiation.

Temperature. Outside Africa, people encountered extreme cold, as well as regions with seasonally low ultraviolet radiation. These new climates were stressful and required biocultural adaptations for survival. Populations with long histories of exposure to extreme cold or heat may have evolved some biological adaptations. Humans conform to some extent to the ecological rules of relatively larger body mass, rounder heads, and short extremities in cold environments and relatively more linear bodies with longer arms and legs in hot environments (Bergmann's and Allen's Rules). These variations in body size and shape help to conserve heat in cold climates and dissipate body heat in hot climates. There are, however, many exceptions to these rules. Behavior and culture mitigate many of the stresses of temperature. Nutritional status, workload, and health can significantly modify human body size, shape, and behavior. Cold-adapted peoples (Eskimo, Inuit, Siberians) can modify the flow of blood to the extremities, especially hands and feet, so as to reduce heat loss without suffering frostbite. Growth and development in the cold as well as some genetic mechanisms are implicated in this adaptation. *See* BIOMETEOROLOGY.

UV radiation. All normal humans have the same skin pigment, melanin. Natural selection acted to produce and maintain the continuous variation (called a cline) in melanization from relatively dark near the Equator to relatively light at higher or lower latitudes (**Fig. 1**). Intense UV radiation destroys folate, an essential nutrient, in the human body, which can prevent successful pregnancy or cause severe damage to the fetus. Moderate UV exposure is needed for the body to produce vitamin D_3, which is also essential for normal reproduction and growth. Heavy melanization prevents folate destruction at tropical latitudes but permits sufficient vitamin D_3 production. At temperate and arctic latitudes, folate destruction is not a problem, and lighter skin color allows for sufficient vitamin D_3 synthesis, even in winter. Some arctic peoples are relatively dark-skinned and get their vitamin D_3 from their animal food diets. The lightest skin color is found in Central and Northwestern Europe and seems to be an adaptation to long, cold winters and cloudy skies, both of which severely limit UV radiation exposure. *See* VITAMIN D.

Altitude. The earliest human ancestors in Africa lived at a moderately high altitude of 1500–2500 m (5000–8000 ft) above sea level. Today, many millions still live at this altitude, and about 25 million people live at high altitude (above 3000 m or 10,000 ft). Life at high altitude is difficult due to the combined stresses of cold temperatures, aridity, nutritional inadequacy, steep terrain, heavy workloads, and hypoxia (low partial pressure of oxygen in the air). There is evidence that some Himalayan populations (for example, the Sherpa) have a genetic adaptation to increase the saturation of oxygen in their arterial blood, which can mitigate hypoxia. Other high-altitude populations (such as the natives in the Andes Mountains) adapt to hypoxia via developmental plasticity (for example, growing larger lungs and chests). All high-altitude peoples use behavioral modification (pacing work) and culture (layered clothing and support from social networks) to accommodate to the other stresses. Even so, short stature due to nutritional inadequacy and heavy work is common.

Human adaptations to disease. Infectious and noninfectious diseases, parasites, trauma, the strains of physical activity, and the lack of physical activity are major factors influencing the direction of human adaptation, variation, and evolution. Some variations in disease are genetic (for example, malaria and sickle-cell alleles), some are the result of developmental plasticity (such as the rates of type II diabetes as a consequence of obesity and physical inactivity),

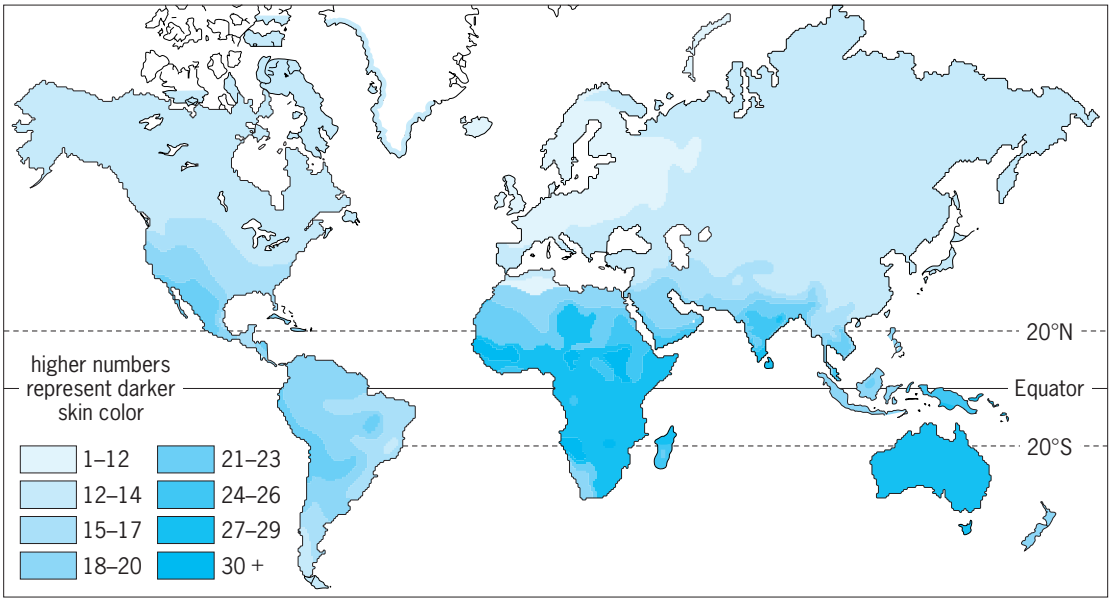

Fig. 1. Map of human skin color distribution around A.D. 1500. (*From D. O'Neil, Distribuzione della Varia Intensitá del Colore della Pelle, in Renato Biasutti, Le Razze e i Popoli della Terra, vol. 1: Razze, Popoli e Culture, 1959*)

while others are due to environmental contaminants (most cancers).

Epidemiology is the study of how, when, where, and why diseases occur. Anthropological epidemiologists focus on the biocultural nature of human diseases, that is, the behavior of humans (the host) and disease-causing agents (pathogens). To better understand the nature of disease, anthropologists also focus on the age, the sex and gender, the social status, lifestyle, ethnic, religious, and other biocultural traits of the hosts. Understanding and combating the worldwide epidemic of human immunodeficiency virus (HIV)/acquired immune deficiency syndrome (AIDS) requires such an approach. The transmission of HIV between people depends on intimate contact between hosts, such as sexual intercourse or blood transfusion. Preventing such contact stops the spread of the disease, but changing human behavior is difficult. Strongly held cultural beliefs and practices regarding sexuality, marriage, and body image (tattooing, scarification) may be vectors of transmission. Some people may believe they are immune to infection because of special social status, and some cultures without an understanding of viral infection may deny the existence of HIV or may ascribe it to spiritual causes. The goal of the anthropologist is to understand these cultural variations and at the same time apply the concepts and methods of human biology to variation in, and problems of, health and disease. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); EPIDEMIOLOGY.

Demography. Demography is the study of fertility, mortality, and migration. Anthropologists study these factors in relation to population growth, family formation, aging, and human ecology.

Fertility. Teenaged girls and women in their late thirties and forties have lower fertility than other age groups (**Fig. 2**). Boys under age 20 rarely produce offspring, but the fertility of men peaks rapidly after age 20, remains level into the fourth decade, and then declines. Biological growth and maturation dictate much of these age trends, but cultural constraints are at work as well. High levels of physical activity and low energy intake decrease female fertility. The rules regarding the formation of reproductive unions and the patterns of intercourse within unions affect the fertility of women and men. Contraceptives are used to some extent in nearly all populations, but their effectiveness varies by the technology employed, social availability and acceptability, and ideological motivation to use them. These factors result in a similar ∩-shaped curve of fertility from ages 15–45 but with differences in amplitude, with some populations achieving fertility rates more than five times greater than other populations (for example, in Fig. 2, the Hutterites versus the Portuguese).

Mortality. Mortality follows a ∪-shaped age pattern in all human populations. Death rates decline from birth to about age seven and remain relatively low throughout the next 20 years of life. Mortality of males rises in the late teens and twenties in militaristic and violent societies, such as the Yanamamö of Venezuela and the United States. By age 30, mortality begins to rise at an ever-increasing rate. In the prehistoric period the major causes of death were likely accidents and other trauma, starvation, and wear and tear on the body from rigorous lifestyles. After the advent of agriculture and settled living, the major causes of death included infectious diseases (interacting with malnutrition) and warfare. With industrialism came greater rates of death from accidents at work, alcohol abuse, and pollution. With the advent of better public health and medicine by the mid-twentieth century, the predominant causes of death in the wealthy nations became chronic degenerative diseases associated with overeating and sedentary lifestyle, automobile accidents, homicides and suicides, and drug abuse (including tobacco and alcohol). *See* AGING.

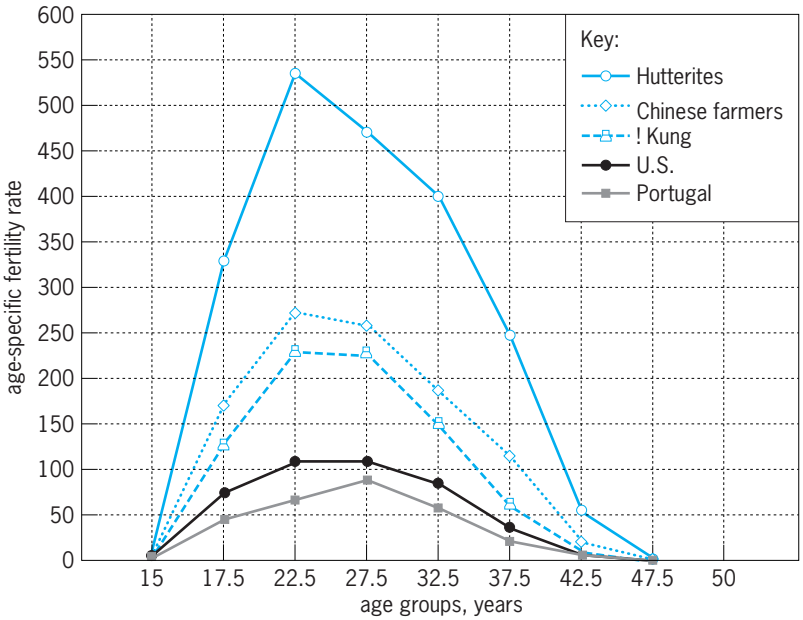

Fig. 2. Age-specific fertility rate of women in several human populations expressed as births per 1000 women. The Hutterites are a religious isolate of North America who prize high fertility and are in generally good health. The "Chinese farmers" represent a group of rural villagers, and the "!Kung" are a society of hunters and gatherers living in the Kalahari desert of Botswana in southern Africa. Lower fertility among the rural Chinese and the !Kung, compared with the Hutterites, may also be due to nutritional stress and disease. The United States is a wealthy industrialized nation in which cultural desire for smaller families combined with effective birth control technology reduces fertility at most ages. The U.S. population is slightly above replacement level, that is, growing very slowly due to in-migration. Portugal is a more extreme case of controlled fertility. The effect of out-migration of young women combined with strong desire for small families lowers fertility and results in a shrinking population over time. (*Adapted from B. Bogin, The Growth of Humanity, 2001*)

Migration. Migration moves people from place to place and redistributes human biocultural variation. Migration injects new genetic, physiological, and morphological variability into the recipient populations. It also may deplete these sources of biological variation from the nonmigrating donor population. Migrants carry diseases from their home to their new places of residence, plus they are exposed to new diseases en route and once they arrive at their destination. Migrants also develop new diseases due to the special circumstances of their life and behavior. Migration rates change due to economic, social, and political circumstances (for example, the slave trade or warfare). Migration rates tend to be highest in

young adults (20 to 30 years old). All of these social and biological factors make the study of human migration difficult, intellectually exciting, and of much practical importance for human welfare. *See* POPULATION DISPERSAL.

Population growth. The balance of fertility versus mortality is responsible for population growth. By the late Neolithic Period (8000 BP) there were about 5 million people on Earth. Today there are more than 6 billion. The exponential growth of the global population is predicted to peak at about 9 billion people by the end of the twenty-first century. *See* POPULATION ECOLOGY.

Conclusion. Today anthropologists and human biologists appreciate the essentially adaptive nature of human variability and the importance of the concept of biocultural adaptability to the study of biological variation. This understanding opens opportunities for greater tolerance, reduced conflict, better health, and greater welfare for the entire human species.

Barry Bogin

Bibliography. B. Bogin, *The Growth of Humanity*, 2001; R. Lewontin, *The Triple Helix: Gene, Organism and Environment*, 2000; J. Marks, *Human Biodiversity: Genes, Race and History*, 1995; S. Stinson et al. (eds.), *Human Biology: An Evolutionary and Biocultural Perspective*, 2000.

Human-computer interaction

An interdisciplinary field focused on the interactions between human users and computer systems, including the user interface and the underlying processes which produce the interactions. The contributing disciplines include computer science, cognitive science, human factors, software engineering, management science, psychology, sociology, and anthropology. Early research and development in human-computer interaction focused on issues directly related to the user interface. Some typical issues were the properties of various input and output devices, interface learnability for new users versus efficiency and extensibility for experienced users, and the appropriate combination of interaction components such as command languages, menus, and graphical user interfaces (GUI). Since the late 1990s, the field of human-computer interaction has broadened and become more attentive to the processes and context in which the users' experience with human-computer interactions takes place. The focus of research and development is now on understanding the relationships among users' goals and objectives, their personal capabilities, the social environment, and the designed artifacts with which they interact. As an applied field, human-computer interaction is also concerned with the development process used to create the interactive system and its value for the human user.

Example of interaction. The interfaces and processes that make up human-computer interaction are understood and advanced through a variety of methods reflecting the field's interdisciplinary nature. Consider, for example, a recent university graduate using a personal computer to search for a job on the Internet. At one level, this interaction can be characterized by the capabilities and processes of the human and the computer to accept input, process that input, and generate output. The computer capabilities include the hardware (input and output devices) such as the monitor, mouse, keyboard, and Internet connection. These devices reflect contributions from computer science and engineering, whereas the human capabilities, both mental and physical, are understood through cognitive science and ergonomics. *See* COMPUTER PERIPHERAL DEVICES.

At another level, the interaction between the computer and the human consists of user interface software which governs the meanings of the inputs and outputs for the computer, as well as the corresponding rules and expectations that the user applies to generate meaningful actions. The software in this example includes an Internet browser with a graphical user interface that reflects recent advances in software engineering and multimedia design. The user's internal model of the interaction is supported by visual cues in the interface and designed in accordance with principles of human factors. *See* HUMAN-FACTORS ENGINEERING.

At a higher level, this interaction includes the context of goals, motivations, and other people and resources that determine what the person is doing, as in this example, searching the Internet for job openings. Understanding the process at this level requires insights from social and organizational sciences.

Advances. Recent advances in knowledge about human-computer interactions across all these levels, from technology to social context, have contributed to the ongoing evolution of the field and to the resulting products and processes available.

Technology. Advances in computer science have significantly increased the processing power of computers while decreasing their size. These advances have provided the underlying technology for creating a wider variety of human-computer interactions. For example, streaming audio and video over the Internet, now common, would not be possible without the increased processing power and network connectivity of computers. These technological developments were influenced by the discovery of useful applications in human-computer interaction. Increasingly sophisticated software has become available to address input through natural speech and immersive environments, providing a virtual reality experience. *See* VIRTUAL REALITY.

Human capability. There is a growing body of knowledge that is providing a good understanding of how humans learn and work with computers. Many factors need to be considered, including models of cognition, motivation, individual differences, and human diversity. Studies on memory, perception, motor skills, attention, learning and skill acquisition, and vigilance are important. Also, there are ergonomic issues that must be considered, such as

sensory limits, fatigue, health, temperature, and environmental noise.

User interface. The user interface includes the input and output devices—such as the monitor, keyboard, and mouse—and the methods by which the users interact with them to carry out tasks. For example, the graphical user interface on most personal computers currently uses a windowing screen designed so that the user can have more than one software application active at any given time. The interface design must provide ways to share both technological and human resources, such as by allocating space on the monitor screen between the active applications and those not in use. These user interface issues are often addressed by interdisciplinary teams. For example, the concept of a windowed environment for multiple applications was investigated by an interdisciplinary team lead by Alan Kay at the Xerox Palo Alto Research Center and popularized in the Macintosh and Windows operating systems. Such empirical studies collect data on how people make use of these facilities, and guide the design of the next generation of the interface to better fit human capabilities and expectations, which are constantly evolving as computer use becomes more ubiquitous. Of particular importance in recent years have been advances in applications fostering human-human collaborations, in visualizing and manipulating multidimensional problem spaces, and in supporting navigation and information retrieval in complex interactive spaces.

Contextual design. Methods from fields such as management sciences, sociology, and anthropology have provided insights into the overall context in which interaction is taking place. For example, people often work in teams to reach goals, so that the interactions taking place between an individual team member and a computer must be considered within the larger framework of human-human communications. The issues involved would be significantly different from the situation of a user who is physically alone but who is participating in a game with fellow players on the Internet, even though the underlying technologies and the software processes might be similar.

Development process. Developing human-computer interactions involves design on both sides of the interaction. On the technology side, the designer must have a thorough understanding of the available hardware and software components and tools. On the human side, the designer must have a good understanding of how humans learn and work with computers, including envisioning new modes of working. The designer's task is to create effective, efficient, and satisfying interactions by balancing factors such as cost, benefits, standards, and the environmental constraints in which the interaction will take place.

Modern prototyping tools allow for the use of an iterative development model where a representative portion of the interface is designed and implemented with each iteration. Feedback from testers is used to enhance the design with each iteration. The final design consists of many elements: the resulting artifacts for use by the target population, as well as supporting elements such as an analysis of needs and tasks, descriptions of the dialog rules and users' conceptual models, expected scenarios of use, and the designer's rationale and reflections from the development process.

Evolution. With the rapid technological advances in interactive computer systems, it is inevitable that new technologies will raise new human-computer interaction issues. For example, the explosive use of multimedia over the Internet has raised issues at various levels. Computer scientists are concerned with providing the highest-quality video possible, given the current computer processing capabilities. User-interface designers are addressing new issues, such as how to index digital video. The types and varieties of human-computer interactions and human-human conversations mediated by the Internet have increased significantly as it has gained wide acceptance at work, school, and home. Understanding the context of these interactions has become more complex as computer use has become more widespread. Perhaps the major challenge facing human-computer interactions is the speed of technological change. The field of human-computer interactions must build the foundations for users to experience incremental change, even though it may not know where those changes will lead. Tom Carey; Kevin Harrigan

Bibliography. M. Helander et al. (eds.), *Handbook of Human-Computer Interaction*, 2d ed., 1997; J. Preece (ed.), *Human Computer Interaction*, 1994; B. Shneiderman, *Designing the User Interface*, 3d ed., 1997.

Human ecology

The study of how the distributions and numbers of humans are determined by interactions with conspecific individuals, with members of other species, and with the abiotic environment. Human ecology encompasses both the responses of humans to, and the effects of humans on, the environment. The lineage leading to modern humans arose in Africa millions of years ago. The oldest fossil remains of members of the genus *Homo* are estimated to be about 2 million years old. The tools associated with human fossils, as well as cave paintings created by early humans, reveal the existence of culture. Cultural learning greatly facilitated the spread of domestic plants and animals and the development of pastoral and agricultural societies. Agriculture, in turn, led to an increasingly sedentary life, greatly expanded food supplies, the development of cities, and the rapid growth of the human population. Human ecology today is the combined result of humans' evolutionary nature and cultural developments. *See* BIOSPHERE; ECOLOGICAL COMMUNITIES; ECOSYSTEM.

Responses to environments. Humans' strong positive and negative emotional responses to components of the environment evolved because our ancestors' responses to environmental information affected survival and reproductive success. Early humans needed to interpret signals from other

organisms and the abiotic environment, and they needed to evaluate and select habitats and the resources there. These choices were emotionally driven.

A vital step in the lives of most organisms, including humans, is selection of a habitat in which to live, gather resources, and seek shelter. Evolutionary theory predicts that habitats that evoke strong positive responses should be those in which survival and reproductive success have been high. Conversely, habitats in which survival and reproductive success have been low should evoke weak or negative emotional responses. The responses of people on several different continents conform to these predictions. Food is one of the most important resources provided by the environment. Gathering food requires decisions of where to forage and what items to select. Anthropologists often use the theory of optimal foraging to interpret how these decisions are made. The theory postulates that as long as foragers have other valuable ways to spend their time or there are risks associated with seeking food, efficient foraging will be favored even when food is not scarce. This approach has facilitated development of simple foraging models and more elaborate models of food sharing and gender division of labor, symbolic communication, long-term subsistence change, and cross-cultural variation in subsistence practices.

Social systems. Humans rely heavily on culture to adapt to their environments. Cultural evolution is a process in which individuals copy others or are taught and then become models for others. But people do not imitate others at random; they pick and choose whom and what to imitate. Thus, cultural evolution proceeds by a process of biased transmission, guided by behavioral rules that have been molded by natural selection. Studies of the diffusion of technical innovations are used to understand the pragmatic decision-making techniques that people use when considering whether to adopt an innovation.

Prior to domestication of plants and animals, humans lived in small bands that shifted locations in response to changing distributions of food resources in the environment. Some of these nomadic people, such as Australian aboriginals, developed elaborate stories and songs that transmitted complex geographical knowledge to young people. Pastoral societies continued with similar nomadic traditions and knowledge.

Domestication of plants and the rise of agriculture, which began about 11,000 years ago in the Mideast, increasingly led to a more sedentary existence because crops needed to be tended and agricultural products could be stored for consumption during lean periods. In addition, by replacing complex ecological communities dominated by species of no food value to humans, with communities dominated by a few species of plants and animals that humans can eat, agriculture increased from 10 to 100 times the amount of food produced per unit land area. Food surpluses made possible the rise of cities, hierarchical societies, and specialized professions. Large domestic mammals revolutionized human society by replacing human backs as the major mode of land transport of goods. Eurasia's horses played a major role in wars of conquest, and maintained their role in military assault until World War I.

Ecological conditions have exerted a powerful influence on human history. The evolution of complex technologies by Eurasians was favored by the locations of continents and mountain ranges. Eurasia offered more readily domesticated plants and animals than other continents. Diffusion of innovations was easier in Eurasia than in other regions because the east-west-oriented mountain ranges facilitated dispersal of domesticated organisms and technological advances across regions at comparable latitudes and, hence, with similar climates. Also, as a result of millennia of living at high densities in association with large domestic animals, Eurasians developed immunity or resistance to an array of diseases with which people on other continents had no experience. Thus, when Europeans explored the other continents they introduced diseases that were lethal to the inhabitants. People from other continents had no comparable diseases to transmit to Europeans.

The rise of cities, a dominant feature of recent human history, was a major result of domestication of plants and animals. A city is supported by resources gathered from regions beyond its area. For example, 29 cities in the Baltic Sea region appropriate the production, in the form of wood, paper, fiber, and food, of ecosystems that occupy about 200 times the total area of the cities themselves. If cities are to be sustainable, extensive supporting ecosystems must continue to provide renewable resources. Most governmental policies pay insufficient attention to these services. Indeed, existing subsidies often deliberately or inadvertently impair the functioning of supporting ecosystems.

Influences on environments. Significant modification of the environment by people was initiated by the domestication of fire, used to change vegetation structure and influence populations of food plants and animals. Vegetation burning is still common in the world, particularly in tropic regions.

By about 1 million years ago, humans had spread from Africa to eastern Asia; Europe was colonized about 500,000 years ago. About 50,000 years ago, humans invented sophisticated tools for capturing and subduing a wide array of prey, including large mammals. Tool-bearing people arrived suddenly in Australia about 40,000 years ago; they crossed the Bering Land Bridge into North America about 12,000 years ago. Within 2000 years humans had spread across the entire New World. Finally, people colonized the islands of the Pacific Ocean, a process that was completed a little more than 1000 years ago.

The arrival of humans with sophisticated tools precipitated the next major transformation of Earth, the extinction of large vertebrates. When humans arrived, Australia and New Guinea had a suite of large birds and mammals, including giant kangaroos, rhinolike marsupials the size of cows, large marsupial carnivores, a 400-lb (182-kg) flightless bird, and giant

lizards, snakes, and land-living crocodiles. All were extinct within a few thousand years of human arrival. Today the largest native mammals of the region are 100-lb (45-kg) kangaroos. The large birds and mammals of New Zealand, Madagascar, and the Hawaiian Islands also became extinct shortly after those islands were colonized by humans.

When humans arrived in North America about 12,000 years ago, they encountered a rich and varied fauna of large mammals that included elephants, horses, giant bison, camels, giant ground sloths, lions, and cheetahs. Within a few centuries, all had become extinct. Debate still surrounds the cause of those extinctions, but those mammals had survived many advances and retreats of glaciers and massive climate changes. Their disappearance precisely when humans arrived strongly suggests that humans sealed their fates. Extinctions of large mammals were far fewer in Africa and Eurasia, where human culture evolved slowly and large animals had ample time to adapt to the presence of this curious predator.

Whatever the cause of the extinctions of the megafaunas of the New World and Australia, humans there lost the opportunity to domesticate large mammals. When horse-riding Spaniards arrived in the New World, they defeated vastly greater numbers of Aztec and Incan foot soldiers. The conquest of the New World was probably the first major sociological consequence of a human-caused loss of biodiversity.

Agriculture drove the third major human modification of environments. Agricultural lands are manipulated so that their productivity is channeled primarily to human needs. Humans appropriate terrestrial production directly as food and fiber and indirectly as food for domestic animals. Today about 35–40% of terrestrial primary production is appropriated by people, and the percentage is rising. Soon, only about half of terrestrial primary production will be available for use by all other species on Earth.

Determining the percentage of aquatic primary production appropriated by people is more difficult. The world fisheries catch was about 131 million metric tons per year in 1990, about 85% of which came from wild stock. By assigning harvested species to trophic levels and using models of the efficiency of trophic energy transfer, ecologists calculated that about 8% of the Earth's primary aquatic production is required to support current fisheries. However, humans use 24–35% of the primary production of continental shelves and upwelling areas where most fish are harvested. Although about 75% of aquatic production occurs in the open oceans, those regions are so unproductive that it is profitable to harvest only the largest predators, such as tuna.

Humans appropriate an even higher percentage of the world's fresh water. Only about 0.77% of the Earth's water exists as fresh water in the atmosphere, aquifers, lakes, streams, soils, and organisms. About 97.5% of global water is salty; the remainder exists as permanent ice. Humans currently use about 54% of available runoff for agriculture, industry, municipal needs, and disposal of wastes. At current per capita rates of usage, the growing human population will need more than 70% of available runoff by 2025.

The high level of human appropriation of ecosystem production, combined with the landscape modifications it engenders, is currently the most important cause of species extinctions. The habitats required by some species are being destroyed. Habitats such as old-growth forests, natural grasslands, and estuaries are being reduced to widely separated patches that are too small to sustain populations of many of the species. Small habitat patches cannot maintain populations of species that require large areas, and they support only small populations of many species that can survive in them. In addition, the fraction of a patch that is influenced by effects originating on adjacent habitats rapidly increases as patch size decreases. As human appropriation of ecosystem production increases, more land will be converted to agriculture, thereby increasing the fragmentation of natural habitats.

The increasingly mobile human population is introducing, either deliberately or inadvertently, thousands of species into some areas. These exotic species often cause the extinction of native species upon which they prey, with which they compete, or with which they share diseases. Biogeographers may come to refer to the current period in human history as the Homogocene.

The magnitude of the current human enterprise is great enough to modify the Earth's biogeochemical cycles. Primarily as a result of burning fossil fuels, the concentration of atmospheric carbon dioxide has risen from a pre–industrial revolution value of about 265 parts per million to 350 ppm today. If current trends continue, concentrations are expected to reach 580 ppm by the middle of the twenty-first century. The buildup of carbon dioxide, which is transparent to sunlight but opaque to radiated heat, is expected to increase the Earth's mean temperature by about 3–5°C (5–9°F), with greater increases at high latitudes. Climate patterns will shift latitudinally, new climates will come into existence, polar icecaps will melt, and sea levels will rise, flooding coastal regions. As world climate changes, the ranges of many species must shift. Individuals will need to disperse in increasingly fragmented landscapes in which only a small fraction of the area is suitable for them. In the past, range shifts took place in landscapes that offered much easier dispersal routes. *See* BIOGEOCHEMISTRY; GREENHOUSE EFFECT.

Other biogeochemical cycles are also being substantially modified. Currently the magnitude of the global sulfur cycle is about twice the preindustrial level. Increases in sulfur emissions increase atmospheric concentrations of dimethyl sulfide, the major source of condensation nuclei, causing increases in cloud cover and additional climate changes. As a result of the massive increases of fixed nitrogen being introduced into the Earth's ecosystems by human activity, concentrations of nitrous oxide and nitric oxide are increasing in the atmosphere. Nitrous oxide absorbs infrared radiation and contributes to global warming. Nitric oxide is involved in reactions

that create peroxyacyl nitrate, a constituent of smog. Nitric oxide is also converted to nitric acid, a principal component of acid precipitation. Acid precipitation causes decreases in soil fertility, acidifies lakes and streams, and percolates into ground water where it causes health problems in communities that consume that water. *See* ACID RAIN.

The future. Humans will continue to exert powerful influences on the functioning of the Earth's ecological systems. The human population is destined to increase for many years. Rising affluence will be accompanied by increased consumption of resources and, hence, greater appropriation of the Earth's primary production. Nevertheless, many future human ecology scenarios are possible, depending on how much the human population grows and how growth is accommodated, the efficiency with which humans use and recycle resources, and the value that people give to preservation of biodiversity. Some of the most important ethical decisions faced by human society today concern relationships with and responsibilities to the other species that share the Earth. The ethical positions that people adopt during the next few decades will determine to a substantial degree the actual ecological scenarios experienced by our distant descendants. *See* ANTHROPOLOGY; ECOLOGY; ENVIRONMENT; SOCIOBIOLOGY. Gordon H. Orians

Bibliography. S. C. Bourassa, *The Aesthetics of Landscape*, Belhaven Press, 1991; J. M. Diamond, *The Third Chimpanzee*, HarperCollins, 1992; J. M. Diamond, *Guns, Germs, and Steel*, Norton, 1997; P. R. Ehrlich, *A World of Wounds: Ecologists and the Human Dilemma*, Ecology Institute, 1997; T. F. Flannery, *The Future Eaters*, Reed Books, 1994; R. T. T. Forman, *Land Mosaics: The Ecology of Landscapes and Regions*, Cambridge University, 1995; G. Hardin, *Living Within Limits*, Oxford University, 1993; E. A. Smith and B. Winterhalder (eds.), *Evolutionary Ecology and Human Behavior*, Aldine de Gruyter, 1992; E. O. Wilson, *Biophilia*, Harvard University, 1984.

Human-factors engineering

The application of experimental findings in behavioral science and physiology to the design and operation of technical systems in which humans are users or operators. This includes design of hardware, software, training, and documentation as well as manufacturing and maintenance. Human-factors professionals are trained in some combination of experimental or cognitive psychology, physiology, and engineering—typically industrial, mechanical, electrical, or software engineering. Human-factors engineering seeks to ensure that humans' tools and environment are best matched to their physical size, strength, and speed and to the capabilities of the senses, memory, cognitive skill, and psychomotor preferences. These objectives are in contrast to forcing humans to conform or adapt to the physical environment.

Human-factors engineering has also been termed human factors, human engineering, engineering psychology, applied experimental psychology, ergonomics, and biotechnology. It is related to the field of human-machine systems engineering but is more general, comprehensive, and empirical and not so wedded to formal mathematical models and physical analysis. *See* HUMAN-MACHINE SYSTEMS.

Among the problems of human-factors engineering are design of visual displays for ease and speed of interpretation; design of tonal signaling systems and voice communication systems for accuracy of communication; design of seats, workplaces, cockpits, and consoles in terms of humans' physical size, comfort, strength, and visibility. Human-factors engineering addresses problems of physiological stresses arising from such environmental factors as heat and cold, humidity, high and low atmospheric pressure, vibration and acceleration, radiation and toxicity, illumination or lack of it, and acoustic noise. Finally, the field includes psychological stresses of work speed and load and problems of memory, perception, decision making, and fatigue. *See* HUMAN-MACHINE SYSTEMS.

History. The foundations of human-factors engineering were laid by Fredrick W. Taylor, who demonstrated that, by proper design of workplaces and procedures, the productivity of workers could be greatly increased. Early systematic studies were made by Frank B. Gilbreth and Lillian M. Gilbreth, whose therblig (an anagram after Gilbreth) system of categorizing hand movements is still a standard motion-and-time analysis technique. With World War II came a great demand for psychologists and physiologists to help engineers design aircraft, ships, tanks, and other weapons to ensure that these devices could be operated under stress by people with relatively little training. All too often it was found that engineers had designed the equipment around themselves and, as a result, large sailors could not fit in the required spaces, small pilots could not reach the required controls, or an appreciable percentage of operators were lacking in sufficient strength or were confused by procedures for operating the complex implements of war. After World War II, formal courses in human-factors engineering were introduced in psychology and engineering departments of colleges and professional organizations in the United States and in many European countries.

Early efforts were directed to providing experimental data for the more obvious gaps in human-factors knowledge, such as the variations in the physical dimensions of men and women of different ages, the strength and speed capacities of humans, and the physical conditions under which they could just barely detect or read certain visual images and hear or discriminate certain sounds. As more powerful techniques come to be developed, other sensory capacities of touch, taste, smell, and motion, as well as the more complex and subtle problems of training, stress, and fatigue, became amenable to laboratory experiment.

Anthropometry. The specialized field dealing with the physical dimensions of the human body is called

anthropometry. In dimensioning a seat, console, workplace, or special piece of personal equipment (such as a space suit), it is important that as large a fraction of the intended users as possible be accommodated. But fitting a console to the largest person usually means the smallest person cannot reach the controls unless various features of the equipment are adjustable. Moreover, people are not shaped in simple proportion; the person with the largest head often does not have the longest legs or greatest girth. There is no average human: the person who is average in one dimension is not usually average in another (**Fig. 1**). To accommodate everyone, many features of a machine need to be adjustable independently. Because this is not always possible, it is important in sizing the equipment to know what percentage of large people and what percentage of small people will not be accommodated, and to weigh the relative costs and advantages of adding adjustments or building the equipment in different sizes. *See* ANTHROPOMETRY.

Displays and controls. Problems of visual displays have been of considerable interest in designing aircraft, spacecraft, ships, submarines, nuclear reactors, electronic equipment, and tools of all kinds. Human-factors engineers have sought to specify the light intensity required for reading different dials and signs in relation to the illumination level of the background. They have recommended, for specific tasks and environments, optimum shape and spacing of numerals and indicator marks on dials and specified what colors give the best contrast and how far away markings of certain size can be read.

Electronics have made possible a number of new visual display techniques: Self-illuminating electroluminescent numerals and holographic photographs, when illuminated by coherent (laser) beams, appear strikingly three-dimensional. Cathode-ray tubes driven by computer can provide a tremendous variety of images, including letters and numbers, graphs, and television-type pictures. In aircraft and other applications, single-purpose dials and indicators have been replaced by relatively few computer displays, which are completely flexible and can display a great variety of kinds of information. Sometimes these displays can be called up by the operator by keying in certain code letters. Alternatively, a number of different indications, both qualitative and quantitative, can be integrated on one display, such as the integrated aircraft landing display. *See* AIRCRAFT INSTRUMENTATION; CATHODE-RAY TUBE; COMPUTER GRAPHICS; ELECTRONIC DISPLAY.

With respect to sound signaling systems (auditory displays), human-factors engineers can specify how loud various sounds must be to be heard against background noises of differing loudness. Because consonants, which have relatively higher frequencies and lower loudness, are known to carry much more information than vowels, the fidelity of voice communication systems at these higher frequencies (2000–4000 Hz) is recognized to be more important than the fidelity at lower frequencies. Electronic reading machines and mobility aids for the blind make considerable use of auditory signals to replace visual senses.

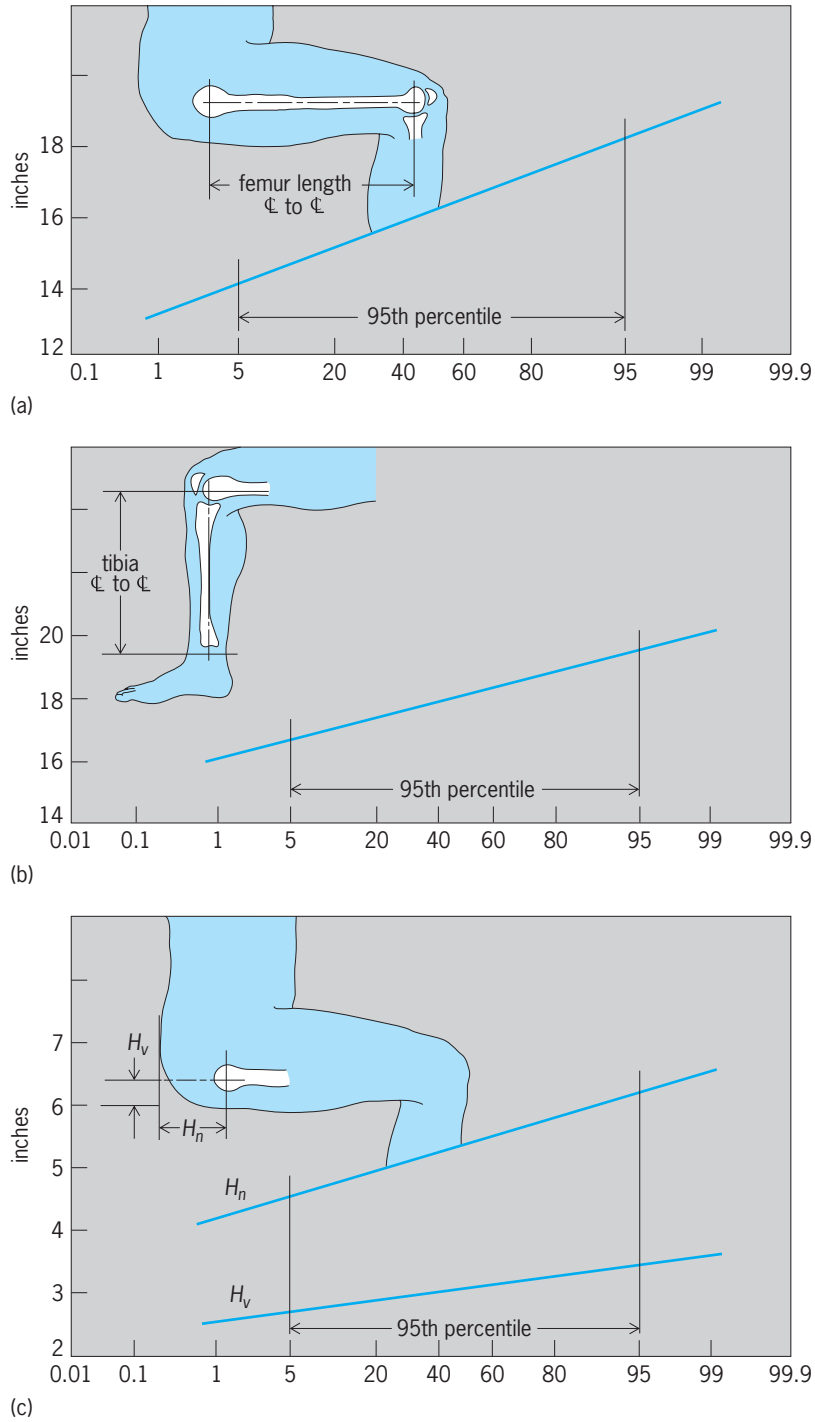

Fig. 1. Variations in the distributions of body dimensions as studied in military personnel. (*a*) Femur length. (*b*) Tibia length. (*c*) Femur-to-seat distances. (*After L. J. Fogel, Biotechnology, Prentice-Hall, 1963*)

An important human-factors engineering problem is coordination of displays with the controls by which responses to these displays are made. In practice, there are many violations of even simple commonsense design principles such as locating controls adjacent to displays when they are used together, arranging for controls to move in the same direction as the associated display, or having all displays and

controls move in the same direction to turn them off. Although the human operator can adapt to a reasonably broad range of forces and sizes of controls, some controls even now are installed that require either more strength or sensitivity than many operators can exercise. Human-factors handbooks are available that summarize various experimental data and give recommendations for display and control design.

Environmental stress. Human-factors engineers are concerned with the design of the environment, both to enhance humans' information-processing ability and to keep the body functioning normally. These concerns require an understanding of human tolerance ranges to heat and cold, to high and low pressures, to varying atmospheric concentrations, and to acceleration and vibration. Engineering solutions can take the form of space suits, diving suits, special clothing for extreme climates, and atmospheric-control units for sealed cabins and tanks. Alternatively, working conditions can be so arranged as to limit exposure to hazard, for example, by determining the least fatiguing distance from the open door of a blast furnace or by recommending the maximum time that can safely be spent at certain altitudes or depths without pressure-breathing equipment.

Bodily comfort is dependent upon a combination of ambient temperature, pressure, humidity, air movement, and amount of clothing—all of which interact to keep the internal body temperature constant at 98.6°F (37°C). Proper respiration depends upon the combination of ambient pressure and oxygen concentration that keep the lungs at the equivalent of breathing air (20% oxygen) at sea level. Constant high accelerations, experienced by astronauts in boost and reentry, can drain the blood from eyes and brain and temporarily prevent the body's proper functioning; such adverse effects can be prevented by orienting the body properly or using g suits, which keep blood from pooling in the veins of the trunk and legs. Vibration is likely to cause motion sickness if it occurs at certain frequencies at which body parts resonate sympathetically (worst at 2–5 Hz). Levels of thermal, pressure, atmospheric, acceleration, and vibration stresses, which are individually experienced as mild, can have violent effects when they occur in combination. *See* AEROSPACE MEDICINE; SPACE BIOLOGY.

Safety. Human-factors considerations often determine whether a particular vehicle, tool, or environment is safe. The ability of the human body to withstand various sudden acceleration forces is an important aspect of highway safety. Seat-belt designs for automobiles are based on rocket sled experiments originally conducted to design crash harnesses and parachute harnesses for aircraft pilots.

Human reaction time is also a large factor in accidents of all types. It has been shown that human reaction under the best conditions, such as those that exist when an alert, expectant test subject is required to push a button in response to a light signal, is about a quarter second. As the number N of response choices from among which the subject must choose increases, the reaction time increases as the logarithm of N. Reaction time can increase many times as a function of boredom and fatigue, but reliable quantitative predictors for those factors are not available.

Work load. Much research has tried to specify the optimum workload for humans. This research is motivated by the gradual disappearance of routine tasks which can be performed at an even pace and somewhat independently of the environment. Humans are used as a monitor of complex semiautomatic systems where the workload for most of the time is relatively light but where detection of certain low-probability contingencies, or failure of the system, requires rapid and dependable overt action. Examples occur in industrial inspection processes, monitoring of nuclear reactors, piloting of spacecraft, and rapid transit trains.

Early studies on ship's watchkeeping and radar watching indicated that humans could remain vigilant only about a half hour for signals that occur rarely if ever (such as a ship on a collision course with another at night or in fog or the approach of enemy aircraft or missiles). After this interval the chances of rare events being detected diminish markedly. One technique employed to keep watchkeepers and monitors alert is to introduce artificial signals that the observer initially cannot discriminate from the real ones. For example, it was reported that in an experiment one soft-drink-bottle inspector passed fewer dirty or defective bottles when, to keep him alert, cockroaches were added to a small percentage of otherwise clean bottles!

The Manned Spacecraft Center of NASA has conducted simulations of long space flights where not only was the workload negligible but in some cases sound, light, temperature, and gravity sensations were reduced to a minimum. Under such conditions humans become decreasingly alert and even hallucinate to substitute imagined sensation in compensation for the lack of real sensation.

At the other end of the workload scale, human-factors research has determined the upper limits on what humans can do with their sense organs and muscles, both in transmitting information and in performing mechanical work. In terms of the unit of information measure, the bit-per second (the average logarithm of the number of choices or binary decisions made per second), humans can transmit up to 35 bits/s in such tasks as piano playing and speaking, but for most routine manipulation skills their rate is much lower. As an engine for mechanical work, humans are rather inefficient because, even while operating at peak efficiency (for example, in pedaling a racing bicycle), experiments show that they can produce 1 hp (750 W) for only about 1 min—and in order to do so their bodies produce several times that amount of energy in wasted heat.

Humans versus machine. A fundamental problem of ever-increasing importance for human-factors engineers is what tasks should be assigned to people and what to machines. It is a fallacy to think that any given whole task can be accomplished best either by a human or by a machine without the aid of the

other, because often some elements of both provide a mixture superior to either alone.

Machines are superior in speed and power; are more reliable for routine tasks, being free of boredom and fatigue; can perform computations at higher rates; and can store and recall specific quantitative facts from memory faster and more dependably. Humans, by contrast, have remarkable sensory capacities which are difficult to duplicate in range, size, and power with artificial instruments (the ratio of the greatest to the least energy which people can either see or hear is about 10^{13}). Humans' ability to perceive patterns, make relevant associations in memory, and induce new generalizations from empirical data remains far superior to that of any computer existing or planned. Thus, while people's overt information-processing rate in simple skills is low, their information-processing rate for these pattern recognition and inductive- reasoning capabilities (of which little is understood) appears far greater.

Problem areas. Below are described four different problem areas of human-factors engineering, and within each, two different specific problems are described. These illustrate the variety of challenges encountered in human-factors engineering. For each problem the design solutions require that the hardware, software, and training considerations be integrated.

Information display. Problems of information display include those of space telepresence and those of displays in nuclear and chemical plants.

Space telepresence. Crewed space flight is very expensive, and currently is not practical beyond the Moon for safety reasons. Exploration of the planets may be accomplished more economically by using robotic space vehicles and telepresence. Telepresence means that a human, in this case on Earth, can look around, see, and have the sense of being located in an environment other than her actual physical location. This is accomplished by means of a head-mounted display, which measures the orientation of the head and communicates this to a video camera in the remote environment, directing it to point in the same direction. The human then sees what she would see were she in that environment looking in that direction. This sense of telepresence is very compelling, and human-factors engineers are studying its limitations and refining this equipment for use in space. One limitation, for example, is the time delay in video and control signal communication between Earth and space, which means the video feedback lags the corresponding head direction by seconds to minutes, depending how far out in space is the planet being explored.

Displays in nuclear and chemical plants. Another type of display problem is the design of computer displays for nuclear power plants and chemical process plants. Because there are literally thousands of variables that must be monitored, for safety reasons it is important that these displays combine information about related variables, and give an overall picture of the health of the plant. There are several types of information to be displayed. First there are alarms, both auditory and visual, indicating what has failed and providing some direction about how to respond. Then there are status lights, showing a pump is on or off, a valve is open or closed. There are also indications of numerical values of key variables, pressures, temperatures, flow rates, and so on. Finally there are electrical and piping diagrams and written information, much of which is computerized to save space and make it more quickly accessible. *See* NUCLEAR REACTOR; PROCESS CONTROL.

Control. Problems of control involve aircraft automation and robots.

Aircraft automation. Modern commercial aircraft have a sophisticated control system called a flight management system. The flight management system is a computer-based system which includes the autopilot, the navigation subsystem, a radar-based traffic collision avoidance subsystem, and many other functions. Mostly the pilot flies the aircraft through the flight management system, and especially on long flights uses the manual controls (yoke and pedals) only for taxiing, takeoff, and landing, and not necessarily even then. A properly programmed flight management system has the capability to fly the aircraft automatically from takeoff to landing without the pilot ever touching the controls. The flight management system is programmed through a keyboard resembling a pocket calculator. However, there remain many problems of pilot error in using the flight management system, requiring significant human-factors effort.

Control of robots. Control of a robot, whether for painting, welding, or parts assembly in a manufacturing plant; for performing inspections and repairs on deep-ocean wellheads or pipes; or for working in chemically toxic or other hazardous environments, is similar to the flight management system in many ways. A human supervisor must plan a task, program its special features, and monitor the robot after it is put into automatic mode, updating its instructions as task conditions change. *See* ROBOTICS.

Safety research. Problems of safety research include those involving field studies in automotive safety and those involving the use of a human-in-the-loop simulator for evaluation.

Field studies in automotive safety. Human-factors specialists are engaged in a variety of field studies to investigate automobile accidents, examine roadway configurations to recommend signing, and test new designs of automobiles and trucks from the driver's point of view. What a normal driver can be expected to see or hear, and how he or she is expected to respond, are issues which must be backed up by statistics.

Human-in-the-loop simulator. Often vehicle safety research must be performed which is too dangerous to do in actual circumstances. In that case, use of a human-in-the-loop simulator is warranted. This is a device that is operated by a real person but in other respects is an electromechanical simulation. The operator uses realistic controls and views a computer-generated display which provides a realis-

tic dynamic view of how the vehicle responds, giving the driver the feeling of being in an actual vehicle. The flight simulator is the best-known example of such a system. These have become so realistic that pilots are checked out in new aircraft types on simulators, and the first time they fly the aircraft is with a load of passengers. Such simulators are used for automobiles as well, enabling study of driver response in crash situations without actual danger to the driver. *See* AIRCRAFT TESTING.

Health care. Problems of health care include those involving doctors' instructions and those of in-home monitoring for the chronically ill.

Doctors' instructions. A well-known source of human error in hospitals has been the misreading of handwritten physician orders to nurses, hospital laboratory technicians, and pharmacists. A solution is to have physicians enter their instructions into computers. Then not only is there immediate feedback to the physician as to what was typed and an unambiguous record which can easily be transmitted to wherever it is needed, but also the computer can perform certain checks to see whether the medicine in the prescribed dosage is appropriate for what is known (in the computerized hospital database for the patient).

In-home monitoring. People who have serious chronic medical conditions such as congestive heart problems often are not aware of when they are in danger and when there is no immediate cause for concern. Some people are not treated until it is too late, while others unnecessarily crowd the emergency rooms of hospitals. An avenue of solution which can avoid time and expense for both the patient and medical personnel is to have monitoring devices at home which are connected by telephone modem to a hospital. Either the patient can attach such devices, or in more serious cases a nurse or health care worker visiting the home regularly can assist the patient in making certain measurements. The measuring device can make a simple analysis, comparing the measurement to readings made on the same patient on previous days, and decide if the situation is safe, that is, not getting worse. Otherwise the data can be telemetered to a hospital database and analyzed there by a specialist for appropriate action. To make such home care systems safe, effective, and user-friendly poses many human-factors challenges. Thomas B. Sheridan

Cognitive engineering. One of the major issues in human-factors engineering is the concern that modern sophisticated hardware and software technology may be too complex for the people who will eventually use it. Requiring people to perform difficult or cognitively complex tasks is perhaps the leading cause of human–machine errors or accidents. Tasks can be cognitively difficult for a number of reasons. The number of steps required to use the system may exceed people's memory limitations, the user may be required to divide his or her attention between several different sources of information, or the person may be required to perform difficult mental operations. All of these factors burden an individual's cognitive capacity and, if that capacity is exceeded, errors may occur.

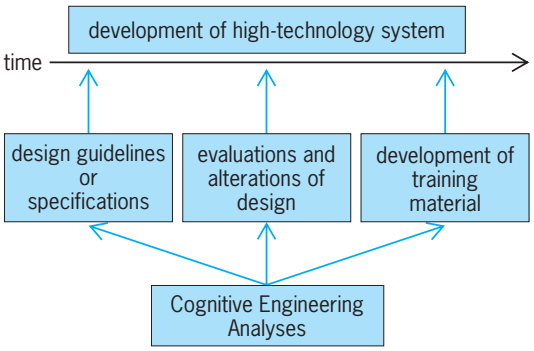

Fig. 2. Cognitive engineering and the development of new technology applications.

Two major trends have led to increased emphasis on the cognitive complexity of human-machine systems. One of these is the move toward larger and more complex systems where human error can have serious consequences for the systems' users and the general public. The other trend is the rapid development of modern information technology based upon powerful yet inexpensive microcomputers.

An important aspect in addressing this problem is the early identification and control of the cognitive complexity or mental difficulty of performing a task required of the new technology application. Identification and control of the mental difficulty of tasks is the goal of cognitive engineering. This aspect of engineering is accomplished in different ways depending on the development status of the new technology application (**Fig. 2**).

The best procedure (in terms of cost and effectiveness) for addressing this problem is to use cognitive analysis to develop specifications or guidelines that can be used during the initial design technology application. Such specifications might include instructions such as "do not ask the user to remember more than two items simultaneously" or "put all relevant information in a single location on the equipment." Successful designs are achieved through planning and sensitivity to user needs. Detailed qualitative models of the cognitive processes of the system's users serve important roles in these analyses. They can suggest the overall complexity of different design options for users, and indicate those aspects of each option that may exceed human cognitive capabilities. Through such design guidance, human cognitive limitations are controlled early in the technology application design when it is easiest and most cost-effective to make changes.

If the technology application has developed to the point where design guides would no longer be useful (for example, much design work has already been completed), an alternative approach is to use cognitive engineering to evaluate the design as it exists. The results of the cognitive analysis will indicate which aspects of the design may be too difficult for people to perform and could lead to human–machine errors. Those aspects of the techonology

application can then be redesigned in order to reduce the likelihood of such error. The disadvantage of this approach is that redesigning hardware or software is often more expensive than designing it correctly in the first place.

The final use of cognitive engineering analysis is in preparing training materials. Cognitive analyses can identify the aspects of a person–machine interface that will be most difficult for people to perform. These aspects can then be given special training aids or more intensive hands-on training in order to reduce the potential of human–machine error. Paul G. Rossmeissl

Engineering psychology. The science of psychology has always been concerned with the basic functions that enable humans to interact effectively with their environment: sensing, attending, perceiving, remembering, choosing among actions, acting. Researchers have typically worked under carefully controlled laboratory conditions, using experimental methods derived from the older physical sciences.

Topics of greatest interest to engineering psychology at any given time are dictated by technological developments and the problems they pose for the designer and operator. Many issues revolve around computer-based systems with their almost limitless potential for automation, information handling, and modes of interfacing with human operators.

Automation. Since almost every system operation can be automated, designers must decide which functions should be left for the human operator. Although the choice seems to depend on the particular system, the trend is clearly toward giving human operators more overseeing and decision responsibility and less moment-to-moment control of system processes. Today's commercial airline pilot, for example, manages and monitors the flight to a great extent rather than actually flying the airplane.

Information processing. Given their immense capability for gathering, manipulating, storing, and displaying information, modern systems can easily swamp the human operator with information. This abundance of information raises important research issues about human capacity limits and ways of handling information, and about the mental effort required by particular tasks and system designs. Systems engineers must determine how mental workload should be measured, what is an acceptable amount of it, and how performance suffers as that amount is exceeded. Ultimately, the goal is to reduce mental workload without sacrificing information that the operator needs in order to perform well. Engineering psychologists have gained considerable insight and made some valuable practical innovations regarding these issues. For example, sound techniques have been developed for measuring the subjective mental workload associated with particular tasks, and these measures can be used to estimate the relative merit of alternative designs from a human performance perspective.

Human-machine interface. One way of coping with task complexity and workload is through improved design of the human-machine interface, that is, the way that information is displayed to the operator and responses are executed. Modern technology offers a host of interface options—from ordinary text to sophisticated graphics on the display side; from keyboards, to menus and touch screens, to "mice," joysticks, trackballs, and speech sensors on the response side. Some combinations of these design options (including applications of techniques such as virtual reality) can make the operator's task considerably easier, in effect lessening the complexity and the resulting mental workload. *See* VIRTUAL REALITY.

The best combinations of such tools, however, are not immediately apparent from a human-performance perspective. What works best is heavily dependent on how the operator is to process the information—the task requirements. The key relationship between the representation of to-be-processed information and processing requirements is known as compatibility. For information-rich, complex tasks, the compatibility of design features with the mental representation of the task (the operator's mental model) is what matters most. Perceived complexity and mental workload are lowest (hence operator and system performance are best) when the system design is compatible with these mental models.

Decision making. Human decision making under the conditions of high stress, uncertainty, and information load that are experienced by firefighters, air-traffic controllers, fighter pilots, and even automobile drivers in an unfamiliar city, requires processing a lot of information quickly and choosing the right course of action when a wrong one could have serious, even catastrophic, consequences.

There are major differences in the way that effective and ineffective decision makers approach this kind of task. Effective ones compare their past experience in similar situations when faced with a new crisis. Rather than trying to analyze the presently available information in depth, they simply scan their memory for a similar pattern and act accordingly. Poor decision makers, however, focus on only the current information and tend to become over-whelmed. Knowing how effective performance is accomplished, designers can build features into new systems that encourage operators to use the proper strategy and that aid them in doing so. Moreover, training can be focused on this approach. William C. Howell

Bibliography. S. Andriole and L. Adelman, *Cognitive Systems Engineering for User-Computer Interface Design, Prototyping and Evaluation*, Lawrence Erlbaum Associates, Mahwah, NJ, 1995; R. W. Bailey, *Human Performance Engineering*, 3d ed., Prentice Hall, Upper Saddle River, NJ, 1996; A. Chapanis, *Human Factors in Systems Engineering*, John Wiley, New York, 1996; W. C. Howell, Engineering psychology in a changing world, *Annu. Rev. Psychol.*, 44:231–263, 1993; J. Rasmussen, A. M. Pejtersen, and L. P. Goodstein, *Cognitive Systems Engineering*, John Wiley, New York, 1994; G. Salvendy (ed.), *Handbook of Human Factors and Ergonomics*, 2d ed., John Wiley, New York,

1997; M. S. Sanders and E. J. McCormick, *Human Factors in Engineering Design*, 7th ed., McGraw-Hill, 1993; B. Shneiderman, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 3d ed., Addison-Wesley, Reading, MA, 1997; C. D. Wickens and J. G. Holland, *Engineering Psychology and Human Performance*, 3d ed., Prentice Hall, Upper Saddle River, NJ, 1999; C. E. Zsambok and G. Klein (eds.), *Naturalistic Decision Making*, Lawrence Erlbaum Associates, Mahwah, NJ, 1997.

Human genetics

A discipline concerned with genetically determined resemblances and differences among human beings. The idea that certain physical and mental characteristics, normal or abnormal, can "run in families" goes back to ancient times, though the mechanism by which heredity operates remained mostly unknown until the twentieth century. Formerly, genetics was thought to be concerned only with the familial transmission of rare and insignificant characteristics, but its fundamental biological role is now apparent. Genes, the units of heredity, have two unique properties: they are self-replicating, and they carry in their biochemical structure the codes for protein synthesis. Consequently, genes play the double role of transmitting genetic information from generation to generation and of governing all the activities of living cells. *See* GENE.

Expansion of human genetic knowledge has come from several different directions and has had major consequences for human health and for the understanding of the place of humans in nature. Technological advances in the visualization of human chromosomes have shown that abnormalities of chromosome number or structure are surprisingly common and of many different kinds, and that they account for birth defects or mental impairment in many individuals as well as for numerous early spontaneous abortions. Progress in molecular biology has clarified the molecular structure of chromosomes and their constituent genes and the mechanisms of deoxyribonucleic acid (DNA) synthesis and protein synthesis, as well as the ways in which change in the molecular structure of a gene can lead to a disease. Concern about possible genetic damage through environmental agents, particularly ionizing radiation, and the possible harmful effects of hazardous substances in the environment on prenatal development has also stimulated research in human genetics. The medical aspects of human genetics have become prominent as nonhereditary causes of ill health or early death, such as infectious disease or nutritional deficiency, have declined, at least in developed countries. It has been estimated, for example, that half of all admissions to pediatric hospitals involve disorders that are completely or partly genetic. Detailed knowledge of the distribution of many genetic traits in individuals, families, and populations has expanded, providing fresh insights into human origins and the process of evolution.

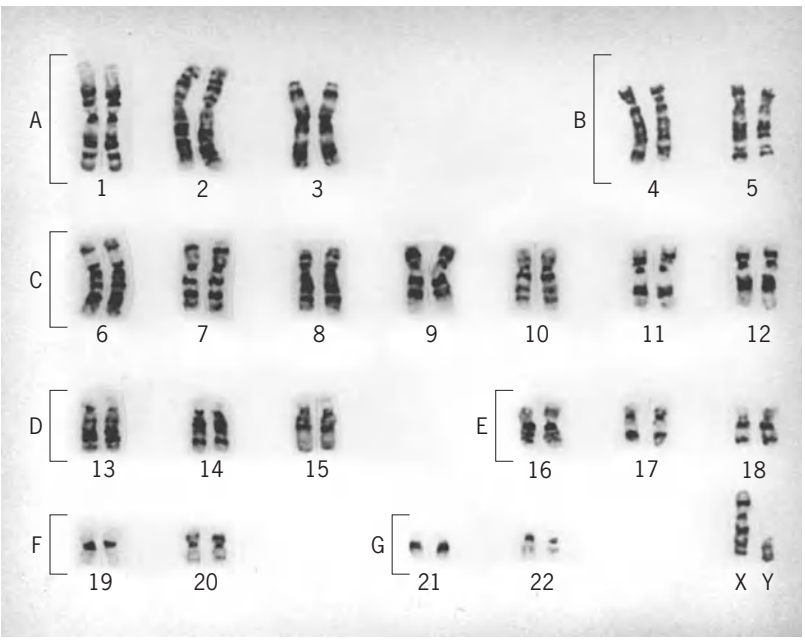

Fig. 1. Normal male karyotype. The chromosomes are classified in seven groups labeled A–G. The chromosome pairs are individually labeled. (*Photomicrograph courtesy of R. G. Worton; from J. S. Thompson and M. W. Thompson, Genetics in Medicine, 3d ed., W. B. Saunders, Philadelphia, 1980*)

Chromosome and gene structure. In normal humans, the nucleus of each normal cell contains 46 chromosomes, which comprise 23 different pairs. A karyotype (**Fig. 1**) may be prepared from any type of cell that will divide in cell culture, such as white blood cells, skin, bone marrow, or cells from amniotic fluid, for example. Of each chromosome pair, one is paternal and the other maternal in origin. In turn, only one member of each pair is handed on through the reproductive cell (egg or sperm) to each child. Thus, each egg or sperm has only 23 chromosomes, the haploid number; fusion of egg and sperm at fertilization will restore the double, or diploid, chromosome number of 46. *See* CHROMOSOME; FERTILIZATION (ANIMAL).

The segregation of chromosome pairs during meiosis allows for a large amount of "shuffling" of genetic material as it is passed down the generation. With 23 pairs of chromosomes, the total possible number of different chromosome combinations in the gametes is 2^{23}, or about 8 million. Two parents can provide $2^{23} \times 2^{23}$ different chromosome combinations. This enormous source of variation is multiplied still further by the mechanism of crossing over, in which homologous chromosomes exchange segments during meiosis. *See* CROSSING-OVER (GENETICS); MEIOSIS.

Twenty-two of the 23 chromosome pairs, the autosomes, are alike in both sexes; the other pair comprises the sex chromosomes. A female has a pair of X chromosomes; a male (Fig. 1) has a single X, paired with a Y chromosome which he has inherited from his father and will transmit to each of his sons. Sex is determined at fertilization, and depends on whether the egg (which has a single X chromosome) is

fertilized by an X-bearing or a Y-bearing sperm. *See* SEX DETERMINATION.

All chromosomes are composed of DNA. It is estimated that the total length of DNA in the haploid set of human chromosomes is 1.7 m (5.7 ft), which is 10,000 times the length of the chromosomes at metaphase. Genes are segments of DNA. There may be as many as 50,000–100,000 genes on the 46 human chromosomes, of which about 3000 have been identified and about 1500 have been mapped to specific chromosomal locations. *See* DEOXYRIBONUCLEIC ACID (DNA); GENETIC MAPPING.

The genetic information is coded in DNA in the form of triplets of four bases: two purines, adenine (A) and guanine (G), and two pyrimidines, thymine (T) and cytosine (C). Each triplet combination (codon) codes for a specific amino acid. The sequence of bases in a specific gene dictates the sequence of amino acids in the specific protein coded by that gene. *See* GENETIC CODE.

A gene initiates synthesis of a protein by transcription of the DNA into messenger ribonucleic acid (mRNA), which is a single-stranded molecule complementary to the DNA. After processing within the cell nucleus, the mRNA moves into the cytoplasm where it binds to ribosomes. The translation of RNA into protein takes place on the ribosomes, which are small particles in the cytoplasm composed of protein and a special type of RNA, ribosomal RNA. Another type of RNA, transfer RNA, of which there is one type for each amino acid, moves amino acids from the cytoplasm to the mRNA molecule, where they are aligned in the sequence dictated by the genetic code carried in the mRNA. The sequence of amino acids forms a polypeptide, which is released from the ribosome when complete. An average protein contains about 500 amino acids; thus an average gene contains about 1500 base pairs. Since the haploid chromosome set contains 3 billion base pairs, the chromosomes must contain a large excess of DNA in addition to the DNA in the constituent genes. *See* PROTEIN; RIBONUCLEIC ACID (RNA); RIBOSOMES.

A typical gene is not a simple uninterrupted length of DNA, but most genes are made up of coding sequences (exons) separated by noncoding regions (intervening sequences, or introns). Transcription of the gene into RNA begins at an initiation site in advance of the first coding sequence and terminates beyond the end of the last sequence. After the entire DNA segment has been transcribed into mRNA, the intervening sequences are removed and the coding sequences are spliced together before the mRNA is translated into a polypeptide (**Fig. 2**). *See* EXON; INTRON.

Any gene occupies a specific chromosomal position, or locus. The alternative genes at a particular locus are said to be alleles. If a pair of alleles are identical, the individual is homozygous; if they are different, the individual is heterozygous. *See* ALLELE.

Mutation. Genetic variation has its origin in mutation. Although in broad terms any change in DNA is a mutation, whether it is a microscopically detectable change in the structure of a chromosome or a single base change in the genetic code, the term is usually applied to stable changes in DNA that alter the genetic code and thus lead to synthesis of an altered protein.

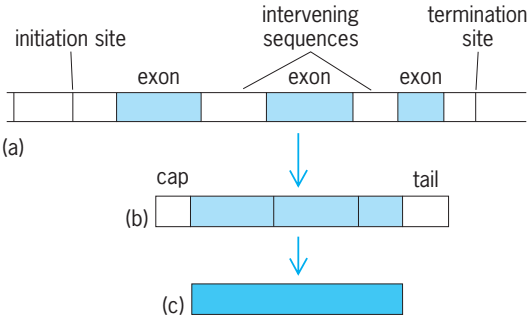

Fig. 2. Simplified diagram showing (*a*) the structure of a gene, (*b*) the mature messenger RNA after the two intervening sequences have been spliced out and the "cap" and "tail" added, and (*c*) the polypeptide translated from the messenger RNA.

Mutation can occur in reproductive cells or somatic cells, but the genetically significant ones are those that occur in reproductive cells and can therefore be transmitted to future generations. Natural selection acts upon the genetic diversity generated by mutation to preserve beneficial mutations and eliminate deleterious ones.

A very large amount of genetic variation exists in the human population. Everyone carries many mutations, some newly acquired but others inherited through innumerable generations. Though the exact number is unknown, it is likely that everyone is heterozygous at numerous loci, perhaps as many as 20%. *See* MUTATION.

Single-gene inheritance. The patterns of inheritance of characteristics determined by single genes or gene pairs depend on two conditions: (1) whether the gene concerned is on an autosome (autosomal) or on the X chromosome (X-linked); (2) whether the gene is dominant, that is, expressed in heterozygotes (when it is present on only one member of a chromosomal pair and has a normal allele) or is recessive (expressed only in homozygotes, when it is present at both chromosomes). *See* DOMINANCE.

Margaret W. Thompson

Quantitative inheritance. A quantitative trait is one that is under the control of many factors, both genetic and environmental, each of which contributes only a small amount to the total variability of the trait. The phenotype may show continuous variation (for example, height and skin color), quasicontinuous variation (taking only integer values—such as the number of ridges in a fingerprint), or it may be discontinuous (a presence/absence trait, such as diabetes or mental retardation). With discontinuous traits, it is assumed that there exists an underlying continuous variable and that individuals having a value of this variable above (or below) a threshold possess the trait.

A trait that "runs in families" is said to be familial. However, not all familial traits are hereditary because relatives tend to share common environments

as well as common genes. It is the major task of quantitative genetics to disentangle the effects of environment and heredity, but this task is not easy in humans where environment and heredity are often confounded.

The variability of almost any trait is partly genetic and partly environmental. A rough measure of the relative importance of heredity and environment is an index called heritability. R. A. Fisher showed that the total variance (a statistical measure of variability) of a trait can be partitioned into a genetic variance and an environmental variance, at least in simple cases. The quantitative model can be expressed as:

$$\text{Variance (phenotypic)} = \text{variance (genetic)} + \text{variance (environmental)} \quad (1)$$

Heritability is then defined as:

$$\text{Heritability} = \frac{\text{variance (genetic)}}{\text{variance (phenotypic)}} \quad (2)$$

In humans, the heritability of height is about 0.75. That is, about 75% of the total variance in height is due to variability in genes that affect height and 25% is due to exposure to different environments. Expressed in another way, the difference in height between two individuals is, on the average, 75% genetic and 25% environmental.

There are actually two kinds of heritability. The one described above, usually called broad heritability, measures the total effect of heredity. The other, narrow heritability, is the proportion of total phenotypic variance resulting from the additive effects of genes. Narrow heritability is a more subtle concept but is actually more useful. For example, the correlation between parent and child is equal to one-half the narrow heritability plus their environmental correlation. Equations of this sort relate heritability to observable phenotypic correlations, thus providing a means of estimating the relative importance of heredity and environment. Carter Denniston

Hereditary Diseases

Medicine is an important field for the practical application of human genetics; medical genetics has become an integral part of preventive medicine (that is, genetic counseling, including prenatal diagnostics). It has contributed increasingly to systematics of disease, diagnostics, and even therapy. Many external causes of disease, such as infections, have been brought under control in the twentieth century; therefore, doctors can devote much of their skill to treating diseases from internal sources, that is, hereditary diseases, or those brought about by interaction of genetic predispositions with certain stress factors in the environment. Widespread use of genetic knowledge in medical practice has important consequences for basic science: problems posed by the numerous and often unexpected observations in medical genetics help in developing basic theory, and suggest new approaches in research.

Hereditary diseases may be subdivided into three classes: chromosomal diseases; hereditary diseases with simple, mendelian modes of inheritance; and multifactorial diseases.

Chromosomal diseases. One out of 200 newborns suffers from an abnormality that is caused by a microscopically visible deviation in the number or structure of chromosomes. Such chromosomal aberrations are much more common (≈40%) among spontaneous miscarriages. About 10–20% of all recognized pregnancies terminate in spontaneous miscarriage and many more embryos die during the first weeks of pregnancy, when fetal loss goes unnoticed. Hence, a large fraction of all human zygotes are abnormal chromosomally and die early. *See* CHROMOSOME ABERRATION.

The most important clinical abnormality among the survivors is Down syndrome—a condition due to trisomy of chromosome 21, one of the smallest human chromosomes. The term trisomy means that this chromosome is present not twice but three times; the entire chromosome complement therefore comprises 47, not 46, chromosomes. Down syndrome occurs one to two times in every 1000 births; its pattern of abnormalities derives from an imbalance of gene action during embryonic development. Down syndrome is a good example of a characteristic pattern of abnormalities that is produced by a single genetic defect. Such patterns, recognizable to the experienced observer as syndromes, are found not only in chromosomal diseases but in hereditary diseases with simple (mendelian) modes of inheritance as well. *See* DOWN SYNDROME.

Other autosomal aberrations observed in living newborns that lead to characteristic syndromes include trisomies 13 and 18 (both very rare), and a variety of structural aberrations such as translocations (exchanges of chromosomal segments between different chromosomes) and deletions (losses of chromosome segments). Translocations normally have no influence on the health status of the individual if there is no gain or loss of chromosomal material (these are called balanced translocations). However, carriers of balanced translocations usually run a high risk of having "unbalanced" offspring—children in whom the same translocation causes gain or loss of genetic material, and who suffer from a characteristic malformation syndrome.

Clinical syndromes caused by specific aberrations vary, but certain clinical signs are common: low birth weights (small for date); a peculiar face; delayed general, and especially mental, development, often leading to severe mental deficiency; and multiple malformations, including abnormal development of limbs, heart, and kidneys. Single malformations in children who otherwise develop normally are not typical for a chromosomal aberration. *See* CONGENITAL ANOMALIES.

Less severe signs than those caused by autosomal aberrations are found in individuals with abnormalities in number (and, sometimes, structure) of sex chromosomes. This is because in individuals having more than one X chromosome, the additional X chromosomes are inactivated early in pregnancy. For example, in women, one of the two X chromosomes

is always inactivated. Inactivation occurs at random so that every normal woman is a mosaic of cells in which either one or the other X chromosome is active. Additional X chromosomes that an individual may have received will also be inactivated; in trisomies, genetic imbalance is thus avoided to a certain degree. However, inactivation is not complete; therefore, individuals with trisomies—for example, XXY (Klinefelter syndrome), XXX (triple-X-syndrome), or XYY—or monosomies (XO; Turner syndrome) often show abnormal sexual development, intelligence, or behavior.

In some individuals, chromosomal aberrations are found only in some cells. Clinical signs in these cases are often milder. Sometimes, a new mutation giving rise to a structural chromosomal aberration may occur in a somatic tissue, leading, for example, to a translocation only in one cell and its descendants. Sometimes, especially when a chromosome break involved in this translocation has affected an oncogene, these cells may have a selective advantage, and develop into a malignant tumor. An example is the translocation between chromosomes 22 and 9 found in chronic myelonic leukemia. *See* MOSAICISM; ONCOGENES.

Diagnosis of chromosomal aberrations, especially in newborns with multiple malformations, individuals with disturbances of sexual development, and parents suffering from multiple miscarriages, are of practical importance, since in many cases monitoring of future pregnancies by prenatal diagnosis is possible.

Diseases with mendelian inheritance. In contrast to chromosomal aberrations, the genetic defects in hereditary diseases with simple, mendelian modes of inheritance cannot be recognized by microscopic examination; as a rule, they must be inferred more indirectly from the phenotype and the pattern of inheritance in pedigrees. The defects are found in the molecular structure of the DNA. Often, one base pair only is altered, although sometimes more complex molecular changes, such as deletions of some bases or abnormal recombination, are involved. Methods of molecular biology have permitted in some cases direct analyses of such defects at the gene level.

Approximately 1% of all newborns have, or will develop during their lives, a hereditary disease showing a simple mendelian mode of inheritance.

Mendel called an allele "dominant" when the homozygote and the heterozygote were indistinguishable phenotypically; in experimental genetics, the same convention is still being used. In medical genetics, the terms dominant and recessive are not used so strictly. A condition is called dominant if the heterozygotes deviate in a clearly recognizable way from the normal homozygotes, in most cases by showing an abnormality. Since such dominant mutations are usually rare, almost no homozygotes are observed and their clinical condition is, in most cases, unknown. In exceptional instances, homozygotes for dominant conditions have been described; they have usually shown more severe clinical signs than the heterozygotes.

Cross between homozygotes of the same allele A			
	AA	Gametes	
AA		A	A
Gametes	A	AA	AA
	A	AA	AA
Genotypes of children: AA			

Backcross between homozygote AA and heterozygote AA′			
	AA′	Gametes	
AA		A	A′
Gametes	A	AA	AA′
	A	AA	AA′
Genotypes of children: 1AA : 1 AA′			

Cross between two heterozygotes AA′			
	AA′	Gametes	
AA′		A	A′
Gametes	A	AA	AA′
	A′	AA′	A′A′
Genotypes of children: 1 AA: 2 AA′ : 1A′A′			

Cross between homozygotes AA and A′A′			
	AA	Gametes	
A′A′		A	A
Gametes	A′	AA′	AA′
	A′	AA′	AA′
Genotypes of children: AA′			

Fig. 3. Examples of mendelian crosses using two alleles, A and A′.

In **Fig. 3** there are four mendelian crosses for one pair of alleles A and A′. Cross no. 2, the backcross between the normal homozygote and the heterozygote, which leads to 1:1 segregation, is found most commonly in autosomal dominant inheritance. Crosses 3 and 4 are extremely rare, since homozygotes for an abnormal allele are usually rare. In autosomal-recessive inheritance, backcrosses between normal homozygotes and heterozygotes (cross no. 2) are also common, but since heterozygotes are normal phenotypically, such crosses will go unnoticed in most instances. The most common cross leading to homozygous, clinically affected offspring is the intercross between two heterozygotes, AA′, which results in 1:2:1 segregation.

Autosomal-dominant inheritance. This type of inheritance pattern is often determined by studying the history of a trait among a group of relatives. Such a history, called a pedigree, can be represented in a standard chart (**Fig. 4**). A pedigree in which a rare autosomal-dominant condition is transmitted through four generations is shown in **Fig. 5**. The mutant allele is located on an autosome; the affected individuals are heterozygous. Since one of the two alleles at this gene locus is altered by mutations and each of the two alleles has a 50% probability of being transmitted to a child (Fig. 3, cross no. 2), each child has a 50% risk of being affected. Since this mutation is autosomal it is transmitted independently of the sex chromosomes and the risk of being affected is not influenced by sex of parents or child.

The pedigree in Fig. 5 demonstrates the characteristic features of autosomal-dominant inheritance. For conditions such as achondroplastic dwarfism, however, such large pedigrees are the exception rather than the rule. In most cases, known pedigrees extend over two generations only. It is also common that one person is the only affected individual in a family. These individuals owe their abnormal allele to a fresh mutation in the germ cell of one of their parents, but the risk for their children to become

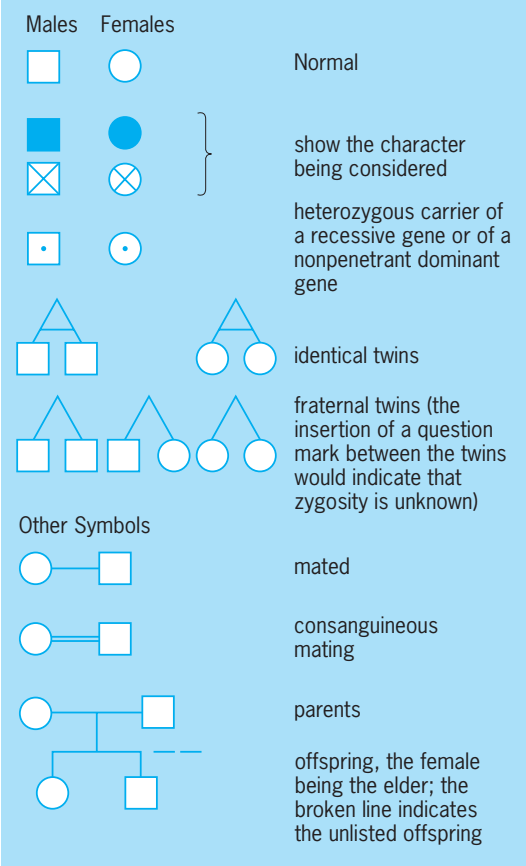

Fig. 4. Some symbols in common use in the United States for presenting human pedigrees.

affected is, nevertheless, 50% (**Fig. 6**). The fraction of new mutants among all carriers of a certain dominant disease must be high if the disease impairs average reproduction of its carriers: the harmful alleles are eventually eliminated from the population since many carriers have no children. When the carriers are so severely affected that they have no children at all, each new mutant will be eliminated in the first generation, and all cases in a population are new mutants. On the other hand, large pedigrees with many affected individuals are usually observed when the anomaly is relatively harmless or manifests itself later in life, at a time when the carriers already have had their children. An example is Huntington's disease, a severe degenerative disease of the brain: in most gene carriers, the first clinical signs become visible only between 40 and 50 years of age; they die after many years of progressive deterioration of brain function.

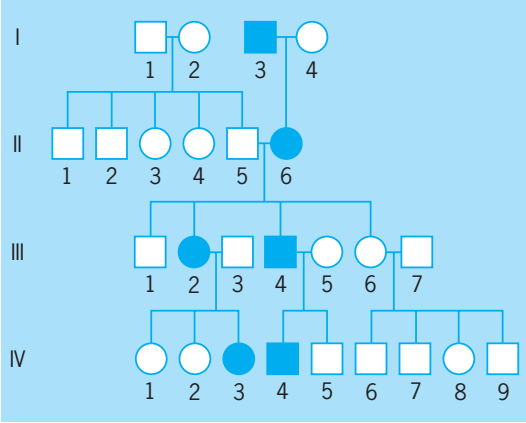

Fig. 5. Typical pedigree pattern of a rare autosomal dominant trait: classical achondroplastic dwarfism. Those with this anomaly, caused by a defect in growth of the long bones, have extremely short arms and legs, but are otherwise normal. I–IV are generations; shaded individuals manifest the trait.

In some dominant conditions, the harmful phenotype may not be expressed in a gene carrier (this is called incomplete penetrance), or clinical signs may vary in severeness between carriers (called variable expressivity). Penetrance and expressivity may be influenced by other genetic factors; sometimes, for example, by the sex of the affected person, whereas in other instances, the constitution of the "normal" allele has been implicated. Environmental conditions may occasionally be important. In most cases, however, the reasons are unknown.

Autosomal-recessive inheritance. A pedigree for albinism, that is, with an autosomal-recessive trait, is shown in **Fig. 7**. The affected individuals (IV, 2 and 6) have two mutant alleles, one from each parent; they are homozygous for the albinism gene. Hence, their unaffected parents, III, 4 and 5, must both be heterozygous for this allele. The probability of two individuals having the same allele increases when some of their genes come from a common ancestor (that is, when they are relatives). Indeed, the mothers of the parents in Fig. 4 are sisters; hence, the parents themselves are first cousins. An increase of matings between relatives, for example, first cousins, in comparison with the population average is characteristic for rare autosomal-recessive diseases. There is a segregation of 1:2:1 between homozygotes AA, heterozygotes Aa, and homozygotes aa. This means that every child from a mating of two heterozygotes (Fig. 3, cross no. 3) has a 25% risk of being homozygous for the mutant allele, and thus affected.

In the first decades of the twentieth century, when the first autonomal-recessive conditions such as albinism were discovered, both criteria—an increase of matings between relatives, and appearance of the disease in 25% of children—were often observed, and were useful for genetic analysis. In the 1980s, both indicators have become much rarer: with the average number of one to two children per marriage that is observed in many industrialized countries, there will be one affected sib only in the great majority of all sibships. Since the fraction of first-cousin marriages is only two to three per 1000 marriages, even a tenfold increase does not lead to an appreciable fraction among parents of such patients. Hence, recognizing an autosomal-recessive mode of inheritance has become more difficult.

As a rule of thumb, autosomal-dominant mutations often lead to structural anomalies such as malformation syndromes, whereas autosomal-recessive mutations can often be traced back to a biochemical abnormality, for example, an enzyme defect.

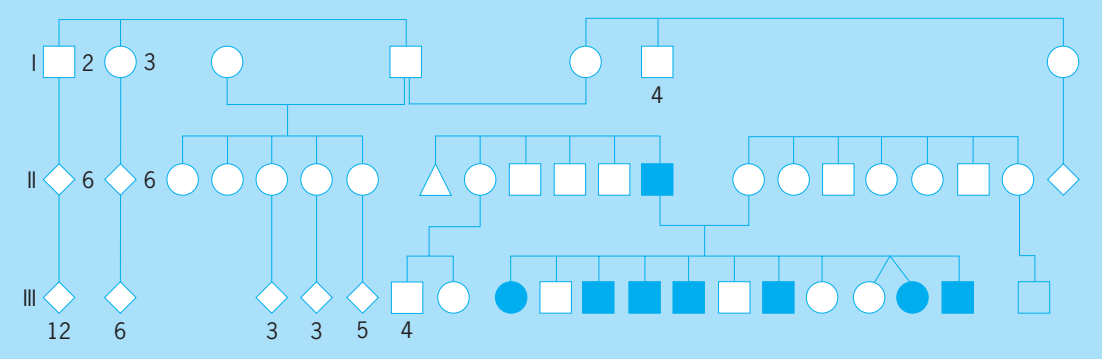

Fig. 6. Pedigree with new mutation to autosomal dominant aniridia (defect of irises). The mutation must have occurred in the germ cell of the father or mother of the patient in generation II. Diamonds represent the indicated number of unaffected offspring. (*After F. Vogel and A. G. Motulsky, Human Genetics: Patterns and Approaches, Springer-Verlag, 1979*)

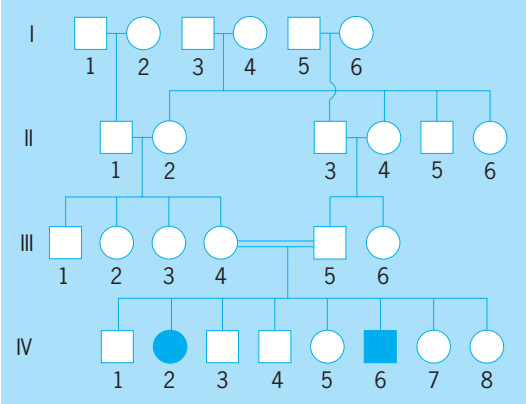

Fig. 7. Typical pedigree pattern of a rare autosomal recessive trait, one form of albinism. Note, however, that sibships with so many sibs as well as marriages between first cousins have become rare in populations of industrialized countries.

X-linked inheritance. X-linked modes of inheritance occur when the mutant allele is located on the X chromosome. The family patterns result from the mechanism of phenotypic sex determination: children receiving a Y chromosome from their fathers become males, and those receiving the paternal X become females. The mothers contribute an X chromosome to children irrespective of their sex. Therefore, all daughters and no sons receive an X-linked mutant allele from their fathers, whereas sons as well as daughters may receive such an allele from the mothers, the chance always being 50%, if the mother is heterozygous for the trait.

The most important X-linked mode of inheritance is the recessive one (**Fig. 8**). Here, the males (referred to as hemizygotes since they have only one allele) are affected, since they have no normal allele. The female heterozygotes, on the other hand, will be unaffected, since the one normal allele is sufficient for maintaining function. A classical example is hemophilia A, in which one of the serum factors necessary for normal blood clotting is inactive or lacking. (The disease can now be controlled by repeated substitution of the deficient blood factor—a good example for phenotypic therapy of a hereditary disease by substitution of a deficient gene product.) As shown in Fig. 8*a*, male family members are affected whereas their sisters and daughters, while being unaffected themselves, transmit the mutant gene to half their sons. Only in very rare instances, when a hemophilic patient marries a heterozygous carrier, are homozygous females observed (Fig. 8*b*). As for autosomal dominant diseases, the classic pedigrees are rare for severe X-linked conditions. Again, many of them are new mutants, and most actual pedigrees are small.

Some X-linked conditions are dominant, such that female heterozygotes express the abnormal trait. As a rule, however, their clinical signs are milder than those found in male hemizygotes. In quite a few such conditions, male hemizygotes are so severely damaged that they die even before birth, and are aborted. Hence, (almost) exclusively female patients are observed; there is a 1:1 ratio of normal and affected daughters of affected mothers, sex ratio among liveborn children is shifted in favor of girls, and

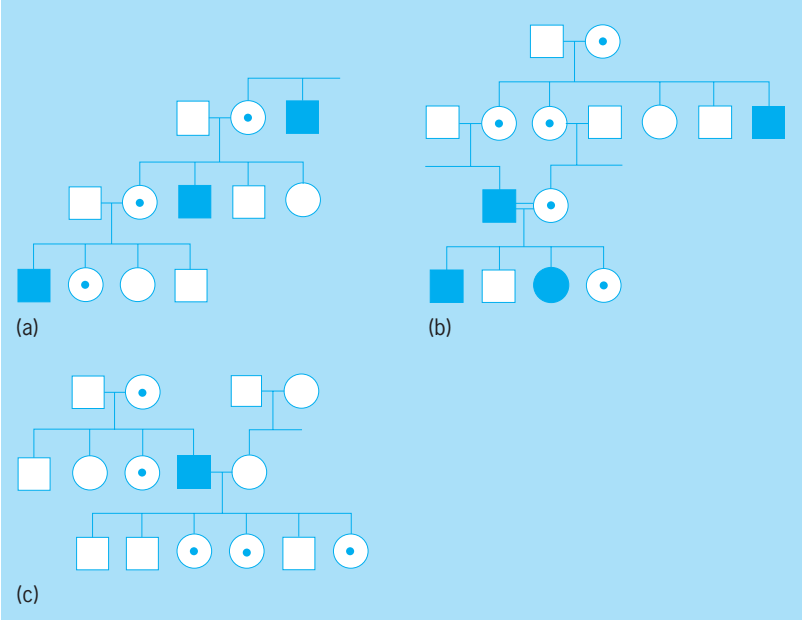

Fig. 8. Typical pedigree patterns of a sex-linked recessive trait, hemophilia A. (*a*) Only males are affected and usually come from unaffected mothers who often have affected fathers or brothers. (*b*) An affected female can result from an affected father and a carrier woman. (*c*) All daughters of affected males are carriers.

miscarriages are common. Mutant genes located on the Y chromosome would be expected to be transmitted from the father to sons only, but to all sons. No such example has been confirmed. *See* SEX-LINKED INHERITANCE.

Multifactorial diseases. There are thousands of hereditary diseases with simple mendelian modes of inheritance, but most common anomalies and diseases are influenced by genetic variability at more than one gene locus. Most congenital malformations, such as congenital heart disease, cleft lip and palate, neural tube defects and many others, fall into this category, as do the constitutional diseases, such as diabetes mellitus, coronary heart disease, anomalies of the immune response and many mental diseases, such as schizophrenia or affective disorders.

All of these conditions are common and often increase in frequency with advanced age. In industrialized societies, almost everybody dies sooner or later of such a multifactorial disease, in which neither a chromosome aberration nor a simple mode of inheritance can be detected. The influence of genetic constitution can be indicated in a variety of ways, including a higher concordance of monozygotic as compared with dizygotic twins; a higher incidence of the same disease among relatives of affected probands than in the general population; and similarity of adopted children with their biological, not their adoptive, parents. However, there is no simple 1:1 relationship between genotype and phenotype in these cases; a variety of genetic factors may be involved. Moreover, the environment usually contributes significantly to the disease risk, which can be concluded from the observation that monozygotic twins, while being affected concordantly more often than dizygotics, are discordant in many cases.

In the absence of a more penetrating analysis, such a situation is described preliminarily by the genetic model of multifactorial inheritance in populations with a threshold effect (**Fig. 9**). The conceptual background of the multifactorial model may be described as follows: many undefined genes cooperate in creating a disease liability, which is distributed normally among individuals of a population. When this liability exceeds a threshold, the individual will be clinically affected. Instead of a sharp threshold, as shown in Fig. 9, a threshold area may be assumed in which manifestation depends on additional environmental factors. From such a genetic model, and the theory of quantitative genetics, some conclusions may be derived that are often found to apply more or less to empirical data; for example, that close relatives are affected more often than remote relatives, or that relatives of more severely affected probands run a higher risk than those of less severely affected ones. However, the multifactorial model does not attempt to identify, much less to characterize, the individual genes involved in a disease liability. The genotype is treated as "black box."

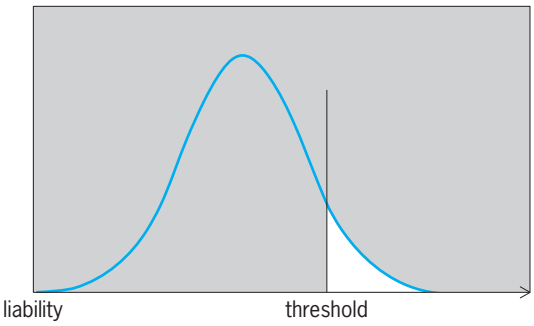

Fig. 9. Multifactorial inheritance in combination with a threshold effect. The figure shows the distribution of disease liabilities in a population (ordinate: number of individuals; abscissa: degree of disease liability). This distribution is assumed to be normal: many individuals have an average liability; in a few, liability is low, and in some, it is high. Individuals having liabilities on the right-hand side of the threshold (white area) are affected. (*After F. Vogel and A. G. Motulsky, Human Genetics: Problems and Approaches, Springer-Verlag, 1979*)

The multifactorial model poses questions rather than answering them. A meaningful answer can come only from analysis of the contribution of single genes (and specific environmental factors) to such disease liabilities. Some such analyses have been successful. The ABO blood groups, for example, influence liability to many common diseases, such as stomach cancer and cancers of some other organs; variants of the major histocompatibility gene (HLA) contribute to the liability for some rheumatic and autoimmune disease; and alpha$_1$-antitrypsin variants are involved in chronic obstructive pulmonary disease. Here, the risk increases when the carriers of such alleles suffer from repeated bronchial infections—for example, heavy cigarette smokers or laborers working in a dusty environment. This is a good example of a specific interaction between a genetic liability and an environmental stress factor. Such examples are studied in a special branch of medical genetics that is known as ecogenetics.

Heterogeneity for autosomal-recessive mutations, while not leading normally to disease, may contribute to disease liabilities under certain external conditions. Often, specific disease units with clearcut mendelian modes of inheritance have been identified by a combination of clinical, biochemical, and genetic methods in a minority of individuals and families within a bulk of conditions hitherto described as multifactorial. There are good chances that future research will help to open the "black boxes" of multifactorial inheritance step by step by analysis of single genes, the mechanisms of their action, and the ways in which they interact with each other and with specific stress factors in the environment.

Friedrich Vogel

Population Genetics

Population genetics is the mathematical basis of evolutionary theory. It is concerned with the frequency of genes and genotypes in a population, their relationship, and how they change over time. A major concern is the frequency of alleles, which are different forms of a gene or DNA sequence. Many principles of population genetics are easily understood using a simple model of a locus with two alleles, A and a. There are three possible genotypes: AA, Aa,

and *aa*. A starting point in population genetics is to determine the relative frequency of each allele. It is conventional to use the symbol p to refer to the relative frequency of the *A* allele, and the symbol q to refer to the relative frequency of the *a* allele. If each genotype can be uniquely identified, the allele frequencies can be determined by simple counting. The number of individuals with each genotype are first counted and expressed as a relative frequency. If, for example, there are 18 individuals with genotype *AA*, 24 with genotype *Aa*, and 8 with genotype *aa*, there will be a total of 50 individuals, and the relative frequencies will be $AA = 18/50 = 0.36$, $Aa = 24/50 = 0.48$, and $aa = 8/50 = 0.16$. The frequency of the *A* allele is determined by taking the frequency of genotype *AA* plus half the frequency of genotype *Aa*, giving $p = 0.36 + 0.48/2 = 0.6$. Likewise, the frequency of the *a* allele is determined by taking the frequency of genotype *aa* plus half the frequency of genotype *Aa*, giving $q = 0.16 + 0.48/2 = 0.4$. Note that $p + q = 1$.

Genotype frequencies. Given allele frequencies p and q, it is possible to predict the genotype frequencies in the next generation under different assumptions regarding mating.

Panmixis. Panmixis refers to random mating with respect to genotype. Under this model, any male of reproductive age is equally likely to mate with any female of reproductive age in terms of the particular gene or trait being analyzed. In other words, a male with genotype *AA* is just as likely to mate with a female of genotype *AA* as with a female of genotype *Aa* or *aa*. Under these conditions, the distribution of genotype frequencies in the next generation is a function of probability, known as Hardy-Weinberg equilibrium.

Hardy-Weinberg equilibrium. Given allele frequencies p and q, corresponding to alleles *A* and *a*, the expected genotype frequencies under panmixis in the next generation are $AA = p^2$, $Aa = 2pq$, and $aa = q^2$. Under Hardy-Weinberg equilibrium, the genotype and allele frequencies will remain constant generation after generation. Hardy-Weinberg equilibrium provides a baseline model from which to make predictions about nonrandom mating and evolutionary change.

Nonrandom mating. If individuals are not mating at random with respect to genotype, then there is no panmixis. There are two basic forms of nonrandom mating: inbreeding and assortative mating. Inbreeding is the mating of individuals that are closely related. In humans, this usually means a couple that are more closely related than third cousins are. Inbreeding changes genotype frequencies by increasing the proportion of homozygotes (*AA, aa*) and decreasing the proportion of heterozygotes (*Aa*). Assortative mating refers to nonrandom mating based on phenotype. Positive assortative mating refers to mating between individuals that are phenotypically similar, such as between two tall people or two people with blond hair. Positive assortative mating also increases the frequency of homozygotes. Negative assortative mating occurs when mating is between individuals that are phenotypically different, and will result in a decrease in homozygotes. Overall, nonrandom mating affects the genotype frequencies but does not change the allele frequencies.

Evolutionary forces. Allele frequencies will remain the same, generation after generation, under both panmixis and nonrandom mating. Evolution, defined here as a change in allele frequency over time, can be caused by four evolutionary forces: mutation, natural selection, gene flow, and genetic drift. Nonrandom mating does not lead to allele frequency changes, although it can affect the rate at which allele frequencies change.

Mutation. Mutation is a random change in the genetic code. It can consist of a small point mutation, where a single nucleotide is changed in the DNA, or larger units of DNA may be changed. Mutation rates are often very difficult to estimate, but most seem to be rather low, with probabilities ranging from about 10^{-4} to 10^{-8} per locus per generation. Mutation introduces new genetic variants into a population and is thus required for any evolutionary change. By itself, however, mutation does not cause rapid evolutionary change; the other evolutionary forces act upon this new variation to increase or decrease the frequency of a mutant allele.

Natural selection. Each genotype can be characterized by its relative probability of survival and reproduction, known as fitness (w), and the probability of not surviving or reproducing, known as the selection coefficient (s), such that $w + s = 1$. For example, if the probability of an individual with genotype *aa* surviving and reproducing is one-fourth that of individuals with genotypes *AA* or *aa*, then the fitness of $aa = 0.25$ and the selection coefficient of $aa = 1 - 0.25 = 0.75$. Fitness values depend on characteristics of the specific locus and the local environment. Hardy-Weinberg equilibrium assumes that the fitness values of all genotypes are the same. When this is not the case, then natural selection can cause changes in allele frequency through the process of differential survival. There are several different types of selection, each with different outcomes.

Selection against recessive homozygotes. This occurs when the allele *a* is recessive and the recessive homozygous genotype (*aa*) has a lower fitness relative to the other genotypes (*AA*, *Aa*), which have equal fitness. Under this condition, the frequency of the *a* allele will decline over time to approach a value of zero (although balanced to some extent by continuing mutation from *A* to *a*). The change in allele frequency per generation is a function of the selection coefficient (s) for *aa* and the initial frequency (q) of the *a* allele. The frequency of *a* will decline each generation by the amount $-spq^2/(1 - sq^2)$. When the allele *a* is lethal, as is the case with many genetic diseases, then $s = 1$ (all individuals with *aa* are selected against), and the decrease in *a* each generation is equal to $-q^2/(1 + q)$. The lethal allele will not be eliminated in a single generation because even though all individuals with *aa* are selected against, individuals with genotype *Aa* continue to contribute *a* alleles into the population.

Selection for recessive homozygotes. In this case, the fitness of the recessive homozygote is higher than that of the other two genotypes, and the frequency of the dominant allele will decline over time. If the dominant allele is completely lethal, then it will be eliminated in a single generation. If it is not completely lethal, then it will take time for the dominant allele to be reduced to near zero.

Selection against the heterozygote. In this case, the fitness of the heterozygote (*Aa*) is lower than the fitness of the homozygous genotypes (*AA*, *aa*). Both alleles will be selected against, and the frequency will change in the direction of the initially most common allele. For example, if selection against the heterozygote starts from a condition where $p = 0.7$ and $q = 0.3$, then selection will drive the population to the state of $p = 1.0$ and $q = 0.0$. If, however, p is initially less than q, then the reverse will occur. For example, if selection against the heterozygote starts with $p = 0.2$ and $q = 0.8$, then selection will ultimately result in $p = 0.0$ and $q = 1.0$. In the unlikely event that the initial allele frequencies are exactly equal ($p = q = 0.5$), the allele frequencies will stay the same until another evolutionary force changes them initially.

Balanced polymorphisms and selection for the heterozygote. All of the above examples involve allele frequencies moving toward 1 or 0. In some cases, the ultimate fate of natural selection is a balance between selective forces producing a set of intermediate allele frequencies that remain in equilibrium. Such balanced polymorphisms result when the heterozygote is selected for because the fitness of the heterozygote is higher than that of the two homozygous genotypes. Selection against *AA* results in the elimination of some *A* alleles, while selection against *aa* results in the elimination of some *a* alleles. At the same time, the selection for the heterozygote results in maintaining some *A* alleles and some *a* alleles. Unlike the previous examples of selection, selection for the heterozygote simultaneously selects for and against both alleles. The result is a balance in allele frequencies determined by the fitness values of the homozygotes. When the heterozygote is the most fit, it has, by definition, a relative fitness of 1. Given a fitness of $1 - s$ for genotype *AA* and a fitness of $1 - t$ for genotype *aa*, selection for the heterozygote will reach an equilibrium where $p = t/(s + t)$ and $q = s/(s + t)$. These values are the allele frequencies where overall survival is maximized. The classic example of a balanced polymorphism is human hemoglobin, a locus with a normal hemoglobin allele (*A*) and the sickle-cell allele (*S*). Individuals with the genotype *SS* have the genetic disease sickle-cell anemia, which is frequently fatal early in life. In many populations, this disease leads to selection against the *SS* genotype and the removal of the *S* allele. In environments with epidemic malaria, heterozygous people (*AS*) are the most fit because they have greater resistance to malaria due to the presence of one *S* allele, but do not suffer from sickle-cell anemia. As a result, the *S* allele is maintained at a relatively high frequency (usually 5–20%) in malarial environments.

Gene flow. Gene flow occurs when populations share genes through the process of mating with someone in another population. An allele absent in one population can be introduced from migrants from another population where the allele is present. Over time, gene flow acts to make populations increasingly genetically similar. Gene flow helps keep populations within a single species; for new species to form, gene flow must be eliminated or severely reduced. Further genetic change must occur for a new species to form, since reduction of gene flow is a necessary, though not sufficient, cause of speciation.

Genetic drift. Genetic drift refers to random fluctuations in allele frequency from one generation to the next because of chance. As an example, consider that if you flip a coin 10 times you expect to get, on average, 5 heads and 5 tails. Because of chance, you might actually get 3 heads and 7 tails, 8 heads and 2 tails, or any other possible combination adding up to 10. Reproduction in a population works in the same manner, so that alleles may not be passed along to the next generation in the exact same frequencies as in the parental generation. If p and q are the allele frequencies in the parental generation, the average expectation in the next generation is also p and q, but the variance around these values is equal to $pq/2N$, where N is the number of reproductive adults. When population size is very large, the variance will be small, so that there is little chance for genetic drift in a single generation. When N is relatively small (generally less than 200), the variance is large and there is a greater chance for genetic drift. Genetic drift is not directional; an allele frequency can increase, decrease, or remain the same. The probability of change is related to population size. Over time, all populations drift, and will continue to do so until, by chance, an allele becomes extinct (frequency = 0) or fixed (frequency = 1).

Interaction of evolutionary forces. The four evolutionary forces have been discussed one at a time. In reality, all four forces can operate simultaneously and can interact with each other in different ways. For example, a new mutation can increase or decrease in frequency because of natural selection and/or genetic drift, and it can be transmitted to another population via gene flow. In some cases, a harmful allele can actually increase in frequency because of genetic drift. New neutral mutations are frequently eliminated by genetic drift, but may occasionally increase because of random chance and become fixed in a population. Some evolutionary forces, such as drift and selection, will act to make populations genetically different, whereas gene flow can counter such differences. The interaction of evolutionary forces, and their net effect on allele frequency change, is the focus of many methods and studies of population genetics.

John H. Relethford

Biochemical Genetics

Biochemical genetics began with the study of inborn errors of metabolism. These are diseases of the body chemistry in which a small molecule such as a sugar

or amino acid accumulates in body fluids because an enzyme responsible for its metabolic breakdown is deficient. This molecular defect is the result of mutation in the gene coding for the enzyme protein. The accumulated molecule, dependent on its nature, is responsible for the causation of a highly specific pattern of disease. There was explosive growth of knowledge in this field with the discovery of many inborn errors of amino acid metabolism when widespread methods of amino acid metabolism followed the recognition that amino acids could be separated chromatographically and that they could be identified and quantified because they became purple when reacted with ninhydrin. More recently the development of gas chromatography–mass spectrometry led to logarithmic growth in the elucidation of organic acidemias (abnormal acidity of the blood) and disorders of fatty acid metabolism.

The field of biochemical genetics expanded during this period with the recognition that similar heritable defective enzymes interfere with the breakdown of very large molecules, such as mucopolysaccharides and the complex lipids that are such prominent components of brain substance. The resultant storage disorders present with extreme alterations in morphology and bony structure and with neurodegenerative disease.

Advances in the methodology of molecular biology have permitted the study and detection of disease at the level of the mutation in the DNA. This capability continues to broaden the scope of biochemical genetics. It has been particularly rewarding in the analysis of mutations in the mitochondrial genome. This technology is elucidating a previously unrecognized spectrum of mitochondrial disorders, in which abnormal energy metabolism is often characterized clinically by lactic acidemia.

Inheritance mechanism. The majority of hereditary disorders of metabolism are inherited in an autosomal recessive fashion. In these families, each parent carries a single mutant gene on one chromosome and a normal gene on the other. Most of these mutations are rare. In populations where consanguinity is common, rare recessive diseases are seen with relative frequency, and affected individuals are homozygous for the same mutation. In populations with more genetic diversity, most affected individuals carry two different mutations in the same gene. Some metabolic diseases are coded for by genes on the X chromosome. Most of these disorders are fully recessive, and so affected individuals are all males, while females carrying the gene are clinically normal. Some genes, such as that for ornithine transcarbamylase (an enzyme of the urea cycle), express as X-linked dominants in which most females are detectable as metabolically abnormal, and some are affected as severely as males. The disorders that result from mutations in the mitochondrial genome are inherited in nonmendelian fashion because mitochondrial DNA is inherited only from the mother. Those that carry a mutation are heteroplasmic; that is, each carries a mixed population of mitochondria, some with the mutation and some without. Expression of clinical disease and its degree of severity are functions of the numbers of abnormal genomes, and they may differ greatly among siblings, because the ovum results from a funneling that may lead to concentration or dilution of the prevalence of the mutation in maternal cells.

Inborn errors of amino acid metabolism. Phenylketonuria (PKU) is a prototypic biochemical genetic disorder. It is an autosomally recessive disorder in which mutations demonstrated in a sizable number of families lead, when present in the genes on both chromosomes, to defective activity of the enzyme that catalyzes the first step in the metabolism of phenylalanine. This results in accumulation of phenylalanine and a recognizable clinical disease whose most prominent feature is severe retardation of mental development. This was the disease with which programs of neonatal screening were developed throughout the world. Determination of the concentration of phenylalanine in a spot of blood obtained from the heel of the baby permits a diagnosis of phenylketonuria early enough to initiate a diet sufficiently restricted in its content of phenylalanine that mental retardation is prevented. This was a wonderful development in public health and preventive medicine. Programs of neonatal screening are expanding throughout the developed world as new methodology permits the detection and treatment of many more inborn errors of metabolism.

The mutations that cause phenylketonuria result in deficient activity of the enzyme phenylalanine hydroxylase, which normally catalyzes the conversion of phenylalanine to tyrosine. Thus, levels of phenylalanine soar and those of tyrosine diminish. When phenylalanine concentrations are high, the amino acid is converted to a number of products including phenylpyruvic acid, a phenylketone. A green color results when ferric chloride is added to solutions containing phenylpyruvic acid, and this once led to the detection of the compound in the urine of two mentally retarded siblings, the index cases of phenylketonuria. Patients with this disease are typically blond and blue-eyed, because the abnormal chemistry interferes with the development of pigment, in addition to affecting brain development. *See* PHENYLKETONURIA.

X-linked metabolic disorders. The most commonly encountered of the inborn errors of purine metabolism is the Lesch-Nyhan disease. The gene is on the end of the long arm of the X chromosome. Very many mutations have been discovered, a distinct one for virtually every affected family, and most mutations lead to a complete absence of enzyme activity. The enzyme is hypoxanthine, guanine phosphoribosyl transferase (HPRT), which catalyzes the conversion of hypoxanthine and guanine to their respective nucleotides, inosinic acid and guanylic acid. Inosinic acid is also converted to guanylic acid and to adenylic acid, and adenylic acid and guanylic acid are converted to adenosine triphosphate and guanosine triphosphate and their deoxynucleotides—the building blocks of the nucleic acids, RNA and DNA. HPRT is called a salvage enzyme because it is used in the

reclaiming of purines resulting from the breakdown of cellular DNA and RNA. The salvage pathways contrast with the de novo pathway of purine nucleotide synthesis in which inosinic acid is made in stepwise fashion from small molecules such as carbon dioxide, ammonia, and glycine. When HPRT is deficient, the phosphoribosylpyrophosphate that serves as the ribose phosphate to form inosinic acid and guanylic acid accumulates and drives the de novo pathway into overactivity; the excess purine synthesized ends up as uric acid.

In this way, patients with HPRT deficiency have high concentrations of uric acid in blood and urine, as do patients with gout. Like patients with gout, these patients may develop acute arthritis, kidney stones, and renal failure. They also have a distinct neurologic disease in which retardation of motor development is associated with spasticity like that of patients with cerebral palsy, and involuntary movements and posturing called chroeoathetosis and dystonia. An even more striking feature of this disease is that the patients develop compulsive-aggressive behavior, the most prominent aspect of which is self-injury through biting. Most patients have loss of tissue about the lips, and most bite their fingers, sometimes with partial amputation. The disease promises to be rewarding in providing chemical explanations of behavior, but so far the linkage between the enzyme defect and the behavioral phenotype has been elusive.

Mitochondrial disease. The diseases that result from mutation in mitochondrial DNA have been recognized as such only since the 1990s. They result from point mutations, deletions, and other rearrangements. It seems likely that there are many more entities yet to be discovered. A majority of these disorders express themselves chemically in elevated concentrations of lactic acid in the blood or cerebrospinal fluid. The importance of the processes of energy metabolism to the central nervous system is underscored by the fact that levels in the cerebrospinal fluid are often higher than those in the blood and may be elevated in cerebrospinal fluid when the blood is normal. Many of the disorders are known as mitochondrial myopathies (diseases of muscles) because skeletal myopathy or cardiomyopathy are characteristic features. The histologic hallmark of these mitochondrial myopathies is the presence of ragged red fibers when the tissue is processed with the Gomori stain. These fibers result from the aggregation of mitochondria, which increase in quantity as compensation for diminished function, and which appear abnormal in size and shape when viewed with electron microscopy. Many of these disorders are known by acronyms, such as MERRF, which stands for mitochondrial encephalomyelopathy with ragged red fibers. An example of disease that results from mutation in mitochondrial DNA is mitochondrial encephalomyelopathy, lactic acidemia, and strokelike episodes (MELAS).

MELAS present classically with strokelike episodes. Myopathy may have been present earlier, as elucidated by a careful history, or the patient may have been unaware of it until it is discovered upon physical examination. The episode is clinically a typical stroke with hemiplegia (paralysis on one side of the body) or other neurologic manifestation dependent on the area of brain infarcted. They are called strokelike episodes because there is no demonstrated occlusion of a vessel. Consistent with this, the symptoms are sometimes transient, but they also may leave permanent neurologic evidence. Patients are characteristically short in stature. Levels of lactic acid, while elevated, are often not very high; levels in the blood may be normal, but those in the cerebrospinal fluid are seldom normal. Some patients have convulsions, and ultimately there is evidence of encephalopathy and neurodegeneration. Dementia is a late consequence. Additional features are migraine and noninsulin-dependent diabetes mellitus. The disease is generally recognized by the occurrence of the classic picture in a family member. Once the mutation is identified, it is then usually found in other members of the family, some of whom are asymptomatic, while others may have diabetes or migraine and no other signs of disease. The disease is caused by a point mutation in the mitochondrial gene for the transfer RNA for the amino acid leucine. The mutation generally interferes with protein synthesis in the mitochondria, disrupting the activity of a number of enzymes of the electron transport chain.

Significance. Biochemical genetics has been a rich source of insight into fundamental principles in biology and medicine. The concept that biochemical and clinical abnormality could result from a defect in a single enzymatic step in metabolism was first enunciated by A.E. Garrod around the turn of the century. This was the first statement of the one gene–one enzyme hypothesis in which genes function to determine the structure of proteins, whose structure determines function. These concepts, developed long before the elucidation of the DNA as the genetic material, have stood the test of time. The functions of mitochondrial DNA and the variety of disease its mutation causes represent a recent chapter in the growth of knowledge. William L. Nyhan

Chromosome Mapping

Human chromosome mapping, the localization of human genes to specific chromosomes or regions of chromosomes, has undergone explosive growth since the mid-1970s. Nearly 1500 genes have been assigned to their respective autosomes, well over a hundred genes to the X chromosome, and several functional genes to the Y chromosome. Assuming that the total number of human genes is about 50,000, this represents slightly more than 3% of the total number. There seems no reason in principle why a complete human gene map should not eventually be achieved. The current state of the map is summarized in **Fig. 10**. For each chromosome only a few genes, usually those of some clinical interest, have been named. A finding of general interest which has emerged is that the genetic maps of the great apes are almost identical to those of humans, and

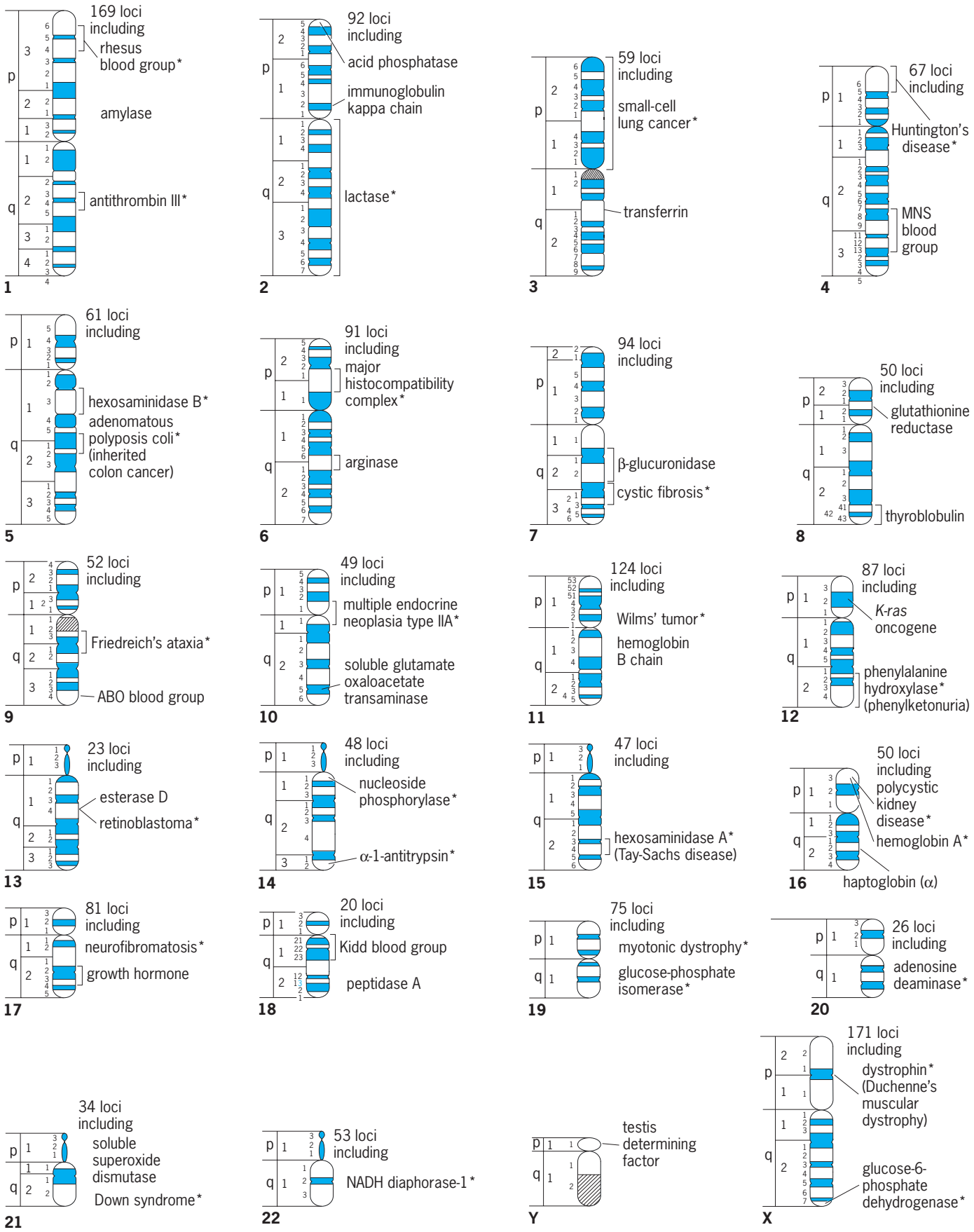

Fig. 10. Diagrammatic representation of the human chromosomes as observed by banding techniques. The number of genes already mapped to each chromosome is indicated beside it, and some examples of particular gene loci are shown. The exact position of these loci is marked where this is known. An asterisk indicates that the locus so marked is the site of one or more mutations associated with inherited disease or is of major clinical importance, such as ABO blood groups.

many groups of genes which are syntenic (on the same chromosome) in humans are syntenic even in the mouse. *See* PROTEINS, EVOLUTION OF.

Most of the information about localization of human genes has been obtained either by somatic cell hybrids or by family studies. Advances in DNA technology have greatly increased the potential of both methods, and have also allowed the development of in situ hybridization, which now accounts for at least one-third of new gene assignments.

Somatic cell hybrids. If human and rodent somatic cells are grown together in tissue culture and treated with certain viruses or chemicals, it is possible to select hybrid uninucleate cells which have resulted from the fusion of human with rodent cells. The most commonly used combinations are human and mouse, or human and Chinese hamster. In both these types of hybrids, on extended subculture, some of the human chromosomes are usually lost preferentially. It is therefore possible to obtain a set of clonal (derived from a single cell) hybrid cell lines, each containing a full set of rodent chromosomes together with a unique subset of human chromosomes.

Any gene whose product is distinguishable in human and rodent and which is expressed in cultured cells can be assigned to a particular chromosome by testing a relatively small number of hybrids for the presence or absence of a given human gene product. This can then be correlated with the presence or absence of some other human gene product or chromosome. The mouse and human gene products, most of which are proteins, can often be separated by electrophoresis. This relies on the fact that the homologous proteins in the two species are not identical but carry slightly different charges owing to differences in their amino acid sequences. **Figure 11** shows an example of electrophoresis of adenylate kinase in human and hamster controls and in six independent human-hamster hybrids. This experiment shows that the two major forms of adenylate kinase on the human cells are coded by different genes on different chromosomes. By chromosome analysis of a number of independently obtained hybrids, one can correlate the synthesis by the hybrid cells of either or both forms of the enzyme with the presence of two particular human chromosomes: actually, one gene (AK_1) is on chromosome 9, the other (AK_2) on chromosome 1. Many gene products distinguished in other ways, such as cell surface antigens, have also been mapped by using such hybrids.

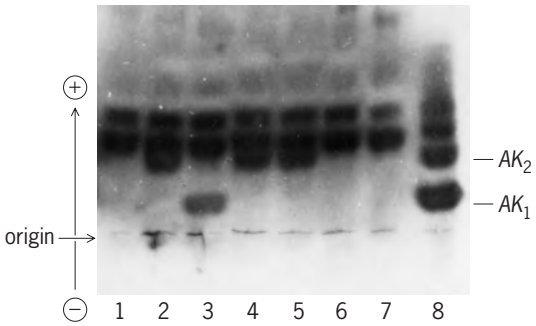

Fig. 11. Photograph of adenylate kinase isozymes separated by starch gel electrophoresis in human control (channel 8), hamster control (channel 7), and six independent hybrid clones, showing independent segregation of human AK_1 (positive in the hybrid in channel 3) and AK_2 (positive in the hybrids in channels 2, 4, and 5).

Originally the only genes that could be mapped in hybrids were those whose products could be detected in these cells. These would not, for example, include genes such as those coding for hemoglobin or insulin which are only produced in specialized cell types, although presumably the genes themselves are present in all cells. But it is now possible to isolate many of the genes themselves. The DNA sequence can then be labeled, usually with radioactivity, and used as a probe to search for similar sequences in the hybrids. The presence of the human sequence is then correlated with that of a particular chromosome in exactly the same way as described above for gene products.

This approach is not confined to DNA sequences of known function, and more than 3300 random DNA fragments have also been assigned to particular chromosomes. Although these are of less intrinsic interest they are proving invaluable as genetic markers in family studies. If random fragments from some particular chromosome are needed, these can be generated either from a hybrid containing the appropriate single human chromosome, or by direct sorting of the human chromosomes using a fluorescent activated cell sorter. In both cases it is necessary to check with a hybrid panel that the fragments obtained are indeed from the right chromosome.

The use of human parental cells containing rearranged chromosomes such as translocations or deletions, or the observation of spontaneous or induced chromosome rearrangements arising in hybrid cells, also allows the assignment of genes to particular regions of chromosomes. *See* SOMATIC CELL GENETICS.

In situ hybridization. The chromosomes of normal human lymphocytes can be readily visualized during the metaphase stage of cell division. In the case of genes involving localized repeated sequences, a DNA probe radioactively labeled to high activity can be used directly for hybridization to specific chromosomes in metaphase spreads, which are then autoradiographed. It has also become possible to label probes with sufficient specific activity to detect sequences present only as single copies (such as most of those coding for proteins), and thus, in one experiment, to localize the gene to a particular region of a particular chromosome. Developments in nonradioactive labeling of DNA with fluorescent dyes and the use of a confocal laser scanning microscope have greatly improved the precision of gene localization by this technique.

Family studies. Unlike somatic cell hybrids, gene mapping by family studies depends on finding genetic differences between people and observing the way these differences are inherited. A special case is the X chromosome because most genes on the X chromosome are inherited in a distinctive way—they never pass from father to son. It is therefore relatively

easy to establish that a particular disease or distinctive trait is X-linked; for a long time, many such diseases have been recognized. A well-known example is the hemophilia gene that was carried by Queen Victoria and caused the disease in many of the royal families of Europe. The assignment of genes to particular autosomes by family studies presents a more formidable problem and it was only in 1969 that the first confident assignment, that of the Duffy blood group to chromosome 1, was made.

For many years, workers have analyzed the patterns of inheritance of two or more gene loci in the same family to establish whether the genes are linked, that is, carried on the same chromosome pair close enough together to show nonindependent segregation. If the genes are linked, a measure of the intergene distance can be derived by the frequency with which they are separated by crossing over. A linkage group found in this way was that of the ABO blood groups with one of the forms of adenylate kinase. The assignment of the adenylate kinase gene to chromosome 9 by somatic cell hybrids thus allowed the indirect assignment of the gene determining the ABO blood group to the same chromosome. *See* CROSSING-OVER (GENETICS).

Direct assignment of genes to chromosomes by using family studies can be accomplished by studying concurrently the inheritance of genetically determined traits and of chromosome polymorphisms or rearrangements. Individuals carrying unbalanced chromosome abnormalities involving deletions or duplications of specific regions of chromosomes can also be useful for mapping by gene dosage. However, quantitative differences indicating one, two, or three copies of a gene are difficult to distinguish from secondary effects, and from unusual normal alleles which code for enzymes of different activity. This method has been most successful in regional localization when the gene is already known to lie on a particular abnormal chromosome.

The efficiency of family studies in assigning genes to chromosomes depends very much on the proportion of the human genome which is within measurable genetic distance of "good" genetic markers. In this context a good genetic marker might be defined as one determined by a single gene locus, easily typed and showing a large amount of normal variation, so that a high proportion of individuals are heterozygous. Before the advent of DNA technology, the genetic markers available have been limited and in practice have been confined to those expressed in readily obtained tissues such as blood. Thus the chance of assigning any new gene to its chromosome even from study of a large number of families was small, one estimate being about 8%. Even in this early work, however, family studies had an essential role in establishing the fine details of gene arrangement, and in the construction of a true genetic map where the known amount of recombination between two genes gives a definite value or map distance. In most chromosome regions it has been found that recombination, and hence map distance, between genes is almost twice as great in females as in males.

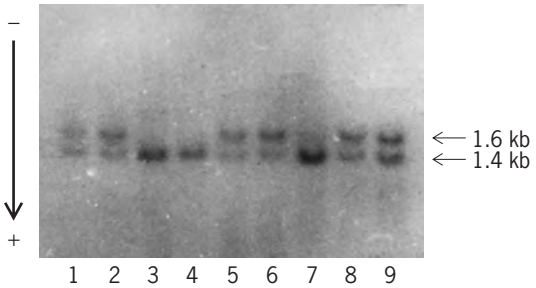

Fig. 12. Autoradiograph showing polymorphism defined by a random DNA fragment from human chromosome 7 and the restriction enzyme *HinfI*. Three individuals (tracks 3, 4, 7) are homozygous for the common allele producing a fragment 1400 bases long (1.4 kb). The remaining samples are from heterozygotes in which one chromosome produces the common pattern and the other a longer fragment (1.6 kb). (*Photograph courtesy of Dr. Ben Carritt*)

In the 1980s there was a dramatic increase in the number of genetic markers available, allowing the construction of at least partial genetic maps of all the human chromosomes. As a result, most of the human genome is now within reach of known genetic markers. This has come about through the use of DNA probes in conjunction with bacterial enzymes known as restriction endonucleases. These enzymes recognize certain specified short sequences of DNA, usually 4 or 5 bases long, and then cut the DNA at all the places where this sequence occurs. Many different enzymes can be extracted from different bacteria, each recognizing a different sequence. Human DNA from different individuals is treated with an appropriate enzyme which breaks it into many fragments. These are then separated by size on an agarose gel and the fragments containing sequences that are complementary to the DNA of the radioactively labeled probe can be identified (**Fig. 12**). The number and size of the fragments depend on the distribution of sites for that particular restriction enzyme in the region of the genome to which the probe hybridizes, and the sizes commonly detected are between 500 bases and 15,000 bases (15 kilobases, or kb). In this way variation in the base sequences can be detected not only in the region of the gene which codes for the protein but also in the large regions of noncoding DNA which form intervening and flanking sequences. Variants that are detected in this way (restriction fragment length polymorphism) behave as single codominant alleles, and it appears that for almost any DNA probe (whether a random fragment or a known gene) such variation can be found if a sufficient number of restriction enzymes are tried. *See* RESTRICTION ENZYME.

Another very useful type of genetic marker usually found in noncoding regions of DNA is the VNTR (variable number of tandem repeats). In this case, sequences recognized by one probe occur at the same place but in variable numbers on the chromosomes of different individuals. If the DNA is digested with any enzyme that does not cut within the repeat, the size of the fragment recognized by probing after agarose gel electrophoresis will vary between individuals and between the two homologous

chromosomes of the same individual chromosomes. At some of these hypervariable loci, 99% of individuals are heterozygous, so every family is informative for linkage analysis with that marker.

Family studies for mapping disease genes. The frequency of the different polymorphisms described above means that virtually the whole genome is within reach of good genetic markers. Family studies on diseases with clear mendelian inheritance, but in which the basic biochemical defect is unknown, have become a very practical undertaking. Such studies have allowed the localization of the genes that cause many autosomal-dominant disorders such as Huntington's disease and adult-onset polycystic kidney disease. It is still much more complex to localize a gene that causes a disease inherited in an autosomal-recessive manner if the heterozygotes cannot be identified, but in the case of cystic fibrosis, a common disease in populations of European origin, this has been accomplished (Fig. 10).

There are several ways in which localizing such genes may help in understanding a disease, in some cases by demonstrating that more than one gene is involved and in others by providing a starting point for reverse genetics. Once the approximate position of a disease gene has been found, it is possible to generate more genetic markers in the same region of the chromosome and to "walk" or "jump" along the chromosome to find the actual disease gene and the mutations that cause the disease, allowing more certain diagnosis and better genetic counseling. A study of the protein coded for by the gene in question should then offer a deeper understanding of the disease and how it may be treated. Understanding a disease by reverse genetics was first successful for Duchenne muscular dystrophy and chronic granulomatous disease, both of which are inherited on the X chromosome. The most dramatic success of this approach has been the identification of the gene that is defective in cystic fibrosis.

A few diseases are inherited only through the mother, but are equally severe in males and females. (This is in contrast to X-linked inheritance, in which genes can pass from father to daughter but not from father to son.) Some of these diseases, including Leber's optic atrophy, are due to mutations of the mitochondria. The mitochondrion is a small circular piece of DNA, only about 16 kb long, that is found in many copies in each cell. The whole DNA sequence of human mitochondrial DNA is known and all the genes have been identified, and so in some ways it can be regarded as a very small chromosome.

Ongoing research. Several techniques should greatly speed up progress in human gene mapping. One is the polymerase chain reaction, which allows the selective amplification of a small region of DNA of interest from a single drop of blood—or even a single cell—without complex purification. One or more particular genes can then be examined quickly and easily in very large numbers of people or in many individual sperms from a single informative individual. The latter technique is the genetic equivalent of sampling hundreds of children from the same father. Another approach, which can be used in conjunction with polymerase chain reaction, is the direct preparation of DNA from a small piece, such as a single band, of a chromosome dissected from a metaphase spread. Therefore the chromosomal origin of clones that are derived from such pieces is precisely known.

Another exciting field of study is that of oncogenes. These are sequences similar to those of ceres, and in some circumstances these DNA sequences can cause changes in mouse cells similar to those seen in cancer cells. Although their relationship to human cancer is still not clear, the coincidence of the map positions of some of these oncogenes and the breakpoints of characteristic chromosome rearrangements seen in some tumors is under investigation. *See* ONCOGENES.

It is important to realize the different scale of mapping that can be achieved by the different approaches. Although the banding patterns shown in Fig. 1 are those usually obtained during the metaphase stage of cell division, higher resolution giving at least 800 bands for the whole genome can be obtained by examining the chromosomes earlier, during prophase, when they are much longer. This has led to a more precise definition of breakpoints and of map positions, although the smallest visible band still represents approximately 1 million base pairs, and two genes that appear to be quite close together in family studies may be several million base pairs apart. Therefore, there is a considerable difference in scale between molecular mapping (for example, of the various defects in beta-hemoglobin) and the physical and genetic chromosome mapping described here. This gap is bridged by a technique known as pulsed-field gel electrophoresis in which so-called rare cutting restriction enzymes are used to generate very large fragments of DNA (up to 1 million base pairs long), which are separated in gels by applying alternating cycles of electric fields in different directions. Some of these fragments are found to carry more than one gene, and so it gives an exact measure of the distance between them. Eventually, overlapping fragments of DNA spanning all the human chromosomes should be identified and most of it should be sequenced. The human genetic map will then be essentially complete and should lead to better diagnosis, genetic counseling, and, eventually, treatment of inherited diseases. The human genetic map should also allow a greater understanding of the genetic changes in cells that underlie other diseases, especially cancers. *See* GENETIC MAPPING; GENETICS.

Sue Povey

Bibliography. K. E. Davies (ed.), *Genome Analysis: A Practical Approach*, 1988; D. S. Falconer, *Introduction to Quantitative Genetics*, 3d ed., 1986; F. C. Fraser and J. J. Nora, *Genetics of Man*, 2d ed., 1986; H. Harris, *The Principles of Human Biochemical Genetics*, 3d ed., 1981; D. L. Hartl and A. G. Clark, *Principles of Population Genetics*, 3d ed., 1997; *Human Gene Mapping 10*, 1989; A. J. Jeffreys, V. Wilson, and S. L. Thein, Hypervariable "minisatellite" regions in human DNA, *Nature*, 314:67–73,

1985; B. S. Kerem et al., Identification of the cystic fibrosis gene: Genetic analysis, *Science*, 245:1073–1080, 1989; M. Levitan, *Textbook of Human Genetics*, 3d ed., 1988; V. A. McKusick, *Mendelian Inheritance in Man*, 8th ed., 1988; *Mapping Our Genes, The Genome Projects: How Big, How Fast*, 1988; J. J. Nora, *Medical Genetics: Principles and Practices*, 4th ed., 1993; W. L. Nyhan and P. T. Ozand, *Atlas of Metabolic Diseases*, 1998; R. R. Race and R. Sanger, *Blood Groups in Man*, 6th ed., 1975; J. H. Relethford, *The Human Species: An Introduction to Biological Anthropology*, 4th ed., 2000; J. R. Riordan et al., Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA, *Science*, 245:1066–1073, 1989; C. R. Scriver et al., *The Metabolic and Molecular Basis of Inherited Disease*, 1995; T. B. Shows et al., Mapping the human genome, cloned genes, DNA polymorphisms, and inherited disease, *Adv. Hum. Genet.*, 12:341–452, 1982; H. E. Sutton, *An Introduction to Human Genetics*, 4th ed., 1988; F. Vogel, Clinical consequences of heterozygosity for autosomal-recessive diseases, *Clin. Genet.*, 25:381–415, 1984; F. Vogel and A. G. Motulsky, *Human Genetics*, 2d rev. ed., 1986.

Human genome

On April 14, 2003, the International Human Genome Sequencing Consortium, led in the United States by the National Human Genome Research Institute and the Department of Energy, announced the successful completion of the Human Genome Project (HGP) more than 2 years ahead of schedule. The DNA sequence produced by the HGP covered about 99% of the human genome's gene-containing regions and was sequenced to an accuracy of greater than 99.99%. In addition, to help researchers better understand the meaning of the human genetic instruction book, the project had taken on a wide range of other goals, from sequencing the genomes of model organisms to developing new technologies to study whole genomes. All of those goals were met or surpassed (see **table**).

Scientific strategy. The scientific strategies that were employed evolved over the years, but the basic concept for the project that was initially proposed by the National Research Council (NRC) of the U.S. National Academy of Sciences in the mid-1980s turned out to be effective. Because the human genome is so big (human DNA consists of about 3 billion nucleotides connected end to end in a linear array; **Fig. 1**), it was necessary to break the task down into manageable chunks. The first step was to create a genetic map of the whole genome. Such a map is generated by analyzing how frequently markers are inherited together in families. Markers that are in proximity are inherited together more frequently than those that are far apart. The resulting map shows a series of signposts along the DNA (**Fig. 2**).

The second step was to create a physical map of the DNA. The DNA was broken into small pieces that could be cloned and replicated in bacteria to generate enough material for study. The pieces were then fitted together by studying how they overlapped with each other. The genetic markers, as well as other types of landmarks, were used for this purpose. Eventually the whole genome was covered with overlapping clones.

The final step was to sequence the clones and piece together the entire genome sequence. Before this could be done, much research and development was needed to make sequencing more efficient and less costly. With a major investment in research, the technology did improve substantially and the costs came down dramatically. The improvements included streamlined methods, the use of robots

Comparison of the first two human chromosomes to be sequenced

Category	Explanation	Chromosome 21	Chromosome 22
Base pairs sequenced	Only the long arms of the chromosomes, representing the euchromatic regions, were sequenced. The heterochromatin cannot be cloned with current technology.	33.5×10^6	33.4×10^6
Number of gaps	The gaps contain sequences that cannot be cloned in any known cloning vector.	3	10
Number of known genes	Genes for which there is evidence that they are expressed in humans, or which are very similar to known genes in model organisms.	127	397
Number of predicted genes	Genes predicted by computer analysis, but unlike any other known genes.	98	148
Number of pseudogenes	Genes that cannot function because they lack a key feature. They are probably evolutionary remnants.	59	134
Percent of repeats	DNA sequences that are repeated over and over throughout the genome.	40	42
Percent of Alu repeats	Two common kinds of repeats.	9.5	16.8
Percent of LINE1 repeats	Two common kinds of repeats.	12.9	9.7
Percent of G + C bases	DNA consists of four bases, abbreviated A, T, G, and C. The frequency of the bases varies between organisms and also in different parts of the genome. Since G always pairs with C and A with T, the composition of particular DNA regions is expressed as % G + C.	41	48

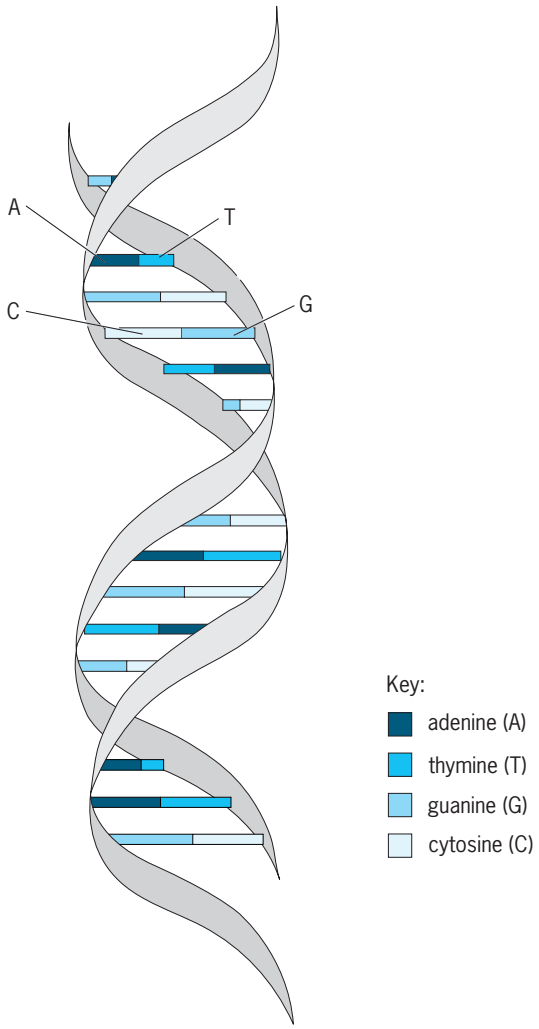

Fig. 1. Structure of DNA. DNA is a long thin molecule made up of chemical units called nucleotide bases. The four bases are adenine, thymine, guanine, and cytosine (A, T, G, and C). They are arranged in two parallel strands that are complementary to each other. An A always appears opposite a T, and a G always appears opposite a C. Sequencing the DNA involves determining the order of the bases in the molecule.

to handle large numbers of samples simultaneously, more accurate and faster sequencing machines, and better computer programs for tracking and assembling the data.

Model organisms. An important element of the overall strategy was to include the study of model organisms in the HGP. There were two reasons for this: (1) Simpler organisms provide good practice material. (2) Comparisons between model organisms and humans yield very valuable scientific information. All life forms have much in common, including some of their DNA sequences. Without the information that can be gained from the study of model organisms, it would be very difficult to know what the human DNA sequence means.

The HGP initially adopted five model organisms to have their DNA sequenced: the bacterium *Escherichia coli*, the yeast *Saccharomyces cerevisiae*, the roundworm *Caenorhabditis elegans*, the fruitfly *Drosophila melanogaster*, and the laboratory mouse *Mus musculus*. Subsequently, the HGP supported the sequencing of additional organisms, including the rat *Rattus norvegicus*, the roundworm *Caenorhabditis briggsae*, and the fruitfly *Drosophila pseudoobscura*.

In December 2002, a study comparing the human genome sequence with the mouse genome sequence found that 99% of human genes have a counterpart in mice, even though the two mammals diverged from a common ancestor more than 75 million years ago. In addition to sharing most gene functions with humans, the mouse is a valuable model for biomedical research because it can be selectively bred and other experiments can be conducted that are not possible on humans.

Project evaluation. Originally, the proponents of the HGP suggested that a 15-year time frame was appropriate for obtaining the human genome sequence. In order to make sure that this schedule was followed, 5-year plans were developed to set intermediate milestones. Remarkably, all the milestones were met, many times well before the planned deadlines. The sequence of yeast was released in April 1996, the sequence of *E. coli* was published in September 1997, *C. elegans* was completed in December 1998, the *Drosophila* sequence was released in early 2000, a first draft of the rat sequence was released in November 2002, and the mouse sequence was published in December 2002.

A first draft of the entire human sequence was released in June 2000 with publications describing it following in February 2001. A high-quality, comprehensive version of the human genome sequence was completed in April 2003, more than 2 years quicker than the 15-year period and coinciding with the 50th anniversary of the discovery of the double-helical structure of DNA. From the outset, sequence information from the HGP was immediately and freely distributed to scientists around the world, with no restrictions on its use or redistribution. In parallel with the publicly funded effort, a private United States company called Celera also generated a draft of the human sequence, which was completed about the same time as the HGP's first draft, though not made freely accessible. The Celera strategy used a different approach. Instead of mapping the cloned DNA pieces first and sequencing them later, random pieces were sequenced directly and subsequently correlated with maps. Some aspects of this approach were incorporated into the strategy for sequencing additional organisms, such as the mouse and the rat.

Findings. How many genes are there is probably the most common question regarding the human genome. Although a definitive count of human genes must await further experimental and computational analysis, the HGP's initial analysis of the human genome in 2001 led to an estimate that humans have only about 30,000 genes. That was quite a surprise because previous estimates were 80,000 to 100,000 genes. The lower gene count does not necessarily mean that the human genome is less complex, because many genes can produce more

than one protein by alternate splicing of their exons (protein-encoding regions of the gene) during translation into the constituents of proteins.

Another fascinating feature of the human genome sequence, as well as the genome sequences of other mammals, is the distribution of genes on the chromosomes. It turns out that mammalian chromosomes have areas with many genes in proximity to one another, but these regions are interspersed with vast expanses of DNA devoid of protein-coding genes. This uneven distribution of genes stands in marked contrast to the more uniform distribution of genes throughout the genomes of many other organisms, such as the roundworm and the fruitfly.

In their analysis, HGP scientists also found that about 50% of the DNA in the human genome is made of repetitive sequences, which is much greater than the percentage of repetitive DNA found in the roundworm (7%), the fruitfly (3%), and a variety of other nonmammalian species. These repeated sequence elements provide a rich record of clues to the evolutionary past of humans. It is possible to date various types of these repeats to the times when they appeared in the evolutionary process and to follow their fates in different regions of the genome and in different species. Because repeated sequences are more likely to misalign when DNA is being copied or repaired, areas with a high concentration of repeats are more prone to mutation than other regions of the genome. Consequently, the various types of repeats have helped to reshape the genome in a multitude of ways by rearranging and modifying existing genes, as well as by creating entirely new genes.

Such repeats and the relatively predictable patterns of DNA shrinkage and growth associated with them are also being used as "DNA dating" tools to explore evolution in a rapidly emerging area of research known as phylogenetics. The pace of evolutionary processes, sometimes referred to as a molecular clock, appears to vary among different types of repeats. However, researchers in some instances have been able to select a specific type of repeat, examine its prevalence and length among the genomes of a wide range of species, and then use that genetic information to construct a family, or phylogenetic, tree of various species showing when in evolution that repetitive element was "born" and how it "moved" among species during the course of evolution.

For example, two types of repeats, called DNA transposons and long terminal repeat (LTR) transposons, present in the genome of the mouse are much rarer in the human genome. Consequently, a phylogenetic tree built using data on DNA transposons and LTR transposons would branch in a manner that reflects how the prevalence of these repeat elements has changed since rodents and primates diverged from a common ancestor an estimated 75 million years ago.

Future research. The availability of large amounts of DNA sequence information as well as other genomic resources has profoundly affected biomedical research and thinking. Information about whole genomes that was previously gained one gene at a time has now been obtained. With the complete sets of genes of organisms available, how genes are turned on and off and how genes interact with each other can be studied. The possibilities are endlessly exciting, and the demand for more sequence is increasing all the time. Many more organisms will be sequenced. But having the sequence is not enough. What the different genes do and how they affect human health must also be learned.

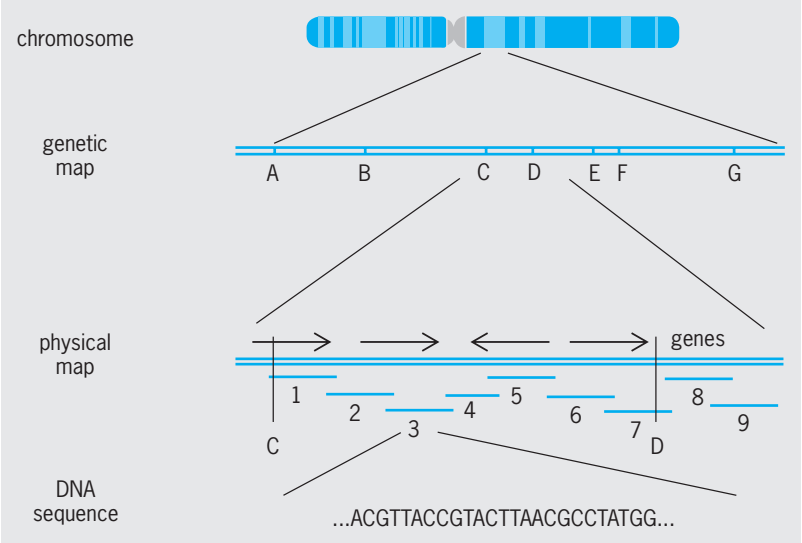

Fig. 2. Steps in analyzing a genome. (1) Markers are placed on the chromosomes by genetic mapping, that is, observing how the markers are inherited in families. (2) A physical map is created from overlapping cloned pieces of the DNA. (3) The sequence of each piece is determined, and the sequences are lined up by computer until a continuous sequence along the whole chromosome is obtained. Steps 2 and 3 can be reversed or done in parallel. As the pieces are sequenced, the sequences at the overlapping ends can be used to help order the pieces. If the sequencing is done before the pieces are mapped, the process is called whole-genome shotgun sequencing.

Among the most important large-scale opportunities for scientists in the genome era are the development of innovative technologies and the creation of informational frameworks to make the results of genomic research applicable to individual health. Although all humans are 99.9% identical in their genetic makeup, the 0.1% variation in DNA sequences is thought to hold key clues to individual differences in susceptibility to disease.

The International HapMap Project was launched in October 2002 with the goal of building a catalog of human genetic variations and determining how these variations are organized into neighborhoods, or haplotypes, along the human chromosomes. The HapMap will serve as a tool of researchers trying to discover the common genetic variations associated with complex diseases, as well as variations responsible for differences in drug response.

In March 2003, scientists also set out to develop efficient ways of identifying and precisely locating all of the functional elements contained in the human DNA sequence. The ultimate goal of the Encyclopedia of DNA Elements (ENCODE) project is to create

a reference work that will help scientists mine and fully utilize the human sequence, gain a deeper understanding of human biology, and develop new strategies for the prevention and treatment of disease. In the first phase of ENCODE, researchers will work cooperatively to develop high-throughput methods for rigorously analyzing a defined set of DNA target regions constituting approximately 1% of the human genome. It is hoped this pilot project will pave the way for scaling up this effort to characterize efficiently and effectively all of the protein-coding genes, nonprotein coding genes, and other sequence-based functional elements contained in human DNA. As has been the case with the public effort to sequence the human genome, data from the ENCODE project will be collected and stored in a database that will be freely available to the entire scientific community.

Other frontiers include the establishment of publicly available libraries of chemical compounds for use by basic scientists in their efforts to chart biological pathways, as well as genomic initiatives to understand the life processes of single-cell organisms, or microbes, with the ultimate aim of using the capabilities of these organisms to address needs relating to health, energy, and the environment.

Many challenges lie ahead. Getting the full sequence of human DNA is a dramatic achievement. But what was viewed as an end point has turned out to be just a beginning: a new era of genomic biology lies ahead. *See* DEOXYRIBONUCLEIC ACID (DNA); GENE; GENETIC CODE; GENETIC ENGINEERING; HUMAN GENETICS; MOLECULAR BIOLOGY; NUCLEIC ACID.

F. Collins; E. Jordan

Bibliography. M. D. Adams et al., The genome sequence of *Drosophila melanogaster, Science*, 287: 2185–2195, 2000 (also related articles, 287:2181–2184, 2196–2224); C. elegans Sequencing Consortium, Genome sequence of the nematode *C. elegans*: A platform for investigating biology, *Science*, 282:2012–2018, 1998 (also related articles, 282:2018–2046); F. S. Collins, Medical and societal consequences of the Human Genome Project, *N. Engl. J. Med.*, 341:28–37, 1999; F. S. Collins et al., A vision for the future of genomics research: A blueprint for the genomic era, *Nature*, 422:835–847, 2003; F. S. Collins, M. Morgan, and A. Patrinos, The Human Genome Project: Lessons from large-scale biology, *Science*, 300:286–290, 2003; A. Goffeau, R. Aert, and M. L. Agostini-Carbone, The Yeast Genome Directory, *Nature*, supplement to vol. 387, May 29, 1997; E. D. Green, The Human Genome Project, in *Metabolic and Molecular Bases of Inherited Disease*, 8th ed., 2001; International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome, *Nature*, 409:860–921, 2001; Mouse Genome Sequencing Consortium, Initial sequencing and comparative analysis of the mouse genome, *Nature*, 420:520–562, 2002; National Research Council, *Mapping and Sequencing the Human Genome*, 1988; J. C. Venter et al., The sequence of the human genome, *Science*, 291:1304–1351, 2001.

Human-machine systems

Complex systems that comprise both humans and machines. Human-machine systems engineering is the analysis, modeling, and design of such systems. It is distinguished from the more general field of human factors and from the related fields of human-computer interaction, engineering psychology, and sociotechnical systems theory in three general ways. First, human-machine systems engineering focuses on large, complex, dynamic control systems that often are partially automated (such as flying an airplane, monitoring a nuclear power plant, or supervising a flexible manufacturing system). Second, human-machine systems engineers build quantitative or computational models of the human-machine interaction as tools for analysis and frameworks for design. Finally, human-machine systems engineers study human problem-solving in naturalistic settings or in high-fidelity simulation environments. *See* HUMAN-COMPUTER INTERACTION; HUMAN-FACTORS ENGINEERING.

Thus, human-machine systems engineering focuses on the unique challenges associated with designing joint technological and human systems. Historically it has grown out of work on cybernetics, control engineering, information and communication theory, and engineering psychology. Subsequently, researchers who focus on cognitive human-machine systems (in which human work is primarily cognitive rather than manual) have also referred to their specialization as cognitive engineering or cognitive systems engineering. *See* CYBERNETICS; INFORMATION THEORY.

The four major aspects of human-machine systems, in roughly historical order, are systems in which the human acts as a manual controller, systems in which the human acts as a supervisory controller, human interaction with artificial-intelligence systems, and human teams in complex systems. This general progression is related to advances in computer and automation technology. With the increasing sophistication and complexity of such technology, the human role has shifted from direct manual control to supervisory control of physical processes, to supervision of intelligent systems, and finally, with an increasing emphasis on the social and organizational aspects of complex systems, to teamwork in complex environments.

Aviation is an example of a human-machine system in which all of these developments have occurred. Early work in aircraft systems focused on manual control models of pilot performance. With increasing levels of automation, the pilot shifted to a more supervisory role in which tasks such as planning and programming the flight management computer became the predominant form of work. *See* AIRCRAFT INSTRUMENTATION; AUTOPILOT; FLIGHT CONTROLS.

Manual control. Manual control means that the individual exerts direct control on a dynamic system. Vehicle control tasks such as driving a car and flying an airplane, process control in chemical plants, and ship handling are examples of human work that have

been modeled as manual control tasks, and in particular, often as closed-loop negative-feedback systems in which the system is continually trying to reduce error (defined as the difference between its goal and its current state) by getting feedback about its current state and then making control actions to compensate for this error. Such human-in-the-loop control involves a display, a human, a controlled system, and a goal (**Fig. 1**). Driving a car may be used as an example of manual control. In this case, the display is the view through the windshield, the human is the driver, the controlled system is the car, and the goal is to stay on the road. The process of driving successfully on the road can then be modeled as a closed-loop negative-feedback system in which the driver is continually getting information via the windshield (display) about the error, which is the difference between the way that the road curves (target state of the controlled system, the input) and the car's current position, velocity, and trajectory (actual state of the controlled system). The driver exerts a change in control, for example by turning the steering wheel. In response to this control input, as well as to any external disturbances such as wind gusts, the car's direction or output changes. This new system output is now fed back to the driver through the windshield, and the cycle continues. The ratio of output to input is called the transfer function, which describes the system dynamics mathematically in terms of a number of concepts such as system order, time lag, and gain. *See* CONTROL SYSTEMS; PROCESS CONTROL.

Classical control theory focuses on system stability in the frequency domain. The original crossover model is based on classical control theory, considers the human-machine system as the invariant, and describes how human manual controllers adjust gain to achieve stability. It provides an elegant mathematical characterization of human manual control: The human exerts control such that the combined human-machine system transfer function behaves like a first-order system with a time delay. Modern or optimal control theory focuses on prediction in the time domain and considers the human controller as a model-based controller that includes a Kalman filter, a model-based state estimator to keep running estimates of state variables, and an optimal controller to make control actions on the basis of predictions derived from these estimates (**Fig. 2**). Thus, the Kalman filter might be viewed as the mental model of the system. More elaborate versions of classical and optimal control models include components such as neuromuscular delays and additional predictors. *See* CONTROL SYSTEM STABILITY; ESTIMATION THEORY; OPTIMAL CONTROL THEORY.

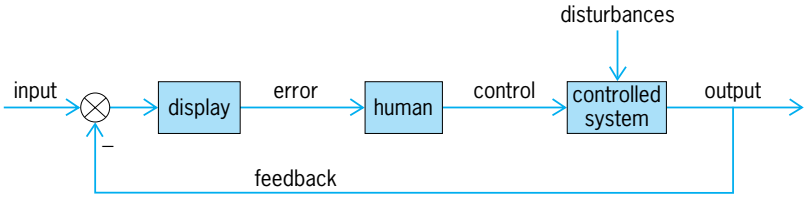

Fig. 1. Simple closed-loop negative-feedback system as a model of manual control.

Work based on control theory looks at the effects of different system dynamics on human performance in manual control (for example, varying gain, time lag, and system order) and different kinds of displays to provide better feedback for manual control. In particular, the use of pursuit, compensatory, and predictor displays is considered in manual tracking tasks (for example, in which the person seeks to minimize the error between a moving target object and the cursor on a computer display). Pursuit displays explicitly show the target and the cursor; compensatory displays show only the difference between the target and cursor. Predictive displays show both the current state of the system and a model-based estimate of its predicted future state. Quickening is a display technique in which higher-order system dynamics are taken into account when generating a predictive display. Teleoperation or remote manipulation, the real-time control of remotely located machines that act as eyes and hands of a person located elsewhere, has been used in undersea and lunar exploration and in microsurgery. *See* MICROMANIPULATION; REMOTE MANIPULATORS; UNDERWATER VEHICLES.

Supervisory control. With the advance of computer and automation technology, in many systems the human no longer acts as a direct manual controller but as a supervisor. This is seen in systems such as nuclear power plants and process-control plants, and is even true in some advanced aircraft systems in which the pilot programs the flight management computer rather than directly flying the airplane. In such systems, human supervisory controllers monitor and only intermittently control a highly automated process. Typical supervisory control activities include equipment configuration (for example, startup and shutdown of the plant), monitoring and fine-tuning system parameters, planning and scheduling control activities, predicting system state, and

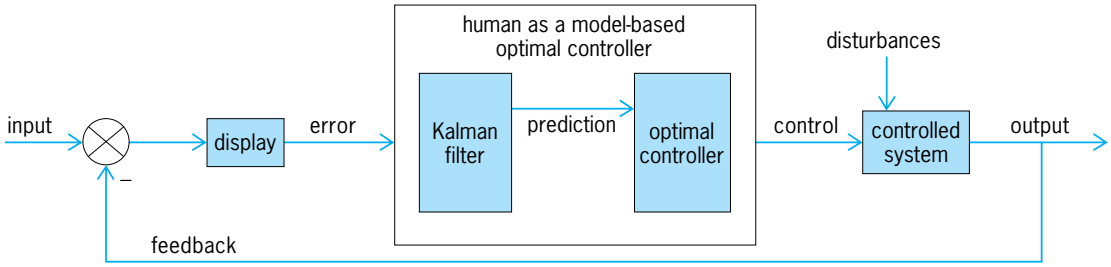

Fig. 2. Modern or optimal control paradigm.

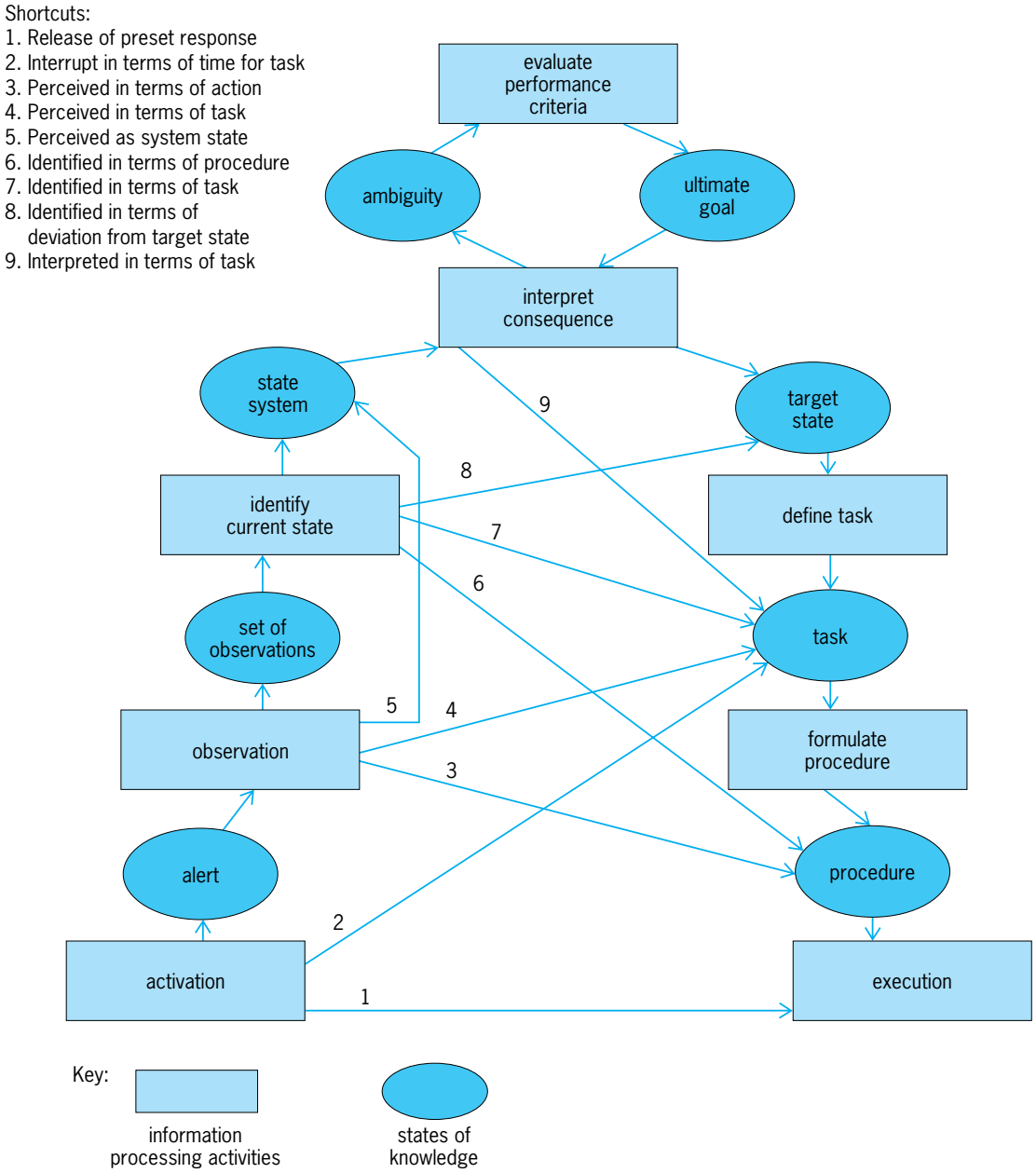

Fig. 3. Decision ladder. Human performance can be described as skill-based (using shortcut 1), rule-based (using shortcuts 2–9), or knowledge-based (traversal of the entire ladder). (*After T. B. Sheridan and G. Johannsen, eds., Monitoring Behavior and Supervisory Control, Plenum Press, 1976*)

compensating for system failures and abnormal conditions. Whereas earlier models of supervisory control tended to focus on monitoring and reactive, compensatory activities, more advanced research has emphasized the active, anticipatory nature of supervisory control. That is, supervisory controllers do not just monitor the system and react to perceived problems; rather, they use so-called mental simulation to predict the future state of the system, actively seek information, set goals, make plans, and perform control actions that anticipate future system state changes.

Human supervisory control differs in many important ways from human manual control. First, human supervisory control is primarily cognitive, that is, concerned with problem solving, planning, decision making, and other kinds of thinking rather than with manual skill and human physical constraints. Second, the systems that are supervised are frequently very complex and highly automated, requiring the human to manage multiple concurrent activities or, typically, requiring a team of human operators (as discussed below). Finally, the level of automation has important implications for the design of human-machine systems with respect to function allocation, human error and human reliability analysis, trust between humans and machines, and the overall integration of human problem-solving skills with potentially obscure workings on the part of the machine. In general, human supervisory controllers are at least one

step removed from the actual process being controlled, and how they can maintain situation awareness and be ready to act quickly and appropriately is a very important and active area of research.

Models. Supervisory control has been modeled by using a wide variety of mathematical and analytic tools, including queueing theory, fuzzy set theory, network models of activity or state changes such as Petri nets, the discrete control model, the operator function model, and a variety of artificial-intelligence knowledge-representation schemes, including rule-based models and blackboard models that reflect activity and decision making in operational tasks. One influential model of human decision making in complex environments is the decision ladder that distinguishes between skill-, rule-, and knowledge-based behavior (**Fig. 3**). Alternative models emphasize the role of situational context and active information seeking, planning, and other anticipatory cognitive actions characteristic of supervisory control. *See* ARTIFICIAL INTELLIGENCE; DECISION ANALYSIS; FUZZY SETS AND SYSTEMS; QUEUEING THEORY.

System design. System design to support human supervisory control has followed three major lines: display design (or representation aiding), telerobotics, and intelligent support systems. Representation aiding considers how to represent and organize system information to support effective human problem solving. One approach is to use the abstraction hierarchy (**Fig. 4**) as a framework for supporting situation assessment. The abstraction hierarchy represents the functional properties of a system according to means-ends relationships; this describes, top-down, why the system works as it does and, bottom-up, how it works as it does. The basic idea of using the abstraction hierarchy as a basis for display design (an approach also called ecological interface design) is to directly show higher-order properties of the system (higher levels as described in the abstraction hierarchy) and thus in a sense change a complex reasoning task into a (presumably) more tractable perceptual one.

Telerobotics refers to the supervisory control of a sophisticated teleoperator that is sometimes called a telerobot. The human supervisor intermittently communicates goals, plans, and other information to a computer to guide the telerobot in carrying out its task. Through the computer intermediary, the human keeps track of the telerobot's progress.

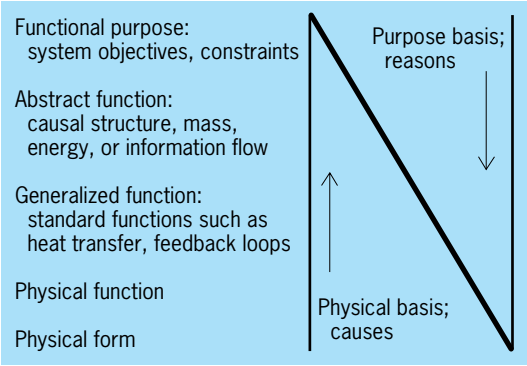

Fig. 4. Abstraction hierarchy. (*After J. Rasmussen, Skills, rules, knowledge: Signals, signs, and symbols and other distinctions in human performance models, IEEE Trans. Sys. Man Cybernet., SMC-13, (3):257–267, 1983*)

Intelligent support systems involve the use of artificial-intelligence programs that act as tutors, assistants, or both to human supervisory controllers. Artificial-intelligence applications in supervisory control systems include expert systems for automatic fault detection and diagnosis; intelligent associate systems that dynamically adapt their behavior to the human in recognition of the intent of operator actions; and intelligent tutoring systems that teach procedural knowledge and diagnostic skills. The concept of intelligent agents offers the prospect of personalized assistants that gradually provide automation in cooperation with the human user: The agent monitors the user's actions, recognizes recurring patterns, and eventually offers to invoke that sequence of actions automatically for the user; alternatively, users can program agents directly by themselves, sometimes via graphical programming languages. However, relatively few such systems have been deployed in actual supervisory control environments. *See* EXPERT SYSTEMS; FAULT ANALYSIS; INTELLIGENT MACHINE.

Human–intelligent-system interaction. In the design of intelligent systems for human use, many issues about trust, authority, and responsibility in decision making are important to consider. For example, in interacting with a typical expert system, the human user inputs data, waits for the expert system to generate a judgment or diagnosis by applying its reasoning algorithms, and then decides if that result makes any sense. This kind of interaction does not support human authority in the decision-making process; this is a particularly important issue in complex dynamic systems in which the costs of errors are potentially catastrophic and it is usually not possible to undo control actions. There are a number of different levels of intelligent automation in decision-making tasks (**Fig. 5**).

In general, emerging theories of how to design good intelligent systems for human use in complex dynamic environments advocate cooperative support for human authority rather than full automation of decision making and reasoning. For example, intelligent systems can support cooperative problem formulation, exploration of alternatives, and dynamic task allocation between human and machine (that is, adaptive automation, an issue that has received much attention).

Besides these general issues of knowledge representation, inference, and trust, control, and responsibility between humans and machines, intelligent systems for supervisory control have the difficulty of supporting problem solving under tight time constraints. A variety of quantitative and computational techniques have been used to provide reasoning under uncertainty in dynamic environments.

Teams in complex systems. Group work plays an important role in human-computer interaction, artificial intelligence, and human supervisory control.

Description of Interaction	Human Functions	Computer Functions
1. Human does the whole job up to the point of turning it over to the computer to implement.	(GETS options from outside) SELECTS action STARTS action	
2. Computer helps by determining the options.	(REQUESTS options) SELECTS action STARTS action	GETS options
3. Computer helps determine options and suggests one, which human need not follow.	(REQUESTS options) (REQUESTS SELECT action) SELECTS action (can be different) STARTS action	GETS options SELECTS action
4. Computer selects action and human may or may not do it.	(REQUESTS options) (REQUESTS SELECT action) APPROVES SELECT action STARTS action if HUMAN APPROVES	GETS options SELECTS action
5. Computer selects action and implements it if human approves.	(REQUESTS options) (REQUESTS SELECT action) APPROVES START action	GETS options SELECTS action STARTS action if HUMAN APPROVES
6. Computer selects action, informs human in plenty of time to stop it.	(REQUESTS options) (REQUESTS SELECT action) APPROVES START action	GETS options SELECTS action STARTS action if HUMAN APPROVES or if $t > T$ and HUMAN HAS NOT DISAPPROVED
7. Computer does whole job and necessarily tells human what it did.	(REQUESTS SELECT action)	GETS options SELECTS action STARTS action TELLS action
8. Computer does whole job and tells human what it did only if human explicitly asks.	(REQUESTS SELECT action) (REQUESTS TELL action)	GETS options SELECTS action STARTS action TELLS action if HUMAN REQUESTS
9. Computer does whole job and tells human what it did if it decides human should be told.	(REQUESTS SELECT action)	GETS options SELECTS action STARTS action TELLS action if COMPUTER APPROVES
10. Computer does whole job if it decides it should be done, and if so tells human if it decides human should be told.	(REQUESTS SELECT action)	GETS options SELECTS action STARTS action if COMPUTER APPROVES TELLS action if COMPUTER APPROVES

Fig. 5. Levels of automation in human-machine interaction. Other variations are possible. (*After T. B. Sheridan and W. L. Verplank, Human and Computer Control of Undersea Teleoperators, Tech. Rep., MIT Man-Machine Systems Laboratory, Cambridge, Massachusetts, 1978*)

The term distributed supervisory control refers to the situation in which a team of human operators works together in the supervisory control of a complex dynamic system. This situation is an example of computer-supported cooperative work, an interdisciplinary area that draws upon computer science, information systems, sociology, psychology, and anthropology to examine the mutual influences of technologies (such as electronic mail and video conferencing) on group work. Research areas include the organizational structures of the workplace, the methods by which the people coordinate and negotiate activity and responsibility, and the use of different technologies in the work environment. Supporting good cooperative work by system design relies on both an understanding of the formal requirements of work (which could be captured in engineering models) and the informal work that is done

to coordinate activity (which cannot be modeled but must be possible). Included are analysis of social, cultural, organizational, and managerial practices as well as the more traditional engineering analyses associated with tasks and physical processes. *See* INFORMATION SYSTEMS ENGINEERING; SYSTEMS ENGINEERING.

Patricia M. Jones

Bibliography. R. Baecker (ed.), *Readings in Groupware and Computer Supported Cooperative Work*, 1993; E. Hollnagel, G. Mancini, and D. Woods (eds.), *Cognitive Engineering in Complex Dynamic Worlds*, 1986; J. Rasmussen, A. Pejtersen, and L. Goodstein, *Cognitive Systems Engineering*, 1994; W. B. Rouse, *Systems Engineering Models of Human-Machine Interaction*, 1980; A. P. Sage (ed.), *System Design for Human Interaction*, 1987; T. B. Sheridan, *Telerobotics, Automation, and Human Supervisory Control*, 1992.

Humidification

The process of increasing the water-vapor content (humidity) of a gas. This process and its reverse operation, dehumidification, are important steps in air conditioning for human comfort and in many industrial operations. For a discussion of the dehumidification of gases, with the exception of air, *see* DRYING

Humidity is normally expressed as the weight of water per unit weight of dry air. This value is also referred to as the absolute humidity or humidity ratio and has the units kg/kg, lb/lb, or grains/lb. The water content of air also is frequently indicated as a relative humidity, which is the ratio (expressed as a percentage) of the partial pressure of water vapor in the air to the vapor pressure of pure water at the same temperature. *See* HUMIDITY.

Air (or other gas) can be humidified by direct injection of water vapor (steam) or, more commonly, by the evaporation of liquid water in contact with the airstream. When evaporation occurs, heat is required to provide the latent heat of vaporization. If no external source of heat is provided, either the water or the air, or both, will be cooled. The cooling of water by this process is the basis of operation for industrial cooling towers, whereas evaporative air coolers often used in hot, dry climates depend upon the air-cooling effect. In both these types of apparatus, humidification of the air occurs, although it is not the prime objective of the operation. In units designed primarily for humidification, the incoming air is usually heated to provide the latent heat of evaporation and to permit the air to leave the unit at controlled levels of both temperature and humidity.

Psychrometric chart. The operation of humidification equipment can best be understood by reference to a psychrometric chart as shown in **Fig. 1**. Every point on this chart represents a specific air condition with regard to temperature (bottom scale) and humidity (right-hand scale). Several other characteristics of the air are indicated by the location of the points on the chart, including the relative humidity and the wet-bulb temperature. The latter is defined as the equilibrium temperature attained by a small surface of liquid evaporating into a large amount of

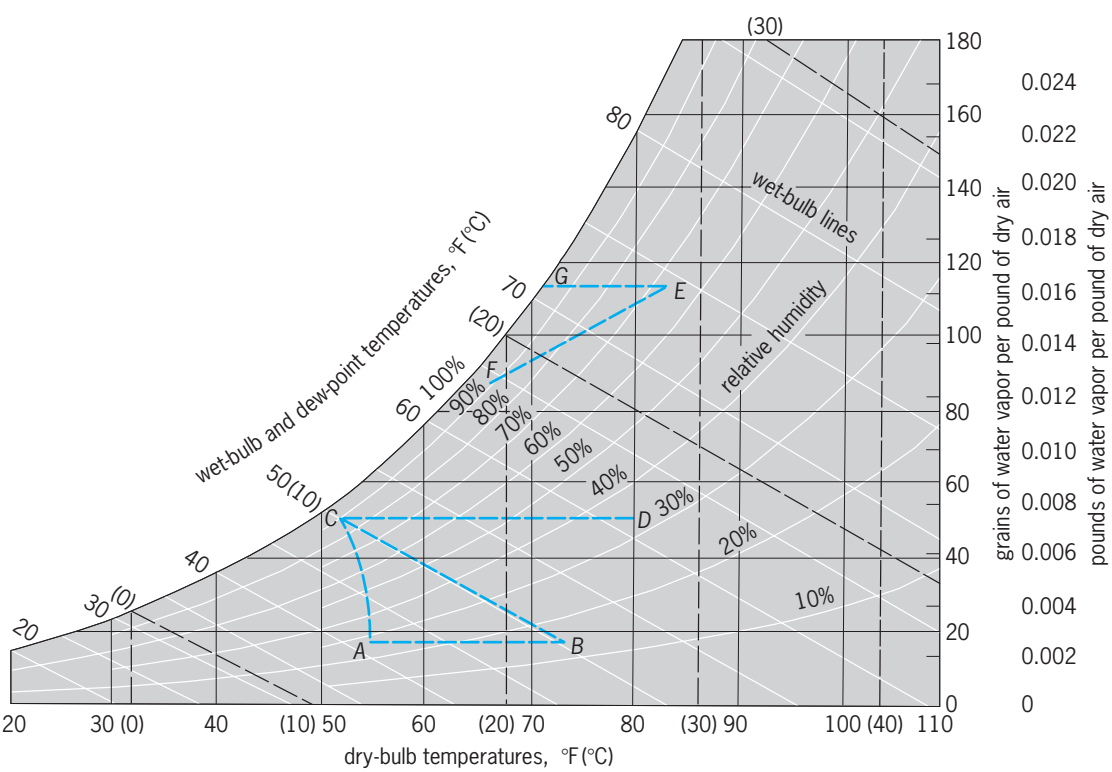

Fig. 1. An example of a psychrometric chart.

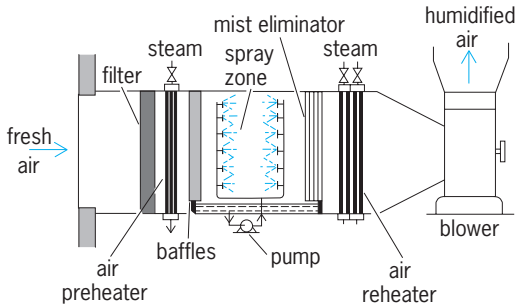

Fig. 2. Schematic of air conditioning humidifier.

unsaturated air. It is normally measured by moving air rapidly past a thermometer bulb covered with a wetted wick—hence the name. For the air-water system, the wet-bulb lines can also be used with sufficient accuracy to indicate the adiabatic saturation temperature, that is, the temperature which a gas will attain when saturated by adiabatic contact with water which is already at the adiabatic saturation temperature. *See* PSYCHROMETRICS.

Operation of an air conditioning humidifier such as shown in **Fig. 2** can be represented by the path *A-B-C-D* on the psychrometric chart. Air entering at the conditions indicated by point *A* is first heated to point *B* without change in humidity by contact with a heated coil. It then passes through a water-spray zone and is adiabatically cooled and humidified to the conditions represented by point *C*. The humidified air is then heated to the desired final outlet conditions of point *D*. Alternatively, the entering air can be contacted with heated water so that humidification occurs with little or no cooling. This path is represented by line *A-C*. *See* AIR CONDITIONING.

Cooling towers. Line *A-C* also represents a typical cooling-tower operation in which warm water is contacted with air for the purpose of cooling the water. Cooling towers are of considerable industrial importance in power plants, refineries, chemical plants, and large air conditioning or refrigeration installations in which considerable quantities of water are used in condensers and coolers to remove process heat. Two general types of water cooling towers are employed—natural circulation and mechanical draft. The former depend primarily upon wind and natural draft effects to provide air circulation, whereas the latter employ fans to move air through the tower. Cooling towers may be further subdivided into crossflow designs, in which the air moves horizontally while the water falls vertically, and counterflow designs (**Fig. 3**), in which the air moves upward countercurrent to the falling water. The crossflow design is useful for towers which must be kept to a minimum height; however, counterflow operation is theoretically more efficient, and this type of tower is capable of producing colder water. *See* COOLING TOWER.

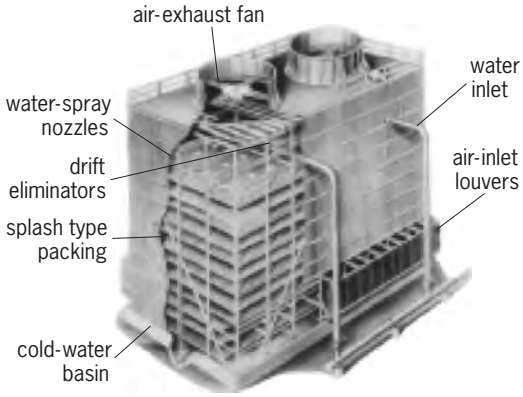

Fig. 3. Cutaway view of large industrial counterflow cooling tower. (*Fluor Products Co.*)

Dehumidification. The dehumidification of air is indicated by line *E-F* on the psychrometric chart. Dehumidification may be accomplished by contacting the air with cold water in a device similar to that of Fig. 2, except that the heater coils are not necessary, or by passing the air across banks of finned tubes through which cold water or refrigerant is passed. Because water condenses on the outside walls of the tubes, the two processes are equivalent in that the warm, moist air is in direct contact with cold water. Condensation results in the release of the latent heat of condensation, and this raises the temperature of the cooling liquid. It is not necessary for the entire gas stream to be cooled to the dew point (point *G*) for condensation to occur; however, no water can be condensed from the gas unless the cooling surface is below the dew point. *See* DEHUMIDIFIER.

Dehumidification can also be accomplished by the use of solid desiccants such as silica gel or alumina, or liquid absorbents such as triethylene glycol or lithium chloride solutions. Both types are commonly employed in regenerative systems in which the spent dehydration agent is stripped of water by applying heat. *See* DESICCANT; HEAT EXCHANGER; REFRIGERATION; STRIPPING (CHEMICAL ENGINEERING); UNIT OPERATIONS.

Arthur L. Kohl

Bibliography. American Society of Heating, Refrigerating, and Air Conditioning Engineers, *ASHRAE Handbook: Systems*, 1992, *Fundamentals*, 1994, *Applications*, 1993; G. W. Brundrett, *Criteria for Moisture Control*, 1990; J. M. Coulson and J. F. Richardson (eds.), *Chemical Engineering*, vol. 1: *Fluid Flow, Heat Transfer, and Mass Transfer*, 4th ed., 1990.

Humidistat

A controller that measures and controls relative humidity. A humidistat may be used to control either humidifying or dehumidifying equipment by the regulation of electric or pneumatic switches, valves, or dampers. Most methods for measuring humidity rely upon the swelling and shrinking of materials, such as human hair, silk, horn, gold-beater's skin, and wood, with increases and decreases in relative humidity.

Human hair is most commonly used because of its small diameter, which contributes to rapid absorption and dissipation of moisture. Strands of hair are

Fig. 1. Human-hair-element humidistat. (*Honeywell Inc.*)

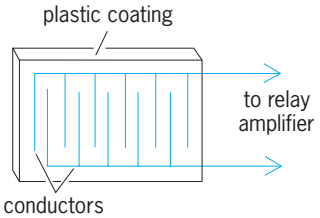

Fig. 2. Electronic humidistat. (*Honeywell Inc.*)

bunched and several such bunches are combined in a ribbonlike element (**Fig. 1**).

As the relative humidity of the air decreases, the strands of hair shorten; this movement is transmitted through a suitable lever mechanism to an electric switch or pneumatic valve, which is part of the humidistat.

An electronic humidistat includes a sensing element and a relay amplifier. The sensing element consists of alternate metal conductors on a small, flat plate with a plastic coating (**Fig. 2**).

An increase or decrease of the relative humidity causes a decrease or increase in the electrical resistance between the two sets of conductors; the change in resistance is measured by the relay amplifier. Small changes in relative humidity can be measured in this way for precise control. *See* HUMIDITY; HUMIDITY CONTROL; PSYCHROMETRICS.

John E. Haines; Richard L. Koral

Bibliography. I. Stepnich, *Humidity Control,* 1988.

Humidity

Atmospheric water-vapor content, expressed in any of several measures, especially relative humidity, absolute humidity, humidity mixing ratio, and specific humidity. Quantity of water vapor is also specified indirectly by dew point (or frost point), vapor pressure, and a combination of wet-bulb and dry-bulb (actual) temperatures. *See* DEW POINT; VAPOR PRESSURE.

Relative humidity is the ratio, in percent, of the moisture actually in the air to the moisture it would hold if it were saturated at the same temperature and pressure. It is a useful index of dryness or dampness for determining evaporation, or absorption of moisture. *See* PSYCHROMETRICS.

Human comfort is dependent on relative humidity on warm days, which are oppressive if relative humidity is high but may be tolerable if it is low. At other than high temperatures, comfort is not much affected by high relative humidity. *See* COMFORT TEMPERATURES.

However, very low relative humidity, which is common indoors during cold weather, can cause drying of skin or throat and adds to the discomfort of respiratory infections. The term indoor relative humidity is sometimes used to specify the relative humidity which outside air will have when heated to a given room temperature, such as 72°F (22°C), without addition of moisture. The indoor relative humidity always has a low value in cold weather and is then a better measure of the drying effect on skin than is outdoor relative humidity. This is even true outdoors because, when air is cold, skin temperature is much higher and may approximate normal room temperature. *See* BIOMETEOROLOGY.

Absolute humidity is the weight of water vapor in a unit volume of air expressed, for example, as grams per cubic meter or grains per cubic foot.

Humidity mixing ratio is the weight of water vapor mixed with unit mass of dry air, usually expressed as grams per kilogram. Specific humidity is the weight per unit mass of moist air and has nearly the same values as mixing ratio.

Dew point is the temperature at which air becomes saturated if cooled without addition of moisture or change of pressure; frost point is similar but with respect to saturation over ice. Vapor pressure is the partial pressure of water vapor in the air. Wet-bulb temperature is the lowest temperature obtainable by whirling or ventilating a thermometer whose bulb is covered with wet cloth. From readings of a psychrometer, an instrument composed of wet- and dry-bulb thermometers and a fan or other means of ventilation, values of all other measures of humidity may be determined from tables. *See* HYGROMETER; MOISTURE-CONTENT MEASUREMENT; PSYCHROMETER.

J. R. Fulks

Bibliography. D. Ahrens, *Meteorology Today: An Introduction to Weather, Climate, and the Environment*, 5th ed., 1994; R. A. Anthes, *Meteorology*, 7th ed., 1996; W. L. Donn, *Meteorology*, 4th ed., 1975; A. Miller and J. C. Thompson, *Elements of Meteorology*, 6th ed., 1999.

Humidity control

Regulation of the degree of saturation (relative humidity) or quantity (absolute humidity) of water vapor in a mixture of air and water vapor. Humidity is commonly mistaken as a quality of air. *See* HUMIDITY.

When the mixture of air and water vapor is heated at constant pressure, not in the presence of water or ice, the ratio of vapor pressure to saturation pressure decreases; that is, the relative humidity falls, but absolute humidity remains the same. If the warm mixture is brought in contact with water in an insulated system, adiabatic humidification takes place; the warm gases and the bulk of the water are cooled as heat is transferred to that portion of the water which evaporates, until the water vapor reaches its saturation pressure corresponding to the resultant water-air-vapor mixture temperature. Relative humidity is then 100% and absolute humidity has increased. Heating of the mixture and use of the heated mixture to evaporate water is typical of many industrial drying

processes, as well as such common domestic applications as hair drying. This same sequence occurs when warm furnace air is passed over wetted, porous surfaces to humidify air for comfort conditioning. *See* AIR CONDITIONING.

To remove moisture from the air-vapor mixture, the mixture is commonly cooled to the required dew point temperature (corresponding to the absolute humidity to be achieved) by passage over refrigerated coils or through an air washer where the mixture is brought in contact with chilled water. The result is a nearly saturated mixture which can be reheated, if required, to achieve the desired relative humidity. *See* DEHUMIDIFIER.

Moisture is also removed without refrigeration by absorption, a process in which the mixture passes through a spray of liquid sorbent that undergoes physical or chemical change as it becomes more dilute. Typical sorbents include lithium and calcium chloride solutions and ethylene glycol. *See* ABSORPTION.

Another means of dehumidification, by adsorption, uses silica gel or activated bauxite which, through capillary action, reduces the vapor pressure on its surface so that the water vapor in its vicinity, being supersaturated, condenses. *See* ADSORPTION; PSYCHROMETRICS. Richard L. Koral

Bibliography. G. W. Brundrett, *Criteria for Moisture Control*, 1990; R. Havrella, *Heating, Ventilating and Air Conditioning Fundamentals*, 1995; Instrument Society of America, *Moisture and Humidity,* 1985; I. Stepnich, *Humidity Control,* 1988.

Humite

A homologous series of magnesium nesosilicate minerals having the general composition $Mg_{2n+1}(SiO_4)_n(F,OH)_2$. The known species include norbergite ($n = 1$), chondrodite ($n = 2$), humite ($n = 3$), and clinohumite ($n = 4$). They are structurally related to forsterite olivine, Mg_2SiO_4, and brucite, $Mg(OH)_2$. All are based on hexagonal close-packed oxygen and fluorine atoms, the Mg atoms occupying octahedral interstices and the Si atoms occupying tetrahedral interstices. The hexagonal close-packed repeat distance is approximately 0.47 nm in these minerals. Forsterite, norbergite, and humite are orthorhombic; chondrodite and clinohumite are monoclinic; brucite is trigonal. Manganese analogs of these minerals occur as pink grains in metamorphosed manganese ores derived from preexisting siliceous carbonates and sedimentary manganese oxides. Other cations which can occur as substituents are Fe^{2+}, Ca^{2+}, Al^{3+}, and Ti^{4+}. Titanoclinohumite, for example, a high-pressure phase, has been found in deep-seated rocks. It may be one of the storage minerals for water at depth.

The minerals of the humite group have similar physical properties. The luster is resinous, and the color usually light yellow, brown, orange, or red. The pure synthetic Mg end members are colorless. Hardness is 6–6$^1/_2$ on Mohs scale, specific gravity is 3.1–3.2. They are very difficult to distinguish visually, and x-ray diffraction, electron microprobe, or optical techniques are required. They are found in regionally crystallized marbles, usually the skarn minerals associated with iron ores. Several species may occur in zoned contact with each other. Typical sources are the Grenville-age marbles in New York and Ontario, marble ejecta from Mount Vesuvius, and marbles from central Sweden. *See* SILICATE MINERALS. Paul B. Moore

Bibliography. W. A. Deer, R. A. Howie, and J. Zussman, *Rock-Forming Minerals: Orthosilicates*, 2d ed., 1997.

Humus

A group of substances that are natural products of earth surface environments. Probably the most widely distributed organic carbon–containing materials in terrestrial and aquatic environments, they are dark-colored, predominantly aromatic, acidic, hydrophilic, molecularly flexible polyelectrolytes. Humic substances constitute 70–80% of the organic matter in inorganic soils and are formed from the chemical and biological degradation of plant and animal residues and from synthetic activities of microorganisms.

Chemistry

Based on their solubility in alkali and acid, humic substances are partitioned into three main fractions: humic acid, which is soluble in dilute alkali but is coagulated by acidification of the alkaline extract; fulvic acid, which is the humic fraction that remains in solution when the alkaline extract is acidified, that is, it is soluble in both dilute alkali and dilute acid; and humin, which is the humic fraction that cannot be extracted from the soil by dilute base or acid.

Analytical characteristics. The elemental composition and functional group content of a typical humic acid and fulvic acid show that (1) the humic acid contains approximately 10% more carbon (C) but 10% less oxygen (O) than the fulvic acid; (2) there is relatively little difference between the two fractions in hydrogen (H), nitrogen (N), and sulfur (S) contents; (3) the total acidity and COOH group content of the fulvic acid are appreciably higher than those of the humic acid; (4) both materials contain per unit weight approximately the same concentrations of phenolic OH and total C═O and OCH_3 groups, but the fulvic acid is richer in alcoholic OH groups than is the humic acid; (5) about 78% of the oxygen in the humic acid is present in functional groups, but all of the oxygen in the fulvic acid is similarly distributed; (6) the E_4/E_6 ratio (the ratio of absorbance at 460 over that at 660 nanometers) of the fulvic acid is twice as high as that of the humic acid, which means that the fulvic acid has a lower particle or molecular weight than the humic acid; and (7) the free-radical content of the humic acid is higher than that of the fulvic acid but the other electron spin resonance

TABLE 1. Analytical characteristics of a typical humic acid and fulvic acid*

Characteristic	Humic acid	Fulvic acid
Element, %		
Carbon	56.2	45.7
Hydrogen	4.7	5.4
Nitrogen	3.2	2.1
Sulfur	0.8	1.9
Oxygen	35.5	44.9
Functional groups, meq/g		
Total acidity	6.7	11.3
COOH	3.6	8.2
Phenolic OH	3.1	3.1
Alcoholic OH	2.6	6.1
Quinonoid C=O	2.9	2.7
Ketonic C=O	2.9	2.7
OCH_3	0.6	0.8
E_4/E_6	4.8	9.6
Free radicals, spins/g $\times 10^{-17}$	2.85	0.64
Line width, gauss	4.5	3.5
g-value†	2.0043	2.0043

*From M. Schnitzer and S. U. Khan (eds.), *Soil Organic Matter*, Elsevier, 1978.
†Spectroscopic splitting constant.

parameters are similar (**Table 1**). *See* ELECTRON SPIN; FREE RADICAL.

Analytical characteristics of humins are similar to those of humic acids. The insolubility of humin in aqueous solutions arises from its being strongly retained or complexed by hydrous oxides and clay minerals.

Analysis of chemical structure. Aside from elemental and functional group analyses, the methods which have been used over the past century for obtaining information on the chemical structure of humic substances can be grouped into nondegradative and degradative methods. Nondegradative methods include different types of spectrophotometry and spectroscopy, x-ray analysis, electron microscopy, colloid-chemical, electrochemical, and radiochemical methods. Among degradative methods employed are oxidative and reductive degradation, hydrolysis, various types of irradiations, thermal analysis, and biological degradation. Of the many methods used, ^{13}C nuclear magnetic resonance spectroscopy and the oxidative degradation of methylated humic preparations have been especially informative.

Carbon-13 nuclear magnetic resonance spectra of humic substances provide inventories of the different components that make up these materials. Especially interesting is the 0–40-ppm region of the spectrum, which indicates the presence of paraffinic materials. The latter may constitute up to 35% of the weights of humic and fulvic acids. Important information on the identities of the major paraffinic components of humic and fulvic acids has been obtained by pyrolysis-field ionization and field desorption mass-spectrometric analyses of organic extracts of these materials. The latter contain *n*-alkanes (C_{17}–C_{101}), *n*-fatty acids (C_{15}–C_{34}), *n*-diols (C_{16}, C_{24}, C_{31}, C_{32}), sterols (C_{28}, C_{29}), *n*-alkyl monoesters (C_{40}–C_{68}), and *n*-alkyl diesters (C_{56}–C_{66}). In addition, humic acid extracts contain *n*-alkyl triesters (C_{75}–C_{93}). The highest molecular weight identified has a mass of close to 1500. The composition of the aliphatics in the humic acid and fulvic acid extracts resembles that of natural waxes. This suggests that humic materials contain significant amounts of waxlike materials that could positively affect the structural stability and water-retaining capacity of soils. The aliphatics present in the humic and fulvic acids could have originated from algae, fungi, bacteria, insects, earthworms, small animals, and plants. Because of their hydrophobic properties, these components would be expected to resist biodegradation and have long residence times in soils. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

The oxidative degradation of humic and fulvic acids and humins produces aliphatic carboxylic, phenolic, and benzene carboxylic acids. The most abundant aliphatic degradation products are monocarboxylic acids (acetic to *n*-caprylic acids); dicarboxylic acids (oxalic to azelaic acids); and other polycarboxylic acids (propanetricarboxylic acid, butanetetracarboxylic acid, and methylfurandicarboxylic acid). Major phenolic acids include those with one to three OH groups and one to five COOH groups per aromatic ring. Prominent benzenecarboxylic acids are the tri, tetra, penta, and hexa forms. Humic and fulvic acids have similar chemical structures; however, the oxidation of humic acids yields more benzenecarboxylic acids but fewer phenolic acids than that of fulvic acid. Major products resulting from the reductive degradation (by zinc-dust fusion) of humic and fulvic acids are methyl-substituted-naphthalene, -anthracene, -phenanthrene, -pyrene, and -perylene. *See* ANALYTICAL CHEMISTRY.

Molecular structure. The chemical structure for humic acid can be characterized as aromatic rings that are joined by alkyl chains of various lengths. Oxygen is present in the form of carboxyls, phenolic and alcoholic hydroxyls, ether and ketone groups, and nitrogen in heterocyclic forms and as nitriles. The oxidative degradation of this structure would produce benzenecarboxylic acids, and the reductive

degradation would yield aromatic polycyclics. The structure contains voids of various dimensions that could trap and bind other organics such as carbohydrates, proteins, lipids, and biocides as well as inorganics such as clay minerals and hydrous oxides. It is assumed that most of the carbohydrates and proteinaceous materials that are usually found in humic substances are adsorbed on the surfaces and in the voids and are not integral structural components. *See* ADSORPTION; CHEMICAL BONDING.

Chemical reactions. One of the most striking characteristics of humic substances is their ability to interact with metal ions, oxides, hydroxides, minerals, and organics, including toxic pollutants, to form water-soluble and water-insoluble associations of widely differing chemical and biological stabilities. The following types of reactions have been observed to occur in terrestrial and aquatic systems: formation of water-soluble simple metal complexes; formation of water-soluble mixed-ligand complexes; adsorption on and desorption from water-soluble mixed-ligand complexes; adsorption on and desorption from water-insoluble humic acids and metal-humate complexes; dissolution of minerals; adsorption on external mineral surfaces; and adsorption in clay interlayers. Through the formation of water-soluble complexes, humic materials can dissolve, mobilize, and transport metals and organics in soils and waters, while the formation of water-insoluble complexes brings about their accumulation in certain soil horizons and in sediments. Adsorption on clay surfaces and in clay interlayers stabilizes humic materials and protects them against chemical and biological decomposition over long periods of time. Thus, the formation of metal-humate and metal-fulvate complexes is important in soil genesis, for the formation of a good soil structure and for the availability of nutrients, especially those present only at microconcentrations. Humic substances can also interact with herbicides (such as atrazine), insecticides (DDT), and plasticizers (dialkylphthalates) by making them soluble in water and thus modifying their mobility and reactivity in terrestrial and aquatic environments. *See* COORDINATION CHEMISTRY; COORDINATION COMPLEXES; SOIL; SOIL CHEMISTRY.

Functions and uses. Humic substances enhance plant growth directly through positive physiological effects and indirectly by affecting the physical, chemical, and biological properties of soils. They also have nutritional effects in that they serve as sources of nitrogen, phosphorus, and sulfur for plants and microorganisms; biological functions in that they affect the activities of microorganisms; and physical functions in that they promote good soil structure, thereby improving tilth, aeration, and retention of moisture. All of these effects increase agricultural productivity. *See* NITROGEN CYCLE.

Humic substances are good chelating agents, have relatively large surface areas, are excellent dispersants, and act as reducing agents. Most of the uses that have been proposed for humic materials take advantage of these properties. Applications in various fields include (1) agriculture: additives to fertilizers and sprays, treatment and coating of seeds, and nutrients in hydroponics; (2) industry: in drilling muds for oil well rigs, as boiler scale removers, pigment extenders, emulsifiers, dispersants, corrosion inhibitors, wood preservatives and flotation reagents; (3) environment: in waste management, deodorizing liquids and gases, stack scrubbing, and absorption of herbicides; and (4) medicine: antimicrobial, anti-inflammatory, antitumor agents, liver stimulants, as well as medications for treating gastric ulcers, bleeding, and skin burns. *See* CHELATION; HYDROPONICS.

Morris Schnitzer

Biochemistry of Humus Formation

Humus is formed during the biodegradation or decay of plant, animal, and microbial organic residues in soils, swamps, and waters. It is a very important constituent of soils and has a number of beneficial properties. These include slow releasing of plant nutrient elements, especially nitrogen; improving soil physical properties; assisting trace-element nutrition of plants through chelation reactions; promoting solubilization of plant nutrients from insoluble minerals; having a high adsorptive or exchange capacity for nutrient cations; promoting growth via certain components; favoring heat absorption because of its dark color; increasing soil buffer capacity; supporting a greater and more varied soil population, which favors biological control; reducing the toxicity of both natural and anthropogenic toxic substances; and promoting increased soil water–holding capacity.

Although humus is a mixture of numerous organic substances, two types of polymers, humic acids and polysaccharides, constitute the major fractions. Humic acids appear to be complex polymers of hydroxyphenols, hydroxybenzoic acids, and other aromatic structures with linked peptides (proteins), amino sugar compounds, fatty acids, microbial cell wall and protoplasmic fractions, and possibly other constituents. The polysaccharide fraction is a complex mixture of polymers of a large number of sugar and sugar-type constituent units. *See* BIOPOLYMER.

The use of ^{14}C-labeled plant residues or specific residue constituents has made it possible to follow more precisely the residue decomposition and transformation processes and the stabilization in humic substances.

Plant residues. As soon as plant residues are returned to the earth and environmental factors such as temperature and moisture are favorable, microorganisms, insects, and worms begin to utilize them as food (carbon), nutrient, and energy sources. Some of the carbon is released as carbon dioxide (CO_2), some is used for synthesis of cell constituents (biomass) and products, and a portion is converted to relatively resistant humic acid–type polymers or to substances which are stabilized by complexing with inorganic metal ions or clays. Initially, about 40–60% of the carbon of readily available plant constituents will be converted to biomass and synthesized organic products. These are subject to decay as they are produced or when the cells die.

TABLE 2. Decomposition and stabilization in humus of organic residue carbons after 1 year in a fertile sandy loam top soil, in percent

Carbon	Glucose	Algal protein	Wheat straw	Polysaccharide from wheat straw	Cornstalk ring lignin C	Cornstalk OCH_3 lignin C	*Aspergillus glaucus* melanin
Added C evolved as CO_2	85	77	69	82	33	58	13
Residual C in biomass	16	11	7.4	10.8	0.5	1.2	0.2
Residual C in humic acid	20	22	34	24	66	58	71
Residual C lost upon 6 *N* HCl hydrolysis	82	78	75	73	9	14	5

Generally after 3 months, about 50–60% of the carbon of most crop residues and plant leaves returned to the soil will have evolved as CO_2. After a year, the carbon loss will have increased to about 55–70%. About 5–15% of the residual carbon will be present in soil biomass and 85–95% in the new humus. Residues having relatively high lignin contents degrade at a slower rate, and young succulent plant tissues and high-nitrogen residues such as legumes decompose a little faster. The bulk of the residual carbon will be present in components which are hydrolyzed in 6 *N* hydrochloric acid (HCl), primarily peptides or proteins, and polysaccharides. A smaller portion will be present in aromatic polymers. The aromatic units are derived primarily from the plant lignins and through microbial synthesis (**Table 2**).

Biodegradable substances. Readily available, small molecular substrates or plant constituents such as sugars, amino acids, aliphatic acids, and pyrimidines are metabolized within a few hours or days and carbon is rapidly evolved as CO_2. After a year, 80–90% of the original carbon will have evolved as CO_2. About 10–20% of the residual carbon will be in the form of soil biomass and 80–90% in new humus. With time, the percentages of residual substrate carbons in biomass will decline and the percentage in humus will increase. In most soils, the biomass constitutes about 2–4% of the organic carbon.

About 20% of the residual carbon from readily biodegradable substrates will be associated with the humic acid fraction of humus, and possibly 10–20% of this carbon is present in aromatic units. The bulk of the residual carbon, however, will be present in the form of peptides and polysaccharides, which are released as sugar or amino acid units upon acid hydrolysis. This would be expected as the major portion of the metabolized carbon not released as CO_2 would be transformed into microbial protoplasm, cell wall polymers, and polysaccharides that could be partially linked into humic molecules during autolysis and decay of the cells. *See* AMINO ACIDS; PEPTIDE; POLYSACCHARIDE.

Polysaccharides, proteins, and lipids. Sixty percent or more of most organic residues consist of cellulose and other polysaccharides. Some residues such as legumes and microbial tissues contain 6–65% protein. Specific plant and microbial tissues may also contain appreciable amounts of lipids. Most of these materials are highly biodegradable, but especially during the early stages of decomposition many decompose a little slower than the simple sugars and acids (Table 2). But after 6 months or a year, about 70–85% of the carbon will have evolved as CO_2 and about 6–16% of the residual carbon will be present in soil biomass and 84–94% in new humus. The small decrease in loss of carbon as CO_2 compared with the more simple substrates indicates that some of the original polymer carbons, or more likely, partially degraded units are stabilized by incorporation into the new humus. *See* CELLULOSE; LIPID.

Phenols. Simple phenolic substances and other aromatic compounds may be present in plant and microbial residues and are released during biodegradation of aromatic polymers such as lignins. Some soil fungi synthesize up to 30 or more phenolic compounds. Phenolic substances are degraded by numerous species of organisms; however, in soil, the carbon losses are not as great as would be expected. After 3 months losses vary from about 10 to 71%. This indicates that a portion of intact ring molecules are stabilized in humus. The more reactive the compound with respect to radical formation or oxidative polymerization reactions, the greater the stability. Also, the more reactive the phenol, the smaller the percentage of residual carbon incorporated into biomass. *See* PHENOL.

Lignins. This is the second most abundant polymer synthesized by plants. It is an important source of structural units for humus formation. During one year, about 15–30% of the ring and 2-side-chain carbons of model and cornstalk lignins are evolved as CO_2. This compares with about 30–50% loss of 1- and 3-side-chain and methoxyl carbons. During the second year, the losses are greatly reduced.

Whereas the bulk of the residual carbons from readily available substrates are associated with peptide and polysaccharide polymers, the major portion of the residual lignin carbons are associated with aromatic complexes. If most of the aromatic rings were cleaved by microbial enzymes, a substantial portion of the ring carbons would have been utilized by the microbes and would be present in the peptide and polysaccharide fractions of the humus. Biomass estimations also support this conclusion: after 6 months or 1 year, about 5–15% of the residual carbons from the readily available substrates are present in biomass compared to about 0.2–1% for the residual lignin carbons. *See* BIOMASS.

Lignin biodegradation and stabilization studies indicate that lignin is indeed an important substrate

for humus formation, but it undergoes very profound changes during humification. In peat bogs and swamps, greater quantities of aromatic lignin–derived structures persist as most organisms cannot readily cleave the benzene ring under anaerobic conditions. *See* LIGNIN.

Melanins. Many organisms, including fungi, bacteria, and actinomycetes, synthesize dark polymers called melanins. The melanins from *Epicoccum nigrum, Aspergillus sydowi, Hendersonula toruloidea,* and *Eurotium echinulatum* are similar to soil humic acids with respect to elemental composition, reactive chemical groups, amino acids released upon acid hydrolysis, phenols released upon sodium-amalgam reductive degradation, resistance to microbial degradation, types of structures released upon oxidative degradation and pyrolysis, nuclear magnetic resonance spectra, and low polysaccharide content. It is therefore very likely that these and other fungal melanins contribute to humic acid formation.

Humic acid formation. The ability of numerous phenolic compounds to undergo enzymatic and autoxidative polymerization reactions is probably of major importance in the formation of humic acid molecules. During the early microbial metabolism of these compounds, transformations such as beta oxidation of side chains, decarboxylation, demethoxylation, and formation of additional OH groups form a large variety of phenolic substances. Many of these such as dihydroxy and trihydroxy phenols readily autoxidize to form polymers at pH values of 6 and above. Also these and less reactive phenols such as ferulic acid, coniferyl alcohol, vanillic acid, and orcinol are readily oxidized by microbial phenolases and peroxidases present in soil or synthesized by soil organisms. Less reactive phenols such as *p*-hydroxybenzoic and 2,5-dihydroxybenzoic acids are linked into the polymers if present at the reaction sites. The phenol oxidation process forms radicals that stabilize by linkage to form dimers, or on further oxidation forms quinones which link other phenols or substances with free amino groups such as peptides or amino sugar polysaccharides, and microbial cell wall structures through nucleophilic addition reactions. Condensed ring aromatic structures and parts of lignin and melanin molecules may be linked into the humic polymer molecules. Model humic acids prepared by the enzymatic polymerization of phenol mixtures and peptides are similar to natural humic acids. *See* AUTOXIDATION; FREE RADICAL.

Polysaccharide fraction. Inasmuch as most plant polysaccharides readily decompose, it is probable that the major portion of this humus fraction originates through microbial synthesis. Also polysaccharides contain amino sugar units which are synthesized by many microorganisms but are not generally present in plant polysaccharides. Microbial polysaccharides are also subject to decomposition, but the rate varies with specific polymers. Several factors may explain how polysaccharides are stabilized in the soil humus. The polysaccharide fraction of humus contains about 10% uronic acid units. These and units with *cis*-hydroxyls or phosphoric acid esters form salts or complexes with di- and trivalent metal ions which may greatly reduce their susceptibility to degradation. Also, polysaccharides may be complexed with clay minerals through metal ion-clay linkages which increase resistance to biodegradation. Complex polysaccharides containing amino acid or amino sugar units may be stabilized by linkage into humic acid polymers through nucleophilic addition to quinones, or strong hydrogen bonding to humic acids. Chitosan, which is a polymer of amino sugar units, and *Chromobacterium violaceum* polysaccharide, which contains glucosamine, are much more resistant to biodegradation than many other plant and microbial polysaccharides. Polysaccharides may further be protected inside dense aggregates, and highly branched polymers may be more resistant than straight-chain structures.

Hydrolysis procedures indicate that soil polysaccharide fractions normally contain 10 or more major sugar units and many others in smaller amounts. Attempts to isolate significant fractions with fewer structural units have been unsuccessful. Generally it has been concluded that it is extremely difficult to separate a complex mixture of polysaccharides into single types. Another possibility is that microbial and plant polysaccharides at all stages of decomposition could be recombined through the action of soil enzymes stabilized in humus or released during cell autolysis. Those structures which resist biodegradation or which readily form complexes with clays and metal ions would persist and contribute to the humus polysaccharide fraction. *See* SOIL MICROBIOLOGY.

James P. Martin

Bibliography. W. R. Jackson, *Humic, Fulvic, and Microbial Balance: Organic Soil Conditioning,* 1993; P. MacCarthy et al. (eds.), *Humic Substances in Soil and Crop Sciences: Selected Readings*, 1990; E. A. Paul and A. D. McLaren (eds.), *Soil Biochemistry*, vol. 4, 1975; F. J. Stevenson, *Humus Chemistry*, 2d ed., 1994; R. L. Tate III, *Soil Organic Matter: Biological and Ecological Effects*, 1987, reprint 1992.

Hunger

A term most commonly used to refer to the subjective feelings that accompany the need for food; however, the study of this topic has come to include consideration of the overall control of food intake. More specifically, experimental work on the problem of hunger has been concerned with the sensory cues that give rise to feelings of hunger, the physiological mechanisms that determine when and how much food will be ingested, and the mechanism governing the selection of the food to be eaten.

Cannon theory. The earliest experimental approach to these problems concentrated almost exclusively on the question of the sensations of hunger or, as they have come to be known, hunger pangs.

The early work of W. B. Cannon, which resulted in the so-called local theory, led to the conclusion that for both hunger and thirst the appropriate sensations arose peripherally in the body. According to theory, hunger sensations (pangs) arose from stomach contractions that stimulated local sensory nerves. Subsequent work substantiated the idea that increased stomach contractions do indeed accompany the state of hunger in ordinary circumstances. It seems unlikely, however, that these contractions contribute substantially to the detailed control of eating behavior. For example, when the sensory nerves from the stomach are cut or even when the stomach is altogether absent, eating behavior can go on in an essentially normal manner. Whatever influence stomach contractions may have on the ingestion of food, it is known that the motility of this organ can be controlled by both a neural and a hormonal route. The hormone involved may be one that is secreted by the stomach itself. *See* HORMONE; NERVOUS SYSTEM (VERTEBRATE).

Physiological mechanisms. Food consumption is basically controlled by the organism's nutritional status. Food deprivation leads to eating, and the ingestion of food materials terminates hunger sensations. The issues are to determine which physiological processes vary quantitatively with nutritional status, and to find out if these changes can be detected by the nervous system in a manner that would instigate and terminate food consumption. Attempts by researchers to find a simple humoral factor that might be involved linearly in this regulation have not been illuminating. In fact, not a single humoral factor has been identified that can be reliably linked to feeding onset as it occurs spontaneously.

Blood-sugar level. Blood-sugar level, which has received more attention than any other factor, can be used as a case in point. The concentration of blood sugar does indeed vary appropriately in a general way with the periodicity of the food cycle. Moreover, hyperglycemia (extremely high blood sugar) and hypoglycemia (extremely low blood sugar) have been observed to decrease and increase hunger, respectively. Detailed analyses of normal life variations of blood sugar, however, reveal that the relation between the concentration of blood sugar and hunger is not sufficiently close for this single humoral factor to be able to control hunger in any simple and direct manner. Moreover, feeding following injections of insulin occurs only after lethal levels of hypoglycemia have been reached. *See* CARBOHYDRATE METABOLISM.

Tissue utilization of food. The evaluation of more local tissue utilization of food has proved a more promising approach to this problem. There is now some evidence suggesting that the status of the liver is pivotal in the control of feeding. Depletion of liver glycogen stimulates feeding; its repletion terminates feeding in rats and rabbits. The control exerted by these hepatic signals may be mediated by the vagus nerve, as complete destruction of the hepatic vagus markedly decreases feeding. *See* LIVER.

Feeding termination. Many stimuli that terminate feeding have been identified. Eating in food-deprived animals is inhibited by the reduction of either cellular water or of plasma fluid. It is also reduced by gastric distension and by infusing nutrients into the intestine and into the systemic, especially venous hepatic, circulation. Satiation produced by nutrient absorption from the intestine may be mediated, in part, by the gut hormone cholecystokinin. Cholecystokinin's effects may be central because injections of this hormone (albeit in very high doses) into the brain ventricles reduce feeding. It is more likely, however, that the cholecystokinin is effective because it reduces the rate at which food passes through the stomach. Generally, inhibitory signals in the control of feeding are more effective when they are brought about in the context of the feeding act than when introduced outside this context.

Neural centers. The previously held notions of discrete neural centers for the onset and termination of feeding have been abandoned, as the complexity of the feeding act and its corresponding neural complexity have become more widely appreciated. Deficits in feeding produced by lesions to the lateral hypothalamus are now recognized to be part of a larger constellation of behavioral impairments. To be sure, destruction of the lateral hypothalamus results in a loss of feeding and drinking behaviors that only slowly recover. But these animals also suffer from a variety of other profound behavioral incapacities. They do not attend to sensory stimuli. They have severe motor impairments (for example, they do not even groom) and cannot be easily aroused. They are deficient in learning. Conversely, although electrical stimulation of the lateral hypothalamic area induces feeding and drinking, it often also leads to participation in other behaviors, such as nest building, carrying, and copulation—even in a given animal.

Other brain areas have been implicated in feeding control. Destruction of the ventromedial nuclei of the hypothalamus leads to overeating, but again the deficits are not specific. Such rats are vicious, hyposexual, and do not attend to their young. There are also metabolic consequences of the lesion that may cause overeating. In short, more recent work has shown that the idea of hypothalamic feeding and satiety centers is an oversimplification and that, seemingly, the entire brain is involved in feeding control. Even decerebrate rats (rats that have had their entire brain removed, from the hypothalamus forward) make very sensitive discriminations about the quality of solutions placed in their mouths. Like intact animals, they reject quinine or very salty solutions and avidly accept the sugars.

Neurochemical control of feeding. It appears that increasing the rate at which the transmitter norepinephrine is taken up by various tissues in the forebrain increases food intake. Conversely, increasing the transmission of serotonin decreases feeding. Again, the cautions raised about the brain lesion studies must be sounded here. The specificity of effects are not known. This is especially problematic because only a small fraction of the neurotransmitters that exist in the central nervous system have been identified.

Specific hungers. Deprivation of certain, specific food substances precipitates an increased appetite for the needed substance. This so-called specific hunger behavior has been demonstrated experimentally with many substances, such as salt, calcium, fats, proteins, and certain vitamins in children and in the lower animals studied. Theories to account for the control of eating by specific food deprivations have suggested the possibilities of learning, differentially lowered taste thresholds, and direct effects upon the brain.

It is now clear that only the hunger for salt in salt-deprived animals appears before the animal has learned about the beneficial consequences of salt ingestion. Specific hungers for other minerals, proteins, and vitamins appear only gradually and reflect the animal's learning that certain foods are no longer beneficial and, in fact, may be harmful. The evidence that animals learn that the "right" food is actually beneficial to them is meager. They appear to stop eating the food associated with the deficit and sample new foods, stopping when the one that corrects the deficit is eaten for a while. This behavior is in contrast with salt appetite, which increases immediately in sodium-deficient animals.

Feeding development. It may be tentatively concluded, in rats at least, that suckling in these animals, until about 2 weeks of age, is not under the control of internal stimuli that manage adult feeding. Suckling does, however, come under the control of these stimuli at about the time that rats start to eat from the environment. Portions of the feeding system appear to be present even in the newborn, however. When tested at very warm temperatures (90°F or 32°C), 3-day-old rats actually eat a liquid diet, and the amount eaten is proportional to the length of time that they have been food-deprived. Such animals, however, do not eat at room temperature. Feeding on a variety of diets at room temperature occurs at about the start of the third week after birth. *See* THIRST AND SODIUM APPETITE.

Elliott M. Blass

Huntington's disease

A hereditary disorder of the basal ganglia causing progressive motor incoordination, abnormal involuntary movements (chorea), and intellectual decline. The disease, which progresses gradually over 15–20 years, is invariably fatal.

Huntington's disease is a rare disorder that affects approximately five to ten of every 100,000 persons. Inherited as an autosomal dominant mendelian trait, it inevitably develops in those who carry the gene if they live long enough. Men and women are affected equally. The average age at onset is between 35 and 40 years, but the disease can begin as early as 2 years or as late as 80 years. The mutation rate is very low, and no confirmed new mutations have ever been reported. The disease appears to have originated in northern Europe, England, or Scandinavia. Pathologically, the nerve cells in the caudate nucleus and putamen degenerate, and although other regions of the brain are affected to a minor extent, these regions are the most seriously damaged.

The disease begins insidiously with restlessness, choreiform (dancelike) movements, and personality changes (irritability, depression, and occasionally psychosis). The symptoms gradually worsen, with severe involuntary, large-amplitude motions, grimacing, poor balance, and very poor fine-motor coordination. Abnormalities of eye movement also occur early in the disease and worsen with its progression; speech and swallowing are always affected. Death frequently results from secondary causes such as pneumonia. About 10% of those affected have a rigid form of the disease in which choreiform movements are less prominent, and stiffness and slowness of movement and abnormal posturing (dystonia) are the dominant features. This variety is more likely to occur when the disease begins before the age of 20 years. Therapy is merely supportive: no medications significantly affect the course of the disease or functional capacity of the sufferer. Depression or psychosis, however, can be temporarily alleviated by antidepressant and antipsychotic medications. A progressive intellectual decline is usually observed, but even in the late stages of the disease, patients can often recognize family members and interact with other people to a limited extent.

The gene for Huntington's disease has been localized at the end of the short arm of chromosome 4. The gene, termed IT15, contains a series of three nucleotides that are repeated from 11 to 30 times. On Huntington's disease chromosomes, this triplet repeat is expanded and occurs in greater numbers (37 to over 86 repeats). The IT15 gene itself is very large (10 kilobases) and is thought to encode a protein of approximately 3144 amino acids. The predicted protein, huntingtin, does not resemble any known protein. Mutations to Huntington's disease are rare but do occur. Offspring of individuals suffering from the disease have a 50% risk of developing it, and can be tested by recombinant genetic technology. *See* CHROMOSOME; MUTATION.

The diagnosis of symptomatic Huntington's disease is usually straightforward if a family history of the disease is confirmed through autopsy. The diagnosis can become more complicated, however, if affected relatives have died prematurely or have been misdiagnosed. In the absence of a clear family history, a tentative diagnosis can be made on the basis of the clinical signs and symptoms after other disorders with similar manifestations have been ruled out. *See* HUMAN GENETICS; NERVOUS SYSTEM DISORDERS.

Anne Young

Bibliography. S. E. Folstein, *Huntington's Disease: A Disorder of Families*, 1989; J. F. Gusella et al., A polymorphic DNA marker genetically linked to Huntington's disease, *Nature*, 306:234–238, 1983; Huntington's Disease Collaborative Research Group, A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes, *Cell*, 72:971–983, 1993; J. B. Martin and J. F. Gusella, Huntington's disease: Pathogenesis and management, *N. Engl. J. Med.*, 315:1267–1276,

1986; J. P. Vonsattel et al., Neuropathological classification of Huntington's disease, *J. Neuropathol. Exp. Neurol.*, 44:559–577, 1985; A. B. Young et al., Huntington's disease in Venezuela: Neurologic features and functional decline, *Neurology*, 36:244–249, 1986.

Hurricane

A tropical cyclone whose maximum sustained winds reach or exceed a threshold of 74 mi/h (119 km/h). In the western North Pacific Ocean it is known as a typhoon. Many tropical cyclones do not reach this wind strength. *See* CYCLONE.

Maximum surface winds in hurricanes range up to about 200 mi/h (320 km/h). However, much greater losses of life and property are attributable to inundation from hurricane tidal surges and riverine or flash flooding than from the direct impact of winds on structures.

Tropical cyclones of hurricane strength occur in low latitudes of all oceans except the South Atlantic and the eastern South Pacific, where combinations of cooler sea temperatures and prevailing winds whose velocities vary sharply with height prevent the establishment of a central warm core through a deep enough layer to sustain the hurricane wind system. *See* WIND.

Impact. Severe hurricanes are responsible for many of the world's greatest natural disasters. Hurricane Camille in 1969 devastated coastal communities of the central Gulf of Mexico; Hurricane Tracy on Christmas 1974 virtually destroyed the Australian city of Darwin; in 1989, Hurricane Hugo, the most destructive atmosphere event of all times, caused $7 billion damage in the Caribbean Sea and South Carolina; and the Bangladesh hurricane of 1970 killed more than 300,000 people. Lesser hurricanes often inflict years of social disruption and economic devastation in small tropical countries.

In the United States, hurricane damage continues to climb, primarily because of increasing exposure of property as the population continues to migrate to the seashores. At the same time, loss of life has decreased sharply because of more effective warnings which reflect the result of extensive research, more complete storm surveillance, and improved programs of public awareness.

On a global scale, the impact of hurricanes may change in the decades ahead as a consequence of increasing carbon dioxide in the atmosphere which may reduce radiative heat losses from the ground—the greenhouse effect. While the impact of global incidence of hurricanes remains controversial, the most likely change is an increase in the number of hurricanes at higher latitudes, owing to possible increases in sea surface temperatures. However, it will remain difficult to distinguish normal climatic aberrations from changes attributable to greenhouse warming for at least several decades. In some regions, the number of hurricanes could decrease. *See* GREENHOUSE EFFECT.

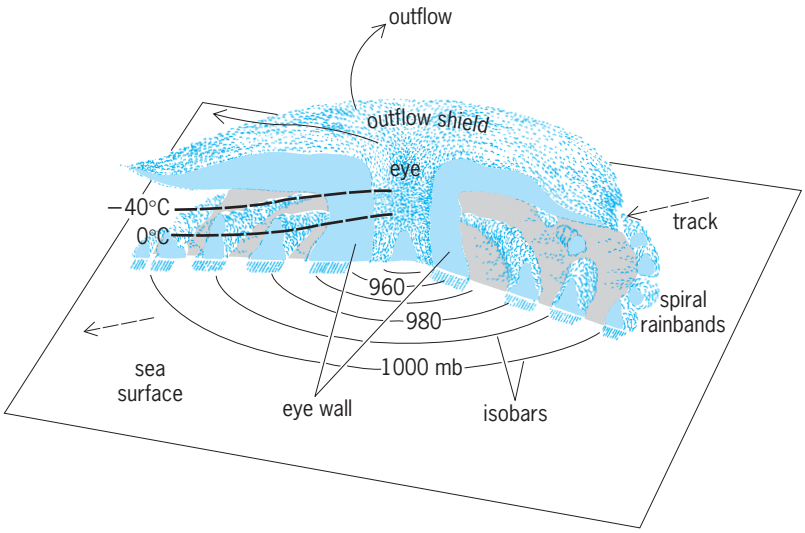

Fig. 1. Model of a hurricane circulation and cloud structure. °F = (°C × 1.8) + 32.

Structure. The characteristic signature of a hurricane (**Figs. 1** and **2**), as viewed by radar or satellite, is of a central, relatively cloud-free eye encircled by an eye wall of towering cumulonimbus clouds 15–50 mi (25–80 km) wide. A dense cloud deck covers the storm at upper levels, and spiral bands of

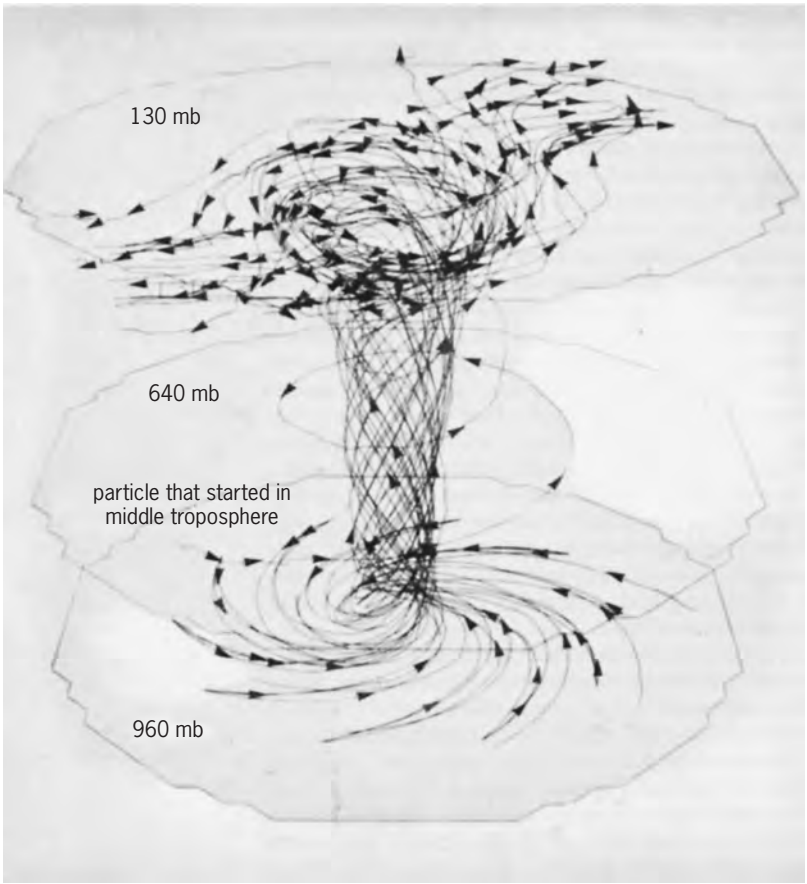

Fig. 2. Particle trajectories calculated over an 8-day period (90–282 h in 9-h intervals) in an experiment with a three-dimensional model hurricane. All particles start in the lower atmospheric boundary layer except one, which starts in the middle troposphere. 1 mb = 10^2 Pa. (*After R. A. Anthes, S. L. Rosenthal, and J. W. Trout, Preliminary results from an asymmetric model of the tropical cyclone, Mon. Weath. Rev., 99:744–758, 1971*)

cumulus clouds wind inward from the environment to the eye wall. *See* RADAR METEOROLOGY; SATELLITE METEOROLOGY.

Near the Earth's surface the hurricane appears as a nearly circular vortex of low pressure, typically 200–400 mi (320–640 km) in diameter. Its dynamic and thermodynamic properties, however, are asymmetrically distributed both at the surface and in upper layers. The cyclonic circulation (counterclockwise in the Northern Hemisphere) extends through most of the troposphere (to an atmospheric depth of about 9 mi or 15 km). In lower layers (up to 2 mi or 3 km), winds spiral inward and accelerate toward lower pressure, reaching peak speeds in a narrow annulus typically 10–20 mi (15–30 km) from the pressure center. Here there is a near balance between the pressure forces acting radially inward and the centrifugal and Coriolis forces acting outward, so that the air, no longer able to move radially, is forced upward, forming and maintaining the cloudy eye wall.

Momentum is transported upward from surface layers, so that wind speeds near the top of the eye wall are almost as strong as peak winds near the surface. Since pressure forces in the relatively warm eye wall must diminish with height, the consequent imbalance of forces causes the air in upper layers to spiral out of the vortex and join environmental circulations, carrying with it a canopy of cloud debris which may extend great distances into the environment. The raging winds in the eye wall, so closely adjacent to the relative calm within the eye, act as a centrifuge, dragging air from the eye which is replaced from above, the sinking motion therefrom gradually filling the eye with very dry, often cloud-free, warm air—warmer, in fact, than in the moist eye wall.

The spiral bands of cumulus clouds extending from the environment and entwining the eye wall result from complex processes. Some portions of these bands grow tall enough to augment the outflow from the eye wall. *See* CLOUD PHYSICS; CORIOLIS ACCELERATION; TROPOSPHERE.

Energy sources and processes. The hurricane draws energy from two main sources, its atmospheric environment and the ocean surface.

In the hurricane vortex, the circulation of mass—inward at low and midlevels, upward and outward aloft—at the rate of some 2.2 million tons (2 metric tons) per second, constitutes an atmospheric heat pump involving the import of angular momentum from the environment. The primary source of fuel for this pump is a combination of the latent heat released by the formation of towering cumulus clouds and heavy precipitation, and the much smaller but crucial flux of sensible heat from the ocean. Since low-level air, flowing from the environment into the storm core, is observed to maintain a nearly constant temperature, an important up-flux of heat energy from the warm ocean surface tends to compensate for the cooling that otherwise would occur during the transit toward lower pressure. Without this oceanic heat source, the pump would be unable to generate and sustain wind speeds of hurricane strength. The loss of this compensating sea-to-air heat flux is the primary reason for rapid decreases in strength usually observed when a hurricane moves over land. *See* ANGULAR MOMENTUM.

The process by which the tall cumuli in the eye wall maintain the warm, light air in the hurricane core, and in turn the pressure forces that determine the strength of hurricane winds, is not a simple matter of releasing latent heat within the clouds. Latent heat released by cumulus clouds does not directly warm the atmosphere, since most of this heat energy is converted to potential energy as air parcels rise. However, the direct role of the latent heat released is to maintain the buoyancy of the rising cloud air and thus to maintain the vertical circulation of the eye wall. The high temperatures at the storm center are created and maintained by adiabatic compression of descending air.

Since latent heat from atmospheric water vapor is the major fuel for the hurricane heat engine, the efficiency and available fuel supply is dependent upon both the atmospheric and oceanic environment. From a purely thermodynamic point of view, the overall hurricane efficiency is a function of the temperature difference between the ocean surface and the cold upper atmosphere where outflowing air, mixing with a colder environment, is further cooled by long-wave radiation to space. Thus warm tropical oceans provide the best medium for generating and sustaining the most intense hurricanes. *See* HEAT; PSYCHROMETRICS.

Hurricane intensity is also constrained by the amount of preexisting rotation in the environment. This rotation is normally quantified in terms of angular momentum, a combination of that from the Earth's rotation and that of environmental flow. If there is a store of cyclonic angular momentum and the necessary processes are triggered to set the development in motion, the low-level inflow described earlier carries this momentum toward the vortex center, generating hurricane-force winds in the same manner that rapid spin is achieved when ice skaters bring their arms close to their bodies. Some momentum is lost to friction at the Earth's surface, so that air flowing out of the eye wall returns to the environment with an anticyclonic motion. *See* ANGULAR MOMENTUM; VORTEX.

The degree of rotation determines the potential for releasing latent heat that can be used to accelerate the winds. In the undisturbed tropics, rotation is very low, so that the latent heat released by daily-observed cumulus clouds and thunderstorms diffuses away, unable to organize a warm core and the consequent pressure drops necessary to produce storm-scale wind systems. Similar amounts of latent heat are released inside hurricanes, but the high rotation ensures that much of this heating is retained locally to lower the pressure.

Formation. An average of about 70 tropical cyclones develops gale or hurricane force somewhere on Earth each year, a figure some 15–20% higher than climatological records indicated before the advent of weather surveillance satellites in the 1960s. The global distribution of tropical cyclone occurrence is shown in **Fig. 3**. *See* METEOROLOGICAL SATELLITES.

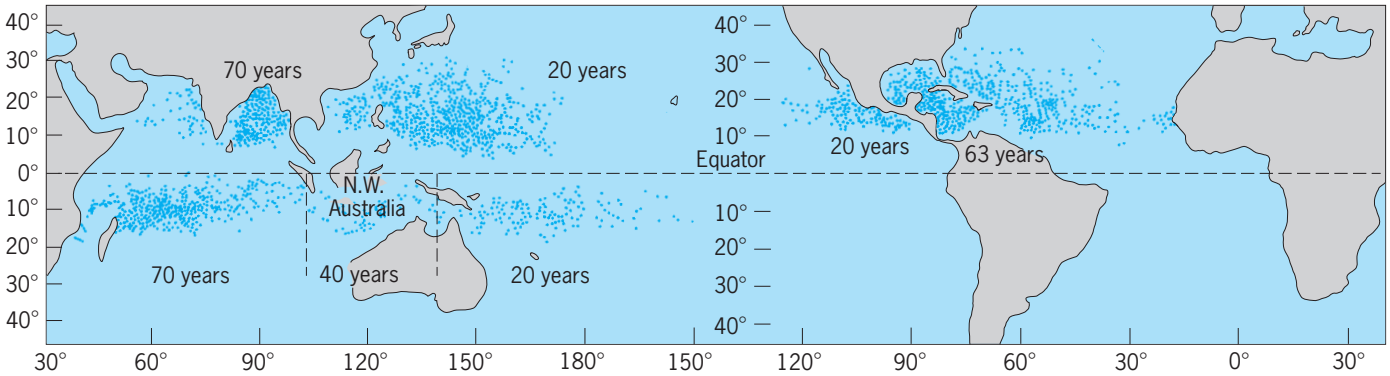

Fig. 3. Points on the globe where tropical cyclones were first detected by weather observers. (*After W. Gray, Global view of the origin of tropical disturbance, Mon. Weath. Rev., 96(10):670, 1968*)

Tropical disturbances, sometimes referred to as seedlings, that breed cyclones of hurricane strength originate mainly in tropical latitudes in the vicinity of the equatorial or monsoonal trough. Some have their origins over continental areas. Seedling disturbances, comprising an agglomeration of convective clouds 180–300 mi (300–500 km) in diameter, often move more than 1200 mi (2000 km) across tropical oceans as benign rainstorms before developing closed circulations and potentially dangerous winds. *See* MONSOON METEOROLOGY; TROPICAL METEOROLOGY.

A number of well-known factors may influence the development of seedlings. The atmosphere must be able to sustain and consolidate the growth of cumulus clouds to release latent heat and cause pressure falls. Also, ocean temperatures must be high enough to enhance the heat content of environmental air entering the system. It was established in the 1950s that this required an ocean temperature of greater than 79°F (26°C). Because of the need for background rotation, hurricanes rarely form or retain their identity equatorward of 5° latitude (Fig. 3).

The variation of prevailing winds with height also is important. If there is a sufficient change in wind velocity with height, the heat released by convection and the warming due to subsidence cannot be stored in vertical columns of sufficient height to produce the pressure drop required to sustain hurricane winds. Finally, conditions favorable to development in the low-level flow must be matched by an upper-layer environment that supports outflow from the convective system.

Nevertheless, even when these necessary conditions are satisfied, many seedlings fail to develop. In the North Atlantic, for example, 100 seedlings form each year (more than half emerging from Africa), yet only 9 develop into tropical cyclones. Knowledge of the complex interactions involved is not sufficient to explain this relatively low incidence of hurricanes.

Movement. The observed movement of hurricanes arises from two mechanisms, advection and propagation. Advection occurs when the environment directs an airflow over the hurricane that moves it in a fashion similar to a cork bobbing in a stream. Forecasters call this the steering current, and considerable effort is made to determine it accurately. Although there are many individual differences, the mean flow throughout the middle levels of the troposphere and over a radial band some 300–400 mi (500–650 km) from the center seems to provide the most consistent results.

Dynamically, the hurricane propagates by internal processes that are quite sensitive to the properties of the environment. A full understanding of many of these processes requires in-depth knowledge of fluid dynamics. One illustration, however, can provide an indication of the essential features. Any cyclone is made up of infinitesimally small rotating elements that are measured as vorticity. For a hurricane, maximum combinations of elements, and thus maximum cyclonic vorticity, is found near the center. Air that is stationary relative to the Earth also has cyclonic rotation and vorticity arising from the component of Earth rotation about the local vertical. The Earth's vorticity component increases from zero at the Equator to a cyclonic maximum at each pole. The flow around a hurricane thus brings increasing cyclonic vorticity to the west side, and decreasing cyclonic vorticity on the east side to produce a poleward flow over the tropical cyclone. This effect combined with associated internal rearrangements causes the hurricane to propagate westward and poleward. The magnitude of this propagation is defined by the outer structure of the hurricane; it is largely insensitive to the hurricane intensity.

The propagation velocity is generally about 3–5 mi/h (5–7 km/h) westward and poleward, but it may vary from near zero to over 10 mi/h (16 km/h) for a hurricane of large diameter. By comparison, most tropical cyclones reach hurricane intensity in the steady trade-wind belt where they are advected westward at a speed of 10–20 mi/h (15–30 km/h). As the hurricane approaches the western extremity of a subtropical ridge and begins moving poleward, the steering current first weakens, and then is replaced by westerlies whose speeds often exceed 40 mi/h (65 km/h). Thus cyclone propagation and advection tend to be in the same direction in the tropics but to act oppositely in midlatitudes.

Numerical simulations. Numerical simulation models have become a powerful tool for the study of hurricane dynamics and energetics. These models consist of finite-difference expressions for the

Fig. 4. Three-dimensional view of Hurricane Gilbert's eye and cloud system viewed from a weather satellite on September 12, 1988. (*F. H. Hasler, NASA Goddard Laboratory for Atmospheres*)

system of partial differential equations that govern the dynamics and energetics of the atmosphere. Pioneer experiments with computer models of hurricanes were carried out in the late 1950s. The limited computing capacity restricted hurricane models to be axisymmetric (assuming that hurricanes had no variations in horizontal azimuthal directions). These models were used to study the effects on hurricane intensity of sea-surface temperature, ambient atmospheric temperature and humidity stratification, and various hurricane modification strategies.

By the early 1970s, more advanced computers permitted development of fully three-dimensional models. Figure 2 shows the hurricane circulation obtained with the first of these models. By 1984, some research groups were using three-dimensional numerical models of the hurricane and its environment covering thousands of miles; they used telescoping, or nested, grids to provide fine resolution of mesoscale features near the eye wall, those with finest grid spacing moving with the hurricane. Topics of investigation ranged from the effects of vertical wind shear on hurricane development to the detailed changes of the hurricane's structure as it made landfall. Numerical models have been developed that cover the entire Earth with sufficient resolution to simulate many of the aspects of tropical cyclones. Some sophisticated models use nested grids which allow global integrations of the equations of motion and thermodynamics with coarse resolution while simulating the detailed circulations and cloud systems of tropical cyclones with very high resolution. An operational global model in use at the European Center for Medium Range Weather Forecasting (at Reading, England) uses powerful enough computers with sufficiently small grid spacings to show useful results in anticipating hurricane development and movement without using nested grids.

Prediction. Prediction of the path of the hurricane center as a function of time is the most important aspect of the hurricane forecast. If the track prediction contains major errors, all other aspects of the forecast (intensity, structure, rainfall, and so forth) are of no consequence. Because the high-energy (damage-producing) portion of the storm has small dimensions (60–90 mi or 100–150 km), the forecaster, in landfall situations, must balance the consequences of warning too small an area and missing the landfall, against overwarning a large area and thereby causing unnecessary expensive preparations by the public. The prediction problem is made especially difficult by the sparse coverage of the oceans by conventional surface and upper-air meteorological data. Meteorological satellites provide some data over these regions. Indeed, when storms are well away from landfall, satellites are often the only source of information concerning location of the storm. For the east and Gulf of Mexico coasts of the United States, center fixes are made several times a day by aircraft penetration when the storm is within 36 h of landfall. *See* STORM DETECTION.

An array of objective models is available to the hurricane forecaster for predicting the path of the hurricane. The National Oceanic and Atmospheric Administration's National Hurricane Center in Miami continues to make regular use of at least four models. However, these methods often yield predictive results that differ by a significant amount. The forecaster is then faced with the problem of deciding which is to be given greatest weight in the official forecast.

Forecast models generally fall into three categories. There are statistical models based solely on climatology and persistence. These produce a most probable hurricane track, the only current data required being the initial location of the storm center, past movement of the center, and the calendar date. A second class of statistical models adds, to climatology and persistence, information concerning the large-scale pressure field in which the storm is embedded. The third class of models is based entirely on atmospheric dynamics. These models are similar to, though less sophisticated than, the simulation models described earlier. They differ in that the initial conditions for simulation studies are generally a simplified idealized state of the atmosphere, whereas the dynamical prediction models have initial conditions based on actual observations of the current state of the atmosphere. Generally speaking, the more advanced dynamical models perform better than statistical models for longer-range forecasts (36–48 h), while the statistical models perform better for shorter-period predictions (12–24 h). *See* MODEL THEORY.

Averaged over the Atlantic Ocean, Caribbean Sea, and Gulf of Mexico, the magnitude of the prediction

error (length of the line connecting the observed position of the storm center to the predicted position) for the official forecast by the National Hurricane Center is 109, 244, and 377 mi (202, 452, and 698 km), respectively, for forecasts of 24, 48, and 72 h. From the mid-1950s through the 1960s, a decrease of about 10% was observed in the average errors of official forecasts. Since then, however, the principal measure of improvement has been a reduction in standard deviations from the mean, as larger individual errors were significantly reduced.

The notable progress in technology and of data-processing capabilities for meteorological satellites offers great promise for hurricane research and for the design of new dynamically based predictions of both development and movement. **Figure 4**, an example of the observational capabilities of modern satellites, is a three-dimensional view of the cloud systems in Hurricane Gilbert, the most powerful Atlantic hurricane of record, displaying details of cloud structure and implications for both the details of circulation and the thermodynamic properties of the hurricane inner core. Under development is an array of microwave sounders to be installed on weather satellites that will "see through" the clouds to sense the three-dimensional thermal and circulation structure of hurricanes and its interacting environment, and for measuring rainfall distributions. Near the end of the century, satellites are planned which will be able not only to measure the winds throughout the storm core but also to observe the core's thermal structure. Such observations will have a significant impact on the understanding of energetic and dynamic processes which control the degree of development and severity of a hurricane; and they may well provide data for initializing more sophisticated prediction models incorporating far more comprehensive physics than was possible with the previous models. With the increasing archives of satellite data from existing sensors, there are already promising opportunities for prediction research, applying well-known methodologies for principal component analyses (statistical procedures for analyzing data sets) and pattern recognition techniques that could reduce prediction errors for forecasts of 24 h or more, the period most critical for issuance of coast warning. *See* REMOTE SENSING; UPPER-ATMOSPHERE DYNAMICS; WEATHER FORECASTING AND PREDICTION.

Greg Holland; Joanne Simpson; Robert Simpson

Bibliography. R. A. Anthes, *Tropical Cyclones: Their Evolution, Structure and Effects*, 1982; C. Neumann and J. M. Pelissier, An analysis of Atlantic tropical cyclone forecast errors, 1970–1979, *Mon. Weath. Rev.*, 109:1248–1266, 1981; C. Neumann and J. M. Pelissier, Models for the prediction of tropical cyclone motion over the North Atlantic: An operation evaluation, *Mon. Weath. Rev.*, 109:522–538, 1981; R. A. Pielke, *The Hurricane*, 1990; R. H. Simpson and H. Riehl, *The Hurricane and Its Impact*, 1981; H. E. Willoughby, J. Clos, and M. Shoreibah, Concentric eyewalls, secondary wind maxima, and the evolution of the hurricane vortex, *J. Atm. Sci.*, 39:395–411, 1982.

Huygens' principle

An assumption regarding the behavior of light waves, originally proposed by C. Huygens in the seventeenth century to explain the fact that light travels in straight lines and casts sharp shadows. Large-scale waves, such as sound waves or water waves, bend appreciably into the shadow. The special behavior of light may be explained by Huygens' principle, which states that "each point on a wavefront may be regarded as a source of secondary waves, and the position of the wavefront at a later time is determined by the envelope of these secondary waves at that time." Thus a wave WW originating at S is shown in **illus.** *a* at the instant it passes through an aperture. If a large number of circular secondary waves, originating at various points on WW, are drawn with the radius r representing the distance the wave would travel in time t, the envelope of these secondary waves is the heavily drawn circular arc $W'W'$. This represents the wave after t. If, as Huygens' principle requires, the disturbance is confined to the envelope, it will be 0 outside the limits indicated by points W'.

Careful observation shows that there is a small amount of light beyond these points, decreasing rapidly with distance into the geometrical shadow. This is called diffraction. *See* DIFFRACTION.

The Huygens-Fresnel principle, a modification of Huygens' original formulation, is capable of explaining diffraction. A. Fresnel in 1814 postulated that the amplitude of any secondary wave decreases in proportion to cos θ, when θ is the angle between the normal to the original wavefront and any point on the secondary wave (see illus. *b*, where the thickness of the arc indicates the amplitude). Fresnel then modified Huygens' requirement that the disturbance be confined to the envelope, by specifying that at any point the disturbance was the resultant of all

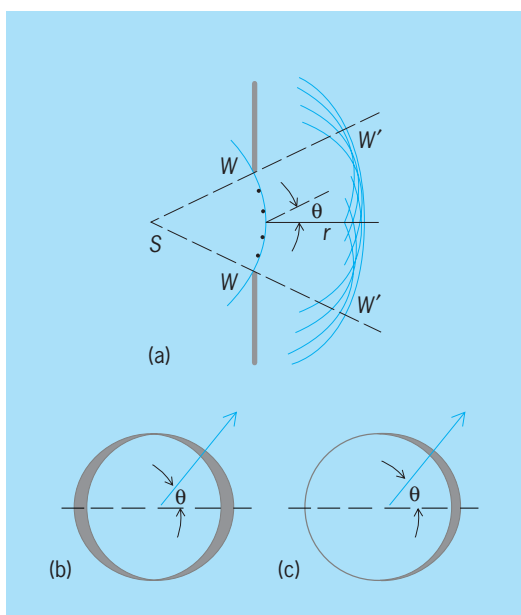

Huygens' principle. (*a*) The construction for a spherical wave. (*b*, *c*) Amplitude of the secondary wave according to Fresnel and Kirchhoff, respectively.

displacements due to secondary waves reaching that point. In this way Fresnel was able to explain the complicated diffraction patterns that are produced by sending light through small apertures. Subsequent theoretical investigations by G. Kirchhoff showed that the correct obliquity factor should be $1 + \cos\theta$ instead of $\cos\theta$ (illus. *c*). Approximations made by Fresnel had compensated for this. A discrepancy in the phase of the resultant wave of one-quarter period was also explained by Kirchhoff.

Francis A. Jenkins; William W. Watson

Bibliography. J. W. Goodman, *Introduction to Fourier Optics*, 3d ed., Roberts, 2005.

Hyades

A small cluster of stars that makes up the nearest well-defined open cluster (galactic cluster) to the Earth. With a total mass of about 300 suns and a population of 400–500 mostly low-mass stars, the Hyades is a typical example of the 2000 or so small star clusters in the Milky Way Galaxy. Most of its stars are located in a loose, roughly spherical system approximately 40 light-years (2.4×10^{14} mi or 3.8×10^{14} km) in diameter. It is located primarily in the constellation Taurus, the Bull, and forms the faint stellar background to the V-shaped asterism that defines the head of the Bull. *See* TAURUS.

Distance. For many years the Hyades was the backbone of the stellar distance scale in the Milky Way Galaxy and beyond. Because of its proximity, its motion through space can be detected from Earth by successive photographs taken years apart. Combining the motion in the sky of its individual stars with motions in the line of sight obtained with spectrographs, astronomers were able to determine a geometric distance to the Hyades. The technique is called the moving cluster method, and the Hyades has until recently been the only cluster close enough to use the method reliably. The galactic distance scale was established by comparing the properties of stars in other, more distant clusters with Hyades stars.

However, it was found in 1967 that certain unexpected errors had entered into the process. Evidence suggested that the distance to the Hyades (and, therefore, almost all other distances in the Milky Way Galaxy and beyond) had been underestimated by 20%. This conclusion was confirmed in 1996, when the European space telescope *Hipparcos* was able to make a more direct and more precise measure of the Hyades distance of 149 light-years (8.7×10^{14} mi or 1.4×10^{15} km). *See* ASTROMETRY; PARALLAX (ASTRONOMY).

Age. The age of the Hyades has been determined by directly comparing its stars, especially its brightest main-sequence stars (stable, long-lived stars), with detailed computer models of stars of different ages. This process must take into account the abundances of the various elements in Hyades stars which, spectroscopy indicates, are similar to the abundances in the Sun, but with all elements heavier that helium being slightly elevated. The age of the cluster is found to be approximately 6.5×10^8 years. Thus the Hyades were formed quite recently, after about 95% of the life of the Milky Way Galaxy had already occurred. *See* ASTRONOMICAL SPECTROSCOPY; STELLAR EVOLUTION.

Chemical composition. The age-measuring process must take into account the abundances of the various elements in Hyades stars. A 2006 study of a large number of individual stars using powerful new telescopes found that the stars of the Hyades are remarkably uniform in their chemical composition, all of them 45% higher than the Sun in their abundances of elements heavier than helium. *See* ASTRONOMICAL SPECTROSCOPY; TELESCOPE.

Stellar membership. The Hyades is made up mainly of normal main sequence stars, but it also contains many binary stars, a few cool giant stars, and many normal cool dwarf stars that are weak x-ray sources. The use of the *Hubble Space Telescope*, together with giant ground-based telescopes using adaptive optics, led to the detection of a population of a new kind of star, called brown dwarfs, in the Hyades. *See* ADAPTIVE OPTICS; BINARY STAR; BROWN DWARF; DWARF STAR; GIANT STAR; HERTZSPRUNG-RUSSELL DIAGRAM; HUBBLE SPACE TELESCOPE; MILKY WAY GALAXY; STAR; STAR CLUSTERS.

Paul Hodge

Bibliography. G. M. De Silva et al., Chemical homogeneity in the Hyades, *Astron J.*, 131:455–471, 2006; P. Hodge, How far are the Hyades?, *Sky Telesc.*, 75(2):138–140, February 1988.

Hybrid control

Control by means of systems that include both digital and analog devices working together to enhance the controller's functionality. As computers and their associated memory have become cheaper, faster, and easier to use, engineers have used them to expand the capabilities of control systems. Most controllers built today are, to some extent, hybrids because digital implementation is usually cheaper and more reliable than analog. The computer in the controller can then be used easily to provide additional capability, such as better response to internal failures. However, the potential benefits of control systems combining the symbol manipulation and decision-making capability of the computer with the precise tracking capability of analog controllers are much greater.

Automobile engine control unit. The current state of the art in hybrid control is well represented by a typical automobile engine control unit. Bosch's Motronic engine management system primarily controls ignition and fuel injection. The ignition map, basically a nonlinear function mapping engine speed and load into spark advance angle, is extremely complex in comparison to analog versions of the same controller. The complexity results from individually optimizing dozens of points on the map and then interpolating smoothly between them. This would not be feasible without the computer in the controller (see **illustration**). *See* AUTOMOTIVE ENGINE.

The rest of the Motronic engine management system includes several such complex, optimized

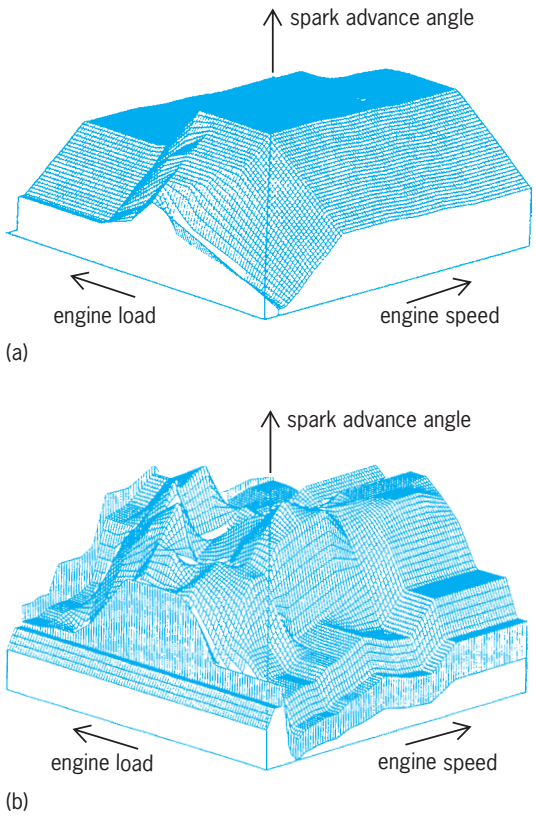

Ignition timing maps. (*a*) Mechanical advance system. (*b*) Electronically optimized system. (*After R. K. Jurgen, ed., Automotive Electronics Handbook, 2d ed., McGraw-Hill, 1999*)

functions. It also includes two other functions that depend on the computer. First, the controller is adaptive. The controller continually monitors certain controlled variables and changes the controller's parameters to maintain the overall system's performance despite changes in the engine's parameters. This also minimizes the need for adjustments during vehicle servicing. Second, the hybrid controller continuously monitors itself for component failures. When a failure is detected, it modifies its operation to protect other components that might be damaged by the failure, and adopts an emergency mode that allows the vehicle to still operate. The computer stores the diagnostic information and supplies it, on demand, to a repair person. Lastly, since the driver may be unaware of the problem, it provides a warning indication.

The engine management system has some decision-making capability. Another automotive example illustrates additional possibilities. The controller that deploys the airbags and seat belt pretensioners in a car is a decision-making feedback control system. The controller senses the vehicle's acceleration, either positive or negative. The analog signals from the sensors are then input to a digital system that must decide whether to deploy the seat belt pretensioners and the airbags and, if so, at what precise instants. The decision is, in principle, difficult. The vehicle is continually accelerated by the driver, by bumps in the road, by the wind, and occasionally by minor impacts. Deployment must not occur unless there is a real crash and, furthermore, it must occur at the right instant. Future systems are expected to have enhanced decision-making capability. They are expected to classify the occupant and adapt the airbag and seat belt pretensioner to the occupants and the crash severity. Current systems also self-test, a very common feature of hybrid control systems.

Incorporating humanlike decision-making into an automatic control system creates very exciting possibilities. For example, there is a great deal of current research on various types of autonomous vehicles. These include pilotless reconnaissance and combat aircraft, driverless earthmoving equipment, and various mobile autonomous robots. With the Global Positioning System (GPS) providing very accurate location sensing; the computer optimizing routing, adjusting speed, and choosing the best lane; and the analog controller maintaining speed and heading, driverless passenger vehicles are nearly feasible. The lack of a good obstacle sensor, cost, and people's willingness to trust their lives to a decision-making machine are all that stand in the way of a true "auto" mobile.

Design. The digital and analog parts of a hybrid control system have different requirements from those imposed on them in nonhybrid situations. Even though step inputs are common tests of control systems, it is generally undesirable for the inputs to a physical system to change suddenly by a large amount. For example, rapidly flooring the accelerator of a car will usually spin the wheels. Rapid changes in the parameters of an analog control system can have even worse effects and should be avoided.

Digital computers can respond almost instantly to a change in an input signal. Thus, they facilitate rapid changes. Hybrid control systems must be designed to prevent harmful jumps in signals and parameters. Alternatively, the controller can be designed to mitigate the effects of sudden changes. In either case the designer has a problem to solve.

A second design challenge results from the fact that digital computers normally act asynchronously. The exact instant at which the computer returns an answer does not matter. Analog systems operate in real time; they need a control signal at every instant. In a hybrid system the computer is constrained to respond no later than a rigid deadline imposed by the analog system. Because the computer is typically capable of many more operations per second than are needed for analog control, it is usually performing other operations, such as a self-test, in its spare time. The designer has to ensure that this does not interfere with the analog operation.

Ensuring that the controller works properly in all situations is an extremely important, and difficult, aspect of hybrid control system design. A software crash that freezes the computer is normally a nuisance; in a hybrid controller it can be a disaster. Testing and verification of the hybrid controller's operation is related to, but substantially more difficult than, the software verification problem.

Theory. Most of the theoretical work on hybrid control has been devoted to the problem of finding a

single, unified description of an arbitrary hybrid system. Without such a framework it is impossible to answer many basic questions about control. For example, is the system stable? This single mathematical model must be a reasonably accurate description of an interesting class of real hybrid systems; otherwise it has only academic interest. The difficulty results from the completely different mathematics associated with the two main parts of any hybrid system.

Analog systems act continuously on continuously applied inputs; they are "time-driven." The driver must keep her foot on the accelerator pedal of her car; the car responds to any change in the pedal angle. The usual mathematical models of analog systems are differential equations or difference equations in which time is an independent variable. Digital systems are more naturally viewed as "event-driven;" they respond to discrete instantaneous inputs. The computer does nothing until the operator strikes the return key; it then executes a string of instructions, stops, and waits for the next command. The mathematical description of discrete-event (event-driven and discrete-state) systems is more like a language—if this event occurs then that event results. In many such systems and their models, time is simply ignored. When time is included, it is just an attribute of an event and not the driving independent variable.

Many mathematical models for hybrid systems have been proposed. There does not seem to be agreement on one such model. The most useful model seems to depend on the specific application. Two examples are given below. They were chosen because they are substantially less abstract that most.

One research project has concentrated on automobile engine control. Motivation includes the importance of the automobile to society and the importance of hybrid controllers to the automobile. An additional motivation is that the four-stroke single-cylinder internal combustion engine can be viewed as a discrete-event system in its own right, even without the hybrid controller. The onset of each stroke is an event. Torque generation, which depends on the stroke, is then a discrete-event system. The power train and air dynamics are described by differential equations; equivalently, they are analog systems. Naturally, the group of researchers involved in this project formulates a hybrid model for the internal combustion engine. They also pose several optimal control problems associated with this engine. A key to the solution of these problems is that they are able to "relax" the original hybrid model of the engine to a purely analog description. Finally, they map the relaxed solution back to the full hybrid engine model to obtain a suboptimal solution to the original problem.

Several aspects of the project illustrate the value of good mathematical modeling. The system specifications can be given at a high level of abstraction so that they are independent of implementation decisions. The rigorous mathematical framework ensures formal correctness of the control algorithms.

Another research project on hybrid control focuses on motion control systems, primarily in robotics. Controlling a movement in such a system requires both discrete event and analog elements. The discrete-event part is obvious—the robot moves until it touches the wall. The analog part is less obvious—it is necessary to describe and control the analog compliance that governs the interaction with the wall. A hybrid model is therefore created for movement description. This model leads to a motion description language, a basis for robot programming that facilitates specification of both the motion and its analog control. Theory such as this that simplifies the programming of hybrid controllers would greatly facilitate their development.

There are many current research projects on various aspects of hybrid control. Two large projects are HYCON (Hybrid Control), a European research consortium, and CHESS (Center for Hybrid and Embedded Software Systems), a cooperative project involving universities in the United States. At a more applied level, in 2006 the U.S. Defense Advanced Research Projects Agency (DARPA) issued a grand challenge to develop an autonomous vehicle that would successfully travel through a 60-mi (96-km) urban course at an average speed of at least 10 mi/h (16 km/h) without hitting any of the numerous obstacles in the environment. The run in which the qualifying vehicles would be expected to meet this challenge was scheduled for November 2007. The critical problem is not the vehicle but the hybrid control system. Another challenge is the RoboCup: to develop a soccer team of fully autonomous humanoid robots that can win against the human world soccer champion team by 2050. *See* CONTROL SYSTEMS; DIGITAL CONTROL; ROBOTICS.

William S. Levine

Bibliography. R. J. Alur and G. Pappas (eds.), *Hybrid Systems: Computation and Control*, Springer-Verlag, Berlin, 2004; M. Egerstedt, Control of autonomous mobile robots, in D. Hristu-Varsakelis and W. S. Levine (eds.), *Handbook of Networked and Embedded Control Systems*, pp. 767–778, Birkhäuser Verlag, Basel, 2005; B. K. Mattes, Passenger safety and convenience, in R. K. Jurgen (ed.), *Automotive Electronics Handbook*, 2d ed., chap. 23, McGraw-Hill, 1999; Robert Bosch GmbH, *Automotive Handbook*, 6th ed., 2005.

Hybrid dysgenesis

A syndrome of abnormal traits that appears in the hybrids between certain strains of the fruit fly *Drosophila melanogaster*. The traits include partial sterility and greatly elevated rates of genetic mutations and chromosome rearrangements. Strains can be classified as P for paternally contributing or M for maternally contributing, so that only the hybrid sons and daughters of M females mated to P males show hybrid dysgenesis.

Cause. Hybrid dysgenesis is caused by the action of a family of transposable genetic elements, that is, segments of the genetic material (deoxyribonucleic acid, or DNA) with the special ability to move from one chromosomal site to another. Such elements range in size from a few hundred to a few

thousand nucleotide pairs and typically occur in 10–100 genetic locations scattered throughout the chromosomes. Altogether, transposable elements are thought to compose 10–15% of the entire genetic complement of *Drosophila melanogaster* and are probably also common in most animal and plant species.

The family of transposable elements that causes most cases of hybrid dysgenesis is called the P family, and the individual elements are called P factors because they occur only in the paternally contributing strains. A typical P strain might have 30 to 50 factors in widely scattered chromosomal locations. The cross of P males to M females activates the transposition mechanisms of these elements, resulting often in mutations and chromosome rearrangements due to the chromosome breakage involved in the transposition process. Transposition occurs preferentially in the cell lines destined to form the gametes (the germ line), which probably accounts for the partial sterility seen in the hybrids. *See* TRANSPOSONS.

Behavior of P factors. P factors can be either complete or defective. The complete copies are 2907 nucleotide pairs long, with the first 31 bases exactly matching the last 31 bases in reverse order. This reverse repeat structure undoubtedly contributes to their ability to transpose. The defective elements differ from the complete ones by the deletion of interior sequences; the end repeats are intact in both types of factor.

The genetic information in the complete elements codes for an enzyme product (known as transposase) that gives the elements their mobility. The defective elements cannot produce their own transposase; however they can still move about if there are complete P factors present in the cell to provide the transposase.

When there is a large number of P factors present in the chromosomes of certain individuals, the transposition process stops and the number of elements does not increase further. This situation is known as the P cytotype. A likely interpretation is that the P factors produce a second gene product that regulates their own transpositional activity. These parental regulator molecules would be packaged in the egg cells of the P strains but not in the sperm cells, thus explaining why only hybrids from P males and M females have active P factors.

P factors in biotechnology. The ability of P factors to transpose at high frequencies and to be regulated by the P cytotype has made them especially useful as tools for manipulating recombinant DNA molecules. For example, P factors commonly cause mutations by inserting into genes and thus inactivating them. When this happens, the DNA of the inactivated gene can then be identified through its association with the P factor. P factors can also be used as vehicles for transporting a gene of interest from the test tube back into the organism, where its expression can be studied. *See* GENETIC ENGINEERING.

Importance of P factors in evolution. The evolutionary significance of P factors and other transposable elements is not known. One possibility is that they act as DNA parasites whose only adaptation is to maintain their presence in the chromosomes. Two arguments that strengthen this interpretation are (1) that the only known effects of P factors on the organism is the hybrid dysgenesis syndrome, which is clearly detrimental; and (2) M strains lack P factors, yet still have normal viability and fertility. However, even in the role of parasites, P factors and other transposable elements might make a positive contribution to the evolution of their hosts by additional mutations and chromosome rearrangements produced as by-products of transposition. Some of these might be favored by natural selection. In addition, the partial reproductive incompatibility between P and M strains could, in some conditions, aid in the process of splitting one species into two. Although no such instance has been observed, it is postulated that if two subspecies are already partly isolated, by geographical barriers for example, one subspecies and not the other could acquire transposable elements such as P factors. In that case, the resulting dysgenesis in some of the hybrids between the two subspecies would yield natural selection that would favor further reproductive isolation and perhaps eventual splitting into separate species. *See* GENE; MUTATION; ORGANIC EVOLUTION.

William R. Engels

Bibliography. W. R. Engels, The P family of transposable elements in *Drosophila, Annu. Rev. Genet.*, 17:315–344, 1983; M. G. Kidwell, J. F. Kidwell, and J. A. Sved, Hybrid dysgenesis in *Drosophila melanogaster*: A syndrome of aberrant traits including mutation, sterility, and male recombination, *Genetics*, 36:813–83, 1977; K. O'Hare and G. M. Rubin, Structures of P transposable elements of *Drosophila melanogaster* and their sites of insertion and excision, *Cell*, 34:25–35, 1983.

Hydatellales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Commelinidae of the class Liliopsida (monocotyledons). The order consists of a single family with five species native to Australia, New Zealand, and Tasmania. The plants are small, submersed or partly submersed aquatic annuals with greatly simplified internal anatomy. The leaves are tufted at the base of the stem, and the inflorescence is a terminal head with two to several bracts, each subtending one to several reduced, unisexual flowers. These plants have sometimes been included within the Restionales, but the structural details of the ovary and seed set them apart. They are of no economic significance. *See* COMMELINIDAE; LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM; RESTIONALES. T. M. Barkley

Hydrate

A particular form of a solid compound which has water in the form of H_2O molecules associated with it. For example, anhydrous copper sulfate is a white solid with the formula $CuSO_4$. When crystallized

from water, a blue crystalline solid which contains water molecules as part of the crystals is formed. Analysis shows that the water is present in a definite amount, and the hydrate may be given the formula $CuSO_4 \cdot 5H_2O$. Four of the water molecules are attached to the copper ion in the manner of coordination complexes, and the fifth water molecule is related to the sulfate and presumably held by hydrogen bonding. *See* HYDROGEN BOND.

Water can also be present in definite proportions in the crystal without being associated directly with the anion or cation. The water occupies a definite place in the crystal lattice. Alums, with their 12 molecules of water, are examples of this. *See* ALUM.

Frank Wagner

Gas hydrates (gas clathrates) are crystalline compounds in which an isometric (cubic) ice (H_2O) lattice contains cages that incorporate small guest gas molecules. They are stable at moderate to high pressures and low temperatures, above and below the ice point. These ice lattices are stable only when the cages contain a gas molecule. The pressure and temperature constraints restrict them to oceanic continental margins in the uppermost few hundred meters of slope and rise sediments where water depths exceed 300–500 m, and to permafrost in polar regions. Under the ocean, the amount of gas hydrates is at least an order of magnitude higher than in permafrost.

Methane (CH_4) hydrate is the dominant natural gas hydrate on Earth, and is probably found on the planets Uranus and Neptune. One cubic meter of methane hydrate when dissociated can contain 165–180 m^3 of methane gas. The total amount of methane in gas hydrates is estimated to be very large; about 10^{19} g of methane carbon is stored in them, approximately twice that in fossil fuels. Geochemical studies of the methane, especially its carbon isotope ratios, indicate that the methane is largely biogenic in origin. But in some regions, for example, the Gulf of Mexico and Caspian Sea, its origin is thermogenic.

Gas hydrate is distributed inhomogeneously in ocean sediment (see **illus.**). It cements coarser sediments, is disseminated in silty and clay-rich sediments, and is abundant in faults or lithologic boundaries.

Methane hydrate core sample taken from the Costa Rica margin slope sediments about 200 m below the sea floor.

Recent interest in natural gas hydrates, most of which are methane hydrates, has resulted from the recognition that global warming may destabilize the enormous quantities of methane hydrate in shallow marine slope sediments and permafrost. The environmental impact of releasing large quantities of methane into the ocean and atmosphere could have important consequences. In the ocean, intense microbial oxidation of methane would reduce the amount released into the atmosphere, but would result in extensive oxygen consumption and somewhat reduce the capacity of the ocean to absorb atmospheric carbon dioxide.

Gas hydrates also affect sediment physical properties. During deglaciation or global warming, the gas hydrate may decompose at the sediment's base layer. The released gas will be trapped in the sediment, causing instability and possibly triggering giant landslides that may release methane to the ocean and atmosphere.

Methane gas is an important contributor to the atmospheric radiation balance. Any flux of methane into the atmosphere beyond the present annual rate of ~0.8% per year would enhance global warming. New evidence has been reported concerning the possible role of methane hydrate decomposition in rapid global warming, in the Santa Barbara Basin (southern California) over the past 50,000–60,000 years and in the deep ocean about 55.6 million years ago. Both episodes were accompanied by thermal maxima—times when the climate on Earth was unusually warm.

The fossil fuel resource potential of the enormous quantities of marine methane hydrates is being evaluated. *See* MARINE SEDIMENTS; METHANE; PERMAFROST.

Miriam Kastner

Bibliography. B. E. Conway, *Ionic Hydration in Chemistry and Biophysics*, 1981; G. R. Dickens, M. M. Castillo, and J. G. Walker, A blast of gas in the latest Paleocene: Simulating first-order effects of massive dissociation of oceanic methane hydrate, *Geology*, 25:259–262, 1997; F. Franks (ed.), *Water: A Comprehensive Treatise*, vol. 2: *Water in Crystalline Hydrates*, 1973; V. Gornitz and I. Fung, Potential distribution of methane hydrates in the world's oceans, *Global Biogeochem. Cycles*, 8:335–347, 1999; H. Kleeberg (ed.), *Interactions of Water in Ionic and Nonionic Hydrates*, 1987; K. A. Kvenvolden, Methane hydrate—a major reservoir of carbon in the shall geosphere?, *Chem. Geol.*, 71:41–51, 1988; E. G. Nisbet and D. J. W. Piper, Giant submarine landslides, *Nature*, 392:329–330, 1998; D. D. Sloan, Jr., *Clathrate Hydrates of Natural Gases*, 2d ed., Marcel Dekker, New York, 1998.

Hydration

The incorporation of molecular water into a complex with the molecules or units of another species. The complex may be held together by relatively weak forces or may exist as a definite compound. Many salts form solid hydrates when exposed to water vapor under certain conditions of temperature

and pressure. Copper sulfate, for example, forms a monohydrate ($CuSO_4 \cdot H_2O$) when exposed at 25°C (77°F) to water vapor at a pressure of 0.8 mm of mercury (1100 pascals). At higher pressures other hydrates are formed. Water is lost from these compounds when they are heated or when the water vapor pressure falls below a minimum value. Solids forming hydrates at low pressures are used as drying agents. *See* DELIQUESCENCE; DESICCANT; EFFLORESCENCE; HYDRATE; SOLUTION; SOLVATION.

Francis J. Johnston

Hydraulic accumulator

A pressure vessel which operates as a fluid source device or shock absorber. It is used to store fluid under pressure or to absorb excessive pressure increases. The hydraulic accumulator is an energy-efficient component, which allows the use of a smaller pump to achieve the same end results in terms of cylinder rod actuation speeds. In certain circuit designs, the accumulator will permit a pump motor to be completely shut down for an extended period of time while the accumulator supplies the necessary fluid to the circuit.

The operation of the hydraulic accumulator is induced by a pressurized gas (usually nitrogen), a spring, or a weighted plunger. The accumulator supplies fluid for actuator movement or to replace fluid lost by leakage. *See* SHOCK ABSORBER.

The gas-charged accumulator and the spring-type accumulator discharge their fluid into the system at pressures which are decreasing as the gas or spring expands. The weighted accumulator allows stored fluid to be discharged into the system at a constant pressure for the entirety of its downward stroke.

The gas-charged accumulator can be bladder type, diaphragm type, or piston type (see **illus.**) The bladder-type accumulator consists of a one-piece outer steel shell whose upper portion contains a bladder of elastomeric material compatible with the hydraulic fluid and charged with nitrogen gas to a pressure of 50–75% of the system pressure. The bottom of the accumulator housing is ported to allow free entry and exit of the fluid as required. When the system needs fluid from the accumulator, the nitrogen pressure expands the bladder and thus forces the fluid out into the system. As the bladder expands, its pressure is reduced, as is the system pressure.

The diaphragm-type accumulator consists of two hemispheres making up the outer shell, with the split being horizontal, and a compatible convoluted elastomeric diaphragm is bolted between the hemispheres. The upper half of the accumulator is charged with nitrogen gas, and the unit functioning is similar to that of the bag-type accumulator.

The piston-type accumulator is similar to a cylinder, except there is no piston rod. The position of the piston is determined by the input of fluid into the accumulator, or the withdrawal of fluid from it. The piston seal prevents gas from leaking into the fluid.

The spring-type accumulator is a piston-type accumulator that uses a powerful spring instead of a gas charge. *See* SPRING (MACHINES).

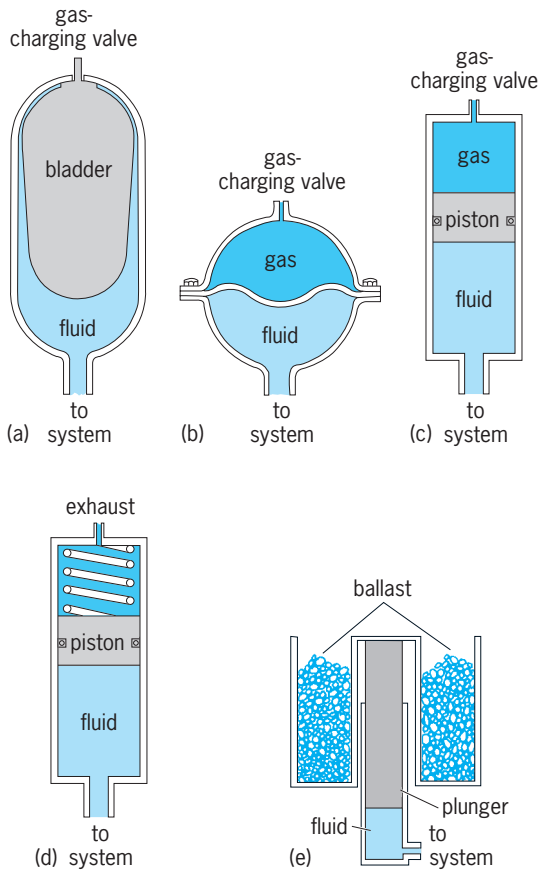

Diagrams showing the five types of hydraulic accumulators. (*a*) Bladder. (*b*) Diaphragm. (*c*) Piston. (*d*) Spring. (*e*) Weighted.

The weighted-type accumulator has a weight pan attached to the plunger's uppermost point. The pan carries an evenly distributed load of scrap metal whose weight is sufficient to induce the desired pressure.

The charging of the accumulator with fluid is accomplished during any period of the system cycle in which the total output of the pump is not needed for actuator function. *See* CONTROL SYSTEMS.

James E. Anders, Sr.

Bibliography. T. C. Frankenfield, *Using Industrial Hydraulics,* 1985; Parker-Hannifin Corp., *Design Engineers Handbook*, 1973; W. W. Vockroth, *Industrial Hydraulics*, 1994.

Hydraulic actuator

A cylinder or fluid motor that converts hydraulic power into useful mechanical work. The mechanical motion produced may be linear, rotary, or oscillatory. Operating pressures range from a few pounds per square inch gage (psig or lb/in.^{2} gage) to several thousand, but usually are about 500–5000 psig (3.4–34 megapascals). Sizes vary from 0.2-in.^{2} (1.3-cm^2) linear actuators capable of a few pounds of force to extremely large linear or rotary actuators capable of exerting hundreds of tons of force. Operation

exhibits high force capability, high power per unit weight and volume, good mechanical stiffness, and high dynamic response. These features lead to wide use in precision control systems and in heavy-duty machine tool, mobile, marine, and aerospace applications. *See* CONTROL SYSTEMS.

Cylinder actuators. To provide a fixed length of straight-line motion, linear cylinder actuators usually consist of a tight-fitting piston moving in a closed cylinder. The piston is attached to a rod that extends from one end of the cylinder to provide the mechanical output. The double-acting cylinder (**Fig. 1**) has a port at each end of the cylinder to admit or return hydraulic fluid. A four-way directional valve functions to connect one cylinder port to the hydraulic supply and the other to the return, depending on the desired direction of the power stroke.

Some examples of the many possible types of linear hydraulic actuators are shown in **Fig. 2**. Single-acting types have a power stroke in one direction only, with the return stroke accomplished by some external means or effected by a spring. Telescopic cylinders are used to accomplish extremely long single-acting motions. A double-ended rod is employed when it is necessary to have equal forces and velocities in each direction of travel. Tandem cylinder actuators consist of two or more pistons and rods fastened together so that they operate as a

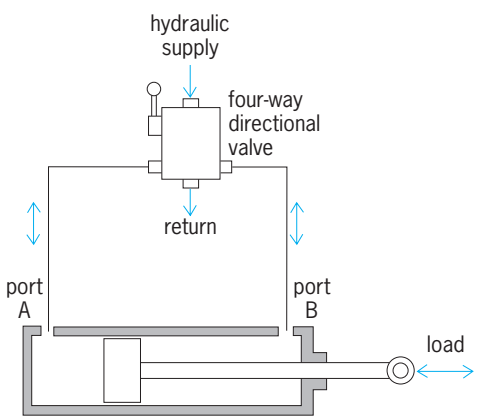

Fig. 1. Function of a hydraulic double-acting cylinder.

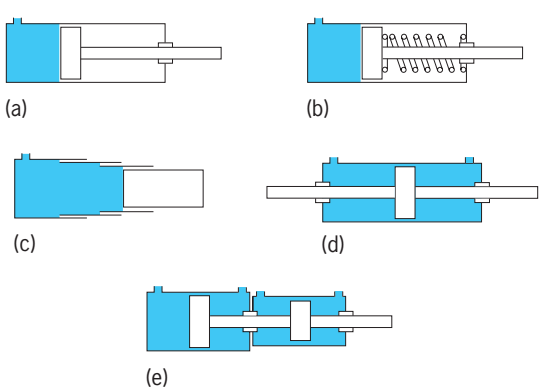

Fig. 2. Types of hydraulic cylinder (linear) actuators. (*a*) Single-acting, external return. (*b*) Single-acting, spring return. (*c*) Telescopic. (*d*) Double-acting double-ended rod. (*e*) Tandem cylinders.

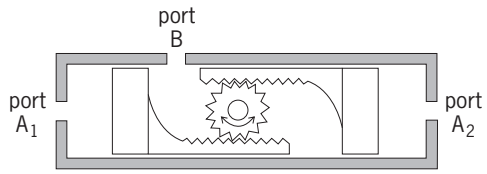

Fig. 3. Piston-rack type rotary actuator.

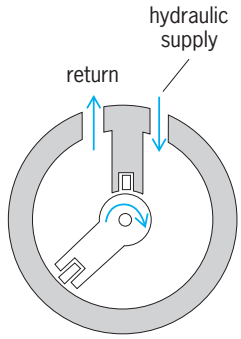

Fig. 4. Single-vane rotary actuator.

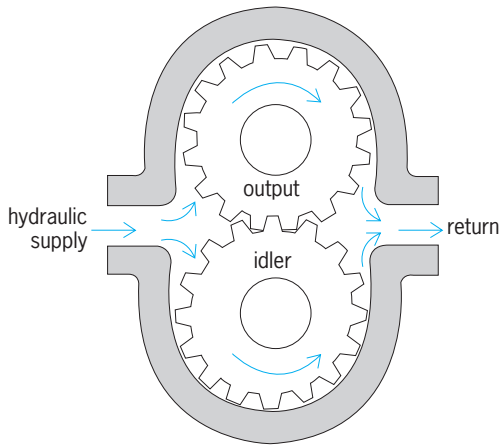

Fig. 5. Gear motor rotary actuator.

unit. In this actuator a combination of low-force and high-speed piston travel can be obtained, followed by high-force and low-speed piston travel.

Limited-rotation actuators. For lifting, lowering, opening, closing, indexing, and transferring movements, limited-rotation actuators produce limited reciprocating rotary force and motion. Rotary actuators are compact and efficient, and produce high instantaneous torque in either direction.

In the piston-rack type of rotary actuator (**Fig. 3**), the pinion gear is attached to the load. Hydraulic fluid is applied to either the two end chambers or the central chamber to cause the two pistons to retract or extend simultaneously so that the racks rotate the pin ion gear. The single-vane type of actuator (**Fig. 4**) has a fixed stationary barrier and a rotating vane that forms two variable volume chambers. Hydraulic fluid is ported to one chamber and returned from the other to effect output shaft rotation of about 280° in either direction. A double-vane actuator has two stationary barriers and two rotating vanes, which limit rotation to about 100°. Its advantages are balanced radial loads on the output shaft

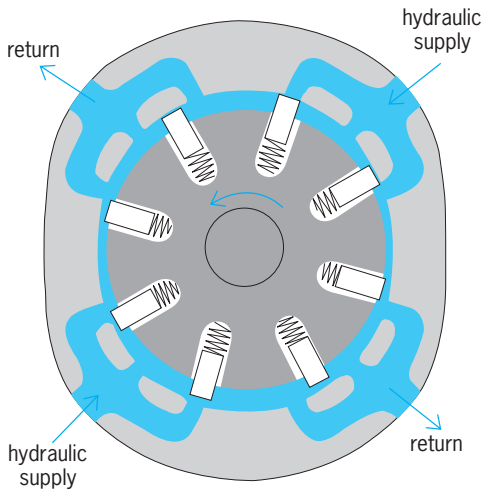

Fig. 6. Vane motor rotary actuator.

and a doubling of the torque output over that of a single-vane unit of comparable dimensions.

Rotary motor actuators. Coupled directly to a rotating load, rotary motor actuators provide excellent control for acceleration, operating speed, deceleration, smooth reversals, and positioning. They allow flexibility in design and eliminate much of the bulk and weight of mechanical and electrical power transmissions.

Motor actuators are generally reversible and are of the gear or vane type. The gear motor actuator (**Fig. 5**) has one gear connected to the output shaft. Supply fluid enters and flows around the chamber as shown, forcing the gears to rotate. The vane motor actuator (**Fig. 6**) consists of a rotor with several spring-loaded sliding vanes in an elliptical chamber. Hydraulic fluid enters the chamber and forces the vanes before it as it moves to the outlets.

Charles Mangion

Bibliography. P. Dransfield, *Hydraulic Control Systems: Design and Analysis of the Dynamics*, 1981; D. A. Pease, *Basic Fluid Power*, 2d ed., 1986; J. Prokes, *Hydraulic Mechanisms in Automation*, 1977; H. L. Stewart and J. M. Storer, *Fluid Power*, 3d ed., 1980; J. Stringer, *Hydraulic Systems Analysis*, 1976.

Hydraulic jump

An abrupt increase of depth in a free-surface liquid flow. A hydraulic jump is characterized by rapid flow and small depths on the upstream side, and by larger depths and smaller velocities on the downstream side. A jump can form only when the upstream flow is supercritical, that is, when the fluid velocity is greater than the propagation velocity c of a small, shallow-water gravity wave ($c = \sqrt{gh}$, where g is the acceleration of gravity and h is the depth). A considerable amount of energy is dissipated in the conversion from supercritical to subcritical flow. *See* OPEN CHANNEL; SURFACE WAVES.

Donald R. F. Harleman

Hydraulic press

A combination of a large and a small cylinder connected by a pipe and filled with a fluid so that the pressure created in the fluid by a small force acting on the piston in the small cylinder will result in a large force on the large piston. The operation depends upon Pascal's principle, which states that when a liquid is at rest the addition of a pressure (force per unit area) at one point results in an identical increase in pressure at all points. Therefore, in **Fig. 1**, the pressure due to the application of force F_1 is shown by Eq. (1), and the equilibrium force F_2 is shown by Eq. (2), where A_1 and A_2 are the areas

$$p = \frac{F_1}{A_1} \tag{1}$$

$$F_2 = pA_2 = F_1 \frac{A_2}{A_1} \tag{2}$$

of pistons 1 and 2, respectively. The mechanical advantage is shown by Eq. (3).

$$\text{MA} = \frac{F_2}{F_1} = \frac{A_2}{A_1} \tag{3}$$

The principle of the hydraulic press is used in lift jacks, earth-moving machines, and metal-forming presses (**Fig. 2**). A comparatively small supply pump creates pressure in the hydraulic fluid. The fluid then acts on a substantially larger piston to produce the action force. In this way forces greater than 15,000 tons (14,000 metric tons) are developed over the entire stroke of hydraulic presses. Heavy objects are accurately weighed on hydraulic scales in which

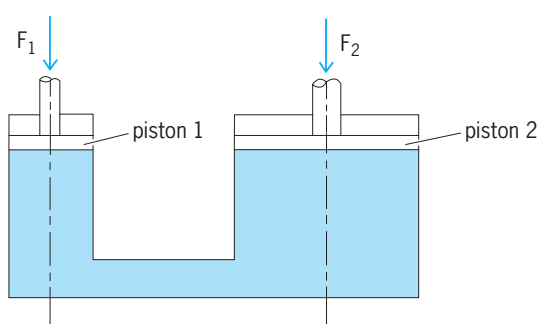

Fig. 1. Principle of hydraulic press.

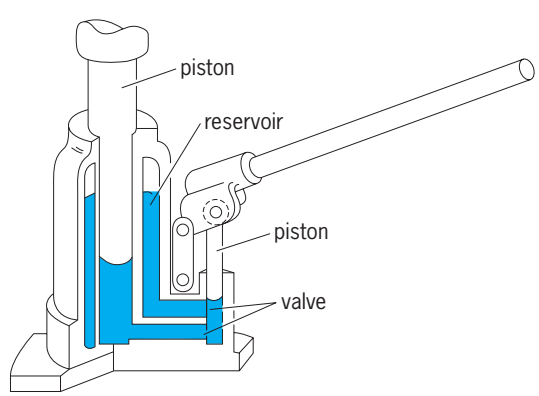

Fig. 2. Hydraulic jack.

precision ground pistons introduce negligible friction. *See* MECHANICAL ADVANTAGE; SIMPLE MACHINE.

Richard M. Phelan

Bibliography. M. Manohar and P. Krishnamachar, *Fluid Mechanics*, vol. 2: *Hydraulic Machinery and Advanced Hydraulics*, 1983; J. P. Pippenger and T. G. Hicks, *Industrial Hydraulics*, 3d ed., 1979.

Hydraulic turbine

A machine which converts the energy of an elevated water supply into mechanical energy of a rotating shaft. Most old-style waterwheels utilized the weight effect of the water directly, but all modern hydraulic turbines are a form of fluid dynamic machinery of the jet and vane type operating on the impulse or reaction principle and thus involving the conversion of pressure energy to kinetic energy. The shaft drives an electric generator, and speed must be of an acceptable synchronous value. *See* GENERATOR.

The impulse or Pelton unit has all available energy converted to the kinetic form in a few stationary nozzles and subsequent absorption by reversing buckets mounted on the rim of a wheel (**Fig. 1**). Reaction units of the Francis or the Kaplan types run full of water, submerged, with a draft tube and a continuous column of water from head race to tail race (**Figs. 2** and **3**). There is some fluid acceleration in a continuous ring of stationary nozzles with full peripheral admission to the moving nozzles of the runner in which there is further acceleration. The draft tube produces a negative pressure in the runner with the propeller or Kaplan units acting as suction runners; the Francis inward-flow units act as pressure runners. Mixed-flow units give intermediate degrees of rotor pressure drop and fluid acceleration. *See* IMPULSE TURBINE.

For many years reaction turbines have generally used vertical shafts for better accommodation of the draft tube whereas Pelton units have favored the horizontal shaft since they cannot use a draft tube. Vertical-shaft Pelton units have found increasing acceptance in large sizes because of multiple jets (for example, 4–6) on a single wheel; these provide reduced runner windage and friction losses and, consequently, higher efficiency. Axial-flow (**Fig. 4**) and diagonal-flow reaction turbines offer improved hydraulic performance and economic powerhouse structures for large-capacity low-head units. Kaplan units employ adjustable propeller blades as well as adjustable stationary nozzles in the gate ring for higher sustained efficiency. Pelton units are preferred for high-head service (1000± ft or 300± m), Francis runners for medium heads (200± ft or

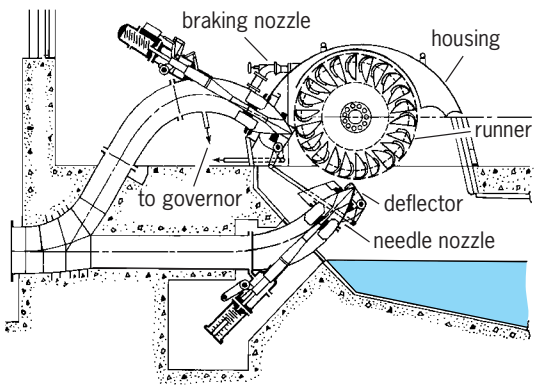

Fig. 1. Cross section of an impulse (Pelton) type of hydraulic-turbine installation.

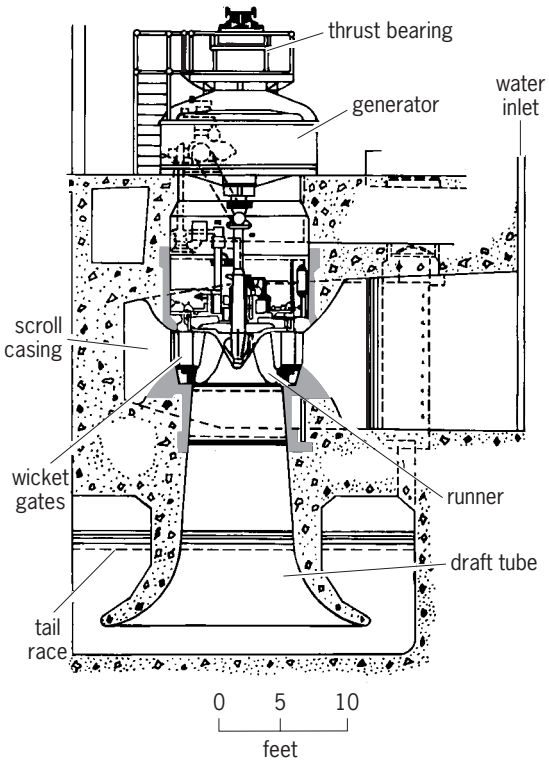

Fig. 2. Cross section of a reaction (Francis) type of hydraulic-turbine installation. 1 ft = 0.3 m.

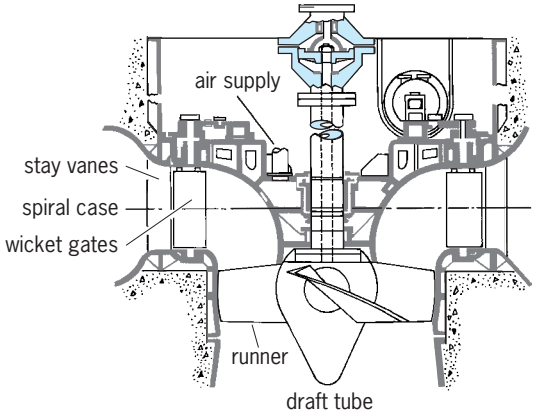

Fig. 3. Cross section of a propeller (Kaplan) type of hydraulic turbine installation.

60± m), and propeller or Kaplan units for low heads (50± ft or 15± m).

Hydraulic-turbine performance is rigorously defined by characteristic curves, such as the efficiency characteristic. The proper selection of unit type and size is a technical and economic problem. The data of **Fig. 5** are significant because they show synoptically the relationship between unit type and site head as the result of accumulated experience on some satisfactory turbine installations. Specific speed N_s is a criterion or coefficient which is uniquely applicable to a given turbine type and relates head, power, and speed, which are the basic performance data in the selection of any hydraulic turbine. Specific speed is defined in the equation below, where rpm is revolutions per

$$N_s = \frac{\text{rpm} \times (\text{shp})^{0.5}}{(\text{head})^{1.25}}$$

minute, shp is shaft horsepower (1 hp = 0.7 kW), and head is head on unit in feet (1 ft = 0.3 m). Specific speed is usually identified for a unit at the point of maximum efficiency. Cavitation must also be carefully scrutinized in any practical selection.

The draft tube (Fig. 2) is a closed conduit which (1) permits the runner to be set safely above tail water level, yet to utilize the full head on the site, and (2) is limited by the atmospheric water column to a height substantially less than 30 ft (9 m), and when made flaring in cross section will serve to recover velocity head and to utilize the full site head.

Efficiency of hydraulic turbine installations is always high, more than 85% after all allowances for hydraulic, shock, bearing, friction, generator, and mechanical losses. Material selection is not only a problem of machine design and stress loading from running speeds and hydraulic surges, but is also a matter of fabrication, maintenance, and resistance to erosion, corrosion, and cavitation pitting.

Governing problems are severe, primarily because of the large masses of water involved, their positive and negative acceleration without interruption of the fluid column continuity, and the consequent shock and water-hammer hazards. *See* PRIME MOVER.

Pumped-storage hydro plants have employed various types of equipment to pump water to an elevated storage reservoir during off-peak periods and to generate power during on-peak periods where the water flows from the elevated reservoir through hydraulic turbines. Although separate, single-purpose, motor-pump and turbine-generator sets give the best hydraulic performance, the economic burden of investment has led to the development of reversible pump-turbine units. Components of the conventional turbine are retained, but the modified pump runner gives optimum performance when operating as a turbine. Compromises in hydraulic performance, with some sacrifice in efficiency, are more than offset by the investment savings with the dual-purpose machines. *See* TURBINE; WATERPOWER.

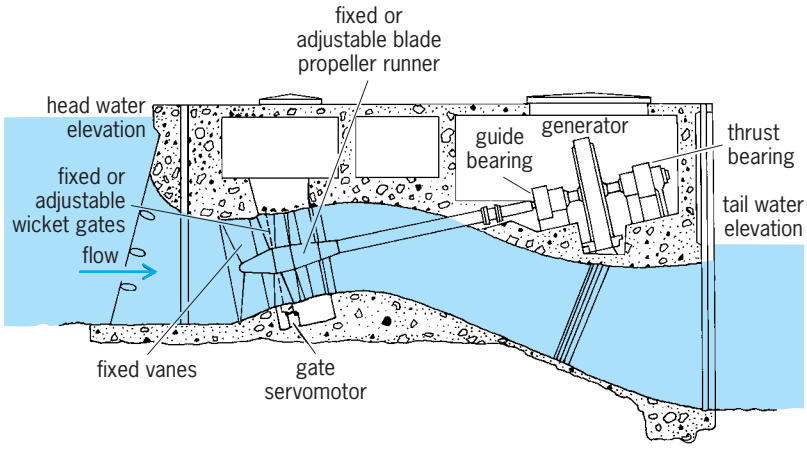

Fig. 4. Axial-flow tube-type hydraulic-turbine installation.

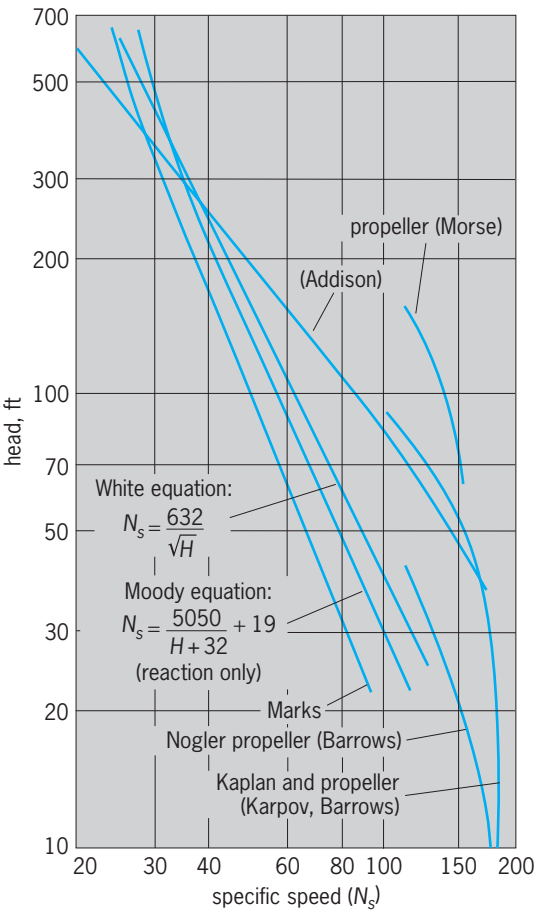

Fig. 5. Hydraulic-turbine experience curves, showing specific speed versus head. 1 ft = 0.3 m.

Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Standard Handbook for Mechanical Engineers*, 10th ed., 1996; D. G. Fink and H. W. Beatty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999.

Hydraulic valve lifter

A device that eliminates the need for mechanical clearance in the valve train of internal combustion engines. Clearance is normally required to prevent the valve's being held open and destroyed as the valve train undergoes thermal expansion. However, clearance requires frequent adjustment and is responsible for much operating noise. The hydraulic lifter is a telescoping compression strut in the linkage between cam and valve, consisting of a piston and cylinder (see **illus.**). When no opening load

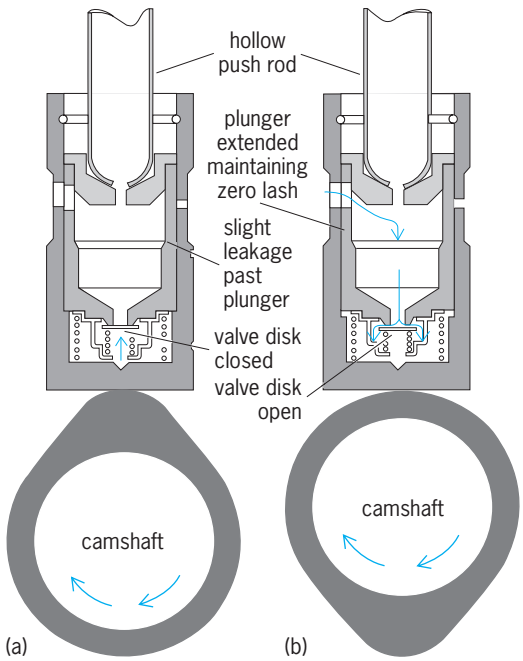

Positions of the hydraulic valve lifter, with engine valve (*a*) open and (*b*) closed. (*After W. H. Crouse, Automotive Mechanics, 5th ed., McGraw-Hill, 1965*)

exists, a weak spring moves the piston, extending the strut and eliminating any clearance. This action sucks oil into the cylinder past a check valve. The trapped oil transmits the valve-opening forces with little deflection. A slight leakage of oil during lift shortens the strut, assuring valve closure. The leakage oil is replaced as the spring again extends the strut at no load. *See* VALVE TRAIN. Augustus R. Rogowski

Hydraulics

The physical description of fluids in motion is based upon the conservation laws of mass, momentum, and energy. The mathematical statement of these laws may be approached from a highly theoretical point of view involving advanced mathematics, and the results obtained from this approach can be very. However, for practical engineering work these physical principles must be supplemented with empiricism and this semi-empirical approach is generally referred to as hydraulics. *See* CONSERVATION OF ENERGY; CONSERVATION OF MASS; CONSERVATION OF MOMENTUM.

Conservation laws. The mathematical form of the conservation laws may be obtained by considering the flow through an arbitrary closed surface drawn in the fluid, which may be either moving or fixed. This surface is called the control surface, and the volume inside it is called the control volume. If the control volume is of finite size (large), the equations governing the flow are obtained in integral form. If the control volume is of infinitesimal size (very small), the equations in differential form are obtained.

Equation of continuity. The continuity equation is the mathematical form of the law of conservation of mass. For a control surface, this law states that the net mass rate of flow out of the control volume through the control surface (mass flow rate out minus mass flow rate in) is equal to the rate at which the mass inside the control volume is decreasing. *See* EQUATION OF CONTINUITY.

For a one-dimensional steady flow (varying over space but not over time) through a fixed control volume with one inlet and one outlet, the equation of continuity becomes Eq. (1), where subscript 1 refers

$$\rho_1 A_1 V_1 = \rho_2 A_2 V_2 \tag{1}$$

to conditions at the entrance and subscript 2 to conditions at the outlet, A_1 is the area through which fluid is entering the control volume, A_2 is the area through which it is leaving, and ρ and V are the fluid density and velocity.

Momentum equation. The momentum equation is obtained by applying Newton's second law to the fluid passing through the control surface. In this case, the second law states that the sum of all the forces acting on the fluid inside the control volume must be equal to the time rate of change of the momentum of the fluid. The forces acting on the fluid include body forces such as gravity, forces that arise from pressure gradients, and viscous stresses. *See* MOMENTUM; NAVIER-STOKES EQUATION; NEWTON'S LAWS OF MOTION; VISCOSITY.

For a steady one-dimensional flow of an incompressible fluid, the momentum equation in integral form is Eq. (2), in which $\Sigma \boldsymbol{F}$ is the resultant force

$$\Sigma F = \rho Q(V_2 - V_1) \tag{2}$$

vector acting on the fluid inside the control volume, ρ is the fluid density, Q is the volume rate of flow, and V_1 and V_2 are the vector velocities of the fluid entering and leaving the control volume.

Energy equation. The energy conservation principle states that the rate at which work is done to the fluid inside the control volume is equal to the rate of change of the energy of the fluid. The energy consists of kinetic and potential forms. *See* ENERGY; THERMODYNAMIC PRINCIPLES.

For steady flow of an incompressible fluid, the integral form of the energy equation is Eq. (3), where

$$\frac{p_1}{\rho g} + \frac{V_1^2}{2g} + z_1 = \frac{p_2}{\rho g} + \frac{V_2^2}{2g} + z_2 + h_s + h_f \tag{3}$$

ρ is the fluid density, g is the acceleration of gravity, h_f is the energy-loss term due to internal fluid

friction, the subscripts 1 and 2 refer to the entrance and exit ports of the control volume, p and V are the fluid pressure and velocity, and z is the vertical coordinate. The term h_s is the shaft work of a pump or turbine per unit weight leaving the control volume. This equation is frequently referred to as the modified Bernoulli equation. *See* BERNOULLI'S THEOREM; FLUID-FLOW PRINCIPLES; FLUID MECHANICS; GAS DYNAMICS.

Applications. Applications of hydrodynamics include the study of closed-conduit and open-channel flow, and the calculation of forces on submerged bodies.

Closed-conduit flow. Flow in closed conduits, or pipes, has been extensively studied both experimentally and theoretically. If the pipe Reynolds number, given

$$\mathrm{Re}_D = \frac{VD\rho}{\mu} \tag{4}$$

by Eq. (4), where V is the average velocity and D is the pipe diameter, is less than about 2000, the flow in the pipe is laminar. In this case, the solution to the continuity, momentum, and energy equations is readily obtained, particularly in the case of steady flows. If Re_D is greater than about 4000, the flow in the pipe is turbulent, and the solution to the continuity, momentum, and energy equations can be obtained only by employing empirical correlations and other approximate modeling tools. The Re_D region between 2000 and 4000 is the transition region in which the flow is intermittently laminar and turbulent. *See* LAMINAR FLOW; REYNOLDS NUMBER; TURBULENT FLOW.

The loss term in Eq. (3) is given by Eq. (5), where

$$h_f = f\frac{L}{D}\frac{V^2}{2g} \tag{5}$$

L is the length of pipe between stations 1 and 2, and f is the friction factor. For laminar flow, the friction factor can be calculated exactly, and is given by Eq. (6). For turbulent flow, f is a function of both Re_D and

$$f_{\mathrm{lam}} = \frac{64}{\mathrm{Re}_D} \tag{6}$$

the relative roughness of the pipe wall, and values of f are obtained from the Colebrook formula, Eq. (7), or various other semiempirical correlations. Here,

$$\frac{1}{f^{1/2}} = -2.0\log\left(\frac{\epsilon/D}{3.7} + \frac{2.51}{\mathrm{Re}_D f^{1/2}}\right) \tag{7}$$

ϵ is the absolute internal surface roughness, whose values have been compiled for various materials. If the pipe has valves and other pipe fittings, the loss due to each of these is given by Eq. (8),where K is

$$h_f = K\frac{V^2}{2g} \tag{8}$$

the loss coefficient for the particular fitting. Values of K are determined experimentally and have been compiled. The force exerted by the fluid on bends, elbows, and reducers is calculated by using the momentum equation. *See* PIPE FLOW.

Open-channel flow. Confined flows that have a liquid surface exposed to the atmosphere (a free surface) are called open-channel flows. Flows in rivers, canals, partially full pipes, and irrigation ditches are examples. The difficulty with these flows is that the shape of the free surface is one of the unknowns to be calculated.

A special case that is sometimes a good approximation is steady, uniform flow. For this type of flow, the free surface of the liquid is parallel to the channel bottom, and the channel discharge is given by Eq. (9), where Q is the volume rate of flow, A is the

$$Q = CA(R_b S_0)^{1/2} \tag{9}$$

cross-sectional area of the flow, C is the Chézy coefficient, R_b is the hydraulic radius (sometimes called the hydraulic mean depth), and S_0 is the slope of the channel bottom. The hydraulic radius R_b is the ratio of the cross-sectional area of the flow to the wetted perimeter, which is the perimeter of the channel cross section in contact with the fluid. The Chézy coefficient is easily obtained from Manning's formula (10), in which n is the boundary roughness factor,

$$C = \frac{1.49}{n}[R_b(\mathrm{ft})]^{1/6} = \frac{1.0}{n}[R_b(\mathrm{m})]^{1/6} \tag{10}$$

whose values have been compiled for various materials.

In most open-channel flows, however, the bottom slope and the water depth change with position, and the free surface is not parallel to the channel bottom. If the slopes are small and the changes are not too sudden, the flow is called a gradually varied flow. An energy balance between two sections of the channel yields a differential equation for the rate of change of the water depth with respect to the distance along the channel. The solution of this equation, which must be accomplished by using one of many available numerical techniques, gives the shape of the water surface.

Flow over spillways and weirs and flow through a hydraulic jump are examples of rapidly varying flows. In these cases, changes of water depth with distance along the channel are large. Here, because of large accelerations, the pressure distribution with depth may not be hydrostatic as it is in the cases of gradually varied and uniform flows. Solutions for rapidly varying flows are accomplished by using approximation techniques. *See* HYDRAULIC JUMP; OPEN CHANNEL.

Forces on submerged bodies. The force exerted by a fluid flowing past a submerged body is in principle calculated by integrating the pressure distribution over the surface of the body. This force is resolved into two components, the lift and the drag. The drag force is the component parallel to the velocity U of the undisturbed stream (flow far away from the body), and the lift force is the component perpendicular to the undisturbed stream. These force components are related to the undisturbed stream

velocity by Eqs. (11), where C_L and $C_D$are the lift and

$$\text{Lift} = C_L A \frac{\rho U^2}{2}$$
$$\text{Drag} = C_D A \frac{\rho U^2}{2} \qquad (11)$$

drag coefficients, respectively. the values of C_d and C_l have been measured and compiled for many shapes ranging from spheres to airfoil sections. For nonlifting bodies, the area A in the drag formula is the projected area of the body on a plane normal to the flow. For lifing bodies A is the planform area. C_L and C_D depend upon the Reynolds number, surface roughness, and the orientation of the object to the stream. *See* BOUNDARY-LAYER FLOW. Warren M. Hagist

Bibliography. D. J. Acheson, *Elementary Fluid Dynamics*, 1990; V. T. Chow, *Open Channel Hydraulics*, 1959; L. M. Milne-Thomson, *Theoretical Hydrodynamics*, 5th ed., 1996; H. Schlichting, *Boundary Layer Theory*, 8th ed., 2000; F. M. White, *Fluid Mechanics*, 4th ed., 1998; F. M. White, *Viscous Fluid Flow*, 2d ed., 1991; E. B. Wylie, V. L. Streeter, and L. Suo, *Fluid Transients in Systems*, 1993.

Hydrazine

A colorless liquid, H_2NNH_2 (boiling point 114°C or 237°F), with a musty, ammonialike odor. Physically hydrazine is similar to water, but chemically it is reducing, decomposable, basic, and bifunctional. Its derivatives range from simple salts to ring compounds, polymers, and coordination complexes. Major uses of hydrazine include such diverse applications as rocket fuels (since combustion of hydrazine is highly exothermic), corrosion inhibition in boilers, and syntheses of biologically active materials.

Hydrazine is manufactured by two routes: the reaction of chloramine with ammonia and the reaction of sodium hypochlorite with urea. Both processes require the presence of glue or gelatin to inhibit catalytic decomposition of the product by unreacted oxidants. Because hydrazine forms an azeotrope containing 31% water (boiling point 121°C or 250°F), anhydrous hydrazine is isolated from aqueous process streams by extractive distillation with aniline.

Added to feedwater, hydrazine reduces rust in boilers to a hard film of magnetic iron oxide and reduces oxygen to water on catalytic metal surfaces. Metal ions such as Cu^{2+} and Ni^{2+} are reduced to free metals, and organic nitro compounds are reduced to amines by hydrazine. The energetic reaction of hydrazine with strong oxidants, such as nitric acid, is utilized in rocket propulsion. The thermal decomposition of hydrazine produces free radicals and gases, useful in rubber curing and foam-rubber production.

A slightly weaker base than ammonia, hydrazine forms most of the analogs of ammonia derivatives as well as distinctive hydrazine derivatives in which both nitrogen atoms are involved. For example, hydrazine forms two series of salts, such as $N_2H_4 \cdot HCl$ and $N_2H_4 \cdot 2HCl$, and forms not only hydrazones, $RCH{=}NNH_2$, but also azines, $RCH{=}NN{=}CHR$, by reaction with aldehydes. In a manner similar to ammonia, hydrazine attacks polar bonds, in one case displacing ammonia from urea to form semicarbazide, a reaction which can be reversed by an excess of ammonia.

Prominent uses of hydrazine derivatives include rocket fuels [1,1-dimethylhydrazine, $(CH_3)_2NNH_2$]; antituberculin drugs (isonicotinic hydrazide, $C_5H_4N \cdot CO \cdot NHNH_2$); plant-growth regulators (maleic hydrazide, $NH \cdot CO \cdot CH{=}CH \cdot CO \cdot NH$, and β-hydroxyethyl hydrazine, $HOCH_2CH_2NHNH_2$); dye and explosive intermediates [aminoguanidine, $NH_2 \cdot C(NH) \cdot NHNH_2$]; algacides and fungicides [copper dihydrazinium sulfate, $CuSO_4 \cdot (N_2H_5)_2SO_4$]; soldering fluxes (hydrazine hydrobromide, $N_2H_4 \cdot HBr$); blowing agents for foam rubber (azides $R \cdot CO \cdot N_3$, and sulfonyl hydrazides, $R \cdot SO_2 \cdot NHNH_2$); insecticides (1,4-diphenylsemicarbazide, $C_6H_5 \cdot NH \cdot CO \cdot NHNH \cdot C_6H_5$); heterocycle syntheses (thiosemicarbazide, $NH_2 \cdot CS \cdot NHNH_2$) and polymers (dihydrazide-formaldehyde resins). *See* AMMONIA; NITROGEN. Theodore H. Dexter

Bibliography. *Kirk-Othmer Encyclopedia of Chemical Technology*, vol. 12, 3d ed., 1980.

Hydride

The isolated atomic hydrogen anion, H^-. It consists of a singly charged positive nucleus and two electrons. The electron-electron repulsion almost overwhelms the nuclear-electron attraction. Thus, the "extra" electron is held weakly and is readily donated. Ionic salts containing this large and easily polarized ion are highly reactive, strongly basic, and powerfully reducing. This makes them important reagents despite the fact that they are readily destroyed by the presence of the relatively acidic compound water (H_2O) or by exposure to the relatively oxidizing dioxygen (O_2) as found in air. *See* ELECTRON CONFIGURATION.

The term hydride also refers to salts containing the H^- anion and a highly electropositive alkali or alkaline-earth metal as the cation. The salt names reflect this high ionic character, for example, sodium hydride (NaH). In such salts the ionic radius of H^- is comparable to that of Cl^-.

There are also complex metal hydrides that are formed from the formal reaction of H^- salts with some more covalent metal or metalloid hydrogen compound. Among the earliest to be investigated were lithium aluminum hydride ($LiAlH_4$) and sodium borohydride ($NaBH_4$). These two species also have numerous derivatives in which one, two, or three hydrogens have been replaced by other univalent groups. All of these species formally transfer the H^- anion to suitable organic hydride acceptors. In these and related reactions, all the components of H^- (that is, two electrons and one proton) are transferred, although the precise electronic details depend on the molecular details.

Ideally, the term hydride should be reserved for those species that contain H^- or that at least formally transfer this ion to another substance in a so-called hydride transfer reaction. Such

reactions are found in the industrial synthesis of 2,2,4-trimethylpentane (isooctane) from isobutylene and isobutane, as well as in many of the classical organic chemistry named reactions, for example, the Cannizzaro reaction. Hydride transfer is also important in most of the biologically important oxidation-reduction reactions of the vitamin niacin (nicotinamide) as found in the forms of nicotinamide adenine dinucleotide/hydrogenated nicotinamide adenine dinucleotide (NAD^+/NADH) and nicotinamide adenine dinucleotide phosphate ($NADP^+$)/NADPH. *See* NIACIN; NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD); NICOTINAMIDE ADENINE DINUCLEOTIDE PHOSPHATE (NADP).

There are compounds with metal-hydrogen bonds that are also referred to as hydrides. For example, hydridocarbonyl [$HCo(CO)_4$] and dihydridotetracarbonyl iron [$H_2Fe(CO)_4$] are often named cobalt tetracarbonyl hydride and iron tetracarbonyl dihydride by analogy to corresponding halides. However, both species are quite acidic. These carbonyl hydrides and derived anions are important for both the academic and industrial chemist. So are other metal-hydrogen complexes with so-called soft ligands. The presence of two or more hydrogens in a metal complex does not necessarily convey two metal-hydrogen bonds of whatever polarity, but may be a complex of molecular hydrogen instead. *See* COORDINATION COMPLEXES; HYDRIDO COMPLEXES; HYDROGEN; LIGAND; METAL CARBONYL; METAL HYDRIDES. Joel F. Liebman

Bibliography. F. A. Cotton and G. Wilkinson, *Advanced Inorganic Chemistry*, 6th ed., 1999; N. N. Greenwood and A. Earnshaw, *Chemistry of the Elements*, 2d ed., 1997; M. Smith and J. March, *March's Advanced Organic Chemistry*, 5th ed., 2000.

Hydrido complexes

Complex hydrides containing a hydride ligand bonded to a central atom. The prefix hydro instead of hydrido is sometimes used. Sodium tetrahydridoborate, $NaBH_4$ (or sodium tetrahydroborate, originally called sodium borohydride), and lithium tetrahydridoaluminate, $LiAlH_4$ (originally called lithium aluminum hydride), are important reducing agents in synthetic and industrial reactions. $NaBH_4$ is employed in aqueous or alcoholic solutions, and $LiAlH_4$ is employed in ethers. Sodium cyanotrihydridoborate, $NaBH_3CN$, can be used in acidic medium. A family of aluminum-based reducing agents is now available commercially, including sodium diethyldihydridoaluminate, $NaAlH_2(C_2H_5)_2$; sodium tri-*tert*-butoxohydridoaluminate, $NaAlH(O\text{-}t\text{-}C_4H_9)_3$; and sodium bis(2-methoxyethoxo)-dihydridoaluminate, $NaAlH_2(OCH_2CH_2OCH_3)_2$. All are soluble in aromatic hydrocarbons. They are rather expensive, but their specific reducing powers make them attractive for synthesizing high-value products, such as pharmaceuticals, flavorings, fragrances, dyes, and insecticides.

Similar hydrides are C_6H_5MgH and $C_6H_5Mg_2H_3$. If zinc is considered a nontransition metal, then complex hydrides such as Li_3ZnH_5, Li_2ZnH_4, $LiZnH_3$, $NaZn_2H_5$, $LiZnH(CH_3)_2$, and $LiZn(CH_3)_2AlH_4$ may be grouped with the above compounds. *See* HYDROBORATION; METAL HYDRIDES.

Transition metals. More than a thousand hydrido complexes of transition metals have been prepared. Excepting Sc, Y, and La, hydrido compounds of all three transition series metals are known. In only two cases do the complex anions contain the central atom bonded to hydride ions and no other ligand; these are K_2ReH_9 and K_2TcH_9 (see below). The rest contain H-M-L linkages, where hydrogen is bonded to a transition metal M which is also bonded to one or more pi-bonded ligands L [for example, CO, CN^-, PF_3, or $P(C_6H_5)_3$]. The metal-hydrogen bonds are stabilized by the pi bonding between the metal and other ligands. The M-H bond lengths are generally 0.16–0.17 nanometer. These compounds are exceedingly diverse in nature and undergo a multitude of reactions. A few industrial processes involve hydrido complexes. Some chemists suspect that the enzymes which convert atmospheric nitrogen to ammonia (fixation) depend on Fe-H or Mo-H functions. *See* COOR-DINATION CHEMISTRY; COORDINATION COMPLEXES.

Enneahydridorhenate ion. A compound of formula K_2ReH_9 was isolated in 1960. Brackets are frequently employed to identify complex ions or molecules, for example, $K_2[ReH_9]$. The structure of the enneahydridorhenate ion, ReH_9^{2-}, is that of a triangular prism with three extra hydrogen atoms bonded through three faces, as shown in **Fig. 1**. One of these facial hydrogen atoms can be replaced by a phosphine molecule, the compound $K[ReH_8PH_3]$ being formed. The technetium compound analogous to $K_2[ReH_9]$, namely $K_2[TcH_9]$, is the second transition metal hydrido complex known which has only hydride ligands. Both compounds are fairly stable in alkaline aqueous solution.

Carbonyl hydrides. Hydrido complexes of metals also bonded to carbonyl ligands were the first such compounds discovered in 1931. The iron

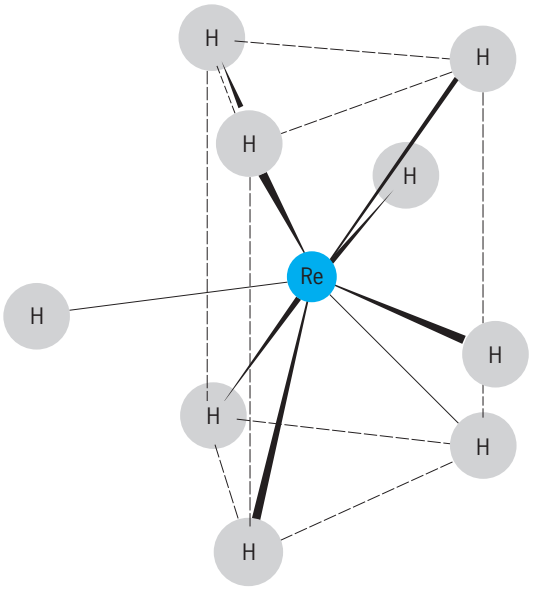

Fig. 1. Structure of ReH_9^{2-}.

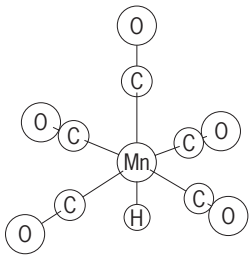

Fig. 2. Structure of [HMn(CO)$_5$].

compounds were prepared by the action of an alkali on iron pentacarbonyl as shown in reaction (1).

$$[Fe(CO)_5] + 3OH^- \longrightarrow [HFe(CO)_4]^- + CO_3{}^{2-} + H_2O \quad (1)$$

Acidification of the product from (1) yields dihydridotetracarbonyliron, [$H_2Fe(CO)_4$]. The corresponding cobalt compound is prepared by hydrogenation of dicobalt octacarbonyl as shown in reaction (2).

$$[Co_2(CO)_8] + H_2 \rightleftharpoons 2[HCo(CO)_4] \quad (2)$$

Some other hydridocarbonyls are [$HMn(CO)_5$] (**Fig. 2**), [$HRe(CO)_5$], and [$HV(CO)_6$]. The effective atomic number rule is valid in these cases. These compounds are acidic, and many derivatives are known, such as [$(CO)_4Co—Hg—Co(CO)_4$], a substance containing metal-metal bonds. The hydridocarbonyls [$HCo(CO)_4$] and [$HV(CO)_6$] are strong acids, while [$H_2Fe(CO)_4$], with $pK_1 = 4.4$, is about as acidic as acetic acid. All of these hydridocarbonyls are volatile, toxic compounds with obnoxious odors. Some hydridocarbonyls are listed below.

[$HZrCl(C_5H_5)_2$]	[$HFe(CO)_4$]$^-$
[$H_2Zr(C_5\{CH_3\}_5)_2$]	[$HFe(C_5H_5)(CO)_2$]$^+$
[$H_3Ta(C_5H_5)_2$]	*cis*-[$H_2Fe(PF_3)_4$]
[$HCr(C_5H_5)(CO)_3$]	[$HCo(CO)_4$]
[$H_2Mo(C_5H_5)_2$]	[$HCo(PF_3)_4$]
[$H_3Mo(C_5H_5)_2$]$^+$	[$HCo(CO)_3P(C_6H_5)_3$]
[$H_2W(C_5H_5)_2$]	[$H_2Co(P\{C_6H_5\}_3)]_3$]
[$HW(C_5H_5)(CO)_3$]	K_3[$HCo(CN)_5$]
[$HMn(CO)_5$]	[$HRu(C_5H_5)_2$]$^+$
[$HMn(PF_3)_5$]	[$HRu(C_5H_5)(CO)_2$]
[$HRe(C_5H_5)_2$]	[$HRuCl(P\{C_6H_5\}_3)_3$]
[$H_2Re(C_5H_5)_2$]$^+$	[$HPtCl(P\{C_2H_5\}_3)_2$

Knowledge of the hydrogen-metal bond in these complexes derives primarily from neutron and x-ray diffraction studies, nuclear magnetic resonance studies, and infrared absorption spectra. Proton (hydride) nuclear magnetic resonance shifts are exceptionally large, and to high field. *See* INFRARED SPECTROSCOPY; NEUTRON DIFFRACTION; NUCLEAR MAGNETIC RESONANCE (NMR); X-RAY DIFFRACTION.

The hydridocarbonyls are chemically reactive substances, as illustrated by olefin and carbon monoxide insertion reactions of the cobalt compound, reaction (3). The last compound reacts with

$$[HCo(CO)_4] \xrightarrow{H_2C=CH_2} [CH_3CH_2Co(CO)_4] \xrightarrow{CO} [CH_3CH_2COCo(CO)_4] \quad (3)$$

hydrogen to regenerate [$HCo(CO)_4$], forming an aldehyde, as shown in reaction (4).Thus hydridotetracarbonylcobalt functions as a catalyst. Further hydrogenation of the aldehyde leads to an alcohol. The above sequence of reactions is known as hydroformylation or the oxo reaction, and is employed industrially to produce long-chain alcohols. *See* HYDROFORMYLATION; METAL CARBONYL.

$$[CH_3CH_2COCo(CO)_4] \xrightarrow{H_2} [HCo(CO)_4] + CH_3CH_2CHO \quad (4)$$

Hydridocyclopentadienyl complexes. The hydrocarbon cyclopentadiene, C_5H_6, readily loses one hydrogen ion to form the cyclopentadienide ion, $C_5H_5{}^-$. The iron(II) derivative, [$Fe(C_5H_5)_2$], is bis(cyclopentadienyl)iron, the famous sandwich compound, ferrocene.

In 1955 hydrido-bis(cyclopentadienyl)rhenium, [$HRe(C_5H_5)_2$], was synthesized, and nuclear magnetic resonance studies proved that the hydrogen atom is bonded to the metal (**Fig. 3**).The compound is a Brönsted base, and neutralized dilute acids to form [$H_2Re(C_5H_5)_2$]$^+$. In 60% dioxane the value of pK_b for [$HRe(C_5H_5)_2$] is 8.5; pK_b for ammonia under the same conditions is 8.85. It is believed that in [$HRe(C_5H_5)_2$] and similar compounds there are three orbitals directed from the metal atom; electrons in these orbitals cause the two C_5H_5 rings to be angular rather than parallel. In [$H_2Re(C_5H_5)_2$]$^+$, hydrogen atoms are bonded through two of these orbitals, and all three are occupied in [$H_3Ta(C_5H_5)_2$] (which is not a base) and in [$H_3W(C_5H_5)_2$] $^+$. Mixed ligand complexes such as [$HCr(C_5H_5)(CO)_3$] were also discovered in 1955.

When attempts were made to prepare [$HM(C_5H_5)_2$], where M is Co, Rh, or Ir, it was found that the hydrogen atom bonds, not to the metal atom, but to one of the five-carbon rings, resulting in [$M(C_5H_5)(C_5H_6)$], which is not a hydrido complex.

Hydridophosphine and hydridocyano complexes. Trifluorophosphine (PF_3), triethylphosphine [$P(C_2H_5)_3$] and triphenylphosphine [$P(C_6H_5)_3$] are important ligands in hydrido-transition metal complexes. Examples of the first type are the colorless liquids [$HMn(PF_3)_5$], *cis*-[$H_2Fe(PF_3)_4$], and [$HCo(PF_3)_4$], which resemble the corresponding carbonyls. The structures of [$HRuCl(P\{C_6H_5\}_3)_3$] and of [$H_3Ir(P\{C_6H_5\}_3)_3$] are sketched in **Fig. 4**. Within a given periodic family the stability of a given type of hydride increases as shown by the following trend: [$HNiCl(P\{C_2H_5\}_3)_2$] can be detected but not isolated.

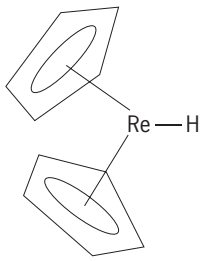

Fig. 3. Structure of [HRe(C_5H_5)$_2$]. The aromatic C_5H_5 groups are represented by pentagons.

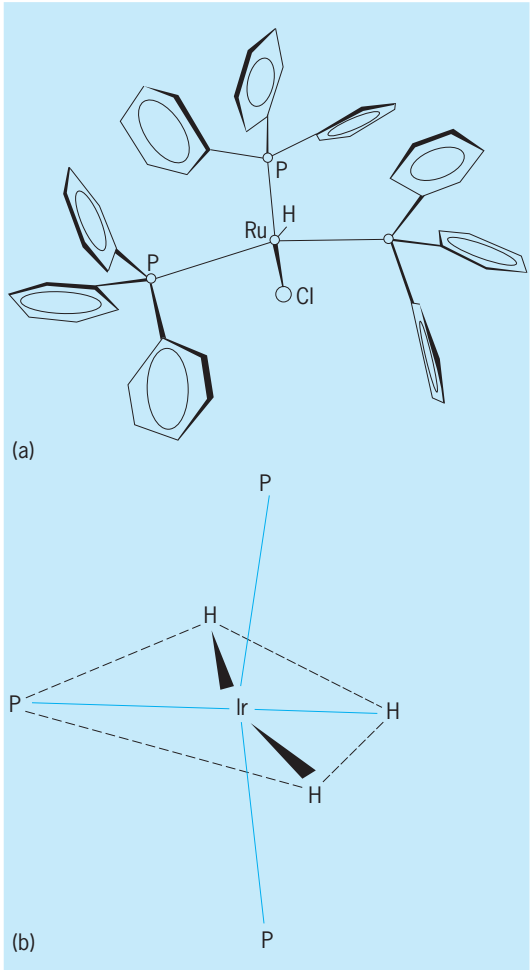

Fig. 4. Hydridophosphine structures, with the C_6H_5 groups not shown. (*a*) $[HRuCl(P\{C_6H_5\}_3)_3]$. (*b*) $[H_3Ir(P\{C_6H_5\}_3)_3]$.

$[HPdCl(P\{C_2H_5\}_3)_2]$ can be isolated, but is unstable. $[HPtCl(P\{C_2H_5\}_3)_2]$ is stable in air even at 100°C (212°F), and can be vacuum-distilled. Certain hydrido complexes of the platinum metals function as hydrogenation catalysts at room temperature and 1 atm (100 kilopascals) pressure. The catalysis is attributed to the ability of the metal atom to coordinate, via a vacant site, the reactant molecules, thereby orienting them and lowering the activation energy for bond making and breaking. The transition metal atom can supply or accept electrons as necessary. The hydrogen atom is facile, and under these circumstances is inserted into the substrate molecule. One well-known compound of this type is chloro-tris(triphenylphosphine)rhodium(I), Wilkinson's catalyst. Its role in catalyzing the hydrogenation of an olefin ($RCH{=}CH_2$) is shown in **Fig. 5**.

The mobility of the hydrido ligand is further shown by the reversible insertion of ethylene into an H-Pt bond. This is shown in reaction (5), where the

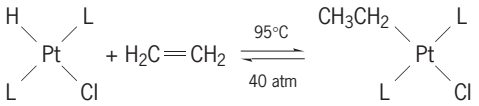

(5)

ligand L is $P(C_6H_5)_3$. The conversion of orange chlorotris (triphenylphosphine)iridium(I) to its colorless hydrido isomer is shown inreaction (6). Hydrido(tri-

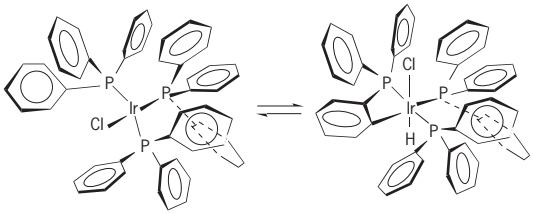

(6)

n-butylphosphine) copper(I), $[HCuP(C_4H_9)_3]$, which might be a polymer, is a mild, selected reducing agent employed in organic syntheses. It and similar complexes are very unstable and must be employed below −20°C (−4°F).

Reversible hydrogenation of the cyanocobalt(II) complex yields the hydridocyanocobalt(III) derivative, shown in reaction (7). Hydrido complexes of

$$2[H_2OCo(CN)_5]^{3-} + H_2 \rightleftharpoons [HCo(CN)_5]^{3-} + 2H_2O \qquad (7)$$

this type are catalysts in selectively reducing diene hydrocarbons such as butadiene, C_4H_6, to the monoene stage, C_4H_8.

Polynuclear hydrido complexes. Reduction of $Cr(CO)_6$ with $NaBH_4$ yields the yellow binuclear

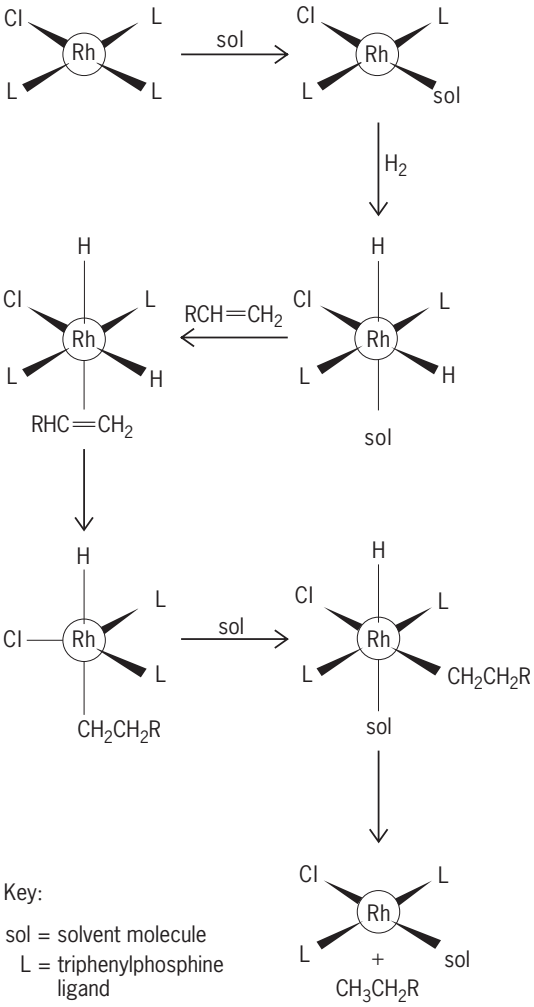

Fig. 5. Reaction scheme for hydrogenation of an olefin using Wilkinson's catalyst.

(two-chromium) anion, $[HCr_2(CO)_{10}]^-$. The earliest structural studies of this ion indicated that the hydrogen atom lies on the Cr-Cr axis, bonding the two $Cr(CO)_5$ units together. More recently it was demonstrated that the Cr-H-Cr linkage is inherently bent, as indicated in **Fig. 6**. The hydrogen atom is about 0.03 nm from the center of the Cr-Cr bond. In $[HW_2(CO)_{10}]^-$ the hydrogen atom is about 0.07 nm off-axis. In these compounds there is an electron-deficient three-center linkage. Only one electron pair is involved, as in the case of diborane. When

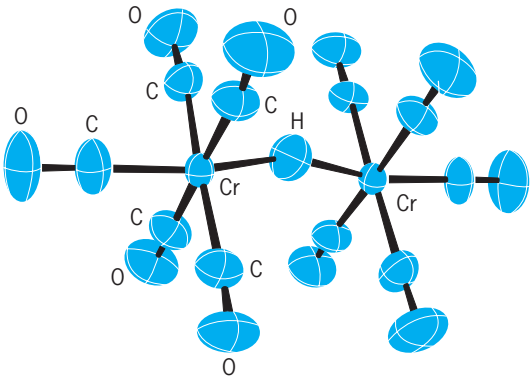

Fig. 6. Structure of $[HCr_2(CO)_{10}]^-$.

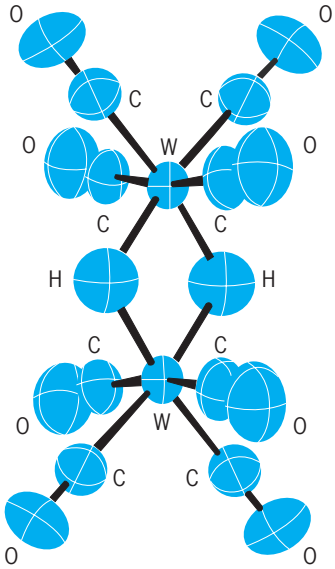

Fig. 7. Structure of $[H_2W_2(CO)_8]^{2-}$.

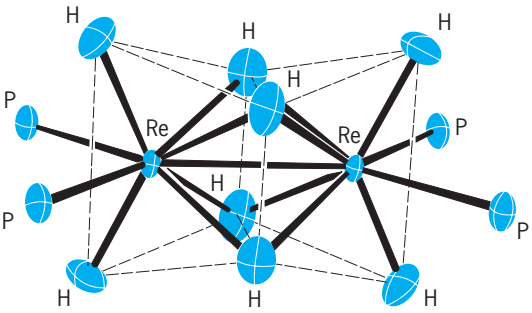

Fig. 8. Core of the $[H_8Re_2(P\{C_2H_5\}_2C_6H_5)_4]$ molecule. All organic groups are omitted for clarity.

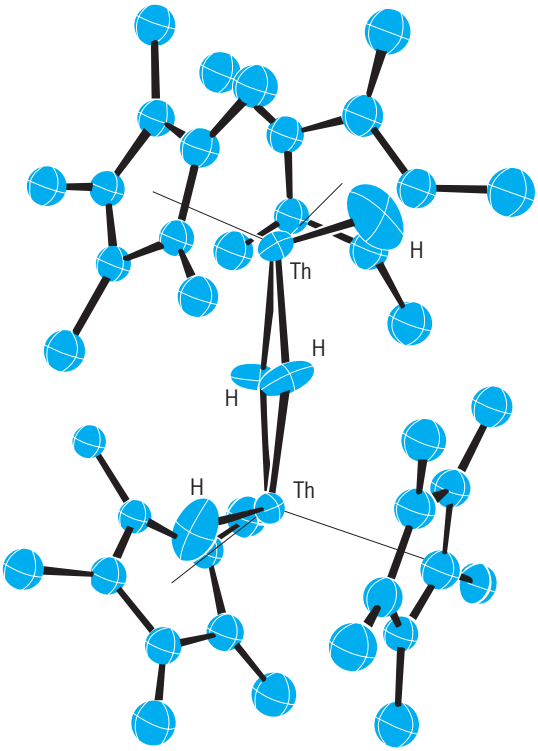

Fig. 9. Structure of $[H_2Th(C_5\{CH_3\}_5)_2]_2$. The hydrogen atoms of the methyl groups are omitted for clarity.

$[C_5H_5Mo(CO)_3]_2$ is dissolved in concentrated sulfuric acid, it is protonated to $[C_5H_5Mo(CO)_3]_2H^+$, which also has a hydrogen bridge bond. In $[H_2W_2(CO)_8]^{2-}$ there are two such bridging atoms, as sketched in **Fig. 7**. The molecule $[H_8Re_2(P\{C_2H_5\}_2C_6H_5)_4]$ has an Re-Re bond and no less than four bridging H atoms, as well as four terminal H atoms. The core of this structure is shown in **Fig. 8**. Each hydrogen atom in $[H_3Ni_4(C_5H_5)_4]$ is bounded to three nickel atoms. The hydrogen atom of $[HCo_6(CO)_{15}]^-$, a cluster ion, is in the center of the octahedron formed by the six cobalt atoms. A recently discovered hydrido complex of thorium is the dimer of $[H_2Th(C_5\{CH_3\}_5)_2]$. The organic ligands are pentamethylcyclopentadienyl groups. As shown in **Fig. 9**, two hydrogen atoms are bridging and two are terminal. Some polynuclear hydrido complexes are listed below.

$[HNb_6I_{11}]$	$[H_2Re_3(CO)_{12}]^-$
$[HCr_2(CO)_{10}]^-$	$[HFe_3(CO)_{11}]$
$[C_5H_5Mo(CO)_3]_2H^+$	$[HCo_6(CO)_{15}]^-$
$[HW_2(CO)_{10}]-$	$[H_2Ni_2(CO)_6]$
$[H_2W_2(CO)_8]^{2-}$	$[H_3Ni_4(C_5H_5)_4]$
$[H_2Mn_2(CO)_9]$	$[HNi_{12}(CO)_{21}]^{3-}$
$[H_8Re_2(PR_3)_4]$, (R = ethylorphenyl)	$[HCuP(C_6H_5)_3]_6$

James C. Warf

Bibliography. A. Dedieu, *Transition Metal Hydrides*, 1992; A. P. Ginsberg, *Transition Metal Chemistry*, vol. 1, 1965; E. L. Muetterties, *Transition Metal Hydrides*, 1971.

Hydroboration

The process of producing organoboranes by the addition of diborane to unsaturated organic compounds. In ether solvents the addition of diborane to such molecules is exceedingly rapid and essentially quantitative. This reaction therefore makes the organoboranes readily available. Such organoboranes find application as intermediates for organic synthesis.

Procedures. Diborane is highly soluble in tetrahydrofuran, where it exists as the addition compound tetrahydrofuran-borane. Such solutions are often used for hydroboration, and merely involve bringing the two reactants together as indicated by reaction (1).

$$3RCH{=}CH_2 + C_4H_8O{:}BH_3 \longrightarrow (RCH_2CH_2)_3B + C_4H_8O \quad (1)$$

Alternatively, sodium borohydride may be utilized to achieve hydroboration by the addition of boron trifluoride etherate. This is shown by reaction (2).

$$12RCH{=}CH_2 + 3NaBH_4 + 4(C_2H_5)_2O{:}BF_3 \longrightarrow 4(RCH_2CH_2)_3B + 3NaBF_4 + 4(C_2H_5)_2O \quad (2)$$

Usually the organoborane is not isolated but is utilized in place, similar to applications of the Grignard reagent in synthesis.

Scope and stoichiometry. Essentially all molecules containing one or more double or triple bonds undergo rapid conversion to organoboranes by this procedure. In general, disubstituted olefins react to give trialkylboranes, trisubstituted olefins react rapidly to give dialkylboranes and only slowly beyond, and tetrasubstituted olefins react rapidly to the monoalkylborane stage. The respective reactions are indicated by (3), (4), and (5).

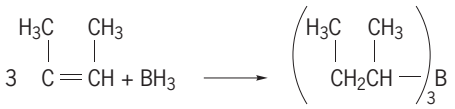
(3)

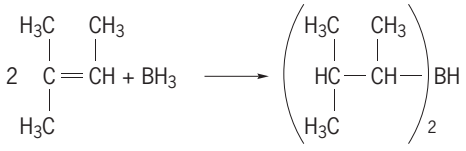
(4)

Disiamylborane

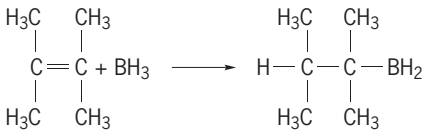
(5)

t-Hexylborane

Such substituted organoboranes are often very useful for controlled hydroborations, as indicated by reactions (6) and (7).

$$R_2BH + H_2C{=}CHCH_2Cl \longrightarrow R_2BCH_2CH_2CH_2Cl \quad (6)$$

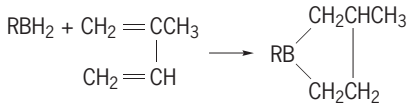
(7)

The hydroboration of unsaturated molecules containing functional groups is usually easily accomplished. Reaction (8) illustrates this technique. In

$$R_2BH + CH_2{=}CHCH_2CO_2C_2H_5 \longrightarrow R_2BCH_2CH_2CH_2CO_2C_2H_5 \quad (8)$$

this way organoboranes containing such functional groups become readily available for organic synthesis.

Asymmetric syntheses. The hydroboration of optically active α-pinene yields an optically active dialkylborane. This can be utilized to introduce an asymmetric center, as illustrated in reaction (9), the synthesis of optically active 2-butanol in high optical purity. In reaction (9) the asymmetric center is indicated by an asterisk and [O] indicates an oxidation step.

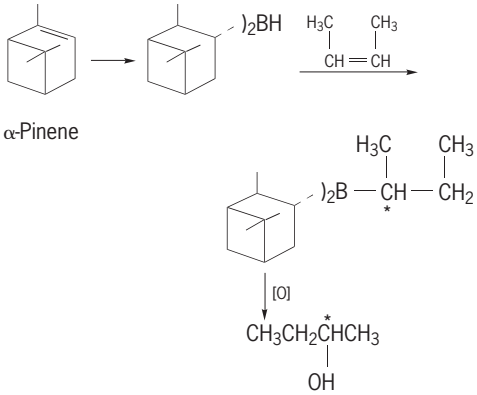
(9)

α-Pinene

2-Butanol

Directive effects. Hydroboration generally proceeds to place the boron atom at the less substituted of the two carbon atoms of a double bond (anti-Markovnikov addition). Since the boron atom may readily be replaced by many functional groups, such as hydroxyl and amino, this makes possible the anti-Markovnikov hydration and amination of double bonds, as shown by reaction (10).

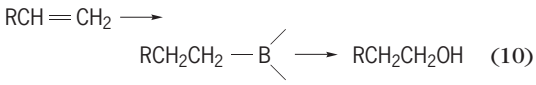
(10)

Cis addition. The hydroboration reaction appears to involve a simple four-center cis addition of the hydrogen-boron bond to the carbon-carbon double bond. Reaction (11) indicates this addition.

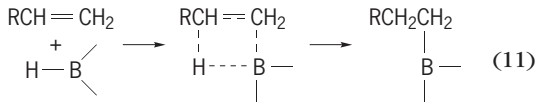
(11)

This oxidation proceeds with retention of configuration. Thus the hydroboration-oxidation of

cyclic olefins, such as of 1-methylcyclohexene in reaction (12), provides the pure trans alcohol.

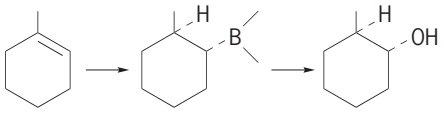

(12)

1-Methyl cyclohexene

Steric effects. The hydroboration reaction is quite sensitive to steric influences. Thus it hydroborates bicyclic olefins, such as norbornene in reaction (13),

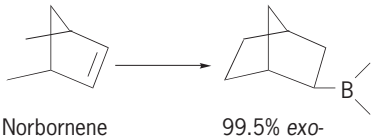

(13)

Norbornene 99.5% *exo*-

predominantly from the less hindered side. Moreover, the reaction does not cause rearrangements even in labile systems.

Stereospecific syntheses. These unusual characteristics—(1) anti-Markovnikoff addition, (2) freedom from rearrangement, (3) cis addition, and (4) high degree of steric control—give the hydroboration reaction major importance for achieving stereospecific syntheses. This is indicated in the reaction sequence (14), which is the synthesis of isopinocampheol from α-pinene.

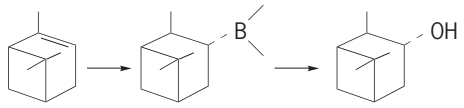

(14)

α-Pinene Isopinocampheol

Isomerization. Although the hydroboration reaction is remarkably free of rearrangements of the carbon structure, even in labile systems, the boron atom is capable of facile migration around the carbon skeleton at temperatures of 100–150°C (212–300°F). The reaction achieves a thermodynamic equilibrium among the possible organoboranes with that particular carbon skeleton. The preferred isomer is the one in which the boron atom is in the least crowded position. Reaction (15) illustrates this isomerization.

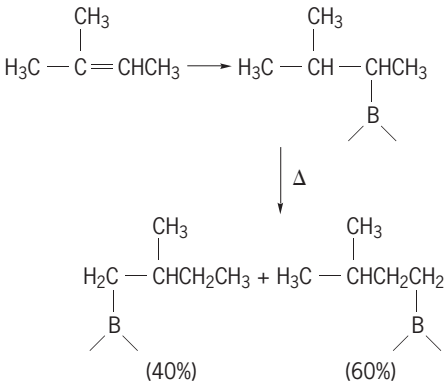

(15)

Displacement reaction. Heating an organoborane with another olefin transfers the boron atom to the new olefin, liberating the original alkyl group as an olefin, as in reaction (16). The reaction can be made

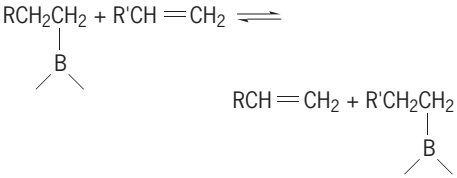

(16)

to proceed to essential completion by taking advantage of differences in volatility of the olefins, by using an excess of the displacing olefin, or by using an olefin, such as ethylene, which forms a very stable organoborane.

Contrathermodynamic isomerization. A combination of isomerization of the organoborane and the displacement reaction makes it possible to move double bonds from the more stable position at an alkyl branch or internal position to the less stable, unbranched terminal position. This is shown by the sequence in reaction (17).

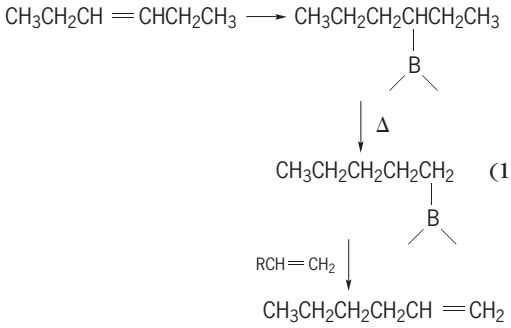

(17)

Oxidation. Organoboranes are readily oxidized by oxygen. Consequently, reactions of organoboranes are generally carried out under an inert atmosphere, such as nitrogen. Organoboranes are oxidized exceedingly readily by alkaline hydrogen peroxide, and this is the reagent of choice for the synthesis of alcohols via hydroboration, as shown by reaction (18).

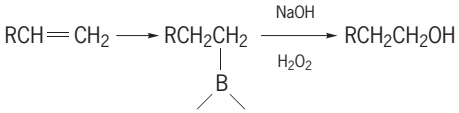

(18)

Amination. Chloramine and *o*-hydroxylaminesulfonic acid convert organoboranes into the corresponding amine. Amination with *o*-hydroxylaminesulfonic acid is shown by reaction (19).

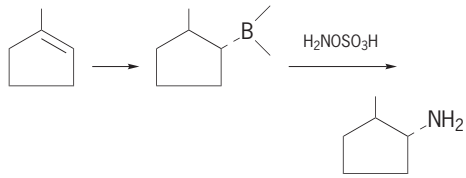

(19)

Coupling. Treatment of an olefin after hydroboration with alkaline silver nitrate produces a

coupled product, reactions (20) and (21), that is a

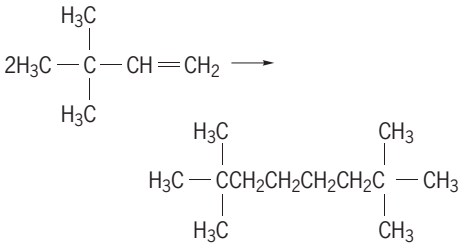

(20)

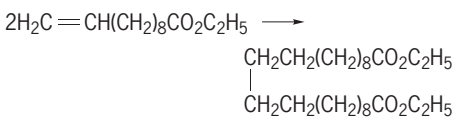

(21)

new compound composed of two molecules of the original.

1,4-Additions. Organoboranes react very rapidly with certain α,β-unsaturated aldehydes and ketones to give saturated products. This is illustrated by reaction (22).

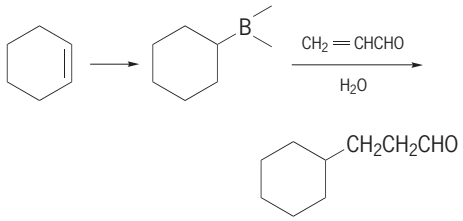

(22)

Condensations. Organoboranes react with α-halo-substituted carbanions to transfer alkyl groups from boron to carbon. This is shown by reactions (23) and (24), where the symbol KO*t*-Bu represents potassium tertbutoxide.

$$R_3B + CH_2BrCO_2C_2H_5 \xrightarrow{KOt\text{-}Bu} RCH_2CO_2C_2H_5 \quad (23)$$

$$R_3B + CH_2BrCO_2C_2H_5 \xrightarrow{KOt\text{-}Bu} RCHBrCO_2C_2H_5 \quad (24)$$

Carbonylation. Organoboranes react with carbon monoxide to produce intermediates which can be converted to tertiary alcohols, as in reaction (25),

$$R_3B + CO \longrightarrow R_3CBO \xrightarrow{[O]} \underset{\text{Tertiary alcohol}}{R_3COH} \quad (25)$$

secondary alcohols, ketones, aldehydes, methylol derivatives, ring ketones, as in reaction (26), and polycyclics, as in reaction (27).

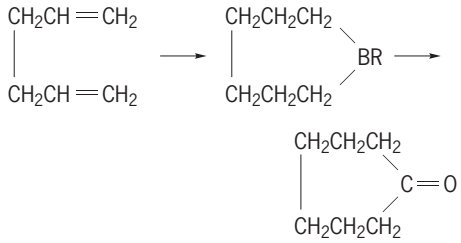

(26)

Ring ketone

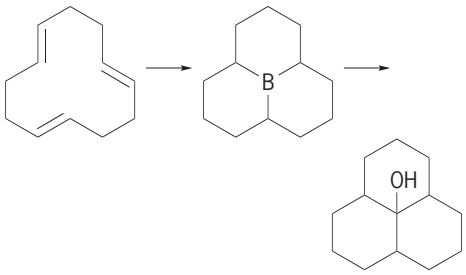

(27)

Polycyclic compound

Future developments. It is already apparent that the organoboranes are among the most versatile synthetic intermediates that are available to the organic chemist. The hydroboration reaction has made these intermediates readily available. *See* BORANE; CARBORANE; METAL HYDRIDES; ORGANIC SYNTHESIS; ORGANOMETALLIC COMPOUND. Herbert C. Brown

Bibliography. H. C. Brown, *Hydroboration*, 1962; H. C. Brown, *Organic Syntheses via Boranes*, 1975; G. E. Coates and K. Wade, *Organometallic Compounds*, vol. 1, pt. 1: *Groups I–III*, 4th ed., 1999; G. Cragg, *Organoboranes in Organic Synthesis*, 1973; T. Onak, *Organoborane Chemistry*, 1975; K. Smith, *Organometallic Compounds of Boron*, 1985.

Hydrocharitales

An order of aquatic flowering plants, division Magnoliophyta Angiospermae), in the subclass Alismatidae of the class Liliopsida (monocotyledons). The order consists of the single family Hydrocharitaceae, with about 100 species. Within the subclass the order is marked by its inferior, compound ovary with basically laminar placentation. The ovules are scattered over the walls of the individual carpels, which are weakly connate to form partial partitions (intruded placentas) in the ovary. The aquarium plant *Vallisneria spiralis*, called tape grass or eelgrass, is not a grass but belongs to the Hydrocharitaceae. *See* ALISMATIDAE; LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM. Arthur Cronquist

Hydrocracking

A catalytic, high-pressure process flexible enough to produce either of the two major light fuels—high octane gasoline or aviation jet fuel. It proceeds by two main reactions: adding hydrogen to molecules too massive and complex for gasoline and then cracking them to the required fuels. The process is carried out by passing oil feed together with hydrogen at high pressure (1000–2500 lb/in.^{2} gage or 7–17 megapascals) and moderate temperatures (500–750°F or 260–400°C) into contact with a bifunctional catalyst, comprising an acidic solid and a hydrogenating metal component. Gasoline of high octane number is produced, both directly and through a subsequent step such as catalytic reforming; jet fuels may also be

manufactured simply by changing conditions with the same catalysts. The process is characterized by a long catalyst life (2–4 years), though a slow decline in activity occurs, caused by the deposition of carbonaceous material on the catalyst. Regeneration at intervals by burning off these deposits restores the activity, but eventually the catalyst porosity is destroyed and it must be replaced.

Generally, the process is used as an adjunct to catalytic cracking. Oils, which are difficult to convert in the catalytic process because they are highly aromatic and cause rapid catalyst decline, can be easily handled in hydrocracking, because of the low cracking temperature and the high hydrogen pressure, which decreases catalyst fouling. Usually, these oils boil at 400–1000°F (200–540°C), but it is possible to process even higher-boiling feeds if very high hydrogen pressures are used. However, the most important components in any feed are the nitrogen-containing compounds, because these are severe poisons for hydrocracking catalysts and must be removed to a very low level.

Hydrocracking was carried out on a practical scale in Germany and England starting in the 1930s. In this early work, a common hydrocracking catalyst was tungsten disulfide on acid-treated clay; thus, both hydrogenation and acidic components were present. Generally, a light oil from coal or coking products was vaporized and passed over the catalyst at high pressure. After separation of gasoline from the products, the unconverted material was returned to the reactor with a fresh portion of feed. Because this catalyst was not very active, the process had to be carried out at very high pressures and temperatures (4000 lb/in.^{2} gage or 28 MPa; 750°F or 400°C). It was costly and the products were not of high quality.

Research in the United States concentrated on the development of much more active catalysts, a different mode of operation, and the use of heavier oil feeds. As a result, the reaction is carried out in two separate, consecutive stages; in each, oil and hydrogen at high pressure flow downward over fixed beds of catalyst pellets placed in large vertical cylindrical vessels.

First stage. In the first, or pretreating, stage the main purpose is conversion of nitrogen compounds in the feed to hydrocarbons and to ammonia by hydrogenation and mild hydrocracking. Typical conditions are 650–740°F (340–390°C), 150–2500 lb/in.^{2} gage (1–17 MPa), and a catalyst contact time of 0.5–1.5 h; up to 1.5 wt % hydrogen is absorbed, partly by conversion of the nitrogen compounds, but chiefly by aromatic compounds which are hydrogenated. It is most important to reduce the nitrogen content of the product oil to less than 0.001 wt% (10 parts per million). This stage is usually carried out with a bifunctional catalyst containing hydrogenation promotors, for example, nickel and tungsten or molybdenum sulfides, on an acidic support, such as silica-alumina. The metal sulfides hydrogenate aromatics and nitrogen compounds and prevent deposition of carbonaceous deposits; the acidic support accelerates nitrogen removal as ammonia by breaking carbon-nitrogen bonds. The catalyst is generally used as $^1/_8 \times {}^1/_8$ in. (0.32 × 0.32 cm) or $^1/_{16} \times {}^1/_8$ in. (0.16 × 0.32 cm) pellets, formed by extrusion.

Second stage. Most of the hydrocracking is accomplished in the second stage, which resembles the first but uses a different catalyst. Ammonia and some gasoline are usually removed from the first-stage product, and then the remaining oil, which is low in nitrogen compounds, is passed over the second-stage catalyst. Again, typical conditions are 600–700°F (300–370°C), 1500–2500 lb/in.^{2} gage (10–17 MPa) hydrogen pressure, and 0.5–1.5 h contact time; 1–1.5 wt % hydrogen may be absorbed. Conversion to gasoline or jet fuel is seldom complete in one contact with the catalyst, so the lighter oils are removed by distillation of the products and the heavier, high-boiling product combined with fresh feed and recycled over the catalyst until it is completely converted.

The catalyst for the second stage is also a bifunctional catalyst containing hydrogenating and acidic components. Metals such as nickel, molybdenum, tungsten, or palladium are used in various combinations, dispersed on solid acidic supports such as synthetic amorphous or crystalline silica-aluminas, such as zeolites. These supports contain strongly acidic sites and sometimes are enhanced by the incorporation of a small amount of fluorine. A long period (for example, 1 year) between regenerations is desirable; this is achieved by keeping a low nitrogen content in the feed and avoiding high temperatures, which lead to excess cracking with consequent deposition of coke on the catalyst. When activity of the catalyst does decrease, it can be restored by carefully controlled burning of the coke.

The catalyst is the key to the success of the hydrocracking process as now practiced, particularly the second-stage catalyst. Its two functions must be most carefully balanced for the product desired; that is, too much hydrogenation gives a poor gasoline but a good jet fuel. The oil feeds are composed of paraffins, other saturates, and aromatics—all complex molecules boiling well above the required gasoline or jet-fuel product. The catalyst starts the breakdown of these components by forming from them carbonium ions, that is, positively charged molecular fragments, via the protons (H^+) in the acidic function. These ions are so reactive that they change their internal molecular structure spontaneously and break down to smaller fragments having excellent gasoline qualities. The hydrogenating function aids in maintaining and controlling the ion reactions and protects the acid function by hydrogenating coke precursors off the catalyst surface, thus maintaining catalyst activity. Any olefins formed in the carbonium ion decomposition are also hydrogenated.

Products. The products from hydrocracking are composed of either saturated or aromatic compounds; no olefins are found. In making gasoline, the lower paraffins formed have high octane numbers; for examble, the 5- and 6-carbon number fractions have leaded research octane numbers of 99–100. The remaining gasoline has excellent properties as a feed to catalytic reforming, producing a highly

aromatic gasoline which, with added lead, easily attains 100 octane number. Both gasolines are suitable for premium-grade motor gasoline. Another attractive feature of hydrocracking is the low yield of gaseous components, such as methane, ethane, and propane, which are less desirable than gasoline. When making jet fuel, more hydrogenation activity of the catalysts is used, since jet fuel contains more saturates than gasoline. *See* GASOLINE.

The hydrocracking process is being applied in other areas, notably, to produce lubricating oils and to convert very asphaltic and high-boiling residues to lower-boiling fuels. Its use will certainly increase greatly in the future, since it accomplishes two needed functions in the petroleum-fuel economy: Large, unwieldy molecules are cracked, and the needed hydrogen is added to produce useful, high-quality fuels. *See* AROMATIZATION; CRACKING; HYDROGENATION; ISOMERIZATION; REFORMING PROCESSES.

Charles P. Brewer

Bibliography. L. W. Hall, *Petroleum Production Operations*, 1986; J. McKetta, *Petroleum Processing Handbook*, 1992.

Hydroelectric generator

A low-speed generator driven by water turbines. Hydrogenerators may have a horizontal or vertical shaft. The horizontal units are usually small with speeds of 300–1200 revolutions per minute (rpm). The vertical units are usually larger and more easily adapted to small hydraulic heads. The rotor diameters range from 2 to 62 ft (0.6 to 19 m) and capacities from 50 to 900,000 kVA. The generators are rated in kVA (kilovolts times amperes). The kilowatt output is the product of kVA and power factor. The normal power-factor rating of small synchronous generators is between 0.8 and 1.0 with 0.9 being common. For large generators a rating of 0.9–0.95 is common with the machines able to operate up to 1.0 when the load requires. The generators may also supply reactive power. *See* ALTERNATING CURRENT; ELECTRIC POWER MEASUREMENT; VOLT-AMPERE.

Structure. The turbine shown in the **illustration** has an adjustable blade propeller, typical of large, low-head units that are common on large river power plants. The water enters the turbine spiral scroll casing, falls down through the turbine, causing rotation, and empties into the river. The shaft transmits the rotation to the generator spider or hub and thence to the rotor rim and poles. The magnetic field of the rotor poles transmits the torque to the stator and changes the mechanical power to electrical power. *See* HYDRAULIC TURBINE; TURBINE.

The poles are spaced around the rotor rim and are magnetized by direct current flowing in the turns of the field coil around each pole. The magnetic field, or flux, crosses the air gap between rotor and stator, flows radially through the stator teeth and thence to the area one pole pitch away, and back to the adjacent pole on the rotor. The magnetic flux is stationary with respect to the rotor poles but sweeps around the stator at the peripheral rotor speed. Coils are installed in the stator slots between the teeth. Thus there is an ever- changing flux linking stator coils, which causes an induced electromotive force in the coils according to Faraday's law. The alternating voltage induced by the changing flux is given by

$$E = 4.4fN\phi$$

where E is the root-mean-square (rms) value of the induced voltage in volts, f is the system frequency in hertz, N is the number of turns in the coil, and ϕ is the maximum flux linkage in the coil in webers (= flux in lines $\times 10^{-8}$). *See* ALTERNATING-CURRENT GENERATOR.

Almost all machines have three groups of coils per pole, making a three-phase generator. The windings are connected; thus the sum of the coil voltages in each phase is equal to the rated phase voltage. The phase voltages have a time distribution: sequential phases reach their positive maximum at intervals of one-third of each cycle. This provides the best practical combination for transmission and a rotating field for driving motors.

The direct current required to magnetize the field poles is supplied by a small exciter generator mounted on the rotor or by an electronically controlled rectifier. The current is fed to the field through brushes on collector rings on the rotor.

The rotating parts of both generator and turbine are supported by the generator thrust bearing, which also supports the hydraulic thrust of the turbine. This thrust may be as much as the total weight of the rotating parts. Guide bearings on the turbine and generator keep the rotating parts on center. These bearings are usually of the shoe type, operating in a bath of oil.

All hydromachines are required to withstand the highest attainable speed (runaway speed) with the governor inoperative after tripping from the line. For good governing, the flywheel effect may have to be larger than inherent in the electrical parts. This may affect the diameter and weight of the unit. *See* GOVERNOR.

Large machines need a housing for reasons of safety, noise reduction and cleanliness. The air is circulated by the fan action of the rotor between the poles and through the radial ducts in the stator core, cooling both the stator iron and the coils. The air then passes through coolers and recirculates.

Standards and characteristics. Most modern machines are built to the 1977 ANSI standard with class F insulation on the windings but temperature limits for class B insulation. The limit for this class of machines is a stator-winding temperature rise of 80°C (144°F), above an ambient temperature of 40°C (104°F), as measured by a resistance detector. The temperature rise observed by the detector is frequently less than the 80°C (144°F) specified as the practical insulation limit. The temperature-rise limit on the rotor field windings is also 80°C (144°F), but as measured by change in field resistance.

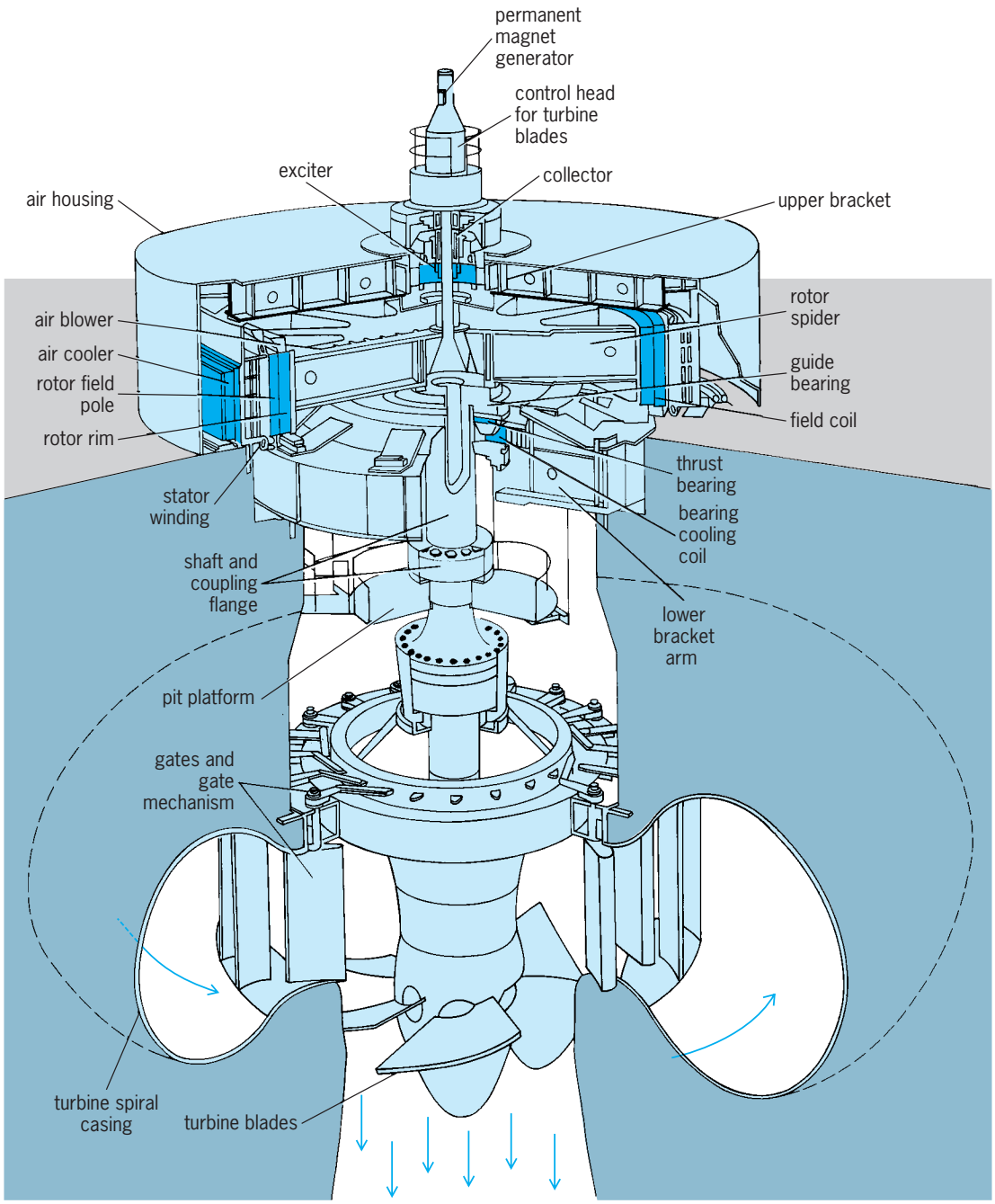

Large hydroelectric generator. (*Westinghouse Electric Corp.*)

In 1977 the standard method of rating hydrogenerators was changed to allow the 80°C (144°F) temperature rise but with no overload rating. The rated kVA of the generator should nearly match the maximum expected turbine capability and should be operated below the name-plate rating most of the time for best winding life. Before this change, the generator nameplate kVA and power factor were based on a 60°C (108°F) rise and about matched the turbine operating point of highest efficiency. With 15% overload and a higher temperature rise, the generator could usually handle the turbine capability.

Hydrogenerators of normal characteristics usually serve for application in the United States with the plant near a large power concentration or substation. However, when long lines are required or special problems exist, subnormal reactances and other characteristics with larger underexcited capability may be required for stability and voltage regulation. These features may be purchased with the equipment.

Induction generator. An induction generator is the same as an induction motor, except that it is driven by a turbine at a few percent above a synchronous speed instead of the motor driving a load a few percent below a synchronous speed. Induction generators are used locally to recover energy at small dams. They receive excitation or magnetizing kVA from the

transmission line; thus their kVA ratings are always higher than those equivalent synchronous machines with the same turbine kilowatt output. They often have step-up gears to raise the generator speed to 720–1800 rpm, and must withstand the higher speed required by the turbine. The power system capability at the machine terminals must be large compared to the machine capacity. The power company may require a penalty to supply the necessary reactive kVA. Some installations use capacitors connected to the line to correct this power factor. This can be dangerous if the machine is tripped free of the system with the capacitors still connected; the machine may overexcite by resonance and fail because of the overvoltage generated. Induction machines are satisfactory for many applications and are a little less expensive than synchronous machines because of the elimination of a voltage regulator and excitation supply. However, induction machines are seldom satisfactory below 400–500 rpm because of their low power factor (0.6–0.8) and low efficiency (2–7% less). In addition, the mechanical deflection of parts with the variable hydraulic thrust from the turbine causes gear and generator misalignment and noise. *See* ELECTRIC POWER GENERATION; ELECTRIC ROTATING MACHINERY; GENERATOR; INDUCTION MOTOR; WATERPOWER. Eugene C. Whitney

Bibliography. American National Standards Institute, *Rotating Electrical Machinery-Synchronous Machines*, ANSI Stand. C50.10-1990, 1990; E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; Institute of Electrical and Electronics Engineers, *IEEE Guide for Operation and Maintenance of Hydro-Generators*, IEEE Stand. 492–1999, 1999; Institute of Electrical and Electronics Engineers, *IEEE Guide: Test Procedures for Synchronous Machines*, IEEE Stand. 115–1995, 1995; Institute of Electrical and Electronics Engineers, *IEEE Standard for Salient-Pole 50 Hz and 60 Hz Synchronous Generators and Generator-Motors for Hydraulic Turbine Applications Rated 5 MVA and Above*, IEEE Stand. C50.12–2005, 2005.

Hydrofoil craft

A form of high-speed ship that supports its weight by means of wings (properly called hydrofoils, or simply foils) beneath the surface of the water. The hydrofoils generate lift by movement in the same manner as an airplane wing. The hydrofoil was conceived in order to produce faster ships. The most effective means of developing a faster ship is to find a way to lift the hull clear of the water. This greatly reduces the drag on the hull, in turn greatly reducing the power required to drive the ship. The hydrofoil ship is one means to this end. *See* AIRFOIL; SHIP POWERING, MANEUVERING, AND SEAKEEPING.

Many inventors, including Alexander Graham Bell, pursued the hydrofoil concept. A patent for a hydrofoil was granted in the United States in the late 1880s. The earliest record of a successful hydrofoil flight was in 1894, when the Meacham brothers demonstrated a 14-ft (4.3-m) test craft in Chicago. This use of a wing to support a boat precedes by nearly 10 years the Wright brothers' first airplane flight in 1903.

TABLE 1. Specifications of U.S. Navy PHM Class hydrofoil craft

Weight	243 metric tons	239 long tons
Length	44 m	145 ft
Speed	48 knots	89 km/h
Power	13.4 MW	18,000 hp
Year first built	1997	
Number built	6	

The hydrofoil concept was not successfully exploited until the development of lightweight engines and structural materials. These developments—many of which came from the aircraft industry—made possible commercial and military hydrofoils.

There are two basic types of hydrofoils: fully submerged and partially submerged.

Fully submerged hydrofoils. The most efficient hydrofoil craft, from a powering standpoint, are those with fully submerged foils; that is, the wings are completely below the surface of the water. They are not subject to surface interactions such as air drawing, or to the danger of broaching through small waves. They are also relatively unaffected by wave action, resulting in excellent ship ride comfort. However, because they are fully submerged, the ship must be actively controlled. A conventional surface ship will stay on the surface, regulating its height by the simple action of floating. A fully submerged hydrofoil, by contrast, has no passive or inherent means to maintain its ride height and attitude. The ship may slowly (or rapidly) rise or fall, subject to small perturbations in the foil angles. This behavior could be wild and chaotic in an uncontrolled hydrofoil. All existing fully- submerged-type hydrofoils overcome this through the use of an active control system which senses the craft's height and attitude and adjusts the

Fig. 1. U.S. Navy Patrol Hydrofoil, Missile (PHM) Class ship. (*U.S. Navy*)

Fig. 2. JetFoil *Kamehameha*. (*Boeing Airplane Co.*)

Fig. 3. Rodriquez hydrofoil in the RHS-160 series.

wings as necessary to maintain the desired condition. *See* CONTROL SYSTEMS.

Partially submerged hydrofoils. The need for an active control system in a fully submerged hydrofoil led to the development of the partially submerged hydrofoil. In a partially submerged hydrofoil, the foils are generally V-shaped when viewed head on, with the apex of the V below the water and the tips of the wings above the water. In this case the amount of lift generated will increase if the foils are more deeply submerged, and it will decrease if the foils are less submerged. (The increase and decrease in lift are due to the change in the wetted area of the foils.)

The result of this behavior is that if the craft rises, the foils will lose some lift, resulting in its settling downward. If the craft settles beyond equilibrium, excess lift will be generated and the craft will rise. In actual designs, this behavior results in a steady ride attitude at an equilibrium balance of lift. A disadvantage of partially submerged foils is that the ride is not as smooth, especially in rough water, since the foils will react to each wave they encounter.

Limitations. Two significant limitations constrain the application of hydrofoil craft in any given mission. The first is a consequence of size. The weight of a ship increases approximately as the cube of a change in ship dimensions. (Double the length, width, and height of the ship, and the weight will increase approximately eightfold.) The lift of a hydrofoil, however, increases only as the square of a dimension change. (Double the span and chord of the foil, and the lift will increase approximately fourfold.) This cube/square relationship is increasingly difficult to overcome as the respective values increase. The variation in size between, say, a 30-m (100-ft) craft and a 40-m (130-ft) craft can be overcome with detailed engineering. This is not easily accomplished, however, for a 100-m-long (330-ft) craft. As a consequence, there are no hydrofoils larger than about 1000 tons.

The second challenge is that the hydrofoil is a one-speed craft. The lift on a foil varies as the square of the speed. Increasing the speed by 20% will lead to a 44% increase in lift. It is very difficult to accommodate a lift variation this large by any system of foil control. As a consequence, there is generally a rather narrow band of speeds in which the boat is best operated. This one-speed characteristic presents challenges for certain missions, such as military operations, which may involve extended periods of time at low or moderate speeds during which the hydrofoil sits down off its foils on the basic hull structure. The hydrofoil is then supported by buoyancy and offers no advantage over a conventional hulled craft.

Significant modern hydrofoils. The following ships are representative of modern hydrofoils.

US Navy PHM Class. The U.S. Navy Patrol Hydrofoil, Missile (PHM) Class ships were developed as the result of a NATO cooperative program. Similar hydrofoils were put in service by Italy as part of this program, and related craft were built for Israel and France. These craft use fully submerged foils, the main one located aft and a secondary one located forward. The U.S. Navy craft (**Table 1**; **Fig. 1**) were retired from service in 1996–1997. *See* NAVAL SURFACE SHIP.

Boeing JetFoil. The USN PHM Class ships were built by the Boeing Company, which used the same technology to produce a line of commercial hydro-

TABLE 2. Characteristics of five types of Rodriquez hydrofoils

	RHS-70	RHS-150	RHS-160	RHS-200	MEC-1
Length, m (ft)	22 (72)	29 (95)	31 (102)	36 (118)	25 (82)
Breadth, m (ft)	7.85 (25.75)	5.85 (19.2)	6.2 (20.3)	7.0 (23.0)	8.4 (27.6)
Weight, metric tons (long tons)	31.5 (31.0)	65.6 (64.6)	85 (84)	120 (118)	47.7 (46.8)
Passengers	69	180	205	300	146
Cruise speed, knots (km/h)	32.5 (60.2)	32.5 (60.2)	38 (70)	35 (65)	36 (67)

TABLE 3. Characteristics of Russian Katran Class hydrofoils

Length	34.5 m	113.2 ft
Breadth	10.3 m	33.8 ft
Weight	72 metric tons	71 long tons
Passengers	147	
Cruise speed	35 knots	65 km/h

Fig. 4. Russian hydrofoil in Katran Class. (*Volga Shipyard*)

foil ferries. The Boeing JetFoil ferry (**Fig. 2**) was produced in a small variety of sizes and capacities, and several units were sold commercially. The majority of them are in service as ferries in Hong Kong. In the late 1980s, Boeing sold the rights and technology for the JetFoil. JetFoils are now available from Kawasaki Jetfoil. *See* FERRY.

Rodriquez ferries. The Boeing JetFoil and the USN PHM use fully submerged foils. Since the 1950s, the Italian shipbuilder Rodriquez Cantieri Navali SpA has been building commercial hydrofoils of the surface-piercing type. The earliest Rodriquez ferries were the PT-20 type, which have all but disappeared from service. These were followed by a range of craft up to 40 m (130 ft) in length (**Table 2**; **Fig. 3**). Rodriquez continues to make incremental improvement and developments in their hydrofoil product line.

Russian hydrofoils. The former Soviet Union sponsored development of several hydrofoils for civilian river transport. These craft were all of the surface-piercing type. A few have been successfully exported for commercial operation, such as the Katran Class (**Table 3**; **Fig. 4**).

Buoyant hydrofoils. Since the mid-1990s, there has been some research on buoyant hydrofoils, also called lifting bodies. These foils are very thick and are intended to provide substantial hydrostatic lift in addition to their hydrodynamic lift. The purpose of these foils is to marry the speed potential of the hydrofoil with the seakindliness of the SWATH (small waterplane area, twin hull) ship concept. These concepts, however, are still developmental. Milestones in this development include the MID-FOIL craft and the HDV-100, both developed by Navatek Ships in Honolulu, and the QUEST HYSWAS developed by Maritime Applied Physics Corporation.

Chris B. McKesson

Bibliography. *Fast Ferry International* (formerly *Hovering Craft and Hydrofoil*), Fast Ferry Information Ltd., Kent, U.K.; *Ferry Technology*, Riviera Maritime Media Ltd, Enfield, Middlesex, U.K.; Capt. R. Johnson (ed.), Hydrofoils, *Naval Eng. J.*, pp. 142–199, February 1985; E. V. Lewis (ed.), *Principles of Naval Architecture*, 2d ed., 3 vols., Society of Naval Architects and Marine Engineers, Jersey City, NJ, 1988.

Hydroformylation

An aldehyde synthesis process that falls under the general classification of a Fischer-Tropsch reaction but is distinguished by the addition of an olefin feed along with the characteristic carbon monoxide and hydrogen. In the oxo process for alcohol manufacture, hydroformylation of olefins to aldehydes is the first step. The second step is the hydrogenation of the aldehydes to alcohols. At times the term "oxo process" is used in reference to the hydroformylation step alone. In the hydroformylation step, olefin, carbon monoxide, and hydrogen are reacted over a cobalt catalyst to produce an aldehyde which has one more carbon atom than the feed olefin. The olefin conversion takes place by the addition of a formyl group (CHO) and a hydrogen atom across the double bond. This is represented by reaction (1).

$$R-CH=CH_2 + CO + H_2 \xrightarrow{Co} R-CH_2-CH_2-CHO \quad (1)$$
$$R-CH=CH_2 + CO + H_2 \xrightarrow{Co} R-CH(CHO)-CH_3$$

See FISCHER-TROPSCH PROCESS.

The aldehyde is then treated with hydrogen to form the alcohol. In commercial operations, the hydrogenation step is usually performed immediately after the hydroformylation step in an integrated system.

A wide range of carbon number olefins, C_2–C_{16}, have been used as feeds. Propylene, heptene, and nonene are frequently used as feedstocks to produce normal and isobutyl alcohol, isooctyl alcohol, and primary decyl alcohol, respectively. Feed streams to oxo units may be single-carbon-number or mixed-carbon-number olefins.

The lower-carbon-number alcohols such as butanols are used primarily as solvents, while the higher-carbon-number alcohols go into the manufacture of plasticizers, detergents (surfactants), and lubricants.

Reactants. Reactants are CO, H_2, and olefins. The H_2 and CO are usually fed in a 1-to-1 ratio, as synthesis gas from methanol conversion. The hydroformylation takes place in the liquid phase. If the olefins are normally gaseous at the reactor conditions, a heavier liquid solvent is used as a suspension medium. The synthesis gas is usually fed in considerable excess of the required stoichiometric amount.

The reaction mass is contacted with the cobalt catalyst either by flowing the liquid reaction mass over a

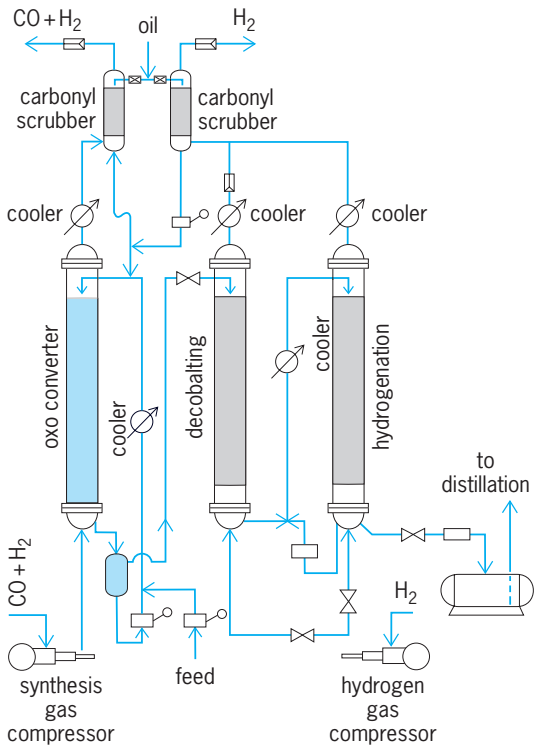

Fig. 1. Oxo process with fixed catalyst bed.

fixed bed of supported cobalt catalyst or by pumping a cobalt catalyst slurry into the liquid reaction mass in a continuous stirred tank or plug flow reactor. In the latter case, the catalyst is carried along with the reaction mass in what is frequently called the slurry system.

The olefin reactants are fed in blocked operation as either single-carbon-number C_7, C_8, or C_9 or in multiple-carbon-number feeds, such as C_{11}–C_{13}. For straight-chain olefins the double bond can be either in the terminal position or internal position. The product distribution is approximately the same in either case, roughly 60–40% normal alcohol and 40–60% alpha branched alcohol. Although the product distribution is nearly the same, the reaction rate when the double bond is in the terminal position is almost 3.5 times faster than the rate with double bond in an internal position.

The reaction rate has been found to be directly proportional to the olefin concentration, the cobalt concentration, and the hydrogen pressure, and inversely proportional to carbon monoxide pressure. The proposed mechanism for the reaction begins with the formation of an olefin-carbonyl complex and carbon monoxide from the reaction of dicobalt octacarbonyl with olefin, as shown in reaction (2). The complex

$$Co_2(CO)_8 + RCH{=}CH_2 \longrightarrow Co_2(CO)_7RCH{=}CH_2 + CO \quad (2)$$

decomposes through a reaction with a hydrogen or cobalt hydrocarbonyl to form the aldehyde and a precursor of dicobalt octacarbonyl. The heat of reaction of the hydroformylation step is about 30 kcal/mole (125,500 joules/mole). The capability of removing the heat of reaction is often the limiting factor in the reactor capacity.

Commercial operations. Two general types of liquid-phase process have been used. The first process employs a fixed bed of cobalt catalyst with the liquid reactant mixture flowing past the catalyst. The second type of process, the slurry type, carries the catalyst along with the liquid reactant. The slurry type of reactors can be either the mechanically agitated, stirred- tank type or the gas-sparged type. A stirred-tank type with mechanical agitation can be used when the operating pressure is not too high. However, many operations are run at high pressure, 100–300 atm (1–3 × 10^6 pascals). For these processes, a gas-sparged slurry-type reactor can be designed to use the gas rising through the liquid to provide mixing and thus eliminate the mechanical agitator and seals.

The fixed-bed process is shown schematically in **Fig. 1**. In this process, soluble cobalt salts of fatty acids or naphthenates are pumped with the olefin to the top of the first reactor and flow countercurrent to the synthesis gas. One type of fixed-bed catalyst consists of 2% metallic cobalt on a pumice carrier. Part of the cobalt is converted to carbonyl, leaves the reactor with the overhead product, and is replaced by cobalt salts in the feed. Unreacted synthesis gas leaving the top of the reactor is cooled, passed through a packed tower countercurrent to the olefin feed to remove cobalt carbonyl, and recycled to the reactor.

The second vessel is a decobalting converter in which cobalt carbonyl, dissolved in the product from the first reactor, is decomposed by treatment with hydrogen at about 200–220 atm (2–2.2 × 10^6 Pa) and 257–300°F (125–150°C). The liquid enters at the top and flows countercurrent to the hydrogen. Metallic cobalt is deposited on the packing. Gas flows from the top of the decobalting unit to the carbonyl scrubber, and the liquid leaving the bottom is sent to the hydrogenation reactor for conversion to alcohol.

The slurry type of process shown schematically in **Fig. 2** begins with cobalt oxide being mixed with

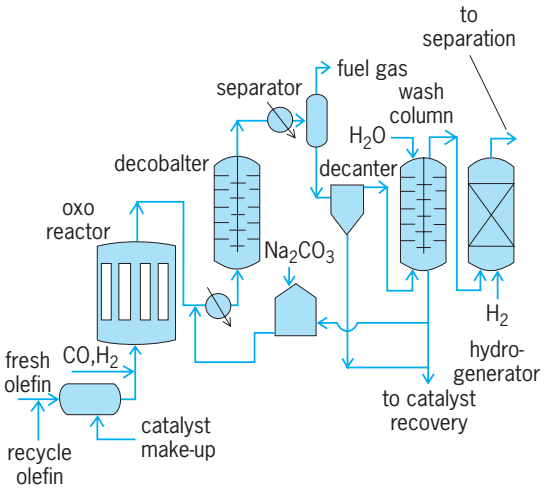

Fig. 2. Slurry process, in which the catalyst is carried with the liquid reactant.

recycle olefins. This slurry is then combined with the main olefin feed stream to give a 3–5 wt % cobalt slurry. This reaction slurry is fed into the bottom of the first reactor along with the synthesis gas. This process may employ up to five reactors in series. In the reactors, the catalyst reacts with the synthesis gas to form cobalt hydrocarbonyl, which is a gas at the reactor conditions of 356°F (180°C) and 220 atm (2.2×10^6 Pa). The liquid-gas mixture exits the last reactor of the series, and an aqueous solution of sodium carbonate is injected into this stream to form the water-soluble sodium cobalt carbonylate. The aqueous phase is separated from the organic stream in the decanter, and the organic stream is then water-washed to remove the last traces of catalyst. The washed aldehyde stream is then fed to the hydrogenator from which the product alcohol goes to the separation train downstream. The catalyst is recovered from the wash water and recycled to the process.

D. L. Holt

Hydrogen

The first chemical element in the periodic system. Under ordinary conditions it is a colorless, odorless, tasteless gas composed of diatomic molecules, H_2. The hydrogen atom, symbol H, consists of a nucleus of unit positive charge and a single electron. It has atomic number 1 and an atomic weight of 1.00797. The element is a major constituent of water and all organic matter, and is widely distributed not only on the Earth but throughout the universe. There are three isotopes of hydrogen: protium, mass 1, makes up 99.98% of the natural element; deuterium, mass 2, makes up about 0.02%; and tritium, mass 3, occurs in extremely small amounts in nature but may be produced artificially by various nuclear reactions. *See* DEUTERIUM; ISOTOPE; PERIODIC TABLE; TRITIUM.

Uses. The largest single use of hydrogen is in the synthesis of ammonia. A rapidly expanding use for hydrogen is in petroleum-refining operations, such as hydrocracking and hydrogen treatment for removal of sulfur. Large quantities of hydrogen are consumed in the catalytic hydrogenation of unsaturated liquid vegetable oils to make solid fats. Hydrogenation is used in the manufacture of organic chemicals. Large quantities of hydrogen are used as a rocket fuel, in conjunction with oxygen or fluorine, and as a propellent for nuclear-powered rockets.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Properties of hydrogen

Property	Value
Melting point	−259.2°C
Boiling point at 1 atm	−252.8°C
Density of solid at −259.2°C	0.0866 g/cm^3
Density of liquid at −252.8°C	0.0708 g/cm^3
Critical temperature	−240.0°C
Critical pressure	13.0 atm
Critical density	0.0301 g/cm^3

Properties. Ordinary hydrogen has a molecular weight of 2.01594. The gas has a density at 0°C and 1 atm of 0.08987 g/liter. Its specific gravity, compared to air, is 0.0695. Hydrogen is the lightest substance known. Some additional properties of hydrogen are listed in the **table**.

Hydrogen is somewhat more soluble in organic solvents than in water. Many metals adsorb hydrogen. The adsorption of hydrogen in steel may cause "hydrogen embrittlement," which sometimes leads to the failure of chemical processing equipment.

At ordinary temperatures hydrogen is a comparatively unre-active substance unless it has been activated in some manner, for example, by a suitable catalyst. At elevated temperatures it is highly reactive.

Although ordinarily diatomic, molecular hydrogen dissociates at high temperatures into free atoms. Atomic hydrogen is a powerful reducing agent, even at room temperature. It reacts with the oxides and chlorides of many metals, including silver, copper, lead, bismuth, and mercury, to produce the free metals. It reduces some salts, such as nitrates, nitrites, and cyanides of sodium and potassium, to the metallic state. It reacts with a number of elements, both metals and nonmetals, to yield hydrides such as NaH, KH, H_2S, and PH_3. With oxygen atomic hydrogen yields hydrogen peroxide, H_2O_2. With organic compounds atomic hydrogen reacts to produce a complex mixture of products. With ethylene, C_2H_4, for example, the products include ethane, C_2H_6, and butane, C_4H_{10}. The heat liberated when hydrogen atoms recombine to form hydrogen molecules is used to obtain very high temperatures in atomic hydrogen welding.

Hydrogen reacts with oxygen to form water. At room temperature this reaction is immeasurably slow, but is accelerated by catalysts, such as platinum, or by an electric spark, and then may take place with explosive violence.

With nitrogen, hydrogen undergoes an important reaction to give ammonia. Hydrogen reacts at elevated temperatures with a number of metals to give hydrides. The oxides of many metals are reduced by hydrogen at elevated temperatures either to the free metal or to lower oxides. Hydrogen reacts at room temperature with the salts of the less electropositive metals and reduces them to the metallic state. In the

presence of a suitable catalyst hydrogen reacts with unsaturated organic compounds and adds to the double bond. *See* HYDROGENATION.

Principal compounds. Hydrogen is a constituent of a very large number of compounds containing one or more other elements. Such compounds include water, acids, bases, most organic compounds, and many minerals. Compounds in which hydrogen is combined with a single other element are commonly referred to as hydrides. For additional details on the compounds of hydrogen *see* ACID AND BASE; HYDRAZINE; HYDRIDE; HYDRIDO COMPLEXES; HYDROGEN FLUORIDE; HYDROGEN PEROXIDE; WATER.

Preparation. A large number of methods may be used to prepare hydrogen gas. The choice of method is determined by such factors as the quantity of hydrogen desired, the purity required, and the availability and cost of raw materials. Among the processes frequently used are the reactions of metals with water or acids, the electrolysis of water, the reaction of steam with hydrocarbons or other organic materials, and the thermal decomposition of hydrocarbons. The principal raw materials for hydrogen production are hydrocarbons, such as natural gas, oil refinery gas, gasoline, fuel oil, and crude oil. Louis Kaplan

Bibliography. G. F. Bassani, M. Inguscio, and T. W. Hansch (eds.), *The Hydrogen Atom*, 1989; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; J. S. Rigden, *Hydrogen: The Essential Element*, 2003.

Hydrogen bomb

A device in which an uncontrolled, self-sustaining, thermonuclear fusion reaction is carried out in heavy hydrogen (deuterium or tritium) to produce an explosion. In a fusion reaction, the collision of two energy-rich nuclei results in a mutual rearrangement of their protons and neutrons to produce two or more reaction products, together with a release of energy of amount E given by A. Einstein's formula $E = mc^2$, where m is the mass difference between the original and produced nuclei, and c is the velocity of light. *See* NUCLEAR FUSION; THERMONUCLEAR REACTION.

For the hydrogen bomb reaction to become self-sustaining, a so-called critical temperature of about 3.5×10^7 K (6.3×10^7°F) must be attained with the aid of the enormous temperature created by a fission explosive. Once this temperature is achieved, the energy released in the initial reaction maintains the temperature, and the chain proceeds either until the supply of fusionable material is exhausted or until sufficient expansion has taken place that the material is cooled below the critical temperature. When the isotopes fuse, the result is a release of energy and radiation. Some of the radiation consists of fast neutrons, whose role is discussed below. *See* ATOMIC BOMB.

Fusion is a more difficult process than fission to start and propagate because of the high temperatures required—temperatures comparable to those at the center of stars (or in nuclear fission explosions). Also, deuterium and tritium can be produced only in highly specialized facilities. Although a fission explosion requires special isotopes carefully arranged in a weapons configuration, a fusion explosive must have separated isotopes of different light elements, and is much more complex to design and build. There are two ways to use fusion: boosting of fusion explosive yields or generating multistage thermonuclear reactions.

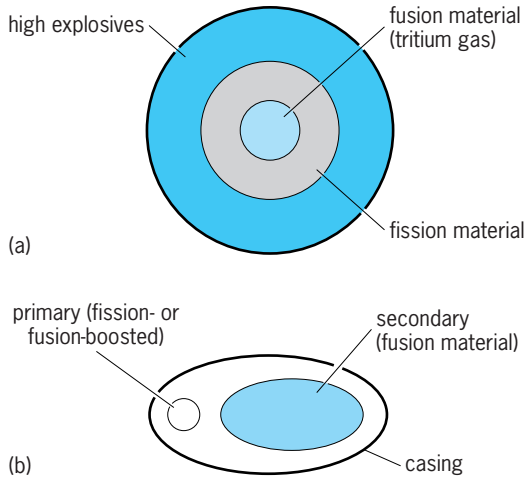

Design configuration of two types of fusion weapons: (*a*) fusion-boosted atomic bomb, (*b*) multistage hydrogen bomb. (*After A. DeVolpi et al., Born Secret: The H-Bomb, the Progressive Case and National Security, Pergamon, 1981*)

Fusion-boosted weapons. In a fusion-boosted warhead (**illus.** *a*), when the sphere of fissile materials is compressed (imploded) by the chemical explosion, an uncontrolled fission chain reaction begins. The fissionable material rapidly (in tenths of a millionth of a second) gets as hot as the center of the Sun. If there is fusionable material inside the device, thermonuclear reactions will boost the fission yield. This type of weapon is called fusion-boosted because the fusion reactions do not directly contribute very much to the explosive energy, but instead enhance the fission rate, due to the release of a large number of additional neutrons. Fusion-boosted weapons are militarily more desirable than pure-fission weapons because they are generally lighter, more efficient, and more powerful. *See* ATOMIC BOMB.

Multistage weapons. Multistage thermonuclear weapons are conceptually quite different from fission and fusion-boosted devices. They contain three essential components, which are physically separated from each other (illus. *b*). One component is a small fission or fusion-boosted explosive called a primary or trigger. Separated from the primary is an assembly of lithium-deuteride fusion material called the secondary. Surrounding the primary and secondary is the third major component, a massive casing. Before thermonuclear ignition, neutrons from the exploding primary convert some of the lithium deuteride in the secondary to a fusionable mixture of deuterium and tritium.

Multistage thermonuclear explosive detonation is basically generated in three major phases: ignition caused by the fission-explosive primary stage, coupling of x-radiation to the physically separated secondary fusion stage, and secondary-stage implosion induced by fission x-rays. The radiation-induced implosion of the secondary compresses the lithium-deuteride.

As neutrons from the primary traverse the compressed lithium-deuteride compound, they are most likely to be absorbed in enriched lithium-6, which immediately decomposes into tritium and an alpha particle. If the tritium thus created fuses with deuterium, an energetic alpha particle and a 14-MeV neutron are released. Alpha particles heat the material when they lose kinetic energy.

The fission triggers are themselves highly sophisticated, efficient nuclear explosive devices. The type, chemical form, isotopic content, density, and geometrical arrangement of the fissile materials are important. The exploding trigger emits a great deal of energy as photons, in blackbody radiation. At the temperatures present (about 5×10^7 K or 9×10^{7}°F), much of that energy is in the soft x-ray region. The photons fill the inside of the casing, traveling hundreds of times faster than the material portions of the exploding primary and other parts of the bomb. The casing behaves at first like a bottle, for a time keeping the energy confined.

There is so much energy released that very large compressive forces are exerted on the secondary, offset by expansive forces on the casing, caused by the shock wave from the primary while it disassembles. Because the casing is massive relative to the fusion package, it moves slowly compared with the rate of compression of the fusion materials. Soon those materials reach densities and temperatures where thermonuclear ignition occurs, liberating many times more energy than that which came originally from the trigger. *See* HEAT RADIATION.

Neutrons from the thermonuclear reactions escape in large numbers from the fusing materials and strike the nuclei in the casing. If the massive casing is made mostly of uranium-238 (natural or depleted uranium), the fusion neutrons will cause the uranium nuclei to undergo fission, giving off still more energy. A device of this sort can be regarded as a three-stage fission-fusion-fission bomb.

Yield. The yield, or total energy, of a hydrogen bomb is expressed in megatons (1 megaton equals 10^{15} calories or 4.18×10^{15} joules). Typical fusion-boosted weapons yield hundreds of kilotons (tenths of megatons), and typical multistage weapons yield megatons. The Soviet Union exploded a bomb with a yield of over 50 megatons in 1961; it is now known that this bomb was designed for a total of 150 megatons. Because a larger proportion of its yield is in radiation than in blast, the neutron-warhead type of thermonuclear weapon can be deployed for use with field artillery. *See* NUCLEAR EXPLOSION.

Complexity. Multistage hydrogen (fusion) bombs are far more complex than the fission weapon or the booster type of fusion weapon. Even with massive national resources, it has required an average of 5 years to develop such weapons after testing a fission explosive. A. DeVolpi

Bibliography. A. DeVolpi et al., *Born Secret: The H-Bomb, the Progressive Case and National Security*, 1981; C. Hansen, *U.S. Nuclear Weapons: The Secret History*, 1988; H. York, *The Advisors: Oppenheimer, Teller, and the Superbomb*, 1976, reprint 1989.

Hydrogen bond

The interaction which occurs when a hydrogen atom, covalently bonded to an electronegative atom (as in A—H), interacts with another atom to form the aggregate A—H···Y. The shortest and strongest bond is indicated as A—H, while the secondary and weaker interaction is written as H···Y. Thus A—H is a proton donor, while (Y) is a proton acceptor which often contains lone pair electrons and can act as a base. The strongest hydrogen bonds are formed between the most electronegative (A) atoms such as fluorine, nitrogen, and oxygen which interact with (Y) atoms having electronegativity greater than that of hydrogen (C, N, O, S, Se, F, Cl, Br, I). The weakest of hydrogen bonds are formed by acidic protons of C—H groups, as in chloroform and acetylene, and by olefinic and aromatic π-electrons acting as (Y).

Bond energies. The majority of hydrogen bonds have energies in the range 4–6 kcal/mole (17–25 kilojoules/mole) and involve those between O—H functional groups (as in water, alcohols, or acids) or N—H groups (as in amides or amines) and oxygen atoms (as in water, alcohols, carbonyls, or esters). The strongest hydrogen bond known is that found in the hydrogen difluoride ion, $(F—H—F)^-$, which has been variously estimated at 37–55 kcal/mole (155–230 kJ/mole). Therefore, the average hydrogen bond is of much lower energy than a normal chemical bond (>100 kcal/mole or 418 kJ/mole). Although hydrogen bonding gives rise to a specific interaction between atoms, resulting in a complex with characteristic A—H···Y distances and angles, especially in the solid state, it is difficult to establish a lower limit for the H-bond enthalpy because experimental methods of detection are becoming increasingly more sensitive and accurate.

The weaker the hydrogen bond, the shorter the lifetime of the complex it forms. The detection of weak hydrogen bonds often amounts to measuring shorter and shorter lifetimes of rapidly associating and dissociating species in equilibrium. This is a difficult problem because proton transfer along hydrogen bonds belongs to the fastest known chemical reactions, and in most experimental studies only mean values and "average" structures are determined.

An important aspect of weak hydrogen bond formation is that the different molecular aggregates which do form can be easily and reversibly transformed. Thus the small energy changes resulting in

the rapid making and breaking of hydrogen bonds in biological systems are of great importance; for example, hydrogen bonding determines the configuration of the famous α-helix, and the structures of most proteins, thereby serving an important function in determining the nature of all living things. *See* DEOXYRIBONUCLEIC ACID (DNA); PROTEIN.

Spectroscopy. Even though slight energy changes are usually involved in hydrogen bond formation, the aggregate once formed changes almost every measureable physical property of the original species. When investigating hydrogen bonding, the most frequently used techniques are infrared and nuclear magnetic resonance spectroscopy. In the infrared method the A—H stretching frequency is shifted to lower values and is accompanied by band broadening and increased intensity. Such changes are usually easily discernible, but in the case of very strong hydrogen bonds, such as $(F—H—F)^-$, the shift and broadening is so drastic that it is difficult to assign the new frequencies correctly. In the case of nuclear magnetic resonance, hydrogen bonding usually shifts the proton resonance to lower fields. *See* INFRARED SPECTROSCOPY; NUCLEAR MAGNETIC RESONANCE (NMR).

Neutron diffraction. While infrared and nuclear magnetic resonance techniques can yield considerable information about hydrogen bond formation, far greater information content is derived from diffraction studies of crystalline solids. The method of choice is neutron diffraction (versus x-ray diffraction), because in the former case hydrogen atoms scatter almost as well as any other atom while their scattering is often swamped out in the x-ray case. Neutron diffraction crystal structure analysis has become the best probe available for the study of the geometry of A—H···Y bonds. Using this technique, it is generally observed that the A—H separation is $\sim 0.10 \pm 0.01$ nanometer and is much less than the H···Y separation; that is, hydrogen bonds are usually asymmetric. In the extreme case the hydrogen atom may be equally bonded to both atoms, as in certain O—H—O and $(F—H—F)^-$ containing systems where the atomic environment around A and Y is identical and symmetric. In such systems strong hydrogen bond formation is indicated when the A···Y separation is ~0.02–0.03 nm less than the sum of the van der Waals radii. However, contrary to prior belief, even the shortest and strongest hydrogen bond known, $(F—H—F)^-$, may be asymmetric. This was demonstrated in a neutron diffraction study of *p*-toluidinium hydrogen difluoride in which the two terminal F atoms exist in vastly different F···H—N hydrogen bonding environments (see **illus.**). The F—H distances are unequal, and the $(F—H—F)^-$ ion is asymmetric, because of the very different N—H···F hydrogen bonding environments around the F atoms. Thus it seems that the H atom of a strong $(X—H—X)^-$ bond is a probe of the X atom environment. Indeed, the hydrogen dichloride ion, $(CHl_2)^-$, with a bond energy of about 12 kcal/mole (49 kJ/mole) appears to be symmetric (centered H atom) in some salts and asymmetric in others. Other types of hydrogen atom interactions such as those of M—H···M and C—H···M, where M is a metal, are becoming increasingly important in catalysis. *See* NEUTRON DIFFRACTION.

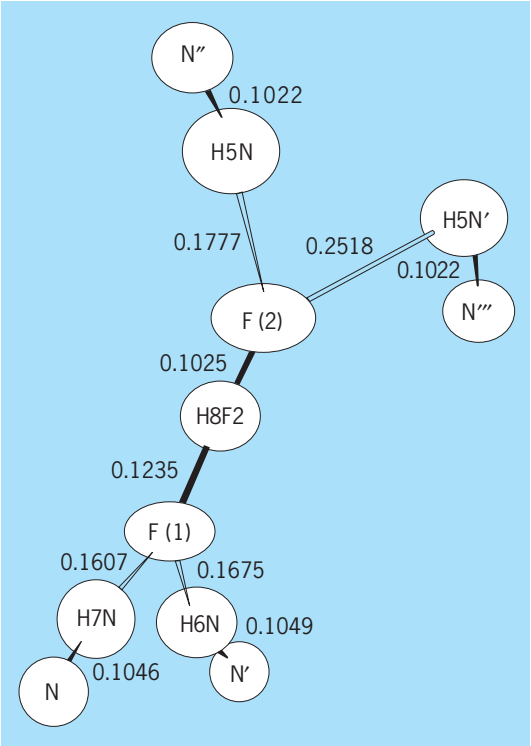

Hydrogen difluoride ion, $(F—H—F)^-$, geometry. Distances are in nanometers, and atoms are represented as ellipsoids of 50% probability. Numbers following symbols for the elements signify that like atoms are not equivalent structurally. The primes on N atoms indicate that they are structurally equivalent and related by symmetry operations.

Theory. Developments in theory have made it possible to better define certain contributions to hydrogen bond energies. The relative importance of forces of different origin (dispersion, polarization, exchange, coulomb, and so on) have become possible to estimate by using both molecular orbital methods and perturbation theory. In general, it appears that quantum theory gives reliable descriptions of isolated imers and trimers, but fails when dealing with large clusters of the type found in condensed phases.

Jack M. Williams

Bibliography. P. Schuster, G. Zundel, and C. Sardorfy (eds.), *The Hydrogen Bond: Recent Developments in Theory and Experiments*, vols. 1–3, 1976; M. J. Winter, *Chemical Bonding*, 1994.

Hydrogen fluoride

The hydride of fluorine and the first member of the family of halogen acids. Anhydrous hydrogen fluoride is a mobile, colorless liquid that fumes strongly in air. It has the empirical formula HF, melts at −83°C (−117°F) and boils at 19.8°C (67.6°F). The vapor is highly aggregated, and gaseous hydrogen fluoride deviates from perfect gas behavior to a greater extent

than any other gaseous substance known. Aggregate formation in both the vapor and liquid phase arises from unusually strong hydrogen-bond interactions. Hydrogen fluoride is prepared on the large industrial scale by treating fluorspar (calcium fluoride, CaF_2) with concentrated sulfuric acid. The crude product is purified by fractional distillation to yield a product containing more than 99.5% HF; the remaining impurities are principally water and small amounts of sulfur dioxide, silicon tetrafluoride, and boron trifluoride. Very dry hydrogen fluoride can be obtained by electrolysis or by treatment with reagents such as fluorine or cobaltic fluoride that react with water. *See* HYDROGEN BOND.

Properties. Anhydrous hydrogen fluoride is an extremely powerful acid, exceeded in this respect only by 100% sulfuric acid. Like water, hydrogen fluoride is a liquid of high dielectric constant that undergoes self-ionization and forms conducting solutions with many solutes. Because anhydrous hydrogen fluoride is a superacid, many organic solutes dissolve in it to form stable carbonium ions. Alkali metal fluorides and silver fluoride dissolve readily in hydrogen fluoride to form conducting solutions. The alkali metal fluorides are bases in the hydrogen fluoride system and correspond to solutions of alkali metal hydroxides in water. Conversely, antimony pentafluoride and boron trifluoride act as acids in hydrogen fluoride and accentuate the already strong acid properties of the solvent. *See* SUPERACID.

Anhydrous hydrogen fluoride dissolves a wide variety of organic compounds. Oxygen-, nitrogen-, and sulfur-containing compounds usually have high solubility in liquid hydrogen fluoride, generally higher than that found in water. Aromatic hydrocarbons are moderately soluble, and even saturated aliphatic hydrocarbons show appreciable solubility. Despite the fact that hydrogen fluoride is a strong dehydrating agent, many organic solutes can be recovered unchanged from hydrogen fluoride solution. Surprisingly, the enzymes trypsin and lysozyme survive dissolution and recovery from solution in anhydrous liquid hydrogen fluoride with full retention of biological activity.

Uses. Hydrogen fluoride is a widely used industrial chemical. It was formerly used in the petroleum refining industry for the isomerization of aliphatic hydrocarbons to form more desirable automotive fuels, but this application has been superseded by other methods. The largest industrial use of hydrogen fluoride is in making fluorine-containing refrigerants (Freons, Genetrons).

Another important use of hydrogen fluoride is in the preparation of organic fluorocarbon compounds by the Simons electrochemical process. In this procedure, an organic compound is dissolved in hydrogen fluoride, and an electric current is passed through the solution, whereupon the hydrogen atoms in the organic solute are replaced by fluorine. Hydrogen fluoride is employed in the electrochemical preparation of fluorine and for the preparation of inorganic fluorides. Thus, hydrogen fluoride is used for the conversion of uranium dioxide to uranium tetrafluoride, an intermediate in the preparation of uranium metal and uranium hexafluoride. With the great increase in nuclear energy–produced electricity, this represents an important use of hydrogen fluoride.

Important organic reactions may in some cases be performed to advantage in hydrogen fluoride solution, and nitration, sulfonation, diazotization, cyclization, and polymerization reactions have been carried out in this medium. Anhydrous hydrogen fluoride can be used for the carboxylation of olefins with carbon monoxide, for the alkylation of isoparaffins with olefins, for the esterification of fatty acids, and for the removal of protective groups in peptide preparation by solid-phase reactions.

Aqueous solutions of hydrogen fluoride (hydrofluoric acid) are relatively weakly acidic as compared to hydrochloric acid. Fluoride salts are formed by reaction of hydrofluoric acid with metal oxides and carbonates. Of particular importance is the rapid reaction of hydrofluoric acid or anhydrous hydrogen fluoride with silica, which leads to the application of these substances as etching agents for glass.

Both hydrogen fluoride and hydrofluoric acid cause unusually severe burns; appropriate precautions must be taken to prevent any contact of the skin or eyes with either the liquid or the vapor. *See* FLUORINE; HALOGENATED HYDROCARBON; ISOMERIZATION.

Joseph J. Katz

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; P. Henderson, *Inorganic Chemistry*, 1982; A. G. Sharpe, *Inorganic Chemistry*, 2d ed., 1987; H. Tsunoda and M. H. Yu (eds.), *Fluoride Research*, 1985.

Hydrogen ion

A proton combined with a number of water molecules. It is often written as H_3O^+ and called the hydronium ion. However, this species is best considered as an excess proton on a tetrahedral group of four water molecules and so would be designated as $H_9O_4^+$. For simplicity, it is most commonly written as H^+ (aq). *See* PROTON.

Chemical activity. Since it is formed by the self-ionization of water according to reaction (1), the hydrogen ion is present in all aqueous solutions. This formation also means that H^+(aq) is always found in the company of the hydroxide ion, OH^-(aq). The relationship between the concentrations of these two species is a very important property of water and at 25°C (77°F) is given by Eq. (2). The equilibrium constant K_w increases rapidly with increasing temperature, and for more precision the concentrations in Eq. (2) should be replaced by the activities of the ions. *See* IONIC EQUILIBRIUM.

$$H_2O \rightleftharpoons H^+(aq) + OH^-(aq) \qquad (1)$$

$$K_w = [H^+(aq)][OH^-(aq)] = 10^{-14} \qquad (2)$$

Reaction (1) and Eq. (2) indicate that the H^+(aq) and OH^-(aq) concentrations in pure water are equal

to each other with a value of 10^{-7} mole/liter. Any aqueous solution with this concentration of $H^+(aq)$ is called a neutral solution. If the $H^+(aq)$ concentration is greater than 10^{-7} mole/liter, the solution is called acidic. Basic solutions are those in which the $H^+(aq)$ concentration is less than 10^{-7} mole/liter. It is clear from Eq. (2) that as the $H^+(aq)$ concentration increases the $OH^-(aq)$ concentration must decrease, and vice versa. In the most straightforward system, acids are substances that can donate an $H^+(aq)$, and bases are substances that can accept one. Therefore, the action of common acids can be written, for example, as reaction (3) or (4), and the action of common bases can be exemplified by reaction (5).

$$HCl(aq) \xrightarrow{H_2O} H^+(aq) + Cl^-(aq) \quad (3)$$

$$H_2SO_4(aq) \xrightarrow{H_2O} H^+(aq) + HSO_4^-(aq) \quad (4)$$

$$NH_3(aq) + H^+(aq) \longrightarrow NH_4^+(aq) \quad (5)$$

Metal cations can also act as acids, particularly when they are small ions with high positive charges. An example of such a hydrolysis reaction is shown in reaction (6). The reaction of an acid and a base is called

$$Al^{3+}(aq) \rightleftharpoons AlOH^{2+}(aq) + H^+(aq) \quad (6)$$

a neutralization reaction and has great importance. A typical example of a neutralization is reaction (7).

$$HCl(aq) + NaOH(aq) \longrightarrow NaCl(aq) + H_2O \quad (7)$$

The products of such a neutralization are always a salt, in this case NaCl, and water. *See* ACID AND BASE.

The reaction of $H^+(aq)$ with bicarbonates, carbonates, sulfites, bisulfites, and sulfides produces the volatile gases carbon dioxide, sulfur dioxide, and hydrogen sulfide, respectively. The hydrogen ion also reacts with metals above hydrogen in the electromotive force series to produce hydrogen gas and the cation of the metal [reaction (8)].

$$Zn(s) + 2H^+(aq) \longrightarrow H_2(g) + Zn^{2+}(aq) \quad (8)$$

See ELECTROCHEMICAL SERIES.

Concentration. Hydrogen ion concentration determines the course of many chemical reactions that occur in living organisms and in the chemical industry. The control of hydrogen ion concentration is achieved in living organisms and in the laboratory by buffer systems. These are chemical mixtures designed to resist change in hydrogen ion concentration. Careful control of acid concentration is crucial in industries such as brewing, pharmaceutical manufacturing, electroplating, and textiles. Water sanitation, whether in a swimming pool or a water treatment plant, demands close attention to acidity. *See* BUFFERS (CHEMISTRY).

Electrical conductivity. Another property of the $H^+(aq)$ ion is important in both theoretical and practical ways. $H^+(aq)$ is the best conductor of electricity of any ion in aqueous solution. Its conductance at 25°C (77°F) is almost five times as large as the next-most-conducting ion. This abnormally large conductance is due to a unique mechanism of electricity conduction. Ordinary ions must travel physically through the solution to conduct electricity. The $H^+(aq)$ conducts electricity by a chain mechanism whereby an excess proton can be attached to one side of a water molecule cluster, and another excess proton lost from the opposite side of the same cluster. This high conductivity is of great practical importance in batteries such as the common lead-acid cell, where the sulfuric acid electrolyte keeps the internal cell resistance low. *See* ELECTROLYTIC CONDUCTANCE.

Relation between ion concentrations and pH in water

$[H^+(aq)]$	$[OH^-(aq)]$	pH	$[H^+(aq)]$	$[OH^-(aq)]$	pH
1	1×10^{-14}	0	1×10^{-7}	1×10^{-7}	7
1×10^{-1}	1×10^{-13}	1	1×10^{-8}	1×10^{-6}	8
1×10^{-2}	1×10^{-12}	2	1×10^{-9}	1×10^{-5}	9
1×10^{-3}	1×10^{-11}	3	1×10^{-10}	1×10^{-4}	10
1×10^{-4}	1×10^{-10}	4	1×10^{-11}	1×10^{-3}	11
1×10^{-5}	1×10^{-9}	5	1×10^{-12}	1×10^{-2}	12
1×10^{-6}	1×10^{-8}	6	1×10^{-14}	1	14

pH. The hydrogen ion concentration can vary over fourteen powers of 10. To avoid dealing with such exponentials, the concept of pH was advanced by S. P. L. Sørensen in 1909. It is defined by Eq. (9). In

$$pH \equiv -\log[H^+(aq)] \quad (9)$$

precise work the activity of the hydrogen ion would be used to replace the concentration in Eq. (9). Since all aqueous solutions contain both hydrogen ion and hydroxide ion, it is possible to define all degrees of acidity and basicity on the pH scale. The **table**, which gives hydrogen ion concentrations, hydroxide ion concentrations, and pH for aqueous solutions at 25°C (77°F), illustrates certain characteristics of the pH scale. Neutral solutions have a pH of 7, while acid solutions have a pH less than 7, and basic solutions have a pH greater than 7. A change of one pH unit corresponds to a tenfold change in acidity. The pH is also useful in pointing out the differences between strong acids such as hydrochloric acid and weak acids such as acetic acid, the main component of vinegar. For a strong acid, reaction (10) is almost 100% complete.

$$HCl(aq) + H_2O \longrightarrow H^+(aq) + Cl^-(aq) \quad (10)$$

Therefore, a 0.1 mole/liter solution of HCl has a pH of 1.0. For the weak acid, reaction (11) occurs to

$$CH_3COOH(aq) + H_2O \rightleftharpoons H^+(aq) + CH_3CO\overline{O}(aq) \quad (11)$$

a very limited extent. A 0.10 mole/liter solution of acetic acid has a pH of 2.9 and thus is almost 100 times less acidic than the hydrochloric acid solution. *See* PH.

Determining concentrations. Two general methods are used for the determination of hydrogen ion concentrations. For relatively crude work, colorimetric methods are commonly used. These methods depend on the fact that certain natural and synthetic

dyes have colors that depend on the hydrogen ion concentration. At times, paper is impregnated with such an indicator. By dipping the paper into the aqueous solution, the color of the paper changes due to the acidity of the solution. The paper can then be compared to a standard printed color series, and a rough idea of the pH obtained. In another application, a small amount of such an indicator dye can be added to a solution whose acidity is being determined. Small increments of a base solution are then added in a titration. If the indicator is chosen correctly, the solution will change color when the exact amount of base has been added to completely neutralize the original acid. A simple calculation will then give the amount of acid initially present. *See* ACID-BASE INDICATOR.

In most precise work, a potentiometric method is used for the determination of hydrogen ion concentration. This method depends on an electrode whose potential is sensitive to hydrogen ion concentration. The only electrode commonly in use for practical pH measurements is the glass electrode. A glass electrode, together with a reference electrode, is placed in the solution of unknown acidity. The potential of the electrochemical cell formed is then measured by a high-input-impedance voltmeter.

Many good commercial pH meters are available. Because it is difficult to relate the measured voltage directly to a hydrogen ion concentration, the pH meter–electrode combination is calibrated with buffer solutions of well-defined pH values. The calibrated instrument can then be used to measure the pH of unknown solutions. The importance of pH measurement led to the creation of very precisely defined pH standards at the National Bureau of Standards. It is now possible to make meaningful pH measurements, even in extreme conditions such as very high pressures (1000 bars or 100 megapascals) or very high temperatures (200°C or 390°F). The technology of pH measurement has advanced to the point where many industrial processes are automatically controlled through pH measurements on process streams. *See* ION-SELECTIVE MEMBRANES AND ELECTRODES; TITRATION.

G. Atkinson

Bibliography. J. O'M. Bockris and A. K. N. Reddy, *Modern Electrochemistry*, vol. 1, 1970; S. A. Borman (ed.), *Instrumentation in Analytical Chemistry*, 1982; G. Eisenman (ed.), *Glass Electrodes for Hydrogen and Other Cations*, 1967; G. K. McMillan, *pH Measurement and Control*, 2d ed., 1993; J. E. Ricci, *Hydrogen Ion Concentration*, 1952; L. Wilson, *Quantitative Analysis: Gravimetric, Volumetric and Instrumental Analysis,* 2d ed., 1989.

Hydrogen peroxide

A binary compound of hydrogen and oxygen, empirical formula H_2O_2, used mostly in dilute aqueous solutions as an oxidizing agent. It was discovered in 1818 by the French chemist Louis-Jacques Thenard, who named it *eau oxygénée*. Its most remarkable feature is its tendency to decompose readily into water and oxygen, the first observed instance of contact catalysis.

Properties. Anhydrous hydrogen peroxide is a clear, colorless liquid, of nearly the same viscosity and dielectric constant as water, but of greater density, 1.442 g · cm^{-3} at 25°C (77°F). Like water, it is strongly associated through hydrogen bonds. It boils at 150°C (302°F) with violent, sometimes explosive decomposition. With considerable supercooling, it can be frozen into needle-shaped crystals which melt at −0.41°C (31°F). The decomposition, strongly exothermic (690 cal · g^{-1} or 2.89 kJ · g^{-1}), is almost always catalytic, becoming homogeneous only in the vapor above 420°C (790°F). The rate varies considerably with the nature of the catalyst, the temperature, and the surface-volume ratio of the sample. Decomposition by light begins only in the near ultraviolet. As a solvent, hydrogen peroxide resembles water, except that acids and bases show much lower electrical conductivity. Although a fairly strong oxidant, it can act as a mild reducing agent, for example, with permanganates and perchromates.

The H_2O_2 molecule has the skew chain configuration shown in the **illustration**, the simplest example of internal rotation or torsion about a single bond. This is a semirigid structure, because the torsion is hindered by a rather low potential barrier (1 kcal · mol^{-1} or 4 kJ · mol^{-1} for the trans versus 7 kcal · mol^{-1} or 29 kJ · mol^{-1} for the cis configuration).

Formation. Hydrogen peroxide occurs only in traces in nature, mostly in rain and snow. It has not yet been detected in interstellar space. Its formation, usually in low concentration, has been studied in a variety of systems: (1) from the elements in the products of the oxyhydrogen flame quickly quenched in liquid air, or in explosion of hydrogen-rich mixtures near 550°C (1020°F); (2) from water, liquid or vapor, irradiated by ultraviolet light of a wavelength shorter than 185 nm, with and without sensitizer, and by ionizing radiation; and (3) from a silent electrical discharge through water vapor at atmospheric pressure, or under reduced pressure (below 10^{-3} atm or 10^2 pascals) in electrodeless discharges through a fast-flowing system, followed by quick chilling in a liquid-nitrogen trap. Under optimum conditions, concentrations up to 50% can be reached. This last

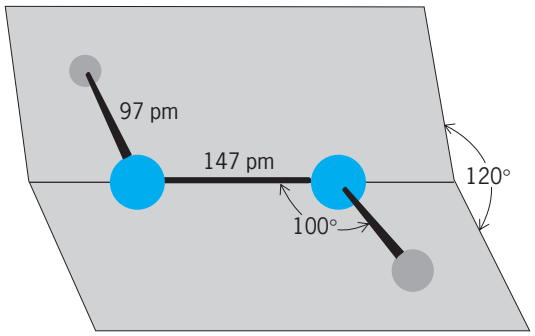

Structural parameters of the H_2O_2 molecule.

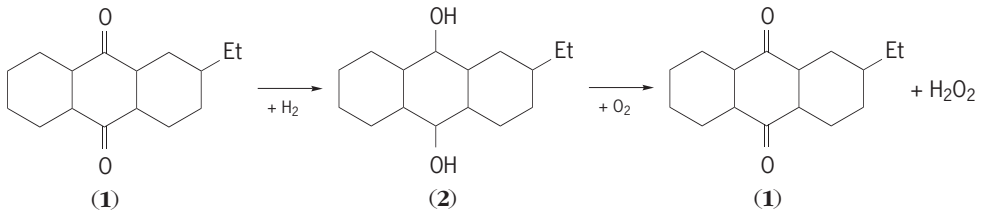

system, developed for the synthesis of deuterium peroxide, D_2O_2, from heavy water, has led to identification of the long-postulated higher oxides, H_2O_3 and H_2O_4, as metastable intermediates.

Manufacture. The original barium peroxide–sulfuric acid method was superseded long ago by the electrochemical process, which involves oxidation of sulfuric acid, or ammonium, or potassium sulfate in concentrated solutions at high current density on a platinum anode. This also is now largely replaced by the anthraquinone process based on the cyclic oxidation-reduction of some substituted anthraquinone as shown in the reaction above. The anthraquinone derivative (**1**) dissolved in an appropriate solvent is first reduced catalytically by hydrogen at atmospheric pressure. The resulting hydroquinone (**2**) is then oxidized by air, and the hydrogen peroxide (5 to 15 g per liter) is removed by countercurrent extraction with water, thereby regenerating the working substance. The efficiency of the latter step and the avoidance of side reactions are essential for economical operation.

Uses. Hydrogen peroxide is used for bleaching cotton and other fibers, natural or synthetic, and in the pulp and paper industry. Because of its milder action on fibers and the fact that it leaves no undesirable residue, it is preferable to chlorine and its compounds. Its cosmetic use as hair bleach consumes relatively little of the commercial 10% (30 volume) solution. In medicine it is useful for cleansing wounds and cuts, although its antiseptic action is rather slow. A limited but important use of the concentrated peroxide is for energy production in rockets, submarines (during submersion), airplanes (at takeoff), and the steering of space vessels. Fast decomposition is achieved by sudden addition of a catalyst or by blowing the vapor through a porous bed of catalyst-impregnated ceramics. The jet from a 90% solution can reach temperatures of about 750°C (1380°F). *See* ANTISEPTIC; BLEACHING; CHEMICAL FUEL; STERILIZATION.

Handling. Hydrogen peroxide, especially when concentrated, requires great care in handling and storing. When dropped on paper or wood, it can start a fire. Contact with the skin causes blotches that can be painful, but they disappear after a few hours without leaving traces. Appropriate containers are made of Pyrex glass, poly(tetrafluoroethylene), or polyethylene; for industrial storage and shipment, tanks or carboys are made of electropolished stainless steel or pure aluminum. All containers must be fitted with a vent for the escape of oxygen, and be of such design as to prevent spilling.

Slight decomposition may be overcome by distillation, or by addition of some stabilizer (such as zinc salts, phosphates, or ascorbic acid). For special purposes, a 90% product, stabilizer-free, is sometimes available. It can be concentrated to over 99% by fractional distillation in an all-Pyrex-glass still. The last traces of water are easily removed by fractional crystallization, provided that the solid is frozen very slowly into a large single crystal. *See* HYDROGEN; OXYGEN; PEROXIDE. Paul A. Giguère

Bibliography. W. C. Schumb, C. N. Satterfield, and R. L. Wentworth, *Hydrogen Peroxide*, ACS Monogr. 128, 1955; G. Strykul (ed.), *Catalytic Oxidations with Hydrogen Peroxide as an Oxidant*, 1993.

Hydrogenation

The chemical reaction of hydrogen with another substance, generally an unsaturated organic compound, and usually under the influence of temperature, pressure, and catalysts. There are several types of hydrogenation reactions. They include (1) the addition of hydrogen to reactive molecules; (2) the incorporation of hydrogen accompanied by cleavage of the starting molecules (hydrogenolysis); and (3) reactions in which isomerization, cyclization, and so on, result. Other reactions that involve molecular hydrogen and catalysts are reductive amination (hydroammonolysis) and hydroformylation (oxo reaction).

Hydrogenation is synonymous with reduction in which oxygen or some other element (most commonly nitrogen, sulfur, carbon, or halogen) is withdrawn from, or hydrogen is added to, a molecule. When hydrogenation is capable of producing the desired reduction product, it is generally the simplest and most efficient procedure.

Hydrogenation is used extensively in industrial processes. Important examples are the synthesis of methanol, liquid fuels, hydrogenated vegetable oils, fatty alcohols from the corresponding carboxylic acids, alcohols from aldehydes prepared by the aldol reaction, cyclohexanol and cyclohexane from phenol and benzene, respectively, and hexamethylenediamine for the synthesis of nylon from adiponitrile.

Hydrogenolysis. This term refers particularly to cleavages in a molecule associated with the addition of hydrogen. Hydrogenolysis is analogous to hydrolysis and ammonolysis, which involve cleavage of a bond induced by the action of water and ammonia, respectively. Chemical bonds which are broken by hydrogenolysis reactions include carbon-carbon, carbon-oxygen, carbon-sulfur, and carbon-nitrogen.

Two examples are hydrodealkylation of toluene to form benzene and methane, reaction (1), and re-

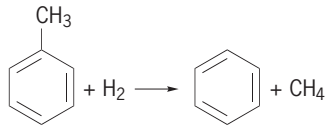

(1)

duction of methyl laurate to form lauryl alcohol and methanol, reaction (2).

$$CH_3(CH_2)_{10}COOCH_3 + H_2 \longrightarrow CH_3(CH_2)_{10}CH_2OH + CH_3OH \quad (2)$$

Catalytic hydrogenation. A variety of organic compounds can be hydrogenated easily in the presence of a catalyst. Acetylenes readily add two moles of hydrogen giving the saturated derivatives, as shown in reaction (3), where R and R′ are aliphatic, aromatic,

$$\begin{aligned} RC \equiv CR' + 2H_2 &\longrightarrow RCH = CHR' + H_2 \\ &\longrightarrow RCH_2CH_2R' \end{aligned} \quad (3)$$

or certain other groups. Under proper conditions the hydrogenation can be stopped at the intermediate olefin stage.

Catalytic hydrogenation of olefins can be carried out either in gas or in liquid phase, depending on their molecular weights. A nickel-containing catalyst and sometimes platinum or palladium catalysts are employed.

Aromatic compounds may be reduced either in the vapor phase at atmospheric pressure or in the liquid phase at hydrogen pressures up to 200 atm (2×10^4 kilopascals). In the latter case, aromatics, such as benzene, toluene, and *p*-cymene, can be hydrogenated readily in the presence of a nickel catalyst. In the case of naphthalene or substituted naphthalenes, the product may be the tetra- or decahydronaphthalene derivative.

The reduction of carbonyl compounds, such as aldehydes and ketones to the corresponding alcohols is represented by reaction (4), where R is an aliphatic

$$RCOR' + H_2 \rightarrow RCHOHR' \quad (4)$$

or aromatic group, and R′ may be the same group or a hydrogen atom. Frequently when R is an aromatic group, it is difficult to stop the reduction at the alcohol. Instead, it proceeds further to yield a hydrocarbon, RCH_2R'. In general, aldehydes are reduced more rapidly than ketones, although there are numerous examples in which both undergo reduction at room temperature and only a few atmospheres of hydrogen pressure. Often a small amount of water (1–10%) is added to the feed to the hydrogenator to suppress ether formation.

Other processes. Hardening of various animal fats and vegetable oils (such as soybean, cottonseed, fish, whale, and peanut) is carried out on a large scale by partial hydrogenation. The resultant plastic fats have a consistency and other properties suitable for the manufacture of shortenings, margarine, soaps, and a variety of other edible and industrial products. Chemically, the process involves the conversion of glycerides of unsaturated fatty acids (for example, oleic and linoleic) to saturated ones. Mild conditions (such as 212–1380°F or 100–750°C, 1–14 atm or $1–14 \times 10^2$ kPa and Ni catalysts) are employed to avoid the hydrogenolysis of the ester linkage. The conversion of olein to stearin may be expressed as reaction (5).

$$(C_{17}H_{33}COO)_3C_3H_5 + 3H_2 \longrightarrow (C_{17}H_{35}COO)_3C_3H_5 \quad (5)$$

Unlike the hardening of fats, which involves only the hydrogenation of ethylene linkages, the hydrogenolysis of the carboxyl group of acids and esters takes place with the formation of alcohols, as shown by reaction (6). The olefinic bonds in the fatty

$$RCOOH + 2H_2 \longrightarrow RCH_2OH + H_2O \quad (6)$$

chain may or may not be reduced. A reduced copper–ammonium chromate catalyst is used at 660–750°F (350–400°C).

The synthesis of methanol from carbon monoxide and hydrogen is carried out at high pressures (3000–5000 lb/in.^{2} or 20,600–34,500 kPa), because the reaction (7) involves a decrease in volume. The

$$CO + 2H_2 \rightarrow CH_3OH \quad (7)$$

practical reaction temperature range is small. Below 570°F (300°C) the rate is slow; above 750°F (400°C) the equilibrium becomes unfavorable. Mixed catalysts consisting of oxides of zinc, chromium, manganese, or aluminum, such as zinc oxide with 10% chromium oxide, are utilized. Carbon monoxide can also be hydrogenated to give various hydrocarbons and higher alcohols.

Petroleum, tar, and coal are hydrogenated to (1) improve existing products; (2) convert low-grade materials such as heavy oils into valuable fuels; and (3) transform solid fuels such as lignites and coal into liquid fuels. Several of these processes are hydrodesulfurization, hydrocracking, and hydrodealkylation. By a proper selection of catalysts and operating conditions, such hydrogenations can be directed to give desired end products and at the same time cause impurities that are common catalytic poisons, such as sulfur, nitrogen, and oxygen, to be detached from their molecular linkages and to be removed as hydrogen sulfide, ammonia, and water. *See* HYDROCRACKING.

Catalytic hydrogenation may continue to increase in importance, size, and the variety of processes. Examples are possible commercialization of coal-hydrogenation processes and the conversion of petroleum residues and shale oils to lighter fractions. *See* COAL GASIFICATION; COAL LIQUEFACTION.

Thermodynamics. Hydrogenation reactions are generally reversible. Catalysts affect the rate or speed of reaction, but have nothing to do with the inherent tendency of the reaction to proceed. To know whether or not the reaction is feasible, the free-energy change of a reaction, ΔG, can be

determined. For example, for the reaction represented by (8), the change in free-energy content per

$$CO(g) + 3H_2(g) \rightleftharpoons CH_4(g) + H_2O(g) \quad (8)$$

atom of carbon in each molecule at 750°F (400°C) is shown in Eqs. (9) and (10).

$$-40,500 + 0 + \Delta G = -3600 - 50,000 \quad (9)$$

$$G = -13,100 \text{ cal at } 400^\circ\text{C } (1 \text{ cal} = 4184 \text{ J}) \quad (10)$$

The decrease in free energy means that, at this temperature, reduction of carbon monoxide to methane is possible. At 1830°F (1000°C), the reverse reaction for the production of hydrogen and carbon monoxide from natural gas and steam is feasible. It so happens that both the forward and the reverse reactions are industrially important. *See* FREE ENERGY.

Hydrogenation reactions are exothermic, that is, heat is released during the reaction. Typically, the heat release per gram mole of hydrogenated material formed is about 28–30 kcal (117–126 kilojoules) for the hydrogenation of alkenes to alkanes and is 50 kcal (209 kJ) for the hydrogenation of benzene to cyclohexane. The heat of reaction must be removed by heat exchangers in the reactor or is utilized to heat the feeds to the reaction temperature.

Effect of temperature. The reaction temperature affects the rate and the extent of hydrogenation as it does any chemical reaction. Practically every hydrogenation reaction can be reversed by increasing temperature. High temperatures often lead to loss of selectivity and, therefore, yield of desired product, if a second functional group is present. As a practical measure, hydrogenation is carried out at as low a temperature as possible compatible with a satisfactory reaction rate. Although the optimum temperature depends on the catalyst type and age, the temperatures for hydrogenation reactions are generally below 930°F (500°C).

Effect of pressure. Hydrogenation rates are generally increased by increasing the hydrogen pressure. Pressure also increases the equilibrium yield in hydrogenations where there is a decrease in volume as the reaction proceeds. For economic reasons, many industrial hydrogenation processes are carried out under an imposed pressure but seldom above 300 atm (3×10^2 kPa).

Catalysts. For industrial applications, hydrogenation catalysts are generally solids consisting of metals, metal oxides, and some salts. These catalysts may be classified in accordance with their customary use. Vigorous catalysts suitable for the hydrogenation of alkyne and alkene linkages, aldehydes, and ketones include nickel and cobalt, and molybdenum and tungsten oxides or sulfides. Mild catalysts, useful for stepwise hydrogenations of aldehydes and ketones include oxides of copper, zinc, and chromium, and metallic platinum and palladium. Molybdenum sulfide and especially tungsten disulfide are active catalysts for operations at 3000 lb/in.^{2} (20,600 kPa). These catalysts are useful for the hydrogenation of unsaturates and to effect the cleavage of C-C, C-O, and C-N bonds. *See* CATALYSIS.

A wide range of metal ions and complexes has been found to catalyze hydrogenation reactions homogeneously in solution. These ions and complexes have been derived from a variety of metals, including platinum, cobalt, rhodium, and copper. Homogeneous catalytic systems are inherently simpler chemically and kinetically, and are often more selective. Judging from the patent activity in this area, the use of homogeneous catalysts to effect hydrogenation shows considerable promise. *See* HOMOGENEOUS CATALYSIS.

Equipment. There are two common types of reaction vessels. The first is used with liquids or solids, as in the hydrogenation of oils and viscous hydrocarbons. Internal agitators bring about an intimate mixing of organic compound, catalyst, and hydrogen. Alternatively, hydrogen is kept dispersed in the oil by recirculating the gas extracted from the head space of the reactor to the bottom by means of a blower. Usually, these are batch processes, although continuous mode of operation is becoming more attractive. The second type of reactor resembles a column or tube containing a fixed bed of catalyst, and is used where the organic compound has sufficient vapor pressure at the reaction temperature, as in the synthesis of methanol from carbon monoxide, to permit gas-phase, continuous operations.

The design and construction of process equipment which can withstand hydrogen gas at high temperature or high pressure, or both, are complicated. Alloy steels are the most common materials of construction. *See* DEHYDROGENATION; FISCHER-TROPSCH PROCESS; HIGH-PRESSURE PROCESSES; HYDROFORMYLATION; HYDROGEN; ORGANIC SYNTHESIS; OXIDATION-REDUCTION.

Roberto Lee

Bibliography. M. Freifelder, *Catalytic Hydrogenation in Organic Synthesis: Procedures and Commentary*, 1978; B. R. James, *Homogeneous Hydrogenation*, 1973; H. P. Patterson, *Hydrogenation of Fats and Oils*, 1983; P. N. Rylander, *Hydrogenation Methods,* 1985; M. J. Satriana (ed.), *Hydroprocessing Catalysts for Heavy Oil and Coal*, 1982.

Hydrography

The science of measuring and describing the physical features and conditions of navigable waters and adjoining coastal areas. Hydrography is an applied science involving the study of marine areas, including oceans, rivers, and lakes. It involves geodesy, physical oceanography, marine geology, geophysics, photogrammetry (in coastal areas), remote sensing, and marine cartography. Basic parameters observed during a hydrographic survey are time, geographic position, depth of water, and bottom type. However, observation, analysis, and prediction of tides and currents area are also normally included in order to reduce depth measurements to a common vertical datum. *See* GEODESY; PHOTOGRAMMETRY.

A principal objective of hydrography is to provide for safe navigation and protection of the marine

environment through the production of up-to-date nautical charts and related publications. In addition, hydrographic data are essential to a multitude of other activities such as global studies, for example, shoreline erosion and sediment transport studies; coastal construction; delimitation of maritime boundaries; environmental protection and pollution control; exploration and exploitation of marine resources, both living and nonliving; and development of marine geographic information systems (GIS). *See* GEOGRAPHIC INFORMATION SYSTEMS; NAVIGATION.

The International Hydrographic Organization (IHO), which is intergovernmental, coordinates the activities of the national hydrographic offices, develops sciences in the field of hydrography and physical oceanography, and achieves the greatest possible uniformity in nautical charts and documents.

Electronic position-fixing systems provide continuous availability of the necessary data, which previously were limited to line-of-sight observations. The electronic systems are usually classified in terms of their range capability, that is, long, medium, or short range. The long-range (Loran) systems have accuracies far offshore to a fraction of a nautical mile. Medium-range systems are typically comparison systems, extending 100–200 mi (160–320 km) offshore, and their accuracies are in the range of several tens of meters. Short-range systems are typically high-frequency, with the data converted to the elapsed time for a pulse to travel between two units into a measured distance; accuracy is in the range of meters. Distances measured from two known stations on land to the mobile unit (ship or launch) then provide the necessary position. Laser measurements provide highly accurate but short-range distance measurements (centimeters). *See* LORAN.

Satellite navigation. The U.S. Department of Defense satellites of the NAVSTAR GPS (Navigation System with Time and Ranging Global Positioning System) provide worldwide, continuous, and precise three-dimensional real-time positioning. Another global satellite system, GLONASS, is operated by Russia. Many satellite navigation receivers are compatible with both GPS and GLONASS. By using the differential technique, wherein a shore-based fixed unit is used to monitor variations of the satellite signal that are broadcast to the mobile unit to refine the ship%s position, positions can be obtained that are accurate to within less than 33 ft (10 m). Submeter accuracies are obtainable by monitoring the phase of the carrier signal from the satellites. This advance has revolutionized hydrography, making it possible to survey the world%s oceans with great accuracy. An added advantage is that satellite systems provide positions referenced with a common horizon datum rather than the regional datum of the past. *See* AIR NAVIGATION; MILITARY SATELLITES; SATELLITE NAVIGATION SYSTEMS.

Sonar. Modern depth information is achieved with sonar measurements. Dual-frequency echo sounders are used, with a high-frequency, narrow beam to measure the depth below the vessel, and a lower-frequency, wider beam to obtain larger coverage of the terrain. Ships used for offshore surveys and boats used for nearshore shallow-water surveys usually make observations along parallel sounding lines that run perpendicular to the depth contours, that is, usually perpendicular to the beach, which provides for better information for interpreting the bathymetry. The spacing of the lines depends mainly on the scale of the survey and the sea bottom morphology. The principal deficiency of this method is the possibility that a hazard may be undetected between the sounding lines.

Side-scan sonar, an instrument that transmits acoustic signals obliquely through the water, is normally towed behind the survey vessel and displays the returning echoes via an onboard graphic recorder. Although this technique does not allow exact determination of position and depth (both can be approximated), it provides excellent resolution with a depiction with what lies to either side of the vessel. As towing must be done at slower-than-normal sounding speeds and the area of coverage varies according to height of the towed sensing element (known as towfish) above the sea floor and the sonar frequency, this technique is limited in its speed of survey. Multibeam side-scan sonars provide for increased numbers of sonar returns, thereby allowing the vessel speed to be increased while still detecting small objects and providing improved depiction of the sea-floor texture.

Multibeam hydrographic survey systems consist of hull-mounted arrays such that a fan-shaped array of sound beams is transmitted perpendicular to the direction of the ship%s track. This provides for the possibility of 100% coverage of the sea floor. The resolution of the system depends on the size of the sonar beam (beam angle) and depth of the water. Since beams are transmitted at an angle through the water column, refraction becomes a problem. While all sonar systems need information about the structure of the water column so as to properly account for the speed of sound in seawater (affecting depth), with multibeam systems both depth and position of the depth measurement are affected. Also, with multibeam systems it is important to measure the attitude of the vessel (heave, roll-pitch, and yaw). Such systems cover a swath width of typically two to six times the water depth. *See* ECHO SOUNDER; SONAR.

Airborne systems. Laser airborne systems mounted in fixed-wing aircraft or helicopters are also available for hydrographic surveys. The system emits a two-color laser beam, usually green and red, such that a return is received from the surface of the water by the red laser and from the bottom by the lower-frequency green laser, allowing the depth to be determined from the time difference. They can be operated in depths down to 165 ft (50 m), but more normally to 66 ft (20 m), depending on water clarity. High turbidity can preclude use of this method. Resolution depends on the scattering of the laser beam as it passes through the water column, and it may not detect the shoal point of a hazard. The great advantage of this system is its speed of survey in coastal waters. It can survey an area at a rate of perhaps

60 times that of a traditional vessel and may be quickly deployed to different areas. *See* LASER.

Hydrographers use tide-coordinated aerial photography to delineate the high and low water lines for charting, which in turn is used for base-line determination of offshore boundaries. Satellite positioning of the aircraft using the Global Positioning System with carrier phase measurement and postprocessing of the data provides for determination of the position of the aircraft of the decimeter level, significantly reducing the need for a ground crew and geodetic control. Thus, metric information can be obtained in areas to which geodetic control has not or cannot (because of ice covering) be extended. *See* AERIAL PHOTOGRAPHY.

Digital data. Hydrographers provide suitable data to feed the Electronic Chart Display and Information System (ECDIS), which may be used instead of paper charts. The resolutions of the International Maritime Organization (IMO), taken from the approved standards of the International Hydrographic Organization, regulate the carriage requirements of electronic charts and their use. *See* MARINE GEOLOGY; OCEANOGRAPHY.

Christian Andreasen; Giuseppe Angrisano

Bibliography. Hydrographic Society, *Hydrographic Journal*, quarterly; A. E. Ingham, *Hydrography for the Surveyor and Engineer*, 3d ed., 1993; IMO Assembly resolutions A.817(19) and MSC 64(67), MSC 86(70), International Hydrographic Bureau, *International Hydrographic Bulletin*, monthly.

Hydroida

An order of the cnidarian which includes the freshwater hydras, the attached and usually colonial hydroids, and many of the smaller jellyfish. It is the largest order of the class Hydrozoa.

Taxonomy. The order Hydroida includes two principal suborders, Gymnoblastea and Calyptoblastea; the descriptions in this article apply mainly to them. The Gymnoblastea are those hydroids which lack protective cups around the hydranths and gonozoids. Jellyfish produced by these athecate hydroids are Anthomedusae. The Calyptoblastea include the hydroids with protective cups around the hydranths (hydrothecae) and around the gonozoids (gonothecae). Jellyfish of these thecate hydroids are called Leptomedusae.

Two minor suborders are Limnomedusae and Chondrophora. Limnomedusae include a few species of fresh-water jellyfish (such as *Craspedacusta*) which have small and simple polyp stages in their life history. The marine genus *Gonionemus* also belongs to this suborder. Chondrophora are animals of the open sea; the best known is *Velella*, the by-the-wind sailor or purple sail. Because they possess floats, they have sometimes been included with the Siphonophora. A direct comparison, however, may be made with the structure of a symnoblastic polyp whose stem region is used for attachment, but for the chondrophoran the stem region is modified to

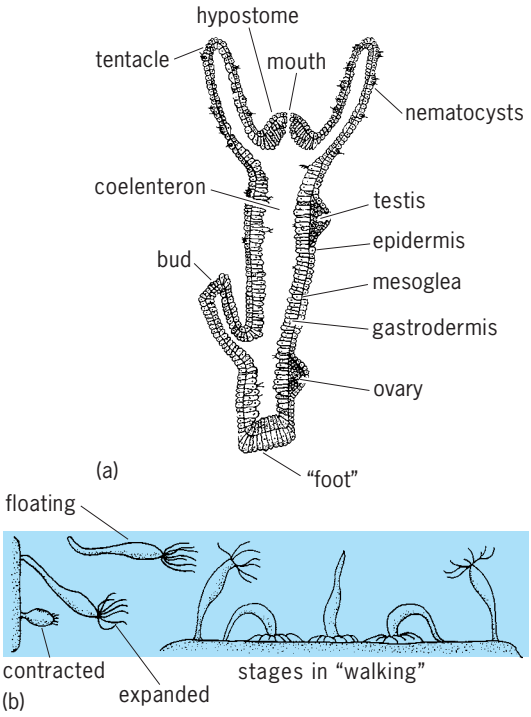

Fig. 1. Hydra. (*a*) Longitudinal section. (*b*) Movements. (*After T. I. Storer and R. L. Usinger, General Zoology, 4th ed., McGraw-Hill, 1965*)

become the float. The medusae which are produced by budding resemble anthomedusae.

Morphology. Anthomedusae are typically ovoid jellyfish, often with eyespots. Leptomedusae are usually flattened or saucer-shaped, have statocysts (sense organs of balance), and lack eyespots. The gonads of Anthomedusae are generally on the wall of the stomach just above the mouth, and those of Leptomedusae below the radial canals.

Young hydranths of gymnoblastic hydroids are small with five or more tentacles, but subsequently grow much larger and add more tentacles. Calyptoblastic hydranths, in contrast, emerge from a bud

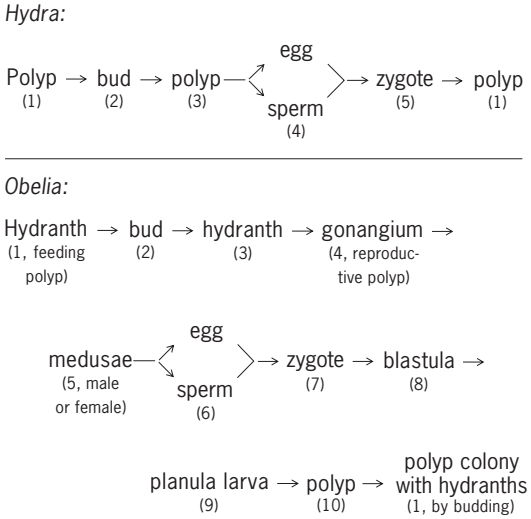

Fig. 2. Life cycle stages of *Hydra* and *Obelia*.

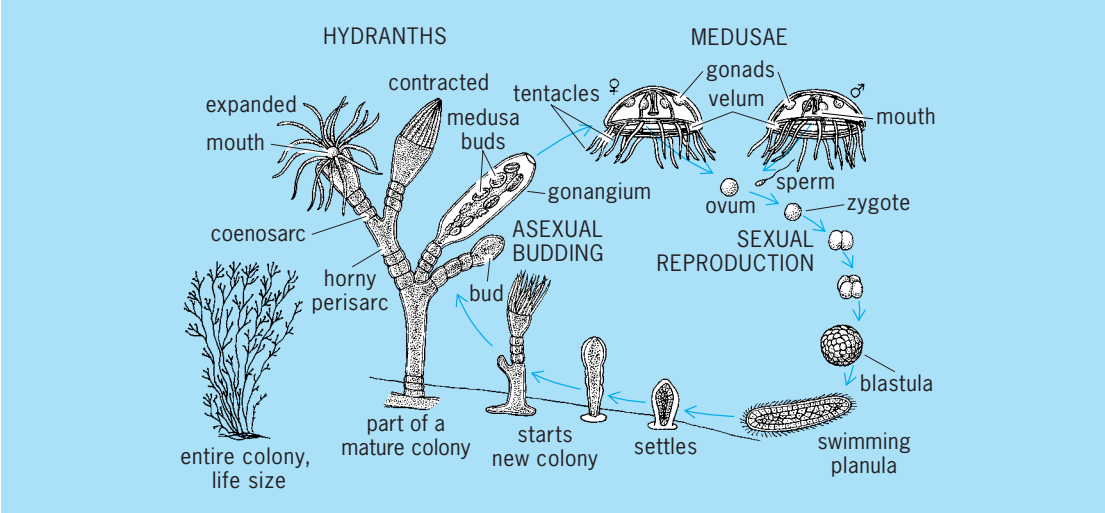

Fig. 3. Diagram illustrating the life cycle of *Obelia*, a polymorphic colony of polyps which has a medusoid state. (*After T. I. Storer and R. L. Usinger, General Zoology, 4th ed., McGraw-Hill, 1965*)

with a full complement of parts; they do not grow, live for only about a week, undergo regression and absorption by the colony, and are then replaced by new hydranths.

The fresh-water hydras are simple, motile polyps which do not produce colonies (**Fig. 1**). Buds separate from the parent and become individual polyps. Simple gonads develop on the body and there is no medusa. Hydras are sometimes included in the Gymnoblastea and sometimes placed in a suborder by themselves, the Hydrida.

Life cycle. The fertilized egg usually develops into a free-swimming ciliated larva, the planula, which soon attaches to some support and develops a mouth surrounded by tentacles at its free end. This attached stage is called a polyp and produces stolons, stems, and further polyps which remain connected to make up a hydroid or hydroid colony. Such colonies may show elaborate patterns of branching and may consist of hundreds of polyps, also called hydranths. Buds develop on the colony and are liberated as jellyfish or medusae which produce the sperm and eggs. *See* INVERTEBRATE EMBRYOLOGY.

Although the life cycle described for *Obelia geniculata* is regarded as typical, many species show reduction of either the polyp or the medusa stage. A comparison showing the differences in the life cycles of *Hydra* and *Obelia* is shown in **Fig. 2**. *Hydra*, a solitary polyp, has no medusoid stage, while *Obelia*, a polymorphic colony of small polyps, has a medusoid state (**Fig. 3**).

The colony. In a hydroid colony there is sometimes only one kind of polyp and the medusae are developed on the hydranth or stem. Often, however, there are special reproductive polyps, with partial or complete reduction of mouth and tentacles. Such reproductive polyps are called gonozooids or gonangia and may produce small medusae which are liberated to grow and reproduce sexually. Often the medusa is incompletely developed except for its gonads, and the gametes ripen while the meduosoid body remains attached to the parent hydroid.

In addition to ordinary nutritive hydranths and reproductive ones, some hydroids have other specialized types which are sensory, defensive, or aid in capturing food.

The stems of a colony are covered by a stiff secreted substance, the perisarc or periderm, which supports and protects it. The periderm may form cups around the hydranths and gonozooids.

Ecology. Hydroids are species in which the polyp stage is usually dominant. Most hydroids are found near the shore attached to various supports such as rocks, wharves, boats, mussels, barnacles, worm tubes, crab and snail shells, and seaweeds. Few occur on mud or sand. Sometimes the medusa stage is well developed and the hydroid stage may be lacking. Medusae are abundant both in coastal waters and in the open sea.

Some hydroids live in direct association with other animals. A few are parasitic on fish; others live inside clams and on snails, hermit crabs, and worm tubes. The larger animal is valuable to the hydroid because its activity produces a movement of water past the hydroid and thus suspended food is made available.

Use in research. Hydras and hydroids have been used extensively in research on problems of growth, development, and regeneration. They have a high capacity for reorganization. Missing parts are quickly replaced, a bit of stem can produce a new hydranth, and completely disorganized masses of cells can reconstitute a new polyp. *See* REGENERATIVE BIOLOGY.

Dried *Sertularia*, a large and delicately branched hydroid, is sold as decorative material known as white weed (air fern when dyed green). *See* HYDROZOA.

Sears Crowell

Bibliography. R. D. Barnes, *Invertebrate Zoology*, 6th ed., 1994; J. G. Engemann, *Invertebrate Zoology*, 3d ed., 1981; C. M. Fraser, *Hydroids of the Atlantic Coast of North America*, 1944; C. M. Fraser, *Hydroids of the Pacific Coast of Canada and the*

United States, 1937; L. H. Hyman, *The Invertebrates*, vol. 1, 1940; F. S. Russell, *The Medusae of the British Isles*, 1953.

Hydrology

The science dealing with all aspects of the waters of the Earth: their occurrence, circulation, and distribution; their chemical and physical properties; and their reaction with the environment, including their relation to living things.

Hydrologic cycle. Water in liquid and solid form covers most of the crust of the Earth. By a complex process powered by gravity and the action of solar energy, an endless exchange of water, in vapor, liquid, and solid forms, takes place between the atmosphere, the oceans, and the crust. Water circulates in the air and in the oceans, as well as over and below the surface of landmasses (**Figs. 1** and **2**). The distribution of water in the planet is uneven. General patterns of circulation are present in the atmosphere, the oceans, and the landmasses, but regional features are very irregular and seemingly random in detail. Therefore, while causal relations underlie the overall process, it is believed that important elements of chance affect local hydrological events. *See* ATMOSPHERIC GENERAL CIRCULATION.

Hydrologic studies. Water is essential for all living things. It also participates in the physical and geochemical evolution of most nonliving matter on Earth. Its adequate supply is a key factor for urban, agricultural, and industrial development. Water can also be the recipient of pollutants that degrade its quality for all uses; and it may be a destructive agent when it inundates valleys and causes death and great damage during floods. The rising of ground-water levels in agricultural lands may cause deterioration of the soils and loss of fertility by waterlogging and increased salinity. Erosion of soils by flowing waters, and the ultimate deposition of the sediment in lakes, reservoirs, stream channels, and harbors, are also serious problems. Thus, the means by which natural waters may be captured and controlled are of utmost importance for the development of human economy. The study of hydrology provides the information necessary for determining those means. *See* BIOSPHERE; EROSION; WATER POLLUTION.

Whereas the global linkages of the hydrologic cycle are recognized, the science of hydrology has traditionally confined its direct concern to the detailed study of the portion of the cycle limited by the physical boundaries of the land; thus, it has generally excluded specialized investigations of the ocean (which is the subject of the science of oceanography) and the atmosphere (which is the subject of the science of meteorology). The heightened interest in anthropogenically induced environmental impacts has, however, underlined the critical role of the hydrologic cycle in the global transport and budgeting of mass, heat, and energy. Hydrology has become recognized as a science concerned with processes at the local, regional, and global scales. This enhanced status has strengthened its links to meteorology, climatology, and oceanography. *See* CLIMATOLOGY; METEOROLOGY; OCEANOGRAPHY.

A number of field measurements are performed for hydrologic studies. Among them are the amount and intensity of precipitation; the quantities of water stored as snow and ice, and their changes in time; discharge of streams; rates and quantities of infiltration into the soil, and movement of soil moisture; rates of production from wells and changes in their water levels as indicators of ground-water storage;

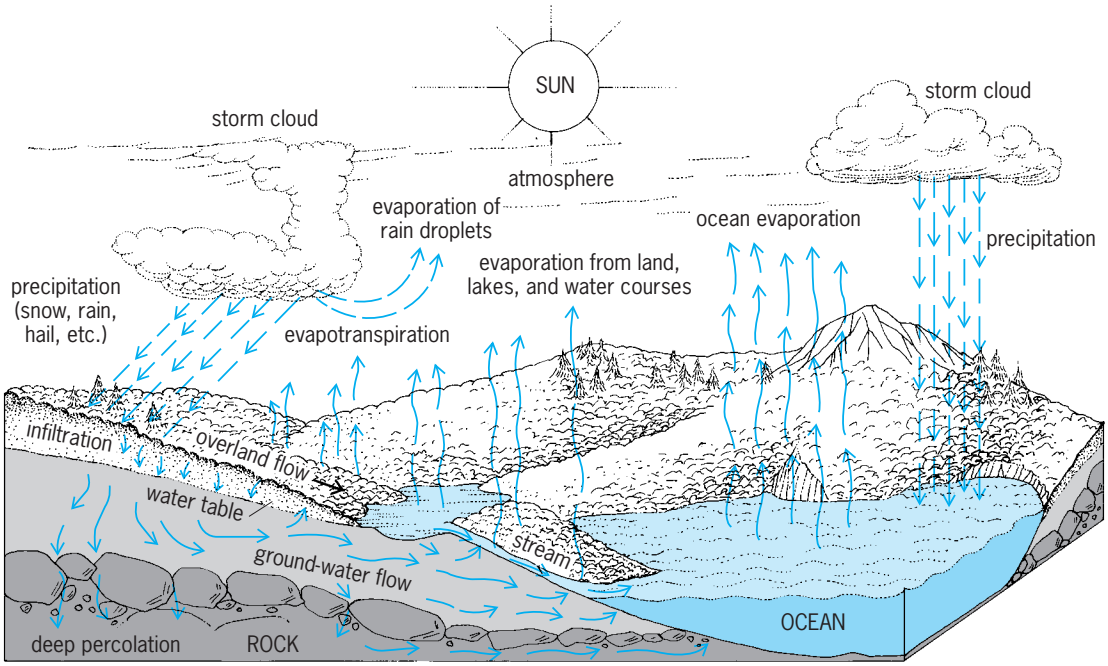

Fig. 1. Diagram of the hydrologic cycle.

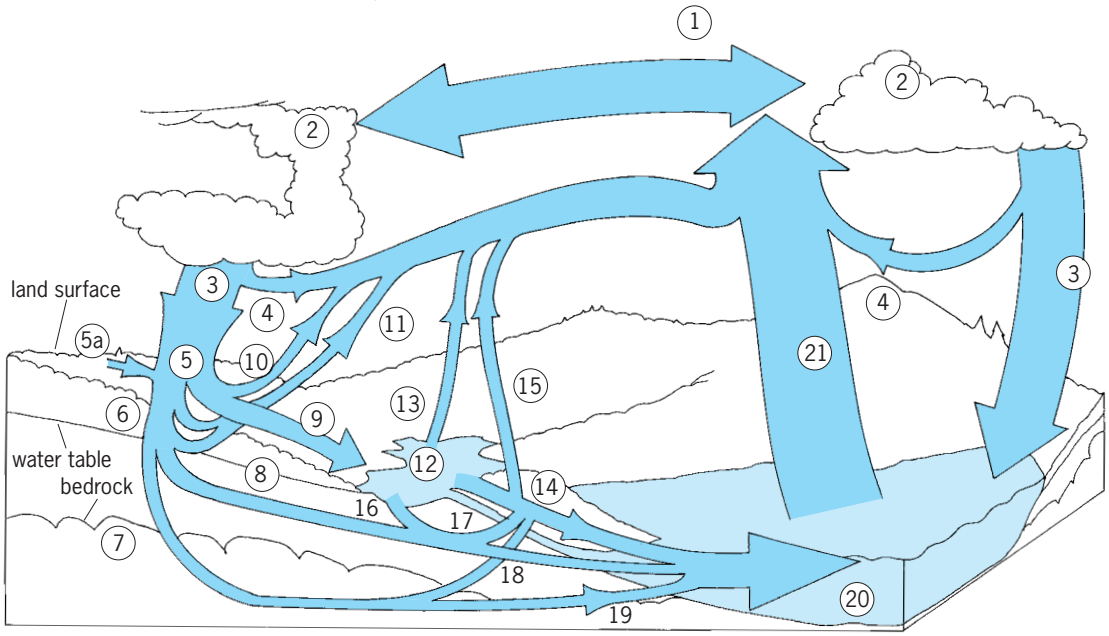

Key:

1. Atmospheric water in circulation
2. Formation of strom clouds
3. Precipitation (snow, rain, hail, etc.)
4. Evaporation of precipitation in transit and moisture diffusion
5. Net precipitation on Earth surface
5a. Snow storage and melt
6. Infiltration
7. Deep percolation
8. Ground-water flow and storage
9. Overland flow and depression storage
10. Evporation of intercepted and surface water
11. Evapotranspiration
12. Surface storage in lakes and reservoirs
13. Evaporation from lake surfaces
14. Surface streams
15. Evaporation from streams
16. Ground-water exchange from channel storage (base flow and return)
17. Ground-water exchange from channel storage (base flow and return)
18. Ground-water flow to ocean
19. Ground-water flow to ocean
20. Storage in the ocean
21. Evaporation from ocean surfaces

Fig. 2. Water flow scheme of the hydrologic cycle.

concentration of chemical elements, compounds, and biological constituents in surface and ground waters; amounts of water transferred by evaporation and evapotranspiration to the atmosphere from snow, lakes, streams, soils, and vegetation; and sediment lost from the land and transported by streams.

In addition to the making of these measurements with specially designed instruments and devices, hydrology is concerned with research on the phenomena and mechanisms involved in all physical and biological components of the hydrologic cycle, with the purpose of understanding them sufficiently to permit quantitative predictions and forecasting. The field investigations and measurements not only provide the data whereby the behavior of each component may be evaluated in detail, permitting formulation in quantitative terms, but also give a record of the historical performance of the entire system. Thus, two principal vehicles for hydrological forecasting and prediction become available: a set of elemental processes, whose operations are expressible in mathematical terms, linked to form deterministic models that permit the prediction of hydrologic events for given conditions; and a group of records or time series of measured hydrologic variables, such as precipitation or runoff, which can be analyzed by statistical methods to formulate stochastic models that permit inferences to be made on the future likelihood of hydrologic events.

Surface water. Hydrology, when specifically applied to the solution of problems, such as the estimation of water supplies from streams, flood control and protection, dam and reservoir design and operation, urban drainage, and changes in flow regimes due to modifications introduced by humans, traditionally has been designated as surface-water hydrology. This does not imply that flow barriers exist between surface water and ground water; however, because ground-water flows usually occur at much slower rates, the processes can be decoupled to some degree when the primary objective is to evaluate surface flows. Both deterministic and stochastic models are employed.

A variety of deterministic models have been proposed; they differ from each other mainly in the degree of simplification and schematization used to describe the components of the hydrologic cycle and their linkages. The common underlying concept is that the overall process is composed of distinct elements of water storage and translation, each following specific rules but all obeying the hydrologic balance equation, expressed, for any finite period of

time, as inflow = outflow + change in water storage. Typically, this water accounting is carried out at short intervals of time, both internally for each element and as a whole for the entire catchment, to yield the surface outflows to be expected as a result of changing precipitation rates. Data on relevant physical and vegetative characteristics of the catchment, as well as on factors affecting evaporation and evapotranspiration, are needed for the calculations, which are generally performed on computers, except in the simplest models for rough approximations. *See* STREAM GAGING; WATER SUPPLY ENGINEERING.

Predictions on the fate of chemical and biological species contained naturally in surface waters or introduced as pollutants due to urbanization and industrial and agricultural developments can be made by appropriate coupling of the equations of chemical and biochemical reactions with the equations representing the dynamics of water movement. Solutions to these coupled equations are typically obtained by using computers.

Stochastic process models are used for purposes such as drawing inferences on the probabilities of occurrence of future extreme events such as floods and droughts, or simulating possible future flow sequences similar in likelihood to those of the historical record. These inferences and synthetic time series can be used in the planning and design of water resource systems. Research on climatic change is an area of study that has a profound influence on the interpretation of historical hydrologic data. All conclusions drawn from recorded data need to be consistent with the possibility that local hydrologies may or may not remain as observed in the recent past. *See* CLIMATE HISTORY; STOCHASTIC PROCESS.

Ground water. Waters contained in porous formations below the ground are extremely valuable sources of supply. They depend for their long-term availability on replenishment (or recharge) from the surface. Water-bearing layers suitable for economical extraction by means of wells are called aquifers; they operate simultaneously as reservoirs and as flow media. The ability of an aquifer material to store water depends primarily on its porosity; and the ability to permit fluid motion, on its permeability or hydraulic conductivity. Because the properties of aquifers vary in space, the flow patterns vary from point to point. These patterns are further altered by the presence of wells, which act as localized sink points. For the rational management of ground-water basins that may contain several nonuniform and nonhomogeneous aquifers, special computer models, based on the equations of flow through porous media, have been developed. These specialized models can be deterministic in nature, where the aquifer properties are assumed to be completely known, or stochastic in nature, where aquifer properties are assumed to vary randomly over an area. Physical, chemical, and biological processes can be incorporated into either type of computer model to predict the quality of the water in the aquifer. Studies can also be undertaken by means of these models for the planning of conjunctive use operations of surface and ground waters, whereby controlled aquifer recharges and discharges can be combined with regulated surface flows and reservoirs for optimizing the management of water resources. *See* AQUIFER; FLUID-FLOW PRINCIPLES; GROUND-WATER HYDROLOGY; RESERVOIR; WATER TABLE; WELL.

Snow. In many parts of the world, snow is an important component of the total water supply. Snow fields constitute natural reservoirs that accumulate water during the winter and release it during the melting season. Snow hydrology is concerned with the study of the conditions of accumulation of snow, the properties of snowpacks as porous media, and the processes of heat and vapor exchange between the atmosphere and the snow as the snowpack evolves and melts under gradually varying conditions of solar radiation. These studies, together with field measurements and remote sensing of snow covered by satellites, permit the formulation and solution of mathematical (computer) models for the prediction of snowmelt as a function of time. Accordingly, it becomes possible to plan the operation of surface reservoir systems in combination with expected snowmelt yields for optimal water resource management. *See* REMOTE SENSING; SNOW; SNOW SURVEYING; SNOWFIELD AND NÉVÉ.

Global hydrology. The role of water in its liquid, gaseous (water vapor), and solid (ice) phases as a critical agent and regulator of biogeochemical and energy transport and balance in the planet Earth has been recognized since the beginning of the twentieth century. The strong coupling of land, atmosphere, and oceans in determining the rates of evapotranspiration, precipitation, ice melting, and biogeochemical cycles, and the importance of water as a greenhouse gas, have resulted in the emergence of hydrology as a scientific discipline with strong emphasis in basic research. Modern hydrologic curricula include training in meteorology, climatology, physical geography, and the basic natural sciences (biology, chemistry, physics), in addition to the traditional instruction in engineering, physical hydrology, mathematics, fluid mechanics, and geology. *See* BIOGEOCHEMISTRY; GREENHOUSE EFFECT; HYDROSPHERE.

Miguel A. Mariño

Bibliography. P. B. Bedient and W. C. Huber, *Hydrology and Floodplain Analysis*, 2d ed., Addison-Wesley, Reading, MA, 1992; S. L. Dingman, *Physical Hydrology*, Macmillan, New York, 1994; P. S. Eagleson (ed.), *Opportunities in the Hydrologic Sciences*, National Research Council, National Academy Press, Washington, DC, 1991; A. D. Feldman (ed.), *Engineering Hydrology: Symposium Proceedings*, Hydraulics Division, American Society of Civil Engineers, New York, 1987; R. A. Freeze and J. A. Cherry, *Groundwater*, Prentice Hall, Englewood Cliffs, NJ, 1979; R. K. Linsley, Jr., et al., *Hydrology for Engineers*, 3d ed., McGraw-Hill, New York, 1982; D. R. Maidment (ed.), *Handbook of Hydrology*, McGraw-Hill, New York, 1993; M. A. Mariño and J. N. Luthin, *Seepage and Groundwater*, Elsevier Scientific, Amsterdam, 1982; W. Stumm and J. J. Morgan, *Aquatic Chemistry*, John Wiley, New York, 1996.

Hydrolysis

A chemical reaction in which splitting of a molecule by water occurs.

Hydrolysis as applied to organic molecules can be considered a reversal of such reactions as esterification and amide formation. The hydrolysis of esters [reaction (1)] and of amides [reaction (2)]

$$\underset{\text{Ethyl acetate}}{CH_3COOC_2H_5} + H_2O \rightarrow \underset{\text{Acetic acid}}{CH_3COOH} + \underset{\text{Ethanol}}{C_2H_5OH} \quad (1)$$

$$\underset{\text{Acetamide}}{CH_3CONH_2} + H_2O \rightarrow \underset{\text{Acetic acid}}{CH_3COOH} + \underset{\text{Ammonia}}{NH_3} \quad (2)$$

is shown. Other classes of organic compounds that are subject to hydrolysis include acetals, acyl and alkyl halides, ketals, and peptides. While the overall hydrolysis reaction [as in reactions (1) and (2)] appears to involve the addition of the water molecule (H_2O), the reaction is in fact more complicated. There are several reaction steps, such as the formation of a complex with either a proton [H^+; acid-catalyzed hydrolysis] or a hydroxyl ion (OH^-; base-catalyzed hydrolysis), followed by elimination of these ions to give the overall equation. The hydrolysis reaction is frequently encountered in biological systems. The kinetics of these reactions are greatly enhanced by the action of enzymes (biological catalysts) such as the esterases and the peptidases. *See* ENZYME; HYDROGEN ION.

In inorganic chemistry, hydrolysis, also called aquation, represents a class of reactions involving metal coordination complexes in which one of the coordinated ligands is displaced by either H_2O or OH^-. Hydrolysis is a special case of the class of reactions termed ligand displacement reactions [reaction (3), where M is a metal, L and X are ligands,

$$L_nMX + Y \rightarrow L_nMY + X \quad (3)$$

and Y is the displacing ligand (H_2O in hydrolytic reactions)]. *See* COORDINATION COMPLEXES; LIGAND.

The phenomena attributed to hydrolysis are commonly explained in terms of Brönsted acid-base theory. An acid is a species capable of donating a proton; a base, one that can accept a proton. A conjugate acid-base pair comprises two species that differ by a single transferable proton. From this, it follows that a solution of a weak acid (HB) will contain some of the conjugate base to an extent governed by the acid dissociation constant, K_a, and its concentration, C_{HB}. As the K_a of a weak acid gets smaller, the extent of dissociation of its conjugate base increases. Not all acids or bases are neutral and uncharged. Positively charged ammonium ion (NH_4^+) is the conjugate acid to ammonia (NH_3), and negatively charged acetate ion is the conjugate base of acetic acid. A solution of sodium acetate will be basic; its pH will be greater than 7. Since the proton must be donated by water, the extent of this reaction is determined by the competition between OH^-, the conjugate base of water, and acetate ion; since acetate ion is a weaker base than OH^-, the extent is small. Similarly, ammonium salts will have pH values somewhat lower than 7. Hence, to say that acetates and ammonium salts undergo hydrolysis is no longer necessary in view of the Brönsted theory. *See* ACID AND BASE; HYDROLYTIC PROCESSES; ION; PH.

Henry Freiser

Bibliography. H. Freiser, *Concepts and Calculations in Analytical Chemistry*, 1992; I. M. Kolthoff et al., *Quantitative Chemical Analysis*, 4th ed., 1969.

Hydrolytic processes

Reactions of both organic and inorganic chemistry wherein water effects a double decomposition with another compound, hydrogen going to one component, hydroxyl to another, as in reactions (1)–(3).

$$XY + H_2O \rightarrow HY + XOH \quad (1)$$

$$KCN + H_2O \rightarrow HCN + KOH \quad (2)$$

$$C_5H_{11}Cl + H_2O \rightarrow HCl + C_5H_{11}OH \quad (3)$$

Although the word hydrolysis means decomposition by water, cases in which water brings about effective hydrolysis unaided are rare, and high temperatures and pressures are usually necessary. *See* HYDROLYSIS.

In the field of organic chemistry, the term "hydrolysis" has been extended to cover the numerous reactions in which alkali or acid is added to water. An example of an alkaline-condition hydrolytic process is the hydrolysis of esters, reaction (4), to produce

$$CH_3COOC_2H_5 + NaOH \rightarrow CH_3COONa + C_2H_5OH \quad (4)$$

alcohol. An example of an acidic-condition process is the hydrolysis of olefin to alcohol in the presence of phospheric acid, reaction (5). The addition of acids

$$C_2H_4 + H_2O \xrightarrow{H_2PO_4} C_2H_5OH \quad (5)$$

of alkalies hastens such reactions even if it does not initiate the reaction.

Hydrolysis reactions may be classified as follows: (1) hydrolysis with water alone; (2) hydrolysis with dilute or concentrated acid; (3) hydrolysis with dilute or concentrated alkali; (4) hydrolysis with fused alkali with little or no water at high temperature.

Alcohol processes. The direct synthesis of ethanol from ethylene in the presence of phosphoric acid has already been mentioned. On a commercial scale this reaction is conducted at high temperature (300°C or 572°F) and high pressure (1000 lb/in.^{2} or 6895 kilopascals) with the reactants in the vapor state and the phosphoric acid supported on pelleted diatomaceous earth. During operation a small stream of phosphoric acid is injected at the inlet of the reactor to replace acid that is carried out by the reactants. Conversion of ethylene per pass is low, about 4%. Unreacted ethylene is recycled, and the overall process yield is 95–98%.

An alternative process for ethanol production is a liquid-phase process which uses sulfuric acid as the catalyst. In this process, the ethylene is absorbed

into concentrated sulfuric acid and forms monoethyl and diethyl sulfate. The esters are transferred to a hydrolyzer vessel along with water to form the alcohol. The stepwise reaction is shown in (6). The sulfuric-

$$C_2H_4 + H_2SO_4 \rightarrow C_2H_5OSO_3 \xrightarrow{H_2O} C_2H_5OH + H_2SO_4 \quad (6)$$

acid-catalyzed process is also used to produce isopropyl alcohol from normal propylene, secondary butanol from normal butylene, tertiary butanol from isobutylene, and secondary and tertiary amyl alcohol from pentenes. The process produces secondary and tertiary alcohols almost exclusively and very little normal alcohols. *See* ALCOHOL.

Phenol processes. The commercial production of phenol may be accomplished by several hydrolysis processes. The processes discussed here represent earlier technology and, although valid, are no longer economically attractive for new plants. Some existing units are still operated. One of the older processes is an example of an alkali fusion process. Benzene is the starting material for the lengthy process. The reactions involved are summarized in (7)–(10).

$$C_6H_6 + H_2SO_4 \rightarrow C_5H_5SO_3H + H_2O \quad (7)$$

$$2C_6H_5SO_3H + Na_2SO_3 \rightarrow 2C_6H_5SO_3Na + H_2SO_3 \quad (8)$$

$$C_6H_5SO_3Na + 2NaOH \rightarrow C_6H_5ONa + Na_2SO_3 + H_2O \quad (9)$$

$$2C_6H_5ONa + H_2SO_3 \rightarrow 2C_6H_5OH + Na_2SO_3 \quad (10)$$

Reactions (7) and (8) represent the formation of benzenesulfonic acid from benzene and sulfuric acid and the subsequent neutralization with sodium sulfite to form sodium benzenesulfonate. The fusion of sodium hydroxide and sodium benzenesulfonate in reaction (9) produces sodium phenoxide, and in reaction (10), sodium phenoxide is neutralized to yield phenol.

A second process is based on the hydrolysis of chlorobenzene in caustic soda solution at 350°C (662°F) and 4000 lb/in.^{2} (27,580 kPa). The reaction is summarized in (11)–(13). In a commercial op-

$$C_6H_5Cl + 2NaOH \rightarrow C_6H_5ONa + NaCl + H_2O \quad (11)$$

$$C_6H_5ONa + C_6H_5Cl \rightarrow C_6H_5{-}O{-}C_6H_5 + NaCl \quad (12)$$

$$C_6H_5{-}O{-}C_6H_5 + H_2O \rightarrow 2C_6H_5OH \quad (13)$$

eration the reaction is carried out in a very long, high-pressure tubular reactor. The overall reaction is exothermic, and external heating is needed only during start-up.

A third phenol process is a regenerative process that takes place in the vapor phase. The two steps in the process are shown in reaction (14) and (15).

$$C_6H_6 + HCl + O \rightarrow C_6H_5Cl + H_2O \quad (14)$$

$$C_6H_5Cl + H_2O \rightarrow 2C_6H_5OH \quad (15)$$

In the first step benzene, hydrogen chloride, and atmospheric oxygen are consumed, and chlorobenzene and water are products. The chlorobenzene and water are consumed in the second-step reaction to yield the desired phenol product and the first-step reactant, hydrogen chloride. The reaction is carried out in two catalyzed vapor-phase reactions. Although the overall reaction is simple and straightforward, in practice the per pass conversion in the first step is only about 12% while the per pass conversion in the second step is only 15%. The separation and recycle equipment required to conserve the reactants complicates the commercial unit considerably. *See* HALOGENATED HYDROCARBON.

Soap manufacture. Perhaps some of the oldest and largest-volume hydrolysis technology is involved in soap manufacture. A two-step process for saponification is shown in reactions (16) and (17). In the

$$(C_{17}H_{35}COO)_3C_3H_5 + 3H_2O \rightarrow 3C_{17}H_{35}COOH + C_3H_5(OH)_3 \quad (16)$$

$$C_{17}H_{35}COOH + NaOH \rightarrow C_{17}H_{35}COONa + H_2O \quad (17)$$

first step, glyceryl stearate acid, a fat, is hydrolyzed with water to yield stearic acid and glycerin. In the second step, the stearic acid is neutralized with caustic soda to give sodium stearate, the soap, and water. Originally performed as a batch operation, the process is now automated to a continuous process. *See* FAT AND OIL.

Products. Most hydrolytic processes take place in acid or alkaline conditions which in some cases are quite severe. The acids used are primarily sulfuric, hydrochloric, and phosphoric. Much of the equipment in hydrolytic processes must be of special construction such as glass-lined steel or special metal alloy.

Hydrolytic processes account for a huge product volume. Conversion of starch such as cornstarch into maltose and glucose (sugar syrups) by treatment with hydrochloric acid is a major industry. Similarly, the production of furfural from pentosans of oat hulls or other cereal by-products such as corn cobs, rice hulls, or cottonseed bran is another commercial hydrolytic process. The production of alcohols, soaps, and industrial intermediate chemicals have already been mentioned. The volume and diversity of the products serve to illustrate the fundamental nature and importance of hydrolytic processes. D. L. Holt

Hydrometallurgy

The extraction and recovery of metals from their ores by processes in which aqueous solutions play a predominant role. Two distinct processes are involved in hydrometallurgy: putting the metal values in the ore into solution via the operation known as leaching; and recovering the metal values from solution, usually after a suitable solution purification or concentration step, or both. The scope of hydrometallurgy is quite broad and extends beyond the processing of ores to the treatment of metal concentrates, metal

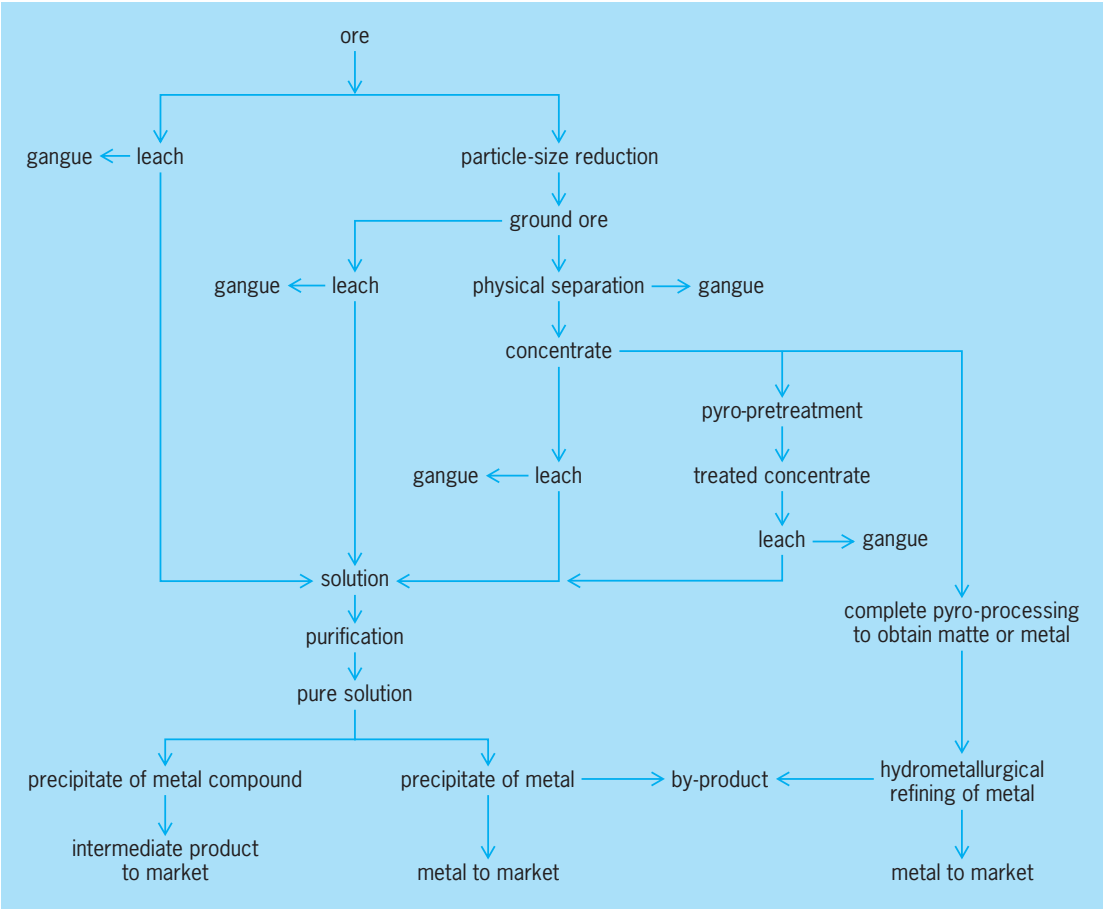

Generalized metallurgical flow sheet.

scrap and revert materials, and intermediate products in metallurgical processes. Hydrometallurgy enters into the production of practically all nonferrous metals and of metalloids, such as selenium and tellurium.

A generalized metallurgical flow sheet (see **illus.**) provides an indication of the nature and extent of the role of hydrometallurgy in metal production. It also shows the manner in which hydrometallurgical and pyrometallurgical processes complement each other. *See* PYROMETALLURGY.

The advantages of hydrometallurgy include applicability to low-grade ores (copper, uranium, gold, silver) and complex sulfide ores, amenability to the treatment of materials of quite different compositions and concentrations, adaptability to separation of highly similar materials (hafnium from zirconium), and flexibility in terms of the scale of operations. Hydrometallurgical operations are amenable to effective control leading to automation and continuous operation.

There are some important disadvantages in hydrometallurgical processes. The processes are generally energy-intensive and can involve the handling of large volumes of dilute solutions that may be corrosive or hazardous. In addition, the processes yield residues and effluents that must be disposed of in an environmentally acceptable manner.

The first commercial hydrometallurgical operation was the application of cyanidation to the treatment of gold ores in 1889, more than 40 years after the discovery that gold can be dissolved in dilute alkaline solutions of sodium cyanide, as shown in reaction (1).

$$4Au + 8NaCN + O_2 + 2H_2O \longrightarrow 4NaAu(CN)_2 + 4NaOH \quad (1)$$

Leaching. A major unit process in hydrometallurgy is the leaching of ores and concentrates. A particularly noteworthy and time-honored dissolution process is the caustic pressure leaching of bauxite ore to produce pure alumina. This process, which was developed in the nineteenth century, is still in use as the principal method of producing alumina. The ever-increasing demand for aluminum and the rising cost of bauxite stimulated research on the development of processes for leaching nonbauxite alumina materials such as a alunite, anorthosite, clays, and coal shales. *See* ALUMINUM; LEACHING.

In the case of sulfide concentrates, leaching in acid or ammoniacal solution under oxygen pressure provides a number of advantages, including accelerated leaching and the recovery of sulfur in a suitable form, such as ammonium sulfate in the ammoniacal leaching of nickel concentrates in the Sherritt-Gordon

process or as elemental sulfur in the acid pressure leaching of zinc concentrates.

The economic recovery of metals from low-grade ores became increasingly important as mining costs rose and as higher-grade materials became more difficult to find. In many operations, what was at one time waste is now considered as ore to be treated by never leaching methods such as heap, dump, and in situ leaching. In the United States, copper in excess of 160,000 tons (145,000 metric tons) is extracted annually by such processes. The same technology is applied in the leaching of uranium ores and in the treatment of gold and silver ores.

Biotechnology has come to play a more significant role in mineral leaching and metal separation processes. Of particular importance is the microbiological leaching of low-grade sulfide copper ores and of uranium ores in the presence of pyrite utilizing the action of the bacterium *Thiobacillus ferrooxidans*. The same organism provides an attractive option in the treatment of refractory gold ores. The bioleaching of minerals such as pyrite and arsenopyrite, in which gold is entrapped, facilitates the subsequent recovery of the gold by cyanide leaching. *See* BIOLEACHING.

The establishment of more stringent regulations in many countries concerning sulfur dioxide emissions from smelters has led not only to the adoption of new and improved smelting processes but also to a great deal of research on hydrometallurgical alternatives to smelting, especially of copper concentrates. In the case of copper concentrates, the principal difficulty has been the realization of a hydrometallurgical process that has sufficiently well-defined economic advantages. From an environmental standpoint, there are definite advantages which are afforded by hydrometallurgy, principally the elimination of sulfur dioxide emissions and the conversion of sulfur to elemental sulfur or a sulfate. The overall economics of copper hydrometallurgy can be improved by the recovery of copper in final form by a modified electrowinning process or by an alternative process such as the hydrogen reduction of cuprous chloride.

Environmental and other considerations led to the cessation of the production of zinc by the carbothermic reduction of zinc calcine at 2200–2400°F (1200–1300°C) in horizontal retorts by reaction (2), where

$$ZnO(s) + C(s) \rightarrow Zn(g) + CO(g) \qquad (2)$$

(s) and (g) denote solid and gas. This process has been replaced by the sulfuric acid leaching of zinc calcine, purification of the resulting leach solution, and the recovery of zinc by electrowinning. An important alternative to the production of zinc calcine by fluid-bed roasting is the acid pressure leaching of zinc concentrates, in which the sulfide is oxidized to yield elemental sulfur.

Recovery. The recovery of metals from leach solutions is often preceded by the purification or concentration of the metal to be isolated. Two processes that can be used to achieve this objective are ion exchange and solvent extraction.

Ion exchange. Ion exchange involves the selective adsorption of metal ions onto a synthetic organic ion-exchange resin (commonly a styrene-divinylbenzene copolymer containing suitable functional groups for ion exchange). The process is stoichiometric and reversible, so that the adsorbed ions are subsequently desorbed or eluted from the resin by a solution known as the eluant. The resulting solution or eluate is treated for metal recovery or for production of a metal compound, such as ammonium diuranate, $(NH_4)_2U_2O_7$, in uranium extraction. Ion exchange, which is widely used in water purification, is employed extensively in uranium recovery. Here continuous ion-exchange processes are in commercial use, with resultant increases in efficiency and productivity. Chelating resins should find increasing use in the separation of metals such as copper, nickel, cobalt, and zinc. *See* CHELATION; ION EXCHANGE.

Solvent extraction. As applied in hydrometallurgy, solvent extraction relates to the selective transfer of metal species from an aqueous solution to an immiscible organic solvent with which it is in contact. The solvent itself can be the active reagent in the process, as in the extraction of uranium by tributylphosphate. On the other hand, the solvent can contain an organic extractant which forms a complex or chelation compound with the metal ion from the aqueous phase. The extraction process is reversible, so that the extracted metal can be stripped from the organic solution and returned to the aqueous phase.

With the development of suitable chelating compounds, solvent extraction became established as an excellent technique for the upgrading of copper leach solutions. The application of solvent extraction is especially suited to the treatment of the dilute solutions resulting from the heap and dump leaching of low-grade copper ores. The selectivity of the organic reagents, notably the hydroxyoximes, is such as to permit an excellent separation from iron. The stripping of the organic phase by spent electrowinning electrolyte leads to a strong solution, from which copper can be electrowon without difficulty.

Solvent extraction is used commercially in the separation of many metals, including copper, nickel, cobalt, zinc, uranium, thorium, zirconium, hafnium, molybdenum, tungsten, niobium, tantalum, and beryllium. *See* SOLVENT EXTRACTION.

Recovery from aqueous solution. Metal recovery from aqueous solution can be effected by a number of reduction processes, the most prominent being electrowinning and gaseous reduction using hydrogen. The latter process yields metal powder, whereas the former results in cathodes which are melted and cast to produce various marketable shapes.

Importance. Hydrometallurgy occupies an important role in the production of aluminum, copper, nickel, cobalt, zinc, gold, silver, platinum metals, selenium, tellurium, tungsten, molybdenum, uranium, zirconium, and other metals. Considering the versatility of hydrometallurgy and the need to process

more complex ores, as well as lower-grade ores and secondary materials, and to produce high-purity advanced materials, hydrometallurgical processes are expected to play an even greater part in the production of metals and materials in the future. *See* ELECTROMETALLURGY; METALLURGY.

W. Charles Cooper

Bibliography. J. H. Canterford, Hydrometallurgy: Winning metals with water, *Chem. Eng.*, pp. 41–48, October 28, 1985; W. C. Cooper, G. E. Lagos, and G. Ugarte (eds.), Hydrometallurgy and electrometallurgy of copper, *Copper 87*, vol. 3, University of Chile, 1988; G. A. Davies (ed.), *Separation Processes in Hydrometallurgy*, Society of the Chemical Industry, U.K., 1987; H. L. Ehrlich and C. L. Brierley, *Microbial Mineral Recovery*, 1990; J. B. Hiskey and C. W. Warren (eds.), *Hydrometallurgy: Fundamentals, Technology, and Innovations*, 1993; K. Osseo-Asare and J. D. Miller (eds.), *Hydrometallurgy: Research, Development and Plant Practice*, AIME, 1983; G. Rossi, *Biohydrometallurgy*, 1990.

Hydrometeorology

The study of the occurrence, movement, and changes in the state of water in the atmosphere. The term is also used in a more restricted sense, especially by hydrologists, to mean the study of the exchange of water between the atmosphere and continental surfaces. This includes the processes of precipitation and direct condensation, and of evaporation and transpiration from natural surfaces. Considerable emphasis is placed on the statistics of precipitation as a function of area and time for given locations or geographic regions.

Water occurs in the atmosphere primarily in vapor or gaseous form. The average amount of vapor present tends to decrease with increasing elevation and latitude and also varies strongly with season and type of surface. Precipitable water, the mass of vapor per unit area contained in a column of air extending from the surface of the Earth to the outer extremity of the atmosphere, varies from almost zero in continental arctic air to about 1.4 oz/in.^{2} (6 g/cm^2) in very humid, tropical air. Its average value over the Northern Hemisphere varies from around 0.46 oz/in.^{2} (2.0 g/cm^2) in January and February to around 0.84 oz/in.^{2} (3.7 g/cm^2) in July. Its average value is around 0.63 oz/in.^{2} (2.8 g/cm^2), an amount equivalent to a column of liquid water slightly greater than 1 in. (2.5 cm) in depth. Close to 50% of this water vapor is contained in the atmosphere's first mile, and about 80% is to be found in the lowest 2 mi (3 km).

Atmospheric water cycle. Although a trivial proportion of the water of the globe is found in the atmosphere at any one instant, the rate of exchange of water between the atmosphere and the continents and oceans is high. The average water molecule remains in the atmosphere only about 10 days, but because of the extreme mobility of the atmosphere it is usually precipitated many hundreds or even thousands of miles from the place at which it entered the atmosphere.

Evaporation from the ocean surface and evaporation and transpiration from the land are the sources of water vapor for the atmosphere. Water vapor is removed from the atmosphere by condensation and subsequent precipitation in the form of rain, snow, sleet, and so on. The amount of water vapor removed by direct condensation at the Earth's surface (dew) is relatively small.

A major feature of the atmospheric water cycle is the meridional net flux of water vapor. The average precipitation exceeds evaporation in a narrow band extending approximately from 10°S to 15°N lat. To balance this, the atmosphere carries water vapor equatorward in the tropics, primarily in the quasi-steady trade winds which have a component of motion equatorward in the moist layers near the Earth's surface. Precipitation also exceeds evaporation in the temperate and polar regions of the two hemispheres, poleward of about 40° latitude. In the middle and higher latitudes, therefore, the atmosphere carries vapor poleward. Here the exchange occurs through the action of cyclones and anticyclones, large-scale eddies of air with axes of spin normal to the Earth's surface.

For the globe as a whole the average amount of evaporation must balance the precipitation **(Table 1)**. This exchange is related to the characteristics of the general circulation of the atmosphere. It seems likely that a similar cycle would be observed even if the Earth were entirely covered by ocean, although details of the cycle, such as the flux across the Equator, would undoubtedly be different.

Complications in the global pattern arise from the existence of land surfaces. Over the continents the only source of water is from precipitation; therefore, the average evapotranspiration (the sum of evaporation and transpiration) cannot exceed precipitation. The flux of vapor from the oceans to the continents through the atmosphere, and its ultimate return to atmosphere or ocean by evaporation, transpiration, or runoff is known as the hydrologic cycle. Its atmospheric phase is closely related to the air mass cycle. In middle latitudes of the Northern Hemisphere, for example, precipitation occurs primarily from maritime air masses moving northward and eastward across the continents. Statistically, precipitation from

TABLE 1. Meridional flux of water vapor in the atmosphere

Latitude	Northward flux, 10^{10} g/s
90°N	0
70°N	4
40°N	70
10°N	−61
Equator	45
10°S	71
40°S	−75
70°S	1
90°S	0

these air masses substantially exceeds evapotranspiration into them. Conversely, cold and dry air masses tend to move southward and eastward from the interior of the continents out over the oceans. Evapotranspiration into these continental air masses strongly exceeds precipitation, especially during winter months. These facts, together with the extreme mobility of the atmosphere and its associated water vapor, make it likely that only a small percentage of the water evaporated or transpired from a continental surface is reprecipitated over the same continent. *See* ATMOSPHERIC GENERAL CIRCULATION; HYDROLOGY.

Precipitation. Hydrometeorology is particularly concerned with the measurement and analysis of precipitation data. Radar plays an important role in estimating precipitation. By relating the intensity of radar echo to rate of precipitation, it has been possible to obtain a vast amount of detailed information concerning the structure and areal distribution of storms. *See* METEOROLOGICAL INSTRUMENTATION; RADAR METEOROLOGY; STORM DETECTION.

Deficiencies in the observational networks over the oceans and over the more sparsely inhabited land areas of the Earth are now being bridged through the use of meteorological satellite observations. Progress toward the development of methods for estimating rainfall amounts from satellite observations of cloud type and distribution is of particular significance to hydrometeorology. *See* METEOROLOGICAL SATELLITES.

Precipitation occurs when the air is cooled to saturation. The ascent of air toward lower pressure is the most effective process for causing rapid cooling and condensation. Precipitation may therefore be classified according to the atmospheric process which leads to the required upward motion. Accordingly, there are three basic types of precipitation: (1) Orographic precipitation occurs when a topographic barrier forces air to ascend. The presence of significant relief often leads to large variations in precipitation over relatively short distances. (2) Extratropical cyclonic precipitation is associated with the traveling regions of low pressure of the middle and high latitudes. These storms, which transport sinking cold dry air southward and rising warm moist air northward, account for a major portion of the precipitation of the middle and high latitudes. (3) Air mass precipitation results from disturbances occurring in an essentially homogeneous air mass. This is a common precipitation type over the continents in mid-latitudes during summer. It is the major mechanism for precipitation in the tropics, where disturbances may range from areas of scattered showers to intense hurricanes. In most cases there is evidence for organized lifting of air associated with areas of cyclonic vorticity, that is, areas over which the circulation is counterclockwise in the Northern Hemisphere or clockwise in the Southern Hemisphere.

The availability of data from geosynchronous meteorological satellites, together with surface and upper-air data acquired as part of the Global Atmospheric Research Program (GARP), is leading to significant advances in the understanding of the character and distribution of tropical precipitation. *See* HURRICANE; STORM; TROPICAL METEOROLOGY.

Precipitation may, of course, be in liquid or solid form. In addition to rain and snow there are other forms which often occur, such as hail, snow pellets, sleet, and drizzle. If upward motion occurs uniformly over a wide area measured in tens or hundreds of miles, the associated precipitation is usually of light or moderate intensity and may continue for a considerable period of time. Vertical velocities accompanying such stable precipitation are usually of the order of several centimeters per second. Under other types of meteorological conditions, particularly when the density of the ascending air is less than that of the environment, upward velocities may locally be very large (of the order of several meters per second) and may be accompanied by compensating downdrafts. Such convective precipitation is best illustrated by the thunderstorm. Intensity of precipitation may be extremely high, but areal extent and local duration are comparatively limited. Storms are sometimes observed in which local convective regions are embedded in a matrix of stable precipitation. *See* PRECIPITATION (METEOROLOGY).

Analysis of precipitation data. Precipitation is essentially a process which occurs over an area. However, despite the use of radar, most observations are taken at individual stations. Analyses of such "point" precipitation data are most often concerned with the frequency of intense storms. These data (such as in **Table 2**) are of particular importance in evaluating

TABLE 2. Record observed point rainfalls*

Duration	Depth, in. (cm)	Station	Date
1 min	1.23 (3.12)	Unionville, Maryland	July 4, 1956
8 min	4.96 (12.6)	Fuüssen, Bavaria	May 25, 1920
15 min	7.80 (19.8)	Plumb Point, Jamaica	May 12, 1916
42 min	12.00 (30.5)	Holt, Missouri	June 22, 1947
2 h 45 min	22.00 (55.9)	Near D'Hanis, Texas	May 31, 1935
24 h	73.62 (187)	Cilaos, La Reunion (Indian Ocean)	March 15–16, 1952
1 month	366.14 (930)	Cherrapunji, India	July 1861
12 months	1041.78 (2646.1)	Cherrapunji, India	August 1860 to July 1861

*From D. R. Maidment (ed.), *Handbook of Meteorology*, 1993.

local flood hazard, and may be used in such diverse fields as the design of local hydraulic structures, such as culverts or storm sewers, or the analysis of soil erosion. Intense local precipitation of short duration (up to 1 h) is usually associated with thunderstorms. Precipitation may be extremely heavy for a short period, but tends to decrease in intensity as longer intervals are considered.

A typical hydrometeorological problem might involve estimating the likelihood of occurrence of a storm of given intensity and duration over a specified watershed to determine the required spillway capacity of a dam. Such estimates can only be obtained from a careful meteorological and statistical examination of large numbers of storms selected from climatological records.

Evaporation and transpiration. In evaluating the water balance of the atmosphere, the hydrometeorologist must also examine the processes of evaporation and transpiration from various types of natural surfaces, such as open water, snow and ice fields, and land surfaces with and without vegetation. From the point of view of the meteorologist, the problem is one of transfer in the turbulent boundary layer. It is complicated by topographic effects when the natural surface is not homogeneous. In addition the simultaneous heating or cooling of the atmosphere from below has the effect of enhancing or inhibiting the transfer process. Although the problem has been attacked from the theoretical side, empirical relationships are at present of greatest practical utility. *See* METEOROLOGY; MICROMETEOROLOGY.

Eugene M. Rasmusson

Bibliography. R. K. Linsley et al., *Hydrology for Engineers*, 3d ed., 1982; F. K. Lutgens and E. J. Tarbuck, *Atmosphere: An Introduction to Meteorology*, 8th ed., 2000; R. R. Maidment (ed.), *Handbook of Meteorology*, 1993; W. D. Sellers, *Physical Climatology*, 1965.

Hydrometer

A direct-reading instrument for indicating the density, specific gravity, or some similar characteristic of liquids. Almost all hydrometers are made of a high-grade glass tubing. The main body is the float section in the bottom of which ballast, such as small shot, is secured. A small-diameter tube, the stem, extends from the upper end of the float section. Inside the stem is the scale, printed on heavy-grade paper, and well-secured within the stem so its position will not change. When the hydrometer is placed in a liquid, the stem extends vertically above the surface for a portion of its length.

The regular, or plain, hydrometer is illustrated in **Fig. 1***a*, and the thermohydrometer, which has a thermometer enclosed in the float section, in Fig. 1*b*.

Hydrometers may be classified according to the indication provided by graduations of the scale as follows: (1) density hydrometers, to indicate densities at a particular temperature, and usually for a particular liquid; (2) specific gravity hydrometers to indicate specific gravity of a liquid, with reference to water, at a particular temperature; (3) percentage hydrometers to indicate, at a particular temperature, the percentage of a substance such as salt, sugar, or alcohol dissolved in water; (4) arbitrary scale hydrometers, indicating the density, specific gravity, or concentration of a liquid in terms of an arbitrarily defined scale, at a defined temperature.

Alcoholometers, which belong in group 3, indicate the percentage of ethyl alcohol, either by volume or by weight, in a water-alcohol mixture, or the scale may be percent "proof spirit." (Percent proof spirit is twice the percentage of ethyl alcohol by volume at 60°F or 15.6°C.) The Brix and Balling (Plato) saccharimeters indicate the percentage by weight of pure sucrose solutions.

Possibly the most widely used arbitrary scale hydrometers are the API (American Petroleum Institute) and the Baumé (Bé). The relation between the

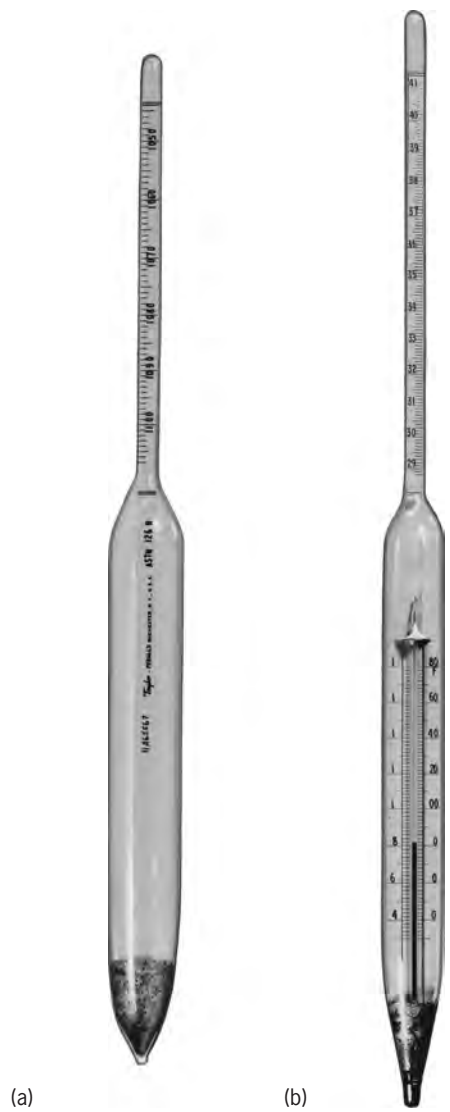

Fig. 1. Types of hydrometer. (*a*) Plain hydrometer. (*b*) Thermohydrometer. (*Taylor Instrument Co.*)

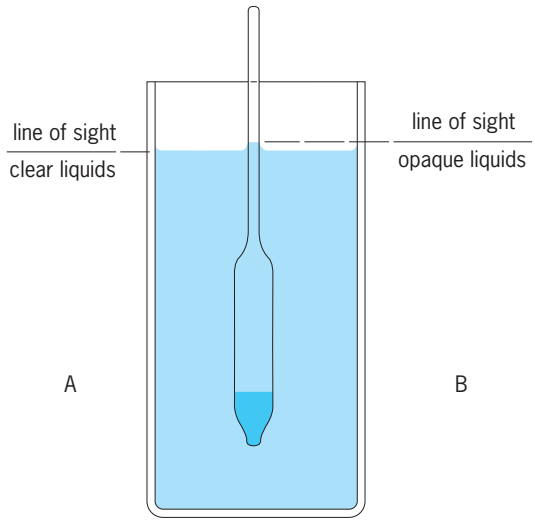

Fig. 2. Illustration of method of reading hydrometers.

API scale and specific gravity is given by Eq. (1). The

$$\text{Degrees API} = \frac{141.5}{\text{sp gr } 60/60\text{ F}} - 131.5 \qquad (1)$$

denominator, sp gr 60/60 F, means the specific gravity of the liquid (an oil) when the temperatures of both the liquid and the water to which it is referred are 60°F (15.6°C).

The relation between the Baumé scale and specific gravity for liquids lighter than water is given by Eq. (2); for liquids heavier than water, by Eq. (3).

$$\text{Degrees Bé} = \frac{140}{\text{sp gr } 60/60\text{ F}} - 130 \qquad (2)$$

$$\text{Degrees Bé} = 145 - \frac{145}{\text{sp gr } 60/60\text{ F}} \qquad (3)$$

Other special scales are the Quevenne, used in lactometers for testing milk, and the barkometer and Twaddle scales used by the leather-tanning industry in testing the strength of tanning extracts.

The best way to read a hydrometer in clear liquids is to start with the eyes slightly below the plane of the liquid surface and slowly to raise the eyes until the surface of the liquid appears as a straight line (**Fig. 2**, side A). The place where the line crosses the scale is the reading. With opaque liquids, such as milk and many oils, it is necessary to read the hydrometer at the top of the meniscus (Fig. 2, side B).

For accurate readings the hydrometer, especially the stem, must be absolutely clean. Also, the surface of the liquid must be clean and free of dust. With a precision grade hydrometer, with a long small-diameter stem, specific gravity values may be read to 0.0001. In the general use of hydrometers, however, the uncertainty of readings will probably be about ±0.001 specific gravity or the equivalent. *See* SPECIFIC GRAVITY.

Howard S. Bean

Bibliography. D. M. Considine, *Process/Industrial Instruments and Controls Handbook*, 4th ed., 1993; R. W. Herschy, *Hydrometry; Principles and Practices*, 1978; J. C. Hughes, *Testing of Hydrometers*, U.S. Department of Commerce, Nat. Bur. Stand. Circ. 555, October 1954.

Hydrophone

The underwater equivalent of a microphone, which generates an electrical signal as a response to the pressure component of an acoustic signal. Modern hydrophones are usually composed of piezoelectric ceramics such as barium titanate and lead zirconate titanate. The piezoelectric effect is the ability of materials to generate a voltage in response to applied mechanical stress. This effect is reciprocal in that an applied voltage will induce a small deformation in a piezoelectric material. Thus, some hydrophones can also be used as acoustic sources or projectors. In general, the capacitance of a hydrophone is relatively small, so a preamplifier is required to be placed near the piezoelectric material. The preamplifier boosts the electrical signal to allow the use of long cables without reducing the sensitivity of the hydrophone due to the added capacitance. When a preamplifier is installed in a hydrophone, the hydrophone is no longer reciprocal and may not be used as a projector. Hydrophone design requirements differ from microphones in that they must be electrically insulated from water and may need to operate at high hydrostatic pressure (deep depths). *See* ELECTRET TRANSDUCER; MICROPHONE; PIEZOELECTRICITY; PREAMPLIFIER.

Acoustic signals involve very small displacements of particles (water molecules, for instance) that propagate with relatively high speeds (1500 m/s or 4900 ft/s for water). These displacements result in local condensations (higher mean density) and rarefactions (lower mean density). Associated with the density fluctuations are changes in pressure, a scalar quantity containing no directional information. In order to obtain directional information, it is necessary to use arrays of hydrophones distributed along a line, over a surface, or throughout a volume. The acoustic signal arriving from a distant source will reach each hydrophone at a different time. This time difference, along with the known location of each hydrophone, provides information on the arrival direction. Modern underwater sensor systems also use sensors to detect the vector component (particle displacement, velocity, or acceleration) of acoustic signals to provide directional information in a more compact array than is achievable in a hydrophone array. *See* SOUND; SOUND PRESSURE.

Hydrophones are currently used to detect and determine the location of submarines, with hydrophone arrays mounted on or towed behind surface ships or submarines, mounted on some underwater weapons, or suspended from air-deployed sonobuoys. Submarines are also detected with stationary bottom-mounted hydrophones, such as the Sound Surveillance System (SOSUS) arrays located

in the Atlantic and North Pacific oceans. Submarine detection systems may be active or passive. In a passive system the hydrophone detects acoustic signals made by the submarine. In an active system the hydrophone detects acoustic signals generated by the system sources which have been reflected by the submarine. Hydrophones are also used to detect and track marine mammals, to study global climate change by detecting and mapping temperature change in the world's oceans by means of the change in sound speed caused by the temperature change, to find schools of fish, to assess ocean biomass, and to prospect for offshore seismic oil. *See* ANTISUBMARINE WARFARE; OIL AND GAS, OFFSHORE; SONOBUOY; UNDERWATER SOUND. Peter H. Rogers; David H. Trivett

Bibliography. R. J. Bobber, *Underwater Electroacoustic Measurements*, Government Printing Office, Washington, DC, 1970, reprint, Peninsula Publishing 1990; J. W. Horton, *Fundamentals of Sonar*, United States Naval Institute Press, Annapolis, 1957; L. E. Kinsler et al., *Fundamentals of Acoustics*, 4th ed., Wiley, New York, 2000.

Hydroponics

Techniques for supplying nutrients and water directly to the roots of plants, without soil or other media. Methods that utilize an inert medium such as sand, gravel, peat, or vermiculite to provide the root environment, with water and nutrients added in solution, are soilless culture but are not hydroponic in the strict sense.

Hydroponic systems range in complexity from a single plant supported above an aerated jar of nutrient solution to thousands of plants supported above a large area of flowing solution in which pH, temperature, and nutrient concentrations are controlled by using a sophisticated computer system and automated chemical analysis. In hydroponic culture, precise control of the pH and the concentrations of elements in the solution is critical; all essential elements must be provided and in the correct ratios for plant growth. Elements known to be essential are nitrogen, phosphorus, potassium, calcium, magnesium, iron, manganese, boron, copper, molybdenum, and zinc. Oxygen in the root zone is also necessary. *See* PLANT MINERAL NUTRITION.

Hydroponic systems. Hydroponic systems offer a number of advantages when compared to soil culture. They reduce water, pH, and nutrient stress; yield clean roots and leaves; and facilitate rapid crop turnaround and automation. The disadvantages are that disease may spread more rapidly, pH and nutrient control are required, and initial expenses are higher. In theory, the growth-limiting factors in hydroponic systems are the availability of photosynthetic light and carbon dioxide. *See* PHOTOSYNTHESIS.

Among typical hydroponic systems are aerated standing culture, intermittent-flow culture, and continuous-flow culture. These techniques require careful preparation of the nutrient solution, continued monitoring, and adjustment or periodic replacement of the solution.

The aerated standing-culture technique (**Fig. 1***a*) is widely used in research and was the first of the hydroponic methods. In this system, plants are supported above an opaque container, with the roots submerged in a nutrient solution that is aerated by bubbling. Light must be excluded from the solution to avoid the growth of algae. Either the pH and nutrients must be adjusted frequently or the solution must be replaced daily to weekly, depending upon the volume per plant. Intermittent- and continuous-flow cultures use similar techniques for delivering the solution and differ only in the length of time that the roots are immersed. Such systems can be open (single-pass) or closed (recirculating), depending on the requirements for reuse of the solution.

The most widely used active solution-delivery techniques are flowing-solution culture with standing solution, the nutrient film technique, and aeroponics. The flowing-standing culture employs intermittent or continuous replacement of the standing solution in the root zone by using a tub or trough with an overflow drain pipe. The solution is pumped from a reservoir to the plant roots, overflows through the drain, and is either discarded or returned to the reservoir (Fig. 1*b*) for aeration, pH control, and replenishment of nutrients. The nutrient film technique is similar, except that the root mass lies at the bottom of the tray or trough (Fig. 1*c*) and is exposed to a thin film of solution that is introduced to the upper end of a slightly tilted tray. The solution flows across the bottom of the tray and leaves at the lower end, to be either discarded or returned to the reservoir. In early applications of this technique, a capillary matting was used to distribute nutrients to the roots. In aeroponics, the plant is suspended, and the nutrient is sprayed on the roots and allowed to drain off to be discarded or recycled (Fig. 1*d*). The last two techniques require less water in the root zone, and with adequate control, the total water requirements can be much reduced. However, there is a greater danger of crop loss in case of pump failure. Regardless of the technique, recirculation increases the danger of spreading root disease.

Research and applications. Because it makes accurate control of the root zone possible, hydroponic culture is widely used in research on plant nutrition and on the effect of temperature and pH on roots. Of specific interest is the role of different mineral nutrients. Hydroponics can also be used to study the effect of microbes on plant health.

Commercial systems have been improved by the use of plastics and computer control, but still tend to be limited to certain crops for which appearance is important, such as lettuce, spinach, and tomatoes. Such growing is typically done in greenhouses or similar controlled-environment buildings.

Hydroponic systems also have obvious value in the field of education and for the amateur horticulturist, who can grow flowers or vegetables in a

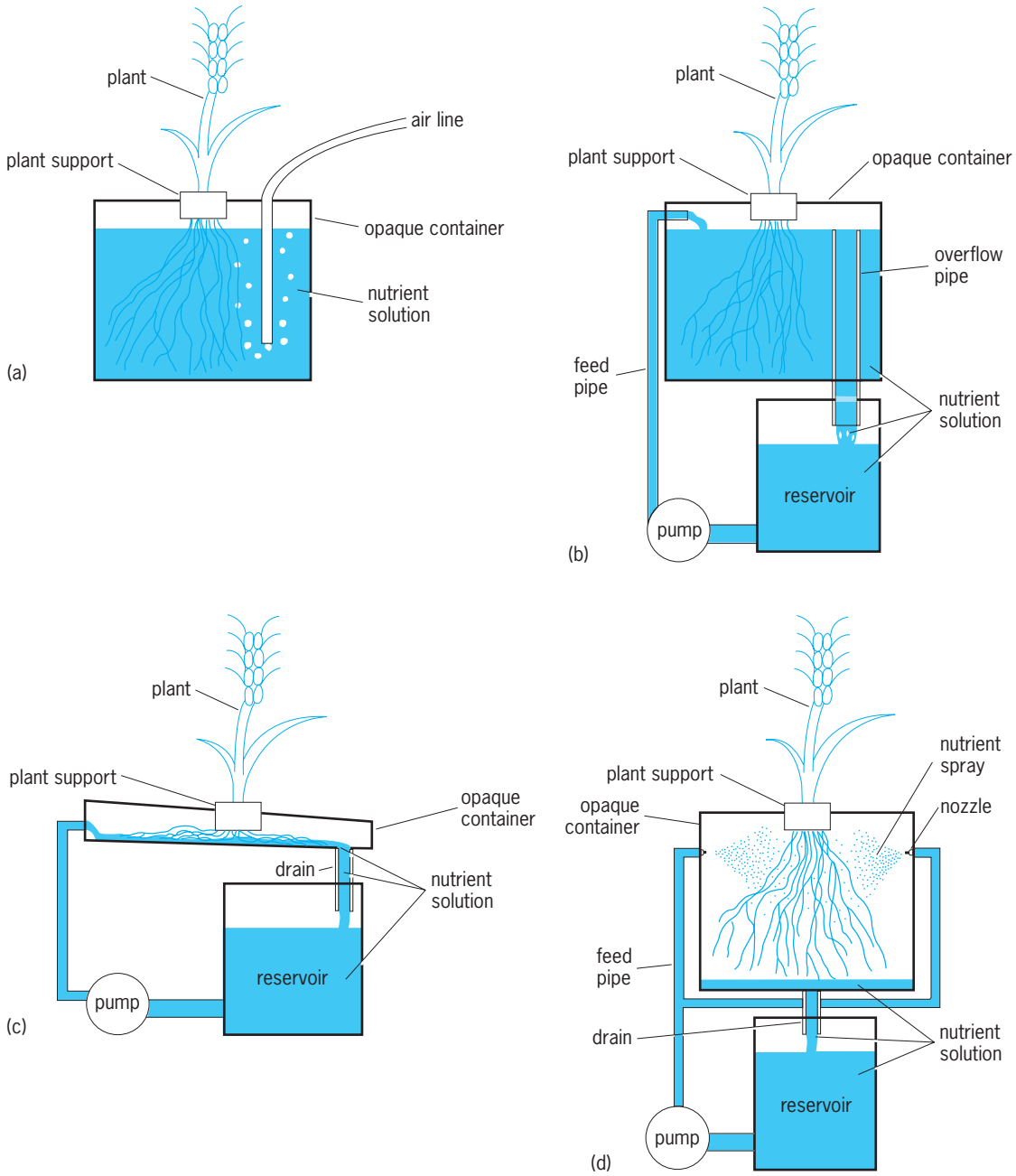

Fig. 1. Hydroponic techniques: (*a*) aerated standing culture; (*b*) flowing-solution culture with standing solution; (*c*) nutrient film technique (NFT); (*d*) aeroponics.

confined space with an indoor hydroponic garden. Hydroponic systems that have been developed for the NASA Controlled Ecological Life Support Systems (CELSS) program are used for crops not typically in hydroponic culture such as wheat, soybeans, white potatoes, and sweet potatoes. In a sealed, controlled-environment growth chamber at the Kennedy Space Center, crops are produced in a 120-ft^2 (16-m^2) growing space (**Fig. 2**), simulating a food production module for space. It is believed that conventional hydroponic systems on Mars or the Moon would produce oxygen and food and remove carbon dioxide, but that such systems could not be used in space because of the absence of the gravity needed for solution flow. Among the alternatives that have been proposed, the most promising system feeds nutrient solution to roots through a hydrophilic, microporous material. The roots grow on the surface of the material but do not penetrate it, and an opaque covering excludes light and retains moisture. A system using this technique at the Kennedy Space Center utilizes a tube of porous material contained within an opaque solid tube that shades and holds the roots, which surround the porous tube (**Fig. 3**). The nutrient solution is fed from a reservoir (bladder) through the porous tube under slight suction. Some solution moves by capillary action through the porous material to the roots; the rest is pumped back to the

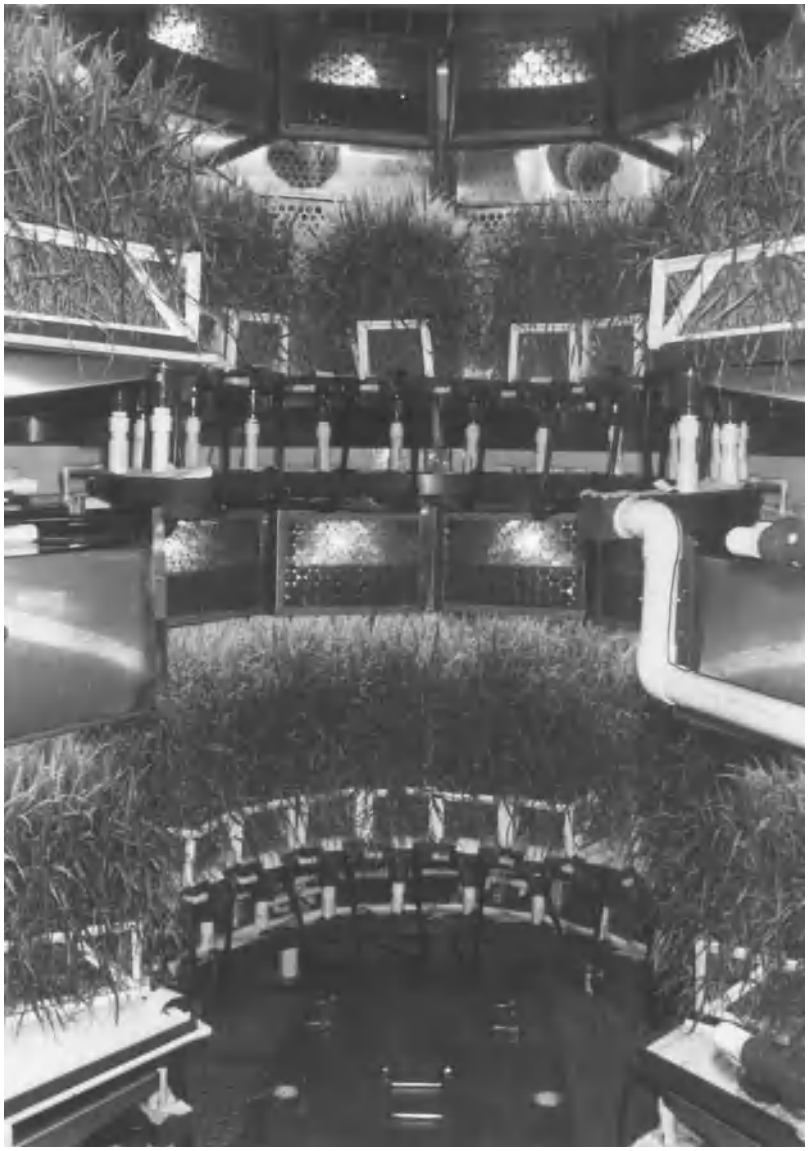

Fig. 2. Wheat grown under artificial lighting by using the nutrient film technique in the Biomass Production Chamber at the Kennedy Space Center. (*NASA*).

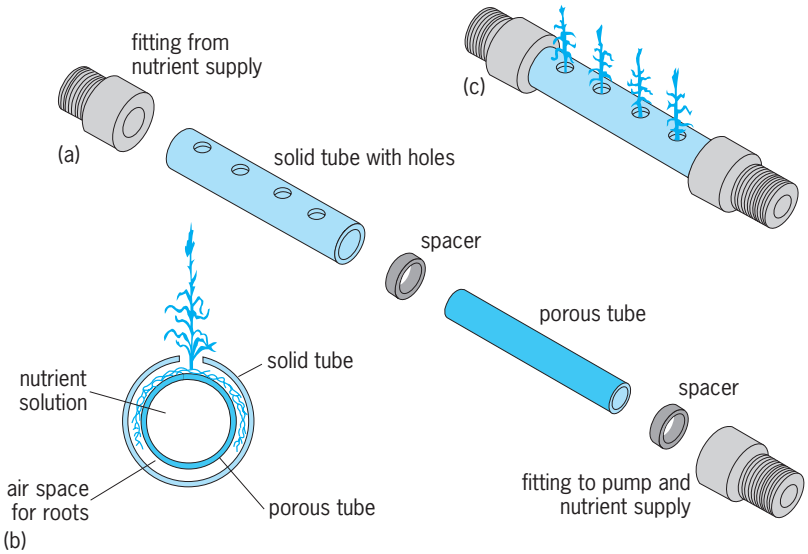

Fig. 3. Porous-tube plant-growth unit shown in (*a*) exploded view, (*b*) cross section, and (*c*) exterior view.

reservoir for control of pH and for replenishment of water and nutrients. Wheat, lettuce, beans, soybeans, and potatoes have been grown by using the system, with varying success. *See* PLANT GROWTH.

T. W. Dreschel

Bibliography. A. Cooper, *The ABC of NFT: Nutrient Film Technique*, 1979; T. W. Dreschel, *The Results of Porous Tube Plant Growth Unit Experiment T6B*, NASA Tech. Mem. 100988, 1988; D. R. Hoagland and D. I. Arnon, *The Water Culture Method for Growing Plants Without Soil*, Calif. Exp. Sta. Circ. 347, 1950; J. B. Jones, *A Guide for the Hydroponic and Soilless Culture Grower*, 1983.

Hydrosphere

The water portion of the Earth as distinguished from the solid part and from the gaseous outer envelope (atmosphere). Approximately 74% of the Earth's surface is covered by water, in either the liquid or solid state. These waters, combined with minor contributions from ground waters, constitute the hydrosphere.

The oceans account for about 97% of the weight of the hydrosphere, while the amount of ice reflects the Earth's climate, being higher during periods of glaciation. There is a considerable amount of water vapor in the atmosphere. The circulation of the waters of the hydrosphere results in the weathering of the landmasses. The annual evaporation from the world oceans and from land areas results in an annual precipitation of 320,000 km^3 (76,000 mi^3) on the world oceans and 100,000 km^3 (24,000 mi^3) on land areas. The rainwater falling on the continents, partly taken up by the ground and partly by the streams, acts as an erosive agent before returning to the seas. *See* GROUND-WATER HYDROLOGY; HYDROLOGY; LAKE.

The unique chemical properties of water make it an effective solvent for many gases, salts, and organic compounds. Circulation of water and the dissolved material it contains is a highly dynamic process driven by energy from the Sun and the interior of the Earth. Each component has its own geochemical cycle or pathway through the hydrosphere, reflecting the component's relative abundance, chemical properties, and utilization by organisms. The introduction of materials by humans has significantly altered the composition and environmental properties of many natural waters.

Rainwater. Rainwater contains small but measurable concentrations of many elements derived from the dissolution of airborne particulate matter and produced by equilibration of rainwater with atmospheric gases.

Total dissolved solids in rainwater range from over 10 parts per million (ppm) in rain formed in marine air masses to less than 1 ppm in rain precipitated over continental interiors. The major dissolved constituents of rainwater are chloride, sodium, potassium, magnesium, and sulfate (**Table 1**). These salts are derived over oceans and coastal areas from the dissolution of aerosol particles formed during the evaporation of sea spray. *See* AEROSOL.

TABLE 1. Chemical composition of waters, in parts per million

Constituent	Average rainwater	Average river water	Seawater
Sodium (Na)	1.98	6.3	10,500
Potassium (K)	0.3	2.3	380
Magnesium (Mg)	0.27	4.1	1,300
Calcium (Ca)	0.09	15.0	400
Chlorine (Cl)	3.79	7.8	19,000
Sulfate (SO_4)	0.58	11.2	2,650
Bicarbonate (HCO_3)	0.12	58.4	140
Silica (SiO_2)	—	13.1	6

A significant portion of the dissolved sodium, potassium, calcium, and sulfate in rain formed over continental areas is introduced by reaction with land-derived dust particles. Additional sulfate comes from the oxidation of sulfur dioxide, produced by the oxidation of hydrogen sulfide and by the burning of fossil fuels and smelting of sulfide ores [reaction (1)].

$$2SO_2 + O_2 + 2H_2O \rightarrow 4H^+ + 2SO_4^{2-} \qquad (1)$$

A map of the average sodium content of rain in the continental United States (**Fig. 1***a*) shows sodium contours subparallel to the coastlines, reflecting mixing of continental air masses with salt-rich marine air. The distribution of sulfate (Fig. 1*b*) is more complex, and reflects significant continental input from dust storms and industrial activity.

Rainwater contains in dissolved state each of the gases present in the lower atmosphere. Carbon dioxide (CO_2) is derived from both biological respiration and the burning of fossil fuels. It reacts with rain to form carbonic acid (H_2CO_3) [reaction (2)].

$$CO_2 + H_2O \rightarrow H_2CO_3 \qquad (2)$$

The presence of free oxygen and carbon dioxide makes rain both a natural oxidizing agent and an acid. Rain equilibrated with normal atmosphere has a pH of 5.7. More highly acidic rains form in areas where the industrial discharge of carbon dioxide or sulfur dioxide is intense. *See* ACID RAIN; PH.

Rain contains variable trace concentrations of many additional elements and compounds. Some of these, such as heavy metals and radionuclides, are derived from industrial pollution and nuclear testing, respectively. Precipitation of rain is the primary process by which many of these materials are transported from the atmosphere to the continents and oceans. *See* PRECIPITATION (METEOROLOGY).

Soil waters. As rainwater percolates downward and laterally through the soils and surface rocks of the continents, a complex group of reactions occurs.

The release of carbon dioxide and organic acids by bacterial processes increases the chemical reactivity of waters passing through the upper part of the soil zone. Weathering of most carbonate or silicate minerals generally involves acid attack, with carbonic

acid dominating [reactions (3) and (4)].

$$\underset{\text{Calcite}}{CaCO_3} + H_2CO_3 \rightarrow Ca^{2+} + 2HCO_3 \quad (3)$$

$$2\underset{\text{Illite}}{KAl_5Si_7O_{20}(OH)_4} + 2H_2CO_3 + 13H_2O \longrightarrow$$

$$5\underset{\text{Kaolinite}}{Al_2Si_2O_5(OH)_4} + 4H_4SiO_4 + 2K^+ + 2HCO_3^- \quad (4)$$

The weathering of carbonates and silicates consumes acid and produces a soil water enriched in cations, bicarbonate, and dissolved silica (H_4SiO_4). Dissolved sulfate is derived from dissolution of sulfate minerals and by reaction between sulfide minerals and dissolved oxygen in soil waters, as in reaction (5). Chloride is derived from the weathering of

$$4\underset{\text{Pyrite}}{FeS_2} + 15O_2 + 8H_2O \longrightarrow$$

$$2\underset{\text{Hematite}}{Fe_2O_3} + 8SO_4^{2-} + 16H^+ \quad (5)$$

fluid inclusions in silicate minerals and dissolution of halite (NaCl). *See* HALITE; SILICATE MINERALS; WEATHERING PROCESSES.

The ease with which a particular element can be accommodated in soil waters depends in part on the ionic radius r and charge Z of the cation that it forms (**Fig. 2**). The ratio of Z to r is known as the ionic potential. Large cations with a small charge, such as potassium (K^+) and calcium (Ca^{2+}), are usually readily accommodated in aqueous solution. In oxidizing environments, elements that form small, highly charged cations, such as the sulfur cation S^{6+}, combine with oxygen to form highly soluble and stable anionic complexes (for example, SO_4^{2-}). Cations of intermediate size and charge, however, including aluminum (Al^{3+}) and iron^{III} (Fe^{3+}), are only sparingly soluble. These elements are usually incorporated in the solid products of weathering, for example, aluminum in kaolinite in reaction (4) and iron in hematite in reaction (5). In soil waters depleted in free oxygen, the more highly soluble, reduced form of iron, Fe^{2+}, may go into solution. Other elements may also be solubilized in the absence of oxygen or in the presence of suitable complexing agents.

River water. River water represents a variable mix of subsurface waters, which enter the river at the ground-water table, and surface runoff. Some of the material in river water is derived from the dissolved sea salts and dust present in rainwater, but most has been introduced through weathering reactions. River waters are higher in bicarbonate and dissolved silica, and the relative abundance of the cations they contain reflects the lithology of the drainage basins from which they are derived (Table 1). Waters draining carbonate terranes are typically enriched in calcium and magnesium [reaction (3)].

Shale terranes, in contrast, will produce waters preferentially enriched in potassium, which is released during weathering of illite [reaction (4)]. The salinity of river waters varies from less than 40 ppm for the Amazon River, which drains a region of exceptionally high rainfall, to over 800 ppm for the Rio Grande, which drains a region of low rainfall and high evaporation. Dissolved organic material is high in tropical streams and rivers, where rates of organic production and decay are high. Organics in many rivers draining the southeastern United States exceed the concentration of dissolved inorganic salts. The composition salinity of a given river may vary seasonally.

Rivers and streams have been used by humans since earliest history as a source of potable water, a place to discard wastes, and a vehicle for the transportation of goods. The concentrations of many metals and organic compounds deliberately or

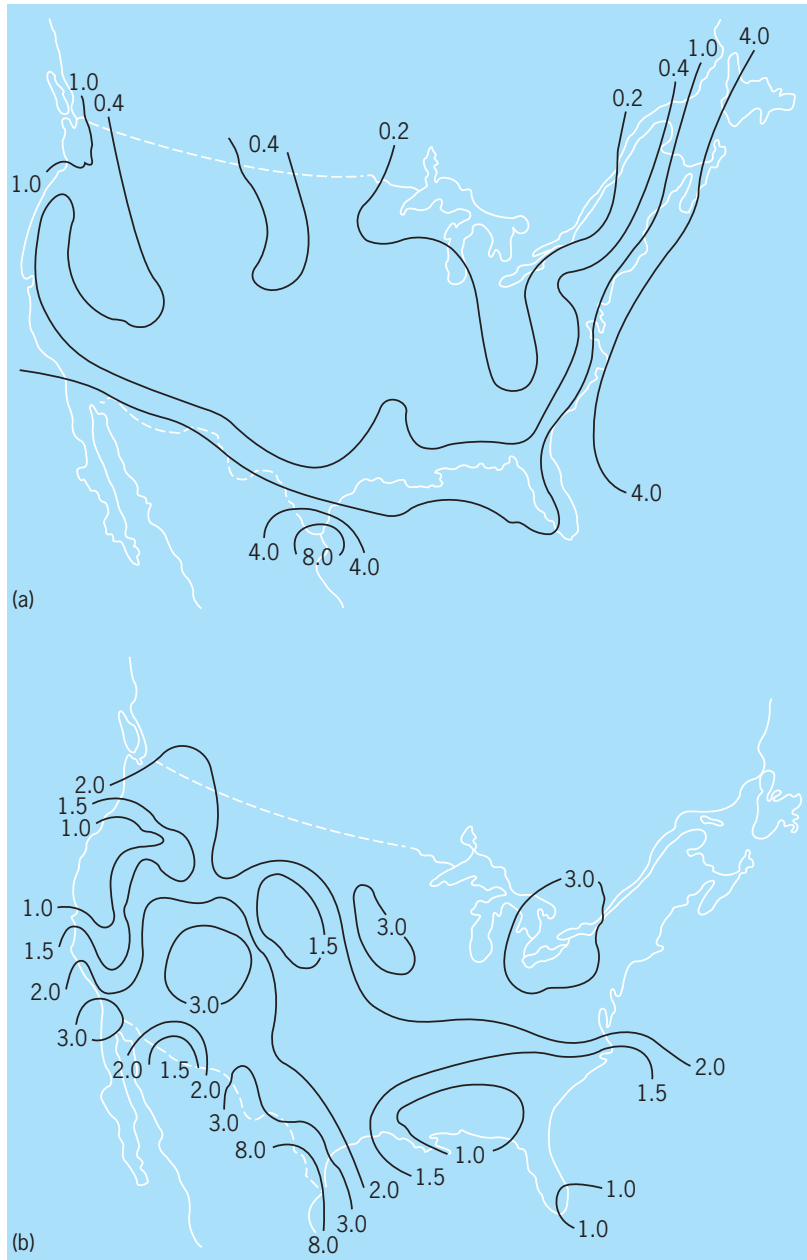

Fig. 1. Dissolved (*a*) sodium and (*b*) sulfate in rain (in parts per million) over the continental United States (*After R. M. Garrels and F. T. Mackenzie, Evolution of Sedimentary Rocks, W. W. Norton, 1971*)

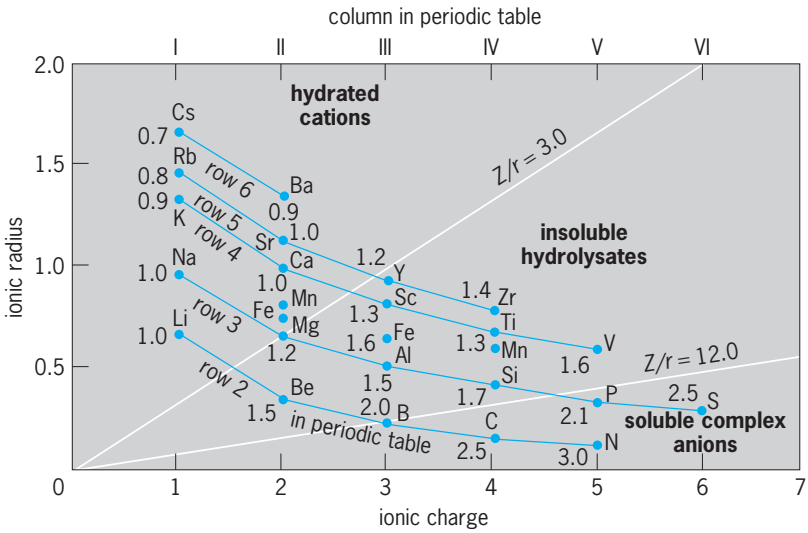

Fig. 2. Accommodation of cations in aqueous solution. (*After H. Blatt, G. Middleton, and R. Murray, Origin of Sedimentary Rocks, Prentice-Hall, 1972*)

accidentally introduced as wastes now often exceed natural river levels of these materials. Humans have also introduced compounds such as chlorinated hydrocarbons, which were unknown in the natural environment.

Most river water eventually mixes with marine waters in coastal and estuarine areas. The concentrations of most of the major cations and anions in these zones of mixing are not affected by processes other than the physical mixing of fresh and marine waters. Such constituents are said to behave conservatively. Many minor and trace constituents, however, behave nonconservatively, and are preferentially introduced into or removed from aqueous solution by chemical or biological processes occurring in the zone of mixing. Significant quantities of barium, for example, are desorbed from river clays when these particles are transported into marine waters. Humic-metal colloids present in river waters are flocculated as they mix with marine waters. The removal of dissolved iron by this process has been extensively documented. Field studies have shown that silica is removed from solution in some river estuaries. However, the question of whether this removal is due to biological uptake, reaction with suspended mineral particles, or both, has not been resolved. *See* RIVER.

Seawater. The dissolved salt content of open ocean water varies between 32,000 and 37,000 ppm. This range reflects dilution of seawater by rain and concentration by evaporation. Chloride, sodium, sulfate, magnesium, calcium, and potassium ions dominate sea salt (Table 1) and, with the exception of calcium, are present in remarkably constant proportions throughout the oceans. Other elements, such as boron, bromine, and fluorine, also show a constant ratio with chloride, but the chloride ratios of many elements vary significantly. The average pH of seawater is 8.2.

Most variations in the composition of seawater arise from the removal of elements by organisms that are living in surface seawater and the later release of these elements by the destruction of biologically produced particles which have sunk downward into deeper waters. Exceptions to this general rule are dissolved gases, whose solubility and concentration in surface seawater increase with decreasing temperatures.

Marine plants can live only in surface seawater, where sufficient light is available for photosynthesis. These organisms give off oxygen and extract carbon dioxide and nearly all of the dissolved nitrate and phosphate from seawater to produce organic matter. Some plants, in addition, secrete solid particles of calcium carbonate ($CaCO_3$) or opaline silica ($SiO_2 \cdot nH_2O$). Marine plants are consumed by animals, some of which also extract dissolved calcium, bicarbonate, and silica to make carbonate or opaline shells or tests. During the downward rain of particles produced by plants and animals in surface waters, destruction of organic matter by bacteria and animals releases dissolved nitrogen, phosphorus, and carbon dioxide back into the water column at depth and consumes dissolved oxygen (**Fig. 3**). Ocean waters are undersaturated with respect to opaline silica, and these particles begin to dissolve, releasing dissolved silica, after the death of their parent organism. Some particles, however, reach the sea floor to accumulate as siliceous oozes. Carbonate particles are stable in surface waters, which are supersaturated, and accumulate readily in shallower areas of the sea floor. Deeper waters are undersaturated because of increased pressure, and below depths of 4000 m (2.5 mi) the degree of undersaturation is such that carbonate dissolves very rapidly.

Thus, in response to biological processes, nitrogen, phosphorus, and silicon are almost totally depleted in surface waters, and marine plant life can flourish only where upwelling currents renew surface water in these biolimiting elements. Elements that show some lowering in concentration in surface waters are carbon, copper, nickel, and cadmium and the alkaline earths calcium, strontium, barium, and radium. The behavior of strontium, barium, and radium may reflect in part their coprecipitation with calcium in carbonate. Dissolved oxygen is unique in that it is produced at the surface and consumed at depth (Fig. 3). Analytical data for many elements are not precise enough to establish patterns of variation in their concentration. *See* UPWELLING.

In closed basins on the sea floor, stagnant bottom waters can become totally depleted in dissolved oxygen. In anoxic waters, anaerobic respiration reduces sulfate and forms hydrogen sulfide. Iron and manganese become more soluble and may increase in concentration, while other metals, such as copper, precipitate out as sulfides. *See* ANOXIC ZONES; SEAWATER.

Surface brines. Evaporation of fresh waters flowing into closed basins on the continents typically produces alkaline brines (**Table 2**; Soap Lake). Calcium and magnesium precipitate out as insoluble carbonates or hydroxysilicates. Sodium and potassium concentrate continuously, and total carbonate and

pH increase. Chemical evolution of marine waters during evaporation follows a different course. Gypsum ($CaSO_4 \cdot 2H_2O$) is the first mineral to precipitate out during continued evaporation, followed by halite (NaCl). Reaction with carbonates to form dolomite [$CaMg(CO_3)_2$] may remove magnesium (Table 2; Dead Sea). *See* SALINE EVAPORITES.

Subsurface waters. Marine waters trapped in the pore spaces of sediments during deposition react with the mineral and organic particles surrounding them and undergo significant changes in composition. Pore waters in organic-rich sediments are quickly depleted in dissolved oxygen, and anaerobic reduction of sulfur destroys dissolved sulfate and produces hydrogen sulfide. Anaerobic reduction of carbon dioxide in the absence of sulfate produces methane. Changes in the relative proportions of dissolved cations occur as a result of diagenetic reactions with silicate minerals. The vertical variation in concentration of dissolved species in sediment pore waters provides valuable information on the nature and rates of diagenetic reactions in fine-grained marine sediments.

TABLE 2. Chemical composition of some saline waters, in parts per million

Constituent	Soap Lake, Washington	Dead Sea, southwest Asia	Subsurface brine, Louisiana
Sodium (Na)	12,500	39,700	63,900
Potassium (K)	12,500	7,590	869
Magnesium (Mg)	23	42,430	1,070
Calcium (Ca)	4	17,180	9,210
Chlorine (Cl)	4,680	219,250	124,000
Sulfate (SO_4)	6,020	420	153
Bicarbonate (HCO_3)	11,270	220	115
Carbonate (CO_3)	5,130	—	—
Silica (SiO_2)	101	—	16

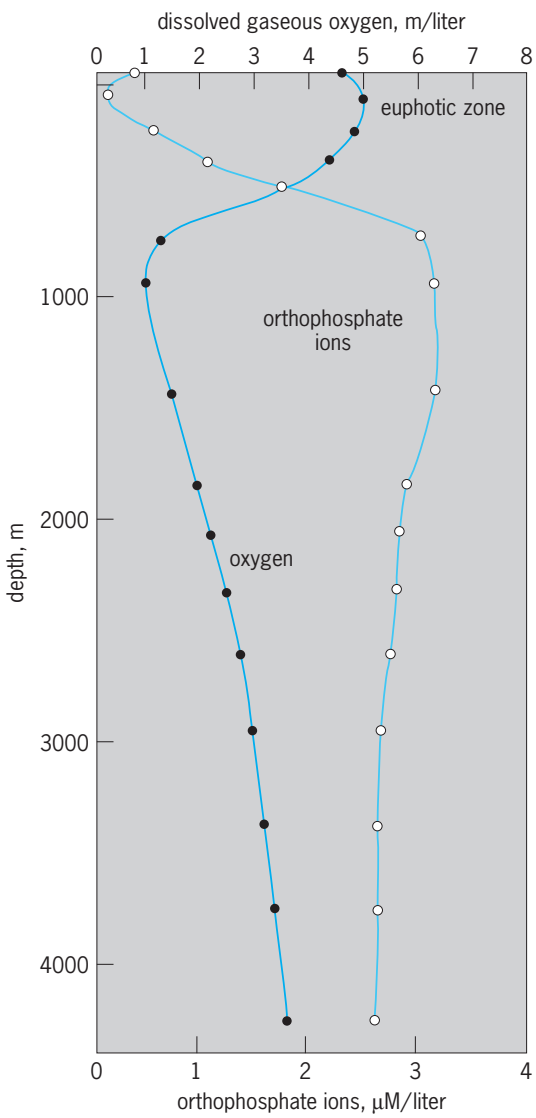

Fig. 3. Distribution of dissolved gaseous oxygen (low values) and nutrient species (high concentrations) of orthophosphate ions at 26°22′.4 N and 168°57′.5 W in the Pacific Ocean. 1 m = 3.3 ft.

Subsurface waters in deep sedimentary basins are often far more saline than the connate pore waters trapped at the time of sediment deposition. Salinities of 100,000–150,000 ppm are common but can exceed 400,000 ppm. Basinal brines are typically dominated by sodium and chloride, but calcium becomes the most abundant cation at extreme salinities. Subaerial evaporation of marine and continental waters and the subsurface dissolution of halite can produce the range of salinities and dissolved chloride concentrations observed for most subsurface basinal brines, but not their major cation compositions (Table 2). The observed systematic increases in individual dissolved major cations and decreases in pH and alkalinity with increasing salinity observed in deep sedimentary waters worldwide support the hypothesis that the approach toward thermodynamic buffering by silicate-carbonate mineral assemblages is the first-order control on subsurface fluid compositions, even at temperatures well below 100°C (212°F).

Dissolved organic acid anions, such as acetate, are associated primarily with low-salinity basinal waters; and dissolved ore-forming metals, such as copper, lead, and zinc, are preferentially found in basinal brines having salinities in excess of 200,000 ppm. The high chloride concentration and low pH of these saline waters may enhance solubilization of metals through chloride complexing. Long-distance migration of brines and other subsurface fluids appears to be a common aspect of the geological evolution of sedimentary basins. *See* BASIN; DIAGENESIS.

Ice. Ice is a nearly pure solid, and in contrast to the solvent power of liquid water, few foreign ions can be accommodated in its lattice. Ice does contain particulate matter, however, and the change in the composition of these particles with time, as recorded in the successive layers of ice that have accumulated in polar regions, has provided much information on the progressive input of lead and other materials into the environment by humans. *See* SEA ICE.

Jeffrey S. Hanor

Bibliography. E. K. Berner and R. A. Berner, *The Global Water Cycle: Geochemistry and Environment*, 1987; J. I. Drever, *The Geochemistry of*

Natural Waters, 3d ed., 1997; J. S. Hanor, *Origin of Saline Fluids in Sedimentary Basins*, 1994.

Hydrostatics

The study of fluids at rest; that is, at every point in the fluid, the velocity and acceleration are zero. The only force acting at a point on the surface of a small volume of fluid is due to the pressure at that point, and the force is perpendicular to the surface. Application of Newton's second law shows that the pressure is independent of the orientation of the surface on which it is acting, and that the pressure increases with depth.

The pressure difference between any two points in the fluid is given by Eq. (1), where p_1 and p_2 are the

$$p_2 - p_1 = \int_{z_1}^{z_2} \rho g \, dz \tag{1}$$

pressures at the two points, z_1 and z_2 are the vertical coordinates (z positive downward) of the two points, ρ is the density of the fluid, and g is the acceleration due to gravity. The integration on the right-hand side can be carried out only if a relationship between density and elevation is known. Variations of g with elevation are negligible in earth-bound engineering problems. If the density of the fluid is constant, Eq. (1) becomes Eq. (2), where $h = z_2 - z_1$ is the

$$p_2 - p_1 = \rho g(z_2 - z_1) = \rho g h \tag{2}$$

vertical distance between the two points. The pressures p_2 and p_1 may be either absolute pressures or gage pressures. The absolute pressure is the sum of the gage pressure and atmospheric pressure. *See* PRESSURE.

Manometers. Equation (2) shows that a pressure difference may be expressed in terms of the height h of a column of fluid of density ρ. A manometer is a device that makes use of this equivalence. An example of a U-tube manometer with its attendant equation is shown in **Fig. 1**. Manometers are available in a wide variety of configurations. Each configuration has its own attendant equation, but all of the equations are derivable from Eq. (2). *See* MANOMETER; PRESSURE MEASUREMENT.

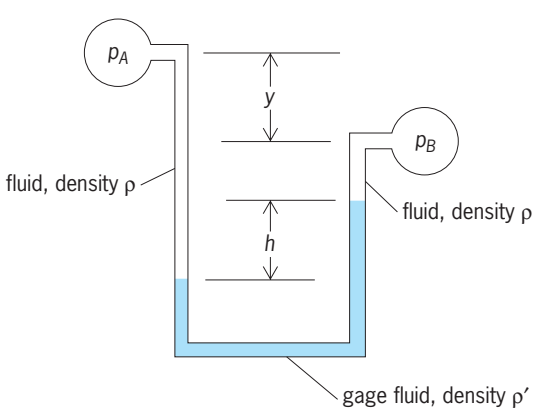

$$p_A - p_B = gh(\rho' - \rho) - \rho g y$$

Fig. 1. U-tube manometer and its attendant equation for equation for determining pressure difference $p_A - p_b$.

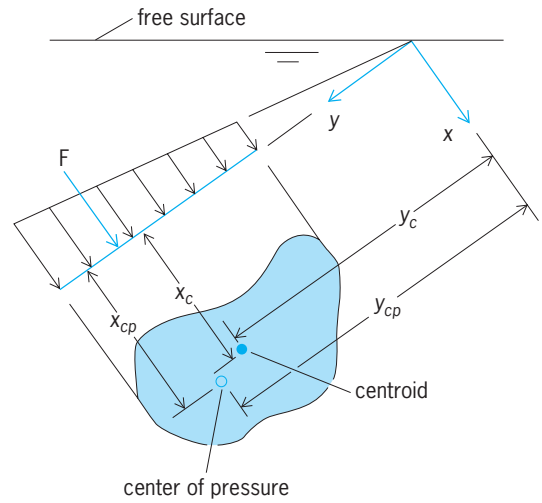

Fig. 2. Fluid force on a plane surface. Symbols are explained in text.

Forces on plane submerged surfaces. The design of tanks, dams, and other devices for containing fluids at rest requires the calculation of the force exerted by the fluid on the walls of the container. The force is calculated by integrating the pressure distribution over the submerged area. **Figure 2** illustrates the pressure distribution on a plane submerged area and shows the notation used. The coordinates are chosen so that the plane of the area is the xy plane, and the intersection of this plane with the free surface is the x axis.

For a fluid of constant density, the magnitude of the force exerted by the fluid is given by Eq. (3), where

$$F = p_c A \tag{3}$$

p_c is the pressure at the centroid of the submerged area and A is the submerged area. The force **F** is distributed over the entire submerged area, but it is convenient to replace it with a concentrated equivalent force. The point of application of the concentrated equivalent force is called the center of pressure. The location of the center of pressure is given by Eqs. (4),

$$\begin{aligned} y_{cp} &= y_c + \frac{I_{yc}}{y_c A} \\ x_{cp} &= x_c + \frac{I_{xyc}}{y_c A} \end{aligned} \tag{4}$$

where x_c and y_c are the coordinates of the centroid of the submerged area, I_{yc} is the moment of inertia of the submerged area with respect to an axis parallel to the x axis through its centroid, and I_{xyc} is the product of inertia with respect to centroidal axes. *See* CENTROIDS (MATHEMATICS); MOMENT OF INERTIA; PRODUCT OF INERTIA; STATICS.

Forces on curved submerged surfaces. The easiest way to calculate the force exerted on a curved submerged surface (**Fig. 3**) is to calculate the horizontal and vertical components of the force and then

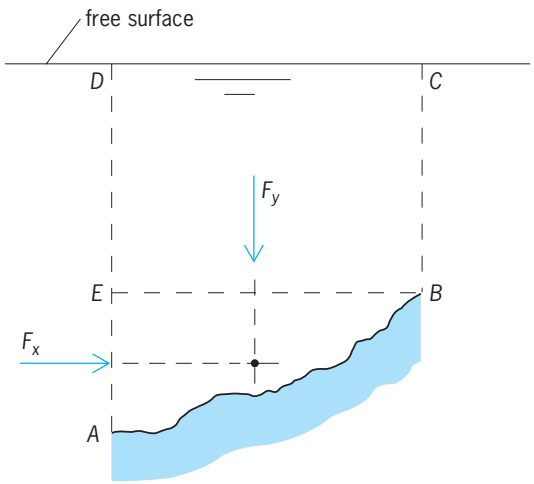

Fig. 3. Fluid force on a curved surface. Symbols are explained in text.

combine them into a single force, if necessary. The horizontal component F_x is determined by calculating the force on a projected area which is equal to the projection of the curved area on a vertical plane (area *EA* in Fig. 3). The horizontal force component is calculated by using Eqs. (3) and (4) applied to this projected area.

If the fluid is above the curved surface, the vertical force component F_y is equal to the weight of the fluid above the surface (*DEABCD* in Fig. 3), and its line of action passes through the center of gravity of this weight of fluid. If the fluid is below the curved surface, the vertical force component acts upward and is equal to the weight of the fluid displaced by the curved surface. The line of action of the upward force passes through the center of gravity of the displaced fluid.

Submerged and floating bodies. The pressure distribution on a floating or submerged body gives rise to a net upward force which is called the buoyant force F_B. This buoyant force is equal to the weight of the fluid displaced by the body. This equality is known as Archimedes' principle. If the weight of the body W is greater than the buoyant force, the body will sink to the bottom of the fluid. If the weight of the body is equal to the buoyant force, the body will float. The line of action of the buoyant force is through the center of gravity of the displaced fluid, which, in this case, is called the center of buoyancy. *See* ARCHIMEDES' PRINCIPLE; BUOYANCY.

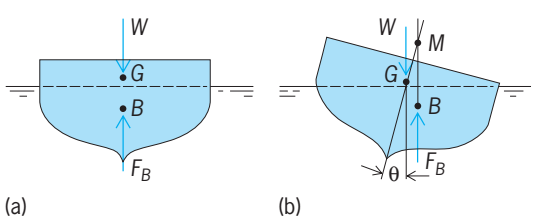

Fig. 4. Stability of a floating body (*a*) at equilibrium position and (*b*) at angular displacement θ from equilibrium position. Symbols are explained in text.

Stability of floating bodies. The stability of a floating body depends on the position of the center of buoyancy B relative to the center of gravity G of the body (**Fig. 4**). The position of the center of gravity of a body depends on its shape and on the distribution of mass within that shape. The position of the center of buoyancy, since it is the center of gravity of the fluid displaced by the body, depends only on the shape of that part of the body which is submerged.

For a particular body in Fig. 4, an angular displacement θ from the equilibrium position produces a couple that tends to return the body to its equilibrium position. This body has a positive righting moment, and the equilibrium is stable. The point M located at the intersection of the originally vertical line through G and the vertical line through the center of buoyancy B is called the metacenter, and the distance GM, the metacentric height. As the angle θ is increased, the shape of the displaced volume of fluid changes, and as a result the position of the center of buoyancy changes. If the angle θ is made large enough, the point M will lie below the point G and the righting moment will become negative. In this position the equilibrium is unstable and the body will capsize. *See* SHIP DESIGN. Warren M. Hagist

Bibliography. R. L. Daugherty, J. B. Franzini, and E. J. Finnemore, *Fluid Mechanics*, 9th ed., 1997; R. W. Fox and A. T. McDonald, *Introduction to Fluid Mechanics*, 5th ed., 1998; V. L. Streeter and E. S. Wylie, *Fluid Mechanics*, 9th ed., 1998; F. M. White, *Fluid Mechanics*, 4th ed., 1998.

Hydrothermal ore deposits

The predominant sources of scarce metals, such as copper, zinc, lead, tin, molybdenum, antimony, tungsten, mercury, bismuth, uranium, silver, and gold. Fascinating because of their variety, they are the subject of intense research aimed at understanding them better and finding them efficiently. *See* ORE AND MINERAL DEPOSITS.

Formation. Hydrothermal ore deposits are inferred to have formed by mineral precipitation from hot hydrous fluids. This deposition occurred within the rocks through which these fluids traveled, or as the fluids entered bodies of water (ocean or lakes), or as they reached the atmosphere at the Earth's surface. These inferences are based on the following observations: (1) Most hydrothermal ores occur in places that favor the passage of fluids: fracture zones (veins), pipelike conduits (chimneys, pipes) in easily soluble rocks (such as limestone), sedimentary beds that are porous or susceptible to chemical replacement (mantos), fractured caps over intrusive cupolas (porphyries), and porous calcsilicates (skarns) at contacts of magmatic intrusives with limestone. Other hydrothermal ores occur where fracture zones reached lake or ocean waters (as along mid-ocean ridges) or the atmosphere (as in hot springs above presently active geothermal reservoirs). (2) Hydrothermal minerals often contain small fluid inclusions consisting

mostly of water with some carbon dioxide; this water has significant concentrations of common salt and of several of the metals commonly found in hydrothermal deposits. These inclusions are interpreted as trapped remnants of the mineralizing fluids. Such fluid inclusions tend to homogenize between 50 and 500°C (120 and 930°F), suggesting that their host minerals formed at temperatures above the homogenization temperatures from one-phase fluids. (3) The original composition of hydrothermal ore host rocks was generally modified by the formation of water-bearing minerals (hydrothermal alteration). (4) Many hydrothermal deposits are in the vicinity of magmatic intrusive rocks, implying that they formed either from fluids emanating from a crystallizing magma or from fluids put in motion by magmatic heat. *See* LIMESTONE; MID-OCEANIC RIDGE; PORPHYRY; SEDIMENTARY ROCKS; SKARN.

Hydrothermal fluids may be generated in different ways: (1) as water expelled from a crystallizing magma; (2) as meteoric water that becomes ground water and is mobilized by gravity, tectonic activity (for example, tilting or folding), or magmatic heat and reacts with continental crustal rocks; (3) as seawater that infiltrates near mid-ocean ridges, is mobilized by ocean ridge magmatism, and reacts with oceanic crust; or (4) as water trapped in sediments (connate water) and released upon compaction and heating (diagenesis and metamorphism). *See* DIAGENESIS; MAGMA; METAMORPHISM.

Composition of ore and host rocks. The ore minerals in hydrothermal deposits are mostly sulfides of copper, zinc, lead, tin, molybdenum, antimony, mercury, bismuth, and silver. However, some metals occur as oxides (tin and uranium), tellurides (gold), tungstates (tungsten), or native (gold). The accompanying nonvaluable gangue minerals consist mostly of sulfides (pyrite, pyrrhotite, orpiment, realgar), oxides (quartz), carbonates (calcite), sulfates (barite), fluorides (fluorite), and silicates (rhodonite). *See* MINERAL; SULFIDE AND ARSENIDE MINERALS.

Hydrothermal alteration. Hydrothermal alteration of the surrounding host rocks due to contact with the hot hydrous fluids may extend from a few meters to a few kilometers from an ore. The alteration generally consists of addition of silica and conversion of anhydrous silicates (orthoclase, plagioclase, olivine) to hydrous silicates (muscovite, sericite, kaolinite, biotite, epidote, chlorite) by a low-pH fluid. Where present, the sulfate mineral alunite indicates that the fluid was oxidized and acidic. Where magma intruded limestone, hydrothermal fluids promoted the formation of porous silicates (garnet, wollastonite)—the skarns that host hydrothermal ores of copper, zinc, gold, or tungsten.

Mineralization pulses and stages. Hydrothermal ores often contain specific minerals or mineral groups (paragenetic associations or assemblages) that either form successive layers or cut prior minerals, thus defining a paragenetic sequence of mineralization stages. Thin sections of transparent minerals and etching or electron microprobe analyses of opaque minerals reveal that each mineral consists of many hundreds of growth bands, proving that an ore deposit is made by hundreds or thousands of mineralization pulses.

Zoning. The chemical composition of hydrothermal ores often varies spatially within a mining district or within a given vein, manto, pipe, or disseminated deposit. Over a century ago, this led to a zonal theory, according to which certain minerals or elements were considered to be characteristic of the upper or peripheral parts (chalcedony, barite, fluorite, mercury, antimony, gold-silver), others of the central part (silver, lead, zinc, copper), and still others of the inner or lower part (pyrrhotite, arsenopyrite, tourmaline, cassiterite, gold, arsenic, bismuth, tungsten, tin). This was only a generalized scheme, because every mineral deposit and district has its own zoning sequence.

Hydrothermal zoning has traditionally been considered to result from progressive changes in the temperature, pressure, or composition of the hydrothermal fluids along their paths. However, field observations have so far failed to provide convincing evidence that a specific paragenetic stage (S-1) evolves chemically (from composition A to compositions B and C) from one part of a deposit to another (**illus.** *a*). It seems more likely that successive stages (S-1, S-2, and S-3) are deposited by different fluids (A, B, C) at various points along the hydrothermal fluid paths, thus generating a zoned deposit (illus. *b*).

Oxidation, leaching, and supergene enrichment. The near-surface part of most hydrothermal ore deposits contains hydroxides and oxides of iron, copper, and manganese (limonite, hematite, cuprite, tenorite, pyrolusite, manganite), sulfate of lead (anglesite), carbonates of copper, lead, and zinc (malachite, azurite, brochantite, cerussite, smithsonite), and chloride of silver (cerargyrite). This oxide zone passes in depth into the underlying hypogene sulfides. Copper was often leached from the oxide zone, but precipitated again as a sulfide (chalcocite, covellite) between the sulfide and oxide zones to form a supergene enrichment zone or blanket. This sequence results from the near-surface atmospheric oxidation and dissolution of most of the original hypogene sulfides, which leads to the formation of acid ground waters. The downward-percolating ground waters are reduced at depth, just below the water table, leading to the precipitation of the supergene copper sulfides. Other factors influencing near-surface processes are the reactivity of the host rock to the acid ground water (for example, limestone is much more effective in neutralizing acid ground water than igneous, metamorphic, or other sedimentary rocks) and changes in the water table due to climatic or sea-level variations. Supergene enrichment of copper played an important role in making it economic to mine disseminated (porphyry) copper deposits in the southwestern United States, Mexico, and northern Chile. Residual enrichment in the oxide zone of relatively insoluble gold, and lead or zinc carbonates, may have played a role in the discovery of some deposits. *See* CARBONATE MINERALS; OXIDE AND HYDROXIDE MINERALS.

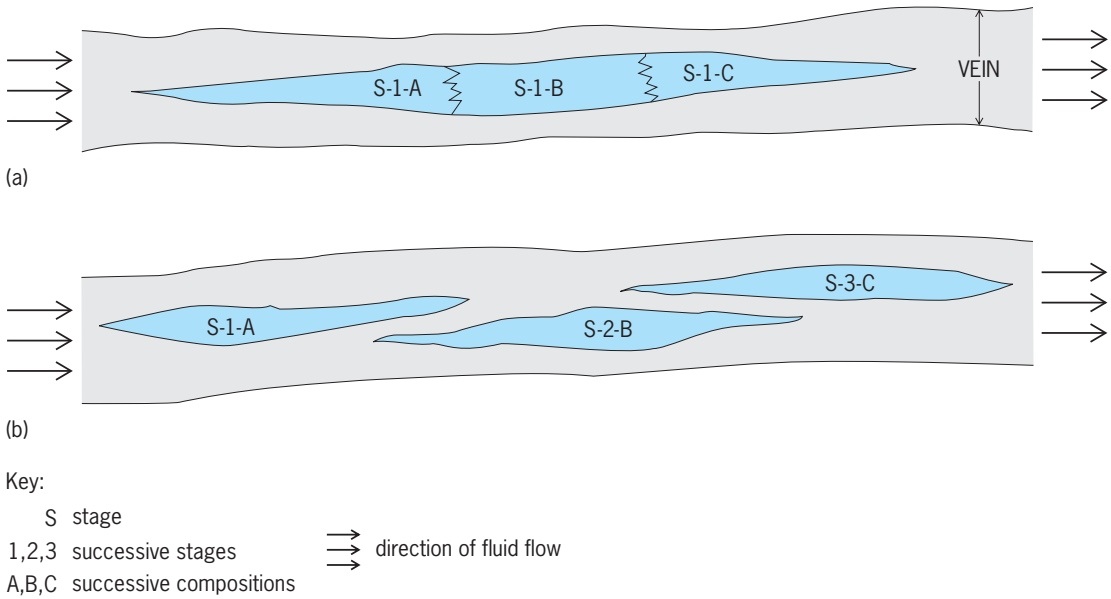

Alternative explanations of zoning in hydrothermal ore deposits. (*a*) A single-stage fluid (S-1) evolves along its path, depositing successively different mineral compositions or assemblages (A, B, and C). (*b*) Three different fluids flow through the vein at various times or stages (S-1, S-2, and S-3), depositing various mineral compositions or assemblages (A, B, and C).

Ore deposition factors. Perhaps more important than the sources of the metals and of the hydrothermal fluids are the factors determining the uptake and transport of ore constituents in hydrothermal fluids, as well as the causes of hydrothermal ore deposition. It has long been known that elevated temperatures and acidity (low pH) promote increased metal solubility. Experimental studies also highlighted the importance of chloride, sulfide, organic, and other complexing agents in metal solubility. Thus, magmas contaminated by major evaporite sequences (which contain chloride in halite, and sulfide in gypsum or anhydrite) would produce more "fertile" hydrothermal fluids than uncontaminated magmas. Similarly, heated sedimentary basin brines or infiltrating seawater would be more likely to dissolve and transport significant amounts of metals than dilute ground waters. In contrast, hydrothermal ore deposition would require cooling or neutralization (increasing pH) of the metal-bearing fluids.

Temperature. In the early 1900s, the temperature of the hydrothermal fluids was considered to largely determine the mineralogical composition of an ore deposit. This led to a temperature classification of hydrothermal ore deposits (epithermal, leptothermal, mesothermal, hypothermal) based mainly on minerals deemed to be indicative of the temperature of ore deposition. However, subsequent laboratory studies of pertinent chemical systems demonstrated that supposedly "diagnostic" minerals are stable over wide ranges of temperature and are, therefore, generally unreliable as geothermobarometers. For example, pyrrhotite (FeS_{1-x}) had been thought to deposit at a higher temperature than pyrite (FeS_2), but the experimental studies demonstrated that both are stable over the general range of hydrothermal ore deposition (and that the deposition of one or the other depends more on the concentrations of sulfur and iron in the hydrothermal fluid). Unfortunately, reliable geothermometers, such as temperature ranges of different polymorphs of a mineral (for example, argentite and acanthite) or incompatible mineral assemblages (pyrite-magnetite versus pyrrhotite-hematite), only indicate large temperature ranges, are generally scarce, or are restricted to unusual ore deposition temperatures. Studies of fluid inclusion homogenization temperatures and of light-isotope fractionations have shown that hydrothermal ore deposition occurs at temperatures from several hundred to a few tens of degrees Celsius. In addition, they showed that in some ore deposits there was a fast initial increase and a subsequent gradual decrease in temperature, as evidenced by the depositional temperatures of successive paragenetic mineralization stages. These studies have so far failed to clearly demonstrate thermal gradients along hydrothermal solution paths for a single depositional stage. At present, the temperature gradient factor is supported mainly by pointing out the temperature gradients observed in geothermal fields and by noting mineral precipitation by cooling solutions in laboratory experiments.

Acidity. During the mid-1900s, it became evident that the hydrothermal alteration of rocks adjoining veins and disseminated ore results largely from attack by acid fluids. This has been reinforced by abundant laboratory evidence, and a pH increase due to the interaction of hydrothermal fluids with host rocks is now seen as a more general cause of hydrothermal ore deposition, especially for lead-zinc ores in limestone. In some geothermal fields (for example, Yellowstone), hot springs of substantially different

pH occur in proximity, suggesting the possibility that fluid mixing may increase the pH of an acid metal-bearing fluid and cause mineral precipitation. Fluid mixing also may be responsible for mineral precipitation by bringing together ions that form insoluble compounds.

Pressure. Decreasing pressure gradients have been invoked as causes for mineral deposition because hydrothermal fluids tend to flow toward low-pressure regions. Fluid inclusion studies have shown that some minerals precipitated from boiling solutions (because different inclusions in a given mineral trapped varying proportions of liquid to water, thus giving an apparent range of homogenization temperatures). However, this appears to be an occasional rather than a predominant factor.

Most chemical equilibria vary with both temperature and pressure (that is, they are really geothermobarometers), so that reliable pressure gradients can be determined only if a pressure-independent geothermometer is available (such as a light-isotope fractionation).

Oxidation and reduction. Reduction of oxidized hydrothermal fluids is generally credited for the deposition of uranium, vanadium, copper, or silver ores in redbeds containing organic matter. Conversely, oxidation of hydrothermal fluids is also considered to be occasionally important in ore precipitation. *See* REDBEDS.

Fugacity. The decrease in emphasis on depositional temperature was supplanted by an increased awareness of the roles played by the fugacities of sulfur, oxygen, and carbon dioxide in hydrothermal solutions. (Fugacity is the partial pressure of a component in a perfect gas that is in equilibrium with a liquid or solid containing this gas in solution.) Eventually, this led to the current tendency to classify ore deposits as low-sulfidation, high-sulfidation, or acid-sulfate. But given that such expressions provide little information on the nature of a specific ore deposit, and given the difficulty involved in pinpointing the exact causes of ore deposition, there has been a parallel development of a hodge-podge terminology of ore deposit types, such as "Cordilleran vein, manto, pipe or breccia," "five-element (Ag-Ni-Co-Fe-As) vein," "disseminated or porphyry," "skarn," "volcanogenic massive sulfide," "redbed or sediment-hosted stratiform Cu-Ag," and "igneous-metamorphic." *See* FUGACITY. Ulrich Petersen

Bibliography. H. L. Barnes, *Geochemistry of Hydrothermal Ore Deposits*, 3d ed., Wiley, 1997; J. M. Guilbert, Linkages among hydrothermal ore deposit types, *Pro-Explo 2001*, Lima, Peru, 2001; J. M. Guilbert and C. F. Park, Jr., *The Geology of Ore Deposits*, W. H. Freeman, 1986; H. D. Holland, Some applications of thermochemical data to problems of ore deposits, I. Stability relations among the oxides, sulfides, sulfates and carbonates of ore and gangue minerals, *Econ. Geol.*, 54:184–233, 1959; H. D. Holland, Some applications of thermochemical data to problems of ore deposits, II. Mineral assemblages and the composition of ore-forming fluids, *Econ. Geol.*, 60:1101–1166, 1965.

Hydrothermal vent

A hot spring on the ocean floor, where heated fluids exit from cracks in the Earth's crust. Most hydrothermal vents occur along the central axes of mid-oceanic ridges, which are underwater mountain ranges that wind through all of the deep oceans. The best-studied vents are at tectonic spreading centers on the East Pacific Rise and at the Mid-Atlantic Ridge. However, vents are also found over hot spots such as the Hawaiian Islands and Iceland, in back-arc basins such as those in the western Pacific, in shallow geothermal systems such as those off the Kamchatka Peninsula, and on the flanks of some underwater volcanoes and seamounts. Major parts of the mid-oceanic ridges, including those of the Indian Ocean and Southern Hemisphere, are poorly explored, but available evidence suggests that these areas should also have extensive hydrothermal venting. Hydrothermal vent sites, or closely grouped clusters of vent deposits and exit ports, may cover areas from hundreds to thousands of square feet (tens to hundreds of square meters). Individual vent sites may be separated along mid-ocean ridges by more than 1000 mi (1600 km). *See* MID-OCEANIC RIDGE; SEAMOUNT AND GUYOT.

Hydrothermal fluid formation. All of the hydrothermal vent sites occur in areas where quantities of magma exist below the sea floor (**Fig. 1**). Cold seawater is drawn down into the oceanic crust toward the heat source. As the seawater is heated and reacts with surrounding rock, its composition changes. Sulfate and magnesium are major components of seawater lost during the reactions; sulfide, metals, and gases such as helium and methane are major components gained. This modified seawater is known as hydrothermal fluid. Buoyant, hot hydrothermal fluid rises toward the sea floor in a concentrated zone of upflow to exit from the sea floor at temperatures ranging from 50°F (10°C) to greater than 750°F (400°C), depending on the degree of cooling and of mixing with seawater during the ascent. If the sea floor is shallow enough and the fluid hot enough, the solution may boil; but it usually does not because of the pressure of overlying seawater at depths of 4900–9800 ft (1500–3000 m) or more. *See* MAGMA.

Fluid venting. Hydrothermal fluid that mixes extensively with seawater below the sea floor surface may reach the sea floor as warm springs, with temperatures of 50–86°F (10–30°C). This outflow is usually detectable as cloudy or milky water, but the flow is slow and no mineral deposits accumulate except for some hydrothermal staining or oxidation of sea floor basalts. When hotter, relatively undiluted hydrothermal fluid reaches the sea floor, it is still buoyant with respect to seawater, so that the hot solution rises out of cracks in the sea floor at velocities up to about 6 ft (2 m) per second, mixing turbulently with seawater as it rises. Mixing of hydrothermal fluid with seawater leads to precipitation of minerals from solution, forming mineral deposits at the exit from the sea floor and so-called smoke, tiny mineral particles suspended in the rising plume of fluid. Black

smoker vents are distinguished by the presence of such large quantities of minute mineral particles that the plumes become virtually opaque (**Fig. 2**).

The roughly cone-shaped plume of hydrothermal fluid continues to rise while mixing with seawater until the density of the mixture is equal to that of the surrounding seawater. It then spreads horizontally as a layer within the water column that is detectable by chemical analysis for trace components such as helium-3, an isotope originating in the Earth's interior, or methane (CH_4). Detection of this hydrothermal layer, which can be traced back to the general area of its source, is one of the techniques for detecting large areas of hydrothermal venting.

Formation and outflow of hydrothermal fluid makes a major contribution to the concentration and balance of elements in the oceans by changing the composition of seawater. The quantities of elements added or removed from the oceans by hydrothermal venting around the world are comparable to quantities contributed by the worldwide flow of rivers into the oceans. Hydrothermal venting also represents a major flow of heat from the Earth's crust and a major mechanism for cooling of new oceanic lithosphere. *See* LITHOSPHERE; SEAWATER.

Mineral deposits. Mineral deposits form at hydrothermal vents as the result of mixing between hot hydrothermal fluids and cold seawater, which leads to mineral precipitation. These deposits provide an opportunity to analyze a major ore-forming process in progress, although the exact relationship between hydrothermal-vent deposits and ore deposits on land is not always clear. The hydrothermal deposits consist largely of iron, copper, and zinc sulfides; calcium and barium sulfates; and silica. Two major structural types of mineral deposits, chimneys and mounds, are found at high-temperature vents.

Chimneys are roughly tubular, vertical structures with clearly defined central channelways through which hydrothermal plumes exit (Fig. 2). They consist of concentric layers of minerals, usually with a monomineralic channel lining of zinc sulfide [wurtzite, (Zn,Fe)S], copper-iron sulfide [chalcopyrite ($CuFeS_2$) or cubanite ($CuFe_2S_3$)] and outer layers dominated by iron sulfides [pyrite (FeS_2) and pyrrhotite (FeS_{1-x})], zinc sulfides [wurtzite or sphalerite, both (Zn,Fe)S], and anhydrite [$CaSO_4$].

Mounds vary from low-lying features about 16 ft (5 m) high and about 50 ft (15 m) in diameter to massive edifices up to 56 ft (17 m) high and 115 ft (35 m) in diameter. The larger edifices often have horizontally oriented outgrowths called flanges, with pools of high-temperature fluid trapped underneath. The flanges consist of iron sulfides [pyrrhotite, pyrite, and marcasite (also FeS_2)], oxyhydroxides, zinc sulfides, rare copper-iron sulfides, anhydrite, barite ($BaSO_4$), and silica. The flanges grow from deposition within the underlying pools and from the overflow of hot fluid from the pools. The large edifices themselves appear to have formed from the coalescence of flanges cemented by late-depositing silica.

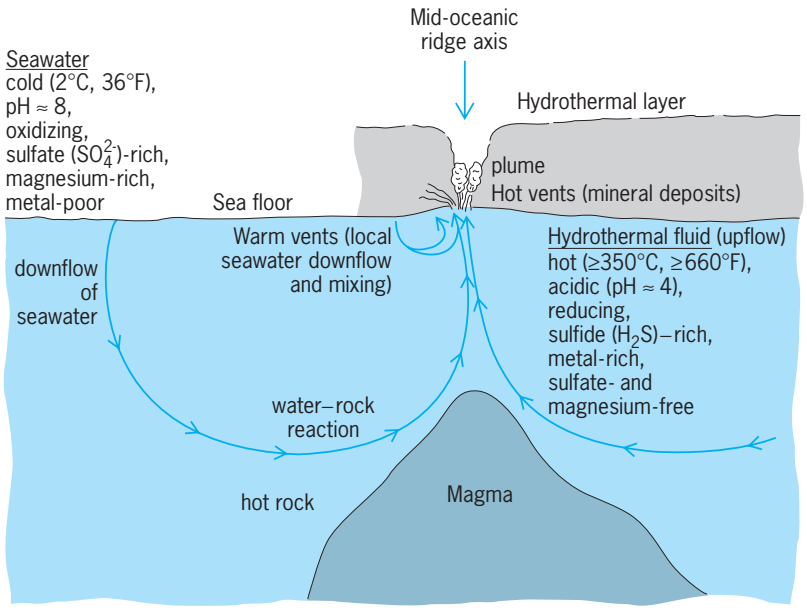

Fig. 1. Schematic view of the seawater circulation system creating hydrothermal vents along a mid-oceanic ridge.

Warm springs rarely show much mineralization at the sea floor, but minerals may be deposited below the surface where the initially hot fluids mix with seawater. *See* ORE AND MINERAL DEPOSITS.

Vent biology. Perhaps the most striking feature of sea-floor hydrothermal vents is their dense biologic communities, easily detectable in sea-floor photographs as indicators of the presence of even merely warm vent fluid. Biologists have compared the discovery of the vents to landing on a new planet.

About 450 animal species have now been identified at hydrothermal vents. About 95% of these have been species new to science, and about 75%

Fig. 2. Black smoker plume at the hydrothermal vent site at 21°N on the East Pacific Rise, off the tip of Baja California. Behind the plume is another hydrothermal chimney; visible in the foreground is the mechanical arm of the submersible *ALVIN* holding a temperature probe into the plume. Chimneys at this site reach a maximum of about 16 ft (5 m); the section in the background is about 6 ft (2 m) high.

Fig. 3. Giant *Riftia* tubeworms from warm vents at the Galápagos Rift vent site. These tubeworms reach lengths of at least 6 ft (2 m).

are known from only a single vent site. Vent faunas tend to be dominated by mollusks, annelids, and crustaceans, whereas faunas on nonvent hard-bottom habitats consist predominantly of cnidarians, sponges, and echinoderms. Large mussels or clams are found in nearly all vent systems, as are a diverse assortment of tiny limpets and other gastropods. Gutless vestimentiferan annelids (tube worms), including the enormous *Riftia pachyptila* (**Fig. 3**) with its brilliant red plume and white tube up to 2 m (6 ft) in length, are characteristic of Eastern Pacific vents, but are completely absent from vents on the Mid-Atlantic Ridge. The latter sites are dominated by several species of caridean shrimp which swarm in massive numbers near areas of active venting.

Biologically, vents are among the most productive ecosystems on Earth—an interesting paradox, since they occur in the deep ocean where food is normally very scarce. Sulfide from hydrothermal fluids provides the energy to drive these productive systems. Whereas most animal life depends on food of photosynthetic origin (inorganic carbon converted to useful sugars by plants using energy from the Sun), the animals at hydrothermal vents obtain most or all of their food by a process of chemosynthesis. Chemosynthesis is accomplished by specialized bacteria residing in hydrothermal fluids, in mats on the sea floor, or in symbiotic relationships with other organisms. The bacteria convert inorganic carbon to sugars by mediating the oxidation of hydrogen sulfide, thereby exploiting the energy stored in chemical bonds. A few vent animals are also known to use methane gas as a source of energy and carbon.

The distribution of organisms at individual vents and at different vent sites seems to be closely related to the details of the distribution of temperature and fluid composition around the vents. *See* DEEP-SEA FAUNA.

The physical and chemical conditions at hydrothermal vents would be lethal to most marine animals, but vent species have adapted to the conditions there. For example, small pompeii worms (*Alvinella pompeiana*) occupy tubes in the sides of chimneys where water temperatures are very high; these worms may experience extreme thermal gradients between their anterior and posterior ends. Vestimentiferan tube worms absorb toxic sulfide into their bloodstreams for use by symbiotic bacteria, which reside in a large internal organ known as the trophosome. The hemoglobin in their blood binds the sulfide, thereby protecting the tissues of the worm, and permitting maximum delivery of sulfide to the symbionts. One of the most remarkable adaptations is found in *Rimicaris exoculata*, a very abundant shrimp at deep vent sites on the Mid-Atlantic Ridge. This shrimp has an eyelike structure on the back of its head capable of detecting the infrared radiation that is emitted from the hot water vents. This is apparently a mechanism for locating black smokers, where there is sufficient sulfide for the chemosynthetic bacteria upon which the shrimp depends for food.

Because vent fields tend to be dynamic and ephemeral, and the life of an individual vent is measured in years to decades, scientists have long wondered how new vents are colonized and how sulfide-dependent organisms survive local extinction when old vents shut down. Most vent species produce planktonic larvae capable of dispersing for long distances without the benefit of food. Eggs of many vent species are buoyant, causing them to move rapidly upward, where they can disperse with the ocean currents. The larvae of many vent species have been captured in the buoyant hydrothermal plumes far above vent fields in the Eastern Pacific, and the larvae of the mid-Atlantic vent shrimps have been taken in plankton samples hundreds of meters above the bottom. In those systems that have been studied, the prevailing currents run along the axes of the mid-ocean ridges, a fortuitous pattern that probably increases the likelihood that larvae will locate another vent. It has also been suggested that vent species may use whale carcasses and other sulfide-rich deep-sea habitats as "stepping stones" between distant vent sites. Some scientists speculate whether life itself could have originated in environments similar to those at hydrothermal vents. Many ubiquitous biochemical mechanisms involving the use of essential metals, heat-shock proteins, and certain enzymes appear to have originated in hydrothermal systems. It has even been suggested that the biochemical pathways of photosynthesis may have begun as part of an infrared heat-detection system such as that used by vent shrimp. *See* PREBIOTIC ORGANIC SYNTHESIS.

In a remarkable discovery, it was shown that chemosynthetic microbes known as Archaea are flushed from cavities deep within the Earth's crust by hydrothermal and volcanic activity. These microbes are hyperthermophilic (hot-water-loving) and thrive at temperatures exceeding 90°C (194°F). It is now suspected that an entire community of such microbes inhabits the rocks deep within the water-saturated portions of the Earth's crust. This discovery has led to the speculation that similar microbial communities could exist on Europa, a moon of Jupiter that may have a liquid ocean with deep-sea volcanism similar to that found on Earth.

The mineral deposits at hydrothermal vents have great commercial potential, and plans to mine vent

deposits are being formulated in some parts of the world. There is increasing concern, however, that these systems are not well enough known for exploitation, and many scientists are therefore opposed to the commercial mining of vents. The genetic heritage of vent organisms may be an equally valuable resource; indeed, unique biochemical features of vent organisms are already having an impact in biotechnology. A heat-tolerant enzyme discovered in hydrothermal vent microbes is now an essential component of the polymerase chain reaction (PCR), a recent technological breakthrough that permits deoxyribonucleic acid (DNA) sequencing and fingerprinting from minuscule biological samples. This vent-dependent technology has become critically important in fields as disparate as medicine, criminology, and archeology. Marjorie Goldfarb; Craig M. Young

Bibliography. T. J. Barrett and J. L. Jambor (eds.), Seafloor hydrothermal mineralization, *Can. Mineralog.*, 26:429–888, September 1988; M. L. Jones (ed.), Hydrothermal Vents of the Eastern Pacific: An Overview, *Bull. Biol. Soc. Wash.*, no. 6, 1985; V. Robigou et al., Large massive sulfide deposits in a newly discovered active hydrothermal system, the High-Rise Field, Endeavour Segment, Juan de Fuca Ridge, *Geophys. Res. Lett.*, 20(17):1887–1890, September 3, 1993; A. P. Rona et al. (eds.), *Hydrothermal Processes at Seafloor Spreading Centers*, NATO Conference Series IV: Marine Sciences, vol. 12, 1983; The Southern Juan de Fuca Ridge: Hydrothermal fluids, sulfides, and geophysical studies, *J. Geophys. Res.*, 92:11,281–11,433, 1987.

Hydroxide

A compound containing the hydroxide ion (OH^-) and having the general formula $M(OH)_n$, where M represents a metal. Hydroxides are a subset of compounds containing the hydroxyl group (—OH) and range in chemical character from strongly basic, to amphoteric (having both acidic and basic characteristics), to essentially acidic. The hydroxide ion has a closed-shell electronic structure with a singlet ground state ($^1\Sigma^+$).

In the Lewis acid-base scheme, where a base is defined as an electron pair donor and an acid as an electron pair acceptor, a typical hydroxide decreases in base strength as attraction for electrons of the cation increases. For example, the hydroxides of electropositive elements such as the alkali metals and alkaline earths tend to be bases. When these ionic compounds are dissolved in water, they form metal ions and hydroxide ions, as inreaction (1), where

$$M^+(OH^-)_n(S) + nH_2O \longrightarrow M^{n+}(aq) + nOH^-(aq) \qquad (1)$$

s represents a solid and aq represents a water solution. The hydroxides of nonmetals (X) such as boron hydroxide [$B(OH)_3$], where the X-O bonds are covalent, are generally acidic, as shown in reaction (2),

$$X(OH)_n + nH_2O \longrightarrow XO_n^{n-}(aq) + nH_3O^+(aq) \qquad (2)$$

where H_3O^+ is the hydronium ion. The amphoteric hydroxides may dissociate by either mechanism, depending on the presence of strong acids or bases. *See* ACID AND BASE; HYDROGEN ION.

The alkali metal hydroxides such as sodium hydroxide (NaOH) are extremely important as reagents in metallurgy and photography and in the manufacture of soaps and detergents. Calcium hydroxide [$Ca(OH)_2$], known as slaked lime, is used in the preparation of mortar for brick laying. Minerals such as brucite [$Mg(OH)_2$] and pyrochroite [$Mn(OH)_2$] are naturally occurring hydroxides.

The lanthanides (Ln) also form hydroxides [$Ln(OH)_3$]; they are not simply hydrated oxides. They are prepared by precipitation of Ln(III) salts in aqueous NaOH and have hexagonal structures with tricapped prismatic coordination (a classification of ligand geometry around the lanthanide). Lanthanides exhibit the lanthanide contraction (that is, a decrease in atomic and ionic size with a corresponding increase in atomic number). Therefore, the hydroxides of lanthanides exhibit a reduction in basic character from lanthanum hydroxide [$La(OH)_3$] to lutetium hydroxide [$Lu(OH)_3$]. *See* HYDRATE; LANTHANIDE CONTRACTION.

The hydroxide ion may behave as a ligand in transition-metal complexes; examples are the hydroxides of cobalt {$[Co(OH)_4]^{2-}$}, nickel [$Ni(OH)_2$], and zinc [$Zn(OH)_2$]. Ferrous hydroxide [$Fe(OH)_2$] is oxidized to red-brown ferrous oxide (Fe_2O_3), commonly known as rust. In addition, hydroxo bridges are well documented; single (**1**), double (**2**), or even triply (**3**) bridged structures are known. The

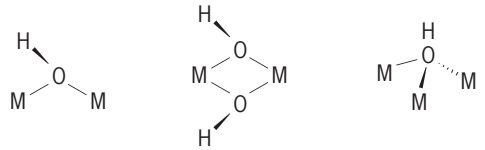

(**1**) (**2**) (**3**)

double-bridged species (**2**) are the most common. *See* LIGAND. Thomas J. Meade

Hydroxyl

A chemical group in which oxygen and hydrogen are bonded and act as a single entity. In inorganic chemistry the hydroxyl group is known as the hydroxide ion (OH^-), and it is frequently bonded to metal cations, for example, sodium hydroxide (NaOH). In organic chemistry it frequently acts as a functional group, for example, in an alcohol (ROH, where R represents an alkyl group). *See* ACID AND BASE.

Oxygen toxicity. Many of the intermediate redox forms of dioxygen are toxic and damage important biomolecules. Much of this toxicity is thought to involve the generation and reactivity of hydroxyl (·OH), which is sometimes called the hydroxy radical. Radicals are chemical species with unpaired electrons (a dot represents an unpaired electron). Radicals can be formed via the addition of a single

electron to a molecule, so superoxide, which is generated by the addition of a single electron to dioxygen, is also a radical and is sometimes written as $O_2^{\cdot -}$. Radicals can also be generated by removal of a single electron or by breaking in half the bonds that are formed by sharing two electrons. *See* FREE RADICAL; OXYGEN TOXICITY.

Production. The most common means for producing hydroxyl is the reaction of a reducing agent with hydrogen peroxide (H_2O_2). Other biologically available radicals, such as nitric oxide or semiquinones, may act as the reducing agent, but it is believed that transition-metal ions, such as ferrous ion (Fe^{2+}), are the most common reducing agents for generating ·OH, according to the Fenton reaction (1). Most of

$$Fe^{2+} + H_2O_2 \longrightarrow Fe^{3+} + OH^- + \cdot OH \qquad (1)$$

the free iron available in biological systems exists as Fe^{3+}, but the superoxide anion is capable of reducing Fe^{3+} to Fe^{2+} [reaction (2)]. The net reaction is

$$O_2^{\cdot -} + Fe^{3+} \longrightarrow O_2 + Fe^{2+} \qquad (2)$$

therefore the reaction of O_2^- with H_2O_2, which is called the Haber-Weiss reaction (3), and is catalyzed

$$O_2^{\cdot -} + H_2O_2 \longrightarrow O_2 + \cdot OH + OH^- \qquad (3)$$

by redox-active metal ions such as Fe^{3+}. Since many transition metals exist in oxidation states that differ by a single electron, many reactions that form radicals use transition metals as an electron source or sink.

Homolytic bond cleavages are reactions in which the two electrons from a bond are divided equally between the two atoms in the bond, so homolytic cleavage always produces two radicals. Electrons are more stable when shared as a pair by two atoms; thus, homolytic cleavage requires the input of energy, often in the form of radiation. Another common means for generation of hydroxyl comes from the action of ionizing radiation on water, which leads to the homolytic cleavage of one of the O-H bonds in water to produce hydroxyl and a free hydrogen atom [reaction (4)]. The action of ultraviolet light on hy-

$$HO{-}H \xrightarrow[\text{radiation}]{\text{ionizing}} H\cdot + \cdot OH \qquad (4)$$

drogen peroxide also produces ·OH via homolytic bond cleavage [reaction (5)]. The O-H bond in water

$$HO{-}OH \xrightarrow[\text{light}]{\text{ultraviolet}} HO\cdot + \cdot OH \qquad (5)$$

is stronger than the O-O bond in hydrogen peroxide, so radiation of higher energy is required for homolytic cleavage of water.

Chemical behavior. Once generated, hydroxyl is a potent one-electron oxidant that forms the very stable OH^- ion [reaction (6)], and it abstracts hy-

$$\cdot OH + e^- \longrightarrow OH^- \qquad (6)$$

drogen atoms from organic molecules that contain C-H bonds to form the stronger O-H bond in water [reaction (7), where R represents an organic

$$\cdot OH + R{-}H \longrightarrow R\cdot + H_2O \qquad (7)$$

molecule that contains C-H bonds and R· is an organic radical with an unpaired electron on a carbon atom]. This reaction demonstrates an important property of radicals: the reaction of a radical, which contains an uneven number of electrons, with a molecule, which contains an even number of electrons paired in bonds, must generate a radical, because the number of electrons cannot change during the reaction. Thus, most reactions of radicals generate new radicals in processes called radical chain reactions. Reactions of radicals with molecules will continue to produce new radicals until other odd-electron species (such as transition-metal ions or other radicals) react with the radicals to produce even-electron molecules via termination reactions.

In biological systems, reaction of hydroxyl with lipids, which are organic molecules that make up cell membranes, leads to chain reactions involving dioxygen in a process called lipid peroxidation. This process is very destructive to cells, and it is a primary pathway of hydroxyl (and hence dioxygen) toxicity. Other deleterious effects of hydroxyl involve damage to nucleic acids [deoxyribonucleic acid (DNA) and ribonucleic acid (RNA)], which also contain many C-H bonds that react with hydroxyl to produce new radicals. *See* CATALYSIS; CHAIN REACTION (CHEMISTRY); DEOXYRIBONUCLEIC ACID (DNA); ENZYME; PROTEIN; RIBONUCLEIC ACID (RNA); TRANSITION ELEMENTS.

H. Holden Thorp

Bibliography. I. Bertini et al. (eds.), *Bioinorganic Chemistry*, 1994; S. J. Lippard and J. M. Berg, *Principles of Bioinorganic Chemistry*, 1994; M. Roberfroid and P. Buc Calderon, *Free Radicals and Oxidation Phenomena in Biological Systems*, 1994.

Hydrozoa

A class of the phylum Cnidaria which includes the fresh-water hydras, the marine hydroids, many of the smaller jellyfish, a few special corals, and the Portuguese man-of-war. The Hydrozoa may be divided into seven orders: the Hydroida, Milleporina, Stylasterina, Trachylina, Chondrophora, Siphonophora, and Spongiomorphida. See separate article on each order.

Hydrozoa differ from the Scyphozoa, which are mostly the large jellyfish, and from the Anthozoa, to which the sea anemones and most corals belong, in the following features: The digestive space is not divided by longitudinal partitions; it lacks nematocyst-bearing structures, and has no stomodeum. The medusa has a velum but lacks the highly specialized sense organs characteristic of the Scyphozoa. *See* ANTHOZOA; SCYPHOZOA.

The form of the body varies greatly among the hydrozoans. This diversity is due in part to the existence of two body types, the polyp and the medusa. A specimen may be a polyp, a medusa, a colony

of polyps, or even a composite of the first two. Polyps are somewhat cylindrical, attached at one end, and have a mouth surrounded by tentacles at the free end. Medusae are free-swimming jellyfish with tentacles around the margin of the discoidal body.

In a representative life cycle, the fertilized egg develops into a swimming larva which soon attaches itself and transforms into a polyp. The polyp develops stolons (which fasten to substrates), stems, and other polyps to make up a colony of interconnected polyps. Medusae are produced by budding and liberated to feed, grow, and produce eggs and sperm.

Not all hydroids conform to the foregoing description. Variation among them is due largely to differences in the pattern of growth and the extent to which the medusa stage is developed. In *Hydra* and a few hydroids, the medusa stage is absent, and new polyps, produced by budding, separate from the parent. In these cases each polyp is solitary. Corallike forms are similar to hydroids with the addition of a calcareous skeleton. In some cases the polyp form is greatly reduced or absent and the medusa stage predominates as in the Trachylina. The greatest complexity occurs in the Siphonophora, whose bodies have several different types of both medusoid and polypoid components.

Most hydrozoans are carnivorous and capture animals which come in contact with their tentacles. The prey is immobilized by poison injected by stinging capsules, the nematocysts. Most animals of appropriate size can be captured, but small crustaceans are probably the most common food. *See* CNIDARIA; FEEDING MECHANISMS (INVERTEBRATE).

Sears Crowell

Bibliography. L. H. Hyman, *The Invertebrates*, vol. 1, 1940; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Hyena

An African carnivore represented by three species of the family Hyaenidae; this family also includes the aardwolf (*Proteles cristatus*). Hyena adults are over 4 ft (1.2 m) high and almost 5 ft (1.5 m) in length, and may weigh up to 165 lb (75 kg). While they resemble dogs superficially, they are more closely related to the felids (cats). All three species are adapted for feeding on carrion, having well-developed foreparts, reduced hindquarters, a rounded head, and short strong jaws. The 34 teeth (dental formula I 3/3 C 1/1 Pm 4/3 M 1/1) are extremely strong and together with the powerful jaw muscles can crush bone.

Hyenas are four-toed digitigrade (walking on the toes) animals with blunt, nonretractile claws that are used for digging and disinterring bodies. The gestation period is about 13 weeks, with two to four young comprising the yearly litter. The maximum life-span of this animal is about 25 years.

The spotted hyena (*Crocuta crocuta*; see **illus.**) ranges south of the Sahara, while the brown hyena

Spotted hyena (*Crocuta crocuta*). (*Photo by H. Vannoy Davis; © 2001 California Academy of Sciences*)

(*Hyaena brunnea*) is restricted to southern Africa. The striped hyena (*H. hyaena*) is found in northeastern Africa and down into southern Asia. The spotted hyena is more active diurnally than the striped hyena, which is nocturnal and solitary in its activities. *Crocuta crocuta* is interesting in that it is difficult to distinguish between the sexes, since the female closely resembles the male in the external appearance of the reproductive organs. *See* CARNIVORA.

Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999.

Hyeniales

An order of Devonian plants considered by some to be related to the spenophytes. The small dichotomously forked leaves tend to be borne in whorls. Some leaves bear terminal sporangia, but these appendages are neither aggregated into a tight cone nor separated by bracts.

The Middle Devonian genus *Hyenia* (see **illus.**) has aerial axes borne on stout rhizomes. Its leaves are two to four times dichotomized (branched repeatedly into two branches) and their arrangement approaches the whorled condition. Fertile leaves are characterized by recurved segments which bear paired, terminal sporangia, and often by erect, sterile segments as well. The anatomy of *Hyenia* is essentially unknown. *Protohyenia,* believed to be a precursor of *Hyenia*, has since proved to be a *Pseudosporochnus*.

Calamophyton, also from the Middle Devonian, consists of a stout axis that divides digitate-fashion into three or more lesser branches which may dichotomize. The axes frequently appear to be jointed, and whorled leaves are borne at the nodes. The leaves dichotomize two to four times in three dimensions. The best-known species, *C. bicephalum*, bears whorls of fertile leaves, each of whose two main subdivisions bear three recurved appendages. Each appendage bears a pair of terminal sporangia. Sporangia are fusiform and dehisce longitudinally along one edge. Erect, sterile portions of the fertile

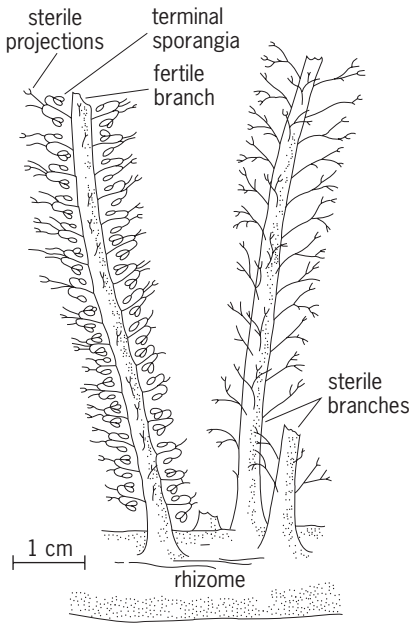

***Hyenia*, showing stout rhizome with fertile and sterile branches.**

appendages project beyond the position of the sporangia. Spores assignable to the genus *Dibolisporites gibberosus* have been found in some sporangia. They range from 86 to 166 micrometers in diameter and are ornamented by coni and spinae up to 4.5 micrometers long. The vascular strand of *Calamophyton* has been regarded as siphonostelic (xylem surrounding a pith). The anatomy of *C. bicephalum* has been shown to consist of a number of separate strands of xylem, thus resembling *Pseudosporochnus*. *See* PALEOBOTANY; PLANT KINGDOM. Harlan P. Banks

Hygrometer

In the most general sense, a hygrometer is any device used to measure humidity in air or other gases. The more modern definition is that hygrometers are only devices that measure humidity via changes in material properties. Humidity may also be measured by other methods, chiefly psychrometry and dewpoint, both of which rely on temperature changes: due to evaporation for psychrometers, and due to condensation for dewpoint. In contrast, hygrometers are generally isothermal. Primitive hygrometers relied on mechanical changes, such as the slackening of hair with increasing humidity. However, the vast majority of modern hygrometers are based on changes in the electrical resistance or capacitance of a material with changes in humidity. Most hygrometers are designed and calibrated to indicate relative humidity, which is the percentage of water in air (or another gas) relative to the saturation condition. When air is saturated with water, further water content appears as mist or frost on exposed surfaces. Many commercial hygrometers are combined with thermometers, sometimes in the same electrical circuit, and these are known as thermohygrometers. *See* DEWPOINT; HUMIDISTAT; HUMIDITY; MOISTURE-CONTENT MEASUREMENT; PSYCHROMETER; PSYCHROMETRICS.

There are three common classes of electronic hygrometer: resistive, capacitive, and piezoelectric. The first two types may also be combined to operate as an inductive circuit. Resistive and capacitive hygrometers are commonly based on changes in resistance or capacitance of a thin polymer film. Sulfonated polystyrene, created by treating polystyrene with sulfuric acid, is an example of a suitable resistive polymer. It is important that the polymer film be well exposed to air. The rate of response of a hygrometer is dependent on the ability of air to reach the sensor. Some hygrometers even incorporate fans to accelerate response time. *See* POLYSTYRENE RESIN.

Figures 1 and **2** illustrate common resistive and capacitive circuits incorporating a thin film. In both circuits, the film is open to air, and humidity can be measured due to changes in the resistance or capacitance of the film. In the capacitive circuit (Fig. 2), the film is shielded by the porous electrode, while in the resistive circuit (Fig. 1) a porous sheath would be required to protect the sensor from damage. Polymer films are preferred over many other materials because they can be chosen to respond to relative humidity. Other materials, including ceramics such as aluminum oxide, tend to respond to absolute humidity (the amount of absolute water content in air, rather than the amount relative to saturation). *See* CAPACITANCE MEASUREMENT; RESISTANCE MEASUREMENT.

Piezoelectric crystals such as quartz can also be used to measure absolute humidity. The mass of the oscillating crystal changes directly with the amount of water absorbed, which is typically proportional to absolute humidity. This effect, in turn, changes the frequency of oscillation, which can be measured extremely precisely. *See* PIEZOELECTRICITY.

Commercial thin-film hygrometers tend to be precise to within an error of ±2% for relative humidity. This accuracy can be refined by calibration against a reference humidity close to the temperature of interest. The simplest calibration references are saturated

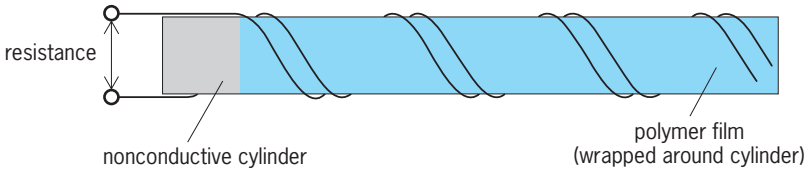

Fig. 1. Simplified resistive hygrometer circuit in which the electrical resistance of the polymer film depends on the ambient humidity. The electrical resistance of the polymer is measured across the electrical leads as shown. The ends of the two electrical leads are not joined but are connected to the film.

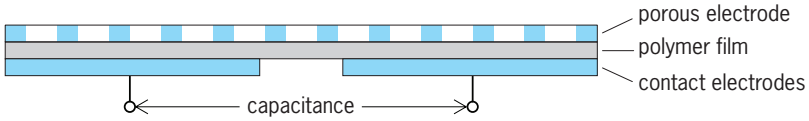

Fig. 2. Simplified capacitive hygrometer circuit in which the electrical capacitance of the contact electrodes across the polymer film depends on the ambient humidity.

Relative humidities for saturated salt soluitions

Salt	Relative humidity*, %
Lithium chloride	11.31
Potassium acetate	23.11
Potassium carbonate	43.16
Sodium chloride	75.47
Potassium sulfate	97.59

*Saturated solution, 20°C (68°F).

salt solutions. These are created by dissolving an excess of a given salt to form a slushy solution. The hygrometer is then sealed in a container with the solution. Relative humidities for saturated salt solutions are shown in the **table** for 20°C (68°F). The salts shown give humidities that are almost constant with temperature. For example, sodium chloride (table salt) produces relative humidities of 75.51% at 0°C (32°F), 75.47% at 20°C (68°F), 74.41% at 55°C (176°F), and 76.29% at 80°C (176°F). For careful calibration, the solution should be allowed to return to the reference temperature after dissolving the salt, as the heat of dissolution can change the temperature significantly. Brian J. Lowry

Bibliography. J. J. Carr, *Sensors and Circuits*, Prentice Hall, 1993; W. C. Dunn, *Fundamentals of Industrial Instrumentation and Process Control*, McGraw-Hill, 2005; H. N. Norton, *Hand-book of Transducers*, Prentice Hall, 1989.

Hymenomycetes

An artificial class of fungi in the phylum Basidiomycota. It was traditionally divided into two subclasses: Holobasidiomycetidae, delimited by nonseptate basidia and the absence of a yeast phase; and Phragmobasidiomycetidae, frequently with septate basidia and often forming a yeast phase. A typical hymenomycete produces a fruit body or basidiome with spore-bearing basidia organized in a membranelike layer called the hymenium. The fruit bodies are designed to allow for basidiospore discharge into air currents either directly off basidia or after falling from elaborate fertile surfaces. Falling spores, indicative of active discharge, accumulate in powdery masses called spore prints. The shape of the hymenium varies from lamellate (gilled as in mushrooms), poroid (as in conk or bracket fungi), toothed (in hedge hog fungi), coralloid (coral fungi), labyrinthoid (daedaleoid fungi), wrinkled (merulioid fungi), or smooth to diffuse (corticioid fungi). Exceptional hymenomycetes may be aquatic, lack a mycelial phase, or lack a fruit body.

Antibiotics have been isolated from many species. Commercially grown edible species include the button mushroom (*Agaricus bisporus*), Shiitake (*Lentinula edodes*), Paddy Straw mushroom (*Volvariella volvacea*), and Wood Ear (*Auricularia polytricha*). Wild harvested species include the Matsutake (*Tricholoma matsutake* and *T. magnivelare*), chanterelles (*Cantharellus cibarius* and allies), and the King Bolete (*Boletus edulis*).

Most genera are either saprophytic (for example, *Agaricus* and *Polyporus*), or mycorrhizal with trees (*Albatrellus, Cortinarius, Ramaria*, and *Thelephora*). Others are parasites. *Heterobasidion* causes destructive tree diseases; *Rhizoctonia* and *Typhula* (snow molds) cause field crop losses; *Mycena citricolor* blights coffee leaves; and *Exobasidium*, an obligate plant pathogen, induces the formation of galls and leaf curls. Other notable pathogens include *Hohenbuehelia* and *Pleurotus*, which capture nematodes; *Serpula*, a major dry-rot agent; *Dictyonema*, a basidiolichen; and *Septobasidium*, which harnesses living scale insects. *See* BASIDIOMYCOTA; EUMYCOTA; FUNGI; PLANT PATHOLOGY. Scott A. Redhead

Bibliography. D. S. Hibbett et al., *Proc. Natl. Acad. Sci.*, 94:12002-12006, 1997; K. K. Nakasone, *Cultural Studies and Identification of Wood-Inhabiting Corticiaceae and Selected Hymenomycetes from North America*, 1990; P. Stamets, *Growing Gourmet and Medicinal Mushrooms*, 1993; E. C. Swann and J. W. Taylor, *Mycologia*, 85:923–936, 1993.

Hymenoptera

The third largest order of insects, containing the sawflies, ants, wasps, bees, and related forms. There are some 15,700 described species, subspecies, and varieties from America north of Mexico. Conservative estimates suggest that the world fauna may comprise well over 100,000 described species of this order, with many thousands still to be described.

This order is of great importance to humans. Some members such as the sawflies, certain chalcidoids, and most cynipoids, feed during the larval stage on foliage or other plant tissues. Many species, such as the ichneumon flies, most chalcid flies, and wasps, are parasites or predators of other insects or spiders during their larval stage. Bees are indispensable in the pollination of many fruits, vegetables, and forage crops. *See* POLLINATION.

Hymenoptera occur in all major faunal zones but are more abundant and have greater diversity of species in the tropical and temperate zones. Representation in the northern parts of the boreal zone is limited to a few sawflies, some Parasitica, and very few Aculeata, of which the bumblebees are the most conspicuous representatives. *See* INSECTA.

Adult Hymenoptera usually may be recognized by having two pairs of membranous wings with reduced venation, the hind pair smaller than the front pair, and by mouthparts formed for biting and often for lapping or sucking. In the higher forms, the abdomen is constricted basally, its first segment fused with the hind part of the thorax. Females always have an ovipositor modified for sawing, piercing, or stinging. Metamorphosis is complete, and sawfly larvae resemble caterpillars except for the abdominal prolegs which lack a series of hooklets. Larvae of higher forms are legless and maggotlike but have a well-developed head. Pupae have the appendages free

TABLE 1. Families of Hymenoptera

Classification	Common name	No. of species
Suborder Symphyta	Sawflies	1009
Superfamily Megalodontoidea		120
Xyelidae		33
Pamphiliidae	Web-spinning sawflies	87
Superfamily Tenthredinoidea		849
Pergidae		13
Argidae		32
Cimbicidae	Cimbicid sawflies	12
Diprionidae	Conifer sawflies	35
Tenthredinidae	Sawflies	757
Superfamily Siricoidea		28
Syntexidae		1
Siricidae	Horntails	15
Xiphydriidae		6
Orussidae		6
Superfamily		12
Cephoidea		12
Cephidae	Stem sawflies	12
Suborder Apocrita		13,346
Superfamily Ichneumonoidea		3814
Stephanidae		7
Braconidae	Braconid wasps	1239
Ichneumonidae	Ichneumon flies	2568
Superfamily Chalcidoidea		2032
Mymaridae	Fairy flies	110
Trichogrammatidae	Minute egg parasites	39
Eulophidae		544
Elasmidae		17
Thysanidae		18
Eutrichosomatidae		2
Tanaostigmatidae		4
Encyrtidae		320
Eupelmidae		89
Eucharitidae		27
Perilampidae		31
Agaontidae	Fig insects	2
Torymidae	Torymids	181
Ormyridae		17
Pteromalidae		321
Eurytomidae	Seed and stem chalcids	203
Chalcididae	Chalcids	101
Leucospidae		6
Superfamily Cynipoidea	Gall wasps	877
Ibaliidae		6
Liopteridae		2
Figitidae		58
Cynipidae	Cynipids of gall wasps	811
Superfamily Proctotrupoidea		985
Evaniidae	Ensign flies	11
Gasteruptiidae		50
Pelecinidae	Pelecinid wasps	1
Vanhorniidae		1
Roproniidae		3
Heloridae		1
Proctotrupidae		54
Ceraphronidae		101
Diapriidae		304
Scelionidae	Scelionid wasps	272
Platygasteridae		182
Trigonalidae		5
Superfamily Bethyloidea		345
Chrysididae	Cuckoo wasps	124
Bethylidae		100
Sclerogibbidae		1
Dryinidae		120
Superfamily Scolioidea		643
Tiphiidae	Tiphiid wasps	185
Sierolomorphidae		2
Mutillidae	Velvet ants	409
Rhopalosomatidae		2
Scoliidae		26
Sapygidae		19
Superfamily Formicoidea		786
Formicidae	Ants	786
Superfamily Vespoidea		368
Vespidae	Hornets, yellow jackets, potter wasps	368
Superfamily Pompiloidea		279
Pompilidae	Spider wasps	279
Superfamily Sphecoidea	Fossorial wasps	1215
Ampulicidae		3
Sphecidae		1212
Superfamily Apoidea	Bees	3310
Colletidae	Colletid bees	149
Andrenidae	Andrenid bees	852
Halictidae	Halictid and sweat bees	472
Melittidae		31
Megachilidae	Leafcutting bees	730
Apidae	Honeybees, bumblebees, and carpenter bees	1076

from the body except for certain chalcidoids.

The first four superfamilies of the Apocrita—the Ichneumonoidea, Chalcidoidea, Cynipoidea, and Proctotrupoidea—are commonly called the Parasitica, and the remaining superfamilies are known as the Aculeata. The Aculeata are stinging forms and the Parasitica are parasites of other insects. It is impossible to demarcate these two groups sharply because some Aculeata are parasites and some Parasitica are phytophagous. However, except for the phytophagous species of Parasitica, these insects lay their eggs in or on an insect or spider host while the Aculeata place theirs in nests with a provision of food. Some connecting forms apparently occur in the Procotrupoidea and Bethyloidea. Some of the Parasitica are more highly specialized, morphologically, than any other Hymenoptera, but the Aculeata show the greatest specialization in behavior. **Table 1** presents the major classification of the order as recognized in North America. The number of described

species, subspecies, and varieties in America north of Mexico is noted for each family, and common names are given.

Morphology

The adult hymenopteron has a clearly differentiated head, thorax, and abdomen. Wings, when present, and legs are attached to the thorax (**Fig. 1**).

Head. Typically the head is so oriented that mouthparts are directed downward; however, all variations occur, and in some species the mouthparts are directed forward. The large compound eyes occupy much of the sides of the head, though they are reduced in size in many ants and some Parasitica. Three ocelli are typically present on the top of the head, but they may be reduced or absent in wingless forms. The paired antennae arise from the face between the eyes, and they may be close to the mouthparts or removed from them. Primitively, the antennae are long, slender, and multisegmented, but in the more highly specialized members of all major groups there has been a tendency toward reduction in the number of segments.

The mouthparts consist of paired mandibles, and a labiomaxillary complex formed by membranous connections between the maxillae and labium (**Fig. 2**). Most sawflies have simple biting mouthparts, but in the higher forms the labiomaxillary complex is variously modified to permit lapping or sucking types of feeding. In many bees and a few wasps, the labiomaxillary complex is greatly elongated to permit these insects to suck nectar from flowers which have deeply seated nectaries. The mandibles, in forms other than sawflies, do not function primarily in feeding but are used in manipulation of parasites or prey, construction of nests, and to escape from the cocoon or host body.

Thorax. The thorax consists of three segments, tightly fused together. Each segment bears a pair of legs, and each of the last two segments bears a pair of wings. In the sawflies, the thorax is broadly joined to the abdomen, but in the Apocrita the true first abdominal segment, the propodeum, has fused firmly with the thorax and is separated from the remainder of the abdomen by a constriction.

Wings. Most species have two pairs of wings, of which the posterior pair is smaller. In flight, the fore- and hindwings are joined by a row of tiny hooks, the hamuli, along the fore margin of the posterior wing, which fit into the downfolded hind margin of the anterior wing. The most complete venation is shown by the sawflies, and extensive reduction of venation in varying degrees has occurred in all higher forms (**Fig. 3**). Flightless species with shortened, nonfunctional wings, or no wings at all occur in most major groups, except the sawflies and bees. The females usually exhibit a greater reduction than do the males.

Legs. Each leg consists of a coxa which articulates with the corresponding thoracic segment, a trochanter, femur, tibia, and segmented tarsus. The tarsus usually has five segments, but these may be reduced to four or three in many Chalcidoidea and males of *Orussus* (Orussidae). The last tarsal segment typically bears a pair of claws. In most Parasitica, there is a constriction near the base of the femur which causes the trochanter to appear two-segmented.

The legs are frequently modified to serve varied specialized uses. Many species which dig through the soil, like the Scolioidea, have thickened femora and tibiae bearing numerous stout spines; while other digging wasps have a comb of bristles on each fore tarsus which aids in raking the soil out of the burrow. Some bees, such as *Andrea*, have groups of curled, branched hairs on the hindlegs to collect

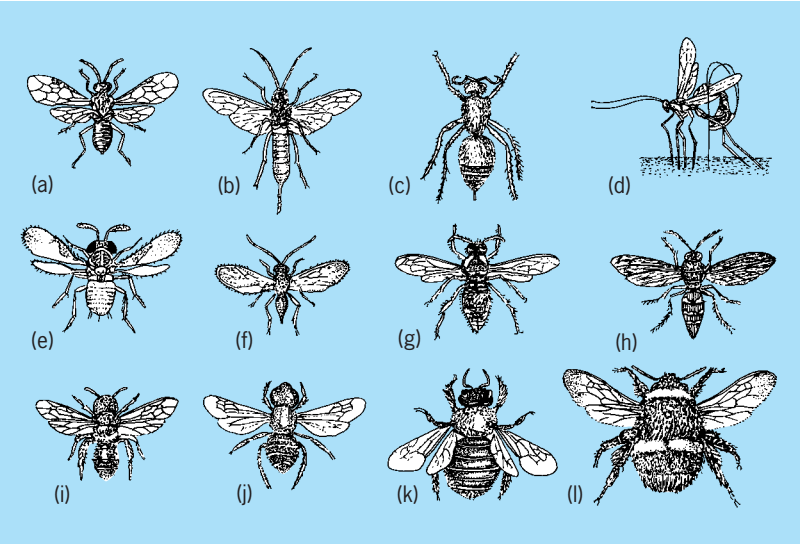

Fig. 1. Representatives of 12 families of Hymenoptera. (*a*) Tenthredinidae. (*b*) Siricidae. (*c*) Mutillidae. (*d*) Ichneumonidae. (*e*) Chalcididae. (*f*) Cynipidae. (*g*) Vespidae. (*h*) Sphecidae. (*i*) Andrenidae. (*j*) Megachilidae. (*k*) Xylocopidae. (*l*) Bombidae. (*After T. I. Storer and R. L. Usinger, General Zoology, 3d ed., McGraw-Hill, 1957*)

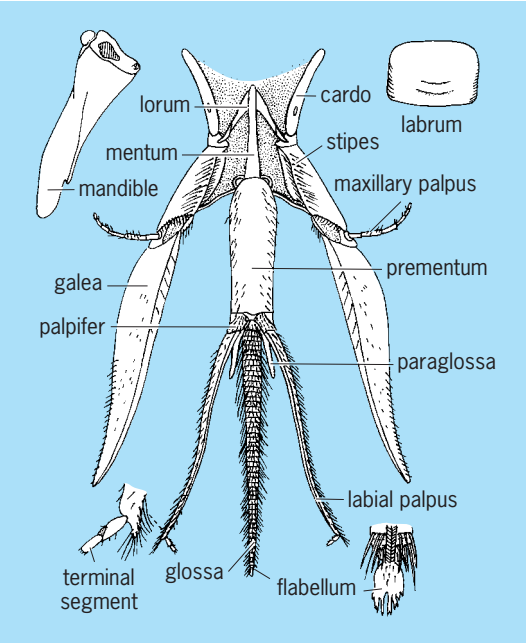

Fig. 2. Mouthparts of *Anthophora edwardsii*. (*After E. O. Essig, College Entomology, MacMillan, 1942*)

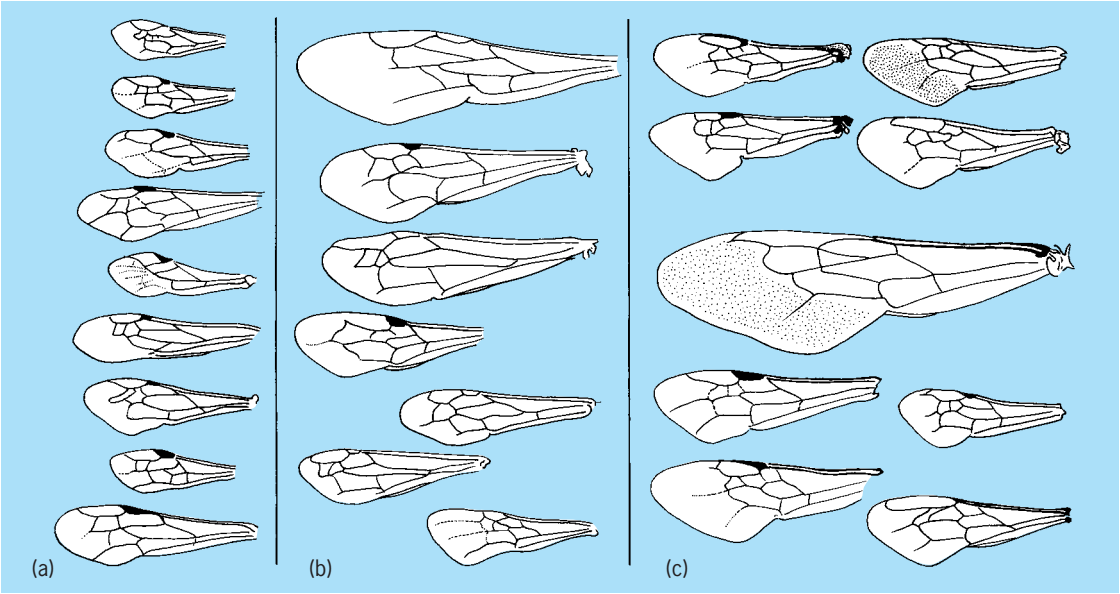

Fig. 3. Wings of Hymenoptera. Most species have two pairs, of which the posterior pair is smaller. (*a*) Sphecoid wasps. (*b*) Vespoid wasps. (*c*) Bees. (*After E. O. Essig, College Entomology, Macmillan, 1942*)

pollen, while others, like the honeybee and bumblebees, have the hind tibia expanded, flattened, and fringed by long curved hairs to form a pollen basket, or corbicula. Females of most Dryinidae have the fore tarsus modified into a chela to grasp the prey.

Abdomen. The abdomen primitively consists of 10 segments, though the number appears to be less because of modification or loss in the higher forms. In the sawflies, 10 terga and 9 sterna may be recognized, the apical parts being modified in connection with the male genitalia or female ovipositor. In the Apocrita, the first abdominal segment has fused with the thorax, and 6 other segments are normally visible in the female, 7 in the male. An exception to this is the Chalcidoidea, where there are 7 terga. Frequently there are fewer visible segments, either because of fusion, as in some Braconidae and Proctotrupidae, or because of retraction of the posterior 2 or 3 segments, as in the Chrysididae.

The female ovipositor, or sting, is formed from processes of the eighth and ninth sterna (**Fig. 4**). In the sawflies, parts of the ovipositor have ridges which terminate ventrally in the teeth of the saw. Vestiges of such ridges in the Apocrita constitute the barbs, which, if well-developed as in the honeybee, cause the sting to remain in the wound. The eggs of both Symphyta and Parasitica pass through the ovipositor during oviposition. In the Aculeata, the ovipositor is purely a stinging mechanism, and the egg issues from an opening at its base. The sting is reduced or lacking in some ants and in the stingless honeybees of the tropics.

Venom. In the Apocrita there is a pair of acid glands opening into a poison sac connected with the ovipositor. The secretion of these glands produces either a temporary paralysis when injected into their hosts by some Parasitica, or, usually, permanent paralysis when injected into their prey by aculeate wasps. Bees use their stings purely for defense. The venom of the Aculeata is a complex substance consisting of a protein and certain enzymes, as well as other constituents. Apparently the composition varies slightly with each species, which complicates the preparation of a desensitizing agent. When a human is stung, the enzymes react with his tissues to release histamine. Death may occasionally result from anaphylactic shock, or from mechanical suffocation due to swelling of the lymphatic system. Medical assistance should be sought if severe swelling occurs following a sting, especially one on the face or throat.

Biology

Practically all hymenopterous adults are terrestrial forms, living in, on, or near the Earth's surface. A few species are secondarily aquatic, the adults swimming or walking under water to search out and parasitize aquatic or subaquatic hosts.

Most adults feed on plant nectar or honeydew secretions of various insects. A few sawflies prey on other insects. Some species of Parasitica and Aculeata imbibe body juices of the host or prey which they attack primarily for oviposition. Not too much is known of the nutritional requirements of newly emerged adults, but it is likely that many of them require some nutrient materials in order for the eggs to mature.

Reproduction. Mating takes place in a variety of situations, but it is always of rather short duration. Customarily, the males emerge from one to several days earlier than the females. Among the aculeates, the males of ground-nesting species may indulge in prenuptial flights over the nesting site while awaiting the emergence of females. Usually they pounce on the females and mating takes place on the ground. Males of some wood-nesting species may hover in front of the burrows harboring the females, and in

this instance mating usually takes place during flight. Other species meet and mate on flowers. Among the ants and social wasps, large numbers of reproductive females and males emerge simultaneously from several colonies and form mating swarms in the air. The honeybee has a similar nuptial flight, but it is composed of only one virgin queen and a number of drones.

Most species are represented by both males and females. Males are usually produced from unfertilized eggs and have half the normal number of chromosomes, while females are produced from fertilized eggs and have the normal number of chromosomes. However, both facultative and obligate parthenogenesis are more common than in any other order of insects. In those species in which facultative parthenogenesis occurs, apparently only males are produced from virgin females. In those species exhibiting obligate parthenogenesis, females are produced from unfertilized eggs. Certain Cynipoidea have alternating generations, in which one kind of female produces both sexes parthenogenetically, and the fertilized females of that generation produce only females.

In addition to males and fertile females found in most Hymenoptera, the social species have a third form called the worker caste. These are females which are sterile except under unusual circumstances. The workers of social wasps and bees are normally very similar in appearance to the queens, differing principally in their smaller size and occasionally in color pattern. The honeybee queen is quite different from the workers because she lacks wax glands and pollen-collecting apparatus, and she has a smooth rather than a barbed sting. Worker ants ordinarily differ more from their queens than do other species of social Hymenoptera—wings are absent, and the thorax is more modified. In some ants, the workers may be dimorphic, some of them being modified into a soldier form with huge head and mandibles for defense of the colony.

Life history. Hymenoptera exhibit complete metamorphosis during development and pass through an egg, larval, and pupal state (**Fig. 5**). So far as is known, Hymenoptera always lay eggs. These are deposited in a protected situation on or near the supply of larval food. The hymenopterous egg is usually ovoid or sausage-shaped, and many species in some groups, like the Cynipoidea and some of the other Parasitica, have stalked eggs. The surface is usually unsculptured and very delicate. The eggs hatch after an incubation of variable length, and the resulting larvae begin a growth period consisting of three or four instars.

The larvae of sawflies are very similar in gross appearance to lepidopterous caterpillars. There is a well-developed head bearing powerful biting mouthparts. There are three pairs of segmented thoracic legs and, in all but the forms that bore in stems or wood, there are from six to eight pairs of unsegmented abdominal prolegs which do not bear tiny hooks, or crochets, as in the Lepidoptera.

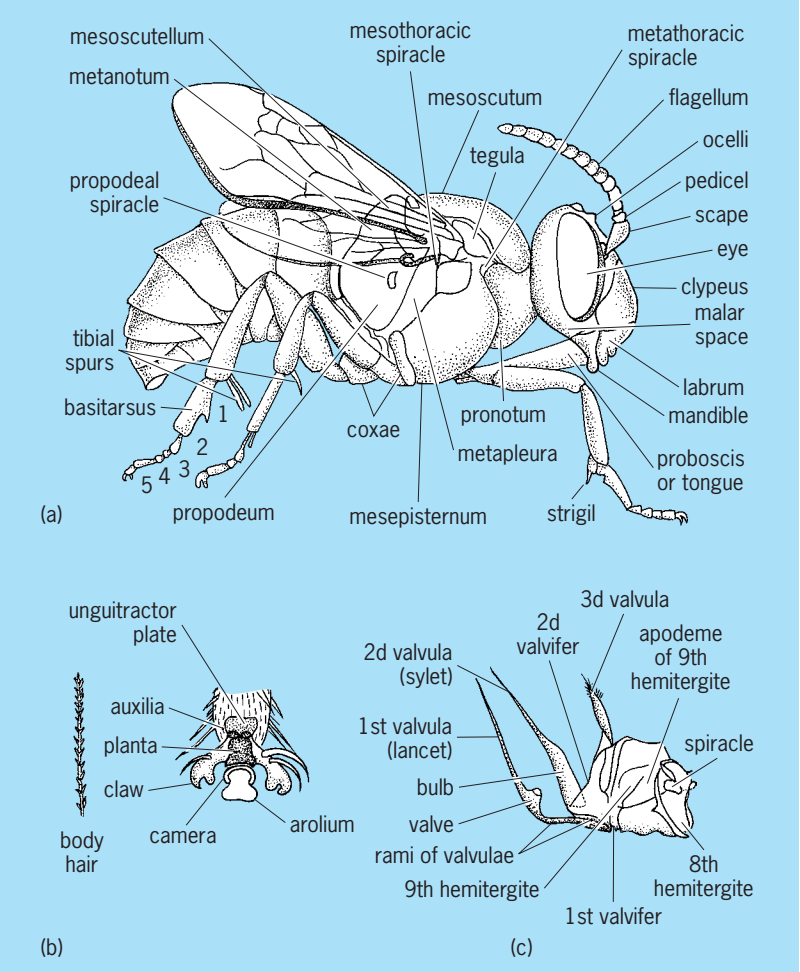

Fig. 4. The bee *Anthophora edwardsii*. (*a*) Adult male. (*b*) Sting. (*c*) Accessory organs of the female showing the important characters used in classifying bees. (*After E. O. Essig, College Entomology, Macmillan, 1942*)

The larvae of Apocrita are legless, maggotlike forms in the later instars with a well-developed head and biting mouthparts. Occasionally, fleshy processes are present which aid in locomotion. The first instar larvae of some Apocrita, particularly those of the Parasitica and of the social parasites among the Aculeata, are very different from the larvae of later instars, a condition known as hypermetamorphosis.

The insect then spins a cocoon, if characteristic of the particular species, and enters the prepupal stage which may be of short or lengthy duration. The pupal stage usually lasts only two or three weeks, and the adult then ecloses, or emerges from the pupal case. The hymenopterous pupa has the appendages free from the body, except in certain Chalcidoidea. It is usually enclosed in a cocoon either constructed of silk alone or with other interwoven substances. The cocoon may be well developed, or it may be just a thin silken lining of the larval cell. It is absent in almost all Chalcidoidea and Cynipoidea, and, occasionally, in all other superfamilies. The adult usually remains in the cocoon for one to several days until the integument is thoroughly hardened,

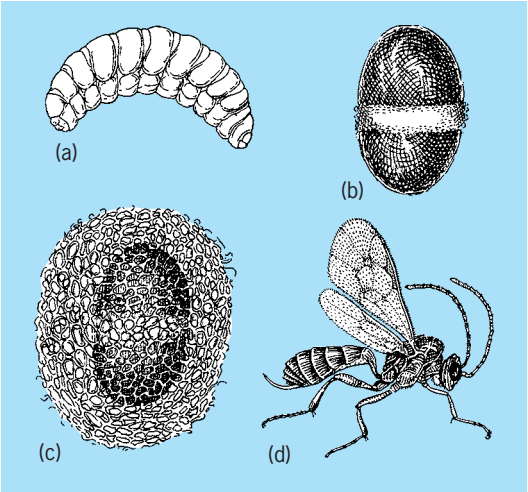

Fig. 5. Life cycle of *Bathyplectes curculionis*, a parasite of alfalfa weevil. (*a*) Larva. (*b*) Pupa. (*c*) Pupa within cocoon of weevil. (*d*) Adult female. (*After E. O. Essig, College Entomology, Macmillan, 1942*)

though occasionally the adult, as in some species of *Osmia* in the Megachilidae, may remain in the cocoon for 10 months before beginning its active adult existence.

Some species have only one generation a year in temperate zones, others have two, and many breed continually during the warmer months. Hymenoptera usually overwinter as prepupae but occasionally ants, social wasps, and some bees overwinter as adults or as larvae, as in some Parasitica. So far as is known, none passes the winter in the egg or pupal stage.

Dimorphism. Sexual dimorphism is often very marked. The two sexes of some species are so dissimilar that earlier students placed them in different genera or families. Even today there are many puzzles, and sexes have not been associated for many of the species having wingless females and winged males. Ordinarily the males are somewhat smaller than females, though the reverse is true in most species having wingless females. *See* SEXUAL DIMORPHISM.

In addition to sexual dimorphism, there are other kinds of variations in adult Hymenoptera. Species of Parasitica or of the social parasites among the Aculeata sometimes vary a great deal in size, depending on the size of the host or on the number of parasites per host. Also, there may be variation in form and size between the generations of species having more than one generation a year.

Behavior. Behavior of the Parasitica is much less complex than in the Aculeata. The primitive pattern is for the female to hunt for her host and then oviposit in or on it, without causing it any great disturbance. The host continues to live a normal life until it is killed by the developing parasite larva. This pattern is typical of many Ichneumonidae. A more complicated pattern is developed by stinging of the host to produce a temporary or permanent paralysis, or even death, before oviposition. This method is typical of many Braconidae and Chalcidoidea. Some species of Parasitica are hyperparasites or secondary parasites; that is, they are parasitic on parasites of the host. Some species may be either primary or secondary parasites, while tertiary parasitism, that is, species parasitic on secondary parasites, also has been reported in a few chalcidoids.

Nests and nesting. The higher Aculeata are remarkable for the amount of care taken in construction of nests for their progeny. The nest may consist of a burrow in the ground or a boring in wood. The nest may also be constructed of clay, paper formed from masticated wood fibers, gums, resins, masticated leaf pulp, pieces of leaves, or wax made in the body of the insect. There may be only one cell per nest, a series of a dozen or so cells arranged in a clump or in a linear series, or there may be a multicellular nest, as in the social wasps and bees.

There are several basic types of nesting behavior among the aculeates. The most primitive pattern is for the wasp to search out its prey, paralyze it by stinging, and then to lay one or several eggs on it, leaving the prey where it was found. This is typical of most Bethylidae and many Scolioidea. A somewhat more complicated pattern, typical of most Pompilidae, is for the wasp to capture and paralyze its prey before constructing a nest for it in a sheltered situation. Another pattern is for the cell to be constructed and an egg deposited in it before the first specimen of prey is brought in. This occurs typically among the Vespidae and a few of the Sphecidae which practice progressive provisioning. A fourth pattern is a variation of the third and consists of preparation of the cell, storing the cell with prey or a pollen-nectar mixture, and then oviposition. This method is typical of most Sphecidae and the Apoidea.

The solitary wasps store varied amounts of prey per cell. The Scolioidea, Pompilidae, and a few Sphecidae use only one specimen of prey per cell. Others, like the Vespidae and most Sphecidae, store from two to more than two dozen specimens of prey per cell depending upon the size of the prey used. Some solitary wasps are quite restrictive in their prey preferences and store only one species. The majority are not so selective and prey upon different species. However, most of these latter prey on species belonging to a single family or order, or to species occurring in a very restricted ecological niche. Thus, in the Sphecidae, species of *Pemphredon* prey on a number of species of aphids (Aphidae), and species of *Ectemnius* on a wide variety of Diptera, but *Symmorphus canadensis* in the Vespidae preys on leafmining larvae belonging to either the Coleoptera or Lepidoptera.

Sociality. Social parasites, also known as cuckoo wasps and bees, have arisen independently in several different families of wasps and bees. Except in the Chrysididae, ordinarily each is closely related, taxonomically, to the particular host wasp or bee which it parasitizes. These parasites either oviposit on the prey of the host wasp as the latter transports it to a nesting site, or they keep the nest under observation and slip into it at some time during the provisioning cycle to deposit the egg while the host female is absent from the nest.

Social life has arisen independently in several stocks of aculeate Hymenoptera, such as the ants, vespid wasps, halictid bees, bumblebees, stingless honeybees, and true honeybees. In the more primitive social insects, the colony dies out at the end of the year except for the recently emerged, fertilized queens which hibernate. These colonies are necessarily less populous than in the more advanced social forms, and are exemplified by the hornets and allies, the halictid bees, and bumblebees. In the higher social forms consisting of the ants, social wasps of the tropics, stingless honeybees, and true honeybees, the colonies are persistent and frequently number several thousands of individuals. The colonies of these wasps and bees divide by swarming, a process by which a queen and a number of workers leave the parent nest and begin a new colony. *See* SOCIAL INSECTS.

Communication. One of the most remarkable facts that has been learned about the honeybee is that the bees employ a sign language to tell other members of the colony that they have located flowers with a copious flow of nectar or abundance of pollen, and approximately how far and in what direction this source is located. If the source is more than 100 yd (90 m) from the hive, the worker performs a tail-wagging dance, usually on the vertical surface of a comb in the darkness. This dance consists of running over a figure-eight pattern with the part between the two loops of the 8 straight. The bee wags her abdomen from side to side when on this straight section. The number of runs over the straight section indicates the distance from the food source, and varies from 9 to 10 runs within 15 s for a distance of 130 yd (120 m) to 3 within 15 s for a distance of 830 yd (750 m). The direction of the source is indicated by the angle of the straight section of the dance. If straight up, the source is toward the Sun; if straight downward, the source is directly away from the Sun; if at an angle to one side or the other of the vertical, then the source is at a corresponding angle and direction away from the Sun. If the food source is less than 100 yd (90 m) from the hive, the worker performs a rapid round dance. The other bees do not receive any clue as to the direction or distance but only an indication that there is a copious food source nearby. *See* ANIMAL COMMUNICATION.

Reproduction. Sawflies deposit their eggs on or in foliage, fruits, stems, or woody tissue, in accordance with the feeding requirements of the larva. Many Parasitica and a few Aculeata deposit more than one egg per host. Some species of Parasitica are polyembryonic, that is, from two to a thousand or more embryos may develop from a single egg. Usually the resulting progeny are all of one sex but exceptions occur. Polyembryony occurs in a few Braconidae, some species of *Macrocentrus*, some species of several genera of Encrytidae, a few Platygasteridae (some *Platygaster*), and one species of Dryinidae, *Aphelopus theliae*.

The size and number of eggs laid is subject to great variation within the order. Most Symphyta and Chalcidoidea deposit a rather small number, from 10 to 50. Most of the Ichneumonoidea deposit larger numbers of eggs, from several hundred to more than a thousand. Some subsocial wasps which practice progressive provisioning may deposit only 6 eggs. Most aculeates lay from 10 to 75 eggs, but apparently some bethylids can lay 150 or more. Honeybees are probably the most fecund of the Aculeata. A queen may lay as many as 1500 eggs per day and 200,000 a year for at least 3 years.

Eggs of the Parasitica are deposited more or less at random on or in the host, but in the Aculeata each species deposits its egg at a particular location on the prey or in the nest.

Sawfly larvae are phytophagous and live on or in foliage, in stems, or in woody tissue. Larvae of the Parasitica may be either internal or external parasites, depending on whether the development takes place inside or outside the host. They can also be classified as solitary if there is only one larva per host, as in most Ichneumonidae, or gregarious, if there are two or more per host as in many Braconidae and Chalcidoidea. Larvae of most wasps are predaceous, feeding on paralyzed insects or spiders stored for them by the mother. Bee larvae feed on pollen-nectar mixture stored by the mother, except for the social parasites, which are predaceous in their first instar on the host egg or young larva and feed on the pollen-nectar mass in the later instars. The larvae of most Aculeata have a blind stomach during most of the larval life so that waste matter is not excreted until the end of the larval life when the connection between stomach and intestine is opened. The larvae of a few wasps, such as Pemphredoninae in the Sphecidae, and of the megachilid bees have this connection opened early in the larval life and excrete small meconial pellets during the last three instars.

Many larvae of the Parasitica and those of the social parasites among the Aculeata are hypermetamorphic, that is, the first instar larva is very different in appearance and behavior from the succeeding instars, which assume the typically maggotlike form of mature larvae of the Apocrita. There are at least 10 different primary larval forms among the species of Apocrita exhibiting hypermetamorphosis. Rarely, the second instar larva differs in appearance from either the first or third.

There are few reliable data on the number of larval instars. Rather meager information indicates that there are three or four instars in various groups of Aculeata.

Silken cocoons are of general occurrence throughout the order, although they are absent in most Chalcidoidea and Cynipoidea, some Formicoidea, and absent sporadically throughout the rest of the Aculeata. The cocoons may be of two layers, as in a few sawflies and wasps, but normally there is only one layer of silk, or the cocoon may consist of only a silken cap. Foreign material may be frequently incorporated in the cocoon. The larva may impregnate the silk with secretions from the gut which cause the cocoon wall to become brittle and varnished. Sand, mud, or prey remnants may be incorporated in the silk of the cocoon. Almost all Hymenoptera

TABLE 2. Economically important Hymenoptera

Family	Scientific name	Common name	Economic importance
Pamphiliidae	*Neurotoma inconspicua*	Plum, web-spinning sawfly	Larvae spin webs and eat foliage of plums
	Pamphilius persicus	Peach sawfly	Larvae roll peach leaves and feed on them
Cimbicidae	*Cimbex americana*	Elm sawfly	Larvae eat foliage of elm, willow, poplar, and maple
Diprionidae	*Diprion hercyniae*	European spruce sawfly	Larvae defoliate spruce; introduced from Europe
	Neodiprion lecontei	Red-headed pine sawfly	Larvae defoliate pines
Tenthredinidae	*Caliroa cerasi*	Pear slug	Larvae skeletonize foliage of pear, cherry, plum; probably introduced from Europe
	Fenusa ulmi	Elm leaf miner	Larvae mine leaves of elm; probably introduced from Europe
	Cladius isomerus	Bristly rose slug	Larvae skeletonize rose leaves
	Hoplocampa cookei	Cherry fruit sawfly	Larvae feed inside cherries
	Pteronidea ribesii	Imported currant worm	Larvae feed on leaves of currants, gooseberries; introduced from Europe
Siricidae	*Tremex columba*	Pigeon tremex	Larvae bore in trunks of weakened or dead maple, oak, elm, other deciduous trees
Cephidae	*Hartigia trimaculata*		Larvae bore in stems of roses, blackberries
	Cephus cinctus	Wheat stem sawfly	Larvae bore in stems of wheat, rye, timothy, and wild grasses
Braconidae	*Aphidius* sp.		Larvae are internal parasites of aphids
	Meteorus sp.		Larvae are internal parasites of lepidopterous and coleopterous larvae
	Macrocentrus sp.		Larvae are internal parasites of lepidopterous larvae
	Apanteles sp.		Larvae are internal parasites of lepidopterous larvae
	Opis sp.		Larvae are dipterous parasites.
	Spathius sp.		Larvae are external parasites of coleopterous larvae
Ichneumonidae	*Scambus* sp.		Larvae are internal parasites of small lepidopterous larvae in leaf mines, leaf rolls, galls
	Polysphincta sp.		Larvae are external parasites of spiders
	Ephialtes sp.		Larvae are internal parasites of lepidopterous pupae
	Megarhyssa macrurus		Larvae are external parasites of wood-boring sawfly larvae
	Gelis sp.		Larvae are parasites in cocoons of other Ichneumonoidea and in egg sacs of spiders
	Acroricnus aequatus		Larvae are parasites in nests of mud dauber wasps
	Diplazon laetatorius		Eggs are laid in syrphid (Diptera) eggs or young larvae, and adults emerge from host puparia
Mymaridae	*Anagrus armatus*		Larvae are parasites in eggs of Hemiptera
Trichogrammatidae	*Trichogramma minutum*		Larvae are parasites in eggs of other insects
Eulophidae	*Sympiesis* sp.		Larvae are parasites of leaf-mining coleopterous and lepidopterous larvae
	Tetrastichus sp.		Larvae are parasites of a wide variety of insect eggs and larvae including those of other Chalcidoidea
Thysanidae	*Thysanus* sp.		Larvae are parasites of Homoptera or of other Chalcidoidea which parasitize Homoptera
Encyrtidae	*Copidosome gelechiae*		Polyembryonic parasites of gall-making lepidopterous larvae
	Aphycus sp.		Larvae are parasites of scale insects
	Ooencyrtus sp.		Larvae are parasites in eggs of other insects
Agaontidae	*Blastophagus psenes*	Fig wasp	Lives within figs and fertilizes them; introduced from Europe
Torymidae	*Megastigmus* sp.		Larvae live in seeds
Eurytomidae	*Harmolita tritici*	Wheat jointworm	Larvae bore in stems of wheat
	Harmolita grandis	Wheat strawworm	Larvae bore in stems of wheat
	Bruchophagus gibbus	Clover seed chalcid	Larvae develop in seeds of clover, alfalfa; possibly an introduced species
Cynipidae	*Diplolepis rosae*	Mossy rose gall	Larvae develop in galls on rose stems; introduced from Europe
	Acraspis erinacei	Oak hedgehog gall	Agamic generation (no males); develops in hedgehog gall on oak leaves; sexual generation develops in soft galls in oak buds
	Amphibolips confluenta	Large oak-apple gall	Larvae develop in large globular galls on oak leaves
Evaniidae	*Evania appendigaster*		Parasitic in egg cases of cockroaches; introduced
Scelionidae	*Telenomus* sp.		Parasitic in eggs of many orders of insects
Chrysididae	*Chrysis* sp.	Cuckoo wasps	Social parasites of other wasps and bees
Bethylidae	*Cephalonomia tarsalis*		Parasites of beetle larvae infesting stored grains
	Scleroderma sp.		Parasites of old-house borer (Coleoptera); have very painful sting
	Goniozus sp.		Parasites of lepidopterous leaf rollers
Tiphiidae	*Tiphia vernalis*	Spring tiphia	Parasites of Japanese beetle larvae; introduced from Japan, Korea, China
Multillidae	*Dasymutilla occidentalis*	Cow killer	Parasite of bumblebee pupae
Formicidae	*Eciton* sp.	Legionary ants	Predacious on other insects
	Pheidole sp.	Harvesting ants	Nest in ground and feed on seeds
	Monomorium pharaonis	Pharaoh ant	A pest ant which nests in buildings; introduced

TABLE 2. Economically important Hymenoptera (*cont.*)

Family	Scientific name	Common name	Economic importance
Formicidae (cont.)	*Solenopsis saevissima var. richteri*	Imported fire ant	Nests in soil; very aggressive; introduced from South America
	Tetramorium caespitum	Pavement ant	Infests houses; introduced
	Atta sp.	Fungus ants	Nest in soil and feed upon fungi which they cultivate on beds of masticated leaves
	Tapinoma sessile	Odorous house ant	Infests houses
	Camponotus sp.	Carpenter ants	Many species nest in wood
	Lasius interjectus	Larger yellow ant	Nests around building foundations and beneath cellar floors
	Myrmecocystus sp.	Honey ants	Nest in soil in Southwest; some workers store large quantities of honey until abdomen is greatly distended
	Formica exsectoides	Allegheny mound ant	Nests in ground beneath large mounds of excavated soil
	Polyergus sp.	Slave-making ants	Enslave workers of other ants to build their nests and feed their young
Vespidae	*Vespa crabro germana*	Giant hornet	Social wasp nesting in sidings of buildings and hollow trees; introduced from Europe
	Vespula (Vespula) sp.	Yellow jacket	Social wasps nesting in ground
	Vespula (Dolichovespula) maculata	Bald-faced hornet	Social wasp which builds large paper nests in trees and shrubs
	Polistes sp.	Paper wasps	Social wasps which build a single umbrella-shaped comb, frequently under eaves or porch roofs
	Eumenes sp.	Jug wasp	Solitary wasps which build small clay jugs in which to rear their young; prey on caterpillars
Pompilidae	*Pepsis* sp.	Tarantula hawks	Prey on tarantulas
Sphecidae	*Trypoxylon politum*		Builds the familiar clay "pipe-organ" nests in which it stores many small spiders
	Chlorion ichneumoneum	Great golden digger wasp	Nests in soil and preys on long-horned grasshoppers
	Sceliphron caementarium	Black and yellow mud dauber	Builds clay cells in which it stores many small spiders
	Sphecius speciosus	Cicada killer	Nests in soil and preys on adult cicadas
	Stictia carolina	Horse guard	Nests in ground and preys on a variety of flies
Colletidae	*Colletes* sp.		Solitary bees which nest in ground, frequently gregariously, and collect pollen and nectar from a variety of flowers
Andrenidae	*Andrena* sp.		As in Colletidae
Halictidae	*Halictus* sp.		As in Colletidae, except that some species are subsocial or social
Megachilidae	*Megachile* sp.	Leaf-cutter bees	Solitary bees which nest in cavities in wood or in ground, and make cells from pieces cut from leaves
Apidae	*Anthophora* sp.	Mining bees	Solitary ground-nesting bees
	Xylocopa virginica	Carpenter bee	Solitary bee which excavates tunnels in solid wood in which it rears its brood
	Bombus sp.	Bumblebees	Social bees which next on or in ground, frequently in old mouse nests
	Trigona sp.	Stingless honeybees	The honeybee of the tropics; does not occur in North America
	Apis mellifera	Honeybee	The beekeeper's bee; escaped bees nest in hollow trees or house sidings; introduced from Europe

that spin cocoons produce the silk from labial glands through the mouthparts, but those Chalcidoidea that construct cocoons produce the silk in the Malpighian tubules and spin it from the anus.;

Economically important species. Many of the Hymenoptera are of economic importance and include both useful and destructive forms. The species listed in **Table 2** are North American except where noted to the contrary. Common names are given, and a brief note is included explaining the importance of each species. *See* ENTOMOLOGY, ECONOMIC.

Phylogeny

The earliest known fossil Hymenoptera are from the Middle Jurassic. Only Symphyta and Parasitica are known from these strata. The absence of Aculeata perhaps substantiates the claim that this section evolved from the Parasitica; or its absence may be only apparent and due to the meager fossil record. Unfortunately, the wing venation of these fossil sawflies is quite specialized and offers no clues to the derivation of the order. The consensus is that the Hymenoptera probably arose from an ancestral type that also gave rise to the Neuroptera and other related orders. The Baltic amber, which dates from the Tertiary, preserved many ants and a few bees and other Hymenoptera. The bees are all members of genera which are now extinct, but some of the ants are thought to be the same as species living now.

Karl V. Krombein

Bibliography. R. H. Arnett and R. L. Jacques, *Insect Life*: *A Field Entomology Manual for the Amateur Naturalist*, 1985; R. J. Elzinga, *Fundamentals of Entomology*, 5th ed., 1999; H. E. Evans et al., *Insect Biology*: *A Textbook of Entomology*, 1984; R. L. Jeanne, *Interindividual Behavior Variability in Social Insects*, 1988; B. Klausnitzer, *Insects*: *Their Biology and Cultural History*, 1987; R. D. Shenfelt, *Hymenopterorum Catalogus*, 1980; F. Taylor and R. Karban (eds.), *The Evolution of Insect Life Cycles*, 1986; A. Wooten, *Insects of the World*, 1985.

Hymenostomatida

An order of the Holotrichia which contains many species that often are of small size and fairly uniform ciliation. Primarily, these protozoa are of importance as the first possessors of a definite, though inconspicuous, buccal ciliature. This ciliature consists of an undulating membrane on the right side of the buccal cavity and an adoral zone of membranelles that is primitively composed of three membranelles on the left side. This tetrahymenal, or four-part, buccal ciliary apparatus is considered the fundamental condition from which the oral ciliature of many subsequent higher groups evolved. *See* CILIOPHORA.

The majority of hymenostomes are free-living fresh-water forms. On the basis of differences in their patterns of stomatogenesis (new-mouth formation), one group of hymenostomes has been authoritatively considered as representing a new and different ordinal group of the Holotrichia; but the better-known, if more conservative, scheme of classification is espoused here.

The parasitic genus *Ichthyophthirius* is often mistakenly considered to be a gymnostome. *Paramecium* (**illus.** *a*) is the best-known genus of ciliates in the entire subphylum. It is a good-sized widely distributed ciliate. Being easy to recognize and culture, it is a much-studied form. It was at one time classified as a trichostome. As an experimental animal, *Paramecium* has played a major role in the advance of protozoan genetics. *Tetrahymena* (illus. *b*), beginning to rival *Paramecium* as a favorite ciliate in much experimental work, owes its scientific popularity primarily to its ability to grow axenically, that is, free from all other organisms, in a chemically defined medium. Its species are the first animal organisms, excluding the green plantlike flagellate protozoans, to be so grown. The achievement of axenic growth represents a great forward step in experimental studies, particularly those of a biochemical nature, in cancer research and in other important fields of direct, immediate interest to humans. *See* AXENIC CULTURE; GYMNOSTOMATIDA; HOLOTRICHIA; PROTOZOA.

John O. Corliss

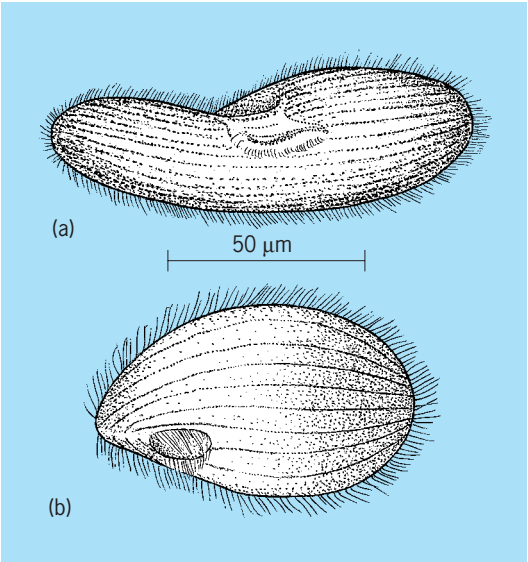

Hymenostomatida. (*a*) *Paramecium*. (*b*)*Tetrahymena*.

Hyperbaric oxygen chamber

A specially equipped pressure vessel used in medicine and physiological research to administer oxygen at elevated pressures.

Basic principle. Under normal conditions the red blood cells provide the main transport mechanism for distributing oxygen through the bodies of warm-blooded animals. In humans less than 5% of the oxygen in the body is dissolved in body fluids. The transport capacity of red blood cells permits warm-blooded animals to maintain high body temperatures even in cold climates, and to supply the heavy oxygen demand of a large active brain. However, circulation of the red cells through the blood vessels requires a great amount of work by the heart and can be reduced or stopped by damage or blockage of the blood vessels.

The amount of oxygen dissolved in the body fluids is related to the pressure of oxygen in the lungs (Henry's law). When a person breathes pure oxygen, the amount of oxygen dissolved in body fluids is about six times that when breathing air. This is still too low to supply the needs of the human body. However, breathing pure oxygen at three times normal air pressure causes the amount of oxygen dissolved in body fluids to be equal to that normally carried by the red cells. Research animals have been kept alive in a hyperbaric chamber for some time with all the red cells removed from their blood. At the end of the experimental period the red cells were returned and the animals subsequently led perfectly normal lives.

A principal advantage of dissolved oxygen is that it can be transported throughout the body wherever there is fluid of any sort. It is not limited to circulation through blood vessels. A second advantage is that a given volume of blood contains twice the normal amount of oxygen. The high level of oxygen in the body aids the patient in the following ways:

(1) Oxygen can be carried past an obstruction in the circulatory system, thus relieving oxygen-starved tissues. (2) The work load on the heart can be reduced, since one-half the normal blood flow will provide the normal amount of oxygen required by the body. (3) Poisons such as carbon monoxide can be eliminated. (4) Anaerobic bacteria such as tetanus can be destroyed. (5) The effectiveness of radiation treatment of cancer is increased. *See* RADIATION BIOLOGY; TETANUS.

Equipment. For physiological studies and for some types of patient treatment chambers just large enough for one person have been built. These chambers are relatively inexpensive to build but have limited usefulness because the patient cannot be treated or cared for while sealed up inside the oxygen-filled pressure chamber.

In order to give the patient adequate care or to perform surgery, it is necessary to use much larger equipment. The **illustration** shows a hyperbaric facility consisting of two large chambers. One chamber contains a fully equipped operating room, while the other contains a small medical ward. These large chambers are filled with ordinary air that is compressed to the appropriate pressure. The patients breathe pure oxygen through masks or tents; the attending doctors and nurses breathe the air that fills the chamber.

Each chamber consists of a steel pressure vessel which is roughly 12 ft (4 m) in diameter and about 45 ft (14 m) long. Larger vessels have been employed, but they are extremely difficult to install and require excessively large compressors to provide a fresh air supply.

Inside, the vessel is divided into three areas. In the front area is a large entry lock in which patients on stretchers can be compressed or decompressed while the main section remains at elevated pressure. The main area of the chamber is a room about 30 ft (9 m) wide because of the curvature of the walls. In one of the chambers this room contains a fully equipped surgical operating table. In the other chamber, beds for six patients can be accommodated. Beyond the main section there is a small lavatory and another small lock through which attending personnel can enter or leave quickly during an operation. There is also an instrument lock through which small equipment and supplies can be passed.

In a nearby room outside the chambers air compressors and air conditioning equipment are located. From a central console the operator can monitor all that is happening in each of the chamber areas and can activate any of the mechanical equipment by remote control.

Proper conditioning of the air under a wide variety

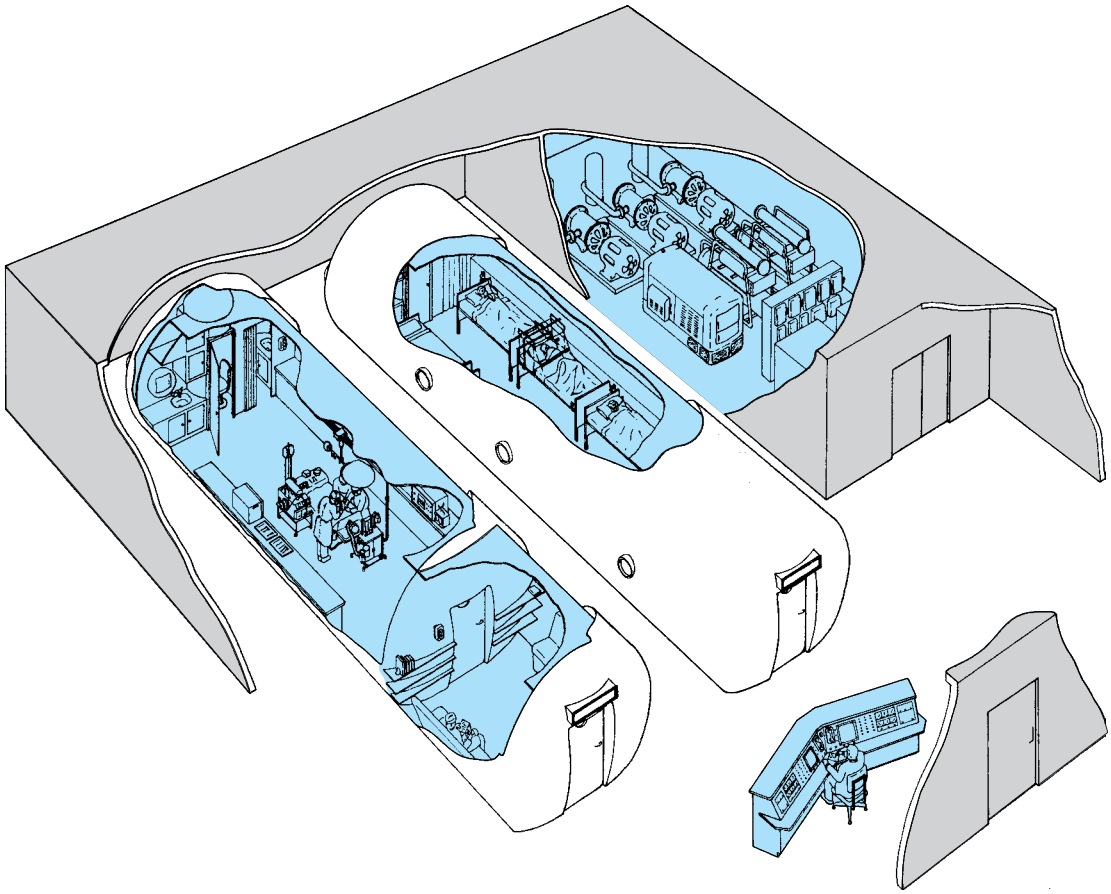

Hyperbaric oxygen treatment facility.

of operating pressures is necessary for the comfort and safety of the people in the chamber. When air is compressed, that part of the total pressure represented by each component increases to the same extent (Dalton's law). For example, on a comfortable day relative humidity is about 50–75%. This means that the water vapor is 50–75% of the saturation pressure. If this air is compressed to three times its normal pressure, the pressure of each component including the water vapor will be increased threefold. On the other hand, the water vapor saturation pressure will remain constant as long as the temperature remains the same. Therefore, the humidity will rise to 150–225% which means that some of the water vapor must condense as fog or even rain. Without any special provision to remove moisture the hyperbaric chamber would always be foggy and damp, and the atmosphere would have a very unpleasant effect on the senses. To keep the atmosphere clean and comfortable, the air in the chamber is changed completely every 20 min. The fresh air is dried and cooled after being compressed.

Providing normal facilities in a hyperbaric chamber presents novel problems. For example, normal water pressure may not be sufficient to make the water flow out of a faucet into the high-pressure environment in the chamber. Booster pumps raise the pressure enough to ensure normal flow. Draining wastewater out of the chamber under 3 atm (300 kilopascals) pressure could have rather spectacular results if the drain lines led directly to the sewers. Special waste-receiving tanks are provided which permit the high pressure to bleed off before releasing the wastewater to the normal drainage system.

Changes in pressure must take place rather gradually for the comfort and safety of the people in the chamber. The ear-popping problems that commonly occur in express elevators and airplanes are considerably magnified because the total change in pressure is considerably greater. Normally a period of 5–10 min is required to pressure up or down.

A special problem exists for the attending personnel who breathe the compressed air. This is the well-recognized problem of the elimination of dissolved nitrogen from the body. For working periods in the chamber up to about 1 h no problem is involved, because nitrogen dissolves slowly. However, for periods of more than 1 h enough nitrogen will accumulate to cause formation of gas bubbles in the bloodstream if decompression takes place too fast. These gas bubbles cause the condition known as caisson disease, or the bends. Tables governing the safe decompression schedules for various working times and pressures have been prepared by the U.S. Navy.

For example, if a surgical team conducts an open-heart operation lasting 3 h in a hyperbaric chamber at three times normal pressure, the Navy tables indicate that decompression should include a 19-min hold at $1^2/_3$ times normal pressure and another hold of 79 min at $1^1/_3$ times normal pressure. Altogether this means that decompression requires almost 2 h after a 3-h working period. This would not be required for the patient, since the pure oxygen administered during the surgery does not present any hazard. *See* DECOMPRESSION ILLNESS.

Because of the potentially long periods of time which may be required to decompress, the people in the chamber are relatively isolated from the outside world. They are highly dependent on the chamber operator, who remains at the control console continually as long as any person is in the chamber. Communication is maintained by closed-circuit television with the main section and each of the personnel locks. The operator continually observes the pressure, temperature, and gas composition in the chamber.

While loss of electric power would not of itself be dangerous, the chamber would soon become rather uncomfortable without the constant flow of fresh air, and it would soon become necessary to decompress and leave the chamber. This could mean stopping an operation before completion, which could pose serious problems. Emergency power supplies are therefore provided, and clean dry air is stored in high-pressure cylinders so that no interruption disrupts the work of the physicians. During the great power failure that blacked out the Northeast on November 9–10, 1965, the hyperbaric facility at Mount Sinai Hospital in New York City was an island of light in otherwise complete darkness.

History. Physicians have long been interested in the medical applications of elevated pressures. Orval Cunningham observed beneficial results from such conditions when treating influenza victims during World War I. However, treatment with oxygen at atmospheric pressure produced the same benefits with much less difficulty. It was not until the 1950s that the combination of elevated pressure and pure oxygen was suggested, especially for open-heart surgery. In 1965 a large chamber was built at the University of Amsterdam in the Netherlands. Since that time facilities have been constructed at a number of medical schools and large hospitals in Great Britain, Canada, and the United States. *See* OXYGEN; RESPIRATION.

Arthur W. Francis

Bibliography. B. Fischer et al., *Handbook of Hyperbaric Oxygen Therapy*, 1988; G. S. Innes (ed.), *The Production and Hazards of a Hyperbaric Oxygen Environment*, 1970.

Hyperbola

A curve cut from a cone or revolution by a plane that intersects both nappes of the cone and does not contain the apex (**Fig. 1**). In analytic geometry it is shown, as shown in **Fig. 2**, that a hyperbola is the locus of points P in a plane, such that $PF = \epsilon \cdot PD$, where PF and PD denote the distances of P from a fixed point F (focus) and a fixed line (directrix) of the plane, respectively, and ϵ is a constant, greater than 1. It is also the locus of points P, the difference of whose distances from two fixed points F, F' (foci)

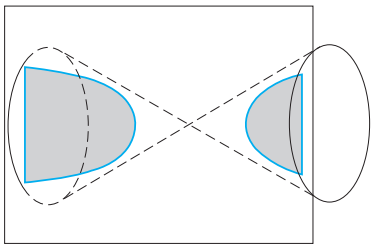

Fig. 1. Hyperbola as a conic section.

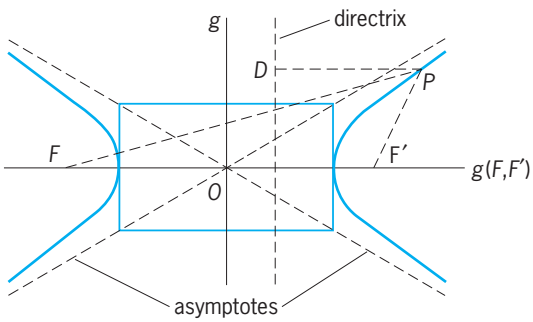

Fig. 2. Hyperbola as a locus of points.

$PF - PF'$ is a constant $2a$ that is less than the distance $2c$ between the foci. The curve is symmetric to the line $g(F, F')$ determined by F, F' and to O, their midpoint. It consists of two branches that are images of each other in the line g through O, perpendicular to $g(F, F')$. There are two lines through O, making equal angles with $g(F, F')$, and to each of which points on each branch get indefinitely close; that is, if point P traverses either branch of the hyperbola, its distance from these lines approaches zero.

The lines are called asymptotes of the hyperbola. If the asymptotes are mutually perpendicular, the hyperbola is called rectangular or equilateral, since then its transverse axis $AA' = 2a$, where A, A' are on $g(F, F')$ with $OA = OA'$, equals its conjugate axis $BB' = 2b = 2(c^2 - a^2)^{1/2}$ (B, B' on g, $OB = OB'$). The eccentricity $\epsilon = c/a$ of every equilateral hyperbola is $\sqrt{2}$. The asymptotes of each hyperbola are the diagonals of a rectangle with center at O, having two sides of length $2a$ parallel to $g(F, F')$, and the other two parallel sides have length $2b$. If an ellipse and a hyperbola have the same foci F, F', they are called confocal. Through each point of the plane there is exactly one ellipse and one hyperbola with the same given foci. They intersect each other at right angles at each of the four points they have in common.

The area of the triangle formed with the asymptotes of a hyperbola by a variable tangent is equal to ab. Conversely, if a line cuts the asymptotes of a hyperbola in two points that lie on the same side of the conjugate axis and forms with them a triangle of area ab, then the line is tangent to the hyperbola. The tangent and normal at a point P bisect the angles formed by the lines joining P to the foci F, F'. *See* ANALYTIC GEOMETRY; CONIC SECTION. Leonard M. Blumenthal

Hyperbolic function

The hyperbolic sine and cosine of a real or complex variable z are defined by Eqs. (1). Both $\sinh z$ and

$$\sinh z = \frac{e^z - e^{-z}}{z} \qquad \cosh z = \frac{e^z + e^{-z}}{z} \tag{1}$$

$\cosh z$ have a period $2\pi i$ of e^z. From $De^z = de^z/dz = e^z$, Eqs. (2) are obtained.

$$D \sinh z = \cosh z \qquad D \cosh z = \sinh z \tag{2}$$

Since

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

by definition, Eqs. (3) hold and the series converge for all z and yield relations (4). Thus relations be-

$$\begin{aligned} \sinh z &= \sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!} \\ \cosh z &= \sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!} \end{aligned} \tag{3}$$

$$\begin{aligned} \sinh iz &= i \sin z \quad & \cosh iz &= \cos z \\ \sin iz &= i \sinh z \quad & \cos iz &= \cosh z \end{aligned} \tag{4}$$

tween the circular functions become hyperbolic (and vice versa) when z is replaced by iz. In particular $\cos^2 z + \sin^2 z = 1$ becomes $\cosh^2 z - \sinh^2 z = 1$; and addition theorems (5) become Eqs. (6). Moreover Eqs. (1) yield eulerian equations (7).

$$\begin{aligned} \sin(z + \zeta) &= \sin z \cos \zeta + \cos z \sin \zeta \\ \cos(z + \zeta) &= \cos z \cos \zeta - \sin z \sin \zeta \end{aligned} \tag{5}$$

$$\begin{aligned} \sinh(z + \zeta) &= \sinh z \cosh \zeta + \cosh z \sinh \zeta \\ \cosh(z + \zeta) &= \cosh z \cosh \zeta + \sinh z \sinh \zeta \end{aligned} \tag{6}$$

$$i \sin z = \frac{e^{iz} - e^{-iz}}{2} \qquad \cos z = \frac{e^{iz} + e^{iz}}{2}, \tag{7}$$

With $z = x + iy$, the addition theorems give Eqs. (8). Hence Eqs. (9) hold and all zeros of sinh

$$\begin{aligned} \sinh z &= \sinh x \cos y + i \cosh x \sin y \\ \cosh z &= \cosh x \cos y + i \sinh x \sin y \end{aligned} \tag{8}$$

$$\begin{aligned} |\sinh z|^2 &= \sinh^2 x + \sin^2 y \\ |\cosh z|^2 &= \sinh^2 x + \cos^2 y \end{aligned} \tag{9}$$

z and $\cosh z$ lie on the axis $x = 0$ at the points $y = 2n\pi i$ and $y = (2n + {}^1\!/_2)\pi i$, respectively. All zeros are simple.

The functions e^z, $\sin z$, $\cos z$, $\sinh z$, $\cosh z$ are analytic throughout the complex plane; all have an essential singularity at $z = \infty$; e^z has no zeros, $\sin z$ and $\cos z$ vanish only on the axis of reals, $\sinh z$ and $\cosh z$ only on the axis of imaginaries.

The hyperbolic tangent and cotangent are defined by Eqs. (10). They have the period πi and have

$$\tanh z = \frac{\sinh z}{\cosh z} \qquad \coth z = \frac{\cosh z}{\sinh z} \tag{10}$$

simple poles at the zeros of $\cosh z$ and $\sinh z$,

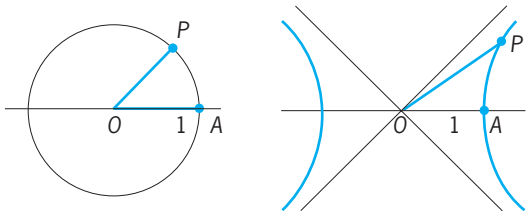

Circle, $x^2 + y^2 = 1$; hyperbola; $x^2 - y^2 = 1$.

respectively. Moreover Eqs. (11) hold true. By definition Eqs. (12) hold.

$$D \tanh z = \frac{1}{\cosh^2 z} \qquad D \coth z = -\frac{1}{\sinh^2 z} \tag{11}$$

$$\operatorname{sech} z = \frac{1}{\cosh z} \qquad \operatorname{csch} z = \frac{1}{\sinh z} \tag{12}$$

The hyperbolic functions are related to the equilateral hyperbola in much the same way that the circular functions are related to the circle (equilateral ellipse). For the hyperbola of Eqs. (13), and for the circle of the area of the sector *OAP* shown in the **illustration** the circuit integral is given by Eq. (15). In both cases of Eqs. (14), parameter $t = \sigma$.

$$x = \cosh t \qquad y = \sinh t \tag{13}$$

$$x = \cos t \qquad y = \sin t \tag{14}$$

$$\sigma = \frac{1}{2} \oint_{OAP} (x\,dy - y\,dx) = \frac{1}{2} \oint_0^t dt = \frac{t}{2} \tag{15}$$

See HYPERBOLA; TRIGONOMETRY. Louis Brand

Hyperbolic navigation system

A navigation system that produces hyperbolic lines of position (LOPs) through the measurement of the difference in times of reception (or phase difference) of radio signals from two or more synchronized transmitters at fixed points. Such systems require the use of a receiver which measures the time difference (or phase difference) between arriving radio signals. Assuming the velocity of signal propagation is relatively constant across a given coverage area, the difference in the times of arrival (or phase) is constant on a hyperbola having the two transmitting stations as foci (**Fig. 1**). Therefore, the receiver measuring time or phase difference between arriving signals must be located somewhere along the hyperbolic line of position corresponding to that time or phase difference. If a third transmitting station is available, the receiver can measure a second time or phase difference and obtain another hyperbolic line of position. The intersection of the lines of position provides a two-dimensional navigational fix (Fig. 1). User receivers typically convert this navigational fix to latitude and longitude for operator convenience.

The choice of frequency and locations of transmitters determines both the utility and the accuracy of hyperbolic navigation systems. In general, longer baselines (Fig. 1) enhance accuracy; however, transmitter power limitations may constrain a system to shorter baselines in order to maintain accurate synchronism.

Accuracy is a primary concern for any navigation system. In hyperbolic systems (as with all navigation systems) there are fundamental limits to accuracy. Important geometric concepts in describing accuracy limits for hyperbolic navigation systems are LOP crossing angles, gradient, and horizontal dilution of precision (HDOP).

LOP crossing angle. This is the angle between lines of tangency at the intersection of the lines of position. The first station's signal may be denoted as a master, and all other time differences can be measured relative to the master signal. Each line tangent to a hyperbolic line of position at the user's position *P* (**Fig. 2**) bisects one of the angles, α or β, formed at *P* between one of the two secondary stations and the master station. As a result, the LOP crossing angle, γ, equals $\alpha/2 + \beta/2$, or one-half the angle between the two secondary stations as observed at the user's position. LOP crossing angle is purely a function of the bearings to the two secondary stations being used, and is independent of the bearing to master. For a two-LOP crossing, angles of 120° are considered to be optimal due to correlation in time differences; however, such cannot be achieved throughout the coverage area. As crossing angles diminish from 90° to about 30°, the fix-error contour increases in area and elongates to form an ellipse. Crossing angles of less than 30° are generally considered unacceptable.

Gradient. This is a measure of the change in calculated geodetic position caused by a change in the measured time or phase difference; that is, gradient can be considered an expression of geodetic error sensitivity. Gradient varies depending on user location within the coverage area. For hyperbolic systems that measure difference in times of reception of two radio signals propagating at a nominal velocity of 983.24 ft/μs (299.69 m/μs), the gradient Δ

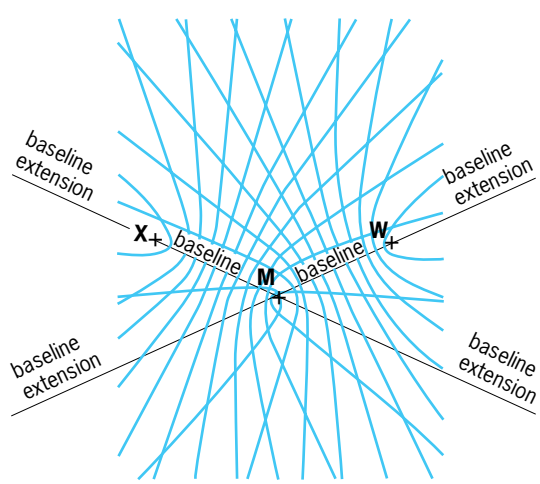

Fig. 1. Typical grid of hyperbolic lines of position formed by a master (M) and two secondary stations (W and X).

can be expressed by the equation below, where ψ

$$\Delta = \frac{491.62}{\sin(\psi/2)}\text{ft}/\mu\text{s} = \frac{149.85}{\sin(\psi/2)}\text{ m}/\mu\text{s}$$

is the angular difference in bearing between the two stations (that is, one master and one secondary) as observed at the user's position. Gradient is purely a function of the difference in bearing to master and bearing to secondary. The minimum gradient occurs for positions located on the baseline, that is, directly between the two stations (Fig. 1), where the angle ψ is 180°. Gradient tends toward infinity for positions located on the baseline extension (Fig. 1), where the angle ψ tends toward zero. For a single pair of transmitting stations, smaller gradient implies improved accuracy, so system planners typically specify that hyperbolic navigation system baselines occur over areas where improved accuracy is needed. For a single pair of transmitting stations, no reliable position information can be obtained by measuring time or phase differences in the vicinity of the baseline extension (Fig. 1), since gradient tends to infinity at these locations.

Horizontal dilution of precision. This is a dimensionless multiplier which takes into account the geometric effects of both gradient and crossing angle, and shows how accuracy is degraded as a function of bearing to the master and two secondary stations. **Figure 3** shows a master and two secondary transmitters along with curves of constant HDOP. Small values of HDOP (as seen within the triangle *MXY*) suggest areas of improved accuracy. HDOP is multiplied by ranging error (σ) to obtain root mean square position error (drms), and by twice ranging error (2σ) to obtain 2 drms. For example, if gaussian signal stability statistics are assumed, then the receiver's actual position will be outside a circle of radius 2 drms, centered at the receiver's computed fix position, between 1.8 and 4.6% of the time. In other words, if a receiver is allowed to record positions over a period of time, between 95.4 and 98.2% of the calculated positions will be within a circle of radius 2 drms centered at the actual position. *See* AIR NAVIGATION; DECCA; LORAN; MARINE NAVIGATION; STATISTICS. Benjamin B. Peterson; Richard J. Hartnett

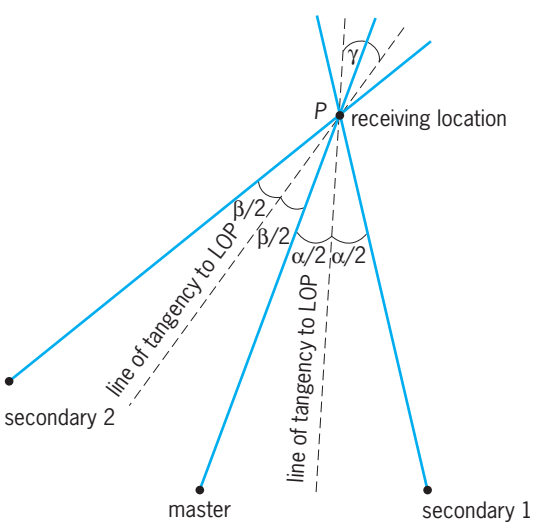

Fig. 2. Individual lines of position from a master and two secondaries at location *P*. γ = LOP crossing angle.

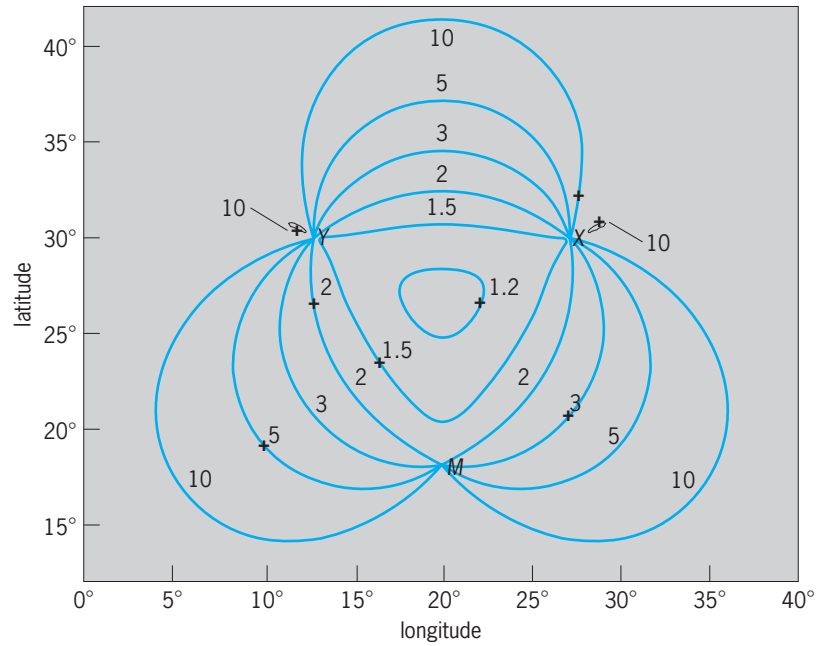

Fig. 3. Typical curves of constant horizontal dilution of precision (HDOP) for a master (*X*) and two secondary stations (*X* and *Y*). Numbers give HDOP values of contours.

Bibliography. Special issue on global navigation systems, *Proc. IEEE*, vol. 71, no. 10, 1983; L. Tetley and D. Calcutt, *Electronic Aids to Navigation*, 2d ed., 1992; U.S. Department of Transportation, *U.S. Coast Guard Loran-C User Handbook*, COMDTPUBP16562.6, 1992; J. C. Whitaker (ed.), *The Electronics Handbook*, CRC Press, 1996.

Hypercharge

A quantized attribute, analogous to electric charge, introduced in the classification of a subset of elementary particles—the so-called baryons—including the proton and neutron as its lightest members. As far as is known, electric charge is absolutely conserved in all physical processes. Hypercharge was introduced to formalize the observation that certain decay modes of baryons expected to proceed by means of the strong nuclear force simply were not observed. *See* ELECTRIC CHARGE.

Unlike electric charge, however, the postulated hypercharge was found not to be conserved absolutely; the weak nuclear interactions do not conserve hypercharge—and indeed can change hypercharge by ± 1 or 0 units. These observations establish certain relationships but do not define the hypercharge. The simplicity of the relationships suggested a hypercharge (symbol *Y*) as follows: for the proton and neutron, $Y = +1$; for the pion, $Y = 0$.

For example, consider the heavier Λ^0 baryon, which in principle is energetically unstable against decay into a proton and a pion in a characteristic time of about 10^{-23}s. But it is observed that the Λ^0

lives about 10^{13} times longer than this characteristic time, suggesting that, for this baryon, $Y = -1$ and thus it can decay into a proton and a pion only by means of the weak nuclear interaction. Systematic observations of this kind have permitted assignment of hypercharge values to all the baryons and antibaryons.

When the known baryons are classified according to their electric charge and their hypercharge, they naturally group into octets in the scheme first proposed by M. Gell-Mann and K. Nishijima. The quarks, hypothesized as the fundamental building blocks of matter, must have fractional hypercharge as well as electrical charge; the simplest quark model suggests values of 1/3 and 2/3, respectively. *See* BARYON; ELEMENTARY PARTICLE; QUANTUM MECHANICS; QUARKS; SYMMETRY LAWS (PHYSICS); UNITARY SYMMETRY.

D. Allan Bromley

Hyperfine structure

A closely spaced structure of the spectrum lines forming a multiplet component in the spectrum of an atom or molecule, or of a liquid or solid. In the emission spectrum for an atom, when a multiplet component is examined at the highest resolution, this component may be seen to be resolved, or split, into a group of spectrum lines which are extremely close together. This hyperfine structure may be due to a nuclear isotope effect, to effects related to nuclear spin, or to both.

Isotope effect. The element zinc, for example, has three relatively abundant naturally occurring nuclear isotopes, ^{64}Zn, ^{66}Zn, and ^{68}Zn. The radius of a nucleus increases with the nuclear mass and, for a given element, the Coulomb interaction of the nucleus with the atomic *s*-electrons will be slightly weaker when the nuclear size is larger. This nuclear size effect causes a slight shift of certain of the spectrum lines, and this shift will be different for each isotope. For a mixture of ^{64}Zn, ^{66}Zn, and ^{68}Zn, certain of the multiplet components will thus consist of three closely spaced lines, one line for each isotope. A study of the isotope effect for an element leads, for example, to information about the dependence of the nuclear size on isotope mass, that is, on the number of neutrons in the isotope. *See* ISOTOPE SHIFT.

Structure due to nuclear spin. For the zinc isotopes discussed above, the nuclear spin $I = 0$, and these nuclei will be nonmagnetic and, in effect, have a spherical shape. If $I \neq 0$, however, two new nuclear properties may be observed. The nucleus may have a magnetic moment, and the shape of the nucleus may not be spherical but rather may be that of a prolate or oblate spheroid; that is, it may have a quadrupole moment. *See* SPIN (QUANTUM MECHANICS).

Atoms and molecules. If the electrons in an atom or a molecule have an angular momentum, the electron system may likewise have a magnetic moment. An electron quadrupole moment may also exist. The magnetic moment of the nucleus may interact with the magnetic moment of the electrons to produce a magnetic hyperfine structure. The quadrupole moments of the nucleus and of the electrons may couple to give an electric quadrupole hyperfine structure. In a simple example, the magnetic and the electric quadrupole hyperfine structure may be described by an energy operator H_{hfs} which has the form below,

$$H_{\text{hfs}} = A\bar{I} \cdot \bar{S} + P\left[\bar{I}_z{}^2 - 3I(I+1)\right]$$

where $\bar{S}$ is an operator describing electron spin, $\bar{I}$ and $\bar{I}_z$ are operators describing the nuclear spin and its z component, and I gives the magnitude of the nuclear spin. A and P are coupling constants which may take positive or negative values, and may range in magnitude from zero to a few hundred meters^{-1}. The term in A describes a magnetic, and the term in P a quadrupole, hyperfine structure.

The measurement of a hyperfine structure spectrum for a gaseous atomic or molecular system can lead to information about the values for A and P. These values may be interpreted to obtain information about the nuclear magnetic and quadrupole moments, and about the atomic or molecular electron configuration.

Important methods for the measurement of hyperfine structure for gaseous systems may employ an interferometer, or use atomic beams, electron spin resonance, or nuclear spin resonance. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; INTERFEROMETRY; MAGNETIC RESONANCE; MOLECULAR BEAMS; NUCLEAR MAGNETIC RESONANCE (NMR).

Liquid and solid systems. Hyperfine structure coupling may also occur and may be measured for liquid and solid systems. For liquids and solids, measurements are often made by electron spin or nuclear spin resonance methods. For solids, and for radioactive nuclei, one may, for example, also employ the Mössbauer effect or the angular correlation of nuclear gamma rays. *See* GAMMA RAYS; MÖSSBAUER EFFECT.

For a diamagnetic solid, $A = 0$ in the equation above, and if the crystalline environment of an atom is cubic, $P = 0$ also. If this environment is not cubic, P may have a finite measurable value. *See* DIAMAGNETISM.

If the solid is paramagnetic, ferromagnetic, or antiferromagnetic, A may be finite and measureable, and again P may or may not be zero depending on whether the atomic environment is cubic or not. *See* ANTIFERROMAGNETISM; FERROMAGNETISM; PARAMAGNETISM.

One may gain information about the nuclear moments and about electron bonding and magnetic structure from measurements of hyperfine structure for liquids and solids. Such measurements are extensively used, for example, in atomic and condensed matter physics, chemistry, and biology. *See* ATOMIC STRUCTURE AND SPECTRA; NUCLEAR MOMENTS.

Louis D. Roberts

Bibliography. A. Abragam, *The Principles of Nuclear Magnetism*, 1961, reprint 1983; D. M. Brink and G. R. Satchler, *Angular Momentum*, 3d ed., 1994; R. S. Raghavan and D. E. Murnick (eds.), *Hyperfine Interactions IV*, 1978.

Hypergeometric functions

The analytic continuation of the function defined by the series in Eq. (1), where the shifted factorial

$$ {}_2F_1(a, b; c; z) = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n n!} z^n \qquad |z| < 1 \tag{1} $$

$(a)_n$ is defined by Eq. (2). It satisfies Eq. (3), a lin-

$$ (a)_n = a(a+1)\cdots(a+n-1) \qquad n = 1, 2, \ldots, \qquad (a)_0 = 1 \tag{2} $$

$$ z(1-z)y'' + [c - (a+b+1)z]y' - aby = 0 \tag{3} $$

ear, homogeneous differential equation whose only singular points in the full complex plane are regular singular points at 0, 1, and ∞. For $|z| < 1$, Re $c >$ Re $b > 0$, it is given by the integral representation of Eq. (4), due to L. Euler, where Γ represents the

$$ {}_2F_1(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \cdot \int_0^1 (1-zt)^{-a} t^{b-1} (1-t)^{c-b-1} dt \tag{4} $$

gamma function. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES; GAMMA FUNCTION; SERIES.

The interest in hypergeometric functions comes from the many important functions which are special cases of the general hypergeometric function, the rich theory which has been developed for the general hypergeometric function, and the many times they occur in applications. Classically hypergeometric functions have arisen in science as solutions to differential equations. As discrete, rather than continuous, models of physical phenomena have become increasingly useful, hypergeometric functions have continued to arise as solutions to the equations governing these models. *See* DIFFERENTIAL EQUATION.

Elementary functions. Among the special cases are the elementary functions in Eqs. (5).

$$ \begin{aligned} \log(1-z) &= z\,{}_2F_1(1, 1; 2; z) \\ \sin^{-1} z &= z\,{}_2F_1\left({}^1/_2, {}^1/_2; {}^3/_2; z^2\right) \\ \tan^{-1} z &= z\,{}_2F_1\left({}^1/_2, 1; {}^3/_2; -z^2\right) \\ (1-z)^{-a} &= {}_2F_1(a, b; b; z) \end{aligned} \tag{5} $$

Continued fraction expansions. An example of a result which is useful, far from obvious, and yet is a simple consequence of the general theory of hypergeometric functions is the continued fraction expansion of $\tan^{-1} z$ in Eq. (6). Since $\tan^{-1} 1 = \pi/4$, the continued fraction in Eq. (6) gives an explicit expression for π.

$$ \tan^{-1} z = \cfrac{z}{1 + \cfrac{\frac{1.1}{1.3} z^2}{1 + \cfrac{\frac{2.2}{3.5} z^2}{1 + \cfrac{\frac{3.3}{5.7} z^2}{1 + \cdots}}}} \tag{6} $$

See CIRCLE.

C. F. Gauss found a continued fraction for the ratio ${}_2F_1(a, b+1; c+1; z)/{}_2F_1(a, b; c; z)$, which reduces to Eq. (6) in the special case $a = {}^1/_2$, $b = 0$, $c = {}^1/_2$. To obtain this continued fraction, Gauss derived three-term recurrence relations which connect three hypergeometric functions, two of which are contiguous if the three parameters, "a," "b," and "c," are equal and the third differs by one. These contiguous relations can be thought of as three-term difference equations which are a discrete analog of differential equation (3). Most of the occurrences of hypergeometric functions arise because they satisfy a second-order differential or difference equation. *See* INTERPOLATION.

Transformations. A simple change of variables ($t = 1 - s$) in the integral of Eq. (4) gives linear (fractional) transformation (7). Iterating this trans-

$$ {}_2F_1(a, b; c; z) = (1-z)^{-a}\,{}_2F_1[a, c-b; c; z/(z-1)] \tag{7} $$

formation, after using the symmetry in "a" and "b," gives Eq. (8).

$$ {}_2F_1(a, b; c; z) = (1-z)^{c-a-b}\,{}_2F_1(c-a, c-b; c; z) \tag{8} $$

See CONFORMAL MAPPING.

There is a very important subclass of hypergeometric functions, depending on two, rather than three, parameters, which has a "quadratic transformation." The two basic ones are Eqs. (9). This class is impor-

$$ \begin{aligned} &{}_2F_1\left(2a, 2b; a+b+{}^1/_2; z\right) \\ &\quad = {}_2F_1\left[a, b; a+b+{}^1/_2; 4z(1-z)\right] \\ &{}_2F_1(2a, a; 2a; z) \\ &\quad = (1-z)^{-b}\,{}_2F_1\left[b, a-b; a+{}^1/_2; z^2/(4z-4)\right] \end{aligned} \tag{9} $$

tant because each function in it can be multiplied by an algebraic function to give a Legendre function, and all Legendre functions arise in this way. *See* LEGENDRE FUNCTIONS.

Confluent hypergeometric functions. Differential equation (3) has regular singular points at the points 0, 1, and ∞. These can be moved to arbitrary points z_1, z_2, z_3 by a linear fractional transformation, and the resulting equation exhibits the symmetries given in linear and quadratic transformations (7)–(9), in a more transparent fashion. Also the resulting singular

Orthogonal polynomials expressed as generalized hypergeometric functions

Polynomial	Hypergeometric representation	Distribution	
Jacobi $P_n^{(\alpha,\beta)}(x)$	$\frac{(\alpha+1)_n}{n!}\ {}_2F_1[-n,\ n+\alpha+\beta+1;\ \alpha+1;\ (1-x)/2]$	$(1-x)^\alpha(1+x)^\beta$	$-1<x<1$
Laguerre $L_n^\alpha(x)$	$\frac{(\alpha+1)_n}{n!}\ {}_1F_1(-n;\ \alpha+1;\ x)$	$x^\alpha e^{-x}$	$x>0$
Hermite $H_n(x)$	$(2x)^n{}_2F_0\left(-\frac{n}{2},\ \frac{1-n}{2};\ -;\ -\frac{1}{x^2}\right)$	e^{-x^2}	$-\infty<x<\infty$
Hahn $Q_n(x;\ \alpha,\ \beta,\ N)$	${}_3F_2\left(\begin{matrix}-x,\ -n,\ n+\alpha+\beta+1\\ -N,\ \alpha+1\end{matrix};\ 1\right)$	$\frac{(\alpha+1)_x}{x!}\ \frac{(\beta+1)_{N-x}}{(N-x)!}$	$x=0,1,\ldots,N$
Meixner $M_n(x;\ \beta,\ c)$	${}_2F_1(-n,\ -x;\ \beta;\ 1-c^{-1})$	$\frac{(\beta)_x c^x}{x!}$	$x=0,1,\ldots$
Krawtchouk $K_n(x;\ p,\ N)$	${}_2F_1(-n,\ -x;\ -N;\ p^{-1})$	$\binom{N}{x}p^x(1-p)^{N-x}$	$x=0,1,\ldots,N$
Charlier $C_n(x;\ a)$	${}_2F_0(-n;\ -x;\ -;\ -a^{-1})$	$a^x/x!$	$x=0,1,\ldots$

points can be made to coalesce. For the ordinary hypergeometric function this procedure is called confluence. The function ${}_2F_1(a,b;c;z/b)$ satisfies a differential equation whose regular singular points are at 0, b, and ∞, and if b is allowed to become large, the resulting function is called a confluent hypergeometric function. Explicitly, it is given by Eq. (10). Transformation formula (8) becomes Eq. (11). Quadratic transformation (9) implies Eq. (12). The function on

$$ {}_1F_1(a;c;z)=\sum_{n=0}^{\infty}\frac{(a)_n}{(c)_n}\frac{z^n}{n!} \tag{10}$$

$$ {}_1F_1(a;c;z)=e^z{}_1F_1(c-a;c;-z) \tag{11}$$

$$ e^{-z}{}_1F_1(a;2a;2z)={}_0F_1\left(-;a+{}^1/_2;z^2/4\right) \tag{12}$$

the right-hand side is a simple multiple of a Bessel function of imaginary argument, or $I_{a-(1/2)}(z)$. Using this limit, differential equation (3) becomes Eq. (13).

$$ zy''+(c-z)y'-ay=0 \tag{13}$$

It has a regular singular point at 0 and an irregular singular point at infinity, and one of its solutions is ${}_1F_1(a;c;z)$. *See* BESSEL FUNCTIONS.

Generalized hypergeometric functions. Since a number of sums, such as ${}_2F_1$, ${}_1F_1$, and ${}_0F_1$, arise in a very natural way, it is useful to consider the more general hypergeometric function defined by Eq. (14), where

$$ {}_pF_q(a_p;b_q;z)=\sum_{n=0}^{\infty}\frac{(a_1)_n\cdots(a_p)_n z^n}{(b_1)_n\cdots(b_q)_n n!} \tag{14}$$

$a_p=a_1,\ \ldots,a_p;\ b_q=b_1,\ \ldots,b_q$. All of these generalized hypergeometric functions satisfy differential equations of order max $(p,q+1)$, where if $p>q+1$ one of the a's is assumed to be a negative integer, since the series does not converge except when $z=0$ without this assumption.

Generalized hypergeometric functions in one variable can also be thought of as coming from the series (15), where the term ratio a_{n+1}/a_n is a rational

$$ \sum_{n=0}^{\infty}a_n \tag{15}$$

function of n. Since the 1980s, significant progress has been made in studying many variable hypergeometric functions through integral representations and differential equations, including the discovery of new beta-function-type integrals.

Orthogonal polynomials. These functions also satisfy difference equations in the parameters a_i and b_j. Some of these difference equations give rise to polynomials which are orthogonal with respect to many of the important distribution functions in statistics (see **table**).

These polynominals are useful in studying statistical and probabilistic problems. Meixner and Laguerre polynomials play an essential role in birth and death processes when the birth and death parameters are linear functions of the size of the population. Jacobi polynomials for certain special values of the parameters α and β are the zonal spherical harmonics on spheres and projective spaces. For the case $\alpha=\beta=0$, the polynomials reduce to Legendre polynomials and the sphere is the unit sphere in three dimensions. This point of view leads to some of the deepest formulas known about hypergeometric functions, especially those known as addition formulas. For example, Jacobi polynomials satisfy formula (16) where notation (17) applies,and $c(m,k,\alpha,\ \beta,n)$

$$ P_n^{(\alpha,\beta)}\left[\frac{(1+x)(1+y)}{2}+\frac{r^2(1-x)(1-y)}{2}+(1-x^2)^{1/2}(1-y^2)^{1/2}r\cos\phi-1\right]$$
$$=\sum_{k=0}^{n}\sum_{m=0}^{k}c(m,k,\alpha,\beta,n)f_{n,m,k}(x)f_{n,m,k}(y)\cdot P_m^{(\alpha-\beta-1,\beta+k-m)}(2r^2-1)r^{k-m}P_{k-m}^{(\beta-1/2,\beta-1/2)}(\cos\phi) \tag{16}$$

$$ f_{n,m,k}(x)=(1-x)^{(k+m)/2}(1+x)^{(k-m)/2}\cdot P_{n-k}^{(\alpha+k+m,\beta+k-m)}(x) \tag{17}$$

is a product of shifted factorials. *See* PROBABILITY; SPHERICAL HARMONICS; STATISTICS.

Krawtchouk polynomials when $p={}^1/_2$ are the zonal spherical harmonics on the space consisting of the vertices of the unit cube in N-space and there is a

corresponding addition formula. The vertices of the unit cube can be considered as a message of zeros and ones, and Krawtchouk polynomials have been shown to play an important role in coding theory. *See* INFORMATION THEORY; ORTHOGONAL POLYNOMIALS.

Special values of the argument. A number of hypergeometric functions can be evaluated as quotients of gamma functions when the argument z takes on special values. Gauss' sum, Eq. (18), occurs often

$$ {}_2F_1(a, b; c; 1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} \qquad (18) $$
$$ c > a + b $$

in applications. For the generalized hypergeometric series, conditions have to be placed on the parameters before the series can be explicitly summed. Two general classes of sums often arise. One, called well-poised, occurs when $p = q + 1$ and the parameters can be paired so that $a_1 + 1 = a_2 + b_1 = \cdots = a_{q+1} + b_q$. The general well-poised ${}_3F_2$ has been summed at $z = 1$. A series is called very well-poised if it is well poised and $a_2 = b_1 + 1$. In the second type, called k-balanced, $p = q + 1$, $a_1 + \cdots + a_{q+1} + k = b_1 + \cdots + b_q$ for some integer k, and one of the a_i's is a negative integer The l-balanced ${}_3F_2$ has been summed at $z = 1$. The most complicated sum of this type which has been discovered is the sum of the very well-poised, 2-balanced ${}_7F_6$ at $z = 1$.

These sums, and related transformation formulas, are fundamental and occur in many applications. Among these are complex spectra, symmetries of the Clebsch-Gordan coefficients, and the fluctuation theory of random walks. *See* ANGULAR MOMENTUM; STOCHASTIC PROCESS.

Generalizations. There are many generalizations of hypergeometric functions which are also very useful.

Basic hypergeometric functions.. The study of these functions started with Euler in the 1740s. From a power-series point of view, these are series (15) with the term ratio a_{n+1}/a_n a rational function of q^n, where q is a fixed parameter. This is equivalent to replacing the shifted factorial in Eq. (1) or Eq. (14), by the quantities in Eq. (19), and introducing appropriate

$$ (q^a)_{q,n} = (1 - q^a)(1 - q^{a+1}) \cdots (1 - q^{a+n-1}) \qquad (19) $$

powers of q as multiplicative factors. The resulting series are connected with theta functions, and with a pair of formulas which are important in number theory and combinatorial analysis. One of these is Eq. (20)

$$ \sum_{n=0}^{\infty} \frac{qn^2}{(q)_{q,n}} = \frac{1}{\prod_{n=0}^{\infty}(1 - q^{5n+1})(1 - q^{5n+4})} \qquad (20) $$

where $$ \prod_{n=0}^{\infty} a_n = a_0 a_1 \cdots a_n \cdots $$

is an infinite product. In terms of partitions of positive integers as the sum of positive integers, this says that the number of ways of writing a positive integer as the sum of positive integers so that the difference between any two of the parts is at least two is equinumerous with the number of ways of writing the same integer as the sum of positive integers so that each of the parts is either one more or one less than an integer divisible by five. An infinite number of sums similar to Eq. (19) have been found, but the sums are now multiple sums. *See* COMBINATORIAL THEORY; ELLIPTIC FUNCTION AND INTEGRAL.

The coefficients $[(q)_{q,n}]/[(q)_{q,k}(q)_{q,n-k}]$ arise when counting the subspaces of dimension k of an n-dimensional vector space over a field of q elements. This property was used to obtain some identities for q-hypergeometric functions. The results were previously known, but these observations opened up new fields for applications of this generalization of hypergeometric functions.

Since the 1970s there have been notable developments in the study of basic hypergeometric functions. These functions have arisen in the study of the hard hexagon model in statistical mechanics. They are used to study and develop codes, in a discrete setting corresponding to the occurrence of Legendre functions on the surface of a sphere. They also occur in the study of quantum groups. These are Hopf algebras rather than groups, but some examples have enough structure so that detailed calculations can be done in the same way as for Lie groups. *See* LIE GROUP.

One example of the type of formula which occurs is a noncommutative extension of the binomial theorem. The binomial theorem is Eq. (21), where the

$$ (x + y)^n = \sum_{k=0}^{n} = \binom{n}{k} x^{n-k} y^k \qquad (21) $$

binomial coefficient is given by Eq. (22). A noncom-

$$ \binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad (22) $$

mutative extension of Eq. (21) assumes that $yx = qxy$, $qx = xq$, and $qy = yq$. Then the expansion (21) becomes Eq. (23), with the q-binomial coefficient a

$$ (x + y)^n = \sum_{k=0}^{n} \binom{n}{k}_q x^{n-k} y^k \qquad (23) $$

polynomial in q with integer coefficients. These coefficients are given by Eq. (24).

$$ \binom{n}{k}_q = \frac{(q)_{q,n}}{(q)_{q,k}(q)_{q,n-k}} \qquad (24) $$

The coefficients in the polynomial

$$ \binom{n}{k}_q $$

count the number of inversions needed to take a given pattern to a fixed one. For example, there are 6 ways to arrange two red and two black balls in four spots. These are rrbb, rbrb, rbbr, brrb, brbr, and bbrr. If the first is taken as the fixed pattern, then there are 2 inversions needed to take each of brrb and brbr

to rrbb. The others require 1, 3, and 4 inversions respectively. Then Eq. (25) is valid.

$$\binom{4}{2}_q = \frac{(1-q^4)(1-q^3)}{(1-q)(1-q^2)}$$
$$= 1 + q + 2q^2 + q^3 + q^4 \qquad (25)$$

Functions of multiple variables and matrices. In 1880 P. Appell introduced four hypergeometric functions of two variables. These functions arise very naturally from a group theoretic point of view, and thus have possible applications in the study of elementary particles by methods employing group theory. *See* ELEMENTARY PARTICLE; GROUP THEORY.

A final generalization is to hypergeometric functions of matrices. These functions have arisen in statistical multivariate analysis and in number theory. *See* MATRIX THEORY; NUMBER THEORY.

Richard Askey

Bibliography. G. Andrews et al. (eds.), *Ramanujan Revisited*, 1988; B. Dwork, *Generalized Hypergeometric Functions*, 1990; A. Erdélyi et al., *Higher Transcendental Functions*, 3 vols., 1953–1955, reprint 1981; G. Gasper and M. Rahman, *Basic Hypergeometric Series*, 1990; J. B. Seaborn, *Hypergeometric Functions and Their Applications*, 1991; N. Ja. Vilenkin, *Special Functions and the Theory of Group Representations*, 1968, reprint 1988.

Hyperiidea

A suborder of amphipod crustaceans. Most hyperiids can be recognized by the large eyes which cover nearly the entire surface of the head. The first maxillae and especially the maxillipeds are greatly reduced in comparison to the suborder Gammaridea. In the prehensile pereiopods, the claw is formed by the fifth and sixth segments, the carpus and propodus, rather than by the sixth and seventh segments or the propodus and dactyl, as in the Gammaridea. The second and third somites of the urosome are always fused, a condition rarely found in the Gammaridea. *See* AMPHIPODA.

Biology. The Hyperiidea are exclusively pelagic and marine. They are found in all the oceans, from the surface to great depths. Most are characteristic of oceanic rather than neritic waters, although some species frequent inshore waters in the tropics. After the Copepoda and Euphausiacea, the Hyperiidea are the most abundant planktonic crustaceans.

Little is known about the habits of the hyperiids. Some species are frequently found in association with other animals. Members of the genus *Vibilia* deposit their larvae in salps and feed on the salp until they are able to swim freely. Some species of *Hyperia* and *Hyperoche* are commonly found on Scyphomedusae, clinging to the surface of the bell. The young of *Hyperia galba* have been found in the subgenital pits. *Eupronoe* also sometimes occurs with medusae.

The adult females of *Phronima sedentaria* are found in deep water in gelatinous barrel-shaped cases open at both ends. These cases are fashioned from the tests of pyrosomes. The young hatch in the cases, and they pass through several molts before leaving them to move nearer to the surface.

Other species are truly free-living. Except in tropical and subtropical seas, species of *Parathemisto*, formerly *Themisto*, are the most abundant hyperiids. *Parathemisto libellula* occurs in large numbers in Arctic waters and is an important food for ringed seals. Japanese sardines feed on *Parathemisto japonica. Phronisa semilunata*, *Phromina sedentaria*, *Streetsia chellengeri*, and *Brachyscelus crusculum* are eaten by albacore.

Some hyperiids are luminescent, like *Scina* and *Parapronoe crustulum*. The whole body glows with a greenish-yellow light. *See* BIOLUMINESCENCE.

Taxonomy. The suborder Hyperiidea is divided into two superfamilies, Physosomata and Genuina. Members of the Physosomata, except for a number of species of *Scina* and one species of *Lanceola*, are bathypelagic. The eyes are small or rarely absent, and the inner plates of the maxillipeds are free at the apex. The Genuina include the more familiar hyperiids with large eyes and completely fused maxillipedal inner plates, and are divided into three groups according to the structure of the male antennae. In the Recticornia, containing the genera *Vibilia*, *Paraphronima*, and *Cystisoma*, the first antennae are straight. They arise from the anterior margin of the head, and have few flagellar segments. The first antennae of the Filicornia, in such genera as *Hyperia*, *Dairella*, *Primno*, and *Phronima*, are also inserted anteriorly, and have many-segmented flagella. The antennae of the Curvicornia originate from the ventral margin of the head. This may be observed in *Eupronoe*, *Lycaea*, *Oxycephalus*, and *Platyscelus*, in which the first segment is large and curved, with the remaining segments few in number. The flagellar segments of the second antennae are folded on themselves. The Curvicornia comprise globular forms with greatly enlarged basal segments on the fifth and sixth legs as well as slender, elongate species, culminating in the rod-shaped *Rhabdosoma*. *See* CRUSTACEA.

Thomas E. Bowman

Hypermastigida

An order of the Protozoa in the class Zoomastigophorea comprising the most complex flagellates, both structurally and in modes of division. All inhabit the alimentary canal of termites, cockroaches, and woodroaches. These organisms are multiflagellate, often with complicated blepharoplast-parabasal-axostylar structures. The nucleus is single and the organisms are plastic and slow-moving, generally ovoid to elongate. Flagella occur in spiral rows, in tufts, or over the entire body. These flagellates vary from 15 to 350 micrometers in size. They may be holozoic or saprozoic. Ingestion is usually pseudopodial, in the posterior region. A high degree of adaptiveness exists, and species peculiar to one host species are not viable in another. Sexual processes

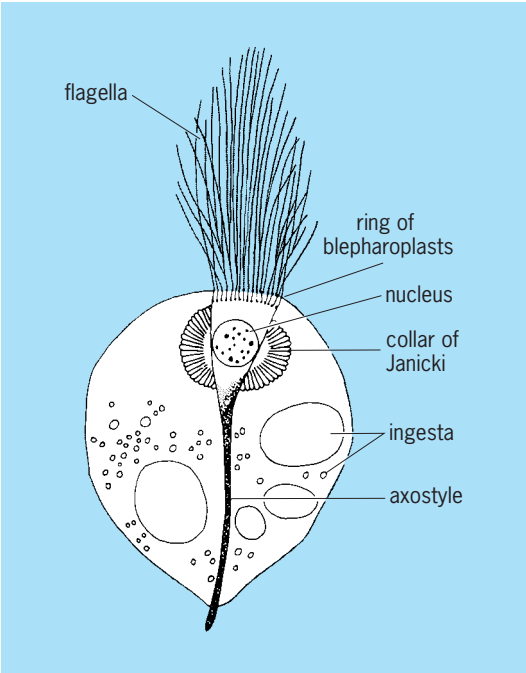

***Lophomonas blattarum*, zylophagous (wood-eating parasite) symbiont of cockroach**

and multiple fission are known, but trinary fission is most common. *Holomastigotoides* has spirochetes attached to its trophic body zone.

Lophomonas inhabits the cockroach intestine and is apparently cosmopolitan. The body is round to pyriform with a tuft of apical flagella arising from a blepharoplast ring which forms a calyx enclosing a nucleus (see **illus.**). Six to eight chromosomes form on a spindle of which the poles are centrioles. Laboratory culture is easy. *See* CILIA AND FLAGELLA; PROTOZOA; SARCOMASTIGOPHORA; ZOOMASTIGOPHOREA. James B. Lackey

Hypernuclei

Nuclei that consist of protons, neutrons, and one or more strange particles such as lambda particles. The lambda particle is the lightest strange baryon (hyperon); its lifetime is 2.6×10^{-10} s. Because strangeness is conserved in strong interactions, the lifetime of the lambda particle remains essentially unchanged in the nucleus also. Lambda hypernuclei live long enough to permit detailed study of their properties. *See* BARYON; ELEMENTARY PARTICLE; HYPERON; STRONG NUCLEAR INTERACTIONS.

There is no bound lambda-nucleon system, demonstrating that the lambda-nucleon force is weaker than the force between nucleons which can bind two nucleons into deuterium. The lightest bound Λ hypernucleus is ${}^{3}_{\Lambda}\mathrm{H}$—lambda hypertriton—which is composed of a proton, a neutron, and a lambda particle. Lambda hypernuclei up to ${}^{16}_{\Lambda}\mathrm{O}$ have been identified experimentally.

Two cases of the double lambda hypernuclei have been found so far; one is ${}_{\Lambda\Lambda}^{6}\mathrm{He}$, which is composed of two lambda particles coupled to the ${}^{4}\mathrm{He}$ nucleus. The bound double-lambda system is still being sought. In some theoretical models of the elementary particles a strong binding for the two lambda particles is predicted.

The lambda particle in the nucleus experiences an attractive potential, the strength of which is about two-thirds that for the nucleon. The spin-orbit force, which is strong in the case of the nucleon as it causes splitting comparable to the energy differences between the nucleon shells, is negligibly small in the case of the lambda particle. Theoretically, the difference between the lambda-nucleus and nucleon-nucleus interaction is explained as reflecting the differences in the internal quark structure of the two baryons.

Sigma, xi, and omega particles, all with a lifetime of about 10^{-10} s as free particles, convert, through the strong interaction, into lambda particles in the nucleus. Nevertheless, sigma hypernuclei have been experimentally observed, the lifetime being of the order of 10^{-21} s. This lifetime is long enough to permit determination of the most important parameters of the sigma-nucleus interaction. There is a good chance that xi and omega hypernuclei could also be investigated experimentally. *See* NUCLEAR STRUCTURE. Bogdan Povh

Bibliography. D. Wilkinson (ed.), *Progress in Particle and Nuclear Physics*, vol. 5, 1981.

Hyperon

A collective name for any baryon with nonzero strangeness number s. The name hyperon has generally been limited to particles which are semistable, that is, which have long lifetimes relative to 10^{-22} s and which decay by photon emission or through weaker decay interactions. Hyperonic particles which are unstable (that is, with lifetimes shorter than 10^{-22} s) are referred to as excited hyperons. The known hyperons with spin $\frac{1}{2}\hbar$ (where $\hbar$ is Planck's constant divided by 2π) are Λ, Σ^-, Σ^0, and Σ^+, with $s = -1$, and Ξ^- and Ξ^0, with $s = -2$, together with the Ω^- particle, which has spin $\frac{3}{2}\hbar$ and $s = -3$. The corresponding antihyperons have baryon number $B = -1$, and opposite values of strangeness s and charge Q; they are all known empirically.

The first excited hyperon was reported in 1960. The state $\Sigma(1385)^+$ was observed as a $\pi\Lambda$ resonance in the final state of reaction (1). The symbol $\Sigma(m)$ or

$$K^- + p \rightarrow \pi^- + \Sigma(1385)^+ \rightarrow \pi^- + \pi^+ + \Lambda \qquad (1)$$

$\Lambda(m)$ indicates that the strangeness is $s = -1$ and that the isospin is $I = 1$ or 0, respectively. The superscript gives the charge and m is the mass in MeV; if no m is given, the symbol refers to the ground state, for example, Λ means $\Lambda(1115.5)$. *See* I-SPIN.

Reaction (1) is an example of an excited hyperon production reaction. Formation reactions are also possible for most excited hyperons with $s = -1$. For example, the properties of $\Lambda(1520)$ are particularly

well known from observations on its formation and decay, for K^- mesons incident on hydrogen, reaction (2). $\Sigma(1385)$ cannot be formed in this way, because

$$K^-p \rightarrow \Lambda(1520) \rightarrow \begin{cases} K^- + p \\ \pi^\pm + \Sigma^\pm \\ \pi^+ + \pi^- + \Lambda, \text{etc.} \end{cases} \quad (2)$$

its mass lies below the K^-p threshold energy, $m_K + m_p \simeq 1432$ MeV.

The name hyperon may be extended to include any baryon for which one or more of the quantum members s, c, and b are nonzero, where c = charm and b = bottom. For example, the hyperon Λ_c^+,which has the quark structure $c(ud - du)/\sqrt{2}$, has zero strangeness but nonzero charm. The heaviest $c = 1$ hyperon in Ω_c^0, which has structure ssc and mass 2719 ± 8 MeV; it has spin-parity $1/2^+$ and is semistable, with decay mode (3). The heaviest $b =$

$$\Omega_c^0 \rightarrow \Xi^- K^- \pi^+ \pi^+ \quad (3)$$

1 hyperon is Λ_b^0, with structure $b(ud - du)/\sqrt{2}$ and mass 5640 ± 50 MeV; it has spin-parity $1/2^+$, like Λ and Λ_c, and is semistable, with decay mode $\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-$.

There is no deep distinction between hyperons and excited hyperons, beyond the phenomenological definition above. Indeed, the hyperon $\Omega(1672)^-$ and the excited hyperons $\Xi(1530)$ and $\Sigma(1385)$, together with the unstable nucleonic states $\Delta(1236)$ are known to form a unitary decuplet of states with spin $3/2\hbar$. *See* BARYON; ELEMENTARY PARTICLE; QUARKS; SYMMETRY LAWS (PHYSICS); UNITARY SYMMETRY.

Richard H. Dalitz

Hypersonic flight

Flight at speeds well above the local velocity of sound. In the span of a few years speeds of aircraft, missiles, and spacecraft increased tremendously, from far less than the speed of sound (subsonic) to speeds required for orbiting or escape from the Earth. By convention, hypersonic flight starts at about Mach 5 (five times the speed of sound) and extends upward in speed indefinitely. A more precise definition, attributed to T. von Kármán, states that hypersonic flight starts when the cross-flow Mach number (or Mach-number component perpendicular to the longitudinal axis) becomes 1 or greater. In any case, hypersonic flight refers to atmospheric flight at hypersonic speed when the characteristics of the flow field about the body are sensibly unchanged as the speed of the body increases further.

Conditions at high speed. Below about 350,000 ft (107 km) altitude, air is sufficiently dense to be considered a continuum and to form viscous laminar and turbulent boundary layers on body surfaces. Below this altitude, the speed of flight divided by the speed of sound—that is, the Mach number—is a characteristic aerodynamic parameter. In the realm of superaerodynamics, air density above about 350,000 ft (107 km) is so low that the mean free path between molecular collisions is large compared to body dimensions, and a new parameter known as the molecular speed ratio (ratio of flight speed to most probable speed of an incident molecule) is used.

Subsonic, transonic, and supersonic vehicles fly at speeds less than, equal to, and greater than the local speed of sound, respectively. However, when the Mach number is high, the flow field around an object exhibits a special behavior, which is worth studying separately from supersonic flight. This behavior is characteristic of hypersonic flight. For long slender bodies and for thin airfoils at small angles of attack, transition from supersonic to hypersonic flight may require a Mach number approximately equal to or greater than 5 (a flight speed of approximately 5000 ft/s or 1500 m/s). On Earth the fastest-moving natural objects seen by observers are meteorites, which travel at 25,000–250,000 ft/s (7500–75,000 m/s), or from Mach 25 to 250.

A body entering the Earth's atmosphere from space (for example, a meteorite, a ballistic reentry vehicle, or a spacecraft) has high velocity and hence large kinetic and potential energy. The kinetic energy KE is shown in Eq. (1) and potential energy PE in Eq. (2), where W is the weight, V is velocity, g is

$$\text{KE} = \frac{WV^2}{2g} \quad (1)$$

$$\text{PE} = WZ \quad (2)$$

acceleration of gravity, and Z is altitude. At hypersonic speeds the potential energy of a body is small compared to its kinetic energy. For instance, the potential energy of a body at an altitude of 200,000 ft (60 km) about equals the kinetic energy of the body if it is moving at 3600 ft/s (1080 m/s).

During reentry, drag forces act upon a reentry vehicle or spacecraft and cause it to decelerate, thus dissipating its kinetic and potential energy. The energy lost by the reentry vehicle is then transferred to the air within the flow field around the reentry vehicle. The flow field around the forward portion of a blunt-nosed vehicle (body of revolution or leading edge of a wing) generally exhibits (1) a distinct bow shock wave, (2) a shock layer of highly compressed hot gas, and (3) a highly sheared boundary layer over the surface (**Fig. 1**). The flow field is defined as the region of disturbed air between the body surface and the shock wave. The temperature of the shock layer is so high that molecules begin to dissociate during collisions at about 2500 K, or at Mach 7. At about Mach 12, gas in the stagnation region reaches a temperature of 4000 K and becomes ionized.

In general, at high hypersonic flight speed the characteristic temperature in the shock layer of a blunted body and in the boundary layer of a slender body is proportional to the square of the Mach number. (Thus the working fluid in high hypersonic flight cannot be considered to behave as a perfect gas.) Heat energy is then transferred from the hot gases to the vehicle surface by conduction and diffusion of chemical species in the boundary layer and by radiation from the shock layer near the nose. Heat

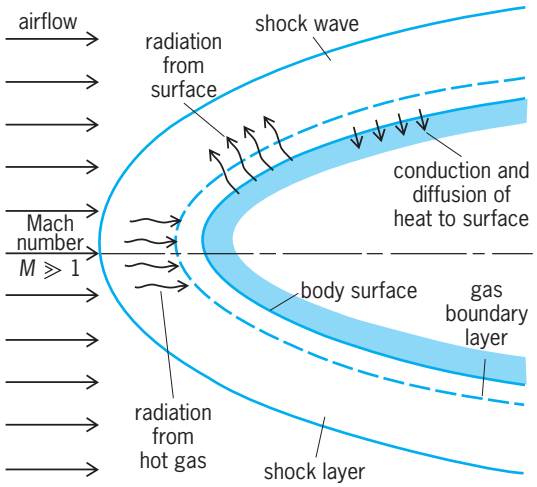

Fig. 1. Heat transfer in stagnation region of a blunted nose cap of a reentry vehicle or spacecraft.

energy is also radiated from the vehicle surface to space or to adjacent objects. One important problem confronting the designer of reentry vehicles or spacecraft is therefore to design a minimum-weight vehicle able to withstand large heat loads from adjacent hot-gas layers during reentry while retaining the ability to carry a given useful payload. *See* ATMOSPHERIC ENTRY; SPACECRAFT STRUCTURE.

Vehicular trajectories. A hypersonic vehicle in sustained powered flight maintains a constant speed (**Fig. 2**). It needs only to be initially accelerated to its operating velocity. An air-breathing engine or a rocket motor may provide the thrust to overcome aerodynamic drag, but the propulsive work done to sustain the flight speed must be dissipated as heat to the surrounding air or radiated to space. A hypersonic vehicle is in accelerating flight when the propulsive thrust is used to accelerate the vehicle as well as to overcome atmospheric drag.

There are three basic trajectories for the unpowered portion of the flight of a reentry vehicle or spacecraft reentering the Earth's atmosphere: ballistic, skip, and glide.

Ballistic trajectory. A trajectory such as that of an IRBM or ICBM is actually a segment of one of Kepler's planetary ellipses (Fig. 2). The range of the ballistic missile is shorter than that of other reentry vehicles for given initial conditions.

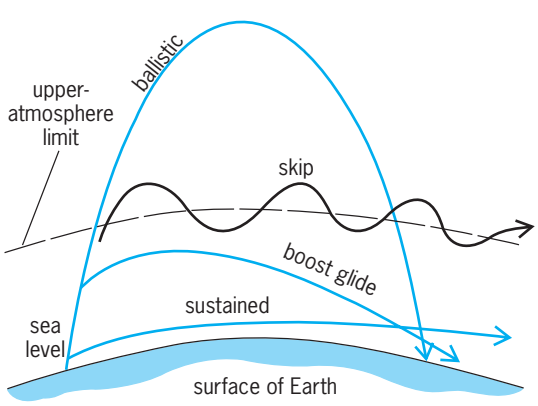

Fig. 2. Typical trajectories for hypersonic flight.

The short transit time of the ballistic missile in the Earth's atmosphere causes kinetic energy to be dissipated at a high rate as the reentry vehicle experiences rapid deceleration due to atmospheric drag. The effect of peak aerodynamic heating on the vehicle can be reduced if a blunt shape is employed to increase the vehicle's drag. This high-drag shape causes the vehicle to expend a large portion of its kinetic energy in pressure drag rather than in frictional forces. The effect of aerodynamic heating on the vehicle can be further reduced by the use of ablative, transpiration, shock generation and by magnetohydrodynamic cooling techniques. *See* BALLISTIC MISSILE; BALLISTICS.

Skip trajectory. If the trajectory is made up of ballistic phases alternating with skipping phases, the flight path resembles that of a stone skipping over a pond (Fig. 2). In the ballistic phase the vehicle experiences only gravitational and inertial forces; accordingly, each phase is simply a segment of one of Kepler's ellipses. During the skipping phase the vehicle experiences large aerodynamic forces, from which it develops lift. The forces tend to be so large when the reentry angle is steep that, for most practical purposes, it is permissible to neglect the gravitational force and to treat such a skipping phase as an impact. The heating effect is also more severe during the skipping process, even if ablation cooling is employed. Because of these effects a skip vehicle on a maximum-range trajectory must have a strong structure. The skip vehicle with a hypersonic lift-to-drag ratio between 1 and 4 appears to be more efficient than both the ballistic and glide types in converting high velocity into long range.

Glide trajectory. In a glide trajectory the vehicle descends and decelerates slowly through the atmosphere in such a manner that the aerodynamic lift plus inertial forces acting on the vehicle just counterbalance the weight and drag in the glide phase. For hypersonic lift-to-drag ratios that are greater than about 1.5, the range of the glide trajectory exceeds the maximum ballistic range for the same payload-to-takeoff mass ratio—that is, for the same velocity at burnout. For lift-to-drag ratios in the neighborhood of 4 and greater, the glide vehicle is comparable to the skip vehicle in its ability to convert velocity into range.

To obtain a relatively high lift-to-drag ratio, a slender shape is necessary. This low-drag configuration will in turn increase the total convective heat transfer to the vehicle because of the long flight time, but the effect is not so severe as for the skip vehicle. For high hypersonic lift-to-drag vehicles, the combination of radiative cooling, structural heat-sink techniques, and glide-trajectory control (gradient descent and deceleration) appears able to reduce skin temperatures to a level feasible for present materials. For glide vehicles with a hypersonic lift-to-drag ratio of about 2 or lower, ablation cooling is more atractive

as a thermal protection technique on surface areas with high heating rates. For flight ranges on the order of the Earth's radius and greater, the performance of the hypersonic vehicle compares favorable with that of the supersonic airplane. *See* AERODYNAMICS; SUPERCRITICAL WING; TRAJECTORY. Shih-Yuan Chen

Bibliography. J. D. Anderson, *Hypersonic and High Temperature Gas Dynamics*, 2d ed., 2006; J. J. Bertin, *Hypersonic Aerothermodynamics*, 1994; W. H. Heiser and D. T. Pratt, *Hypersonic Airbreathing Propulsion*, 1994; C. Park, *Nonequilibrium Hypersonic Aerothermodynamics*, 1990.

Hypertension

High blood pressure. The prevalence of this clinical disease in acculturated societies is sufficiently high to warrant its being referred to as a public health problem. Blood pressure is expressed in two numbers: the higher number is the systolic blood pressure, which is the pressure exerted by the blood against the walls of the blood vessels while the heart is contracting. The lower number is the diastolic blood pressure, which is the residual pressure that exists between heart contractions, or while the heart is relaxing. Normal blood pressure provides sufficient blood flow to the vital organs, including the brain, heart, kidneys, intestine, and skeletal muscle.

It is not entirely accurate to think of high blood pressure as a distinct disease. Epidemiological and actuarial data suggest that blood pressure is distributed unimodally within populations and is directly related to mortality (see **illus.**). Simply stated, high blood pressure appears to be both a disease and a risk factor for other diseases. At the highest end of the blood pressure distribution, there is an increased probability of premature death secondary to stroke, heart disease, or kidney failure. Lower on the distribution curve (for example, diastolic blood pressure of 90–104 mmHg, which is referred to as mild hypertension), the absolute risk of premature mortality is lower and continues to decline with further decreases in blood pressure. High blood pressure is thus a disease when its value is very high and a risk factor throughout its distribution. For diagnostic purposes, blood pressure is considered high when persistently above 140/90 mmHg.

Some cases of very high blood pressure are due to specific causes that may be surgically remediable. Most hypertension, however, results from the combination of a genetic predisposition and an environmental factor such as excessive sodium intake, sedentary habits, and stress. *See* ARTERIOSCLEROSIS; STRESS (PSYCHOLOGY).

High blood pressure can be controlled. Mild cases are treated by losing excess weight and reducing the intake of sodium and alcohol. More serious cases are treated with drugs such as diuretics, beta blockers, calcium antagonists, angiotensin-converting enzyme inhibitors, alpha blockers, and centrally acting compounds that affect regulatory centers in the brain.

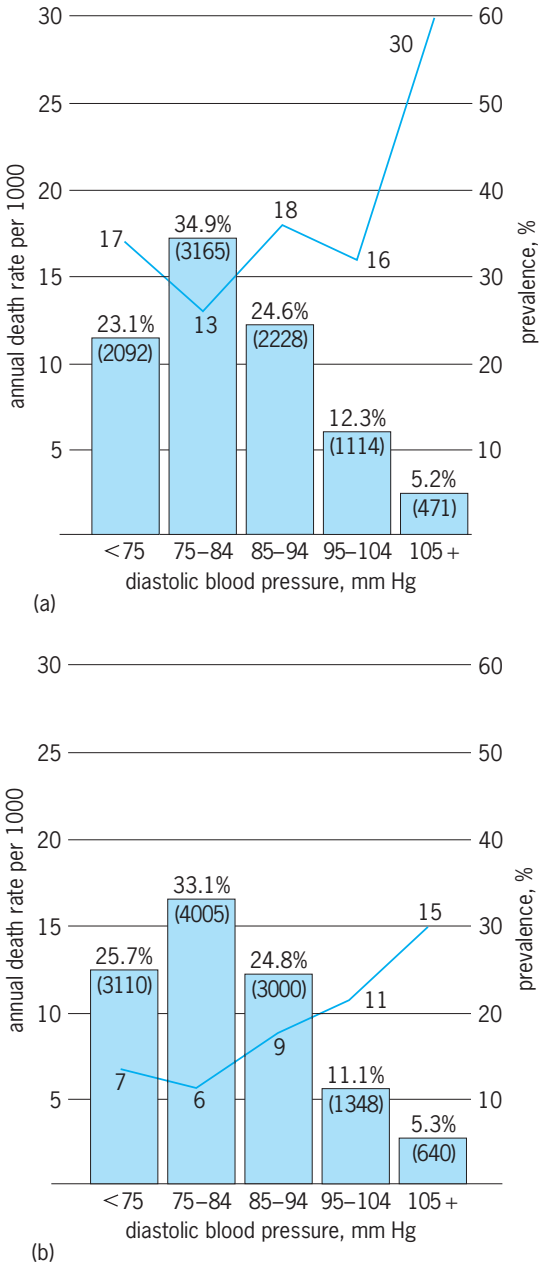

Mortality curves derived from the Framingham Heart Study show that in (*a*) men and (*b*) women aged 55–64 years, there is a correlation between diastolic blood pressure and mortality. The number of subjects examined is indicated in parentheses.

Treatment can usually assure a normal life. *See* CIRCULATION; HEART DISORDERS. Michael Horan

Bibliography. P. R. Hebert et al., The community-based randomized trials of pharmacologic treatment of mild-to-moderate hypertension, *Amer. J. Epidemiol.*, 127:581–590, 1988; W. B. Kannel, Role of blood pressure in cardiovascular morbidity and mortality, *Prog. Cardiovasc. Dis.*, 17:5–24, 1974; The 1988 Report of the Joint National Committee on Detection, Evaluation, and Treatment of High Blood Pressure, *Arch. Internal Med.*, 148:1023–1038, 1988; Society of the Association of Life Insurance Medical Directors and the Society of Actuaries, *Blood Pressure Study 1979*, 1980.

Hypervalent compounds

Group 1, 2, and 13–18 compounds which contain a number (*N*) of formally assignable electrons of more than eight (octet) in a valence shell directly associated with the central atom (X) in direct bonding with a number of ligands (*L*). The designation *N*-X-*L* is conveniently used to describe hypervalent molecules. *See* ELECTRON CONFIGURATION; LIGAND; VALENCE.

Compounds of main group elements in the second row (such as carbon, nitrogen, and oxygen) have eight valence electrons. As such, the fundamental shapes of their atoms are linear (such as acetylene, sp orbital), triangular, (such as ethylene, sp^2 orbital), and tetrahedral (such as methane, sp^3 orbital). In contrast, main group elements in the third row of the periodic table (such as silicon, phosphorus, and sulfur) may contain more than eight electrons in a valence shell. These are called hypervalent compounds. Fundamental shapes of *10*-X-*5* molecules (including *10*-X-*4* molecules bearing a pair of unshared electrons) are trigonal bipyramid (TBP) or square pyramid (SP), and those of *12*-X-*6* (including *12*-X-*5* bearing a pair of unshared electrons) are octahedral. Hence, there is an apparent similarity in shape between hypervalent compounds and organotransition metal compounds. *See* MOLECULAR ORBITAL THEORY.

The hydrogen bond is one of the best examples of a hypervalent bond. The covalent nature of the hydrogen bond of the [N—H···N] system has recently been established experimentally and theoretically. The structure of $(HF)_5$ is pentagonal,

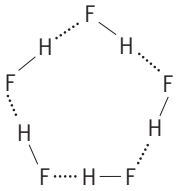

with a bond length for F—H of 1.0 Å and H···F of 1.5 Å. The hydrogen atom shifts rapidly between the two fluorines in the range of 0.5 Å. However, $[F—H—F]^-K^+$, (*4*-H-*2*), is linear, and the two F—H bond lengths are elongated to 1.13 Å, which is typical for a hypervalent bond. *See* CHEMICAL BONDING; HYDROGEN BOND.

In order to accept extra electrons in a valence shell, an electron-rich and polarized sigma bond constitutes an apical bond (three-center, four-electron bond, which is defined as a hypervalent bond by molecular orbital theory) on the central atom of *10*-X-*5*. One of the most unstable hypervalent molecules is $[F—F—F]^-K^+$, (*10*-F-*2*). It is linear and the F—F bond is calculated to be 1.701 Å, which is elongated from that of F—F (1.412 Å) by accepting a fluoride ion. This is essentially the same as that of $[F—H—F]^-Li^+$, (*4*-H-*2*).

Novel and fundamental characteristics of hypervalent compounds are summarized as follows:

1. For TBP molecules of *10*-X-*5*, intramolecular positional isomerization between apical bonds and equatorial bonds (sp^2 orbital) is very rapid. This is demonstrated by a Berry pseudorotation, Sub-

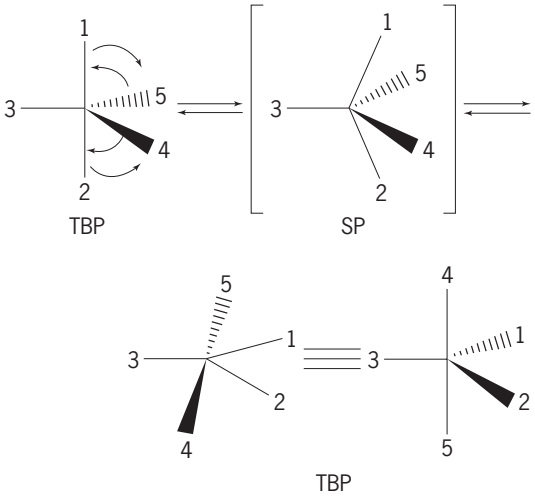

stituents (ligands) with larger electron-withdrawing ability preferably occupy apical positions (apicophilicity).

2. Ligand exchange reactions (LER) and ligand coupling reactions (LCR) take place for hypervalent compounds of *10*-X-*5* and *10*-X-*4*. These are formally similar to LER and reductive elimination of organometallic compounds.

3. The stereochemistry of nucleophilic substitution at *4*-coordinate main group compounds of higher rows is complex, unlike inversion at a carbon atom by a nucleophilic substitution reaction. Both retention and inversion occur due to the stability of intermediate *10*-X-*5* species, where rapid pseudorotation takes place to change the relative positions of substituents. *See* STEREOCHEMISTRY.

4. Edge inversion can be the mechanism of inversion for main-group element compounds in the higher-numbered rows,

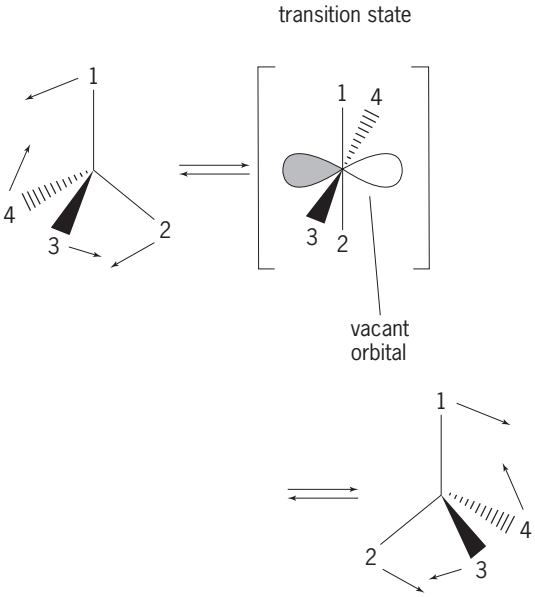

This is applicable for *10*-X-*3* compounds bearing lone pair electrons. This mechanism involves the appear-

ance of a vacant *p*-type orbital at the transition state which can be coordinated by two nucleophiles to help stabilize it. This is an essentially different mechanism from that of ammonia, which has a vertex inversion.

5. The stereochemistry of *12*-X-*6* compounds is stable in contrast to *10*-X-*5* compounds. This is composed of three hypervalent bonds. Kin-ya Akiba

Bibliography. K.-y. Akiba (ed.), *Chemistry of Hypervalent Compounds*, 1999; S. Borman, Hydrogen bonds revealed by NMR, *C&EN*, pp. 36–38, May 10, 1999.

Hyphochytriomycota

A phylum of microscopic, fungus-like organisms that reproduce with anteriorly uniflagellate zoospores (independently motile spores). In Hyphochytriomycota, two opposite rows of hairlike structures (mastigonemes) cover the flagellum. The mastigonemes are hollow and split into three parts at their tips. The zoospore swims with the flagellum directed straightforward and the tip moving back and forth, propelling the oval to pyriform zoospore forward. Functioning in dispersal, the zoospore uses stored food reserves until it reaches a substrate that it can use. After a period of swimming, between 30 min to 2 h, the zoospore rounds up, absorbs its flagellum, secretes a wall around itself, and develops into a thallus. The thallus can be a branching filament with spindle-shaped swellings and sporangia (spore-containing structures) connected by empty isthmuses (polycentric, eucarpic, **Fig. 1**); a sporangium with rootlike rhizoids (monocentric, eucarpic, **Fig. 2**); or simply an endobiotic (living in the cells or tissues of a host) sporangium (monocentric, holocarpic). [Polycentric means having a number of centers of growth and development and with more than one reproductive organ; eucarpic, using only part of the thallus for the fruit body; monocentric, having one center of growth and development; holocarpic, using the entire thallus for the fruit body.] The thallus is the growth and heterotrophic phase, acquiring nutrients by absorption. As the thallus grows, the sporangium becomes multinucleate. When mature, the sporangial cytoplasm cleaves out uninucleate, uniflagellate zoospores. Zoospore formation can occur within the sporangium, but in some genera, such as *Rhizidiomyces*, zoospores are formed externally. A long discharge tube grows from the sporangium (Fig. 2) as the sporangial protoplast extrudes over about a 15-min period and comes to lie as a pulsating mass at the tip of the exit tube. Zoospores are then cleaved, swarm as a mass for about 5 min, and eventually swim away. Hyphochytrids also produce resistant sporangia that aid in surviving adverse conditions. Most resistant sporangia are produced asexually, and there have been only a few reports of sexual reproduction in this group. *See* EUMYCOTA; FUNGI; MASTIGOMYCOTINA; REPRODUCTION (PLANT).

Taxonomy and phylogenetic relationships. Three families are generally recognized and are based on thallus type: Hyphochytriaceae (polycentric, eucarpic), Rhizidiomycetaceae (monocentric, eucarpic), and Anisolpidiaceae (monocentric, holocarpic). Taxonomists debate the classification of

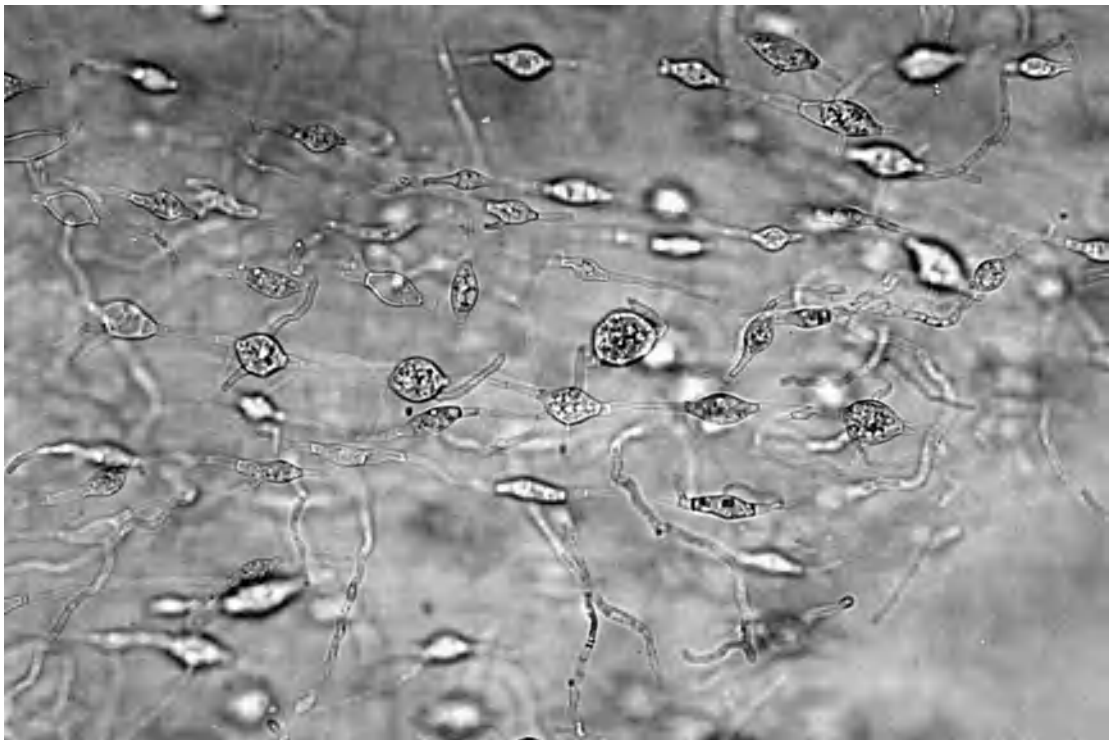

Fig. 1. Polycentric, filamentous thallus of *Hyphochytrium catenoides*. Fusiform swellings, separated by empty isthmus cells, will round up and become sporangia that later produce anteriorly uniflagellate zoospores.

Fig. 2. Eucarpic, monocentric thallus of *Rhizidiomyces apophysatus* with a long, meandering exit tube through which the sporangial protoplast will extrude and cleave externally into zoospores. Recently encysted zoospores with threadlike rhizoids surround the mature thallus.

genera in this phylum, which encompasses approximately 29 described species in five to seven genera. There have been only limited molecular phylogenetic analyses of genes in this group, and it is not known whether classification based on thallus structure results in natural groupings or whether all taxa included in the phylum are actually related. Biochemical and molecular phylogenetic studies, however, do show that the hyphochytrids, *Hyphochytrium* and *Rhizidiomyces*, are closely related.

Hyphochytrid thallus forms parallel to those of the zoosporic fungi, Chytridiomycota, but hyphochytrids are actually more closely related to photosynthetic (phototrophic) and nonphotosynthetic (heterotrophic) organisms classified in the kingdom Straminopila or the monophyletic supergroup Chromalveolata. Hyphochytrids are clearly related to the water molds (Oomycetes, Peronosporomycetes) and to heterokont (having unequal flagella) algae, forming a monophyletic group in analyses of ribosomal sequences. Hyphochytrids share a number of biochemical characteristics with their biflagellate relative, the Oomycetes. Their cells walls contain cellulose, their mitochondrial cristae are tubular, and they produce lysine by the diaminopimelate (DAP) pathway. However, hyphochytrids are still distinctive. Some contain chitin as well as cellulose in their cell walls. Unlike Oomycetes, nuclear division in hyphochytrids is not totally closed, but fenestrae form at the poles of the nuclear envelope during mitosis and intranuclear vesicles coalesce around condensed separating chromatin, reconstituting new nuclear envelopes. *See* CHYTRIDIOMYCOTA; OOMYCOTA.

In phylogenetic analyses of ribosomal genes, the sequence of divergences among groups in the stramenopiles is controversial. However, most analyses reveal that the heterotrophic members are basal to the more derived phototrophic stramenopiles (such as diatoms, brown algae, chrysophytes, and yellow-green algae). The hyphochytrids and oomycetes are sister to each other (also known as pseudofungi) and cluster with other basal heterotrophic stramenopiles, including *Developayella elegans*, bicosoecids, thraustochytrids, labyrinthulids, and opalinids.

Zoospores. The most distinctive feature of the hyphochytrids is the ultrastructure of its zoospore (**Fig. 3**). Although its relatives have two flagella, one smooth and the other tinsel (covered with lateral filaments), hyphochytrid zoospores bear a single anterior tinsel flagellum (with tubular tripartite mastigonemes). The transition region of the flagellum resembles that of the oomycete *Saprolegnia*, consisting of a set of five to seven rings and a zone of struts perpendicular to the flagellar microtubules. The flagellum extents from a kinetosome (basal body), and a secondary centriole is positioned at a 150° angle. The backward projection of the secondary centriole suggests loss of the posteriorly directed flagellum and an earlier divergence from a

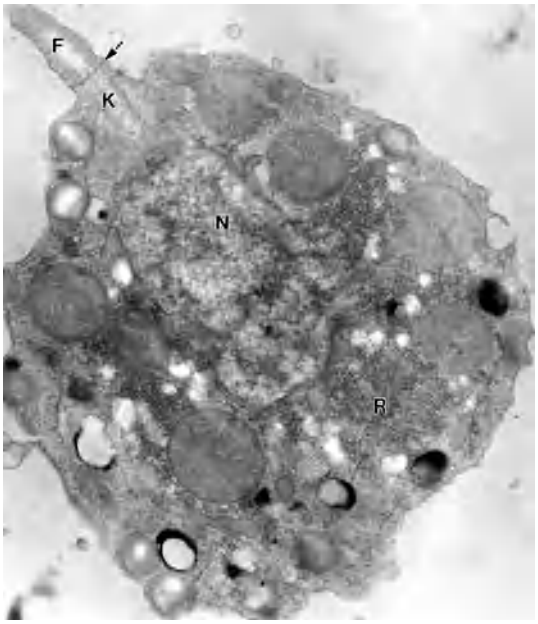

Fig. 3. Transmission electron microscopy image of a longitudinal section of a zoospore of *Rhizidiomyces apophysatus*. The flagellum (F) extends from the kinetosome (K). A cluster of coarsely granular ribosomes (R) surround the posterior portion of the single nucleus (N). The arrow points to the basal plate, with the transition zone located more anterior, with stacked rings (transitional helix) visible in longitudinal section as two stacks of dots.

biflagellate ancestor shared with oomycetes. With the loss of the flagellum, there was a loss of a root system. In the hyphochytrid zoospore, three of the four systems of microtubular roots found in oomycetes extend from the region around the kinetosome. The zoospore contains a single nucleus with ribosomes clustered over the posterior portion.

Importance. Despite being a relatively small group in terms of number of named species, hyphochytrids are ecologically diverse. Hyphochytrids can be found in soils and in freshwater and marine habitats as decay-causing organisms that break down plant and insect parts or as parasites of algae and fungi. As common parasites of algae in freshwater and marine environments, up to 50% of a population has been reported to be infected with hyphochytrids. *Hypochytrium catenoides* is weakly parasitic on root hairs of corn, and in soil it may be a component of suppressive soils inhibitory to the root rot oomycete, *Phytophthora cinnamomi*. Because hyphochytrids can parasitize spores of oomycetes and endomycorrhizae, they may have a role in natural control of microbial populations. Recent analyses of molecular sequences from marine environmental samples have revealed novel lineages closely related to hyphochytrids, and the prevalence of cryptic related taxa suggests hyphochytrids may be more diverse and more vital in microbial food webs than previously recognized.

Martha J. Powell

Bibliography. C. J. Alexopoulos, C.W. Mims, and M. Blackwell, *Introductory Mycology*, 4th ed., 1996; M. W. Dick, *Straminipilous Fungi*, 2001; M. S. Fuller, Phylum Hyphochytriomycota, in L. Marqulis et al. (eds.), *Handbook of Protoctista*, pp. 380–387, 1990; J. S. Karling, *Chytridiomycetarum Iconographia*, Lubrecht & Cramer, Monticello, NY, 1977; F. K. Sparrow, *Aquatic Phycomycetes*, 2d rev. ed., University of Michigan Press, Ann Arbor, 1960.

Hyphomycetes

Fungi that form spores asexually (via mitosis) and that produce their spores externally on threadlike hyphae (as opposed to bearing them inside a fruiting body). In traditional taxonomy, Hyphomycetes was a formal class, but because this group is nonphylogenetic (it is composed of more than one evolutionary lineage), many mycologists now either attach a prefix, as in "form-class Hyphomycetes," or use the term colloquially, as in "hyphomycetes." The mitotic spores of hyphomycetes (as well as those of coelomycetes) are called conidia. *See* AGONOMYCETES; COELOMYCETES; DEUTEROMYCOTINA; FUNGI.

Taxonomy and identification. Most hyphomycetes are life stages of ascomycetes, but some are related to basidiomycetes; for example, *Ingoldiella* is a relative of some of the jelly fungi, and *Nematoctonus* is closely related to a genus of mushrooms. Both these genera have clamp connections, buckle-shaped structures common in the hyphae of many basidiomycetes. However, most hyphomycetes are phylogenetically ascomycetes. For example, some *Penicillium* species are anamorphic (asexual) stages in the teleomorphic (sexual) ascomycete genus *Talaromyces*. Other species of *Penicillium* have sexual stages in a different ascomycete genus, *Eupenicillium*, or have no known sexual stage at all. Some hyphomycetes may appear to lack a sexual state simply because the teleomorph is as yet undiscovered, but in other instances the sexual state may no longer exist.

Hyphomycetes can be further divided on the basis of gross morphology, color and shape of conidia, and characteristics of their conidiogenous (conidium-producing) cells. Hyphomycetes which produce conidiophores (conidium-bearing hyphae) aggregated into cushion-shaped structures called sporodochia are in the form-order Tuberculariales. Some hyphomycetes have conidiophores aggregated into stalk- or club-shaped structures (synnemata), or head-shaped structures (coremia). These latter hyphomycetes belong to the form-order Stilbellales. The hyphomycetes whose conidiophores occur singly, or at least are not aggregated in the above manners, are in the form-order Moniliales. None of these form-orders is phylogenetic, but all are useful for identification. Moniliales has been further divided into dark- (dematiaceous) and light-spored groups, which themselves can be subdivided on the basis of conidial shape and number of septa (cell-dividing walls) in the conidia. All of these categories represent classifications of convenience, and some fungi display characters intermediate between these categories. For this reason, it has been very useful to utilize attributes of the conidiogenous cells and formation of conidia. Some conidia are produced through the tips of bottle-shaped cells called phialides (enteroblastic conidiogenesis), whereas others are produced from swellings of the conidiophore wall (holoblastic conidiogenesis). Sometimes conidia are produced through holes in a conidiophore wall (tretic conidiogenesis) or by formation and dissolution of many cross-walls at the apex of a conidiophore (thallic conidiogenesis). There are many variations of the above, including the direction and pattern of conidiogenesis (acropetal, basipetal, sympodial, etc.). This system of identification and classification by conidiogenesis, initially introduced by the mycologist S. J. Hughes, is the basis of modern identification of hyphomycetes. Identification efficiency is increased when characters from conidiogenesis are combined with information on spore color, septation, and ornamentation (spines, warts, grooves, etc.). Increasingly, identification is enhanced by use of molecular methods such as random amplified polymorphic DNAs (RAPDs), restriction fragment length polymorphisms (RFLPs), amplified fragment length polymorphisms (AFLPs), simple sequence repeats (SSRs), DNA sequence determination (often obtained via polymerase chain reaction, PCR), or metabolic profiles.

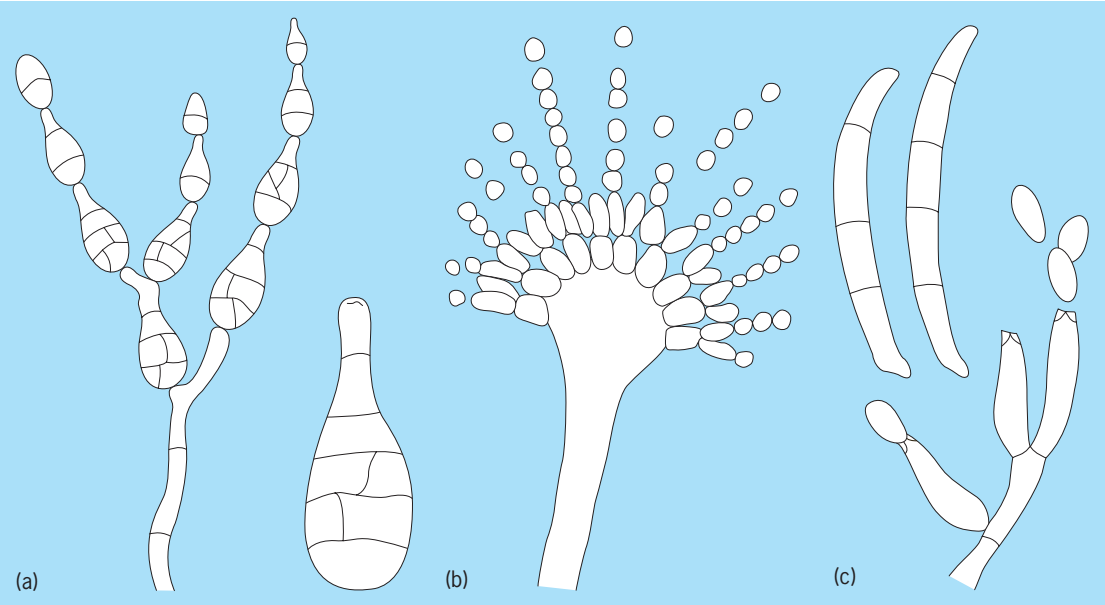

Three hyphomycete genera. (*a*) *Alternaria*. (*b*) *Aspergillus*. (*c*) *Fusarium*.

See ASCOMYCOTA; BASIDIOMYCOTA; FUNGAL GENETICS; GENE AMPLIFICATION.

Cultivation and preservation. Most hyphomycetes can be cultivated on artificial media in the laboratory, a distinct asset in diagnosis and identification. In fact, cultivation under standardized conditions is a prerequisite for accurate identification of species in several genera such as *Alternaria*, *Aspergillus*, *Fusarium*, *Penicillium*, and others (see **illustration**). Production of conidia, a prerequisite for successful identification, is usually enhanced on media of low sugar content and sometimes by exposure to periodic fluorescent or near-ultraviolet light. Occasionally, it is necessary to experiment with plant materials, sterilized in an autoclave, by radiation, or by exposure to propylene oxide, and used as an amendment to the medium. For example, radiation-sterilized carnation leaves are used to promote formation of sporodochia by *Fusarium* isolates. Hyphomycetes, like other fungi, are best preserved in an inert (nongrowing) condition. One popular method is mixing conidia into vials with a dilute solution of glycerol and placing the vials into liquid nitrogen vapor or into an ultralow-temperature freezer [approximately −80°C (−112°F)]. Lyophilization (freeze-drying) in hermetically sealed vials can be used to preserve some species. Less technologically intensive approaches include putting conidia in silica gel or sterile soil, or drying them on filter paper, followed by storage in a conventional freezer [approximately −20°C (−4°F)]. Some laboratories prefer to excise several small (1–2 mm^3) cubes from an actively growing culture and store them at room temperature in vials of sterile water. Some fungi will keep well when grown on "slants" (test tubes, filled partly with medium and placed at an angle so that the medium hardens on a slant), then covered with sterile mineral oil and stored in a refrigerator [approximately 4°C (39°F)]. Whichever method is adopted, it is advisable to assess viability at periodic intervals by reculturing the fungi onto fresh artificial media. Although many fungi (including hyphomycetes, whose spores are often easily dispersed in air) are innocuous, some are pathogenic or allergenic to humans, so it is advised that manipulation of fungi take place in a biological safety cabinet or in a transfer box. *See* CRYOBIOLOGY.

Biology, economic importance, control, and utilization. In general, sexual stages of fungi provide for recombination of genes (via meiosis) and often serve to initiate new growth cycles in the spring. The subsequent asexual stages produce more numerous spores; hence they function strongly in the spread to new hosts, substrates, or environments. For example, Fusarium head blight, a disease of small grains (predominantly wheat and barley), is often initiated by ascospores from the sexual stage *Gibberella zeae*. Growth resulting from ascospores can produce conidia of the hyphomycetous stage, *Fusarium graminearum*, further fueling the epidemic. Species lacking a sexual cycle (such as *F. oxysporum*) may overwinter as resistant structures (chlamydospores, sclerotia) or by inhabiting buried debris. Some dark-spored hyphomycetes (the dematiaceous hyphomycetes) have conidia with thick melanized walls and contain ample nutrient reserves. In such cases, the conidia themselves are effective survival structures, as well as functioning in dissemination.

Numerous genera of hyphomycetes are well known as plant pathogens. Prominent examples include *Alternaria* (on many plants but especially fruits and vegetables), *Botrytis* (a serious problem on various fruits, vegetables, and cut flowers), *Curvularia* (often on grasses), *Cylindrocarpon* (a relative of *Fusarium*, and causing leaf spots, and shoot, stem, and root diseases of miscellaneous plants), *Fusarium* (on many plants but especially

cereal grains), *Penicillium* (on grains, fruits, and vegetables), *Ramularia* (and its relatives *Cercospora*, *Cercosporella*, and *Pseudocercosporella* on miscellaneous hosts), *Stemphylium* (several species on forage legumes), and *Verticillium* (on various plants, with one species attacking cultivated mushrooms). Some of these pathogens are very host specific. For instance, *Fusarium oxysporum* f. sp. *cepae* is specialized to attack host plants in the genus *Allium* (onion and garlic). Others are opportunists with a broad host range, such as *Cladosporium herbarum*, generally a saprophyte (an organism living on dead or decaying organic matter) but also attacking ripe or overripe fruit. Some genera have very few pathogenic members. For example, *Trichoderma*, a generally saprophytic genus, is often used in biological control of plant diseases, but one species causes disease in cultivated mushrooms. Disease management for these and other fungi is most often exerted by combining agronomic practices with use of fungicides. *See* AGRONOMY; PLANT PATHOLOGY.

Some hyphomycetes cause disease in humans or animals. Especially well known are *Epidermophyton*, *Microsporum*, and *Trichophyton*; the first is known as the cause of athlete's foot, and the latter as agents of ringworm. *Blastomyces*, *Coccidioides*, *Histoplasma*, and *Paracoccidioides* are hyphomycetes causing serious diseases in humans. Several dematiaceous hyphomycetes are the cause of chromoblastomycosis (a skin disease characterized by warty nodules that may ulcerate); mycetoma (a chronic infection, usually of the feet, resulting in swelling), and other diseases in both normal and immunocompromised hosts. In fact, many other hyphomycetes (for example, *Acremonium*, *Chrysosporium*, *Fusarium*, and *Paecilomyces*) are well documented in clinical situations. Some species of *Aspergillus*, especially *A. fumigatus*, can cause respiratory diseases. Various species in *Alternaria*, *Aspergillus*, *Cladosporium*, and *Penicillium* are documented as allergens affecting a high proportion of humans, but the hyphomycete which has most captured the public imagination is *Stachybotrys*, one of the agents of sick building syndrome. Several of these same fungi, especially species in *Aspergillus*, *Fusarium*, and *Penicillium*, are well documented as producing mycotoxins in human or animal foods. Animals can also be affected by allergenic or toxigenic fungi in forage, moldy hay, or silage (fodder). Facial eczema in sheep is caused by a species of *Pithomyces*. Mycotic abortion in cows and heaves in horses are attributable to adverse effects of toxigenic hyphomycetes, especially *Aspergillus fumigatus*. Control of invasive fungal diseases is most often with antifungal compounds, sometimes combined with surgery. Indoor allergy problems are best resolved by structural improvements which deprive fungi of moisture. Quality control for production of food, feeds, and fodder is necessary to protect humans and animals from adverse effects of fungal contamination. *See* ALLERGY; MEDICAL MYCOLOGY; MYCOTOXIN.

Hyphomycetes also play a significant role in degradation or discoloration of manufactured products such as paper, textiles, building materials (including wood, stone, rubber, and plastic), and fossil fuels. Notable examples are *Hormoconis resinae* in diesel fuel, *Aureobasidium pullulans* on some plastics, *Phialophora* species in pulp and paper, and *Epicoccum* on stone. Modern building materials can be impregnated with a variety of fungi-resistant compounds, but restoration of deteriorating monuments and artifacts presents multiple challenges to conservators.

Although some hyphomycetes are undesirable, many species undoubtedly play beneficial roles as saprophytes in decomposition of plant residues and in nutrient recycling. Fungal mycelia are also an important component in soil structure. Many hyphomycetes are opportunistic phytopathogens (causing disease only on weakened or senescent plants) and are otherwise beneficial saprophytes. Other species cause disease on weeds (for example, *Myrothecium* on kudzu vine) or arthropod pests (for example, *Beauveria* on whiteflies and *Metarhizium* on locusts), or are used to counteract phytopathogenic fungi (for example, *Trichoderma* against *Rhizoctonia*, *Pythium*, and wood-rotting fungi). *Arthrobotrys* species are predators of nematodes. Not only are hyphomycetes increasingly useful in biological control, but several hyphomycetes play major roles in biotechnology. *Aspergillus niger* is used commercially to produce cellulases and citric acid, and *Aureobasidium pullulans* for the production of pullulan, a food additive. *Aspergillus oryzae* produces proteases used in the manufacture of breads and cheeses. Lipases for food manufacture are derived from some *Penicillium* species, and some *Penicillium* species have played crucial roles in the production of penicillin antibiotics. Perhaps the most famous example of fungi in food biotechnology is the manufacture of soy sauce with the hyphomycete *Aspergillus sojae*. *See* FUNGAL BIOTECHNOLOGY; FUNGAL ECOLOGY.

Frank M. Dugan

Bibliography. H. L. Barnett and B. B. Hunter, *Illustrated Genera of Imperfect Fungi*, 1998; J. W. Carmichael et al., *Genera of Hyphomycetes*, 1980; G. S. de Hoog et al., *Atlas of Clinical Fungi*, 2d ed., 2000; F. M. Dugan, *The Identification of Fungi*, 2006; E. Kiffer and M. Morelet, *The Deuteromycetes*, 2000; J. R. von Arx, *The Genera of Fungi Sporulating in Pure Culture*, 1981.

Hypnales

An order of the true mosses (subclass Bryidae). Also known as the Hypnobryales, this order consists of 14 families and some 135 genera, primarily put together because they share a hypnoid peristome (as described below). The families show considerable sporophytic unity; it is gametophyte structure that determines family membership.

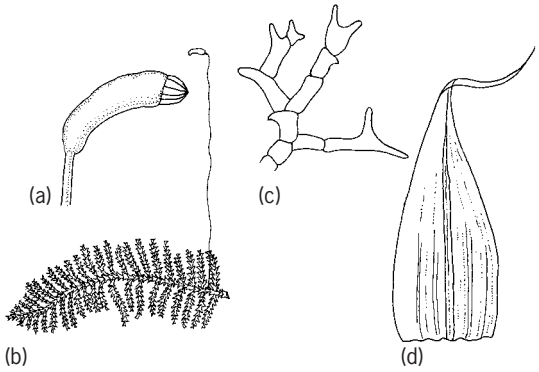

Fig. 1. ***Thuidium recognitum*. (*a*) Urn and peristome. (*b*) Portion of plant. (*c*) Paraphyllium. (*d*) Perichaetial leaf. (*After W. H. Welch, Mosses of Indiana, Indiana Department of Conservation, 1957*)**

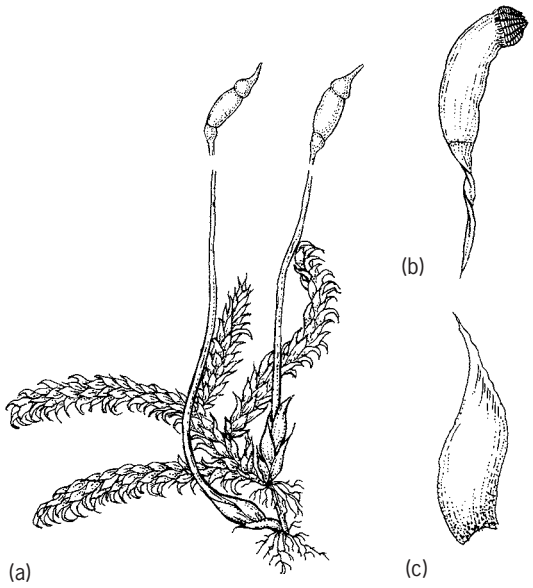

Fig. 2. ***Hypnum reptile*. (*a*) Portion of plant, with stems shortened. (*b*) Urn and peristome. (*c*) Faintly bicostate leaf. (*After W. H. Welch, Mosses of Indiana, Indiana Department of Conservation, 1957*)**

Hypnales are often known as feather mosses owing to their freely branched stems, with branches regularly or irregularly arranged in two rows. The plants form mats, especially in woodlands, with pseudoparaphyllia and paraphyllia often present. The leaves are most commonly acute or acuminate, and the costa is generally short and double or nearly lacking (**Figs. 1** and **2**). The upper cells are mostly elongate and smooth, but they may be variously papillose, and the alar cells are often differentiated. The sporophytes are lateral with elongate, smooth setae, and the capsules are generally inclined and asymmetric with well-developed double peristomes. The peristome generally has exostome teeth abruptly tapered, cross-striate below, bordered, and trabeculate, and endostome segments keeled, alternating with cilia, and rising from a high basal membrane. The calyptrae are cucullate and usually naked. The chromosome numbers range from 5 to 22, often in polyploid series. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA.

Howard Crum

Hypnosis

A presumed altered state of consciousness in which the hypnotized individual is usually more susceptible to suggestion than in his or her normal state. In this context, a suggestion is understood to be an idea or a communication carrying an idea that elicits a covert or overt response not mediated by the higher critical faculties (that is, the volitional apparatus).

Effects. Hypnosis cannot be physiologically distinguished from the normal awake state of an individual, and for this reason its existence has been questioned by some investigators. There are few phenomena observed in association with hypnosis, if any, that are specific to the hypnotic state. Most are directly or indirectly produced by suggestions. It is usual to assign a depth or degree to hypnosis that is proportional to the extent to which the presumably hypnotized person is found to be suggestible. Through suggestions given to hypnotized individuals, it is possible to induce alterations in memory, perception, sensation, emotions, feelings, attitudes, beliefs, and muscular state. Such changes can be, and usually are, incorporated into the complex behavior of the individual, resulting in amnesias and paramnesias, fugue-like conditions, paralysis, loss of sensory functions, changes in attention, personality alterations, hallucinatory and delusional behavior, and even physiological changes. Enhanced recall is sometimes possible. Although sometimes remarkable, the effects produced through hypnosis with the majority of individuals are much less spectacular than popularly believed.

Many of the phenomena seen in association with hypnosis and suggestion can be observed to occur in situations that are not obviously ones involving hypnosis or suggestion. Some investigators have viewed this as indicating that there is no hypnosis or suggestion, whereas others have interpreted this as indicating that all these situations effectively involve hypnosis or suggestion. However, there is just as much reason to conclude that neither is the case, and that the effects in question are simply not unique to hypnosis and suggestion. Hypnosis and suggestion are then simply convenient means for producing these effects.

Theories. The theories of hypnosis attempt to explain its essential nature. They all are inadequate in some respects. These theories, arranged approximately in order of their chronological development, adduce that hypnosis is: a form of subtle emanation between the hypnotist and subject; a modified form of sleep; a form of dissociation; a conditioned response resulting in ideomotor action (that is, nonvoluntary movement resulting from some idea); a form of role playing or goal-directed activity; a form of regression to a level representing the relation between parent and child; a form of regression in which the subject assumes the submissive role similar to that assumed by the female in the sexual role; a form of interpersonal relationship between the subject and hypnotist, but not necessarily of the kinds represented in the theories of regression. Finally, one

theory, which is an extension of the theory of role playing, takes the position that there is no state of hypnosis, and that all hypnotic phenomena can be accounted for without this concept; this view is held by only a small number of investigators. Although the foregoing statements represent an oversimplification, they do give some idea concerning the nature of the theories.

Applications. The applications of hypnosis to the field of medicine and psychology are exceedingly wide in scope.

Diagnostic tool. Hypnosis has been employed successfully in distinguishing functional from organic disorders. For example, functional deafness may be distinguished from organic deafness; hysterical conversion paralysis may be differentiated from organic paralysis; psychogenic convulsions may be contrasted with organic convulsions. Hypnosis may also be used in determining a normal baseline for the functioning of certain endocrine glands. Anxiety may bring about dysfunction of the thyroid; consequently a metabolic or other test would not distinguish between a pathological and a functional disorder. The anxiety can be reduced by hypnotic relaxation, and a comparative relationship can be established between the output produced by an anxiety component and possible pathology.

Hypnosis can also be used in teasing out the part that a functional overlay plays in what appears to be an organic condition. For example, in a limb in which there is a partial paralysis due to any organic cause, there may be superimposed a further dysfunction due to psychological causes that can be demonstrated by hypnotic suggestion.

Aid in examination. Hypnosis can be used in controlling pain in certain types of examination, as in cystoscopy, and leaves the person able to assist the physician by giving information concerning the location of the instrument. It has an advantage over analgesics or anesthetics in such situations. Hypnosis may be useful in nerve block infiltrations and in making spinal punctures. It can also be employed successfully to reduce gagging when stomach tubes are inserted, and has been utilized in the control of bleeding, salivation, and muscle spasms.

Anesthesiology and surgery. Hypnosis has been used effectively for the control of pain. In cases where an analgesic or anesthetic is inadvisable, it may be used in place of, or in conjunction with, chemoagents. Hypnosis may also be used where tolerance becomes a factor in long-term cases. Pain control in all kinds of surgical operations is dependent on the susceptibility of the hypnosis subject. Tooth extractions, cranial operations, amputations, cardiac and lung surgery, as well as abdominal and genitourinary operations, have been performed without any kind of anesthesia except hypnosis. *See* ANALGESIC; ANESTHESIA; PAIN.

Symptom control. Intractable pain, such as that involved in trigeminal neuralgia, phantom limb, carcinoma, and burns, has been allayed. Hypnosis has been used in the treatment of hiccoughs, neurodermatitis, the dumping syndrome (weakness, nausea, vertigo, and palpitation occurring immediately after partial or complete removal of the stomach), coughing spasms, headaches, frigidity, impotence, dysmenorrhea, discomfort from contact lenses, stuttering, tics, and the pain of childbirth. Many organic symptoms, as well as most functional symptoms, can be suppressed in appropriate cases. The suppression may be complete or only partially controlled, so that the individual is still able to report dangerous developments that may be occurring. The above symptoms are only a partial list of those in which hypnosis has been applied.

Motivational tool. Hypnosis has been utilized to motivate speech retraining in asphasics, to increase the desire for psychotherapy, to enhance or decrease movement of limbs that are involved in surgical procedures, to increase motivation for early ambulation following surgery, and to enhance feeling tone in some very depressed individuals. Hypnosis has also been used to alter attitudes, likes, and dislikes, as in the case of obesity, for various foods and eating; in this connection, it has been used as an uncovering technique, with the person hypnotized only part of the time. *See* MOTIVATION.

Conjunction with analysis. Hypnosis is used in a number of ways with psychiatric analysis. Analysis may proceed entirely in the hypnotic state, or frequently hypnosis is used only as an aid in uncovering repressed material, and has been particularly useful in overcoming blocks in free association and in producing dreams. Hypnosis has been used to create artificial neurotic behavior that permits the individual to better understand his or her own neurotic symptoms and control them. Another application is to create illusory situations through which the individual works, thus enabling the therapist to gain a better understanding of the person's ability to cope with certain problems. Hypnosis has been utilized to enhance the feelings of schizophrenics who may develop some intellectual understanding of their problem but who otherwise remain unresponsive. *See* PSYCHOTHERAPY.

Conjunction with behavior therapy. One of the earliest uses of hypnosis was in the production of aversive reactions in the treatment of smoking and alcoholism. Hypnosis is an excellent tool for producing the state of relaxation which is called for by certain behavior-modification techniques, and it is also an effective way of creating various experiences which can be used for purposes of desensitization and of reinforcement.

Pharmacological substitute. Through the effects that may be produced in its presence, hypnosis may be used to provide substitutes for various pharmacological agents or may be employed with them to increase their effects. Hypnotic suggestions have thus been particularly used in place of or with soporifics, hypnotics, anesthetics, analgesics, and tranquilizers. Some lesser uses have included those of a stimulant, mood elevator, laxative, appetite reducer or enhancer, antiallergenic agent, lactation inhibitor or stimulant, and agent for control of menstrual

discharge and cramps. In all of these instances, the advantages of hypnosis lie in the absence of tolerance problems, side effects, and other complications. The main disadvantage lies in the fact that such uses of hypnosis are probably limited to 20% of the population. Also, psychobiological problems can arise in unselected individuals. *See* CONSCIOUSNESS; PSYCHOPHARMACOLOGY. André M. Weitzenhoffer

Bibliography. S. J. Lynn and J. W. Rhue (eds.), *Theories of Hypnosis: Current Models and Perspectives*,1991; R. H. Rhodes, *Hypnosis: Theory, Practice, and Applications*, 1989, reprint 1998; A. Weitzenhoffer, *Practice of Hypnotism*, 2 vols., 2d ed., 1999; W. C. Wester, II and H. Alexander, Jr., *Clinical Hypnosis: A Multidisciplinary Approach*, 1991.

Hypodermis

The outermost cell layer of the cortex of plants. It forms a prominent layer immediately under the epidermis in many but not all plants (see **illus.**). In shoots, the hypodermis may be composed of parenchyma, collenchyma, or sclerenchyma and be from one to several cells thick. In roots, the hypodermis is often called the exodermis; it resembles the endodermis, and it develops Casparian strips, suberin deposits, and cellulose deposits impregnated with phenolic or quinoidal substances. Thus the root hypodermis is similar to the endodermis in cell wall anatomy and in its reaction to histochemical tests. The hypodermis is the mirror image of the endodermis in appearance; in the endodermis, wall deposits develop from the inner tangential wall outward; in the hypodermis, from the outer tangential wall inward. *See* CELL WALLS (PLANT).

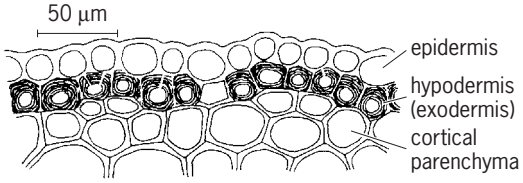

Transection from the outer part of a *Smilax* root, illustrating a thick-walled hypodermis (exodermis) beneath the epidermis. One hypodermal cell is not thickened. (*After K. Esau, Plant Anatomy, 2d ed., John Wiley and Sons, 1965*)

The hypodermis may produce substances that act as a barrier to the entry of pathogens, and in some plants it may function in the absorption of water and the selection of ions that enter the plant. Young cells of the hypodermis in the growing tip of the root contain cytochrome and peroxidase, enzymes associated with respiration and centers of metabolic activity. There is a large amount of protein and nucleic acid in the cells of the hypodermis, with unusually large nuclei, as in the endodermis. Similarity in substance content of the hypodermis and endodermis, with a similarity in anatomical form, suggests a common function for both tissues, but the function of the hypodermis has yet to be determined. *See* CORTEX (PLANT); ENDODERMIS. D. S. Van Fleet

Hypohalous acid

An oxyacid of a halogen [fluorine (F), chlorine (Cl), bromine (Br), iodine (I), or astatine (At)] possessing the general chemical formula HOX, where X is the halogen atom. The chemical behavior of hypofluorous acid (HOF) is dramatically different from the heavier hypohalous acids which, as a group, exhibit similar properties. These differences are attributed primarily to the high electronegativity and small size of the fluorine atom, which cause HOF to be an extremely strong oxidant with an anomalously weak O-F bond. Thus, the molecule is highly reactive and relatively unstable. (Gaseous HOF decomposes to HF and O_2 at room temperature with a half-life of about 1 h, and the liquid has a tendency to explode.) Because the most electronegative element in HOF is fluorine, whereas the other halogen atoms are less electronegative than oxygen, the O-X bond polarities are reversed in HOF and the heavier congeners. HOF therefore acts primarily as an oxygenating and hydroxylating agent, whereas the other hypohalous acids are electrophilic halogenating agents. For example, HOF hydroxylates aromatic compounds to form phenols and reacts instantaneously with water to give hydrogen peroxide [reaction (1)], whereas

$$HOF + H_2O \longrightarrow HF + H_2O_2 \qquad (1)$$

hypochlorous acid (HOCl) chlorinates aromatic compounds and is unreactive toward water. *See* ASTATINE; HALOGEN ELEMENTS.

Hypochlorous, hypobromous, and hypoiodous acid solutions are formed by disproportionation of the corresponding halogen [reaction (2)]. In aque-

$$X_2 + H_2O \rightleftharpoons H^+ + X^- + HOX \qquad (2)$$

ous solutions, the equilibrium position lies increasingly to the left as the halogen size increases through the series Cl < Br < I. The reaction can be driven to completion by adding base or removing the halide ion by ion exchange or with precipitants such as silver or mercury salts. HOF is prepared at low temperature from F_2 and ice by the same reaction. Chloride-free solutions of HOCl can be prepared by hydrolysis of its anhydride, Cl_2O, which is in turn obtained by reaction of Cl_2 with mercuric oxide. Hypoiodous acid (HOI) can also been generated by oxidation of iodide ion or by hydrolysis of the polyhalide anion, ICl_2^-; these methods allow preparation of relatively high concentrations of HOI, which is unstable in aqueous solutions. Astatine is a radioactive compound formed during the decay of uranium and thorium. Its short lifetime permits isolation in only vanishingly small quantities, precluding quantitative investigation of its properties. Qualitatively, it appears to behave similarly to iodine. Evidence for formation of hypoastatous acid (HOAt) is that it reacts with weak oxidants and coprecipitates with iodinium dipyridine

$[I(C_6H_5N)_2^+]$ salts, implying an At(I) oxidation state. *See* CHEMICAL EQUILIBRIUM.

Hypohalous acids are weak acids, that is, $HOX \rightleftharpoons H^+ + OX^-$, whose dissociation constants increase in the order HOI < HOBr < HOCl. Both the acids and their conjugate bases are thermodynamically unstable with respect to disproportionation [reactions (3) and (4)]. The decomposition mechanisms are

$$5HOX \longrightarrow 2X_2 + XO_3^- + H^+ + 2H_2O \qquad (3)$$

$$3OX^- \longrightarrow 2X^- + XO_3^- \qquad (4)$$

complex with rates that vary widely, increasing in the order $X = Cl^- < Br^- < I^-$. Hypochlorite solutions (for example, commercial bleach) are stable at room temperature and below for months but decompose at elevated temperatures; hypobromite solutions are stable only at reduced temperatures; and hypoiodite generally decays within minutes of its formation. Decomposition reactions of hypobromous acid (HOBr) and HOI are second-order in HOX and strongly buffer-catalyzed. Textbooks often incorrectly describe OI^- as so unstable that it appears only as a reaction transient with fleeting existence; under favorable medium conditions, it will decay at room temperature with reaction half times exceeding several minutes. *See* BLEACHING.

Reactivity. HOCl and HOBr and their anions are powerful oxidants that react rapidly with a wide range of organic and inorganic reductants. In addition, both acids halogenate aromatic compounds and form halohydrins with unsaturated organic compounds, and N-chloro compound with nitrogen bases. In general, HOCl reacts much more rapidly than OCl^-. Rate constants with a series of inorganic ions have been shown to increase proportionately with their nucleophilicity and, qualitatively, molecules that are good nucleophiles are oxidized much more rapidly than molecules that contain no nucleophilic centers. This high reaction selectivity is also observed in reaction environments as complex as living cells, where electron-rich compounds such as iron-sulfur clusters, metalloporphyrins, nitrogen heterocycles such as the adenosine nucleotides, conjugated polyenes such as carotenes, and biological amines and thiol-containing compounds are preferentially oxidized. This overall reaction behavior is taken to indicate that HOCl and, by inference, the other hypohalous acids react by electrophilic attack of the electron-deficient halogen atom at the nucleophilic center, forming an incipient halogen-nucleophile bond. For example, the reaction between HOCl and SO_3^{2-} is thought to occur by a stepwise mechanism [reactions (5) and (6)], in-

$$HOCl + SO_3^{2-} \longrightarrow ClSO_3^- + OH^- \qquad (5)$$

$$ClSO_3^- + H_2O \longrightarrow SO_4^{2-} + Cl^- + 2H^+ \qquad (6)$$

volving transfer of Cl^+ from HOCl to the sulfite sulfur atom as the initial step. This general mechanism is supported in favorable cases by detection of chlorine-nucleophile intermediates. *See* HALOGENATION; OXIDATION-REDUCTION.

Cytotoxicity. HOCl and HOBr are potent cytotoxins. Hypochlorite was first used as a disinfectant around the beginning of the nineteenth century by the French pharmacist A. G. Labarraque, and has subsequently been widely applied to problems in public sanitation. Boric acid buffered solutions were prominent pharmaceutical agents in the early 1900s but have now been almost completely replaced by antibiotics. Recent studies with bacteria indicate that death is accompanied by disruption of metabolic functions associated with the plasma membrane, including inactivation of the adenosine triphosphate synthase and proteins involved with active transport of metabolites, and (in aerobes) inhibition of respiration. This pattern of reactivity suggests that toxicity arises from disruption of the energy-transducing capabilities of the cell. Remarkably, white blood cells also appear to use hypohalous acids to fight infection. Specifically, neutrophils and monocytes contain the enzyme myeloperoxidase, which catalyzes the oxidation of chloride ion to HOCl by respiration-generated hydrogen peroxide, and eosinophils contain a similar peroxidase which oxidizes bromide ion to HOBr. Many mammalian secretory fluids also contain peroxidases capable of oxidizing halides to their +1 oxidation state; prominent among these enzymes is lactoperoxidase, which in physiological environments is thought to oxidize the pseudohalide thiocyanate (SCN^-) to the hypohalite analog hypothiocyanite (HOSCN). Other chloroperoxidases and bromoperoxidases that catalyze organic halogenations by oxidation of chloride and bromide to Cl(I) and Br(I) are found in fungi and marine organisms. Some of these are being evaluated with respect to potential applications in medicine and as industrial biocatalysts.

James K. Hurst

Bibliography. J. C. Bailar et al. (eds.), *Comprehensive Inorganic Chemistry*, vol. 2, Pergamon Press, Oxford, 1973; J. Everse, K. E. Everse, and M. B. Grisham (eds.), *Peroxidases in Chemistry and Biology*, CRC Press, Boca Raton, FL, 1990; D. W. Johnson and D. W. Margerum, Non-metal redox kinetics: A reexamination of the mechanism of reaction between hypochlorite and nitrite ions, *Inorg. Chem.*, 30:4845–4851, 1991.

Hypothermia

A condition in which the internal temperature of the body is at least 3.6°F (2.0°C) below an internal temperature of 98.6°F (37°C).

Heat Regulation

Hypothermia represents a continuum of effects that vary with the severity of cold on physiological systems. Even though the human body can do without food for a number of weeks or water for a number of days, it needs a specific internal temperature that is regulated on a minute-by-minute basis to maintain all normal body functions. The many physiological and

behavioral processes involved in maintaining the internal temperature constant are called thermoregulation. *See* THERMOREGULATION.

Core and peripheral temperatures. Core and peripheral temperatures are the two major categories of body temperature that are sensed by the brain and spinal cord in order to coordinate various body processes and maintain a constant internal, or core, temperature. The internal temperature is that of the brain, spinal cord, heart, and lungs. It must be maintained within strict limits; otherwise the heart and brain will be compromised, eventually leading to death. This temperature is usually monitored clinically by recording the temperature of the rectum, but a more accurate site would be the esophagus, because the esophagus is anatomically close to the heart, and the temperature of the heart is considered to be the overall average body temperature. Another area that can be used to monitor internal temperature is the eardrum, which is anatomically close to the brain and has the same temperature as the brain.

The peripheral, or shell, temperature is that of the skin and the muscles. It can vary over a much wider range (32–95°F or 0–35°C) than the core temperature. That is possible because the cells of the skin, muscles, and blood vessels are not as sensitive to decreased temperatures as are those of the vital organs. Skin temperatures, such as those recorded in the axilla, are not indicative of an internal temperature.

Physiological control of temperature. Temperature regulation involves the brain and spinal cord, which monitor the difference between the internal and the peripheral temperature, with subsequent psychological and physiological adjustments to maintain a constant internal temperature. The brain records the various body temperatures by means of specialized nerves throughout the body that respond specifically to temperature. These specialized nerve endings, called thermal receptors, are dedicated to responding either to cold or to heat. One theory as to how the system operates holds that the hypothalamus is programmed to maintain a specific temperature (the set point) in the internal parts of the body. If the brain records that the internal temperature is falling relative to the set point, certain automatic reactions occur to minimize or reverse the fall in internal temperature.

Striving for some form of thermal comfort is instinctive. For a human to be comfortable in a thermal environment, a relationship (known as the thermal comfort zone) must be maintained between the core–shell temperature profile and the temperature of the environment. The thermal comfort zone is about 68–72°F (20–23°C). Even on a relatively warm day of 70°F (22°C), the body must keep producing enough heat to minimize the heat transfer that occurs as a result of the difference between the internal temperature and the environment. This process explains how hypothermia can occur even in relatively moderate climates; cases of hypothermia have been reported in the tropics among malnourished individuals. *See* COMFORT TEMPERATURES.

Behavioral and physiological responses to cold stress. The initial response to a cold stress is a behavioral one: Leave the cold environment, button a jacket, or try to decrease body size by curling up into a ball to minimize heat loss. Unlike animals (which grow more fur, deposit fat in anticipation of the winter, or hibernate), humans must maintain a warm internal temperature by means of clothing and by behavior modification. Human response to cold stress is highly variable, however.

Cold stress can be either acute or chronic. People who live in a cold climate usually do not respond to a cold stress as vigorously as those who live in warmer areas. Tolerance to cold temperatures can be developed by constant exposure to the cold, but even in cold environments an internal temperature of 98.6°F (37°C) must be maintained: falling below a critical temperature has dire consequences. Hypothermic individuals as cold as 53.6°F (12°C) have been reported to have survived after medical treatment, however. In contrast with modern, urban populations, some societies whose members have been constantly exposed to the cold, such as the aborigines of Australia, are able to drop their internal temperature a number of degrees to minimize the metabolic stresses required to shiver and stay warm. These adaptations seem to become blunted or lost following a transition to a more contemporary, or urban, existence.

Physical processes influencing heat transfer. In general, urban populations take advantage of certain physical processes of heat transfer (radiation, conduction, convection, and evaporation) in their behavior, selection of clothing materials, and design of cold-weather gear in order to maintain the internal temperature in various climates. These same physical processes explain the body's physiological responses to the cold. In addition, food and exercise are also important in maintaining warmth.

Radiation. For general considerations, the body radiates approximately as much heat as a 70-watt light bulb, and that heat transfer occurs between the skin surface and the environment. Whether heat is exchanged by radiation is determined by the temperature difference between the radiating surfaces, their emissivity, and their surface characteristics. Thus, if a number of people are in a room, the room temperature increases because the people are radiating heat. Conversely, if the surrounding walls are colder than the human skin, the people lose heat. That principle can be demonstrated at night by standing in a room near a large window surface. Even though the room air is warm, the colder surface of the window causes a significant heat transfer away from the body.

Heat loss can be minimized by using the concept of radiation. Usually clothing decreases the temperature difference between the skin and the environment and thus minimizes heat transfer. Thin blankets made of aluminum also minimize radiant heat loss. Such coverings are used by runners after a long race and patients who have undergone heart surgery in cold operating rooms. A decrease in the surface area

of the radiating objects minimizes heat loss, and so curling up into a ball or assuming the fetal position minimizes heat transfer. *See* HEAT RADIATION.

Conduction. Conduction is heat transfer between two thermal objects in physical contact. All substances have thermal conductivity values. If a substance is a poor thermal conductor, it is a good insulator and thus minimizes heat transfer. Fat is an example of a poor conductor: The more fat that is accumulated, the less heat is transferred from the core—and the smaller the chance of becoming hypothermic. In addition to its being a poor thermal conductor and a good insulator, fat is a major energy depot for animals. Hibernating forms therefore gorge themselves before winter arrives to protect against the ravages of the oncoming cold and to accumulate a built-in supply of food. Studies suggest that under certain conditions of cold stress, humans also accumulate fat deposits. *See* ADIPOSE TISSUE; HIBERNATION AND ESTIVATION.

Materials that act as poor conductors, such as wool, feathers, and synthetic fibers, minimize heat transfer. The challenge is to find the right kind of material (conductor) so that the person is simultaneously comfortable and not too hot or cold while exercising or relaxing. An example is the wet suit, which is worn by those who swim in cold water or who work in cold environments and might accidentally fall into cold water. The suits allow a certain amount of water to enter, and it is continually warmed by the heat generated by the activity of the person and the body's metabolism. The layer of warm water between the skin and the suit provides an additional layer of insulation between the skin and the cold water, and thus minimizes heat transfer, and increases dramatically the time a person can stay in cold water.

Because water has a very high rate of conduction, heat transfer occurs faster in cold water than in cold air (see **illus.**). A person surrounded by air at a temperature of 50°F (10°C) experiences a drop in internal temperature that takes place more slowly than a person in 50°F (10°C) water. A nude person in 50°F (10°C) air becomes hypothermic in 24 h, whereas the same person in water of the identical temperature becomes hypothermic in 2 h. *See* CONDUCTION (HEAT).

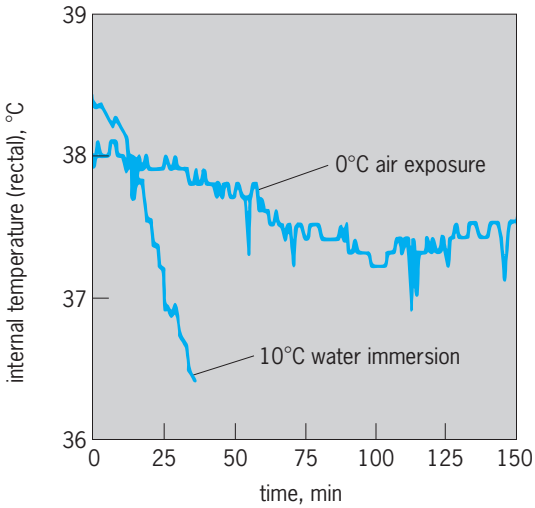

Body heat lost with exposure to 0°C air and 10°C water. °C = (°F − 32)/1.8.

Convection. Convection is a process of heat gain or loss that takes place between two objects in physical contact when one object is in motion. As air moves over a thermal object, heat is transferred to the molecules of the air or fluid. These molecules then move out of contact and are replaced by cooler molecules, which again absorb heat and then move out of contact. This process takes place when riding a motorcycle or jogging, for example. Wind streams across the body and causes a major heat transfer from the body, which cools quickly.

This physical phenomenon is the basis for the term wind chill. Cold, dry wind has severe chilling effects because it increases the rate at which heat and moisture are drawn off the skin. The wind chill index is an attempt to quantify the combined effects of air speed and temperature on heat transfer. Because of the efficiency of convection as a heat transfer process, the cold wind becomes a continual heat drain. *See* CONVECTION (HEAT).

Evaporation. The process of evaporative heat loss can also be important in terms of facilitating heat transfer and predisposing a person toward hypothermia. Evaporation is effective as a heat dissipator since heat from the body is used to evaporate the sweat that is secreted from the body with exercise. If the sweat does not evaporate, it is ineffective in cooling the body and so clothing that does not allow the sweat to evaporate causes the sweat to cling to the body. If this sweat is cooled by cold wind, it cools the body and sets up the conditions that can lead to hypothermia. In addition to heat loss, evaporation promotes a net water loss from the body. Water is critical for maintaining normal body function, especially in the cold.

Although overt sweating occurs during exercise, insensible evaporation takes place when the body is at rest. It serves as a fine control of the body's temperature but can lead to hypothermia since in certain situations, when functioning in conjunction with other body processes, insensible sweating may promote a significant amount of body water and heat loss.

A significant amount of water is lost by way of the lungs, which humidify and heat the inspired air. When cold air is inspired, it cools the blood in the lungs and simultaneously removes some fluid. Considering that the lungs have the surface area of a tennis court, breathing in a cold environment can be a significant factor in inducing hypothermia. *See* EVAPORATION.

Metabolism. In addition to the physical processes that determine heat transfer, the body generates heat as a result of cellular reactions that produce energy. This metabolic heat, which is expressed in terms of internal temperature and which maximizes the chemical reactions of the body, is ultimately derived from the food that is consumed. The caloric value

of food is a quantification of the amount of energy contained in that food. When food or caloric intake is insufficient, the body gets its energy from its own tissues. In extreme cases, the body uses all of its fat and then begins to digest its protein. In a state of malnourishment (for example, in cases of starvation or anorexia nervosa), the body feels cold since it is not taking in enough energy to stay warm. *See* ANOREXIA NERVOSA.

Metabolism is controlled by various hormones, among which are the thyroid hormone and epinephrine. The thyroid hormone increases overall body metabolism. In extreme cases, if the thyroid is not functioning, the body's metabolism is slowed and consequently the body temperature decreases. Epinephrine increases cell metabolism and thereby creates a feeling of warmth. Emotions such as fright or drugs such as amphetamines cause an increase in this hormone and thus increase the body's metabolism and the internal temperature. *See* EPINEPHRINE; THYROID HORMONES.

Exercise also raises internal temperature by increasing metabolism and generating more heat. Shivering, which generates heat by way of muscle contractions that are similar to exercise, represents an effort on the part of the nervous system to keep the body warm by causing involuntary muscle contractions. *See* METABOLISM.

Temperature clock. The physiological factors that affect overall heat transfer are under the control of the nervous system. The brain has one or more internal clocks, which regulate many of the heat-generating processes of the body over a 24-h period. During sleep, the internal temperature drops 0.9°F (0.5°C) and then rises and falls over the sleep period in a programmed manner. The relationship between the temperature clock and personality is a subject of interest in psychiatry, specifically with respect to depression, because a decrease in sleep time influences the clock.

Types of Hypothermia

Various environmental situations predispose humans to hypothermia, which can occur even in the absence of cold. In fact, hypothermia is more common in temperate regions than in the colder climates. Because of the uniqueness of the situations in which hypothermia can occur, various kinds of hypothermia have been classified, all of which can prove fatal.

Primary hypothermia. Primary hypothermia is a decrease in internal temperature that is caused by environmental factors in which the body's physiological processes are normal but thermoregulation capability is overwhelmed by environmental stress.

Air hypothermia. Air (formerly exposure) hypothermia is thought to be the most common form. A person exposed to cold air experiences the same processes as a person in cold water, but air hypothermia occurs more slowly. The induction of air hypothermia is more subtle and therefore more dangerous since it can occur over a number of weeks. People who exercise in cold air and who sweat excessively cool themselves faster than those who do not sweat, because water removes the body's heat. Cold air is a special concern to those who are confined and are unable to walk or exercise. A room with a temperature of 60°F (15.6°C) can become lethal over time for an individual who cannot exercise. On the other hand, mountain climbers, joggers, runners, back-packers, hikers, and victims of natural disasters in which they are exposed to cold are also likely victims of air hypothermia. Mountain climbers represent a subgroup since they have three major stressors: the cold environment, the decrease in oxygen at high elevations, and dehydration.

The degree to which a person reacts to a cold air stress is dependent on such factors as age, physical stamina, the intensity of the cold stress, and the responsiveness of the thermoregulatory system. One of the most convenient ways to determine whether someone is suffering from hypothermia is a noted change in personality: Complaints of fatigue, sluggish speech, and confusion are common, and in some cases the behavior resembles that of intoxication.

Initially, skin temperature falls rapidly, blood vessels to the skin constrict, and shivering begins. After 5–10 min, shivering ceases for about 10–15 min, but this is followed by uncontrollable shivering. The cardiovascular system, specifically the blood vessels to the extremities, circulates blood to satisfy the oxygen demands of the body organs and to distribute body heat. In a cold situation, the nervous system causes the blood to be redistributed away from the skin as the blood vessels of the skin close down to minimize heat transfer to the cold environment. The decrease in skin temperature coupled with vasoconstriction makes the person feel cold, and sometimes the fingers and toes can become painful. Internally, there is an increase in the levels of hormones that control metabolism, and blood is shunted primarily to the lungs, heart, and brain. The person becomes dehydrated as the inspired air is warmed and humidified. If the tense and shivering muscles do not generate enough heat, the hypothermic process begins and progresses for at least 3–5 h. As hypothermia continues, the arms become rigid, and the person loses the ability to make fine movements.

During this period of time the heart rate initially increases, then stabilizes and as the person's internal temperature becomes progressively colder, the heart rate and respiration slow. In severely hypothermic persons, it is very difficult to detect a slow heart rate or determine if the person is breathing. A temperature of 95°F (35°C) is only the beginning of mild hypothermia and shivering can continue for hours, depending on the muscle and fat supplies available. Eventually, the environment becomes overwhelming. At 86°F (30°C), the person loses consciousness and shivering ceases. Death does not occur until the internal temperature drops further: Death results at 68–77°F (20–35°C) because of cardiac standstill.

Immersion hypothermia. When a person falls into cold water, a gasping response is triggered by the thermal

receptors on the skin. For some individuals, the cold stress may trigger a heart attack. Although as much of the body as possible should be kept out of the water, many victims of immersion hypothermia stay in the cold water because they cannot tell how cold they are. Shivering becomes generalized and, unlike its effect in cold air, may in fact cause a faster drop in internal temperature since the water layer closest to the body is stirred and convective heat loss is promoted. As a rule, any movement in cold water promotes heat loss due to heat convection. A person who is unable to come out of the water experiences a rapid drop in internal temperature as a result of movement in the water, the amount of clothing worn, the amount of body fat, and the temperature and state of the water.

The misconception persists that a person can die in 10 min in 50°F (10°C) water. Although the greater conductive property of water relative to air is a major heat sink, physiological and behavioral responses act to minimize the heat loss. Survival in 50°F (10°C) water is possible for several hours at most if the person is dressed in street clothes and a life jacket. Death within minutes may be attributed to drowning, not hypothermia.

In divers' hypothermia, a variation of immersion hypothermia, the person is properly suited and carries an air supply. Even with a wet suit to minimize the cold, cooling of the periphery of the body takes place because the cold water is a heat sink for body heat. In addition, breathing a cold, dry air mixture cools the core of the body, and movement in cold water increases heat loss. Gas mixtures such as helium–oxygen combinations promote an even greater loss of body heat due to their convective heat transfer characteristics. Thus these divers are cooling themselves both internally and externally. By breathing various gas mixtures that cool and dehydrate the core of the body, a drop in internal temperature can precede a drop in peripheral temperature, thus overriding the normal warning signs of approaching hypothermia. *See* DIVING.

Submersion hypothermia. The cooling of the body in this form of hypothermia (actually a situation in which a person drowns in cold water) allows the brain and heart to withstand approximately 45 min of oxygen debt. This is most operative for young children, and people making rescue attempts should always be aware of this. A child can survive for an extended period of time while completely submerged because of the fact that the body is undergoing both internal and external cooling. As the child is drowning, cold water is swallowed and enters the lungs, which cools the core. At the same time, the cold water that bathes the skin rapidly cools the periphery. The multiple effects of the internal and external cooling decrease the metabolic rate and give the child a window of safety of approximately 45 min. In warm water, survival is possible for only 5–7 min.

In addition to the physical factors associated with water cooling, a reflex may play a significant factor in a child's chance for survival. This reflex, called dive reflex, is triggered when cold water bathes the face. This stimulus causes the peripheral blood vessels to constrict, the heart rate to decrease, and the blood to move to the core, giving vital organs additional oxygen. This reflex is quite pronounced in children and, because of its oxygen-conserving results, plays a role in improving the submersed hypothermic victim's chance for survival.

Secondary hypothermia. A decrease in core temperature caused by an underlying pathology that prevents the body from generating enough core heat is referred to as secondary hypothermia. If any of the thermoregulatory systems are altered, the body's ability to generate heat subsequently decreases and hypothermia can then develop without warning. Insufficient muscle mass to generate heat, medications that interfere with metabolism, an underlying systemic infection, decreased thyroid hormone production, and paralysis predispose to hypothermia. Premature infants with low body fat and a large surface-to-volume ratio lose heat rapidly and are at risk for becoming hypothermic. The elderly are perhaps the most susceptible to secondary hypothermia. However, whether the process of aging with no associated debility also alters the thermoregulatory system in the elderly remains to be determined.

Ethanol is often imbibed to ward off the effects of the cold, and this leads to the mistaken belief that it is effective in warming victims of hypothermia. Actually, ethanol causes decreased sensitivity to the cold and results in depressed sensory input from the thermal receptors. Consequently, there is no vasoconstriction or shivering. Many drugs and chemicals affect the thermoregulatory system. Sedatives and antidepressants can decrease a person's ability to sense cold and respond to it, indirectly predisposing the user to hypothermia. Some antidepressants also cause peripheral vasodilation, giving the user a sense of warmth even though the person is losing heat to the environment.

Clinical hypothermia. Some cardiac surgical procedures require clinically induced cooling to stop the heart from beating. Induced hypothermia lowers the oxygen demand of the body tissues, so that oxygenated blood need not circulate. In the case of coronary bypass surgery, the entire body is cooled, enabling the surgeons to work for an extended period of time on the cold heart. During such procedures, the heart is actually the coldest part of the body. Also, cooling the heart cools the blood and thus minimizes hemorrhage.

Frostbite. In hypothermia, the body's internal temperature decreases, but no solid freezing takes place. In frostbite, which is freezing of the digits or the limbs, there is actual formation of ice crystals. Basically the digits go through various stages of cooling. Initially, in the prefreeze phase, the finger temperature is 37.4–50°F (3–10°C). Next, at 24.8°F (−4°C) ice crystals form outside the cells of the digits, circulation is limited, and cell death takes place if the process is allowed to continue. The cells of the digits and limbs can tolerate low temperatures that would

be lethal to brain or nerve cells. However, once they are rewarmed and thawed, they develop an increased sensitivity to the cold and become more susceptible to frostbit Any part of the body can become frostbitten, but the fingers, toes, ears, nose, and cheeks are most often affected.

Cold-induced vasodilation. Associated with cold exposure is a phenomenon that occurs when fingers are placed in freezing water. When the fingers are first put into cold water, the blood vessels constrict and, therefore, the temperature of the fingers falls. After falling to 39.2–42.8°F (4–6°C), the blood vessels reopen and the finger temperature rises again. This sequence of vasoconstriction followed by vasodilation continues as long as the finger remains in cold water. This same phenomenon can also localize in areas such as the toes, fingers, earlobes, and buttocks. The process, which is thought to be protective in that vasodilation does not allow the finger to get too cold, is augmented among those who are often exposed to the cold for long periods of time so that the individuals can spend more time working in the cold.

Rewarming strategies. Even when the body becomes so cold that all function stops, revival with proper rewarming techniques is still possible. In some cases, physiologically dead patients have been revived as long as 45 min after any sign of life could be detected, because hypothermia stops the cells' metabolism and consequently diminishes any cellular demand for oxygen or glucose. Thus, sensitive cells in the brain and heart can survive cold temperatures for a period of time. This period, known as the metabolic icebox, is the primary reason that victims of hypothermia can survive. When a person becomes hypothermic, safe rewarming is the major concern. Since the brain and other vital organs are in a hypothermic state and their oxygen and glucose demands are small, the most important parameters to monitor in hypothermic individuals are heart rate, blood pressure, respiration, and fluid and electrolyte levels. Although rewarming would appear to be the first consideration, monitoring the heart and minimizing any mechanical or chemical damage to it are, in fact, the most critical.

Careful monitoring is important during rewarming. When a person is being rewarmed, the body temperature continues to fall for some time. This phenomenon, called after-drop, may contribute to further manifestations of hypothermia. Because of the effects of hypothermia, cells may lose their physiological or structural integrity, and when the body is warmed too fast, lethal concentrations of some ions, such as potassium, may stop the heart from beating.

For mild hypothermia, physical exercise is recommended because it effectively warms the core. Conversely, a campfire increases heat transfer to the skin but does not substantially warm the core. Similarly body-to-body rewarming warms only the periphery, not the core.

When severe hypothermia renders a person unconscious but with a detectable pulse and respiration, transfer to an emergency center or hospital for warming under the supervision of a physician is recommended. Rewarming in the field is not recommended because of the associated cardiovascular problems. *See* HOMEOSTASIS.

Robert S. Pozos; L. E. Wittmers; James Hodgdon

Bibliography. K. Bowler and B. J. Fuller (eds.), *Temperature and Animal Cells*, 1987; E. Kano (ed.), *Current Researches in Hyperthermia Oncology*, 1988; R. S. Pozos and L. E. Wittmers (eds.), *The Nature and Treatment of Hypothermia*, 1983; J. R. Sutton (ed.), *Hypoxia and Cold*, 1987; J. A. Wilkerson et al. (eds.), *Hypothermia, Frostbite and Other Cold Injuries: Prevention, Recognition, Pre-Hospital Treatment*, 1986.

Hypothesis

A tentative supposition with regard to an unknown state of affairs, the truth of which is thereupon subject to investigation by any available method, either by logical deduction of consequences which may be checked against what is known, or by direct experimental investigation or discovery of facts not hitherto known and suggested by the hypothesis. A classical example is the nebular hypothesis with regard to the origin of the solar system. The "truth" of this hypothesis may be checked by its ability to explain the properties of the solar system as now observed.

Percy W. Bridgman; Henry Margenau

Hypotrichida

An order of Spirotrichia. These protozoa are commonly considered to represent the pinnacle of specialized development in the evolution of all ciliates. Somatic ciliature of the ventral surface has been replaced by cirri; on the dorsal surface it is absent or represented by inconspicuous sensory bristles. The adoral zone of membranelles is very prominently developed, and the buccal area may occupy

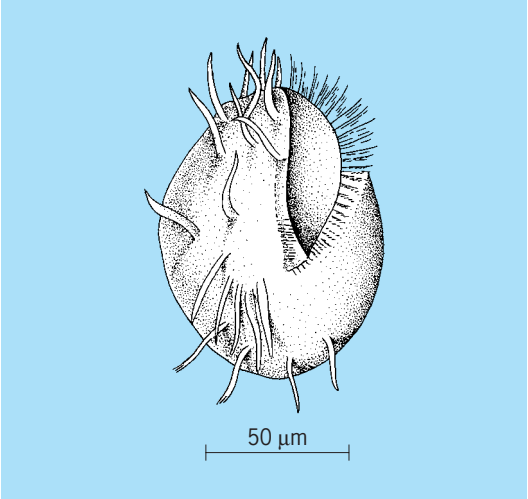

***Euplotes*, an example of Hypotrichida.**

a large part of the ventral surface of the body. The whole body is generally rigid and is pronouncedly flattened dorsoventrally. Hypotrichs occur ubiquitously in fresh- and salt-water habitats and have been studied quite extensively by experimental biologists. Common examples are *Euplotes* (see **illus.**) and *Diophrys*, as well as *Oxytricha* and *Stylonychia*. *See* CILIOPHORA; PROTOZOA; SPIROTRICHIA.

John O. Corliss

Bibliography. S. Anderson and J. K. Jones, Jr., *Recent Mammals of the World*, 1967; W. N. McFarland, *Vertebrate Life*, 1979; A. S. Romer, *Vertebrate Paleontology*, 1966.

Hypoxia

Failure of an adequate amount of oxygen to gain access to, or to be utilized by, the body. When the lack of oxygen is extreme, the term anoxia is used.

Oxygen deprivation may result from interference with inspiration of oxygen, passage of oxygen from the lungs into the blood, transport of oxygen to tissues, and use of oxygen by cells. Interference at any of these stages can produce damage that leads to further hypoxia. This is seen most dramatically when the brain is deprived of the necessary oxygen for more than a few minutes. Nerve cell degeneration begins quickly, and although the original cause of anoxia is removed, damage to the respiratory centers prevents resumption of breathing. *See* RESPIRATORY SYSTEM.

Oxygen may fail to enter the lungs in adequate amounts because of lowered concentrations in the atmosphere or because of obstructions in the airways. High altitudes, drowning, strangulation, aspiration of a foreign body, and lung disease are extrinsic causes of hypoxia. *See* RESPIRATION.

Inadequacy or failure of the respiratory mechanism may result from trauma, poisoning, progressive hypoxia, and other causes.

The passage of oxygen from the lung alveoli to the adjacent blood capillaries may be prevented or decreased by such conditions as chronic lung disease, infections, presence of foreign materials, and developmental defects.

A great number of conditions may interfere with the blood transport of oxygen, most of which is accomplished by the red blood cells. Examples include various forms of anemia, heart disease, trauma, hemorrhage, and circulatory diseases. In brief, anything that decreases blood supply or oxygenation of the blood cells may produce hypoxia. *See* CIRCULATION DISORDERS; RESPIRATORY SYSTEM DISORDERS.

Oxygen transfer from the blood into tissue cells may be prevented by various disturbances. Cellular damage that interferes with oxygen use may result from a host of circumstances, including previous anoxia or hypoxia. The most notable causes of such damage are various microorganisms, noxious chemicals, and physical injury. Failure of cellular respiration is sometimes termed histiotoxic anoxia.

Edward G. Stuart; N. Karle Mottet

Hyracoidea

An order of mammals closely related to elephants. Hyracoids were a diverse and successful group of mammals in Africa during the Eocene and Oligocene epochs (early part of the Age of Mammals, 55 to 34 million years ago), and they are still represented by a few living species. The early hyracoids ranged from animals as small as rabbits to ones as large as modern Sumatran rhinoceroses. The fossil skeletons of the early hyracoids indicate that some species were active runners and leapers, while others were heavy, piglike quadrupeds. Their teeth suggest herbivorous diets, ranging from fibrous leaves in some species to pulpy fruits and roots in others. Hyracoids originated in Africa but later extended their range into Europe and Asia, eventually attaining a distribution encompassed by China, Spain, and South Africa. In contrast to the early diversity of the order, the only living species are a few small-bodied animals (1.5–5.5 kg or 3.3–12 lb) that inhabit forests, scrubby brushlands, and rocky deserts in Africa and the Middle East. Hyracoidea represents a classic case of a spectacular adaptive radiation on an isolated continent, now reduced to a few remnant living taxa. *See* EOCENE; OLIGOCENE.

Classification and phylogeny. The modern hyracoids include about 7 to 12 species classified in three genera, *Procavia, Heterohyrax*, and *Dendrohyrax*. All are members of the family Procaviidae. *Gigantohyrax* is a fourth, extinct genus of Procaviidae known from Plio-Pleistocene caves in southern and eastern Africa. All of the remaining fossil hyracoids, representing at least 18 genera, are often classified in a separate, very diverse family, Pliohyracidae, pending better understanding of the phylogenetic relationships among forms. Important subfamilies include the Geniohyinae, a group of primitive, bunodont (lumpy toothed), piglike forms; the Saghatheriinae, the most abundant and diverse hyracoids of the Eocene and Oligocene; the Titanohyracinae, larger cursorial and graviportal (weight-bearing) hyracoids with teeth resembling those of early perissodactyls; and the Pliohyracinae, an assemblage of robust hyracoids, at least some of which were partially aquatic.

Morphological and molecular lines of evidence indicate that hyracoids are close relatives of elephants (order Proboscidea) and manatees and dugongs (order Sirenia). Hyracoids and elephants share several peculiar specializations, including continuously growing upper incisors, or tusks; mammary glands that are located on the chest rather than low on the abdomen; testes that remain in the body cavity rather than descending during development into a scrotum; and a peculiar linear arrangement of the wrist and foot bones (taxeopody). Molecular studies of blood and eye lens proteins further support a phylogenetic relationship among these three orders. Therefore, the orders Hyracoidea, Proboscidea, and Sirenia are best classified together in the superorder Paenungulata. *See* PROBOSCIDEA; SIRENIA.

Fossil record. Among the earliest hyracoids are *Seggeurius* and *Microhyrax*, both known only by

jaws from the middle Eocene of Algeria (about 50 million years old). *Seggeurius* is a small geniohyine that may retain the most primitive dental morphology known within the order. *Microhyrax* is an even smaller hyracoid with crestier teeth than *Seggeurius*. Both are poorly known. Much more complete fossil material, including numerous jaws, crania, and limb bones, are known for three Eocene saghatheriines, *Saghatherium, Thyrohyrax*, and *Megalohyrax*. These genera differ from modern hyraxes in having long snouts, complete eutherian dentitions, and distinctive patterns of sexual dimorphism. Small *Thyrohyrax* and large *Megalohyrax* had a hollow chamber inside the mandible of one sex (perhaps the females), while the other sex did not. The hollow chamber is not understood but may have been a resonating device for loud vocalizations. Both sexes of *Saghatherium* lacked a chamber, but the males were much larger in size than the females. These patterns of sexual dimorphism hint that fossil hyracoids must have had complex social systems and behaviors. Another well known genus is *Titanohyrax*, a very large cursorial (running) hyracoid with teeth that indicate a leafy diet.

The best fossil record of early hyracoids comes from the Fayum, Egypt, where eight genera have been found in upper Eocene and lower Oligocene sediments (36–32 million years ago). After the early Oligocene, the diversity and importance of hyracoids began to decline, possibly due to competition with the newly arriving rhinos, bovids, and other groups migrating into Africa from the northern continents. By the early Miocene, only a few species are known from East Africa, but two evolutionary radiations of hyracoids were still to come. The distinct subfamily Pliohyracinae differentiated in Africa by the middle Miocene (*Parapliohyrax* and *Prohyrax*) and later spread to Eurasia. Some of the heavily built, aquatic pliohyracines survived until the Plio-Pleistocene in Europe (*Pliohyrax*) and China (*Postschizotherium*). These large pliohyracines had extremely hypsodont (high-crowned) teeth, indicating a very tough, fibrous plant diet. The modern family Procaviidae, characterized by relatively small size, reduced dentition (loss of canines and sometimes other teeth), and the development of a large gap between incisors and premolars, first appeared in the late Miocene of Namibia (*Heterohyrax auricampensis*). An abundance of Pleistocene procaviids is found throughout southern and eastern Africa, in habitats that are reconstructed by paleontologists as being similar to modern arid savannas and scrublands.

Habits and distribution. *Heterohyrax brucei* (bush hyrax) and species of *Procavia* (rock hyraxes) (see **illustration**) inhabit scrubby arid country in eastern Africa, from low desert to mountainous, alpine habitats. In addition, the range of *Procavia* extends into the deserts of northern and southern Africa, and the Middle East. *Procavia* and *Heterohyrax* prefer rocky outcrops, or kopjes, where they can hide in cracks and crevices, or bask on sunlit boulders. The elastic soles of their feet and the linear alignment of their wrist bones allow them to scramble and leap

Cape hyrax (*Procavia capensis*). (*Photo by Lloyd Glenn Ingles; © 1999 California Academy of Sciences*)

nimbly among the rocks. They live in colonies numbering from just a few to up to 40 individuals. A colony's territorial male, who is larger than the females, defends a rocky range inhabited by several females and immatures. Warning vocalizations, made loud with the help of an expanded chamber in the eustachian tube (the connection between throat and ear), alert the colony to predators (hyraxes are among the common prey of several carnivores and large birds). Female hyracoids produce few offspring (one to four per annual litter) which develop slowly for mammals of their small size, taking more than a year to reach sexual maturity. *Procavia* is a grazer that consumes a wide range of grasses and other coarse plant foods, an adaptation reflected by its high-crowned molar teeth, which are reminiscent of miniature rhino molars. Hyraxes do not ruminate, but they do have complex guts and utilize the help of gut microbes in digesting their plant food. As a consequence, they are able to digest some plants that are toxic to other mammals. *Heterohyrax* differs from *Procavia* in having a more general diet of softer leaves, grasses, and other plant parts. In contrast to its open country relatives, *Dendrohyrax* is typically an arboreal climber in equatorial tropical forests. It does not live in concentrated colonies but is more evenly distributed throughout the forest, communicating socially by loud calls at night. Because of the destruction of African forests, *Dendrohyrax* is the most threatened of the living hyraxes. *See* DENTITION; EUTHERIA; MAMMALIA. D. Tab Rasmussen

Bibliography. J. Kingdon, *East African Mammals: An Atlas of Evolution in Africa*, University of Chicago Press, 1974; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford, 1984; R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999; D. T. Rasmussen, The evolution of Hyracoidea: A review of the fossil evidence, in D. R. Prothero and R. M. Schoch (eds.), *The Evolution of Perissodactyls*, Oxford University Press, 1989.

Hysteresis

A phenomenon wherein two (or more) physical quantities bear a relationship which depends on prior history. More specifically, the response Y takes

on different values for an increasing input X than for a decreasing X.

If one cycles X over an appropriate range, the plot of Y versus X gives a closed curve which is referred to as the hysteresis loop. The response Y appears to be lagging the input X.

Hysteresis occurs in many fields of science. Perhaps the primary example is of magnetic materials where the input variable H (magnetic field) and response variable B (magnetic induction) are traditionally chosen. For such a choice of conjugate variables, the area of the hysteresis loop takes on a special significance, namely the conversion of energy per unit volume to heat per cycle. In a similar way the dielectric energy loss per unit volume per cycle can be obtained by using the variables E (electric field) and D (electric displacement). For mechanical hysteresis, it is customary to take the variables stress and strain (see **illus.**), where the energy density loss per cycle is related to the internal friction. In the framework of particle rather than continuum mechanics, the energy loss per cycle is given by the hysteresis loop for the applied force and the particle displacement. Thermal hysteresis is characteristic of many systems, particularly those involving phase changes, but here the hysteresis loops are not usually related to energy loss. *See* FERROELECTRICS; STRESS AND STRAIN; THERMAL HYSTERESIS.

Hysteresis is intimately associated with microscopic irreversibility. In the electric and magnetic situations the motion of domain walls in ordered structures is impeded by microscopic inhomogeneities in the matrix which catch and then let go under increasing stress. Such effects are clearly irreversible and in the magnetic case are evident as the Barkhausen noise. On the other hand, any domain rotation is largely reversible and does not contribute to the hysteresis loss. In the mechanical situation, the pinning and unpinning of dislocations gives rise to the hysteresis. *See* BARKHAUSEN EFFECT; CRYSTAL DEFECTS; DOMAIN (ELECTRICITY AND MAGNETISM).

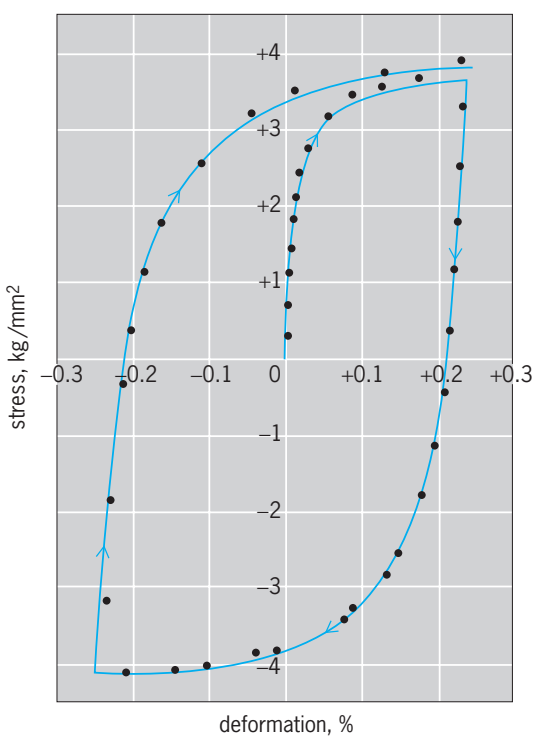

Plot of stress versus strain (deformation) for a single crystal of brass. The positive and negative strains correspond to elongation and compression, respectively. (*After G. Sachs and H. Shoji, Zug-Druckversuche an Messing-Kristallen (Bauschinger effect), Z. Physik, 45:776–796, 1927*)

Time-dependent effects, where the final equilibrium value of the history-independent response Y occurs some time after the input X is applied, are complex and are not usually referred to as hysteresis effects. However, cycled experiments (resonance and relaxation) whose frequencies are comparable with the natural frequencies of the system will display an X-versus-Y dependence of the hysteretic form. The largest values of the hysteresis area will occur when the applied frequency coincides closely with natural frequencies of resonance or relaxation. H. B. Huntington; R. K. MacCrone

Hysteresis motor

A type of synchronous motor in which the rotor consists of a central nonmagnetic core upon which are mounted rings of magnetically hard material. The rings form a thin cylindrical shell of material with a high degree of magnetic hysteresis. The cylindrical stator structure is identical to that of conventional induction or synchronous motors and is fitted with a three-phase or a single-phase winding, with an auxiliary winding and series capacitor for single-phase operation. *See* INDUCTION MOTOR; SYNCHRONOUS MOTOR.

When the motor is running at synchronous speed, the hysteresis material is in a constant state of magnetization and acts as a permanent magnet. Full-speed performance is therefore exactly the same as in a permanent-magnet synchronous motor.

The outstanding special feature of a hysteresis motor is the production of nearly constant, ripple-free torque during starting. At this time the stator current is large, and the rotor hysteresis material is continuously cycled around its hysteresis loop. The resulting rotor flux is delayed in time (and hence in space) behind the rotating magnetomotive force (mmf) produced by the stator currents. This spatial angle between the stator mmf and rotor flux determines the starting torque. Because the angle depends only on the hysteresis loop shape and not on the rate at which the loop is traversed, the starting torque is nearly independent of rotor speed. When synchronous speed is attained, the rotor flux is synchronous with the stator mmf and the hysteresis material retains its last operating point as a constant state of magnetization, resulting in operation as a permanent-magnet synchronous motor. The stator current falls abruptly to its normal operating levels as the machine reaches synchronous speed.

Hysteresis motors are widely used in synchronous motor applications where very smooth starting is

required, such as in clocks and other timing devices and record-player turntables, where smooth starting torque reduces record slippage. Hysteresis motors are limited to small size by the difficulty of controlling rotor losses caused by imperfections in the stator mmf wave. *See* ALTERNATING-CURRENT MOTOR; MOTOR.

Donald W. Novotny

Bibliography. P. L. Cochran, *Polyphase Induction Motors: Analysis, Design, and Application*, 1989; S. A. Nasar, *Permanent Magnet, Reluctance and Self-Synchronous Motors*, 1993; S. A. Nasar and L. E. Unnewehr, *Electromechanics and Electric Machines*, 1979.

Hysteriales (lichenized)

An order of the Ascolichenes shared by the Ascomycetes. This order formerly included all lichens with linear elongate ascocarps called hysterothecia. But species with true paraphyses are now classified in the family Graphidaceae of the Lecanorales, leaving in the lichenized Hysteriales only those species with a so-called ascolocular structure. *See* LECANORALES.

The hymenium consists of densely interwoven, branched pseudoparaphyses and irregularly scattered locules in which the asci are located. At times it is difficult to distinguish branched paraphyses from pseudoparaphyses, and a few genera with hysterothecia are transitional between the Hysteriales and the Graphidaceae.

The growth form of the Hysteriales is either crustose or fruticose. There are five or six families now included here. Arthoniaceae is the largest family, with two genera, *Arthonia* and *Arthothelium*, both widespread on bark, rarely on leaves, in temperate and tropical regions. The ascocarps are exceedingly small and irregularly branched. The Opegraphaceae includes four major genera with larger elongate ascocarps, crustose on bark and rocks. The Roccellaceae is a family of conspicuous fruticose species that grow profusely on trees and rocks along the coastlines of Portugal, California, Baja California, and parts of western South America. The ascocarps are long and linear. These plants were collected by the ton in the Middle Ages for use as dyestuffs. These same lichens, on a smaller scale, now yield the dye used in litmus paper. *See* ASCOMYCOTA.

Mason E. Hale

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

9 I - LEV

www.MHEST.com

I

I–spin — Ixodides

I-spin

A quantum-mechanical variable or quantum number applied to quarks and their compounds, the strongly interacting fundamental hadrons, and the compounds of those hadrons (such as nuclear states) to facilitate consideration of the consequences of the charge independence of the strong (nuclear) forces. This variable is also known as isotopic spin, isobaric spin, and isospin.

The many strongly interacting particles (hadrons) and the compounds of these particles, such as nuclei, are observed to form sets or multiplets such that the members of the multiplet differ in their electric charge and magnetic moments, and other electromagnetic properties but are otherwise almost identical. For example, the neutron and proton, with electric charges that are zero and plus one (in units of the magnitude of the electron charge), form a set of two such states. The pions, one with a unit of positive charge, one with zero charge, and one with a unit of negative charge, form a set of three. It appears that if the effects of electromagnetic forces and the closely related weak nuclear forces (responsible for beta decay) are neglected, leaving only the strong forces effective, the different members of such a multiplet are equivalent and cannot be distinguished in their strong interactions. The strong interactions are thus independent of the different electric charges held by different members of the set; they are charge-independent. *See* ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS; HADRON; STRONG NUCLEAR INTERACTIONS.

The I-spin (I) of such a set or multiplet of equivalent states is defined such that Eq. (1) is satisfied,

$$N = 2I + 1 \tag{1}$$

where N is the number of states in the set. Another quantum number I_3, called the third component of I-spin, is used to differentiate the numbers of a multiplet where the values of I_3 vary from $+I$ to $-I$ in units of one. The charge Q of a state and the value of I_3 for this state are connected by the Gell-Mann-Okubo relation, Eq. (2), where Y, the charge offset, is called

$$Q = I_3 + \frac{Y}{2} \tag{2}$$

hypercharge. For nuclear states, Y is simply the number of nucleons. Electric charge is conserved in all interactions; Y is observed to be conserved by the strong forces so that I_3 is conserved in the course of interactions mediated by the strong forces. *See* HYPERCHARGE.

Similarity to spin. This description of a multiplet of states with I-spin is similar to the quantum-mechanical description of a particle with a total angular momentum or spin of j (in units of $\hbar$, Planck's constant divided by 2π). Such a particle can be considered as a set of states which differ in their orientation or component of spin j_z in a z direction of quantization. There are $2j + 1$ such states, where j_z varies from $-j$ to $+j$ in steps of one unit. To the extent that the local universe is isotropic (or there are no external forces on the states that depend upon direction), the components of angular momentum in any direction are conserved, and states with different values of j_z are dynamically equivalent.

There is then a logical or mathematical equivalence between the descriptions of (1) a multiplet of states of definite I and different values of I_3 with respect to charge-independent forces and (2) a multiplet of states of a particle with a definite spin j and different values of j_z with respect to direction-independent forces. In each case, the members of the multiplet with different values of the conserved quantity I_3 on the one hand and j_z on the other are dynamically equivalent; that is, they are indistinguishable by any application of the forces in question. *See* ANGULAR MOMENTUM; SPIN (QUANTUM MECHANICS).

Relative intensities of virtual transitions determined by isobaric spin symmetry

Transition	Relative intensity
$p \rightarrow n + \pi^+$	2/3
$p \rightarrow p + \pi^0$	1/3
$n \rightarrow n + \pi^0$	1/3
$n \rightarrow p + \pi^-$	2/3

Importance in reactions and decays. The charge independence of the strong interactions has important consequences, defining the intensity ratios of different charge states produced in those particle reactions and decays which are mediated by the strong interactions. A simple illustration is provided by the virtual transitions of a nucleon to a nucleon plus a pion, transitions which are of dominant importance in any consideration of nuclear forces. The neutron and proton form a nucleon I-spin doublet with $I = {}^1/_2$; the pions form an I-spin triplet with $I = 1$. If initially there are one neutron and one proton, with no bias in charge state or I_3, and the strong forces responsible for the virtual transitions do not discriminate between states with different charge or different values of I_3, then it follows that in the final system there must be equal probabilities of finding each charge member of the pion triplet and equal probabilities of finding a neutron or proton. This condition, that the strong interactions cannot differentiate among the members of an isobaric spin multiplet, determines the relative intensities of the transitions (see **table**).

Using the same kind of argument, it is easy to see that the conditions of equal intensity of each member of a multiplet cannot be fulfilled in a transition from an initial doublet to a final state of a doublet and a quartet. Therefore, none of the individual transitions is allowed by charge independence, though charge or I_3 is conserved in the decays. In general, decays are allowed for a transition $A \rightarrow B + C$ only if inequality (3) is satisfied. This is analogous to the

$$|I(B) + I(C)| \geq |I(A)| \geq |I(B) - I(C)| \qquad (3)$$

vector addition rule for spin or angular momentum; the strong interactions conserve I-spin in the same manner as angular momentum is conserved. *See* SELECTION RULES (PHYSICS).

Classification of states. I-spin considerations provide insight into the total energies or masses of nuclear and particle states. The fundamental constituents of nuclei are the nucleons, the neutron and proton–spin-${}^1/_2$ fermions which must obey the Pauli exclusion principle to the effect that the sign of the wave function that describes a set of identical fermions must change sign upon exchange of any two fermions. Similarly, hadrons are described as compounds of quarks, which are also spin-${}^1/_2$ fermions. The two fermions that make up an I-spin doublet can be considered as different charge states of a basic fermion, even as states with the spin in the plus and minus direction of quantization are considered as different spin states of the fermion. The extended Pauli exclusion principle then requires that the wave function amplitude change sign upon exchange of spin direction, charge, and spatial coordinates for two (otherwise) identical fermions. *See* EXCLUSION PRINCIPLE.

A space state $u(r)$ of two identical fermions, where r is the vector distance between the two particles, will be even upon exchange of the two particles if $u(r)$ has an even number of nodes, and will be odd under exchange if there is an odd number of nodes. With more nodes, the space wavelength is smaller, and the momentum and energy of the particles are larger. The lowest energy state must then have no spatial nodes and must be even under spatial interchange. From the Pauli principle, the whole wave function must be odd, and then the exchange under spin and I−spin coordinates must be odd. Using this kind of argument, the low-mass (low-energy) states of light nuclei can be classified in terms of their I-spin symmetries.

An application of the same principle, that the space wave function of the lowest state must be even under exchange, was an important element in the discovery of a new quantum number (labeled color) for quarks, the elementary constituent of the hadrons, and concomitantly the unfolding of a deeper understanding of fundamental particles.

Basis for charge independence. The basis for the symmetry described by I-spin is to be found in the quark structure of the strongly interacting particles and the character of the interactions between quarks. All of the strongly interacting particles are quark compounds; the conserved fermions, such as the neutron and proton, are made up of three quarks; the bosons, such as the π mesons, are quark-antiquark pairs. There are six significantly different quarks arranged in three pairs of charge $+{}^2/_3$ and charge $-{}^1/_3$ particles, ($u^{2/3}$, $d^{-1/3}$), ($c^{2/3}$, $s^{-1/3}$), and ($t^{2/3}$, $b^{-1/3}$), called up and down, charm and strange, and, top and bottom. The quarks interact through their strong interacting color charges that couple the quarks to gluons in a manner analogous to the coupling of electrical charge to photons. Rather than one kind of electrical charge, and its negative, quarks have three kinds of color charges (conventionally labeled r for red, y for yellow, and b for blue), and their negatives. Even as there are three different colors, there are $3 \times 3 - 1 = 8$ different gluons labeled with color and anticolor (with the so-called white combination ruled out), rather than the one photon that carries the electromagnetic forces.

Since each kind of quark carries exactly the same sets of color charge, the strong forces between two quarks are exactly the same as the forces between two other quarks. However, the simple consequences of this color independence of the strong forces are largely obviated by the fact that the six quarks have different masses. Hence, the Σ^+ hyperon, (uus), is appreciably heavier than the proton, (uud), even as the s-quark is heavier than the d-quark, though the quark-quark forces holding the systems together are the same.

However, the masses of the two lightest quarks, the *u* and *d*, are almost the same, differing by only a few electron masses. Then, for the many situations that this small difference can be neglected, the *u* and *d* quarks cannot be differentiated by the strong forces; that is, the neutron (*duu*) and proton (*ddu*) are identical under the strong forces, which is just the symmetry described by I-spin.

There are some effects where the mass difference, equal to about two pion masses, between the *s* and *u* quark, or the *s* and *d* quark, can be neglected, or for which corrections can be calculated. For such effects twofold symmetries like I−spin, called, respectively, V-spin and U-spin, are useful. In principle, similar broken symmetries exist for compounds of the heavier quarks, but the symmetry breaking that follows from their much larger mass differences so obscures the underlying symmetry that it is not useful. *See* COLOR (QUANTUM MECHANICS); GLUONS; QUANTUM CHROMODYNAMICS; QUARKS; SYMMETRY LAWS (PHYSICS).

Robert K. Adair

Bibliography. F. Halzen and A. Martin, *Quarks and Leptons*, 1984; N. A. Jelley, *Fundamentals of Nuclear Physics*, 1990; G. Kane, *Modern Elementary Particle Physics*, updated ed., 1993; L. B. Okun, *Leptons and Quarks*, 1980, paper 1985; S. S. Wong, *Introductory Nuclear Physics*, 2d ed., 1999.

Ice cream

A commercial dairy food made by freezing while stirring a pasteurized mix of suitable ingredients. The product may include milk fat, nonfat milk solids, or milk-derived ingredients; other ingredients may include corn syrup, water, flavoring, egg products, stabilizers, emulsifiers, and other non-milk-derived ingredients. Air incorporated during the freezing process is also an important component.

The structure of ice cream is complex. It consists of solid, gaseous, and liquid phases; ice crystals and air cells are dispersed throughout the liquid phase, which also contains fat globules, milk proteins, and other materials.

Composition. In the United States, ice cream composition is regulated by Federal Frozen Dessert Standards of Identity, which set forth minimum composition requirements of not less than 10% milk fat and not less than 20% total milk solids. Ice cream must weigh not less than 4.5 lb/gal (0.54 kg/liter) and must contain not less than 1.6 lb food solids/gal (0.26 kg/liter). In the case of bulky flavors, the fat may be not less than 8%, nor can the total milk solids be less than 16% in the finished food. Ingredient and nutritional-requirements labeling is included in the standards.

The composition of ice cream may vary depending on whether it is an economy brand satisfying minimal requirements, a trade brand of average composition, or a premium brand of superior composition. The components by weight of an average-composition ice cream are 12% fat, 11% nonfat milk solids, 15% sugar, and 0.3% vegetable gum stabilizer.

An average serving of 4 fl oz (120 ml) of vanilla ice cream with 10% milk fat provides about 135 calories, 0.4 oz (11.2 g) fat, 0.8 oz (22.4 g) protein, 0.56 oz (15.9 g) carbohydrate, 0.003 oz (88 mg) calcium, 0.0024 oz (67 mg) phosphorus, 270 International Units vitamin A, and 0.0000058 oz (0.164 mg) riboflavin. Thus, a serving of ice cream provides more than 10% of the U.S. Recommended Daily Dietary Allowance (USRDA) for riboflavin and calcium and about 5% of the USRDA for protein for adults and teenagers. *See* NUTRITION.

French ice cream may contain a relatively high fat content, have a slight yellow color, and contain egg yolk solids. Both this product and ice cream carry the same requirements. French ice cream must contain not less than 1.4% egg yolk solids for plain flavor and not less than 1.12% for bulky flavors. It is usually sold as a soft-serve product but may be sold as a prepackaged hardened product. Gelato is Italian for ice cream; it commonly increases in volume by only one-third on freezing.

The composition of ice milk is similar to that of ice cream except that it must contain more than 2% but not more than 7% milk fat, not less than 11% total milk solids, and not less than 1.3 lb food solids/gal (0.15 kg/liter).

Sherbet is made with about 1 part ice cream mix to 4 parts water ice mix and is manufactured in the same way as ice cream. It must weigh not less than 6 lb/gal (0.7 kg/liter). Sherbet contains not less than 1% or more than 2% fat, and total milk solids of not less than 2% or more than 5% by weight of the finished food; it must also contain 0.35% edible citric or natural fruit acid. A sherbet made with addition of egg yolk is known as frappe.

Frozen dairy desserts. Products classed as frozen dairy desserts include French ice cream, frozen custard, ice milk, sherbet, water ice, frozen dairy confection, dietary frozen dairy dessert, and mellorine (imitation ice cream). Most of these desserts can be sold in the soft-serve form, as they are when first removed from the freezer at about 20°F (−7°C), or in the hard form as they are when the temperatures are reduced to about 8°F (−13°C). A frozen dessert mix is the combination of ingredients, usually in the liquid form and ready to be frozen. Dry mixes must be dissolved in water before freezing. Unflavored mixes may be flavored before freezing, as with vanilla and chocolate, or may have flavorings added to the soft frozen product as it exits the freezer, as with fruit, nut, and candy type desserts.

Water ice is made from water, sugar, corn sweeteners, fruit juices, flavorings, and stabilizers. It contains no milk solids. Sorbet is similar; in addition it contains finely divided fruit and is whipped while being frozen. There is no federal standard for sorbet in the United States.

Dietary frozen desserts are comparatively low in calories. Ice milk commonly contains about one-third the fat of ice cream. Proposed U.S. Federal Standards permit 2–5% fat in ice milk and establish a category of light (lite) ice cream with 5–7% fat. To provide frozen desserts for diabetics, sugar (sucrose)

is replaced with nonnutritive sweeteners, and complex carbohydrates, such as sorbitol, are added to provide desirable texture.

Mellorine is manufactured in a process similar to that for ice cream, but is not limited to milk-derived nonfat solids and may be composed of animal or vegetable fat or both, only part of which may be milk fat. It contains not less than 6% fat and not less than 2.7% protein by weight.

Frozen dairy confections are produced in the form of individual servings and are commonly referred to as novelties, including bars, stick items, and sandwiches. These constitute an important segment of the industry.

Soft frozen dairy products make up about 9.6% of the total frozen desserts produced in the United States. Per capita consumption is 5 qt (4.7 liters) per year. The soft frozen product is usually drawn from the freezer at 18 to 20°F (−8 to −7°C) and is served directly to the consumer. Ice milk accounts for about three-fourths of the soft-serve products that are manufactured in the United States. These products usually vary from 2 to 7% fat, 11 to 14% nonfat milk solids, 13 to 18% sugar, and have about 0.4% stabilizer. Frozen custard and ice cream are also important soft-serve products.

Commercial manufacture. In ice cream manufacture the basic ingredients are blended together (see **illus.**). The process ranges from small-batch operations, in which the ingredients are weighed or measured by hand, to large automated operations, where the ingredients, all in liquid form, are metered into the mix-making equipment. The liquid materials, including milk, cream, concentrated milk, liquid sugar syrup, and water, are mixed. The dry solids, such as nonfat dry milk, dried egg yolk, stabilizer, and emulsifier, are blended with the liquid ingredients. This liquid blend is known as the mix. Following the blending operation, the mix is pasteurized, homogenized, cooled, and aged. Pasteurization destroys all harmful microorganisms and improves the storage properties of the ice cream. *See* PASTEURIZATION.

The hot mix is pumped from the pasteurizer through a homogenizer, which reduces the fat globules to less than 2 micrometers in diameter. Homogenization is accomplished by forcing the mix through the homogenization valves under pressure in the range 1500–3000 lb/in.^{2} (103–207 kilopascals). This phase of the manufacturing process results in a uniform, smooth product and limits churning of the fat during the freezing process.

The mix is cooled immediately to 32–40°F (0–4°C); then it is aged 4–12 h in order to permit development of favorable hydration and particle orientation effects, which serve to improve the body and texture in the finished ice cream.

Soluble flavoring materials are usually added to the mix just before the freezing process, but fruits, nuts, and candies are not added until the ice cream is discharged from the freezer. In the United States, popular flavors of ice cream include vanilla, chocolate, strawberry, butter pecan, and "cookies 'n cream" (broken cookies are blended into the ice cream), and there are many other flavors. Neapolitan is one-third each of vanilla, chocolate, and strawberry extruded simultaneously into one container.

Ice cream is frozen in batch or continuous

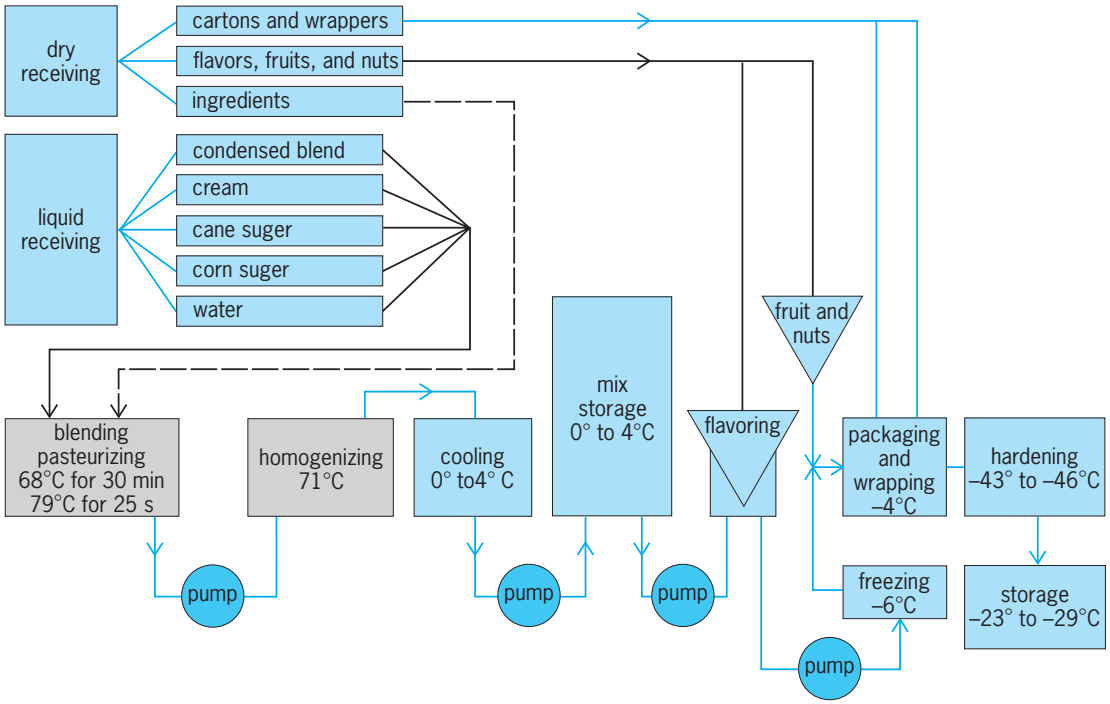

Flow chart for ice cream processing. °F = (°C × 1.8) + 32. (*Arbuckle and Co.*)

freezers. The continuous freezing process is more effective, as it freezes more rapidly and produces a product with finer texture. Continuous freezers vary in capacity from 85 gal (323 liters) to more than 2000 gal (7600 liters) per hour.

During freezing, air is incorporated into the mix, resulting in increased volume. This increase in volume is called overrun. The drawing temperature of the ice cream from the freezer is about 21°F (−6°C). The ice cream is packaged at this temperature and cooled further to about −20°F (−29°C) in a hardening room, where it is stored until marketed.

The ice cream industry of the United States has been described as a high-volume, highly automated, progressive, very competitive industry composed of large and small factories. It has developed on the basis of abundant and economical supply of ingredients. Also contributing to the growth of the industry are factors such as achievement of high quality; favorable consumer response; excellent transportation, refrigeration, and distribution; advanced packaging; and abundant home refrigeration.

The annual per capita production in the United States is about 15 qt (14 liters) of ice cream. Combined production of ice cream, ice milk, sherbet, and mellorine amounts to about 22 qt (21 liters) per person. This amounts to 1.32 billion gal (5 billion liters) per year of ice cream and related products.

Wendell S. Arbuckle; Robert T. Marshall

Sources of microorganisms. The cream, condensed- and dry-milk products, egg products, sugar, stabilizers, flavoring and coloring materials, and fruits and nuts used as ingredients in ice cream may be contaminated with microorganisms. Therefore the ice cream mix must be pasteurized before freezing because a low temperature in itself is not sufficient to kill microorganisms. Other possible sources of microorganisms are equipment and utensils used in manufacture, scoops and dippers, vendors, employees, and air. Stringent regulations must be established for all manipulations of ice cream from producer to consumer in order to avoid recontamination, especially with regard to pathogens. Once incorporated, microorganisms survive a long time in ice cream.

Kinds of microorganisms. A wide variety of microorganisms may be found, including streptococci, micrococci, coliform bacteria, spore-formers, yeasts, and molds. No selection of species by freezing takes place. Since no growth and only slow destruction of microorganisms in the final product occur, the flora of ice cream consists of organisms surviving pasteurization or incorporated by contamination. With well-controlled sanitation during production, plate counts may be as low as a few hundred; with careless methods they may be in the millions. Normally a slight increase of the bacterial count takes place during manufacture, especially after homogenization and freezing, due to the breaking up of clumps of bacteria. A plate count not greater than 50,000–100,000 per gram is regarded as acceptable from a sanitary point of view in most states.

Hygienic measures. These include proper pasteurization of the mix and adequate sanitizing of equipment. Pasteurization must be more intensive with the mix than with milk because increased milk solids and sugar protect microorganisms to some degree. At least 155°F (68.4°C) for 30 min or 175°F (79.5°C) for 25 s is applied to obtain a sufficient margin of safety. In this way all pathogens are killed and total bacterial counts of 10,000 or less per gram are usually reached. Higher counts after pasteurization are due to thermoduric organisms. The source of these organisms is inadequately cleaned farm or plant equipment. Sanitizing of equipment is achieved by cleaning and disinfecting operations with heat or chemicals. *See* DAIRY MACHINERY; STERILIZATION.

Bacteriological control. For determining the hygienic quality of ice cream, the laboratory tests used are similar to those employed for milk. Of these, the standard plate count and the test for coliform bacteria have found widespread application. *See* FOOD MANUFACTURING; MILK.

William C. Winder

Bibliography. W. S. Arbuckle, *Ice Cream*, 5th ed., 1996; International Ice Cream Association, *The Latest Scoop*: *Facts and Figures on Ice Cream and Related Products*, annually; R. K. Robinson (ed.), *Modern Dairy Technology*, 2d ed., 1993.

Ice field

A network of interconnected glaciers or ice streams, with common source area or areas, in contrast to ice sheets and ice caps. The German word *Eisstromnetz*, which translates literally to ice-stream net, is sometimes used for glacial systems of moderate size (such as less than 3000 mi^2 or 8000 km^2) and is most applicable to mountainous regions. Being generally associated with terrane of substantial relief, ice-field glaciers are mostly of the broad-basin, cirque, and mountain-valley type. Thus, different sections of an ice field are often separated by linear ranges, bedrock ridges, and nunataks.

Contrast with ice sheet. An ice sheet is a broad, cakelike glacial mass with a relatively flat surface and gentle relief. Ice sheets are not confined or controlled by valley topography and usually cover broad topographic features such as a continental plateau (for example, much of the Antarctic ice sheet), or a lowland polar archipelago (such as the Greenland ice sheet). Although ice sheets are generally of very large dimension, in some regions small, rather flat ice bodies have been called ice sheets because they are thinned remnants of once large masses of this form. Small ice sheets and even ice fields are sometimes incorrectly referred to as ice caps, even though their configurations have been well characterized.

Contrast with ice cap. Ice caps are properly defined as domelike glacial masses, usually at high elevation. They may, for example, make up the central nourishment area of an ice field at the crest of a mountain range, or they may exist in isolated positions as

separate glacial units in themselves. The latter type is characterized by a distinctly convex summit dome, bordered by contiguous glacial slopes with relatively regular margins not dissected by outlet valleys or abutment ridges.

Similarities and gradations. There are all gradations between ice caps, ice fields, and ice sheets. Over a period of time, a morphogenetic gradational sequence may also develop in any one region. Major ice sheets, for example, probably originate from the thickening and expansion of ice fields and the coalescence of bordering piedmont glaciers. Conversely, ice fields can develop through the thinning and retraction of a large ice sheet overlying mountainous terrane. *See* GLACIATED TERRAIN; GLACIOLOGY.

Maynard M. Miller

Ice manufacture

Commercial production of manufactured artificial ice from water, or of dry ice from the solidification of carbon dioxide, by mechanical refrigeration. Of great economic importance was the manufacture of water ice for use in refrigeration units, fishing boats, fish- and meat-packing plants, and dairies. However, there was a sharp increase in the trend toward cubing, sizing, and packaging of ice. The domestic ice market practically vanished with development of the automatic household refrigerator. *See* DRY ICE.

Most ice is made in galvanized cans that are partially immersed in brine in an ice-making tank (**Fig. 1**). Brine made of sodium chloride or calcium chloride is used. The brine is cooled by ammonia as the refrigerant evaporates in pipe coils or brine coolers submerged in the ice tank. The ice cans are filled with raw water, or treated water if it initially contains large amounts of impurities. The water is usually precooled. Cold brine circulates around the ice cans to freeze the water. Commercial ice is frozen in 300- or 400-lb (135- or 180-kg) blocks for which the freezing time with brine at 12°F (−11°C) is 38–42 h. Freezing time depends largely on the thickness of the ice cake and the brine temperature. In large plants, cans are arranged in group frames for harvesting as many as 34 cans at a time. A traveling crane picks up a frozen group, transports and drops it into a dip tank for thawing, then moves it to a can dump where the ice cakes slide out and into a storage room. The empty cans are refilled with fresh water and are returned to the ice tank by the crane.

If clear ice is required, the water in each can must be agitated with air during freezing; otherwise opaque ice is formed. Because the water must be cooled to 32°F (0°C) before freezing can start and because of system losses, about 1.6 tons (1.5 metric tons) of refrigeration is required to make 1 ton (0.9 metric ton) of ice when the raw water enters at 70°F (21°C).

The manufacture of ice in slush, flake, or cube form by continuous ice-makers (**Fig. 2**) constitutes a large portion of the commercial market. Also, small, fully automatic, self-contained, cube and flake ice-makers are used in bars, restaurants, hotels, and hospitals.

Ice is also frozen artificially in a flat horizontal sheet for ice-skating. Such skating rinks may be indoor or outdoor, permanent or portable. The rink floor is covered with pipe coils through which cold brine is circulated and over which water is sprayed

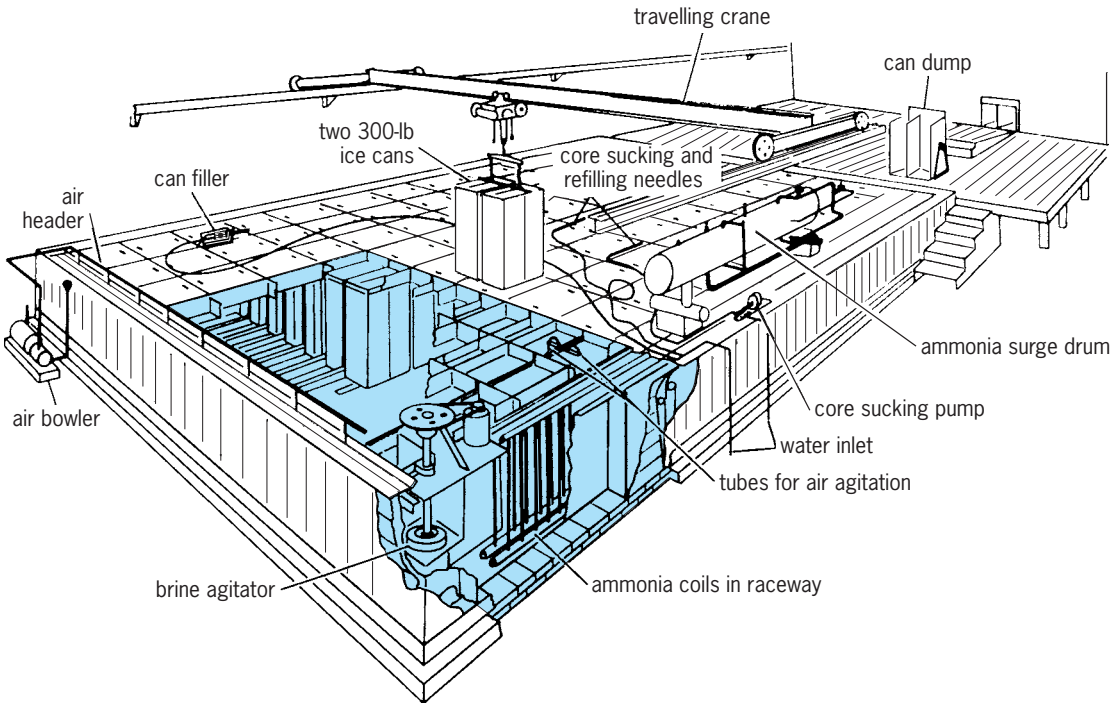

Fig. 1. Typical can ice plant. 300 lb = 135 kg. (*Worthington Pump and Machinery Corp.*)

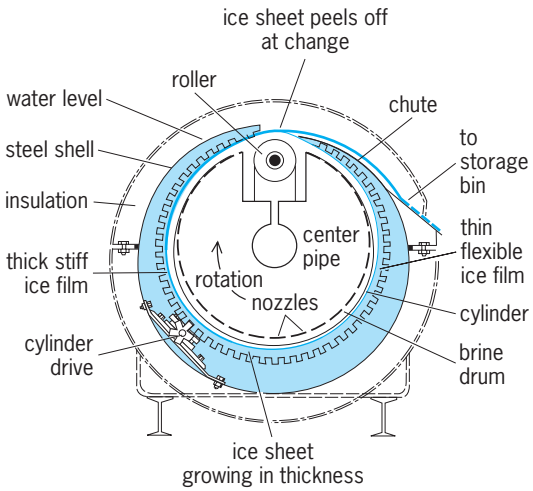

Fig. 2. Commercial flake ice machine.

until a frozen sheet $^1/_2$–$^3/_4$ in. (13–19 mm) thick is obtained. The brine is cooled in coolers by a refrigerant such as ammonia or Freon-12. *See* REFRIGERATION.

Carl F. Kayan

Ice point

The temperature at which liquid and solid water are in equilibrium under atmospheric pressure. The ice point is by far the most important "fixed point" for defining temperature scales and for calibrating thermometers. It is 273.15 K (0°C or 32°F). A closely related point is the triple point, where liquid, solid, and gaseous water are in equilibrium. It is 0.01 K (0.018°F) higher on the Kelvin scale than the ice point. The triple point has gained favor as the primary standard since it can be attained with great accuracy in a simple closed vessel, isolated from the atmosphere. Readings are reproducible to about 0.0001 K (0.0002°F), but dissolved gases or other foreign matter may raise the error to 0.001 K (0.002°F) or more. *See* TEMPERATURE; TRIPLE POINT.

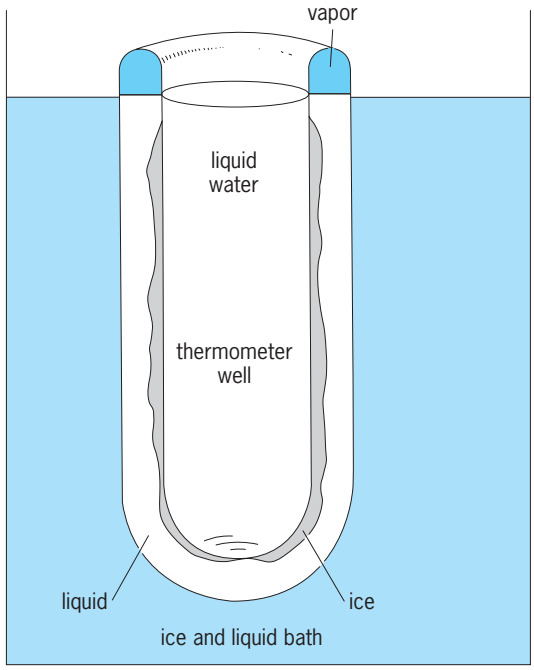

Arrangement for determining triple point.

The triple-point apparatus shown in the **illustration** consists of a thermometer well that is filled with liquid water and jacketed by a cavity containing the three phases of water under a pressure of about 0.006 atm (0.09 lb/in.^{2} or 600 pascals). The ice, initially deposited by prechilling the well, melts during heat transfer from the thermometer. *See* PHASE EQUILIBRIUM.

Ralph A. Burton

Bibliography. J. F. Schooley (ed.), *Temperature: Its Measurement and Control in Science and Industry*, 1982; K. Wark, *Thermodynamics*, 6th ed., 2000.

Iceberg

A large mass of glacial ice broken off and drifted from parent glaciers or ice shelves along polar seas. Icebergs are distinguished from polar pack ice, which is sea ice, and from frozen seawater, whose rafted or hummocked fragments may resemble small icebergs. *See* GLACIOLOGY; SEA ICE.

Characteristics and types. The continental or island icecaps of both the Arctic and Antarctic regions produce icebergs where the icecaps extend to the sea as glaciers or ice shelves. The "calving" of a large iceberg is one of nature's great spectacles, considering that a Greenland iceberg may weigh over 5 million tons and that Antarctic icebergs are many times larger. An iceberg consists of glacial ice, which is compressed snow, having a variable specific gravity that averages about 0.89 gram per cubic centimeter. An iceberg's above-water volume is about one-eighth to one-seventh of the entire mass. However, spires and peaks of an eroded or weathered iceberg results in height-to-depth ratios of between 1:6 and 1:3. Tritium age experiments on melted ice from Greenland icebergs indicate that they may be around 50,000 years old. Minute air bubbles entrapped in glacial ice impart a snow-white color to it and cause it to effervesce when immersed. *See* SEAWATER; TRITIUM.

Icebergs are classified by shape and size. Terms describing icebergs include arched, blocky, domed, pinnacled, tabular, valley, and weathered, as well as bergy-bit and growler for iceberg fragments that are smaller than cottage size above water. An iceberg may last indefinitely in cold polar waters, eroding only slightly during the summer months. But an iceberg that drifts into warmer water disintegrates in a matter of weeks in sea temperatures between 40 and 50°F (4 and 10°C) or in days in sea temperatures over 50°F. A notable feature of icebergs is their long and distant drift, under the influence of ocean currents, into shipping lanes where they become navigation hazards. The normal extent of iceberg drift is shown in **Fig. 1**.

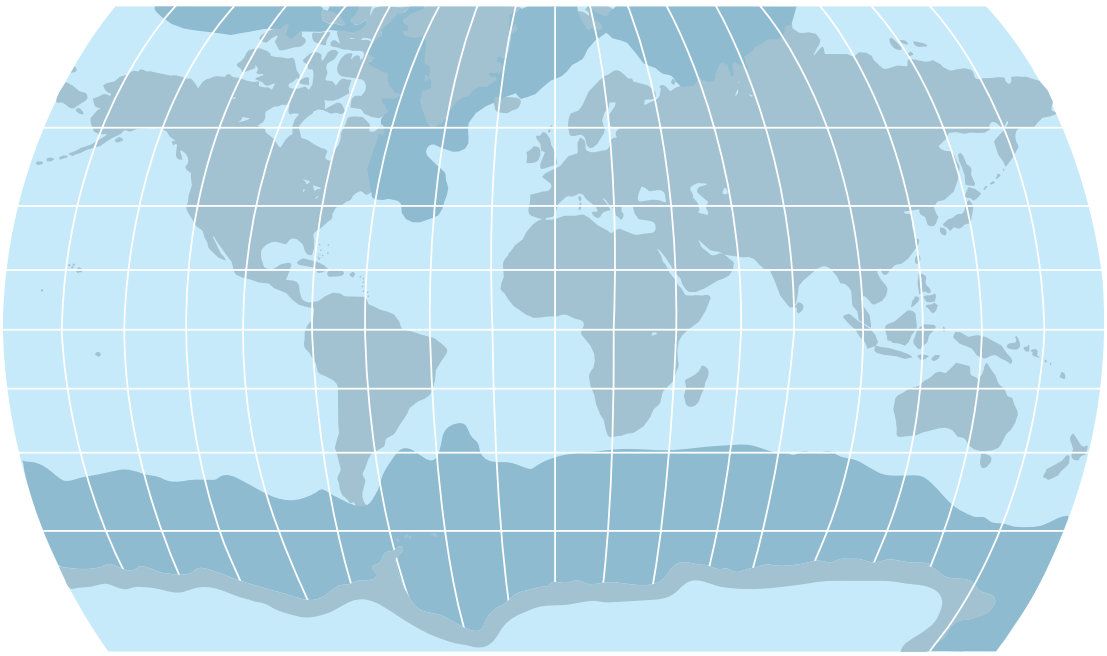

Fig. 1. Normal extent of iceberg drift, indicated by darker color tint.

Arctic icebergs. In the Arctic, icebergs originate chiefly from glaciers along Greenland's coasts. It is estimated that a total of about 16,000 icebergs are calved annually in the Northern Hemisphere, of which over 90% are of Greenland origin. But only about half of these have a size or source location to enable them to achieve any significant drift. The majority of the latter stem from some 20 glaciers along the west coast of Greenland between the 65th and 80th parallels of latitude. The most productive glacier is the Jacobshavn at latitude 68°N, calving about 1400 icebergs yearly, and the largest is the Humboldt Glacier at latitude 79° with a seaward front extending 65 mi (104 km). The remainder of the Arctic icebergs come from East Greenland and the icecaps of Ellesmere Island, Iceland, Spitsbergen, and Novaya Zemlya, with almost no sizable icebergs produced along the Eurasian or Alaskan Arctic coasts. No icebergs are discharged or drift into the North Pacific Ocean or its adjacent seas, except for a few small ones that calve each year from the piedmont glaciers along the Gulf of Alaska. These achieve no significant drift. *See* ARCTIC OCEAN; ATLANTIC OCEAN.

The Arctic ocean currents and adjacent seas determine the drift and ultimately, the distribution of icebergs—wind having only short-term effects. Icebergs along the East Greenland coast drift predominantly southward around the tip of Greenland (Cape Farewell) and then northward along its west coast. Here they join the West Greenland icebergs and travel in a counterclockwise gyre across Davis Strait and Baffin Bay. The icebergs are then carried southward under the influence of the Labrador Current along the coasts of Baffin Island, Labrador, and Newfoundland. This drift terminates along the Grand Banks of Newfoundland, where the waters of the Labrador Current mix with the warm northeastward flowing Gulf Stream. Here even the largest icebergs melt within 2–3 weeks. *See* GULF STREAM.

Freak iceberg drifts, outside the limits shown in Fig. 1, have been reported in the mid-Atlantic and off Scotland, Nova Scotia, and the Azores. In 1926 the southernmost known iceberg (a growler) reached 30–20°N, 62–32°W (about 150 mi from Bermuda). In 1912 a growler was sighted 75 mi east of Chesapeake Bay, U.S.A. Such sightings, however, are extremely rare.

About 500 icebergs a year are carried past Newfoundland (latitude 48°N) and into the North Atlantic, usually between February and June. These are survivors of an estimated 3-year journey from West Greenland. Most icebergs become stranded along the Arctic coasts and are ultimately destroyed by wave action and summer melting. The number arriving at the Grand Banks each year varies from near zero (in 1924, 1940, 1941, 1958) to 1500 or more (in 1972, 1984, 1991, 1994).

Icebergs in the Northern Hemisphere may reach proportions as large as 2000 ft in width and 400 ft in height above the water (**Fig. 2**). Such icebergs weigh upward of 15 million tons. More commonly, icebergs are 200–500 ft in width and 200 ft in height. The highest iceberg ever sighted (1959, in Melville Bay, Greenland) was 550 ft above the water. Shipping hazards, however, come not so much from large icebergs as from smaller ones, which can be difficult to detect by radar.

Ice islands. An icecap extending seaward on a broad front may form a landfast floating ice shelf rather than as an iceberg-producing glacier. Such ice shelves are common in the Antarctic but relatively rare in the Arctic. Ice broken from the shelf drifts freely as large tabular masses, ranging in size

Fig. 2. Arctic iceberg, eroded to form a valley or drydock type; grotesque shapes are common to the glacially produced icebergs of the North. Note the brash and small floes of sea ice surrounding the berg.

from a few square miles to several hundred square miles (**Fig. 3**). The term "ice island" is limited to the Arctic. Similar masses in the Antarctic are termed icebergs. Ice islands originate from the ice shelves of the northern shores of the Canadian archipelago and Greenland, and usually circulate in the Arctic Ocean for an indefinite time. The principal source of large ice islands is the Ward Hunt Ice Shelf on Northern Ellesmere Island. Aerial reconnaissance since 1946 has observed about 100 ice islands. They occasionally drift into the Norwegian Sea and Baffin Bay, where they erode and become icebergs. During

Fig. 3. Antarctic iceberg, tabular type. Such bergs develop from great ice shelves along Antarctica and may reach over 100 mi (160 km) in length. The U.S. Coast Guard icebreaker *Westwind* is in the foreground.

World War II, Winston Churchill proposed establishing airfields on ice islands in the Norwegian Sea. This was not done because of the faster rate of melting in those waters. Scientific camps have been set up on ice islands mostly by United States and Russian research agencies.

International Ice Patrol. The sinking of the passenger liner *Titanic* in 1912 with a loss of over 1500 lives resulted in the formation of the International Conferences on the Safety of Life at Sea. This included the establishment of the International Ice Patrol. The Patrol, operated by the U.S. Coast Guard, maintains an ice observation and reporting service in the North Atlantic Ocean. It is supported by 20 nations and has operated each year (except wartime) since 1913. First by ships, then search aircraft, the Patrol now uses satellite imagery and airborne radar as observation tools. It also has carried out oceanographic surveys, Greenland glacier mapping, and iceberg destruction experiments.

Other ice reporting services are maintained by Canada, Denmark, Iceland, Finland, Japan, and Russia for their respective waters, and the U.S. National Ice Center for Polar Regions.

Antarctic icebergs. In the Southern Ocean, icebergs originate from the giant ice shelves along the Antarctic continent and from glaciers which may flow into the shelves or form floating ice tongues. Masses breaking from either can result in huge, tabular icebergs 200 ft high and several hundred square miles in area. The world's largest glacier and iceberg producer is the Lambert Glacier (70°E longitude), which merges with the Amery Ice Shelf. In 1963 a giant iceberg measuring 50 by 70 mi was calved. This iceberg was tracked until 1976, by which time it had broken into thousands of smaller icebergs. The largest iceberg ever reported was calved in 1992 from the Thwaites Ice Tongue (110°W longitude). It was larger than the state of Rhode Island, and was still being tracked in 2000 with a size of about 25 by 50 mi.

Antarctic icebergs drift throughout the Southern Ocean under the influence of the Antarctic Circumpolar Current, and into the South Atlantic, Indian, and Pacific oceans, where they have been sighted as far north as latitude 35°S (Fig. 1). When weathered, Antarctic icebergs attain a deep bluish hue of great beauty rarely seen in the Arctic.

Water volume. The total volume of ice calved each year as icebergs is estimated to equal about half the volume of the world's total water usage. The feasibility of towing or otherwise moving large icebergs to water-deficient lands has been proposed and studied. But the difficulties of towing large icebergs over 6 million tons (as determined by actual tests), and the rapid deterioration of icebergs in warm waters appear to make such a venture impractible at this time.

Robertson P. Dinsmore

Bibliography. N. Bowditch, Ice in the sea, The American Practical Navigator, chap. 34, *U.S. Defence Mapping Agency Publ.*, no. 9, 1995; T. F. Budinger, Iceberg detection by radar, *Proceedings of the International Maritime Consultative Organization Conference on Safety of Life at Sea*, London, 1960; R. P. Dinsmore, Ice and Its Drift into the North Atlantic Ocean, *International Commission for Northwest Atlantic Fisheries (ICNAF) Spec. Publ.*, no. 8, 1972; J. Sater, Arctic Drifting Stations, *Report on Activities Supported by the Office of Naval Research*, 1968; E. H. Smith, The Marion Expedition to Davis Strait and Baffin Bay, 1928-29-30-31, 3 vols., *U.S. Coast Guard Bull.*, no. 19., 1932; U.S. Coast Guard, Annual Reports of the International Ice Patrol, *USCG Publ.*, no. 188, 1919 onward; U.S. Defense Mapping Agency, Sailing Directions for Antarctica, *DMA Publ.*, no. 200, 1992; World Meteorological Organization, Sea Ice Nomenclature, *WMO Publ.*, no. 259, 1971.

Icebreaker

A ship designed to break floating ice. Icebreaker technology has evolved rapidly since the 1960s as a result of potential resource development in the Arctic regions of Russia, Canada, and the United States. This led to the construction of icebreaking ships that can transit to all areas of the world, including the North Pole. The Arctic Ocean and Antarctica play a large role in shaping the global climate. With the Arctic Ocean and the ice-covered waters of Antarctica being the least studied of all the oceans, icebreakers provide the platforms from which polar science and research can be conducted on a year-round basis. On August 22, 1994, the first U. S. and Canadian icebreakers, USCGC *Polar Sea* and CCGS *Louis S. St. Laurent*, arrived at the North Pole after having conducted scientific research in areas never previously studied. *See* ANTARCTIC OCEAN; ARCTIC OCEAN.

Types. Icebreakers may be classed as polar or subpolar, depending on their primary geographic area of operation. Polar icebreakers have a capability for independent operation in first-year, second-year, and multiyear ice in the Arctic or the Antarctic. Subpolar icebreakers are capable of icebreaker operations in the ice-covered waters of coastal seas and lakes outside the polar regions. *See* MARITIME METEOROLOGY; SEA ICE.

Resistance. The speed that an icebreaker can achieve in ice depends upon the resistance of the hull to break the ice and submerge the broken ice pieces. The resistance increases approximately as the square of the level ice thickness. As an example, the resistance in breaking 4 ft (1 m) of ice is basically four times that for breaking 2 ft (0.5 m) of ice. Resistance increases almost linearly with an increase in the ship's beam. Other factors that influence resistance are the amount of friction developed between the hull surface and the ice, and, to a lesser degree, the length of the icebreaker.

Hull form. The combined shapes of the bow, midbody, and stern constitute the hull form. Each component has unique requirements to make it efficient in icebreaking and ice transiting. The icebreaker's bow must be designed to break ice efficiently, and

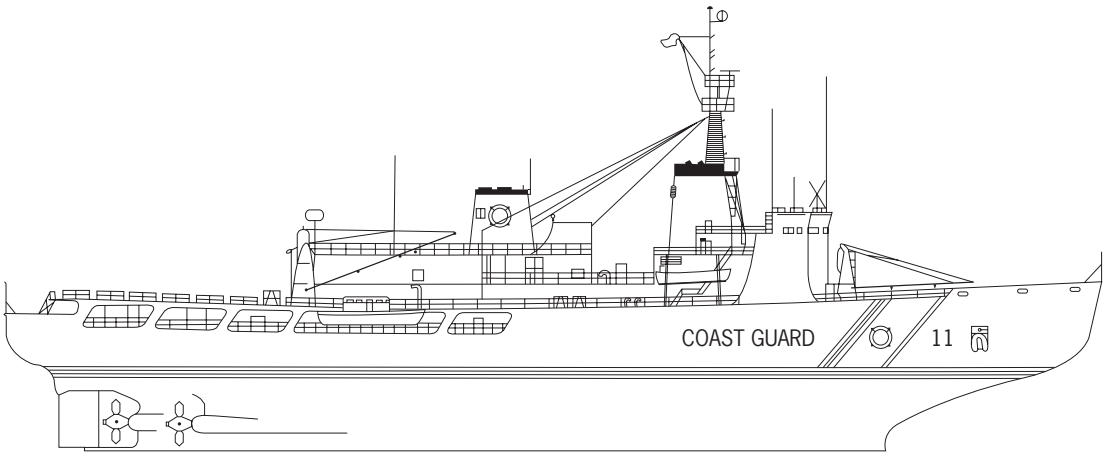

Outboard profile of the U.S. Coast Guard icebreaker *Polar Sea*.

it consists of an inclined wedge that forms an angle of about 20° with the ice surface (see **illus.**). As the icebreaker advances, the bow rides up on the edge of the ice until the weight of the ship becomes sufficiently large that the downward force causes the ice to fail. As the icebreaker advances, the broken ice pieces move under the hull and the process is repeated. The sides of the icebreaker can be vertical or sloped. Traditionally, icebreakers have sloped sides of 5–20° to reduce the frictional resistance. Sloped sides can also improve the turning ability (maneuverability) of the ship in ice. The stern of the icebreaker is designed in a manner similar to the bow so that the ship can break ice while going astern. It must also be designed for the proper flow of water into the propellers in both the ahead and astern directions. Because of the many factors that influence icebreaking performance, the design involves considerable compromise and judgment, particularly when icebreakers may spend 50% of their time in open water. *See* SHIP DESIGN.

Hull structure. The hull of a polar icebreaker is built very strong so that it can withstand tremendous ice forces. Many icebreakers have to withstand impacts with ice that is 20 ft (6 m) thick. Because these frequent and high loads occur at very low air temperatures of −50°F (−45°C), specialty steels are used on polar icebreakers. It is common to have bow and stern plating of 1.5–2.0 in. (4–5 cm) in thickness. The side and bottom plating thickness are somewhat less. Subpolar icebreakers, with only a limited ice transiting capability, have only a portion of the hull strengthened. It is called an ice belt and runs several feet above and below the waterline of the hull.

Propulsion. Icebreaker propulsion systems consist of engines, transmission systems, and propellers. Most icebreakers are powered with diesel engines. However, when the horsepower required for icebreaking is sufficiently great, gas turbines or nuclear steam turbines are needed. Transmission systems are either mechanical (primarily gears) or electrical (generators and motors). Propellers on an icebreaker can number from one to four, and the propellers can be either fixed pitch or controllable pitch. A single propeller is frequently used on those ships requiring only limited icebreaking capability. Most icebreakers, however, have either two or three propellers for purposes of system reliability and improved maneuverability. All propulsion systems must be capable of withstanding the shock and high torque loads that are transmitted to the system when the propeller blades impact or mill ice pieces. *See* MARINE ENGINE; MARINE MACHINERY; PROPELLER (MARINE CRAFT); SHIP POWERING, MANEUVERING, AND SEAKEEPING.

Reduction of icebreaking resistance. Many auxiliary methods have been developed to improve the performance of icebreakers by reducing the friction between the hull and the ice. One widely used method is to roll or heel the ship from side to side. The rolling motion is achieved by pumping large quantities of water from side to side in specially configured heeling tanks. In addition to heeling systems, most icebreakers use a low-friction coating or a stainless-steel ice belt to reduce icebreaking resistance.

Another concept is the use of air bubbler systems. In these systems, icebreaking resistance is reduced by discharging large volumes of air underwater at the bow and sides of the icebreaker. The rising air bubbles carry large quantities of water to the surface and reduce the friction at the hull-ice interface. Another method to reduce friction is the water wash system, involving pumping large quantities of water through nozzles located in the hull onto the ice surface at the bow and forward sides of the icebreaker.

Tests in ice model basins. The ability to evaluate the icebreaking performance of new designs before construction is very important. A scaled model of the ship is made and tested in a towing basin inside a large refrigerated space. Ice is grown on the water surface, and the model is towed or self-propelled through the ice to determine the resistance. *See* TOWING TANK.

Richard P. Voelker

Bibliography. L. W. Brigham (ed.), *The Soviet Maritime Arctic*, 1991; R. A. Dick and J. E. Laframboise, An empirical review of the design and performance of icebreakers, *Mar. Technol.*, 26(2):145–159, 1989; E. V. Lewis (ed.), *Principles of Naval*

Architecture, 2d ed., 3 vols., Society of Naval Architects and Marine Engineers, 1988; R. P. Voelker et al., Eight years of performance measurements aboard *Polar* class icebreakers, *Soc. Nav. Arch. Mar. Eng. Trans.*, 100:271–292, 1992; H. Yamaguchi, Experimental Voyage through Northern Sea Route, INSROP Symposium Tokyo '95, 1996.

Ichneumon

The common name for medium- to large-sized insects (wasps) comprising 25,000 species and belonging to the family Ichneumonidae (Hymenoptera); together with the Braconidae (with an additional 15,000 species) they form the superfamily Ichneumonoidea. The majority of ichneumons are slender wasps with long filiform antennae, freely rotating head, and permanantly extruded ovipositor. The ichneumons, found worldwide, are most abundant in heavy vegetational areas such as the edges of woodlands, hedgerows, and in secondary growth vegetation, and are often seen in the woodland canopy or in the litter of the forest floor. Ichneumons are more commonly seen in the spring and fall, as these wasps avoid hot, dry weather.

All members of the family Ichneumonidae are parasitic during their immature stages. They attack a wide variety of insects, usually the larval stages of beetles, butterflies, moths, flies, and other Hymenoptera. Since many of their insect hosts are considered pests of forest and agricultural crops, ichneumons are beneficial and important in biological control. *See* ENTOMOLOGY, ECONOMIC.

Female ichneumons use chemical and physical cues emanating from the host (the insect on or within which the ichneumon's progeny will develop) to locate it for egg laying. The length of the ovipositor is variable, being relatively short in species attacking exposed hosts, and extremely long in species attacking hosts enclosed in galls or stem cavities, or boring in wood. Some species, such as in the genus *Megarhyssa* (see **illus.**), may have ovipositors longer than the body.

Long-tailed ichneumon (*Megarhyssa lunator*), about 6 in. (15 cm), of which more than 4 in. (10 cm) is "tail."

The female ichneumon deposits an egg on or within the host. Most species are solitary; that is, only one egg is laid. A few species are gregarious as larvae, with the female laying several eggs in the host which all develop into adults. Species that lay their eggs on the host are referred to as ectoparasites; those which deposit their eggs within the host, endoparasites. The eggs hatch into white, fleshy, grub-like larvae which slowly consume the host. The ichneumon larvae feed through the cuticle if they are ectoparasites or feed internally on tissues if they are endoparasites. The mature larvae usually kill the host as they emerge and subsequently pupate in a cocoon.

The females of some species of ectoparasitic ichneumons inject a venom that paralyzes the host prior to laying an egg. Some endoparasites have been found to inject a symbiotic virus along with an egg into the host. The virus appears to alter the physiology of the host for the parasite's benefit.

Most ichneumons are host-specific; that is, they only attack and develop in a particular species of host. However, a few species are polyphagous, and attack a number of different hosts. The ability to attack and destroy a wide variety of insects makes ichneumons particularly important to humans in suppressing pest insects. *See* HYMENOPTERA; INSECTA.

S. B. Vinson

Bibliography. R. P. Askew, *Parasitic Insects*, 1971.

Ichthyopterygia

A supraorder of extinct marine reptiles from the Mesozoic, or the Age of Dinosaurs, belonging to the subclass Diapsida. The group consists of the order Ichthyosauria and several basal forms from the Early Triassic. *See* DIAPSIDA; DINOSAUR; REPTILIA.

Fossil record. Ichthyopterygians existed for about 160 million years. Their oldest definitive record comes from the latest part of the Early Triassic (latest Olenekian), whereas the youngest record is from the strata formed during the earliest part of the Late Cretaceous (Cenomanian). Ichthyosaurs became extinct about 25 million years before dinosaurs disappeared at the end of the Cretaceous. The timing approximately corresponds to an extinction event of marine organisms at the Cenomanian/Turonian boundary. *See* EXTINCTION (BIOLOGY).

Body characteristics. Ichthyopterygians are noted for their tuna-shaped (tunniform) body outline in combination with the presence of a well-defined tail fin or caudal fluke. Cetaceans (whales and dolphins) were the only other tetrapods (vertebrates with four limbs) to evolve this body form. Early ichthyosaurs were elongate and could be loosely described as "lizards with flippers." These early forms probably propelled themselves using body undulation, as in living cat sharks. Various body outlines evolved during the Middle Triassic, but only a single lineage, which gave rise to a demarcated caudal fluke, survived the Late Triassic. A tuna-shaped body outline also became established by mid Late Triassic

time (middle Norian), and this group diversified into a large number of species during the Jurassic and Cretaceous.

Body length varied from approximately 0.5 m (2 ft) to over 20 m (67 ft) in ichthyosaurs. Body size of the first ichthyosaurs (Early Triassic) never exceeded 3 m (10 ft), but giants reaching 15 m (50 ft) or more appeared during the Late Triassic (*Shonisaurus*), and some fragments suggest the possibility of a second whale-sized form in the Early Jurassic. The largest species from the Jurassic and Cretaceous attained maximum lengths of 10 m (33 ft).

Eyes. Many Jurassic ichthyosaurs possessed disproportionately large eyeballs, including the largest eyes that have ever been measured in vertebrates, which reached over 26 cm (10 in.) in diameter in *Temnodontosaurus*, whose body length was about 9 m (30 ft). *Ophthalmosaurus* had eyes measuring 23 cm (9 in.) and had a body length of less than 4 m (13 ft). *See* EYE (VERTEBRATE).

Limbs. Early ichthyosaurs had five digits (fingers) in their hands and feet (that is, their front and back flippers), as in most other tetrapods. However, some later species had as many as 10 digits per hand, whereas others had as few as three. These ichthyosaurs also had disk-shaped phalanges and metacarpals (finger and palm bones), unlike most other vertebrates with limbs.

Diet. Ichthyopterygian diet included fish, cephalopods, shells, and some other vertebrates such as sea turtles and birds. Possibilities also existed for feeding upon other marine reptiles. R. Motani

Bibliography. C. McGowan and R. Motani, Ichthyopterygia, *Handbuch der Palaerpetologie Part 8*, Verlag Dr. Friedrich Pfeil, Munich, 2003; C. McGowan, *Dinosaurs, Spitfires, and Sea Dragons*, Harvard University Press, Cambridge, MA, 1991; R. Motani, Ichthyosauria: Evolution and physical constraints of fish-shaped reptiles, *Annu. Rev. Earth Planet. Sci.*, 33:12.1–12.26, 2005; R. Motani, Rulers of the Jurassic seas, *Sci. Amer.*, 283(6):52–59, December 2000.

Ichthyornithiformes

A group of extinct flying birds known exclusively from the latest stages of the Cretaceous (80–65 million years ago) in North America. Along with Hesperornithiformes, Ichthyornithiforms are some of the earliest nonmodern bird taxa known to science. Indeed, alongside *Archaeopteryx* (from much earlier Jurassic rocks in Germany, about 140 million years old), the ichthyornithiformes *Ichthyornis*, *Apatornis*, and their kin formed the basis for knowledge of avian evolution for almost a century. Although the fossil record of these birds has improved in recent years, Ichthyornithiformes are still poorly understood. *See* AVES; HESPERORNITHIFORMES; ODONTOGNATHAE.

The discovery of *Ichthyornis* in 1875 caused excitement—here was a flying bird that possessed sharp pointed teeth. Charles Darwin wrote about Ichthyornithiformes (and their relatives, Hesperornithiformes) in 1880 and noted that they appeared to provide excellent evidence in support of his ideas on the origin of species. Most recently, these archaic avians were placed within the clade Ornithurae, which also contains Hesperornithiformes and all modern birds (Neornithes). *See* NEORNITHES.

Much controversy has surrounded these birds: only recently have they been well described anatomically, and at least one familiar skeleton appears to comprise several individuals. Although all original finds of these birds were from the Cretaceous marine rocks of Kansas, more recent discoveries have extended their known range across the Northern Hemisphere, into Europe and Asia. Ichthyornithiformes appear to have been important and diverse constituents of late Cretaceous ecosystems, dying out along with most lineages of dinosaurs at the end-Cretaceous mass extinction event. *See* EXTINCTION (BIOLOGY). Gareth Dyke

Bibliography. S. Chatterjee, *The Rise of Birds*, Johns Hopkins University Press, Baltimore, 1997; L. M. Chiappe and L. M. Witmer, *Mesozoic Birds: Above the Heads of Dinosaurs*, University of California Press, Berkeley, 2002; A. Feduccia, *The Origin and Evolution of Birds*, 2d ed., Yale University Press, New Haven, 1999; S. L. Olson, *The Fossil Record of Birds*, in *Avian Biology*, vol. 8, Academic Press, New York, 1985.

Ichthyostega

Four-legged vertebrates (basal tetrapods) that evolved from their lobe-finned fish ancestors during the later Devonian Period (400–350 million years ago) [see **illus.**]. *Ichthyostega* was the first Devonian tetrapod to be described and was found in East Greenland during the 1930s. For many decades it remained the only representative of the "fish-tetrapod transition" which was known from articulated skeletal fossils. Subsequently, a second genus from the same beds, *Acanthostega*, was more fully described, and during the 1990s Devonian tetrapod genera were recognized in other parts of the world.

Like other very early tetrapods, *Ichthyostega* had a skull composed of an outer casing of bone (the skull roof) with a separate box containing the brain (the braincase) inside it. The skull roof bore the teeth, and the teeth of the upper jaw were substantially larger but fewer in number than those of the lower jaw. This suggests some specialization of diet, but it is not clear for what purpose. *Ichthyostega* was undoubtedly a predator. The braincase shows some highly unusual features, especially in the ear region, which are difficult to interpret and do not resemble those of any other early tetrapod.

The postcranial skeleton shows some primitive and some very specialized features. The tail bore a fringe of fin-rays like those of a fish; these are lost in all other known early tetrapods except *Acanthostega*. The vertebrae consisted of several parts. The centra were formed of hoops of bone surrounding the

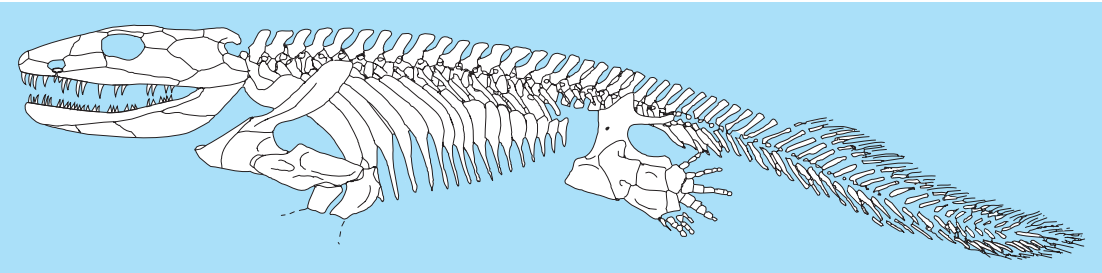

Reconstruction of the skeleton of *Ichthyostega* based on recent information. The animal was about 1 m (3 ft) long.

flexible supporting notochord. The notochord was the main point of attachment between the vertebral column and the head, inserting into the rear of the braincase. The neural arches supporting the spinal cord articulated with one another via well-developed joint surfaces (zygapophyses). The ribs were massive overlapping blades which formed a corset around the body and gave it rigidity. The animal would not have been able to bend very much.

The massive shoulder girdle bore the large complex humerus which articulated with the shoulder at right angles to the body. The elbow was also more or less permanently fixed in a right angle, allowing the forearm to be used as a prop but permitting very little other movement. The hand is still unknown. The large pelvic girdle was attached to the vertebral column only by ligaments and muscles—no bony joint surfaces were involved as they are in modern tetrapods. The femur appears to have been about half the length of the humerus, and the hindlimb was paddlelike with tibia, fibula, and ankle bones all broad and flattened. The foot bore seven toes, with three tiny ones at the leading edge and four stout ones behind. It seems likely that the toes were enclosed in a web of skin.

Because of its peculiarities, *Ichthyostega* has not been helpful in working out many ideas about the origin of tetrapods. Now however, it can be seen in the context of other Devonian tetrapods, and its specializations will help provide more information about how, where, and when early tetrapods first radiated onto the land. *See* AMPHIBIA. Jennifer A. Clack

Bibliography. J. A. Clack, Devonian tetrapod trackways and trackmakers: A review of the fossils and footprints, *Palaeogeog. Palaeoclimatol. Palaeoecol.*, 130:227–250, 1997; E. Jarvik, The Devonian tetrapod *Ichthyostega, Fossils and Strata*, 40:1–213, 1996.

Igneous rocks

Rocks that are formed from melted rock, usually by cooling. The melted rock is called magma when it is below the surface, and lava when it flows out onto the surface. When the magma or lava cools or moves into areas of lower pressure, crystals or glass or both form. Thus, igneous rocks are composed of crystals, glass, or combinations of these materials. Magmas sometimes erupt explosively, creating ash that is composed of broken crystals, glass, and rock materials called pyroclastic material. Rocks formed at or very near the surface, including pyroclastic rocks, are called volcanic rocks, whereas those formed from magma at depth are called plutonic rocks. In the past, some rocks that formed below the surface but near it were called hypabyssal rocks. *See* LAVA; MAGMA.

Like all rocks, igneous rocks are classified on the basis of their composition and texture. Both mineral and chemical compositions are used for classification, but the rocks are first subdivided by texture into three categories. (1) Volcanic rocks are dominated by interlocking grains too small to see with the unaided eye or a simple lens. (2) Plutonic rocks are composed almost entirely of interlocking grains big enough to see. (3) Pyroclastic rocks are composed of fragments (clasts) of glass, minerals, and preexisting rocks. *See* PETROLOGY.

Composition. The range of igneous rock chemistries is relatively large. The common igneous rocks consist of silica (SiO_2), 40 to 77%; alumina (Al_2O_3), <1% to nearly 20%; magnesia (MgO), <1 to 50%; total iron oxide, <1% to more than 10%; and lime (CaO), soda (Na_2O), and potash (K_2O), <1% to 9% (see **table**).

The minerals in common igneous rocks include quartz, alkali and plagioclase feldspars, micas (muscovite, biotite, and phlogopite), amphiboles (such as hornblende and riebeckite), pyroxenes (including augite and the orthopyroxenes), and olivines. In addition, minerals occurring in small amounts include apatite, zircon, epidote, titanite, magnetite, and ilmenite. *See* AMPHIBOLE; FELDSPAR; MICA; OLIVINE; PYROXENE; QUARTZ.

For purposes of description, igneous minerals are divided into three groups. Essential minerals are those on which the rock name is based. Characterizing accessory minerals are nonessential minerals that occur in amounts of 5% or more. Minor accessory minerals occur in amounts of less than 5%. *See* MINERAL.

Classification. Traditionally, classification of all igneous rocks was based on minerals and textures. In volcanic rocks, however, the minerals are often so small that identification is difficult or impossible, and glass contains no minerals. Thus, in recent years volcanic rock classification has been based largely on rock chemistry, except for classifications used in field studies.

Volcanic flow and pyroclastic rocks. The International Union of Geological Sciences (IUGS) volcanic rock

Chemical compositions of some igneous rocks in wt %

Compound	Dunite[a]	Basalt[b]	Andesite[a]	Foid-syenite[c]	Granite[a]	Rhyolite[d]
SiO_2	40.08	49.8	58.97	59.54	69.22	77.24
TiO_2	0.01	2.6	1.04	0.14	0.48	0.20
Al_2O_3	0.29	14.0	17.17	18.60	15.50	10.81
Fe_2O_3	0.31	2.5	4.36	2.86	1.03	1.66
FeO	7.62	8.5	2.02	2.09	1.42	0.27
MgO	49.69	7.2	1.51	0.10	0.73	0.33
CaO	0.11	11.3	4.90	1.16	1.93	1.48
Na_2O	0.05	2.2	4.23	8.96	4.15	2.59
K_2O	0.01	0.62	2.90	4.24	4.42	4.12
P_2O_5	0.00	0.32	0.51	0.16	0.15	0.06
Other	0.58	0.35	1.55	1.80	0.30	0.65
TOTAL	98.86	99.6	99.26	99.87	99.37	99.80

[a]F. J. Flanagan (ed.), Descriptions and Analyses of Eight New USGS Rock Standards, *U.S. Geol. Surv. Prof. Pap.*, no. 840, 1976.
[b]F. J. Flanagan et al., Basalt, BHVO-I, from Kilauea Crater Hawaii, in F. J. Flanagan (ed.), *U.S. Geol. Surv. Prof. Pap.*, no. 840, 1976.
[c]P. D. Snavely, Jr., et al., Nepheline syenite, STM-1, from Table Mountain, Oregon, in F. J. Flanagan (ed.), *U.S. Geol. Surv. Prof. Pap.*, no. 840, 1976.
[d]M. H. Staatz and W. J. Carr, Geology and Mineral Deposits of the Thomas and Dongway Ranges, Juab and Tooele Counties, Utah, *U.S. Geol. Surv. Prof. Pap.*, no. 415, 1964.
SOURCE: From L. A. Raymond, *Petrology: The Study of Igneous, Sedimentary, and Metamorphic Rocks*, 2d ed., 2002.

classification defines the common volcanic rock types on the basis of silica (SiO_2) content versus alkalis ($Na_2O + K_2O$). The most common types of volcanic rocks have less than 7% total alkalis and include basalt (45–52% SiO_2), basaltic andesite (52–57% SiO_2), andesite (57–63% SiO_2), dacite (63% to 69–77% SiO_2), and rhyolite (>69% SiO_2). *See* ANDESITE; BASALT; DACITE; RHYOLITE.

Pyroclastic volcanic rocks are classified on the basis of grain size, in addition to the compositional name. If the dominant grain size is less than 2 mm, the volcanic material is called ash and the rock composed of these grains is called tuff. If the rock contains a significant percentage of grains between 2 and 64 mm, it is called a lapilli tuff. Rocks dominated by fragments larger than 64 mm are called volcanic breccias. Thus, a complete pyroclastic rock name might be lapilli tuff rhyolite. *See* BRECCIA; PYROCLASTIC ROCKS; TUFF.

Some volcanic rocks are composed entirely or almost entirely of glass. If the glass largely lacks vesicles (gas bubble holes), the rock is called obsidian (or rhyolite obsidian). If the glassy rock contains abundant, small, aligned, tubular holes, it is called pumice. Larger vesicles, characteristic of the less siliceous rocks, require the modifier "vesicular," as in vesicular basalt. Rocks dominated by vesicles are called (basalt) scoria. *See* OBSIDIAN; PUMICE.

Plutonic rock. Classification of plutonic rocks is based primarily on mineralogy. The IUGS classification is the most widely used plutonic rock classification. In this classification, which features two triangles joined at their bases, the essential minerals are quartz, alkali feldspar, plagioclase, and feldspathoid minerals (foids). All other minerals are ignored for classification purposes, except in classification of low-silica rocks, for which separate sets of triangles are used. *See* FELDSPATHOID.

In order to classify a rock, the essential mineral percentages are observed and totaled, and the individual percentages are divided by the total to obtain a percentage of the three essential minerals. The single point on the triangle represented by the three mineral percentages determines the name for the rock. Among the most common plutonic rocks are granite, composed of quartz, alkali feldspar, and some plagioclase feldspar; and gabbro, composed of augite and plagioclase (**Fig. 1**). In the United States, some geologists use alternative classifications that subdivide the "granite" field of the IUGS classification. *See* GABBRO; GRANITE.

For low-silica rocks—the gabbros and ultrabasic rocks—separate triangles are used for classification (**Fig. 2**). The essential minerals for these classifications are plagioclase, olivines, pyroxenes, and hornblende. *See* HORNBLENDE.

Plutonic rocks are also classified by some geologists on the basis of their chemistry. Thus, there are peraluminous, metaluminous, and peralkaline rocks, which have chemistries high, moderate, and low in

Plutonic rocks	GRANITE	QUARTZ DIORITE	DIORITE	GABBRO	PERIDOTITE
Volcanic rocks	Rhyolite	Dacite	Andesite	Basalt	Komatüte

Fig. 1. Some plutonic rock names and their volcanic equivalents.

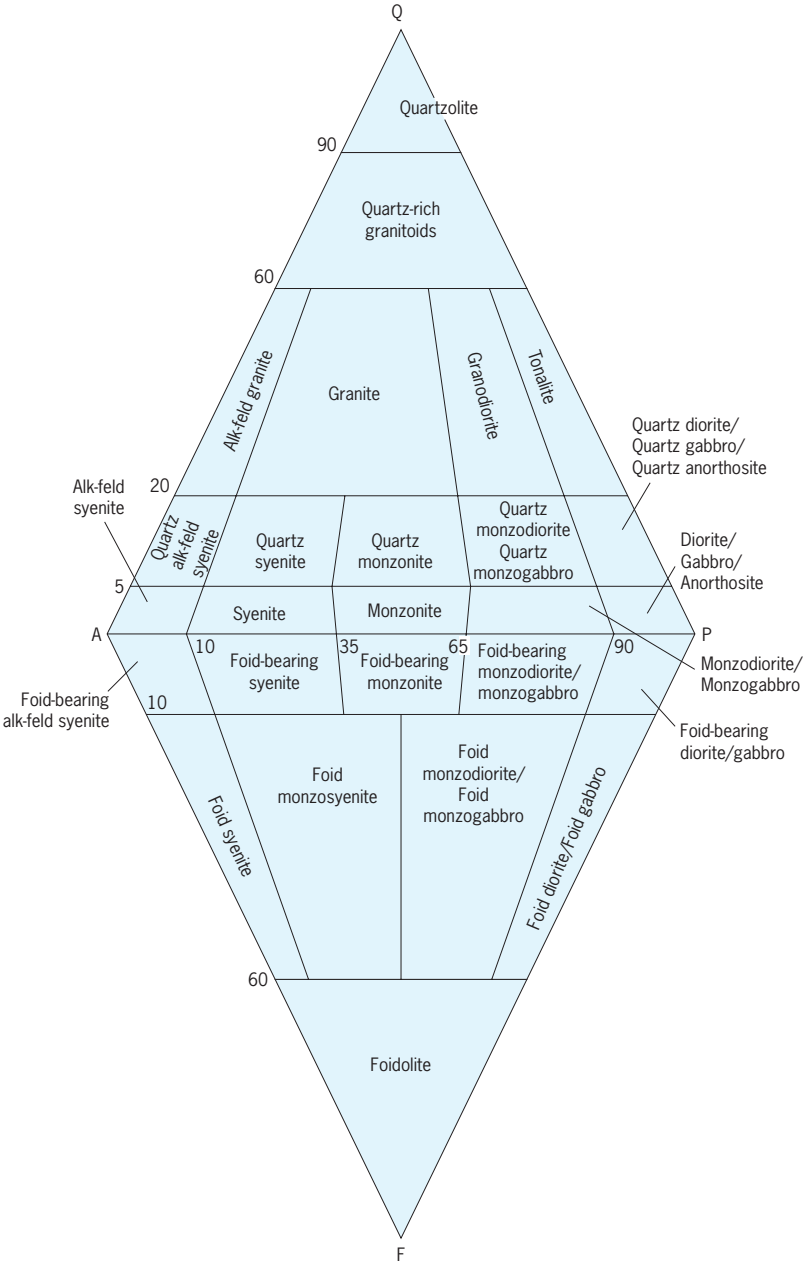

Fig. 2. IUGS classifications of gabbroic rocks (triangle AQP) and ultramafic rocks (triangle AFP). (*From A. Streckeisen, To each plutonic rock its proper name, Earth Sci. Rev., 12:1–33, 1976*)

alumina, respectively. And there are S-, I-, and A-type granitoid rocks formed from melts of sedimentary, igneous, and relatively high-alkali materials.

Textures and structures. The texture of a rock includes the size, shape, arrangement, and grain boundary features. Many geologists also include abundant holes and the relationship between glass and crystals in volcanic rocks as part of the texture. In general, a feature is textural if there are many of them in a hand-sized sample. Structures are generally larger features. There may be a few or only one in a hand specimen, or they may be much larger than a hand specimen. *See* PETROGRAPHY.

All of the plutonic rocks are crystalline in texture. This means that the grain boundaries in the rock are interlocking (**Fig. 3***a*). Rocks with grains large enough to see are said to have phaneritic texture. If the grain length exceeds 3 cm in length, the rock has pegmatitic texture. Volcanic rocks include those with crystalline textures too small to see without magnification (aphanitic textures), those with aphanitic crystals surrounding larger visible crystals (porphyritic textures) [Fig. 3*b*] and those composed of fragments (clasts) formed by explosive eruption of magmas (pyroclastic textures) [Fig. 3*c*]. In the last, clasts may include fragments of glass, fragments of crystals, or fragments of earlier-formed rocks.

Grain size, shape, and arrangement. Grain (crystal) sizes in igneous rocks vary widely. The smallest crystals

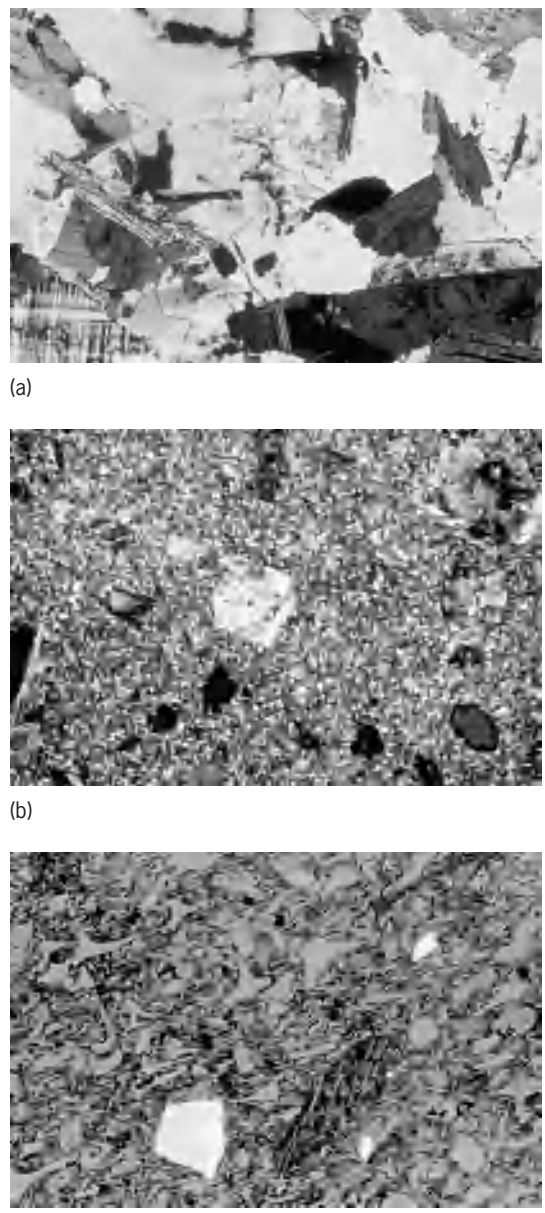

Fig. 3. Photomicrographs of common igneous rock textures. (*a*) Hypidiomorphic-granular texture. (*b*) Porphyritic texture. (*c*) Pyroclastic texture. Width of field of view is 3.2 mm (0.125 in.).

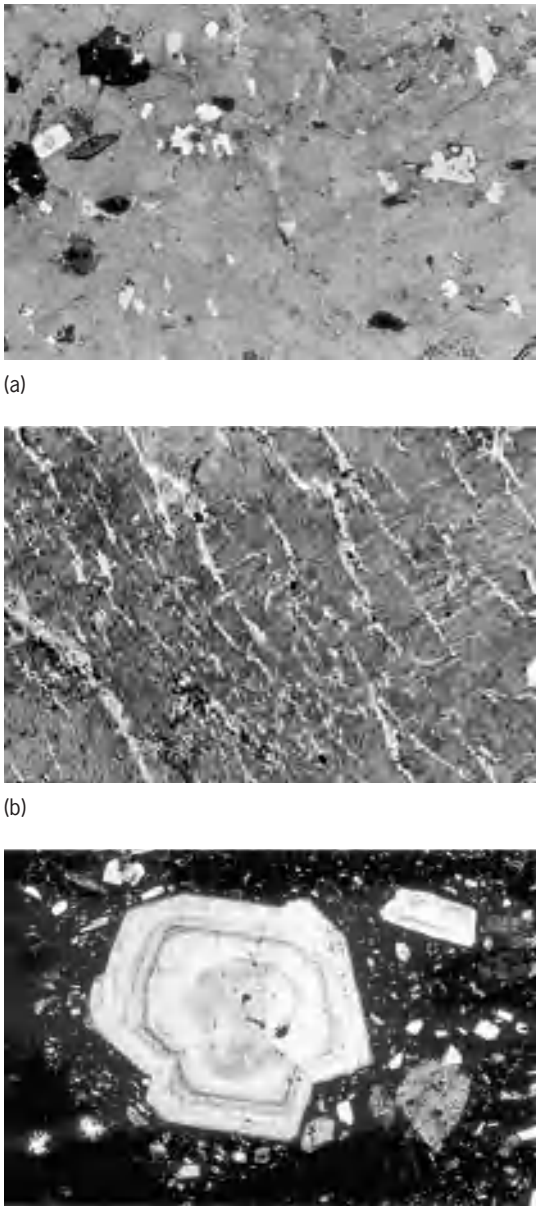

(a)

(b)

(c)

Fig. 4. Photomicrographs of some mineral textures in igneous rocks. (*a*) Poikilitic texture in alkali feldspar. (*b*) Perthitic texture in alkali feldspar. (*c*) Zoned texture in plagioclase. Width of field of view is 3.2 mm (0.125 in.).

are less than 0.01 mm, whereas the largest crystals are meters long. These largest crystals form in pegmatitic rocks, especially those rich in SiO_2. Fine-grained rocks have grains less than 1 mm in length. In medium-grained rocks, the average grains are 1 to 5 mm in length. Coarse-grained rocks have average grain lengths of 5 mm to 3 cm. In pegmatitic rocks, the dominant grain size exceeds 3 cm. In any given rock, there may be a range of crystal sizes for each mineral, called a crystal size distribution (CSD), which reflects the crystallization history. *See* PEGMATITE.

The crystalline grains of igneous rocks have shapes that vary and depend on a variety of factors controlling the growth history of the crystal. In addition, the crystal form of each mineral (the natural shape of the particular mineral) is a dominant influence on the shape. Mineral grains that have flat surfaces (crystal faces) on all sides, which formed as the crystal grew, are said to be euhedral. Crystals partially bounded by crystal faces are subhedral, and those not bounded by crystal faces are anhedral.

Early or rapid crystal growth from a melt yields partial, incomplete, or skeletal crystals. In volcanic glass, tiny beginning crystals of this type are called crystallites and microlites. In plutonic rocks, they are called skeletal or snowflake grains. *See* VOLCANIC GLASS.

Where growth of crystals is prevented by explosive eruption of magma, glass fragments erupt into the atmosphere. Tuffs form from accumulations of these variably sized, typically curved pieces of glass, mixed with fragments of crystals and rocks (Fig. 3*c*).

Minerals may form in random orientations, or they may have specific arrangements. In volcanic rocks in which elongate crystals are aligned (usually because of flow of the magma), the rock has a trachytic texture, but similar textures in plutonic rocks are called trachytoidal textures. Randomly oriented grains that are subhedral, on average, yield a hypidiomorphic-granular texture in plutonic rocks (Fig. 3*a*). This is the typical texture of the granites. The similar texture at a smaller scale in volcanic rocks is called intergranular texture.

Many specialized textures exist and have been named based on the kinds, shapes, or arrangement of grains. Porphyritic texture consists of two distinct sizes of grains, with substantially larger grains (phenocrysts) enclosed by a matrix of smaller grains (Fig. 3*b*). Porphyritic textures can be either volcanic or plutonic. Ophitic texture consists of large grains of pyroxene enclosing smaller, rectangular grains of plagioclase. Graphic texture, a texture that occurs within some pegmatitic rocks, consists of V- and L-shaped grains of quartz enclosed in a large alkali feldspar grain. The quartz in graphic texture is usually part of a single but incomplete skeletal grain. Corona textures are those in which a rim of one mineral surrounds a core of another mineral. *See* PHENOCRYST; PORPHYRY.

Mineral textures. Individual minerals within igneous rocks may also have special textures. Commonly in igneous rocks, larger grains of one mineral, such as alkali feldspar, will enclose many smaller grains of other minerals to form a poikilitic texture (**Fig. 4***a*). Ophitic and graphic textures are poikilitic textures that occur throughout and characterize particular rock types.

In granites, alkali feldspar grains also commonly have a mineral texture called perthitic. Perthitic texture results when small, tabular, irregular, or worm-shaped grains of plagioclase feldspar separate from, but typically occur as subparallel blobs within, larger alkali-feldspar grains (Fig. 4*b*). Myrmekitic and granophyric textures involve wormlike quartz intergrowths within plagioclase and alkali feldspar, respectively.

Fig. 5. Some volcanic structures. (*a*) Lava plateau with flat-lying basalt lava flows, near Sprague, Washington. (*b*) Composite volcanoes, Atitlan (left) and Toliman in Guatemala. (*c*) Columnar jointing in andesite near Lost Creek Dam, Oregon. (*d*) Pillow basalts, Golden Gate National Recreation Area, California.

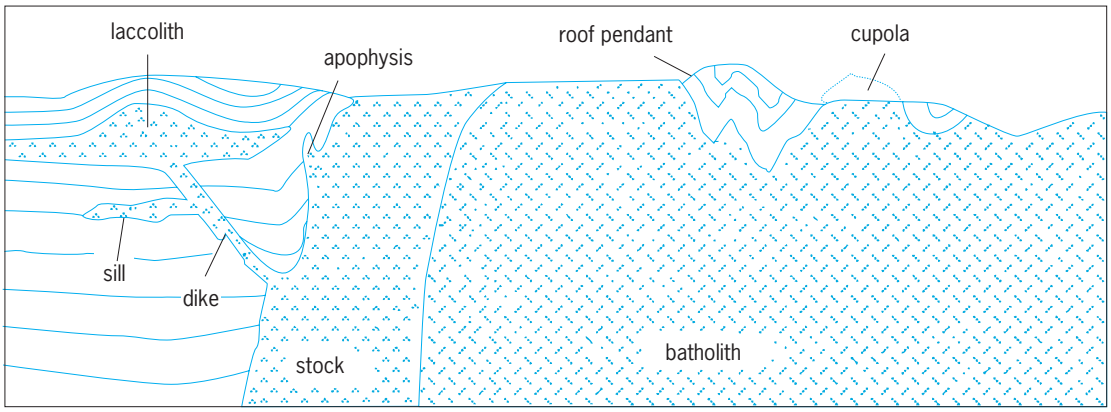

Fig. 6. Some plutonic structures.

Another important mineral texture is zoning (Fig. 4*c*). In this texture, common in plagioclase feldspar, each successive layer of the growing crystal has a different composition, so that the mineral appears to have rings of different colors or brightness under polarized light and magnification.

Texture formation. Various factors control the formation of crystals (mineral grains) in a magma. The temperature at which the crystal nucleates (begins to form), how fast it grows, the nature of the "neighborhood" of growth (that is, whether the crystal is surrounded entirely by melt or mostly by crystals, and whether the crystal is able to complete or nearly complete its growth) influence the shapes of crystals.

Crystal formation in a rock occurs in two stages: nucleation and growth. Some magmas retain unmelted crystal fragments of the rock from which the magma formed, and these form nuclei (clusters of the particular atoms needed to form specific minerals). When atoms simply attach themselves to preexisting surfaces or nuclei to form a growing nucleus for a mineral grain, this is called heterogeneous nucleation. If no nuclei exist in a magma, nuclei must form by establishing bonds between atoms. Nucleation from a magma containing no remnants of former crystals is called homogeneous. Heterogeneous nucleation is more common than homogeneous nucleation.

Once nuclei form, mineral grains grow by adding atoms to the nucleus. Growth is influenced by factors such as the composition of the melt, the temperature of the melt, the cooling rate during growth, and the rate of migration of atoms through the melt to the growing crystal. In high-silica magmas, crystals grow more slowly than in low-silica magmas, if all other factors are equal.

Some minerals begin to form earlier than others as a magma cools. For example, iron-and magnesium-rich minerals tend to nucleate at higher temperatures than potassium- and sodium-rich minerals. Once adequate cooling occurs, different minerals may grow simultaneously but at different rates. The final texture of a rock results from the growth of different crystals under the influence of the various controlling influences.

Volcanic structures. Volcanic and plutonic structures vary widely in size. Some are less than 1 cm in diameter, and others are thousands of meters across.

Major volcanic structures include lava plateaus and plains, pyroclastic sheets, shield volcanoes, composite volcanoes, and calderas (**Fig. 5**). Lava plains and plateaus are large sheets of lava that cover up to a million or more square kilometers. Pyroclastic sheets are similarly shaped masses of ash and cinder that have fallen from the sky as a result of explosive eruptions. Shield and composite volcanoes are cone-shaped piles of lava. Shield volcanoes are flat cones dominated by low-silica flows of basalt, whereas composite volcanoes are steep cones dominated by alternating lava flows and layers of pyroclastic material (Fig. 5*b*). Massive explosive eruptions sometimes produce large, crudely circular holes; these calderas may be tens to thousands of square kilometers in area. *See* CALDERA; VOLCANO; VOLCANOLOGY.

Intermediate-sized structures include small pyroclastic sheets, lava flows, pyroclastic (cinder) cones, and domes. Lava flows are sheets of lava that have flowed across the Earth's surface after erupting from a vent. Domes are piles of lava in the shape of an inverted cup that have not flowed far from the site of eruption because of their sticky character (high viscosity).

Small-sized structures range from tiny piles of lava (spatter cones) or cones of pyroclasts (small cinder cones) a few meters high to features less than a centimeter across that occur in hand specimens. Many of these structures are associated with lava flows. Small blobs of lava that squeeze out through cracks in a lava flow and solidify are called squeeze-ups. Larger bumps on lava flows are called pressure ridges. Where lavas cool in tabular sheets, they commonly form polygonal, columnar joints (Fig. 5*c*); but if they are extruded or flow into water, they may form pillow structures—tubular masses that are elliptical in cross section (Fig. 5*d*).

Other relatively small volcanic structures include craters, small lava flows, bombs, xenoliths, autoliths, and flow bands. Craters are the circular holes formed by eruption of magma. They range from a few meters to about 1 km (0.6 mi) across. Bombs are masses of lava of various shapes that have been blown into the air and have fallen to the ground as volcanic rock. Xenoliths (foreign stones) and autoliths (earlier crystallized parts of a magma) occur as inclusions in volcanic rocks. Flow bands are millimeter to centimeter thick layers that form during flow from concentrations of bubbles, crystals, or magma of a different composition. Vesicles (holes) in lava flows and amygdules (holes filled with minerals) are considered structures by some geologists. *See* XENOLITH.

Plutonic structures. Plutonic structures occur as both large- and small-sized features in magmas that intrude into rocks of the crust. The large features include batholiths, stocks, lopoliths, roof pendants, cupolas, and various types of dikes and sills. Batholiths are bodies of plutonic rock that are exposed at the surface over an area of 100 km^2 (40 mi^2) or more. Stocks are plutonic bodies of smaller dimension (**Fig. 6**). Batholiths are commonly lens-shaped. Lopoliths are special types of batholiths that are dish-shaped and typically consist of layers of silica-poor rock. Roof pendants are masses of country rock (the rock surrounding a batholith or stock) that hang down into the top of the plutonic rock body (the pluton). Cupolas are dome-shaped masses of the pluton that stick up into the country rock (Fig. 6). *See* PLUTON.

Various types of dikes and sills represent masses of magma that intruded and solidified within country rocks. Dikes are crosscutting, typically sheetlike structures that range from centimeters to kilometers of scale (Fig. 6). Some dikes are cone-shaped or cylindrical and are called cone sheets and ring

dikes. Short fat dikes are called apophyses. Sills are intruded sheets of rock that parallel the layers in the country rocks (Fig. 6). When sills are relatively short, thick, and convex on their tops, they are called laccoliths.

Small-scale plutonic structures include dikes, xenoliths, and autoliths, as well as layers and foliation. Gabbroic magmas may undergo a process of fractional crystallization in which crystals form and settle to the bottom of the magma body to form layers of alternating composition. Where early formed crystals are dragged along and aligned by flow of the magma, the aligned crystals define foliation, a structure that also occurs in metamorphic rocks. *See* METAMORPHIC ROCKS.

Magmas. In order to form igneous rocks, older rock must melt to form a magma. Usually only part of a rock melts, a process called anatexis. The magma never has the same chemical composition as the original rock, because some minerals melt before others. Generally, the newly formed magma is richer than the rock from which it came in elements such as silicon, aluminum, sodium, and potassium.

Origins. Magmas form in two general regions: at the base of the crust and in the mantle below the crust. Magmas form when rocks melt as temperature rises (thermal melting), as pressure falls (decompression melting), and as water is added to hot rock (flux melting). Flux melting typically occurs where wet ocean crust is carried down into the mantle (subducted) as a result of plate movements, and water from suducting rocks migrates into hot mantle rocks above the subducted crust. The water causes flux melting that creates basaltic magma. *See* EARTH INTERIOR.

Mantle rocks are hot and plastic. In some places they are cooler than in others, and the hotter rocks tend to rise or flow plastically upward. As they rise, the pressure from the weight of the overlying rock becomes lower. The drop in pressure causes decompression melting of mantle rocks, forming basaltic magmas. Decompression melting is most common in the mantle beneath mid-ocean ridges (spreading centers). *See* MID-OCEANIC RIDGE.

Thermal melting occurs primarily at the bottom of the Earth's crust. Hot basaltic magmas rising from the mantle have extra heat. If they form a pool or pond at the base of the crust, they may have enough heat to melt the crust to form a rhyolitic or granitic magma.

Movement and modification. Once magmas form, they tend to rise because they are hotter and lighter than the rocks around them. Basaltic magmas are more fluid than rhyolitic magmas and can move more rapidly. Through fractures, they can move up 100 km (62 mi) from the mantle in one to a few days. Sometimes they move as diapirs—blobs shaped like an upside-down teardrop. Diapirs move more slowly than magmas in fractures because they must force their way through the overlying rocks. *See* DIAPIR.

Rhyolitic/granitic magmas also move through cracks and as diapirs, but because they are thicker (stickier) they move more slowly than basaltic magmas. In many cases, they do not reach the surface. Although they solidify if they move too slowly, calculations show that if rhyolitic/granitic magmas did not solidify, it might take them more than a million years to move 100 km.

As magmas move upward, they are commonly modified in composition. The three most common processes of modification are fractional crystallization, assimilation, and mixing. Mixing occurs where two magmas of different compositions come together at the same time and place. For example, where basaltic magma melts the bottom of the crust to form a rhyolitic magma, these two magma types may mix to form an andesitic magma.

Assimilation occurs if a magma melts or dissolves the rocks on the sides or walls of a diapir, fracture, or magma chamber, or if pieces of the walls fall into the magma and are absorbed. Generally, the chemistry of wall rocks is different from that of the magma, so assimilated material changes the chemistry of the magma. The chemistry of a magma can also change through fractional crystallization, one of several processes of differentiation. In this process, part (a fraction) of the magma crystallizes to form certain minerals. The minerals have a different chemistry than the whole magma, so when they form they are removing particular chemicals from the magma, changing its composition.

When modified magmas crystallize into rocks, they form rocks of compositions different from those that would have formed from the original magmas. Thus, from the two basic magma types—rhyolitic and basaltic—a whole range of rock types may form.

Loren A. Raymond

Bibliography. M. G. Best and E. H. Christiansen, *Igneous Petrology*, 2001; L. A. Raymond, *Petrology: The Study of Igneous, Sedimentary, and Metamorphic Rocks*, 2d ed., 2002; H. Sigurdsson (ed.), *Encyclopedia of Volcanoes*, 2000; J. D. Winter, *An Introduction to Igneous and Metamorphic Petrology*, 2001.

Ignimbrite

A pyroclastic rock deposit formed by one or more ground-hugging flows of hot volcanic fragments and particles, essentially synonymous with pyroclastic-flow deposit, ash-flow tuff, flood tuff, or welded tuff. When coined by the New Zealand geologist P. Marshall in 1935, the term "ignimbrite" referred only to flows with emplacement temperatures high enough to cause the fragments forming the deposit to become adhered together or welded; this restriction no longer applies, and the term now includes any pyroclastic-flow deposit, welded or not.

Ignimbrites are commonly produced during explosive eruptions and are associated with most of the world's volcanic systems. They vary in size by orders of magnitude (10^{-3} to 10^{3} km^3 of erupted material) and have chemical compositions that span the entire range commonly exhibited by igneous rocks (basaltic

A stack of ignimbrites erupted from the Valles Caldera forms the extensive Pajarito Plateau in the Jemez Mountains, New Mexico. (*Photograph by R. A. Bailey, USGS*)

to rhyolitic). An ignimbrite can be of any form and size, but most deposits have sheetlike shapes and cover many thousands of square kilometers. The largest ignimbrite deposits tend to have dacitic to rhyolitic compositions and are associated with powerful, caldera-forming explosive eruptions—usually in the distant geologic past—involving huge eruptive volumes. For example, nearly 4800 km^3 appeared 28 million years ago in a single short-duration eruption from La Garita caldera (San Juan Mountains, Colorado) to form the Fish Canyon Tuff. By comparison, the largest eruption in recorded history (1815, Tambora Volcano, Indonesia) produced only about 30 km^3. Large ignimbrite deposits have sufficient volume and thickness to bury the topography of the underlying terrain and construct nearly flat-topped accumulations, forming distinctive volcanic plateaus such as Yellowstone Plateau (Wyoming and Montana), much of North Island, New Zealand, and the region around Valles Caldera (Jemez Mountains, New Mexico) [see **illus.**].

Ignimbrite deposits are characterized by a poorly sorted aggregate of ash (crystals and glass shards) and pumice. In the larger deposits, the pumice fragments may be flattened and stretched to yield ovoid-to-lenticular shapes, reflecting the compaction and welding of the deposit after or during emplacement. The degree of welding depends on several factors, the most important of which are the thickness of the flow, its temperature, and the amount of contained gas; in general, the greatest welding is seen in the lower half of the thickest part of the deposit.

Specialists in ignimbrite recognize a distinction between flow units and cooling units. A flow unit is a layer that represents a single pyroclastic flow, and the boundaries between flow units are marked by differences—sometimes subtle—in grain size, mineral composition, concentration and nature of pumice, and textural and structural features. If several high-temperature flow units are deposited in rapid succession, they may cool as a single cooling unit. *See* CALDERA; IGNEOUS ROCKS; PUMICE; PYROCLASTIC ROCKS; TUFF; VOLCANIC GLASS; VOLCANO.

Robert I. Tilling

Bibliography. R. A. F. Cas and J. V. Wright, *Volcanic Successions: Modern and Ancient*, Chapman & Hall, London, 1988; R. V. Fisher and H.-U. Schmincke, *Pyroclastic Rocks*, Springer-Verlag, Berlin, 1984; P. Francis, *Volcanoes: A Planetary Perspective*, Oxford University Press, Oxford, U.K., 1993; G. Heiken, *An Atlas of Volcanic Ash*, Smithsonian Contributions to the Earth Sciences, Smithsonian Institution, Washington, DC, 1974; G. A. Macdonald, *Volcanoes*, Prentice Hall, Englewood Cliffs, NJ, 1972; W. E. Scott, Volcanic and related hazards, in R. I. Tilling (ed.), *Short Course in Geology*, vol. 1: *Volcanic Hazards*, American Geophysical Union, Washington, DC, 1989; H. Sigurdsson et al. (eds.), *Encyclopedia of Volcanoes*, Academic Press, San Diego, 2000.

Ignition system

The system in an internal combustion engine that initiates the chemical reaction between fuel and air in the cylinder charge by producing a spark. An ignition system for a multicylinder internal combustion engine has three basic functions: (1) to provide a sufficiently energetic spark to initiate the burning of the fuel-air mixture within each cylinder; (2) to control spark timing for optimum efficiency so that cylinder pressure reaches its maximum value shortly after the piston reaches the top of its compression stroke; and (3) to select the correct cylinder fired.

The ignition system consists of the ignition module, ignition coil, distributor, spark plugs, high-voltage wires, and spark-timing control system. The spark timing function is one of the functions

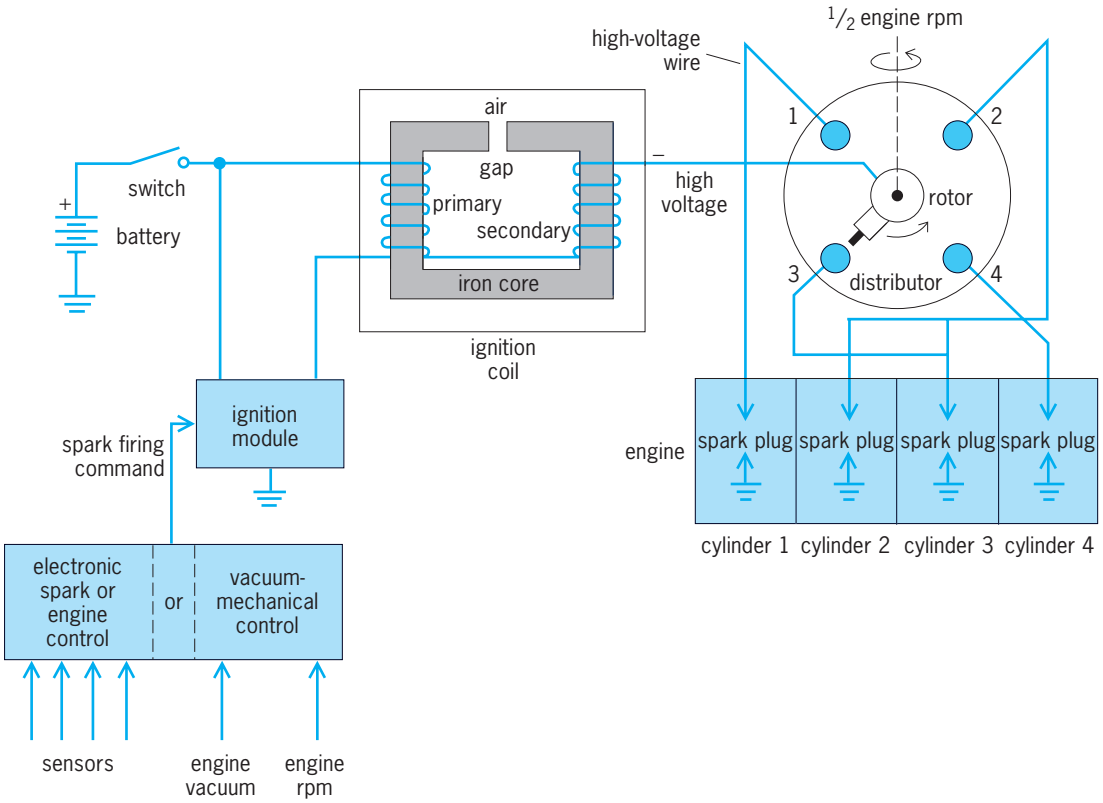

Fig. 1. Inductive discharge ignition system for a four-cylinder engine. The cylinder firing order is 1–3–4–2.

performed by an electronic engine control system or may be performed by a vacuum-mechanical mechanism.

In an inductive ignition system (**Fig. 1**), there are three possible types of control-vacuum-mechanical, electronic spark, or full electronic engine. Prior to spark discharge, electrical energy is stored inductively in the coil primary. The current to the coil primary winding is turned on and off by the ignition module in response to the spark-timing trigger signal. The current-off time marks the beginning of the sparking event. An accurate spark-timing schedule is a complex function of many engine variables, such as fuel-air composition, engine revolutions per minute (rpm), temperature, cylinder pressure, exhaust gas recirculation rate, knock tendency, and engine design. Optimum spark timing is determined from experimental engine testing. A closed-loop electronic feedback control system may dynamically adjust spark timing according to sensor feedback data and internal algorithms.

The ignition coil stores electrical energy during the dwell (current-on) period and acts as a transformer at the end of dwell by converting the low-voltage–high-current energy stored in the primary to high-voltage–low-current energy in the secondary. The distributor selects the fired spark plug by positioning the rotor opposite the terminal connected to one spark plug. The plug selected depends on the cylinder firing order, which in turn depends on the engine design. The distributor is driven at one-half engine speed from the camshaft. *See* SPARK PLUG.

When high voltage (10–30 kV) is created in the coil secondary, a spark jumps from the rotor to a distributor cap terminal, establishing a conducting path from the ignition coil high-voltage terminal along a high-voltage wire to the spark plug. Each cylinder usually has one spark plug. (High-efficiency engines may have two spark plugs per cylinder and two complete ignition systems.) The plug electrodes project as far into the cylinder as possible. After high voltage is applied to the plug, an electrical discharge is generated between its two electrodes. The energy and temperature of this discharge must be sufficient to reliably ignite the fuel-air mixture under all encountered conditions of composition, temperature, and pressure. Otherwise, misfires occur, leading to loss of engine power, excessive emissions, and unsteady power delivery. Reliable ignition results in a self-sustaining flame that travels quickly and completely across the entire cylinder, consuming all the combustible mixture. *See* FLAME.

Inductive discharge ignition. Figure 1 is an example of a standard inductive discharge ignition system. (The term inductive refers to the method of energy storage, which in this case is in the primary coil winding.) Electrical energy is obtained from a 6- or 12-V storage battery, which is connected to the ignition coil primary winding, consisting of several hundred turns of large-diameter wire around an iron core. Wound on the same core are several tens of thousands of turns of small-diameter wire constituting the secondary winding. Energy is stored in the magnetic core with primary current present. When

the primary current (1–6 A) is rapidly switched off by the electronic module, current is immediately created in the secondary, since the stored magnetic energy cannot change instantaneously. The ratio of the initial secondary current to the initial primary current is equal to the ratio of primary to secondary turns. The secondary current charges up any capacitance in the secondary circuit, resulting in high spark plug voltage and the creation of a spark. An inductive discharge ignition system has a duration of 0.001 to 0.002 s and a spark energy of 0.005 to 0.050 joule.

In simple ignition systems, the control of spark timing is based only on engine rpm and manifold vacuum. Timing is advanced to occur earlier in the compression stroke as the value for rpm increases and as manifold vacuum increases. Higher manifold vacuum results in lower cylinder pressure during compression and slower combustion. The rpm advance compensates for the ignition delay period occurring between spark firing and the start of rapid combustion. The spark-timing control in vacuum-mechanical ignition systems is accomplished by a combined centrifugal and vacuum advance mechanism that phase-shifts the ignition trigger signal as required.

Other ignition systems. Among the several other types of ignition systems for internal combustion engines are capacitive discharge, multiple-firing capacitive discharge, continuous sustaining, magneto, and distributorless ignitions. The input energy for capacitive discharge systems is stored on a capacitor at several hundred volts (generated by a dc-dc converter). A semiconductor switch (thyristor) controls the discharge of the capacitor into the primary winding. In a multiple-firing capacitive discharge ignition, the ignition module repetitively fires a capacitive discharge ignition during one spark event, increasing both the energy and effective time duration of the spark. In a continuous sustaining ignition, supplemental electrical power is added to the spark after it is established, resulting in electronically controlled extended duration rather than uncontrolled duration as for conventional ignitions. In a magneto ignition, electric current and energy are generated in the primary by relative rotational motion between a magnet and a coil (electromagnetic induction). High voltage is generated in the secondary when a set of contacts in the primary circuit is mechanically opened. Magnetos require no external source of electrical power.

The distributorless ignition system eliminates the need for mechanical distribution of spark energy by using a single coil for one, two, or four cylinders. For the two-cylinder–single-coil system (**Fig. 2***a*),

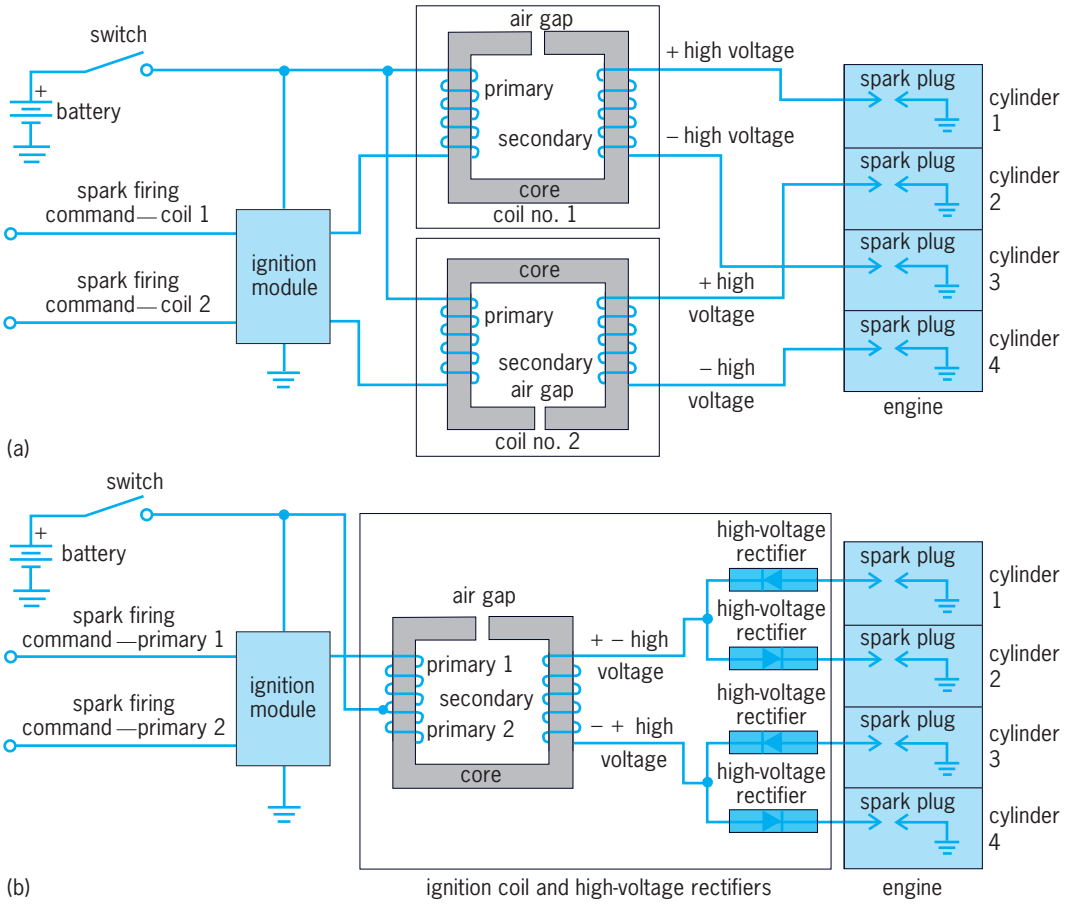

Fig. 2. Distributorless inductive discharge ignition systems for four-cylinder engines. (*a*) Design with two double-ended coils. (*b*) Design with one double-ended coil and four high-voltage rectifiers.

a double-ended ignition coil simultaneously fires a cylinder in a compression stroke together with a second in an exhaust stroke. The exhaust stroke cylinder accepts the waste spark to complete the electrical circuit through the engine block. A design variation (Fig. 2*b*) uses alternating polarity high voltage from a special type of double-ended coil and four high-voltage rectifiers to fire four plugs. The rectifiers steer the voltage to the correct pair of plugs.

Preignition and knock. These are two ignition-related problems that may adversely affect the efficiency of spark ignition engines and even cause damage. Preignition occurs when a hot spot in the cylinder, rather than the spark, initiates combustion, resulting in uncontrolled firing. The hot spot is associated with either cylinder deposits or the spark plug tip itself. Knock is a very rapid, explosive type of combustion that may result in engine damage if severe enough. Knock is often signaled by an audible thudding or pinging. It occurs when an unburned fraction of the premixed fuel-air burns abnormally rapidly. For good efficiency, engines are often operated at light knock (pinging) during low-power operation without incurring damage. *See* SPARK KNOCK.

Compression ignition. In diesel or compression ignition engines, sparkless ignition occurs almost immediately after fuel injection into the cylinder due to high in-cylinder air temperatures. High temperatures result from the high compression ratio of diesel engines. Mechanical or electronic injection timing systems determine ignition timing. *See* COMBUSTION CHAMBER; DIESEL ENGINE; INTERNAL COMBUSTION ENGINE. Joseph R. Asik

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; W. B. Ribbens and N. P. Mansour, *Understanding Automotive Electronics*, 5th ed., 1998.

Illiciales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Magnoliidae of the class Magnoliopsida (dicotyledons). The order consists of two families, Illiciaceae and Schisandraceae, with fewer than a hundred species in all. They are all woody, with scattered spherical cells containing volatile oils. The leaves are alternate and simple. The flowers are solitary or a few clustered together, regular, and hypogynous; the perianth has five to many segments that are not clearly differentiated into sepals and petals. The Illiciales are related to the Magnoliales and have sometimes been included in that order, but they are apparently more advanced in having fundamentally triaperturate pollen. *Illicium*, with about 40 species in southeastern Asia and in the Caribbean region, is the source of the spice star anise. *See* MAGNOLIIDAE; MAGNOLIOPSIDA. T. M. Barkley

Illite

A clay-size, micaceous mineral; a common component of soil, sediments, sedimentary rocks, and hydrothermal deposits. Generally, the term refers to any relatively nonexpanding clay that possesses a *c* dimension of approximately 1.0 nanometer. Specifically, illite is considered to possess a smaller layer charge and potassium (K) content than muscovite, its potassium (K) content generally being about 0.8 equivalent per half-unit cell as opposed to a content of −1.0 equivalent for muscovite. It is characterized by the ideal formula $Al_2(Si_{3.2}Al_{0.8})O_{10}(OH)_2K_{0.8}$.

Structure. Illite possesses a micalike structure (see **illus.**) consisting of an octahedral sheet sandwiched between two tetrahedral sheets. Apical oxygens from the tetrahedra replace hydroxyls in the alumina octahedra. Negative charge on the 2:1 layer may arise from isomorphous substitutions in both the tetrahedral and octahedral sheets, but tetrahedral substitutions generally predominate. This layer charge is electrically balanced by interlayer potassium ions which, as in muscovite, are fixed in hexagonal cavities between adjacent 2:1 layers in approximate 12-fold coordination with oxygen. Illites may be either dioctahedral or trioctahedral, but the dioctahedral variety is most common. It has been found in five polytypes: disordered one-layer monoclinic ($1M_d$), one-layer monoclinic (1M), two-layer monoclinic ($2M_1$ and $2M_2$), and three-layer triclinic (3T).

Illite frequently is interlayered with smectite, and can be considered the nonswelling end member in an illite-smectite mixed-layer clay series. Pure end-member illite, with no interlayer smectite, is rare.

Properties. Pure illite is characterized by a cation exchange capacity of about 10 milliequivalents/100 g, this capacity arising from broken bonds on crystal edges. Because most illites contain some

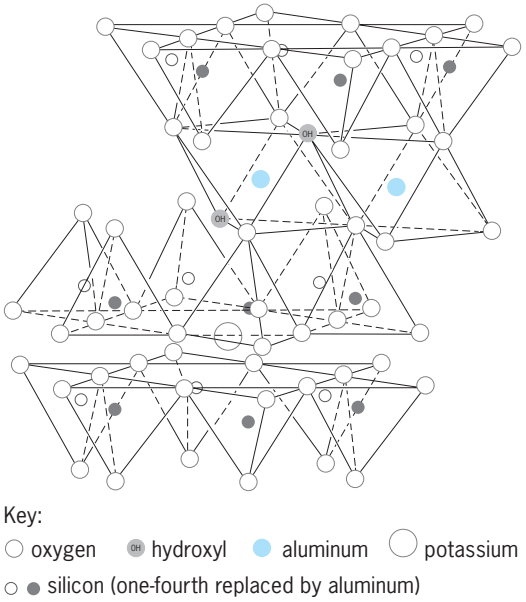

Structure of illite. (*After R. E. Grim, Clay Mineralogy, McGraw-Hill, 2d ed., 1968*)

smectite layers, however, measured cation exchange capacities are generally greater than this figure.

Fixed interlayer potassium can be leached with difficulty from illite and replaced by cations of greater hydration energy, thereby transforming illite into vermiculite. This structure can be converted back into illite by adding potassium. This reconstitution can occur during crop fertilization or when clays are transported into the ocean. *See* VERMICULITE.

Formation. Illite forms abundantly by the reaction of smectite during burial diagenesis in sedimentary basins such as the Gulf of Mexico. The reaction is detectable when burial temperature reaches 140°F (60°C); at this depth smectite begins to react to form mixed-layer illite-smectite by dissolving potassium feldspar and fixing interlayer potassium. The reaction is K-feldspar + smectite → illite-smectite + chlorite + quartz. The extent of this reaction is a function of temperature, pressure, solution composition, and time. Illites also can form in potassium-rich weathering environments and by hydrothermal action. An illite like structure also can be formed by subjecting smectite to wetting and drying cycles in the presence of potassium compounds, and by saturating high-charge smectites or vermiculites with K^+. *See* DIAGENESIS; WEATHERING PROCESSES.

Uses. Illitic clays are used for manufacturing structural clay products such as brick and tile. Some degraded illites (vermiculites) are used for molding sands. Illite may also be useful for storing certain types of radioactive wastes, because it is less subject to transformation by heat than are other common clays and because it is highly specific for the sorption of cesium. *See* CLAY, COMMERCIAL; CLAY MINERALS.

Dennis D. Eberl

Bibliography. S. W. Bailey (ed.), *Reviews in Mineralogy*, vol. 13, 1984; G. W. Brindley and G. Brown (eds.), *Crystal Structures of Clay Minerals and Their X-Ray Identification*, 1980; R. E. Grim, *Clay Mineralogy*, 2d ed., 1968.

Illuminance

A term expressing the density of luminous flux incident on a surface. This word has been proposed by the Colorimetry Committee of the Optical Society of America to replace the term illumination. The definitions are the same. The symbol of illumination is E, and the equation is $E = dF/dA$, where A is the area of the illuminated surface and F is the luminous flux. *See* ILLUMINATION; LUMINOUS FLUX; PHOTOMETRY.

Russell C. Putnam

Illumination

In a general sense, the science of the application of lighting. Radiation in the range of wavelengths of 0.38–0.76 micrometer produces the visual effect commonly called light by the response of the average human eye for normal (photopic) brilliance levels. Illumination engineering pertains to the sources of lighting and the design of lighting systems which distribute light to produce a comfortable and effective environment for seeing. In a specific quantitative sense, illumination is the combination of the spatial density of radiant power received at a surface and the effectiveness of that radiation in producing a visual effect.

Illuminance, the alternate term for illumination, is receiving growing worldwide acceptance, particularly in science, but it is not generally used by illuminating engineers in the United States. *See* ILLUMINANCE.

Subjective color. Subjective evaluation of the surfaces of objects as viewed by the human eye has been treated in three classifications. All three are usually grouped by the nonspecialist as "color." *See* COLOR VISION.

One of the classifications is the attribute of brilliance or subjective brightness, in respect to which every subjective evaluation of the seeing process may be classified according to its subjective intensity, ranging from very dim to very bright.

A second classification is hue, which describes color according to the common color names, such as red, yellow, and green.

The third classification is called variously saturation or purity or vividness of hue. Two colors may be subjectively evaluated as having the same brilliance and the same hue, but the first may seem to contain more white than the second. Conversely, one would evaluate the second color as being a more vivid or pure color than the first.

The hue designation of pure (completely saturated) spectral colors may be approximately assigned to regions of the wavelength spectrum.

Spectral response function of human eye. The spectral response of the human eye to radiation at normal (photopic) brilliance levels was tested by K. S. Gibson and E. P. T. Tyndall. The data from these tests were the source from which the Commission Internationale de l'Eclairage in 1924 established by international agreement a standard brilliance response of the human eye to radiant flux. The function was first known as the CIE values of visibility. The letters ICI (International Commission on Illumination) for the English name are also used. Later the function was named the relative luminosity function, but is now called the spectral luminous efficiency of radiant flux or simply the spectral luminous efficiency. The ordinates of this function are plotted against the wavelength of radiation in **Fig. 1** and are shown in the **table**. *See* LUMINOUS EFFICIENCY.

At threshold (scotopic) levels of seeing, the spectral response function of the human eye is shifted toward lower wavelengths. This is known as the Purkinje effect.

Photometric definitions. Illumination (or illuminance) is the density of radiant power incident upon a receiving surface evaluated throughout the spectrum in terms of the spectral luminous efficiency. Illumination is symbolized by E and is the density of luminous flux per unit of surface area. The unit of illumination in the International System of Units

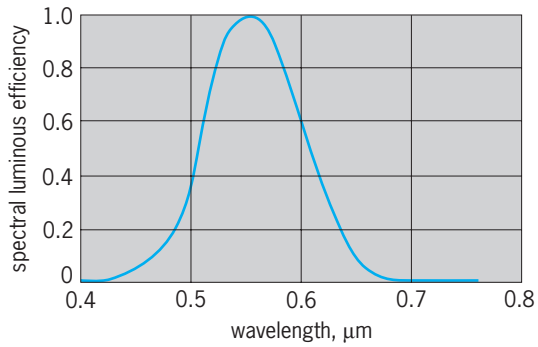

Fig. 1. Spectral response of human eye to radiation. (*After W. B. Boast, Illumination Engineering, 2d ed., McGraw-Hill, 1953*)

(SI) is the lux, which is equal to the lumen per square meter (1 lx = 1 lm/m^2). Units of illumination for other area measures are the phot (ph), equal to lumen per square centimeter, and the footcandle (fc), equal to lumen per square foot. *See* UNITS OF MEASUREMENT.

For a line-spectrum component of irradiation, a simple multiplication of the irradiation component G_i and the value of the spectral luminous efficiency υ_i taken at the proper value of wavelength i yields the component illumination. For continuous-spectrum irradiation, G_λ is a function of wavelength; therefore, the product of the two functions υ_λ and G_λ must be integrated. Equation (1) states mathematically a combination of line spectra and a continuous-spectrum source. Here G_i is the irradiation due to the ith component of the line spectra, υ_i is the spectral luminous efficiency at the wavelength of the ith component irradiation, n is the number of the component line-spectra radiations, G_λ is the spectral irradiation of the continuous spectrum at λ wavelength, υ_λ is the spectral luminous efficiency at the corresponding wavelengths, and k is a constant equal to 683 lm/W.

$$E = k\sum_{i}^{n} \upsilon_i G_i + k\int_0^\infty \upsilon_\lambda G_\lambda \, d\lambda \qquad (1)$$

Spectral luminous efficiency of radiant flux (value of unity at 0.554 μm wavelength)

Wavelength, μm	Efficiency
0.38	0.00004
0.39	0.00012
0.40	0.0004
0.41	0.0012
0.42	0.0040
0.43	0.0116
0.44	0.023
0.45	0.038
0.46	0.060
0.47	0.091
0.48	0.139
0.49	0.208
0.50	0.323
0.51	0.503
0.52	0.710
0.53	0.862
0.54	0.954
0.55	0.995
0.56	0.995
0.57	0.952
0.58	0.870
0.59	0.757
0.60	0.631
0.61	0.503
0.62	0.381
0.63	0.265
0.64	0.175
0.65	0.107
0.66	0.061
0.67	0.032
0.68	0.017
0.69	0.0082
0.70	0.0041
0.71	0.0021
0.72	0.00105
0.73	0.00052
0.74	0.00025
0.75	0.00012
0.76	0.00006

Colorimetry. The experiments from which the curve of Fig. 1 was obtained involved matching the brilliance of radiant flux from narrow spectral bands of adjacent wavelengths by 52 observers. Many scientists had concluded that by adding the radiations from three distinctly different spectral sources properly distributed in the visible spectrum, combinations could be obtained that would uniquely define measurements for all colors. The science of such measurements is called colorimetry. *See* COLOR.

Luminous flux. As previously implied, the lumen is the SI unit of luminous flux that results when radiant power is evaluated through the spectral luminous efficiency function. Related to illumination (or illuminance), the luminous flux Φ received over an area A is given by Eq. (2). The lumen is the only presently used unit of luminous flux and is the SI standard. In Eq. (2) the units E and A must be consistent combinations between lux and square meter, phot and square centimeter, or footcandle and square foot. If E is uniform over the area A, then $\Phi = EA$. *See* LUMINOUS FLUX.

$$\Phi = \int_A E \, dA \qquad (2)$$

Luminous intensity. The luminous intensity I of a source is the radiant intensity J of that source evaluated throughout the spectrum in terms of the spectral luminous efficiency (υ_i or υ_λ or both) in the same manner as illumination was derived from radiant power density upon a receiving surface in Eq. (1).

The concepts of radiant intensity and luminous intensity are strictly applicable only to a point source. In practice, however, flux (radiant or luminous) emanating from a source whose dimensions are negligibly small in comparison with distances from which it is observed may be considered as coming from a point.

The luminous intensity I of a small source, when viewed in a particular direction, is related to the differential of luminous flux $d\Phi$ emitted in that direction in a differential solid angle $d\omega$ by Eq. (3). The

$$I = \frac{d\Phi}{d\omega} \qquad (3)$$

candela, the SI unit of luminous intensity, is related to the lumen as 1 candela equals 1 lumen per unit solid angle (often called steradian).

Quantitative standardization in the field of illumination has been established in various manners by international agreement over the years. At a Paris meeting in October 1979, the General Conference on Weights and Measures redefined the base SI unit candela as the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} Hz and of which the radiant intensity in that direction is 1/683 watt per steradian (W/sr).

The frequency of 540×10^{12} Hz corresponds to a wavelength of 0.555 μm in free space where the velocity of light is essentially 3×10^8 m/s (1.86×10^5 mi/s). This frequency and wavelength is essentially that where the peak of unity occurs for the spectral luminous efficiency function of Fig. 1. Thus Eq. (4) relates J_i in W/sr for the monochromatic ra-

$$I_i = kv_iJ_i \tag{4}$$

diation to the resulting monochromatic intensity I_i in candela. The constant k is 683 lm/W as before in Eq. (1). For the source defining the candela, v_i is unity at $f = 540 \times 10^{12}$ Hz, and J_i is 1/683 W/sr, giving Eq. (5). Consistency thus exists among Eqs. (3) and

$$I_i = kv_iJ_i = \left(683\frac{\text{lm}}{\text{W}}\right)(1.000)\left(\frac{1}{683}\frac{\text{W}}{\text{sr}}\right) = 1\frac{\text{lm}}{\text{sr}} = 1 \text{ candela} \tag{5}$$

(4) and the constant k from Eq. (1). *See* LUMINOUS INTENSITY; PHYSICAL MEASUREMENT.

Luminance. Many sources of luminous flux are not small relative to the distance between the luminous source and points on receiving surfaces where illumination results. Consequently, another concept called luminance, or photometric brightness, is very convenient.

The luminance at a point on a radiating surface, acting either from self-emission or by reflection of luminous flux, can be related to a differential segment of the radiating surface dA and the differential luminous intensity dI when the differential area is observed from a particular direction by Eq. (6). The

$$L = \frac{dI}{\cos\alpha\ dA} \tag{6}$$

angle α is that between the particular direction of observation and the normal to the differential surface area dA. The denominator $\cos\alpha\ dA$ is the projected area of the differential element from the viewed direction. The equation states that the luminance L is the ratio of the differential luminous intensity to the projected differential area from the direction of observation.

The metric unit of candela per square meter (cd/m^2) has been named the nit (nt) and is recognized as the SI unit of luminance. The fractional value (also metric) of cd/cm^2 is called the stilb (sb). Additional units involving a factor π are the apostilb (asb = cd/π m^2, lambert (L = cd/π m^2), and the foot-lambert (ftL = cd/π ft^2). The reason for the π factor in these units is discussed below. *See* LUMINANCE.

Inverse-square law. Consider a differential area dA of a luminous surface possessing a differential intensity dI in a specific direction. At a distance D in that specific direction, the differential illumination dE_P upon a receiving surface at a point P is given by Eq. (7), where β is the angle between the normal to

$$dE_P = \frac{dI}{D^2}\cos\beta \tag{7}$$

the receiving surface and D. Equation (7) indicates that the differential illumination varies inversely as the square of the distance between source and receiving surface. Substituting dI from Eq. (6) into Eq. (7) yields Eq. (8). The total illumination from an

$$dE_P = \frac{L\ \cos\alpha\cos\beta\ dA}{D^2} \tag{8}$$

extended surface may be obtained by integrating the right-hand side of Eq. (8) over the surface.

If the total luminous source is negligibly small relative to D, the inverse square law is given by Eq. (9),

$$E_P = \frac{I}{D^2}\cos\beta \tag{9}$$

where I is the intensity in the direction of D toward P. *See* INVERSE-SQUARE LAW.

Light sources. Nature's source of radiation, the Sun, produces radiant power on the Earth extending from wavelengths below 0.3 μm in the ultraviolet region to well over 3 μm in the infrared region of the spectrum. The Sun's spectral radiation per unit of wavelength is greatest in the region of 0.4–0.9 μm. The response of the human eye is well matched to this range of wavelengths. *See* SOLAR RADIATION; VISION.

Oil-flame and gas lighting were used before the advent of the electric light, but since about 1900 electric energy has been the source of essentially all modern lighting devices.

Incandescent lamps. These electric lamps operate by virtue of the incandescence of a filament heated by electric current. The filament usually is composed of tungsten and in the conventional lamp is contained in either a vacuum or inert-gas-filled glass bulb. More recent designs use a halogen regenerative cycle. Currently, iodine or bromide vapors are used in a quartz bulb, which in some designs is internal to an outer, conventional inert-gas-filled bulb. The evaporated tungsten reacts chemically to form a halogen compound, which returns the tungsten to the filament. Bulb blackening is prevented, higher-temperature filaments with increased luminous efficacy result, and higher luminous output can be achieved in relatively small bulbs. Much more efficient optical designs can be achieved for projection equipment, where the source size of light should be small for efficient utilization of the luminous flux. *See* INCANDESCENT LAMP; INFRARED LAMP.

Vapor lamps. These operate by the passage of an electric current through a gas or vapor. In some lamps the

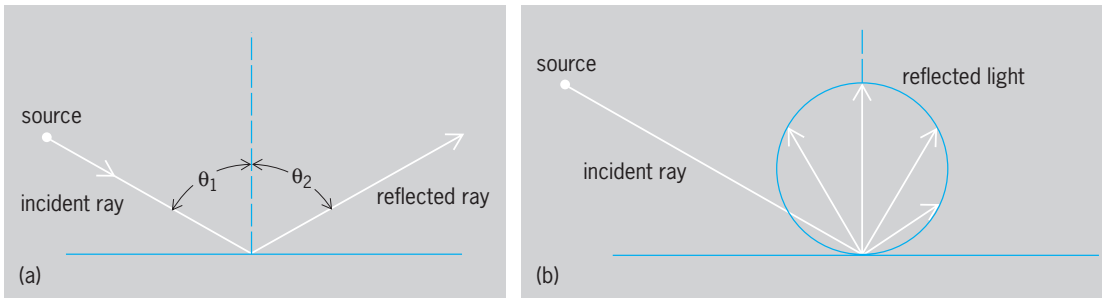

Fig. 2. Ideal reflection. (*a*) Specular, such as from a polished surface. (*b*) Diffuse, such as from a that surface of microscopic roughness. (*After W. B. Boast, Illumination Engineering, 2d ed., McGraw-Hill, 1953*)

light may be produced by incandescence of one or both electrodes. In others the radiation results from luminescent phenomena in the electrically excited gas itself. *See* ARC LAMP; VAPOR LAMP.

Fluorescent lamps. These are usually in the form of a glass tube, either straight or curved, coated internally with one or more fluorescent powders called phosphors. Electrodes are located at each end of the tube. The lamp is filled to a low pressure with an inert gas, and a small amount of mercury is added. The electric current passing through the gas and vapor generates ultraviolet radiation, which in turn excites the phosphors to emit light. If the emission of light continues only during the excitation, the process is called fluorescence. If the materials continue to emit light after the source of excitation energy is removed, the process is called phosphorescence. Phosphors for fluorescent lamps are chosen to accentuate the fluorescent action. *See* FLUORESCENT LAMP.

Reflection and transmission. The control of light is of primary importance in illumination engineering because light sources rarely have inherent characteristics of distribution, brightness, or color desirable for direct application. Modification of light may be provided in a number of ways, all of which may be grouped under the general topics of reflection and transmission. Reflection from a surface and transmission through it each may be classified according to their spatial and spectral characteristics.

Spatial characteristics. Spatially a surface may exhibit reflection conditions ranging from a regular, or specular, reflection to an ideally diffuse characteristic. Similarly, transmission may range spatially from complete transparency to an idealized diffuse transmission.

Figure 2 illustrates the extremes of specular reflection such as would be obtained from polished metal or silvered glass, and of an ideal mat-finished surface possessing microscopic roughness of minute crystals or pigment particles. The specular reflector gives a direct image of the source, with the angle of reflection equal to the angle of incidence (Fig. 2*a*). The plot shown for the diffuse reflector (Fig. 2*b*) illustrates the differential luminous intensity dI of a differential area dA as observed from different directions in the plane of the incident ray. For perfect diffusion, where L is constant for all angles of observation, it follows from Eq. (6) that dI must possess a spatial distribution with α which contains the cos α function. The ideal diffuser illustrated in Fig. 2*b* has such a distribution. The dI distribution with α for a perfect diffuser is the same cosine distribution for all planes passing through the normal to the reflecting surface element dA.

The differential flux density, $d\Phi_E/dA$, emitted by any reflecting surface is symbolized by M and called luminous exitance (luminous emittance is used also but is deprecated). For the perfectly diffusing surface of constant luminance L (called a Lambert surface), differential luminous intensity can be integrated over α from 0 to $\pi/2$ and shown to produce a differential flux $d\Phi_E$ of $\pi L/dA$. The ratio of $d\Phi_E/dA$ is the luminous exitance M. Thus for the perfectly diffusing surface, Eq. (10) is valid.

$$M = \frac{d\Phi_E}{dA} = \pi \tag{10}$$

The luminance units apostilb, lambert, and footlambert, which contain the factor π in their denominators, are sometimes called rationalized units because they are convenient for expressing the luminance of a perfectly diffusing surface. It is unfortunate that the use of units of luminance established

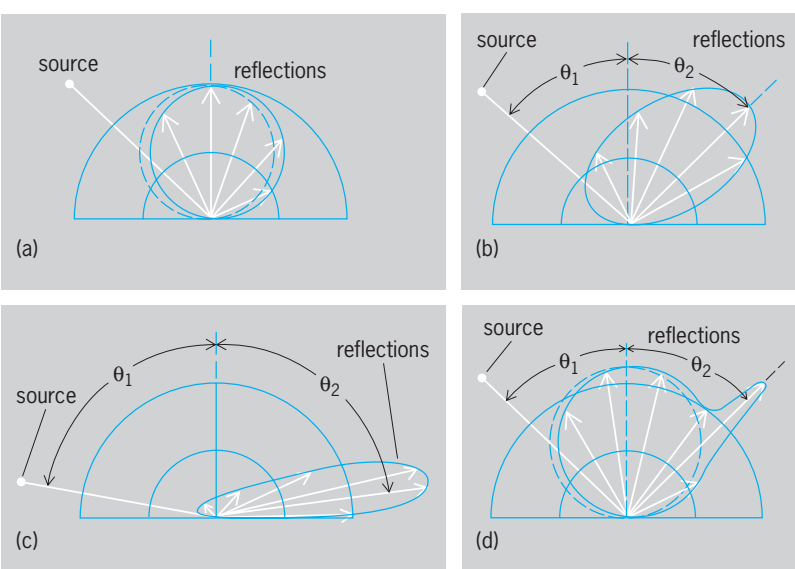

Fig. 3. Typical reflection characteristics. (*a*) Practical mat surface. (*b*) Semimat surface. (*c*) Semimat surface at large angle of incidence. (*d*) Porcelain-enameled steel. Broken-line curves in *a* and *d* illustrate ideal cosine distribution. (*After W. B. Boast, Illumination Engineering, 2d ed., McGraw-Hill, 1953*)

to match for a special situation of perfect diffusion has persisted.

Practical surfaces possess spatial-reflection characteristics intermediate between the ideals of diffuse and specular reflection, particularly at the diffuse end of the range. Typical distribution curves are shown in **Fig. 3**. In all of these illustrations the light source is small. Figure 3*a* demonstrates that even with the best practical mat surfaces, such as dull-finished metals or those painted with flat paint, the location of the source of light has some slight influence upon the intensity distribution as evidenced by an irregularity of the cosine distribution of intensity of the small area of the surface Figure 3*b* and *c* demonstrates the pronounced influence of the location of the light source upon the intensity distribution for surfaces covered with semimat materials, such as satin-finish paint. Figure 3*d* illustrates a mixed-reflection phenomenon resulting when a diffusing surface is overlaid with a surface possessing sheen.

Spatial characteristics of transmission are illustrated in **Fig. 4**. Figure 4*a* shows regular or transparent transmission, such as occurs with clear glass or plastic. Figure 4 *b* shows the diffuse transmission that would result with an idealized diffusing material. In Fig. 4*a* a refraction of the transparent image occurs, but angles θ_1 and θ_2 are equal.

Practical transmission materials possess spatial transmission characteristics intermediate between these ideals and exhibit intensity distribution curves illustrated by **Fig. 5**. Figure 5*a* demonstrates that a perfect cosine intensity distribution is not obtained even with a solid opal- or milk-glass material. A more direct transmission is shown in Fig. 5*b* for a flashed opal-glass medium, in which the majority of the base material is clear glass. Sandblasted or frosted-glass material is much less diffusing, and the location of the light source greatly influences the intensity distribution curves of Fig. 5*c* and *d*.

Spectral characteristics. If a homogeneous radiation of wavelength λ impinges upon a surface, the reflection characteristic or the transmission characteristic will result geometrically according to the spatial considerations of the preceding section. The magnitude of total power density reflected, or transmitted, to that received is defined as the spectral reflectance ρ_λ or the spectral transmittance τ_λ respectively. Thus $\rho_\lambda = J_\lambda/G_\lambda$ and $\tau_\lambda = J'_\lambda/G_\lambda$, where J_λ is the reflected spectral emission at wavelength λ, J'_λ is the transmitted spectral emission at wavelength λ, and G_λ is the spectral irradiation at wavelength λ.

A device for measuring ρ_λ or τ_λ is called a spectrophotometer. Comparison of J_λ or J'_λ with G_λ (or its essential equivalence) may be accomplished by visual or photoelectric methods of photometry.

Either continuous- or line-spectra sources of irradiation may constitute a practical source of radiant-power density. The resulting secondary source of radiant-power density J_λ or J'_λ will be modified, compared with the original, depending on spectral reflection or transmission of the reflecting or transmitting material.

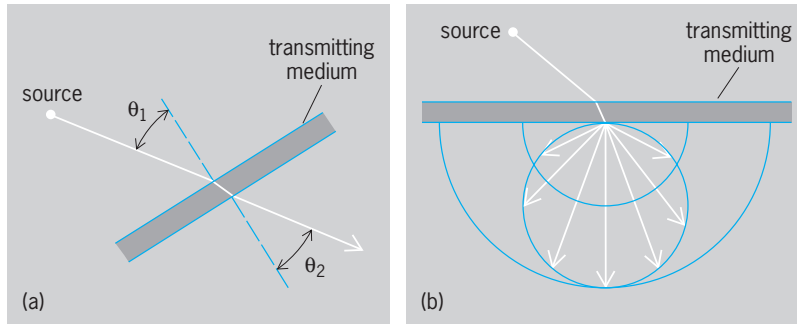

Fig. 4. Spatial characteristics of ideal transmission media. (*a*) Specular. (*b*) Diffuse. (*After W. B. Boast, Illumination Engineering, 2d ed., McGraw-Hill, 1953*)

An evaluation of J_λ or J'_λ using the spectral luminous efficiency v_λ may be effected for the secondary sources of radiant-power density in the same manner as is done for incident radiant-power density conditions. Both line- and continuous-spectrum effects may be included as demonstrated for the evaluation of the illumination. The evaluation of the reflected or transmitted luminous flux density yields the luminous exitance, symbolized by M for the reflecting surface and M' for the transmitting surface. The functional operations are shown graphically in the series of curves of **Fig. 6** for the reflection resulting from a continuous-spectrum source. The ratio of M or M' to the illuminating E results in $\rho = M/E$ and $\tau = M'/E$, where ρ is the reflectance of the surface and τ is the transmittance of the material. The reflectance and transmittance are dependent upon the source

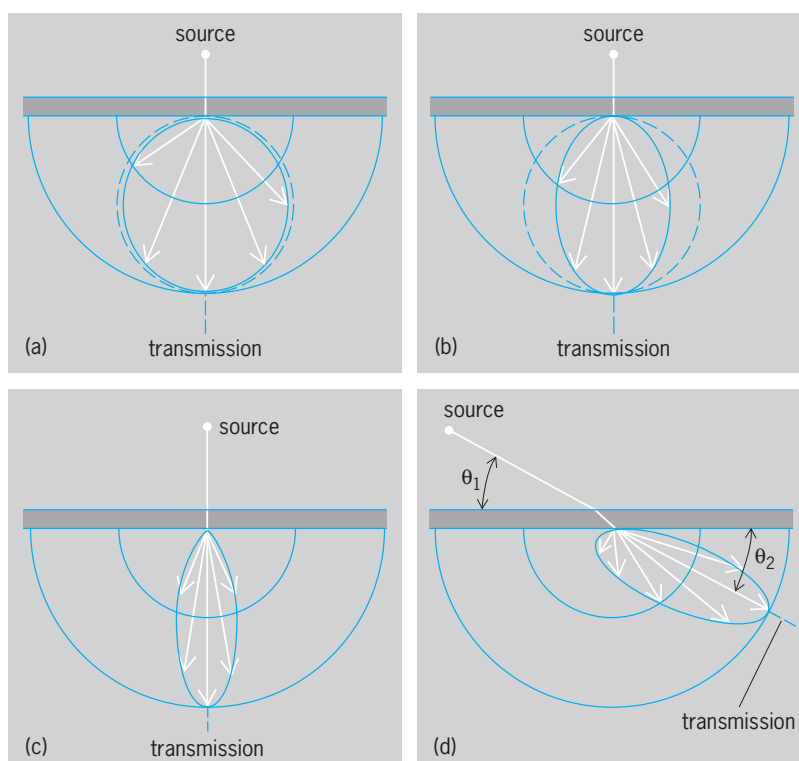

Fig. 5. Transmission characteristics. (*a*) Solid opal glass. (*b*) Flashed opal glass. (*c*) Sand-blasted glass. (*d*) Sand-blasted glass, incident light at 30° from plane at surface. (*After W. B. Boast, Illumination Engineering, 2d ed., McGraw-Hill, 1953*)

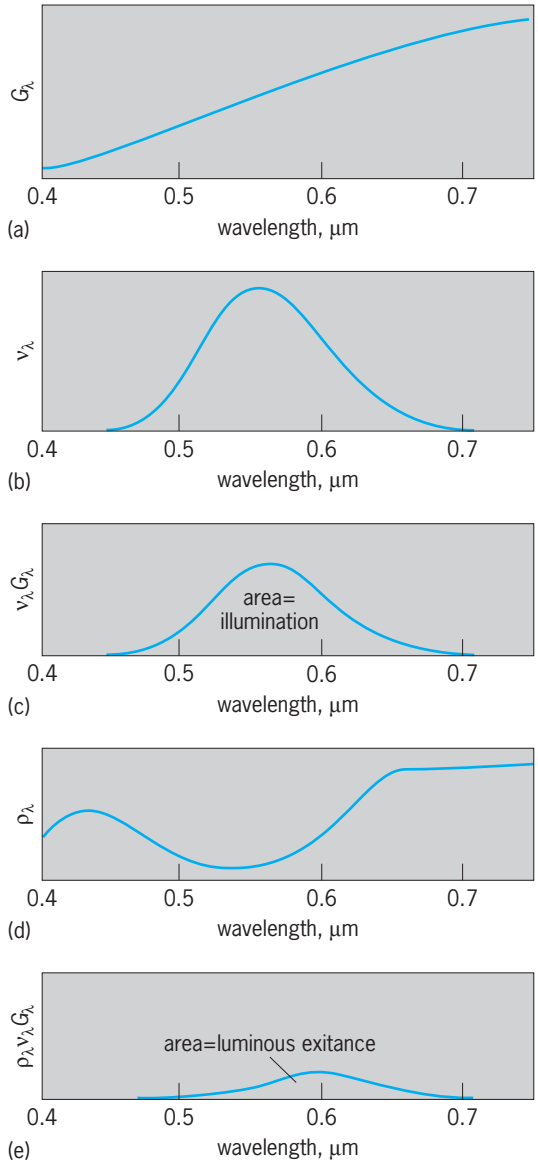

Fig. 6. Process of evaluating the luminous exitance of a reflecting surface after evaluating the illumination. (*a*) Spectral irradiation $G_\lambda$curve. (*b*) Spectral luminous efficiency v_λ curve. (*c*) Curve of $v_\lambda G_\lambda$. Illumination is the area under this curve, the integration of $v_\lambda G_\lambda$. (*d*) Spectral reflectance $_\lambda$ curve. (*e*) Curve of $\rho_\lambda v_\lambda G_\lambda$. Luminous exitance is the area under this curve, the integration of $\rho_\lambda v_\lambda G_\lambda$. (*After W. B. Boast, Illumination Engineering, 2d ed., McGraw-Hill, 1953*)

of illumination as well as the spectral characteristics of the surface or the transmitting meedium. Reflectances and transmittances of surfaces and media are frequently published for the case of a tungsten-filament light source (often without reference to this source) and are useful as guides and indications of magnitude if the sources of illumination do not deviate too greatly from this special distribution.

Light-control methods. Either a control of the character of the reflecting surface or transmitting material or a control of the primary source of luminous flux determines the degree of light control possible. If a light ray is reflected or transmitted diffusely, the shape of the reflecting or transmitting surface is of little importance, although its size may be of great importance. A larger diffuse surface reduces the brightness of the primary source. If a light ray is reflected from a specular surface, more accurate control of the light can be effected.

Parabolic reflectors. These are probably the most useful of all reflector forms. The rays emerge essentially parallel from such a reflector if the light source is placed at its focal point. Searchlights use parabolic mirror reflectors.

Elliptic reflectors. A relatively large amount of light may be made to pass through a small opening before spreading out by using an elliptic reflector. A light source is placed at one of the focal points of a nearly complete elliptical reflector. A small opening is placed at the second focal point where the light is focused. Such a design is used for the pinhole spotlight. For further discussion of reflectors *see* LIGHT PROJECTOR.

Lens control. Refraction of light rays at the boundary between glass and air may be used to control transmitted light in a manner similar to reflector control. Simple thick lenses may be cut away if the cut-away piece is duplicated in the new equivalent surface. Slight irregularities in the field of the light beam occur. These may be of little consequence for searchlights or floodlights, but they would be undesirable for high-quality picture projection. *See* LENS (OPTICS).

Polarization of light. Light rays emitted by common sources may be considered as waves vibrating at all angles in planes at right angles to the direction of the ray. As they pass through some substances or are reflected from specular surfaces, particularly at certain angles, wave components in some directions are absorbed more than those in other directions. Such light is said to be polarized.

Polarized light may be controlled by a transmitting section of a polarizing material. When the absorbing section is oriented to permit passage of the polarized light, minimum control is exercised. If the absorbing section is rotated 90°, the polarized light is essentially completely absorbed. Because sunlight reflected from water or specular surfaces is highly polarized, glare from such sources is effectively controlled by polarized sunglasses or camera filters. Polarized light is used also for detection of strains and defects in glass or plastics. The strained areas appear as color fringes in the material. *See* POLARIZED LIGHT.

Control of primary source. Control of the source of luminous flux can be achieved by controlling the voltage supplied and thus controlling the magnitude of current for some types of light sources or by controlling the time at which the current is permitted to begin on each half cycle of the electric supply for other types of light sources.

For simple incandescent lamps the control of the magnitude of alternating voltage and current can be controlled by variable transformer-type dimmers. Very small silicon controlled rectifiers (SCRs), a modern solid-state development, can accomplish the same control function with equipment which is so small that a unit can be installed in the same space

as would be occupied by a conventional, residential "on-off" switch and with very small inherent heat loss caused by the control itself. SCRs are also used in larger capacities for control of large lighting installations. *See* SEMICONDUCTOR RECTIFIER.

The dimming control of fluorescent lamps or other sources employing a gas or vapor discharge is more involved because of the relatively high striking voltage of the vapor discharge. To achieve a wide range of dimming control, the timing on each half-cycle at which current is permitted to begin can be controlled through electronic means and thus produce dimming to low levels. The complexity and cost of such control equipment, however, is greater than the control equipment needed for incandescent lamp dimming.

Interior lighting design. Attainment of a comfortable visual environment is the ultimate goal of interior lighting design. This requires an adequate distribution of illumination and control of brightness of all visible surfaces to avoid excessive contrasts. The adequate distribution of illumination requires that a sufficient number of lighting sources be installed in a manner that produces an adequate level and a proper distribution of illumination. Brightness is controlled by adequate surface reflection or transmission areas associated with the lighting sources and with proper reflecting surfaces within the room itself. Control should ensure that the brightest surfaces of the room are not more than three times the brightness of the work, and the brightness of the work not more than three times that of the darkest surfaces in the room. When such brightness-ratio limitations are satisfied, glare is absent and a satisfactory visual environment is achieved. Proper application of color in the interior enhances overall appearance.

The function of a building greatly influences the manner in which lighting is applied. A particular visual task may require a certain level of illumination, but the lighting design is also influenced by the economics, appearance, and quality of desired results. For example, application techniques are designated as industrial lighting, commercial lighting, and school lighting. Certain application experience and consumer acceptance assume importance in the methods of achieving the level of illumination and control of brightness contrasts. The *IES Lighting Handbook* treats in detail many of these specialized concepts.

Location and spacing of luminaires. Many interior lighting designs utilize general lighting to distribute luminous flux throughout the room on an imaginary working plane, which is usually 30 in. (77 cm) above the floor. To accomplish a reasonably uniform illumination on this plane, certain limitations in spacing between luminaires and in spacing between luminaires and walls must not be exceeded. The height of luminaires above the plane may be an important factor for some interiors. For indirect-type luminaires the distance from the ceiling to the luminaire must provide a proper distribution of light over the ceiling.

Room coefficients. The shape of a room influences the fraction of the emitted luminous flux that will be received on the working plane. Equations and empirical tabulations have been developed to classify rooms according to their shape. Large, broad rooms are more efficient than small or narrow rooms for transferring luminous flux from the luminaire to the working plane, particularly if the reflectance of the side walls is low.

Lumen method of illumination design. This empirical method of design gives the average illumination on the working plane for a particular number of luminaires arranged in a symmetric pattern in a room. The illumination that may be expected to be maintained in the room is given in Eq. (11), where E is

$$E = \frac{N\phi_L K_u M_f}{A} \tag{11}$$

illumination, N is number of luminaires, ϕ_L is initial luminous flux per luminaire in lumens, K_u is coefficient of utilization, M_f is maintenance factor, and A is area of the room.

Tables of the coefficient of utilization are obtained empirically. Usually these tables pertain to a particular luminaire. The coefficient of utilization is a function of the room coefficient and the reflectances of principally the ceiling and side walls. Also called the flux-of-light method, the lumen method was originally designed as a component method whereby the intensity distribution curve of any luminaire could be resolved into three component curves and each calculated according to an equation similar to that shown. However, modern use of the method is principally through coefficients of utilization for each particular luminaire.

The maintenance factor M_f is a number less than one to account for lamp darkening and the collection of dirt within the luminaire, which reduce the illumination emitted from the luminaire.

Zonal-cavity method. The interreflection of luminous flux within a room is accounted for in the empirical data of the lumen method. An exact mathematical calculation of the distribution of luminous flux within rectangular rooms becomes extremely involved. Computer-derived flux-transfer data have resulted in a design method called the zonal-cavity method, which was the method recognized by the Illuminating Engineering Society (*IES*) as of 1969. Calculation procedures are presented in the *IES Lighting Handbook*.

Luminous architectural elements. The original emphasis in the lumen method was upon discrete luminaires arranged symmetrically within the room. The development of the interreflection method of design presumed surfaces of initial luminance on ceiling, walls, and floor as the sources of light within the room. Ingenuity in the application of both methods permits them to be applied both to designs involving luminaires within the room and to designs incorporating many types of luminous architectural elements, such as louvered ceilings, luminous ceilings, and cove systems of lighting.

Natural lighting. Daylight may be an important factor in building design. For comfortable visual conditions any openings for the admission of daylight

are planned in position, size, and shape as carefully as any other part of a building. A window can be treated as any other type of light source and included as a part of the illumination and luminance design according to the conventional design procedures. It is usually necessary to provide electric lighting systems when daylight is not available.

Supplementary lighting. Local lighting, such as spotlighting, is frequently used to supplement general lighting where a high level of illumination is desirable on an isolated task.

Exterior lighting design. The nightime lighting of exterior areas serves many useful purposes. Such lighting extends into the fields of advertising and recreational pursuits as well as utilitarian fields, such as the lighting of industrial yards, parking areas, airports, streets, and highways. The *IES Lighting Handbook* and many textbooks treat in detail many of these specialized applications.

Recommended standards. Illumination levels have been established by the IES to bring all visual tasks to the same level of visibility for normal observers. Some visual tasks are more arduous than others, and higher illuminations provide a compensating factor in bringing such tasks to a satisfactory level of visibility. Standards are established not only for interior lighting designs but also for exterior systems, including recreational lighting. For tables of recommended illumination levels, consult the references in the bibliography.

Warren B. Boast

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; B. Fitt and J. Thornley, *The Control of Light*, 1993; C. Gardner and B. Hannaford, *Lighting Design: An Introduction*, 1993; Illumination Engineering Society, *IES Lighting Handbook, Reference and Application Volume*, 1993; A. Thumann, *Lighting Efficiency Applications*, 2d ed., 1991; L. H. Watson, *Lighting Design Handbook*, 1990.

Image processing

Manipulating data in the form of an image through several possible techniques. An image is usually interpreted as a two-dimensional array of brightness values, and is most familiarly represented by such patterns as those of a photographic print, slide, television screen, or movie screen. An image can be processed optically, or digitally with a computer. This article focuses on digital processing of an image to perform one of five tasks: enhancement, restoration, compression, image fusion, or automated image analysis.

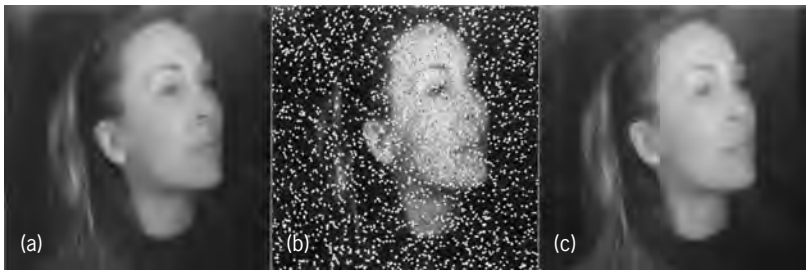

Fig. 1. Noise smoothing. (*a*) Original 256 × 256 pixel image. (*b*) Original image corrupted by spike noise. (*c*) Corrupted image filtered by a 3 × 3 pixel median filter.

Digitization. To digitally process an image, it is first necessary to reduce the image to a series of numbers that can be manipulated by the computer. If, for example, the image is in the form of a photographic transparency, this digitization can be performed by a scanning device that moves a spot of light over the transparency in a rasterlike fashion (from side to side and top to bottom) similar to a television scan. When the light falls on a particular area of the transparency, the amount of light transmitted at that point can be recorded by a photodetector, the signal digitized to, say, eight bits (yielding a possible 2^8, or 256, gray levels), and stored. The spot is then moved a distance usually equal to its width, and the process is repeated, recording the entire transparency in this way as a long string of numbers. The smaller the spot and the finer the scan (up to the limit of the film grain), the more accurate the description of the transparency will be. Each number representing the brightness value of the image at a particular location is called a picture element, or pixel. A typical digitized image may have 512 × 512 or roughly 250,000 pixels, although much larger images are becoming common.

Once the image has been digitized, there are three basic operations that can be performed on it in the computer. For a point operation, a pixel value in the output image depends on a single pixel value in the input image. For local operations, several neighboring pixels in the input image determine the value of an output image pixel. In a global operation, all of the input image pixels contribute to an output image pixel value. These operations, taken singly or in combination, are the means by which the image is enhanced, restored, or compressed.

Enhancement. An image is enhanced when it is modified so that the information it contains is more clearly evident, but enhancement can also include making the image more visually appealing. An example is noise smoothing. Given an image (**Fig. 1***a*), by replacing the value of randomly chosen pixels by 150 (bright gray) or 0 (black), a noisy image (Fig. 1*b*) results. This kind of noise might be produced, for example, by a defective television camera. To smooth the noisy image, median filtering can be applied with a 3 × 3 pixel window. This means that the value of every pixel in the noisy image is recorded, along with the values of its nearest eight neighbors. These nine numbers are then ordered according to size, and the median is selected as the value for the pixel in the new image. As the 3 × 3 window is moved one pixel at a time across the noisy image, the filtered image is formed (Fig. 1*c*). This is a local operation, since the value of a pixel in the new image depends on the values of nine pixels in the old image.

Another example of enhancement is contrast manipulation, where each pixel's value in the new image depends solely on that pixel's value in the old image; in other words, this is a point operation.

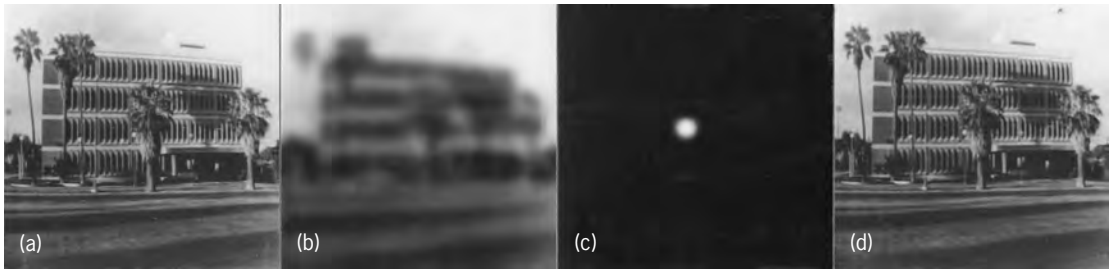

Fig. 2. Example of restoration. (*a*) Original 256 × 256 pixel image (*courtesy of Richard Murphy*). (*b*) Blurred version of the original image, produced by computer simulation. (*c*) The image of a point subject to the same blurring that produced the blurred version of the original image. (*d*) The results of the inverse filter applied to the blurred version of the original image.

Contrast manipulation is commonly performed by adjusting the brightness and contrast controls on a television set, or by controlling the exposure and development time in printmaking. Another point operation is that of pseudocoloring a black-and-white image, by assigning arbitrary colors to the gray levels. This technique is popular in thermography (the imaging of heat), where hotter objects (with high pixel values) are assigned one color (for example, red), and cool objects (with low pixel values) are assigned another color (for example, blue), with other colors assigned to intermediate values.

Restoration. The aim of restoration is also to improve the image, but unlike enhancement, knowledge of how the image was formed is used in an attempt to retrieve the ideal (uncorrupted) image (**Fig. 2***a*). Any image-forming system is not perfect, and will introduce artifacts (for example, blurring, aberrations) into the final image that would not be present in an ideal image. A computer can simulate an out-of-focus lens, producing a blurred image (Fig. 2*b*). The image of a point object through this system (Fig. 2*c*) is called the point spread function of the system. Such blurring can be referred to as a local operation, in that the value of a pixel in the blurred image is an average of the values of the pixels around (and including) that pixel location in the original image. The blurring process is mathematically referred to as a convolution of the original object with the point spread function.

Another point spread function, called a filter, can be constructed that undoes the blurring caused by the original point spread function. Such a point spread function has both positive and negative values, even though the blurring point spread function has only positive values. By imaging the blurred image (Fig. 2*b*) with the filter point spread function, the restored image (Fig. 2*d*) results. The filter point spread function is spread out more than the blurring point spread function, bringing more pixels into the averaging process. This is an example of a global operation, since perhaps all of the pixels of the blurred image can contribute to the value of a single pixel in the restored image. This type of deblurring is called inverse filtering, and is sensitive to the presence of noise in the blurred image. By modifying the deblurring filter according to the properties of the noise, performance can be improved. An example of the need to deblur images from an optical system is the Hubble Space Telescope before its spherical aberration was corrected with new optics. *See* SATELLITE ASTRONOMY.

Mathematical transformations can be performed on images to facilitate deblurring, as well as other operations. The Fourier transform decomposes the image into two-dimensional sine waves of varying frequency, direction, and phase. Convolution of an image with a point spread function (the imaging process) can be carried out in the Fourier domain by simply multiplying the Fourier transform of the image by the Fourier transform of the point spread function, and taking the inverse Fourier transform of this product to get the final image. *See* FOURIER SERIES AND TRANSFORMS.

Compression. Compression is a way of representing an image by fewer numbers, at the same time minimizing the degradation of the information contained in the image. Compression is important because of the large quantities of digital imagery that are sent electronically and stored. Digital high-definition television relies heavily on image compression to enable transmission and display of large-format color images.

Compression can be performed in the image space or in transform space (for example, Fourier). In wavelet decomposition, a particularly useful type of transformation, the image is broken (that is, decomposed) into successively smaller representations of itself by a downsizing process, with each representation containing information about the original image at different size scales. The smaller representations are then wavelet-transformed. The wavelet transform reduces the correlation and dynamic range of the image representations, so that after the transformation each representation requires considerably fewer bits to encode the information that it contains. *See* WAVELETS.

All of the smaller representations (also referred to as subbands) are included in the compressed image. The dynamic range of each subband is different; smaller subbands have a larger dynamic range (that is, require more bits per pixel) than larger subbands, and the compression algorithm automatically takes this into account.

Once the image is compressed for storage or transmission, it must be uncompressed for use, by the inverse of the compression operations. There is a trade-off between the amount of compression and

(a) (b) (c)

Fig. 3. Example of compression. (*a*) Original 1024 × 768 pixel image, at 3.3 bits per pixel (bpp), typical of the compression provided by digital cameras. (*b*) Compressed version of the original image using JPEG compression, at 0.7 bpp. There is little visual degradation, yet the file size is over 4 times smaller. (*c*) Further compressing the image to 0.4 bpp; noticeable degradation has occurred with this lossy algorithm.

the quality of the uncompressed image. High compression rates (giving rise to so-called lossy compression) are acceptable with television images, for example. However, where high image quality must be preserved (lossless compression, as in diagnostic medical images), only compression rates as low as three to four may be acceptable. Figure 3 is an example showing the tradeoff between image compression and image quality, for the JPEG compression scheme. *See* DATA COMPRESSION.

Many image formats have been developed for the storage and transmission of image data, and both lossy and lossless compression methods are incorporated into them. Familiar examples of such formats are JPEG (lossy), GIF (lossless LZW algorithm, limited to 256 colors), PNG (lossless, an open-source updated GIF format supporting millions of colors), TIFF (lossy or lossless), and Exif (a common format used with digital cameras that supports JPEG compression).

Image fusion. This term refers to the combination of information from two different images to produce a new image that provides more information to the viewer than either original image by itself. Such combination can provide several advantages: low-resolution infrared information (which provides information about the temperature of an object) can be mapped onto a higher-resolution visible image or a road map to aid in object location and recognition; objects invisible in one image may be visible in the other, so the fused image allows object relationships to be clarified; and the addition of elevation data to a visible, infrared, or radar image allows a three-dimensional stereo view of the scene to be produced. In medicine, low-resolution (5–10-mm) positron emission tomography (PET) data, which represent the functioning of the brain, can be combined with high-resolution (1-mm) magnetic resonance image (MRI) data that represent anatomical information. *See* MEDICAL IMAGING.

A basic requirement of image fusion is that the original images be registered. This usually means that a set of control points must be identified in each set of images that correspond to the same objects. One image is called the reference image, and the other image is geometrically distorted until the position of its control points matches the control points in the reference image.

Simple techniques for image fusion add or multiply the pixel values of the registered images together to form the new image. More sophisticated techniques involve nonlinear combinations of pixel values or involve transformations of the registered images to form the fusion product. The coherent fusion of two synthetic-aperture radar (SAR) images of the ground taken days apart can allow the imaging of the motion of ice flows. *See* RADAR; REMOTE SENSING.

Geometrical morphing of one image to another can be loosely considered another kind of image fusion, in that a temporal progression of one image into another is formed by geometrically distorting one image into another, with corresponding grayscale and color changes to the overlapping pixels. The effect is to make one image flow into another. This is used extensively in the advertisement and entertainment industry, though not as frequently in science and engineering.

Automated image analysis. This technique, also referred to as image understanding and image exploitation, involves automatically extracting information from an image. This information extraction can take many forms, such as identifying a tumor in a medical image; finding a flaw in an integrated-circuit chip; looking for misshapen candy on an assembly line; spotting a small hole in a sheet of metal moving at high speed; identifying an enemy tank, airplane, or missile in milliseconds; or identifying a person's face.

In the automatic extraction of image information, first the image is usually enhanced or restored to facilitate the analysis. Next, measurable quantities, called features, are extracted from the image which represent the structures that are being sought.

One use of the extracted features is to label each pixel in the image, forming a segmentation of the image. A segmentation represents a homogeneous area of the image (that is, an area with similar feature characteristics) as a single color or gray level. Areas with different properties, based upon their features, are represented by different colors. Once all the pixels of an image are labeled in this manner, this new image can be presented to a human observer for a decision, or higher-level automated decision

processing can take place. For example, in computer-aided diagnosis in medicine, a tumor in an MRI of the brain can be labeled a different color than the surrounding healthy tissue, based upon subtle textural differences of the tumor tissue. This automatic labeling can assist the radiologist in diagnosis or can be the jumping-off point for further automated processing. *See* CHARACTER RECOGNITION; COMPUTER VISION.

Applications. Image processing is an active area of research in such diverse fields as medicine, astronomy, microscopy, seismology, defense, industrial quality control, security, and the publication and entertainment industries. The concept of an image has expanded to include three-dimensional data sets (volume images), and even four-dimensional volume-time data sets. An example of the latter is a volume image of a beating heart, obtainable with x-ray computed tomography (CT). CT, PET, single-photon emission computed tomography (SPECT), MRI, ultrasound, SAR, confocal microscopy, scanning tunneling microscopy, atomic force microscopy, and other modalities have been developed to provide digitized images directly. Digital images are widely available from the Internet, CD-ROMs, and inexpensive charge-coupled-device (CCD) cameras, scanners, and frame grabbers. Software for manipulating images is also widely available. *See* CHARGE-COUPLED DEVICES; COMPACT DISK; COMPUTER STORAGE TECHNOLOGY; COMPUTERIZED TOMOGRAPHY; CONFOCAL MICROSCOPY; DATA COMMUNICATIONS; DIGITAL COMPUTER; ELECTRON MICROSCOPE; MEDICAL ULTRASONIC TOMOGRAPHY; NUCLEAR MEDICINE; SCANNING TUNNELING MICROSCOPE; TELEVISION CAMERA.

Warren E. Smith

Bibliography. G. A. Baxes, *Digital Image Processing: Principles and Applications*, 1994; K. R. Castleman, *Digital Image Processing*, 1996; R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*, 2003; W. K. Pratt, *Digital Image Processing: PIKS Inside*, 3d ed., 2001; J. C. Russ, *The Image Processing Handbook*, 4th ed., 2002.

Image tube (astronomy)

A photoelectric device for intensifying faint astronomical images. In these devices a photoemissive surface, called the photocathode, emits electrons through the photoelectric effect. In most tubes the photocathode is semitransparent and is deposited on the inside of a transparent window that is mounted on the end of an evacuated glass or ceramic cylinder. When light from a telescope or spectrograph is imaged on the photocathode, electrons are ejected into the vacuum inside the tube. Electric fields or electric and magnetic fields in combination then accelerate and direct the photoelectrons through the device. In an image intensifier tube the photoelectrons are ultimately reimaged on a phosphor-coated output window that converts them back into visible light. An external television camera or charge-coupled device (CCD) is used to record the image. Optical gain is provided by the phosphor, which may release as many as a thousand photons for each photoelectron which crashes into it. In other devices the electron image is projected onto a solid-state diode array or a grid of anodes that is built directly into the tube. *See* CATHODE-RAY TUBE; PHOSPHORESCENCE; PHOTOELECTRIC DEVICES; PHOTOEMISSION; PHOTON.

The quantum efficiency of an image tube is the probability that an incoming photon of light will release a detectable photoelectron inside the tube. Although a single photon of light is all that is necessary to create a photoelectron, some photons pass through the photocathode without being absorbed. Others are absorbed but do not produce a free electron. In spite of these losses, photocathodes with quantum efficiencies above 20% are routinely fabricated. Bialkali and multialkali photocathodes are useful in the optical and ultraviolet and are composed of antimony and two or more alkali metals. Solar-blind photocathodes are made from cesium and iodine or tellurium. They have high quantum efficiencies in the ultraviolet but are insensitive at optical wavelengths. All photocathodes emit a few electrons even in the absence of light. This dark emission is a source of noise and must be taken into account when observing faint sources. Because modern CCDs can have quantum efficiencies over 60% in the optical wavelength range, photoemissive devices have largely been phased out of ground-based astronomy. The extended red sensitivity of CCDs effectively precludes their use in the far-ultraviolet, however, since even the best ultraviolet filters have substantial red leaks. Photoemissive detectors have thus been flown on virtually all ultraviolet astronomy satellites such as the *International Ultraviolet Explorer* and the Hubble Space Telescope. *See* CHARGE-COUPLED DEVICES; ULTRAVIOLET ASTRONOMY.

Electrostatically focused tubes. Each tube in an electrostatically focused image intensifier (**Fig. 1**) is constructed with fiber-optic input and output windows that are flat on the outside and spherical on the inside. Light falling on the front of each detector is piped to the photocathode by the optical fibers of the input window. The ejected photoelectrons then pass through a small hole in the tip of a conical electrode and come to focus on a phosphor screen that has been deposited on the inner surface of the output window. A reflective coating over the phosphor prevents light from traveling back to the photocathode. High gain is achieved by having successive stages of amplification. While electrostatically focused image tubes have low dark emission rates, they have poor distortion characteristics. They are also not useful in the ultraviolet because of the fiber-optic input window. *See* OPTICAL FIBERS.

Proximity-focused tubes. Proximity-focused image intensifiers are very short and have plane-parallel input and output windows. There are no electrodes for imaging the photoelectrons; a strong electric field simply accelerates them into the phosphor before any significant defocusing can take place. While proximity-focused image tubes have almost no image

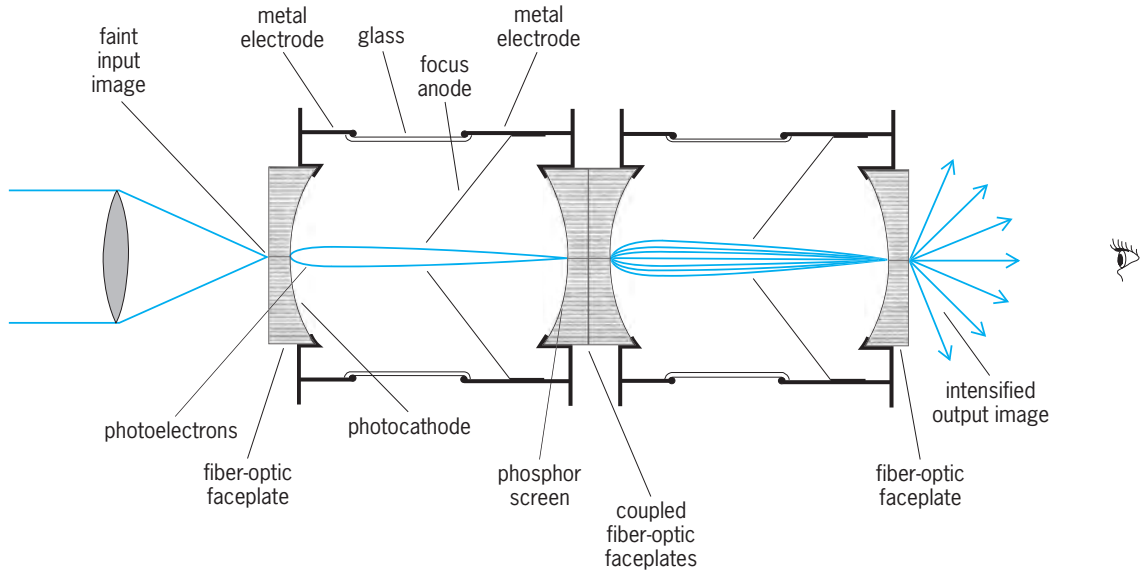

Fig. 1. Two-stage electrostatically focused image tube, fiber-optically coupled.

distortion, they have high dark emission rates and relatively low gain.

Microchannel-plate tubes. A microchannel plate (**Fig. 2**) consists of an array of closely spaced tubes, each of which is coated on the inside with a secondary-emitting material. When a free electron enters a channel, it eventually collides with the tube wall and releases several secondary electrons. These electrons then travel farther down the tube and produce secondary electrons of their own. Since this process is repeated several times down the channel, thousands of electrons exit the rear face of the microchannel plate. A microchannel plate can be added to a proximity-focused image intensifier to obtain higher gain at an acceptable level of dark emission. In such an application, the microchannel plate is simply sandwiched between the photocathode and the phosphor. For space applications, the photocathode is usually deposited directly on the microchannel plate. Since this technique eliminates the photocathode window, such a device can be used in the extreme ultraviolet. In a multianode microchannel plate device, the phosphor screen is replaced by a grid of anodes that records the position of each pulse of electrons that exits the plate.

Magnetically focused tubes. Magnetically focused image tubes use electric and magnetic fields for electron imaging and have plane-parallel input and output windows. Electrons from the photocathode are accelerated down the tube by a series of ring-shaped electrodes. As the electrons are accelerated, an axial magnetic field from an external permanent magnet or solenoid forces the photoelectrons to follow spiral paths that reconverge at the output window of the tube. In an image intensifier the electron image is converted back into light. In other types of tubes the electron image is focused directly on a silicon diode array, and each incoming electron produces a large number of electron-hole pairs in the silicon instead of a burst of photons. While magnetically focused devices have some distortion, they usually have very low dark emission rates. *See* ASTRONOMICAL PHOTOGRAPHY; ELECTRON MOTION IN VACUUM.

Richard G. Allen

Bibliography. A. Boksenberg and D. L. Crawford (eds.), *Proceedings of SPIE*, vol. 445: *Instrumentation in Astronomy V*, 1983; D. L. Crawford (ed.), *Proceedings of SPIE*, vol. 331: *Instrumentation in Astronomy IV*, 1982; P. W. Hawkes (ed.), *Advances in Electronics and Electron Physics*, vols. 64A and 64B, 1985.

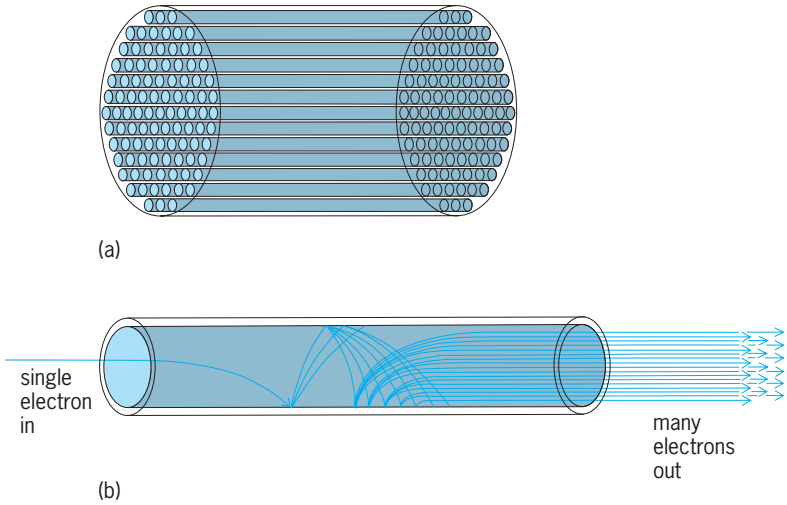

Fig. 2. Microchannel plate showing (*a*) array of tubes and (*b*) details of a single channel.

Immittance

The impedance or admittance of an alternating-current circuit. It is sometimes convenient to use the term immittance when referring to a complex number which may be either the impedance (ratio of voltage to current) or the admittance (ratio of

current to voltage) of an electrical circuit. The units of impedance and admittance are, of course, different and so units cannot be assigned to an immittance. However, in certain theoretical work it may be necessary to deal with general functions which afterward will be specialized to become either an impedance or an admittance by the assignment of suitable units; in such cases it is convenient to refer to the functions as immittances. *See* ADMITTANCE; ALTERNATING-CURRENT CIRCUIT THEORY; ELECTRICAL IMPEDANCE. J. O. Scanlan

Immune complex disease

Local or systemic tissue injury caused by the vascular deposition of products of antigen-antibody interaction, termed immune complexes. Immune complex formation with specific antibodies causes the inactivation of potentially dangerous antigenic substances and their elimination by phagocytic cells. There are harmful consequences only when immune complexes deposit in tissues, inciting various mediators of inflammation, leading to injury and disease. When the reaction takes place in the extravascular fluids near the site of origin of the antigen (by injection, secretion, and such), focal injury can occur, as exemplified by the Arthus reaction or such conditions as experimental immune thyroiditis. Systemic disease may occur when soluble antigens combine with antibodies in the vascular compartment, forming circulating immune complexes that are trapped nonspecifically in the vascular beds of various organs, causing such clinical diseases as serum sickness or systemic lupus erythematosus with vasculitis and glomerulonephritis. The term immune complex disease usually signifies this systemic immune complex formation and vascular deposition. In the widely used classification originated by Gell and Coombs, immune complex disease is Type III immunopathology. *See* ANTIGEN-ANTIBODY REACTION; CONNECTIVE TISSUE DISEASE.

Pathophysiology. Circulating immune complex disease occurs when the host's antibody production, relative to the amount of antigens, is inadequate for prompt elimination of antigen. Normally, excess amounts of antibody are formed which generate large immune complexes that are removed very rapidly from the circulation and are disposed of by the mononuclear phagocytic system. If the antibody response is very poor, only a few very small complexes are formed which are not prone to vascular deposition. When the relative antibody production is such that complexes of intermediate size form, vascular trapping can occur and injury results from the effects of inflammation. In addition to immune complex size, other factors influence vessel deposition, including the efficiency of systemic clearance of immune complexes, the hemodynamics of blood flow, and vasoactive amine-influenced changes in vascular permeability. Through dynamic equilibrium, continual modification of the deposits occurs as antigen and antibody fluctuate in the body fluids.

Human immune complex disease. Circulating immune complexes play a prominent role in several types of human disease. They can follow the production of antibody to exogenous antigens (iatrogenic, infectious, and so on), or to endogenous ones during an autoimmune response. Some individuals with streptococcal infections of the throat or skin develop poststreptococcal glomerulonephritis with evidence of circulating immune complexes, and complex deposition in the kidney. Although usually short-lived, this condition can progress to renal insufficiency. The more chronic infection with hepatitis B virus can result in immune complex disease. A transient serum sickness–like illness may occur as a prodrome to the hepatic illness. In individuals with prolonged circulation of hepatitis B viral antigens, severe immune complex vasculitis and glomerulonephritis may develop. Chronic parasitic diseases such as malaria and schistosomiasis are often complicated by complex-related problems. Foreign serum and drugs used in treatment or illicitly can cause immune complex injury.

Immune complexes are often responsible for some of the systemic manifestations of autoimmune diseases, for example, in systemic lupus erythematosus. Tissue biopsy is used to identify complex deposition. Vasculitis (inflammation of the blood vessel walls) results from the vascular immune complex deposit in individuals with systemic lupus erythematosus as well as in those with other so-called autoimmune diseases such as rheumatoid arthritis and polyarteritis nodosa. In most cases of vasculitis, the offending antigen is unknown, although nuclear antigens including DNA are involved in lupus, as is hepatitis B antigen in some cases of polyarteritis nodosa.

Circulating complexes occur in individuals with autoimmune thyroiditis as well as individuals with neoplasms, most commonly carcinomas. Signs of immune complex tissue damage are infrequent in individuals with neoplasia except for occasional glomerulonephritis in which tumor-associated antigens (for example, carcinoembryonic antigen) have been demonstrated in immune complex deposits in glomeruli.

Because of its unique role as a filter, the kidney is the most common organ in which immune complexes deposit. Most immunologic glomerular injury is considered to be a consequence of complex deposition, and this mechanism is recognized in tubulointerstitial renal disease as well. In some immune complex-mediated glomerular diseases, the relevant antigens are known, as in postinfectious disease, malignancy, and systemic lupus erythematosus.

Diagnosis. The diagnosis of immune complex disease in compatible clinical situations is aided by tests including the detection of complement consumption and the presence of anti-DNA antibodies in systemic lupus erythematosus.

Tissue deposition of complexes is routinely determined by immunofluorescence microscopy using fluorescein conjugated antibodies specific for components of the immune complex, namely various classes of immunoglobulin, complement

components, and, when known, the antigen itself. The random deposition of complexes in tissue results in irregular granular collections of immunoglobulin within vessel walls or glomeruli. Although positive immunofluorescent demonstration of granular immunoglobulin and complement deposits in tissue is suggestive, it is not pathognomonic of immune complex localization, unless antigen and antibody can be demonstrated in a similar distribution. Alternately, antibody may be eluted from tissue and tested for antigen specificity. Electron microscopy often reveals electron-dense deposits corresponding to sites of complex deposition, allowing detailed morphologic assessment. Immunoelectron microscopy using labeled antibodies (such as ferritin or peroxidase) may provide additional information as to the precise location of deposited complexes. *See* IMMUNOFLUORESCENCE.

A number of methods have been described which detect circulating immune complexes with varying degrees of success. In general, the assays more often detect complexes in acute forms of disease when the levels of complex are presumably higher than in more indolent disease processes. The tests may be divided into those based on physicochemical properties and on immunochemical properties. Physicochemical methods take advantage of size and configurational changes in complexed immunoglobulin. Ultracentrifugation and column chromatography are relatively insensitive in separating the immune complexes from smaller uncomplexed immunoglobulin, so that large amounts of complexes must be present for detection. Precipitation of immune complexes with poly(ethylene glycol) of appropriate concentration is relatively more sensitive. Cryoprecipitation may also detect some immune complexes.

The immunochemical methods tend to be more sensitive and employ the ability of complexed immunoglobulin to interact with complement, bovine conglutinin, rheumatoid factor, anti-antibody sera or cellular elements. The complement-based assays, most often relying on binding of a portion of the first complement component, Clq, to the immune complexes (in solid-phase or fluid-phase radioimmunoassay) are relatively sensitive. *See* COMPLEMENT; IMMUNOGLOBULIN.

Assays in which immune complexes cause platelet aggregation or bind to immunoglobulin Fc, or complement, receptors on cells have been used quite successfully to detect small quantities of complex in sera. The most common cell line used is the lymphoblastoid Raji cell line derived from a person with Burkitt's lymphoma. This line binds complexes attached by the complement components. The immune complexes are subsequently quantified by radioimmune assay. Advantage may be taken of the different reactivities of the various assays for identification and more importantly isolation and characterization of the immune complex. *See* RADIOIMMUNOASSAY.

Treatment. Treatment of immune complex disease can be divided into nonspecific and more specific modalities. Primary among the specific measures is the identification and elimination of the offending antigen. This may be possible with some infections when specific therapies are available, and in certain instances where the antigenic source can be removed, such as a neoplasm. More frequently, nonspecific anti-inflammatory (corticosteroids) and immunosuppressive agents (such as cyclophosphamide and azathioprine) are used to attempt to blunt the person's immune response, thereby lessening the amount of immune complexes produced. Removal of the circulating plasma by plasmapheresis is also being tried as a way of removing circulatory immune complexes and perhaps circulating mediators of inflammation as well. *See* AUTOIMMUNITY; IMMUNOLOGY.

Edward H. Cole; Curtis B. Wilson

Bibliography. E. Braunwald et al. (eds.), *Harrison's Principles of Internal Medicine*, 15th ed., 2001; B. M. Brenner and F. C. Rector, Jr. (eds.), *The Kidney*, 2 vols., 6th ed., 2000; C. G. Cochrane and D. Koffler, Immune complex disease in experimental animals and man, *Adv. Immunol.*, 16:185–264, 1973; A. N. Theofilopoulos and F. J. Dixon, The biology and detection of immune complexes, *Adv. Immunol.*, 28:89–220, 1979; C. B. Wilson, B. M. Brenner, and J. H. Stein (eds.), *Contemporary Issues in Nephrology*, vol. 3, 1979.

Immunity

A state of resistance to an agent, the pathogen, that normally produces an infection. Pathogens include microorganisms such as bacteria and viruses, as well as larger parasites. The immune response that generates immunity is also responsible in some situations for allergies, delayed hypersensitivity states, autoimmune disease, and transplant rejection. *See* ALLERGY; AUTOIMMUNITY; TRANSPLANTATION BIOLOGY.

Immunity is engendered by the host immune system, reacting in very specific ways to foreign components (such as proteins) of particular parasites or infective agents. It is influenced by many factors, including the environment, inherited genes, and acquired characteristics. Reaction to a pathogen is through a nonadaptive or innate response as well as an adaptive immune response. The innate response is not improved by repeated encounters with the pathogen. An adaptive response is characterized by specificity and memory: if reinfection occurs, the host will mount an enhanced response.

The components of the pathogen that give rise to an immune response, to which antibodies are generated, are called antigens. There are two types of specific responses to an antigen, antibodies and the cellular response. Antibodies help to neutralize the infectious agent by specifically binding it. A series of proteins in the blood (called complement) act in conjunction with antibodies to destroy pathogenic bacteria. In the cellular response, cytotoxic T cells are recruited to kill cells infected with intracellular agents such as viruses. Helper T cells may also be generated, which influence B cells to produce appropriate

antibodies. Inflammatory responses and activation of other kinds of cells, such as macrophages, in conjunction with lymphocytes, is another important aspect of the immune response, as in delayed hypersensitivity. This kind of response seems to be common in certain chronic infections. *See* ANTIBODY; ANTIGEN; COMPLEMENT.

Complex immune systems (antibody and specific cellular responses) have been demonstrated in mammals, birds, amphibians, and fish, and are probably restricted to vertebrates. The evolutionary strategy of eliminating diseased individuals is not feasible in vertebrates because they have so few progeny. Many other organisms have self and nonself discrimination mechanisms and primitive defense strategies, most likely lacking the flexibility as well as the specificity and memory of vertebrate systems.

Natural or innate immunity. There are natural barriers to infection, both physical and physiological, which are known collectively as innate immunity, and include the effects of certain cells (macrophages, neutrophils and natural killer cells) and substances such as serum proteins, cytokines, complement, lectins, and lipid-binding proteins. The skin or mucous membranes of the respiratory tract are obvious barriers and may contain bacteriostatic or bactericidal agents (such as lysozyme and spermine) that delay widespread infection until other defenses can be mobilized. Fatty acids provide a good example of natural immunity, such as in foot rot (athlete's foot) in which infection tends to occur between the small toes where there are no sebaceous glands.

If organisms manage to enter tissues, they are often recognized by molecules present in serum and by receptors on cells. Bacterial cell walls, for example, contain substances such as lipopolysaccharides that activate the complement pathway or trigger phagocytic cells. Some receptors on cells, called Toll receptors, are related to proteins found on plants or insects that trigger the production of antimicrobial substances. These innate responses to infectious agents are obviously ancient and may have originated before common ancestors of humans, plants, and insects.

Host range is dramatic in its specificity. Animals and plants are generally not susceptible to each other's pathogens. Within each kingdom, infectious agents are usually adapted to affect a restricted range of species. For example, mice are not known to be susceptible to pneumococcal pneumonia under natural conditions. If a small number of organisms are injected, bypassing the natural barriers, a fatal infection results. Domestic cats never contract measles. The lack of infectivity of rodents with human immunodeficiency virus (HIV) has meant that monkeys have to be used to study the behavior of this retrovirus. In some cases, disparate species lack the receptors at their cell surfaces for the agent to gain access. In other situations, the agent is not able to adapt the host's biochemistry to its own ends, for self-replication, or the environment in the host lacks essential growth factors. Active destruction of anthrax organisms apparently accounts for the natural resistance of dogs to the disease. The health of the host and environmental conditions may also make a difference to susceptibility. This is readily apparent in fish that succumb to fungal infections if their environment deteriorates. Genetic factors have an influence on susceptibility. Some of these genes have been identified, in particular the genes of the major histocompatibility complex which are involved in susceptibility to autoimmune diseases as well as some infectious disorders. There is a case for natural killer cells, as well as macrophages, being involved in natural immunity to some diseases; Chagas' disease (infection with a protozoon) is an example. Immunity to diverse diseases such as leprosy, acquired immune deficiency syndrome (AIDS), malaria, and allergies are clearly influenced by genetic background. In the plant kingdom, immunity involves a different set of mechanisms; for example, disease-resistant crops have a genetic basis. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); HISTOCOMPATIBILITY.

Once parasites gain entry, phagocytic cells attack them. They may engulf and destroy organisms directly, or they may need other factors such as antibody, complement, or lymphokines, secreted by lymphocytes, which enhance the ability of the phagocytes to take up antigenic material. In many cases these cells are responsible for alerting cells involved in active immunity so there is two-way communication between the innate and adaptive responses. Certain organisms, such as gonococcus, remain in the cell and, being refractory to treatment, invoke a chronic infection. Nonspecific agglutinins and lysins may also influence early development of the parasite once in the host. For obvious reasons, the innate system of host defense operates during the early phases of an immune response, and in higher organisms is inextricably associated with the adaptive immune response. *See* PHAGOCYTOSIS.

Adaptive immune response. Adaptive immunity is effected in part by lymphocytes, which are derived from bone marrow stem cells. Lymphocytes are of two types: B cells, which develop in the bone marrow or fetal liver and may mature into antibody-producing plasma cells, and T cells, which develop in the thymus. T cells have a number of functions, which include helping B cells to produce antibody, killing virus-infected cells, regulating the level of immune response, and controlling the activities of other effector cells such as macrophages.

Each lymphocyte carries a different surface receptor that can recognize a particular antigen. The antigen receptor expressed by B cells consists of membrane-bound antibody of the specificity that it will eventually secrete; B cells can recognize unmodified antigen. However, T cells recognize antigen only when parts of it are complexed with a molecule of the major histocompatibility complex. The principle of the adaptive immune response is clonal recognition: each lymphocyte recognizes only one antigenic structure, and only those cells stimulated by antigen respond. Initially, in the primary response, there are few lymphocytes with the appropriate receptor for an antigen, but these cells proliferate. If the

antigen is encountered again, there will be a proportionally amplified and more rapid response. Primed lymphocytes either differentiate into immune effector cells or form an expanded pool of memory cells that respond to a secondary challenge with the same antigen.

The acquired or adaptive immune response is characterized by exquisite specificity such that even small pieces of foreign proteins can be recognized. This specificity is achieved by the receptors on T cells and B cells as well as antibodies that are secreted by activated B cells. The genes for the receptors are arranged in multiple small pieces that come together to make novel combinations, by somatic recombination. Each T or B cell makes receptors specific to a single antigen. Those cells with receptors that bind to the foreign protein and not to self tissues are selected out of a large pool of cells. For T cells, this process takes place in the thymus. The extreme diversity of T- and B-cell receptors means that an almost infinite number of antigens can be recognized. It has been calculated that potentially about 3×10^{22} different T-cell receptors are made in an individual. Even if 99% of these are eliminated because they bind to self tissues, 3×10^{20} would still be available.

Inflammation takes place to activate immune mechanisms and to eliminate thoroughly the source of infection. Of prime importance is the complement system, which consists of tens of serum proteins. A variety of cells are activated, including mast cells and macrophages. Inflammation results in local attraction of immune cells, increased blood supply, and increased vascular permeability. *See* CELLULAR IMMUNOLOGY.

Tolerance. The immune system is primed to react against foreign antigens while avoiding responses to self tissue by immunological tolerance. The neonatal immune system is particularly susceptible to tolerance induction. Although most T cells which might activate against host proteins are deleted in the thymus, these self-reactive cells are not always destroyed, and a chronic autoimmune condition results. *See* ACQUIRED IMMUNOLOGICAL TOLERANCE.

Immunization. Adaptive immunity is characterized by the ability to respond more rapidly and more intensely when encountering a pathogen for a second time, a feature known as immunological memory. This permits successful vaccination and prevents reinfection with pathogens that have been successfully repelled by an adaptive immune response.

Vaccination against smallpox was introduced by Edward Jenner in 1796, using a living agent, calf pox, that is antigenically related to smallpox. It was well known that a case of smallpox conferred life-long immunity. In fact, inhalation or dermal inoculation with smallpox, resulting in mild disease and consequent immunity, had been common in parts of Europe for decades but was not readily accepted by the medical establishment. Modern inoculation programs have been so effective that smallpox seems to have been completely eradicated.

Vaccination has been one of immunology's greatest practical successes, and mass immunization programs have led to the virtual eradication of several very serious diseases, although not always on a worldwide scale. Living attenuated vaccines against a variety of agents, including poliomyelitis, tuberculosis, yellow fever, and bubonic plague, have been used effectively. Nonliving vaccines are commonly used for prevention of bacterial diseases such as pertussis, typhoid, and cholera as well as some viral diseases such as influenza and bacterial toxins such as diphtheria and tetanus. *See* CANCER (MEDICINE); VACCINATION.

Passive immunization. Protective levels of antibody are not formed until some time after birth, and to compensate for this there is passive transfer of antibody across the placenta. Alternatively, in some animals antibody is transferred in the first milk (colostrum). Antibody may also be passively transferred artificially, for example, with a concentrated preparation of human serum gamma globulin containing antibodies against hepatitis. Protection is temporary. Horse serum is used for passive protection against snake venom. Serum from the same (homologous) species is tolerated, but heterologous serum is rapidly eliminated and may produce serum sickness. On repeated administration, a sensitized individual may experience anaphylactic shock, which in some cases is fatal. Cellular immunity can also be transferred, particularly in experimental animal situations when graft and host reactions to foreign tissue invariably occur unless strain tissue types are identical.

Hypersensitivity reactions. Reactions mediated by antibody to foreign antigens are not always beneficial to the host. Some antigen-antibody reactions produce extensive host pathology and can result in alarmingly rapid death. Such hypersensitivity reactions include anaphylaxis; serum sickness and Arthus reactions [mediated by conventional antibodies such as immunoglobulin G (IgG)]; atopic allergies, such as pollen sensitivities [mediated by immunoglobulin E (IgE)]; and various delayed reactions such as tuberculin-type reactions, or to poison ivy (a response mediated by T cells), and graft-rejection reactions (mediated by cell-bound antibodies). Toxic shock, the overstimulation of the immune system upon infection with some bacteria, is usually due to binding of their toxins to major histocompatibility complex molecules, widely and inappropriately activating T cells. *See* ANAPHYLAXIS.

In allergy, an immune response can occur to an innocuous foreign substance (allergen), such as pollen, bee stings, poison ivy, or some foods. To become allergic, an individual first needs to be exposed to the allergen and is thus sensitized. The response on reexposure is the allergic reaction. An allergy is a normal immune response, but to a substance that is not usually harmful. *See* ALLERGY.

Autoimmunity. Individuals sometimes respond immunologically to certain self antigens, so-called autoantigens, resulting in an autoimmune response. These exceptions to self tolerance are frequently associated with disease, the autoimmune diseases, which are widespread pathological conditions,

including Addison's disease (adrenal insufficiency of nontuberculous origin), celiac disease, Goodpasture's syndrome, Hashimoto's thyroiditis, juvenile-onset diabetes mellitus, multiple sclerosis, myasthenia gravis, pemphigus vulgaris, rheumatoid arthritis, Sjögren's disease, and systemic lupus erythematosus. In these diseases, antibodies or T cells activate against self components.

The mechanisms responsible for autoimmune diseases are not known and may be complex. Although these diseases share some features, it is unlikely that they have a common cause. It is normally assumed that in autoimmune disease tolerance to self breaks down, eventually leading to both effector cells and antibodies against self. *See* AUTOIMMUNITY.

Suppression of immunity. In cases of unfavorable antigen-antibody reactions in autoimmunity, allergy, and transplantation, it may be desirable to suppress immune reactions. Various agents have been employed, such as x-irradiation, folate analogs, corticosteroid hormones, and antilymphocyte antisera. *See* IMMUNOSUPPRESSION. John Trowsdale

Bibliography. C. A. Janeway et al., *Immunobiology: The Immune System in Health and Disease*, 4th ed., Elsevier/Garland, 1999; J. Kuby, *Immunology*, 3d ed., Freeman, 1997; I. Roitt, J. Brostoff, and D. Male, *Immunology*, 5th ed., Mosby, 1998.

Immunoassay

An assay that quantifies antigen or antibody by immunochemical means. The antigen can be a relatively simple substance such as a drug, or a complex one such as a protein or a virus. *See* ANTIBODY; ANTIGEN.

The reactants are first mixed so that a varying quantity of one (A) is added to a constant amount of the other (B). The formation of an immune (antigen-antibody) complex is measured as a function of the varied reactant (A). The result is represented by a "standard curve" for reactant A. An unknown sample is tested by adding it to reactant B. The extent of the measured change is referred to the standard curve, and thereby is obtained the amount of reactant A which produces a comparable change. That amount is represented as the content of reactant A in the unknown sample.

This type of immunoassay requires a direct detection of only the amount of immune complex or only the amount of residual, free reactant. The immune complex can be measured by the formation of large aggregates (immunoprecipitation), of small aggregates detected by light scattering (nephelometry), or by the evolution of heat in the reaction (microcalorimetry). If one of the reactants is initially fixed to a solid support such as the wall of a test tube or a fragment of plastic, the support can be washed free of contaminants at the end of the reaction. The product of the reaction can then be assayed, for example, by immunofluorescence. *See* IMMUNOFLUORESCENCE; IMMUNONEPHELOMETRY.

Determination of residual, free reactant requires that the substance have some special property which is lost in the immune reaction. An example of this is an enzyme, such as streptococcal hemolysin O, which is inhibited by a specific antibody. The degree of inhibition can be used to quantify either the enzyme or the antibody to the enzyme.

Other forms of immunoassay require that one of the reactants (A or B) be labeled. The sensitivity of the assay, that is, its ability to detect very small amounts of a reactant, is related to the ability to detect the label. Commonly employed labels include enzymatic, fluorescent, and radioactive; others may be detected by such methods as electron spin resonance and nuclear magnetic resonance. The reactant to be measured in a test sample (AU) is of the same type as the labeled reactant (AL), but, of course, is unlabeled. AU and AL may be mixed, or AU may be added first and AL later to a constant amount of B. The reaction product of AU + AL + B now contains an amount of bound AL which varies inversely with that of added AU. By contrast, the content of free or unreacted AL varies directly with the amount of added AU. The bound AL, the free material, or both can be assayed and the results used to construct a standard curve. *See* RADIOIMMUNOASSAY.

Assays which use one reactant that adheres to a solid support again facilitate the separation of the reaction product from the unreacted material. Otherwise, the bound reactant must be separated from the free material by a subsequent procedure. That procedure is commonly precipitation with either antibody, directed against the first (specific) antibody used in the assay or substances such as ammonium sulfate or polyethyleneglycol. The label is then assayed by an appropriate procedure, for example, immunofluorescence or enzyme-linked immunoassay.

Enzymatic labels can achieve great sensitivity, similar to that of radioimmunoassay. They can be used in two ways: (1) The enzyme is initially reacted with a low-molecular-weight substrate that is to be quantified. The antibody to this substance inhibits the action of the enzyme. The addition of a free substance reverses the inhibition of the enzyme by the antibody. The extent of reversal of the inhibition, or of the residual activity of the enzyme, is related to the quantity of the measured substance. (2) The enzyme is coupled to a macromolecular reactant. The activity of the enzyme then reflects the concentration of the labeled reactant. This, in turn, is related to the concentration of the unlabeled reactant in the test sample.

The enzymatic label can be measured by various assays. The simplest are colorimetric, which use a color change. Other changes can be measured by absorptiometry, chemiluminescence, fluorometry, nephelometry, potentiometry, or radioactive assay. *See* IMMUNOLOGY. Alexander Baumgarten

Bibliography. J. W. Kimball, *Introduction to Immunology*, 3d ed., 1990; W. E. Paul, *Fundamental Immunology*, 4th ed., 1998; D. M. Weir (ed.), *Handbook of Experimental Immunology*, 5th ed., 1996; D. Wild (ed.), *Immunoassay Handbook*, 2d ed., 2000.

Immunochemistry

A discipline concerned both with the structure of antibody (immunoglobulin) molecules and with their ability to bind an apparently limitless number of diverse chemical structures (antigens); with the structure, organization, and rearrangement of the genes coding for the immunoglobulin molecules; and with the structure and function of molecules on the surface of animal cells, such as the transplantation (histocompatability) antigens, which recognize antibodies and the thymus-derived lymphocytes mediating the cellular immune response. A development of immunochemistry that is of medical, scientific, and commercial importance uses the binding specificity of immunoglobulins to measure complex (but sometimes simple) chemical structures such as alkaloids, hormones, proteins, peptides, complex carbohydrates, and lipids. The tests used for such measurements employ the binding of radioactive antigens (the radioimmunoassay) or the ligation of an antigen bound to an enzyme (the enzyme-linked immunoassay). These assay methods are highly sensitive and specific.

The impact of immunochemistry has been very strong on medicine, giving pathophysiological insights into the mechanism of diseases such as multiple myeloma, chronic infections, and various disorders of immune complex formation. It seems likely that precise delineation of cell recognition and cell-cell interactions may yield information which will make easier the transplantation of organs between incompatible donors and recipients, and will also aid the elimination of aberrations of the immune system which at present cause diseases in humans. The applications of immunoassays for measuring contaminants in industry and for the detection of toxic agents in the environment is also likely to expand considerably. *See* CELLULAR IMMUNOLOGY; IMMUNOASSAY; IMMUNOLOGY; RADIOIMMUNOASSAY; TRANSPLANTATION BIOLOGY. W. H. Konigsberg; F. F. Richards

Bibliography. L. E. Glynn and M. D. Steward (eds.), *Immunochemistry: An Advanced Textbook*, 1977; J. W. Kimball, *Introduction to Immunology*, 3d ed., 1990; C. J. Van Oss and M. H. V. Van Regenmortel (eds.), *Immunochemistry*, 1994.

Immunoelectrophoresis

A combination of the techniques of electrophoresis and immunodiffusion used to separate the components of a mixture of antigens and make them visible by reaction with specific antibodies.

Most biological substances in solution carry either an excess of protons (+1, +2, etc.) or a deficiency (−1, −2, etc.) and therefore can be separated according to their charge in an electrical field by electrophoresis. Substances which carry a similar net charge tend to occupy a similar position after separation. These substances, however, are generally antigenic and can react with antibodies to form immunoprecipitates which, in clear media (such as agar or gelatin), appear as opaque, white lines because they scatter light, whereas the media themselves transmit light without hindrance. The latter technique is known as immunodiffusion. *See* ELECTROPHORESIS; PRECIPITIN TEST.

Method. A medium such as agar is deposited on a convenient base, for example, a microscopic slide. A small well (0.04–0.20 in. or 1–5 mm in diameter) is cut in the medium. A test solution is deposited in the well, and the contained substances are separated by electrophoresis along one axis of the plate. A trough is then cut in the medium parallel to, but at some distance (0.12–0.20 in. or 3–5 mm, for example) from, the line of the separated substances. The trough is filled with antiserum which contains antibodies to one or more of the separated substances. The antiserum and substances diffuse toward one another and, where they meet, form curvilinear patterns of precipitation. These patterns can be seen directly in clear media or can be visualized after washing out unreacted materials and staining in opaque media.

Variants. Many variants of the technique have been developed.

Crossed electrophoresis. This method uses an initial electrophoretic separation along one axis of the plate. After separation, a strip of medium along the axis is preserved, while the remainder is excised. New medium, which contains antiserum, is then poured beside the central strip. The plate is turned 90°, and electrophoresis is resumed. Lines of precipitation form as the initially separated substances (antigens) are driven by their electrical charge out of the central strip and into the antibody-containing medium.

"Rocket" electrophoresis. This is a variant of crossed electrophoresis in which the medium typically contains only one antibody. No initial electrophoresis is done, but the test substances are driven directly into the medium, which contains the antibody. "Λ"- or "rocket"-shaped trails of precipitation appear within suitable ratios of concentration of the reactants. The height of the rocket is proportional to the concentration of the test substances and can be used to quantify as little as 0.1 μg/ml of protein.

Counterimmunoelectrophoresis. This technique uses two wells or points of application, one above the other, along the electrical axis. The anodal well is filled with antibody, and the cathodal with a negatively charged antigen. Electrophoresis results in the antibody migrating cathodally, while the antigen moves anodally. A line of precipitation appears where the two meet. This technique has allowed rapid detection of substances, for example, those of infectious origin.

Further variants. Other variants of the above techniques can be subdivided into those which: (1) substitute another method of separation for the initial electrophoresis, such as isoelectric focusing, gel filtration, or chromatography; (2) enhance resolution by the addition before separation of a marked tracer antigen or antibody, such as a radioactively labeled one, followed by autoradiography; or (3) use a vertical overlay of the plate with, or with incubation in, an antibody with or without a label (radioactive,

enzymatic, and so on) to enhance visibility after separation. In immunofixation the initial separation (which may be electrophoretic) is followed by an overlay with antibody, instead of its application to a trough at the side. The separated components are precipitated in place, thus helping to identify variously charged forms of the same antigen prior to their undue diffusion.

Immunoelectrophoresis or its variants can also be combined with methods which exploit an inherent property of the antigen to enhance visibility and resolve constituents that possess the selected property. For instance, the precipitates can be stained with dyes that react only with lipid or carbohydrate constituents, or the antigens in the precipitates can be identified by their enzymatic properties. *See* IMMUNOASSAY. Alexander Baumgarten

Bibliography. O. G. Bier et al., *Fundamentals of Immunology*, 1986; A. J. Crowle, *Immunodiffusion*, 1973; J. J. Scheidegger, *Int. Arch. Allergy Appl. Immunol.*, 7:103, 1955.

Immunofluorescence

A technique that uses a fluorochrome to indicate the occurrence of a specific antigen-antibody reaction. The fluorochrome labels either an antigen or an antibody. The labeled reactant is then used to detect the presence of the unlabeled reactant, for example, a component of a section of body tissue seen under a microscope. The use of a labeled reactant (such as an antibody which both detects and indicates the antigen) to reveal the presence of an unlabeled one is termed direct immunofluorescence. The use of a labeled indicator antibody, which reacts with an unlabeled detector antibody that has previously reacted with an antigen, is termed indirect immunofluorescence. Both techniques can be used qualitatively to indicate the presence of a specific reactant in a preparation. They can also be used semiquantitatively, for example, to determine the highest dilution (lowest concentration) of a tested detector reactant that can detect a given substance. Substitution of a light meter for the human eye permits a quantitative measurement in immunofluorometry. *See* FLUORESCENCE.

Background fluorescence. Many tissues and substances of biological origin are capable of autofluorescence. The emitted light is often blue or blue-green, but may be another color. Autofluorescence contributes to the background fluorescence of the examined material. Another contributor to the background is the residual deposit of the indicator material which cannot be removed by washing but is retained by nonspecific adherence to the examined material. The specific fluorescence must increase above the background in order to be perceived. This can be obtained in several ways. For instance, the indicator may be a substance which emits light of a different wavelength than the natural autofluorescence of the tested material, or the indicator may concentrate over accumulations of the tested reactant so that its brightness there is much greater than the general level of background fluorescence.

Efficiency of detection. Direct immunofluorescence generally produces low background, but also a low level of specific illumination. Indirect immunofluorescence tends to increase both the background and specific illumination, but because several molecules of the detector can react with each molecule of the tested reactant and, in turn, several molecules of the indicator can react with each molecule of the detector, a marked enhancement occurs in the efficiency of detection. Therefore, indirect immunofluorescence can detect less reactant than direct. Furthermore, the indirect technique is more convenient because one indicator antibody (such as goat antirabbit immunoglobulin) can be used with various detector antibodies (such as those of rabbit origin). The successive application of antibodies in the indirect technique produces a "sandwich" of which only the last layer fluoresces. Theoretically, the layers in the sandwich could be increased indefinitely, with a concomitant augmentation in sensitivity. In practice, each layer deposits some nonspecifically bound material, to which the next layer reacts specifically, as well as adding its own nonspecific load. The background therefore becomes excessively great within three or four layers.

Fluorochromes. The most common fluorochromes used for immunofluorescence are fluorescein isothiocyanate (with maximal absorption at 495 nanometers), which emits yellow-green light, and tetramethylrhodamine isothiocyanate (with maximum absorption at 544 nm), which emits red light. Both react spontaneously with the amino groups of proteins in alkaline solution. Other fluorochromes include DANS (1-dimethylaminonaphthalene-5-sulfonic acid) and RB200 (lissamine rhodamine B) sulfonyl chlorides. Two fluorochromes, each marking a different antibody, can be used for simultaneous identification of two antigens. Separate localization of each antigen is visualized by one color of fluorescence (such as green or red), but congruence of the antigens produces a mixed color (such as orange).

Light source. Observation of immunofluorescence by transmitted light requires an exciting beam of light below the specimen and a barrier filter above to absorb the exciting but not the emitted light. Alternatively, the exciting light is refracted down the objective of a microscope and passed through the specimen. The emitted light, however, travels back up the objective to the eyepiece (incident light fluorescence). Excitation is usually provided by ultraviolet light sources, such as a mercury lamp, a xenon arc, or halogen-filled lamps. It should be noted that sources of ultraviolet light can damage unprotected eyes.

Adaptations. Novel adaptations of immunofluorescence include immunoenzymatic assays and quenching. The former use a reactant coupled to an enzyme (β-galactosidase, for example) which cleaves a nonfluorescent compound (such as a derivative of umbelliferone) to yield a fluorescent derivative

(umbelliferone, for example). Quenching uses, say, two fluorochromes, one of which absorbs light emitted by the other. One fluorochrome labels the antigen, another the antibody. Antigen-antibody complexes retain both fluorochromes, so that the initially emitted light is absorbed and so quenched by the second compound. The extent of quenching is assayed. *See* IMMUNOASSAY. Alexander Baumgarten

Bibliography. C. C. Curtain and A. Baumgarten, *Immunology*, 10:499, 1966; J. W. Kimball, *Introduction to Immunology*, 3d ed., 1990; G. Wick, K. N. Traill, and K. Schaunstein (eds.), *Immunofluorescence Technology*, 1983.

Immunogenetics

A scientific discipline that uses immunological methods to study the inheritance of traits. Traditionally, immunogenetics has been concerned with moieties that elicit immune response, that is, with antigens (antigenic determinants). It has now broadened its scope to study also the genetic control of the individual's ability to respond to an antigen.

Methods. The immunological methods used in immunogenetics are of two principal kinds, serological and histogenetical. In serological methods, antibodies are used to detect antigens, either in solution or on a cell surface. In histogenetical methods, immune cells (lymphocytes) are used to detect antigens on the surface of other cells. *See* ANTIBODY; ANTIGEN.

Serological methods. To produce antibodies or to stimulate immune cells, an antigen is introduced into the blood of an animal—usually a mammal, but other vertebrates that have developed an immune system can also be used. Blood delivers the antigen to the lymphoid organs, where it is picked up by specialized cells which then present it to B and T lymphocytes. The B lymphocytes react to the stimulation by secreting antibodies (immunoglobulins) into the blood plasma; T lymphocytes react by differentiating into effector cells. The antibodies are then collected by bleeding the animal and separating the serum from the clotted blood. Serum containing antibodies specific for a given antigen is called immune serum or antiserum. The activated B lymphocytes can also be explanted from the animal and immortalized by fusion with an appropriate tumor cell. The resulting hybridomas then continuously produce antibodies of one kind (such antibodies are called monoclonal because they derive from a single lymphocyte clone). Monoclonal antibodies can be collected from the supernatant of a tissue culture in which the hybridoma is growing, or collected from the ascites fluid filling the peritoneal cavity of an animal into which the hybridoma has been inoculated. *See* CELLULAR IMMUNOLOGY; MONOCLONAL ANTIBODIES; SERUM.

The antibody binds only to the antigen that elicited its secretion, or at most to a few related compounds; it does not bind to any other antigens. There are five principal ways of measuring this antigen-antibody binding: (1) In the precipitation test, the soluble antigen and soluble antibody form insoluble complexes that fall out as visible or visualizable precipitate from a solution. (2) In the complement-fixation test, the antigen-antibody complexes bind (fix) complement, and the depletion of complement is then determined with the help of a test system. (3) In the hemolytic and cytotoxic tests, antibody binds to antigen on a cell surface and the complex activates complement components, some of which then kill erythrocytes or other cells. (4) In the agglutination test, the antibodies bind with one site to antigen on one cell and with the other site to antigen on another cell and thus glue the cells together. (5) In the fluorescence test and the radioimmunoassay, a label is attached to the antibody (fluorescein dye or radioactive atoms), which then signals to the investigator that the antibody has remained attached to the cell surface. *See* AGGLUTINATION REACTION; COMPLEMENT-FIXATION TEST; IMMUNOFLUORESCENCE; PRECIPITIN TEST; RADIOIMMUNOASSAY.

Histogenetical methods. The reaction of the immune lymphocytes with the target antigens can be revealed by three principal assays. In one, graft rejection methods, the lymphocytes trigger a reaction that destroys a foreign tissue or organ transplanted to a recipient. In the second assay, the immune lymphocytes proliferate in response to an antigen: this is called the mixed lymphocyte reaction if the antigen is present in the membrane of a foreign lymphocyte; it is called the T-lymphocyte proliferation assay when a soluble antigen is picked up and presented by a specific cell, which is usually genetically identical with the lymphocyte. In the third assay, the cell-mediated lympholysis assay, the activated lymphocytes differentiate into cytolytic cells which kill appropriate target cells. *See* TRANSPLANTATION BIOLOGY.

Genetic situations. There are four genetic situations in which immune cells and antibodies can be elicited. In the xenogeneic situation, the antigen derives from a different animal species than the responding animal. In the allogeneic situation, the donor is one individual and the responder another, genetically disparate individual of the same species. In the syngeneic situation, the donor and the recipient are genetically the same (for example two individuals of an inbred strain). In the autologous situation, an individual responds to its own antigens. Responses in syngeneic and autologous situations occur only under pathological conditions when the individual fails to distinguish between self and nonself entities and regards its own molecules as foreign antigens (autoimmunity). *See* AUTOIMMUNITY.

Molecular methods. In modern immunogenetics research, the serological and histogenetical methods are combined with molecular methods in which the researcher isolates and works with the genes that code for the traits. Usually, the antigens are identified with antibodies or lymphocytes, then messenger ribonucleic acid (mRNA) coding for the antigen-bearing protein is isolated and used to identify the deoxyribonucleic acid (DNA) segment (gene) from which it has been transcribed. The isolated gene, in turn, can be inserted into a suitable cell and activated, and the expressed molecules used to elicit

new sets of antibodies or immune cells for a refined analysis of the gene product. This approach of going back and forth from classical to molecular methods has proved to be very successful and has led to the elucidation of several complex genetic systems.

Tools. Animal immunogenetics relies heavily on the use of inbred, congenic, and recombinant inbred strains.

Inbred strains. Inbred lines result when individuals that are more closely related to each other than randomly chosen individuals mate together, for many generations. In a randomly mating population, each individual is homozygous for most of its genes (carries identical alleles at each of the loci) but heterozygous at some 20 to 40% of loci. Random mating propagates this status, whereas inbreeding tends to increase the proportions of homozygous loci.

The reason for this increase is easy to understand. First, assume that the two parents of an initial mating are both *Aa* heterozygotes (where *A* and *a* are alleles at a given locus). The *Aa* × *Aa* mating produces *AA*, *Aa*, and *aa* progeny. If brothers are mated with sisters from this litter, sooner or later two *AA* or two *aa* individuals will be selected by chance and mated together. When this happens, all the progeny in all subsequent brother × sister matings will always be *AA* or *aa*. Homozygosity has been achieved at the *A* locus, and as long as the individuals in this genealogical line are not outcrossed to unrelated individuals, the line remains homozygous for this locus. The same thing happens sooner or later for other loci, *B*, *C*, *D*, and so on—for all those loci that were originally in a heterozygous state. The mathematical theory of inbreeding predicts that after 20 generations of brother × sister matings, more than 97% of the loci become homozygous. By convention, this figure is considered sufficient for a strain to qualify as inbred.

When the matings are not between brothers and sisters but between some more distant relatives, it takes proportionally longer to achieve this degree of homozygosity. Because each individual carries a certain load of deleterious genes (that is, genes that in a homozygous state lower the viability of their bearer), and because inbreeding also increases the homozygosity of these genes, an inbreeding depression may set in after a few inbreeding generations. In this phase of the inbreeding program, many individuals either die or do not breed. However, if the program has been initiated on a sufficiently large scale, there are usually enough genetically healthy individuals to carry the program over this hurdle. After 20 generations of brother × sister mating, individuals within a group become genetically homogeneous, resembling one another the way identical twins do; an inbred strain (line) is thus established. A number of inbred strains have been developed in the laboratory mouse, rat, domestic fowl (chicken), guinea pig, Syrian hamster, and rabbit.

Congenic lines. The advantage in working with inbred strains rather than outbred animals is that inbred strains restrict the variability of the conditions of an experiment. However, when two strains are compared and it is found that they respond differently to a treatment, it is not known to what gene this difference should be attributed. The strains may differ at as many genetic loci as two unrelated individuals in an outbred population do. To study the effect of single, defined genes, immunogeneticists have developed congenic lines. These lines always come in groups, the smallest group being a pair, which consists of a congenic line and its inbred partner strain. The two are homozygous at more than 97% of their loci (that is, they are inbred) and are identical except, ideally, at one locus—the locus that is to be studied. At this differential locus, the two strains carry different alleles. This ideal situation is achieved only rarely, if ever; most of the time immunogeneticists are satisfied when the two strains in a pair differ in a short chromosomal segment that includes the differential locus. In addition to pairs, trios, quartets, or quintets of congenic lines can also be produced, in which all the lines differ from the inbred partner strain and from one another in that they carry different alleles at the same differential locus.

The simplest way of producing a pair of congenic lines involves two strains, 1 and 2, that differ at locus *D* (see **illus.**). Strain 1 is a *dd* homozygote and strain 2 is a *DD* homozygote. To study the effect of the *D* gene, a line is produced that is genetically identical with strain 1 but carries the *D* allele. To this end, the two strains are crossed, and the resulting F_1 hybrid is

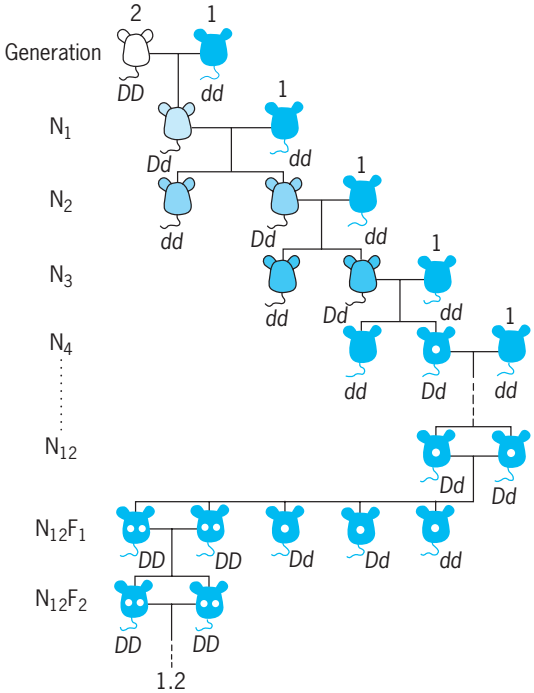

A scheme of congenic-line production by repeated backcrossing to strain 1 and selection in each segregating generation for the *D* allele donated by strain 2. The successive generations of matings of *Dd* individuals with strain H-1 are designated N_1, N_2, N_3, and so on. After 12 such generations, two *Dd* individuals are intercrossed and *dd* individuals are selected and mated among themselves. The intercrossing generations are then designated F_1, F_2, F_3, and so on. (*After J. Klein, Immunology, John Wiley and Sons, 1982*)

backcrossed to strain 1. In the backcross generation, the *D* gene segregates so that half the progeny is *Dd* and the other half *dd*. The animals are then typed and one of the *Dd* individuals is selected for further mating (backcrossing) to strain 1; this procedure is repeated for several generations. At the end, two *Dd* individuals are intercrossed and *DD* animals are obtained. By brother × sister mating of *DD* animals, a 1.2 line is obtained which is very similar to strain 1, but which carries the *D* instead of the *d* allele at the *D* locus. According to calculations, 99% of the loci not linked to *D* (that is, on a different chromosome than *D*) are of strain 1 origin when the backcross is repeated 12 times.

The largest number of congenic lines is available in the laboratory mouse. Lines which are known to differ from the inbred partner strain at a single locus (rather than a chromosomal segment) are called coisogenic. They arise when one of the loci in an inbred strain mutates and thus changes to another allele. Individuals homozygous for the mutation are then coisogenic with the original strain.

Recombinant inbred strains. To find out whether two loci are on the same or on different chromosomes, two individuals that differ in the traits controlled by these loci are mated and then the F_1 hybrids are intercrossed. In the F_2 generation that results from this intercross, the genes assort either independently, if they are on different chromosomes, or nonrandomly, if they are on the same chromosome—that is, when they are linked. Each time the strains are tested for linkage, this laborious procedure must be repeated. To avoid this repetition, immunogeneticists have prepared a "frozen" F_2 generation by establishing separate inbred lines from the different F_2 individuals. Such lines are called the recombinant inbred strains. Similarity of recombinant inbred strain-distribution patterns of two traits suggests that they are controlled by loci on the same chromosome. A number of recombinant inbred strains are available in the laboratory mouse—one of the most favored experimental animals.

Scope. Contemporary immunogenetic research concentrates on two main categories of antigenic substances—those present in body fluids, primarily blood serum or plasma, and those expressed on surfaces of various cells.

Body-fluid antigens. In this category, a prominent position is occupied by immunoglobulins (Ig), the proteins that constitute antibodies. Although antibodies are usually used to detect antigens, they themselves may also serve as antigens, and antibodies can be produced against them. These antibodies against antibodies detect three principal kinds of antigenic determinants: isotypic, allotypic, and idiotypic.

1. *Isotypic determinants*. These are shared by all individuals of a given species but absent in individuals of other species. Antibodies against isotypic determinants can normally be produced by xenogeneic immunization, that is, by injecting immunoglobulins of one animal species into individuals of another species. Isotypic determinants differentiate classes and subclasses of immunoglobulins. There are five classes of human immunoglobulins: IgG, IgM, IgA, IgD, and IgE. Some of the classes (IgG, IgA) consist of several subclasses (IgG 1–4; IgA 1, 2) which are more related to one another than two different classes are. Both classes and subclasses differ from one another not only in their serological but also in their physiochemical properties. *See* IMMUNOGLOBULIN.

2. *Allotypic determinants (allotypes)*. These are determinants present in some individuals of a given species and absent in other individuals. Antibodies against allotypic determinants arise normally during allogeneic immunizations; abnormally, they are present in the sera of patients with certain diseases, in particular, rheumatoid arthritis. In the latter case, the antibodies react with self components: they are autoantibodies. Most identified allotypic determinants are present in the constant region of the Ig molecule, but a few have also been located in the variable region. In many instances, allotypic differences have been correlated with amino acid substitutions in the primary structure of the Ig polypeptide chain. One or two substitutions are usually enough to create an allotypic difference between two Ig chains. Allotypic determinants have been studied extensively in human, rabbit, and mouse. The best known of the human allotypes is the Gm system, expressed on the different subclasses of the IgG class. A number of Gm determinants have been identified and their distribution in human populations mapped.

3. Idiotypic determinants (idiotypes). These are unique antigenic determinants restricted to antibodies of a single species or of only a few related species. They can be identified with antibodies produced during xenogeneic, allogeneic, or syngeneic (autologous) immunization. They reside in the variable region of the Ig molecule on either the heavy or the light chain. Differences in idiotypic determinants arise because individuals of the same species differ in the set of genes controlling the variable-region polypeptide chains, and because in ontogeny of the B lymphocyte these genes undergo rearrangements that change the structure of the encoded proteins. Because of their location in the V region, idiotypic determinants are associated with antibody specificity in that Ig molecules reacting with different antigens usually express different idiotypes. *See* IMMUNOLOGICAL ONTOGENY.

Cell-surface antigens. The main categories of cell-surface molecules studied by immunogenetical methods are blood group antigens, histocompatibility antigens, tissue-restricted antigens, and receptors.

1. *Blood-group antigens*. These are alloantigens found on erythrocytes. Historically, they were the first antigens studied by immunogenetical methods, and a great number of them have been identified in both humans and animals. The blood-group variations usually reflect differences in the structure of cell-surface carbohydrates. *See* BLOOD GROUPS.

2. *Histocompatibility (H) antigens*. These are antigens capable of inducing cellular immune responses and hence are detectable by histogenetical methods. However, at least some of them also stimulate antibody responses. They are divided into two

groups, minor and major, but the two groups have nothing else in common than the fact that they are detected by the same method.

The minor *H* loci are quite numerous (there may be several hundred of them), and they are scattered all over the genome. They are referred to as minor because they act as weak antigens in inducing cellular immune response. However, they act synergistically and their combined effect can be quite strong. Their function is not known.

In contrast, most of the major *H* loci reside in a single chromosomal region and are potent inducers of cellular immune response. The function of the major histocompatibility complex (MHC) is to introduce antigens to the T lymphocyte. The T lymphocyte does not recognize antigens if they are not accompanied by MHC molecules on the antigen-presenting cell. The MHC of different species has different names. In humans it is called the human leukocyte antigen (HLA) complex.

The MHC genes can be divided into two classes, which differ in their organization and function. Class I and II genes comprise polypeptide chains that are folded in such a way that a groove is created in the external portion of the molecule. During the synthesis of the MHC molecule, the groove binds short peptides derived from other proteins of the cell. The molecule then displays the peptides on the cell surface. Peptides derived from the body's own proteins are ignored by the immune system; those derived from invading parasites are recognized (together with portions of the MHC molecule itself) by T lymphocytes, the recognition activates the lymphocytes, and the activated lymphocytes initiate the specific form of the immune response. The class I MHC molecules bind peptides derived from proteins manufactured in the cell (either the cell's own proteins or those of intracellular parasites). The bound foreign peptides are recognized by cytotoxic T lymphocytes, whose function in the immune response is to kill other cells. The class II MHC molecules, however, bind peptides derived from proteins that have been taken up by the cell from the extracellular milieu. The bound foreign peptides are recognized by helper T lymphocytes, which regulate the immune response and facilitate the production of antibodies by B lymphocytes.

The functional class I and class II genes are highly polymorphic, which means that they occur in a number of allelic forms carried by different individuals. This is apparently unique and characteristic of the MHC. Its significance is unclear, but the polymorphism may represent a way of coping with MHC-associated nonresponsiveness to certain antigens. Certain combinations of antigen and MHC molecules are not recognized by T lymphocytes, so that an individual carrying the particular MHC molecules (alleles at a particular MHC locus) is unresponsive to the particular antigen. Because the nonresponsiveness depends on the presence of a particular MHC allele, when responders mate with nonresponders, the ability to respond segregates in the progeny with the MHC. Outwardly it appears as if there were special immune response (*Ir*) genes linked to the MHC, while, in fact, the MHC themselves are the *Ir* genes. Because certain MHC molecules confer nonresponsiveness to certain antigens, they may make certain individuals vulnerable to certain pathogens. The MHC polymorphism may be the consequence of selection for MHC alleles that allowed individuals to survive past epidemics. *See* POLYMORPHISM (GENETICS).

In the human population, certain HLA alleles or combinations of alleles (haplotypes) are associated with disease. Individuals suffering from a specific disease often possess a particular HLA allele more frequently than healthy individuals. For example, nearly all individuals suffering from narcolepsy (a disease characterized by a sudden, uncontrollable compulsion to sleep during waking hours) have the *HLA-DRB1*02* allele. *See* HISTOCOMPATIBILITY.

3. *Tissue-restricted antigens*. These are expressed on some tissues but not on others and therefore serve as markers for cell sorting. A number of such antigens have been identified in humans and the mouse, particularly on blood cells (many blood-group antigens fall into this category). There are antigens that are expressed on T but not B lymphocytes and others for which the opposite is true. Some antigens are expressed only on a subpopulation of T or B lymphocytes.

4. *Receptors*. These are molecules that are capable of specifically interacting with certain other molecules. The interaction often leads to activation or inhibition of the receptor-bearing cell. The receptors themselves can serve as antigens and can elicit antibody production or activation of immune lymphocytes. The most extensively studied receptors are those that recognize antigens and are located on the surface of B and T lymphocytes. The B-lymphocyte receptor closely resembles the antibody molecules that are secreted by the cell upon activation and differentiation. The T-lymphocyte receptors are related to but different from the immunoglobulin molecules. *See* GENETICS; IMMUNOLOGY. Jan Klein

Bibliography. J. Klein, *Immunology*, 2d ed., 1997; J. Klein, *Natural History of the Major Histocompatibility Complex*, 1986; R. Srivastava, G. P. Ram, and P. Tyle (eds.), *Immunogenetics of the Major Histocompatability Complex*, 1991.

Immunoglobulin

Any of the glycoproteins in the blood serum that are induced in response to invasion by foreign antigens and that protect the host by eradicating pathogens. Antibodies belong to this group of proteins. An antigen is any substance capable of inducing an immune response. Intact antigens are able to specifically interact with the induced immunoglobulins. Normally, the immune system operates in a state known as self-tolerance, and does not attack the host's own tissues, but occasionally the immune system targets host-specific antigens, resulting in autoimmune disease. *See* AUTOIMMUNITY.

Immunoglobulins are composed of two identical heavy (H) and two identical light (L) polypeptide chains. Each H and L chain has an amino-terminal variable (V) region and a carboxyl-terminal constant (C) region. Although V regions from different antibodies exhibit considerable sequence variation, there is a large degree of sequence similarity among C regions of different antibodies.

In the living animal, antibodies first bind to antigen and then, ideally, eliminate it as a threat to the host. The antigen is bound to the antibody at the antigen-combining site, a cleft or groove formed by the association of the amino-terminal variable regions of both chains, which are complementary to the size, shape, and charge characteristics of the antigen. Bound antigen is then inactivated by one or more of the effector functions of the antibody.

Human myeloma proteins, the first structurally discrete immunoglobulins known, result from the malignant expansion of a normal clone of antibody-producing cells. The antigenic specificity of myeloma proteins, however, is generally unknown.

Structure. An immunoglobulin molecule has a Y-shaped appearance with two arm segments which are the Fab (fragment antigen binding) region connected by a flexible hinge to a stem composed of the Fc (fragment crystallizable) region. The basic structural unit of immunoglobulins is the immunoglobulin domain. The L chain is composed of two domains, and the H chains are divided into four domains for IgG, IgA, and IgD and a fifth domain for both IgM and IgE. Each domain contains approximately 110 amino acids and is encoded by a separate deoxyribonucleic acid (DNA) exon. *See* EXON.

Classes of immunoglobulins. Immunoglobulins are heterogeneous with respect to charge, size, antigenicity, and function. There are three categories of antigenic determinants present on immunoglobulins: isotypes are found in all individuals, allotypes are found in some individuals, and idiotypes are associated with the amino-terminal variable region. Isotypic determinants are located on the carboxyl-terminal constant region and are used to group immunoglobulin H and L chains into isotypes or classes. In total, there are five human H-chain classes. IgM contains mu (μ) H chains, IgG contains gamma (γ) H chains, IgA contains alpha (α) H chains, IgD contains delta (δ) H chains, and IgE contains epsilon (ϵ) H chains. IgG has four subclasses, IgG1, IgG2, IgG3, and IgG4, while IgA has two subclasses, IgA1 and IgA2. There are two L-chain isotypes named kappa (κ) and lambda (λ). Kappa and lambda chains may be associated with H chains of any isotype, and a complete description of an immunoglobulin molecule requires identification of both H and L chains.

IgG. IgG is the most abundant immunoglobulin class in the serum. The four IgG subclasses, based on chemical distinctions between the H chains, are named numerically in order of their relative concentration in normal serum. All four subclasses contain approximately 3% carbohydrate. However, the main structural difference among them is in the hinge region, where the number of amino acids and the number of interchain disulfide bonds vary. IgG isotypes are associated with complement fixation, opsonization (that is, rendering more susceptible to phagocytosis), fixation to macrophages, and membrane transport. All four IgG subclasses bind streptococcal protein G. IgG1, IgG2, and IgG4 isotypes bind protein A from *Staphylococcus aureus*. *See* STAPHYLOCOCCUS; STREPTOCOCCUS.

IgA. Of the two IgA subclasses, IgA1 is the predominant subclass of IgA in human serum, representing 7–15% of serum immunoglobulin. IgA1 is the dominant subclass in all external secretions, including milk, saliva, tears, and bronchial fluids. The percentage of subclass IgA2 is higher in these fluids than in serum. Secretory IgA consists of two IgA monomers, one J chain, and one glycoprotein called secretory component. In IgA, J chain may be necessary for polymerization of the two monomeric units; however, it is not required for the polymerization or secretion of IgM antibodies. The secretory component is cotransported with the antibody. Its function is unknown, but possibly it stabilizes the IgA molecule and provides protection from the enzymatic activity of the gastrointestinal tract.

IgM. IgM is the first immunoglobulin to appear during the primary immune response. IgM is also a minor component of secretory immunoglobulins at mucosal surfaces and in breast milk. IgM antibodies are secreted from B lymphocyte cells as pentameters or hexameters.

IgD and IgE. IgD and IgE are present in minute amounts in normal human serum. No function has been clearly attributed to IgD. IgE is active against parasites and acts as a mediator of immediate hypersensitivity. Circulating IgE binds the high-affinity receptor on the surface of basophils in the circulation and mast cells in tissues. Subsequent to the interaction of bound IgE with antigen, basophils disintegrate, releasing inflammatory agents such as histamine. This type of response leads to allergy and anaphylaxis (hypersensitivity following injection of an antigen) in susceptible individuals. Atopic dermatitis, an inflammatory skin disorder of unknown etiology, is characterized by excessive production of IgE. *See* ANAPHYLAXIS; ANTIBODY; ANTIGEN; ANTIGEN-ANTIBODY REACTION; IMMUNOLOGY; PROTEIN.

Kathleen N. Potter; J. Donald Capra

Bibliography. W. E. Paul (ed.), *Fundamental Immunology*, 4th ed., 1998; F. Shakib, *Molecular Aspects of Immunoglobulin Subclasses*, 1990; G. P. Smith, *The Variation and Adaptive Expression of Antibodies*, 1990.

Immunoglobulin genes

The genes encoding the genetic information to make antibodies (immunoglobulins) of different specificities and classes. Specificity is encoded in the variable region of the antibody, which is generated by assembling gene segments through a process called rearrangement. Biological activity of an antibody is established through a process called class switching

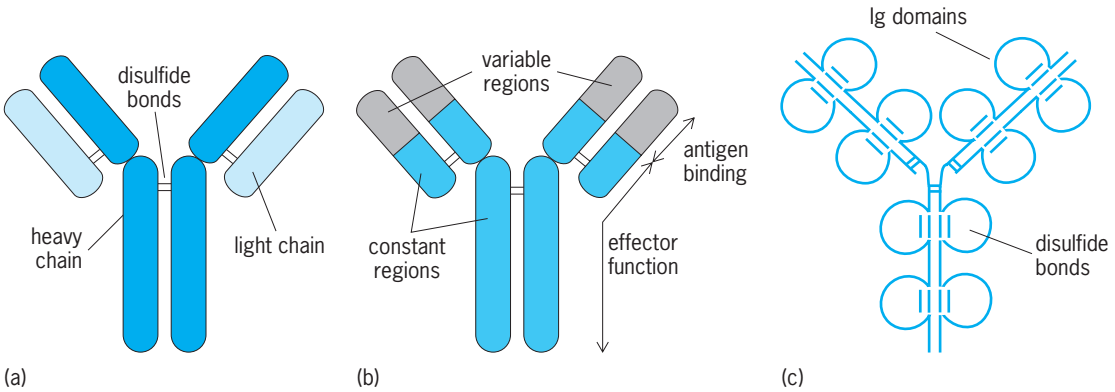

Fig. 1. Structure of a generic antibody (IgG). (*a*) A heavy chain is connected to a light chain by a disulfide bond. This heterodimer is linked to another heavy/light-chain heterodimer by another disulfide bond. (*b*) Amino acid sequencing showed two regions within one chain, a constant region conveying the biological activity (effector function) of the antibody and a variable region that specifies the antigen binding. (*c*) More detailed amino acid sequencing identified Ig domains, shown here as loops that are created by disulfide bonds. Variable regions consist of one domain; constant regions of the heavy chain have three or four Ig domains, whereas light chain constant regions have only one.

in the constant part of the antibody. *See* ANTIBODY; GENE; IMMUNOGENETICS; IMMUNOGLOBULIN.

Structure. Once researchers observed that mice injected with mineral oil would develop B cell (bone marrow–derived lymphocyte) tumors that secreted immunoglobulins (antibodies, Ab) that appeared to be all the same (monoclonal), the groundwork was laid for solving the structure of antibodies. These uniform immunoglobulins (which are glycoproteins) could be purified and studied to determine the order of amino acids in the protein or, by treating the protein with chemicals or enzymes, to deduce the structure of the protein of interest.

Rodney Porter and Gerald Edelman chose the latter approach and solved the structure of immunoglobulins by treating the immunoglobulin (Ig) with a reducing agent (β-mercaptoethanol). They found that immunoglobulins consist of two identical small protein chains and two identical large protein chains. These are generally referred to as the light and heavy chains of an antibody, respectively. To form the final structure, one light chain pairs with one heavy chain, they are linked together by disulfide bonds (S-S bonds between two amino acids), and these heterodimers (larger protein structures made of two different protein chains) are then paired with another heterodimer and held together also by disulfide bonds making a final 4-chain structure (**Fig. 1***a*).

Amino acid sequencing also revealed extreme heterogeneity in the first (amino-terminal) 110–125 amino acids of the heavy and light chains. Investigators concluded that this was the part responsible for the unique antigenic specificity of an Ab. These areas were called variable (V) regions (Fig. 1*b*). The two variable regions of the heavy and light chains together create the binding site of the antibody. This binding site recognizes the antigen in a very specific way, similar to the way one specific key fits one particular lock. *See* ANTIGEN.

The more conserved part of the sequence was named the constant (C) region (Fig. 1*b*). The C region of the heavy chain in particular mediates the biological activity or effector function of the antibody.

The amino acid sequence also revealed that the variable and the constant regions shared a common domain structure. These Ig domains (Fig. 1*c*) were roughly 110 amino acids long and showed some level of sequence and structure conservation, such as a disulfide bond creating a loop of ~60 amino acids, and the secondary structure of the protein forming a barrel structure, in which the walls are formed by beta sheets (in which hydrogen bonds are formed between two parts of the protein chain that can be far apart). These Ig domains were later found in many more proteins, creating an "Ig superfamily" that encompasses a multitude of immunologically important proteins [for example, T-cell (thymus-derived lymphocyte) receptors, major histocompatibility complex (MHC) molecules, and even adhesion molecules].

More detailed analyses of these monoclonal antibodies detected further differences in the variable regions of antibodies. There were stretches of amino acids that were more conserved between the different monoclonal immunoglobulins, and others that showed a particular high variability. The three conserved regions basically build the backbone of the variable region (the barrel) and were named framework regions (FR), while the three sections of amino acids that showed high variability formed amino acid loops with higher flexibility. They were called complementarity-determining regions (CDRs) because they participate in the actual binding of the antigen.

In summary, an antibody is built of two heterodimers, each consisting of one heavy and one light chain. Each of the two chains has a variable region, and together the two chains constitute the binding site for the antigen, which is formed by the complementarity-determining regions of the two chains. Thus, the antibody has two binding sites for antigen.

Variability. For a long time, the dogma "one gene–one protein" ruled molecular biology. This would mean two genes are required to encode one antibody, since an antibody consists of two different protein chains (heavy and light). The question then arises as to how one human being can make millions and millions of antibodies with different specificities that react against basically everything in this world when the human genome encodes only 30,000–40,000 genes. The solution came in 1979 when Susumu Tonegawa discovered that immunoglobulin genes are assembled from gene segments. Heavy chains are assembled from three different segment types, variable (VH), diversity (D), and joining (JH) segments (where H stands for heavy), to be rearranged into one functional VHDJH variable region. The rearrangement of the variable light-chain region is more simple; only two gene segment types are rearranged, VL and JL (where L stands for light), omitting the D region (**Fig. 2**).

In the human germline for the heavy chain (before rearrangement happens), there are about 50 VH genes, 27 D genes, and 5 JH genes, all lined up on the DNA and interrupted by stretches of noncoding DNA. All these gene segments then can be combined freely, as long as they join to make a VDJ region in the end. To get to this point, first a D region is attached to a JH region, making a DJH fragment. Next, a VH is rearranged to attach to a DJH fragment. The V, D, and J segments are flanked by small recognition sequences (called RSS for recombination signal sequences) that are mandatory for the rearrangement. These guarantee that the VH rearranges only to attach to the D, and the D to the JH, but not one JH to another JH, or a VH to a JH. On the molecular level, two recombination signal sequences are brought together and the intervening noncoding DNA is looped out. Two enzymes called RAG1 and RAG2 (recombination-activating genes) then cut the DNA at the outer ends of the coding sequences (VH, D, or JH). This cut is not precise and can add to the variability of an antibody. Before ligation (sewing them back together) of the two DNA ends, another enzyme (TdT: terminal deoxynucleotidyl transferase) can add random nucleotides. This too is a major contributor to add variability to the sequence and eventually the antibody. Together these nucleotides encode amino acids that had not been encoded by the germline DNA but increase the variability of antibodies extremely. While TdT and RAG1/2 are the only enzymes that are specific for B and T cells, there are many more enzymes involved in the process of rearrangement that are expressed ubiquitously. *See* DEOXYRIBONUCLEIC ACID (DNA).

The light chains are more simple in structure; they consist only of V and J segments, which are brought together in one simple rearrangement. In contrast to the heavy chain, there are two light chains to choose from, called the kappa and the lambda chain. In humans, about 50% of antibodies have a kappa light chain. The one heavy chain expressed can pair with only one type of light chain per cell, kappa or lambda. The order of the gene segments differs slightly between the kappa and lambda chain. While in both cases all V segments are together, the V segments in the kappa chain are followed by a cluster of five J segments and one constant region. The lambda locus has three available J segments with a compulsory lambda constant region in each.

The last way to increase variability is the ability of every heavy chain to combine with every single light chain when forming an antibody.

In summary, this makes a total of four ways to increase variability, which can be differentiated as combinatorial and junctional diversity. The free choice of picking V, D, and J segments for rearrangement and the fact that all generated heavy chains can pair with any light chain are considered combinatorial diversity. The junctional diversity is generated by a sloppy mechanism when cutting out the noncoding DNA, and by the addition of random nucleotides before ligation of the DNA. Four mechanisms, though, do not seem powerful enough to increase the variability to the amount that a few genes could cover millions of antigens. However, a running example with numbers will give a better picture. Assuming 51 VH, 27 D, and

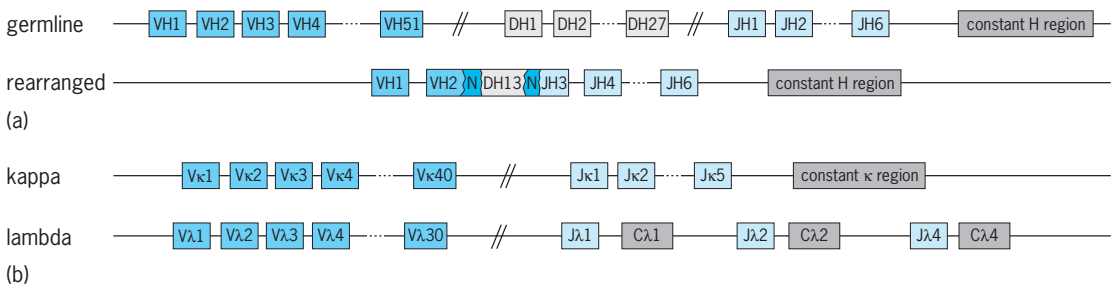

Fig. 2. Human immunoglobulin loci of heavy and light chains. (*a*) Heavy-chain immunoglobulin locus showing the lineup of the V, D, and J segments. The more complex constant region is depicted as a box (for more details, see Fig. 3). The germline configuration applies before rearrangement happens, or in a cell that does not undergo rearrangement (a liver cell, for example). In the example of a rearranged locus, VH2 has rearranged to DH13 and JH3. In the borders between the fragments, the cutting mechanism showed some sloppiness (wavy lines), and some N nucleotides have been added between the segments. Note that the intervening sequences from VH14 to DH12 and from DH14 to JH3 have been deleted from the genome. Double bars between segments indicate many kilobases of intervening noncoding sequence. (*b*) Light-chain immunoglobulin loci in germline configuration. While the kappa chain resembles that of the heavy-chain lineup (the only exception being that it does not have D elements), the lambda chain shows three constant regions with one linked J for each. Rearrangement to J always leads automatically to the expression of the linked constant region.

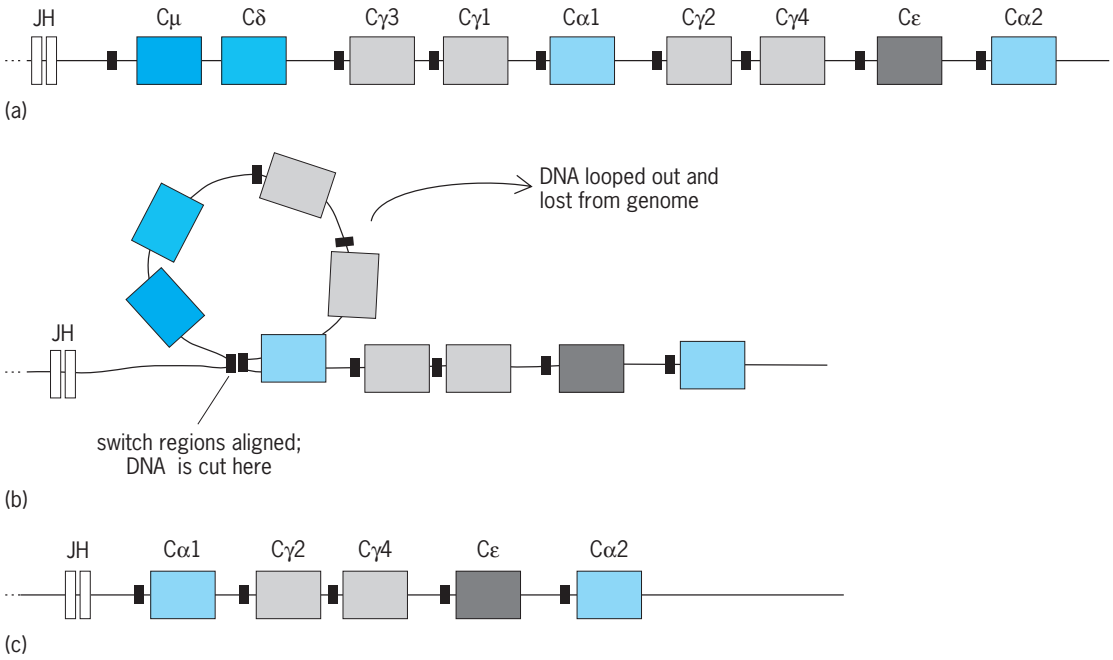

Fig. 3. Heavy-chain constant region. (*a*) Locus before class switch. The Greek letters in the constant-region names indicate the class of antibodies that will be expressed, μ for IgM, γ for IgG, etc. Some of the classes have multiple subclasses: in humans IgG actually is IgG1–4. Note that all classes have a switch region (little box) in front of their gene, except for Cδ, which is cotranslated with Cμ and then alternatively spliced. Thus it is possible to switch from IgM to any other class except IgD. (*b*) During class switch, the switch regions are brought into proximity, and DNA is looped out, cut, and religated. The intervening sequence is circularized and lost from the chromosome. (*c*) Constant region after class switch. Due to the loss of Cμ, the cell is never able again to make an IgM.

6 JH segments that can combine freely, the heavy-chain rearrangement alone gives $51 \times 27 \times 6 = 8262$ possibilities. If the light chain has $40\ \text{V} \times 5\ \text{J} = 200$ combinations, the free combination of 8262 heavy chains with 200 light chains gives $8262 \times 200 = 1.65$ million antibodies with different specificities. As we take into account the sloppiness of the process and the addition of N nucleotides, let us add an example: Assume that 10 different sequences are possible at each of the segment boundaries for any one combination of Vκ, Jκ, VH, D, and JH gene segments (this is an underestimate). This would multiply diversity by a thousandfold. Hence, the 1.65 million binding sites generated by combinational diversity alone is now amplified to 1.65 billion antibody specificities, starting out with less than 200 gene fragments.

Ig classes. While the light chain has only one constant region, the heavy chain has different constant regions available, representing different antibody classes (**Fig. 3**). These classes are called IgM, IgG, IgA, IgE, and IgD, and all have specialized functions. However, a B cell always starts out making IgM and IgD by default. After a cell has been activated and participates in an immune response, it can switch to other classes. This is a DNA recombination process that is very similar to VDJ rearrangement. Again, there are small regions that direct the reorganization of the DNA (here just called switch regions); there is one specific enzyme known to date that participates (AID: activation-induced cytidine deaminase); and as before, the DNA is looped, cut, and religated. Once IgG is made, the genetic information for IgM is lost and the cell can never again make IgM. Crucial in this class switch is that while the antibody changes the effector function by changing the class, it does not change the specificity.

Alternative splicing. Immunoglobulins not only are secreted as antibodies, but also serve a second purpose on the surface of B cells as B cell receptors (BCRs). The B cell receptor binds antigen, and once many of these are cross-linked by antigen, they signal the cell that it has been activated to participate in an immune response. But how can the same protein on one hand be secreted and thus be soluble in a watery solution (the serum), and on the other hand be anchored in the cell membrane and thus have a fatty region to stick to the membrane? The answer is alternative splicing, a process in which a long primary RNA is cut in different ways to encode for different proteins. In the case of immunoglobulins, alternative splicing determines if an immunoglobulin becomes a secreted antibody (the RNA is cut to encode for a hydrophilic protein tail so the antibody is soluble in serum) or a membrane-bound BCR (it encodes a stretch of lipophilic amino acids that can anchor the immunoglobulin in the membrane). This is the case for all classes, as all primary RNAs have the information for secreted and membrane-bound forms enclosed.

Another special case concerns the IgM and IgD classes. These two classes are actually encoded as a very long combined RNA transcript, and alternative splicing in this case not only determines whether it will be secreted but also whether IgM or IgD is made. All the other classes make only primary

transcripts for their class. *See* CELLULAR IMMUNOLOGY; IMMUNOCHEMISTRY; IMMUNOLOGICAL ONTOGENY; RIBONUCLEIC ACID (RNA). Katja Aviszus

Bibliography. G. M. Edelman, Biochemistry and the sciences of recognition, *J. Biol. Chem.*, 279(9):7361–7369, 2004; T. J. Kindt, B. A. Osborne, and R. A. Goldsby, *Kuby Immunology*, 6th ed., W. H. Freeman, 2006; R. R. Porter, Structural studies of immunoglobulins, *Science*, 180:713–716, 1973; S. Tonegawa, Somatic generation of immune diversity: The Nobel lectures in immunology, *Scand. J. Immunol.*, 38:303–319, 1993.

Immunologic cytotoxicity

The mechanism by which the immune system destroys or damages foreign or abnormal cells. Immunologic cytotoxicity may lead to complete loss of viability of the target cells (cytolysis) or an inhibition of the ability of the cells to continue growing (cytostasis). Immunologic cytotoxicity can be manifested against a wide variety of target cells, including malignant cells, normal cells from individuals unrelated to the responding host, and normal cells of the host that are infected with viruses or other microorganisms. In addition, the immune system can cause direct cytotoxic effects on some microorganisms, including bacteria, parasites, and fungi.

Immunologic cytotoxicity is a principal mechanism by which the immune response copes with, and often eliminates, foreign materials or abnormal cells. Cytotoxic reactions are frequently observed as a major component of an immune response that develops following exposure to foreign cells or microorganisms. In addition, there is increasing evidence that cytotoxic reactions represent a major mechanism for natural immunity and resistance to such materials. In most instances, cytotoxicity by immune components involves the recognition of particular structures on the target cells; also, the targets need to be susceptible to attack by the immune components. Some cells are quite resistant to immunologic cytotoxicity, and this appears to represent a major mechanism by which they can escape control by the immune system.

There are a variety of mechanisms for immunologic cytotoxicity. The two main categories are antibody- and cell-mediated cytotoxicity. Within cell-mediated cytotoxicity, there is a multiplicity of effector cell types and mechanisms that can be involved.

Antibody-mediated cytotoxicity. Cytotoxicity by antibodies is usually complement-dependent. Certain antibodies are of the classes of immunoglobulins that have the ability to bind to the series of serum proteins that collectively are known as complement. Immunoglobulin M and some subclasses of immunoglobulin G (IgG1 and 3 strongly, IgG2 weakly, in humans) have the ability to bind complement components, and only these are cytotoxic. *See* COMPLEMENT; IMMUNOGLOBULIN.

As part of an immune response to foreign materials, cytotoxic antibodies are often induced. Moreover, most normal individuals have natural antibodies against a wide variety of cells and microorganisms, and some of these natural antibodies can be cytotoxic in the presence of complement. For cytotoxicity by antibodies to occur, the antibody needs to first bind to the relevant antigens on the target cells, and this is followed by binding of a series of complement components to the immune complex. After binding of the full array of complement components, a susceptible target cell is rapidly lysed. The mechanism for such damage is thought to be related to the creation by the complement components of pores in the surface membrane of the cells, with subsequent loss of intracellular contents. *See* ANTIBODY; ANTIGEN.

Cytotoxic T lymphocytes. Immune T cells are the best-known effectors of cell-mediated cytotoxicity. Upon immunization with cells that differ from the host in their major histocompatibility antigens and are recognized as foreign, cytotoxic T cells are generated which have a potent ability to lyse a variety of target cells bearing the foreign histocompatibility antigens. In addition, cytotoxic T cells can be generated against a variety of other foreign antigenic structures on the surface of cells. This can include antigens associated with tumors of the host, and also host cells infected with various microorganisms, particularly viruses. In general, cytotoxicity against cells with these other antigens is displayed only when the cells share the same major histocompatibility antigens with the responding host. Only some subpopulations of T lymphocytes have the ability to develop cell-mediated cytotoxicity, presumably because they contain the necessary intracellular machinery for such damage.

The actual mechanism of killing by immune T cells is not well defined, but it is clear that binding of the immune cells to the targets is a requisite first step, and that this is followed by a series of metabolic events both in the immune T lymphocytes and in the target cells that eventually leads to lysis of the targets. The immune T cells themselves are not damaged during the interaction, and have the ability to go on to kill other target cells. Such unidirectional cytotoxicity of only the target cells and not the effector cells is characteristic also of the other effector cells discussed below. *See* CELLULAR IMMUNOLOGY.

Macrophage-mediated cytotoxicity. Macrophages and the related monocytes, the principal scavenger cells in the body, have the ability not only to ingest foreign cells or materials but also to cause cytolytic or cytostatic effects on a variety of target cells. Cytotoxic macrophages appear to play a major role in the elimination of various microorganisms. They also have an ability to discriminate between tumor cells and normal cells and to selectively damage a wide variety of tumor cells. Macrophages and monocytes from normal individuals frequently have low levels of cytotoxic reactivity. They can be activated to a potently cytotoxic state, mainly by interaction with a product of immune T cells that is referred to as the macrophage-activating factor. In addition, interferon that is produced by immune T cells or by other

cell types can activate macrophages to become cytotoxic.

As with cytotoxic T cells, the first step in the cytotoxic reaction by macrophages involves binding to structures recognized on the target cells. This appears to be followed by either of two types of metabolic events that can lead to damage of the targets. One major mechanism is the development in the activated macrophage of hydrogen peroxide and other reactive oxygen species, which can be very toxic against some microorganisms and other cells. Moreover, it appears that cytotoxic macrophages can release proteolytic enzymes which can cause damage and lysis of target cells. *See* BLOOD.

Natural killer cells. In the course of studies designed to find immune cytotoxic T cells in humans or animals with tumors, it was found that normal individuals also have lymphocytes with the ability to recognize and kill a variety of tumor cells. Such effector cells have become known as natural killer (NK) cells. It has been shown that natural killer cells represent a discrete subpopulation of lymphocytes which vary morphologically from typical lymphocytes and are known as large granular lymphocytes, because of their larger size and the presence of characteristic granules in their cytoplasm. Natural killer cells have the ability to recognize and lyse a wide variety of target cells, including tumor cells, some cells of the embryo and of normal bone marrow and thymus, and a variety of microorganisms. Normal fibroblasts or other cells, upon infection with viruses, become highly susceptible to lysis by natural killer cells. There is increasing evidence that natural killer cells play a major role in natural resistance against tumors, particularly against the metastatic spread of tumor cells. They have also been shown to be involved in natural resistance against infections by certain viruses, fungi, and parasites. In addition, they appear to play an important role in the rejection of grafts of foreign bone marrow cells. *See* TRANSPLANTATION BIOLOGY.

Granulocytes. Granulocytes are the most common type of white blood cell in the blood and tissues. These cells have the ability to mediate immunologic cytotoxicity against tumor cells and microorganisms. During the course of an immune response, proteins produced by immune T cells or other cells can activate the granulocytes to become highly cytotoxic. In addition, some components of microorganisms themselves have the ability to activate granulocyte-mediated cytotoxicity. A major mechanism for such cytotoxicity by granulocytes is the release of hydrogen peroxide or other reactive oxygen species, similar to that described above for macrophages.

Antibody-dependent cell-mediated cytotoxicity. Immunologic cytotoxicity also can be mediated by a collaboration between immune or natural antibodies and certain immunologic effector cells. It is known as antibody-dependent cell-mediated cytotoxicity, and it appears to be a potent mechanism for elimination of some microorganisms; it may also be involved in the destruction of tumors and other foreign cells. The initial event in antibody-dependent cell-mediated cytotoxicity is an attachment of antibodies, usually immunoglobulins of the G class, to target cells. In turn, certain types of effector cells are bound by means of cell surface receptors for a portion of immunoglobulin G that is shared by most antibodies (the Fc moiety). After binding of the effector cells to the targets, a series of metabolic events occurs which leads to the lysis of the targets. Antibody-dependent cell-mediated cytotoxicity can be mediated by a variety of lymphoid cells, particularly macrophages, natural killer cells, and granulocytes. Therefore, this represents an alternative mechanism by which these cytotoxic effector cells can recognize and destroy target cells. *See* IMMUNOLOGICAL DEFICIENCY; IMMUNOLOGY.

Ronald B. Herberman

Bibliography. E. S. Golub (ed.), *The Cellular Basis of the Immune Response: An Approach to Immunobiology*, 1980; R. B. Herberman (ed.), *NK Cells and Other Natural Effector Cells*, 1982; D. P. Stites and A. I. Terr, *Medical Immunology*, 9th ed., 1997.

Immunological deficiency

A state wherein the immune mechanisms are inadequate in their ability to perform their normal function, that is, the elimination of foreign materials (usually infectious agents such as bacteria, viruses, and fungi). Immune mechanisms are also responsible for the rejection of transplanted organs. The processes described above are accomplished by white blood cells known as lymphocytes, of which there are two major types, T lymphocytes (thymus-derived) and B lymphocytes (bone marrow–derived). *See* CELLULAR IMMUNOLOGY.

Development of competent immunity. T and B lymphocytes are derived from immature cells known as stem cells, which normally reside in the bone marrow. In order to perform their protective functions effectively, both T and B lymphocytes must undergo processes of maturation. For T cells, this development takes place in the thymus gland, while B cells mature in the bone marrow. *See* IMMUNOLOGICAL ONTOGENY.

In general, a given lymphocyte reacts to only one antigen. From this specialized lymphocyte descends a family of cells, which respond to that specific antigen. It has been estimated that normal humans have the capacity to recognize and respond to over a million different antigens; thus, a million different lymphocyte families or clones could be generated. Human lymphocytes recognize different antigens by having specific receptors on their cell surfaces. A lymphocyte responds to the antigen only if there is a fit between the lymphocyte receptor and the shape of the antigen. The reaction is similar to the fit of a lock and key. *See* ANTIGEN.

Upon encountering an antigen, the responsive lymphocytes increase markedly in number; thereafter, the two types of lymphocytes respond differently. The T-cell family forms a class of helper cells which enhance the response to the antigen,

and another set, called suppressor-killer cells, which perform a double duty. The killer function gives a cell the ability to destroy target cells, whereas the suppressor capacity slows down the immune response when it is sensed that the emergency is over and the immune system can be put to rest. The response to antigen by the B lymphocytes is to generate a large number of cells that manufacture and secrete proteins (immunoglobulins) that can physically combine with the antigen. *See* IMMUNOLOGIC CYTOTOXICITY.

The five major classes of immunoglobulins, IgM, IgG, IgA, IgE, and IgD, are known as isotypes. IgM is thought to be the first line of defense and responds very quickly to infections. It is mainly contained in the bloodstream, where it can readily neutralize agents attempting to gain entrance through the blood. IgG is the immunoglobulin which confers long-lasting protection and is transferred to the fetus through the placenta. In this way, the baby is protected by the mother's IgG for the first few months of life. Other Ig classes do not cross the placenta.

IgA is localized primarily in the lungs and intestine, where it protects from organisms which are inhaled or eaten. IgE causes the symptoms of hay fever and asthma. It is considered to have evolved as a component of the protective mechanism against parasites. For instance, its release causes the intestine to contract violently, which may expel an intestinal worm. *See* IMMUNOGLOBULIN.

When either a killer T cell or an immunoglobulin molecule combines with the antigen, other body reactions are brought into play to make the elimination of the antigen more effective and complete. T cells secrete a number of molecules known as lymphokines which attract other cells that are also able to destroy target cells and in general make the body more resistant to the antigenic material. The immunoglobulins can combine with another set of proteins, known as the complement system. The complement cascade consists of 11 proteins which act in a sequential manner to destroy whatever target is coated by the immunoglobulin. *See* COMPLEMENT.

Causes. Immunodeficiency diseases (IDDs) result from a failure at any point in the complex set of interactions described above. In general, the diseases are due to absence of cell populations; failure of cells to mature; failure to secrete the products (such as lymphokines) necessary for effective cell interactions; or failure of accessory cell populations or protein systems (for example, complement) which are necessary for the complete competence of B-cell immune function. An immune response begins when immune system cells are triggered by an antigen to which the individual has been immunized. This response consists of recognition of the stimulus (often an infection) by external cell receptors which leads to a series of internal cell signals that initiate the process of cell synthesis and secretion. These responses to external stimuli are known as signal transduction. Defects of enzymes vital in signal transduction account for many immunodeficiency diseases. Some of the diseases are carried on the X chromosome and affect only males, being carried by females. *See* SEX-LINKED INHERITANCE.

In recent years, impressive progress has been made in defining causes of IDDs. The specific protein defect and the chromosomal location of its gene have been defined (**Table 1**).

Symptoms. It is apparent from a consideration of the functions of the immune system that the prime symptom of immunodeficiency would be an increased susceptibility to infections. Since T cells and B cells differ in the effectiveness with which they protect against various kinds of organisms, some insight into which system (T cell or B cell) is primarily

TABLE 1. Genes involved in immunodeficiency disease

Disease	Gene product affected	Chromosome location of gene
Severe combined immunodeficiency (X-linked)	γ_c; component of several cytokine receptors	Xq13.1–13.3
Severe combined immunodeficiency (non-X-linked)	Jak3, necessary in signal transduction	19p13.1
Ataxia-telangiectasia	Ataxia-telangiectasia-mutated (ATM), which is part of mechnisms that repair damaged DNA, a process necessary for the continued production of normal cells	11q23.1
Wiskott-Aldrich syndrome	Wiskott-Aldrich syndrome protein (WASP), part of a complex protein system which is critical for antigen recognition and signal transduction	Xp11.23
Adenosine deaminase deficiency	Adenosine deaminase, necessary for catabolism of cellular breakdown product; in its absence, toxic metabolites develop which destroy and B cells	20q13-ter
Nucleoside phosphorylase deficiency	Nucleoside phosphorylase, an enzyme with functions similar by to adenosine deaminase but acts at a different site in the cellular breakdown metabolic pathway. the accumulated toxic metabolites destroy T cells primarily	14q13.1
DiGeorge anomaly*	Specific protein defect not defined as yet	22q11.2[2]
X-linked hypogammaglobulinemia	Bruton tyrosine kinase, an intracellular protein important in cell signaling pathways	Xq21.3–22
X-linked hyper IgM syndrome	CD40 ligand (CD40L), a protein on T cells which reacts with the B-cell membrane protein CD40, causing B cells to switch from IgM to IgG and IgA synthesis	Xq26.3–27.1

*Some cases are associated with deletions in chromosome 10p.

TABLE 2. Combined T- and B-cell diseases

Disease	Distinguishing clinical features*	Remarks
Severe combined immunodeficiency	Total lack of hair or severe skin infections may be an early sign	Often fatal by age 2
Short linked dwarfism or cartilage hair hypoplasia	Very short legs Characteristic x-rays Children have fine thin hair Double-jointed	There may be susceptibility only for chickenpox
Ataxia-telangiectasia	Prominent blood vessels in white of eyes, ear lobes, and cheeks Poor coordination Ultimately mental retardation Poor tolerance to x-ray therapy	High incidence of cancer
Wiskott-Aldrich syndrome	Severe eczema Frequent severe herpes infection Low blood platelets cause excessive bleeding X-linked	High incidence of cancer
Adenosine deaminase deficiency	May have characteristic x-rays with disease involving ribs, pelvis, arms, and legs	B-cell deficiency may be present early or occur with time; first disease to be treated by gene therapy

*For general symptoms common to all diseases, see text.

affected can be gained by considering the kind of organisms involved. T cells primarily fight off infections caused by yeasts, viruses (causing chickenpox or measles), fungi, and certain bacteria, such as those causing tuberculosis. B cells are particularly effective against most bacteria (for example, staphylococci, streptococci, and pneumococci). There is also evidence that tumor cells are eliminated by cells of the immune system (probably T cells and a related type known as the natural killer cell). In any event, the incidence of some types of cancer (largely of the reticuloendothelial system) in immune-deficient populations is 10,000 times that of normals.

Many of the organisms to which people are constantly exposed do not ordinarily have the capability to cause infections in immunocompetent individuals because these organisms are so weak that they cannot establish themselves in normals. In immunodeficients, however, they can cause fatal infections. One such organism is a protozoon, *Pneumocystis carinii*, which can cause debilitating pneumonia. In other situations, agents which usually cause only mild infections lead to fatal disease. Chickenpox, cytomegalovirus, and measles will cause fatal pneumonia in deficient individuals. Chronic parasitic infections leading to severe diarrhea are also characteristic of this deficient group. Thus, immunodeficient individuals respond severely or fatally to very mild or ordinarily innocent infectious agents.

Specific diseases. There are many varieties of immunological deficiency diseases. In general, it is sufficient to think of the diseases as involving primarily the T-cell system, the B-cell system, or both cell lines. Involvement of both T and B cells is known as combined immunodeficiency disease and is the most severe form. It is often fatal, usually before the age of 2 years. Defects which involve the T cells but not the B cells are the next most severe. Pure B-cell deficiencies are the mildest and are known as hypogammaglobulinemias. In some of the variants, abnormalities of the red blood cells, white blood cells, or blood platelets are associated. Occasionally, one also finds biochemical abnormalities and disorders of the bones (**Tables 2–4**).

It is possible to have single immunoglobulin deficiencies. The most common immune defect is an isolated deficiency of IgA, thought to occur in approximately 1 out of 600 people. In most individuals with selective IgA deficiency, compensatory mechanisms keep those affected individuals from having serious disease. In fact, many people with IgA deficiency are unaware of their abnormality.

Most of these problems are due to inherited faults. In general, immunodeficiency states are recessive traits, inherited as either autosomal or X-linked traits. As a result, approximately 25% of all children in such families can be expected to be affected. In the case of X-linked disorders, females will not be affected but

TABLE 3. Diseases predominantly involving T cells*

Disease	Distinguishing clinical features*	Remarks
Nucleoside phosphorylase deficiency	Severe anemia is often associated	
Chronic mucocutaneous candidiasis	Skin lesions involve areas on face, and trunk, nails, and often lead to severe disfigurement	Often associated with diseases of endocrine glands causing hypothyroidism and diabetes May be due to biotin defect
DiGeorge anomaly	Associated with characteristic facial appearance (receding chin, low-set ears, widely spaced eyes), severe heart disease, and low calcium	Occurs with all degrees of severity and often has associated B-cell defect

*For general symptoms common to all diseases, see text.

TABLE 4. Diseases predominantly involving B cells

Disease	Distinguishing clinical features	Remarks
X-linked hypogammaglobulinemia	Disease present at birth	B cells absent
Congenital hypogammaglobulinemia (Bruton's syndrome)	Tonsils and lymph nodes very small or absent	Occasionally associated autoimmune phenomena (rheumatoid arthritis, skin and muscle disorders)
		Responds well to gamma globulin injections or infusions
Common variable immunodeficiency	Disease usually begins later in life	Variable degree of T-cell defect more common B cells present but nonfunctional One variety associated with thymus gland tumor (thymoma)
X-linked hyper IgM	Neutropenia common IgG and IgA absent, IgM normal or at elevated levels	Pneumocystis pneumonia may occur Low platelet number: severe intestinal and liver disease may be seen
Chronic lymphatic leukemia	Large spleen	Cancer of the lymphocytes; 25% have B-cell deficiency
Multiple myeloma	Gammaglobulin levels paradoxically very high	Deficiency due to replacement of effective immunoglobulin molecules by the abnormal myeloma protein
	Disease commonly produces holes in bones due to involvement by cancer	

will carry the trait. In this instance, the 25% figure is derived from a 50% incidence in male births, which occur 50% of the time.

Immunodeficiency can also be acquired as a complication of other disease processes. One of the most common forms of deficiency is caused by aggressive treatment of leukemia. Unfortunately, when the cancer cells are destroyed, the normal lymphoid cells are killed as well. As yet there are no known agents which are selective only for cancer cells; in most cases they preferentially react against the malignant cell, but in the high dose required, the normal cells ultimately are lost, too. Another cause of induced immunodeficiency is seen with transplant rejection therapy. To prevent organ rejection, drugs which destroy lymphocytes must be administered. The strategy is to employ enough medicine to prevent rejection without completely losing all lymphocyte function. This fine line is frequently crossed, however; and serious, even fatal, infection can complicate organ transplantation. *See* LEUKEMIA; TRANSPLANTATION BIOLOGY.

A fascinating variety of immunodeficiency is secondary to virus infections. Certain viruses, such as the Epstein-Barr virus (EBV), which causes infectious mononucleosis, infect lymphocytes. Involvement of the lymphoid system is nearly always only temporary, and most individuals can clear the virus after a few weeks. In a small number of individuals, probably as a result of a genetic defect, the virus cannot be eliminated, and a chronic infection of B cells leads to the loss of normal lymphocyte function. Hypogammaglobulinemia has been described after EBV infection. An overwhelming immunodeficiency disorder, known as the acquired immune deficiency syndrome (AIDS), has been observed primarily in homosexuals, intravenous drug users, and hemophiliacs. The cause is infection with a virus termed the human immunodeficiency virus (HIV). *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); EPSTEIN-BARR VIRUS.

Immunodeficiency can be observed secondary to dietary deficiency. Two main varieties are seen. In protein-calorie malnutrition, serious deficiency primarily involving the T-cell system predisposes affected individuals to overwhelming infection by the agents of measles or tuberculosis. In the second variety, deficiency of single substances is the cause; the two most commonly observed deficiencies are those of zinc and biotin. It should be pointed out that deficiency disorders are caused by extremely severe dietary restriction or rare absorption defects, and that among the general population it is not necessary to enhance the diet with massive dietary supplements to enhance immunity. In fact, ingestion of tremendous amounts of many supposedly innocuous foods or vitamins can be dangerous. *See* BIOTIN; MALNUTRITION; VITAMIN.

Tests of immune function. The capability of the immune system can be precisely estimated by examination of the blood. The T cells and B cells can be stained with special reagents that are highly specific for each cell type. It is even possible to differentiate between the special varieties of T cells, that is, the helper and the suppressor-killer types. Helper cells are designated CD4, and suppressor-killers are known as CD8 according to their reactivity with corresponding monoclonal antibodies. B cells are designated CD19. The different classes of immunoglobulins are measured by chemical tests. By taking samples from thousands of normal persons, values in healthy individuals have been established. One can then compare the values observed in a patient to determine if a deficiency exists. *See* MONOCLONAL ANTIBODIES.

Sometimes the cells or immunoglobulin proteins are present in normal amounts, but their functional capability is weakened. That is, they will not combine with, or respond to, antigens. Special tests can be employed to actually mimic an immune response in the test tube. These tests make it possible to determine whether a person's lymphocytes can behave in a normal manner when stimulated by antigens.

A form of immunodeficiency associated with enzyme defects has been described. The enzymes

involved are adenosine deaminase and nucleoside phosphorylase (NP), necessary, respectively, for the catabolism of the purines adenosine and inosine. In the absence of these enzymes, abnormal metabolites accumulate and are toxic to the lymphocytes. The T cells are more susceptible to the toxicity, and hence the defect primarily involves that cell line. Ultimately, in adenosine deaminase deficiency, B cells as well become involved, so that combined deficiency occurs with time. For unknown reasons, the B cells remain uninvolved in nucleoside phosphorylase deficiency. These biochemical defects are inherited as autosomal recessive traits.

Further information can be obtained by studying the sites of the immune reaction by taking a sample of a lymph node. A small sample of the thymus gland can also be studied. Lymphoid tissues which are functioning normally show characteristic features under the microscope.

Diagnosis of immunodeficiency diseases can be made in utero in many cases. This has been greatly facilitated by molecular definition of the defect and determination of the chromosomal localization of the gene involved. In addition, fetal blood samples can be obtained and the cells tested in the same manner as cells taken from patients in the clinic.

For X-linked diseases, females at risk (aunts and sisters of patients) can be studied and the carrier status determined.

Treatment. The treatment depends upon the defect. Pure B-cell disorders are the simplest to treat and usually respond to immunoglobulin injections or infusions. These are required for the rest of the life of the individual.

Pure T-cell disorders are very rare but can be treated by thymus transplantation. It is not necessary to place the organ in its natural site (in the chest underneath the breastbone); the organ can be transplanted in a large muscle. The rich blood supply of the muscle will bring the stem cells from the bone marrow to the transplant, where maturation of T cells occurs, allowing replacement of the missing cell population. Since individuals with T-cell deficiency are unable to reject organs, it is easy to perform the transplant without any concern about rejection of the graft. *See* THYMUS GLAND.

In combined immune deficiency states, it is usually necessary to provide both B and T cells. This is best accomplished by a bone marrow transplant. The theory is that the bone marrow provides stem cells which can mature in the recipient. However, it is now apparent that the bone marrow contains mature as well as immature cells, and exactly how much of the reconstitution is due to immature cells and how much to mature cells is presently unknown. It is not possible to transplant bone marrow between any two individuals chosen randomly. The reason is that the mature T cells in the bone marrow suspension can react vigorously against any foreign substance. In the case of a transplant, that foreign substance would be the recipient and the donor bone marrow could actually "reject" the individual.

This is exactly the same reaction which occurs when an immunocompetent individual rejects a transplant. It is called host-versus-graft reaction; in this case, the graft dies. In the bone marrow situation, the marrow is the graft, and when it rejects the recipient, it is called graft-versus-host reaction. This complication can be fatal for the host. Graft-versus-host disease is controlled in two ways. First, the donor and recipient must be matched as closely as possible according to bone marrow types. In principle, the matching is similar to matching people for blood transfusions but, in practice, is much more difficult because the lymphocyte system is so much more complex than the red blood cell system. There are only four major types of blood, and it is relatively easy to find compatible types among the population at large. There are thousands of lymphocyte variants, and in general, only family members can be considered as possible donors. As a practical matter, usually only brothers or sisters are checked, but even then, a match will be found only 25% of the time. On a rare occasion, a parent or more distant relative may match, especially if there has been marriage among close relatives (for example, cousins). The chance of finding a match between two unrelated individuals in the population at large is approximately 1 in 10,000. *See* BLOOD GROUPS.

It is also possible to eliminate the mature T cells from the marrow before transfusing it into the patient. This strategy has been used successfully in a number of transplants between parent and child. Because of the nature of the transplantation antigens, in nearly all cases, a parent will be matched to the child for one-half of all antigens. (Each person carries two sets of transplantation antigens—one set is inherited from each parent.) If a bone marrow transplant can be successfully accomplished, the individual can be expected to be cured.

Umbilical cord blood from closely matched donors is now used with great success. The T cells from umbilical cord blood are much less aggressive in causing graft versus host disease.

With the cloning of genes which cause immunodeficiency disorders, genetic reconstitution becomes possible. Adenosine deaminase (ADA) deficiency was chosen for the first clinical attempt. The gene was well suited for this trial, since ADA is a single-chain enzyme and its production does not require strict control. Thus, complicated problems of protein assembly are avoided, and increased levels are not harmful to the patient. Stem cells were isolated from the umbilical cord blood of three patients with ADA deficiency, the gene inserted, and the stem cells infused back into the patients. After 2 years, genetically corrected cells were still present in the patients' blood and showed levels of ADA cDNA at 1/1000 to 1/10,000. However, injectable ADA which had been given to support the patients until sufficient enzyme could be produced by the corrected cells could not be decreased, so the levels of enzyme produced are not sufficient to correct the defect, as assessed 4 years after the cell infusion.

It was stated above that lymphocytes secrete various molecules which augment the immune

response. Some of the substances have been purified and can be made in large quantity either by artificial synthesis or by gene-cloning methods. Substances which have been used include interleukin-2 (IL-2), interferon, and many of the hormones elaborated by the thymus gland (such as thymosin, thymopoietin, and TP-5). *See* GENETIC ENGINEERING; IMMUNOLOGY; THYMOSIN.

Richard Hong

Bibliography. A. Fischer and B. Malissen, Natural and engineered disorders of lymphocyte development, *Science*, 280:237–242, 1998; F. S. Rosen et al., Primary immunodeficiency diseases, *Clin. Exper. Immunol.*, 109 (suppl. 1):1–28, 1997; E. R. Stiehm (ed.), *Immunologic Disorders in Infants and Children*, 4th ed., W. B. Saunders, Philadelphia, 1996; K. I. Weinberg (ed.), Molecular basis of lymphoid immunodeficiency, *Sem. Hematol.*, 35(4):280–367, 1998.

Immunological ontogeny

The origin and development (ontogeny) of the lymphocyte system, from its earliest stages to the two major populations of mature lymphocytes: the thymus-dependent or T lymphocytes, and the thymus-independent or B lymphocytes. The T lymphocytes carry out those aspects of function which are called cell-mediated immunity, including graft rejection, elimination of tumor cells, and delayed hypersensitivity. B cells are responsible for humoral or antibody-mediated immunity. *See* IMMUNITY.

For both systems, development or differentiation proceeds in discrete stages (see **illus.**). In the first stage, pluripotent hematopoietic stem cells, which originate in the yolk sac in the embryo and then successively in fetal liver and fetal bone marrow, develop into precursor cells committed to becoming T or B cells. Hematopoiesis in human fetal liver begins at about 4 weeks of gestational age and in fetal bone marrow after 20 weeks. *See* HEMATOPOIESIS.

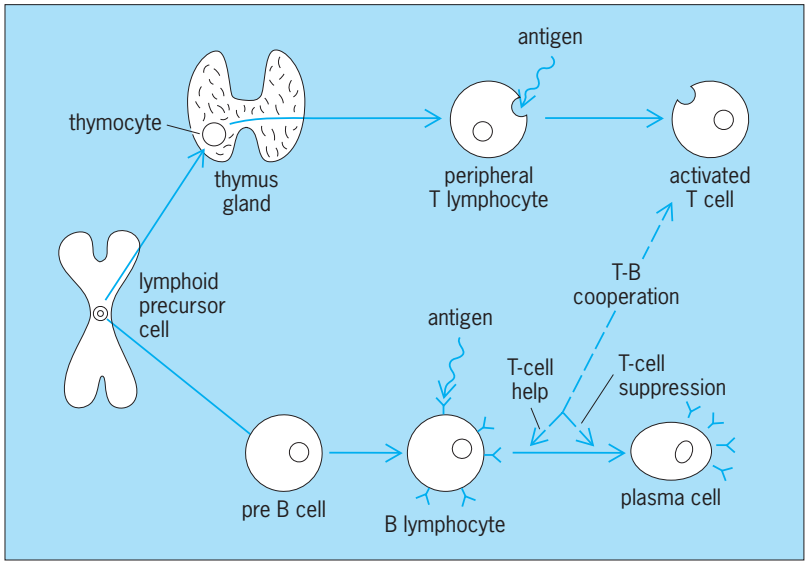

Schematic representation of the differentiation of the immune system.

Ontogeny of T lymphocytes. The thymus plays a strategic role in the development of T lymphocytes. It arises from the third and fourth pharyngeal pouches, and between the seventh and tenth week of gestation it descends from the neck into the anterior mediastinum. The thymus is highly organized with a network of reticuloepithelial cells. T-lymphocyte precursor cells are attracted into the thymus where, under the influence of this microenvironment, they undergo rapid proliferation and maturation. In humans, stem cells enter the thymus beginning in the eighth week of gestation and then begin to undergo differentiation. The epithelial cells within the thymic stroma are primarily responsible for this development, because they contain hormonelike substances which trigger T-cell differentiation. By the ninth or tenth week, fetal thymocytes proliferate in response to phytohemagglutinum and allogeneic cells, indicating T-cell maturation. These cells also begin to express a variety of cell-surface markers, which parallels developing immunocompetence. From the thymus, these maturing T cells are exported to the peripheral lymphoid tissues, including lymph nodes, gut, spleen, and peripheral blood. In the developing peripheral lymphoid tissues, these T cells populate the deep and midcortical areas of the lymph nodes and periarteriolar regions of the spleen. Immunocompetent T cells can be identified in these tissues by the time of 13–14 weeks' gestation. *See* THYMUS GLAND.

As maturation proceeds, the T cells become specialized and different subsets can be identified through the detection of different cell-surface antigens by monoclonal antibodies or by measuring distinct functions. These functions include T-helper-cell activity which, on interacting with B cells, promotes their maturation into antibody-secreting cells, and T-suppressor-cell activity, which regulates this response.

Ontogeny of B lymphocytes. The developmental pathway of B lymphocytes has been studied in greatest detail in the chicken. In this species, the maturation of B cells begins with migration of progenitor cells to a specific microenvironment within the bursa of Fabricius (an outpouching of the hindgut, hence the term B or bursa-dependent lymphocytes). In mammals, there does not appear to be a discrete and specialized environment analogous to the bursa of Fabricius. The bone marrow seems to be the major site of B-cell differentiation in mature mammals. Similarly to the T cells, B lymphocytes pass through several differentiation steps in their ontogeny, which may be characterized by the expression of specific markers.

The earliest B cells identified in fetal liver are pre-B cells. As the cells mature, they express immunoglobulin M (IgM) and subsequently IgD on their surface. At this stage, the cell is a specific, competent B cell ready to interact with an antigen. In the course of B-lymphocyte differentiation, diversity of immunoglobulin classes is generated by an orderly switch from IgM to IgG to IgA with expression of the respective immunoglobulin on the cell surface.

This ontogenic sequence is supported by the expression of constant-region heavy-chain genes. *See* IMMUNOGLOBULIN.

Cells bearing surface IgM, IgG, and IgA have been demonstrated by immunofluorescence in peripheral blood, liver, spleen, and bone marrow at 11–12 weeks' gestation, and cells containing cytoplasmic immunoglobulins by 20 weeks in human fetuses. Synthesis of IgM has been noted as early as 10–11 weeks' gestation, and synthesis of IgE and IgG shortly thereafter. *See* IMMUNOFLUORESCENCE.

For B cells, the surface immunoglobulin serves as a specific receptor for antigen. Once activated by antigen and following interaction with T lymphocytes and monocytes, B lymphocytes terminally differentiate into mature antibody-secreting plasma cells. The immunoglobulin-producing cells differentiate relatively early in gestation, and the human fetus can produce significant amounts of antibody if stimulated by antigen. The development of a fetal antibody response to a specific antigen depends on the nature of the antigen and on the maturation of specific clones of antibody-producing cells, or may be limited by the ability of fetal monocytes to process certain antigens. Furthermore, active infection of the fetus may hasten or retard the process of developing immunocompetence.

Immunoglobulin gene expression. The initial step in B-cell differentiation is a genomic rearrangement of the immunoglobulin gene complex that brings a variable-region gene segment into the site of a constant-region gene segment. Once the immunoglobulin genes have undergone the rearrangements that are necessary to encode a complete molecule, immunoglobulin is synthesized. On the surface of the cell, the immunoglobulin can bind antigen and signal the next step in differentiation. This involves a switch from surface to secreted immunoglobulin. *See* IMMUNOGENETICS.

Diversity of immunoglobulin classes with different heavy chains (IgM, IgG, IgA, IgD, and IgE) is generated through a heavy chain constant region switching mechanism. Analysis of the heavy-chain locus of both humans and mice has demonstrated that the eight different types of heavy-chain constant-region genes are encoded on a single stretch of DNA. The signals responsible for heavy-chain class switch are not understood. The different classes of heavy chain confer different biological properties on the immunoglobulin molecule. *See* CELLULAR IMMUNOLOGY.

Erwin W. Gelfand

Bibliography. A. Baumgarten and F. Richards (eds.), *CRC Handbook Series in Clinical Laboratory Science*, vol. 1, 1978; M. D. Cooper et al., The functions of the thymus system and the bursa system in the chicken, *J. Exp. Med.*, 123:75–102, 1966; E. W. Gelfand and H.-M. Dosch (eds.), *Biological Basis of Immunodeficiency*, 1980; A. R. Lawton et al., Ontogeny of B-lymphocytes in the human fetus, *Clin. Immunol. Immunopath.*, 1:84–93, 1972; P. Leder, Genetic control of immunoglobulin production, *Hosp. Pract.*, 18:73–82, 1983; O. Stutman, Two main features of T-cell development: Thymus traffic and post-thymic maturation, *Contemp. Top. Immunobiol.*, 7:1–46, 1977.

Immunological phylogeny

The study of immunology and the immune system in evolution. All vertebrates can recognize and respond to nonself-molecular configurations on microorganisms, cells, or organic molecules by utilizing a complex recognition system termed the immune response. The presence of lymphocytes and circulating antibodies has been documented in all extant vertebrate species, ranging from the ostracoderm-derived cyclostomes, such as hagfish and lampreys, to mammals. However, the existence of induced, specific reactions directly homologous to the immune repertoire of vertebrates has not been clearly established in invertebrates. Studies have been done which focused on the presence and characterization of cells of the T and B lymphoid series in lower vertebrate species; the nature of the antibodies produced by lower vertebrate species and their relationship to well-characterized immunoglobulin classes of mammals; the extent of diversity in the antibodies of lower vertebrates with particular respect to the degree of variable-region heterogeneity within lower species compared with that exhibited by mammals; and the deoxyribonucleic acid (DNA) sequence of genes encoding variable regions of immunoglobin of lower species. Studies designed to pinpoint the origin of immune diversification have been actively pursued in the protochordates, such as tunicates, and there is also considerable interest in the defense mechanisms of invertebrate phyla not directly related to the evolutionary stream which gave rise to the chordates.

Cellular immune responses. The role of phagocytic cells in engulfing foreign pathogens has been documented in virtually all metazoan organisms. Phagocytic cells possess a limited capacity to discriminate self from nonself, and this is due in part to the presence of lectins (molecules capable of binding specifically to various sugars) on their surface. This conclusion has been directly shown for phagocytic cells of the oyster, and apparently correlates with the fact that most invertebrates possess lectins of various specificities circulating in their hemolymph. Although there is no evidence to suggest that invertebrate lectins and vertebrate immunoglobulins are homologous structures, sufficient diversity exists within lectins of certain species to indicate that these types of molecules and their cellular expression on phagocytes might serve as a primitive and universal recognition mechanism. *See* IMMUNOGLOBULIN; LECTINS.

Colonial tunicates (protochordates) have shown the capacity to reject allografts, a phenomenon comparable to T-cell–mediated graft rejection of true vertebrates, although the existence of true lymphocytes in protochordates has not been firmly established. All true vertebrates possess cells clearly recognizable as lymphocytes and can carry out T-cell

functions, such as graft rejection, and show the capacity of B cells to synthesize and secrete immunoglobulins. The teleost fish and amphibians possess lymphocytes of distinct functional and cell-surface–maker phenotypes which resemble T and B lymphocytes of birds and mammals. In addition, functional assays indicate that teleost fishes and amphibians possess helper T lymphocytes which are capable of interacting specifically with B cells in the production of circulating antibody. Lampreys and hagfish lack discrete lymphoid organs characteristic of higher vertebrates but contain diffuse collections of lymphoid cells in the area of the gills, which may correspond to a primitive thymus, and in the gut, which may correspond to the spleen of higher vertebrates. By contrast, all placoderm-derived vertebrates (sharks, bony fishes, amphibians, reptiles, birds, and mammals) have clearly demarcated thymuses and spleens. True lymph nodes are not present in vertebrate species more primitive than mammals, but birds possess aggregates of lymphoid tissue probably serving a similar function. The thymus is a primary lymphoid organ which generates cells of the T-lymphocyte series. In birds, a discrete organ, the bursa of Fabricius, is the primary source of B cells. The precise anatomical location of the bursa is not known in mammals or vertebrate species distinct from birds, but some sort of functional equivalent must exist. *See* LYMPHATIC SYSTEM; THYMUS GLAND.

B cells of teleosts and amphibians express surface immunoglobulin, which has been isolated and characterized. It resembles in overall properties the 7S IgM (IgMm) which is associated with the membrane of B lymphocytes of mammals. Mammalian B lymphocytes express, in addition, IgD membrane immunoglobulin. This has not yet been conclusively shown on the surfaces of B lymphocytes of teleost and amphibian species. The nature of the antigen receptor on T lymphocytes is not exactly known, but studies with T lymphocytes of teleost fish suggest that a molecule related to immunoglobulin might serve as the specific antigen receptor on these cells. *See* CELLULAR IMMUNOLOGY.

Immunoglobulins of lower species. Humans possess five major classes or isotypes of immunoglobulin: IgG, IgM, IgA, IgE, and IgD. These classes are defined by the type of heavy chain (γ, μ, α, ϵ, or δ) which occurs in disulfide-bonded association with light chains (either of κ or λ type). The usual structure of an immunoglobulin consists of two heavy chains joined to two light chains; in order to form the combining site to antigen, a light chain must be associated with the heavy chain. The combining site for antigen is made up by interaction of the variable (V) regions of a light chain in association with that of a heavy chain. A molecule consisting of two light (L) and two heavy chains would thus have two combining sites. The variable regions constitute approximately the first 100 N-terminal residues (amino acids at the amino terminus of the peptide chain) of light and heavy chains, and the remainder of the molecules are common (or constant) within a given isotype.

Some immunoglobulins, notably IgM (immune macroglobulin), exist as polymers of the basic four-chain unit. The IgM molecule normally consists of a cyclic pentamer of $L^2\mu^2$ subunits. This is the first immunoglobulin to appear in ontogeny, and the first to appear in phylogeny. Immunoglobulins of cyclostomes, elasmobranchs (sharks and rays), and many teleost fishes consist only of IgM polymers (5-mers in elasmobranchs, 4-mers in teleosts) and in some case IgM subunits (μ^2L^2). Immunoglobulins possessing heavy chains distinct from the μ-like heavy chains of those groups are present in some lungfish (Dipnoi) and in anuran amphibians (frogs and toads). Dipnoi have a low-molecular-weight non-IgM immunoglobulin (termed IgN) that possesses a heavy chain with a molecular weight of approximately 36,000. The μ chain has a weight of approximately 70,000, and the γ chain approximately 50,000. Anuran amphibians show a classical progression of antibody activity from high-molecular-weight molecules (IgM) to molecules of approximate weight 180,000, which is apparent with time following immunization. The low-molecular-weight immunoglobulin is not IgM, and it does not correspond directly to IgG or other of the non-IgM immunoglobulins of mammals. Amino acid sequence data are required to allow definitive conclusions regarding its direct homology to mammalian classes, but the data suggest that the heavy chain of the low-molecular-weight immunoglobulins of anuran amphibians represents a parallel duplication of a heavy-chain gene distinct from that which generated the gamma chain of mammals. Birds possess IgM and IgA immunoglobulins, but also possess a non-Ig immunoglobulin similar to that of amphibians as their major immunoglobulin class. This immunoglobulin has been termed IgY. IgG immunoglobulins containing gamma chains clearly homologous to those of humans and of true mammals are found only within the three subclasses of living mammals: eutherians, metatherians (marsupials), and monotremes (for example, the echidna).

Duplication of constant-region genes appears to be an essential feature of immunoglobulin evolution both on the large scale of emergence of vertebrate classes and on the smaller scale of speciation where distinct subclasses still arise from the major classes which might have emerged early in vertebrate evolution. The constant regions consist of domains or internal homology units of approximately 100 amino acids and contain an internal disulfide bond. These domains are specified by individual gene segments or exons occurring within the heavy-chain chromosome. For example, the μ chain consists of five domains, one variable and four constant, each of which is encoded by a separate exon. The emergence and evolutionary variation of the constant regions in immunoglobulin has resulted from duplication or deletion of gene segments corresponding to the domains of the actual proteins. Selective forces have acted upon the proteins produced because each domain carries out a particular effector function, such as fixation of complement or binding to cell surfaces. *See* EXON; IMMUNOGENETICS.

Variable-region diversity in lower vertebrates. Although all vertebrates can produce recognizable immunoglobulin, and sharks and higher vertebrates possess immunoglobulin clearly similar to that of mammals, the degree of diversity within the V-region compartment of lower vertebrates is probably less than that of mammals. Sharks, for example, can make specific antibodies to a wide variety of antigens, including small organic molecules (haptens), viruses, and foreign cells, but the affinity of these antibodies is low and does not increase with the duration of immunization. Although the size of the antibody can shift from IgM pentamer to IgM monomer with prolonged immunization, only the IgM class is represented and the affinity does not increase. By contrast, immunization of a common mammal, such as the rabbit, with a hapten usually generates first a primary response which is predominantly IgM with low affinity in the combining site. However, with prolonged immunization, the predominance of antibody activity is shifted to the IgG class and the affinity of the antibody for the hapten may increase by three or four orders of magnitude. Failure of affinity maturation is also a property of IgM antibodies of higher vertebrates and may reflect a lack of somatic mutation in the V_b genes that become associated with the μ constant regions. In addition, analyses, such as isoelectric focusing (which provides a means of counting the number of distinct antibodies on the basis of charge) and amino acid sequence analysis, indicate that in the shark and in the anuran amphibian *Xenopus*, the number of antibodies capable of responding to a hapten, such as dinitrophenol, is much less than that of mammals. The evidence suggests that the degree of heterogeneity of amphibian and fish antibodies is less than that of mammals. Although the basic event underlying immune recognition must be ancient and possibly preceded the origin of vertebrates, subsequent developments for increasing diversity, such as somatic mutation, apparently do not occur in lower vertebrates.

The primal event in the emergence of the classic immune system was the generation of V-region genes, probably by tandem duplication of a precursor and the joining of one of these genes to a constant-region gene in order to allow formation of an immunoglobulin heavy or light chain. It is possible that V-region genes and their products preceded constant regions which are necessary for effector functions because specific alloreactive lymphocytes, comparable to T cells, occur in species, such as tunicates, which lack the capacity to form circulating antibodies. Substantial evolutionary conservation has been found by comparing the sequences of heavy-chain variable-region genes of a reptilian species, the caiman, with those of humans and mouse. It is interesting to note that the caiman gene sequence clearly fits into the third heavy-chain variable-region subgroup of humans. This indicates that the generation of these subgroups preceded the events of speciation which led to the demarcation between reptiles and mammals.

Although the precise nature of the precursors of the specific elements of the immune system in evolution remains to be determined, the genetic and cellular events which lead to the capacity for specific immune recognition, diversification, and reactivity occurred early in vertebrate evolution. V- and C-region genes and a mechanism for joining them must have emerged prior to the emergence of true vertebrates, although somatic diversification mechanisms resulting in amplified repertoires and affinity maturation may not have arisen until relatively late in vertebrate evolution corresponding to the phylogenetic levels of birds and mammals. The specific immune elements were grafted onto a primitive and universal mechanism involving phagocytic cells and lectins. *See* IMMUNITY; IMMUNOLOGY. John J. Marchalonis

Bibliography. N. Cohen and M. M. Sigel (eds.), *The Reticuloendothelial System*, vol. 3: *Phylogeny and Ontogeny*, 1982; L. Du Pasquier, Antibody diversity in lower vertebrates—Why is it so restricted?, *Nature*, 296:311–313, 1982; M. Feldmann and N. A. Mitchison (eds.), *Immune Regulation*, 1985.

Immunology

The division of biological science concerned with the native or acquired response of complex living organisms to the intrusion of other organisms or foreign substances. The immune system allows the host organism to distinguish between self and nonself and to respond to a target (termed an antigen). An immunological response is specific and can be genetically determined or learned, that is, developed as a consequence of exposure. Examples of immune responses are neutralization of toxic protein molecules, killing of infecting microorganisms, rejection of foreign tissue grafts, and memory by the protective effects of vaccination. The last, discovered by E. Jenner toward the end of the eighteenth century, was the primary immunological finding.

However, it was not until the germ theory of infectious disease was established that the full implication of immunology was realized. First came the recognition that certain bacteria caused corresponding diseases; for example, *Salmonella typhi* caused typhoid fever and not whooping cough. Second came the recognition that it was a specific resistance to that bacterium or its toxins that prevented recurrence of the same disease. Third came the discovery that after recovery from an infectious disease, protective substances called antibodies could be found in the blood of animals and humans that would react with bacteria and their products in the living body and in the test tube. A host of serological reactions were devised to detect the presence of specific antibodies in serum. Antigens, such as bacteria and their products, triggered the production of antibodies and indeed all kinds of chemical and biological molecules.

Immune system. Lymphocytes are the principal cells of the immune system. There are two chief groups of lymphocytes, the thymic or T lymphocytes (T cells) and the B cells. T cells participate in most immunological responses. The thymus matures

helper T cells that assist immune recognition and other (cytotoxic) T cells that are capable of suppressing immune reactions as well as of attacking other cells. T cells recognize antigen by its association with cell membrane–bound receptors. B lymphocytes, after stimulation with antigen molecules, and usually under the influence of factors released by helper T cells, proliferate and later secrete antibody molecules. B cells recognize antigen by means of cell membrane–bound antibodies. The bone marrow has been identified as the ultimate repository and source of a small number (perhaps a thousand) of stem cells that give rise to all the varieties of lymphocytes. *See* CELLULAR IMMUNOLOGY; IMMUNOGLOBULIN.

The mechanism enabling the recognition of self and producing the enormous variety of molecules by which nonself is recognized and countered has been found to reside in the genetic code. A basic code forms the foundation: one set of code corresponds to the immunoglobulins and is expressed in the B cells; another set corresponds to the membrane-fixed receptors on the T cells. *See* IMMUNOGENETICS.

After a foreign antigenic substance, such as a microorganism, protein, or carbohydrate, is recognized as being nonself, it is degraded by phagocytic cells. A portion of the antigen becomes associated in the cell with histocompatibility antigen molecules (antigen processing) and is routed to the cell surface, where it is presented for detection by lymphocytes. The specialized cells with this function are part of the lymphoid reticuloendothelial system and include dendritic cells and macrophages as well as lymphocytes. Lymphoid cells which recognize antigens may multiply in number as well as generate long-lived descendants that provide for memory. The latter allows an accelerated response in the event of subsequent exposure to the same antigen.

A variety of signaling molecules, such as the interleukins, assist the communication between cells. Effector cells which implement some immune reactions or participate in the removal of immunogenic materials include phagocytes, cytotoxic lymphocytes, mast cells, and natural killer cells. The effector mechanisms that kill and remove targeted agents also include a complex of blood proteins known collectively as complement, which has numerous effects of its own, as well as being able to interact with effector cells. *See* COMPLEMENT.

The action of these effector mechanisms, however, has come to be recognized as being not always protective or conferring immunity, but sometimes becoming grossly exaggerated or inappropriate, or capable of turning upon the host in a destructive fashion that causes disease. Exaggerated or inappropriate responses are typically associated with the production of certain antibodies to inhalants, contactants, foods and other normally innocuous products, as well as therapeutic materials or drugs. These responses are classified as allergies. Illnesses associated with a misguided response of the immune system that is directed against the self and results from a breakdown in the normal immunological tolerance of, or unresponsiveness to, self antigens are termed autoimmune, and include systemic lupus erythematosus and rheumatoid arthritis. The mechanisms responsible for these disorders are unknown but probably include the intervention of factors such as viruses that either modify or naturally resemble self molecules. Subsequently, the immune response, in seeking out what is foreign, proceeds to attack the self. Knowledge that would permit control of these aberrant responses would also illuminate the mechanisms involved in the rejection of transplants, thereby enhancing the exploitation of such procedures. *See* ALLERGY; AUTOIMMUNITY.

Related disciplines. Immunology is an eclectic science drawing on many other branches of knowledge. The comprehensive study of infection and resistance may be divided arbitrarily into data relating to the pathogen, the host, and their interactions during infection.

Important information on the host includes its species and the anatomical, physical, genetic, and nutritional factors that may influence infection and resistance. Host-parasite interactions include some classification of disease types, because one microbial species (such as the hemolytic streptococcus, *Streptococcus pyogenes*) may cause multiple diseases (septicemia, boils, scarlet fever, pneumonia, and rheumatic fever), with diverse immunological consequences. Many manifestations of pathology (for example, fever and inflammation) influence the reactivity of both host and parasite. Of prime concern to immunology are the cellular factors (lymphocyte accessory cells) and the humoral factors (antibodies and complement) that are significant aids in host resistance. *See* BIOASSAY; COMPLEMENT; FEVER; PATHOLOGY.

Immunology is also concerned with assaying the immune status of the host through a variety of serological procedures, and in devising methods of increasing host resistance through prophylactic vaccination. There has also been much important investigation of induced resistance and tolerance to transplants of skin and organs, including tumors. *See* ACQUIRED IMMUNOLOGICAL TOLERANCE; BLOOD GROUPS; IMMUNITY; IMMUNOASSAY; ISOANTIGEN; PHAGOCYTOSIS; PRECIPITIN; SEROLOGY; TRANSPLANTATION BIOLOGY; VACCINATION.

Alexander Baumgarten; Margaret J. Polley

Bibliography. D. J. Bibel, *Milestones in Immunology: A Historical Exploration*, 1988; F. H. Garrison, *An Introduction to the History of Medicine*, 4th ed., 1929; W. E. Paul, *Fundamental Immunology*, 4th ed., 1998; I. M. Roitt, *Essential Immunology*, 9th ed., 1997.

Immunonephelometry

An application of nephelometry to the quantification of antigen or antibody. Owing to the antigenic property of many macromolecules, immunonephelometry is capable of very broad application. In addition to its use for the quantification of soluble materials, it has been applied to the assay of

microparticulate substances. The method permits the assay of antigen concentrations as low as 1 microgram per milliliter. *See* ANTIBODY; ANTIGEN; LIGHT-SCATTERING TECHNIQUES.

Method. The technique depends on the light-scattering properties of microparticles. An intense beam of collimated white light or a monochromatic beam (such as a laser) is passed through a solution. Scattering of the light is measured by a suitable detector (such as a photomultiplier tube) placed at an angle to the beam of light. As the number of particles which scatter light increases, so does the intensity of the detected light. A mixture of soluble antigen and antibody initially scatters no light. The initial antigen-antibody complexes are macromolecular in size and do not scatter light. However, such complexes have the ability to aggregate. The size of the aggregates increases with time until they do scatter light. This process takes from several minutes to hours before it attains maximum, depending on various properties of the reactants. If other aspects are kept constant, including, for example, the concentration of one reactant, then both the rate at which the scatter of light increases and its maximal value increase with the concentration of the other reactant. Either the increase in the rate of scatter or its maximal value can be calibrated for various concentrations of a reactant to yield a concentration-response calibration curve. Unknown concentrations of that reactant can then be quantified by measuring the response and comparing it with the calibration curve.

Although the maximal (equilibrium) values are maintained for a time (minutes to hours), they eventually decrease. This happens because, as the size of the aggregates continues to increase, their number and total reflective surface area decrease. This process eventually results in the formation of conventional antigen-antibody precipitates. Accordingly, rate immunonephelometry, which measures the rate of increase in light scatter, may be preferable to a technique which depends on the determination of "equilibrium" or maximal light scatter, because the former method is less sensitive to variation in the absolute duration of the reaction or precise geometry of the container and can generally be performed much more rapidly.

Reactants. The process of aggregation can be markedly accelerated by polyionic substances such as polyethylene glycol, which are therefore commonly added to the solution of reactants. It should be noted that the concentration of the reactants must be adjusted within certain limits, because antigen-antibody complexes are soluble in high relative concentrations of either reactant. The actual limits depend chiefly on the behavior of the antibody, so that some antisera are better than others for immunonephelometric use. The behavior of each antiserum must be individually assessed. The specificity of the immunonephelometric measurement of antigen depends entirely on the monospecificity of the antiserum, that is, its ability to react with only one antigen in a mixture of many.

Inhibition immunonephelometry. A variant of the immunonephelometric technique, inhibition immunonephelometry, has been used for the assay of substances of low molecular weight (haptens). The assay uses a constant amount of antibody and a predetermined quantity of the same reactant as that assayed, chemically coupled to a macromolecular carrier. The mixture yields a high nephelometric value of known magnitude. An introduction of the low molecular substance in the tested sample reduces the known magnitude of the nephelometric value in proportion to the concentration of the substance added in the tested sample. The extent of the reduction is used to determine the concentration of the assayed material in the sample by reference to a calibration curve. *See* IMMUNOASSAY. Alexander Baumgarten

Bibliography. C. L. Cambiaso et al., *J. Immunol. Meth.*, 5:293, 1974; F. W. Putnam (ed.), *The Plasma Proteins: Structure, Function and Genetic Control*, 5 vols., 2d ed., 1987; R. F. Ritchie, in F. W. Putnam (ed.), *The Plasma Proteins*, p. 375, 1975.

Immunopathology

The manifestations of conditions in which normally protective antibody or T cell immune mechanisms are overactive or expressed against self antigens, resulting in damage to cells or organs; also the study of these conditions. *See* ANTIGEN; ANTIBODY; IMMUNOLOGY.

The function of the immune system is to recognize molecular structures that are not normally part of the host's body, and to arrange for their inactivation and removal. Foreign structures may be pathogens; therefore this basic role of immunity is vital to survival. The innate immune system recognizes a few structures called pathogen-associated molecular patterns (PAMP), while the adaptive immune system can respond to billions of molecular configurations. When immunity is activated, the events are intense: cells divide rapidly and release powerful proinflammatory chemokines and cytokines. Such a system needs two types of control: it must avoid responding to normal, "self" structures, and it must regulate the intensity of its reaction so that tissues and organs are not seriously damaged as "innocent bystanders." When the controls work well, the functioning of the immune system is almost unapparent, although it is essential for survival in a world full of pathogens. When the controls fail, we observe immunopathology. *See* IMMUNITY; PATHOGEN; PATHOLOGY.

At least 200 conditions have been described in humans and animals in which the immune system is involved in tissue damage. This array would be bewildering had not P. G. H. Gell and R. R. A. Coombs in 1963 shown that almost all examples of immunopathology (excluding immunodeficiency and tumors of the immune system) could be assigned to one of four categories, based on the underlying immunological mechanism. Although modifications and additions to these categories have been suggested over the years, they remain the

fundamental way in which clinicians look at immunopathology. *See* CELLULAR IMMUNOLOGY; CLINICAL IMMUNOLOGY.

Type I immunopathology. Type I immunopathology is also called allergy or immediate hypersensitivity. The mechanism involves excessive production of the immunoglobulin E (IgE) class of antibody, usually in response to harmless environmental antigens. IgE seems to play its major normal role in resistance to infestations by parasites such as intestinal worms and protozoans such as *Giardia*, *Cryptosporidium*, *Trypanosoma*, and *Schistosoma*. In type I immunopathology the antigen may be a harmless substance entering through the mucous membranes, such as a food or pollen. Why some people react this way to these allergens is unknown, though there is a strong genetic predisposition to it. Allergic people make strong IgE responses, and the antibody adheres to IgE receptors on the surface of mast cells (connective-tissue cells with numerous granules rich in histamine and heparin) in the tissues, especially those of the skin, eyes, nose and throat, and gastrointestinal tract. When further allergen enters the body, the allergen binds to the specific IgE that was generated against it, and crosslinks it on the mast cell surface. This constitutes a signal to the mast cell to rapidly release granules full of preformed histamine, and to begin synthesis of other inflammatory mediators, including prostaglandins and leukotrienes (thus antagonists of these classes of mediators are mainstays of allergy treatment). Mast cells also release substances that are chemotactic for (attract) blood eosinophils (granular leukocytes, or white blood cells). In response to protozoans, for example, these white blood cells are essential, as they contain powerful agents that can destroy them, something the more common neutrophils cannot achieve. In allergy, the eosinophil infiltration may help diagnosis, but if anything it worsens the inflammatory condition. *See* ALLERGY; HYPERSENSITIVITY; IMMUNOGLOBULIN.

Allergic syndromes are categorized on the basis of the organ most affected. Histamine in general causes contraction of smooth muscle in the respiratory and gastrointestinal tracts, and relaxation of smooth muscle in blood vessels, which leads to vessel leakiness. In the skin, release of histamine produces the intensely itchy hive, or wheal-and-flare reaction. In the intestine histamine speeds up peristalsis, which can result in diarrhea and vomiting. Constriction of smooth muscle in the respiratory tract leads to sinus blockage, wheezing, or asthma, as well as the runny eyes and nose that are called (incorrectly, since hay is rarely an allergen) hay fever. If there is sudden systemic histamine release, as might follow ingestion of certain foods or an insect sting, there can be a fall in blood pressure so pronounced as to result in anaphylactic shock and death. Agents that constrict blood vessels and dilate airways, such as epinephrine, can be life-saving. *See* ASTHMA; HISTAMINE.

Allergies are detectable by intradermal skin tests with the suspect agent; a wheal-and-flare that develops within 15 min (hence, it is also called immediate hypersensitivity) helps in making the diagnosis. Lab tests for elevated levels of specific IgE antibody to many known allergens, called RASTs (radioallergosorbent tests), are available. Total blood IgE may be elevated as well. *See* SKIN TEST.

Type II immunopathology. In type II immunopathology, autoimmune antibodies are made that cause damage not to pathogens but to normal body cells. T and B lymphocytes are exposed during their maturation to a wide variety of self molecules. Any lymphocyte whose receptors happen to bind self with high affinity are deleted from the repertoire, since they would be potentially autoimmune. This phenomenon is known as negative selection, and results in us being "tolerant" to our own tissues. In rare individuals there are vital genetic defects in the process and severe autoimmunity results. In most of the more common autoimmune diseases such as rheumatoid arthritis, systemic lupus erythematosus (SLE), and Graves' disease, multiple genes and environmental influences have been implicated. *See* AUTOIMMUNITY.

A case in which the cause is known is rheumatic heart disease. When certain individuals develop an infection with *Streptococcus*, the antibodies made against the bacteria cross-react with a structure expressed on the lining of the heart, especially around the valves. The antibodies bind to the heart, complement is activated, neutrophils are attracted, and damage to the tissues results. All of this represents a normal immune mechanism misdirected against a body structure. In developed countries, children are treated aggressively with antibiotics for streptococcal infections, leading to a steep decline in rheumatic heart disease, but it is still common in the developing world. *See* HEART DISORDERS; RHEUMATIC FEVER.

The expression of autoimmunity may be focal or general. In autoimmune hemolytic anemia the red blood cells throughout the body are the target, and in myasthenia gravis it is the acetylcholine receptor, involved in neural stimulation of muscles. Type 1 diabetics make antibodies that damage the insulin-producing beta cells in the pancreas. An unusual mechanism is seen in Graves' disease, an autoimmune hyperthyroid condition. Here the antibodies are directed against the thyroid-stimulating hormone receptor, but instead of causing cell destruction they act as powerful stimulators at the receptor, driving the thyroid to overproduce its hormones. *See* ANEMIA; DIABETES; MYASTHENIA GRAVIS; THYROID GLAND DISORDERS.

Treatment involves immunosuppressive drugs and powerful anti-inflammatory agents. Side effects are thus common, including increased susceptibility to infection. Sophisticated new drugs and biological response modifiers are coming onto the market with promising results.

Type III immunopathology. Type III immunopathology is also called immune complex disease, and again involves a normal mechanism that becomes dysregulated. Because antibodies have two combining sites for antigen, and antigens may have many antigenic determinants, antibodies can form immune

complexes, three-dimensional lattices of many antigens and antibodies strung together. Large complexes are readily cleared from the blood by the phagocytic cells. Smaller ones may get trapped in the basement membranes, a sort of molecular mat that underlies the cells lining small blood vessels. The trapped complexes activate complement and an inflammatory response ensues. This is most common where there is a net outflow of fluid from blood to tissues: in the joints, pleura, choroid plexus of the brain, skin, and most importantly the kidneys, in which glomerulonephritis (inflammation of the kidney) can develop. *See* ANTIGEN-ANTIBODY REACTION; IMMUNE COMPLEX DISEASE.

The antigens involved may be external to the body or they may be self. The condition was first described in 1905 in children who, in the preantibiotic era, had received large doses of serum from diphtheria-immunized horses. Although the diphtheria was controlled, in about a 7 to 10 days the patients would develop rash, joint pain, fever, and signs of glomerulonephritis. They had made antibodies against the foreign horse serum proteins, and these were appearing in the blood at a time when there were still considerable amounts of the horse proteins circulating. Immune complexes formed and deposited in the joints, skin, and kidneys. This condition is still seen in people who receive foreign sera, such as for the treatment of snake bite, and is called serum sickness.

If a person's immune system mistakenly forms antibody against a protein normally found in the circulation, chronic type III immunopathology may result. Thus people with systemic lupus erythematosus make antibodies to DNA, of which there is always a small amount in the blood from the normal turnover of cells. In rheumatoid arthritis, antibody is actually made against antibody.

Type III immunopathology is treated by decreasing the concentration of antigen if possible, and by the same anti-inflammatory and suppressive regimens used for type II.

Type IV immunopathology. This is the only one of the four types that involves damage to body tissues by T cells. It can be the result of an excessive response or true autoimmunity. One of the classic type IV conditions is due to poison ivy. When a person is first exposed to the resin of the plant, one of its constituents, urushiol, penetrates intact skin and associates with antigen-presenting cells which carry it to the local lymph nodes. Th1 (T helper 1) cells respond, proliferate, and become active. By the time (a week or two) these T cells are circulating around the body, the resin is usually gone from the skin, and nothing is observed, so the victim is unaware that he or she has become "sensitized." However, a subsequent exposure means the T cells are already there and they do what they are designed to: call up a powerful, macrophage-type, inflammatory response. The skin is damaged, although it is an innocent bystander. This condition is called delayed hypersensitivity or contact hypersensitivity, and can be seen with many antigens, including nickel, cosmetics and dyes, and latex. *See* IMMUNOLOGICAL ONTOGENY; POISON IVY.

Other T cell–mediated conditions involve true autoimmunity. For example, in multiple sclerosis, clones of T cells attack antigens derived from the myelin sheaths of central nervous system neurons. What causes this—possibly cross reactions between myelin and an unknown environmental agent—remains unknown.

Many of the autoimmune diseases involve more than one immunopathologic mechanism at the same time: people with rheumatoid arthritis have both antibodies against antibody and T cells against joint collagen, for example, and both antibodies and T cells are implicated in type 1 diabetes. Because the immune system must be able to respond to a near-infinity of antigens, it sometimes becomes confused; and because it is so complex, there is still much research and clinical testing to be done. J. John Cohen

Bibliography. H. Chapel et al. (eds.), *Essentials of Clinical Immunology*, 5th ed., Blackwell, London, 2006; C. Janeway, Jr., et al. (eds.), *Immunobiology*, 6th ed., Garland Science, New York, 2005; D. Male et al. (eds.), *Immunology*, 7th ed., Mosby Elsevier, Philadelphia, 2006.

Immunosuppression

The natural or induced active suppression of the immune response, as contrasted with deficiency or absence of components of the immune system. Like many other complex biological processes, the immune response is controlled by a series of regulatory factors. A variety of suppressor cells play a role in essentially all of the known immunoregulatory mechanisms, such as maintenance of immunological tolerance; limitation of antibody response to antigens of both thymic-dependent and thymic-independent types, as well as to antigens that stimulate reaginic antibody (antibodies involved in allergic reactions); genetic control of the immune response; idiotype suppression; control of contact and delayed hypersensitivity; and antigenic competition. All of the major cell types [macrophages (monocytes), bone marrow–derived lymphocytes that normally mature into antibody-producing cells (B cells), and especially thymus-derived lymphocytes (T cells)] involved in the positive side of cellular interactions required for an immune response have also been found capable of functioning as suppressors in different regulatory systems. *See* IMMUNOLOGICAL DEFICIENCY.

Suppressor cells. Some suppressor functions are antigen- or carrier-specific. (A carrier is a molecule that can be chemically bound to another small molecule, called a hapten, in such a way that the combination induces an immune response that the hapten alone would not induce.) Others may not be carrier-specific, but may be specific for the type of response, such as immunoglobulin production but not delayed hypersensitivity. In the case of immunoglobulin production, the suppressor T cell may regulate the production of all immunoglobulin classes (IgG,

IgM, IgA, and so on), a single class of immunoglobulins, or molecules that bind only a given antigen. Other suppressors may affect only cellular immunity and not humoral immunity. *See* CELLULAR IMMUNOLOGY; IMMUNOGLOBULIN.

Regulation of normal immune response. Suppressor cells are critical in the regulation of the normal immune response. Immunological tolerance refers to the ability of an individual's immune system to distinguish between its own and foreign antigens and to mount a response only to foreign antigens. A major role has been established for suppressor T lymphocytes in this phenomenon, although a complex array of other mechanisms that are not mutually exclusive are also implicated.

Suppressor cells play a role in regulating the magnitude and duration of the specific antibody response to an antigenic challenge. This regulatory role was first demonstrated with antigens that are thymic-dependent, requiring helper T cells as well as B cells for an optimal response. Subsequently it was shown that antigens that are relatively thymic-independent are also regulated in both a positive and negative manner by different sets of T cells.

Immune-response genes determine the capacity to mount an immune response to certain well-defined antigens in a number of species. There is evidence that in some cases suppressor cells play a major part in inducing the nonresponder status, but that they cannot account for it in most situations. Other mechanisms must be active as well.

A mitogen is a substance, usually plant-derived, that induces cells to divide. The induced mitosis is accompanied by changes identifiable under the microscope from a mature-appearing cell to a "blast" or immature cell. Specific mitogens have been found to activate T cells in antigen-nonspecific ways. Concanavalin-A is especially effective in stimulating the suppressor T-cell system that can suppress the generation of lymphocytes which are cytotoxic to allogeneic cells (cells from another animal of the same species), the mixed lymphocyte reaction, mitogen-induced blastic transformation, and antibody and immunoglobulin production.

Reagin, or IgE, is the class of immunoglobulin that mediates allergic reactions such as asthma and urticaria. The reaginic antibody response depends heavily on nonspecific cooperator T cells and specific helper T cells as well as the B cells that produce the antibody. In a negative direction, IgG-blocking antibodies regulate the response, but antigen-specific and antigen-nonspecific suppressor T cells also play a critical role in regulating this response. *See* ALLERGY.

Allotype refers to inherited variations in the genes coding for certain amino acid sequences in the constant region of the immunoglobulin molecules. An animal can be treated in a manner that prevents production of an inherited allotype. In some circumstances, suppressor cells maintain this suppression. Cells capable of producing the allotype can be demonstrated if the suppressors are removed. Idiotype refers to the specific structure of the hypervariable regions that bind antigen, and in a similar manner it has been shown that suppressor cells maintain suppression of idiotype production. *See* IMMUNOGENETICS.

Antigenic competition refers to the observation that immunization with one antigen interferes in a nonspecific manner with the immune response to a subsequently administered second antigen. This phenomenon is short-lived and is mediated by macrophages rather than by T lymphocytes.

Mechanisms of suppressor cell action. T cells are the major cells involved in immunosuppression, although activated phagocytic mononuclear cells are also significant as nonspecific suppressors in many systems. Helper T cells and suppressor T cells are different cell populations that are distinguished to a considerable extent by surface antigens that react with monoclonal antibodies (such as T_4 and T_8) or receptors for specific substances such as histamine. *See* MONOCLONAL ANTIBODIES.

The suppressor T cell emerges from the thymus as an inactive precursor and matures into an effector of suppression after antigen stimulation. The developmental process does not involve antigen stimulation alone, but also requires the interaction of interleukin I and interleukin II with their receptors on T cells. Interleukins I and II are antigen-nonspecific soluble factors produced by monocytes and T cells, respectively, that are involved in the expansion of the numbers of T lymphocytes including suppressor cells as well as helper and cytotoxic T-cell populations.

No single model explains the entire array of cellular suppressor phenomena. In different systems, other T cells, macrophages, or even B cells may be the immediate targets of the suppressor cells and their secretions. Some suppression requires direct cell–cell interaction, whereas other suppression may be mediated by suppressor lymphokines (soluble factors produced by lymphocytes). Both antigen-specific and antigen-nonspecific factors are known, and they may be secreted to act upon other cells, or especially in the case of antigen-specific factors, they may be integral parts of the cell membrane. The soluble immune-response suppressor factor, produced by activated T cells and then activated by monocytes, inhibits B-cell proliferation and immunoglobulin production in response to antigens. Macrophages also secrete suppressor factors, including prostaglandins that act on T cells and other soluble factors that are B-cell–specific.

Disorders of suppressor cell regulation. There is a variety of disorders of immunoglobulin production in humans. In many cases these involve intrinsic defects in the bone marrow stem cells that normally mature into immunoglobulin-producing plasma cells. However, some individuals have been found to have B lymphocytes capable of maturing into cells that secrete immunoglobulins when separated from their own suppressor T cells. The T cells from such individuals have been demonstrated to suppress the production of immunoglobulins by B cells of normal humans. In animal models of immunoglobulin deficiency, the development of suppressor T cells may occur as a secondary event that maintains

suppression of immunoglobulin production even when the primary event is a defect in the B-cell line. A subgroup of individuals with deficiency of only the IgA class of immunoglobulins has suppressor T cells that prevent B cells of normal humans from producing IgA.

Defects in cell-mediated immunity occur in individuals who are infected with various fungal organisms. Suppressor T cells have been implicated, although it is not clear whether the appearance of suppressor cells is the initial event allowing development of the fungal infection or whether they develop secondarily after infection. Those individuals found to have suppressor T cells are at high risk for dissemination of the fungal infection and relapse following therapy.

Although probably only one of many mechanisms, suppressor cells interfere with the host tumor-growth–inhibiting immune response to the foreign tumor-specific transplantation antigens that occur on malignant cells, thus allowing the tumors to progress. In cases of multiple myeloma, there is a significant reduction in polyclonal immunoglobulin synthesis, and serum levels of polyclonal immunoglobulins are low, though there is a high production of the monoclonal immunoglobulin by the malignant plasma cells. Affected individuals are particularly susceptible to infection by pathogenic encapsulated bacteria. One mechanism has been shown to be a blockage of polyclonal lymphocyte maturation by host regulatory macrophages. The acquired immunodeficiency principally affecting cell-mediated immunity in Hodgkin's disease is also caused by macrophage suppressors. Hypogammaglobulinemia associated with benign thymoma has been demonstrated to be associated with circulating suppressor T cells.

Both animal and human studies indicate a major role for both an activation of immunoglobulin-producing B cells as well as the absence or reduced numbers or function of suppressor T cells in autoimmune disorders such as Coombs-positive hemolytic anemia, systemic lupus erythematosus, rheumatoid disorders, and thyroid disorders in which antithyroid antibodies appear in the serum. *See* AUTOIMMUNITY.

Katharine S. Waldmann; Thomas A. Waldmann

Immunosuppressors. Suppression of the immune response may be specific to a particular antigen or may be a response to a wide range of antigens encountered. The whole immune response may be depressed, or a particular population of immunologically active lymphocytes may be selectively affected. In some cases, the effect may be preferentially on T cells rather than B cells. If B cells are affected, it may be on a specific subclass of antibody-producing cells. Antigen-specific immunosuppression may be the result of deletion or suppression of a particular clone of antigen-specific cells, or the result of enhanced regulation of the immune response by antigen-specific suppressor cells. It can also be the result of increased production of anti-idiotypic antibody.

Nonspecific suppression of the immune response occurs in a number of rare immunological deficiency diseases of childhood. These include severe combined immune deficiency syndrome (SCID), in which both T lymphocytes and B lymphocytes are affected; deficiencies of development of the third and fourth branchial arches, affecting the thymus and thus T cells only; or deficiencies of B-cell development, resulting in defects in antibody synthesis. Other conditions are the IgA deficiency associated with ataxia telangiectasia and the variable immunodeficiency states. Acquired deficiency states affecting mainly T-cell function occur in states of malnutrition and in the presence of tumors, particularly those of the lymphoreticular system both in Hodgkin's disease and non–Hodgkin's type lymphoma. Acquired deficiences may also occur secondary to a number of infectious diseases, particularly those of viral origin, but including malaria and typhoid fever. The acquired immune deficiency syndrome (AIDS) is probably of similar origin; its manifestations are similar although more severe and more deramatic. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); IMMUNE COMPLEX DISEASE.

The effect of total body irradiation on the lymphoreticular system also produces an acquired immunodeficiency. Irradiation with x-rays or gamma rays has a direct effect on all rapidly dividing cell populations, though B lymphocytes and T lymphocytes differ in their susceptibility to irradiation. Circulating antibody production may be eliminated by as low as 200–300 roentgens of total body irradiation, but it takes 400–800 R to affect cell-mediated immunity. Depression of T-cell–mediated immunity results in increased susceptibility to opportunistic infections.

Drug-induced immunosuppression. There are a number of compounds capable of suppressing the immune response. The main stimulus for studies designed to identify these substances has been to devise means for controlling organ graft rejection. However, there has also been considerable activity in looking for compounds that will suppress the immune response and reduce the inflammatory process in experimental models of rheumatoid arthritis.

For a number of years it has been known that irradiation by x-rays or gamma rays suppressed the immune response, so it was logical to look first at compounds with a radiomimetic profile. These were mainly alkylating agents and derivatives of nitrogen mustards. Further research uncovered purine and pyrimidine analogs and antimetabolites such as the folic acid antagonists. Another group of agents with immunosuppressive properties is the fungus-derived antibiotics that suppress nucleic acid and protein synthesis in mammalian cells. Corticosteroids, antilymphocyte globulin (ALG), and cyclosporin-A (cyclosporine) complete the list.

In the treatment of recipients of organ graft transplants (kidney, liver, heart), there is an extremely fine balance between the dose which is necessary to inhibit graft rejection alone and that which will suppress the body's resistance to infection particularly with opportunistic organisms such as *Candida albicans, Staphylococcus albus*, *Pneumocystis carinii*, and *Cytomegalovirus*. A reduction in dose will allow

development of a rejection crisis, whereas an increase may precipitate an overwhelming and uncontrollable infection with these and similar organisms. *See* TRANSPLANTATION BIOLOGY.

The ideal immunosuppressive drug should fulfill five main requirements: (1) There should be a wide margin of safety between a toxic and a therapeutic dose. (2) The drug should have a selective effect on lymphoid cells and not cause damage to the rest of the body. (3) If possible, this effect should be only on those cells which are involved in the specific immune process to be suppressed. (4) The drug should need to be administered for only a limited period until the immunological processes become familiar with the foreign antigen and begin to recognize it as part of "self." (5) The drug should be effective against immune processes once they have developed.

Alkylating agents. Not all alkylating agents employed in antitumor chemotherapy have a satisfactory therapeutic index when used as immunosuppressive agents, though they may produce immunosuppression as a side reaction in the treatment of cancer. The most powerful immunosuppressant in this group is cyclophosphamide, which is a phosphoric ester diamide. Cyclophosphamide is not active in the test tube and is broken down in the body in liver microsomes to its active components. These metabolites remain active in the body for no longer than 4 h following injection. As do other alkylating agents, cyclophosphamide produces single-strand breaks in deoxyribonucleic acid (DNA) following alkylation and also forms cross-links binding two strands of DNA. It acts therefore on cells during mitosis, and on more rapidly dividing than slowly dividing cells. Cyclophosphamide has been shown to be most active against the precursors of B cells and T suppressor cells, and thus if given before immunization can sometimes enhance the immune response. If given during the development of the immune response, it will block both antibody production and cell-mediated immunity, as it hits at T cells particularly during the proliferative phase of the immune response. Clinically, cyclophosphamide has been used mainly in the treatment of systemic lupus erythematosus, severe rheumatoid arthritis, and nephritis. It may also have a powerful nonspecific anti-inflammatory effect. The clinical use of this agent is limited by its extremely toxic side effects, which include hair loss and exfoliation of the bladder epithelium.

Thiopurines. The main thiopurines used as immunosuppressive agents are the guanine analogs 6-mercaptopurine and azathioprine. Azathioprine has the advantage of solubility over 6-mercaptopurine and is also considered to be therapeutically more active. Purine analogs compete with free purines for incorporation into nucleic acid and thus inhibit nucleic acid biosynthesis and protein synthesis. Azathioprine has been the standard immunosuppressant used in transplantation surgery since the early 1960s, and is usually supplemented by corticosteroid therapy. It may also be supplemented successfully by antilymphocytic globulin, which has proved particularly valuable in cardiac transplantation. Azathioprine has also been used in the treatment of severe rheumatoid arthritis and systemic lupus erythematosus. *See* NUCLEIC ACID.

Antimetabolites. The most commonly encountered antimetabolite with an immunosuppressive effect is methotrexate, a folic acid inhibitor. This blocks the reduction of dihydrofolinic acid to tetrahydrofolinic acid, which is a coenzyme used in the formation of thymidine from uracil deoxyriboside. Methotrexate is a potent immunosuppressive agent in a number of animal models.

Fungal metabolites. The actinomycins were used temporarily to supplement azathioprine in transplantation surgery. However, the most promising fungal metabolite is cyclosporin-A. Cyclosporin-A has a preferential effect on T lymphocytes and thus on cell-mediated immune processes. It is highly effective in transplantation surgery and can be used alone without azathioprine. It has a much wider margin of safety between the therapeutic and toxic levels and does not have a strong depressive effect on host resistance to infection.

Corticosteroids. There are considerable differences in the susceptibility of lymphoid tissue of various species to the action of corticosteroids. Whereas the hamster, mouse, rat, and rabbit are strongly steroid-sensitive and the action of these compounds is lympholytic, the ferret, monkey, guinea pig, and humans are relatively resistant. Moreover, whereas the action of corticosteroids in the mouse is primarily on lymphocytes, in the guinea pig a major effect is on cells of the mononuclear phagocyte series. The major use of corticosteroids in medicine is as nonspecific and anti-inflammatory agents, and it is likely that they act in this way when used to supplement azathioprine in transplantation surgery. *See* STEROID.

Antilymphocytic globulin. Antilymphocyte serum was first prepared by E. Metchnikoff in 1899, but it was not until the advent of transplantation surgery that it was shown to be such a potent immunosuppressive agent with a preferential action on cell-mediated immunity. Its action is directed mainly against T lymphocytes. The commercial preparation of a purified reagent for use in transplant surgery in humans has presented a number of problems, the most important of which is the difficulty in standardization. Different batches produced over the years have varied considerably in immunosuppressive activity and toxicity. The source of the antigen is critical. Thus thymus and thoracic duct lymphocytes seem to provide the best source. Although horses have been used in a number of commercial preparations, rabbits appear to provide a better serum. A further problem is that there is little correlation between tests of activity outside the body and efficacy in the body. Toxicity of preparations may be due to sensitization to heterologous serum proteins, anti–red cell antibodies, antiplatelet antibodies, and nephrotoxic antibodies (particularly antiglomerular basement membrane), all of which must be removed.

In addition to all these problems in preparation,

purification of the antibody is necessary either by globulin fractionation or by absorbtion to lymphocytes followed by elution at low pH. However, there are a number of preparations available that have proved extremely valuable, especially in cardiac transplants in humans.

Drug-induced immunological tolerance. Specific immunological unresponsiveness can be enhanced by drug therapy, including cyclophosphamide, 6-mercaptopurine, and antilymphocytic globulin. The ease by which unresponsiveness can be induced in adult animals in which the numbers of both T and B cells are reduced confirms the impression from earlier work that this form of tolerance is produced by the action of antigen on immature cells. One might suppose that by reducing the number of available immature cells and most of the immmunologically mature cells, there develops a higher ratio of antigen to immature cells present. In this way the animal is pushed from the zone of immunization to that of high-zone tolerance, not by increasing the level of antigen, but by reducing the number of potentially active cells.

Drug-induced potentiation of immune response. The result of any immune response is a balance between the action of effector cells mediating the phenomenon and suppressor cells regulating the response. Anything that reduces the regulatory function of suppressor cells will functionally increase the immune response. As suppressor cells are derived from rapidly turning-over precursor cells, and effector cells of T-cell–mediated immunity are derived from slowly dividing precursors, it is possible preferentially to depress the action of suppressor cells without affecting effector cells. This may be done by the use of alkylating agents such as cyclophosphamide given before immunization. Cyclophosphamide used in this way can increase a normal cell-mediated immune response, reverse immunological tolerance caused by increased regulatory activity of suppressor cells, and even reverse antigenic competition. It is likely that the chemotherapeutic effect of alkylating agents which are used extensively in the treatment of cancer in humans is partially due to these agents modifying the biological response to the tumor, producing an immunopotentiating action. *See* CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; IMMUNITY; IMMUNOLOGY.

J. L. Turk

Bibliography. A. Allison, K. Lafferty and H. Fliri (eds.), *Immunosuppressive and Antiinflamatory Drugs*, 1993; J. F. Bach, *The Mode of Action of Immunosuppressive Agents*, 1986; G. Möller (ed.), Immunosuppressive agents, *Immunol. Rev.*, vol. 65, 1982; J. L. and D. Parker (eds.), *Drugs and Immune Responsiveness*, 1979; E. I. Reinherz and S. F. Schlossman, The characterization and function of human immunoregulatory T lymphocyte subsets, *Pharmacol. Rev.*, 34:17–22, 1982; R. S. Schwartz, *Progress in Clinical Immunology*, vol. 3, 1977; A. W. Thomson (ed.), *Cyclosporin: Mode of Action and Clinical Application*, 1989; J. L. Turk, *Delayed Hypersensitivity*, 3d ed., 1980.

Immunotherapy

The treatment of cancer by improving the ability of a tumor-bearing individual (the host) to reject the tumor immunologically. There are molecules on the surface of tumor cells, and perhaps in their interior, that are recognized as different from normal structures by the immune system and thus generate an immune response. These molecules, called tumor-associated antigens, have been clearly demonstrated with animal tumors and are strongly suggested with human tumors.

The two components of the immune response are cell-mediated and antibody-mediated immunity, which must work in concert to overcome tumor cells. One type of thymus-derived lymphocyte (also called a cytotoxic T cell) can destroy tumor cells directly, while another recruits other white blood cells, the macrophages, that do the killing. Natural killer cells and perhaps other white blood cells may also participate. Antibodies produced in reaction to a tumor probably work mostly by binding to white blood cells at one end and to tumor cells at the other, bringing the two cells into close proximity and facilitating destruction of the tumor cells. However, elements that normally regulate immunity, such as suppressor T cells, are stimulated excessively by the tumor, which leads to an immune response that is deficient and unable to reject the growing tumor. Thus the strategy of immunotherapy is to stimulate within or transfer to the tumor-bearing individual the appropriate antitumor elements while avoiding further stimulation of suppressor elements. *See* CELLULAR IMMUNOLOGY; IMMUNOLOGIC CYTOTOXICITY; IMMUNOSUPPRESSION.

There are four broad categories of immunotherapy: active, adoptive, passive, and restorative.

Active immunotherapy. This form of treatment attempts to stimulate the host's intrinsic immune response to the tumor, either nonspecifically or specifically.

Nonspecific. Nonspecific active immunotherapy utilizes materials that have no apparent antigenic relationship to the tumor, but have modulatory effects on the immune system, stimulating macrophages, lymphocytes, and natural killer cells. A large and rather heterogeneous group of agents have been studied, the best-known of which are microbial agents, such as bacillus Calmette-Guérin (BCG), and also pyran, interferon, and the vitamin A derivatives called retinoids.

Clinical trials with nonspecific active immunotherapy alone have largely failed to improve the survival of cancer patients, with only a few exceptions, such as in nodular lymphoma and in ovarian cancer. Interferon-alpha is a low-molecular-weight protein produced by leukocytes in response to viral infection. Initial studies with partially purified, extracted interferon-alpha reported several remissions in patients with multiple myeloma and nodular, poorly differentiated lymphocytic lymphoma, and a prolonged survival of patients with osteosarcoma which was treated prophylactically after amputation. Partial

responses were also observed in late-stage breast cancer patients. Trials of recombinant interferon-alpha have noted significant responses of patients with renal cell (kidney) carcinoma, multiple myeloma (a bone marrow tumor), and Kaposi's sarcoma, a tumor often found in patients with acquired immune deficiency syndrome (AIDS). However, it is uncertain whether the nonspecific immunomodulatory effects of interferon were responsible, since interferon also has a direct antitumor effect. Studies on retinoids were prompted by epidemiological data showing a decreased incidence of lung and bladder cancer by the high intake of vitamin A. While retinoids do augment the immune response, like interferon they have direct effects on neoplastic cells. Retinoids also promote the maturation of preneoplastic cells toward normal development. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); VITAMIN A.

Specific. Specific active immunotherapy attempts to stimulate specific antitumor responses with tumor-associated antigens as the immunizing materials. Tumor cells from the same patient or from another with the same type of tumor, irradiated or altered in tissue cultures to make them more antigenic, have been used to try to augment specific immunity. Also, the potential usefulness of antigenic extracts, both soluble and particulate, has been investigated in such dissimilar diseases as malignant melanoma (malignant skin mole) and lung cancer. Rapid progress in identifying important tumor-associated antigens has been made with these and several other tumors, especially through the use of monoclonal antitumor antibodies. The use of extracts of tumor antigens for immunotherapy must be approached with great caution, since soluble tumor antigens can generate suppressor T cells when combined with a small amount of host antibodies. This could lead to enhanced tumor growth. The content of antigens in specific tumor vaccines, the form in which antigens should be presented to the patient (for example, particulate, together with adjuvant materials, and so on), and the schedule and route of administration are all critical questions. *See* MONOCLONAL ANTIBODIES.

Adoptive immunotherapy. This denotes the transfer of immunologically competent white blood cells or their precursors into the host. Bone marrow transplantation, while performed principally for the replacement of hematopoietic stem cells, can also be viewed as adoptive immunotherapy. Some success in the treatment of acute leukemia has been noted by allogeneic (second-party) bone marrow transplantation after sublethal irradiation of the patient. The results from syngeneic (twin) bone marrow transplants have been less encouraging, perhaps because the identical twin's lymphocytes do not recognize the leukemic cells as sufficiently different from self to reject them. It is, however, possible to grow T cells in tissue culture with the aid of T-cell growth factor (interleukin-2), stimulate (or restimulate) them with tumor antigens, and transfuse the cells into the patient. This has been effective therapy in mice. Transfusion of leukocytes armed with antitumor antibodies is another approach that appears feasible for the treatment of human cancers.

Restorative immunotherapy. This approach comprises the direct and indirect restoration of deficient immunological function through any means other than the direct transfer of cells. Repletion of competent effector-cell populations is one form of restorative immunotherapy. Thymic hormones, such as thymopoietin and thymosin, can convert T-cell precursors into helper T cells, a property shared by the antihelminthic agent Levamisole and several of its congeners. In a study of small-cell carcinoma of the lung, thymosin increased the median survival of patients who had sustained a complete response from chemotherapy. The subgroup that benefited most was the one with a deficient complement of T cells. Deficient functioning of T cells may be a problem in most patients with a tumor. Deficient percentages of T cells relative to other leukocytes have been noted in several types of tumor, the most noteworthy of which is Kaposi's sarcoma. Repletion of number or function is an important strategy which might by itself be curative in certain tumors. Antagonism of suppressor influences is another form of restorative therapy. Low doses of the antitumor drug cyclophosphamide selectively inhibit suppressor T cells. Inhibitors of prostaglandin synthesis, such as indomethacin, may be useful in overcoming other suppressor cells of the macrophage series, which act principally through prostaglandins. *See* EICOSANOIDS.

Passive immunotherapy. This means the transfer of antibodies to tumor-bearing recipients. This approach has been made feasible by the development of hybridoma technology, which now permits the production of large quantities of monoclonal antibodies specific for an antigenic determinant on tumor cells. Monoclonal antibodies made against tumor-specific antigens, if they in fact exist, or against tumor-associated antigens, which may be shared by certain normal cells, are of potential importance therapeutically. These antibodies probably work through the arming of white blood cells, especially macrophages, of the host. They may be most effective in treating dispersed tumors such as leukemia, and ovarian cancer in ascites (abnormal accumulation of serous fluid in the abdominal cavity). The potential for enhancement of tumor growth, as a result of the formation of antigen-antibody complexes and the excessive stimulation of suppressor cells, exists and must be carefully monitored. The development of human monoclonals should obviate some of the allergic reactions that have occurred with mouse antibodies, and promises still greater specificity of attack on tumor cells. *See* CANCER (MEDICINE); GENETIC ENGINEERING; IMMUNOLOGY.

June Kan-Mitchell; Malcolm S. Mitchell

Bibliography. P. Calabresi and P. S. Schein (eds.), *Medical Oncology: Basic Principles and Clinical Management of Cancer*, 2d ed., 1993; S. Cryz, *Immunotherapy and Intravenous Immunoglobulins*, 1991; T. J. Hamblin, *Immunotherapy and Disease*, 1990; P. Imbach (ed.), *Immunotherapy with Intravenous Immunoglobulins*, 1991; M. S. Mitchell, The advent of monoclonal antibodies, *Consultant*, 23:127–134, 1983.

Impact

A force which acts only during a short time interval but which is sufficiently large to cause an appreciable change in the momentum of the system on which it acts. The momentum change produced by the impulsive force is described by the momentum-impulse relation. For a discussion of this relation *see* IMPULSE (MECHANICS)

The concept of impulsive force is most useful when the time in which the force acts is so short that the system which it acts on does not move appreciably during this time. Under these conditions the momentum of the system is changed rapidly by a finite amount. The details of the way in which the force varies with time are unimportant, since only the impulse determines the momentum change.

Ordinarily the forces occurring in collisions are impulsive. In the processes in which impulsive forces occur, mechanical energy can be dissipated, and attempts to apply the conservation of mechanical energy may lead to incorrect results. *See* COLLISION (PHYSICS); CONSERVATION OF ENERGY.

A phenomenon known as impulsive loading occurs when materials are subjected to high-speed impacts or explosive charges. The study of failure of materials under impulsive loads is increasing in technological importance. *See* FORCE. Paul W. Schmidt

Bibliography. F. J. Bueche, *Introduction to Physics for Scientists and Engineers*, 4th ed., 1986; D. Halliday, R. Resnick, and K. Krane, *Fundamentals of Physics*, 6th ed., 2000; H. D. Young and R. A. Freedman,*Sears and Zemansky's University Physics*, 10th ed., 2000.

Impedance matching

The use of electric circuits and devices to establish the condition in which the impedance of a load is equal to the internal impedance of the source. This condition of impedance match provides for the maximum transfer of power from the source to the load. In a radio transmitter, for example, it is desired to deliver maximum power from the power amplifier to the antenna. In an audio amplifier, the requirement is to deliver maximum power to the loudspeaker. *See* ELECTRICAL IMPEDANCE.

The maximum power transfer theorem of electric network theory states that at any given frequency the maximum power is transferred from the source to the load when the load impedance is equal to the conjugate of the generator impedance. Thus, if the generator is a resistance, the load must be a resistance equal to the generator resistance for maximum power to be delivered from the generator to the load. When these conditions are satisfied, the power is delivered with 50% efficiency; that is, as much power is dissipated in the internal impedance of the generator as is delivered to the load.

Impedance matching network. In general, the load impedance will not be the proper value for maximum power transfer. A network composed of inductors and capacitors may be inserted between the load and the generator to present to the generator an impedance that is the conjugate of the generator impedance. Since the matching network is composed of elements which, in the ideal case of no resistance in the inductors and perfect capacitors, do not absorb power, all of the power delivered to the matching network is delivered to the load. An example of an L-section matching network is shown in the **illustration**. Matching networks of this type are used in radio-frequency circuits. The values of inductance and capacitance are chosen to satisfy the requirements of the maximum power transfer theorem. The power dissipated in the matching network is a small fraction of that delivered to the load, because the elements used are close approximations of ideal components.

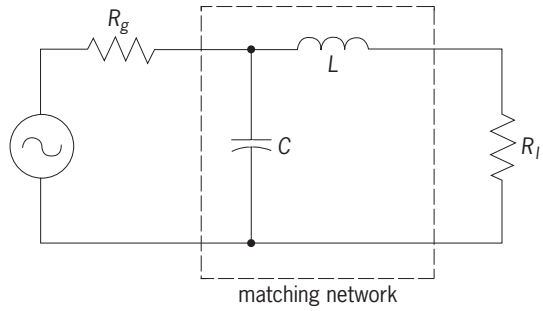

L-section impedance matching network.

Transformers. The impedance measured at the terminals of one winding of an iron-cored transformer is approximately the value of the impedance connected across the other terminals multiplied by the square of the turns ratio. Thus, if the load and generator impedances are resistances, the turns ratio can be chosen to match the load resistance to the generator resistance for maximum power transfer. If the generator and load impedances contain reactances, the transformer cannot be used for matching because it cannot change the load impedance to the conjugate of the generator impedance (the L-section matching network can). The turns ratio can be chosen, however, to deliver maximum power under the given conditions, this maximum being less than the theoretical one.

Iron-cored transformers are used for impedance matching in the audio and supersonic frequency range. The power dissipated in the core increases with frequency because of hysteresis. Above the frequency range at which iron-cored transformers can be used, the air-core transformer or transformers with powdered-iron slugs can be used effectively. However, in these cases the turns-ratio-squared impedance-transforming property is no longer true. Since the transformer is usually part of a tuned circuit, other factors influence the design of the transformer.

The impedance-transforming property of an iron-cored transformer is not always used to give maximum power transfer. For example, in the design of power-amplifier stages in audio amplifiers, the impedance presented to the transistor affects distortion. A study of a given circuit can often show that at a given output power level, usually the maximum

expected, there is a value for the load resistance which will minimize a harmonic component in the harmonic distortion, such as the second or third harmonic. The transformer turns ratio is selected to present this resistance to the transistor. *See* TRANSFORMER.

Emitter follower. In electronic circuitry a signal source of large internal impedance must often be connected to a low-impedance load. If the source were connected directly to the load, attenuation of the signal would result. To reduce this attenuation, an emitter follower is connected between the source and the load. The input impedance of the emitter follower is high, more nearly matching the large source impedance, and the output impedance is low, more nearly matching the low load impedance. If the object were the delivery of maximum power to the load, it might be possible to design the emitter follower to have an output resistance equal to the load resistance, assuming that the load is a resistance. (Special audio amplifiers have been designed to use emitter followers, rather than a transformer, to connect the loudspeaker to the power amplifier.) In many cases, maximum power transfer is not the goal; the emitter follower is introduced primarily to reduce to a minimum the attenuation of the signal.

There exist a number of applications where the emitter follower is not useful as an impedance matching circuit. For example, if a very low-impedance source must be matched to a high-impedance load, then a transistor is used in the common base configuration. *See* EMITTER FOLLOWER; VOLTAGE AMPLIFIER.

Christos C. Halkias

Bibliography. C. A. Schuler, *Electronics: Principles and Applications*, 5th ed., 1998; R. L. Thomas, *A Practical Introduction to Impedance Matching*, 1976; M. E. Van Valkenburg, *Network Analysis*, 3d ed., 1974.

Impsonite

A black, naturally occurring carbonaceous material having specific gravity 1.10–1.25 and fixed carbon 50–85%. Impsonite is infusible and is insoluble in carbon disulfide. A vein of impsonite 10 ft (3 m) wide occurs in La Flore County, Oklahoma, and there are other deposits in Arkansas, Nevada, Michigan, Peru, Argentina, Brazil, and Australia. The Peruvian impsonite may be derived from grahamite by weathering. The origin of impsonite is not well understood, but it appears to be derived from a fluid bitumen that polymerized after it filled the vein in which it is found. Impsonites are known to contain unusually high percentages of vanadium. *See* ASPHALT AND ASPHALTITE; WURTZILITE.

Irving A. Breger

Impulse (mechanics)

The integral of a force over an interval of time. For a force **F**, the impulse **J** over the interval from t_0 to t_1 can be written as Eq. (1). The impulse thus

$$\mathbf{J} = \int_{t_0}^{t_1} \mathbf{F}\, dt \tag{1}$$

represents the product of the time interval and the average force acting during the interval. Impulse is a vector quantity with the units of momentum.

The momentum-impulse relation states that the change in momentum over a given time interval equals the impulse of the resultant force acting during that interval. This relation can be proved by integration of Newton's second law over the time interval from t_0 to t_1. Let **P** represent the momentum at time t, with $\mathbf{P}_0$ and $\mathbf{P}_1$ being the values of **P** at times t_0 and t_1, respectively. Then Eq. (2) holds. If, as is

$$\begin{aligned}\mathbf{J} &= \int_{t_0}^{t_1} \mathbf{F}\, dt = \int_{t_0}^{t_1} (\mathbf{dP}/dt)\, dt \\ &= \int_{\mathbf{P}_0}^{\mathbf{P}_1} \mathbf{dP} = \mathbf{P}_1 - \mathbf{P}_0\end{aligned} \tag{2}$$

ordinarily true, the mass m is constant, the momentum change can be expressed in terms of the velocities $\mathbf{v}_1$ and $\mathbf{v}_0$ at times t_1 and t_0, respectively, giving Eq. (3).

$$\mathbf{J} = m(\mathbf{v}_1 - \mathbf{v}_0) \tag{3}$$

The concept of impulse is ordinarily most useful when the forces are large but act only for a short period. In most of these cases it is necessary to know only the momentum change, which is determined by the impulse. The relation between momentum and impulse thus has the advantage of eliminating the need for a detailed knowledge of how the forces, which can be very complicated, change with time. Forces which occur during collisions are of this type. *See* COLLISION (PHYSICS).

A further simplification of this type of system is obtained if the time interval is short enough for the system to be considered essentially stationary during the time of action of the force, so that the momentum change occurs almost instantaneously. Except for the changed momentum, the motion of the system is treated as it would have been if no impulse had occurred. The effect of the impulse can thus be considered to provide a set of initial conditions for the motion. In a ballistic galvanometer, for example, there is an electric current for only a short time, during which the galvanometer coil, originally at rest, is given an impulse by the force associated with the current. The only effect of the current on the later motion of the coil, which can be considered stationary during the time there is a current, is to provide an initial velocity for the subsequent motion. *See* GALVANOMETER; IMPACT; MOMENTUM.

Paul W. Schmidt

Bibliography. F. Bueche, *Introduction to Physics for Scientists and Engineers*, 4th ed., 1986; D. Halliday, R. Resnick, and K. Krane, *Fundamentals of Physics*, 6th ed., 2000.

Impulse generator

An electrical apparatus which produces very short high-voltage or high-current surges. Such devices can be classified into two types: impulse voltage generators and impulse current generators. High impulse voltages are used to test the strength of electric power equipment against lightning and switching surges. Also, steep-front impulse voltages are sometimes used in nuclear physics experiments. High impulse currents are needed not only for tests on equipment such as lightning arresters and fuses but also for many other technical applications such as lasers, thermonuclear fusion, and plasma devices. *See* FUSE (ELECTRICITY); LASER; LIGHTNING AND SURGE PROTECTION; NUCLEAR FUSION; PARTICLE ACCELERATOR; PLASMA (PHYSICS).

A 1-MV impulse voltage generator. (*Hipotronics, Inc.*)

Impulse voltage generator. An impulse voltage generator (sometimes called a Marx generator, after E. Marx who first proposed it in 1923) consists of capacitors, resistors, and spark gaps. The capacitors are first charged in parallel through charging resistors by a high-voltage, direct-current source and then connected in series and discharged through a test object by a simultaneous spark-over of the spark gaps. The output voltage of such a generator is determined by the number of capacitors and the charging voltage. It is also determined by the efficiency of the generator. The waveform of the impulse voltage can be controlled by suitably adjusting the circuit components of such a generator.

The fundamental parameters of an impulse voltage generator are its rated voltage, energy, waveform T_f/T_t, and polarity. The wavefront time T_f, in microseconds, represents the steepness of the front of the impulse wave. The wave-tail time T_t, defined as the time, in microseconds, required for the tail of the wave to decline to half its peak value, represents the speed of attenuation of the impulse wave. The value of T_f/T_t for a standard lightning impulse wave is 1.2/50, and for a switching impulse wave it is 250/2500. The highest rated voltage of an impulse voltage generator is 10 MV.

The 1-MV impulse voltage generator shown in the **illustration** consists of 10 stages. Each stage has one capacitor. The maximum charging voltage of these capacitors is 100 kV.

Impulse current generator. The impulse current generator comprises many capacitors that are also charged in parallel by a high-voltage, low-current, direct-current source, but it is discharged in parallel through resistances, inductances, and a test object by a spark gap. It provides very high current. Its parameters are rated amperage, waveform T_f/T_t, and polarity. For standard impulse current waves, the values of T_f/T_t, in microseconds, are 4/10 and 8/20. The waveform can be controlled by suitably adjusting the value of the circuit resistances and inductances. The highest rated amperage of an impulse current generator is above several hundred kiloamperes.

Changchang Wang; T. C. Cheng

Bibliography. T. J. Gallagher and A. J. Pearmain, *High Voltage Measurement, Testing and Design*, 1983; H. M. Ryan, *High Voltage Engineering and Testing*, 1994.

Impulse turbine

A turbine in which fluid is deflected without a pressure drop in the blade passages. A turbine is a power-producing machine fitted with shaft-mounted wheels. Turbine blades, attached to the wheels' periphery, are driven by the through-flow of water, steam, or gas (**Fig. 1**). The rotary motion of the wheel is maintained by forces imparted to the blades by the impingement against them of high-speed fluid streams. Before the stream of fluid reaches the moving turbine blades, it is accelerated in stationary

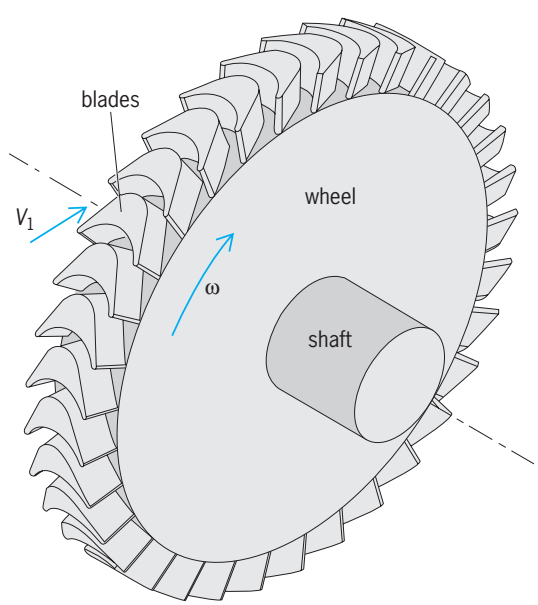

Fig. 1. Impulse turbine wheel. Symbols are explained in text.

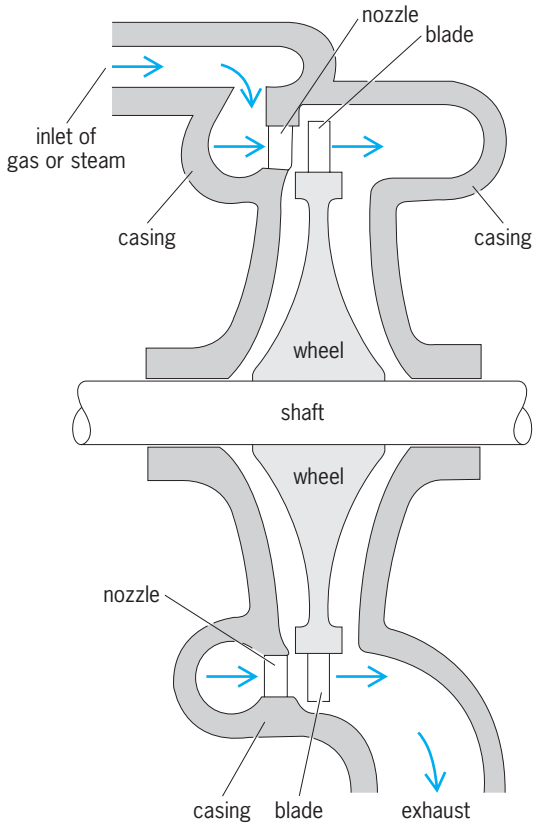

Fig. 2. Longitudinal cross section of a steam or gas impulse turbine.

passages called nozzles (**Fig. 2**). The nozzles are shaped to convert mechanical or thermal energy of the fluid into kinetic energy; that is, the nozzles increase the fluid's velocity while decreasing its pressure and temperature. Upon leaving the nozzles the high-speed fluid strikes the moving blades, and a force is imparted to the blades as the fluid is deflected by them. If the fluid's deflection in the blade passage is accompanied by a pressure drop and a relative velocity rise, the turbine is called a reaction turbine; if the fluid is deflected without a pressure drop in the blade passages, it is called an impulse turbine. *See* NOZZLE; REACTION TURBINE.

The principle of the impulse turbine derives from Newton's laws of motion as applied to the particles of fluid deflected by the blades. From application of the laws of motion to the impulse turbine's blades, it follows that the force of the fluid on the blades is equal and opposite to that of the blades on the fluid. Furthermore, Newton's principle states that the force on the fluid equals the time rate of change of its linear momentum, where linear momentum is defined as the fluid's mass m times its velocity V relative to the moving blade. This principle can also be stated in another way, namely, that the impulse given to the fluid is equal to its change of linear momentum, where impulse is defined as the product of average blade force F by the time t that the fluid resides in the blade passage. The mass m of the parcel will be the mass of fluid entering the blade passage during the time t.

If the fluid approaches the passage between the blades with a relative velocity V_1 and is moving, relative to the blades, in a direction which makes an angle θ with a tangent line to the circumference of the wheel (**Fig. 3**), then the tangential component of the momentum of an entering parcel of fluid of mass m is simply $mV_1 \cos\theta$. Similarly, the tangential component of momentum of the same fluid mass is $-mV_2 \cos\theta$ when it leaves the blade passages. Typically, impulse turbine blades are symmetrical in shape (Figs. 2 and 3), and there is no change in the magnitude of the relative velocity in the blade passage (if frictional effects are neglected); thus, $V_1 = V_2$, and the tangential component of the impulse, Ft, created by the deflection of the fluid in the blade passage is $2mV_1 \cos\theta$.

When an impulse turbine is driven with water, it is called a Pelton wheel (**Fig. 4**). In this case the impulse blade is shaped differently from that used in steam and gas turbines (Figs. 1 and 3), and the water stream impinges tangentially on the blade's center. Nevertheless, Pelton wheel blades are also

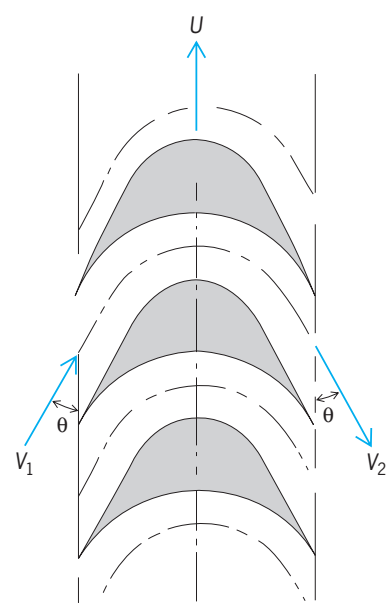

Fig. 3. Cross section through blades mounted at the periphery of the turbine wheel, showing impulse blading. Symbols are explained in text.

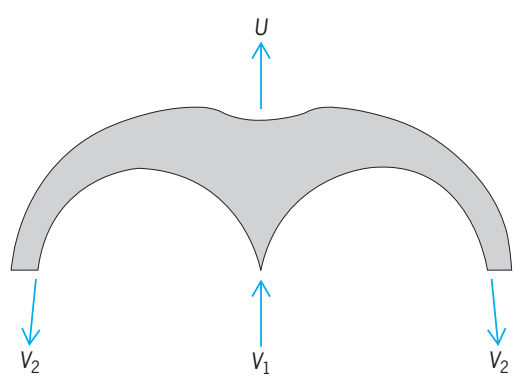

Fig. 4. Hydraulic impulse turbine blade of a Pelton wheel. Symbols are explained in text.

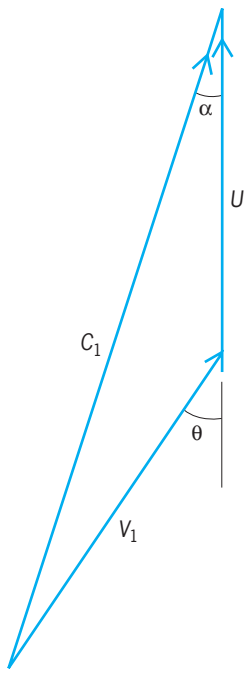

Fig. 5. Velocity diagram, showing relation of blade velocity *U*, absolute velocity C_1 of the stream entering the blade passages, and velocity V_1 of the entering stream relative to the blades.

symmetrical, and the equality, $V_1 = V_2$, applies to them as well.

If all the blades of the turbine wheel are considered together, then the the total blade force F created on all the blades by the total flow is given by $2mV_1 \cos \theta/t$; m/t is the mass rate of flow of fluid through the turbine. Since the velocity V_1 of the stream entering the blade passages is a velocity relative to the moving blade, it is related to the blade velocity U by the equation below. Here, the entering

$$V_1 \cos \theta + U = C_1 \cos \alpha$$

stream has an absolute velocity whose magnitude is C_1 and whose direction makes an angle α with a tangent to the circumference of the blades (**Fig. 5**). Since the stream entering the blade passages has just emerged from the nozzle, C_1 is also the nozzle exit velocity. Thus, $V_1 \cos \theta$ will be zero when the tangential component of the absolute velocity of the stream, $C_1 \cos \alpha$, equals the blade velocity U. The blade velocity is related to the shaft's angular velocity ω and the blade's average radial distance R from the shaft's axis of rotation; thus, the blade velocity is given by $U = \omega R$. When a turbine reaches a certain maximum shaft speed ω_m at which $\omega_m R = C_1 \cos \alpha$, no blade force is created, and the turbine cannot produce any work at all. Since the power P produced by the turbine is the blade force times the blade velocity, that is, $P = FU$, the turbine power is zero at both zero blade velocity ($\omega = 0$) and at zero blade force, that is, where the shaft speed $\omega = \omega_m$. Maximum power and maximum efficiency are realized between these two extremes of rotational speed. In fact, ideal turbine speed occurs when the blade velocity is approximately half the nozzle exit velocity C_1. *See* GAS TURBINE; HYDRAULIC TURBINE; PRIME MOVER; STEAM TURBINE; TURBINE.

Earl Logan, Jr.

Bibliography. H. Cohen, G. F. C. Rogers, and H. I. H. Saravanamuttoo, *Gas Turbine Theory*, 4th ed., 1996; E. Logan (ed.), *Handbook of Turbomachinery*, 1994; F. M. White, *Fluid Mechanics*, 4th ed., 1999; D. G. Wilson and T. Korakianitis, *The Design of High-Efficiency Turbomachinery and Gas Turbines*, 2d ed., 1998.

Inadunata

One of three Paleozoic subclasses of the Crinoidea consisting of over 350 genera ranging from the Ordovician to the Permian. Its members generally lack fixed brachials such that the arms are free above the level of the radials. The calyx may be monocyclic, consisting of a circlet of basal and a circlet of radial plates (**illus.** *a*), or dicyclic, having an additional circlet of infrabasal plates (illus. *b*). The arms, constructed of serially arranged brachials, may be simple or branched. Each arm plate may be bare or give rise to a slender, unbifurcated branchlet called a pinnule. Articulating surfaces of adjacent brachials may be smooth, forming a tight, inflexible suture, or each may possess a central ridge. In the latter case, abutting ridges of adjacent brachials form a mechanical fulcrum allowing a substantial degree of movement and flexibility. The anus may lie flush with the tegmen or be elevated on a tube of variable length and diameter.

The generic diversity of the Inadunata reached its peak during the Late Paleozoic, when it became the dominant crinoid group; all post-Paleozoic crinoids, members of the subclass Articulata, descended from one of the clades of the Inadunata. Cladistic analyses of the Inadunata reveal a lack of shared derived characters (synapomorphies), implying that

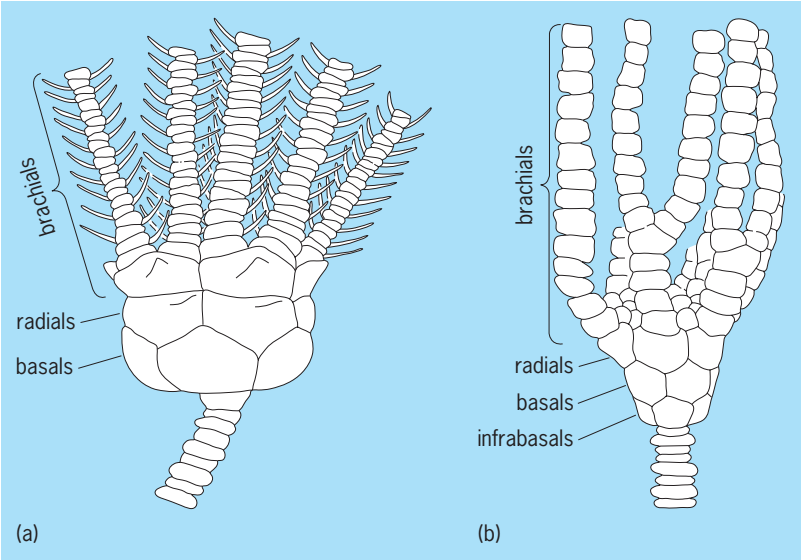

Inadunates. (*a*) Monocyclic. (*b*) Dicyclic.

the taxon represents a polyphyletic grouping of distantly related crinoid clades. *See* CRINOIDEA; ECHINODERMATA. Tomasz K. Baumiller

Bibliography. W. I. Ausich, Phylogeny of Arenig and Caradoc crinoids (phylum Echinodermata) and supergeneric classification of the Crinoidea, *Univ. Kans. Paleontol. Contrib. New Ser.*, 9:1–36, 1998; H. Hess et al. (eds.), *Fossil Crinoids*, Cambridge University Press, Cambridge, 1999; R. C. Moore and C. Teichert (eds.), *Echinodermata 2: Crinoidea*, pt. T of *Treatise on Invertebrate Paleontology*, 1978; C. R. C. Paul and A. B. Smith (eds.), *Echinoderm Phylogeny and Evolutionary Biology*, Clarendon Press, London, 1988; M. Simms and G. D. Sevastopulo, The origin of articulate crinoids, *Palaeontology*, 36:91–109, 1993.

Inarticulata

Formerly, one of the two classes of the phylum Brachiopoda. The phylum comprises solitary, exclusively marine, coelomate [having a true coelom or mesodermally lined body cavity], lophophorate [having a lophophore, a food-gathering ciliated organ], bivalved animals, with the two valves symmetrical about a median longitudinal plane. Now the former Inarticulata comprises two of the three subphyla currently recognized, Linguliformea and Craniiformea, together with a few taxa considered to be basal rhynchonelliform brachiopods. These two subphyla encompass approximately 6% of the generic diversity of the Recent (Holocene, from 10,000 years ago to the present day) and fossil brachiopods. The name Inarticulata refers to the lack of tooth-and-socket articulatory structures on the two valves of these brachiopods. The name was abandoned because of recent changes made in the classification thought to reflect a better understanding of phylogenetic relationships among several groups of Cambrian brachiopods. Because some linguliform and craniform species have primitive forms of articulation that may or may not be homologous with articulated brachiopods, and because some early rhynchonelliforms lack articulation, the term Inarticulata no longer accurately characterizes this group of brachiopods. However, molecular and morphological evidence suggests that Linguliformea and Craniiformea are each likely to be monophyletic (evolved from a single interbreeding population), are sister groups to one another, and may form a clade, together with Phoronida at the base, that is the sister group to Rhynchonelliformea. *See* BRACHIOPODA; PHORONIDA; RHYNCHONELLIFORMEA.

Classification is as follows:

Phylum Brachiopoda
- Subphylum Linguliformea
  - Class Lingulata
    - Order: Lingulida
      - Acrotretida
      - Siphonotretida
  - Class Paterinata
    - Order: Paterinida
- Subphylum Craniiformea
  - Class Craniata
    - Order: Craniida
      - Craniopsida
      - Trimerellida
- Subphylum Rhynchonelliformea

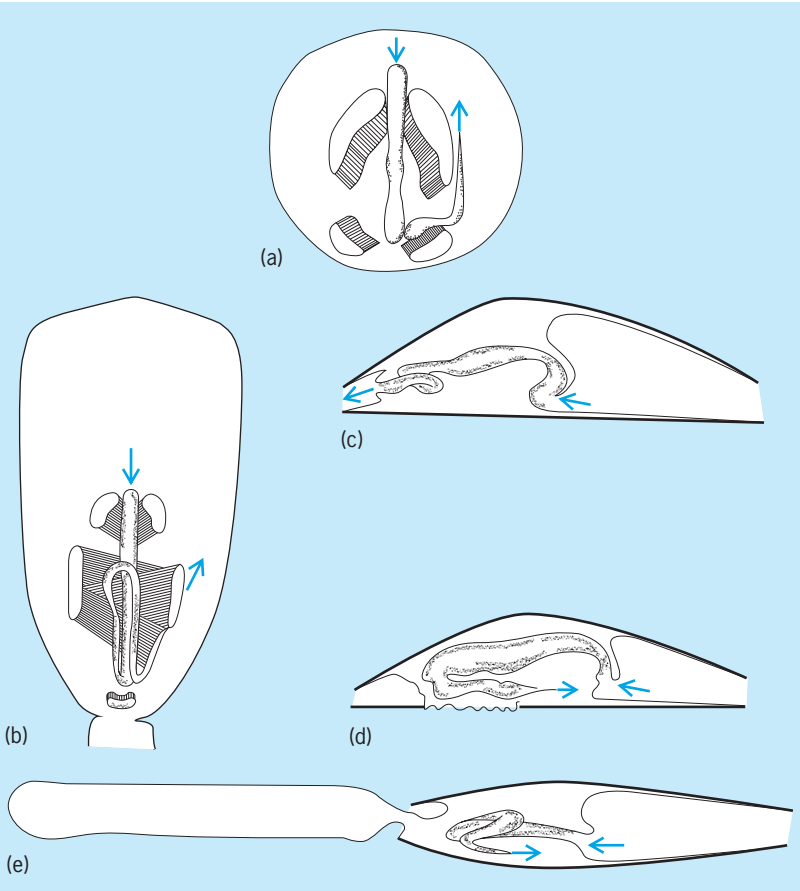

Fig. 1. Position of the gut in relation to the major muscles between the valves; position of mouth and anus and direction of movement of material indicated by arrows. (*a*) *Discinisca*, dorsal view, as if dorsal valve is transparent. (*b*) *Lingula*, dorsal view, as if dorsal valve is transparent. (*c–e*) Comparisons between lateral views of three of the four main types of brachiopod gut arrangement: (*c*) *Novocrania*; (*d*) *Discinisca*; (*e*) *Lingula*. (*After R. L. Kaesler, ed., Treatise on Invertebrate Paleontology, Part H, Revised, vol. 1, Geological Society of America and University of Kansas Press, 1997; adapted from Nielsen, 1991*)

Living linguliforms and craniiforms share a number of morphological features that can distinguish them from rhynchonelliforms. These features include a mantle (ectodermal epithelium) that does not reverse (turn "inside out") in development; discrete, not continuous mantle epithelium at the posterior of the animal; valve articulation that is largely but not exclusively absent; lack of mineralized lophophore supports; and presence of an anal opening (these last two features appear to be primitive characters shared with a common ancestor).

Linguliformea. Linguliformea may (artificially) unite two distinct groups, lingulates and paterinates, and is further characterized by the presence of a pedicle (a tough but flexible cuticle-covered structure at or near the posterior of the animal) that has a coelomic core that develops from the ventral

mantle, not as a larval rudiment as in rhynchonelliforms. An organophosphatic exoskeleton (of two halves) with a stratiform structure with vertical canals is present in all linguliforms.

Lingulata. The lingulates are represented in the modern marine fauna by lingulids and discinids; several genera of each are cosmopolitan (having worldwide range) in their biogeographical distribution. *Lingula* and *Glottidia* have a free-living lifestyle, unusual for brachiopods, and burrow in soft substrata, apparently by contracting dermal muscles and utilizing changes in hydrostatic pressure within the coelom. *Discinisca* lives attached to hard substrata by a pedicle, and possesses a biminerallic (silica and organophosphate) shell, with minute siliceous tablets formed as an early precursor to the organophosphatic shell. Lingulates have planktotrophic (feeding) larvae, which are essentially swimming juveniles with shells, spending months in the water column, and have gonads located in the body cavity, rather than the mantle canals as in rhynchonelliforms. Lingulids can tolerate reduced salinity and can be found today and in the fossil record in marginal marine-to-brackish, shallow-water environments. Lingulate valves can vary from subcircular to elongate oval to subrectangular in outline, and from gently to steeply conical (**Fig. 1***a*, *b*, *d*, *e*; **Fig. 2***b*, *d*, *e*) in lateral view. Their muscle systems (**Fig. 3**) involve several pairs of oblique muscles, in addition to adductor (closing) muscles; they typically do not have diductor (opening) muscles, as rhynchonelliforms do. Acrotretides and siphonotretides are two extinct groups of early Paleozoic linguliforms. Acrotretides have conical ventral valves with large pseudo-interareas (somewhat flattened, posterior shell sectors) [Fig. 2*e*] and an apical pedicle foramen (opening). Siphonotretides have hollow spines on their valve exterior and a large apical pedicle foramen. *See* ACROTRETIDA; LINGULIDA.

Paterinata. Paterinates, the first brachiopods to appear in the fossil record [in the Lower Cambrian (Tommotian)], are known only as fossils from the Cambrian and Ordovician periods. They are

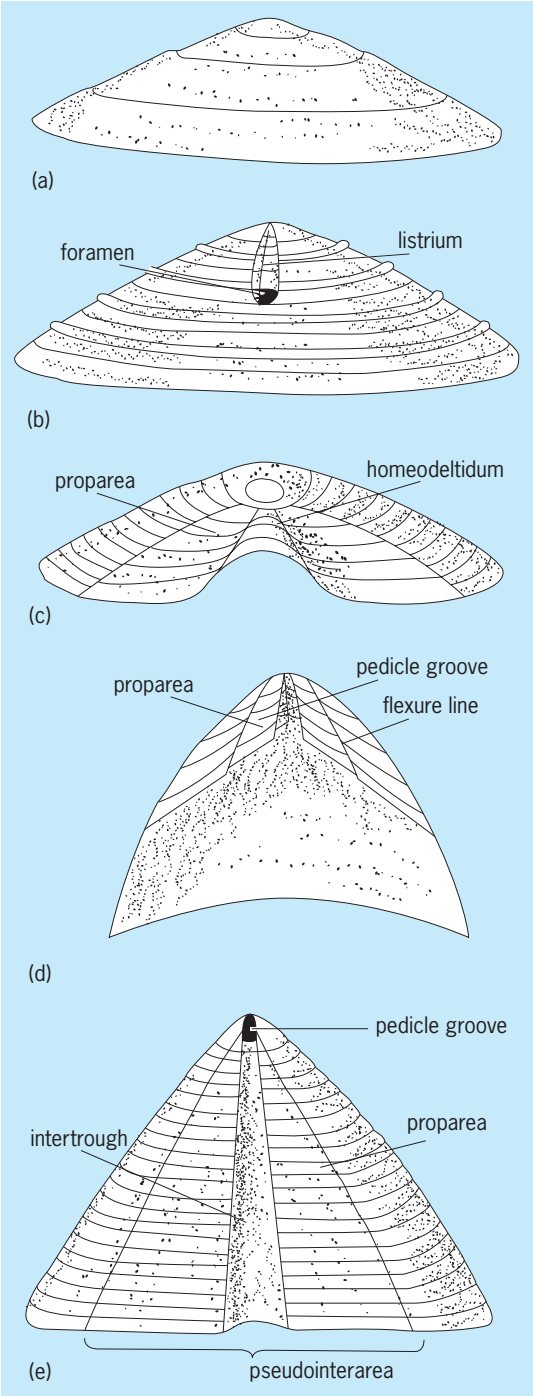

Fig. 2. Modifications of the posterior sector of the ventral valve of some linguliform and craniiform brachiopods. (*a*) Craniiform *Pseudocrania*. (*b*) Discinoid *Orbiculoidea*. (*c*) Paterinate *Paterina*. (*d*) Linguloid *Lingulella*. (*e*) Acrotretide *Prototreta*. (*After R. L. Kaesler, ed., Treatise on Invertebrate Paleontology, Part H, revised, vol. 1, Geological Society of America and University of Kansas Press, 1997*)

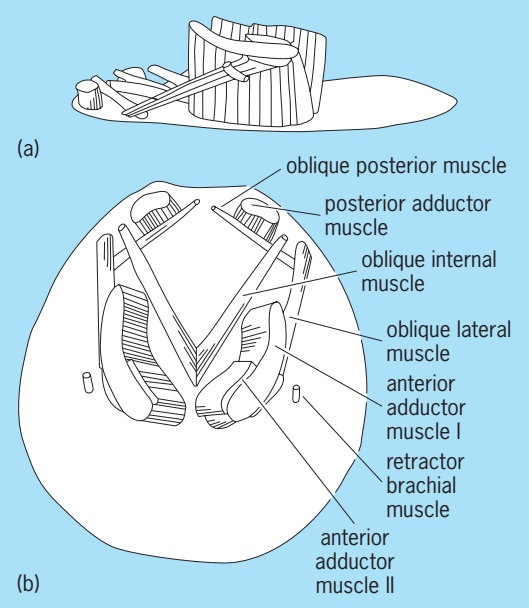

Fig. 3. Muscle system of the discinoid linguliform *Discinisca*. (*a*) Lateral view. (*b*) Dorsal view. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, Part H, Geological Society of America and University of Kansas Press, 1965*)

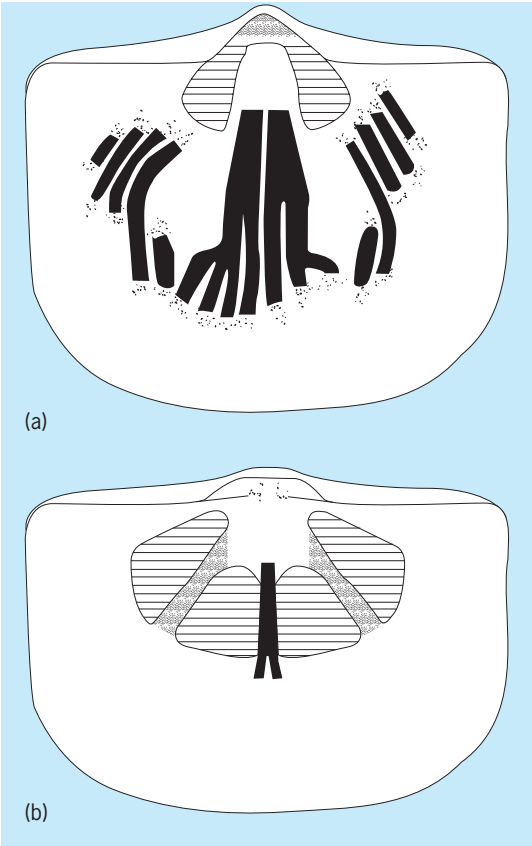

Fig. 4. Interpretation of musculature (muscle scars are hatched) and mantle canal systems (black) of the paterinate brachiopod *Askepasma*. (*a*) Ventral valve interior. (*b*) Dorsal valve interior. (*After R. L. Kaesler, ed., Treatise on Invertebrate Paleontology, Part H, revised, vol. 2, Geological Society of America and University of Kansas Press, 2000*)

morphologically more similar to early rhynchonelliforms than to other linguliforms, differing primarily in the composition of the shell. Their valves have a straight (strophic) posterior margin (Fig. 2*c*; **Fig. 4**) and lack articulation, but appear to have diductor (valve opening) muscle scars (impressions or elevations on valves representing the final sites of attachment of muscles); their valves may well have rotated relative to one another about a hinge axis in life. Paterinates may prove to be more closely related to rhynchonelliforms if organophosphatic mineralogy has evolved more than once in brachiopods. This is certainly a possibility given our incomplete understanding of metazoan mineralization early in the Cambrian. *See* PATERINIDA.

Craniiformea. Craniiformea is characterized by the lack of a pedicle at any stage in development. Recent adults are cemented to hard substrates by their ventral valves. Both valves are typically subcircular in outline and are very slightly conical (Fig. 1*c*, 2*a*) to planar in lateral view. They have lecithotrophic (nonfeeding) rather than planktotrophic larvae, and endopunctate calcareous, not canaliculate phosphatic, valves with a distinctive laminar (tabular) shell structure. All living craniates lack articulation, but many have strophic posterior valve edges. *Novocrania* appears to have a cosmopolitan distribution. Craniopsides, extinct Paleozoic craniiforms, are cemented or free-lying, with impunctate (lacking punctate) valves. Trimerellides, extinct free-lying early Paleozoic craniiforms, have rudimentary articulatory structures in the valves, muscle scars on raised platforms, and unusually large and thick shells, possibly of aragonitic rather than calcitic composition.

Sandra J. Carlson

Bibliography. S. J. Carlson, Recent research on brachiopod evolution, Part H: Brachiopoda, revised, vol. 6: Supplement, in R. L. Kaesler (ed.), *Treatise on Invertebrate Paleontology*, Geological Society of America and University of Kansas, 2007; B. L. Cohen and A. Weydmann, Molecular evidence that phoronids are a subtaxon of brachiopods (Brachiopoda: Phoronata) and that genetic divergence of metazoan phyla began long before the Early Cambrian, *Organisms, Diversity & Evolution*, 5(4):253–273, 2005; B. L. Cohen, A. Gawthrop, and T. Cavalier-Smith, Molecular phylogeny of brachiopods and phoronids based on nuclear-encoded small subunit ribosomal RNA gene sequences, *Phil. Trans. Roy. Soc. London*, Series B, 353(1378):2039–2061, 1998; A. Williams et al., Part H: Brachiopoda, revised, vol. 2: Linguliformea, Craniiformea, and Rhynchonelliformea (part), in R. L. Kaesler (ed.), *Treatise on Invertebrate Paleontology*, Geological Society of America and University of Kansas, 2000; A. Williams and S. J. Carlson, Affinities of brachiopods and trends in their evolution, Part H: Brachiopoda, Revised, vol. 6: Supplement, in R. L. Kaesler (ed.), *Treatise on Invertebrate Paleontology*, Geological Society of America and University of Kansas, 2007; A. Williams and A. J. Rowell, Classification, Part H: Brachiopoda, vol. 1, in R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, Geological Society of America and University of Kansas, H214–H236, 1965.

Incandescence

The emission of visible radiation by a hot body. A theoretically perfect radiator, called a blackbody, will emit radiant energy according to Planck's radiation law at any temperature. Prediction of the visual brightness requires additional consideration of the sensitivity of the eye, and the radiation will be visible only for temperatures of the blackbody which are above some minimum. The relation between brightness and temperature is plotted in the **illustration**. As shown, the minimum temperature for incandescence for the dark-adapted eye is about 730°F (390°C). Under these ideal observing conditions, the incandescence appears as a colorless glow. The dull red light commonly associated with incandescence of objects in a lighted room requires a temperature of about 930°F (500°C). *See* BLACKBODY; HEAT RADIATION; VISION.

Not all sources of light are incandescent. A cold gas under electrical excitation may emit light, as in the so-called neon tube or the low-pressure mercury-vapor lamp. Ultraviolet light from mercury vapor may excite visible light from a cold solid, as in the fluorescent lamp. Luminescence is the term used to refer to the emission of light due to causes other

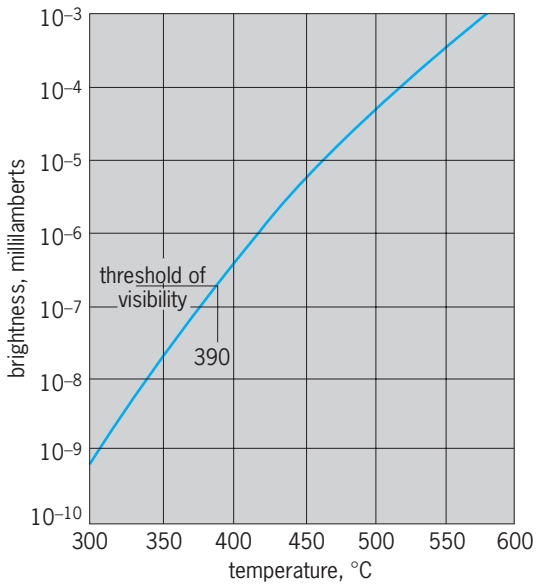

Graph showing the relation between the brightness of a blackbody and temperature. °F = (°C × 1.8) + 32.

Approximate color temperatures of common light sources

Source	Color temperature, °F (K)
Candle	3005 (1925)
Kerosine lamp	3140 (2000)
Common tungsten-filament 100-watt electric light bulb	4600 (2800)
Carbon arc	6700 (4000)
Sun	10,000 (5800)

than high temperature, and includes thermoluminescence, in which emission of previously trapped energy occurs on moderate heating. *See* FLUORESCENT LAMP; MERCURY-VAPOR LAMP; THERMOLUMINESCENCE.

Flames are made luminous by incandescent particles of carbon. Gas flames can be made to produce intense light by the use of a gas mantle of thoria containing a small amount of ceria. This mantle is a good emitter of visible light, but a poor emitter of infrared radiation. As less heat is lost in the long waves, the mantle operates at a higher temperature than a blackbody would and, hence, produces more intense visible light.

A useful criterion of an incandescent source is its color temperature, the temperature at which a blackbody has the same color, although not necessarily the same brightness. The color temperature of common light sources depends upon operating conditions. Approximate values are given in the accompanying **table**. *See* INCANDESCENT LAMP.

H. W. Russell; George R. Harrison

Bibliography. F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976; J. R. Meyer-Arendt, *Introduction to Classical and Modern Optics*, 4th ed., 1994; Optical Society of America, *Handbook of Optics*, 2d ed., 1994.

Incandescent lamp

A lamp that creates radiant energy when its metallic filament is heated by an electric current. The filament is designed to produce radiant energy in the visible portion of the electromagnetic spectrum (light). The filament is of a special material that is supported in an envelope (bulb) that has been evacuated or filled with an inert gas such as argon, nitrogen, or krypton. In addition to light, the heated filament emits infrared and ultraviolet energy. When either of these radiations is accentuated, the lamp may be used as a source of that energy but with a reduction in luminous efficacy. The luminous efficacy (formerly called light-source efficiency) is the ratio of the units of light produced (lumens) to the power (watts) required, and is expressed in units of lumens per watt (lm/W). *See* LIGHT; LUMINOUS EFFICACY.

Advantages and disadvantages. The advantages of incandescent lamps include the following: (1) Lamps are low in cost compared to other illumination devices. (2) Lamps are simple to use and to connect into the electrical system. (3) Lamps need no auxiliary, such as a ballast. (4) The light source may be kept very compact, making the light easily controlled by reflectors and lenses. (5) Many shapes and sizes are available for special uses. (6) Color rendition of skin colors is pleasing to the observer. (7) Lamps can be designed to operate at low voltage, that is, 1.5 V, and are therefore readily used with battery sources of power (flashlights and autos). (8) Inexpensive dimmer control of light output is possible.

The disadvantages include the following: (1) Efficacy is low compared to fluorescent or high-intensity-discharge lamps. (2) Lamps are very inefficient when colored light at the blue end of the spectrum is required. (3) Lamps have short life compared to fluorescent and high-intensity-discharge lamps. (4) A large amount of heat is produced.

Lamp construction. The important parts of an incandescent lamp are the bulb (envelope), the filament, and the base (**Fig. 1**).

The bulb may be clear, colored, inside-frosted, or coated with diffusing or reflecting material. Most lamps have soft-glass bulbs; hard glass is used when

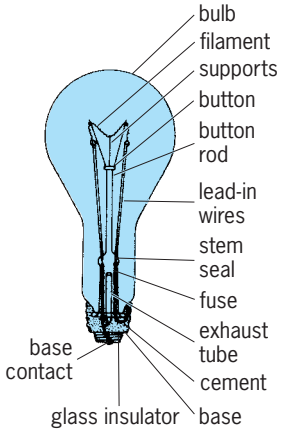

Fig. 1. Parts of an incandescent lamp. (*General Electric Co.*)

the lamp will be subjected to sudden and severe temperature changes. Lamps have a variety of bulb shapes, base types, and filament structures (**Fig. 2**). These vary according to the type of service planned, the need for easy replacement, and other environmental and service conditions.

The efficient design of an incandescent lamp centers about obtaining a high temperature at the filament without the loss of heat or disintegration of the filament. The early selection of carbon, which has the highest melting point of any element (3872 K or 6510°F) was a natural one. (It is standard practice to express filament temperatures in kelvins. The Kelvin temperature scale has degrees identical in size to the Celsius degree; however, the scale, instead of starting at the melting point of ice, as does the Celsius scale, starts at absolute zero, which occurs at −459.67°F or −273.15°C.) Carbon evaporates from its solid phase (sublimates) below this temperature, so carbon filaments must be operated at relatively low temperatures to obtain reasonable life. With the advent of ductile tungsten (1911) a nearly perfect filament material was discovered. Ductile tungsten has a tensile strength four times that of steel, its melting point is high (3655 K or 6120°F), and it has relatively low evaporation. Hot tungsten is an efficient light radiator; it has a continuous spectrum closely following that of a blackbody radiator with a relatively high portion of the radiation in the visible spectrum. Because of its strength, ductility, and workability, it may be formed into coils, these coils again recoiled (for coil-coil filaments), and these again recoiled for cathodes in fluorescent lamps. If tungsten could be held at its melting point, 52 lm/W would be radiated. Because of physical limitations, however, 22 lm/W is the highest practical radiation for general-service lamps; some special lamps can reach as high as 35.8 lm/W. In a given incandescent lamp, a higher filament temperature will result in an increased efficacy and color temperature (whiter light) and a reduction in lamp life.

Vaporization of the filament is reduced as much as possible. A small amount remains, however, and causes blackening of the bulb. The evaporated tungsten particles are carried to the upper part of the bulb by convection currents. In a base-down position the blackening reduces the output a few percent. To reduce blackening, the inner atmosphere of the lamp is maintained as clean as possible by use of a getter, which combines chemically with the tungsten particles. In some lamps a grid is placed to attract and hold the evaporated tungsten particles.

The electric power supplied to the filament of a general service lamp is converted to radiant energy; however, only 4–6% of the energy is light. The remaining energy is heat radiation, except for a very minor part which is in the form of ultraviolet radiation. This better than 90% conversion to heat results in portions of the lamp operating at rather high temperatures. The 100-W general service lamp may have glass temperatures of almost 450°F (230°C) operated base-down and base temperatures over 225°F (108°C) operated base-up. Placing the lamp in some types of luminaires (light fixtures) or other devices that tend to hold the heat in may produce lamp temperatures well above these. Excessive heat may cause the cement holding the base to the glass to fail. This failure can be prevented by selecting lamps with skirted or bipost bases designed for use at high ambient temperatures (Fig. 2). Lamp filaments that require large electric currents, such as high-wattage 120-V lamps or moderate-wattage 6-V lamps, are designed with bipost or other bases that can be solidly connected to the electric circuit.

Slide and movie projector lamps are examples of lamps required to operate at high temperatures. A common failure of such lamps results from improper adjustment of the projector optics. The filament image is focused on the bulb wall instead of the film and the glass softens and blisters. Another common failure results when a lamp is exposed to a spray or mist of cold water, causing thermal cracks in the glass. To protect the electric circuit from some lamp failures, a fuse is built into the lead wire that connects the base to the filament.

Lamp ratings. Most lamps are rated in watts at a specified voltage. The most common voltage is 120 V for lamps to be used for general lighting service. General-service lamps are also available with voltage ratings of 125, 130, 230, 250, and 277. Lamps for special lighting service where the voltage may not be relatively constant may have ratings such as 115–125 V. Typical of such lamps are the rough-service lamps for use on extension cords where supplementary lighting is needed. Lamps for various special uses have voltage ratings which range from 1.5 V for flashlight lamps, 6 V for projector lamps, 12 V for automotive lamps, to 300 V for mines and special industrial usage.

Decorative lamps used in strings for Christmas trees and similar uses are available in two forms. The first uses 7–15 W 120-V lamps connected so the lamps operate in parallel from a 120-V source of power. The second form uses 35–50 miniature lamps connected in series to a 120-V source. The lamps are rated at 2.5–3.5 V and normally have currents in the range 0.04–0.06 A. A device which shorts the lamp when a filament burns out keeps the entire string from going out when one or two lamps burn out, but if three or more lamps are burned out the entire string is extinguished. The power for an entire string of these miniature lamps is only 5–7 W.

Lamp characteristics. Among the more important lamp characteristics are life, lumen (light) output, efficacy or lumens per watt of input power, color, color temperature, lumen depreciation, and the effect of variations in the applied voltage on the above characteristics of lamps.

Life and efficacy. The life of an incandescent lamp may range from a low 2–6 h for a photoflood lamp, where life has been sacrificed to obtain high efficacy and high color temperature, to small long-life lamps designed for exit fixtures. The latter lamps have low efficacy but a very long 10,000–100,000-h life to assure continuity of service. General-service

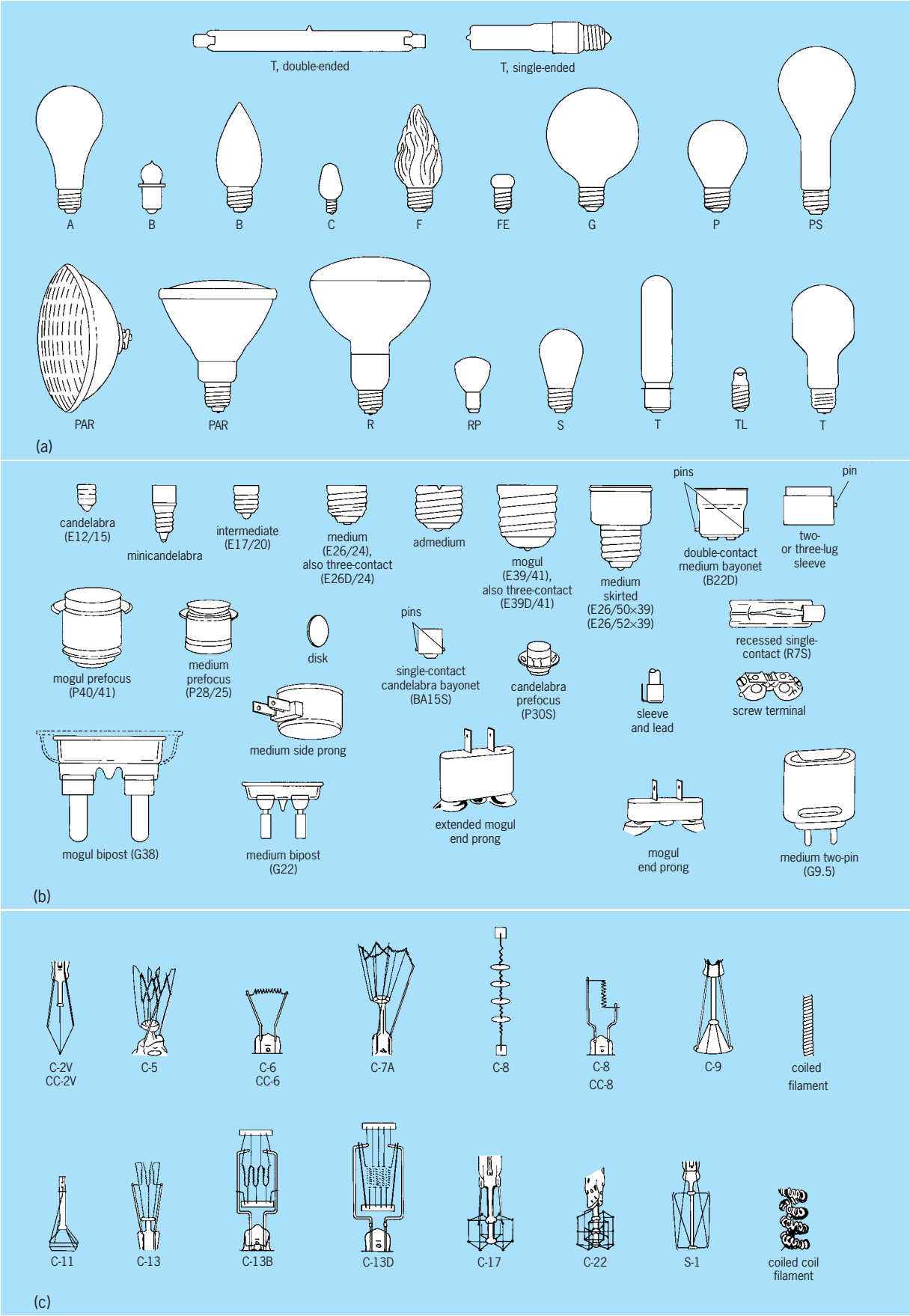

Fig. 2. Variations of incandescent lamps. Drawings are not to scale. (*a*) Bulb shapes and designations. (*b*) Lamp bases. Designations of International Electrotechnical Commission are shown, where available. (*c*) Filament structures. (*After Illuminating Engineering Society*, *IES Lighting Handbook*, *7th ed.*, *1984*)

lamps have a normal life of 750 h, while extended-service and long-life lamps offer 1500–5000 h of life by reducing the efficacy and light output to increase the life (**Table 1**). Users of the longer-life lamps save by buying fewer lamps, but they get much less light and still pay the same amount for electric energy.

The life of lamps is always given on an average basis. Thus, of ten 750-h lamps, two would probably be burned out at 600 h, or 80% life, five after 750 h, or 100% life, and eight after 900 h, or 120% life. The foregoing assumes the lamps operate at rated voltage. The lamp life varies rapidly with voltage. Because utilities cannot supply a constant voltage to all parts of their systems and because there is a variance in wiring from one house to the next, the average life of lamps in homes can vary from the rated value.

Vacuum and gas-filled lamps. Two major classes of lamps are vacuum and gas-filled. Vacuum lamps have had most of the air pumped out of the bulb, and the bulb is sealed. A "getter" chemical in the bulb is then heated inductively, causing it to combine with most of any gas remaining, producing a very high vacuum. Most of the lamps rated 10 W or less are vacuum lamps. Vacuum lamps operate at a lower filament temperature and have a lower efficacy than gas-filled lamps.

Gas-filled lamps have most of the air removed, and the bulb is then filled with an inert gas that acts to slow down the rate at which tungsten molecules evaporate from the filament, enabling the filament temperature to be increased without decreasing the life. The efficacy of the lamp is increased, and it is further increased by coiling the filament in a helix. The effect is enhanced by using heavier gases. The fill gas in standard gas-filled lamps is a mixture of 50–99+% argon (atomic weight 39.91) and nitrogen (at. wt 14.008). The use of the gas krypton (at. wt 82.9) as a fill gas produces an improvement of 7–20% in light output. Krypton lamps are sold at a premium price as energy-saving lamps. Lamps above 75 W are normally gas-filled, while 10–75-W lamps are either vacuum or gas-filled, depending on their design and intended usage.

Color and glare. Vacuum lamps have a higher red radiation and a lower blue radiation than gas-filled lamps (**Fig. 3**). This means that the vacuum lamps appear more red and enhance red objects and darken blue objects more than the same objects when lighted by gas-filled lamps. The daylight lamp has a blue glass bulb which yields light of a color more nearly approaching pure white theoretical emission.

TABLE 1. Comparison of life versus light output of 100-W lamps*

Lamp type	Life	Initial lumens	% light output
General-service lamp, inside frost	750 h	1740	100
Extended-service lamp, white	2500 h	1460	84
Long-life lamp, inside frost	5000 h	1050	60

*Energy consumption of these three lamps is identical.

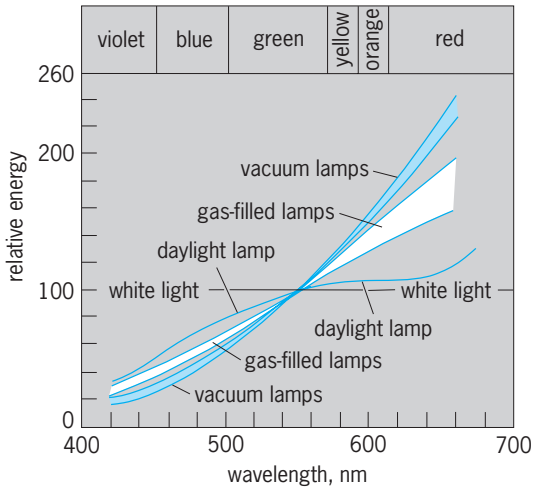

Fig. 3. Spectral energy distribution for important types of incandescent lamps.

The color of the light emitted and the glare produced by the bright filament can be controlled by adding a surface treatment or coat to the clear glass bulb. Lamps where a normal filament color is desired, but where glare reduction is needed, are etched or frosted on the inside surface of the glass. The lumen loss is about 0.5% maximum compared to a clear-glass lamp. Where still greater filament hiding is desired, the inside of the glass is coated with a white ceramic powder with a resultant light loss of about 2%. The improved seeing due to a glare reduction far outweighs the light loss, and these are the recommended lamps to use if the lamp is within the field of view.

When light of a particular color such as red, blue, green, or pink is required, the bulbs may be made of colored glass or coated with plastic or ceramic. Some lamps with reflectors built in, such as the PAR (parabolic-aluminized-reflector) lamp (Fig. 2*a*), have dichroic filters for the reflector, so that most of the light passes out the front side of the lamp and the heat, or infrared radiation, passes out the back of the lamp, thus helping to keep the illuminated objects cool. These lamps are often used to illuminate meat cases in stores.

Lumen depreciation. The lumen output of incandescent lamps decreases with age, a characteristic referred to as lumen depreciation. Lumen values stated for lamps are always initial lumens. The rate of depreciation of a lamp is determined by its design and the manner in which it is operated (that is, base-up or base-down, and normal or abnormal voltage). At rated life, most general service lamps emit 75–85% of their initial lumen output.

Effects of voltage variation. Lamps operated at other than rated voltage have their characteristics altered (**Table 2** and **Fig. 4**). The lamp life varies rapidly with voltage, as mentioned above (Fig. 4).

Applications and special types. Incandescent lamps have been developed for many services. Most

common are those used in general service and the miniature lamp. Special types have been developed for rough service applications, bake-oven use, severe vibration applications, showcase lamps, multiple lights (three-way lamp), sign lamps, spotlights, floodlights, and insect-control lamps.

Reflector-type lamps are designed with built-in light control, that is, silvering placed on a portion of the lamp's outside surface. The reflecting surface is protected and is effective for the life of the lamp.

Projector-type lamps have molded-glass reflectors, silvered inside the lamp cavity, with either a clear glass cover or a molded control-lens cover. The reflector and the cover are sealed together, forming a lamp with an internal reflector. The parts are of hard glass, and the lamps may be used for outdoor service in floodlight and spotlight installations. This is also the type of lamp used in the sealed-beam headlight for automobiles, locomotives, and airplanes. The contours of the reflector are molded for accurate beam control. With this sturdy structure the filament can be positioned for the best use of the lens, and the lamp has little depreciation during its life. Being constructed of hard glass, it lends itself to high-wattage use.

Miniature incandescent lamps have diverse applications, from the ordinary flashlight to the "grain of wheat" lamps used in surgical and dental instruments. These lamps are designed to give the highest efficiency consistent with the nature of the power source employed.

Special picture-projection lamps are designed for accurate filament location in the focal plane of the optical system, with the filament concentrated as much as possible in a single plane and in a small area. These precision lamps use a prefocus base for accurate positioning of the filament with respect to the base. Projector lamps run at high temperatures, and forced ventilation is frequently required. *See* OPTICAL PROJECTION SYSTEMS.

A special class of lamp is designed for the photographic field, where the chief requirement is actinic quality. Frequently the most important rating is the color temperature, with little regard for economic efficency or life. Photoflood lamps give high illumination for a short life, obtaining twice the lumens of general-service lamps from high filament temperature, and high color temperature with three times the photographic effectiveness. The "daylight blue" photoflood lamp gives a very white light at 4800 K color temperature and 35.8 lm/W.

TABLE 2. Typical performance of large incandescent lamps burned below rated voltage

	% normal operation		
Voltage	Light output	Watts	% initial efficacy
100.0	100.0	100.0	100.0
99.2	97.3	98.8	98.5
98.3	94.4	97.4	96.9
97.5	91.8	96.1	95.5
96.7	89.2	95.0	93.9
95.8	86.4	93.6	92.3
95.0	84.0	92.4	90.9
94.2	81.5	91.2	89.4
93.3	78.0	89.8	86.8
92.5	76.6	88.7	86.3
91.7	74.1	87.5	84.7
90.0	69.5	85.0	81.8
88.3	65.0	82.5	78.8
86.7	60.8	80.3	75.7
85.0	56.6	77.9	72.7
83.3	52.0	75.5	68.8

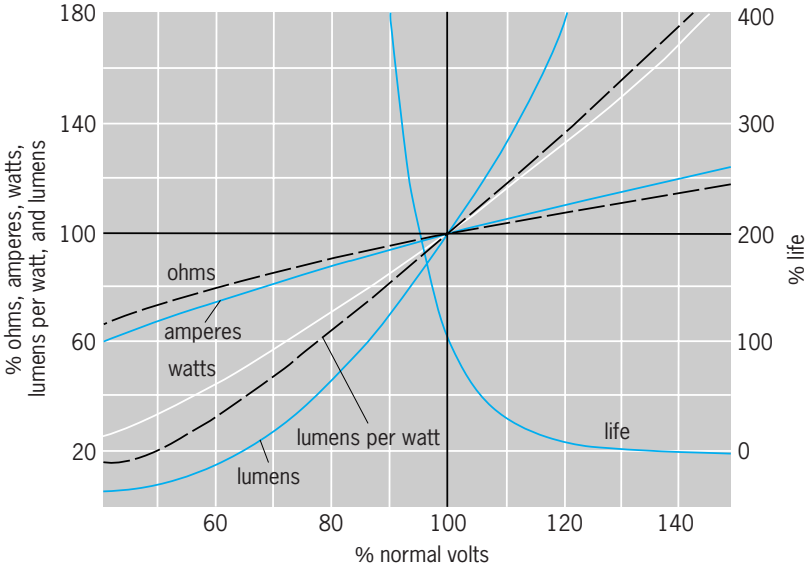

Fig. 4. Characteristic curves of general-service lamps. (*After Illuminating Engineering Society, IES Lighting Handbook, 8th ed., 1993*)

Tungsten-halogen lamps. These lamps are made with a fill gas that includes a small amount of one of the halogen elements such as iodine, bromine, or chlorine. The special changes that result from the halogen addition are: (1) the filament temperature can be increased, giving a whiter light output; (2) the depreciation in light output with time is greatly decreased; and (3) the lumen output and the life are increased.

Operation. The filament is enclosed in a small-diameter tubing made of fused quartz instead of glass to withstand the 500°F or 260°C bulb wall temperature required for proper functioning of the halogen gas fill. In operation, as the tungsten evaporates from the filament, it combines with the halogen forming a tungsten-halogen gas. This gas circulates within the lamp, but instead of blackening the bulb wall with tungsten deposits, as happens in regular lamps, the tungsten remains as a gas, until coming in contact with the very high-temperature filament it separates into halogen and tungsten with the tungsten being deposited back on the filament. Dimming the lamp impedes this process and if continued for a long period will result in the blackening of the bulb wall. Occasional operation of the lamp at rated voltage and current will alleviate this problem. This lamp develops a larger amount of ultraviolet radiation than general-service lamps. This may cause problems when lighting objects sensitive to ultraviolet, so precautions may be required.

Configurations. Tungsten-halogen lamps are available in four basic configurations, as replacement lamps, mirrored-reflector (MR) lamps, encapsulated lamps, and tubular lamps.

Replacement lamps are available in sizes of 50 W and less and are used where reflectors are permanently installed. Mirrored-reflector lamps are available in several sizes, with the 2-in.-diameter (5-cm) lamp (MR-16) being the most widely used. They are available in a few different wattages, usually at 12 V, and with a variety of distribution characteristics. They are an excellent choice for illuminating small targets, such as jewelry, to a high illuminance.

Encapsulated lamps, may be used as a substitute for conventional incandescent lamps of the same shape. The tungsten-halogen source is installed (**Fig. 5**) in a conventional lamp shape such as PAR-38,

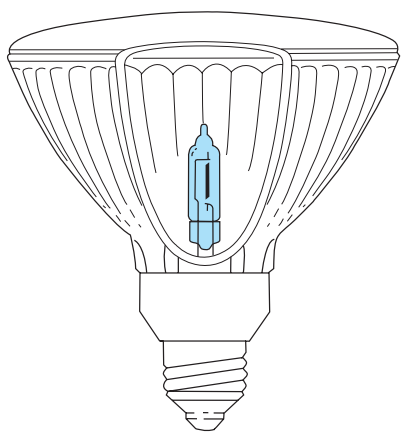

Fig. 5. Encapsulated tungsten-halogen lamp with a section cut away, showing installation of the tungsten-halogen source.

a pressed reflector lamp, 4.75 (38/8) in. (12 cm) in diameter. When replacing a conventional incandescent, a smaller-wattage tungsten-halogen will usually provide equal illuminance.

Tubular lamps are available in a double-ended configuration or with a candelabra base. These lamps are available in sizes from 50 to over 500 W. They are commonly used in indirect floor-standing luminaires, wall scones, and other applications where a compact source size is desirable. When replacing the lamps, touching the tube with the fingers should be avoided, as this may truncate the lamp life.

For other types of incandescent lamps *see* ARC LAMP; INFRARED LAMP. G. R. Peirce; Robert Leroy Smith

Bibliography. Illuminating Engineering Society, *IES Lighting Handbook*, 8th ed., 1993.

Inclined plane

A plane surface inclined at an angle with the line of action of the force that is to be exerted or overcome. The usual application is diagrammed in the **illustration**. A relatively small force acting parallel to the surface holds an object in place, or moves it at a constant speed.

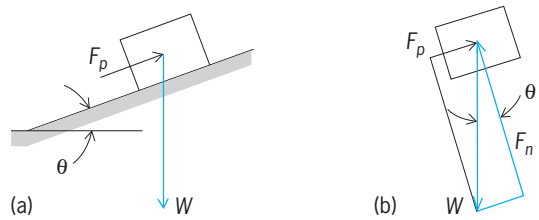

Weight resting on an inclined plane (*a*) with principal forces applied, and (*b*) their resolution into normal force.

In the free-body diagram shown here, three forces act on the object when no friction is present. The forces are its weight W, the force F_p parallel to the surface, and a force F_n normal to the surface. The summation of the forces acting in any direction on a body in static equilibrium equals zero; therefore, the summation of forces parallel to and forces normal to the surface are given by Eqs. (1) and (2). A force

$$F_p - W \sin\theta = 0 \qquad (1)$$

$$F_n - W \cos\theta = 0 \qquad (2)$$

slightly greater than $W \sin\theta$ moves the object up the incline, but the inclined plane supports the greater part of the weight of the object. The principal use of the inclined plane is as ramps for moving goods from one level to another. Wheels may be added to the object to be raised to decrease friction but the principle remains the same. The wedge and screw are closely related to the inclined plane and find wide application. *See* SIMPLE MACHINE; WEDGE. Richard M. Phelan

Incompressible flow

Fluid motion with negligible changes in density. No fluid is truly incompressible, since even liquids can have their density increased through application of sufficient pressure. But density changes in a flow will be negligible if the Mach number, Ma, of the flow is small. This condition for incompressible flow is given by Eq. (1), where V is the fluid velocity and a is the

$$\mathrm{Ma} = \frac{V}{a} < 0.3 \qquad (1)$$

speed of sound of the fluid. It is nearly impossible to attain Ma = 0.3 in liquid flow because of the very high pressures required. Thus liquid flow is incompressible. *See* MACH NUMBER.

Speed ranges in gas dynamics. Gases may easily move at compressible speeds. Doubling the pressure of air—from, say, 1 to 2 atm—may accelerate it to supersonic velocity. In principle, practically any large Mach number may be achieved in gas flow, as shown in the **illustration**. As Mach number increases above 0.3, the four compressible speed ranges occur: subsonic, transonic, supersonic, and hypersonic flow. Each of these has special characteristics and methods of analysis. *See* COMPRESSIBLE FLOW.

Air at 68°F (20°C) has a speed of sound of 760 mi/h (340 m/s). Thus inequality (1) indicates that air flow

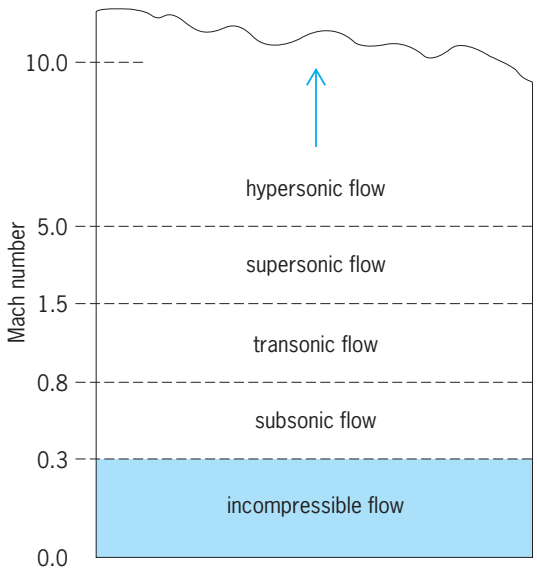

Schematic of various speed ranges of fluid motion. Incompressible flow means small but not zero Mach number.

will be incompressible at velocities up to 228 mi/h (102 m/s). This includes a wide variety of practical air flows: ventilation ducts, fans, automobiles, baseball pitches, light aircraft, and wind forces. The result is a wide variety of useful incompressible flow relations applicable to both liquids and gases.

Mass, momentum, and energy relations. The incompressible flow assumption yields very useful simplifications in analysis. For example, a mass balance reduces to a velocity-space relation, independent of fluid density. In steady flow through a variable-area duct, the volume flow Q (m^3/s or ft^3/s) is constant as given in Eq. (2), where V is the average velocity and

$$Q = VA = \text{constant} \tag{2}$$

A the cross-section area. Constricted areas (nozzles) yield high velocities; for example, in a hypodermic syringe the thumb moves slowly but the jet emerges rapidly from the needle.

The momentum equation for steady, inviscid, incompressible flow reduces to the Bernoulli theorem, given by Eq. (3), where p is pressure, ρ is density

$$\frac{p}{\rho} + \frac{V^2}{2} + gz = \text{constant} \tag{3}$$

(assumed constant), g is the acceleration of gravity, and z is the elevation. Combining Eqs. (2) and (3) gives, for example, the formula for flow rate of a Bernoulli-type flowmeter. *See* BERNOULLI'S THEOREM; FLOW MEASUREMENT.

Similarly, the general energy equation reduces for incompressible flow to a form of the heat conduction equation, for which many practical solutions are known. *See* CONDUCTION (HEAT); CONVECTION (HEAT).

The incompressible flow assumption also greatly simplifies analysis of viscous flow in boundary layers. Pressure and density are eliminated as variables, and velocity becomes the only unknown, with many accurate solution techniques available. *See* BOUNDARY-LAYER FLOW.

Any disturbance in either a gas or liquid causes the propagation of pressure pulses or "sound waves" in the fluid. This is not a flow but rather a problem in wave propagation and particle oscillation in an otherwise nearly still fluid. *See* SOUND. Frank M. White

Bibliography. I. G. Currie, *Fundamental Mechanics of Fluids*, 2d ed., 1993; V. L. Streeter and E. B. Wylie, *Fluid Mechanics*, 9th ed., 1998; B. Thwaites, *Incompressible Aerodynamics*, 1960, reprint 1987; F. M. White, *Fluid Mechanics*, 4th ed., 1998; C. S. Yih, *Fluid Mechanics*, 1969, reprint 1988.

Index fossil

The skeletal remains (body fossils) or, rarely, the traces of activity (trace fossils) of animals, plants, protistans, and bacteria that can be used to uniquely identify the ages of sedimentary rocks. The term "index fossil" is nearly synonymous with guide fossil or zone fossil. *See* FOSSIL; TRACE FOSSILS.

With the rare exception of fossils that are reworked, that is, eroded from older beds and incorporated into younger sediments, the simple occurrence of particular fossils gives an unambiguous age bracket for the containing strata. Biostratigraphy, the study of the stratigraphic distribution of index or zone fossils, is the primary means by which most sedimentary rocks are assigned to a particular age and then regionally correlated. Index fossils are an important tool in the search for petroleum, coal, and other key geologic resources. *See* GEOLOGIC TIME SCALE; STRATIGRAPHIC NOMENCLATURE; STRATIGRAPHY.

Basis of biostratigraphy. The idea of using index fossils to identify and correlate—or match—strata of a particular geologic age dates to the early 1800s, when the British geologist William Smith established the "law of faunal and floral succession" and used this principle to help trace and map the strata of England, producing the first major geological map. This guiding principle notes that fossil species, genera, and larger groups appear and disappear in the geological record in a definite order or succession (**Fig. 1**); implicit in this notion is the assumption that the first and last appearances of index fossil species are approximately synchronous across large regions.

The "law" of faunal succession makes the tacit assumption that once a species or larger group becomes extinct it can never reappear. The true last appearance of a species, its extinction, is final and global. Thus, species should not be of markedly different ages in different places: finding the species in a particular stratum or horizon at any locality places that stratum into a narrow time slot and, within limits, finding that species in any other location confines that stratum within the same bracket. This concept was developed empirically by Smith and other stratigraphers well before the basic processes of evolution were understood. Darwin's recognition of evolution

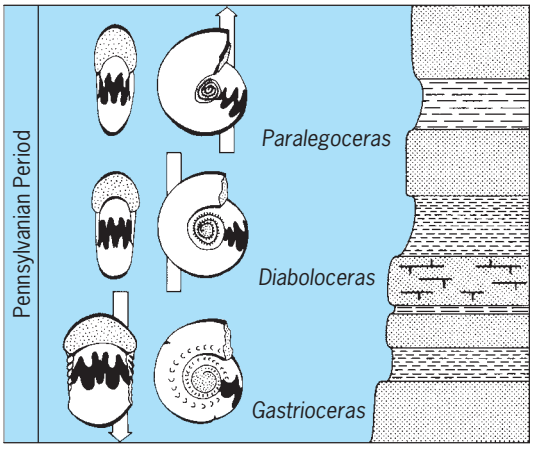

Fig. 1. Three cephalopod genera forming an evolutional sequence. Their short ranges (indicated by bars) and distinctive suture lines (the area between two sutures is shown in black) make them ideal index fossils for the early Pennsylvanian Period. (*After A. K. Miller and W. M. Furnish, Middle Pennsylvanian Schistoceratidae (Ammonoidea), J. Paleontol., 32(2):254, 1958*)

by natural selection provided a theoretical basis for biostratigraphy: evolution is an irreversible process, species arise uniquely from other species at particular times, and once species become extinct they do not reappear later in the geological record; evolution provides an "arrow" of time. Gradually evolving lineages might potentially pose a problem for recognition of index species, raising the question of how to define the time limits of a species that is gradually and continuously evolving. However, the fact is that index fossils are typically recognizable with little or no change for prolonged periods. This empirical observation fostered the concept of punctuated equilibrium. As recognized in the early 1970s by Niles Eldredge and Stephen Jay Gould, evolution commonly does not occur gradually and continuously. Rather, in the model of punctuated equilibrium, most morphological transformation occurs geologically rapidly at the time of speciation—the splitting of new species. Following this burst of change, species lineages may change very little (stasis) for millions of years. Therefore, punctuated equilibrium provides a theoretical basis for the use of index fossils to define biozones. *See* EXTINCTION (BIOLOGY); ORGANIC EVOLUTION; SPECIATION.

The concept of index fossils was originally developed to identify particular stratigraphic units, especially formations, which are basic mappable units. Commonly the formation name was incorporated into the index fossil name: for example, the brachiopod *Gypidula coeymanensis* was identified as unique to the Lower Devonian Coeymans Formation. Paleontologists have compiled lists and catalogs of index fossils and associated formations. A classic example is H. W. Shimer and R. R. Shrock's *Index Fossils of North America*, which illustrates and lists the ages and formations of occurrence of several thousand species of invertebrate fossils. However, an overreliance on single index fossils led to some errors. There was an implicit assumption that the fossils and the formation that contained them represented the same precise time intervals wherever found. In some instances, the formations represented facies (records of particular environments) that were diachronous, that is, of different ages in different places. If the characteristic index species migrated across environments through time, their appearance in local sections would also be diachronous. In some cases, biostratigraphers tended to split species on very minor taxonomic features that represented slight ecological or local differences, not characteristic of particular ages. Species were assigned ranges coincident with the thicknesses of formations, and too often their actual ranges were not critically evaluated. Also, the application of index fossils in correlation was typically limited by the restriction of those fossils to local areas. *See* FACIES (GEOLOGY).

From the original conception of index fossils evolved the concept of zone fossils. Zone fossils are index fossils that characterize a particular biozone, the fundamental intervals of biostratigraphy; biozones are based on the geological age ranges of particular index fossils (**Fig. 2**). Such zonal fossils ideally should be widespread, abundant, and short-lived (see below). The true range represents the total span of a species from its true first appearance (also called first appearance datum or FAD) to its true last appearance (last appearance datum or LAD). In practice, the local range of a given index fossil is almost never the full range of that species (**Fig. 3**). The local FAD may represent the immigration of the species into the local area in response to favorable environmental conditions; the local LAD (extermination) of the species may record the onset of unfavorable conditions and/or emigration of the species out of the local area. Nonetheless, when gauged against independent indicators of age such as volcanic ash

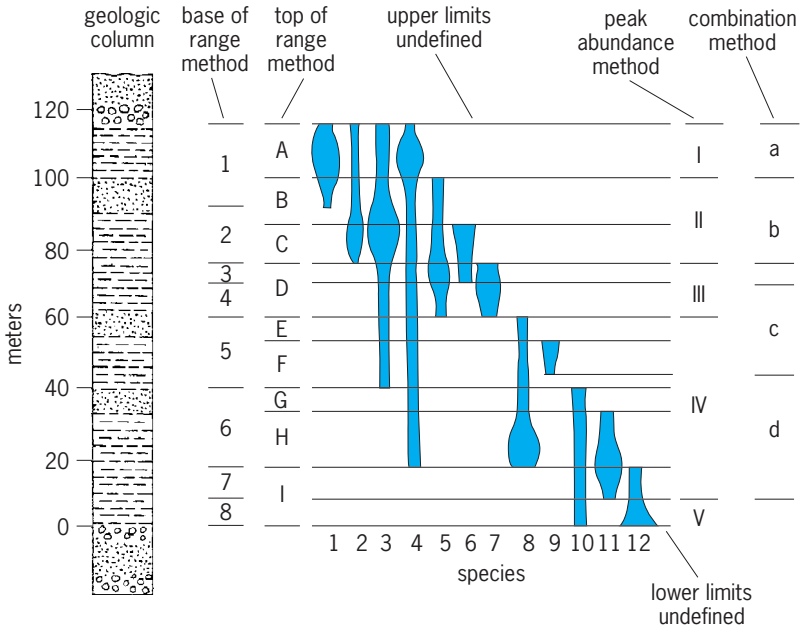

Fig. 2. Different concepts of range zones based on various combinations of overlap of two or more index species. (*From D. R. Prothero and F. Schwab, Sedimentary Geology: An introduction to Sedimentary Rocks, Freeman, 1996*)

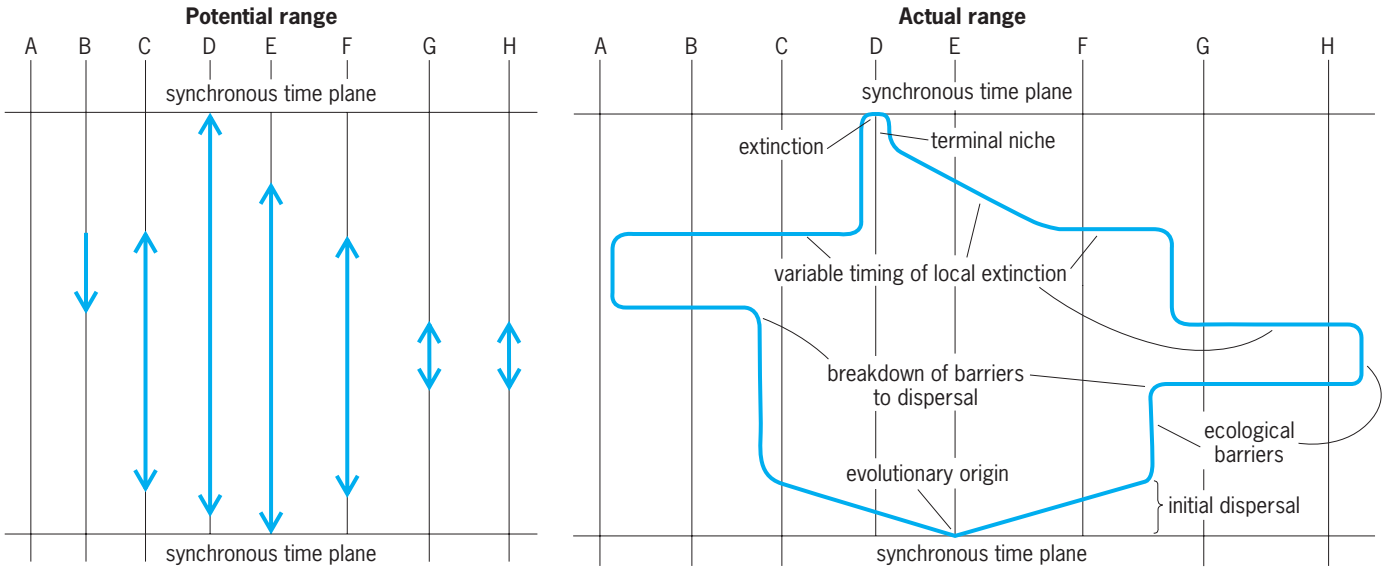

Fig. 3. Potential range versus actual range of an index fossil species. Note that the actual time of occurrence at a given locality is generally some fraction of the actual range of the species because of the limitations of environment, immigration, and emigration. (*From M. Bassett, in D. E. G. Briggs and P. R. Crowther, Palaeobiology: A Synthesis–Part 1, 1991*)

layers or magnetic reversal horizons, many well-chosen index species are demonstrated to have ranges that are roughly synchronous over considerable areas. The biozone is technically defined as a body of strata representing the range of a given species. However, biozones are also a reflection of the time slice representing the duration of the species; this somewhat more abstract concept is referred to as the chronozone.

Biozones are named for their characteristic index fossils, although the names may be abbreviated by just the species name. Separate zonal systems have been developed for different taxonomic groups. For example, the *kockelianus* zone of the Middle Devonian is named for the conodont subspecies *Tortodus kockelianus kockelianus*; the *Agoniatites vanuxemi* Zone is based on a Middle Devonian ammonoid and overlaps in time with the *kockelianus* zone. This system may seem redundant, but it is necessary because various groups evolve at different times and at different rates, and some rocks may carry one fossil and not another. Moreover, the overlap of different index fossils may more precise subdivisions of the rocks. However, calibration among zones is necessary, and generally one zonation scheme, for example, that based on conodonts for Devonian time, is favored for purposes of establishing international stratigraphic boundaries.

In the 1960s, the International Union of Geological Sciences (IUGS) established an International Commission on Stratigraphy and many subcommissions, charged with establishing international boundaries of geological periods and their subdivisions on the basis of biostratigraphy. This commission established the practice of defining particular boundaries, at specific reference sections called global stratotypes, at well-studied, readily accessible localities. At these localities, the boundaries of specific zonal or index species can be precisely located in continuous sections without major gaps. For example, the Silurian-Devonian system boundary was defined precisely at the first appearance (FAD) of the fossil graptolite species *Monograptus uniformis* at the stratotype, a section of alternating shale and limestone in a quarry at Klonk, Czech Republic. Not only was a monument erected at this section in 1972, but a golden spike was driven into the outcrop at the precise point where the lowest specimen of *M. uniformis* was collected, symbolically establishing a standard for global comparison. This notion of the "golden spike" has been figuratively applied in numerous other cases for which global standards were identified based on globally recognizable index fossils at boundary stratotype sections. This provides a needed objectivity to biostratigraphic zonation.

Biostratigraphers long recognized that time increments could be defined more precisely as the overlap of two or more index species. This concept of overlapping of first and/or last appearances is sometimes referred to as Oppel zones in honor of the nineteenth-century paleontologist Carl Albert Oppel, who advocated using overlapping ranges to correlate strata. A more quantitative approach to using index fossils in correlating strata was developed in the 1960s by the American stratigrapher Alan Shaw. Shaw suggested that the ranges of many index fossils in stratigraphic sections to be matched could be plotted against one another as the *X*/*Y* axes of a graph. The FADs and LADs of various species, measured in meters (or feet) from the base of each section in each of the two sections, could then be plotted as coordinates of points on the graph (**Fig. 4**). By plotting a series of points, a line of correlation could be determined. Most importantly, the process could be used iteratively to refine the first and last appearances of species on an idealized standard reference section, derived from the most complete and typically thickest stratigraphic section. Such a method is

quantitative and makes use of entire assemblages of fossils to define time slices.

Criteria. The identification of fossil species as good index fossils is based on several criteria. In general, the species should represent a group that evolved rapidly. Obviously, the precision with which geologic time can be subdivided is based upon the ranges of the index species: The faster the rate of evolution, the shorter the period of time represented by a given index or zonal fossil species and the narrower the limits of relative age assigned to the rocks containing the species.

An index fossil also must be present in the rocks in sufficient numbers to be found with reasonable

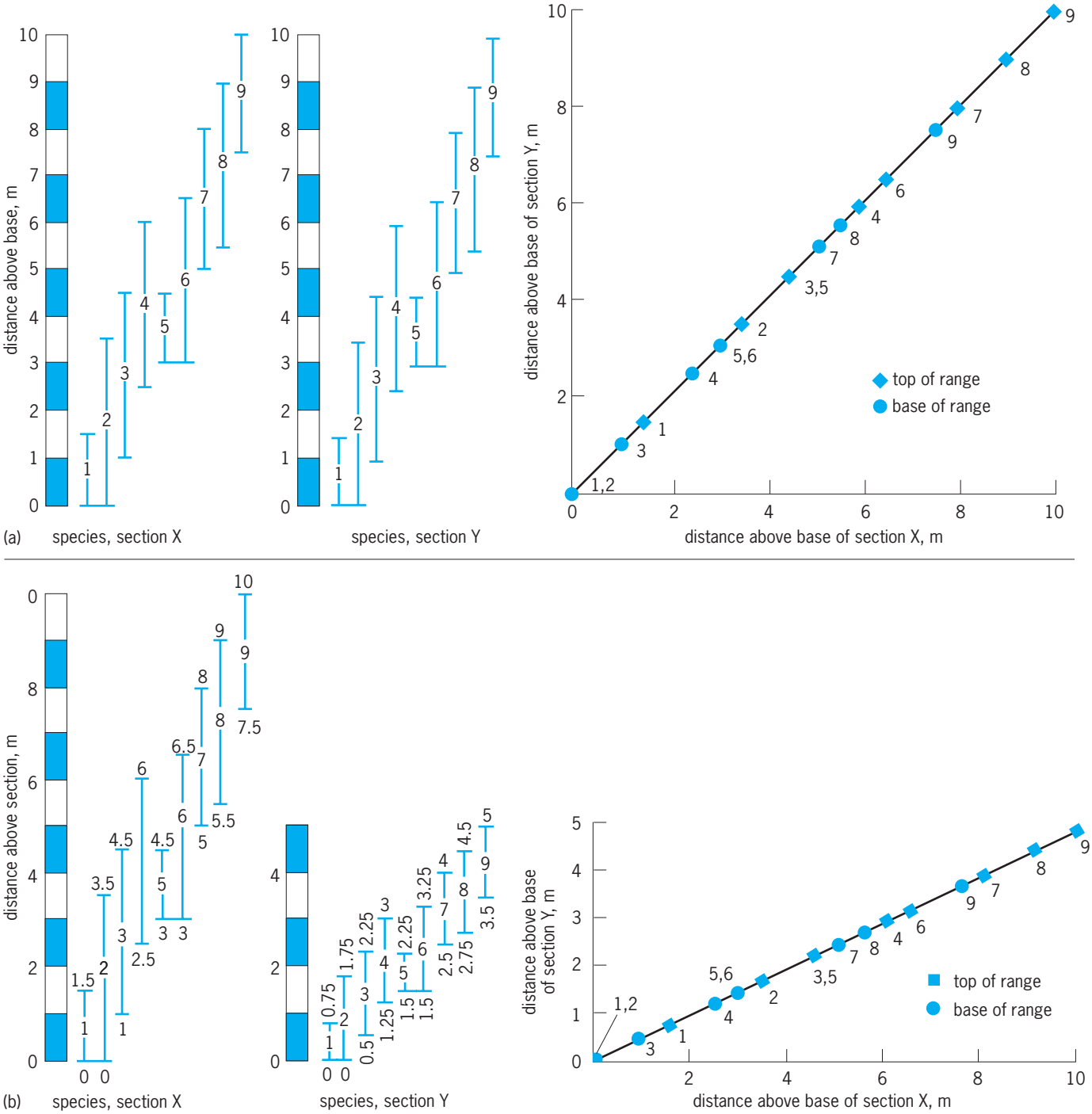

Fig. 4. Graphic correlation method of Alan Shaw. Note that the *X* and *Y* axes of the graphs are scaled diagrams of measured sections of the two different outcrops to be correlated. The points plotted on the graphs are the corresponding heights of the first or last appearances of various index fossils in each of the stratigraphic sections. The line of correlation (LOC) is a best-fit line drawn through these points. (*a*) Two outcrops with identical stratigraphic sections; note that the heights or first or last appearances of a given species are identical and the angle of the LOC is 45°. (*b*) Correlation plot of two outcrops in which one had a higher rate of sedimentation than the other; note the more tightly spaced FAD and LAD points for the section with lower sedimentation rate. The slope of the LOC is less than 45°. (*From D. R. Prothero and F. Schwab, Sedimentary Geology: An Introduction to Sedimentary Rocks, Freeman, 1996*)

effort, must be relatively easy to collect and identify, and must be geographically extensive so that the zone it defines is widely recognizable. Ideally, the fossil should be resistant to dissolution or surface weathering. Features that permit the fossil to be identified under conditions of incomplete preservation or weathering are advantageous. Moreover, the usefulness of the fossil is increased if it can be extracted readily from the rock matrix. For example, the minute "teeth" or elements of primitive vertebrates, called conodonts, are composed of apatite (calcium phosphate), which is relatively insoluble in acetic or muriatic acid; these fossils can thus be extracted by dissolving limestones in these acids. Organic-walled fossils, such as pollen, spores, and acritarchs (algal resting cysts), can be released from shales by dissolving the shales in hydrofluoric acid. The study of palynology deals with such organic-walled microfossils (palynomorphs), which are extractable using hydrofluoric acid. Small size in such cases is an advantage as many fossils must be identified from small samples of drill core (material removed from the earth by drilling) or rock chips of drilled well cuttings. Such microfossils, including foraminiferans, radiolarians, conodonts, and spores and pollen, must be studied and identified using microscopy. Micropaleontology is widely used in the petroleum industry. *See* CONODONT; FORAMINIFERA; MICROPALEONTOLOGY; PALYNOLOGY.

The best index fossils would thus be remains of organisms that either live over a wide array of environments or may be distributed widely on water currents or wind. This is most likely to be true for fossils of planktonic (free-floating) organisms, such as many foraminiferans and radiolarians, or nektonic (free-swimming) forms, such as ammonites.

The presence or absence of particular index fossils should not reflect local environments, as the best index fossils are those completely independent of environmental factors. Ideally, index and especially zone fossils should thus occur independent of facies, the records of different environments.

This is not to say that benthic (sea-bottom-dwelling) organisms may not be useful as biostratigraphic markers. Many benthic forms do evolve rapidly and can provide time markers, at least within a particular range of environments. Moreover, many invertebrate animals, even those which are attached in one place (sessile) throughout their adult lives, as is the case for many corals, brachiopods, or clams, have a planktonic or nektonic larval stage in their life cycles. This stage, especially if it involves larvae that feed in the plankton (planktotrophic larvae), may last for many days, and the larvae may thus be dispersed widely on oceanic currents, thereby aiding in the distribution of species over large areas.

Other bottom-dwelling animals are much more tied to specific environmental conditions. Such organisms will typically migrate over time in response to the shifting of their favored environmental conditions, and some may persist for long periods of time with little change. For example, the brachiopod *Lingula* is well adapted to tidal flats and has survived in these environments with little change for more than 500 million years. As such, these fossils may be useful for reconstructing ancient environments and tracing their movements through time; thus they are sometimes termed facies fossils. However, they are not as useful as index fossils, as their occurrence is governed more by environment than by time.

A particularly difficult problem is the correlation of terrestrial and marine environments because few, if any, organisms live in both types of settings. Fortunately, there are some fossils that may enable the connection between these disparate realms. For example, many land plants produce spores and pollen that are adapted for wide dispersal because of their compact, lightweight nature; thus, pollen and spores may be borne on air currents over considerable distances. These palynomorphs may then be transported into and buried within floodplain muds, lake sediments, or even offshore marine sediments. This dispersal potential makes palynomorphs attractive for linking ancient land and sea environments.

Typical index fossils through time. Few types of organisms meet all criteria for an ideal index fossil, and for good reason. Those organisms with widespread or even cosmopolitan (worldwide) distributions and those which are tolerant of a wide variety of environments are also buffered against evolution. A wide distribution and broad environmental tolerance favor survival of species even during times of crisis and major extinction. Thus, widespread species tend to evolve slowly, and vice versa. That said, there are many fossil organisms that may be useful as biostratigraphic markers, at least over particular regions or within particular types of environments. Several groups have proved widely usable as index fossils and have been widely accepted as a basis for global biochronology. Foremost among these forms are acritarchs, spores and pollen, coccoliths, foraminiferans, radiolarians, graptolites, ammonoid cephalopods, and conodonts.

Older Precambrian [prior to 600 million years before present (BP)] sedimentary rocks typically contain few age-diagnostic fossils. During the Archean to middle Proterozoic eons, from about 3.5 to 1.4 billion years ago, most life-forms were simple unicellular organisms such as bacteria. These are often difficult to isolate and identify and have not proved highly useful in biostratigraphy, although there have been attempts to use forms of stromatolites, layered mounds built by cyanobacteria (or blue-green "algae"), to identify particular time increments in the Proterozoic. In the later Proterozoic Eon, from about 1.0 to 0.5 billion years ago (10 to 5×10^8 years BP), age-diagnostic fossils—especially the reproductive and resting cysts of marine algae termed acritarchs—are increasingly abundant and provide excellent index fossils for relative dating of sedimentary rocks. In the Phanerozoic Eon, about 5.43×10^8 years BP to the present time, a wide array of marine and nonmarine invertebrate and vertebrate animals and plants provide excellent material for subdividing geologic time (see below). Single-celled eukaryotic organisms (cells with nuclei and

organelles), such as planktonic algae, first appeared in the early to middle Proterozoic eon about 20 to 15×10^8 years BP. These organisms are well represented by acritarchs. Recent advances in the study of acritarchs have established the first detailed biostratigraphic zonation of sediments dating back to at least late Proterozoic time. Acritarchs are also useful microfossils for early Paleozoic biostratigraphy.

The common base of the Phanerozoic Eon, the Paleozoic Era, and the Cambrian Period (about 5.4×10^8 years BP) is established as the first appearance of a trace fossil of *Trichophycus pedum*, the oldest complex burrow, in a stratotype section in Newfoundland—one of the rare cases of the use of a trace fossil as an index fossil. Trilobites (extinct, segmented arthropods) provide critical index fossils for zonation in the Cambrian and Ordovician periods. Proteinaceous skeletons of planktonic, colonial organisms called graptolites have proved useful in correlating Ordovician (5.0 to 4.4×10^8 years BP) and Silurian (4.4 to 4.17×10^8 years BP) rocks. Conodonts, recently recognized as parts of the feeding apparatus of small eellike marine vertebrates, are among the most important index fossils for much of the Paleozoic Era (5.4 to 2.5×10^8 years BP) and the Triassic Period (2.5 to 2.0×10^8 years BP). Global zonation of the Devonian Period, for example, includes more than 40 successive conodont zones. Ammonoids (coiled cephalopods with variably complex suture patterns reflecting the line of junction of crenulated septa and the external walls of the shells) first appeared in the Devonian Period (about 3.9×10^8 years BP) and evolved rapidly throughout their history to final extinction at the end of the Cretaceous Period (6.5×10^8 years BP). Goniatite ammonoids with relatively simple zigzag sutures are useful in Late Paleozoic biostratigraphy. Ceratitic ammonoids with intermediate sutural complexity and ammonites with highly complex sutures have proved to be very useful index fossils for the Triassic Period and later Mesozoic Era, respectively. Foraminiferans (protozoans with agglutinated or calcareous tests) evolved in the Cambrian and diversified into many forms during the Phanerozoic Eon; various groups of foraminiferans have been very useful index fossils at different times. Large coiled foraminiferans about the size and shape of rice grains, the fusulinids, are critical index fossils during the Pennsylvanian and Permian periods (3.1 to 2.5×10^8 years BP); various smaller benthic foraminiferans are important biostratigraphic microfossils in the Mesozoic (2.5 to 0.65×10^8 years BP) and Cenozoic (0.65 to 0×10^8 years BP) eras, whereas planktonic foraminiferans are exceptionally important microfossils for correlation in rocks of the late Cretaceous onward. In addition, the beautiful silica lattice skeletons of radiolarians, another planktonic protozoan group, have been used extensively in biostratigraphy of deep-sea sediments throughout the Phanerozoic. From the Cretaceous Period (about 1.2 to 0.65×10^8 years BP) onward, the skeletons of planktonic algae become important as index fossils in deep-sea sediments. These microfossils include the tiny diatoms, and skeletal plates of calcite called coccoliths; the latter, which form a dominant component of chalks, belong to a category of ultramicroscopic fossils, sometimes termed nannofossils, that require a scanning electron microscope for study. *See* CEPHALOPODA; FUSULINACEA; GRAPTOLITHINA; RADIOLARIA.

A wide variety of shelly invertebrate groups, including brachiopods, mollusks, corals, ostracodes (tiny bivalved crustaceans), and echinoderms are also used as index fossils for marine deposits. From the Devonian onward, spores and pollen of land plants are useful indices for nonmarine to marginal marine zonation. Freshwater mollusks, ostracodes, and even mammal teeth have also proved useful for zonation of nonmarine sediments in the Cenozoic Era. *See* PALEOECOLOGY; PALEONTOLOGY.

Carlton E. Brett

Bibliography. M. Brasier, *Micropaleontology*, 1986; D. E. G. Briggs and P. R. Crowther, *Palaeobiology: A Synthesis–Part 1*, 1991, Part 2, 2001; E. N. K. Clarkson, *Invertebrate Palaeontology and Evolution*, 3d ed., 1992; D. R. Prothero and F. Schwab, *Sedimentary Geology: An Introduction to Sedimentary Rocks*, Freeman, 1996; H. W. Shimer and R. R. Shrock, *Index Fossils of North America*, 9th printing, MIT Press, 1972.

Indian Ocean

The smallest and geologically the most youthful of the three oceans. It differs from the Pacific and Atlantic oceans in two important aspects. First, it is landlocked in the north, does not extend into the cold climatic regions of the Northern Hemisphere, and consequently it is asymmetrical with regard to its circulation. Second, the wind systems over its equatorial and northern portions change twice each year, causing an almost complete reversal of its circulation.

General Features

The eastern and western boundaries of the Indian Ocean are 147 and 20°E, respectively. In the southeastern Asian waters the boundary is usually placed across Torres Strait, and then from New Guinea along the Lesser Sunda Islands, across Sunda Strait and Singapore Strait. The surface area within these boundaries is about 21% of the surface of all oceans. The average depth is about 12,630 ft (3850 m).

The ocean floor is divided into a number of basins by a system of ridges (**Fig. 1**). The largest is the Mid-Oceanic Ridge, the greater part of which has a rather deep rift valley along its center. It lies like an inverted Y in the central portions of the ocean and ends in the Gulf of Aden. Most of the ocean basins separated by the ridges reach depths in excess of 16,000 ft (5000 m). The Sunda Trench, stretching along Java and Sumatra, is the only deep-sea trench in the Indian Ocean. The Andaman Basin is separated from the open ocean by a 14,600-ft (1400-m) sill. The Red Sea has a maximum depth of 9301 ft (2835 m), but its entrance at the Strait of Bab-el-Mandeb is only 410 ft

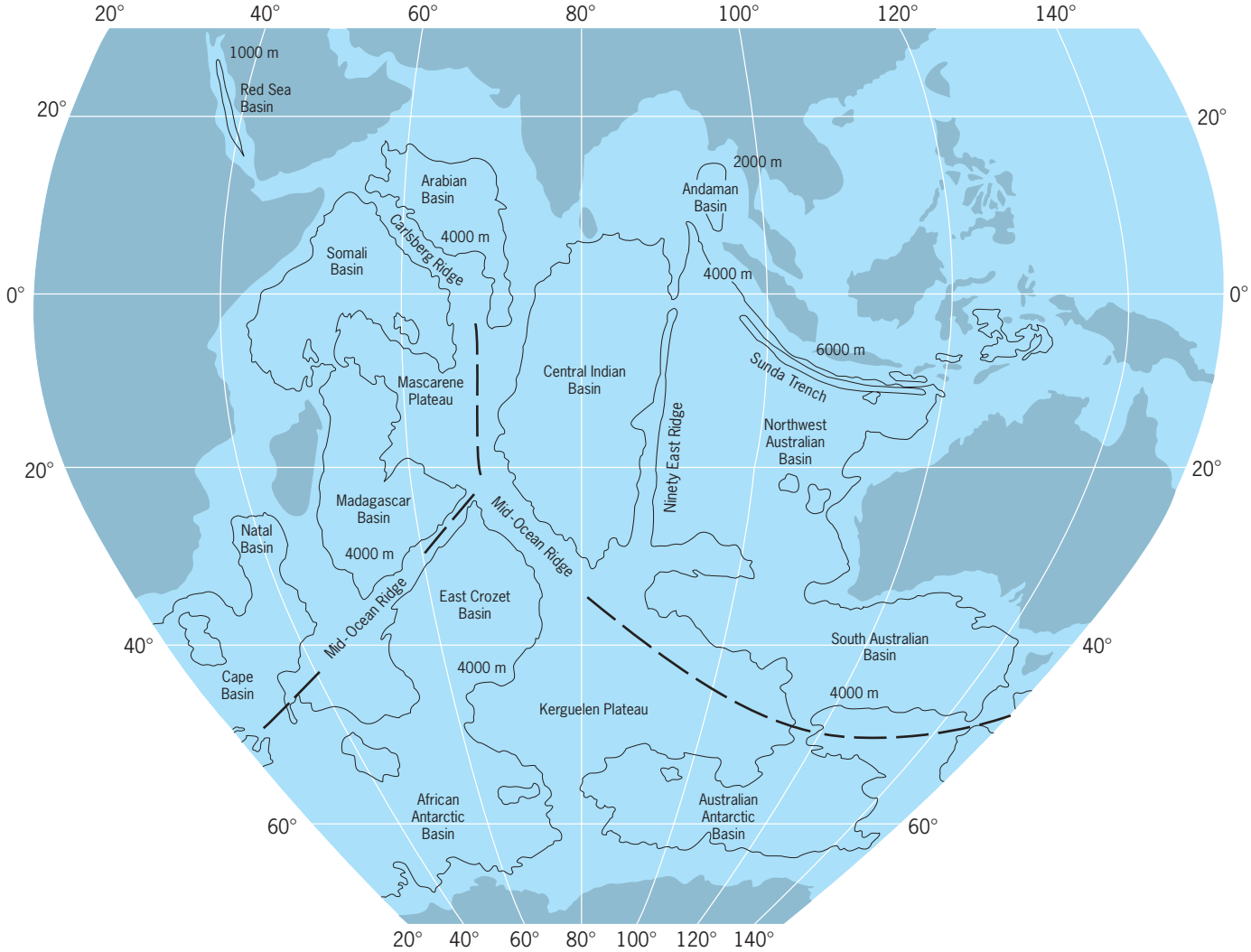

Fig. 1. Physiographic features of the Indian Ocean.

(125 m) deep. East of the Mid-Oceanic Ridge, deep-sea sediments are chiefly red clay; in the western half of the ocean, globigerina ooze prevails and, near the Antarctic continent, diatom ooze. *See* BASIN; MARINE GEOLOGY; MARINE SEDIMENTS; MID-OCEANIC RIDGE.

Wind systems. Atmospheric circulation over the northern and equatorial Indian Ocean is characterized by the changing monsoons. In the southern Indian Ocean, atmospheric circulation undergoes only a slight meridional shift during the year (**Fig. 2**). During winter in the Northern Hemisphere, from December through February, the equatorial low-pressure trough is situated at about 10°S and continues into a low-pressure system over northern Australia. A strong high-pressure system lies over Asia. This situation causes the Northeast Monsoon to blow everywhere north of the Equator. The winds cross the Equator from north to south and then usually become northwest winds before reaching the Intertropical Convergence. The subtropical high-pressure ridge of the Southern Hemisphere is situated near 35°S. North of it, the Southeast Trades blow. South of 40°S, winds from the west prevail and are associated with cyclones traveling around Antarctica.

During summer in the Northern Hemisphere, from June through August, a low-pressure system is developed over Asia with the center around Iran. The subtropical high-pressure ridge of the Southern Hemisphere has shifted slightly northward and continues into a high-pressure system over Australia. The Southeast Trades are more strongly developed during this season, cross the Equator from south to north, and become the Southwest Monsoon, bringing rainfall and the wet season to India and Burma. Atmospheric circulation during June through August is much stronger and more consistent than during February. In the Southern Hemisphere the West Wind Belt has shifted about 5° to the north, with westerly winds starting from 30°S and storms becoming stronger and more frequent during winter in that hemisphere. *See* MONSOON METEOROLOGY; TROPICAL METEOROLOGY.

Circulation. The surface circulation is caused largely by winds and changes in response to the wind systems (**Fig. 3**). In addition, strong boundary currents are formed, especially along the western

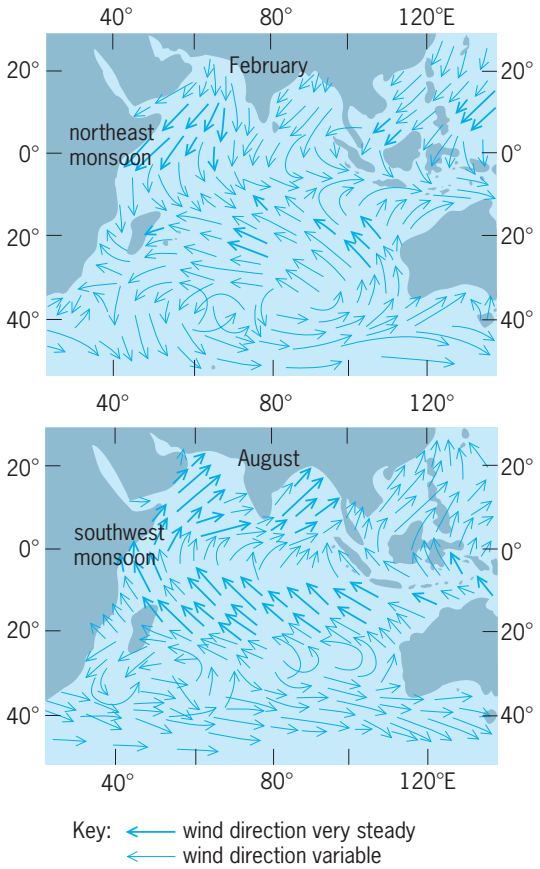

Fig. 2. Winds over the Indian Ocean.

coastline, as an effect of the Earth's rotation and of the boundaries created by the landmasses. In the Southern Hemisphere south of 35°S, a general drift from west to east is found as a result of the prevailing west winds. Near 50°S, where west winds are strongest, the Antarctic Circumpolar Current is embedded in the general West Wind Drift. In the subtropical southern Indian Ocean, circulation is anticyclonic. It consists of the South Equatorial Current flowing west between 10 and 20°S, a flow to the south along the coast of Africa and Madagascar, parts of the West Wind Drift, and flow to the north in the eastern portions of the ocean, especially along the coast of Australia. The flow to the south between Madagascar and Africa is called the Mozambique Current. It continues along the coast of South Africa as the Agulhas Current, most of which turns east into the West Wind Drift. The circulation in the Southern Hemisphere south of 10°S changes only slightly with the seasons.

North of 10°S the changing monsoons cause a complete reversal of surface circulation twice a year. In February, during the Northeast Monsoon, flow north of the Equator is mostly to the west and the North Equatorial Current is well developed. Its water turns south along the coast of Somali and returns to the east as the Equatorial Countercurrent between about 2 and 10°S. In August, during the Southwest Monsoon, the South Equatorial Current extends to the north of 10°S; most of its water turns north along the coast of Somali, forming the strong Somali Current. North of the Equator flow is from west to east and is called the Monsoon Current. Parts of this current turn south along the coast of Sumatra and return to the South Equatorial Current. During the two transition periods between the Northeast and the Southwest monsoons in April-May and in October, a strong jetlike surface current flows along the Equator from west to east in response to the westerly winds during these months.

Although the Antarctic Circumpolar Current reaches to great depths, probably to the bottom, most of the other currents are much shallower. The Agulhas Current and the Somali Current are examples of deep currents, but the remainder of the circulation is limited to the upper surface of the ocean. The deep circulation of the Indian Ocean consists of western boundary currents that carry water of Antarctic origin northward into the various deep-sea basins. This water ascends slowly, is integrated into the surface circulation, and returns south with the Agulhas Current. At the Equator, during the time of the Northeast Monsoon, an Equatorial Undercurrent is found flowing as a subsurface current from west to east. This current is weaker than the corresponding currents in the Pacific and Atlantic oceans. A massive flow of water takes place from the Pacific to the Indian Ocean, called the Indonesian Throughflow. *See* OCEAN CIRCULATION.

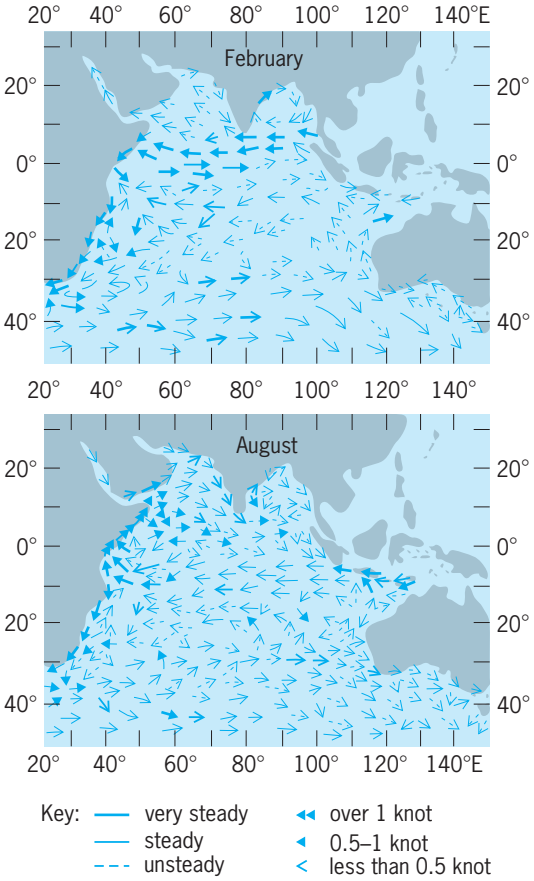

Fig. 3. Surface currents of the Indian Ocean.

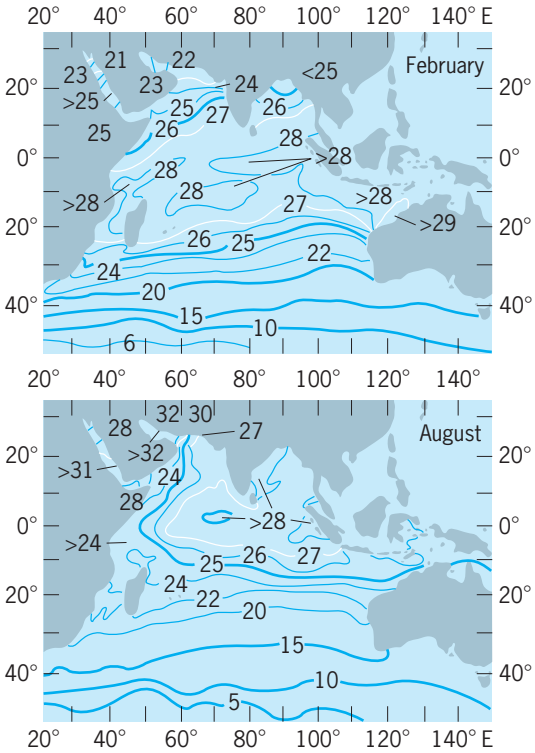

Fig. 4. Sea-surface temperature (°C) of Indian Ocean.

Surface temperature. The pattern of sea-surface temperatures changes considerably with the seasons (**Fig. 4**). During February, when the Intertropical Convergence is near 10°S, the heat equator is also in the Southern Hemisphere and most of the area between the Equator and 20°S has temperatures near 82°F (28°C). The water in the northern parts of the Bay of Bengal and of the Arabian Sea is much cooler, and temperatures below 68°F (20°C) can be found in the northern portions of the Persian Gulf and the Red Sea. In the Southern Hemisphere temperatures decrease gradually from the tropics to the polar regions. Surface circulation affects the distribution of temperature, and warm water spreads south along the coast of Africa and cool water north off the west coast of Australia, causing the isotherms to be inclined from west to east.

During August high temperatures are found in the Northern Hemisphere and in the equatorial region. The Somali Current advects cool water along the coast of Africa to the north. Simultaneously the Southwest Monsoon causes strong upwelling of cooler subsurface water, lowering average water temperature in August to less than 72°F (22°C) along Somalia, and to less than 75°F (24°C) along Arabia. In the Southern Hemisphere isotherms are almost 10° of latitude farther north.

Surface salinity. The distribution of surface salinity is controlled by the difference between evaporation and precipitation and by runoff from the continents (**Fig. 5**).

High surface salinities are found in the subtropical belt of the Southern Hemisphere, where evaporation exceeds rainfall. In contrast, the Antarctic waters are of low salinity because of heavy rainfall and melting ice. Another area of low salinities stretches from the Indonesian waters along 10°S to Madagascar. It is caused by the heavy rainfall in the tropics. The Bay of Bengal has very low salinities, as a result of the runoff from large rivers. In contrast, because of high evaporation the Arabian Sea has extremely high salinities. High salinities are also found in the Persian Gulf and in the Red Sea, representing the arid character of the landmasses surrounding them. The salinity distribution changes relatively little during the year; however, south of India from the Bay of Bengal to the west, a flow of low-salinity water, caused by the North Equatorial Current, can be noticed during February.

Surface water masses. The different climatic conditions over various parts of the Indian Ocean cause the formation of characteristic surface water masses. The Arabian Sea water is of high salinity, has a moderate seasonal temperature variation, and can be classified as subtropical. The water in the Bay of Bengal is of low salinity and always warm, representing tropical surface water. Another type of tropical surface water stretches from the Indonesian waters to the west and is called the Equatorial Surface Water. Subtropical Surface Water has a seasonal temperature variation of between 59 and 77°F (15 and 25°C) and is found in the subtropical regions of the Southern Hemisphere. Its southern boundary is the Subtropical Convergence coinciding with temperatures of about 59°F (15°C). From there, temperature and salinity decrease in the area of the transition water to

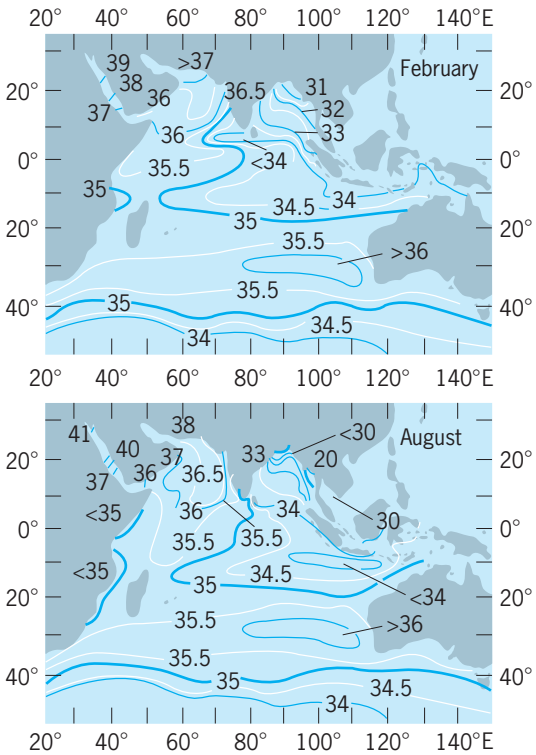

Fig. 5. Sea-surface salinity of the Indian Ocean in parts per thousand by weight (‰).

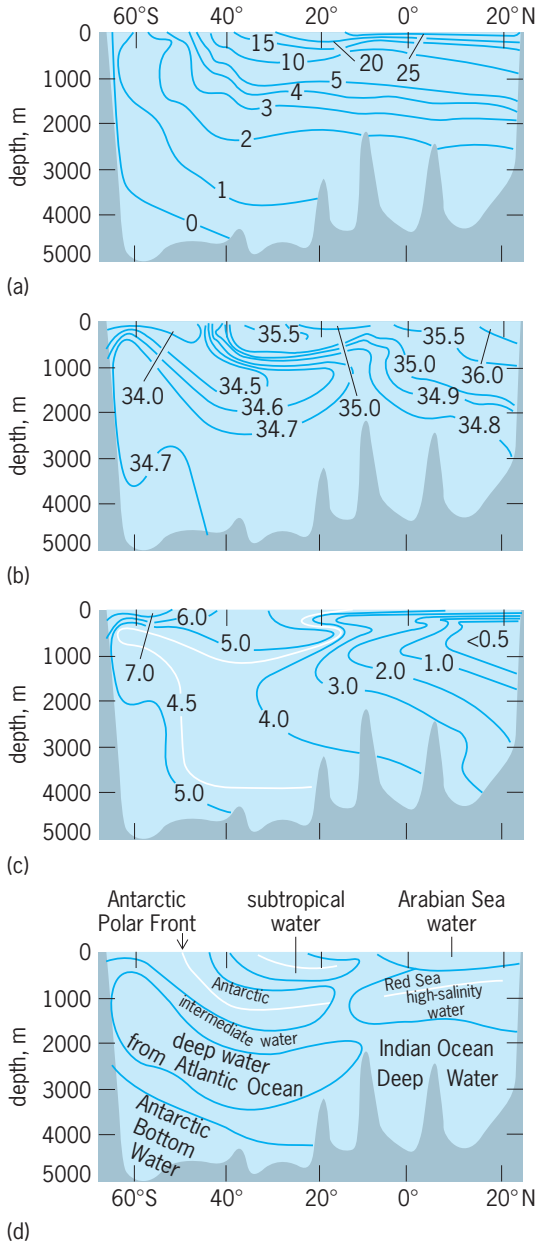

Fig. 6. Indian Ocean from the Arabian Sea to Antarctica showing (*a*) temperature (°C), (*b*) salinity (‰), (*c*) oxygen content (ml/liter), and (*d*) water masses.

about 39°F (4°C) at the Antarctic Polar Front. South of the Antarctic Polar Front, Antarctic Surface Water of low salinity is found; its temperature is near the freezing point in winter and is approximately 36°F (2°C) in summer.

Subsurface water masses. In the Indian Ocean, warm water of more than 59°F (15°C) occupies only a very thin surface layer (**Fig. 6***a*). This layer contains the tropical and subtropical water masses. Subtropical water of the Southern Hemisphere spreads as a subsurface salinity maximum toward the Equator at depths between 330 and 660 ft (100 and 200 m) beneath the lighter tropical surface water.

Water of high salinity formed in the Red Sea and in the Persian Gulf leaves these basins and spreads as a subsurface layer of high salinity throughout the Arabian Sea (Fig. 6*b*). To the south it can be traced as far as Madagascar, and to the east as far as Sumatra. Water of low salinity and a temperature of approximately 39°F (4°C) sinks near the Antarctic Polar Front and spreads north as a salinity minimum. It is called the Antarctic Intermediate Water.

The deep and bottom water of the Indian Ocean is of external origin. South of Africa, water of rather high salinity, originating south of Greenland in the North Atlantic Ocean, enters the Indian Ocean, filling the layers between 4900- and 9800-ft (1500- and 3000-m) depths. The water below the 9800-ft (3000-m) depth is Antarctic Bottom Water originating in the Weddell Sea. Its temperature is lower than 32°F (0°C).

Because of their origin at the sea surface, both the Antarctic Intermediate Water and the Antarctic Bottom Water are of rather high oxygen content (Fig. 6*c*). Since their residence time is long, the water masses of the northern Indian Ocean below the surface layer have a very low oxygen content. An oxygen minimum is associated with the Red Sea water and extends south between the Antarctic Intermediate Water and the North Atlantic Deep Water. *See* ANTARCTIC OCEAN; RED SEA; SEAWATER.

Upwelling. Several areas of upwelling are found along the shores of the Indian Ocean, but all of them are seasonal in character, in contrast to upwelling areas in the other oceans. They are especially developed during the Southwest Monsoon season from May through September, when upwelling takes place in the Banda Sea, south of Java, along the coast of Somali, and off Arabia.

The strongest upwelling is found along the Somaliland coast, where surface temperatures during this season may be as low as 57°F (14°C). Water of high nutrient content ascends from deep levels and is integrated into the surface flow. These areas have a high biological production. During the Northeast Monsoon, some upwelling takes place in the Andaman Sea and off the west coast of India. *See* UPWELLING.

Tides. Both semidiurnal and diurnal tides occur in the Indian Ocean. The semidiurnal tides rotate around three amphidromic points situated in the Arabian Sea, southeast of Madagascar, and west of Perth. The diurnal tide also has three amphidromic points: south of India, in the Mozambique Channel, and between Africa and Antarctica. It has more the character of a standing wave, oscillating between the central portions of the Indian Ocean, the Arabian Sea, and the waters between Australia and Antarctica.

Along the coast of Africa, in the Bay of Bengal, and along the northwest coast of Australia semidiurnal tides prevail, with two high waters each day. Mixed tides are found around the Arabian Sea and along the coasts of Sumatra and Java. Prevailing diurnal tides with only one high water each day occur only along southwest Australia. The ranges of spring tides are high in the Arabian Sea, for example, at Aden and Bombay. The Bay of Bengal has high tidal ranges, chiefly along the coast of Burma, particularly at Rangoon. Moderately high tides are found along

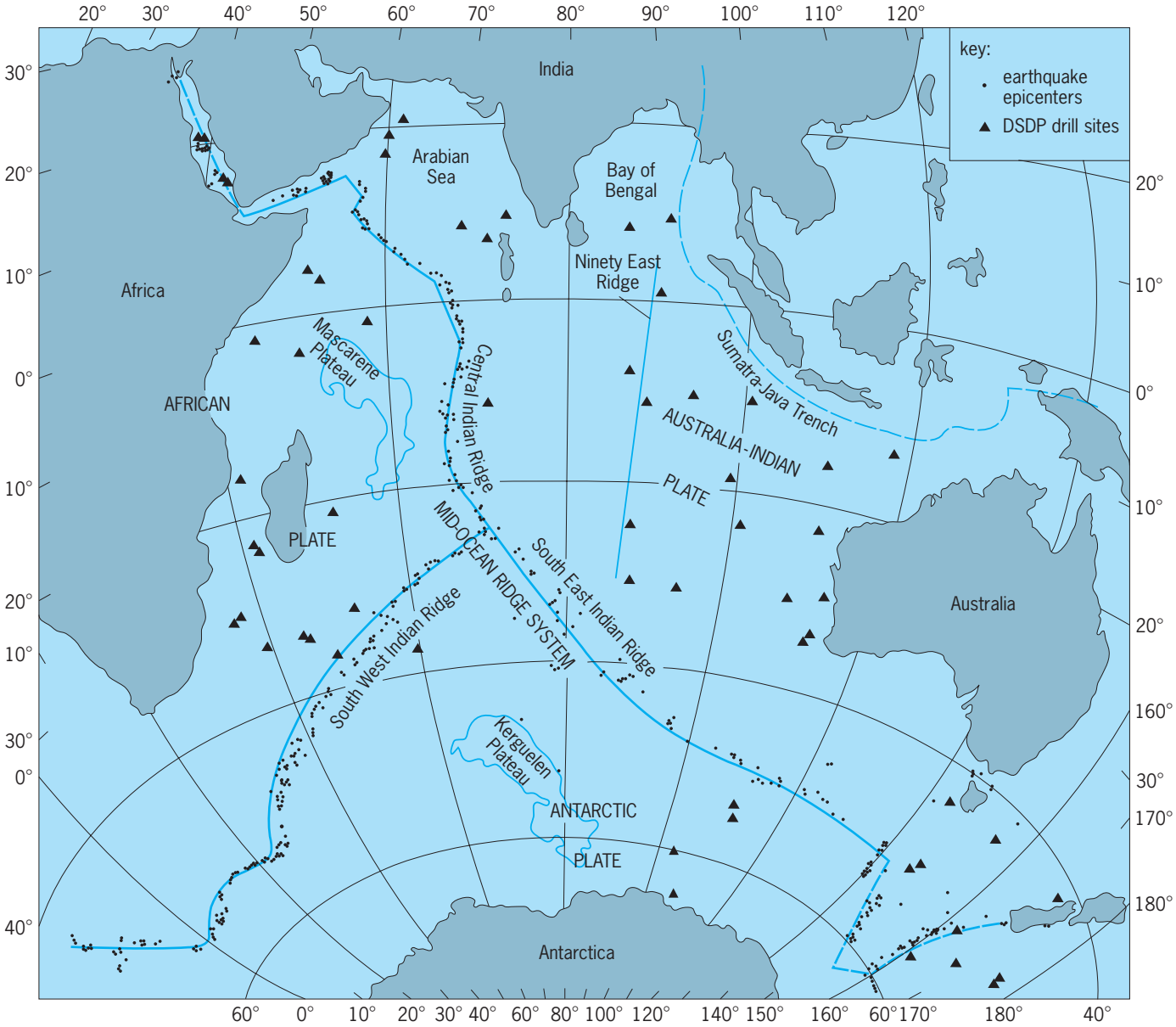

Fig. 7. Mid-ocean ridge system of the Indian Ocean. DSDP = Deep Sea Drilling Project.

the coast of Sumatra and Java, but spring tides along the northwest coast of Australia are much higher. *See* TIDE.

Klaus Wyrtki

Geology and Geophysics

The Indian Ocean, the most youthful of the major ocean basins, was probably formed entirely within the last 1×10^8 years. This ocean, in effect, represents that area created during the fragmentation (through processes of sea-floor spreading and plate tectonics) of the ancient supercontinent of Gondwanaland. *See* CONTINENTAL DRIFT; JURASSIC; PLATE TECTONICS.

Major plates. The Indian Ocean is composed of parts of three major plates: on the west, the African plate; on the east and north, the Australian-Indian plate; and on the south, the Antarctic plate. There are many spectacular morphologic features within the Indian Ocean. The major one is the system of mid-ocean ridges that consists of three branches (**Fig. 7**). The mid-ocean ridge system is seismically active and marks the sites of the creation of new oceanic crust. Between the mid-ocean ridges and the surrounding continents lie the major basins of the Indian Ocean. With one exception, the boundaries of the Indian Ocean are characterized by passive continental margins, that is, margins that are not seismically active and that lie well away from the boundaries between major plates. *See* CONTINENTAL MARGIN.

In addition to the mid-ocean ridge systems, ocean basins, and continental margins, the Indian Ocean contains several relatively shallow ridges and plateaus that are free from earthquake activity and whose origins remain enigmatic. One prominent example of these features is the Ninety East Ridge, a 2400-mi-long (4000-km), spectacularly linear feature

which may represent a scar left in the ocean floor in the wake of India's northward drift before it collided with the Asian mainland. Another major topographic plateau is the Mascarene Plateau. It lies to the northeast of Madagascar and includes the islands of the Seychelles and Mauritius. Although the origin of the entire plateau remains controversial, there is general agreement that the Seychelles Islands represent a microcontinental fragment that was left behind when the major continents were dispersed. Lying well to the south is the Kerguelen Plateau, whose origin is also unknown. The available data, however, suggest that at least the northern portion of the plateau was formed by submarine volcanism and uplift within a previously existing ocean basin.

The northeastern boundary of the Indian Ocean is clearly defined by an active arc-trench system lying south of Sumatra and Java. This trench system marks the locus of the destruction of Indian Ocean crust that is being transported northward (part of the Australian-Indian plate) as the result of active sea-floor spreading along the Southeast Indian Ocean Ridge.

Distribution of sediments. The distribution of sediments within the Indian Ocean is controlled primarily by three factors: the age of the underlying sea floor, the proximity of the continents, and the barrier effects of the major ridges that are so prevalent throughout the Indian Ocean. Thick sedimentary deposits characterize those parts of the ocean lying near the continental margins. In particular, the major drainage basins of eastern and western India bring immense quantities of sediment from the Himalayas to the northern Indian Ocean floor via the Ganges and Indus rivers. Hence, great thicknesses of sediment are found in the Bay of Bengal and in the Arabian Sea, where accumulations exceed several miles.

The active mid-ocean ridge systems, because of their relative youth, elevation, and great distance from terrigenous sediment sources, mark the sites of little or no sediment accumulation.

No drill holes and very few near-surface sediment samples exist for the high southern latitudes between about 0° longitude and 100°E longitude. Samples from other parts of the circum-Antarctic Ocean region indicate the dominance of sediments composed of the remains of silicic organisms. Therefore, it is expected that similar sediments characterize the extreme Southwest Indian Ocean. As the Antarctic continental margin is approached, significant quantities of terrigenous materials derived from the continent are also expected.

There has been much seismic research activity focused on the Indian Ocean segment of the Antarctic continental margins. While these studies are designed primarily to obtain more information about the early geologic history of the region, they are also of great interest to various national groups who would like to better assess the hydrocarbon potential of the offshore Antarctic.

Magnetic reversals. The study of magnetic reversals and the lineation stripes formed in association with the spreading mid-ocean ridge system has provided the primary tool for dating the underlying crust in the Indian Ocean. These data are also used to determine the time at which the various continental fragments were split away from Antarctica, and the rates and directions they migrated. The magnetic lineation data are rather convincing, and their interpretations are generally accepted for the southeast and northwest branches of the Indian Ocean mid-ocean system. However, the section of the mid-ocean ridge lying between Africa and Antarctica, the Southwest Indian Ridge, is characterized by numerous closely spaced offsets. This fact, coupled with a paucity of data, leaves much of the Southwest Indian Ocean still poorly understood in terms of its sea-floor spreading history. *See* PALEOMAGNETISM.

Drilling programs. An extensive program of drilling carried out as part of the Deep Sea Drilling Project (DSDP) (Fig. 7) confirmed the predicted ages of the sea floor obtained from the analysis of marine magnetic anomalies. They also contributed to present understanding of the plate tectonic evolution, the paleoenvironment, and the paleocirculation patterns of the Southeast Indian Ocean.

Fig. 8. Geotectonic image of the Indian Ocean as reflected in the satellite-derived gravity field of the region. A few specific structural features are identified as follows: Spreading centers: A, Carlsberg Ridge; B, Southwest Indian Ridge; C, Southeast Indian Ridge. Fracture zones: D, Indian-Antarctic; E, Ninety East Ridge (tentatively identified as a fracture zone). Plateaus: F, Kerguelen; G, Chagos-Lacadive; H, Mascarene; J, Agulhas. Zone of intraplate deformation, K. (*W. Haxby*, *Lamont-Doherty Geological Observatory of Columbia University*)

History. Unfortunately, major portions of the high-latitude regions of the Indian Ocean, particularly the region between Africa and Antarctica, remain poorly explored. Although there is general agreement about the reconstruction of major parts of the continental masses surrounding the Indian Ocean to a prerifted Gondwanaland configuration, the history prior to about 1×10^8 years remains controversial.

It has been postulated that the fragmentation of the Gondwanaland continent, especially the final departure of Australia from Antarctica in the early Cenozoic, allowed the formation of a major oceanic circulation pattern around the continent of Antarctica. Paleomagnetic evidence clearly indicates that the Antarctic continent resided in a near-polar position long before it became glaciated; hence an additional factor or condition was needed to cause glaciation. The removal of major topographic barriers by sea-floor spreading allowed the circumpolar current pattern to form and led to complex processes of latitudinal heat exchange. This may have provided the triggering mechanism that led to the eventual glaciation of the Antarctic continent about 2.5×10^7 years ago. *See* CENOZOIC.

Mapping the tectonic features of the ocean basins is possible by utilizing minute variations in the average height of the sea surface as measured from satellites; thus, a much more detailed map of the ocean gravity field can be deduced. The gravity field in turn reflects the subtle lateral density variations caused by the sea-floor topography and also by tectonic structures that are buried. The satellite observations have thus led to a synthetic representation of sea-floor relief, including fracture zone "traces," which reveal the pathways along which the Gondwanaland continents drifted apart (**Fig. 8**). The potential for reconstructing the details of the formation of the Indian Ocean sea floor has been greatly enhanced.

Geotectonic imagery has shown that the trends of the fracture zones in the northwest Indian Ocean are not parallel to the Ninety East Ridge (Fig. 7), thus making it unlikely that this ridge represents the trace of an old fracture zone, a previously popular hypothesis. Also, in the area south of India but north of the mid-ocean ridge system, both the satellite data and other geophysical data collected from ships indicate that the Indian plate is undergoing major compressional deformation within its boundaries. Continued deformation there could eventually lead to the formation of a new collision plate boundary and the partition of the Indian plate into two plates. It is thus clear that a knowledge of the detailed tectonic history of the region can have an important impact on such far-ranging topics as climate evolution, ocean circulation, and mineral resource predictions. *See* PALEOCEANOGRAPHY.

Dennis E. Hayes

Bibliography. S. Chandra, B. Arunachalam, and V. Suryanarayan (eds.), *The Indian Ocean and Its Islands: Strategic, Scientific, and Historical Perspectives*, 1993; D. N. Desai (ed.), *Oceanography of the Indian Ocean*, 1992; C. A. Geller, J. K. Weissel, and R. N. Anderson, Heat transfer and intraplate deformation in the central Indian Ocean, *J. Geophys. Res.*, 88(B2):1018–1032, 1983; S. Hastenrath, P. J. Lamb, and L. L. Greischer, *The Climatic Atlas of the Indian Ocean*, 3 vols., 1989; W. F. Haxby, Geotectonic imagery: Quantum leap in mapping power, *Lamont Newsl.*, Fall 1982; K. Wyrtki, *Oceanographic Atlas of the International Indian Ocean Expedition*, 1989.

Indium

A chemical element, In, atomic number 49, a member of group 13 and the fifth period of the periodic table. Indium has a relative atomic weight of 114.82. Indium occurs in the Earth's crust to the extent of about 0.000001% and is normally found in concentrations of 0.1% or less. It is widely distributed in many ores and minerals but is largely recovered from the flue dusts and residues of zinc-processing operations. *See* PERIODIC TABLE.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Indium is used in soldering for electronics and as a component of the intermetallic semiconductor used for germanium transistors. Indium arsenide, antimonide, and phosphide are semiconductors with unique properties. Other uses of indium are sleeve-type bearings to reduce corrosion and wear, glass-sealing alloys, and dental alloys. *See* GERMANIUM.

Edwin M. Larsen

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; A. J. Downs (ed.), *Chemistry of Aluminium, Gallium, Indium and Thallium*, 1993; W. F. Gale and T. C. Totemeier, *Smithells Metals Reference Book*, 8th ed., 2004; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; K. T. Tsen (ed.), *Ultrafast Physical Processes in Semiconductors: Semiconductors and Semimetals*, vol. 67, 2000.

Indole

The parent compound of a group of organic heterocyclic compounds containing the indole nucleus, which is a benzene ring fused to a pyrrole ring as in indole itself. Indole can exist in two tautomeric forms, the more stable enamine form and the

3-H-indole or imine form. Unsubstituted 3-H-indoles (sometimes called indolenines) and a structural isomer of indole, isoindole, are not stable, but have been shown to be reaction intermediates. They are isolable when properly substituted. *See* PYRROLE.

The importance of the indole ring lies in its presence in a large number of naturally occurring compounds. Some of these are the plant growth hormone indole-3-acetic acid, the animal tissue constitutent serotonin, the amino acid tryptophan, the pigment indigo, and the antitumor antibiotic mitomycin A. In addition, the indole nucleus is found in the indole alkaloids and in a number of synthetic drugs and pigments.

Physical properties. Indole is a steam-volatile, colorless solid, melting point 52.5°C (126°F), boiling point 253°C (487°F). It is found in small amounts in coal tar, feces, and flower oils. Indole has a characteristic ultraviolet spectrum with absorption maxima at 216, 266, 287, and 276 nanometers. It fluoresces (shines) when irradiated with ultraviolet light. These, and other spectral properties, vary widely for the many derivatives of the enamine form.

Preparations. Indoles are generally synthesized by constructing a five-membered nitrogen ring on a preformed benzene ring. The Fischer synthesis is the most common and versatile of many known methods. In this method, the phenylhydrazone of some ketone is treated with acid to give the substituted indole. The R groups in the ketone may be varied widely, and the phenylhydrazone ring may be substituted. Other syntheses include the Bischler synthesis from anilines and haloketones and the ring closure of *o*-toluides.

Chemical properties. Despite the presence of nitrogen, indole is not basic in the sense that it dissolves in aqueous acid or turns litmus blue. In fact, the hydrogen on the nitrogen is about as acidic as an aliphatic alcohol hydrogen (pK_a = about 17). Indole is an aromatic compound and undergoes electrophilic substitutions much like benzene, although it is much more reactive than benzene. Its reactivity is comparable to that of phenol, and it undergoes a number of reactions similar to those of phenol. Much of the chemistry of indole can be summarized mechanistically by reaction scheme. Essentially, indole is an efficient provider of electrons which can react with an electrophile (E^+) to yield a substituted product, or which can take part in the displacement of a nucleophile (Nu^-) from carbon to yield an alkylated product. When the electrophile is a polarized double bond such as a carbonyl or an imine, addition reactions take place.

Alkylation and acylation are more facile when the nucleophilic character of the indole is enhanced by converting it into a metallic salt, using the acidic proton on nitrogen. Thus, the nitrogen Grignard reagent of indole can be prepared from and any simple Grignard reagent. The metallic salt or complex is then alkylated or acylated by a nucleophilic displacement. Both the N- and the 3-substituted products are formed.

In the third type of reaction, indole reacts with aldehydes or ketones, with imines, and with activated alkenes.

Indole itself reacts slowly with air and rapidly with most oxidizing agents to give intractable polymeric tars. However, substituted indoles undergo a number of smooth oxidations to give isolable products. Air oxidations appear to involve a peroxide intermediate formed from 1,2,3,4-tetrahydrocarbazole. The peroxide can then undergo a number of reactions. Indole can be reduced to give 2,3-dihydroindoles (indolines) or octahydro derivatives.

Indole alkaloids. The indole alkaloids are a large group of substances containing the indole nucleus and can be isolated from plants. They contain a number of physiologically active materials, such as strychnine, reserpine, some forms of curare poison, and the rye-fungus drug, ergot. The indole alkaloids are synthesized in tissue from the amino acid tryptophan by a series of complex enzyme-moderated processes. The processes involve the decarboxylation of to give tryptamine, which reacts with some aldehyde to give a compound containing the B-carboline ring system. When the R on the aldehyde is methyl, the simple B-carboline alkaloids, such as harman, result. When R is a more complex fragment, the more complex indole alkaloids such as strychnine and reserpine result. When tryptamine is combined with a five-carbon unit, a key component of ergot, lysergic acid, arises. When the acid group of lysergic acid is converted to the *N*,*N*-diethylamide, lysergic acid diethylamide (LSD;) is obtained. LSD is a synthetic, and not a naturally occurring, substance. There are many hundreds of other indole alkaloids. *See* ALKALOID; HETEROCYCLIC COMPOUNDS. James M. Bobbitt

Bibliography. Atta-ur-Rahman and A. Basha, *Indole Alkaloids*, 1998; J. A. Joule and K. Mills, *Heterocyclic Chemistry*, 4th ed., 2000; R. J. Sundberg, *Indoles*, 1996; K. P. C. Vollhardt and N. E. Schore, *Organic Chemistry*, 5th ed., 2006.

Inductance

That property of an electric circuit or of two neighboring circuits whereby an electromotive force is induced (by the process of electromagnetic induction) in one of the circuits by a change of current in either of them. The term inductance coil is sometimes used as a synonym for inductor, a device possessing the property of inductance. *See* ELECTROMAGNETIC INDUCTION; ELECTROMOTIVE FORCE (EMF); INDUCTOR.

Self-inductance. For a given coil, the ratio of the electromotive force of induction to the rate of change of current in the coil is called the self-inductance L of the coil, given in Eq. (1), where $\mathscr{E}$ is the elec-

$$L = -\frac{\mathscr{E}}{dI/dt} \tag{1}$$

tromotive force at any instant and dI/dt is the rate

of change of the current at that instant. The negative sign indicates that the induced electromotive force is opposite in direction to the current when the current is increasing (dI/dt positive) and in the same direction as the current when the current is decreasing (dI/dt negative). The self-inductance is in henrys when the electromotive force is in volts, and the rate of change of current is in amperes per second.

An alternative definition of self-inductance is the number of flux linkages per unit current. Flux linkage is the product of the flux Φ and the number of turns in the coil N. Then Eq. (2) holds. Both sides of

$$L = \frac{N\Phi}{I} \tag{2}$$

Eq. (2) may be multiplied by I to obtain Eq. (3), which may be differentiated with respect to t, as in Eqs. (4).

$$LI = N\Phi \tag{3}$$

$$\begin{aligned} L\frac{dI}{dt} &= N\frac{d\Phi}{dt} = -\mathscr{E} \\ L &= -\frac{\mathscr{E}}{dI/dt} \end{aligned} \tag{4}$$

Hence the second definition is equivalent to the first.

Self-inductance does not affect a circuit in which the current is unchanging; however, it is of great importance when there is a changing current, since there is an induced emf during the time that the change takes place. For example, in an alternating-current circuit, the current is constantly changing and the inductance is an important factor. Also, in transient phenomena at the beginning or end of a steady unidirectional current, the self-inductance plays a part. *See* ELECTRIC TRANSIENT.

Consider a circuit of resistance R and inductance L connected in series to a constant source of potential difference V. The current in the circuit does not reach a final steady value instantly, but rises toward the final value $I = V/R$ in a manner that depends upon R and L. At every instant after the switch is closed the applied potential difference is the sum of the iR drop in potential and the back emf $L\,di/dt$, as in Eq. (5), where i is the instantaneous value of the current. Separating the variables i and t, one obtains Eq. (6). The solution of Eq. (6) is given in Eq. (7).

$$V = iR + L\frac{di}{dt} \tag{5}$$

$$\frac{di}{(V/R) - i} = \frac{R}{L}\,dt \tag{6}$$

$$i = \frac{V}{R}(1 - e^{-(R/T)t}) \tag{7}$$

The current rises exponentially to a final steady value V/R. The rate of growth is rapid at first, then less and less rapid as the current approaches the final value.

The power p supplied to the circuit at every instant during the rise of current is given by Eq. (8).

$$p = iV = i^2R + Li\frac{di}{dt} \tag{8}$$

The first term i^2R is the power that goes into heating the circuit. The second term $Li\,di/dt$ is the power that goes into building up the magnetic field in the inductor. The total energy W used in building up the magnetic field is given by Eq. (9). The energy used in

$$\begin{aligned} W &= \int_0^t p\,dt = \int_0^t Li\,\frac{di}{dt}\,dt \\ &= \int_0^I Li\,di = {}^1/_2LI^2 \end{aligned} \tag{9}$$

building up the magnetic field remains as energy of the magnetic field. When the switch is opened, the magnetic field collapses and the energy of the field is returned to the circuit, resulting in an induced emf. The arc that is often seen when a switch is opened is a result of this emf, and the energy to maintain the arc is supplied by the decreasing magnetic field.

Mutual inductance. The mutual inductance M of two neighboring circuits A and B is defined as the ratio of the emf induced in one circuit $\mathscr{E}$ to the rate of change of current in the other circuit, as in Eq. (10).

$$M = \frac{\mathscr{E}_B}{-(dI/dt)_A} \tag{10}$$

The SI unit of mutual inductance is the henry, the same as the unit of self-inductance. The same value is obtained for a pair of coils, regardless of which coil is taken as the starting point.

The mutual inductance of two circuits may also be expressed as the ratio of the flux linkages produced in circuit B by the current in circuit A to the current in circuit A. If Φ_{BA} is the flux threading B as a result of the current in circuit A, Eqs. (11) hold. Integration leads to Eq. (12).

$$\mathscr{E}_B = -N_B\frac{d\Phi_{BA}}{dt} = -M\frac{dI_A}{dt} \tag{11}$$

or

$$N_B\,d\Phi_{BA} = M\,DI_A$$

$$M = \frac{N_B\Phi_{BA}}{I_A} \tag{12}$$

See INDUCTANCE MEASUREMENT. Kenneth V. Manning

Bibliography. B. I. Bleaney and B. Bleaney, *Electricity and Magnetism,* 3d ed., 1976, reprint 1989; E. M. Purcell, *Electricity and Magnetism,* 2d ed., 1985.

Inductance measurement

The measurement of self- or mutual inductance. An electrical reactance such as the angular frequency ($2\pi f$, where f is the frequency) times self- or

mutual inductance is the ratio of the alternating voltage having the appropriate phase, which appears across specified terminals, to the current through the device. Commercial instruments often measure inductance from this ratio by comparing it with the voltage-to-current ratio associated with a noninductive resistor. *See* ELECTRICAL IMPEDANCE.

Precautions. Some practical precautions must be taken if accurate results are to be obtained. Any magnetic field associated with the inductor must not interact significantly with magnetic or conducting material in the vicinity of the inductor, inasmuch as the field, and therefore the inductance, would be altered. The varying magnetic field of an inductor will induce eddy currents in any nearby conducting material, which will in turn produce a magnetic field which interacts with the inductor and measuring system. Errors in a measurement of inductance may also arise from the interaction of the magnetic field of an inductor with the rest of the measuring circuit. Capacitance to other parts or to the surroundings of an inductor arising from its associated electric field will inevitably affect the impedance or apparent inductance of an inductor by a frequency-dependent amount but capacitive currents associated with screening of the measuring circuit can be routed in such a way as not to affect the measurement. *See* EDDY CURRENT; INDUCTOR.

If the magnetic circuit of an inductor includes magnetic material whose permeability depends on its previous magnetic history, or the magnetic flux caused by a direct current flowing simultaneously in the coil, its inductance will also be current- or history-dependent, and these conditions must be specified if the measurement is to be meaningful.

The electrical property of self- or mutual inductance is only defined for complete circuits. Since a measuring device or network forms part of the complete circuit when it is connected to an inductor to perform a measurement, it is necessary to ensure that either the inductance associated with the measuring circuit is negligible or that the measured quantity is defined as the change in inductance occurring when the unknown is replaced by a short circuit. The former procedure is usual for mutual inductors, and the latter for self-inductors.

Phase defect. In a mutual inductor the changing magnetic flux resulting from a varying current i in one coil induces a varying voltage V, 90° out of phase, in another coil. A perfect mutual inductor would have negligible energy losses associated with the interactions between the changing electric and magnetic fields of the device and nearby resistive conductors, including the conductors of the coils themselves. The energy loss per cycle in a practical air-cored mutual inductor amounts to only a few parts per million of the total stored magnetic energy of the device. The effect of this loss is that the phase of the voltage induced in the secondary will lead that of the current in the primary by slightly less than 90°. The departure from 90° is called the phase defect of the device. It is usual to define a configuration for a mutual inductor in which one specified end of the primary coil is connected to a specified end of the secondary coil; the effect of a phase defect can then be represented by an equivalent circuit (**Fig. 1**), in which the primary and secondary share a common fictitious path of small resistance r. This resistance has nothing to do with the actual resistance of the primary coil, which is irrelevant as far as the mutual inductance of the device is concerned.

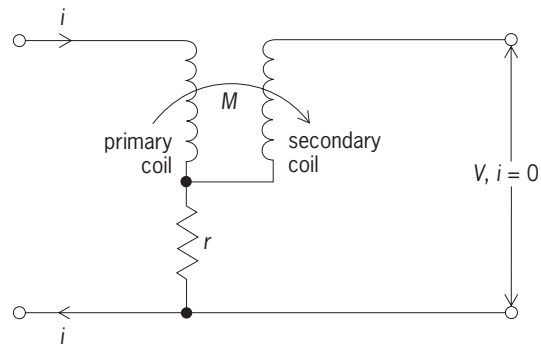

Fig. 1. Equivalent circuit representing loss in a mutual inductor.

In a self-inductor the functions of primary and secondary are combined in one coil, and therefore the voltage across the coil includes the potential drop of the current as it flows through the actual resistance of the coil. Self-inductors therefore have large, frequency-dependent phase defects, and this makes measuring them difficult. Their phase defect may be represented by an equivalent circuit in which a resistance R (approximately the actual resistance of the coil) is in series with an inductance L.

Comparing mutual inductors and deriving their values in terms of other electrical standards requires different methods from those used for self-inductors. The measurement of each will be discussed separately, taking account of the above considerations.

Measuring self-inductance. Capacitance and resistance standards have replaced inductance standards as the primary standards of impedance, and a scale of self-inductance is now established by finding the value of a self-inductor in terms of the reactance of a known capacitance C. For example, if these components are connected in parallel so that the combination appears to be purely resistive when measured in a suitable bridge at a frequency f, then Eq. (1) is satisfied.

$$2\pi f L' = \frac{1}{2\pi f C} \tag{1}$$

See CAPACITANCE MEASUREMENT; RESISTANCE MEASUREMENT.

This procedure measures the inductance parameter associated with a parallel equivalent circuit in which a resistance R' is in parallel with the inductance of value L'. The conversion between this representation and the more usual series equivalent circuit is given by Eqs. (2), where $D = 2\pi f L'/R' = R/2\pi f L$

$$L = \frac{L'}{1 + D^2} \qquad R = \frac{R' D^2}{1 + D^2} \tag{2}$$

is the reciprocal of the Q factor of the inductor.

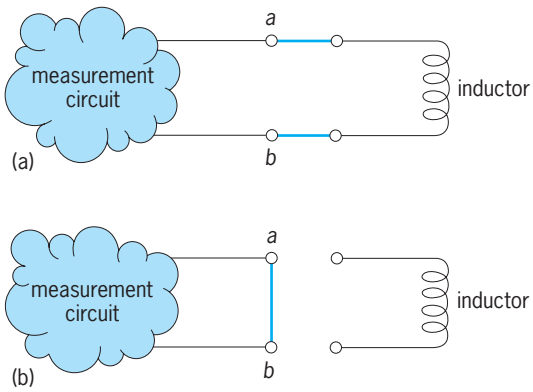

Fig. 2. Definition of the value of a self-inductor. (*a*) Circuit with self-inductor. (*b*) Circuit with self-inductor replaced by shorting link between terminals *a* and *b*

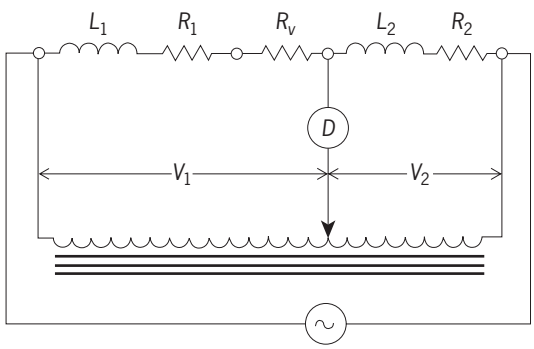

Fig. 3. Comparison of self-inductor using an inductive voltage divider.

Self-inductance should be measured as the change obtained by breaking the circuit at terminals *a* and *b* (**Fig. 2***a*) and substituting a shorting link for the inductor.

Self-inductors can be compared in a bridge in which the ratio of voltages V_1/V_2 across them is equated to the voltage ratio of a tapped autotransformer, otherwise called an inductive voltage divider (**Fig. 3**). Adjusting a noninductive resistance R_v in series with the inductor having the higher value of the ratio L/R enables the magnitude and phase of these voltage ratios to be equated when the detector is nulled. Then Eqs. (3) are satisfied.

$$L_1 = \frac{V_1}{V_2} L_2 \qquad R_1 = \frac{V_1}{V_2} R_2 - R_v \tag{3}$$

See INDUCTIVE VOLTAGE DIVIDER.

If the inductors have any significant external magnetic field, interaction of the one with the other or with the measuring circuit can be avoided by placing the inductors far apart with their coil axes in perpendicular directions, and by constructing the measuring circuit, including the leads to the inductors, as twisted conductor pairs or as coaxial cables whose inner and outer conductors carry equal and opposite currents.

The effect on the measurement of interaction via alternating magnetic fields emanating from a measuring instrument and the leads to the inductor can be eliminated by reversing the connections to the inductor, repeating the measurement, and taking the mean of the two results. Additionally, in a bridge of the type described above which compares an unknown inductor with a standard, the effect of any coupling between the inductors can be eliminated by also reversing the connections to the standard. If the reversal of both inductors is repeated, the mean of the resulting four measurements will eliminate all unwanted magnetic couplings.

Inaccuracy of a measuring instrument might be due, among other possibilities, to insufficient rejection by the measuring instrument of the resistive component of the inductor's impedance. This can be tested by connecting a resistor of value comparable with that of the inductor and of negligible inductance in series or parallel with the inductor to be measured. Having paid due regard to whether a series or parallel equivalent circuit representing the inductor is assumed by the instrument and any necessary consequent conversion made using Eq. (2), the inductance value given by the instrument ought to be unaffected by the presence of the resistor.

Measuring mutual inductance. If losses and capacitative effects between windings and to other parts of the circuit are neglected, mutual inductance is linearly additive. That is, the voltages induced in two or more secondary coils of a mutual inductor can be added simply by connecting the windings in series. Moreover, a mutual inductor is inherently a four-terminal device having separate current and potential terminals. Its impedance is defined as the ratio of the voltage across the potential terminals to the angular frequency times the current flowing into and out of the current terminals. These two considerations make measuring and comparing mutual inductances easier and more exact than comparing self-inductances, but it is still important that no significant magnetic fields other than those of the devices being compared are produced by the measuring circuits, and that the field of one device does not interact with another.

Two mutual inductors M_1 and M_2 can be compared by using a circuit (**Fig. 4**), which again uses an inductive voltage divider as a device of known voltage ratio.

The same current can be made to flow in the primary windings of two mutual inductors by

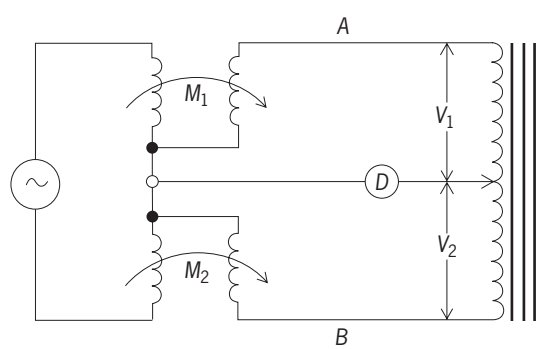

Fig. 4. Comparison of mutual inductors using an inductive voltage divider.

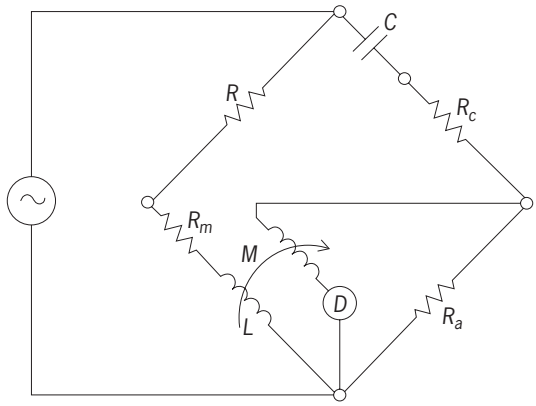

Fig. 5. Diagram of Heydweiller bridge.

connecting them in series across the bridge source. The ratio of the two voltages induced in the secondary coils can then be compared with the voltage ratio of the inductive voltage divider. Should the input impedance of the inductive voltage divider be insufficiently large so that the secondary coils are loaded by it to an unacceptable extent, it will be necessary to supply the energizing current for the inductive voltage divider directly from the source in such a way that the defining condition of negligible current in the conductors at A and B is realized. Then the ratio of the inductors is given by Eq. (4)

$$\frac{M_1}{M_2} = \frac{V_1}{V_2} \tag{4}$$

to an accuracy of the order of $(r/2\pi fM)^2$, where r represents the loss of a mutual inductor M as in the equivalent circuit of Fig. 1. This accuracy will usually be sufficient for all ordinary purposes.

The Heydweiller or Carey-Foster bridge (**Fig. 5**) is a convenient classical four-armed bridge for mutual inductance; L and R_m are respectively the inductance and resistance of the current-carrying windings of the mutual inductor. As can be seen, the source and detector share a common terminal, which simplifies the problem of using shields to prevent stray capacitive impedances from affecting the measurement.

If the parameter r which represents the phase defect of a mutual inductor can be neglected, the bridge balance equations are Eqs. (5) and (6).

$$M = R_a(R_m + R)C \tag{5}$$

$$L = (R_c + R_a)(R_m + R)C \tag{6}$$

R_m must be known, or separately measured, but R can often be chosen to be much larger than R_m, so that no great accuracy for R_m is required.

A mutual inductor can also be measured by connecting the primary and secondary coils in series in the same and opposing senses, and measuring the respective total self-inductances L_1 and L_2. The mutual inductance M can then be calculated from Eq. (7).

$$M = \frac{L_1 - L_2}{4} \tag{7}$$

Shielding of capacitive currents. In all the circuits described above, appropriate electrical screens should surround the conductors and components of the measuring circuit. These screens should be maintained at the potential of the detector terminals by connecting them to the output of a suitably adjusted potential divider connected across the source. Capacitive currents from all parts of the circuit are then intercepted and diverted back to the source so that they do not flow through the detector and influence the bridge balance.

A further improvement in technique is to construct the measurement system with fully screened components interconnected with coaxial cables. There are methods to ensure that the currents in the coaxial inner conductors are balanced by equal and opposite currents in the coaxial outer conductors. In this way unwanted magnetic interactions between the various parts of the network can also be avoided. *See* COAXIAL CABLE; ELECTRICAL MEASUREMENTS; ELECTRICAL SHIELDING; INDUCTANCE.

Bryan P. Kibble

Bibliography. B. Hague and T. R. Foord, *Alternating Current Bridge Methods*, 6th ed., 1971; B. P. Kibble and G. H. Rayner, *Coaxial A.C. Bridges*, 1984; L. Schnell (ed.), *Technology of Electrical Measurements*, 1993.

Induction coil

A device for producing a high-voltage alternating current or high-voltage pulses from a low-voltage direct current. The largest modern use of the induction coil is in the ignition system of internal combustion engines, such as automobile engines. Devices of similar construction, known as vibrators, are used as rectifiers and synchronous inverters. *See* IGNITION SYSTEM.

Figure 1 shows a typical circuit diagram for an induction coil. The primary coil, wound on the iron core, consists of only a few turns. The secondary coil, wound over the primary, consists of a large number of turns.

When the switch S is closed, the iron core becomes magnetized and therefore attracts the armature A. This automatically breaks the circuit to the coil through contact B and the armature. The armature is returned to its initial position by a spring and again makes contact with the contact B, restoring

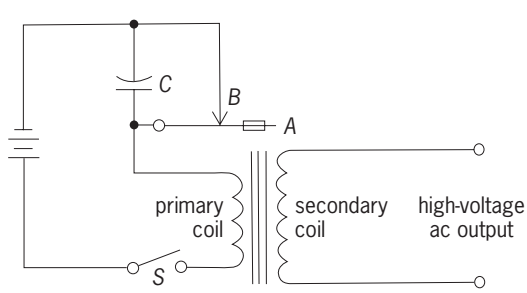

Fig. 1. Typical circuit for an induction coil.

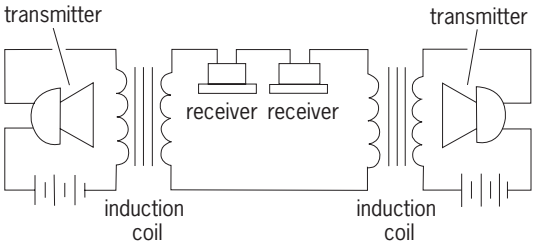

Fig. 2. Induction coils in telephone circuit.

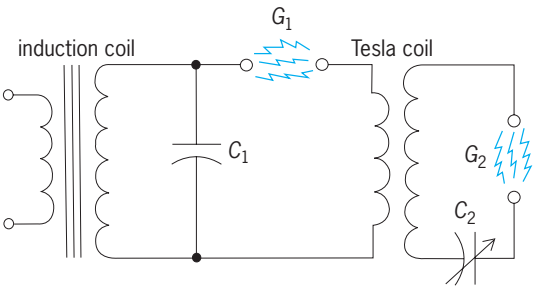

Fig. 3. Circuit diagram of Tesla coil.

the circuit to the primary coil. The cycle is then repeated rapidly.

While current is flowing in the primary coil, a magnetic field is produced. When the contact between A and B is broken, the magnetic field collapses and induces a high voltage in the secondary coil, similar to transformer action. The self-inductance of the coil must be limited; therefore, the core is a straight bundle of iron wires, which minimize eddy-current losses, rather than a closed iron circuit as is used in a transformer. *See* TRANSFORMER.

The capacitor C is placed across the breaker contacts to reduce the voltage across the contacts at the moment of their opening and thus reduce sparking. Sparking is caused by the induced voltage in the primary winding resulting from the collapsing magnetic field. The capacitor allows some of the energy of the magnetic field to be converted to electrostatic energy in the capacitor, rather than into heat at the contacts.

Induction coils of a different type are used in telephone circuits to step up the voltage from the transmitter and match the impedance of the line. The direct current in the circuit varies in magnitude at speech frequencies; therefore, no interrupter contacts are necessary. The battery and primary winding are connected in series with the transmitter as in **Fig. 2**. The secondary winding and the receiver are connected in series with the line. This circuitry reduces the required battery voltage.

Still another type of induction coil, called a reactor, is really a one-winding transformer designed to produce a definite voltage drop for a given current. *See* REACTOR (ELECTRICITY).

In 1892 Nicola Tesla used a form of induction coil to obtain currents of very high frequencies and high voltages. The oscillatory discharge of a Leyden jar was used as the interrupter. The terminals of the secondary of an induction coil are connected, one to the inner coating and the other to the outer coating of an insulated Leyden jar C_1 (**Fig. 3**). The circuit is completed through the primary winding of the Tesla coil, and the primary gap G_1. The coils and the Leyden jar act as a resonant circuit in the production of the high-frequency oscillation.

The primary of the Tesla coil consists of a half-dozen turns of wires wound on a nonmagnetic core. The secondary consists of many turns. The two coils are separated by air or oil as insulation. The alternation from the Leyden jar may have a frequency of several million hertz. Hence, the current induced in the secondary is not only of high voltage but also of high frequency.

Norman R. Bell

Induction heating

The heating of a nominally electrical conducting material by eddy currents induced by a varying electromagnetic field. The principle of the induction heating process is similar to that of a transformer. In **Fig. 1**, the inductor coil can be considered the primary winding of a transformer, with the workpiece as a single-turn secondary. When an alternating current flows in the primary coil, secondary currents will be induced in the workpiece. These induced currents are called eddy currents. The current flowing in the workpiece can be considered as the summation of all of the eddy currents.

In the design of conventional electrical apparatus, the losses due to induced eddy currents are minimized because they reduce the overall efficiency. However, in induction heating, their maximum effect is desired. Therefore close spacing is used between the inductor coil and the workpiece, and high-coil currents are used to obtain the maximum induced eddy currents and therefore high heating rates. *See* CORE LOSS; EDDY CURRENT.

Applications. Induction heating is widely employed in the metalworking industry for a variety of industrial processes. While carbon steel is by far the most common material heated, induction heating is also used with many other conducting materials such as various grades of stainless steel,

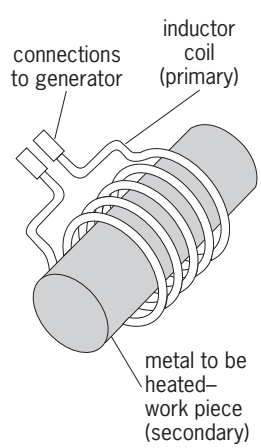

Fig. 1. Basic elements of induction heating.

aluminum, brass, copper, nickel, and titanium products. Induction heating is widely employed in (1) melting, holding, and superheating of ferrous and nonferrous metals in coreless and channel furnaces; (2) forging, forming, and rolling of slabs, billets, bars, and portions of bars; (3) heat treatment, such as hardening, annealing, and tempering applications; these processes are performed as continuous in-line systems or with automated handling of discrete parts to improve the metal's physical properties; (4) surface conditioning, such as curing of coatings, sintering, and semiconductor processing; (5) metal joining such as welding, brazing, and soldering. *See* BRAZING; HEAT TREATMENT (METALLURGY); SOLDERING; WELDING AND CUTTING OF MATERIALS.

The advantages of induction heating over the conventional processes (like fossil furnace or saltbath heating) are the following: (1) Heating is induced directly into the material. It is therefore an extremely rapid method of heating. It is not limited by the relatively slow rate of heat diffusion in conventional processes using surface-contact or radiant heating methods. (2) Because of skin effect, the heating is localized and the heated area is easily controlled by the shape and size of the inductor coil. (3) Induction heating is easily controllable, resulting in uniform high quality of the product. (4) It lends itself to automation, in-line processing, and automatic-process cycle control. (5) Startup time is short, and standby losses are low or nonexistent. (6) Working conditions are better because of the absence of noise, fumes, and radiated heat.

Heating process. The induced currents in the workpiece decrease exponentially from the surface toward the center of the material being heated. This current flowing through the material's own resistivity generates heat. The depth of heat penetration is defined as the depth to which 87% of the heat is generated. It is proportional to the square root of the material's resistivity ρ divided by the square root of its permeability μ and frequency f of the ac power applied to the inductor ($\delta_W = \sqrt{\rho/\mu f}$). The effective depth of heat penetration is therefore greater for lower frequencies and higher-resistivity materials. In magnetic materials such as steel, the depth of current penetration is less below the Curie temperature (approximately 1350°F or 730°C where the steel is magnetic) than it is above the Curie temperature. *See* CURIE TEMPERATURE; SKIN EFFECT (ELECTRICITY).

The rate at which thermal energy is required (P_{th}) can easily be calculated by the formula below.

$$P_{th} = MC\Delta T \times 10^{-3}\ \text{kW}$$
$$= 17.6M'C'\Delta T' \times 10^{-3}\ \text{kW}$$

M = kilograms of product heated per second
C = specific heat, (J/kg)/°C
ΔT = temperature rise in °C
M' = pounds of product heated per minute
C' = specific heat, (Btu/lb)/°F
$\Delta T'$ = temperature rise in °F

Frequencies used in induction heating

Frequency	Power source	Application
50/60 Hz 180 Hz–3 kHz	Line Solid-state	Melting
50/60 Hz 180 Hz–10 kHz	Line Solid-state	Forging, forming, rolling
1–10 kHz 200–500 kHz	Solid-state Vacuum tube oscillators	Heat treatment and surface conditioning
10–50 kHz 200–500 kHz	Solid-state Vacuum tube oscillators	Metals joining

When the proper frequency is employed, the overall heating efficiency averages 75% for magnetic materials, and 30–50% for nonmagnetic materials. For efficient heating, the frequency employed must be high enough so that the depth of current penetration is less than one-third the diameter or cross section of the material being heated. The **table** shows the range of frequencies used for applications of the induction heating process.

Since the capital equipment cost (in dollars per kilowatt) increases with the frequency of the power supply, the lowest frequency that will properly heat the part is employed. Where high-volume production is involved, as in forging applications, line frequency and dual frequencies with separate induction coils are often employed to maximize production at minimum equipment costs. When the workpieces are small, it is necessary to use higher frequencies to efficiently heat the part. Likewise, higher frequencies must be used when it is

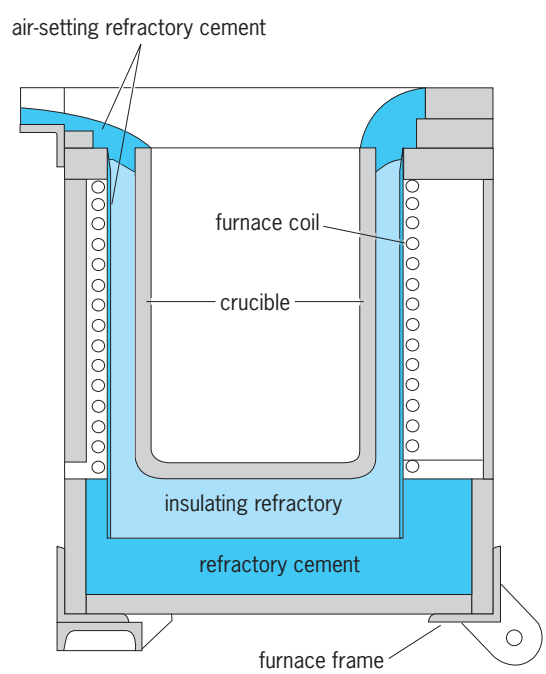

Fig. 2. Cross section of a typical induction melting furnace. (*Inductotherm Corp.*)

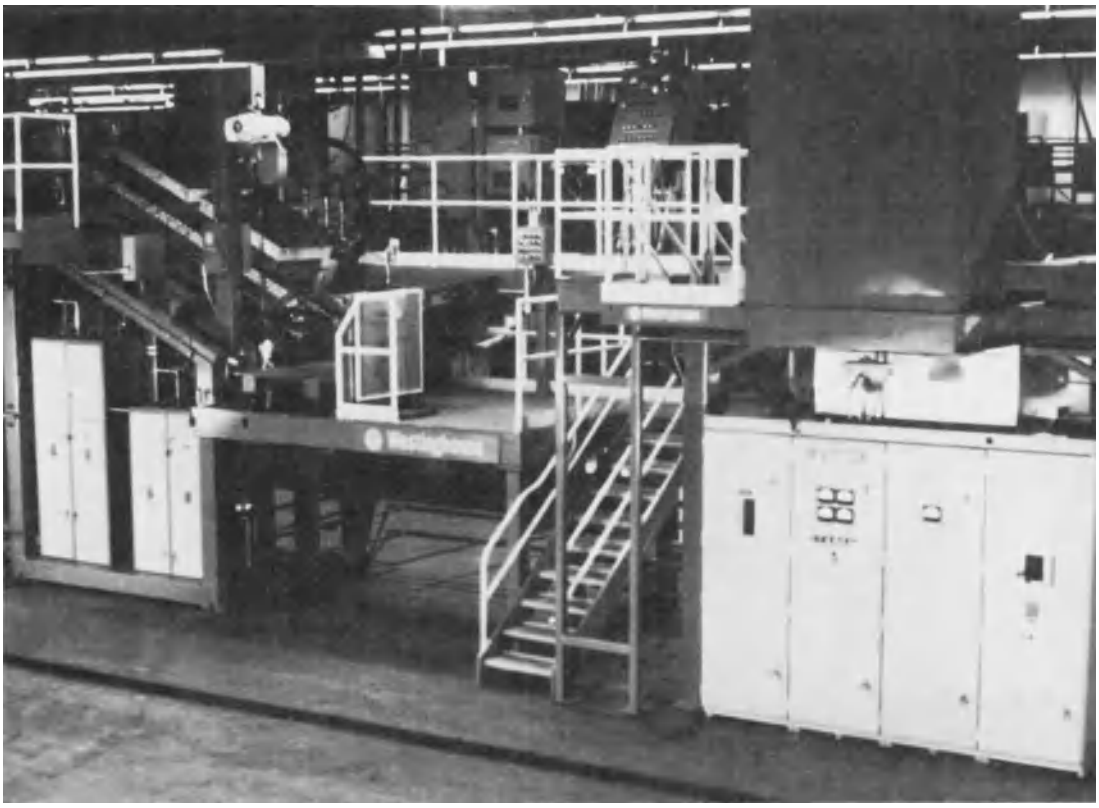

Fig. 3. A 1600-kW bar end heating system for feeding an automatic upsetter for forging flanges on automotive axles at a production rate of 720 per hour. (*Westinghouse Electric Corp.*)

necessary to concentrate the heat near the surface, as in surface hardening applications.

Power sources. The equipment used as power sources depends on the frequency required for the application. When line frequencies (generally 60 Hz) are used, suitable transformers, power factor correction capacitors, and control equipment are employed.

For higher frequencies, ranging up to 50,000 Hz, solid-state power supplies have replaced rotating equipment (motor generators). These power supplies utilize high-power thyristor solid-state devices to generate frequencies of 180 Hz to 50 kHz and ratings from 30 to 2000 kW. The conversion efficiencies of these power supplies average 90% or better as compared with the motor generator equipment with conversion efficiencies of only about 75%. *See* SEMICONDUCTOR RECTIFIER.

For frequencies over 50 kHz, power sources utilize vacuum tube oscillators to generate frequencies up to 4 MHz with power ratings from 1 to 800 kW. The conversion efficiencies of these power sources average only 60%, and therefore solid-state devices and equipment have been developed to replace them. However, the turn-off time characteristics and economics of silicon controlled rectifiers dictate the use of transistors, and therefore solid-state equipment is limited to low power ratings. *See* OSCILLATOR.

Process use. Induction heating is used for many processes as shown in the table. The construction of a typical melting furnace is shown in **Fig. 2**. A typical bar-heating system is shown in **Fig. 3**. Many different types of material-handling equipment such as scanners, indexing fixtures, conveyors, and lift and carry mechanisms are used with automated induction heating systems. *See* FURNACE CONSTRUCTION.

Induction heating differs from other methods in its high speed of heating and in its ability to generate heat from within the metal itself. In applications such as hardening, time and temperature affect the end product. Therefore slightly different heating cycles, temperatures, and materials are selected to provide the desired metallurgical properties in the end product. For other heating methods *see* ELECTRIC HEATING.

George F. Bobart

Bibliography. American Society of Metals, *Source Book on Heat Treating*, vol. 1, 1975; E. J. Davies and P. G. Simpson, *Induction Heating Handbook*, 1978; S. Zinn and S. L. Semiatin, *Elements of Induction Heating*, 1988.

Induction motor

An alternating-current motor in which the currents in the secondary winding (usually the rotor) are created solely by induction. These currents result from voltages induced in the secondary by the magnetic field of the primary winding (usually the stator). An induction motor operates slightly below synchronous speed and is sometimes called an asynchronous (meaning not synchronous) motor.

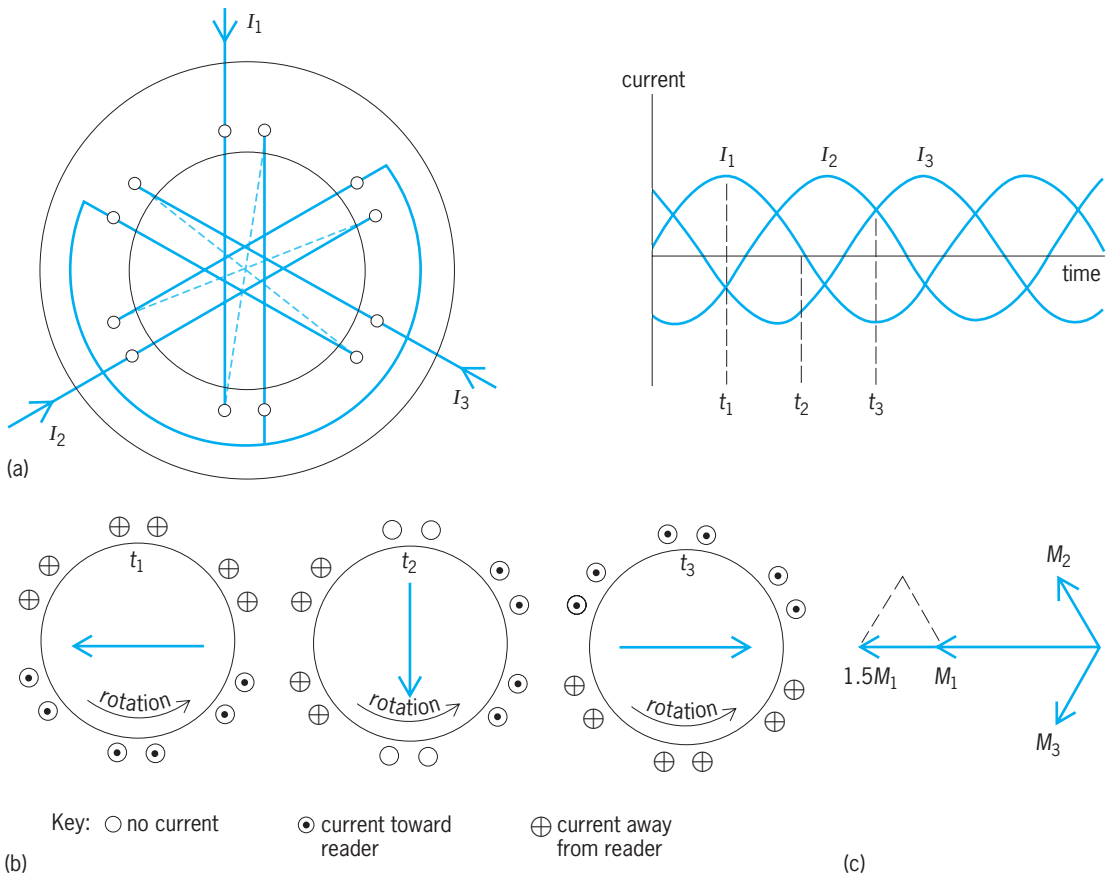

Fig. 1. Three-phase, two-pole, Y-connected stator of induction motor supplied with currents I_1, I_2, and I_3. (*a*) Stator windings and currents. (*b*) Rotating field. (*c*) Magnetomotive forces produced by stator winding.

Induction motors are the most commonly used electric motors because of their simple construction, efficiency, good speed regulation, and low cost. Polyphase induction motors come in all sizes and find wide use where polyphase power is available. Single-phase induction motors are found mainly in fractional-horsepower sizes (1 hp = 0.75 kW), and those up to 25 hp (19 kW) are used where only single-phase power is available.

Polyphase Induction Motors

There are two principal types of polyphase induction motors: squirrel-cage and wound-rotor machines. The differences in these machines is in the construction of the rotor. The stator construction is the same and is also identical to the stator of a synchronous motor. Both squirrel-cage and wound-rotor machines can be designed for two- or three-phase current.

Stator. The stator of a polyphase induction motor produces a rotating magnetic field when supplied with balanced, polyphase voltages (equal in magnitude and 90 electrical degrees apart for two-phase motors, 120 electrical degrees apart for three-phase motors). These voltages are supplied to phase windings, which are identical in all respects. The currents resulting from these voltages produce a magnetomotive force (mmf) of constant magnitude which rotates at synchronous speed. The speed is proportional to the frequency of the supply voltage and inversely proportional to the number of poles constructed on the stator.

Figure 1 is a simplified diagram of a three-phase, two-pole, Y-connected stator supplied with currents I_1, I_2, and I_3. Each stator winding produces a pulsating mmf which varies sinusoidally with time. The resultant mmf of the three windings (Fig. 1*c*) is constant in magnitude and rotates at synchronous speed. Figure 1*b* shows the direction of the mmf in the stator for times t_1, t_2, and t_3 shown in Fig. 1*a* and shows how the resultant mmf rotates. The synchronous speed N_s is shown by Eq. (1), where f is the frequency

$$N_s = \frac{120f}{p}\text{rpm} \tag{1}$$

in hertz and p is the number of stator poles. For any given frequency of operation, the synchronous speed is determined by the number of poles. For 60-Hz frequency, a two-pole motor has a synchronous speed of 3600 rpm; a four-pole motor, 1800 rpm; and so on. For details of stator windings *see* WINDINGS IN ELECTRIC MACHINERY

Squirrel-cage rotor. Figure 2 shows the bars, end rings, and cooling fins of a squirrel-cage rotor. The bars are skewed or angled to prevent cogging (operating below synchronous speed) and to reduce noise. The end rings provide paths for currents that result from the voltages induced in the rotor bars by the stator flux. The number of poles on a squirrel-cage

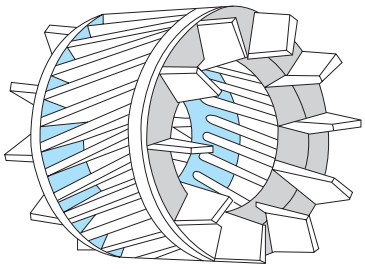

Fig. 2. Bars, end rings, and cooling fins of a squirrel-cage rotor.

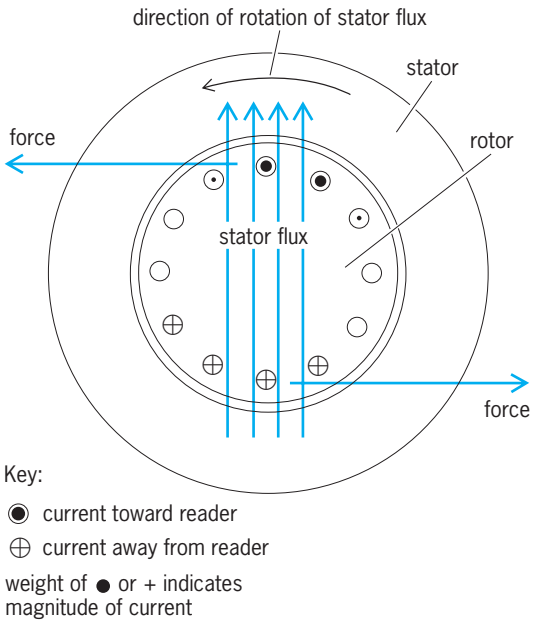

Fig. 3. Forces on the rotor winding.

rotor is always equal to the number of poles created by the stator winding.

Figure 3 shows how the two motor elements interact. A counterclockwise rotation of the stator flux causes voltages to be induced in the top bars of the rotor in an inward direction. These currents interact with the stator flux and produce a force on the rotor bars in the direction of the rotation of the stator flux.

When not driving a load, the rotor approaches synchronous speed N_s. At this speed there is no motion of the flux with respect to the rotor conductors. As a result, there is no voltage induced in the rotor and no rotor current flows. As load is applied, the rotor speed decreases slightly, causing an increase in rotor voltage and rotor current and a consequent increase in torque developed by the rotor. The reduction in speed is therefore sufficient to develop a torque equal and opposite to that of the load. Light loads require only slight reductions in speed; heavy loads require greater reduction. The difference between the synchronous speed N_s and the operating speed N is the slip speed. Slip s is conveniently expressed as a percentage of synchronous speed, as in Eq. (2).

$$s = \frac{N_s - N}{Ns} \times 100\% \tag{2}$$

When the rotor is stationary, a large voltage is induced in the rotor. The frequency of this rotor voltage is the same as that of the supply voltage. The frequency f_2 of rotor voltage at any speed is shown by Eq. (3), where f_1 is the frequency of the supply

$$f_2 = f_1 s \tag{3}$$

voltage and s is the slip expressed as a decimal. The voltage e_2 induced in the rotor at any speed is shown by Eq. (4), where e_{2s} is the rotor voltage at standstill.

$$e_2 = (e_{2s})s \tag{4}$$

The reactance x_2 of the rotor is a function of its standstill reactance x_{2s} and slip, as shown by Eq. (5). The

$$x_2 = (x_{2s})s \tag{5}$$

impedance of the rotor at any speed is determined by the reactance x_2 and the rotor resistance r_2. The rotor current i_2 is shown by Eq. (6). In the equation, for

$$\begin{aligned} i_2 &= \frac{e_2}{\sqrt{r_2^2 + x_2^2}} \\ &= \frac{(e_{2s})s}{\sqrt{r_2^2 + (x_{2s})^2 s^2}} = \frac{e_{2s}}{\sqrt{\left(\frac{r_2}{s}\right)^2 + (x_{2s})^2}} \end{aligned} \tag{6}$$

small values of slip, the rotor current is small and possesses a high power factor. When slip becomes large, the r_2/s term becomes small, current increases, and the current lags the voltage by a large phase angle. Standstill (or starting) current is large and lags the voltage by 50–70°. Only in-phase, or unity-power-factor, rotor currents are in space phase with the air-gap flux and can therefore produce torque. The current i_2 contains both a unity power-factor component i_p and a reactive component i_r. The maximum value of i_p and therefore maximum torque are obtained when slip is of the correct value to make r_2/s equal to x_{2s}. If the value of r_2 is changed, the slip at which maximum torque is developed must also change. If r_2 is doubled and s is doubled, the current i_2 is not changed and the torque is unchanged.

This feature provides a means of changing the speed-torque characteristics of the motor. In **Fig. 4**, curve 1 shows a typical characteristic curve of an

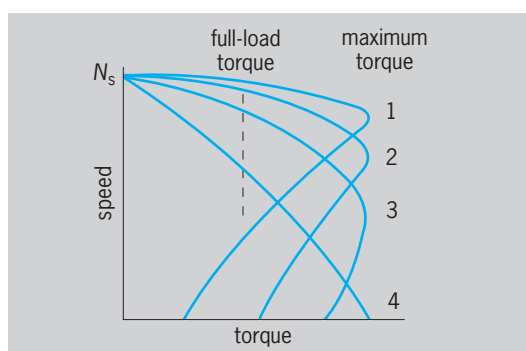

Fig. 4. Speed-torque characteristic of polyphase induction motor.

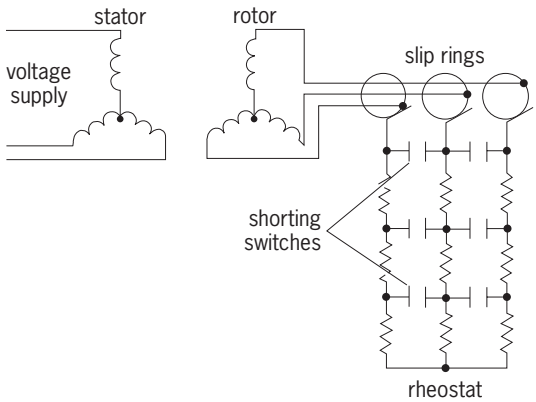

Fig. 5. Connections of wound-rotor induction motor.

induction motor. If the resistance of the rotor bars were doubled without making any other changes in the motor, it would develop the characteristic of curve 2, which shows twice the slip of curve 1 for any given torque. Further increases in the rotor resistance could result in curve 3. When r_2 is made equal to x_{2s}, maximum torque will be developed at standstill as in curve 4. These curves show that higher resistance rotors give higher starting torque. However, since the motor's normal operating range is on the upper portion of the curve, the curves also show that a higher-resistance rotor results in more variation in speed from no load to full load (or poorer speed regulation) than the low-resistance rotor. Higher-resistance rotors also reduce motor efficiency. Except for their characteristic low starting torque, low-resistance rotors would be desirable for most applications.

Wound rotor. A wound-rotor induction motor can provide both high starting torque and good speed regulation. This is accomplished by adding external resistance to the rotor circuit during starting and removing the resistance after speed is attained.

The wound rotor has a polyphase winding similar to the stator winding and must be wound for the same number of poles. Voltages are induced in these windings just as they are in the squirrel-cage rotor bars. The windings are connected to slip rings so that connections may be made to external impedances, usually resistors, to limit starting currents, improve power factor, or control speed.

Figure 5 shows the connection of a rheostat used to bring a wound-rotor motor up to speed. The rheostat limits the starting current drawn from the supply to a value less than that required by a squirrel-cage motor. The resistance is gradually reduced to bring the motor up to speed. By leaving various portions of the starting resistances in the circuit, some degree of speed control can be obtained, as in Fig. 4. However, this method of speed control is inherently inefficient and converts the motor into a variable-speed motor, rather than an essentially constant-speed motor. For other means of controlling speed of polyphase induction motors and for other types of ac motors *see* ALTERNATING-CURRENT MOTOR

Single-Phase Induction Motors

Single-phase induction motors display poorer operating characteristics than polyphase machines, but are used where polyphase voltages are not available. They are most common in small sizes ($^1/_2$ hp or 0.4 kW or less) in domestic and industrial applications. Their particular disadvantages are low power factor, low efficiency, and the need for special starting devices.

The rotor of a single-phase induction motor is of the squirrel-cage type. The stator has a main winding which produces a pulsating field. At standstill, the pulsating field cannot produce rotor currents that will act on the air-gap flux to produce rotor torque. However, once the rotor is turning, it produces a cross flux at right angles in both space and time with the main field and thereby produces a rotating field comparable to that produced by the stator of a two-phase motor.

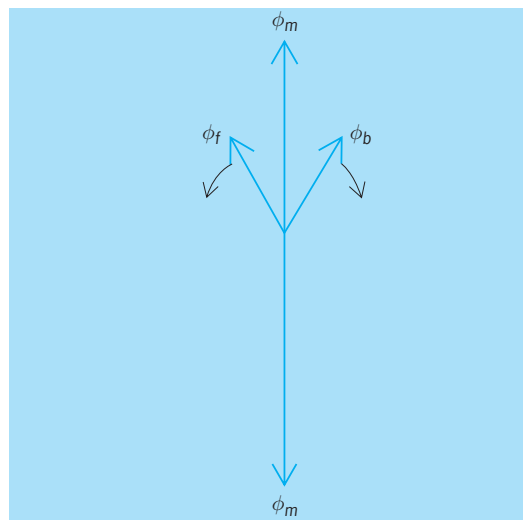

Fig. 6. Fluxes associated with the single-phase induction motor.

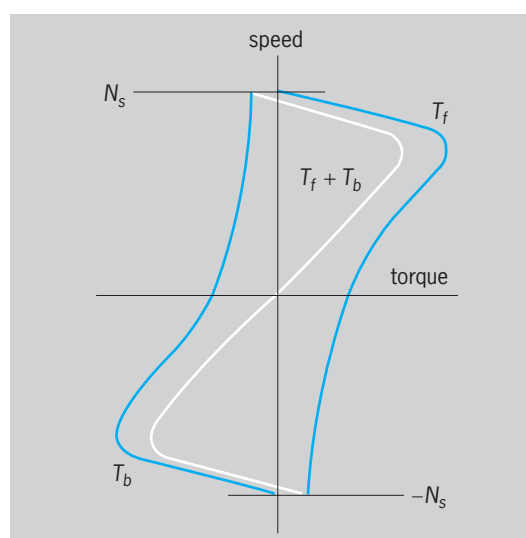

Fig. 7. Torques produced in the single-phase induction motor.

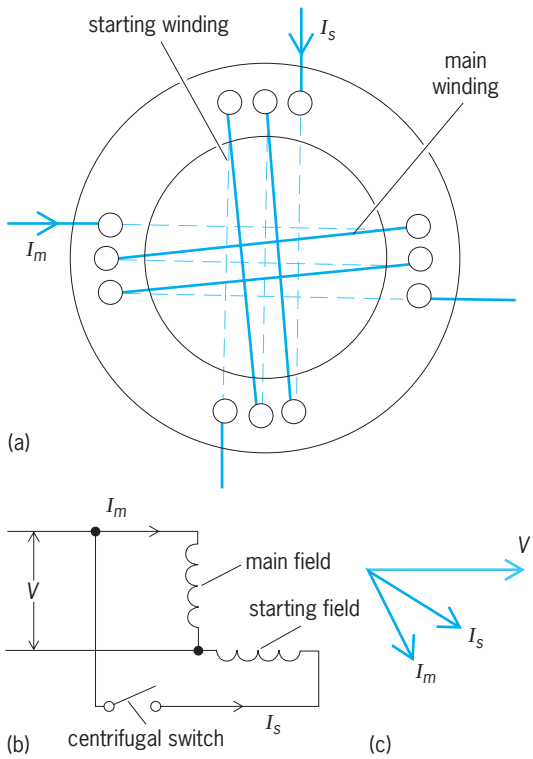

Fig. 8. Split-phase motor. (*a*) Windings. (*b*) Winding connections. (*c*) Vector diagram.

An explanation of this is based on the concept that a pulsating field is the equivalent of two oppositely rotating fields of one-half the magnitude of the resultant pulsating field. In **Fig. 6**, ϕ_m is the maximum value of the stator flux ϕ, which is shown only by its two components ϕ_f and ϕ_b, which represent the two oppositely rotating fields of constant equal magnitudes of $\phi_m/2$. Each component ϕ_f and ϕ_b produces a torque T_f and T_b on the rotor. **Figure 7** shows that the sum of these torques is zero when speed is zero. However, if started, the sum of the torques is not zero and rotation will be maintained by the resultant torque.

This machine has good performance at high speed. However, to make this motor useful, it must have some way of producing a starting torque. The method by which this starting torque is obtained designates the type of the single-phase induction motor.

Split-phase motor. This motor has two stator windings, the customary main winding and a starting winding located 90 electrical degrees from the main winding, as in **Fig. 8***a*. The starting winding has fewer turns of smaller wire, to give a higher resistance-to-reactance ratio, than the main winding. Therefore their currents I_m (main winding) and I_s (starting winding) are out of time phase, as in Fig. 8*c*, when the windings are supplied by a common voltage V. These currents produce an elliptical field (equivalent to a uniform rotating field superimposed on a pulsating field) which causes a unidirectional torque at standstill. This torque will start the motor. When sufficient speed has been attained, the circuit of the starting winding can be opened by a centrifugal switch and the motor will operate with a characteristic illustrated by the broken-line curve of Fig. 7.

Capacitor motor. The stator windings of this motor are similar to the split-phase motor. However, the starting winding is connected to the supply through a capacitor (**Fig. 9***a*). This results in a starting winding current which leads the applied voltage. The motor then has winding currents at standstill which are nearly 90° apart in time, as well as 90° apart in space. High starting torque and high power factor are therefore obtained. The starting winding circuit can be opened by a centrifugal switch when the motor comes up to speed. A typical characteristic is shown in Fig. 9*c*.

In some motors two capacitors are used. When the motor is first connected to the voltage supply, the two capacitors are used in parallel in the starting circuit. At higher speed one capacitor is removed by a centrifugal switch, leaving the other in series with the starting winding. This motor has high starting torque and good power factor.

Shaded-pole motor. This motor is used extensively where large power and large starting torque are not required, as in fans. A squirrel-cage rotor is used with a salient-pole stator excited by the ac supply. Each salient pole is slotted so that a portion of the pole face can be encircled by a short-circuited winding, or shading coil.

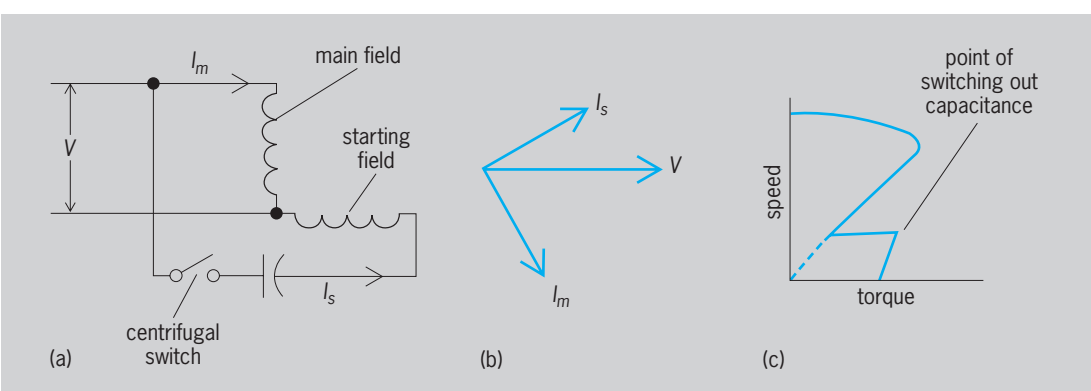

Fig. 9. Capacitor motor. (*a*) Winding connections. (*b*) Vector diagram. (*c*) Characteristic.

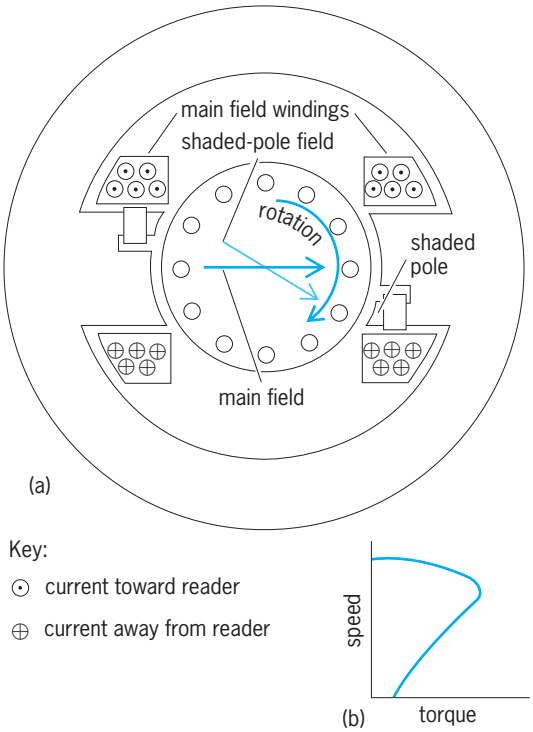

Fig. 10. Shaded-pole motor. (*a*) Cross-sectional view. (*b*) Typical characteristic.

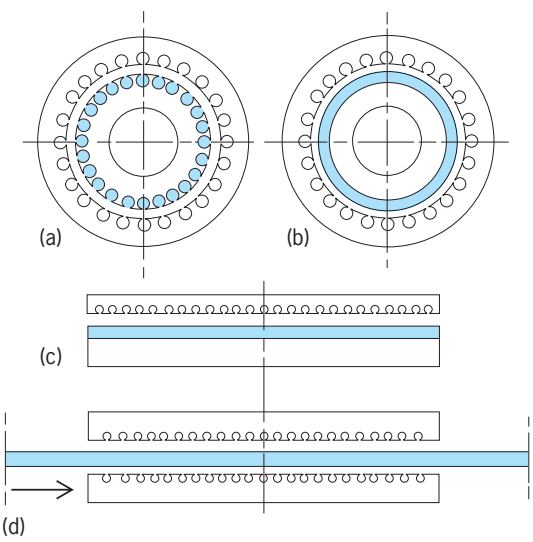

Fig. 11. Evolution of linear induction motor. (*a*) Polyphase squirrel-cage induction motor. (*b*) Conduction sheet motor. (*c*) One-sided long-secondary linear motor with short-flux return yoke. (*d*) Double-sided long-secondary linear motor.

The main winding produces a field between the poles as in **Fig. 10**. The shading coils act to delay the flux passing through them, so that it lags the flux in the unshaded portions. This gives a sweeping magnetic action across the pole face, and consequently across the rotor bars opposite the pole face, and results in a torque on the rotor. This torque is much smaller than the torque of a split-phase motor, but it is adequate for many operations. A typical characteristic of the motor is shown in Fig. 10*b*. *See* HYSTERESIS MOTOR; RELUCTANCE MOTOR; REPULSION MOTOR.

Linear motor. Figure 11 illustrates the arrangements of the elements of the polyphase squirrel-cage induction motor. The squirrel cage (secondary) is embedded in the rotor in a manner to provide a close magnetic coupling with the stator winding (primary). This arrangement provides a small air gap between the stator and the rotor. If the squirrel cage is replaced by a conducting sheet as in Fig. 11*b*, motor action can be obtained. This machine, though inferior to that of Fig. 11*a*, will function as a motor. If the stator windings and iron are unrolled (rectangular laminations instead of circular laminations), the arrangement of the elements will take a form shown in Fig. 11*c*, and the field produced by polyphase excitation of the primary winding will travel in a linear direction instead of a circular direction. This field will produce a force on the conducting sheet that is in the plane of the sheet and at right angles with the stator conductors. A reversal of the phase rotation of the primary voltages will reverse the direction of motion of the air-gap flux and thereby reverse the force on the secondary sheet. No load on the motor corresponds to the condition when the secondary sheet is moving at the same speed as the field produced by the primary. For an arrangement of Fig. 11*c*, there is a magnetic attraction between the iron of the stator and the iron of the secondary sheet. For some applications, this can be a serious disadvantage of the one-sided motor of Fig. 11*c*. This disadvantage can be eliminated by use of the double-sided arrangement of Fig. 11*d*. Here the primary iron of the upper and lower sides is held together rigidly, and the forces that are normal to the plane of the sheet do not act on the sheet.

Applications. In conventional transportation systems, traction effort is dependent on the contact of the wheels with the ground. In some cases, locomotives must be provided with heavy weights to keep the wheels from sliding when under heavy loads. This disadvantage can be eliminated by use of the linear motor. With the linear motor with air cushions, friction loss can be reduced and skidding can be eliminated. **Figure 12** illustrates the application of the double-sided linear motor for high-speed transportation.

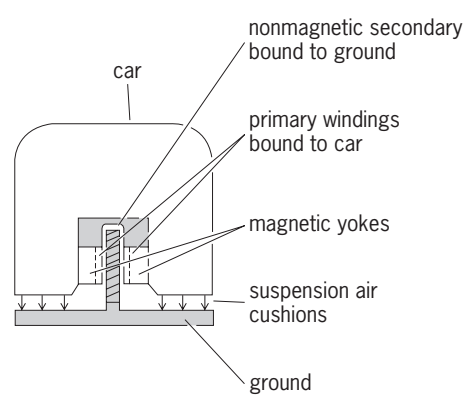

Fig. 12. Drawing of linear induction-motor configuration. Magnetic interaction supplies the forces between the car and the ground. (*Le Moteur Linéaire and Société de l'Aérotrain*)

Conveying systems that are operated in limited space have been driven with linear motors. Some of these, ranging from $^1/_2$ to 1 mi (0.8 to 1.6 km) in length, have worked successfully.

Because of the large effective air gap of the linear motor, its magnetizing current is larger than that of the conventional motor. Its efficiency is somewhat lower, and its cost is high. The linear motor is largely in the experimental stage. Albert G. Conrad

Bibliography. P. L. Alger, *Induction Machines: Their Behavior and Use*, 2d ed., 1970, revised 1995; S. J. Chapman, *Electric Machinery Fundamentals*, 3d ed., 1998; P. L. Cochran, *Polyphase Induction Motors*, 1989; A. E. Fitzgerald, C. Kingsley and S. D. Umans, *Electric Machinery*, 5th ed., 1990; J. F. Gieras, *Linear Induction Drives*, 1994; M. Poloujadoff, *The Theory of Linear Induction Machinery*, 1980.

Inductive coordination

The avoidance of inductive interference. Electric power systems, like almost everything run by electricity, depend on internal electric and magnetic fields; some of these fields find their way into the environment. The strongest of these fields can then induce voltages and currents in nearby devices and equipment and, in some cases, can interfere with the internal fields being used by electrical equipment in the vicinity. These induced voltages and currents, which are due to the coupling between the energized source and the electrical equipment, are called inductive interference. *See* ELECTRIC FIELD; ELECTROMAGNETIC INDUCTION.

In the process of the generation and transmission of large amounts of power, power systems necessarily utilize high currents and high voltages and, further, use overhead power lines that must be in proximity to the equipment they supply. This situation can cause potential inductive coordination problems. Avoiding such problems puts constraints on the design, operation, maintenance, and siting of the various electric power system components. *See* TRANSMISSION LINES.

Any current-carrying conductor (or device) generates electric and magnetic fields, which may radiate to the environment and interfere with nearby devices. Total shielding is sometimes possible, as with underground or overhead cables. In most cases, however, shielded cables are not practical and air-insulated overhead power lines are used. They create electric and magnetic fields that couple with objects in the immediate area, producing induced voltages and currents.

Figure 1 shows the most common example of electric and magnetic induction. Objects coupled with the power lines include the power lines themselves, telephone lines, fences, tracks, and other objects within the power-line right-of-way or corridor space. Coupling is stronger near the source. Induced voltages and currents are affected by many design parameters, such as conductor size, phase spacing, geometry line configuration, object size, distance from the source, line load currents, and operating conditions. Certain induced voltages and currents could reach unsafe values.

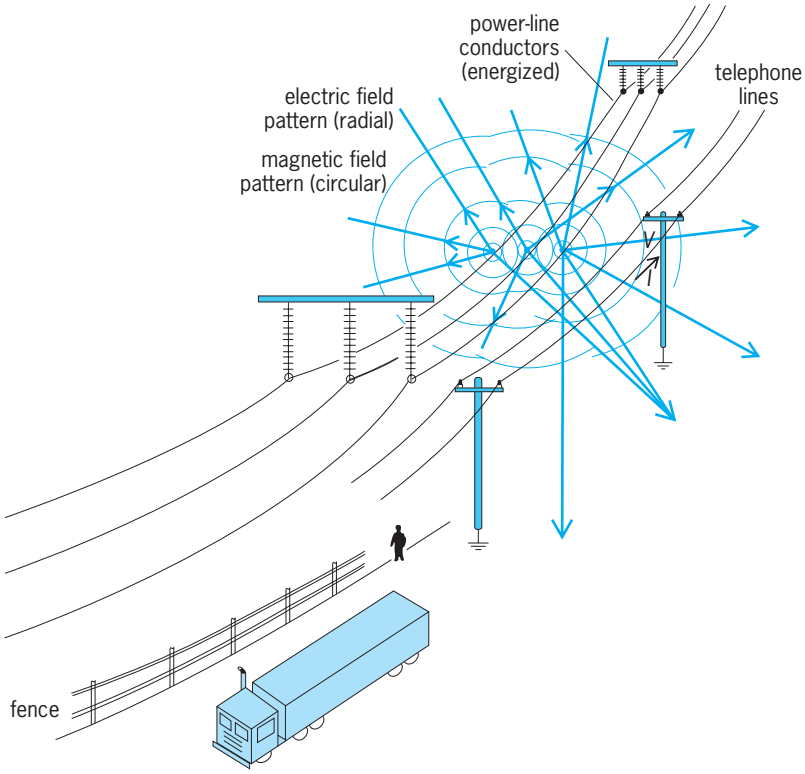

Fig. 1. Power-line field pattern and inductive coupling. Coupling produces induced voltages (*V*) and currents (*I*).

Design considerations. Inductive coupling can be reduced by proper configuration, circuit design, and siting. Generating stations and large switching stations built on remote sites and well shielded by enclosures do not produce significant stray fields. Well-shielded underground substations, manholes, cables, and distribution equipment also do not cause interference. Overhead power lines are, however, air-insulated, not shielded, and must of course be built in inhabited areas. These overhead lines cause practically all of the problems due to inductive coupling. For this reason and for safety considerations, power lines are restricted as far as possible to specific corridors or rights of way. Spacing between them and the requirements of their surroundings are considered and carefully calculated to minimize possible interference. These corridors are often shared by telephone lines, communication circuits, railroads, and sometimes trolley buses, each of which must be considered for possible inductive coupling.

Power-line inductive coordination. Modern telephone and communication circuits are well shielded and rarely encounter interference from nearby power lines. In some communication systems extensive transposition is used to reduce such interference. However, where a long parallel exposure exists, inductive coupling can be reduced by balancing the operation of the power line and by transposition of power and communication lines (**Fig. 2**).

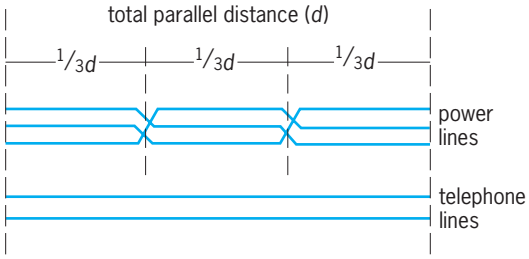

Fig. 2. Transposition: the induced voltages and currents are canceled by transposing the power line as shown.

Transposition of power lines reduces the electrical imbalances with significant reduction in induced voltages and currents in the exposed parallel telephone circuits. Fences, long irrigation pipes, and large ungrounded objects within the right of way may experience considerable inductive coupling and must be grounded for safety. *See* GROUNDING.

Special cases. In some cases, such as electric trains or unbalanced operating conditions, the ground must be used as a partial return conductor and transposition is impossible. The electric loop in such cases is composed of the wires supplying the train, the rail, and the earth. Other cases, such as solid-state devices in power conditioners, motor controllers for energy conservation, or variable-speed applications, have also created inductive coupling problems. In these cases, it would be desirable to isolate the power circuits geographically. This is usually not feasible, however, and other measures must be found, such as properly distributed capacitor banks within the power system, addition of specially designed filters, or employment of specifically designed, low-harmonic-content devices. *See* ELECTRICAL INTERFERENCE; ELECTRICAL SHIELDING. Aly A. Mahmoud

Bibliography. Electric Power Research Institute, *Transmission Line Reference Book—345 KV And above*, 2d ed., 1982; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; Institute of Electrical and Electronics Engineers, *IEEE Recommended Practice for Inductive Coordination of Electric Supply and Communication Lines*, rev. ed., 1993.

Inductive voltage divider

An autotransformer that has its winding divided into a number of equal-turn sections (usually 10) so that when an alternating voltage V is applied to the whole winding, the voltage across each section is nominally $V/10$. The progressive voltages from one end to the section junctions are thus $V/10$, $2V/10$, $3V/10$, . . . , and $9V/10$. This voltage division can be realized with errors considerably less than 1 part per million of V, and such units therefore find wide use as standards of ac voltage ratio in the discipline of electrical measurements. *See* TRANSFORMER.

The division of voltage will be in error if there are differences of resistance and leakage inductance from section to section, and these errors will be significant if the differences are significant in relation to the input impedance of the winding. Leakage inductance is caused by that very small fraction of the flux from one section's winding which fails to thread the rest of the windings. The most commonly used constructional technique for minimizing such errors is to take 10 equal lengths of insulated copper wire and twist them into a "rope." The rope is wound onto a toroidal core made of thin, high-permeability, low-iron-loss magnetic material. The strands of the winding are then connected in series so that each strand forms the winding of one section of the 10-section divider (**Fig. 1**). The resistances of the sections are very nearly equal since the strands are the same length, and the leakage inductances are also closely equal and small because of the close flux coupling of this type of winding. The low-reluctance magnetic path of the core ensures a very high value of input impedance. Thus, voltage division at low audio frequencies can be accurate to a few parts in 10^8 of V.

If an impedance is connected across one section of an inductive voltage divider, currents will flow (**Fig. 2**) so as to produce near-zero ampere-turns on the core. Because of the close coupling of the sectional windings, only a very small proportion of the total opposing fluxes fails to thread the other sections. Hence the ratio of output to input voltage of an inductive voltage divider with one section loaded is only slightly affected by the very small leakage inductance (typically a few microhenrys in a rope-type winding) plus the ohmic resistance of the wire (typically a few tens of milliohms). Thus a single-decade

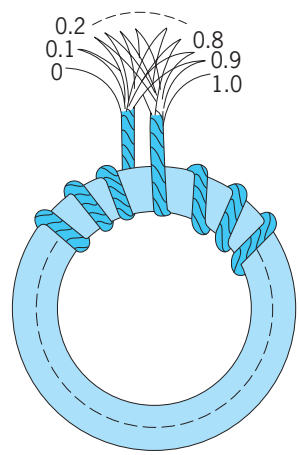

Fig. 1. Rope winding on toroidal core making single-decade inductive voltage divider.

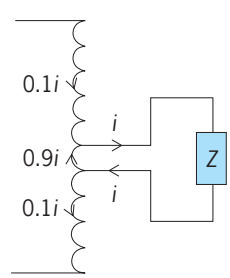

Fig. 2. Inductive voltage divider with an impedance *Z* across one section.

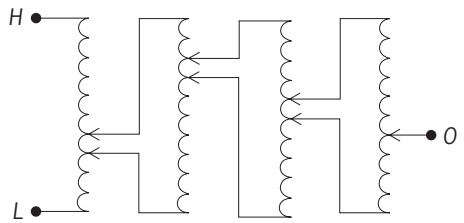

Fig. 3. Four-decade inductive voltage divider. Input is voltage between *H* and *L*; output is voltage between *O* and *L*.

inductive voltage divider has the characteristic of high-input impedance with low-output impedance, and because of this it is possible to connect units in tandem (**Fig. 3**) to provide a multidecade system, giving very fine resolution of the voltage division.

A common commercial application of inductive voltage dividers is as ratio arms in impedance bridges where one impedance is a transducer of a physical property (for example, temperature, force, or displacement) which is compared with a fixed reference impedance by using an inductive voltage divider as the variable ratio device. *See* INDUCTANCE MEASUREMENT.

Inductive voltage dividers operate most accurately in the frequency range 20–1592 Hz but can be constructed to operate at frequencies up to 1 MHz. They are usually designed to accept an input voltage of up to about 0.25 times the frequency in hertz. *See* VOLTAGE MEASUREMENT. T. A. Deacon; Bryan P. Kibble

Bibliography. T. A. Deacon and J. J. Hill, Two-stage inductive voltage dividers, *Proc. Inst. Elec. Eng.* (*UK*), 115(6):888–892, June 1968; J. J. Hill and T. A. Deacon, Theory, design and measurement of inductive voltage dividers, *Proc. Inst. Elec. Eng.* (*UK*), 115(5):727–735, May 1968.

Inductor

A device for introducing inductance into a circuit. The term covers devices with a wide range of uses, sizes, and types, including components for electric-wave filters, tuned circuits, electrical measuring circuits, and energy storage devices.

Inductors are classified as fixed, adjustable, and variable. All are made either with or without magnetic cores. Inductors without magnetic cores are called air-core coils, although the actual core material may be a ceramic, a plastic, or some other nonmagnetic material. Inductors with magnetic cores are called iron-core coils. A wide variety of magnetic materials are used, and some of these contain very little iron. Magnetic cores for inductors for low-frequency, or high-energy storage, use are most commonly made from laminations of silicon steel. Some iron-core inductors with cores of compressed powdered iron, powdered permalloy, or ferrite are more suitable for higher-frequency applications.

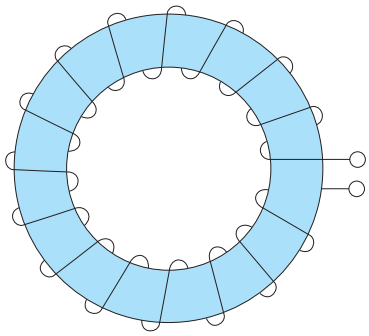

Toroidal coil, a type of fixed inductor.

Fixed inductors. In fixed inductors coils are wound so that the turns remain fixed in position with respect to each other. If an iron core is used, any air gap it has is also fixed and the position of the core remains unchanged within the coil.

A toroidal coil is a fixed inductor wound uniformly around a toroidal form (see **illus.**). Because of the closed magnetic circuit, such an inductor has practically no leakage flux and is little affected by the presence of stray magnetic fields. High-accuracy standard inductors are commonly made in this form. Powdered cores are used to increase the Q of the coil and reduce the size required for a specified inductance. Ceramic-core toroidal coils supported in cork are used as standard inductors of high stability and accuracy.

Adjustable inductors. These either have taps for changing the number of turns desired, or consist of several fixed inductors which may be switched into various series or parallel combinations.

Variable inductors. Such inductors are constructed so that the effective inductance can be changed. Means for doing this include (1) changing the permeability of a magnetic core; (2) moving the magnetic core, or part of it, with respect to the coil or the remainder of the core; and (3) moving one or more coils of the inductor with respect to one or more of the other coils, thereby changing mutual inductance. *See* ALTERNATING-CURRENT CIRCUIT THEORY; INDUCTANCE. Burtis L. Robertson; Wilson S. Pritchett

Industrial ecology

The multidisciplinary study of industrial and economic systems and their linkages with fundamental natural systems. Industrial ecology incorporates research involving energy supplies, materials, technologies, and technological systems; physical, biological, and social sciences; economics; law; and business management. Industrial and economic systems are viewed not in isolation but in their cultural and ecological context. Both demand-side (consumer) and supply-side (producer) activities are included, as are all sectors of economic activity, such as mining, agriculture, forestry, fisheries, manufacturing, and service activities. Industrial ecology also includes subsistence activities at the fringes of formal economic systems, which generate a number of important impacts on natural systems. Industrial ecology provides the understanding to support the reasoned improvement of the economic,

environmental, and social efficiency of current industrial systems.

History. There is an obvious appeal to thinking of industrial systems as analogous to natural ecological systems, especially when considering manufacturing activities and their associated environmental impacts. Ecological systems are characterized by a high degree of cycling of materials: the wastes from one organism become food for another. Until recently, however, most manufacturing activities were more linear: materials were used in processes and products only once, then became wastes with associated environmental impacts. Thus, the early development of industrial ecology reflected a desire to apply the cycling model characteristic of biological communities to manufacturing, with the residual materials from one process or product feeding into another. It was in this spirit that a few uses of the term "industrial ecology" began to occur in the 1970s and 1980s. In Japan, for example, the Ministry of Industry and International Trade explored such a concept in the early 1970s to suggest a model for structuring the country's industrial system.

In the late 1980s and early 1990s, industrial ecology began to evolve from an intuitively appealing metaphor to an actual field of study. Among the reasons was the recognition of the growing inability of existing environmental policy, focused on end-of-pipe emissions and remediation, and reductionist environmental science research, to address complex, systemic environmental perturbations such as global climate change, loss of biodiversity and habitat, and degradation of air, water, and soil resources. Moreover, private firms were being affected by rapidly increasing costs of environmental compliance and associated liabilities, combined with increasing public and customer pressure for enhanced environmental performance. Traditional compliance activities were not sufficient to respond to these pressures, especially in sectors such as heavy manufacturing and chemical production, which had significant environmental issues. The fundamental problem with the traditional approach is that it treated environmental issues as "overhead," that is, as ancillary to the principal activities of an individual, a firm, or, indeed, society as a whole. Industrial ecology became the framework within which environmental issues began to be treated as strategic, that is, as integral to the primary activities of firms and society.

Given these incentives, significant progress in both the theory of industrial ecology and its application through specific tools and methodologies soon followed. Among the methodologies which have become more sophisticated and accepted over this time are life-cycle assessment (LCA), developed initially in rigorous form by the Society of Environmental Toxicology and Chemistry. Life-cycle assessment is an increasingly standardized methodology which identifies the environmental impacts associated with the life cycle of a material in a specific application or, in some cases, a product, thus identifying opportunities for improvement in environmental performance. This improvement may be accomplished in a number of ways, including reductions of waste through better engineering of manufacturing processes, or substitution of less toxic materials in processes and products. Another industrial ecology application is through design for environment (DFE), developed initially by the United States electronics sector. Design for environment is a module of a broader concurrent engineering approach, design for X (DFX), where X is a desirable product characteristic such as manufacturability, safety, testability, or simplicity of materials and parts. Design for X methodologies are often used in the design and manufacture of complex articles such as computers or consumer electronics, so developing a design for environment approach made adoption of environmentally conscious manufacturing and product design easier in those manufacturing sectors. In general, life-cycle assessment methodologies are useful in evaluating a particular material in a specific application, while design for environment tends to emphasize the importance of artifact design choices and technologies, and is used more in sectors manufacturing complex articles such as electronic devices.

Industrial ecology is still a nascent field. While there are some textbooks and courses in industrial ecology, primarily in engineering schools, it still somewhat reflects its origins in manufacturing organizations. Thus, significant foundational work remains to be done to understand the industrial ecology of service economies in developed countries, on the one hand, and its application to subsistence economic activity in developing countries, on the other. Biomass sectors, such as forestry, agriculture, and fisheries, are increasingly beginning to adopt industrial ecology approaches, but in general lag behind the manufacturing sectors in doing so.

Conceptual framework. Industrial ecology can best be understood by placing it in a framework illustrating its relationships to concepts such as sustainable development, and methodologies such as design for environment and life-cycle assessment (**Fig. 1**). The highest level of the framework is the concept of sustainability, frequently framed in terms of sustainable development, or development that meets the needs of the current generation without compromising the ability of future generations to meet theirs. While many people find this a desirable concept, it is important to recognize that a number of potentially sustainable worlds can be defined, some more desirable than others. Thus, while a desirable sustainable state is important to provide a general goal toward which environmentally preferable activities can be directed, it cannot yet be defined with any specificity, or assumed as an outcome, especially in the absence of continued work in industrial ecology and related fields.

Evolution along at least three dimensions—the ethical dimension, the institutional dimension, and the science and technology dimension—is required if a desirable sustainable state is to be defined and achieved. Industrial ecology focuses on scientific and technological, not ethical or institutional, issues; it supports the evolution of the science and technology

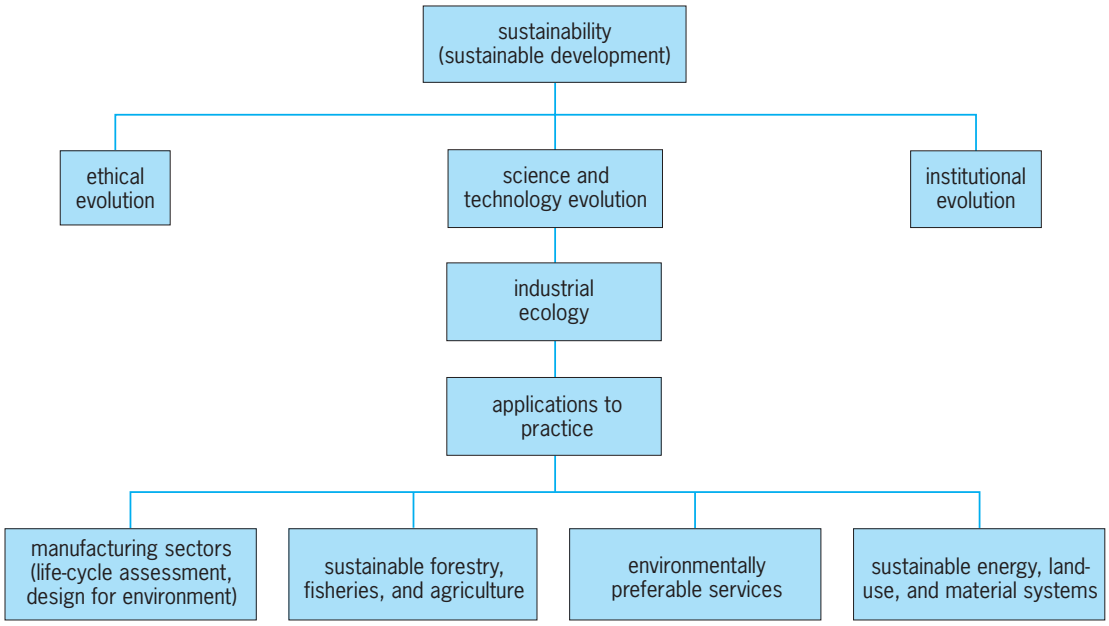

Fig. 1. Conceptual framework illustrating the relationship of industrial ecology to other concepts.

base necessary to enable progress toward sustainability. It does not replace existing disciplines and fields of study, such as engineering, environmental science, economics, or ecology; rather, it integrates them into a new and different understanding at a higher level (**Fig. 2**). In doing so, it links together results from existing, reductionist approaches, augmenting but not replacing such research. Thus, it supports a better understanding of the couplings between natural and human systems that exist in reality.

A number of different applications of, and methodologies based on, industrial ecology principles are being developed. They range from those appropriate to the firm level, such as life-cycle assessment and design for environment, to those that operate at a sector or social level, such as integrated materials management, environmentally preferable services, or sustainable fisheries. Further detail arises as each family of methodologies expands; the number of analytic tools and technologies being developed is rapidly increasing. Thus, for example, an important initiative in sustainable agriculture is integrated pest management, and intensive efforts are under way in the forestry industry to define the characteristics of sustainable forestry operations.

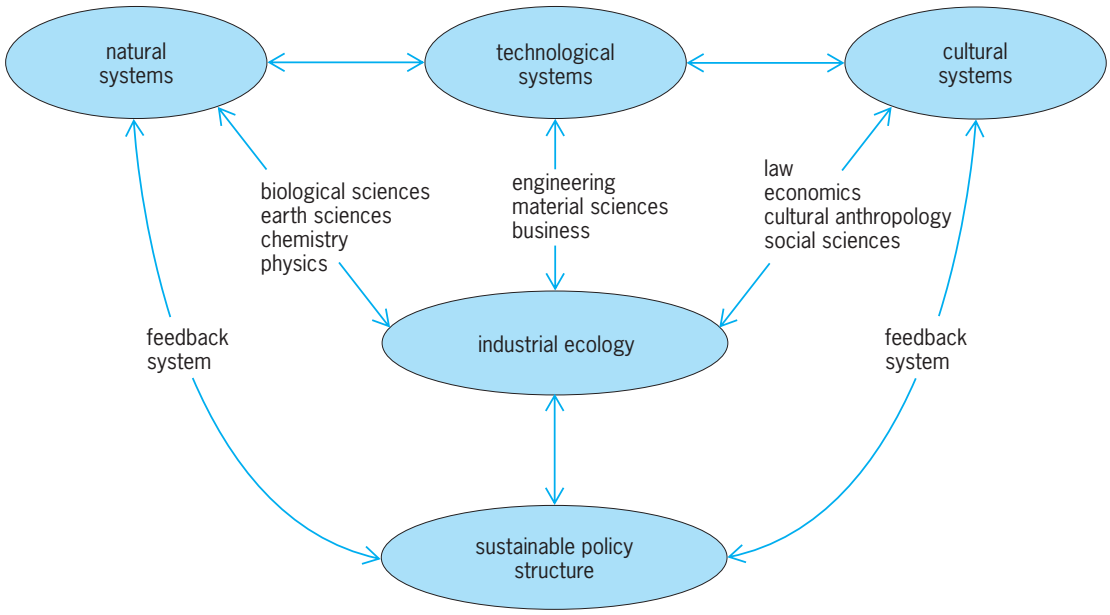

Fig. 2. Industrial ecology as an integrative multidisciplinary field. It pulls together results from existing fields of study to support an understanding of environmental perturbations in the context of complex human-natural systems.

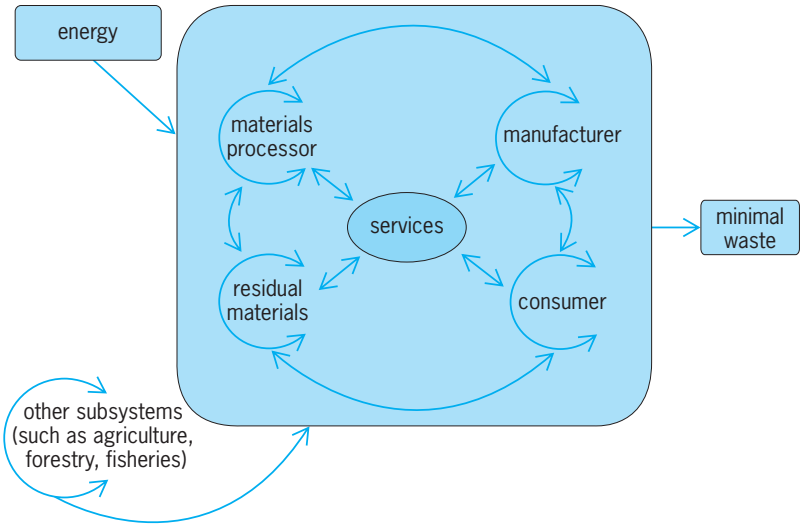

Fig. 3. Energetically open, low-waste economy. Materials are cycled through the economy at different scales and in different sectors to minimize overall waste production while quality of life is enhanced.

Principles. Some of the obvious principles of industrial ecology can be illustrated by analogy to biological communities. For example: (1) Creation of economic systems where material flows are reused rather than becoming waste is a critical element of industrial ecology (**Fig. 3**). (2) The focus is on systems and integrative analysis rather than specific elements. Key techniques which support such an approach are a focus on material stocks and flows, and energy consumption, throughout the system. (3) Industrial ecology is concerned not just with static analyses of systems, but with their responses over time, and particularly with their resilience (how well they maintain system coherence and function when stressed).

Five key concepts of industrial ecology are:

1. Design of products, processes, facilities, infrastructure, services, and technology systems so that they can be easily adapted to environmentally preferable innovation with minimal waste. Modular design of complex technologies is an example.

2. Minimization of waste production and energy consumption in all activities.

3. Use of the least toxic alternatives whenever possible, particularly when the materials may be dispersed into the environment. This principle would have argued against adding lead to gasoline, since predictably the pollutant subsequently would be dispersed throughout the environment.

4. Design of products, facilities, infrastructure, and technology systems to preserve the embedded utility of materials and energy used in initial manufacture. Thus, in many cases designs which extend the life of products and support the recycling of subassemblies or components, rather than materials, are preferable.

5. Design of physical products at all scales not just to perform their intended function but also to be used in creating other useful products at the end of their current life.

The field of industrial ecology is at present underdeveloped, and in many cases researchers are not yet even asking the right questions. The rapid development of methodologies and tools based on industrial ecology principles, the pressures for environmentally preferable behavior by firms and other institutions, and global agreements addressing complex environmental perturbations such as global climate change, however, indicate that continued growth of industrial ecology as a field of study is critical. *See* CLIMATE MODIFICATION; CONSERVATION OF RESOURCES; ECOLOGY; ECOLOGY, APPLIED; ENVIRONMENT; ENVIRONMENTAL ENGINEERING; HUMAN ECOLOGY; SYSTEMS ENGINEERING. Braden R. Allenby

Bibliography. B. R. Allenby, *Industrial Ecology: Policy Framework and Implementation*, Prentice Hall, Upper Saddle River, NJ, 1999; R. U. Ayres and U. E. Simonis (eds.), *Industrial Metabolism*, United Nations University Press, Tokyo, 1994; T. E. Graedel and B. R. Allenby, *Industrial Ecology*, Prentice Hall, Upper Saddle River, NJ, 1995; R. Socolow et al. (eds.), *Industrial Ecology and Global Change*, Cambridge University Press, 1994.

Industrial engineering

A branch of engineering dealing with the design, development, and implementation of integrated systems of humans, machines, and information resources to provide products and services. Industrial engineering encompasses specialized knowledge and skills in the physical, social, engineering, and management sciences, such as human and cognitive sciences, computer systems and information technologies, manufacturing processes, operations research, production, and automation. The industrial engineer integrates people into the design and development of systems, thus requiring an understanding of the physical, physiological, psychological, and other characteristics that govern and affect the performance of individuals and groups in working environments.

Functions and activities. Industrial engineering is a broad field compared to other engineering disciplines. The major activities of industrial engineering stem from manufacturing industries and include work methods analysis and improvement; work measurement and the establishment of standards; machine tool analysis and design; job and workplace design; plant layout and facility design; materials handling; cost reduction; production planning and scheduling; inventory control, maintenance, and replacement; statistical quality control; scheduling; assembly-line balancing, systems, and procedures; and overall productivity improvement. Computers and information systems have necessitated additional activities and functions, including numerically controlled machine installation and programming; manufacturing systems design; computer-aided design/computer-aided manufacturing, design of experiments, quality engineering, and statistical process control; computer simulation, operations research, and management science methods; computer applications, software development, and

information technology; human-factors engineering and ergonomics; systems design and integration; and robotics and automation.

The philosophy and motivation of the industrial engineering profession is to find the most efficient and effective methods, procedures, and processes for an operating system, and to seek continuous improvement. Thus, industrial engineering is utilized to help organizations grow and expand efficiently during periods of prosperity, and streamline costs and consolidate and reallocate resources during austere times. Industrial engineers, particularly those involved in manufacturing and related industries, work closely with management. Therefore, some understanding of organizational behavior, finance, management, and related business principles and practices is needed.

Functional areas. Industrial engineering can be classified into four areas: human factors and work systems, production and distribution systems, manufacturing systems, and operations research. This classification is in accordance with the various functions and activities in the design, development, and improvement of methods, operations, processes, and products.

Human factors and work systems. These combine the traditional areas of work measurement and design with the analysis and evaluation of the human element, which includes the physiological and psychological considerations in designing products and work systems. Technology continues to challenge the limitations and role of individuals and groups in production and service systems. Specifically, there is the challenge of matching capabilities such as the speed, touch, strength, and decision-making skills of humans with sophisticated computers and machine components. The influence and effects of noise, heat, humidity, lighting, and other environmental factors are dealt with in terms of the safety and welfare of the worker, as well as the evaluation of conditions for achieving optimum performance. Related issues deal with identifying and evaluating the information requirements and constraints associated with a system or design in order to achieve effective operation or service. *See* CONTROL SYSTEMS; HUMAN-FACTORS ENGINEERING; HUMAN-MACHINE SYSTEMS; INDUSTRIAL FACILITIES; INDUSTRIAL HEALTH AND SAFETY; METHODS ENGINEERING; MODEL THEORY; OPTIMIZATION; PRODUCTION METHODS; PRODUCTIVITY; SIMULATION; SYSTEMS ANALYSIS; SYSTEMS ENGINEERING; WORK MEASUREMENT; WORK STANDARDIZATION.

Production and distribution systems. These deal with the planning, scheduling, and flow of materials and information in production and manufacturing. This area includes the design, control, and implementation of production systems, including analysis procedures for determining where to locate people, machines, warehouses, and production centers; how to control and regulate the flow of parts and services to provide timely arrival of materials, acceptable quality, and customer satisfaction; scheduling and routing of operations; and maintenance and replacement procedures for ensuring cost effectiveness and overall productivity. *See* INVENTORY CONTROL; MATERIALS HANDLING; PRODUCTION PLANNING.

Manufacturing systems. These consist of organized configurations of machines, workstations, information, and processes to transform materials into products, which include parts, assemblies, and subassemblies. The processes include machining, forming, assembly, inspection and testing, and transfer. They can be achieved through various levels of automation and hierarchical control. Design and development issues involve the specification and allocation of cells, buffer storage, computer integration, degree of automation, and inspection for overall effectiveness in producing high-quality, low-cost products to meet customer expectations. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING; COMPUTER-INTEGRATED MANUFACTURING; FLEXIBLE MANUFACTURING SYSTEM.

Operations research. Operations research deals with the development and application of quantitative techniques and methods for analyzing and evaluating systems that aid in making engineering and management decisions. Management science, which is more often thought to be associated with business applications, is frequently used synonymously with operations research. This area includes the development and application of analytical, experimental, and simulation models for describing complex systems and prescribing conditions for efficient and optimal performance. Mathematical programming algorithms for allocating resources; methods for analyzing queueing and facility bottlenecking problems, inventory, and machine replacement decisions; and computer computation and implementation methods are among the issues dealt with in the field of operations research. *See* ARTIFICIAL INTELLIGENCE; OPERATIONS RESEARCH; QUEUEING THEORY.

Related subdisciplines. With its extensive breadth, industrial engineering has several subfields that overlap with neighboring disciplines, such as psychology, mechanical engineering, materials science, mathematics, computer science, business, and electrical engineering. Among the larger and more recognized subdisciplines are ergonomics, human factors, manufacturing engineering, operations research, operations management, statistics, and systems engineering. In many cases the extent of overlap is so significant, particularly in relation to research programs, that industrial engineering methodologies and techniques have been adopted and assimilated by other disciplines. *See* PRODUCTION ENGINEERING.

Marlin U. Thomas

Bibliography. A. B. Badiru (ed.), *Handbook of Industrial and Systems Engineering*, 2005; H. P. Emerson and D. C. E. Naehring, *Origins of Industrial Engineering*, 1988; M. P. Groover, *Automation, Production Systems, and Computer-Aided Manufacturing*, 2d ed., 2000; G. Salvendy (ed.), *Handbook of Industrial Engineering: Technology and Operations Management*, 3d ed., 2001; K. B. Zandin and H. B. Maynard, *Maynard's Industrial Engineering Handbook*, 5th ed., 2001.

Industrial health and safety

An interdisciplinary field that focuses on preventing occupational illnesses and injuries. The disciplines of engineering, epidemiology, toxicology, medicine, psychology, and sociology provide the methods for study and prevention. In the United States, prevention efforts have resulted in a reduction of accidental work deaths per year; nevertheless, the accident toll remains high.

Occupational Hazards

Tens of thousands of occupational hazards exist. Occupational hazards can be organized by separately considering plants and equipment, the physical work environment, hazards of materials, and task demands. Significant interactions occur between these categories. For example, equipment can modify the work environment by producing noise, potentially hazardous materials, or heat, but will be hazardous only if inappropriate procedures are followed.

Plants and equipment. Plant hazards are often associated with energy sources and power transmission, processes at the point of operation, vehicles and materials-handling systems, walking and climbing surfaces, ingress-egress, and confined spaces.

Energy sources and power transmission. Potentially catastrophic results are associated with the rapid uncontrolled release of energy from high-energy sources, such as fuel supplies, pressure vessels and boilers, cylinders containing compressed gases, and electrical power supplies. Uncontrolled releases of energy may occur at the source or during power transmission to the point of operation. The nature of such hazards depends upon the energy transmitted. For example, electrical power lines pose serious dangers when they are inadvertently contacted. At high voltage levels, the electromagnetic field surrounding power lines may induce currents or cause voltage buildups. Piping systems containing fuels such as natural gas, heated liquids, steam, compressed air, or toxic materials analogously pose serious risks, if they break or are inadvertently cut by other equipment. Moving parts in mechanical power-transmission equipment are dangerous if contacted. Rotating shafts, gears, pulleys, wheels, impacts, and the turbulent flow of gases or liquids produce noise and vibration hazards. *See* ELECTRIC POWER SYSTEMS.

Point-of-operation hazards. Machinery and tools may pose a potential for abrading, cutting, shearing, or crushing injuries when people contact sharp or moving parts. Machinery may also propel flying objects and emit dusts and potentially hazardous materials, resulting in eye or other injuries. Other point-of-operation hazards include high voltages, nonionizing radiation, and hot or cold surfaces. Point-of-operation hazards usually decrease when industries shift away from manual machining tasks.

Vehicles and materials-handling systems. Collisions between vehicles or involving pedestrians are a particularly significant problem in many industries. Overhead cranes, hoists, conveyors, forklifts, carts, and other materials-handling systems can also cause impact-related injuries. Collisions and impacts cause substantial losses. *See* MATERIALS-HANDLING EQUIPMENT.

Unsafe walking or climbing surfaces. Many falls occur while people are climbing or maneuvering on elevated surfaces, such as stairs, escalators, ladders, and scaffolds, or near floor openings; but most falls involve slipping or tripping while walking, pushing, or pulling objects on level ground. The coefficient of friction between a shoe and floor surface is a primary determinant of slipping hazard. This coefficient varies between floor surfaces and shoe soles, and as a function of contaminants.

Ingress-egress and confined spaces. Inadequate or blocked passageways may impede ingress or egress from work areas or allow insufficient clearance for passing vehicles. Poorly designed ingress-egress systems can impede emergency egress from vehicles and cause falls. Confined spaces that are entered during maintenance or repair tasks, such as tanks, boilers, or ducts, can pose special ingress-egress problems, and may concentrate hazardous vapors, fumes, and gases.

Physical work environment. Hazards in the physical work environment include vibration and noise, thermal extremes, pressure extremes, and ionizing or nonionizing radiation.

Vibration and noise. Excessive vibration and noise is a potential safety problem in vehicles and many work environments. Whole-body vibration at certain resonant frequencies is annoying and fatiguing to workers and drivers and may cause chronic pain. Localized vibration produced by some hand-tools can interfere with blood circulation, resulting in Raynaud's phenomenon, damage to the soft tissues of the hand, bursitis, or arthritis.

Noise becomes a safety problem at certain levels and frequencies. The most commonly used scale [the adjusted decibel (dBA) scale] attenuates low frequencies not normally used for speech communication. Even at low levels, noise can annoy or distract workers. At moderate levels of 80–90 dBA, noise interferes with communication and causes hearing loss in susceptible people. Higher levels (above 85 dBA) significantly increase the chance of hearing loss. The effects of long-term exposure to noise are cumulative and nonreversible. In the United States the Occupational Safety and Health Act specifies allowable levels and durations of exposure to noise (**Table 1**). The act also requires that employers monitor noise levels, perform audiometric tests, furnish hearing protection, and maintain exposure records whenever employee noise exposure equals or exceeds an 8-h time-weighted average level of 85 dBA. *See* ACOUSTIC NOISE; AUDIOMETRY; HEARING (HUMAN); HEARING IMPAIRMENT.

Thermal extremes. Excessively hot or cold working environments can adversely affect workers. Working at high temperatures can result in heat cramps, heat exhaustion, or heat stroke. Such effects become more likely when the work is strenuous, the exposure is prolonged, the humidity is high, or the

TABLE 1. Permissible noise exposure

Duration per day	Slow response, decibel level, adjusted decibels (dBA)
8	90
6	92
4	95
3	97
2	100
$1^1/_2$	102
1	105
$^1/_2$	110
$^1/_4$ or less	115

workers are using respirators or wearing protective clothing that interferes with normal evaporative cooling. Working at low temperatures can result in localized tissue damage, due to restriction of circulation, or hypothermia at temperatures well above freezing; Frostbite can occur below freezing. Air movement increases the potential for frostbite and hypothermia. Moisture also increases the rate of cooling and the risk associated with cold temperatures. *See* COMFORT TEMPERATURES; HYPOTHERMIA.

Pressures. Environments above or below normal atmospheric pressure can be a significant hazard. Hyperbaric pressures are found in diving, underwater, or in tunneling operations where compressed gas is used to prevent the entry of water. Decompression injuries may occur in such tasks, when gases expand in body cavities or when nitrogen dissolved in the bloodstream or other tissues changes to gaseous form. One solution is to slowly carry out the decompression process. A second is to provide mixtures of compressed oxygen and helium, rather than oxygen and nitrogen (helium is less soluble in blood and other tissues). Hypobaric pressures are found at high altitudes. Decompression rapid enough to cause pain or injury to the inner ear and sinuses may occur when airplanes rapidly lose cabin pressure or reduce altitude by a few thousand feet. *See* AIR PRESSURE; DECOMPRESSION ILLNESS; DIVING.

Ionizing and nonionizing radiation. Radioactive materials that emit alpha and beta particles, neutrons, x-rays, or gamma rays are sources of ionizing radiation. X-rays and radioactive materials are commonly used in medicine, chemical laboratories, and other industries to scan or obtain information from within living or inert objects. They are also used to generate power. Ionizing radiation has sufficient energy to ionize components of human cells. Radioactive materials consequently pose a potential hazard if excessive amounts are ingested, inhaled, or otherwise introduced into the body. Sources of alpha particles pose a hazard only when within the body. Sources of beta particles are capable of burning the skin but are primarily a hazard when within the body. Sources of neutrons, x-rays, and gamma rays pose hazards both within and outside the body. Control measures include limits on exposure time, adequate separation distance from the source, and use of barriers. *See* RADIATION BIOLOGY; RADIATION INJURY (BIOLOGY).

Forms of nonionizing radiation include visible, ultraviolet, infrared, and microwave radiation. Visible light is focused by the lens of the eye onto the retina, while the other forms are beyond the visible spectrum. High-intensity visible light may burn the retina, resulting in color blindness or blind spots. Ultraviolet light can seriously burn human skin and the cornea of the eye. Repeated exposure has been implicated as a cause of skin cancer and cataracts. Welding equipment, lasers, and the Sun emit potentially dangerous intensities of visible and ultraviolet light. Infrared radiation is emitted by hot objects and may excessively heat workers, resulting in heat exhaustion, heat stroke, or heat cramps. At high levels, it can burn the skin and eyes and cause cataracts. Microwave radiation is emitted by equipment such as radar, microwave ovens, and communication apparatus. Microwaves can cause serious tissue damage by excessive heating of exposed body parts and inductive heating of metals. *See* ELECTROMAGNETIC RADIATION; LASER; MICROWAVE; RADAR.

Hazards of materials. Materials used in industrial processes vary greatly in nature and form. Mists, vapors, gases, liquids, dusts, and fumes from certain materials may be hazardous. Some materials pose fire and explosion hazards. Others are chemically or biologically active when they contact or enter the human body. Even chemically inert materials can cause injuries or illness.

Fire and explosion hazards. Fires and explosions are major hazards of certain materials. Numerous materials can burn and create fire hazards. Airborne gases, vapors, fumes, or dusts are particularly dangerous, since they can travel significant distances to reach ignition sources and often burn explosively. Beyond the hazards of heat and potentially explosive forces, combustion of materials can result in highly toxic products. Fires also deplete the oxygen content of air, creating the potential for suffocation. *See* COMBUSTION; EXPLOSIVE; FIRE TECHNOLOGY.

Chemical and biological hazards. Certain chemicals cause immediate or delayed injury or even destruction of various organs upon bodily contact or entry. Biologically active agents, such as bacteria, viruses, fungi, mycoplasma, or parasitic organisms, cause disease. Chemicals and biologically active agents can enter the body through the skin, by ingestion, or by inhalation. Some agents readily pass through unbroken skin; others may enter through skin abrasions or cuts. Workers may ingest contaminated food or fluids while eating or smoking without first washing their hands. Dusts, fumes, mists, vapors, gases, and biologically active agents may also be inhaled.

The effects of chemicals are related to toxicity, particle size, and mode of entry into the body. Several categories of toxic chemical effects have been determined. Asphyxiation refers to the effects of chemicals that cause an inadequate supply of oxygen to be delivered through the blood to organs within the body. Simple asphyxiants dilute the air that is breathed so that not enough oxygen is present to support life. Chemical asphyxiants interfere with the ability of blood to deliver oxygen. *See* HYPOXIA.

Irritation occurs when contact with a chemical causes an inflammatory or allergic response by the body. Fine particulate matter or dusts can cause irritation of the respiratory tract and lungs, even when the matter is nontoxic or chemically inert. Symptoms of irritation include an increased flow of blood to the location of contact, swelling, and the activation of immune system response. Severe cases may involve dermatitis, extensive blistering, pulmonary edema, or interference with breathing. *See* ALLERGY; INFLAMMATION; RESPIRATORY SYSTEM DISORDERS.

Corrosive or chemical burns occur when a chemical agent reacts with the contacted surface to break down the tissues. Systemic poisons interfere with the metabolic functions of tissues and often affect only particular organs; neurotics, depressants, anesthetics, stimulants, and hypnotics target the nervous system; and carcinogens and mutagens cause various forms of cancer. *See* MUTAGENS AND CARCINOGENS; TOXICOLOGY.

Task demands. The task performed by a worker can be hazardous. Lifting, pushing, pulling, and other physical activity can cause injury when applied or reactive forces, pressures, or torques exceed the tolerance of the body. Repeated performance of manual tasks over prolonged periods, excessive reaches, twisting motions, rapid movements, and postures that concentrate forces can significantly increase the risk of injury. Tasks that are stressful or monotonous can also contribute to human error. Changes in work conditions requiring deviations from ordinary routines, such as when equipment is being repaired, are particularly likely to increase the chance of errors.

Safety Activities

Safety and health activities performed within industry include identifying and analyzing hazards, ensuring compliance with standards and codes, implementing controls, and monitoring controls.

Hazard identification and analysis. A fundamental safety and health activity is to identify potential hazards and then to analyze them in terms of severity and probability. This process allows the cost of control measures to be compared with expected loss reduction and helps justify choices between control alternatives.

Hazard identification is guided by past experience, codes and regulations, checklists, and other sources. This process can be organized by separately considering each step in the making of a product (**Fig. 1**). Numerous complementary hazard analysis methods are available, including failure modes and effects analysis, work safety analysis, human error analysis, and fault tree analysis. Failure mode and effects analysis systematically documents the effects of malfunctions on work sheets that list the components of a system, their potential failure modes, the likelihood and effects of each failure, and potential countermeasures. Work safety analysis and human error analysis are related techniques that organize the analysis around tasks rather than system components. This process involves an initial division of tasks into subtasks. For each subtask, potential effects of product malfunctions and human errors are then documented. Fault tree analysis takes an approach that begins with a potential accident and then works down to its fundamental causes. Fundamental causes may be system malfunctions, human errors, or ordinary nonmalfunction states. Probabilities are often assigned to the fundamental causes, allowing the probability of accidents to be calculated. *See* OPERATIONS RESEARCH; RISK ASSESSMENT AND MANAGEMENT; SYSTEMS ANALYSIS.

Compliance with standards and codes. Determining which standards, codes, and regulations are

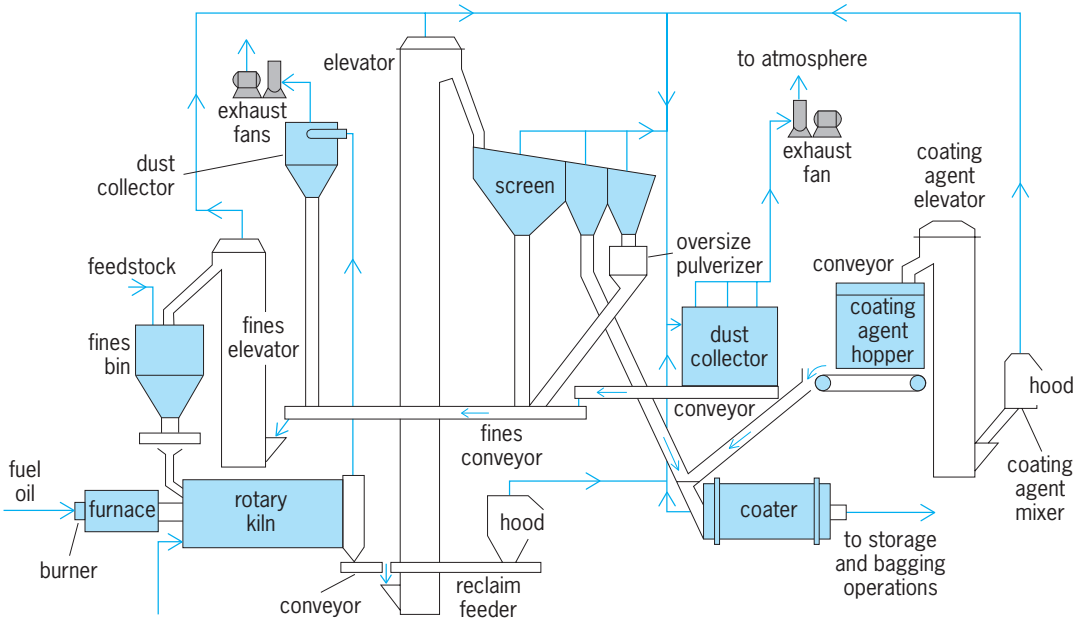

Fig. 1. Process flow sheet showing the stepwise introduction of raw materials and the product of each step. The extent of chemical or physical hazards that may arise out of any step in the operation should be determined.

U.S. DEPARTMENT OF LABOR
Occupational Safety and Health Administration

Form Approved
OMB No. 44-R1387

MATERIAL SAFETY DATA SHEET

Required under USDL Safety and Health Regulations for Ship Repairing, Shipbuilding, and Shipbreaking (29 CFR 1915, 1916, 1917)

SECTION I

MANUFACTURER'S NAME	EMERGENCY TELEPHONE NO.
ADDRESS *(Number, Street, City, State, and ZIP Code)*	
CHEMICAL NAME AND SYNONYMS	TRADE NAME AND SYNONYMS
CHEMICAL FAMILY	FORMULA

SECTION II - HAZARDOUS INGREDIENTS

PAINTS, PRESERVATIVES, & SOLVENTS	%	TLV (Units)	ALLOYS AND METALLIC COATINGS	%	TLV (Units)
PIGMENTS			BASE METAL		
CATALYST			ALLOYS		
VEHICLE			METALLIC COATINGS		
SOLVENTS			FILLER METAL PLUS COATING OR CORE FLUX		
ADDITIVES			OTHERS		
OTHERS					

HAZARDOUS MIXTURES OF OTHER LIQUIDS, SOLIDS, OR GASES	%	TLV (Units)

SECTION III - PHYSICAL DATA

BOILING POINT (°F.)		SPECIFIC GRAVITY (H_2O=1)	
VAPOR PRESSURE (mm Hg.)		PERCENT, VOLATILE BY VOLUME (%)	
VAPOR DENSITY (AIR = 1)		EVAPORATION RATE (_____ = 1)	
SOLUBILITY IN WATER			
APPEARANCE AND ODOR			

SECTION IV - FIRE AND EXPLOSION HAZARD DATA

FLASH POINT (Method used)	FLAMMABLE LIMITS	Lel	Uel
EXTINGUISHING MEDIA			
SPECIAL FIRE FIGHTING PROCEDURES			
UNUSUAL FIRE AND EXPLOSION HAZARDS			

SECTION V - HEALTH HAZARD DATA

THRESHOLD LIMIT VALUE
EFFECTS OF OVEREXPOSURE
EMERGENCY AND FIRST AID PROCEDURES

SECTION VI - REACTIVITY DATA

STABILITY	UNSTABLE		CONDITIONS TO AVOID
	STABLE		
INCOMPATABILITY *(Materials to avoid)*			
HAZARDOUS DECOMPOSITION PRODUCTS			
HAZARDOUS POLYMERIZATION	MAY OCCUR		CONDITIONS TO AVOID
	WILL NOT OCCUR		

SECTION VII - SPILL OR LEAK PROCEDURES

STEPS TO BE TAKEN IN CASE MATERIAL IS RELEASED OR SPILLED
WASTE DISPOSAL METHOD

SECTION VIII - SPECIAL PROTECTION INFORMATION

RESPIRATORY PROTECTION *(Specify type)*		
VENTILATION	LOCAL EXHAUST	SPECIAL
	MECHANICAL *(General)*	OTHER
PROTECTIVE GLOVES	EYE PROTECTION	
OTHER PROTECTIVE EQUIPMENT		

SECTION IX - SPECIAL PRECAUTIONS

PRECAUTIONS TO BE TAKEN IN HANDLING AND STORING
OTHER PRECAUTIONS

Fig. 2. Sample Material Safety Data Sheet, used to obtain information from the manufacturer. (*Occupational Safety and Health Administration*)

relevant and then ensuring compliance are essential health and safety activities. Consensual standards developed by the American National Standards Institute (ANSI) are particularly important; they address topics such as chemical labeling, material safety data sheets, personal protective equipment, workplace signs and product labels, and pressure vessels. Other widely applicable standards developed in the United States by nongovernmental groups include the National Electric Code, the Life Safety Code, and the American Council of Industrial and Governmental Hygienists guidelines for threshold limit values. Thousands of consensual standards contain safety provisions.

In the United States, the best-known governmental standards are the general industry standards specified by the Occupational Safety and Health Administration (OSHA). OSHA also specifies standards for the construction, maritime, and agriculture industries. Other standards include those specified by the Environmental Protection Agency (EPA) on disposal and cleanup of hazardous materials, the Federal Aviation Administration (FAA) standards on worker safety in air travel, the Federal Highway Administration (FHWA) standards regarding commercial motor carriers, the Mine Safety and Health Administration (MSHA) standards for mine workers, the Nuclear Regulatory Commission's and Department of Energy's standards regarding employees working with radioactive materials, and the U.S. Coast Guard standards regarding safety of workers on tank and passenger vessels. State and local governments may also implement safety and health standards.

The scope of governmental regulations is illustrated by the OSHA general industry standard. Besides specifying a large set of detailed and mandatory health and safety design standards, OSHA addresses noncovered hazards with a general duty clause that places an obligation on employers to provide a workplace free from recognized hazards. OSHA also requires that industries implement comprehensive hazard communication programs for workers involving labeling, material safety data sheets, and employee training (**Fig. 2**). OSHA enforces these standards by conducting inspections and imposing fines if they are not met. Employers are required to maintain records of work-related injuries and illnesses and to prominently post both an annual summary of injuries and illness and notices of noncompliance with standards. Other functions of OSHA include monitoring the performance of state agencies which implement health and safety standards and coordinating such activity with that of other regulatory agencies.

Control methods. Methods of controlling or eliminating hazards include plant or process design; job design; employee selection, training, and supervision; personal protective equipment; and warnings.

Plant or process design. Plant and process design can eliminate or control hazards in numerous ways. Eliminating hazards at their source by substituting less hazardous materials, equipment, or processes is an ideal solution, if feasible. For example, hazardous mists can be eliminated by brush painting instead of spray painting. Processes can be modified to eliminate hazards at the source, such as wetting-down grinding operations to minimize dust production. Another solution is to eliminate hazardous jobs through automation. For example, robot welders and painters have eliminated hazards in the automobile industry. *See* AUTOMATION; DUST AND MIST COLLECTION; HAZARDOUS WASTE; ROBOTICS.

Eliminating hazards, however, may be technically infeasible. An alternative approach is to isolate them, thereby reducing the number of potentially exposed workers; the conflicts between processes, such as welding and paint spraying operations; the potential contact with power transmission equipment, such as power lines; and the catastrophic potential of high-energy systems. Isolation may be attained through barriers and enclosures that block or absorb hazardous flows of energy or material. Plant layouts can contribute to this goal by spatially separating hazards and susceptible structures. Barriers such as fire doors and fire walls, along with fire suppression and warning systems, provide time to escape and respond to fires. Process enclosures prevent the release of hazardous chemicals, noise, heat, or other hazards into the workplace. Appropriate passageways separate workers from vehicles and hazardous processes. Under emergency conditions, they also speed egress of workers and access by emergency equipment and personnel. Detailed provisions regarding passageways and emergency exits include appropriate passage dimensions, markings, signage, lighting, stairs, walking surfaces, door locations, and door designs. *See* AIR FILTER; INDUSTRIAL ENGINEERING; INDUSTRIAL FACILITIES.

The removal or dilution of airborne contaminants through local or dilution ventilation is another important control measure. Local ventilation removes contaminants near their source, often from within enclosed processes, making it more effective than dilution ventilation. OSHA specifies extensive ventilation requirements for tasks producing airborne contaminants. *See* VENTILATION.

Job design. Job design involves the specification of appropriate methods, tools, and equipment. Focus may also be on safety procedures and safety devices. Each aspect of job design can play a critical role in eliminating hazards, preventing contact with hazards, and minimizing effects of contact.

In some cases, hazards can be entirely eliminated through job design. For example, the need for heavy lifting, pushing, or pulling may be eliminated by providing hoists, conveyors, or other materials-handling equipment. Automation may eliminate the need for repetitive manual operations in assembly tasks. More often, job design reduces hazards. For example, providing grounded tools, nonconductive mats, and nonconductive ladders reduces, but does not eliminate, electrocution hazards. Job design can also keep physical demands within reasonable bounds through job rotation, rest intervals, and ergonomic guidelines. Among such guidelines are those that describe limits on lifted weights as a function of load location, lifting height, task duration, and lifting frequency. Guidelines oriented toward preventing carpal tunnel syndrome specify acceptable wrist orientations and exerted forces when using keyboards, hand tools, and other devices. Guidelines for preventing falls in climbing systems specify appropriate handholds, steps, or rungs. Guidelines for preventing falls in walking, lifting, pushing, or pulling tasks specify slip-resistant work surfaces, appropriate footwear, and limits on exerted forces.

Contact with hazards often can be prevented by modifying tools and equipment. One approach is to modify contact surfaces by eliminating sharp edges, points, rough surface finishes, pinch points, and temperature extremes. Guards also can prevent workers from contacting hazardous components of equipment or tools, but are less effective than eliminating the hazard since they can be removed by workers.

Other approaches address unsafe behavior due to perceptual errors and memory lapses. Examples include delineating risk zones with barriers or markings, providing unobstructed vision, and reducing levels of noise. Inadvertent activation of equipment and entry into risk zones can be minimized by providing temporary barriers, interlocks, and lockouts. Errors can also be reduced by providing appropriate displays and controls, reducing demands on short-term memory, and lowering the cognitive complexity of tasks. Job performance aids, such as checklists, instruction manuals, and computer programs, can be particularly useful for preventing errors when performing nonroutine tasks, such as maintenance, repair, and troubleshooting. *See* HUMAN-MACHINE SYSTEMS.

Employee selection, training, and supervision. Modification of the employee through employee selection, training, and supervision is necessary in almost all occupations. Selecting employees with the necessary physical and cognitive abilities is essential. For example, airplane pilots must have good visual acuity, the safe use of hand tools requires some minimal level of dexterity, and lifting tasks require some minimal level of strength.

Training workers and supervisors is important. A training program must teach workers what the hazards are and their severity, how to identify and avoid them, and what to do after exposure. Trained workers should follow safe procedures, including steps to take after exposure, such as washing and rinsing after skin or eye contact with hazardous materials; methods of proper disposal of contaminated clothing, hazardous chemicals, or biologically active agents; necessary maintenance activities; and housekeeping procedures, such as cleaning up spills or

accumulations of materials in the work environment. Housekeeping procedures are particularly essential to prevent falls, fires, and exposure to hazardous materials.

Supervision, as the final element related to modification of the employee, should focus on consistently reinforcing safe procedures.

Personal protective equipment. Personal protective equipment, such as gloves, respirators or masks, protective clothing, hearing protectors, eye protectors, foot protectors, helmets, seat belts, and safety harnesses, may be useful when other control methods are inadequate or infeasible. Training in proper use of personal protective equipment is important and supervision may be necessary to ensure that the equipment is used properly when needed. Plant and task layouts should provide convenient access to personal protective equipment. *See* RESPIRATOR.

Warnings. Warnings are the method of last resort, and they should be viewed only as a supplement to proper training and supervision. Warnings are of primary value as a means of alerting knowledgeable users to follow an established safe behavior pattern. When established behavior patterns are unsafe, supervision becomes the most important option.

Monitoring controls. Accident investigations, plant inspections, and monitoring are complementary ways of ensuring that implemented control strategies are fulfilling their intended function. They can uncover deficiencies in existing controls and help formulate needed changes.

Accident investigation. Accident investigation provides a means of determining why existing control methods failed. Accident investigations vary from simple documenting of what happened on a supervisor's report to in-depth studies involving governmental agencies and experts in forensics, medicine, human-factors engineering, and other fields. Supervisors' reports document who was injured, the nature and extent of injury, location and time of the accident, and what was being done at the time of the accident. The reports also include a narrative description of cause and corrective actions. Accident investigations almost always reveal multiple causes for accidents. An accident investigation should document violations of safety standards and operating procedures, equipment malfunctions, errors, and unsafe conditions. More sophisticated investigations focus on the root causes of equipment malfunctions and human errors, such as management oversights (inadequate facilities, training, supervision, emergency procedures and so on), complex interactions between processes, or error-provocative situations.

Plant inspections. Plant inspections are usually performed on a periodic basis to determine whether equipment, machinery, and tools are within design specifications. Routine quick walkthroughs by in-house safety engineers and supervisors can uncover unsafe conditions and violations of operating procedures. Inspections should focus on environmental control systems, such as ventilation systems; pollution control systems; walking and climbing surfaces; tools, machinery, and materials-handling systems; accumulations of flammable or toxic materials; violations of operating procedures, such as failure to use personal protective equipment; disabling or removal of guards; unsafe operations of equipment and vehicles; lack of maintenance; failures to lock out equipment during repair or maintenance; and smoking in restricted areas.

Monitoring. Monitoring is a continuing program of observation, measurement, and judgment of the exposure of workers to chemical and physical hazards of the work environment (**Table 2**). Monitoring is essential to determine whether control methods are successfully reducing exposure. The four most

TABLE 2. Sampling procedures for industrial hygiene surveys

Chemical hazards (airborne)		Physical hazards	
Gases and vapors	May be determined by use of approved calibrated field indicator tubes yielding direct readings	Pressure	May be measured barometrically
	May be collected in containers or absorbed on charcoal for laboratory evaluation	Temperature	May be measured by thermometer, thermocouple, or radiometer; determination of heatstress, however, requires in some form the measurement of evaporation rate; usually heat stress is inferred from humidity and air velocity
	Organic vapor may be absorbed on charcoal or determined chromatographically	Ionizing radiation	May be measured by survey meter, personal dosimetry, or film badge techniques
Fumes and mists	May be absorbed and measured in the field	Noise levels	May be measured with sound-level meters or octave-band analyzers; vibration may be determined with additional sound-level equipment
	May be absorbed and evaluated in the laboratory	Nonionizing radiation	May be measured by a number of direct-reading meters
	May be collected on filter media and analyzed in the laboratory		
Dusts	May be collected by a personal air sampler and fractionated into respirable size by a cyclone separator, and the fractions weighed to determine the concentration		
	May be collected on an open-faced filter and weighed		
	May be collected in an appropriate manner and counted		

common methods are environmental, personal, biological, and medical monitoring. Environmental monitoring is performed at locations where employees are likely to be exposed to hazards. Noise levels, radiation levels, concentrations of airborne contaminants, and temperatures are routinely measured to determine risks. Technically advanced systems may use remote sensors to detect transient and potentially hazardous changes in the work environment. Certain systems may respond to unacceptable deviations by giving warnings, triggering fire suppression systems, or adjusting flow rates in ventilation systems.

Personal monitoring entails placing a dosimeter on a worker as close as possible to the port of entry. Thus, the amount of exposure that the employee receives over the course of a work interval can be measured. For example, a dosimeter placed in the breathing zone can be used to estimate the amount of hazardous material inhaled. Biological monitoring involves the measurement of hazardous materials or metabolic by-products of hazardous materials in bodily fluids, tissues, hair, or expired air. Examples include the measurement of lead, fluoride, cadmium, or mercury in blood or urine, or of phenol in urine to measure excessive benzene absorption. Medical monitoring is a related technique used to measure health effects on workers after exposure; an example is the routine chest x-rays of workers. *See* CLINICAL PATHOLOGY; DOSIMETER; HUMAN-FACTORS ENGINEERING.

Mark R. Lehto

Bibliography. D. B. Chaffin et al., *Occupational Biomechanics*, 4th ed., 2006; L. J. DiBerardinis (ed.), *Handbook of Occupational Safety and Health*, 2d ed., 1998; H. W. Heinrich, D. Peterson, and N. Roos, *Industrial Accident Prevention*, 5th ed., 1980; M. R. Lehto and J. M. Miller, *Warnings*, vol. 1: *Fundamentals, Design, and Evaluation Methodologies*, 1986; National Institute for Occupational Safety and Health, *Work Practices Guide to Manual Lifting*, NIOSH 81–122, 1981; National Safety Council, *Accident Prevention Manual for Industrial Operations*, vol. I: *Administration and Programs*, vol. II: *Engineering and Technology*, 9th ed., 1988; G. Salvendy (ed.), *Handbook of Industrial Engineering*, 3d ed., 2002; G. A. Taylor, K. Easter, and R. Hegney, *Enhancing Occupational Safety and Health*, 2004.

Industrial meteorology

The application of meteorological information to industrial, business, or commercial problems. Generally, industrial meteorology is a branch of applied meteorology, which is the broad field where weather data, analyses, and forecasts are put to practical use. In recent years, the term "private sector meteorology" has taken on the broader context of traditional industrial meteorology, expanding to include the provision of weather instrumentation/remote sensing devices, systems development and integration, and various consulting services to government and academia as well as value-added products and services to markets in industry (such as media, aviation, and utilities). Some areas in which industrial meteorology may be applied include environmental health and air-pollution control, weather modification, agricultural and forest management, and surface and air transportation. *See* AERONAUTICAL METEOROLOGY; AGRICULTURAL METEOROLOGY; METEOROLOGICAL INSTRUMENTATION.

Specific examples of the uses of industrial meteorology include many in the public sphere. For example, electric utilities need hourly predictions of temperature, humidity, and wind to estimate system load. In addition, they need to know when and where thunderstorms will impact their service area, so that crews can be deployed to minimize or correct disruptions to their transmission and distribution systems. Highway departments need to know when and where frozen precipitation will affect their service areas so that crews can be alerted, trucks loaded with sand and salt, and, if necessary, contractors hired to assist. Since a few degrees' change in temperature, or a slight change in intensity of snow or ice, determines the type of treatment required, early prediction and close monitoring of these parameters are critical.

Agricultural enterprises, from farmers to cooperatives to food manufacturers, rely on precise weather information and forecasts. Weather is the single most important factor in determining crop growth and production. Thus, monitoring and prediction of drought, floods, heat waves, and freezes is of extreme importance.

Professionals involved with the meteorological aspects of air pollution are generally concerned with the atmospheric transport, distribution, transformation, and removal mechanisms of air pollutants. They are often called upon to evaluate the effectiveness of pollution control technologies or regulatory (policy) actions used to achieve and maintain air-quality goals.

History. The genesis of industrial meteorology began at the end of World War II. During the war, meteorologists within the armed forces were responsible for providing site- and time-specific forecasts to support military operations. With the introduction of radar and other observing and forecasting methods, meteorologists became problem solvers for their military clients. After the war, some of these meteorologists remained in the field and developed private sector businesses. Through the 1950s and 1960s the primary service provided by these companies was weather forecasting: short- and medium-term predictions of site-specific weather conditions.

The 1970s and 1980s saw a substantial increase in the availability of real-time weather information, fueled by increases in computing power and communications technology, such as radar and satellite imagery, numerical model output, and surface observations. Eventually the traditional methods of transmitting these data, such as analog phone lines and slow-speed teletypes, became insufficient. High-speed computer-to-computer connections, satellite distribution, and local (computer) processing replaced the hand-plotted paper weather maps,

wet-paper facsimile machines, and miles of teletype paper rolls that had been in use since World War II. This achievement belonged to a new segment of industrial meteorology—the value-added weather information industry. Seeing the need to capture the megabytes of weather data available every minute and transform them into usable information, the value-added industry blossomed from a few small start-up companies, in the late 1970s, to many organizations with hundreds of employees and annual sales in the tens of millions of dollars by the mid-1990s.

Value-added meteorology. These companies absorb vast amounts of digital data provided by government and other sources, such as the United Nations' World Meteorological Organization and aircraft pilot reports, and turn it into useful information for their clients. Value-added products include animated graphics, interactive model manipulation, high-powered workstations with associated software, and Internet site content. The value-added sector also provides satellite transmission and delivery of its products, customer service and training, and other user-specific services. Similarly, systems integration organizations design, develop, install, and maintain customized, high-powered meteorological processing and display systems for clients in government, industry, research, and academia. Today, the value-added weather information and systems integrations sectors are substantially larger (in personnel employed and sales volume) than the traditional weather forecasting and consulting companies, though these latter sectors remain strong.

Consulting and media services. At present, demand for personalized consulting services continues to be strong. The American Meteorological Society (AMS) has certified over 600 consulting meteorologists since the inception of the Certified Consulting Meteorologist (CCM) program in 1957. Certified Consulting Meteorologists provide services in a number of areas, including the provision of expert testimony in legal cases (forensic meteorology). The National Council of Industrial Meteorologists was founded in 1968 to further the development and expansion of the professional practice of industrial meteorology.

Media services are also an important part of industrial meteorology. It is estimated that about 75% of the weathercasters on television and radio in the United States are professional meteorologists. Since 1959 the American Meteorological Society has awarded its Seal of Approval to over 150 individuals in radio and over 1000 in television. The public's appetite for frequent, reliable weather information is evidenced by the popularity of full-color newspaper weather pages and a cable television national weather network, both products of the private sector.

Weather forecasting. With the significant improvements in short- and long-range forecasting capabilities in the 1990s, the insurance and financial communities have created risk-hedging "weather derivative contracts" for weather-sensitive industries such as utilities and agriculture. Private-sector meteorologists are assisting in the development of these markets through the refinement and delivery of temperature data and seasonal forecasts. *See* WEATHER FORECASTING AND PREDICTION.

In the United States, the federal government provided weather forecasting and prediction services for decades through local offices of the National Weather Service (NWS). Surplus demand for these services, combined with the desire of many in the private sector to provide the services themselves, led to the formulation of a mixed economic approach to weather forecasting. In 1989 the NWS (formerly the U.S. Weather Bureau, and now an agency of the National Oceanic and Atmospheric Administration within the U.S. Department of Commerce) developed Circular A-06, *Policy and Guidelines Governing National Weather Service and Private Sector Roles*.

The primary mission of the NWS, according to these policy guidelines, is the protection of life and property and the enhancement of the national economy. Hence, the basic functions of the NWS are the provision of forecasts and warnings of severe weather, flooding, hurricanes, and tsunami events; the collection, exchange, and distribution of meteorologic, hydrologic, climatic, and oceanographic data and information; and the preparation of hydrometeorological guidance and core forecast information. The NWS is the single official voice when issuing warnings of life-threatening situations and is the source of a common national hydrometeorological information base. This national information base forms an infrastructure on which the private sector can build and grow. *See* HYDROMETEOROLOGY.

The NWS does not compete with the private sector when a service is currently provided or can be provided by commercial enterprises, unless otherwise directed by applicable law. The private weather industry is ideally suited to put the basic data and common hydrometeorological information base from the NWS into a form, and with the necessary detail, that is most useful to specific weather and water resources–sensitive users. Todd S. Glickman

Bibliography. J. R. Fleming (ed.), *Historical Essays on Meteorology 1919–1995*, pp. 417–442, American Meteorological Society, Boston, 1996; T. S. Glickman, *Glossary of Meteorology*, 2d ed., American Meteorological Society, Boston, 2000; D. D. Houghton, *Handbook of Applied Meteorology*, John Wiley, New York, 1985; D. D. Houghton, Results of the 1995 AMS Private Sector Survey, *Bull. Amer. Meteorol. Soc.*, 77(2):325–333, 1996.

Industrial microbiology

A field concerned with the development of technologies to control and manipulate the growth and activities of selected biological agents to create desirable products and economic gain or to prevent economic loss. In addition to bacteria and yeasts, animal and plant cell cultures are now used to produce sophisticated products such as monoclonal

antibodies, immunomodulating compounds, and complex plant metabolites.

The fermentation of foods and beverages was the earliest application of industrial microbiology. Although such products have been consumed for thousands of years, only in the nineteenth century was microbial activity associated with the fermentation process. Soon after that discovery, microorganisms, especially bacteria, were selectively introduced on the commercial level. During World War I, industrial microbiology was directed at the production of chemicals, especially glycerol, acetone, and butanol. Techniques were developed gradually for pure-culture fermentation and strain improvement, but the major advance in industrial microbiology occurred during World War II with the large-scale production of penicillin by submerged-culture fermentation.

In the 1950s, at a time when microbial processes for chemicals production were being replaced by petrochemical processes, industrial microbiology shifted its focus to the production of therapeutic agents, especially antibiotics. Advances in molecular biology have greatly increased the potential applications of industrial microbiology in areas such as therapeutics, diagnostics, environmental protection, and agriculture. The techniques of genetic engineering, along with technology developments in bioprocessing, make possible large-scale production of complex natural compounds that would otherwise be very difficult to obtain.

Strain development. With the exception of the food industry, few commercial fermentation processes use wild strains of microorganisms isolated directly from nature. Of the many thousands of microbial species, few are used commercially, and fewer still are used as hosts for genetically engineering genes. Process development occurs in large part by strain improvement directed at increasing product yield, enhancing growth on cheaper substrates, and simplifying purification.

Strain development is achieved by either a traditional mutation and selection program or direct genetic manipulation. The dominant reason for strain improvement by mutation and selection is the alteration of the regulatory mechanisms of an organism so that maximum metabolic energy is directed to a single product, either a protein or a product from a metabolic pathway. The approach, which is relatively rapid, is inexpensive, and has shown good results, involves successive rounds of mutagenesis and screening for the desired mutants. However, it is limited by the initial genetic potential of the organism.

The recombinant deoxyribonucleic acid (DNA) approach has succeeded in introducing new genetic material into a convenient host microorganism and amplifying genetic material. About 20% of the synthesizing capacity of a bacterium can be devoted to a single polypeptide or protein. *See* RECOMBINATION (GENETICS).

Fermentation. Commercial microbial compounds are produced in two distinct phases: fermentation and product recovery. Fermentation procedures are developed by determining optimal conditions for product synthesis. Commercial production usually occurs in a batch fermentor, where gas of controlled composition and flow is bubbled through a stirred pure microbial culture suspended in a liquid medium of optimum nutrient composition. The maintenance of sterile conditions during the fermentation is crucial, and so all liquids and gases are sterilized before they are fed to the fermentor. Key operating variables such as pH, temperature, and dissolved oxygen content are controlled at their optimal values during fermentation for rapid cell growth and product formation.

Product recovery and purification involves a series of operations. The first steps usually involve cell disruption or the separation of the cell or cellular debris from the fluid medium, typically through centrifugation and filtration. Later stages of purification include finer membrane filtration, extraction, precipitation, and chromatography. These techniques must be optimized with respect to the specific biochemical and physical differences between product and impurities. *See* CENTRIFUGATION; FERMENTATION; FILTRATION; FOOD FERMENTATION; STERILIZATION.

Applications. The production of certain foods and beverages was an early application of industrial microbiology. Now, the fields of mineral recovery, medicine, environmental protection, and food and agriculture are using similar techniques.

Mineral recovery. Bacterial leaching reactions have been used to alter metal-bearing minerals, usually converting them to more soluble forms before the metals are extracted. Such operations can result in improved extraction rates in comparison to those of conventional processes, which are usually conducted on ore waste dumps and heaps. Large-scale commercial applications of biochemical mining and extraction have been limited mainly to copper and uranium. Through genetic engineering techniques, microorganisms are being developed that are more compatible with high concentrations of metal ions, low levels of oxygen, and high temperatures. For some applications, a mixture of several strains is the best, because some are more resistant to high temperature, whereas others tolerate higher concentrations of heavy metals. The most common strain, involved in copper and iron sulfide leaching, is *Thiobacillus ferrooxidans*.

Bacteria derive all their metabolic needs from inorganic substances, such as carbon dioxide and metal sulfides, and organic compounds, such as carbon dioxide, sugars, and organic acids. Certain combinations of these organics can enhance the effectiveness of the bacteria, as can nitrogen-enriched leachates and surfactants. In metallurgical leaching reactions, bacteria aid the oxidation of complex metal sulfides to form water-soluble metal sulfates. In addition, bacterial leaching has been used to extract other metal values from sulfide ores: zinc from sphalerite, lead from galena, and cobalt and nickel from base-metal sulfide ores. Nonmetal ores have also been processed with that method, particularly coal, which often

contains high concentrations of sulfur that cause air pollution when burned.

Besides enhancing or inhibiting the recovery of metal values from ores, bacteria are being used to precipitate or accumulate metal. The process, known as bioaccumulation, normally involves the adsorption of metal ions on the bacteria, which are then chemically transformed to an insoluble precipitate. *See* METAL AND MINERAL PROCESSING.

Medical applications. Diagnostic applications have developed rapidly, beginning with the introduction of diagnostic tests that use monoclonal antibodies. A monoclonal antibody is produced by a hybridoma, a cell formed by the fusion of an antibody-producing cell with a cancerous myeloma cell. Monoclonal antibodies specifically combine with their complementary antigen and form the diagnostic basis for many drug-monitoring tests, syphilis and rubella testing, blood testing, cancer screening, and pregnancy tests. *See* MONOCLONAL ANTIBODIES.

The most visible products of industrial microbiology are therapeutics for human health. Microbial synthesis is the preferred method of production for most health care drugs with complex chemistry. Microorganisms still have a remarkable ability for producing new commercial antibiotics, the largest class of drugs, and for continued yield improvement. With recombinant DNA technology, many proteins and polypeptides that are present naturally in the human body in trace amounts can be produced in large amounts during fermentation of recombinant microorganisms. Full regulatory approval for new biotechnology products usually takes 8 to 9 years from initial synthesis, but already the following products have been approved: recombinant insulin for the treatment of diabetes; recombinant factor VIII for hemophilia; somatropin human growth hormone for growth deficiencies in children; tissue plasminogen activator for acute myocardial infarction; interferon for hairy cell leukemia; a vaccine for hepatitis B; and a monoclonal antibody to counteract rejection of transplanted kidneys. Other areas of possible application are immunomodulators, including interferons, colony stimulating factors, and interleukins; blood proteins, including anticoagulants and thrombolytic agents; and monoclonal antibodies with specific anticancer activity. *See* ANTIBIOTIC; INSULIN.

Pharmaceutical agents produced from recombinant microorganisms must meet stringent purity criteria. These products must be completely free of microbial cells, have no more than 10 picograms of DNA contamination per human dose, and be free of viral contamination. Recombinant products are usually proteins that must be separated from other contaminating proteins through several different and complementary methods of purification. Licenses issued for their production are generally process-specific, in contrast to more simple antibiotics that can be precisely defined chemically.

Environmental applications. Microorganisms have important roles in the natural cycling of elements because of the wide range of microbial degradation activities. Microbial activities have long been the basis for sewage treatment facilities, and industrial and hazardous waste cleanup, or bioremediation, has become important. The usual first step in bioremediation is the isolation of microorganisms, often from an environment contaminated by the toxic compounds, that are capable of growing on and degrading a specific toxic chemical waste. Environmental control during bioremediation of wastewater or contaminated soil is crucial because proper nutrient supply, aeration, and mixing can greatly enhance degradation rates. *See* HAZARDOUS WASTE; SEWAGE.

Many recalcitrant synthetic compounds are released into the environment. Some of these persist in nature because of their low rates of biodegradation by indigenous microorganisms. Bioremediation successes have been achieved by using native bacteria to degrade petroleum products (such as kerosine, benzene, and toluene that have leaked from underground storage tanks), toxic chlorinated herbicides (such as 2,4-dichlorophenoxyacetic acid [2,4-D]), and toxic biocides (such as pentachlorophenol).

Food and agricultural applications. Some of the oldest and most established areas of industrial microbiology concern food and beverage products, such as the production and use of brewer's yeast and baker's yeast. Food and agricultural applications are more limited by cost constraints than are medical applications, with the cost of raw materials often the major economic factor. The beer industry, which produces the highest-value product of all the biotechnology industries, has developed cost-containing strategies such as spent mash and carbon dioxide recovery, efficient flocculation and clarification, nonaseptic fermentation, and continuous large-scale fermentation. *See* CHEESE; DISTILLED SPIRITS; MALT BEVERAGE; WINE.

The food industry and the detergent industry are the major users of industrial enzymes produced by microbial fermentation. Detergents, especially in Europe, often contain protein-degrading enzymes (proteases). In the food industry, amylases convert starch to glucose, and glucose isomerase converts glucose to fructose. *See* DETERGENT; FOOD MANUFACTURING.

Agricultural applications use technology similar to that of medical applications. There are many agricultural applications for genetic manipulation of plants and possibly animals. Plant cultures have been applied to the production of important natural compounds and the improvement and propagation of crop plants. Because the production of natural compounds by plant cells must compete with less complex extraction from whole plants, the only commercial plant-cell culture process thus far appears to be the production of the red dye shikonin in Japan. Mass propagation by tissue culture and improvement of plants have received more commercial attention. *See* BIOTECHNOLOGY; BREEDING (PLANT); GENETIC ENGINEERING; PETROLEUM MICROBIOLOGY. Roger Korus

Bibliography. E. Antébi and D. Fishlock, *Biotechnology Strategies for Life*, 1986; C. Brown, I. Campbell, and F. Priest, *Introduction to Biotechnology*, 1987; A. L. Demain and N. A. Solomon (eds.), *Biology of Industrial Microorganisms*, 1985; H. L. Ehrlich and

C. L. Brierley, *Microbial Mineral Recovery*, 1990; R. Hack (ed.), *The International Biotechnology Handbook*, 1989; D. P. Labeda, *Isolation of Biotechnological Organisms from Nature*, 1990; P. Prave et al. (eds.), *Fundamentals of Biotechnology*, 1987; I. K. Vasil (ed.), *Biotechnology: Perspectives, Policies and Issues*, 1987.

Industrial trucks

Manually propelled (handtrucks) or powered carriers for transporting materials over level, slightly inclined, or slightly declined running surfaces. Some industrial trucks can lift and lower their loads, and others can also tier them. In any event, all such trucks maintain contact with the running surface over which they operate and, except when towed by a chain conveyor, follow variable paths of travel as distinct from conveying machines or monorails.

In industry the principle of handling materials in unit loads has developed in parallel with the increased use of powered industrial trucks, particularly the forklift type. These mobile mechanical handling aids have removed the limitations that existed when the weight and size of a load for movement and stacking depended mainly on the ability of a person to lift it manually. The unit-load principle of materials handling underlies the skid-platform and the pallet-forklift methods of operation. Both methods, especially the latter, have revolutionized handling techniques and equipment and even production equipment. *See* MATERIALS HANDLING.

Arthur M. Perrin

Bibliography. Icon Group International, Inc., *The 2006–2011 World Outlook for Self-Propelled, Electric, Gasoline and Other Power Work Trucks, Fork Lifts, and Tractors Fitted or Not Fitted with Lifting and Handling Equipment*, 2005; D. E. Mulcahy, *Materials Handling Handbook*, 1998.

Industrial wastewater treatment

A group of unit processes designed to separate, modify, remove, and destroy undesirable substances carried by wastewater from industrial sources. United States governmental regulations have been issued for volatile organic substances, designated priority pollutants, aquatic toxicity as defined by a bioassay, and in some cases nitrogen and phosphorus. As a result, sophisticated technology and process controls have been developed for treating industrial wastewater. It is also necessary to implement better water management and waste minimization practices to reduce waste loads and operating cost of treatment facilities.

Wastewater characterization. Wastewater must be characterized to determine if it will meet regulatory requirements. Characterization is also necessary to collect the information needed to design processes that can effectively reduce and treat the wastewater. Before choosing the treatment technologies, some key points must be considered, including understanding the production processes and how wastewater is generated, measuring the wastewater volumes and flow rates, and sampling and analyzing the physical, chemical, and biological characteristics.

Water-pinch technology emerged as a significant new approach to the problems of industrial water reuse, wastewater minimization, and effluent treatment system design. This technology allows chemical, civil, and environmental engineers to analyze water-using processes before design and operation, as well as after, to minimize freshwater consumption and wastewater generation.

Zero emission, zero discharge, cleaner production, waste minimization, pollution prevention, design for environment, material substitution, and dematerialization are relevant terms. The U.S. Environmental Protection Agency (EPA), the United Nations Industrial Development Organization (UNIDO), and other national and international organizations have promoted these terms.

Source treatment. Wastewater streams that are toxic or resistant to conventional treatment methods (refractory) should be treated at the source, and there are a number of source treatment technologies available. Metal precipitation enables the removal of bioinhibitors or biotoxicants. Chemical oxidation removes bioinhibitors and improves biodegradation rates. Wet air oxidation, which oxidizes organic materials at high temperature and pressure [14 kilopascals (2000 lb/in.^{2}) and 288°C (550°F)], is restricted to very high concentrations of these substances. Macroreticular (macroporous) resins are specific for removing particular organic materials, and the resin is regenerated and reused. Membrane processes, particularly reverse osmosis, are high-pressure operations in which water passes through a semipermeable membrane, leaving the contaminants in a concentrate. Other source treatment processes include stripping, filtration, and carbon adsorption. *See* HAZARDOUS WASTE; MEMBRANE SEPARATIONS.

Pretreatment and primary treatment. These processes address the problems of equalization, neutralization, removal of oil and grease, removal of suspended solids, and precipitation of heavy metals (**Fig. 1**).

Equalization. Effective operation of a wastewater treatment plan requires controlling the flow and loading to the facility. Equalization is used to achieve this objective. If the wastewater flow rate is relatively constant, as in a pulp and paper mill, a constant-volume basin is used in which the concentration is equalized. If both flow and concentration change, a variable-volume basin is used in which the inflow is variable and the outflow is constant. In this way, the mass load is equalized. Many industries are subject to spills and accidental discharges, for example, black liquor spills in the pulp and paper industry. In these cases, a spill basin is provided to retain spills and discharge them at a controlled rate to the equalization basin.

Neutralization. Wastewater permits usually require a pH in the range of 6 to 9. Biological treatment

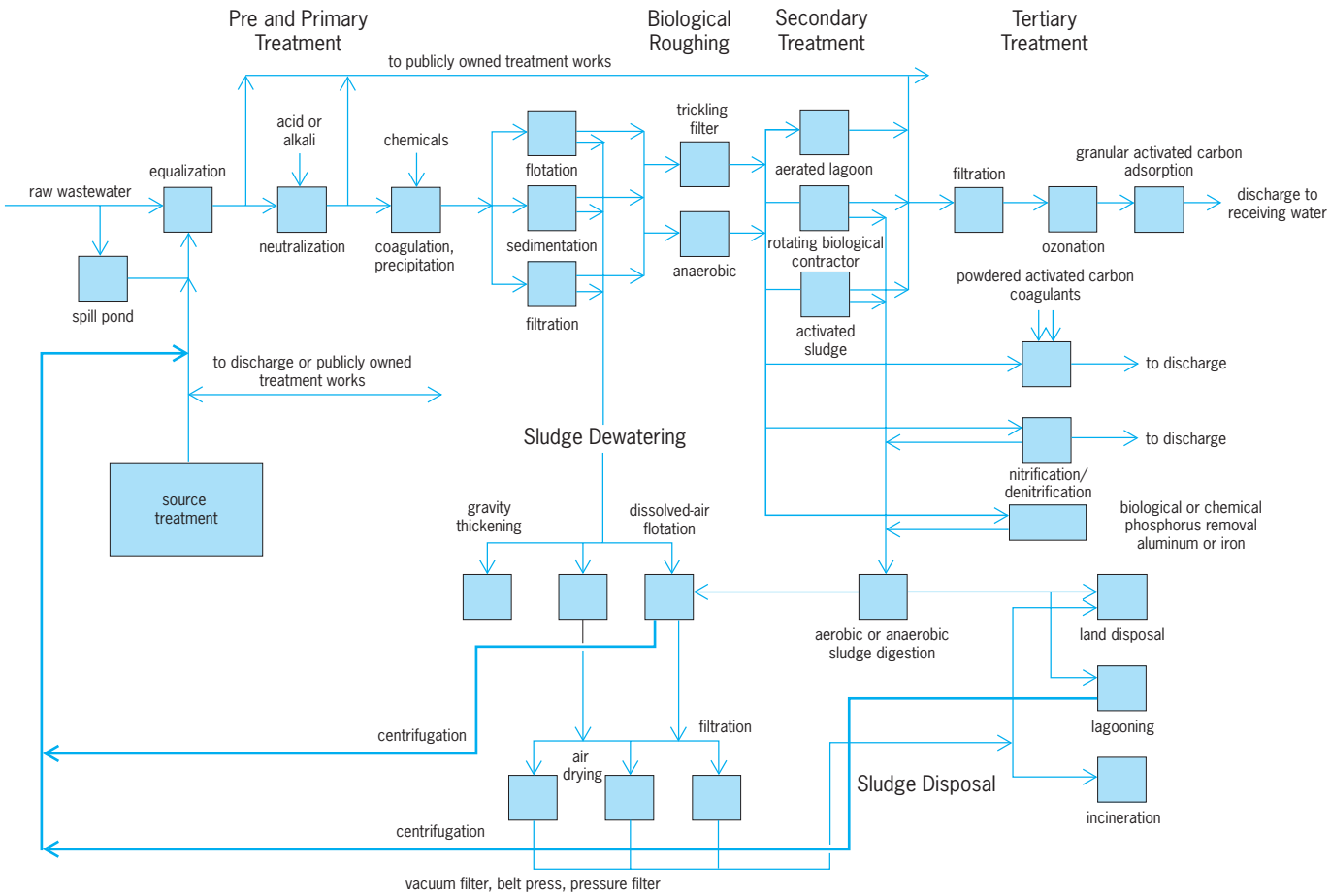

Fig. 1. Flowchart showing alternative wastewater treatment technologies. Flow returns are shown by heavy lines.

requires a pH of 6.5 to 8.5, and chemical coagulation processes have defined pH ranges. For this reason, acidic or alkaline wastewater streams require neutralization. Acidic wastewaters can be neutralized with lime ($CaOH_2$), caustic (NaOH), or magnesium hydroxide [$Mg(OH)_2$]. Magnesium hydroxide frequently has the advantage of producing less but more concentrated sludge. Since pH is a logarithmic function, highly acidic wastewaters require a two-stage neutralization process to maintain an effluent with a stable pH. Where the wastewater pH is fairly constant, a limestone ($CaCO_3$) bed can be used. Alkaline wastewaters are neutralized with sulfuric acid (H_2SO_4) or flue gas (CO_2), which forms carbonic acid when dissolved in water. *See* PH.

Oil, grease, and suspended solids. Free oil is removed by gravity separation. Large wastewater volumes containing oil, such as those issuing from a petroleum refinery, use an American Petroleum Institute (API) separator. The separator is a long, rectangular shallow basin in which the lower-specific-gravity oil rises to the surface and is skimmed off. Effluent oil concentrations are about 50 mg/liter. Lower-volume wastewater is treated in a corrugated plate separator that contains closely spaced plates, permitting a higher efficiency of oil removal. Effluent oil concentrations of 10 mg/liter are possible. In the case of emulsified oil, the emulsion must be broken using polymers or surfactants, enabling the creation of oily aggregates (flocs) that can be separated by physical means. *See* EMULSION; SURFACTANT.

Suspended solids can be removed by sedimentation or flotation. The efficiency of suspended-solids removal by sedimentation is related to the overflow rate [m^3/m^2 (gal/ft^2)], and for most wastewaters removal efficiencies will range 65–90%. Flotation is used to remove low concentrations of oil, grease, and nonsettleable solids. In dissolved air flotation, a portion of the effluent is pressurized in the presence of excess air, and the additional air goes into solution. When the pressurized liquid is mixed with the influent wastewater in the flotation unit at near-atmospheric pressure, the dissolved air comes out of solution in the form of small bubbles that attach themselves to and become enmeshed in globules and particles that rise to the surface and are skimmed off. Induced air flotation serves the same purpose, but it generates air bubbles through rotary inductors. *See* SEDIMENTATION (INDUSTRY).

Colloidal particles, turbidity, and color can be removed by coagulation. In this process, a coagulant, usually alum [$Al_2(SO_4)_3$] or iron(III) chloride [$FeCl_3$], is added to the wastewater. The positive aluminum and iron ions neutralize the negatively charged particles and generate a floc that is subsequently removed in a settling basin. *See* COLLOID; PARTICULATES.

Heavy-metal precipitation. Heavy metals are conventionally removed by precipitation. Most metals will precipitate as the hydroxide, carbonate, or sulfide. Carbonate and hydroxide precipitates are highly pH-dependent, according to the metal. Therefore, close pH control is essential. The sulfide is less pH-dependent. But because sulfide in the form of hydrogen sulfide (H_2S) presents an odor nuisance and a health hazard, sulfide precipitation is mainly used as a polishing step. In co-precipitation, heavy-metal ions are adsorbed onto an aluminum hydroxide or iron hydroxide floc. While this process is highly effective for removing most heavy metals, a large amount of sludge is generated that requires disposal. In the treatment of inorganic wastewater, emphasis is placed on metal recovery using techniques such as ion exchange, reverse osmosis, and electrochemical processes. A number of technologies are available for removing various metals from the wastewater stream. *See* ELECTROCHEMICAL PROCESS; ION EXCHANGE; PRECIPITATION (CHEMISTRY).

Biological treatment. Aerobic biological treatment is used to remove biodegradable organics. Depending on the effluent requirements, the nature of the wastewater, and the availability of land, several process options are available. In the activated-sludge process, the biomass generated is settled in a clarifier and recycled back to the aeration basin to build a high concentration of biomass. The process performance is related to the organic loading. Effluent concentrations of biodegradable organics in the order of 25 mg/liter are readily achievable. To maintain good sludge quality, a plug-flow configuration or batch treatment should be used in the case of readily degradable wastewaters, and a complete mix configuration for the more refractory wastewaters. Excess sludge generated in the process is disposed of from the final clarifier. Inhibition and poor settleability of the biosolids are common disturbances reported in cases of biological treatment of wastewater from petrochemical, pharmaceutical, paper, and other industries. In some cases, membrane separation technologies are used instead of sedimentation, and these systems are known as membrane biological reactors (MBR).

Where large land areas are available, an aerated lagoon system is applicable for treating nontoxic wastewaters, such as generated by pulp and paper mills. Three basins in series are used. The first basin is completely mixed in order to maximize biodegradation. In the second basin, at reduced power level, solids are permitted to settle, and then undergo anaerobic degradation and stabilization. A third settling basin is usually employed to reduce effluent suspended solids. The advantages of an aerated lagoon are simplicity of operation and the need to remove the sludge only about every 10 years.

Fixed-film processes include the trickling filter and the rotating biological contactor. In these processes, a biofilm is generated on a surface, usually plastic. As the wastewater passes over the film, organics diffuse into the film and are biodegraded. The trickling filter is primarily used as a pretreatment process for high concentrations of biodegradable organics in readily degradable wastewaters. Removal efficiencies of about 50% are usually achieved. The rotating biological contactor has been successfully used for treating wastewaters with low-to-moderate concentrations of biodegradable organics, such as generated by petroleum refineries and paper mills. *See* BIOFILM.

Anaerobic processes are sometimes used before aerobic processes for treating high-strength, readily degradable wastewaters. The primary advantages of the anaerobic process are low sludge production and the generation of energy in the form of methane (CH_4) gas. In the anaerobic process, organic materials are broken down to volatile organic acids, primarily acetic acid, which in turn is broken down to methane and carbon dioxide (CO_2). High-rate processes include the anaerobic filter, in which the anaerobic organisms are grown on a support medium; the anaerobic contact process, which is similar to the activated-sludge process and consists of a reactor, a clarifier, and a sludge recycle; and the upflow anaerobic sludge bed, which generates a granular sludge through which the wastewater is passed (**Fig. 2**). Where land area is available, low-rate processes have the advantage of greater process stability against spilling high concentrations of organics or accidentally dumping wastewater. *See* BIODEGRADATION; SEWAGE DISPOSAL; SEWAGE TREATMENT.

Physical-chemical treatment. Biological processes can remove only degradable organics. Nondegradable organics can be present in the influent wastewater or generated as oxidation by-products in the biological process. Many of these organics are toxic to aquatic life and must be removed before discharging the effluent. The most common technology used is adsorption on activated carbon. Granular carbon filters can be used as a posttreatment in which the refractory organics are adsorbed on the carbon. The granular carbon is regenerated for reuse. Granular media filtration is usually employed prior to the carbon columns. Alternatively, powdered activated carbon can be used in conjunction with the activated-sludge process. In this case, the powdered activated carbon is added prior to or into the activated-sludge aeration basin. The biomass degrades the degradable organics, and the carbon adsorbs the nondegradable organics. The carbon-biosludge can be regenerated or disposed of in a landfill. *See* ACTIVATED CARBON; ADSORPTION.

In some cases, toxic and refractory organics can be pretreated by chemical oxidation using ozone, catalyzed hydrogen peroxide, or advanced oxidation processes. In this case, the objective is not mineralization of the organics but detoxification and enhanced biodegradability.

Biological nitrogen removal, including both nitrification and denitrification, is used to remove ammonia from wastewaters. While this process is predictable in the case of municipal wastewaters, many industrial wastewaters inhibit the nitrifying organisms. For example, while nitrification will occur at

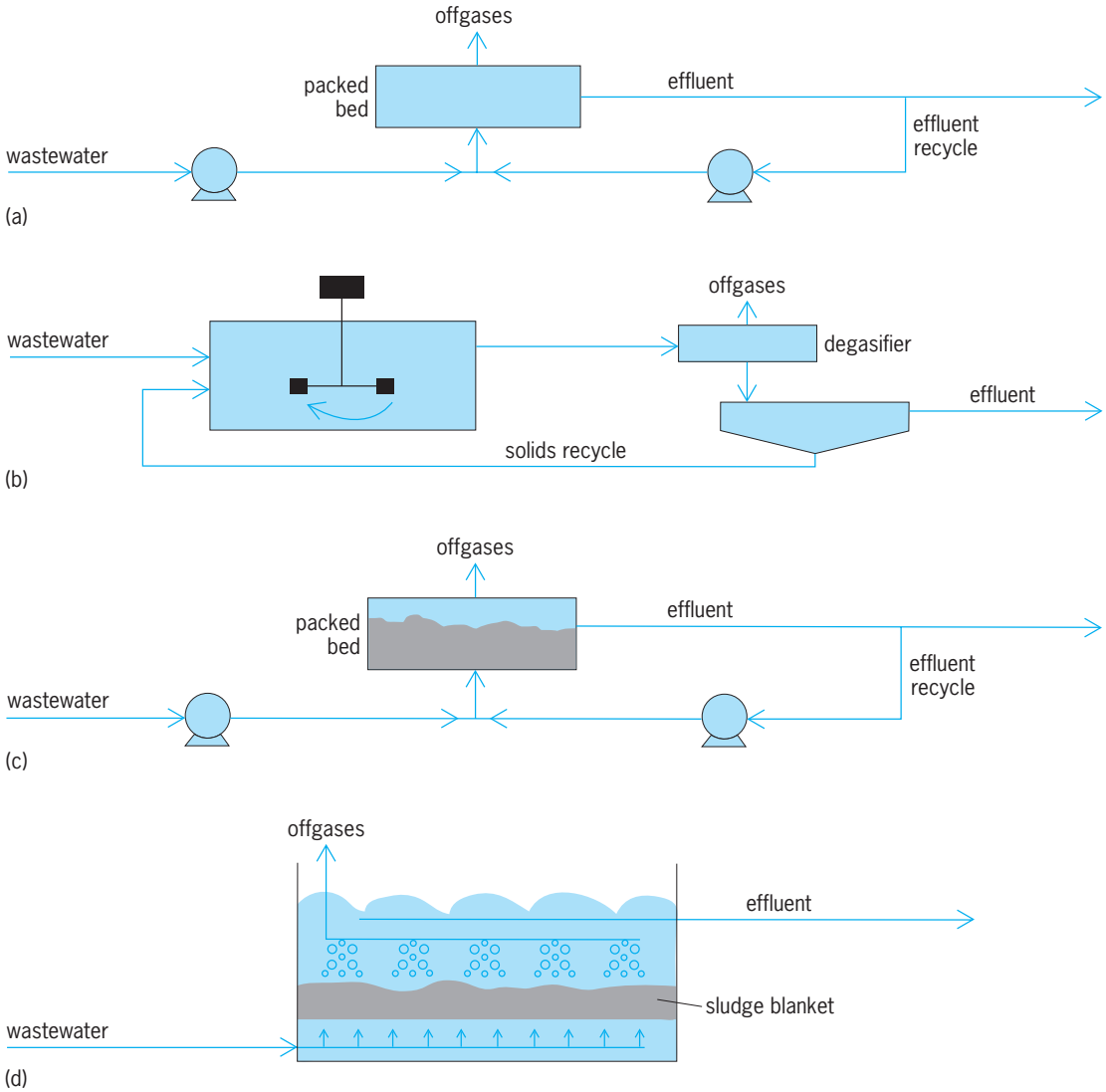

Fig. 2. Anaerobic wastewater treatment processes. (*a*) Anaerobic filter reactor. (*b*) Anaerobic contact reactor. (*c*) Fluidized-bed reactor. (*d*) Upflow anaerobic-sludge blanket.

a sludge age of 3 days at 20°C (68°F) in municipal wastewater, as much as 25 days may be required in the case of wastewater from a chemical plant. Sludge age is defined as the average length of time that the organism is under aeration. Inhibited conditions slow the growth rate of the organisms, and the sludge age must be increased. This step, in turn, increases the required aeration basin volume. In extreme cases of toxicity to the nitrifiers, powdered activated carbon has been required to adsorb the toxic molecules to permit nitrification to proceed. *See* BACTERIA.

Volatile organics can be removed by air or steam stripping. Air stripping is achieved by using packed or tray towers in which air and water counterflow through the tower. In cases where air stripping is used, the offgas containing the volatiles frequently must be treated by using vapor-phase carbon adsorption or a combustion chamber. In steam stripping, the liquid effluent from the column is separated as an azeotropic mixture. *See* AZEOTROPIC MIXTURE; STRIPPING.

Residuals management. Residuals are defined as any waste that is solid, liquid, or gas. Wastewater treatment operations usually produce sludge and solids. Disposal of those materials can be a major consideration when it comes to selecting wastewater unit processes. Every plant should perform residuals inventory, which should go back to the source of the residuals. A residuals management plan can be formulated to find the most cost-effective method for disposing of these wastes.

Sludge mixtures frequently contain less than 5% solids and therefore need to be thickened and dewatered for cost-effective disposal. Waste-activated sludge is usually stabilized by anaerobic digestion or by aerobic digestion, in which the degradable solids are oxidized by prolonged aeration. Ultimate disposal can be on land, by lagooning, or by incinerating the dewatered material. Return flows from thickening and dewatering are returned to the plant influent. It is frequently cost-effective to dry or evaporate sludges to reduce the volume and

increase heat content in case of incineration.

It is important to separate hazardous from nonhazardous sludges to minimize disposal costs. There are two types of landfills for industrial waste-treatment residuals: those that accept only nonhazardous waste and those that accept hazardous waste. Landfills are designed to prevent the contamination of ground water and the migration of contaminants from the landfill. Therefore, landfills must meet very stringent technical requirements. *See* SEWAGE SOLIDS.

Landfill leachates. Most landfill leachates have high and variable concentrations of organic and inorganic substances. All municipal and most industrial landfill leachates are amenable to biological treatment and can be treated anaerobically or aerobically, depending on the effluent quality desired. Activated carbon has been used to remove nondegradable organics. In Europe, some plants use reverse osmosis to produce a high-quality effluent. *See* WATER POLLUTION.

W. Wesley Eckenfelder; Noah I. Galil

Bibliography. J. B. Blackburn, *The Industrial Wastewater Systems Handbook*, 1998; W. W. Eckenfelder, *Industrial Water Pollution Control*, 3d ed., 1999; W. W. Eckenfelder, *Principles of Water Quality Management*, 1980, reprint 1991; J. D. Edwards, *Industrial Wastewater Treatment*, 1995; N. Galil, M. Rebhun, and Y. Brayer, Disturbances and inhibition in biological treatment of an integrated oil refinery, *Water Sci. Technol.*, 20(10):21–29, 1988; P. W. Lankford and W. W. Eckenfelder, *Toxicity Reduction in Industrial Effluents*, 1990; J. G. Mann and Y. A. Liu, *Industrial Water Reuse and Wastewater Minimization*, 1999; Metcalf & Eddy, Inc., G. Tchobanoglous et al., *Wastewater Engineering: Treatment, and Reuse*, 4th ed., 2003; R. K. Miller and M. E. Nemerow, *Industrial Wastewater Treatment*, 1991; J. W. Patterson, *Industrial Wastewater Treatment Technology*, 1985; L. K. Wang et al., *Handbook of Industrial and Hazardous Waste Treatment*, 2004; F. Woodard, *Industrial Waste Treatment Handbook*, 2001.

Inert gases

The inert gases, listed in the **table**, constitute group 0 of the periodic table of the elements. They are now better known as the noble gases, since stable compounds of xenon have been prepared. The term rare gases is a misnomer, since argon is plentiful and helium is not rare in the United States and some other countries.

The inert gases

Name	Symbol	Atomic number	Atomic weight
Helium	He	2	4.0026
Neon	Ne	10	20.183
Argon	Ar	18	39.948
Krypton	Kr	36	83.80
Xenon	Xe	54	131.30
Radon	Rn	86	(222)

All these gases occur to some extent in the Earth's atmosphere, but the concentrations of all but argon are exceedingly low. Argon is plentiful, constituting almost 1% of the air.

All isotopes of radon are radioactive, the longest lived having a half-life of about 4 days. Each of the other inert gases has at least two stable (nonradioactive) isotopes, in addition to one or more radioactive isotopes.

All the gases are colorless, odorless, and tasteless. They are all slightly soluble in water, the solubility increasing with increasing molecular weight. They can be liquefied at low temperatures, the boiling point being proportional to the atomic weight. All but helium can be solidified by reducing the temperature sufficiently, and helium can be solidified at temperatures of 0 to 1 K (−459.67 to −457.87°F) by the application of an external pressure of 25 atm (2.5×10^3 kilopascals) or more.

The noble gases are all monatomic. The outer shell of each of the atoms (unless strongly excited by radiation, electron bombardment, or other disturbing effects) is completely filled with electrons, and the gases are generally chemically inert. However, since the outer electrons in xenon are relatively far from the nucleus, they are held rather loosely; and they are capable of interacting with the outer electrons of fluorine atoms and the atoms of some other elements to form fairly stable compounds, such as XeF_2 and XeF_4. Radon can presumably form compounds similar to those of xenon. A fluoride of krypton which is stable at −8°C (17.6°F) has been prepared. No compounds of any of the lighter noble gases have been obtained. *See* XENON.

In an electric discharge, as in a mass spectrometer, very short-lived ions and molecules of all the noble-gas atoms can be formed; examples are Ar_2^+, HgHe, HgNe, HgAr, $(ArKr)^+$, and $(NeNe)^+$. In the presence of methane, $(XeCH_3)^+$ and other highly unstable ions can be formed.

All the inert gases except helium and radon are produced in concentrated form by the liquefaction and distillation of air, followed by special purification processes. Helium is obtained from certain natural gases containing 0.5% or more He. Radon is obtained by collecting the gas, called radium emanation, given off in the radioactive decay of radium. *See* AIR SEPARATION; ARGON; HELIUM; KRYPTON; NEON; PERIODIC TABLE; RADON; XENON.

A. W. Francis

Bibliography. V. G. Fastovskii, *Inert Gases*, 1967; B. L. Smith and J. P. Webb, *The Inert Gases: Model Systems for Science*, 1971.

Inertia

That property of matter which manifests itself as a resistance to any change in the motion of a body. Thus when no external force is acting, a body at rest remains at rest and a body in motion continues moving in a straight line with a uniform speed (Newton's first law of motion). The mass of a body is a measure of its inertia. *See* MASS.

Leo Nedelsky

Inertia welding

A welding process used to join similar and dissimilar materials at very rapid speed. It is, therefore, a very attractive welding process in mass production of good-quality welds. The ability to join dissimilar materials provides further flexibility in the design of mechanical components.

Principles. Inertia welding is a type of friction welding which utilizes the frictional heat generated at the rubbing surfaces to raise the temperature to a degree that the two parts can be forged together to form a solid bond. The energy required for inertia welding comes from a rotating flywheel system built into the machine. **Figure 1** illustrates the principle of the process. Like an engine lathe, the inertia welding machine has a headstock and a tailstock. One workpiece held in the spindle chuck (usually with an attached flywheel) is accelerated rapidly, while the other is clamped in a stationary holding device of the tailstock. When a predetermined spindle speed is reached, the drive power is cut and the nonrotating part on the tailstock is pushed against the rotating part under high pressure. Friction between the rubbing surfaces quickly brings the spindle to a stop. At the same time the stored kinetic energy in the flywheel is converted into frictional heat which raises the temperature at the interface high enough to forge the two parts together without melting.

During the process the material near the interface is softened and expelled from the original contacting surfaces to form a flash. This brings the nascent sublayers of the material close together to form a metallurgical bond. Therefore, the welding process is autogenous and carried out in solid state without the need of flux or filler material.

Weld characteristics. Inertia welds of similar materials consist of a thin layer of a severely deformed but highly refined grain structure at the weld interface. The heat-affected zone is very narrow as compared with any fusion welding process. For a properly welded joint, the weld strength is significantly higher than that of the base material. This is because during inertia welding the forging process improves the hardness of the material at the weld zone by means of work hardening. An inertia weld of carbon steels can have an average hardness at the weld some 25% higher than that of the base material. A tensile test will result in a fracture occurring at the base material instead of at the weld. If the specimen is annealed after welding, the interface would completely disappear as if the materials were not welded.

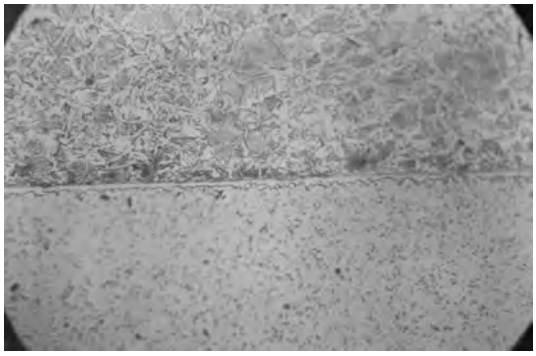

Fig. 2. Micrograph of a joint of SAE 1020 steel welded to tungsten carbide.

One of the principal advantages of inertia welding is its ability to join dissimilar metals which would be unweldable or uneconomical to join by any other means. **Figure 2** is a micrograph of a specimen of SAE 1020 steel welded to tungsten carbide. The process is also used to join aluminum to steel and stainless steel, aluminum to copper, copper to steel, copper to titanium, and zirconium alloys to each other. The weldability chart (**Fig. 3**) shows that some metal combinations can be welded to achieve strength at the weld equal to or better than the strength of the weaker material in the pair. In some cases, however, it may be necessary to perform an appropriate postweld heat treatment to realize the full weld strength. The chart indicates that there are other combinations that can be welded, but with strengths usually less than the strength of the weaker material.

The mechanisms which cause bonding of dissimilar metals are rather complex. Factors such as physical properties, surface energy, crystalline structure, mutual solubility, and possible formation of intermetallic compounds from the materials being welded may all play a role in the bonding mechanism. Since the key to achieving a metallurgical bond in solid state is to break up and remove the contaminated surface layers, metals with a substantial difference in hardness are generally more difficult to weld. In such cases, the hard material incurs practically no deformation, while the soft one has already been severely deformed or extruded. In

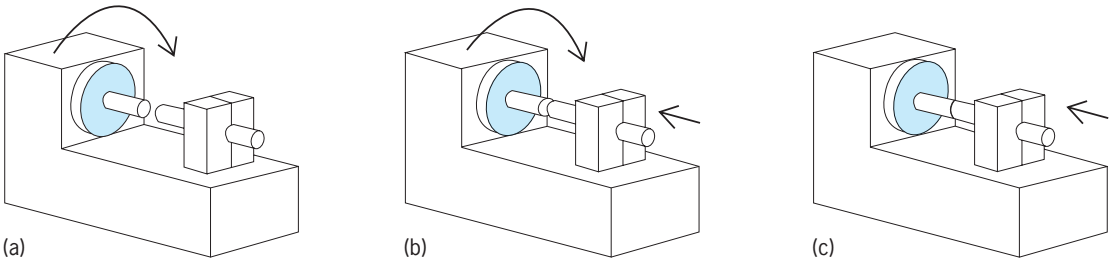

Fig. 1. Work sequence of inertia welding process. (*a*) Acceleration of flywheel and chuck-holding workpiece to preset speed. (*b*) At speed, drive is cut and workpieces are thrust together. (*c*) Workpieces bond together to stop flywheel rotation; weld is complete.

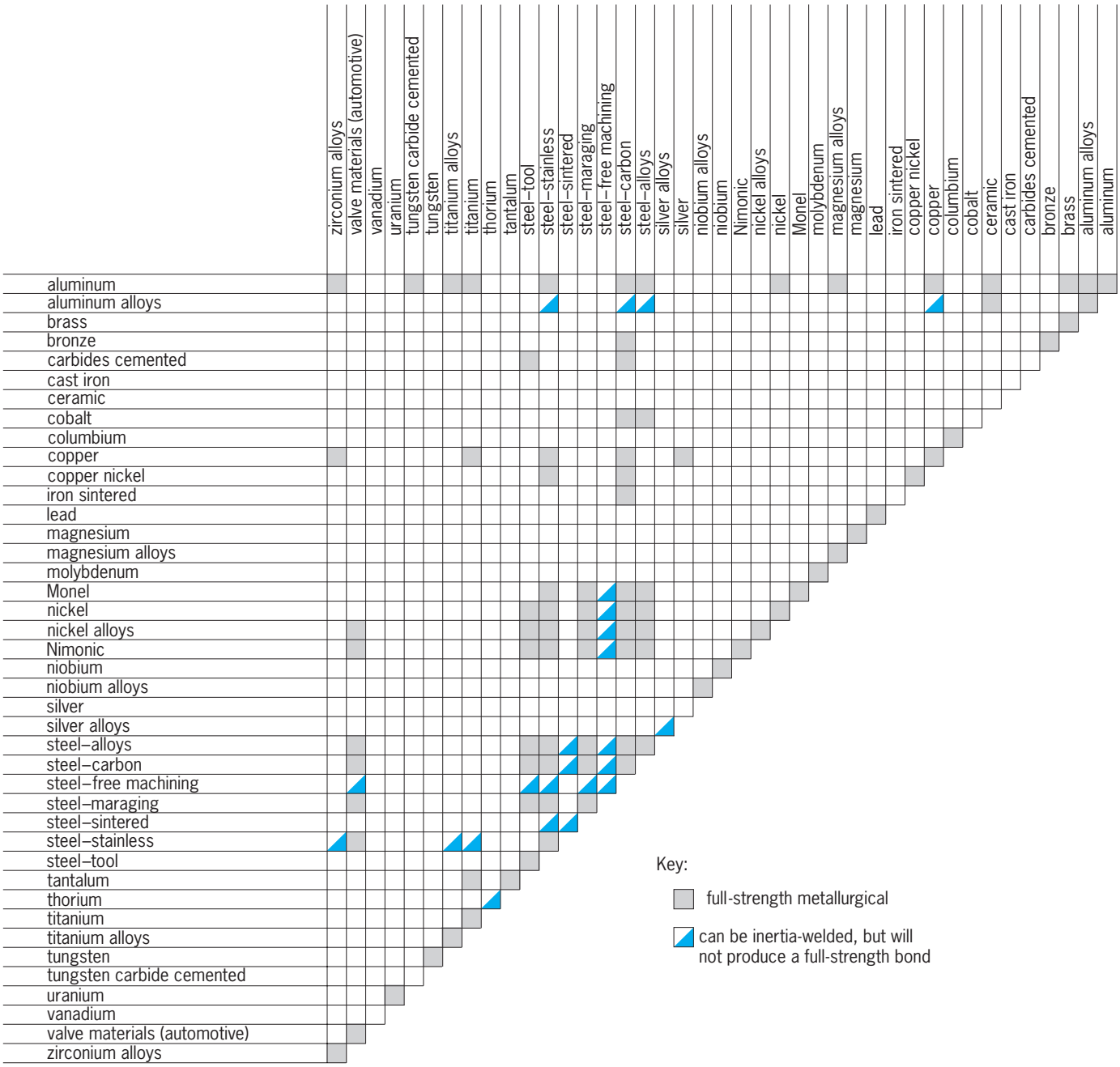

Fig. 3. Weldability between metals by inertia welding.

general, inertia welding can be applied with steels, aluminum, titanium, cobalt, nickel, and their alloys, copper, brass, and some bronzes, refractory materials, precipitation-hardened metal alloys, sintered materials, and heat-treated materials. The process cannot be successfully applied to cast irons, dry bearing materials, beryllium, and other materials where a distinct weak phase is present in the microstructure, such as in graphite, manganese sulfide, or there is a large amount of free lead or tellurium.

Applications. In addition to the advantages of high weld strength and short welding-cycle time, inertia welding machines can be readily automated for mass production. The automotive and truck industry is the major user of the process. Typical applications are engine valves (a medium-carbon-steel bar stem welded to a forged heat-resistant-alloy steel head), base-cup shock absorber assemblies, universal joints, axle housings, and hydraulic cylinder components. Auto makers, as well as other manufacturers, are constantly looking for ways to reduce the weight of components without sacrificing strength. An inexpensive way of achieving this goal is to replace standard open-die forgings or castings by weldments of precision closed-die forgings or castings welded to less expensive tubes, bar stocks, or stampings. Inertia welding of a tool pin joint to a drill pipe is one typical application in the oil drill industry (**Fig. 4**).

Other areas of application include welding of stainless steel to plain carbon steel for outboard motor

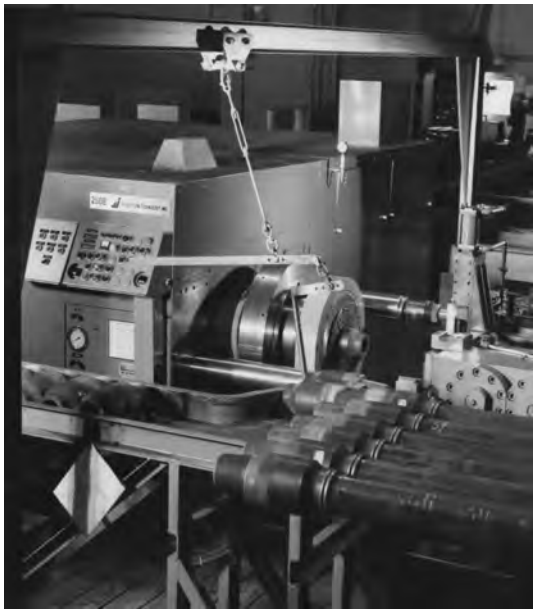

Fig. 4. Inertia welding machine for oil drill pipe sections.

shafts or electric motor shafts. Inertia welding is used to replace electron-beam welding for joining low-alloy steel shafts to nickel-base super-alloy turbine wheels or turbochargers in the aircraft industry. The process is also for making form-cutting tools by welding carbon-steel shanks to tool-steel heads. If one part has a round cross section and the materials are weldable by this process, the number of possible applications is virtually unlimited.

Other developments include design of multiple-spindle machines and radial inertia welding for joining a sleeve onto a ring in the radial direction. *See* METAL, MECHANICAL PROPERTIES OF; WELDED JOINT; WELDING AND CUTTING OF MATERIALS. K. K. Wang

Bibliography. American Welding Society, *Welding Handbook*, 8th ed., vol. 2, 1991; W. M. Hallett, *A Review of Inertia Welding*, SME Pap. AD76-090, 1976; K. K. Wang, *Friction Welding*, Welding Res. Counc. Bull. 204, April 1975.

Inertial navigation system

A self-contained system that can automatically determine the position, velocity, and attitude of a moving vehicle for the purpose of directing its future course. Based on prior knowledge of time, gravitational field, initial position, initial velocity, and initial orientation relative to a known reference frame (coordinate system), an inertial navigation system (INS) is capable of determining its present position, velocity, and orientation without the aid of external information. The generated navigational data are used to determine the future course for a vehicle to follow in order to bring it to its destination. Such systems have found application in the guidance and control of submarines, ships, aircraft, missiles, and spacecraft. The sensors making these measurements, based on one of the basic properties of mass—inertia, are gyroscopes (gyros) and accelerometers. *See* MASS.

Inertial navigation systems are used in many different types of applications that require vastly different performance from the gyros and accelerometers, and they are broadly classified by the quality (accuracy limitations) of their gyros and accelerometers (**Table 1**). Here, the gyro bias drift rate is given in units of degrees/unit time, and the accelerometer bias is given in terms of milli-g or micro-g, where $1\ g = 9.8\ \text{m/s}^2 = 32\ \text{ft/s}^2$.

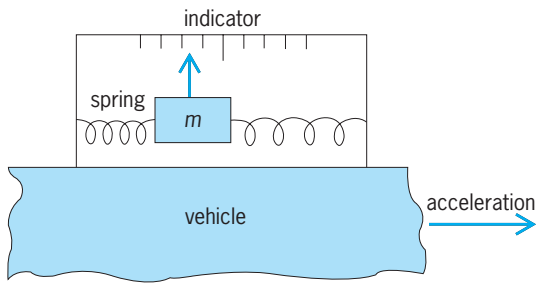

Fig. 1. Elementary accelerometer.

Principle of operation. An inertial navigation system integrates Newton's laws of motion to determine velocity and distance traveled from an initial position in three dimensions (vectorially), and continually updates these calculations. For a vehicle with initial velocity $\mathbf{v_o}$ at distance $\mathbf{s_o}$ from the starting point and acceleration $\mathbf{a}$, the subsequent velocity $\mathbf{v}$ and position $\mathbf{s}$ are given by Eqs. (1) and (2).

$$\mathbf{v} = \mathbf{v_o} + \int \mathbf{a}\,dt \quad (= \mathbf{v_o} + \mathbf{a}\,t \text{ for constant } \mathbf{a}) \qquad (1)$$

$$\mathbf{s} = \mathbf{s_o} + \mathbf{v_o}t + \iint \mathbf{a}\,dt^2$$

$$(= \mathbf{s_o} + \mathbf{v_o}t + {}^1\!/_2\mathbf{a}\,t^2 \text{ for constant } \mathbf{a}) \qquad (2)$$

See ACCELERATION; DYNAMICS; NEWTON'S LAWS OF MOTION; VELOCITY.

The basic measurement of an inertial guidance system is thus acceleration, which is measured by an accelerometer. A notional accelerometer is shown schematically in **Fig. 1**. The acceleration is determined by measuring the inertial force $\mathbf{F}$ required to

TABLE 1. Classification of inertial navigation systems

Classification	Application	Gyro bias drift rate	Accelerometer bias
Commercial grade	Automobiles	>1°/s	>50 mg
Tactical grade	Short-range missiles	1°/h	1 mg
Navigation grade	Aircraft	0.01°/h	25 μg
Strategic grade	Long-range missiles	0.001°/h	1 μg

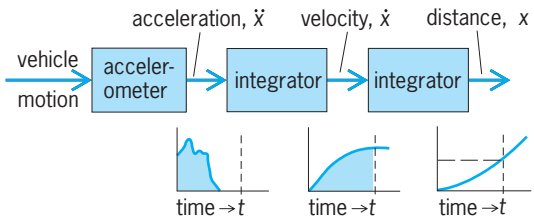

Fig. 2. Computation of distance by two integrations. Integration of acceleration gives velocity; integration of velocity gives distance.

move [or restore to its original (null) position] the proof mass m and inserting this measurement into the expression $\mathbf{a} = \mathbf{F}/m$. However, an accelerometer cannot distinguish between acceleration and gravitation. This can be visualized by considering a vehicle in free fall toward the Earth; an accelerometer would show zero output even though the vehicle is accelerating at 1 g. Conversely, an accelerometer sitting on a fixed table would indicate that the table is accelerating vertically at 1 g, whereas it is actually stationary. This means that gravity (gravitation plus centrifugal acceleration effects from the Earth's rotation) must be compensated in an inertial navigation system. *See* FORCE; RELATIVITY.

If three accelerometers are placed on a stable base and arranged so that their sensitive axes are, respectively, north, east, and down, and the vertical accelerometer is compensated for gravity, then any motion of the base can be sensed and its velocity and the position determined from these three measurements. The velocity is determined by integrating the accelerometer output, and the position is determined by integrating velocity (**Fig. 2**). This idealized concept is complicated by the necessity to maintain the base level and to stay pointed in one direction (leading to a gimbaled inertial system), or alternatively to precisely track the angular motion of the sensor about each axis and computationally determine the effective acceleration in each direction (leading to a strap-down inertial system). In either case, gyroscopes are used to determine the angular motion of the vehicle and to provide the attitude (vector direction) information. Depending on the type of gyro used, the output is either a change in angular rate (rate gyro) or change in angle (integrating gyro). The gyro characteristics, more than any other component, determine the quality of the inertial system. While various types of gyros can be used in either system, generally the floated integrating gyro and the dynamically tuned gyro find their major application in gimbaled systems, and the electrostatic, optical (ring laser and fiber optic), and micromechanical (MEMS) gyros are commonly used in strap-down applications. Depending upon which navigation coordinates are being used, the Earth's rotation (15°/h) has to be compensated in the gyro's output.

A precision clock is required to perform the integration of acceleration to velocity, velocity to distance, and angular rate to angle. *See* ACCELEROMETER; GYROSCOPE.

Gimbaled inertial system. There are several ways that a gimbaled system can be mechanized:

1. The system can be inertially stabilized, in which case the platform is kept fixed relative to inertial space (that is, relative to distant star directions), except for gyro drift.
2. In local level stabilization, the platform is kept aligned with a local level coordinate frame, which rotates with the Earth.
3. In rate biasing, the platform has a large rotation rate about an axis due to deliberate gimbal or vehicle rotation, in order to, for example, prevent laser gyro lock-in.
4. In thrust following, one axis of the platform is kept along the thrust vector of a vehicle.
5. In a carouseled or maytagged system, the platform is continuously rotated about one or more axes to reduce navigation errors due to gyro and accelerometer bias errors.
6. In an indexed system, the platform is rotated in discrete increments about one or more axes to reduce navigation errors by estimating gyro and accelerometer error terms or reducing the effect of bias errors.

For an inertially stabilized three-gimbal platform (**Fig. 3**), the inner gimbal or controlled member is the stable platform on which the three gyros and accelerometers are mounted. As the vehicle to which the system is attached moves, the gyros sense any rotation and drive the gimbals in the opposite direction to null the gyro rate. As a result, the inner platform remains fixed in inertial space. Most inertial platforms are mechanized, however, in the "local vertical" mode, which means that the perpendicular to the stable platform (inner gimbal) always points toward the center of the Earth. Under horizontal accelerations, the pendulum of the accelerometer measuring the vertical direction will tilt and indicate a false vertical. To correct for this, the integrated outputs of the north and east accelerometers (vehicle velocity) are divided by the radius of the Earth and used as an additional signal to drive the gimbal torquers.

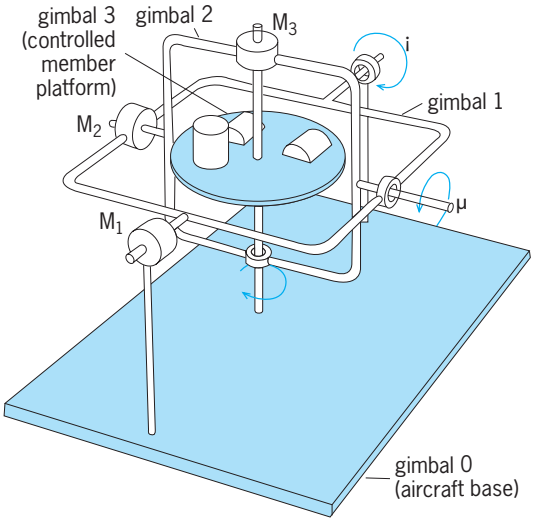

Fig. 3. Schematic of a three-gimbal platform.

This is known as a Schuler loop. If the output errors of such a system are observed, they will contain an oscillation with a period of the order of 84 min, known as the Schuler frequency. The vertical position channel of a stable platform is unstable; hence, the vertical accelerometer is sometimes deleted. *See* SCHULER PENDULUM.

Strap-down inertial system. In a strap-down inertial system, the three accelerometers and three gyros are hard-mounted to the vehicle (**Fig. 4**). As the vehicle moves, the gyro and accelerometer outputs are read approximately every 1/100 of a second, the new direction is established, and the velocity change and position change are calculated for the past 1/100 s. These changes are then summed with previous data to give total velocity and position changes. The computational burden is very heavy, especially for the coordinate transformations necessary to maintain the mathematical equivalent of the gimbal position in inertial space. Strap-down systems also need to be Schuler-tuned, where this operation is accomplished mathematically in the navigation computer. Two major contributors to the practicality of the modern strap-down system were the solid-state high-speed computer and the development of gyros that do not require torquers and have outputs that are insensitive to vehicle accelerations, such as the ring laser gyro. Nearly all applications now use strap-down systems. However, at present the most precise navigators are still gimbaled systems. In a strap-down system, the "vertical accelerometers" cannot be deleted since the accelerometers are tied to the vehicle and may assume any attitude. A schematic of the inertial navigation process for a strap-down system navigating in the local level (rotating with the Earth) coordinate frame is shown in **Fig. 5**.

Typical aircraft-quality strap-down navigators are relatively large (around 7500 cm^3 or 450 $in.^3$) and expensive. Navigators for short-range missiles and artillery shells are much smaller (150–500 cm^3 or 10–30 $in.^3$). **Figure 6** shows examples of strap-down inertial navigation systems.

Other terminologies. Navigation is only part of the process of arriving at the desired destination. The first requirement for achieving this objective is a knowledge of where you are and where your destination is (using maps or surveys). The second is navigation (using sensors or a navigation system). The third is guidance—that is, based on the navigation information, what needs to be done to get to the destination or target (using a guidance computer). Finally, there is the control stage, in which the guidance commands are implemented (using actuators). Often guidance and navigation are used interchangeably. Other terminologies commonly used in navigation are given in **Table 2**. *See* GUIDANCE SYSTEMS.

Aiding. Most inertial navigation systems incorporate some form of aiding (also called augmentation) to improve their accuracy by periodically updating the inertial navigation system to overcome its drift over time. Typical aiding systems could be one or more of the following: airspeed indicator, magnetic compass, magnetometer, odometer, baroaltimeter, infrared (IR) seeker, radar, Doppler radar, star tracker, Loran, or Global Positioning System (GPS). Kálmán filtering techniques are used to update the navigation state to provide the "optimal" solution. **Figure 7** shows a schematic of an automobile navigation system with several augmentation systems. *See* AIRSPEED INDICATOR; ALTIMETER; DOPPLER RADAR; ESTIMATION THEORY; INFRARED RADIATION; LORAN; MAGNETIC COMPASS; MAGNETOMETER; RADAR; SATELLITE NAVIGATION SYSTEMS; STAR TRACKER.

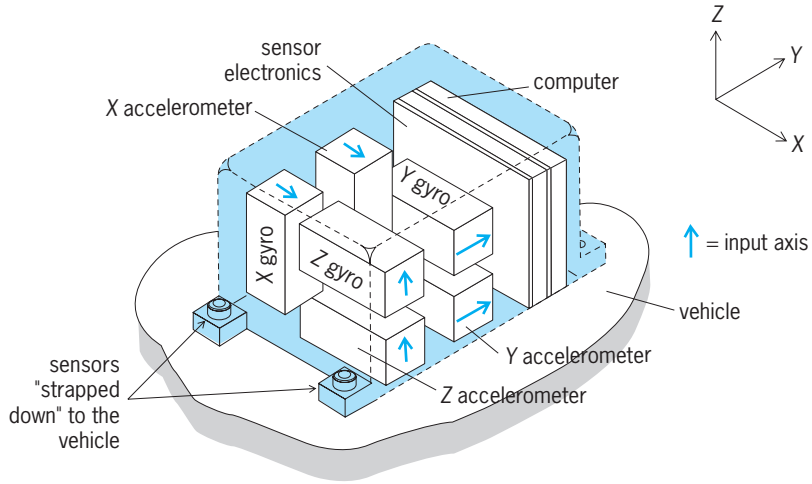

Fig. 4. Strap-down inertial navigation system. (*After A. Lawrence, Modern Inertial Technology, 2d ed., Springer, 1998*)

System initialization. When an inertial system is first turned on, it must be initialized. This is commonly done by electronically inserting the present latitude and longitude into the navigation computer via a key pad or from the GPS signal. The inertial unit itself must determine the vertical since its precise alignment relative to the vehicle is unknown, as is the precise alignment of the vehicle itself. In a gimbaled system, the vertical is determined by nulling the north and east accelerometer signals (that is, leveling the platform). The north alignment is done by

TABLE 2. Navigation terminologies

Term	Definition
Inertial sensor assembly (ISA)	A structure containing multiple inertial sensors in fixed orientations relative to one another
Attitude heading reference system (AHRS)	Estimates the body angles of a vehicle with respect to the local level coordinate system; performs direction finding (for example, north), but is not a navigator
Inertial reference unit (IRU)	Measures angular motion in three dimensions without external reference (usually used in satellites)
Inertial measurement unit (IMU)	Measures linear and angular motion in three dimensions without external reference (provides output from an ISA to the navigation computer)

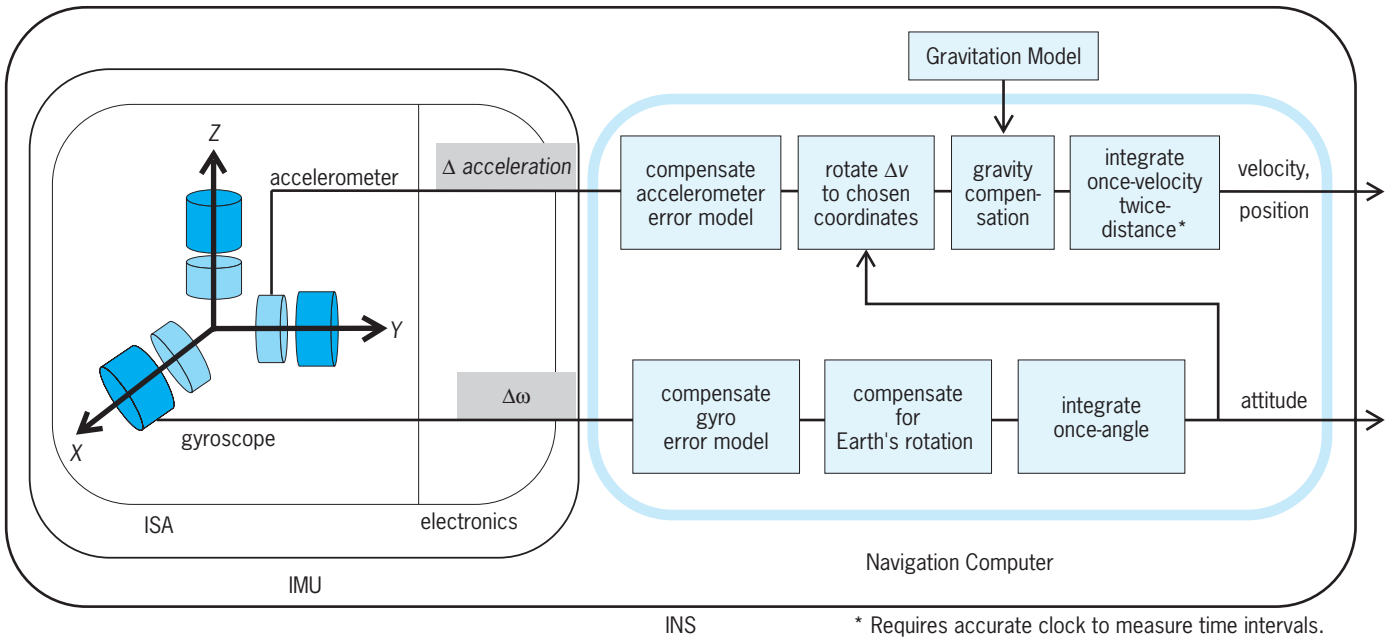

Fig. 5. Inertial navigation process. The inertial sensor assembly (ISA) and inertial measurement unit (IMU) are defined in Table 2.

gyrocompassing the platform until no Earth rate is sensed by the east gyro and also until the north and east accelerometers have zero output. This process may take from 10 to 30 min, depending upon the quality of the instruments and the vehicle's latitude. Because of this alignment technique, the inertial system aligns north instead of magnetic north in order to be parallel to the Earth's spin vector. In strap-down systems, the mathematical equivalent of aligning the gyros and accelerometers takes place in the computer. Down is defined to be along the accelerometer triad vector (gravitation), east is defined to be along the vector cross product of the gyro triad vector (Earth's rotation) and up, and north is defined to be the vector cross product of the first two. In many military systems, the vehicle is initialized with a transfer alignment from a more accurate navigation system (for example, from an aircraft's navigation system to the inertial navigation system of the missile being carried). *See* CALCULUS OF VECTORS.

Fig. 6. Strap-down inertial navigation systems. (*a*) Mock-up of system based on laser gyro technology. The three laser gyros, hermetically sealed, are on the left (*Honeywell*). (*b*) 130-cm^3 (8-in.^{3}) system that uses silicon MEMS technology (*Draper Laboratory*).

System accuracies. The accuracies of aircraft inertial systems are usually stated as a drift rate in nautical miles per hour. Typical commercial systems such as those that fly in transoceanic airliners have accuracies slightly better than 1 nautical mile per hour (1.85 km/h); therefore, in a 5-h flight the total expected error would be less than 5 nmi (9.3 km), which would allow other navigation aids to be acquired at the destination. The major contributor to this error is gyro drift; as a rule of thumb, a gyro with a drift rate of 0.01° per hour could be expected to produce an error of 1 nmi/h when incorporated into an inertial system. Below a capability of approximately 0.25 nmi/h (0.5 km/h), the uncertainties in the gravitational field vector begin to be significant sources of error. Nevertheless, systems have been built which consistently exhibit errors less than this.

For a short-range-missile inertial navigation system with an accelerometer error of 1 milli-*g* and perfect gyros, the distance error over a 1-min flight would be around 20 m (60 ft).

Aiding the navigation system by using information from the GPS allows improved navigation accuracy when using lower-performing gyros and

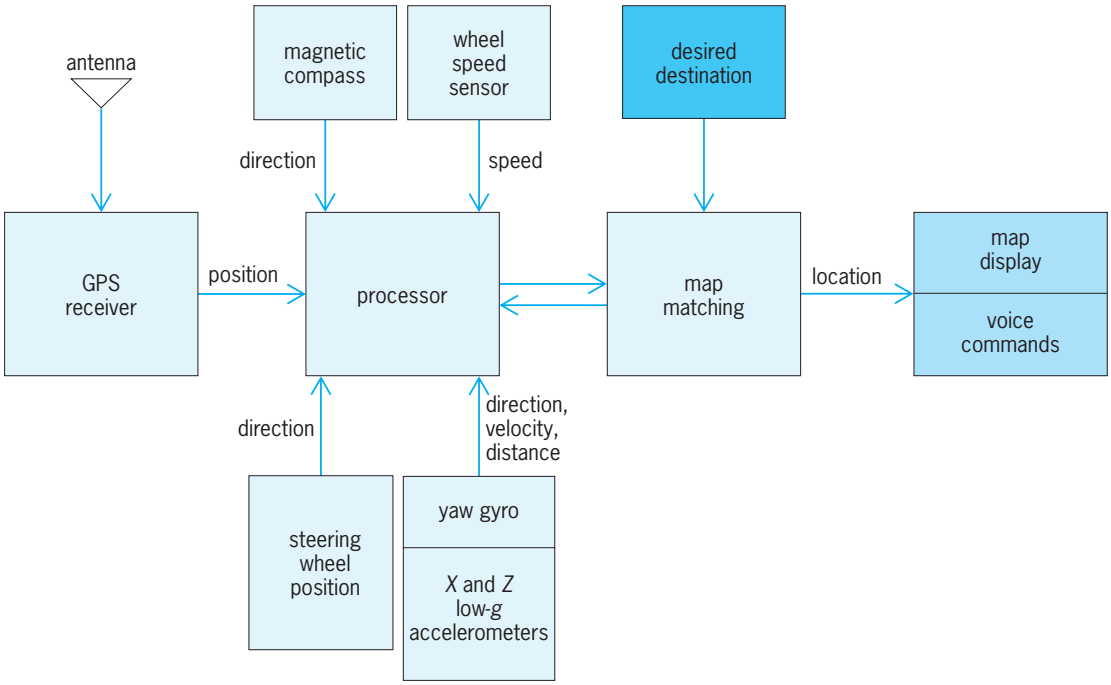

Fig. 7. Automobile navigation system.

accelerometers, and has been responsible for the introduction of inertial navigation to many new applications, such as automobiles and guided bombs. Each technology counteracts the weakness of the other. GPS errors are bounded, but updates are available at a lower rate, and there is the possibility of jamming or of signals becoming unavailable (for example, in parts of city centers or in buildings). Inertial navigation system errors are unbounded (grow over time), but they provide higher rate updates, and cannot be jammed. Also, the inertial navigation system can perform the autopilot function (stabilization and control) for the vehicle. Most navigation systems are no longer purely inertial but are now INS/GPS systems, and would not be able to achieve the desired performance with the inertial navigation system alone. The 1-nmi/h aircraft navigator described above, when coupled with GPS, would be accurate to the order of meters per hour. Neil M. Barbour; William E. Howell

Infant diarrhea

Diarrhea and its complications are the most important causes of infant death in most developing regions. The causes of the illness vary from dietary incompatibilities to intestinal infection. The most important infectious causes are, in approximate order of importance, rotavirus, the bacteria *Campylobacter jejuni*, *Shigella* (often causing dysentery or passage of bloody stools), *Salmonella*, and *Escherichia coli*, and the parasites *Giardia lamblia*, *Cryptosporidium*, and *Entamoeba histolytica*. The microorganisms are spread to the infant from other persons or through food or water. Breast-feeding is associated with a decreased occurrence of diarrhea and represents a major means of preventing infantile diarrhea in the developing world.

Since dehydration is the complication which must be prevented in diarrhea, attention to fluid and electrolyte replacement is the cornerstone of therapy. Oral electrolyte solutions administered under supervision are generally effective even in severely dehydrating forms of diarrhea. In a small percentage of cases, oral fluids cannot keep up with body losses, particularly if vomiting complicates the illness. In these situations, hospitalization and intravenous fluid and electrolyte replacement are lifesaving. Antibiotics are useful in the therapy of infants with more serious forms of intestinal infection. Hospital nursery outbreaks represent serious public health problems, occurring because of the appearance of infectious organisms in the hospital environment together with the presence of a confined nursery population of infants which are highly susceptible to these agents. *See* BACILLARY DYSENTERY; DIARRHEA; GIARDIASIS. Herbert L. Dupont

Infant respiratory distress syndrome

A common disorder of premature (or preterm) birth, defined by respiratory difficulty, which requires oxygen or assisted ventilation, and characteristic changes on chest x-ray. It is primarily the result of deficiency in surface-active lipids (contained within pulmonary surfactant) which serve to stabilize the air spaces (alveoli) within the lung. Although most cases of infant respiratory distress syndrome rarely result in a fatal outcome in developed countries, all require specialized care in a neonatal intensive care unit.

Incidence and pathophysiology. The incidence of respiratory distress syndrome increases with greater prematurity as well as with maternal diabetes and asphyxia around the time of delivery. In contrast,

growth retardation, other prolonged stress in the uterus, and maternal steroid administration decrease the incidence of infant respiratory distress syndrome. In addition, male infants are more predisposed to the complications of prematurity, including infant respiratory distress syndrome. *See* DIABETES.

The disease was initially termed hyaline membrane disease because autopsies demonstrated a membranous material lining the air spaces of these lungs, which are also very solid secondary to poor aeration. However, these membranes are not the cause of poor gas exchange, but are the result of lung injury which produces a layer of debris from blood products and injured lung epithelial cells. Rather, the primary cause of respiratory distress syndrome is deficiency of pulmonary surfactant. Pulmonary surfactant is manufactured in specialized lung cells and secreted to line the surfaces of alveoli. It has unique physical properties that lower surface tension and prevent the alveoli from collapsing. Surfactant deficiency and the resultant collapse of alveoli cause the lung to fail in its primary role of oxygen uptake and carbon dioxide elimination. Severe untreated infant respiratory distress syndrome therefore results in respiratory failure. Moreover, failure of lung aeration has cardiovascular consequences secondary to constriction of pulmonary blood vessels and decrease in blood flow to the lungs. Respiratory failure results in poor oxygenation of tissues and impaired metabolism throughout the body.

Clinical features. Infants with respiratory distress syndrome show signs of respiratory difficulty (including rapid respiration, retractions of the rib cage, nasal flaring, grunting or whining upon exhaling, and a bluish discoloration of the skin and mucous membranes) if adequate supplemental oxygen is not provided. Confirmation of the diagnosis using radiographic methods is essential, because a variety of other disorders (such as infection) may have a similar clinical picture. A characteristic x-ray shows a uniform, diffuse haziness of both lungs due to their poor aeration. Biochemical confirmation of surfactant deficiency in lung secretions would ideally confirm the diagnosis; however, this is not practical or widely available. Infant respiratory distress syndrome is usually well advanced 5–7 days after birth; however, complications of the disease and its therapy may prolong this course in the smallest and sickest babies.

Prevention and treatment. Prevention of prematurity and administration of good prenatal care reduce the incidence of infant respiratory distress syndrome. If fetal well-being is in jeopardy and the mother must undergo induced labor before the end of her term, pulmonary maturity of the fetus can be determined by analysis of amniotic fluid (which is in part composed of fetal lung fluid), sampled by means of amniocentesis. Corticosteroids given before birth to the mother at risk for premature delivery is a well-established therapy for reducing the incidence of respiratory distress syndrome; this treatment enhances pulmonary maturation in the infant.

Surfactant therapy is now widely used for treatment of infants with moderate or severe respiratory distress syndrome. This requires introducing a tube into the trachea for direct placement of surfactant into the infant's lungs; this is done one or more times over the first 48 h of life. Both synthetic and animal based preparations of surfactant are effective in reducing the mortality of preterm infants, and are free of known complications.

Surfactant therapy must be performed in conjunction with conventional treatment such as maintenance of an optimal-temperature environment and use of intravenous fluids. Oxygen is administered directly to the spontaneously breathing infant or by means of assisted ventilation. Blood gas status is closely checked by noninvasive monitoring combined with frequent blood sampling. Throughout the course of respiratory distress syndrome, complications of this disease, its management, or prematurity in general may arise at any time.

Prognosis. Infants with respiratory distress syndrome can be successfully treated, especially with the widespread use of surfactant therapy. However, the outcome remains heavily influenced by the infant's degree of prematurity. Very immature infants and those with severe respiratory distress syndrome may develop a more prolonged form of neonatal respiratory disease called bronchopulmonary dysplasia. This has been attributed to lung injury caused by exposure of an immature respiratory system to high inspired oxygen. Thus, the most severe cases of infant respiratory distress syndrome, requiring the most intensive oxygen and ventilator therapy, are consequently the most prone to prolonged lung injury. Fortunately only a small proportion of these infants remain symptomatic throughout childhood. *See* RESPIRATORY SYSTEM DISORDERS. Richard J. Martin

Bibliography. M. E. Avery and J. Mead, Surface properties in relation to atelectasis and hyaline membrane disease, *Amer. J. Dis. Child.*, 97:517–523, 1959; A. A. Fanaroff and R. J. Martin (eds.), *Neonatal-Perinatal Medicine*, 5th ed., 1992; T. Fujiwara et al., Artificial surfactant therapy in hyaline membrane disease, *Lancet*, 1:55–59, 1980; A. H. Jobe, Pulmonary surfactant therapy, *N. Engl. J. Med.*, 328:861–868, 1993; M. H. Klaus and A. A. Fanaroff (eds.), *Care of the High-Risk Neonate*, 5th ed., 2001.

Infarction

The process of anoxic tissue death. The usual cause is occlusion of an artery by a thrombus or embolus and sometimes by severe atherosclerosis. The development of the infarct depends to a great extent on the collateral circulation. If the collateral blood supply is inadequate or if the vessel is the sole source of blood supply to the region, an infarct results. *See* ARTERIOSCLEROSIS; THROMBOSIS.

Following an initial period of hyperemia, the infarcted region becomes pale. Because of the branching arrangement of the vascular system, infarcts of the spleen, kidney, and lung are usually wedge-shaped, with the base of the wedge at the periphery and the apex toward the point of the vascular obstruction (**Fig. 1**).

With occlusion of an artery the existing collateral

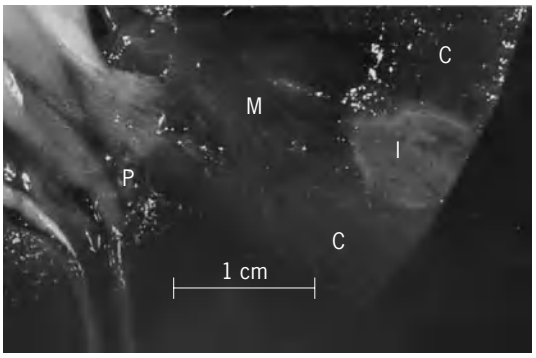

Fig. 1. Magnified gross photograph of a wedge-shaped infarct of kidney (I). Renal cortex (C), medullary pyramid (M), and renal pelvis (P) can be seen.

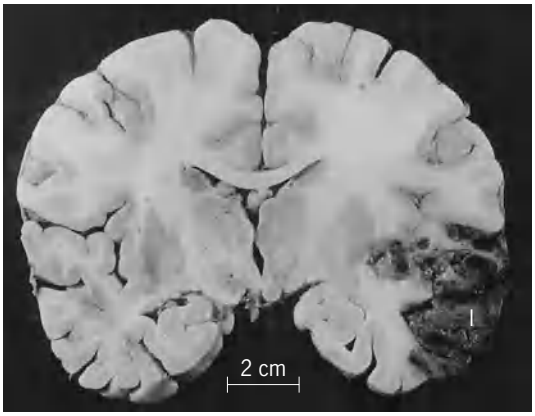

Fig. 2. Gross photograph of a brain section, showing a wedge-shaped infarct (I) that resulted from a thrombus. Secondary hemorrhage has occurred.

circulation attempts to bring more blood to the region. Depending on the extent of the process, the infarcted area will be red or pale. This process varies with the organ involved. Those of the kidney and heart are pale, whereas those of the spleen and lung are usually red. With death of the tissue a secondary hemorrhage may occur (**Fig. 2**).

The tissue in the affected area undergoes coagulative necrosis and dies while the tissue around the infarcted region becomes inflamed. In time the dead tissue is replaced by a shrunken depressed scar. *See* INFLAMMATION.

Infarcts commonly occur in the lungs, heart, brain, spleen, and kidneys. The common cause of infarcts of the heart is thrombosis of the coronary artery, usually secondary to atherosclerosis. Hemorrhage into an atherosclerotic plaque can also result in thrombosis of an artery.

Myocardial infarction is usually accompanied by severe prolonged substernal oppression or pain, electrocardiographic changes, fever, leukocytosis, and serum enzyme changes. Myocardial infarcts usually involve the left ventricle or interventricular septum and only rarely involve the right ventricle or atria. The location of the infarct usually depends on the coronary artery occluded. If death does not occur immediately, the infarct heals and is replaced by scar tissue in 5–8 weeks, depending on its size.

Complications of myocardial infarction include rupture of the necrotic portion of the heart with hemorrhage into the pericardium and rapid death; mural thrombi which form on the endocardial surface of the healing infarct and which may give rise to embolism; congestive heart failure as a result of an insufficient amount of remaining healthy muscle to carry on the work of the heart (occurring most commonly with repeated infarcts); and disturbances in heart rhythm if the pacemaker or conduction system is involved.

Infarction of a portion of bowel will result in death unless surgical intervention is forthcoming. Embolization to the lungs is a rather frequent occurrence. However, because of the collateral blood supply of the lungs, infarction follows only when there is some interference with the circulation, such as chronic pulmonary venous congestion. An extensive collateral circulation also exists in the liver, hence the rarity of infarcts in this organ. *See* CIRCULATION DISORDERS. Romeo A. Vidone; Irwin Nash

Bibliography. E. Braunwald et al. (eds.), *Harrison's Principles of Internal Medicine,* 15th ed., 2001; R. S. Cotran et al., *Robbins' Pathologic Basis of Disease*, 6th ed., 1999; J. M. Kissane (ed.), *Anderson's Pathology*, 9th ed., 1989; M. D. Silver, *Cardiovascular Pathology*, 2 vols., 3d ed., 2000.

Infection

A term considered by some to mean the entrance, growth, and multiplication of a microorganism (pathogen) in the body of a host, resulting in the establishment of a disease process. Others define infection as the presence of a microorganism in host tissues whether or not it evolves into detectable pathologic effects. The host may be a bacterium, plant, animal, or human being, and the infecting agent may be viral, rickettsial, bacterial, fungal, or protozoan.

A differentiation is made between infection and infestation. Infestation is the invasion of a host by higher organisms such as parasitic worms. *See* EPIDEMIOLOGY; HOSPITAL INFECTIONS; MEDICAL BACTERIOLOGY; MEDICAL MYCOLOGY; MEDICAL PARASITOLOGY; OPPORTUNISTIC INFECTIONS; PATHOGEN; VIRUS. Daniel N. Lapedes

Infectious disease

Infectious diseases are any of those pathological conditions spread among all biological species. Infectious diseases, although varied in their effects, are always associated with viruses, bacteria, fungi, protozoa, multicellular parasites and aberrant proteins known as prions. A complex series of steps, mediated by factors contributed by both the infectious agent and the host, is required for microorganisms or prions to establish an infection or disease. Worldwide, infectious diseases are the third leading cause of human death.

Pathogenic microorganisms. It is difficult to group microorganisms as exclusively commensal or pathogenic species. The most common relationship

between a host and a microorganism is a commensal one, in which advantages exist for both organisms. For example, hundreds of billions of bacteria of many genera live in the human gastrointestinal tract, coexisting in ecological balance without causing disease. These bacteria help prevent the invasion of the host by more virulent organisms such as *Salmonella*, *Shigella*, or *Campylobacter* species. In exchange, the host provides an environment in which harmless bacteria can readily receive nutrients to fulfill their metabolic needs. There are very few microorganisms that cause disease every time they encounter a host. Instead, many factors of both host and microbial origin are involved in infectious disease. These factors include the general health of the host, previous exposure of the host to the microorganism, and the complement of molecules produced by the bacteria. Pathogenic microorganisms are frequently categorized by the host site where they most commonly cause disease. For example, *Bacteroides*, *Prevotella*, and *Haemophilus* species are usually associated with respiratory tract infections, while poxviruses and *Trichophyton* species are frequently encountered on the skin. Some microorganisms, such as *Staphylococcus* species, can cause a myriad of infections involving the respiratory, gastrointestinal, and genitourinary tracts, as well as skin, eyes, or ears.

Infectious disease transmission. Spread of a pathogenic microorganism among individual hosts is the hallmark of an infectious disease. This process, known as transmission, may occur through four major pathways: contact with the microorganism, airborne inhalation, through a common vehicle such as blood, or by vector-borne spread.

Contact spread. In contact-spread disease, the host is in direct or indirect contact with the infectious agent. Most sexually transmitted diseases are spread by direct contact with the microorganism from an infected host. Indirect transmission occurs via an inanimate object that carries the infecting organism, as in spread of infections resulting from the sharing of needles among intravenous drug abusers. Another route of transmission involves spread by droplets, relatively large particles coughed or sneezed into the air. The droplets are usually broadcast within a few feet of the infected host and transmitted to a susceptible individual nearby; examples include streptococcal pharyngitis (strep throat) and measles. *See* STREPTOCOCCUS.

Airborne spread. Transmission through the air (beyond a few feet from the infected host) occurs when microorganisms attach to dust particles or when they are contained in aerosols. Particles may remain suspended in the air for hours or days, and if they are of the appropriate size they can be inhaled. Tuberculosis (*Mycobacterium tuberculosis*) and chickenpox (varicella zoster virus) are examples of diseases transmitted by the airborne route. *See* CHICKENPOX AND SHINGLES; TUBERCULOSIS.

Common-vehicle spread. Persons can become infected following contact with a contaminated common vehicle, such as food, water, or blood. Poisoning by *Escherichia coli* toxins, mycotoxins (fungal toxins), *Clostridium botulinum* toxins, and the *Staphylococcus aureus* enterotoxins is primarily through the ingestion of contaminated food products. The infectious agents of spongiform encephalopathies, such as mad cow disease, are known as prions and are transmitted through contaminated feed. Typhoid fever, cholera, and amebic dysentery are examples of illnesses acquired through contaminated drinking water. Hepatitis B and acquired immune deficiency syndrome (AIDS) are examples of viral diseases spread by transfusion with infected blood. *See* BOTULISM; FOOD POISONING; HEPATITIS; WATERBORNE DISEASE.

Vector-borne spread. Transmission of microorganisms can occur through an intermediate vector, such as an insect of arthropod. Often the intermediate host is a necessary agent, serving an essential biological role in the microorganism's life cycle. For example, mosquitoes spread the malaria-causing protozoon *Plasmodium falciparum* and the yellow fever arbovirus as they transmit infected blood from one mammal to another. Spotted fever, caused by *Rickettsia rickettsii*, and Lyme disease, caused by the spirochete *Borrelia burgdorferi*, are transmitted via the bite of an infected tick. *See* LYME DISEASE; MALARIA; YELLOW FEVER.

Pathogenesis of infection. The manner in which an infectious disease develops, or its pathogenesis, usually follows a consistent pattern. To initiate an infection, there must be a physical encounter as which the microorganism enters the host. The most frequent portals of entry are the respiratory, gastrointestinal, and genitourinary tracts as well as breaks in the skin. Surface components on the invading organism determine its ability to adhere and establish a primary site of infection. For example, the influenza A virus targets epithelial cells of the respiratory tract. After infection, the virus interacts with the host to cause typical symptoms of the flu. In another example, the protozoon *Entamoeba histolytica* adheres to cells lining the colon to produce the diarrhea associated with amebic dysentery. The cellular specificity of adherence of microorganisms often limits the range of susceptible hosts. For example, although measles and distemper viruses are closely related, dogs do not get measles and humans do not get distemper.

From the initial site of infection, microorganisms may directly invade further into tissues or travel through the blood or lymphatic system to other organs. For example, *Neisseria meningitidis* colonizes the mucous membranes of the upper respiratory tract of humans. However, the bacteria must spread through the bloodstream into the protective membranes that cover the brain and spinal cord for a life-threatening disease known as meningitis to occur.

Microorganisms produce toxins that can cause tissue destruction at the site of infection, can damage cells throughout the host, or can interfere with the normal metabolism of the host. The damage that microorganisms cause is directly related to the toxins they produce. Toxins are varied in their mechanism of action and host targets. For example,

Staphylococcus aureus can express several toxins that destroy tissue near its site of entry into the body, leading to the formation of boils, carbuncles, and abscesses. The tissue-damaging toxins collegenase, coagulase, hemolysins, proteases, and hyaluronidases serve either to protect the bacteriaL host defenses or to promote the spread of bacteria within the host. For example, coagulase, in conjunction with serum factors from the host, contributes to the formation of a fibrin clot that surrounds the bacteria, insulating them from destruction by the host. Alternatively, the enzyme hyaluronidase can weaken the host's connective tissue, facilitating the spread of the bacteria. Another bacterium, *Clostridium tetani*, produces a powerful toxin with a systemic effect. This toxin blocks the normal relaxation impulses of the nerves and thus causes the severe muscle spasms characteristic of tetanus. Cholera can interfere with the metabolism of the host. This ancient disease is contracted by ingesting water contaminated by the bacterium *Vibrio cholerae*. The bacterium releases a protein (an enterotoxin) that causes severe diarrhea by stimulating the secretion of enormous quantities of watery fluid from the small intestine. Patients lose so much body fluid that they go into shock and die. *See* CHOLERA; STAPHYLOCOCCUS; TETANUS.

Response to infection. The host's reaction to an infecting organism is the inflammatory response, the body's most important internal defense mechanism. Although the inflammatory response is also seen as secondary to physical injury and nonspecific immune reactions, it is a reliable indicator of the presence of pathogenic microorganisms. Inflammation is characterized by the cardinal signs of heat, redness, swelling, and pain. The warmth and redness are due to the dilation of minute blood vessels near the site of infection. The increased permeability of those vessels causes them to release fluid into the area causing swelling, and the irritation of nearby nerve endings causes pain. The dilated small vessels allow more blood to flow to the site of infection, while the enhanced permeability of the blood vessels permits increased numbers of disease-fighting substances to the area. Immune cells known as lymphocytes and granulocytes are carried by the blood to the site of infection. These cells either engulf and kill, or secrete substances which inhibit and neutralize, microorganisms. Other white blood cells, primarily monocytes, recognize foreign organisms and transmit chemical signals to other cells of the host's immune system, triggering the production of specific antibodies or specialized killer cells, both of which are lethal to the infecting microorganism. Any influence that reduces the immune system's ability to respond to foreign invasion, such as radiation therapy, chemotherapy, or destruction of immune cells by an immunodeficiency virus such as HIV, increases the likelihood that a organism will cause disease within the host. *See* INFLAMMATION.

Infectious disease control. Chemical compounds that are more toxic to microorganisms than to the host are commonly employed in the prevention and treatment of infectious disease; however, the emergence of drug-resistant organisms has led to increases in the morbidity and mortality associated with some infections. Other methods for controlling the spread of infectious diseases are accomplished by breaking a link in the chain of transmission between the host, microorganism, and mode of spread by altering the defensive capability of the host. Overall, the three most important advances to extend human life are clean water, vaccination, and antibiotics (in that order of importance).

Vector control. Water-borne infections are controlled by filtration and chlorination of municipal water supplies. Checking food handlers for disease, refrigeration, proper cooking, and eliminating rodent and insect infestation have markedly reduced the level of food poisonings. The transmission of vector-borne diseases can be controlled by eradication of the vector. Control of malaria in the United States succeeded after mosquito-breeding areas were drained and sprayed with insecticides. Blood-borne infections are reduced by screening donated blood for antibodies specific for HIV and other viruses and by rejecting donations from high-risk donors. For diseases such as tuberculosis, the airborne spread of the causative agent, *Mycobacterium tuberculosis*, can be reduced by quarantining infected individuals. The spread of sexually transmitted diseases, including AIDS, syphilis, and herpes simplex, can be prevented by inhibiting direct contact between the pathogenic microorganism and uninfected hosts. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS).

Vaccination. The spread of an illness can be halted by activating the immune system of the host. The most beneficial implementation of that concept has been vaccination, which renders an individual immune to an illness. Children in the United States are routinely immunized against diphtheria, whooping cough, tetanus, measles, mumps, rubella, chickenpox, and poliomyelitis. Because of a global vaccination programs, smallpox has been eradicated, and soon polio may be eliminated throughout the world.

Antimicrobial drugs. Drug therapy encompasses many classes of compounds that are active against common pathogens. These drugs are the "magic bullets" envisioned by Paul Erhlich to kill bacterial, viral, fungal, or protozoan pathogens without harming their human host. The pencillins, cephalosporins, and vancomycin are selective inhibitors of synthesis of the rigid cell wall that surrounds bacteria. Other antibacterial drugs, such as the tetracyclines and the aminoglycosides, block protein synthesis by the bacteria but not by the cells of the host. Still other compounds can alter the cell wall of bacterial cells, blocking the transport of nutrients into the cell. Viruses, because of their unique structure and method of replication, can be susceptible to drugs such as amantadine hydrochloride, which blocks the penetration of certain myxoviruses such as influenza A, or drugs that inhibit enzymes such as deoxyribonucleic acid (DNA) polymerases, proteases, or kinases that viruses use to replicate. Compounds that inhibit or suppress the growth of pathogenic fungi and protozoa have been

used as effective treatment for infectious diseases caused by these agents, although side effects can limit their utility. *See* DIPHTHERIA; EPIDEMIOLOGY; MUMPS; POLIOMYELITIS; RUBELLA; SMALLPOX; VACCINATION; WHOOPING COUGH. Peter J. McNamara

Bibliography. J. V. Bennet and P. S. Brachman, *Hospital Infections*, 2d ed., 1986; P. D. Hoeprich, *Infectious Diseases: A Modern Treatise of Infectious Processes*, 3d ed., 1983; P. R. Murray et al. (eds.), *Manual of Clinical Microbiology*, 6th ed., 1995; G. P. Youmans, P. Y. Paterson, and H. M. Sommers, *The Biological and Clinical Basis of Infectious Diseases*, 3d ed., 1986.

Infectious mononucleosis

A disease of children and young adults, characterized by fever and enlarged lymph nodes and spleen. Only in the mid-1970s did accumulated evidence permit clear linkage of this disease to EB (Epstein-Barr) herpesvirus as the causative agent.

Onset of the disease is slow and nonspecific with variable fever and malaise; later, cervical lymph nodes enlarge, and in about 50% of cases the spleen also becomes enlarged. The disease lasts 4–20 days or longer. Epidemics are common in institutions where young people live. EB virus infections occurring in early childhood are usually asymptomatic. In later childhood and adolescence, the disease more often accompanies infection—although even at these ages inapparent infections are common.

Seroepidemiologic studies indicate that infection with EB virus is common in different parts of the world. In some areas, including the urban United States, about 50% of children 1 year old, 80–90% of children over age 4, and 90% of adults have antibody to EB virus. The mechanism of virus transmission is unknown, but it could take place through oropharyngeal excretion, which has been found to be very common, especially in those with overt disease, and can occur in the absence of symptoms and be prolonged.

For diagnosis, the heterophile antibody test is useful. This test is based on a nonspecific serologic reaction present at high levels in patients with infectious mononucleosis. Total white blood cell count and differential blood count are also useful in diagnosis. *See* CLINICAL PATHOLOGY; HETEROPHILE ANTIGEN.

The nature of the immune response, which terminates the disease and the shedding of virus, is not fully understood; the virus appears to persist in the host at a location inaccessible to neutralizing antibody. Unlike short-lived heterophile antibodies, neutralizing antibodies against EB virus persist for years.

The presence of EB virus, an antigenically distinct member of the herpes group, was initially detected by means of electron microscopy in a small proportion of cells in continuous lymphoblastoid cell lines derived from Burkitt's lymphoma, a tumor indigenous to children in central Africa. EB virus has been detected in peripheral blood leukocytes of persons with infectious mononucleosis, leukocytes from normal individuals, and also in lymphoblastoid cell lines derived from nasopharyngeal carcinomas. A causative role of the virus in the cancers has not been unequivocally shown; it is possible that the virus is merely a "passenger" that grows well in the type of cell in the neoplasm. *See* EPSTEIN-BARR VIRUS; HERPES. Joseph L. Melnick

Bibliography. A. S. Evans (ed.), *Viral Infections of Humans: Epidemiology and Control*, 4th ed., 1997; G. Miller, Epstein-Barr herpesvirus and infectious mononucleosis, *Prog. Med. Virol.*, 20:84–112, 1975.

Infectious myxomatosis

A viral disease of European rabbits (*Oryctolagus*) and domestic rabbits, spread mainly by biting insects (mosquitoes and rabbit fleas) and characterized by edematous swellings of the skin, particularly on the head and anogenital area. The disease is caused by infection with myxoma virus, a pox virus, which occurs naturally in certain species of the genus *Sylvilagus* in North, Central, and South America. In these rabbits, infection results generally in localized, nonmalignant tumors that disappear in weeks or months.

Myxomatosis was introduced into Australia in 1950 as a method of controlling rabbits, which were causing great economic and environmental damage. It was also introduced into France in 1952 and spread throughout western Europe, including the British Isles. Although the virus strains introduced into Australia and France killed virtually every rabbit infected and rabbit populations were dramatically reduced, they were not eradicated, as less virulent strains appeared. Myxomatosis now has much less effect on rabbit populations: one estimate is that only 12–19% of rabbits die from the disease.

In typical wild rabbit populations, most of the adults are immune, following recovery from earlier infection. In domestic and commercial rabbits, there is no buildup of immunity or resistance, because of the common policy of isolating and slaughtering infected animals. There is no effective treatment once clinical signs have appeared. Preventive measures include restriction of contact with insect vectors and vaccination with active or inactivated myxoma virus or with Shope fibroma virus. *See* ANIMAL VIRUS. John Ross

Bibliography. F. Fenner and F. N. Ratcliffe, *Myxomatosis*, 1965; H. V. Thompson and C. M. King, *The European Rabbit*, 1994.

Infectious papillomatosis

A nonfatal viral disease that occurs naturally in cottontail rabbits in states bordering the Mississippi River as well as in Oklahoma and Texas, and in brush rabbits in California. Cottontail rabbits from other parts of the United States and domestic rabbits (*Oryctolagus*) are also susceptible. The disease is caused by infection with Shope papilloma virus, one of a number of papilloma viruses that infect a range of

animals, including humans; it is spread naturally by contamination of broken skin or by the rabbit tick.

The disease is characterized by cutaneous warts (papillomas) which can persist for months or years, sometimes becoming very prominent. Some papillomas eventually disappear, but in 25% of infected cottontail rabbits and 75% of infected domestic rabbits these warts become malignant carcinomas.

Rabbits can be immunized against papillomatosis by injection of active or inactivated virus. The virus can readily be recovered from papillomas on cottontail rabbits but rarely from those on domestic rabbits, indicating that it exists in mature fully infective virus and immature noninfective virus, which will proliferate in the host and stimulate an immune response. The disease is unlikely to spread between domestic rabbits because infection does not produce infective virus. *See* ANIMAL VIRUS. John Ross

Infertility

Inability to conceive or induce conception. Of all cases of infertility, 35% may be attributed to the male and 55% to the female; the remaining 10% is undetermined.

The female. The principal cause of increasing rates of infertility is the postponement of pregnancy: adverse effects of increasing age on reproductive capacity include decreased conception rates and increased pregnancy losses. The increasing incidence of pelvic inflammatory disease is also thought to be a major cause: infertility increases from 11% after one episode of infection to 54% after three episodes. Pelvic infections lead to scar formation around the ovaries and the fallopian tubes, thereby impeding the transport of oocytes for fertilization.

Ovulatory dysfunction accounts for approximately 20% of cases of infertility. Ovulation, the release of the oocyte from the ovary, is dependent on the timely secretion of the two gonadotropic hormones from the pituitary gland, follicle-stimulating hormone (FSH) and luteinizing hormone (LH). This process can be monitored by determining basal body temperature or the serum concentration of the hormone progesterone, which is secreted by the ovary after ovulation. Gonadotropic hormone release can be impeded by certain medications and by many disease states, including stress, anorexia nervosa, weight loss, and thyroid disease. Anovulation can be treated by replacing the deficient gonadotropins with human menopausal gonadotropin, a mixture of follicle-stimulating and luteinizing hormones. In addition, the drug clomiphene citrate increases the endogenous secretion of those hormones from the pituitary gland. *See* ENDOCRINE MECHANISMS; PITUITARY GLAND.

Fallopian tube and uterine abnormalities account for 25% of the cases of infertile couples. These abnormalities include defective development as well as scar formation after surgery or infection. The initial evaluation of this portion of the reproductive tract includes a hysterosalpingogram, that is, an x-ray of the uterus and fallopian tubes. After injection of a radiopaque dye through the cervix, the intrauterine contour and patency of the fallopian tubes are revealed. Not uncommonly the region of the fallopian tube closest to the uterus may be obstructed by scar tissue, but the obstruction can be microscopically removed or the fallopian tube can be surgically cut and reconnected. Occasionally, this area is closed by mucous plugs or tubal spasm. This condition can be successfully treated by directing a fine catheter into the fallopian tube to relieve the spasm or dislodge the plugs. This approach eliminates the need for microscopic surgery and yields comparable pregnancy rates. Uterine defects can be resected, but sometimes several attempts are necessary.

In endometriosis, tissue that normally lines the uterus is found on the pelvic lining. Severe forms of endometriosis involve the ovaries and fallopian tubes, which doubles the likelihood of infertility in those women. Identification of endometriosis is made at surgery, and the condition is treated either by surgical excision of the aberrant tissue or by subsequent medical treatment.

The cervix is more than a conduit for the passage of sperm to the uterus. Sperm may reside in cervical crypts for as long as 48 h, possibly with some enhancement of function. Abnormalities in cervical mucus production may result from trauma, surgery, and diethylstilbestrol (DES) exposure before birth. The cervical environment may be evaluated by examining the mucus 4–6 h after intercourse for the presence of viable sperm. The treatment of cervical abnormalities remains controversial, but may include estrogen; or another approach is to bypass this obstruction by using intrauterine transfer of semen (artificial insemination).

The male. Male infertility resulting from abnormal semen may be due to developmental defects, genitourinary infections, or varicocele. An evaluation of sperm count, motility, and morphology is helpful, but in most cases the cause of the abnormality remains undetermined. No treatment other than donor insemination is available for these couples.

Unexplained infertility. In approximately 10% of infertile couples, a thorough evaluation reveals no cause, and unexplained infertility is diagnosed. In most of these cases, diagnostic laparoscopy is recommended, involving direct visualization of the woman's pelvis through a laparoscope inserted through the abdominal wall. In nearly 50% of women with unexplained infertility, pelvic abnormalities are found and are most often caused by endometriosis and scar formation.

Assisted reproduction. In vitro fertilization takes place through the controlled hyperstimulation of ovulation followed by surgical extraction of mature oocyte from the ovaries. The oocytes are allowed to incubate in a petri dish with the semen specimen. After the oocytes are fertilized, the embryos are allowed to mature and are then transferred into the uterus. In vitro fertilization may be appropriate when unexplained infertility, nonfunctioning fallopian tubes, endometriosis, or decreased sperm counts

prevent conception; several thousand children have been conceived by using such a procedure.

Another method of assisted reproduction is gamete intrafallopian tube transfer (GIFT). Ovulation is stimulated, as with in vitro fertilization, but the extracted oocytes are placed directly into the fallopian tube with the semen specimen. Incubation occurs within the fallopian tube rather than outside the body. The procedure is undertaken for the same reasons as those for in vitro fertilization; however, in this case the fallopian tubes must be normal. In micromanipulation, microscopic probes hold oocytes and embryos, so that procedures to improve fertilization or genetic diagnostic tests on embryos can be performed. With severely low sperm numbers, even in vitro fertilization is generally not a viable option. Intracytoplasmic sperm injection can result in a viable pregnancy when a single sperm is injected into the oocyte by micromanipulation. This approach offers an option to childless couples, in addition to donor insemination. Also, micromanipulation allows sophisticated chromosomal analysis of embryos. *See* FERTILIZATION (ANIMAL); PREGNANCY; REPRODUCTIVE SYSTEM; REPRODUCTIVE SYSTEM DISORDERS. Richard E. Leach; Roger D. Kempers

Bibliography. S. J. Behrman et al. (eds.), *Progress in Infertility*, 4th ed., 1993; J. Cohen et al., *Micromanipulation of Human Gametes and Embryos*, 1992; D. Mishell, *Infertility, Contraception and Reproductive Endocrinology*, 3d ed., 1991; G. Sher, V. A. Marriage, and J. Stoess, *In Vitro Fertilization*, 1989.

Infinity

The terms infinity and infinite have a variety of related meanings in mathematics. The adjective finite means "having an end," so infinity may be used to refer to something having no end. In order to give a precise definition, the mathematical domain of discourse must be specified.

Infinity in sets. A simple and basic example of an infinite collection is the class of natural numbers, or positive integers. A fundamental property of positive integers is that after each integer there follows a next one, so that there is no last integer. Now it is necessary in mathematics to treat the collection of all positive integers as an entity, and this entity is the simplest infinity, or infinite collection.

Suppose that two collections A and B of objects can be set into one-to-one correspondence; that is, each object in A is paired with one and only one object in B, and each object in B belongs to just one of these pairs. Then A and B are said to contain "the same number" of objects, or to have "the same cardinal number." This is equivalent to the ordinary meaning of "the same number," obtained by counting, when the collections A and B are finite, and is taken as the definition of this phrase in the general case.

Now let A consist of all the positive integers, and let B consist of the even integers. Then A and B have the same number of elements, since each element of A can be paired with its double in B; thus elements can be removed from an infinite collection without reducing the number of its elements. This property distinguishes infinite collections from finite ones, and indicates that the statement "the number of elements of a collection A is greater than the number of elements of a collection B" should be taken to mean "B can be put into one-to-one correspondence with part of A, but not with the whole of A."

Consider next the number of subsets of a given collection A. For example, let $A = \{a, b, c\}$; that is, A consists of the three letters a, b, and c. Then the sets $\{a\}$, $\{b\}$, $\{c\}$, $\{a, b\}$, $\{a, c\}$, $\{b, c\}$, are subsets of A. To these, for convenience, can be adjoined the null set, containing no elements, and the set A itself. This makes a total of $8 = 2^3$ subsets. It is readily proved that a set A having a finite number n of elements has 2^n subsets. Thus the number of subsets of a finite set is always greater than the number of elements. By a different method, the same property may be proved to hold also for infinite sets. Hence for every infinite number there is a greater number. In other words, there exist infinitely many distinct infinities. Sets having the same number of elements as the collection of positive integers are called countably infinite, and this is the smallest infinity.

Infinity in limits of functions. The term infinity appears in mathematics in a different sense in connection with limits of functions. For example, consider the function defined by $y = 1/x$. When x tends to 0, y approaches infinity, and the expression may be written as

$$\lim_{x \to 0^+} y = \infty$$

Precisely, this means that for an arbitrary number $a > 0$, there exists a number $b > 0$ such that when $0 < x < b$, then $y > a$, and when $-b < x < 0$, then $y < -a$. This example indicates that it is sometimes useful to distinguish $+\infty$ and $-\infty$. The points $+\infty$ and $-\infty$ are pictured at the two ends of the y axis, a line which has no ends in the proper sense of euclidean geometry.

Infinity in geometry. In geometry of two or more dimensions, it is sometimes said that two parallel lines meet at infinity. This leads to the conception of just one point at infinity on each set of parallel lines and of a line at infinity on each set of parallel planes. With such agreements, parts of euclidean geometry can be discussed in the terms of projective geometry. For example, one may speak of the asymptotes of a hyperbola as being tangent to the hyperbola at infinity. Note that the points at infinity which are adjoined to a euclidean line or plane are chosen in a manner dictated by convenience for the theory being discussed. Thus only one point at infinity is adjoined to the complex plane used for geometric representation in connection with the theory of functions of a complex variable.

Other concepts. Other types of infinities may be distinguished when properties other than the mere cardinal number of a set are being considered. For

example, a line and a plane contain the same number of points, but when continuity considerations are important the line is said to be made up of a single infinity of points, whereas the plane has a double infinity, or ∞^2, points. This is because there exists a one-to-one continuous correspondence between the line and a subset of the plane, but there does not exist such a correspondence between the whole plane and the line or any part of it. As another example, an infinite set may be ordered in different ways so as to have different ordinal numbers. *See* CALCULUS.

Lawrence M. Graves

Bibliography. R. G. Bartle and D. R. Sherbert, *Introduction to Real Analysis*, 3d ed., 1999; W. Rudin, *Real and Complex Analysis*, 3d ed., 1991.

Inflammation

The local response to injury, involving small blood vessels, the cells circulating within these vessels, and nearby connective tissue.

The early phases of the inflammatory response are stereotyped: A similar sequence of events occurs in a variety of tissue sites in response to a diversity of injuries. The response characteristically begins with hyperemia, edema, and adherence of the circulating white blood cells to endothelial cells. The white cells then migrate between the endothelial cells of the blood vessel into the tissue. The subsequent development of the inflammatory process is determined by factors such as type and location of injury, immune state of the host, and the use of therapeutic agents.

Cardinal signs. The discomfort of inflammation may be attendant on a sunburn, a mosquito bite, a cut, an abscess, or a vaccination, and so it is not surprising that historical records of inflammation date to the earliest known medical writings, in the Egyptian papyri. It remained, however, for Celsus (about 30 B.C.–A.D. 38), a Roman, to enumerate the four cardinal signs of inflammation: redness, swelling, heat, and pain. Redness is caused by hyperemia, that is, an increased number of red blood cells in the local capillary bed. Heat results from increased flow of blood in the small blood vessels, and the edema, or swelling, represents the accumulation of extracellular fluid in the connective tissue. Stimulation of nerve endings by agents released during the inflammatory process causes pain. *See* EDEMA.

While it is possible to observe the events of inflammation in human skin, a detailed study of the dynamic cellular events requires a more convenient site in an experimental animal. Tissues such as rat mesentery, frog tongue, hamster cheek pouch, or bat wing have the virtues of accessibility, vascularity, and transparency necessary for microscopic study of the inflammatory process.

Vascular changes. Following a mild injury, there is a fleeting constriction of the smallest arterial branches (arterioles) in the viable tissue close to the injury, lasting from 5 s to about 5 min, followed by dilation of the same arterioles. That leads to engorgement and dilation of the capillaries, a change that extends over an area larger than the constriction site. Constriction of the smallest veins (venules) may also contribute to capillary engorgement.

Along with those blood vessel changes, the steady-state flow of water back and forth across the vessel walls is disrupted; when the outflow of water from the blood vessel exceeds the return, extravascular extracellular fluid collects. An excess of fluid in the extracellular compartment of the connective tissue is termed edema. The vessel walls also become more permeable to the large protein molecules that circulate in the plasma; those molecules leak into the tissue and induce the outward movement of water. Blood flow is slowed in the immediate vicinity of the injury by constriction of the venules and concentration of the blood cells. The latter phenomenon is a result of loss of water from the circulation. Cessation of blood flow (stasis) may occur and be associated with formation of a thrombus, a blood clot in the small vessels, or may be only transitory with prompt restoration of flow. *See* CIRCULATION; THROMBOSIS.

The leukocytes, at first predominantly neutrophilic granulocytes, adhere to endothelial cells that line the capillaries and venules, and then migrate through the vessel wall into the edematous connective tissue, drawn there by chemotactic agents that are released at the site of the injury.

If the tissue injury breaks small blood vessels, another mechanism comes into play. Exposure of circulating blood to collagen leads to clumping of platelets, and the resulting aggregate forms a temporary hemostatic plug that prevents free bleeding from the broken vessel. The clotting mechanism is set in action somewhat more slowly, and the fibrin formed thereby bolsters the plug. *See* COLLAGEN.

Cellular changes. In contrast to the vascular response, the cellular response in inflammation is varied and serves to characterize the different types of inflammation. Participating cells come from two sources, the circulating blood and the local connective tissue. The circulating leukocytes are divided into six distinguishable cell types: granulocytes (eosinophil, neutrophil, and basophil), small and large lymphocytes, and monocytes. Fibroblasts and mast cells are solely connective-tissue cells. Macrophages and giant cells arise in the inflammatory locus from monocytes. *See* BLOOD.

Neutrophils are normally the most abundant leukocytes in the circulation and are the first cells to accumulate in an inflammatory site. They have the capacity to ingest and kill bacteria; the cytoplasmic granules, which identify the neutrophil, contain the enzymes responsible for killing and digesting the bacteria. When the number of circulating neutrophils is greatly decreased, as in patients after treatment with certain drugs or exposure to nuclear radiation, frequent and severe infections may occur. In spite of antibiotic therapy, such patients often succumb to infection. *See* PHAGOCYTOSIS.

Eosinophils are much rarer than neutrophils but are also phagocytic and may be considerably

increased in patients with allergies. The basophil, least abundant of the granulocytes, is similar to the tissue mast cell; both store and release histamine and are thus responsible for hives and the symptoms of hay fever. There is evidence that the beneficial effects of both eosinophils and basophils are affected in dealing with multicellular parasites rather than bacteria or viruses. Lymphocytes are second to neutrophils in abundance and are very important in immune responses, including the rejection of grafts of foreign tissues, and the attack of donor cells. Monocytes, like lymphocytes, lack specific granules. This cell type when stimulated has the potential to change into a macrophage, which, like a neutrophil, can ingest and kill bacteria. Fibroblasts are responsible for synthesis of collagen and other components of the extracellular connective tissue during the healing process. *See* HISTAMINE.

Cause-effect relationship. An inflammatory response may be induced in a variety of ways. Causes of inflammation include trauma, heat, ultraviolet light, x-rays, bacteria, viruses, many chemicals such as turpentine and croton oil, and certain inherently innocuous but foreign substances (antigens) which evoke immune responses. Although many of the components of inflammation are common to a variety of inflammatory responses, particularly in the early stages, there is considerable diversity in fully developed responses; the response to an abscess is very different from that of a burn, for instance. The character of the injury, its severity, and the site of injury modify the progress of the inflammatory response, as does therapeutic intervention. *See* ANTIGEN; IMMUNITY.

A local inflammatory response is usually accompanied by systemic changes: fever, malaise, an increase in circulating leukocytes (leukocytosis), and increases in specific circulating proteins called acute-phase reactants. Such signals and symptoms are often helpful to the physician, first as clues to the presence of inflammation and later as an indication of its course.

Mediators of inflammation. The process of inflammation, both vascular and cellular, is orchestrated by an array of molecules produced locally. They variously induce the adhesion of circulating leukocytes to the endothelium, increase small blood vessel permeability to proteins, increase blood flow, direct the migration of inflammatory cells, enhance phagocytic activity, incite proliferation and differentiation of inflammatory cell types locally, and elicit the systemic correlates of inflammation. These mediators include histamine, leukotrienes, prostaglandins, complement components, kinins, antibodies, and interleukins. Many anti-inflammatory drugs function by preventing the formation of those mediators or by blocking their actions on the target cells whose behavior is modified by the mediators.

Types. Inflammation is frequently described in terms of its time course. Acute inflammation develops rapidly, in a matter of hours to days, and is of relatively short duration. In acute inflammation the neutrophil is the predominant cell type, and hyperemia and edema are prominent. Chronic inflammation develops over a period of weeks to months and is characterized by an infiltrate of lymphocytes, monocytes, and plasma cells—the chief antibody-producing cells. Local differentiation and proliferation of macrophages is characteristic of chronic inflammation.

Granulomatous inflammation is a specific type of chronic inflammation in which a discrete nodular lesion (granuloma) is formed of macrophages, lymphocytes, plasma cells, and giant cells arranged around a central mass of noncellular material. Granulomas, typical of tuberculosis and fungus infection, also occur in rheumatoid arthritis. *See* GRANULOMA INGUINALE; TUBERCULOSIS.

Abscesses and cellulitis are specific forms of acute inflammation; the former term denotes a localized collection of pus composed of necrotic debris derived from dead tissue cells and neutrophils. Cellulitis is characterized by diffuse hyperemia and edema with an extensive neutrophil infiltrate, and often little tissue destruction. When inflammation causes erosion of an epithelial surface, the lesion is termed an ulcer. *See* ULCER.

Function. Inflammation is basically a protective mechanism. The leakage of water and protein into the injured area brings humoral factors, including antibodies, into the locale and may serve to dilute soluble toxic substances and wash them away. The adherence and migration of leukocytes brings them to the local site to deal with infectious agents. There are also instances in which no causative toxic substance or infectious agent can be found to account for the inflammation. This is the case in rheumatoid arthritis and rheumatic fever. Such diseases may be examples in which an uncontrolled or misdirected inflammatory response with an autoimmune component is turned against the host. *See* ARTHRITIS; AUTOIMMUNITY; INFECTION; RHEUMATIC FEVER. David Lagunoff

Bibliography. I. L. Bonta and M. A. Bray (eds.), *The Pharmacology of Inflammation*, 1985; J. I. Gallin et al. (eds.), *Inflammation: Basic Principles and Clinical Correlates*, 3d ed., 1999; G. P. Lewis, *Mediator in Inflammation*, 1986; J. G. Lombardino (ed.), *Nonsteroidal Antiinflammatory Drugs*, 1985; P. Venge and A. Lindbom, *Inflammation: Basic Mechanisms, Tissue Injuring Principles, and Clinical Models*, 1985.

Inflammatory bowel disease

Inflammatory bowel disease is a general term for two closely related conditions, ulcerative colitis and regional enteritis or Crohn's disease. The diseases can affect the colon, distal, small intestine, and sometimes other portions of the gastrointestinal tract as well as several sites outside the gastrointestinal tract. In 15–25% of cases limited to the colon, ulcerative colitis and Crohn's disease cannot be distinguished by clinical manifestations, x-ray examination,

or even pathology. For this reason the broad term inflammatory bowel diseases is useful. The cause of these diseases is unknown.

Ulcerative colitis. Ulcerative colitis, an inflammatory condition limited to the colon, primarily affects the mucosa or lining of the colon. Marked inflammation gives rise to small ulcerations and microscopic abscesses that produce bleeding. The condition tends to be chronic, alternating between periods of complete remission and episodes of active and even life-threatening disease. Colitis may involve a varying extent of the colon; when limited to the rectum, it is known as ulcerative proctitis; it may be confined to the descending colon; or it may affect the entire colon, being known in this case as universal colitis.

The most common symptoms include rectal bleeding and diarrhea; blood is almost always present during a flare-up. Occasionally, there may be signs of toxicity, with other constitutional symptoms. Manifestations of ulcerative colitis, as well as Crohn's disease, outside the gastrointestinal tract include eye and joint inflammation and skin disorders. Extraintestinal disease occurs in 15–20% of patients with ulcerative colitis.

The greatest concern in long-term management is the risk of colon cancer; ulcerative colitis must therefore be monitored as a precancerous condition. When limited to the rectum and sigmoid, the disease does not seem to increase the risk of colon cancer, but when it is present beyond those most distal segments, the risk begins after about 10 years of disease and progresses steadily. The risk of colon cancer can be evaluated by microscopic examination of the epithelial cells of the colon lining. Biopsies can be examined for cell changes known as dysplasia or precancer. If there is no dysplasia on multiple samples, the risk of cancer remains low; if dysplasia is found, the risk is high.

Ulcerative colitis is treated with corticosteroids and a salicylate–sulfa drug combination known as sulfasalazine. If medication is not effective, surgical removal of the colon eliminates the disease but necessitates the creation of an opening in the abdominal wall from which the contents of the intestine can pass. The procedure, known as an ileostomy, redirects the lower end of the small intestine through the surface of the body where a collection appliance is attached to the skin. Alternatives to this method include an internal pouch or a valve mechanism. Other surgical techniques modify the distal rectum or anus so that evacuation can take place through the normal anatomic site.

Crohn's disease. Crohn's disease, also known as regional enteritis, granulomatous colitis, and terminal ileitis, affects the colon and small intestine, and rarely the stomach or esophagus. Like ulcerative colitis, it is chronic and of unknown etiology. The two diseases have many similarities, including their treatment. Pathologically, however, the findings are usually distinct. In Crohn's disease, chronic inflammation is present and is usually accompanied by granulomas. The inflammation involves the full thickness of the intestinal wall, often with bowel narrowing and obstruction of the lumen. Abdominal cramps, alteration of bowel function, and diminished food intake are common.

In the young, retardation of growth and of sexual maturation are frequently observed. Crohn's disease involving different parts of the gastrointestinal tract produces symptoms related to those regions. For example, extensive disease of the distal small intestine can interfere with absorption of nutrients, particularly vitamin B_{12} and bile salts, which are absorbed somewhat specifically in that region. Deficiency of bile salts can lead to malabsorption of fats.

Corticosteroid medications are utilized to treat the condition. Sulfapyridine used in ulcerative colitis can be effective, particularly when Crohn's disease affects the colon. Metronidazole is effective for treating inflammation in and around the anal region, which is more common in Crohn's disease than in ulcerative colitis. Surgery may be required in cases of bowel obstruction, hemorrhage, or severe debility. Although Crohn's disease carries a slight increase in the risk of cancer, the risk is modest when compared to ulcerative colitis. *See* DIGESTIVE SYSTEM; GASTROINTESTINAL TRACT DISORDERS. Leonard A. Katz

Bibliography. A. Anagostides, H. J. Hodgson, and J. B. Kirsner (eds.), *Inflammatory Bowel Disease*, 1991; S. B. Hanauer and J. B. Kirsner (eds.), *Inflammatory Bowel Disease*: *A Guide for Patients and Their Families*, 1985; D. Rachmilewitz (ed.), *Inflammatory Bowel Diseases*, 1986.

Inflationary universe cosmology

A theory of the evolution of the early universe, motivated by considerations from elementary particle physics as well as certain paradoxes of standard big bang cosmology, which asserts that at some early time the observable universe underwent a period of exponential, or otherwise superluminal, expansion.

In order to resolve the various paradoxes of the big bang theory, inflationary theories predictthat during this inflationary epoch the scale of the universe increased by at least 28 orders of magnitude. The period of superluminal expansion (inflation) in most cases is caused by the appearance of a nonzero constant energy density in the universe associated with the postulated existence of a phase transition between different ground-state configurations of matter that occurs as the universe expands and cools. After the transition is completed, the constant energy density is converted into the energy density of a gas of relativistic particles. At this point, inflationary scenarios match the standard big bang cosmological model.

Origin of inflationary models. The suggestion of an inflationary period during the early universe, in connection with a specific model, was first made in 1980 by A. Guth. (Somewhat earlier a less concrete

but not entirely unrelated possibility was discussed by A. A. Starobinskii.) Guth proposed, based on a consideration of recently proposed grand unification theories in particle physics, that phase transitions—associated with the breaking of certain symmetries of the dynamics governing the interactions of matter—could occur at sufficiently high temperatures and have important consequences in the early universe.

Symmetry breaking and phase transitions. The concept of symmetry is of fundamental importance in physics. In particular, a symmetry can be associated with each of the known forces in nature, and each symmetry is, in turn, associated with the existence of certain conserved quantities, such as electric charge in the case of electromagnetism. Surprisingly, a physical system in its ground state may not possess a symmetry that could be fundamental to the basic physics governing its dynamics. The most familiar example is a ferromagnet made up of many individual elementary magnetic dipoles. The equations of electromagnetism are manifestly rotationally invariant: there is no intrinsic north or south. However, the ground state of such a spin system will involve all the spins aligned in one particular direction, yielding the familiar case of a permanent magnet. In this case, it is said that the rotational symmetry of the equations of electromagnetism has been spontaneously broken. The signal for this is that the average value of the total spin of the system points in a certain direction. If the system is heated, so that all the spins become excited and randomly oriented, the net spin of the system approaches zero, its magnetic field vanishes, and rotational symmetry is restored. The total spin of the system is referred to as an order parameter, because its value tells which state the system is in, and a change in the order parameter is an indication of the presence of a phase transition between different ground states of the system. *See* PHASE TRANSITIONS; SYMMETRY BREAKING; SYMMETRY LAWS (PHYSICS).

Phase transitions in particle physics. A situation similar to the case described above occurs in particle physics. At the temperatures and energies that occur in the universe now, there isavast difference in the nature of the known fundamental forces. For two of these forces—the electromagnetic force, and the so-called weak force, which governs the beta decay processes important to energy production in the Sun—it is now known that the perceived differences are the result of spontaneous symmetry breaking. Above a certain energy or temperature, the weak and electromagnetic interactions appear exactly the same. However, at a critical temperature, the ground state of matter, which in particle physics is called the vacuum state, breaks the symmetry relating the two interactions, and below this temperature they appear quite different. The signal for this symmetry breaking is again the appearance of a nonzero value in a certain order parameter of the system; in this case it is referred to as a vacuum expectation value. In the case of the ferromagnet, the order parameter was a spin, and the ground state after symmetry breaking had a net magnetization. In the case of the particle physics system, the order parameter describes the ground state expectation value of a certain elementary particle field, and the vacuum state after symmetry breaking has a net charge that would have been zero had the symmetry not been broken.

In particle physics, if a symmetry is broken at a certain scale of energy, particles that transmit the force of nature associated with that symmetry can have masses characteristic of this energy scale. This is indeed what happens in the case of the weak interactions, which are weak precisely because the particles that transmit the weak force are so massive. However, once the temperature of a system is so great that the energy it takes to produce such a massive particle is readily available in thermal energy, the distinction between this force and electromagnetism, which is transmitted by massless particles (photons), disappears and the two forces become unified into one.

Shortly after it was recognized that these two forces could be unified into one, it was proposed that all three observed forces in nature outside of gravity might be combined into one grand unified theory. There are reasons to believe that the scale of energy at which all the forces would begin to appear the same is truly astronomical—some 16 orders of magnitude greater than the mass of the proton. Such energies are not achieved in any terrestrial environment, even in the laboratory. The only time when such a scale of energy was common was in the earliest moments of the big bang fireball explosion. *See* ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS; GRAND UNIFICATION THEORIES.

"Old" inflationary model. Based on these facts, Guth reasoned that if grand unification symmetries are indeed broken at some large energy scale, then a phase transition could occur in the early universe as the temperature cooled below the critical temperature where symmetry breaking occurs. According to the standard big bang model of expansion, the time at which this would occur would be about 10^{-35} s after the initial expansion had begun (**Fig. 1**). This

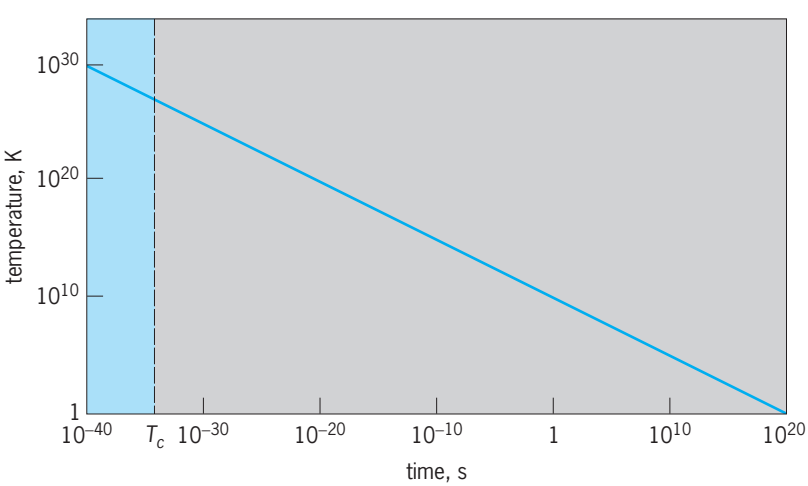

Fig. 1. Temperature versus time after beginning of expansion for the standard big bang model. Before T_c, inflationary scenarios differ from the standard model.

time is far earlier than the times previously considered in cosmological analyses using the big bang model. For instance, the time at which it is predicted that many of the light elements were first formed in the process of primordial nucleosynthesis is of the order of 1 s, some 35 orders of magnitude later in time.

First-order phase transitions. As Guth demonstrated, the effects of such a phase transition in the early universe could be profound. In order to calculate the dynamics of a phase transition, it is necessary to follow the behavior of the relevant order parameter for the transition. This is done by determining the energy (actually the thermodynamic free energy in the case of systems at finite temperatures) of a system as a function of the order parameter, and following the changes in this energy as the temperature changes. Typically (**Fig. 2**), at some high temperature T, the minimum of the relevant energy function is at zero value of the order parameter, the vacuum expectation value of a certain field. Thus the ground state of the system, which occurs when this energy is a minimum, will be the symmetric ground state. This case is analogous to the spin system, which at high temperatures is disordered, so that no preferred direction is picked out. As the temperature is decreased, however, at a certain critical temperature T_c a new minimum of the energy appears at a nonzero value of the order parameter. This is the symmetry-breaking ground state of the system. *See* FREE ENERGY.

How the system makes a transition between the original, symmetric ground state and the new ground state depends on the shape of the energy curve as a function of the order parameter. If there is a barrier between the two minima (Fig. 2), then classically the system cannot make a transition between the two states. However, it is a well-known property of quantum mechanics that the system can, with a certain very small probability, tunnel through the barrier and arrive in the new phase. Such a transition is called a first-order phase transition. Because the probability of such a tunneling process is small, the system can remain for a long time in the symmetric phase before the transition occurs. This phenomenon is called supercooling. When the transition finally begins, "bubbles" of new phase locally appear throughout the original phase. As more and more bubbles form, and the original bubbles grow (in the case of such a phase transition in particle physics the bubbles grow at the speed of light), eventually they combine and coalesce until all of the system is in the new phase. Another example of this type of phenomenon is the case of a supercooled liquid such as water that has been continually stirred as it has been cooled. At a certain point, ice crystals will spontaneously form throughout the volume of water, and grow until all of the water has turned to ice. *See* NONRELATIVISTIC QUANTUM THEORY.

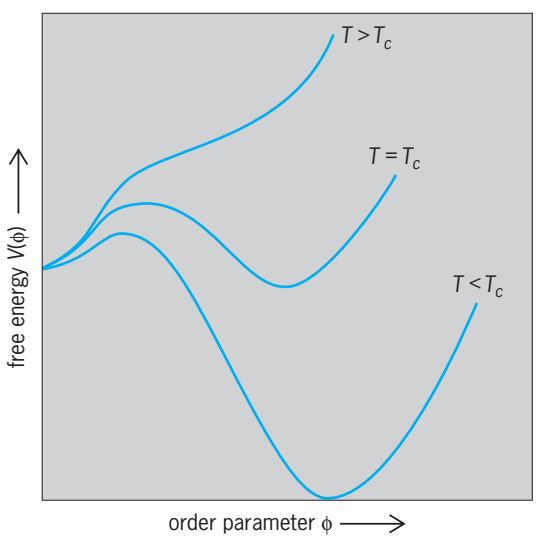

Fig. 2. Free energy as a function of the order parameter for the case of a first-order transition that is typical of symmetry breaking in grand unification theories.

Inflation. Such a phase transition, when treated in the context of an expanding universe, can result in remarkable behavior. The dynamics of an expanding universe are governed by Einstein's equations from the theory of general relativity. When these equations are solved by using the properties of ordinary matter, it is seen that the rate of expansion of the universe, characterized by the Hubble parameter H—which gives the velocity of separation of objects—slows down over time in a way that is dependent on the temperature of the universe, T. (Specifically, in a radiation-dominated universe, H is dependent on T^2.) This slowdown occurs because the energy density of matter, which is driving the expansion, is decreased as the matter becomes more dilute because of the effects of expansion.

However, once the temperature is below the critical temperature for the transition (Fig. 2), the metastable symmetric phase has a higher energy than the new lower-energy symmetry-breaking phase. Until the transition occurs, this means that the symmetric phase has associated with it a large constant energy density, independent of temperature. When this constant energy density is placed on the right-hand side of Einstein's equations, where the energy density of matter appears, it is found that the resultant Hubble parameter describing expansion is a constant. Mathematically, this implies that the scale size of the universe increases exponentially during this supercooling phase. This rapid expansion is what is referred to as inflation. Once the phase transition is completed, the constant energy density of the original phase is converted into the energy density of normal matter in the new phase. This energy density decreases with expansion, so that the universe reverts to the slower rate of expansion characteristic of the standard big bang model. *See* HUBBLE CONSTANT; RELATIVITY.

Successes of inflation. Guth pointed out that a period of exponential expansion at some very early time could solve a number of outstanding paradoxes associated with standard big bang cosmology.

Flatness problem. In 1979, it was pointed out that present observations seem to imply that either the present era is a unique time in the big bang

expansion or the initial conditions of expansion had to be fine-tuned to an incredible degree. The observation involves the question of whether the universe is now open, closed, or flat. A closed universe is one where there is sufficient mass in the universe to eventually halt and reverse the observed expansion because of net gravitational attraction; an open universe will go on expanding at a finite rate forever; and a flat universe forms the boundary between these two cases, where the expansion rate will continually slow down and approach zero asymptotically. Measurements of the observed expansion rate, combined with measurements of the observed mass density of the universe, yield a value for the density parameter Ω that rapidly approaches 0 for an open universe, infinity for a closed universe, and is exactly equal to 1 for a flat universe. All measurements of Ω yield values between about 0.1 and 2. What is so strange about this measurement is that theory suggests that once the value of Ω deviates even slightly from 1, it very quickly approaches its asymptotic value far away from 1 for open or closed universes. Thus, it is difficult to understand why, after 10^{10} years of expansion, the value of Ω is now so close to 1. At a quantitative level, the problem is even more mysterious. In order to have its present value in the range given above, at the time of nucleosynthesis the value of Ω would have had to have been equal to 1 within 1 part in 10^{15}.

Inflation naturally explains why Ω should exactly equal 1 in the observable universe today. Einstein's equations for an expanding universe can in the cases of interest be written in the form of the equation below, where t is the time after beginning of expansion, R is the cosmic scale factor, related to the size of the universe in closed universe models, ρ is the energy density of matter, and K is a constant that is equal to 0 for flat universe models. Normally, for an expanding universe, the energy density of matter decreases with scale at least as fast as $1/R^3$, so that the left-hand side deviates from 1 as time goes on. This is the quantitative origin of the flatness problem. However, in an inflationary phase, $\rho(t)$ remains constant, while R increases exponentially. Thus, during inflation $\Omega(t)$ is driven arbitrarily close to 1 within the inflated region. If there are some 28 orders of magnitude of exponential expansion, which is possible if supercooling lasts for some time, then $\Omega(t)$ need not have been finely tuned to be close to 1, even if inflation occurred as early as the Planck time (10^{-45} s).

$$1 + \frac{K}{R(t)^2 \rho(t)} = \Omega(t)$$

Horizon problem. An equally puzzling problem that inflationary cosmology circumvents has to do with the observed large-scale uniformity of the universe. On the largest observable scales, the universe appears to be largely isotropic and homogeneous. In particular, the 3-K microwave radiation background, which in different directions has propagated from distances separated by 10^{10} light-years, is known to be uniform in temperature to about 1 part in 100,000. *See* COSMIC BACKGROUND RADIATION.

This observed uniformity may not seem so puzzling, until an attempt is made to derive it in the context of any set of reasonable initial conditions for the big bang expansion. Physical effects can propagate at best at the speed of light. The distance a light ray could have traveled since the big bang explosion, which is also thus the farthest distance at which one object can affect another, is called the horizon. The horizon size increases linearly with time, since light travels with a constant velocity even in an expanding universe. Thus, the size of the horizon today, which is about 10^{10} light-years, is much larger than the horizon size at the time the radiation in the microwave background was emitted. In particular, in the standard big bang model the sources of the radiation observed coming from opposite directions in the sky were separated by more than 90 times the horizon distance at the time of emission. Since these regions could not possibly have been in physical contact, it is difficult to see why the temperature at the time of emission was so uniform in all directions.

Even if some very isotropic initial conditions are postulated for the big bang expansion to account for this uniformity, a quantitative problem similar to the flatness problem is encountered in attempting to account for the degree of uniformity now observed. Small fluctuations in energy density tend to grow as the universe evolves because of gravitational clumping. Indeed, observed clumping on the scale of galaxies and smaller attests both to this fact and to the existence of some initial irregularities in the distribution of matter. However, since irregularities grow but the universe is relatively smooth on large scales, any initial conditions for the big bang expansion imposed at early times, say 10^{-45} s, would have to involve an absurdly uniform distribution of matter on the largest scales.

Inflation solves the horizon problem very simply. If the observed universe expanded by 28 orders of magnitude in a short period at a very early time, then in inflationary cosmology it originated from a region 10^{28} times smaller than the comparable region in the standard big bang model extrapolated back beyond that time. This makes it quite possible that at early times the entire observed universe was contained in a single horizon volume (**Fig. 3**).

Problems with "old" inflation. The list of cosmological problems resolved by the original inflationary universe model is quite impressive. Moreover, its origin in current elementary-particle-physics ideas made it more than an artificial device designed specifically to avoid cosmological paradoxes. Unfortunately, the original inflationary scenario was fundamentally flawed. The central problem for this scenario was how to complete the phase transition in a uniform way. The phase transition began by the formation of bubbles of one phase nucleating amidst the initial metastable phase of matter. While these bubbles grow at the speed of light once they form, the space between the bubbles is expanding exponentially. Thus, it is extremely difficult for bubbles to

eventually occupy all of space, as is required for the completion of the transition. Moreover, even if bubbles are forming at a constant rate, each region of space quickly becomes dominated by the largest bubble, which formed earliest. Collisions with much smaller bubbles in this region will not adequately or uniformly dissipate the energy of this largest bubble, so that even if the transition does manage to percolate through space, the final state will be very nonuniform on all scales. Finally, topologically stable objects such as magnetic monopoles and domain walls can form at the intersection of bubbles, and end up with densities after inflation well above those allowed by observation. *See* MAGNETIC MONOPOLES.

While these problems motivated the development of a "new" inflationary scenario, described below, in 1989 it was pointed out that in certain previously examined special models related to so-called Brans-Dicke models of gravitation, the gravitational constant can itself vary during an inflationary phase transition. In this case, the inflationary expansion rate, while very fast, need not be exponentially fast, so that bubbles growing at the speed of light can eventually percolate throughout all of space, thereby completing the phase transition and ending the inflationary expansion. However, many details of these rather special scenarios, which may still leave large remnant inhomogeneities on the scale where bubbles coalesce, remain to be worked out. *See* GRAVITATION.

"New" inflationary cosmology. In 1981 a new approach was developed by A. D. Linde, and independently by A. Albrecht and Steinhardt, which has since become known as new inflation. They suggested that if the energy function (potential; Fig. 2) were slightly changed, then it might be possible to maintain the successful phenomenology of the old inflationary model while avoiding its problems. In particular, they considered a special form of the potential that is extremely flat at the origin (**Fig. 4**). Most important, such functions have essentially no barrier separating the metastable from the stable phase at low temperature. Now, when the universe cooled down below the critical temperature, the order parameter could continuously increase from zero instead of tunneling discretely to a large nonzero value. As long as the potential is sufficiently flat near the origin, however, it can take a long time before the order parameter approaches its value at the true minimum of the potential (a so-called slow-rollover transition), at which point the region will be in the new phase. During this time the region of interest can again be expanding exponentially because of the large constant energy density that is maintained while the order parameter remains near the origin. Thus, in some sense a single bubble can undergo inflation in this scenario. If the amount of inflation in this region is sufficient, the whole observable universe could have originated inside a single inflating bubble.

This key difference between old and new inflationary cosmology accounts for the ability of the latter to bypass the problems for the former. Because in new inflation the observable universe grew from a single inflating bubble, problems of inhomogeneity and topological defects resulting from bubble percolation and collisions are avoided, at least as far as observational cosmology is concerned. Moreover, the problem of completing the phase transition is naturally avoided. Once the order parameter in the inflating region approaches the steep part of the potential, the large constant energy density of the initial phase is reduced, and inflation slows. This energy density is converted into a form of kinetic energy of the changing order parameter. This motion of the order parameter is reflected in a changing charge of the new ground state of matter. Just as a time-varying electric charge produces electromagnetic radiation, so the variation of the order parameter produces a thermal background of real particles. By the time the order parameter has settled at the new minimum of the potential, all of the original constant energy density of the symmetric phase of matter has been converted into energy of real particles at a finite temperature. (Any particle density that existed before inflation has, of course, been diluted away by the vast expansion of the region during inflation.) The evolution from this point on is that of the standard

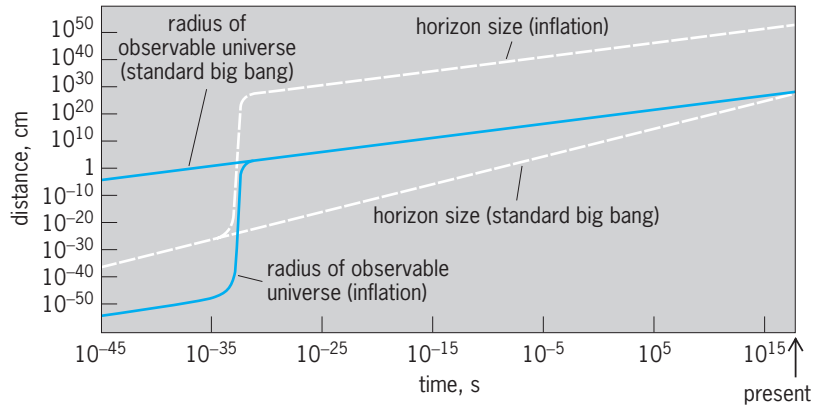

Fig. 3. Horizon size and radius of the observable universe as functions of time after beginning of expansion, in the standard big bang and inflationary models. (*After A. H. Guth and P. J. Steinhardt, The inflationary universe, Sci. Amer., 250(5):116–128, May 1984*)

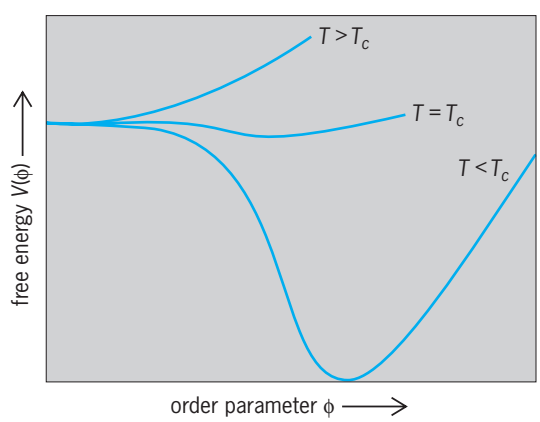

Fig. 4. Free energy as a function of the order parameter for the case of a so-called slow-rollover transition that appears in the new inflationary cosmology.

big bang model, with the initial conditions being a uniform density of matter and radiation at some finite temperature.

Problems with new inflation. New inflation, too, is not without its problems. The type of potential (Fig. 4) needed for a new inflationary scenario is not at all as generic as that postulated in the old inflationary model (Fig. 2). To have such a slow-rollover transition, the parameters of particle physics models must be finely tuned to some degree. Moreover, the temperature to which the universe is reheated after inflation must be large enough that, at the very least, big bang nucleosynthesis can proceed afterward. This again rules out some possible models. No clear candidate model for new inflation has emerged from particle physics.

Another potential problem for any inflationary scenario concerns initial conditions. As discussed above, if an inflationary phase precedes the standard big bang expansion, then it is possible to resolve problems of the standard big bang model related to the unphysical fine tunings that seem necessary at time zero in order for the big bang to evolve into its presently observed form. However, there is also the question of how generic such an inflationary phase is—namely, how special the initial preinflationary conditions must be so that space-time will undergo an inflationary transition in the first place. If these too are unphysical, then inflation may not have really solved any fundamental problems. While Guth suggested in his initial work that the preconditions for inflation were not severe, this view has been questioned by some. Attempts have been made to develop an existence proof that not all initial configurations of the universe can be made isotropic by inflation. This, of course, does not address the fundamental issue of whether such configurations are at all generic.

In 1983, Linde proposed a version of inflation, which he called chaotic inflation, that may in principle address this issue. In particular, he argued that inflation may have occurred very generally in the early universe, even if there was no related phase transition. As long as a scalar field existed with a potential whose scale of energy was sufficiently close to the Planck scale, then, if quantum fluctuations carried the field expectation value far enough away from the origin, the field would relax slowly enough that its energy density would remain nonzero and roughly constant for some time, causing an inflationary phase of expansion. Such a possibility led Linde to suggest that the early universe may have been arbitrarily inhomogeneous. In some regions, inflation may have successfully taken place, even without the spontaneous symmetry breaking associated with a grand unified theory, and in other regions it may not have. He then suggested that the regions in which inflation did take place are in fact most probable. In particular he argued that those values of parameters that allow large amounts of inflation to take place are most probable, if an initial random distribution of regions in the preinflation universe governed by different metric parameters is allowed. In addition, he pointed out that life would form only in those regions that became sufficiently isotropic so that, if many different regions of the universe now exist, it is not surprising that humans live in one that has undergone inflation. This argument, a reformulation of the so-called anthropic principle, is, however, difficult to quantify. In any case, the issue of initial conditions for inflation is the subject of much research, but may require an understanding of quantum gravity for its eventual resolution. *See* CHAOS; QUANTUM GRAVITATION.

Predictions. Whether or not there exists a specific model for new inflationary cosmology, and whether or not understanding inflation may require further refinements in the understanding of quantum gravity, the implications of an inflationary phase in the early universe are profound. As discussed above, many of the fundamental paradoxes of standard big bang cosmology can be resolved. Moreover, it has been demonstrated that new inflationary cosmology naturally allows a derivation from first principles of the spectrum of primordial energy density fluctuations responsible for galaxy formation. Before new inflation, there existed no method, even in principle, for deriving this spectrum for the standard big bang model; it always had to be input as an artificial addition. Remarkably, the spectrum that emerges from inflationary models is a so-called scale-invariant spectrum of perturbations. It is exactly this type of spectrum that had been postulated 10 years before the formulation of the new inflationary model to account for the observed properties of galaxies while maintaining agreement with the observed anisotropy of the microwave background.

In a remarkable discovery, the first observation of anisotropies in the microwave background was announced in April 1992, based on the analysis of 2 years of data on the microwave background structure using the differential microwave radiometer experiment aboard the *Cosmic Background Explorer* (*COBE*) satellite launched by the National Aeronautics and Space Administration (NASA) in 1989. A quadrupole anisotropy in the background was observed at a level of about 5×10^{-6}, just in the range that might be expected for primordial fluctuations from inflation that might also result in the observed distribution of galaxies. Moreover, the correlation of temperature deviations observed across the sky on scales greater than about 10° is remarkably consistent with the scale-invariant spectrum predicted by inflationary models. While neither of these observations conclusively proves the existence of an inflationary phase in the early universe, the fact that they are clearly consistent with such a possibility, and at present with no other scenario, gives great confidence that inflationary models may provide the correct description of the universe in the era preceding the present observed isotropic expansion.

Theoretical and observational factors. Since the early 1990s several factors, both theoretical and observational, have dominated thinking about possible inflationary models.

Hybrid inflation. As the theoretical notion of supersymmetry—a symmetry relating the two different kinds of elementary particles in nature, bosons (integer-spin particles) and fermions (half-integer-spin particles)—has begun to dominate the model building of elementary particle physics, the kind of flat potentials which allow long slow rollover periods of inflation have become more widely used. The problem in such models is not how to have sufficient inflation but rather how to end inflation. A new mechanism, which can be relatively natural in some supersymmetric models, has been proposed in this regard. Called hybrid inflation, this mechanism relies not on the evolution of a single-field order parameter but on several interacting fields. In this case, the evolution of these fields as the universe evolves can be more complex. In certain cases, the preferred, lowest free-energy state of matter can evolve along the flat direction, but as the value of the corresponding field evolves, its couplings to other fields can, at a certain critical value, cause the lowest free-energy state to evolve along a nonflat direction, abruptly ending inflation. The advantages of these models are that inflation can end without fine tuning, primordial density fluctuations that are of an acceptable magnitude can be generated without a great deal of fine tuning, and the models are reasonably well motivated on particle physics grounds. Unfortunately, model building in this area is still sufficiently murky that no obvious set of phenomenologically acceptable models yet exists. *See* SUPERSYMMETRY.

It has also been proposed that the generation of the observed matter-antimatter asymmetry in nature might not occur at temperatures corresponding to the grand unified scale, where the symmetry breaking which motivated Guth's original model is thought to occur, but at much lower energies, corresponding to the scale where the symmetry between the weak and electromagnetic interactions breaks. In this case, it may be possible that following inflation the universe thermalizes at a temperature far below that associated with the original inflationary models. This relaxes some of the constraints on inflationary models, and again favors the possible existence of long flat directions in the potential that might be associated with supersymmetric models. *See* ANTIMATTER; ELECTROWEAK INTERACTION; STANDARD MODEL; WEAK NUCLEAR INTERACTIONS.

Preheating. Original calculations suggested that, following inflation, the universe would quickly thermalize to a temperature comparable to the energy density stored in the original symmetric preinflationary phase of matter. Such a scenario was called reheating. Subsequent calculations have shown that a very different behavior can take place. As the order parameter relaxes about the new minimum of the potential following inflation, its vibrations can release energy in certain nonthermal modes. As a result, immediately following inflation a finite fraction of the released energy might exist in such nonthermal modes. Thus, estimates of particle number densities based on the assumption of thermal equilibrium may be dramatically wrong. Among the possible consequences of this situation, called preheating, is the generation of a large density of topological defects such as cosmic strings, and possibly magnetic monopoles in the postinflationary phase. Since one of the original purposes of inflation was ridding the universe of such defects, it is not yet clear what the implications of this new possibility are for the health of inflationary models. *See* COSMIC STRING.

Dark energy, branes, and a flat universe. Finally, since 1995 it has been increasingly clear, based on measurements of the total clustered mass in the universe in galaxies and clusters of galaxies, that there is not sufficient matter to result in a flat universe today. However, it is now recognized that matter accounts for only part of the total energy density of the universe, with the remainder being associated with some new form of energy, possibly energy stored in the ground state of empty space, called a cosmological constant. This energy may in fact be identical to the energy that is stored during the inflationary phase itself. If this is true, we live in an inflationary universe today. Measurements of the expansion rate of the universe as a function of cosmic time, by observing the recession velocities of certain types of exploding stars, called type 1a supernovae, in very distant galaxies, argue that this is precisely the case, with 30% of the energy of a flat universe in matter and 70% in dark energy. It will probably be necessary to rethink many of the current microphysical notions about the expanding universe to accommodate this newly discovered dark energy. *See* ACCELERATING UNIVERSE; COSMOLOGICAL CONSTANT; DARK ENERGY; SUPERNOVA.

A great deal of theoretical work has been devoted to exploring the possibility that our four-dimensional universe might be embedded in a universe of higher dimensions. In such models—in which our four-dimensional universe, which forms a structure called a brane, can be impacted upon by higher-dimensional physics—the possibility exists that inflation might be caused by new physics associated with other branes embedded in higher-dimensional spaces, instead of being associated with phase transitions due purely to physics on our brane. However, these ideas remain highly speculative. *See* SUPERSTRING THEORY.

Results and prospects. The next-generation advanced cosmic microwave background satellite, the *Wilkinson Microwave Anisotropy Probe* (*WMAP*), has been able to probe many of the fundamental parameters of cosmology, including the Hubble expansion rate, the overall matter density, the spectrum of primordial density fluctuations, and the possible existence of a cosmological constant, to an accuracy of better than 10%. *WMAP* has confirmed many of the generic predictions of inflation, including a flat universe and a roughly constant spectrum of primordial fluctuations. Even more significant perhaps, it achieved sufficient accuracy to determine that the spectrum is not precisely scale-invariant, but changes with wavelength very slightly in a way that is precisely

that expected by many inflationary models. (Inflation predicts only a roughly scale-invariant spectrum when detailed estimates are performed. The fluctuations spectrum actually increases slightly with increasing wavelength.) While this is not proof that inflation actually occurred, the consistency between theoretical expectations and observations provides further support for this idea. Moreover, detailed measurements of the spectrum of primordial fluctuations can actually rule out some inflationary models. *See* WILKINSON MICROWAVE ANISOTROPY PROBE.

While all of the current data supports the idea of inflation, because many different inflationary models exist, the question remains whether there is a prediction that can be probed that is unique to inflation. Interest currently is focused on the possibility of measuring a primordial spectrum of gravitational waves, which is also generated during an inflationary phase of the early universe. The spectrum of these waves would also be close to scale-invariant, and the amplitude of the waves would depend upon the energy scale at which inflation occurred. Such a spectrum might produce a measurable polarization in the cosmic microwave background at a level that might be probed by *Planck*, the next-generation cosmic microwave background satellite, or by future ground-based detectors. A positive signal could provide very strong evidence for inflation, but unfortunately the absence of a signal would not provide definitive evidence against it. *See* GRAVITATIONAL RADIATION.

For the moment, inflation remains a beautiful and well-motivated theoretical possibility, which is now more strongly supported by all existing data, and in fact is the only such idea that is consistent with all existing cosmological observations. Hopefully, theoretical and observational developments will help settle the question of whether or not it actually occurred. *See* BIG BANG THEORY; COSMOLOGY; UNIVERSE.

Lawrence M. Krauss

Bibliography. C. L. Bennett, M. S. Turner, and M. White, The cosmic Rosetta stone, *Phys. Today*, 50(11):32–38, November 1997; A. H. Guth, Inflationary universe: A possible solution to the horizon and flatness problems, *Phys. Rev.*, D23:347–356, 1981; A. H. Guth and P. J. Steinhardt, The inflationary universe, *Sci. Amer.*, 250(5):116–128, May 1984; L. Krauss, Cosmological antigravity: The case for a cosmological constant, *Sci. Amer.*, 282(1):52–59, January 1999; L. Krauss, *Quintessence: The Mystery of the Missing Mass in the Universe*, 2000; A. D. Linde, Eternally existing self-reproducing chaotic inflationary universe, *Phys. Lett.*, B175:395–400, 1986; A. D. Linde, The inflationary universe, *Rep. Prog. Phys.*, 47:925–986, 1984; D. N. Spergel et al., First-year *Wilkinson Microwave Anisotropy Probe* (*WMAP*) observations: Determination of cosmological parameters, *Astrophys. J. Suppl.*, 148:175–194, 2003; D. N. Spergel et al., *Wilkinson Microwave Anisotropy Probe* (*WMAP*) three year results: Implications for cosmology, *Astrophys. J.*, in press, 2006.

Inflorescence

A flower cluster segregated from any other flowers on the same plant, together with the stems and bracts (reduced leaves) associated with it. Certain plants produce inflorescences, whereas others produce only solitary flowers. The stalks of the individual flowers are the pedicels.

Inflorescences are easily identified in certain species, though not in others. For example, it is often difficult, if not arbitrary, to determine whether flowers on a plant are sufficiently segregated from each other to be interpreted as solitary flowers or components of one or more inflorescences. In other cases, such as species of *Solidago* (goldenrod), a plant produces flowers in small clusters conventionally designated as inflorescences, but these inflorescences are aggregated into a very conspicuous inflorescencelike cluster at the tip of the plant.

Branching. Patterns of branching in inflorescences, as well as other portions of the plant, are classified under two general categories, monopodial and sympodial. Distinctions between these categories are not absolute, and a given inflorescence may exhibit one or both patterns. In monopodial branching, each branch is subordinate to that which bears it, and stems may bear numerous lateral branches (including flowers; **illus.** *a* and *o*). In sympodial branching, each branch is more or less equivalent to the branch which bears it, generally terminates in a flower, and tends to be short, bearing one or a few lateral branches (illus. *h–n′*). In some cases of sympodial branching, a structure which superficially appears as a single branch (that is, one derived from a single apical meristem) is actually composed of several branches.

Classification. Inflorescences are classified according to gross appearance at maturity or pattern of branching. Depending on the species, an inflorescence may be regarded as belonging to one or more types; for example, the inflorescence of *Limnocharis flava* (velvet leaf) is properly classified as an umbel (based on gross appearance), yet also as a compound monochasium or bastryx (based on branching pattern).

Types of inflorescences include: raceme (illus. *a*), composed of a monopodial central axis, along the sides of which arise pedicel-bearing flowers; corymb (illus. *b*), a raceme whose pedicels are proportionally elongated so as to bring all the flowers more or less to the same level; umbel (illus. *c*), composed of several flowers which radiate from almost the same point; compound umbel (illus. *d*), composed of several branches which radiate from almost the same point, each branch terminating in an umbellike flower cluster; spike (illus. *e*), composed of flowers which lack or virtually lack pedicels, arranged singly or in contracted clusters along a central axis; catkin (ament; illus. *f*), a slender, usually scaly spike bearing unisexual, apetalous flowers, which often hangs downward and is deciduous as a whole; spadix (illus. *g*), a spike with a thickened fleshy axis, usually enveloped by one or more spathes (large modified leaves); dichasium (simple dichasium; illus. *h*),

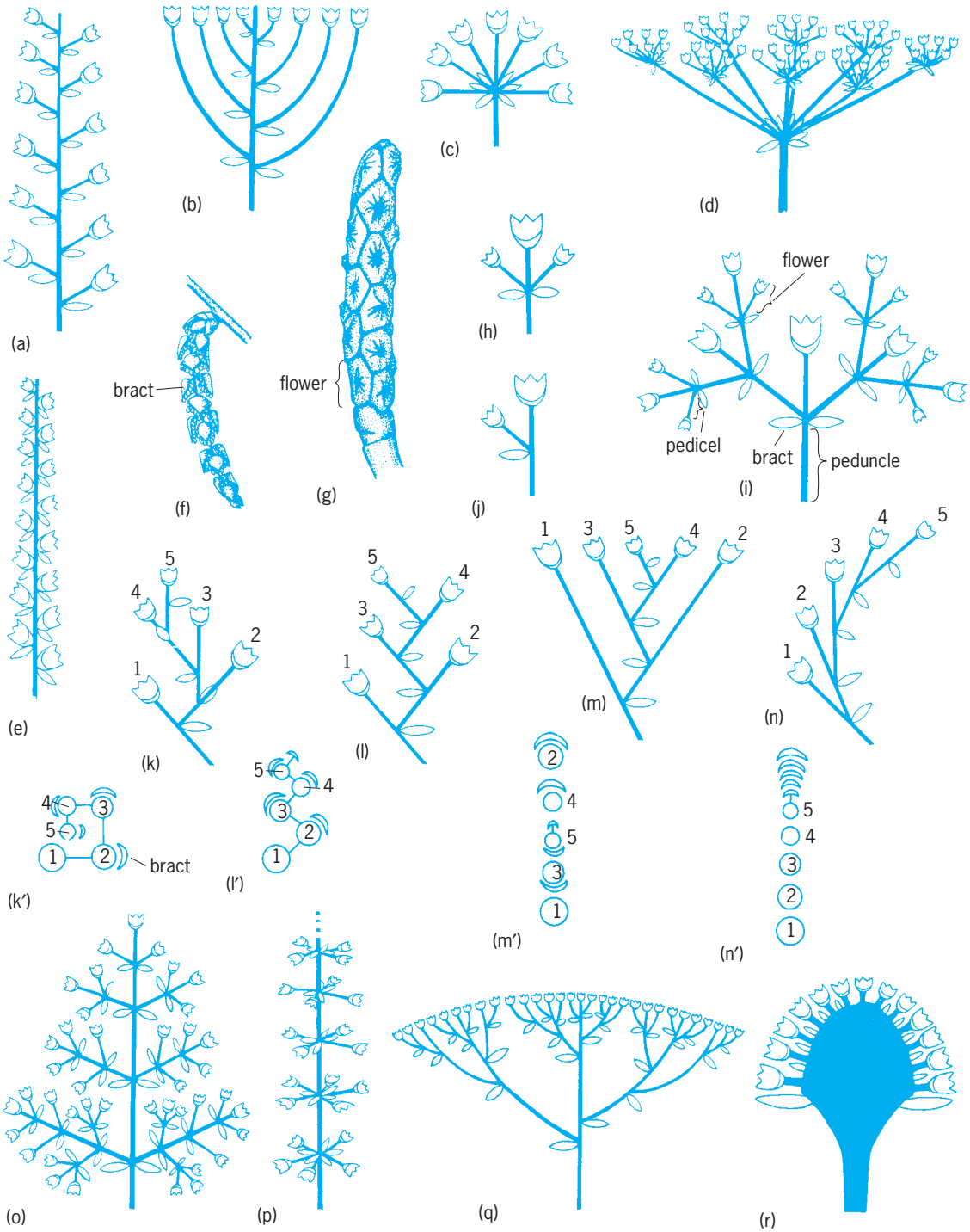

Examples of inflorescence types. All are side views except *k′*, *l′*, *m′*, and *n′* which are top views. Numbers in *k–n′* indicate flowers in order of development. (*a*) Raceme. (*b*) Corymb. (*c*) Umbel. (*d*) Compound umbel. (*e*) Spike. (*f*) Catkin. (*g*) Spadix (fruiting, *Ludovia integrifolia*). (*h*) Dichasium. (*i*) Compound dichasium. (*j*) Monochasium. (*k*, *k′*) Bostryx. (*l*, *l′*) Cincinnus. (*m*, *m′*) Rhipidium. (*n*, *n′*) Drepanium. (*o*) Panicle. (*p*) Verticillate inflorescence. (*q*) Cyme. (*r*) Head.

a three-flowered cluster composed of an axis which bears a terminal flower and, below it, two bracts which subtend lateral flowers; compound dichasium (illus. *i*), composed solely of dichasia; monochasium (simple monochasium; illus. *j*) a two-flowered cluster composed of an axis which bears a terminal flower and, below it, one bract which subtends a lateral flower; compound monochasium (illus. *k–n′*), composed solely of monochasia; bostryx (helicoid cyme; illus. *k* and *k′*), a compound monochasium or modified compound dichasium in which successive pedicels describe a spiral; cincinnus (scorpiod cyme, illus. *l* and *l′*), like a bostryx, but successive pedicels follow a zigzag path; rhipidium (illus. *m* and *m′*), a bostryx compressed so that all flowers lie in one plane; drepanium (illus. *n* and *n′*), a cincinnus

compressed so that all flowers lie in one plane; panicle (illus. *o*), composed of a monopodial central axis which bears flowers indirectly on branches of higher order and which also may bear some flowers directly; thyrse, a compact panicle of more or less cylindrical form; verticillate inflorescence (illus. *p*), like either a raceme or panicle, but the central axis bears flowers in whorls, approximate whorls, or contracted clusters which simulate whorls; cyme (illus. *q*), a compound, more or less flat-topped sympodial inflorescence, sometimes simulating a compound umbel but having branches less regularly disposed; and head (capitulum; illus. *r*), compact inflorescence having a very short, often discoid or globose flower-bearing portion.

Formerly, the aforementioned inflorescence types were classified under two main categories, indeterminate (racemose) and determinate (cymose); however, this classification is now obsolete. *See* FLOWER.

George J. Wilder

Bibliography. H. C. Bold et al., *Morphology of Plants and Fungi*, 1980; H. W. Rickett, The classification of inflorescences, *Bot. Rev.*, 10:187–231, 1944; H. W. Rickett, Materials for a dictionary of botanical terms, III: Inflorescences, *Bull. Torrey Bot. Club*, 82:419–445, 1955.

Influenza

An acute respiratory viral infection characterized by fever, chills, sore throat, headache, body aches, and severe cough; it is commonly known as the flu. While many viruses can cause respiratory infections such as the common cold, influenza viruses are more likely to cause severe illness and to result in serious medical complications. Although gastrointestinal symptoms may sometimes accompany influenza infection, especially in children, the term "stomach flu" is a misnomer for gastrointestinal illnesses caused by a variety of other viruses, bacteria, or other agents. *See* VIRUS.

Another unique feature of influenza is its distinctive seasonal pattern. Outbreaks of influenza occur during cold weather months in temperate climates, and typically most cases cluster during a period of 1–2 months, in contrast to broader periods of circulation with many other respiratory viruses. During annual influenza epidemics, an increased death rate due to influenza-related complications is often observed, and is one of the ways that public health authorities monitor epidemics. During an average influenza season in the United States, more than 200,000 people are hospitalized and 36,000 die as a result of such complications. *See* PUBLIC HEALTH.

Viral agents. The bacterium *Hemophilus influenzae* was once mistakenly considered the cause of influenza, based on finding the microbe in the sputum of individuals with the disease. However, viruses were shown to be the cause in 1933, when the first influenza virus, designated type A, was isolated. The second major type, designated type B, was identified in 1940, and the third, type C, in 1949.

The three types of influenza viruses are classified in the virus family Orthomyxoviridae, and they are similar, but not identical, in structure and morphology. Types A and B are more similar in physical and biologic characteristics to each other than they are to type C; and unlike type C viruses, they are responsible for widespread illness during epidemics. Influenza viruses may be spherical or filamentous in shape, and they are of medium size among the common viruses of humans. There are two important viral protein antigens on the virion surface, hemagglutinin and neuraminidase. Hemagglutinin functions in the attachment of the virus to cell receptors and initiates infection by anchoring the virus to respiratory epithelial cells. Neuraminidase is an enzyme that facilitates release of virus from infected cells and aids in the spread of virus in the respiratory tract. *See* ANTIGEN.

Influenza type A viruses are divided into subtypes based on the hemagglutinin and neuraminidase proteins. There are 16 different hemagglutinin subtypes and 9 neuraminidase subtypes. Currently there are two influenza A subtypes circulating among humans, but all hemagglutinin and neuraminidase subtypes have been found in wild birds, which are believed to be the natural reservoir for all influenza A subtypes. Influenza type B does not have an animal reservoir and is not divided into subtypes.

Antigenic drift and shift. Influenza type A and B viruses undergo changes over time due to an accumulation of point mutations in the viral proteins, especially hemagglutinin and to a lesser extent neuraminidase. This is referred to as antigenic drift, and necessitates changes to at least one of the three virus strains in the influenza vaccine almost every year. These changes are also responsible for the fact that people remain susceptible to influenza infections during their lives, since antibodies produced in response to infection with a particular influenza strain may not recognize the mutated viruses and therefore may provide little or no protection against them as the strains continue to mutate. Influenza viruses that currently cause epidemics among humans are influenza types A(H1N1), A(H3N2), and B. *See* MUTATION.

Influenza A viruses may also undergo sudden changes known as antigenic shift, which results in a virus with a new hemagglutinin, neuraminidase, or both. When this occurs, a new virus subtype emerges to which most or all people have no immunity, resulting in a worldwide influenza epidemic, or pandemic. Influenza pandemics occurred in 1918–1919, 1957–1958, and 1968–1969. The 1918–19 pandemic (the Spanish influenza pandemic) was by far the most severe, and is estimated to have caused 20–40 million deaths worldwide with 675,000 in the United States alone. The reasons for the unusually severe illnesses and high death rates during this pandemic have long puzzled influenza researchers, and for many years little was known about the characteristics of the causative virus. However, in recent years a team of scientists has recovered influenza virus genetic material (RNA) from formalin-fixed autopsy samples

taken from soldiers who died of influenza during the 1918–1919 pandemic and from frozen lung tissues of a person who died of influenza and was buried in the permafrost (perennially frozen ground) in Alaska in 1918. The genetic material recovered from these sources has been used to reconstruct the 1918 Spanish influenza pandemic strain. Studies of this reconstructed virus will greatly enhance the understanding of that pandemic virus, and perhaps will help scientists better detect viruses with pandemic potential and develop strategies to help mitigate the effects of pandemic influenza in the future. *See* EPIDEMIC.

Animal influenza. In addition to infecting humans, strains of influenza A infect many other species, including pigs, horses, seals, whales, and a wide variety of avian species. The severity of influenza infection in animals ranges from asymptomatic to fatal.

Animal influenza viruses are believed to play an important role in the emergence of human pandemic strains. One way that antigenic shift can occur is when a human and an animal influenza A virus strain infect the same host. When a cell is infected by two different type A viruses, various combinations of the original parental viruses may be assembled into the new progeny; thus, a progeny virus may be a mixture of gene segments from each parental virus and therefore may gain a new characteristic, for example, a new hemagglutinin. Hence a new virus is introduced into a population that has no immunity against it.

Since pigs are susceptible to infection with both human and avian influenza strains, this type of antigenic reassortment could occur when a pig is simultaneously infected with human and avian influenza strains. This can happen in settings where people, pigs, and a variety of avian species live in proximity, a circumstance common in many parts of the world. Since in recent years it has been observed that humans can also become infected with some strains of avian influenza viruses, it is also theoretically possible for reassortment to occur if a human is simultaneously infected with an avian and a human influenza virus.

The first human illnesses associated with an outbreak of a highly pathogenic A(H5N1) virus in birds were reported in 1997. This virus was widespread among poultry throughout live bird markets in Hong Kong, and led to 18 known human illnesses and 6 deaths. Further spread of this virus in Hong Kong was prevented by destroying 1.5 million birds. In 2003 the A(H5N1) virus was discovered in poultry in Vietnam and has since spread to other countries in Asia, Europe, and Africa, where it has been found in poultry and wild birds. By July 2006 a total of 231 cases in humans had been confirmed in 10 countries, of which 133 were fatal. Although most human cases of "bird flu" are believed to have been contracted by contact with infected birds, it is possible that the virus might change, either by reassortment or mutation of the viral genes, and become easily transmitted from human to human, causing a pandemic. This has prompted governments and health agencies to rush development of vaccines and to develop pandemic preparedness plans as a precaution, as well as to intensify surveillance of this and other avian influenza viruses with similar potential, and to destroy infected poultry when outbreaks are detected. *See* ANIMAL VIRUS.

Pathogenesis. When influenza virus enters the respiratory tract, usually by inhalation of aerosolized (suspended in air) viruses from an infected person, the virus hemagglutinin attaches to epithelial cells and invades the cells. After attachment, the cell's normal defense mechanisms actually help the virus gain entry and replicate. The cell engulfs the virus and attempts to destroy it with enzymes. However, instead of destroying the virus, the enzymes allow the viral RNA proteins to spill into the cell and move to the cell nucleus. Replication of viral RNA begins rapidly in the nucleus of the host cell. Final assembly of new virus occurs at the plasma membrane, and new viruses spread to nearby cells. The virus has a short incubation period. There is only a period of 1–3 days between infection and illness, and this leads to the abrupt development of symptoms that is a hallmark of influenza infections. The virus is typically shed in the throat for 5–7 days, during which time the sick person can infect others. Complete recovery from uncomplicated influenza usually takes several days to a week, and the individual may feel weak and exhausted for a week or more after the major symptoms disappear.

Numerous medical complications can occur as a result of influenza infection. Pulmonary complications include secondary bacterial and primary viral pneumonias. Viral pneumonia is far less common than bacterial pneumonia, but the mortality rate is much higher. However, because it occurs much more frequently, bacterial pneumonia is responsible for many more hospitalizations and deaths during influenza epidemics. Other pulmonary complications of influenza include croup in infants and young children and worsening of chronic lung conditions. Even previously healthy people can show alterations in pulmonary function that can last for weeks after the initial infection. Influenza can also cause inflammation of the heart muscle (myocarditis) or of the sac around the heart (pericarditis), and may worsen chronic heart conditions. It can also cause encephalitis (inflammation of the brain), kidney failure, and Reye's syndrome, which affects the liver and central nervous system. Reye's syndrome occurs almost exclusively in children, and has been associated with the use of aspirin to treat influenza (as well as chickenpox). For this reason it is recommended that aspirin substitutes such as acetaminophen be used when treating children for suspected viral infections. *See* ASPIRIN; HERPES; PNEUMONIA.

Prevention and control. Vaccination is the primary means of preventing influenza. Because at least one of the three virus strains in the influenza vaccine is changed during most years due to viral mutations, and because vaccine-induced antibody can wane over the course of a year, the vaccine must be taken

every year before the influenza season. Currently there are two types of influenza vaccines. One is made from inactivated, or killed, viruses, and the other from living viruses that have been attenuated, or weakened. Anyone over 6 months of age can take the inactivated vaccine, but the live attenuated vaccine is recommended only for those between the ages of 5 and 49 years who are not pregnant. The inactivated vaccine is administered by injection, but the live attenuated vaccine is delivered as a nasal spray. With both vaccines, antibodies develop within about 2 weeks of vaccination, but the live virus vaccine may better stimulate the cell-mediated immune response and provide broader and longer-lasting immunity, especially in children and younger adults. How effective the vaccine is in preventing influenza also depends on how healthy the immune system of the recipient is, and on how closely the virus strains in the vaccine resemble those that circulate during the influenza season. Since the vaccine strains must be chosen well in advance of the influenza season, one or more viruses may continue to mutate to an extent that causes the vaccine to be less effective. However, even when this happens, people who have been vaccinated usually have a shorter and less severe illness if they are infected, and are less likely to develop complications than people who were not vaccinated.

Antiviral medications can be used to prevent or to treat influenza. These drugs act by inhibiting replication of the virus and, to prevent influenza, must be taken daily during the period of possible exposure. The drugs can also decrease the severity and lessen the duration of the illness if taken soon after symptoms begin. Two of these drugs, amantadine (Symmetrel) and rimantadine (Flumadine), classed as M2 inhibitors (since they inhibit viral protein M2, which uncoats the virus's protein shell, a process needed for viral replication in the host cell), are effective only against type A viruses. While the drugs were used effectively for many years, influenza viruses can develop resistance during treatment, and in recent years so many resistant strains have been detected that the two are no longer the drugs of choice. Newer drugs, Osletamivir (Tamiflu) and Zanamivir (Relenza), classified as neuraminidase inhibitors, are used to treat influenza with relatively few problems of resistance to date. These drugs are active against both type A and type B viruses. *See* BIOTECHNOLOGY; VACCINATION. John M. Quarles; Nancy H. Arden

Bibliography. J. A. Cabezas et al., New data on influenza virus type C confirm its peculiarities as a new genus, *Intervirology*, 32:325–326, 1991; R. B. Couch et al., Influenza: Its control in persons and populations, *J. Infect. Dis.*, 153:431–440, 1986; J. M. Katz, Preparing for the next influenza pandemic, *ASM News*, 70:412–419, 2004; E. D. Kilbourne, *Influenza*, 1987; M. W. Shaw, N. H. Arden, and H. F. Maassab, New aspects of influenza viruses, *Clin. Microbiol. Rev.*, 5:74–92, 1992; T. M. Tumpey et al., Characterization of the reconstructed 1918 Spanish influenza pandemic virus, *Science*, 30:77–80, 2005.

Information management

The functions associated with managing the information assets of an enterprise, typically a corporation or government organization. Increasingly, companies are taking the view that information is an asset of the enterprise in much the same way that a company's financial resources, capital equipment, and real estate are assets. Properly employed, assets create additional value with a measurable return on investment. Forward-looking companies carry this view a step further, considering information as a strategic asset that can be leveraged into a competitive advantage in the markets served by the company.

In some respects, the term "information management" is interchangeable with management information systems (MIS); the choice of terminology may reflect the attitude of a company regarding the value of information. Companies using the traditional MIS descriptor are likely to hold the view that management information systems represent a cost to be minimized, in contrast to the strategic-asset view with the objective of maximizing return on investments.

For the large enterprise with multiple business units and functional elements distributed over a broad geographic area, information management can be an enormously complex task. Its successful accomplishment is dependent not only on the diligent application of professional skills, but even more on leadership and management abilities strongly rooted in a thorough knowledge of and insight into the enterprise.

Scope. The scope of the information management function may vary between organizations. As a minimum, it will usually include the origination or acquisition of data, its storage in databases, its manipulation or processing to produce new (value-added) data and reports via application programs, and the transmission (communication) of the data or resulting reports. While many companies may include, with good reason, the management of voice communications (telephone systems, voice messaging, and, increasingly, computer-telephony integration or CTI), and even intellectual property and other knowledge assets, this article focuses primarily on the data aspects of information management.

Information versus data. Executives and managers frequently complain of "drowning in a flood of data, but starving for information." There is a significant difference between these two terms. Superficially, information results from the processing of raw data. However, the real issue is getting the right information to the right person at the right time and in a usable form. In this sense, information may be a perishable commodity. Thus, perhaps the most critical issue facing information managers is requirements definition, or aligning the focus of the information systems with the mission of the enterprise. The best technical solution is of little value if the final product fails to meet the needs of users.

Information engineering. One formal approach to determining requirements is information engineering. By using processes identified variously as business systems planning or information systems planning, information engineering focuses initially on how the organization does its business, identifying the lines from where information originates to where it is needed, all within the context of a model of the organization and its functions. While information systems personnel may be the primary agents in the information engineering process, success is critically dependent on the active participation of the end users, from the chief executive officer down through the functional staffs.

An enterprise data dictionary is a principal product of the information engineering effort. This dictionary is a complete catalog of the names of the data elements and processes used within the enterprise along with a description of their structure and definition. The dictionary may be part of a more comprehensive encyclopedia, which is typically a formal repository of detailed information about the organization, including the process and data models, along with descriptions of the design and implementation of the organization's information systems applications.

A major advantage of the application of information engineering is that it virtually forces the organization to address the entire spectrum of its information systems requirements, resulting in a functionally integrated set of enterprise systems. In contrast, ad hoc requirements may result in a fragmented set of systems (islands of automation), which at their worst may be incompatible, contain duplicate (perhaps inconsistent) information, and omit critical elements of information.

Information systems architecture. Information engineering develops a logical architecture for an organization's information systems. Implementation, however, will be within the context of a physical systems architecture. Information engineering ensures functional integration, whereas systems engineering provides for the technical integration of information systems. Together, functional and technical (logical and physical) integration lay the architectural foundation for the integrated enterprise. *See* SYSTEMS ARCHITECTURE; SYSTEMS ENGINEERING.

Open versus proprietary architectures. Many critical decisions are involved in establishing a physical systems architecture for an organization. Perhaps the single most important decision is whether the architecture will be based on a single manufacturer's proprietary framework or on an open systems framework. Although the former may be simpler to implement and may offer less technical risk, it is far more restrictive with regard to choices of equipment. An open systems architecture allows for the integration of products from multiple vendors, with resulting increased flexibility in functional performance and the potential for significant cost savings in competitive bidding. The flexibility to incorporate new products is particularly important in an era of rapid technological advancement, where new technology may become obsolete in 18 months.

Client-server and network computing. One of the most widely used architectural models is the client-server model. In this model, end users, for example, may employ personal computers as clients to request data from an enterprise database resident at another location on a minicomputer or mainframe (the server). Clients and servers interconnect and communicate with one another over networks, from local-area networks (LANs) within a single facility, to wide-area networks (WANs) which may span one or more continents. A variation on this model (sometimes referred to as network computing) employs "thin" clients, with the applications running on servers accessed over the network. *See* CLIENT-SERVER SYSTEM; DISTRIBUTED SYSTEMS (COMPUTERS); LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

Operating systems. An extremely important architectural consideration is the choice of operating systems. This selection may determine the options available in acquiring off-the-shelf software packages. Alternatively, an existing (legacy) set of software application programs (normally representing a major investment) may severely constrain the choice of hardware on which the programs will run. The issue here is whether the operating systems are proprietary to a single vendor's hardware or are vendor-independent (such as UNIX). Because of the portability it offers, UNIX (and more recently a derivative of UNIX called LINUX) has become a de facto standard in the industry. Although it is actually proprietary, Microsoft's Windows family of operating systems has also achieved de facto standard status due to its increasingly ubiquitous availability on most major hardware platforms. *See* OPERATING SYSTEM.

Databases. Another major architectural decision is the structure of the databases to be utilized. Most database products follow a hierarchical, relational, or object-oriented paradigm. Although the most widely used structure has been relational, increasing attention is being given to object-oriented systems. Information engineering tends to separate data from applications, while object-oriented methods couple the two. *See* DATABASE MANAGEMENT SYSTEM; OBJECT-ORIENTED PROGRAMMING.

Internet and related technologies. The rapid rise of the Internet since the mid-1990s has profoundly impacted information systems architectures as well as the role of information systems within an organization. With tens of millions of individuals, corporations, and even governments tied in to the Internet, there are few businesses that do not reach out to their customer base via a Web site, thus expanding their information systems architecture to encompass the extended enterprise. Equally important, the same technology that drives the Internet (the Internet Protocol or IP) has made possible intranets, which function like the Internet but are private to the organization. The corporate Web site involves the marketing and public relations organizations as major information systems users, and even the sales organization may be involved, as actual sales are made over the Web,

giving rise to "electronic commerce" (e-commerce). *See* INTERNET; WORLD WIDE WEB.

Software applications. The information engineering process can produce a complete set of requirements for the applications needed to run the enterprise. When all architectural parameters have been determined, unique applications can be developed, preferably employing computer-aided software engineering (CASE) tools to automate the process to the maximum degree possible. In general, however, the cost and time involved in developing custom software, and the even greater cost of supporting (maintaining) that software throughout its life cycle, dispose strongly toward the maximum use of off-the-shelf software packages. *See* SOFTWARE ENGINEERING.

There are many functions that are common to almost any organization, such as personnel (human resources) and finance. Applications that support these and other basic infrastructure functions, such as manufacturing, are readily available off-the-shelf from major vendors in the form of enterprise resource planning (ERP) software packages that have become a major element of the software industry. Although considerable work is required either to configure the vendor's software to match the organization's internal processes, or to alter the processes to match the software, or a combination of the two, the net cost in time and dollars (including maintenance) often makes this option more attractive than developing custom applications from scratch.

Functions that are unique to an organization may justify the cost and risk associated with custom software development, but these decisions should not be made without careful analysis. Strategic applications that produce a compelling competitive advantage are most likely to meet this test.

Enterprise application integration. The proliferation of Web and e-commerce applications, as well as the trend toward enterprise resource planning systems, has posed an additional major challenge for information managers: successfully integrating these new applications with the enterprise's legacy systems, or enterprise application integration (EAI). *See* INFORMATION SYSTEMS ENGINEERING; SYSTEMS INTEGRATION.

Alan B. Salisbury

Bibliography. S. Alter, *Information Systems: A Management Perspective*, 1999; J. Martin, *Information Engineering*, 1989; A. Sage, *Systems Management for Information Technology and Software Engineering*, 1996; T. Wheeler, *Open Systems Handbook*, 1992.

Information processing (psychology)

The coding, retrieval, and combination of information in perceptual recognition, learning, remembering, thinking, problem solving, and performance of sensory-motor acts.

Coding. Information is coded in some type of memory by having something in the memory represent the information. In books, recording tapes, and almost all digital computer memories, the memory consists of a long one-dimensional series of memory locations. The encoding of a piece of information such as the proposition "John plays tennis" consists of imposing the patterns representing "John," "plays," and "tennis" into adjacent locations in the memory. Any other piece of information, such as the proposition "Tennis is a sport," learned at a different time would be coded in a similar manner in a totally different set of locations in memory. The same pattern would be used to represent the concept "tennis," which is common to both propositions, but this pattern would be stored in two separate places, and an information processor using the memory would not have direct access to both facts about "tennis" when it had retrieved one fact. An additional search of memory would be required. Such a memory is called nonassociative, because there are no direct associations linking all instances of the same concept or idea in memory.

Human memory is not at all like a tape recorder or computer memory. Human memory is associative. From a coding standpoint, the defining property of an associative memory is the use of specific node encoding, in which each idea is represented only once by a designated node (location). The representative of an idea is a particular location in memory, rather than a particular pattern that can be, and often is, stored in a variety of locations. There are many different types of associative memories; two of these are considered below: horizontal associative memory and vertical associative memory.

In a horizontal associative memory, the coding of propositions consists of forming a strong association directly from each concept node to the following concept node (**Fig. 1**). From a semantic viewpoint, the association from one node to another means "is followed by." Careful study of Fig. 1 reveals the two primary deficiencies of a horizontal associative memory. First, there is no encoding of where one proposition ends and another begins, a difficulty which might be referred to as "run-on propositions." Second, there is a serious associative interference problem. For example, there is no way to know whether John plays tennis, basketball, or both, and similarly for Peter. The ability of an associative memory to integrate information that involves the same concepts can produce excessive and erroneous integration in the case of a horizontal associative memory.

Vertical associative memories solve these problems while retaining the desirable feature of specific node encoding. They associate atomic concepts to higher-order chunk (phrase) concepts. These associations mean that the lower-order concept "is a constituent of" the higher-order (phrase or propositional) concept. All of the concepts in a single proposition are associated directly or indirectly to a single higher-order node that represents the entire proposition (**Fig. 2**). In this way both the run-on proposition and associative interference problems are solved by a hierarchical tree-structure encoding of exactly which concepts belong in each proposition.

Associations are assumed to be unidirectional, that is, from one node to another. Information processing in an associative memory requires the ability to get from a chunk node to its constituents, and vice versa. Semantically, the associations from a chunk to its constituents mean "has as a constituent." Since the meaning of vertical associations in the two directions is logically opposite, it seems likely that these associations are represented in the mind (brain) by structurally dissimilar types of links that function somewhat differently when information is retrieved from a person's memory. Many theories of associative memory assume an even larger variety of links than this, but at least these two types appear to be necessary.

Order and context-sensitive coding. There are many chunks whose constituents have no unique ordering, such as the members of a family or the features of many objects. Although the sentences used to communicate propositions to others have a definite word order, the same proposition can be communicated using virtually any ordering of the constituent concepts, provided the proper syntactic plan is used. Semantically, there does not appear to be any unique ordering of the constituents of a proposition. Hence, it is likely that the encoding of propositions is another example of the encoding of an unordered set.

However, there are numerous examples of chunks encoded in the mind whose constituents are uniquely ordered, for example, the episodes of a story or the letters in a word. How are these ordered sets coded? A number of possibilities have been suggested, but the one that appears to have the most evidence supporting it is context-sensitive coding. According to this theory, the constituents of a chunk representing an ordered set such as the word "clank" are nodes representing overlapping triples of the component segments. In the case of "clank" the constituents would be "#Cl," "cLa," "lAn," "aNk," and "nK#." The reason for distinguishing the middle element in each triple is that it is the central or focal element and the two on each side are the context. In speech production, the node "lAn" would control the speech musculature to produce the particular "a" that is preceded by "l" and followed by "n." If horizontal associations are added to the constituents of a chunk as a third type of link in vertical associative memory, then horizontal associations between pairs of context-sensitive nodes will produce a unique ordering of constituents when there is one and no such ordering when there is not. In the "clank" example, sufficient information has been encoded to know the first constituent of the set, given that the entire set is active in the mind due either to bottom-up activation by stimulus presentation of the word "clank" or to top-down activation of the segmental constituents by the word node for "clank." In both recognition and production of the work "clank," there is only one segmental node of the form "#Xy." Since "#" represents the junction between words, this must be the first node in the word. Given that "#Cl" is the first node, "cLa" will have the strongest horizontal association from "#Cl," so one knows it is the second node, and so forth.

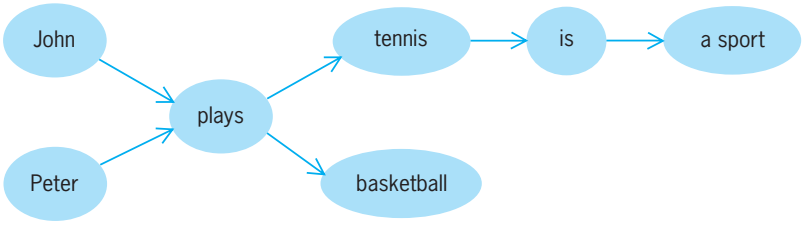

Fig. 1. Horizontal associative memory with a serious associative interference problem (who plays tennis—John, Peter, or both?) and no encoding of where each proposition begins and ends.

Learning. All learning in an associative memory can be considered to be the strengthening of different associations between nodes. Thinking of a set of ideas at about the same time (within approximately a second) activates the nodes representing the ideas at about the same time. In a way that is not understood, this strengthens vertical associations gluing these idea nodes to higher-order chunk nodes and probably also strengthens horizontal associations among the idea nodes. If a set of idea nodes is simultaneously activated, then all horizontal associations are equally strong on the average (all orders are encoded), and no unique ordering is encoded. If the set of idea nodes is activated in a simple linear order, than a unique set of horizontal associations will be encoded to represent the sequence (ordered set).

For certain purposes it may be useful to distinguish among three types of learning: chunking (forming

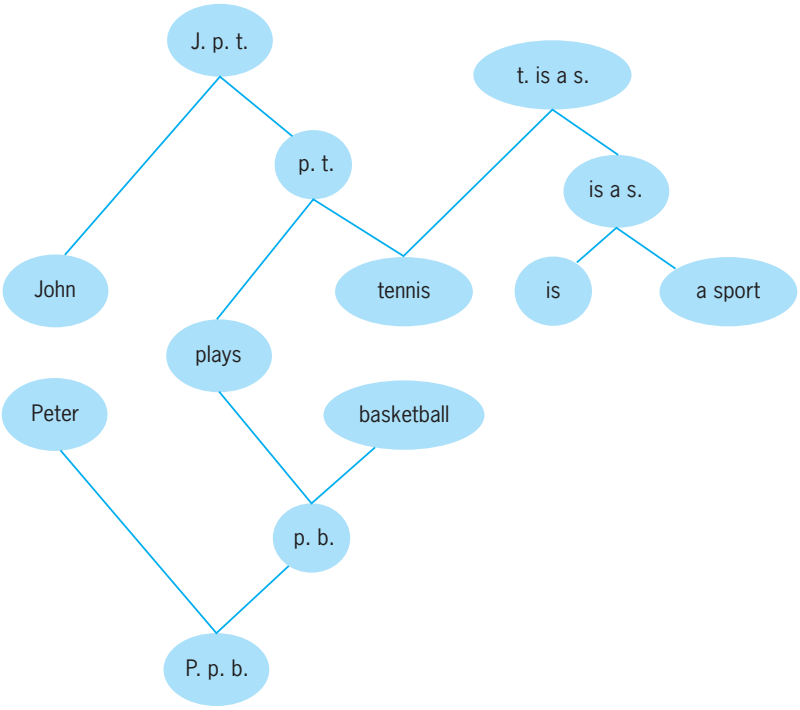

Fig. 2. Vertical associative memory where concepts are grouped into phrases and propositions by associating each concept to a higher-order node that represents the entire group.

new vertical associations from a set of old nodes to a new node, specifying that node to represent the set of old nodes); sequencing (forming new horizontal associations between adjacent pairs of nodes in an ordered set); and strengthing (increasing the strength of existing associations). Forming new associations refers to the initial increase in strength above the baseline, or zero, value, so all learning consists of the strengthening of associations in the most general sense.

However, the initial strengthening of associations appears to require more thinking than does the further strengthening of previously learned associations. Chunking and sequencing are greatly facilitated by coding new material so as to integrate it with stored material as much as seems relevant and as consistently as possible. When such integration uses relationships that are true and of considerable general value, it is called understanding. When such integration uses fictional relationships or relationships of little general value, it is called a mnemonic trick. Either way, new learning is made easier by integration with prior knowledge. The difference is that all associations formed or strengthened by understanding are true and useful, as opposed to only a few of the associations formed or strengthened by mnemonic tricks.

Further strengthening of learned associations requires much less thinking effort—just retrieval of the associations to be rehearsed. However, this does not mean that such review is of no value. Review can serve to counteract the effects of prior forgetting, to overlearn so as to retard future forgetting, to increase retrieval speed for the reviewed material in the future, and to reduce the attentional demands of such retrieval.

Retrieval. Information that is currently being thought about is said to be in active memory. All of the rest of the information that is stored in long-term memory is said to be in passive memory. Retrieval refers to the cognitive processes that allow one to make use of stored information by converting the memory traces for particular pieces of information from the passive to the active state. Retrieval can be analyzed into two components: recognition, which is the activation of the nodes in memory that encode the cue information. Every act of retrieval requires recognition of the retrieval cues, and probably every act of retrieval involves recall of some information beyond that given in the retrieval cues.

The basic information-retrieval process in an associative memory is direct-access, that is, direct activation of the nodes encoding the cue and closely associated information. No search is required when the retrieval cues are sufficient to identify a unique memory node (location). When the retrieval cues are inadequate, there is sometimes information stored in long-term associative memory that would be adequate to retrieve a unique answer to the question. In such cases, a sequence of direct-access retrieval processes is initiated. For example, sometimes the associations from an acquaintance's face to his or her name is too weak to retrieve the name from the face cues. However, when the backward associations from the name to the face cues are added in, the information stored in memory is sufficient to recognize which name goes with that face. So one can engage in a memory search process that consists of recalling various names and judging the strength of association of each name with the face. The entire sequence is a search process, but each elementary act of retrieval in this complex sequence is a direct-access process. The whole purpose of having an associative memory is to make direct-access retrieval possible, by encoding every idea in one specific location that can be identified from the cue input without having to search through sets of alternative locations.

Retrieval dynamics. While retrieval from memory is direct, it is not instantaneous. One elementary act of retrieval requires approximately 1 s of time to reach completion, with a range from 0.3 to 3 s. Retrieval can be analyzed into three temporal phases: latent, transition, and terminal (**Fig. 3**). The latent phase extends from the time the retrieval cues are first presented until the initial time (I) that any response can be made which is at all influenced by the memory traces activated by the retrieval cues. The terminal phase is the period of time after retrieval is complete during which the retrieved set of memory traces is fully activated and able to control behavior and thought. In the terminal phase, retrieved memory strength is equal to the limit set by the passive strength level stored in long-term memory (L). The transition phase is the period in between, when retrieved memory strength increases from zero to L. This transition phase is not abrupt, as if retrieval occurred all at once. Instead, memory retrieval appears to occur incrementally over a period of many tenths of a second. If people respond to retrieval cues after some retrieval has occurred, but before the retrieval of the memory trace is complete, they respond at a level of accuracy which is greater than chance but less than they could achieve had they delayed their response until retrieval was complete. A plot

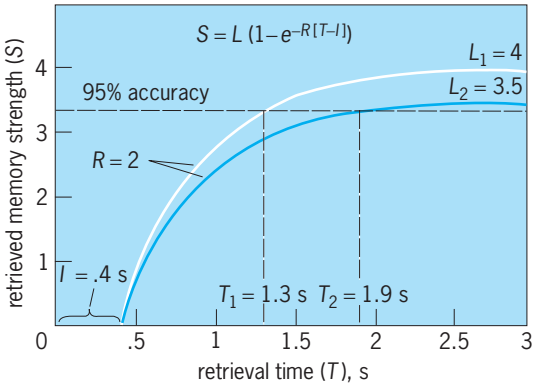

Fig. 3. Memory-retrieval functions for conversion of stored long-term memory strength *L* into active short-term memory strength *S* as a function of retrieval time *T*. Note that the time to achieve 95% accuracy is greatly affected by stored strength *L*, even when retrieval dynamics parameters (*I* + *R*) are identical.

of accuracy as a function of reaction time is called a speed-accuracy tradeoff function. The accuracy of a recognition or recall response to retrieval cues measures the strength of the retrieved memory trace. The increase in the strength of the retrieved memory trace as a function of the time allowed for processing the retrieval cues (retrieval time) is called the retrieval function. Empirical speed-accuracy tradeoff functions for recognition or recall thus provide measures of the theoretical retrieval functions for these tasks.

The dynamics of memory retrieval appears to be closely approximated by an exponential approach to a limit L at rate R after the initial delay time I (Fig. 3). Exponential approach to a limit means that in every very small increment of time (after delay time I), a constant percentage of what stored trace remains to be retrieved is retrieved. That percentage is the retrieval rate R. In Fig. 3, two retrieval functions are displayed that have identical retrieval dynamics (identical I and R parameters) but different limits set by their different levels of stored strength in long-term memory (L_1 and L_2). Note that the time required for memory retrieval to reach a certain level of accuracy (95% in Fig. 3) is considerably less for the trace with the greater strength in storage ($T_1 < T_2$). Thus, even without altering the dynamics of memory retrieval, one can decrease reaction time considerably in recall or recognition simply by increasing the strength of the stored trace in long-term memory, with no loss in accuracy.

Retrieval dynamics (I and R parameters) is not affected by memory strength. It is primarily affected by the hierarchical level of encoding of the memory trace being retrieved—the higher the coding level (that is, the longer the chain of upward vertical associations from the sensory periphery to the highest level chunk node), the slower the retrieval dynamics (that is, the greater the I value and the smaller the R value). Thus, the retrieval dynamics for recognizing letters is faster than that for words, which is in turn faster than that for sentences. The more simply one codes material in terms of hierarchical depth, the faster it can be retrieved, provided the strength of the trace is the same in both cases.

Attention and parallel processing. There are limits to how many things a person can attend to at once. The laws of attention are not completely understood, but the following appear to be true: A very large number of retrieval cues can be processed at once, provided they all converge on a single stored memory trace, regardless of complexity, which has been previously unified by vertical associations to a single chunk node at the top of its trace hierarchy; this is called convergent parallel processing. Two or three separate sets of retrieval cues can probably be processed, provided that the traces they retrieve are very strong; this is called divergent parallel processing. Divergent parallel processing is easier if the traces are stored in different memory modalities, for example, verbal versus spatial. Thus, a person can carry on a conversation and sign letters simultaneously, with little loss of effectiveness in either task, because the memory traces are in somewhat different modalities and the signature trace is very strong. One of the important aspects of increasing skill in sports, for instance, is the increasing ability to perform several information-processing tasks in parallel. Part of this skill is due simply to increased practice at each component skill and part is due to explicit coordination of these components under unified, higher-order chunk nodes. *See* COGNITION; MEMORY. Wayne A. Wickelgren

Bibliography. J. R. Anderson, *Cognitive Psychology and Its Implications*, 4th ed., 1999; K. Haberlandt, *Cognitive Psychology*, 2d ed., 1996; D. E. Rumelhart and J. L. McClelland, *Parallel Distributed Processing*, 1987; A. Trehub, *The Cognitive Brain*, 1991.

Information systems engineering

The process by which information systems are designed, developed, tested, and maintained. The technical origins of information systems engineering can be traced to conventional information systems design and development, and the field of systems engineering. Information systems engineering is by nature structured, iterative, multidisciplinary, and applied. It involves structured requirement analyses, functional modeling, prototyping, software engineering, and system testing, documentation, and maintenance.

Information systems. Modern information systems solve a variety of data, information, and knowledge-based problems. In the past, most information systems were exclusively data-oriented; their primary purpose was to store, retrieve, manipulate, and display data. Application domains included inventory control, banking, personnel record keeping, and the like. The airline reservation system represents the quintessential information system of the 1970s. Since then, expectations as to the capabilities of information systems have risen considerably. Information systems routinely provide analytical support to users. Some of these systems help allocate resources, evaluate personnel, and plan and simulate large events and processes. This extremely important distinction defines the range of application of modern information systems. The users expect information systems to perform all the tasks along the continuum shown in **Fig. 1**. In the 1980s the applications expanded to data-oriented and analytical computing.

Systems engineering. Systems engineering is a field of inquiry unto itself. There are principles of applied systems engineering, and a growing literature defines a field that combines systems analysis, engineering, and economics. Systems engineering extends over the entire life cycle of systems, including requirement definitions, functional designs, development, testing, and evaluation. The systems engineer's perspective is different from that of the product engineer, software designer, or technology developer. The product engineer deals with detail, whereas the systems engineer takes an overall viewpoint. Where the product engineer deals with internal operations,

Data-Oriented Computing		Analytical Computing	
Physical tasks	Communicative tasks	Perceptual tasks	Mediational tasks
• file • store • retrieve • sample	• instruct • inform • request • query	• search • identify • classify • categorize	• plan • evaluate • prioritize • decide
Analytical Complexity Continuum			

Fig. 1. Data-oriented and analytical computing, suggesting the range of information systems applications.

the systems engineer deals more extensively with the external viewpoint, including relations to other systems, users, repairers, and managers. Systems engineering is based upon the traditional skills of the engineer combined with additional skills derived from applied mathematics, psychology, management, and other disciplines. The systems engineering process is a logical sequence of activities and decisions that transform operational needs into a description of system performance configuration. *See* SYSTEMS ENGINEERING.

Figure 2 is a flowsheet for the design, development, and testing of information systems. The process is by nature iterative and multidisciplinary. Modern information systems problems are frequently analytical rather than data-oriented, and therefore require an iterative design, as the feedback loops in Fig. 2 suggest.

Requirements modeling and prototyping. The first and most important step is the identification of user requirements, on which software requirements are based; the two are by no means the same. Requirements analysis may be based on questionnaires and surveys, interviews and observation, and simulations and games.

Questionnaire and survey-based requirements analysis includes the use of importance-rating questionnaires, time-estimate questionnaires, and Delphi and policy capture techniques. Questionnaire and survey methods assume that insight into user requirements can be obtained from indirect queries.

Interview and field observation requirements analysis assumes that requirements can be determined by asking users what they do and how, and by observing them at the tasks that the system is to perform. These methods include the use of structured and unstructured interviews, the ad hoc convening of working groups, the use of critical incident techniques, and formal job analysis.

Simulation and gaming methods include the use of scenario-driven simulations of the functions of the system. Some simulations are paper-based, others are computer-based, and still others combine human and computer-based aspects of a problem. *See* SIMULATION.

Requirements analysis extends to the users of the system and the organization it is intended to support. **Figure 3** presents a three-dimensional requirements matrix comprising tasks, users, and organizational characteristics needed by information systems engineers.

Identified requirements should be modeled for validation; Fig. 2 suggests several methods. Popular techniques offer users a working demonstration of the fully programmed system. This is referred to as prototyping and may result in evolutionary and temporary systems. When requirements can be specified with some confidence, it is often possible to develop an evolutionary prototype that can be improved with time. When requirements are poorly defined, prototypes may be discarded repeatedly until requirements are clarified. *See* MODEL THEORY; PROTOTYPE.

System sizing. Once requirements have been modeled and validated, the information system concept can be specified. It requires formulation of a user–computer relation that will determine the user–system interaction, the specification of databases, the selection of analytical methods to drive data-oriented or analytical tasks, the identification of pertinent engineering principles, and the specification of hardware. *See* COMPUTER.

An important decision is the selection of a programming language. This may be determined by existing investments, experience of personnel, and limitations of hardware. Sometimes, however, the choice of language can be based upon requirements identified early in the process. Comparative programming languages is an emerging field that helps to make the choice. The growing range of high-level languages includes FORTRAN, C, Basic, COBOL, LISP, Prolog, Pascal, and Ada. *See* PROGRAMMING LANGUAGES.

Software engineering. Several techniques are available for the specification of software requirements and programming. There are data-flow, data-structure, and object-oriented tools and techniques, and a variety of methods for the generation of knowledge structures, modules, interfaces, and control structures. *See* DATAFLOW SYSTEMS; SOFTWARE ENGINEERING.

Testing, documentation, and maintenance. The information systems engineering process calls for testing throughout design and development. Of special importance is the testing of the software to determine if it satisfies user requirements, is consistent with specifications, and works efficiently. Quality assurance, fault tolerance, and redundancy techniques are available. *See* FAULT-TOLERANT SYSTEMS.

Information systems documentation consists of system specifications, functional descriptions, user manuals, and training materials.

System testing determines if the pieces of the system work harmoniously together; multiattribute utility and traditional cost–benefit methods are available. Systems are tested to determine whether they satisfy user requirements, are cost-effective, and can be maintained.

Information systems engineers also develop maintenance plans comprising schedules, procedures, and personnel.

Implementation. The introduction of a new information system into a computer-free environment can

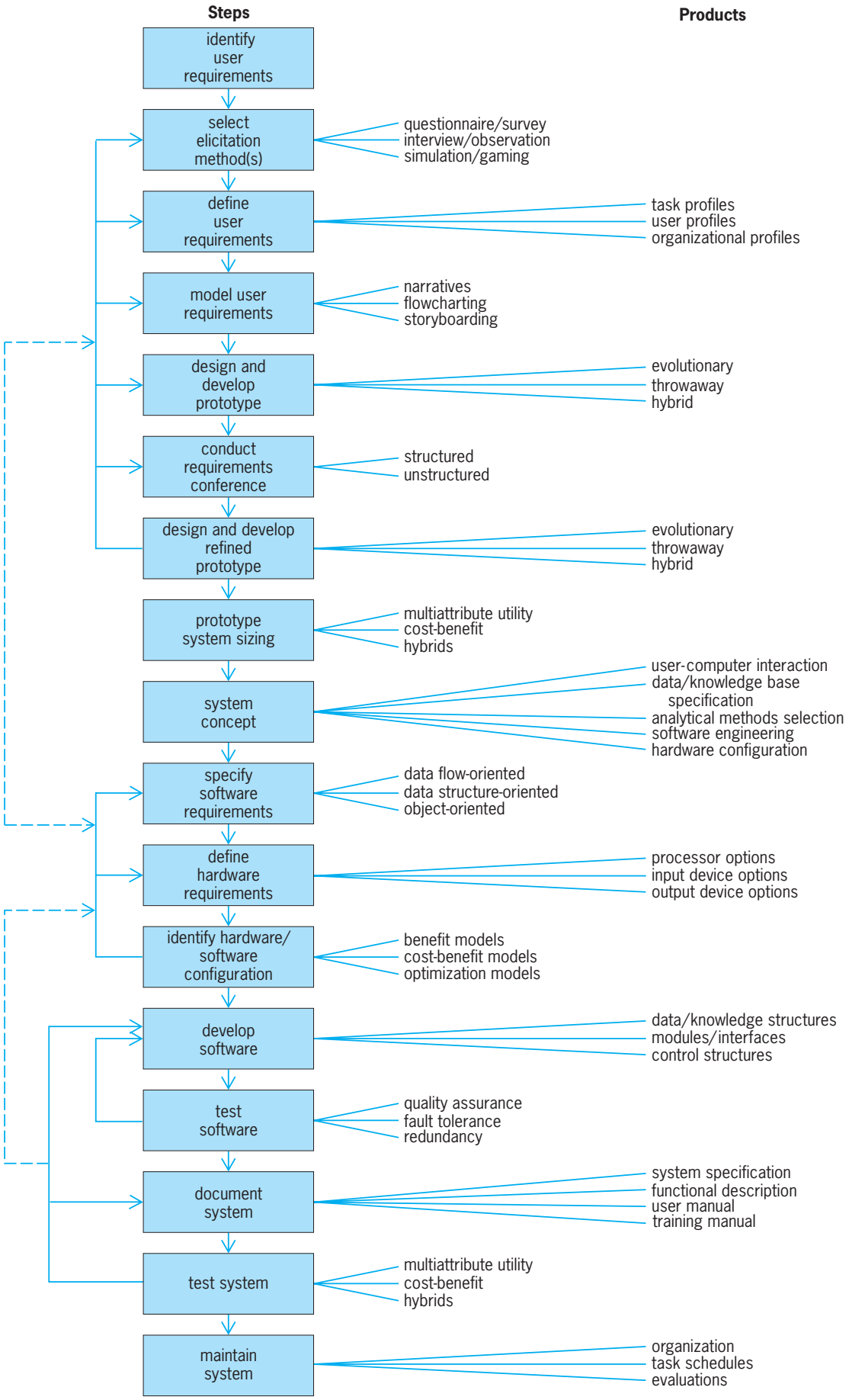

Fig. 2. Information systems engineering process, with steps, methods, and products.

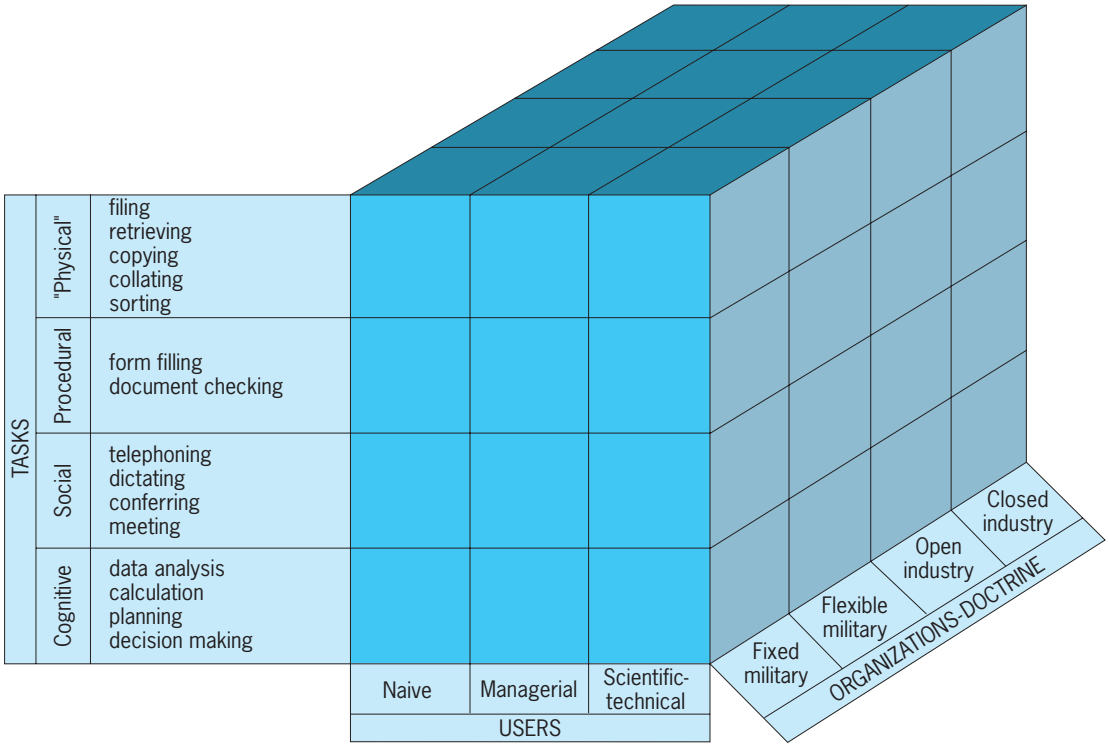

Fig. 3. Three-dimensional requirements matrix comprising information about tasks, users, and the organization that the system is intended to support.

create personnel problems. The information systems engineer adheres to procedures for the management of change developed by organizational theorists and sociologists.

Support. Information systems must be supported by an infrastructure group. This group is responsible for defining a maintenance and support process and providing the necessary tools to keep the application running and to plan for its modification, and, eventually, its replacement, triggering a whole new instance of the information systems engineering process.

Stephen J. Andriole

Bibliography. S. J. Andriole, *Managing Systems Requirements: Methods, Tools, and Cases*, McGraw-Hill, 1996; S. J. Andriole, *Storyboard Prototyping for Systems Design: A New Approach to User Requirements Validation and System Sizing*, 1989; R. S. Pressman, *Software Engineering: A Practitioner's Approach*, McGraw-Hill, 5th ed., 2000; A. P. Sage, *Systems Engineering*, Wiley-Interscience, 1992; A. Solvberg and D. C. Kung, *Information Systems Engineering: An Introduction*, Springer Verlag, 1993; K. M. Van Hee, *Information Systems Engineering: A Formal Approach*, Cambridge University Press, 1994.

Information technology

The field of engineering involving computer-based hardware and software systems, and communication systems, to enable the acquisition, representation, storage, transmission, and use of information. Successful implementation of information technology (IT) is dependent upon being able to cope with the overall architecture of systems, their interfaces with humans and organizations, and their relationships with external environments. It is also critically dependent on the ability to successfully convert information into knowledge.

Information technology is concerned with improvements in a variety of human and organizational problem-solving endeavors through the design, development, and use of technologically based systems and processes that enhance the efficiency and effectiveness of information in a variety of strategic, tactical, and operational situations. Ideally, this is accomplished through critical attention to the information needs of humans in problem-solving tasks and in the provision of technological aids, including electronic communication and computer-based systems of hardware and software and associated processes. Information technology complements and enhances traditional engineering through emphasis on the information basis for engineering.

Tools. Information technology at first was concerned with the implementation and use of new technologies to support office functions. These technologies have evolved from electric typewriters and electronic accounting systems to include very advanced technological hardware to perform functions such as electronic file processing, accounting, and word processing. Among the many potentially critical information technology–based tools are databases, e-mail, artificial intelligence systems, facsimile transmission (FAX) devices, fourth-generation programming languages, local-area networks (LANs), integrated service digital networks (ISDNs), optical

disk storage (CD-ROM) devices, personal computers, parallel processing algorithms, word processing software, computer-aided software engineering packages, and accounting software.

Management information systems and decision support. The development of support through management information systems (MIS) became possible through the use of mainframe computers. These systems have become quite powerful and are used for a variety of purposes, such as scheduling airplane flights and booking passenger seats, and registering university students in classes. As management information systems began to proliferate, it soon was recognized that, while they were very capable of providing support for organizing data and information, they did not necessarily provide much support for tasks involving human judgment and choice. These tasks range from providing support in assessing situations, such as better detection of issues or faults, to supporting diagnosis in order to enable the identification of likely causative or influencing factors. This capability was provided by linking the database management systems (DBMS) so common in the MIS era with a model base management system (MBMS) capability and a visualization and interactive presentation capability made possible through dialog generation and management systems (DGMS). The resulting systems are generally known as decision support systems (DSS). Continuing developments in microchip technology have led to electronic communications–based networking, a major facet of information technology. Many would argue that the major information technology development in recent times is the Internet. *See* DECISION SUPPORT SYSTEM; INFORMATION SYSTEMS ENGINEERING.

Activities. The knowledge and skills required in information technology come from the applied engineering sciences, especially information, computer, and systems engineering sciences, and from professional practice. Professional activities in information technology and in the acquisition of information technology systems range from requirements definition or specification, to conceptual and functional design and development of communication and computer-based systems for information support. They are concerned with such topics as architectural definition and evaluation. These activities include integration of new systems into functionally operational existing systems and maintenance of the result as user needs change over time. This human interaction with systems and processes, and the associated information processing activities, may take several diverse forms. *See* REENGINEERING; SYSTEMS ARCHITECTURE; SYSTEMS ENGINEERING.

The hardware and software of computing and communications form the basic tools for information technology. These are implemented as information technology systems through use of systems engineering processes. While information technology and information systems engineering does indeed enable better designs of systems and existing organizations, it also enables the design of fundamentally new organizations and systems such as virtual corporations. Thus, efforts in this area include not only interactivity in working with clients to satisfy present needs but also awareness of future technological, organizational, and human concerns so as to support transition over time to new information technology–based services.

Systems integration. Often, it is very difficult to cope with the plethora of new information technology–based support systems. The major reason is the lack of systems integration across the large variety of such products and services. This has led to the identification of an additional role for information technology professionals, one involving support through information systems integration engineering. An information systems integration engineer is responsible for overall systems management, including configuration management, to ensure that diverse products and services are identified and assembled into total and integrated solutions to information systems issues of large scale and scope. There are many contemporary technological issues here. There is a need for open systems architectures, or open systems environments, that provide, for example, for interoperability of applications software across a variety of heterogeneous hardware and software platforms. The key idea here is the notion of open, or public, that is, intended to produce consensus-based developments that will ameliorate difficulties associated with lack of standards and the presence of proprietary interfaces, services, and protocols. *See* SYSTEMS INTEGRATION.

Distributed collaboration. In this network age of information and knowledge, a major challenge is to capture value. Associated with this are a wide range of new organizational models. Distributed collaboration across organizations and time zones is increasingly common. The motivation for such collaboration is the desire to access sources of knowledge and skills not usually available in one place. The result of such changes has been a paradigm shift that has prompted the reengineering of organizations; the development of high-performance business teams, integrated organizations, and extended virtual enterprises; as well as the emergence of loosely structured organizations that have enhanced productivity. *See* DATA COMMUNICATIONS; INTERNET; LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

Organizational challenges. There are a number of substantial challenges associated with use of information technology with respect to enhancing the productive efforts of an individual, a group, or an organization. It has been suggested that the command and control model of leadership is poorly suited to the management of organizations where the participants are not bound by traditional incentive and reward systems. A collaborative effort has to continue to make sense and to provide value to participants for it to be sustained. Otherwise, knowledge and skills are quite portable, and the loss of knowledge workers poses a major potential risk for organizations.

Decrease in costs. The information technology revolution is driven by technology and market

considerations. Information technology costs declined in the 1990s due to the use of such technologies as broadband fiber optics, spectrum management, and data compression. The power of computers continues to increase; the cost of computing declined by a factor of 10,000 during from 1975–2000. Large central mainframe computers have been augmented, and in many cases replaced, by smaller, powerful, and more user-friendly personal computers. There has, in effect, been a merger of the computer and telecommunications industries into the information technology industry, and it now is possible to store, manipulate, process, and transmit voice, digitized data, and images at very little cost.

Benefits. As a consequence of the information technology revolution, information and knowledge have become powerful factors for socioeconomic development on a global scale. This knowledge has the potential to provide comprehensive support for enhanced production of goods and services, educational and employment opportunities for all peoples, institutions and infrastructures that enable better management in the private sector and governance in the public sector, natural resource conservation and environmental preservation, and global sustainable development. There are a number of complex issues that require attention, such as the impact of the information technology revolution on social, ethical, cultural, and family values. Most importantly, the benefits of the information technology revolution are overwhelmingly large and lead to new opportunities for world progress through the dissemination of sustainable development knowledge. Indeed, perhaps sustainable human development is not a realistic possibility at this time without major reliance on the support provided by information technology.

Andrew P. Sage

Bibliography. B. H. Boar, *Aligning Information Technology with Business Strategy*, John Wiley, New York, 1994; S. P. Bradley and R. L. Nolan, *Sense and Respond: Capturing Value in the Network Age*, Harvard Business School Press, Boston, 1998; F. Cairncross, *The Death of Distance: How the Communications Revolution Will Change Our Lives*, Harvard Business School Press, Boston, 1997; M. Dertouzos, *What Will Be: How the New World of Information Will Change Our Lives*, Harper Collins, San Francisco, 1997; P. F. Drucker, *Managing in a Time of Great Change*, Dutton, New York, 1995; R. K. Lester, *The Productive Edge: How U. S. Industries are Pointing the Way to a New Era of Economic Growth*, W. W. Norton, New York, 1998; D. Tapscott and A. Caston, *Paradigm Shift: The New Promise of Information Technology*, McGraw-Hill, New York, 1993.

Information theory

A branch of communication theory devoted to problems in coding. A unique feature of information theory is its use of a numerical measure of the amount of information gained when the contents of a message are learned. Information theory relies heavily on the mathematical science of probability. For this reason the term information theory is often applied loosely to other probabilistic studies in communication theory, such as signal detection, random noise, and prediction. *See* ELECTRICAL COMMUNICATIONS.

Information theory provides criteria for comparing different communication systems. The need for comparisons became evident during the 1940s. A large variety of systems had been invented to make better use of existing wires and radio spectrum allocations. In the 1920s the problem of comparing telegraph systems attracted H. Nyquist and R. V. L. Hartley, who provided some of the philosophy behind information theory. In 1948 C. E. Shannon published a precise general theory which formed the basis for subsequent work in information theory. *See* MODULATION; RADIO SPECTRUM ALLOCATION.

In information theory, communication systems are compared on the basis of signaling rate. Finding an appropriate definition of signaling rate was itself a problem, which Shannon solved by the use of his measure of information, to be explained later. Of special interest are optimal systems which, for a given set of communication facilities, attain a maximum signaling rate. Optimal systems provide communication-systems designers with useful absolute bounds on obtainable signaling rates. Although optimal systems often use complicated and expensive encoding equipment, they provide insight into the design of fast practical systems.

Communication systems. In designing a one-way communication system from the standpoint of information theory, three parts are considered beyond the control of the system designer: (1) the source, which generates messages at the transmitting end of the system, (2) the destination, which ultimately receives the messages, and (3) the channel, consisting of a transmission medium or device for conveying signals from the source to the destination. Constraints beyond the mere physical properties of the transmission medium influence the design. For example, in designing a radio system only a given portion of the radio-frequency spectrum may be available. The transmitter power may also be limited. If the system is just one link in a larger system which plans to use regenerative repeaters, the designer may be restricted to pulse-transmission schemes. All such conditions are considered part of the description of the channel. The source does not usually produce messages in a form acceptable as input by the channel. The transmitting end of the system contains another device, called an encoder, which prepares the source's messages for input to the channel. Similarly the receiving end of the system will contain a decoder to convert the output of the channel into a form recognizable by the destination. The encoder and decoder are the parts to be designed. In radio systems this design is essentially the choice of a modulator and a detector.

Discrete and continuous cases. A source is called discrete if its messages are sequences of elements (letters)

taken from an enumerable set of possibilities (alphabet). Thus sources producing integer data or written English are discrete. Sources which are not discrete are called continuous, for example, speech and music sources. Likewise, channels are classified as discrete or continuous according to the kinds of signals they transmit. Most transmission media (such as transmission lines and radio paths) can provide continuous channels; however, constraints (such as a restriction to use pulse techniques) on the use of these media may convert them into discrete channels.

The treatment of continuous cases is sometimes simplified by noting that a signal of finite bandwidth can be encoded into a discrete sequence of numbers. If the power spectrum of a signal $s(t)$ is confined to the band zero to W hertz (cycles per second) then Eq. (1) applies. Equation (1) reconstructs

$$s(t) = \sum_{n=-\infty}^{\infty} s\left(\frac{n}{2W}\right) \frac{\sin 2\pi W\left(t - \frac{n}{2W}\right)}{2\pi W\left(t - \frac{n}{2W}\right)} \qquad (1)$$

$s(t)$ exactly from its sample values (Nyquist samples), at discrete instants $(2W)^{-1}$ s apart. Thus, a continuous channel which transmits such signals resembles a discrete channel which transmits Nyquist samples drawn from a large finite set of signal levels and at the rate of $2W$ samples per second.

Noiseless and noisy cases. The output of a channel need not agree with its input. For example, a channel might, for secrecy purposes, contain a cryptographic device to scramble the message. Still, if the output of the channel can be computed knowing just the input message, then the channel is called noiseless. If, however, random agents make the output unpredictable even when the input is known, then the channel is called noisy. *See* COMMUNICATIONS SCRAMBLING; CRYPTOGRAPHY.

Encoding and decoding. Many encoders first break the message into a sequence of elementary blocks; next they substitute for each block a representative code, or signal, suitable for input to the channel. Such encoders are called block encoders. For example, telegraph and teletype systems both use block encoders in which the blocks are individual letters. Entire words form the blocks of some commercial cablegram systems. The operation of a block encoder may be described completely by a function or table showing, for each possible block, the code that represents it.

It is generally impossible for a decoder to reconstruct with certainty a message received via a noisy channel. Suitable encoding, however, may make the noise tolerable, as may be illustrated by a channel that transmits pulses of two kinds. It is customary to let binary digits 0 and 1 denote the two kinds of pulse. If the source has only the four letters A, B, C, D, it is possible to simply encode each single-letter block into a pair of binary digits (code I; see **table**). In that case the decoder would make a mistake every time that noise produced an error. If each single-letter block is encoded into three digits (code II), the decoder can at least recognize that a received triple of digits must contain errors if it is one of the triples (001, 010, 100, or 111) not listed in the code. Because an error in any one of the three pulses of code II always produces a triple that is not listed, code II provides single-error detection. Similarly, a five-binary-digit code (code III) can provide double-error detection, because errors in a single pulse or pair of pulses always produce a quintuple that is not listed.

Three possible binary codes for four-letter alphabet

Letter	Code I	Code II	Code III
A	0 0	0 0 0	0 0 0 0 0
B	0 1	0 1 1	0 0 1 1 1
C	1 0	1 0 1	1 1 0 0 1
D	1 1	1 1 0	1 1 1 1 0

As an alternative, code III may provide single-error correction. In this usage, the decoder picks a letter for which code III agrees with the received quintuple in as many places as possible. If only a single digit is in error, this rule chooses the correct letter.

Even when the channel is noiseless, a variety of encoding schemes exists and there is a problem of picking a good one. Of all encodings of English letters into dots and dashes, the Continental Morse encoding is nearly the fastest possible one. It achieves its speed by associating short codes with the most common letters. A noiseless binary channel (capable of transmitting two kinds of pulse 0, 1, of the same duration) provides the following example. In order to encode English text for this channel, a simple encoding might just use 27 different five-digit codes to represent word space (denoted by #), A, B, . . . , Z; say # 00000, A 00001, B 00010, C 00011, . . . , Z 11011. The word #CAB would then be encoded into 00000000110000100010. A similar encoding is used in teletype transmission; however, it places a third kind of pulse at the beginning of each code to help the decoder stay in synchronism with the encoder. The five-digit encoding can be improved by assigning four-digit codes 0000, 0001, 0010, 0011, 0100 to the five most common letters #, E, T, A, O. There are 22 quintuples of binary digits which do not begin with any of the five four-digit codes; these may be assigned as codes to the 22 remaining letters. About half the letters of English text are #, E, T, A, or O; therefore the new encoding uses an average of only 4.5 digits per letter of message. *See* TELETYPEWRITER.

More generally, if an alphabet is encoded in single-letter blocks, using $L(i)$ digits for the ith letter, the average number of digits used per letter is shown in Eq. (2), where $p(i)$ is the probability of the ith letter.

$$L = p(1)L(1) + p(2)L(2) + p(3)L(3) + \cdots \qquad (2)$$

An optimal encoding scheme will minimize L. However, the encoded messages must be decipherable, and this condition puts constraints on the $L(i)$. The code lengths of decipherable encodings must satisfy

the relationships shown in inequality (3). The real

$$2^{-L(1)} + 2^{-L(2)} + 2^{-L(3)} + \cdots \leq 1 \qquad (3)$$

numbers $L(1), L(2), \ldots,$ which minimize L subject to inequality (3) are $L(i) = -\log_2 p(i)$ and the corresponding minimum L is shown in Eq. (4), which

$$H = -\sum_i p(i) \log_2 p(i) \qquad (4)$$

provides a value of H equal to a number of digits per letter.

The $L(i)$ must be integers and $-\log_2 p(i)$ generally are not integers; for this reason there may be no encoding which provides $L = H$. However, Shannon showed that it is always possible to assign codes to letters in such a way that $L \leq H + 1$. A procedure for constructing an encoding which actually minimizes L has been given by D. A. Huffman. For (27-letter) English text $H = 4.08$ digits per letter, as compared with the actual minimum 4.12 digits per letter obtained by Huffman's procedure.

By encoding in blocks of more than one letter, the average number of digits used per letter may be reduced further. If messages are constructed by picking letters independently with the probabilities $p(1)$, $p(2), \ldots,$ then H is found to be the minimum of the average numbers of digits per letter used to encode these messages using longer blocks. *See* DATA COMPRESSION.

Information content of message. The information contained in a message unit is defined in terms of the average number of digits required to encode it. Accordingly the information associated with a single letter produced by a discrete source is defined to be the number H. Some other properties of H help to justify using it to measure information. If one of the $p(i)$ equals unity, only one letter appears in the messages. Then nothing new is learned by seeing a letter and, indeed, $H = 0$. Second, of all possible ways of assigning probabilities $p(i)$ to an N-letter alphabet, the one which maximizes H is $p(1) = p(2) = \cdots = 1/N$. This situation is the one in which the unknown letter seems most uncertain; therefore it does seem correct that learning such a letter provides the most information. The corresponding maximum value of H is $\log_2 N$. This result seems reasonable by the following argument. When two independent letters are learned, the information obtained should be $2H = 2 \log_2 N$. However, such pairs of letters may be considered to be the letters of a larger alphabet of N^2 equally likely pairs. The information associated with one of these new letters is $\log_2 N^2 = 2 \log_2 N$. Although H given by Eq. (4) is dimensionless, it is given units called bits (a contraction of binary digits). Occasionally the information is expressed in digits of other kinds (such as ternary or decimal). Then bases other than 2 are used for the logarithm in Eq. (4).

The majority of message sources do not merely pick successive letters independently. For example in English, H is the most likely letter to follow T but is otherwise not common. The source is imagined to be a random process in which the letter probabilities change, depending on what the past of the message has been. Statistical correlations between different parts of the message may be exploited by encoding longer blocks. The average number of digits per letter may thereby be reduced below the single-letter information H given by Eq. (4). For example, by encoding English words instead of single letters, 2.1 digits/letter suffice. Encoding longer and longer blocks, the number of digits needed per letter approaches a limiting minimum value. This limit is called the entropy of the source and is interpreted as the rate, in bits per letter, at which the source generates information. If the source produces letters at some fixed average rate, n letters/s, the entropy may also be converted into a rate in bits per second by multiplying by n. The entropy may be computed from tables giving the probabilities of blocks of N letters (N-grams). If in Eq. (4) the summation index i is extended over all N-grams, then the number H represents the information in N consecutive letters. As $N \to \infty$, H/N approaches the entropy of the source. The entropy of English has been estimated by Shannon to be about 1 bit/letter. However, an encoder might have to encode 100-grams to achieve a reduction to near 1 digit/letter. Comparing English with a source that produces 27 equally likely letters independently (hence has entropy $\log_2 27 = 4.8$ bits/letter), this result is often restated: English is 80% redundant. *See* ENTROPY.

In computer applications, coding can reduce the redundancy of data for storage, saving space in memory. Large white areas make black-and-white diagrams very redundant. To store diagrams or send them by facsimile, coding white runs instead of individual black-and-white samples will remove much redundancy. *See* FACSIMILE.

Universal codes. Even when a probabilistic description of the source is unavailable, coding can remove redundancy. In a simple illustration, the encoder begins by parsing a binary message into blocks, now of varying lengths. Each new block B is made as long as possible, but B must be a block B', seen earlier, followed by one extra digit x. A typical parsed message is 0,01,00,1,000, . . . , with commas separating the blocks. When transmitting the kth block B, the transmitter first sends the last digit x of B. It next sends a string of digits, of length about $\log_2 k$, to identify the earlier block B'. Redundancy makes block sizes tend to grow faster than $\log_2 k$; then the encoding reduces message length. A. Lempel and J. Ziv have shown that codes like this one compress data about as well as codes that use given source statistics. In effect, the parsing gathers statistics to help encode future blocks.

Capacity. The notion of entropy is more widely applicable than might appear from the discussion of the binary channel. Any discrete noiseless channel may be given a number C, which is called the capacity. C is defined as the maximum rate (bits per second) of all sources that may be connected directly to the channel. Shannon proved that any given source (which perhaps cannot be connected directly to the

channel) of entropy H bits/letter can be encoded for the channel and run at rates arbitrarily close to C/H letters/s.

By using repetition, error-correcting codes, or similar techniques, the reliability of transmission over a noisy channel can be increased at the expense of slowing down the source. It might be expected that the source rate must be slowed to 0 bits/s as the transmission is required to be increasingly error-free. On the contrary, Shannon proved that even a noisy channel has a capacity C. Suppose that errors in at most a fraction ϵ of the letters of the message can be tolerated ($\epsilon > 0$). Suppose also that a given source, of entropy H bits/letter, must be operated at the rate of at least $(C/H) - \delta$ letters/s ($\delta > 0$). No matter how small ϵ and δ are chosen, an encoder can be found which satisfies these requirements.

For example, the symmetric binary channel has binary input and output letters; noise changes a fraction p of the 0's to 1 and a fraction p of the 1's to 0 and treats successive digits independently. The capacity of this channel is shown by Eq. (5), where m is

$$C = m[1 + p \log_2 p + (1 - p) \log_2(1 - p)] \qquad (5)$$

the number of digits per second which the channel transmits.

Shannon's formula, shown by Eq. (6), gives the capacity C of a band-limited continuous channel. The channel consists of a frequency band W Hz wide, which contains a gaussian noise of power N. The noise has a flat spectrum over the band and is added to the signal by the channel. The channel also contains a restriction that the average signal power may not exceed S.

$$C = W \log_2 \left(1 + \frac{S}{N}\right) \qquad (6)$$

Equation (6) illustrates an exchange relationship between bandwidth W and signal-to-noise ratio S/N. By suitable encoding a signaling system can use a smaller bandwidth, provided that the signal power is also raised enough to keep C fixed. *See* BANDWIDTH REQUIREMENTS (COMMUNICATIONS).

Typical capacity values are 20,000 bits/s for a telephone speech circuit and 50,000,000 bits/s for a broadcast television circuit. Speech and television are very redundant and would use channels of much lower capacity if the necessary encodings were inexpensive. For example, the vocoder can send speech, only slightly distorted, over a 2000-bits/s channel. Successive lines or frames in television tend to look alike. This resemblance suggests a high redundancy; however, to exploit it the encoder may have to encode in very long blocks.

Not all of the waste in channel capacity can be attributed to source redundancies. Even with an irredundant source, such as a source producing random digits, some channel capacity will be wasted. The simplest encoding schemes provide reliable transmission only at a rate equal to the capacity of a channel with roughly 8 dB smaller signal power (the 8-dB figure is merely typical and really depends on the reliability requirements). Again, more efficient encoding to combat noise generally requires larger-sized blocks. This is to be expected. The signal is separated from the noise on the basis of differences between the signal's statistical properties and those of noise. The block size must be large enough to supply the decoder with enough data to draw statistically significant conclusions. *See* ELECTRICAL NOISE.

Algebraic codes. Practical codes must use simple encoding and decoding equipment. Error-correcting codes for binary channels have been designed to use small digital logic circuits. These are called algebraic codes, linear codes, or group codes because they are constructed by algebraic techniques involving linear vector spaces or groups.

For example, each of the binary codes I, II, and III discussed above (see table) contains four code words which may be regarded as vectors $C = (c_1, c_2, \ldots, c_n)$ of binary digits c_i. The sum $C + C'$ of two vectors may be defined to be the vector $(c_1 + c'_n, \ldots, c_n + c'_n)$ in which coordinates of C and C' are added modulo 2. Codes I, II, and III each have the property that the vector sum of any two code words is also a code word. Because of that, these codes are linear vector spaces and groups under vector addition. Their code words also belong to the n-dimensional space consisting of all 2^n vectors of n binary coordinates. Codes II and III, with $n = 3$ and 5, do not contain all 2^n vectors; they are only two-dimensional linear subspaces of the larger space. Consequently, in Codes II and III, the coordinates c_i must satisfy certain linear homogeneous equations. Code II satisfies $c_1 + c_2 + c_3 = 0$. Code III satisfies $c_3 + c_4 = 0$, $c_2 + c_3 + c_5 = 0$, $c_1 + c_2 = 0$, and other equations linearly dependent on these three. The sums in such equations are performed modulo 2; for this reason the equations are called parity check equations. In general, any r linearly independent parity check equations in $c_1, \ldots, c_n$ determine a linear subspace of dimension $k = n - r$. The 2^k vectors in this subspace are the code words of a linear code. *See* GROUP THEORY; LINEAR ALGEBRA; NUMBER THEORY.

The r parity checks may be transformed into a form which simplifies the encoding. This transformation consists of solving the original parity check equations for some r of the coordinates c_i as expressions in which only the remaining $n - r$ coordinates appear as independent variables. For example, the three parity check equations given for Code III are already in solved form with c_1, c_4, c_5 expressed in terms of c_2 and c_3. The $k = n - r$ independent variables are called message digits because the 2^k values of these coordinates may be used to represent the letters of the message alphabet. The r dependent coordinates, called check digits, are then easily computed by circuits which perform additions modulo 2. *See* LINEAR SYSTEMS OF EQUATIONS.

At the receiver the decoder can also do additions modulo 2 to test if the received digits still satisfy the parity check equations. The set of parity check equations that fail is called the syndrome because it contains the data that the decoder needs to diagnose the errors. The syndrome depends only on the error

locations, not on which code word was sent. In general, a code can be used to correct e errors if each pair of distinct code words differ in at least $2e + 1$ of the n coordinates. For a linear code, that is equivalent to requiring the smallest number d of "ones" among the coordinates of any code word [excepting the zero word $(0, 0, \ldots, 0)$] to be $2e + 1$ or more. Under these conditions each pattern of $0, 1, \ldots, e - 1$, or e errors produces a distinct syndrome; then the decoder can compute the error locations from the syndrome. This computation may offer some difficulty. But at least it involves only r binary variables, representing the syndrome, instead of all n coordinates.

Hamming codes. The r parity check equations may be written concisely as binary matrix equation (7). Here C_T is a column vector, the transpose of

$$HC^T = 0 \tag{7}$$

$(c_1, \ldots, c_n)$. H is the so-called parity check matrix, having n columns and r rows. A Hamming single-error correcting code is obtained when the columns of H are all $n = 2^r - 1$ distinct columns of r binary digits, excluding the column of all zeros. If a single error occurs, say in coordinate c_i, then the decoder uses the syndrome to identify c_i as the unique coordinate that appears in just those parity check equations that fail. *See* MATRIX THEORY.

Shift register codes. A linear shift register sequence is a periodic infinite binary sequence $\ldots, c_0, c_1, c_2, \ldots$ satisfying a recurrence equation expressing c_j as a modulo 2 sum of some of the b earlier digits $c_{j-b}, \ldots, c_{j-1}$. A recurrence with two terms would be an equation $c_j = c_{j-a} + c_{j-b}$, with a equal to some integer $1, 2, \ldots,$ or $b - 1$. The digits of a shift register sequence can be computed, one at a time, by very simple equipment. It consists of a feedback loop, containing a shift register to store $c_{j-b}, \ldots, c_{j-1}$ and a logic circuit performing modulo 2 additions. This equipment may be used to implement a linear code. First, message digits $c_1, \ldots, c_b$ are stored in the register and transmitted. Thereafter the equipment computes and transmits successively the $n - b$ check digits c_j obtained from the recurrence equation with $j = b + 1, \ldots, n$. By choosing a suitable recurrence equation, the period of the shift register sequence can be made as large as $2^b - 1$. Then, with n equal to the period $2^b - 1$, the code consists of the zero code word $(0, 0, \ldots, 0)$ and $2^b - 1$ other code words which differ from each other only by cyclic permutations of their coordinates. These latter words all contain $d = 2^b - 1$ "ones" and so the code can correct $e = 2^{b-2} - 1$ errors. *See* SWITCHING CIRCUIT.

Intermediate codes. The Hamming codes and maximal period shift register codes are opposite extremes, correcting either one or many errors and having code words consisting either mostly of message digits or mostly of check digits. Many intermediate codes have been invented. One of them requires $n + 1$ to be a power of 2; say $n + 1 = 2^q$. It then uses at most qe check digits to correct e errors.

Perfect codes. Although each pattern of $0, 1, \ldots, e$ errors produces a distinct syndrome, there may be extra syndromes which occur only after more than e errors. In order to keep the number of check digits small, extra syndromes must be avoided. A code is called perfect if all 2^r syndromes can result from patterns of $0, 1, \ldots, e - 1$, or e errors. Hamming codes are all perfect. M. J. E. Golay found another perfect binary code having $n = 23$, $r = 11$ check digits, and correcting $e = 3$ errors.

Orthogonal parity codes. Orthogonal parity codes are codes with especially simple decoding circuits which take a kind of majority vote. Suppose the parity check equations can be used to derive $2e + 1$ linear equations in which one digit, say c_1, is expressed with each of the remaining digits $c_2, \ldots, c_n$ appearing in at most one equation. If at most e errors occur, then the received digits satisfy a majority of the $2e + 1$ equations if and only if c_1 was received correctly. For example, the recurrence $c_j = c_{j-2} + c_{j-3}$ generates a maximal period shift register code with $n = 7$, $r = 4$, $e = 1$. With $j = 1$, 3, and 4 in the recurrence equation, three of the parity check equations, $c_1 = c_5 + c_6$, $c_1 = c_3 + c_7$, and $c_1 = c_2 + c_4$, are obtained, after using the fact that the shift register sequence has period 7. These three equations are already in the form required for decoding c_1 by majority vote. Similar equations, obtained by permuting $c_1, \ldots, c_7$ cyclically, apply for $c_2, \ldots, c_7$. Then the decoder can be organized so that most of the equipment used to decode c_1 can be used again in decoding $c_2, \ldots, c_7$. *See* COMBINATORIAL THEORY.

Wide-band signaling. A band-limited continuous channel has capacity C, given by Eq. (6), very near $1.44\ WS/N$ if the signal-to-noise ratio S/N is small. Then high capacity is obtained even with low signal power S if the bandwidth W is wide enough. Wide bands are used in satellite systems and in military applications where a weak signal must hide in background noise or penetrate jamming. Or, many users can share the same wide radio channel effectively even though their signals spread over the entire band and interfere with one another.

One method of wide-band signaling modulates a kind of noise carrier, a stream of short pulses of random sign, by changing signs in accordance with the digits of a binary message. The receiver has a synchronous detector that must regenerate the carrier locally in order to demodulate the signal. A carrier kept secret from other listeners can serve as the running cryptographic key of a stream cipher. Noiselike carriers may take the form of long periodic sequences (for example, Gold codes) generated by linear recurrences in the manner of shift-register codes. *See* AMPLITUDE-MODULATION DETECTOR; SPREAD SPECTRUM COMMUNICATION.

Nonblock codes. Many useful codes do not send messages block by block. Simple convolutional codes intersperse check digits among message digits in a regular pattern. An example, designed to correct errors that cluster together in isolated bursts, transmits the binary message $\ldots, x_1, x_2, x_3, \ldots$ as $\ldots, c_1, x_1, c_2, x_2, c_3, x_3, \ldots$, with $c_i = x_{i-3} + x_{i-6}$ (modulo 2).

It can correct any burst of length 6 or less that lies more than 19 digits away from other errors.

Convolutional codes are commonly decoded by sequential or dynamic programming methods such as the Viterbi algorithm. Using the given probabilistic description of the noise process, the decoder calculates the posterior probabilities of messages that might possibly have produced the data received up to the present time. The decoder then chooses the most likely message as its decoded message. The calculation is done digit by digit, using the probabilities computed for the first k received digits to simplify the calculation for $k + 1$ digits. Although the number of possible messages grows very quickly with k, one can eliminate highly unlikely messages to keep the calculation reasonably short at each step. *See* ESTIMATION THEORY.

For signaling over radio channels with additive gaussian noise, trellis codes, which are particular kinds of nonblock codes, often have an advantage over block codes of equal complexity. Expressed in terms of signal-to-noise ratio, the advantage can be 3 dB or more even for simple codes. Edgar N. Gilbert

Bibliography. E. R. Berlekamp, *Algebraic Coding Theory*, rev. ed., 1984; R. W. Hamming, *Coding and Information Theory*, 2d ed., 1986; R. W. Lucky, *Silicon Dreams: Information, Man, and Machine*, 1991; M. Schwartz, *Information Transmission*, *Modulation*, *and Noise*, 4th ed., 1990; N. J. A. Sloane and F. J. MacWilliams, *Theory of Error-Correcting Codes*, 9th reprint, 1998; C. E. Shannon, *Collected Papers*, 1993.

Information theory (biology)

The application to biological systems of the theory of encoding effects on signal transmission and communications efficiency. In everyday language, information is associated with knowledge, meaning (semantics), and the influence of information on behavior (pragmatics). Information theory considers information to be quantitative and capable of being expressed in binary digits, or bits. The measure of information is based on its structure or its representation, taking into account the statistical properties of its structure. Mathematical information theory was created to describe the transmission of information in telecommunications, but it has been widely applied also in mathematics, physics, and biology.

Representation of information. A physical quantity that carries information is called a signal. A signal may be represented in space and time. Examples include speech and radio waves, time-dependent voltage, a still or moving picture, and the pulse sequence within a nerve fiber. In order to calculate the amount of information it contains, a signal must be quantized in a two-step process for symbolic presentation. First, the time or space coordinate has to be discrete, appearing only at precise, usually equidistant points on the coordinates. For signals with limited variations (limited bandwidth), this is possible without a loss of generality because of the sampling theorem. Second, the amplitude of the signal must be quantized, or expressed in (usually equal) steps. The size of the steps, which is arbitrary, determines the accuracy of the representation, but it is usually chosen at the limit of resolution to detect a certain variation of the amplitude. The signal can then be represented by a finite number of discrete points in time or space and by amplitude levels at each point.

An information source is a device that produces signals that carry information. In the symbolic representation of information theory, these signals are discrete in time or space and quantized in amplitude, and the information source is usually described by a stochastic process. Written language is a typical example of the signal sequence produced by a stationary information source.

Coding theorem of information. The coding theorem of information shows that it is possible to code the output of an information source with a number of bits equal to its entropy, or uncertainty with respect to the signals of the source. Symbols or symbol sequences with a high probability are represented by a short code word, and those with low probability are represented by a longer code word, much like the Morse alphabet.

The coding theorem of information restricts the complexity needed to represent the information produced by an information source, and that restriction is important in telecommunications. Source coding makes use of the redundancy of the information source to reduce the amount of data to be transmitted. In the case of speech transmission with telephone quality, a reduction factor of about 1:10 has been achieved. For a television or picture phone signal, reduction factors of 1:100 or even 1:1000 are hypothesized. *See* ENTROPY.

Information channel. The information channel is the connection of an information source with a receiver. The information source is connected by way of a coder to the channel. The coder transforms the signals of the source into transmissible signals, eliminates redundancy, and reduces the effect of noise introduced by the channel (channel coding). Suitable modulation may be included, and when performed simultaneously with channel coding, is called codulation. The transmission channel may have frequency dispersion (channel with memory). The channel typically introduces noise, which in the simplest case is additive to the signal and has a gaussian amplitude distribution, but the scheme also holds for more general cases, such as multiplicative noise and Poisson distributions. The decoder delivers the decoded signal to the receiver or destination. Entropy at the input and entropy at the output of the channel can then be calculated: if only noise without dispersion is present, perfect transmission without errors or loss of information is possible.

Coding theorem of transinformation. The coding theorem of transinformation states that in the limiting case of very long sequences a transmission with vanishing error probability is possible. The ideal case may be approximated by coding very long signal

sequences, which also implies a very long delay between transmitter and receiver. However, error-free transmission is possible even in noisy surroundings. Error-correcting codes can approach the ideal error-free case, and many optimized binary coding schemes exist.

The decoder, which can be considered equivalent to a pattern recognition device, may have a very high complexity, but it can be reduced by algorithms. Its task is to classify or detect the desired signal in a set of possible received signals. The human receiver, or decoder, has the same task of pattern recognition, which is performed by the nervous system. *See* CYBERNETICS.

Human information channel. Information theory can be applied to human information processing (**Fig. 1**). The number of receptors for each of various sensory organs has been determined, and the numerical values of maximum information flow for some of the sensory organs has been estimated. The data given in Fig. 1 show that television and telephone channels are well adapted to human receptors. *See* INFORMATION PROCESSING (PSYCHOLOGY).

Information flow for various conscious human activities can also be estimated. Counting requires 3 bits/s; reading, 40 bits/s; hearing, 40 bits/s; and piano playing, 20 bits/s. The maximum human channel capacity for conscious output is 50 bits/s, which indicates an apparent reduction of information in the central nervous system. Such reduction is necessary because not all details of the external world need to penetrate the consciousness. A good deal of this information is superfluous or redundant and therefore can be eliminated; otherwise memory would be unable to store the entire flow of sensory information. If 50 bits/s is multiplied by a life expectancy of 70 years, the information received totals about 10^{10} bits. The storage capacity of human memory is estimated to be of that order of magnitude. *See* MEMORY.

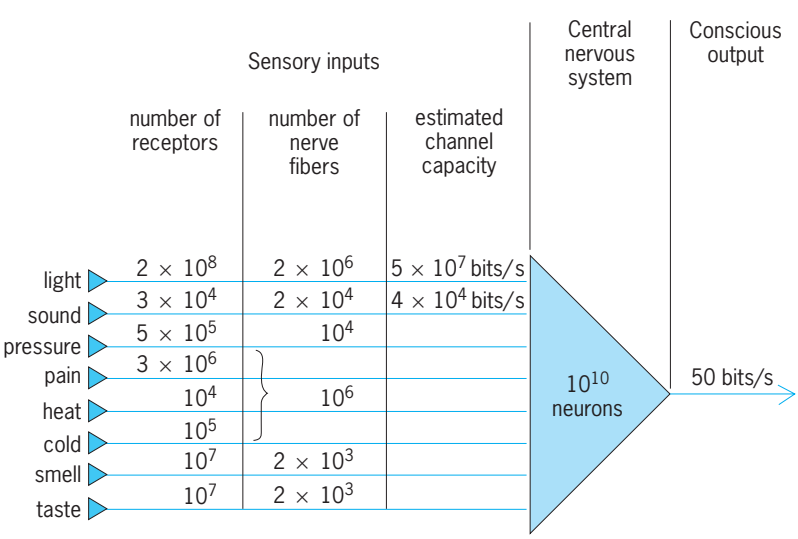

Fig. 1. Human information processing. The estimates of channel capacity are based on the number of rods in the retina of the eye and the number of hair cells in the inner ear's organ of Corti.

Controlled source and bidirectional communication. In human communication, a unidirected information channel is a monologue. The usual situation, however, is a dialogue, which involves bidirectional communication. In bidirectional communication theory, the independent information source is replaced by a controlled or dependent information source. In a dialogue, the transmitter and the receiver are both described by controlled information sources. The behavior of the living being may be described statistically by a dependent information source, in which controlling input represents the sensory signals and the output is the behavior. The mathematical description of a controlled information source is given by the conditional probability $p(x|x_n y_n)$ where x is the produced symbol, x_n sequence of previously produced symbols, and y_n the sequence of previously received symbols. The expression $p(x|x_n y_n)$ also represents generator probability independent of the combination of the transmitted and received sequences $x_n y_n$. If behavior changes because of a learning or adaptive process, the probability changes.

For the dependent information source, two entropy quantities may be defined. The first is the usual entropy, given by Eq. (1). This takes into account

$$H(x) = \lim_{n\to\infty} E[-\log p(x|x_n)] \tag{1}$$

only the produced signal sequence x_n. If both the produced and the received signal sequences x_n and y_n are considered, the free entropy is determined by Eq. (2).

$$F(x) = \lim_{n\to\infty} E[-\log p(x|x_n y_n)] \tag{2}$$

Clearly $F(x) \le H(x)$ because it contains the knowledge of the received signals. The difference between the entropies, which is known as directed transinformation $y \to x$, is given by Eq. (3), where $T(x|y)$ represents the influence of the controlling (or received) signals on the information production of the source, and is thus the transinformation from the y sequence to the x sequence. The whole entropy $H(x)$ of the source thus consists of a free part $F(x)$ and a dependent part $T(x|y)$ transmitted by the y sequence.

$$T(x|y) = H(x) - F(x) \tag{3}$$

A bidirectional communication takes place when two dependent information sources are coupled. The entropy quantities for that case are shown in the **table**. The sum of the two directed transinformation quantities equals the transinformation calculated by the original information theory. If one of the two transinformations, $T(x|y)$ or $T(y|x)$, is zero, the problem is reduced to the special case of the unidirectional information channel. The two equations for the directed transinformation lead to the flow diagram of the communication shown in **Fig. 2**. This diagram also shows that only part of the produced entropy $H(x)$, namely $T(y|x)$, reaches the counterpart. Thus the coupling may be small because of noise or other limitations of the communication channels.

Quantities associated with entropy of bidirectional communication

Quantity	Equation
Entropy of x	$H(x) = \lim_{n\to\infty} E[-\log p(x/x_n)]$
Entropy of y	$H(y) = \lim_{n\to\infty} E[-\log p(y/y_n)]$
Free entropy of x	$F(y) = \lim E[-\log p(y/x_n y_n)]$
Free entropy of y	$F(y) = \lim_{n\to\infty} E[-\log p(y/x_n y_n)]$
Directed transinformation $y \to x$	$T(x/y) = H(x) - F(x)$
Directed transinformation $x \to y$	$T(y/x) = H(y) - F(y)$
Total transinformation (coincidence, synentropy) $x \leftrightarrow y$	$T = T(x/y) + T(y/x)$

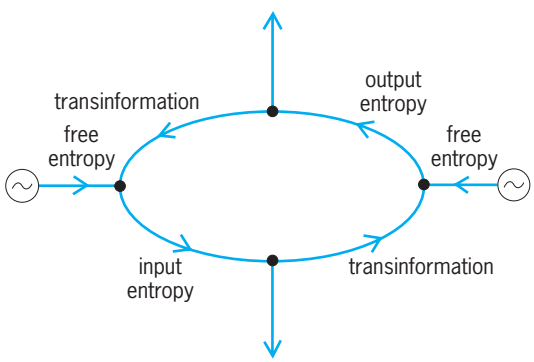

Fig. 2. Flow diagram of communication.

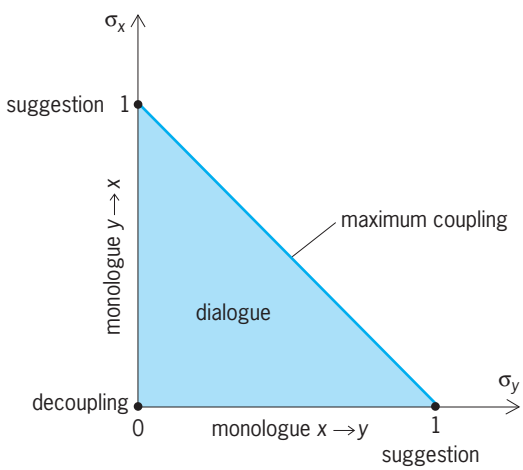

Fig. 3. Coupling diagram of communication.

Maximal coupling, however, is limited. If the coupling coefficients G_x and G_y are defined by Eqs. (4) and (5), then the limiting condition $\sigma_x + \sigma_n \leq 1$ holds.

$$G_x = \frac{T(x|y)}{H(x)} \tag{4}$$

$$G_y = \frac{T(y|x)}{H(y)} \tag{5}$$

According to that inequality, all possible coupling states of a communication are shown in the coupling diagram of **Fig. 3**. The maximum coupling occurs when $\sigma_x + \sigma_y = 1$. Complete (quasi-deterministic) coupling can obviously occur in only one direction. With respect to a human communication, it is referred to as suggestion. Bidirectional communication theory has been applied to the group behavior of monkeys, but it may also be used to describe the theoretical aspects of user–machine interaction. *See* HUMAN-MACHINE SYSTEMS; INFORMATION THEORY.

Hans Marko

Bibliography. R. E. Blahut, *Communications, Probability, Information Theory, Coding*, 1987; C. Cherry, *On Human Communication: A Review, a Survey and a Criticism*, 3d ed., 1978; R. G. Gallager, *Information Theory and Reliable Communication*, 1968; D. L. Kincaid (ed.), *Communication Theory: Eastern and Western Perspectives*, 1987; S. Oyama, *Ontogeny of Information: Developmental Systems and Evolution*, 2d ed., 2000; J. R. Sampson (ed.), *Biological Information Processing: Current Theory and Computer Simulation*, 1989; C. Sybsema, *Biophysics: An Introduction*, 1989.

Infrared astronomy

The field of astronomical observations specializing in detecting photons from the infrared portion of the electromagnetic spectrum. Detection technology has traditionally defined the domains of the electromagnetic spectrum. The infrared portion of the spectrum spans the range from the red limit of human vision (approximately 0.7 micrometer) to the shortest wavelengths accessible to heterodyne radio receivers (several hundred micrometers). *See* ELECTROMAGNETIC RADIATION; INFRARED RADIATION.

Differences from optical astronomy. Astronomers observing the universe with infrared light encounter a number of fundamental differences relative to those observing with visible light:

1. Infrared observations are more sensitive to cooler objects than visible-wavelength observations. Blackbodies at a temperature cooler than 2000 kelvins (3100°F) radiate virtually all of their light in the infrared part of the spectrum (**Fig. 1**). Infrared observations are particularly well suited to detect both forming stars and evolved stars (that is, stars in the final stages of their lives) since both classes of objects are cool. Since starlight often heats nearby dust grains to temperatures of tens or hundreds of degrees, reprocessing the visible starlight into exclusively infrared radiation, warm dust is also a common target for infrared observations. *See* HEAT RADIATION.

2. Interstellar dust is substantially more transparent at infrared wavelengths than at visible wavelengths. The dust grains constituting the interstellar medium tend to be smaller than 1 μm in diameter and are mainly composed of silicate and carbon grains. The theory of the scattering of electromagnetic radiation by dielectric particles dictates that such particles efficiently scatter wavelengths of light which

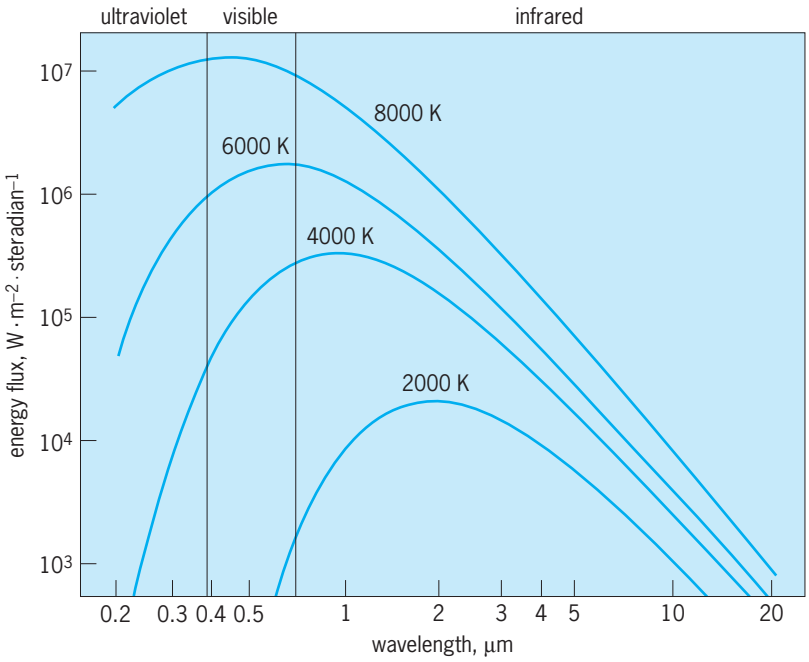

Fig. 1. Plots of the energy flux emitted by a blackbody as a function of wavelength for several temperatures. Hotter objects are substantially brighter and emit most of their energy at shorter wavelengths. Objects cooler than a few thousand kelvins emit predominantly infrared radiation. (The vertical axis is the product of the energy flux per unit wavelength and the wavelength itself.)

are equal to or smaller than the particle's size. Since infrared light has a wavelength substantially larger than the size of interstellar grains, the grains are ineffective at extinguishing the infrared light of stars and galaxies which lie behind the dust. Infrared observations thus enable astronomers to view distant objects through the obscuring dust that permeates the Milky Way Galaxy. Forming and evolved stars often reside in dense clouds of interstellar dust grains and can be observed only at infrared and even longer radio wavelengths. *See* INTERSTELLAR EXTINCTION; INTERSTELLAR MATTER; SCATTERING OF ELECTROMAGNETIC RADIATION.

3. The quantized energies of molecular rotational and vibrational transitions, which give rise to molecular spectral lines, fall largely in the infrared part of the spectrum, as do many hyperfine lines of individual atoms. Infrared spectral lines of water, molecular hydrogen, and carbon monoxide probe dense interstellar environments and cool stellar atmospheres. Very small interstellar grains can be considered as immense molecules. The bending and stretching modes of (CH) and (SiO) on the surface of these grains produce broad infrared spectral features diagnostic of the grains and, in cold environments such as dark interstellar clouds, of their icy mantles. *See* INFRARED SPECTROSCOPY.

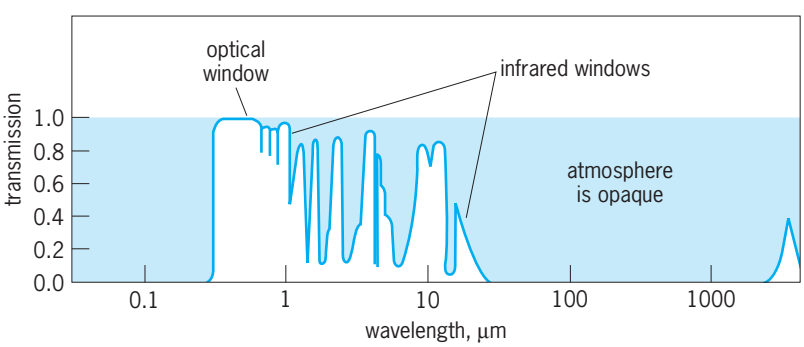

Fig. 2. Atmospheric transmission as a function of wavelength throughout the visible and infrared parts of the spectrum, showing atmospheric windows.

In addition to these astrophysical differences, the technology and practice of infrared astronomy differs from visible-wavelength astronomy in fundamental ways:

1. The Earth's atmosphere is opaque to infrared radiation through a substantial fraction of the spectrum (**Fig. 2**). This opacity arises largely from water in the Earth's lower atmosphere. Water's infrared spectral absorption lines blanket the infrared spectrum, particularly when blurred together by the pressure broadening which occurs in the dense terrestrial atmosphere. At the shorter infrared wavelengths, the infrared spectrum has several windows of transmission which permit ground-based telescopes to observe celestial sources. At the longer infrared wavelengths, the atmosphere is nearly opaque, and astronomers must resort to high-flying aircraft or orbiting satellites in order to evade the atmospheric opacity. *See* SATELLITE (ASTRONOMY).

2. The energy of chemical bonds and the work function for the liberation of electrons from a metal via the photoelectric effect is of order a few electronvolts. Planck's law, $E = h\nu = hc/\lambda$, which relates the energy E of a photon to its frequency ν or its wavelength λ via Planck's constant ($h = 6.63 \times 10^{-34}$ joule-second) and the speed of light ($c = 3 \times 10^8$ m/s), dictates that only photons with wavelength shorter than about 1 μm can induce chemical reactions or liberate free electrons. Thus physics limits photography and photon-counting photomultiplier tubes to operation mainly in the visible-wavelength domain. Infrared detection technology relies largely on either mimicking the photoelectric effect inside a crystalline semiconductor material or monitoring the temperature change of a semiconductor under the influence of infrared radiation. *See* PHOTOEMISSION.

3. Objects of room temperature (300 K or 80°F) emit radiation throughout most of the infrared spectrum. This glow interferes with the detection of faint astronomical sources, limiting the sensitivity of observations. Rapid comparison and differencing of the signal from the source direction versus the adjacent sky can mitigate some of these effects. Cooling of the entire telescope—impractical on the ground but possible in the vacuum of space—can substantially reduce the thermal glow of the telescope optics and result in unprecedented sensitivity to faint astronomical sources.

Infrared technology. Detectors of infrared radiation divide into two classes: bolometers and photovoltaic or photoconductive devices. Bolometer detectors have temperature-sensitive electrical conductivity. Incident radiation warms the detector, and

the resulting subtle change in the electrical resistance of the detector is measured. Infrared photodetectors are crystalline semiconductors in which some electrons in the crystal lattice lie only a short distance in energy away from becoming unbound and behaving like metallic conducting electrons. Infrared light with energy in excess of the binding energy creates free charge carriers, either changing the bulk conductivity of the device (photoconductors) or charging or discharging a semiconductor junction (photovoltaics). The detector's response to infrared radiation depends on the chemistry-dependent solid-state physics of the detector material. In a mixture of indium and antimony (InSb), for example, the binding energy of the electrons corresponds to an infrared wavelength of 5.5 μm. This material is responsive to photons of this wavelength or shorter. Different semiconductor materials have different long-wavelength cutoffs. By doping the semiconductor material with impurities, it is possible to produce even more weakly bound electrons, yielding detectors with sensitivity to wavelengths as long as 100 μm or more. *See* BOLOMETER; PHOTOCONDUCTIVITY; PHOTOVOLTAIC EFFECT.

Since the mid-1980s, large-scale integration of semiconductor components has permitted the production of arrays of infrared detectors. These arrays now exist in formats as large as 2048 × 2048 elements (4 million detectors on a single device). Each detector on such an array is also substantially more sensitive than its 1980s counterpart. Infrared astronomy has blossomed thanks to the multiple advantages of observing with so many detectors simultaneously.

Systems of mirrors, gratings, prisms, and lenses direct light to infrared detectors in applications such as photometers, cameras, and spectrographs. Ordinary glass lenses and prisms are insufficiently transmissive at wavelengths longer than 2 μm to be useful in an infrared optical system. Instead, infrared transmissive materials such as calcium fluoride, zinc selenide, silicon, and germanium can be fabricated into lenses and prisms. Calcium fluoride, for example, is transparent to infrared light of wavelengths as long as 10 μm. Mirrors and diffraction gratings in infrared optical systems are often gold-coated since gold is a particularly efficient reflector across the entire infrared region of the spectrum. Many instruments depend on the use of wavelength-band-limiting filters for their operation. These filters are produced by coating an infrared transmissive substrate with multiple alternating layers of material of differing refractive index. With proper tuning of the thickness and number of layers, the constructive and destructive interference of the electromagnetic waves reflected from the layer boundaries can efficiently transmit a well-defined range of wavelengths. *See* INTERFERENCE FILTERS; OPTICAL MATERIALS; REFLECTION OF ELECTROMAGNETIC RADIATION.

Astronomical targets. The targets of infrared observations include ordinary stars and galaxies, planets, brown dwarfs, young stellar objects, evolved stars, starburst galaxies, and redshifted radiation.

Ordinary stars and galaxies. Although popular interest in infrared astronomy focuses on exotic objects observable only using infrared light, infrared observations continue to play a fundamental role in understanding the more pedestrian stars and galaxies that constitute most of the visible-wavelength universe. Just as at visible wavelengths, where the colors and luminosities of stars divide them into distinct spectral classes via the Hertzsprung-Russell diagram, stars can be classified as to their temperature and luminosity class via their infrared "colors." In this case, color refers to the ratio of the brightness of an object compared at two different infrared wavelengths. Sources with proportionally more flux at longer infrared wavelengths are referred to as "red" by infrared astronomers in analogy to objects which literally appear red using visible light. Color-based distinctions can also be made among classes of galaxies. At infrared wavelengths, interstellar dust within the Milky Way Galaxy becomes only a minor hindrance to seeing stars across or galaxies beyond the galactic plane, permitting a more complete census of these objects. At infrared wavelengths longer than 10 μm, infrared light emitted from dust warmed by stars in normal galaxies begins to dominate the infrared light from stars, making dusty spiral galaxies prime targets for mid-infrared observations. *See* GALAXY, EXTERNAL; HERTZSPRUNG-RUSSELL DIAGRAM; MILKY WAY GALAXY; STAR.

Planets. At the shortest infrared wavelengths, planets and asteroids are observed via reflected sunlight. Their infrared spectra are diagnostic of their surface mineralogy and atmospheric chemistry. The outer gas giants have near-infrared reflected spectra which are dominated by the absorption bands of molecular hydrogen and methane in their atmospheres. At infrared wavelengths longer than 3 μm, thermal radiation from the planets begins to dominate over reflected sunlight. Jupiter, still cooling from its initial formation, emits twice as much energy as it receives from the Sun—nearly all of it at infrared wavelengths. *See* ASTEROID; JUPITER; PLANET; PLANETARY PHYSICS.

Brown dwarfs. An object with a mass less than 8% that of the Sun (equivalent to 80 times the mass of Jupiter) is incapable of burning hydrogen in its core to sustain its luminosity—the hallmark of being a star. At the time of their formation, the interiors of such substellar objects are warmed by their gravitational contraction from an initially diffuse interstellar cloud of gas. Immediately following their formation, their surface temperature mimics those of the coolest stars (2000 K or 3100°F). Even at this most luminous point in their evolution, most of the light emerges in the infrared part of the spectrum. Since they have no internal source of energy, the radiation emerging from their surface causes them to cool over time. An object with 3% of the mass of the Sun cools to a surface temperature of 300 K (80°F) after only about 10^9 years. Infrared surveys of the entire sky are beginning to reveal large numbers of these brown dwarfs, and they appear to be more common than ordinary stars in the Milky Way Galaxy.

Young stellar objects. Young stellar objects (YSOs) are newly formed or forming stars. These stars are being assembled by gravity out of dense (1000 atoms per cubic centimeter) interstellar clouds. The environment surrounding the forming star is naturally very dusty, and the slight rotation of the natal cloud combined with gravity drives the infalling material to form a thin flattened disk around the star. This disk may be heated either by the friction of the infalling material or by radiation from the central star. In either case, the disk temperatures range from 2000 K (3100°F) close to the star to 20 K (−424°F) in the outer region of the disk. Virtually all of the disk emission emerges at infrared wavelengths. Planets accrete within these disks, and infrared observations provide the primary astrophysical insight into the process of planetary formation. *See* PROTOSTAR.

Evolved stars. After exhausting the initial supply of hydrogen in the stellar core, stars reconfigure themselves to liberate the energy of hydrogen shell burning around the core and helium burning within the core. Stellar physics dictates that the star must grow to large size and become cool at its surface in order to dissipate the energy being produced within. At this stage, some fraction of stars become unstable to pulsation. Their outer envelopes are weakly bound and easily lost. Dust grains condensing in the expanding envelope can completely enshroud the star. Under these circumstances, the observed radiation emerges entirely in the infrared part of the spectrum as the dense surrounding dust absorbs the short-wavelength starlight and, warmed only to a temperature of a few hundred kelvins, emits exclusively infrared radiation. Stars slightly more massive than the Sun, which are common throughout the Milky Way Galaxy, evolve through this stage. Since these stars are intrinsically luminous and numerous, and because the emergent infrared luminosity penetrates the dust which permeates the Milky Way Galaxy, these stars are ideal tracers of the structure of the Galaxy. *See* GIANT STAR; STELLAR EVOLUTION.

Starburst galaxies. Although normal galaxies emit most of their light at the boundary between the visible and infrared portions of the spectrum, galaxies undergoing active bursts of star formation can produce most of their radiation at wavelengths of 10 μm or longer. The radiation emerges in this part of the spectrum because the star-forming regions are embedded in dust clouds which absorb the starlight and, having been warmed to temperatures of tens of kelvins, reradiate energy largely in the infrared portion of the spectrum. *See* STARBURST GALAXY.

Ultraluminous infrared galaxies. The gravitational interaction between two gas-rich galaxies can induce both objects to undergo an extensive burst of star formation. The resulting energy release can augment the flux of the galaxy by a factor of 10 or more with most of the radiation arising in the infrared part of the spectrum. Such "ultraluminous" infrared galaxies are among the most luminous galaxies in the universe.

Redshifted radiation. The apparent Doppler shift due to the expansion of the universe causes the ultraviolet and visible light originally emitted by extremely distant stars and galaxies to be shifted into the infrared part of the spectrum. Redshifts this large originate from objects at distances of 10^{10} light-years or more. Since the light collected from these objects was emitted by them 10^{10} years ago, these observations probe the state of the universe at the earliest times. Many observational investigations of the early universe focus on collecting and analyzing infrared light from the most distant objects known. Bolometer arrays capable of imaging at a wavelength of 300 μm are now able to detect the first galaxies which formed following the big bang. *See* BIG BANG THEORY; COSMOLOGY; REDSHIFT.

Ground-based infrared astronomy. Telescopes dedicated to visible-wavelength astronomy are also effective collectors of infrared radiation. Equipped with infrared focal-plane-array imagers and spectrographs, these telescopes can observe the infrared universe through the accessible atmospheric windows (Fig. 1). Such observations are particularly effective shortward of 2 μm, where thermal emission from the telescope is negligible relative to other backgrounds. Nearly every telescope larger than 2 m (80 in.) in aperture has one or more dedicated infrared instruments that share time with visible-wavelength images and spectrographs. Since moonlight interferes strongly with visible observations but hardly affects infrared observations, instruments are scheduled in synchrony with the phases of the Moon. Near full moon, most large telescopes conduct infrared observations.

Ground-based infrared observations are less susceptible to blurring due to atmospheric turbulence than visible-wavelength observations. Under ideal circumstances, only diffraction of light through the telescope aperture should blur stellar images. At visible wavelengths, atmospheric turbulence broadens images by a factor of 50 or more compared with the diffraction limit. Modern computers have enabled the incorporation of flexible mirrors into telescope optical paths which compensate for the blurring effects of the atmosphere. At visible wavelengths, the atmosphere changes too rapidly to permit optimal removal of the effects of atmospheric turbulence. The changes occur more slowly at infrared wavelengths, and such adaptive-optics systems can compensate for the atmosphere's effects and yield images with diameters limited by diffraction alone (**Fig. 3**). In addition to providing images of finer detail, the sharper images concentrate more light into a smaller region, making it possible to detect fainter objects. *See* ADAPTIVE OPTICS.

Infrared space missions. The opacity of the Earth's atmosphere at most infrared wavelengths and the need to place a telescope in an environment where the entire telescope structure can be cooled to cryogenic temperatures have motivated a number of extremely successful satellite missions largely devoted to infrared astronomy.

IRAS. The *Infrared Astronomy Satellite* (*IRAS*) conducted the first survey of nearly the entire sky at wavelengths of 12, 25, 60, and 100 μm. The

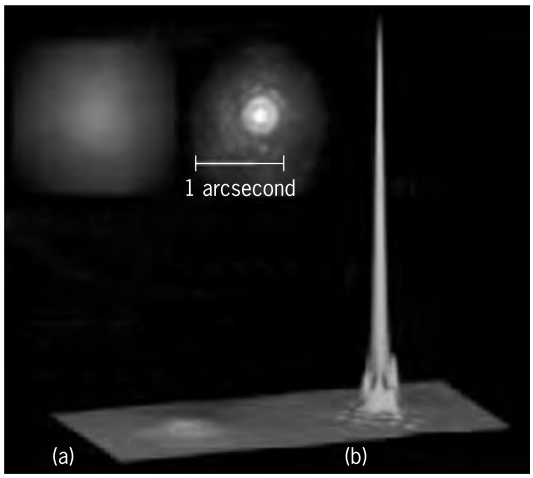

Fig. 3. Example of the improvement in both sharpness and intensity of a stellar image with the application of adaptive optics to compensate for atmospheric turbulence. Plots of image intensity with adaptive optics (*a*) turned off and (*b*) on show improvement both in resolution from 0.6 to 0.068 arcsecond (as measured by the full width of the peak at half its maximum amplitude) and in peak intensity, enabling the detecction of fainter objects. Inset shows the corresponding images. Images were obtained at the Canada-France-Hawaii Telescope (aperture of 3.3 m or 130 in.) at a wavelength of 0.936 μm with an exposure time of 30 s. (*University of Hawaii*)

three longest wavelengths are accessible almost exclusively from space. Launched by the National Aeronautics and Space Administration in 1983, *IRAS* contained a 0.6-m (24-in.) telescope cooled to liquid helium temperature and a focal plane of dozens of individual infrared detectors which scanned across the sky. The resulting catalogs of 350,000 stars and galaxies, along with accompanying images of most of the sky, remain a fundamental scientific resource. As might be expected when opening an entirely new wavelength regime to sensitive observation, the *IRAS* mission revealed a number of fundamentally new phenomena, including residual dust disks around main-sequence stars (presumably the debris of evolving solar systems), ultraluminous infrared galaxies, and tenuous filaments of interstellar dust now known as infrared cirrus.

ISO. The *IRAS* mission placed the infrared sky in context with an all-sky survey. In 1995, the European Space Agency launched the *Infrared Space Observatory* (*ISO*), the first infrared observatory with the capability of making detailed photometric and spectroscopic studies of individual objects at wavelengths ranging from 2.5 to 240 μm. The *ISO*'s 0.6-m (24-in.) liquid-helium-cooled telescope operated for $2^1/_2$ years, making more than 20,000 independent observations of astronomical sources. The telescope's focal plane included an infrared camera (ISOCAM), with sensitivity from 2.5 to 17 μm, which imaged the sky onto two different 32 × 32-element infrared focal-plane arrays; an infrared photometer and polarimeter (ISOPHOT), which used individual detectors to measure precisely the quantity of infrared radiation from celestial sources with detectors sensitive between 2.5 and 240 μm; and a pair of spectrometers observing at wavelengths between 2.4 and 200 μm. *ISO* observations have contributed to nearly every area of astrophysics, ranging from spectroscopic studies of planetary atmospheres to infrared imaging of forming galaxies in the early universe.

SIRTF. Scheduled for launch into an Earth-trailing orbit in 2001, the *Space Infrared Telescope Facility* (*SIRTF*) will place a sophisticated battery of infrared cameras and spectrographs behind a 0.85-m (33.5-in.) liquid-helium-cooled telescope with a 5-year cryogen lifetime. *SIRTF*'s infrared array camera (IRAC) will contain four 256 × 256-element infrared arrays spanning the wavelength range from 3.6 to 8.0 μm. These arrays are substantially larger and more sensitive than those available on *ISO* and illustrate the rate of advance of infrared imaging technology. *SIRTF* will have an imaging capability to wavelengths as long as 160 μm and spectroscopic coverage between 4 and 40 μm. *SIRTF* joins the *Compton Gamma-Ray Observatory*, the *Chandra X-ray Observatory*, and the Hubble Space Telescope (observing at visible and near-infrared wavelengths) as one of NASA's "great observatories," which operate across the electromagnetic spectrum. *SIRTF* will illuminate nearly every area of astrophysics and is particularly well suited to revealing the abundance of cool brown dwarfs and ultraluminous infrared galaxies.

Prospects. More ambitious plans for space telescopes emphasize infrared observational capability because of the range of astrophysics addressed in this wavelength regime as well as relaxed design constraints on telescopes, since infrared wavelengths are several times larger than visible-light wavelengths. NASA's long-term goal involves detecting and imaging planets around other stars. Infrared imaging and interferometry play a central role in the realization of this objective. Since the planets of interest naturally reside next to an extremely bright star, an interferometer—which combines the light from two or more well-separated telescopes—is required to isolate the planet's feeble light from that of the star.

Interferometry has been pioneered largely at infrared wavelengths from the ground, since the Earth's atmosphere appears more stable when observing infrared light than when observing visible light. In addition, the contrast between the star and adjacent planet is substantially improved at infrared wavelengths, making the planet easier to discern in the star's glare. Infrared spectroscopic observations are diagnostic of water and ozone in planetary atmospheres and will be fundamental in assessing the probability that life could arise or even has developed on extrasolar planets. Michael F. Skrutskie

Bibliography. G. Basri, The discovery of brown dwarfs, *Sci. Amer.*, 282(4):76–83, April 2000; H. Freudenreich, Deconstructing the Milky Way, *Amer. Sci.*, 87(5):418–427, September–October 1999; C. Kraan-Korteweg and O. Lahav, Galaxies behind the Milky Way, *Sci. Amer.*, 279(4):50–57, October 1998; M. Skrutskie, Charting the infrared sky, *Sky Telesc.*, 94(2):46–47, August 1997.

Infrared imaging devices

Devices that convert an invisible infrared image into a visible image. Infrared radiation is usually considered to span the wavelengths from about 0.8 or 0.9 micrometer to several hundred micrometers; however, most infrared imaging devices are designed to operate within broad wavelength regions of atmospheric transparency, that is, the atmospheric windows. At sea level, for horizontal paths of a few kilometers' length, these are approximately at 8–14 μm, 3–5 μm, 2–2.5 μm, 1.5–1.9 μm, and wavelengths shorter than 1.4 μm. The radiation available for imaging may be emitted from objects in the scene of interest (usually at the longer wavelengths called thermal radiation) or reflected. Reflected radiation may be dominated by sunlight or may be from controlled sources such as lasers used specifically as illuminators for the imaging device. The latter systems are called active, while those relying largely on emitted radiation are called passive. Active optical imaging systems were developed to achieve a nighttime aerial photographic capability, and work during World War II pushed such systems into the near-infrared spectral region. Development of passive infrared imaging systems came after the war, but only the advent of lasers allowed creation of active infrared imagers at wavelengths much longer than those of the photographic region. Striking advances have been made in active infrared systems which utilize the coherence available from lasers, and hybrid active-passive systems have been studied intensively. *See* INFRARED RADIATION; LASER.

Although developed largely for military purposes, infrared imaging devices have been valuable in industrial, commercial, and scientific applications. These range from nondestructive testing and quality control to earth resources surveys, pollution monitoring, and energy conservation. Infrared images from aerial platforms are used to accomplish "heat surveys," locating points of excessive heat loss. An example is shown in **Fig. 1***a*. As discussed below, calibration allows association of photographic tones in this figure with values of apparent (that is, equivalent blackbody) temperatures. Dark areas in the figure are "colder" than light ones. *See* NONDESTRUCTIVE EVALUATION; REMOTE SENSING.

Scanning systems. Infrared imaging devices may be realized by electrooptical or optomechanical scanning systems. All have an optical head for receiving the infrared radiation and a display for the visible image. These are connected by electronics for the passage of electrical signals from the detector element(s) to the display input. Signal processing may be incorporated in the electronics to selectively enhance or reduce features in the produced visible image. For example, in Fig. 1*b* a "level-slice" technique presents in white all areas (mainly rooftops) with apparent temperatures between −7.9 and −8.9°C or 17.4 and 15.6°F. (The ambient air temperature was −5°C or 23°F.) Black regions in the figure correspond to apparent temperatures below or above the narrow "sliced" temperature range of the white regions.

Fig. 1. Thermal imagery in the wavelength range 10.4–12.5 μm obtained during flights over Ypsilanti, Michigan, at 2400 hours, November 23, 1975, by the Airborne Multispectral Scanner operated by the Environmental Research Institute of Michigan. (*a*) Calibrated thermal imagery. (*b*) Signal-processed thermal imagery of same scene. (*From F. Tannis, R. Sampson, and T. Wagner, Thermal imagery for evaluation of construction and insulation conditions of small buildings, Environmental Research Institute of Michigan, ERIM Rep. 116600-12-F, July 1976*)

Optomechanical methods such as rotating prisms or oscillating mirrors may be used to sample or scan the spatial distribution of infrared radiation in either the object or image plane. Electrooptical imaging devices may use an electron beam (for example, vidicons) or charge transport in solids (for example, charge-coupled devices, or CCDs) to scan the infrared image formed by the optics of the device. This image-plane scanning places more stringent requirements upon the optics for image quality off-axis than does use of mechanically moved optical elements placed before the entrance aperture of the system. Intensive development of pyroelectric vidicons, detector arrays, and infrared charge-coupled devices (IRCCDs) has taken place, reflecting the critical role played by the detector element in all infrared systems. The spectral, spatial, and temporal responses of detectors are the major factors in determining the wavelength regions, the spatial resolution, and the frequency response (that is, the time constant) of imaging devices. *See* CHARGE-COUPLED DEVICES.

Detector arrays. Optomechanical scanning methods often stress the response time of the detector-electronics package. As a result, multiple detectors or detector arrays are sometimes incorporated in the focal planes, resulting in partially electronically scanned optomechanical systems. The technology for use of a linear array of detector elements (often lead selenide and indium antimonide detectors for the 3–5-μm region, and mercury-doped germanium and mercury cadmium telluride for the 8–14-μm window) is well developed, and the use of a two-dimensional array or matrix of detectors has

been studied. Optomechanical imagers incorporating such arrays allow the use of time delay and integration (TDI) of the signals to improve the resulting signal-to-noise ratios.

Solid-state components such as charge-coupled devices afford the opportunity for implementation of signal processing directly at the focal plane. Two approaches have been undertaken to attain the focal-plane array technology of infrared charge-coupled devices. In one, the development of a hybrid device, an infrared detector matrix of any suitable photodetector material, for example, indium antimonide, mercury cadmium telluride, and lead tin telluride, is mated with a conventional silicon charge-coupled device. Thus two solid-state wafers or "chips" are integrated to obtain an infrared charge-coupled device. In the other, the goal is a monolithic chip, one incorporating the photodetection, charge generation, and charge transfer in a structure made of essentially one material. Candidate materials include impurity-doped silicon, indium antimonide, and mercury cadmium telluride. The hybrid device technology can be implemented more readily than that needed for monolithic infrared charge-coupled devices. The development of infrared charge-coupled devices with the number of detecting elements in a sufficiently closely packed array required for high-performance infrared imaging devices involves additional difficulties.

Scanning motion. Some optomechanical imagers produce a two-dimensional scan entirely by movement of components of the device itself; others utilize the motion of a platform such as an aircraft or satellite. The first kind of system includes the forward-looking infrared (FLIR) or framing imagers which usually scan in televisionlike raster patterns and display, synchronously if done in real time, a visible image corresponding to the spatial distribution of infrared radiation. These visible image outputs have been named thermographs and thermograms. Commercially available imaging devices of this type have used horizontally and vertically rotating germanium prisms, mirrors oscillated in orthogonal directions, two mirrors and a six-sided rotating prism, and other schemes to produce images at relatively slow rates from 16 per second to less than a quarter of a frame per second. Higher-performance systems have been produced for military purposes. The second class of imaging systems includes those often called line scanners or thermal mappers. One such system, the 12-channel Airborne Multispectral Scanner operated by the Environmental Research Institute of Michigan (ERIM; **Fig. 2**), includes two thermal radiation channels, at 8.2–9.3 μm and 10.4–12.5 μm whose magnetic-tape recorder output was processed to produce the thermal imagery in Fig. 1.

Multispectral imaging systems can be used to collect large amounts of data that may be processed to create special-purpose images. Advanced image data processing techniques are being used in remote sensing applications such as mineral detection and assessment, vegetation cover classification, agricultural crop monitoring and management, and urban analyses and planning. If the system has more than about 30 spectral channels with relatively fine spectral resolution, imaging spectroscopy can be carried out. These hyperspectral imaging systems, compared with multispectral ones, have data outputs better described by more complex representations than simple images. Hypercubes, many-dimensional vectors that need multidimensional data analysis methodologies, have become especially useful. Industrial and laboratory applications now include spatially resolved infrared microspectroscopy. Such systems are becoming available as a direct result of the creation of modern focal-plane arrays (FPAs) for use in, for example, fast Fourier transform infrared (FT-IR) imaging systems. A simple illustration of a system for remote sensing applications from airborne or space-based platforms would look similar

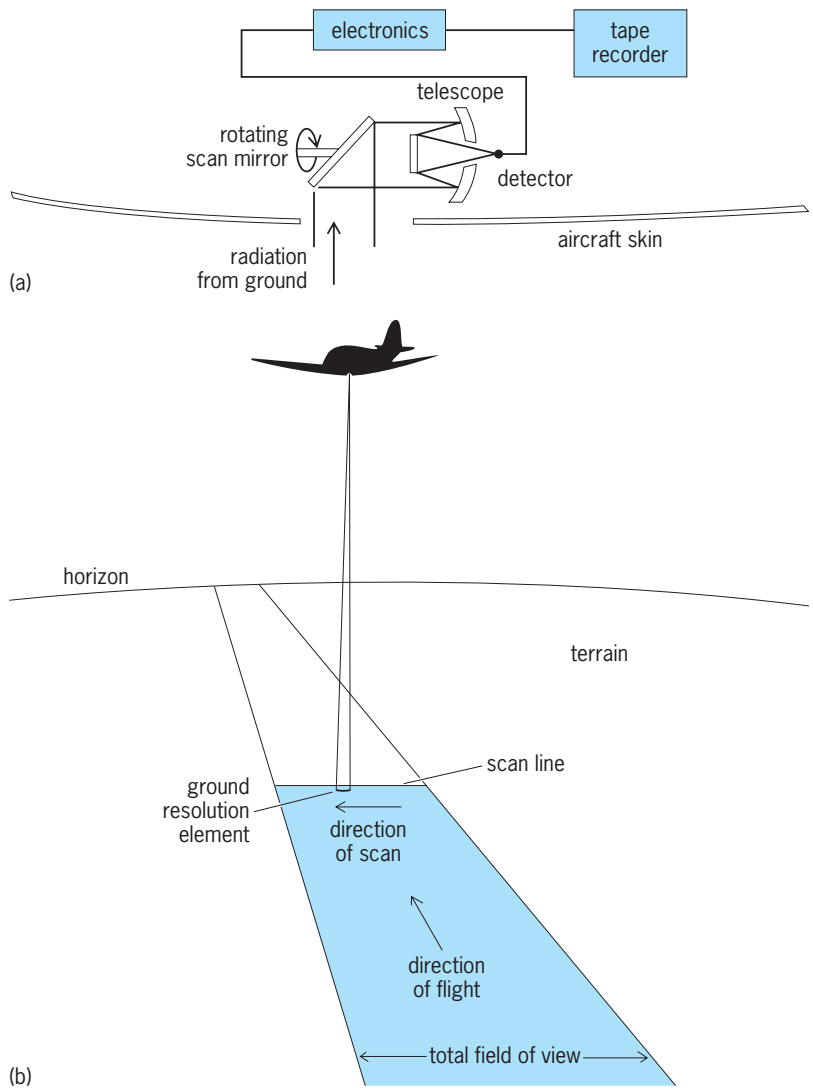

Fig. 2. Airborne Multispectral Scanner. (*a*) Schematic diagram of equipment. (*b*) Scanning operation, utilizing motion of aircraft. (*After F. Tannis, R. Sampson, and T. Wagner, Thermal imagery for evaluation of construction and insulation conditions of small buildings, Environmental Research Institute of Michigan, ERIM Rep. 116600-12-F, July 1976*)

to Fig. 2*a*, with the detector replaced by an entrance aperture for a multispectral or hyperspectral dispersing element, such as a prism or grating or interferometric subsystem. *See* FOURIER SERIES AND TRANSFORMS.

Characterization of output. The instantaneous field of view (IFOV) or resolution element of imaging systems is always geometrically filled by the radiating source, so that the output of the device is a response to changes in amount of radiation from field of view to field of view. These changes are best characterized in terms of radiance L, the radiant flux per unit area per unit solid angle, usually in a selected spectral band of wavelengths, $\Delta\lambda$. Even in the infrared regions, the radiance variation may be ascribed to changes in reflectance, incident radiation, emissivity, or temperature. By restriction to wavelengths dominated by emission, the so-called thermal wavelengths longer than 3.5 μm, the radiance change can be described by the equation below, where T is the

$$\Delta L = \frac{\partial L}{\partial T}\,\Delta T + \frac{\partial L}{\partial \epsilon}\,\Delta\epsilon$$

absolute temperature and ϵ is the emissivity. Contributions due to $\Delta\epsilon$ are usually treated as changes in an equivalent blackbody temperature by setting $\epsilon = 1$ and $\Delta\epsilon = 0$. Then T represents an equivalent blackbody temperature, and the radiance variation can be ascribed entirely to a value of ΔT. That value of ΔT corresponding to a radiance difference which will just produce a signal-to-noise ratio of 1 is called the noise equivalent temperature difference (NETD). One can also characterize the performance of an imaging system by a noise equivalent emissivity difference or even a noise equivalent reflectivity difference. The use of noise equivalent temperature difference as a figure of merit for thermal imagers is obviously more appropriate. For the higher-performance forward-looking infrared systems, a useful figure of merit is the minimum resolvable temperature difference (MRTD), a figure of merit which includes the characteristics of the display and of the observer as well. *See* EMISSIVITY; HEAT RADIATION.

Display. If a visible image is the desired final output of an infrared imaging device, it may be displayed as a conventional television picture by means of a cathode-ray tube (CRT). Cathode-ray-tube technology has been developed to a level that is suitable, and research has been undertaken toward creation of satisfactory flat panel displays using liquid crystal elements, light-emitting diodes, or plasma panels. Systems not requiring a real-time image display may utilize analog or digital data storage or transmission systems, which then are used to produce permanent visual records such as photographs. High-resolution "hard copy" images can be produced by sophisticated systems using electron-beam or laser recording on film. Complex signal-processing techniques are easily introduced before the final image recording is made. *See* CATHODE-RAY TUBE; LIGHT-EMITTING DIODE; LIQUID CRYSTALS.

George J. Zissis

Bibliography. J. S. Accetta and D. L. Shumaker (eds.), *The Infrared and Electro-Optical Systems Handbook*, vols. 1–8, ERIM and SPIE, 1993; M. Bass with E. Stryland, D. R. Williams, and W. L. Wolfe (eds.), *Handbook of Optics*, 2d ed., vols. 1 and 2, Optical Society of America/McGraw-Hill, 1995; G. C. Holst, *CCD Arrays, Cameras, and Displays*, 2d ed., SPIE and JCD, 1998; G. C. Holst, *Testing and Evaluation of Infrared Imaging Systems*, 2d ed., SPIE and JCD, 1998; W. L. Wolfe, SPIE Tutorial Texts: *Introduction to Infrared System Design*, 1996, *Introduction to Imaging Spectrometers*, 1997, *Introduction to Radiometry*, 1998, and with G. J. Zissis (eds.), *The Infrared Handbook*, 1978, SPIE/ERIM-IRIA.

Infrared lamp

A special type of incandescent lamp that is designed to produce energy in the infrared portion of the electromagnetic spectrum. The lamps produce radiant thermal energy which can be used to heat objects that intercept the radiation. All incandescent lamps have radiation in three regions of the electromagnetic spectrum, the infrared, the visible, and the ultraviolet. An infrared lamp with a filament operating at 2500 K will release about 85% of its energy in the form of thermal radiant energy, about 15% as visible light, and a tiny fraction of a percent as ultraviolet energy. The amount of infrared radiation produced by a lamp is a function of the input wattage of the lamp and the temperature of its tungsten filament. For most infrared lamps, the thermal radiation will be about 65–70% of the input wattage. A typical 250-W infrared lamp radiates about 175 W of thermal energy.

Temperature and energy spectrum. The temperature of the incandescent filament determines the portion of the energy output allocated to each spectral region. As the filament temperature increases, the wavelength at which peak energy output occurs shifts toward the visible region of the spectrum (**Fig. 1**). Lamp filament temperatures are always stated in kelvins (K). *See* HEAT RADIATION; TEMPERATURE.

Types. The lamps are supplied in two shapes (**Fig. 2**). The most common shape, for general use, is the R lamp (Fig. 2*a*), since the reflector unit is built in and the lamp needs only a suitable socket to form an infrared heating system. These lamps are available in 125-, 250-, and 375-W ratings, all in glass bulbs with a nominal reflector diameter of 12.5 cm (5.0 in.) and voltage rating of 115–125 V.

The other type of infrared lamp, the tubular quartz bulb lamp (Fig. 2*b*), is used with a separate external reflector designed to distribute the heat as desired. More than 30 sizes of the tubular quartz bulb lamps are available in 375–5000-W ratings. Lighted lengths range from 12.5 cm (5 in.) to 127 cm (50 in.), overall lengths (including connecting terminals) from 22.38 cm (8.8 in.) to 138.68 cm (53.8 in.), and tube diameter is 0.95 cm (0.375 in.). Voltage ratings are a

function of length, with nominal values ranging from 120 V for short tubes to 960 V for the longest tubes.

Because the lamp bases become quite hot, R lamps (Fig. 2*a*) should be used only in porcelain sockets, rather than in the common brass shell socket which would fail with prolonged use, posing a fire hazard. The R lamp is available with the glass face of the bulb stained red or ruby, reducing the amount of light emitted by the lamp. This reduces the undesirable glare in situations where the lamps must be in sight of the workers.

Uses. The major advantage of infrared heating is that it is possible to heat a surface that intercepts the radiation without heating the air or other objects that surround the surface being heated. Infrared lamps have many uses, some of which will be discussed.

Paint drying. Many industrial infrared ovens are used to dry painted surfaces as they pass by on a conveyor belt. Drying times are much shorter than for other methods of drying paint.

Evaporative drying. Infrared lamps can be used to remove moisture or solvents from the surface of objects by speeding the evaporation process. Porous materials with internal moisture require longer exposure, since the entire object must be heated by conducting heat from the surface to the interior.

Farm heating of animals. The lamps can be used to heat brooders and to keep other livestock warm without having to heat a large, poorly insulated shed.

Food. Many R lamps are used in restaurant kitchens to keep food warm while it is waiting to be served to customers.

Comfort heating. Areas under theater marquees can be made more pleasant for patrons waiting in line to buy tickets, and sidewalk cafes and similar areas can be made comfortable in relatively chill weather by installing suitable lamp systems.

Other uses. These include therapeutic heating of portions of the body to relieve muscle strains, tobacco drying, textile and wallpaper drying, and drying of printing inks so that press operation can be speeded up without causing image offset. *See* THERMOTHERAPY.

Design of heating systems. Most infrared heating ovens or systems are designed by calculating the number and wattage of lamps, plus experimenting with the system after it is constructed. The latter is necessary because the number of variables is large, and precise data on the magnitude of the variables are often difficult to obtain.

Voltage and lifetime. Moderate variations in voltage will alter lamp operation somewhat. A 1% increase in voltage will increase both the input and radiated power in watts by 1.5–2%, and it will also increase the filament temperature and the light from the lamp. Most lamps are rated for use over a range of voltage, such as 115–125 V for a lamp intended for use on a 120-V power system. The life of all infrared lamps exceeds 5000 h. The greatest cause of life reduction is subjecting the lamp to vibration.

Exposure to radiation. Exposure of persons to moderate levels of infrared is not known to be harmful.

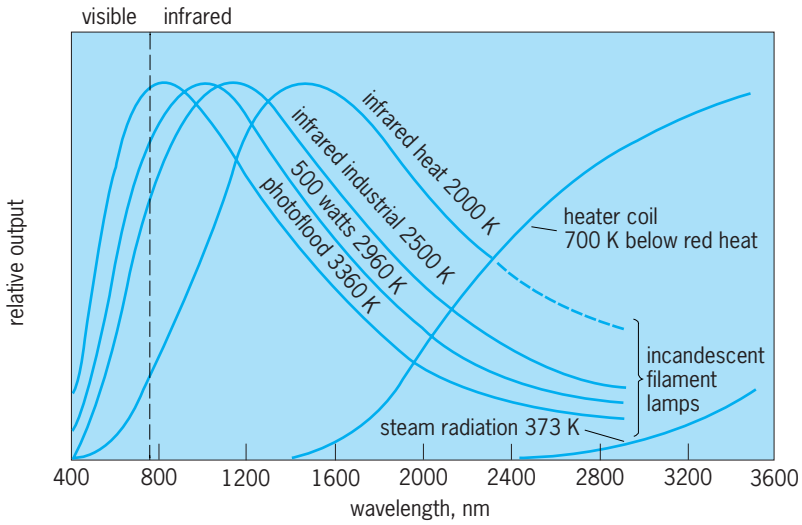

Fig. 1. Relative energy output of four incandescent lamps as a function of wavelength and filament temperature. As temperature increases, more of the lamp's energy shifts to shorter wavelengths. °F = (K × 1.8) − 460. (*After Illuminating Engineering Society, IES Lighting Handbook, 8th ed., 1993*)

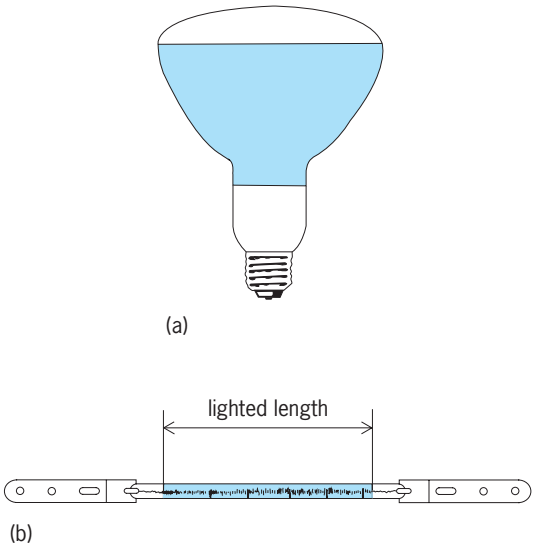

Fig. 2. Typical shapes of infrared lamps. (*a*) R lamp, with built-in reflector unit. (*b*) Tubular configuration with a quartz bulb.

High levels could cause burns, but the exposed individual feels the heat and would normally move away before a burn actually occurred. *See* INCANDESCENT LAMP; INFRARED RADIATION.

G. R. Peirce; Robert Leroy Smith

Bibliography. K. J. Button (ed.), *Infrared and Millimeter Waves: Sources of Radiation*, vol. 1, 1979; J. D. Hall, *Industrial Applications of Infrared*, 1947; Illuminating Engineering Society, *IES Lighting Handbook*, 8th ed., 1993; M. La Toison, *Infrared and Its Thermal Applications*, 1964; W. L. Wolfe and G. J. Zissis (eds.), *The Infrared Handbook*, 1978, reprint 1985.

Infrared radiation

Electromagnetic radiation in which wavelengths lie in the range from about 1 micrometer to 1 millimeter. This radiation therefore has wavelengths just a little longer than those of visible light and cannot be seen with the unaided eye. The radiation was discovered in 1800 by William Herschel, who used a prism to refract the light of the Sun onto mercury-in-glass thermometers placed just past the red end of the visible spectrum generated by the prism. Because the techniques and materials used to collect, focus, detect, and display infrared radiation are different from those of the visible, and because many of the applications of infrared radiation are also quite different, a technology has arisen, and many scientists and engineers have specialized in its application. *See* ELECTROMAGNETIC RADIATION.

Infrared techniques. A complete infrared system consists of a radiating source and background, intervening medium, optical system, detector, electronics, and display.

Sources. The source can be described by the spectral distribution of power emitted by an ideal radiating body. This distribution is characteristic of the temperature of the body. A real body is related to it by a radiation efficiency factor also called emissivity. It is the ratio at every wavelength of the emission of a real body to that of the ideal under identical conditions. **Figure 1** shows curves for these ideal bodies radiating at a number of different temperatures. The higher the temperature, the greater the total amount of radiation. The total number of watts per square meter is given by $5.67 \times 10^{-8}T^4$, where T is the absolute temperature in kelvins (K). Higher temperatures also provide more radiation at shorter wavelengths. This is evidenced by the maxima of these curves moving to shorter wavelengths with higher temperatures. Ideal radiators are also called blackbodies. *See* BLACKBODY; EMISSIVITY; HEAT RADIATION.

The sources can be either cooperative or uncooperative. Some examples of the former include tungsten bulbs (sometimes with special envelopes), globars, and Nernst glowers. These are made of rare-earth oxides and carbons. They closely approximate blackbodies and are used mostly for spectroscopy. Lasers have been used in special applications. Although they provide very intense, monochromatic, and coherent radiation, they are limited in their spectral coverage. The principal infrared lasers have been carbon dioxide (CO_2) and carbon monoxide (CO) gas lasers and lead-tin-tellurium (PbSnTe) diodes. *See* INFRARED LAMP; LASER.

Transmitting medium. The radiation of one of these sources propagates from the source to the optical collector. This path may be through the vacuum of outer space, 3 ft (1 m) of laboratory air, or some arbitrary path through the atmosphere. **Figure 2** shows a low-resolution transmission spectrum of the atmosphere and the transmissions of the different atmospheric constituents. Two of the main features are the broad, high transmission regions between 3 and 5 μm and between 8 and 12 μm. These are

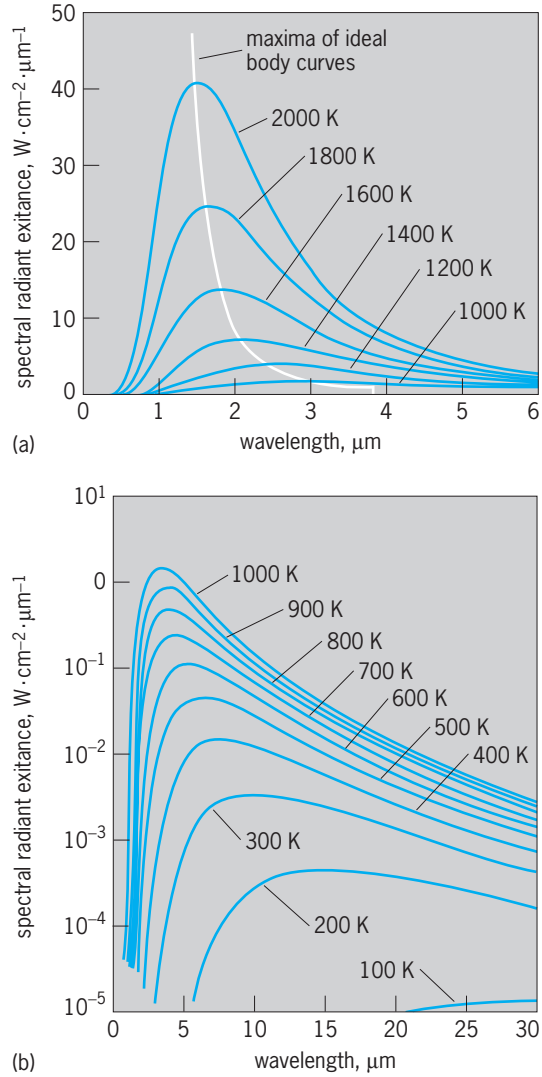

Fig. 1. Radiation from ideal bodies (blackbodies) at different temperatures, shown on (*a*) a linear scale and (*b*) a logarithmic scale. °F = (K × 1.8) − 460.

the spectral regions chosen for most terrestrial applications. **Figure 3** shows a small section of this spectrum, illustrating its complexity. The radiation from the source is filtered by the atmosphere in its propagation so that the flux on the optical system is related to both the source spectrum and the transmission spectrum of the atmosphere.

Optical system. A lens, mirror, or a combination of them is used to focus the radiation onto a detector. Since glass is opaque in any reasonable thickness for radiation of wavelengths longer than 2 μm special materials must be used. The **table** lists the properties of some of the most useful materials for infrared instrumentation. In general, these are not as effective as glass is in the visible, so many infrared optical systems use mirrors instead. The mirrors are characterized by the blank and by its coating. Blanks are usually made of aluminum, beryllium, or special silica materials. The choices are based on high strength, light weight, and good thermal and mechanical stability. They are coated with thin evaporated layers of

aluminum, silver, or gold. The reflectivities of thin-film metallic coatings increase with wavelength, and the requirements for surface smoothness are also less stringent with increasing wavelength. *See* MIRROR OPTICS; OPTICAL MATERIALS.

Detectors. Photographic film is not useful for most of the infrared spectrum. The response of the silver halide in the film, even when specially processed, is only sensitive to about 1.2-μm. It cannot respond to radiation in the important 3–5-μm and 8–12 μm atmospheric transmission bands. If there were a film for the infrared, it would have to be kept cold and dark before and after exposure. Accordingly, infrared systems have used point (elemental) detectors or arrays of them. These detectors are based either on the generation of a change in voltage due to a change in the detector temperature resulting from the power focused on it, or on the generation of a change in voltage due to some photon-electron interaction in the detector material. This latter effect is sometimes called the internal photoelectric effect. Electrons which are bound to the different atomic sites in the crystal lattice receive a quantum of photon energy. They are freed from their bound lattice positions and can contribute to current flow. The energy in electronvolts required to do this is $1.24/\lambda$ (where λ is the wavelength in micrometers). Thus only a very small binding energy, about 0.1 eV, is permitted in photon detectors. The thermal agitation of the lattice could cause spontaneous "emission," so most infrared photodetectors are cooled to temperatures from 10 to 100 K (−424 to −280°F). This does not affect the speed of response of photodetectors, which depends upon photon-electron interactions, but it does slow down thermal detectors. These are sometimes cooled because they generally respond to a relative change in temperature (and for a lower temperature, a small absolute change gives a larger relative change), and thermal noise is also reduced at low temperatures. *See* BOLOMETER; PHOTOCONDUCTIVE CELL; PHOTODIODE; PHOTOTRANSISTOR; PHOTOVOLTAIC CELL; RADIOMETRY; THERMOCOUPLE.

Electronic circuitry. The voltage or current from the detector or detector array is amplified by special circuitry, usually consisting of metal oxide semiconductor field-effect transistors (MOSFETs) which are designed for low-temperature operations. The amplified signals are then handled very much like television signals. One important feature of most of these systems is that the system does not yield a direct response; only changes are recorded. Thus a "dc restore" or absolute level must be established with a thermal calibration source. The black level of the display can then be chosen by the operator. *See* TRANSISTOR.

Reticle. A reticle or chopper is an important feature of nonimaging systems. A typical reticle and its use are shown in **Fig. 4**. An image of the scene is portrayed on a relatively large fraction of the reticle—anywhere from 10 to 100%. The lens just behind it collects all the light that passes through the reticle and focuses it on a detector. The reticle is rotated.

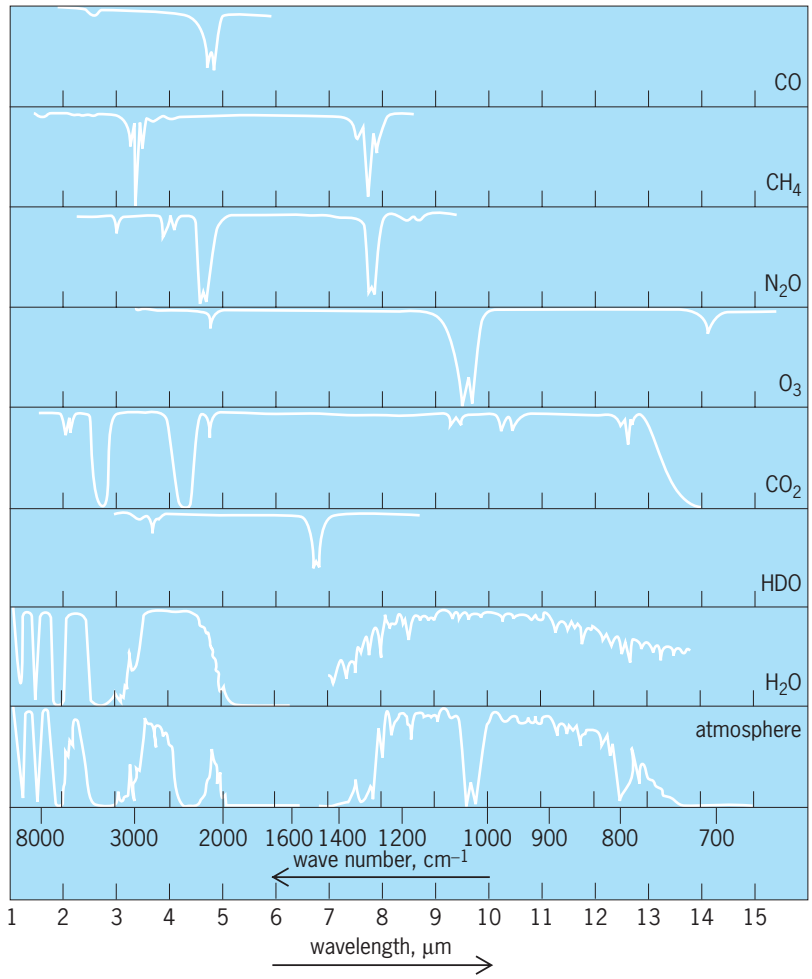

Fig. 2. Low-resolution transmission spectra of the atmosphere and of constituent gases.

If the scene is uniform, there will be no change in the detector output. However, if the scene has a point source (like the image of the hot exhaust of an aircraft engine), then a point image is formed on the reticle and a periodic detector output is generated. The phase and frequency of this output can be used with properly designed reticles to obtain the angular coordinates of the point source. The reticle pattern can also be used to reduce the modulation obtained from almost uniform scenes which radiate from large areas. Reticles are used in most infrared tracking systems, although other schemes are sometimes employed.

Applications. Infrared techniques have been applied in military, medical, industrial, meteorological, ecological, forestry, agricultural, chemical, and other disciplines.

Meteorological applications. Weather satellites use infrared imaging devices to map cloud patterns and provide the imagery seen in many weather reports. Substances at temperatures between 200 and 300 K (−100 and +80°F) emit copious amounts of infrared radiation, but are not hot enough to emit in the visible. The Earth has a temperature of approximately 300 K (80°F), and high-altitude clouds are colder (about 250 K or −10°F). An infrared sensor placed

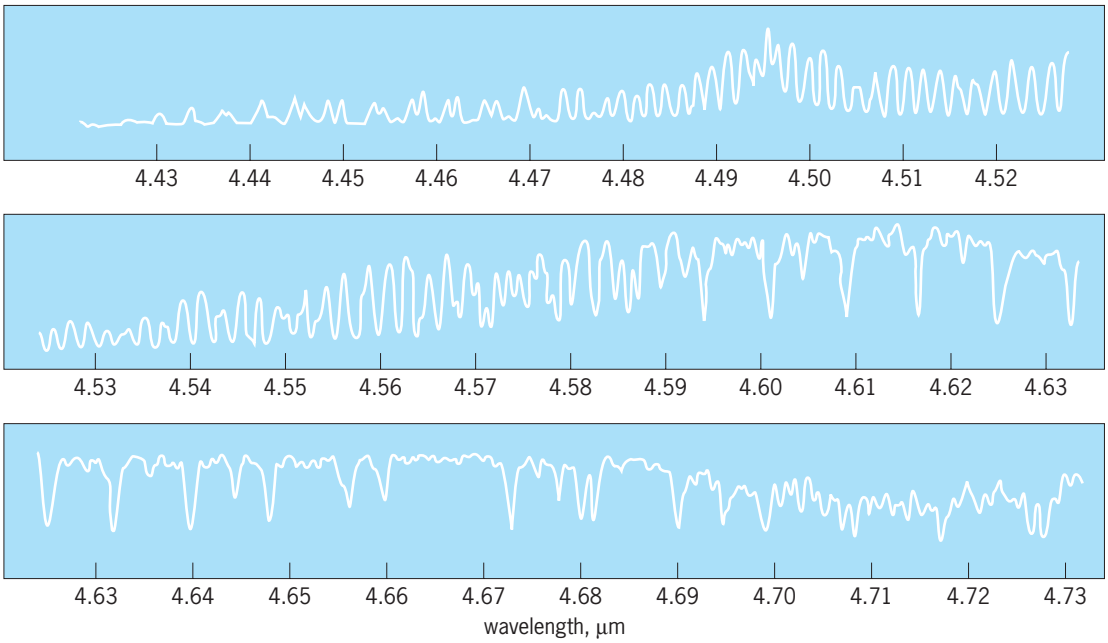

Fig. 3. High-resolution atmospheric transmission spectrum between 4.43 and 4.73 μm.

on a synchronous-orbit satellite can easily sense this difference and present a picture of these temperature patterns. Although the technique is not very widely known, radiometric sensors on lower-altitude satellites can determine the vertical temperature distribution along any path to the Earth. The infrared emission of the Earth's atmosphere is a function of wavelength; a space-borne radiometer senses the emitted radiation at several different wavelengths. The different wavelength bands correspond to greater atmospheric transmission and some "look" deeper into the atmosphere than others. Calculations determine the atmospheric temperature distribution that is required to produce the observed wavelength distribution based on these facts. Although there can be problems with the uniqueness and conversion of such an inversion or fitting process, the results obtained have been accurate to within a few degrees and about a kilometer of altitude. *See* METEOROLOGICAL SATELLITES; REMOTE SENSING.

Medical applications. Infrared imaging devices have also been used for breast cancer screening and other medical diagnostic applications. In most of these applications, the underlying principle is that pathology produces inflammation, and these locations of increased temperature can be found with an infrared imager. The major disadvantage is that localized increases of about 1 K usually need to be detected, and a fair number of nonpathological factors (or at least not the ones in question) can cause equivalent heating. In addition to breast cancer detection (which detects 80% or more cases and has about a 15% false-alarm rate), infrared techniques have been used in the analysis of severe burns, faulty circulation, back problems, and sinus ailments, and has even been proposed to test for medical malingerers.

Airborne infrared imagers. Airborne infrared imagers have been used to locate the edge of burning areas in forest fires. Typically, a forest fire in the western United States is ignited by the lightning of a late afternoon thunderstorm; the valley becomes filled with smoke by the next morning when the crews arrive. The longer wavelengths of the emitted infrared radiation penetrate the smoke better than the visible wavelengths, so the edges of fire are better delineated.

Properties of materials for infrared instrumentation

Material	Transmission region, μm	Approximate refractive index	Comment
Germanium	2–20	4	Opaque when heated
Silicon	2–15	3.5	Opaque when heated
Fused silica	0.3–2.5	2.2	Strong, hard
Zinc selenide	0.7–15	2.4	Expensive
Magnesium fluoride	0.7–14	1.6	Not very strong
Arsenic sulfur glass	0.7–12	2.2	Not always homogeneous
Diamond	0.3–50	1.7	Small sizes only
Salt	0.4–15	1.5	Attacked by moisture

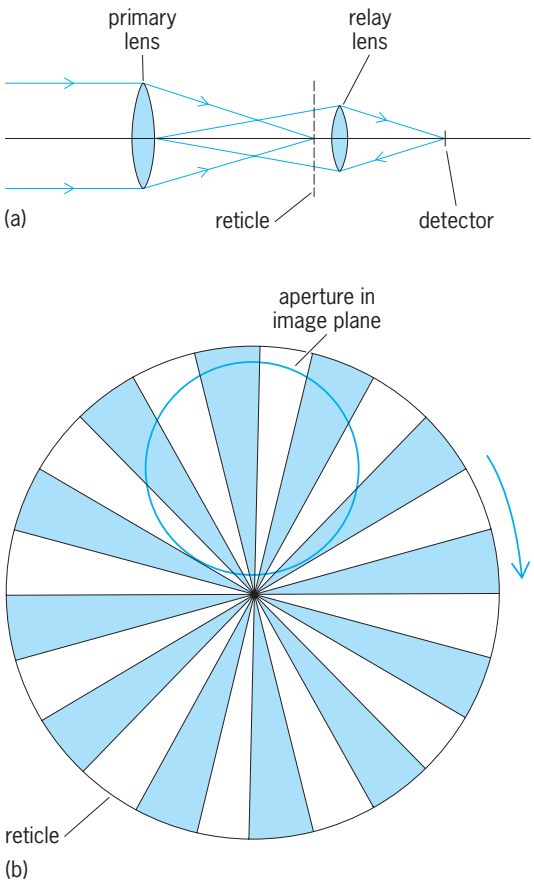

Fig. 4. Reticle system. (*a*) Configuration of components of system. (*b*) Reticle, showing area covered by image.

Thermal pollution contained in the power-plant effluent into rivers has been detected in various locations of the United States, and the viability of crops by assessment of the moisture content has also been accomplished with some degree of success. *See* INFRARED IMAGING DEVICES.

Infrared spectroscopy. Infrared spectrometers have long been a powerful tool in the hands of the analytical chemist. The spectrum of a substance in either absorption or emission provides an unmistakable "fingerprint." The only problem is to find the spectrum of the substance that matches the unknown. *See* INFRARED SPECTROSCOPY.

Military and space applications. The best-known military techniques are probably the Sidewinder air-to-air missile, which detects the radiation from the exhaust of a jet aircraft, and the spaceborne systems which detect the extremely large amounts of radiation from the plume of launching intercontinental rockets at great distances. Infrared systems have also been used for stabilizing satellites by sensing the Earth's horizon, for night vision applications, and for perimeter warning and gas detection. William L. Wolfe

Bibliography. P. Klocek (ed.), *Handbook of Infrared Optical Materials*, 1991; J. L. Miller, *Principles of Infrared Technology*, 1994; M. Schlessinger, *Infrared Technology Fundamentals*, 2d ed., 1994; J. D. Vincent, *Fundamentals of Infrared Detector Operation and Testing*, 1990; W. L. Wolfe and G. J. Zissis, *The Infrared Handbook*, 1978, reprint 1985.

Infrared radiation (biology)

Infrared radiations occupy the span between the visible spectrum and radio waves, and encompass wavelengths 780–400,000 nanometers, neither boundary being precisely delimited. All bodies above absolute zero in temperature are potential sources of infrared radiations; therefore, all organisms are continually exposed to them. About 60% of the Sun's rays are infrared. Water absorbs infrared radiations strongly, except for the band of transparency between 780 and 1400 nm. Since protoplasm contains much water, it absorbs infrared radiations readily. A large animal absorbs infrared radiations at its surface, only the span from 780 to 1400 nm penetrating as far as the blood vessels. *See* INFRARED RADIATION.

While many substances and even tissues selectively absorb infrared rays, and one might therefore postulate selective effects of these radiations, none have been unequivocally demonstrated except possibly in conjunction with x-rays. The reason is perhaps because quanta of infrared radiation do not excite energy states in molecules other than those excited by conducted heat. *See* X-RAYS.

The essential biological effect of infrared rays depends primarily upon the rise in temperature produced following their absorption, which in turn increases the rate of biological activities in proportion to the temperature change. Because of the prominence of infrared in sunlight, organisms show many adaptations to dissipate or to avoid the heat. The temperature of a large animal or plant may rise temporarily, but the heat is dissipated by transpiration in the plant and by perspiration in the animal. Submerged animals and plants are protected by the water, the temperature change of which depends upon the heat capacity of the particular body of water.

Treatment of biological materials with infrared radiations (780–1150 nm) either before or after x-ray treatment increases the rearrangements in chromosomes induced by the x-ray in tissues of plants and animals tested. The way in which the infrared radiations do this is unknown, but a comparable amount of conducted heat does not have the same effect.

Medical practitioners make use of infrared radiations to treat sprains, strains, bursitis, peripheral vascular diseases, arthritis, muscle pain, and many other pains for which heating gives relief, probably because of vasodilation of peripheral vessels. For this purpose, glow coil radiators are generally employed, but for some purposes, the more penetrating radiations (780–1400 nm) obtainable from incandescent sources are preferable to the glow coil to stimulate circulation deeper in the tissues. *See* BIOPHYSICS; THERMOTHERAPY. Arthur C. Giese

Bibliography. J. R. Cameron and J. G. Skofronick, *Medical Physics*, 1978; W. R. Hendee, *Medical Radiation Physics*, 2d ed., 1979.

Infrared spectroscopy

The spectroscopic study of the interaction of matter with infrared radiation. Electromagnetic waves from the long-wavelength limit of visible light at 800 nanometers to the shortest microwaves at 1 mm are used. In the wave-number units usually employed (oscillations per centimeter, read as reciprocal centimeters), this corresponds to 12,500–10 cm^{-1}. *See* INFRARED RADIATION.

Infrared regions. The infrared is conveniently divided into three regions, according to the different types of atomic and molecular energy levels that can be studied and the experimental techniques.

Radiation in the near-infrared region, at 4000–12,500 cm^{-1} (wavelengths of 2.5–0.8 micrometers), can excite harmonics of molecular vibrations as well as low-energy electronic transitions in molecules and crystals. The usual source is a quartz-envelope tungsten-halogen lamp, with detection by photomultipliers or photodiodes (for the shorter wavelengths) and lead sulfide (PbS) or indium antimonide (InSb) photoconductors. *See* INCANDESCENT LAMP; PHOTOELECTRIC DEVICES.

The mid-infrared region, at 200–4000 cm^{-1} (50–2.5 μm), is where most stretching and bending fundamental vibrations of molecules occur. Blackbody sources (Nernst glowers and globars) are used. Detectors are thermal (thermocouples, bolometers), pyroelectric (deuterated triglycine sulfate, DTGS), or photoconductive InSb, doped germanium, and mercury cadmium telluride (HgCdTe). *See* BOLOMETER; HEAT RADIATION; PHOTOCONDUCTIVITY; PYROELECTRICITY; THERMOCOUPLE.

The far-infrared region, at 10–200 cm^{-1} (1000–50 μm), is where rotational transitions in gaseous molecules; certain low-frequency bending, torsional, and skeletal vibrations; and lattice modes in solids are observed. Plasma emission from a high-pressure mercury arc is the usual source; DTGS and doped-germanium bolometers are useful detectors, as is the Golay pneumatic cell. For still lower frequencies, microwave techniques are used. *See* CRYSTAL ABSORPTION SPECTRA; MICROWAVE SPECTROSCOPY; MOLECULAR STRUCTURE AND SPECTRA; RADIOMETRY.

Development. Although infrared radiation was discovered in 1800, the development of spectroscopy was slow until the 1880s. By then the bolometer detector had been invented and solar spectra recorded as far as 10 μm. The recognition that molecular vibrational frequencies are characteristic of specific chemical groups was emerging, based on studies in the near-infrared. Investigations in the early 1900s of the infrared spectra of organic compounds demonstrated their usefulness for identification, and established correlations between molecular structure and mid-infrared frequencies. The invention of the echelette or blazed grating, which concentrates most of the diffracted radiation into a single order, was a major experimental advance.

After the rapid development of infrared and electronics technology during World War II, recording infrared spectrometers became available commercially, and were widely adopted by industrial analytical laboratories in the 1940s. Initially, these used prisms, but were soon followed by higher-performance grating instruments of increasing sophistication. The development and commercialization of Fourier-transform spectrometers and tunable infrared lasers have further expanded the capabilities of infrared spectroscopy.

Instrumentation and techniques. The broad wavelength range of infrared radiation, and the few transparent optical materials available, require that infrared instruments be designed with reflecting optics: radiation is focused with front-surface aluminized mirrors rather than lenses. Because of strong infrared absorption by water vapor and carbon dioxide, operation takes place in a vacuum or the optical path is purged with dry nitrogen. Absorption spectroscopy is the principal method, where attention centers on the frequencies absorbed by the sample from a broadly emitting source. However, spectrometers and interferometers can easily be adapted to emission spectroscopy.

Dispersive spectrometers. In **Fig. 1**, infrared radiation from the source Q is focused by a spherical mirror M_2 onto the entrance slit S_1 of the monochromator, after passing through the sample cell SC. The beam is collimated by the off-axis paraboloid mirror M_3, dispersed by refraction through the prism P, and focused in the plane of the exit slit S_2 by a second reflection from M_3. A narrow spectral region of the dispersed radiation passes the exit slit and is focused by the ellipsoidal mirror M_7 onto the detector D, which converts the radiant energy into an electrical signal. Since the beam has been chopped at constant frequency by a rotating mechanical chopper C, this signal is an alternating current that is amplified by the lock-in amplifier A, controlling the pen of the chart recorder R. To scan the spectrum, M_4 is rotated by the drive mechanism DM, which also drives the recorder. Successive frequencies are thus moved across the exit slit, producing a record of signal intensity as a function of mirror position; with a proper mechanical linkage, this record is made linear in wavelength or wave number. The same arrangement can be used with a diffraction grating as the dispersive element; the prism is removed and mirror M_4 is replaced by the grating.

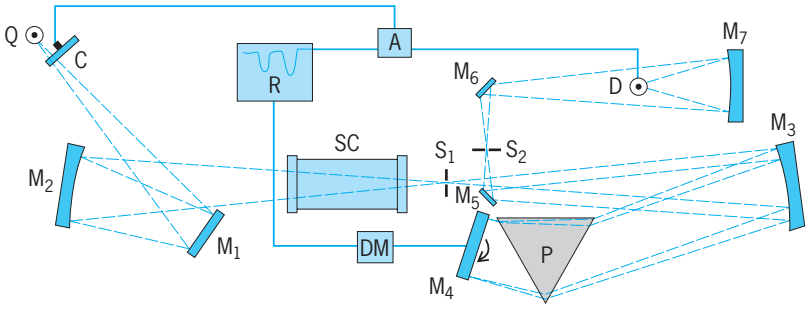

Fig. 1. Basic single-beam recording infrared prism spectrometer. The monochromator (the portion from the entrance slit S_1 to the exit slit S_2) is a Littrow mounting, a common arrangement for infrared instruments.

Even with an empty sample cell, the signal from such a single-beam instrument is affected by atmospheric absorption, variations in source output, absorption and scattering from the cell, and so forth. This is avoided with double-beam spectrometers, which use a beam-switching system of rotating mirrors to compare (typically some 10 times per second) the energy transmitted by the sample to that passing through a reference beam (which may contain an empty cell), and plot the ratio.

In the near-infrared, quartz prisms are used, and in the mid-infrared, alkali-metal halide or calcium fluoride (CaF_2) prisms, but no prism material is suitable beyond about 50 μm. Diffraction gratings can be used in all regions with the advantage, for equivalent optical configurations, of significantly higher resolving power than prisms. Prism instruments have resolutions of little better than 1 cm^{-1} near wavelengths of maximum dispersion, and much poorer than this elsewhere. Grating resolution can be several tenths of a reciprocal centimeter for commercial spectrometers; some specially built instruments can resolve a few hundredths of a reciprocal centimeter. In most laboratories these instruments are being replaced by other techniques, although inexpensive double-beam grating spectrometers are still manufactured. *See* DIFFRACTION GRATING; OPTICAL PRISM; RESOLVING POWER (OPTICS).

Fourier-transform spectroscopy. In a Michelson interferometer, light from the source strikes a thin, partially reflecting beam splitter at an angle of 45° and is divided into two beams that are returned by mirrors and recombined at the beam splitter. The intensity of the recombined and interfering beams, recorded as a function of the optical path difference (or retardation) as one mirror is moved, yields an interferogram. From this, the desired spectrum (that is, the source intensity as a function of wave number) can be recovered by the mathematical procedure known as a Fourier transform. *See* FOURIER SERIES AND TRANSFORMS; INTEGRAL TRANSFORM; INTERFEROMETRY.

The components of a Fourier-transform spectrometer (FTS) are shown in **Fig. 2**. Radiation from the source Q is chopped at C, collimated by mirror M_1, and divided at the beam splitter BS, with half sent to a fixed mirror M_2 and half to a movable mirror M_3. The recombined beams, with an optical-path difference $2L$, pass through a sample cell SC and are focused by M_4 onto the detector D. The signal is amplified (A) and stored in the digital computer DC. The interferogram is scanned by moving M_3 from zero path difference ($L = 0$) to some distance L_{max}; it is then Fourier-transformed and displayed as a spectrum on the recorder R. The spectral band pass of a Fourier-transform spectrometer is approximately $1/(2L_{max})$, and achieving a resolution of 0.1 cm^{-1}, for example, requires a mechanically very reasonable 5-cm mirror movement. At sufficiently high mirror velocities (about 1 mm/s for mid-infrared spectra), the signal is automatically modulated at audio frequencies and can be amplified directly, eliminating the chopper with its 50% signal loss. *See* DIGITAL COMPUTER.

Fourier-transform spectroscopy offers several advantages over dispersive methods; these are especially important for effective use of the limited radiant power available from most infrared sources. Whereas a spectrometer samples only one small frequency range at any given instant and must scan these frequencies sequentially to produce a spectrum, an interferometer processes information from all frequencies simultaneously (multiplex or Fellgett advantage). Furthermore, an interferometer passes a much greater light flux for a given resolving power than a spectrometer, which can accept only a very limited solid angle of source radiation because of the narrow slits required (throughput or Jacquinot advantage). These advantages can be translated into improvements of orders of magnitude in any one of the three interrelated important parameters of resolution, signal-to-noise ratio, and scan time. Another advantage is that the mirror movement is monitored to high precision with a fixed-frequency laser, so that the wave-number scale of the transformed spectrum is highly accurate compared with spectrometer readings.

Until the discovery of the Cooley-Tukey fast-Fourier-transform (FFT) algorithm in 1964, interferometric spectroscopy was largely impractical because of computational difficulties. The ready availability of inexpensive computers has now made it the accepted technique for high-performance infrared spectroscopy.

Commercial Fourier-transform spectrometers are marketed with resolutions from a few reciprocal centimeters for routine qualitative analyses to research instruments that can resolve better than 0.002 cm^{-1}. Typically, they operate from 4000 to 400 cm^{-1} on one beam splitter such as germanium-coated potassium bromide (KBr); this range can be extended broadly in either direction with different beam-splitter materials. For the far-infrared, Mylar films in various thicknesses are used. These instruments are controlled

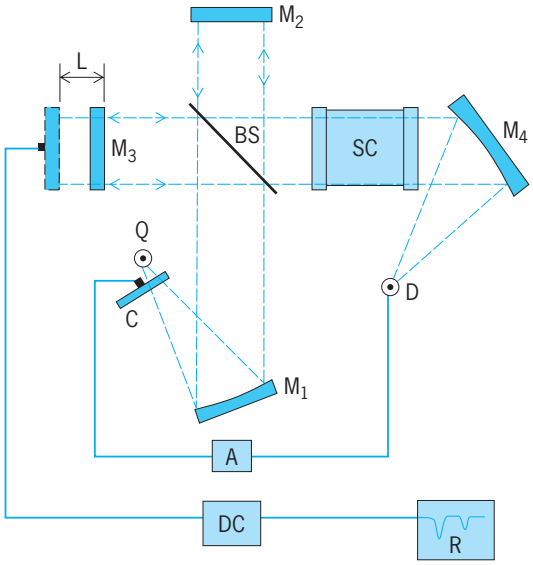

Fig. 2. Michelson interferometer used as a Fourier-transform spectrometer.

by microprocessors and include a digital computer to handle the Fourier transform. This computing power allows data manipulation such as repetitive scanning and signal averaging; background subtraction; spectral smoothing, fitting, and scaling; and searching digitized spectral libraries and databases to identify unknowns. Although backgrounds can be subtracted with software, some Fourier-transform instruments are designed for true optical double-beam operation. Many offer rapid-scan capability for the study of short-lived species and can be adapted to recording Fourier-transform Raman spectra. *See* MICROPROCESSOR; RAMAN EFFECT.

Tunable infrared lasers. In most experiments the lower limit to the linewidths of molecular transitions in gases is the Doppler broadening, the width caused by the translational (thermal) motion of the molecules. In the mid-infrared this is typically 10^{-2}–10^{-4} cm^{-1}; better resolution than even interferometers provide is desirable for spectroscopy of gases. Continuously tunable and nearly monochromatic laser sources, introduced about 1970, have met this need. *See* LASER; LASER SPECTROSCOPY.

Characteristics of some tunable lasers are listed in **Table 1.** Many of these emit beyond the ranges listed, which are limited to regions where useful high-resolution spectroscopy has actually been demonstrated. Other techniques, such as four-wave mixing, vibronic transition lasers, and polariton lasers, have been little used for spectroscopy.

Ideally, tunable infrared lasers require only a simple experimental arrangement: the source, a sample cell, and a detector with recorder. A low-resolution monochromator is often included to eliminate unwanted laser modes. In practice, many of these lasers continuously tune over very limited wavelength ranges, require constant adjustment, and tend to be complex and troublesome devices. However, they are indispensable when the highest resolution is required, and they have opened exciting new areas of infrared research.

If a tunable laser has enough power to saturate molecular transitions, as do carbon dioxide (CO_2) lasers, a technique known as saturation spectroscopy can be used to record spectra below the Doppler width, and thus exploit the full potential of these sources. This ultrahigh resolution can reveal a wealth of spectral detail that is otherwise lost in the Doppler broadening. *See* DOPPLER EFFECT.

Wave-number calibration. The accuracy of low-resolution instruments can be quickly checked by recording the spectra of standards such as polystyrene film or a solution of camphor and cyclohexanone in indene; these provide sharp absorption peaks with positions known within ± 0.5 cm^{-1}. For high-resolution spectra, tabulations are available giving the frequencies of thousands of rotation-vibration lines of light gaseous molecules such as hydrogen chloride (HCl), carbon monoxide (CO), water vapor (H_2O), hydrogen cyanide (HCN), carbon dioxide (CO_2), nitrous oxide (N_2O), ammonia (NH_3), and ethylene (C_2H_2), with accuracies of ± 0.005 cm^{-1} or better.

Sample handling. Gases are examined in simple cylindrical cells with infrared-transmitting windows; a typical path length is 10 cm. For trace species or weak absorptions, compact multiple-reflection White cells (named after their inventor) are made with adjustable optical path lengths up to 100 m (300 ft) and more. Liquid cells are available for sample thicknesses of 0.01 mm to several centimeters. Solids can be observed directly in thin sections; as mulls in mineral or fluorocarbon oils; as finely ground dispersions in disks of an infrared-transparent salt such as potassium bromide; or in solutions of solvents such as carbon tetrachloride (CCl_4) or carbon disulfide (CS_2) that have few infrared bands. (The usefulness of water is limited by its strong, broad infrared absorptions.)

Many accessories have been designed for sampling problems. Infrared microscopes with reflecting optics can focus on objects as small as 10 μm to record the spectra of subnanogram samples; this is useful in particulate and fiber analysis and forensic work. Attenuated total reflection, where the infrared beam is transmitted inside a thin, highly refractive crystal [such as zinc selenide (ZnSe) or germanium] by multiple internal reflections, and is

TABLE 1. Tunable mid-infrared laser sources*

Device	Demonstrated spectral coverage, μm	Maximum resolution, cm^{-1}	Typical continuous-wave power, W
Semiconductor diode lasers	3–27	2×10^{-6}	10^{-3}
Gas lasers			
Waveguide CO_2	9–11†	3×10^{-7}	1
Zeeman-tuned neon or xenon	~3.5†	1×10^{-3}	10^{-3}
	7–8†	1×10^{-3}	?
Spin-flip Raman lasers	5–6	1×10^{-6}	0.1
Color-center lasers	1.5–4	3×10^{-8}	10^{-2}
Optical parametric oscillators	1–16	$<1 \times 10^{-3}$	10^{-3}
Nonlinear optical mixing techniques			
Difference-frequency generation	2–6	1×10^{-4}	10^{-6}
Tunable-sideband generation	~ 3.4†	2×10^{-10}	?
	9–11†	3×10^{-6}	10^{-3}

*Adapted from R. S. McDowell, Infrared laser spectroscopy, *Proc. SPIE*, 380:494–507, 1983.
†Tunable only near discrete lines in these regions.

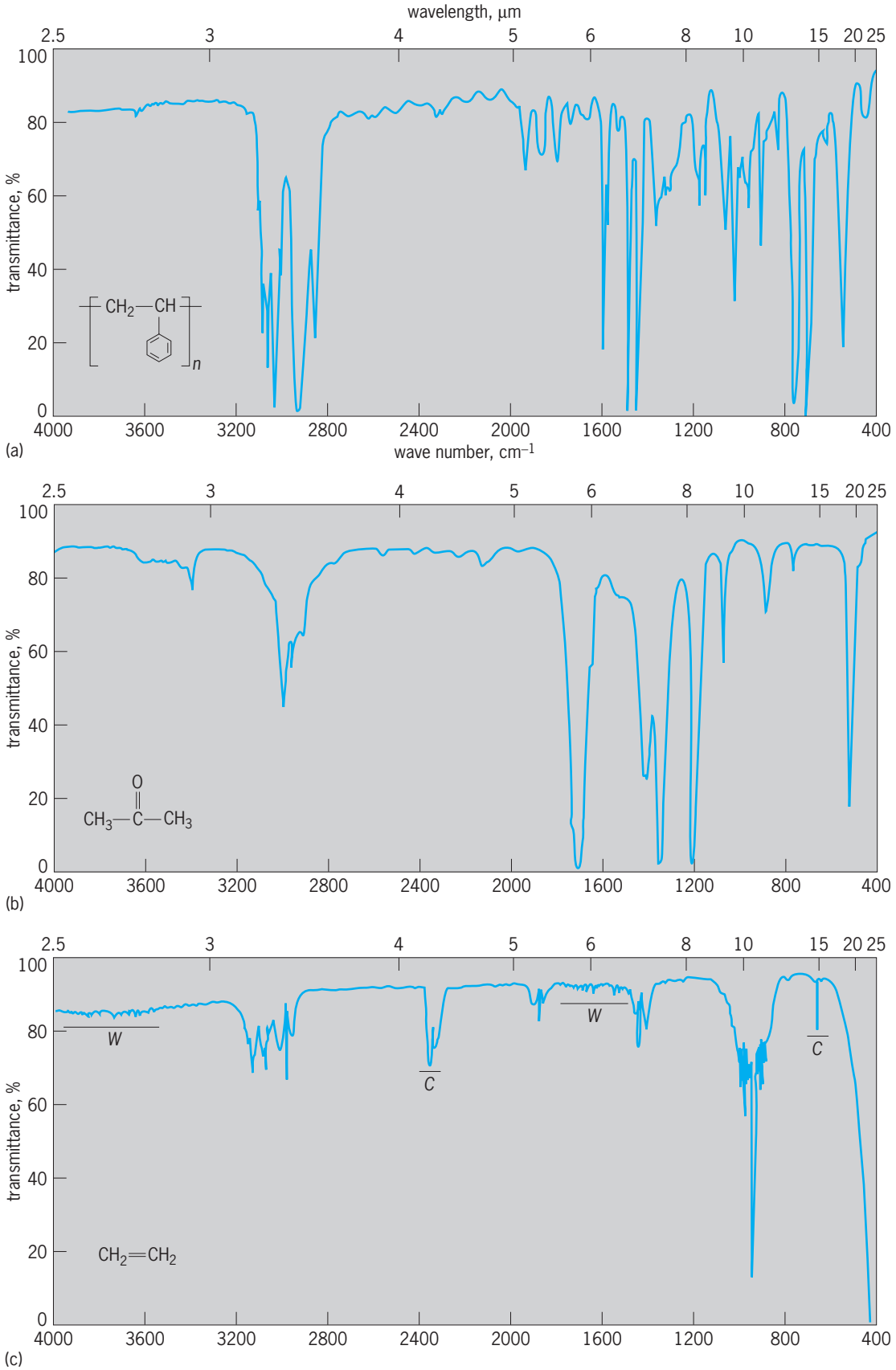

Fig. 3. Mid-infrared absorption spectra recorded on a Fourier-transform spectrometer at a resolution of 1 cm^{-1}. (*a*) Polystyrene film, 0.05 mm thick. The weak intensity fluctuations between 2000 and 2800 cm^{-1} are interference fringes, caused by reflections from the surfaces of the plastic, and can be used to determine the film thickness. (*b*) Liquid acetone, capillary film between cesium iodide (CsI) plates. (*c*) Ethylene gas at about 30 torr (4 kilopascals) in a 10-cm (4-in.) cell. The absorptions marked W and C are due to traces of water vapor and carbon dioxide (CO_2), respectively. Below 600 cm^{-1} the potassium chloride (KCl) windows absorb. (*V. T. Hamilton, Los Alamos National Laboratory*)

absorbed by a sample in optical contact with the crystal, may be appropriate for difficult materials such as gels and strongly absorbing liquids. Photoacoustic detectors can be used for nearly opaque samples. Other accessories provide spectra of specularly reflected light from coatings and films; of diffuse reflectance from inhomogeneous samples such as powders; of samples under high pressure in diamond-anvil cells; and so forth. Many devices are available for holding samples at cryogenic or elevated temperatures. *See* HIGH-PRESSURE PHYSICS; PHOTOACOUSTIC SPECTROSCOPY.

Typical spectra. Infrared spectra are usually plotted as percent transmittance T or absorbance A on a scale linear in wave number ν (less commonly, in wavelength λ). Transmittance is the ratio of the intensity of radiation transmitted by the sample (I) to that incident on the sample (I_0), expressed as a percentage, so that $T = 100I/I_0$; this is related to absorbance by Eq. (1). Transmittance and absorbance,

$$A = \log_{10}\left(\frac{I_0}{I}\right) = \log_{10}\left(\frac{100}{T}\right) \qquad (1)$$

of course, vary with frequency and should formally be expressed as $T(\nu)$ and $A(\nu)$.

Examples of mid-infrared spectra recorded on a commercial instrument are shown in **Fig. 3.** The very strong absorption in Fig. 3*b* (liquid acetone) at 1705–1725 cm^{-1} is characteristic of the carbonyl C═O stretch in saturated aliphatic ketones. The rotational lines in Fig. 3*c* (ethylene gas) between 850 and 1050 cm^{-1} arise from quantized changes in the amount of rotational excitation that occurs simultaneously with the ═CH_2 out-of-plane "wagging" vibrational transition at 949 cm^{-1}. A similar structure is seen on the C—H stretch between 3050 and 3200 cm^{-1}, and would be evident in all bands at higher resolution. In liquids and solids, this rotational structure is suppressed.

A single rotation–vibration band of a gaseous molecule, the infrared-active stretch (designated ν_3) of sulfur hexafluoride (SF_6), is shown in **Fig. 4** at four resolutions from 2 to 10^{-6} cm^{-1}, to illustrate the additional detail revealed at higher resolution. The emission linewidth of the tunable semiconductor diode laser used to record Fig. 4*c* was less than 10^{-5} cm^{-1}, but the effective resolution is the sulfur hexafluoride Doppler width of 0.0010 cm^{-1}. Figure 4*d* is a sub-Doppler saturation spectrum taken inside the linewidth of a carbon dioxide (CO_2) laser line at an effective resolution of less than 10^{-6} cm^{-1}. This was recorded in the derivative mode, where the ordinate is $dA/d\nu$, the derivative of A with respect to ν, instead of A or T. Many thousands of features like those in Fig. 4*c* and *d* have been assigned to specific sulfur hexafluoride transitions.

Applications. Among the more important applications of infrared spectroscopy are chemical analysis and the determination of molecular structures.

Qualitative analysis. Infrared spectra are ideal for identifying chemical compounds because every molecule [except homonuclear diatomics such as nitrogen (N_2), oxygen (O_2), and chlorine (Cl_2)] has an infrared spectrum. Since the vibrational frequencies depend upon the masses of the constituent atoms and the strengths and geometry of the chemical bonds, the spectrum of every molecule is unique (except for optical isomers). Pure unknowns can thus be identified by comparing their spectra with recorded spectra; catalogs are available in digitized versions, and searches can be made rapidly by computer.

Simple mixtures can be identified with the help of computer software that subtracts the spectrum of a pure compound from that of the unknown mixture. More complex mixtures may require fractionation

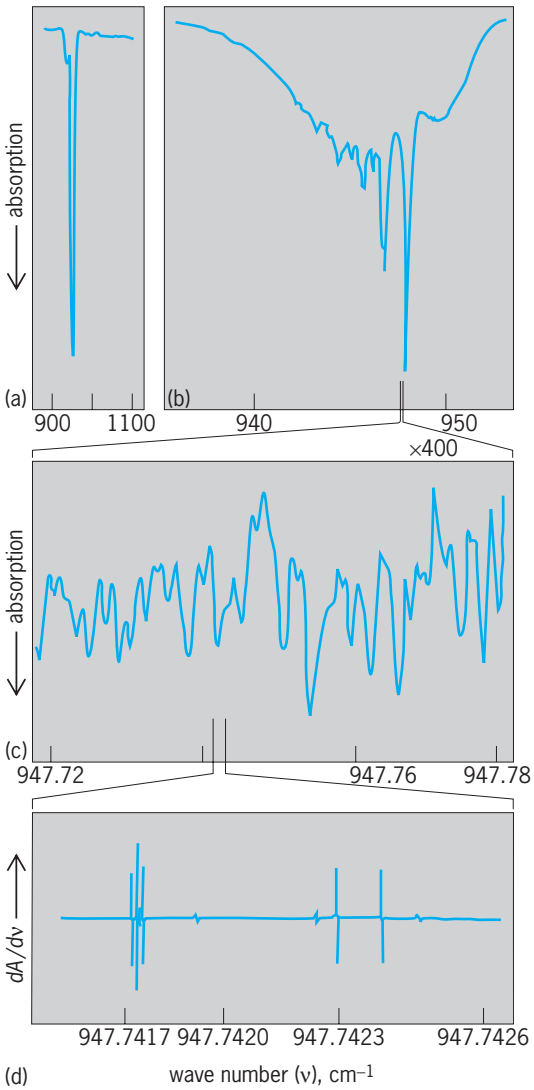

Fig. 4. S—F stretching fundamental ν_3 of sulfur hexafluoride (SF_6) as it appears with increasing resolving power. (*a*) Fourier-transform spectrum at resolution of 2 cm^{-1}. (*b*) Fourier-transform spectrum at resolution of 0.06 cm^{-1}. (*c*) Expansion of small portion of the strong central peak in part *b*, recorded with a tunable semiconductor diode laser at an effective resolution of 0.0010 cm^{-1}. (*d*) Further resolution of the absorption line at 947.742 cm^{-1} into five strong components with sub-Doppler saturation spectroscopy at an effective resolution of $<10^{-6}$ cm^{-2}. (*After R. S. McDowell, in J. R. Durig, ed., Vibrational Spectra and Structure, vol. 10, pp. 1–151, 1981*)

TABLE 2. Group frequencies for O—H, N—H, and C—H stretches*

Functional group	Frequency, cm^{-1}
Alcohol and phenol —OH	3593–3643 (medium)
Carboxylic acid —CO—OH	3500–3560 (medium)
Most NH	3310–3540 (medium)
Hydrazine $>N—NH_2$	3139–3327, 3354–3427 (medium)
Amino acid $—NH_3^+$	3030–3130 (medium)
Alkane $—CH_3$ (methyl)	2872 ± 10, 2962 ± 10 (strong)
Alkane $>CH_2$ (methylene)	2853 ± 10, 2926 ± 10 (strong)
Alkane —CH (methylidyne)	2890 ± 10 (weak)
Alkene $—CH{=}CH_2$ (vinyl)	3010–3040, 3075–3095 (medium)
Alkene $>C{=}CH_2$ (vinylidene)	3075–3095 (medium)
Alkene —CH = (methine)	3010–3040 (medium)
Alkyne —C≡CH (ethynyl)	3315 5 (medium)
Aldehyde —CH=O (formyl)	2700–2900 (weak; usually 2)
Aromatic CH	~3030 (medium), ~3070 (weak)
Heterocyclic (azine) CH	3010–3050, 3060–3090 (weak)

*Hydrogen bonding lowers the OH and NH frequencies by 100–1000 cm^{-1} from these values, which apply to free molecules.
SOURCE: After L. J. Bellamy, *The Infra-Red Spectra of Complex Molecules,* vol, 1, 3d ed., 1975.

first. This has led to the development of combinations of analytical techniques, such as gas chromatography used together with Fourier-transform infrared spectroscopy; instruments that combine these functions are available commercially.

Many functional groups have characteristic infrared frequencies (**Table 2**) that are relatively independent of the molecular environment. Often, specific conclusions can be drawn from frequencies, and it may be possible to identify even a new compound from its spectrum. Group frequencies are most useful above about 1500 cm^{-1}; below this the absorptions are due to, or are influenced more by, the skeletal vibrations of the molecule. This is the "fingerprint" region, where even similar molecules may have quite different spectra.

With the help of Table 2 and the spectrum of Fig. 3*a*, the absorptions at 2851 and 2924 cm^{-1} can be immediately assigned to methylene. That at 3027 cm^{-1} could be alkene, heterocyclic, or aromatic C—H; a more extensive table of group frequencies shows that the last is correct, from the phenyl C—C ring "breathing" stretches at 1400–1650 cm^{-1} and out-of-plane C—H bends at 600–800 cm^{-1}. The latter bands are characteristic of monosubstitution, as is the pattern of weak combinations between 1650 and 2000 cm^{-1}. There is no indication of other functional groups. This evidence suggests polystyrene, and the identification would be confirmed by comparison with a known spectrum. *See* QUALITATIVE CHEMICAL ANALYSIS.

Quantitative analysis. Quantitative infrared analysis is possible because the absorbance defined in Eq. (1) is directly proportional to the number of absorbing molecules, and in mixtures is the sum of the absorbances of the individual components at the frequency of interest. The Bouguer-Lambert-Beer law is given in Eq. (2), where $a(\nu)$ b, and c are the absorp-

$$A(\nu) = a(\nu)\, bc \qquad (2)$$

tivity (a frequency-dependent molecular property), path length, and concentration, respectively. When the absorptivity of a compound is known at a given frequency, for example, at an absorption peak, the concentration of that compound in a mixture can be determined from Eq. (2). Mixtures with n components can be analyzed by measuring absorbances at n selected frequencies and solving the resulting system of simultaneous linear equations.

Traditional infrared analysis, although highly specific, is not notably sensitive: typical detection limits are 0.1–1% with precisions of a few percent of the quantity measured, except in especially favorable cases. These limits improve with increasing resolution, until the resolution becomes somewhat less than the width of the feature being measured. Tunable-laser sources offer much greater sensitivity because of their narrow linewidths and easy adaptability to first- and second-derivative detection (see also Fig. 4*d*). Lasers are used for concentration measurements of trace atmospheric gases and pollutants; detection limits in the parts-per-trillion range have been reported. *See* QUANTITATIVE CHEMICAL ANALYSIS.

Molecular structure. Many details of molecular structure and dynamics can be extracted from an infrared spectrum, especially for light molecules that can be examined in the gas phase and therefore exhibit rotational structure. The near-infrared absorption of acetylene (H—C≡C—H) in **Fig. 5** provides an example. This is a combination band, in which two fundamentals are excited, ν_1 (symmetric C—H stretch) at 3373 cm^{-1} and ν_5 (bend) at 730 cm^{-1}. The difference between the band center at 4091 cm^{-1} (indicated by A in Fig. 5) and the sum of $3373 + 730 = 4103$ cm^{-1} is a measure of the anharmonicity of the vibrations involved.

The 3:1 intensity alternation in the rotational lines (B) is due to nuclear spin statistics, and indicates that acetylene is linear and symmetric. This molecular symmetry can also be deduced from the number of fundamental vibrations in the infrared and Raman spectra, and the number of infrared–Raman coincidences. From the spacing of the rotational lines (C), the molecular moment of inertia can be obtained; this and the moment of the deuterated species yield the bond lengths, 0.106 nm for C—H and 0.121 nm for C≡C. The small change in the rotational line spacing across the band is due to vibration–rotation

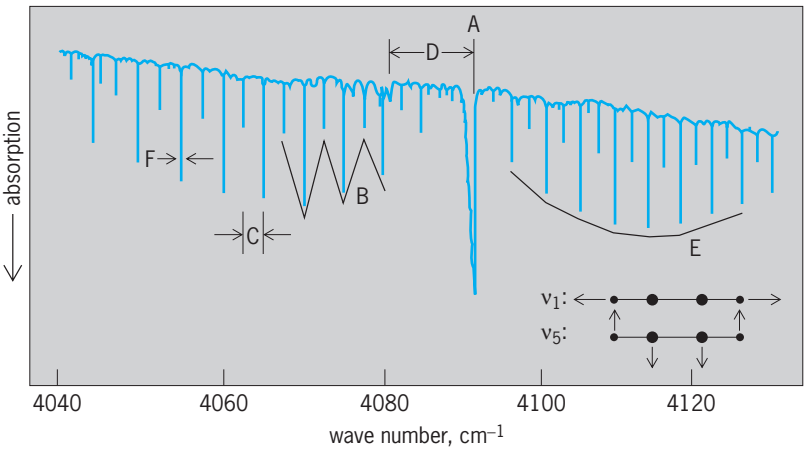

Fig. 5. Spectrum of the $\nu_1 + \nu_5$ rotation–vibration band of acetylene recorded with a grating spectrometer, resolution 0.05 cm^{-1}; gas at 25°C (77°F) and 100-torr (13.3-kPa) pressure, 25-cm (10-in.) path length. The letters indicate features that are discussed in the text. (*After A. R. H. Cole, Tables of Wavenumbers for the Calibration of Infrared Spectrometers, 2d ed., Pergamon Press, 1977*)

interaction, that is, to the fact that the bond lengths differ slightly in the ground and excited vibrational states.

The weak absorption at 4080 cm^{-1} is the center of the corresponding band of the isotopic species $H^{12}C^{13}CH$, present in its natural abundance of 2.2%. From this frequency shift (D), and the shifts for the deuterated species, the molecular force constants can be calculated, which are measures of the resistance of the chemical bonds to deformations. For acetylene, the stretching force constants are 0.64 micronewton/nm (6.4 millidyne/Å) for the C—H bond and 1.63 μN/nm (16.3 mdyn/Å) for C≡C, which are typical values for single and triple bonds, respectively. Knowledge of the bond strengths helps in understanding the electronic structure.

Finally, the spectrum provides information on the conditions of the sample. The overall intensity contour of the rotational lines (E) is characteristic of the temperature of the gas; the individual linewidths (F) increase with pressure over their minimum (Doppler) value. Therefore, good estimates of temperature and pressure can be made if these are not otherwise known. The usefulness of such information in remote sensing is obvious.

Other applications. Other applications include calculation of thermodynamic quantities from vibrational and rotational energy levels; studies of intermolecular forces in condensed phases; distinguishing between free or hindered internal rotation (as of methyl groups); quantitative intensity measurements to obtain bond dipole moments; studies of molecular interactions in adsorption, surface chemistry, and catalytic processes; time-resolved monitoring of transient species and chemical reaction kinetics; characterization of reactive molecules isolated at cryogenic temperatures in rare-gas lattices (matrix isolation); electronic energy states in semiconductors and superconductors; studies of biological molecules and membranes (effects of hydrogen bonding); and analysis of energy levels in laser systems. Double-resonance techniques, in which molecules are excited by an intense pulse of ultraviolet, infrared, or microwave energy, and simultaneously probed by infrared or microwave spectroscopy, are useful in monitoring population changes in molecular energy levels, energy transfer and relaxation, and multiphoton absorption processes. *See* ADSORPTION; CATALYSIS; CHEMICAL DYNAMICS; HYDROGEN BOND; INTERMOLECULAR FORCES; SEMICONDUCTOR; SPECTROSCOPY; SUPERCONDUCTIVITY. Robin S. McDowell

Bibliography. J. Chalmers and P. Griffiths (eds.), *Handbook of Vibrational Spectroscopy*, 5 vols., 2002; A. A. Chirstty, Y. Ozaki, and V. G. Gregoriou, *Modern Fourier Transform Infrared Spectroscopy*, 2001; N. B. Colthup, L. H. Daly, and S. E. Wiberley, *Introduction to Infrared and Raman Spectroscopy*, 3d ed., 1990; H.-U. Gremlich and B. Yan, *Infrared and Raman Spectroscopy of Biological Materials*, 2000; B. Schrader, *Infrared and Raman Spectroscopy: Methods and Applications*, 1994; J. I. Steinfeld, *Molecules and Radiation: An Introduction to Modern Molecular Spectroscopy*, 2d ed., 2005; B. H. Stuart, *Infrared Spectroscopy: Fundamentals and Applications*, 2004.

Infrasound

Low-frequency sound waves that cannot be sensed by the human ear. Infrasound was first detected on barometers around the world following the eruption of the volcano Krakatoa in 1883. We now know subaudible sounds are generated by a wide variety of natural phenomena, such as large storms and meteors, and some human-produced sources, such as large chemical or nuclear explosions. Infrasound is particularly useful for basic research, for example, in the study of atmosphere phenomena; as it decays slowly across thousands of kilometers. How infrasound is generated, how it propagates through our turbulent atmosphere, and how it can be recorded amid noise due to air circulation are the subjects of ongoing research.

Origin. Humans can hear sounds from 20 to 20,000 Hz, in the middle of the acoustic spectrum. Infrasonic signals (below 20 Hz) are subaudible. Infrasound is generated by myriad sources, either impulsive, that is, relatively short-lived, or ongoing. To generate infrasound, these sources must be energetically significant and move a large volume of air. Some sources, such as hurricanes and earthquakes, are very large. Earthquakes compress the air and excite infrasound waves by vibrating large areas of the Earth's surface in an up-down pistonlike fashion or by shaking nearby topography. Volcanoes, tornadoes, avalanches, surf, aurora, sprites, projectiles (such as supersonic aircraft, rockets, and meteors), and explosive events (such as atmospheric nuclear tests and surface mine blasts) also generate infrasound waves. Some large animals, including elephants, rhinoceros, and whales, are known to use infrasound to communicate over distances of many kilometers.

The frequency of infrasound waves scales with the size of the source, with the largest sources generating the lowest-frequency signals (**Fig. 1**). Thunder and sonic booms, for example, generate

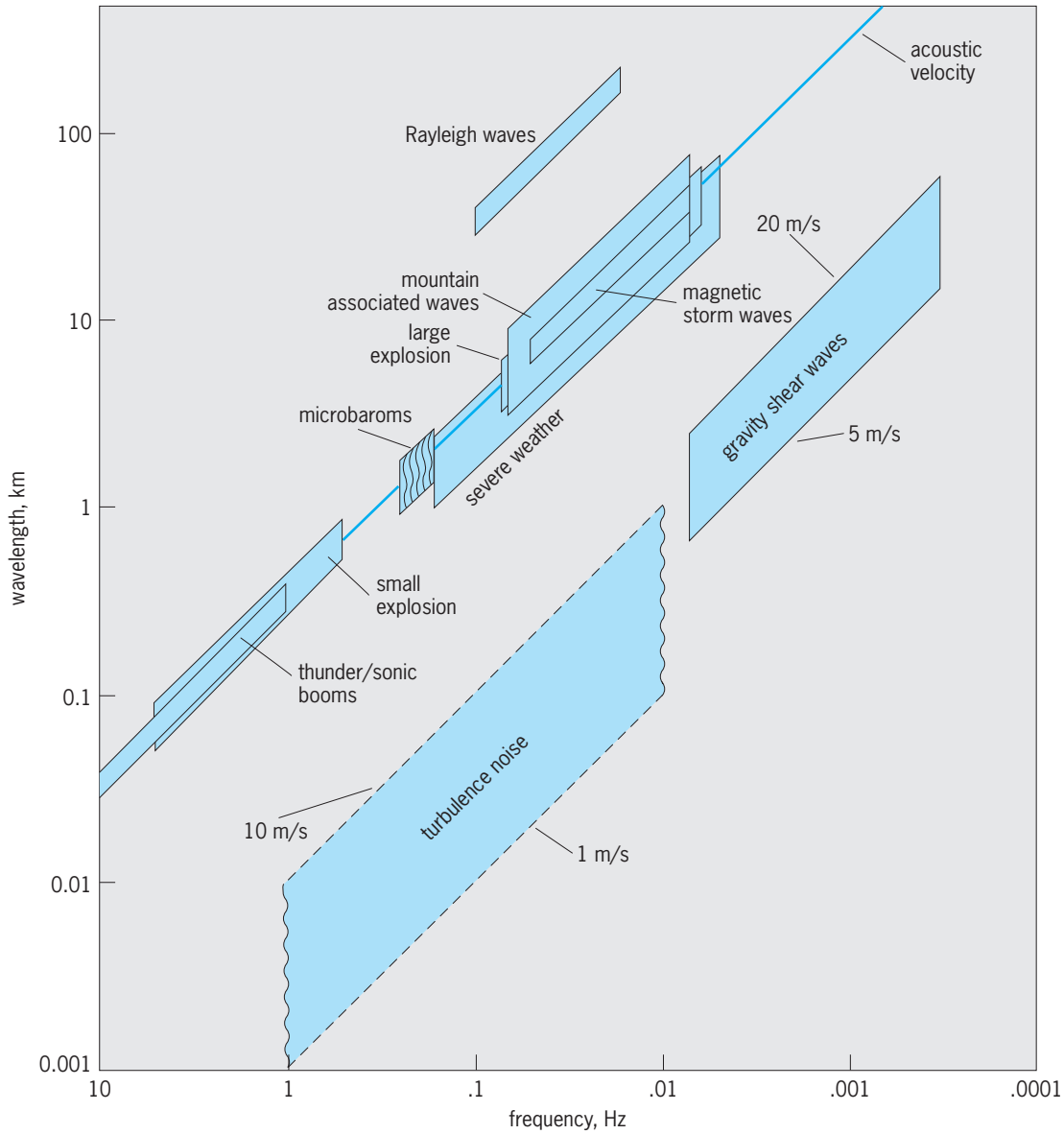

Fig. 1. Wavelength of subaudible acoustic waves as a function of frequency. Turbulent noise exists at the same frequency as acoustic signals, but at much shorter spatial scales. (*Adapted from Alfred Bedard with permission*)

higher-frequency infrasound waves. Lower-frequency infrasound is generated by larger sources, including large explosions and severe storms. Very low frequency infrasound is due to turbulence from air flowing over mountains. Mountain-associated waves are due to turbulence from air flowing over mountains. Earth's gravity becomes important at very low frequencies (below ~0.01 Hz). *See* TURBULENT FLOW.

Propagation. Infrasound is a mechanical wave phenomenon similar to seismic waves that travel through the solid Earth. The propagation of infrasound waves, however, differs significantly from seismic waves in that the infrasound velocity structure of the atmosphere changes rapidly. As with any wave phenomenon, infrasound refracts, or changes direction, due to changes in the infrasound velocity of the medium through which it passes. Infrasound refracts away from regions of high velocity. The velocity of infrasound is directly proportional to the square root of the absolute temperature of air, and increases in the direction of wind due to advection. Both temperature and wind depend on altitude, geographic location, time of day, and time of year. *See* REFRACTION OF WAVES; SEISMOLOGY.

To be useful for research, infrasound must return to sensors on the ground. With favorable winds, infrasound is efficiently refracted back to the Earth's surface from within the stratosphere (**Fig. 2**). Infrasound returns to the Earth's surface from within the thermosphere due to the very significant increase in temperature with increasing altitude. Temperature decreases with increasing altitude in the troposphere, the lowest atmospheric layer; however, infrasound can be refracted back to the ground from within this layer in the presence of strong temperature inversions or if the infrasound propagates in the direction of strong winds. *See* ATMOSPHERE; ATMOSPHERIC ACOUSTICS.

Infrasound is valuable for monitoring nuclear

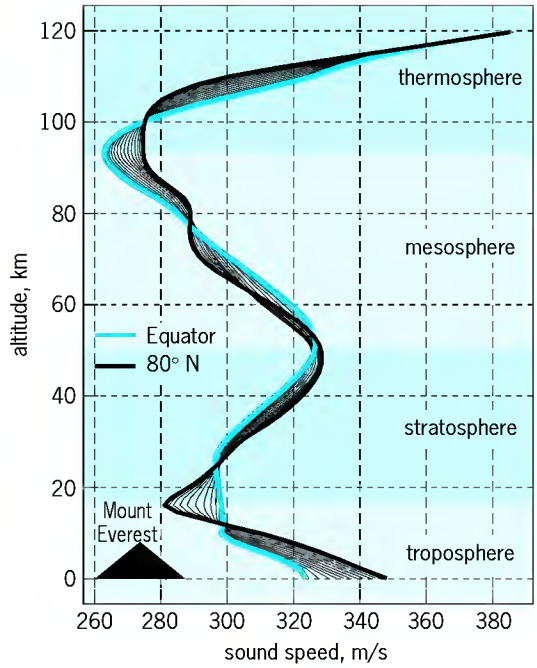

Fig. 2. Annual mean sound speed (Equator to 80° N). The speed of sound tracks air temperature and varies with altitude. Sound speed decreases with increasing altitude in the troposphere as the air temperature decreases away from Earth's surface. Sound speed increases in the stratosphere, decreases in the mesosphere, and increases sharply with increasing altitude in the thermosphere. Sound waves refract away from regions where the sound velocity is high, and thus turn back to the Earth's surface within the thermosphere and, with favorable winds, from within the stratosphere. Waves ducted within the troposphere have been observed and are believed to be due to temperature inversions or unusually strong winds in the direction of propagation of the sound waves.

testing and other atmospheric phenomena, as it propagates through the atmosphere with relatively little dissipation. The rate at which energy is lost scales with frequency, with the highest frequencies attenuated relatively rapidly. Infrasound waves above and below 1 Hz are termed near- and far-infrasound, respectively.

The use of infrasound for the study of atmospheric phenomena requires that the dynamic structure of the atmosphere be taken into account. Atmospheric models based on climatologies, numerical weather prediction models, and pinpoint measurements, such as radiosondes, are used to synthesize the propagation of infrasound.

Detection. It is possible to feel infrasound waves near a particularly intense source, such as a jet airplane on take-off. The eruption of Krakatoa generated the loudest sound in recorded history. Infrasound waves from this cataclysmic event registered several times on barometers worldwide due to repeated circuits around the globe. This was highly unusual, as infrasound waves, in general, are very faint and can be detected only with sensitive instruments or indirectly through the effect on structures.

The ability to detect infrasound waves is limited by noise from atmospheric turbulence. Most infrasonic noise in the frequency band of interest to infrasound researchers is coherent across distances of less than several meters, while infrasonic signals can be coherent over distances in excess of 100 m (330 ft; Fig. 1). The ratio of coherent signal to incoherent noise can be increased by averaging the infrasonic pressure field over an area smaller than the coherence limit of signals but greater than the coherence limit of noise. The most common method used to incoherently sum atmospheric noise while coherently summing the signal is to sample atmospheric pressure at many points distributed spatially (across an area ranging up to 100 m), transmitting the sound through an array of pipes to a microbarometer where they are summed acoustically. *See* ACOUSTIC NOISE; COHERENCE.

Some of the drawbacks resulting from the mechanical reduction of noise can be avoided by integrating sound waves at the speed of light. One such system under development uses fiber-optic cables for this purpose. Another approach is to record infrasound using a spatial grid of instruments, and then sum the data from the instruments optimally to reject noise while preserving coherent signals. *See* OPTICAL FIBERS.

Global monitoring network. A global infrasound network under construction (as of mid-2006) will comprise 60 continuously operating arrays (**Fig. 3**). The network is designed to provide uniform coverage of the Earth's atmosphere, given the constraint that all arrays must be located on land. Each array in the network will comprise four to eight microbarometers with apertures of 1–3 km (0.6–1.8 mi). Each microbarometer in an array is equipped with a noise-reducing filter averaging noise due to turbulence at spatial scales up to about 70 m (230 ft). The use of arrays of sensors, rather than single instruments, allows the detection of more faint infrasound signals and estimation of the speed and direction of the incident waves. The network is designed for optimal detection of signals between 0.1 and 1.0 Hz.

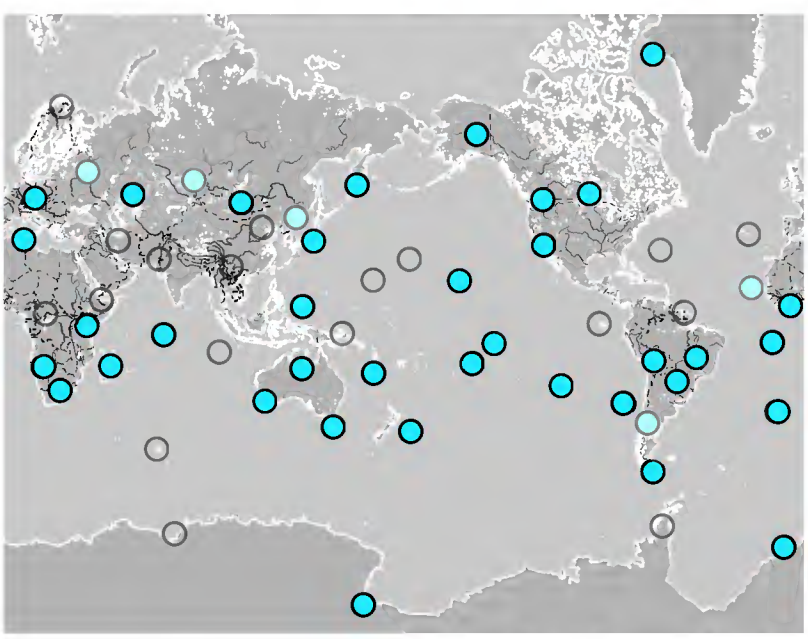

Fig. 3. Global infrasound network. Of the 60 planned operating stations, over half are built and recording data (dark color circles), several are under construction (light color circles), and a third have not yet been built (open circles).

Array processing. Array processing methods take advantage of the coherence, or similarity, of a signal recorded at different elements in an array to provide information about the signal, such as the velocity at which the signal crossed the array. The velocity can be used to infer the direction to the source and the path that the signal took through the atmosphere. The most commonly used method in the infrasound field is based on the progressive multichannel correlation (PMCC) algorithm. The PMCC method is based on the cross-correlation of data from different elements in an array. Cross-correlation is a mathematical measure of similarity of two segments of data. If two elements in an array record the same signal, the cross-correlation procedure will provide an estimate of how similar the recordings are, and how much the signal is shifted in time between the two recordings due to the propagation delay. The PMCC method extends this idea to subarrays of several elements and defines a closure relation as a more demanding test of whether a signal is present in the data. *See* SIGNAL PROCESSING; STATISTICS.

Consider a subarray of three elements (i, j, k). The PMCC method defines signal consistency (r_{ijk}) for the subarray as

$$r_{ijk} = \Delta t_{ij} + \Delta t_{jk} + \Delta t_{ki}$$

where Δt_{ij} is the time delay between the arrival of a signal at sensors i and j. A signal is considered consistent if the sum of the time delays is below a predefined threshold. The PMCC method applies this test to several small subarrays within an array. If a highly correlated and consistent signal is found, the test is repeated for progressively larger subarrays to refine the estimate of the velocity of the signal across the array. This procedure is repeated in several frequency bands and in numerous overlapping windows spanning the time of interest. The use of small subarrays avoids ambiguity problems inherent in correlating distant signals.

A single array recording can be a very accurate estimate of the direction to the source of the signal. The velocity of the signal across the array can be used to infer the path through the atmosphere taken by the infrasound; however, pinpointing the location of the source requires recordings made by several arrays.

Continuing basic research. In monitoring the atmosphere continuously, the infrasound network offers an unprecedented opportunity to better understand human-produced and natural atmospheric phenomena on a global scale. Research continues on how sound propagates through our unsteady atmosphere and how clear recordings of distant events can be made despite noise due to atmospheric turbulence. Researchers are improving models of the atmosphere and are collecting information about significant atmospheric sources to provide a basis for this research. *See* SOUND. Michael A. H. Hedlin

Bibliography. A. J. Bedard and T. M. Georges, Atmospheric infrasound, *Phys. Today*, March, pp. 32–37, 2000; M. A. H. Hedlin et al., Listening to the secret sounds of Earth's atmosphere, *Eos Trans. Amer. Geophys. Union*, 83:557, 564–565.

Inhibitor (chemistry)

A substance which is capable of stopping or retarding a chemical reaction. To be technically useful, such compounds must be effective in low concentrations, usually under 1%. The type of reaction which is most easily inhibited is the free-radical chain reaction. The study of inhibitor action is often used as a diagnostic test for free-radical chain character of a reaction. Vinyl polymerization and autoxidation are two important examples of the class. Another reaction type for which inhibitors have been found is corrosion, particularly in aqueous systems. The economic importance of corrosion inhibition can scarcely be overestimated. An understanding of inhibitor action depends on an understanding of the processes to be interrupted.

Inhibition of vinyl polymerization. This type of inhibitor action must be considered in terms of the accepted mechanism for the polymerization process, which may be summarized as reaction sequence (1).

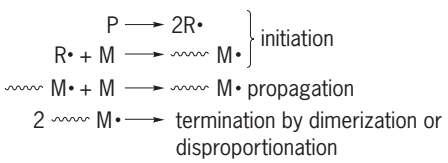

(1)

The symbol P represents a catalyst, often a peroxide, R· is a free radical derived from the catalyst, M is a monomer, and ∼M· is a growing polymer chain. The polymerization will be stopped or inhibited if some added substance (an inhibitor) reacts more readily than does the monomer with R· to yield a product which will not sustain the polymerization. Every reaction chain is stopped until the inhibitor is consumed. If the added substance (a retarder) is somewhat less reactive, the monomer can compete more successfully for the initiating radicals, so that the result is retardation rather than total inhibition. The difference between inhibition and retardation is one of degree rather than of kind.

Phenolic compounds and quinones. These interact with an initiating radical or a growing polymer chain either by hydrogen atom abstraction or by radical addition to an unsaturated linkage. These interactions are represented by sequence (2).

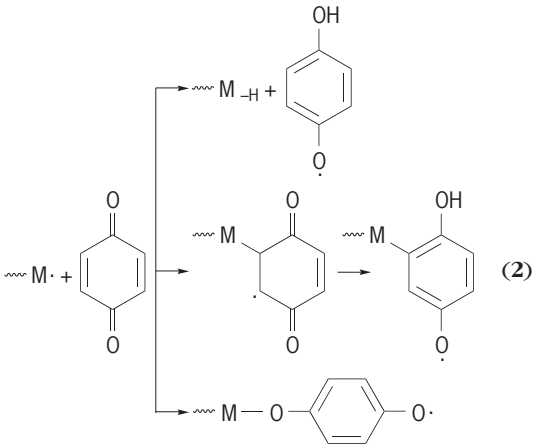

(2)

The phenoxy radicals produced are stabilized by

resonance and are not sufficiently reactive to add to the vinyl linkage of another monomer molecule. The usual fate of these radicals is further hydrogen atom loss by reaction with a second polymer radical or by disproportionation with another phenoxy radical to yield a quinone and a hydroquinone, both of which may continue to act as inhibitors. All the reaction possibilities shown have been demonstrated experimentally. Although the efficiencies of phenols and quinones for the interruption of polymerizations vary with the structures of these molecules, they may be classed as inhibitors. Aromatic amines react in a similar manner.

Nitroaromatics. As typified by trinitrobenzene, nitroaromatics function as retarders rather than as inhibitors in most polymerizations. It is necessary, however, to consider the specific reaction involved. Thus polynitroaromatics inhibit the polymerization of vinyl acetate, retard that of styrene, and have no effect on that of methyl methacrylate. No clear-cut mechanism has been established for the interaction of nitro compounds with free radicals.

Monomers. Both monomers and the radicals derived from them differ greatly among themselves in reactivity. Thus, although certain monomers may copolymerize with one another, others may actually function as inhibitors. Styrene and vinyl acetate, for example, both polymerize well when alone. Styrene, however, inhibits the polymerization of vinyl acetate. This occurs because the vinyl acetate radical and the styrene monomer are highly reactive, whereas the styrene radical and the vinyl acetate monomer are not. A small amount of styrene added to vinyl acetate will rapidly react with any vinyl acetate radicals formed when polymerization is initiated. The resulting styrenelike radical will react only very slowly with the vinyl acetate monomer. In the overall process the chain-carrying radical is converted to one which is too unreactive to carry on the chain.

Autoinhibition. This action, sometimes called allylic termination, is exhibited by monomers which contain the highly reactive allylic C—H linkage. Free radicals are capable of hydrogen atom abstraction from copresent molecules as well as of addition to an unsaturated linkage. The ease with which this abstraction reaction is carried out by a given radical is a function of the reactivity of the C—H linkage which is attacked. The reactivity of radicals containing these C—H linkages increases in the order aryl < primary alkyl < secondary alkyl < tertiary alkyl < allyl < benzyl. Because of this high reactivity, a monomer containing an allylic C—H, allyl acetate, for example, functions as its own retarder, as shown by sequence (3). The resonance-

$$\begin{array}{ccc} & \sim\!\sim M\cdot + CH_2{=}CHCH_2OAc & \\ \swarrow & & \searrow \\ \sim\!\sim MCH_2CHCH_2OAc & & \sim\!\sim MH + CH_2{=}CHCHOAc \end{array} \qquad (3)$$

stabilized allylic radical will react with the monomer only very slowly. The predominant further reaction is dimerization. Not only is polymerization slowed in this case, but also the molecular weight of the polymer formed in the reaction is low.

Miscellaneous inhibitors. Oxygen, iodine, and nitric oxide interact rapidly with free radicals to yield stable products. Inclusion of these materials in polymerizing systems thus leads to effective inhibition. It is of particular interest that oxygen will copolymerize, under carefully controlled conditions, with certain monomers, styrene, for example, to yield high-molecular-weight polymeric, peroxides. The repeating unit is shown in formula (4). Iodine and nitric

$$[{-}CH(C_6H_5)CH_2OO{-}]_n \qquad (4)$$

oxide have been used extensively to detect and in some cases to identify alkyl free radicals in nonchain as well as in chain reactions. This method has proved to be of great value in defining primary processes in photochemical decompositions of aldehydes and ketones. *See* POLYMERIZATION.

Inhibition of corrosion. Metallic corrosion in conducting media is electrochemical in nature. Local electrolytic cells are set up because of the presence of impurities, crystal lattice imperfections, or strains within the metal surface. The result is dissolution of the metal from the anodic regions. Corrosion inhibitors now in use may operate at the anodes or the cathodes or provide physical protection over the entire surface.

Anodic inhibitors. These are mild oxidants which reduce the open-circuit potential difference between local anodes and cathodes and increase the polarization of the former. Sodium chromate and sodium nitrite are most commonly used. The former is used in air conditioners, refrigeration systems, automobile radiators, power plant condensers, and similar equipment. Sodium nitrite finds special use in the protection of petroleum pipelines. It is effective even on rusty, mild steel. An extension of the nitrite type is the use of nitrite salts of secondary amines as vapor-phase inhibitors. The inclusion of a salt such as dicyclohexyl-ammonium nitrite with a packaged steel object provides effective protection against corrosion.

Cathodic inhibitors. Compounds such as calcium bicarbonate and sodium phosphate, in an aqueous medium, deposit on metal surfaces films that provide physical protection against corrosive attack.

Organic inhibitors. These are usually long-chain aliphatic acids and the soaps which are derived from them. Adsorption of these compounds on metal surfaces gives a hydrophobic film which protects the metal from corrosion by many agents. As little as 0.1% of palmitic acid, for example, is sufficient to protect mild steel from attack by nitric acid. *See* ANTIOXIDANT; CATALYSIS; CORROSION; FREE RADICAL.

Lee R. Mahoney

Bibliography. D. Braun et al., *Polymer Synthesis: Theory and Practice: Fundamental, Methods, Experiments*, 4th ed., 2004; V. S. Sastri, *Corrosion Inhibitors: Principles and Applications*, 1998.

Ink

A dispersion of a pigment or a solution of a dye in a vehicle (carrier), yielding a fluid, paste, or powder that can be applied to and dried on a substrate. Printing, writing, marking, and drawing inks are applied by several methods to paper, metal, plastic, wood, glass, fabric, or other substrates. Inks perform communicative, decorative, and even protective functions.

Printing Ink

Printing ink is by far the most widely used and technically developed type of ink.

Classification. Printing inks may be classified according to their characteristic physical properties (such as paste or liquid), the method of application, or other considerations such as the manner of drying. The composition of an ink can be oil-, solvent-, or water-based. The ink may be a high-viscosity paste, a low-viscosity liquid, or even a powder, as is the case for electrophotographic toner inks. The various drying manners are oxidation, evaporation, penetration, precipitation, polymerization, and radiation curing. Over 100,000 new ink formulations are created each year to meet the needs and requirements of printers.

The most important commercial printing methods that use inks are lithography, flexography, gravure, and letterpress. Economic factors as well as efficiency and speed of press preparation and setup are the key factors for their use. Lithography is currently the leading method of printing in the world and has been so for many decades, particularly in the publication and commercial fields. The next most widely used method is flexography, particularly in packaging printing, followed by gravure printing for both publication and packaging. Letterpress (nonflexographic) is declining in use and is now well behind the other three methods. Other printing methods include screen printing, inkjet, electrophotographic, stencil, and electrostatic. While inkjet and electrophotographic inks are widely used for computer hardcopy output, the quantities of ink consumed are still relatively low as compared to commercial graphic arts printing methods. *See* INKJET PRINTING; PRINTING.

Inks can be applied in four-color (screened) processes, as special spot colors (solid or screened), or as type or line work. Some end uses are newspapers, magazines, commercial, folding cartons, books, corrugated boxes, paper bags, wrappers, labels, metal containers, plastic containers, plastic films, foils, food inserts, sanitary papers, and textiles.

Formulation. Inks are composed of four major components: colorants, vehicles, additives, and solvents. Colorants (pigments, toners, and dyes) provide the desired color and contrast with the substrate. Vehicles, or varnishes, act as carriers for the colorants during the printing operation and, upon drying, bind the colorants to the substrate. Additives influence the printability, film characteristics, drying speed, and end-use properties. Solvents (including water) are used in the formation of the vehicles, and as diluents to reduce ink viscosity and adjust the drying rate. The ink components are weighed, mixed, and dispersed to form the finished product.

Colorants. These are used in the form of a dry powder, predispersed paste (called a flush), or liquid concentrate. Their selection critically influences the color, hue, permanency, opacity, gloss, rheology, and print quality. Thousands of colorants are available for matching the desired print color. Pigments containing cadmium, arsenic, mercury, antimony, lead, selenium, or significant impurities are considered undesirable for food packaging or children's products. *See* DYE; PIGMENT (MATERIAL); RHEOLOGY.

Vehicles. These are usually natural or synthetic resins or polymers dissolved in solvents, oils (hydrocarbon or vegetable), or liquid resins. Many ink resins are of moderate molecular weight to achieve good solubility. Synthetic resins with higher molecular weights have been developed for inks used on high-speed presses. *See* POLYMER; RESIN.

Inks solvents typically are emitted to the atmosphere and often are classified as volatile organic compounds and hazardous air pollutants. The primary solvent used in publication gravure inks is toluene because it will dissolve the high-molecular-weight resins used to make high-performance inks. However, publication gravure avoids solvent emissions by using an activated carbon system to recover the toluene. Water has become an important solvent for liquid flexographic and gravure printing inks that are used for printing porous substrates such as paper, particularly in packaging printing.

Drying methods. Vehicle selection or its composition is dictated by the ink drying process and the substrate and end-use requirements. The ink vehicles are best categorized according to their drying mechanism and application. Both physical and chemical reactions may take place during the solidification, or drying, step.

Evaporation. Solvent evaporation is accomplished by convective or radiant heat. An obstacle associated with converting to water-based inks is the drying process. Water has a higher latent heat of evaporation than organic solvents, resulting in slower drying and lower press speed. To alleviate this problem, a certain amount of organic solvent, such as an alcohol, is added to the ink. In addition, water-based inks require much higher color strength than solvent-based inks in order to facilitate drying by reducing the wet-ink film thickness.

Letterpress and lithographic heatset inks are based on vehicles consisting of modified rosin and hydrocarbon-type resins. These are dissolved mainly in aliphatic hydrocarbons. Heatset vehicles are often gelled for faster and sharper tone printing. *See* ROSIN.

Flexographic solvent-based inks use one or more synthetic polymers dissolved in lower alcohols, often in combination with small amounts ($<10\%$) of esters, hydrocarbons, or glycol ethers. Emulsions or microdispersions of acrylic copolymers, polyamide resins, and polyvinyl copolymers, along with nitrocellulose, are used in flexographic inks for printing

films and foils. Since flexographic printing presses may use rubber plates and rollers, the tolerance for many solvents is often below 10%. *See* ALCOHOL; COPOLYMER; ESTER; HYDROCARBON; POLYACRYLATE RESIN; POLYAMIDE RESINS; POLYVINYL RESINS.

Gravure inks for packaging applications contain a variety of polymers dissolved in blends of organic solvents. The use of these solvents has been reduced, or even eliminated in many cases, since they contributed significantly to the volume of solvent emissions. One goal has been to replace most or all the solvent content with water. However, water-based inks have not been developed with all of the physical properties required by many packaging applications, especially resistance to water and heat.

Gravure systems for publication printing use calcium or zinc salts of rosins, rosin esters, and phenolic modified rosin or hydrocarbon resins dissolved in aliphatic or aromatic hydrocarbon solvents.

Precipitation. Binder (resin and oil) precipitation occurs when water is absorbed by a hygroscopic solvent. Steam or moisture from the air causes the binder to separate from the solvent; the solvent then either penetrates a porous substrate or evaporates. Typical binders used in letterpress and dry-offset printing inks are fumarated rosin and occasionally a maleated drying oil.

Oxidation. This is the reaction of atmospheric oxygen with unsaturated fatty acids present in vegetable oils or in their alkyds. Cobalt and manganese salts of fatty and aromatic acids act as catalysts for the reaction. Letterpress and lithographic sheetfed inks dry (solidify) by this mechanism. *See* DRYING OIL.

Thermal polymerization. This reaction normally proceeds via a polyol-amine condensation path or through the Diels-Alder diene-addition mechanism. Lithographic, dry-offset, letterpress, and screen-printing inks for metal, paper, and glass frequently use this method. Alkyds or polyesters with free hydroxyl groups condense with melamine resins in the presence of acid catalysts. Alkyds or oils with diene-conjugated unsaturation polymerize via the addition and oxidation mechanism simultaneously. *See* AMINE; DIELS-ALDER REACTION; POLYESTER RESINS; POLYMERIZATION; POLYOL.

Gel-fusion. Gel-fusion takes place when the binder, composed of a nonsolvent plasticizer and a dispersed polymer, is heated to the temperature at which polymer solubility is achieved, and upon cooling a solid ink film results. Letterpress, dry-offset, and some lithographic inks contain a poly(vinyl chloride) resin dispersed in a suitable plasticizer.

Quick setting. This method relies on a two-step mechanism involving the critical solubility of a resinous binder and the porosity of the substrate. A vehicle with borderline solubility releases part of its solvent to the substrate, precipitating the resins, and becomes tack-free. The second drying step proceeds via the conventional oxidation process. Lithographic and letterpress paper and folding-carton inks for sheetfed presses use this drying principle.

Absorption. Absorption setting is achieved through the penetration (sponge effect) of the total vehicle or its liquid portion into a porous substrate. Newspaper inks, based on mineral (hydrocarbon) oil or on rosin, or gilsonite (variety of asphalt) in mineral oil varnishes, set, although never fully dry (**Fig. 1**).

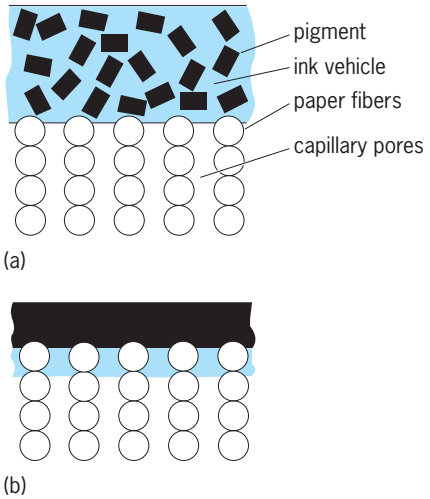

Fig. 1. Ink drying by absorption setting. (*a*) Ink film on uncoated paper immediately after printing. (*b*) Ink film after absorption of the vehicle by the paper.

Cold setting. In this method, inks are applied in a molten state and solidify on cooling. Cold-setting inks (also called hot-melt inks) contain no solvents. Instead, they are based on a combination of waxes (petroleum and synthetic) and colorants. *See* WAX, PETROLEUM.

Radiation curing. This method uses ultraviolet (UV) or electron-beam radiation to cure (dry) the ink. The principal advantages of this method are lower energy consumption, instant curing, low-temperature curing, and reduced solvent emissions.

Ultraviolet light, generated by medium- or high-pressure mercury lamps, activates photoinitiators to generate reactive species, called free radicals, which initiate the polymerization of acrylic monomers and oligomers in the vehicle. Market growth of UV inks has been mainly in printing sheetfed litho cartons and plastic cups, and in metal decorating. UV curing of overprint coatings, which impart gloss and abrasion resistance to book covers, magazines, and packaging materials, is very popular and growing in use. UV curing is also widely used in narrow-web flexographic printing of labels and in screen printing. *See* PHOTOCHEMISTRY; ULTRAVIOLET RADIATION.

Electron-beam inks are chemically similar to ultraviolet inks except that they do not require photoinitiators to polymerize acrylic monomers and oligomers since high-energy electrons are able generate free radicals. *See* CHARGED PARTICLE BEAMS.

Infrared drying. Infrared heating is occasionally used to help set lithographic inks. Although infrared heating is not used with web offset presses, it finds some utility on sheetfed litho presses. Inks containing some volatile solvents are required. *See* INFRARED RADIATION.

Microwave drying. This method is used occasionally in liquid inks that contain water as the main solvent,

but finds very limited use because of poor energy efficiency in creating heat. *See* MICROWAVE.

Additives. A simple suspension of pigment in varnish or solvent usually does not produce a commercially satisfactory ink. Many other materials must be added in order to provide good performance. These additives include plasticizers, wetting agents, waxes, shortening compounds, reducers, thickening agents, antiskinning agents, and antipinhole compounds.

Plasticizers are commonly used in flexographic printing inks. They make flexographic inks softer, more flexible, and more adherent to the substrate. Resins are often rigid and brittle, and adhere poorly to flexible substrates. Selection of the plasticizer depends on the printed product and the resin in the ink. Since the boiling points of plasticizers are generally high, they do not evaporate in the dryer but become a permanent part of the ink film. Plasticizers include various phosphates, sulfonamides, chlorinated materials, citrates, adipates, polyglycol derivatives, resinates, and phthalates.

Most printing inks contain waxes to improve rub and scuff resistance. The waxes used in inks are often a compound comprising a fine dispersion of several waxes in a vehicle. They may also be a micronized (very finely divided) dry powder. Common classifications of waxes include animal (beeswax, lanolin), vegetable (carnauba, candellila), mineral (paraffin, microcrystalline), and synthetic [polyethylene, poly(ethylene glycols), polytetrafluoroethylene (PTFE or Teflon)]. A wax that is too soluble in an ink will come out of solution when the ink cures, form a waxy film on the surface, and prevent wet ink from trapping (overprinting) on dry ink films. This problem is commonly referred to as crystallization. Excessive use of wax should be avoided to prevent problems associated with ink piling. Too much wax in the ink may cause increased drying time, trapping problems, and decreased gloss.

Wetting agents promote the dispersion of pigments in ink varnishes. Wetting agents for lithographic inks must be carefully selected; otherwise, they can cause excessive emulsification of the dampening solution into the ink and other associated problems. Wetting agents may also reduce the surface tension of aqueous flexographic and gravure inks in order to improve their wetting properties on plastic and foil substrates.

Shortening or antimisting agents reduce ink flying or misting of ink at high speeds. The addition of a wax or an organo-clay compound shortens an ink and can reduce misting. Improper amounts of antimisting agents can interfere with ink flow on the press and aggravate ink piling.

Reducers, such as heatset oils or other petroleum solvents, are occasionally added to soften and reduce the tack of a litho ink. A thin litho varnish, grease, boiled linseed oil, or a light-oil-modified alkyd will also reduce the tack of an ink.

Thickening agents are used to modify the viscosity and tack of paste inks to adjust them to specification.

Antiskinning agents are antioxidants that slow the drying of sheetfed offset inks, so that they will not skin (dry) in the can. If they are sufficiently volatile, they will not greatly retard the drying of the printed ink.

Defoamers and antifoamers for aqueous flexo or gravure inks are blends of surface-active organic liquids, surfactants, metal soaps, hydrophobic silica, and other ingredients, with or without silicone modification. They may be used during the preparation of aqueous inks to prevent foam buildup, and may be added during use for the same purpose.

Solvents. Solvents dissolve the oils, resins, and additives used to produce the vehicles that carry the pigment. A wide range of solvents are used in printing inks, including hydrocarbons, heatset oils, alcohols, glycol ethers, ketones, and esters. Polar solvents such as alcohol, ethers, ketones, and esters are useful in flexographic and gravure vehicles for dissolving polar resins such as shellac, cellulose esters, phenolics, and acrylics. The hydrocarbons, which are nonpolar solvents, are suitable in lithographic and letterpress vehicles for dissolving compounds such as drying oils and rosin-modified phenolics and maleics.

Two classes of hydrocarbons are used: aliphatic and aromatic. Aromatic hydrocarbons are more powerful resin solvents than aliphatic hydrocarbons. Toluene is used extensively in publication gravure inks. Mineral spirits, naphtha, and kerosene consist principally of aliphatic hydrocarbons. These solvents are used in inks for lithographic and letterpress printing. In flexographic inks, aliphatic hydrocarbons are primarily used as cosolvents in polyamide formulations or as inexpensive diluents. Aromatic hydrocarbons are used sparingly in flexographic inks because they tend to swell rubber and plastics.

Carefully fractionated (narrow-boiling) hydrocarbon heatset oils are solvents for most of the resins found in lithographic and letterpress inks. The choice of solvent with the proper boiling range depends on the type of ink. The boiling points of aliphatic hydrocarbon solvents for heatset inks vary from about 400 to 600°F (200 to 300°C), but each designated fraction boils over a much narrower range. The choice of heatset oil depends principally upon the type of resin used and the printing conditions for which the ink is intended. The lower the boiling point, the lower the temperature required to dry the ink. Selection of the solvent depends not only on the boiling point but also on the characteristics of the dryer, press, press speed, ink coverage, and paper stock.

Alcohols are used extensively in flexographic and packaging gravure inks and in some screen-printing inks. Their dissolving power, evaporation range, and mild effect on rubber make them important flexographic ink solvents. Alcohols include ethyl, normal propyl, and isopropyl alcohols. Ethyl alcohol and isopropyl alcohol in flexographic inks are used primarily in polyamide-based vehicles and as additives in water-based formulations. Normal propyl alcohol is used to slow down the evaporation rate of inks containing alcohols that are more volatile.

Glycol ethers are alcohols with ether linkages and are excellent solvents for most synthetic resins. The glycol ethers used in printing inks are soluble in both water and alcohols. Glycol ethers are used in flexographic and packaging gravure inks, as well as screen printing inks. They are good solvents for nitrocellulose and many other resins, including shellac, vinyl acetate, and rosin derivatives.

Ketones are strong solvents. Acetone is used in nitrocellulose lacquers and packaging gravure inks. It is miscible with water, alcohols, ethers, oils, hydrocarbons, halogenated hydrocarbons, fatty acid esters, and most organic solvents. Methyl ethyl ketone and methyl isobutyl ketone are used in gravure and screen printing inks. Cyclohexanone and isophorone are used in some screen-printing inks. Ketones are rarely used in flexographic printing inks because they swell rubber and plastic plates.

Esters are generally good solvents for nitrocellulose. The rapidly drying solvents ethyl and propyl acetates are used extensively as cosolvents in flexographic and packaging gravure inks. *See* SOLVENT.

Testing. Inks are subjected to numerous laboratory tests during and after their manufacture. The most common properties that are checked are fineness of grind, color and strength, gloss, abrasion resistance, drying speed, and rheology. Other characteristics important for special types of inks are tack, misting, water emulsification, transferability, and overprint (trap) quality for lithographic inks; adhesion, odor, and solvent retention for flexographic or gravure laminating inks; ice water, oil, and soap sensitivity for surface-printed packaging inks.

Fineness of grind. The most precise method for testing fineness of grind is the actual measurement of the pigment particle size, the size range, and its population. Several approaches are available to test the particle size and distribution, which indirectly influence an ink's color strength, printability, gloss, and rheology. Many light-scattering instruments are now available to determine the particle size distribution of inks, but they are expensive and time-consuming and used mainly in research and development. Particle size and distribution can also be determined by optical and electron microscopy. For routine quality-control tests, a fineness of grind test for oversize particles is done on a grindometer (**Fig. 2**). *See* LIGHT-SCATTERING TECHNIQUES.

Color. This is one of the key properties of an ink because of its functional and esthetic value. There are several methods for controlling the printed color, which depends on both the thickness of the ink film and its intrinsic color. Reflection densitometers are used in conjunction with color filters to measure print optical density at press side, which is proportional to printed ink film thickness. The actual printed ink color can be measured with a portable computerized spectrophotometer, which yields absolute values for hue, lightness, and chromaticity. These data are often specified as a quality-control parameter by the printer for critical jobs.

Rheology. Ink rheology influences press transfer, print fidelity, overprinting, and gloss characteristics.

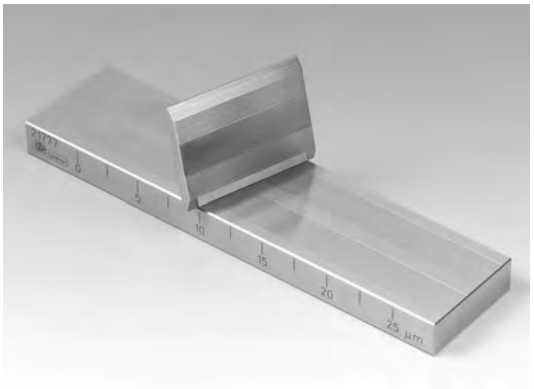

Fig. 2. Grindometer.

Fig. 3. Brookfield digital viscometer. (*BYK-Gardner USA*)

Several instruments are available to measure viscosity, yield value, and flow characteristics (**Fig. 3**). Special instruments have been designed for dynamic high- speed press applications, where determination of the viscous and elastic components in ink rheology has become more critical. Viscometers usually are interfaced with a computer to handle the calculations, program the tests, and analyze the output. *See* MICROCOMPUTER; RHEOLOGY.

Tack. This measure of film splitting (separation) force applies to lithographic and letterpress paste inks. It controls the transfer of ink from plate or rubber blanket to substrate and influences the overprinting characteristics. The tack is measured with roller distribution instruments known as inkometers and tackoscopes (**Fig. 4**). Higher press speeds require lower ink tacks to prevent tearing, or picking up lint from, the paper.

Transfer. The laboratory determination of ink transfer and printability is normally done by print proofing. This test is used to predict the color of an ink

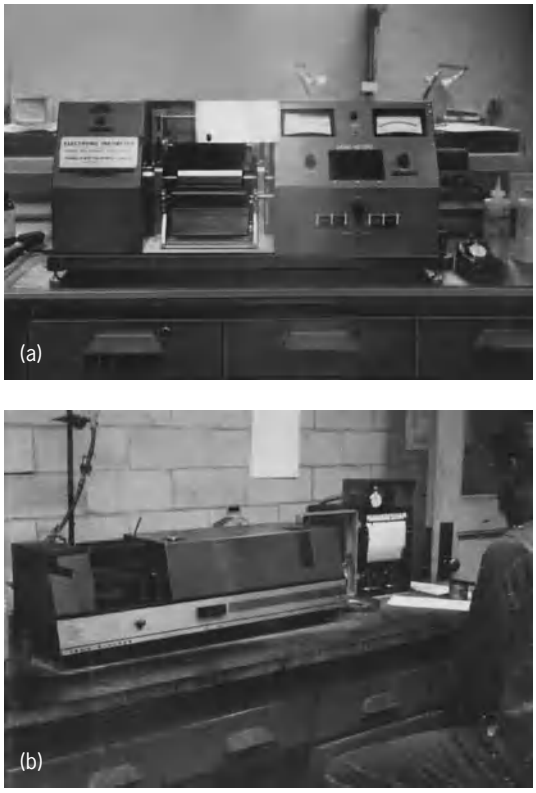

Fig. 4. Instruments for the measurement of ink tack. (a) Inkometer. (b) Precision tack tester. (*Prüfbau Instruments*)

applied to a specific substrate or to a predeposited ink film.

For lithographic inks, proofs can be made with a simple press, where a hand-cranked impression roller transfers ink from a plate to a substrate. For accurate high-speed tests, automatic proof presses are used (**Fig. 5**). Here both ink and fountain solution are applied in a controlled amount to a substrate.

For flexographic inks, a manually operated tool consisting of an engraved (anilox) cylinder and a rubber roller is used. The ink is metered through the nip between the cylinder and the roller. The substrate is put on a flat resilient surface, and the ink is applied at hand speed. Several motor-driven flexographic proofers are also available, which provide more consistent and accurate prints.

For gravure inks, small gravure presses are used. Here the engraved cylinder is partially immersed in an ink fountain. The excess ink is wiped off the cylinder by a metal blade and printed onto a substrate in web or sheet form.

Ink-water balance. In lithographic printing, the most important press settings involve the ratio of the ink feed rate to the fountain solution feed rate. This is called the ink-water balance and is critical to good printing. Since a lithographic plate has a flat printing surface, the image must be chemically differentiated by keeping the image areas wetted with ink and the nonimage areas wetted with water at all times. At present, there is no automatic control for balancing the two feed rates, even though the ink feed can be controlled to achieve a given optical density. The press operator's skill in setting the fountain solution feed is critical to good printing.

Computer color control. Several computer color-matching systems are used in the ink industry for replicating ink (color) formulas, matching new colors, reducing costs, and quality control of pigments. They automate color matching mathematically by blending stored spectral curves of pigments of different color. In these systems, a spectrophotometer is coupled to a computer to produce very rapid results.

Manufacturing processes. Production of most inks is predominantly a batch process, often consisting of wetting, mixing, milling (dispersion), and filtering steps. It must be efficient and flexible, capable of producing a consistent and high-quality product. Some larger-volume inks (such as newspaper, web offset, and publication gravure inks) are mass-produced in a semicontinuous mode. Such lines are automatically controlled, employing microprocessors and process sensors.

The three main steps in ink making are mixing, dispersion, and filtering. Mixing and its intensity differ with the characteristics of various inks and depend on the mixer speed, size, and configuration. Mixing speeds range 100–5000 revolutions per minute, which requires proper sizing of the motor for adequate energy input. Dispersion reduces agglomerated pigment particles and even crushes larger pigment crystals. This step controls the color strength (and value) and influences ink rheology.

Dispersion of paste (letterpress and offset) inks normally is done with three-roller mills or media mills (**Fig. 6**). Liquid (flexographic and gravure) inks are dispersed with closed-head media mills and with ball mills. Colloid mills are used occasionally for

Fig. 5. Multipurpose printability tester. (*Prüfbau Instruments*)

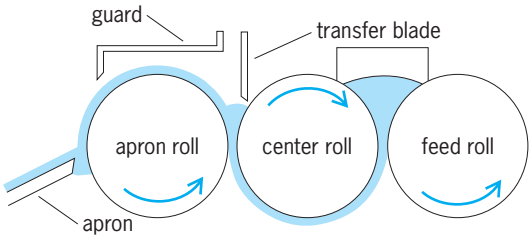

Fig. 6. Three-roll mill.

newpaper inks and some liquid inks. *See* GRINDING MILL.

Filtering is done with special cloth bags for liquid inks and vehicles, and with bags or cartridges for paste inks. This step removes impurities, agglomerated pigment, fibers, and grit. Media milling and filtering are often used in the continuous production of large-volume process color inks. Magnetic filters are also used for gravure inks.

Ink preparation. Typically, an ink can be prepared by either of two processes. In the dry grind process, all the ingredients (such as varnish, pigment, compounds, talc, oil, and solvent) are placed in a mixer. After mixing and dispersion, the ink is tested and compared to a standard for tack, viscosity, grind dispersion, color, strength, and drying. The ink is adjusted as needed and then packaged. The ink is distributed in sizes ranging from small cans to commercial-size reusable tote bins or tank trucks.

In the base process, the pigment is mixed with only that part of the varnish that imparts the proper viscosity to fit the equipment being used, forming a concentrated base. The base is then mixed with the other ingredients, tested, and packaged. For certain lithographic inks, the base is not made from dry pigments but from pigment presscake. Here the water-wet presscake is mixed with varnish in a heavy-duty dough mixer; the water is flushed out via the shearing action and a flushed base is produced. Another method for flexographic or gravure inks is the use of color chips. The dry pigment is mixed with resin plus some plasticizer and milled on a two-roll mill. The resulting dispersion is dissolved in solvent, and other ingredients are added to make the final inks. Liquid pigment dispersions for the preparation of finished inks are also available. They generally are prepared from pigment, solvent, resin, and a dispersant. They are diluted with a vehicle to prepare an ink.

Special Inks

A number of special-purpose inks have been developed; these may be used for complex applications other than simple printing.

Laminating inks. Technically sophisticated formulations are required for producing laminating inks for flexible film packaging. Unlike other applications, where the ink is applied onto the outside surface of a substrate, a laminating ink is applied via flexography or gravure onto the reverse side of a thin film, which in turn is laminated to other films, foils, or papers. Such inks and structures are used for packaging snack foods, beverages, and medicines. Special ink formulations based on synthetic resins are used for more demanding products such as aseptic food packages. *See* FILM (CHEMISTRY).

Electrophotographic inks. The digital color imaging process technology that led to the development of on-demand printing requires different types of ink than those used in conventional printing. Two special inks have been developed: dry toners and liquid toners.

In these systems, images and characters are generated and stored in computers and then transferred via digitally driven laser beams onto photosensitive drums, resulting in a latent image of the desired object. Dry or liquid toner inks are then fed onto the surface of the drum. Next, electrostatic forces within the latent image on the drum attract the toner, as it possesses an opposite charge. The inks are transferred from the drum to an appropriate substrate and fixed by heating. In some instances, an intermediate tool (offset blanket) is used in the transfer process between the drum and the substrate.

Electrophotographic inks must be capable of carrying the electrical charge required by the specific type of drum in a given printing system. Dry toner inks are composed of fine resinous thermoplastic powders, and liquid toners are composed of finer powders dispersed in an inert dielectric carrier such as an isoparaffinic hydrocarbon. *See* PHOTOCOPYING PROCESSES.

Engraving (intaglio) inks. These are inks designed for high-quality and security printing, as required for paper currency, stock certificates, and formal letterheads. The image is engraved below the plate surface in multiple lines on the steel plate. A thick layer of ink is applied to the plate, filling the recesses of the engraving, and the excess ink is removed by wiping with paper or a soft plastic-covered roll and a special washing fluid. The rheology of the ink is such as to allow for easy wiping (a property called short body). In addition to colored pigment, a large amount of colorless extender is used to assure that the thick film of ink applied will wipe properly from the nonimage area (plate surface). In the case of paper currency, special ingredients are used to assure such properties as durability, stiffness to accommodate money-changing machines, magnetic properties, launderability, and chemical and abrasion resistance.

Transfer inks. These inks may be printed by the flexographic, gravure, or lithographic process onto a temporary substrate (usually paper) for later transfer onto the final objects they are intended to decorate. Heat-transfer inks that contain sublimable dyes are printed on paper and are transferred to other substrates using heat and pressure. The most popular application is in imaging of fabrics, especially polyester. Decalcomania inks are normally transferred from paper onto objects such as ceramics, windows, or outdoor signs. *See* TEXTILE PRINTING.

Waterless lithography. Waterless offset inks contain many of the same ingredients found in conventional lithographic inks. The difference is that waterless inks have vehicles that produce higher initial viscosities than conventional inks. Waterless offset inks

also may have different tack than conventional offset inks in order to increase their cohesion. The viscosity and tack differences of these inks relate to differences of the printing plates used. Conventional offset lithographic plates have a hydrophilic nonimage area that attracts water or dampening solution, and an oleophilic image area that repels water and attracts oily ink. Waterless offset plates use a silicone polymer for nonimage areas of the plate and a metal or ink-accepting image area. The silicone polymer's low surface energy resists ink, as long as the ink's viscosity is high enough for the ink to be more cohesively attracted to itself than to the silicone.

The viscosity of ink decreases rapidly with increasing temperature (about 10% per degree Celsius). In conventional lithography, the presence of water in the press's dampening solution cools the ink, allowing it to maintain a more stable viscosity. The viscosity of waterless offset inks is maintained on the press by a temperature control system, either a plate cylinder cooling system or an ink oscillator (vibrator) cooling system. Waterless offset inks are formulated at varied viscosities for different press speeds.

Duplicating inks. Essentially, these are modified sheetfed lithographic inks for special small presses used by offices and print shops. Duplicating printing presses are characterized by two different and relatively simple ink-distribution roller arrangements. In the integrated ink-water system, a single roller is used to apply both ink and fountain solution to the plate. As a result, the ink must be capable of releasing the water easily under modest contact pressure. The ink formula is similar to other lithographic inks, except that it is formulated to be fast-setting without heat.

Other inks. Inks developed for ballpoint pens are thin fluids of high color tint (pigment or dye) strength. They usually contain dye solutions or fine pigment dispersions in various low-viscosity vehicles. The inks must be free of particles and resist premature drying to feed to the paper without clogging. Rapid penetration into the paper accomplishes drying.

Stamp-pad inks are supplied in cloth or foam rubber pads and transferred by pressure to raised rubber type or image, which is then stamped or pressed against the substrate. The inks must remain nondrying on the pad, yet rapidly penetrate into the stamped substrate. Dyes are often dissolved in glycols or other low-volatility solvents to form these inks.

Regulations. In the United States, federal and local governments regulate ink manufacture and use. The most important regulation affecting the industry is the Clean Air Amendment of 1990, administered by the Environmental Protection Agency (EPA). This regulation monitors emission of volatile organic compounds and hazardous air pollutants. In addition, the industry must conform to two complex federal government regulations: the Resource Conservation and Recovery Act (RCRA) and the Comprehensive Environmental, Response, Compensation, and Liability Act (CERCLA), including its self-standing Superfund Amendments and Reauthorization Act (SARA). These laws, also administered by EPA, have led to major changes in the industry such as avoidance of pigments containing heavy metals, development of inks capable of easy deinking for recycling of news and office paper, and recycling of ink waste. In the more complex area of toxicology, the ink industry is regulated by the Toxic Substance Control Act (TSCA).

Worker safety is ensured by the Occupational Safety and Health Act (OSHA), administered by the Department of Labor. This regulation is mainly concerned with the vapors (solvents) and liquids (resins) with which the ink manufacturer or the printer comes into contact during routine work. *See* ENVIRONMENTAL ENGINEERING; INDUSTRIAL HEALTH AND SAFETY.

Robert W. Bassemir

Bibliography. T. M. Destree (ed.), *Lithographers Manual*, 9th rev. ed., GATF Press, 1994; N. R. Eldred, *What the Printer Should Know about Ink*, 3d ed., GATF Press, 2000; N. R. Eldred, *Chemistry for the Graphic Arts*, 3d ed., GATF Press, 2001; G. G. Field, *Color and Its Reproduction*, 2d ed., GATF Press, 1999 Flexographic Technical Association, *Flexography: Principles and Practices*, 5th ed., FFTA, 1999; R. H. Leach (ed.), *The Printing Ink Manual*, 5th ed., Blueprint-Chapman & Hall, 1993; J. MacPhee, *Fundamentals of Lithographic Printing*, vol. 1, GATF Press, 1998.

Inkjet printing

A method of forming a hardcopy image (text or graphics) by generating very small drops of ink and precisely depositing them on a substrate. This concept was introduced commercially in the mid-1980s to replace dot-matrix (low cost but slow and noisy) and laser-based (high speed and high resolution but also high hardware cost) printing systems for home and office applications. As inkjet color inks and photographic-quality paper development progressed, a new goal of supplanting traditional silver halide photographs was established. Today, the home photographer can produce inkjet prints that exceed the quality and light-fade stability of color photographic prints. *See* PRINTING.

Inkjet print quality has improved continuously by using a systems approach of providing smaller drop volumes along with better media (paper) and inks. The colorants used in the inks may be dyes (water/solvent-soluble) or pigments (insoluble). Dye-based inks historically provided more vivid color and were somewhat easier to manufacture and keep in suspension, but pigment-based inks are closing the gaps in these areas. Pigmented inks, which do not penetrate the media as readily as dye-based inks, have been developed with the equivalent quality of laser printers or typewriters for black text printed on plain paper. Compared to dye-based inks, pigmented inks provide water-fastness, light-fade resistance, and better smear resistance. Recently, color pigmented inks have found use in large format inkjet printers for outdoor signs. Photo imaging and color adoption in the office applications have been instrumental

Comparison of inkjet printhead technologies

		Drop-on-demand	
	Continuous (piezo)	Thermal (TIJ)	Piezo (PIJ)
Advantages	High volume High frequency Ink flexibility Long throw distance Many substrates	High efficiency Low cost Integrated chip processes Integrated drive High frequency Nozzle density	High efficiency Ink flexibility Long throw distance Many substrates Multiple drop volumes Long-life head
Disadvantages	Nozzle density High head cost	Ink flexibility Unwanted heat Head reliability	Nozzle density High head cost Electrical integration

in increasing use of inkjet products. *See* DYE; INK; PAPER; PIGMENT (MATERIAL).

The human eye can detect dots on paper 25 micrometers (0.001 in.) or smaller if the color is dark (high optical density). A dot of this size on photo paper requires an ink drop of 2 picoliters or less, assuming typical dot spreading on the media. Therefore, drop volume reductions beyond 1 pL are probably not necessary for grain-free photos. Color printing speed and quality also have been improved by adding more colors, some with reduced dye or pigment loads of the black, yellow, cyan, and magenta colorants, but this adds product cost. In the future, important product features will be speed and ease of use.

Printhead technologies. There are two popular techniques used to jet inks with adequate drop ejection velocity, size, and directional control. They are continuous and drop-on-demand (DOD), as seen in the **table**. The earliest inkjet (continuous) used a piezoelectric material that changes shape when a voltage is applied. This shape change is used to increase or decrease the volume of ink confined in the ink chamber. A column of ink exits the firing chamber through a small hole or nozzle and then breaks into a droplet. The ink flow is continuous as a periodic voltage is supplied to the piezoelectric material. Unwanted drops are deflected into an ink-recycling channel. Piezoelectric materials are also used in DOD printheads. Thermal inkjet (TIJ) is another DOD technology and has gained over 75% of the market. Here the ink is superheated, and the resulting large volumetric expansion is used to propel the droplet. Typical drop volumes range from 2 to 50 pL, with velocities of 10 m/s, and a head-to-media spacing of 1 mm. *See* PIEZOELECTRICITY.

Continuous inkjet. This technology is used primarily for high-volume applications such as magazine labels, can-dating codes, bar-code labels, and carton marking. Inkjetting frequencies of 100 kHz or more can be attained since the head is operated in a stable mode. The ability to switch between ink systems (such as dye and pigmented inks), known as ink flexibility, and the short distance from the head to the substrate are advantages. However, the nozzle packing density (number of nozzles per printhead area) and the high printhead cost have prevented this technology from penetrating consumer printing applications. New applications for continuous inkjet are emerging as new inks with improved durability and specific attributes that allow direct printing on glass, ceramics, fabrics, and other substrates are developed. Systems using continuous inkjet technology typically cost from $50,000 to nearly $1 million.

Drop-on-demand inkjet. In digital printing, the most efficient use of energy is to place ink only when and where needed, or DOD. The volume, velocity, and direction of each ejected drop are critical for maintaining consistent quality and operating conditions. A computer generates the desired image and sends the dot firing sequence to the printer, which either energizes a piezo crystal or fires a thermal resistor in a printhead with hundreds or thousands of small nozzles lined up in several rows. The printhead scans over the media, printing a swath of one or more colors. This process is repeated until the entire document is printed.

Thermal DOD inkjet. Since the mid-1980s, thermal inkjet has shown continuous improvement in quality, speed, and cost. A major breakthrough in the technology was the development of the replaceable printhead. This allowed both the ink and printhead to be placed in one inexpensive cartridge. There are two primary technologies used in thermal inkjet: edge and top shooter (see **illus.**). Both techniques use the physical principle of converting the electrical energy delivered to the firing resistor into heat in order to superheat a small volume of ink. Most thermal inkjet now use the top shooter design since it is more thermally efficient. Thermal inkjet primarily uses established integrated circuit manufacturing processes and equipment, so costs are generally lower than piezoelectric inkjet (PIJ). Its silicon-based substrate allows the integration of electronic drive circuitry, which reduces the overall printing system costs and improves performance. The latest thermal inkjet technology uses silicon etching for ink feed slots, thermal inkjet resistor formation, and ink channel/nozzle formation. Drop ejection frequencies of up to 36 kHz have been achieved. The upper limits for firing frequency are determined by the overheating and fluid refill characteristics of the chamber. Firing chamber densities are currently at 600 dots per inch (dpi). For most applications, acceptable text

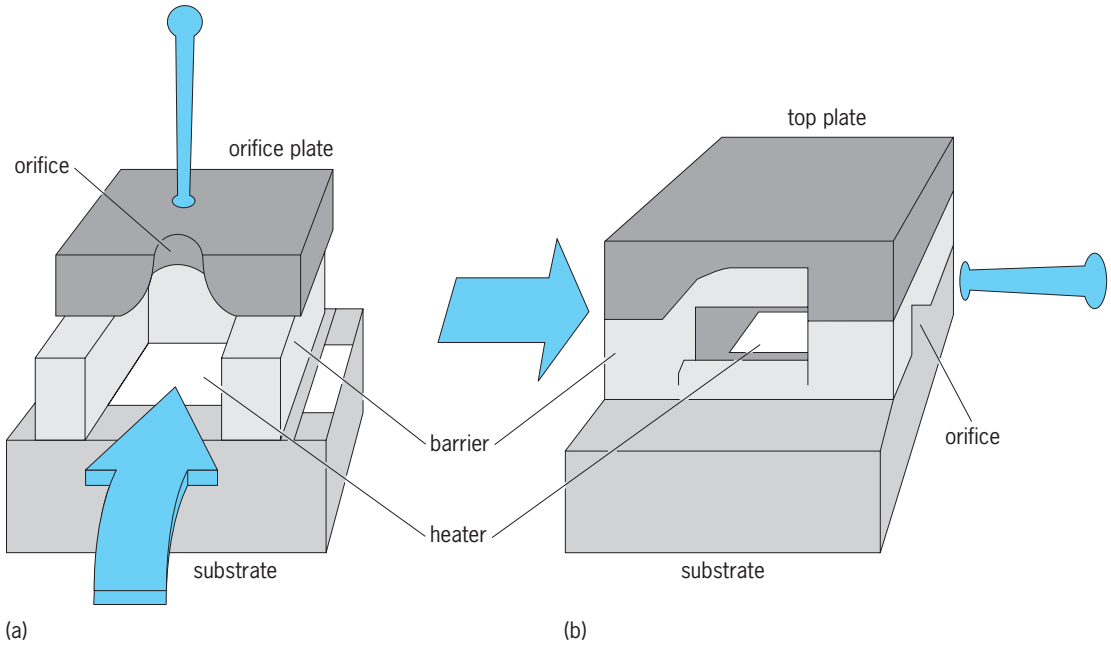

Thermal inkjet printhead designs: (*a*) top shooter and (*b*) edge shooter.

and graphics quality can be achieved at 600 dpi, so a single pass of the printhead is sufficient. Higher-quality printing is achieved with several passes of the printhead at a sacrifice in printing speed. Thermal inkjet is lower-cost, and allows a higher degree of circuit integration and higher nozzle packing density than piezoelectric inkjet.

Piezoelectric DOD inkjet. Piezoelectric inkjet did not find wide acceptance until the mid-1990s, when costs, image quality, and ease of use began to match thermal inkjet. Early DOD piezoelectric inkjet piezoceramic printheads required very costly machining operations. The development of low-cost thick-film piezo elements and high-volume manufacturing contributed to this technology. Materials used in piezoelectric inkjet printhead construction are more robust than thermal inkjet, which has enabled them to jet a wider range of inks. In contrast to thermal inkjet, the piezoelectric inkjet printhead is permanent and only the ink cartridge is changed. One popular use is a phase change ink, or wax-based ink. The printhead is operated above the melting temperature of the wax, so the jetted drops "freeze" on the media. This allows a wide choice of substrates, but the printhead and maintenance costs are high. These products are used in the office and low-volume commercial markets. Since the ink is "squeezed" from the chamber, more control of the fluid dynamics is possible. As many as three different drop volumes ranging in an order of magnitude are possible from a single firing chamber, reducing thermal inkjet's advantage of higher nozzle packing density. Drop dynamics are similar to thermal inkjet with fewer unwanted satellite drops formed, but slightly lower drop velocity and operating frequency. Rob Beeson

Bibliography. R. Beeson, Is print quality an issue anymore—Or is it performance, performance?, *GIGA 25th Global Ink Jet and Thermal Printing Conference*, 2002; H. Le, Fabrication materials and process on ink jet printheads, *IS&T Non-impact Printing Conference Tutorial*, 2000; R. Mills, Ink jet in a competitive market place, *IMI 10th Annual European Ink Jet Printing Conference*, 2002.

Inland waterways transportation

The movement of cargo on inland and intracoastal waterways by barge. Barges move over 800 million tons of raw materials and finished goods each year just in the network of inland and coastal waterways of the United States. This waterways network is a "super water highway" that is an important part of the intermodal transportation system. Barges directly serve 87% of all major United States cities, accounting for 78% of all domestic waterborne traffic.

In the United States, barges move on the 25,543 mi (41,107 km) network of commercially navigable inland and intracoastal waterways (**Fig. 1**). Generally, barges are either pushed or towed by towboats on the inland waters. Barges require standard operating depths of at least 9 ft (about 3 m). Such depths are provided in 15,675 mi (25,226 km) of channel; the other 9868 mi (15,881 km) are less than 9 ft deep. Certain sections of the inland waterways are deep enough to support oceangoing vessels, including the Hudson River to Albany, New York; the Lower Mississippi north to Baton Rouge, Louisiana; the Houston Ship Channel; the Delaware River; the James River; and portions of the Columbia-Snake River.

Vessels. Four major types of barges are used on the inland and intracoastal waterways: the open hopper, the covered dry cargo, the tank barge, and the deck barge.

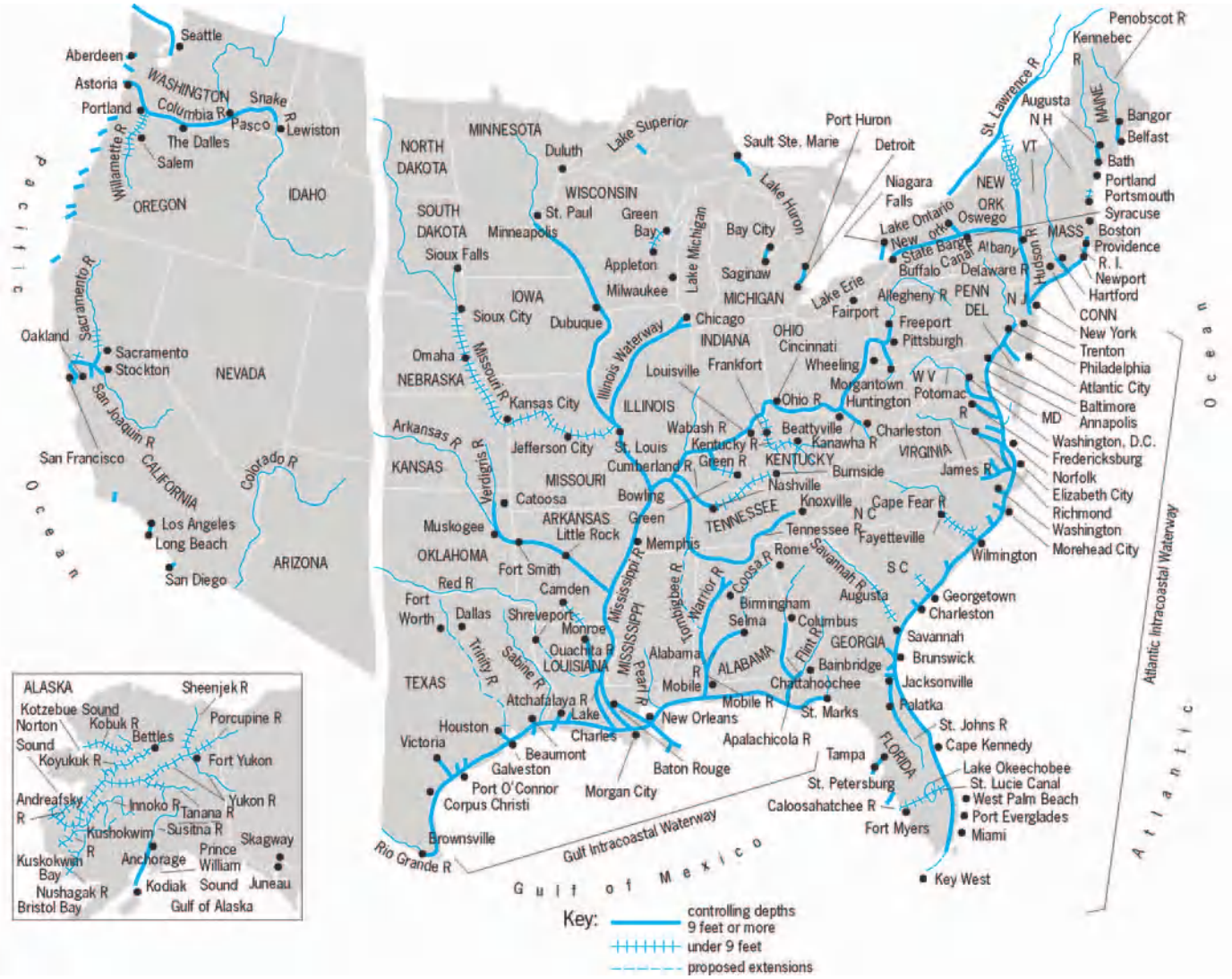

Fig. 1. Commercially navigable inland and intracoastal waterways of the United States.

The open hopper barge is used to transport cargo that does not need to be protected from the elements, such as coal, sand, and gravel. The capacity of this type of barge ranges from 1000 to 1300 tons (900 to 1200 metric tons), and the range in length from 175 to 290 ft (53 to 88 m).

Cargoes that need protection, such as grain, are shipped on covered dry cargo barges, which have 1000-1500-ton (900–1350-metric-ton) capacities and range in length from 175 to 200 ft (53 to 61 m). The covers of these barges may be the lift-on, lift-off type, or have a telescoping design, where the covers roll fore and aft on tracks installed on the barge.

Tank barges carry liquid commodities, such as petroleum and chemicals. The capacity range is 1000-3000 tons (900-2700 metric tons), and lengths vary from 175 to 290 ft (53 to 88 m). Many tank barges are divided into compartments—some heated to high temperatures, some below freezing—so that a variety of liquid cargoes can be carried in one tow.

Deck barges have flat tops, conducive to carrying containers, construction and drilling equipment, oversized equipment, and machinery, such as turbines for power plants. Deck barges are even used to carry components of the space shuttle from their construction site to the launch pad area. The roll on/roll off, or RO/RO, barge (**Fig. 2**) is a deck barge with multiple decks that can accommodate enormous numbers of containers, trailers, or rail cars, which can be wheeled directly onto or off the barge.

Over 27,000 of these non-self-propelled barges ply the commercial navigation system, offering a combined capacity of 5×10^7 tons (4.5×10^7 metric tons). These vessels receive propulsion power from approximately 4000 towboats and tugboats.

Towboats with barge flotillas commonly operate on the inland waterways, where they are fully protected by land on either side. The barges are kept together by wire rigging and are lashed against the towing knees of the towboat that pushes them ahead

(**Fig. 3**). The flotillas vary in size depending on the waterway on which they are operating. On the Upper Mississippi River above St. Louis, and on other waterways that are controlled by locks and dams, the standard-size flotilla is 15 barges and one towboat. On the Lower Mississippi, where the river is free of locks, flotillas can reach a half mile (0.8 km) in length, with 40 or more barges offering a cargo capacity of over 58,000 tons (52,000 metric tons).

As the flotilla grows in size, so does the amount of horsepower required from the towboat. The standard for a towboat operating on the Upper Mississippi is 6000 hp (4.5 MW). Heavy-weight-class towboats operating on the Lower Mississippi, where the load is much larger, have up to 10,500 hp (8 MW) and are about 200 ft (60 m) long.

The design of a towboat is vastly different from that of a tug. The towboat (**Fig. 4**) has a relatively flat bottom with a square bow on which heavy, upright towing knees are fixed. Towboats have conventional doors and windows, and on some the upper decks are retractable for easy passage under bridges. Tugboats, on the other hand, have curved hulls and bows, and are designed with watertight doors and portholes to prevent the entry of water (**Fig. 5**).

Fig. 2. Roll on/roll off (RO/RO) deck barge. (*The American Waterways Operators*)

Fig. 3. Inland tow, consisting of open and closed dry cargo barges, being pushed ahead by a towboat. (*The American Waterways Operators*)

Fig. 4. Inland towboat. (*The American Waterways Operators*)

Fig. 5. Articulated tug-barge used in coastal towing. (*The American Waterways Operators*)

Barges and towboats cannot withstand the wind and wave action encountered on open water. Therefore, on open channels, such as the Atlantic and Gulf intracoastal waterways, tugboats are used, which pull barges on a long towline, or hawser. The number of barges that can be handled by this method is less than can be handled in push-towing operations; however, barges involved in coastal towing often greatly exceed the inland barges in size. The state-of-the-art in coastal towing is the articulated tug-barge, intended to increase efficiency and safety by eliminating the long towline. Instead, the tugboat fits into a notch in the stern of the barge with the two units tightly connected. This allows for more control in steering the barge.

Carrying capacity. One barge carries up to five times its weight in freight. One 1500-ton (1350-metric-ton) barge load of soybeans equals 1167 acres (472 hectares) of farm production at 45 bushels per acre (4 m^3 per hectare). In relation to other modes of transportation, one 24-barge flotilla has the cargo capacity of 180 rail cars or 1440 trucks.

Cargo. Many of the cargoes carried by barge are essential materials, such as petroleum, coal, iron

and steel, chemicals, forest products, fertilizer, sand, and salt. Most commodities shipped by barge are those that lend themselves to bulk shipments and efficient cargo-handling techniques. Petroleum and petroleum products account for the largest percentage (31.9%) of cargo shipped by barge in the United States. Coal is second-largest, representing 25% of barge tonnage, followed by crude materials (forest products, minerals, ores, sand, stone, and gravel) and then food and farm products. Barges also transport ethanol, a new commodity market created by demand for alternative fuels. Other products shipped by barge include industrial chemicals, iron and steel products, sulfur, fertilizer, logs, and paper products. Cumbersome cargoes that cannot easily be shipped by other means are often moved by barge, including nuclear reactors and drilling platforms.

Barges transport over 60% of the nation's export grain, a significant amount of the coal that is delivered to domestic power plants and utilities (enough to produce 10% of the electricity used each year in the United States), and most of New England's home heating oil and gasoline.

Barge transportation is an important link in the nation's intermodal transportation network. For example, petroleum products are transported from Gulf Coast refineries by pipeline to New York and barged to New England power plants, where they are used to provide electricity, heat homes, and keep industry moving. Dry whey is trucked to Lewiston, Idaho, barged to Portland, Oregon, and shipped to Asia. Food, equipment, and supplies from the lower 48 states are moved by rail to Seattle, barged to Alaska, and trucked on the Alaskan Highway to the oil fields of the North Slope. The strength of these kinds of intermodal links to transport cargo is critical to the flow of the nation's commerce.

Waterways. Except for the Lower Mississippi, the Missouri, the Gulf Coast, and the Atlantic and Pacific waterways, the navigable rivers are slack water routes which have been improved for navigation by dredging and the construction of locks (**Fig. 6**) and dams. The St. Lawrence Seaway is a deep draft waterway that connects the Atlantic Ocean to the Great Lakes. It contains 13 Canadian and two United States locks. With the exception of the New York State Barge Canal, which is state-run, the inland waterways of the United States are maintained on a federal level by the Army Corps of Engineers. This includes the maintenance and modernization of the inland waterways locks and dams, the infrastructure that enables waterborne commerce. This infrastructure is critical, as failure of any of the aging locks on the Ohio River System or the Upper Mississippi River–Illinois Waterway System, for example, affects the flow of commerce on the entire interconnected waterways system. Many of these locks were built over 50 years ago when even the most optimistic projections did not match the traffic volume of today. Back-ups at these antiquated structures are common, and attention is being turned toward replacement and repair of these facilities. Since 1980, barge operators have paid a fuel tax that finances a percentage of the construction, operation, and maintenance costs for inland waterway projects. Before this time, all projects were entirely federally funded with appropriations from Congress.

Fig. 6. Barge tow going through a lock in the Pacific Northwest. (*The American Waterways Operators*)

Most of the navigable inland rivers are open to navigation year-round, with the exception of the northern waterways. Commonly, ice conditions close the Upper Mississippi, portions of the Ohio River System, the Missouri River, and the New York State Barge Canal from December through March. Occasionally, ice impedes navigation on the other waterways as well. *See* CANAL; DAM; RIVER ENGINEERING.

Thomas J. Allegretti; Joseph Farrell

Bibliography. J. Leventhal, *Tugs: The World's Hardest Working Boats*, Black Dog & Leventhal Publishers, New York, 1999; *National Transportation Statistics*, U.S. Government Printing Office, annually; W. Salinger, *Tugboats 'n Towlines: The Men and Women Who Give Them Life*, Twin Lights Publishers, Rockport, MA, 2003; U.S. Army Corps of Engineers, *Waterborne Commerce of the United States*, 2004.

Inoculation

The process of introducing a microorganism or suspension of microorganism into a culture medium. The medium may be (1) a solution of nutrients required by the organism or a solution of nutrients plus agar; (2) a cell suspension (tissue culture); (3) embryonated egg culture; or (4) animals, for example,

rat, mouse, guinea pig, hamster, monkey, bird, or human being. When animals are used, the purpose usually is the activation of the immunological defenses against the organism. This is a form of vaccination, and quite often the two terms are used interchangeably. Both constitute a means of producing an artificial but active immunity against specific organisms, although the length of time given by such protection may vary widely with different organisms. *See* VACCINATION.

Inoculation is the natural process of acquiring protection against disease, in that most persons are exposed to some organisms at times when no severe symptoms are displayed. The protective mechanisms of the body, especially antibody production, are stimulated by such a mild or insignificant exposure. An example of this is the discovery that the majority of adults have antibodies to poliomyelitis present, despite the absence of a history of the severe or recognizable disease form.

Inoculation may also refer to the deliberate seeding of organisms into culture media, and the introduction of fermenting bacteria, yeasts, or molds into various industrial processes that employ the chemical reactivity of these organisms. *See* BACTERIOLOGY; CULTURE; IMMUNITY; INDUSTRIAL MICROBIOLOGY.

Edward G. Stuart; N. Karle Mottet

Inorganic and organometallic polymers

Macromolecules that contain metals, metalloids, or other elements. The inorganic element may lie along the polymer backbone where it is covalently bonded to the adjacent atoms (**Fig. 1***a, e*); the inorganic element may be ligand-coordinated (Fig. 1*b, c*); or the inorganic element may be pendant to the polymeric chain (Fig. 1*d*). Organometallic polymers are a special class of inorganic polymers in which the metal, metalloid, or other element is covalently bonded to carbon (Fig. 1*d–f*).

Many examples of inorganic polymers are found in nature having important physiological functions such as respiration, photosynthesis, energy transfer, and metal-ion storage. Small proteins called metallothioniens store metals such as copper and zinc in higher animals. For example, cd_7-metallothionein-2 is a rat protein rich in cysteine that is bonded to four zinc atoms through sulfur-containing amino acids. Ferrodoxins are bacterial proteins containing up to eight iron-sulfur complexes that are responsible for biological oxidation-reduction and electron transfer reactions. In addition, mammalian ferritin (a high-molecular-weight protein containing approximately 4500 iron atoms) is essential for the storage and transfer of Fe(II) and oxidation to Fe(III).

Synthetic inorganic and organometallic polymers were developed in the early part of the twentieth century. One of the first useful inorganic coordination polymers was prepared by impregnating a plastic film with dithiooxamide (a tetradentate ligand), and treating the film with a solution of Cu(II) or Ni(II) salts. The resulting linear polymer [reaction (1)] had

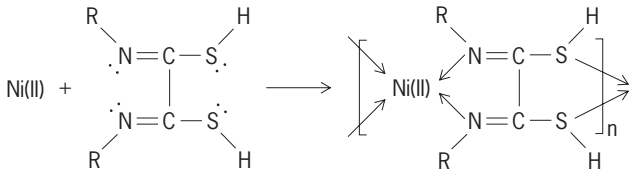

(1)

light-polarizing properties. In the 1960s, new developments in synthetic methods led to a number of polymers that contain main-group, transition, and inner transition elements. Many of these materials were designed to withstand high temperatures and resist erosion, while maintaining desirable attributes such as oxidative stability, elasticity, electric conductivity or resistivity, optical properties, fire retardancy, catalytic properties, low density, high tensile strength, and inertness to chemicals.

Classification. Inorganic and organometallic polymers may be classified in terms of morphology (chains, sheets, and networks) or backbone elemental composition (homoatom and heteroatom).

Chain polymers. A chain polymer is a one-dimensional (1D) macromolecule composed of linear repeating units that lack inter- or intrachain crosslinks. Examples are poly(dialkylstannane) (structure **1**) and polythiazyl (**2**). Under certain circumstances, the

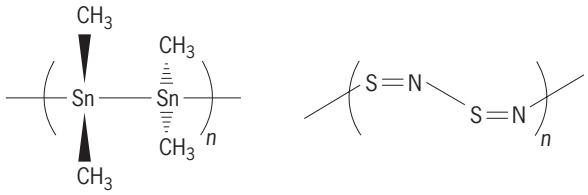

(**1**) (**2**)

ends of linear chains connect to form macrocyclic structures. Thus, cyclic poly(methylhydrosiloxane) with as many as 50 —CH_3SiH— repeat units is known. Cyclic chains with a relatively small number of repeat units (oligomers) are very common. Some well-known cyclic oligomers are octamethylcyclotetrasiloxane, $(Me_2SiO)_4$ (**3**), and several allotropic forms of sulfur, for example, S_8 (**4**). Linear 1D chain

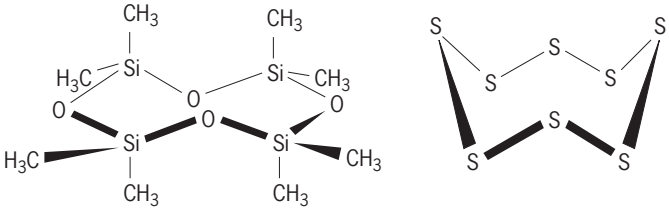

(**3**) (**4**)

polymers are also known to assume random coil, helical (proteinlike), and rigid rod structures.

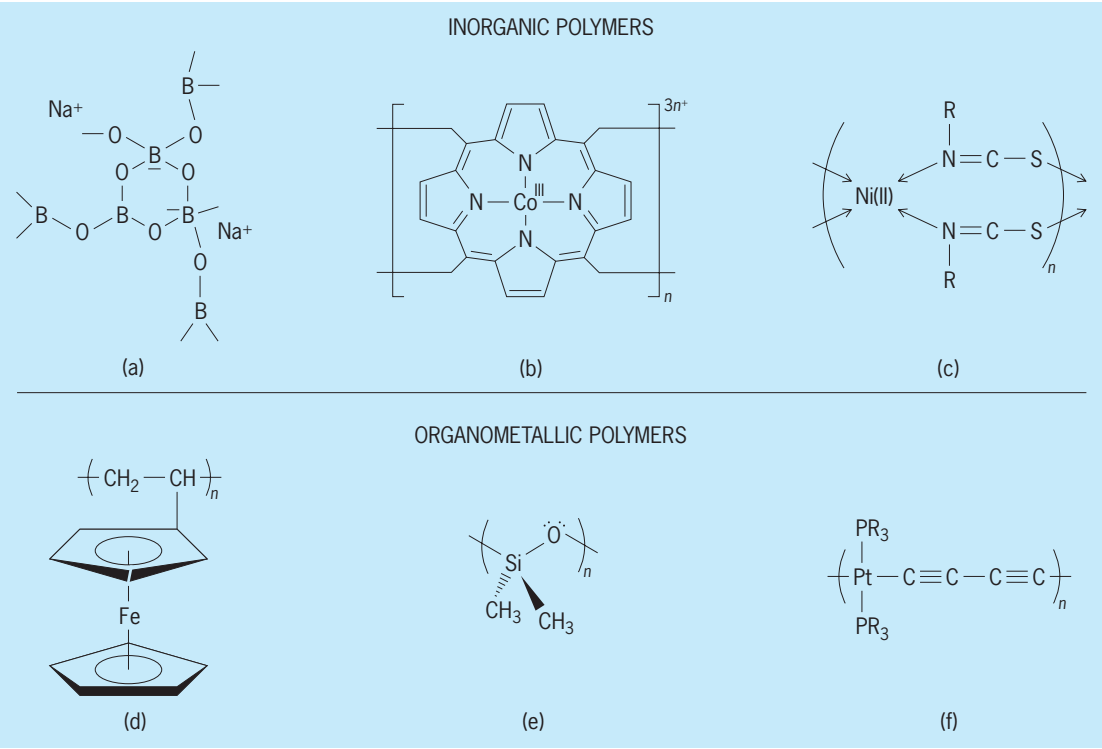

Fig. 1. Inorganic polymers: (*a*) borate polymer (metalloid in the backbone), (*b*) cobaltoporphyrin polymer (ligand-coordinated metal), and (*c*) nickel(II) dithiooxamide polymer (ligand-coordinated metal). Organometallic polymers: (*d*) ferrocene polymer (metalloid in backbone), (*e*) poly(dimethylsiloxane) (metalloid in backbone), and (*f*) platinum-diacetylene polymer (metal in backbone).

Sheet polymers. Sheet polymers are two-dimensional (2D) macromolecules in which the primary valences of the metal/metalloid atoms are satisfied by covalent bonds. Boron nitride (α-BN) is an example of a 2D sheet polymer (**Fig. 2***a*) in which each atom is bonded to three other atoms in the same layer. The stacked layers (lamellae) interact through van der Waals forces. Another variation of a 2D sheet occurs if two chains are crosslinked at regular intervals to give a ladder structure such as the crosslinked siloxane polymer (Fig. 2*b*). Ladder polymers are unusually high-melting solids with oxidative, hydrolytic, thermal, and chemical stability.

Network polymers. Network polymers are highly crosslinked materials in which the metal/semimetal atom valences are satisfied by bonds that result in a three-dimensional (3D) structure. Such polymers are usually difficult to characterize because of their refractory (ceramic-like) nature. Network polymers are often hard, infusible, and insoluble substances such as silicon dioxide and silicon nitride (Fig. 2*c*). These materials are added to organic polymers to enhance their strength and thermal stability.

Homoatom polymers. Poly(diphenylsilane), $(Ph_2Si)_n$, is an example of a homoatom polymer because it contains only silicon atoms along the backbone. Some

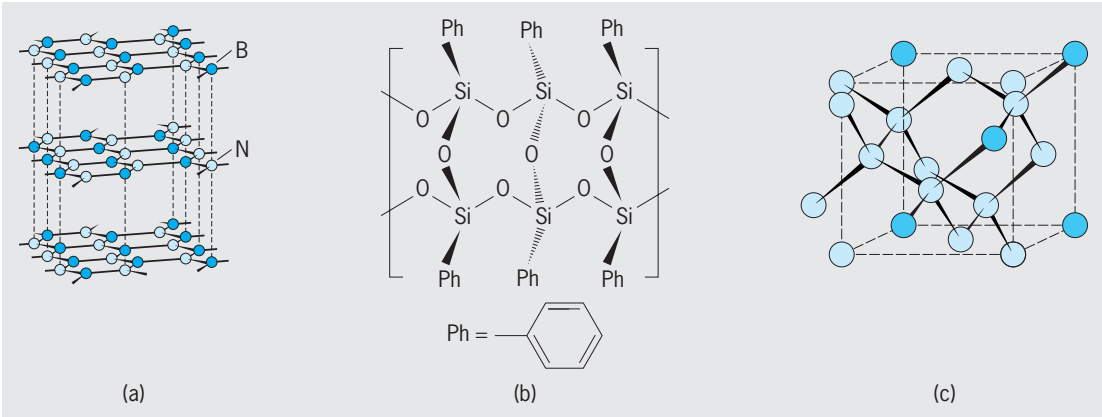

Fig. 2. Morphological classes of polymers: (*a*) boron nitride (α-BN), 2D parallel sheet; (*b*) poly(diphenylsiloxane), 2D ladder; and (*c*) silicon dioxide (silica), 3D network.

TABLE 1. Polymeric metal boride structures

Formula	Metal	Morphology	Anion structure
MB	Fe, Ni	Zigzag chains	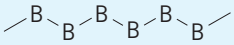
$M_{11}B_8$	Ru	Branched chains	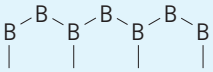
M_3B_4	Cr, Ta	Ladder	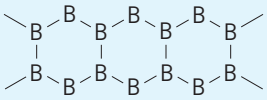
MB_2	Mg, Ti, Zr, Gd	Sheet	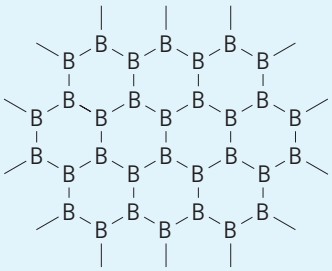
MB_4	La	Network (octahedron)	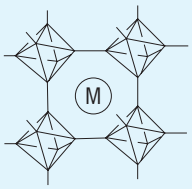 M = metal; B at each vertex of octahedra

metalloid anions have homoatomic polymeric structures. For instance, metal borides of formula M_xB_y contain the anion as chains, ladders, 2D sheets, or 3D networks (**Table 1**). These anionic polymers possess high thermal stability, excellent chemical resistance, and potentially useful electrical properties.

Certain elements, such as phosphorus, prefer to aggregate as low-molecular-weight oligomers. For example, white phosphorus is a tetramer (P_4) with phosphorus atoms at the corners of a tetrahedron. Black phosphorus, which is obtained by heating P_4 under pressure [200°C (390°F), 12,000 atm (1220 MPa)] with a catalyst, is a sheet polymer with each phosphorus atom bound to three neighbors in double layers. This giant molecule consists of stacked double layers, analogous to the boron nitride lamellae. In addition, black phosphorus is a good electrical conductor. Red phosphorus is made by heating P_4 at 400°C (750°F) in a sealed vessel. One crystalline form of red phosphorus has a linear, tubular arrangement of eight phosphorus atoms in a pentagonal wedge structure.

Elemental sulfur is also a polyatomic and polymorphic substance. Oligomeric sulfur rings consisting of 6, 7, 8, and 12 sulfur atoms have been identified. The orthorhombic α-S_8 form is the most thermodynamically stable and consists of eight-member puckered rings in a crown conformation stacked together in "crankcase" fashion. The chain polymer form of sulfur (catena or plastic sulfur) is obtained by quenching molten S_8 in water. If stretched, catena sulfur has a helical structure. Selenium forms similar polymers.

Heteroatom polymers. Heteroatom polymers contain a metal or semimetal in addition to one or more other atoms (such as oxygen, sulfur, nitrogen, phosphorus, and carbon) in the backbone. In general, polymer stability and inertness increases when the heteroatom is electron-rich. Presumably, such atoms serve as electron sources toward low-lying vacant orbitals on the metal or metalloid (structure **5**). The resulting

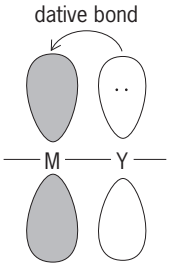

M = metal, semimetal
Y = heteroatom with a lone pair of electrons

(5)

back-bonding (dative bond) generally decreases the electrophilicity of the metal or metalloid and reduces the susceptibility of the polymer to attack by Lewis bases. In addition, the interaction tends to stabilize the polymer toward thermal degradation and aids in the regulation of structure, chain mobility, and interchain forces that enhance its dynamic and mechanical properties.

Synthetic methods. A number of methods have been developed to synthesize inorganic and organometallic polymers. Representative examples of the more important and commonly used preparative techniques include step-growth (condensation) polymerization, ring-opening polymerization, chain-growth (addition) polymerization, and reductive (Wurtz-type) coupling polymerization.

Step-growth polymerization. In step-growth polymerization, two molecules combine with the evolution of a small molecule such as water, ammonia, or a salt (for example, hydrogen chloride). For example, dihydroxymetallophthalocyanine [prepared from a metalloid halide such as silicon tetrachloride ($SiCl_4$) or germanium tetrachloride ($GeCl_4$) and phthalocyanine followed by hydrolysis] undergoes dehydration under vacuum at elevated temperatures to give a linear, 1D "shish-kebab" polymer with an —M—O—M—O— chain backbone [reaction (2)]. This material has excellent thermal and chemical stability, solubility in strong acids, interesting magnetic and optical properties, as well as electrical (semiconducting) properties when doped with iodine.

Another step-growth polymerization uses an inorganic or organometallic compound that contains a difunctional organic ligand. For example, diiodoferrocene reacts with a dilithioferrocene [reaction (3)] to give a polyferrocene. The driving force for the reaction is the precipitation of the salt, lithium iodide. Similarly, condensation with bridging difunctional reagents, such as the sodium salt of an organic diol, can produce metal-containing polymers [reaction (4)].

Ring-opening polymerization. Ring-opening polymerization is a principal method of preparing inorganic and organometallic polymers. The method is convenient and usually requires heat, light, or a catalyst. The reaction yields high-molecular-weight polymers without condensation products. An example of a ring-opening polymerization is the synthesis of polysiloxanes (silicones), which have many

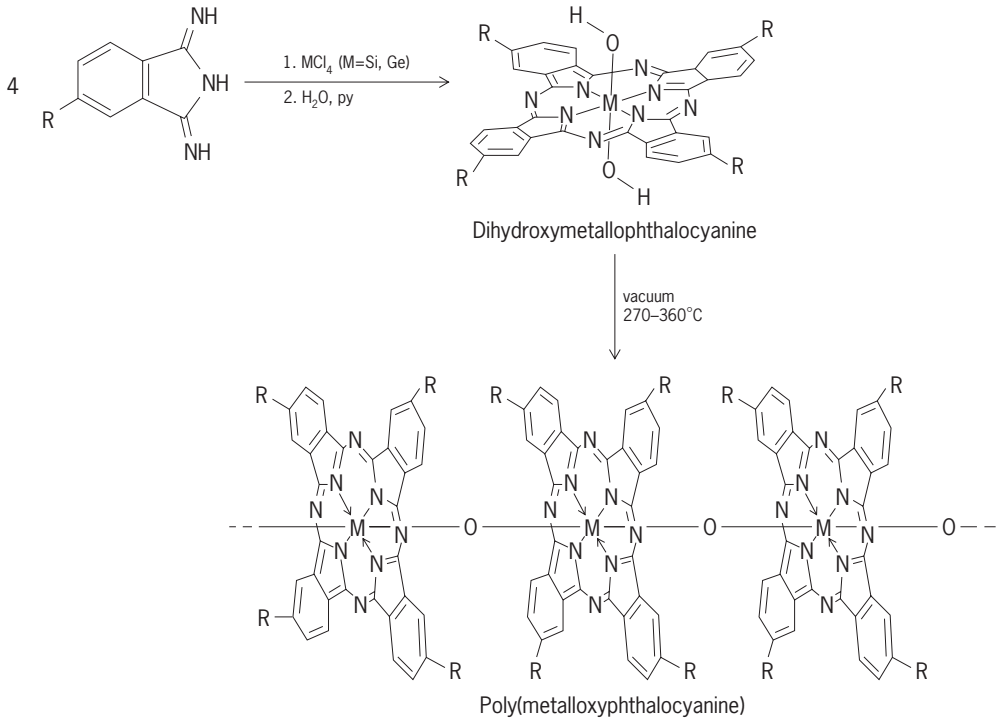

(2)

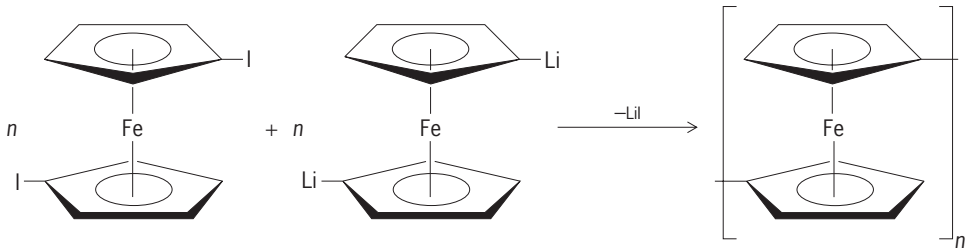

(3)

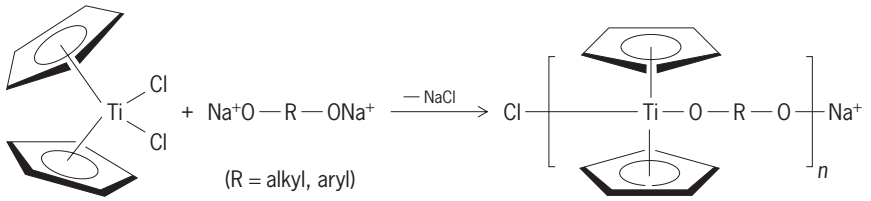

(4)

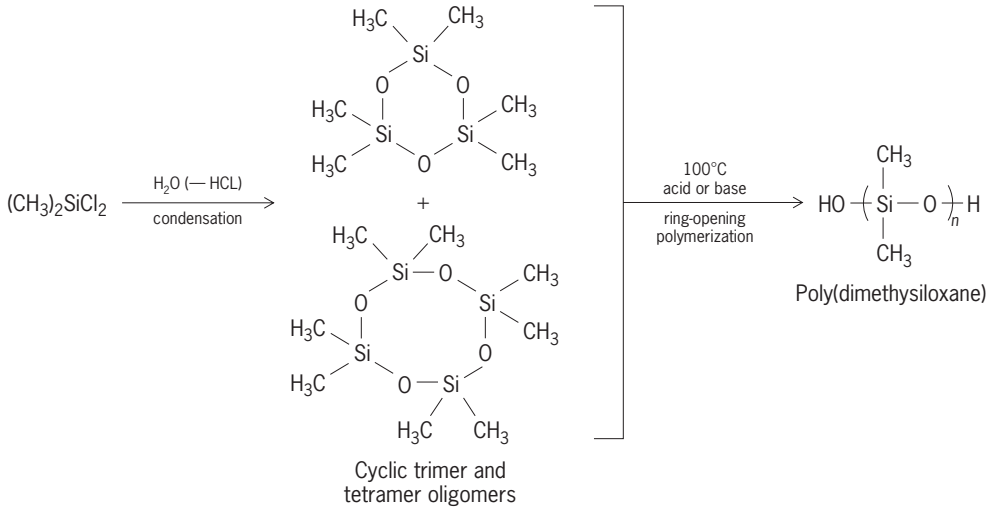

(5)

industrial and consumer applications, including:

Insulators	Masonry additives
Dielectric materials	Surfactants
Heat-exchange fluids	Paper release coatings
Seals and gaskets	Coupling agents
Caulking agents	Water repellants
Emulsifying agents	Lubricants
Surgery implants	Hydraulic fluids
Mold-forming agents	Ceramic composites
Antifoaming agents	Pressure-sensitive adhesives

The industrial process involves the hydrolysis of dimethyldichlorosilane followed by acid- or base-catalyzed ring-opening polymerization [reaction (5)] to give poly(dimethylsiloxane). The ring-opening polymerization method permits excellent control of molecular weight and polydispersity (molecular-weight distribution). **Table 2** summarizes other examples of rings and polymers that can be prepared by ring-opening polymerization.

Chain-growth polymerization. Chain-growth polymerization is initiated with free radicals, which may be generated chemically, thermally, photochemically, electrochemically, as well as with ions (cations or anions) or coordination complexes (for example, Ziegler-Natta or ferrocene catalysts). Invariably, chain-growth polymerization requires that the monomer contain an unsaturated functionality, usually an olefin. As a result, the polymers contain carbon in the chain backbone. Some examples of chain-growth polymerization that incorporate metals or metalloids into the polymer are given in **Fig. 3**.

Reductive (Wurtz-type) coupling polymerization. Reductive coupling polymerization uses active metals (such as sodium or potassium) in a reaction with difunctional, halogen-containing organometallic compounds (for example, organosilanes, organogermanes, organostannanes, carbometallanes, and ferrocenes). Typically, the reactions are carried out at a temperature that melts the alkaline metal and disperses it in a nonreactive solvent such as toluene or tetrahydrofuran. Alternatively, the reducing agent is made soluble with a chelating (complex-forming) agent such as crown ether.

For background information *see* BORON; COORDINATION COMPLEXES; GERMANIUM; INORGANIC POLYMER; LIGAND; METALLOID; NONMETAL;

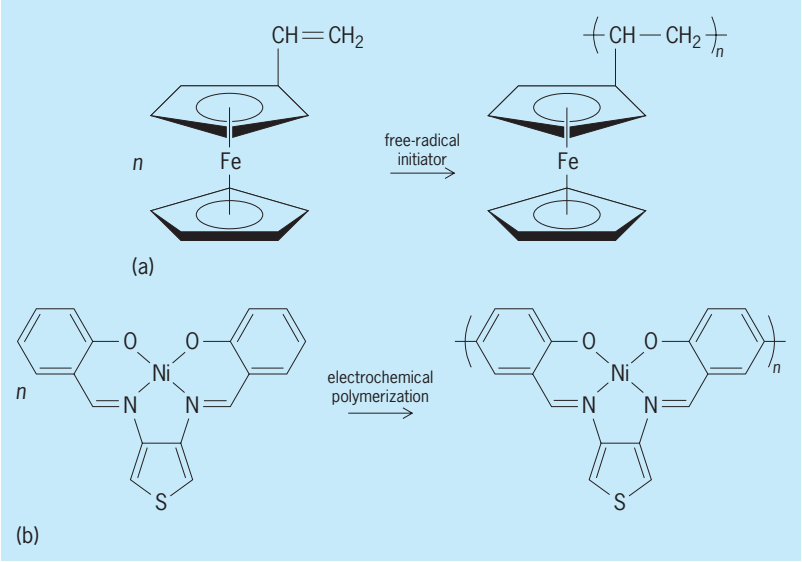

Fig. 3. Chain-growth polymerization synthetic routes. (*a*) Free-radical initiated. (*b*) Electrochemical polymerization.

TABLE 2. Some rings and polymers formed by ring-opening polymerization

Ring	Polymer	Name
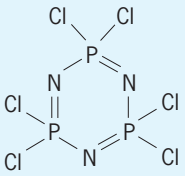	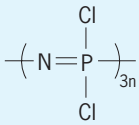	Polyphosphazene
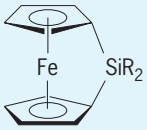	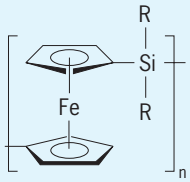	Poly(diorganosilyl ferrocene)
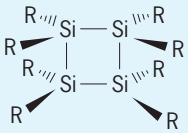	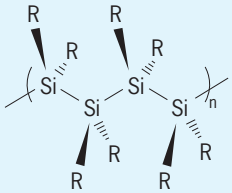	Poly(diorganosilane)
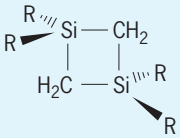	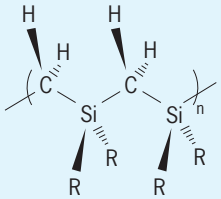	Poly(diorganocarbosilane)

ORGANOMETALLIC COMPOUND; POLYMER; POLYMERIZATION; SILICON; SULFUR in the McGraw-Hill Encyclopedia of Science & Technology. Martel Zeldin

Bibliography. R. D. Archer, *Inorganic and Organometallic Polymers*, 2001; I. Bertini et al. (eds.), *Bioinorganic Chemistry*, 1994; V. Chandrasekhar, *Inorganic and Organometallic Polymers*, 2005; J. E. Mark, H. R. Allcock, and R. West, *Inorganic Polymers*, 2d ed., 2005; C. U. Pittman, Jr., et al. (eds.), *Metal-Containing Polymeric Materials*, 1996.

Inorganic chemistry

The chemical reactions and properties of all the elements in the periodic table and their compounds, with the exception of the element carbon. The chemistry of carbon and its compounds falls in the domain of organic chemistry. The boundaries of inorganic chemistry with the other major areas of chemistry are not precisely defined, and it is often a matter of taste as to whether a particular topic is to be included in the field of inorganic chemistry or is to be considered physical or even organic chemistry. Physical chemistry may be defined as the application of quantitative and theoretical methods to chemical problems, and is a methodology rather than a specific body of knowledge. Investigations into theoretical inorganic chemistry or the study of problems in inorganic chemistry by quantitative and sophisticated physical methods may be considered either inorganic or physical chemistry quite arbitrarily. In similar fashion, organometallic compounds may be considered to be in the sphere of either inorganic or organic chemistry. The modern inorganic chemist is concerned with problems that once were considered the prerogative of physical chemists, organic chemists, or even biochemists. *See* PHYSICAL CHEMISTRY.

Synthetic inorganic chemistry. The reactivity of the elements of the periodic table varies enormously, and over a much wider range than is encountered in organic chemistry. Consequently, the inorganic chemist must frequently employ unusual apparatus and techniques. The elements range from the rare gases, which are unreactive and form very few chemical compounds, to the extremely reactive halogens and alkali metals. Fluorine is perhaps the most reactive element known; it forms compounds with all other elements, including the rare gases. Because of the great reactivity of fluorine and the closely related halogen fluorides, special metal and plastic apparatus must be used in experimentation. Both fluorine and the important compound hydrogen fluoride attack glass, and this common material of construction

cannot be used for experimentation with these substances. There are methods to effect direct reaction between elemental fluorine and organic compounds without completely destroying the organic reactant, and this procedure has greatly increased the number of fully fluorinated organic compounds that can be readily obtained. *See* FLUORINE; FLUOROCARBON; HYDROGEN FLUORIDE; PERIODIC TABLE.

Another element important in synthetic inorganic chemistry since World War II is boron. The hydrides of boron were first obtained by reaction of a metal boride with a solution of an aqueous acid. The procedure was very difficult and tedious, and it required weeks or months of labor to obtain a few cubic centimeters of the gaseous product, which turned out to be the simplest boron hydride, B_2H_6 (diborane). Many years later it was discovered that yields could be increased by passing boron trichloride and hydrogen through an electric discharge, but the yields were still distressingly low. Under the impetus of wartime urgency, chemical syntheses for the boron hydrides were developed. Using the readily available alkali metal hydrides and boron trifluoride as starting materials, boron hydrides were obtained in very good yield. The metal borohydrides, such as $Al(BH_4)_3$ (bp 44.5°C or 112.1°F), when they can be prepared, are likely to be the most volatile compounds of the metal. *See* BORANE; BORON.

The carboranes are a relatively new class of organoboron compounds with a rich chemistry and many potential applications. Hydroboration, the addition of diborane to organic compounds containing a carbon-carbon double bond, is a preparative reaction that has many important applications in organic synthesis. Many compounds of elements such as silicon, germanium, gallium, phosphorus, and nitrogen, some of which were at one time considered too unstable to be synthesized, have been prepared by the synthetic inorganic chemist. An outstanding achievement in modern inorganic synthesis was the preparation of the noble gas fluorides. The discovery that the noble gases could form compounds with fluorine and oxygen resulted in a whole new area of inorganic research. *See* CARBORANE; HYDROBORATION; INERT GASES.

Many boron compounds react violently with air and water. It is necessary, therefore, to use all-glass vacuum systems for carrying out experiments with these substances. Vacuum-line techniques are widely used in inorganic chemistry for the manipulation of volatile, highly reactive compounds such as the hybrides of phosphorus, silicon, and related compounds, and an array of vacuum lines often is the hallmark of a laboratory for synthetic inorganic chemistry.

The synthesis of inorganic compounds as an end in itself, which at one time was perhaps the most characteristic of the activities of the inorganic chemist, has been largely superseded by synthesis directed to specific goals; among these are compounds required to elucidate fundamental laws of chemical bonding, reactivity, and structure; inorganic polymers with unusual thermal and mechanical properties; compounds or systems with superconductor or electron conductor properties; and compounds with particular virtues as reagents in organic syntheses.

Coordination chemistry. The transition elements in the periodic table form coordination compounds with a great variety of organic donor molecules. Traditionally, coordination chemistry has been a major interest of inorganic chemists. However, interest in maintaining a large flow of new coordination compounds whose principal virtue is that they can be easily prepared has considerably diminished in most quarters. The principal interest in coordination chemistry is in characterizing the physical properties of coordination compounds with unusual properties or uses. Thus, the synthesis of multidentate ligating agents such as the crown ethers and the cryptates, which have high specificity for binding particular ions, became a very important activity. Macrocyclic polyethers and macrobicyclic polyethers with tertiary amino groups at the bridgeheads mimic many of the properties of ion-transporting agents in living systems. The template properties of transition-metal ions for the synthesis of multidentate ligands such as the porphyrins and corrins that occur in nature have also received considerable attention. These and other initiatives in coordination chemistry have become an important component in the discipline of bioinorganic chemistry. *See* COORDINATION CHEMISTRY; MACROCYCLIC COMPOUND.

Organometallic compounds. Organometallic compounds constitute a borderline area of study between inorganic and organic chemistry. Until 1950, the organometallic compounds known were entirely compounds of the principal families in the periodic table; no stable derivatives of subgroup metals or transition metals had been prepared. Today a very large number of organometallic compounds of the transition and subgroup elements with metal-carbon bonds to cyclopentadiene, C_5H_6, aromatic hydrocarbons such as benzene, C_6H_6, as well as other organic compounds, have been synthesized. These compounds form interesting derivatives with carbon monoxide, CO, and can also undergo numerous reactions, and the structure and properties of these compounds pose interesting theoretical problems. Homogeneous catalysis by organometallic compounds has become one of the most active areas of catalysis research. Hydride complexes of ruthenium, rhodium, and iridium are catalysts for hydrogenation, olefin isomerization, hydroformylation, and olefin polymerizations. Much effort has been devoted to the development of homogeneous organometallic catalysts for the reduction of carbon monoxide by hydrogen; these catalysts are intended to be more selective and to function under milder conditions than those required in heterogeneous catalysis. Organometallic reagents are also being sought for highly selective syntheses of complicated molecules such as pharmaceuticals employed in therapy. Palladium complex compounds are being used in the synthesis of alkaloids, and the preparation of such important classes of compounds as the insect pheromones and the prostaglandins is being facilitated by

the use of nickel and aluminum organometallic compounds. *See* ALKALOID; EICOSANOIDS; HOMOGENEOUS CATALYSIS; ORGANOMETALLIC COMPOUND; PHEROMONE.

Solid-state chemistry. The elements and compounds that form the subject matter of inorganic chemistry exhibit a very wide range of physical properties which run the gamut from helium, the substance of lowest known melting and boiling point, to elements such as tungsten, titanium, and carbon, which are among the most refractory high-melting substances known. The study of solid-state reactions and compounds that can be prepared by them has assumed great importance, and is another major research area in inorganic chemistry. Unlike gaseous compounds and most organic compounds, which obey the law of definite proportions, many solid compounds, particularly those of the transition elements, exhibit variability of composition, or as it is frequently designated, nonstoichiometry. When a solid compound deviates from simple stoichiometric relations, it contains an excess of either positively charged metal cations or negatively charged anions. Such solid systems frequently show unusual electronic properties, which are made use of in many solid-state devices. Photovoltaic cells, for example, which convert light to electricity, belong to this group of nonstoichiometric solid-state devices. *See* NONSTOICHIOMETRIC COMPOUNDS; REFRACTORY.

Transistors, thermistors, phosphors, and light-emitting diodes are also important solid-state devices. Because these materials are frequently prepared by solid-phase reactions, high temperatures may be required for the reaction. High-temperature chemistry has greatly expanded in scope, not only in the preparation of nonstoichiometric compounds, but also in the preparation of refractories useful at very high temperatures in space and nuclear technology. Although many reactions can be forced to proceed in the desired direction by only increasing the temperature, this is not always adequate in all instances. Another dimension in solid-state chemistry was added by the simultaneous use of high temperature and extremely high pressures. With the equipment available, it is possible to carry out solid-state reactions at temperatures of approximately 2500°C (4532°F) and pressures of the order 10^5 millipascals (10^6 bars). Under conditions of high temperature and pressure, ordinary carbon (graphite) can be converted into diamond, and since the pressure is actually sufficient to distort the electron orbitals, new varieties of matter can be prepared. For example, liquid hydrogen has been converted to a metallic form at these ultrahigh pressures. *See* HIGH-PRESSURE CHEMISTRY; HIGH-TEMPERATURE CHEMISTRY; LIGHT-EMITTING DIODE; THERMISTOR; TRANSISTOR.

The solid-state systems used for semiconductors, photovoltaic devices, and laser photodiodes have traditionally been highly ordered, crystalline materials. However, substances in the amorphous or glassy state have been found to possess very interesting properties for similar uses. Amorphous silicon appears to have considerable potential in solar energy conversion cells for direct production of electricity, and amorphous silicon and germanium have semiconductor properties quite analogous to their crystalline counterparts. Amorphous vanadium or molybdenum sulfides can store lithium ions in nonaqueous lithium storage batteries better than crystalline materials can. Amorphous systems are only imperfectly understood, but they possess much scientific and technical interest and can, therefore, be expected to attract continued scientific investigation. *See* AMORPHOUS SOLID; METALLIC GLASSES; SEMICONDUCTOR.

Layered materials were discovered that have unusual properties, and this discovery revived interest in the synthesis and study of layered compounds. Graphite intercalation compounds, for example, consist of layers of carbon atoms that have halogen or other atoms or molecules inserted between the carbon layers. Certain of these intercalation compounds have an electrical conductivity comparable to that of metallic copper, and the possibility that such systems could replace metallic electrical conductors has aroused much interest. Intercalated room-temperature superconductors are also being sought. Intercalation compounds derived from graphite or from transition metal sulfides are effective and highly selective reagents in organic synthesis. Many layered compounds can accommodate layers of ions with high mobility. These systems are in effect solid electrolytes or charge carriers and may contribute significantly to the long-sought lightweight high-capacity electric storage battery.

Still another class of electronic conductors comprises chains and clusters of metal atoms. Linear chains of transition-metal atoms (Krogmann salts) are one-dimensional electron conductors. Metal cluster compounds also have extremely interesting electrical properties. Organometallic compounds containing metal clusters with as many as 28 platinum metal atoms and 44 carbonyl ligands have been made. The metal atoms in the cluster are directly bonded to each other and exhibit many of the properties of solid metal catalysts. The clusters often exhibit superconductivity at higher than customary temperatures, and have, therefore, become a point of departure in the search for new superconductive materials that may be suitable for electric power transmission at liquid hydrogen rather than liquid helium temperatures. Metal cluster chemistry ranks among the most promising of the approaches being explored in the search for new catalysts. *See* METAL CLUSTER COMPOUND; SOLID-STATE CHEMISTRY; SUPERCONDUCTIVITY.

Geochemical aspects. Many of the synthetic procedures in inorganic chemistry carried out at high temperatures and pressures have considerable interest in geochemistry. Some of the earliest inorganic chemistry was practiced in connection with mineralogy. Mineral syntheses, or the preparation of inorganic compounds identical with those found in nature, provide important information to the geochemist. Not only has such research helped to explain the sequence of chemical reactions and conditions

responsible for the formation of minerals in nature, but many minerals and gems such as diamond, ruby, sapphire, quartz, and corundum are now manufactured on the large industrial scale. High-temperature reactions at very high pressures or hydrothermal reactions at high temperatures and pressure are employed for this purpose. With the advent of ultrahigh-pressure equipment, it became feasible to study chemical reactions under conditions approximating those many miles below the Earth's surface, and such studies are expected to add greatly to the understanding of geochemical phenomena. *See* GEM; GEOCHEMISTRY; HIGH-PRESSURE MINERAL SYNTHESIS; HIGH-PRESSURE PROCESSES.

Nuclear science and energy. The development of nuclear energy since World War II has provided a great impetus to inorganic chemistry. The discovery of the transuranium elements was one of the outstanding events in chemical science, and opened up an entirely new area of the periodic table for investigation. The chemistry of the actinide 5*f* elements has revealed many surprises, and despite the great amount of new chemistry that has already been acquired, much still remains to be learned about the very complex phenomena associated with the oxidation-reduction behavior, the ions in solution, and the metallic states of the actinide elements. A particularly active field of actinide element research has been the synthesis of organometallic compounds of the 5*f* elements. The solvent-extraction and ion-exchange procedures developed for separating the actinide elements have had widespread applications in other areas of inorganic chemistry. Nuclear technology has also provided the impetus for the development of other separation procedures, for example, the separation of zirconium and hafnium, and of the rare-earth elements from each other, and has generally served to reinforce the traditional interest of the inorganic chemist in separations procedures. The development of safe, effective processing methods for the intensely radioactive spent nuclear fuels from breeder reactors, and the solution to the problem of safe storage and disposal of nuclear waste, remain paramount challenges to the inorganic chemist. *See* ACTINIDE ELEMENTS; ION EXCHANGE; NUCLEAR CHEMISTRY; NUCLEAR FUELS REPROCESSING; RARE-EARTH ELEMENTS; SOLVENT EXTRACTION; TRANSURANIUM ELEMENTS.

Applications in organic chemistry. Many of the most important advances in organic chemistry since 1900 resulted from the introduction of inorganic substances as reagents. Synthetic reactions based on magnesium metal gave rise to the vast corpus of Grignard chemistry, and the investigation of the metal carbonyls provided the impetus for the development of acetylene chemistry. Hydroboration, a process for producing organoboranes, can be used for stereospecific syntheses, asymmetric syntheses, and isomerizations. Hydroboration must be reckoned as among the most versatile synthetic procedures added to the armory of the organic chemist. Other inorganic substances that have found important use in inorganic chemistry are selenium for dehydrogenation reactions; lead tetraacetate and thallium compounds for selective oxidations; aluminum chloride as a catalyst for alkylation, acylation, and ring-closure reactions; anhydrous hydrogen fluoride for diazotization, nitration, and sulfonation; and lithium aluminum hydride (and various of its derivatives) and alkali-metal borohydrides for selective reduction reactions. Metal or carbon atoms in the gas phase are remarkable reagents for gas-phase organic reactions. *See* GRIGNARD REACTION; METAL CARBONYL; ORGANIC CHEMISTRY; ORGANIC SYNTHESIS.

Reaction kinetics and mechanisms. Organic reactions generally proceed with the skeleton of the molecule remaining intact. Inorganic gas-phase reactions, on the contrary, are usually characterized by a complete disruption of molecular structure followed by reorganization to form the products of the reaction. Gas-phase reactions of inorganic compounds are thus in principle more complicated than are the usual organic reactions. Electron transfer in the oxidation-reduction reactions of transition-metal ions and compounds in solution has been a particularly important topic of theoretical interest. As is the case with electron transfer reactions in general, the role of electron tunneling in inorganic redox reactions still has many obscure features. This is particularly true of the very complex redox reactions of the actinide elements. Hydrolytic reactions of highly charged ions in aqueous solutions are also important in the mechanisms of many inorganic reactions and are important aspects of modern transition-element and actinide element chemistry. These subjects have become the concern of physical and theoretical chemists, but they also continue to be an important part of contemporary inorganic chemistry. *See* CHEMICAL DYNAMICS; OXIDATION-REDUCTION.

Bioinorganic chemistry. Inorganic chemistry is under pressure, as are other areas of chemical research, to escalate attention from simple molecules to more complicated systems. Like the organic chemist, the inorganic chemist has turned to biology as a new field to explore, and much of the most interesting contemporary work in organic chemistry is directed to the inorganic chemical aspects associated with living organisms.

Metal ions are very important participants in many biological phenomena. These range from the role of trace elements as essential nutrients to such questions as the function of metal ions in respiratory and photosynthetic pigments. Other important questions relate to the function of the metal ion in metal-containing enzymes involved in oxidation or reduction reactions, and to the transport of metal ions across membranes. Platinum and other metal coordination compounds have been investigated for possible use in the therapy of cancer. The basis for the use of lithium ion to control manic-depressive fluctuations in mood also poses important questions for the inorganic chemist.

A major activity in bioinorganic chemistry is the preparation and study of model systems intended to mimic the behavior of important biological entities. Thus, the construction of model systems

for the prosthetic group of respiratory pigments (hemoglobin, myoglobin), oxidative enzymes (peroxidase), electron transfer proteins (cytochrome, ferredoxin, plastocyanin), and enzymes such as nitrogenase (the nitrogen-fixing enzyme) is under very active investigation by inorganic chemists.

The role of zinc ion in such enzymes as carbonic anhydrase and proteolytic peptidases has also focused attention on the inorganic components of these and other important enzyme systems. The application of modern spectroscopic techniques, particularly infrared and magnetic resonance spectroscopy, has suggested a role for the central magnesium atom of chlorophyll in the light-energy conversion step of photosynthesis. *See* BIOINORGANIC CHEMISTRY; SPECTROSCOPY.

Inorganic technology. The production of inorganic chemicals is a basic aspect of the chemical industry, and the heavy inorganic chemicals sulfuric acid, ammonia, chlorine, and phosphoric acid provide indispensable materials for many industries. The production of glass, ceramic, cement, fertilizer, and metals essentially involves reactions of inorganic chemistry carried out on the large scale. These industries have become less traditional in nature because of the contributions from inorganic chemical research, and it is altogether likely that modern inorganic chemistry will continue to make critically important contributions to the continued evolution of inorganic technology. Joseph J. Katz

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; G. L. Eichhorn and L. G. Marzelli (eds.), *Advances in Inorganic Biochemistry,* vol. 6, 1985; N. N. Greenwood and A. Earnshaw, *Chemistry of the Elements*, 2d ed., 1997; C. E. Housecroft and A. G. Sharpe, *Inorganic Chemistry*, 2d ed., 2004; J. E. Huheey, E. A. Keiter, and R. L. Keiter, *Inorganic Chemistry: Principles of Structure and Reactivity*, 4th ed., 1997; K. Mackay, A. Mackay, and B. Henderson, *Introduction to Modern Inorganic Chemistry*, 6th ed., 2002; G. L. Miessler and D. A. Tarr, *Inorganic Chemistry*, 3rd ed., 2003.

Inorganic photochemistry

The study of the light-induced behavior of various metal compounds. The physical and chemical properties of substances are generally altered by the absorption of light. Typical metal compounds have a characteristic number (coordination number) of molecules or ions (ligands) directly bonded to the metal center. Many of these compounds are colored, and it is possible that some metal compounds could mediate the transformation of solar radiation into useful chemical or electrical energy, or that they might scatter light in such a way that its frequency is increased. *See* LIGAND; NONLINEAR OPTICS.

The photochemistry of metal compounds has grown in concert with modern theories of the electronic structure of molecules and of chemical bonding in molecules; photochemical studies are often designed to probe and test these theories. The range of studies spans some of the subdisciplines of chemistry, and includes or bears on such topics as photophysics, the development of laser materials, catalysis, physics, photosynthesis, oxidation-reduction chemistry, acid-base chemistry, organometallic chemistry, metalloenzyme chemistry, solid-state chemistry, and surface chemistry. *See* CHEMICAL BONDING.

Excited states. The absorption of light results in a rearrangement of electrons within a molecule. In many molecules, the new electron configuration can persist for a significant period of time. Chemical properties depend on electronic configurations, and the ground-state and excited-state electronic configurations differ. In general, the bond lengths (angles and so on) of the excited state of a molecule will be different from those of the ground-state (or thermally equilibrated) molecule. Consequently it is useful to regard these excited states of molecules as new chemical species with chemical properties distinctly different from those of the ground state. Some of the energy used to generate an excited state is degraded to heat when the excited molecule relaxes to a bonding arrangement compatible with the new electronic configuration. In addition, the initial excited state (*X) may rapidly convert (or cross) to a lower-energy excited state (*Y) with yet another electronic configuration (see **illus.**). After light absorption in metal compounds, the time required for generation of the lowest-energy excited state is usually very short (less than a nanosecond). As a consequence, any chemistry due to the higher-energy excited states (*X; see illus.) must occur very quickly, and either is intramolecular or involves the nearest neighbors of the excited molecule in condensed phases. The lowest-energy excited states of some metal compounds (*Y; see illus.) can exist for nearly a millisecond. This is long enough for many

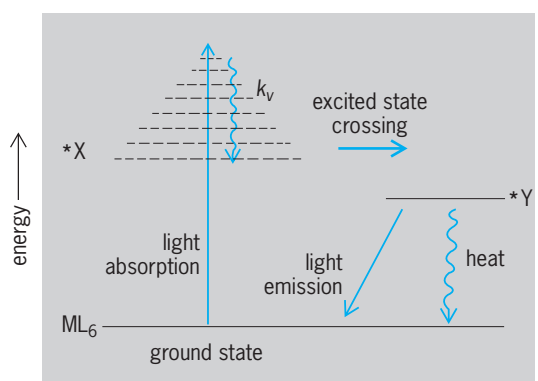

Qualitative electronic energy level scheme for a six-coordinate metal compound (ML_6). Absorption of light by ML_6 results in an excited state $^*(ML_6)$ in which electrons are rearranged (either *X or *Y). Any bond length difference between ML_6 and $^*(ML_6)$ results in the intial excited state being vibrationally excited (bonds stretched or compressed from the equilibrium position). Vibrational relaxation k_v is rapid in condensed phases.

collisions in condensed phases, and such excited-state species can often react with other molecules present in the medium (bimolecular reactions). The lifetimes of excited molecules are limited by the probability of chemical reaction, by the probability of return to the ground state with the emission of light, and by the probability of return to the ground state without light emission but with the generation of heat energy. *See* ELECTRON CONFIGURATION; EXCITED STATE.

Unimolecular reactions. The initial steps of the simplest excited-state chemical reactions involve only the excited molecule or solvent molecules, and may be classified as excitation [reaction (1), where M is a metal and L and X are ligands]; decrease in coordination number [reaction (2)]; increase in coordination number for a solvent species [S; reaction (3)]; and homolysis, as in oxidation-reduction, in which bonding electrons shared equally in the products [reaction (4), with ·X as a free radical].

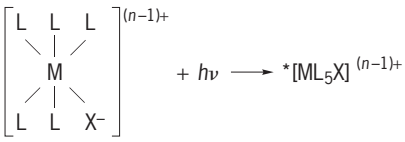

(1)

$$^{*}[ML_5X]^{(n+1)+} \longrightarrow ML_5^{n+} + X^{-}(\text{or L}) \qquad (2)$$

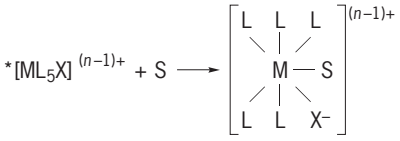

(3)

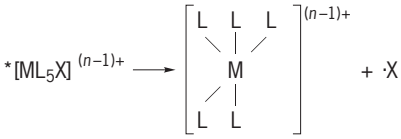

(4)

Each of the chemical processes (2), (3), or (4) gives rise to an unstable, therefore reactive, chemical intermediate. Such intermediates are often useful in the synthesis of new compounds, as in reaction (5).

$$ML_5^{n+} + Y^{-} \longrightarrow [ML_5(Y^{-})]^{(n-1)+} \qquad (5)$$

When these intermediates are generated on a surface or in an inert matrix, their lifetimes may be enhanced, and they may function as reactive catalytic sites. *See* CATALYSIS; FREE RADICAL; REACTIVE INTERMEDIATES.

The changes in the number of ligands coordinated to the central metal [reaction (2) or (3)] can be systematically treated by using theoretical models of the relationship between the electronic configuration of the molecular species and the nature of the coordinate covalent bond. The net chemical changes observed depend on the chemistry of the intermediate species $[(ML_5)^{n+}$ or $(ML_5SX)^{(n-1)+}]$ as well as the chemistry of the reactive excited states $[(^{*}ML_5X)^{(n-1)+}]$.

While excited-state reactions that result in changes in coordination number produce a single reactive species, the intermediate metallofragment, homolytic cleavage of a metal-ligand bond results in formation of a pair of very reactive species: a reduced metallofragment and a one-electron oxidized ligand species. These very reactive substances tend to recombine to regenerate the original metal compound, in addition to reacting with themselves or other solution species. The recombination reaction may be very rapid (in less than a nanosecond) when these species are formed in proximity, or it may be somewhat slower (on the order of 10^{-6} s) if the species manage to separate a few molecular diameters. In order for a simple homolytic process to be observed, the minimum energy difference between the reactive excited state and the ground state must exceed the M^{n+}-(X^{-}) bond energy. However, there are some systems in which the solvent can assist the homolytic process by means of a concerted displacement of the departing radical fragment. The net quantum yields for photohomolytic processes are reasonably large (that is, greater than 0.1) only when the recombination reactions are slow compared to other reactions of the reactive fragments. Homolytic reactions tend to dominate the chemistry of the excited states of cobalt(III) complexes.

Homolytic reactions turn out to be very efficient for methylcobalamin and other organocobalt complexes related to vitamin B_{12}. These materials are very highly colored (deep red), and most of the light photons absorbed result in homolytic cleavage of the cobalt-carbon bonds. Various recombination reactions decrease the observed product yields to less than 30%. This photosensitivity results in degradation of these natural complexes, and photoinduced reactions are not important to their enzymatic function. However, the small photonic energy required for cobalt-carbon homolysis is a manifestation of the weakness of this chemical bond. Thermal cleavage of the cobalt-carbon bond has been postulated as a key step in the enzymatic functioning of coenzyme-B_{12}.

Inevitably, many of the photoinduced unimolecular processes found for metal compounds cannot be neatly placed in the above categories. Among the most intriguing are processes that involve bond-breaking or bond-making processes on a ligand. For example, irradiations of some rhodium(III) or iridium(III) azide complexes (M^{III}-N_3^{-}) result in cleavage of the N^{-}-N_2 bond forming a coordinated nitrene, M^{III}-N^{-}. Another remarkable example is cleavage of the O_2C-(CO_2^{2-}) bond in some oxalate complexes (**1**) to form a metal-carbon bonded formate complex (**2**).

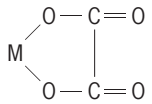

(**1**)

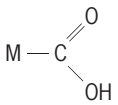

(**2**)

There are a number of photoinduced isomerizations [reaction (6)] that might appear to fall outside the cat-

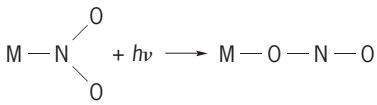

(6)

egories listed above but in which the primary photoprocess appears to involve metal-ligand homolysis, followed by very rapid recombination of the fragments before they can be separated by one or more solvent molecules. *See* NITROGEN COMPLEXES.

Bimolecular excited-state reactions. Reactions between excited states and molecules other than solvent species are very important when the electronically excited molecules are long-lived. Metal compounds are frequently facile oxidation-reduction reagents, and metal-compound excited states can be employed to displace oxidation-reduction equilibria. The important categories of bimolecular excited-state reactions are electron transfer [reaction (7*a*) or (7*b*)] and quenching by electronic energy transfer [reaction (8)].

$$^{*}[ML_6^{n+}] + A \longrightarrow ML_6^{(n-1)+} + A^{+} \qquad (7a)$$

$$^{*}[ML_6^{n+}] + B \longrightarrow ML_6^{(n+1)+} + B^{-} \qquad (7b)$$

$$[ML_6^{n+}]^{*} + C \longrightarrow ML_6^{n+} + C^{*} \qquad (8)$$

The transfer of an electron from one species to another constitutes a net redistribution of electrical charge. As a consequence, there is necessarily some difference in the solvation of reactants and products; in addition, there will often be differences in the bond lengths of reactants and products. The larger these differences in bond length and solvation, the slower will be the rate of the electron-transfer process. Excited-state lifetimes are necessarily short, so that electron-transfer reactions must occur rapidly if they are to be of any consequence. Very rapid excited-state electron-transfer reactions occur most often between large molecules (thus minimizing solvation energies); most often between molecules in which the electron transfer does not result in large bond-length changes; and in part because the excitation energy stored in the excited molecules helps make the reactants appreciably less stable than the products. Typical of the metal compounds employed in the study of such reactions are *tris*-bipyridyl complexes, $M(LL)_3^{n+}$, where LL equals structure (**3**) and

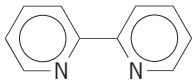

(**3**)

M^{n+} is Ru^{2+} or Cr^{3+}. Enough energy is released in the electron-transfer process to overcome some of the intrinsic factors limiting the rate. Indeed, typical metal-compound excited states can store enough energy that even the recombination reactions $ML_6^{(n-1)+}, + A^{+} \rightarrow ML_6^{n+} + A$ or $ML_6^{(n+1)+} + B^{-} \rightarrow ML_6^{n+} + B$, could ideally produce 1–2 V of electrical energy in a battery. Such excited-state electron transfer reactions afford a convenient means of generating very reactive intermediate species (A^{+}, B^{-}). *See* BOND ANGLE AND DISTANCE; ELECTRON-TRANSFER REACTION.

Quenching by electronic energy transfer is a theoretically complex process since it requires simultaneous relaxation of a donor electron and excitation of an electron in the acceptor molecule. Although such reactions tend to be degradative in most systems, they can be employed to generate chemically interesting acceptor-molecule excited states that cannot be easily populated by direct light absorption.

Photoinduced electron transfer and supramolecular clusters. Attempts to utilize metal compounds to mediate the transformation of light energy from the Sun into a useful chemical fuel have often focused on the cleavage of water, shown in reaction (9).

$$H_2O + ML_6^{n+} + h\nu \longrightarrow H_2 + {}^{1}/_{2}O_2 + ML_6^{n+} \qquad (9)$$

Several metal-compound excited states store enough energy to promote this reaction as it is written. However, detailed consideration of reaction (9) reveals that it must involve at least two water molecules (H_2O) and four electrons per product molecule formed; and the energy requirements for this reaction cannot be simply equated to the energy available from single-electron-transfer processes of molecular excited states [as in reactions (7*a*) and (7*b*)]. However, reactive intermediates generated in electron-transfer processes might be capable of transforming more than one equivalent of electrons (for example, by reduction of H^{+} to a coordinated hydride H^{-}). Appropriate intermediates might be homogeneous (for example, low-valent complexes in solution) or heterogeneous (for example, colloidal metals or metal oxides).

Covalently linked aggregates of transition-metal complexes can be synthesized in which a specific photoreaction is favored. For example, one complex many function as a light absorber and sensitizer, and a second complex may function as the site of an excited-state reaction following intramolecular electron transfer. In principle, the acceptor center could be a semiconductor electrode or particle [for example, of titanium oxide (TiO_2)] as well as a coordination complex. Some such clusters of light-absorbing complexes and acceptor centers can generate reasonably good photocurrents in solar cells, while others may be useful as efficient photocatalysts. In some such clusters, it is possible to accumulate sufficient charge separation to drive reactions as complex as (9). *See* CHEMICAL DYNAMICS; COORDINATION CHEMISTRY; COORDINATION COMPLEXES; LASER PHOTOCHEMISTRY; PHOTOCHEMISTRY; SUPRAMOLECULAR CHEMISTRY.

John F. Endicott

Bibliography. A. W. Adamson and P. D Fleischauer (eds.), *Concepts of Inorganic Photochemistry*, 1975, reprint 1985; V. Balzani and F. Scandola, *Supramolecular Photochemistry*, 1991; G. J. Ferraudi, *Elements of Inorganic Photochemistry*, 1988; M. A. Fox and M. Chanon

(eds.), *Photoinduced Electron Transfer*, 1988; V. Ramamurthy and K. S. Schanze (eds.), *Organic and Inorganic Photochemistry: Molecular and Supramolecular Photochemistry*, vol. 2, 1998; C. E. Wayne and R. P. Wayne, *Photochemistry*, 1996.

Inositol

The common name for 1,2,3,4,5,6-hexahydroxycyclohexane, which is classified as a carbohydrate. Inositol has nine different naturally occurring structural isomers. However, the most common form used biologically is D-*myo*-inositol (**Fig. 1***a*), commonly referred to as *myo*-inositol. Biologically, *myo*-inositol is used in the formation of the phosphatidylinositols in organisms, including plants, animals, and microorganisms. Phosphatidylinositols are used in a variety of cellular functions such as cell signaling, anchoring of proteins to the membrane, membrane trafficking, cellular development, and interaction with cytoskeletal elements. *See* CARBOHYDRATE.

Synthesis. The cellular biosynthesis of *myo*-inositol occurs by the cyclization of D-glucose-6-phosphate, an intermediate of carbohydrate metabolism, followed by dephosphorylation by a specific phosphatase. The cyclization reaction requires nicotinamide adenosine dinucleotide (NAD) as a cofactor and is catalyzed by the enzyme glucose-6-P cyclase. Mammals have a very limited ability to synthesize inositol, so it is considered an essential nutrient in their diets. *myo*-Inositol is commercially produced from phytic acid (*myo*-inositol hexaphosphate), which is found in corn-steep liquor following wet milling of corn. *See* CARBOHYDRATE METABOLISM.

Role in phosphatidylinositol synthesis. The most important biological function of *myo*-inositol is its incorporation into phosphatidylinositol (PI; Fig. 1*b*), a key phospholipid regulating cellular biochemical pathways. The biosynthesis of PI is a two-step process. In the first step, phosphatidic acid is converted to cytidine diphosphate diacylglycerol (CDP-DAG) by the condensation of cytidine triphosphate with phosphatidic acid. This reaction is catalyzed by CDP-DAG synthase and is freely reversible. The second step involves the action of the microsomal enzyme PI synthase to catalyze the condensation of *myo*-inositol with the CDP-DAG to form PI. This reaction requires Mg^{2+} as a cofactor and is selective for 1-steroyl,2-arachidonyl-glycerol-3-phosphate, a specific phosphatidic acid, thereby accounting for accumulation of this molecular species of PI within cells. *See* PHOSPHOLIPID.

Cell signaling. The *myo*-inositol moiety of PI has multiple sites which are sequentially phosphorylated by a family of specific kinases. The most important phosphorylated form of PI is phosphatidylinositol-4,5-bisphosphate (PIP_2). This form is localized inside the cell in two different sites: the inner leaflet of the plasma membrane and the nuclear membrane (**Fig. 2**).

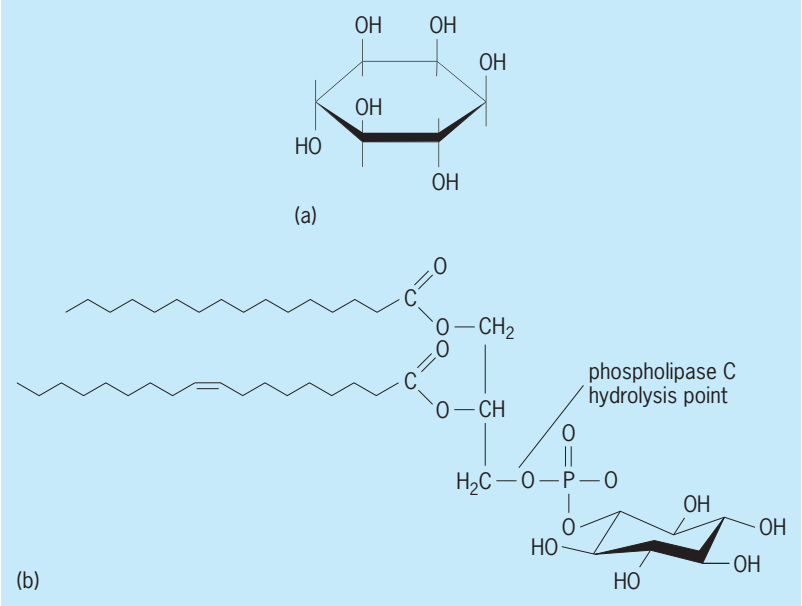

Fig. 1. Chemical structure of (*a*) D-*myo*-inositol and (*b*) phosphatidylinositol.

Cell signaling is turned on by the binding of various ligands to one of three different membrane receptor families. Upon ligand binding, phospholipase C (PLC) is activated, thereby initiating a potent intracellular signaling cascade. PLCs hydrolyze the PIP_2, releasing 1,4,5-inositol trisphosphate (IP_3) and diacylglycerol (DAG). These molecules potentiate the signal by their own downstream signaling events. IP_3 binds to specific intracellular receptors, releasing calcium from intracellular stores. DAG and the released Ca^{2+} then activate the protein kinase C (PKC) family, which then phosphorylate multiple different proteins, eliciting a variety of cellular responses. DAG lipase hydrolyzes the first fatty acid off the released DAG, making 2-arachidonyl glycerol (2-AG), a potent cannabinoid ligand that can act outside the cell (Fig. 2). Hence, the activation of a given cell surface receptor has multiple downstream effects, resulting in an amplification of the initial signal.

For every event that turns a signaling cascade on, there are a number of key steps that turn a signaling cascade off. The process described above is tightly regulated because of the tremendous amplification of the signal. Either IP_3 is phosphorylated to IP_4 by PI3 kinase or a phosphate is hydrolyzed by specific phosphatases, forming IP_2. Both of these inositol phosphates are inactive at the intracellular IP_3 receptors, so calcium release is stopped. Calcium is removed from the cell by intracellular calcium pumps and is returned primarily to the endoplasmic reticulum (ER). Eventually, through the sequential action of a family of phosphatases, the inositol phosphates have all of the phosphates removed, reforming *myo*-inositol. DAG is rapidly phosphorylated by DAG kinase to form phosphatidic acid. This phosphatidic acid and the *myo*-inositol are reused to form PI. This sequential set of events is commonly referred to as the PI cycle. *See* CALCIUM METABOLISM; SIGNAL TRANSDUCTION.

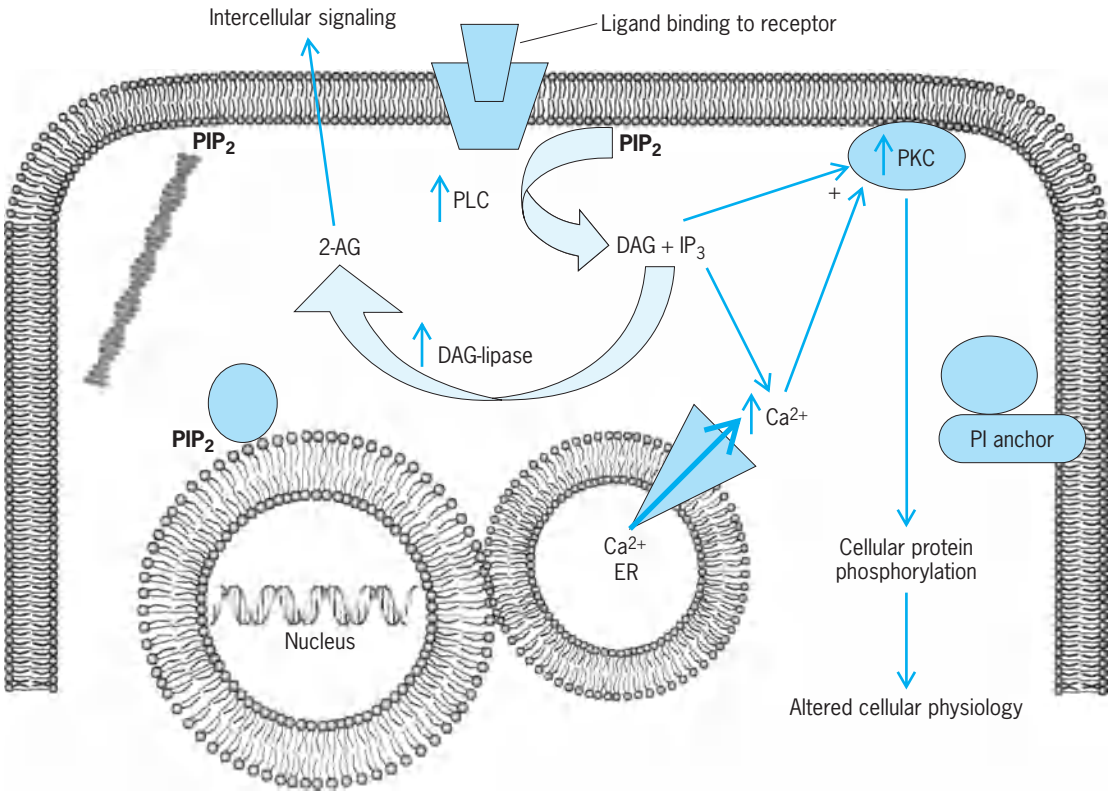

Fig. 2. Process of cell signaling.

Other functions. Another important function of phosphatidylinositols is the anchoring of membrane proteins through three distinct mechanisms.

(1) PI is glycosylated with a number of sugar moieties, including glucosamine, to form a polysaccharide chain on the PI. This glycosyl-PI is found in the plasma membrane inner leaflet, and free amine reacts with a number of proteins to anchor the proteins to the membrane.

(2) PIP_2 interacts directly with cellular proteins containing a certain sequence of amino acids called the pleckstrin homology domain. Recent discoveries demonstrate that this form of interaction is important in lipid-mediated signaling in the nuclear membrane, which aids in regulating gene expression.

(3) PI3 kinase phosphorylates the 3 position of the *myo*-inositol ring of PI, resulting in a family of PI's that have a unique function in cells. These PI3P's interact with a number of different proteins within the cell to promote membrane trafficking, cytoskeleton maintenance, PKC regulation, and cell survival. This includes regulation of important cellular processes such as endocytosis and exocytosis. *See* CELL MEMBRANES.

Eric J. Murphy

Bibliography. M. A. J. Ferguson and A. F. Williams, *Annual Review of Biochemistry*, 1988; P. W. Majerus, *Annual Review of Biochemistry*, 1992; C. K. Matthews and K. E. van Holde, *Biochemistry*, 2d ed., 1996; W. Pigman and D. Horton (eds.), *The Carbohydrates IA*, 1972; G. J. Siegel et al. (eds.), *Basic Neurochemistry: Molecular, Cellular, and Medical Aspects*, 6th ed., 1999; D. E. Vance and J. E. Vance (eds.), *Biochemistry of Lipids, Lipoproteins, and Membranes*, 4th ed., 2002.

Insect control, biological

The use of parasitoids, predators, and pathogens to reduce injurious pest insect populations and consequently the damage they cause. Viruses and bacteria are the most commonly used pathogens, but fungi, protozoa, and nematodes may also be important biological control agents. *See* PARASITOLOGY.

Ecological foundations. Three ecological assumptions underlie biological control. First, natural enemies are among the prime factors responsible for the regulation, or control, of pest populations. Second, the influence exerted by parasitoids, predators, and pathogens is density-dependent. Density dependence means that the killing power of the natural enemy increases as the prey or host density increases. Conversely, the mortality induced by density-dependent natural enemies decreases as host density increases. Both of these assumptions have become the subject of debate, and scientists have proposed alternative ecological assumptions that do not require density-dependent responses or stable interactions between natural enemy and prey or host. In the dominant, or classical, form of biological control the third assumption is found: when an insect species escapes into a new area without its natural enemies, it reaches outbreak levels and becomes a pest. Biological-control practitioners

believe regulation can be reestablished by importing the natural enemies of the pest from its area of origin.

Classical approach. In classical biological control, all efforts are typically directed toward establishing the natural enemies that were left behind in the area of origin. Such an approach demands the verification of the area of origin, foreign exploration for natural enemies in the area of origin, importation of those natural enemies, quarantine, mass rearing, and release of the biological control agents. Quarantine procedures ensure that no natural enemy of the biological control agent selected—including a hyperparasitoid, which is a parasitoid of a parasitoid—is accidentally released. Mass rearing attempts to produce for release the largest number of vigorous and competitive biological control agents possible.

Classical biological control is by far the most frequently used form of biological control, assuming one excludes the use of resistant plant varieties as biological control. The Commonwealth Agricultural Bureau's International Institute of Biological Control has recorded 860 successful establishments of 393 species of parasitoids against 274 pest insects in 99 countries. In 216 cases, parasitoids by themselves or aided by predators achieved complete or satisfactory pest suppression, and an additional 52 cases are considered to have provided a useful reduction.

Conservation. Conservation involves manipulation of the environment in order to favor survival, reproduction, or any other aspect of the natural enemy's biology that affects its function as a biological control agent. Conservation can be undertaken regardless of whether the biological control agent is introduced or native, and generally involves tactics such as maintaining floral food sources, spraying sugar or yeast-sugar solutions for parasitoids and invertebrate predators, or providing nest boxes for vertebrate or invertebrate predators. Modification of agronomic practices or the timing of activities such as tilling or harvest residue destruction is another approach that may enhance the survival of natural enemies. For example, spraying insecticides in strips rather than across entire fields may provide an important degree of protection for natural enemies. Some natural enemies may become established but fail to spread beyond the area of release. That problem often may be solved by redistributing the biological control agent—collecting samples of the agent and releasing them at a distant, appropriate site. *See* INSECTICIDE.

Trends. Aspects of research on and application of biological control may provide new or improved approaches. Successes in the development of pesticide-resistant predatory mites suggest that the improvement of biological control agents through selection, hybridization, or genetic engineering techniques may play an important role. A major strategy for control of pest insects may involve the use of genetic engineering to introduce traits into natural enemies that enhance their performance, or mortality-causing traits of natural enemies, such as insect pathogens, into plants. Corn, tobacco, tomato, potato, and cotton have already been genetically engineered to incorporate the gene that regulates the production of the toxin normally produced by the bacterium *Bacillus thuringiensis*. Insect parasitoids and pathogens such as viruses are also the subject of genetic engineering research, which has attempted to clone and characterize the DNA encoding genes responsible for their toxicity or for characters that determine survival in the environment. *See* BREEDING (PLANT); GENETIC ENGINEERING.

Pedro Barbosa

Bibliography. M. A. Hoy and D. C. Herzog, *Biological Control in Agricultural IPM Systems*, 1985; G. C. Papavizas, *Biological Control in Crop Protection*, 1981; J. Waage and D. Greathead, *Insect Parasitoids*, 1986.

Insect diseases

A biological discipline embracing the general principles of pathology (disease in the broadest sense) as applied to insects. It refers to human observations and actions concerning the cause, symptomatology, gross pathology, histopathology, pathogenesis, and epizootiology of the diseases of insects; it is concerned with whatever can go wrong or become abnormal in an insect. A diseased insect may be suffering from an infectious disease caused by a microorganism or a noninfectious disease, such as metabolic disturbance, a genetic abnormality, a fetal malformation, a nutritional deficiency, a physical or chemical injury, or an injury caused by parasites or predators.

Insect pathology draws upon and contributes to the general field of microbiology and provides understanding of certain of the biological relationships existing between insects and microorganisms not pathogenic to them. Insect pathology finds applications in agriculture, medicine, and biology generally. Microbial control, the use of microorganisms in biological control, is one area of applied insect pathology. Microorganisms are introduced to control insect pests for the protection of humans and other animals and agricultural crops. However, the suppression of disease in beneficial insects, such as the silkworm, honeybee, and ladybird beetle, is also of significant practical importance. *See* INSECT CONTROL, BIOLOGICAL.

Pathogenesis and epizootiology. Many factors contribute to the condition or state described as a disease. It is important when diagnosing a disease to understand its pathogenesis, that is, how the pathology develops and how the individual host responds. Identifying the infectious microorganism or physical, chemical, or genetic factors affecting the insect is only a part of the knowledge needed for determining pathogenesis. Distinguishing one disease from another is as basically important in insect pathology as it is in the study of disease in other forms of life. One must know what is abnormal about or has killed an insect before the disease can be properly studied,

controlled, or suppressed, used as a microbial control measure, its potentialities for natural spread determined, or its role in the ecology of an insect species known. *See* POPULATION ECOLOGY.

An important facet of insect pathology is concerned with the epizootiology of diseases, which involves the study of disease in populations of insects rather than in individual animals. Any epizootic affecting an insect population is concerned with three primary natural entities: the infectious agent, the insect host, and the environment. Each of these factors has certain attributes that, when properly related to the attributes of the others, play their appropriate role in determining the initiation, rise, and decline of an epizootic. Knowledge concerning the nature of epizootics is extremely important in insect ecology generally and in an understanding of insect-microbe ecosystems. The degree to which disease-producing microorganisms are dependent upon the density of the host population, the susceptibility of the host insect to the disease, the influence of weather and other environmental conditions on epizootics, the mode of transmission of insect pathogens, the success of the microorganism remaining alive, and other factors are important aspects of the epizootiology of insect disease. In the microbial control of an injurious insect, humans can initiate and increase the rapid rate of development of an epizootic by controlling the amount, method of distribution, and time of introduction of an effective microorganism.

Infectious diseases. While the noninfectious diseases and abnormalities of insects are very important and of numerous types, most of the activity in insect pathology has been with the infectious diseases. The infectious agents responsible for diseases in insects belong to the same major groups as those that cause such diseases in other organisms: the bacteria, fungi, viruses, rickettsiae, protozoa, and nematodes. In general, however, insects are not normally susceptible to those particular microorganisms that cause diseases of man, other animals, and plants. Moreover, most of the microorganisms that cause fatal diseases in insects are harmless to plants and to higher animals. Insects often can resist infection by humoral and cellular types of immunity; acquired humoral immunity is manifested in the hemolymph. However, it appears that the antibody immunity characteristic of higher animals does not occur in insects.

A convenient way to consider the various infections of insects is on the basis of the etiologic agent, as will be done in the sections to follow.

Bacterial Diseases

The characteristics of a bacterial disease in an insect may vary, but in general the following may be observed. As the disease develops, the insect becomes less active, has a smaller appetite, and discharges fluids from the mouth and anus. Infection may begin as a dysenteric condition with an accompanying diarrhea, but in most instances the invading bacterium eventually enters the body cavity of the insect and causes a septicemia that terminates in the death of the host. Following death, the insect's body usually darkens to brown or black as melanin pigment forms in the tissues. This is especially true of larvae and pupae in which the disintegration takes place rapidly.

The freshly dead insect is usually soft and flaccid. Internal tissues may disintegrate to a viscid consistency, sometimes accompanied by odor, but ordinarily they do not "melt," or liquefy, as do insects dying of certain virus infections. The cadaver usually dries and becomes shriveled, the integument remaining intact. Microscopic examinations of smears or of histological sections of an insect dead or dying of a bacterial disease usually disclose large numbers of the causative bacterium. These bacteria, if they are pathogens, must be differentiated from similar-appearing saprophytes that are normally in the insect intestine and that may flourish in the tissues of the dead insect. In some cases the bacterial pathogen may not be present in large numbers if it has killed the insect primarily through the production of toxins.

Spore-forming bacteria, which remain dormant for long periods of time if necessary and which are resistant to destructive environmental factors, have increased the probability, when compared to other bacteria, of contributing to epizootics with or without human manipulation.

Foulbrood. The first bacterial diseases of insects to receive concentrated study were flacherie of the silkworm, which Louis Pasteur considered at the time he was studying the protozoan disease pébrine of this same insect, and European foulbrood of the honeybee, which was studied by F. R. Cheshire and W. W. Cheyne in 1885. Later it was found that, in addition to European foulbrood, caused by a group of nonspore-forming bacteria, American foulbrood caused by the spore-forming bacterium *Bacillus larvae* also afflicted this beneficial insect. After the role of these bacteria in causing diseases of larval bees (hence the term foulbrood) was elucidated, it was possible to institute measures of control; however, in some areas these and other bee diseases still cause losses to beekeepers.

Milky diseases. In addition to the diseases of bees, two of the most important bacterial diseases of insects are the so-called milky diseases, which affect grubs (larvae) of the Japanese beetle (*Popillia japonica*) and are caused by *Bacillus popilliae* and *B. lentimorbus*; and the disease of numerous lepidopterous larvae caused by *B. thuringiensis*. All of these bacteria are sporeformers and are effective in the control of certain insect pests. Infection of Japanese beetle grubs occurs through ingestion of the bacterial spore, which germinates, penetrates the insect's intestinal wall, and develops in the body cavity, where the bacteria sporulate. The blood of the infected grub assumes the milky-white appearance that gives the disease its common name. Eventually the insect dies and decomposes, thus providing new spores which may be consumed by other grubs. In the eastern United States a dust containing milky-disease spores is used, at times in conjunction with chemicals and insect parasites, to control the Japanese beetle in

lawns and turf. Distribution of this material has been carried on by the U.S. Department of Agriculture, which did much of the original work on the disease. Products containing the spores are commercially available.

Crystalliferous bacteria. *Bacillus thuringiensis* is one of the most intensively studied bacterial pathogens of insects because it has proved to be an effective microbial control agent against a number of insect pests, such as the alfalfa caterpillar, larvae of the cabbage butterfly, and larval forms of certain mosquitoes, black flies, and beetles. A number of closely related species of this bacillus are known to cause diseases of beneficial insects, such as silkworms. Several different strains or varieties of *B. thuringiensis* are known. Certain types affect only Lepidoptera, others only Diptera or Coleoptera. These bacteria are characterized by the formation of a protein crystal in the sporangium at the time of spore formation. The protein material comprising the crystal is highly toxic for certain insects, but apparently harmless for other forms of life. The crystal usually appears bipyramidal but may occur in other forms. Ordinarily each sporangium contains but one crystal; the crystal is freed from the cell wall with the spore and appears to persist indefinitely.

Bacillus thuringiensis and related species also produce a heat-stable, water-soluble substance, distinct from the crystal and from lecithinase, that is toxic for certain insects when injected, and to larval Diptera when they are exposed to it in a rearing medium.

Preparations containing the spores and crystals of *B. thuringiensis* or its varieties may be sprayed or dusted on crops or into aquatic habitats to destroy lepidopterous, dipterous, or coleopterous pests. Commercial products containing the bacillus are available. One of the primary advantages of this type of control is that this and other entomophilic microorganisms are nontoxic and noninfectious for plants and higher animals.

Bacillus sphaericus complex. Approximately 70 isolates of *Bacillus sphaericus* are described, from soil, mud, and vertebrate and invertebrate animals. Eight of these isolates release substances from the walls of spores in the alimentary tract of larval forms of mosquitoes. They are toxic for the insects and apparently harmless to all other forms of life. The bacteria can reproduce in the dead insect, as well as use other sources in habitats of mosquito larvae to provide additional spores.

Fungal Diseases

There are a large number of species of different groups of fungi that are pathogenic for insects. Their development on and in susceptible hosts depends upon the presence of adequate moisture to provide for the germination of conidia and spores, or on the presence of another invertebrate as an alternative host for development of the form of the fungus infectious for the insect. Whereas bacteria, viruses, and protozoa infect their insect hosts by way of the mouth and digestive tract, most entomophilic fungi enter their hosts directly through the integument or body wall.

Many of the fungi associated with insects are not true pathogens, or are pathogenic only under certain conditions. These saprophytic species frequently develop in or on insects that have died of other causes, since the conidia of the fungi may be closely associated with the living healthy insect. Some fungi appear to be parasitic but not lethal in their association with insects; thus the large order Laboulbeniales (Ascomycetes) is made up of species that live primarily on the external surfaces of insects. The genus *Septobasidium* consists of species associated with scale insects which, although some are parasitized, live beneath the stromata of the fungi. Some fungi, such as the ambrosia fungi of certain termites and wood-boring beetles, are definitely mutualistic in their relationship with insects. *See* ASCOMYCOTA; BASIDIOMYCOTA.

Fungal development. Once the germ tube from a conidium is within the body cavity of an insect, the threadlike hypha of the fungus proliferates, invades the tissues, and fills the body of the insect. In most cases the fungus then develops conidiophores to the exterior of the host body, where conidia develop to provide a means for initiating development of the fungus in new hosts. The infected insect usually assumes a dried mummylike appearance, often is covered with a conidia-bearing mycelium, or can contain resting spores which enable the fungus to survive periods of adverse environmental conditions or absence of a suitable host. Toxins may be produced by the fungus during the initial period of hyphal growth.

Entomophilic fungi. Some entomophilic fungi occur in the orders Mucorales, Blastocladiales, and Chytridiales, but the most important fungi of the class Zygomycetes belong to the order Entomophthorales. In this order the single family Entomophthoraceae is made up of several genera, of which *Entomophthora* and *Massospora* are composed primarily of entomophilic species. One of the most commonly known species is *E. muscae*, which attacks the housefly. A large number of *Entomophthora* species are known. They infect many species of Diptera, Orthoptera, Lepidoptera, and other insects.

Infected houseflies are frequently found indoors attached to the walls and ceilings of buildings in a lifelike position. Close inspection of flies killed by the fungus usually reveals on the wall or windowpane a distinct halo of discharged spores (conidia) encircling the insects. This halo results from the fact that, after the conidiophores have emerged through the integument of the insect, conidia are formed which are discharged violently into the air from the terminal portion of the conidiophores, thus forming a ring of conidia about the dead insect. When the discharged conidum comes into contact with a susceptible insect in the presence of adequate moisture, it begins to germinate, sending out a conidial hypha (germ tube) that penetrates through the integument into the body cavity.

Blastocladiales. The entomophilic Blastocladiales, class Chytridiomycetes, are confined largely to one group, the family Coelomomycetaceae, which parasitizes mainly mosquito larvae. About 50 species are known, all in the genus *Coelomomyces*. Investigations of the biology and life cycle of *Coelomomyces* species have demonstrated the requirement for the presence of invertebrate animals, recognized as copepods, as alternative hosts for the fungus pathogen. Infectious forms from the larvae of mosquitoes infect the copepod, and the infectious agents from the copepod initiate the fatal infection in the mosquitoes. Certain regions of the body cavity or virtually the entire hemocoel of the mosquito may become filled with the spores and mycelium of the fungus.

Hypocreales. Species of the genus *Cordyceps*, an ascomycete of the order Hypocreales, are among the first recorded entomophilic fungi. They are frequently of large size and very colorful in appearance. About 250 species are known. They are cosmopolitan in distribution and occur on representatives of several orders of insects, principally Hemiptera, Diptera, Lepidoptera, Hymenoptera, and Coleoptera. In members of the genus *Cordyceps*, the stroma arises from a sclerotium formed within the body of the infected insect. At the end of this stroma, or stem, is a fertile portion, the "head," which may be brightly colored. The life history of most *Cordyceps* is similar to that of entomophilic fungi in general.

Muscardine. One of the most important groups of entomophilic fungi is responsible for muscardine diseases, in which the fungus emerges from the body of the insect and covers the animal with a characteristic fungus mat resembling, in a way, a French bonbon or candy mint. The word muscardine was first used for white muscardine, a well-known disease of the silkworm caused by the fungus *Beauveria bassiana*. It has also been used in reference to green muscardine, a disease of the European corn borer, the wheat cockchafer, and other insects caused by *Metarhizium anisopliae*. The mycelium from conidia can enter the respiratory siphon of mosquitoes at the water surface and grow, resulting in the death of larvae. These fungi belong to the Fungi Imperfecti. *Beauveria bassiana* and closely related species have been identified from a large number of other insects. In certain situations the fungus has shown some promise as a means of control of pest insects. The fungi produce toxin in the host, accelerating the death of the insect. Among their best known hosts, in addition to the silkworm, are the European corn borer, codling moth, chinch bug, and the alfalfa weevil. *See* DEUTEROMYCOTA.

Viral Diseases

The insect species infected by viruses are principally in the order Lepidoptera, although a few Hymenoptera and Diptera and one or two Coleoptera and Neuroptera are known to suffer infection with these agents. Only the immature stages, larva and pupa, are killed by virus. Adults may carry the virus, but they are not killed by it. In general, many insect viruses exhibit a fairly high degree of host specificity, but numerous instances of cross infectivity have been recorded. *See* VIRAL INCLUSION BODIES; VIRUS.

Polyhedroses. The polyhedroses are characterized by the formation of polyhedron-shaped inclusion bodies in the infected tissues of the host. These proteinaceous inclusion bodies contain the virus particles embedded in their matrices. In the nuclear polyhedroses the virus multiplies in the nuclei of the infected cells, and in the cytoplasmic polyhedroses the virus multiplies in the cytoplasm of such cells. As far as is known, viruses causing nuclear polyhedroses are rod-shaped and contain deoxyribonucleic acid (DNA); those causing cytoplasmic polyhedroses are more or less spherical and contain ribonucleic acid (RNA).

Larvae infected with the virus of a nuclear polyhedrosis usually show few distinctive symptoms until a few hours before death. The incubation period varies in different host species between 5 and 20 days; usually it covers a period of about a week. In some species of insects, the infected larvae may cease feeding, become somewhat sluggish in movement, and become yellowish or pale in color. They may swell slightly, then become limp and flaccid. The tissues most prominently affected are the fat-body cells, epidermal cells, blood cells, and the cells of the tracheal matrix. Shortly before and after death the integument is very fragile and easily ruptured, emitting the liquefied contents, which are filled with disintegrating tissue and polyhedra. Polyhedra increase in size and number in the nucleus of the cell until the nuclear and cell membranes are disrupted. The dead larvae are usually found hanging by their prolegs from the host plant or other support. Eventually, they may burst or dry to a dark-brown or black cadaver.

The polyhedra are insoluble in water, alcohol, ether, and acetone but soluble in weak acids and alkalies. They range in size from 0.5 to 15 micrometers. If the polyhedra are treated properly with dilute alkali, the rod-shaped particles within them can be demonstrated. The size of the virus rods ranges from about 20 to 50 nanometers in width and from about 200 to 350 nm in length. The virus rods may occur singly or in bundles of up to as many as eight rods.

Production in living larval Lepidoptera of nuclear polyhedroses in large quantities to control agriculture and forest insects is an accepted practice throughout the world. In the United States, the Environmental Protection Agency has approved virus products for use against the cotton bollworm, the gypsy moth, and the tussock moth.

The cytoplasmic polyhedroses are, in general, not so fulminating as the nuclear polyhedroses. The host integument does not become fragile, as with the nuclear polyhedroses. In the cases known, the infection is limited largely to the midgut epithelium, where the polyhedra are formed in the cytoplasm of the cells. The polyhedra are morphologically similar to those present in nuclear polyhedroses. The virus particles, in contrast, tend to be spherical in shape, about

20–70 nm in diameter, and near the surface of the polyhedron. One of the best-known cytoplasmic polyhedroses is that of the silkworm, which also suffers from a nuclear polyhedrosis. *See* DEOXYRIBONUCLEIC ACID (DNA); RIBONUCLEIC ACID (RNA).

Granuloses. The granuloses have so far been found only in larvae of Lepidoptera. They are characterized by the accumulation of small granular inclusion bodies, called capsules, in the infected cells. Each capsule contains a single virus rod composed of DNA. Compared to the nuclear polyhedroses, the symptoms associated with the granuloses are more subdued; the diseased larvae usually become less active, somewhat flaccid, and assume a pallid or whitish translucent aspect. The integument does not become fragile. The period from infection to death ranges from 6 to 20 days. The inclusion bodies, or capsules, are about 200 by 500 nm in size, and the virus rod ranges, depending on the virus, from 40 to 80 nm in width and from 200 to slightly over 300 nm in length.

Other viruses. Among the insect viruses so far discovered that are not associated with an inclusion of any kind are those which infect certain Lepidoptera (army worms), Hymenoptera (honeybees), Coleoptera (beetle grubs), and Diptera (mosquitoes, crane flies). A number of the noninclusion viruses exhibit a unique physical characteristic when concentrated in large quantity. Centrifugate pellets of the virus particles appear orange or amber in color by transmitted light and iridescent or turquoise by reflected light. Another peculiarity is the fact that an exceptionally large amount of virus is produced in the host tissues. A noninclusion virus called "sigma virus" is known to occur in *Drosophila*. This virus renders infected flies sensitive to carbon dioxide. Among acarines, noninclusion viruses have been found in at least two species of mites.

There is considerable evidence that insect viruses frequently remain occult and cause latent infections in their hosts. Some authorities believe that a virus can survive in a host, even for several generations, without causing recognizable symptoms, but that under the influence of certain stresses or incitants, such as cold, crowding, and certain chemicals, can be triggered into causing an active or frank infection.

Rickettsial Diseases

A number of rickettsiae or rickettsialike organisms are capable of causing frank infections in insects. Three or four have been found in Coleoptera and one in Diptera. These infections appear to be slow to develop and kill their hosts in 1–4 months. The fat body of grubs assumes a characteristic bluish-green coloration, and in some host species crystalline bodies occur in association with the rickettsie.

Protozoan Diseases

Protozoan infections in insects may be benign and cause little morbidity or mortality, but many are severe and fatal. Some appear in epizootic proportions, while others may be represented by only a local invasion of tissue by the protozoon. The success of reproducing populations of insects can be seriously affected by destruction of gonads without death of the host. Generally, protozoan diseases are relatively slow in developing, and frequently become somewhat chronic in nature; but some sporozoan infections develop rapidly and kill the insect host within a short time. Infected insects may show few, if any, external signs of infection, or may be stunted in growth and development or not undergo metamorphosis, or show a change in translucency and color. Usually they become opaque and whitish because of the accumulation of spores or cysts in the internal tissues or fluids. Further, infected insects may exhibit a loss of appetite, abnormality of movement, and may remain moribund for long periods prior to death.

Only occasionally do flagellates cause acutal disease in insects. A few amebas cause disease. There is an important amebic disease in the honeybee and another in grasshoppers. In both of these maladies the insect's Malpighian tubes are especially involved. Several species of Ciliata have been found infecting insects; other are frequently found adhering to the exterior of insect larvae and pupae living in aquatic environments. Perhaps the best known are ciliates associated with mosquitoes. *See* AMOEBIDA; CILIATEA; MASTIGOPHORA; TRYPANOSOMATIDAE.

Sporozoa. The most important protozoan diseases in insects are caused by Sporozoa, of which the Microsporida as a group are the most serious. Many Gregarinida are found in insects. Most of the so-called eugregarines from the alimentary canal are not pathogenic for their hosts, but the schizogregarines found in the body cavity may be highly virulent. A number of Coccidia are also pathogenic for insects. *See* SPOROZOA.

Microsporida. Of the infections caused by members of the order Microsporida, perhaps the most famous is pébrine of the silkworm, caused by *Nosema bombycis*. Another well-known infection is nosema disease of the honeybee, caused by *Nosema apis*. Intensive studies of these two diseases have revealed a great deal about the Microsporida. Microsporidan infections have been found in species of at least 14 orders of Insecta. The most common insect hosts are Diptera, Lepidoptera, Coleoptera, and Ephemerida. Some of these are rather specific as to their insect hosts, while others are capable of infecting insects in different genera or even in different orders.

Spore stage. The most commonly seen form of any Microsporida is the spore stage, the characteristics of which are important in distinguishing the different species. The spore serves as the resistant stage of the organism and is able to tide the pathogen over periods of unfavorable environmental conditions and during the period between the change of host individuals. The average microsporidan spore is 3–6 μm long and 1–3 μm broad: some species exhibit a range of spore sizes. The shape of the spore is usually oval-spheroidal, although it may be spherical or bacilliform. The spore consists of a spore membrane or covering surrounding a sporoplasm and a polar filament coiled within the spore. The polar filament

is capable of extrusion as a very fine, long thread that may be several hundred micrometers in length in some species.

Microsporidan development. Although the life cycles of Microsporida vary from species to species, in general, soon after the spore is ingested by a susceptible host, the polar filament is extruded. A successful infection is initiated when the tip of the polar filament deposits sporoplasm from the spore either within or between the epithelial cells of the insect's intestine. The parasite, now called a planont, may remain in the intestinal cell or pass into the body cavity, where in either location it multiplies by binary fission. Various tissues of the body can be invaded, and the developing protozoon is now designated as a schizont or meront. In their intracellular location the spherical-to-oval meronts divide actively by fission, budding, or multiple division, filling the cytoplasm of the host cell. This period of schizogony ends in the formation of sporonts, each of which produces one sporoblast or many, depending on the genus of the microsporidian. Each sporoblast becomes a spore. This is the sporogony part of the life cycle. The entire life cycle is usually completed in about 4 days. Insects are usually infected with Microsporida through the mouth. However, some of these protozoa are transmitted from one generation to the next by way of the eggs.

Evidence is provided in India and the United States in support of the utility of *Nosema algerae* to infect larvae and for the microsporidia to interfere with the development of the malaria parasite in adult female mosquitoes. *Nosema locustae* infects grasshoppers and is a microsporidian which can be mixed with a bait to increase the consumption of the parasite by grasshoppers and consequently their mortality.

Nematode Diseases

Those roundworms for which insects serve as primary hosts are included in the classes Nematoda and Nematomorpha. Some members of the class Acanthocephala spend their larval stage in an insect host and the adult stage in mammals. More than 1000 species of roundworms have been reported from insects; most of these belong to Nematoda that frequently kill or seriously harm their hosts. At least 16 orders of insects are involved as hosts to nematodes. *See* ACANTHOCEPHALA; NEMATA (NEMATODA); NEMATOMORPHA.

All nematodes are similar in their general external appearance. Their elongate, unsegmented body has little or no variation in diameter, although it does taper toward one or both ends. There are three main stages in the developmental cycle of most nematodes: egg; juvenile, or larval, which includes four growth stages; and adult. Usually the young larvae spend a short period as free-living organisms, frequently in an aquatic environment.

The biological relations between nematodes and insects vary all the way from mere fortuitous association to obligate parasitism. For convenience these relationships may be separated into three general groups: (1) nematodes that live in the alimentary tract of the insect in a more or less commensal association; (2) nematodes that combine saprophagous and parasitic habits, that is, semiparasites; and (3) nematodes that parasitize the body cavity or tissues of their hosts, called obligate parasites. Examples of the first relationship are found in the alimentary tracts of crickets and cockroaches. These nematodes cause little, if any, harm to the insect. Of the second group, one is *Neoaplectana glasseri*, which infects the Japanese beetle. *Romanomermis culicivorax* is an obligate parasite which has shown some promise as a biological control agent against larval mosquitoes. Several commercial attempts in the United States to develop economical production of nematodes have not been successful. Among the best-known obligate parasites are the Mermithidae for safe and effective use in terrestrial habitats.

Noninfectious Diseases

Although most of the present activity in insect pathology appears to be concerned with the infectious diseases, increasing attention is paid to the noninfectious diseases of insects. These noninfectious diseases and abnormalities may be grouped as follows: (1) mechanical injuries such as traumata, bruises, and torn tissues; (2) injuries caused by physical agents such as burning, freezing, and drought; (3) injuries caused by poisons or chemical agents such as insecticides; (4) injuries caused by parasitization or infestation by other insects or arachnids; (5) diseases caused by a nutritional disturbance or a deficiency of proper nutrients or vitamins; (6) diseases caused by deranged physiology and metabolism; (7) inherited abnormal conditions or genetic diseases; (8) congenital anomalies and malformations; (9) certain tumors and neoplasms; and (10) disturbances in development and in regenerative capacity of tissues. *See* INSECT PHYSIOLOGY; INSECTA; INSECTICIDE.

John D. Briggs

Bibliography. N. E. Beckage et al., *Parasites and Pathogens of Insects*, 1993; A. E. Hajek and R. J. St. Leger, Interactions between fungal pathogens and insect hosts, *Annu. Rev. Entomol.*, 39:293–322, 1994; H. K. Kaya and R. Gaugler, Entomopathogenic nematodes, *Annu. Rev. Entomol.*, 38:181–206, 1993; P. R. Shewry and S. Gutteridge (eds.), *Plant Protein Engineering*, 1992; Y. Tanada and H. K. Kaya, *Insect Pathology*, 1997.

Insect physiology

The study of the functional properties of insect tissues and organs. Insects are one of the most diverse groups of animals on Earth. They have evolved adaptations that allow them to live in a variety of ecosystems. The basic body plan comprises a hard exoskeleton forming a protective covering over internal tissues. The interior of the exoskeleton is referred to as the hemocoel. Insects have an open circulatory system with hemolymph (blood) freely moving within the hemocoel and bathing all of the internal tissues. A basement membrane composed

of a noncellular matrix of proteins, lipids, and carbohydrates separates the internal tissues from the hemolymph. The fat body is a specialized tissue within the hemocoel. It is usually a whitish tissue found throughout the body of the insect and is involved in storing fat reserves. It is also involved in general metabolism and regulates the composition of the hemolymph. The fat body is, in fact, the site of production for the major proteins found in the hemolymph. *See* INSECT PATHOLOGY; INSECTA.

Development and Growth

Insects undergo development in stages. In general, development is either hemimetabolous or holometabolous. Hemimetabolous insects have three life stages: the egg, larva (or nymph), and adult. Holometabolous insects have four life stages: the egg, larva, pupa, and adult. The larval stage or nymph of hemimetabolous insects is similar in appearance to the adult. Holometabolous larval stages do not resemble the adult stage, and a pupal stage is present. Holometabolous development is considered complete metamorphosis because transformation from larva to adult occurs in the pupal stage. Hemimetabolous development is referred to as incomplete metamorphosis because of the gradual change from an immature nymph to an adult. The life stages are characterized by progressive development in which the cuticle is shed. Typically, the larval stages have more than one stage and undergo successive molts to become larger. *See* METAMORPHOSIS.

Embryonic development takes place within an eggshell called the chorion. Fertilization takes place within the egg. In some cases, parthenogenesis can occur, in which development takes place without fertilization. Many species of ants, wasps, and bees produce males through parthenogenesis. However, in these insects the majority of individuals are female and are produced by fertilization. In these insects the females are diploid (having two complete chromosome pairs in a nucleus), and the males are haploid (having half of the diploid or full complement of chromosomes). The summer generations of aphids also reproduce without fertilization, but all are female. Males can be produced in the fall for sexual reproduction. Most insects lay their eggs externally, after which they undergo embryogenesis and hatch. In some insects the eggs are retained in the female until hatching, and the mother gives birth to live young; aphids, tsetse flies, and some cockroaches are examples.

Upon hatching, the larva starts to consume food and grow. When growth reaches a certain size, the insect must shed its old cuticle, and a new cuticle is formed. This process is called molting or ecdysis. The cuticle of insects is part of the hard exoskeleton that is characteristic of the arthropods. The cuticle is an acellular layer that is produced by the underlying cellular epidermal cells. It is composed of several layers that are made up of protein, chitin, catecholamines, and lipid. Chitin is a complex polymer composed of repeating units of *N*-acetylglucosamine linked together. The chitin and proteins are cross-linked together by catecholamines such as *N*-acetyldopamine, and this cross-linking creates the rigidity found within the cuticle. The cuticle can become quite hard, or it can remain soft and flexible. The different types of protein, amount of chitin, and catecholamines are thought to contribute to the varied types of cuticle that are found in the insects. Based on the chemical properties of cuticle proteins, as determined by structural analysis and molecular genetic analysis, it is now known that the more acidic proteins are found in flexible cuticle. Molecular techniques have also revealed that epidermal cells and hemocytes (cellular elements of the hemolymph) are the source of these proteins, and at what time during development particular proteins are produced and become incorporated into the cuticle.

Lipids are located on the outer surface of the cuticle. These form a barrier that is primarily responsible for protecting against water loss in terrestrial insects and against water gain in freshwater insects. In addition, the nonpolar lipids help the insect resist attack by microbes. Some of the lipids found on the surface are also used in chemical communication. Hydrocarbons especially are utilized by a variety of insects for communicating.

The old cuticle is shed at each molt in a process that can occur over several days. Prior to ecdysis, the epidermal cells begin to divide and become close-packed and columnar in form. The old cuticle then breaks away from the epidermal cells in a process called apolysis. This creates a space between the epidermal cells and the old cuticle. New cuticle is then produced on top of the epidermal cells, and a molting fluid is secreted into the subcuticular space. After a portion of the new cuticle is produced and hardened, the molting fluid is activated and digests away some of the old cuticle. Many of the products of digestion are absorbed back into the epidermal cells. The next step is ecdysis, which is the actual shedding of the old cuticle. The old cuticle splits open, and the insect draws itself out. The old cuticle is then called an exuviae. The new cuticle is expanded and then hardens over several hours. The hardening process is partly due to the cross-linking of proteins. When they first emerge, many insects have a white cuticle, which slowly hardens and changes color to black or brown. The peptide hormone bursicon regulates the hardening of the cuticle after an ecdysis.

The overall process of molting is under the regulation of several hormones. The main ones are juvenile hormone and ecdysone. The juvenile hormone is responsible for maintaining the juvenile or larval stage upon subsequent molts. The actual molting hormone is ecdysone. Juvenile hormone ($Cl_6H_{26}O_3$) has a sesquiterpene type of structure, and is produced by endocrine glands called corpora allata. Ecdysone is a steroid that is produced by the prothoracic glands. It is a prohormone (a hormone precursor) that is converted to the active hormone, 20-hydroxyecdysone, by epidermal or fat body cells. In

general, it is the action of juvenile hormone that governs what type of cuticle and internal tissues will occur at the next molt. If juvenile hormone is present during a critical period of a larval stage, the next stage will be larval. For hemimetabolous insects, if juvenile hormone is absent during the last nymphal stage, the next molt will result in a change to the adult. Holometabolous insects have a pupal stage, and if juvenile hormone is absent during a critical period but present during a subsequent critical period, the next molt will result in a pupal stage. During the pupal stage, juvenile hormone is completely absent and the adult is formed within the pupal case. Ecdysone controls the molting process, and it is present before the molt to transform the tissues for the actual ecdysis.

A receptor protein for ecdysone has been identified through molecular genetic techniques. The receptor for ecdysone also binds to another protein called ultraspiracle. This complex of proteins bound to ecdysone is found within the nucleus and initiates gene transcription. Both ecdysone and juvenile hormone act by regulating the transcription of specific genes. The receptor for juvenile hormone has not yet been characterized. Other hormones are involved in regulating the complex process of actual ecdysis. These include the peptides eclosion hormone and ecdysis-triggering hormone. *See* ECDYSONE; ENDOCRINE SYSTEM (INVERTEBRATE); MOLTING (ARTHROPODA).

Diapause

Insects that live in temperate climates must be able to overwinter. In some cases, insects in the Northern Hemisphere migrate south for the winter. For example, the monarch butterfly overwinters in Mexico and southern California. However, most insects do not migrate and have developed a means to help them survive adverse conditions. This is called diapause, a physiological state into which the insect enters in advance of the actual onset of adverse environmental conditions. In most cases, diapause is also under genetic control. During diapause, insects become quiescent, metabolic rates decline, and the insects develop various physiological and biochemical adaptations to the cold environment.

The environmental conditions that induce diapause are primarily the photoperiod and temperature. Usually a critical day length induces diapause development. In temperate climates, diapause is triggered by shorter day length, which indicates that winter is approaching. Cooler temperatures act synergistically with photoperiod. Declining food availability is another indicator to some insects to enter diapause.

Diapause can occur in all of the different life stages of insects, including embryos in the egg stage. The silkmoth has an embryonic diapause that is determined by whether or not the mother underwent diapause. This insect has two generations per year, with the first generation diapausing as an egg. These eggs hatch and develop into adults that lay eggs that do not diapause. However, the second generation lays diapause eggs that survive the winter. Embryonic diapause in the silkmoth is under the control of a peptide hormone (diapause hormone).

Diapause can also occur in the larval, pupal, and adult stages. The same hormones that regulate metamorphosis (juvenile hormone and ecdysone) usually regulate these types of diapause. The cessation of feeding and an alteration in behavior to find a place to undergo diapause is characteristic of these stages. The stages also start producing cold-protective agents, usually alcohols such as glycerol that act as antifreezes. In insects that have an obligatory diapause, a certain period of cold exposure is required to break diapause and continue development. How an insect measures this cold period is probably coupled to the internal biological clock.

System Physiology

Insects, like all higher animals, have physiological mechanisms to support life. These functions are effected by the respiratory system for acquiring oxygen and releasing carbon dioxide, the muscular system for movement, the digestive system for acquiring nutrients, the circulatory system for movement of blood, the nervous system for perception of the environment and internal regulation, the reproductive system, as well as other systems.

Respiration. Most insects acquire oxygen and release carbon dioxide through a series of air-filled tubes called tracheae. The tracheae are connected to the cuticle and open externally through valves called spiracles. Insects can have as few as one pair or up to ten pairs of spiracles. The opening and closing of the spiracles in part regulates the airflow in the tracheae. Tracheae are composed of cuticle with underlying cells. They are large in diameter and supported by spiral thickenings of the cuticle called taenidia, which help to prevent collapse of the tracheae. Tracheae become progressively smaller as they integrate with the tissues, and are known as tracheoles. Almost every cell in the insect is associated with a tracheole. Diffusion accounts for the movement of oxygen within the tracheal system. Respiring cells acquire the oxygen, which creates a diffusion gradient. Carbon dioxide is released from the cells and can be converted first to bicarbonate for intracellular storage and buffering. When the bicarbonate concentration becomes too high, carbon dioxide is released as a gas and diffuses to the outside environment through the tracheae and spiracle system. Carbon dioxide causes the spiracle muscles to relax, resulting in opening of the spiracle.

In addition to the tracheae and tracheoles, insects have air sacs. These are characterized as large-diameter tracheae with reduced taenidia to allow for expansion and deflation. During periods of activity when an insect may need more oxygen, these air sacs expand and fill with air. To force air through the tracheal system, the spiracles close and the air sacs are forced to deflate. The rhythmic expansion and contraction of the abdomen along with the timing of spiracle opening and closing create active ventilation.

Aquatic insects also have a spiracle and trachea system. These insects, however, have adaptations to living under water. Some insects have a gill system that allows them to acquire oxygen dissolved in water, but the tracheal tubes are still gas-filled. Very few insects utilize hemoglobin for transport of oxygen. Diving insects acquire oxygen by periodically returning to the surface to replenish their air supply. In some cases, these insects create an air bubble that is maintained against the outer surface of the exoskeleton. This air bubble is held in place by hairs or under the wing, and is called a plastron. *See* RESPIRATION.

Nutrition and digestive system. Insects are the main competitors for food and fiber consumed by humans; indeed, they consume almost any type of food source. Various herbivorous insects eat almost every kind of plant material, either directly or by ingesting plant fluids. Many insects prey upon other animals; carnivore insects primarily consume other insects. Liquid blood is also consumed by mosquitoes, other biting flies, fleas, and some true bugs (of the order Hemiptera), which may make them transmitters of blood-borne pathogens. Larval stages of some insects, notably some Hymenoptera, are parasitic on other insects. The larvae feed on the living host but eventually kill it. These types of parasitic insects are called parasitoids.

Despite the varied dietary lifestyles of insects, nutritional requirements are similar. Insects require the 10 essential amino acids, vitamins, and choline. A protein source is usually required for development of eggs. Sugars are usually found in insect diets. Insects have a dietary requirement for cholesterol or another steroid, since they cannot make the cholesterol that is necessary for cell membranes and is the precursor for the steroid molting hormone ecdysone. Most insects also require polyunsaturated fatty acids for cell membrane dynamics and for biosynthesis of prostaglandins. However, some insects have the ability to make the diunsaturated fatty acid linoleic acid. Plants were thought to be the only source for this fatty acid until it was demonstrated that some insects could biosynthesize it.

Insects usually have a symbiotic relationship with microorganisms, which are found within the gut lumen and help the insect in digestion. In some cases, there are specialized cells called mycetocytes that house the symbionts. These cells are usually found within the hemocoel. The blood-feeding insects require the presence of the symbionts for survival, since they provide the vitamin B complex. These microorganisms can be passed on to the next generation by feeding to the young or through transovariole transmission, passing the symbiont to the developing egg.

The digestive tract of insects consists of a tube that runs from the mouth to the anus (**Fig. 1**). It is usually divided into the foregut, midgut, and hindgut. The foregut and hindgut are lined with cuticle. In most insects the midgut includes the peritrophic matrix, which is a chitin-protein complex that surrounds the food. Salivary glands release digestive enzymes and other factors in the mouth for extra-oral digestion and to begin the digestive process. The foregut takes in the food and in some cases can store food (in the crop). Most of the digestive enzymes are released from the midgut, and absorption of nutrients occurs in the ventriculus. The release of digestive wastes occurs in the hindgut. *See* DIGESTION (INVERTEBRATE).

Excretory system. The excretory system in insects involves the malpighian tubules, which produce the primary urine that flows into the digestive system and hindgut. These closed-end tubes connect to the gut at the intersection of the midgut and hindgut, and they are involved in secretion of nitrogenous and other waste products into the hindgut. The number of malpighian tubules varies from 2 to 200. They are usually free-moving within the hemocoel and clear

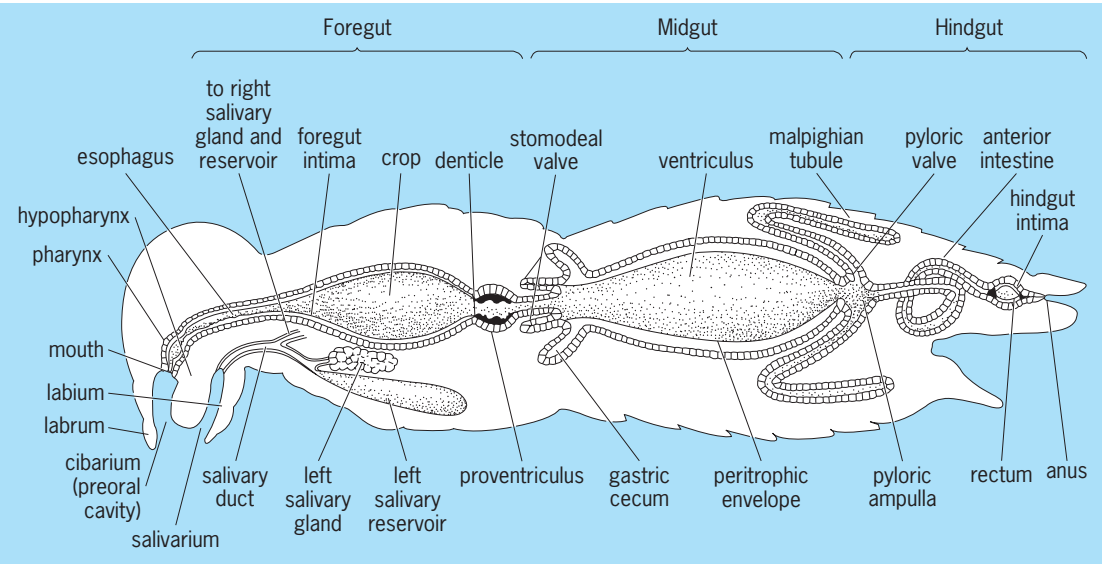

Fig. 1. Generalized insect digestive system. (*After W. S. Romoser and J. G. Stoffolano, Jr., The Science of Entomology, 4th ed., McGraw-Hill, 1998*)

the hemolymph of unwanted products. The primary urine formed in these tubules is water-based and contains nitrogenous waste products, usually in the form of uric acid for terrestrial insects and ammonia for aquatic insects. Other forms of nitrogenous waste are allatonin and urea. Other substances in the primary urine include ions (chloride, sodium, potassium), amino acids, and sugars. The ions are involved in providing the water flow through the malpighian tubules to the hindgut. Water is conserved in the hindgut and is absorbed back into the hemolymph. The hindgut also absorbs most of the amino acids and sugars and any ions that are in short supply. The feces consist of undigested food products and concentrated urine. *See* EXCRETION.

Water regulation. Water balance is maintained in insects by taking in water from the diet. The cuticle provides a very good barrier against water transport. The excretory system is also involved in water balance by regulating the amount of water lost in the feces. For example, if water needs to be conserved, the hindgut absorbs more water from the feces into the hemolymph. Two types of peptide hormones control water regulation in the excretory system. Diuretic hormones are released when the insect senses that too much water is present in the hemolymph. These act on the malpighian tubules to secrete more water and on the hindgut to inhibit the reabsorption of water. Antidiuretic hormones have been identified in a few insects, and these act primarily on the hindgut to increase the uptake of water from the feces. Some diuretic hormones have been identified; they are similar in sequence to the corticotropin-releasing factor of vertebrates. A receptor found on malpighian tubules of *Manduca sexta* (tobacco hornworm) was isolated and cloned. It possesses seven putative transmembrane domains, typical of the G-protein [GTP (guanosine 5′-triphosphate)–binding protein] coupled receptors. With the relatively recent sequencing of several insect genomes, other peptide hormones and their receptors have been identified. These diuretic hormones act on the malpighian tubules by stimulating the production of cyclic adenosine monophosphate (cAMP), which in turn regulates the ion transport machinery of the cells. Fluid secretion is stimulated by the active transport of ions and by a proton pump. The active transport of ions sets up a concentration gradient which water follows.

Circulation, hemolymph, and hemocytes. Insects have an open circulatory system. The exoskeleton forms the outer protective covering; internal tissues are found within the open hemocoel. The internal cavity is bathed in hemolymph or blood. Thus, all of the internal tissues are bathed in a hemolymph, with only a basement membrane separating the hemolymph from the tissues. Hemolymph flows backward and mixes within the insect by movement of the skeletal and visceral muscles. A dorsal vessel (essentially a hollow tube that runs the length of the insect) then circulates hemolymph forward again. This tube is composed of a one-cell-layer-thick heart that circulates hemolymph from the abdomen to the head. The heartbeat is produced by rhythmic contractions of the heart muscle. The heart is innervated by neurons, but since it will continue to beat in the absence of neural input the heartbeat is said to be myogenic. The heartbeat increases during activity, such as flight, and during the molting period. Several compounds affect the heartbeat, including biogenic amines such as octopamine and cardioacceleratory peptides. Special additional pulsatile organs are used to circulate hemolymph into the wings and antennae.

The hemolymph is composed of a variety of compounds. It can be colorless, or yellowish or green due to the presence of pigments. The most common sugar present in insect hemolymph is trehalose (a disaccharide of glucose units). Lipids are transported through the hemolymph in the form of diacylglycerol. Lipophorin is a protein carrier that holds the lipid and hydrocarbons for transport. A variety of proteins are present, most of which are made by the fat body. Insect hemolymph also contains high levels of amino acids, which are used to help maintain the proper osmotic pressure. Insects that are exposed to cold temperatures can also increase the levels of antifreeze alcohols, such as glycerol, in the hemolymph.

Hemocytes are the insect's blood cells. These are free-floating in the hemolymph, and are composed of several different types. Hemocytes are primarily involved in the insect immune response, which can be divided into cellular and humoral. Cellular responses involve the hemocyte first recognizing an invading pathogen. How this recognition occurs is not well understood but is most likely based on cell surface chemistry of the pathogen. The next response is the phagocytosis of pathogens that are smaller than the hemocyte, such as bacteria. If phagocytosis does not clear the infection, nodulation occurs as the next response. This is a process whereby hemocytes bind to the pathogen and form aggregates of hemocytes and pathogens. The aggregate grows to form a small nodule within the hemocoel, within which pathogens are trapped and killed. Larger pathogens, such as parasitoids, are encapsulated within several layers of hemocytes. The humoral immune response is characterized by the production of antibacterial peptides; it is similar to the innate response of vertebrates. Hemocytes or the fat body produces these peptides, and their production continues for several weeks after the primary infection. This provides the insect with a degree of immunity against further attack. *See* CIRCULATION.

Muscles. Insect muscles are similar to those found in other organisms except that all insect muscles are striated; there are no smooth muscles. The skeletal muscles are connected to the exoskeleton by cuticle structures called apodemes. These muscles move the sclerites (hardened outer plates) of the insect, allowing the insect to move. Visceral muscles are found in most internal organs, including the gut, heart, and reproductive tissues. Skeletal muscles can be further differentiated into synchronous and asynchronous muscles. These differ based on motor neuron

activity and contraction of the muscle. Synchronous muscles contract once for every neuronal impulse, whereas asynchronous muscles have several contractions per neural impulse. Synchronous muscles are found throughout the body and can be both skeletal and visceral. Asynchronous muscles are found only as flight muscles and occur in pairs. A feature of asynchronous muscles is that they are stretch-activated; thus one set of muscles contracts slightly, stretching the other pair. This stretching then activates the muscle for a contraction. In this manner the asynchronous muscles work together and can maintain a contraction-relaxation rate independent of neural input for several cycles. However, a neural input is required to maintain the contraction rate of asynchronous muscles over longer periods.

Asynchronous muscles are primarily used as flight muscles that are of the indirect kind, meaning that they do not directly move the wing; instead they are connected to the exoskeleton of the thorax. Contraction of the indirect muscles moves the body wall of the thorax, resulting in movement of the wing. There are two types of indirect flight muscles, dorsal-longitudinal and tergosternal muscles (**Fig. 2**). When the dorsal-longitudinal muscles contract, the top of the thorax moves up; and when the tergosternal muscles contract, the top of the thorax moves down. Thus these muscles operate alternately to cause movement of the wing. Direct muscles are attached to the wing in the hinge region and provide control of aerodynamic forces placed on the wing during flight. Insects that have asynchronous flight muscles also have the highest wing beat frequencies; for example, some small flies have wing beat frequencies greater than 500 beats per second. The indirect asynchronous muscles contract and relax once during each wing beat; thus there is the need for a discontinuous neural input. In contrast, insects that use synchronous flight muscles have wing beat frequencies typically lower than 50 beats per second.

A consequence of the high contraction rate of flight muscles is that they also have the highest metabolic demands of any muscles. High metabolic rates require the availability of oxygen, which is supplied by the tracheal system. In some cases the tracheoles actually penetrate deep within the muscle. The fuel to supply muscle contraction comes from adenosine triphosphate (ATP), but the amount of ATP present in resting muscles would last only 1 second during flight. Therefore, muscles must regenerate the ATP, and this takes place in the mitochondria. For most insects, sugar in the form of glycogen is broken down to provide energy for muscles. In those insects that are constantly taking in sugar, such as foraging honeybees, the diet supplies the necessary energy for the flight muscles. In those insects that undertake long-distance flight, such as migrating insects like the locust or monarch butterfly, stored lipid is utilized as a fuel source. The lipid is stored in the fat body as triacylglycerol and transported to the muscles through the hemolymph as diacylglycerol, which is carried by the protein lipophorin. This process is regulated by peptide hormones (adipokinetic hormones) produced in the corpora cardiaca near the brain.

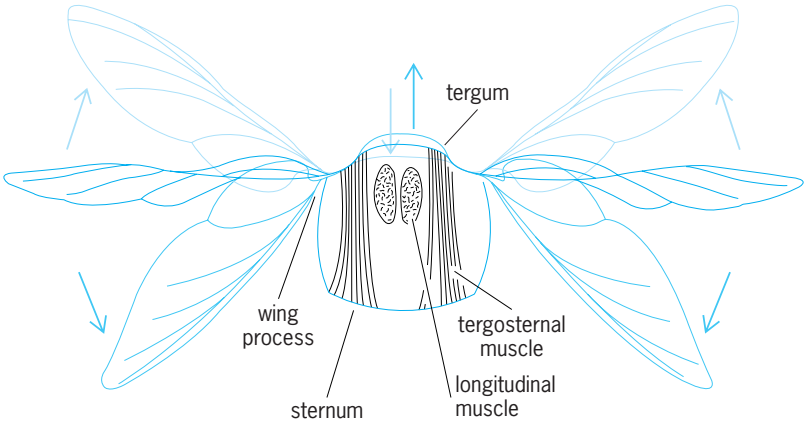

Fig. 2. Wing movements in the flight of an insect. (*After T. I. Storer and R. L. Usinger, General Zoology, 4th ed., McGraw-Hill, 1965*)

In some larger insects the thoracic flight muscles generate a considerable amount of heat, necessary to maintain the optimal temperature for muscle action. A preflight warm-up period is used to heat the muscles to this optimal temperature. In some large sphingid moths, this temperature is around 35°C (95°F), which can be considerably higher than the environmental temperature. During flight the thoracic temperature is kept from rising too high by a modification of the circulatory system in which excess heat is transferred to the hemolymph and pumped back to the abdomen. The abdomen is cooler, and the excess heat is lost to the environment. Thus, some insects can thermoregulate body temperature before and during flight. *See* FLIGHT; MUSCULAR SYSTEM.

Programmed cell death. Larval insects have intersegmental muscles that connect abdominal segments and are used in larval locomotion. During the ecdysis to the pupal stage, some of these muscles degenerate and die. After the adult emerges from the pupal case, the remaining intersegmental muscles die. The muscle cells that undergo programmed cell death (apoptosis) are regulated by the decline in the amount of ecdysone and 20-hydroxyecdysone. As in *Manduca sexta*, the decline in ecdysteroid amount results in the activation of certain genes and the repression of other genes in these muscles. Two of the repressed genes encode for actin and myosin, which are the contractile proteins of muscles. Two of the stimulated genes encode for apolipophorin and polyubiquitin. Polyubiquitin genes encode multiple forms of ubiquitin, which is a 76-amino-acid protein found in all eukaryotic cells. The main mechanism for the specific degradation of proteins is an ATP-requiring addition of ubiquitin to proteins. These tagged proteins are then rapidly degraded by the proteasome. Thus, cell death is not simply the cessation of muscle growth, but is due to the programmed development of specific pathways that lead to cell death. The loss of intersegmental muscles in *M. sexta* is a model system that has been used

to study programmed cell death in eukaryotes. Other cells within the insect also undergo programmed cell death during metamorphosis, for example, specific neurons within the central nervous system.

Reproductive system. The internal structures of reproduction are the testis in the male and the ovaries in the female. The external reproductive structures are quite varied among the different types of insects and in some cases are the only structures that separate species.

Sperm and eggs. The internal male reproductive structures consist of the testis and tubes for carrying the sperm to the ejaculatory duct. In most cases there is also a seminal vesicle for storage of sperm. There are accessory organs that produce the seminal fluid that is transferred to the female with the sperm. In some insects the sperm and seminal fluid are encapsulated into a structure called a spermatophore. This spermatophore can be preformed before copulation or can be made during copulation. In some insects the spermatophore is used as a mating plug that blocks the genital tract of the mated female to help prevent other males from fertilizing the female.

The internal female reproductive structures consist of the paired ovaries connected to a common oviduct for laying eggs. Females also have accessory organs for producing material that is laid along with the eggs. A spermatheca is present for the storage of sperm. During mating, the spermatophore is first placed in a bursa copulatrix and then the sperm is transferred to the spermatheca. In some insects the sperm can be stored and remain viable in the spermatheca for the entire life of the female. This is especially so for the social queen female, which mates once and then fertilizes thousands to millions of eggs using sperm from the one male.

The process of egg development is called oogenesis. Each ovary is made up of a few to 100 ovarioles. In some insects each ovariole produces one egg at a time, while in other insects each ovariole has continual egg development. During egg maturation, yolk is formed inside the developing eggs. Most of this yolk is protein and lipid. The main protein that is found in the yolk is called vitellogenin. It is produced by the fat body, secreted into the hemolymph, and taken up by the ovaries through a process of selective endocytosis. This occurs before the chorion or eggshell is formed. The receptors on the ovarioles that aid in the uptake of vitellogenin have been identified at the molecular level, and they are similar to the low-density lipoprotein receptors found in mammals.

In general, there are two types of reproductive patterns of females. Females can continuously develop eggs, or they can produce eggs in discrete batches. In the latter group the synchronization of egg development is usually triggered by a meal or the laying of a previous batch of eggs. For example, some mosquitoes develop eggs in batches, and this step is governed by the taking of a blood meal. In most females, juvenile hormone is the main hormone involved in regulating egg development. Juvenile hormone stimulates the production of vitellogenin by the fat body and also the uptake of vitellogenin by the ovaries. In the Diptera (mosquitoes and flies), ecdysone is the hormone that stimulates the fat body to produce vitellogenin.

Fertilization. Once the egg has been fully developed, fertilization takes place just prior to being laid. The sperm enter the egg through the shell or chorion through special openings called micropyles. Oviposition, or the process of laying eggs, is a complex behavior that requires the right environmental conditions, and also the availability of a suitable substrate. For most insects, this involves a lot of visual, physical, and chemical cues to find the proper egg-laying substrate.

Mating also has a great influence on egg development and oviposition. The male in most cases is passing not only sperm to the female but also chemical information that will change her behavior. This behavior change usually results in increased oviposition and decreased receptivity toward other males. For example, *Drosophila* males pass a sex peptide to the female that changes her behavior and results in decreased sexual receptivity toward other males. Part of the decreased receptivity is a result of lower sex pheromone levels of the female. This helps ensure that the male's sperm will fertilize the eggs that the female lays. In crickets it has been shown that prostaglandins will stimulate oviposition in females. The male transfers both the substrate for prostaglandin formation and the enzyme prostaglandin synthetase to the female to stimulate an increased oviposition rate. In those insects where the female mates more than once, evolution has produced mechanisms whereby the male can increase the probability that his sperm will fertilize the eggs and his genes will be passed on to the offspring.

Most insects are oviparous, reproducing by laying eggs. A few insects are viviparous, reproducing by bearing live young. For example, in the tsetse fly, eggs hatch within the mother and the larvae receive nourishment from her until they are ready to pupate. The mother develops one larva at a time. The aphids are also viviparous and give birth to live young. Aphids are also parthenogenetic, meaning that reproduction can occur without fertilization. In the case of aphids, females are produced during the summer months parthenogenetically and then males are produced in the fall. Males and females mate and the females lay eggs that overwinter. The next year females hatch from the eggs to start the cycle over again.

Nervous system. Insects, like other higher organisms, have complex nervous systems. The nervous system processes information coming in from the environment via sensory systems mediating vision, touch, taste, smell, and hearing; and it produces a reaction to that information resulting in a typical behavior. The processing may result in motor neuron activity stimulating sets of muscles. For example, the tactile information that the feet are losing touch with the ground is sufficient to start the neuronal flight pattern. Wind-sensitive hairs on the tail

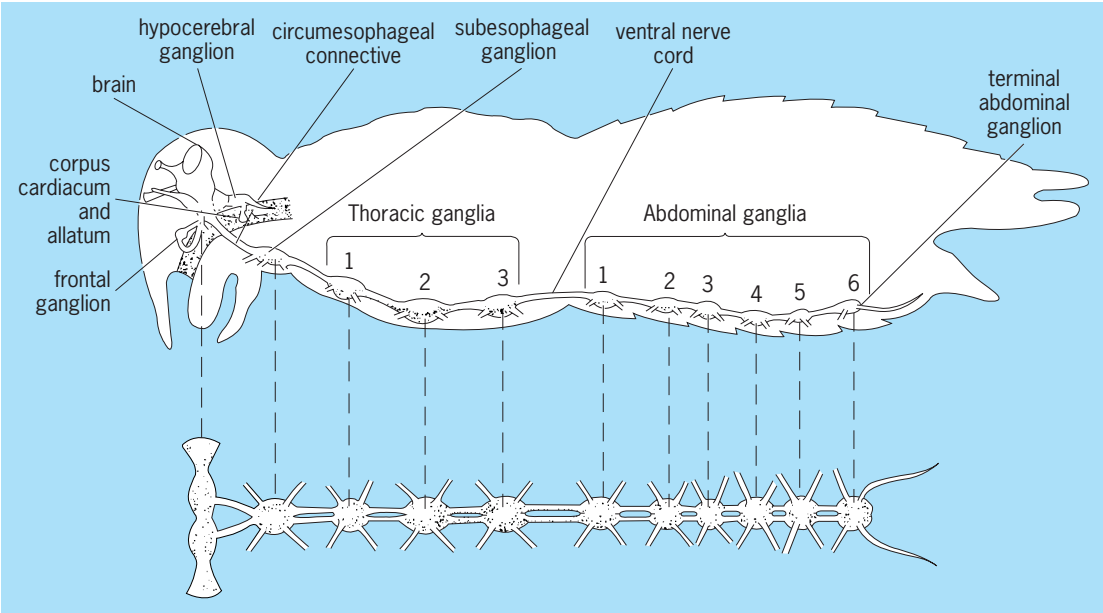

Fig. 3. Generalized insect nervous system. (*After W. S. Romoser and J. G. Stoffolano, Jr., The Science of Entomology, 4th ed., McGraw-Hill, 1998*)

end of a cockroach detect slight air movement and prompt the cockroach to run away. These events are integrated within the nervous system of the insect.

Morphologically the nervous system consists of segmental ganglia that are connected together by a paired solid nerve cord (**Fig. 3**). These ganglia are found ventral in the hemocoel, and in the primitive state there is one per segment. However, in most insects some of these ganglia have fused together. The brain is composed of several ganglia and is responsible for information coming from the eyes (vision), antennae (smell), and mouthparts (taste). The thorax usually has three ganglia, responsible for movements of the wings and legs. Each abdominal segment has smaller ganglia, which have control over the internal tissues and abdominal movement. These ganglia are all connected together by a paired ventral nerve cord. The ganglia themselves contain the individual nerve cells (neurons), with the axons traveling either through the ventral nerve cord or out to the periphery through nerves. Sensory neurons are located in the periphery and send their axons into the central nervous system. Sensory neurons are coupled together with motor neurons by interneurons. Together all of these neurons are involved in integrating information and regulating the behavior of the insect. *See* NERVOUS SYSTEM (INVERTEBRATE).

Vision. Adult insects usually have two kinds of light-detecting organs: compound eyes and ocelli (small simple eyes composed of photoreceptor cells and pigment cells). Hemimetabolous insects have compound eyes throughout development, whereas holometabolous insects have compound eyes as adults and stemmata during the larval stages. Stemmata are simple compound eyes that can discriminate shapes and detect movement. However, they probably are not involved in composing a visual image, and holometabolous larval insects rely more on chemical and mechanical receptors to learn about their environment.

The compound eyes of insects are composed of a few to thousands of individual light-gathering structures called ommatidia (**Fig. 4**). If only a few ommatidia are present, they are usually widely spaced and circular in pattern. If there are many ommatidia, they make up the hexagonal faceted eyes of most insects.

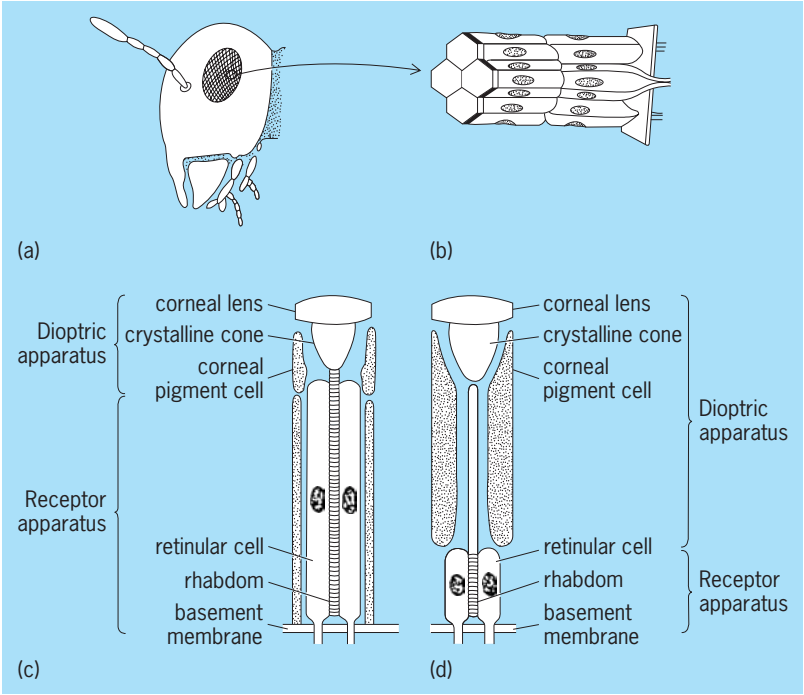

Fig. 4. Compound eye structure. (*a*) Head with compound eye. (*b*) Four ommatidia removed and magnified. (*c*, *d*) Representative types of ommatidia. (*After W. S. Romoser and J. G. Stoffolano, Jr., The Science of Entomology, 4th ed., McGraw-Hill, 1998*)

Light strikes photopigments in the ommatidia, and this information is transduced into a neural signal that is sent to the optic lobe of the brain. Neural centers within the brain process this information into an image that is thought to be right side up (superposition). Each ommatidium has a visual field of view of about 3–8°. However, the large number of ommatidia can cover a greater field of view, in some cases over 180°. Thus with both compound eyes an insect can detect images in a wide field of view without the need to turn the head.

The outer portion of an individual ommatidium has a biconvex corneal lens that is actually cuticle and an underlying crystalline cone (Fig. 4). These structures focus the light onto sensory cells called retinular cells. The corneal pigment cells are primarily for isolation of light between individual ommatidia to prevent light from bouncing around below the cornea. These pigments can move up or down to allow more light to be detected in low-light conditions such as at night. There are usually eight retinula cells arranged radially with light-detecting portions located in a central structure called a rhabdom. The retinular cells are monopolar sensory neurons that send their axons into the optic lobe of the brain. Light-sensing pigments known as rhodopsins are located in the rhabdom and can detect different wavelengths of light. Most insects detect blue, green, and ultraviolet wavelengths, with some detecting red wavelengths.

Some insects that fly at night, such as the moths, have developed a clear-zone type of eye, or a superposition eye (Fig. 4*d*). The pigment in the corneal pigment cells moves up around the crystalline cone, leaving a clear area between the cone and the retinular cells. This happens in the dark-adapted eye to allow more light to be detected. Light entering one ommatidial lens can cross over and be detected by adjacent ommatidia. Moths also have a tapetum, in which tracheae run parallel to the retinular cell and reflect light back into it. This allows more light to be detected but is thought to decrease the resolution of the final image. However, the image resolution is apparently still good enough to allow the insect to fly at night.

The other type of eye is an apposition eye, and it is characterized by the retinular cells located just below the crystalline core (Fig. 4*c*). The corneal pigment cells help to concentrate light onto the rhabdom of individual ommatidia. Apposition eyes are typically found in insects that fly during the day (diurnal).

Visual acuity is hard to judge in insects, but in general it is thought to be inferior to that of humans. However, insects such as the honeybee can discriminate certain shapes and patterns. Most insects can detect polarized light coming from the Sun using the ultraviolet light receptors and the orientation of the rhabdoms in the retinular cells. Detecting polarized light is important in helping the insect determine direction. Insects also have the ability to detect a high frequency of light flashes. This is important for fast-flying insects to detect objects and for stationary insects to detect other insects flying.

Ocelli are visual organs found in most insects. They are usually paired on the dorsal surface of the head with one located medially. These have a single corneal lens with many retinular cells located beneath the lens. They do not form images but are used to detect the intensity of light. Flying insects use them to help determine the position of the horizon. *See* EYE (INVERTEBRATE).

Chemical sensilla and pheromones. The senses of taste and smell are involved in detecting chemicals in the environment. Smell is involved with detecting airborne chemicals (olfaction), and taste is involved with detecting chemicals in the food. Each has a unique type of sensory structure called sensilla. A sensillum is similar to a tactile hair where the cuticle forms a hair with an underlying sensory neuron and accessory cells. Contact sensilla involved in taste have one pore in the cuticle at the tip of the hair, whereas the sensilla detecting airborne chemicals have multiple pores in the cuticle hair. The shape of the hair can be varied: from hairlike structures found on antennae of moths, to short peglike structures on antennae of locusts, to pits on antennae of flies. Physiologically these sensory cells act in a similar manner. The compound being detected crosses the cuticle through a pore and comes in contact with a receptor lymph fluid. The receptor, which is a monopolar neuron, senses the compound, and this causes the neuron to depolarize and a signal is sent down the axon to the brain. It is unclear how the receptor neuron actually senses a specific compound, but it probably involves the compound binding to a receptor on the sensory neuron. Putative receptor proteins have been identified recently in insects using molecular cloning techniques. These proteins belong to the superfamily of receptors named G-protein-coupled receptors (GPCR), with characteristic seven-transmembrane domains. Identification of a wide variety of chemoreceptors has been aided by the recent sequence identification of several insect genomes.

Contact chemosensilla involved in the sense of taste can be found on various body parts: antennae, mouthparts, and feet. For example, the tarsi (feet) of flies have over 3000 contact chemosensilla for the detection of sugar in food. If these sensilla come in contact with a high enough concentration of sugar, the proboscis (tubular organ) extends and the insect probes with additional sensilla on the mouthparts. If sufficient stimulatory sensory information is present, the insect feeds. In addition to detecting stimulatory compounds, insects can detect deleterious compounds. For example, the fly readily feeds on a sugar solution, but if a certain concentration of salt is also present the fly does not feed. Most insects have similar types of sensilla for detecting a wide variety of compounds, and these compounds can stimulate or inhibit feeding. This is important for phytophagous (plant-feeding) insects that usually feed on only one type of plant. Plants have produced a wide variety of compounds to inhibit insect

feeding, and in turn insects have developed ways of detecting these compounds. In some cases the insect eats the plant containing the deleterious compound anyway but must inactivate the compound in some manner.

Chemosensilla detecting airborne or volatile compounds are usually found on the antennae. These chemosensilla have hairlike cuticle structures with an underlying sensory neuron. The volatile compound, usually hydrophobic, enters the multiporous hair and comes in contact with the aqueous sensory lymph. In this sensory lymph are proteins that bind to the volatile compound. These odorant-binding proteins probably function in the transport of the volatile compound across the sensory lymph to the sensory neuron cell, where the volatile compounds are detected. Just like the contact chemosensory cell, the neuron then depolarizes and a signal is sent down the axon to the antennal lobe of the brain. The brain integrates this information and produces a behavioral response. The brain antennal lobe morphology is such that the integration areas are formed into ball-like structures called glomeruli. These neuronal glomeruli are groups of axons that originate from specific sensory hairs on the antennae. These sensory neurons relay the information to other neurons that send the information to higher-order processing centers within the brain. In moths a set of these glomeruli is found only in males and receives the information coming from specific sensory neurons on the antennae for detecting sex pheromones.

Some of the sensory neurons on the antennae are involved in detecting general types of odors such as host-plant odors. The other sensory neurons are involved in detecting specific types of chemicals. Pheromones are chemicals used by insects to communicate between individuals of the same species. In general there are four different types of pheromones: sex, aggregation, alarm, and trail. Sex pheromones are chemicals produced by one sex to attract an individual of the opposite sex. Aggregation pheromones are used to attract both sexes to a particular site for feeding or for mating. Alarm pheromones are released to signal other individuals of danger; these are primarily released by social insects and by aphids. Trail pheromones are released to mark a given area or territory and are usually less volatile than the other types of pheromones.

A considerable amount of research has been conducted on sex pheromones. Moths represent a group of insects that utilize sex pheromones, and in most cases it is the female that is producing the volatile sex pheromone to attract the male. Many factors are important in order for a male moth to find the female, including environmental factors such as time of day, wind speed, temperature, and humidity. The female produces a blend of chemicals that are usually species-specific, and the male moth uses this blend of chemicals to find the female. The female has a special pheromone gland that makes and releases the chemicals into the air. These chemicals float downwind, forming a plume of odors. The male detects these chemicals with specific chemosensilla on the antennae, which results in a typical upwind flight pattern to search for the female. Males are very sensitive to the sex pheromone molecules, and in some cases only a few molecules are detected. In the silkmoth it was calculated that only about 300 molecules need to be detected in order for the male to initiate the upwind flight behavior.

Most pheromones are a blend of components, most of which are needed to maintain a search flight for the female. There are also compounds, some of which are pheromones for other species, which inhibit male flight. An example is the gypsy moth and the nun moth. The gypsy moth female produces a chemical called (+)-disparlure (7*R*,8*S*-2-methyl-7,8-epoxyoctadecane). The nun moth female produces both (+)-disparlure and (−)-disparlure. These compounds are the same compound except the epoxide is an optical isomer (+ is 7*R*,8*S*; − is 7*S*,8*R*). The gypsy moth male can detect both isomers, with the (+)-isomer producing upwind flight and the (−)-isomer inhibiting upwind flight. In contrast, the nun moth male initiates upwind flight only if both isomers are present. This is a way to separate the two species so the male will find the right female, and sex pheromones help to ensure species separation.

There have been a variety of chemicals identified that insects utilize as pheromones. These can be quite complex to relatively simple structures. The first was identified in the silkmoth in 1959 as an alcohol with 16 carbons and two double bonds (*E*10,*Z*12-hexadecadien-1-ol). In some cases the chemical structure is of sufficient complexity that no other chemical in nature would have the same structure. Another strategy that insects utilize is to use a blend of chemicals that are specific to one species. The blend of chemicals becomes the pheromone. These chemicals can be volatile and airborne as occurs with the moths and beetles, or they can be nonvolatile contact chemicals as occurs with cockroaches and flies. In most cases the female produces the pheromone chemical in specialized glands, and the production of the pheromone can be under endocrine regulation. Juvenile hormone has been shown to regulate pheromone production in adult cockroaches and beetles. In adult flies it is either juvenile hormone or ecdysone. In the moths it is a peptide hormone produced in the subesophageal ganglion of the brain. *See* PHEROMONE.

Touch and hearing. The sense of touch in insects is concerned with detecting forces put on the body by mechanical energy. This energy may originate either internally through the action of muscles or externally by the environment. Perception of the various forms of mechanical energy results in the sensing of gravity, pressure, touch, position, vibration, and hearing. Sensory cells called mechanosensilla detect these mechanical forces. These sensilla are quite variable in structure, but they all possess a neuron with dendrites reaching the cuticle, and an axon that sends an impulse to the central nervous system. The variability is due to the number of neurons present

and the type of cuticular structure. In addition, there can be both mechanosensilla and chemical sensilla in one structure. Essentially all the mechanosensilla can detect mechanical forces placed on the cuticle. When the cuticle is deformed, the dendrite senses this, resulting in depolarization of the membrane, and a neural impulse is sent down the axon to the central nervous system.

The cuticle is shaped differently for different types of mechanosensilla. The most common are the hair-shaped or trichoid sensilla. These are found all over the insect's body and are primarily involved as tactile receptors for touch. They also can detect air movement. In general a neural impulse is sent when the hair is bent and again when it is straightened. Hairs associated with joints have mechanosensilla where an impulse is sent continuously whenever the hair is bent. These are involved in determining positioning of the appendages. The head contains groups of these sensilla that are used to detect air movement. They are used to help determine wind direction when the insect is on the ground, and in flight to help control wing movement in relationship to the wind. Some of these hairs on the head are also involved in helping the insect detect the gravitational pull of the Earth. This is important in all insects to determine up and down, but especially important in the communication dance of honeybees, where up or down movement relays information to nest mates.

The campaniform sensilla are another type of mechanosensilla. These are characterized by the cuticle forming a cap to which the sensory neuron is attached. Again, these sensilla are detecting forces placed on the cuticle. These are usually found at the base of the wing and around other joints. Campaniform sensilla are important in helping the insect coordinate position of the legs while walking and wing position during flight.

A further type of mechanosensilla are the chordotonal organs that are responsive to vibrations. These sensilla may have no external sign of their presence. They are located beneath the cuticle but are attached to the cuticle by some kind of cell. If these sensilla are found positioned on the cuticle, in general they are involved in determining position of body parts. In some cases, these sensilla are tuned to detect certain sound frequencies produced by flying insects. For example, the larvae of the cabbage moth has eight sensilla on its thorax that can detect the vibrations produced by a flying yellow jacket wasp. When the larva hears that particular vibration, it stops moving or falls to the ground to avoid predation by the wasp. The other types of chordotonal organs are found in specific locations on the insect body. Some are found on the leg of crickets (subgenual organ) that function in hearing. These can detect vibrations coming from the substrate that the cricket is standing on. Others are found at the base of the antennae, forming the Johnston's organ. Insects use these to help determine position of the antennae, and in male mosquitoes they are used to detect the sound of flying female mosquitoes. Tympanal organs are the true organs of hearing in insects, and they are found on the thorax, legs, or abdomen. These organs have thin areas of cuticle to which chordotonal sensilla are attached. Sound waves vibrate the thin cuticle, and the vibration is picked up by the sensilla. These types of hearing organs are found in moths, grasshoppers, and cicadas. Moths use these organs to detect the echolocation cries of hunting bats, allowing the moths to take evasive action in flight. *See* PHONORECEPTION.

Sound and light production. Many insects communicate by producing sound. The sound can travel either through the air or through the substrate (the surface upon which an insect is standing). Many insects simply strike the substrate with a body part in a regular fashion to produce a vibration in the substrate. This is used as some kind of signaling mechanism to other insects of the same species. Another way of producing sound involves modifications of the cuticle surface. Stridulation refers to sound production by moving a cuticular scraper over a toothed file (a rib providing wing support, also called a vein). Crickets and grasshoppers produce their calling sounds in this way. Sound production in some crickets is produced by rubbing the scraper on one forewing along the file of the other forewing. Grasshoppers produce sound by rubbing the femur of the leg against the forewing. Another way to produce sound is exemplified by the high-pitched sound of the cicadas. These insects use a tymbal mechanism located in the first abdominal segment. The tymbal is a thin cuticle composed primarily of a flexible protein called resilin and a series of cuticular ridges. It is connected to a muscle that contracts to buckle the tymbal inward. This inward buckling produces the sound and, with repeated muscular contractions, yields sound patterns. Air sacs behind the tymbal amplify the sound and send it outward. Usually only the male produces sound to attract a conspecific (referring to individuals of a single species) female and for courtship behavior.

Several insects produce light without the aid of luminescent bacteria. Primarily several groups of beetles are known to produce their own light, notably the fireflies or lightning bugs. These beetles have a special organ usually found on the ventral side of the abdomen below a clear cuticle. The cells have a rich tracheal supply, providing the oxygen required for the light-producing reaction. A low-molecular-weight compound, known as luciferin, is the substrate for an enzyme called luciferase. This enzyme catalyzes a reaction using energy and oxygen to oxidize luciferin. The oxidized luciferin spontaneously decays and in the process gives off light. Millions of these reactions occurring simultaneously produce one flash. The regular flashes of the firefly are regulated by the nervous system and the amount of oxygen presented to the light organ cells. These regular flashes are used primarily in communicating to the opposite sex for mating purposes. *See* BIOLUMINESCENCE.

Russell Jurenka

Bibliography. R. F. Chapman, *The Insects: Structure and Function*, 1998; P. J. Gullan and P. S. Cranston, *The Insects: An Outline of Entomology*, 3d ed., 2005; M. J. Klowden, *Physiological*

Systems in Insects, 2002; J. L. Nation, *Insect Physiology and Biochemistry*, 2001; H. F. Nijhout, *Insect Hormones*, 1994; W. S. Romoser and J. G. Stoffolano, Jr., *The Science of Entomology*, 4th ed., 1998.

Insecta

A class of the phylum Arthropoda, sometimes called the Hexapoda. In fact, Hexapoda is a superclass consisting of both Insecta and the related class Parainsecta (containing the springtails and proturans). Class Insecta is the most diverse group of organisms, containing about 900,000 described species (see **table**), but there are possibly as many as 5 million to perhaps 20 million actual species of insects. Like other arthropods, they have an external, chitinous covering. Fossil insects dating as early as the Early Devonian have been found. *See* PARAINSECTA.

Approximate numbers of described species of insects in the modern world

Insect order	World species
Diplura	750
Archaeognatha	415
Zygentoma	320
Ephemeroptera	2,200
Odonata	5,100
Plecoptera	2,000
Grylloblattodea	25
Orthoptera	18,700
Phasmatodea	2,300
Mantodea	1,800
Blattodea	3,500
Isoptera	2,000
Dermaptera	1,500
Embioptera	2,000
Zoraptera	25
Psocoptera	2,600
Phthiraptera	3,500
Thysanoptera	8,800
Hemiptera	80,000
Megaloptera	260
Raphidioidea	100
Neuroptera	4,200
Coleoptera	340,000
Strepsiptera	370
Mecoptera	450
Siphonaptera	1,700
Diptera	102,200
Trichoptera	6,500
Lepidoptera	137,200
Hymenoptera	141,100

Classification

The class Insecta is divided into orders on the basis of the structure of the wings and the mouthparts, on the type of metamorphosis, and on various other characteristics. There are differences of opinion among entomologists as to the limits of some of the orders. The orders of insects (and their relatives the parainsects) are shown below.

Superclass Hexapoda
- Class Parainsecta
  - Order: Protura; Proturans
    - Collembola; springtails
- Class Insecta
  - Subclass Monocondylia
    - Order: Diplura; telsontails
      - Archaeognatha; bristletails
  - Subclass Dicondylia
    - Infraclass Apterygota
      - Order: Zygentoma; silverfish, firebrats
    - Infraclass Pterygota
      - Section Palaeoptera
        - Order: Ephemeroptera; mayflies
          - Odonata; damselflies and dragonflies
      - Section Neoptera
        - Hemimetabola
          - Order: Plecoptera; stoneflies
            - Grylloblattodea; rockcrawlers
            - Orthoptera; grasshoppers, katy-dids, crickets
            - Phasmatodea; walkingsticks
            - Mantodea; mantises
            - Blattodea; cockroaches
            - Isoptera; termites
            - Dermaptera; earwigs
            - Embioptera; webspinners
            - Zoraptera; zorapterans
            - Psocoptera; psocids, booklice
            - Phthiraptera; lice
            - Thysanoptera; thrips
            - Hemiptera; cicadas, hoppers, aphids, whiteflies, scales
        - Holometabola
          - Order: Megaloptera; dobsonflies, alderflies
            - Raphidioidea; snakeflies
            - Neuroptera; lacewings, antlions
            - Coleoptera; beetles
            - Strepsiptera; twisted-wing parasites
            - Mecoptera; scorpionflies
            - Siphonaptera; fleas
            - Diptera; true flies
            - Trichoptera; caddisflies
            - Lepidoptera; moths, butterflies
            - Hymenoptera; sawflies, wasps, ants, bees

Morphology

Insects are usually elongate and cylindrical in form, and are bilaterally symmetrical. The body is segmented, and the ringlike segments are grouped into three distinct regions: the head, thorax, and abdomen (**Fig. 1**). The head bears the eyes, antennae, and mouthparts; the thorax bears the legs and wings, when wings are present; the abdomen usually bears no locomotor appendages but often bears some appendages at its apex. Most of the appendages of an insect are segmented.

Body wall. In vertebrate animals the skeleton is on the inside of the body and is spoken of as the endoskeleton; in an insect the skeleton is primarily on the outside of the body and is called an exoskeleton. However, important endoskeletal structures occur, particularly in the head.

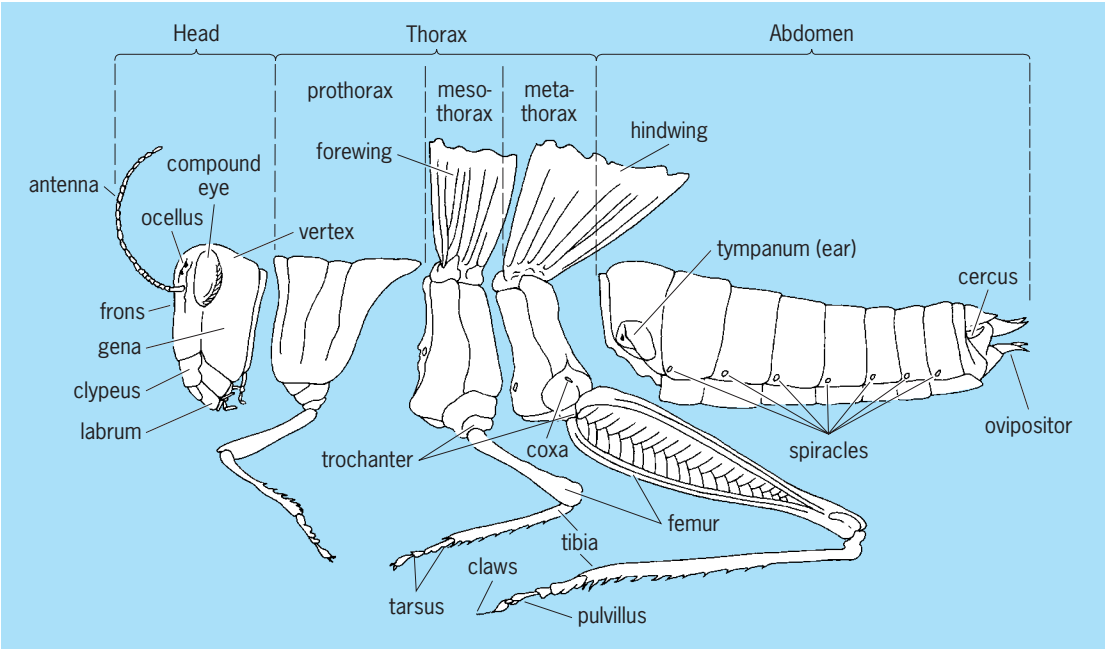

Fig. 1. Grasshopper, in lateral view, showing three main body divisions: head, thorax, and abdomen. (*After C. L. Metcalf, W. P. Flint, and R. L. Metcalf, Destructive and Useful Insects, 3d ed., McGraw-Hill, 1951*)

The body wall of an insect serves not only as a covering, but also as a supporting structure to which many important muscles are attached. The body wall of an insect is composed of three principal layers: the outer cuticula, which contains, among other chemicals, chitin; a cellular layer, the epidermis, which secretes the chitin; and a thin noncellular layer beneath the epidermis, the basement membrane. Chitin is a nitrogenous polysaccharide, $C_{32}H_{54}N_4O_{21}$, and is resistant to most weak acids and alkalies. It is secreted as a liquid and hardens. The surface of an insect's body consists of a number of hardened plates, or sclerites, separated by sutures or membranous areas, which permit bending or movement.

Head. On the head of an adult insect are the eyes, of which there are usually two types; the mouthparts, which are variously modified according to the feeding habits of the individual, and the associated digestive glands; and the antennae, which usually consist of several segments and vary greatly in size and form.

Eyes. A pair of compound eyes usually covers a large part of the head surface. The surface of each compound eye is divided into a number of circular or hexagonal areas called facets, each of which serves as a lens for a single eye unit, called an ommatidium. The compound eye may contain from a few hundred to several thousand facets. An eye of this type is believed to have mosaic vision, forming a great number of tiny similar images. In addition to having compound eyes, most insects also possess two or three simple eyes, the ocelli, usually located on the upper part of the head between the compound eyes; each of these has a single lens. The ocelli apparently do not serve as organs of vision, but rather are able to detect the degree of light intensity and its source. Insect larvae do not possess compound eyes, but usually have several ocelli. Adult insects never possess more than three ocelli. *See* EYE (INVERTEBRATE).

Mouthparts. Insect mouthparts typically consist of labrum, or upper lip; a pair each of mandibles and maxillae; a labium, or lower lip; and a tonguelike structure, the hypopharynx (**Fig. 2**). These structures are variously modified in different insect groups and are often used in classification and identification. The type of mouthparts an insect has determines how it feeds and what type of damage it is capable of causing. A cricket is said to have chewing mouthparts because it has heavily sclerotized mandibles that move sideways and can bite off and chew particles of food. A bedbug has long, slender, needlelike mandibles which can pierce the skin. Also, a pair of maxillae form a tube which is inserted through the puncture and used to suck blood.

Certain glands, such as the salivary glands, may secrete fluids which aid in digestion or which may serve as a defense mechanism against attack by other insects. These glands may also serve as a reservoir for disease organisms, such as malaria parasites. Modified salivary glands in silkworm larvae produce a fluid which, upon hardening, is known as silk.

Antennae. Several forms of antennae are recognized, to which various names are applied; they are used extensively in classification. The antennae are usually located between or below the compound eyes and are often reduced to a very small size. They are sensory in function and act as tactile organs, organs of smell, and in some cases organs of hearing.

Thorax. The thorax, or middle region, of the body bears the legs and the wings and is composed of

three segments: the prothorax, mesothorax, and metathorax. Each thoracic segment usually bears a pair of legs; the wings, when present, are borne by the mesothorax and the metathorax. If only one pair of wings is present, they are usually borne by the mesothorax. The prothorax never bears wings, but in some fossil forms displays winglike structures. In some adult insects there are no wings, and in many immature and a few adult insects there are no legs. The thorax is usually connected to the head by a membranous neck region, the cervix.

Legs. The legs of modern insects typically consist of five general parts: the coxa, the basal segment next to the thorax; the trochanter, a small segment, occasionally two segments; the femur, usually a long segment and often quite large; the tibia, which may also be quite long; and the tarsus, the last portion, a series of small segments. The number of tarsal segments varies from one to five in different insects. The last tarsal segment usually bears a pair of claws and often a padlike structure between the claws. The legs may be variously modified and fitted especially for running, crawling, swimming, jumping, digging, or clinging to mammalian hair. The different segments of the leg may vary in size, shape, and spination. The crickets and long-horned grasshoppers have an eardrum, or tympanum, at the basal end of the front tibiae.

Wings. Insects are the only winged invertebrates, and their dominance as a group is probably due to their wings. Immature insects do not have fully developed wings, except in the mayflies, where wings develop in the subimago stage. The wings are saclike outgrowths of the body wall. In the adult they are solid structures and may be likened to the two sides of a cellophane bag that have been pressed tightly together. The form and rigidity of the wing are due to the stiff chitinous veins which support and strengthen the membranous portion. At the base are small sclerites which serve as muscle attachments and produce consequent wing movement. The wings vary in number, placement, size, shape, texture, and venation, and in the position at which they are held at rest. Adult insects may be wingless or may have one pair of wings on the mesothorax, or, more often two pairs, one each on the mesothorax and on the metathorax. The membranous wings may bear tiny hairs or scales. In some insects the front wings are thickened, leathery, or hardened and brittle. Most insects are able to fold the wings over the abdomen when at rest, but the butterflies, mayflies, dragonflies, and damselflies cannot do this; they hold the wings outstretched or together above the body when at rest. Some insects, such as the male crickets and grasshoppers, are able to produce characteristic sounds by rubbing the two front wings together or by rubbing the front wings with the hindlegs.

There is a common basic pattern of wing venation in insects which is variously modified and in general quite specific for different large groups of insects. Much of insect classification depends upon these variations. A knowledge of fossil insects depends largely upon the wings, because they are among the more readily fossilized parts of the insect body.

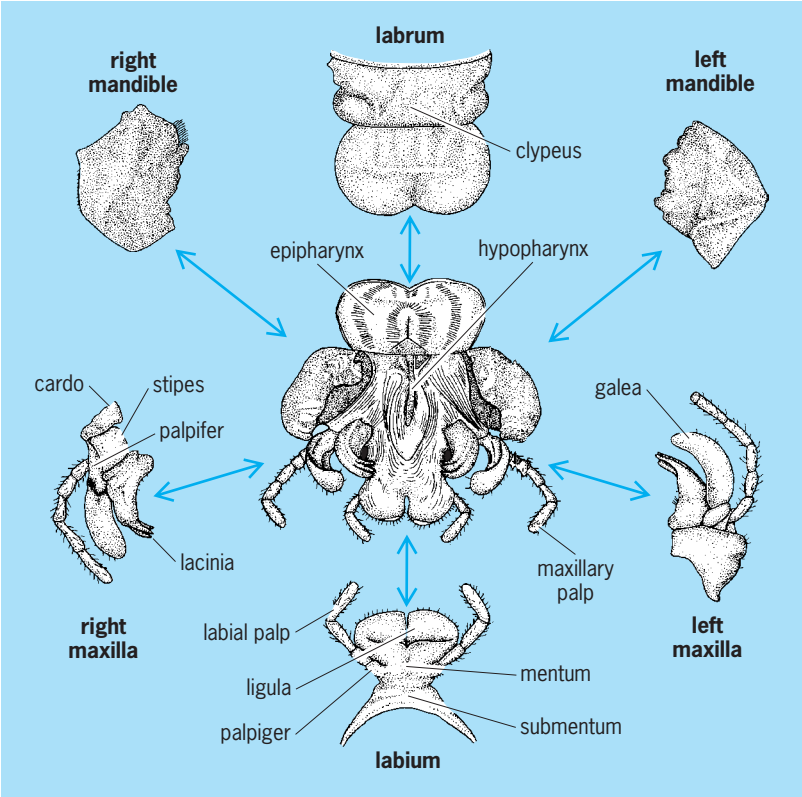

Fig. 2. Grasshopper mandibulate mouthparts. (*After C. L. Metcalf, W. P. Flint, and R. L. Metcalf, Destructive and Useful Insects, 3d ed., McGraw-Hill, 1951*)

Abdomen. The third division of the insect body is the abdomen; it has typically 11 segments. The eleventh segment is usually greatly reduced and is represented only by appendages; therefore, the maximum number of segments rarely appears to be more than 10. In many insects this number is reduced, either by a fusion of segments or by a telescoping of the terminal segments. Segments 1–7 bear appendages in various immature insects and in adult Archaeognatha, Zygentoma, and male Odonata. The abdominal appendages of immature insects may consist of tracheal gills (mayfly nymphs), lateral filaments (Neuroptera larvae), or prolegs (Lepidoptera larvae). These appendages are represented as styli in the Archaeognatha and Zygentoma and as copulatory structures in male Odonata. When 10 complete segments are present, a pair of clasperlike or feelerlike abdominal appendages, the cerci, arise from segment 10; however, they are absent in many species.

The genital structures arise from segments 8 and 9. In the male these function as copulatory structures for the transfer of sperm to the female. These external genital structures are quite variable and complex and frequently serve as excellent taxonomic characters. When not in use, these organs are usually withdrawn into the tip of the abdomen, often called the genital chamber, and frequently cannot be observed unless dissected. In the female these structures form the ovipositor or egg-deposition organ

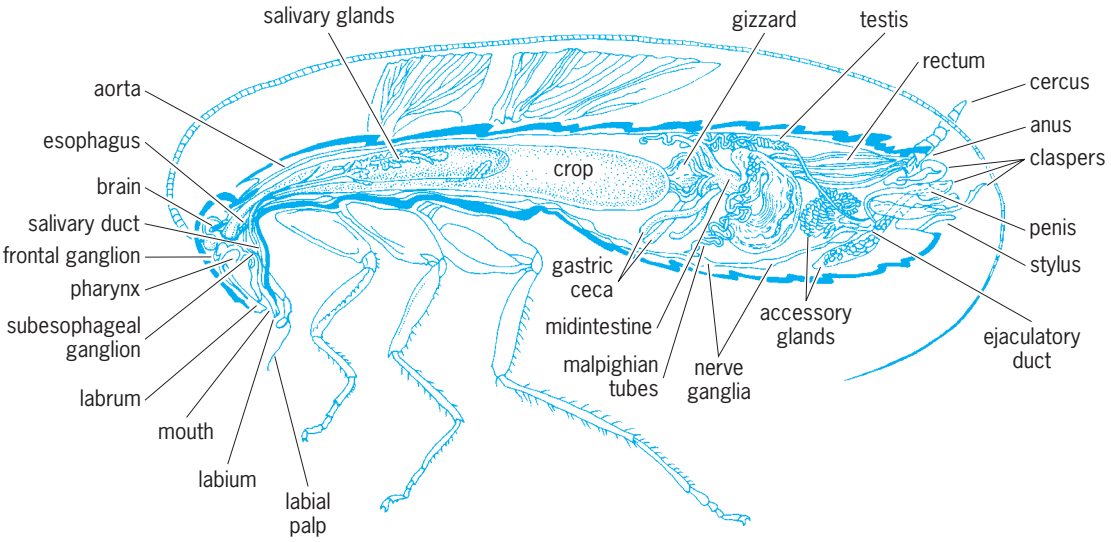

Fig. 3. Sagittal section of a cockroach showing the internal anatomy of a male specimen. (*After C. L. Metcalf, W. P. Flint, and R. L. Metcalf, Destructive and Useful Insects, 3d ed., McGraw-Hill, 1951*)

concerned with oviposition. Ovipositors are often fitted for slitting open plant or animal tissues and depositing eggs in the tissue. They may be modified to form a sting in the bees and wasps. In parasitic Hymenoptera they often are quite long and are used to drill into wood, where the eggs are deposited. In some of the meadow grasshoppers the ovipositor is a bladelike structure as long as the body.

Internal Anatomy

The internal structures of insects (**Fig. 3**) may be considered as six major systems: respiratory, digestive, excretory, circulatory, nervous, and reproductive.

Respiratory system. The intake of oxygen, its distribution to the tissues, and the removal of carbon dioxide are accomplished by means of an intricate system of tubes called the tracheal system. The principal tubes of this system, the tracheae, open externally at the spiracles. Internally they branch extensively, extend to all parts of the body, and terminate in simple cells, the tracheoles. The tracheae are lined with a thin layer of cuticle, which is thickened to form spiral rings that give the tracheae rigidity. The spiracles are located laterally and vary in number from 1 to 10. Usually there is a pair of spiracles on the mesothorax, a pair on the metathorax, and one pair on each of the first seven or eight abdominal segments. The oxygen enters through the spiracles, and carbon dioxide is eliminated principally through them; probably one-fourth of the carbon dioxide is eliminated through the body wall. In some insects the oxygen is taken in the body through the thoracic spiracles and is eliminated through the abdominal spiracles. The spiracles may be partially or completely closed for extended periods by some insects. Many adaptations for carrying on respiration are known. Aquatic insects may have tracheal gills located either on the external portion of the body or in the rectal cavity. The larvae of certain insects, such as mosquitoes, have respiratory tubes at the posterior end of the body. Some aquatic insects with normal tracheal systems and active spiracles can dive beneath the water and remain for extended periods of time by carrying with them a film or bubble of air on the surface of the body.

Digestive system. Insects possess an alimentary tract consisting of a tube, usually coiled, which extends from the mouth to the anus. It is differentiated into three main regions: the foregut, midgut, and hindgut. The foregut is usually composed of four parts: the pharynx, esophagus, crop, and proventriculus. The salivary glands are evaginations of the foregut. The foregut is an invagination of the ectoderm and is lined with cuticle. The midgut is usually undifferentiated, except for evaginations called gastric ceca, and is endodermal in origin and lined with epithelial cells. Some of the epithelial cells produce enzymes, and others absorb digested food. The hindgut may be differentiated into the small intestine, or ileum; the large intestine, or colon; and the rectum. The Malpighian tubules, which are excretory in function, connect with the alimentary canal at the point of union of the hindgut and midgut. Valves between the three main divisions of the alimentary canal regulate the passage of food from one region to another.

The major functions of the digestive system are the ingestion and digestion of food. Digestion is the process of changing food chemically so that it may be taken up by the blood to supply nutrients to various parts of the body. The food habits may vary in a given order. Larvae and adults of holometabolous insects usually have entirely different food habits and different types of digestive systems. Because of the great variety of insect foods, both the specific structures and the functions of the different parts vary greatly. When food is ingested, the crop may serve as a temporary storage chamber, where partial digestion may take place; before reaching the midgut, food passes through the proventriculus, which is often fitted with large or sharp needlelike teeth which may strain out or crush the larger particles of food. Digestion is completed and most of the absorption takes

place in the midgut. After food is digested and for the most part absorbed, the residue passes into the hindgut, where some absorption may take place.

Excretory system. The excretory system consists of a group of tubes with closed distal ends, the Malpighian tubules, which arise as evaginations of the anterior end of the hindgut. They vary in number from 1 to over 100, and extend into the body cavity. Various waste products are taken up from the blood by these tubules and passed out via the hindgut and anus.

Circulatory system. The circulatory system of an insect is an open one, as compared with a closed system composed of arteries, veins, and capillaries, as in the vertebrates. The only blood vessel is a tube located dorsal to the alimentary tract and extending through the thorax and abdomen. The posterior portion of this tube, the heart, is divided into a series of chambers, each of which has a pair of lateral openings called ostia. The anterior part of the tube which lacks ostia is called the dorsal aorta. Pulsations of the heart produce the circulation by pumping the blood forward, and out of the aorta in the neck region. The increased pressure in this region then causes the blood to move posteriorly through the body cavity to the ostia.

Nervous system. The nervous system consists of a brain, often called the supraesophageal ganglion, located in the head above the esophagus; a subesophageal ganglion, connected to the brain by two commissures that extend around each side of the esophagus; and a ventral nerve cord, typically double, extending posteriorly through the thorax and abdomen from the subesophageal ganglion. In the nerve cords there are enlargements, called ganglia. Typically, there is a pair to each body segment, but this number may be reduced by the fusion of ganglia.

From each ganglion of the chain, nerves extend to each adjacent segment of the body, and also extend from the brain to part of the alimentary canal. There are several minor ganglia connected directly with the brain.

Reproductive system. Reproduction in insects is nearly always sexual, and the sexes are separate. Variations from the usual reproductive pattern occur occasionally. In many social insects, such as the ants and bees, certain females, the workers, may be unable to reproduce because their sex organs are undeveloped; in some insects, individuals occasionally occur that have characters of both sexes, called gynandromorphs. Also, parthenogenesis—the process of females giving rise to females—is known in some species.

The female reproductive system consists primarily of a pair of ovaries, each with a number of tubes called ovarioles. The ovarioles from each ovary unite into a common tube, and the two tubes fuse to form the oviduct. This discharges through an external opening, the vulva, located on the ventral side of the eighth segment. There are special glands and other chambers that open into the oviduct. The germ cell forms at the inner end of each ovariole and passes through a series of chambers, receiving its supply of food material as a consequence. In the last chamber it receives its chorion, or shell, which has a minute opening, the micropyle. As the completed egg proceeds into the oviduct, it passes the opening of a chamber, the spermatheca, where sperm previously received from the male is released to enter the egg and cause fertilization. The other glands may, at this point, contribute various substances used for attaching the egg to a substrate or for covering it.

In the male, there are paired testes in which the sperm are formed. From these, a pair of slender ducts, called the vasa deferentia, unite into a common ejaculatory duct, and this extends into a variously sclerotized penis. The sperm is commonly stored in swellings of the vasa deferentia to await copulation. Connected with these structures are various accessory glands.

Among the many groups of insects, various modifications exist in these basic structures and in their methods of operation. In fact, almost every method of reproduction that is known in the animal kingdom can be found among the insects. Copulation normally occurs through the vulva, but in the bedbug family Cimicidae of the order Hemiptera, it takes place by means of a special structure occurring at various points on the female body. The sperm is received in this structure and then passed through the body cavity to fertilize the egg in the ovariole. In some insects the sperm must pass through the oviduct into the ovariole to fertilize the egg before it enters the oviduct. The fertilized immature egg develops to maturity in the ovariole. In some forms the ovarioles are not tubular but are shaped like a bunch of grapes. In a few insects the embryo is retained within the body until it is fully grown, being nourished by glandular secretions. There is even something of a placental development in some forms of insect. *See* INSECT PHYSIOLOGY.

Metamorphosis

After insects hatch from an egg, they begin to increase in size (growth) and will also usually change, to some degree at least, in form and often in appearance. This developmental process is metamorphosis. The growth of an insect is accompanied by a series of molts, or ecdyses, in which the cuticle is shed and renewed. Prior to the actual shedding process, a new cuticular layer is secreted beneath the old layer, and a molting fluid is secreted which separates the new cuticle from the old.

The molt involves not only the external layers of the body wall, the cuticula, but also the cuticular linings of the tracheae, foregut, and hindgut; the cast skins often retain the shape of the insects from which they were shed. The shedding process begins with a splitting of the old cuticle, usually along the midline on the dorsal side of the thorax and head. This split grows and the insect eventually wriggles out of the old cuticle. The new skin, at first wrinkled and folded, remains soft and pliable long enough for the body to expand to its fullest capacity before hardening.

Insects differ regarding the number of molts during their growing period. Many have as few as four molts; a few species have 40 or more, and the latter continue to molt throughout life.

Insects have been grouped or classified upon the basis of the type of metamorphosis which they undergo. All entomologists, however, do not agree upon the same classification. These types of metamorphosis have been placed in three groups, under the headings ametaboly, hemimetaboly, and holometaboly. In order to cover the more detailed pattern, the following outline is used.

1. Ametabolous or primitive: No distinct external changes are evident with an increase in size.

2. Hemimetabolous: Direct metamorphosis that is simple and gradual; immature forms resemble the adults except in size, wings, and genitalia. Immatures are referred to as nymphs or naiads if aquatic.

3. Holometabolous: Complete, or indirect, metamorphosis; stages in this developmental type are: egg→larva→pupa→adult (or imago).

In the ametabolous forms, the insects never obtain wings and have no visible external changes. In the hemimetabolous forms, each insect gradually acquires wings during a period of growth and several moltings. In the holometabolous, or complete, form of metamorphosis, there is a complete change from the larva (which has only simple eyes, no wings, and frequently no legs or antennae) to the adult, usually of entirely different appearance, with wings, compound eyes, well-developed legs and antennae, and in certain groups entirely different mouthparts. Also, in this type of metamorphosis there is an extra life stage, the pupa, usually called a resting stage, in which the complete transformation takes place. The larva is the growth stage, the pupa the transformation stage, and the adult the reproductive stage. In the insects having complete metamorphosis, the feeding habits of the larva and the adult of the same species are usually different. Dwight M. De Long

Fossils

Insects and parainsects have a rich fossil record that extends to 415 million years, representing all taxonomic orders and 70% of all families that occur today. In addition, the insect fossil record documents an assemblage of extinct, higher-level taxa centered principally in the late Paleozoic. At a first approximation, the fossil insect record is divisible into a Paleozoic Insect Fauna of Early Devonian to late Permian age, and a Modern Insect Fauna beginning during the Triassic and continuing to their current unbridled diversity. These two evolutionary faunas are separated by the terminal Permian extinction of 251 million years ago, during which most species on Earth were extinguished, including many insect lineages. It also is likely that Permian taxa were being ecologically replaced by the progenitors of the Modern Insect Fauna, many of which survived the Permian extinction to continue as the familiar insect lineages of today (**Figs. 4** and **5***a*).

Insect deposits are characterized by an abundance of exceptionally well-preserved deposits known as Lagerstätten. Lagerstätten refer not only to the familiar amber deposits that entomb insects in hardened tree resin (**Fig. 6***a*), but more importantly to a broad variety of typically laminar, sedimentary deposits (Fig. 6*b*, *c*) that range from internally drained, fault-bounded montane lakes to lowland basins formed by broad tectonic subsidence. These lacustrine deposits, formed in lake basins, are the most persistent of insect-bearing deposits and document the evolution of insect biotas during the past 300 million years. By contrast, the oldest amber is approximately 120 million years old and extends modern lineages and associated taxa to the Early Cretaceous. Other major types of insect deposits include terrestrial shales and fine-grained sandstones marginal to marine deposits during the Early and Middle Devonian, a proliferation of nodular ironstone-bearing strata of late Carboniferous age from the equatorial lowlands of the paleocontinent Euramerica, and distinctive lithographic limestones worldwide from the Middle Jurassic to Early Cretaceous. More modern deposits are Miocene to Recent sinter deposits created by hydrothermal zones with mineral-rich waters, and similarly aged asphaltum, representing the surface accumulation of tar (Fig. 6*d*). Lastly, insects are abundant in many Pleistocene glacial deposits of outwash and stranded lake sediments, formed by the waxing and waning of alpine and continental glaciers.

Fossil record. Almost all the information of insect fossil history originates from approximately 150 major localities representing all the major types of deposits mentioned above. However, these deposits are distributed very unevenly in time and space. No insects are known from a 55-million-year hiatus between the Middle Devonian and the latest Mississippian. Another gap from the late Permian to Early Triassic features 15 million years of very poor, scrappy occurrences and an absence of any significant faunas. Described fossil insect deposits overwhelmingly occur at north temperate latitudes, principally North America and northern to central Eurasia, especially western Europe. Very few deposits have been found in the tropics, and only a few major deposits have been examined from south temperate latitudes, principally from southern Africa and eastern Australia.

Paleozoic. The earliest known insect is a bristletail (order Archaeognatha) that occurs in approximately 405-million-year-old Early Devonian deposits of Gaspé, Quebec; younger bristletails also are known from the Middle Devonian of New York state. An early Devonian collembolan from the 415-million-year-old Rhynie Chert of Scotland documents the closest relative of the insects, the Parainsecta. The Parainsecta (orders Collembola and Protura), when combined with the Insecta (all orders in Fig. 4), constitute a larger clade, Hexapoda, ancestors of which are thought to have originated from a marine, probably crustacean, arthropod sometime during the Late Silurian. During the Early Devonian, terrestrial ecosystems consisted of algae, mosses, primitive vascular plants, massive fungi, and arthropods that

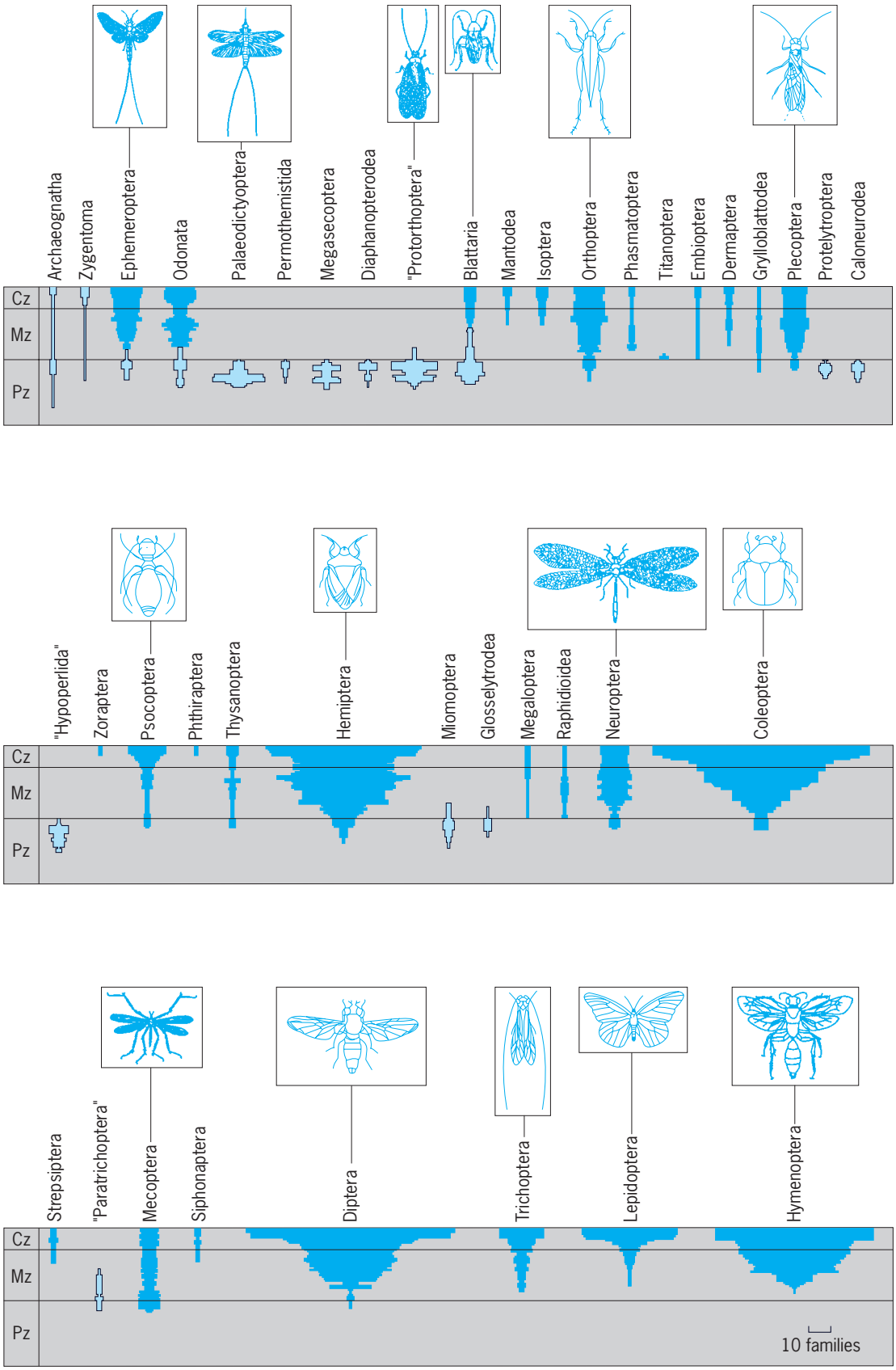

Fig. 4. Spindle diagrams showing diversities of fossil families within insect orders in stratigraphic stages of the Phanerozoic. Note the scale bar in the lower right. Abbreviations are Pz, Paleozoic (Silurian through Permian); Mz, Mesozoic; and Cz, Cenozoic. Boxed illustrations (not to scale) depict typical adult representatives of the more important orders; light-colored spindles or portions are Paleozoic Insect Fauna; dark-colored spindles or portions are Modern Insect Fauna. The appearance of angiosperms is approximately two-thirds of the way up for the Mesozoic band. (*Modified slightly from C. C. Labandeira and J. J. Sepkoski, Jr., Insect diversity in the fossil record, Science, 261:310–315, 1993*)

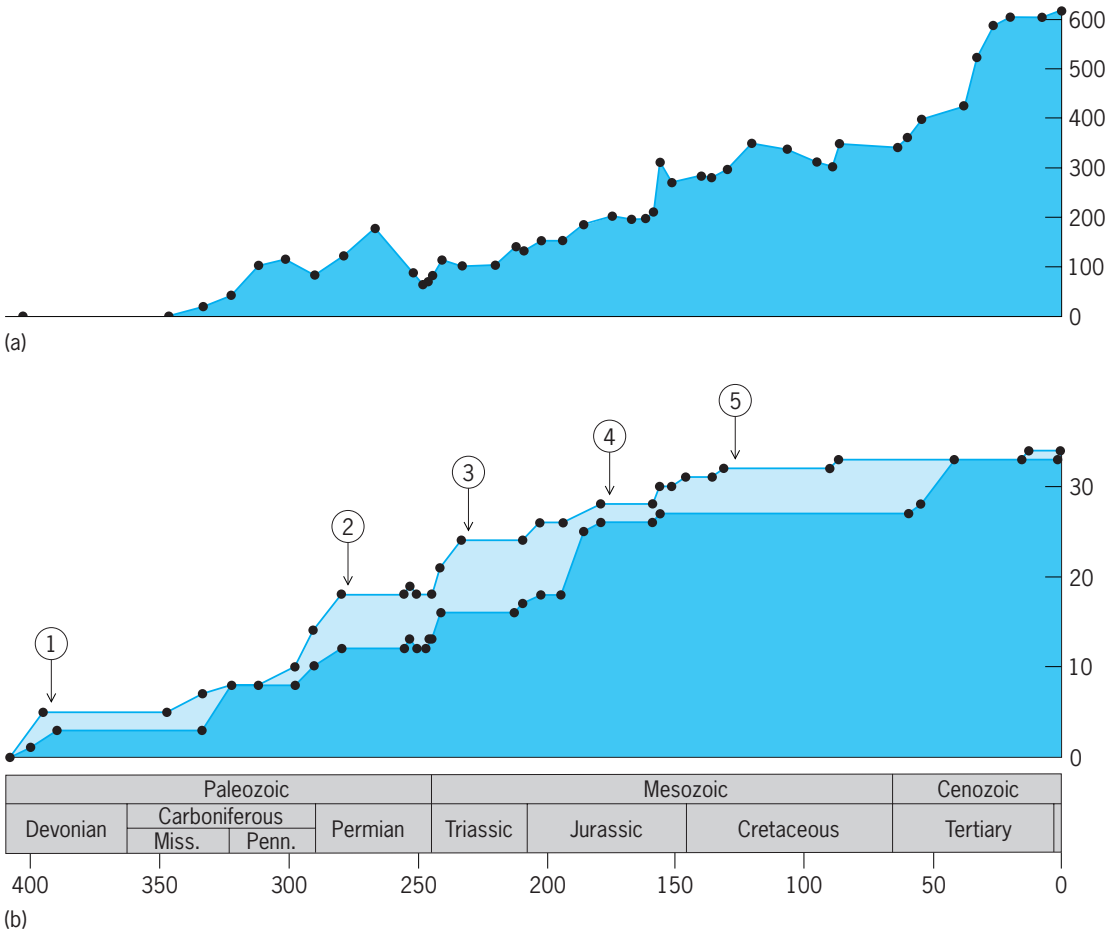

Fig. 5. Comparison of (*a*) family-level taxonomic diversity and (*b*) mouthpart class diversity since the Early Devonian. Data for both graphs are resolved at the geologic stage level, shown in *b* sequentially as data points but included in *a* only to reflect changes in diversity. For mouthpart diversity, in *b*, the darker pattern indicates strong evidence and the lighter pattern indicates less reliable evidence; numbers with arrows designate the five phases of mouthpart class diversification discussed in the text. Abbreviations are Miss., Mississippian; Penn., Pennsylvanian. (*Reproduced with permission from C. C. Labandeira, Insect mouthparts: Ascertaining the paleobiology of insect feeding strategies, Annu. Rev. Ecol. Sys., 28:153–193, 1997*)

included myriapods, arachnids, and primitively nonwinged collembolan and archaeognathan hexapods. Ecologically these early land habitats harbored photosynthesizing plants, detritivores, especially carnivores, but also significant herbivores, judging from all available evidence.

Permian insects are principally known from two deposits: Elmo, Kansas; and Chekarda, from the Ural Mountains of Russia. Many taxa from these localities occur in the same taxonomic families, suggesting that there were paleobiogeographical connections between the western equatorial and eastern north temperate margins of the vast, Northern Hemisphere continent of Laurussia. This Permian fauna represented a blend of scaled-down late Carboniferous lineages that were centered in the equatorial lowlands, and newly emergent, basal lineages of cicada-like hemipteroids and holometabolans occurring in the more arid higher latitudes. [Holometabolous insects (orders Miomoptera to Hymenoptera in Fig. 4) are characterized by immature larval stages separated from the adult stage by a typically resting stage, the pupa.] Hemipteroids (orders Psocoptera to Hemiptera in Fig. 4) and the holometabolans preferentially survived the end-Permian extinctions, forming the core of the Modern Insect Fauna during the Triassic.

After a 55-million-year hiatus, the known insect fossil record does not resume until the boundary between the early Carboniferous (Mississippian) and the late Carboniferous (Pennsylvanian), at 323 million years ago. At that time, numerous winged insect lineages appear in profusion and geochronologically suddenly (Figs. 4 and 5*a*), suggesting that there was either a significant but unrecorded earlier history, or a preceding spurt of rapid evolutionary diversification. By 300 million years, 15 orders and at least 100 families of insects were present, constituting the basis of the Paleozoic Insect Fauna. Most prominent by virtue of their large size (up to 76-cm wingspan) were dragonflies, mayflies, and four extinct palaeodictyopteroid orders with piercing-and-sucking mouthparts, known collectively as the Palaeoptera. With one exception, the Palaeoptera lacked the ability to swivel their laterally outstretched wings behind their abdomen. Other winged insects possess wing-folding mechanisms, including the second most abundant

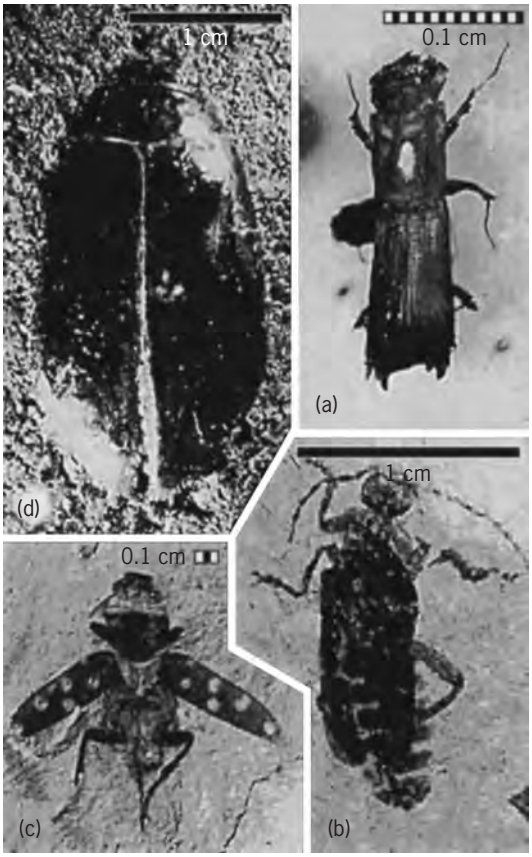

Fig. 6. Examples of the preservational variety in four well-known Lagerstätten deposits, illustrating Cenozoic beetles (Coleoptera). (*a*) Pinhole borer of the family Platypodidae, from early Miocene amber of the Dominican Republic. (*b*) Probable blister beetle of the family Meloidae, from the early Oligocene Antero Formation near Florissant, central Colorado. (*c*) Longhorn beetle of the family Cerambycidae, from the middle Eocene Green River Formation near Rifle, northwestern Colorado. (*d*) Predacious diving beetle of the family Dytiscidae, from the late Pleistocene asphalt deposit at McKittrick, south-central California. (*Paleobiology collections of the National Museum of Natural History, Washington, DC*)

Carboniferous group, the generalized and mandibulate Protorthoptera. The latter are an assemblage of many unrelated lineages that are difficult to define taxonomically, some of which have been segregated to the Hypoperlida and related Caloneurodea that combine a grasshopperlike body facies with mouthpart and head structures that resemble modern booklice (Psocoptera). Cockroaches were the third dominant constituent during the Carboniferous; although they were locally abundant, they had also a greater potential for preservation.

Mesozoic. Most insect clades present during the late Carboniferous did not survive the terminal Permian extinctions; of those that did, many eventually were extinguished during the early Mesozoic. Several clades present during the Permian, however, have survived to the present, either in decreased (Blattoidea, Mecoptera) or modestly expanded (Ephemeroptera, Odonata, Orthoptera, Thysanoptera, Neuroptera) diversity (Fig. 4). Notably, clades that originated during the late Permian to Middle Triassic, namely the Hemiptera and the three preeminent holometabolous orders with important fossil records (Coleoptera, Diptera, and Hymenoptera), overwhelmingly contributed to the Modern Insect Fauna, and their fossil diversity approximately matches proportionally their current global abundance. An exception is the Lepidoptera, whose earliest occurrence is Early Jurassic, and which possess a relatively poor fossil record when compared to their extant diversity.

This rebound is attributable to three major invasions of ecospace. (1) There was expansion during the Late Triassic to Late Jurassic of insect subadult stages (naiads and larvae) into fresh water as filterers, scrapers, shredders, and predators. This invasion provided a more complex trophic structure in aquatic ecosystems, reaching a modern type of organization during the Early Cretaceous. (2) On land, newly emergent lineages of seed plants, particularly advanced seed ferns, bennettitaleans, cycads, conifers, and ephedroids, were colonized by new holometabolous clades of phytophagous insects. These insect taxa consumed leaves, fructifications (fruiting bodies), and pollen, and included feeding types such as leaf miners, gallers, and wood borers that penetrated inner tissues. The establishment of virtually all of the broad categories of herbivory on seed plants during the Jurassic to Early Cretaceous was enlarged by the diversification and ecologic expansion of flowering plants during the mid-Cretaceous, which promoted at lower taxonomic levels further ecologic partitioning of food resources, and propelled new coevolutionary mutualisms. (3) The radiation of small wasp taxa during the Jurassic was associated with the emergence of new feeding guilds in which insects attacked other insects as food resources. This was done either by parasitoids eventually killing their prey, or by parasites feeding on body tissues without inducing death of the host.

Cenozoic. The terminal Cretaceous extinction, which extirpated planktonic foraminifera, ammonites, and dinosaurs, had no detectable effect on family-level insect diversity. However, like flowering plants, there now is evidence for a dieback of insect species, with rebound during the early Cenozoic. For herbivores, many modern-style associations with angiosperms are known from the early Cenozoic and even the Late Cretaceous. Additional evidence of prolonged longevity of lower-level insect taxa, when compared to other common invertebrates, includes the presence of extant genera in late Eocene (about 37 million years ago) Baltic amber, and the identification of species in early Miocene (about 21 million years ago) Dominican amber that are hardly distinguishable from modern species. Examination of fossil beetle species in Pleistocene glacial deposits reveals that there has been virtually no subsequent extinction, and that accommodation to major environmental perturbations has been by geographic dislocation rather than evolutionary adaptation.

Significance. Other types of fossil documentation are important for understanding insect paleobiology,

such as the body-fossil history of mouthparts. A recent study of insect mouthparts reveals a fivefold phase of increasing mouthpart disparity through time (Fig. 5*b*). This geochronologic deployment of the 34 basic types of modern insect mouthparts began during the Early Devonian with a few generalized types (Phase 1), was expanded during the late Carboniferous to early Permian to include major modifications of mandibulate and piercing-and-sucking types (Phase 2), and increased significantly again during the Late Triassic to Early Jurassic (Phase 3) to include filter-feeding mouthpart types and others involved in the ecologic penetration of aquatic ecosystems, and also intricate interactions with other animal and seed-plant hosts. During the Late Jurassic, there was expansion of mouthpart types involved in fluid-feeding on plant, fungal, and animal tissues (Phase 4) and during the Early Cretaceous mouthpart innovation was completed by the addition of a few specialized mouthpart types involved in blood-feeding and other specialized associations (Phase 5). A comparison of taxonomic diversity and mouthpart disparity (Fig. 5) reveals that the generation of taxa has proceeded overall in a semilogarithmic increase reflected in a concave curve, whereas morphologic innovation, as revealed by mouthpart disparity, is a logistic process evidenced by a convex curve. This suggests that the deployment of basic morphologic types typically precedes taxonomic diversification in insect fossil history. Conrad C. Labandeira

Bibliography. S. A. Elias, *Quaternary Insects and Their Environments*, 1992; J. Kukalová-Peck, Fossil history and the evolution of hexapod structures, in I. D. Naumann et al. (eds.), *The Insects of Australia: A Textbook for Students and Research Workers*, 2d ed., vol. 1, pp. 141–179, 1991; C. C. Labandeira, Insects and other hexapods, in R. Singer and M. K. Diamond (eds.), *The Encyclopedia of Paleontology*, 1999; C. C. Labandeira, Plant-insect associations from the fossil record, *Geotimes*, 43:18–24, 1998; G. O. Poinar, Jr., *Life in Amber*, 1992.

Insecticide

A material used to kill insects and related animals by disruption of vital processes through chemical action. Insecticides may be inorganic or organic chemicals. The principal source is from chemical manufacturing, although a few are derived from plants.

Insecticides are classified according to type of action as stomach poisons, contact poisons, residual poisons, systemic poisons, fumigants, repellents, attractants, insect growth regulators, or pheromones. Many act in more than one way. Stomach poisons are applied to plants so that they will be ingested as insects chew the leaves. Contact poisons are applied in a manner to contact insects directly, and are used principally to control species which obtain food by piercing leaf surfaces and withdrawing liquids. Residual insecticides are applied to surfaces so that insects touching them will pick up lethal dosages. Systemic insecticides are applied to plants or animals and are absorbed and translocated to all parts of the organisms, so that insects feeding upon them will obtain lethal doses. Fumigants are applied as gases, or in a form which will vaporize to a gas, so that they can enter the insects' respiratory systems. Repellents prevent insects from closely approaching their hosts. Attractants induce insects to come to specific locations in preference to normal food sources. Insect growth regulators are generally considered to act through disruption of biochemical systems or processes associated with growth or development, such as control of metamorphosis by the juvenile hormones, regulation of molting by the steroid molting hormones, or regulation of enzymes responsible for synthesis or deposition of chitin. Pheromones are chemicals which are emitted by one sex, usually the female, for perception by the other, and function to enhance mate location and identification; pheromones are generally highly species-specific.

Inorganic insecticides. Prior to 1945, large volumes of lead arsenate, calcium arsenate, paris green (copper acetoarsenite), sodium fluoride, and cryolite (sodium fluoaluminate) were used. The potency of arsenicals is a direct function of the percentage of metallic arsenic contained. Lead arsenate was first used in 1892 and proved effective as a stomach poison against many chewing insects. Calcium arsenate was widely used for the control of cotton pests. Paris green was one of the first stomach poisons and had its greatest utility against the Colorado potato beetle. The amount of available water-soluble arsenic governs the utility of arsenates on growing plants, because this fraction will cause foliage burn. Lead arsenate is safest in this respect, calcium arsenate is intermediate, and paris green is the most harmful. Care must be exercised in the application of these materials to food and feed crops because they are poisonous to humans and animals as well as to insects.

Sodium fluoride has been used to control chewing lice on animals and poultry, but its principal application has been for the control of household insects, especially roaches. It cannot be used on plants because of its extreme phytotoxicity. Cryolite has found some utility in the control of the Mexican bean beetle and flea beetles on vegetable crops because of its low water solubility and lack of phytotoxicity.

The use of inorganic insecticides has declined to almost nil. Domestic production of inorganic arsenicals has apparently ceased, with current uses being supplied from existing stocks and by some limited importation. Probably the largest current uses are as constituents of formulations for wood impregnation and preservation.

Synthetic organic insecticides. After 1945, with the introduction of DDT, organic insecticides began to supplant the arsenicals. There are three major classes: the organochlorines, the organophosphorus insecticides, and the carbamates.

Organochlorines. The first of the chlorinated hydrocarbons to be widely used was DDT, followed by lindane. Since 1968, however, in the United States all

nongovernmental and nonprescription uses of DDT have been canceled. Coincidentally, the use of lindane has declined both through the development of resistant strains of insects and through official restriction. Over the years, other chlorinated hydrocarbon insecticides have become available, although a number of these exhibited undesirable characteristics. They include TDE, methoxychlor, Dilan, chlordane, heptachlor, telodrin, aldrin, dieldrin, endrin, toxaphene, endosulfone, Kepone, mirex, chlordecone, and Perthane. Registrations for some have been canceled.

The U. S. Environmental Protection Agency (EPA) monitors the risks to human health and the environment of available insecticides and issues registrations specifying permitted uses. For example, registrations for all uses of TDE were canceled in 1971 because it is a metabolite of DDT. Most uses of aldrin and dieldrin were canceled by the EPA in 1975, and their legal use is limited to control of subterranean termites. Kepone and mirex were prohibited for all uses in 1976. Most uses of chlordane and heptachlor were canceled in 1978. Toxicity reviews for insecticides are scheduled as needed.

Organophosphorus insecticides. The development of this type of insecticide paralleled that of the chlorinated hydrocarbons. Since 1947, many thousands of organophosphorus compounds have been synthesized in academic and industrial laboratories throughout the world for evaluation as potential insecticides. Parathion and methyl parathion are two examples that have achieved significant use.

A great diversity of activity against insects is found among organophosphorus insecticides. It has been estimated that more than 100,000 different organophosphorus insecticides have been synthesized and evaluated. In general, they are very reactive species and thus readily degrade, having low soil persistence. Unfortunately, many can be quite poisonous to humans and other warm-blooded animals, as well as to insects. A few materials possess the desirable attribute of low mammalian toxicity, however. The gamut of available commercial products includes azinphosmethyl, bromophos, carbophenothion, chlorpyrifos, crotoxyphos, diazinon, dichlorvos, EPN, ethion, ethoprop, fensulfothion, fonofos, malathion, methidathion, naled, phosmet, sulfotep, stirofos, sulprofos, temephos, terbufos, and trichlorfon.

Schradan was unique among organic insecticides in that it showed systemic properties when applied to plants. Its activity upon direct contact with insects was relatively low, but following spray application to plants, it was absorbed, transported, and altered chemically by enzymatic processes to one or more products which were toxic to pests, such as aphids, which suck juices from plants. Schradan was thus a selective insecticide because it killed pests which fed on fluids from treated plants but did not affect predators which lived on the surfaces and preyed only upon the pests; for example, aphids are the target pests, and ladybird beetles and their larvae which prey upon the aphids are not affected by the insecticide retained within the plant. Unfortunately, schradan was ultimately found to have environmental properties, the risks of which exceeded benefits to the extent that in 1976 all registrations for use in the United States were canceled.

The following materials which exhibit a varied spectrum in degree and breadth of activity have become available commercially for insect control through systemic action following application to plants: acephate, demeton, dimeton methyl, dicrotophos, dimefox, dimethoate, disulfoton, methamidophos, mevinphos, monocrotophos, oxydemeton-methyl, oxymeton methyl, phorate, phosalone, and phosphamidon.

Coumaphos was one of the first products to be discovered which possessed systemic properties. In addition to coumaphos and dimethoate, chemicals useful as animal systemic insecticides include: chlorofenvinphos, chlorothion, crufomate, dialiflor, diaxothion dicapthon, diphos, famphur, fenthion, fenithrothion, menazon, methamiprophos, and ronnel.

Activity of organic phosphate insecticides results from the inhibition of the enzyme cholinesterase, which performs a vital function in the transmission of impulses in the nervous system. Inhibition of some phenyl esterases occurs also. Inhibition results from direct coupling of phosphate with enzyme. Phosphorothionates are moderately active, but become exceedingly potent upon oxidation to phosphates.

Carbamates. These insecticides are relatives of the alkaloid physostigmine. Individual compounds tend to exhibit selective insecticidal activity. The first commercial carbamate insecticide, carbaryl was introduced during the mid-1950s and continues to enjoy broad use not only in the United States but worldwide. The more important products include aminocarb, bendiocarb, bufencarb, dimetilan, dioxacarb, formetanate, methiocarb, mexacarbate, promecarb, and propoxur.

Considerable success has been achieved in discovering carbamates which exhibit systemic action when applied to plants. The more important products with this property include aldicarb, carbofuran, methomyl, oxamyl, pirimicarb, and thiofanox.

Carbamates, like organophosphates, are cholinergic. Several not only interfere with cholinesterase, but may also inhibit one or more enzymes known as aliesterases. The binding between carbamate inhibitor and enzyme appears to be much more readily reversible than does that between enzyme and organophosphate inhibitor.

Other types of insecticides. During the mid-1960s, much interest developed in the possibilities of utilizing certain of their own secretions and hormones for control of insects. The secretions, named pheromones, are released by one sex, usually the female, and perceived by the other. All pheromones so far discovered are highly species-specific chemicals whose primary function is to facilitate mate finding and therefore maintenance and propagation of the species. Pheromones are volatile materials, and chemically are frequently unsaturated esters of

low-molecular-weight acids in which the molecules tend to lie within a range of 8–20 carbon atoms. The number of unsaturations and the isomerization associated with each are largely responsible for both their specificity and the difficulty of chemical characterization. Attempts to use pheromones to control insects have so far met with limited success. A combination of pheromone and a delivery system has been registered for use against the pink bollworm in cotton. Its performance seems to indicate that the procedure offers considerable promise. *See* CHEMICAL ECOLOGY.

Hormones from insects have been investigated intensively and given careful consideration as potential insect control agents; examples are the molting hormone and the juvenile hormone. The molting hormone is a steroidal molecule which is generally considered too complex and costly for manufacture and use as a pesticide. The juvenile hormone, however, is a more simple aliphatic unsaturated triterpenoid methyl ester. Juvenile hormones II and III have been discovered and identified. The natural hormones have thus far not proved suitable for use in practical pest control. However, some promising synthetic derivatives of the juvenile hormones have been registered, including neotenin, R-20458, hydroprene, and methopreme. *See* ENDOCRINE SYSTEM (INVERTEBRATE).

Diflubenzuron, another chemical with a different mode of action, interferes with the action of chitin synthetase, an enzyme necessary for the synthesis of chitin, and as a consequence of that process, disrupts the orderly deposition of that vital element of the exoskeleton. It is registered for the control of the gypsy moth in hardwood forests of the eastern United States. A conditional registration was granted for the use, beginning during the 1979 season, of diflubenzuron for the control and possibly the eradication of the boll weevil from cotton in the United States.

Pyrethrum is one of the oldest insecticides. Because it is highly active and quite safe for humans and animals, extensive effort was made during the period 1910–1945 to determine its chemical structure. The analytical problem was complicated greatly because a number of active principles were ultimately found to be present. Those identified were designated pyrethrins I and II, cinerins I and II, and jasmolins I and II. Efforts began immediately after identification of these natural products to produce chemicals with modified structures through laboratory synthesis. These efforts culminated in the discovery of some of the most active insect control agents yet devised. The first product, allethrin, which was only partially synthetic, was introduced commercially during the early 1950s. Subsequent, entirely synthetic products have included dimethrin, fenothrin, resmethrin, and tetramethrin.

There are two analogs which are highly active with residues on plant surfaces showing significant degrees of persistence. Their structures differ considerably from those of the natural product. Both fenvalerate and permethrin have been granted conditional registrations in the United States for use in the control of pests of cotton.

Insecticides obtained from plants include nicotine, rotenone, pyrethrins, sabadilla, and ryanodine, some of which are the oldest known insecticides. Nicotine was used as a crude extract of tobacco as early as 1763. The alkaloid is obtained from the leaves and stems of *Nicotiana tabacum* and *N. rustica*. It has been used as a contact insecticide, fumigant, and stomach poison and is especially effective against aphids and other soft-bodied insects.

Rotenone is the most active of six related compounds found in a number of plants, including *Derris elliptica, D. malaccensis, Lonchocarpus utilis*, and *L. urucu*. *Derris* is a native of eastern Asia, and *Lonchocarpus* occurs in South America. The highest concentrations of active principles are found in the roots. Rotenone is active against a number of plant-feeding pests and has found its greatest utility where toxic residues are to be avoided. Rotenone is known also as derris or cubé.

The principal sources of pyrethrum are *Chrysanthemum cinerariaefolium* and *C. coccineum*. Pyrethrins, which are purified extracts prepared from flower petals, contain chemically different active ingredients. The pyrethrins find their greatest use in fly sprays, household insecticides, and grain protectants because they are the safest insecticidal materials available.

Synergists. These materials have little or no insecticidal activity, but increase the activity of chemicals with which they are mixed. Piperonyl butoxide, sulfoxide, MGK, propyl isomer, and sesamex are commercially available. These synergists have their greatest utility in mixtures with the pyrethrins. Some have been shown to enhance the activity of insecticides as well.

Insect resistance. The resistance of insects to DDT was first observed in 1947 in the housefly. By the end of 1967, 91 species of insects had been proved to be resistant to DDT, 135 to cyclodienes, 54 to organophosphates, and 20 to other types of insecticides, including the carbamates. By 1975, 203 species were known to be resistant to DDT, 225 to cyclodienes, 147 to organophosphates, 36 to carbamates, and 35 to other insecticides. Since numerous insects are resistant to more than one type of compound, the total number of species involved is 364 worldwide. Almost every country has reported the presence of resistant strains of the housefly.

During 1957 and 1958, many growers of cotton in the Southern states changed from toxaphene and γ-BHC to organic phosphorus chemicals because of the resistance of the cotton boll weevil to chlorinated hydrocarbon insecticides. By 1967, resistant strains of the cotton bollworm and the tobacco budworm had developed and proliferated to the extent that in numerous areas the use of chlorinated hydrocarbon insecticides was of doubtful value. During the same decade, larvae of three species of corn rootworms also developed resistance to three potent chlorinated insecticides—aldrin, dieldrin, and heptachlor—active against a wide spectrum of

soil-inhabiting insects. The development of resistance to chlorinated hydrocarbon insecticides among several species of disease-transmitting mosquitoes not only continues to pose a threat to world health, but led a World Health Organization committee to report in 1976 that efforts to cope with vector-borne diseases and to eliminate malaria were being hampered most of all by problems arising from resistance to insecticides.

The need to replace the highly active chlorinated materials stimulated the development and use of organophosphorus and carbamate insecticides. Strains of several species developed which were resistant to widely used organophosphates. One of the notable examples was that of the cotton pest, *Heliothis virescens* (the tobacco budworm), and methyl parathion, for which laboratory tests showed a greater than 100-fold difference between susceptible and resistant strains. The production of cotton in parts of northeastern Mexico was practically abandoned and was seriously threatened in the lower Rio Grande Valley of Texas by the mid-1970s due to difficulties and costs of attempting to control this pest.

Certain of the synthetic pyrethroids have been shown under controlled laboratory conditions to induce the selection of resistant strains of mosquitoes. This knowledge has been obtained prior to their registration and release for large-scale commercial development. Perhaps with this forewarning, these products will be employed in the field in ways that will minimize selection for resistance and thus prolong their usefulness for insect control. The synthetic pyrethroids, like the natural pyrethrins, have a very high unit activity, but differ from the briefly lasting natural products in that they persist for moderate periods of time following application in the field. These two attributes—high activity and persistence—are important factors associated with the selection of resistant strains of insects, regardless of the chemical involved.

Formulation and application. Formulation of insecticides is extremely important in obtaining satisfactory control. Common formulations include dusts, water suspensions, emulsions, and solutions. Accessory agents, including dust carriers, solvents, emulsifiers, wetting and dispersing agents, stickers, deodorants or masking agents, synergists, and antioxidants, may be required to obtain a satisfactory product.

Insecticidal dusts and formulated for application as powders. Toxicant concentration is usually quite low. Water suspensions are usually prepared from wettable powders, which are formulated in a manner similar to dusts except that the insecticide is incorporated at a high concentration and wetting and dispersing agents are included. Emulsifiable concentrates are usually prepared by solution of the chemical in a satisfactory solvent to which an emulsifier is added. They are diluted with water prior to application. Granular formulations are an effective means of applying insecticides to the soil to control insects which feed on the subterranean parts of plants.

Proper timing of insecticide applications is important in obtaining satisfactory control. Dusts are more easily and rapidly applied than are sprays. However, results may be more erratic, and much greater attention must be paid to weather conditions than is required for sprays. Coverage of plants and insects is generally less satisfactory with dusts than with sprays. It is best to make dust applications early in the day while the plants are covered with dew, so that greater amounts of dust will adhere. If prevailing winds are too strong, a considerable proportion of dust will be lost. Spray operations will usually require the use of heavier equipment, however. Whatever the technique used, the application of insecticides should be correlated with the occurrence of the most susceptible or accessible stage in the life cycle of the pest involved. By and large, treatments should be made only when economic damage by a pest appears to be imminent.

Attention has focused sharply on the impact of the highly active synthetic insecticides upon the total environment—humans, domestic animals and fowl, soil-inhabiting microflora and microfauna, and all forms of aquatic life. Effects of these materials upon populations of beneficial insects, particularly parasites and predators of the economic species, have been critically assessed.

The study of insect control by biological means has expanded. The concepts and practices of integrated pest control and pest management involve an insect control strategy that employs ecologically based procedures which rely on multiple interventions with minimal disturbances to the ecosystem.

Among problems associated with insect control are the development of strains of insects resistant to insecticides; the assessment of the significance of small, widely distributed insecticide residues in and upon the environment; the development of better and more reliable methods for forecasting insect outbreaks; the evolvement of control programs integrating all methods—physical, physiological, chemical, biological, and cultural—for which practicality was demonstrated; the development of equipment and procedures to detect chemicals much below the part-per-million and microgram levels. As a consequence of the provisions of the Federal Insecticide, Fungicide, and Rodenticide Act as amended by the Federal Environmental Pesticide Control Act of 1972, there have been increased efforts to obtain data delineating mammalian toxicology, persistence in the environment, and immediate chronic impact of pesticides upon nontarget invertebrate and vertebrate organisms occupying aquatic, terrestrial, and arboreal segments of the environment.

The registration of pesticides is a detailed, highly technical process. Pesticides must be selected and applied with care. Recommendations as to the product and method of choice for control of any pest problem—weed, insect, or varmint—are best obtained from county or state agricultural extension specialists. Recommendations for pest control and pesticide use can be obtained from each state agricultural experiment station office. In addition, it is

necessary to read carefully and to follow explicitly the directions, restrictions, and cautions for use which are on the label attached to the product container. Insecticides are a boon to the production of food, feed, and fiber, and their use must not be abused in the home, garden, farm field, forest, or stream. *See* FUMIGANT; INSECT CONTROL, BIOLOGICAL; INSECT PHYSIOLOGY; INSECTA; PESTICIDE.

George F. Ludvik

Bibliography. A. J. Burn et al. (eds.), *Integrated Pest Management*, 1988; H. J. Cottrell (ed.), *Pesticides on Plant Surfaces*, 1987; P. Hurst, N. Dudley, and A. Hay, *The Pesticide Handbook*, 1989; D. H. Hutson and T. R. Roberts (eds.), *Insecticides*, vol. 5, 1986; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., vol. 13, 1994; A. R. Leslie, *Integrated Pest Management*, 1994; R. L. Metcalf and R. A. Metcalf, *Destructive and Useful Insects: Their Habits and Control*, 5th ed., 1992; F. H. Nijhout, *Insect Hormones*, 1994, reprint 1998; H. H. Shorey and J. T. McKelvey (eds.), *Chemical Control of Insect Behavior*, 1977; C. Worthing and R. Hance, *Pesticide Manual*, 1991.

Insectivora

An order of placental mammals including shrews, moles, and hedgehogs. The tree shrews (Tupaiidae) and elephant shrews (Macroscelididae) are now recognized as unrelated, and they are placed in separate orders (Scandentia and Macroscelidea). Formerly thought to be the basal placental order, from which other orders were derived, the Insectivora is now restricted to members of the former suborder Lipotyphla. It evolved side by side with the other placental orders, with a fossil record going back to the Paleocene. A number of fossil families from the Cretaceous and early Tertiary, formerly included in the Insectivora, are classified as Proteutheria. *See* MACROSCELIDEA.

Description. Living lipotyphlous insectivores are small animals: the largest (*Potamogale*) weighs about 1 kg (2 lb). Most eat insects, worms, and other invertebrates, for which they search in ground litter and vegetation, using their highly developed olfactory sense and their mobile, sensitive snouts. Some burrow, such as moles; some are aquatic, such as the desman. Anatomically, they are distinguished by the absense of a cecum on the intestine, reduction of the pubic symphysis (fibrocartilaginous union of bones), and characters of the skull. The cheek teeth typically have sharp cusps and crests, and the incisors are often enlarged to act as forceps. Insectivores are found on all continents except Australia and Antarctica, but only one genus (*Cryptotis*, a shrew) has reached South America.

Classification. Three suborders can be distinguished: Erinaceomorpha, Soricomorpha, and Chrysochloromorpha. Living erinaceomorphs belong to the family Erinaceidae, comprising the spiny hedgehogs (Erinaceinae) of Eurasia and Africa and the hairy moonrats (Echinosoricinae) of Southeast Asia. Probably originating in Asia, both subfamilies appeared in Europe in the Oligocene and in North America and Africa in the Miocene. *Proterix* (late Oligocene) and *Brachyerix* (late Miocene–Pliocene) are aberrant North American erinaceids which represent other invasions. Erinaceids died out in North America in the Pliocene. A giant echinosoricine, *Deinogalerix* (**Fig. 1**), with a skull 21 cm (8 in.) long, evolved in the late Miocene in isolation on a Mediterranean island, now part of Italy. A variety of erinaceomorphs existed in the early Tertiary of North America, Europe, and Asia, but most are known only by teeth and jaws. Exceptionally well preserved material of *Macrocranion* from the middle Eocene of Messel, Germany, shows the presence of spines on the back.

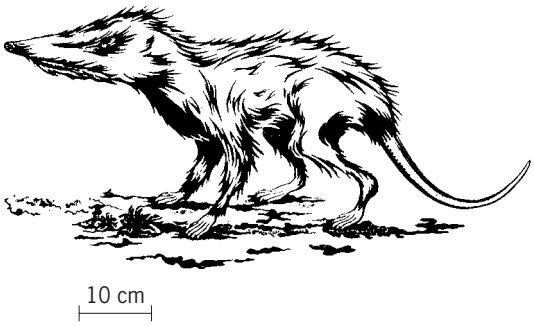

Fig. 1. Reconstruction of the giant erinaceid *Deinogalerix*, from the Miocene of Italy. (*From M. Freudenthal, Scripta Geologica, 14:1–19, 1972*)

Four living families are included in the Soricomorpha: Soricidae (shrews), Talpidae (moles, desman), Tenrecidae (Madagascan tenrecs and African otter shrews), and Solenodontidae (*Solenodon*, confined to Cuba and Hispaniola). The phyletic relationships between these families remain to be elucidated. Dentitions of primitive Paleocene soricomorphs from North America and Europe resemble those of primitive erinaceomorphs. Soricids and talpids were of northern, probably Asiatic origin, and they were distinct before the end of the Eocene. Most early soricids belong to the subfamily Heterosoricinae, which became extinct in the Miocene. Shrews did not enter Africa until the middle Miocene. Among the talpids, the subfamily Talpinae is the most advanced in modification of the forelimbs for burrowing, the Desmaninae have aquatic adaptations, and the Uropsilinae (represented by the extant Chinese shrew mole) retained shrewlike limbs. True moles did not reach North America until the Miocene; they were preceded in the Oligocene by the Proscalopidae, which burrowed in a different way, with the hands moving longitudinally beneath the body, instead of laterally as in true moles. The Dimylidae, from the Oligocene and Miocene of Europe, are poorly known forms with crushing dentition; they may be talpid relatives.

The living *Solenodon*, in the West Indies, is one of the largest insectivores, but nevertheless shrewlike. In the Pleistocene it was accompanied by *Nesophontes* (Nesophontidae), which may be related.

They are believed to have become isolated in the West Indies since the early Tertiary. Attempts to relate them to mainland North American Eocene and Oligocene soricomorphs such as *Apternodus* and *Centetodon* have been unsuccessful. *See* SOLENODON.

The Tenrecidae are far removed geographically from other soricomorphs. They probably evolved in Africa, where their fossil history goes back to the early Miocene, and the otter shrews (Potamogalinae) survive on the continent of Africa today. The remainder are in Madagascar where, like the lemurs, they have evolved in diverse directions during a long time of isolation. In some ways, for example brain size, they have remained more primitive than other insectivores.

The golden moles (Chrysochloridae) are put into a separate suborder, Chrysochloromorpha. They are highly specialized burrowers, using large claws on the forefoot. They are confined to Africa, where fossils show that they were already specialized in the early Miocene.

Proteutherians. A number of more or less primitive placentals, from the Cretaceous and early Tertiary, that could not be incorporated in the better-defined orders were formerly classified in the Insectivora. With restriction of the Insectivora to the Lipotyphla, the order Proteutheria was proposed to include these fossils. This is an artificial and temporary expedient; as the phyletic relations of early placentals become better known, the proteutherians are being separated into more clearly defined orders.

The placentals had a long history in Asia prior to the extinction of the dinosaurs, going back into the Early Cretaceous of Mongolia (*Prokennalestes*). The best-preserved specimens are from the Late Cretaceous of the Gobi Desert. These include the primitive *Kennalestes* (**Fig. 2**) and *Asioryctes*, and the Zalambdalestidae, which were more specialized. The latter have been grouped on dental grounds with some lesser-known genera from central Asia in a Cretaceous suborder, Mixotheridia. Other central Asiatic fossils have been interpreted as early condylarths.

The Palaeoryctidae appeared in North America near the end of the Cretaceous and survived to the Eocene; they have been thought to be ancestral to creodonts, and some similarities to soricomorph insectivores have been noted. The Leptictidae (Late Cretaceous–Oligocene, mainly in North America) were insectivorous, ground-living animals, which have been claimed as distant relatives of lipotyphlan insectivores. The Pseudorhynchocyonidae, from the European Eocene, had long hindlegs, like kangaroo rats; they are possibly related to leptictids. The Apatemyidae (Paleocene–Oligocene) had enlarged incisors, probably used to obtain wood-boring insects, as in the living aye-aye; they are often put in a separate order, Apatotheria. The Pantolestidae (Paleocene–Oligocene), aquatic forms that reached a relatively large size and may have fed on mollusks, have also been separated as the order Pantolesta. The Paleocene Mixodectidae have been linked with the Dermoptera. *See* MAMMALIA; MOLE (ZOOLOGY); SHREW; TENREC. Percy M. Butler; Malcolm C. McKenna

Bibliography. P. M. Butler, Phylogeny of the insectivores, in M. J. Benton (ed.), *The Phylogeny and Classification of the Tetrapods*, vol. 2, pp. 117–141, Clarendon Press, Oxford, 1988; G. C. Gould, Hedgehog phylogeny (Mammalia, Erinaceidae)—the reciprocal illumination of the quick and the dead, *Amer. Mus. Novit.*, 3131:1–45, 1995; S. B. McDowell, The Greater Antillean insectivores, *Bull. Amer. Mus. Nat. Hist.*, 115:113–214, 1958; R. D. E. MacPhee and M. J. Novacek, Definition and relationships of Lipotyphla, in F. S. Szalay, M. J. Novacek, and M. C. McKenna (eds.), *Mammal Phylogeny*, Part 2: *Placentals*, pp. 13–31, Springer-Verlag, New York, 1992; R. D. E. MacPhee, M. J. Novacek, and G. Storch, Basicranial morphology of Early Tertiary erinaceomorphs, and the origin of primates, *Amer. Mus. Novit.*, 2921:1–42, 1988; R. M. Nowack, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999; C. A. Reed, Locomotion and appendicular anatomy in three soricoid insectivores, *Amer. Midland Natural.*, 45:513–671, 1951.

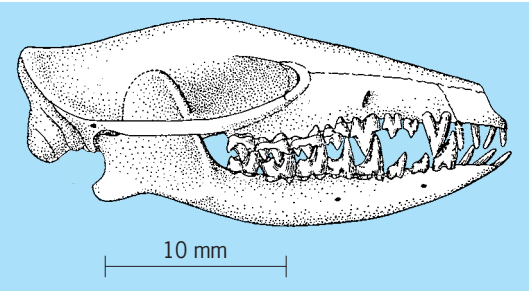

Fig. 2. Skull of *Kennalestes*, a proteutherian from the Cretaceous of Mongolia.

Insectivorous plants

Plants having variously modified, highly specialized leaves which capture and digest insects. The proteins of the digested insect bodies supply nitrogen, which otherwise may be unavailable to these plants in the places where they grow. Sometimes they are also called carnivorous (flesh-eating) plants. *See* NEPENTHALES; PITCHER PLANT; SECRETORY STRUCTURES (PLANT); SUNDEW; VENUS' FLYTRAP. W. W. Watson; Perry D. Strausbaugh

Insolation

The incident radiant energy emitted by the Sun, which reaches a unit horizontal area of the Earth's surface. The term is a contraction of incoming solar radiation. About 99.9% of the Sun's energy is in the spectral range of 0.15–4.0 micrometers. About 95% of this energy is in the range of 0.3–2.4 μm; 1.2% is below 0.3 μm, and 3.6% is above 2.4 μm. The bulk of the insolation (99%) is in the spectral region of 0.25–4.0 μm; about 40% is found in the visible region of 0.4–0.7 μm and only 10% is in wavelengths shorter

than the visible. Energy of wavelengths shorter than 0.29 μm is absorbed high in the atmosphere by nitrogen, oxygen, and ozone.

Insolation depends on several factors: (1) the solar constant—that is, the amount of energy that in a unit time reaches a unit plane surface perpendicular to the Sun's rays outside the Earth's atmosphere, when the Earth is at its mean distance from the Sun; (2) the Sun's elevation in the sky; (3) the amount of solar radiation returned to space at the Earth–atmosphere boundary; and (4) the amount of solar radiation absorbed by the atmosphere and the amount of solar radiation reflected at the lower boundary of the Earth. Insolation is commonly expressed in units of watts per square meter, or calories per square centimeter per minute, also known as langley/min. For instance, the mean value of the solar constant has been estimated as 1368 W/m^2 ($\sim$1.96 ly/min), and the average insolation in summer for a midlatitude clear region could be 340 W/m^2 (700 ly/day), while for a cloudy region it is only about 120 W/m^2 (250 ly/day).

In the past, the solar constant was measured from balloons, rockets, aircraft, or high-altitude observations. Continuous extraterrestrial measurements of the solar constant are mainly from the instruments on the *Nimbus 6* and 7 satellites launched in mid-1975 and late 1978, respectively. These measurements gave the first clear evidence of short-term solar irradiance changes of the order of a few tenths of a percent. Longer-term variability is difficult to establish, since the longest continuous record from *Nimbus* 7 is about 10 years.

Since 1980, long-term solar monitoring has been done with the Active Cavity Radiometer Irradiance Monitor (ACRIM) on board the *Solar Maximum Mission* satellite. It is believed that the most accurate solar constant monitoring is from the self-calibrating ACRIM instrument. The ACRIM data have shown that the Sun's luminosity is negatively correlated with sunspots, which have an 11-year cycle. This is in contradiction to what was believed previously. Since the solar irradiance supplies the energy to drive the atmospheric and oceanic motions, it regulates the Earth's mean temperature and climate. A 1% change in the solar constant was found by some modelers to lead to a different climate. *See* CLIMATE HISTORY; METEOROLOGICAL SATELLITES; SOLAR CONSTANT.

The Sun's elevation, the complement of the zenith distance, depends on latitude, solar declination, and local hour angle. Therefore, the irradiance at the top of the atmosphere will have for each time of the year a certain diurnal variation and a meridional distribution which, in the annual mean, decreases toward the poles from a maximum at the Equator.

The Earth–atmosphere system reflects to space a certain fraction of the incident solar radiation, referred to as the planetary albedo. Even clear sky reflects to space by Rayleigh scattering from molecules, by reflection from the surface, and by scattering from aerosols (Mie scattering). The scattered radiation is called diffuse or sky radiation, and the radiation arriving directly from the Sun is called extraterrestrial direct or beam radiation. However, clouds affect the reflected portion of the insolation more than any other atmospheric factor. The cloud albedo depends on the drop size, liquid water content, water vapor content and thickness of the cloud, and the Sun's elevation. The smaller the drops and the greater the liquid water content, the greater the cloud albedo. The albedo of clouds varies from less than 10% to above 90%. From satellite observations it was determined that, when averaged over the globe, about 32% of the insolation is returned to space at the top of the atmosphere. *See* ALBEDO.

Approximately 20% of the solar radiation is absorbed in the atmosphere by water vapor, carbon dioxide, ozone, oxygen, aerosols, and clouds. The spectral variability of the solar irradiance will affect the amount and altitude of solar radiation absorbed by the Earth's atmosphere and can impact upper-atmospheric dynamical and photochemical processes. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; UPPER-ATMOSPHERE DYNAMICS.

The fraction of insolation that is reflected upward at the ground (surface albedo) depends strongly on the solar zenith angle, the spectral distribution of the radiation reaching the ground, and the type of the underlying surface. For instance, the near-infrared albedo may be three times larger than the visible albedo, especially for vegetation. The standard technique is to measure the total albedo integrated over wavelength. Typical surface albedos range from 5% for oceans to 12% for forests and 30% for deserts; fresh snow can reflect up to 90% of the incident energy.

About 2000 stations around the world routinely measure insolation, and about 1000 of them report to the Radiation Data Center in Leningrad. In the United States an existing network of solar radiation–monitoring stations was upgraded in 1977. It is operated by the National Oceanic and Atmospheric Administration (NOAA) and consists of 38 stations. The measuring instrument is known as a pyranometer. To supplement ground observations of insolation, models have been developed to estimate it from ground-based or satellite-cloud observations. It is interesting to note that the gradient of absorbed radiation, which is the driving force of the circulation, is similar to the gradient of the extraterrestrial insolation. *See* ATMOSPHERE; SOLAR RADIATION; TERRESTRIAL RADIATION.

Rachel T. Pinker

Bibliography. M. Igbal, *An Introduction to Solar Radiation*, 1983.

Instanton

A solution to the (imaginary-time) nonlinear field equation which arises in Yang-Mills field theory—a modern nonlinear generalization of Maxwell electromagnetic theory that is believed to give a fundamental description of elementary particles and forces. The instanton solution carries information about quantum tunneling. (Initially, the instanton was called a pseudoparticle.) *See* GAUGE THEORY.

A dynamical model when analyzed within quantum mechanics gives rise to processes different from those seen in an analysis based on classical mechanics. Of course, since nature is quantum mechanical, only the former processes are truly physical, but the latter can be used to give an approximate, semiclassical description. Of specific interest here is the tunneling phenomenon whereby a quantum-mechanical system evolves along paths that are classically forbidden, in that the classical equation of motion does not possess solutions which follow the forbidden paths. Nevertheless, there does exist in the classical theory a spur of quantum-mechanically allowed transitions: if, in the classical equation of motion, time t is replaced by imaginary time ($t \to \sqrt{-1}t$), solutions which follow tunneling paths can be found. In other words, processes classically forbidden in real time become allowed in imaginary time. Moreover, the quantum-mechanical tunneling amplitude Γ is approximately given by the equation below, where

$$\Gamma \approx \exp\left(-\frac{1}{\hbar}I\right)$$

I is the (imaginary-time) action for the (imaginary-time) tunneling solution, and $\hbar$ is Planck's constant. Obviously only solutions of noninfinite action are relevant, since Γ vanishes if I diverges. Also, I must be positive, so that Γ disappears in the classical limit, $\hbar \to 0$. *See* ACTION; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

While the above ideas, as they relate to point particles, were widely appreciated from the early days of quantum mechanics under the names semiclassical and WKB theory, their relevance to fields, such as the Maxwell electromagnetic field and its generalization the Yang-Mills field, was not realized until the mid-1970s, when an imaginary-time, finite-action solution to the Yang-Mills nonlinear wave equation was found. First called a pseudoparticle, then renamed an instanton, this solution was quickly interpreted as evidence for quantum tunneling in the Yang-Mills theory. This tunneling in turn served to explain the mass difference between the pi and eta mesons, and predicted that nucleons (protons and neutrons) are not stable, although their decay rate is exponentially small, hence negligible. [The instanton (tunneling) mediated proton decay is independent from that predicted by theories that try to unify all fundamental interactions.] *See* FUNDAMENTAL INTERACTIONS; GRAND UNIFICATION THEORIES; QUANTUM FIELD THEORY; WENTZEL-KRAMERS-BRILLOUIN METHOD.

The instanton solution is also of interest to mathematicians, who independently had come to a study of Yang-Mills fields from a purely mathematical context called fiber-bundle theory, which combines differential geometry and topology. The original solution was generalized, and it was found that the (imaginary-time) Yang-Mills action of an arbitrary solution is an integral multiple of the action for the original solution. The integer is called instanton number by physicists and Chern-Pontryagin number by mathematicians. The reason for this quantization with classical field theory derives from the topological properties of Yang-Mills fields. The most general N-instanton solution has been found. *See* FIBER BUNDLE.

In addition to calling attention to unexpected physical processes, namely tunneling, the instanton solution has put into evidence the rich topological structure that is present in Yang-Mills theory. Contemporary research is concerned with understanding this structure, an activity in which both physicists and mathematicians are engaged in a rare collaborative effort. *See* ELEMENTARY PARTICLE; TOPOLOGY.

Roman Jackiw

Bibliography. D. S. Freed and K. K. Uhlenbeck, *Instantons and Four-Manifolds*, 2d ed., 1991; R. Rajamaran, *Solitons and Instantons*, 2d ed., 1989, paper 1987; M. A. Shifman, *Instantons in Gauge Theories*, 1994.

Instinctive behavior

A relatively complex response pattern which is usually present in one or both sexes of a given species. These responses have a genetic basis, are essentially unlearned, and are generally adaptive.

The qualifying terms are an important part of the definition, because each of the defining characteristics of instinctive behavior can be shown in certain circumstances and under certain conditions to be inappropriate. Because of this, some theorists have suggested that the terms instinct and instinctive behavior be stricken from the lexicon. However, instinctive behavior is still a useful term to describe intricate actions such as the complex behavior of the courting bower bird and the efficient and precise killing behavior of the predatory cat.

Instinctive behavior occurs when an animal has a particular internal state while it is in the presence of a specific external stimulation called a releaser or a sign stimulus. Neither the internal state nor the external stimulus alone is adequate for the elicitation of the response. Many animals show particular instinctive behaviors only during the mating season, when hormonal changes associated with sexual behavior sensitize specific portions of the central nervous system, which will then be active in the presence of the releaser. The external stimulus may be relatively simple or incredibly complex. In forms lower on the phylogenetic scale, the releasing stimulus may be deceptively simple. When the small fresh-water fish, the three-spined stickleback (*Gasterosteus*), is hormonally prepared for breeding, it attacks any other male stickleback that comes near its nesting area. The fish's attack is stimulated by the opponent's red belly. The fact that there is another male attached to that red stimulus is really irrelevant. Experimental work has shown that the nesting male will attack a very crude cardboard model of a fish as long as it has a belly area that is painted red. During the spring and summer, the European robin (*Erithacus rubecula*) threatens or attacks another that approaches the breeding area. Once again, the releasing or eliciting

stimulus is red, a tuft of red feathers. The robin attacks a tuft of feathers though it stands alone on a post with no bird near. Some sign stimuli (releasers) can be designed so that they elicit an exaggerated response which may actually be maladaptive. The eggs of the oystercatcher (*Haematopus ostralegus*) are spotted and of modest size, befitting a modest-sized bird. These eggs elicit particular behaviors from the bird, including sitting on them and nestling over them. However, if the oystercatcher is presented with a many-spotted egg several times larger than its own eggs, the bird shows a preference for the larger egg and repeatedly attempts nesting behavior on it, while its own eggs are neglected. *See* REPRODUCTIVE BEHAVIOR.

The innate neural mechanisms which permit the animal to recognize stimuli as complex as a spotted egg, or even the redness of a belly, cardboard or real, must be complex indeed. The linking of the neural mechanisms for perception with those that control the rather complicated responses are, of course, far more complex. A number of instinctive behaviors have been subjected to physiological analysis. Predation in the cat may serve as an example. One obvious releasing stimulus for predatory behavior in the cat is a rat. If the neural system for predation is in an active state, the prey will elicit the typical attack response. It should be noted, however, that the apparently simple stimulus "rat" is, in fact, remarkably complex. It does not matter whether the rat is lying down, curled up in a ball, walking toward or away from the cat, or standing on its hindlegs and washing its whiskers. It also makes no difference if the rat is white, brown, or spotted black and brown; the cat reacts to the vast complexity of those stimuli with a predatory response. Further, the cat may respond to those complex stimulus patterns even though it has never seen or been near a rat during its entire life.

In the wild cat a combination of internal neural sensitivity and relevant external stimuli combine to produce neural action which results in the killing response. However, many pet house cats show no inclination to attack small animals and may live with a rat for a prolonged period. It would appear that prey catching is not in their behavioral repertoire. Yet it is possible to produce relatively precise prey catching by direct electrical stimulation of a number of points in that portion of the brain called the limbic system. The chase and kill occur to direct brain stimulation even though the experimental animal has been completely isolated from the time of its birth.

The prey-catching and killing response shows considerable consistency from one cat to another. The prey is chased until caught, usually with the paws, and is held down while the killing bite is made to the back of the neck. This relative behavioral constancy has led some investigators to refer to instinctive behavior as fixed action patterns. Such terms cannot be taken literally, however. It is true that the end result is the killing bite, but the actual behaviors shown between the sighting of the prey and the final kill may vary widely and are a function of the behavior of the prey. If the prey jumps up on something, the cat follows. The cat's muscle action pattern is very different if it chases the prey to the right or to the left. If the prey runs under something, the cat will show flexible behavior by reaching under the obstacle and attempting to strike the animal with its paw. Certainly these behavior patterns are fixed only in the broadest sense.

Within limits, the instinctive behaviors can be modified by learning. There is evidence, for example, that some predators learn to attack their prey at the back of the neck because when held in that position the prey cannot counterattack. *See* MIGRATORY BEHAVIOR.

Kenneth E. Moyer

Bibliography. J. Alcock, *Animal Behavior: An Evolutionary Approach*, 6th rev. ed., 1997; N. R. Carlson, *Physiology of Behavior*, 7th ed., 2000; R. A. Hinde and J. S. Hinde, *Instinct and Intelligence*, 3d ed., 1987.

Instrument landing system (ILS)

A collection of discrete radio navigation aids used by pilots of all types of aircraft for approach guidance to a specific airport runway, especially during times of limited visibility. Typically, three components constitute the instrument landing system: a localizer, which provides lateral guidance; a glide slope, which gives vertical guidance; and one to three marker beacons, which give position fixes along the approach path (**Fig. 1**). A distance-measuring-equipment (DME) readout or a nondirectional, low-frequency radio beacon (NDB) is sometimes substituted for the outer marker beacon, usually located about 5 mi (8 km) from the runway. Details essential for use of the instrument landing system by the pilot are peculiar to the airport and runway. *See* DISTANCE-MEASURING EQUIPMENT.

The localizer, glide slope, and markers function by radiating continuous-wave, horizontally polarized, fixed beams of radio-frequency energy into the air space. The localizer operates in the band of 108–112 MHz, the glide slope in the band of 329–335 MHz, and the marker beacons on the discrete frequency of 75 MHz. The localizer and glide-slope frequencies are paired; for example, every localizer operating at 109.5 MHz has a glide slope operating at 329.6 MHz. If there is a DME present, it is also paired. Only the localizer receiver in the aircraft needs to be tuned.

Use. Over 1000 instrument landing systems operate in the United States, with an additional 300 systems in use worldwide. There are approximately 117,000 aircraft with ILS receivers, many with two for redundancy purposes. Engineering advances have made it possible to install the instrument landing system at all sites where the physical obstructions allow an instrument procedure to be developed. A major contender for ultimate replacement of the instrument landing system appears to be a Global Navigation Satellite System (GNSS), in which the Global Positioning System (GPS) would play a major role. (The microwave landing system may be used for

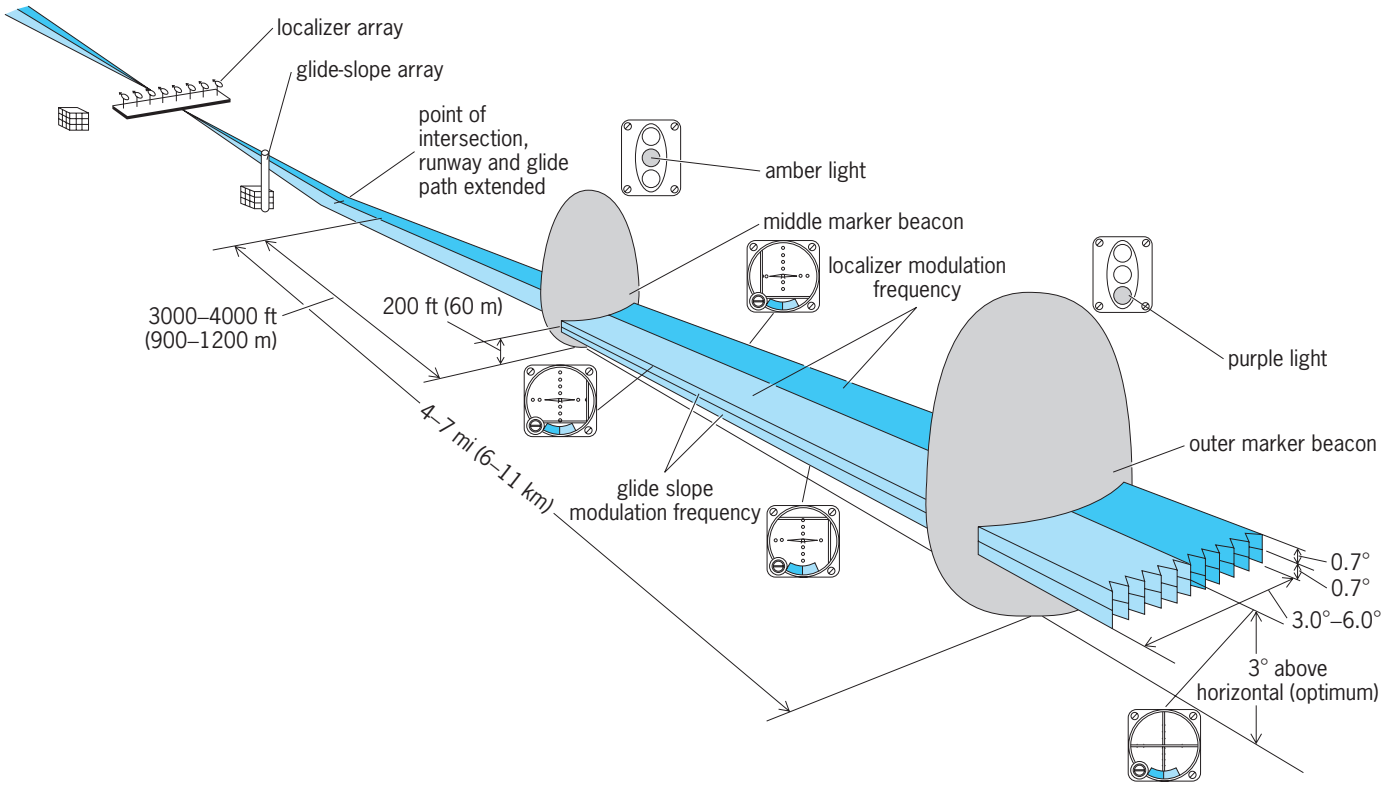

Fig. 1. Components of instrument landing system (ILS), and resulting signals used by the pilot. (*Federal Aviation Administration*)

special-purpose landing needs, but its widespread deployment is no longer planned.) *See* MICROWAVE LANDING SYSTEM (MLS); SATELLITE NAVIGATION SYSTEMS.

Standards. The International Civil Aviation Organization (ICAO) publishes the standards and recommended practices for the instrument landing system. The U.S. Federal Aviation Administration (FAA) publishes numerous manuals for the installation, maintenance, operation, monitoring, and calibration of the instrument landing system. A need for greater dependability of flight schedules has motivated the government to achieve even higher quality of ILS performance. Higher-quality ILS courses permit automatic flight controls to be used, especially when visibilities from the cockpit are measured to be only in hundreds of feet horizontally and tens of feet vertically. Another requirement for operation with lower visibilities is high system reliability. This reliability has been obtained largely through the use of solid-state transmitting equipment. All instrument landing systems are electronically monitored, and should out-of-tolerance conditions exist the transmitter is shut off automatically. *See* AUTOMATIC LANDING SYSTEM.

Spectrum availability. A limitation with ILS deployment is the availability of spectrum. Frequency assignments in some locations are difficult to make, and signals from the commercial frequency-modulation (FM) band, immediately below the ILS localizer band, have necessitated abandonment of some frequency channels. Relief has been obtained by making use of the directivity of the antenna arrays, and increasing the technical specifications for airborne receivers in order to make them more capable of rejecting undesirable signals. *See* RADIO SPECTRUM ALLOCATION.

Localizer. The localizer provides lateral guidance to the pilot by radiating its signal from an array of antennas usually placed about 600–1000 ft (183–305 m) beyond the stop end of the runway. There can be both a front and back course. The back course is an inexpensive means for providing guidance signals for approaches from the opposite direction when the wind direction dictates.

The localizer antenna array axis is perpendicular to the runway centerline. The horizontal extent and number of antennas in the array are determined by the need to confine the course guidance energy to narrow areas over the runway. Larger array dimensions are related inversely to the beam width of the signal in space. Narrower beams of energy prevent signal energy from being incident on buildings and then reflecting to cause course quality derogation. The dimensions of the array range from 40 to 130 ft (12 to 40 m) and the number of antennas from 6 to more than 20.

The three main types of localizer antenna are the V ring, a rather omnidirectional element; the traveling-wave types, which are unidirectional; and the log-periodic dipole (LPD) antenna, which is broad-band, usually unidirectional, and physically smaller (9 ft or 2.7 m long) and is located 6 ft (2.7 m) above ground. The LPD is the antenna in common use, typically

in arrays of eight elements. *See* ANTENNA (ELECTROMAGNETISM).

Glide slope. The glide-slope array, located 300–400 ft (90–120 m) to the side of the desired touchdown point on the runway, in some respects can be considered a localizer on its side. There are five types of glide-slope arrays in use. Three are image systems; that is, they use the ground to form the path, and the space distribution patterns discussed below are the result of signals from the ground combining with direct signals from the antennas. The null-reference type, the simplest and most common of the image systems, uses two antennas 14 and 28 ft (4.3 and 8.6 m) above ground. The elements in the image arrays are three collinear dipoles backed by a corner reflector.

The performance of the image glide slope critically depends on the smoothness of the Earth's surface in front of the array out to 3 or 4 mi (4.8 or 6.4 km) into the approach region. A minimum of 1000 ft (305 m) of ground in front of the array with variations no greater than 12 in. (0.3 m) is desired. To overcome these limitations, special nonimage systems are used at approximately 20 sites. The two types of nonimage systems are the waveguide type and the endfire glide slope.

Theory. The localizer and glide-slope courses are formed by a technique called space modulation. A carrier frequency is generated in the transmitter, then amplitude-modulated with discrete 90-Hz and 150-Hz audio tones, thus giving sidebands 90 Hz and 150 Hz above and below the carrier. Some of the sideband energy is separated from the carrier by use of a passive, radio-frequency bridge. The resulting two types of energy, namely, carrier with sidebands (CSB) and sidebands only (SBO), are fed to the array such that the CSB main beam covers the general course area and SBO beams exist on either side of the centerline (**Fig. 2**). A sharp null in SBO energy is placed precisely on the centerline. The superposition of these signal patterns provides the space modulation that is interpreted by the airborne receiver as course guidance information.

In particular, the SBO lobes differ in phase by 180° on opposite sides of the null. In the case of the localizer (Fig. 2*a*), the 90-Hz sidebands of the CSB and SBO signals are in phase to the left of the centerline and therefore add, while the 150-Hz sidebands of these signals are 180° out of phase there and therefore subtract, thereby tending to cancel each other. To the right of the centerline the reverse is true: the 150-Hz sidebands are in phase and add, while the 90-Hz sidebands are 180° out of phase and subtract. Since only the CSB signal is received on the centerline, the 90-Hz and 150-Hz sidebands have equal amplitude there. The glide-slope theory is similar (Fig. 2*b*). The ILS receivers thus produce 90-Hz signals when the aircraft is above the glide-slope course and to the left of the localizer course, 150-Hz signals when the aircraft is below and to the right, and equal 90-Hz and 150-Hz signals when the aircraft is on course.

Receivers and display. ILS receivers are commonly double-conversion, superheterodyne types with amplitude-modulation (AM) detectors. The output of the detector is fed to 90-Hz and 150-Hz filters, after which the signal is rectified. In the most basic circuit this signal is used to drive zero-centered, dc microammeters. *See* AMPLITUDE-MODULATION DETECTOR; ELECTRIC FILTER; RADIO RECEIVER.

The receiver output is displayed in the cockpit on what is called a course deviation indicator (CDI). This instrument contains four basic elements: a vertical needle or pointer for displaying localizer course information, a horizontal needle for glide-slope information, and two flags which show when indications from the needles should be disregarded. (Marker beacon lights are usually on another display.) The needles deflect in a direction corresponding to the greater of the two detected modulations. The design and calibrations are for edge-to-edge deflection of the localizer needle to represent ±1.5 to 3.0° in azimuth depending on runway length, and for edge-to-edge deflection of the glide-slope needle to represent ±0.7° in elevation. For example, when the aircraft is left and below course, the pilot sees the needles deflected to the right and upward (Fig. 1) and should steer toward the needles to reach the on-courses. The on-courses turn out to be a vertical plane for the localizer and a cone open upward with a 184° vertex angle for the glide slope, making a 3° path angle for flight. In an aircraft flying on the localizer and glide-slope centerlines, the CDI shows centered needles, with the warning flags out of view.

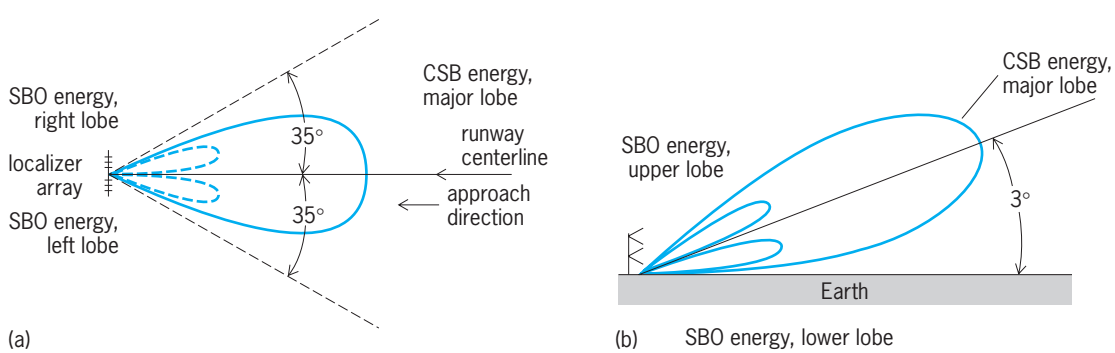

Fig. 2. Signal distribution space patterns for carrier-plus-sideband (CSB) energy and sideband-only (SBO) energies. (*a*) Localizer patterns in horizontal plane. (*b*) Glide-slope patterns in vertical plane. Angles are exaggerated.

Total localizer coverage is provided 35° on either side of the centerline. If narrower coverage is needed to protect against reflections, a two-frequency system is used and is based on the capture principle. The pilot need not be aware of any difference.

Marker beacons. The 75-MHz marker beacons are typically placed at two or three locations under the approach path on the extended runway centerline. They radiate a beam directly upward to indicate when the aircraft is at a precise location along the approach path. Because of the increase in the availability of other quality navigation information, the use of the markers is declining.

Sometimes associated with the instrument landing system is a low-frequency (200–400-kHz band), nondirectional radio beacon called a compass locator (COMLO), colocated with the outer marker. While very common and useful in the past, it too is finding less use, with other navigation signals supplanting it. *See* AIR NAVIGATION; ELECTRONIC NAVIGATION SYSTEMS.

Richard H. McFarland

Bibliography. K. W. Bose, *Aviation Electronics*, 4th ed., 1990; Federal Aviation Administration, *Airman's Information Manual*, rev. ed. 1994; M. Kayton and W. Fried, *Avionics Navigation Systems*, 2d ed., 1997.

Instrument science

The systematically organized body of general concepts and principles underlying the design, analysis, and application of instruments and instrument systems. Instruments are devices for the acquisition and processing of information acquired from the physical world. They embrace, in particular, measuring and observation instruments.

Nature and scope. Instruments are very diverse in function and form. They differ according to measurands, range of magnitudes and the dynamic variations of the measurand, required accuracy, and nature and environment of application. Equally diverse is the range of technologies that can be used to realize a particular instrument function. To enable this information to be handled in such a way that it can be usefully applied to the design, analysis, or application of instruments, it must be organized on the basis of a systematic framework of general concepts and principles. This framework is termed instrument science. The term measurement science is also used for this field, when it is desired to emphasize measuring instruments and the measurement processes for which they are used.

The basis of instrument science is, first, the consideration of instruments as members of the general class of information machines; and second, an analysis and synthesis of the instruments as systems, applying the principles of system engineering.

Information machines. There exists a wide class of machines whose function is to acquire, process, and feed out information. This class includes measuring instruments, control apparatus, communication equipment, and computers. These machines function by the transformation of an input physical variable into an output variable in such a way that the output is functionally related to the input. Thus the output carries information about the input. This basic principle of functioning determines the general features of the analysis and synthesis of information machines. *See* COMPUTER; CONTROL SYSTEMS; ELECTRICAL COMMUNICATIONS.

Since information machines operate by maintaining a functional relation between a physical input and a physical output, they can be viewed as a material realization of a mathematical relation or model. Thus the analysis and synthesis of information machines relies on systematic methods of mathematical description of signals (information-carrying physical variables) and on mathematical modeling of signal transformation by machines. *See* INFORMATION THEORY.

More generally, the characteristics of the magnitude and time variations of a physical variable, which carry information about the objects and events of the real world, are termed symbols for those objects or events. The principles of symbolic representation form an important aspect of instrument science.

Instruments as systems. Instruments and indeed other forms of information machines are conveniently analyzed and synthesized as systems. A system is a set of interconnected components functioning as a unit. There is a set of general principles and techniques for treating complex entities as systems. It is known as the systems approach, and is based on the decomposition of the complex whole into individual components. Thus an instrument would be considered, for example, as an interconnection of a sensing element, a signal amplification element, a signal transmission element, a display, and so on.

The individual components are considered in the first instance in terms of function rather than form. Thus a particular instrument sensor is considered in the first instance as a displacement-to-voltage transducer rather than as a strain gage or a differential transformer.

Considering systems, such as instruments, as structures of functional building blocks makes clear that a very wide variety of instruments can be constructed from a much smaller variety of building blocks organized into a small variety of basic structures, such as a chain or loop connection. Further, it makes clear that apparently different systems—say, electrical and mechanical, or physical and biological—are essentially analogous.

An essential feature of the systems approach, following from its treatment of components of systems as functional blocks, is the use of mathematical models for systems analysis and design. Systematic methods of modeling are employed. The systems approach is further based on the application of systematic methods of design.

Finally, the systems approach to instruments and similar machines stresses the fact that instruments in their application are always embedded in a supersystem such as a control system or information system. It is the requirement of the supersystem which

determines the design of the instrument. The instrument approach also stresses the significance of the environment in which the instrument is to operate, and the significance of all the phases of its life—from design and manufacture, through installation, commissioning, operation, and ultimate modification or disposal. *See* SYSTEMS ANALYSIS; SYSTEMS ENGINEERING.

Instrument system architecture. A typical measuring or observation instrument consists of a number of subsystems. First there is the system under measurement. This is connected to a sensor system and acts on the sensor by a flow of matter or energy. The sensor converts this flow into a signal, maintaining a functional relation between the input flow and the information-carrying characteristics of the signal. There is usually a signal-conditioning bloc, which converts this signal into a symbol. The symbol, in turn, may be conveniently handled by the following bloc, which performs any required functions of information transformation and communication. This system passes the information to the effector bloc, to further processing, or to the human operator. The measuring instrument system operates under the control of a control bloc. *See* TRANSDUCER.

An important part of the system is the human-machine interface. Through this interface, the operator effects supervisory control of the measurement process. The interface embodies also any displays. In the great majority of modern systems, and to a rapidly increasing extent, once information has been acquired by a sensor and conditioned, it is processed and effectuated by standard computing equipment. Thus, much of instrumentation is implemented by modern computer equipment. *See* DIGITAL COMPUTER; HUMAN-MACHINE SYSTEMS.

Models. An essential feature of instrument science is the development of systematic schemes for mathematical modeling or representation of the information-carrying physical variables or signals, and of the transformation performed by them on machines.

There are several levels of models of instruments employed: embodiment, power flow, signal flow, and information flow.

Embodiment models. These models relate the physical inputs and outputs of an instrument, or instrument component, in terms of features of its construction, that is, geometry and material properties.

Power-flow models. These models are based on considering power flow as a product of two conjugate variables, one an effort and one a flow. In mechanical systems they are force and velocity; in electrical systems, voltage and current; in fluid systems, pressure and flow; and in thermal systems, temperature and heat flow. All physical elements perform one of the following functions on power flow: storage, conversion, or transmission. The main power flow in elements may be modulated by controlling signals.

A modeling scheme based on these principles leads to a convenient classification of instrument elements and the recognition of analogies. Instrument elements such as sensors are basically one of three types: (1) power converters, such as a coil moving in a magnetic field, which transforms either a velocity into a voltage or a current into a force; (2) stores of two forms of energy, such as a moving-plate capacitor which stores mechanical and electrical energy and in which, for example, a mechanical velocity input produces an electrical voltage output; or, (3) a modulator, such as a resistance strain gage, in which the displacement straining the gage modulates the voltage-current relation of the gage.

Signal-flow models. Signal models are based on representing the signal as a time function (time domain) or as transforms such as Fourier and Laplace (frequency domain). The transformations performed by machines are modeled either in the time domain by convolution integrals or in the frequency domain by transfer functions. *See* FOURIER SERIES AND TRANSFORMS; LAPLACE TRANSFORM.

Information-flow models. The information level of modeling is concerned with information carrying characteristics of signals and their transformation. For example, in the case of digitally encoded information the signals are modeled in terms of digits, and their transformation in terms of truth tables and the like.

Signals. The principles underlying the analysis and design of signals, which carry information in information machines, and of signal transformation processes are a significant aspect of instrument science.

Intelligence. The processing of information in modern instruments is performed by subsystems of high information and knowledge processing power, which may be termed intelligent. The study of machine intelligence and the human cognitive processes which modern instruments emulate are a significant aspect of instrument science. *See* INTELLIGENT MACHINE.

Errors and uncertainty. The essential function of an instrument is the realization and maintenance of a prescribed functional transformation between input and output. In practice, the transformation cannot be realized or maintained perfectly. The result is error. Instrument science is concerned with providing a systematic framework for the description and analysis of errors.

This framework makes use of mathematical models of instruments. Errors are classified into two types, intrinsic errors and influence or disturbance errors. Intrinsic errors are the result of imperfections in the instrument input-output relations resulting, for example, from its imperfect determination or drift with time. Influence or disturbance errors are due to the effect of disturbing inputs acting on the instrument, such as changes of environmental temperature.

Errors may be determinate, that is, always the same for a particular value of input; or random, that is, different for different applications of a particular value of input. Random errors are modeled and analyzed statistically by assuming that they follow a gaussian or normal probability distribution. *See* DISTRIBUTION (PROBABILITY); PROBABILITY; STATISTICS.

Instrument design methodology. The great diversity of instrumentation means that instrument science cannot attempt to provide comprehensive catalogs of instrumentation. Rather it gives a systematic methodology of instrument design, based on the methodology of systems engineering, which enables appropriate instruments or instrument systems to be designed for a particular requirement.

The essential features of the methodology are primarily as pointed out above: the synthesis of instruments as systems of functional building blocks described by mathematical models. The other aspects are quantitative formulation of specifications in terms of objective criteria, the use of formal methods of optimization, and the application of formal techniques of decision in the consideration of alternative candidate designs. *See* DECISION THEORY; OPTIMIZATION.

Organization of technical knowledge. Knowledge about the working principles of instruments and instrument systems must be systematically classified and organized, so that it can be conveniently stored and retrieved. Three approaches may be adopted in classifying technical knowledge about instruments. The first is based on decomposing instruments into functional components and providing organized catalogs of such components and the principles of their combination. The second is organizing knowledge about measuring instrumentation on the basis of measurands. The third is organizing the knowledge about instrumentation on the basis of applications.

L. Finkelstein

Bibliography. L. Finkelstein, Measurement and instrumentation science—an analytical review, *Measurement*, 14:3–14, 1994; L. Finkelstein and K. T. V. Grattan (eds.), *Concise Encyclopedia of Measurement and Instrumentation*, Pergamon Press, Oxford, 1994; P. H. Sydenham (ed.), *Handbook of Measurement Science*, 3 vols., 1982, 1983, 1992.

Instrument transformer

A device that serves as an input source of currents and voltages from an electric power system to instruments, relays, meters, and control devices. The basic design is that of a transformer with the primary winding connected to the power system, and the secondary winding to the sensing and measuring equipment. Data from these devices are necessary for the operation, control, and protection of the power system. The primary reason for setting up the instrument-transformer interface is to provide analog information at low voltage levels, insulated from the higher system voltages. The range of use is from 480 V through the maxima of the 765–1000-kV power systems. *See* ELECTRIC POWER SYSTEMS; TRANSFORMER.

Current transformers are connected in series with the power conductor. In many cases this conductor serves as the one-turn primary. The principal types are the window type, where the power conductor is passed through a hole in the center of the current transformer; the bar type, where the power conductor is fastened to a one-turn bar which is part of the current transformer; and the bushing type, a toroidal core and winding that is slipped over the insulating bushings of circuit breakers, transformers, and so forth. At voltages of 345 kV and above, the current transformer may be mounted on top of a tall porcelain insulator. The standard secondary rating is 5 A, except for a few 1-A current transformers used for relaying, mostly outside the United States.

Voltage transformers and coupling capacitor voltage transformers are connected in parallel from one conductor to another or to ground. The coupling capacitor voltage transformer is widely used at the higher system voltages of 115 kV and above. It is a voltage transformer tapped across part of a capacitor unit connected from the conductor to ground. The standard secondary ratings are 120 (115) V line to line, or 69.3 (66.5) V line to neutral.

There are other means of providing primary system information to the secondary equipment, involving electronics, light pipes, optical fibers, and so forth. Some of these are in service on power systems. They are used primarily for instrumentation and data acquisition.

A typical equivalent circuit diagram for these transformers is shown in **Fig. 1**. The objective of the design of instrument transformers is to have the primary quantities precisely reproduced in the secondary, in both magnitude and phase angle, and under all steady-state and transient conditions. In practice, this goal can only be approached since the iron in the transformer core must be magnetized (by the exciting impedance and current, Z_e and I_e, in Fig. 1) resulting in losses through the device. The actual secondary current available I_s is equal to the secondary value I_r of the primary current I_p minus the exciting current I_e. Air-core (linear coupler) transformers, which do not suffer from these losses, are employed in relaying, but their use is quite limited.

Accuracy requirements. For instrumentation and especially for metering used for billing and revenue, very high accuracy is important for load current (which ranges from 0 to 5 A in the secondary) and load voltages (which range from 90 to 110% of the rated value). On the other hand, for protective relaying, reasonable ratio accuracy is important over a

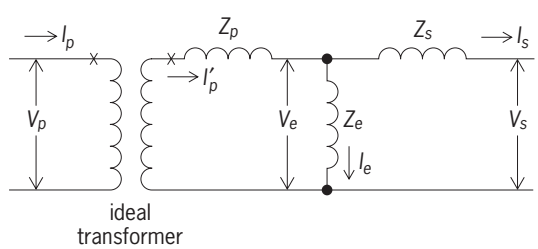

Fig. 1. Equivalent diagram for instrument transformers. The ideal transformer is required to change volts and amperes from primary to secondary. With per-unit values, the ideal transformer ratio is 1:1, and so for simplification is normally omitted. The ideal transformer has the ratio 1:*n* for current transformers and *n*:1 for voltage transformers, where *n* is the number of turns.

very wide range from just above zero to a maximum of 100 to 300 A in the secondary. Phase-angle error as such is not directly a problem in relaying. Thus, particularly for current transformers, metering and relaying designs are quite different and are not interchangeable. For voltage transformers, the values of impedance in the primary and secondary, Z_p and Z_s, are kept low to minimize voltage loss and phase shift from the primary to the secondary. For current transformers, Z_e is kept high to minimize current loss from primary to secondary. Z_p is negligible when the primary has one or a small number of turns.

Load. The equipment (load) connected to the instrument transformer secondary is commonly called the burden Z_B and is often expressed in volt-amperes (VA_B) for each device. Thus Eq. (1) is valid, where

$$Z_B = \frac{VA_B}{I^2} \quad \text{ohms} \tag{1}$$

I is the current in amperes at which the burden is measured. The total load Z_L is the sum of all the connected equipment, including the connecting leads. The phase angle is important, especially for metering.

Performance standards. The performance of instrument transformers is standardized in the United States by American National Standards Institute Standard C57.13 for both metering and relaying service. Accuracy classes are specified. For metering, transformer correction factors for standard burdens provide the means to determine performance for both current transformers and voltage transformers. For relaying, current transformer performance can be determined by several methods.

Current transformer. For current transformers connected in the three phases of a three-phase power system, the current transformer ratio is selected by determining the maximum current that can flow for any normal or emergency operating condition. A current transformer ratio R_c is selected so that this current, divided by the ratio, is around 4 but less than 5 A. Most indicating instruments used have 5-A movements regardless of the scale, so this avoids pegging the indicator. Also, much of the equipment to be connected has a 5-A rating.

For relaying, the next step is to determine the maximum fault current I_F at which the relays must operate. Assume that the current transformer will pass this current through the total burden, including the current transformer secondary impedance Z_s with negligible exciting current I_e. To accomplish this, the current transformer must develop an exciting voltage V_e given by Eq. (2), where Z_s has a value of

$$V_e = \frac{I_F}{R_c}(Z_s + Z_{B,\text{total}}) \tag{2}$$

0.0025 ohm per turn.

There are two types of current transformers used in relaying: ANSI class C, where the leakage flux in the core does not significantly affect the ratio (Z_s is small); and ANSI class T, where the leakage flux has appreciable effect on the ratio (Z_s is significant). Many relaying current transformers are class C, and their performance can be calculated and evaluated from excitation curves supplied by the manufacturer or determined by test. Typical curves for a multiratio-bushing-type current transformer are shown in **Fig. 2**. The value of V_e, determined from

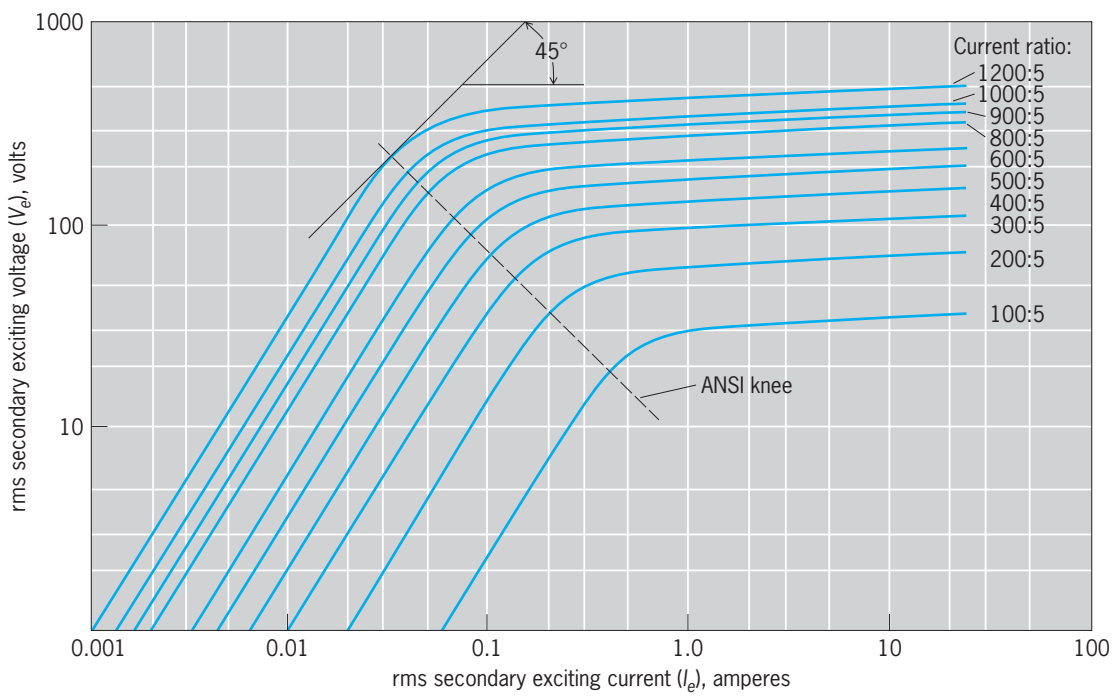

Fig. 2. Typical excitation curves for a multiratio class C current transformer. Below the line labeled ANSI knee, the exciting current for a given voltage for any unit will not exceed the curve value by more than 25%. Above this line, the voltage for a given exciting current for any unit will not be less than 95% of the curve value. (*After American National Standards Institute, Requirements for Instrument Transformers, ANSI/IEEE Stand. C57.13, 1978*)

Standard burdens for current transformers with 5-A secondaries*

ANSI burden designation	Impedance, ohm	Volt-amperes at 5 A	Power factor	Service
B-0.1	0.1	2.5	0.9	Metering
B-0.2	0.2	5.0	0.9	Metering
B-0.5	0.5	12.5	0.9	Metering
B-0.9	0.9	22.5	0.9	Metering
B-1.8	1.8	45.0	0.9	Metering
B-1	1.0	25.0	0.5	Relaying
B-2	2.0	50.0	0.5	Relaying
B-4	4.0	100.0	0.5	Relaying
B-8	8.0	200.0	0.5	Relaying

*After American National Standards Institute, *Requirements for Instrument Transformers,* ANSI/IEEE Standard C57.13, 1978.

Eq. (2), provides the value of I_e required. If, for the maximum fault, V_e for the connected burden is below the knee of the curves, I_e will be small and I_s essentially equal to the primary current divided by R_c.

For relaying class T current transformers, overcurrent ratio curves are supplied (**Fig. 3**) in terms of standard burdens connected to the terminals of the current transformer. The **table** shows these burdens. From the total burden connected, the current transformer performance can be evaluated from these curves.

As an example, consider a current transformer with a ratio of 800:5 and an external burden of 1.6 ohms. The maximum fault is 10,000 A primary. If the transformer is class C, V_e is given by Eq. (3),

$$V_e = \frac{10,000}{160}(0.4 + 1.6) = 125 \text{ V} \qquad (3)$$

where, from Fig. 2, I_e is approximately 0.045 A, quite small for the 62.5 A secondary for the fault. If the current transformer is type T, the burden for Fig. 3 is 1.6 ohms, and the ratio is essentially unity for the fault current of 62.5/5 in the secondary or 10,000/800 in the primary, both equal to 12.5 on the curve.

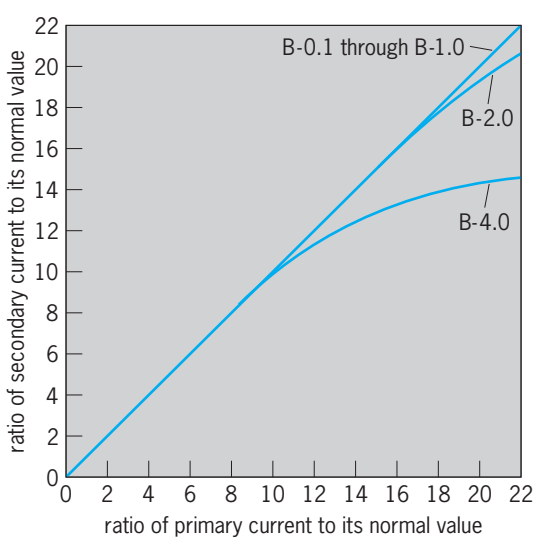

Fig. 3. Typical overcurrent ratio curves for class T current transformers. (*After American National Standards Institute, Requirements for Instrument Transformers, ANSI/IEEE Stand. C57.13, 1978*)

In many cases the ANSI accuracy class designation can be used without detailed information on the current transformer. In relaying, classes C and T are followed by a number indicating the voltage that the transformer will deliver to a standard burden at 20 times rated secondary current without exceeding 10% ratio error. This voltage is the terminal voltage V_s, and for multiratio current transformers applies at the full winding tap. Proportionately lower voltages are developed at the other taps. In the example above where $V_s = 62.5 \times 1.6 = 100$ V, any current transformer rated C100 or T100 or higher could be used.

Transient performance. Transient performance of instrument transformers is a complex subject, not covered in the present standards. In general, modern transformers transform the various harmonics and wave shapes very adequately except for the direct-current decrement. Fault currents usually are offset by various amounts. This occurs because at the instant of the fault the flux relations in the circuit are not correct for the sudden change in current. As a result, a dc component is produced which then decays according to the inductance/resistance (L/R) ratio of the power system. This is a random effect so that the dc peak is variable and may or may not reach its maximum possible value. This direct current inhibits current transformer performance by saturating the core so that I_e becomes very large, and I_s is reduced. It is not practicable to build a current transformer to minimize this effect, so it must be managed by relay design and application to avoid incorrect operations. As a generality, with the maximum dc offset, a current transformer will reproduce about a half-cycle or so of the wave before saturation reduces the secondary current to very low values. As the direct current decays, the current transformer recovers to eventually accurately reproduce the primary current. Thus relays operating in the order of the half-cycle may operate before saturation, while slower relays will be delayed. In differential-type protection where several current transformers are interconnected and must operate as a team, application rules are established to prevent misoperations.

Voltage transformers. Voltage transformers are seldom a problem in performance and in capacity as

long as the connected burden is within the rating of the units. In three-phase power systems, the common connections are three units connected wye-grounded-wye-grounded or two units in open delta across the three phases. The connected load is generally line to line. Special applications may require only one voltage transformer connected usually line to ground. Coupling capacitor voltage transformers and some voltage transformers have two secondaries. In such cases the primaries are connected wye-grounded. One secondary is connected wye-grounded to provide three-phase voltages for relays, instruments, and so forth. The other windings are connected in delta with one corner open (broken delta) and connected to ground relays.

Faults reduce the voltage, and power system overvoltage is carefully controlled so that coupling capacitor voltage transformers and voltage transformers operate only from very low voltages to around 100% of their ratings, so that saturation is not a problem. *See* ELECTRICAL MEASUREMENTS. J. Lewis Blackburn

Bibliography. American National Standards Institute, *Requirements for Instrument Transformers*, ANSI/IEEE Standard C57.13, 1978; J. S. Blackburn, *Protective Relaying*, *Principles and Applications*, 2d ed., 1997; W. A. Elmore (ed.), *Protective Relaying Theory and Practice*, 1994; Westinghouse Electric Corp., *Applied Protective Relaying*, 1982.

Instrumental conditioning

Learning based upon the consequences of behavior. For example, a rat may learn to press a lever when this action produces food. Instrumental or operant behavior is the behavior by which an organism changes its environment. The particular instances of behavior that produce consequences are called responses. Classes of responses having characteristic consequences are called operant classes; responses in an operant class operate on (act upon) the environment. For example, a rat's lever-press responses may include pressing with the left or with the right paw, sitting on the lever, and so on, but all these taken together constitute the operant class called lever pressing.

History. Instrumental conditioning was first studied systematically by Edward L. Thorndike in the late nineteenth century. Thorndike constructed puzzle boxes in which he placed cats and other animals. The animal could get out of the box (and, usually, get fed) by operating a latch or some other device that released the door. Thorndike found that the time it took the animal to get out of a particular box typically decreased gradually over many trials. He spoke of this as the law of effect; behavior was "stamped in" by the effects it had or, in other words, by its consequences. Instrumental learning in these types of situations was sometimes called trial-and-error learning.

Subsequently, instrumental learning was studied with many different species in many devices, such as mazes and runways. Experimenters asked questions such as whether this learning took place in discrete steps or continuously, and whether organisms learned movements (response learning) or properties of the environment (stimulus learning). The experiments had the common result that animal response became more likely when it had certain consequences. Graphs showing how behavior changed during an experiment were called learning curves: these plot time to respond as a function of number of trials, percentage of correct responses, or proportion of animals performing successfully. But the shapes of the curves depended so much on differences in experimental detail that the progress of learning could not be defined in any consistent way.

The performances of the animals were too complicated to be clearly interpreted. For example, measuring the time it took a rat to get through a maze showed little of how learning proceeded at a single choice point. An important innovation was B. F. Skinner's design of a chamber in which an animal could repeatedly emit some easily specified response without the experimenter's intervention (the response was said to be emitted because it was not produced by any particular stimulus). A chamber typically included lamps and speakers for presenting stimuli, switches attached to response devices such as levers or disks on the chamber wall, and dispensers that presented food or water as consequences for responses. These chambers were used to demonstrate reinforcement (an increase in behavior that occurs because of the consequences of responses), extinction (the return of behavior to its previous lower levels when these consequences are discontinued), and the effects of schedules of reinforcement (arrangements of intermittent consequences for responses, as, for example, after every tenth response or on the average once every 2 min). The relations between responses and their consequences are called contingencies.

Reinforcers and punishers. Consequences of responding that produce increases in behavior are called reinforcers. For example, if an animal has not eaten recently, food is likely to reinforce many of its responses. The effectiveness of a reinforcer depends on the response that is reinforced. Suppose that a rat is deprived of water but has had unlimited access to a running wheel. If access to water is made a consequence of running in the wheel, the opportunity to drink will reinforce wheel running. But suppose next that the rat is given unlimited access to water but is prevented from running in the wheel. If access to the wheel is made a consequence of drinking, the opportunity to run will now reinforce drinking. In other words, the reinforcement relation is reversible. Whether a particular event will reinforce a response or not depends on the relation between the response and the behavior for which its consequences provide an opportunity. Consequences of responding that are effective as reinforcers may include many kinds of events: simple sensory events like making contact with something as a result of reaching for it or seeing something that one has turned to look at; biologically significant events such as sexual contact

or the opportunity to eat or drink; and the removal or prevention of something, as in escape from or avoidance of painful stimuli.

Some consequences of responding called punishers produce decreases in behavior. The properties of punishment are similar to those of reinforcement, except for the difference in direction. The effects of reinforcement are temporary in that responding decreases to earlier levels when reinforcement is discontinued (extinction); the effects of punishment are also temporary in that punished responding returns to its earlier higher levels when punishment is discontinued. As with reinforcers, many types of events can serve as punishers, and their effects depend on their relation to the response that is punished. For example, if a painful stimulus makes a rat jump, the stimulus will probably not be an effective punisher of jumping, though it might work for other responses such as grooming.

Discriminative or occasion-setting stimuli. The consequences of its behavior are perhaps the most important properties of the world about which an organism can learn, but few consequences are independent of other circumstances. Organisms learn that their responses have one consequence in one setting and different consequences in another (stepping on the gas, for example, has some consequences when the traffic light is green that differ from those when the traffic light is red). Stimuli that set the occasion on which responses have different consequences are called discriminative stimuli (these stimuli do not elicit responses; their functions are different from those that are simply followed by other stimuli). When organisms learn that responses have consequences in the presence of one but not another stimulus, their responses are said to be under stimulus control. For example, if a rat's lever presses produce food when a light is on but not when it is off and the rat comes to press only when the light is on, the rat's presses are said to be under the stimulus control of the light. The relations among a discriminative stimulus, a response, and a consequence have been described by Skinner as a three-term contingency, in that all three terms must enter into what the organism learns: in a particular "environmental setting," its "behavior" produces certain "outcomes." These contingencies do not explain instrumental learning; they are what the organism learns. *See* CONDITIONED REFLEX; LEARNING MECHANISMS. A. Charles Catania

Bibliography. A. C. Catania, *Learning*, 4th ed., 1997; M. Sidman, *Tactics of Scientific Research: Evaluating Experimental Data in Psychology*, 1988; B. F. Skinner, *The Behavior of Organisms*, 1938; B. F. Skinner, *Contingencies of Reinforcement*, 1969; E. L. Thorndike, *Animal Intelligence*, 1911.

Instrumentation amplifier

A special-purpose linear amplifier, used for the accurate amplification of the difference between two (often small) voltages, often in the presence of much larger common-mode voltages. An instrumentation amplifier has a pair of differential (usually high-impedance) input terminals, connected to sources V_{in1} and V_{in2}; a well-defined differential-mode gain A_{DM}; and a voltage output V_{out}, satisfying the relationship given in Eq. (1). It differs from an operational

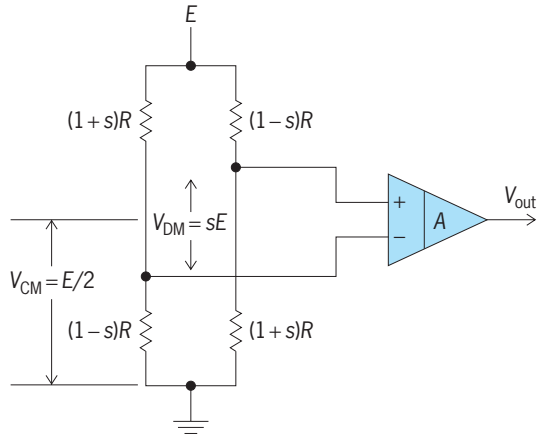

Fig. 1. Strain-gage bridge which produces a small wanted signal on a large bias, connected to an instrumentation amplifier that responds only to the wanted signal. The amplifier symbol indicates that the differential gain is controlled, unlike that of an operational amplifier.

$$V_{out} = A_{DM}(V_{in1} - V_{in2}) \tag{1}$$

amplifier (op-amp), which ideally has infinite open-loop gain and must be used in conjunction with external elements to define the closed-loop transfer function. At one time instrumentation amplifiers were built in discrete or hybrid form using operational amplifier and resistor networks, but they are now readily available as inexpensive monolithic integrated circuits. Typical commercial amplifiers provide preset gains of 1, 10, 100, and 1000. In some cases, the gain may be set to a special value by one or more external resistors. The frequency response invariably is flat, extending from 0 (dc) to an upper frequency of about 1 kHz to 1 MHz. *See* INTEGRATED CIRCUITS; OPERATIONAL AMPLIFIER; RESISTOR.

Instrumentation amplifiers are used to interface low-level devices, such as strain gages, pressure transducers, and Hall-effect magnetic sensors, into a subsequent high-level process, such as analog-to-digital conversion. **Figure 1** shows a typical strain-gage bridge, which generates a small differential-mode output V_{DM}, proportional to the strain factor s, superimposed on a large common-mode voltage V_{CM}. The desired output level might be several volts, and is typically ±10 V full-scale. *See* HALL EFFECT; PRESSURE TRANSDUCER; STRAIN GAGE.

An important requirement of an instrumentation amplifier is that it responds only to the difference signal and rejects other signal components. A real amplifier has some finite response to its common-mode input, as defined in Eq. (2). The common-mode

$$V_{CM} = \frac{V_{in1} + V_{in2}}{2} \tag{2}$$

rejection ratio (CMRR) is a measure of how well the amplifier performs this rejection. If A_{CM} is the common-mode gain, Eq. (3) gives the common-mode

$$\text{CMRR (dB)} = 20 \log_{10} \frac{A_{\text{DM}}}{A_{\text{CM}}} \tag{3}$$

rejection ratio, conventionally expressed in decibels.

At low frequencies (0 to a few kilohertz), common-mode rejection ratios between 90 and 120 dB are possible. At higher frequencies, the common-mode rejection ratio is eroded by small imbalances in the differential signal path. For example, differences in the source impedances may result in the conversion of a common-mode signal to a differential-mode one. Many effects in the amplifier cause degradation of the common-mode rejection ratio at high frequencies.

A classical topology uses two operational amplifiers as voltage followers (to provide the required high-input impedance) and a third operational amplifier in a differencing configuration (**Fig. 2***a*). The output voltage is given by Eq. (4), so that the differential-

$$V_{\text{out}} = \frac{R_2}{R_1}(V_{\text{in1}} - V_{\text{in2}}) \tag{4}$$

mode gain is simply the resistor ratio R_2/R_1. If like resistors match perfectly, the common-mode rejection ratio is limited only by the active elements. However, a very high degree of resistor matching is required in this topology to achieve high rejection of the common-mode signal; for example, a 0.1% mismatch would limit the common-mode rejection ratio to 60 dB.

A more accurate topology is used in many high-performance instrumentation amplifiers (Fig. 2*b*),

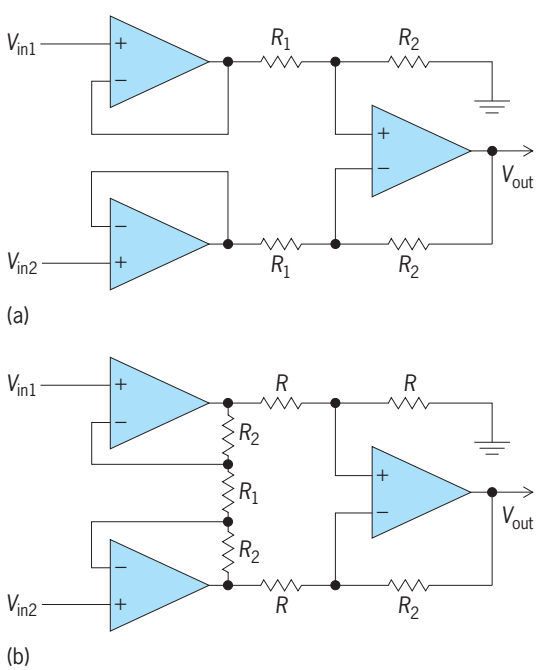

Fig. 2. Instrumentation amplifier topologies. (*a*) Simple topology based on three operational amplifiers. (*b*) Improved topology used in a modern integrated-circuit instrumentation amplifier.

with output voltage given by Eq. (5). By raising the

$$V_{\text{out}} = \left(\frac{2R_2}{R_1} + 1\right)(V_{\text{in1}} - V_{\text{in2}}) \tag{5}$$

gain ahead of the differencing amplifier (now a unity-gain structure), the effect of all resistor mismatches is minimized, and the amplifier has a high common-mode rejection ratio even when resistor matching is moderately poor, particularly at high gains. Many other configurations have been used with varying success. As with conventional operational amplifiers, modern monolithic instrumentation amplifiers have low input offset voltage (as low as 50 microvolts), low input-bias current (0.1–50 nanoamperes), and low voltage and current noise. Laser trimming of the integral resistors is often employed to achieve very high-gain accuracy. *See* AMPLIFIER; DIFFERENTIAL AMPLIFIER.

Barrie Gilbert

Bibliography. E. W. Greeneich, *Analog Integrated Circuits*, 1996; J. C. C. Nelson, *Operational Amplifier Circuits: Analysis and Design*, 1995; D. H. Sheingold (ed.), *Transducer Interfacing Handbook*, 1980.

Instrumented buoys

Unattended, floating structures equipped with systems for collecting, processing, and transmitting meteorological and oceanographic data. Such information is useful for many purposes, including storm warnings and forecasts, coastal engineering, climatology, and oceanographic and atmospheric research.

In earlier times, merchant vessels supplied almost all our knowledge of ocean weather and climate. Ships continue to play a major role in reporting weather and surface conditions (waves, swell, and temperature), but the irregular and sporadic nature of these observations, especially in the sparsely traveled Southern Hemisphere, makes it difficult to obtain a systematic picture of the state of the atmosphere and how it changes with time. Apart from crude estimates of surface currents through dead reckoning, ships in transit tell us nothing about the ocean. So, almost all our knowledge of ocean currents and their variability today comes from instrumented buoys, which measure ocean currents and related properties, particularly temperature. Moored systems, anchored to the ocean bottom, record currents and water properties as the flow passes the buoy (the eulerian framework). Drifting buoys move with the waters, indicating where they go (the lagrangian framework). In both cases, additional sensors may record a variety of parameters, such as temperature, pressure, acceleration, and water properties.

Moored systems. Surface moorings come in many shapes and sizes depending upon a number of factors, such as research versus operational requirements, duration of deployment, surface versus subsurface measurements, strength of currents, and weather considerations. To operate for a long time,

Fig. 1. Atlas toroid buoy with its meteorological instrumentation. (*Pacific Marine and Environmental Laboratory, NOAA*)

moored buoys require highly reliable components, including a strong mooring line to prevent failure due to wear and tear in heavy seas. This in turn requires large buoys with substantial surface flotation to support the weight of the line. These buoys support a suite of atmospheric sensors and power systems for their operation and data telemetry.

Surface buoys have a long history, particularly along the United States Atlantic, Gulf, and Pacific coasts in reporting offshore weather and sea surface conditions for shipping, fishing, and recreational use. Their size and shape vary enormously, depending upon their application. Doughnut-shaped toroid, discus, and boat-shaped NOMAD (Navy Oceanographic and Meteorological Automatic Device) buoys can survive very harsh weather and current conditions (**Fig. 1**). In response to the impact of the equatorial Pacific on year-to-year variations in the Earth's weather (El Niño-La Niña), an extensive array of moored toroids have been deployed across the equatorial Pacific (about 70) and Atlantic (about 10) oceans. They transmit weather and water temperature from the surface down through the shallow thermocline (the boundary between the warm upper and cold deeper waters) to 500 m. The weather service uses the real-time data sent from these buoys for their forecasting activities. The public can also access the data via Internet websites. To ensure reliable performance, the buoys get serviced or replaced every year. *See* EL NIÑO; METEOROLOGICAL INSTRUMENTATION; TELEMETERING; WEATHER FORECASTING AND PREDICTION.

Studies of the subsurface ocean typically employ moorings with flotation below the surface. Subsurface moorings experience much less fatigue of the hardware and mooring line due to the absence of wave motion, permitting the use of lighter-weight, less expensive hardware. More importantly, the absence of surface wave motion results in a much calmer mooring line, greatly facilitating accurate current measurements. Subsurface moorings also eliminate the risk of piracy and entanglement with fishing nets, a major source of equipment and data loss for surface moorings. However, these moorings require costly subsurface flotation devices because they must have the strength to withstand the higher hydrostatic pressure. Large steel spheres are used for near-surface flotation applications, and glass balls are used at greater pressures. Syntactic foam, consisting of a matrix of glass microspheres embedded in epoxy, has proven very effective in applications where the flotation must have a special shape to support instruments, yet remain streamlined to minimize drag forces in strong currents.

A tall mooring reaching close to the surface in deep water, as in the Gulf Stream, experiences very strong drag forces due to current flow past the wire, the weight of the instrumentation, and the weight of the flotation device at the top of the mooring. In the absence of sufficient flotation, these drag forces can pull the mooring over many hundreds of meters and can exert a very large drag force on the anchor, which therefore must be very heavy (typically an assembly of discarded railroad wheels weighing 1–3 tons). Subsurface moorings require considerable design attention because the expected drag forces determine the required flotation, the minimum necessary strength for the mooring wire, and the diameter of the wire. The mooring line typically consists of a jacketed steel wire for strength and for protection against fish bite in the upper ocean, and lighter-weight synthetic fiber-rope at greater depths.

For many years the use of subsurface flotation precluded the possibility of real-time telemetry since radio waves cannot pass through the electrically conducting seawater. New technologies avoid this limitation if the top of the mooring reaches close to the surface. In one design, a small winch at the top of the mooring sends a small transmitter module to the surface that transmits its data to satellites, and then pulls itself back below the surface. In another design, a small float rides the surface but remains attached to the top of the mooring and its flotation device by means of a stiff yet flexible tether known as an S-tether. The stiffness prevents kinking, while the flexibility lets the small surface buoy with its embedded radio transmitter follow the surface motion without affecting the rest of the mooring.

A novel mooring design called the moored profiler consists of a taut subsurface mooring along which a module drives itself up and down, recording water properties and currents between the bottom and near the surface. The profiler can also send the data to the top of the mooring inductively via a magnetic coupling between the profiler and the wire. Using an S-tether or other surface module, the system can telemeter the data with little delay for analysis. Real-time telemetry also means data are available up to the point of failure.

Drifting buoys. These buoys drift with the waters, either as surface drifters or as subsurface floats. In order to follow the waters while minimizing the effect of winds, surface drifters have very little exposure at the surface but have large drag elements hanging beneath. Today, the nearly universal drifter design includes a spherical flotation element and a

long, large-diameter canvas tube with large holes, known as a holey sock. Its large size guarantees that the drifter moves with the waters around the sock, readily pulling the small surface sphere with it. Depending upon the application, the sock might hang just to 10 m or as deep as 100 m. The movement of the drifters are tracked by satellites. These drifters have made major contributions to our knowledge of the surface circulation of all the oceans. Some new designs come equipped with temperature sensors, and increasingly, with barometric pressure gauges. They also telemeter valuable sea surface weather information, including acoustic measurements of winds. These buoys cost-effectively complement the functions of the surface buoys.

Subsurface drifters also have a long history, beginning with the development of the neutrally buoyant float equipped with a small pinger (a sound source) so that a ship can follow it as it drifts. The floats achieve their neutral buoyancy at depth by adjusting their weight very carefully to match that of the displaced waters (Archimedes' principle). A float must also have less compressibility than water for it to remain at depth. Suppose, for example, a disturbance such as an internal wave deflects the float upward. The float will expand less than the surrounding waters due to its compressive stiffness, making it slightly heavier than the surrounding waters, causing it to sink back toward its equilibrium depth. These early float-based studies revealed unexpectedly energetic subsurface currents. *See* ARCHIMEDES' PRINCIPLE.

The acoustic transparency of the ocean, especially at low frequencies (a few hundred hertz), led to the development of low-frequency sound sources, enabling acoustic tracking over more than 2000 km though the deep sound or SOFAR (sound fixing and ranging) channel. Earlier float designs had the floats transmitting signals on a precisely timed basis to two or more underwater listening stations or hydrophones. Given the speed of sound in water, one can estimate the position of the float to a few kilometers' accuracy, and hence its drift as a function of time. Today, acoustic tracking works in the opposite direction: moored sound sources function as an underwater navigation system for drifting floats, which listen to and record the arrival times of the signals. At the end of a float's underwater mission, it surfaces and telemeters the stored data via satellite to the users, who can then reconstruct the trajectory of the float.

In addition to acoustic tracking, floats always measure temperature and pressure. Recent float designs include oxygen sensors. These RAFOS (SOFAR spelled backward) floats can either float at constant depth or follow density surfaces as they change depth. In the latter case, the floats incorporate a compressible element to give them the same compressibility as that of seawater, freeing them to follow the waters in the vertical. To ensure that they stay with the same waters, the floats have virtually no coefficient of thermal expansion (they are made of borosilicate glass). Using the same argument as above with respect to pressure, imagine a float temporarily displaced upward into warmer waters. The lack of thermal expansion compared to that of seawater makes it slightly heavier than the surrounding waters, forcing it back to the water temperature it came from. More generally, it follows that once the float has settled on a given density surface, it will stay

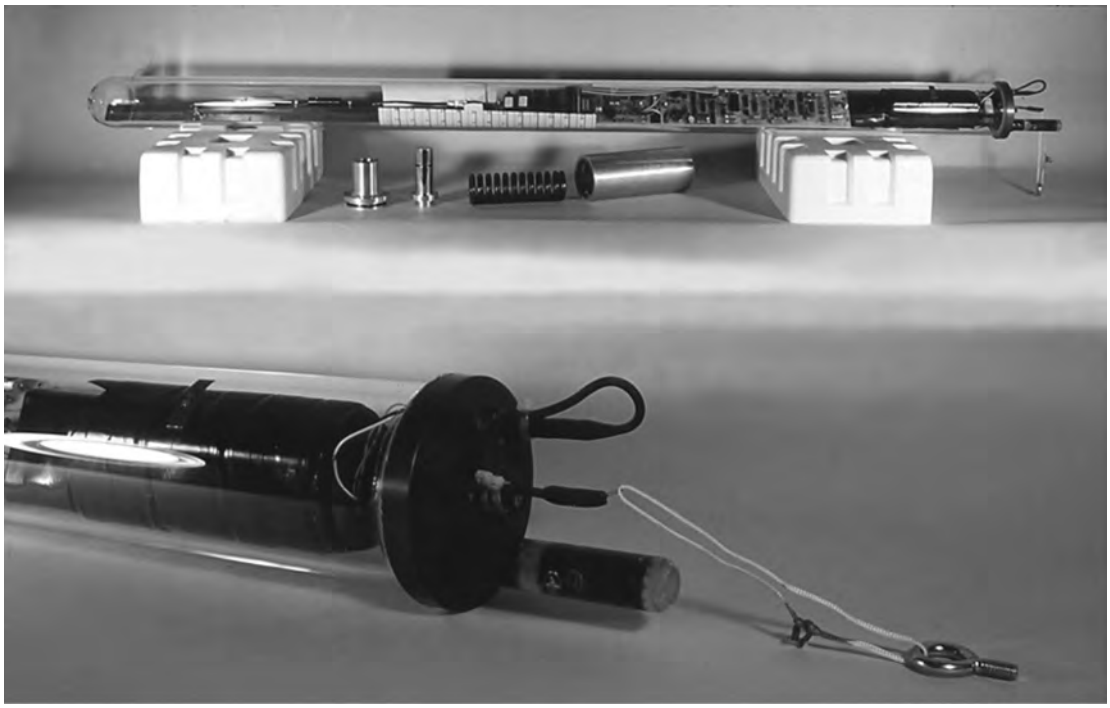

Fig. 2. RAFOS float showing the whole instrument, with a detail of the right end in the foreground. (*Graduate School of Oceanography, University of Rhode Island*)

with that surface regardless of temperature and salinity, since neither can affect the density of the float. This feature confers major advantages for the study of dispersion and mixing processes in the ocean. These acoustically tracked floats have yielded much insight into the horizontal movements of the ocean over a wide range of space and time scales and at all depths (**Fig. 2**).

Another class of floats focuses on the vertical structure of the ocean. Drifting at some preset depth (typically 1000–1500 m), they ascend to the surface while sampling pressure, temperature, and salinity. While at the surface, the floats use the Global Positioning System (GPS) to determine their position, which they broadcast, along with the profile data, to a satellite. Alternatively, some satellites can determine their positions. The floats then return to depth and continue their drift for a predetermined period (typically a couple of weeks) before repeating the process. Their change in position between surface reports gives valuable information on the subsurface circulation. Satellite navigation allows their deployment anywhere in the global ocean. The combination of wide-area deployment of these profiling floats and telemetry provides a very effective means for obtaining wide-area information on the interior state of the ocean in real time (**Fig. 3**). Advanced planning calls for their deployment around the world in large numbers as part of the Global Ocean Observing System. *See* NAVIGATION; SATELLITE NAVIGATION SYSTEMS; SOFAR; SONAR; UNDERWATER SOUND.

Sensors. Buoy sensors can be categorized in many ways. For example, they can be grouped according to whether they measure scalar or vector properties. Scalar sensors measure the state of a fluid, such as temperature, pressure, humidity, salinity, and light intensity. Vector sensors measure speed and direction of winds and currents. Considerable development continues to take place in all categories of sensors, including research into the chemical and biological properties of seawater. Atmospheric temperature and humidity sensors must have shielding to protect them from heating by solar radiation, yet be open enough to ensure adequate ventilation. Pyroheliometers (radiometers) must remain clean and not become covered with dust, salt, or bird droppings. Wind anemometers must withstand exposure to salt spray. In the past, precipitation measurement presented formidable problems, but the tropical moorings across the Pacific have successfully measured rainfall. An acoustic method that measures the noise generated by raindrops as they hit the surface also shows promise. These new methods will help fill a serious gap in our knowledge of ocean climates. High winds and complex airflow patterns around ships have prevented the accurate measurement of precipitation over the oceans from commercial vessels.

Fig. 3. An autonomous profiler float, PALACE type. (*Webb Research Corp.*)

Underwater sensors traditionally measure temperature and velocity, also measuring pressure as an indication of depth. They must operate at high hydrostatic pressure and resist the corrosive effects of salt. These conditions continue to challenge ocean engineers, especially for multiyear deployments. Temperature sensors often sit inside the thermally conducting metal instrument housing. Velocity sensors come in two groups: directly measuring mechanical devices and acoustic sensors. The former come in many configurations. A common design consists of a rotor that spins around a vertical axis in response to the flow past it while a separate vane measures the direction of flow relative to the instrument and a compass inside monitors the orientation of the instrument with respect to magnetic north.

Acoustic sensors come in two types: local and profiling. Local sensors measure the difference in acoustic travel time between two points caused by the flow moving with the sound in one direction and opposite it in the other. With two or three such pairs, one can obtain the horizontal or all three components of motion. Profiling acoustic current meters depend upon the backscatter of signals from suspended matter in the water, typically zooplankton and small nekton. These instruments, called acoustic Doppler current profilers, send out short pulses in four oblique (relative to the vertical) beams and then immediately listen to the backscattered echoes. From the slight Doppler shifts of the return signals, such a sensor can determine the magnitude of flow in the direction of each beam, from which one can then calculate the corresponding horizontal and vertical motions. These instruments have proven very effective in studies of upper ocean currents by looking upward from subsurface moorings and downward from surface platforms. Drifters, floats, and profilers do not require current sensors since that information comes directly from their movement.

The widespread use of instrumentation has had

the beneficial effect of focusing attention on the development of reliable and stable conductivity sensors for estimating salinity. While conductivity sensors used for moored buoys have existed for a long time, the need to ensure that they remain stable for years while operating over a wide range of pressures, as well as at the surface (where biological fouling remains a particular problem), has led to redoubled engineering efforts. These efforts have already led to considerable improvement. Temperature and salinity constitute the two most important indicators of water type. Changes in these properties along float trajectories provide valuable information on the mixing and exchange of waters between different regions. Progress in the development of other biologically important sensors for particulate matter, downwelling light from the surface, chlorophyll, dissolved oxygen and carbon dioxide, nutrients, and pH holds much promise for the future. *See* OCEAN CIRCULATION; OCEANOGRAPHIC VESSELS; REMOTE SENSING; SEAWATER; THERMISTOR. H. Thomas Rossby

Bibliography. H. O. Berteaux, *Coastal and Oceanic Buoy Engineering*, 1991; *Sea Technology*, monthly.

Insulation resistance testing

The measurement of direct current or voltage drop across an electrical insulating material to determine its resistance by the application of Ohm's law. The main purpose of electrical insulation is to isolate the energized components of a system from each other and from ground. This requires that the insulation resistance to the flow of current be as high as possible. Good insulators have resistances in the range of 10^6 to 10^{20} ohms, and must also have acceptable mechanical, chemical, and thermal properties. *See* ELECTRIC INSULATOR; ELECTRICAL INSULATION; ELECTRICAL RESISTANCE; OHM'S LAW.

Although insulating materials in practical use can be solids, liquids, or gases, resistance measurements generally refer to solid insulation. Insulating liquids, as well as gases, are tested by applying an electric field by means of two electrodes submerged in the fluid, and raising the voltage to determine its dielectric strength. The insulating properties of a liquid can also be stated in terms of its specific resistance (or resistivity in ohm-centimeters). *See* DIELECTRIC MATERIALS; ELECTRICAL RESISTIVITY.

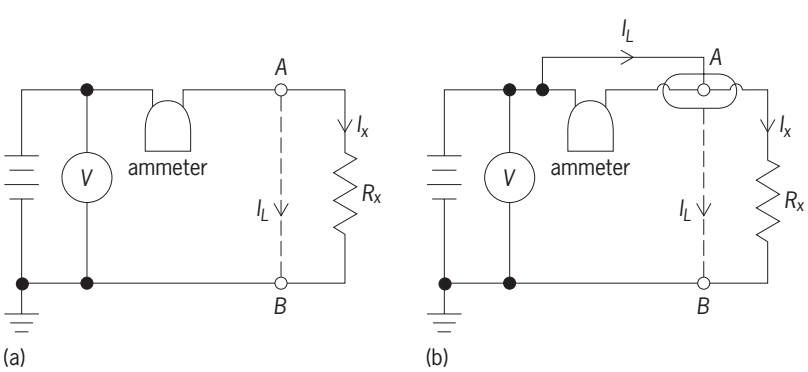

Fig. 1. Direct deflection method of resistance measurement (*a*) with no guard, and (*b*) with guard circuit applied at terminal *A* to eliminate the effects of leakage current.

Insulation resistance testing is performed either on samples of the insulation material or on the electrical equipment which uses the insulation. Tests on specimens are made to determine the material properties, while tests on the equipment are performed as a quality-control check at the factory or to determine the insulation condition while in service.

Insulation resistance in solids is made up of a combination of volume resistance and surface resistance. These two types of resistances arise because the current can physically flow through the volume of the insulation material or it can find other parallel paths over the surface of the insulation. Volume resistance can be used as an indirect measure of the quality of the material with regard to processing or to detect impurities. For example, indications of moisture content, degree of cure, mechanical continuity, or degradation can be obtained. Surface resistance measurements can be used to characterize the condition of the insulation surface, such as deposition of contaminants and insulation breakdown strength. The applicability of these measurements depends on the agreement of this type of data with actual field experience.

Factors affecting tests. In all cases, the conditions under which resistance measurements on insulators are performed must be fully specified. Several factors can significantly affect the test results. One of the most important of these factors is the geometry of the electrode system. With the correct selection of electrode and specimen dimensions, the surface and volume resistances can be measured separately. Temperature and humidity can change the resistance values by several orders of magnitude. To avoid this problem, the sample being tested can be conditioned prior to testing and the measurements can be made in the conditioning environment. A large error can also be introduced by stray surface currents. A guard system can be connected by forming a network of conductors which will shunt the stray currents away from the measuring circuit. The guard system is maintained at the same potential as that of the electrode to be guarded, but is insulated from it. Among the factors that may affect insulation resistance measurements are the nonuniformity of the material, the time period that the sample is energized, the magnitude and polarity of the voltage, the time required for charge to build up or decay in the sample, and the contour of the specimen.

Apparatus and methods. There is an extensive range of insulation test methods for both specimens and apparatus, and there are various accepted standards. Some typical tests are outlined below. Reproducible standard procedures should be used, and special care should be exercised in selecting the specimens and in their conditioning.

The measured value of insulation resistance will be most useful when the test samples and the electrodes have the same arrangement and are tested in a similar environment as that expected under

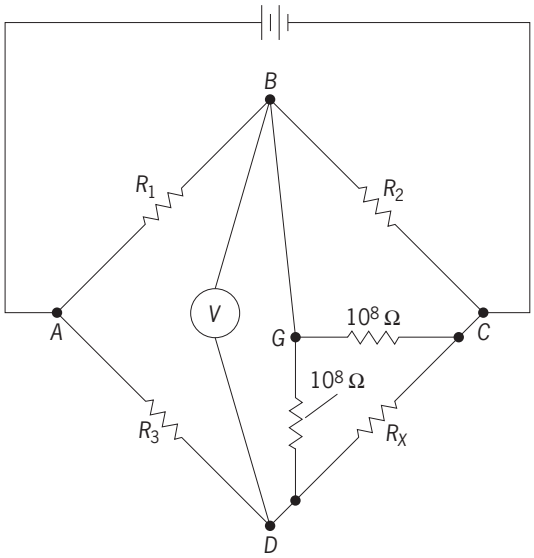

Fig. 2. High-resistance measurement with a guarded Wheatstone bridge.

actual use. A very stable direct current voltage source such as batteries is required. Instrumentation with a sensitivity of ±10% is generally considered adequate. Current-measuring devices commonly employed include galvanometers, direct-current (dc) amplifiers, and electrometers. Modern electronic instrumentation has simplified and improved insulation resistance measurements because of the high input impedance available with these devices, their fast response time, their memory storage capability, and their ease of operation. *See* ELECTROMETER; GALVANOMETER.

Direct deflection method. **Figure 1** shows a simplified version of the circuit used in this method with and without the application of a guarded terminal at A. The stray or leakage current is represented by I_L, and the insulation sample's resistance is R_X. The guard conductor surrounds terminal A and is connected to the battery side of the meter, so that no leakage current will flow through the meter, thus eliminating this source of error. Normally, the meter used is a sensitive galvanometer with a high-precision decade box (Ayrton shunt) in parallel to protect it and to provide several scale ranges. A switch can also be added to discharge a capacitive-type specimen after the test. The readings obtained are a measure of the volume resistance.

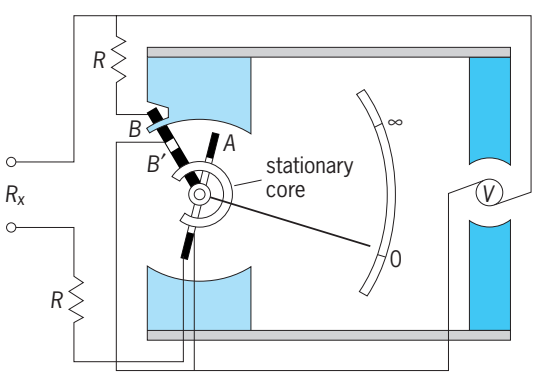

Fig. 3. Crank-driven Meggar circuit.

Loss-of-charge method. Insulation resistance is measured in this method by discharging a capacitor C through the sample to be measured. The charge Q remaining on the capacitor at the end of a specified time is measured with an electrometer or a galvanometer. Repeated measurements are necessary, making the procedure lengthy. Based on the voltage relationship $E = Q/C$, the value of resistance is found by plotting log E versus time. *See* CAPACITANCE; TIME CONSTANT.

Bridge method. A Wheatstone bridge circuit is utilized as shown in **Fig. 2**. The guard circuit in this case is made up by the two 10^8-ohm resistors connected to point G. The resistor from G to D shunts the detector, and the resistor from G to C shunts the BC arm, reducing stray current errors. *See* WHEATSTONE BRIDGE.

Meggar method. The Meggar is a megohmmeter that is widely used to measure the insulation resistance of rotating machinery, line insulators, bushings, and other equipment in the factory and on the field. Periodic measurements can provide valuable information to anticipate a serious insulation fault. The tests are made with an applied high voltage on the insulation of up to 5000 V direct current from a hand-driven generator or a rectified alternating-current (ac) supply. The Meggar consists of a permanent magnet crossed-coil meter (**Fig. 3**). A current coil A is connected in series with the unknown insulation resistance R_X. B and B' are the potential coils connected to the generator and in series with the resistor R.

All coils are mounted on the pointer mechanism and move in the field set up by the generator's permanent magnets. The stationary ring core is mounted so that coils B' and A rotate about it. When the generator is cranked, B and B' will be energized. If A has no current through it, the pointer and coils mechanism will move to where the minimum flux will pass through them. This position is opposite the open portion or slot of the core, and the pointer will indicate infinity on the scale. If current flows in A, a torque will result and the three coils will move away from the slot of the core. Coils B and B' produce opposing torques to this movement until a balance is reached. The balance point is determined by the current through A, which is directly dependent on the insulation resistance being tested. The reading is not dependent on the generator's voltage. *See* RESISTANCE MEASUREMENT.

T. C. Cheng

Bibliography. American Society for Testing and Materials, *Annual Book of ASTM Standards*, 1990; A. D. Helfrick and W. D. Cooper, *Modern Electronic Instrumentation and Measurement Techniques*, 1990; Institute of Electrical and Electronic Engineers, *Standard No. 118: Test Code for Resistance Measurements*, 1978; L. Schnell (ed.), *Technology of Electrical Measurements*, 1993; L. Thompson, *Electrical Measurements and Calibration*, 2d ed., 1994.

Insulin

One of the hormones produced by the beta cells of the pancreas, which are located in the islets of Langerhans.

Biosynthesis. Insulin participates in carbohydrate and fat metabolism. Along with insulinlike growth factors (IGF1 and IGF2), it regulates a variety of metabolic and growth-related effects in target tissues, including stimulation of glucose transport, glycogen synthesis, lipogenesis, gene transcription, and deoxyribonucleic acid (DNA) synthesis. Moreover, the insulin/IGF signaling system promotes cell growth and survival, and is essential for normal reproductive capacity. When insulin becomes unavailable or fails to function, diabetes ensues. *See* CARBOHYDRATE METABOLISM; HORMONE; LIPID METABOLISM; PANCREAS.

Human insulin is composed of 51 amino acids with a molecular mass of 6000. It is synthesized from the proinsulin gene located on human chromosome 11 (11p15.5). The proinsulin gene encodes a protein called pre-proinsulin, which is targeted to the endoplasmic reticulum where the amino-terminal signal peptide (referred to in **Fig. 1** as the signal sequence) is removed to produce proinsulin. Proinsulin displays only 10% of the biological activity of insulin. The A-chain and the B-chain that form insulin are produced by removal of the C-peptide (residues 33–63) from the middle of proinsulin (Fig. 1). The C-peptide facilitates proper folding of proinsulin so that three disulfide bonds form correctly between the A- and B-chains to stabilize the shape of insulin needed for biological activity. No biological function has been assigned to the C-peptide, although it is packaged and secreted with insulin in equimolar amounts. *See* GENE ACTION.

Physiological activity. The consumption of a carbohydrate meal causes an immediate rise in blood glucose concentration from 80 mg/dl to 130 mg/dl. Immediate insulin secretion then lowers blood glucose to the original level within 90 min. In healthy people, the blood contains about 0.2–1 ng/ml insulin before a meal and 1–5 ng/ml after the ingestion of a carbohydrate meal. Antibodies that specifically bind to insulin are used to determine precisely the concentration of insulin in blood. Similar methods measure the concentration of C-peptide in blood, which is used as an indicator that insulin is being produced in the beta cells, especially in people receiving insulin injections.

The blood glucose level is an important signal for insulin secretion. Glucose enters the pancreatic beta cells and increases the production of adenosine triphosphate (ATP), which directly stimulates insulin secretion. Other hormones released from cells in the small intestine during a meal travel to the beta cells to further promote insulin secretion. Upon secretion from the beta cells, insulin passes through the liver on its way to the peripheral tissues. Insulin-degrading enzyme is abundant in the liver, destroying about 50–60% of the secreted insulin as it passes through: the "half-life" of insulin in the body has been estimated to be 10–30 min. Insulin inhibits the production of glucose by the liver and increases the uptake of glucose by muscle and adipose tissues. Depending upon the tissue type, the glucose is used for energy (ATP production) or stored as glycogen or fat. *See* GLUCOSE.

Diabetes. The role of pancreatic secretions to control blood sugar levels was demonstrated in 1889. Insulin purified from the pancreas was first used to treat diabetic patients in 1921. Over the following decades, clinicians and scientists revealed the system-wide effects of insulin in liver, muscle, and adipose tissues. By the middle 1930s, the widespread use of insulin revealed important differences between insulin-dependent (type 1) and non-insulin-dependent (type 2) diabetes. Antibodies against insulin were used in the 1950s to confirm that hyperinsulinemia—elevated insulin in the blood—is characteristic of type 2 diabetes. Cell surface insulin receptors were first identified in the 1970s, and their tyrosine kinase activity was discovered 10 years later. Remarkably, this steady scientific progress has not stemmed the worldwide diabetes epidemic that will take a huge toll in premature morbidity and mortality in the twenty-first century.

The inability of beta cells to detect glucose or secrete enough insulin causes diabetes mellitus. Diabetes mellitus is a complex disorder that arises from various causes, including dysregulated glucose sensing or insulin secretion [maturity-onset diabetes of youth (MODY)], autoimmune-mediated β-cell destruction (type 1), or insufficient compensation for peripheral insulin resistance (type 2). MODY is a rare and usually mild form of diabetes caused by one of many well-defined genetic mutations. By contrast, the complete lack of insulin production that occurs in type 1 diabetes and in the late stages of type 2 diabetes is catastrophic. The most obvious problem is a 5- to 10-fold elevated blood glucose level, which causes life-threatening symptoms of diabetes mellitus—polydipsia (excessive thirst), polyphagia (excessive hunger), and polyuria (excessive urine production). The complete absence of insulin is accompanied by the breakdown of triglyceride into fatty acids. Fatty acids are useful fuel materials, but when present in large amounts they form beta-ketoacids in the liver, which acidifies the blood resulting in coma. In addition to an increased fat breakdown, an increase in the breakdown of tissue proteins occurs. Insulin injection reverses all of these acute problems by favoring storage and swift intake of glucose into the tissues, by decreasing the breakdown of stored fat, and by promoting protein synthesis. Optimized insulin treatment also reduces the occurrence of long-term complications that impair eye and kidney function and damages nerves and blood vessels.

Type 2 diabetes is the more prevalent form of the disease, affecting 18 million people in the United States alone. It develops more slowly and usually occurs in middle age, but is appearing with greater frequency in children and adolescents. Insulin resistance in muscle, liver, and adipose tissues is always associated with type 2 diabetes, and

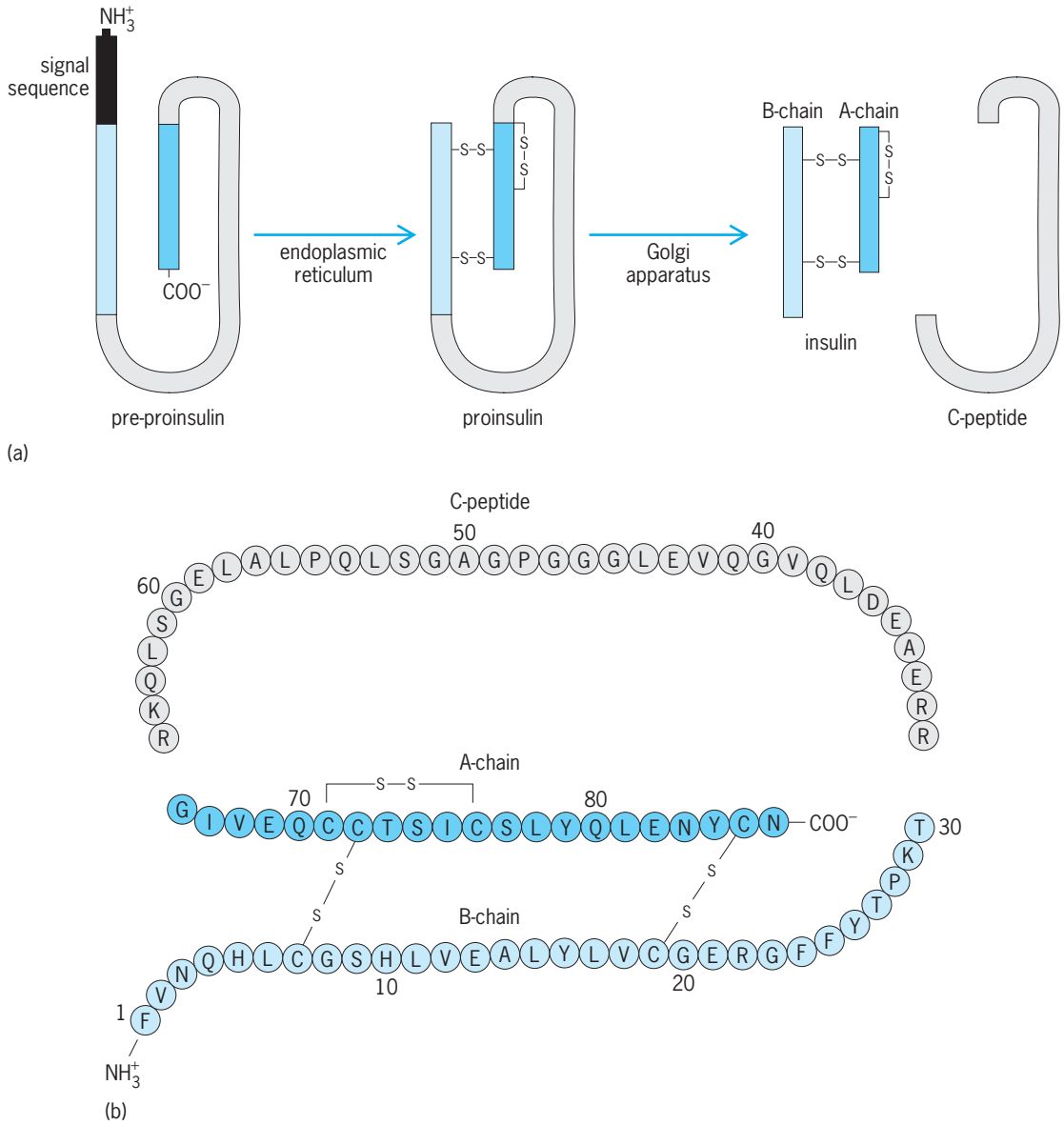

Fig. 1. Insulin biosynthesis. (*a*) Schematic diagram. Pre-proinsulin is synthesized from the insulin gene and directed into the endoplasmic reticulum by the signal sequence, where it folds into a proper conformation that is stabilized by three disulfide bonds (—S—S—). The signal sequence is removed, and proinsulin is further processed in the Golgi apparatus, where the C-peptide is removed and packaged with insulin for secretion. (*b*) Primary amino acid sequence of human insulin and the C-peptide.

hyperinsulinemia prevents gross decompensation of glucose homeostasis in the early stages of type 2 diabetes. However, insulin resistance is associated with a cohort of systemic disorders—dyslipidemia, hypertension, cardiovascular disease, female infertility and neurodegeneration—which usually develop before the disease is diagnosed. *See* DIABETES.

Commercial production. Human insulin was the first commercially available protein produced by recombinant DNA technology. The production is accomplished by inserting the gene for human insulin or proinsulin into a microorganism such as yeast or bacteria. While the microorganism grows, it produces the recombinant protein, which is purified and processed into biologically active insulin for injection. Treatment of diabetes with human insulin avoids the problems created by using insulin purified from animal sources, including the production of neutralizing antibodies.

The chemical characteristics of human insulin are exploited to produce insulin preparations that suit the treatment requirements of individual patients. Recombinant human insulin forms dimers in solution due to hydrogen bonding. Insulin dimers diffuse into the blood, where they separate into monomers that are biologically available 30–45 min after injection. The addition of zinc to insulin promotes the formation of insulin hexamers that require extra time to dissolve and form biologically active monomers: Insulin hexamers display a slower onset but a longer duration of activity (up to 24 h), resulting in less frequent need for insulin injections. The duration

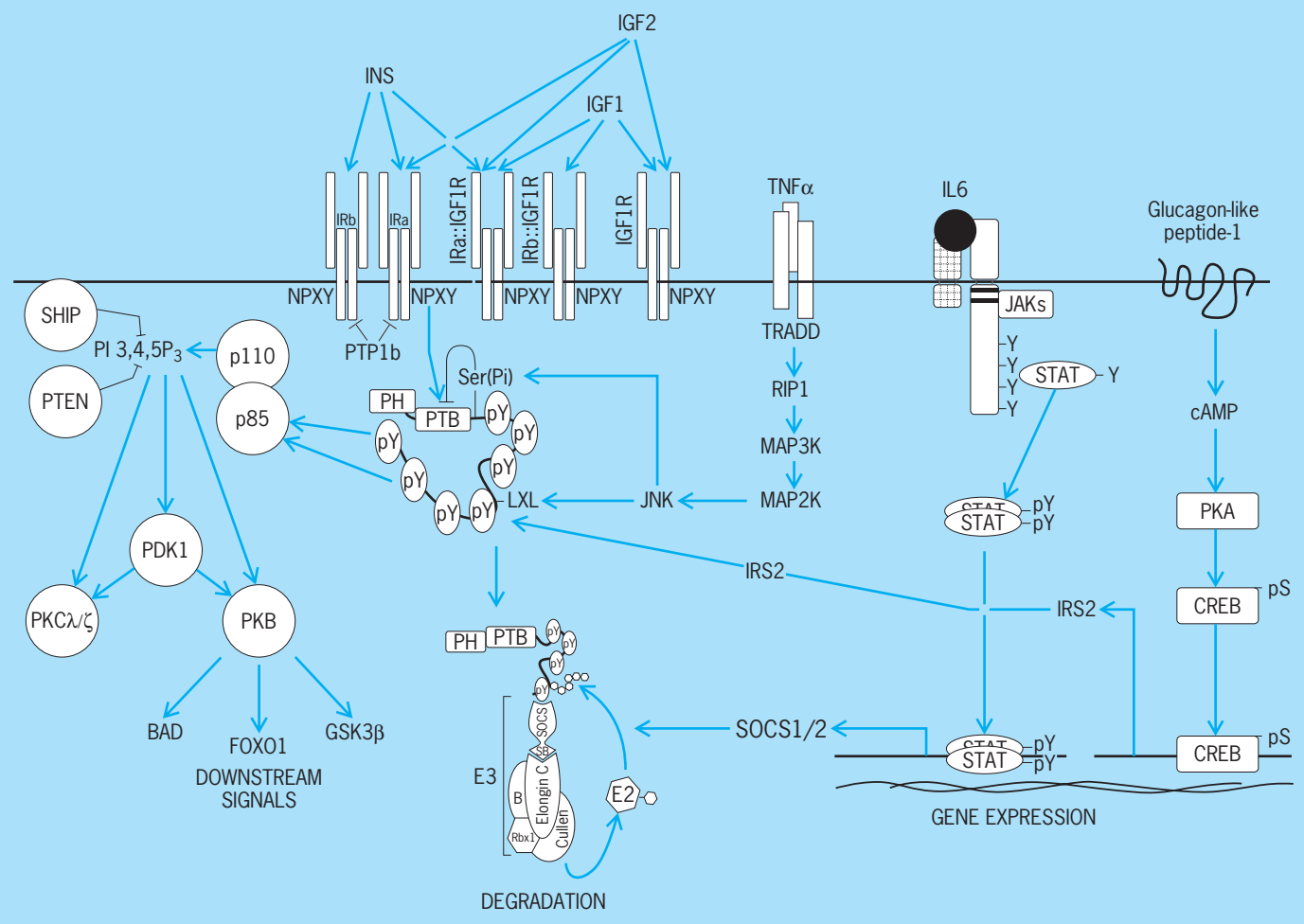

Fig. 2. Regulation of the insulin/IGF signaling pathway. The diagram shows the relation between the insulin receptor and its principal substrates, including some of their downstream elements and mechanisms of inhibition. The IRb predominates in adult tissues and binds insulin with highest affinity (<1 nM); IRa predominates during fetal life and in adult brain, binds insulin with lower affinity (~1 nM), but has a similar affinity for IGF2. The homologous IGF1 receptor binds IGF1/2 with high affinity (<1 nM). Hybrid receptors display intermediate binding properties: IRb::IGF1R binds IGF1 better than insulin, whereas IRa::IGF1R binds insulin and IGF1/2 with approximately equal affinity. The receptor β-subunits are tyrosine-phosphorylated during ligand binding and recruit substrates for phosphorylation, including IRS1/2 through an interaction between the phosphorylated NPXY motif on the receptor and the PTB domain. This interaction is inhibited by JNK-mediated phosphorylation of Ser[312], which causes insulin resistance. SOCS1/2 is an adapter protein that recruits IRS1/2 to an elongin B/C-based ubiquitin ligase, which promotes ubiquitination and degradation of the IRS-protein; proinflammatory cytokines like IL6 upregulate SOCS1/2. While proinflammatory cytokines inhibit IRS-protein signaling, cAMP agonists upregulate IRS2 through the activity of phosphorylated CREB. Tyrosine phosphorylation of IRS1/2 recruits and activates various SH2-domain-containing proteins, including the 85-kilodalton regulatory subunit of the PI 3-kinase. Activation of PI 3-kinase during association with IRS1/2 strongly stimulates the PKB cascades that phosphorylate various substrates, including BAD, GSK3β, and Foxo1.

of action of all forms of insulin is dependent on dose, site of injection, blood supply, temperature, and physical activity. Excess injected insulin causes severe hypoglycemia—blood sugar below its normal value—which damages neurons in the brain. For this reason, insulin should be injected subcutaneously and never intravenously or intramuscularly.

"Designer" insulin molecules with modified amino acid sequences can be produced by selectively changing the DNA sequence of the gene used in the process. These modifications alter the chemical behavior of insulin. For example, switching the order of the lysine and proline residues at the C-terminal end of the B-chain minimizes the formation of dimers and hexamers (Fig. 1). This designer insulin, called lyspro, displays almost immediate biological activity with short duration. It may be injected during a meal to prevent an excessive increase in plasma glucose after eating in certain patient populations. *See* GENETIC ENGINEERING.

Insulin receptors. Insulin action is initiated in cells by the binding of insulin to its cell receptor (**Fig. 2**). The insulin receptor is the prototype for a family of homologous integral membrane proteins composed of an extracellular insulin-binding domain that controls the activity of an intracellular tyrosine kinase. The 150-kilobase gene on chromosome 19, composed of 22 exons, encodes the human receptor precursor. During translation, two pro-receptors form disulfide-linked dimers that are glycosylated and cleaved to form a heterotetramer of two extracellular α-subunits and two transmembrane β-subunits.

Insulin binds to the juxtaposed α-subunits, facilitating ATP binding and tyrosine autophosphorylation in the β-subunit, which activates the kinase and recruits cellular substrates: the first steps of signal transduction.

Selective insulin binding is complicated by the production of two types of insulin receptors (IRa and IRb) by tissue-specific alternative splicing of exon 11. The type b insulin receptor that lacks exon 11 is mainly expressed in adult peripheral tissues (most tissues except the brain), where it selectively binds insulin with high affinity. The type a isoform that includes exon 11 is expressed in fetal tissues and the adult brain, where it binds insulin or insulinlike growth factor-2 (IGF2). Dysregulated splicing alters fetal growth patterns and contributes to rare forms of insulin resistance in adults. Homologous IGF1 receptors selectively bind IGF1 or IGF2, and are essential during development for about 50% of brain and body growth. However, hybrid receptors composed of an $\alpha\beta$-dimer of the IGF1 receptors and the IRb selectively bind IGF1. By contrast, hybrids between the IGF1R and the IRa bind all three ligands with similar affinities (Fig. 2).

Signaling. Most if not all insulin signals are produced or modulated through tyrosine phosphorylation of IRS1 or IRS2 (insulin receptor substrate 1 or 2). IRS-proteins are not enzymes but are composed of multiple interaction domains and phosphorylation motifs that generate regulatory signals. All IRS-proteins contain an NH_2-terminal pleckstrin homology (PH) domain (a domain of about 100 residues involved in intracellular signaling) adjacent to a phosphotyrosine-binding (PTB) domain, followed by a variable-length COOH-terminal tail with numerous tyrosine and serine phosphorylation sites. The PH and PTB domains couple the IRS-proteins to the activated insulin receptors. Tyrosine phosphorylation sites in the COOH-terminal part of IRS1/2 bind common effector proteins, including enzymes (PI 3-kinase, SHP2, and Fyn) or adapters (Grb-2, nck, crk, Shb, and others). Through these generic signaling proteins, IRS1/2 controls gene expression needed for carbohydrate and lipid metabolism, cell growth and survival, and the function of specialized neuroendocrine tissues that promote nutrient homeostasis and female fertility.

Tyrosine phosphorylation of IRS1/2 leads to the stimulation of the phosphatidylinositol 3-kinase (PI-3), which is composed of p85 and p110 subunits. PI-3 generates phosphatidylinositol 3,4,5-trisphosphate (PI 3,4,5P_3) in the plasma membrane, which stimulates the activity of certain serine kinases in cells. One of these serine kinases, called protein kinase B (PKB), is especially important because it promotes the phosphorylation of other proteins, including BAD (important for cell survival), GSK3β (cell growth and glycogen synthesis), and Foxo1 (regulation of gene expression) [Fig. 2].

Resistance. Since insulin resistance is an underlying cause of metabolic dysregulation and diabetes, understanding its molecular basis is an important goal. Genetic mutations are obvious sources of lifelong insulin resistance, but they are difficult to identify and usually associated with rare metabolic disorders. The idea that inflammation is associated with insulin resistance has been known for a long time, which provides a framework to understand how diet, acute or chronic stress, and obesity might cause insulin resistance. Familiar proinflammatory cytokines, including interleukin-6 (IL6) and tumor necrosis factor alpha (TNFα), ordinarily secreted from leukocytes during inflammation or from adipocytes in obese people, promote insulin resistance. These cytokines stimulate serine phosphorylation of IRS1/2, which inhibit their ability to undergo tyrosine phosphorylation. These cytokines also promote ubiquitination of IRS1/2, which targets them for degradation. Insulin resistance is also caused by protein or lipid phosphatases, including PTP1B, SHIP2, or PTEN. PTP1B is especially interesting because it resides in the endoplasmic reticulum, where it dephosphorylates the insulin receptor. This mechanism of action appears to confer selectivity toward the insulin receptor, so inhibitors of PTP1B might be least likely to cause unregulated cell growth.

Peripheral insulin resistance alone rarely causes diabetes, as β-cell failure is an essential feature of the disease. Why β-cells frequently fail to compensate for insulin resistance was difficult to determine until the IRS2-branch of the insulin/IGF signaling system was found to be essential for β-cell growth, function, and survival. Transgenic mouse experiments clearly show that upregulation of IRS2 potentates compensatory insulin secretion in obese mice and prevents β-cell destruction. Recent work shows that the IRS2 gene in mice is strongly upregulated by cAMP agonists, including glucose and glucagon-like peptide-1 in β-cells, through pathways that activate CREB (Fig. 2). While many cAMP-mediated pathways oppose the action of insulin, the upregulation of IRS2 reveals an unexpected intersection of these important signals that prepares cells to respond during the next round of insulin secretion (Fig. 2). IRS2 upregulation might play a role in the beneficial health effects of many drugs that increase cellular cAMP levels. The positive effects of glucagon-like peptide-1 on β-cell function, peripheral insulin action, and neuronal function might be mediated in part by upregulation of IRS2.

Diabetes is a serious illness, but only one of the consequences of insulin resistance. The clinical effects of insulin resistance can be expected to increase, as the people become heavier and less fit. Whether better management of inflammatory responses can attenuate insulin resistance and diminish its consequences is an important area of investigation. Understanding how the IRS-proteins negotiate the conflicting signals generated during insulin and cytokine action might be a valuable starting point. However, too much insulin action might shorten our lives, so future work must better resolve the network of insulin responses that are generated in various tissues, and pinpoint the ones that prolong health as well as the ones that diminish it.

Morris F. White

Bibliography. M. Bliss, *The Discovery of Insulin*, University of Chicago Press, 2003; M. F. White, Insulin signaling in health and disease, *Science*, 302:1710–1711, 2003; P. Zimmet, K. G. Alberti and J. Shaw, Global and societal implications of the diabetes epidemic, *Nature*, 414:782–787, 2001.

Intaglio (gemology)

The name given to the type of carved gemstone in which the figure is engraved into the surface of the stone, rather than left in relief by cutting away the background, as in a cameo. Intaglios are almost as old as recorded history, for this type of carving was popular in ancient Egypt in the form of cylinders. The cylinder, as well as the more familiar form of intaglio, was popular for impressing seals on sealing wax. Intaglios have been carved in a variety of gem materials, including emerald, crystalline quartz, hematite, and the various forms of chalcedony. *See* CAMEO; GEM. Richard T. Liddicoat, Jr.

Integral equation

An equation of the form typified by Eq. (1). The major

$$\int_a^b K\left(x, y, \phi(y)\right) dy + f(x) = a(x)\phi(x) \qquad (1)$$

problem is to decide when there is a function $\Phi(x)$ which is a solution to the equation. Equations such as (1) arise from the analysis of ordinary differential equations. For example, it is easy to verify that the function $\Phi(t)$ is a solution of differential equation (2) if and only if Φ is a solution of integral equation (3).

$$\frac{d\phi}{dt} = F(t, \phi) \qquad \phi(0) = A \qquad (2)$$

$$\phi(t) = A + \int_0^t F(s, \phi(s))ds \qquad (3)$$

This equation is the same type as Eq. (1), with the function K given by Eq. (4).

$$K(t, s, \phi) = \begin{cases} F(s, \phi) & 0 \le s \le t \\ 0 & t < s \end{cases} \qquad (4)$$

Integral equation (1) is an equation of the first kind if $a(x) \equiv 1$, and an equation of the second kind if $a(x) \equiv 0$.

The domain of integration, which in Eq. (1) is the interval $[a, b]$, can be a domain in two or more variables, a curve, or even a surface. The unknown function Φ and the kernel K can be vector-valued functions, in which case Eq. (1) is called a system of integral equations.

Volterra equations. If $K(x, y, \Phi) = 0$ for $x \leqslant y \leqslant b$, Eq. (1) is called a Volterra equation. Thus Eq. (3) is an example of a Volterra equation. Under mild assumptions about the smoothness of $K(x, y, \Phi)$, for example, if K is continuous and if $|K(x, y, \Phi) - K(x, y, \Psi)| \leqslant A|\Phi - \Psi|$, where A is a constant, for all x, y in the domain of integration and for all Φ and Ψ—then the Volterra equation of the second kind, Eq. (5), can be

$$\phi(x) = \int_a^x K\left(x, y, \phi(y)\right) dy + f(x) \qquad (5)$$

solved by the method of successive approximations (also known as the method of Picard). The idea is to define a sequence of functions $\Phi_0, \Phi_1, \ldots$ according to the scheme of Eqs. (6). Under the assumed

$$\begin{aligned} \Phi_0(x) &= f(x) \\ \Phi_1(x) &= \int_0^x K\left(x, y, \Phi_0(y)\right) dy = f(x) \\ &\ldots \\ \Phi_{n+1}(x) &= \int_0^x K\left(x, y, \Phi_n(y)\right) dy + f(x) \end{aligned} \qquad (6)$$

conditions, it can be shown that this sequence converges to a function $\Phi(x)$ which is the only solution to Eq. (5). The method of Picard is fundamental to the study of integral equations, but occurs also in the study of many other functional equations.

Linear integral equations. A special case of some interest is that of linear integral equations. The function $K(x, y, \Phi)$ is a linear function of Φ. The linear equations of the first and second kind are shown in Eqs. (7) and (8) respectively. The function

$$f(x) = \int_a^b K(x, y)\phi(x)dy \qquad (7)$$

$$\phi(x) = f(x) + \lambda \int_a^b K(x, y)\phi(y)dy \qquad (8)$$

$K(x, y)$ is called the kernel, and the complex number λ is called the parameter. The equation of the second kind, Eq. (8), is homogeneous if $f(x) \equiv 0$. In typical cases, the homogeneous equations will have only the trivial solution $\Phi(x) \equiv 0$. For some values of the parameter λ, however, there will be nontrivial solutions. Such a value of the parameter λ is called a characteristic value for K, and the corresponding function Φ is called a characteristic function for K. These concepts are related to corresponding concepts in linear algebra and operator theory. The operator L_K acting on the function Φ is defined by Eq. (9). Then L_K is an example of a linear operator,

$$L_K\phi(x) = \int_a^b K(x, y)\phi(y)dy \qquad (9)$$

that is, if Φ and Ψ are functions and α and β are complex constants, then $L_K(\alpha\Phi + \beta\Psi) = \alpha L_K\Phi + \beta L_K\Psi$. If the function Φ is a characteristic function for K with corresponding characteristic value λ, then the homogeneous equation can be rewritten as Eq. (10).

$$L_K\phi = \frac{1}{\lambda}\phi \qquad (10)$$

In terms of linear operator theory, Eq. (10) means that $1/\lambda$ is an eigenvalue (proper value) for the linear operator L_K, and Φ is a corresponding eigenvector (proper vector), or in this case, eigenfunction.

A linear Volterra equation has no characteristic values. *See* EIGENFUNCTION; MATRIX CALCULUS; MATRIX THEORY; OPERATOR THEORY.

Fredholm equations. The integral equation, Eq. (7) or (8), is called a Fredholm equation, and $K(x, y)$ is called a Fredholm kernel if $\|K\| < \infty$, where $\|K\|$, the norm of K, is defined by Eq. (11). This is certainly

$$\|K\|^2 = \int_a^b \int_a^b |K(x, y)|^2 \, dx \, dy \tag{11}$$

true if K is continuous or even bounded. Even some infinite discontinuities are allowable: for example, if $|K(x,y)| \leq M|x - y|^{-\alpha}$, where M is a constant and $\alpha < {}^1/_2$, then K is a Fredholm kernel.

Fredholm equations of the second kind for which the parameter satisfies the inequality $|\lambda| \, \|K\| < 1$ can be shown to have unique solutions by the method of Picard. The result is Eq. (12) where the resolvent $R_K(x, y; \lambda)$ is given by the Neumann series, Eqs. (13) and (14). Equation (12) can be rewritten in terms

$$\phi(x) = f(x) + \lambda \int R_K(x, y; \lambda) f(y) dy \tag{12}$$

$$R_K(x, y; \lambda) = \sum_{n=1}^{\infty} \lambda^{n-1} K^{(n)}(x, y) \qquad K^1(x, y) = K(x, y) \tag{13}$$

$$K^{(n+1)}(x, y) = \int_a^b K(x, z) K^{(n)}(z, y) \, dz \tag{14}$$

$$\phi = \sum_{n=0}^{\infty} \lambda^n L_K^n f \tag{15}$$

of the operator L_K as Eq.(15), when $L^2{}_K f = L_K(L_K f)$, $\ldots, L^{n+1}{}_K(f) = L_K(L^n{}_K f)$.

As a result there are no characteristic values for which $|\lambda| \leqslant 1/\|K\|$. In general, for any constant $M > 0$ there are only finitely many characteristic values which satisfy $|\lambda| \leqslant M$. Further, only finitely many linearly independent characteristic functions are associated with each characteristic value.

For the kernel K the conjugate kernel is defined to be $K^*(x, y) = \overline{K(y, x)}$, where the bar denotes complex conjugation. It can be shown that λ is a characteristic value for the kernel K if and only if $\overline{\lambda}$ is a characteristic function for the conjugate kernel K^*. Furthermore, the number of linearly independent corresponding characteristic functions is the same for λ and K as for $\overline{\lambda}$ and K^*. Finally, inhomogeneous equation (16) has a solution if and only if f satisfies Eq. (17) for every solution ψ of conjugate homogeneous equation (18).

$$\phi(x) = f(x) + \lambda \int_a^b K(x, y) \phi(y) dy \tag{16}$$

$$\int_a^b f(x) \overline{\psi(x)} \, dx = 0 \tag{17}$$

$$\psi(x) = \lambda \int_a^b K^*(x, y) \psi(y) dy \tag{18}$$

These results are a statement of the Fredholm alternative which may be stated more succinctly as follows: For a given value of λ, either λ is a characteristic value and the homogeneous equation has a nontrivial solution, or λ is not a characteristic value and the inhomogeneous equation has a solution for every choice of the function f. *See* DIFFERENTIAL EQUATION; INTEGRAL TRANSFORM; INTEGRATION. John Polking

Bibliography. C. Corduneanu, *Integral Equations and Applications*, 1991; J. Kondo, *Integral Equations*, 1992; A. C. Pipkin, *A Course on Integral Equations*, 1991; D. Porter and D. S. Stirling, *Integral Equations: A Practical Treatment, from Spectral Theory to Applications*, 1991.

Integral transform

An integral relation between two classes of functions. For example, a relation such as Eq. (1) is said

$$f(x) = \int_{-\infty}^{\infty} G(x, y) \phi(y) \, dy \tag{1}$$

to define an integral transform. More generally, the integral may be a multiple integral, and the functions f, G, ϕ may depend on a larger number of variables. Equation (1) is thought of as transforming a whole class, or space, of functions $\phi(y)$ into another class of functions $f(x)$. The function $G(x, y)$ is the kernel of the transform. One of the important uses of such a transform is based on the fact that a problem posed in one of the two spaces in question may be more easily solved in the other. For example, a differential equation to be solved for the function $\phi(y)$ may become an algebraic equation for the unknown function $f(x)$. *See* LAPLACE TRANSFORM.

The two basic problems for any integral transform are inversion and representation. In inversion the aim is to recover $\phi(y)$ from $f(x)$, the kernel $G(x,y)$ being known. That is, Eq. (1) is thought of as an integral equation (of the first kind) to be solved for the unknown function $\phi(y)$. A means of calculating $\phi(y)$ from $f(x)$ is called an inversion formula, and in its presence the transform achieves maximum utility, since explicit passage from each space to the other is thus assured. In representation the question is which functions $f(x)$ may be written or represented in the form (1). That is, one asks which functions $f(x)$ will make Eq. (1) solvable for $\phi(y)$. Usually this problem becomes more tractable when the solutions $\phi(y)$ are restricted to some subspace such as the class of positive or bounded functions.

Inversion theory. In many cases where the inversion problem has been solved, an inversion operator, such as a differential or integral operator, has been found, depending perhaps on a parameter, which accomplishes the inversion. It may be denoted by $O_t[f(x)]$. This means that for each value of the parameter t some operation O on $f(x)$ is defined. For example, t might be 3 and the operation might consist of computing the third derivative of $f(x)$. Suppose

that O_t applied to Eq. (1) produces Eq. (2), where

$$O_t[f(x)] = \int_{-\infty}^{\infty} G_t(x, y)\phi(y)\, dy \tag{2}$$

$G_t(x, y)$ is a new kernel for each t. In favorable cases the integral (2) approaches $\phi(x)$ as t approaches a suitable limit, for example, when G_t approaches the Dirac δ-function $\delta(x - y)$.

An important special case of Eq. (1) is the convolution transform, when the kernel is a function of $(x - y)$, Eq. (3). An equivalent form of Eq. (3) is Eq. (4), since the change of variable $x = e^t$, $y = e^{-u}$

$$f(x) = \int_{-\infty}^{\infty} G(x - y)\phi(y)\, dy \tag{3}$$

$$F(x) = \int_{0}^{\infty} K(xy)\Phi(y)\, dy \tag{4}$$

carries Eq. (4) into Eq. (3) after a suitable change in notation.

Named transforms. Many of the classical transforms have been named for the original or principal investigator. A few of the more important ones can be listed with the name ordinarily attached to them.

For Eq. (1):		
A. Bilateral Laplace	$G(x, y) = e^{-xy}$	
B. Fourier	$G(x, y) = e^{ixy}$	
C. Mellin	$G(x, y) = y^{x-1}$	$y > 0$
	$= 0$	$y < 0$
D. Stieltjes	$G(xy) = (x + 1)^{-1}$	$y > 0$
	$= 0$	$y < 0$
For Eq. (3):		
E. Dirichlet	$G(x) = x^{-1} \sin x$	
F. Fejér	$G(x) = x^{-2}(\sin x)^2$	
G. Fractional (Weyl form	$G(x) = (-x)^{\mu-1}$	$x < 0$
	$= 0$	$x > 0$
H. Hilbert	$G(x) = (x)^{-1}$	
I. Picard	$G(x) = e^{-\lvert x\rvert}$	
J. Poisson	$G(x) = (1 + x^2)^{-1}$	
K. Laplace (unilateral)	$G(x) = e^x e^{-e^x}$	
L. Stieltjes	$G(x) = \operatorname{sech}(x/2)$	
M. Weierstrass or Gauss	$G(x) = e^{-x^2}$	
For Eq. (4):		
N. Fourier cosine	$K(x) = \cos x$	
O. Fourier sine	$K(x) = \sin x$	
P. Hankel	$K(x) = \sqrt{x} J_v(x)$	
Q. Laplace (unilateral	$K(x) = e^{-x}$	
R. Meijer	$K(x) = \sqrt{x} K_v(x)$	

Here $J_v(x)$ is a Bessel function and $K_v(x)$ a modified Bessel function. Certain transforms are listed twice because of their frequent appearance in two equivalent forms. In fact, all of the above are convolution transforms except A, B, and C. Also A and C are the same transform (set $y = e^{-t}$ in C).

Convolution inversion. The inversion theory for a large class of convolution transforms has been completed. Operational calculus is a guide to the method. If the symbol D stands for differentiation with respect to x but is nonetheless treated as a number in the familiar series represented by Eq. (5), one is led to make the definition in Eq. (6), since the series (6)

$$e^{tD} = \sum_{k=0}^{\infty} \frac{t^k}{k!} D^k \tag{5}$$

$$e^{tD} f(x) = \sum_{k=0}^{\infty} \frac{t^k}{k!} D^k f(x) = f(x + t) \tag{6}$$

is precisely the Maclaurin expansion of $f(x + t)$ if D^K is allowed to mean a kth derivative. *See* OPERATOR THEORY.

Now consider the bilateral Laplace transform of $G(t)$ in Eq. (1), denoting it by $1/E(s)$, as shown in Eq. (7). Replacing s by D and using Eq. (6), one

$$\frac{1}{E(s)} = \int_{-\infty}^{\infty} e^{-st} G(t)\, dt \tag{7}$$

has Eq. (8). If $E(D)$ were a number, one could solve Eq. (8) to obtain Eq. (9). Finally, reverting to the original meaning of D as a derivative, one has in Eq. (9) an inversion of Eq. (1) by means of a differential operator $E(D)$.

$$\begin{aligned} \frac{1}{E(D)}\phi(x) &= \int_{-\infty}^{\infty} e^{-tD}\phi(x)G(t)\, dt \\ &= \int_{-\infty}^{\infty} \phi(x - t)G(t)\, dt \\ &= \int_{-\infty}^{\infty} G(x - y)\phi(y)\, dy \\ &= f(x) \qquad x - t = y \end{aligned} \tag{8}$$

$$\phi(x) = E(D)f(x) \tag{9}$$

This argument is meant to be exploratory only, but the result is accurate for a large class of kernels G and their corresponding inversion functions E. In summary, the inversion function is the reciprocal of the bilateral Laplace transform of the kernel. It has been shown that the result is correct if, for example, $E(s)$ is the infinite product in Eq. (10), where $c \geqq 0$

$$E(s) = e^{bs - cs^2} \prod_{k=1}^{\infty} \left(1 - \frac{s}{a_k}\right) e^{s/a_k} \tag{10}$$

and the series of real constants

$$\sum_{k=1}^{\infty} a_k$$

converges.

For example, if $K(x) = e^{-x}$, then Eq. (4) is the Laplace transform. Expressed as a convolution transform as in Eq. (3), it becomes Eq. (11), where G

$$e^x F(e^x) = \int_{-\infty}^{\infty} G(x - y)\Phi(e^{-y})\, dy \tag{11}$$

is given in the above list as entry K. The bilateral Laplace transform of this kernel is the familiar gamma function, Eq. (12), whose reciprocal has

$$\Gamma(1 - s) = \int_{-\infty}^{\infty} e^{-st} G(t) dt = \int_{0}^{\infty} e^{-t} t^{-s}\, dt \tag{12}$$

a well-known expansion in the form of Eq. (10). In Eq. (13) γ is Euler's constant. In the present

$$E(s) = \frac{1}{\Gamma(1-s)} = e^{-\gamma s}\prod_{k=1}^{\infty}\left(1-\frac{s}{k}\right)e^{s/k} \qquad (13)$$

example Eq. (9) becomes Eq. (14), or if $e^{-x} = t$, Eq. (15) may be written. This familiar inversion for-

$$e^{-\gamma D}\prod_{k=1}^{\infty}\left(1-\frac{D}{k}\right)e^{D/k}e^{x}F(e^{x}) = \Phi(e^{-x}) \qquad (14)$$

$$\lim_{k\to\infty}\frac{(-1)^k}{k!}F^{(k)}\left(\frac{k}{t}\right)\left(\frac{k}{t}\right)^{k+1} = \Phi(t) \qquad (15)$$

mula also serves to illustrate the operator O_t appearing in Eq. (2). In the present case the operator is a differential one, and the parameter t is an integer k which tends to ∞. *See* CONFORMAL MAPPING; INTEGRATION.

David V. Widder

Bibliography. M. I. Akhiezer, *Lectures on Integral Transforms*, 1992; L. C. Andrews and B. K. Shivamoggi, *Integral Transforms for Engineers and Applied Mathematicians*, 1987; M. Ya. Antimirov, A. A. Kolyshkin, and R. Vaillancourt, *Applied Integral Transforms*, 1993; A. J. Jerri, *Integral and Discrete Transforms with Applications and Error Analysis*, 1992.

Integrated-circuit filter

An electronic filter implemented as an integrated circuit, as contrasted with filters made by interconnecting discrete electrical components. The design of an integrated-circuit filter (also called simply an integrated filter) is constrained by the unavailability of certain types of components, such as piezoelectric resonators, that are often valuable in filtering. However, integrated filters can benefit from small size, close integration with other parts of a system, and the low cost of manufacturing very complex integrated circuits.

Mathematically, there are two broad types of filters: Continuous-time filters implement systems of linear ordinary differential equations, while discrete-time filters implement systems of difference equations. Both types can be integrated. *See* DIFFERENCE EQUATION; DIFFERENTIAL EQUATION.

Filters have many applications. A particularly important one is to smooth signal waveforms sufficiently to allow accurate sampling or to interpolate smoothly between given samples of a signal. Since the analog-to-digital converters that sample signals are usually made as integrated circuits (chips), it is often convenient to put the associated filters on the same chip. *See* ANALOG-TO-DIGITAL CONVERTER.

Passive filters, made by interconnecting inductors and capacitors, are not easily integrated because integrated-circuit inductors are usually of poor quality. They are compromised by being restricted to two-dimensional construction, by the resistance of the very thin metal available, and by proximity to a conductive substrate in which currents can be induced. These problems are less damaging at frequencies above 1 GHz, so microwave filters can be passive. This technology is most common when using gallium arsenide (GaAs) substrates, because higher speeds, insulating substrates, and high-conductivity metal are available there. These circuits are known as monolithic microwave integrated circuits (MMICs), where monolithic is a synonym for integrated. *See* MICROWAVE FILTER.

Active filters. Because amplifiers are very cheap in integrated-circuit technology, active filters are widely implemented. The five main types of active filter—active-RC, MOSFET-C, transconductance-C, switched-capacitor (or switched-C), and active-RLC—are distinguished by their frequency-sensitive components. The switched-capacitor filters operate on samples of signals, while the other types operate without sampling (in continuous time). There is also a trend toward digital filters, which are easily integrated but require that analog signals be converted to digital form, which in turn requires filtering. *See* DIGITAL FILTER.

The filter examples discussed below are second order, meaning that they implement a system of two linear differential or difference equations, or equivalently that their transfer functions can be written as ratios of polynomials of order two (quadratics). They are called biquadratic circuits because their transfer functions can be so written. Practical filters are often of higher order, but a filter of arbitrary order can be made by cascading (placing one after another) second-order circuits.

Active-RC filters. Discrete active-RC filters were widely used in the 1970s, and modern integrated filters are derived from them. The frequency-sensitive mechanism in active-RC filters is the charging of a capacitor C through a resistor R, giving a characteristic frequency $\omega_0 = 1/RC$ radians per second, at which the impedances of the resistor and capacitor are equal. Unfortunately, integrated-circuit manufacturing techniques do not control the product RC at all accurately, with variations of 20–50% being possible. This limits active-RC filtering to those applications where accuracy is unimportant, where external passive components are tolerable, or where tuning circuitry is available.

MOSFET-C filters. These filters replace the resistors of an active-RC filter with metal-oxide-semiconductor (MOS) transistors, in which a conducting channel along the surface can be enhanced or depleted by applying an electrical field from a gate electrode, thereby changing the resistance of the channel. The result is a tunable variant of an active-RC filter. A simple example (**Fig. 1**) consists of a loop of two integrators (V_2 is the integral of V_1, for example) implemented by amplifiers, which may be labeled A1 and A2, and their associated devices. An additional amplifier, A3, inverts the sign of the output of A2 to set loop gain correctly. Replacing all transistors with resistors (and ignoring the tuning control V_c) reduces this circuit to the Tow-Thomas biquad. *See* TRANSISTOR.

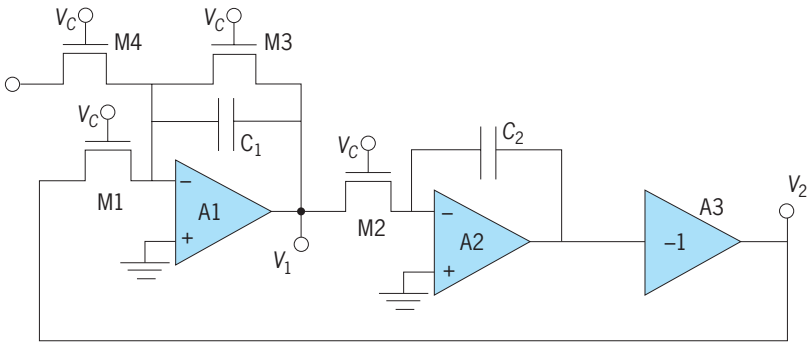

Fig. 1. MOSFET-C filter, consisting of a loop of two integrators.

If two of the transistors, M1 and M2, are equal in size, and each consists of Q parallel copies of the two other transistors in the circuit, M3 and M4, then, for a given control voltage V_c, M1 and M2 will simulate resistors of some value R and M3 and M4 will simulate resistors of value QR. Then when the two capacitors in the circuit, C_1 and C_2, are designed with $C_1 = C_2 = C$, the output of amplifier A1 will be a band-pass-filtered version of the input, with a center frequency $\omega_0 = 1/RC$, at which the input is passed unattenuated, and a bandwidth ω_0/Q between the frequencies at which the output voltage gain is $1/\sqrt{2}$. Band-pass filters are used to detect tones, while the output of amplifier A2 is a low-pass filter of the kind used to smooth signals before sampling. The control voltage V_c is generated with a separate circuit that adjusts the behavior of a reference MOSFET-C circuit with a simple known input.

MOSFETs are nonlinear resistors in that their resistances change significantly for voltage inputs of a few hundred millivolts. This behavior is partially controlled by using a differential or balanced version of the circuit discussed above, in which all signals are carried in both positive and negative polarities; this type of circuit tends to cancel quadratic nonlinearities, and also partially cancels interference from noise on the power supply and external sources. It also makes the amplifier A3 redundant, since differential signals can be inverted just by exchanging two wires.

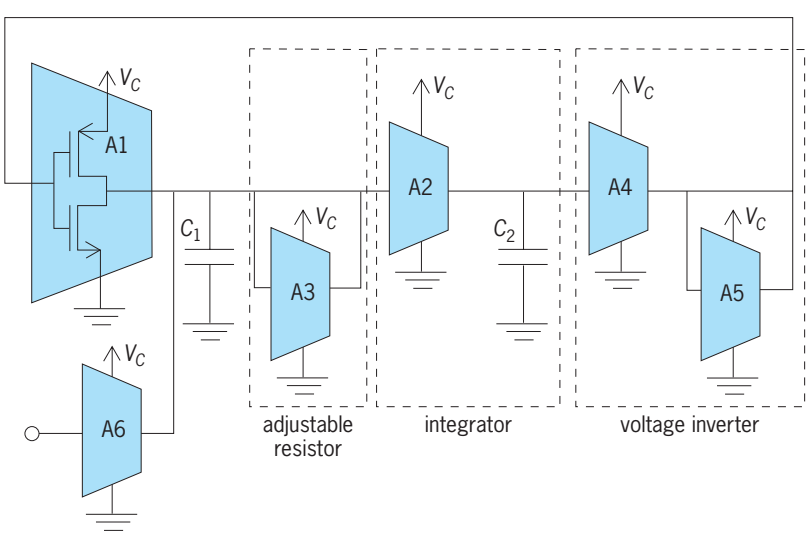

Fig. 2. Transconductance-C filter with amplifier A1 shown in detail.

Transconductance-C filters. These filters combine the functions of the amplifier and the simulated resistor into a transconductance amplifier, which is an amplifier whose output current (rather than output voltage) is proportional to its input voltage. The effect of M2 and A2 in the MOSFET-C filter (Fig. 1) is to force a current V_1/R into the feedback capacitor on A2; a transconductance amplifier with gain $1/R$ and input voltage V_1 has an output current V_1/R, so if it is connected to a capacitance C the same effect will be obtained. A simple transconductance-C filter (**Fig. 2**) can be made entirely from transconductance amplifiers (which may be labeled A1 to A6) and capacitors. Amplifiers that have their inputs and outputs connected together (A3 and A5) just act as grounded resistors, since they have a single signal terminal and it conducts a current proportional to the applied voltage. Since transconductors are usually adjustable, so is this simulated resistor. When one of these simulated resistors loads another transconductor (as A5 loads A4), an inverting amplifier is formed.

There are many circuits that approximate an ideal transconductance amplifier. One of the simplest (A1 in Fig. 2) consists of a complementary pair of n-channel and p-channel MOS transistors. Other, more practical implementations have a balanced form in which every signal is represented by a pair of equal and opposite voltages; just as for the MOSFET-C case, this reduces the effect of interference and even-order nonlinearities while eliminating the need for voltage inverter formed by A4 and A5. *See* AMPLIFIER; DIFFERENTIAL AMPLIFIER.

Transconductance amplifiers can be very simple and hence are capable of high-frequency operation (up to approximately 1 GHz) but tend to have poor linearity when designed for high speeds. The complementary MOS pair (Fig. 2) is one of the simplest and fastest but is difficult to adjust because the power-supply voltage acts as the control signal.

The transconductance-C filter can also be described in terms of gyrators, which are circuit-theoretic elements that simulate inductors at one port when loaded by capacitors at another (Fig. 2). In this view, A1, A2, A4, and A5 are a gyrator loaded by C_2, and the result is that the capacitor C_1 appears to be in parallel with a simulated inductor and a resistor (A3). *See* GYRATOR.

A technique known as active-RLC filtering combines the ideas of active filtering with the use of physical inductors (made as spirals of metallization on the top layer of the chip). In this method, amplifiers, connected to simulate negative resistors, are used to enhance the performance of the inductors, whose losses can be modeled (to a first approximation) as being caused by a parallel positive resistance.

Switched-capacitor filters. The basic Tow-Thomas loop can also be implemented in discrete time, that is, on samples of a signal (**Fig. 3**). Each of the

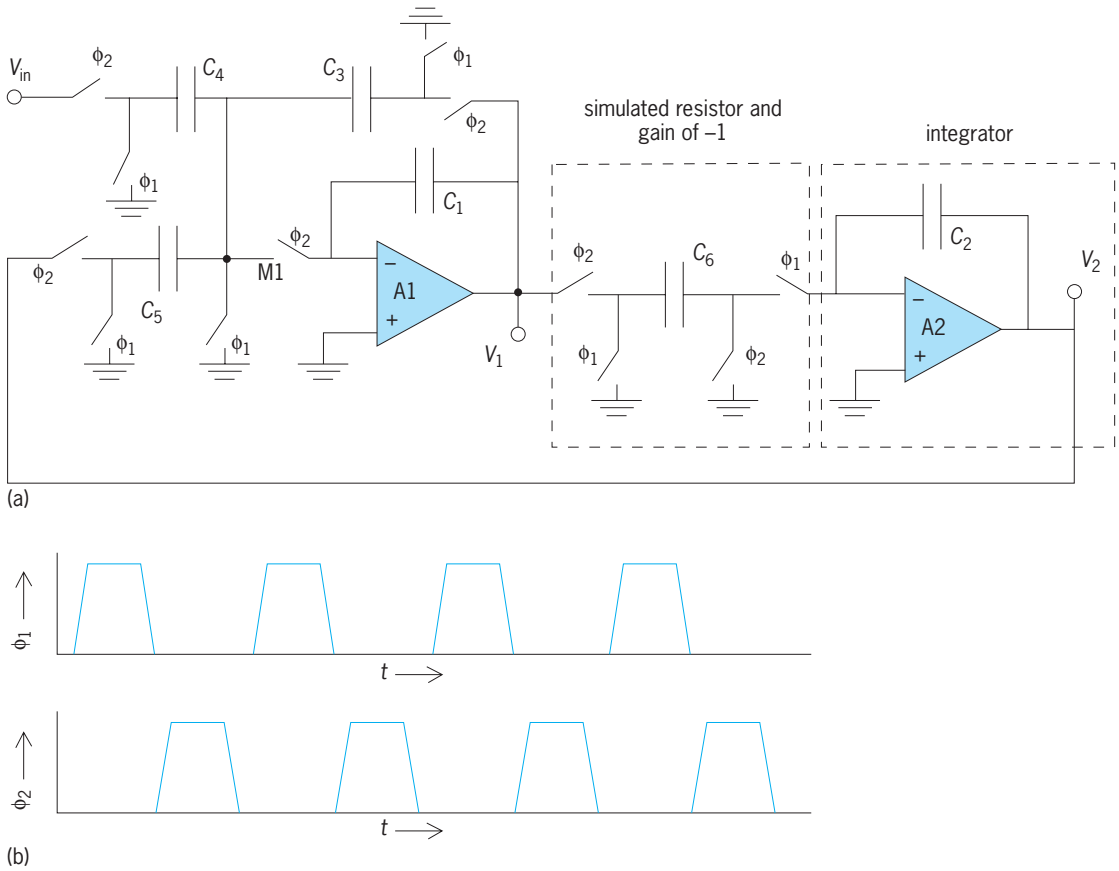

Fig. 3. Switched-capacitor filter. (*a*) Circuit diagram. (*b*) Nonoverlapping clock phases ϕ_1 and ϕ_2.

simulated resistors is now replaced with a capacitor and an arrangement of switches clocked on one or the other of two nonoverlapping phases, ϕ_1 and ϕ_2. A capacitor C_6, for example, is charged up by a voltage V_1 on each phase ϕ_2, then discharged into the input of an amplifier A2 on each phase ϕ_1. Transferring a packet of charge on every clock cycle is a way of making current flow, and the fact that the charge transferred on each cycle is proportional to voltage makes the switched capacitor C_6 behave like a resistor. Thus C_6 and its switches replace the transistor M2 in the MOSFET-C filter (Fig. 1). The switch arrangement on C_6 also inverts the capacitor between phases, thus getting the effect of the voltage inverter in the MOSFET-C filter without using an amplifier.

In a similar way, a capacitor C_5 and its switches replace the transistor M1 in the MOSFET-C filter, while two capacitors C_3 and C_4 replace the transistors M3 and M4, respectively. The switching arrangement at the input of A1 is slightly different, so that voltage inversion does not occur. A simple design technique applies the idea that a switched capacitor simulates a resistor of value $R = 1/(Cf_s)$, where f_s is the sampling (clock) frequency. Under this first-order approximation the center frequency of the filter is $\omega_0 = (C_5C_6/C_1C_2)f_s$. An exact design style is now preferred, in which the circuit is seen to implement a set of difference equations, rather than differential equations. The voltage on capacitor C_2, for example, changes by $(C_5/C_2)V_1$ from the end of one phase ϕ_2 to the end of the next. However, it is difficult to predict the precise value of the voltage V_2 at intermediate times, because there are large transient currents at the beginning of each clock phase that cause the operational amplifiers to behave nonideally. Thus switched-capacitor filters produce known outputs at discrete times (the end of phase ϕ_2, in this example), while the MOSFET-C and transconductance-C filters of (Figs. 1 and 2) provide well-defined outputs continuously (in continuous time).

The primary advantage of switched-capacitor filters is that they can be very accurate, since critical frequencies are determined by the product of a clock frequency and a ratio of capacitors (rather than a single capacitor). Switched-capacitor filters are probably the most prevalent integrated filters. Most telephone systems, for example, use them to smooth signals before sampling them for digital transmission.

A key problem for switched-capacitor filters is that the switches, implemented as MOSFETs, inject small amounts of charge into the circuit when they turn off. This creates constant offset errors, couples noise in from power supplies, and can produce nonlinear distortion. Just as for MOSFET-C and transconductance-C filters, the trend is to implement switched-capacitor filters with fully balanced circuits, in which these effects are partially cancelled. *See* SWITCHED CAPACITOR CIRCUIT.

Current-mode filters. In these filters, the signals are represented by current levels rather than voltage

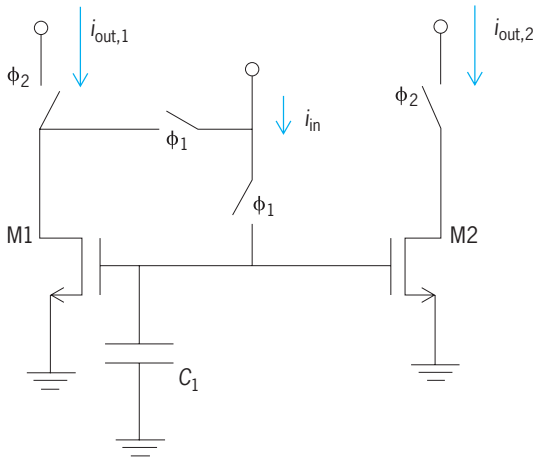

Fig. 4. Sampled-current circuit.

levels. A key building block of these filters is the current copier (**Fig. 4**). During phase ϕ_1, any difference between the input current i_{IN} and transistor drain current is fed back to correct the gate voltage of a transistor M1, so that a voltage is established on a parasitic capacitance C_1 that forces the transistor to operate at the input current. Then, during phase ϕ_2, the transistor is switched to drive this same current into the output. Placing a second transistor M2 in parallel with the copier device M1 allows for a second output that can be scaled by a constant factor (the ratio of device sizes) from the first; and signals represented by currents can be added just by connecting them together. Given the operations of scaling, addition, and copying, an arbitrary set of difference equations (and hence an arbitrary filter) can be implemented. The principal advantage of this technology over that of switched capacitors is that it can make use of the nonlinear gate-source parasitic capacitance that exists in any process, rather than needing a high-quality capacitor. A disadvantage comes from the lower accuracy of ratios set by transistor sizes. *See* ELECTRIC FILTER; INTEGRATED CIRCUITS. Martin Snelgrove

Bibliography. S. K. Mitra and C. F. Kurth (eds.), *Miniaturized and Integrated Filters*, 1989; R. Schaumann, M. S. Ghausi, and K. R. Laker, *Design of Analog Filters*, 1990; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Integrated circuits

Miniature electronic circuits produced within and upon a single semiconductor crystal, usually silicon. Integrated circuits (ICs) range in complexity from simple logic circuits and amplifiers, about 1.3 mm (0.05 in.) square, to very large scale integrated circuits (VLSI), up to about 25 mm (1 in.) square. They can contain over 10^9 transistors and other components that provide computer memory circuits and complex logic subsystems such as microcomputer central processor units (microprocessors) and digital signal processing (DSP). *See* SEMICONDUCTOR; SILICON.

Key inventions leading to the invention of the integrated circuit include the invention of the bipolar transistor by John Bardeen, Walter Brattain, and William Shockley in 1947; the invention of junction isolation by Kurt Lehovec; and the invention of silicon planar technology by Jean Horni.

Jack Kilby was the first to demonstrate an integrated circuit in 1958, using germanium and attaching small wires to connect transistors, diodes, and capacitors; however, it was the approach by Robert Noyce, which included patterned aluminum metal interconnections on silicon (**Fig. 1**), that made the integrated circuit revolution possible. Noyce may be considered the leader of this revolution, not only for his role in inventing the integrated circuit but also for his role in promoting integrated circuit development and applications.

Since the mid-1960s, integrated circuits have become the primary components of most electronic systems, and they are used in almost every piece of hardware that has some electronics. Their low cost, high reliability, and speed have been essential in furthering the wide use of digital applications. Microprocessors have spread the use of computer technology to personal computers, instruments, business machines, automobiles, consumer electronics, and other equipment. Other common uses of very large scale integrated circuits are in pocket calculators, electronic watches, cell phones, digital cameras, personal digital assistants (PDA's), appliance controls, and Internet circuits. For analog signal processing, integrated subsystems such as FM stereo demodulators and switched-capacitor filters are made. As improvements in the technology enable the development of more complex circuits, the challenge becomes not only one of how to build these circuits but also one of what to build, so that there is a demand for the new features and performance that will continue to drive the industry's growth. *See* CALCULATORS; DIGITAL COMPUTER; ELECTRONICS; MICROCOMPUTER; MICROPROCESSOR; MOBILE COMMUNICATIONS; SIGNAL PROCESSING; SWITCHED CAPACITOR CIRCUIT; WATCH.

Integrated circuits consist of the combination

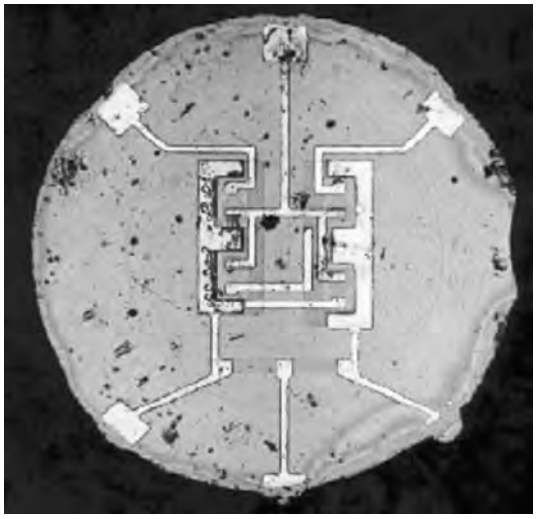

Fig. 1. Original integrated circuit, with aluminum interconnections on silicon. (*G. Moore, ISSCC '03, Intel Corp.*)

of active electronic devices such as transistors and diodes with passive components such as resistors and capacitors, within and upon a single semiconductor crystal. The construction of these elements within the semiconductor is achieved through the introduction of electrically active impurities into well-defined regions of the semiconductor. The fabrication of integrated circuits thus involves such processes as wet chemical etching, vapor-phase deposition of semiconductors and insulators, oxidation, ion implantation, vacuum deposition, sputtering, plasma etching, chemical mechanical polishing (CMP), rapid thermal processing (RTP), and lithography patterning techniques involving thin films of resist and optical exposure. *See* CRYSTAL GROWTH; DIFFUSION; ION IMPLANTATION; MICROLITHOGRAPHY; SEMICONDUCTOR HETEROSTRUCTURES; SPUTTERING.

Generally, integrated circuits are not straightforward replacements of electronic circuits assembled from discrete components. They represent an extension of the technology by which silicon planar transistors are made. Because of this, transistors or modifications of transistor structures are the primary devices of integrated circuits. Methods of fabricating good-quality resistors and capacitors have been devised, but the third major type of passive component, inductors, must be simulated with complex circuitry or added to the integrated circuit as discrete components. *See* TRANSISTOR.

Digital logic circuits were the easiest to adapt to these design changes. The first of these, such as inverters and gates, were produced in the early 1960s primarily for miniaturization of missile guidance computers and other aerospace systems. Analog circuits, often called linear integrated circuits, did not become commercially practical until several years later because of their heavy dependence on passive components such as resistors and capacitors. The first good-quality operational amplifiers for analog computers and instruments were produced in 1966. *See* ANALOG COMPUTER; LOGIC CIRCUITS; OPERATIONAL AMPLIFIER.

Types of Circuits

Silicon integrated circuits can be classified into two groups on the basis of the type of transistors that they employ: bipolar integrated circuits, in which the principal element is the bipolar junction transistor; and metal oxide semiconductor (MOS) integrated circuits, in which the principal element is the MOS transistor. Both depend upon the construction of a desired pattern of electrically active impurities within the semiconductor body, and upon the formation of an interconnection pattern of metal films on the surface of the semiconductor.

Bipolar circuits are generally used where highest logic speed is desired, and MOS for largest-scale integration or lowest power dissipation. Linear circuits have used bipolar technology extensively, but MOS devices are becoming much more important. High-performance bipolar transistors and complementary MOS (CMOS) transistors have been combined on the same chip (BiCMOS) to obtain circuits combining high speed and high density.

Integrated circuits are also made using other semiconductor materials, especially gallium arsenide (GaAs). These circuits tend to be more expensive but have much higher speeds. It is also much harder to integrate the large number of transistors on GaAs integrated circuits than can be done with silicon MOS. This article focuses on silicon MOS circuits, which probably account for over 95% of the integrated circuits currently produced (2006).

MOS integrated circuits. This major class of integrated circuits is called MOS because its principal device is a metal oxide semiconductor field-effect transistor (MOSFET). It is more suitable for very large scale integration than bipolar circuits, because MOS transistors are self-isolating and can have an average size of less than $6\lambda^3$, where λ is the minimal pattern that can be defined lithographically. This has made it practical to use over 10^9 transistors per chip. Because of this high-density capability, MOS transistors are used for high-density random-access memories (RAMs; **Fig. 2**), nonvolatile flash memories, and microprocessors. *See* COMPUTER STORAGE TECHNOLOGY; SEMICONDUCTOR MEMORIES.

Several major types of MOS device fabrication technologies have been developed since the mid-1960s: (1) metal-gate *p*-channel MOS (PMOS), which uses aluminum for electrodes and interconnections; (2) silicon-gate *p*-channel MOS, which employs polycrystalline silicon for gate electrodes and the first interconnection layer; (3) *n*-channel MOS (NMOS), which is usually silicon gate; and (4) complementary MOS (CMOS), which employs both *p*-channel and *n*-channel devices. The silicon gate CMOS is currently (2006) the dominant technology, and the term MOS usually refers to CMOS technology.

Both conceptually and structurally, the MOS transistor is a much simpler device than the bipolar transistor. In fact, its principle of operation has been known since the late 1930s, and the research effort

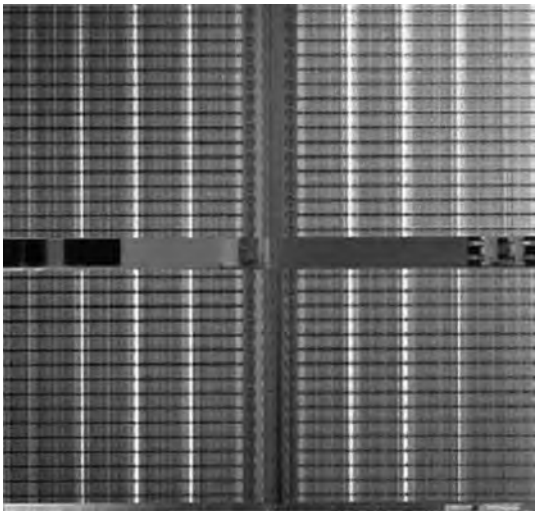

Fig. 2. Photomicrograph of metal oxide semiconductor (MOS) very large scale integrated circuit used as 153-megabit static random-access memory, using 45-nm technology (that is, with 45-nm feature sizes). Cell area = 0.346 μm^2, chip area = 119 mm^2. Chip has over 10^9 transistors. (*M. Bohr, Intel Corp.*)

that led to the discovery of the bipolar transistor was originally aimed at developing the MOS transistor. What kept this simple device from commercial utilization until 1964 is the fact that it depends on the properties of the semiconductor surface for its operation, while the bipolar transistor depends principally on the bulk properties of the semiconductor crystal. Hence MOS transistors became practical only when understanding and control of the properties of the oxidized silicon surface had been perfected to a very great degree.

CMOS technology. A simple CMOS inverter circuit is shown in **Fig. 3** and a circuit schematic in **Fig. 4**. The gates of the *n*-channel and *p*-channel transistors are connected together, as are the drains. The common gate connection is the input node, while the common drain connection is the output node.

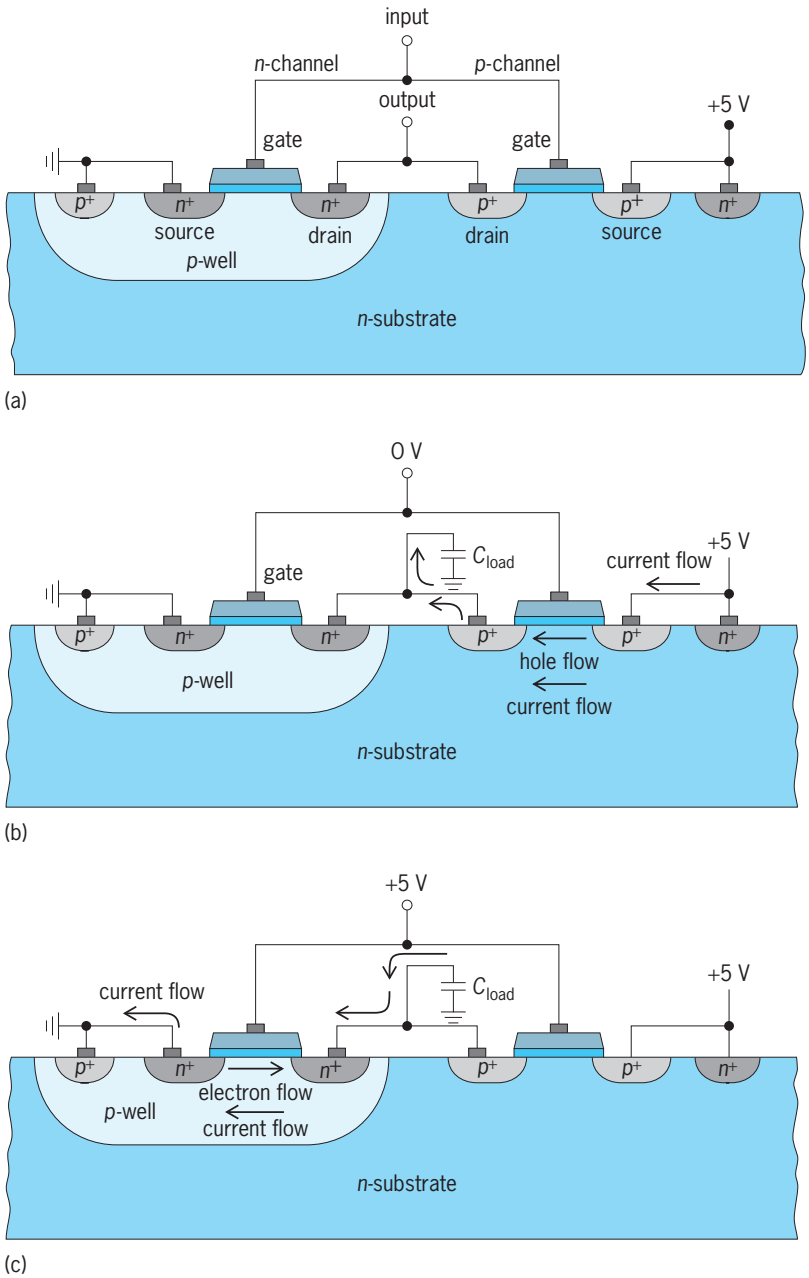

Fig. 3. Simple CMOS inverter circuit. (*a*) Schematic cross section. (*b*) Current flow when input is low at 0 V. (*c*) Current flow when input is high at 5 V.

A capacitor is added to the output node to model the loading expected from the subsequent stages on typical circuits.

When the input node is in the low state (0 V), the *n*-channel gate-to-source voltage is 0 V, while the *p*-channel gate to source voltage is −5 V. The *n*-channel transistor requires a positive gate-to-source voltage that is greater than the transistor threshold voltage (typically 0.5–1 V) before it will start conducting current between the drain and source. Thus, with a 0-V gate-to-source voltage it will be off, and no current will flow through the drain and source regions. The *p*-channel transistor, however, requires a negative voltage between the gate and source that is less than its threshold voltage (typically −0.5 to −1 V). When the input is put to the low state at 0 V, the −5-V gate-to-source potential is clearly less than the threshold voltage, and the *p*-channel will be turned on, conducting current from the source to the drain, and thereby charging up the loading capacitor. Once the capacitor is charged to the high state at 5 V, the transistor will no longer conduct because there will no longer be a potential difference between the source and drain regions.

When the input is now put to the high state at 5 V, just the opposite occurs. The *n*-channel transistor will be turned on, while the *p*-channel will be off. This will allow the load capacitor to discharge through the *n*-channel transistor, resulting in the output voltage dropping from a high state at 5 V to a low state at 0 V. Again, once there is no potential difference between the drain and source (capacitor discharged to 0 V), the current flow will stop, and the circuit will be stable.

This simple circuit illustrates a very important feature of CMOS circuits. Once the loading capacitor has been either charged to 5 V or discharged back to 0 V, there is no current flow since one of the transistors will be in the off state, and the standby power is very low. The important invention of CMOS digital logic was proposed as early as 1963 and became the dominant technology after 1980. The small stand-by power consumption is the reason for the high popularity of CMOS because of heat dissipation concerns. None of the other MOS technologies offers this feature without complex circuit techniques, and even then will typically not match the low standby power of CMOS. Bipolar circuits require even more power than these other MOS technologies. The price for the lower power of CMOS is the additional fabrication steps required (5–10% more) when compared to NMOS.

Integrated Optical Devices

Semiconductors have long been used as light sensors. Advances in very large scale integration have enabled large pixel arrays of light sensors, such as digital cameras and camcorders, to be made commercially. MOS is now preferred for sensor arrays because it permits large amounts of control logic to be fabricated in the same circuit as the sensors. *See* TELEVISION CAMERA; TELEVISION CAMERA TUBE.

In MOS optical arrays, there are two basic cell types: MOS devices or charge-coupled devices

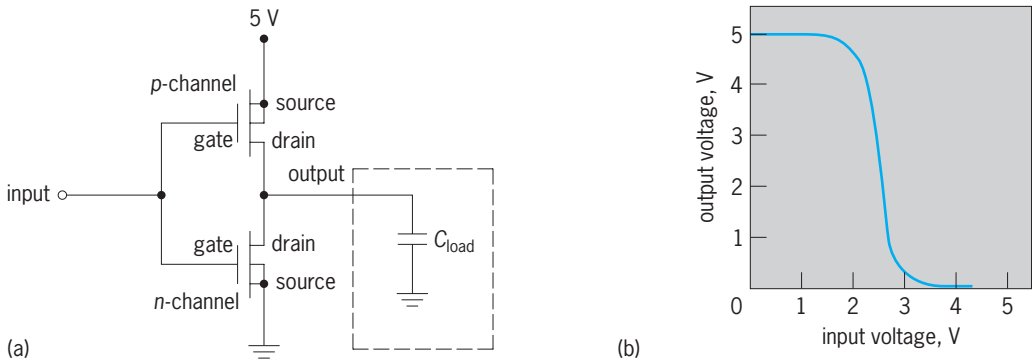

Fig. 4. Characteristics of the CMOS inverter circuit of Fig. 3. (*a*) Circuit diagram. (*b*) Transfer characteristics.

(CCDs). The sensor portion of the circuit views the scene through a transparent window. Light variations in the scene viewed cause variations in charge, producing signals transferable in shift-register fashion to processing stages such as amplifiers. These arrays have numerous applications, from measuring and sorting objects on production lines to optical character readers, which automatically translate printed and written information to digital computer input codes, such as the bar-code system and optical character recognition (OCR). *See* CHARACTER RECOGNITION; CHARGE-COUPLED DEVICES; COMPUTER VISION; INTEGRATED OPTICS; OPTICAL DETECTORS.

Microprocessors

The VLSI development having the most profound effect on electronic equipment economics and design in general is the microprocessor. The complexity and capability of these circuits, introduced in 1971, have expanded greatly. Initially these integrated circuits were applied to dedicated functions, with the tailoring of that function being done with the ROM circuits on the chip. However, there are now many different types of microprocessor circuits, ranging from the simple microcontroller types of functions to very complex computers on a chip such as the 64-bit microprocessor (**Fig. 5**).

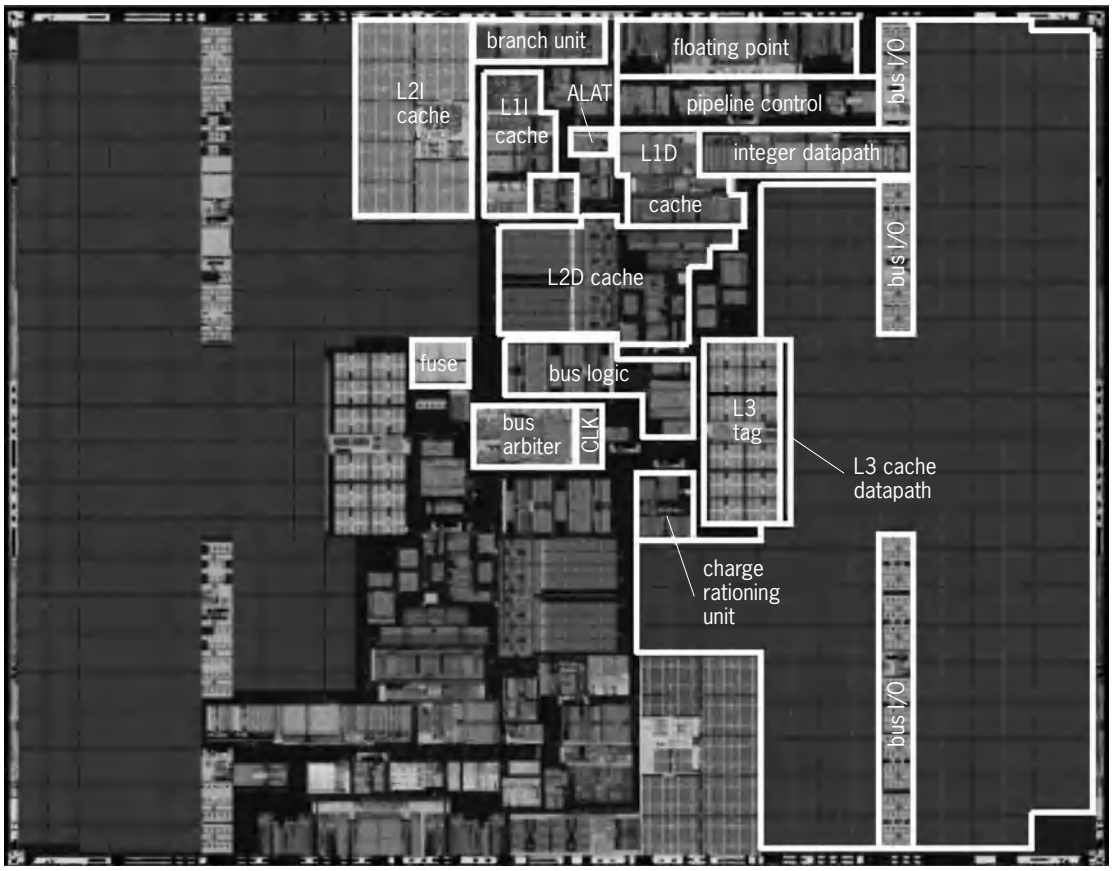

Fig. 5. Photomicrograph of a 64-bit MOS microprocessor, the Dual Core Intel® Itanium® processor (codenamed Montecito), which uses 1.7 × 10^9 transistors. ALAT = advanced local address table, CLK = clock. (*Intel Corp.*)

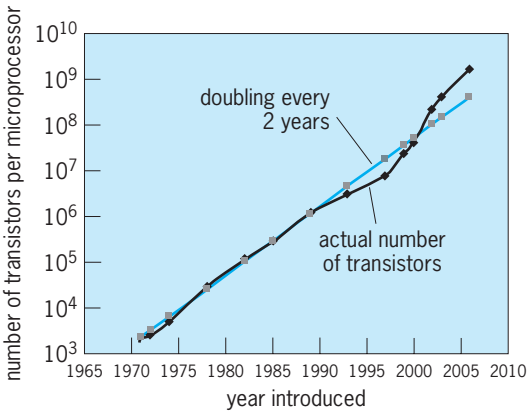

Fig. 6. Exponential growth in number of transistors for microprocessors. Actual number of transistors per microprocessor is compared with number that would have resulted from doubling every 2 years. (*Intel Corp.*)

This rapid development has been a result of both the technology which allows more circuits to be placed on a chip, and circuit innovation. The technology has increased the number of transistors on a chip from a few thousand in 1971 to over 10^9 in 2006, corresponding to a doubling every 2 years (**Fig. 6**). The circuit innovations include both digital techniques and the combination of analog and digital processing functions on the same chip. Both analog-to-digital and digital-to-analog data conversion circuits have been combined with the digital processor into one integrated circuit. The major digital innovations include more cache memory and parallel processing, with the example in Fig. 5 using two core processors.

Fabrication

Integrated-circuit fabrication begins with a thin, polished slice of high-purity, single-crystal semiconductor (usually silicon) and employs a combination of physical and chemical processes to create the integrated-circuit structures described above. Junctions are formed in the silicon slice by the processes of thermal diffusion and ion implantation. Electrical isolation between devices on the integrated circuit is achieved with insulating layers grown by thermal oxidation or deposited by chemical deposition. Conductor layers to provide the necessary electrical connections on the integrated circuit are obtained by a variety of deposition techniques. Precision lithographic processes are used throughout the fabrication sequence to define the geometric features required.

Requirements. The integrated circuit fabrication process is very sensitive to both particulate and impurity contamination. Airborne particulates must be minimized during the fabrication sequence, since even small ($\leq$0.05-micrometer) particles on the wafer surface can cause defects. A particulate-free fabrication ambient is normally achieved by the use of vertical laminar-flow clean rooms along with wafer transport systems that are clean, self-contained units that interlock with the equipment, reducing exposure to the environment. Lint-free garments are worn to minimize operator-borne particles (**Fig. 7**). To minimize impurity contamination effects, the chemicals, solvents, and metals that are used must be of the highest possible purity (electronic grade).

The precision and cleanliness requirements of integrated-circuit processing necessitate high discipline throughout the process sequence. This is achieved by extensive operator training, in-process tests and inspection with continual feedback, and a high degree of equipment calibration and control. The physical environment and operator attitude in an integrated-circuit fabrication facility are important factors for successful operation.

Simplified CMOS process flow. The principal steps in fabricating the simple CMOS inverter circuit of Fig. 3 are illustrated in **Fig. 8**, using an older (1980s) process flow. The process starts with the highly polished single slice of silicon crystal. For CMOS, the first step is typically to put in the *p*-well regions for the *n*-channel devices, if the original silicon wafer is *n*-type (Fig. 8*a*). It is also possible to start with a *p*-type wafer in which *n*-well regions are introduced for the *p*-channel transistors. While this discussion will be confined to the *p*-well case, most CMOS integrated circuits currently use wells for both the *n*-channel and *p*-channel transistors, with a *p*-type substrate starting material.

After the wells have been introduced, thick oxide layers (0.5–1.0 μm thick, called the field oxide) are grown in specific regions to isolate the transistor junctions from each other (Fig. 8*b*). Next the transistor gate oxides (typically 10–50 nanometer thick) are

Fig. 7. Person with lint-free garments in a vertical laminar-flow clean room for integrated-circuit fabrication with 300-mm (12-in.) wafer. (Personnel do not handle wafers in this manner. This was done just for the photograph.) (*Intel Corp.*)

grown. Prior to the silicon gate deposition, implants may be used to control the threshold voltages of the transistors. Polysilicon is then deposited for the gate regions and doped with impurities to reduce the resistance. The gate regions are then defined by using photoresist and lithography, and the polysilicon and thin gate oxide is removed where it is not wanted. The transistors are now ready for the introduction of impurities for the source and drain junctions. With CMOS, separate masking steps are required to introduce the phosphorus or arsenic for the *n*-channel source and drain junctions, and the boron for the *p*-channel source and drain junctions (Fig. 8*c*).

An insulating layer (typically a glass) to isolate the polysilicon gate conductor from the aluminum metallization is now deposited. Openings in the glass layer are made for the metal to make contact to the polysilicon lines and source/drain junctions. The metal is then deposited, and the regions to be removed are etched away after using photoresist and lithography techniques (Fig. 8*d*). Finally a scratch-protection passivation layer is deposited over the metal, with openings made at the bond pad areas for the bond wire connections.

The isolation of the transistors is accomplished by reverse-biased junctions. The *n*-substrate (or *n*-well) is always at the highest chip voltage, while the *p*-well (or *p*-substrate) is always at the lowest chip voltage.

While the previous example is useful for developing the concepts of an integrated-circuit process flow, it does not match the current technology (**Fig. 9**), which includes about seven copper metal layers. The newer technologies use chemical-mechanical polishing (CMP) extensively throughout the process to keep the surface flat, starting with field oxide region. Chemical mechanical polishing is also used to remove the unwanted copper metal in the dual damascene process.

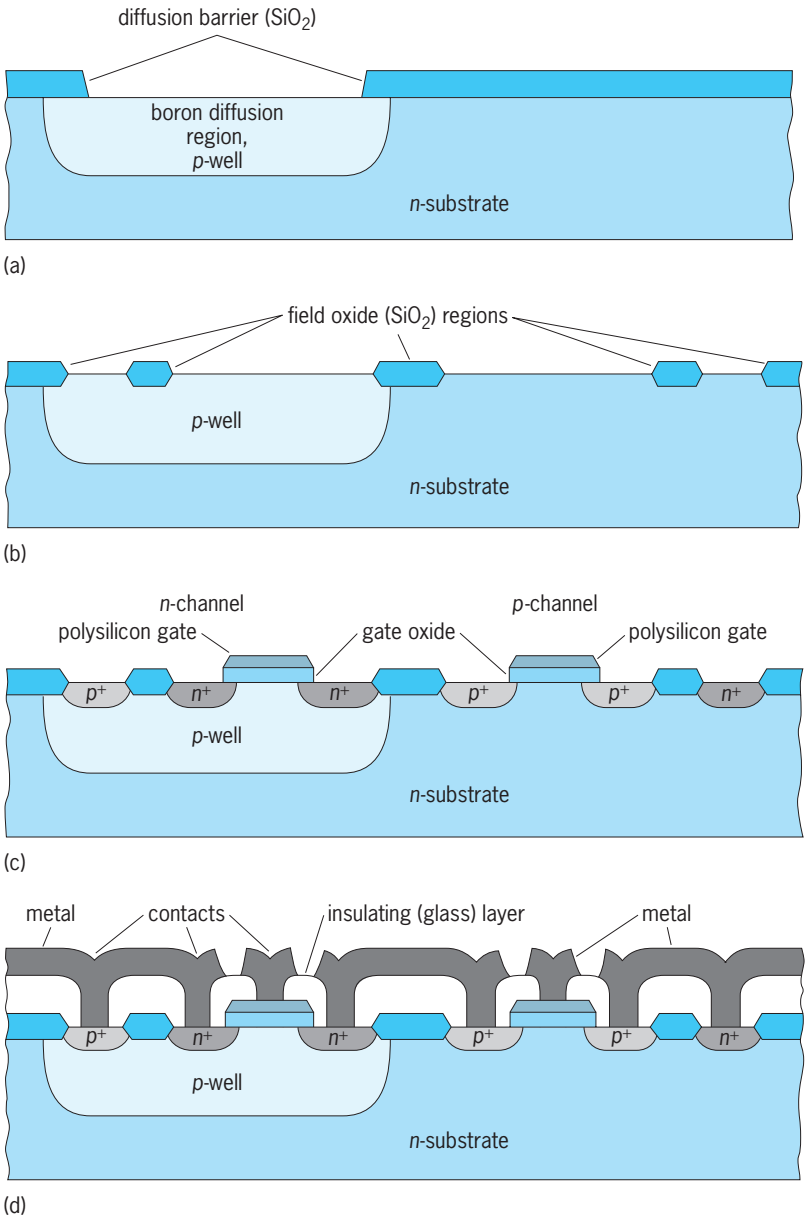

Fig. 8. Steps in fabrication of CMOS inverter circuit. (*a*) Introduction of *p* well. (*b*) Growth of field oxides. (*c*) Addition of gate oxides, polysilicon gates, and junctions. (*d*) Deposit of insulating layer, opening of contacts, and deposit of metal.

Trends and Process Changes

While the previous CMOS process flow is useful as an educational example to help highlight the fabrication process, the reality of current integrated-circuit technology is much more complex. The complexity of integrated circuits has been increasing exponentially since they were introduced in the early 1960s. This exponential growth was first recognized by Gordon Moore in 1965 and, as mentioned above, has been proceeding with a doubling about every 2 years, which has become known as Moore's law (**Fig. 10**).

Integrated-circuit fabrication technology changes rapidly due to the steadily decreasing feature size of the individual circuit elements, the improvements in yields allowing large chip sizes, the advances in equipment required for the finer features, and the increasing wafer sizes. Because of these rapid changes, a fabrication facility will remain state of the art in capability for no more than 2 years without making major equipment and process changes.

The size of the circuit elements has been decreasing as improvements in the lithography and etching techniques are incorporated (**Fig. 11***a*), amounting to a 30% reduction of the feature size every 2 years. This results in the area of the static RAM (SRAM) cell decreasing by about 50% every 2 years (Fig. 11*b*).

In addition to the increase in density, the wafer size has been increasing (**Fig. 12**), which makes it more economical to produce the large die sizes. Whereas in 1970 50-mm (2-in.) wafers were common, by 1980 100-mm (4-in.) wafers were the choice for new manufacturing lines. During the mid-1980s, the 150-mm (6-in.) wafer was the direction of new manufacturing lines, while in 1989 the 200-mm (8-in.) wafer began to be used. Starting about 2000, 300-mm (12-in.) wafers were being used in production. The larger wafer, plus the smaller feature sizes of the newer technologies, continues to decrease the cost of integrated circuits.

The economic drivers of the integrated circuit

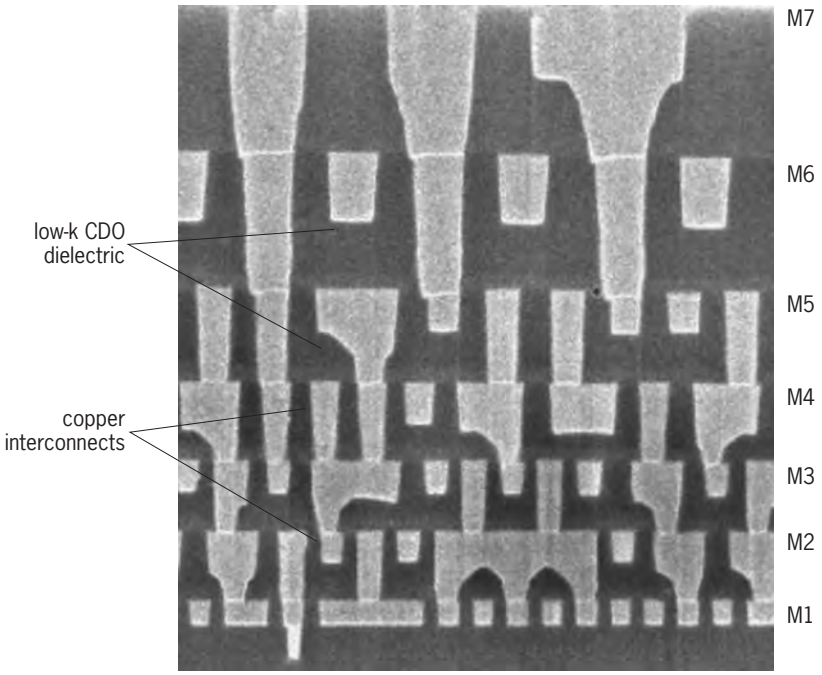

Fig. 9. Cross-sectional view using scanning electron microscope of chip manufactured by 90-nm process (that is, with 90-nm feature sizes), using seven copper metal layers. Low-permittivity (low-*k*) carbon-doped oxide (CDO) dielectric is used. (*G. Moore, ISSCC '03, Intel Corp.*)

industry have been cost and performance. As the technology has allowed faster and more complex functions on a chip, and feature size and yield have kept the cost from rising, the customer has been able to get more performance for the same or less cost. This is illustrated in **Figs. 13** and **14**, which show that while the cost of lithographic tools and the overall fabrication cost are rising the transistor cost and the performance cost, as measured in millions of instructions per second (MIPS), have been decreasing due to the smaller geometries and the efficiencies of scaling.

Historic process developments. These scaling trends could not have been achieved without major improvements in the technology. Along with improved lithography and transistor gate oxides, there were other major technology changes that made the scaling possible, and this trend must continue for the scaling to continue into the future.

The first major development in MOS integrated circuits was the understanding and development of quality gate oxides in the 1960s. This was from the work of many people, and resulted in an understanding of the silicon surface and the development of stable thermal oxides for the gate insulator. Initially, gate oxide thicknesses were greater than 100 nm (larger than the typical minimum feature size on today's technology), while currently (2006) they are of the order of 1–2 nm.

Another important development in the late 1960s was the use of polycrystalline silicon (polysilicon) for the metal gate. Previously, aluminum was used, which severely limited the processing temperatures that could be used after deposition. With polysilicon, self-aligned source-drain junctions could be fabricated, and polysilicon doping could better match threshold voltages for CMOS transistors. Polysilicon is still being used with a nickel salicide cap (discussed below).

Field oxides using the local oxidation process (LOCOS) with the self-align twin-well process were developed during the 1970s. This process uses a silicon nitride layer to prevent oxidation. This practice greatly lowered the surface step between the field oxide and the active region.

Ion implantation of elements was developed in the 1960s and 1970s, and became the primary method to introduce dopant impurities into the silicon in the 1980s. Both the dose and depth of the impurities can be controlled much better with implantation. It also leads naturally to source and drain tilt implants to

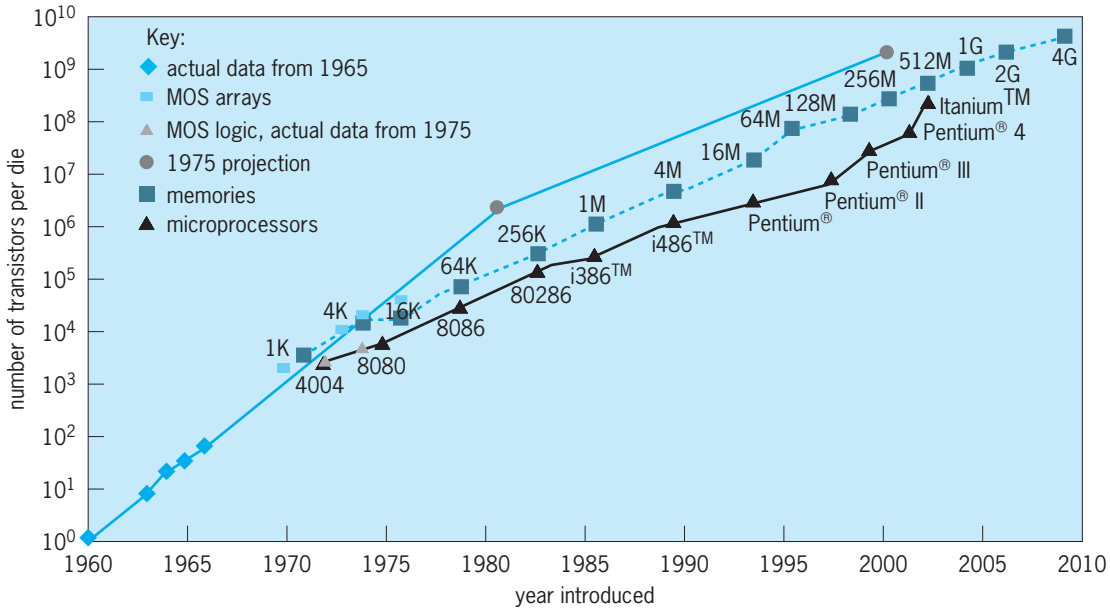

Fig. 10. Exponential growth in integrated circuit complexity, which is doubling about every 2 years. (*G. Moore, ISSCC '03, Intel Corp.*)

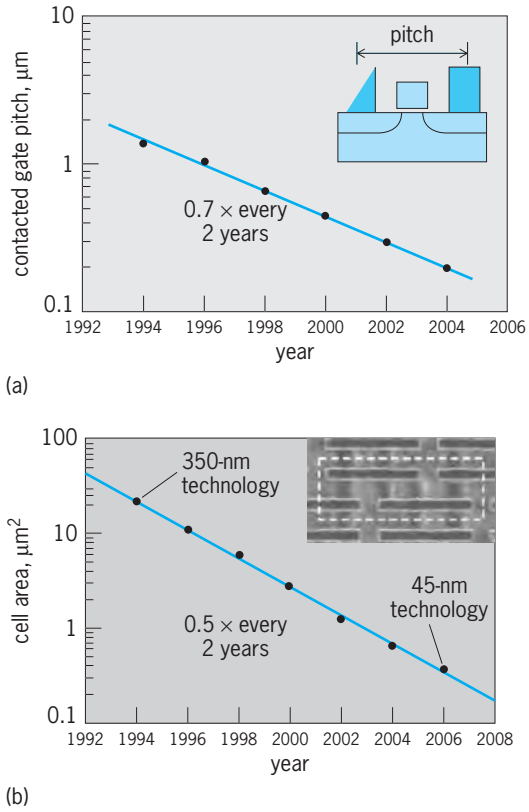

Fig. 11. Exponential trends of decreasing size and area. (*a*) Trend in contacted gate pitch (defined in the inset). Solid line indicates multiplication by factor of 0.7 every 2 years. (*b*) Trend in area of static random-access memory cell. Solid line indicates multiplication by factor of 0.5 every 2 years. Inset shows cell of 45-nm generation, introduced in 2006, with area of 0.346 μm, manufactured by dry-etching technique with ultraviolet light of 193-nm wavelength. (*P. Bai, Intel Corp.*)

control the punchthrough and short-channel effects (SCE) of submicrometer transistors.

Lithography changed significantly with the introduction of 5–10x step and repeat projection tools. Now one could make the mask 5–10 times larger and have better resolution and alignment control on the wafer. Instead of exposing the whole wafer at once, the exposure is into a small area on the wafer that can better be optimized, and this exposure is followed by stepping and repeating to the other areas on the wafer, until the total wafer has been exposed. Combining this technology with laser alignment, Ebeam mask generation, optical proximity correction, phase-shift masks, shorter wavelengths, and high-index immersion has allowed the printing of 22-nm feature sizes using optical techniques.

During the 1980s, increasing the die yields on the wafer became a major effort. Prior to this, the defect densities were greater that 1 defect/cm^2. Such a defect density meant that a die that was approximately 1 cm (0.4 in.) square could not be manufactured, since almost all the dice would then have a defect. The area of the microprocessor chip of Fig. 5 is almost 6 cm^2, which requires the defect density to be of the order of 0.1 defect/cm^2 or less for reasonable yields to be achieved. Thus, defect densities have decreased, at the same time that feature sizes have decreased and the complexity of the process has increased (the number of masking steps has more than doubled since 1980)—which is a remarkable achievement. This accomplishment required the introduction of new monitoring techniques during the processing of wafers (including new laser scanning systems), major modifications to processing equipment (to reduce defects introduced during processing), and new analysis techniques [such as fine-focused ion beam (FIB) techniques].

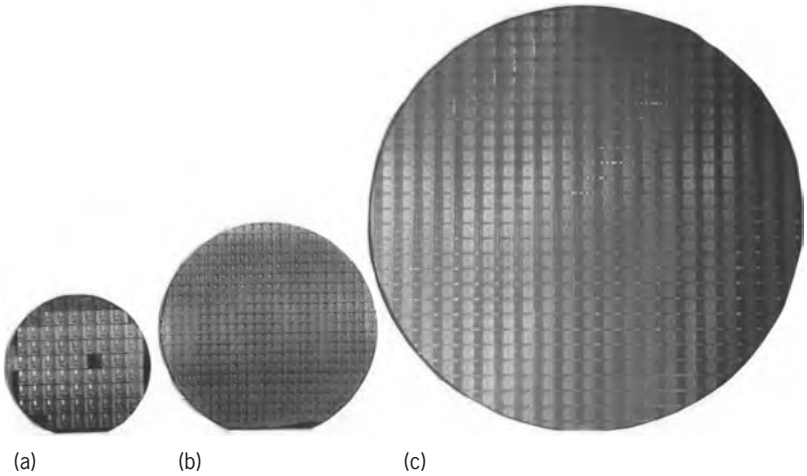

Fig. 12. Increase in wafer sizes, showing the increased number of dice (chips) per wafer available when increasing the wafer area. (*a*) 100-mm (4-in.) wafer. (*b*) 150-mm (6-in.) wafer. (c) 300-mm (12-in.) wafer. (*Intel Corp.*)

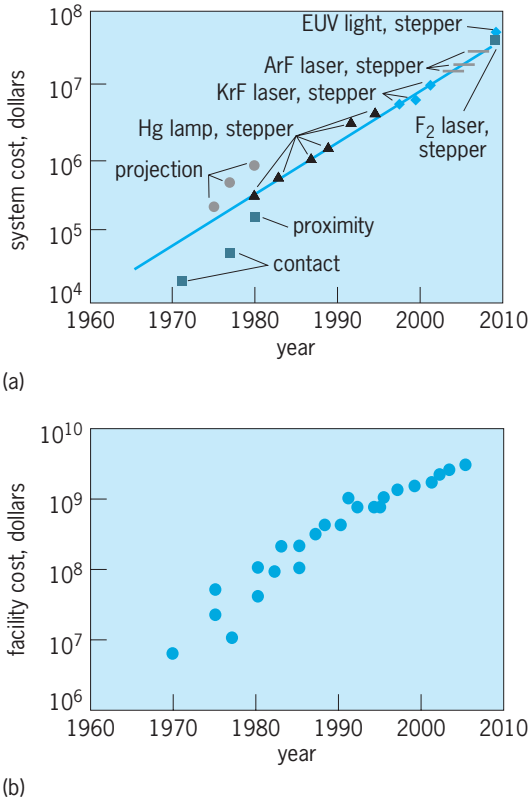

Fig. 13. Trends in (*a*) lithographic tool cost and (*b*) new fabrication costs in the integrated circuit industry. (*IC Knowledge LLC*)

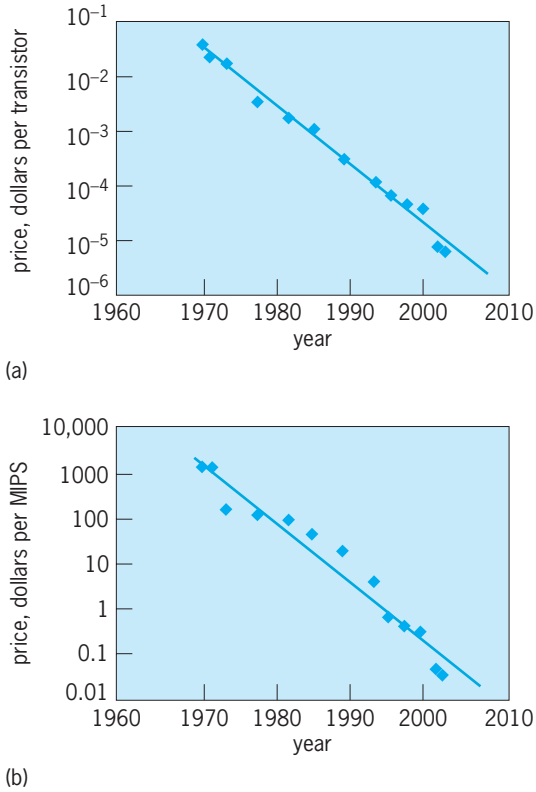

Fig. 14. Exponential trends in the price of the integrated circuit. (*a*) Price per transistor. (*b*) Price per MIPS (millions of instructions per second). (*S. Borkar, Intel Corp.*)

Dry-etching techniques using plasma became available in the 1980s and have become the primary etching techniques. In comparison with wet-etching techniques, their advantage lies in their ability to create large-aspect-ratio features without undercut. Without them, it would be impossible to pattern features smaller than 0.5 μm.

To reduce the polysilicon electrical resistance and improve the source/drain junction contact resistance, self-aligned silicide (salicide) formation on the polysilicon and junctions was developed during the 1980s. Initially, the salicide was formed using titanium deposition. However, currently nickel is the preferred metal. The silicide formation requires rapid thermal processing (RTP), which has been another important technology development and has been added to other parts of the process.

During the 1980s, a planarization technique involving chemical mechanical polishing (CMP) of the glass insulating layers between the metal layers was developed. This advance combined with tungsten plugs for the interconnection of the aluminum metal layers allowed the number of metal layers to increase from the two-layer practical limit at the time and dramatically reduced the lithographic and etching problems for metal layers.

During the 1990s, plasma etching into the silicon allowed for deeper field isolation without a large step on the surface, which is called the deep trench. This practice has since been combined with chemical mechanical polishing to remove the surface step altogether. The technique is also used in dynamic RAMs (DRAMs) to increase the area for the capacitors in the cells.

Because of its lower resistance and higher current capability, copper metal has always been preferred over aluminum metal. However, due to the limited number of good dry etch techniques for copper and the need for a diffusion-barrier layer. the dual damascene process had to be developed before copper could be practical. During this process, trenches are etched into the insulating dielectric layers between the metal layers (Fig. 9), which are first covered with a thin barrier seed layer, and then filled with copper using an electroplating process. The excess copper is then removed using chemical mechanical polishing techniques. Thus, copper can be removed without etching, which was always a major problem.

Prospective process developments. As the technology continues to shrink, the printing of these small features is becoming a major problem with the existing lithography tools. The wavelength of the light used in lithography has not decreased as significantly in recent technologies as previously (**Fig. 15**). Only with the special techniques of subresolution optical proximity correction (OPC), phase-shifted masks, and high-index immersion is it possible to print 45-nm features using 193-nm wavelength. Looking forward, it appears that extreme ultraviolet (EUV) radiation at a wavelength of 13 nm will be required at some point.

Another area of scaling concern is the transistor performance. As the gate oxide has been scaled to less that 2 nm, the gate current leakage has been increasing due to quantum tunneling effects. Also, surface effects increase the effective electrical thickness so that the capacitance coupling into the channel is less than classically predicted from physical oxide thickness, resulting in fewer free carriers, which means less channel current. To overcome these problems, high-*k* (high-permitivity) dielectrics are being investigated. In **Fig. 16**, the thicker experimental

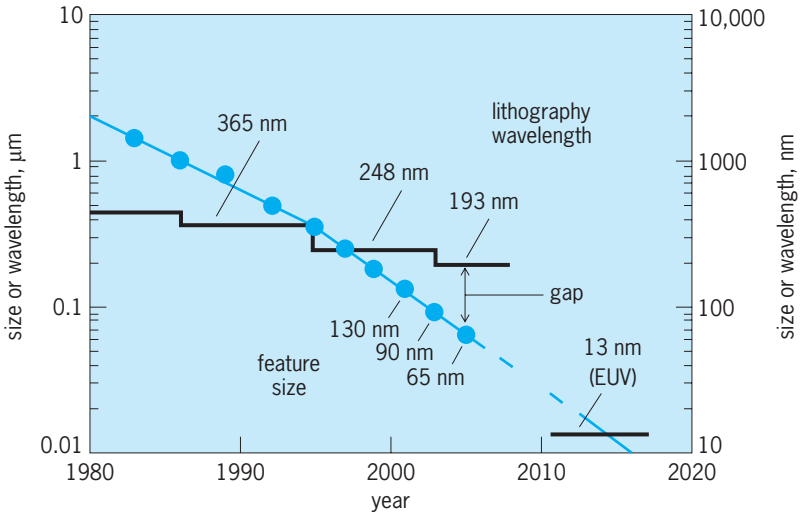

Fig. 15. Trends of feature size with lithography, compared with the wavelength of light used in lithography. (*P. Bai, Intel Corp.*)

high-*k* dielectric has more capacitance and over 100 times less leakage than the thinner silicon dioxide layer. *See* DIELECTRIC MATERIALS; PERMITTIVITY; TUNNELING IN SOLIDS.

Special efforts have been made to improve the carrier mobility of the transistors. As the carrier mobility increases, the current drive of the transistors increases, reducing delay, which increases the speed. The use of enhanced strain to improve mobility is illustrated in **Fig. 17** for both *n*-channel and *p*-channel transistors, where the NMOS strain is introduced through a tensile capping layer on the edges of the gate and PMOS strain is introduced using germanium-silicon (GeSi) in the channel region. Additional methods may need to be developed that continue to improve the mobility.

Another approach to improve the transistor performance is to use the additional sides of the transistor to carry current. An example is the tri-gate transistor (**Fig. 18**), whose drive currents are 50% higher than those of the planar transistor.

The delay in sending signals across the chip is related not just to the transistor drive but also to the load, which is mostly capacitance from the metal lines. As the dimensions between metal lines decrease, the edge capacitance per unit length will increase if the thickness does not scale. To reduce this capacitance and the capacitance between metal layers, low-*k* dielectric layers are being introduced. A processing example is shown in Fig. 9, and improvements will continue to be developed. The use of optical techniques to reduce delays on integrated circuits is also being studied.

The future technology challenges are still paramount, and the complexity of integrated circuits can not continue to increase at the present rate forever; however, it is very difficult to predict when Moore's law will become invalid. For example, in the early 1980s, no one could foresee the impact that chemical mechanical polishing would have on the technology. Today, new materials such as carbon fibers and other carbon structures are being investigated, which may also have a large impact in the future.

Design

VLSI chips containing over 10^9 transistors and operating at a few gigahertz (10^9 cycles/second) have been designed and fabricated, and are commercially available. Projections indicate that silicon chips containing as many as 10^{10} transistors are feasible for digital applications. The design of these new circuits starts with computer simulations that are done with computers using chips from the previous cycle. Advanced integrated circuits from current manufacturing are used to develop, design, test, and analyze the next generation of integrated circuits. *See* COMPUTER-AIDED CIRCUIT DESIGN.

The process development cycle for a new technology begins with the assumption that Moore's law continues to be valid. This assumes an approximate 30% reduction in linear dimensions and the introduction of products with the new technology 2 years after the previous generation. This does not mean that the development starts just 2 years before product introduction, for it may take more than 2 years. Some developments take many years, such as EUV lithography, which has been under development for more than 5 years (as of 2006) and may require another 5 years before it will used in manufacturing. Thus, the industry needs to anticipate technology and design requirements many years in advance to have enough time to adequately insert the new technologies and designs.

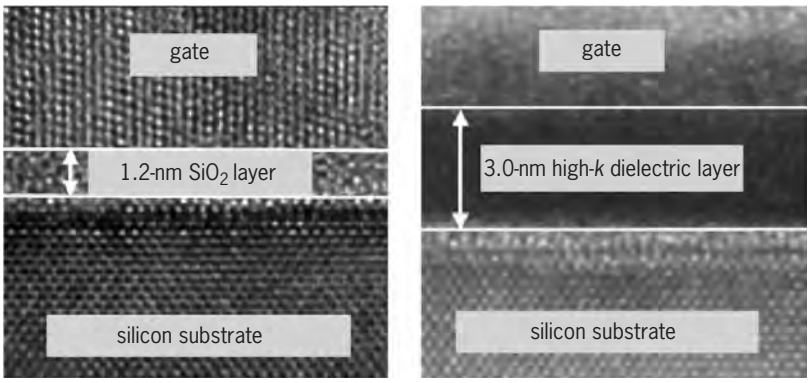

(a) (b)

Fig. 16. Comparison of transmission electron microscope cross sections of (*a*) standard 90-nm process chip with oxide layer and (*b*) experimental chip with thicker high-*k* (high-permittivity) dielectric. (*G. Moore, ISSCC '03, Intel Corp.*)

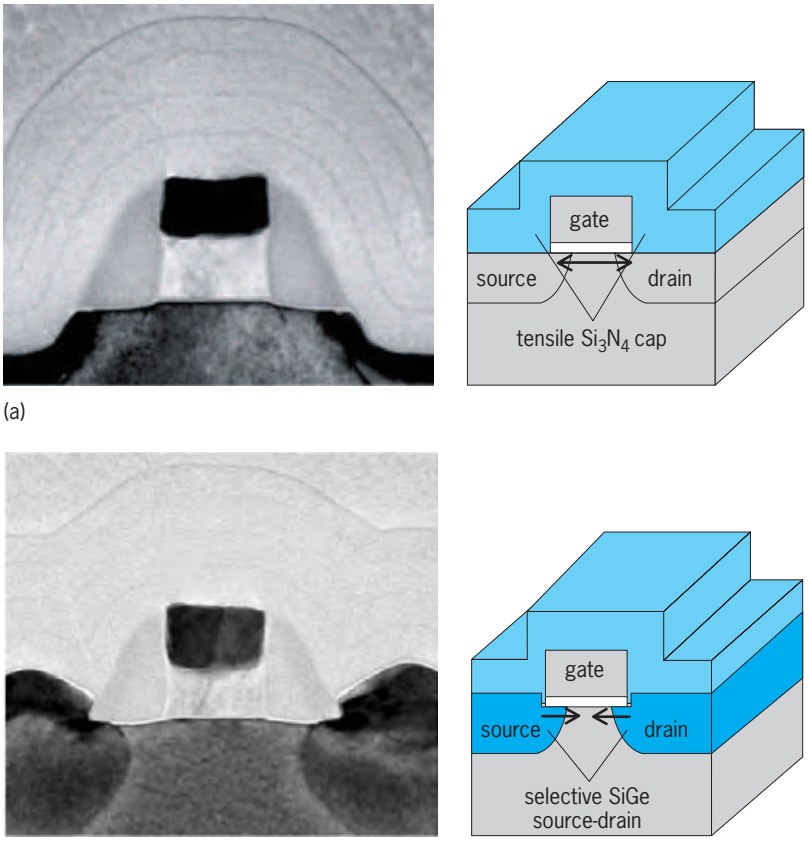

(a)

(b)

Fig. 17. Cross-sectional views of (*a*) NMOS and (*b*) PMOS transistors with enhanced strain to improve carrier mobility. (*C.-H. Jan, Intel Corp.*)

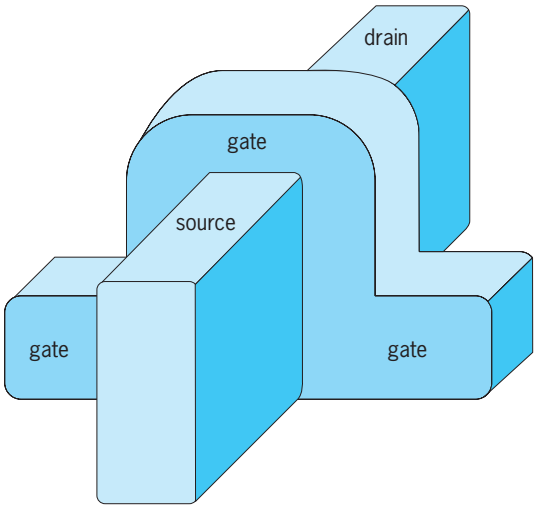

Fig. 18. Tri-gate transistor. (*R. Chau, Intel Corp.*)

The design process starts with computer simulations of the process and devices. As process flows are developed, the design rules and goals for electrical parameters become available to start the design of standard cells (smaller blocks of logic and cells), which are optimized to the design rules and performance of the new technology. A library of these cells is made, which is used by logic design teams. In technologies developed for microprocessors, SRAMs like that in Fig. 2 are often used to evaluate the process during development. There is also a lot of SRAM on microprocessors (Fig. 5), which makes it a good yield test vehicle. Chip designs from the previous technology are also used to evaluate the new technology, after shrinking the design as much as possible to match the new design features. These are often memory products (RAMs) or microprocessors, with known performance and yield indicators.

Computer-aided engineering (CAE) systems provide the environment, specific computer tools, data management, and other services that are intended to support the design of these very complex, high-performance products. In many cases, the design of complex chips requires the cooperative endeavors of large design teams; thus the CAE system must also manage the design process to ensure that proper documentation has occurred, needed changes in the design database are made, and a chosen design methodology is enforced. The design process must be adapted to the very short design cycle times from product conceptualization to manufacturing of a salable product. This is a distinctive characteristic of the semiconductor industry. *See* COMPUTER-AIDED ENGINEERING.

As it is becoming more difficult and more expensive to continue the performance improvements on the processing side, the architecture, logic, and circuit design areas become even more important to continue the overall performance improvements needed for the growth of the industry. As noted in **Fig. 19**, the computing performance as measured in MIPS has recently been going up faster than expected from just technology improvements. Some of the design features that increase overall performance include increased bit length, more memory on chip, instruction-level parallel processing (superscalar, speculative OOO), thead-level parallel processing (multithreaded), multicore parallel processing, more prediction in pipelines, and optimizing the layout to reduce delay times. In addition, special purpose

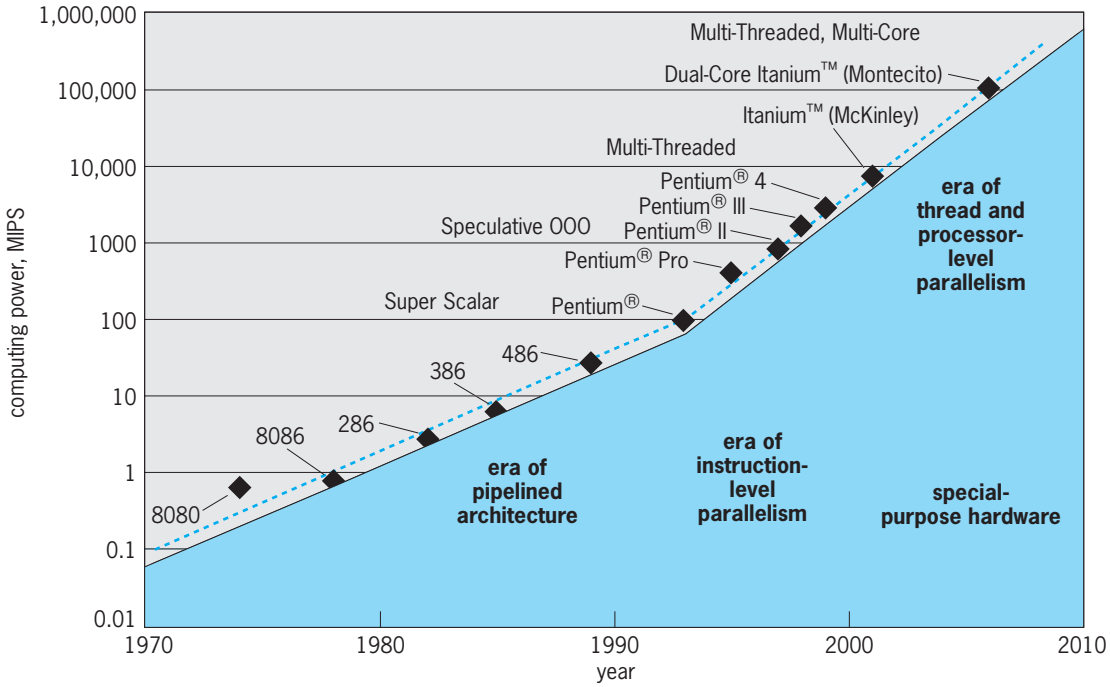

Fig. 19. Increase in computing power [measured in millions of instructions per second (MIPS)] of integrated circuits versus time and design enhancements. (*S. Borkar, Intel Corp.*)

hardware is used, which optimizes the performance of a specific function. *See* CONCURRENT PROCESSING.

Ron Burghard; Youssef El-Mansy; Neil Berglund

Bibliography. L. Berlin, *The Man Behind The Microchip*, 2005; N. G. Einspruch et al. (eds.), *VLSI Electronics: Microstructure Science*, vols. 1–24, 1981–1994; D. J. Elliott, *Integrated Circuit Fabrication Technology*, 2d ed., 1989; R. L. Geiger, P. E. Allen, and N. Strader, *VLSI Design Techniques for Analog and Digital Circuits*, 1990; P. R. Gray et al., *Analysis and Design of Analog Integrated Circuits*, 4th ed., 2001; D. A. Hodges, H. G. Jackson, and R. A. Saleh, *Analysis and Design of Digital Integrated Circuits*, 3d ed., 2004; S. I. Long and S. E. Butner, *Gallium Arsenide Digital Integrated Circuit Design*, 1990; G. Moore, *Presentation at the International Solid State Circuits Conference (ISSCC)*, 2003; J. Plummer, M. D. Deal, and P. B. Griffin, *Silicon VLSI Technology*, 2000; T. E. Price, *Introduction to VLSI Technology*, 1994; D. A. Pucknell and K. Eshraghian, *Basic VLSI Design*, 3d ed., 1994; W. S. Ruska, *Microelectronic Processing*, 1987; S. M. Sze, *Physics of Semiconductor Devices*, 2d ed., 1981; S. M. Sze, *VLSI Technology*, 2d ed., 1988; P. Van Zant, *Microchip Fabrication*, 5th ed., 2004; S. Wolf, *Silicon Processing for the VLSI Era*, vol. 4: *Deep-Submicron Process Technology*, 2002; E. S. Yang, *Microelectronic Devices*, 1988; M. Zambuto, *Semiconductor Devices*, 1989.

Integrated optics

The study of optical devices that are based on light transmission in planar waveguides, that is, dielectric structures that confine the propagating light to a region with one or two very small dimensions, on the order of the optical wavelength. The principal motivation for these studies is to combine miniaturized individual devices through waveguides or other means into a functional optical system incorporated into a small substrate. The resulting system is called an integrated optical circuit (IOC) by analogy with the semiconductor type of integrated circuit. Alternative names for an IOC include opto-electronic integrated circuit (OEIC), optical integrated circuit (OIC), and photonic integrated circuit (PIC). An integrated optical circuit could include lasers, integrated lenses, switches, interferometers, polarizers, modulators, detectors, and so forth. Important uses envisioned for integrated optical circuits include signal processing (for example, spectrum analysis and analog-to-digital conversion) and optical communications through glass or plastic fibers, which are themselves circular (or elliptical) waveguides. Integrated optical circuits could be used in such systems as optical transmitters, switches, repeaters, and receivers. *See* INTEGRATED CIRCUITS; OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

The advantages of having an optical system in the form of an integrated optical circuit rather than a conventional series of components include (apart from miniaturization) reduced sensitivity to air currents and to mechanical vibrations of the separately mounted parts, low driving voltages and high efficiency, robustness, and (potentially) reproducibility and economy. As in the case of semiconductor integrated circuits, an integrated optical circuit might be fabricated on or just within the surface of one material (the substrate) modified for the different components by shaping structures (using etching, for example) or incorporating suitable substitutes or dopants, or alternatively, by depositing or expitaxially growing additional layers. This form of IOC is known as a monolithic optical integrated circuit (MOIC). It is also possible to construct independent components that are then attached to form the integrated optical circuit. This option, called hybrid, has the advantage that each component could be optimized, for example, by using gallium aluminum arsenide lasers as sources for an integrated optical circuit and silicon detectors. In the former case, the MOIC is expected to have the advantage of ease of processing, similar to the situation for monolithic semiconductor integrated circuits. The most promising materials for monolithic integrated optical circuits are direct band-gap semiconductors composed of III-V materials such as gallium aluminum arsenide (GaAlAs) and indium gallium arsenide phosphide (InGaAsP) since with suitable processing they may perform almost all necessary operations as lasers, switches, modulators, detectors, and so forth. Recently, MOICs have been fabricated by integrally bonding chips of III-V semiconductors to chips of silicon in order to combine the above optical devices with well-developed silicon electronic circuits such as CMOS (complementary metal oxide semiconductor) computer logic chips. *See* BAND THEORY OF SOLIDS; SEMICONDUCTOR.

Guided waves. The simplest optical waveguide is a three-layer or sandwich structure with the index of refraction largest in the middle or waveguiding layer (**Fig. 1**). The lower and top layers are usually the substrate and superstrate, respectively. Often, the top layer is air and the waveguide layer is referred to as a film. Sometimes, the outer regions are called cladding layers. A guided wave does not have light distributed uniformly across the waveguide, but is a pattern that depends on the indices of refraction of all three layers and the guide thickness. The waveguide is usually designed by selecting its refractive index and thickness, so that only one such characteristic pattern propagates with no change in

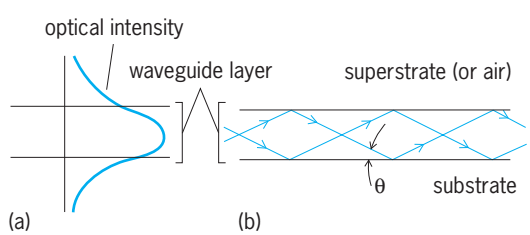

Fig. 1. Optical waveguide. (*a*) Optical intensity of the fundamental guided mode. (*b*) Ray picture of the guided mode.

shape. This pattern, referred to as the fundamental or lowest-order mode, travels down the guide with a characteristic velocity.

A mode pattern such as that of Fig. 1*a* may be viewed as the superposition of two progressive waves moving at a particular slight angle to the propagation direction as shown in Fig. 1*b*. If the ray directions are sufficiently near grazing incidence, the large refractive index of the waveguide layer will cause total internal reflection at the interfaces, and the mode will propagate along the waveguide without loss. *See* WAVEGUIDE.

Materials and fabrication. Waveguides have been made of many different materials, most of which may be categorized as ferroelectric, semiconductor, or amorphous. Examples of these classes are lithium niobate, gallium arsenide/gallium aluminum arsenide [GaAs/(GaAl)As], and glass, respectively. Methods for fabricating a waveguide layer at the surface of lithium niobate include heating the crystal in a vacuum to drive off lithium oxide, diffusing titanium metal into the crystal, or replacing lithium ions with hydrogen ions in an acid bath. These processes create a region of high refractive index near the surface; air is the superstrate. Although the refractive index of the waveguide layer decreases with depth rather than being constant, the guided mode is similar to that in Fig. 1. A semiconductor waveguide may be fabricated, for example, by successively growing crystalline layers of $(Ga_{0.7}Al_{0.3})As$, GaAs, and $(Ga_{0.7}Al_{0.3})As$. The thin GaAs waveguide layer of high refractive index is thus interposed between thicker cladding regions of the lower-index $(Ga_{0.7}Al_{0.3})As$. High-energy ion implantation of suitable dopant atoms into a semiconductor may also be used to create the high-refractive-index waveguide layer. Glass waveguides may be formed, for example, by sputter deposition of a relatively high-refractive-index glass on a lower-index glass substrate. Deposition of polymer materials also may be used to form amorphous waveguides. The advantage that polymer waveguides have over glass is that they are flexible. In recent years, waveguides and other integrated optic devices have been fabricated using polymer materials by a number of companies in order to reduce costs. Semiconductors such as GaInAsP and dielectrics such as lithium niobate are relatively expensive, and complicated processes must be used to create devices in them. The demand for inexpensive integrated optic devices has grown markedly as telecommunication networks worldwide have converted to lightwave signals transmitted over optical fibers and processed by integrated optic circuit chips. Many of the devices that formerly were made only in semiconductors or lithium niobate are now also made in polymers. These include waveguides, couplers, gratings, modulators, light detectors, light-emitting diodes (LEDs), and even lasers. In many cases, a polymer device can be substituted for one made in a semiconductor or crystalline dielectric, resulting in a significant cost reduction without excessive loss of performance. *See* FERROELECTRICS; GLASS; ION IMPLANTATION; LIGHT-EMITTING DIODE; OPTICAL MODULATORS; SEMICONDUCTOR HETEROSTRUCTURES.

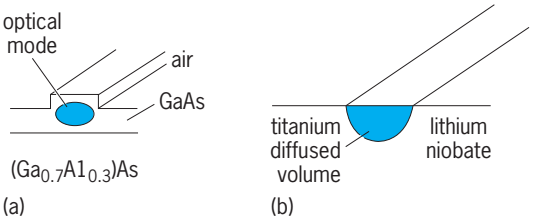

Fig. 2. Two-dimensional transverse waveguiding. (*a*) Rib waveguide. (*b*) Diffused channel guide.

Waveguides that confine light in two dimensions, rather than one as shown in Fig. 1, use refractive index differences in both transverse directions. Examples of these are the rib guide (**Fig. 2***a*) and the titanium-diffused channel guide (Fig. 2*b*). In fabricating a rib guide, photolithography is employed to delineate the stripe, followed by chemical or dry etching to remove the undesired material. The channel guide is produced by etching away all but a strip of metal prior to diffusion. *See* WAVEGUIDE.

Coupling of external light beams. An external light beam may be coupled into a waveguide by introducing the light at the end of the guide at an edge of the substrate or through the surface of the waveguide. The former approach may employ a lens to focus the light beam onto the guide end. Alternatively, the laser or an optical fiber is placed against, or in proximity to, the guide end, which is referred to as butt coupling. Light may also be injected through the guide surface with an auxiliary element. Simply illuminating that surface will not suffice since the refractive-index difference needed for effective waveguiding prevents light from entering at the required angle. The necessary angle, which is exactly that of the propagating waves shown in Fig. 1*b*, may be obtained by passing the light through a high-refractive-index prism whose surface is so close to the guide layer that waveguiding is disturbed in that region. The prism not only permits the introduction of light but also acts to couple light from the waveguide. For this reason, the light must illuminate the guide near the end of the prism (**Fig. 3**). The angle of light incident on the surface of the waveguide layer can also be modified to coincide with the internal angle of the propagating wave by a periodic structure or

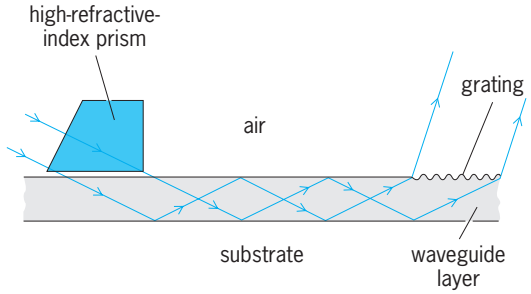

Fig. 3. Prism input coupling and grating output coupling. With the ray directions reversed, the grating acts as an input coupler and the prism as an output coupler.

grating on the waveguide surface. Such a grating may be made by a multiple-step process that begins by depositing a light-sensitive material called photoresist and concludes with etching and cleaning. The grating also functions as an output coupler. By reversing the incident and output beam directions in Fig. 3, the roles of the prism and grating couplers are interchanged.

Lasers and distributed feedback. The diode laser is already an integrated optics device in a sense since the lasing medium is a waveguide laser interposed between two cladding layers. The waveguide layer may be GaAs, in which case the outer regions are composed of $(Ga_{0.7}Al_{0.3})As$, for example. When used as a separate source, the crystal facets act as end reflectors. The diode laser's high-efficiency low-voltage operation, small size, and physical integrity cause it to be the laser of choice in many hybrid integrated optics applications. Moreover, it lends itself to integration in a monolithically integrated optical circuit where reflectivity may be provided by introducing an appropriate periodic structure. This could be a thickness variation in the waveguide layer. Devices utilizing such structures are called distributed feedback or distributed Bragg reflection (DBR) lasers. *See* LASER.

Switches and modulators. Both lithium niobate and gallium arsenide belong to the family of electro-optically active crystals. When an electric field is applied to these materials, their refractive indices are modified. This effect is employed in integrated optical circuit switching and modulation applications. To construct a switch, gold or other conducting electrodes are deposited on a lithium niobate integrated optical circuit surface parallel to two closely spaced waveguides. If the electrodes and waveguides are suitably designed, the applications of specific small voltages to the electrodes will cause the transfer of optical power from one waveguide to its neighbor with high efficiency and little residual power in the initial guide. Shown in **Fig. 4** is a "4 by 4" switching network in which each of four input optical signals, possibly from optical fibers butt-coupled to the waveguide inputs, may be switched to any one of the four output ports. Such an integrated optical circuit serves to interconnect four computers through optical fibers, for example. *See* ELECTROOPTICS.

A switch is in effect a modulator as well. Modulation is a process in which information is encoded onto an optical or other electromagnetic wave. Pulse modulation results simply by interrupting or connecting a lightwave in a manner intelligible to the receiver. (Morse code is an unsophisticated example.) By transferring light into or out of a waveguide in response to an electric signal at a switching electrode, the output optical wave becomes modulated; that is, the switch operates as a modulator. Devices of these types are available commercially. *See* OPTICAL MODULATORS.

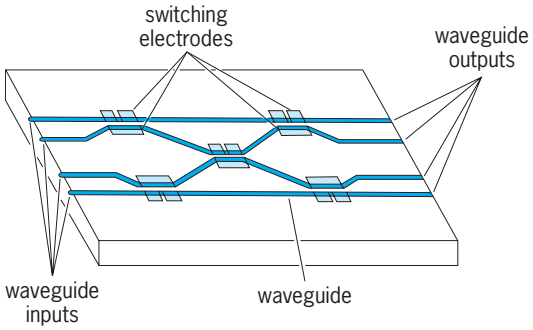

Fig. 4. A "4 by 4" directional coupler switch.

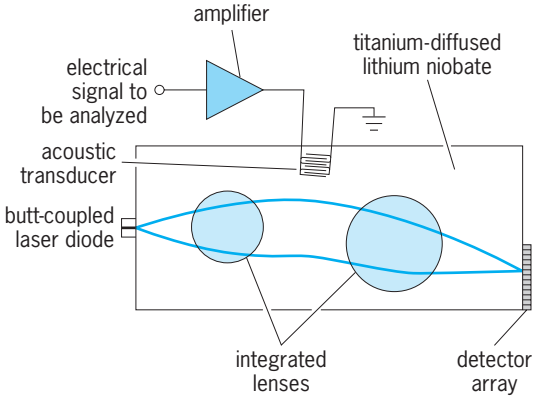

Fig. 5. Top view of an integrated optics spectrum analyzer.

Spectrum analyzer. Light propagating in an integrated optical circuit within a one-dimensional waveguide may be deflected by interacting with an acoustic wave. The deflection angle is a function of the acoustic frequency. This interaction forms the basis of an integrated optics spectrum analyzer (**Fig. 5**). An acoustic wave is generated by an electric signal whose frequencies are being analyzed. The acoustic wave interacts with a guided optical wave, and the deflected light is focused onto a detector array. The one or more detectors that respond determine the previously unknown frequencies. The integrated optical circuit incorporates lenses that focus the light and a transducer that generates the acoustic wave. The lenses may be fabricated by depositing additional high-refractive-index material in an area to decrease locally the propagation velocity of the guided wave. The transducer is made of a piezoelectric material (such as lithium niobate) which has the property of converting an electric field into a strain field. If lithium niobate is used for the integrated optical circuit, the transducer consists of a sophisticated interdigital electrode structure deposited on the waveguide surface. *See* ACOUSTOOPTICS; SPECTRUM ANALYZER.

Wavelength division multiplexing. A major application of integrated optics has taken place in the field of telecommunications, where signals are now mostly transmitted over long distances by modulated lightwaves guided by optical fibers, and OICs are often used for signal switching and processing. One of the key advantages of this technology is that a multiplicity of individual signals (telephone conversations or computer data transfers) can be transmitted simultaneously over a single optical fiber

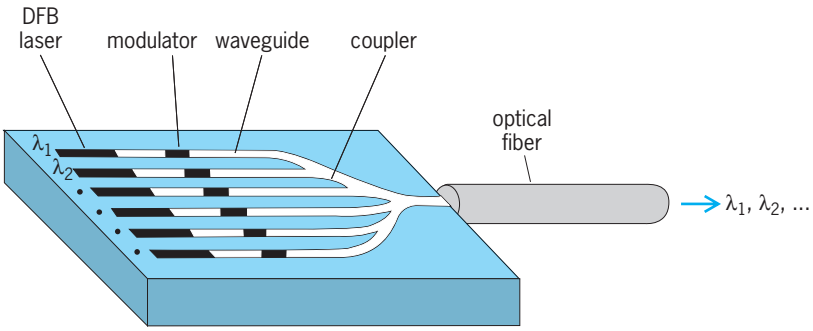

Fig. 6. A wavelength-division-multiplexed transmitter chip. DFB indicates distributed feedback.

without interfering with one another. This process is known as multiplexing. There are several different methods of multiplexing: time domain multiplexing (TDM), frequency division multiplexing (FDM), and wavelength division multiplexing (WDM). WDM is the most effective for lightwave signal transmission. The electromagnetic fields of lightwaves of different wavelength are orthogonal; that is, they do not interact. Thus, separate information channels operating at different wavelengths can be carried over a single fiber. Using increasingly sophisticated methods of WDM, the number of individual telephone channels carried by a single fiber has been increased from 1920 in 1980 to 240,000 in 2000. The signal capacity of an optical fiber is often described in terms of its bandwidth or digital data rate in gigabits per second (Gb/s). [A gigabit contains 1 billion (10^9) bits of digital information.] Currently, the better WDM systems have bandwidths of hundreds of gigabits per second transmitted over each fiber. They are called densely wavelength division multiplexed (DWDM) systems. *See* MULTIPLEXING AND MULTIPLE ACCESS.

The lightwave signals for WDM systems are usually generated by distributed feedback lasers designed to operate at different wavelengths that are accurately set by the periodicity of their reflective gratings. These lasers can be discrete devices or monolithically integrated into an OIC (**Fig. 6**). The signal information is impressed on the lightwave beam by a modulator, which can be of the electro-optic, acousto-optic, or electro-absorption type. All of these devices can be fabricated in waveguide-coupled, multilayer structures of GaAlAs or GaInAsP on a GaAs or InP substrate. At the receiving end, the multiplexed signal channels are separated by wavelength selective filters into individual waveguides, which lead to semiconductor photodiode detectors. This process is usually done by an OIC. The filters can be either gratings or array waveguides. The array waveguide (AWG) is a device that performs the spatial dispersion of light of different wavelengths in a manner similar to that of a prism. However, it is easier to fabricate in an OIC and is less sensitive to vibration than a prism. *See* FIBER-OPTIC CIRCUIT.

Photonic crystals. In both electronic integrated circuits (ICs) and optical integrated circuits (OICs), designers continually strive to decrease device size in order to increase the number of devices that can be fabricated on a single chip. High density of devices yields both improved performance and lower cost. However, electronic device density is reaching a fundamental limit because of inadequate heat dissipation. Conventional optical waveguides based on total internal reflection (TIR) are not subject to the heating effect of electrical current flowing through a wire, but they also have a fundamental limit in that tight bends in the waveguide cause excessive loss of the light. This has led researchers to investigate OICs based on photonic crystals. Photonic crystals guide light by the distributed Bragg reflection mechanism because of their periodic nature. This permits device sizes to approach that of the wavelength of the light, and 90° bends in waveguides are possible. A photonic crystal can be created in a dielectric material by drilling or etching a regularly spaced pattern of air holes that is analogous to the pattern of uniformly spaced atoms of a semiconductor crystal. Alternatively, an ordered array of columns or "fingers" can be fabricated in the dielectric by etching. The spacing of the holes or columns must be approximately the wavelength of the light in the dielectric material (~0.0001 mm). In that case, the photonic crystal responds to the guiding of light in the same way that a semiconductor crystal responds to the flow of electrons. There is an optical band gap over a range of wavelength (that is, photon energy) in which no light can propagate, just as there is a band gap in a semiconductor crystal over a range of energy in which there are no allowed electron states. It is possible to create energy levels within the photonic band gap by changing the size of some of the air holes or columns. This is the photonic equivalent of introducing dopant atoms into the semiconductor crystal. Thus, the properties of the photonic crystal can be modified to produce different devices. The critical parameters are the diameter of the air holes or columns and the contrast in refractive index. Because of the extremely small dimensions of photonic crystals, state-of-the-art semiconductor nanofabrication techniques—such as e-beam lithography, x-ray interference lithography, and ion beam etching—must be used to fabricate them. Photonic crystals have been used to make waveguides, couplers, splitters, and resonant cavities. *See* MICROLITHOGRAPHY.

Robert Hunsperger

Bibliography. J. T. Boyd (ed.), *Integrated Optics: Devices and Applications*, 1991; K. J. Ebeling, *Integrated Optoelectronics*, 1993; R. G. Hunsperger, *Integrated Optics: Theory and Technology*, 5th ed., 2002; S. I. Najafi (ed.), *Glass Integrated Optics and Optical Fiber Devices*, 1996; H. Nishihara, M. Haruna, and T. Suhara, *Optical Integrated Circuits*, 1989.

Integrated services digital network (ISDN)

A generic term referring to the integration of communications services transported over digital facilities such as wire pairs, coaxial cables, optical fibers, microwave radio, and satellites. ISDN provides

end-to-end digital connectivity between any two (or more) communications devices. Information enters, passes through, and exits the network in a completely digital fashion.

Characteristics. Since the introduction of pulse-code-modulation (PCM) transmission in 1962, the worldwide communications system has been evolving toward use of the most advanced digital technology for both voice and nonvoice applications. Pulse-code modulation is a sampling technique which transforms a voice signal with a bandwidth of 4 kHz into a digital bit stream, usually of 64 kilobits per second (kbps). *See* PULSE MODULATION.

Many aspects of telecommunications are improved with digital technology. For example, digital technology lends itself to very large-scale integration (VLSI) technology and its associated benefits of miniaturization and cost reduction. In addition, computers operate digitally. Digital transport provides for human-to-human, computer-to-computer, and human-to-computer interactions. The ISDN is capable of transporting voice, data, graphics, text, and even video information over the same equipment. *See* DIGITAL COMPUTER; INTEGRATED CIRCUITS.

Service integration and standard interface. The customer has access to a wide spectrum of communications services by way of a single access link. This is in contrast to existing methods of service access, which segregate services onto specialized lines (see **illus.**).

Associated with integrated access and ISDN is the concept of a standard interface. The objective of a standard interface is to allow any ISDN terminal to be plugged into any ISDN interface, resulting in terminal portability, flexibility, and ease in operation. The benefits of standard interfaces are readily visible in the United States electrical industry, where almost any electrical outlet can be used by any electrical device.

The customer's perspective of the service is through the interface. In fact, ISDN is recognized solely by the service characteristics (performance, bandwidth, and protocol) offered at the user-network boundary, rather than by any internal network components.

Signaling. ISDN provides a signaling control channel. This is a common channel, transported out of band, that provides advanced digital signaling capabilities.

Digital signaling allows computers, which do not understand voice messages, to react to the varied states which occur in a typical call. Furthermore, the time required to set up a call is reduced to the order of a few seconds, and additional messages are easily introduced when required for new service applications.

Out-of-band signaling means that the messages do not use the same channel as the customer information but are carried independently. This allows messages to be sent to the users while the call is active without interfering with the call. Thus, new services are available. For example, the identity of the calling party, sent to the called party in machine-readable form, can prevent invalid users from accessing private computers. Out-of-band signaling also enables functions to be performed at the central office much more rapidly. *See* COMPUTER SECURITY.

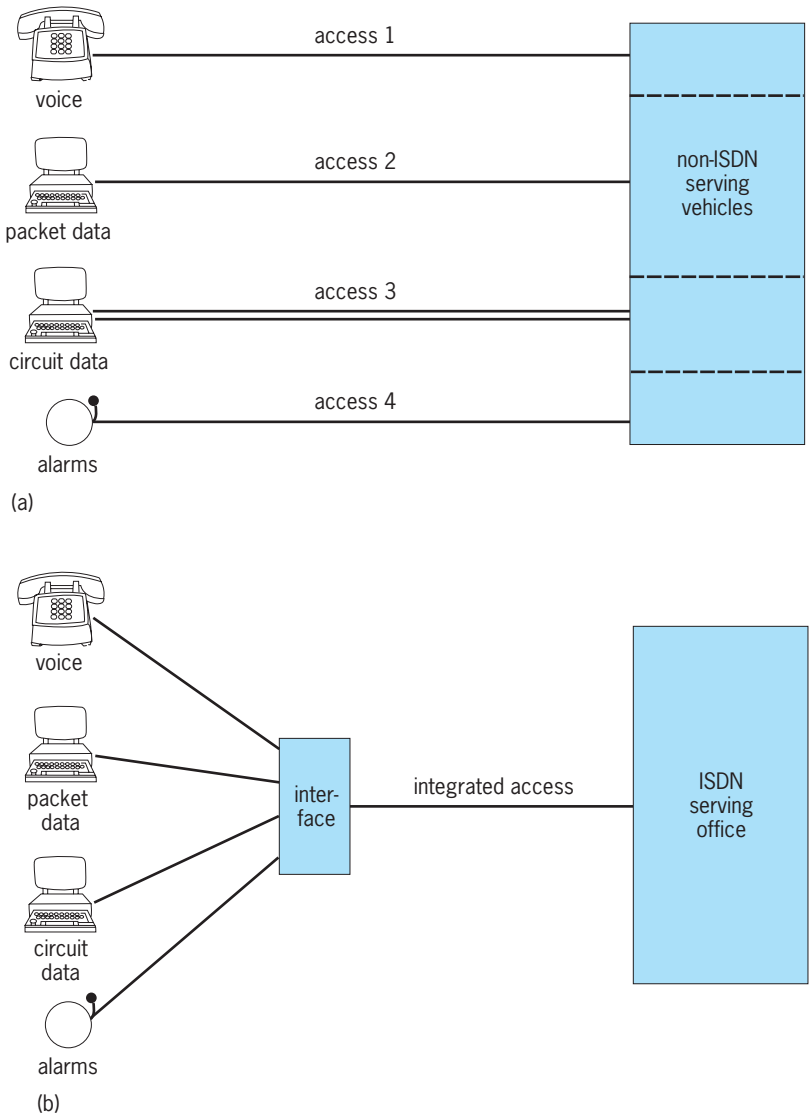

Comparison of segregated and integrated access. (*a*) For segregated access customer-network interfaces differ whereas (*b*) for integrated access there is one customer interface.

The signaling between switching centers in an ISDN is carried on a separate digital data network using packet-switching techniques. Not only is the call setup time decreased, but also additional information can be exchanged between originating and terminating offices for such service enhancements as caller identity and ring back.

Standards. In order to establish ISDN, it is necessary to define a set of interfaces between the customer premises equipment and the telecommunications transport network. An important goal is to make these interfaces internationally compatible. To realize this, an international team of experts in customer premises equipment and telecommunications planners has been formed under the auspices of the International Telegraph and Telephone Consultative Committee (CCITT).

Interfaces. Currently, two ISDN interfaces are defined in the CCITT, each an outgrowth of existing capabilities. Since the existing network encodes voice into a 64-kbps stream, the building block of ISDN is a transmission rate of 64 kbps.

The basic rate interface operates at 144 kbps, providing two 64-kbps channels plus a 16-kbps packet-oriented digital channel. The 64-kbps channels, referred to as B channels (for bearer channels), transport customer information, either voice or data. The 16-kbps channel, referred to as a D channel (for data channel), carries signaling and packet data.

The primary rate interface at 1.544 megabits per second (Mbps) can provide access to digital private branch exchanges (PBXs) or other large communications systems. This interface consists of twenty-four 64-kbps channels, and is an outgrowth of existing multiplexed digital carrier communication. One of these 64-kbps channels, called a D channel, operates as a signaling and packet data channel. The other twenty-three channels are B channels, as in the basic access. This interface is flexible in that several of the 64-kbps B channels can be combined to provide higher-speed services such as 384 kbps.

Additional work is under way in the CCITT to standardize bearer service subrates, such as 8, 16, and 32 kbps. Thus, a wide spectrum of customer data rates will be provided for numerous applications. ISDN is designed to provide bandwidth on demand, which means that the user will be able to specify the data rate desired on a call-by-call basis. Also, since different applications require different performance levels, performance on demand will be another key component of ISDN.

The channel designation H is used for bit rates from 384 kbps up to 45 Mbps. This is known as wideband ISDN.

As end-to-end digitalization of networks becomes commonplace at these rates, a need is expected to emerge for the ability to transmit digital information at even higher rates. Optical fiber lines have been installed that operate at 1.2×10^9 bits per second. At these rates, applications that require large amounts of information, such as high-definition real-time motion pictures or television, become practical.

Communications network. In addition to integrated-access transport from the customer to the switching node, ISDN must provide digital switching, interoffice transport, and signaling functions. Switching includes circuit switching and packet switching.

Circuit switching is used in the existing voice network where end-to-end paths are dedicated to a single call during the call duration. Packet switching is intended for data calls which, unlike voice, tend to be "bursty" in information content. Packet-switched networks bundle customer data into self-contained messages that use the transport facility only during the transmission of the packet and share the medium with packets from other users. *See* DATA COMMUNICATIONS; PACKET SWITCHING.

B channels are carried over the integrated-access line to the ISDN serving office. Here, they are routed over the circuit-switched network or are circuit-switched to a packet-switched network.

The D channel is the integrated-access control link between the customer and the ISDN serving office, ending at the office. It carries signaling information and packet data. The packet data are routed to a packet-switched network, while the signaling information is either interpreted by the local switch or sent into the network over a signaling network.

A method called asynchronous transfer mode (ATM) packages digital information to be transmitted into uniform cells of 53 bytes. These cells are then transmitted over appropriate media at will, a procedure known technically as asynchronous transmission. Each cell includes address information to guide it to its ultimate destination. Cells from diverse sources with different priorities of transmission can be mixed serially on the same media. While the basic transmitted element is a packet-like cell, switching for ATM is a mixture of packet and circuit techniques. *See* ELECTRICAL COMMUNICATIONS; SWITCHING SYSTEMS (COMMUNICATIONS).

Robert M. Wienski; Amos E. Joel, Jr.

Bibliography. G. C. Kessler, *ISDN: Concepts, Facilities, and Services*, 2d ed., 1993; B. Kumar, *Broadband Communications: A Professional's Guide to ATM, Frame Relay, SMDS, SONET, and B-ISDN*, 1995; W. Stallings, *ISDN and Broadband ISDN with Frame Relay and ATM*, 3d ed., 1994; C. K. Summers, *ISDN Implementor's Guide: Standards, Protocols and Services*, 1995.

Integration

An operation of the infinitesimal calculus which has two aspects. The roots of one go back to antiquity, for Archimedes and other Greek mathematicians used the "method of exhaustion" to compute areas and volumes. A simple example of this is the approximation to the area of a circle obtained by inscribing a regular polygon of known area, and then repeatedly doubling the number of sides. The areas of the successive polygons are computable with the help of elementary geometry. The limit of the sequence of these areas gives the area of the circle. The area of each polygon can be regarded as being made up of the sum of the areas of triangles with vertices at the center of the circle, and so the process described is a constructive definition of an integral which is the limit of a sum. Modern definitions of integrals as limits of sums are discussed in this article.

The other aspect of integration is the process of finding antiderivatives, that is, for a given function $f(x)$ to find another function $g(x)$ whose derivative is $f(x)$. The aspect is related to the first by the fundamental theorem of integral calculus, so both processes are called integration.

Isaac Newton emphasized the antiderivative aspect of integration, and his work shows how much can be done in the applications of integral calculus without introducing limits of sums. However,

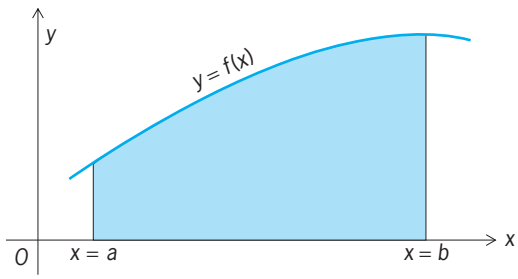

Fig. 1. Graph of $y = f(x)$.

limits of sums lead to very fruitful theoretical developments in the theory of integration, as in the notion of multiple integrals, for example, and hence lead to a wider variety of applications. Leibnitz, a seventeenth-century mathematician, inspired the development in this direction, but many years elapsed before the theory was given a firm logical foundation. In the early nineteenth century A. L. Cauchy gave a clear-cut definition of the definite integral for continuous functions and a proof of its existence. Later, G. F. B. Riemann discussed the integral for discontinuous functions and gave a necessary and sufficient condition for its existence. Thus the most generally used definition of the integral as the limit of a sum has come to be called the Riemann integral.

Riemann integral. The precise definition of the Riemann integral for a real function f of one real variable x on a finite interval $a \leqq x \leqq$ may be formulated as follows. Let P be a partition of the interval $[a, b]$ into n subintervals by points t_i, where $t_{i-1} < t_i$, $t_0 = a$, $t_n = b$, and consider a sum S of the form of Eq. (1),

$$S = \sum_{i=1}^{n} f(x_i)(t_i - t_{i-1}) \tag{1}$$

where $t_{i-1} \leqq x_i \leqq t_i$. The sum S depends not only on the partition P but on the choice of the intermediate points x_i. It may happen that the sum S approaches a definite limit I when the maximum of the numbers $(t_i - t_{i-1})$ tends to zero, and in this case I is called the Riemann integral (or the definite integral) of f from a to b, and is denoted by the Leibnitzian symbol (2).

$$\int_a^b f(x)\,dx \tag{2}$$

Also, f is said to be integrable on $[a,b]$. When f is a continuous function with positive values on the interval $[a,b]$, the integral has a simple geometrical interpretation as the area bounded by the x axis, the ordinates $x = a$ and $x = b$, and the graph of $y = f(x)$ [**Fig. 1**]. It is convenient to use Eq. (3) as a definition

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx \tag{3}$$

of the symbol on the left.

For functions f and g which are integrable on $[a, b]$, the integral has the following properties:

(i) f is integrable on every subinterval of $[a, b]$.

(ii) For every triple of points c, d, and e in $[a, b]$,

$$\int_c^d f(x)\,dx + \int_d^e f(x)\,dx = \int_c^e f(x)\,dx$$

(iii) $f(x) + g(x)$ is integrable, and

$$\int_a^b [f(x) + g(x)]\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$$

(iv) $f(x)g(x)$ is integrable, and in particular $cf(x)$ is integrable for every real number c, and

$$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx$$

(v) If $f(x) \leqq g(x)$ on $[a,b]$, then

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx$$

(vi) $|f(x)|$ is integrable, and

$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx$$

It can be proved that a function f is integrable on $[a, b]$ if and only if the following two conditions are satisfied: f is bounded on $[a, b]$; and the set of points where f is discontinuous can be enclosed in a series (possibly infinite) of intervals, the sum of whose lengths is arbitrarily near to zero.

Antiderivatives. To develop the fundamental theorem of integral calculus let u be a variable in the interval $[a, b]$, on which f is integrable; then formula (4) defines a function of u which may be noted by

$$\int_a^u f(x)\,dx \tag{4}$$

$I(u)$. If f is continuous on $[a, b]$, then f is integrable, and it is also true that $I(u)$ has a derivative $I'(u) = f(u)$. Now let h be any antiderivative of f, that is, $h'(u) = f(u)$ on $[a, b]$. Then $I'(u) - h'(u) = 0$, so $I(u) - h(u) = \text{constant} = -h(a)$, by the theorem of the mean for derivatives and the fact that $I(a) = 0$, so relation (5) can be written. This is the fundamental theorem of integral calculus, and it shows that definite integrals may be calculated by the process of finding antiderivatives. For this reason antiderivatives are frequently called indefinite integrals and denoted by $\int f(x)\,dx$, and special methods of finding indefinite integrals for frequently occurring functions occupy a large part of elementary calculus. The principal methods are outlined in the next section. The standard notation for an indefinite integral of $f(x)$ is formula (6).

$$I = I(b) = h(b) - h(a) \tag{5}$$

$$\int f(x)\,dx \tag{6}$$

Elementary methods of integration. Obviously, each formula for differentiation yields a formula for indefinite integration.

Method of substitution. From these indefinite integrals, additional formulas can be obtained by the method of substitution, which is based on the chain rule for differentiating composite functions. For example, from formula (7) one obtains Eq. (8), by the substitution $v = a^2 + x^2$, and Eq. (9), by the substitution $v = \sin x$.

$$\int x^m dx = \frac{x^{m+1}x}{m+1}, m \neq -1 \tag{7}$$

$$\int (a^2 + x^2)^m x\, dx = \frac{(a^2 + x^2)^{m+1}}{2(m+1)}, m \neq -1 \tag{8}$$

$$\int \sin^m x \cos x\, dx = \frac{\sin^{m+1} x}{m+1}, m \neq -1 \tag{9}$$

Method of partial fractions. In case $f(x)$ is a quotient of two polynomials in x, $f(x)$ may be represented as a sum of a polynomial and terms of the form, shown as notations (10) and (11), where m is a positive integer and the coefficients are real. This is called the method of partial fractions. Its actual execution requires a knowledge of the factors of the denominator of $f(x)$. Terms of the form (10) have known indefinite integrals. Terms of the form (11) with $m = 1$ can also be integrated by elementary formulas. When $m > 1$, a term of the form (11) can be written as a sum of terms of the forms (12) and (13),where $\alpha =$

$$\frac{1}{(ax+b)^m} \qquad a \neq 0 \tag{10}$$

$$\frac{Ax+B}{(cx^2+dx+e)^m} \qquad c \neq 0, d^2 - 4ce < 0 \tag{11}$$

$$\frac{\alpha}{(\gamma x^2 + dx + e)^{m-1}} \tag{12}$$

$$\frac{(\beta x + \gamma)(2cx + d)}{(cx^2 + dx + e)^m} \tag{13}$$

$(2dA - 4cB)/(d^2 - 4ce)$, $\beta = -\alpha/2$ and $\gamma = (dB - 2eA)/(d^2 - 4ce)$. By the process of integration by parts, the indefinite integral of a term of the form (13) can be expressed in terms of the integral of a term such as (11), with m replaced by $m - 1$. Reduction formulas of this type are given in standard tables of integrals.

A rational function of $\sin x$ and $\cos x$ can be integrated by the method just described after the substitution $u = \tan(x/2)$ has been applied. This substitution reduces the problem to the integration of a rational function of u. Various functions involving radicals can be integrated by means of trigonometric substitutions or by other substitutions which reduce the problem to the integration of a rational function. For example, by setting $u = a + bx$, one finds Eq. (14),

$$\int x^m (a+bx)^p\, dx = \int \left(\frac{u-a}{b}\right)^m u^p \frac{du}{b} \tag{14}$$

and when m is a nonnegative integer the right side can be multiplied out and integrated by the power formula, even though p is a fraction.

Integration by parts. This is a very powerful and important method. It follows from the formula for differentiating a product, namely Eq. (15). After the terms

$$d(fg) = f\, dg + g\, df \tag{15}$$

are rearranged and integrated, Eq. (16) or (17) may

$$\int f\, dg = fg - \int g\, df \tag{16}$$

$$\int fg'\, dx = f(x)g(x) - \int gf'\, dx \tag{17}$$

be written, where the arbitrary constant, as usual, has been omitted. As an example, consider Eq. (18),

$$I = \int \frac{2x^2 dx}{(x^2+1)^m} \tag{18}$$

which is a special case of the form (13). Let $f = x$, $g' = 2x/(x^2+1)^m$. Then $f' = 1$, and g can be written as Eq. (19), and by Eq. (17), Eq. (20) follows.

$$g = \frac{1}{(1-m)(x^2+1)^{m-1}} \tag{19}$$

$$I = \frac{x}{(1-m)(x^2+1)^{m-1}} - \frac{1}{(1-m)} \int \frac{dx}{(x^2+1)^{m-1}} \tag{20}$$

The indefinite integrals of many of the commonly occuring functions are given in tables of integrals in handbooks and textbooks.

The elementary functions are those expressible by means of a finite number of algebraic operations and trigonometric and exponential function and their inverses. The integrals of many elementary functions are known to be not elementary, so they define new functions. A number of these are sufficiently important that their values have been tabulated. Examples are Eqs. (21) and (22). $F(k, x)$ is called an elliptic integral of the first kind. The inverse of $u = F(k, x)$ is an elliptic function called the sine amplitude of u and denoted by $x = \text{sn}\, u$. Si x is called the sine integral. Some nonelementary functions have been included in tables of integrals. *See* ELLIPTIC FUNCTION AND INTEGRAL.

$$F(k, x) = \int_0^x \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}} \tag{21}$$

$$(k^2 \neq 1)$$

$$\text{Si}\, x = \int_0^x \frac{\sin x\, dx}{x} \tag{22}$$

Improper integrals. This term is used to refer to an extension of the notion of definite integral to cases in which the integrand is unbounded or the domain of integration is unbounded. Consider first the case when $f(x)$ is integrable (in the sense defined above) and hence bounded on every interval $[a + \epsilon, b]$ for $\epsilon > 0$, but is unbounded on $[a,b]$. Then by definition, Eq. (23) is written, provided the limit on the right exists. For example, if $f(x) = x^{-2/3}$, relation (24)

$$\int_a^b f(x)\, dx = \lim_{\epsilon \to 0} \int_{a+\epsilon}^b f(x)\, dx \tag{23}$$

follows. Similarly, Eq. (25) follows. When $f(x)$ is in-

$$\int_0^1 x^{-2/3}\,dx = \lim_{\epsilon\to 0}\int_\epsilon^1 x^{-2/3}\,dx$$

$$= \lim_{\epsilon\to 0} 3(1-\epsilon^{1/3}) = 3 \qquad (24)$$

$$\int_{-1}^0 x^{-2/3}\,dx = \lim_{\epsilon\to 0}\int_{-1}^{-\epsilon} x^{-2/3}\,dx$$

$$= \lim_{\epsilon\to 0} 3[(-\epsilon)^{1/3}+1] = 3 \qquad (25)$$

tegrable on every finite subinterval of the real axis, by definition Eqs. (26) and (27) hold, provided the

$$\int_a^{+\infty} f(x)\,dx = \lim_{b\to+\infty}\int_a^b f(x)\,dx \qquad (26)$$

$$\int_{-\infty}^a f(x)\,dx = \lim_{b\to-\infty}\int_b^a f(x)\,dx \qquad (27)$$

limits on the right exist. More general cases are treated by dividing the real axis into pieces, each of which satisfies one of the conditions just specified. Equation (28) is an example. Because the second and

$$\int_{-\infty}^{+\infty} x^{-5/3}\,dx = \lim_{a\to-\infty}\int_a^{-1} x^{-5/3}\,dx + \lim_{\delta\to 0}\int_{-1}^{-\delta} x^{-5/3}\,dx + \lim_{\epsilon\to 0}\int_\epsilon^1 x^{-5/3}\,dx + \lim_{b\to+\infty}\int_1^b x^{-5/3}\,dx \qquad (28)$$

third of the limits on the right do not exist, integral (29) does not exist. However, it is sometimes use-

$$\int_{-\infty}^{+\infty} x^{-5/3}dx \qquad (29)$$

ful to assign a value to it, called the Cauchy principal value, by replacing definition (28) by Eq. (30). In this

$$\int_{-\infty}^{+\infty} x^{-5/3}\,dx = \lim_{e\to 0}\left(\int_{-1}^{-\epsilon} x^{-5/3}\,dx + \int_\epsilon^1 x^{-5/3}\,dx\right) + \lim_{b\to+\infty}\left(\int_{-b}^{-1} x^{-5/3}\,dx + \int_1^b x^{-5/3}\,dx\right) \qquad (30)$$

particular case the value is zero.

The preceding definitions apply to cases when the integrand $f(x)$ is bounded, except in arbitrarily small neighborhoods of a finite set of points. In the closing years of the nineteenth century various extensions were made to more general cases. Then Henri Lebesgue produced a comparatively simple general theory for the case of absolutely convergent integrals. The integral of Lebesgue will be discussed below. The integral of Denjoy includes both nonabsolutely convergent integrals and the integral of Lebesgue.

Multiple integrals. The concept called the Riemann integral can be extended to functions of several variables. The case of a function of two variables illustrates sufficiently the additional features which arise. To begin with, let $f(x, y)$ denote a real function defined on a rectangle of the form (31). Let P be a

$$R: a \le x \le b \qquad c \le y \le d \qquad (31)$$

partition of R into n nonoverlapping rectangles R_i with areas A_i, and let (x_i, y_i) be a point of R_i. Define S by Eq. (32). In case the sum S tends to a definite

$$S = \sum_{i=1}^n f(x_i, y_i)A_i \qquad (32)$$

limit I when the maximum diagonal of a rectangle R_i tends to zero, then f is said to be integrable over R, and the limit I is called the Riemann integral of f over R. It will be denoted here by the abbreviated symbol $\int\int_R f$. Properties corresponding to those numbered (i) to (iv) can be proved for these double integrals, except that property (ii) should now read as follows:

(ii′) If the rectangle R is the union of two rectangles S and T, then Eq. (33) can be written. Necessary and

$$\int\int_R f = \int\int_S f + \int\int_T f \qquad (33)$$

sufficient conditions for a function $f(x, y)$ to be integrable over R can be stated as follows: f is bounded on R; and the set of points where f is discontinuous can be enclosed in a series of rectangles, the sum of whose areas is arbitrarily near to zero.

To define the integral of a function $f(x, y)$ over a more general domain D where it is defined, suppose D is enclosed in a rectangle R, and define $F(x, y)$ by Eqs. (34). Then f is integrable over D if case F is

$$\begin{aligned} F(x,y) &= f(x,y) \text{ in } D \\ F(x,y) &= 0 \text{ outside } D \end{aligned} \qquad (34)$$

integrable over R, and Eq. (35) holds, by definition.

$$\int\int_D f = \int\int_R F \qquad (35)$$

When the function $f(x, y)$ is continuous on D, and D is defined by inequalities of the form (36), where

$$\begin{aligned} a \le x \le b \\ a(x) \le y \le \beta(x) \end{aligned} \qquad (36)$$

the functions $\alpha(x)$ and $\beta(x)$ are continuous, the double integral of f over D always exists, and may be represented in terms of two simple integrals by Eq. (37).

$$\int\int_D f = \int_a^b \left[\int_{\alpha(x)}^{\beta(x)} f(x,y)\,dy\right] dx \qquad (37)$$

In many cases, this formula makes possible the evaluation of the double integral.

Improper multiple integrals have been defined in a variety of ways. There is not space to discuss these definitions and their relations here.

Line, surface, and volume integrals. A general discussion of curves and surfaces requires an extended treatise. The following outline is restricted to the simplest cases.

A curve may be defined as a continuous image of an open interval. Thus a curve C in the xy plane is given by a pair of continuous functions of one variable, Eq. (38), whereas a curve in space is given by

$$x = f(u) \quad y = g(u) \quad a < u < b \tag{38}$$

a triple of such functions. The mapping (38) is more properly called a representation of a curve, and a curve is then defined as a suitable class of such representations. Note that a curve such as Eq. (38) is a path rather than a locus. For example, notation (39) is the path going twice around the unit circle, Eq. (40),

$$x = \cos u \quad y = \sin u \quad 0 < u < 4\pi \tag{39}$$

$$x^2 + y^2 = 1 \tag{40}$$

in the counterclockwise direction. A curve is called smooth in case the function f and g have continuous derivatives $f'(u)$ and $g'(u)$ which are never simultaneously zero.

The length of a smooth curve C may be defined by Eq. (41) when this integral exists (as a proper or

$$L(C) = \int_a^b \sqrt{[f'(u)]^2 + [g'(u)]^2 du} \tag{41}$$

improper integral). When the integral does not exist, the length of C is said to be infinite. This definition of length in terms of the integral may be shown to coincide with the geometric definition in a much larger class than the class of smooth curves. When C has finite length, the position vector $[f(u), g(u)]$ approaches definite limits when u tends to a or to b, which are the ends of the curve C. A smooth curve has at every point a nonzero tangent vector $\mathbf{T}(u)$, with components $f'(u)$ and $g'(u)$, so that the integrand of $L(C)$ is the length $|\mathbf{T}(u)|$ of this vector. Thus one may write $ds = |\mathbf{T}(u)|\, du$, where s is the arc length measured from some convenient point on C.

Let $A(x, y)$ be a bounded continuous function defined on a set containing a plane curve C having finite length. Then A determines three functions of C called line integrals, whose symbols and definitions, Eqs. (42), follow. If $-C$ denotes the curve

$$\begin{aligned} \int_C A\, dx &= \int_a^b A[f(u), g(u)]f'(u)\, du \\ \int_C A\, dy &= \int_a^b A[f(u), g(u)]g'(u)\, du \\ \int_C A\, ds &= \int_a^b A[f(u), g(u)]\, |T(u)|\, du \end{aligned} \tag{42}$$

C traversed in the opposite direction, as shown in Eqs. (43), a reversal of the orientation of C changes

$$\begin{aligned} \int_{-C} A\, dx &= -\int_C A\, dx \\ \int_{-C} A\, dy &= -\int_C A\, dy \\ \int_C A\, ds &= \int_C A\, ds \end{aligned} \tag{43}$$

the signs of the first two but not of the third of these functions.

The extension of the preceding definitions to curves in three-space is made in an obvious way. An important application of line integrals is to express the work done by a force field on a moving particle. Thus, if a force field $\mathbf{F}(x, y, z)$ has components $F_1(x, y, z)$, $F_2(x, y, z)$, and $F_3(x, y, z)$ in the directions of the x, y, and z axes, then the work done by the field on a particle moving on a curve C is given by formula (44). When $\mathbf{T}$ is the tangent vector of C, expressed

$$W = \int_C F_1 dx + F_2 dy + F_3 dz \tag{44}$$

in terms of a parameter u, the work W can be expressed in terms of the dot product by Eq. (45). If $\mathbf{T}_1$ denotes the tangent vector of unit length, notation (46) then can be written.

$$W = \int_a^b \mathbf{F} \cdot \mathbf{T}\, du \tag{45}$$

$$W = \int_C \mathbf{F} \cdot \mathbf{T}_1\, ds \tag{46}$$

To pass from curves to surfaces replace the open interval $a < u < b$ by a bounded connected open set D in a uv plane, which may be called the parameter plane. The domain D can be restricted to be of sufficiently simple shape so that every function which is continuous and bounded in D is integrable over D, as a multiple integral. For example, D may be the interior of a circle, or of a rectangle. Then a surface S in three-space is defined by a triple of functions continuous on D, Eqs. (47). Such a surface is called smooth

$$x = f(u, v) \quad y = g(u, v) \quad z = h(u, v) \tag{47}$$

in case the functions f, g, and h have continuous first partial derivatives in D, and the three Jacobians listed in formulas (48), are never simultaneously zero. A

$$J_1 = \frac{\partial(y, z)}{\partial(u, v)} \quad J_2 = \frac{\partial(z, x)}{\partial(u, v)} \quad J_3 = \frac{\partial(x, y)}{\partial(u, v)} \tag{48}$$

smooth surface has at every point a nonzero normal vector $\mathbf{J}(u, v)$ with components J_1, J_2, J_3, and length $|\mathbf{J}|$.

The area of a smooth surface S may be defined by the integral (49), and it is always finite when the

$$\sigma(S) = \int\int_D |\mathbf{J}(u, v)| \tag{49}$$

vector $\mathbf{J}$ has bounded length.

If $A(x, y, z)$ is a bounded continuous function defined on a set containing a surface S with finite area, four surface integrals can be defined as in notation

(50). If $\mathbf{F}(x,y,z)$ is a vector field with components

$$\begin{aligned}
&\int\int_S A\,dy\,dz = \\
&\qquad \int\int_D A[f(u,v), g(u,v), h(u,v)]J_1(u,v) \\
&\int\int_S A\,dz\,dx = \\
&\qquad \int\int_D A[f(u,v), g(u,v), h(u,v)]J_2(u,v) \\
&\int\int_S A\,dx\,dy = \\
&\qquad \int\int_D A[f(u,v), g(u,v), h(u,v)]J_3(u,v) \\
&\int\int_S A\,d\sigma = \\
&\qquad \int\int_D A[f(u,v), g(u,v), h(u,v)]|\mathbf{J}(u,v)
\end{aligned} \tag{50}$$

F_1, F_2, F_3, and $\mathbf{n}$ denotes the unit vector normal to S, with components $J_1/|\mathbf{J}|, J_2/|\mathbf{J}|, J_3/|\mathbf{J}|$, then a combination, formula (51), of the first three kinds of surface

$$\int\int_S F_1\,dy\,dz + F_2\,dz\,dx + F_3\,dx\,dy \tag{51}$$

integral may also be written in the form of expression (52),where $\mathbf{F}\cdot\mathbf{n}$ is the dot product. Expression

$$\int\int_S \mathbf{F}\cdot\mathbf{n}\,d\sigma \tag{52}$$

(51) or (52) is referred to as the integral of the vector $\mathbf{F}$ over the side of the surface S to which the vector $\mathbf{n}$ points. The integral over the opposite side of S is denoted by formula (53), whose value is the negative of that in (51), by Eq. (54). The last integral in

$$\int\int_S F_1\,dz\,dy + F_2\,dx\,dz + F_3\,dy\,dx \tag{53}$$

$$\frac{\partial(z,y)}{\partial(u,v)} = -J_1, \text{ and so on} \tag{54}$$

Eq. (50), however, is independent of a choice of side of the surface.

The important formulas of Stokes and of Gauss, Eqs. (55), are expressed in terms of line and surface

$$\begin{aligned}
&\int_C \mathbf{F}\cdot\mathbf{T}\,ds = \int\int_S \text{curl}\,\mathbf{F}\cdot\mathbf{n}\,d\sigma \quad \text{(Stokes)} \\
&\int\int_S \mathbf{F}\cdot\mathbf{n}\,d\sigma = \int\int\int_V \text{div}\,\mathbf{F} \quad \text{(Gauss)}
\end{aligned} \tag{55}$$

integrals and the triple or volume integral analogous to the double integral. In the formula of Stokes, $\mathbf{F}$ is a vector field, S is a smooth surface having a piecewise smooth boundary curve C, directed so that S lies to the left of an observer proceeding along C on the side of S on which the unit normal vector $\mathbf{n}$ is chosen, and $\mathbf{T}$ is the unit tangent vector to C. In terms of the components of F, Stokes' formula may be written as Eq. (56), where $F_{3y} = \partial F_3/\partial\gamma$, and so on. When the

$$\begin{aligned}
\int_C F_1\,dx + F_2\,dy + F_3\,dz &= \int\int_S (F_{3y} - F_{2z})\,dy\,dz \\
&+ (F_{1z} - F_{3x})\,dz\,dx + (F_{2x} - F_{1y})\,dx\,dy
\end{aligned} \tag{56}$$

surface S lies in the xy plane, this equation reduces to Green's formula, Eq. (57). The latter is readily proved

$$\int_C F_1 dx + F_2 dy = \int\int_S (F_{2x} - F_{1y})\,dx\,dy \tag{57}$$

for regions S of simple shape, and from it Stokes' formula may be deduced.

In the formula of Gauss (also called the divergence theorem), V is a space domain bounded by a smooth or piecewise smooth surface S, and $\mathbf{n}$ is the exterior normal of S. In terms of the components of F, Gauss' formula may be written as Eq. (58).

$$\begin{aligned}
&\int\int_S F_1\,dy\,dz + F_2\,dz\,dx + F_3\,dx\,dy \\
&\qquad = \int\int\int_V (F_{1x} + F_{2y} + F_{3z})
\end{aligned} \tag{58}$$

See CALCULUS OF VECTORS.

Functions defined by integrals. Mention has already been made of elliptic functions, which may be defined as the inverse functions of elliptic integrals. Certain definite integrals may also be used to define important nonelementary functions which occur frequently in applications. For example, the gamma function may be defined by Eq. (59) for $x > 0$. The

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1}\,dt \tag{59}$$

Bessel functions of integral order may be defined by Eq. (60) for all values of x. The properties of these

$$J_n(x) = \frac{1}{\pi}\int_0^\pi \cos(nt - x\sin t)\,dt \tag{60}$$

functions may be derived from these expressions with the help of various methods of advanced analysis, such as the theory of functions of a complex variable. *See* BESSEL FUNCTIONS; GAMMA FUNCTION.

In general, if $f(x, t)$ is an integrable function of t on $a \leqq t \leqq b$ for each x, then Eq. (61) is a well-

$$g(x) = \int_a^b f(x,t)\,dt \tag{61}$$

defined function of x. When suitable conditions are satisfied by the function f, the derivative $g'(x)$ may be calculated by formula (62). This process is called

$$g'(x) = \int_a^b \frac{\partial}{\partial x} f(x,t)\,dt \tag{62}$$

differentiation under the integral sign. It is valid in case f and $\partial f/\partial x$ are continuous in (x,t) for $c \leqq x \leqq d$, $a \leqq t \leqq b$, where a and b are finite, and even in more

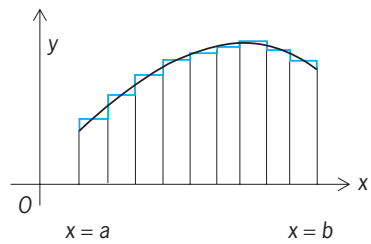

Fig. 2. Division of interval [*a*, *b*] into parts.

general cases. However, in the simple example given in Eqs. (63), formula (62) leads to a wrong result.

$$f(x, t) = x^3 e^{-tx^2}$$
$$g(x) = \int_0^\infty f(x, t)\, dt \tag{63}$$

Approximate and mechanical integration. The definition of the definite integral itself gives a means of calculating its value approximately. There remains the question of a suitable selection of the functional values $f(x_i)$ for use in formula (1). It is usual to divide the interval $[a, b]$ into a number n of equal parts of length h (**Fig. 2**). If the points x_i are taken at the midpoints of the subintervals, one obtains the midpoint formula, Eq. (64). For a curve that is concave downward, this gives too large a value.

$$S = h \sum_{i=1}^{n} f[a + (i - 1/2)h] \tag{64}$$

Another formula, called the trapezoidal rule, is derived by calculating the area below a polygon inscribed in the graph of $f(x)$ [**Fig. 3**]. This formula is Eq. (65). It gives too small a value for a curve that is concave downward.

$$T = \frac{h}{2}\left[f(a) + f(b) + 2\sum_{i=1}^{n-1} f(a + ih)\right] \tag{65}$$

The error in the approximation may sometimes be reduced without increasing the number of subintervals by use of the parabolic rule (Simpson's rule). To obtain the formula, the subarcs of the graph of $f(x)$ are replaced by arcs of parabolas rather than by line segments. Since three points determine a parabola (with vertical axis), the interval $[a, b]$ is divided into an even number n of subintervals. The area under the parabola passing through the points on the graph of $f(x)$ having abscissas given in expressions (66) is

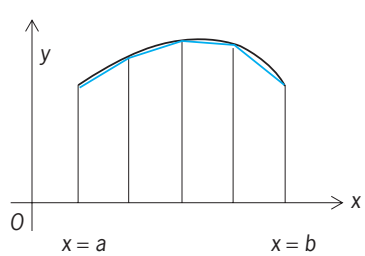

Fig. 3. Polygon inscribed in graph of *f*(*x*).

given by expression (67). Therefore, the parabolic rule gives approximation (68).

$$a + (2i - 2)h \quad a + (2i - 1)h \quad a + 2ih \tag{66}$$

$$\frac{h}{3}\{f[a + (2i - 2)h] + 4f[a + (2i - 1)h] + f(a + 2ih)\} \tag{67}$$

$$P = \frac{h}{3}\left\{f(a) + f(b) + 2\sum_{i=1}^{(n/2)-1} f(a + 2ih) + 4\sum_{i=1}^{n/2} f[a + (2i - 1)h]\right\} \tag{68}$$

When a function $f(x)$ is given only by a table, its integral is most simply computed by one of these formulas or a similar formula.

When a function $f(x)$ is given by a graph, mechanical means may be used to calculate associated areas. One such means is the polar planimeter, which registers on a rotating wheel the area enclosed by a closed curve around which a tracing point is passed. The integraph, invented by Abdank Abakanowicz, is designed to draw the graph of an indefinite integral of $f(x)$ when a tracing point is passed over the graph of $f(x)$. These simple devices were the forerunners of more complex machines designed to solve differential equations, such as the differential analyzer of Vannevar Bush. For large-scale computations involving formulas such as (64), (65), or (68), a digital computer may be preferred.

Other methods of integration. The method of differentiation under the integral sign is often convenient for the evaluation of definite integrals, even when other methods are available. For example, Eq. (69) leads to Eq. (70). Or, if one has Eq. (71), Eq. (72) follows.

$$g(x) = \int_0^a e^{xt}\, dt = \frac{e^{xa} - 1}{x} \tag{69}$$

$$g'(x) = \int_0^a te^{xt}\, dt = \frac{ae^{xa}}{x} - \frac{e^{xa} - 1}{x^2} \tag{70}$$

$$g(x) = \int_0^a \frac{dt}{t^2 + x^2} = \frac{1}{x}\arctan(a/x) \tag{71}$$

$$g'(x) = -2x\int_0^a \frac{dt}{(t^2 + x^2)^2} = -\frac{1}{x^2}\arctan(a/x) - \frac{a}{x(a^2 + x^2)} \tag{72}$$

Another example in which the conditions for validity of the method still hold, though they are more difficult to verify, is given by Eq. (73), by means of which Eq. (74) can be derived,

$$g(x) = \int_0^\infty e^{-xt}\frac{\sin t}{t}\, dt \quad 0 \le x < \infty \tag{73}$$

$$g'(x) = -\int_0^\infty e^{-xt}\sin t\, dt = -\frac{1}{1 + x^2} \tag{74}$$

by integration by parts twice. Hence $g(x) = -\arctan x + C$, and $C = \pi/2$, since Eq. (75) can be proved. In particular, Eq. (76) holds.

$$\lim_{x\to\infty} g(x) = 0 \tag{75}$$

$$g(0) = \int_0^\infty \frac{\sin t}{t}\,dt = \frac{\pi}{2} \tag{76}$$

In case an indefinite integral is not elementary but the integrand is representable by an infinite series, the method of integrating term by term gives a representation for the integral which may be used for purposes of approximation. This can be seen in Eq. (77).

$$\text{Si}\,(x) = \int_0^x \frac{\sin t\,dt}{t} \sum_0^\infty \frac{(-1)^n x^{2n+1}}{(2n+1)^2(2n)!} \tag{77}$$

Certain definite integrals may be readily evaluated by means of contour integrals, that is, integrals of analytic functions taken along curves in the complex plane.

Integral of Lebesgue. In the study of the "space" of real functions defined, for example, on the interval $a \leqq x \leqq b$, it is frequently useful to take formula (78) as the distance between the functions f and g.

$$\int_a^b \left|f(x) - g(x)\right| dx \tag{78}$$

This distance already has a meaning when f and g are Riemann-integrable, that is, bounded and not too discontinuous, in the sense specified for the Riemann integral. There is no generally useful extension of the concept of integral to apply to all real functions on $[a, b]$, but it is desirable to extend it to apply to the functions obtained from the continuous ones by certain limiting processes. In particular, it is desirable to have correspond to each sequence (f_n) of functions satisfying the Cauchy condition for convergence in terms of the distance (78), namely Eq. (79),

$$\lim_{\substack{m\to\infty \\ n\to\infty}} \int_a^b |f_m(x) - f_n(x)|\,dx = 0 \tag{79}$$

a function g which is integrable (in the extended sense), and for which Eq. (80) holds. An extended

$$\lim_{n\to\infty} \int_a^b \left|f_n(x) - g(x)\right| dx = 0 \tag{80}$$

definition of integral having this property was given by H. L. Lebesgue in his thesis. It made obsolete many of the complicated extensions of the Riemann integral which had been previously proposed. Following Lebesgue, various mathematicians have proposed other ways of defining the integral which are equivalent to that of Lebesgue.

F. Riesz devised a definition which can be stated quite simply, at least for the case of a bounded function $g(x)$. As a first definition, a point set S in the interval $[a, b]$ has measure zero in case it can be enclosed in a sequence (finite or infinite) of intervals, the sum of whose lengths is arbitrarily small. Also, a

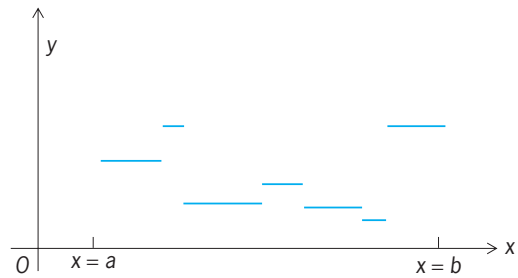

Fig. 4. Step function *f*(*x*).

step function $f(x)$ is defined as one which is constant on each interval of a partition of $[a, b]$, as in **Fig. 4**. The Riemann integral of a step function is expressible as a finite sum. Then a bounded function $g(x)$ is integrable in Lebesgue's sense in case it is the limit of a uniformly bounded sequence of step functions $f_n(x)$, at each point of $[a, b]$ except those in a set S with measure zero, and by definition Eq. (81) can be

$$\int_a^b g(x)\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx \tag{81}$$

written. In case the step functions f_n in Eq. (81) are replaced by Lebesgue-integrable functions forming a uniformly bounded sequence, no new functions g are obtained.

The integral of Lebesgue is also defined for unbounded functions, but the points of infinite discontinuity do not need to be considered one by one. For each function $g(x)$ and each positive integer N let $g_N(x)$ denote the lesser of $g(x)$ and N. If each $g_N(x)$ is Lebesgue-integrable in the sense already defined, expression (82) forms a nondecreasing sequence, and

$$\int_a^b g_N(x)\,dx \tag{82}$$

so if expression (82) is bounded, it tends to a finite limit, which is taken as the value of integral (83). An

$$\int_a^b g(x)\,dx \tag{83}$$

arbitrary function $g(x)$ is the difference of two nonnegative functions, Eqs. (84) and (85), and one may write Eq. (86) whenever both terms on the right have

$$g^+(x) = \frac{|g(x)| + g(x)}{2} \tag{84}$$

$$g^-(x) = \frac{|g(x)| - g(x)}{2} \tag{85}$$

$$\int_a^b g(x)\,dx = \int_a^b g^+(x)\,dx - \int_a^b g^-(x)\,dx \tag{86}$$

a meaning according to the definition just given. It is provable that the class of all functions for which integrals exist according to the definitions just given does indeed have the property of completeness with respect to the Cauchy condition (79). The theory of Lebesgue yields many other useful results, including an extension of the fundamental theorem of integral calculus. One important restriction on the class

of Lebesgue-integrable functions which is implicit in the definition is that when $g(x)$ is integrable so is $|g(x)|$. This restriction does not hold for the improper integrals which were described earlier.

The various methods of defining Lebesgue integrals are extensible also to functions of several variables. This extension throws light on the properties of multiple Riemann integrals and on the reduction of multiple integrals to repeated integrals.

Other definitions of integration. If $\alpha(x)$ is a fixed function defined on the interval $[a, b]$, the sum S defined by Eq. (1) may be replaced by that in Eq. (87).

$$S = \sum_{i=1}^{n} f(x_i)[\alpha(t_i) - \alpha(t_{i-1})] \tag{87}$$

Then if the limit I of S exists in this case, it is called the Stieltjes integral of f with respect to α, and is denoted by integral (88). It has many of the proper-

$$\int_a^b f(x)\, d\alpha(x) \tag{88}$$

ties of the Riemann integral, especially in case the function α is nondecreasing. In addition, it may take special account of the values of f at a finite or countable set of points. For example, if $\alpha(x) = c_i$ on $u_{i-1} < x < u_i$, where $u_0 < a$, $u_n > b$, $a < u_i < b$ for $i = 1, \ldots, n - 1$, and if $f(x)$ is continuous at each u_i, then Eq. (89) may be written. With some restrictions to

$$\int_a^b f(x)\, d\alpha(x) = \sum_{i=1}^{n-1} f(u_i)\,[c_{i+1} - c_i] \tag{89}$$

ensure convergence, this may be extended to the case in which α has infinitely many discontinuities.

In the case in which α is nondecreasing, the Stieltjes integral has an extension which is similar to that of the Lebesgue integral and is called the Lebesgue-Stieltjes integral.

Other extensions of the Lebesgue integral are the integrals of Denjoy. They apply to certain functions f for which $|f|$ is not integrable, and include the improper integrals, but not the Cauchy principal value. Numerous other types of integrals have been defined. In particular the integral of Lebesgue has been extended to cases in which the independent variable lies in a suitable space of infinitely many dimensions, or in which the functional values of f lie in such a space. *See* CALCULUS; FOURIER SERIES AND TRANSFORMS; SERIES.

Lawrence M. Graves

Bibliography. S.-B. Chae, *Lebesque Integration*, 2d ed., 1995; H. B. Dwight, *Tables of Integrals and Other Mathematical Data*, 4th ed., 1961; J. S. Gradshteyn and J. M. Ryzhjk, *Table of Integrals, Series, and Products*, 6th ed., 2000; P. R. Halmos, *Measure Theory*, 1950, reprint 1991; J. L. Kelley and T. P. Srinivasan, *Measure and Integral*, 1987; E. J. McShane, *Unified Integration*, 1983; W. Rudin, *Real and Complex Analysis*, 1987; C. W. Schwartz, *Measure, Integration and Function Spaces*, 1994; D. W. Stroock, *A Concise Introduction to the Theory of Integration*, 3d ed., 1998.

Integumentary patterns

All the features of the skin and its appendages that are arranged in designs, both in humans and other animals. Examples are scales, hairs, and feathers; coloration; and epidermal ridges of the fingers, palms, and feet. In its common usage, the term applies to the configurations of epidermal ridges, collectively named dermatoglyphics. Dermatoglyphics are characteristic of primates. *See* EPIDERMAL RIDGES; FEATHER; HAIR; SCALE (ZOOLOGY).

The superficial ridges are associated with a specific inner organization of skin. Skin is composed of two chief layers, the epidermis on the outside and the dermis underlying it (**Fig. 1**). These two layers are mortised by pegs of dermis, a double row of pegs corresponding to each ridge; these pegs accordingly form a patterning like that of the ridges.

Ridge patterns. The patterning of ridges, including that of the epidermal-dermal mortising, is determined during the third and fourth fetal months. All characteristics of single ridges and of their alignments are then determined definitively. Ridge alignments reflect directions of stress in growth of the hand and foot at the critical period of ridge differentiation. An important element in the production of localized patterns, for example, on the terminal segment of each digit, is the development in the fetus of a series of elevations, the volar pads.

Volar pads. The pads are homologs of the prominent pads on the paws of some mammals, but in primates they attain little elevation and soon tend to subside. The volar pads are disposed in a consistent topographic plan. Localized patterns have the same placement because growth of the pad is the determiner of the specific local pattern. When a pad has subsided before ridges are formed, its area does not present a localized pattern, and the ridges follow essentially straight, parallel courses. Variations in contours of the pads are accompanied by wide variations in the designs formed by the ridges overlying them.

Pattern variability. Variability of patterning is a major feature of dermatoglyphics and the basis for

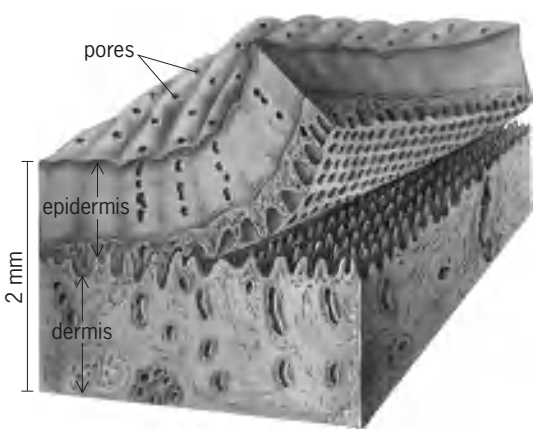

Fig. 1. Structure of ridged skin showing the chief two layers. (*From H. Cummins and C. Midlo, Finger Prints, Palms and Soles, McGraw-Hill, 1943*)

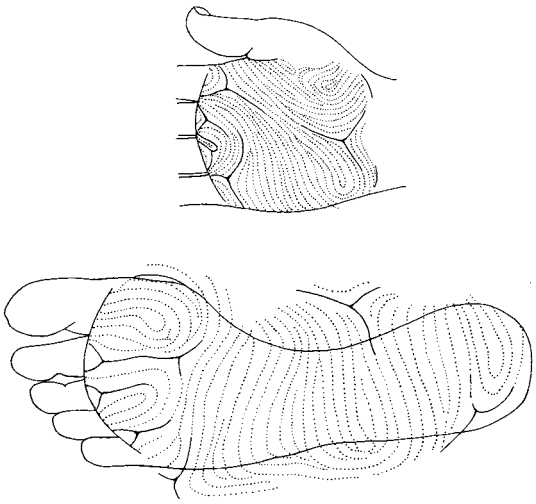

Fig. 2. Dermatoglyphics of palm and sole. (*After H. Cummins and C. Midlo, Finger Prints, Palms and Soles, McGraw-Hill, 1943*)

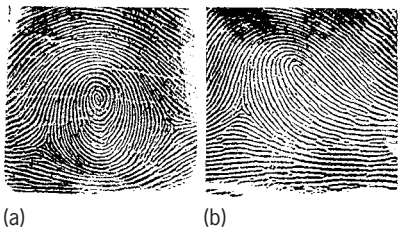

Fig. 3. Fingerprints. (*a*) Whorl pattern. (*b*) Loop pattern. (*From H. Cummins and C. Midlo, Finger Prints, Palms and Soles, McGraw-Hill, 1943*)

various applications (**Fig. 2**). In personal identification, prints (**Fig. 3**), customarily of fingers, are classified for filing in accordance with variables of pattern type and counts of ridges. Systematic filing makes it possible to locate readily and compare the sets of prints corresponding in classification to a set for which an identification is sought. In anthropological and medical investigations, groups of individuals are compared statistically in reference to the occurrence of these variables. Deductions may be drawn in accordance with likeness or unlikeness in the directions of variation. A few examples are cited. Trends of inheritance have been demonstrated in family groups and in comparisons of the two types of twins, fraternal (two-egg) and identical (one-egg). Dermatoglyphics thus are useful in diagnosing the types of twins and in analyzing cases of questioned paternity. Among different racial groups, the similar or discrepant trends of variation have been used to analyze racial affinities. Trends of variation are unlike in right and left hands, and the fact that they differ in accordance with functional handedness indicates an inborn predisposition of handedness. Departures from normal trends occur in many conditions associated with chromosomal aberrations, such as Down syndrome. *See* FINGERPRINT; HUMAN GENETICS; SKIN.

Harold Cummins

Bibliography. M. S. Elbauly and J. D. Schindler, *Handbook of Clinical Dermatoglyphs*, 1971; R. D. Olsen, Sr., *Scott's Fingerprint Mechanics*, 1978; R. I. Spearman, *The Integument: A Textbook of Skin Biology*, 1973.

Intelligence

General mental ability due to the integrative and adaptive functions of the brain that permit complex, unstereotyped, purposive responses to novel or changing situations, involving discrimination, generalization, learning, concept formation, inference, mental manipulation of memories, images, words and abstract symbols, eduction of relations and correlates, reasoning, and problem solving.

Theory. Verbal definitions or descriptions of intelligence, of examples of intelligent behavior, or of the characteristics of tests that best measure intelligence are all inadequate for a scientific understanding of intelligence. This is because "intelligence" is not a denotative noun, but a hypothetical construct or theory that is needed to account for one of the most striking and best-established phenomena in all of psychology: the ubiquitous positive intercorrelations among measurements of individual differences on all tests of "mental ability," however diverse. By "mental" is meant only that individual differences in the test scores are not the result of individual differences in sensory acuity or motor dexterity and coordination. By "ability" is meant conscious, voluntary performance on a task (for example, a test item) for which it can be reliably determined whether or not the performance meets some objective standard of correctness or efficiency.

The British psychologist Charles E. Spearman was the first to discover that all mental tests show positive correlations with one another when the scores are obtained in a representative sample of the general population. From this he inferred that all mental tests of every kind must measure some factor of ability that they all measure in common. He termed this common factor "general mental ability" and gave it the label *g*. Spearman, and most psychologists following him, have identified "general intelligence" with *g*.

By means of factor analysis, a quantitative method developed by Spearman for analyzing a matrix of all of the intercorrelations among a number of tests, it became possible to determine the proportion of the total variation in all of the test scores that can be attributed to variation in the general factor *g* common to all of the tests, and to determine the degree to which each test measured the *g* factor. A test's *g* loading, as derived from factor analysis, can be thought of as the test's coefficient of correlation with the *g* factor. Factor analyses of more than a hundred varieties of tests or test items reveal a wide range of *g* loadings. By comparing the tests having high, moderate, and low *g* loadings, one can gain an inkling of the nature of *g*. It cannot be identified with any specific type of informational content or skill required by the test. In these respects, highly *g*-loaded items are as diverse as all items in general. For example,

tests of vocabulary, problem arithmetic, and block designs (copying designs with multicolored blocks) are all highly *g*-loaded (and by their very nature are highly correlated with one another), whereas tests of spelling ability and arithmetic computation are only moderately *g*-loaded. Spearman, after inspecting hundreds of types of tests in relation to their *g* loadings, concluded that the only common features of the most highly *g*-loaded tests are that they involve "the eduction of relations and correlates" and "abstractness."

In more general terms, it is found that *g* is related to the complexity of the mental operations that are required by the task. Tasks involving reasoning and problem solving are much more *g*-loaded than tasks involving rote learning, routine skills, or the recall of specific information. When diverse tests are rank-ordered by their *g* loadings, they fall into four groups, from highest to lowest *g* loadings; the test items in these groups can be characterized as relational, associative, perceptual, and sensory-motor.

As yet there is no satisfactory or generally accepted theory of the neural mechanisms responsible for *g* or for individual differences in *g*, although there has been no paucity of speculation involving such notions as the total neural energy available to the brain, the number of brain cells and synaptic connections, and the speed of neural conduction.

The sexes, on average, do not differ in *g*; however, females slightly excel in verbal ability, and males slightly excel in spatial and quantitative abilities.

Factor analysis has revealed about 30 other factors of mental ability besides *g*. The most important of these factors, in terms of the amount of individual variation they account for, are verbal ability, numerical ability, spatial-visualization ability, and memory. A number of persons all with the same level of general intelligence or *g* may differ markedly from one another in these other ability factors. *See* FACTOR ANALYSIS.

Unity or modularity of intelligence. Some of the highly correlated abilities identified as factors probably represent what are also referred to as modules. Modules are distinct, innate brain structures that have developed in the course of human evolution. They are especially characterized by the various ways that information or knowledge is represented by the neural activity of the brain. Thus, the main modules are linguistic (verbal, auditory, lexical, and semantic), visuo-spatial, object recognition, numerical-mathematical, musical, and kinesthetic.

Although modules generally exist in all normal individuals, they are most strikingly in those with highly localized brain damage or pathology, and idiot savants. Savants evince striking discrepancies between a particular narrow ability and nearly all other abilities, often showing the greatest inconsistency with the individual's low level of general ability. There are some savants who are too mentally retarded to take care of themselves, yet can perform feats of mental calculation, or play the piano by ear, or memorize pages of a telephone directory, or draw objects from memory with nearly photographic accuracy. The modularity of these abilities is evinced by the fact that rarely, if ever, is more than one of them seen in a given savant. In contrast, there are individuals whose tested general level of ability is within the normal range, yet who, because of a localized brain lesion, show a severe deficiency in some particular ability, such as face recognition; receptive or expressive language dysfunctions (aphasia); or inability to form long-term memories of events. Again, modularity is evinced by the fact that these functional deficiencies are quite isolated from the individual's, total repertoire of abilities. In individual's with a normally intact brain, a module's efficiency can be narrowly enhanced through extensive experience and practice in the particular domain served by the module.

Measurement. Intelligence tests are diverse collections of *g*-loaded tasks (or items), graded in difficulty. The person's performance on each item can be objectively scored (for example, pass or fail); the total number of items passed is called the raw score. Raw scores are converted to some form of scaled scores which can be given a statistical interpretation.

The first practical intelligence test for children, devised in 1905 by the French psychologist Alfred Binet, converted raw scores to a scale of "mental age," defined as the raw score obtained by the average of all children of a given age. Mental age (MA) divided by chronological age (CA) yields the well-known intelligence quotient or IQ. When multiplied by 100 (to get rid of the decimal), the average IQ at every age is therefore 100, with a standard deviation of approximately 15 or 16. Because raw scores on mental tests increase linearly with age only up to about 16 years, the conversion of raw scores to a mental-age scale beyond age 16 must resort to statistical artifices. Because of this problem and the difficulty of constructing mental-age scales which preserve exactly the same standard deviation of IQs at every age, all modern tests have abandoned the mental-age concept and the calculation of IQ from the ratio of MA to CA. Nowadays the IQ is simply a standardized score with a population mean of 100 and a standard deviation (σ) of 15 at every age from early childhood into adulthood. The population distribution of IQ conforms closely to the normal curve in the range of $\pm 2\sigma$ from the mean (that is, from IQ 70 to 130), but there are excesses beyond that range. The middle 50% considered "average," fall between IQs of 90 and 110. IQs below 70 generally indicate "mental retardation," and above 130, "giftedness." The vast majority of students in selective colleges range above IQ 110, and graduates of 4-year colleges average about 120–125. *See* MENTAL RETARDATION.

Correlates of IQ. The IQ, as a measure of *g*, gains its importance from its substantial correlations with many educationally, occupationally, and socially valued variables. No other single measurement that can be obtained on a child (including parents' education and socioeconomic status) better predicts the child's scholastic performance. Childhood IQ, because of its relationship to later educational attainment, also predicts adult occupational status.

The fact that the IQ correlates with physical variables such as brain size and the amplitude and speed of evoked electrical potentials of the brain indicates that IQ tests reflect something biological, and not merely knowledge or skills acquired in school or in a cultured home.

Genetics. Individual differences in IQ involve polygenic inheritance, much as other continuous characteristics such as height and weight. The proportion of the total phenotypic variance in the population attributable to genetic factors, termed heritability, can be estimated by the methods of quantitative genetics from the correlations between various kinships—identical and fraternal twins, siblings, parent-child, and adopted children. Virtually all of the methodologically defensible estimates of the heritability of IQ range between.50 and.80, with a median value of about.70. Thus genetic factors are about twice as important as environmental factors as a cause of individual variation in intelligence.

All levels of intelligence are found within every major racial group. There is no definitive scientific evidence or consensus on the relative contributions of genetic and environmental factors to any undisputed average differences in intelligence between races. *See* BEHAVIOR GENETICS. Arthur R. Jensen

Bibliography. M. Anderson, *Intelligence and Development: A Cognitive Theory*, 1992; N. Brody, *Intelligence*, 2d ed., 1997; J. B. Carroll, *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*, 1993; D. K. Detterman (ed.), *Current Topics in Human Intelligence*, vol. 2: *Is Mind Modular or Unitary?*, 1992; J. A. Fodor, *The Modularity of Mind*, 1983; R. J. Sternberg (ed.), *Encyclopedia of Intelligence*, 1994.

Intelligent machine

Any machine that can accomplish its specific task in the presence of uncertainty and variability in its environment. The machine's ability to monitor its environment and then adjust its actions based on what it has sensed is a prerequisite for intelligence. The term intelligent machine is an anthropomorphism in that intelligence is defined by the criterion that the actions would appear intelligent if a person were to do it. A precise, unambiguous, and commonly held definition of intelligence does not exist.

Examples of intelligent machines include industrial robots equipped with sensors, computers equipped with speech recognition and voice synthesis, self-guided vehicles relying on vision rather than on marked roadways, and so-called smart weapons, which are capable of target identification. These varied systems include three major subsystems: sensors, actuators, and control. The class of computer programs known as expert systems is included with intelligent machines, even though the sensory input and output functions are simply character-oriented communications. The complexity of control and the mimicking of human deductive and logic skills makes expert systems central in the realm of intelligent machines. *See* EXPERT SYSTEMS; GUIDANCE SYSTEMS; ROBOTICS; VOICE RESPONSE.

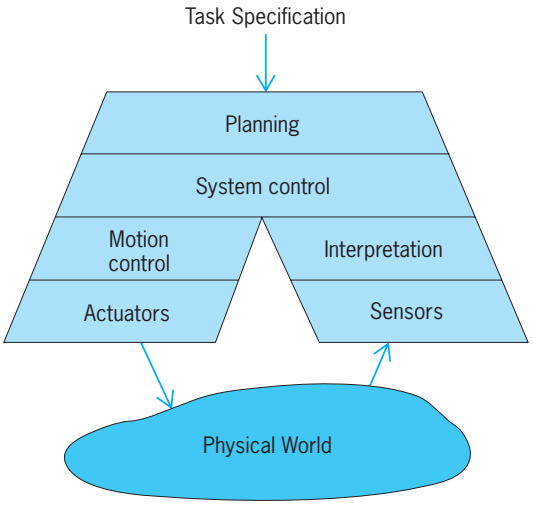

Fig. 1. Flow of information and data in a typical intelligent machine.

Intelligent control. Since the physical embodiment of the machine or the particular task performed by the machine does not mark it as intelligent, the appearance of intelligence must come from the nature of the control or decision-making process that the machine performs. Given the centrality of control to any form of intelligent machine, intelligent control is the essence of an intelligent machine. The control function accepts several kinds of data, including the specification for the task to be performed and the current state of the task from the sensors. The control function then computes the signals needed to accomplish the task. When the task is completed, this also must be recognized and the controller must signal the supervisor that it is ready for the next assignment (**Fig. 1**).

Automatic, feedback, or regulatory systems such as thermostats, automobile cruise controls, and photoelectric door openers are not considered intelligent machines. Several important concepts separate these simple feedback and control systems from intelligent control. While examples could be derived from any of the classes of intelligent machines, robots will be used here to illustrate five concepts that are typical of intelligent control. (1) An intelligent control system typically deals with many sources of information about its state and the state of its environment. A robot may contain a vision system and tactile sensors in addition to the internal position sensors that allow calculation of its spatial location. Given multiple sources of information, the intelligent control function evaluates them as needed based on the nature of the functions that must be performed at any instant. (2) An intelligent control system can accommodate incomplete or inconsistent information. For example, a robot expecting to pick up a single object from a conveyor belt may be confronted with two overlapping objects or no objects. (3) Intelligent control is characterized by the use of heuristic methods in addition to algorithmic control methods. A

heuristic is a rule of thumb, a particular solution or strategy to be used for solving a problem that can be used for only very limited ranges of the input parameters. A possible heuristic for solving a particular robot-implemented insertion task would be to move the part incrementally along an outward spiral path, retrying the insertion, if the original attempt should fail. This heuristic could be repeated many times until an area large compared to positional uncertainties was covered. Clearly, such a heuristic is applicable only to a very small fraction of robot tasks. (4) An intelligent machine has a built-in knowledge base that it can use to deal with infrequent or unplanned events. An intelligent robot would have mechanisms, both sensors and heuristics, ensuring its own and its operator's safety. (5) An algorithmic control approach assumes that all relevant data for making decisions is available. The heuristic approach is predicated on the knowledge that all relevant data cannot be made available, even in principle, and that the control system will have to resolve ambiguous cases. *See* ADAPTIVE CONTROL; CONTROL SYSTEMS.

Intelligent robots. To clarify these very general concepts and to illustrate them with more detail, a sensor-equipped robot performing an electronics assembly task will be described. The task is to pick a part from an egg-carton-like pallet, inserting it into a test fixture, and, if it tests acceptably, to insert it into the appropriate place on the printed wiring board being assembled. A human worker, especially one with some experience in production line assembly, could be instructed in this task in a few minutes and would also know, without further instruction, what actions to take for the many exceptional conditions that can arise in this type of activity. Outlining the steps that must be taken to program a robot to accomplish the same task will illustrate the amount of prior knowledge possessed by a person and the general level of complexity of programming in an environment where all relevant data and actions must be stated completely and unambiguously. The robot program will be given in verbal outline form rather than coded in a particular robot programming language. The numbers are used both for specifying branches in the flow of control and for referencing in the discussions that follow.

1. Wait for a printed wiring board to enter the assembly station.
2. Take an image of the parts presentation pallet: an egg-carton arrangement that separates and loosely orients the parts.
3. Locate a part in the image.
4. If the pallet is empty, move it to the stacking area and signal the conveyor to move a fresh pallet into position. Resume at step 2.
5. If a part is present, pick up the part and position it in the visual inspection station. Take an image of the bottom of the part. Determine the location and orientation of the pin field.
6. If an incorrect pin field is seen, deposit the part in the reject bin and resume at step 2.
7. If the pin field is correct, orient the robot hand based on the observed orientation of the part and insert the part into the electrical test station.
8. Signal the electrical test to begin.
9. If the test fails, remove the part from the test fixture and deposit it in the reject bin. Then resume at step 2.
10. If the electrical test was passed, regrasp the part to ensure it is accurately held by the robot. Remove it from the test fixture and move it to a point just above the insertion location on the printed wiring board.
11. Insert the part in the printed wiring board by moving it vertically downward while monitoring the force. When the vertical force reaches a specified value, stop the robot's vertical motion.
12. Check the vertical position of the robot hand when it stopped moving. If it is at a height indicating that the part was inserted, release the part, retract the robot hand, and resume at step 1.
13. If the part was not inserted correctly, move the robot hand in the plane of the printed wiring board by a small increment along a prescribed path and resume at step 11.

Several comments can be made about a robot that is following the above program and the program itself. Such a robot would appear to be intelligent given the definition in the beginning of the article. There is variability in the robot's environment, and it is using a variety of sensory data to guide its actions. However, the program is incomplete. There are many things that can go wrong that have not been accounted for: If a hole in the printed wiring board for the part is missing, the insertion cannot be successfully accomplished. In this case, program steps 11, 12, and 13 form an apparent infinite loop. Many other similar problems lurk in this program. The program also presupposes significant capabilities and data in the robot and related vision system. The vision system and robot must be calibrated; that is, a procedure must be periodically executed that calculates the transformations between the robot and vision system coordinates. The vision system must be programmed to recognize first what the part looks like in the context of the pallet (step 3) and second what a correct pin field looks like (step 5). The precise locations of all points of reference for the robot must be determined. The search strategy to be followed if the part is not inserted correctly (step 13) must be carefully defined. The calibration and teaching programs required to accomplish these steps may be as complex as the primary task.

Execution of the assembly program and all of the implied subsidiary functions defines the control capability of the robot. The threshold of intelligent control is reached when the system can deal with many of the infrequent and exceptional conditions without requiring outside intervention or assistance. A second part of intelligent control is the development of the detailed programs given a high-level statement of the task requirements. This is the planning function (Fig. 1). Systems of this level of complexity are

primarily found in robotics research laboratories. The automatic planning of complex robot control programs given high-level task descriptions is an area of research.

Control hierarchy. The control function in an intelligent machine is generally implemented as a hierarchy of processes. The upper levels are responsible for global decision-making processes and for the planning phases of the task. The lower levels implement critical time-dependent subtasks. In **Fig. 2** the interconnection of a group of processors forming a vision-controlled robot is diagrammed. The manipulator arm is itself controlled by a two-level hierarchy. The lower level consists of a processor that performs the servo function for each joint of the arm. The robot controller converts world coordinates supplied by the system controller into the values for each joint required to place the robot hand in the desired location. The robot controller also divides complete movement specifications into a series of short steps that result in coordinated motion and controlled velocity of the robot. *See* SERVOMECHANISM.

The control program outlined in the robot example would be executed on the system-level control processor. In addition to the vision, manipulator, and safety subsystems, the system-level controller would also be the interface with the robot operator and the factory communications system. The partitioning of subtasks onto autonomous communicating processors is a natural and economic way to obtain increased system performance by using multiple microprocessors. *See* MICROPROCESSOR.

Computer vision. A computer vision system interprets the contents of a large array (256 × 256 is common) of intensity values that is the computer representation of a frame of a television image. Interpret in this context means to find, recognize, or inspect a prespecified object or group of objects. The first processing step is to isolate or segment the image of the object from the background image. The most common technique in use is to create a high-contrast image, typically by backlighting the object of interest, which allows a simple threshold to be used for the segmentation. Objects in the image are found at those locations where the intensity value is less than some prespecified constant. Lower-contrast images, such as a part in a pallet, require segmentation techniques that rely on rapid changes in brightness (edges) or differences in intensity between two images, or techniques that look for specific visual features by using any of a large number of matched filter techniques. The amount of computation required to segment a low-contrast image is significantly greater than that required by thresholding. Once an object image is segmented from the background, a set of parameters known as features is calculated. The set of features, which may include the object area, perimeter, number of holes and any other enumerable parameter, is compared to a model of the object either to ascertain its identity or to compare it to some inspection criteria. The specific set of features needed for a specific visual recognition problem depends

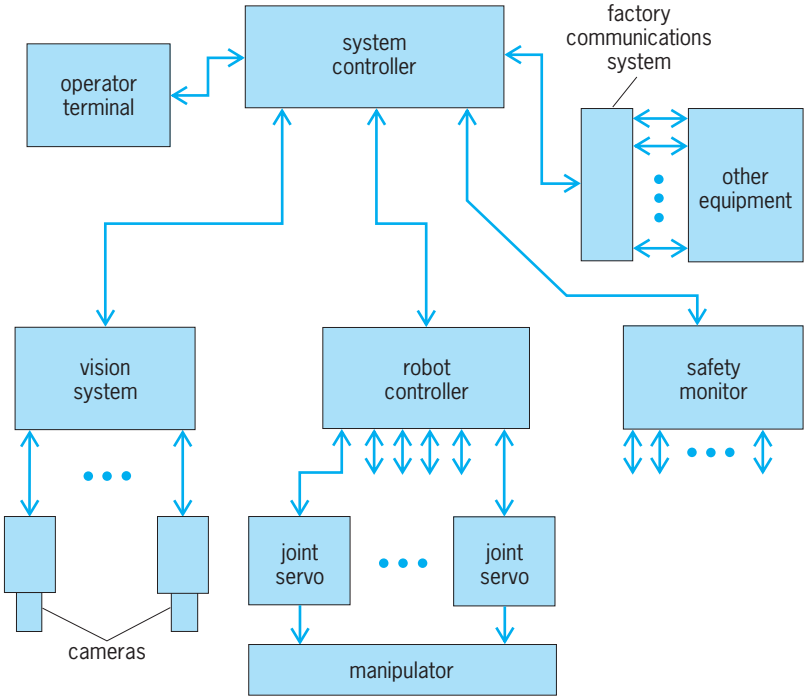

Fig. 2. Interconnection of processors forming a sensor-based robot.

on the exact appearances of the objects in complex ways with the result that the feature set is normally determined empirically. Similarly, orientation and location quantities are also computed for the segmented image. Generally, an image represents a two-dimensional projection of a three-dimensional object. Correction to the perceived shape and location may have to be made based on its location relative to the viewing camera. No matter what complexity of techniques are used in its construction, a computer vision system is invariably implemented with a dedicated processor. *See* COMPUTER VISION.

Interconnection of machines. Another trend expected in the area of intelligent machines is their interconnection into cooperating groups to address larger tasks. Standardization of information exchange protocols and languages describing various fields of work are prerequisites to any extensive development of interconnected, cooperating systems.

Capabilities. Slow, steady improvements may be expected in intelligent machine capabilities. Progress will result from advances in microprocessors and the implementation of various functions as very large-scale integrated (VLSI) circuits. *See* INTEGRATED CIRCUITS.

Expert systems. The control schemes outlined above are algorithmic even if the resulting system behavior appears intelligent. The explicit programming methods described do not lend themselves to situations where several alternative actions could be rationally considered. In the explicit programming of a complex task, there is very little opportunity to use default actions for many infrequent conditions. The method of expert systems is expected to provide a paradigm for a better way of robot task planning and

execution. Expert systems have a knowledge base and an inference procedure for obtaining a result consistent with the contents of the knowledge base, given an arbitrary set of input conditions. The knowledge base is represented by a collection of condition-action statements. These statements are called rules and the inference procedure is known as the rule interpreter. The other major components of an expert system (discussed below) are a user interface, a database, and a rule structure.

The primary difficulties are in obtaining an adequate collection of rules for the task and in defining an appropriate strategy for resolving inconsistencies or ambiguities when applying rules to a given problem. A specific robot task would be programmed by adding a set of rules, describing the particular operation to be formed, to an existing, and probably much larger, set of rules that defined the default and exception-handling behavior of the robot. John F. Jarvis

Applications. Since the mid-1970s, expert systems have been continuously developed and applied by industry, business, and commerce. Expert systems are the most successful implementation of artificial intelligence. Applications include medical diagnosis, mineral analysis, control of complex processes, and fault diagnosis and monitoring; systems for fault diagnosis and monitoring account for about half the applications. *See* ARTIFICIAL INTELLIGENCE.

Goal-directed systems. In many expert systems, the rule interpreter, or inference engine as it is sometimes called, uses an exhaustive backward-chaining strategy, in which the rules are expressed in a highly focused "IF-condition-THEN-result" form, the result is known or concluded, and, knowing the answer, the system searches backward for the rule whose condition produced the result. Such an expert system is termed goal-directed, since the search is, in fact, directed toward a goal (the known result). Goal-directed systems are therefore the result of going from the general to the particular.

An inference mechanism based on a backward-chaining strategy is commonly equipped with an extensive explanation facility to reduce the amount of conflict resolution needed. To be user-friendly, an expert system must also possess a user interface that interrogates the human expert in a manner that is unambiguous. The information obtained from the user interface must then be stored as facts in a database, and the relationship between individual facts must be stated. The rule structure determines which rule should be examined next, in a forward-chaining expert system, or it further questions the user in a backward-chaining system. Finally, the representational structure must also include an uncertainty-handling mechanism that is based on measurements of belief. *See* COMPUTER; DATABASE MANAGEMENT SYSTEM; FUZZY SETS AND SYSTEMS.

While the use of rules having the form "IF-condition-THEN-result," known as production rules, has been particularly successful when used in goal-directed applications such as diagnostics, such rules have limitations. The major limitations associated with production rules are the so-called knowledge acquisition bottleneck and the difficulty of determining the correct method of representing the rules when there is a large number of them. In order to overcome these limitations, alternative representations have been attempted. Production rules have been represented in the form of inference nets, causal mechanisms, and frame-based abduction mechanisms. The inference net is overlaid by a semantic net to connect different objects. As with production rules, an inference-net approach can be backward-chained, but it requires a different mechanism for handling uncertainty. For inference nets, a mechanism based on Bayes' rule, a statistical rule, has been used. However, if an expert system is constructed by using causal or frame-based abduction mechanisms it must have inference and control mechanisms suited to those representations. Also, the relational mechanism for storing facts in the database must alter accordingly. For causal nets the database is divided into findings and hypotheses. In frame-based abduction the database holds domain knowledge in the form of descriptive frames of information. Also, with frame-based abduction the inference mechanism uses hypotheses-and-test cycles. *See* ESTIMATION THEORY.

Machine learning. The difficulties surrounding production rules, including knowledge acquisition, descriptive knowledge representation, and context handling, have led to the development of a subdivision of expert system technology known as machine learning, which is a method for automatically generating rules from examples. Thus, machine learning is an approach based on the idea of going from the particular to the general.

Machine learning can be classified into three areas, artificial-intelligence–type learning, neural nets, and genetic algorithms. An effective machine-learning system must use sampled data to generate internal updates, be capable of explaining its findings in an understandable way (for example, symbolically), and communicate its explanation to a human expert in a manner that improves the human's understanding. Expert systems based on artificial-intelligence learning have surpassed the performance of human experts, and they have been able to communicate what they have learned to human experts for the purpose of verification.

There are two schools of artificial-intelligence-type learning. One is based on logic and attempts to find minimal representations of finite functions, while the other concentrates on learning via knowledge representation, acquisition, compaction, and learning from example. The theoretical base of machine learning is multivariate analysis in statistics, and the use of decision trees. Attribute-based learning systems use a common form of knowledge representation to learn explicit rules. Input attributes can be discrete, continuous, or tree-structured, or they can be example sets that contain attribute values. In example sets, each distinguished attribute is the

classification attribute for each example. The output is a decision rule that is capable of classifying further, unseen examples that are presented to the system. *See* DECISION THEORY; STATISTICS.

Artificial-intelligence–type learning originated from an investigation into the possibility of using decision trees or production rules for concept representation. Subsequently the work was extended to use decision trees and production rules in order to handle the most conventional data types, including those with noisy data, and to function as a knowledge acquisition tool. To construct a rule from a training set, the algorithm must first examine all the examples in the training set to see if they are all of the same class. If this is the case, classification is achieved. If the training sets are not all of the same class, then the most significant attribute is selected from the training set, the set is partitioned into two sets, and subtrees are constructed. The evaluation function used for selecting an attribute is based on information theory, and recursive methods are used to construct the subtrees. Based on this approach, in 1978 the first expert system generated exclusively by inductive learning from example was used to diagnose 19 soybean diseases with a degree of accuracy higher than that of a human expert.

Interest in neural-net (connectionist) learning algorithms stems from their ability to construct a network model based on training examples. When finally constructed, the neural network becomes the knowledge base for an expert system. Neural networks are capable of automatically generating behavior based on example sets of types described in the discussion of artificial-intelligence–type learning above. Their major limitation is that they are useful only in diagnostic systems that have fixed input and output sets. *See* NEURAL NETWORK.

Genetic algorithms are based on iteration. They use general-purpose search strategies that have their foundations in genetics, in particular, in the process of natural selection. The purpose of iteration is to weed out weaker representations and give strengthened support to those that provide better results. Genetic algorithms attempt to balance exploration with exploitation, yet at the same time they try to avoid the problems associated with other goal-directed methods. For example, they attempt to avoid the local-minima problems found in hill climbing methods and the inefficiencies of random search. Since they make no prior assumption about the problem domain, they can be applied generally, and, in particular, in areas whose characteristics are unknown. They have been applied successfully in domains where optimization is a problem, but they started to be used for learning systems only in the late 1980s. The iterative procedure of a genetic algorithm maintains a set of knowledge structures, called a population. This structure is a string of symbols chosen from a fixed set of examples; this is a candidate solution. At each generation the structure is evaluated and is given a fitness value that is dependent on how close it is to the required solution. While genetic algorithms and neural nets have contributed to the study of machine learning, a disadvantage of these methods is that learning is a slow process when they are used. *See* EXPERT SYSTEMS. Edward Grant

Bibliography. D. A. Bradley et al., *Mechatronics and the Design of Intelligent Mechanics and Systems*, 2000; R. Kurzweil, *Age of Intelligent Machines*, 1992; M. W. Spong, S. Hutchinson, and M. Vidyasagar, *Robot Modeling and Control*, 2005; G. Zobrist and C. Y. Ho, *Intelligent Systems and Robotics*, 2000.

Intercalation compounds

Crystalline or partially crystalline solids consisting of a host lattice containing voids into which guest atoms or molecules are inserted. Candidate hosts for intercalation reactions may be classified by the number of directions (0 to 3) along which the lattice is strongly bonded and thus unaffected by the intercalation reaction. Isotropic, three-dimensional lattices (including many oxides and zeolites) contain large voids that can accept multiple guest atoms or molecules. Layer-type, two-dimensional lattices (graphite and clays) swell up perpendicular to the layers when the guest atoms enter (**illus.** *a*). The chains in one-dimensional structures (polymers such as polyacetylene) rotate cooperatively about their axes during the intercalation reaction to form channels that are occupied by the guest atoms (illus. *b*). In the intercalation family based on solid C_{60} (buckminsterfullerene), the zero-dimensional host lattice consists of 60-atom carbon clusters with strong internal bonding but weak intercluster bonding. These clusters pack together like hard 1-nm-diameter spheres, creating interstitial voids which are large enough to accept most elements in the periodic table (illus. *c*). The proportions of guest and host atoms may be varied continuously in many of these materials, which are therefore not true compounds. Many ternary and quaternary substances, containing two or three distinct guest species, are known. The guest may be an atom or inorganic molecule (such as an alkali metal, halogen, or metal halide), an organic molecule (for example, an aromatic such as benzene, pyridine, or ammonia), or both. *See* CRYSTAL STRUCTURE; FULLERENE; GRAPHITE; POLYMER; ZEOLITE.

Competing interactions. The stability of a particular guest-host configuration depends on the balance between positive and negative contributions to the formation energy of the so-called compound. For example, in carbon-based hosts such as graphite, both guest and host become ionized after intercalation, and the attraction between oppositely charged guest and host ions helps to offset the elastic energy penalty associated with the swelling of the lattice. Some of these contributions vary with temperature, leading to structural phase transitions such as internal melting of the guest layers in graphite intercalation compounds. In some cases the guest layers

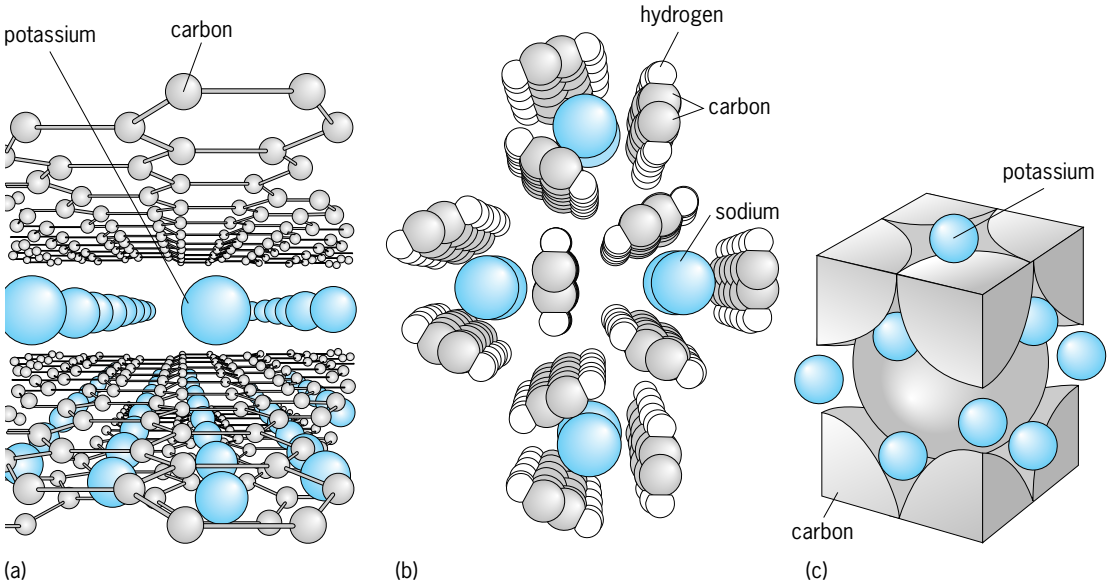

Typical intercalation compounds. (*a*) Potassium in graphite, KC_8, a prototype layer intercalate. (*b*) Sodium in polyacetylene, $[Na_{0.13}(CH)]_x$, where *x* denotes infinitely repeating polymer chains. (*c*) Potassium in solid buckminsterfullerene, K_3C_{60}. The internal structure of the fullerene molecule is not shown. (*After J. E. Fischer and P. Bernier, Les cristayx des fullerènes, La Recherche, 24(250):46–55, January 1993*)

can be far enough apart to be considered independent, providing important prototypes for studying phenomena such as magnetism, superconductivity, and phase transitions in two dimensions. The spatial distribution of guest atoms or molecules within the host network can be regular or disordered, depending on temperature, concentration, and the chemical nature of the guest. *See* COHESION (PHYSICS); SOLID-STATE CHEMISTRY.

An example of temperature-dependent behavior is exhibited by alkali atoms inserted in graphite. If precisely half the number of alkali atoms that could be accommodated at the limit of close-packed alkali monolayers is added between each pair of graphite sheets, the system displays a phase transition. At low temperature, the elastic penalty will be minimized if the alkali ions are crowded into only half the galleries, resulting in a periodic dense-stage-2 configuration: maximally dense alkali monolayers between every other pair of host sheets. At high temperature the mixing entropy between filled and empty sites drives a transition to a site-disordered dilute-stage-1 configuration, in which all galleries contain some ions, on average 50% of the hard-sphere limit. The dense-stage-2 configuration may be recovered by applying hydrostatic pressure to the high-temperature configuration, because the volume is minimized by segregating the filled and empty sites into completely filled and empty galleries respectively. *See* ENTROPY; PHASE TRANSITIONS; THERMODYNAMIC PRINCIPLES.

Properties. The physical properties of the host are often dramatically altered by intercalation. In carbon-based hosts such as graphite, polyacetylene, and solid C_{60}, the charge donated to the host by the guest is delocalized in at least one direction, leading to large enhancements in electrical conductivity. Some hosts can be intercalated by electron-donating or -accepting guests, leading respectively to *n*-type or *p*-type synthetic metals, analogous to doped semiconductors. The chemical bonding between guest and host can be exploited to fine tune the chemical reactivity of the guest. In many cases, superconductors can be created out of nonsuperconducting constituents. *See* SEMICONDUCTOR; SUPERCONDUCTIVITY.

Applications. Many applications of intercalation compounds derive from the reversibility of the intercalation reaction. The best-known example is pottery: Water intercalated between the silicate sheets makes wet clay plastic, while driving the water out during firing results in a dense, hard, durable material. Many intercalation compounds are good ionic conductors and are thus useful as electrodes in batteries and fuel cells. A technology for lightweight rechargeable batteries employs lithium ions which shuttle back and forth between two different intercalation electrodes as the battery is charged and discharged: vanadium oxide (three-dimensional) and graphite (two-dimensional). Zeolites containing metal atoms remain sufficiently porous to serve as catalysts for gas-phase reactions. Many compounds can be used as convenient storage media, releasing the guest molecules in a controlled manner by mild heating. *See* BATTERY; CLAY, COMMERCIAL; FUEL CELL; SOLID-STATE BATTERY.

J. E. Fischer

Bibliography. P. Bernier et al. (eds.), *Chemical Physics of Intercalation II*, 1993; T. Enoki, M. Suzuki, and M. Endo, *Graphite Intercalation Compounds and Applications*, 2003; C. Julien, J. P. Pereira-Ramos, and A. Momchilov (eds.), *New Trends in Intercalation Compounds for Energy Storage*, 2002; A. P. Legrand and S. Flandrois (eds.), *Chemical Physics of Intercalation*, 1987.

Intercommunicating systems (intercoms)

A privately owned system that allows voice communication between a limited number of locations, usually within a relatively small area, such as a building, office, or residence. Intercommunicating systems are generally known as intercoms. Intercom systems can vary widely in complexity, features, and technology. Though limited in size and scope, intercom systems can provide easy and reliable communication for their users.

Wire systems. An extremely simple intercom is a two-station arrangement in which one station is connected to the other via a dedicated wire. Simple twisted-pair cabling is the most common type of wiring for this type of intercom. An example in a business application is an executive-secretary system, in which the executive presses a button to signal the secretary and initiate the connection. When the secretary answers, the voice link is automatically established.

Other systems have multiple stations, as many as 10 to 20, any of which can connect with any other station. The user must dial a one- or two-digit code to signal the intended destination. When the called party answers, the talk path is established between the two stations. In a residence, this type of system permits calling between extension telephones through abbreviated dialing codes; the called extension alerts the user through a distinctive ringing signal.

Telephone intercoms. Still other intercom systems work in conjunction with key and hybrid key telephone/private branch exchange telephone systems. They support internal station-to-station calling rather than access to outside lines. Many intercoms possess some or all of the features of that particular telephone system, such as call hold, call pickup, and conferencing. *See* KEY TELEPHONE SYSTEM; PRIVATE BRANCH EXCHANGE.

Normally the telephone intercom is incorporated in the same telephone instrument that is used to access the public switched network. Two primary types of intercom systems exist: the single-link and the multilink. The single-link system provides a single dedicated intercom talk path in the telephone system, no matter how many stations might utilize the intercom system. The multilink intercom system provides two or more dedicated intercom paths that allow more than one simultaneous intercom conversation to be in progress. *See* SWITCHING SYSTEMS (COMMUNICATIONS); TELEPHONE SERVICE.

Wireless intercoms. A third type of intercom is the wireless intercom system for intrabuilding communications, which consists of a base unit radio transmitter, equipped with an antenna, and a number of roving units tuned to different frequencies. The base can selectively communicate with individual roving units by dialing the code corresponding to each roving unit's specific frequency. Commercial systems usually have a range of 1–2 mi (2–3 km). Structural components of the building, such as steel or concrete, can interfere with transmission quality and limit effective range. *See* MOBILE COMMUNICATIONS.

Byron W. Battles; Vincent F. Rafferty

Bibliography. The Aries Group·MPSG, *Glossary of Terminology: Key Telephone Systems and Hybrid KTS/PBX Systems*, 1990; Center for Communications Management, Inc., *Telecommunications Dictionary and Fact Book*, 1984; G. Langley, *Telephony's Dictionary*, 2d ed., 1986.

Interface of phases

The boundary between any two phases. Among the three phases, gas, liquid, and solid, five types of interfaces are possible: gas-liquid, gas-solid, liquid-liquid, liquid-solid, and solid-solid. The abrupt transition from one phase to another at these boundaries, even though subject to the kinetic effects of molecular motion, is statistically a surface only one or two molecules thick.

A unique property of the surfaces of the phases that adjoin at an interface is the surface energy which is the result of unbalanced molecular fields existing at the surfaces of the two phases. Within the bulk of a given phase, the intermolecular forces are uniform because each molecule enjoys a statistically homogeneous field produced by neighboring molecules of the same substance. Molecules in the surface of a phase, however, are bounded on one side by an entirely different environment, with the result that there are intermolecular forces that then tend to pull these surface molecules toward the bulk of the phase. A drop of water, as a result, tends to assume a spherical shape in order to reduce the surface area of the droplet to a minimum.

Surface energy. At an interface, there will be a difference in the tendencies for each phase to attract its own molecules. Consequently, there is always a minimum in the free energy of the surfaces at an interface, the net amount of which is called the interfacial energy. At the water-air interface, for example, the difference in molecular fields in the water and air surfaces accounts for the interfacial energy of 7.2×10^{-6} joule/cm^2 (11.15×10^{-6} cal/in.^{2}) of interfacial surface. The interfacial energy between the two liquids, benzene and water, is 3.5×10^{-6} J/cm^2 (5.4×10^{-6} cal/in.^{2}), and between ethyl ether and mercury is 37.9×10^{-6} J/cm^2 (58.67×10^{-6} cal/in.^{2}). These interfacial energies are also expressed as surface tension in units of millinewtons per meter.

The surface energy at an interface may be altered by the addition of solutes that migrate to the surface and modify the molecular forces there, or the surface energy may be changed by converting the planar interfacial boundary to a curved surface. Both the theoretical and practical implications of this change in surface energy are embodied in the Kelvin equation, Eq. (1) where P/P_0 is the ratio of the vapor pressure of

$$\ln \frac{P}{P_0} = \frac{2M\gamma}{RT\rho r} \qquad (1)$$

a liquid droplet with diameter r to the vapor pressure

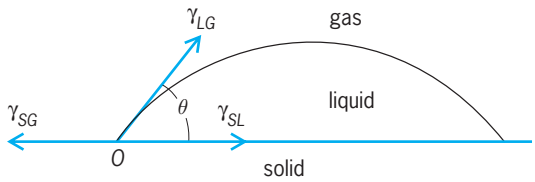

Contact angle θ at interface of three phases.

of the pure liquid in bulk, ρ the density, γ the surface energy, and M the molecular weight. Thus, the smaller the droplet the greater the relative vapor pressure, and as a consequence, small droplets of liquid evaporate more rapidly than larger ones. The surface energy of solids is also a function of their size, and the Kelvin equation can be modified to describe the greater solubility of small particles compared to that of larger particles of the same solid. *See* ADSORPTION; SURFACTANT.

Contact angle. At liquid-solid interfaces, where the confluence of the two phases is usually termed wetting, a critical factor called the contact angle is involved. A drop of water placed on a paraffin surface, for example, retains a globular shape, whereas the same drop of water placed on a clean glass surface spreads out into a thin layer. In the first instance, the contact angle is practically 180°, and in the second instance, it is practically 0°. The study of contact angles reveals the interplay of interfacial energies at three boundaries. The **illustration** is a schematic representation of the cross section of a drop of liquid on a solid. There are solid-liquid, solid-gas, and liquid-gas interfaces that meet in a linear zone at O. The forces about O that determine the equilibrium contact angle are related to each other according to Eq. (2), where the γ terms represent free energies at

$$\gamma_{SG} = \gamma_{SL} + \gamma_{LG} \cos\theta \qquad (2)$$

the interfaces and θ is the contact angle. Since only γ_{LG} and θ can be measured readily, the term adhesion tension is defined by Eq. (3). Adhesion tension,

$$\gamma_{LG} \cos\theta = \gamma_{SG} - \gamma_{SL} = \text{adhesion tension} \qquad (3)$$

which is the free energy of setting, is of critical importance in detergency, dispersion of powders and pigments, lubrication, adhesion, and spreading processes.

The measurement of interfacial energies is made directly only upon liquid-gas and liquid-liquid interfaces. In measuring the liquid-gas interfacial energy (surface tension), the methods of capillary rise, drop weight on pendant drop, bubble pressure, sessile drops, Du Nuoy ring, vibrating jets, and ultrasonic action are among those used. There is a small but appreciable temperature effect upon surface tension, and this property is used to determine small differences in the surface tension of a liquid by placing the two ends of a liquid column in a capillary tube whose two ends are at different temperatures. The determination of interfacial energies at other types of interfaces can be inferred only by indirect methods.

See FLOTATION; FOAM; FREE ENERGY; PHASE EQUILIBRIUM; SURFACE TENSION. Wendell H. Slabaugh

Bibliography. A. W. Adamson and A. P. Gast, *Physical Chemistry of Surfaces*, 6th ed., 1997; H.-J. Butt, K. Graf, and M. Kappl, *Physics and Chemistry of Interfaces*, 2d ed., 2006.

Interference filters

Optical filters that are based on the interference of optical radiation in thin films—the phenomenon responsible for the iridescent colors observed in soap bubbles, oil slicks, shells, bird feathers, and butterfly wings. An optical filter is any material or device that is used to change the intensity, spectral distribution, phase, or polarization properties of the transmitted or reflected components of the optical radiation incident upon it. Large and complicated instruments, such as interferometers and spectrophotometers, are not usually included within this definition. Some of the many other different physical principles that have been used in the past to produce optical filters are absorption, reflection, scattering, diffraction, and polarization. The term "coating" is commonly used as a synonym for "filter" and, historically, has been applied to thin-film filters of certain functionalities. *See* COLOR FILTER; INTERFERENCE OF WAVES.

Filters and coatings. Filters and coatings based on thin-film interference are remarkably versatile. Although most are currently produced for the visible and adjacent parts of the spectrum, thin-film interference filters are also made for spectral regions extending from the x-ray to submillimeter wavelengths. Thus, the ratio of the highest to the lowest wavelengths for which optical interference coatings are made is of the order of $1{:}10^6$. *See* ELECTROMAGNETIC RADIATION.

In addition to their use in optical instruments for science and engineering, thin-film interference filters are used in eyeglasses and photographic cameras, energy conversion and conservation, architecture, communications, and display devices. New applications for optical interference filters appear regularly.

One disadvantage of filters based on thin-film interference is that unless special measures are taken, their properties change with the angle of incidence of light. In some cases, even that can be turned into an advantage. For example, it is possible to tune the filters by tilting them, and the iridescent property of the thin-film coatings can be used for protection from counterfeiting as well as for decorative purposes.

Thin-film thicknesses and materials. Optical thin films are mostly made of inorganic materials, including dielectrics (such as Al_2O_3, CaF_2, HfO_2, MgF_2, PbF_2, Sc_2O_3, SiO_2, TiO_2, Y_2O_3, ZnS, and ZrO), semiconductors (such as GaAs, Ge, and Si), metals (such as Ag, Al, Au, Cr, Ni, and Mo), and alloys (such as the Ni-Cr-Fe alloys, Inconel and Nichrome), although organic materials are increasingly being used. The thicknesses of the transparent films vary, but they are usually within one or two orders of magnitude of the wavelengths

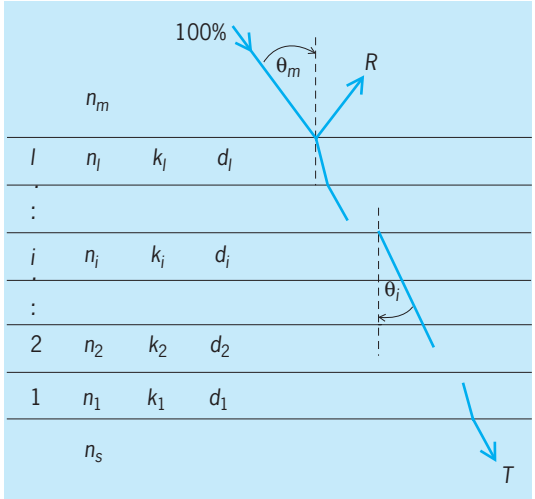

Fig. 1. Schematic representation of a homogeneous multilayer coating.

for which the filter is to be effective. The absolute thickness will vary a great deal, depending on whether the filter is intended for the x-ray, visible, or submillimeter spectral regions. For a filter designed for the visible part of the spectrum, a typical thickness might be around 100 nanometers. The thicknesses of metal and alloy layers within a multilayer system have to be thin enough that the layers are substantially transmitting. Metal layer thicknesses are typically around 5 nm.

Deposition processes. The most precise optical interference filters are currently produced by physical deposition processes, such as thermal or electron-beam evaporation with or without ion assist, by ion plating methods, by various forms of direct-current (dc), radio-frequency (rf), or alternating-current (ac) magnetron sputtering, or by ion-beam sputtering. All these methods require expensive vacuum equipment. Less precise coatings can be deposited by chemical vapor deposition or from a liquid phase.

Equipment for these processes is usually considerably cheaper, and is used for depositing thin-film filters for consumer products whenever the accuracy is adequate. When choosing the deposition process, other considerations are whether it lends itself for the manufacture of large areas of coatings with uniform thickness, whether the properties of the resulting coatings are chemically stable and do not change with time, whether the produced coating's electrical properties are damaged by the high power of laser systems, and whether the mechanical properties of the coatings will withstand the expected wear and tear. Meeting the nonoptical requirements is often more challenging than obtaining the required optical performance. *See* CHEMICAL VAPOR DEPOSITION; LASER; PHYSICAL VAPOR DEPOSITION; RADIATION DAMAGE TO MATERIALS; SPUTTERING.

Theory. In its simplest manifestation, a coating can consist of a single layer deposited on a glass surface. Some of the light incident on this film will be reflected from the air–film interface, some from the glass–film interface, and the remainder will be transmitted into the glass. Depending on the value of the refractive index of the coating material relative to that of the glass and the relative thickness of the film to the wavelength of the incident light, constructive or destructive interference will take place between the two reflected electromagnetic waves, and the reflectance of the glass–air interface onto which the layer has been deposited can be enhanced or reduced. For such a simple system, it is possible to write down explicit expressions for the resulting reflectance. But most applications require filters that consist not of one thin film but tens or hundreds of thin films. A schematic representation of an optical thin-film interference filter is shown in **Fig. 1**. The filter consists of l layers deposited onto a substrate of refractive index n_s and immersed in a medium of refractive index n_m. The parameters available to the thin-film designer to achieve the desired performance include n_s, n_m, and the refractive indices n_i, as well as extinction coefficients k_i, and thicknesses d_i of the l layers. Clearly, the more complicated the required spectral performance of a filter is, the greater the number of layers l and the higher the overall thickness of the multilayer coating that is needed to yield the required result. For accurate results during the design process, it is necessary to take into account not only the primary reflections from each interface but also the effects of multiply reflected beams from the various interfaces. The performance of such systems is best evaluated with the aid of computer programs that are based on Maxwell's equations for the propagation of electromagnetic waves. Depending on the design method used, the solution to the problem may consist of a single, thick inhomogeneous layer in which the optical constants vary gradually as a function of distance from the substrate, or of a multilayer system consisting of many homogeneous layers of two or more different materials. The latter are more readily implemented. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; MAXWELL'S EQUATIONS; OPTICAL MATERIALS; REFLECTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES.

Optical interference filters and coatings. The following are examples of the most important types of optical interference filters and coatings.

Antireflection coatings. Antireflection coatings are the first application of optical thin-film interference. Around 1817, Joseph Fraunhofer discovered that tarnished glasses reflected less light than freshly polished glasses. He then proceeded to develop a process for artificially aging glasses to achieve this effect. The results were the first antireflection coatings. To this day, antireflection coatings are one of the most important optical thin-film components in both scientific and consumer instrumentation. The reflection of light from the surface of a transparent substrate at normal incidence can be as little as 4% for glass and as much as 30% for germanium, a material that is used extensively in infrared (IR) optics. Thus, germanium optics without antireflection coatings would lead to huge light losses. In

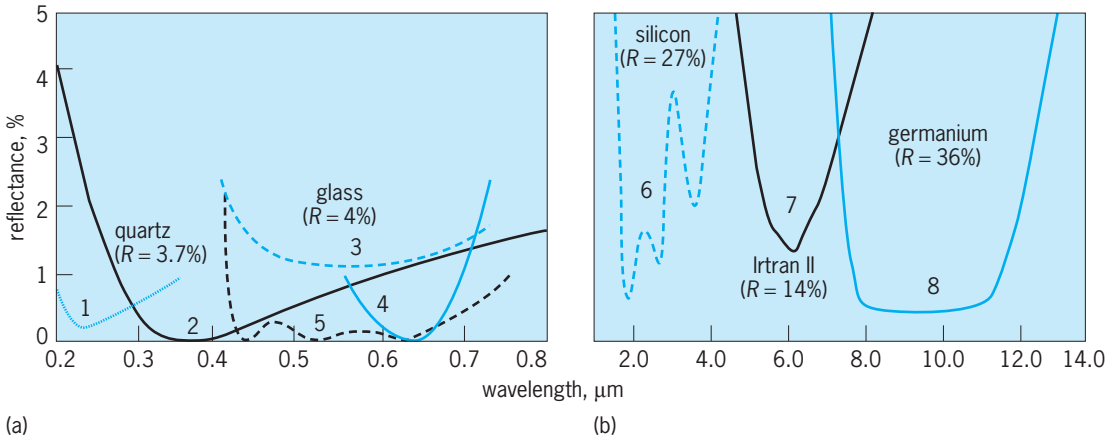

Fig. 2. Measured performance of antireflection coatings for quartz (curves 1, 2), glass (3–5), silicon (6), Irtran II [ZnS] (7), and germanium (8) surfaces in the (*a*) ultraviolet-visible and (*b*) infrared spectral regions. The numbers in brackets represent the reflectances of the surfaces without the antireflection coatings.

many instances, even glass surfaces must be antireflection-coated. For example, telephoto lenses for television and personal cameras can consist of many glass–air interfaces that must be antireflection-coated if the lenses are to be useful. This not only increases the light throughput of the objectives but also prevents the formation of stray light though multiple reflections that would reduce the contrast in the image plane. **Figure 2** shows the reflectances of substrate–air interfaces of various materials before and after treatment with antireflection coatings. Unless otherwise stated, the curves shown in Fig. 2 and subsequent diagrams are measured values. At the present state of thin-film deposition technology, the difference between theory and experiment is not very significant in most cases. Another special type of antireflection coatings, not shown in Fig. 2, is black layer coatings. Such coatings do not transmit or reflect any of the incident light in a given spectral region but absorb all the incident light. Such coatings find uses in high-contrast display devices and energy conversion coatings.

High-reflection coatings. These coatings are used whenever surfaces are required to have a high reflectance. At one time, metallic surfaces were the only way to achieve this goal; that is, metal substrates were polished until they formed mirrors. More recently, opaque metallic layers have been deposited onto polished glass surfaces to achieve the same effect. Some typical spectral reflectance curves of metals are shown in **Fig. 3**. They have broad regions of high reflectivity. However, when the highest reflectances are required, interference effects in stacks of nonabsorbing multilayers are used. In this way, reflectances as high as 99.9998% can be achieved in narrow parts of the visible and near-infrared spectral regions. Such values are required for some laser applications. Using special designs consisting of a larger number of layers of suitable thicknesses, high reflectances can also be achieved over broader spectral regions. For femtosecond lasers, reflectors are required whose phase dispersions must meet certain criteria. High-reflectance coatings are made for wavelengths ranging from the x-ray to the submillimeter spectral regions, although the reflectances that can be achieved depend greatly on the optical constants of the coating materials available in a given region. The performances of some typical high-reflectance

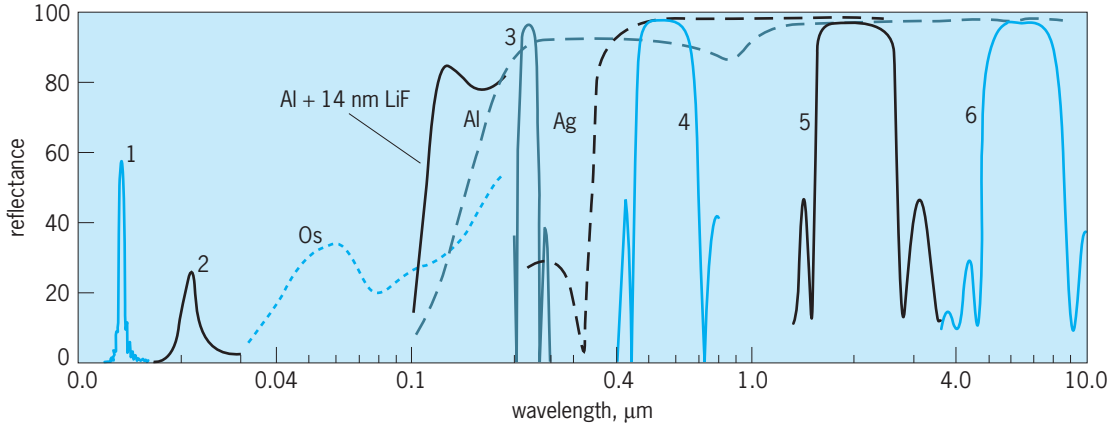

Fig. 3. Reflecting coatings. The solid curves represent typical measured reflectances of multilayer coatings for the x-ray (1, 2), ultraviolet (3), visible (4), and infrared (5, 6) spectral regions. The reflectances of opaque osmium, aluminum, and silver films (broken-line curves) are shown for comparison. The reflectances of metal layers can be extended or enhanced by additional thin-film coatings (for example, opaque aluminum with a 14-nanometer-thick LiF layer).

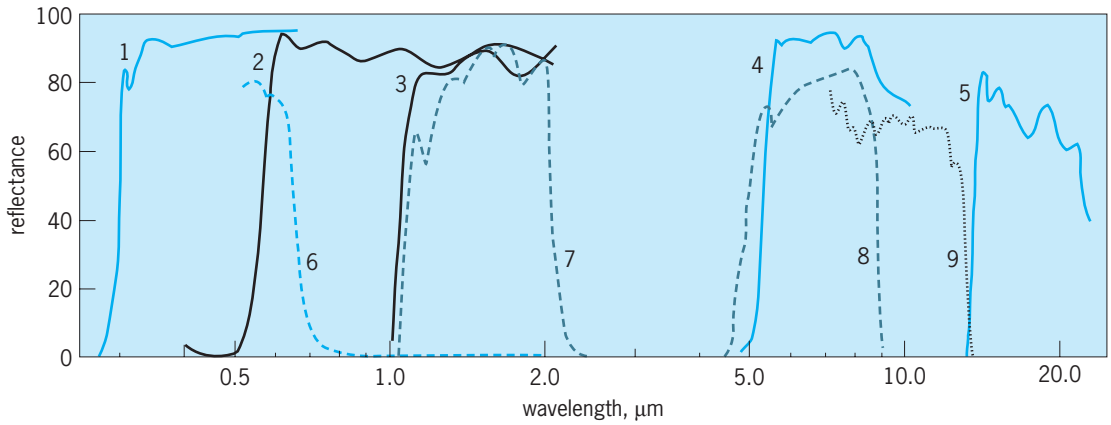

Fig. 4. Measured performances of short-wavelength (curves 1–5) and long-wavelength (6–9) cut-off filters for the ultraviolet, visible, and infrared spectral regions.

coatings for different wavelengths are shown in **Fig. 3**.

Long- and short-wavelength cut-off filters. These cut-off filters have an important technological application. For the visible and adjoining spectral regions, short-wavelength cut-off filters can be more easily produced from colored glasses. However, long-wavelength cut-off filters are difficult to produce except by thin-film interference. Even with this technique, the spectral ranges over which the transmittance remains high and the rejection low leave much to be desired. The performance of typical thin-film interference short- and long-wavelength cut-off filters is shown in **Fig. 4**.

Achromatic and color-selective beam splitters. Beam splitters are used in many scientific instruments, as well as in some consumer products. Suitable thin-film systems are deposited onto a plate or are located between two right-angled prisms. They split the obliquely incident light beam into two beams, one of which is transmitted and the other reflected. Usually when light is incident onto a multilayer system at nonnormal incidence, the transmittances and reflectances will be different for incident light that is polarized parallel (T_p, R_p) and perpendicular (T_s, R_s) to the plane of incidence. For beam splitters, this is a disadvantage and measures are often taken to reduce the effect. When the two beams have a uniform energy distribution in the spectral region of interest, the beam splitters are said to be achromatic. The beams may be of equal intensity, or one beam may be more intense than the other. Examples of the performance of such coatings are shown in **Fig. 5**. At other times, beam splitters are required in which the incident light is divided into beams of different color. Such beam splitters are called color-selective, and they are used in cameras and projection equipment for television and cinematography. The measured mean transmittances $(T_p + T_s)/2$ of one set of color selective beam splitters are shown in Fig. 5*c*.

Narrow-band-transmission interference filters. These are used in many scientific and engineering applications. In general, the specifications for narrow-band filters not only include the center wavelength and width at the points where transmittance is half of the maximum transmittance and the width of the rejection region where the reflectance is low, but also the widths at 0.1 and 0.01 of the maximum transmittance. The relative values of these widths are a measure of the

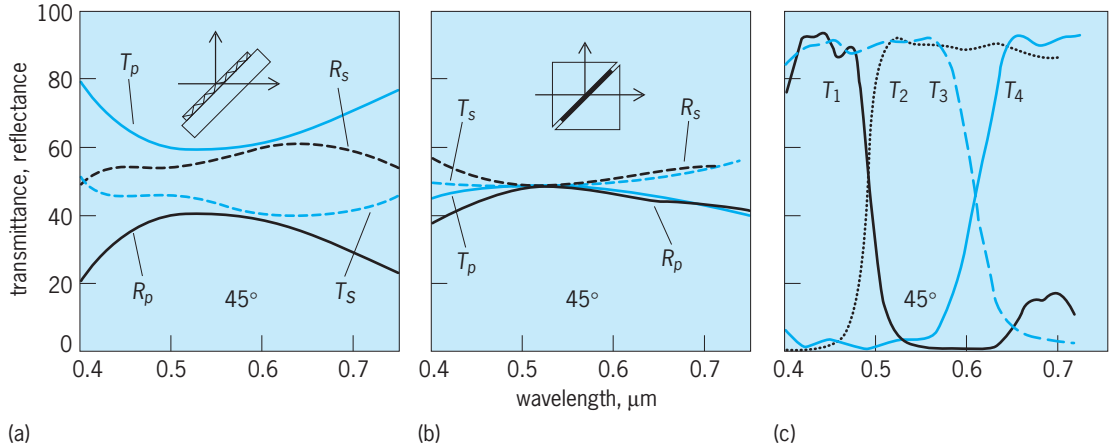

Fig. 5. Measured performance of (*a*) achromatic plate, (*b*) cube, (*c*) and color-selective beam splitters. In each case, the light was incident at 45° to the multilayer. The transmittance and reflectance may exhibit a strong dependence on the polarization of the incident light, as in *a* and *b*.

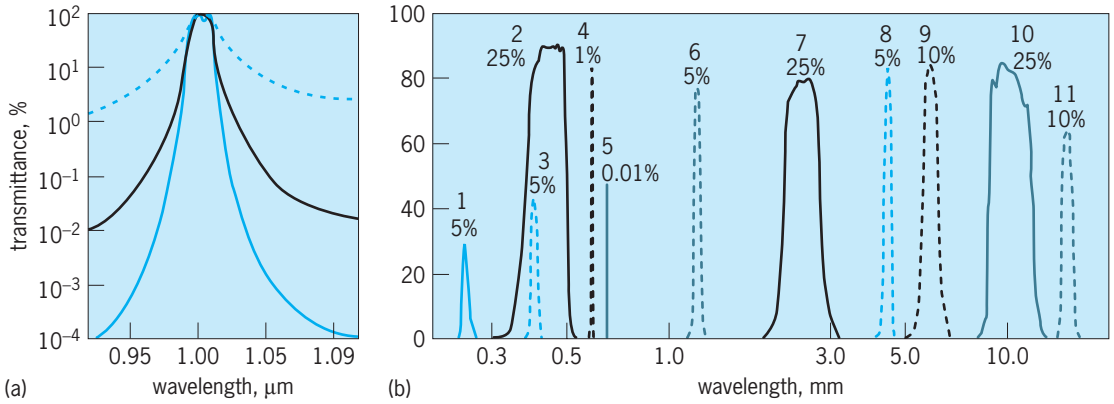

Fig. 6. Narrow-band transmission filters. (*a*) Calculated transmittance curves of filters having progressively squarer transmission bands. (*b*) Measured transmittances of a number of typical narrow-band interference filters for the ultraviolet, visible, and infrared spectral regions. The widths of the filters at half-maximum transmission, expressed as a percentage of the center wavelength, are indicated. All these filters do not transmit at shorter and longer wavelengths, except for the narrowest filter (curve 5), which requires additional blocking.

squareness of the transmission band (**Fig. 6***a*). The specifications are most stringent for filters for telecom applications, where the transmission in the rejection region may have to be as small as 10^{-5} and the transmission in the pass band may have to be in excess of 98%. To meet such stringent requirements, filters are required that often consist of more than 150 layers. A few representative narrow-band filters are shown in Fig. 6*b*. Narrow-band filters can be constructed with half-widths as narrow as 1 nanometer or as wide as 25% of the center wavelength.

Narrow-band rejection filters. Sometimes called minus filters, these filters should transmit all the incident radiation except for one or more specified wavelengths. Such filters are used in safety goggles for scientists and technologists working with lasers or for rejecting specific spectral lines in spectroscopic laboratories. Narrow-band rejection filters can consist of many layers made of two materials of close refractive indices, or they can consist of an inhomogeneous layer in which the refractive index oscillates gradually between two close values. Such coatings, often called rugate filters, have the advantage that the transmittance stays high over a broader range of wavelengths on the short-wavelength side of the rejection band. Examples of the performance of narrow-band rejection filters are shown in **Fig. 7**.

Polarizers and polarizing beam splitters. These find application in scientific instruments, as well as in liquid-crystal projection displays. Polarizers allow only light of one polarization to pass and block the light of the orthogonal polarization, while polarizing beam splitters separate unpolarized light into two orthogonally polarized transmitted and reflected beams of light. The purity of the polarized beam is measured by the extinction ratio, for example T_p/T_s or R_s/R_p, depending on whether the polarizer acts in transmission or reflection. Depending on the application, the polarizing devices might need to be effective over a narrow or broad spectral region. Most practical polarizers and polarizing beam splitters are based on isotropic interference multilayer coatings and need to operate with light incident at a specified oblique angle θ onto the multilayer. Most thin-film polarizers and polarizing beam splitters are effective only for light that is incident within a small angular field $\delta\theta_{air}$ around θ, or operate only in a narrow spectral

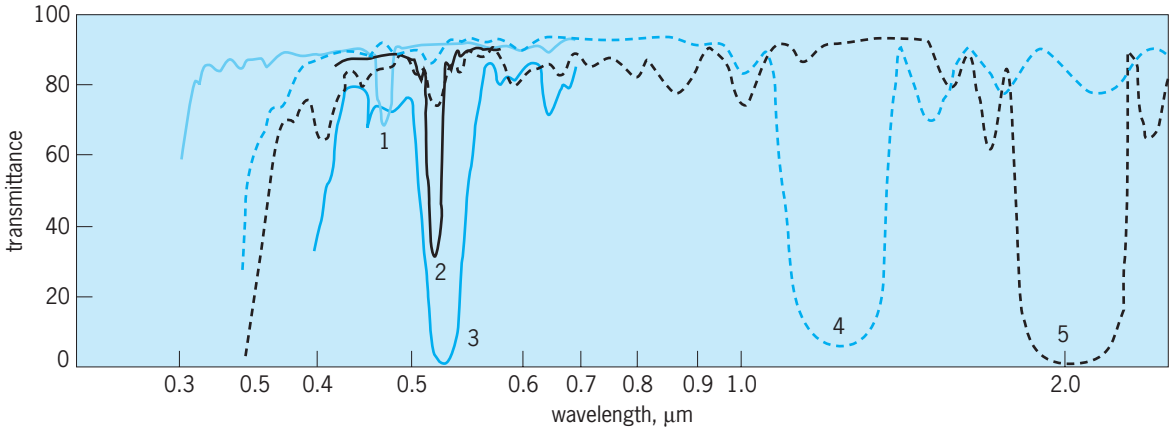

Fig. 7. Measured performance of narrow-band rejection filters based on homogeneous multilayers (curves 1–3) and inhomogeneous layers (4, 5).

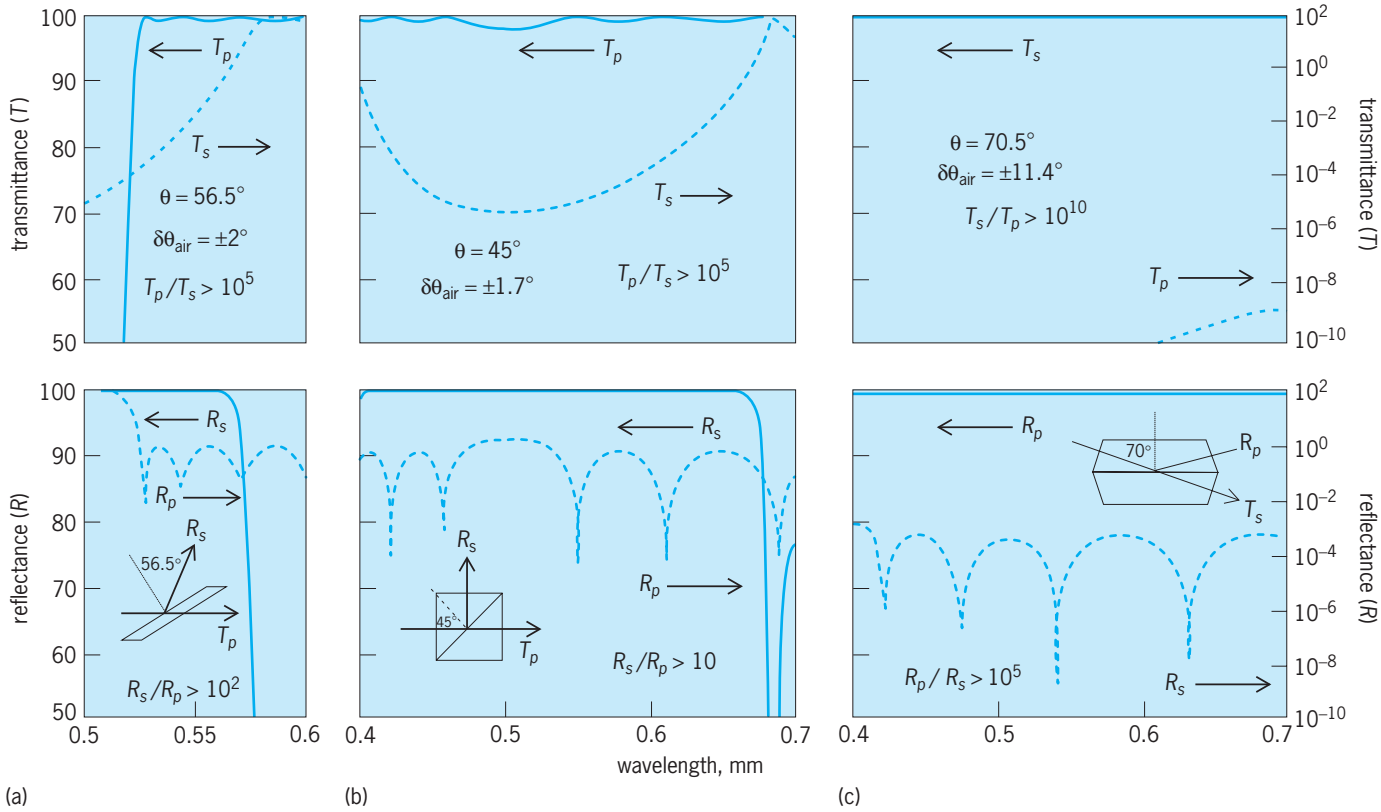

Fig. 8. Calculated performance of polarizers and polarizing beam splitters at the design angle θ. (*a*) Plate polarizers operate in a narrow spectral region with well-collimated incident light and can withstand high laser power. (*b*) MacNeille polarizing beam splitters operate over a broad spectral range, but their performances deteriorate rapidly with the departure of the incident beam from the design angle. (*c*) The Li-Li polarizing beam splitter operates over a broad spectral range and has a wide angular field for both the transmitted and reflected beams.

region. An exception is the Li-Li polarizing beam splitter, which is not yet (mid-2006) available commercially. The calculated spectral transmittances T_p, T_s and reflectances R_p, R_s of three types of polarizers and polarizing beam splitters at the design angles of incidence are shown in **Fig. 8**. Also given in the diagrams are the approximate extinction ratios for the transmitted and reflected beams at the edges of their angular fields measured in air. *See* POLARIZATION OF WAVES; POLARIZED LIGHT.

Correction and gain-flattening filters. Correction and gain-flattening filters are used when the spectral energy distribution of the light source or sensitivity of a detector vary too much for the satisfactory operation of a spectral or telecommunications device. The spectral transmittance of a correction or gain-flattening filter is such that in combination with the source and the detector, it will produce a more uniform response in the spectral region of interest. Correction filters often need to have very irregular spectral transmittances with values varying by orders of magnitude, especially when the light sources used have sharp emission lines.

State of the art. The thin-film interference filters and coatings whose properties are shown in Figs. 2–8 are only a few examples of what can be produced, since filters with properties intermediate to those in the graphs are possible. However, the performance that can be achieved depends also on the spectral region. At the outer edges of the spectrum for which interference filters are designed, the performance can be much worse than that in the visible and near infrared because of the lack of suitable coating materials.

The state of the art of thin-film optical interference filters has made great strides since the 1950s. It is now possible to design filters with almost any spectral transmittance or reflectance characteristics for normally incident light, provided that the spectral region over which they are specified is not too wide. Deposition processes exist now for the manufacture of very stable, mechanically durable coatings. Process controls have been developed that make it possible to produce quite complex multilayer coatings automatically. The accuracy of experimentally produced thin-film interference filters is limited by the accuracy of the measurements made during their deposition. In the future, the use of nanostructured films, with properties that supplement those found in natural materials, should lead to filters with improved properties. J. A. Dobrowolski

Bibliography. P. W. Baumeister, *Optical Coating Technology*, 2004; J. A. Dobrowolski, Optical properties of films and coatings, in *Handbook of Optics*, vol. 1, 2d ed., ed. by M. Bass, pp. 42.1–42.130, 1995; H. A. Macleod, *Thin-film Optical Filters*, 3d ed., 2001; J. D. Rancourt, *Optical Thin Films Users' Handbook*, 1987.

Interference-free circuitry

Circuitry that is designed to ensure that external sources of interference do not degrade its performance. The design principles involved are very simple, and their application in a systematic manner will reduce interference from external sources to an acceptable level, even for the most sensitive circuits and applications.

Electrical currents flow around complete, closed paths. These paths contain the energy sources which generate the currents. In complicated circuit networks comprising many meshes, a current from a given source may well divide and subdivide as it encounters circuit junction points, but these partial currents must recombine to give the original total current on return to a source. Unfortunately, most diagnostic equipment (such as oscilloscopes, voltmeters, and spectrum analyzers) responds to voltage differences, and the currents which cause these must be inferred. Nevertheless, attention must be concentrated on the currents flowing around circuits rather than voltages observed between points.

Coupling modes. Consider the simplest possible circuit containing an impedance and a voltage or current detector, as illustrated on the left of Figs. 1, 2, and 3. Suppose it is close to another circuit consisting of a source of interference and a load impedance, as on the right of these diagrams. There are only three ways in which interference from the source can couple in to the left-hand circuit and be registered as an unwanted signal by the output device D.

1. In **Fig. 1***a*, the two circuits have a portion of conductor A-B which is common to both. The current *I* in the right-hand circuit, which is the source of the interference, flows through this shared conductor, and this results in a voltage difference between A and B. This voltage in turn causes an unwanted interfering current to flow around the left-hand circuit, and to be registered by D. We could term this mode of coupling interference "common-conductor." It will be effective over a wide range of interfering frequencies and for both high and low values of the impedances in both circuits. Supposing the circuits must be connected at all, the measure to take to prevent interference coupling-in via this mode is to reduce the length of common conductor to a single point, as in Fig. 1*b*. If the circuits have to be spatially separated, a single conductor can be connected between the two circuits, as in Fig. 1*c*; and it is obviously still true that however large the current in the interfering circuit, no voltage difference or consequent interference current is produced in the left-hand circuit.

2. In **Fig. 2***a*, there is a connection between the circuits, and a stray capacitance *C* is shown between them. In Fig. 2*b*, there is no connection, but two stray capacitances are shown. Currents can circulate via capacitances to cause interference in the left-hand circuit. This capacitive coupling mode is usually significant at higher frequencies for circuits containing high impedances, say of the order of 100 Ω or more.

This coupling mode can be eliminated by strategically placed and connected conducting screens between the circuits, as shown in Fig. 2*c*.

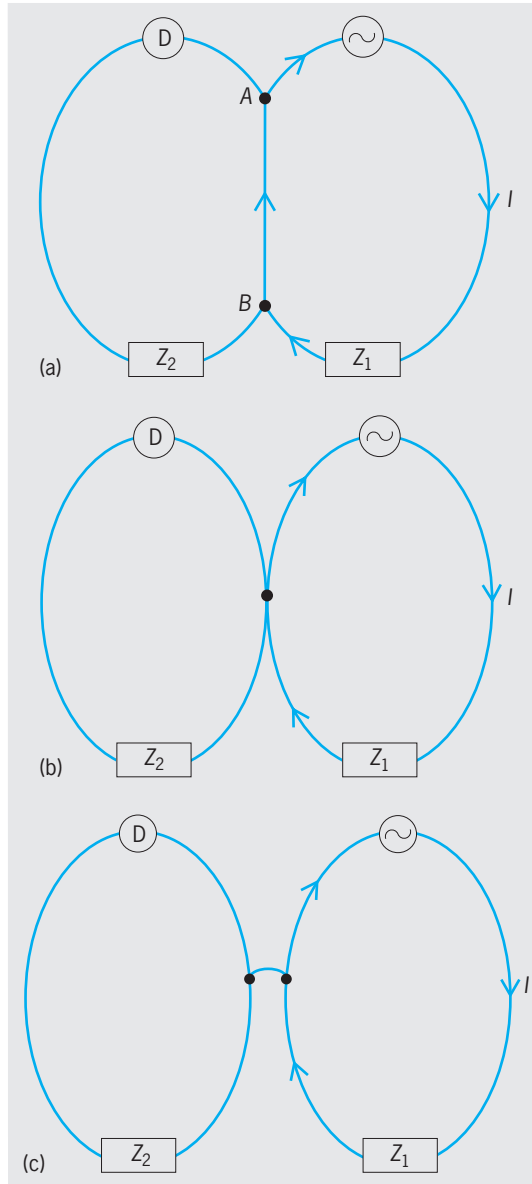

Fig. 1. Common-conductor coupling of two circuits. (*a*) Two circuits sharing a common conductor *AB*. *Z* represents load impedance, D is a detector, ~ represents a source of interference, and *I* is current. Coupling is prevented by (*b*) reducing the common conductor to a point or (*c*) connecting the circuits with a length of conductor.

3. In **Fig. 3***a*, the circuits have no common connection, but some of the magnetic flux *B* created by the interfering circuit threads the left-hand circuit and induces an interfering emf in it and a consequent interference current around it. This mode of interference is most significant where the left-hand circuit is of low impedance, that is, of the order of 10,000 Ω or less.

The way to avoid this magnetic coupling mode is to reduce the coupling between the circuits by either increasing the distance between them or decreasing to a very small value the area enclosed by one or, preferably, both circuits. Figure 3*b*, *c*, and *d* illustrates three ways of accomplishing the latter. In Fig. 3*b*, the conductors of both circuits are routed

close to each other in go-and-return pairs. In Fig. 3*c*, the area enclosed by the go-and-return pairs is further reduced by being twisted, with the advantage that the small flux generated is in opposing directions from successive loops of the twist and thereby cancels at a little distance from the circuit to an exceedingly small value. In Fig. 3*d*, the go-and-return conductors are arranged to be the inner and outer conductors of a coaxial cable. The line integral of

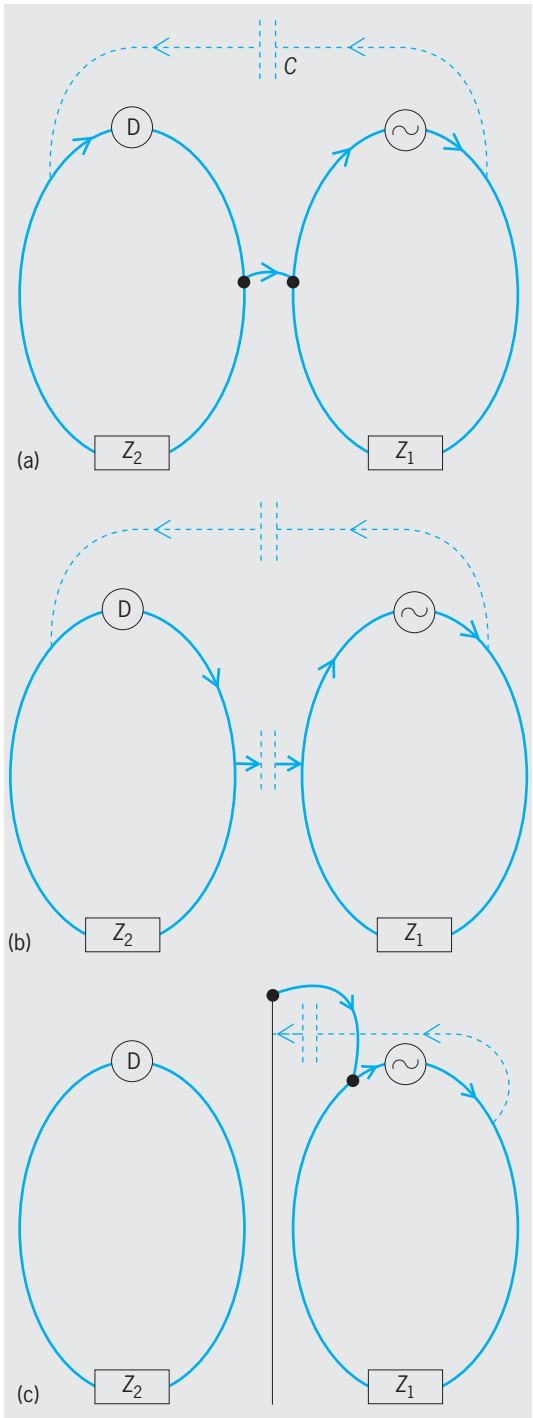

Fig. 2. Capacitive coupling of two circuits. (*a*) An interference current circulating via a stray capacitance *C*. (*b*) An interference current circulating via two stray capacitances. (*c*) Elimination of coupling by adding a screen to intercept capacitive currents.

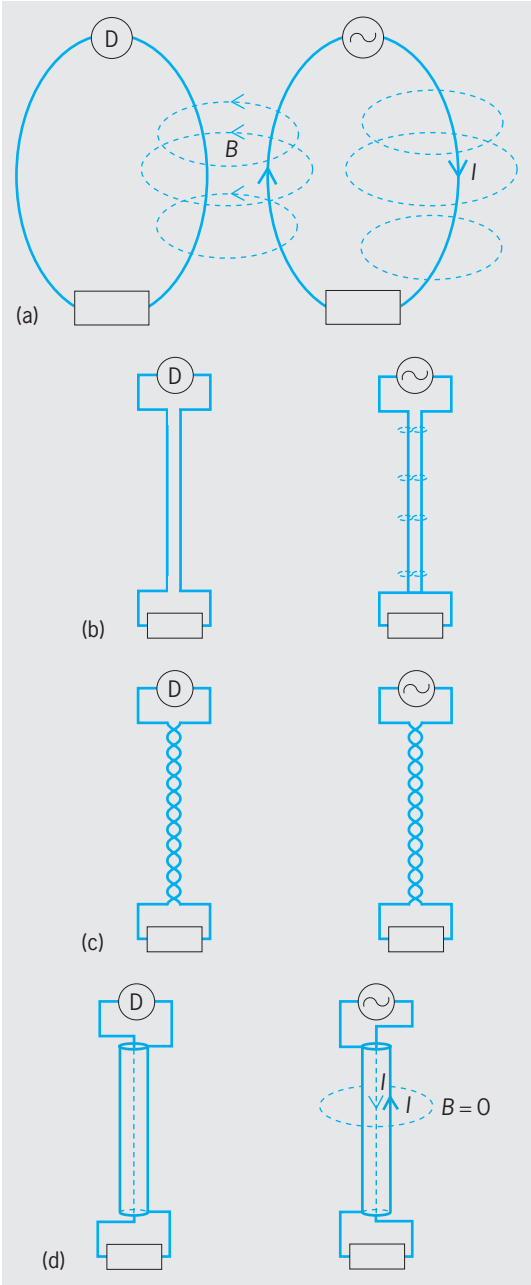

Fig. 3. Magnetic coupling of two circuits. (*a*) The magnetic field *B* created by the interference current *I* threads the left-hand circuit. (*b*) Decreasing the loop area of both circuits reduces magnetic coupling. (*c*) Twisting the conductors further reduces coupling. (*d*) An alternative solution employs coaxial connecting cables.

magnetic flux around a concentric circular path of any radius outside the cable is zero because the total current threading it is zero. Because of the cross-sectional symmetry of a circular coaxial cable, this line integral can be zero only if the flux itself is zero everywhere. Although this result is strictly true only for a straight cable of infinite length, flux elimination is good enough for all practical purposes for short lengths of cable if the cable is not too sharply bent anywhere. A further advantage of the coaxial solution is that if, as is usual, the outer conductors of the sending and receiving circuits are at more or less the same potential and involve no appreciable impedance, they act as screens, and capacitive

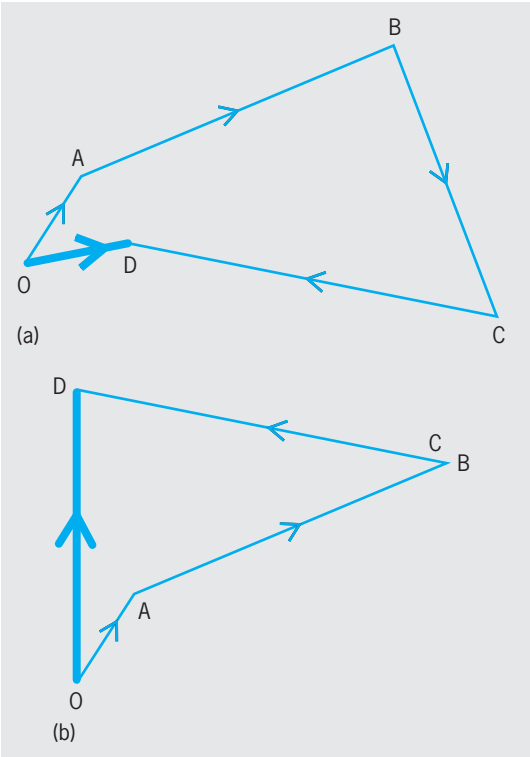

Fig. 4. Vector diagrams illustrating an unexpected result. (*a*) Summing vectors *OA*, *AB*, *BC*, and *CD* which represent interference at a single frequency provides the resultant *OD*. (*b*) The resultant *OD* is increased when the interference represented by *BC* is eliminated.

coupling is also reduced to negligible proportions. If, in an actual situation, the outer conductors are themselves interconnected as a network, at lower frequencies additional measures may need to be taken to ensure that the currents in the inner and outer conductors of each cable are equal and opposite. One way of doing this is to thread one cable in each mesh several times through a high-permeability magnetic core to ensure that the total impedance of the mesh is much greater than the go-and-return impedance. The latter is unaffected by the presence of the core. *See* COAXIAL CABLE; COUPLED CIRCUITS.

Need for systematic approach. In any actual situation, more than one of the three modes will be acting simultaneously, and in a complicated network there will be several coupling routes associated with each mode. Therefore a systematic approach is needed whereby each and every possible route for the ingress of interference is considered in turn and appropriate action taken.

Unfortunately, eliminating interference systematically in the common situation where interference at a single frequency enters circuitry through several routes or through more than one mode can cause puzzling results. An example of single-frequency coherent interference is that from mains or line frequency and its harmonics. In this case there may well be partial cancellation at the final output device of the interference entering via the various routes. This situation is illustrated in the vector diagram of **Fig. 4***a*, where the vector representing the sum of the interference is *OD*. Then if one of the routes of entry is eliminated, say that corresponding to the vector *BC*, the remaining interference may well sum to larger value, as in Fig. 4*b*. This will make a trial-and-error approach impossible, and will force a more considered and systematic approach to be taken. Sometimes, coherent interference of this kind may well accidentally sum to an acceptably small value, or another route of adjustable phase and magnitude may be deliberately created to oppose the total interference. This situation is not a good one, because any change in the circumstances or environment of the circuitry will alter the magnitude and phase of the various coupling modes and the interference will reappear. Besides, interference in one part of sensitive circuitry may cause it to become saturated and to malfunction. Therefore the aim should be reduction to acceptable levels of interference via all possible routes. In the example of mains/line interference, this will have the added advantage that any other interference carried and propagated via mains/line circuitry will also be eliminated.

Isolation. Isolation is often an important concept in eliminating interference. An instrument or section of circuitry is said to be isolated if there are no routes by which significant currents, interfering or otherwise, can flow through the instrument or circuitry from a source onward to other circuitry. Circuitry entirely contained within a conduction enclosure apart from go-and-return current paths to supply power to it will be isolated. This description applies to many test and measurement instruments, even when connected to circuitry which is itself isolated, such as a resistance-measuring instrument connected to a resistance thermometer whose leads and sensing element are enclosed in a continuous shield. This shield should be connected only to the shield of the instrument, thus forming an extension of it. Unfortunately, isolation is usually compromised by the mains/line and power supply connections if the instrument or circuitry is connected into a larger, overall system of other instruments and circuits. One way of maintaining isolation in these circumstances is to make use of differential voltage-sensing input circuits which exhibit high input impedance. Battery-operated instruments whose batteries are contained entirely within the conducting case of the instrument can also offer a way out of some difficult situations. *See* ELECTRICAL SHIELDING.

Misconceptions. Finally, clear thinking about interference elimination will be aided if two common misconceptions are recognized and avoided.

1. Interference does not enter circuitry at localized points or circuit nodes. A complete current path must be involved.

2. For brevity in circuit descriptions and diagrams, many connections are said to be made to a conductor called ground. The term is best avoided, but if old habits die hard, at least it should be recognized that this name is shorthand for a whole network of conductors which return currents to their sources, thereby completing circuit loops for interference coupling modes. This network may well not be an equipotential conductor system because of the currents flowing in it, and may itself give

rise to common-conductor coupling, which can be avoided by the star or single-conductor connections described above. *See* ELECTRICAL INTERFERENCE; GROUNDING; ELECTROMAGNETIC COMPATIBILITY.

Bryan P. Kibble

Bibliography. *A Guide To Measuring Direct and Alternating Current and Voltage Below 1 MHz*, Institute of Measurement and Control, U.K., 2003; P. Horowitz and W. Hill, *The Art of Electronics*, 2d ed., Cambridge University Press, 1989; R. Morrison, *Grounding and Shielding Techniques in Instrumentation*, 3d ed., Wiley-Interscience, 1986.

Interference microscope

An instrument for visualizing and measuring differences in the phase of light transmitted through or reflected from microscopic specimens. It is closely allied to the phase-contrast microscope. *See* PHASE-CONTRAST MICROSCOPE.

Simple theory. A microscopic image is formed by the interaction of all light waves passing through the optical system. In a phase-contrast microscope, light diffracted by a transparent object is spatially separated from nondiffracted light in the back focal plane of the objective, where a phase plate alters the relative phases of the diffracted and nondiffracted beams so that they interfere in the image plane to produce a visible, intensity-modulated image. In an interference microscope, the diffracted and nondiffracted waves are not spatially separated, but light (the object beam) that passes through or is reflected from the object interferes with light (the reference beam) that passes through or is reflected from a different region of the specimen plane or is reflected from a comparison (reference) surface. For interference to be visible, the light beams must be coherent; in other words, the beams must maintain a constant relationship of wavelength, phase, and polarization over a relatively long period. The easiest way to achieve coherence is by using a device such as a semireflecting mirror, which splits a light beam into two beams. Random changes in the properties of successive photons from a given point in the source then affect both beams simultaneously. For a Mach-Zehnder type of interferometer (**Fig. 1**), light from a source is split into two beams, which after total reflection at the mirrors are recombined. If the optical paths are identical, the light waves arrive in phase at the point where an image is formed, and interfere constructively. A retarding specimen in one beam throws the beams out of phase and produces partial or complete destructive interference where the image is formed. Differences in optical path introduced by various parts of the object can be seen as variations in intensity or color. It may be noted that energy is conserved: light "missing" from the image appears at a right angle to it. The optical path difference introduced by the specimen expresses a relationship between the geometric thickness, and the refractive indices of the reference space and the specimen. The optical path difference can be measured by a calibrated compensator. It restores the relative phase of the two beams. *See* INTERFERENCE OF WAVES.

Many forms of interferometer have been described, but few have been used in commercial microscopes. A Leitz interference microscope using the Mach-Zehnder principle had two condensers, two objectives, and wide separation of the specimen and reference beams. Maintenance of the necessary mechanical rigidity was difficult and expensive, however, and so several compact, single-objective polarizing microinterferometers have been devised based on the Jamin-Lebedeff principle. In this type of instrument, plane-polarized light from a substage polarizer is split by a birefringent plate cemented to the condenser into two parallel beams polarized at right angles to each other and separated, or sheared, by a distance of about 30–200 micrometers, depending on the magnification of the objective used. After passing through the specimen plane, the beams are recombined by a birefringent plate cemented to the objective. The relative phase of the beams can be adjusted with a polarizing compensator such as a quartz wedge or a compensator that combines a fixed quarter-wave plate and a rotating analyzer. *See* BIREFRINGENCE.

Because of the beam splitters, an interference

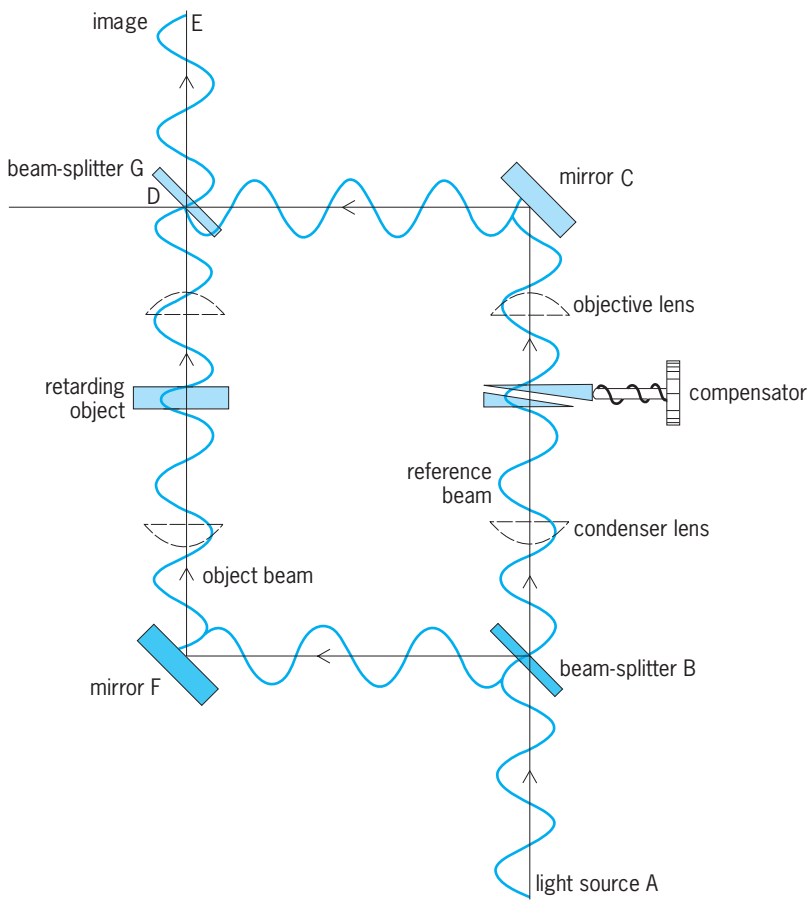

Fig. 1. Diagram of a Mach-Zehnder type of interference microscope (eyepiece not shown). Light from source A is split at B into coherent object and reference beams, which are reflected at F and C, respectively, and recombined at D. If the optical paths are identical, constructive interference occurs at E and complete destructive interference occurs at G. That relationship, if disturbed by a retarding specimen in the object beam, can be restored by a compensator in the reference beam. (*After D. J. Goldstein, Scanning microinterferometry, in G. A. Meek and H. Y. Elder, eds., Analytical and Quantitative Methods in Microscopy, Cambridge University Press, 1977*)

microscope displays two separate images of the specimen. In some instruments, one of these, the so-called ghost image, is astigmatic and out of focus. If the ordinary and ghost images overlap, as can occur with closely packed or large specimens, reliable measurements or observations in the overlapping region are impossible.

Measurement of phase changes. To measure the retardation of a relatively thick specimen, a gradient of retardation can be superimposed on the microscopic field by tilting an optical element or by inserting a retarding wedge in one beam. This produces a series of dark fringes across the field, uniformly spaced at intervals of one wavelength retardation. The deviation of fringes crossing a specimen is proportional to its retardation; whole wavelengths are readily counted, and fractions of a wavelength can be estimated with a precision of about 0.1 wavelength. To measure the retardation more precisely, a calibrated compensator (Fig. 1) is adjusted so that the background is maximally dark by using monochromatic light or some readily identifiable color such as sensitive purple. The compensator setting is recorded and then readjusted until the specimen appears with the same intensity or color as the background had. The difference between the two settings indicates the specimen retardation. The precision of visual setting can be improved appreciably by a half-shade device in the eyepiece, and under ideal circumstances is about 0.01 wavelength. With automatic microinterferometers that use mechanically phase-modulated light and a photoelectric detector, the reproducibility of repeated measurements of optical path difference exceeds 0.001 wavelength. The methods so far described measure the retardation at an arbitrary point in the object plane. In cell biology, however, it is often more useful to measure the integrated optical path difference of a discrete, nonhomogeneous specimen, which is the product of the specimen area and the mean optical path difference. The integrated optical path difference can be estimated by measuring the specimen area and by taking many measurements of local optical path differences or—more rapidly and precisely—by using a microinterferometer in its scanning and integrating mode.

Differential interference–contrast microscope. In Mach-Zehnder and Jamin-Lebedeff interference microscopes, the ordinary and ghost images of a discrete object are relatively widely separated, or sheared, and the specimen is compared with a clear region of background. In differential interference–contrast (DIC) instruments, on the other hand, the shear is very small and the images overlap almost completely; a given point of the specimen is compared not with empty mounting medium but with an adjacent object point, and contrast depends on the gradient of optical path in the direction of shear (**Fig. 2**). Consider, for example, a uniform sphere mounted in a medium of different refractive index. In the center of the image the optical path gradient is zero and the intensity is identical with the background. At one edge the gradient is maximal and the image is brighter than the background,

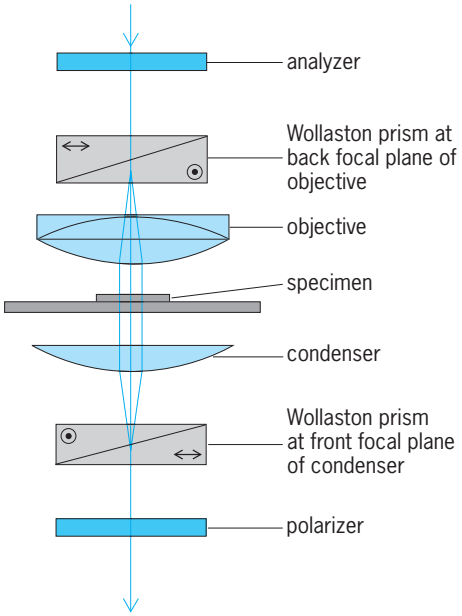

Fig. 2. Principle of a differential interference microscope with Wollaston prisms at the focal planes of the condenser and objective. Nomarski devised modified Wollaston prisms that can be used if the focal planes are not physically accessible. (*After W. Lang, Nomarski differential interference–contrast microscopy, Zeiss Inform., no. 70, pp. 114–120, 1969*)

whereas at the other edge the gradient is opposite in sign and the image is darker than the background. The object appears almost as if shadow-cast. With a given objective and a large condenser aperture, differential interference–contrast gives good lateral resolution and a small depth of focus; as a result, a degree of optical sectioning through the specimen can be achieved. Reducing the condenser aperture increases the contrast somewhat, but differential interference–contrast is less sensitive in general than phase contrast for detecting small phase differences. Note that in transmitted-light differential interference contrast, the contrast is related not to the gradient in geometrical object thickness but to the gradient in optical path, and so the image does not necessarily indicate true surface structure. For example, an internal cell organelle such as a nucleus, with a refractive index higher than that of the cytoplasm, may incorrectly seem to project from or to be an indentation in the surface of the cell. In spite of those interpretation problems, the striking three-dimensional appearance of differential interference contrast images has made the method popular in the field of biology.

Reflecting specimens. Surface structure can be studied with the Linnik microscope. If its mirror is tilted slightly relative to the reflecting surface, interference fringes appear and irregularities in the reflecting surface are seen as displacements in the fringe system. If the two surfaces are exactly parallel, the fringe will be of infinite width so that irregularities in the reflecting surface appear as intensity variations against a uniformly illuminated background, as in phase contrast. Nomarski differential interference contrast is also eminently suitable for

reflected-light work. Excellent results have been obtained relatively inexpensively by using very narrow interference fringes produced by multiple reflection with an ordinary incident illuminator and a partially reflecting cover glass over the specimen.

Nomarski differential interference contrast is eminently suitable for reflected-light work. Also, there are phase-stepping or phase-shifting interference microscopes with which highly reproducible and accurate measurements of a surface can be made automatically and rapidly. This type of instrument typically uses a Mirau interference objective to illuminate the specimen with incident light, and a charge-coupled-device camera to capture an image with superimposed interference fringes. Intensities in the image are measured automatically, and the data are transferred to a computer. A small phase difference is introduced between the object and reference beams by moving either the reference surface or the specimen with a piezoelectric transducer controlled by the computer; this alters the fringes, and the intensities are again recorded. Provided at least three sets of intensity measurement are available, it is possible to compute the phase change at, and hence the height of, each image point. Some confocal scanning microscopes are able to function as microinterferometers, with the characteristic benefits of improved lateral resolution and optical sectioning ability given by confocal scanning. *See* CONFOCAL MICROSCOPY.

Comparison with phase-contrast microscope. The phase-contrast microscope can be regarded as an imperfect interference microscope in which the necessarily finite width of the phase annulus produces incomplete separation of the two interfering beams. The resulting halo effect and accentuation of edges are absent in some types of interference microscope. Although eliminating optical artifacts reduces the risk of incorrect image interpretation, it is not necessarily advantageous for purely observational work. Ghost images may hamper considerably the use of the interference microscope to view extended or densely packed specimens. In addition, the relative sensitivity, low cost, and ease of adjustment of the phase-contrast microscope make it the instrument of choice for routine observations and for measurement of refractive indices by the immersion method. The interference microscope is, however, essential for measuring phase change.

Biological applications. Knowing the thickness of the section being examined is often valuable in stereological work. Plastic sections more than 1 μm in thickness can readily be measured with a visual microinterferometer, and with an automatic instrument even ultrathin sections (about 50–70 nanometers) for electron microscopy can be measured with a reproducibility better than 5%.

The primary biological use of microinterferometry, however, is probably the determination of cellular dry mass. With an automatic scanning and integrating microinterferometer, the dry mass of a living mammalian cell can be precisely and nondestructively determined in a few seconds. Because measured optical path difference depends on the product of thickness and refractive index difference, only one of those quantities need be known to calculate the other. Thus the combination of interference microscopy and immersion microrefractometry is capable of yielding much useful quantitative information. In some cases it has proved possible to determine cellular volume, thickness, concentration of solids and water, and total dry and wet mass. Interference microscopy has also been successfully used in quantitative histochemistry for studying changes in dry mass produced by enzymic digestion or specific extraction procedures.

Nonbiological procedures. Like the phase-contrast microscope, the transmission interference microscope can be applied to the qualitative study of transparent specimens such as fibers and crystals. Incident-light instruments are employed in metallurgy, engineering, and the electronic industry to study specimens such as metallic surfaces, metalized replicas of surfaces, and electronic circuits. In that type of work it is a common practice to operate the instrument with fringes in the field of view, the shape of the fringes giving a contour map of the surface. The sensitivity is such that irregularities of the order of 0.01 wavelength or even less can be observed. The method gives valuable information about the quality of ground, polished, or etched surfaces. The differential (Nomarski) type of interference microscope has proved particularly useful for studying the surface structure of reflecting objects, since with that type of specimen it gives a true relief image uncomplicated by variations in refractive index. Vertical resolution on the order of 0.1 nm has been obtained with incident-light differential interference contrast microscopes combined with electronic image-analysis devices. *See* DIFFRACTION; OPTICAL MICROSCOPE; POLARIZED LIGHT; POLARIZED LIGHT MICROSCOPE; RESOLVING POWER (OPTICS). D. J. Goldstein

Bibliography. M. Francon and S. Mallick, *Polarization Interferometers: Applications in Microscopy and Macroscopy*, 1971; W. Lang, Nomarski differential interference–contrast microscopy, *Zeiss Inform.*, nos. 70 and 71, 1969; G. A. Meek and H. Y. Elder (eds.), *Analytical and Quantitative Methods in Microscopy*, 1977; M. Pluta, *Advanced Light Microscopy*, vol. 3, 1993; S. Tolansky, *Multiple-Beam Interference Microscopy of Metals*, 1970.

Interference of waves

The process whereby two or more waves of the same frequency or wavelength combine to form a wave whose amplitude is the sum of the amplitudes of the interfering waves. The interfering waves can be electromagnetic, acoustic, or water waves, or in fact any periodic disturbance.

The most striking feature of interference is the effect of adding two waves in which the trough of one wave coincides with the peak of another. If the two waves are of equal amplitude, they can cancel each other out so that the resulting amplitude is zero. This

is perhaps most dramatic in sound waves; it is possible to generate acoustic waves to arrive at a person's ear so as to cancel out disturbing noise. In optics, this cancellation can occur for particular wavelengths in a situation where white light is a source. The resulting light will appear colored. This gives rise to the iridescent colors of beetles' wings and mother-of-pearl, where the substances involved are actually colorless or transparent.

Two-beam interference. The quantitative features of the phenomenon can be demonstrated most easily by considering two interfering waves. The amplitude of the first wave at a particular point in space can be written as Eq. (1), where A_0 is the peak am-

$$A = A_0 \sin(\omega t + \varphi_1) \tag{1}$$

plitude, and ω is 2π times the frequency. For the second wave Eq.(2) holds, where $\varphi_1 - \varphi_2$ is the phase

$$B = B_0 \sin(\omega t + \varphi_2) \tag{2}$$

difference between the two waves. In interference, the two waves are superimposed, and the resulting wave can be written as Eq. (3). Equation (3) can be expanded to give Eq. (4).

$$A + B = A_0 \sin(\omega t + \varphi_1) + B_0 \sin(\omega t + \varphi_2) \tag{3}$$

$$A + B = (A_0 \sin\varphi_1 + B_0 \sin\varphi_2)\cos\omega t + (A_0 \cos\varphi_1 + B_0 \cos\varphi_2)\sin\omega t \tag{4}$$

By writing Eqs. (5) and (6), Eq. (4) becomes Eq. (7), where C^2 is defined in Eq. (8). When C is less than A

$$A_0 \sin\varphi_1 + B_0 \sin\varphi_2 = C \sin\varphi_3 \tag{5}$$

$$A_0 \cos\varphi_1 + B_0 \cos\varphi_2 = C \cos\varphi_3 \tag{6}$$

$$A + B = C \sin(\omega t + \varphi_3) \tag{7}$$

$$C^2 = A_0{}^2 + B_0{}^2 + 2A_0B_0 \cos(\varphi_2 - \varphi_1) \tag{8}$$

or B, the interference is called destructive. When it is greater, it is called constructive. For electromagnetic radiation, such as light, the amplitude in Eq. (7) represents an electric field strength. This field is a vector quantity and is associated with a particular direction in space, the direction being generally at right angles to the direction in which the wave is moving. These electric vectors can be added even when they are not parallel. For a discussion of the resulting interference phenomena *see* POLARIZED LIGHT; SUPERPOSITION PRINCIPLE.

In the case of radio waves or microwaves which are generated with vacuum tube or solid-state oscillators, the frequency requirement for interference is easily met. In the case of light waves, it is more difficult. Here the sources are generally radiating atoms. The smallest frequency spread from such a light source will still have a bandwidth of the order of 10^7 Hz. Such a bandwidth occurs in a single spectrum line, and can be considered a result of the existence of wave trains no longer than 10^{-8} s. The frequency spread associated with such a pulse can be written as

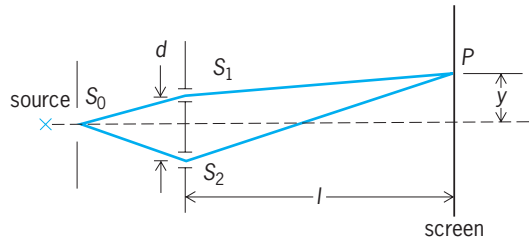

Fig. 1. Young's two-slit interference.

notation (9), where t is the pulse length. This means

$$\Delta f \simeq \frac{1}{2\pi t} \tag{9}$$

that the amplitude and phase of the wave which is the sum of the waves from two such sources will shift at random in times shorter than 10^{-8} s. In addition, the direction of the electric vector will shift in these same time intervals. Light which has such a random direction for the electric vector is termed unpolarized. When the phase shifts and direction changes of the light vectors from two sources are identical, the sources are termed coherent.

Splitting of light sources. To observe interference with waves generated by atomic or molecular transitions, it is necessary to use a single source and to split the light from the source into parts which can then be recombined. In this case, the amplitude and phase changes occur simultaneously in each of the parts at the same time.

Young's two-slit experiment. The simplest technique for producting a splitting from a single source was done by T. Young in 1801 and was one of the first demonstrations of the wave nature of light. In this experiment, a narrow slit is illuminated by a source, and the light from this slit is caused to illuminate two adjacent slits. The light from these two parallel slits can interfere, and the interference can be seen by letting the light from the two slits fall on a white screen. The screen will be covered with a series of parallel fringes. The location of these fringes can be derived approximately as follows: In **Fig. 1**, S_1 and S_2 are the two slits separated by a distance d. Their plane is a distance l from the screen. Since the slit S_0 is equidistant from S_1 and S_2, the intensity and phase of the light at each slit will be the same. The light falling on P from slit S_1 can be represented by Eq. (10) and from S_2 by Eq. (11), where f is the frequency, t the

$$A = A_0 \sin 2\pi f\left(t - \frac{x_1}{c}\right) \tag{10}$$

$$B = A_0 \sin 2\pi f\left(t - \frac{x_2}{c}\right) \tag{11}$$

time, c the velocity of light; x_1 and x_2 are the distances of P from S_1 and S_2, and A_0 is the amplitude. This amplitude is assumed to be the same for each wave since the slits are close together, and x_1 and x_2 are thus nearly the same. These equations are the same as Eqs. (1) and (2), with $\varphi_1 = x_1/c$ and $\varphi_2 = x_2/c$. Accordingly, the square of the amplitude or the

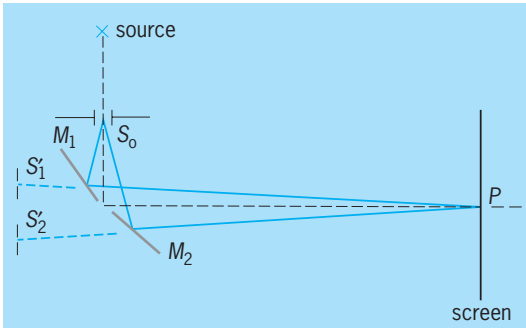

Fig. 2. Fresnel's double-mirror interference.

intensity at P can be written as Eq. (12).

$$I = 4A_0^2 \cos^2 \frac{2\pi f}{c}(x_1 - x_2) \tag{12}$$

In general, l is very much larger than y so that Eq. (12) can be simplified to Eq. (13).

$$I = 4A_0^2 \cos^2 \pi \left(\frac{yd}{l\lambda}\right) \tag{13}$$

Equation (13) is a maximum when Eq. (14) holds and a minimum when Eq. (15) holds, where n is an integer.

$$y = n\lambda \frac{l}{d} \tag{14}$$

$$y = (n + {}^1/_2)\lambda \frac{l}{d} \tag{15}$$

Accordingly, the screen is covered with a series of light and dark bands called interference fringes. If the source behind slit S_0 is white light and thus has wavelengths varying perhaps from 400 to 700 nm, the fringes are visible only where $x_1 - x_2$ is a few wavelengths, that is, where n is small. At large values of n, the position of the nth fringe for red light will be very different from the position for blue light, and the fringes will blend together and be washed out. With monochromatic light, the fringes will be visible out of values of n which are determined by the diffraction pattern of the slits. For an explanation of this *see* DIFFRACTION.

The energy carried by a wave is measured by the intensity, which is equal to the square of the amplitude. In the preceding example of the superposition of two waves, the intensity of the individual waves in Eqs. (1) and (2) is A^2 and B^2, respectively. When the phase shift between them is zero, the intensity of the resulting wave is given by Eq. (16).

$$(A + B)^2 = A^2 + 2AB + B^2 \tag{16}$$

This would seem to be a violation of the first law of thermodynamics, since this is greater than the sum of the individual intensities. In any specific experiment, however, it turns out that the energy from the source is merely redistributed in space. The excess energy which appears where the interference is constructive will disappear in those places where the energy is destructive. This is illustrated by the fringe pattern in the Young two-slit experiment. The energy on the screen from each slit alone is given by Eq. (17), where A_0^2 is the intensity of the light from

$$E_1 = \int_0^\infty A_0^2 \, dy \tag{17}$$

each slit as given by Eq. (10). The intensity from the two slits without interference would be twice this value. The intensity with interference is given by Eq. (18). The comparison between $2E_1$ and E_3 need

$$E_3 = \int_0^\infty 4A^2 \cos^2 \left[2\pi \left(\frac{yd}{l\lambda}\right)\right] dy \tag{18}$$

be made only over a range corresponding to one full cycle of fringes. This means that the argument of the cosine in Eq. (18) need be taken only from zero to π. This corresponds to a section of screen going from the center to a distance $y = l\lambda/2d$. From the two slits individually, the energy in this section of screen can be written as Eq. (19).

$$2E_1 = 2\int_0^{l\lambda/2d} A_0^2 \, dy = \frac{A_0^2 l\lambda}{d} \tag{19}$$

With interference, the energy is given by Eq. (20). Equation (20) can be written as Eq. (21).

$$E_3 = \int_0^{l\lambda/2d} 4A_0^2 \cos^2 \left[\left(\frac{yd}{l\lambda}\right)\right] dy \tag{20}$$

$$E_3 = \frac{l\lambda}{2\pi d}\int_0^\pi 4A_0^2 \cos^2 \varphi \, d\varphi = \frac{A_0^2 l\lambda}{d} \tag{21}$$

Thus, the total energy falling on the screen is not changed by the presence of interference. The energy density at a particular point is, however, drastically changed. This fact is most important for those waves of the electromagnetic spectrum which can be generated by vacuum-tube oscillators. The sources of radiation or antennas can be made to emit coherent waves which will undergo interference. This makes possible a redistribution of the radiated energy. Quite narrow beams of radiation can be produced by the proper phasing of a linear antenna array. *See* ANTENNA (ELECTROMAGNETISM).

The double-slit experiment also provides a good illustration of Niels Bohr's principle of complementarity. For detailed information on this *see* QUANTUM MECHANICS

Fresnel double mirror. Another way of splitting the light from the source is the Fresnel double mirror (**Fig. 2**). Light from the slit S_0 falls on two mirrors M_1 and M_2 which are inclined to each other at an angle of the order of a degree. On a screen where the illumination from the two mirrors overlaps, there will appear a set of interference fringes. These are the same as the fringes produced in the two-slit experiment, since the light on the screen comes from the images of the slits S_1' and S_2' formed by the two mirrors, and these two images are the equivalent of two slits.

Fresnel biprism. A third way of splitting the source is the Fresnel biprism. A sketch of a cross section of this device is shown in **Fig. 3**. The light from the slit at S_0 is transmitted through the two halves of the prism to the screen. The beam from each half will strike the

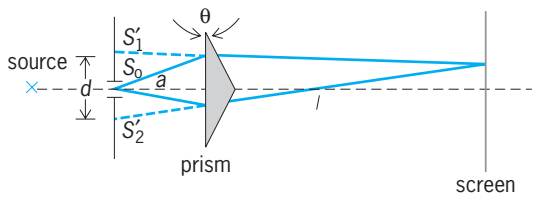

Fig. 3. Fresnel biprism interference.

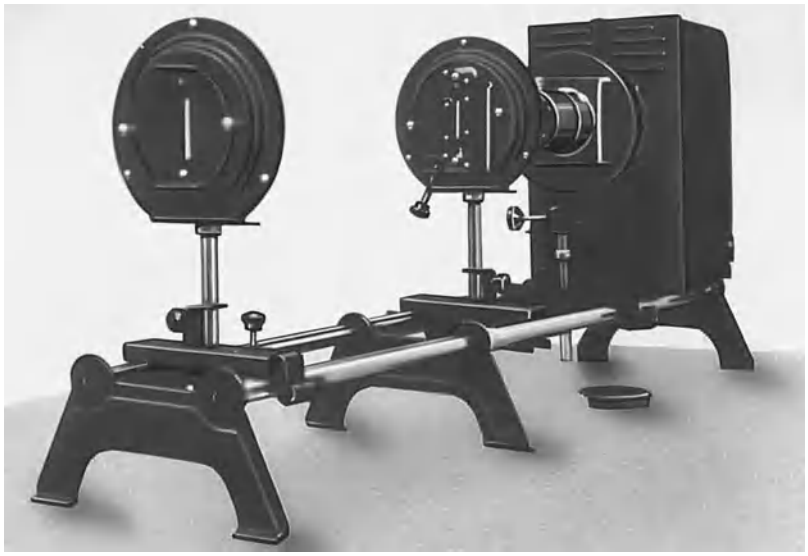

Fig. 4. Equipment for demonstrating Fresnel biprism interference.

screen at a different angle and will appear to come from a source which is slightly displaced from the original slit. These two virtual slits are shown in the sketch at S'_1 and S'_2. Their separation will depend on the distance of the prism from the slit S_0 and on the angle θ and index of refraction of the prism material. In Fig. 3, a is the distance of the slit from the biprism, and l the distance of the biprism from the screen. The distance of the two virtual slits from the screen is thus $a + l$. The separation of the two virtual slits is given by Eq. (22), where μ is the refractive index

$$d = 2a(\mu - 1)\theta \tag{22}$$

of the prism material. This can be put in Eq. (14) for the two-slit interference pattern to give Eq. (23) for the position of a bright fringe.

$$y = n\lambda \frac{a + l}{2a(\mu - 1)\theta} \tag{23}$$

A photograph of the experimental equipment for demonstrating interference with the Fresnel biprism is shown in **Fig. 4**. A typical fringe pattern is shown in **Fig. 5**. This pattern was obtained with a mercury-arc source, which has several strong spectrum lines, accounting in part for the intensity variation in the pattern. The pattern is also modified by diffraction at the apex of the prism.

Billet split lens. The source can also be split with the Billet split lens (**Fig. 6**). Here a simple lens is sawed into two parts which are slightly separated.

Lloyd's mirror. An important technique of splitting the source is with Lloyd's mirror (**Fig. 7**). The slit S_1 and its virtual image S'_2 constitute the double source. Part of the light falls directly on the screen, and part is reflected at grazing incidence from a plane mirror. This experiment differs from the previously discussed experiments in that the two beams are no longer identical. If the screen is moved to a point where it is nearly in contact with the mirror, the fringe of zero path difference will lie on the intersection of the mirror plane with the screen. This fringe turns out to be dark rather than light, as in the case of the previous interference experiments. The only explanation for this result is that light experiences a 180° phase shift on reflection from a material of higher refractive index than its surrounding medium. The equation for maximum and minimum light intensity at the screen must thus be interchanged for Lloyd's mirror fringes.

Amplitude splitting. The interference experiments discussed have all been done by splitting the wavefront of the light coming from the source. The energy from the source can also be split in amplitude. With such amplitude-splitting techniques, the light from the source falls on a surface which is partially reflecting. Part of the light is transmitted, part is reflected, and after further manipulation these parts are recombined to give the interference. In one type of experiment, the light transmitted through the surface is reflected from a second surface back through the partially reflecting surface, where it combines with the wave reflected from the first surface (**Fig. 8**). Here the arrows represent the normal to the wavefront of the light passing through surface S_1 to surface S_2. The wave is incident at A and C. The section

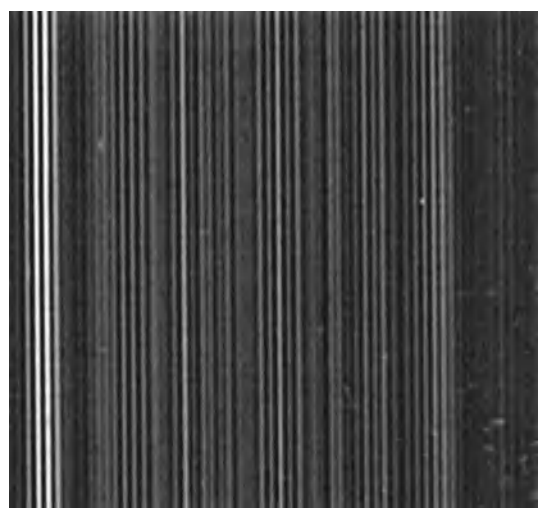

Fig. 5. Interference fringes formed with Fresnel biprism and mercury-arc light source.

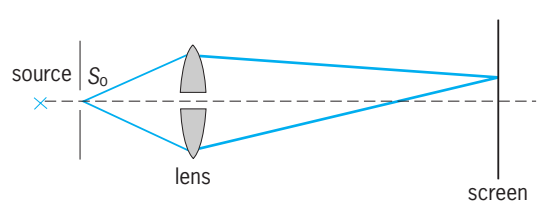

Fig. 6. Billet split-lens interference.

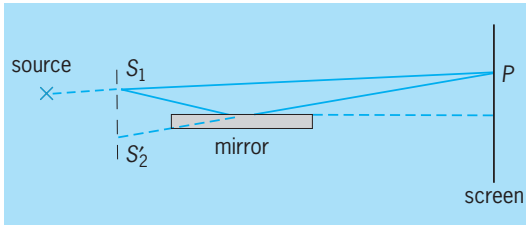

Fig. 7. Lloyd's mirror interference.

at A is partially transmitted to B, where it is again partially reflected to C. The wave leaving C now consists of two parts, one of which has traveled a longer distance than the other. These two waves will interfere. Let AD be the perpendicular from the ray at A to the ray going to C. The path difference will be given by Eq. (24), where μ is the refractive index of the

$$\Delta = 2\mu(AB) - (CD) \tag{24}$$

medium between the surfaces S_1 and S_2, and AB and CD are defined in Eqs. (25) and (26).

$$(AB) = \frac{d}{\cos\ r} \tag{25}$$

$$(CD) = 2(AB)\sin r \cos i \tag{26}$$

From Snell's law, Eq. (27) is obtained and thus Eq. (28) holds.

$$\sin i = \mu \sin r \tag{27}$$

$$\Delta = \frac{2\mu d}{\cos r} - \frac{2\mu d}{\cos r}\sin^2 r = 2\mu d \cos r \tag{28}$$

The difference in terms of wavelength and the phase difference are, respectively, given by Eqs. (29) and (30).

$$\Delta' = \frac{2\mu d \cos r}{\lambda} \tag{29}$$

$$\Delta\varphi = \frac{4\pi\mu d \cos r}{\lambda} + \pi \tag{30}$$

The phase difference of π radians is added because of the phase shift experienced by the light reflected at S_1. The experimental proof of this 180° phase shift

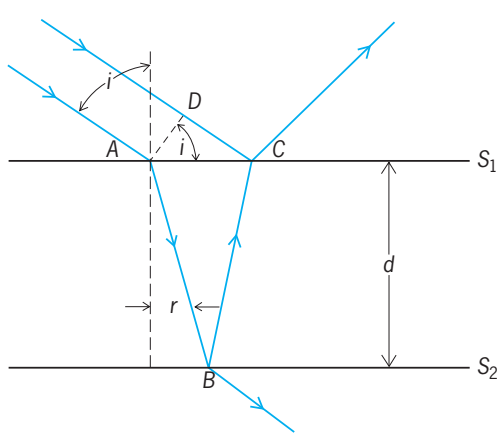

Fig. 8. Dielectric-plate-reflection interference.

was shown in the description of interference with Lloyd's mirror. If the plate of material has a lower index than the medium in which it is immersed, the π radians must still be added in Eq. (30), since now the beam reflected at S_2 will experience this extra phase shift. The purely pragmatic necessity of such an additional phase shift can be seen by considering the intensity of the reflected light when the surfaces S_1 and S_2 almost coincide. Without the extra phase shift, the two reflected beams would be in phase and the reflection would be strong. This is certainly not proper for a film of vanishing thickness. Constructive interference will take place at wavelengths for which $\Delta\varphi = 2m\pi$, where m is an integer. If the surfaces S_1 and S_2 are parallel, the fringes will be located optically at infinity. If they are not parallel, d will be a function of position along the surfaces and the fringes will be located near the surface. The intensity of the fringes will depend on the value of the partial reflectivity of the surfaces.

Testing of optical surfaces. Observation of fringes of this type can be used to determine the contour of a surface. The surface to be tested is put close to an optically flat plate. Monochromatic light is reflected from the two surfaces and examined as in Fig. 8. One of the first experiments with fringes of this type was performed by Isaac Newton. A convex lens is pressed against a glass plate and illuminated with monochromatic light. A series of circular interference fringes known as Newton's rings appear around the point of contact. From the separation between the fringes, it is possible to determine the radius of curvature of the lens.

Thin films. Interference fringes of this two-surface type are responsible for the colors which appear in oil films floating on water. Here the two surfaces are the oil-air interface and the oil-water interface. The films are close to a visible light wavelength in thickness. If the thickness is such that, in a particular direction, destructive interference occurs for green light, red and blue will still be reflected and the film will have a strong purple appearance. This same general phenomenon is responsible for the colors of beetles' wings.

Channeled spectrum. Amplitude splitting shows clearly another condition that must be satisfied for interference to take place. The beams from the source must not only come from identical points, but they must also originate from these points at nearly the same time. The light which is reflected from C in Fig. 8 originates from the source later than the light which makes a double traversal between S_1 and S_2. If the surfaces are too far apart, the spectral regions of constructive and destructive interference become so close together that they cannot be resolved. In the case of interference by wavefront splitting, the light from different parts of a source could only be considered coherent if examined over a sufficiently short time interval. In the case of amplitude splitting, the interference when surfaces are widely separated can only be seen if examined over a sufficiently narrow frequency interval. If the two surfaces are illuminated with white light

and the eye is used as the analyzer, interference cannot be seen when the separation is more than a few wavelengths. The interval between successive wavelengths of constructive interference becomes so small that each spectral region to which the eye is sensitive is illuminated, and no color is seen. In this case, the interference can again be seen by examining the reflected light with a spectroscope. The spectrum will be crossed with a set of dark fringes at those wavelengths for which there is destructive interference. This is called a channeled spectrum. For large separations of the surfaces, the separation between the wavelengths of destructive interference becomes smaller than the resolution of the spectrometer, and the fringes are no longer visible.

Fresnel coefficient. The amplitude of the light reflected at normal incidence from a dielectric surface is given by the Fresnel coefficient, Eq. (31), where

$$A = A_0 \frac{n_1 - n_2}{n_1 + n_2} \tag{31}$$

A_0 is the amplitude of the incident wave and n_1 and n_2 are the refractive indices of the materials in the order in which they are encountered by the light. In the simple case of a dielectric sheet, the intensity of the light reflected normally will be given by Eq. (32), where B is the amplitude of the wave which

$$C^2 = A^2 + B^2 + 2AB \cos\varphi \tag{32}$$

has passed through the sheet and is reflected from the second surface and back through the sheet to join A. The value of B is given by Eq. (33), where the

$$B = \frac{n_2 - n_3}{n_2 + n_3} \tag{33}$$

approximation is made that the intensity of the light is unchanged by passing through the first surface and where n_3 is the index of the material at the boundary of the far side of the sheet.

Nonreflecting film. An interesting application of Eq. (32) is the nonreflecting film. A single dielectric layer is evaporated onto a glass surface to reduce the reflectivity of the surface to the smallest possible value. From Eq. (32) it is clear that this takes place when $\cos\varphi = -1$. If the surface is used in an instrument with a broad spectral range, such as a visual device, the film thickness should be adjusted to put the interference minimum in the first order and in the middle of the desired spectral range. For the eye, this wavelength is approximately in the yellow so that such films reflect in the red and blue and appear purple. The index of the film should be chosen to make $C^2 = 0$. At this point Eqs. (34)–(36) hold.

$$(A - B)^2 = 0 \tag{34}$$

$$\frac{n_1 - n_2}{n_1 + n_2} = \frac{n_2 - n_3}{n_2 + n_3} \tag{35}$$

$$n_1 n_2 - n_2{}^2 + n_1 n_3 - n_2 n_3 = n_1 n_2 - n_1 n_3 + n_2{}^2 - n_2 n_3 \tag{36}$$

Equation (36) can be reduced to Eq. (37). In the

$$n_2 = \sqrt{n_1 n_3} \tag{37}$$

case of a glass surface in air, $n_1 = 1$ and $n_3 \cong 1.5$. Magnesium fluoride is a substance which is frequently used as a nonreflective coating, since it is hard and approximately satisfies the relationship of Eq. (37). The purpose of reducing the reflection from an optical element is to increase its transmission, since the energy which is not reflected is transmitted. In the case of a single element, this increase is not particularly important. Some optical instruments may have 15–20 air-glass surfaces, however, and the coating of these surfaces gives a tremendous increase in transmission.

Haidinger fringes. When the second surface in two-surface interference is partially reflecting, interference can also be observed in the wave transmitted through both surfaces. The interference fringes will be complementary to those appearing in reflection. Their location will depend on the parallelism of the surfaces. For plane parallel surfaces, the fringes will appear at infinity and will be concentric rings. These were first observed by W. K. Haidinger and are called Haidinger fringes.

Multiple-beam interference. If the surfaces S_1 and S_2 are strongly reflecting, it is necessary to consider multiple reflections between them. For air-glass surfaces, this does not apply since the reflectivity is of the order of 4%, and the twice-reflected beam is much reduced in intensity.

In **Fig. 9** the situation in which the surfaces S_1 and S_2 have reflectivities r_1 and r_2 is shown. The space between the surfaces has an index n_2 and thickness d. An incident light beam of amplitude A is partially reflected at the first surface. The transmitted component is reflected at S_2 and is reflected back to S_1 where a second splitting takes place. This is repeated. Each successive component of the waves leaving S_1 is retarded with respect to the next. The amount of each retardation is given by Eq. (38).

$$\varphi = \frac{4\pi nd}{\lambda} \cos\theta \tag{38}$$

Equation (7) was derived for the superposition of two waves. It is possible to derive a similar expression for the superposition of many waves. From

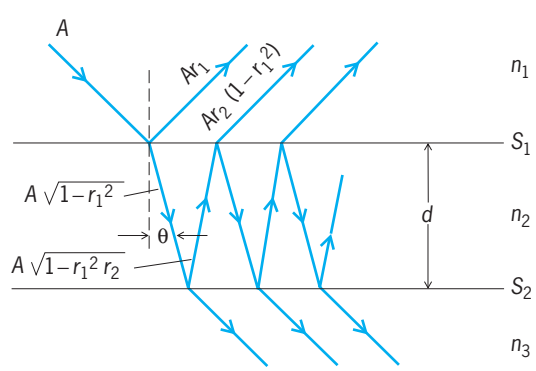

Fig. 9. Multiple reflection of wave between two surfaces.

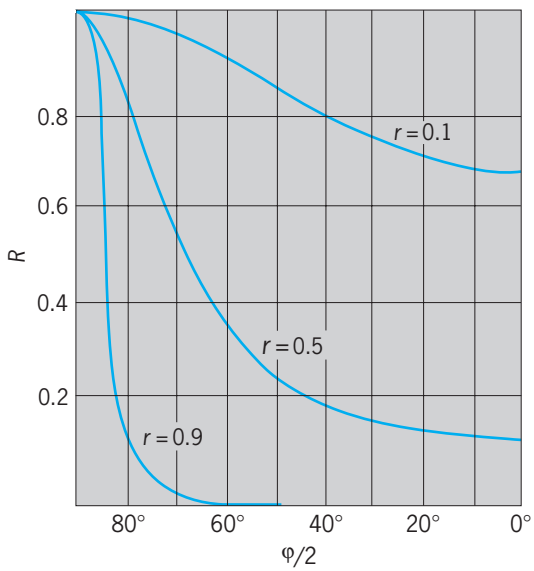

Fig. 10. The shape of multiple-beam fringes for different values of surface reflectivity.

Fig. 9, the different waves at a plane somewhere above S_1 can be represented by the following expressions:

$$\begin{aligned}
&\text{Incoming wave} &&= A\ \sin\omega t\\
&\text{First reflected wave} &&= Ar_1\ \sin\omega t\\
&\text{Second reflected wave} &&= A(1-r_1{}^2)r_2\ \sin(\omega t+\varphi)\\
&\text{Third reflected wave} &&= -A(1-r_1{}^2)^2 r_1 r_2{}^2\\
& && \quad\times\sin(\omega t+2\varphi)
\end{aligned}$$

By inspection of these terms, one can write down the complete series. As in Eq. (3), the sine terms can be broken down and coefficients collected. A simpler method is to multiply each term by $i=\sqrt{-1}$ and add a cosine term with the same coefficient and argument. The individual terms then are all of the form of expression (39), where m is an integer.

$$Be^{-i\omega t}e^{-im\varphi} \tag{39}$$

The individual terms of expression (39) can be easily summed. For the reflected wave one obtains Eq. (40). Again, as in the two-beam case, the minimum in the reflectivity R is obtained when $\varphi = N\pi$, where N is an odd integer and $r_1 = r_2$. The fringe shape, however, can be quite different from the earlier case, depending on the values of the reflectives r_1 and r_2. The greater these values, the sharper become the fringes.

$$R=\frac{r_1+r_2e^{-i\varphi}}{1+r_1r_2e^{-i\varphi}} \tag{40}$$

It was shown earlier how two-beam interference could be used to measure the contour of a surface. In this technique, a flat glass test plate was placed over the surface to be examined and monochromatic interference fringes were formed between the test surface and the surface of the plate. These two-beam fringes have intensities which vary as the cosine squared of the path difference. It is very difficult with such fringe to detect variations in fringe straightness or, in other terms, variations of surface planarity that are smaller than 1/20 wavelength. If the surface to be examined is coated with silver and the test surface is also coated with a partially transmitting metallic coat, the reflectivity increases to a point where many beams are involved in the formation of the interference fringes. The shape of the fringes is given by Eq. (40). The shape of fringes for different values of r is shown in **Fig. 10**. With high-reflectivity fringes, the sensitivity to a departure from planarity is increased far beyond 1/20 wavelength.

It is thus possible with partially silvered surfaces to get a much better picture of small irregularities than with uncoated surfaces. The increase in sensitivity is such that steps in cleaved mica as small as 1 nm in height can be seen by examining the monochromatic interference fringes produced between a silvered mica surface and a partially silvered glass flat. *See* INTERFERENCE FILTERS; INTERFEROMETRY.

Bruce H. Billings

Bibliography. M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976; J. R. Meyer-Arendt, *Introduction to Classical and Modern Optics*, 4th ed., 1995; F. L. Pedrotti and L. S. Pedrotti, *Introduction to Optics*, 3d ed., 2006.

Interferometry

The design and use of optical interferometers. Optical interferometers based on both two-beam interference and multiple-beam interference of light are extremely powerful tools for metrology and spectroscopy. A wide variety of measurements can be performed, ranging from determining the shape of a surface to an accuracy of less than a millionth of an inch (25 nanometers) to determining the separation, by millions of miles, of binary stars. In spectroscopy, interferometry can be used to determine the hyperfine structure of spectrum lines. By using lasers in classical interferometers as well as holographic interferometers and speckle interferometers, it is possible to perform deformation, vibration, and contour measurements of diffuse objects that could not previously be performed.

Basic classes of interferometers. There are two basic classes of interferometers: division of wavefront and division of amplitude. **Figure 1** shows two arrangements for obtaining division of wavefront. For the Young's double pinhole interferometer (Fig. 1*a*), the light from a point source illuminates two pinholes. The light diffracted by these pinholes gives the interference of two point sources. For the Lloyd's mirror experiment (Fig. 1*b*), a mirror is used to provide a second image S_2 of the point source S_1, and in the region of overlap of the two beams the interference of two spherical beams can be observed. There are many other ways of obtaining division of wavefront; however, in each case the light leaving the source is spatially split, and then by use of diffraction, mirrors, prisms, or lenses the two spatially separated beams are superimposed.

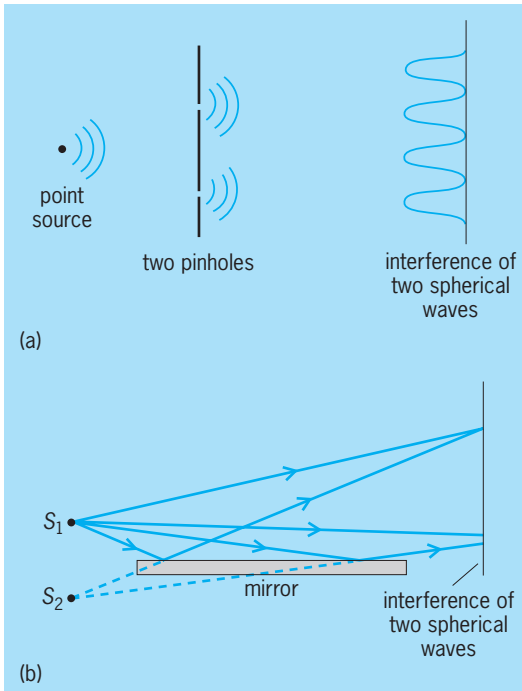

Fig. 1. Interference produced by division of wavefront. (*a*) Young's two-pinhole interferometer. (*b*) Lloyd's mirror.

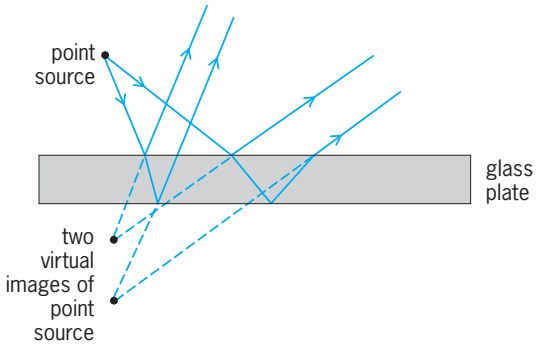

Fig. 2. Division of amplitude.

Figure 2 shows one technique for obtaining division of amplitude. For division-of-amplitude interferometers a beam splitter of some type is used to pick off a portion of the amplitude of the radiation which is then combined with a second portion of the amplitude. The visibility of the resulting interference fringes is a maximum when the amplitudes of the two interfering beams are equal. *See* INTERFERENCE OF WAVES.

Michelson interferometer. The Michelson interferometer (**Fig. 3**) is based on division of amplitude. Light from an extended source S is incident on a partially reflecting plate (beam splitter) P_1. The light transmitted through P_1 reflects off mirror M_1 back to plate P_1. The light which is reflected proceeds to M_2 which reflects it back to P_1. At P_1, the two waves are again partially reflected and partially transmitted, and a portion of each wave proceeds to the receiver R, which may be a screen, a photocell, or a human eye. Depending on the difference between the distances from the beam splitter to the mirrors M_1 and M_2, the two beams will interfere constructively or destructively. Plate P_2 compensates for the thickness of P_1. Often when a quasimonochromatic light source is used with the interferometer, compensating plate P_2 is omitted.

The function of the beam splitter is to superimpose (image) one mirror onto the other. When the mirrors' images are completely parallel, the interference fringes appear circular. If the mirrors are slightly inclined about a vertical axis, vertical fringes are formed across the field of view. These fringes can be formed in white light if the path difference in part of the field of view is made zero. Just as in other interference experiments, only a few fringes will appear in white light, because the difference in path will be different for wavelengths of different colors. Accordingly, the fringes will appear colored close to zero path difference, and will disappear at larger path differences where the fringe maxima and minima for the different wavelengths overlap. If light reflected off the beam splitter experiences a one-half-cycle relative phase shift, the fringe of zero path difference is black, and can be easily distinguished from the neighboring fringes. This makes use of the instrument relatively easy.

The Michelson interferometer can be used as a spectroscope. Consider first the case of two close spectrum lines as a light source for the instrument. As the mirror M_1 is shifted, fringes from each spectral line will cross the field. At certain path differences between M_1 and M_2, the fringes for the two spectral lines will be out of phase and will essentially disappear; at other points they will be in phase and will be reinforced. By measuring the distance between successive maxima in fringe contrast, it is possible to determine the wavelength difference between the lines.

This is a simple illustration of a very broad use for any two-beam interferometer. As the path length L is changed, the variation in intensity $I(L)$ of the light coming from an interferometer gives information on the basis of which the spectrum of the input light can be derived. The equation for the intensity of the emergent energy can be written as

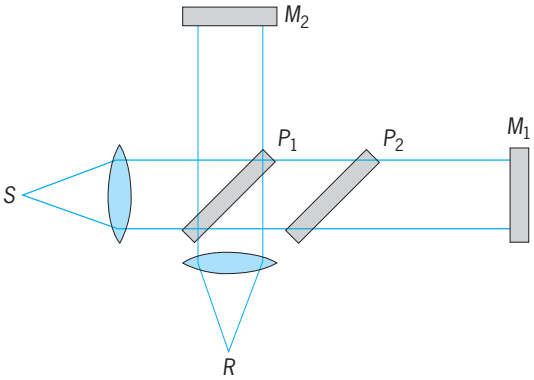

Fig. 3. Michelson interferometer.

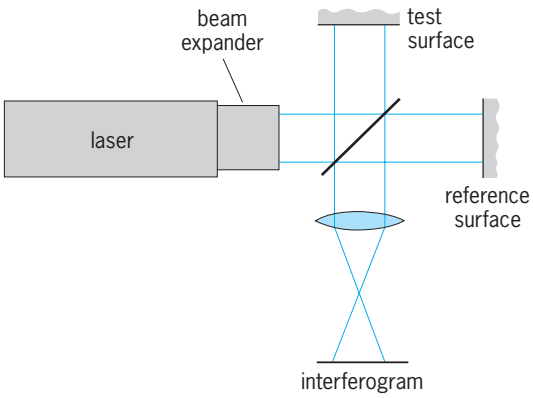

Fig. 4. Twyman-Green interferometer for testing flat surfaces.

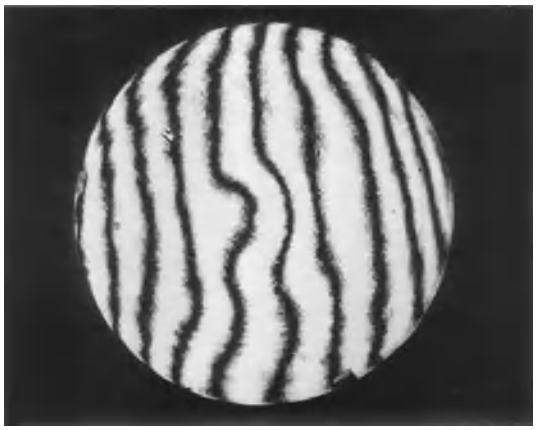

Fig. 5. Interferogram obtained with the use of a Twyman-Green interferometer to test a flat surface.

Eq. (1), where β is a constant, and $I(\lambda)$ is the inten-

$$I(L) = \int_0^\infty I(\lambda) \cos^2\left(\frac{\beta L}{\lambda}\right) d\lambda \qquad (1)$$

sity of the incident light at different wavelengths λ. This equation applies when the mirror M_1 is moved linearly with time from the position where the path difference with M_2 is zero, to a position which depends on the longest wavelength in the spectrum to be examined. From Eq. (1), it is possible mathematically to recover the spectrum $I(\lambda)$. In certain situations, such as in the infrared beyond the wavelength region of 1.5 micrometers, this technique offers a large advantage over conventional spectroscopy in that its utilization of light is extremely efficient. *See* INFRARED SPECTROSCOPY.

Twyman-Green interferometer. If the Michelson interferometer is used with a point source instead of an extended source, it is called a Twyman-Green interferometer. The use of the laser as the light source for the Twyman-Green interferometer has made it an extremely useful instrument for testing optical components. The great advantage of a laser source is that it makes it possible to obtain bright, good-contrast, interference fringes even if the path lengths for the two arms of the interferometer are quite different. *See* LASER.

Figure 4 shows a Twyman-Green interferometer for testing a flat mirror. The laser beam is expanded to match the size of the sample being tested. Part of the laser light is transmitted to the reference surface, and part is reflected by the beam splitter to the flat surface being tested. Both beams are reflected back to the beam splitter, where they are combined to form interference fringes. An imaging lens projects the surface under test onto the observation plane.

Fringes (**Fig. 5**) show defects in the surface being tested. If the surface is perfectly flat, then straight, equally spaced fringes are obtained. Departure from the straight, equally spaced condition shows directly how the surface differs from being perfectly flat. For a given fringe, the difference in optical path between light going from laser to reference surface to observation plane and the light going from laser to test surface to observation plane is a constant. (The optical path is equal to the product of the geometrical path times the refractive index.) Between adjacent fringes (Fig. 5), the optical path difference changes by one wavelength, which for a helium-neon laser corresponds to 633 nm. The number of straight, equally spaced fringes and their orientation depend upon the tip-tilt of the reference mirror. That is, by tipping or tilting the reference mirror the difference in optical path can be made to vary linearly with distance across the laser beam.

Deviations from flatness of the test mirror also cause optical path variations. A height change of half a wavelength will cause an optical path change of one wavelength and a deviation from fringe straightness of one fringe. Thus, the fringes give surface height information, just as a topographical map gives height or contour information.

The existence of the essentially straight fringes provides a means of measuring surface contours relative to a tilted plane. This tilt is generally introduced to indicate the sign of the surface error, that is, whether the errors correspond to a hill or a valley. One way to get this sign information is to push in on the piece being tested when it is in the interferometer. If the fringes move toward the right when the test piece is pushed toward the beam splitter, then fringe deviations from straightness toward the right correspond to high points (hills) on the test surface and deviations to the left correspond to low points (valleys).

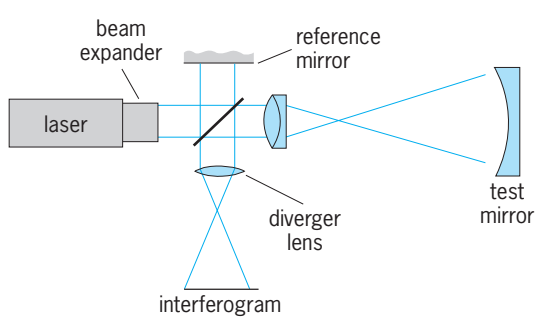

Fig. 6. Twyman-Green interferometer for testing spherical mirrors or lenses.

The basic Twyman-Green interferometer (Fig. 4) can be modified (**Fig. 6**) to test concave-spherical mirrors. In the interferometer, the center of curvature of the surface under test is placed at the focus of a high-quality diverger lens so that the wavefront is reflected back onto itself. After this retroflected wavefront passes through the diverger lens, it will be essentially a plane wave, which, when it interferes with the plane reference wave, will give interference fringes similar to those shown in Fig. 5 for testing flat surfaces. In this case it indicates how the concave-spherical mirror differs from the desired shape. Likewise, a convex-spherical mirror can be tested. Also, if a high-quality spherical mirror is used, the high-quality diverger lens can be replaced with the lens to be tested.

Fizeau interferometer. One of the most commonly used interferometers in optical metrology is the Fizeau interferometer, which can be thought of as a folded Twyman-Green interoferometer. In the Fizeau, the two surfaces being compared, which can be flat, spherical, or aspherical, are placed in close contact. The light reflected off these two surfaces produces interference fringes. For each fringe, the separation between the two surfaces is a constant. If the two surfaces match, straight, equally spaced fringes result. Surface height variations between the two surfaces cause the fringes to deviate from straightness or equal separations, where one fringe deviation from straightness corresponds to a variation in separation between the two surfaces by an amount equal to one-half of the wavelength of the light source used in the interferometer. The wavelength of a helium source, which is often used in a Fizeau interferometer, is 587.56 nm; hence one fringe corresponds to a height variation of approximately 0.3 μm.

Mach-Zehnder interferometer. The Mach-Zehnder interferometer (**Fig. 7**) is a variation of the Michelson interferometer and, like the Michelson interferometer, depends on amplitude splitting of the wavefront. Light enters the instrument and is reflected and transmitted by the semitransparent mirror M_1.

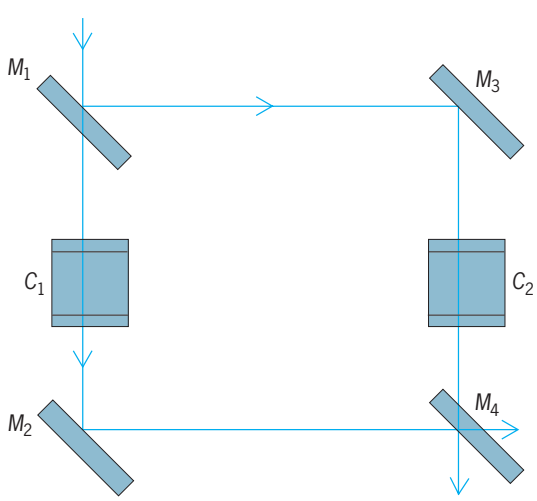

Fig. 7. Mach-Zehnder interferometer.

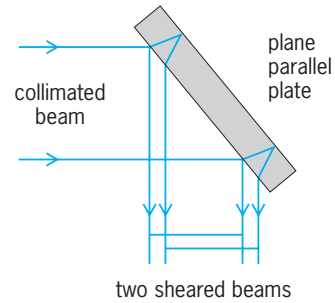

Fig. 8. Lateral shear interferometer.

The reflected portion proceeds to M_3, where it is reflected through the cell C_2 to the semitransparent mirror M_4. Here it combines with the light transmitted by M_1 to produce interference. The light transmitted by M_1 passes through a cell C_1, which is similar to C_2 and is used to compensate for the windows of C_1.

The major application of this instrument is in studying airflow around models of aircraft, missiles, or projectiles. The object and associated airstream are placed in one arm of the interferometer. Because the air pressure varies as it flows over the model, the index of refraction varies, and thus the effective path length of the light in this beam is a function of position. When the variation is an odd number of half-waves, the light will interfere destructively and a dark fringe will appear in the field of view. From a photograph of the fringes, the flow pattern can be mathematically derived.

A major difference between the Mach-Zehnder and the Michelson interferometer is that in the Mach-Zehnder the light goes through each path in the instrument only once, whereas in the Michelson the light traverses each path twice. This double traversal makes the Michelson interferometer extremely difficult to use in applications where spatial location of index variations is desired. The incoming and outgoing beams tend to travel over slightly different paths, and this lowers the resolution because of the index gradient across the field.

Shearing interferometers. In a lateral-shear interferometer, an example of which is shown in **Fig. 8**, a wavefront is interfered with a shifted version of itself. A bright fringe is obtained at the points where the slope of the wavefront times the shift between the two wavefronts is equal to an integer number of wavelengths. That is, for a given fringe the slope or derivative of the wavefront is a constant. For this reason a lateral-shear interferometer is often called a differential interferometer.

Another type of shearing interferometer is a radial-shear interferometer. Here, a wavefront is interfered with an expanded version of itself. This interferometer is sensitive to radial slopes.

The advantages of shearing interferometers are that they are relatively simple and inexpensive, and since the reference wavefront is self-generated, an external wavefront is not needed. Since an external reference beam is not required, the source

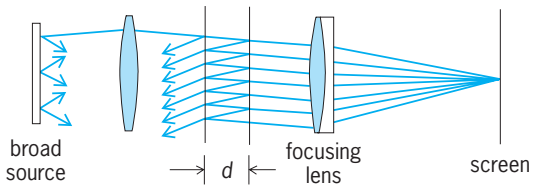

Fig. 9. Fabry-Perot interferometer.

requirements are reduced from those of an interferometer such as a Twyman-Green. For this reason, shearing interferometers, in particular lateral-shear interferometers, are finding much use in applications such as adaptive optics systems for correction of atmospheric turbulence where the light source has to be a star, or planet, or perhaps just reflected sunlight. *See* ADAPTIVE OPTICS.

Michelson stellar interferometer. A Michelson stellar interferometer can be used to measure the diameter of stars which are as small as 0.01 second of arc. This task is impossible with a ground-based optical telescope since the atmosphere limits the resolution of the largest telescope to not much better than 1 second of arc.

The Michelson stellar interferometer is a simple adaptation of Young's two-slit experiment. In its first form, two slits were placed over the aperture of a telescope. If the object being observed were a true point source, the image would be crossed with a set of interference bands. A second point source separated by a small angle from the first would produce a second set of fringes. At certain values of this angle, the bright fringes in one set will coincide with the dark fringes in the second set. The smallest angle α at which the coincidence occurs will be that angle subtended at the slits by the separation of the peak of the central bright fringe from the nearest dark fringe. This angle is given by Eq. (2), where d is the

$$\frac{\lambda}{2d} = \alpha \tag{2}$$

separation of the slits, λ the dominant wavelength of the two sources, and α their angular separation. The measurement of the separation of the sources is performed by adjusting the separation d between the slits until the fringes vanish.

Consider now a single source in the shape of a slit of finite width. If the slit subtends an angle at the telescope aperture which is larger than α, the interference fringes will be reduced in contrast. For various line elements at one side of the slit, there will be elements of angle α which will cancel the fringes from the first element. By induction, it is clear that for a separation d' such that the slit source subtends an angle as given by Eq. (3) the fringes from a single

$$\alpha' = \frac{\lambda}{d'} \tag{3}$$

slit will vanish completely. For additional information on the Michelson stellar interferometer *see* DIFFRACTION

Fabry-Perot interferometer. All the interferometers discussed above are two-beam interferometers. The Fabry-Perot interferometer (**Fig. 9**) is a multiple-beam interferometer since the two glass plates are partially silvered on the inner surfaces, and the incoming wave is multiply reflected between the two surfaces. The position of the fringe maxima is the same for multiple-beam interference as two-beam interference; however, as the reflectivity of the two surfaces increases and the number of interfering beams increases, the fringes become sharper.

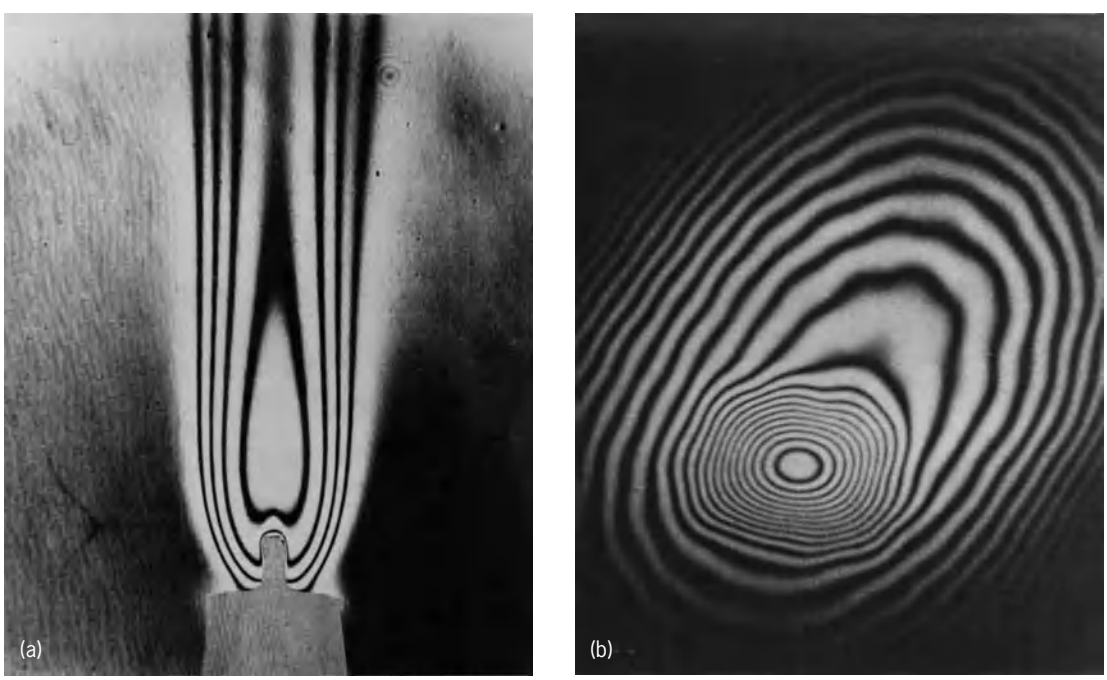

Fig. 10. Double-exposure holographic interferograms. (*a*) Interferogram of candle flame. (*b*) Interferogram of debanded region of honeycomb construction panel. (*From C. M. Vest, Holographic Interferometry, John Wiley and Sons, 1979*)

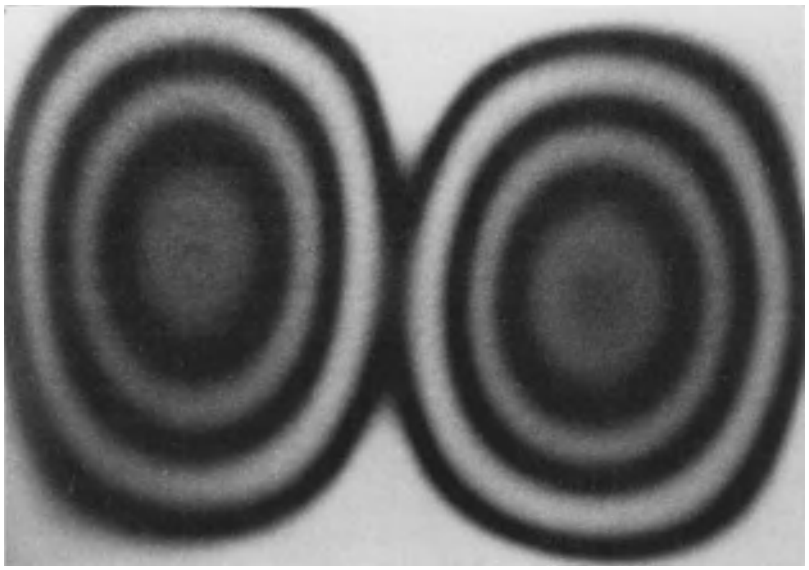

Fig. 11. Photograph of time-average holographic interferogram. (*From C. M. Vest, Holographic Interferometry, John Wiley and Sons, 1979*)

A quantity of particular interest in a Fabry-Perot is the ratio of the separation of adjacent maxima to the half-width of the fringes. It can be shown that this ratio, known as the finesse, is given by Eq. (4), where

$$\mathscr{F} = \frac{\pi\sqrt{R}}{1-R} \tag{4}$$

R is the reflectivity of the silvered surfaces.

The multiple-beam Fabry-Perot interferometer is of considerable importance in modern optics for spectroscopy. All the light rays incident on the Fabry-Perot at a given angle will result in a single circular fringe of uniform irradiance. With a broad diffuse source, the interference fringes will be narrow concentric rings, corresponding to the multiple-beam transmission pattern. The position of the fringes depends upon the wavelength. That is, each wavelength gives a separate fringe pattern. The minimum resolvable wavelength difference is determined by the ability to resolve close fringes. The ratio of the wavelength λ to the least resolvable wavelength difference $\Delta\lambda$ is known as the chromatic resolving power $\mathscr{R}$. At nearly normal incidence it is given by Eq. (5), where n is the refractive index between the

$$\mathscr{R} = \frac{\lambda}{(\Delta\lambda)_{\min}} = \mathscr{F}\frac{2nd}{\lambda} \tag{5}$$

two mirrors separated a distance d. For a wavelength of 500 nm, $nd = 10$ mm, and $R = 90\%$, the resolving power is well over 10^6. *See* REFLECTION OF ELECTROMAGNETIC RADIATION; RESOLVING POWER (OPTICS).

When Fabry-Perot interferometers are used with lasers, they are generally used in the central spot scanning mode. The interferometer is illuminated with a collimated laser beam, and all the light transmitted through the Fabry-Perot is focused onto a detector, whose output is displayed on an oscilloscope. Often one of the mirrors is on a piezoelectric mirror mount. As the voltage to the piezoelectric crystal is varied, the mirror separation is varied. The light output as a function of mirror separation gives the spectral frequency content of the laser source.

Holographic interferometry. A wave recorded in a hologram is effectively stored for future reconstruction and use. Holographic interferometry is concerned with the formation and interpretation of the fringe pattern which appears when a wave, generated at some earlier time and stored in a hologram, is later reconstructed and caused to interfere with a comparison wave. It is the storage or time-delay aspect which gives the holographic method a unique advantage over conventional optical interferometry.

A hologram can be made of an arbitrarily shaped, rough scattering surface, and after suitable processing, if the hologram is illuminated with the same reference wavefront used in recording the hologram, the hologram will produce the original object wavefront. If the hologram is placed back into its original position, a person looking through the hologram will see both the original object and the image of the object stored in the hologram. If the object is now slightly deformed, interference fringes will be produced which tell how much the surface is deformed. Between adjacent fringes the optical path between the source and viewer has changed by one wavelength. While the actual shape of the object is not determined, the change in the shape of the object is measured to within a small fraction of a wavelength, even though the object's surface is rough compared to the wavelength of light.

Double-exposure. Double-exposure holographic interferometry (**Fig. 10**) is similar to real-time holographic interferometry described above, except now two exposures are made before processing: one exposure with the object in the undeformed state and a second exposure after deformation. When the hologram reconstruction is viewed, interference fringes

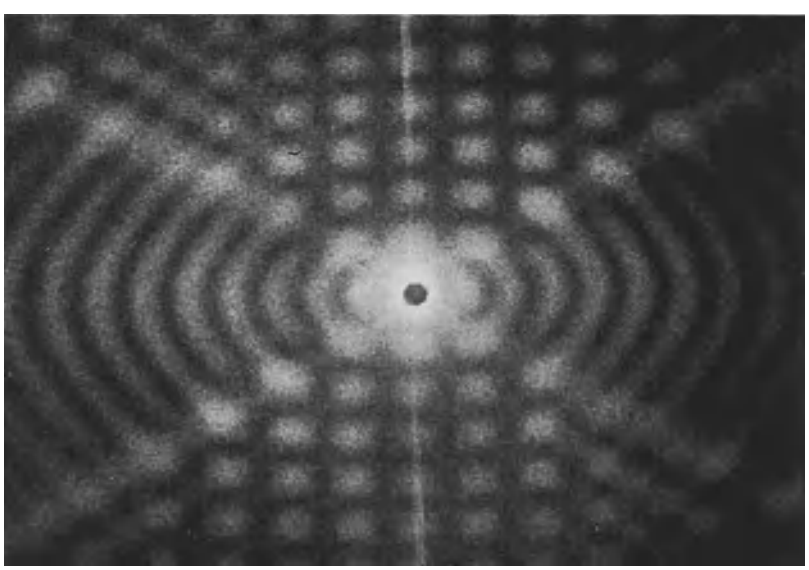

Fig. 12. Diffraction patterns from time-averaged speckle interferogram of a surface vibrating in its own plane with a figure-eight motion. (*From J. C. Dainty, Laser Speckle and Related Phenomena, Springer-Verlag, 1975*)

will be seen which show how much the object was deformed between exposures.

The advantage of the double-exposure technique over the real-time technique is that there is no critical replacement of the hologram after processing. The disadvantage is that continuous comparison of surface displacement relative to an initial state cannot be made, but rather only the difference between two states is determined.

Time-average. In time-average holographic interferometry (**Fig. 11**) a time-average hologram of a vibrating surface is recorded. If the maximum amplitude of the vibration is limited to some tens of light wavelengths, illumination of the hologram yields an image of the surface on which is superimposed several interference fringes which are contour lines of equal displacement of the surface. Time-average holography enables the vibrational amplitudes of diffusely reflecting surfaces to be measured with interferometric precision. *See* HOLOGRAPHY.

Speckle interferometry. A random intensity distribution, called a speckle pattern, is generated when light from a highly coherent source, such as a laser, is scattered by a rough surface. The use of speckle patterns in the study of object displacements, vibration, and distortion is becoming of more importance in the nondestructive testing of mechanical components. For example, time-averaged speckle photographs can be used to analyze the vibrations of an object in its plane. In physical terms the speckles in the image are drawn out into a line as the surface vibrates, instead of being double as in the double-exposure technique. The diffraction pattern of this smeared-out speckle-pattern recording is related to the relative time spent by the speckle at each point of its trajectory (**Fig. 12**). *See* NONDESTRUCTIVE EVALUATION.

Speckle interferometry can be used to perform astronomical measurements similar to those performed by the Michelson stellar interferometer. Stellar speckle interferometry is a technique for obtaining diffraction-limited resolution of stellar objects despite the presence of the turbulent atmosphere that limits the resolution of ground-based telescopes to approximately 1 second of arc. For example, the diffraction limit of the 200-in.-diameter (5-m) Palomar Mountain telescope is approximately 0.02 second of arc, $1/\sqrt{50}$ the resolution limit set by the atmosphere.

The first step of the process is to take a large number, perhaps 100, of short exposures of the object, where each photo is taken for a different realization of the atmosphere. Next the optical diffraction pattern, that is, the squared modulus of the Fourier transform of all the short-exposure photographs, is added. By taking a further Fourier transform of each ensemble average diffraction pattern, the ensemble average of the spatial autocorrelation of the diffraction-limited images of each object is obtained. *See* SPECKLE.

Phase-shifting interferometry. Electronic phase-measurement techniques can be used in interferometers such as the Twyman-Green, where the phase distribution across the interferogram is being measured. Phase-shifting interferometry is often used for these measurements since it provides for rapid precise measurement of the phase distribution. In phase-shifting interferometry, the phase of the reference beam in the interferometer is made to vary in a known manner. This can be achieved, for example, by mounting the reference mirror on a piezoelectric transducer. By varying the voltage on the transducer, the reference mirror is moved a known amount to change the phase of the reference beam a known amount. A solid-state detector array is used to detect the intensity distribution across the interference pattern. This intensity distribution is read into computer memory three or more times, and between each intensity measurement the phase of the reference beam is changed a known amount. From these three or more intensity measurements, the phase across the interference pattern can be determined to within a fraction of a degree. James C. Wyant

Bibliography. P. Hariharan, *Basics of Interferometry*, 1991; E. Hecht and A. Zajac, *Optics*, 3d ed., 1997; G. Hernandez, *Fabry-Perot Interferometers*, 1988; D. Malacara (ed.), *Optical Shop Testing*, 2d ed., 1992; P. K. Rastogi, *Holographic Interferometry: Principles and Methods*, 1994; R. S. Sirohi (ed.), *Speckle Metrology*, 1993; W. H. Steel, *Interferometry*, 2d ed., 1985.

Interhalogen compounds

The elements of the halogen family (fluorine, chlorine, bromine, and iodine) possess an ability to react with each other to form a series of binary interhalogen compounds (or halogen halides) of general composition given by XY; where n can have the values 1, 3, 5, and 7, and where X is the heavier (less electronegative) of the two elements. All possible diatomic compounds of the first four halogens have been prepared. In other groups a varying number of possible combinations is absent. Although attempts have been made to prepare ternary interhalogens, they have been unsuccessful; there is considerable doubt that such compounds can exist. *See* HALOGEN ELEMENTS.

Formation. In general, interhalogen compounds are formed when the free halogens are mixed as gases, or, as in the case of iodine chlorides and bromides, by reacting solid iodine with liquid chlorine or bromine. Most of the nonfluorinated interhalogens also readily form when solutions of the halogens in inert solvent (for example, carbon tetrachloride) are mixed. It is also possible to form them by the reaction of a halogen with a salt of a more electropositive halogen, such as $KI + Cl_2 \rightarrow KCl + ICl$. Higher polyhalides can also be prepared by reacting more electronegative halogen with a corresponding halogen halide, for example, $ICl + Cl_2 \rightarrow ICl_3$ or $ClF_3 + F_2 \rightarrow ClF_5$. Chlorine pentafluoride can also be prepared by reacting a $MClF_4$ salt with fluorine (M = alkali metal), $MClF_4 + F_2 \rightarrow MF + ClF_5$. A list of known interhalogen compounds and some of their

TABLE 1. Known interhalogen compounds

	XY	XY_3	XY_5	XY_7
	ClF	ClF_3	ClF_5	IF_7
mp	−154°C (−245°F)	−76°C (−105°F)	−103°C (−153°F)	
bp	−101°C (−150°F)	12°C (54°F)	−14°C (6.8°F)	
				4.77°C (40.6°F) (sublimes)
	BrF	BrF_3	BrF_5	
mp	≈−33°C (91°F)	8.77°C (47.8°F)	−62.5°C (−80.5°F)	
bp	≈20°C (68°F)	125°C (257°F)	40.3°C (105°F)	
	IF	IF_3	IF_5	
mp	—	−28°C (−18°F)	10°C (50°F)	
bp	—	—	101°C (214°F)	
	BrCl	ICl_3*		
mp	≈−54°C (−65°F)	101°C (214°F)		
bp	—	—		
	ICl†			
mp	27.2 C(α)			
bp	≈100°C (212°F)			
	IBr			
mp	40°C (104°F)			
bp	119°C (246°F)			

*In the solid state the compound forms a dimer.
†Unstable β-modification exists, mp 14° C (57° F).

physical properties is given in **Table 1**. Interhalogen compounds containing astatine have not been isolated as yet, although the existence of AtI and AtBr has been demonstrated by indirect measurements.

Stability. Thermodynamic stability of the interhalogen compounds varies within rather large limits. In general, for a given group the stability increases with increasing difference in electronegativity between the two halogens. Thus for XY group the free energy of formation of the interhalogens, relative to the elements in their standard conditions, falls in the following order: IF > BrF > ClF > ICl > IBr > BrCl. It should be noted, however, that the fluorides of this series can be obtained only in minute quantities since they readily undergo disproportionation reaction, for example, $5IF \rightarrow 2I_2 + IF_5$. The least-stable compound, bromine chloride, has only recently been isolated in the pure state. Decrease of stability with decreasing difference of electronegativity is readily apparent in the higher interhalogens since compounds such as $BrCl_3$, IBr_3, or ICl_5 are unknown. The only unambiguously prepared interhalogen containing eight halogen atoms is IF_7. *See* ELECTRONEGATIVITY.

Reactivity. The reactivity of the polyhalides reflects the reactivity of the halogens they contain. In general, they behave as strong oxidizing and halogenating agents. Most halogen halides (especially halogen fluorides) readily attack metals, yielding the corresponding halide of the more electronegative halogen. In the case of halogen fluorides, the reaction results in the formation of the fluoride, in which the metal is often found in its highest oxidation state, for example, AgF_2, CoF_3, and so on. Noble metals, such as platinum, are resistant to the attack of the interhalogens at room temperature. Halogen fluorides are often handled in nickel vessels, but in this case the resistance to attack is due to the formation of a protective layer of nickel(II) fluoride.

All halogen halides readily react with water. Such reactions can be quite violent and, with halogen fluorides, they may be explosive. Reaction products vary depending on the nature of the interhalogen compound. For example, in the case of chlorine trifluoride, reaction with an excess of water yields HF, Cl_2, and O_2 as reaction products.

The same reactivity is observed with organic compounds. The nonfluorinated interhalogens do not react with completely halogenated hydrocarbons, and solutions of ICl, IBr, and BrCl are quite stable in dry carbon tetrachloride or hexachloroethane, as well as in fluorocarbons, as long as the solvents are very dry. They readily react with aliphatic and aromatic hydrocarbons and with oxygen- or nitrogen-containing compounds. The reaction rates, however, can be rather slow and dilute solutions of ICl can be stable for several hours in solvents such as nitrobenzene. Halogen fluorides usually react vigorously with chlorinated hydrocarbons, although IF_5 can be dissolved in carbon tetrachloride and BrF_3 can be dissolved in Freon 113 without decomposition. All halogen fluorides react explosively with easily oxidizable organic compunds. *See* FLUORINE.

Halogen halides, like halogens, act as Lewis acids and under proper experimental conditions may form a series of stable complexes with various organic electron donors. For example, mixing of carbon tetrachloride solutions of pyridine and of iodine monochloride leads to the formation of a solid complex, $C_5H_5 \cdot ICl$. The same reaction can occur with other heterocyclic amines and ICl, IBr, or ICl_3. Addition compounds of organic electron donors with IF, IF_3, and IF_5 have been reported. In all cases it is the

iodine atom which is directly attached to the donor atom.

A number of interhalogen compounds conduct electrical current in the liquid state. Among these are ICl and BrF_3. For example, electrical conductance of molten iodine monochloride is comparable to a concentrated aqueous solution of a strong electrolyte (4.52×10^{-3} ohm^{-1} cm^{-1} at 30.6°C or 87.08°F). The conductances, however, are much smaller than those of fused salts and, therefore, it can be concluded that the bonding in these compounds is largely covalent. Electrical conductance is due to self-ionization reactions, as shown in reactions (1) and (2).

$$3ICl \rightleftharpoons I_2Cl^+ + ICl_2^- \quad (1)$$

$$2BrF_3 \rightleftharpoons BrF_2^+ + BrF_4^- \quad (2)$$

The above behavior leads to the possibility of studying acid-base reactions. In these systems an acid is any compound which generates the solvo-cation, while a base would generate solvo-anions. Thus SbF_5 would be an acid in liquid bromine trifluoride, reaction (3), while an electrovalent fluoride would be a base, reaction (4).

$$SbF_5 + BrF_3 \rightleftharpoons BrF_2^+ + SbF_6^- \quad (3)$$

$$KF + BrF_3 \rightleftharpoons K^+ + BrF_4^- \quad (4)$$

See SUPERACID.

Analogy with acid-base reactions in water is obvious, as shown by reactions (5) and (6). Such relations

$$K^+OH^- + H_3O^+Cl^- \rightarrow K^+Cl^- + 2H_2O \quad (5)$$

$$K^+BrF_4^- + BrF_2^+SbF_6^- \rightarrow KSbF_6 + 2BrF_3 \quad (6)$$

have been studied in BrF_3, ClF_3, and IF_5. Numerous salts of the interhalogen acid-base systems have been isolated and studied.

Thus, numerous compounds have been formed containing either interhalogen anions (solvo-anions) or interhalogen cations (solvo-cations) simply by adding the appropriate acid or base to a liquid halogen halide or a halogen halide in an appropriate nonaqueous solvent. In addition, cations derived from previously unknown compounds can be prepared by using powerful oxidizing agents, such as KrF^+ salts. For example, even though to date, a compound containing BrF^+_6 has been prepared according to reaction (7).

$$BrF_5 + KrF^+AsF_6^- \rightarrow BrF_6^+AsF_6^- + Kr \quad (7)$$

Pentahalides can also be formed by the addition of an interhalogen compound to a trihalide ion as shown in reaction (8).

$$ICl + ICl_2^- \rightarrow I_2Cl_3^- \quad (8)$$

A compilation of interhalogen anions and cations which have been previously prepared is given in **Tables 2** and **3**.

Terry Surles

Bibliography. V. Gutmann (ed.), *MTP International Review of Science: Inorganic Chemistry*, Ser. 1, vol. 3, 1975.

TABLE 2. Known interhalogen anions

Three-membered	Five-membered	Seven-membered	Nine-membered
ClF_2^-	ClF_4^-	ClF_6^-	IF_3^-
BrF_2^-	BrF_4^-	BrF_6^-	
ICl_2^-	IF_4^-	IF_6^-	
IBr_2^-	ICl_4^-		
$IBrCl^-$	$I_2Cl_3^-$		
$BrCl_2^-$	$I_2Cl_2Br^-$		
I_2Cl^-	$I_2ClBr_2^-$		
Br_2Cl^-	I_4Cl^-		
I_2Br^-			

TABLE 3. Known interhalogen cations

Three-membered	Five-membered	Seven-membered
Cl_2F^+	ClF_4^+	ClF_6^+
ClF_2^+	BrF_4^+	BrF_6^+
BrF_2^+	IF_4^+	
ICl_2^+		
I_2Cl^+		
IBr_2^+		
I_2Br^+		
$BrCl_2^+$		
Br_2Cl^+		
$IBrCl^+$		

Intermediate-frequency amplifier

An amplifying circuit in a radio-frequency (RF) receiver that processes and enhances a downconverted or modulated signal. Signal frequency spectrum downconversion is achieved by multiplying the radio-frequency signal by a local oscillator signal in a circuit known as a mixer. This multiplication produces two signals whose frequency content lies about the sum and difference frequencies of the center frequency of the original signal and the oscillator frequency. A variable local oscillator is used in the receiver to hold the difference-signal center frequency constant as the receiver is tuned. The constant frequency of the downconverted signal is called the intermediate frequency (IF), and it is this signal that is processed by the intermediate-frequency amplifier.

Unfortunately, radio-frequency signals both higher and lower than the local oscillator frequency by a difference equal to the intermediate frequency will produce the intermediate frequency. One of these is the desired signal; the undesired signal is called an image. *See* MIXER; OSCILLATOR.

Superheterodyne receiver. Aside from demodulation and conversion, the purpose of each stage of a radio receiver is to improve the signal-to-noise ratio (SNR) through a combination of signal amplification and noise/interference suppression. The first stage of a superheterodyne receiver is generally a broadband tuned radio-frequency amplifier intended to improve the signal-to-noise ratio of a selected signal whose

frequency lies anywhere within the receiver's tuning range. The second stage is a variable-frequency mixer or downconverter that shifts the frequency of the signal of interest to the predetermined intermediate frequency by use of a local oscillator (LO). Waveform impurities in the LO signal and nonlinearities in the RF amplifier and mixer may interact with the received signal to generate undesirable false signals that are replicas or images of the received signal but displaced in frequency. The third stage of that receiver is a high-quality intermediate-frequency amplifier that has been optimized to improve the quality of the signal in a very narrow frequency band about the fixed intermediate frequency by greatly attenuating out-of-band noise, adjacent-channel signals, and images. More sophisticated digital intermediate-frequency stages can also reduce broadband interference.

Unlike the broadband tunable radio-frequency amplifier, the intermediate-frequency amplifier is designed to operate over a narrow band of frequencies centered about a dedicated fixed frequency (the intermediate frequency); therefore, the intermediate-frequency amplifier can be an extremely efficient stage. If the intermediate frequency is on the order of a few megahertz, the undesirable images may be efficiently rejected, but narrow-band filtering for noise and adjacent-channel-signal rejection is difficult and expensive because of the high ratio of the intermediate frequency to the bandwidth of the intermediate-frequency amplifier. If the intermediate frequency is much smaller, say, on the order of a few hundred kilohertz, then inexpensive and more selective filters are possible that can separate the desired signal from closely packed adjacent signals, but they do not reject images very well. A high-quality double-conversion receiver combines the best of both approaches by cascading both high- and low-frequency intermediate-frequency stages that are separated by a second fixed-frequency mixer. The last intermediate-frequency stage of a digital or analog superheterodyne receiver is followed by the detector, which extracts the information from the signal. *See* HETERODYNE PRINCIPLE; RADIO-FREQUENCY AMPLIFIER; RADIO RECEIVER; TELEVISION RECEIVER.

Applications. The superheterodyne structure is common for television, ground-based and satellite communications, cell phones, ground-based and airborne radar, navigation, and many other receivers. The intermediate-frequency amplifier function is ubiquitous.

Gain and bandwidth. The measure of quality of an intermediate-frequency amplifier is how well it forms a "window" in the frequency spectrum for observing signals of interest and rejecting all other signals (including images) and noise. Maximum gain is usually at the intermediate frequency. The intermediate-frequency-amplifier gain decreases at frequencies both higher and lower than the intermediate frequency. The difference between the upper and lower frequencies at which the signal power drops to one-half of the maximum signal power (usually at the intermediate frequency) is known as the selectivity or the 6-dB bandwidth. The difference between the upper and lower frequencies at which the signal power drops to one-thousandth of the maximum signal power is known as the 60-dB bandwidth. The ratio of the 60-dB bandwidth and the 6-dB bandwidth is known as the shape factor, a number that must always be greater than one. A small shape factor indicates good noise rejection because the filter rejects any signal outside of the bandwidth of interest. The best selectivity matches the bandwidth of the signal.

Because a single receiver may be used for several different applications, a different intermediate-frequency filter may be employed for each application. The filter requirements vary widely with application. For example, signal bandwidth for broadcast-quality frequency modulation (FM) may be in the tens of kilohertz; for communications-quality FM, 7–10 kHz; for broadcast-quality amplitude modulation (AM), 5 kHz; for communications-quality AM, 3 kHz; for single sideband (SSB), 1–2 kHz; and for continuous-wave (CW) communications (radio-telegraphy), under 100 Hz. *See* BANDWIDTH REQUIREMENTS (COMMUNICATIONS); GAIN; SELECTIVITY.

Analog IF amplifiers. Analog (or continuous-time) intermediate-frequency amplifiers consist of narrow-band tuned circuits that filter out nearly all signals and noise except the desired signal at the intermediate frequency, and amplification to increase the signal strength. The original tuned-circuit structures were made of capacitors and inductors. Today virtually all high-Q intermediate-frequency filters are only a part of a monolithic integrated circuit that contains the entire radio receiver, including fully integrated RC/op-amp active filters, piezoelectric crystal filters, magnetostrictive lines, and micro-electro-mechanical systems (MEMS). *See* ELECTRIC FILTER; INTEGRATED CIRCUITS; MAGNETOSTRICTION; MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS); OPERATIONAL AMPLIFIER; PIEZOELECTRICITY; Q (ELECTRICITY).

Digital IF amplifiers. This misnomer came to be because the concept is easy to grasp in view of its parallel to analog IF amplifiers. Either digital IF stage or digital IF filter would be a more appropriate name.

The current goal of radio-frequency system design is to make an analog-to-digital converter (ADC) directly connected to the antenna to permit realization of an all-digital radio receiver. Fast devices permit analog-to-digital converters that operate at near-terahertz (10^{12} Hz) speed. In 1999 such devices were still costly and power-hungry, but the same was true of bit-serial near-megahertz devices around 1970. There now exist reasonably priced and powered radio receivers that consist of an analog radio-frequency amplifier and first converter, followed by a high-speed analog-to-digital converter, which is connected to the remainder of the digital-receiver processing, beginning with the intermediate-frequency stage. The digital-receiver function can be produced as a dedicated integrated circuit or as a part of a programmed microprocessor, depending on the frequency range of interest. *See* MICROPROCESSOR.

A digital intermediate-frequency stage is far more versatile and controllable than its analog counterpart. For example, while changing the selectivity of an analog filter may require switching to different physical signal paths, changing the selectivity of a digital filter requires only the change of stored numerical filter coefficients. Therefore, selective filtering, provided by a digital intermediate-frequency stage, can be more accurate than that of its analog counterpart, absolutely repeatable, and even program-controlled. Such filtering can be used to attenuate out-of-band noise and to sharply reduce in-band interference by use of individual (or a combination of) programmed, adaptive, and manually tuned notch filters. Furthermore, broadband noise can be greatly attenuated by using a digital adaptive noise suppressor. *See* AMPLIFIER.

Stanley A. White

Bibliography. R. C. Dorf (ed.), *The Electrical Engineering Handbook*, 2d ed., 1997; S. Gibilisco, *Handbook of Radio and Wireless Technology*, 1998; R. F. Graf, *Converter and Filter Circuits*, 1997; J. E. Paul, Adaptive Noise-Cancelling Receiver, U.S. Patent No. 4, 177, 430; M. S. Roden, *Analog and Digital Communications Systems*, 4th ed., 2000; M. E. Van Valkenburg (ed.), *Reference Data for Engineers: Radio, Electronics, Computer, and Communications*, 8th ed., 1996.

Intermediate vector boson

One of the three fundamental particles that transmit the weak force. (An example of a weak interaction process is nuclear beta decay.) These elementary particles—the W^+, W^-, and Z^0 particles—were discovered in 1983 in very high energy proton-antiproton collisions. It is through the exchange of W and Z bosons that two particles interact weakly, just as it is through the exchange of photons that two charged particles interact electromagnetically. The intermediate vector bosons were postulated to exist in the 1960s; however, their large masses prevented their production and study at accelerators until 1983. Their discovery was a key step toward unification of the weak and electromagnetic interactions. *See* ELECTROWEAK INTERACTION; ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS; WEAK NUCLEAR INTERACTIONS.

Production and detection. The W and Z particles are roughly 100 times the mass of a proton. Therefore, the experiment to search for the W and the Z demanded collisions of elementary particles at the highest available center-of-mass energy. Such very high center-of-mass energies capable of producing the massive W and Z particles were achieved with collisions of protons and antiprotons at the laboratory of the European Organization for Nuclear Research (CERN) near Geneva, Switzerland. The $p\bar{p}$ collisions were monitored in two underground experimental areas (UA1 and UA2). *See* PARTICLE ACCELERATOR; PARTICLE DETECTOR.

Properties of W and Z particles. Striking features of both the charged W and the Z^0 particles are their large masses. The charged boson (W^+ and W^-) mass is measured to be about 80 GeV/c^2, and the neutral boson (Z^0) mass is measured to be about 91 GeV/c^2. (For comparison, the proton has a mass of about 1 GeV/c^2.) Prior to the discovery of the W and the Z, particle theorists had met with some success in the unification of the weak and electromagnetic interactions. The electroweak theory as it is understood today is due largely to the work of S. Glashow, S. Weinberg, and A. Salam. Based on low-energy neutrino scattering data, which in this theory involves the exchange of virtual W and Z particles, theorists made predictions for the W and Z masses. The actual measured values are in agreement (within errors) with predictions. The discovery of the W and the Z particles at the predicted masses is an essential confirmation of the electroweak theory, one of the cornerstones of the standard model of particle physics. *See* STANDARD MODEL.

Only a few intermediate vector bosons are produced from 10^9 proton-antiproton collisions at a center-of-mass energy of 540 GeV. This small production probability per $p\bar{p}$ collision is understood to be due to the fact that the bosons are produced by a single quark-antiquark annihilation. The other production characteristics of the intermediate vector bosons, such as longitudinal and transverse momentum distributions (with respect to the $p\bar{p}$ colliding beam axis), provide support for this theoretical picture. *See* QUARKS.

Decay. The decay modes of the W and Z are well predicted. The simple decays $W^+ \rightarrow e^+\nu$, $W^- \rightarrow e^-\bar{\nu}$, $Z^0 \rightarrow e^+e^-$, and $Z^0 \rightarrow \mu^+\bar{\mu}$ are spectacular signatures of the intermediate vector bosons (**Fig. 1**). These leptonic decays are only a few percent of the total number of W and Z decays, however, because the W and Z also decay into

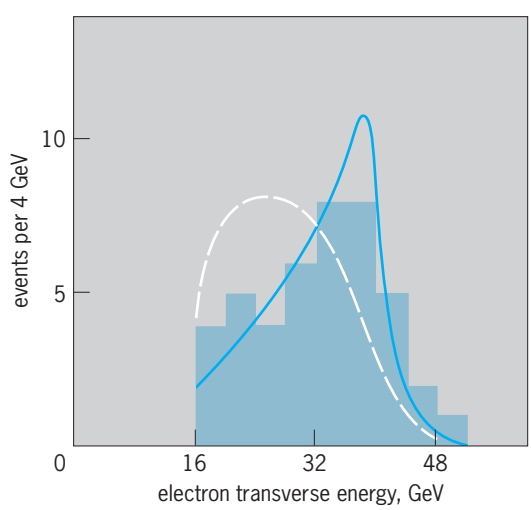

Fig. 1. Observation of the *W* particle in the UA1 data from the electron transverse energy distribution. The solid curve is the expected distribution for the decay, $W \rightarrow e\nu$, while the broken curve is the distribution expected for the decay of a different particle "X" into an electron and two neutrinos.

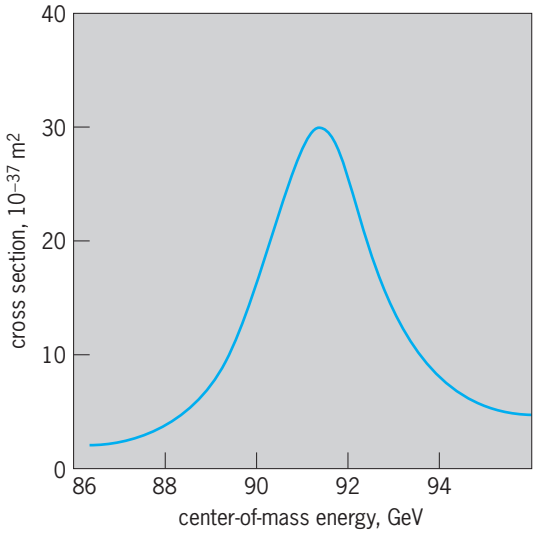

Fig. 2. Rates from the L3 experiment at LEP for the process $e^+e^- \rightarrow Z^0$ as a function of total center-of-mass energy of the electron plus positron. The rate is measured as a total cross section. (*After O. Adriani et al., Results from the L3 Experiment at LEP, Phys. Rep., 236:1–146, 1993*)

quark-antiquark pairs. The W particle also has an identifying feature in its decay, which is due to its production through quark-antiquark annihilation. The quark and antiquark carry intrinsic angular momentum (spin) that is conserved in the interaction. This means that the W is polarized at production; its spin has a definite orientation. This intrinsic angular momentum is conserved when the W decays, yielding a preferred direction for the decay electron or positron. This distribution is characteristic of a weak interaction process, and it determines the assignment of the spin quantum number of the W to be 1, as expected from theory. (Spin-1, or more generally integer-spin, particles are called bosons; another example is the photon.) *See* QUANTUM STATISTICS; SPIN (QUANTUM MECHANICS).

Results from electron-positron annihilation. In 1989, the Z^0 particle was produced in electron-positron annihilations at both the Stanford Linear Collider (SLC) and the CERN Large Electron-Positron (LEP) accelerator. The Z^0 was first observed in these experiments by its decay into a quark-antiquark ($q\bar{q}$) pair. The quark and antiquark are not observed as free particles; however, the energetic hadrons originating from them give a clear experimental signature. This process, $e^+e^- \rightarrow Z^0 \rightarrow q\bar{q}$, is the inverse of the fundamental process in which the Z^0 particle was discovered, $q\bar{q} \rightarrow Z^0 \rightarrow e^+e^-$, where the quark and antiquark are constituents of the energetic proton and antiproton. Electron-positron annihilations offer an exceptionally background-free method for the study of the Z^0 particle because it is created alone as a pure state. Thus, the LEP machine with its high collision rate is appropriately referred to as a Z^0 factory. Since the Z^0 particle couples to all known quarks and leptons, the study of its decays also provides information about the properties of all known particles.

The mass of the Z particle has been extremely accurately determined at LEP by varying the energy of the electrons and positrons by a few percent and observing the rate of the process $e^+e^- \rightarrow Z^0 \rightarrow q\bar{q}$ as a function of energy (**Fig. 2**). The result from four LEP experiments is that $m_Z = 91.1876$ GeV/c^2 with an experimental error of 0.0021 GeV/c^2. Many details of the decay of Z particles have been measured at LEP with high statistics. All are in agreement with the standard model. The final stages of LEP (LEP II) allowed an increase in machine energy so that pairs of W^+ and W^- were produced.

The W and Z particles have also been extensively studied at the Fermilab Tevatron proton-antiproton collider. Detection of the W particles was the key element in the observation of the top quark at Fermilab. The data from LEP II and Fermilab allowed accurate determination of the W mass. The present value is $m_W = 80.423$ GeV/c^2 with an experimental error of 0.039 GeV/c^2.

The Z^0 particle has an extremely short lifetime of about 10^{-25} s. This is due in part to the large number of channels into which it may decay. The result of this short lifetime is that the Z^0 particle is observed as a resonance having a natural width (Γ) of 2.495 GeV/c^2 (Fig. 2). There is a partial contribution to this width for each decay channel available to the Z^0. These partial widths can be calculated theoretically. The particles that are known to contribute to the Z^0 width are the electron, muon, and tau leptons ($Z^0 \rightarrow e^+e^-$, $\mu^+\mu^-$, or $\tau^+\tau^-$); the up, down, strange, charm, and bottom quarks ($Z^0 \rightarrow q\bar{q}$); and the electron-, muon-, and tau-neutrinos ($Z^0 \rightarrow \nu\bar{\nu}$). There is good agreement between the measured width and the sum of the theoretical partial widths for each known decay channel. The existence of possible new leptons, quarks, or neutrinos with masses less than about $m_Z/2$ is therefore ruled out. *See* LEPTON; NEUTRINO.

Now the focus in particle physics is to understand why the weak interaction mediators (W and Z) are massive while the electromagnetic interaction mediator (photon) is massless, since the two forces are unified. A new proton-proton accelerator, the Large Hadron Collider (LHC), will come into operation in 2008 to address this puzzle. The center-of-mass collision energy of 14 TeV and the high luminosity of 10^{34} cm^{-2} s^{-1} will provide, for the first time, interactions of W and Z particles at energies much larger than the mass energies of the bosons, providing the first window to probe the mechanism by which elementary particles obtain their masses.

James W. Rohlf

Bibliography. G. Arnison et al., Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV, *Phys. Lett.*, 122B:103–116, 1983; G. Arnison et al., Experimental observation of lepton pairs of invariant mass around 95 GeV/c^2 at the CERN SPS collider, *Phys. Lett.*, 126B:398–410, 1983; The LEP collaborations: Electroweak parameters of the Z^0 resonance and the standard model, *Phys. Lett.*, B276:247–253, 1992; J. W. Rohlf, *Modern Physics from α to Z^0*, 1994.

Intermetallic compounds

All metal–metal compounds, including ordered, disordered, binary, and multicomponent. The metal–metal aspect of the definition is often relaxed to include some metal–metalloid compounds such as silicides and tellurides and certain compound semiconductors such as InSb. Such inclusion is appropriate since the phenomenology of many of these compounds is similar to the metal-metal ones. Intermetallic compounds constitute an intriguing field for scientific inquiry and a rewarding source of materials for diverse applications. *See* METAL; METALLOID; SEMICONDUCTOR.

Binary systems. Some of the formation characteristics of intermetallics may be appreciated from examination of the phase diagram such as that for the binary system Al-Ni in **Fig. 1**. Here it is seen that intermetallics can melt congruently (for example, AlNi), and they form by solid–liquid reactions (such as Al_3Ni, Al_3Ni_2, or $AlNi_3$) or interaction between solids (such as Al_3Ni_5). The phase diagram also shows that intermetallics can exhibit broad ranges of composition, extending on one or both sides of stoichiometry (for example, AlNi), or may show essentially no deviation from stoichiometry (for example, Al_3Ni). These same features also appear in ternary and higher-order systems. Stoichiometric deviation implies the presence of point defects (isolated atoms), including substitutions, vacancies, or interstitials. *See* CRYSTAL DEFECTS; EUTECTICS; NONSTOICHIOMETRIC COMPOUNDS; PHASE EQUILIBRIUM; SOLID-STATE CHEMISTRY; STOICHIOMETRY.

Crystal structure. Another characteristic of intermetallic compounds is that their crystal structures are idiosyncratic, not to be inferred from the structures of the component metals. Aluminium and nickel are face-centered-cubic (fcc) metals, but the five intermetallics shown in Fig. 1 have unique structures, none of them the simple fcc of the component metals. In this system, $AlNi_3$ exhibits an ordered variant of the fcc structure, a so-called superstructure (**Fig. 2***a*), but the others are quite different—orthorhombic, trigonal, or ordered body-centered-cubic (bcc). In this particular system, all the intermetallics remain ordered until the melting point is reached, but in other cases (for example, compounds in the Au-Cu system) the ordering disappears well before the melting point. In a few cases, a discrete compound exists in the system over a limited composition range and exhibits no ordering at all (for example, Ag_3Al). Some intermetallic crystal structures are quite simple with a low number of atoms per unit cell. For example, AlNi has the CsCl structure (ordered bcc) with only two atoms per unit cell (Fig. 2*b*), while others, such as $NaCd_2$ with a simple formula, have very complex structures with hundreds of atoms per unit cell. In general, intermetallic crystal structures tend to be high-symmetry, high-space-filling, and high-coordination-number. Although some 2750 intermetallic crystal structures are known, ~5000 are potentially possible. The 100 most common crystal structures account for about

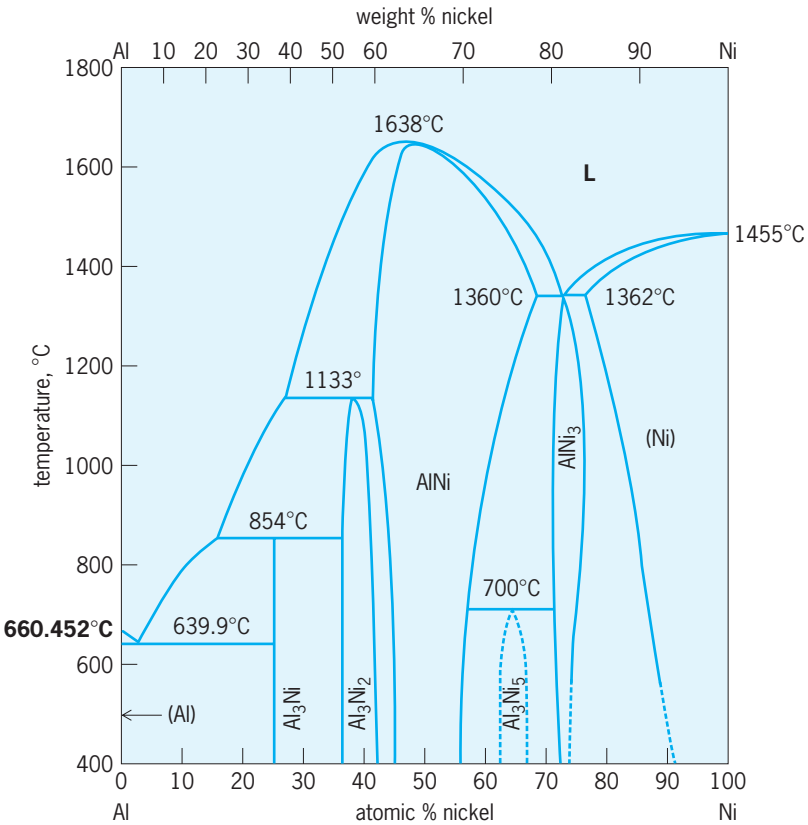

Fig. 1. Al-Ni phase diagram. Note that the peritectic formation of $AlNi_3$ lies only 2° above the eutectic between AlNi and $AlNi_3$. (*After J. H. Westbrook, ed., Handbook of Binary Phase Diagrams, Genium Publishing, 2005*)

72% of all known intermetallics (~25,000). The particular structure adopted by a pair of metal atoms depends on their relative size, the atomic number factor, the electrochemical factor, the valence electron factor, and the angular valence–orbital factor. The relative importance of these factors varies from case to case. *See* ATOM; COORDINATION NUMBER; CRYSTAL STRUCTURE; VALENCE.

Ternary systems. For ternary systems, the situation is even more complex and little is known. About 15,000 ternary compounds have been identified, but this is only about 3% of the expected number of the possible ternary intermetallics. Ternary compounds are not just of conceptual interest but also of commercial significance, as seen by the important magnetic phase, $Nd_2Fe_{14}B$, and the embrittling G-phase, $Ni_{16}Ti_6Si_7$, sometimes encountered in high-temperature alloys. Complications do not end with ternaries. A seven-component intermetallic ($Cu_{11.5}Zn_{23}Ga_{7.5}Ge_8As_{16}Se_{30}Br_4$) with a disordered zinc blende structure has been successfully synthesized. Quasicrystals, structures lacking long-range order but exhibiting quasiperiodicity and icosahedral or decagonal structures, and intermetallic clusters (for example, $Ba_{16}Na_{204}Sn_{310}$) are also intriguing intermetallics. *See* ALLOY; ALLOY STRUCTURES; QUASICRYSTAL.

Formation. Intermetallic compounds can be prepared in various ways, depending on the particular compound and the intended product. Possible methods include melting and casting, powder metallurgy,

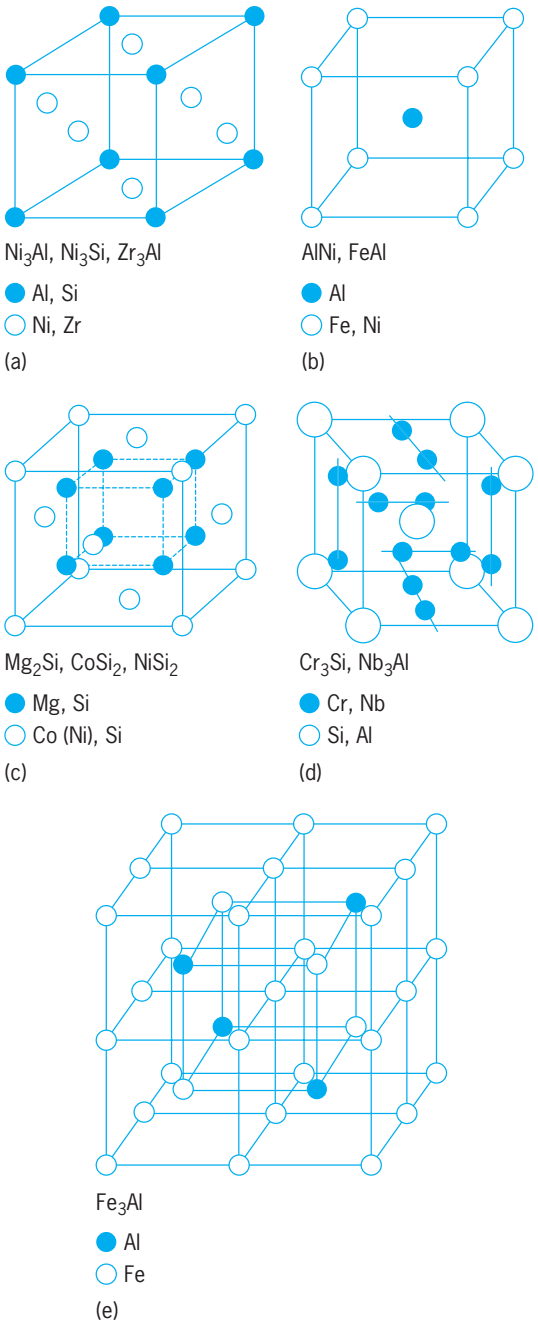

Fig. 2. General cubic lattices of aluminides and silicides. (*a*) $L1_2$ [cP4], (*b*) B2 [cP2], (*c*) C1 [cF12], (*d*) A15 [cP8], (*e*) $D0_3$ [cF16]. (*After K. H. J. Buschow et al., eds., The Encyclopedia of Materials: Science and Technology, vol. 5, p. 4177, Pergamon, 2001*)

mechanical alloying, diffusion, nanotech synthesis, and molecular-beam epitaxy. They also can occur naturally as minerals—for example, braithauptite (NiSb), domeykite (Cu_3As), froodite ($PdBi_2$), drysdallite ($MoSe_2$), nigglite (PtSn), petzite [$(Ag,Au)Te_2$], kolymite (Cu_7Hg_6), and hapkeite (Fe_2Si), to cite a single example in each family. *See* DIFFUSION; MECHANICAL ALLOYING; POWDER METALLURGY.

Applications. Practical use of intermetallics goes far back in time. In prehistoric times, meteoritic NiFe was used for tools and Cu_3Au (tumbaga) for fishhooks. In Roman times, CuSn was used for mirrors and Cu_3Sn in bronze swords and bells. In the eighteenth and nineteenth centuries, CuZn was used for ship sheathing, SbSn in printing type and bearings, and Ag_2Hg_3 and Sn_8Hg in dental amalgams. Today, intermetallics are used in many high-tech applications, such as Ni_3Al in jet engine alloys, $FeZn_3$ in automotive galvanized steel sheet, FeAl as a resistive heating element, GaAs in light-emitting diodes and semiconducting substrates, $LaNi_5$ as a battery electrode, $(Tb,Dy)Fe_2$ as a magnetoelastic transducer, and $Ni_{0.46}Ti_{0.54}$ as shape memory alloy in the biomedical field. Other specialized applications include protective coatings, catalysts, heat storage materials, thermoelectrics, pyrophoric alloys, and jewelry. The market for intermetallics in the United States has been estimated at $10 billion per year.

Other intermetallics. There are a few cases where intermetallics exhibit truly exceptional behavior. $LaGa_3$ is a "plastic" crystal with orientational ordering. Cs_4Pb_4 undergoes a two-step melting (first loss of rotational order, then loss of positional order). CeSb has the most complex magnetic system known (>17 magnetic structures). Colored intermetallics include CoGa (yellow), $CoSi_2$ (dark blue), $AuAl_2$ (violet), $AuLi_2Sn$ (pink), and FeAl (brown), to name some. Pd_8U is a thermoelectric with the very large positive thermopower for a metal. $SmMn_2Ge_2$ is a magnetoresistor. NaTl is a substance having ionic, covalent, and metallic bonds, all in the same crystal. Mg_3Bi_2 is a superionic conductor with mobile Mg^{2+} cations in a metallic framework. And Cu_2MnAl is a ferromagnet with no ferromagnetic constituents.

J. H. Westbrook

Bibliography. F. R. de Boer and D. G. Pettifor (eds.), *Cohesion and Structure*, North-Holland, 1991; M. Doyama and M. Yabe, *Databook of Intermetallic Compounds*, Science Forum, Tokyo, 1988; R. D. Dudley and P. D. Desai, *Properties of Intermetallic Alloys III*, CINDAS, Purdue University, 1995; R. L. Fleischer and J. H. Westbrook, *Intermetallic Compounds; Principles and Practice*, Wiley, vol. 1, 1996, vol. 2, 1996, vol. 3, 2002; B. P. Gilp and P. D. Desai, *Properties of Intermetallic Alloys II*, CINDAS, Purdue University, 1994; C. T. Liu, R. W. Cahn, and G. Sauthoff (eds.), *Ordered Intermetallics: Physical Metallurgy and Mechanical Behavior*, Kluwer, 1992; N. M. Matveeva and E. V. Kozlov, *Ordered Phases in Metallic Systems*, Nova Science, 1996; National Materials Advisory Board, *Intermetallic Alloy Development: A Program Evaluation*, NMAB-487-1, National Academy Press, 1997; J. E. Payne and P. D. Desai, *Properties of Intermetallic Alloys I*, CINDAS, Purdue University, 1994; H. Ringpfeil (ed.), *Intermetallische Phasen*, VEB Deutscher Verlag, Leipzig, 1976; G. Sauthoff, *Intermetallics*, VCH, 1995; N. S. Stoloffand and V. K. Sikka, *Physical Metallurgy and Processing of Intermetallic Compounds*, Chapman and Hall, 1996; P. Villars, *Pearson's Handbook: Crystallographic Data for Intermetallic Phases*, Desk ed., 2 vols., ASM, 1997; H. Warlimont (ed.), *Order-Disorder Transformations in Alloys*, Springer, 1974.

Intermolecular forces

Attractive or repulsive interactions that occur between all atoms and molecules. Intermolecular forces become significant at molecular separations of about 1 nanometer or less, but are much weaker than the forces associated with chemical bonding. They are important, however, because they are responsible for many of the physical properties of solids, liquids, and gases. These forces are also largely responsible for the three-dimensional arrangements of biological molecules and polymers.

Attractive and repulsive forces. The observation that gases condense into liquids and solids at low temperatures, rather than continuing to occupy all the space available to them, is evidence of a force of attraction between their molecules. This attractive force is overcome only when the thermal energy is high enough. The further observation that liquids and solids occupy a finite volume and cannot be compressed easily is evidence of a repulsive force that comes into play when the molecules are close together.

Accordingly, we can deduce that the energy of interaction between two molecules or atoms takes a form like that in **Fig. 1**, which shows the interaction between argon atoms as deduced from experimental data. Argon atoms are electrically neutral and do not form chemical bonds with each other. When two argon atoms are far apart, the energy of interaction is essentially zero, and the atoms do not exert a force on each other. (The force is the negative gradient, or slope, of the potential energy.) As the two atoms approach, the energy between them decreases, reaching a minimum at a distance of 0.38 nm and then increasing again. At distances greater than 0.38 nm, the energy increases with increasing distance, and the force between the molecules is attractive; at shorter distances, the energy increases rapidly with decreasing distance, and the force becomes strongly repulsive. At the minimum, the interaction energy between the argon atoms is about 1.2 kJ/mol, which may be compared with typical chemical bond energies, of the order of 300 kJ/mol. *See* ARGON; CHEMICAL BONDING.

Description. Intermolecular forces can be classified into several types, of which two are universal (as illustrated by the same example of two interacting argon atoms). The attractive force known as dispersion arises from the quantum-mechanical fluctuation of the electron density around the nucleus of each atom. At long distances—greater than 1 nm or so—the electrons of each atom move independently of the other, and the charge distribution is spherically symmetric (**Fig. 2***a*). At shorter distances, an instantaneous fluctuation of the charge density in one atom can affect the other. If the electrons of one atom move briefly to the side nearer the other (Fig. 2*b*), the electrons of the other atom are repelled to the far side. In this configuration, both atoms have a small dipole moment, and they attract each other electrostatically. At another moment, the electrons may move the other way, but their motions are correlated so that an attractive force is maintained on average. Molecular orbital theory shows that the electrons of each atom are slightly more likely to be on the side nearer to the other atom, so that each atomic nucleus is attracted by its own electrons in the direction of the other atom. Fritz London demonstrated in 1930 that at long range the dispersion energy is

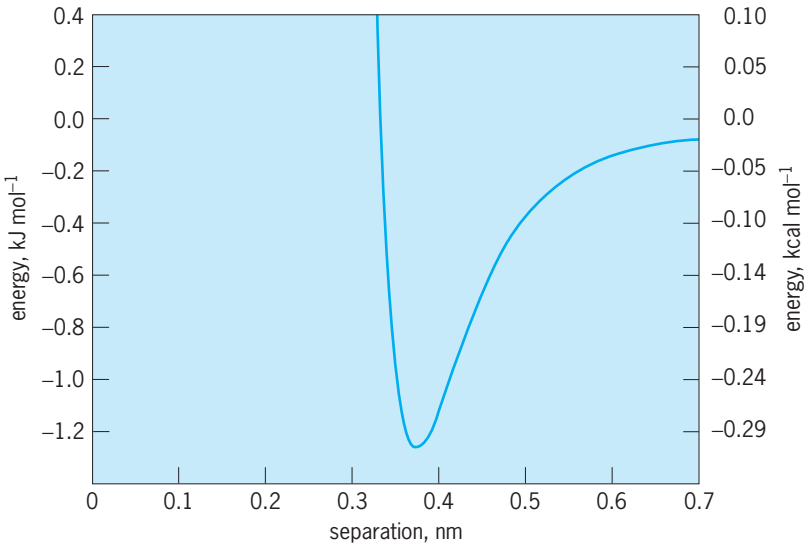

Fig. 1. Intermolecular potential energy of two argon atoms. (*After R. A. Aziz and H. H. Chen, An accurate intermolecular potential for Ar, J. Chem. Phys., 67:5719-5726, 1977*)

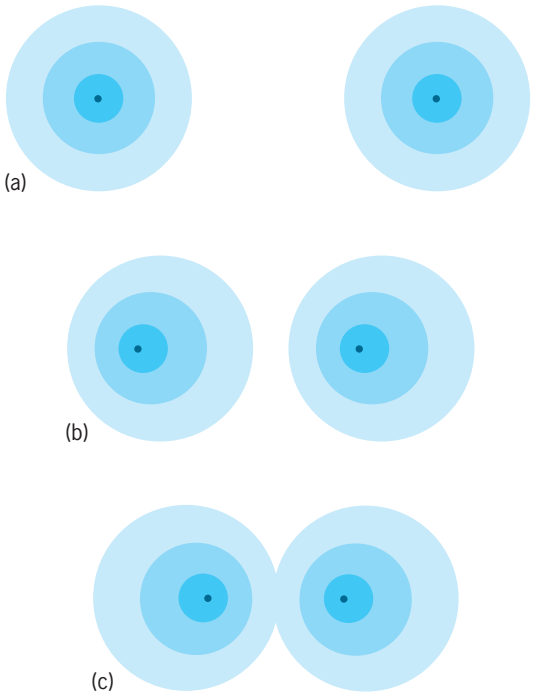

Fig. 2. Schematic diagram of intermolecular interaction. (*a*) There is negligible interaction between atoms that are 1 nm apart. (*b*) For atoms separated by about 0.8 nm or less, attractive dispersion forces result from correlated fluctuations of the electron charge distributions of the atoms. (*c*) For atoms closer than 0.3 nm or so, exchange effects cause a distortion of the charge distribution that leads to repulsion.

proportional to the inverse sixth power of the distance between the atoms. This London dispersion energy is larger for atoms and molecules whose electron densities fluctuate freely—in other words, those atoms and molecules that are easily polarized by external electric fields. *See* DIPOLE MOMENT; MOLECULAR ORBITAL THEORY.

At small separations the electron clouds can overlap, and repulsive forces arise. These forces are described as exchange–repulsion and are a consequence of the Pauli exclusion principle, a quantum-mechanical effect which prevents electrons from occupying the same region of space simultaneously (Fig. 2*c*). To accommodate it, electrons are squeezed out from the region between the nuclei, which repel each other as a result. Each element can be assigned, approximately, a characteristic van der Waals radius; that is, when atoms in different molecules approach more closely than the sum of their radii, the repulsion energy increases sharply. It is this effect that gives molecules their characteristic shape, leading to steric effects in chemical reactions. *See* EXCLUSION PRINCIPLE; STERIC EFFECT (CHEMISTRY).

The other important source of intermolecular forces is the electrostatic interaction. When molecules are formed from atoms, electrons flow from electropositive atoms to electronegative ones, so that the atoms become somewhat positively or negatively charged. In addition, the charge distribution of each atom may be distorted by the process of bond formation, leading to atomic dipole and quadrupole moments. The electrostatic interaction between these is an important source of intermolecular forces, especially in polar molecules, but also in molecules that are not normally thought of as highly polar. The interaction between two ethyne (acetylene) molecules, for example, is dominated by the quadrupole moment associated with the carbon-carbon triple bond. A further effect is that the electrostatic field of a molecule may cause polarization of its neighbors, and this leads to a further induction contribution to the intermolecular interaction. An important feature is that an induction interaction can often polarize both molecules in such a way as to favor interactions with further molecules, leading to a cooperative network of intermolecular attractions. This effect is important in the network structure of water and ice. *See* WATER.

An example of intermolecular forces is the hydrogen bond, in which a hydrogen atom attached to an electronegative atom such as oxygen or nitrogen is weakly bonded to another electronegative atom in the same molecule or a different molecule. The dominant attractive contribution in a hydrogen bond is the electrostatic interaction, but there are also significant contributions from dispersion and induction. *See* HYDROGEN BOND.

Occurrence. Intermolecular forces are responsible for many of the bulk properties of matter in all its phases. A realistic description of the relationship between pressure, volume, and temperature of a gas must include the effects of attractive and repulsive forces between molecules. The viscosity, diffusion, and surface tension of liquids are examples of physical properties which depend strongly on intermolecular forces. Intermolecular forces are also responsible for the ordered arrangement of molecules in solids, and account for their elasticity and properties (such as the velocity of sound in materials). The detailed description of the effects of intermolecular forces in condensed phases is complicated, since every molecule interacts not just with its immediate neighbors but with more distant ones too. *See* DIFFUSION; ELASTICITY; GAS; SURFACE TENSION; VAN DER WAALS EQUATION; VIRIAL EQUATION; VISCOSITY.

Intermolecular forces are important between atoms within a molecule. The repulsions between atoms constrain the possible three-dimensional structures that can be taken up by a large molecule such as a polymer or protein. Hydrogen bonds are especially important in biological molecules. The energy associated with a hydrogen bond is typically of the order of 10–20 kJ/mol, so such bonds are much weaker than normal chemical bonds. In a protein, however, there are many hydrogen bonds between the amino acid residues, and they are responsible for the stability of the alpha helix and other elements of protein structure. The twin strands of DNA are composed of adenine, cytosine, guanine, and thymine molecules, connected by a phosphate–sugar backbone. An adenine molecule in one strand can pair only with a thymine in the other, and cytosine can pair only with guanine. This selectivity, which is responsible for the duplication of the genetic code, arises from the pattern of hydrogen bonds. *See* DEOXYRIBONUCLEIC ACID (DNA); PROTEIN.

Atoms and molecules may be held to a solid surface by intermolecular forces. This weak bonding, called physisorption, has many important applications. The trapping of molecules from the gas phase onto cooled surfaces is the basis of pumps for producing high vacuums. Undesirable odors or colors in food or water may sometimes be removed by filters which capture the offending contamination by physisorption. The selective adsorption of molecules by surfaces is a useful method for separation of mixtures of molecules. Some solid catalysts function by attracting molecules from the gas phase onto the surface, where they can react with each other. *See* ADSORPTION; CATALYSIS; CHROMATOGRAPHY; VACUUM PUMP.

Experimental techniques. Because of their importance, intermolecular forces have been studied extensively for many years. Until the early 1970s, most of the information on intermolecular forces was inferred from the study of matter in bulk. For example, measurements of the viscosity of gases or the crystal structure of solids were used for this purpose. However, such properties depend rather indirectly on the intermolecular forces, and the conclusions of such studies were often ambiguous. However, the intermolecular potential energy curve shown in Fig. 1 was derived from experimental data on the viscosity, thermal conductivity, diffusion, and second pressure virial coefficient of argon gas, together with spectroscopic data on the argon dimer, Ar_2.

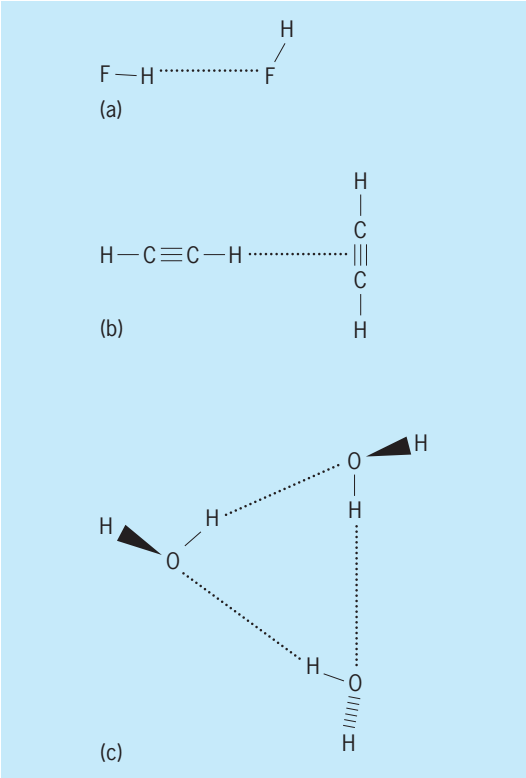

Fig. 3. Structures of some van der Waals molecules. (*a*) Hydrogen fluoride dimer. (*b*) Ethyne dimer. (*c*) Water trimer; two of the non-hydrogen-bonded H atoms are on one side of the plane of the O atoms, and the third is on the other.

Modern methods study interactions between individual molecules. In molecular-beam collision experiments, low-density streams of atoms or molecules are directed so that individual particles collide. The way in which the molecules rebound from the collision depends on their initial velocities, which can be controlled, and on the forces acting between them. Information about intermolecular forces can be extracted from the experimental data. *See* MOLECULAR BEAMS.

A method which has been increasingly important in recent years, with the development of sophisticated spectroscopic techniques using lasers, involves van der Waals molecules. A van der Waals molecule is a cluster of two or more atoms or molecules in which the molecules retain their individual identity and are held together by intermolecular forces. Such molecules are usually made by expanding a mixture of gases, usually with an inert carrier gas such as argon, into a high vacuum through a narrow nozzle. The resulting molecular beam contains van der Waals molecules, which can be studied spectroscopically. Very detailed information about intermolecular forces can be obtained in this way. In the case of the Ar-HCl cluster, for instance, the interaction energy is now known accurately and in detail as a function of the separation and orientation of the argon and hydrogen chloride components. Other van der Waals molecules have been studied in this way, including clusters of up to six water molecules, $(HF)_2$, $(C_2H_2)_2$, and $(C_2H_2)_3$. The structures of some van der Waals molecules are shown in **Fig. 3**. The T-shaped structure of ethyne dimer is a consequence of the quadrupole–quadrupole interaction mentioned above, while the bent structure of HF dimer is also attributed to electrostatic forces. Indeed, it has become apparent in recent decades that the electrostatic forces are dominant in determining the structures of van der Waals molecules, primarily because they are repulsive in some orientations and attractive in others, and so are much more orientation-dependent than other intermolecular forces. *See* LASER.

Many of these molecules can rearrange into equivalent but different structures. The HF dimer, for instance, can rearrange in the way shown in **Fig. 4**, the hydrogen atom in the hydrogen bond moving aside and the other taking its place. The symmetrical intermediate structure is only about 4 kJ/mol higher in energy than the initial and final structures, and tunneling can occur between them, leading to splittings in the spectrum that can be measured. These tunneling splittings provide further information about the nature of the intermolecular potential energy surface. *See* MOLECULAR STRUCTURE AND SPECTRA.

Computational methods. An important source of information about intermolecular potential energy surfaces is quantum theory. The underlying general theory of intermolecular forces has been understood since the 1930s, but only in recent years has it become possible to explore it quantitatively and to understand its finer details. Calculations ab initio, starting from Schrödinger's equation without any empirical information and using quantum-mechanical perturbation theory, can provide accurate details of all features of the forces between molecules, though this approach is limited to relatively small molecules. In this way, it is possible to obtain a deeper

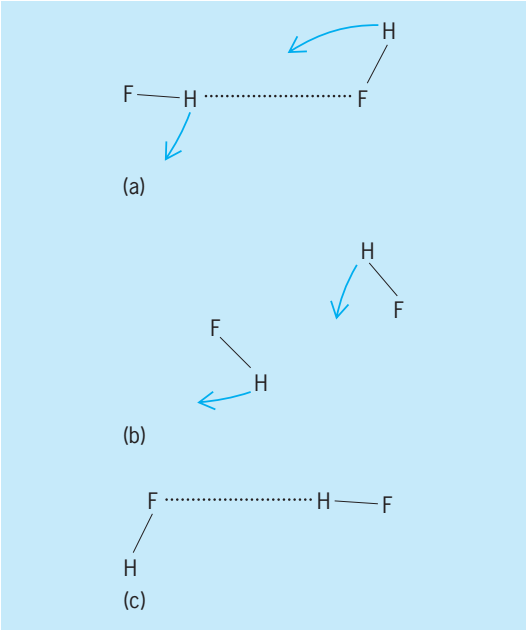

Fig. 4. Donor–acceptor exchange in HF dimer.

understanding of the quantum-mechanical origin of intermolecular forces and to calculate them to an accuracy comparable with experiment. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM CHEMISTRY.

Anthony J. Stone

Bibliography. J. Israelachvili, *Intermolecular and Surface Forces*, Academic Press, 1992; J. S. Rowlinson, *Cohesion: A Scientific History of Intermolecular Forces*, Cambridge University Press, 2002; A. J. Stone, *The Theory of Intermolecular Forces*, Oxford University Press, 1996; D. J. Wales, *Energy Landscapes*, Cambridge University Press, 2003.

Internal combustion engine

A prime mover that burns fuel inside the engine, in contrast to an external combustion engine, such as a steam engine, which burns fuel in a separate furnace. *See* ENGINE.

Most internal combustion engines are spark-ignition gasoline-fueled piston engines. These are used in automobiles, light- and medium-duty trucks, motorcycles, motorboats, lawn and garden equipment, and light industrial and portable power applications. Diesel engines are used in automobiles, trucks, buses, tractors, earthmoving equipment, as well as marine, power-generating, and heavier industrial and stationary applications. This article describes these types of engines. For other types of internal combustion engines *see* GAS TURBINE; ROCKET PROPULSION; TURBINE PROPULSION.

The aircraft piston engine is fundamentally the same as that used in automobiles but is engineered for light weight and is usually air-cooled. *See* RECIPROCATING AIRCRAFT ENGINE.

Engine types. Characteristics common to all commercially successful internal combustion engines include (1) the compression of air, (2) the raising of air temperature by the combustion of fuel in this air at its elevated pressure, (3) the extraction of work from the heated air by expansion to the initial pressure, and (4) exhaust.

Four-stroke cycle. William Barnett first drew attention to the theoretical advantages of combustion under compression in 1838. In 1862 Beau de Rochas published a treatise that emphasized the value of combustion under pressure and a high ratio of expansion for fuel economy; he proposed the four-stroke engine cycle as a means of accomplishing these conditions in a piston engine (**Fig. 1**). The engine requires two revolutions of the crankshaft to complete one combustion cycle. The first engine to use this cycle successfully was built in 1876 by N. A. Otto. *See* OTTO CYCLE.

Otto's engine, like almost all internal combustion engines developed at that period, burned coal gas mixed in combustible proportions with air prior to being drawn into the cylinder. The engine load was generally controlled by throttling the quantity of charge taken into the cylinder. Ignition was by a device such as an external flame or an electric spark,

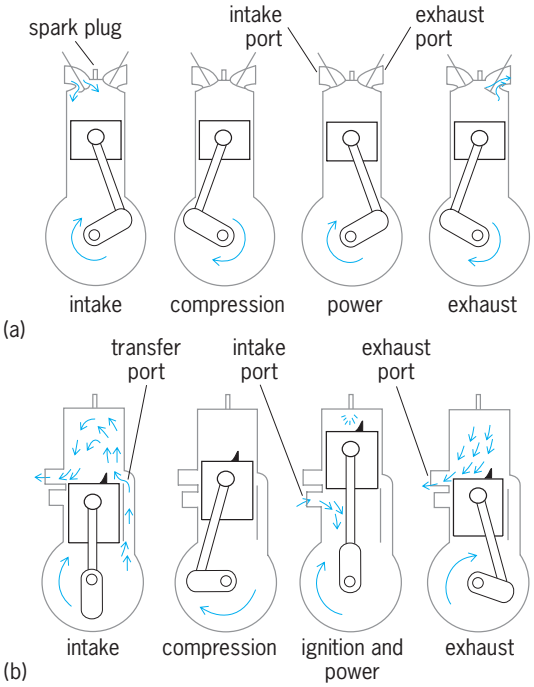

Fig. 1. Engine cycles (*a*) The four strokes of a four-stroke engine cycle. On intake stroke, the intake valve (left) has opened and the piston is moving downward, drawing air and gasoline vapor into the cylinder. On compression stroke, the intake valve has closed and the piston is moving upward, compressing the mixture. On power stroke, the ignition system produces a spark that ignites the mixture. As it burns, high pressure is created, which pushes the piston downward. On exhaust stroke, the exhaust valve (right) has opened and the piston is moving upward, forcing the burned gases from the cylinder. (*b*) Three-port two-cycle engine. The same action is accomplished without separate valves and in a single rotation of the crankshaft.

so that the timing was controllable. These are essential features of what has become known as the Otto or spark-ignition combustion cycle.

Two-stroke cycle. In 1878 Dougald Clerk developed the two-stroke engine cycle by which a similar combustion cycle required only one revolution of the crankshaft. In this cycle, exhaust ports in the cylinder were uncovered by the piston as it approached the end of its power stroke. A second cylinder then pumped a charge of air to the working cylinder through a check valve when the pump pressure exceeded that in the working cylinder.

In 1891 Joseph Day simplified the two-stroke engine cycle by using the crankcase to pump the required air. The compression stroke of the working piston draws the fresh combustible charge through a check valve into the crankcase, and the next power stroke of the piston compresses this charge. The piston uncovers the exhaust ports near the end of the power stroke and slightly later uncovers intake ports opposite them to admit the compressed charge from the crankcase. A baffle is usually provided on the piston head of small engines to deflect the charge up one side of the cylinder to scavenge the remaining burned gases down the other side and out the exhaust ports with as little mixing as possible.

Modern engines using this two-stroke cycle have a third cylinder port known as the transfer port (Fig. 1*b*), instead of the crankcase check valve used by Day. Small engines of this type are widely used where fuel economy is not as important as mechanical simplicity and light weight. They do not need mechanically operated valves, and they develop one power impulse per cylinder for each crankshaft revolution.

Two-stroke-cycle engines do not develop twice the power of four-stroke-cycle engines with the same size of working cylinders at the same number of revolutions per minute (rpm). Principal reasons are (1) reduction in effective cylinder volume due to the piston movement required to cover exhaust ports; (2) appreciable mixing of burned (exhaust) gases with the combustible mixture; and (3) loss of some combustible mixture through the exhaust ports with the exhaust gases.

Otto-cycle engines. In the idealized four-stroke Otto cycle, combustion is instantaneous and at constant volume. This simplifies thermodynamic analysis, but combustion takes time. Gas pressure during the four strokes of the Otto cycle varies with the piston position as shown by the typical indicator card in **Fig. 2***a*. This is a pressure–volume (PV) card for an 8.7:1 compression ratio.

Engine power. To simplify calculations of engine power, the average net pressure during the working stroke, known as the mean effective pressure (mep), is frequently used. It may be obtained from the average net height of the card, which is found by measurement of the area and then division of this area by its length. Similar pressure–volume data may be plotted on logarithmic coordinates as in Fig. 2*b*, which develops expansion and compression relations as approximately straight lines. The slopes show the values of exponent n to use in equations for PV relationships.

The rounding of the plots at peak pressure, with the peak developing after the piston has started its power stroke, even with the spark occurring before the piston reaches the end of the compression stroke, is due to the time required for combustion. Changes in design can vary charge turbulence in the compression space prior to and during combustion. The greater the turbulence, the faster the combustion and the lower the antiknock or octane number required of the fuel, or the higher the compression ratio that may be used with a given fuel without knocking. The amount to which the turbulence can be raised is limited by the increased rate of pressure rise, which increases engine roughness. This must not exceed a level acceptable for automobile or other service. *See* AUTOMOBILE; AUTOMOTIVE ENGINE; COMBUSTION CHAMBER; COMPRESSION RATIO; MEAN EFFECTIVE PRESSURE; OCTANE NUMBER; SPARK KNOCK.

Detonation of a small part of the charge in the cylinder, after most of the charge has burned progressively, causes knock. This limits the compression ratio of an engine with a given fuel.

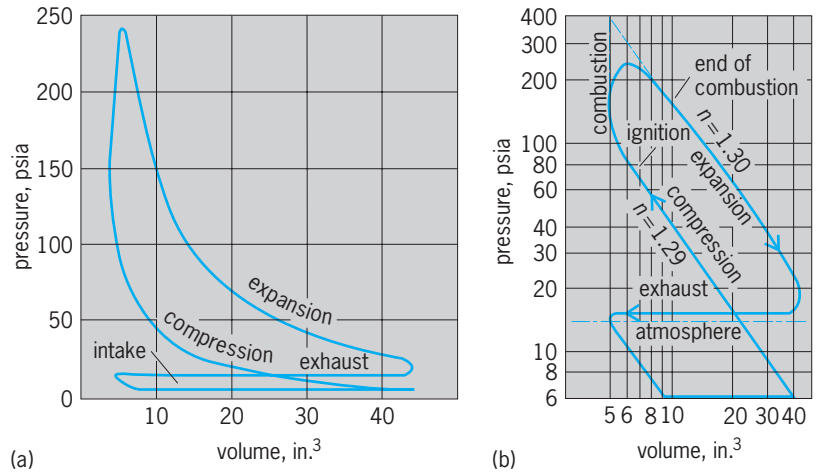

Fig. 2. Typical Otto-cycle pressure–volume indicator card plotted on (*a*) rectangular coordinates and (*b*) logarithmic coordinates.

Compression ratio. According to classical thermodynamic theory, thermal efficiency η of the Otto combustion cycle is given by Eq. (1), where the com-

$$\eta = 1 - \frac{1}{r^{n-1}} \tag{1}$$

pression ratio r_c and expansion ratio r_e are the same ($r_c = r_e = r$). When theory assumes atmospheric air in the cylinder for extreme simplicity, exponent n is 1.4. Efficiencies calculated on this basis are almost twice as high as measured efficiencies. Logarithmic diagrams from experimental data show that n is about 1.3 (Fig. 2*b*). Even with this value, efficiencies achieved in practice are less than given by Eq. (1), because it assumes instantaneous combustion and 100% volumetric efficiency. This exponent should vary with the fuel-air mixture ratio, and to some extent with the compression ratio. For an 8:1 compression ratio, the exponent should vary from about 1.28 for a stoichiometric (chemically correct) mixture to about 1.31 for a lean mixture. Actual practice gives even lower thermal efficiencies. This is because of the assumed instantaneous changes in cyclic pressure (during combustion and exhaust) and the disregard of heat losses to the cylinder walls.

A change in compression ratio causes little change in the mechanical efficiency, or the volumetric efficiency resulting from raising the compression ratio provides a corresponding increase in torque or mean effective pressure. This is frequently of more practical importance than the actual efficiency increase. *See* THERMODYNAMIC CYCLE.

Engine load has little effect on indicated thermal efficiency, provided the fuel–air ratio remains constant and the ignition time is suitably advanced at reduced loads. This compensates for the slower rate of burning that results from dilution of the combustible charge with the larger percentages of burned gases remaining in the combustion space and the reduced turbulence at lower speeds.

High compression improves fuel economy because of improved thermal efficiency. However, the increased peak combustion temperature increases

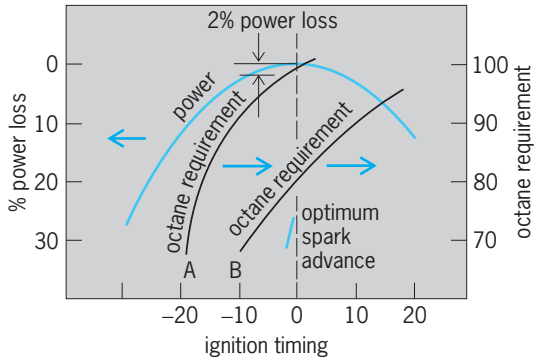

Fig. 3. Effects of advancing or retarding ignition timing from optimum on engine power and resulting octane requirement of fuel in an experimental engine with a combustion chamber having typical turbulence (A) and a highly turbulent design (B) with the same compression ratio. Retarding the spark 7° for 2% power loss reduced octane requirement from 98 to 93 for design A.

emissions of oxides of nitrogen in the exhaust gas. *See* AIR POLLUTION; SMOG.

Ignition timing. High thermal efficiency is obtained from high compression ratios at part loads, where engines normally run at automobile cruising speeds, with optimum spark advance. To avoid knock on available gasolines at wide-open throttle, a reduced or compromise spark advance is used. The tendency of an engine to knock at wide-open throttle is reduced appreciably when the spark timing is reduced 5–10° from optimum (**Fig. 3**). Advancing or retarding the spark timing from optimum results in an increasing loss in mean effective pressure for any normal engine, as shown by the heavy curve in Fig. 3. The octane requirement falls rapidly as the spark timing is retarded, the actual rate depending on the nature of the gasoline as well as on the combustion chamber design. Curves A and B show the effects on a given gasoline of the use of moderate- and high-turbulence combustion chambers, respectively, with the same compression ratio. Because the curve for mean effective pressure is relatively flat near optimum spark advance, retarding the spark for a 1–2% loss is normally acceptable because of the reduction in octane requirement.

In addition to the advantages of the higher compression ratio at cruising loads with optimum spark advance, the compromise spark at full load may be advanced toward optimum with higher-octane fuels. Then there is a corresponding increase in full-throttle mean effective pressure.

Many automotive engines have an electronic engine control system with an onboard microprocessor that controls spark timing electronically. The microprocessor continuously adjusts ignition timing for optimum fuel economy and drivability, while minimizing exhaust emissions. The microprocessor may be capable of retarding the timing if spark knock occurs. This allows the benefits of using a higher compression ratio and gasoline with a lower octane number without danger of engine-damaging detonation. In some systems, a digital map stored in memory provides a wide range of predetermined ignition settings (**Fig. 4**). *See* CONTROL SYSTEMS; IGNITION SYSTEM; MICROPROCESSOR.

Fuel–air ratio. A fuel–air mixture richer than that which develops maximum knock-free mep will permit use of higher compression ratios. However, the benefits derived from compromise or rich mixtures vary so much with mixture temperature and the sensitivity of the octane value of the particular fuel to temperature that this method is not generally practical. Nevertheless, piston-type aircraft engines may use fuel–air mixture ratios of 0.11 or even higher during takeoff, instead of about 0.08, which normally develops maximum mep in the absence of knock.

In automotive engines with an electronic engine control system, the microprocessor usually controls the amount of fuel delivered by either a feedback carburetor or a single-point (throttle-body) or a multipoint (port) electronic fuel-injection system. This maintains the mixture at or near the stoichiometric ratio, which minimizes exhaust emissions of hydrocarbons, carbon monoxide, and oxides of nitrogen. However, spark-ignition engines deliver maximum power with an air deficiency of 0–10%, and minimum fuel consumption with about 10% excess air (**Fig. 5**).

Stroke–bore ratio. The ratio of the length of the piston stroke to the diameter of the cylinder bore has no appreciable effect on fuel economy and friction at corresponding piston speeds. Practical advantages that result from the short stroke include the greater rigidity of crankshaft from the shorter crank cheeks, with crankpins sometimes overlapping main bearings, and the narrower as well as lighter cylinder block that is possible. However, the higher rates of crankshaft rotation for an equivalent piston speed necessitate greater valve forces and require stronger valve springs. Also, the smaller depth of the compression space for a given compression ratio increases the surface-to-volume ratio and the proportion of heat lost by radiation during combustion. In automotive engines, stroke–bore ratios have decreased over the years.

Valve timing. The times of opening and closing the valves of an engine in relation to piston position are usually selected to develop maximum power over a desired speed range at wide-open throttle. The

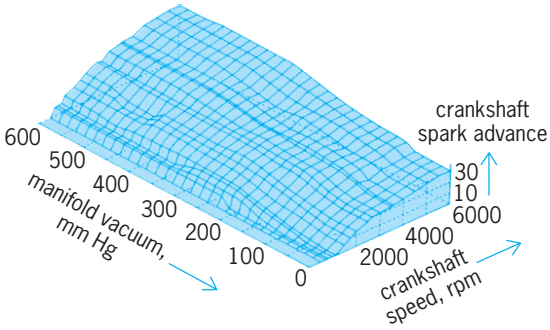

Fig. 4. Three-dimensional ignition map showing 576 timing points stored in the memory of a microprocessor system for a four-cylinder automotive engine. 1 mmHg = 133 Pa. (*Ford Motor Co.*)

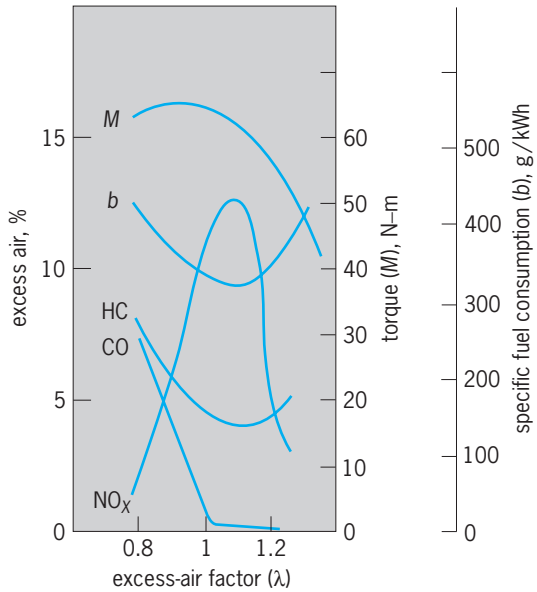

Fig. 5. Graph showing how the excess-air factor affects exhaust-gas [composition carbon monoxide (CO), nitrogen oxides (NO_x), and hydrocarbons (HC)], torque (*M*), and specific fuel consumption (*b*) in the part-load range of an automotive spark-ignition engine running at a constant midrange speed and cylinder charge. (*Robert Bosch Corp.*)

timing of these events is usually expressed as the number of degrees of crankshaft rotation before or after the piston reaches the end of one of its strokes.

Because of the time required for the burned gas to flow through the exhaust valve at the end of the power stroke of a piston, the exhaust valve usually starts opening considerably before the end of the stroke. If the valve opens when the piston is nearer the lower end of its stroke, power is lost at high engine speeds because the piston on its exhaust stroke has to move against gas pressure remaining in the cylinder. If the valve opens before necessary, the burned gas is released while it is still at sufficient pressure to increase the work done on the piston.

For any engine, there is an optimum time for opening the exhaust valve that will develop the maximum power at some particular speed. The power loss at other speeds does not increase rapidly. Therefore when an engine is throttled at part load, there is less gas to discharge through the exhaust valve and less need for the valve to be opened as early as at wide-open throttle.

The timing of intake valve events is normally selected to trap the largest possible quantity of combustible mixture (air in a diesel engine) in the cylinder when the valve closes at some desired engine speed and at wide-open throttle. The intermittent flow through the intake valve undergoes alternate accelerations and decelerations, which require time. During the intake stroke, the mass of air moving through the passage to the intake valve is given velocity energy that may be converted to a slight pressure at the valve when the air mass still in the passage is stopped by its closure. Advantage of this phenomenon may be obtained at some engine speed to increase the air mass which enters the cylinder.

The engine speed at which the maximum volumetric efficiency is developed varies with the relative valve area, closure time, and other factors, including the diameter and length of the passage. The curves in **Fig. 6** show the characteristic falling off at high speeds from the inevitable throttling action as air flows at increased velocities through any restriction such as a valve or intake passage and venturi of a carburetor.

Volumetric efficiency has a direct effect on the mean effective pressure developed in a cylinder, on the torque, and on the power that may be realized at a given speed. Since power is a product of speed and torque, the peak power of an engine occurs at a higher speed than for maximum torque, where the rate of torque loss with any further increase in speed will exceed the rate of speed increase. An engine develops maximum power at a speed about twice that for maximum torque.

To obtain maximum torque and power, intake-valve closing may be delayed until the piston has traveled almost half the length of the compression stroke. At engine speeds below those where maximum torque is developed by this valve timing, some of the combustible charge that has been drawn into the cylinder on the intake stroke will be driven back through the intake valve before it closes. This reduces the effective compression ratio at wide-open throttle. The engine has an increasing tendency to develop spark knock as the speed and the resulting gas turbulence are reduced. *See* ENGINE MANIFOLD.

Supercharging spark-ignition engines. Volumetric efficiency and thus the mep of a four-stroke spark-ignition engine may be increased over a part of or the whole speed range by supplying air to the engine intake at higher than atmospheric pressure. This is usually accomplished by a centrifugal or rotary pump. The indicated power (power developed in the cylinder) of an engine increases directly with the absolute pressure in the intake manifold. Because fuel consumption increases at the same rate, the indicated specific fuel consumption (fuel flow rate per unit

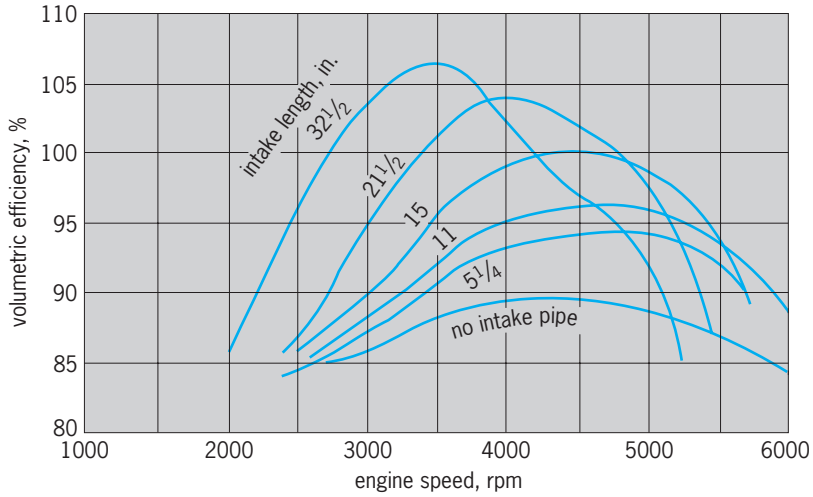

Fig. 6. Effect of intake-pipe length and engine speed on volumetric efficiency of one cylinder of a six-cylinder engine. 1 in. = 2.5 cm.

power output) is generally not altered appreciably by supercharging.

The three principal reasons for supercharging four-cycle spark-ignition engines are (1) to lessen the tapering off of mep at higher engine speed; (2) to prevent loss of power due to diminished atmospheric density, as when an airplane (with piston engines) climbs to high altitudes; and (3) to develop more torque at all speeds.

In a normal engine characteristic, torque rises as speed increases but falls off at higher speeds because of the throttling effects of such parts of the fuel intake system as valves and carburetors. If a supercharger is installed so as to maintain the volumetric efficiency at the higher speeds without increasing it in the middle-speed range, peak horsepower can be increased.

The rapid fall of atmospheric pressure at increased altitudes causes a corresponding decrease in the power of unsupercharged piston-type aircraft engines. For example, at 20,000 ft (6 km) the air density, and thus the absolute manifold pressure and indicated torque of an aircraft engine, would be only about half as great as at sea level. The useful power developed would be still less because of the friction and other mechanical power losses which are not affected appreciably by volumetric efficiency. By the use of superchargers, which are usually of the centrifugal type, sea-level air density may be maintained in the intake manifold up to considerable altitudes. Some aircraft engines drive these superchargers through gearing which may be changed in flight, from about 6.5 to 8.5 times engine speed. The speed change avoids oversupercharging at medium altitudes with corresponding power loss. Supercharged aircraft engines must be throttled at sea level to avoid damage from detonation or excessive overheating caused by the high mep which would otherwise be developed. *See* SUPERCHARGER.

Normally an engine is designed with the highest compression ratio allowable without knock from the fuel expected to be used. This is desirable for the highest attainable mep and fuel economy from an atmospheric air supply. Any increase in the volumetric efficiency of such an engine would cause it to knock unless a fuel of higher octane number were used or the compression ratio were lowered. When the compression ratio is lowered, the knock-limited mep may be raised appreciably by supercharging but at the expense of lowered thermal efficiency. There are engine uses where power is more important than fuel economy, and supercharging becomes a solution. The principle involved is illustrated in **Fig. 7** for a given engine. With no supercharge this engine, when using 93-octane fuel, developed an indicated mean effective pressure (imep; an average pressure forcing the piston down the cylinder) of 180 pounds per square inch (psi; 1240 kilopascals) at the borderline of knock at 8:1 compression ratio. If the compression ratio were lowered to 7:1, the mep could be raised by supercharging along the 7:1 curve to 275 imep before it would be knock-limited by the same fuel. With a 5:1 compression ratio it could be

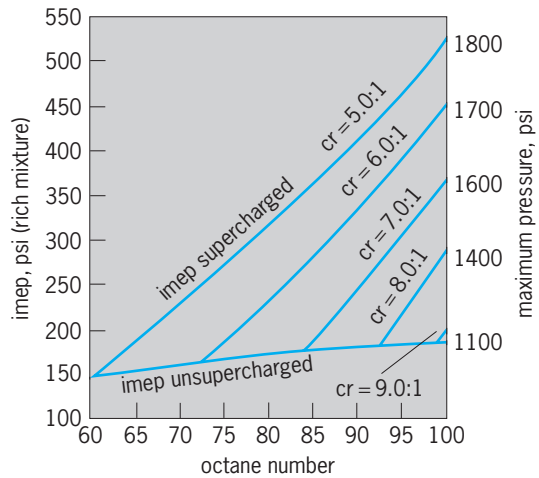

Fig. 7. Graph showing the relationship between compression ratio and knock-limited imep for given octane numbers, obtained by supercharging a laboratory engine. (*After H. R. Ricardo, The High-Speed Combustion Engine, 4th ed., Blackie, 1953*)

raised to 435 imep. Thus the imep could be raised until the cylinder became thermally limited by the temperatures of critical parts, particularly of the piston head.

Engine balance. Rotating masses such as crank pins and the lower half of a connecting rod may be counterbalanced by weights attached to the crankshaft. The vibration which would result from the reciprocating forces of the pistons and their associated masses is usually minimized or eliminated by the arrangement of cylinders in a multicylinder engine so that the reciprocating forces in one cylinder are neutralized by those in another. Where these forces are in different planes, a corresponding pair of cylinders is required to counteract the resulting rocking couple.

If piston motion were truly harmonic, which would require a connecting rod of infinite length, the reciprocating inertia force at each end of the stroke would be as in Eq. (2), where W is the total weight

$$F = 0.000456WN^2s \qquad (2)$$

of the reciprocating parts in one cylinder, N is the rpm, and s is the stroke in inches. Both F and W are in pounds. But the piston motion is not simple harmonic because the connecting rod is not infinite in length, and the piston travels more than half its stroke when the crankpin turns 90° from firing dead center. This distortion of the true harmonic motion is due to the so-called angularity a of the connecting rod, shown by Eq. (3), where r is the crank radius,

$$a = \frac{r}{l} = \frac{s}{2l} \qquad (3)$$

s the stroke, and l the connecting rod length, all in inches.

Reciprocating inertia forces act in line with the cylinder axis and may be considered as combinations of a primary force—the true harmonic force from Eq. (2)—oscillating at the same frequency as the crankshaft rpm and a secondary force oscillating

at twice this frequency having a value of Fa, which is added to the primary at top dead center and subtracted from it at bottom dead center. In general, harmonics above the second order may be disregarded. Therefore, for a connecting rod with the angularity $a = 0.291$, the inertia force caused by a piston at top dead center is about 1.29 times the pure harmonic force, and at bottom dead center it is about 0.71 times as large.

Where two pistons act on one crankpin, with the cylinders in 90° V arrangements, the resultant primary force is radial and of constant magnitude, and it rotates around the crankshaft with the crankpin. Therefore, it may be compensated for by an addition to the weight required to counterbalance the centrifugal force of the revolving crankpin and its associated masses. The resultant of the secondary force of the two pistons is 1.41 times as large as for one cylinder, and reciprocates in a horizontal plane through the crankshaft at twice crankshaft speed.

In four-cylinder inline engines with crankpins in the same plane, the primary reciprocating forces of the two inner pistons in cylinders 2 and 3 cancel those of the two outer pistons in cylinders 1 and 4, but the secondary forces from all pistons are added. Therefore, they are equivalent to the force resulting from a weight about $4a$ times the weight of one piston and its share of the connecting rod, oscillating parallel to the piston movement, having the same stroke, but moving at twice the frequency. A large a for this type of engine is advantageous. Where the four cylinders are arranged alternately on each side of a similar crankshaft, and in the same plane, both primary and secondary forces are in balance. Six cylinders in line also balance both primary and secondary forces.

V-8 engines use a crank arrangement in which the crankpins are in two planes 90° apart. Staggering the crankpins for pistons 1 and 2 90° from each other equalizes secondary forces, but the forces are in different planes. The couple this introduces is canceled by an opposite couple from the pistons operating on the crankpins for pistons 3 and 4.

Torsion dampers. In addition to vibrational forces from rotating and reciprocating masses, vibration may develop from torsional resonance of the crankshaft at various critical speeds. The longer the shaft for given bearing diameters, the lower the speeds at which these vibrations develop. On automotive engines, such vibrations are dampened by a viscous vibration damper or by a bonded-rubber vibration damper that is similar to a small flywheel coupled to the crankshaft through a rubber ring. The vibration damper may be combined with the pulley for an engine-accessory drive belt. *See* MECHANICAL VIBRATION.

Firing order. The firing order is the sequence in which the cylinders deliver their power impulses to the crankshaft. It is determined by such factors as engine design, ignition intervals, and crankshaft loading. Cylinder arrangements are generally selected for even firing intervals and torque impulses, as well as for balance.

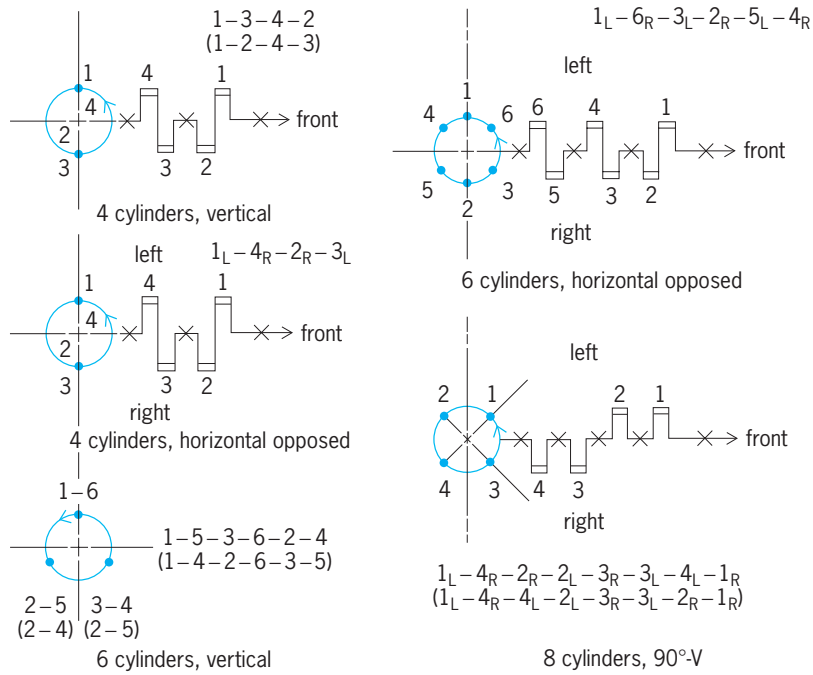

Fig. 8. Various cylinder arrangements and firing orders.

Figure 8 shows various cylinder arrangements and firing orders that have been used in automotive engines. The Society of Automotive Engineers (SAE) Standard for engine rotation and cylinder numbering provides that standard rotation is counterclockwise rotation of the crankshaft as viewed from the principal output end of the engine. If power can be delivered from either end, rotation shall be as viewed from the flywheel end.

Excluding radial engines or those with coplanar cylinder bore axes, cylinders may be numbered by either of two methods: (1) In single or multibank engines, the cylinders are numbered in the sequence in which the connecting rods are mounted along the crankshaft, beginning with the cylinder farthest from the principal output end. (2) In multibank engines, the cylinders may be numbered in sequence in each bank, starting with the cylinder farthest from the principal output end and designated right or left bank by suffixes R and L, such as 1R and 1L.

Cylinder bank and accessory locations are described as right or left when the engine is viewed from the flywheel or principal output end.

Compression-ignition engines.. In 1897, about 20 years after Otto first ran his engine, Rudolf Diesel successfully demonstrated an entirely different method of igniting fuel. Air is compressed to a pressure high enough for the adiabatic temperature to reach or exceed the ignition temperature of the fuel. Because this temperature is 1000°F (538°C) or higher, compression ratios of 12:1 to 23:1 are used commercially with compression pressures from about 440 to 800 psi (3 to 5.5 megapascals). The fuel is injected into the cylinders shortly before the end of the compression stroke, at a time and rate suitable to control the rate of combustion.

Compression ratio and combustion. The idealized diesel engine cycle assumes combustion at constant pressure. Like the Otto cycle, thermal efficiency increases with compression ratio, but also varies with the amount of heat added (at the constant pressure) up to the cutoff point where the pressure begins to drop from adiabatic expansion. *See* DIESEL CYCLE; DIESEL ENGINE.

Fuel injection. Early diesel engines used air injection of the fuel to develop extremely fine atomization and a good distribution of the spray. But the need for injection air at pressures of about 1500 psi (10 MPa) required expensive and bulky multistage air compressors and intercoolers.

A simpler fuel-injection method was introduced by James McKechnie in 1910. He atomized the fuel as it entered the cylinder by use of high fuel pressure and suitable spray nozzles. After considerable development, it became possible to atomize the fuel sufficiently to minimize the smoky exhaust that had been characteristic of the early airless or solid-injection engines. By 1930, solid injection had become the generally accepted method of injecting fuel in diesel engines.

During the 1980s, electronically controlled fuel injection began replacing the mechanical system. Eletronically controlled mechanically actuated unit injectors allowed injection pressure of 22,000 psi (150 MPa). *See* FUEL INJECTION.

Supercharged diesel engines. Combustion in a four-stroke diesel engine is improved by supercharging. Fuels that would smoke heavily and misfire at low loads will burn otherwise satisfactorily with supercharging. The indicated mean effective pressure rises directly with the supercharging pressure, until it is limited by the rate of heat flow from the metal parts surrounding the combustion chamber, and the resulting temperatures.

When superchargers of either the centrifugal or positive-displacement type are driven mechanically by the engine, the power required becomes an additional loss to the engine output. There is a degree of supercharge for any engine that develops maximum efficiency. A supercharge that is too high absorbs more power in the supercharger than is gained by the engine, especially at low loads. Another means of driving the supercharger is by an exhaust turbine, which recovers some of the energy that would otherwise be wasted in the exhaust. This may be accomplished with so small an increase of back pressure that little power is lost by the engine. The result is an appreciable increase in efficiency at loads high enough to develop the necessary exhaust pressure. *See* TURBOCHARGER.

Supercharging a two-cycle diesel engine requires some means of restricting or throttling the exhaust to build up cylinder pressure at the start of the compression stroke, and is used on a few large engines. Most medium and large two-stroke diesel engines are usually equipped with blowers to scavenge the cylinders after the working stroke and to supply the air required for the subsequent cycles. These blowers, in contrast to superchargers, do not build up appreciable pressure in the cylinder at the start of compression. If the capacity of such a blower is greater than the engine displacement, it will scavenge the cylinder of practically all exhaust products, even to the extent of blowing some air out through the exhaust ports. Such blowers, like superchargers, may be driven by the engine or by exhaust turbines.

Contrast between diesel and Otto engines. There are many characteristics of the diesel engine which are in direct contrast to those of the Otto engine. The higher the compression ratio of a diesel engine, the less the difficulties with ignition time lag. Too great an ignition lag results in a sudden and undesired pressure rise which causes an audible knock. In contrast to an Otto engine, knock in a diesel engine can be reduced by use of a fuel of higher cetane number, which is equivalent to a lower octane number. *See* CETANE NUMBER.

The larger the cylinder diameter of a diesel engine, the simpler the development of good combustion. In contrast, the smaller the cylinder diameter of the Otto engine, the less the limitation from detonation of the fuel.

High intake-air temperature and density materially aid combustion in a diesel engine, especially of fuels having low volatility and high viscosity. Some engines have not performed properly on heavy fuel until provided with a supercharger. The added compression of the supercharger raised the temperature and, what is more important, the density of the combustion air. For an Otto engine, an increase in either the air temperature or density increases the tendency of the engine to knock and therefore reduces the allowable compression ratio.

Diesel engines develop increasingly higher indicated thermal efficiency at reduced loads because of leaner fuel-air ratios and earlier cutoff. Such mixture ratios may be leaner than will ignite in an Otto engine. Furthermore, the reduction of load in an Otto engine requires throttling, which develops increasing pumping losses in the intake system. Neil MacCoull; Donald L. Anglin

Bibliography. *Bosch Automotive Handbook*, 1986; W. H. Crouse, *Automotive Engines*, 8th ed., 1994; J. B. Heywood, *Internal Combustion Engine Fundamentals*, 1988; L. C. R. Lilly (ed.), *Diesel Engine Reference Book*, 1984; K. Newton, W. Steeds, and T. K. Garrett, *The Motor Vehicle*, 1989; Society of Automotive Engineers, *SAE Handbook*, 4 vols., annually; C. F. Taylor, *The Internal Combustion Engine in Theory and Practice*, vols. 1 and 2, 1985.

Internal energy

A characteristic property of the state of a thermodynamic system, introduced in the first law of thermodynamics. For a static, closed system (no bulk motion, no transfer of matter across its boundaries), the change ΔU in internal energy for a process is equal to the heat Q absorbed by the system from its surroundings minus the work W done by the system on

its surroundings. Only a change in internal energy can be measured, not its value for any single state. For a given process, the change in internal energy is fixed by the initial and final states and is independent of the path by which the change in state is accomplished.

The internal energy includes the intrinsic energies of the individual molecules of which the system is composed and contributions from the interactions among them. It does not include contributions from the potential energy or kinetic energy of the system as a whole; these changes must be accounted for explicitly in the treatment of flow systems. Because it is more convenient to use an independent variable (the pressure P for the system instead of its volume V), the working equations of practical thermodynamics are usually written in terms of such functions as the enthalpy $H = U + PV$, instead of the internal energy itself. *See* CHEMICAL THERMODYNAMICS; ENTHALPY.

Paul J. Bender

International Date Line

The 180° meridian, where each day officially begins and ends. As a person travels eastward, against the apparent movement of the Sun, 1 h is gained for every 15° of longitude; traveling westward, time is lost at the same rate. Two people starting from any meridian and traveling around the world in opposite directions at the same speed would have the same time when they meet, but would be 1 day apart in date. If there were no international agreement as to where each day should begin and end, there could be any number of places so designated. To eliminate such confusion, the International Meridian Conference, in 1884, designated the 180° meridian as the location for the beginning of each day. Thus, when a traveler goes west across the line, a day is lost; if it is Monday to the east, it will be Tuesday immediately as the traveler crosses the International Date Line. In traveling eastward a day is gained; if it is Monday to the west of the line, it will be Sunday immediately after the line is crossed.

An interesting example can be taken from conditions now nearly attainable with jet aircraft. If one could board such a plane, say at noon in Washington, D.C., and fly westward at that latitude around the world to return in 24 h, the rate would match the rotation of the Earth. Although constantly under the noontime Sun, this traveler would need to adjust her calendar 1 day ahead upon crossing the International Date Line, because she would arrive in Washington at noon, 24 h after embarking. Thus her calendar day would agree with that of the Washingtonians.

The 180° meridian is ideal for serving as the International Date Line (see **illus.**). It is exactly halfway around the world from the zero, or Greenwich, meridian, from which all longitude is reckoned. It also falls almost in the center of the largest ocean; consequently there is the least amount of inconvenience as regards population centers. A few deviations in the alignment have been made, such as swinging the line east around Siberia to keep that area all in the same day, and westward around the Aleutian Islands so that they will be within the same day as the rest of Alaska. Other minor variations for the same purpose have been made near the Fiji Islands, in the South Pacific. *See* MATHEMATICAL GEOGRAPHY.

Van H. English

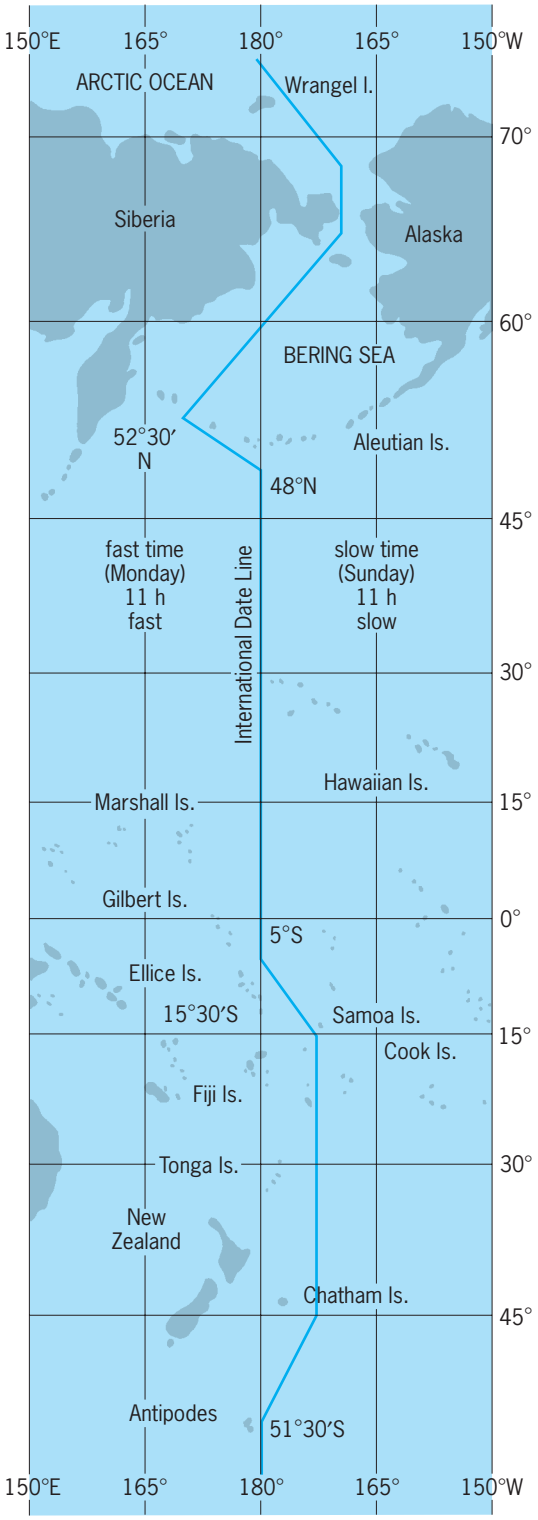

The International Date Line.

Internet

A worldwide system of interconnected computer networks. The Internet has begun to transform the way that communication systems operate. The origins of the Internet can be traced to the creation of ARPANET (Advanced Research Projects Agency Network) as a network of computers under the auspices of the U.S. Department of Defense in 1969. Today, the Internet connects millions of computers around the world in a nonhierarchical manner unprecedented in the history of communications. The Internet is a product of the convergence of media, computers, and telecommunications. It is not merely a technological development but the product of social and political processes as well, involving both the academic world and the government (the Department of Defense). From its origins in a nonindustrial, noncorporate environment and in a purely scientific culture, it has quickly diffused into the world of commerce. While the Internet has had a sudden and dramatic impact on the global economic and social order, it took almost 30 years to emerge as a major technological force (*see* **table**).

The Internet is a combination of several media technologies and an electronic version of newspapers, magazines, books, catalogs, bulletin boards, and much more. This versatility gives the Internet its power. However, it is difficult to make precise predictions about the success of the Internet because of the complicated relationships it has created among technologies, markets, and political systems. Contributing to the development of the Internet were a wide variety of users who adopted the Internet to communicate among themselves in a semiprivate way. These users were not bound by any customs and practices, but initiated a culture of communication that was flexible, quick, unbounded, and nonhierarchical, where there was no superior authority in control of the communication system and no standards to govern performance.

The Internet has grown dramatically since about 1988, although its impact was not felt until about 1995. By mid-2000, estimates of the number of connected computers were around 80 million.

The average user gets an Internet connection with the help of an Internet service provider (ISP). An Internet service provider is a commercial organization that provides an Internet dial-up account for the user. In order to utilize the service, the user needs to have a dial-up program, an e-mail program, and a Web browser program. These generally come as a standard package with the purchase of a computer.

Technological features. The Internet's Technological success depends on its principal communication tools, the Transmission Control Protocol (TCP) and the Internet Protocol (IP). They are referred to frequently as TCP/IP. A protocol is an agreed-upon set of conventions that defines the rules of communication. TCP breaks down and reassembles packets, whereas IP is responsible for ensuring that the packets are sent to the right destination.

Data travels across the Internet through several levels of networks until it reaches its destination. E-mail messages arrive at the mail server (similar to the local post office) from a remote personal computer connected by a modem, or a node on a local-area network. From the server, the messages pass through a router, a special-purpose computer ensuring that each message is sent to its correct destination. A message may pass through several networks to reach its destination. Each network has its own router that determines how best to move the message closer to its destination, taking into account the traffic on the network. A message passes from one network to the next, until it arrives at the destination network, from where it can be sent to the recipient, who has a mailbox on that network. *See* LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

TCP/IP. TCP/IP is a set of protocols developed to allow cooperating computers to share resources across the networks. The TCP/IP establishes the standards and rules by which messages are sent through the networks. The most important traditional TCP/IP services are file transfer, remote login, and mail transfer.

Early history of the Internet

Year	Event
1969	ARPANET (Advanced Research Projects Agency Network) online, connecting four university computer centers using NCP (network control protocol)
1972	With 30 host computers in network, public demonstration at a computer conference in Washington, DC; Internetworking Group (INWG) formed to establish standard protocols
1977	THEORYNET on line, providing e-mail to over 100 non-defense-contracting computer researchers at the University of Wisconsin
1979	UNIX User Network started, to send technical messages, and later adapted for other discussion groups
1981	Corporate funding of BITNET, providing e-mail to some academics
1982	INWG establishes TCP/IP (Transmission Control Protocol/Internet Protocol) as the standard on ARPANET
1983	MILNET online, devoted to operational military nodes formerly on ARPANET
1984	Over 1000 host computers connected to the Internet
1986	NSFNET (National Science Foundation Network) online; management of regional nets by private firms begins
1988	Private networks and for-profit service providers establish links to the Internet for commercial users
1989	Over 100,000 host computers connected to the Internet
1991	World Wide Web software distributed by the European Laboratory for Particle Physics (CERN)
1993	Contract to manage various functions of NSFNET awarded to three corporations; a graphical browser of the Web (MOSAIC) distributed by the National Science Foundation
1995	Widespread popularity of the Internet, particularly e-mail and Web, extending to small businesses, homes, and schools around the world
1997	Internet becomes truly global

The file transfer protocol (FTP) allows a user on any computer to get files from another computer, or to send files to another computer. Security is handled by requiring the user to specify a user name and password for the other computer.

The network terminal protocol (TELNET) allows a user to log in on any other computer on the network. The user starts a remote session by specifying a computer to connect to. From that time until the end of the session, anything the user types is sent to the other computer.

Mail transfer allows a user to send messages to users on other computers. Originally, people tended to use only one or two specific computers. They would maintain "mail files" on those machines. The computer mail system is simply a way for a user to add a message to another user's mail file.

Other services have also become important: resource sharing, diskless workstations, computer conferencing, transaction processing, security, multimedia access, and directory services.

TCP is responsible for breaking up the message into datagrams, reassembling the datagrams at the other end, resending anything that gets lost, and putting things back in the right order. IP is responsible for routing individual datagrams. The datagrams are individually identified by a unique sequence number to facilitate reassembly in the correct order. The whole process of transmission is done through the use of routers. Routing is the process by which two communication stations find and use the optimum path across any network of any complexity. Routers must support fragmentation, the ability to subdivide received information into smaller units where this is required to match the underlying network technology. Routers operate by recognizing that a particular network number relates to a specific area within the interconnected networks. They keep track of the numbers throughout the entire process.

Internet connection speeds. In the conventional system, a stand-alone personal computer accesses the Internet through a dial-up telephone connection via a modem. This provides a temporary connection, known as a SLIP (single-line Internet protocol) or PPP (point-to-point protocol) connection, that enables full access to the Internet as long as the telephone connection is maintained. *See* MODEM.

Cable modem. A more advanced system is via the cable modem, which guarantees Internet service at much higher speeds. For example, a 2-magabyte file, which might take 9 min to download in the conventional system with a 28.8-kilobit-per-second telephone modem, can be downloaded in 10 s using a cable modem. The cable modem is faster because it uses the digital fiber-optic network, which is superior to the available conventional telephone system. Since the cable modem is always connected, there is no need to log in or dial up. However, the cable modem tends to be rather slow at peak hours when a number of subscribers try to send or receive data at the same time. In addition, the number of lines available via cable modem is limited, and is not likely to increase appreciably in the near future, despite a rapidly growing demand for higher speeds.

DSL. The digital subscriber line (DSL) offers a faster Internet connection than a standard dial-up connection or the cable modem connection. DSL uses the existing telephone line and in most cases does not require an additional line. Once installed, the DSL router provides the customer site with a continuous connection to the Internet without any fear that it will cause a busy signal to occur over the telephone line. In its various forms, DSL offers users a choice of speeds ranging from 32 kilobits per second to more than 50 megabits per second. Most Internet service providers offer symmetric DSL data services at a series of speeds so that customers can choose the rate that meets their specific business needs. Over any given link, the maximum DSL speed is determined by the distance between the customer site and the central office. At the customer premises, a DSL router or modem connects the DSL line to a local-area network or an individual computer. DSL takes existing voice cables that connect customer premises to the telephone company's central office and turns them into a high-speed digital link.

ADSL. Asymmetric digital subscriber line (ADSL) is a broadband communication technology designed for use on regular telephone lines. It is called asymmetric because more bandwidth is reserved for receiving data than for sending data. It has the ability to move data over telephone lines at speeds up to 140 times faster than the fastest analog modems available.

Domain Name System. The addressing system on the Internet generates IP addresses, which are usually indicated by numbers such as 128.201.86.290. Since such numbers are difficult to remember, a user-friendly system has been created known as the Domain Name System (DNS). This system provides the mnemonic equivalent of a numeric IP address and further ensures that every site on the Internet has a unique address. For example, an Internet address might appear as crito.uci.edu. If this address is accessed through a Web browser, it is referred to as a URL (Uniform Resource Locator), and the full URL will appear as http://www.crito.uci.edu.

The Domain Name System divides the Internet into a series of component networks called domains that enable e-mail (and other files) to be sent across the entire Internet. Each site attached to the Internet belongs to one of the domains. Universities, for example, belong to the "edu" domain. Other domains are gov (government), com (commercial organizations), mil (military), net (network service providers), and org (nonprofit organizations).

An Internet address is made up of two major parts separated by an @ (at) sign. The first part of the address (to the left of the @ sign) is the username, which usually refers to the person who holds the Internet account and is often the person's login name or in some way identifies him or her. The second part of the address (to the right of the @ sign) contains the hostname (which can refer to a server on a network), followed by the Internet address, which together identify the specific computer where the person has an Internet e-mail account. For example, the address jlin@dee.uxyz.edu can be diagrammed as shown in the **illustration**.

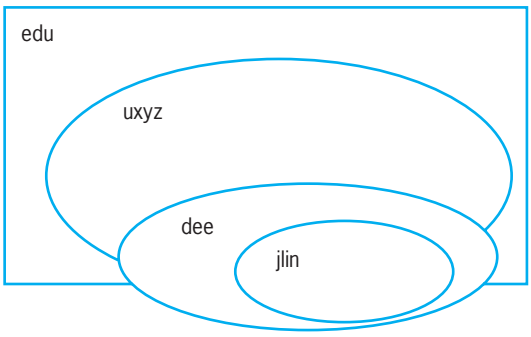

Diagram of the relationship of the sets of computers specified by the Internet address jlin@dee.uxyz.edu.

International domains. Because the Internet has become a global medium, many questions arise as to how the domains are to be managed and controlled, and by whom. Currently, a nongovernmental organization, the Internet Corporation for Assigned Names and Numbers (ICANN), performs a supervisory or advisory function in regard to standards, multilingual sites, domain names, and dispute resolution relative to international domain issues, and in particular, cybersquatting. Generally speaking, cybersquatting refers to the practice of registering a domain name of a well-known trademark in hopes that the owner of the trademark will eventually pay to get it back. Cybersquatting can also be considered an infringement of property right even if it does this unintentionally with the same effect. Since there is no case law or precedents governing Internet maintenance and management at the international level (or for that matter at the national level), many areas are open to question. Besides ICANN, other organizations specialize in various technical and management issues relating to Internet use and dissemination at the global as well as the local level. They work in consultation with ICANN.

Communication on the Internet. Electronic mail, or e-mail, may be the most heavily used feature of the Internet. Binary files, such as pictures, videos, sounds, and executable files, can be attached to e-mail messages. Because the Internet is not able to directly handle binary files in e-mail, the file must be encoded in one of a variety of schemes. Popular schemes are MIME and uuencode. The person who receives the attached binary file (called an attachment) must decode the file with the same scheme that was used to encode the file. Many e-mail software packages do this automatically. *See* ELECTRONIC MAIL.

Usenet. This is the world's largest electronic discussion forum, and it provides a way for messages to be sent among computers across the Internet. People participate in discussions on thousands of topics in specific areas of interest called newsgroups. There are at least 20 major hierarchies of newsgroups, such as recreation (identified by the letters "rec") and computers (identified by the letters "comp"). Within these major hierarchies are subcategories (such as rec.arts) and further subcategories (such as rec.arts.books).

Chats. One of the most immediate ways to communicate with others via the Internet is to participate in "chats" in which a person holds live keyboard "conversations" with other people on the Internet. A person can hold chats with many people simultaneo0usly.

Internet telephony. It is also possible to make telephone calls on the Internet. This can be done in two ways. A computer and special hardware and software can be used to make calls, so that communication is through the computer. In the second method, often referred to as Internet telephony, the user can make calls in the usual way, except that the call is routed over the Internet rather than the normal telephone service. This usually cuts the cost of making long-distance calls. In fact, the Internet telephone call is free; the user pays only for the Internet connection.

Intranets. An intranet is a packet of services that are accessible only within an organization. An intranet works over an organization's internal network and is set up for a wide variety of purposes, including e-mail, group brainstorming, group scheduling, access to corporate databases and documents, and videoconferencing, as well as buying and selling goods and services. Intranets use TCP/IP networks and technologies as well as Internet resources such as the World Wide Web, e-mail, Telnet, and FTP. An intranet is separated from the rest of the Internet by a firewall, a hardware and software combination that prohibits unauthorized access to the intranet. *See* COMPUTER SECURITY.

Telnet. A remarkable feature of the computer is that it lets an individual use the resources of a distant computer. A person can log onto another computer, issue commands as if he or she were at that computer's keyboard, and then gain access to all the computer's resources. This is done with an Internet resource called the Telnet. Telnet follows a client-server model, which allows a piece of software running on the user's own personal computer (the client) to use the resources of a distant server computer. The host allows many clients to access its resources at the same time; it is not devoted to a single user. *See* CLIENT-SERVER SYSTEM.

Many hosts on the Internet can be accessed by using Telnet. They are all different computers, so many of them do not work or look alike. For example, some might be UNIX-based systems, some might be Windows NT/2000®-based computers, some might be Macintoshes®, and so on. To connect using Telnet, the client must use terminal emulation, in order to ensure, in essence, that the client's keyboard and monitor function as the host expects. The most common terminal emulation is the VT-100, which is safe to use.

World Wide Web. The World Wide Web (WWW) is based on technology called hypertext. Most of the Web development has taken place at the European Laboratory for Particle Physics (CERN). The Web may be thought of as a very large subset of the Internet, consisting of hypertext and hypermedia documents. A hypertext document is a document that has a reference (or link) to another hypertext document, which may be on the same computer or in a different computer that may be located anywhere in the world. Hypermedia is a similar concept except that

it provides links to graphic, sound, and video files in addition to text files.

In order for the Web to work, every client must be able to display every document from any server. This is accomplished by imposing a set of standards known as a protocol to govern the way that data are transmitted across the Web. Thus data travel from client to server and back through a protocol known as the HyperText Transfer Protocol (http). In order to access the documents that are transmitted through this protocol, a special program known as a browser is required, which browses the Web. *See* WORLD WIDE WEB.

Commerce on the Internet. Commerce on the Internet is known by a few other names, such as e-business, Etailing (electronic retailing), and e-commerce. The strengths of e-business depend on the strengths of the Internet. The Internet is available practically all over the world, at all times. It is simple to use, and the transaction costs for the end user are low. The costs are also extremely low for the vendors on the Internet, compared to traditional distribution channels. The Internet allows for two-way communication and is built on open standards. The two-way communication, in turn, allows for direct feedback of the customers, and the open standards mean interoperability between companies. It is fairly easy to integrate processes, services, and products, once they have been digitized.

Internet commerce is divided into two major segments, business-to-business (B2B) and business-to-consumer (B2C). In each are some companies that have started their businesses on the Internet, and others that have existed previously and are now transitioning into the Internet world. Some products and services, such as books, compact disks (CDs), computer software, and airline tickets, seem to be particularly suited for online business. Alladi Venkatesh

Bibliography. U. D. Black, *Internet Architecture: An Introduction to IP Protocols*, Prentice Hall PTR, Upper Saddle River, NJ, 2000; D. Comer, *The Internet Book: Everything You Need To Know about Computer Networking and How the Internet Works*, 2d ed., Prentice-Hall, Upper Saddle River, NJ, 3d ed., 2000; H. M. Deitel, P. J. Deitel, and T. R. Nieto, *Internet and the World Wide Web: How To Program*, Prentice Hall, Upper Saddle River, NJ, 2000; C. Dhawan, *Remote Access Networks: PSTN, ISDN, ADSL, Internet and Wireless*, McGraw-Hill, New York, 1998; E. A. Hall, *Internet Core Protocols: The Definitive Guide*, O'Reilly, Cambridge, MA, 2000; D. Minoli and A. Schmidt, *Internet Architectures*, Wiley, New York, 1999; M. Stefik, *The Internet Edge: Social, Legal, and Technological Challenges for a Networked World*, MIT Press, Cambridge, MA, 1999.

Interplanetary matter

Low-density dust or gas that fills the space in a planetary system around or between the planets. Most interplanetary matter in the inner solar system is dust created by collisions among asteroids or released by comets as they pass by the Sun. Ionized gas, launched at high speeds from the Sun as the solar wind, also permeates the solar system and creates a variety of important electromagnetic effects where it interacts with planets. Viewed from a nearby star, the interplanetary matter around the Sun would outshine the Earth. Recent detections of large quantities of similar material around other stars—many of which appear to harbor planets—have rekindled widespread interest in this subject.

Zodiacal dust. A cloud of dust, called zodiacal dust, fills the plane of the solar system interior to the asteroid belt. An observer on the Earth can see this dust with the unaided eye on some clear moonless nights because it scatters sunlight in the Earth's direction. It appears as a faint triangle of light in the ecliptic plane, which connects the constellations of the zodiac. Sometimes, backscattered sunlight from the dust can be seen as the Gegenschein, an additional faint patch of light in the antisolar direction. *See* ZODIACAL LIGHT.

Two satellites, the *Infrared Astronomical Satellite* (*IRAS*) and the *Cosmic Background Explorer* (*COBE*), made detailed maps of the zodiacal cloud as seen from Earth orbit, using detectors that were sensitive at a wide range of infrared wavelengths (1–200 micrometers). While a human being, who observes light at visible wavelengths, can hardly discern the zodiacal cloud among the stars and the galactic background, instruments in space that sense midinfrared wavelengths (7–30 μm) detect a bright glare from this dust cloud that dominates all other background sources. This glare is thermal emission akin to the glow from a red-hot stove. *See* HEAT RADIATION; INFRARED ASTRONOMY.

Most of this zodiacal dust comes from collisions among asteroids. Prominent families of asteroids, groups of asteroids that appear to have fragmented from a single larger body, create bands of dust that both *IRAS* and *COBE* detected. Comets also contribute large amounts of dust as they pass near the Sun. *IRAS* and *COBE* detected trails of dust left by individual comets, mixing into the background cloud. *See* ASTEROID; COMET.

The cloud dust takes the shape of a disk whose thickness is proportional to the distance from the Sun. At the asteroid belt, most of the dust sits closer than 0.2 astronomical unit from the midplane of the solar system. (One astronomical unit, or AU, equals the mean distance from the Earth to the Sun, 1.496×10^8 km or 9.3×10^7 mi.) Zodiacal dust has been measured within a few solar radii from the Sun in the F corona, where it reaches sublimation temperatures. In between the asteroid belt and the F corona, the number density of the particles falls off as $r^{-1.3}$, where r is the distance from the Sun. *See* SUN.

The sizes of the particles lie on a continuum from 1-μm-diameter dust particles to the size of asteroids, the largest of which is 950 km (590 mi) in diameter. The larger objects contain most of the interplanetary mass, but the smaller objects combined have most of the surface area, and so emit and reflect the most light. Near the Earth, the density of the cloud is about 3×10^{-23} g cm^{-3}.

A parameter called β, the ratio of the magnitude of the radiation pressure force on a particle to the magnitude of the Sun's gravitational force on the particle, determines the dynamical behavior of each particle. In general, small particles, which have a high mass-to-surface area ratio, have high β values. Particles smaller than about 1 μm tend to have β values near 1; radiation pressure rapidly blows them out of the solar system. These tiny particles that stream away from the Sun are called β-meteoroids. *See* RADIATION PRESSURE.

Solar photons add energy to the orbits of dust grains, but because the radiation pressure from these photons is outward, they do not add angular momentum. When the particles reradiate this energy, however, the radiation is beamed in the direction of the particle's motion, according to special relativity. The reaction to the force of the beamed radiation, the Poynting-Robertson drag, removes angular momentum from the orbit. Due to Poynting-Robertson drag, zodiacal dust with β somewhat less than 1 spirals in from the asteroid belt, reaching the Sun in roughly 10,000 years for a 10-μm particle, 10 times as long for a 100-μm particle. In general, because of the effects of radiation pressure and Poynting-Robertson drag, the steady presence of large amounts of dust in a planetary system implies the existence of a population of larger orbiting bodies to supply it. *See* POYNTING'S VECTOR.

The shape and composition of zodiacal dust particles vary depending on their origin, and none is homogeneous and spherical. However, these particles fly so fast through the inner solar system, up to 30 km s^{-1} (20 mi/s), that any attempt to catch them for close study destroys many particles completely, and biases the sample. One such biased sample, micrometeoroids collected at high altitudes in the Earth's atmosphere, consists of small particles that mostly resemble carbonaceous chondrites, the most primitive meteorites. They are aggregates of micrometer-sized silicate grains, with roughly solar composition, minus the volatile gases, plus some extra carbon and water. Particles that originate from comets are thought to be fluffy and porous, like a bird's nest; the spaces in the structure are where ice used to be before the comet neared the Sun and began to sublimate.

Interplanetary gas: the solar wind. The inner solar system also contains some interplanetary gas, or more specifically, plasma, most of which comes from the Sun in the form of the solar wind. The solar wind flows supersonically in all directions out of the corona of the Sun at typical speeds of 450 km s^{-1} (280 mi/s) and temperatures of 100,000 K. Like the corona, the solar wind consists mostly of ionized hydrogen. At the Earth's orbit, the density of the wind is roughly 5 particles/cm^3 (80 particles/in.^{3}), or about 8×10^{-24} g cm^{-3}. *See* SOLAR WIND.

Although the wind is electrically neutral, it supports a magnetic field with a strength of about 5 nanotesla (5×10^{-5} gauss) at the orbit of the Earth, which flows outward from the Sun along with the gas. Due to the rotation of the Sun, the path of the solar wind takes the shape of an archimedean spiral, and the magnetic field lines stretch so that they are locally parallel to this spiral. Interaction with the solar wind charges interplanetary dust particles to typical potentials of +5 volts. This charge allows the magnetic field carried by the wind to entrain the smallest β-meteoroids. This magnetic field also protects the solar system from low-energy galactic cosmic rays.

The zone near the Earth where the planet's magnetic field dominates the field in the solar wind is the Earth's magnetosphere. Mercury, Jupiter, Saturn, Uranus, and Neptune also have strong magnetic fields, and therefore large magnetospheres. The fields of these planets punch holes in the solar wind, and the solar wind sweeps these magnetospheres into long magnetotails stretching away from the Sun, removing ionized hydrogen and other gases from the atmospheres of these planets in the process. The interaction of the Earth's magnetic field and the solar wind causes the aurorae. *See* AURORA; JUPITER; MAGNETOSPHERE; MERCURY (PLANET); NEPTUNE; PLANETARY PHYSICS; SATURN; URANUS.

Comets also release gas, which sublimates from ice on their surfaces when they are near the Sun. This gas is mostly water and water-derived molecules such as OH and H_3O^+. However, in situ measurements of Comet Halley with the *Giotto* space probe suggest that comets may release up to six times as much mass in solids as in gas. The solar wind soon drags the gas released by comets into long plasma tails and finally pushes it out of the solar system.

The solar wind flows all the way out to about 50 AU, where the thermal pressure of the interstellar gas which surrounds the solar system overcomes the pressure of the solar wind, which is mostly ram pressure due to its high velocity flow. This interaction occurs at a shock front called the heliopause. Many other stars have winds and other outflows; these phenomena are not associated with the existence of planets. *See* INTERSTELLAR MATTER.

Outer solar system. The region between the asteroid belt and the orbit of Neptune appears to be relatively devoid of interplanetary dust. The *Pioneer 10* and *11* spacecraft and the *Ulysses* space probe found two minor populations of dust particles in this zone: the β-meteoroids ejected from the inner solar system, and a flow of interstellar dust grains on hyperbolic orbits. Beyond Neptune, a ring of comet-sized icy bodies, called the Kuiper Belt, orbits the Sun. Undoubtedly, Kuiper Belt Objects collide with one another and produce a disklike cloud of dust. Neptune's gravitational effects are expected to confine most of this dust to the region beyond 30 AU from the Sun. Dust so far from the Sun is cold, may be 50 K or less, so that it radiates most of its thermal energy at far-infrared wavelengths (30–200 μm). The low temperature makes the cloud too dim to detect from near the Earth, even though it may be just as massive as the zodiacal cloud. *See* KUIPER BELT; SPACE PROBE.

Other planetary systems. Other stars have circumstellar clouds of dust that may be analogous to the interplanetary dust in the solar system. The *Spitzer* satellite and other telescopes have imaged dusty disks around many nearby sunlike stars. Some of

these stars are known to host planets as well as dust; the dust we see is interplanetary matter in these extrasolar planetary systems.

Surveys with the *Spitzer Space Telescope* have found many more examples of this phenomenon since its launch in 2003. The strong infrared emission is thought to come from starlight absorbed and reradiated by large quantities of circumstellar dust.

Circumstellar dust clouds around stars older than roughly 10 million years are often referred to as debris disks. Several stars in this age group with strong infrared emission, notably Beta Pictoris, AU Microscopii, HR 4796, Fomalhaut, Vega, Epsilon Eridani, and Tau Ceti, appear to be surrounded by rings or disks. These disks are thought to represent the "debris" left-over from the process of planet formation.

Some debris disks appear to harbor planetary systems. Some have planets that have been detected by close monitoring of the radial velocity of the star to look for the wobble caused by gravitational pull of an orbiting body. Other debris disks have planetary systems inferred from the structure asymmetries of the disks themselves. The debris disk around the star Beta Pictoris has been found to contain small amounts of gas, possibly arising from cometlike bodies. *See* EXTRASOLAR PLANETS.

The dust in the solar system emits and reflects at least six times as much light as the Earth. However, circumstellar dust around other stars is much more difficult to see than dust in the solar system. Known debris disks around other stars typically have hundreds or thousands of times as much dust as the solar system; less massive disks would be too faint to detect. Imaging disks often requires a special observational tool, like a coronagraph or an interferometer.

Debris disks often have relatively little dust interior to a distance of roughly 10 AU from their host stars. When a debris disk contains dust orbiting within a few astronomical units of the star, this dust is called exozodiacal dust, in analogy with the zodiacal dust in the solar system. Exozodiacal dust is considered one of the chief obstacles in the search for extrasolar planets that resemble the Earth.

Planets and interplanetary dust. The planets in the solar system interact with interplanetary dust in several important ways. For one, they can directly intercept some of the particles. Some 400 metric tons (4×10^5 kg, or 8×10^5 pounds-mass) of meteoritic material fall on the Earth every day. Objects more massive than about 1 kg (2.2 lbm) have enough material to survive the passage through the Earth's atmosphere and reach the ground; these are called meteorites. Particles less massive than 10^{-10} g (2×10^{-13} lbm), smaller than about 1 μm, can also survive the fall because they slow down so fast; these are called micrometeorites. Meteoroids larger than 1 g (2×10^{-13} lbm) produce visible trails as they disintegrate overhead; these are shooting stars (meteors). *See* METEOR; METEORITE; MICROMETEORITE.

Planets also affect interplanetary dust grains indirectly through their gravitational perturbations. For example, as zodiacal dust spirals past the Earth on its way into the Sun, some of the grains become trapped in resonant orbits with the Earth. If the number of years it takes for the particle to orbit the Sun is approximately a ratio of two small integers (like 3/2 or 5/3), the slight gravitational tugs the particle feels when it passes near the Earth tend to add up, reinforcing each other. The sum of all these tugs can temporarily counteract Poynting-Robertson drag and halt the dust grain's inward spiral. This process creates a ring of enhanced dust density around the Sun at a distance of roughly 1 AU. The *IRAS* and *COBE* missions detected both this ring and an asymmetry in the ring, a wake of further density enhancement trailing the Earth in its orbit.

No existing telescope can directly image a planet around another star. But the large-scale dynamical effects of a planet on a cloud of circumstellar dust are often easier to detect from afar than the planets themselves. For example, resonant phenomena like the Earth's dust ring may cause the asymmetries seen in maps of the dust around the nearby star Vega. Large planets may maintain some of the central holes in disks debris, just as Neptune probably prevents most dust generated in the Kuiper Belt from spiraling past it toward the Sun. Marc J. Kuchner

Bibliography. E. Grün et al. (eds.), *Interplanetary Dust*, 2001; A. C. Levasseur-Regourd and H. Hasegawa (eds.), *Origin and Evolution of Interplanetary Dust*, 1991.

Interplanetary propulsion

Means of providing propulsive power for flight to the Moon or to a planet. A variety of different propulsion systems can be used. The space vehicles for these missions consist of a series of separate stages, each with its own set of propulsion systems. When the propellants of a given stage have been expended, the stage is jettisoned to prevent its mass from needlessly adding to the inertia of the vehicle. *See* ROCKET STAGING.

By expelling mass at high velocities, the propulsion systems provide impulse which permits the vehicle to execute its flight trajectory and the necessary maneuvers. Interplanetary or lunar trajectories, as well as the associated maneuvers, are complex, and different kinds of propulsion systems are desirable to meet the requirements of the various phases of flight. Although all propulsion systems actually used in interplanetary flight have been chemical rockets, several basically different propulsion systems have been studied. *See* ION PROPULSION; ROCKET PROPULSION; SPACE FLIGHT; SPACE NAVIGATION AND GUIDANCE; SPACE PROBE; SPACE TECHNOLOGY; SPACECRAFT PROPULSION; SPACECRAFT STRUCTURE. George P. Sutton

Interpolation

A process in mathematics used to estimate an intermediate value of one (dependent) variable which is a function of a second (independent) variable when values of the dependent variable corresponding to

several discrete values of the independent variable are known.

Suppose, as is often the case, that it is desired to describe graphically the results of an experiment in which some quantity Q is measured, for example, the electrical resistance of a wire, for each of a set of N values of a second variable υ representing, perhaps, the temperature of the wire. Let the numbers Q_i, $i = 1, 2, \ldots, N$, be the measurements made of Q and the numbers υ_i be those of the variable υ. These numbers representing the raw data from the experiment are usually given in the form of a table with each Q_i listed opposite the corresponding υ_i. The problem of interpolation is to use the above discrete data to predict the value of Q corresponding to any value of υ lying between the above υ_i. If the value of υ is permitted to lie outside these υ_i, the somewhat more risky process of extrapolation is used. *See* EXTRAPOLATION.

Graphical interpolation. The above experimental data may be expressed in graphical forms by plotting a point on a sheet of paper for each pair of values (υ_i, Q_i) of the variables. One establishes suitable scales by letting 1 in. represent a given number of units of υ and of Q. If υ is considered the independent variable, the horizontal displacement of the ith point usually represents υ_i and its vertical displacement represents Q_i.

If, for simplicity, it is assumed that the experimental errors in the data can be neglected, then the problem of interpolation becomes that of drawing a curve through the N data points P_i having coordinates (x_i, y_i) that are proportional to the numbers υ_i and Q_i, respectively, so as to give an accurate prediction of the value Q for all intermediate values of υ. Since it is at once clear that the N measurements made would be consistent with any curve passing through the points, some additional assumptions are necessary in order to justify drawing any particular curve through the points. Usually one assumes that the υ_i are close enough together that a smooth curve with as simple a variation as possible should be drawn through the points.

In practice the numbers υ_i and Q_i will contain some experimental error, and, therefore, one should not require that the curve pass exactly through the points. The greater the experimental uncertainty the farther one can expect the true curve to deviate from the individual points. In some cases one uses the points only to suggest the type of simple curve to be drawn and then adjusts this type of curve to pass as near the individual points as possible. This may be done by choosing a function that contains a few arbitrary parameters that may be so adjusted as to make the plot of the function miss the points by as small a margin as possible. For a more complete discussion of this topic *see* CURVE FITTING.

For many purposes, however, one uses a French curve and orients it so that one of its edges passes very near a group of the points. Having drawn in this portion of the curve, one moves the edge of the French curve so as to approximate the next group of points. An attempt is made to join these portions of the curve so that there is no discontinuity of slope or curvature at any point on the curve.

Tabular interpolation. This includes methods for finding from a table the values of the dependent variable for intermediate values of the independent variable. Its purpose is the same as graphical interpolation, but one seeks a formula for calculating the value of the dependent variable rather than relying on a measurement of the ordinate of a curve.

In this discussion it will be assumed that x_i and y_i ($i = 1, 2, \ldots, N$), which represent tabulated values of the independent and dependent variables, respectively, are accurate to the full number of figures given. Interpolation then involves finding an interpolating function $P(x)$ satisfying the requirement that, to the number of figures given, the plot of Eq. (1) pass

$$y = P(x) \tag{1}$$

through a selected number of points of the set having coordinates (x_i, y_i). The interpolating function $P(x)$ should be of such a form that it is made to pass readily through the selected points and is easily calculated for any intermediate value of x. Since many schemes are known for determining quickly the unique nth degree polynomial that satisfies Eq. (1) at any $n + 1$ of the tabulated values and since the value of such a polynomial may be computed using only n multiplications and n additions, polynomials are the most common form of interpolating function.

If the subscripts on x and y are reassigned so that the points through which Eq. (1) passes are now (x_0, y_0), (x_1, y_1), $\ldots$, (x_n, y_n), the polynomial needed in Eq. (1) may be written down by inspection, and one has Eq. (2), where Eq. (3) applies. Equation (3) is

$$y = \sum_{k=0}^{n} \frac{L_k(x)}{L_k(x_k)} y_k \tag{2}$$

$$L_k(x) = \frac{(x - x_0)(x - x_1) \cdots (x - x_n)}{(x - x_k)} \tag{3}$$

Lagrange's interpolation formula for unequally spaced ordinates. Since $L_k(x)$ vanishes for all x_s in the set $x_0, x_1, \ldots, x_n$ except x_k, substituting $x = x_s$ in the right-hand side of Eq. (1) gives rise to only one nonzero term. This term has the value y_s, as required.

For $n = 1$ Eqs. (2) and (3) give rise to Eq. (4),

$$y = \frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x - x_1} y_1 \tag{4}$$

whose plot is a straight line connecting the points (x_0, y_0) and (x_1, y_1). Such an interpolation is referred to as linear interpolation and is used in all elementary discussions of interpolation; however, another equivalent form of this equation, given below in Eq. (12), is more often used in these cases.

Suppose the table were obtained from the equation $y = f(x)$, in which $f(x)$ is some mathematical function having continuous derivatives of all order up to and including the $(n + 1)$th. It is then possible to obtain an accurate expression for intermediate values of $f(x)$ by adding to the right-hand side of Eq. (2)

a so-called remainder term, expression (5), where

$$\frac{(x-x_0)(x-x_1)\cdots(x-x_n)}{(n+1)!}f^{(n+1)}(\xi) \qquad (5)$$

$f^{(n+1)}(\xi)$ is the $(n+1)$th derivative of $f(x)$ at some point $x=\xi$ lying between the smallest and largest of the values $x_0, x_1, \ldots, x_n$. Since the value of ξ is not known, the remainder term is used merely to set an upper limit on the truncation error introduced by using Lagrange's interpolation formula.

If the ordinates are equally spaced, that is, $x_s = x_0 + sh$ where h is the interval of tabulation, Lagrange's interpolation formula simplifies considerably and may be written as Eq. (6), where n is the degree of in-

$$y = \sum_{k=-p}^{n-p} A_k(u)y_k = A_{-p}(u)y_{-p} + \cdots + A_0(u)y_0 + \cdots + A_{n-p}(u)y_{n-p} \qquad (6)$$

terpolating polynomial that now passes through the $n+1$ points $(x_{-p}, y_{-p}), \ldots, (x_{n-p}, y_{n-p})$. Here p is the largest integer less than or equal to $n/2$, and the $A_{-p}(u), \ldots, A_{n-p}(u)$ are polynomials in the variable shown as Eq. (7). The polynomials of the variable in

$$u = \frac{x-x_0}{h} \qquad (7)$$

Eq. (7) have been tabulated as functions of u.

Inverse interpolation. If the value of y is known and the value of the corresponding independent variable x is desired, one has the problem of inverse interpolation. Since the polynomials $A_k(u)$ in Eq. (6) are known functions of u, and the values of y and y_k are also known, the only unknown in this equation is u. Thus the problem reduces to that of finding a real root u of an nth-degree polynomial in the range $0 < u < 1$. Having found u, one may find x from Eq. (7). For a discussion of numerical methods for solving for such a root and for more information on interpolation *see* NUMERICAL ANALYSIS.

One may also perform an inverse interpolation by treating x as a function of y. Since, however, the intervals between the y_i are not equal it is necessary to employ the general interpolation formula of Eq. (2) with the xs and ys interchanged.

Round-off errors. In the tabulated values, round-off errors y_i resulting from the need to express the entries y_i of the table as finite decimals will cause an additional error in the interpolated value of y that must be added to the truncation error discussed before. The effect of these errors on the application of Lagrange's interpolation formula is seen by Eq. (6) to be a total error ϵ_T in y given by Eq. (8). Letting e

$$\varepsilon_T = \sum_k A_k(u)\varepsilon_k \qquad (8)$$

be the smallest positive number satisfying the condition $e > |\epsilon_k|$ for all k, one knows from Eq. (8) that relation (9) holds. Since the sum of the $A_k(u)$ is equal

$$|\varepsilon_T| \le \sum_k |A_k(u)| \, |\varepsilon_k| \le e\sum_k |A_k(u)| \qquad (9)$$

to 1, the factor

$$\sum_k |A_k(u)|$$

in Eq. (9) is usually not much larger than 2 or 3 and thus the interpolated value of y has about the same round-off error as the individual entries.

Use of finite differences. For some purposes it is more convenient to use an interpolating formula based not so much on the entries y_i of a central difference table as upon their differences:

x_{-2}	y_{-2}				
		$\delta y_{-3/2}$			
x_{-1}	y_{-1}		$\delta^2 y_{-1}$		
		$\delta y_{-1/2}$		$\delta^3 y_{-1/2}$	
x_0	y_0		$\delta^2 y_0$		$\delta^4 y_0$
		$\delta y_{1/2}$		$\delta^3 y_{1/2}$	
x_1	y_1		$\delta^2 y_1$		$\delta^4 y_1$
		$\delta y_{3/2}$		$\delta^3 y_{3/2}$	
x_2	y_2		$\delta^2 y_2$		
		$\delta y_{5/2}$			
x_3	y_3				
⋮	⋮				

Each difference $\delta^k y_s$ is obtained by subtracting the quantity immediately above and to the left of it from the quantity immediately below and to the left; thus, Eq. (10) can be written, where k and $2s$

$$\delta^k y_s = \delta^{k-1} y_{s+1/2} - \delta^{k-1} y_{s-1/2}. \qquad (10)$$

are required to be integers. For example, $\delta y_{1/2} = y_1 - y_0$ and $\delta^2 y_0 = \delta y_{1/2} - \delta y_{1/2}$.

An interesting property of a difference table is that, if y, the dependent variable tabulated, is a polynomial of the nth degree in x, its kth difference column will represent a polynomial of degree $n-k$. In particular, its nth differences will all be equal and all higher differences will be zero. For example, consider a table of cubes and the difference table formed from it by the rule given above:

χ	$y = \chi^3$				
0	0				
		1			
1	1		6		
		7		6	
2	8		12		0
		19		6	
3	27		18		0
		37		6	
4	64		24		0
		61		6	
5	125		30		
		91			
6	216				

Most functions $f(x)$, when tabulated at a small enough interval $\Delta x = h$, behave approximately as polynomials and therefore give rise to a difference table in which some order of difference is nearly constant. Consider, for example, the difference table of $\log x$, in which the third differences fluctuate between 7 and 9 times 10^{-7}:

Experimental data, if taken at small enough interval of the independent variable, would be expected

x	$y = \log x$	δy	$\delta^2 y$	$\delta^3 y$
1.00	0.0000 000			
		43 214		
1.01	0.0043 214		−426	
		42 788		8
1.02	0.0086 002		−418	
		42 370		9
1.03	0.0128 372		−409	
		41 961		8
1.04	0.0170 333		−401	
		41 560		7
1.05	0.0211 893		−394	
		41 166		
1.06	0.0253 059			

to exhibit much the same behavior as a mathematical function except for the presence of experimental error. The presence of the latter will cause the differences to have a more or less random fluctuation. The size of the fluctuation may, in fact, be used to indicate the number of significant figures in the data.

The constancy of the third differences for log x indicates that for the accuracy and the interval used, a third-degree polynomial may be employed as an interpolating function. Since such a polynomial is determined by the choice of four coefficients, one would expect the interpolation formula to involve four numbers derivable from the difference table. Thus the forward-interpolation formula of Gauss, Eq. (11), can be written. If terminated after the fourth

$$\begin{aligned} y = y_0 &+ u\delta y_{1/2} + \frac{1}{2}u(u-1)\delta^2 y_0 \\ &+ \frac{1}{3!}u(u^2-1)\delta^3 y_{1/2} \\ &+ \frac{1}{4!}u(u^2-1)(u-2)\delta^4 y_0 \\ &+ \frac{1}{5!}u(u^2-1)(u^2-4)\delta^5 y_{1/2} + \cdots \end{aligned} \tag{11}$$

term, it represents a third-degree polynomial in $u = (x - x_0)/h$, and hence in x. It involves the four constants y_0, $\delta y_{1/2}$, $\delta^2 y_0$, and $\delta^3 y_{1/2}$. Since any one of the entries in the y column may be chosen as y_0, the differences required are picked from a central difference table, for example, in relationship to this entry. The interpolating polynomial obtained passes through the four points (x_{-1}, y_{-1}), (x_0, y_0), (x_1, x_2), (x_2, y_2). In general, the interpolating polynomial will pass through only those points whose y coordinate is needed to form the differences used in the formula.

If one terminates the series in Eq. (11) after the second term, one obtains Eq. (12). This is the linear

$$y = y_0 + u\delta y_{1/2} = y_0 + u(y_1 - y_0) \tag{12}$$

interpolation formula most often used when making a simple interpolation in a table.

There are a great variety of interpolation formulas, such as Gregory-Newton's, Stirling's, and Bessel's, that differ mainly in the choice of differences used to specify the interpolating polynomial.

Difference equations. Repeated application of Eq. (10) may be used to express any difference in terms of the tabulated values, for example, Eqs. (13). Expressed in a more general form Eqs. (13) become Eqs. (14). If one sets the second difference equal to

$$\begin{aligned} \delta y_{1/2} &= y_1 - y_0 \\ \delta^2 y_0 &= y_1 - 2y_0 + y_{-1} \end{aligned} \tag{13}$$

$$\begin{aligned} \delta f(x) &= f\left(x + \frac{h}{2}\right) - f\left(x - \frac{h}{2}\right) \\ \delta^2 f(x) &= f(x+h) - 2f(x) + f(x-h) \end{aligned} \tag{14}$$

zero, one obtains a so-called difference equation for $f(x)$. In general, a difference equation is any equation relating the values of $f(x)$ at discrete values of x.

Difference equations play much the same role in analytical work as differential equations. Because they can be interpreted in terms of only those values of $f(x)$ tabulated at some interval $\Delta x = h$, they are admirably adapted to numerical computations. For this reason most numerical solutions of differential equations involve approximating the equation by a suitable difference equation.

For ordinary differential equations the transformation to a difference equation can be made by replacing each derivative in the equation by the appropriate difference expression according to formula (15),

$$f^{(2k)}(x) \to \frac{1}{h^{2k}}\delta^{2k} f(x) f^{(2k+1)}(x) \to \frac{1}{2h^{2k+1}}\left[\delta^{2k+1} f\left(x + \frac{h}{2}\right) + \delta^{2k+1} f\left(x - \frac{h}{2}\right)\right] \tag{15}$$

where $f^{(n)}(x)$ designates the nth derivative of $f(x)$. The difference equation resulting can then, as mentioned before, be used to express the relationship between the values of $f(x)$ at the discrete points $x_s = x_0 + sh$, $s = 0, 1, 2, \ldots, n$.

Partial difference equations. Suppose one chooses to specify a function of two variables $f(x, y)$ by giving its value at some regular array of points in the xy plane having coordinates $(x_m, y_n$. Then, in place of a linear partial differential equation for $f(x, y)$ one has a linear partial difference equation, Eq. (16), where

$$\sum_{s,\,t} A_s f(x_{i+s}, y_{j+t}) = g(x_i, y_j) \tag{16}$$

$g(x, y)$ is a known function and i and j any of the set of integers for which the difference equation has significance. A difference equation in which some of the $f(x_{i+s}, y_{j+t})$ occur to a power other than the first is termed a nonlinear difference equation.

If one employs a square lattice makeup of the points, Eqs. (17), then Laplace's differential equation is approximated by difference equation (18), where, for simplicity, Eq. (19) holds.

$$x_m = x_0 + mh \qquad y_n = y_0 + nh \tag{17}$$

$$f_{i+1,j} + f_{i,j+1} + f_{i-1,j} + f_{i,j-1} - 4f_{ij} = 0 \tag{18}$$

$$f_{mn} = f(x_m, y_n) \tag{19}$$

See GRAPHIC METHODS; LATTICE (MATHEMATICS).

Kaiser S. Kunz

Bibliography. K. E. Atkinson, *Elementary Numerical Analysis*, 2d ed., 1993; R. L. Burden and

J. D. Faires, *Numerical Analysis*, 7th ed., 2001; J. Gregory and D. Redmond, *Introduction to Numerical Analysis*, 1994; J. Szabados and P. Vertesi (eds.), *Interpolation of Functions*, 1990.

Interstellar extinction

Dimming of light from the stars due to absorption and scattering by grains of dust in the interstellar medium. In the absorption process the radiation disappears and is converted into heat energy in the interstellar dust grains. In the scattering process the direction of the radiation is altered. Interstellar extinction produces a dimming of the light from stars situated beyond interstellar clouds of dust according to the equation $F_\lambda = F_\lambda(0)e^{-\tau_\lambda}$, where F_λ is the observed flux of radiation from the star at a wavelength λ, $F_\lambda(0)$ the flux that would be observed in the absence of interstellar extinction, and τ_λ the dimensionless optical depth for interstellar extinction at λ. Measures of the radiation from pairs of stars of similar intrinsic properties but with differing amounts of interstellar extinction can be used to obtain information about τ_λ, which can then be used to provide clues about the nature of the interstellar dust grains. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; SCATTERING OF ELECTROMAGNETIC RADIATION.

An interstellar extinction curve is a plot of the extinction optical depth τ_λ versus the wavelength or frequency of the radiation. It is common for astronomers to express extinction in the astronomical magnitude system. In that case the extinction in magnitudes, A_λ, and the extinction optical depth are related by $A_\lambda = 1.086\tau_\lambda$. A typical interstellar extinction curve for a 500-parsec path (1630 light-years, 9.58×10^{15} mi, or 1.542×10^{16} km) through the interstellar medium of the Milky Way disk is shown in the **illustration.** The spectral regions spanned by the curve are indicated. The extinction curve exhibits a nearly linear increase from infrared to visual wavelengths. In the ultraviolet there is a pronounced extinction peak near $1/\lambda = 4.6\ \mu m^{-1}$ or $\lambda = 217.5$ nm, followed by an extinction minimum and rapidly rising extinction to the shortest ultraviolet wavelengths for which data exist. Details not shown in the illustration include weak enhancements in extinction in the infrared near 20 and 9.7 μm and a very large number of such enhancements at visible wavelengths which are referred to as the diffuse interstellar features. A feature near 3.1 μm and other weaker infrared features appear in the extinction curves for sight lines passing through exceptionally dense interstellar clouds.

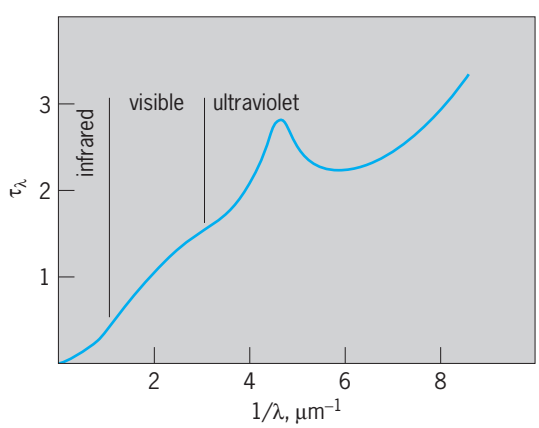

An interstellar extinction curve for a typical 500-parsec path through the interstellar medium of the Milky Way disk. Extinction optical depth τ_λ at wavelength λ is plotted versus $1/\lambda$ in μm^{-1}. The curve spans three segments of the spectrum.

The absorbing and scattering properties of solid particles depend on their size and composition. A detailed interpretation of the interstellar extinction curve of the illustration and other data relating to interstellar dust suggests that the interstellar grains of dust range in size from about 0.01 to 1 μm and are composed of silicate grains (to explain the 20- and 9.7-μm features) and probably some type of carbon grain (to explain the 217.5-nm extinction peak). The 3.1-μm feature in the extinction curve implies that the interstellar dust acquires coatings of water ice in the densest regions of interstellar space. Other ices, including those of ammonia (NH_3) and carbon monoxide (CO), are likely responsible for weaker infrared features observed in the densest interstellar clouds. A comparison of interstellar extinction with the absorption by interstellar atomic hydrogen reveals that the dust contains about 1% of the mass of the interstellar medium. *See* INTERSTELLAR MATTER.

Blair D. Savage

Bibliography. B. T. Draine, Interstellar dust grains, *Annu. Rev. Astron. Astrophys.*, 41:241–289, 2003; J. S. Mathis, Interstellar Dust and Extinction, *Annu. Rev. Astron. Astrophys.*, 28:37–70, 1990; D. C. B. Whittet, *Dust in the Galactic Environment*, 2d ed., 2002.

Interstellar matter

The gaseous material between the stars, constituting several percent of the mass of stars in the Milky Way Galaxy. Being the reservoir from which new stars are born in the Galaxy, interstellar matter is of fundamental importance in understanding both the processes leading to the formation of stars, including the solar system, and ultimately to the origin of life in the universe. Among the many ways in which interstellar matter is detected, perhaps the most familiar are attractive photographs of bright patches of emission-line or reflection nebulosity. However, these nebulae furnish an incomplete view of the large-scale distribution of material, because they depend on the proximity of one or more bright stars for their illumination. Radio observations of hydrogen, the dominant form of interstellar matter, reveal a widespread distribution throughout the thin disk of the Galaxy, with concentrations in the spiral arms. The disk is very thin (scale height 135 parsecs for the cold material; 1-pc is equal to 3.26 light-years, 1.92×10^{13} mi, or 3.09×10^{13} km) compared to its radial extent (the distance from the Sun to the galactic center is about 8000 pc, for example). Mixed in with the gas are small solid particles, called dust grains, of

characteristic radius 0.1 micrometer. Although by mass the grains constitute less than 1% of the material, they have a pronounced effect through the extinction (absorption and scattering) of starlight. Striking examples of this obscuration are the dark rifts seen in the Milky Way, and the Coalsack Nebula in the southern sky. On average, the density of matter is only 15 hydrogen atoms per cubic inch (1 hydrogen atom per cubic centimeter; in total, 2×10^{-24} g · cm^{-3}), but because of the long path lengths over which the material is sampled, this tenuous medium is detectable. Radio and optical observations of other spiral galaxies show a similar distribution of interstellar matter in the galactic plane.

A hierarchy of interstellar clouds, concentrations of gas and dust, exists within the spiral arms. Many such clouds or cloud complexes are recorded photographically. However, the most dense, which contain interstellar molecules, are often totally obscured by the dust grains and so are detectable only through their infrared and radio emission. These molecular clouds, which account for about half of the interstellar mass, contain the birthplaces of stars. *See* GALAXY, EXTERNAL; MILKY WAY GALAXY.

Gas. Except in the vicinity of hot stars, the interstellar gas is cold, neutral, and virtually invisible. However, collisions between atoms lead to the production of the 21-cm radio emission line of atomic hydrogen. Because the Milky Way Galaxy is quite transparent at 21 cm, surveys with large radio telescopes have produced a hydrogen map of the entire Galaxy. Different emission regions in the Galaxy along the same line of sight orbit in the Galaxy with systematically different velocities relative to Earth and so are distinguishable by their different Doppler shifts. Supplemental information is obtained from 21-cm absorption-line measurements when hydrogen is located in front of a strong source of radio emission. These radio studies show that the gas is concentrated in clouds within the spiral arms, with densities typically 100 times the average of 15 atoms per cubic inch (1 atom per cubic centimeter). A typical size is 10 pc, encompassing the equivalent of 500 solar masses. These cold clouds (−316°F or 80 K) appear to be nearly in pressure equilibrium with a more tenuous warmer (14,000°F or 8000 K) phase of neutral hydrogen (whose mass totals that in clouds) and with even hotter ionized coronal-type gas (discussed below) in which they are embedded. *See* RADIO ASTRONOMY.

Other species in the gas are detected by the absorption lines they produce in the spectra of stars, and so are observable only in more local regions of the Milky Way Galaxy. Interstellar lines are distinguished from stellar atmospheric lines by their extreme narrowness and different Doppler shifts. High-dispersion spectra show that the lines are composed of several components possessing unique velocities which correspond to the individual clouds detected with the 21-cm line. Elements such as calcium, sodium, and iron are detected in optical spectra, but the more abundant species in the cold gas, such as hydrogen, carbon, nitrogen, and oxygen, produce lines only in the ultraviolet and require observations from satellites outside the Earth's atmosphere.

A broad ultraviolet line of O VI (oxygen atoms from which five electrons have been removed) has also been discovered; collisional ionization to such an extent requires a very hot gas ($\sim 3 \times 10^5$ K). Such hot gas emits soft x-rays; a Local Bubble of hot gas, extending to about 100 pc from the Sun, is a dominant contributor to the diffuse x-ray background radiation. A similar hot ionized component, called the coronal gas, occupies about half of the volume of interstellar space at the galactic midplane and extends significantly above the disk, but because of its low density (0.045 in.$^{-3}$ or 0.003 cm^{-3}) its contribution to the total mass density is small. *See* X-RAY ASTRONOMY.

When the Galaxy formed, only the elements hydrogen and helium (and traces of lithium) were present. In the course of the evolution of a star, heavier elements are built up by nuclear burning processes in the hot interior. Obviously, the interstellar gas will be gradually enriched in these heavy elements if some of this processed material can be expelled by the star. The most significant such event identified is the supernova explosion due to core collapse, which can occur late in the lifetime of a star whose original mass exceeds about 8 solar masses. A large fraction of the stellar mass is ejected, and in the course of the violent detonation further nucleosynthesis can occur. Also important are supernova explosions of white dwarfs and perhaps nova explosions on white dwarf surfaces. Computer simulations of the burning in supernova shocks predict remarkably large relative abundances of the isotopes. Lower-mass stars might be needed to explain the abundance of some elements such as carbon and nitrogen, the relevant mass loss being by stellar winds from red giants, culminating in the planetary nebula phase. *See* NOVA; NUCLEOSYNTHESIS; PLANETARY NEBULA; SUPERNOVA.

Through many cycles of star formation and mass loss from dying stars, the present interstellar element abundances have been built up to the levels seen in the atmospheres of relatively recently formed stars, including the Sun. The elements other than hydrogen and helium (principally carbon, nitrogen, and oxygen) constitute about 2% by mass. Direct examination of the cool interstellar gas using interstellar absorption lines reveals a considerable depletion of heavy elements relative to hydrogen. The atoms missing in the interstellar gas can be accounted for by the interstellar matter seen in solid particle form.

Molecules. In interstellar space, molecules form either on the surfaces of dust particles or by gas-phase reactions. In dense (greater than 1600 in.$^{-3}$ or 100 cm^{-3}) clouds where molecules are effectively shielded from the ultraviolet radiation that would dissociate them, the abundances become appreciable. These are called molecular clouds. Hydrogen is converted almost completely to its molecular form, H_2. The next most abundant molecule is carbon monoxide (CO), whose high abundance may be attributed to the high cosmic abundances of carbon and oxygen and the great stability of carbon monoxide.

The presence of molecules in interstellar space

was first revealed by optical absorption lines. Unfortunately, most species produce no lines at optical wavelengths, and so the discovery of large numbers of molecules had to await advances in radio astronomy, for it is in this spectral band that molecular rotational transitions take place. Most lines are seen in emission. Dozens of molecules, with isotopic variants, are now known, the most widespread being carbon monoxide (CO), hydroxyl (OH), and formaldehyde (H_2CO). The most abundant molecule, H_2, has no transitions at radio wavelengths, but its presence in molecular clouds is inferred by the high density required for the excitation of the emission lines of other species and the lack of 21-cm emission from atomic hydrogen. In localized regions which have been heated by a shock or stellar wind to several thousand kelvins, the quadrupole infrared rotational-vibrational transitions of H_2 have been detected. Molecular hydrogen has also been observed in less opaque clouds by ultraviolet electronic transitions seen in absorption (carbon monoxide is also seen this way).

Molecular clouds. The existence of molecular clouds throughout the galactic disk has been traced by using the 2.6-mm emission line of carbon monoxide in a manner analogous to mapping with the 21-cm line. About half of the interstellar hydrogen is in molecular rather than atomic form, with enhancements in the inner 500 pc of the Galaxy and in a 3–7 kiloparsec ring. Molecular clouds exist primarily in the spiral arms, where they are thought to form behind the spiral shocks. The molecular clouds are quite cold (−441°F or 10 K), unless they have an internal energy source such as a newly forming protostar. They exist in complexes with the following typical characteristics: radius, 20 pc; density, 1.5×10^4 in.$^{-3}$ (1×10^3 cm^{-3}); mass, 10^5 solar masses. Within the complexes are more dense regions, whose counterparts are locally identified with the optically defined dark nebulae (discussed below). Only in the largest and most dense condensations can the rarer polyatomic (largely organic) molecules be seen. One of the chief such regions is the Orion Molecular Cloud OMC1, of which the Orion Nebula H II region is a fragment ionized by newly formed massive stars. The other region is the 5–10-pc core of the Sagittarius B2 cloud near the galactic center. Altogether, 114 molecules and molecular ions have been detected, many of which could be classed as prebiotic. It is interesting that the first stages of organic evolution can occur in interstellar space.

Unlike diffuse clouds, which are confined by the pressure of the surrounding medium, molecular clouds are confined by their own gravity. Molecular clouds are therefore susceptible to collapse into progressively denser phases, a sequence leading to the formation of new stars—usually in clusters of hundreds or thousands. *See* MOLECULAR CLOUD; ORION NEBULA; STAR FORMATION.

Masers. An intriguing phenomenon seen in some molecular emission lines is maser amplification to very high intensities. The relative populations of the energy levels of a particular molecule are determined by a combination of collisional and radiative excitation. The population distribution is often not in equilibrium with either the thermal gas or the radiation field; and if a higher-energy sublevel comes to have a higher population than a lower one, the electromagnetic transition between the states is amplified through the process of stimulated emission. Such nonequilibrium occurs in interstellar space in two distinct environments. *See* MASER.

The best-understood masers are the hydroxyl (OH), water (H_2O), and silicon monoxide (SiO) masers in the circumstellar envelopes of cool mass-losing red supergiant stars, because the relevant physical conditions relating to density, temperature, and radiation and velocity fields can be estimated independently from other observations. Positional measurements made possible by very long baseline interferometry show that the silicon monoxide and water masing regions are within a few stellar radii of the star, whereas the hydroxyl masers arise in less dense regions more than 100 stellar radii out. Even here the identification of the detailed pump (inversion-producing) mechanism is difficult, but it seems that the 1612-MHz hydroxyl maser is pumped by far-infrared emission by warm circumstellar dust grains and that the silicon monoxide masers are pumped by collisions.

The other masers are found in dense molecular clouds, in particular near compact sources of infrared and radio continuum emission identified as massive stars just being formed. Interstellar hydroxyl and water masers are widespread throughout the Galaxy, methyl alcohol (CH_3OH) masing is seen in OMC1, and silicon monoxide masers are seen in OMC1 and two other star-forming regions. The strongest hydroxyl masers are identified with underlying compact H II regions, probably occurring near the boundary. Although water masers are usually close to such compact H II regions, they are not coincident with the regions of radio or infrared continuum emission. Interstellar masers appear to require collisional pumping, which is thought to be due to powerful jets emitted from stars as they accrete. Peter G. Martin; Christopher D. Matzner

Chemical composition. At present, 123 interstellar molecular species are identified (see **table**). Mainly diatomic species are known, but recently a few more complex species have been seen in diffuse clouds (extinction $\lesssim 1$ magnitude); translucent clouds (1–5 magnitudes) contain molecules as complex as C_3H_2; over half the known species reside in cold dense clouds (>5 magnitudes); while nearly all species exist in warm dense star-forming cores in giant molecular clouds. The known molecules are largely organic. Only 21 are inorganic and stable. There are 36 stable organic species, and 67 unstable species, mostly organic, which include 33 free radicals, 14 (positive) ions, 4 simple rings, and 12 carbon chains as complex as $HC_{11}N$. Most of the unstable species were unknown terrestrially before their interstellar identification. The most rapidly growing list is that of the ions and radicals, because of new laboratory synthesis techniques.

Carbon chemistry is dominant, and long carbon-chain species are most prominent among organic

Known interstellar molecules

Inorganic species (stable)*

Diatomic		Triatomic	4-atom	5-atom
H_2	HCl	H_2O	NH_3	SiH_4*
CO	HF	H_2S		
CS	NaCl*	SO_2		
SiO	AlCl*	OCS		
SiS	KCl*	HNO		
PN	AlF*	N_2O		
		SiH_2 ?		

Organic molecules (stable)*

Alcohols	Aldehydes and ketones	Acids	Hydrocarbons
CH_3OH (methanol)	H_2CO (formaldehyde)	HCN (hydrocyanic)	CH_2 (methylene)
CH_3CH_2OH (ethanol)	CH_3CHO (acetaldeyde)	HCOOH (formic)	CH_4* (methane)
CH_2CHOH (vinyl alcohol)	H_2CCO (ketene)	HNCO (isocyanic)	C_2H_2* (acetylene)
$HOCH_2CH_2OH$ (ethylene glycol)	$(CH_3)_2CO$ (acetone)		C_2H_4* (ethylene)
	HCCCHO (propynal)		C_6H_6 (benzene)?
	$HOCH_2CHO$ (glycol aldehyde)		

Amides	Esters and ethers	Organosulfur
NH_2CHO (formamide)	CH_3OCHO (methyl formate)	H_2CS (thioformaldehyde)
NH_2CN (cyanamide)	$(CH_3)_2O$ (dimethyl ether)	HNCS (isothiocyanic acid)
NH_2CH_3 (methylamine)		CH_3SH (methyl mercaptan)

Paraffin derivatives	Acetylene derivatives	Rings	Other
CH_3CN (methyl cyanide)	CH_3N (cyanoacetylene)	C_2H_4O (ethylene oxide)	CH_2NH (methylenimine)
CH_3CH_2CN (ethyl cyanide)	CH_3C_2H (methylacetylene)		CH_2CHCN (vinyl cyanide)
	CH_3C_4H (methyldiacetylene)		NaCN (sodium cyanide)

Unstable molecules*

Radicals		Ions	Rings	Carbon chains	Isomers
OH	C_2S	H_3^+	SiC_2	ℓ-C_3H	HNC
NO	HCCN	H_2D^+ ?	c-C_3H_2	C_3S	CH_3NC
NS	C_3*	CH^+	c-C_3H	HC_5N	HCCNC
SO	C_3N	N_2H^+	SiC_3	HC_7N	HNCCC
SiN*	C_5N	H_3O^+		HC_9N	MgCN*
NH_2	C_3O	SO^+		$HC_{11}N$	HOC^+
CH	C_4H	CO^+		CH_3C_3N	
SH	C_5*	HCO^+		CH_3C_5N ?	
CN	C_5H	HCS^+		C_4Si	
CSi*	C_6H	$HOCO^+$		H_2CCC	
CP*	C_7H	$HCNH^+$		H_2CCCC	
HCO	C_8H	CH_2D^+ ?		$H_2CCCCCC$	
CH_3	CH_2CN	HC_3NH^+			
C_2	MgNC*	H_2COH^+			
C_2H	ScCN				
C_2O	AlNC*				

* Molecules followed by asterisks are detected only in circumstellar envelopes (surrounding evolved, mass-losing, red giant stars), not yet in the interstellar medium.

compounds. Single, double, and triple carbon bonds are equally prevalent. No branched-chain species are known. Carbon-oxygen and carbon-nitrogen species are of comparable number, while nitrogen-oxygen bonds are rare. The molecules mostly contain carbon, oxygen, and nitrogen, reflecting the high cosmic abundances of these elements ($\sim 10^{-4}$ that of hydrogen). At $\sim 10^{-5}$ of hydrogen, sulfur and silicon are seen in fewer species, although organosulfur compounds (such as C_2S and C_3S) are noteworthy. Phosphorus, magnesium, aluminum, and sodium are seen in only one or two interstellar species, while those containing aluminum, chlorine, fluorine, and potassium arise also in unusual thermochemical equilibrium conditions in the innermost parts of circumstellar envelopes. Despite cosmic abundances similar to sulfur and silicon, the refractory elements magnesium and iron are underrepresented and are presumed locked in interstellar dust grains. *See* ELEMENTS, COSMIC ABUNDANCE OF; FREE RADICAL; ORGANOSULFUR COMPOUND.

Relative to H_2, into which virtually all hydrogen combines in molecular clouds, all other species occur only in trace amounts, ranging in abundance from carbon monoxide (CO) at 2×10^{-4} down to $HC_{11}N$ just detectable at 2×10^{-12}. The simplest species are usually the most abundant, although many complex species are surprisingly abundant also. As interstellar chemistry is far from equilibrium, abundances generally bear no relation to the

constituent atomic abundances or to any thermodynamic properties.

Astrochemistry. Because of the very low temperatures and densities of the interstellar medium, molecules cannot form from atoms by normal terrestrial processes. Astrochemists have focused on three different chemistries. Most important is gas-phase chemistry, particularly involving molecular ions. Ion-molecule reactions satisfy the requirements of minimal activation energy and of rapid two-body rates. The (positive) ions are initiated by the cosmic-ray ionization of H_2, producing H_3^+, which then reacts with other abundant species such as carbon monoxide (CO) and N_2 to produce a large number of the observed species such as HCO^+ and N_2H^+. Ion fragments themselves react rapidly at low temperatures to form larger ions, which eventually recombine with electrons or neutralize by reaction with easily ionized metals. Slower reactions involving neutral atoms and molecules also produce several species. All of the observed molecular ions and a sizable number of the neutral species are successfully modeled by these processes, both in diffuse clouds where photochemistry is also important, and in cold and warm dense clouds opaque to ultraviolet radiation, which harbor many more species. Most of the observed patterns of species are explained. However, several complex molecules are not readily explained, nor is the mere existence of a gas phase in the cold dense clouds, given the rapid rate at which all molecules are expected to adsorb onto dust grains at low temperatures.

The chemistry of interstellar grains takes several forms. Grains can act as passive repositories of frozen molecular material while a cloud is in a cold dense phase, and the material later evaporates without chemical modification if the cloud warms as a result of nearby star formation. The same frozen material can undergo chemical reactions within the icy mantle, modifying the chemical composition before subsequent evaporation. Finally, molecules may be catalyzed from interstellar atoms arriving on dust grains. The ubiquitous H_2 molecule is the best example of a species that cannot form in the gas phase at sufficient rate but that catalyzes on grains and desorbs efficiently because of its high volatility. Details of the catalysis mechanism and the nature of the grain surfaces and composition are lacking, but experiments and models suggest many species can form readily on both ice-covered grains (cold clouds) and bare silicate or carbonaceous grains (warm clouds). In cold clouds especially, desorption is problematic, and the difficulty of desorption probably accounts for the lack both of refractory-element compounds and of several of the more complex species that also present problems for gas-phase formation. Recent searches for new species have targeted large complex stable molecules, such as are expected to form on grains. The two new alcohols and one new aldehyde suggest a chemistry that forms aldehydes and their corresponding reduced alcohols—for example, H_2CO/CH_3OH, long accepted as forming on grains from CO and H. Whether the next steps—CH_3CHO/CH_3CH_2OH or $HOCH_2CHO/HOCH_2CH_2OH$—can occur on grains is unknown. No known gas-phase formation is possible. It is believed that large molecules do form on grains, but desorb only by evaporation driven by nearby star formation, and then react in the gas phase with other gaseous species to form even larger molecules. *See* HETEROGENEOUS CATALYSIS.

Strong shocks abound in the interstellar medium, resulting from supernovae and expanding H II regions. These shocks briefly heat and compress the gas, producing required conditions for many high-temperature chemical reactions. The resulting shock chemistry has proved difficult to simulate. Models depend strongly on the type of shock assumed (dissociative or nondissociative, with or without a magnetic precursor) and do not yet provide a firm basis for comparison with observations. Certain highly abundant species such as hydroxyl (OH) and water (H_2O) in their maser forms, some inorganic sulfur species, and the refractory element species (SiO, SiS, PN) are the best candidates for shock formation.

Besides shocks and heating effects in star-forming regions, other processes also give rise to time dependence in interstellar chemistry. Ion-molecule chemistry in quiescent cold clouds predicts that the abundances of many species, especially the complex carbon-rich compounds (such as HC_3N and HC_5N), are up to 1000 times greater at ages of the order of 10^5 years than after the gas-phase chemistry reaches steady state ($\geq 10^7$ years), when essentially all carbon ends up as carbon monoxide (CO). Many clouds are therefore chemically young, implying a short lifetime, a dynamical mixing mechanism which renews the chemistry, or an overabundance of carbon. *See* SHOCK WAVE.

Isotopic composition. All of the most abundant isotopes of the elements occurring in interstellar molecules have been observed: H, D; ^{12}C, ^{13}C; ^{16}O, ^{18}O, ^{17}O; ^{14}N, ^{15}N; ^{32}S, ^{34}S, ^{33}S; ^{28}Si, ^{29}Si, ^{30}Si; and ^{25}Mg, ^{26}Mg. The relative abundances of the isotopically substituted molecules may differ from the nuclear isotopic ratios because of chemical fractionation, which favors the incorporation of heavier isotopes because of the reduced zero-point vibrational energy of the corresponding molecule. The effect is small for elements whose isotopic mass ratios are small, but becomes noticeable for $^{13}C/^{12}C$, and is enormous for deuterium/hydrogen (D/H), where the concentration factor favoring deuterium may approach 10^4, depending strongly upon the temperature and ionization fraction. Thus D/H cannot be determined from molecular observations; a value of $\sim 1.5 \times 10^{-5}$ is found for the local interstellar medium from ultraviolet observations of the atomic species. The other isotopic ratios can be found reliably from the molecules. Several of these ratios vary with distance from the galactic center.

Because the deuterium fractionation strongly increases with decreasing temperature, the deuterated ratio X_D/X_H has proved an important thermometer at the very lowest temperatures. Many multiply deuterated species have been recently detected, including

D_2CO, D_2H^+, D_2O, D_2S, ND_2H, D_3^+, ND^3. To fully replace all the H atoms with D atoms when D/H $\sim 1.5 \times 10^{-5}$, it appears that H_2D^+ is dramatically enhanced in gas depleted at least by 98% by molecules other than H_2.

Starting with the primordial elements hydrogen and helium, nuclear processing in stars forms the heavier elements and their isotopes, which are continuously cycled between stars and the interstellar medium. The concentrations of these elements and isotopes in the gas as a function of time depend on the rates of star formation, the stellar mass distribution, nuclear processes within the stars, and the rates of ejection of stellar material into the interstellar medium. The common isotopes ^{12}C and ^{16}O are formed first, and ^{13}C, ^{17}O, ^{18}O, ^{14}N, and ^{15}N are formed from them by secondary processes of stellar nucleosynthesis. All grow in abundance with time. For fast rates of stellar processing appropriate to the galactic center, the $^{12}C/^{13}C$ ratio is expected to decrease to a steady value of ~ 20 after 5×10^9 years. For slower rates of stellar processing characteristic of the galactic disk, $^{12}C/^{13}C$ should decrease more slowly, reaching ~ 90 when the Earth was formed 4.5×10^9 years ago, and ~ 70 at the present age of the Galaxy. An observed decrease in $^{18}O/^{17}O$ and a decrease in ^{15}N toward the galactic center are also consistent with advanced nucleosynthesis there. Stellar processing is not the only source of new elements: several isotopes of Li, Be, and B are produced primarily via spallation of interstellar C, N, and O nuclei by cosmic rays. *See* ISOTOPE. Barry E. Turner

Particles. The light of stars near the galactic disk is dimmed by dust grains, which both absorb the radiation and scatter it away from the line of sight. The amount of extinction at optical wavelengths varies approximately as the reciprocal of the wavelength, resulting in a reddening of the color of a star, much as molecular scattering in the Earth's atmosphere reddens the Sun, especially near sunrise and sunset. The dependence of extinction on wavelength is much less steep than it is for molecular scattering, indicating the solid particles have radii about 0.1 micrometer (comparable to optical wavelengths). Satellite observations show a continued rise in the extinction at ultraviolet wavelengths, which seems to require a size distribution extending to smaller interstellar grains. *See* INTERSTELLAR EXTINCTION.

By comparison of the observed color of a star with that predicted from its spectral features, the degree of reddening or selective extinction can be fairly accurately determined, but the total extinction at any given wavelength is more difficult to measure. On average, extinction by dust across 1000 pc in the galactic plane reduces a star's visual brightness by 80%. This requires the mass density of grains to be about 1% of the gas density. Since pure hydrogen or helium grains cannot exist in the interstellar environment, a major fraction of the heavier elements must be in the solid particles. The characteristic separation of micrometer-sized grains is about 8 m (13 km). Studies of reddening in conjunction with measurements of the 21-cm and ultraviolet Lyman-α lines of hydrogen and the 2.6-mm line of carbon dioxide show that dust and gas concentrations are well correlated. This correlation is borne out in infrared maps of diffuse 100-μm thermal emission from dust, made in the all-sky survey by the *Infrared Astronomical Satellite* (*IRAS*).

Because of the high concentration of interstellar material toward the galactic plane, it is extremely difficult to detect radiation with a wavelength less than 1 μm coming a large distance through the plane. Conversely, a line of sight to a distant object viewed out of the plane is much less obscured because the disk is so thin. The zone of avoidance, corresponding roughly to the area occupied by the Milky Way, is that region of the sky in which essentially no extragalactic object can be seen at optical wavelengths because of intervening dust. The dark rifts in the Milky Way result from the same obscuration. The component of starlight scattered rather than absorbed by the grains can be detected as a diffuse glow in the night sky near the Milky Way. However, it must be carefully separated from other contributions to the night sky brightness: the combined effect of faint stars, zodiacal light from dust scattering within the solar system, and airglow (permanent aurora). *See* AIRGLOW; AURORA; INTERPLANETARY MATTER; ZODIACAL LIGHT.

Interstellar dust particles are heated by the optical and ultraviolet interstellar radiation field and, being relatively cool, reradiate this absorbed energy in the infrared. Such thermal radiation accounts for a third of the bolometric luminosity of the Milky Way Galaxy. Measurements by *IRAS* at 100 and 60 μm are in accord with emission expected from the grains whose properties are deduced from extinction. On the other hand, the unexpectedly large portion (25%) at shorter wavelengths (25 and 12 μm) has been taken to indicate a significant population of very small (~ 0.001 μm) grains or large molecules. These have such small heat capacity that they can be heated transiently to a relatively high temperature by the absorption of a single photon.

The light of reddened stars is partially linearly polarized, typically by 1% but reaching 10% for the most obscured stars. The broad peak in polarization at yellow light, together with the correlation of the degree of polarization with reddening, suggests that the polarization and extinction are caused by the same dust grains. The grains must be both nonspherical and spinning about a preferred direction in space to produce polarization. The direction of this polarization is ordered across the sky, often parallel to the galactic plane. This is believed to be due to the organized galactic magnetic field. The field strength is very small; measurements based on the Zeeman effect in the 21-cm line of atomic hydrogen and the 18-cm line of hydroxyl indicate a few microgauss ($1\ \mu G = 10^{-10}$ tesla), with some compression in regions of higher density. Minute amounts (0.01%) of circular polarization have also been used to study the topology of the magnetic field.

The possible types of grain material can be restricted through considerations of relative cosmic abundances, but detailed identification is difficult because of the paucity of spectral features in the

extinction curve. Silicates are suggested by an excess of absorption at 10 μm in front of strong infrared sources, and in the ultraviolet an absorption peak at 220 nanometers could be explained by a component of small graphite particles. Spatially extended red emission in excess of that expected from scattering in reflection nebulae has been interpreted as evidence for fluorescence of hydrogenated amorphous carbon solids. A popular theory of grain formation begins with the production of small silicate and carbon particles in the extended atmospheres of red supergiant stars. While in the circumstellar region, these grains are warmed by the starlight, so that they are detectable by their thermal emission on the near infrared (10 μm). Radiation pressure ejects these particles into the interstellar gas, where they become much colder (−433°F or 15 K). A dielectric mantle is then built up by accretion of the most abundant elements, hydrogen, carbon, nitrogen, and oxygen. Ices of water and carbon monoxide are detected by 3.1- and 4.67-μm absorption bands, respectively, but only deep within molecular clouds; these are probably in the form of volatile coatings accreted on more refractory grains.

Dark nebulae. A cloud of interstellar gas and dust can be photographed in silhouette if it appears against a rich star field; however foreground stars make distant clouds indiscernible. Many large dark nebulae, or groups of nebulae, can be seen in the Milky Way where the material is concentrated; these are coincident with molecular clouds. The distance to a dark nebula can be estimated by using the assumption that statistically all stars are of the same intrinsic brightness. When counts of stars within a small brightness range are made in the nebula and an adjacent clear region, the dimming effect of the cloud will appear as a sudden relative decrease in the density of stars fainter than a certain apparent brightness, which corresponds statistically to a certain distance. Alternatively, a lower limit to the distance is provided by the distance to the most distant unreddened stars in the same direction. One of the best-known and nearest dark nebulae is the Coal Sack, situated at a distance of 175 pc and visible in the Southern Hemisphere. Another example is the "Gulf of Mexico" area in the North America Nebula.

Obscuring clouds of all sizes can be seen against the bright H II regions described below. In many cases the H II regions and dark nebulae are part of the same cloud. The bay in the Orion Nebula is one such region, as well as the spectacular Horsehead Nebula. Even smaller condensations, called globules, are seen in the Rosette Nebula (**Fig. 1**) and NGC 6611 (M16; **Fig. 2**). The globules, which are almost completely opaque, have masses and sizes which suggest they might be the last fragments accompanying the birth of stars. *See* GLOBULE; NEBULA.

Bright nebulae. An interstellar cloud can also become visible as a bright nebula if illuminated by a nearby bright star. Whether an H II region or reflection nebula results depends on the quantity of ionizing radiation available from the star. To be distinguished from H II regions, but often also called bright gaseous nebulae, are shells of gas that have been ejected from stars. Included in this latter category are planetary nebulae and nova shells which have a bright emission-line spectrum similar to that of an H II region, and supernova remnants. *See* CRAB NEBULA.

Fig. 1. Quadrant of the shell-shaped Rosette Nebula, showing dense globules of obscuring dust and gas silhouetted on the bright emission-line background of an H II region. The central hole is swept clear of gas by winds, supernovae, and ionization from newly formed massive stars; radiation pressure from the central star (lower left) acting on dust grains has also been suggested. The quadrant was photographed in red light with the 48-in. (122-cm) Schmidt telescope of the Palomar Observatory. (*California Institute of Technology/Palomar Observatory*)

H II regions. A star whose temperature exceeds about 45,000°F (25,000 K) emits sufficient ultraviolet radiation to completely ionize a large volume of the surrounding hydrogen. The ionized regions (Figs. 1 and 2), called H II regions, have a characteristic red hue resulting from fluorescence in which hydrogen, ionized by the ultraviolet radiation, recombines and emits the H α line at 656.3 nm. Optical emission lines from many other elements have been detected, including the "nebulium" line of oxygen at 500.7 nm. H II regions are also conspicuous sources of free-free-radio emission characteristic of close

Fig. 2. NGC 6611 (M16), a complex H II region in which the exciting stars are members of a cluster. Note the dark globules and elephant-trunk structures, and the bright rims where ionizing radiation is advancing into more dense neutral gas. The nebula was photographed in Hα + [N II] with the 200-in. (508-cm) telescope of the Palomar Observatory. (*California Institute of Technology/Palomar Observatory*)

electron-proton encounters in the 14,000°F (8000 K) gas. Some H II regions are seen only as radio sources because their optical emission is obscured by dust grains. Radio recombination lines of hydrogen, helium, and carbon, which result when the respective ions recombine to highly excited atoms, are also important.

An H II region can be extended with a relatively low surface brightness if the local density is low, as in the North America Nebula. However, the best-known regions, such as the Orion Nebula, are in clouds that are quite dense (1.5×10^4 to 1.5×10^5 H atoms per cubic inch, or 1×10^3 to 1×10^4 H atoms per cubic inch) compared to the average; dense clouds absorb ionizing radiation in a smaller region, and hence produce a brighter fluorescence. Since the brightest stars are also the youngest, it is not surprising to find them still embedded in the dense regions from which they formed.

Ionization of molecular gas increases its temperature from 30 K to 8000 K, and increases its pressure correspondingly. Enhanced pressures cause the H II region to explosively expand away from the forming star, often leading it to blowout through the side of the molecular cloud (the "blister phase," Fig. 2). The cloud mass unbound in this process is typically ten times that of the stellar association causing it, so H II regions are responsible for much of the inefficiency of star formation.

Reflection nebulae. In the absence of sufficient ionizing flux, the cloud may still be seen by the light reflected from the dust particles in the cloud. The scattering is more efficient at short wavelengths, so that if the illuminating star is white the nebula appears blue. The absorption or emission lines of the illuminating star appear in the nebular spectrum as well. Reflection nebulae are strongly polarized, by as much as 40%. Both the color and the polarization can be explained by dust grains similar to those which cause interstellar reddening. Extended near-infrared (2-μm) emission in some bright reflection nebulae provides evidence for thermal emission from transiently heated small grains. There are also numerous near-infrared spectral emission features consistent with C-H and C-C bending and stretching vibrations, and suggestive of polycyclic aromatic hydrocarbons (PAHs). Whether such compounds exist as free-flying molecules (or small grains), as coatings on larger dust grains, or both, is undecided.

Some reflection nebulae, such as those in the Pleiades, result from a chance encounter between a cloud and an unrelated bright star, providing a unique look at interstellar cloud structure. Other reflection nebulae appear to be intimately related to stars in early or late stages of stellar evolution. The Orion Nebula has an underlying reflection nebula arising from the dust in the gas cloud which produced the H II region. However, in this and other H II regions the emission-line radiation rather than the reflected light dominates the nebulosity.

Supernova remnants. Long after an exploding star has faded away, its supernova remnant (SNR) can still be seen. These are composed of gas expelled from the star itself and interstellar gas impelled into motion by it. Supernova remnants provide material to enrich the interstellar gas in heavy elements. The gas shell initially expands at up to 6000 mi $\cdot$ s^{-1} (10,000 km $\cdot$ s^{-1}); during this phase the supernova remnant is a strong emitter of radio waves through the synchrotron emission process in which internally accelerated relativistic electrons (cosmic rays) spiral in a magnetic field which has been amplified by turbulent and compressive motions. Later the supernova remnant is decelerated as it plows up more and more of the surrounding interstellar gas. Compression of the ambient magnetic field and cosmic-ray electron components of this gas leads again to a radio synchrotron source, with a characteristic shell structure. The supernova remnant generates a shock where it impinges on the interstellar medium, and because of the high relative velocity, temperatures of 10^5–10^6 K are reached in the post shock gas. This hot gas has been detected by its x-radiation in both emission lines and the free-free continuum. As the gas cools behind the shock, it produces optical emission lines too, giving rise to the bright shells seen in wide-field optical photographs (for example, the Veil Nebula which is part of the Cygnus Loop supernova remnant).

Supernovae caused by the collapse of a massive star's core produce a compact remnant, either a black hole or a neutron star (depending on the mass of the core before collapse). In the case of a magnetized, spinning neutron star, a pulsar will be observed within the supernova remnant. Pulsars emit a wind of relativistic particles and tightly wound magnetic fields, which illuminate the pulsar wind nebulae found within some supernova remnants via synchrotron emission (such as the Crab Nebula around the Crab Pulsar). Notably, no pulsar has been detected to date from the supernova 1987A in the Large Magellanic Cloud.

Supernova remnants provide an important source of energy for heating and driving turbulent motions in the cooler interstellar gas, and perhaps for accelerating cosmic rays. In low-density regions their influence can propagate through large regions of space, and they are probably responsible for maintaining the high-temperature coronal gas seen throughout the galactic disk. *See* COSMIC RAYS.

Magnetic fields and cosmic rays. The interstellar magnetic field, initially inferred from polarization of starlight, has been measured by means of the Zeeman splitting of the circularly polarized 21-cm emission, by the Faraday rotation of linearly polarized radio signals from pulsars and external galaxies, and from the synchrotron emission due to cosmic-ray particles spiraling about the field lines. The mean interstellar field strength is roughly 5 microgauss, or 10^{-5} of Earth's surface field. Magnetic field lines are efficiently coupled to charged particles, causing fields and gas to behave as a single fluid with properties of both components—notably, shearing of field lines engenders an opposing tension force. Magnetic forces contribute significantly to the support of gas against gravity, both above and below the midplane of the galaxy, and within individual molecular clouds where field strengths are 30 μG.

Cosmic rays are energetic charged particles—electrons, protons, and the nuclei of heavy elements—that pervade the interstellar medium. Cosmic rays observed at Earth include a low-energy variety produced by the Sun (which cause aurorae), interstellar particles of a wide range of energies (which cause genetic mutations and seed cloud formation) produced most likely by supernova remnants, and an extremely high energy population of mysterious extragalactic origin. Like interstellar gas, cosmic rays are effectively coupled to the magnetic field and affect the motion of the combined fluid. Indeed, cosmic-ray pressure is comparable to magnetic forces and turbulent motions in supporting gas away from the midplane. Cosmic rays contribute to the ionization and heating of the gas, but outside of dark clouds they are not as effective as the photoelectric absorption of far ultraviolet photons by dust grains. Their collisions with atomic nuclei produce light elements via spallation, and also produce π^0 mesons which decay to observable gamma-ray photons. *See* COSMIC RAYS.

Star formation. Superluminous stars such as those exciting H II regions cannot be very old because of the tremendous rate at which they are exhausting their supply of hydrogen for nuclear burning; the most luminous are under 100,000 years old. With this clear evidence for recent star formation, observations have been directed toward discovering stars even closer to their time of formation. Compact H II regions, such as the Orion Nebula, appear to be the first fragments of much larger molecular clouds to have formed stars. Ultracompact H II regions, seen in molecular clouds at radio wavelengths but totally obscured at optical wavelengths, appear to be an earlier stage in which the protostellar core has just become highly luminous. These are often called cocoon stars. Even earlier still are the compact infrared sources in which hot dust grains in a protostellar cloud are detected at wavelengths of 5–100 μm. These earliest phases are often associated with intense water and hydroxyl molecular maser emission. Examples of all of these stages are often found in the same region of space. In addition to the Orion molecular cloud, many regions such as the W3 radio and infrared sources associated with the visible H II region IC 1795 have been studied extensively.

Both the broad spectral energy distribution of thermal infrared emission and the spatial distribution of molecular emission from dense gas indicate that the final gravitational collapse of an interstellar cloud fragment commonly proceeds through an accretion disk. A flattened geometry is indeed expected from conservation of angular momentum during the collapse. In the protostar stage, radiant energy is derived from gravitational potential energy released as material accretes onto the star through the disk. A commonly associated phenomenon is the ejection of substantial amounts of gas in high-velocity bipolar outflows along in the direction perpendicular to the disk. These outflows are observed as optical emission lines from collimated jets and from shocked knots called Herbig-Haro objects, and as radio emission from carbon monoxide in extended lobes. Like H II regions, violent protostellar outflows are responsible for casting molecular gas away from forming stars and limiting the efficiency with which molecular clouds are converted into stars. *See* HERBIG-HARO OBJECTS; PROTOSTAR.

Although the details are far from understood, there is evidence that star formation is a bimodal process, with more massive stars forming only in some giant molecular cloud complexes which aggregate in the spiral arms of the Milky Way Galaxy, and less massive stars occurring independently over the whole range of molecular cloud sizes. Low-mass stars are less luminous and cooler and so do not produce the diagnostic compact H II regions. In the nearby Taurus molecular cloud, numerous dense (greater than 1.5×10^5 in.$^{-3}$ or 10^4 cm^{-3}) stellar mass cores, detected by ammonia (NH_3) emission, appear to be the immediate precursors of low-mass stars. This association is corroborated by their coincidence with other signposts of star formation: *IRAS* infrared sources, bipolar outflows, and T Tauri stars, which are low-mass pre-main-sequence stars.

Molecular clouds are supported against wholesale gravitational collapse by turbulent motions, probably associated with waves in the interstellar magnetic field, and are eventually dispersed by mechanical energy which results from the star formation process itself (bipolar outflows, expanding H II regions, stellar winds). On average, the star formation rate corresponds to a 2% conversion efficiency of molecular cloud material, but occasionally a higher efficiency, in excess of 30%, can lead to a gravitationally bound cluster of stars such as the Pleiades or NGC 6811 (M16; Fig. 2). Overall, the Milky Way Galaxy processes about five solar masses of interstellar matter per year into new stars, somewhat larger than the rate at which dying stars replenish the gas, so that the present interstellar medium is decaying on a time scale of about 1×10^9 years. A supply of fresh gas from outside the galactic disk is suggested by the fact that this is one-tenth of the Galaxy's age. *See* INFRARED ASTRONOMY; STELLAR EVOLUTION.

Peter. G. Martin; Christopher D. Matzner

Bibliography. W. B. Burton, B. G. Elmegreen, and R. Genzel, *The Galactic Interstellar Medium*, 1992; L. D'Hendecourt, C. Joblin, and J. Anthony (eds.), *Solid Interstellar Matter: The ISO Revolution*, Springer-Verlag, 1999; E. Levy, J. I. Lunine, and M. S. Matthews (eds.), *Protostars and Planets III*, 1993; V. Mannings, A. Boss, and S. Russell (eds.), *Protostars and Planets IV*, Unversity of Arizona Press, 2000; D. E. Osterbrock and G. J. Ferland, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*, 2d ed., 2005; J. Palous, W. B. Burton, and P. O. Lindblad (eds.), *Evolution of Interstellar Matter and Dynamics of Galaxies*, 1992; B. E. Turner, Recent progress in astrochemistry, *Space Sci. Rev.*, 51:235–337, 1989; E. F. van Dishoeck and G. A. Blake, Chemical evolution of star-forming regions, *Annu. Rev. Astron. Astrophys.*, 36:317–368, 1998; T. L. Wilson and R. T. Rood, Abundances in the interstellar medium, *Annu. Rev. Astron. Astrophys.*, 32:191–226, 1994.

Intestine

The tubular portion of the digestive tract, usually between the stomach and the cloaca or anus. The detailed functions vary with the region, but are primarily digestion and absorption of food.

The structure of the intestine varies greatly in different vertebrates (see **illus.**), but there are several common modifications, mainly associated with increasing the internal surface area. One, seen in many fishes, is the development of a spiral valve; this turns the intestine into a structure resembling a spiral staircase. Another, seen in some fish and most tetrapods, is simply elongating and then coiling the intestine. This can reach extremes in large herbivores: Oxen have intestinal lengths of over 150 ft (45 m). In numerous forms there are blind pouches, or ceca, off part of the intestine. In fish these are commonly at the anterior end; in tetrapods they generally lie at the junction between the large and small intestines. In all vertebrates the inner surface of the intestine is irregular, with ridges and projections of various sorts; these reach their maximum development in the extremely fine and numerous finger-shaped villi found in mammals.

In humans the intestine consists of the small and large intestines. The small intestine is further divided into three major parts: the duodenum, the jejunum, and the ileum. The duodenum, 10–12 in. (25–30 cm) long, begins at the pyloric sphincter of the stomach and curves around the head of the pancreas on the right side of the anterior part of the abdomen. It receives the ducts of the biliary system and the pancreas. The jejunum and ileum are about 19 ft (6 m) long and form a much-coiled tube that empties at right angles into the large intestine through the ileocolic valve. The large intestine, or colon, consists of five parts: the ascending, transverse, descending, and sigmoid regions, and the terminal rectum which empties into the anal canal.

The microscopic structure of the intestine comprises an inner glandular mucosa, a muscular coat, and an outer serosa of connective tissues which is covered in most areas by peritoneum.

The intestine is supported by dorsal mesenteries of varying extent, which contain an extensive system of arteries, veins, lymphatics, and nerves to the various regions. *See* DIGESTIVE SYSTEM.

Thomas S. Parsons

Intra-Americas Sea

That area of the tropical and subtropical western North Atlantic Ocean encompassing the Gulf of Mexico, the Caribbean Sea, the Bahamas and Florida, the northeast coast of South America, and the juxtaposed coastal regions, including the Antillean Islands.

Meteorology. Meteorologically, the Intra-Americas Sea (IAS) is a transition zone between truly tropical conditions in the south and a subtropical climate in the north. The annual variability of the intertropical convergence zone (ITCZ) controls the easterly direction (that is, from the east) of the trade winds, where low surface pressures and rising air create the characteristic cumulus and cumulonimbus clouds. Except for winter when frontal passages bring westerly winds to the northern Intra-Americas Sea, the intertropical convergence zone dominates the weather. The Intra-Americas Sea is also the region that either spawns or interacts with the intense tropical storms known locally as the West Indian Hurricane. Typically there are 9 tropical storms per year, six of which become intense enough (74-mi/h or 119-km/h surface winds) to be classified as hurricanes. These storms are destructive because of their winds and storm surge, but they provide significant amounts of the annual rainfall to many of the Intra-Americas Sea islands and coastal zones. Air flowing over the Intra-Americas Sea acquires moisture that is the source of much of the precipitation over the central plains of North America. Its quality has caused some meteorologists to give it the name American Monsoon. *See* HURRICANE; MONSOON METEOROLOGY; TROPICAL METEOROLOGY.

Circulation. Ocean currents of the Intra-Americas Sea are dominated by the Gulf Stream System. Surface waters flow into the Intra-Americas Sea through the passages of the Lesser Antilles, and to a lesser extent through the Windward Passage between Cuba and Haiti, and the Anegada Passage between Puerto Rico and Anguilla. These inflowing waters form the Caribbean Current, which flows westward and northward into the Gulf of Mexico through the Yucatán Channel, where it is called the Yucatán Current. Here it is clearly recognized as a western boundary current because of its high surface speed, depth, volume transport, and very clear waters.

As the water passes through the Gulf of Mexico, it is known as the Gulf Loop Current because it flows far to the north, loops anticyclonically (clockwise) and flows south along the continental slope off Florida, and finally loops again cyclonically as it turns to exit into the Straits of Florida. In the Straits of Florida the current is called the Florida Current, a name it retains until about the latitude of Cape Canaveral, after which is it usually called the Gulf Stream. The Gulf Loop Current forms large (186-mi or 300-km diameter) anticyclonic eddies at an average rate of every 10–11 months; these eddies drift into the western Gulf of Mexico, spinning down and releasing their momentum, heat, and salt into the basin. Unlike the European Mediterranean whose deep waters are formed locally, the subsurface waters of the Intra-Americas Sea can be traced upstream as far south as Antarctica (Antarctic Intermediate Water) and as far east as the Canary Islands (Subtropical Underwater). Deep water flowing southward along the eastern boundary of the Intra-Americas Sea is known as the Deep Western Boundary Current, which has its source in the Arctic. Thus the Intra-Americas Sea's oceanic circulation is truly part of a global pattern. *See* GULF OF MEXICO; GULF STREAM; OCEAN CIRCULATION.

Surrounding land areas. Land surrounding the Intra-Americas Sea is often fringed by coral reefs, and there are five tectonic plates that influence its geophysics: the North American, South American,

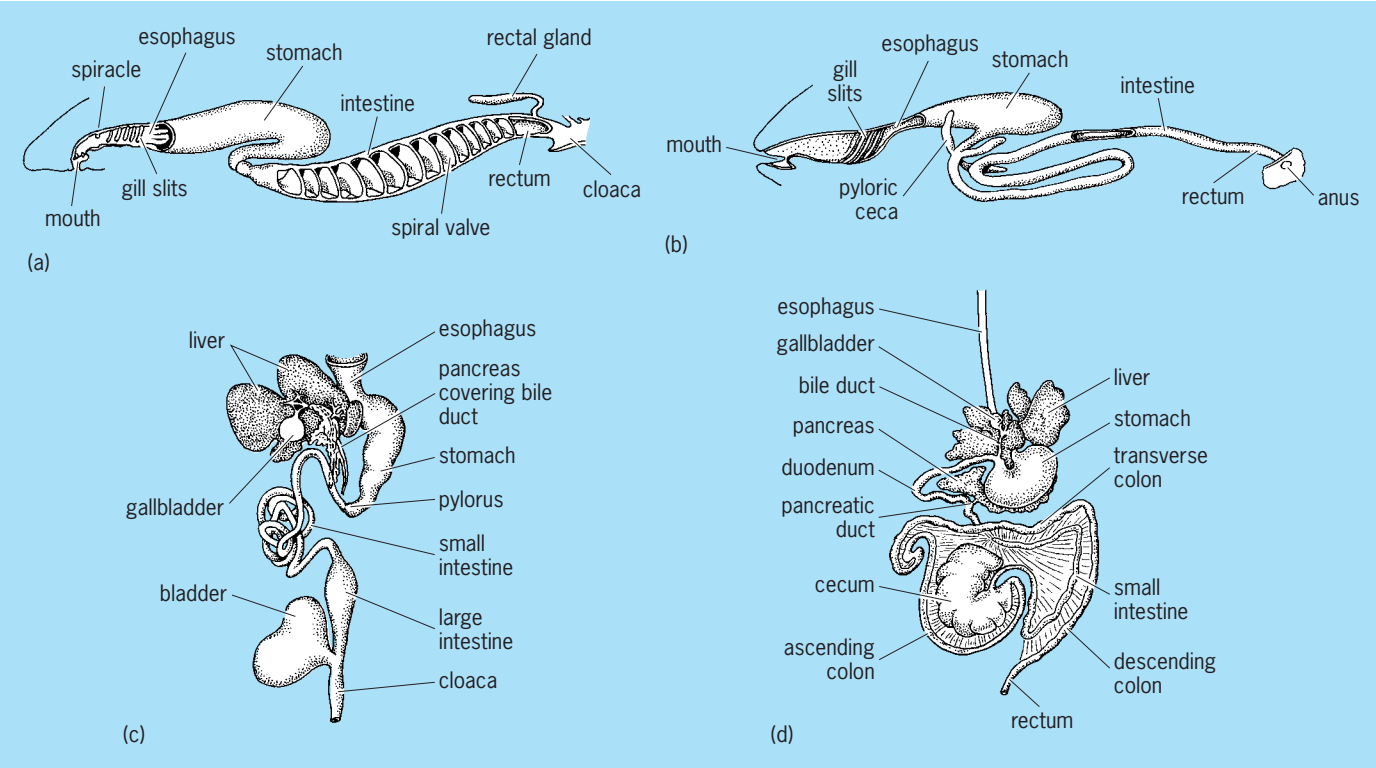

Vertebrate digestive tracts, showing structure of fish, amphibian, and mammalian intestines. (*a*) Shark. (*b*) Perch. (*c*) Frog. (*d*) Guinea pig. (*After A. S. Romer, The Vertebrate Body, 3d ed., Saunders, 1962*)

Caribbean, Cocos, and Nazca plates. Earthquakes are not uncommon, nor are the associated seismic sea waves (tsunamis) that often cause extensive coastal destruction but that are usually considered to be only Pacific Ocean phenomena. The coastal lands are often the site of extensive mangrove forests; the coastal waters are highly productive in shrimp, demersal fishes, mollusks, and lobster; and the coastal and littoral zones are rich in mineral resources, particularly petroleum. River discharge from several South American rivers, notably the Orinoco and Amazon, drifts through the Intra-Americas Sea and carries materials thousands of kilometers from the deltas. The large deltas are heavily impacted by anthropogenic activities, but they remain the source of rich fisheries and plankton communities. Because of the small tidal range in the Intra-Americas Sea, most deltas are wind-dominated geological features. *See* DELTA; EARTHQUAKE; MARINE FISHERIES; PLATE TECTONICS; REEF; TIDE; TSUNAMI. George A. Maul

Bibliography. G. A. Maul (ed.), *Climatic Change in the Intra-Americas Sea*, United Nations Environment Programme, 1993; C. N. K. Mooers and G. A. Maul, Intra-Americas Sea Circulation, *The Sea*, vol. 2, New York, pp. 183-208, 1998.

Intron

A segment of deoxyribonucleic acid (DNA) transcribed into ribonucleic acid (RNA) as part of a longer strand of RNA, called heterogeneous nuclear RNA (hnRNA), but that does not survive the processes of messenger RNA (mRNA) maturation. Hence, genes are split into portions, termed exons, that appear in mRNAs or in structural RNAs, and introns. During the maturation of pre-mRNA to mature mRNA, a series of control processes occur that specify and narrow down the set of mRNAs that become functionally active in a given cell. These processes multiply the species of transcripts through the differential combination of exons and the inclusion of parts of introns into the functional mRNAs; thus, several mRNA products can be generated from a single gene in a controlled manner. For example, certain exons may be included or spliced out during mRNA maturation (alternative splicing of the same hnRNA molecule), or two different hnRNA molecules may form during the splicing reaction of one mRNA transcript, a process termed trans-splicing. *See* DEOXYRIBONUCLEIC ACID (DNA); EXON; RIBONUCLEIC ACID (RNA).

Split genes. Split genes were discovered in 1977 during analyses of mammalian-virus DNAs and their mRNA products. Shortly thereafter, it was shown that an intrinsic mammalian gene, for beta hemoglobin, also has protein-coding regions separated by lengths of DNA that are not translated into protein. Such discoveries were made through techniques of recombinant DNA research, and since 1978 it has become clear that most genes in eukaryotes, and a few in prokaryotes, are split. These include not just a large number of different protein-coding genes but also genes encoding transfer RNAs (tRNAs) in such diverse eukaryotes as yeast and frogs, and genes encoding structural RNAs of ribosomes in some protozoa. Introns are not limited to genes in cell nuclei but are also found in mitochondrial genes of lower

eukaryotes and in some chloroplast genes.

Characteristics. The number of introns in a gene varies greatly, from one in the case of structural RNA genes to more than 50 in the collagen gene. The lengths, locations, and compositions of introns also vary greatly among genes. However, in general these sizes and locations, but not the DNA sequence, are comparable in homologous genes in different organisms. The implication is that introns became established in genes early in the evolution of eukaryotes; further, while their nucleotide sequence appears to be unimportant, their existence, positions, and sizes are significant. As an example, the genes of kinases involved in important signal transduction pathways in eukaryotes—for example, the mitogen-activated protein kinases (JNK and p38)—are interrupted by introns localized primarily in the conserved kinase subdomains. An analysis of these genes in sponges (Porifera) revealed that the positions of the introns are highly conserved from sponges to humans; moreover, in *Caenorhabditis elegans* and *Drosophila melanogaster* they are found, if present, at the same positions.

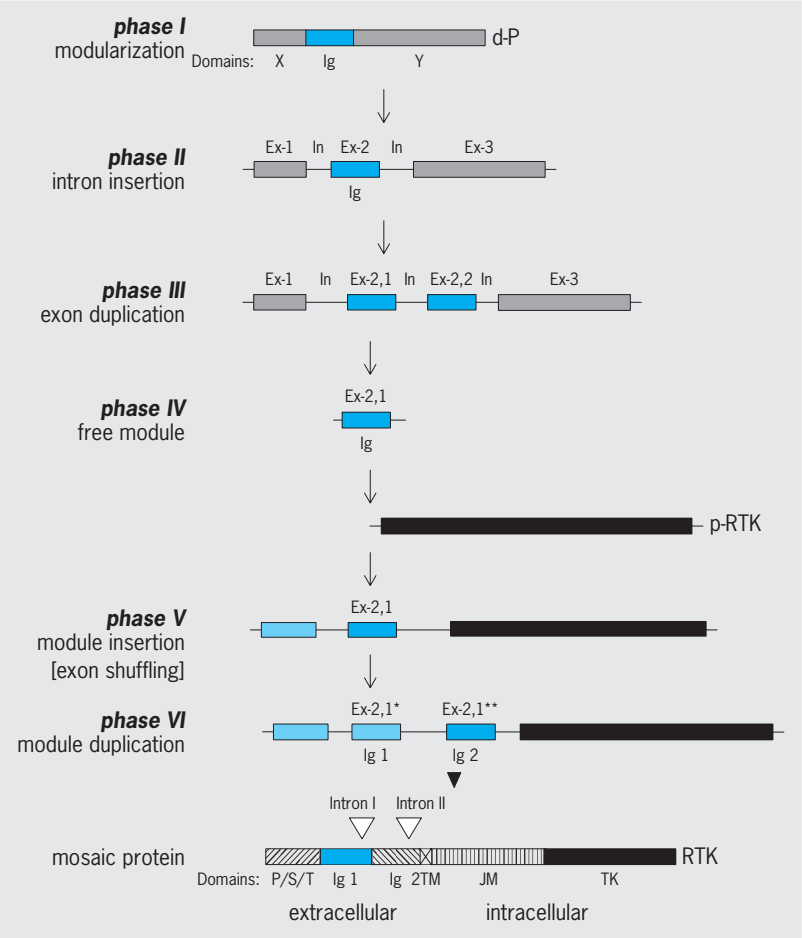

Proposed generation of mosaic proteins for sponge RTK. The gene coding for a donor protein d-P which consists of an Ig-like domain [phase I] undergoes modularization [phase II] by insertion of introns [In]. One of the created exons, Ex-2, is duplicated (forming Ex-2,1 and Ex-2,2) [phase III], and the product Ex-2,1 is liberated [phase IV] and inserted as a module into the precursor gene of RTK [p-RTK]; the process of module insertion by exon shuffling is completed [phase V]. After duplication of the two modules, the sponge RTK gene, composed of two Ig-like domains [Ex-2,1* and Ex-2.1], is completed [phase VI] and the mosaic protein is formed.**

RNA splicing. Three types or mechanisms of RNA splicing have been identified. One involves tRNA, wherein removal of a single intron from tRNA precursors is accomplished by a splicing enzyme complex that recognizes a particular RNA secondary structure, or shape. Another was discovered in studies of protozoan ribosomal RNA (rRNA), and also has been shown to be a part of the maturation of both rRNA and some mRNAs in yeast mitochondria. Excision of the intron from the precursor of *Tetrahymena* rRNA, for example, is an autocatalytic process in which the precursor RNA folds into a structure in which the intron is a loop, then cleaves itself at the ends of the intron and ligates the two parts of the mature rRNA. The reactions proceed without the involvement of protein.

Processing of RNA in mitochondria takes one of two forms. The first is the autocatalytic mechanism mentioned above, and the second involves splicing enzymes, called RNA maturases, that are actually coded for within introns of genes for other mitochondrial proteins. In an interesting form of regulation of gene expression, a maturase mRNA is translated into an enzyme that destroys its own mRNA while splicing together another mRNA. *See* MITOCHONDRIA.

The third splicing mechanism has been characterized for eukaryotic mRNAs of nuclear origin. Although little DNA sequence homology exists among introns in genes that are expressed as mRNA, there is a consistent relatedness among them at their ends (their junctions with flanking protein-coding sequences) that is involved in the removal of intron sequences from gene transcripts during RNA processing. A class of particles known as small nuclear ribonucleoproteins (SnRNPs) have been identified as additional participants in the processing of mRNAs from larger precursors. They interact with exon–intron junctions to loop out introns and facilitate RNA cleavage and exon ligation. In bacteria, certain RNAs in bacteriophage-infected cells result from the same sort of autocatalytic splicing described for protozoan rRNA.

Role. Speculation on the roles and the evolution of introns is mostly based on correlations that have been seen between domains of protein structure and the exons of genes that are defined by intervening introns. Domains in proteins are regions of a molecule that can be distinguished from others on structural and functional grounds. For example, the enzyme tyrosine kinase (TK) has several domains: among them is one that binds the substrate ATP and another that transfers the phosphate group to the protein. The presence of introns in eukaryotic structural genes and the absence in prokaryotes were explained by two different hypotheses. First, the "intron-late" theory suggests that introns arose from transposable elements which were inserted later into genes of eukaryotes. Second, the "intron-early" theory assumes that introns are the relics of the RNA world, implying that the genes-in-pieces organization of the eukaryotic genome reflects a primitive original form. The general conclusion of the intron-early theory was that the exons correspond to the building

blocks from which the genes were composed by intronic recombination.

Mosaic/modular proteins and exon shuffling. Mosaic proteins or modular proteins are formed by fusion of two or more gene segments that encode different protein domains. They are assumed to be young proteins and unique to eukaryotes, especially to multicellular animals. Seeking to explain the evolution of mosaic proteins, the modularization hypothesis was formulated. It was assumed that modules suitable for exon shuffling are created by insertion of introns into the genome adjacent to exons. Members of mosaic proteins contain extracellular domains, such as receptor proteins, including the receptor protein kinases. Mosaic proteins have been identified and described from metazoan phyla including sponges. In contrast to the observations with ancient protein genes, there is strong evidence that in higher eukaryotes, in Metazoa, exon shuffling has been used to create the modular assembly of many multidomain proteins.

To test the hypothesis that the process of exon shuffling contributed to the explosive nature of metazoan radiation, studies on the gene organization of sponges, which are extant metazoans from the period prior to the Cambrian Explosion, have been performed. In one study, it was established that the sponge genes of the receptor tyrosine kinases (RTKs), enzymes found exclusively in Metazoa, contain only two introns located in the region encoding the extracellular part of the RTKs. In contrast, all metazoan genes for RTKs found in phyla younger than the Porifera contain several introns in the TK domain. This surprising fact provided experimental support for the view that ancient genes were not "in pieces." It has been suggested that introns became established in the genes of eukaryotes (and to a limited extent in bacteria) because they facilitate a genetic shuffling or rearrangement of portions of genes which encode various functional units, thus creating new genes with new combinations of properties.

The RTK gene of the sponge *Geodia cydonium* codes for a mosaic protein which is composed of several domains/modules: the Pro-Ser-Thr-rich domain, the two Ig-like domains, and the TK domain. The Ig-like domains are modules according to strict nomenclature. A schematic representation of the formation of the Ig-like module and its subsequent insertion into the RTK precursor gene is shown in the **illustration**. *See* PORIFERA.

Recombination. Genetic recombination within introns, that is, between coding units rather than within them, provides a means of genetic evolution via wholesale reassortments of functional subunits or building blocks, rather than by fortuitous recombinations of actual protein-coding DNA sequences. *See* GENE; GENETIC CODE; RECOMBINATION (GENETICS).

Werner E. G. Müller; Peter M. M. Rae

Bibliography. W. F. Doolittle, Genes in pieces: Were they ever together?, *Nature*, 272:581–582, 1978; V. Gamulin et al., Experimental indication in favor of the introns-late theory: The receptor tyrosine kinase gene from the sponge *Geodia cydonium*, *J. Mol. Evol.*, 44:242–252, 1997; W. Gilbert and M. Glynias, On the ancient nature of the introns, *Gene*, 135:137–144, 1993; B. Lewin, *Genes VII*, Oxford University Press, 2000; W. E. G. Müller (ed.), *Molecular Evolution: Towards the Origin of Metazoa* (*Progress in Molecular and Subcellular Biology 21*), Springer-Verlag, Berlin, 1998; W. E. G. Müller (ed.), *Molecular Evolution: Evidence for Monophyly of Metazoa* (*Progress in Molecular and Subcellular Biology 19*), Springer-Verlag, Berlin, 1998; W. E. G. Müller et al., Conservation of the positions of metazoan introns from sponges to humans: Evolutionary implications, *Gene*, 295:299–309, 2002; L. Patty, *Protein Evolution by Exon-Shuffling*, Springer-Verlag, New York, 1995.

Invasion ecology

The study of the establishment, spread, and ecological impact of species translocated from one region or continent to another by humans. Biological invasions have gained attention as a tool for basic research, used to study the ecology and evolution of populations and of novel biotic interactions; and as a conservation issue tied to the preservation of biodiversity. The invasion of nonindigenous (also called exotic, alien, or nonnative) species is a serious concern for those charged with managing and protecting natural as well as managed ecosystems. For example, nonindigenous species negativley affect over one-third of the U.S. Endangered Species List. *See* ECOLOGY; POPULATION ECOLOGY.

Species. Ecologists make a distinction between introduced species, meaning any species growing outside its natural habitat including cultivated or domesticated organisms, and invasive species, meaning the subset of introduced species that establish free-living populations in the wild. The great majority of introduced species (approximately 90% as estimated from some studies) do not become invasive. While certain problem invaders, such as the zebra mussel (*Dreissena polymorpha*), exact enormous economic and ecological costs, other introduced species are generally accepted as beneficial additions, such as most major food crops.

Routes of entry. Species arrive in a new region by a variety of means, including both intentional and accidental introductions. Intentional plant introductions have been promoted primarily by the horticulture industry to satisfy the public's desire for novel landscaping. However, plants have also been introduced for agriculture, for silviculture, and for control of soil erosion. Intentional animal introductions include game species brought in for sport hunting or fishing [such as ringneck pheasant (*Phasianus colchicus*) in North America, water buffalo (*Bubalus bubalis*) in Australia, and brown trout (*Salmo trutta*) in New Zealand]. Unlike these examples, intentional introductions can also include species that are not necessarily intended to form self-sustaining populations, such as those promoted by the aquarium or

pet trade. Pets may be released into the wild when their disenchanted owners look for an easy method of disposal.

Species introduced accidentally are "hitchhikers." Shipping ballast has been a major vector, first in the form of soil carrying terrestrial invertebrates and plant seeds or rhizomes, and more recently in the form of ballast water carrying planktonic larvae from foreign ports. While many species are introduced in ballast or by similar means (such as packing material before the use of styrofoam), hitchhikers can also be explicitly unwanted parasites that bypass importation and quarantine precautions. For example, many nonindigenous agricultural weeds have been imported in contaminated seed lots.

With the burgeoning of international travel and the rapid development of transportation technology, there are many new routes that species may take. In the past, global patterns of established nonindigenous species reflected the dominance of particular trade routes, for example directional movement from Europe to North America. However, trade and travel now occur between almost any two points on the planet, opening up entirely new patterns of species interchange.

Predicting invaders. Given that most introduced species will not become invaders, an important question concerns the extent to which predictions can be made as to which species will invade. The fact that the same species appear over and over invading different regions [such as the black rat (*Rattus rattus*) and water hyacinth (*Eichornia crassipes*)] is evidence that invasions are not simply random events; rather, there is such a thing as a "good invader." Extrapolating from groups of successful invaders, biologists have compiled lists of traits that seem to confer invasiveness such as rapid growth, high reproductive rates, and tolerance to a broad range of conditions. The flaw with this approach is that saying most invaders grow rapidly is not equivalent to saying most species that grow rapidly will be invasive. To assert the latter, one has to compare the probability of invasion success for species that grow rapidly compared to species that grow slowly. This is a much more difficult task because it requires knowledge of how many species were introduced but were not successful, information that is particularly elusive for accidental introductions. However, being able to predict which species will invade has critical implications for policy regulating the importation of nonindigenous species.

Sites of invasions. Certain types of habitats seem to have higher numbers of established nonindigenous species than others. The characteristics that make a site open to invasion must be determined. For example, islands are notably vulnerable to invasions. Islands usually have fewer resident species to begin with, leading to the conjecture that simpler systems have less biotic resistance to invaders. That is, an introduced species is less likely to be met by a resident competitor, predator, or pathogen capable of excluding it. The idea of biotic resistance is also consistent with the idea that complexity confers stability in natural systems. *See* INVASION ECOLOGY.

A second generalization about invasibility is that ecosystems with high levels of anthropogenic disturbance, such as agricultural fields or roadsides, also seem to be more invaded. Increased turnover of open space in these sites could provide more opportunities for the establishment of new species. An alternative explanation is that many species that adapted to anthropogenic habitats in Europe simply tagged along as humans re-created those habitats in new places. Those species would naturally have an advantage over native species at exploiting human disturbances. A final suggestion by proponents of ecosystem management is that disturbance (including, in this context, a disruption of natural disturbance regimes, for example, fire suppression) weakens the inherent resistance of ecosystems and promotes invasion.

Spread. Early ecological models represented the spread of an invading species across a landscape with a summation of two terms: one describing growth in numbers, and the other describing movement in space, for example:

$$\begin{pmatrix}\text{Rate of}\\ \text{change in}\\ \text{population}\\ \text{density}\end{pmatrix} = \begin{pmatrix}\text{rate of}\\ \text{population}\\ \text{growth due to}\\ \text{reproduction}\end{pmatrix} + \begin{pmatrix}\text{rate of}\\ \text{dispersal}\end{pmatrix}$$

or

$$\frac{\partial N}{\partial t} = rN\left(1 - \frac{N}{K}\right) + D\frac{\partial^2 N}{\partial x^2}$$

where N is population density, t is time, r is the intrinsic rate of population increase, K is the carrying capacity, D is movement by diffusion, and x is position from the origin or release point. For such simple models, mathematicians have shown that the speed of an invasion measured in terms of a new area occupied is given by

$$\begin{pmatrix}\text{Kilometers/}\\ \text{year}\end{pmatrix} \propto \begin{pmatrix}\text{rate of}\\ \text{population}\\ \text{growth at}\\ \text{low density}\end{pmatrix} \times \begin{pmatrix}\text{rate of}\\ \text{dispersal}\end{pmatrix}$$

or

$$C = 2\sqrt{rD}$$

and will be constant after an invasion has been allowed to proceed for a sufficiently long time. The prediction of a constant rate of spread has been compared to empirical data on the movement of invading species over time, and it successfully describes the range expansion of a surprisingly diverse range of organisms, including the collared dove (*Streptopelia decaocto*) and the small white cabbage butterfly (*Pieris rapae*).

A consistent result derived from the theoretical work on range expansion is that long-distance dispersal events can have an enormous effect on the

rate of spread. Moreover, the growth of newly established satellite populations can dominate the overall increase of a species in numbers or area. The obvious implication for management is that small, seemingly insignificant patches or outlying populations of a noxious weed or pest should be the focus of control efforts.

Impact. Invasive species can have several different types of impacts. First, they can affect the traits and behavior of resident organisms (for example, causing a shift in diet, size, or shape of the native species they encounter). Second, impacts can occur at the level of the population, either by changing the abundance of a native population or by changing its genetic composition. Hybridization between an invader and a closely related native can result in introgression and genetic pollution. As in the case of several endangered fish species, the endpoint can be the de facto extinction of the native species when the unique aspects of its genome are overwhelmed. Third, impacts can occur at the level of ecological communities. When individual populations are reduced or even driven extinct by competition or predation by an invasive species, the result is a decrease in the overall biodiversity of the invaded site. Finally, invaders can impact not only other species but the physical characteristics of an ecosystem as well. Shrubs that fix nitrogen with the help of symbiotic bacteria can significantly increase the nutrient pools in soil, while some deep-rooted species can dry down the water table such that native species can no longer survive.

There are two main contributing factors in determining which species have the biggest impacts: abundance and special characteristics. Invaders that reach extremely high density, such as the zebra mussel at up to 4500 per square meter, simply overwhelm all other organisms. Other species have special traits that result in an impact out of proportion to their numbers. For example, the cane toad (*Bufo marinus*) is a noxious pest in Australia in part because its toxic skin can be lethal to potential native predators.

Control. Because of the economic and conservation importance of nonindigenous species, much of invasion ecology focuses on the prevention, eradication, and control of invaders, and the restoration of sites after control. Research has emphasized the importance of early detection and eradication of problem species because of the sheer difficulty of manual and chemical control methods once populations have reached a large size.

Biological control has been touted as an environmentally friendly alternative to herbicides and pesticides. One type of biological control uses native species of insects or microbes as biopesticides to turn back invasive host populations. A second type, termed classical biological control, involves the introduction of new nonindigenous species that attack the invader in its home range. Therefore, classical biological control is also accompanied by the same risks as any intentional introduction. The uncertainties due to the current poor predictive ability of the science of invasion ecology have produced an active debate on the wisdom of using introductions to stem previous introductions. *See* ALLELOPATHY; ECOLOGICAL COMMUNITIES; ECOLOGICAL SUCCESSION; SPECIATION; SPECIES CONCEPT.

Ingrid M. Parker

Bibliography. R. Hengeveld, *Dynamics of Biological Invasions*, Chapman and Hall, 1989; H. A. Mooney and J. A. Drake, *Ecology of Biological Invasions of North America and Hawaii*, Springer-Verlag, 1986; J. L. Ruesink et al., Guilty until proven innocent: Reducing the risks of non-indigenous species introductions, *Bioscience*, 45(7):465–477, 1995; M. Williamson, *Biological Invasions*, Chapman and Hall, 1996.

Inventory control

The process of managing the timing and the quantities of goods to be ordered and stocked, so that demands can be met satisfactorily and economically. Inventories are accumulated commodities waiting to be used to meet anticipated demands. Inventory control policies are decision rules that focus on the tradeoff between the costs and benefits of alternative solutions to questions of when and how much to order for each different type of item.

Benefits of carrying inventories. For a firm, some of the possible reasons for carrying inventories are: uncertainty about the size of future demands; uncertainty about the duration of lead time for receiving deliveries; provision for greater assurance of continuing production, using work-in-process inventories as a hedge against the failure of some of the machines feeding other machines; and speculation on future prices of commodities. Some of the other important benefits of carrying inventories are reduction of ordering costs and production setup costs (these costs are less frequently incurred as the size of the orders are made larger which in turn creates higher inventories); price discounts for ordering large quantities; shipping economies; and maintenance of stable production rates and work-force levels which otherwise could fluctuate excessively due to variations in seasonal demand.

Holding costs. The benefits of carrying inventories have to be compared with the costs of holding them. Holding costs include the following elements: cost of capital for money tied up in inventories; cost of owning or renting the warehouse or other storage spaces; materials handling equipment and labor costs; costs of potential obsolescence, pilferage, and deterioration (these also involve the costs of insurance, security, and protection from natural causes such as humidity or extreme temperatures); property taxes levied on inventories; and cost of installing and operating an inventory control policy.

Pareto analysis. Inventories, when listed with respect to their annual costs, tend to exhibit a similarity to Pareto's law and distribution. A small percentage of the product lines may account for a very large

share of the total inventory budget (they are called class A items, or sometimes the vital few). Aside from the class A items, and in the opposite direction, there exists a large percentage of product lines which tend to constitute a much smaller portion of the budget (they are called class C items). The remaining 20 to 30% of the items in the middle are called class B items.

The ABC analysis may help management direct more of its attention to important issues and may lead to greater cost-effectiveness. For example, if the inventory levels are checked at fixed time intervals, then the status of type A items may be reported weekly, type B items biweekly, and type C items monthly. *See* PARETO'S LAW.

Mathematical inventory theory. Mathematical inventory theory dates back to 1915 to the work of F. Harris of the Westinghouse Corporation, who derived the simple lot size formula. For independent items, economic order quantity (EOQ) means the replenishment quantity (Q) which has to be ordered to minimize the sum of the costs of holding (h dollars per item per period) and setups (and/or ordering costs: K dollars per order). If quantity discounts are not available, and if the demand (d items per period) is known and continues at the same rate through time, then the average cost per period, $Kd/Q + hQ/2$, is minimized by setting $Q = \sqrt{2Kd/h}$. This basic model was later extended to allow for various other conditions. For example, if the replenishment of the lot cannot take place instantaneously, but rather happens at a speed of r items per period, then the optimal value of Q becomes $\sqrt{2Kd/[h(1 - d/r)]}$. Other extensions of the EOQ model have been made to cover cases such as: quantity discounts in purchase prices; allowance for back orders at a cost; limited resources or facilities shared by otherwise independent items; inflation; time-varying deterministic demands and production capacities; and multiple echelons.

As for problems where the demands involve uncertainty, substantial research work has been carried out to develop inventory policies which minimize the expected values of various cost functions. Provided that information regarding the probability density function of demand is known, elegant mathematical solutions are available for independent items on a single echelon. There have also been some findings for certain multiitem, multiechelon problems; however, the amount of further research that needs to be carried out to find optimal solutions for complicated everyday inventory problems is substantial. Inventory control problems of everyday life involve many complications that can include various combinations of sequence-dependent setup times, multiitems, multiechelons with stochastic lead times, joint orders, dependent probabilistic demands, and situations where adequate information on probability density functions is not available. Even worse, the shape of the unknown probability density function may be time-varying. Regardless, mathematical inventory theory is valuable because by using the insight it provides to simpler problems, good heuristics for more complicated everyday problems can be designed.

Simulation. Computer simulation is also used for such purposes. By simulating inventory systems and by analyzing or comparing the performance of different decision policies, further insights can be acquired into the specific problem on hand and a more cost-and service-effective inventory control system can be developed.

Ordering policies. Continuous-review and fixed-interval are two different modes of operation of inventory control systems. The former means the records are updated every time items are withdrawn from stock. When the inventory level drops to a critical level called reorder point (s), a replenishment order is issued. Under fixed-interval policies, the status of the inventory at each point in time does not have to be known. The review is done periodically (every t periods).

Many policies for determining the quantity of replenishment use either fixed-order quantities or maximum-order levels. Under fixed-order quantities for a given product, the size of the replenishment lot is always the same (Q). Under maximum-order levels, the lot size is equal to a prespecified order level (S) minus the number of items (of that product) already in the system. Different combinations of the alternatives for timing and lot sizes yield different policies known by abbreviations such as (s, Q), (s, S), (s, t, S), and (t, S). Other variations of the form of inventory control policies include coordination of timing of replenishments to achieve joint orders, and adjustment of lot sizes to the medium of transportation.

Forecasting. Uncertainties of future demand play a major role in the cost of inventories. That is why the ability to better forecast future demand can substantially reduce the inventory expenditures of a firm. Conversely, using ineffective forecasting methods can lead to excessive shortages of needed items and to high levels of unnecessary ones.

Product design. Careful product design can also reduce inventory costs. Standardization, modularity, introduction of common components for different end products, and extension of the use of interchangeable parts and materials can lead to substantial savings in inventory costs.

MRP system. Material requirements planning (MRP) systems (which are production-inventory scheduling softwares that make use of computerized files and data-processing equipment) have received widespread application. Material requirements planning systems have not yet made use of mathematical inventory theory. They recognize the implications of dependent demands in multiechelon manufacturing (which includes lumpy production requirements). Integrating the bills of materials, the given production requirements of end products, and the inventory records file, material requirements planning systems generate a complete list of a production-inventory schedule for parts, subassemblies, and end products, taking into account the lead-time requirements. Material requirements planning has proved to be a useful tool for manufactur-

ers, especially in assembly operations. *See* MATERIAL RESOURCE PLANNING.

Kanban and just-in-time (JIT). While material requirements planning systems were being developed in the United States, some Japanese manufacturers achieved widely acclaimed success with a different system. By producing components "just in time" to be used in the next step of the production process, and by extending this concept throughout the production line so that even the finished goods are delivered just in time to be sold, they obtained substantial reductions in inventories. One of the key factors for establishing just-in-time is altering the manufacturing process to drastically reduce the setup times and simplifying the ordering and procurement process so that ordering costs are cut down. The idea is to enable the producer to operate with small lot sizes, which get produced when the need arises (and not before).

Once just-in-time is established, an information system is used to determine the timing and quantities of production. Card signals—that is, visible records (in Japanese, *Kanban*)—are used to specify withdrawals from preceding production stages, and to order for production the number and type of items required. Because small batches of production have become economical, the production orders can be filled just in time. Advocates of *Kanban* characterize it as a pull process and criticize material requirements planning as a push system. Though *Kanban* is a simple idea and yields an adaptive–flexible production system, its appropriateness hinges on whether setup and ordering costs have been drastically reduced so as to allow small production batches.

Other systems. An example of application of computerized inventory control systems is IBM's Communication Oriented Production and Informational Control Systems (COPICS), which covers a wide scope of inventory-control-related activities, including demand forecasting, materials planning, and even simulation capabilities to test different strategies.

Taking into account the continuing decreases in the costs of computer hardware, the advances in computerized inventory control systems can be expected to receive even wider applications. *See* COMPUTER-INTEGRATED MANUFACTURING.

Ali Dogramaci

Bibliography. L. S. Aft, *Production and Inventory Control*, 1987; R. G. Brown, *Materials Management Systems*, 1977, reprint 1984; D. W. Fogarty, J. H. Blackstone, and T. R. Hoffmann, *Production and Inventory Management*, 2d ed., 1991; R. T. Lubben, *Just-in-Time Manufacturing: An Elegant Solution*, 1988; S. L. Narasimhan, D. N. McLeavey, and P. J. B. Billington, *Production Planning and Inventory Control*, 1995; R. J. Schonberger, *Japanese Manufacturing Techniques: Nine Hidden Lessons in Simplicity*, 1982; R. J. Schonberger, *World Class Manufacturing Casebook*: *Implementing JIT & TQC*, 1995; E. Silver and R. Petersen, *Decision Systems for Inventory Management and Production Planning*, 2d ed., 1995.

Inverse scattering theory

A theory whose objective is to determine the scattering object, or an interaction potential energy, from the knowledge of a scattered field. This is the opposite problem from direct scattering theory, where the scattering amplitude is determined from the equations of motion, including the potential. The equations of motion are usually linear (operator-valued) equations. An example of an inverse problem was given in an article by M. Kac entitled "Can One Hear the Shape of a Drum?" For some drums this question can be answered yes, for yet others no, but for many the question remains unanswered. *See* SCATTERING EXPERIMENTS (ATOMS AND MOLECULES); SCATTERING EXPERIMENTS (NUCLEI).

Inverse scattering theories can be divided into two types: (1) pure inverse problems, when the data consist of complete, noise-free information of the scattering amplitude; and (2) applied inverse problems, when incomplete data which are corrupted by noise are given. Many different applied inverse problems can be obtained from any pure inverse problem by using different band-limiting procedures and different noise spectra.

The difficulty of determining the exact object which produced a scattering amplitude is evident. It is often a priori information about the scatterer that makes the inversion possible.

Pure inverse scattering theory. Much of the basic knowledge of systems of atoms, molecules, and nuclear particles is obtained from inverse scattering studies using beams of different particles as probes. These probes include beams of charged particles utilizing electromagnetic interactions, weakly interacting particle beams such as neutrinos and antineutrinos, and beams of strongly interacting particles such as protons, neutrons, pions, kaons, and others.

For the Schrödinger equation with spherical symmetry or in one dimension, there is an exact solution of the inverse problem by I. M. Gel'fand and B. M. Levitan. Given complete spectral data, including (1) the energy E_i, and a normalization constant C_i for each bound state, $i = 1, 2, \ldots, N$, and (2) a weight function $w(k)$ for the continuous states, a formula exists to calculate a function $H(r, t)$. From this function the Gel'fand-Levitan equation (1)

$$K(r,t) = -H(r,t) - \int_0^r K(r,s)H(t,s)\,ds \qquad (1)$$

can be solved for K. Then the potential V sought can be calculated from Eq. (2). Furthermore, given

$$V(r) = 2\frac{dK(r,r)}{dr} \qquad (2)$$

suitable discrete and continuous approximate solutions ϕ_0, the exact solutions are given by Eq. (3)

$$\psi(r,k) = \phi_0(r,k) + \int_0^r K(r,t)\phi_0(t,k)\,dt \qquad (3)$$

where k is the wave vector. This completely solves

these inverse problems. *See* ELEMENTARY PARTICLE; INTEGRAL EQUATION; NONRELATIVISTIC QUANTUM THEORY; QUANTUM FIELD THEORY; QUANTUM MECHANICS.

An application of inverse scattering to one-dimensional nonlinear evolution equations was given by C. S. Gardener and colleagues. They studied one-dimensional reflectionless potentials $V(x) = V(x, 0)$ in a time-independent linear Schrödinger equation with one, two, and so forth bound states. The time independence assures that the spectrum remains invariant under translations in time, so that the corresponding flow is isospectral. Then a second linear operator was used to generate translations in time. At large spatial distances x, the regular solution to the second time-dependent equation gave a time-dependent reflection coefficient for scattering from the potential. Inverse scattering theory was then used to convert this reflection coefficient to a time-dependent potential $V(x, t)$. The identification $V(x, t) = \phi(x, t)$ was observed to provide a solution to the nonlinear, dispersive Kortweg–de Vries equation (4), where $\phi_t = \partial\phi/\partial t$, $\phi_{xxx} = \partial^3\phi/\partial x^3$, and

$$\phi_t + 6\phi\phi_x + \phi_{xxx} = 0 \tag{4}$$

so forth. Each bound state of the time-independent Schrödinger equation corresponds to the velocity of one soliton, a localized finite-energy, solitary wave with a super stability. In ordinary stability, a solitary wave would retain its identity after a small enough perturbation; two solitons retain their size, shape, and speed with only a shift in phase even after propagating through one another. This, their superstability, requires a computer calculation to test. A number of other nonlinear equations in one space dimension and one time dimension have soliton solutions, and they are experimentally observed in plasmas, quantum optics, fluids, superconductors, and many other low-dimensional systems. *See* SOLITON.

Applied inverse scattering theory. A number of high-technology areas (nondestructive evaluation, medical diagnostics including acoustic and ultrasonic imaging, x-ray absorption and nuclear magnetic resonance tomography, radar scattering and geophysical exploration) use inverse scattering theory. Several classical waves including acoustic, electromagnetic, ultrasonic, x-rays, and others are used, extending the ability to "see" beyond visible light and beyond electromagnetic waves. For example, acoustic waves readily propagate through a thick metal container while scattering from imperfections in the wall. *See* BIOACOUSTICS, ANIMAL; COMPUTERIZED TOMOGRAPHY; MEDICAL ULTRASONIC TOMOGRAPHY; NONDESTRUCTIVE EVALUATION.

All of the inverse scattering technologies require the solution to ill-posed or improperly posed problems. A model equation is well posed if it has a unique solution which depends continuously on the initial data. It is ill posed otherwise. The ill-posed problems which are amenable to analysis, called regularizable ill-posed problems, are those which depend discontinuously upon the data. This destroys uniqueness, although solutions (in fact, many solutions) exist.

Direct scattering problems are well posed and involve solutions for a scattering amplitude function T from an integral equation of the form of Eq. (5),

$$T = V + VG_0^+T \tag{5}$$

where V and $G^+{}_0$ are given. The pure inverse problem requires the solution of Eq. (6), with T and G_0^+

$$V = -T - TG_0^+V \tag{6}$$

given. However, applied inverse problems involve general equations (7), where Y is the measured quan-

$$Y = R \cdot T \tag{7a}$$

$$Y = R \cdot T + n \tag{7b}$$

tity, T is the direct scattering amplitude, n is noise, and R is a resolution function. The resolution function represents the effects of the measuring apparatus, especially the limit of the range of frequencies accepted. It is Eqs. (7) which are regularizable ill-posed problems.

The discontinuous dependence on data has as a consequence that two given input signals Y_1 and Y_2 can be as close together as desired everywhere, but can lead to solutions T_1 and T_2 which are as far apart as desired. The nonuniqueness means that these equations have (possibly infinite) families of solutions which are near the exact solution of Eq. (5) or (6). It is necessary to have prior knowledge about the scatterer to decide which of these solutions is the best (and even what "best" means here). The regularizer is a mathematical device for choosing one of this family. The solution chosen is called a quasisolution.

The regularizer both removes the discontinuity of the dependence of T on Y and selects a unique quasisolution. The cost of regularization is that the quasisolution selected must have some nonzero minimum spatial resolution. It is impossible to determine any finer details about the scatterer than this spatial resolution.

This limitation arises as follows: Let $E_t(N)$ be the truncation error, that is, the spatial resolution due to truncation of the spatial solutions at N functions of some chosen set (sines and cosines in Fourier series expansions). For a wide variety of scattering problems, it behaves as Eq. (8), where a and b are positive

$$E_t = \frac{a}{N^b} \tag{8}$$

constants which depend upon the experiment being analyzed but not N. The error due to noise E_n behaves as Eq. (9), where c and d are positive constants which

$$E_n = cN^d \tag{9}$$

depend on the experiment but not on N. For pure inverse problems, E_t is the sole source of error so that if N is large enough the error is negligible. For applied inverse problems, both errors occur so that

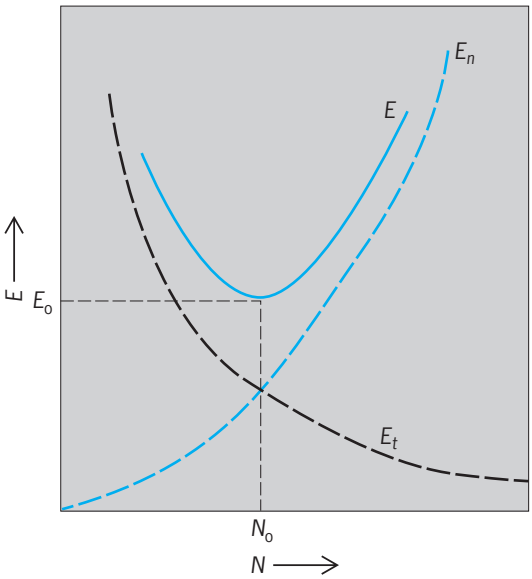

Example of use of Eqs. (8)–(10) to obtain the point of optimal truncation and minimum spatial resolution in an applied inverse scattering problem.

the total error is given by Eq. (10). The **illustration**

$$E = aN^{-b} + cN^{d} \tag{10}$$

shows the truncation error E_t, the noise error E_n; and the total error E, such as that given in Eq. (10). N_0 is the point of optimal truncation. The corresponding value $E_0 = E(N_0)$ is the minimum spatial resolution.

The minimum of Eq. (10) is obtained from Eq. (11), so that N_0 is given by Eq. (12). Since $(d^2E/dN^2)_{N_0} <$

$$\left(\frac{dE}{dN}\right)_{N_0} = 0 = -abN_0^{-b-1} + cdN_0^{d-1} \tag{11}$$

$$N_0 = \left(\frac{ab}{cd}\right)^{1/(d+b)} \tag{12}$$

0, this is a minimum. As an example, let $a = 10^4c$ and $d = b = 2$, which gives that $N_0 = 10$ upon using Eq. (12). In this case, less than 10 terms will not yield all of the information on the scatterer available in the data; more than 10 will lose detail by allowing noise to pollute the signal.

It is the goal of applied inverse scattering theory to understand these issues and to optimize performance for each technology. Each is different, and each serves as a rich source of challenging problems. Brian DeFacio

Bibliography. M. Bertero and E. M. Pike (eds.), *Inverse Problems in Scattering and Imaging*, 1992; K. Chadan and P. C. Sabatier, *Inverse Problems in Quantum Scattering Theory*, 1977, reprint 1989; D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, 1992; G. M. Gladwell, *Inverse Problems in Scattering: An Introduction*, 1993; K. I. Hopcraft, *An Introduction to Electromagnetic Inverse Scattering*, 1992; D. Sattinger et al., *Inverse Scattering and Applications*, 1991.

Inverse-square law

Any law in which a physical quantity varies with distance from a source inversely as the square of that distance. When energy is being radiated by a point source (see **illus.**), such a law holds, provided the space between source and receiver is filled with a nondissipative, homogeneous, isotropic, unbounded medium. All unbounded waves become spherical at distances r, which are large compared with source dimensions so that the angular intensity distribution on the expanding wave surface, whose area is proportional to r^2, is fixed. Hence emerges the inverse-square law.

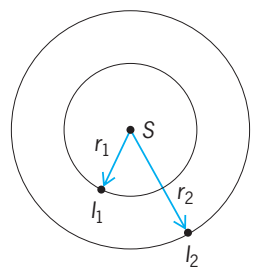

Point source S emitting energy of intensity I. The inverse-square law states that $I_2/I_2 = r_2^2/r_2^1$.

Similar reasoning shows that the same law applies to mechanical shear waves in elastic media and to compressional sound waves. It holds statistically for particle sources in a vacuum, such as radioactive atoms, provided there are no electromagnetic fields and no mutual interactions. The term is also used for static field laws such as the law of gravitation and Coulomb's law in electrostatics. William R. Smythe

Invertebrate embryology

The study of the development or morphogenesis and growth of the invertebrates. The same general principles of development apply to the invertebrates as to the vertebrates. Actually, much of the basic knowledge of embryology has been the result of studies on invertebrates. A common phenomenon in the invertebrates is the release of a free and independent form, the larva, before development is completed. The larvae vary considerably and are characteristic of the different animal groups.

Embryonic development begins with the formation of the gametes in a specialized cell bearing the haploid or N number of chromosomes. *See* GAMETOGENESIS.

Spermatogenesis. Number, compactness, and mobility are important for spermatozoa. Toward this end, the process of spermatogenesis consists of a stage of cell proliferation, followed by a period of progressive concentration and streamlining. The essential, heredity-determining material of the chromosomes is packed tightly into a tiny nucleus. The cytoplasm forms the locomotor apparatus, usually

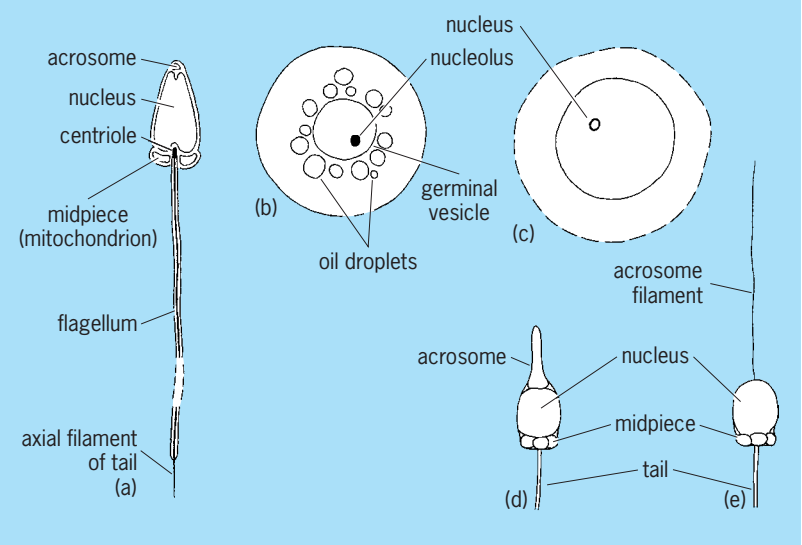

Fig. 1. Spermatozoa and fertilizable eggs. (*a*) Sea urchin spermatozoon. (*b*) Marine annelid (*Nereis*) egg, with intact germinal vesicle containing a nucleolus. Spheres surrounding germinal vesicle are oil droplets. (*c*) Marine mussel (*Mytilus*) egg, germinal vesicle broken down but polar bodies not formed. (*d*) Mussel (*Mytilus*) spermatozoa, acrosome intact. (*e*) *Mytilus* spermatozoa, acrosomal reaction.

a single long flagellum with a centriole at its base and a mitochondrion nearby, as well as an organelle (acrosome) for penetrating the egg coverings. Excess cytoplasm is finally discarded, and the mature spermatozoon (**Fig. 1***a*), ready to take part in fertilization, is a self-contained, stripped-down unit, carrying the hereditary message in code, and provided with enough energy source to propel it in a sustained burst of activity on its one-way trip to join an egg of its species. Millions upon millions of such cells are produced in the testis, where they remain quiescent until they are spawned. *See* SPERM CELL; SPERMATOGENESIS.

Oogenesis. The egg is designed for a very different role. It must contain enough substance to provide structural material for the formation of a set of simple organs of locomotion and digestion so that the young animal can secure food to carry on its further growth. It must also contain enough energy-yielding material to perform the work of dividing the single egg cell into body cells from which such organs can be formed and to synthesize the complex chemical substances needed to provide each of these new cells with a new nucleus. *See* OOGENESIS; OVUM.

The egg, therefore, is specialized for large size and protection of its contents, with less concern for numbers and none at all for motility. In addition, its cytoplasm possesses intrinsic capacities for differentiation and building in exact accordance with the specifications contained in its chromosomes, so that a spider egg, for example, always produces a spider and never a fly. The fact that the physical bases of these capacities have so far eluded most of the efforts directed toward their detection in no way casts doubt on their existence.

The reserve building and energy-yielding materials are stored in the egg cytoplasm as minute spheres or platelets of yolk, a stable lipoprotein substance. Eggs are large cells even without this inert material. At the end of their growth period, when they have accumulated the full amount of yolk, they are huge in comparison to the body cells of the parent animal. No invertebrate eggs, however, achieve the spectacular dimensions of bird eggs. The largest are found among the arthropods [crayfish eggs are 0.1 in. (2.5 mm) in diameter], while some marine animals have very small eggs (oyster eggs are about 65 micrometers).

During the growth period, while the egg cell is actively synthesizing yolk and increasing the amount of cytoplasm, it has a very large nucleus, the germinal vesicle (Fig. 1*b*). When it reaches full size, however, and this synthetic activity subsides, the nuclear membrane breaks down, releasing its contents into the cytoplasm. The two successive nuclear divisions of meiosis follow, but the cytoplasm, instead of dividing equally, pushes out one of the daughter nuclei each time as a polar body. These two minute bodies have no further function in development. The chromosome material left in the egg forms the egg pronucleus, which is ready to unite with the sperm pronucleus. The zygote nucleus, formed by their union, is comparable in size to those of the body cells.

Egg polarity. Many types of eggs show structural departures from radial symmetry which indicate that the unfertilized egg is organized around a bipolar axis, one end of which is called the animal pole and the other the vegetal pole. The polar bodies are given off from the animal pole, and the egg pronucleus remains in this region. When an egg contains conspicuous amounts of yolk, it is usually concentrated in the vegetal half of the egg.

Egg membranes. Since the eggs of invertebrates are often shed directly into the water of oceans and streams, or laid to develop in places where they are exposed to the drying action of air and sunlight, they are always surrounded by a protective covering. In some forms the eggs are laid in batches which may be enclosed in a leathery sac or embedded in a mass of jelly. In other cases each egg has its own separate membranous case, a layer of jelly, or a more complex system of protective structures.

Sperm and egg of each individual species have been shown by light and electron microscopy to be characteristic of its particular species. Mechanisms have evolved that are not fully understood which normally prevent the egg of one species from being fertilized by the sperm of another. Factors responsible for this species specification include egg coats, surface antigens and receptors, and environmental barriers such as substrate, temperature, and variations in salinity.

Preliminaries to fertilization. Sperm cells must complete all the nuclear and cytoplasmic changes and be fully mature before they can take part in fertilization. On the other hand, while this is true of the cytoplasm of egg cells, in most species the nuclear preparation for fertilization is incomplete when sperm entrance takes place. Moreover, the degree of incompleteness varies widely and apparently at random.

The marine annelid *Nereis* sheds its eggs with the germinal vesicle intact (Fig. 1*b*). The entrance of a spermatozoon stimulates the egg to begin the reduction divisions, and the sperm pronucleus waits within the egg cytoplasm while the egg nucleus carries out its preparation for union.

Sea urchin eggs are at the other extreme in this respect. Their reduction divisions are completed in the ovary, and the union of egg and sperm pronuclei follows immediately upon sperm entry. In many other species meiosis begins shortly before the eggs are spawned and stops at one stage or another of the process until the entrance of a spermatozoon sets it going again.

Reproduction among the invertebrates takes place in a variety of ways which differ widely from phylum to phylum. For example, following copulation in *Drosophila* the spermatozoa are stored in a special part of the female reproductive tract and released a few at a time as each ovum passes down the oviduct to be laid. The spermatozoa of squids are packaged in small bundles (spermatophores). These are placed by the male, using one of its arms which is modified for the purpose, in spermatophore receptacles on the body of the female. The eggs are fertilized as they leave the oviduct and are laid singly or in finger-shaped egg masses fixed to underwater objects.

Most of the echinoderms and many mollusks, ascidians, annelids, and coelenterates shed their eggs in tremendous numbers into the seawater where fertilization takes place. The young larvae must usually fend for themselves. In these same groups, however, some species shelter the young in seawater-containing chambers and pockets which are not actually within the body of the parent animal. This is also the case with arthropods such as crabs, which carry the larvae in masses fixed to the abdomen of the female, and some bivalves in which the young develop for a time within the mantle chamber. But whether the fertilized eggs are thus protected by the parent animal, laid in jelly masses or leathery cases or carefully constructed brood cells (bees, hunting wasps), or simply thrown into the water, each egg is an independent unit capable of developing into an adult without any further contribution from the parents.

Fertilization. The course of events of fertilization has been studied in several invertebrates, but especially in the sea urchin. Egg and sperm of the sea urchin are released into the seawater. The eggs are covered with a jelly coat to which a receptor on the plasma membrane of the fertilizing sperm binds. The plasma and outer acrosomal membranes of the sperm break down and fuse with each other as a Ca^{2+} influx occurs; the hydrolytic enzymes within the acrosome are released to lyse the egg coat. Next the inner acrosomal membrane everts by the polymerization beneath it of actin, and forms the acrosomal process which makes contact and fuses with the egg plasma membrane. The egg responds to the sperm by forming a fertilization cone. The sperm nucleus enters the egg, and its DNA swells to form the male pronucleus.

As the sperm binds to the receptors on the egg plasma membrane, the electrical potential of the egg membrane changes and establishes a rapid block to polyspermy. This prevents further sperm from making contact and fusing with the egg. With sperm-egg membrane fusion, Ca^{2+} is released to activate a series of changes in the egg. Oxygen consumption changes, a mitotic factor is released, and cortical granule release is initiated. The cortical granules are compact granules beneath the surface of the egg. At their release (**Fig. 2**), the membranes of egg and cortical granules fuse, and the contents of the granules are released into the perivitelline space surrounding the egg. One component of the granules fuses with the vitelline membrane, becomes strengthened, and lifts off the egg to become the fertilization membrane. A second component of the granules takes up Ca^{2+} and forms the hyaline layer close to the egg surface. As changes occur at the egg surface, the egg pronucleus and the sperm pronucleus with associated astral rays move toward the center of the egg, where they fuse.

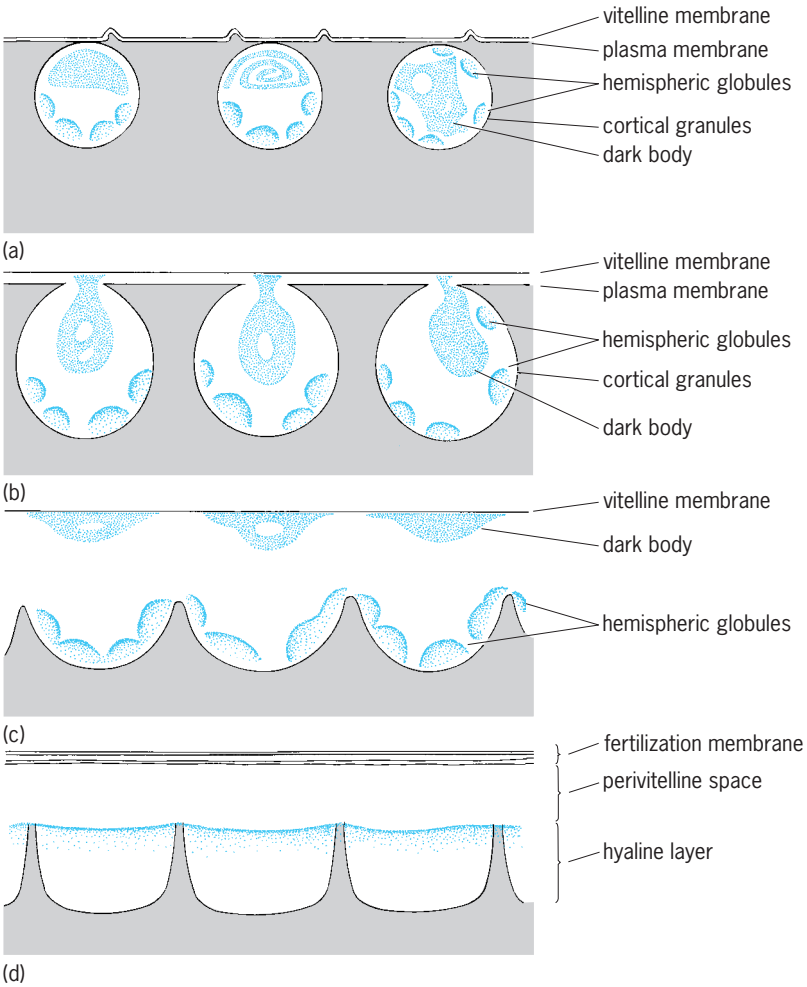

Fig. 2. Formation of the fertilization membrane in the sea urchin egg. (*a*) Surface of the unfertilized egg. (*b*) Explosion of the cortical granules: the vitelline membrane begins to be lifted up while the dark bodies are extruded, and the egg plasma membrane has become continuous with the membrane bounding the cortical granules. (*c*) The dark bodies have joined the vitelline membrane; the hemispheric globules begin to build up a layer over the new egg surface; this will then become the hyaline layer as indicated in *d*. (*d*) The dark bodies have become fused with the vitelline membrane, thus giving rise to the definitive fertilization membrane.

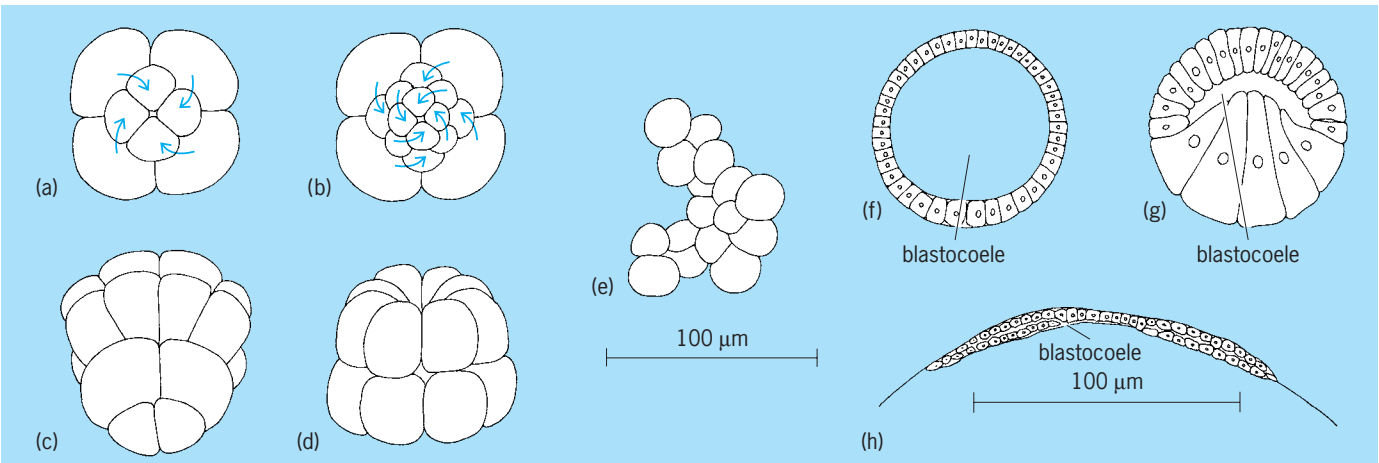

Fig. 3. Symmetry of cleavage patterns and invertebrate blastulae. (*a*, *b*) Spiral cleavage. (*c*) Bilateral cleavage. (*d*) Radial cleavage. (*e*) Irregular cleavage. (*f*) Sea urchin blastula. Relatively uniformly sized blastomeres and large blastocoele. (*g*) Annelid blastula. Blastomeres at vegetal side large, yolky; small blastocoele. (*h*) Squid blastula. Blastomeres in animal pole region only; slitlike or no blastocoele.

The union of the two pronuclei (syngamy) marks the completion of the fertilization process. The fusion forms the zygote nucleus, with the full complement of chromosomes, and the dormant egg cell has been aroused to start the series of changes which will produce a new sea urchin (**Fig. 3**).

A number of the events which have just been described are peculiar to sea urchins. With different time schedules and allowance for the individual characteristics of each species, however, these basic processes of sperm entry, aster formation, and syngamy make up the complex phenomenon of the fertilization reaction as it occurs in all animals.

Such a descriptive presentation suggests many questions about what is actually going on in terms of cellular actions and reactions. One such question, for example, concerns the nature of the contact between spermatozoon and egg surface. It is known that in many invertebrates the acrosome at the anterior tip of the sperm cell undergoes an explosive reaction at the egg surface which transforms it into a very slender filament (Fig. 1*d* and *e*). In at least some cases an enzymelike substance is also released from the acrosome which has a dissolving effect on the egg membrane. There is evidence that the penetration of the egg surface by this filament activates the egg mechanically. Whether it is also useful in drawing the spermatozoon inside the egg cytoplasm has yet to be proved.

The nature of the specificity which ensures that eggs will be entered only by spermatozoa of their own species has also been the object of a great deal of research and has not received a thoroughly satisfactory explanation. The mechanism by which an egg that has received one spermatozoon rejects all later arrivals is another problem that resists solution, but perhaps more difficult to discover than the answers to any of these questions are those concerning the cytoplasmic differences between unfertilized and fertilized eggs. *See* FERTILIZATION (ANIMAL).

Cleavage. The fertilized egg, or zygote, sets about at once to divide the huge mass of the egg into many small cells in order to restore the usual ratio between the amounts of nuclear and cytoplasmic substances. The energy for these repeated mitoses comes from the yolk, which also furnishes at least part of the materials required for synthesis of new nuclear structures. During this cleavage period, which commonly occurs during the first 12 h after fertilization, the blastomeres, as the cleavage stage cells are called, divide more or less synchronously. Generally, cleavage follows one of several patterns characteristic for large groups of animals and often correlated with the amount and mode of distribution of the yolk.

Whatever cleavage pattern is followed, the plane of the first cleavage passes through the animal pole. When the vegetal region contains a large proportion of yolk, cleavage is retarded in this area and the blastomeres tend to be larger than in the animal pole region.

Small eggs, which contain little yolk, divide completely and usually very regularly, forming a mass of cells that shows spiral (mollusks, Fig. 3*a* and *b*), bilateral (ascidians, Fig. 3*c*), or radial symmetry (echinoderms, Fig. 3*d*). Some coelenterates, however, cleave into what appear to be random masses of cells (Fig. 3*e*).

The very large eggs of squid contain a great deal of yolk concentrated at the vegetal pole. The cleavage furrows do not cut all the way through this part but restrict their activity to the living cytoplasm at the animal pole.

Insect eggs also contain a large store of yolk, which occupies the center of the elongate cells and is surrounded by a thin layer of living cytoplasm containing the egg pronucleus. Following fertilization, the nuclei alone divide and move apart in the layer of cytoplasm after each division so that they distribute themselves all around the egg. After nine such nuclear divisions have taken place (producing

512 nuclei), the cytoplasm also cleaves at the next division, forming a single layer composed of about 1000 cells surrounding the central yolk mass. *See* CLEAVAGE (DEVELOPMENTAL BIOLOGY).

Blastula stage. Among all the invertebrate forms except the insects, the result of 6–10 successive cleavage cycles is the formation of a sphere (blastula) composed of small cells which lie in a single compact layer around a central cavity (blastocoele). If the egg has contained relatively little yolk, the blastocoele is rather large (Fig. 3*f*), while it may be very small if the egg includes much yolk (Fig. 3*g*) and is little more than a slit in the squid blastula (Fig. 3*h*). *See* BLASTULATION.

Gastrula stage. The end of the brief blastula stage occurs when the process of gastrulation begins. In its simplest form, this consists in an indenting (invagination) of the blastula wall in the vegetal region (**Fig. 4***a*). Meanwhile cell division is going on steadily, and since the larva has as yet no way of taking in solid food from the outside, all the form changes which occur during this period are accomplished with the material originally present in the fertilized egg. The only addition is water (blastocoele fluid) and such dissolved substances, mostly salts, from the environment as can enter through the cell membranes. As the blastomeres become smaller and the blastular wall becomes correspondingly thinner, cells are provided to extend the vegetal indentation into a pocket (Fig. 4*b*). With the appearance of this structure (primitive digestive tract) the larva becomes two-layered, possessing an outer layer, the ectoderm, which will later produce the nervous system as well as the outermost body covering, and an inner layer, the endoderm, from which will be formed the lining of the functional digestive tract and its associated organs and glands. As the primitive digestive tract extends into the blastocoele, its opening to the outside becomes smaller and is known as the blastopore.

A modification of this process of endoderm formation occurs among some species having large, yolk-filled vegetal blastomeres (Fig. 4*c*). The small, actively dividing cells of the animal pole region spread down to cover these more inert blastomeres (Fig. 4*d*), which become the endoderm and later form the digestive organs, while the overlying ectoderm leaves a small opening in the vegetal region which corresponds to the blastopore (Fig. 4*e*). *See* GASTRULATION.

Mesoderm formation. At this time the first few cells belonging to a third body layer, the mesoderm, make their appearance by slipping from the ectoderm layer into the blastocoele. These early mesoderm cells are of a primitive sort (mesenchyme), possessing pseudopodia and often moving about freely between the ectoderm and endoderm. In sponges and coelenterates, no more highly organized middle layer is formed even in adult animals, but in the other phyla the "true" mesoderm is endodermal in origin, either being formed by successive divisions of a cell which originally belonged to the endoderm (Fig. 4*f* and *g*), as in annelids and mollusks, or separating off from

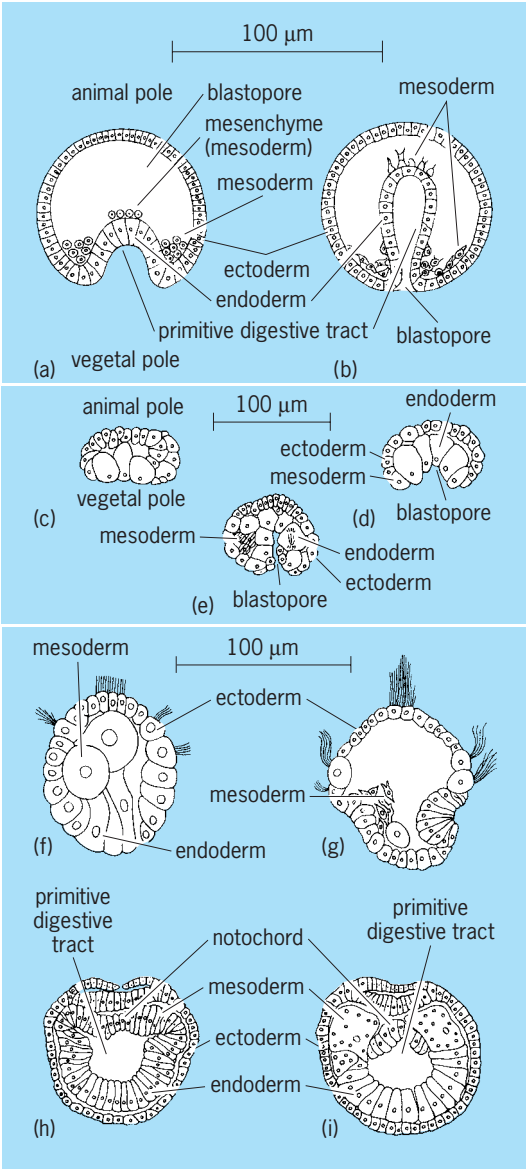

Fig. 4. Gastrulation and larval mesoderm formation. (*a*) Early and (*b*) later stage in nonyolky eggs. (*c*) Late blastula, (*d*) early gastrula, and (*e*) late gastrula in yolky egg of the snail. (*f*) Mesoderm formation in the limpet (*Patella*) shown in sections through center of blastula and (*g*) trochophore larva. (*h*) Cross section of *Branchiostoma* embryo immediately following gastrulation and (*i*) somewhat later.

the primitive digestive tract, as in *Branchiostoma* (*Amphioxus*) [Fig. 4*h* and *i*].

In either case this mesodermal tissue spreads out between the ectoderm and endoderm, and in all phyla more advanced than the flatworms it splits through its center into an inner and an outer layer. The cavity thus formed within the mesoderm is the true body cavity in which the various internal organs lie. The outer layer of mesoderm becomes closely applied to the inner side of the ectoderm, forming body-wall muscles and other supporting layers, while the inner layer of mesoderm surrounds the endoderm with layers of muscle. The organs of circulation, excretion, and reproduction, as well as all muscles and connective tissue, are eventually formed

from this mesodermal layer which surrounds the endoderm.

Later development. So far it is possible to summarize the development of invertebrate animals as a group but beyond this point each subgroup follows its own course, and these are so widely divergent that every one must be considered separately. Meaningful generalizations are not even possible within a single class in some cases, as attested to by the various modes of development occurring among the Insecta, some of which proceed directly from egg to adult form, while others go through an elaborate series of changes. *See* INSECT PHYSIOLOGY; INSECTA.

In very many species there is a sharp break in the life history when the larva, after passing through a number of morphological phases which lead from one to the next with a steady increase in size and complexity, abruptly forms a whole new set of rudimentary adult organs which take over the vital functions. This metamorphosis represents the end of the larval period. The tiny animal which it produces is for the first time recognizable as the offspring of its parents.

For more or less arbitrary reasons, the developmental processes of certain invertebrate forms have been studied very carefully so that their life histories are fully known. A few of these will be outlined in the following sections.

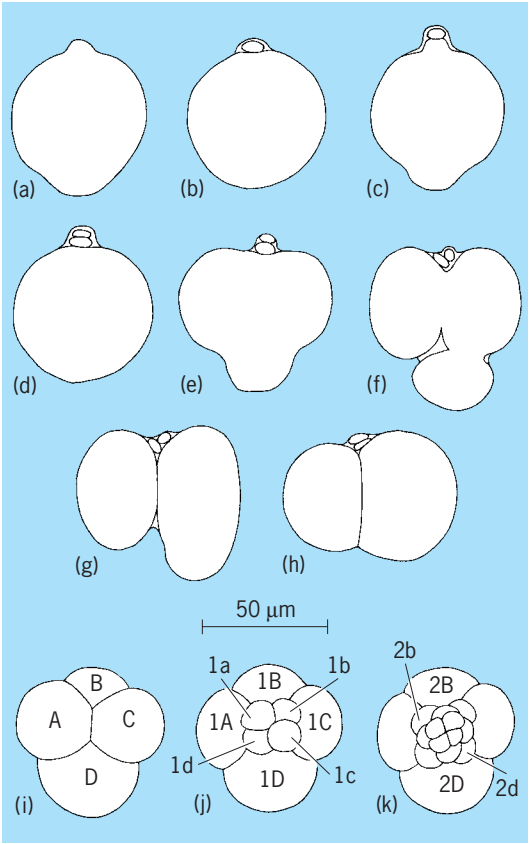

Fig. 5. Maturation and early cleavage in the mussel (*Mytilus*). (*a*) First polar body being formed at animal pole; polar lobe at vegetal pole. (*b*) First polar body completely extruded; polar lobe withdrawn. (*c*, *d*) Second polar body formation. (*e–g*) First cleavage. (*h*) Two-cell stage. (*i*) Four-cell stage. (*j*) Eight-cell stage; first quartet of micromeres (1a, 1b, 1c, and 1d). (*k*) Sixteen-cell stage with three quartets of micromeres.

Molluscan Development

The eggs of *Mytilus*, the common mussel, are fertilizable just after the germinal vesicle breaks down (Fig. 1*c*). As the first polar body is given off from the animal pole, the vegetal surface of the egg bulges out to form the so-called polar lobe (**Fig. 5***a–d*). The bulge disappears shortly, but will reappear at the time of the second polar body formation.

Cleavage. When the egg cleaves, the vegetal cytoplasm is segregated into a more extreme polar lobe (Fig. 5*e* and *f*) and the cleavage furrow divides the remaining material equally between two blastomeres. The constriction forming the polar lobe disappears, returning the polar lobe material to one of the blastomeres (Fig. 5*g* and *h*). The vegetal material is again segregated at the second cleavage and again mixed with one of the four blastomeres.

It is characteristic of this type of cleavage that the mitotic spindles lie aslant in the blastomeres and, moreover, regularly change the direction of their slant by 90° at each division so that a spiral pattern of blastomeres results. Such spiral cleavage is found in the mollusks and in the flat, round, and segmented worms. In modified form it is also found in the crustaceans.

Since the animal-vegetal axis is easy to recognize in such eggs, it has been possible to record the course of cleavage very accurately and to determine the role of particular blastomeres in normally developing embryos. The four-cell stage blastomere containing the polar lobe material is designated as D, and proceeding in a clockwise direction, the others become A, B, and C (Fig. 5*i*). At the third cleavage, these divide very unequally (Fig. 5*i* and *k*) into four large vegetal macromeres (1A, 1B, 1C, and 1D) and four micromeres at the animal side (1a, 1b, 1c, and 1d). *See* CELL LINEAGE.

Blastula stage. After two more such unequal divisions the resulting 28 micromeres have formed a hollow blastula with the four macromeres, 3A, 3B, 3C, and 3D, at its vegetal side. These then extend into the blastular cavity where their descendants will form the digestive tract, except for one of the D daughter cells produced at the next cleavage, 4d, which is set aside as the mesoderm mother cell (Fig. 4*f*).

Trochophore stage. During the succeeding cleavages some of the cells develop cilia, the blastular symmetry becomes bilateral instead of radial, and the micromeres extend down almost to the vegetal pole, thus covering the macromeres except at the small opening of the blastopore. After 24 h of development the cilia are organized into an encircling girdle and an apical tuft at the animal pole, and the larva, now called a trochophore, begins its free-swimming stage (**Fig. 6***a*). The blastopore is shifted forward by the faster proliferation of the ectodermal cells of the other side (Fig. 6*a*) and then closed, but the larval mouth is later formed at this place. Behind

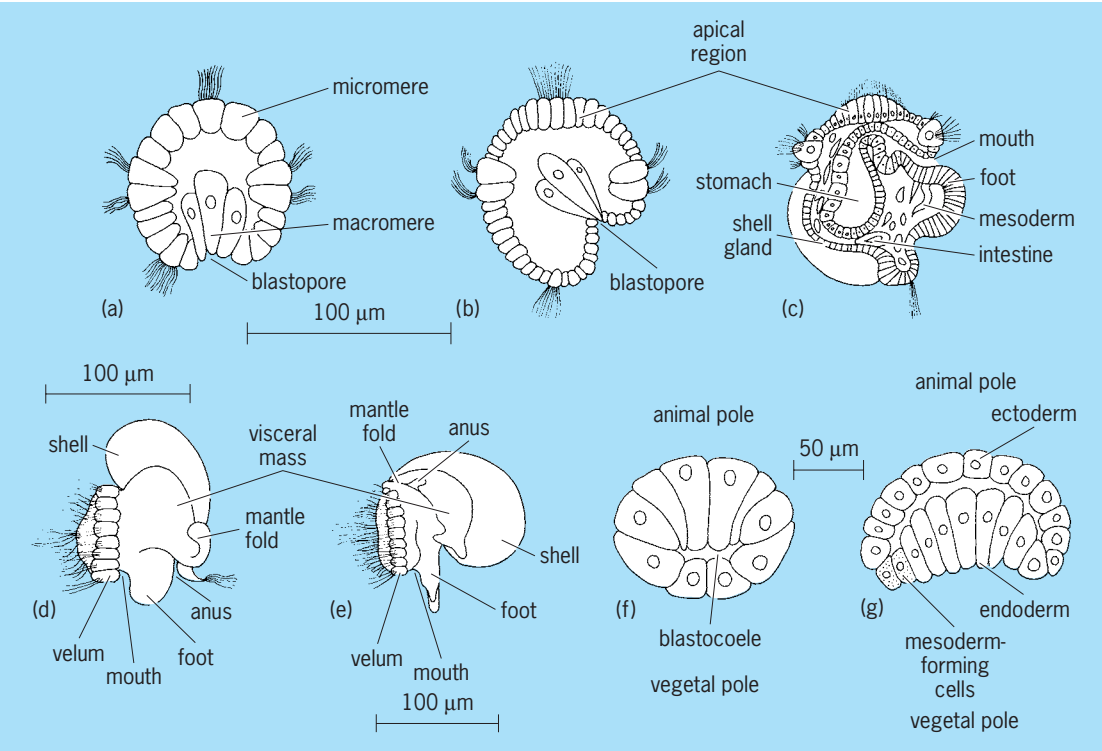

Fig. 6. Stages in the development of *Patella* and gastrulation in a tunicate. (*a*) Late blastula of *Patella*; macromeres being covered by micromeres, and blastopore at vegetal pole. (*b*) Early trochophore of *Patella*; blastopore shifted forward. (*c*) Matured trochophore larva of *Patella*. (*d*) Veliger larva of *Patella*. (*e*) Late veliger stage of *Patella*. Twisting of body has carried mantle fold and position of anus forward and turned shell about. (*f*) Late blastula of tunicate embryo. (*g*) Beginning of gastrulation in tunicate embryo.

it the endoderm forms a stomach, and a narrow tube gradually extends from this to make the intestine. The anus forms later at the place where the intestine reaches the ectoderm.

At this stage a group of ectodermal cells is forming the shell gland which will secrete the shell (Fig. 6*c*). Two small protuberances will unite and develop into the foot, and a pair of elongated pits beside the mouth will form the balancing organs. The 4d blastomere has cleaved into two cells located on either side of the mouth which are giving rise, at this stage, to two rows of mesoderm cells called the mesodermal bands.

Veliger stage. Within a week the shell gland has grown and begun to secrete the shell, and the foot is projecting prominently. The stomach increases in size and bulges into the shell cavity, and cells from the ends of the mesodermal bands form muscular attachments for the stomach and esophagus. The girdle of ciliated cells (velum) enlarges, and the rudiments of a nervous system, including eye cups, appear near the apical tuft. The larva is now called a veliger (Fig. 6*d*).

Metamorphosis. Following further development (especially of the alimentary tract, which becomes U-shaped with the mouth and anus separated from each other only by the foot), there is a sudden period of unequal growth in the two sides of the larva so that the anus is moved around to open on its neck (Fig. 6*e*). Eyes and tentacles have already been formed, and finally the young animal discards its cilium-bearing velum and takes up the adult habit of creeping about on its foot.

Sea Urchin Development

In sea urchins, the first and second cleavage planes divide the fertilized egg (Fig. 2) into four equal blastomeres, intersecting each other at right angles through the animal-vegetal axis. The third cleavage cuts through these four blastomeres horizontally. The fourth cleavage plane divides the upper four cells into eight equal-sized mesomeres, and the lower group into four very small micromeres at the vegetal pole and four large macromeres (**Fig. 7**). These 16 blastomeres are each divided about equally at the fifth cleavage, forming 32 cells, but the eight micromeres fail to divide at the sixth cleavage, giving 56 instead of 64 blastomeres at this stage. By removing certain of these blastomeres (such as the micromeres) and following later development, it has been found that the eight micromeres of 56-cell stage give rise to the first group of mesenchyme cells, the ring of eight cells just above them produces mesenchyme and endoderm, and all other blastomeres form ectoderm.

Blastulation. The blastomeres continue to divide successively, forming the hollow sphere of the blastula stage (Fig. 3*f*). By the tenth cleavage, each of the thousand or so blastomeres has developed a cilium, and the blastula has also secreted enough hatching enzyme to dissolve the fertilization membrane. After about 12 h of development at 72°F (20°C),

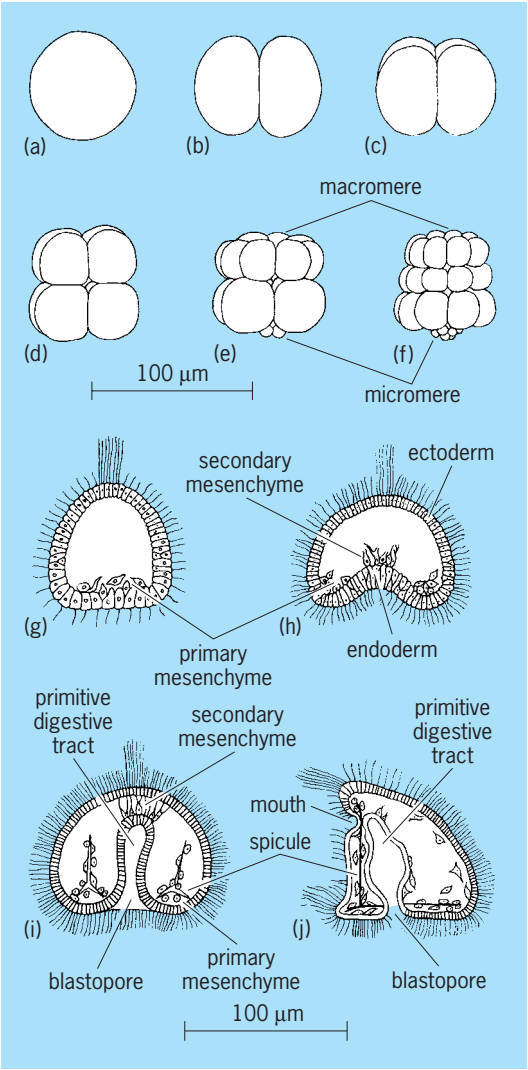

Fig. 7. Cleavage and gastrulation in the sea urchin. (*a*) Zygote. (*b*) Two-cell stage. (*c*) Four-cell stage. (*d*) Eight-cell stage. (*e*) Sixteen-cell stage. (*f*) Thirty-two–cell stage. (*g*) Vegetal region of blastula flattened; primary mesenchymal cells in blastocoele. (*h*) Beginning of invagination; secondary mesenchymal cells appearing in blastocoele. (*i*) Primitive digestive tract deepening; beginning of spicules. (*j*) Late gastrula, from side.

the larva begins its free-swimming period.

Gastrulation. Shortly afterward the larva elongates somewhat toward the animal pole, where a tuft of long, immobile cilia appears, and flattens on the vegetal side. Just before invagination begins, the cells descended from the eight micromeres slip out of the vegetal blastular wall into the blastocoele (Fig. 7*g*), forming the primary mesenchyme cells.

The gastrula stage begins about 20 h after fertilization when the center of the vegetal wall bulges inward (Fig. 7*h*). Most of the cells which form this pocket will become endoderm, but there are also some mesenchymal cells among them (secondary mesenchyme) which will develop pseudopodia that stretch across the blastocoele and, making contact with the opposite body wall, direct the inward expansion of the primitive digestive tract (Fig. 7*i*). As this deepens, the primary mesenchyme cells are building calcareous spicules with calcium taken from the seawater. These skeletal rods extend in three directions from two points beside the blastopore; as they lengthen they determine the characteristic shapes of the larval stages (Fig. 7*j*).

Pluteus stage. In the pluteus stage (**Fig. 8***a* and *b*), a mouth opening is formed where the tip of the primitive digestive tract joins the body wall. The tract begins to function, and the blastopore is changed into an anus as the larva is first able to take in food from outside.

Metamorphosis. During the month which the larva spends as a pluteus, the body increases markedly in size and two more pairs of arms are added. In preparation for metamorphosis, a structure called the echinus rudiment forms the beginnings of the adult organ systems. Within a relatively short time most of the larval body is incorporated into these organs, which are recognizable as those of a young sea urchin. Its metamorphosis is complete when it casts off the unusable parts of its larval skeleton (**Fig. 9**).

Tunicate Development

The fact that certain structures, characteristic of vertebrates and found in no other invertebrates, appear during the larval life of the tunicates forms the basis for giving these otherwise unprepossessing animals their high status at the top of the invertebrate subkingdom and makes them especially interesting from the evolutionary aspect.

Fertilization. The eggs of the tunicate *Styela* begin meiosis as they are laid, going as far as the metaphase

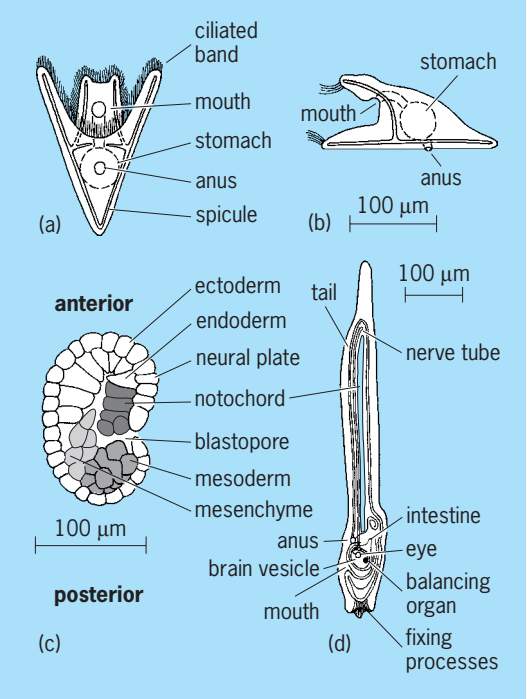

Fig. 8. Larval stages. (*a*) Ventral and (*b*) side view of pluteus larva of sea urchin. (*c*) Section through tunicate gastrula. (*d*) Tadpole stage of tunicate larva.

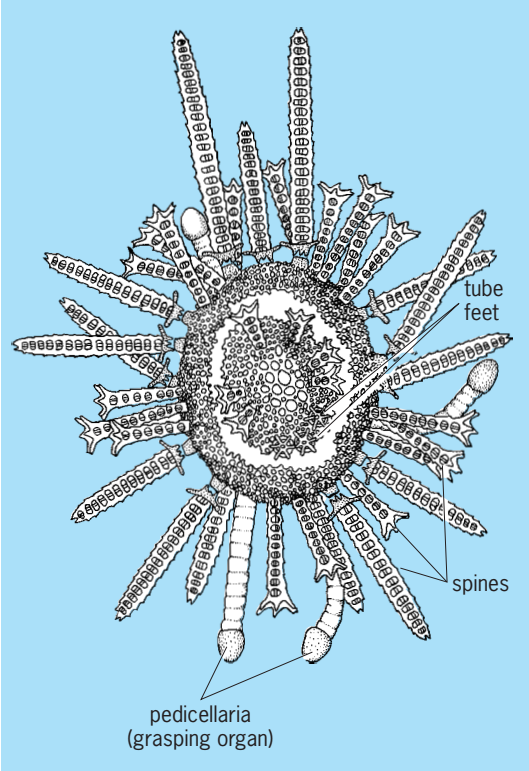

Fig. 9. Young echinoid *Peronella japonica*, after metamorphosis. (*After K. Okazaki and J. C. Dan, Metamorphosis of partial larvae of Peronella japonica Morteusen, a sand dollar, Biol. Bull., 106:83–99, 1954*)

of the first reduction division where they stop until they are fertilized. The spermatozoon penetrates the thick chorion, enters the egg at the vegetal pole, and stimulates it to proceed with meiosis. While the polar bodies are being given off, cytoplasmic streaming segregates the cell components into a yellow-pigmented region, a clear yolk-free region, and a gray yolky mass. It is possible to recognize these differently colored materials later in development and determine the role of each in body layer formation.

Cleavage. The first cleavage divides the egg into similar blastomeres. Because of the arrangement of the colored cytoplasm, its bilateral symmetry is already visible. The 16-cell stage consists of two layers of eight cells each (Fig. 3*d*), with the yellow cytoplasm contained in four of the vegetal cells. At the stage with about 40 cells a blastula is formed. The prospective ectoderm making up its animal side consists of thick columnar cells, while the future endoderm cells at the vegetal side are relatively flat (Fig. 6*f*). This difference is reversed before gastrulation begins (Fig. 6*g*).

Gastrulation. The gastrula is formed by the movement into the blastocoele of the vegetal-side cells, followed by an overlapping growth of the prospective ectoderm. Within this enveloping layer the yellow cells produce mesoderm; the other vegetal cells form endoderm. As the gastrula develops, the surface layer anterior to the blastopore (Fig. 8*c*) forms neural tissue which is organized into a brain and spinal cord, while the mesoderm beneath it forms a notochord, a precursor of the vertebral column characteristic of vertebrate animals. This notochord elongates as the axis of a tail, and the larva hatches from its chorion and begins a free-swimming stage.

Tadpole stage. During this stage (Fig. 8*d*), the tadpole acquires an extensive but nonfunctional digestive tract, two pairs of gill slits (also characteristic of vertebrates), a "cerebral eye," and a balancing organ. At its anterior end it has three papillae with which it will fix itself to a substratum when its short tadpole stage ends.

Metamorphosis. When metamorphosis begins (**Fig. 10**), the tail ectoderm contracts strongly, bending and breaking up the notochord, nerve cord, and tail muscles which are consumed by phagocytes. The "chin" region next to the organ of fixation elongates greatly, carrying the mouth upward. A new nervous system replaces the larval one. The intestine elongates, fuses with the ectoderm to open as an anus, and forms a stomach and a liverlike gland.

A circulatory system which was started during the tadpole stage develops a muscular heart. Four new pairs of gill slits open into the pharynx. These later divide and give rise to further rows of smaller slits. The reproductive organs are formed from two masses of mesoderm cells lying near the pharynx. These develop into an ovary and a testis. Gertrude Hinsch

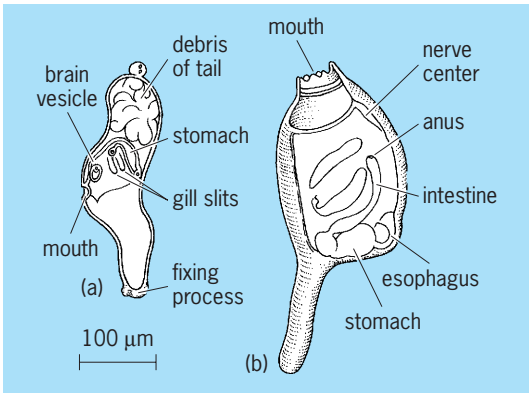

Fig. 10. Metamorphosis in tunicate. (*a*) Beginning of metamorphosis. (*b*) Young adult.

Involute

A term applied to a curve C' that cuts at right angles all tangents of a survey C (see **illus.**). Each curve C has infinitely many involutes and the distance between corresponding points of any two involutes is constant. If $x_i = x_i(s)$, with $i = 1, 2, 3$, are parametric equations of C, with parameter s arc length on C, all involutes C' of C have parametric equations $X_i = x_i(s) + (k - s)x_i'(s)$, with $i = 1, 2, 3$, where $x_i' = dx_i(s)/ds$, with $i = 1, 2, 3$, and k denotes an arbitrary constant. Let a length of string be coincident with a curve C, with one end fastened at a point P_0 of C. If the string is unwound, remaining taut, the other end

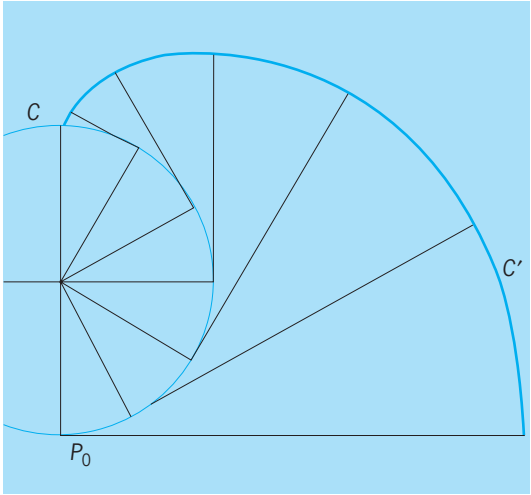

An involute *C′* of a curve *C*.

of the string traces an involute C' of C. By varying the length of the string, all involutes of C are obtained. *See* ANALYTIC GEOMETRY. Leonard M. Blumenthal

Iodine

A nonmetallic element, symbol I, atomic number 53, relative atomic mass 126.9045, the heaviest of the naturally occurring halogens. Under normal conditions iodine is a black, lustrous, volatile solid; it is named after its violet vapor. *See* HALOGEN ELEMENTS; PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

The chemistry of iodine, like that of the other halogens, is dominated by the facility with which the atom acquires an electron to form either the iodide ion I^- or a single covalent bond —I, and by the formation, with more electronegative elements, of compounds in which the formal oxidation state of iodine is +1, +3, +5, or +7. Iodine is more electropositive than the other halogens, and its properties are modulated by: the relative weakness of covalent bonds between iodine and more electropositive elements; the large sizes of the iodine atom and iodide ion, which reduce lattice and solvation enthalpies for iodides while increasing the importance of van der Waals forces in iodine compounds; and the relative ease with which iodine is oxidized, Some properties of iodine are listed in the **table**. *See* ASTATINE; BROMINE; CHEMICAL BONDING; CHLORINE; FLUORINE.

Some important properties of iodine

Property	Value
Electronic configuration	[Kr]4d105s25p5
Relative atomic mass	126.9045
Electronegativity (Pauling scale)	2.66
Electron affinity, eV	3.13
Ionization potential, eV	10.451
Covalent radius,—I, nm	0.133
Ionic radius, I_-, nm	0.212
Boiling point, °C	184.35
Melting point, °C	113.5
Specific gravity (20/4)	4.940

Iodine occurs widely, although rarely in high concentration and never in elemental form. Despite the low concentration of iodine in sea water, certain species of seaweed can extract and accumulate the element. In the form of calcium iodate, iodine is found in the caliche beds in Chile. Iodine also occurs as iodide ion in some oil well brines in California, Michigan, and Japan.

The sole stable isotope of iodine is ^{127}I (53 protons, 74 neutrons). Of the 22 artificial isotopes (masses between 117 and 139), the most important is ^{131}I, with a half-life of 8 days. It is widely used in radioactive tracer work and certain radiotherapy procedures. *See* RADIOACTIVE TRACER.

Iodine exists as diatomic I_2 molecules in solid, liquid, and vapor phases, although at elevated temperatures (>200°C or 390°F) dissociation into atoms is appreciable. Short intermolecular I. . . I distances in the crystalline solid indicate strong inter-molecular van der Waals forces. Iodine is moderately soluble in nonpolar liquids, and the violet color of the solutions suggests that I_2 molecules are present, as in iodine vapor.

Although it is usually less vigorous in its reactions than the other halogens, iodine combines directly with most elements. Important exceptions are the noble gases, carbon, nitrogen, and some noble metals. The inorganic derivatives of iodine may be grouped into three classes of compounds: those with more electropositive elements, that is, iodides; those with other halogens; and those with oxygen. Organoiodine compounds fall into two categories: the iodides; and the derivatives in which iodine is in a formal positive oxidation state by virtue of bonding to another, more electronegative element. *See* GRIGNARD REACTION; HALOGENATED HYDROCARBON; HALOGENATION.

Iodine appears to be a trace element essential to animal and vegetable life. Iodide and iodate in sea water enter into the metabolic cycle of most marine flora and fauna, while in the higher mammals iodine is concentrated in the thyroid gland, being converted there to iodinated amino acids (chiefly thyroxine and iodotyrosines). They are stored in the thyroid as thyroglobulin, and thyroxine is apparently secreted by the gland. Iodine

deficiency in mammals leads to goiter, a condition in which the thyroid gland becomes enlarged. *See* THYROID GLAND.

The bactericidal properties of iodine and its compounds bolster their major uses, whether for treatment of wounds or sterilization of drinking water. Also, iodine compounds are used to treat certain thyroid and heart conditions, as a dietary supplement (in the form of iodized salt), and for x-ray contrast media. *See* ANTIMICROBIAL AGENTS; ANTISEPTIC; SALT (FOOD).

Major industrial uses are in photography, where silver iodide is a constituent of fast photographic film emulsions, and in the dye industry, where iodine-containing dyes are produced for food processing and for color photography. *See* DYE; PHOTOGRAPHIC MATERIALS. Chris Adams

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; A. Varvoglis, *The Organic Chemistry of Polycoordinated Iodine*, 1992; T. Wirth (ed.), *Hypervalent Iodine Chemistry*, 2003.

Ion

An atom or group of atoms that bears an electric charge. Positively charged ions are called cations, and negatively charged ions are called anions. When a single atom gains or loses an electron, monoatomic ions are formed. For example, reaction of the element sodium (Na) with the element chlorine (Cl), which occurs as the diatomic gas Cl_2, leads to the transfer of electrons from Na to Cl to form Na^+ cations and Cl^- anions, as shown in the reaction below.

$$2Na + Cl_2 \rightarrow 2Na^+ + 2Cl^-$$

Removal of an electron from an atom generates a species that has one more proton than electrons. For example, a sodium atom contains 11 negatively charged electrons and 11 positively charged protons. If one electron is removed, there will be 10 electrons and 11 protons, generating a positive charge on the sodium ion. Likewise, addition of an electron to an atom generates a species with more electrons than protons, which is therefore a negatively charged anion. In general, atoms of metallic elements (on the left side of the periodic table) lose electrons to form cations, while atoms of nonmetallic atoms (on the right side of the periodic table) gain electrons to form anions. Ions can bear multiple charges, as in the magnesium ion (Mg^{2+}) or the nitride ion (N^{3-}). The charge on monoatomic ions is usually the same for elements in the same column of the periodic table; for example, hydrogen (H), Na, lithium (Li), potassium (K), rubidium (Rb), and cesium (Cs) all form +1 ions. *See* PERIODIC TABLE.

Ions can also comprise more than one atom and are then called polyatomic ions. For example, the ammonium ion (NH_4^+) carries a positive charge and is composed of one nitrogen atom and four hydrogen atoms. The nitrate ion (NO_3^-) is composed of one nitrogen atom and three oxygen atoms and carries a single negative charge. Polyatomic ions are usually depicted inside brackets with superscripted charges, as shown in the structure below.

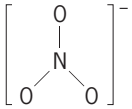

Nitrate ion

Anions and cations can combine to form solid materials called salts, which are named by the cation name followed by the anion name. For a salt composed of the polyatomic ions ammonium and nitrate, the formula is NH_4NO_3 and the name is ammonium nitrate. For monoatomic ions, the cation name is the same as the element and the anion name is the element name with the ending -ide. Thus, common table salt, NaCl, is called sodium chloride. The ratio of anions to cations must always be such that an electrically neutral material is produced. Thus, magnesium nitrate must contain one magnesium for every two nitrates, giving the formula $Mg(NO_3)_2$. *See* SALT (CHEMISTRY). H. Holden Thorp

Ion beam mixing

A process in which bombardment of a solid with a beam of energetic ions causes the intermixing of the atoms of two separate phases originally present in the near-surface region. In the well-established process of ion implantation, the ions are incident instead on a homogeneous solid, into which they are incorporated over a range of depths determined by their initial energy. In the simplest example of ion beam mixing, the solid is a composite consisting of a substrate and a thin film of a different material (**Fig. 1***a*). Ions with sufficient energy pass through the film into the substrate, and this causes mixing of the film and substrate atoms through atomic collision processes (Fig. 1*b*). If the ion dose is large enough, the original film will completely disappear (Fig. 1*c*). This process may result in the impurity doping of the substrate, in the formation of an alloy or two-phase mixture, or in the production of a stable or metastable solid phase that is different from either the film or the substrate. *See* ION IMPLANTATION.

Like ion implantation, ion beam mixing is a solid-state process that permits controlled change in the composition and properties of the near-surface region of solids. Although still being exploited experimentally and not yet employed commercially, it is expected to be useful for such applications as surface modification of metals and semiconductor device processing. Along with thin-film deposition technology, ion beam mixing should make it possible to introduce many impurity elements at concentrations that would be too high for ion implantation to be practical.

Mixing mechanisms. The ion beam mixing of a film and substrate takes place by the following

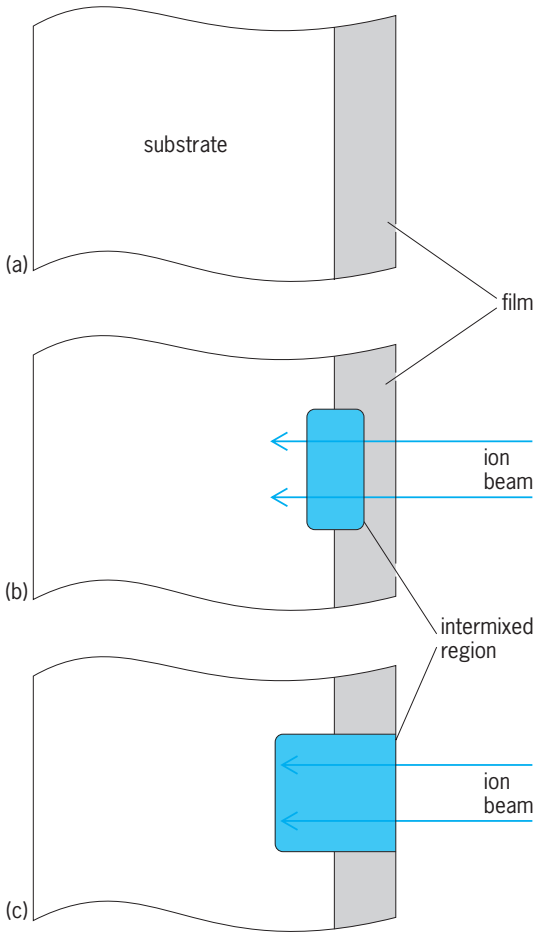

Fig. 1. Ion beam mixing of film and substrate. (*a*) Before ion bombardment. (*b*) Partial intermixing. (*c*) Complete intermixing.

mechanisms, in increasing order of complexity: recoil implantation, cascade mixing, and radiation-enhanced diffusion. In recoil implantation, an atom is driven from the film into the substrate as a result of a direct collision with an incident ion. Cascade mixing takes place when the recoil of an atom from such a collision initiates a series of secondary collisions among the film and substrate atoms. These collisions produce both forward and backward displacements and therefore lead to transfer of atoms from the substrate into the film as well as from the film into the substrate. Both recoil implantation and cascade mixing are kinematic mechanisms whose effectiveness generally increases with the mass, energy, and charge state of the incident ions and with the masses of the film and substrate atoms but is relatively insensitive to temperature. These phenomena usually yield a ratio of intermixed atoms to incident ions no greater than about 10 to 1. Since they are coincident with the atomic collisions, they occur in only about 10^{-11} s, roughly the transit time of the incident ions through the film of about 100 nm in thickness. Mixing of the film and substrate atoms by the mechanism of radiation-enhanced diffusion requires much longer times, and the rate is strongly temperature-dependent because diffusion is a thermally activated process. This mechanism is a consequence of the increase in interdiffusion coefficients that results because additional lattice defects (of which the simplest are vacancies and interstitial atoms) are formed by the atomic displacements produced by ion bombardment. Intermixing of several hundred atoms per incident ion can be achieved by heating a bombarded sample to temperatures where the defects become sufficiently mobile for rapid diffusion. *See* ION-SOLID INTERACTIONS; RADIATION DAMAGE TO MATERIALS.

Stable- and metastable-phase formation. When a sample is subjected to ion beam mixing, the nature of the solid phase obtained depends upon the composition of the intermixed region. For samples initially consisting of a thin film on a substrate, the composition obtained by complete mixing is determined by the thickness of the film. An alternate sample configuration is prepared by depositing alternating thin films of different elements on a substrate (**Fig. 2**), and passage of the energetic ions homogenizes the films on an atomic scale. For this configuration the intermixed composition can be adjusted by varying the relative thickness of the different films.

If the composition produced by ion beam mixing corresponds to a solid phase that is thermodynamically stable, this is the phase that will be obtained. Such a phase could be a solid solution of one constituent in the other, or it could be a chemical compound of the two constituents. In addition, ion beam mixing can also produce metastable phases, because it is a highly nonequilibrium process in which the kinetic energy of the bombarding ions is transferred to the target atoms and then very rapidly dissipated by atomic collision and rearrangement. For example, this technique has been used to prepare a continuous series of silver-copper solid solutions, in which the lattice constant varies almost linearly with alloy composition, and the solid solubility of nickel in silver has been increased to more than 40 times the equilibrium value. Splat cooling, the extremely rapid solidification of a liquid phase, is another nonequilibrium technique for producing metastable materials such as supersaturated solid solutions or amorphous phases. However, single-phase solid compositions can be achieved only to the extent that the constituents are miscible in the liquid phase. Thus in the silver-nickel system the effectiveness of splat cooling in increasing the solid solubility is limited

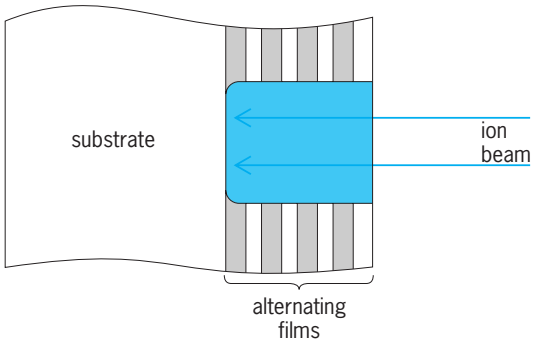

Fig. 2. Ion beam mixing of alternating films.

by the immiscibility in the liquid phase. As a solid-state process, ion beam mixing is not subject to this limitation.

Applications. While ion implantation is a well-established technique for modifying the near-surface physical and chemical properties of metals in order to improve such characteristics as wear and corrosion resistance, this technique has the disadvantage that relatively high ion doses may be necessary for effective surface modification. Thus doses of 10^{17}–10^{18} cm^{-2} are usually required to achieve impurity concentrations of 1–10 at. % in the near-surface region of about 100 nm in depth. To obtain such doses in a reasonable length of time requires high beam intensities that may be difficult to achieve for metallic elements, which are generally the ones used for direct implantation. This disadvantage can be overcome by the use of ion beam mixing, not only because the ratio of intermixed ions to bombarding ions is much greater than one, but also because the bombarding species can be ions of gases such as nitrogen and argon, for which high beam intensities are easily produced in long-life sources that are commercially available. For example, in initial demonstrations of surface modification by ion beam mixing, the wear resistance of titanium alloys has been substantially improved by the intermixing of a tin film bombarded with nitrogen ions, and the corrosion resistance of copper has been improved by the intermixing of a chromium film bombarded with argon ions.

In silicon integrated circuits, metal silicides such as titanium silicide ($TiSi_2$) and tungsten silicide (WSi_2) are widely used for low-resistance contacts and interconnects. These materials have conventionally been formed by deposition of a metal film on the silicon wafer and subsequent thermal reaction between the metal and silicon. Ion beam mixing offers the advantage of reducing the temperature required for silicide formation, since penetration of the bombarding ions through the metal-silicon interface disperses interfacial impurities (particularly native oxides of silicon), which generally form a diffusion barrier that impedes reaction. Silicides formed by ion beam mixing exhibit good surface morphology and electrical properties. *See* INTEGRATED CIRCUITS. Bor-Yeu Tsaur

Bibliography. R. E. Benenson et al., *Ion Beam Modification of Materials*, 1981; J. J. Cuomo, S. M. Rossnagel, and H. R. Kaufman (eds.), *Handbook of Ion Beam Processing Technology*, 1989; G. K. Hubler et al., *Ion Implantation and Ion Beam Processing of Materials*, 1984; F. Komarov, *Ion Beam Modification of Metals*, 1992; S. T. Picraux and W. J. Choyke, *Metastable Materials Formation by Ion Implantation*, 1982.

Ion exchange

The reversible exchange of ions of the same charge between a solution and an insoluble solid in contact with it; or between two immiscible solvents, one of which contains a soluble material with immobilized ionic groups. Ions are atoms or molecules containing charge-bearing groups. Their interactions are dominated by the electrostatic forces between charge centers. These interactions are attractive when the ions are of opposite charge, or repulsive when the ions have the same charge. Ions with a net negative charge are called anions, and those with a net positive charge are cations.

A unique property of ions is their capacity to render gases and liquids conducting, and conductivity is a universal method of detecting ions. Ions in solution are in rapid motion and have no distinct partners. Ions in an electric field migrate to the electrode of opposite charge with a velocity roughly proportional to their charge-to-size ratio. This process is known as electrophoresis, and it is one method used to separate and identify ions. *See* ELECTROPHORESIS.

Ions can also be separated on the basis of their equilibrium with a system containing immobilized ions of opposite charge. Ions can be immobilized by virtue of their location in a rigid matrix. Associated with these fixed ionic sites are mobile counterions of opposite charge. Solution ions with a higher affinity than the counterions for the fixed sites will displace them from the fixed sites and remain localized in the vicinity of the fixed sites. Simultaneously the solution is enriched in the counterions originally localized at the fixed sites. This exchange process for ions of the same charge type is called ion exchange. In a column containing the immobilized ions as part of the stationary phase and the solution of competing ions as the mobile phase, the sample ions can be separated by the repeated equilibrium steps involved as they are transported through the column until they exit it, and are detected. This is an example of ion-exchange chromatography, an important method of separating and identifying ions.

Materials

Ion exchange was first recognized over a hundred years ago, when it was found that certain clay minerals in soil can remove potassium and ammonium ions from water, with the release of an equivalent amount of calcium ions. Since that time many naturally occurring ion-exchange substances have been identified, most of them inorganic substances. Also, attention shifted from naturally occurring inorganic ion exchangers to synthetic organic polymers for most practical uses of the ion-exchange process. These materials can be prepared in large quantities with physical and chemical properties custom-made for specific applications.

Ion-exchange polymers. Phenol-formaldehyde condensation polymers have largely been replaced by materials based on styrene and divinylbenzene and, to a lesser extent, polymers prepared from divinylbenzene, or a similar cross-linking agent, and acrylic, methacrylic, or hydroxyalkyl methacrylic acids and esters. These are usually prepared in bead form by using the suspension process in which the monomers, catalysts, surfactant, and processing aids are suspended as oil droplets, which after reaction separate from the solvent as spherical beads with a

narrow size distribution. By varying the experimental conditions and concentration of reagents, beads with a wide range of physical properties can be prepared. Microporous beads consist of a core of entangled polymer chains and are essentially solid beads with narrow pores possessing a small surface area. Macroporous beads are prepared by bead polymerization in the presence of a solvent that is a good solvent for the monomers but a poor swelling agent for the polymer. In this case, each polymer bead is formed from many microbeads joined together during the polymerization process to create a network of holes and channels. The macroporous beads have greater mechanical stability, permeability, and surface area than the microporous beads.

Ion exchangers prepared for the isolation or separation of cations must have negatively charged functional groups incorporated into the polymer backbone. The most common groups are sulfonic and carboxylic acids. Sulfonic acid groups are introduced by reacting the polymer beads with fuming sulfuric acid or a similar reagent. Similarly, carboxylic acid groups can be introduced by a number of common chemical reactions or by hydrolysis of the ester group or oxidation of hydroxyalkyl groups in methyl methacrylate or hydroxyalkyl methacrylate polymers, respectively. Other common functional groups used in cation exchangers include phosphoric acid and phenol and, to a lesser extent, phosphinic, arsonic, and selenonic acids. Specialty polymers containing functional groups such as 8-hydroxyquinoline or sulfoguanidine [$—CH_2S(:NH)NH_2$] are prepared for the isolation of metals by chelation.

A common approach for the preparation of anion exchangers is to react the styrene-divinylbenzene polymer with chloromethylmethyl ether in the presence of a catalyst, which adds the side chain, $—CH_2Cl$; then this chloromethylated product is treated with an amine to introduce the charged functional group. A tertiary amine produces a quaternary ammonium group, while primary and secondary amines give products that are charged only in contact with solutions of low pH. As well as simple alkyl and benzyl amines, hydroxyalkyl amines are used to introduce functional groups of the type [$—CH_2N(CH_3)_2C_2H_4OH]^+$. *See* QUATERNARY AMMONIUM SALTS.

Tentacle ion exchangers consist of an insoluble matrix, copolymerized from oligoethyleneglycol, glycidylmethacrylate, and pentaerythroldimethacrylate, to which are grafted polymerized chains of acrylamide derivatives that are approximately 15–50 units in length. An average of 18 charged groups is covalently bound to each of these tentacles. This arrangement markedly reduces the contact between the analyte and the matrix, thus suppressing nonspecific support interactions, while also allowing ionic interactions between the ion exchanger and the analyte that are sterically impossible with conventional fixed-site ion exchangers. These materials are used for the isolation of biopolymers without loss of activity.

Ion exchangers with a low exchange capacity and fast mass-transfer properties are required for ion chromatography that uses conductivity detection. The time and temperature of the chemical reactions used to introduce the ionic functional groups into the polymer are controlled such that the reactions are limited to the external surface of the beads. Surface-functionalized anion-exchange polymer beads containing quaternary ammonium ions are difficult to prepare directly. An alternative approach is the preparation of agglomerated anion-exchange beads, prepared by taking surface-sulfonated beads and coating them with a layer of quaternary ammonium functionalized latex particles. The core particles are 5–30 micrometers in diameter, and the latex particles are much smaller at about 0.6–0.9 nanometer. The negatively charged sulfonate groups attract the positively charged quaternary ammonium ions and bind the latex particles very strongly to the surface of the polymer bead. Since only a small fraction of the quaternary ions on the latex particles are involved in binding of the anion-exchange layer, the remaining groups are free to function as an anion exchanger. This type of structure creates a thin anion-exchange pellicular layer within which the ion-exchange separation takes place, and an oppositely charged underlayer that excludes anions from the core by repulsion.

Silica-based materials. These are used primarily in chromatography because of the favorable mechanical and physical properties of the silica (SiO_2) gel support matrix. Ion-exchange groups are introduced by reacting the surface silanol groups of the porous silica particles with silanizing reagents containing the desired functional group (R). This produces a siloxane-type (Si—O—Si—R) bonded phase, where the bonded layer can be monomolecular or polymeric depending on the choice of reagent and experimental conditions. Typical R groups include propylsulfonic acid [$—(CH_2)_3SO_3^-$], benzenesulfonic acid [$—C_6H_4SO_3^-$], butyric acid [$—(CH_2)_3CO_2^-$], propylbenzoic acid [$—(CH_2)_3C_6H_4CO_2^-$], aminopropane [$—(CH_2)_3NH_2$], and dimethylaminopropane [$—(CH_2)_3N(CH_3)_3^+$]. Alternatively, ion exchangers based on silica and other inorganic oxides can be prepared by polymer encapsulation. In this approach a prepolymer such as poly(butadiene-maleic acid) is deposited as a thin film over the porous support and is subsequently immobilized by radical-induced in-place cross-linking of the polymer chains. This creates an ion-exchange polymer as a thin skin covering all or a sizable part of the support surface.

Hydrous oxides. Hydrous oxides of elements of groups 14, 15, and 16 of the periodic table can be used as selective ion exchangers. The most important hydrous oxides used for the separation of organic and inorganic ions are alumina ($Al_2O_3 \cdot nH_2O$), silica ($SiO_2 \cdot nH_2O$), and zirconia ($ZrO_2 \cdot nH_2O$). Silica, by virtue of the presence of surface silanol groups, is used as a cation exchanger at pH > 2. Alumina is amphoteric and can be used as an anion

exchanger at low pH and a cation exchanger at high pH. Alumina has the advantage over silica of being chemically stable over a wide pH range. The ion-exchange capacity of silica and alumina is controlled by the pH of the solution in contact with the oxides, since this controls the number of ionized surface functional groups. Alumina is used to isolate nitrogen-containing drugs and biochemically active substances from biological fluids, thus minimizing matrix interferences in their subsequent chromatographic analysis. Group 1 and 2 cations can be separated by chromatography on silica, with aqueous lithium salts as eluants. Zirconium phosphate is not a hydrous oxide, but it has many properties in common with the hydrous oxides. It has a high selectivity for the ions rubidium (Rb^+), cesium (Cs^+), strontium (Sr^{2+}), and barium (Ba^{2+}) and is used to remove long-lived $^{90}Sr^{2+}$ and $^{137}Cs^+$ from radioactive wastes. *See* PERIODIC TABLE; PH; RADIOACTIVE WASTE MANAGEMENT.

Synthetic carbohydrate polymers. Ion-exchange polymers based on cross-linked dextrans, agarose, and cellulose are used primarily in biochemistry for the isolation and separation of polyelectrolytes such as proteins, enzymes, and nucleotides. The ion-exchange groups are introduced into the hydrophobic matrix by reacting a suitable reagent with the hydroxyl groups on the saccharide subunits. Typical functional groups include carboxymethyl, phosphoric acid, propylsulfonate, diethylaminoethyl, and quaternary amines.

Characteristic Properties

Ion exchangers are broadly categorized as belonging to one of two groups. Strong ion exchangers, such as those containing sulfonate or quaternary amine groups, remain ionized over the full range of normal operating conditions. The ionization of weak ion exchangers, such as those containing carboxyl or secondary amine groups, depends on the pH of the medium. Their degree of ionization governs their ion-exchange capacity.

Capacity. The capacity of an ion exchange is the quantity of exchangeable ions per unit quantity of exchanger. It is usually expressed in gram-equivalents of ions per kilogram of exchanger, or more conveniently, milliequivalents per gram (meq/g) on a dry basis. General-purpose polymeric ion exchangers have capacities of about 3–5 meq/g. For ion chromatography a high capacity is undesirable because of the need to separate ions with eluants of low ionic strength to facilitate their detection by conductivity, and typical exchange capacities are in the range 0.001–0.1 meq/g.

Volume. The amount of cross-linking reagent used in the preparation of bead polymers controls the extent to which the dry polymer swells when placed in solution, the speed with which ions can diffuse in and out of the ion exchanger, and to some extent the selectivity of the ion exchanger for ions of different charge-to-size ratios. Styrene-divinylbenzene polymers containing less than about 8 mole % of divinylbenzene are classified as soft. When suspended in a solvent, the dry polymer absorbs the solvent, changing its volume. The observed change in volume depends on properties of the solvent as well as those of the polymer. When the polymer comes into contact with solvent, the outermost functional groups are solvated, and the randomly arranged polymer chains unfold to accommodate the larger solvated ions. A very concentrated internal solution of fixed ions and counterions exists, and the mobile counterions tend to diffuse out of the exchanger to be replaced by solvent in an attempt to reduce the difference in ionic strength between the internal and external solutions. This generates the swelling pressure, which may be as high as several hundred atmospheres for a high-capacity exchanger. The cross-linking of the polymer chains resists this swelling with different degrees of success that depend on structural and environmental factors. High-pressure liquid chromatography requires materials that are rigid and do not swell or change volume in different solvents. A high level of cross-linking is invariably required in this case.

Size. Polymer beads are available in a wide range of sizes and in narrow size ranges. Large beads (2–0.1 mm in diameter) are used in industrial processing, beads of intermediate size (0.2–0.04 mm) for sample processing and chromatography at limited pressures, and small beads (0.04–0.004 mm) for high-performance liquid chromatography at high pressures. Silica-based materials are available in a similar size range and are widely used in high-pressure liquid chromatography. The carbohydrate-based polymers are generally unstable to high pressures and are usually available in the large and intermediate size ranges only. The superior mass-transfer and favorable flow anisotropy properties of small-diameter particles are essential for modern liquid chromatography. Diffusion of ions through polymer gels is perhaps only one-tenth as fast as in liquids, and the only practical way to reduce this contribution to band broadening in chromatography is to minimize the particle diameter to reduce the average path length for migration of ions through the beads. In the surface-functionalized polymers used in ion chromatography, the migration distance for the exchangeable ions is restricted to the thickness of the active layer at the bead surface, which is only a few micrometers thick. From the perspective of performance characteristics, the silica-based materials are usually better, but the difference in properties between the polymer bead exchangers and the silica-based materials is no longer as significant as before.

Stability and adsorption. The stability and adsorption properties of the matrix are important in selecting an ion exchanger. Silica-based materials, for example, are restricted to use within the pH range 2–8 to avoid dissolution of the silica backbone at high pH and hydrolysis of the surface-bonded siloxane ligands at lower pH. Poly(styrene-divinylbenzene) polymers can be used at all pHs. For biopolymers the choice of matrix may be critical to the recovery of materials with preservation of biological activity. *See* LIGAND.

Applications

Ion exchange has numerous applications for industry and for laboratory research. By the quantity of materials used, water conditioning is the most important. Ion exchange is one of the primary analytical methods used to identify and quantify the concentration of ions in a wide range of environmental, biological, and industrial samples.

Water softening. Natural water from rivers and wells is never pure; it is usually hard, that is, it contains calcium and magnesium salts that form curds with soap and leave hard crusts in pipes and boilers. Hard water is softened by passage through a cartridge or bed of cation exchanger in the sodium form (the mobile counterions are sodium in this case). The calcium and magnesium ions have a higher affinity for the fixed ion-exchange sites than sodium ions and replace the sodium ions on the exchanger. The sodium ions do not cause hardness, so their increased concentration in the conditioned water is not a problem. Eventually the cation exchanger becomes exhausted, that is, fully loaded with magnesium and calcium ions, and must be regenerated or replaced. The cation exchanger is regenerated by passing a concentrated solution of sodium chloride through the bed of ion exchanger. *See* WATER SOFTENING.

Deionized water. Many industrial and laboratory processes require a supply of pure water with a very low concentration of salts. This can be achieved by passing water through a bed of mixed strong cation exchanger in the hydrogen form and a strong anion exchanger in the hydroxide form. The cation exchanger removes all the cations from the water by replacing them by hydrogen ions. The anions are removed by the anion exchanger and replaced by hydroxide ions. The hydrogen and hydroxide ions combine to form water. The purity of the water can be monitored by its conductivity. Inclusion of a cartridge containing carbon to adsorb neutral organic compounds produces a reliable supply of very pure water containing only parts-per-billion levels of ions and organics. Mixed-bed ion exchangers tend to be more efficient, but they are difficult to regenerate. The series coupling of cation-exchange and anion-exchange cartridges allows for ease of regeneration (**Fig. 1**).

Environmental reclamation. Toxic ions such as mercury (Hg^{2+}), lead (Pb^{2+}), chromate (CrO_4^{2-}), and ferrocyanide $(Fe(CN))_6^{4-}$ are removed by ion exchange from industrial wastewaters prior to their discharge into the environment. Ion exchangers are used to recover precious metals such as gold (Au^+), platinum (Pt^+), and silver (Ag^+) in a useful form from mine workings and metalworking factories. Ion exchange is frequently used to decontaminate waste and concentrate radioactive elements from the nuclear industry. Uranium can be recovered from low-grade ores by leaching with dilute sulfuric acid, then absorbing the uranium sulfate complex ions on a strong anion exchanger, which has a high affinity for uranium sulfate. *See* HYDROMETALLURGY; SOLUTION MINING; URANIUM METALLURGY.

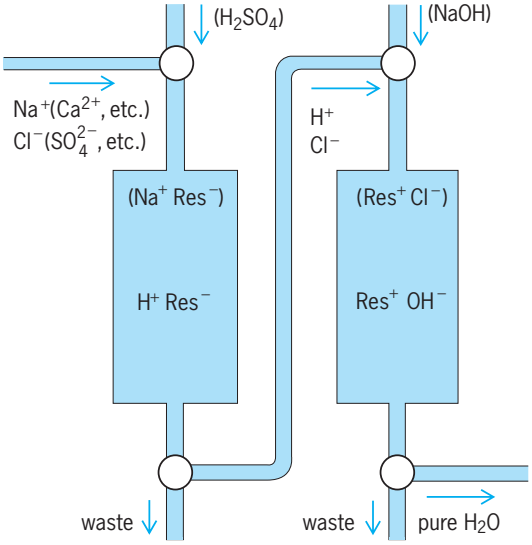

Fig. 1. Two-stage deionization by ion exchange. Valves are for regeneration with sulfuric acid (H_2SO_4) and sodium hydroxide (NaOH).

Chemical analysis. Ion exchange is used on the laboratory scale for isolation and preconcentration of ions prior to instrumental analysis and to obtain preparative scale quantities of material for use in laboratory studies. For example, organic acids can be isolated from urine with a strong cation exchanger free of interference from the bulk of nonionic water-soluble organic compounds normally present in urine. Trace quantities of ions in ice from the polar regions can be concentrated by ion exchange for subsequent separation and identification by ion chromatography. Ion exchange is often employed in conjunction with activation analysis to isolate individual elements for quantification by radiochemical detection. Isolation schemes for individual elements in complex mixtures by ion exchange employing multiple changes in buffers and complexing agents are well known. These methods are slow and tedious but reliable, and they serve as reference methods to check the validity of fast routine methods of analyzing complex materials. *See* ACTIVATION ANALYSIS; ANALYTICAL CHEMISTRY.

Biotechnology. Biotechnology requires reliable, efficient methods to purify commercial-scale quantities of proteins, peptides, and nucleic acids for use in the pharmaceutical, agricultural, and food industries. Ion exchange is widely used in the isolation and purification of these materials. Typical applications include the removal of ionic compounds used in the production process, the elimination of endotoxins and viruses, the removal of host-cell proteins and deoxyribonucleic acid (DNA), and the removal of potentially hazardous variants of the main product. Since the biological activity of proteins is closely related to their three-dimensional structure and this structure is labile, one of the main reasons for choosing ion-exchange methods for the isolation and purification of proteins is that weakly denaturing conditions can be employed throughout the process steps. For commercial exploitation of biotechnology

products, the retention of biological activity is as important as recovery; and if the product is destined for human or animal consumption, biological impurities with adverse reactions must be reduced to a specified control range.

Chromatography. Modern chromatographic techniques employ ion exchangers of small particle size and favorable mass-transfer characteristics, and operate at high pressures, providing better resolution of mixtures in a shorter time than with conventional gravity-flow-controlled separations. In clinical laboratories, ion exchange has long been employed as the basis for the routine, automated separation of amino acids and other physiologically important amines used to identify metabolic disorders and to sequence the structure of biopolymers. Under typical conditions the amino acids are separated in the protonated form on a strong cation exchanger with stepwise changes in pH by using various combinations of citrate and borate buffers. The amino acids are usually detected as their fluorescent products after a postcolumn reaction with ninhydrin or *o*-phthalaldehyde. *See* AMINO ACIDS.

Probably the oldest and most common method of separating carbohydrate mixtures is by ligand-exchange chromatography using polymeric cation exchangers loaded with metal ions, for example calcium, for separating sugar and silver to determine the oligomer distribution of oligosaccharides, with aqueous organic solvents as the mobile phase. Retention in this case is governed by a combination of size exclusion and electrostatic attraction between the electronegative sugar oxygen atoms and the electropositive metal cations. Carbohydrates are also separated by anion-exchange chromatography after they are converted to negatively charged borate complexes or as their negative ion form at high pH in aqueous solution.

Ion chromatography is an ion-exchange-based separation system that changed the practice of ion analysis by replacing many tedious wet-chemical procedures with a simple automated instrument that can determine several ions simultaneously in a single method (**Fig. 2**). Its evolution into a major instrumental method of analysis was a synergistic development of efficient low-capacity ion-exchange column packings that allowed the use of low-concentration eluants containing competing ions with a high affinity for the ion exchanger, together with the provision of continuous on-line detection based on conductivity monitoring of the column effluent. Ion chromatography has been used extensively for the determination of ions in drinking water, wastewater, plating baths, detergents, foods, and many other applications. It is almost always the method of choice for determining anions and organic cations and has become popular for the separation of metal ions. *See* CHROMATOGRAPHY; LIQUID CHROMATOGRAPHY.

Ion-pair chromatography. Ion-pair chromatography provides an alternative method of separating ions without the use of an ion-exchange column. A temporary ion-exchange layer is formed by the adsorption of hydrophobic ions onto the surface of a reversed-phase column packing material. The hydrophobic ions are added to the mobile phase in a fixed concentration in equilibrium with the ions adsorbed onto the stationary phase, and are present throughout the separation. Suitable hydrophobic ions for separating anions are tetraalkylammonium salts, and for the separation of cations, alkanesulfonate salts. The most important variables in optimizing the separation of ions by ion-pair chromatography are the concentration of the ion-pair reagent, the concentration of organic solvent in the mobile phase, and the mobile-phase pH. An advantage of ion-pair chromatography is that ions and neutral molecules can be separated at the same time by using the efficient column packings developed for reversed-phase chromatography.

Colin F. Poole

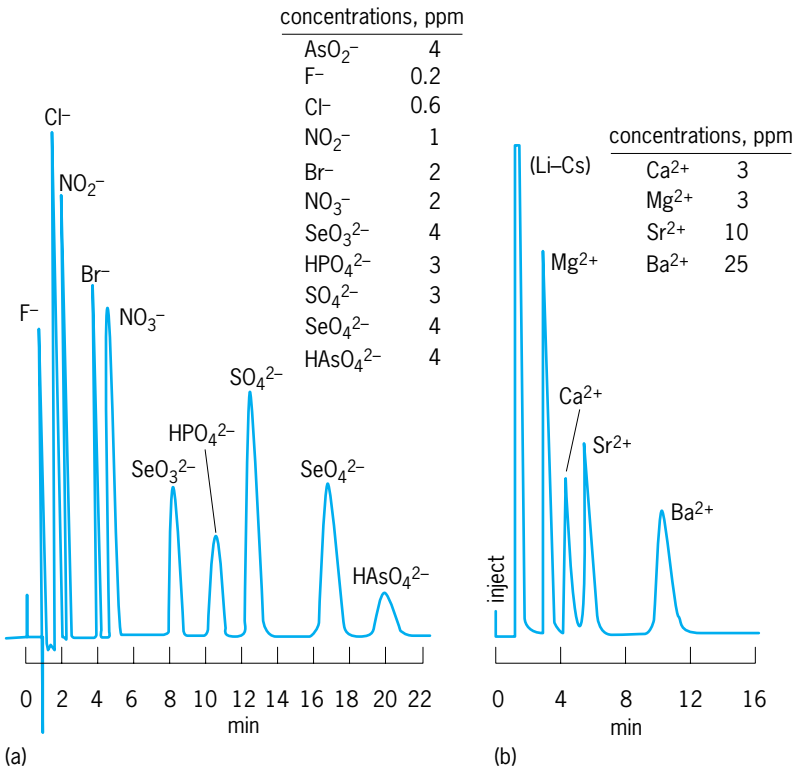

Fig. 2. Chromatograms showing separation of (*a*) common anions and (*b*) alkaline-earth cations by ion chromatography with conductivity detection. (*After C. F. Poole and S. K. Poole, Chromatography Today, Elsevier, 1991*)

Membranes

Ion-exchange membranes are a class of membranes that bear ionic groups and therefore have the ability to selectively permit the transport of ions through themselves. In biological systems, cell membranes and many other biological membranes contain ionic groups, and the conduction of ions is essential to their function. Synthetic ion-exchange membranes are an important field of technology. They are used in fuel cells, electrochemical processes for chlorine manufacture and desalination, membrane electrodes, and separation processes.

Ion-exchange membranes typically consist of a thin-film phase, usually polymeric, to which have

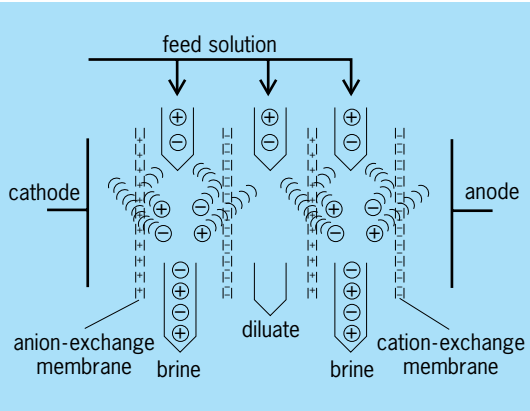

Fig. 3. Electrodialysis process, in which the ions of the saline water are separated under the influence of an electric field. (*After H. Strathmann, The principle of electrodialysis, in P. M. Bungay et al., eds., Synthetic Membranes: Science, Engineering and Applications, D. Reidel Publishing Co., 1986*)

been attached ionizable groups. Numerous polymers have been used, including polystyrene, polyethylene, polysulfone, and fluorinated polymers. Ionic groups attached to the polymer include sulfonate ($—SO_3^-$), carboxylate ($—COO^-$), tetralkylammonium ($—N(CH_3)_4^+$), phosphonate ($—PO_3H^-$), and many others. If the polymer is sufficiently hydrophobic that it will not dissolve in water, it may be formed into a film without further processing, but more hydrophilic polymers may require cross-linking to avoid excessive water solubility.

As a result of their ionic charge, ion-exchange membranes always contain counterions (gegen-ions) that have the opposing charge, and they maintain electrical neutrality. On contact with water, the membranes swell as the ions are hydrated, and they frequently have sufficient water in their structures to form a second continuous phase. This water is beneficial, since it allows ionic mobility and, by plasticizing the membrane polymers, some polymer mobility also. The water-swollen membrane has more space between adjacent polymer chains (greater free volume), and this allows the polymer chains to adjust their positions to allow ions to pass.

The Donnan effect is important to the operation of ion-exchange membranes. The Donnan equilibrium may be written for the simple case of univalent ions [Eq. (1)], where the o subscripts indicate

$$[M^+]_o[X^-]_o = [M^+]_m[X^-]_m = [MX] \qquad (1)$$

the solution outside the membrane; and the m subscripts indicate the interior of the membrane; the bracketed terms indicate the equilibrium concentrations. Since the membrane contains a large number of fixed charges, the concentration of counter-ions is also high, and the concentration of co-ions is therefore low. Thus an ion-exchange membrane containing negative fixed charges will readily permit transport of positive counterions, but it will act as a barrier to negative co-ions. The driving force for this transport may be a concentration gradient or an applied electric field. *See* DONNAN EQUILIBRIUM.

Membrane separators. Ion-exchange membranes are frequently used as membrane separators in electrochemical processes. Important among these are fuel cells, batteries, and chloralkali cells. In these applications, the membrane must be a selective conductor for ions, allowing the circuit to close. It thus separates the two half cells, while allowing the passage of ions necessary for the overall process. *See* FUEL CELL.

An example is the process for manufacturing chlorine. A sodium chloride brine is fed to the anode side of the cell, and an electric current is passed through it. Chloride ions are converted to chlorine gas. Sodium ions pass through the membrane to the cathode side, where water is electrolyzed to hydrogen gas, leaving sodium hydroxide. Since the membrane does not pass chloride ions, the resulting sodium hydroxide is quite pure. *See* ELECTROCHEMICAL PROCESS.

Electrodialysis. Electrodialysis is the process whereby saline or brackish water is rendered potable by separating the ions under the influence of an electric field. It is normally carried out by using a membrane stack (**Fig. 3**), in which cation-exchange membranes and anion-exchange membranes are alternated, and a cathode and anode apply the electric field on opposing sides. The saline solution is fed to all of the intermembrane gaps. Under the influence of the electric field, cations pass through the cation-exchange membranes, and anions through the anion-exchange membranes in opposite directions. The result is that alternate cells are depleted in these ions, while the remaining cells experience a net concentration. The depleted streams form the desalinated product. *See* DIALYSIS; WATER DESALINATION.

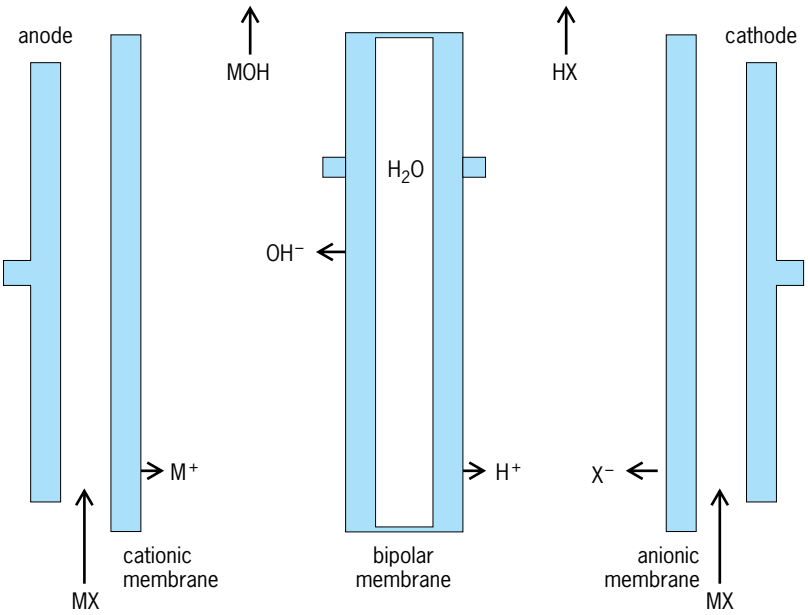

Fig. 4. Bipolar membrane salt-splitting process; water within the membrane splits into hydrogen (H^+) and hydroxide (OH^-) ions, which combine with the anions and cations to form acid and alkaline products.

Salt splitting with bipolar membranes. Salt splitting is related to electrodialysis in technology, but has the goal of converting a salt (MX) into an acid-base pair [HX + MOH; Eq. (2)] for the univalent case. The key

$$MX + H_2O \rightarrow HX + MOH \qquad (2)$$

added element is a bipolar membrane that has positive charges on one side and negative charges on the other. This membrane is placed between anion- and cation-exchange membranes (**Fig. 4**). In operation, water (H_2O) is split, forming hydrogen (H^+) and hydroxide (OH^-) ions within the bipolar membrane. The hydrogen ions combine with anions passing the anion-exchange membrane to form the acid product, while the products combine with cations passing the cation-exchange membrane to form the alkaline product.

Membrane electrodes. Membrane electrodes are widely used in analytical chemistry, the best-known example being the pH electrode. The key to their success is an ion-exchange membrane that is highly selective for the ion of interest. Membranes may be composed of glass or crystalline materials, or they may be made of polymeric membranes similar to those used in other membrane applications. Liquid membranes are also in common use in this application. The membrane is placed between the solution to be analyzed and a second solution containing the solution of interest. The difference in activity between these two solutions creates an electrical potential that can be measured. High flux is not necessary for this application, but rapid response is essential. *See* ELECTRODE; ION-SELECTIVE MEMBRANES AND ELECTRODES.

Liquid membranes. Liquid membranes differ in configuration from the previous applications. Instead of a solid or polymeric membrane, a liquid phase that is immiscible with the external solution is used. Liquid membranes may be supported, in which case the immiscible liquid wets a porous solid that holds it in place; or they may be unsupported, in which case the membrane consists of emulsion globules dispersed in an external phase. Regardless of configuration, differences in concentration between the two nonmembrane phases create a driving force for transport of the ions across the membrane. In the simplest case, these concentration differences are for the ion of interest. However, it is also possible to drive the transport process by using a concentration difference in another ion linked to the first by equilibrium processes. For example, in hydrometallurgical processes a high concentration of acid on the product side (H_p) of a liquid membrane system has been used to drive the concentration of a metal to this side of the membrane. N. N. Li; S. F. Yates

Bibliography. P. M. Bungay, H. K. Lonsdale, and M. N. de Pinho (eds.), *Synthetic Membranes: Science, Engineering and Applications*, 1983; P. R. Haddad and P. E. Jackson, *Ion Chromatography: Principles and Applications*, 1990; W. S. Hancock, *High Performance Liquid Chromatography in Biotechnology*, 1990; N. N. Li and J. M. Calo, *Separation and Purification Technology*, 1992; P. Meares, *Membrane Separation Processes*, 1976; C. F. Poole and S. K. Poole, *Chromatography Today*, 1991; R. Rautenbach and R. Albrecht, *Membrane Processes*, 1989; O. Samuelson, *Ion Exchange Separations in Analytical Chemistry*, 1963; H. Small, *Ion Chromatography*, 1989; H. F. Walton and R. D. Rocklin, *Ion Exchange in Analytical Chemistry*, 1990.

Ion implantation

A process that utilizes accelerated ions to penetrate a solid surface. The implanted ions can be used to modify the surface composition, structure, or property of the solid material. This surface modification depends on the ion species, energy, and flux. The penetration depth can be controlled by adjusting the ion energy and the type of ions used. The total number of ions incorporated into the solid is determined by the ion flux and the duration of implantation. This technique allows for the precise placement of ions in a solid at low temperatures. It is used for many applications such as modifying the electrical properties of semiconductors and improving the mechanical or chemical properties of alloys, metals, and dielectrics. *See* ALLOY; DIELECTRIC MATERIALS; METAL; SEMICONDUCTOR.

Wide ranges of ion energy and dose are applied. For ion energy ranging from 1 keV to 10 MeV, the ion penetration depth varies from 10 nanometers to 50 micrometers. In general, it is difficult to get deeper penetration since extremely high energy ions are required. As such, ion implantation is a surface modification technique and not suitable for changing the entire bulk property of a solid. Ion dosage also varies depending on the applications. Doses ranging from 10^{10} to 10^{18} ions/cm^2 are typically applied. For high-dose applications, ion sources providing high ion currents are needed to keep the implantation time reasonable for production purposes. *See* ION.

System. An ion implantation system consists of an ion source, an analyzer, an accelerator, a scanner, and a dose integrator at the target. In the ion source, a gaseous mixture is excited to form a plasma containing ions, electrons, and neutral particles. The ions enter the mass analyzer, where they are separated by an adjustable magnetic field for ion selection. The desired ions are chosen based on the charge-to-mass ratio of the ions, and the analyzer is sensitive enough to discriminate against adjacent mass numbers. The ions are then accelerated to the desired ion energy. The ion beam is scanned over the entire solid surface either electrostatically or mechanically. The total ion dose is determined by the charge integrator near the target. When ions are implanted into insulators, an electron source near the solid can be used to supply the surface with electrons to prevent charge buildup. *See* ELECTRON; PARTICLE ACCELERATOR.

Alternative ion sources include focused ion beam

systems and immersed plasma ion systems. In focused ion beam sources, ion lenses are used to form a convergent beam of ions with a small spot. Such sources have the flexibility of depositing ions in the desired areas without the use of a masking layer. Since the spot size can be focused down to 10 nm, selective ion implantation is possible with very high resolution. The drawbacks for focused ion beam sources are long processing times (since it is a "direct-write" serial process) and the limited types of available ions. In an immersed plasma ion source, the solid to be implanted is placed inside the plasma where ions are generated. The electric field between the source and the target accelerates the ions to the desired energy. This technique has the advantages of achieving higher ion density, covering a larger area, and allowing solids with different shapes to be implanted. *See* ELECTRIC FIELD; ION SOURCES; PLASMA (PHYSICS).

Projected ion range. As ions enter into a solid, they undergo a number of collisions with the atoms in the solid until they finally come to rest at some depth below the surface, which is the projected ion range. The projected range is proportional to the ion energy and inversely proportional to the ion mass. The projected ranges (see **illus.**) for different ion species implanted into silicon increase proportionally to the energy (10 keV–1 MeV). For boron implanted in silicon, which is often used as a *p*-type dopant for transistors, a high ion energy of 400 keV is required to obtain a 1-μm-deep projected range. Much higher ion energy is needed to achieve projected ranges beyond 1 μm, and it is difficult and expensive to accomplish. *See* BORON; SILICON.

The interactions between the implanted ions and the solid include elastic collisions between pairs of nuclei and inelastic collisions with electrons in the solid. High-energy ions initially lose their energy mostly by electron collisions and travel in a nearly straight path. As energy is reduced, nuclear collisions start to dominate at low energy and a zigzag path begins. Ions stop in the solid when ion energy is decreased to zero after a large number of collisions and changes of ion path. The projected range is the average depth of the implanted ions. The ion distribution can be approximated as a gaussian distribution with a standard deviation. The concentration of the implanted ions is lower near the surface, peaks at the projected range, and decreases beyond the projected range. The ion dispersion occurs in the vertical as well as the horizontal directions. Typically, ion dispersion decreases as the ion energy or the mass of the implanted ion increases. *See* ION-SOLID INTERACTIONS.

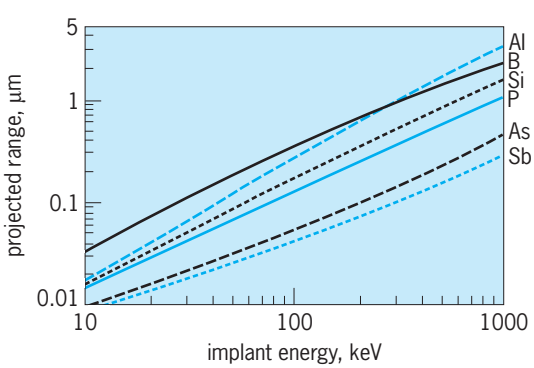

Projected range as a function of implant energy for ions (aluminum, boron, silicon, phosphorus, arsenic, and antimony) in silicon. (*After J. F. Gibbons, ed., Projected Range Statistics, 2d ed., 1975*)

In crystalline materials, such as single-crystal semiconductors, the atoms are arranged in regular patterns. Ions moving in certain directions can encounter rows or planes of atoms that form long-range open spaces through which the ions can travel without significant collision. Ions can travel down these channels to a much deeper distance, which is called ion channeling. For ions inside a channel, the major energy loss is due to electronic scattering since they have very few collisions with the nuclei. Since the stopping power for electronic scattering is low, the ion channeling path can be much deeper, showing up as a long tail to the implanted distribution. To avoid ion channeling and keep the ion implantation shallow, the angle between the incoming ions and the solid can be tilted, or the ions can be implanted through an amorphous layer. Ion implantation into a silicon wafer is typically carried out by tilting the solid 7° to minimize ion channeling. *See* CHANNELING IN SOLIDS.

Ion damage and annealing. As ions are implanted into a solid, they undergo a series of nuclear collisions. Every time an ion is scattered, part of the energy is transferred to the atoms in the solid, possibly displacing atoms in its path. In addition, the implanted ions generally do not rest in the proper sites, and implant-induced damage often occurs (due to the impact by high-energy ions). This damage effect is especially critical for crystalline solids. Thermal annealing is used both to repair the crystalline defects produced by the implanted ions and to allow the ions to move to their proper sites (so that the desired surface modifications can be realized). Annealing at temperature above 600°C (1112°F) for a few minutes is usually needed to remove the ion damage. In some cases, it may be helpful to anneal the damaged solid to cause surface melting, providing regrowth by liquid phase. During thermal annealing, diffusion (or distribution) of the implanted ions can occur. To avoid ion diffusion, rapid thermal annealing is used to remove damage at higher temperatures, but only by using a very short annealing time of a few seconds.

Applications. Ion implantation is used extensively in the semiconductor industry. The fabrication of integrated circuits in silicon often requires many steps of ion implantation with different ion species and energies. The implanted ions serve as dopants in semiconductors, changing their conductivity by more than a factor of 10^8. To introduce more electrons in silicon, *n*-type dopants such as arsenic, phosphorus, and antimony are often used, whereas boron is usually used as a *p*-type dopant to provide more holes in silicon. By implanting high doses of

oxygen ions in silicon, an insulating oxide layer can be formed underneath the surface, providing a silicon-on-insulator (SOI) structure. These SOI wafers can be used to produce integrated circuits with high speed and power. *See* ANTIMONY; INTEGRATED CIRCUITS; PHOSPHORUS.

Ion implantation is also used to change the surface properties of metals and alloys. It has been applied successfully to improve wear resistance, fatigue life, corrosion protection, and chemical resistance of different materials. Even though the ion projected range is less than 1 μm, surface treatment by ion implantation can extend the lives of metal or ceramic tools by 80 times or more. Ion implantation can form new compounds such as nitrides on the surface, and the implanted ions can be found at much greater depths than the projected range due to diffusion or mechanical mixing. *See* CERAMICS. Stella W. Pang

Bibliography. J. F. Gibbons (ed.), *Projected Range Statistics*, 2d ed., 1975; F. Komarov, *Ion Beam Modification of Metals*, 1992; J. W. Mayer (ed.), *Ion Implantation in Semiconductors, Silicon, and Germanium*, 1970; E. Rimini (ed.), *Ion Implantation: Basics to Device Fabrication*, 1994; J. F. Ziegler (ed.), *Handbook of Ion Implantation Technology*, 1992; J. F. Ziegler (ed.), *Ion Implantation: Science and Technology*, 1988.

Ion propulsion

Vehicular propulsion caused by the high-speed discharge of a beam of electrically charged minute particles. These particles, usually positive ions, are generated and accelerated in an electrostatic field produced within an ion thruster attached to a spacecraft. Because positive ions cannot be ejected from the thruster without leaving a substantial negative charge on the thruster and spacecraft, electrons must be ejected at the same rate. Ion propulsion systems are attractive because they expel the ions at very high speeds and, therefore, require much less propellant than other thrusters, such as chemical rockets.

The three principal components of an ion propulsion system (**Fig. 1**) are the power-generation and -conditioning subsystem, the propellant storage and feed subsystem, and one or more ion thrusters.

Power generation. This portion of the system can be broken down into power source and power-conversion and -conditioning equipment.

Power source. The power source can be a nuclear reactor or a radiant-energy collector. In the former, thermal power is released by fission or fusion reactions. Mass-minimization considerations generally define spacecraft subsystem configurations. The dominant masses in the case of the nuclear power source are generally those for the main radiator required to reject waste heat from the thermodynamic cycle, and shielding required to protect equipment and personnel against ionizing radiation from the reactor.

Solar radiation as the power source does not require shielding. It can be used to provide electric power directly through photovoltaic (solar) cells or indirectly through a solar collector–heat exchanger system similar to that for a nuclear system (Fig. 1). A single thermodynamic-cycle working fluid suffices, but power levels are generally considered limited to 100 kW by collector-size limitations. Radiant energy beamed to a spacecraft at nonoptical frequencies (for example, as microwaves) and converted to electrical power in appropriate photovoltaic cells may also be attractive in some applications.

Solar cells are a simple and useful power source for propelling small probes into geolunar space, the inner solar system, and in some instances as far as Jupiter. Their principal disadvantage is low voltage. Power-conditioning equipment, to achieve the high voltage required for ion propulsion, adds to the system mass, although considerable advances continue to be made in reducing the complexity and specific mass (mass per unit power) of solar cells and

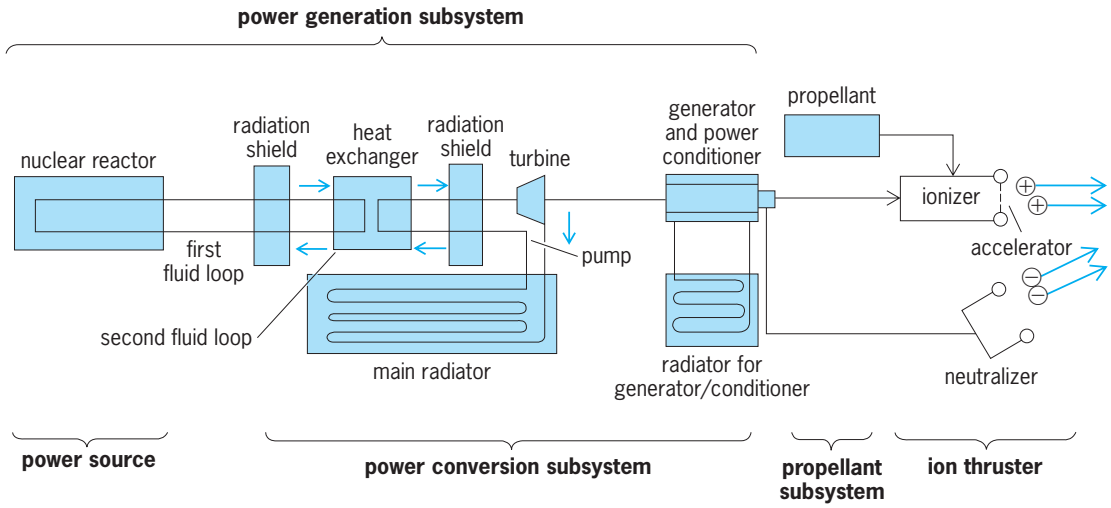

Fig. 1. Nuclear-powered electrostatic (ion) propulsion system.

power-conditioning equipment. Solar cells are damaged by exposure to particle radiation environments, which limits their useful lifetimes to values that may be tolerable in the Earth's radiation belt but are probably too low in the harsher environments at Jupiter or near the Sun.

As solar-powered generators approach the Sun, the energy flux density incident upon their collectors increases, their specific mass decreases, and they become increasingly attractive. Unfortunately, the range of distances from the Sun over which they are truly attractive is not large, because of particle radiation close to the Sun. Additional disadvantages of solar energy systems include the requirement to maintain a collector–concentrator alignment relative to the Sun, the loss of power in the shadow of a celestial body, and the difficulty of collector maintenance in an environment containing micrometeoroids. *See* SOLAR CELL; SOLAR ENERGY.

Power conversion. If the power-generation system involves a nuclear reactor or a solar-thermal subsystem, thermal-to-electric conversion subsystems are required. Those most highly developed involve thermodynamic conversion cycles based on turbine generators. Although most traditional systems have operated on the Brayton gas cycle or the Rankine vapor cycle, more recent efforts include the Stirling gas-cycle system.

In the Rankine cycle, a liquid is heated and evaporated either in a heat exchanger (two-loop system; Fig. 1) or directly in a nuclear reactor or solar-heated boiler (single-loop system). The vapor is then expanded through the turbine, condensed in the main radiator, and pumped back into the heat exchanger, reactor, or boiler. Electrical power is drawn from the generator which is driven by the turbine. Where there is concern about radioactivation of the working fluid (low-melting-temperature metals typically), the two-loop system is used and the heat exchanger serves to limit the spread of radioactivity. Heat transfer is complicated by the weightless conditions of space. Other thermodynamic cycles such as the Brayton or Stirling cycles, which do not require vaporization and condensation, mitigate this problem, but they typically have lower efficiencies. *See* BRAYTON CYCLE; GAS TURBINE; GENERATOR; RANKINE CYCLE; STIRLING ENGINE.

The main radiator, which rejects the waste heat from the conversion subsystem to space, is the largest and heaviest of the power-generation-subsystem components. Its size and mass are directly proportional to the thermal power rejection rate. High efficiency is critical to low radiator mass, but operating temperature, physical configuration, and micrometeor protection are also important.

Alternative methods of power conversion employ thermoelectric (semiconducting thermocouples) and thermionic devices to effect direct conversion of heat to electricity. These devices have a history of lower efficiencies than fluid cycle systems operating between the same temperatures, but they are typically more compact and, because they typically have no moving parts, they are potentially more reliable. *See* THERMIONIC POWER GENERATOR; THERMOELECTRICITY.

Propellants. Ion-thruster propellants are selected according to performance, technical suitability, availability, cost, and storage characteristics. For geospace transportation, the effect of ions that may spiral along the Earth's magnetic field lines to the magnetic poles must also be considered.

High thruster efficiency requires a high-molecular-mass propellant which facilitates operation at high utilization efficiency where the loss of unaccelerated propellant is small. A high molecular mass also facilitates operation at a high acceleration voltage, which assures a high thruster electrical efficiency. Low first-ionization potential and high first-excitation potential for the propellant are important. Such potentials assure minimal radiation losses and low ion-production power.

High thrust-to-power ratio is also promoted by a high propellant molecular mass. Operation at this condition is desirable because spacecraft are usually power-limited.

Thruster lifetime can be extended by a high second ionization potential and a low charge-exchange cross section to limit the rate at which doubly charged and charge-exchange ions erode thruster surfaces.

Compatibility of thruster with spacecraft materials and functions is essential. Propellants that are, for example, likely to interfere with the operation of the thruster or spacecraft systems by forming coatings are undesirable.

Operational controllability is important. Monatomic propellants are preferred because they ionize in a predictable way and the thrust created by their expulsion is, therefore, more readily controlled.

Propellants that have been investigated include argon, xenon, cesium, mercury, and fullerenes such as C_{60} (see **table**). Although mercury received most of the early attention, xenon is now being used on all space missions because of toxicity concerns with mercury. Cesium is not used because it tends to migrate and coat surfaces on components such as insulators. C_{60}, which is attractive because of its high molecular mass, is probably unacceptable because it tends to fragment and to form negative ions, which degrade thruster performance.

Thrust device. Ion or electrostatic thrust devices contain three functional elements: an ionizer that generates the ions; an accelerator providing an electric field for accelerating the ions and forming them into a well-focused beam; and a neutralizer or electron emitter that neutralizes the electrical charge of the exhaust beam of ions after they have been ejected. Electrostatic thrusters are classified according to the type of ions they eject and by the method utilized to produce them.

Ionizer. The positive ions needed for acceleration are produced in a strong electric field, by contact with a surface having a work function greater than the ionization potential of the propellant, or by electron-bombardment ionization. The last method

Ion propellant characteristics

Propellant	Atomic weight	Availability	Storability	Spacecraft thruster material compatibility	Ionization potentials, eV 1st	Ionization potentials, eV 2d	Thruster technology status
Argon	39.95	Good	Cryogenic liquid	Excellent	15.76	27.5	Ground tests
Xenon	131.3	Scarce	High-pressure gas	Excellent	12.13	21.2	Space flight
Cesium	132.9	Limited	Solid	Poor	3.90	23.4	Space flight
Mercury	200.6	Good	Liquid	Good	10.44	18.8	Space flight
Fullerene	720	Good	Solid	Marginal	7.6	?	Preliminary tests

has received the most attention and appears to be the most promising.

In an electron-bombardment ionizer, high-energy electrons collide with relatively massive neutral atoms or heavy molecules like fullerenes being supplied in the discharge chamber and induce ionization (**Fig. 2**). In the direct-current ionizer, the electrons are accelerated from a cathode to kinetic energies of a few tens of electronvolts via a potential difference sustained by an anode. These electrons are usually obtained from a hollow cathode which has the long lifetime, high reliability, and low power consumption needed for space operation. Various magnetic-field configurations, which are usually induced by permanent magnets, are employed to confine the electrons away from the anode until after they have expended most of their kinetic energy in collisions with neutral atoms or molecules. The magnetic field also serves to confine and direct the ions produced toward the accelerating electrodes. This class of thruster has been operated in space with cesium, mercury, and xenon propellants.

In another type of electron-bombardment ionizer, electrons that result from the ionization process itself are accelerated in high-frequency electric fields induced within the discharge chamber by external antennae. These electrons acquire sufficient kinetic energy to ionize propellant atoms; they then migrate toward an electrode where they are collected. In contrast to dc ionizers, these use high-frequency rather than direct-current power to accelerate electrons and require no cathode. *See* ION SOURCES.

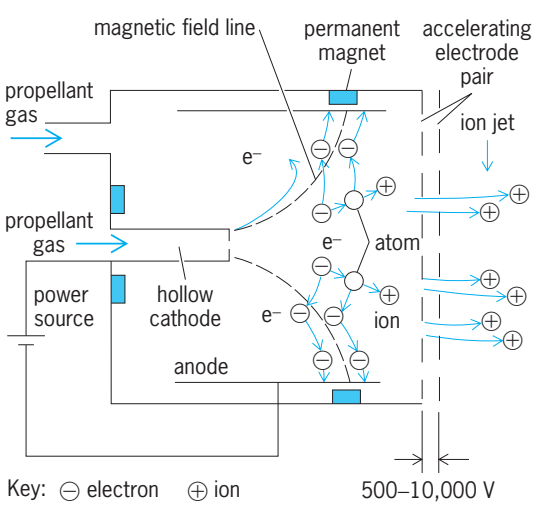

Fig. 2. Direct-current electron-bombardment ion thruster.

Accelerator. Some of the ions produced are directed toward the ion-accelerating subsystem which typically consists of two plates containing large numbers of aligned hole pairs (Fig. 2). The upstream plate and the body of the ionizer are maintained at a potential V (in volts) positive with respect to the space downstream from the thruster, whereas the downstream plate is biased negative at a smaller value. The positive ions are therefore first accelerated between the plates and then decelerated after they pass through the downstream plate until they have been accelerated through the net (or beam) voltage V. In order to ensure that ions do not strike the downstream plate directly, hole size and alignment and plate spacing must be set and maintained properly. For a high extracted ion current density, the plates should be as close together as possible.

Charge neutralization. Electrons are released near the ion beam downstream from the accelerator at the same rate as the ions. These electrons neutralize the space charge of the positive ions and prevent gross spacecraft charging. They are prevented from migrating upstream into the positively biased ionizer by the negatively biased downstream accelerator plate. A hollow cathode, similar to the one used as an electron source in the dc ionizer but optimized for low propellant flow rate, is used as a neutralizer because of its high efficiency and long life.

Stationary plasma thruster. This thruster combines the functions of ionization, acceleration, and neutralization in a simple open chamber without a planar electrode pair. It utilizes a magnetic field to confine electrons while ions are accelerated electrostatically, and current designs have a limited specific impulse capability (1000–2000 s). *See* SPECIFIC IMPULSE.

Applications. Ion propulsion is characterized by high specific impulse and low thrust. Thrusters have generally been designed for operation in the 2500–10,000-s specific impulse range, but the upper limit could be extended. Because high specific impulse means low propellant consumption, ion propulsion is attractive for a wide variety of applications.

Satellites. One functional category includes the use of ion thrusters on satellites for orbit control (against weak perturbation forces) and for station keeping (position maintaining of satellite in a given orbit). Substantial commercial use of ion thrusters in this application began at the end of the twentieth century.

An ion propulsion system can also be used advantageously for changing the satellite's position in a given orbit, especially shifting a satellite to different longitudes over the Earth in an equatorial geostationary orbit. In current systems, the satellite's reaction control system is fired for 2–3 min, causing the satellite to drift slowly to a different place, where it is stopped by a second reaction control system maneuver. The slow drift can be time-consuming, several times the few days required by an ion thruster for a 180° shift. *See* SATELLITE (SPACECRAFT); SPACECRAFT PROPULSION.

Orbit raising and related missions. Ion propulsion in geospace can be advantageous, but it is complicated by three factors: strong gravity forces, which typically result in near-Earth to geosynchronous transfer times of the order of 100 days depending on available power-vehicle mass ratio; the Van Allen radiation belt; and eclipsing. The last two factors are problematical when the Sun's power is used, especially when silicon photovoltaic cells, which suffer radiation damage in the Van Allen belt, are involved. The exclusive use of nuclear power or the use of a reusable high-thrust interorbital stage to boost the solar-electric stage from near-Earth to an orbit above the most damaging region of the Van Allen belt (18,000 km or 11,000 mi) can eliminate the last two concerns. Orbit plane changes accomplished along with orbit raising are also attractive, although they increase transfer times. Geospace missions might involve delivery of cargo, single- and multiple-satellite servicing, and space debris retrieval. This last mission is fast becoming important as the probability of collisions with debris, especially at the geosynchronous altitude, increases. *See* VAN ALLEN RADIATION.

Geolunar transfer. Earth-Moon transfer can be accomplished with ion propulsion starting at any place in the near-Earth to geosynchronous altitude range, but the complicating factors cited above make the higher-altitude starting point more attractive.

Interplanetary transfer. A major functional application of ion propulsion is interplanetary transfer. Here, thrust has to overcome only very weak solar gravitational forces. Because of this, and the long powered flight times of which ion propulsion is capable, transfer times to Venus or Mars need not be longer than transfer times in comparable flights with high thrust drives capable only of short powered flight. At the very large distances to objects in the outer solar system, ion propulsion would yield shorter transfer times than chemical and most high-thrust nuclear concepts. The National Aeronautics and Space Administration (NASA) *Deep Space 1* mission, started at the end of the twentieth century, uses a 30-cm-diameter (12-in.) xenon ion thruster to propel a spacecraft to encounters with the asteroid Braille (in 1999) and the comet Borrelly (2001). *See* ELECTROTHERMAL PROPULSION; INTERPLANETARY PROPULSION; PLASMA PROPULSION; SPACE PROBE.

Paul J. Wilbur

Bibliography. G. R. Brewer, *Ion Propulsion*, 1970; M. S. El-Genk et al., System design optimization for multimegawatt space nuclear power applications, *J. Propuls., Power*, 6:194–202, 1990; H. R. Kaufman, Technology of electron-bombardment ion thrusters, *Adv. Electr. Electr. Phys.*, 36:273–275, 1974; E. J. Lerner, Plasma thrusters from Russia, *Aerosp. Amer.*, 30(9):51, September 1992; H. Mark and H. J. Smith, Fast track to Mars, *Aerosp. Amer.*, 29:36–41, 1991; A. K. Sinha, B. Agrawal, and W. W. Wu, Trends in satellite communications technology, techniques and applications, *Int. J. Satel. Commun.*, 8:283–294, 1990; P. J. Wilbur, J. R. Beattie, and J. Hyman, Approach to the parametric design of ion thrusters, *J. Propuls., Power*, 5:575–583, 1990; P. J. Wilbur, R. G. Jahn, and F. C. Curran, Space electric propulsion plasmas, *IEEE Trans. Plasma Sci.*, 19:1167–1179, 1991; P. J. Wilbur, V. K. Rawlin, and J. R. Beattie, Ion thruster development trends and status in the United States, *J. Propuls., Power*, 14:708–715, 1998.

Ion-selective membranes and electrodes

Membrane-based devices, involving permselective, ion-conducting materials, used for the measurement of activities of species in liquids or partial pressures in the gas phase. Permselective means that ions of one sign may enter and pass through a membrane.

Properties. Ion-selective electrodes are generally used in the potentiometric mode, and they superficially resemble the classical redox electrodes of types 0 (inert), 1 (silver/silver ion; Ag/Ag^+), 2 (silver/silver chloride, chloride ion; Ag/AgCl/Cl-, and 3 (lead/lead oxalate/calcium oxalate/calcium ion; $Pb/PbC_2O_4CaC_2O_4/Ca^{2+}$). The last, while ion-selective, depends on a redox couple (electron exchange) rather than ion exchange as the principal origin of interfacial potential difference. Ion-selective electrodes have a typical form that can be expressed in shorthand form, the "ionic-contact" membrane configuration: [lead wire (Cu; Ag); inner reference electrode (AgCl); inner filling solution (electrolyte: M^+Cl^-); membrane permselective to M^+] and the "all-solid-state" membrane configuration [Cu; Ag membrane permselective to M^+].

In the former, both membrane interfaces are ion-exchange-active, and the potential response depends on M^+ activities in both the test solution and the inner filling solution. In the latter, the membrane must possess sufficient electronic conductivity to provide a reversible, stable electron-exchange potential difference at the inner interface, with ion exchange only at the test solution side.

Potentiometric responses of ion-selective electrodes take the form in Eq. (1) when an ion-selective

$$V(\text{measured}) = V^0 + \frac{RT}{F}\ln\left[a_i^{1/z_i} + \sum_j (k_{ij}a_j)^{1/z_j}\right] \qquad (1)$$

$R =$ the universal gas constant
$T =$ the absolute temperature
$F =$ the Faraday constant (96,487 coulombs/equivalent)
$V_0 =$ formal reference potential

electrode is used with an external reference electrode, typically a saturated-calomel reference electrode with a salt bridge, to form a complete electrochemical cell. Activities of the principal ion, a_i, and interfering ions, a_j, are in the external "test" solution and correspond to ions M^{z_i} and M^{z_j}, where z_i and z_j are charges with sign. The ion M^{z_j} is written first because it is the principal ion favored in the membrane, for example, high ion-exchange constant and high mobility. Values known as selectivity coefficients (k_{ij}) are experimentally determined, but they can be related to extrathermodynamic quantities such as single-ion partition coefficients and single-ion activity coefficients and mobilities. When only one ion is present in a test solution, or the membrane is ideally permselective for only one ion, this equation simplifies as Eq. (2). V^0 can be

$$V(\text{measured}) = V^0 + \frac{RT}{z_i F}(\ln a_i) \qquad (2)$$

written explicitly in terms of activities of species at the inner interface or in terms of solid-state activities for the all-solid-state configuration. Equation (1), variously known as the Horovitz, Nicolsky, or Eisenman equation, resembles the Nernst equation (2), but the former originates from different factors. Equation (1) cannot be derived from first principles for ions of general charge. However, when $z_i = z_j$, the equation can be derived by various means including, mainly, thermodynamic (Scatchard) equations and transport models (Nernst-Planck equations).

The response slope $2.303RT/z_iF$ is $59.14/z_i$ mV per decade of activity at 77°F (25°C). Measurements reproducible to ± 0.6 mV are typically achieved, and activities can be reproducible to ± 2% for monovalent ions. A normal slope is considered nernstian, and can persist over a wide activity range, especially for solid electrodes, for example, 24 decades in the Ag^+,S^{2-} system, and 12 decades of H^+ using Li^+-based glass membrane electrodes. Less than nernstian slopes, and in the limit zero slope, can occur at low activities of sensed species, and can occasionally occur at very high activities. Ultimate low-level response (detection limit) is determined by the solubility of the membrane ion-exchanging material, although impurities may cause premature failure. Because of the logarithmic dependence of potential response on activities, activity measurements using ion-selective electrodes are particularly suited to samples with wide activity variations. Standardization against pure samples or samples of the type to be determined is required. Precise measurements of concentrations over narrow ranges are not favorable, but are possible by elaborate standardization schemes: bracketing, standard additions, and related methods involving sample pretreatment. *See* DONNAN EQUILIBRIUM.

Ion-selective electrodes are most often cylindrical, 6 in. (15 cm) long and 0.25 or 0.5 in. (6 or 13 mm) in diameter, with the lead wire existing at the top and the membrane sensor at the lower end. However, inverted electrodes for accommodating single-drop samples, and solid electrodes with a drilled hole or a special cap for channeling flowing samples past a supported liquid membrane, are possible configurations. The conventional format is intended for dip measurements with samples large enough to provide space for an external reference electrode. Single combination electrodes are useful for smaller samples, because the external electrode is built-in nearly concentrically about the ion-selective electrode. Drilled and channeled-cap electrodes are intended for use with flowing samples. Microelectrodes with membrane-tip diameters of a few tenths of a micrometer have been constructed for single-cell and other measurements in the living body (see **illus.** for construction details).

Electrodes can also be flat-form stacks of active layers on an inert support, that is, Ag/AgCl/KCl (aqueous gel)/membrane. Membranes can also be placed on an inert metal gate or insulator of a field-effect device. These flat devices are susceptible to miniaturization by methods of silicon chip technology. So-called smart sensors have the signal-processing electronics on the chip as well.

Ion-selective electrodes are intended to be used to monitor and measure activities of flowing, or stirred, solutions because electrodes detect and respond to activities only at their surfaces. The time responses of solid and liquid membrane electrodes to ideal step activity changes of the principal ion (already present in the membrane) can be very rapid: 200 milliseconds for glass to about 30 microseconds for silver bromide (AgBr). Generally this fast response cannot be observed or used, because sample mixing or diffusion of a fresh sample to an electrode surface is the limiting process. Also, almost any time two ions are simultaneously determining response, interior diffusion potential generation is involved in reaching a new steady-state potential. Similarly, formation of hydrated surface layers or layers of adsorbed matter or inhomogeneities introduces diffusion barriers. Response times from 2 to 20 s can be expected. About 20–30 samples per hour may be analyzed manually by the dip method, and about 60 per hour when samples are injected into a flowing stream of electrolyte. *See* ELECTRODE POTENTIAL; ELECTROLYTIC CONDUCTANCE.

Classification and responsive ions. Ion-selective electrodes are classified mainly according to the physical state of the ion-responsive membrane material, and not with respect to the ions sensed. It has also proved superfluous to distinguish between homogeneous membranes and those that are made from a homogeneous phase supported physically in voids of an inert polymer, or from two homogeneous phases intimately mixed, so-called heterogeneous membranes.

Glass membrane electrodes. These are used for hydrogen-ion activity measurements. Glass electrodes are based on alkali ion silicate compositions. Superior pH-sensing glasses (pH 1 to 13 or 14) result from lithium silicates with addition of di-, tri-, and tetravalent heavy-metal oxides. The latter are not chain formers. Membranes responsive to

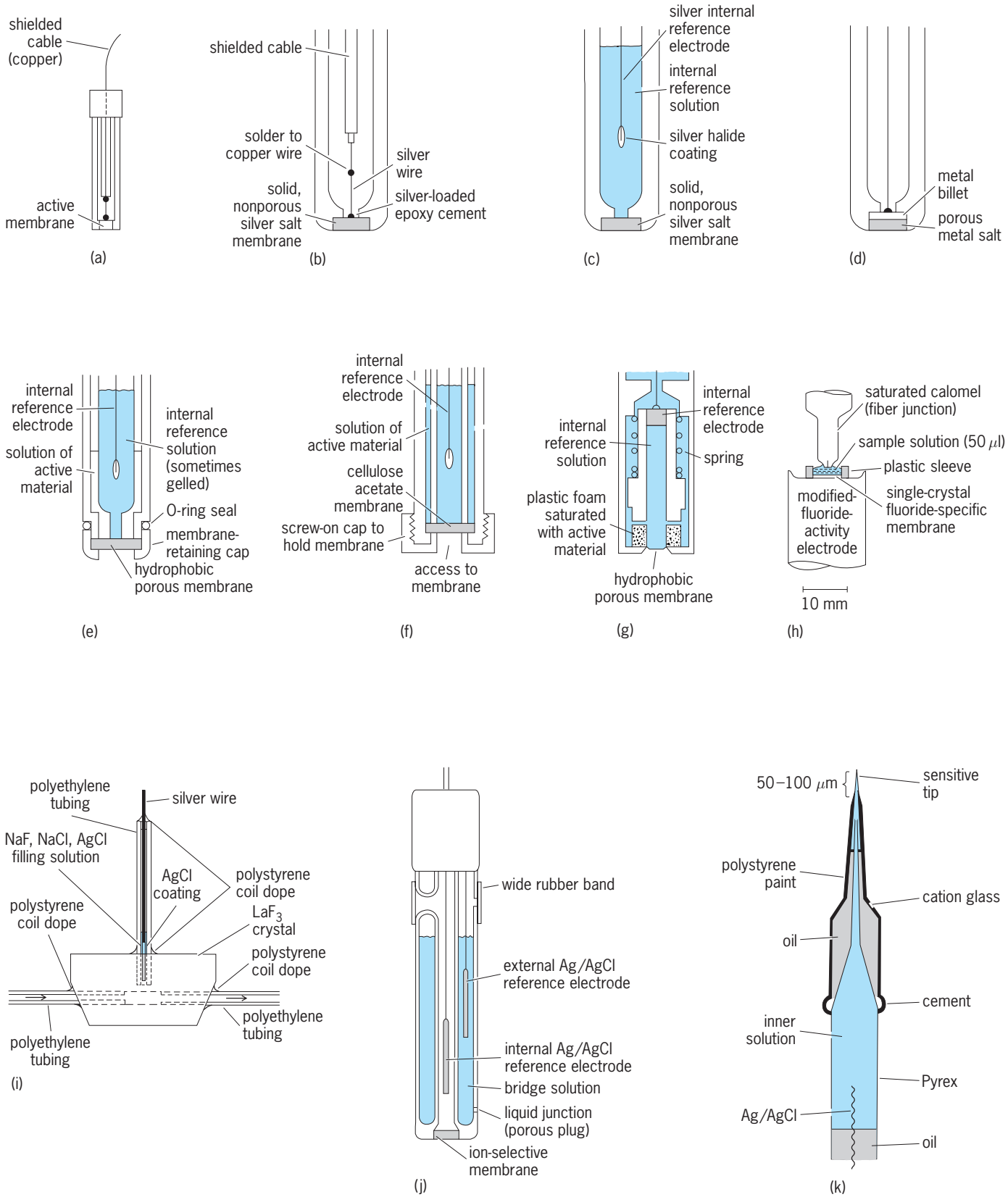

Diagrams of ion-selective electrodes. (*a*) Typical electrode configuration—in this example, an all-solid-state ion-selective electrode. (*b*) Enlarged view of the construction for metal-contacted-membrane ion-selective electrodes. (*c*) Enlarged view of the construction for internal electrolyte-contacted-membrane ion-selective electrodes. (*d*) Enlarged view of the construction for an electrode. (*e–g*) Enlarged views of constructions for liquid ion-exchanger-membrane ion-selective electrodes. (*h*) An inverted electrode microcell using a fluoride-sensing material; the reference electrode (external) is saturated calomel [SCE] (*after R. A. Durst and J. K. Taylor, Anal. Chem., 39:1483, 1967*). (*i*) Construction of a flow-through crystal electrode (*after H. I. Thompson and G. A. Rechnitz, Anal. Chem., 44:300, 1972*). (*j*) A combination electrode illustrating the usual active membrane surrounded by an attached Ag/AgCl external reference electrode. (*k*) An example of a cation-sensing microelectrode used in biological research (*after R. N. Khuri, W. J. Flanagan, and D. E. Oken, J. Appl. Physiol., 21:1568, 1966*).

the sodium (Na^+), potassium (K^+), and ammonium (NH_4^+) cations, and some other cations use additional aluminum oxide (Al_2O_3), boric oxide (B_2O_3), or both. The pH glasses are highly selective for hydrogen ion (H^+) over other monovalent ions. The Na^+-sensing glasses are not intrinsically very selective for Na^+ over H^+, but useful pNa measurements can be made, even in excess K^+ at pH 7 or above. No glasses with high selectivity of K^+ over Na^+ have been found. Chalcogenide glasses containing low contents of copper(III) [Cu^{2+}] or iron(III) [Fe^{3+}], while called glasses, are thought to be semiconductor electrodes with a high component of electron exchange, rather than ion exchange, for establishment of interfacial potential responses to Cu^{2+} and Fe^{3+}.

Electrodes based in water-insoluble inorganic salts. These electrodes include sensors for the ions fluoride (F^-), chloride (Cl^-), bromide (Br^-), iodide (I^-), cyanide (CN^-), thiocyanate (SCN^-), sulfide (S^{2-}), Ag^+, Cu^{2+}, cadmium (Cd^{2+}), and Pb^{2+}. The compounds used are silver salts, mercury salts, sulfides of Cu, Pb, and Cd, and rare-earth salts. All of these are so-called white metals whose aqueous cations (except lanthanum; La^{3+}) are labile. The salts are Frenkel-defect solids which possess the necessary ionic conductivity. Ag_i^+ (interstitials) or Ag_i^- (vacancies) are the mobile species in the silver salts, while F^- interstitials are mobile in LaF_3. These materials are ion exchangers, and show no diffusion potential. Single crystals, doped and undoped, may be used as membranes. Pressed pellets using inert binders such as polyethylene or an insoluble salt such as Ag_2S (for the silver halide electrodes) are popular.

In addition, powdered salts may be suspended in silicone rubber or polyvinyl chloride (about 50:50% by weight) to form heterogeneous flexible membranes. CuS-Ag_2S, CdS-Ag_2S, and PbS-Ag_2S pressed pellets formed at about 480°F (250°C) are indirectly responsive to the divalent metal ion activities through control of Ag^+ activities at the electrode surface and in leached layers or surface pores by means of the common ion effect.

Electrodes using liquid-ion exchangers. These are electrodes supported in the voids of inert polymers such as cellulose acetate, or in transparent films of polyvinyl chloride, and provide extensive examples of devices for sensing. Fewer cation-sensing liquid-ion exchanger systems have been found. The principal example (and among the most important) is the Ca^{2+}-responsive electrode based on calcium salts of diesters of oil-soluble phosphonic acids. Oil-soluble metal salts of hydrophobic acids containing chelating nitrogen (N) and sulfur (S) groups can be used in electrodes for heavy-metal ion sensing. The condition for development of an ion-selective potential signal is rapid exchange of the sensed metal ion between membrane (dissolved chelate or salt) and the same ion in the test solution. If this exchange is slow, as in magnesium ion (Mg^{2+}) exchange between porphyrins and solutions, the sensor is anion-selective. Anion-sensing electrodes typically use an oil-soluble cation Aliquat (methyltricaprylammonium) or a metal ion-uncharged organic chelating agent (Ni^{2+} or Fe^{2+} phenanthroline or substituted phenanthroline cations) in a support matrix. Sensitivity is virtually assured if the salt is soluble in a mediator solvent, typically a nitro aromatic or esters of difunctional carboxylic acids: adipic, sebacic, or phthalic. Selectivity poses a severe problem since these electrodes, based on hydrophobic materials, tend to respond favorably to many oil-soluble anions. Thus construction of electrodes for the simple inorganic anions F^-, hydroxide (OH^-), bicarbonate (HCO_3^-), and hydrogen phosphate (HPO_4^{2-}) is difficult. Yet many electrodes respond to SCN^-, I^-, Br^-, nitrate (NO_3^-), perchlorate (ClO_4^-), and fluoroborate (BF_4^-) in accordance with the Hofmeister lyotropic series. Surfactant anion sensors use salts such as hexadecylpyridinium dodecylsulfate in *o*-dichlorobenzene; surfactant cation sensors use a picrate salt of the species to be measured. Acetylcholine may be measured in the presence of choline, Na^+, and K^+ using the tetra-*p*-chlorophenylborate salt in a phthalate ester in polyvinyl chloride, for example. Cation and anion drug sensors can be made by dissolving a cation pair (with oil-soluble anion such as tetraphenylborate) in a plasticized polyvinyl chloride membrane. Anion drug sensors often use methyltridodecylammonium ion (Aliquat) to form the oil-soluble ion pair.

Neutral carrier-based sensors for monovalent and divalent cations are closely related to ion-exchanger-based electrodes. Both systems involve ion-exchange sites, particularly negative mobile sites arising from mediators and negative fixed sites arising from support materials. Some of the available neutral carriers are hydrophobic complex formers with cations, and they can be either cyclic or open-chain species that form ion-dipole complexes. These compounds permit selective extraction (leading to permselectivity) for ions such as K^+, Na^+, NH_4^+, and Ca^{2+} that would ordinarily not dissolve as simple inorganic salts in the hydrocarbonlike membrane phase. Valinomycin is the best-known example, and its use in supported solvents such as dioctylsebacate provides an electrode with sensitivity of 10^5 for K^+/Na^+. Neutral carriers that form adducts using weak covalent bonds are particularly useful for constructing selective anion sensors without Hofmeister series limitations. Trifluoroacetyl-*p*-butylbenzene$_2$, selective for carbonate (CO_3^{2-}); bis(*p*-chlorobenzyl)tin dichloride, selective for HPO_4^{-2}; and some special porphyrins are examples. *See* ION EXCHANGE.

Electrodes with interposed chemical reactions. These electrodes, with chemical reactions between the sample and the sensor surface, permit an enhanced degree of freedom in design of sensors for species which do not directly respond at an electrode surface. Two primary examples are the categories of gas sensors and of electrodes which use enzyme-catalyzed reactions. Gas sensors for carbon dioxide (CO_2), sulfur dioxide (SO_2), ammonia (NH_3), hydrogen sulfide (H_2S), and hydrogen chloride (HCl), and others can be made from electrodes responsive to H^+, S^{2-}, or Cl^-. By enclosing a pH glass membrane in a thin

layer of dilute sodium bicarbonate ($NaHCO_3$), an electrode for partial pressure of CO_2 is formed, since H^+ increases in a known way with increasing dissolved CO_2. Similarly, immobilized enzymes convert a substrate such as urea or an amino acid to ammonia, which can be sensed and monitored by the underlying electrode. However, increased sensitivity is accompanied by an increased response time. Each diffusion and diffusion-reaction barrier slows the transport and increases the time constant of the overall sensor electrode. *See* ELECTRODE; ENZYME.

Applications. Electrodes for species identified above are, for the most part, commercially available. In addition, electrodes have been made and reported that are responsive to many other species. A few of these are cesium (Cs^+), thallium (Tl^+), strontium (Sr^{2+}), Mg^{2+}, zinc (Zn^{2+}), nickel (Ni^{2+}), (UO_2^{2+}), mercury(II) [Hg], hydrogen sulfite (HSO_3^-), sulfate (SO_4^{2-}), periodate (IO_4^-), perrhenate (ReO_4^-), halide anion complexes of heavy metals (for example, $FeCl_4^-$), pyridinium, pyrocatechol violet, vitamins B_1 and B_6, and many cationic drugs, aromatic sulfonates, salicylate, trifluoroacetate, and many other organic anions. Applications may be batch or continuous. Important batch examples are potentiometric titrations with ion-selective electrode end-point detection, determination of stability constants of complexes and speciation identity, solubility and activity coefficient determinations, and monitoring of reaction kinetics, especially for oscillating reactions. Ion-selective electrodes serve as liquid chromatography detectors and as quality-control monitors in drug manufacture. Applications occur in air and water quality (soil, clay, ore, natural-water, water-treatment, seawater, and pesticide analyses); medical and clinical laboratories (serum, urine, sweat, gastric-juices, extracellular-fluid, dental-enamel, and milk analyses); and industrial laboratories (heavy-chemical, metallurgical, glass, beverage, and household-product analyses). *See* ANALYTICAL CHEMISTRY; CHROMATOGRAPHY; QUALITY CONTROL; TITRATION. Richard P. Buck

Bibliography. P. L. Bailey, *Analysis with Ion-Selective Electrodes*,1988; J. Koryta and K. Stulik, *Ion-Selective Electrodes*, 2d ed., 1984; E. Pungor (ed.), *Ion-Selective Electrodes* vol. 5, 1989; R. L. Solsky, *Analytical Chemistry*, *Fundamental Reviews*, vol. 60, pp. 106R–113R, 1988.

Ion-solid interactions

Physical processes resulting from the collision of energetic ions, atoms, or molecules with condensed matter. These include elastic and inelastic backscattering of the projectile, penetration of the solid by the projectile, emission of electrons and photons from the surface, sputtering and desorption of neutral atoms and ions, production of defects in crystals, creation of nuclear tracks in insulating solids, and electrical, chemical, and physical changes to the irradiated matter resulting from the passage or implantation of the projectile (**Fig. 1**). Ion-solid interactions are also known as particle-solid interactions.

When an energetic ion which has been produced in a plasma, a particle accelerator, or an astrophysical process impinges upon the surface of condensed matter, it experiences a series of elastic and inelastic collisions with the atoms which lie in its path. These collisions occur because of the electrical forces between the nucleus and electrons of the projectile and those of the atoms which constitute the solid target. They result in the transformation of the energy of the projectile into internal excitation of the solid. The precise nature of this excitation and the resulting physical processes are determined largely by the bombarding conditions. The principal determining factors are the species of ion, its energy and direction of incidence, the target composition and crystal structure, the target temperature, and the condition of the target surface. Such factors as the quantum-mechanical state of the projectile may also influence specific processes, especially if they are known to occur very near the target surface.

Backscattering. One of the most basic interactions occurs when the projectile collides with a surface atom and bounces back in generally the opposite direction from which it came. This process is known as backscattering. It was first observed in 1911, when Ernest Rutherford bombarded gold foils with alpha particles from radioactive decay, and its observation led Rutherford to conclude that most of the matter in atoms is concentrated in a small nucleus. Now it is used as an analytical technique, Rutherford backscattering analysis, to measure the masses and locations of atoms on and near a surface. By measuring the energy and direction of a backscattered particle whose initial energy and direction are known, the mass of a surface atom that was struck can be inferred. Moreover, when the scattering occurs below the surface, small but predictable shifts in the measured energy can be used to infer the depth of the collision. For projectiles such as alpha particles with energies of a few MeV, the cross section for backscattering is well known. As a result, the fraction of backscattered particles is an absolute measure of the number of targets on the surface. This technique is most commonly performed with alpha particles of about 2 MeV. It is widely used in the study of other ion-solid phenomena and in analyzing the thickness and composition of thin layers, a measurement often of considerable value in research on semiconductor devices. *See* SCATTERING EXPERIMENTS (NUCLEI).

Another backscattering technique, known as ion-scattering spectrometry, uses projectiles with energies of perhaps 2 keV and thus achieves significantly greater surface specificity. However, results are more qualitative since the absolute cross sections for these collisions are not well known. Many of the advantages of Rutherford-backscattering and ion-scattering spectrometries can be combined by the use of ions in the 100–300-keV energy range. This medium-energy backscattering spectrometry has been made possible

by advances in time-of-flight particle detection technology.

Penetration phenomena. Although backscattering events are well enough understood to be used as analytical tools, they are relatively rare because they represent nearly head-on collisions between two nuclei. Far more commonly, a collision simply deflects the projectile a few degrees from its original direction and slows it somewhat, transferring some of its kinetic energy to the atom that is struck. Thus, the projectile does not rebound from the surface but penetrates deep within the solid, dissipating its kinetic energy in a series of grazing collisions.

The capacity of a solid to slow a projectile is called the stopping power $S(E)$, and is defined as the amount of energy lost by the projectile per unit length of trajectory in the solid. The stopping power is a function of the projectile's energy and has different values for different projectiles. It is commonly measured in units of eV/cm or an equivalent unit of eV/(g/cm^2) obtained by dividing the former value by the density of the solid. Stopping power is of central importance for many phenomena because it measures the capacity of a projectile to deposit energy within a thin layer of the solid. Since this energy drives secondary processes associated with penetration, the stopping power of the projectile is an important scaling parameter.

In order to understand the graph of the stopping power of an energetic nucleus in matter as a function of its energy (**Fig. 2**), it is necessary to distinguish collisions involving the nuclei and the electrons of the target.

When one charged particle moves past another at high velocity, the average force acting on the stationary particle is perpendicular to the trajectory of the moving particle. This force acts for a brief period of time and gives the stationary particle a small momentum. To simplify calculations, it is assumed that the target particle, whose mass is m, does not move very much during the collision and that the moving particle is not significantly deflected by it. This is called the impulse approximation. The magnitude Δp of the momentum acquired by the struck particle is independent of its mass so that the energy that it receives in the collision is $(\Delta p)^2/2m$. Thus, an electron whose mass is about 1/2000 that of even a single nucleon will receive much more energy from a collision than a nucleus. The portion of the graph (Fig. 2) for which this simple explanation is satisfactory has a slope proportional to $1/E$ and is labeled the Bethe region. Here a projectile penetrates the target as a bare nucleus devoid of electrons because they have been stripped away in collisions.

At relativistic velocities the electric field of the projectile becomes increasingly concentrated in the plane perpendicular to the trajectory where it can interact with the target more efficiently. As a result, the stopping power eventually reaches a minimum, called the ionization minimum, and then begins to rise again.

For energies above the ionization minimum the stopping power increases, but at substantially higher energies increases less rapidly. This flattening is known as the density effect and is the result of the rapid rearrangement or polarization of the electrons in the medium by the increasingly strong transverse electric field of the projectile. This rearrangement partially shields the projectile and so slows the rate of increase of the stopping power, but the suddenness of the motion results in the emission of electromagnetic radiation. When the speed of the projectile

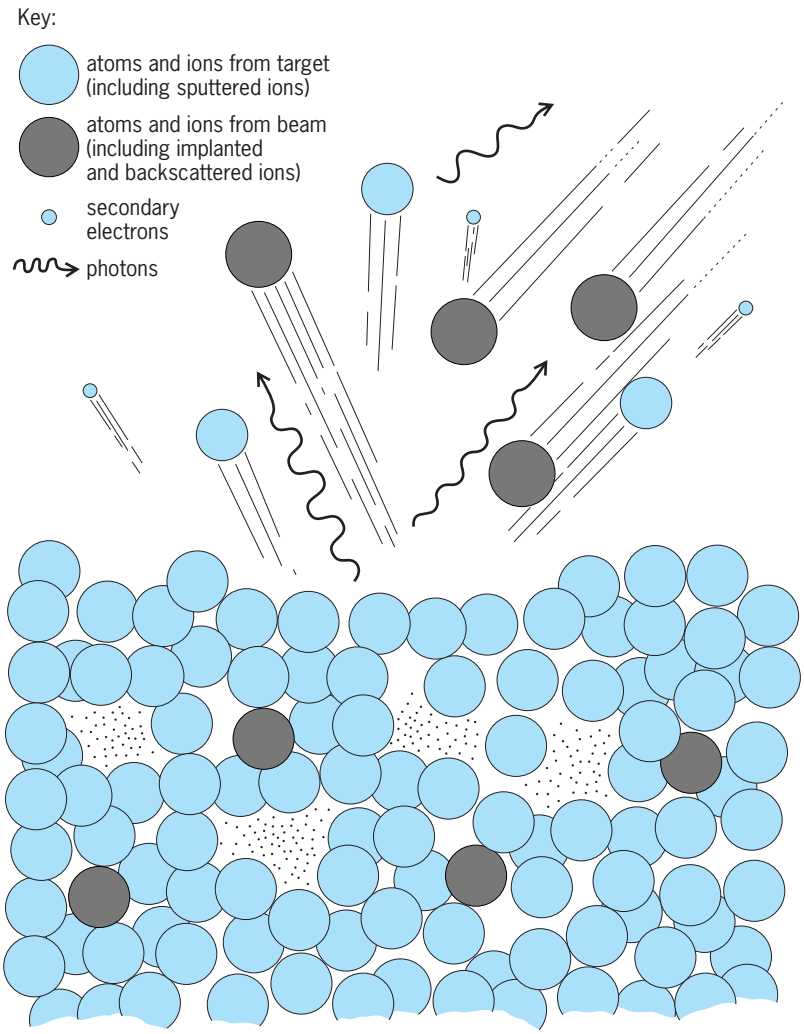

Fig. 1. Processes that can occur when a material is subjected to particle bombardment.

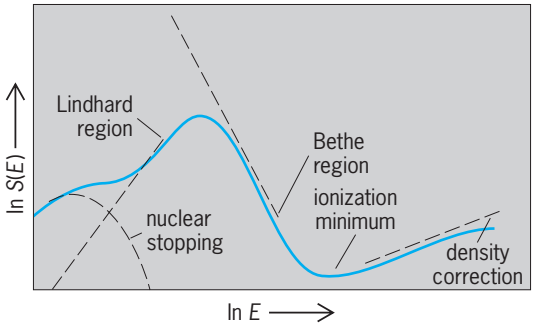

Fig. 2. Graph of the logarithm of the stopping power *S*(*E*) of a heavy ion in matter as a function of the logarithm of its energy *E*.

exceeds the speed of light in the medium, this radiation is called Cerenkov radiation, and its direction is determined by the velocity of the projectile and the medium's index of refraction. *See* CERENKOV RADIATION; RELATIVITY.

As a projectile's energy is lowered below the Bethe region, it begins to capture electrons from the medium. Very roughly, a new electron is captured when the velocity of the projectile in the medium is comparable to the velocity of the electron in its quantum-mechanical orbit about the projectile. The presence of these captured electrons surrounding the penetrating nucleus shields it from the electrons of the medium and thus reduces the stopping power, resulting in a peak in the curve (Fig. 2). However, collisions with target electrons continue to dominate the stopping power below this peak. This portion of the graph is labeled the Lindhard region and has a slope proportional to $E^{1/2}$.

For projectiles with energies typically below a few hundred keV, the collisions are not simply with electrons but are elastic and inelastic collisions between whole atoms. This is called nuclear stopping since the nuclei of struck atoms acquire significant amounts of kinetic energy.

The stopping mechanisms for ions in solids result in their having well-defined ranges. This behavior is distinctly different from the penetration of photons (light, x-rays, and so forth) in matter. Photon intensity in matter decreases exponentially with depth with a characteristic length, but individual photons at any depth have the same energy as they had at the surface. Thus, photons are destroyed but do not lose energy, while ions lose energy but are not destroyed.

In crystalline solids it is possible to pick certain directions where the electron density is significantly lower. Ions penetrating in these channels between crystal rows or planes experience fewer collisions than ions penetrating in other directions. In fact, because an ion's potential energy increases in all transverse directions away from the axis of such a channel, the ion will tend to be constrained to move along it. This phenomenon, known as channeling, has been studied extensively. At higher energies the ion can be excited to higher quantum-mechanical states through its resonant coherent interaction with the channel. When the ion relaxes, it emits radiation. *See* CHANNELING IN SOLIDS.

Radiation effects. The dissipation of a projectile's kinetic energy in a solid can result in a number of secondary phenomena. Near the surface, electrons given sufficient energy may directly escape. These are often accompanied by Auger electrons and x-rays emitted by atoms which have lost inner-shell electrons through collisions. *See* AUGER EFFECT.

In many insulating solids (including mica, glasses, and some plastics) the passage of an ion with a large electronic stopping power creates a unique form of radiation damage known as a nuclear track. The track is actually a cylindrical volume surrounding the trajectory of a single ionizing projectile which has been rendered much more susceptible to chemical attack. As a result, when the substance is chemically etched, conical pits visible under an ordinary microscope are produced where ionizing particles have penetrated. This is particularly significant since the passage of single projectiles is registered and may be subsequently observed. Although the mechanism of track formation is still the subject of theoretical research, track-forming materials are widely used to detect ionizing radiation. *See* PARTICLE TRACK ETCHING.

When the projectiles have energies in the nuclear stopping range, a number of additional phenomena may occur. Of particular significance is the formation of a collision cascade. In the nuclear stopping region it is relatively likely that the projectile will transfer significant amounts of energy to individual target atoms. These atoms will subsequently strike others, and eventually a large number of atoms within the solid will be set in motion. This disturbance is known as a collision cascade. Collision cascades may cause permanent damage to materials, induce mixing of layers in the vicinity of interfaces, or cause sputtering if they occur near surfaces. The dynamical description of collision cascades is extremely complicated because of the large numbers of particles involved. Computer modeling and a branch of statistical mechanics called transport theory have been most successful, although both have limitations. The comprehensive mathematical description of collision cascades with arbitrary energy densities, realistic interatomic potentials, and consistent treatment of target surfaces is an active area of research. *See* BOLTZMANN TRANSPORT EQUATION; CRYSTAL DEFECTS; ION BEAM MIXING; RADIATION DAMAGE TO MATERIALS; SIMULATION; SPUTTERING.

Although it is a much less ubiquitous phenomenon than collisional sputtering, particle emission from insulating surfaces can also be initiated by ions with energies in the electronic stopping region by mechanisms which are collectively termed desorption. Where sputtering is primarily a statistical process understandable in terms of classical mechanics, desorption is inherently quantum mechanical, involving specific electronic transitions. As a result, desorption can be initiated equally efficiently by ions, electrons, and photons. Very large and fragile biological molecules can be removed from surfaces virtually intact following the impacts of MeV heavy ions or pulsed laser irradiation. This has made it possible to mass-analyze a large number of biological molecules which could not previously be studied by conventional mass spectrometry because they were too fragile to be thermally evaporated. *See* DESORPTION.

Surface interactions. Much work in the field of ion-beam interactions with matter has dealt with effects to the material and the projectile which occur in the region of the surface. Projectiles with energies ranging from chemical bond energies to penetration energies can be used to study surface structure, prepare unique deposition layers, include unusual chemical reactions, and even select specific reaction mechanisms. For example, bombardment of a surface by 150 keV C^+ ions has been shown to be one route to the formation of diamondlike carbon. Similarly, by studying the atomic state of hyperthermal ions

such as Li^+ which scatter from the surface, information about both the structure of the surface and the nature of the interaction itself can be inferred, thus providing valuable additional insights to those which have been available previously from beam-foil spectroscopy. *See* BEAM-FOIL SPECTROSCOPY.

Complex and fascinating phenomena have been observed in low-energy ion-surface collisions. They include instances where ions appear to make multiple impacts in much the same way as a stone skips on the water. The electron-cyclotron-resonance (ECR) ion source has made it possible to study the impacts of extremely slow, yet very highly charged ions with surfaces. One result of this work is the recognition that at distances greater than 1 nm from the surface these nearly bare nuclei may capture significant numbers of electrons into their high-lying quantum states while lower-lying states may remain essentially empty. The resulting structures, called hollow atoms, may survive in this condition briefly after impact with the surface. *See* ION SOURCES.

Advances in computing have made it possible to simulate ion-surface impacts with far-reaching consequences for the development of intuition about these events (**Fig. 3**). Such simulations represent essential tools in the design and interpretation of experiments. *See* FULLERENE; MOLECULAR MECHANICS.

Applications. The most ambitious use of ion beams in industry is in the manufacture of integrated circuits. Ion implantation can be used to modify the electrical properties of semiconductors far more precisely than other techniques. As a result, much more elaborate structures can be produced. Ion implantation is an essential element of the manufacture of virtually all digital integrated circuits in use. *See* INTEGRATED CIRCUITS; ION IMPLANTATION.

Another commercial application of ion beams with significant promise is for materials modification. It has been demonstrated that ion implantation can greatly improve the wear and corrosion resistance of metals. One area that is expected to benefit from these advances is the technology of human prosthetic implants. *See* PROSTHESIS.

Ion-solid processes are used extensively as tools in many areas of research. They permit highly sensitive analyses for trace elements, the characterization of materials and surfaces, and the detection of ionizing radiation. Techniques employing them include secondary ion mass spectrometry (SIMS) for elemental analysis and imaging of surfaces, proton-induced x-ray emission (PIXE), ion-scattering spectrometry (ISS), medium-energy and Rutherford backscattering spectrometries (MEBS and RBS), and nuclear reaction analysis (NRA) for elemental and isotopic depth profiling. Ion-solid interactions are also fundamental to the operation of silicon surface-barrier detectors which are used for the measurement of particle radiation, and of nuclear track detectors which have been used in research as diverse as the dating of meteorites and the search for magnetic monopoles. Among the notable products of ion-solid research are the development of ion-beam assisted deposition (IBAD) for the rapid production of thick protective coatings, and the demonstration of thin-film waveguides and waveguide lasers produced in insulators by ion implantation. *See* ACTIVATION ANALYSIS; CHARGED PARTICLE BEAMS; MAGNETIC MONOPOLES; PARTICLE DETECTOR; PROTON-INDUCED X-RAY EMISSION (PIXE); SECONDARY ION MASS SPECTROMETRY (SIMS); SURFACE PHYSICS.

Robert A. Weller

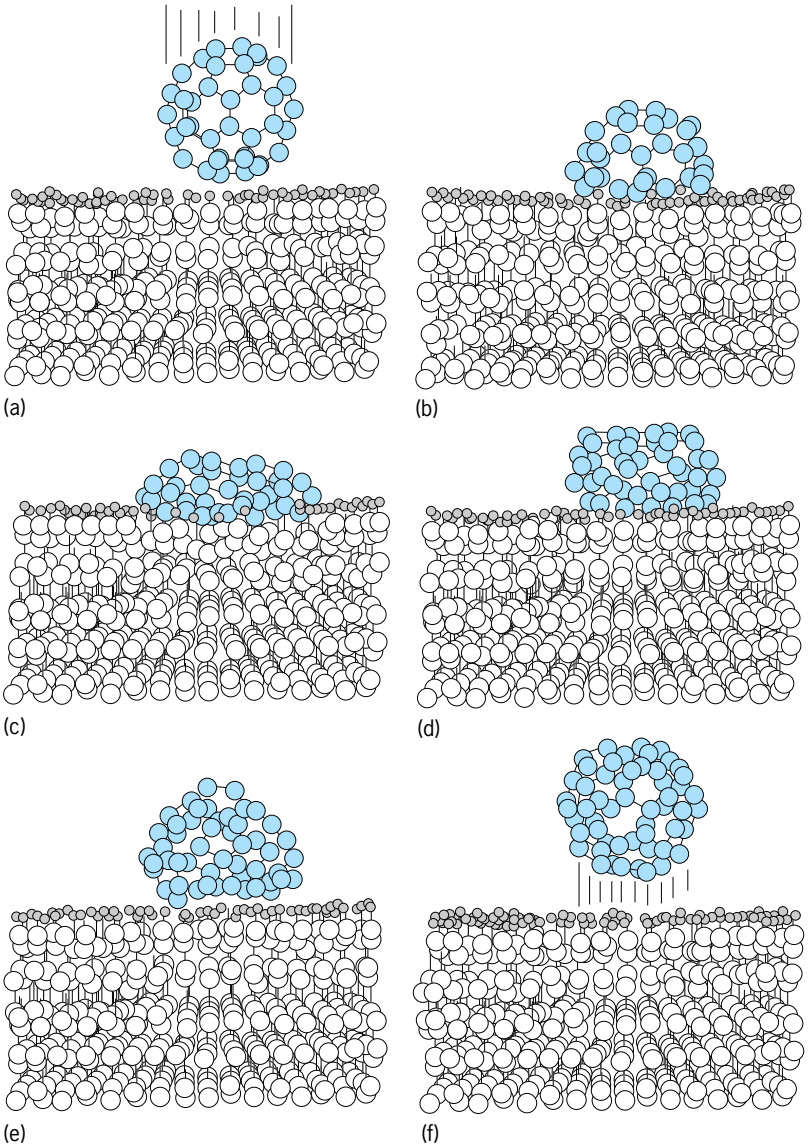

Fig. 3. Molecular dynamics simulation of a 250-eV C_{60}^+ molecular ion (buckminsterfullerene or buckyball) scattering from a hydrogen-terminated diamond surface. Atomic positions are shown at various times. (*a*) 0 femtosecond (*b*) 40 fs. (*c*) 100 fs. (*d*) 160 fs. (*e*) 201 fs. (*f*) 280 fs. (*After R. C. Mowrey et al., Simulations of C_{60} collisions with a hydrogen-terminated diamond [111] surface, J. Phys. Chem., 95:7138–7142, 1991*)

Bibliography. R. Behrisch (ed.), *Sputtering by Particle Bombardment*, 3 vols., 1981, 1983, 1991; W. K. Chu, J. W. Mayer, and M. A. Nicolet, *Backscattering Spectrometry*, 1978; J. P. Eberhart, *Structural and Chemical Analysis of Materials*, 1991; L. C. Feldman and J. W. Mayer, *Fundamentals of Surface and Thin Film Analysis*, 1986; J. R. Tesmer et al. (eds.), *Handbook of Modern Ion Beam Materials Analysis*, 1995; J. F. Ziegler, J. P. Biersack, and V. Littmark, *The Stopping and Range of Ions in Solids*, 1985.

Ion sources

Devices that produce positive or negative electrically charged atoms or molecules. *See* ION.

In general, ion sources fall into three major categories: those designed for positive-ion generation, those for negative-ion generation, and a highly specialized type of source designed to produce a polarized ion beam. The positive-ion source category may further be subdivided into sources specifically designed to generate singly charged ions and those designed to produce very highly charged ions.

Desirable qualities of an ion source are large yield, high ionization efficiency, low energy spread, high brightness, and low emittance. Practical considerations such as reliability, long source life, and ease of changing ion species are also important.

Ion sources have acquired a wide variety of applications. They are used in a variety of different types of accelerators for nuclear research; have application in the field of fusion research; and are used for ion implantation, in isotope separators, in ion microprobes, as a means of rocket propulsion, in mass spectrometers, and for ion milling. *See* ION IMPLANTATION; ION PROPULSION; ISOTOPE SEPARATION; MASS SPECTROSCOPE; NUCLEAR FUSION; PARTICLE ACCELERATOR; SECONDARY ION MASS SPECTROMETRY (SIMS).

Methods of Positive-Ion Formation

The principal methods of positive-ion formation are electron impact, surface ionization, spark discharge, laser ionization, field ionization, thermal ionization, and sputtering.

Electron impact. A common method of ionizing a gas or vapor is to pass high-velocity electrons through it, with the ions being formed as a result of electron-atom collisions. Electron energies are typically a few 100 eV but in some special sources may be as high as 20 keV. An externally applied magnetic field is frequently used to cause the electrons to travel along a helical path and thereby increase the ionization efficiency. Examples of ion sources utilizing this concept are the duoplasmatron and the Penning ion source.

Positive surface ionization. Atoms possessing low ionization potentials can be ionized by allowing them to strike a heated surface having a high work function. Provided that the ionization potential of the atom is less than or about equal to the work function of the surface, there is a high probability that the atom will be thermally desorbed as a positive ion. The method is particularly well suited, though not entirely restricted, to producing ions of the alkali metals, all of which have ionization potentials of less than 5.4 eV. Some high work function metals are platinum (approximately 5.7 eV), tungsten (approximately 4.5–5.5 eV), and rhenium (approximately 5 eV). *See* IONIZATION POTENTIAL; WORK FUNCTION (ELECTRONICS).

Spark discharge. There are several variations of this technique, but basically a spark is induced between two electrodes, one of which, at least, contains the element to be ionized. Generally speaking, the spark consists of a high-density, high-temperature plasma from which ions can be extracted. The spark can be produced by applying a high alternating potential between two fixed electrodes or by mechanically breaking contacting electrodes. *See* ELECTRIC SPARK.

Laser ionization. A focused beam from a high-power pulsed laser can be used to produce a small ball of dense plasma from essentially any solid, and positive ions can be extracted from this plasma. The high temperature of the plasma results in the formation of many multiply stripped ions and thus may prove a very effective method of generating highly charged positive ions.

In principle, lasers or other strong sources of electromagnetic radiation can be used to produce ions by photoionization. A photon can directly ionize an atom if its energy exceeds the ionization potential of the atom. Unfortunately, the probability of photoionization is low (the cross section is of the order of 10^{-19} cm^2), making it difficult to design efficient ion sources based on this process. *See* LASER.

Field ionization. If an atom passes close to or gets absorbed on a very sharp point where the electric field exceeds a few times 10^{10} V/m, there is a probability that it will get ionized; the phenomenon is known as field ionization. Such large electric fields can be achieved in the vicinity of a specially sharpened tungsten needle placed close to an annular electrode, and gas or vapor passing close to the tip of the needle can be ionized.

Field emission or ionization is generally believed to be the underlying operating principle of a novel type of ion source known as the electrohydrodynamic source. In this source a conducting liquid, usually a metal, is allowed to be drawn down a fine-bore tube by capillary action. When an electric field is applied to the tip of the tube, the liquid meniscus, normally spherical, distorts and becomes conical. As the electric field is increased, the tip of the cone becomes sharper, and eventually the field at the tip becomes sufficiently large to cause field emission. *See* FIELD EMISSION.

Thermal ionization. Although the term thermal ionization is ill-defined, it is generally used in the context of heating certain complex compounds, resulting in positive-ion emission. An example is the emission of lithium ions from a heated surface coated with β-eucryptite (a lithium aluminosilicate mineral). The technique has found extensive application in mass spectroscopy to produce a wide variety of ions. The sample to be ionized is usually mixed with silica gel and phosphoric acid and deposited on a rhenium filament. After a preliminary baking, the filament is introduced into the ion source and heated to the point of positive-ion emission.

Sputtering. When a solid is bombarded with energetic heavy ions, a fraction of the sputtered particles leaves the surface as ions. This fraction is usually too low for direct application in an ion source, but the sputtering process is frequently used to introduce

solids into an electron impact source such as a Penning source. *See* SPUTTERING.

Methods of Negative-Ion Formation

All elements can be positively ionized, but not all form stable negative ions. For example, none of the noble gases forms negative ions. However, helium is an exception in that it does have a metastable negative ion with a lifetime of about 1 millisecond. The noble gases are not the only elements that do not form stable negative ions, but most form metastable ones with lifetimes long enough to permit acceleration. Nitrogen has an exceptionally short-lived metastable negative ion, and it is customary to accelerate either NH^- or NH_2^- molecular ions, both of which are stable.

Direct extraction. Most discharge sources, such as the duoplasmatron (**Fig. 1**), yield negative ions when the polarity of the extraction voltage is reversed. However, the yield is usually low, the electron current is high, and there are difficulties when operating them with elements other than hydrogen. These sources are now used almost exclusively to generate intense beams of hydrogen and its isotopes, and several important changes have been made to improve their performance. The negative-ion-to-electron yield from a direct extraction duoplasmatron can be greatly improved if the intermediate electrode is displaced by a millimeter or so off the source axis (Fig. 1*b*). The introduction of cesium vapor into the discharge plasma greatly enhances the negative-ion yield. Negative hydrogen currents from such sources have been increased from a few microamperes to several tens of milliamperes.

Charge exchange. When a positive-ion beam, ranging in energy from a fraction of a kiloelectronvolt to several tens of kiloelectronvolts, is passed through a gas or vapor, some of the ions emerge negatively charged. At the optimum energy (depending upon the ion), and with vapors of lithium, sodium, and magnesium, the negatively charged fraction is quite high, ranging from 1 to 90%. The technique is also highly effective for the creation of metastable negative ions, such as helium.

Cesium-beam sputtering. When a solid surface is sputtered with cesium positive ions, a surprisingly large fraction of the sputtered particles emerge as negative ions. This fraction can be increased by almost an order of magnitude by overlaying the sputter surface with additional neutral cesium. Highly efficient and versatile negative-ion sources involving cesium sputtering have been developed.

Negative surface ionization. Just as positive ions of elements possessing low ionization potentials can be generated by thermally desorbing them from high work-function surfaces, negative ions having large electron affinities can be similarly generated on a low work-function surface. The method is particularly suited to generating negative ions of the halogens, all of which have electron affinities in excess of 3 eV. A particularly suitable low-work-function surface is lanthanum hexaboride, which is reported to have a work function of about 2.6 eV.

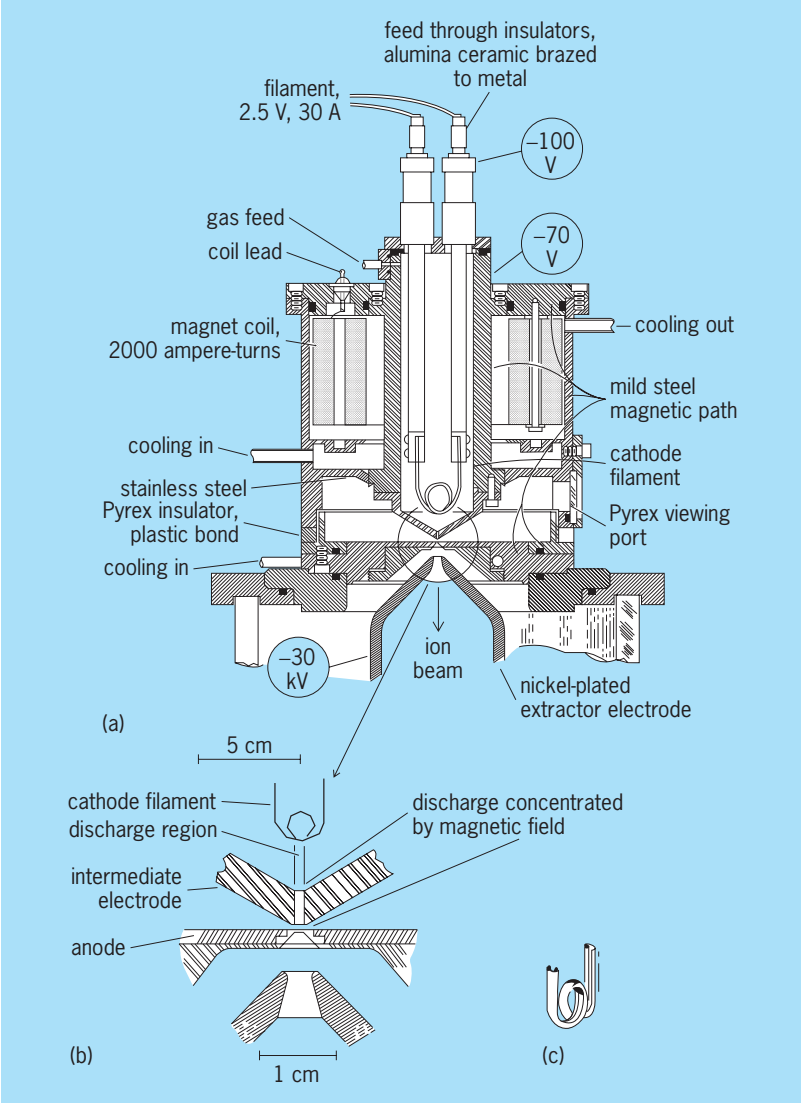

Fig. 1. Duoplasmatron ion source for accelerators. (*a*) Diagram of entire source. (*b*) Detail of discharge and extraction regions. (*c*) Drawing of cathode filament. 1 cm = 0.4 in.

Positive-Ion Source Concepts

Positive-ion source concepts include the duoplasmatron source for protons and the Penning ion source and ion confinement sources for multiply charged heavy ions.

Duoplasmatron. The duoplasmatron (Fig. 1) is a high-current proton source. It makes use of an arc discharge which is constricted as it passes into a very strong magnetic field shaped by iron or mild steel inserts in an intermediate electrode and anode. The beam is extracted at the point where the arc has reached a very small diameter and a very high brilliance. Sources of this type have been developed for accelerators. *See* ARC DISCHARGE.

Heavy-ion sources. The term heavy ion is used to designate atoms or molecules of elements heavier than helium which have been ionized. As was mentioned earlier, ion sources that are used to generate such ions fall into two categories: those intended to form singly charged ions and those designed to

produce multiply charged ions. A heavy ion can be singly ionized (one electron removed or added), can be fully stripped as in argon 18+, or can have any intermediate charge state.

Singly charged heavy ions are most frequently used in isotope separators, mass spectrographs, and ion implantation accelerators. These are much easier to generate than multiply charged ions, and frequently the experimenter has several source concepts to choose from, depending upon the application and the physical characteristics of the element to be ionized.

Penning ion sources. This source is based on a high-current gaseous discharge in a magnetic field with gas at a relatively low pressure (10^{-3} torr or 0.1 pascal). The source (**Fig. 2**) consists of a hollow anode chamber, cathodes at each end, a means for introducing the desired element (usually a gas), and electrodes for extracting the ions (not shown). The cathode may be heated to emit electrons, which then help to initiate the arc discharge current, creating the plasma in which the atoms are ionized. The discharge column between the cathodes (the plasma) consists of approximately equal numbers of low-energy electrons and positive ions. The electron density is much larger than can be accounted for by the primary electrons from the cathodes. The average energy of plasma electrons may range from a few volts to a few tens of volts. Electrons travel parallel to the magnetic field, are reflected from the opposite cathode, and make many traversals of the length of the hollow chamber. The electrons confined by the magnetic field and the cathode potential thereby have a high probability of making ionizing collisions with any gas present in the chamber.

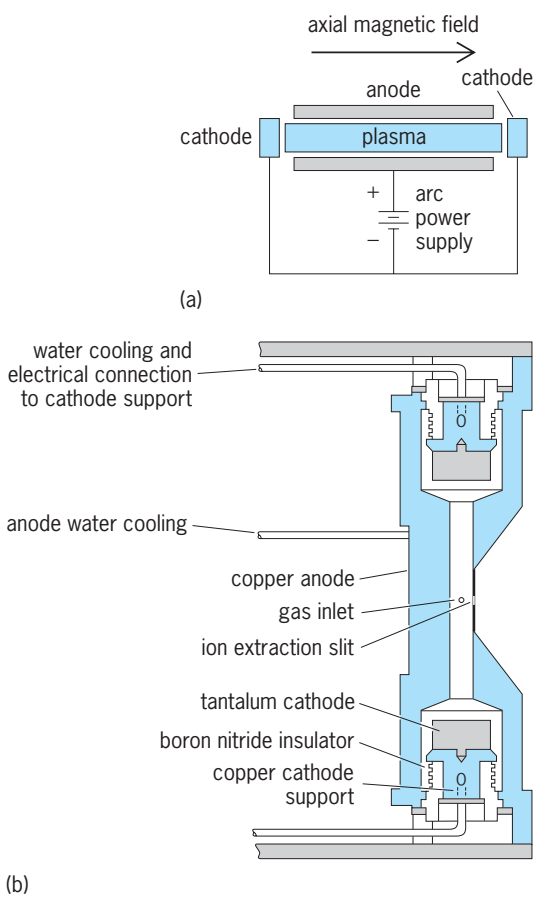

Fig. 2. Penning ion source. (*a*) Schematic diagram illustrating basic principles. (*b*) Section showing geometry. (*After J. R. J. Bennett, A review of PIG sources for multiply charged heavy ions, IEEE Trans. Nucl. Sci., NS-19(2):48–68, 1972*)

TABLE 1. Penning heavy-ion source performance

	Charge state distributions (I^{Q+}/I^{1+})								
Element	1+	3+	4+	5+	6+	7+	8+	9+	12+
Argon	1	8.5	8	3	0.8	0.09	0.01		
Calcium	1	23	22	15	3	0.3	0.035		
Krypton	1	5	5.5	6	4.5	4	2.2	0.6	
Xenon	1	7	9	9	7.5	5.5	4	1.5	0.025

The net result of all the processes in the arc plasma is that some partially stripped atoms diffuse perpendicular to the magnetic field out of the arc, experience the field of the accelerating electrode, and are moved into the accelerator. The arc potential may be constant with time, or it may be pulsed so that ions are produced as needed by the accelerator. High yields of charge states 1+ through 8+ (and for heavier elements, perhaps up to 12+) have been obtained for many elements of the periodic table (**Table 1** shows four of these). *See* ARC DISCHARGE; PLASMA (PHYSICS).

Most studies with Penning sources were made with gases, thus restricting nuclear research programs to a rather limited number of projectiles, until it was discovered that some of the krypton ions from the source were partially accelerated and then returned back into the source where they sputtered the source body material (copper) into the discharge, resulting in a prolific beam of highly charged copper ions. It was a short step to introduce small pieces of various solids immediately behind the ion extraction slit (Fig. 2*b*) and to use the source to produce beams of a wide range of elements. Roy Middleton

Electron cyclotron resonance sources. An electron cyclotron resonance source for highly charged ions consists of a metallic box occupied by low-pressure vapors, microwaves, and specific magnetic fields for plasma confinement. The box can be of any shape but must be large with respect to the wavelength of the microwaves in order to act as a multimode cavity. Plasma confinement is generally obtained by superimposing solenoidal and multipolar fields such that the modulus of the magnetic field is minimal in the center of the box and maximal near the walls. In between, there should be a closed magnetic surface where the Larmor frequency of the electrons equals the frequency of the injected microwaves. Electrons crossing this surface are then energized by the electron cyclotron resonance. When they pass many times through the resonance, according to their random phase, they acquire a global stochastic electron cyclotron resonance heating, yielding

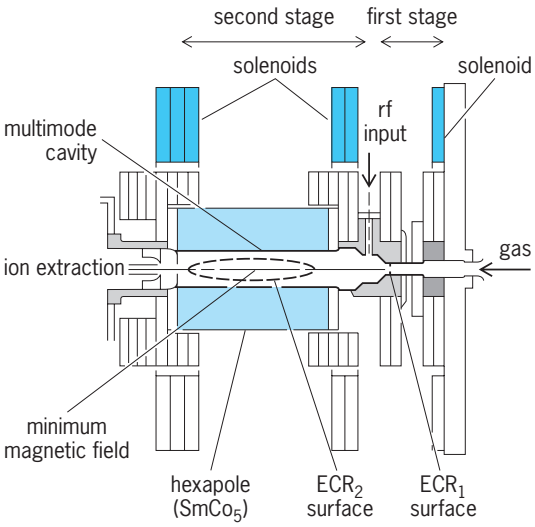

Fig. 3. Electron cyclotron resonance ion source.

energies of tens of kiloelectronvolts, thus exceeding the ionization potentials of many highly charged ions. However, since the ion charges are obtained step by step through successive electron–ion collisions, the ions need long exposure times to many electron impacts (that is, long plasma confinement). The source (**Fig. 3**) generally has two stages in series. In the first, a cold plasma is created at ECR_1, which diffuses toward a very low-pressure second stage, where ECR_2 energizes the electrons inside the confined plasma. Owing to the absence of electrodes, the longevity of the source is unlimited. The magnetic fields are obtained with ordinary or superconducting coils, permanent magnets,or combination of these. The operation of the source can be continuous or pulsed. Most sources utilize frequencies between 2.4 and 16 GHz. The performance improves with increasing frequency. *See* CYCLOTRON RESONANCE EXPERIMENTS; MAGNET; SUPERCONDUCTING DEVICES.

R. Geller

Electron-beam sources. The electron-beam ion source (EBIS) produces the most highly charged ions, up to fully stripped U^{92+}. It can be operated in many different modes depending on the application, the desired ion species, charge state, and ion-beam characteristics, including its duty cycle. The EBIS normally produces ion beams with a small emittance and a narrow energy spread. The highest charge states are normally produced in a batch mode, yielding ion pulses with a charge as high as 1 nC, which allows for instantaneous peak currents as high as 1 μA. The time-averaged currents are normally in the range of nanoamperes to picoamperes because of the required batch production time, which increases with the desired charge state.

The majority of EBISs or electron-beam ion traps (EBITs) are used to study how the produced highly charged, low-energy ions interact with the internal electron beam (in the EBIT) or with external gaseous or solid targets (in the EBIS). Others are used as injectors for pulsed accelerators, storage rings, or ion traps. A few are used to study further developments or new applications.

The ion source (**Fig. 4**) is based on an electron beam which successively strips the ions until the desired high charge state is reached. The electron beam is launched from an electron gun, confined and normally compressed by a strong magnetic solenoid, and later adsorbed by a collector. Ions trapped inside

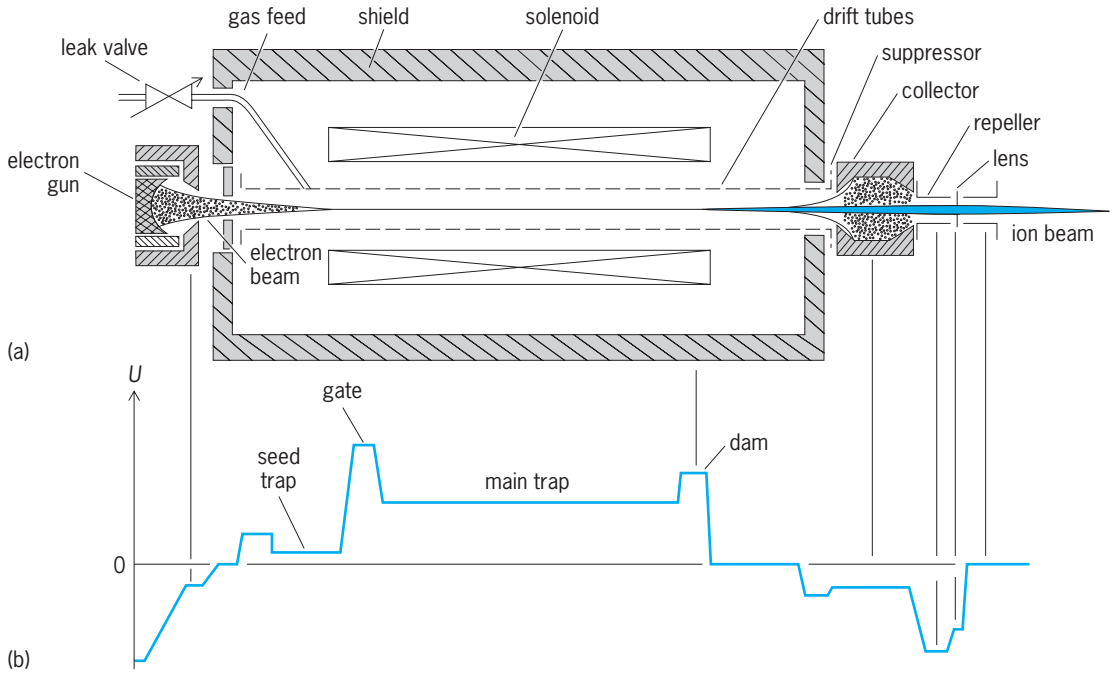

Fig. 4. Electron-beam ion source. (*a*) Cross section. (*b*) Electric potential (*U*) applied to the drift tubes during the production of highly charged ions.

the space charge of the electron beam are ionized at a rate which depends on the electron beam's energy and density, with high-power electron beams and high-field solenoids producing high rates. The charge states produced increase with increasing confinement time until the ionization rate is matched by the rate of electron capture mainly in collisions with the residual gas. Hence the highest charge states are produced with cryogenic EBISs which use a superconducting high-field solenoid (3–6 tesla) to achieve high electron-beam compression and use the required liquid helium to cryopump the production region to reduce the residual gas densities. The electron beam is surrounded by a series of drift tubes, with their potential being used to adjust the electron-beam energy and to control the flow of the ions, allowing for the variation of the confinement time. The highest charge states are normally produced in a batch mode which begins with the injection of low-charged seed ions either from an external ion source or from an internal seed trap in the fringing fields of the solenoid, where the seed gas is ionized by the lower-density electron beam. After the gate or the dam has been raised, the ions are confined in the axial direction with potentials that exceed the potential applied to the intermediate drift tubes, the main trap. After the confinement time required to reach the desired charge state (10^{-4} to 50 s), the main-trap potentials is raised until the ions escape the source. A slow rise of the main-trap potential decreases the ion beam's energy spread and simultaneously decreases the instantaneous current while increasing the duty cycle. *See* MAGNET; SOLENOID (ELECTRICITY); VACUUM PUMP. Martin P. Stockli

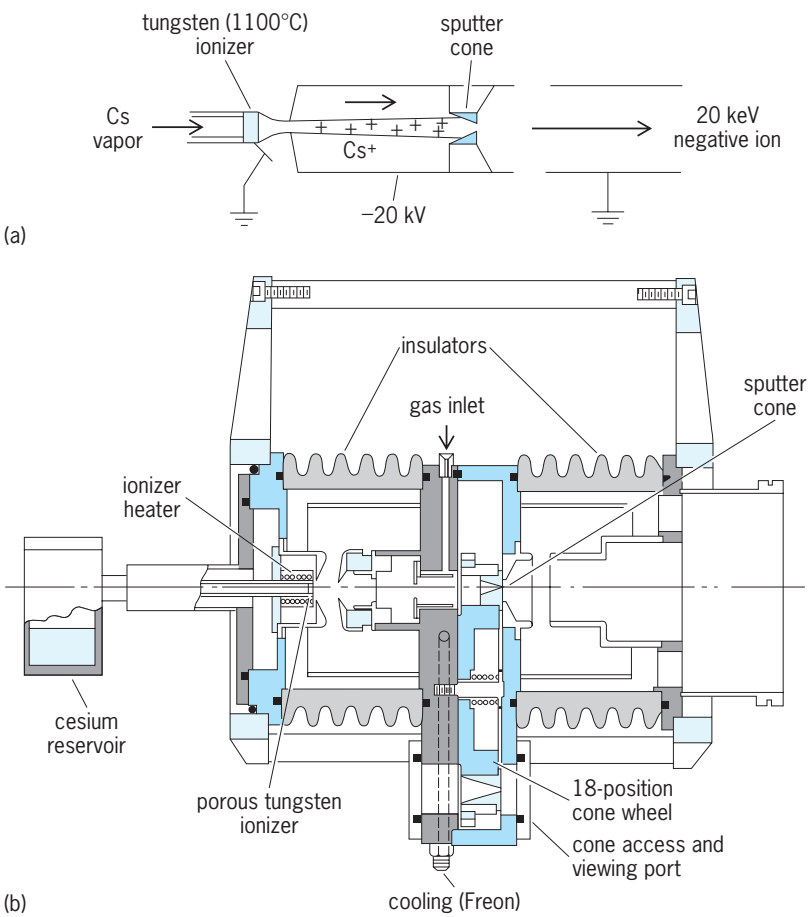

Fig. 5. Cesium-beam sputter source: (*a*) schematic showing operating principles; (*b*) section of a typical source.

Negative-Ion Source Concepts

Negative-ion source concepts will be discussed, including the methods of negative-ion formation discussed above.

Charge exchange source. Although charge exchange can be used to provide a very large variety of negative ions, it is relatively infrequently used, with the exception of producing the metastable negative ion of helium. Indeed, charge exchange is the sole method of producing negative helium beams.

In a typical charge exchange source, about 1 milliampere of positive helium ions is generated in a duoplasmatron ion source which is usually at ground potential. The ions are extracted by an electrode at a potential of about −20 kV, focused by an electrostatic lens, and directed through a donor canal usually containing lithium vapor. In the vapor a fraction of the positively charged helium ions sequentially picks up two electrons to form negative ions. The fraction is usually a little less than 1%, resulting in a negative helium beam of several microamperes.

Cesium-beam sputter source. The cesium-beam sputter source is the negative-ion source most widely used on tandem accelerators. Much of its success is due to the fact that ion species can be changed rapidly, the ion output is usually high (**Table 2**), and the source will operate for several hundreds of hours before a major cleanup is necessary.

TABLE 2. Negative-ion currents from cesium-beam sputter source

Element	Negative-ion current, μA	Element	Negative-ion current, μA
Lithium	2	Sulfur	20
Boron	2	Nickel	6
Carbon	>50	Copper	3
Oxygen	>50	Gold	10
Silicon	20	Lead	0.1

Cesium positive ions are formed in the source (**Fig. 5*a***) by passing cesium vapor through a porous disk of tungsten, heated to about 2012°F (1100°C), by the surface ionization process. These ions are accelerated through a potential difference of between 20 and 30 kV and are allowed to impinge upon a hollow cone that is either fabricated from or contains the element whose negative ion is required. Sputtering of the cesium-coated inner surface of the cone results in a large fraction of the sputtered particles leaving the surface as negative ions. An appreciable fraction of these are extracted from the rear hole of the sputter cone because of electric field penetration and are accelerated toward the ground electrode as a beam.

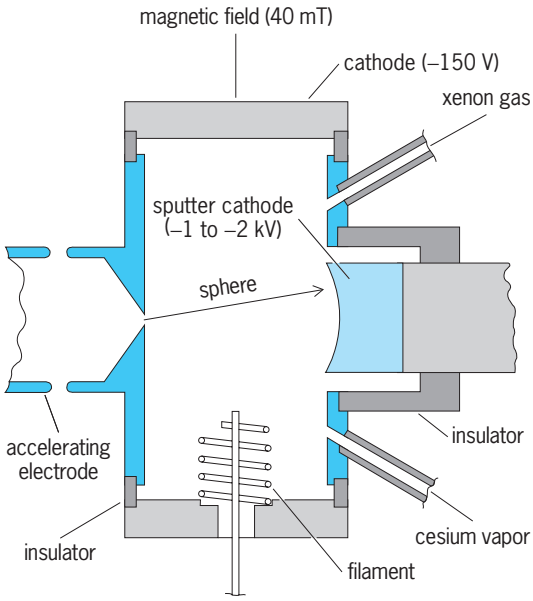

Fig. 6. Cesium-vapor Penning source for generation of negative ions.

The sputter cones or targets can be inserted into a cooled copper wheel resembling the chamber of a revolver (Fig. 5*b*). Thus, by rotating the wheel, sputter cones can be rapidly changed, enabling the negative-ion species to be quickly changed. In addition, negative-ion beams of gaseous elements, such as oxygen, can be formed by leaking the gas into the source and directing the flow onto a suitable getter surface such as titanium.

Cesium-vapor Penning source. The cesium-vapor Penning source (**Fig. 6**) is a direct-extraction negative-ion source. Basically this is a conventional Penning source but with two important modifications making it suitable for negative-ion generation. The first is the introduction of a third sputter cathode, which is the source of negative ions. This cathode is made from or contains the element of interest, has a spherical face centered on the extraction aperture, and is operated at a higher negative potential than the normal cathodes. The second change involves introduction of cesium vapor into the arc chamber in addition to the support gas (usually xenon).

The source operates in the normal Penning mode, and some of the cesium vapor introduced into the arc chamber becomes ionized and is accelerated toward the third sputter cathode. The negative ions that are formed as a result of sputtering are focused and accelerated toward the extraction aperture, and under the influence of the strong electric field generated by the acceleration electrode are extracted as an intense low-divergence beam. The negative-ion yield of the source is quite good and is comparable to that of the cesium-beam sputter source. Roy Middleton

Polarized Ion Sources

A polarized ion source is a device that generates ion beams in such a manner that the spins of the ions are aligned in some direction. The usual application is to inject polarized ions into a particle accelerator; however, it also has applications in atomic physics. The possible types of polarized sources are numerous, for in theory the nuclei of all kinds of atoms can be polarized, provided their spin is not zero. *See* SPIN (QUANTUM MECHANICS).

The original type of source, developed in 1956, generates only positive ions, while the metastable-state or Lamb-shift type of polarized ion source produces a high-quality negative ion beam with a high degree of polarization. The older type of source is referred to as the conventional or ground-state type of polarized ion source. Its output current of positive ions is an order of magnitude larger than the negative-ion output from the Lamb-shift source. With these two types of sources and their variants, polarized ions have been obtained from hydrogen, deuterium, tritium, helium-3, both isotopes of lithium, and others. The extra complication involved in producing polarized ions is such that the output is a factor of a thousand or more below the output of a moderately sized unpolarized ion source.

Conventional or ground-state source. In this type of source (**Fig. 7**), the first step consists in forming a beam of atoms in the ground state by a technique similar to that used in molecular beams. In the case of hydrogen or deuterium, this is done by dissociating the molecular gas in a glass discharge tube and allowing the atoms to escape through a nozzle. The atoms escaping at thermal energies are collimated

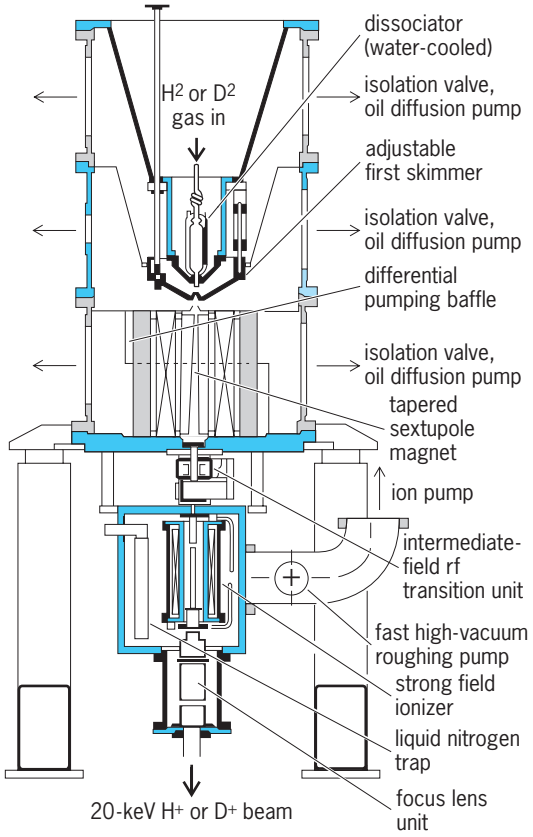

Fig. 7. Conventional or ground-state polarized ion source. ***(After H. F. Glavish, Polarized ion sources, in Proceedings of the 2d Symposium on Ion Sources and Formation of Ion Beams, Berkeley, pp. IV-1-1 through IV-1-7, 1974)***

into a well-directed beam by plates with holes or an adjustable skimmer. High-capacity diffusion pumps sweep away the large quantity of excess hydrogen or deuterium. *See* MOLECULAR BEAMS.

The beam is then passed along the axis of an inhomogeneous magnetic field, which is most commonly generated by a sextupole magnet. This type of magnet consists of six magnets arranged in a circular pattern with alternating polarities.

In a sextupole magnet the absolute magnitude of the field increases as the distance squared from the axis. The atoms are subjected to a radial force that is proportional to their magnetic moment times the gradient of the absolute magnitude of the field strength. In the case of a sextupole magnet the force is proportional to the first power of the distance from the axis. The sign of the force does not depend upon the direction of the magnetic lines, but only on the projection of magnetic moment along these lines of force or m_j, where j is the spin value of the electron. (The atomic magnetic moment results almost entirely from the electron). The result is that atoms with a positive value of m_j are subjected to a force that is directed radially inward and pass through the sextupole magnet, while atoms with a negative m_j experience a force that is directed outward and are rapidly lost from the beam. Out of the sextupole comes a beam of atomic hydrogen that is polarized with respect to the orientation of its electrons but is, as yet, unpolarized in its nuclear spin. *See* ELECTRON SPIN; MAGNETON.

Since aligned nuclei rather than aligned electrons are desired, it is necessary to subject the atomic beam to other fields. Each hydrogen atom is in one of two pure states. It is possible to apply an oscillating magnetic field in combination with a dc magnetic field that will flip the sign of m_I (the projection of the nuclear spin) of one of the pure states and not the other. That aligns the spins of the nuclei but may depolarize the electrons. That does not matter, however, since they will be removed.

The final stage is to send the atomic beam into a strong solenoidal magnetic field. As the atoms from the sextupole field—having all orientations in each cross-sectional plane—enter the solenoid, they adiabatically come into alignment with the parallel lines of force within the solenoid since their m_j components of spin are conserved. In the solenoid the atoms are ionized by energetic electrons as in an arc discharge. The ionizer is actually the most difficult part of this type of polarized source to make function efficiently, even though it is conceptionally simple. The ionizer is followed by electric fields that accelerate and focus the ions to get a beam that can be accepted by the accelerator.

Lamb-shift or metastable-atom source. The polarization process in the Lamb-shift type of source is also performed upon atoms, in this case, metastable ones. The process is most efficient if the atoms have a velocity of approximately 10^{-3} of light rather than thermal velocity as in the case of the ground-state type of source. To get the beam, hydrogen, deuterium, or tritium can be used, but only hydrogen is discussed in this article. The hydrogen is ionized in a conventional ion source such as a duoplasmatron. The H^+ ions are then accelerated and focused into a beam at about 500 eV. The beam is passed through cesium vapor where cesium atoms donate electrons which are resonantly captured in an $n = 2$ state by the hydrogen ions.

Atoms are formed in both the $2p$ and the $2s$ states in the cesium vapor. However, those in the $2p$ state decay almost immediately to the ground state. The small energy difference between the $2p$ and the $2s$ states is the Lamb shift. The lifetime of the $2p$ atoms is 1.6×10^{-9} s, while the lifetime of the $2s$ atoms is 0.15 s because two photons must be emitted simultaneously in their decay to the ground state. Actually few $2s$ atoms decay by emission of two photons for they are necessarily subjected to small electric fields which mix into the $2s$ the $2p$ wave function and its tendency to decay to the ground state. To take advantage of this tendency to decay to the ground state, apparatus can be built so that those atoms which have the undesired value of m_I are stimulated to decay, while those with the desired value of m_I are allowed to pass on without decay. *See* ATOMIC STRUCTURE AND SPECTRA; FINE STRUCTURE (SPECTRAL LINES).

The polarized H^- ions are formed in argon because its atoms are capable of donating electrons to metastable atoms but have a very weak capability of forming H^- ions out of ground-state atoms. The ground-state charge-changing cross section appears to be lower by a factor of about 400; however, ground-state atoms outnumber the metastable atoms at this region by a factor of 40 so that the net polarization is 90%. The remainder of the apparatus consists of electric fields that accelerate and focus the beam so it can be accepted by an accelerator.

The electron spins are polarized by applying a transverse field of about 100 kV/m while the atomic beam of metastables is passing along the axis of a solenoid at a field strength of about 57.5 millitesla. The transverse electric field couples the $2s$ and $2p$ levels through the Stark effect, and the magnetic field is just sufficient to bring the levels with $m_j = -1/2$ very close together in energy, while those with $m_j = +1/2$ have their energy separation doubled, so that $2s$ atoms with $m_j = +1/2$ are transmitted without loss. *See* STARK EFFECT; ZEEMAN EFFECT.

There are several methods of going on to polarize the nuclei, including a device known as the spin filter. To produce the spin filter, a longitudinal electric field of about the same strength as the transverse field is added to the apparatus that polarizes the electrons, with the longitudinal field oscillating at about 1.60 GHz. If the magnetic field is adjusted so the Larmor frequency of the electron in the metastable atom is made equal to the oscillating electric field, then the lifetime of the atom for decay becomes very long exactly at resonance, yet short not far off resonance. The magnetic field that determines the Larmor frequency of the electron in the metastable atom is the sum of that due to the solenoid and that due to the proton aligned in the solenoidal field. These two fields have opposite signs in the case of $m_I = -1/2$,

and it is found that the two resonances for transmission are at 54.0 mT for $m_I = +1/2$ and at 60.5 mT for $m_I = -1/2$. In the case of deuterons, there are three resonances, and they are well resolved even though $m_I = +1$ is at 56.5 mT, 0 is at 57.5 mT, and -1 is at 58.5 mT. Joseph L. McKibben

Colliding-beam source. A novel method of making polarized beams of negative hydrogen and deuterium is based on the direct conversion of polarized neutral hydrogen (H^0) or deuterium (D^0) atoms into polarized negative ions using the reaction below.

$$H^0 \text{ (polarized)} + Cs^0 \rightarrow H^- \text{ (polarized)} + Cs^+$$

Although conceptually very attractive, this idea presents some severe experimental difficulties. Undoubtedly the greatest of these is that if the polarized H^0 or D^0 atoms are produced at thermal energies, where production is greatly facilitated, the cross section or probability of the above reaction proceeding is extremely small. To circumvent this difficulty and to capitalize on a much higher cross section, it was proposed to accelerate a positively charged cesium (Cs) beam to an energy of about 40 keV, neutralize it by passing it through a canal containing cesium vapor, and allow this high-velocity beam of neutral cesium atoms to collide with the polarized atomic beam.

Such a source has been built and demonstrated to yield 2.9 microamperes of polarized hydrogen negative ions and 3.1 μA of deuterium ions. These currents are about five times larger than those obtainable from the best Lamb-shift sources. Roy Middleton

Bibliography. L. W. Anderson (ed.), *Polarized Ion Sources and Polarized Gas Targets*, 1994; I. G. Brown (ed.), *The Physics and Technology of Ion Sources*, 1989; A. Hershcoritz (ed.), *International Symposiumon Electron Beam Ion Sources and Their Applications*, 1989; R. Marrus (ed.), *Physics of High-Ionized Atoms*, 1989; J. Parker (ed.), *International Conference on ECR Ion Sources*, Michigan State University, C.P.47, 1987; G. Roy and P. Schmor (eds.), *Polarized Proton Ion Sources, Conference Proceedings*, TRIUMF, Vancouver, 1983.

Ion transport

Movement of salts and other electrolytes in the form of ions from place to place within living systems.

Animals

Ion transport may occur by any of several different mechanisms: electrochemical diffusion, active-transport requiring energy, or bulk flow as in the flow of blood in the circulatory system of animals or the transpiration stream in the xylem tissue of plants.

Sodium/potassium pump. The best-known system for transporting ions actively is the sodium/potassium (Na/K) exchange pump which occurs in plasma membranes of virtually all cells. The pump is especially concentrated in the cells of membranes such as frog skin, the tubules of vertebrate kidneys, salt glands in certain marine reptiles and birds, and the gills of fresh-water fishes and crustaceans, which transport sodium ion (Na^+) actively.

The physical basis of the Na/K pump has been identified as a form of the enzyme adenosine triphosphatase (ATPase) which requires both Na^+ and K^+ for its activity of catalyzing the hydrolysis of adenosine triphosphate (ATP). This reaction provides the energy required for the extrusion of Na^+ from cells in exchange for K^+ taken in, but the mechanism of coupling of the energy liberation with the transport process has not been established. The ATP which provides energy for this and other processes of cellular work is formed by synthesis coupled with cellular oxidations. Again, the precise coupling mechanism linking oxidation with synthesis of ATP has not been established heretofore.

Chemiosmotic hypothesis. The basis of the hypothesis was the concept that, if the enzyme system responsible for the transport process is embedded in the plasma membrane with an orientation such that the direction of the reaction catalyzed is parallel to the directional or vectorial orientation of the transport across the membrane, then the vectorial character of transport processes may be explained.

Proton pumps. Although this hypothesis was not generally accepted, experimental studies revealed that many transport processes, such as in bacterial cells and in the mitochondria of eukaryotic cells, are associated with a transport of protons (hydrogen ions, H^+). This fact led to the concept of proton pumps, in which the coupling or transfer of energy between oxidation processes and synthesis of ATP and between hydrolysis of ATP and transport or other cellular work is explained in terms of a flow of protons as the means of energy transfer.

Electron transport system. The processes of oxidation in the citric acid cycle of reactions in mitochondria are known to be coupled with the synthesis of ATP, which is formed from adenosine diphosphate (ADP) and inorganic orthophosphate (P_i), through the system of enzymes and cytochromes known as the electron transfer chain or electron transport system. This system transports electrons, removed in dehydrogenation from the organic molecules of the citric acid cycle on one side of the mitochondrial membrane, to the site of their incorporation into water, formed from two hydrogen ions and an atom of oxygen on the other side of the membrane (**Fig. 1**). The flow of electrons from a relatively high potential level in the organic substrate to a level of lower potential in water constitutes, in effect, a current of negative electricity, and it was proposed that the flow drives a flow of protons in the opposite direction, as a current of positive electricity. This proton flow in turn is proposed as the force that drives the synthesis of three molecules of ATP for every two electrons flowing through the electron transport system. In effect, this is the machinery of the cellular power plant.

Na/K ATPase pump. The Na/K ATPase pump (**Fig. 2**) then provides an example of a way in which a proton pump may transfer energy between the

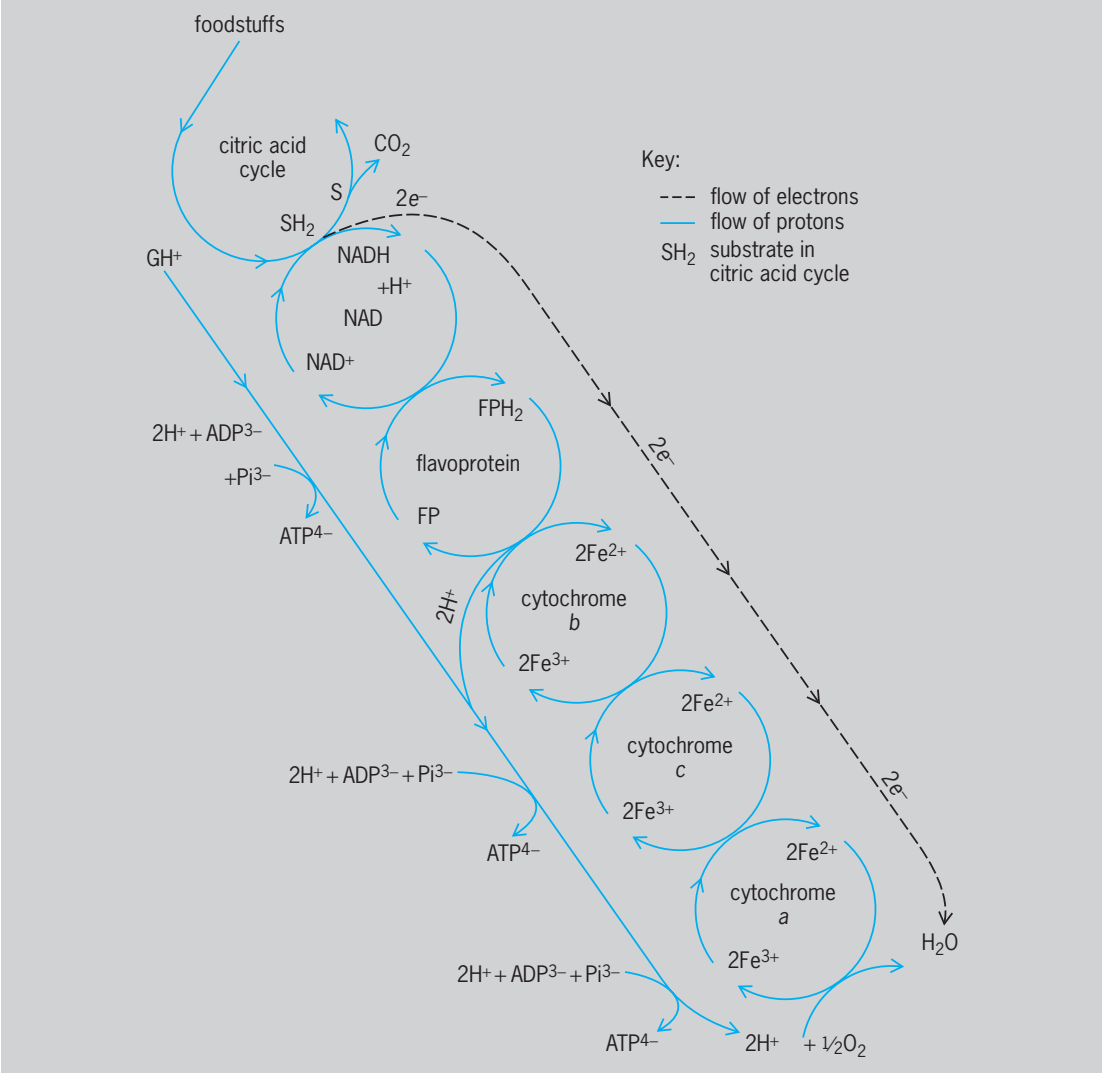

Fig. 1. Electron transport system as a proton pump.

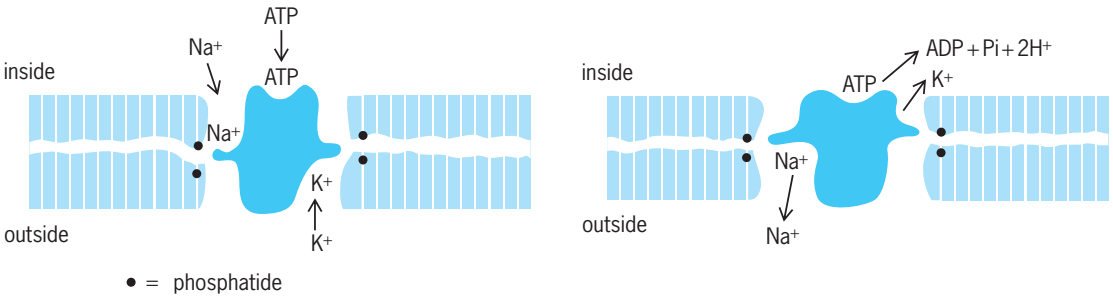

Fig. 2. Hypothetical model for the Na/K ATPase pump.

hydrolysis of ATP and a process of cellular work. The enzyme which is the basis of the pump is known to be bound to the lipid bilayer of the plasma membrane through phosphatides and to function only when so bound. The binding of Na^+, K^+, H^+, and ATP to active sites on the enzyme presumably has an allosteric effect, changing the shape of the enzyme molecule, activating the hydrolysis of ATP, and opening pathways of exchange of Na^+ and K^+. If this exchange were to be accompanied by a flow of H^+ as suggested in Fig. 2, the pump would be electrogenic, contributing to the internal negativity characteristic of nearly all cells by a net outward transport of positive electricity.

Bradley T. Scheer

Plants

Transport processes are involved in uptake and release of inorganic ions by plants and in distribution of ions within plants, and thus determine ionic relations of plants.

Scope. Photoautotrophic plants require nutrition with inorganic ions for growth and maintenance,

that is, for osmoregulation, for sustainment of structure, and for control of metabolism. Ion transport physiology is concerned with the mechanisms of movement along various pathways of transport on organizational levels differing in their degree of complexity. Furthermore, ion transport is the basic functional component of some physiological reactions in plants. In several cases, ion transport processes are the physiological basis for ecological adaptations of plants to particular environments.

Organization levels. In unicellular, filamentous, or simple thalloid algae, in mosses, in poorly differentiated aquatic higher plants, and experimentally in cell suspension cultures of higher plants, ion transport can be considered on the cellular level. In these systems, all cells take up ions directly from the external medium. The cell wall and the external lipid-protein membrane (plasmalemma) have to be passed by the ions. Intracellular distribution and compartmentation are determined by transport across other membranes within the cells. The most important one is the tonoplast separating the cell vacuole from the cytoplasm.

Within tissues the continuous cell walls of adjacent cells form an apoplastic pathway for ion transport. A symplastic pathway is constituted by the cytoplasm extending from cell to cell via small channels of about 40 nanometers diameter (plasmodesmata) crossing the cell walls. Transport over longer distances is important in organs (roots, shoots, leaves, fruits), which are composed of different kinds of tissues, and in the whole plant. Xylem and phloem serve as pathways for long-distance transport. Roots take up ions from the soil and must supply other plant organs. But there is also circulation within the plant; for example, potassium, an inorganic cation of major importance in plants, is readily transported both from root to shoot and in the opposite direction. A cooperation of roots and shoots is observed during reduction of sulfur and nitrogen. The major nutrient elements sulfur, nitrogen, and phosphorus are taken up in the form of the inorganic anions, sulfate, nitrate, and phosphate. Phosphate is used in metabolism in the oxidized form in which it is available to the plant. NO_3^- and SO_4^{2-} must be reduced. The reduction capacity of roots can be limited. The anions are transported via the xylem to the leaves, where reduction proceeds using photosynthetic energy, reduction equivalents, and carbon skeletons for the attachment of reduced nitrogen and sulfur. Reduced nitrogen and sulfur compounds can be exported via the phloem. The nutritional status of roots and shoots regarding both inorganic anions and organic substrates plays a large role in regulation of ionic relations of whole plants. Phytohormones affect transport mechanisms; they are produced in particular tissues, are distributed via the transport pathways, and thus exert a signaling function. *See* PHLOEM; PLANT HORMONES; PLANT TISSUE SYSTEMS; XYLEM.

Transport mechanisms. There are a variety of basic and special mechanisms for the different pathways of ion transport.

Apoplastic and symplastic transport. The cell wall contains amorphic hemicellulosic and pectic substances and proteins. These constituents provide fixed electrical charges for ion exchange, which are largely negative because of the carboxyl groups of pectic acids; positive charges arise from the amino groups of proteins. The cellulose micelles and fibrils of the cell wall have free spaces of up to 10 nm diameter between them, allowing transport in an aqueous phase. Thus transport in the cell wall, that is, the apoplastic space, is based on physical mechanisms and driving forces such as ion exchange, diffusion in solution, and flow of solution.

After uptake via the plasmalemma, ion transport in the symplast is also maintained by physical driving forces. But metabolic energy is required for maintenance of the transport pathway and of concentration gradients. Cytoplasmic streaming can also play a role. *See* CELL WALLS (PLANT).

Membrane transport. A major issue in ion transport is movement across lipid-protein membranes. It is governed by physical and metabolism-dependent driving forces. Diffusion of ions leads to a gradient of electrical potential and of ion concentration at the membrane.

In plant cells, several mechanisms of primary active transport use metabolic energy (ATP or pyrophosphate, PP_i) directly to pump protons or calcium ions across membranes. The proton pumps establish electrochemical gradients of protons across the membranes. The calcium pumps participate in keeping cytosolic Ca^{2+} concentrations at a low level, which is important for the role of Ca^{2+} as a secondary messenger regulating cellular functions. The proton gradients serve secondary active transport since they can drive transport of other ions and of nonelectrolyte solutes. In this way, most ion transport processes at membranes of plants can be explained. Catalyzers of transport or carriers are involved. Transport often shows saturation kinetics with increasing concentration of the transported solute.

Transport in organs. Several mechanisms, corresponding to the various pathways, are involved in transport in organs. The most important example are roots (**Fig. 3**), although similar evaluations can be made for leaves or for specialized systems such as salt glands. Ions are absorbed from the soil by ion exchange at the root surface or by diffusion. The apoplastic pathway is blocked at a cell layer somewhat inside the root (the endodermis) where the cell walls are encrusted with an amorphic hydrophobic polymer (Casparian strips of endodermal cells). Thus, ions must cross the plasmalemma and pass the cytoplasm on their way to the xylem in the interior of the roots. Depending on the relative sizes of the various driving forces and resistances, both the apoplastic and the symplastic pathways can be used up to the endodermis. Two membrane transport processes involving metabolism-dependent ion pumping are involved: uptake from the medium or apoplast into the symplast at the plasmalemma in the root tissue peripheral to the endodermis; transport from the root cells lining the xylem vessels into these vessels.

Long-distance transport in whole plant. The pipe system of the xylem in its mature transporting state is

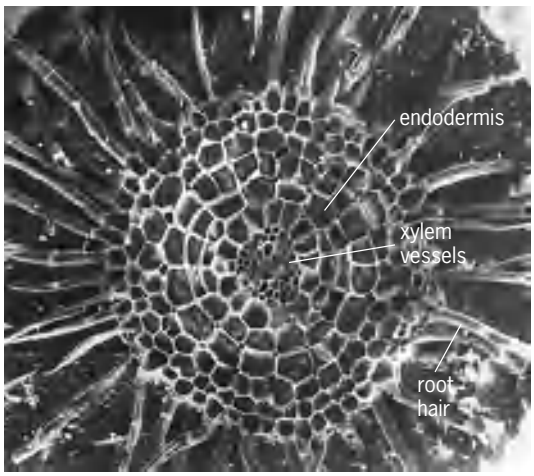

Fig. 3. Scanning electron micrograph of the cross section of a sunflower root, about 1 mm in diameter.

composed of rows of dead cells (tracheids, tracheary elements) whose cross-walls are perforated or removed entirely. The driving force for long-distance transport in the xylem is very largely passive. Transport is caused by transpiration, the loss of water from the aerial parts of the plant, driven by the water potential gradient directed from soil to roots, leaves, and atmosphere. A normally much smaller component driving the ascent of sap in the xylem is osmotic root pressure due to the pumping mechanisms concentrating ions in the root xylem, with water following passively. This can be demonstrated by attaching capillaries to the cut ends of isolated roots; it is the reason for the phenomenon of guttation, the exit of little droplets of sap from small openings in the leaves (**Fig. 4**). Root pressure amounts to only a few atmospheres, and, taking the resistance of the transport pathway into account, it can only supply shoots of small plants.

Fig. 4. Guttation droplets at the leaf dents of lady's-mantle (*Alchemilla*).

In a simplifying way the xylem can be considered as pathway for long-distance transport of ions from root to shoot, and the phloem for metabolite transport from photosynthesizing source leaves to various sinks in the plant. But inorganic ions can also move in the phloem, and this is important for recirculation in the plant. Only a limited number of ions are exempt from translocation in the phloem (such as Ca^{2+}). The long-distance transport pathways of the phloem are the sieve tubes, pipe systems with porous structures in the cross-walls (sieve plates) but, in contrast to vessels of the xylem, having living cytoplasm. Concentration and pressure gradients built up by active loading and unloading of sieve tubes in the source and sink regions, respectively, are the driving forces for transport.

Special adaptations. Ion transport mechanisms can provide the physiological basis for the adaptation to various environmental stress situations. For instance, iron-efficient dicotyledonous plants can increase the electron transport capacity at the plasmalemma of the cells on their root surfaces for reduction of Fe^{3+} to Fe^{2+} and hence for improved uptake of iron. Calcicole plants live only on calcareous soils, calcifuge plants exclusively on acid low-Ca^{2+} soils. These preferences are genetically fixed. There are also genotypic adaptations to high levels of various metal ions such as aluminum, zinc, copper, nickel, and lead. Particular transport properties of plants are of agricultural importance and thus interesting for breeders. There is also responsiveness to changes in the environment. In response to iron deficiency, some plants can thicken their root tissue behind the tips and form infoldings of the walls of peripheral root cells, increasing the plasmalemma surface for facilitation of transport. These structures disappear again as stress is released.

The most important example of stress is salinity due to NaCl. Some plants are moderately resistant by excluding NaCl at the level of the roots. The uptake into the root has K^+/Na^+ selectivity; K^+ can be taken up preferentially by exchanging for Na^+, which is released from the root tissue. Na^+ can also be reabsorbed from the xylem fluid along the pathway to the shoot, mostly in the upper root and lower shoot parts. At stronger salt stress, that is, in an environment of very low water potential, resistant plants must absorb the salt for osmoregulation. This causes problems because large Na^+ levels are not compatible with the cytoplasmic machineries of metabolism. Thus, NaCl is sequestered in the vacuoles of the cells by transport across the tonoplast. To avoid shrinkage and dehydration, and maintain the water structures of the cytoplasm, compatible organic compounds are synthesized and accumulated in the cytoplasm, which often occupies only a few percent of the total cell volume of such rather succulent plants. Compatible solutes are, for example, polyalcohols, quaternary ammonia, and sulfonium compounds. *See* PLANTS OF SALINE ENVIRONMENTS.

Special physiological mechanisms. Growing plants depend not only nutritionally on ion transport. Extension growth of young organs is driven by an

H^+-K^+ exchange mechanism. The proton extrusion pump of cells is stimulated by the phytohormone β-indole acetic acid. K^+ is taken up in exchange for H^+ to provide electrical charge balance. But, without a mechanism regulating pH, the cytoplasm would become unduly alkaline. This is controlled by CO_2 dark fixation giving malic acid via phosphoenolpyruvate and oxaloacetate. The protons of the newly formed carboxyl groups are extruded; the organic anion malate together with K^+ is transported into the vacuoles, acting as osmotically active material building up turgor pressure, which drives cell extension.

In this way, hydrogen ion pumps quite generally can be parts of pH-regulating mechanisms and turgor-dependent processes. Other examples are movements of stomatal guard cells and of pulvini bringing about nyctinastic and seismonastic changes of leaf positions. Turgor changes in the guard cells can operate via H^+-K^+ exchange and malic acid synthesis; increased turgor leads to opening of, and decreased turgor to closing of, stomata.

Movements and other reactions in plants are accompanied by action potentials. The ionic basis for plant action potentials has been investigated in some algae (Characeae, *Acetabularia*). Depolarization and repolarization of the membranes during action potentials are due to efflux of Cl^- followed by K^+. In other cases, Ca^{2+} is also involved. *See* PLANT OF TRANSPORT SOLUTES; PLANT-WATER RELATIONS.

Ulrich Lüttge

Bibliography. F. M. Harold, The 1978 Nobel Prize in chemistry, *Science*, 202(4373), December 15, 1978; N. Higinbotham, Electropotential of plant cells, *Annu. Rev. Plant Physiol.*, 24:25–46, 1973; U. Lüttge and N. Higinbotham, *Transport in Plants*, 1979; U. Lüttge and M. G. Pitman (eds.), Transport in plants II, part A: Cells, part B: Organs and tissues, *Encyclopedia of Plant Physiology*, *New Series*, 1976; M. G. Pitman, Ion transport in the xylem, *Annu. Rev. Plant Physiol.*, 28:71–88, 1977; L. Reinhold and A. Kaplan, Membrane transport of sugars and amino acids, *Annu. Rev. Plant Physiol.*, 35:45–83, 1984.

Ionic crystals

A class of crystals in which the lattice-site occupants are charged ions held together primarily by their electrostatic interaction. Such binding is called ionic binding. Empirically, ionic crystals are distinguished by strong absorption of infrared radiation, good ionic conductivity at high temperatures, and the existence of planes along which the crystals cleave easily. *See* CRYSTAL STRUCTURE.

Compounds of strongly electropositive and strongly electronegative elements form solids which are ionic crystals, for example, the alkali halides, other monovalent metal halides, and the alkaline-earth halides, oxides, and sulfides. Crystals in which some of the ions are complex, such as metal carbonates, metal nitrates, and ammonium salts, may also be classed as ionic crystals.

As a crystal type, ionic crystals are to be distinguished from other types such as molecular crystals, valence crystals, or metals. The ideal ionic crystal as defined is approached most closely by the alkali halides. Other crystals often classed as ionic have binding which is not exclusively ionic but includes a certain admixture of covalent binding. Thus the term ionic crystal refers to an idealization to which real crystals correspond to a greater or lesser degree, and crystals exist having characteristics of more than one crystal type. *See* CHEMICAL BONDING.

Ionic crystals, especially alkali halides, have played a very prominent role in the development of solid-state physics. They are relatively easy to produce as large, quite pure, single crystals suitable for accurate and reproducible experimental investigations. In addition, they are relatively easy to subject to theoretical treatment since they have simple structures and are bound by the well-understood Coulomb force between the ions. This is in contrast to metals and covalent crystals, which are bound by more complicated forces, and to molecular crystals, which either have complicated structures or are difficult to produce as single crystals. Being readily available and among the simplest known solids, they have thus been a frequent and profitable meeting place between theory and experiment. These same features of ionic crystals have made them attractive as host crystals for the study of crystal defects: deliberately introduced impurities, vacancies, interstitials, and color centers. *See* COLOR CENTERS; CRYSTAL DEFECTS.

Crystal structure. The simplest ionic crystal structures are those of the alkali halides. At standard

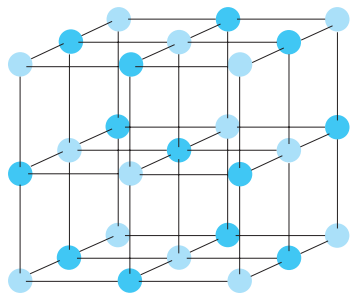

Fig. 1. Sodium chloride lattice. The darker circles represent positive ions and the lighter circles negative ions. (*After F. Seitz, The Modern Theory of Solids, Dover, 1987*)

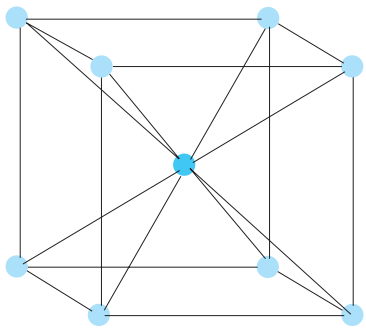

Fig. 2. Cesium chloride lattice. (*After F. Seitz, The Modern Theory of Solids, Dover, 1987*)

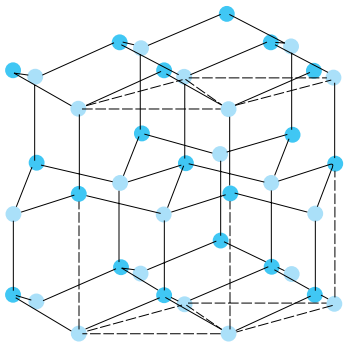

Fig. 3. Wurtzite lattice. (*After F. Seitz, The Modern Theory of Solids, Dover, 1987*)

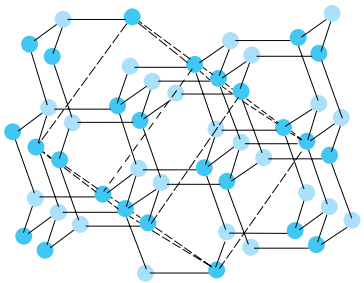

Fig. 4. Zinc blende lattice. (*After F. Seitz, The Modern Theory of Solids, Dover, 1987*)

temperature and pressure the 16 salts of lithium (Li), sodium (Na), potassium (K), and rubidium (Rb) with fluorine (F), chlorine (Cl), bromine (Br), and iodine (I), have the sodium chloride structure of interpenetrating face-centered cubic lattices (**Fig. 1**). Cesium fluoride (CsF) also has this structure but otherwise the cesium halides have the cesium chloride structure of interpenetrating simple cubic lattices (**Fig. 2**). The sodium chloride structure is also assumed by the alkaline-earth oxides, sulfides, and selenides other than those of beryllium (Be) and by silver fluoride (AgF), silver chloride (AgCl), and silver bromide (AgBr). Other crystal structures, such as the wurtzite structure (**Fig. 3**) assumed by beryllium oxide (BeO), β-zinc sulfide (β-ZnS, also known as wurtzite), and zinc oxide (ZnO) and the zinc blende structure (**Fig. 4**) assumed by copper(I) chloride (CuCl), copper(I) bromide (CuBr), copper(I) iodide (CuI), beryllium sulfide (BeS), and α-zinc sulfide (α-ZnS, also known as sphalerite or zinc blende) are also typical of the ionic crystals of salts in which the atoms have equal positive and negative valence. Ionic compounds consisting of monovalent with divalent elements crystallize typically in the fluorite structure (**Fig. 5**) assumed by calcium fluoride (CaF_2), barium fluoride (BaF_2), cadmium fluoride (CdF_2), lithium oxide (Li_2O), lithium sulfide (I_2S), sodium monosulfide (Na_2S), copper(I) sulfide (Cu_2S), and copper(I) selenide or the rutile structure (**Fig. 6**) assumed by titanium(IV) oxide (TiO_2, also known as rutile), zinc fluoride (ZnF), and magnesium fluoride (MgF_2).

Cohesive energy. It is possible to understand many of the properties of ionic crystals on the basis of a simple model originally proposed by M. Born and E. Madelung. In the simplest form of this model, the lattice sites are occupied by spherically symmetric ions having charges corresponding to their normal chemical valence. These ions overlap only slightly the ions at neighboring sites and interact with one another through central forces. In sodium chloride (NaCl), for example, the spherically symmetric closed shell configurations which the free Na^+ and Cl^- ions possess are considered to be negligibly altered by the crystalline environment and to have charges $+e$ and $-e$, respectively, where $-e$ is the charge on the electron. Using this model, together with certain assumptions about the forces between the ions, Born and M. Göppert-Mayer calculated the cohesive energy of a number of ionic crystals. This cohesive energy is defined as the energy necessary to take an ionic crystal from an initial stage, in which the crystal is at 0 K ($-459.67°$F) and zero pressure. While it cannot be measured directly, this cohesive energy can be deduced from experimental quantities by the use of the Born-Haber cycle. Thus the validity of the simple model of Born and Madelung can be tested by comparing the calculated cohesive energy of Born and Mayer with values which have been experimentally determined. *See* COHESION (PHYSICS).

Born-Haber cycle. This is a sequence of processes leading from the initial to the final state specified in the definition of the cohesive energy. Because in most of the processes in this cycle heat changes at constant pressure are measured, it is convenient to consider the change in heat content or enthalpy

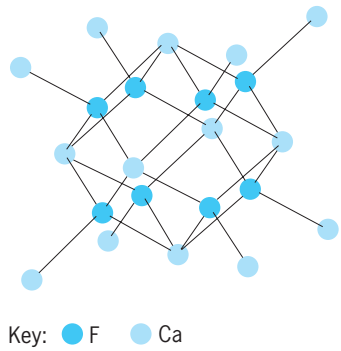

Fig. 5. Calcium fluoride lattice. (*After F. Seitz, The Modern Theory of Solids, Dover, 1987*)

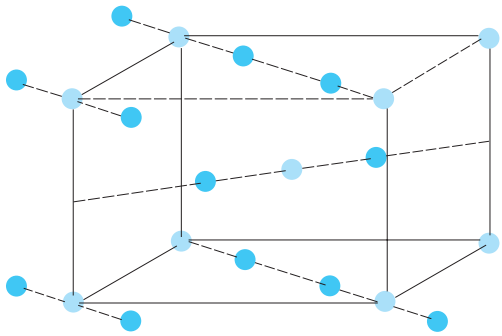

Fig. 6. Rutile lattice. (*After F. Seitz, The Modern Theory of Solids, Dover, 1987*)

$H = U + PV$, where P is the pressure, V is the volume, and U is the cohesive energy, rather than the change in E in each step. Since the total change in H is independent of the intermediate steps taken to accomplish it, the ΔH for the change in state specified in the definition of cohesive energy will be given by the sum of the ΔH in the steps of the cycle. Furthermore, because when $P = 0$, $\Delta H = \Delta U$, the ΔH thus calculated will be the cohesive energy. In the following enumeration of the steps in the Born-Haber cycle $[A]$ indicates element A in a monatomic gaseous state and $[A]$ indicates A in a solid state, and so on. The B without brackets in step 3 refers to the natural form of B at the given temperature and pressure. The ionic compound is taken to be AB, where A is the electropositive and B the electronegative element.

The steps of the Born-Haber cycle (all temperatures in K and pressures in atmospheres) are:

1. $[AB]^{0^\circ}_{P=1} \rightarrow [AB]^{0^\circ}_{P=1}$. The value of ΔH_1 in this isothermal compression is very small and can be neglected in comparison with other heat content changes in the cycle.

2. $[AB]^{0^\circ}_{P=1} \rightarrow [AB]^{298^\circ}_{P=1}$. In this step the crystal is warmed to room temperature. The value of ΔH_2 can be calculated from the specific heat at constant pressure for the crystal.

3. $[AB]^{0^\circ}_{P=1} \rightarrow [A]^{298^\circ}_{P=1} + B^{298^\circ}_{P=1}$. The value of ΔH_3 is given by the heat of formation of the compound AB which is referred to substances in their natural forms at standard temperature and pressure.

4. $B^{298^\circ}_{P=1} \rightarrow (B)^{298^\circ}_{P=1}$. The value of ΔH_4 is the dissociation energy necessary to form a monatomic gas from B in its natural state at standard temperature and pressure. For chlorides, for example, this is the dissociation energy of a Cl_2 molecule into Cl atoms.

5. $[A]^{298^\circ}_{P=1} \rightarrow (A)^{298^\circ}_{P=1}$. In this step, ΔH_5 is the heat of sublimation of the metal A. It can be deduced from the heat of fusion, the specific heats of the solid, liquid, and gaseous phases, and the vapor pressure data for the metal.

6. $(A)^{298^\circ}_{P=1} \rightarrow (A)^{0^\circ}_{P=0}$; $(B)^{298^\circ}_{P=1} \rightarrow (B)^{0^\circ}_{P=0}$. An adiabatic expansion of the gases, considered as ideal, to a very large volume results in a state in which $P = 0$, $T = 0$, and $\Delta H_6 = -{}^5\!/_2\, RT$/mole, where R is the gas constant.

7. $(A)^{0^\circ}_{P=0} \rightarrow (A^+)^{0^\circ}_{P=0} + e^-$. The ionization of the A atoms gives a ΔH_7 per atom equal to their first ionization energy.

8. $(B)^{0^\circ}_{P=0} + e^- \rightarrow (B^-)^{0^\circ}_{P=0}$. The electrons from step (7) are placed on the B atoms. The value of ΔH_8 per atom is given by the electron affinity of the B atom. *See* ELECTRONEGATIVITY; SUBLIMATION.

As an example, for sodium chloride $\Delta H_1 \cong 10^{-4}$ (kilocalories/mole), $\Delta H_2 = 2.4$, $\Delta H_3 = 98.3$, $\Delta H_4 = 26.0$, $\Delta H_5 = 28.8$, $\Delta H_6 = -2.9$, $\Delta H_7 = 11.9$, and $\Delta H_8 = -80.5$. Experimental cohesive energies for a number of the ionic crystals are given in the **table**. *See* STRUCTURAL CHEMISTRY.

Cohesive energies*

Crystal	Structure	U_{exp}, kcal/mole†	U_{calc}, kcal/mole†	U_{calc} refined, kcal/mole†
LiCl	NaCl	201.5	196.3	200.2
LiBr	NaCl	191.5	184.4	189.5
LiI	NaCl	180.0	169.1	176.1
NaCl	NaCl	184.7	182.0	183.5
NaBr	NaCl	175.9	172.7	175.5
NaI	NaCl	166.3	159.3	164.3
KCl	NaCl	167.8	165.7	167.9
KBr	NaCl	161.2	158.3	161.3
KI	NaCl	152.8	148.2	152.4
RbCl	NaCl	163.6	159.1	162.0
RbBr	NaCl	159.0	151.9	156.1
RbI	NaCl	149.7	143.1	148.0
CaF_2	Fluorite	618.0	617.7	
CuCl	Zinc blende	226.3	206.1	
ZnS	Wurtzite	851	816	
PbO_2	Rutile	2831	2620	
AgCl	NaCl	207.3	187.3	

*The cohesive energies in the last two columns are calculated using the Born-Mayer equation and the refined Born-Mayer theory, respectively. The refined calculations for the last five crystals have not been made.
†1 kcal = 4.184 kJ.

Born-Mayer equation. By use of the Born-Madelung model, the cohesive energy of an ionic crystal can be related to its measured compressibility and lattice spacing. Because of the opposite signs of electric charges which they carry, the unlike ions in such a crystal model attract one another according to Coulomb's law. However, such a charge distribution cannot be in equilibrium if only Coulomb forces act. In addition to their Coulomb interaction, the ions exhibit a repulsion which, compared with the Coulomb interaction, varies rapidly with interionic separation. The repulsion becomes strong for small separations and diminishes rapidly for increasing separation. The static equilibrium configuration of the crystal is determined by a balance of these forces of attraction and repulsion. *See* COULOMB'S LAW.

The short-range repulsion between ions must be described by quantum mechanics. When the electron orbits of two ions overlap, the electron charge density in the region of overlap is diminished as a consequence of the Pauli exclusion principle. This charge redistribution results in a repulsion between the ions in addition to the Coulomb interaction, which they have at all interionic distances. In early work, the energy V_{rep} due to repulsion of two ions at a distance r was assumed to have the form of Eq. (1), where B and n are constants to be determined. Quantum-mechanical calculations of the interaction of atoms with closed shells of electrons indicate that the interaction of repulsion is better approximated by an exponential dependence on interionic distance given in Eq. (2), where A and ρ are constants. Both forms for V_{rep} give almost the same calculated cohesive energy; the exponential form gives slightly better agreement with experiment. *See* EXCLUSION PRINCIPLE.

$$V_{rep} = \frac{B}{r^n} \tag{1}$$

$$V_{rep} = Ae^{-r/\rho} \tag{2}$$

Using the exponential form for the repulsive interaction energy, the potential energy $\varphi(r_{ij})$ of a pair

of ions i and j can be written as Eq. (3), where $Z_i e$

$$\varphi(r_{ij}) = \frac{Z_i Z_j e^2}{r_{ij}} + Ae^{-r_{ij}/\rho} \tag{3}$$

and $Z_j e$ are the net charges of the ions i and j, and r_{ij} is the distance of separation of their centers. The assumption that the ions are spherically symmetric has been used here in writing the Coulomb interaction as that of point charges.

The cohesive energy U of an ionic crystal due to the Coulomb and repulsive interactions of its ions is the sum taken over all pairs of ions in the crystal as in Eq. (4), where in the summation the lattice site

$$U = \frac{1}{2} \sum_{i,j}{}' \varphi(r_{ij}) \tag{4}$$

indices, i and j, range over all sites of the crystal. The prime on the summation sign indicates the exclusion from the sum of terms for which $i = j$ and the factor of $^1/_2$ avoids counting pairs of ions twice.

For crystals in which there are only two types of ion, the Coulomb or electrostatic part of U, U_e, can be written in a simple form given in Eq. (5), where

$$U_e = \frac{1}{2} \sum_{ij}{}' \frac{Z_i Z_j e^2}{r_{ij}} = -\frac{N\alpha_M (Z_+ Z_-) e^2}{r} \tag{5}$$

$+Z_+e$ and $-Z_-e$ are the charges of the positive and negative ions, N is the number of ion pairs in the crystal, r is the nearest anion-cation separation, and α_M is the Madelung constant. *See* MADELUNG CONSTANT.

By anticipating that ρ will be small compared to the nearest neighbor separation, the interactions of repulsion may be neglected for pairs of ions other than nearest neighbors. The energy of the crystal model for arbitrary nearest neighbor separation r is then given by Eq. (6), where M is the number of

$$U(r) = N\frac{-\alpha_M Z_+ Z_- e^2}{r} + MAe^{-r/\rho} \tag{6}$$

nearest neighbors which each ion has in the crystal.

The parameter ρ may be evaluated for a given crystal by requiring that (1) U be a minimum for the observed value of r and (2) that the compressibility of the model equal the measured compressibility of the crystal. It follows from these requirements that Eq. (7) holds. This is the Born-Mayer equation for

$$U(r_0) = \frac{N\alpha_M Z_+ Z_- (1 - \rho/r_0) e^2}{r_0} \tag{7}$$

the cohesive energy, where r_0 refers to the nearest neighbor distance at static equilibrium. Further, in this equation, ρ is given in terms of experimental quantities by Eq. (8), where K is the measured compressibility of the crystal.

$$\frac{r_0}{\rho} = \frac{18 r_0^4}{\alpha_M e^2 K} + 2 \tag{8}$$

Cohesive energies for some alkali halides and crystals of other structures calculated in this way are shown in the table, where they can be compared with the experimental values for the cohesive energy. The agreement is considered to be support for the essential validity of the Born-Madelung model. The model has been applied with some success even to the ammonium halides, assuming spherically symmetric ions.

The Born-Mayer theory has been refined, with resulting improvement in the agreement between the calculated and experimental cohesive energies for alkali halides. The refinements have considered the small (a few kilocalories per mole or less) corrections to the cohesive energy arising from van der Waals interactions and zero-point vibrational energy. The van der Waals forces are weak attractive forces between ions due to mutually induced fluctuating dipoles. Similar forces, even weaker, due to dipole-quadrupole interactions have also been considered. Both these interactions make small positive contributions to the cohesive energy. At 0 K (−459.67°F) the lattice is not in static equilibrium but as a consequence of quantum mechanics is in a state of zero-point vibration with nonzero energy. The energy of these vibrational modes cannot be further reduced. The zero-point vibration energy gives a small negative contribution to the cohesive energy. The results of these refinements of the Born-Mayer theory are also shown in the table. *See* LATTICE VIBRATIONS.

While it has had success in enabling the cohesive energy to be calculated, the shortcomings of this simple model become evident in its failure to predict correctly the elastic shear constants of ionic crystals. This requires interionic forces of a noncentral character, which are absent from the model. There are also other instances in which the simple model is found to be inadequate.

More elaborate models which take into account features absent from the Born-Mayer model and which may be regarded as extensions of it have had considerable success in accounting for the elastic and dielectric properties of alkali halides. More important has been the ability of these models to account for the lattice phonons in these crystals. This has given the alkali halides a prominent place in the study of phonons in insulators, where they are among the few reasonably well-understood solids.

B. Gale Dick

Ionic conductivity. Most ionic crystals have large band gaps, and are therefore generally good electronic insulators. However, electrical conduction occurs by the motion of ions through these crystals. The presence of point defects, that is, deviations from ideal order in the crystalline lattice, facilitates this motion, thus giving rise to transport of electric charge. In an otherwise perfect lattice where all lattice sites are fully occupied, ions cannot be mobile.

Mechanism. The two most common types of point defects that give rise to conductivity in ionic crystals are vacancies and interstitials. These intrinsic defects are in thermodynamic equilibrium, and their concentrations depend exponentially upon the temperature. The most common mechanisms by which

ions are conducted are vacancy diffusion and interstitial diffusion. In the former, a mobile ion positioned close to a vacancy may move into it. In the latter, the interstitial ion can change sites by hopping. Ionic crystals with Frenkel disorders (cation vacancy-cation interstitial pairs), such as silver halides, where the predominant defects are silver interstitials, have substantially greater values of ionic conductivity than those with Schottky disorders (anion-cation vacancy pairs), such as alkali halides and alkaline-earth halides, where the predominant defects are vacancies.

Vacancies and interstitials can also be generated extrinsically by doping the ionic crystal with a suitable amount of an aliovalent (that is, of different valence from the host ion) species. This leads to the formation of vacancies or interstitial defects to counterbalance the excess charge brought in by the aliovalent ion since the crystal must preserve its charge neutrality.

Superionic conductors. Many so-called normal ionic crystals possess conductivities of about 10^{-10} (ohm·cm)$^{-1}$ or lower at room temperature. However, a relatively small number of ionic materials, called superionic conductors or fast ionic conductors, display conductivities of the order of 10^{-1} to 10^{-2} (ohm · cm)$^{-1}$, which imply ionic liquidlike behavior. In most of these crystals, only one kind of ionic species is mobile, and its diffusion coefficient and mobility attain values such as found otherwise only in liquids. Due to their high value of ionic conductivity as well as their ability to selectively transport ionic species, superionic conductors have successfully been employed as solid electrolytes in many applications. *See* DIFFUSION; SOLID-STATE BATTERY.

The high value of conductivity and diffusion coefficient for the mobile ionic species is due to the high degree of structural disorder in these materials. This disorder is so great that the mobile ions assume a quasimolten state, swimming in the open spaces in a rigid framework of the immobile ions. For example, the high-temperature α-phase of silver iodide (α-AgI), stable above 147°C (297°F), attains unusually high values of ionic conductivity, in excess of 1 (ohm · cm)$^{-1}$. The iodide ions occupy the body-centered cubic (bcc) sites and form the rigid framework. The silver ions are distributed among sites whose numbers are much greater than the number of silver ions. Similarly, $RbAg_4I_5$ has 56 tetrahedral interstitial positions for the 16 silver ions present in the structure to occupy. [This material possesses the highest ionic conductivity yet reported for a solid at room temperature, 0.3 (ohm · cm)$^{-1}$ at 300 K or 80°F.]

The preexponential factor in the Arrhenius expression for the temperature dependence of the conductivity is appreciably smaller in superionic conductors than that exhibited by normal ionic materials. The activation enthalpy for conduction is generally small in superionic conductors, but can range over a wide spectrum. For example, doped zirconium dioxide (ZrO_2), which transports oxide ions by vacancy diffusion and is used at elevated temperatures between 600 and 1400°C (1100 and 2500°F), possesses a high activation enthalpy for ionic conduction. The oxygen vacancies in the lattice are produced by dissolution of dopants such as calcium oxide (CaO), magnesium oxide (MgO), or yttrium oxide (Y_2O_3) in zirconium dioxide. Since the Ca^{2+} ion substitutes for the Zr^{4+} site, one oxygen vacancy is produced for every Ca^{2+} ion. The typical amount of doping in this material is of the order of 10 mol %. This results in a very high concentration of oxygen vacancies of about 10^{22} cm^{-3}, comparable to the electron density in metals. *See* ENTHALPY.

Parameters and measurements. If charge and mass are transported by the same mobile species, then independent measurements of the ionic conductivity and the diffusion coefficient allow calculation of the concentration as well as the electrical mobility of that species. Even the most mobile ions in superionic conductors have mobilities that are orders of magnitude lower than that of electrons in metals. Hence the ionic conductivity of superionic conductors is still much below that of metals.

An important consideration is the extent to which a mobile ionic species contributes to the total conduction process. The transference number t_i for a species i is defined as the ratio of the partial (or specific) conductivity due to species i to the total conductivity. It can be determined by measuring the open-circuit potential across the ionic crystal in an electrochemical cell arrangement, such that a known difference in the chemical potential of species i is imposed on each side of the crystal at a suitable temperature. The ratio of the measured open-circuit potential to the theoretical potential calculated from the known chemical potential difference gives the transference number. Another method is to use an electrochemical titration technique with suitable slab electrodes that contain the mobile species. With passage of current, the mobile species i is tranported across the ionic material. The weight change in the electrode material due to deposition or depletion of i is compared with the expected value calculated from Faraday's law. However, in all these measurements the interface between the electrode and the ionic material is very important. *See* CHEMICAL THERMODYNAMICS; ELECTROCHEMICAL TECHNIQUES; ELECTROLYSIS.

In order to measure the ionic conductivity or the transference number of the mobile species i in an ionic material, an electrode material that contains species i is required to obtain an interface with the ionic material that is completely nonblocking (sometimes called reversible) to species i but completely blocking (that is, nonreversible) to all other species. By using suitable electrode materials the conductivities and the transference numbers of the minority carriers as well as the mobile species can be obtained.

The easiest way to measure ionic conductivity is a two-probe dc technique with an electrochemical cell arrangement that has nonblocking interfaces for the species i on each side of the ionic material. In

practice, there may be problems with slow electrode reactions that must be isolated and accounted for. A better way is the four-probe dc technique using four nonblocking electrodes, two as voltage probes and the other pair to pass a known and constant current.

If a suitable nonblocking electrode material is not available, ac techniques may be used. A small-amplitude sinusoidal signal at reasonably high frequencies can be applied to the ionic material in an electrochemical cell arrangement, and the conductivity value extracted from the real part of the measured cell impedance. However, if the frequency used in the measurements is not high enough, the conductivity value may be in error. A more involved technique to eliminate this problem is ac impedance spectroscopy. Here a wide range of frequencies covering many decades is imposed upon the ionic material, and the frequency dispersion is analyzed in the complex impedance plane. Experience and insight are required for meaningful analyses of the frequency spectrum. *See* ALTERNATING-CURRENT CIRCUIT THEORY. Turgut M. Gür

Dielectric properties. The polarization of an ionic crystal in an applied time-varying electric field depends on the frequency v of the field. If $v = 0$ the field is static, and a static equilibrium will be achieved in which the ions are displaced from their lattice sites a distance determined by a balance between the force of repulsion due to their nearest neighbors and the electrical force on the ions. Oppositely charged ions move in opposite directions with a resulting ionic polarization. In addition, the electron clouds of the ions are deformed in the local electric field at the ions and, to a much lesser extent, by the repulsive forces, to give an electronic polarization. As the frequency v increases, inertial effects must be considered. When v is sufficiently high, that is, when $v > v_0$ (v_0 is the natural frequency of vibration of ions about lattice sites), the heavy ions can no longer follow the rapid variations of the applied field. The ionic polarization ceases to contribute to the total polarization, although the less massive electron clouds continue to give an electronic polarization. Eventually, with increasing frequency, v exceeds the frequency at which the electron cloud deformations can follow the field, and the electronic polarization becomes altered. The frequency v_0, called the restrahlung or residual-ray frequency, lies in the infrared region, and the associated vibrations of the charged ions are responsible for the strong absorption of infrared radiation in ionic crystals. The term residual-ray frequency comes from the fact that the reflectivity of visible and infrared radiation is highest for radiation of a frequency near to that of maximum absorption, v_0. Thus multiple reflection will yield radiation composed predominantly of residual rays of frequency near to v_0. For additional information on residual rays *see* REFLECTION OF ELECTROMAGNETIC RADIATION

At frequencies below v_0 and above it up to the ultraviolet, an ideal alkali halide is transparent to electromagnetic radiation. Impurities or defects often introduce absorption bands into these extensive and otherwise transparent regions and can thus be studied optically. This fact accounts for much of the wealth of experimental detail available on defects and impurities in alkali halides and for the interest which they have attracted.

Ionic radii. It has been found that ionic radii can be chosen such that the lattice spacings of ionic crystals are given approximately by the sum of the radii of the constituent ions. Thus, for instance, eight radii suffice to give the approximate lattice spacings of the 16 alkali halides having sodium chloride structure. This is true because of the rapid variation of the repulsive interaction with distance, so that the ions interact somewhat like hard spheres. This implies that a given ion, having the same ionic radius in a variety of crystals, has an electron cloud that is only slightly altered by the differing crystalline environments of these crystals. B. Gale Dick

Bibliography. N. W. Ashcroft and W. D. Mermin, *Solid State Physics*, 1976; H. Boettger et al., *Hopping Conduction in Solids,* 1986; S. Chandra, *Superionic Solids: Principles and Applications,* 1981; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; L. C. Pauling, *The Nature of the Chemical Bond*, 3d ed., 1960; F. Seitz et al., *Solid State Physics*, 40 vols. and 15 suppls., 1955–1987; A. Stoneham (ed.), *Ionic Solids at High Temperatures*, 1989; T. Takahashi (ed.), *High Conductivity in Solid Ionic Conductors: Recent Trends and Applications*, 1989.

Ionic equilibrium

An equilibrium in a chemical reaction in which at least one ionic species is produced, consumed, or changed from one medium to another.

Types of equilibrium. A few examples can illustrate the wide variety of types of ionic equilibrium which are known.

Dissolution of an un-ionized substance. The dissolution of hydrogen chloride (a gas) in water (an ionizing solvent) can be used to illustrate this type. Reactions (1), (2), and (3) all represent exactly the same equi-

$$HCl(g) \rightleftharpoons H^+ + Cl^- \quad (1)$$

$$HCl(g) + H_2O \rightleftharpoons H_3O^+ + Cl^- \quad (2)$$

$$HCl(g) + 4H_2O \rightleftharpoons H_9O_4{}^+ + Cl^- \quad (3)$$

librium. Equation (1) ignores the hydration of the proton and is preferred for many purposes when the hydration (or solvation) of the proton is irrelevant to a particular discussion. Equation (2) is written in recognition of the widely held belief that free protons do not exist in aqueous solution. Equation (3) indicates that another three molecules of water are very firmly bound to the H_3O^+ ion (the hydronium ion). There is no implication in Eq. (3), however, that the total number of molecules of water attached to, or weakly affected by, the hydronium ion may not be considerably larger than three.

Not much is known about the solvation of ions, although it has been proved that each chromic ion, Cr^{3+}, in dilute aqueous solution holds at least six water molecules. Only a few other similar data have been clearly established. Hence equations for the other examples cited below are written without regard to solvation, except that the hydrogen ion is usually written in accordance with common practice as H_3O^+. *See* ACID AND BASE.

Dissolution of a crystal in water. The dissociation of solid silver chloride, reaction (4), illustrates this type of

$$AgCl(crystal) \rightleftharpoons Ag^+ + Cl^- \qquad (4)$$

equilibrium. *See* SOLUBILITY PRODUCT CONSTANT.

Dissociation of a strong acid. Nitric acid, HNO_3, dissociates as it dissolves in water, as in reaction (5). At

$$HNO_3 + H_2O \rightleftharpoons H_3O^+ + NO_3^- \qquad (5)$$

77°F (25°C) about one-half the acid is dissociated in a solution containing 10 (stoichiometric) moles of nitric acid per liter.

Dissociation of an ion in water. The bisulfate ion, HSO_4^-, dissociates in water, as in reaction (6). About one-

$$HSO_4^- + H_2O \rightleftharpoons H_3O^+ + SO_4^{2-} \qquad (6)$$

half the HSO_4^- is dissociated in an aqueous solution containing about 0.011 mole of sulfuric acid per liter at 77°F (25°C).

Dissociation of water itself. In pure water at 77°F (25°C) the concentration of each ion is about 10^{-7} mole/liter, but increases rapidly as temperature is increased. This equilibrium is represented by reaction (7).

$$2H_2O \rightleftharpoons H_3O^+ + OH^- \qquad (7)$$

Formation of a complex ion. In water or in a mixture of fused (chloride) salts, complex ions, such as $ZnCl_4^{2-}$, may be formed, as in reaction (8).

$$Zn^{2+} + 4Cl^- \rightleftharpoons ZnCl_4^{2-} \qquad (8)$$

See COORDINATION COMPLEXES.

Dissociation of a weak acid. In water acetic acid dissociates to form hydrogen (hydronium) ion and acetate ion, as in reaction (9).

$$CH_3CO_2H + H_2O \rightleftharpoons H_3O^+ + CH_3CO_2^- \qquad (9)$$

Electrochemical reaction. Reaction (10) takes place "al-

$$\tfrac{1}{2}H_2(g) + AgCl(s) + H_2O(l) \rightleftharpoons H_3O^+ + Cl^- + Ag(s) \qquad (10)$$

most reversibly" when the equilibrium shown exists. A small current is allowed to flow through an electric cell consisting of an aqueous solution of HCl saturated with silver chloride, a hydrogen electrode, and a silver electrode. Saturation is maintained by an excess of solid silver chloride which for convenience is sometimes mixed with the silver or plated, as a coating, on the metal. The electrode is then called a silver–silver chloride electrode.

Many additional types of equilibria could be mentioned, including those reactions occurring entirely in the gaseous phase and those reactions occurring between substances dissolved in two immiscible liquids.

Quantitative relationships. Each reaction obeys an equilibrium equation of the type shown as Eq. (11).

$$\frac{[H^+][Cl^-]}{[HCl(g)]}\frac{f_+f_-}{\gamma_g} = Q_cQ_f = K \qquad (11)$$

The activity coefficient γ_g can be ignored here because it is very nearly unity. The terms f_+ and f_- are the respective activity coefficients of H^+ and Cl^- but cannot be determined separately. Their product can be determined experimentally and can also be calculated theoretically for very dilute solutions by means of the Debye-Hückel theory of interionic attraction. Because γ_g and f_+f_- are nearly unity, Eq. (11) demands that the pressure of HCl gas above a dilute aqueous solution be proportional to the square of the concentration of the solute. Q_c is called the concentration quotient and Q_f the quotient of activity coefficients.

Similarly, the dissociation of acetic acid obeys Eq. (12), where f_u is the activity coefficient of the

$$\frac{[H^+][CH_3CO_2^-]}{[CH_3CO_2H]}\frac{f_+f_-}{f_u} = Q_cQ_f = K \qquad (12)$$

unionized acetic acid. Early work on electrolytes revealed that Q_c, the concentration quotient, was constant within the limits of accuracy attainable at the time. Later work revealed that the measured concentration quotient Q_c first increases as concentration is increased from very small values and then decreases sharply. The initial increase is due largely to the electrical forces between the ions which reduce the product f_+f_-. There is some evidence that the subsequent decrease in Q_c and the concomitant rise in Q_f are due to removal of some of the monomeric acetic acid by the formation of dimeric acetic acid, $(CH_3CO_2H)_2$. The fact is, however, that knowledge concerning activity coefficients in solutions other than very dilute ones is not yet understood. Even the experimental methods for the measurement of the molecular species involved in some equilibria were not evolved until recently. *See* ELECTROLYTIC CONDUCTANCE; HYDROLYSIS.

Thomas F. Young

Bibliography. G. M. Barrow, *Physical Chemistry*, 6th ed., 1996; H. Russotti, *The Study of Ionic Equilibria: An Introduction*, 1978; G. H. Schenk and D. D. Ebbing, *Qualitative Analysis and Ionic Equilibrium*, 2d ed., 1990.

Ionic liquids

In the simplest sense, the liquid phase that forms on the melting of an ionic compound (that is, a salt) and consists entirely of discrete ions (**Fig. 1**). More specifically, ionic liquids are the liquid phase of organic salts (organic, complex cations) which

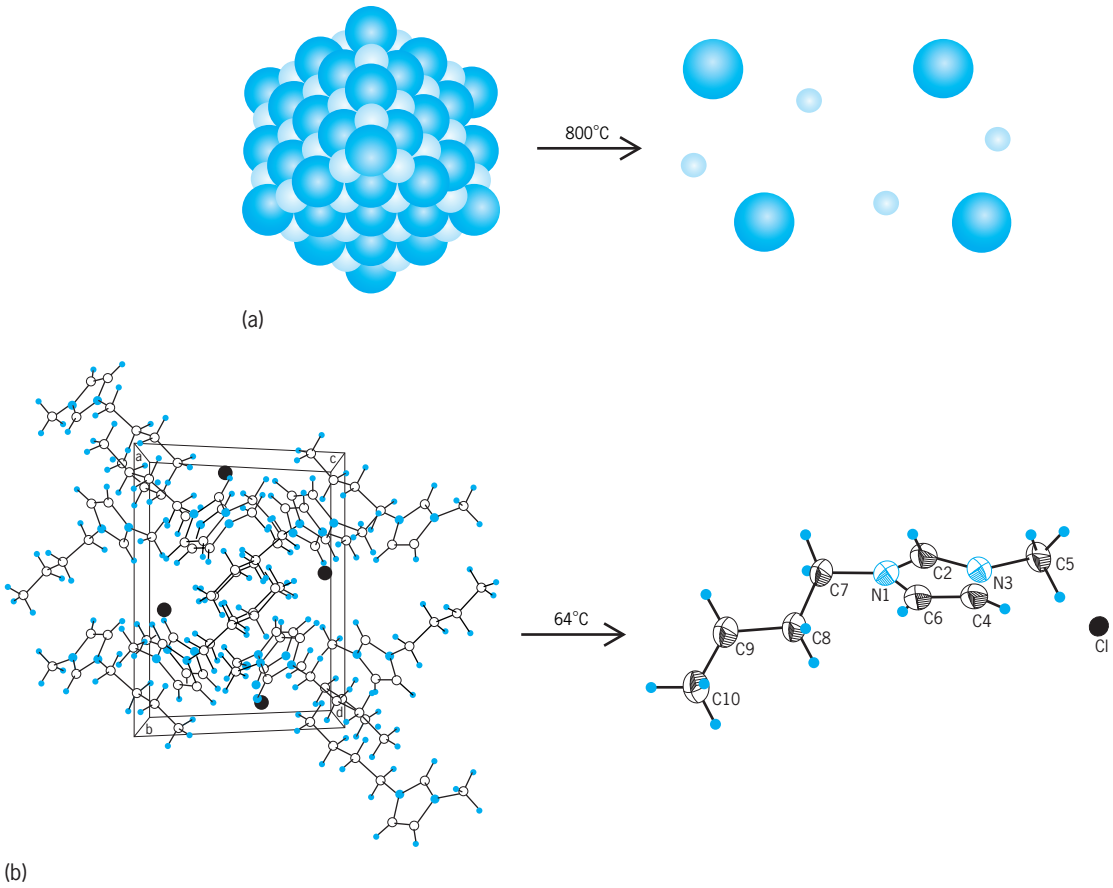

Fig. 1. Examples of (*a*) inorganic and high-melting salts and (*b*) organic and low-melting salts known as ionic liquids. (*After J. D. Holbrey et al., Crystal polymorphism in 1-butyl-3 methylimidazolium halides: Supporting ionic liquid formation by inhibition of crystallization, Chem. Commun., 2003:1636–1637, 2003*)

melt at, or near, ambient temperature (below 100–150°C; 212–302°F). Ionic liquids chemically resemble both phase-transfer catalysts and surfactants and, in some cases, have been used for both purposes. *See* PHASE-TRANSFER CATALYSIS; SALT (CHEMISTRY); SURFACTANT.

The most extensively studied ionic liquids are systems derived from 1,3-dialkylimidazolium (structure **1**), tetraalkylammonium **(2)**, tetraalkylphosphonium **(3)**, and *N*-alkylpyridinium **(4)** cations.

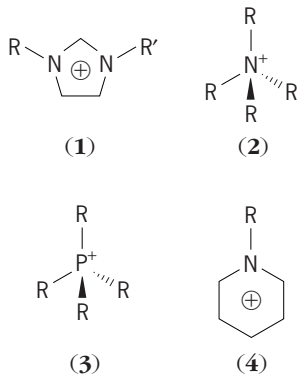

Derivatives made from these cations can increase or decrease functionality, such as alkyl chains, branching, or chirality. Common anions that allow for the formation of low-melting ionic liquids typically have diffuse (delocalized) charge and can range from simple inorganic anions such as chloride (Cl^-), bromide (Br^-), and iodide (I^-), through larger pseudospherical polyatomic anions, including hexafluorophosphate $[PF_6]^-$ and tetrafluoroborate $[BF_4]^-$, to larger, flexible fluorinated anions such as bis(trifluoromethylsulfonamide) $[(CF_3SO_2)N]^-$ and tri(perfluoroX) trifluorophosphate. A wide range of simple polyhalometallate and halometallate/halide complex anion ionic liquids have also been extensively investigated and have important application in electrochemistry and electrodisposition. The anionic component of the ionic liquid typically controls the solvent's reactivity with water, coordinating ability, and hydrophobicity. Anions can also contain chiral components or can be catalytically active, such as carboranes, polytungstates, and tetrachloroaluminate anions. Manipulation of the rheological properties through mixing of cation-anion pairs to obtain materials that support or enhance reactions makes the use of ionic liquids in organic synthesis an intriguing possibility.

Synthesis. Preparation of the cationic portion of an ionic liquid can be achieved through the quaternization of phosphines or amines with a haloalkane or through protonation with a free acid. Quaternization is generally regarded as the

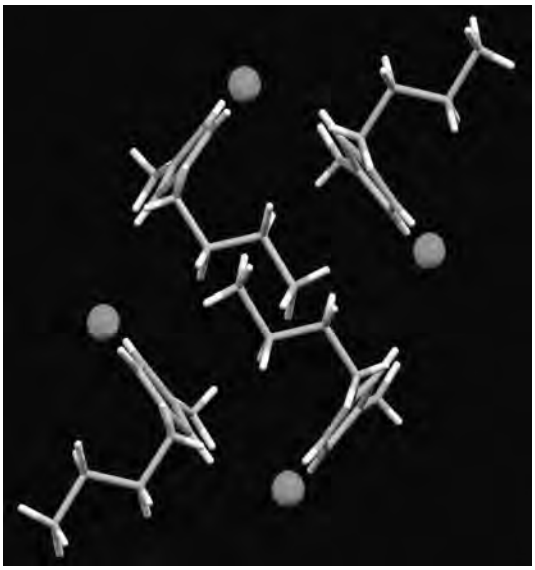

Fig. 2. X-ray crystal structure of 1-butyl-3-methylimidazolium chloride ([C_{4mim}]Cl) in its ortho polymorph (its most common conformation). (*After J. D. Holbrey et al., Crystal polymorphism in 1-butyl-3 methylimidazolium halides: Supporting ionic liquid formation by inhibition of crystallization, Chem. Commun., 2003:1636–1637, 2003. Reproduced by permission of The Royal Society of Chemistry.*)

more sound approach because cations prepared through protonation reactions can be degraded easily through deprotonation, leading to the breakdown of the solvent. Reaction time and temperature for typical quaternization reactions depend upon both the haloalkane and the cation "backbone" used. The most widely used ionic liquids in research include 1-butyl-3-methylimidazolium chloride [C_4mim]Cl, [C_4mim][PF_6], and [C_4mim][BF_4] (**Fig. 2**). Ionic liquids prepared using this method can be modified through anion exchange reactions, acid treatment, or metathesis reactions to prepare ionic liquids containing a desired anion. (The choice of anion is important, as it is the most influential factor on the physical properties of an ionic liquid.) Through these basic reactions, it is possible to prepare a wide range of ionic liquids possessing varied physical properties that can be used to aid in creating optimal reaction conditions. *See* QUATERNARY AMMONIUM SALTS.

Purification of ionic liquids is of utmost importance, specifically those used in organic synthesis, as impurities from preparation could adversely affect the reactions. To synthesize pure ionic liquids, it is necessary to begin with purified starting materials. However, the use of pure starting materials does not ensure pure product. Common impurities in ionic liquids include halide anions from metathesis reactions, the presence of color and water, and acid impurities. Dichloromethane extraction is the most widely accepted method for removing excess halide anions from ionic liquids. Since ionic liquids typically have no measurable vapor pressure, dichloromethane can be removed easily under vacuum. As with organic solvents, water can also be removed from an ionic liquid under vacuum with heat to drive off unwanted moisture. Ionic liquids normally exist as colorless solvents. Impurities that can cause an ionic liquid to become colored can be avoided by using pure starting materials, avoiding the use of acetone in the cleaning of glassware, and keeping temperatures as low as possible during the quaternization step of its synthesis.

Properties. Interest in ionic liquids can be attributed to a number a unique qualities inherent to these materials. First, ionic liquids are excellent conducting materials that can be used in a number of electrochemical applications, including battery production and metal deposition. In fact, they were first designed to be used in electrochemical applications. Second, ionic liquids typically have no measurable vapor pressure, making them an attractive replacement for volatile organic solvents in synthesis. Third, ionic liquids are composed of ions that can be varied to create materials with vastly different physical and chemical properties, compared with conventional solvents. The ability to "fine-tune" the solvent properties of ionic liquids has been exploited to create solvents ideal for a range of organic syntheses. Lastly, ionic liquids exhibit varying degrees of solvation and solubility in a range of organic solvents, allowing simple separations and extractions. *See* SOLVENT.

Use in organic reactions. The replacement of volatile organic solvents with ionic liquids in organic reactions is very exciting. Ionic liquids are polar solvents which can be used as both solvents and reagents in organic synthesis, and the possibility exists that ionic liquids may positively affect the outcome of the reactions.

Certain ionic liquids are suitable for particular classes of organic reactions. For example, neutral ionic liquids are commonly used in Diels-Alder and condensation reactions, as well as nucleophilic displacement. Ionic liquids that possess Lewis acid properties are used in acid-catalyzed reactions, including Friedel-Crafts alkylations and acylations and electrophilic substitutions or additions. *See* ACID AND BASE; ACYLATION; DIELS-ALDER REACTION; FRIEDEL-CRAFTS REACTION; SUBSTITUTION REACTION.

Product isolation from ionic liquids is generally accomplished through extraction with organic solvents. Because most ionic liquids have no measurable vapor pressure, simple vacuum techniques can be used to recover the product in the organic solvent. Biphasic systems composed of ionic liquids and water or organic phases have been used in catalysis reactions exploiting the properties of some ionic liquids to easily recover both product and catalyst. As stated previously, the ionic constituents of an ionic liquid can be "tuned" to exhibit desired physical properties to support such biphasic reactions.

The unique properties of ionic liquids, such as their stability and nonvolatility, make them good candidates for use as solvents in homogeneous catalysis systems. Polymerization reactions in ionic liquids, using transition-metal catalysts and conventional

organic initiators, have been studied and have demonstrated increased efficiency compared with traditional polymerization solvents. *See* HOMOGENEOUS CATALYSIS.

Green chemistry. Sustained interest in ionic liquids can be attributed to their desirable physical properties, such as their electron conductivity, as well as their novelty and the possibility of enhanced reactions. The perceived potential to eliminate volatile organic solvents in synthetic and separation processes has also driven interest and investigation into ionic liquids.

Ionic liquids have many characteristics relevant to a general "green chemistry" approach, including the lack of volatility; however, almost all other properties (such as toxicity, stability, and reactivity) vary with the cation and anion components and cannot readily be generalized. The utility and interest in ionic liquids, as a class of fluids, rests with individual examples displaying new, improved, or different combinations of solvent properties. *See* GREEN CHEMISTRY.

M. B. Turner; J. D. Holbrey; S. K. Spear; R. D. Rogers

Bibliography. R. D. Rogers and K. R. Seddon (eds.), *Ionic Liquids; Industrial Applications to Green Chemistry*, ACS Symp. Ser. 818, American Chemical Society, Washington, DC, 2002; R. D. Rogers and K. R. Seddon (eds.), *Ionic Liquids as Green Solvent, Progress and Prospects*, ACS Symp. Ser. 856, American Chemical Society, Washington, DC, 2003; P. Wasserscheid and T. Welton (eds.), *Ionic Liquids in Synthesis*, Wiley-VCH, Weinheim, 2002; T. Welton, Room temperature ionic liquids. Solvents for synthesis and catalysis, *Chem. Rev.*, 99:2071, 1999.

Ionization

The process by which an electron is removed from an atom, molecule, or ion. This process is of basic importance to electrical conduction in gases, liquids, and solids. In the simplest case, ionization may be thought of as a transition between an initial state consisting of a neutral atom and a final state consisting of a positive ion and a free electron. In more complicated cases, a molecule may be converted to a heavy positive ion and a heavy negative ion which are separated.

Ionization may be accomplished by various means. For example, a free electron may collide with a bound atomic electron. If sufficient energy can be exchanged, the atomic electron may be liberated and both electrons separated from the residual positive ion. The incident particle could as well be a positive ion. In this case the reaction may be considerably more complicated, but may again result in a free electron. Another case of considerable importance is the photoelectric effect. Here a photon interacts with a bound electron. If the photon has sufficient energy, the electron may be removed from the atom. The photon is annihilated in the process. Other methods of ionization include thermal processes, chemical reactions, collisions of the second kind, and collisions with neutral molecules or atoms. *See* ELECTRICAL CONDUCTION IN GASES; ELECTRODE POTENTIAL.

Glenn H. Miller

Bibliography. R. Gomer, *Field Emission and Field Ionization*, 1961, reprint 1993; H. Russotti, *The Study of Ionic Equilibria: An Introduction*, 1978; D. Veza (ed.), *Physics of Ionized Gases*, 1995; A. Von Engel, *Ionized Gases*, 2d ed., 1965, reprint 1994.

Ionization chamber

An instrument for detecting ionizing radiation by measuring the amount of charge liberated by the interaction of ionizing radiation with suitable gases, liquids, or solids. These radiation detectors have played an important part in the development of modern physics and have found many applications in basic scientific research, in industry, and in medicine.

Principle of operation. While the gold leaf electroscope (**Fig. 1**) is the oldest form of ionization chamber, instruments of this type are still widely used as monitors of radiation by workers in the nuclear or radiomedical professions. In this device, two thin flexible pieces of gold leaf are suspended in a gas-filled chamber. When these are electrically charged, as in Fig. 1, the electrostatic repulsion causes the two leaves to spread apart. If ionizing radiation is incident in the gas, however, electrons are liberated from the gas atoms. These electrons then drift toward the positive charge on the gold leaf, neutralizing some of this charge. As the charge on the gold leaves decreases, the electrostatic repulsion decreases, and hence the separation between the leaves decreases. By measuring this change in separation, a measure is obtained of the amount of radiation incident on the gas volume. While this integrated measurement may be convenient for applications such as monitoring the total radiation exposure of humans, for many purposes it is useful to measure the ionization pulse produced by a single ionizing particle. *See* ELECTROSCOPE.

The simplest form of a pulse ionization chamber consists of two conducting electrodes in a container

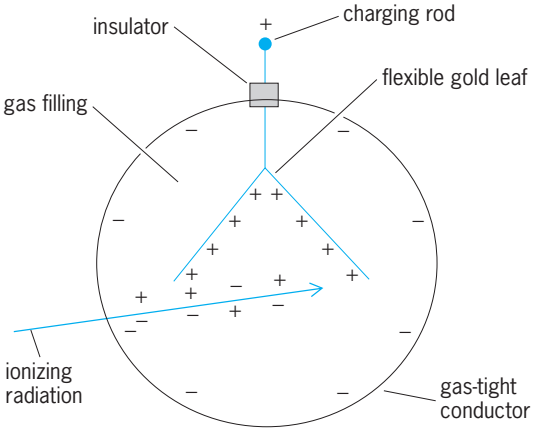

Fig. 1. Gold leaf electroscope used as a radiation detector.

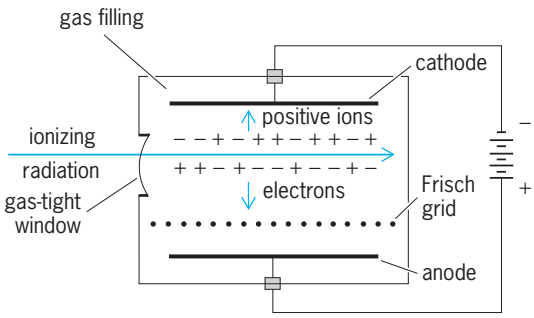

Fig. 2. Parallel-plate ionization chamber.

filled with gas (**Fig. 2**). A battery, or other power supply, maintains an electric field between the positive anode and the negative cathode. When ionizing radiation penetrates the gas in the chamber—entering, for example, through a thin gas-tight window—this radiation liberates electrons from the gas atoms leaving positively charged ions. The electric field present in the gas sweeps these electrons and ions out of the gas, the electrons going to the anode and the positive ions to the cathode.

The basic ion chamber signal consists of the current pulse observed to flow as a result of this ionization process. Because the formation of each electron-ion pair requires approximately 30 eV of energy on the average, this signal is proportional to the total energy deposited in the gas by the ionizing radiation.

Because the charge liberated by a single particle penetrating the chamber is small, very low-noise high-gain amplifiers are needed to measure this charge. In the early days, this was a severe problem, but such amplifiers have become readily available with the development of modern solid-state electronics. *See* AMPLIFIER.

In a chamber, such as that represented in Fig. 2, the current begins to flow as soon as the electrons and ions begin to separate under the influence of the applied electric field. The time it takes for the full current pulse to be observed depends on the drift velocity of the electrons and ions in the gas. These drift velocities are complicated functions of gas type, voltage, and chamber geometry. However, because the ions are thousands of times more massive than the electrons, the electrons always travel several orders of magnitude faster than the ions. As a result, virtually all pulse ionization chambers make use of only the relatively fast electron signal. The electron drift velocities for a few gases are given in **Fig. 3**. Using one of the most common ion chamber gases—argon with a small amount of methane—with electrode spaces of a few centimeters and voltages of a few hundred volts, the electron drift time is of order a microsecond, while the positive-ion drift time is of order milliseconds. By using narrow-bandpass amplifiers sensitive only to signals with rise times of order a microsecond, only the electron signals are observed.

Energy spectrum. One of the most important uses of an ionization chamber is to measure the total energy of a particle or, if the particle does not stop in the ionization chamber, the energy lost by the particle in the chamber. When such an energy-sensitive application is needed, a simple chamber geometry such as that shown in Fig. 2 is not suitable because the fast electron signal charge is a function of the relative distance that the ionization event occurred from the anode and cathode. If an ionization event occurs very near the cathode, the electrons drift across the full electric potential V_0 between the chamber electrodes, and a full electron current pulse is recorded; if an ionization event occurs very near the anode, the electrons drift across a very small electric potential, and a small electron pulse is recorded. This geometrical sensitivity is a result of image charges induced by the very slowly moving positive ions. It can be shown that if the electrons drift through a potential difference ΔV, the fast electron charge pulse is $q' = (\Delta V/V_0)q$, where q is the total ionization charge liberated in the gas.

This geometrical dependence can be eliminated by introducing a Frisch grid as indicated in **Fig. 4**. This grid shields the anode from the positive ions and, hence, removes the effects of the image charges. By biasing the anode positively, relative to the grid, the electrons are pulled through the grid and collected on the anode. Now no signal is observed on the anode until the electrons drift through the grid, but the signal charge which is then observed is the full ionization charge q.

While the ionization chamber generates only small quantities of signal charge for incident particles or photons of megaelectronvolt energies, the resulting signals are nevertheless well above the noise level of modern low-noise electronic amplifiers. When the signals generated by many incident particles of the same energy are individually measured, and a histogram is plotted representing the magnitude of a signal pulse versus the total number of pulses with that magnitude, then an energy spectrum results.

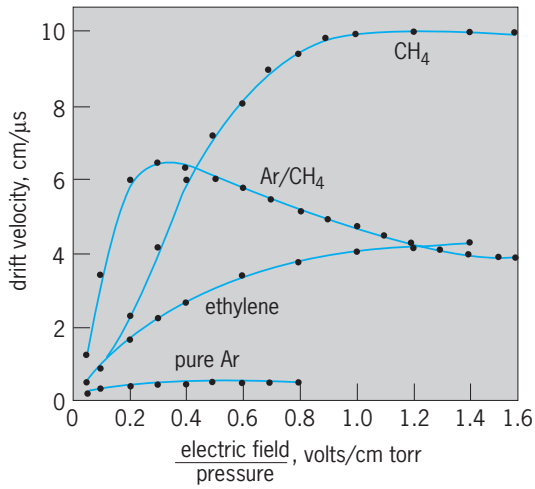

Fig. 3. Electron drift velocity in four different gases as a function of the ratio of the applied electric field strength in volts/centimeter to gas pressure in torrs. 1 torr = 133 Pa. (*After H. W. Fulbright, Ionization chambers, Nucl. Instrum. Meth., 162(1979):21–28,1979*)

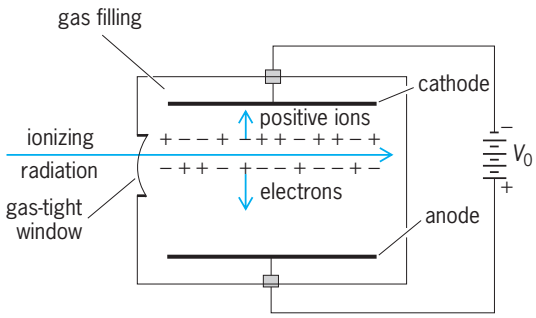

Fig. 4. Frisch grid parallel-plate ionization chamber.

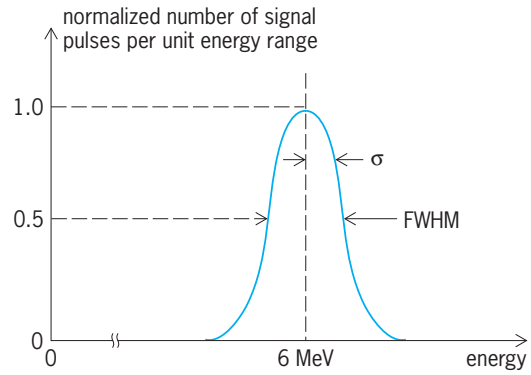

Fig. 5. Idealized energy spectrum produced by monoenergetic (6-MeV) alpha particles incident on an ideal gridded ionization chamber. The spectrum consists of a single gaussian peak with standard deviation σ.

Such a spectrum, smoothed out, consists of an essentially gaussian distribution with standard deviation σ (**Fig. 5**). Assuming a negligible contribution from amplifier noise, it might at first sight appear that σ should correspond to the square root of the average number of electron-ion pairs produced per incident particle. In fact, σ is usually found to be less than this by a substantial amount, usually designated F, where F is the Fano factor. *See* PROBABILITY; STATISTICS.

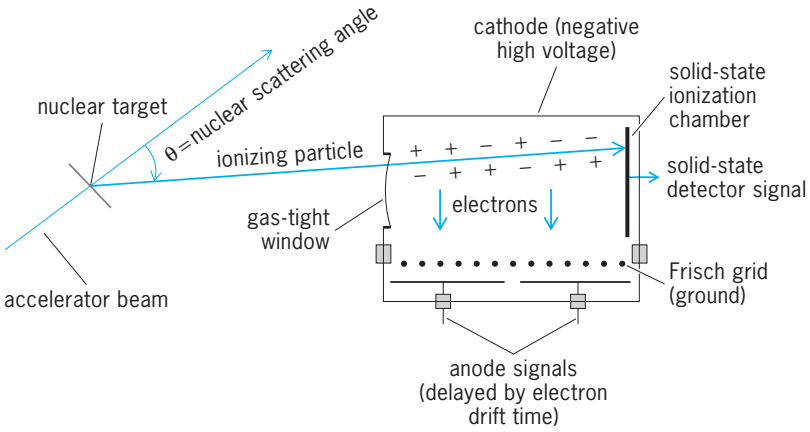

Fig. 6. Heavy-ion detector telescope used to study nuclear reactions.

It is usual to express the width of an energy distribution such as that of Fig. 5 not in terms of σ but in terms of the "full width at half maximum," usually designated FWHM, or Δ. It can be shown that, for situations in which the width of the energy spectrum is governed by statistics alone, the FWHM is given by the equation $\sigma = 2.36F\epsilon E$, where ϵ is the average energy required to create an electron-ion pair and E is the energy deposited in the chamber by each incident particle or photon. Values of the Fano factor F as low as 0.1 have been observed for certain gases.

Gaseous ionization chambers. Because of the very few basic requirements needed to make an ionization chamber (namely, an appropriate gas with an electric field), a wide variety of different ionization chamber designs are possible in order to suit special applications. In addition to energy information, ionization chambers are now routinely built to give information about the position within the gas volume where the initial ionization event occurred. This information can be important not only in experiments in nuclear and high-energy physics where these position-sensitive detectors were first developed, but also in medical and industrial applications.

This position-sensitivity capability results from the fact that, to a good approximation, electrons liberated in an ionizing event drift along the electric field line connecting the anode and cathode, and they drift with uniform velocity (Fig. 3). Hence, a measure of the drift time is a measure of the distance from the anode that the ionization occurred. A simple illustration of this use is a heavy-ion nuclear physics detector "telescope" used in the basic study of nuclear reactions (**Fig. 6**). Ionizing charged particles (such as ^{1}H, ^{4}He, and ^{12}C) produced in nuclear collisions enter the detector telescope through a thin gas-tight window at the left, pass through two Frisch grid ionization chambers, and then stop in a solid-state ionization detector. Measurement of the ionization produced in the gas versus the total energy of the particle as determined by the solid-state ionization detector gives sufficient information to uniquely identify the mass and atomic charge of the incident particle. Because the response of the solid-state detector is fast relative to the electron drift time, the difference in time of the signals from the solid-state detector and the anode determines the electron drift time and, hence, the distance above the grid that the particle entered the ionization chamber. This distance can be used to determine the nuclear scattering angle. Hence, a very simple device can be designed which gives several pieces of useful information. While this example illustrates the principles, very complex ionization chambers are now routinely used in heavy-ion and high-energy physics where tens or a hundred signals are recorded (using a computer) for a single ionization event. Position-sensitive heavy-ion detectors with active surfaces as large as a square meter have been developed.

Aside from applications in basic nuclear physics research, ionization chambers are widely used in

other applications. Foremost among these is the use of gas ionization chambers for radiation monitoring. Portable instruments of this type usually employ a detector containing approximately 1 liter of gas, and operate by integrating the current produced by the ambient radiation. They are calibrated to read out in convenient units such as milliroentgens per hour.

Another application of ionization chambers is the use of air-filled chambers as domestic fire alarms. These employ a small ionization chamber containing a low-level radioactive source, such as ^{241}Am, which generates ionization at a constant rate, the resulting current being monitored by a small solid-state electronic amplifier. On the introduction of smoke into the chamber (which is open to the ambient air), the drifting charged ions tend to attach themselves to the smoke particles. This reduces the ionization chamber current, since the moving charge carriers are now much more massive than the initial ions and therefore exhibit correspondingly reduced mobilities. The observed reduction in ion current is used to trigger the alarm. *See* FIRE DETECTOR.

Another development in ion chamber usage was that of two-dimensional imaging in x-ray medical applications to replace the use of photographic plates. This imaging depends on the fact that if a large flat parallel-plate gas ionization chamber is illuminated with x-rays (perpendicular to its plane), the resulting charges will drift to the plates and thereby form an "image" in electrical charge of the point-by-point intensity of the incident x-rays. This image can be recorded xerographically by arranging for one plate to be a suitably charged insulator. This insulator is preferentially discharged by the collected ions. The resulting charge pattern is recorded by dusting the insulator with a fine powder and transferring this image to paper in the usual xerographic technique. Alternatively, the xerographic insulator may be a photoconductor, such as selenium, which is preferentially discharged by the ionization produced in the solid material. This is then an example of a solid ionization chamber, and its action closely parallels the operation of the optical xerographic copying machines. Such x-ray imaging detectors provide exceedingly high-quality images at a dosage to the patient substantially less than when photographic plates are used. *See* PHOTOCOPYING PROCESSES.

Gaseous ionization chambers have also found application as total-energy monitors for high-energy accelerators. Such applications involve the use of a very large number of interleaved thin parallel metal plates immersed in a gas inside a large container. An incident pulse of radiation, due for example to the beam from a large accelerator, will produce a shower of radiation and ionization inside the detector. If the detector is large enough, essentially all of the incident energy will be dissipated inside the detector (mostly in the metal plates) and will produce a corresponding proportional quantity of charge in the gas. By arranging that the plates are alternately biased at a positive and negative potential, the entire device operates like a large interleaved gas ion chamber. The total collected charge is then a measure of the total energy in the initial incident pulse of radiation.

Solid ionization chambers. Ionization chambers can be made where the initial ionization occurs, not in gases, but in suitable liquids or solids. In fact, the discovery of extremely successful solid-state ionization detectors in the early 1960s temporarily diverted interest from further developments of gas-filled chambers.

In the solid-state ionization chamber (or solid-state detector) the gas filling is replaced by a large single crystal of suitably chosen solid material. In this case the incident radiation creates electron-hole pairs in the crystal, and this constitutes the signal charge. In practice, it has been found that only very few materials can be produced with a sufficiently high degree of crystalline perfection to allow this signal charge to be swept out of the crystal and collected. Although many attempts were made in this direction in the 1940s in crystal counters, using such materials as AgCl, CdS, and diamond, these were all failures due to the crystals not having adequate carrier transport properties. In the late 1950s, however, new attempts were made in this direction using single crystals of the semiconductors silicon and germanium. These were highly successful and led to detectors that revolutionized low-energy nuclear spectroscopy.

There are two important differences between solid and gas-filled ionization chambers. First, it takes much less energy to create an electron-hole pair in a solid than it does to ionize gas atoms. Hence, the intrinsic energy resolution obtainable with solid-state detectors is better than with gas counters. Gamma-ray detectors with resolutions better than 180 eV are commercially available. Second, in the case of solid semiconductors, the positive charge is carried by electron holes whose mobilities are similar to those of electrons. Hence, both the electrons and holes are rapidly swept away by the electric field and, as a result, no Frisch grid is needed to electrically shield the anode from the image charge effects of slow-moving positive ions as in the case of gas- or liquid-filled ionization chambers. *See* CRYSTAL COUNTER; GAMMA-RAY DETECTORS; SEMICONDUCTOR.

Liquid ionization chambers. The use of a liquid in an ionization chamber combines many of the advantages of both solid and gas-filled ionization chambers; most importantly, such devices have the flexibility in design of gas chambers with the high density of solid chambers. The high density is especially important for highly penetrating particles such as gamma rays. Unfortunately, until the 1970s the difficulties of obtaining suitable high-purity liquids effectively stopped development of these detectors. During the 1970s, however, a number of groups built liquid argon ionization chambers and demonstrated their feasibility. A Frisch grid liquid argon chamber achieved a resolution of 34 keV (FWHM).

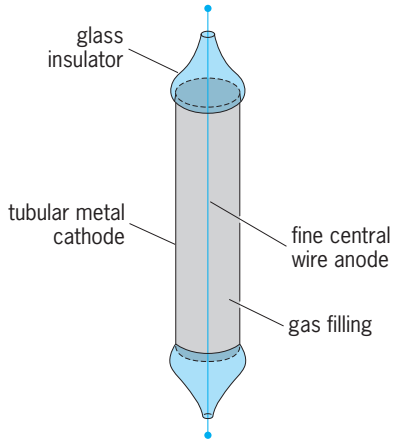

Fig. 7. Basic form of a simple single-wire gas proportional counter.

Proportional counters. If the electric field is increased beyond a certain point in a gas ionization chamber, a situation is reached in which the free electrons are able to create additional electron-ion pairs by collisions with neutral gas atoms. For this to occur, the electric field must be sufficiently high so that between collisions an electron can pick up an energy that exceeds the ionization potential of the neutral gas atoms. Under these circumstances gas multiplication, or avalanche gain, occurs, thereby providing additional signal charge from the detector.

A variety of electrode structures have been employed to provide proportional gas gain of this type. The most widely used is shown in **Fig. 7**. Here a fine central wire acts as the anode, and the avalanche gain takes place in the high field region immediately surrounding this wire. In practice, under suitable circumstances, it is possible to operate at gas gains of up to approximately 10^6.

The gas gain is a function of the bias voltage applied to the proportional counter and takes the general form shown in **Fig. 8**.

Similar avalanche multiplication effects can occur in semiconductor junction detectors, although there the situation is less favorable, and such devices have not found very widespread use except as optical detectors. *See* JUNCTION DETECTOR.

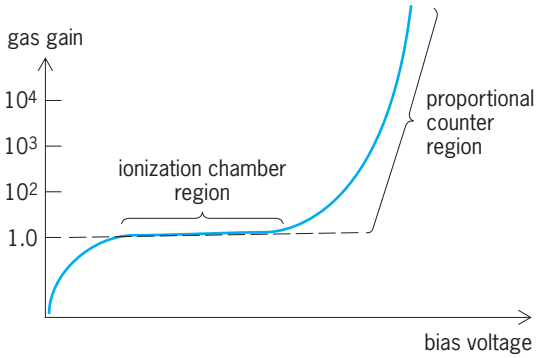

Fig. 8. Plot of gas gain versus applied voltage for a gas-filled radiation detector.

The large gas gains realizable with proportional counters have made them extremely useful for research applications involving very low-energy radiation. In addition, their flexibility in terms of geometry has made it possible to construct large-area detectors, of the order of 10 ft^2 (1 m^2), suitable for use as x-ray detectors in space. Essentially all that has been learned to date regarding x-ray astronomy has involved the use of such detectors aboard space vehicles. *See* X-RAY TELESCOPE.

Further exceedingly useful applications of gas proportional counters involve their use as position-sensitive detectors. In Fig. 7, for example, if the anode wire is grounded at both ends, then the signal charge generated at a point will split and flow to ground in the ratio of the resistance of the center wire between the point of origin and the two ends of the wire. This device therefore comprises a one-dimensional position-sensitive detector. Such devices are widely used as focal plane detectors in magnetic spectrographs. Similar position-sensitive operation can be obtained by taking account of the rise time of the signals seen at each end of the wire. Further extension of such methods allows two-dimensional detectors to be produced, a wide variety of which are under investigation for medical and other imaging uses.

The relatively large signals obtainable from gas proportional counters simplifies the requirements of the subsequent amplifiers and signal handling systems. This has made it economically feasible to employ very large arrays, on the order of thousands, of such devices in multidimensional arrays in high-energy physics experiments. By exploiting refinements of technique, it has proved possible to locate the tracks of charged particles to within a fraction of a millimeter in distances measured in meters. Such proportional counter arrays can operate at megahertz counting rates since they do not exhibit the long dead-time effects associated with spark chambers.

Geiger counters. If the bias voltage across a proportional counter is increased sufficiently, the device enters a new mode of operation in which the gas gain is no longer proportional to the initial signal charge but saturates at a very large, and constant, value. This provides a very economical method of generating signals so large that they need no subsequent amplification. *See* GEIGER-MÜLLER COUNTER.

The most widespread use of Geiger counters continues to be in radiation monitoring, where their large output signals simplify the readout problem. They have also found extensive use in cosmic-ray research, where again their large signals have made it feasible to use arrays of substantial numbers of detectors without excessive expenditures on signal-processing electronics. *See* PARTICLE DETECTOR.

William A. Lanford

Bibliography. D. A. Bromley (ed.), *Detectors in Nuclear Science*, 1979; C. F. Delaney and E. C. Finch, *Radiation Detectors: Physical Principles and Applications*, 1992; K. Kleinknecht, *Detectors for Particle Radiation*, 2d ed., 1999; G. F. Knoll, *Radiation Detection and Measurement*, 3d ed., 1999.

Ionization potential

The potential difference through which a bound electron must be raised to free it from the atom or molecule to which it is attached. In particular, the ionization potential is the difference in potential between the initial state, in which the electron is bound, and the final state, in which it is at rest at infinity.

The concept of ionization potential is closely associated with the Bohr theory of the atom. Although the simple theory is applicable only to hydrogenlike atoms, the picture furnished by it conveys the idea quite well. In this theory, the allowed energy levels for the electron are given by the equation below,

$$E_n = -\frac{k}{n^2} \qquad n = 1, 2, 3, \ldots$$

where E_n is the energy of the state described by n. The constant k is about 13.6 eV for atomic hydrogen. The energy approaches zero as n becomes infinite. Thus zero energy is associated with the free electron. On the other hand, the most tightly bound case is given by setting n equal to unity. By the definition given above, the ionization potential for the most tightly bound, or ground, state is then 13.6 eV. The ionization potential for any excited state is obtained by evaluating E_n for the particular value of n associated with that state. For a further discussion of the energy levels of an atom *see* ATOMIC STRUCTURE AND SPECTRA; ELECTRONVOLT.

The ionization potential for the removal of an electron from a neutral atom other than hydrogen is more correctly designated as the first ionization potential. The potential associated with the removal of a second electron from a single ionized atom or molecule is then the second ionization potential, and so on.

Ionization potentials may be measured in a number of ways. The most accurate measurement is obtained from spectroscopic methods. The transitions between energy states are accompanied by the emission or absorption of radiation. The wavelength of this radiation is a measure of the energy difference. The particular transitions that have a common final energy state are called a series. The series limit represents the transition from the free electron state to the particular state common to the series. The energy associated with the series limit transition is the ionization energy.

Another method of measuring ionization potentials is by electron impact. Here the minimum energy needed for a free electron to ionize in a collision is determined. The accuracy of this type of measurement cannot approach that of the spectroscopic method. *See* ELECTRON CONFIGURATION; ELECTRONEGATIVITY; STRUCTURAL CHEMISTRY. Glenn H. Miller

Bibliography. H. C. Ohanian, *Modern Physics*, 2d ed., 1995; J. W. Rohlf, *Modern Physics from A to Z*, 1994; T. R. Sandin, *Essentials of Modern Physics*, 1989.

Ionophore

A substance that can transfer ions from a hydrophilic medium such as water into a hydrophobic medium, such as hexane or a biological membrane, where the ions typically would not be soluble, also known as an ion carrier. The ions transferred are usually metal ions [for example, lithium (Li^+), sodium (Na^+), potassium (K^+), magnesium (Mg^{2+}), or calcium (Ca^{2+})]; but there are ionophores that promote the transfer of other ions, such as ammonium ion (NH_4^+) or amines of biological interest. *See* ION.

Ionophores were discovered in the early 1960s, when it was found that certain antibiotics, whose mechanism of action was unknown at the time, depended on the presence of ions such as potassium for their biological activity. Mitochondria isolated from cells would swell when placed in an aqueous medium containing small amounts of potassium ions and the antibiotic valinomycin. Measurement of the pH and potassium-ion concentration of the medium revealed that the potassium concentration decreased while the hydrogen concentration increased correspondingly. Thus, valinomycin was catalyzing the exchange of potassium ions in the medium for hydrogen ions present in the mitochondria. The swelling of the mitochondria was found to be due to an increase in their water content, which was necessary to rehydrate the potassium ions. The effect was reasonably specific for potassium ion, as smaller ions such as sodium and lithium were without effect; and, if either valinomycin or the potassium ions were independently eliminated from the experiment, no swelling was observed. Thus valinomycin had promoted the transfer of potassium ion across a hydrophobic cell membrane.

Mechanism of transfer. There are two different mechanisms by which ionophores promote the transfer of ions across hydrophobic barriers: ion-ionophore complex formation and ion channel formation (see **illus.**).

In complex formation, the ion forms a coordination complex with the ionophore in which there is a well-defined ratio (typically 1:1) of ion to ionophore. In these complexes the ionophore wraps around the ion so that the ion exists in the polar interior of the complex while the exterior is predominantly hydrophobic in character, and as such is

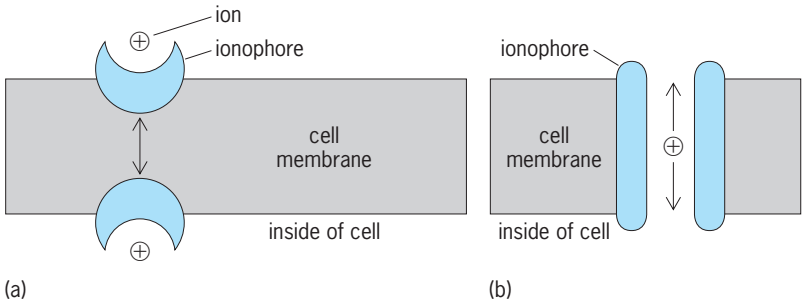

Two mechanisms by which ionophores promote ion transfer across cell membranes: (*a*) ion-ionophore complex and (*b*) ion-channel-formation ion transfer.

soluble in nonpolar media. In these complexes the ion is coordinated by oxygen atoms present in the ionophore molecule through ion-dipole interactions. The ionophore molecule essentially acts as the solvent for the ion, replacing the aqueous solvation shell that normally surrounds the ion. An example of an ionophore that forms ion-ionophore complexes is nonactin (**1**). *See* COORDINATION COMPLEXES.

Ionophores that act via ion channel formation are found in biological environments. In this type of ionophore the molecule forms a polar channel in an otherwise nonpolar cell membrane, allowing passage of small ions either into or out of the cell. An example of an ion-channel-forming ionophore is amphotericin-B (**2**).

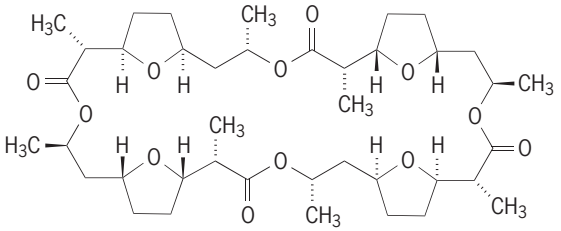

(**1**)

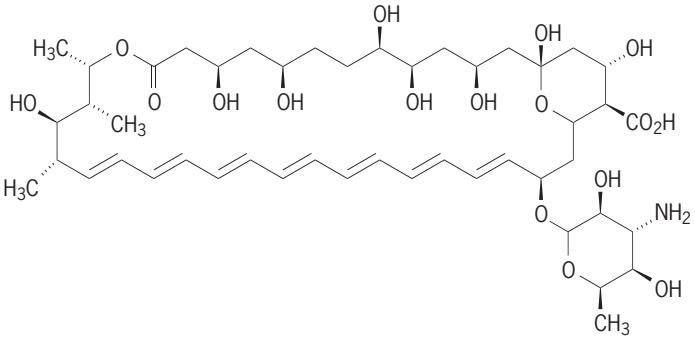

(**2**)

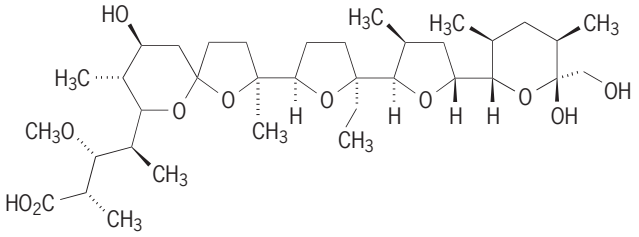

(**3**)

Types. Ionophores were discovered as metabolites of microorganisms, but as interest in these compounds has increased, compounds possessing ionophoric activity have been designed and synthesized by chemists.

Naturally occurring compounds. There are several different types of naturally occurring ionophores. These have been classified into four classes, each of which has antibiotic activity.

1. Peptide ionophores are exemplified by the gramicidins, a family of polypeptide-containing antibiotics. These compounds are thought to form ion channels via the dimerization of two gramicidin molecules to form a helical conformation in a cell membrane. Ions may then pass down the polar interior of the helix.

2. Valinomycin is an example of the cyclic depsipeptide class of ionophores. The molecule is a 36-membered macrocycle that is made up of a repetition of valine, hydroxyvaleric acid, and lactic acid residues. The structure of valinomycin complexed to potassium ion has been determined by x-ray crystallography. In this complex the conformation of the molecule is such that the potassium ion resides in the center of the complex coordinated to six carbonyl oxygens that are arranged in an octahedral configuration about the metal ion.

3. Nonactin (**1**) is a representative of the macrotetrolide class of ionophores. In this macrocycle the tetrahydrofuranyl oxygen atoms and two of the four lactone carbonyl oxygen atoms are responsible for chelation of metal ions. *See* CHELATION.

4. The polyether ionophores form a large family of structurally complex compounds. Monensin (**3**) is the best-known member of this class and was the first polyether ionophore in which the structure was determined by x-ray crystallography. Members of this class of ionophores have in common the presence of a carboxylic acid functional group (COOH) on one end of the molecule and a hydroxyl group (OH) on the other end. In solution these groups link to each other with hydrogen bonds, folding the molecule into a cyclic array. The metal ion sits in the center of the complex and is coordinated to the ether oxygen atoms of the tetrahydrofuran and tetrahydropyran rings.

Synthetic compounds. In addition to the naturally occurring ionophores, a large number of compounds possessing ionophoric activity have been synthesized. The best-known synthetic ionophores are the crown ethers, which are cyclic polyethers containing the repeating units of (—O—CH_2—CH_2—). The name crown ethers is due to the resemblance of the molecule's conformation to a king's crown in which the oxygen atoms form the points of the crown. Simple crown ethers are named by the size of the ring and the number of oxygen atoms contained. For example, the name 18-crown-6 refers to an 18-membered ring containing 6 ether oxygen atoms. As with other ionophores, crown ether complex ions are coordinated through a series of ion-dipole interactions occurring between the ion and the ether oxygen atoms present in the ionophore. *See* MACROCYCLIC COMPOUND.

Synthetic ionophores such as the crown ethers have proven to be useful reagents in synthetic organic chemistry. These reagents can solubilize inorganic salts, for example, potassium hydroxide (KOH), in nonpolar solvents such as toluene, in which they normally would be insoluble. This capability allows reactions involving these salts to be successfully conducted in the absence of water. An example of the utility of these reagents is the very efficient hydrolysis of the ester group in *tert*-butyl

mesitoate [reaction (1)], which is readily accom-

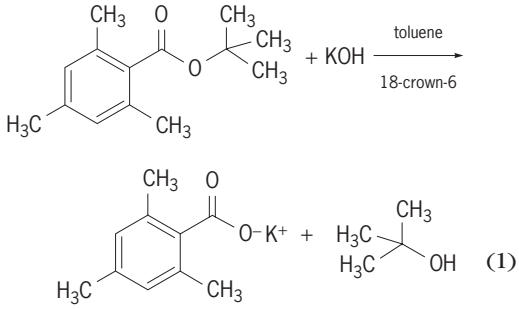

(1)

plished by potassium hydroxide in toluene in the presence of a potassium-complexing crown ether. No reaction occurs under identical conditions in the absence of the crown ether.

Many inorganic reagents (generically denoted as M^+X^-) are found to be much more reactive when their reactions are conducted in nonpolar solvents containing crown ethers. The increased reactivity is due to the poor solvation of the anion (X^-). Potassium fluoride (KF) provides an example of this effect. Normally, in aqueous solution the fluoride ion (F^-) is heavily solvated by water molecules and is not very reactive. However, when potassium fluoride is dissolved in nonpolar solvents by employing a crown ether as the cation-solvating agent [reaction (2)], the

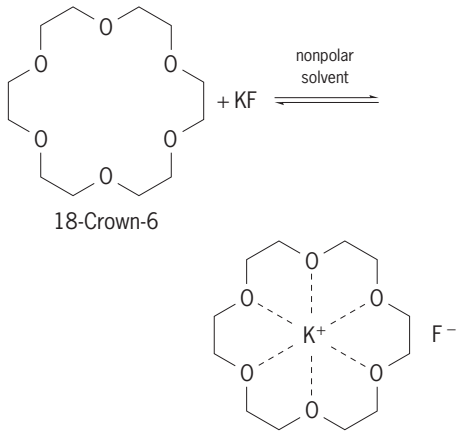

(2)

unsolvated fluoride ion becomes a very powerful nucleophile. *See* ELECTROPHILIC AND NUCLEOPHILIC REAGENTS.

Biological activity antibiotics. The biological activity of ionophore antibiotics is due to their ability to disrupt the flow of ions either into or out of cells. Under normal conditions cells have a high internal concentration of potassium ions but a low concentration of sodium ions. The concentration of ions in the extracellular medium is just the reverse, high in sodium ions but low in potassium ions. This imbalance, which is necessary for normal cell function, is maintained by the presence of a specific transport protein (sodium-potassium ATPase) present in the cell membrane that pumps sodium ions out of the cell in exchange for potassium ions. Ionophore antibiotics possess the ability to disrupt this ionic imbalance by allowing ions to penetrate the cell membrane as ion-ionophore complexes or via the formation of ion channels. Gram-positive bacteria appear to be particularly sensitive to the effect of ionophores perturbing normal ion transport. *See* ION TRANSPORT.

The pharmacological use of many of the ionophore antibiotics in humans is very restricted because of their high toxicity and multitude of physiological effects. *See* ANTIBIOTIC. Chris A. Veale

Bibliography. M. Dobler, *Ionophores and Their Structures*, 1981.

Ionosphere

The part of the upper atmosphere that is sufficiently ionized that the concentration of free electrons affects the propagation of radio waves. Existence of the ionosphere was suggested simultaneously in 1902 by O. Heaviside in England and A. E. Kennelly in the United States to explain the transatlantic radio communication that was demonstrated the previous year by G. Marconi; and for many years it was commonly referred to as the Kennelly-Heaviside layer. The existence of the ionosphere as an electronically conducting region had been postulated earlier by B. Steward to explain the daily variations in the geomagnetic field. *See* IONIZATION; RADIO-WAVE PROPAGATION.

The ionosphere has been extensively explored by rockets, satellites, and remote-sensing tools. The earliest technique involved the ionosonde, which utilizes a pulsed transmitter to send radio signals vertically upward while slowly sweeping the radio frequency. The ionosphere is a type of plasma, and such a system has a free oscillation mode. When the frequency of the signal transmitted by the ionosonde reaches a height where it equals the free oscillation frequency of the plasma, it reflects to Earth. Pulses reflected by the ionosphere are received at the transmitter and recorded; the elapsed time between pulse transmission and reception can be measured, and by using the speed of light it can be converted to an apparent distance to the point of reflection. The plasma density content can be found in this way by using an ionosonde, but no information is obtained above the point where the plasma density is maximum, so properties of the ionosphere above about 180 mi (300 km) altitude were impossible to determine. Major breakthroughs occurred in space science with the development of the artificial space satellite and sounding rockets. *See* METEOROLOGICAL ROCKET; OSCILLATION; PLASMA (PHYSICS).

Another tool for ionospheric research is the incoherent scatter radar. Powerful radio sources, such a radar, can probe out to many thousand kilometers' altitude and detect the weak signal scattered back to Earth by thermal fluctuations in the plasma. The first two of these instruments were built in 1960 near Arecibo, Puerto Rico, and Jicamarca, Peru; both are roughly 1000 ft (300 m) across. The Arecibo system is so sensitive that it is used for radio astronomy as well

as ionospheric and atmospheric science. Eight radars of this type form an observatory network stretching north-south in the Americas from Peru to Greenland and east-west from Japan to Europe. Using the Doppler technique, much like the radar in a police car, such a system can be used to measure the velocity of the plasma as well as its temperature and density. *See* DOPPLER RADAR; RADAR.

Regions. The ionosphere is highly structured in the vertical direction. It was first thought that discrete layers were involved, referred to as the D, E, F_1, and F_2 layers; however, the layers actually merge with one another to such an extent that they are now referred to as regions rather than layers. The very high temperatures in the Earth's upper atmosphere are colocated with the upper ionosphere (see **illus.**), since both are related to the effect of x-rays from the Sun. That is, the x-rays both ionize and heat the very uppermost portion of the Earth's atmosphere.

The ionosphere also displays striking geographic and temporal variations; the latter include regular diurnal, seasonal, and sunspot-cycle components and irregular day-to-day components that are associated mainly with variations in solar activity and atmospheric motion. *See* SUN.

D region. The D region is the lowest ionospheric region, extending approximately from 35 to 50 mi (60 to 85 km). The upper portion is caused mainly by the ionization of nitric oxide by Lyman-alpha radiation in sunlight, and the lower portion is mainly due to ionization by cosmic radiation. The daytime electron concentrations are about 10^8–10^9 per cubic meter. The region virtually disappears at night even though the cosmic radiation continues, because attachment of electrons to molecules (that is, formation of negative ions) quickly removes free electrons. This effect is suppressed during the daytime by photo detachment. The D region is the only ionospheric region in which negative ions are thought to be significant. *See* COSMIC RAYS.

The collision frequency for electrons with heavier particles in the D region is relatively high, which causes absorption of energy from radio signals traveling through the region. This severely limits radio propagation and is responsible for the very limited daytime range for stations in the AM broadcast band; FM radio and television signals are at such a high frequency that they pass directly through the ionosphere with no reflection possible. At night, when the D region disappears, lower-frequency radio signals can reflect many times from the ionosphere and the Earth, leading to the vast propagation distances used by amateur radio operators for communication. *See* RADIO-WAVE PROPAGATION.

The D region is located within the Earth's atmospheric region known as the mesosphere (see illus.). The coldest temperatures on Earth and the highest clouds occur here. The latter occur in the summer polar regions where temperatures below −280°F (100 K) have been recorded. These clouds are called noctilucent since they can be seen only at twilight, illuminated by sunlight at their height near 48 mi (80 km) and viewed from the dark surface of the Earth. The clouds are composed of ice; they are charged electrically and form a different type of

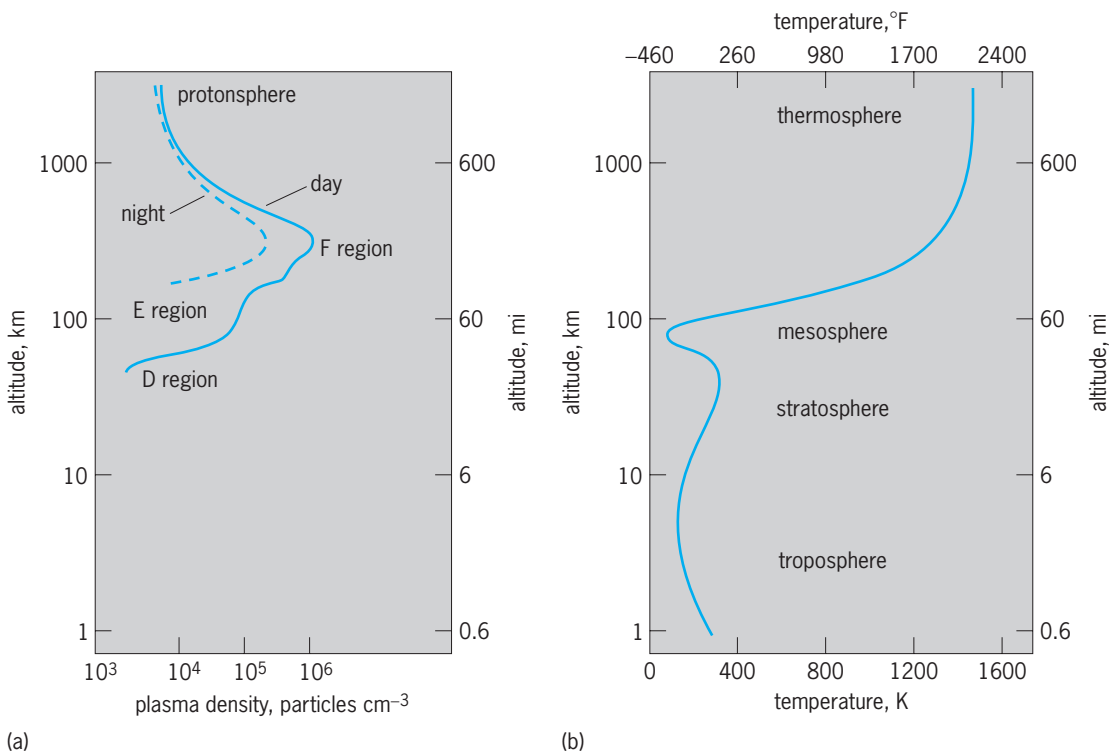

Atmospheric profiles. (*a*) Ionosphere, nighttime plasma density. (*b*) Thermal structure of the atmosphere. (*After M. C. Kelley, The Earth's Ionosphere: Plasma Physics and Electrodynamics, Academic Press, 1989*)

material known as dusty or icy plasma. *See* MESOSPHERE.

E region. Soft x-rays and the more penetrating portions of the extreme ultraviolet radiation from the Sun are absorbed in the altitude region from 50 to 85 mi 85 to (140 km), where they cause daytime electron concentrations of the order of 10^{11} per cubic meter in the E region (see illus.). This is the region from which ionospheric reflections were first identified. The principal ions have been observed to be the oxygen ions O_2^+ and O^+ and the oxynitrogen ion NO^+, the last presumably being formed in chemical reactions involving the primary oxygen ions. The soft x-rays that are principally responsible for the formation of the E region must also produce nitrogen ions (N_2^+); however, these are not observed because they are removed very rapidly by chemical reactions. When the Sun sets, these molecular ions react very rapidly with equal numbers of electrons, yielding oxygen and nitrogen atoms. These reactions are so fast that the E region virtually disappears at night unless there is an aurora present. This explains the low level of nighttime plasma density in the ionospheric profile. (illus. *a*).

The Earth is continually bombarded by meteors which burn up and become ionized when they hit the top of the atmosphere. This leaves a layer of dust, metallic atoms, and metallic ions in the E region. Ions such as iron (Fe^+), sodium (Na^+), and silicon (Si^+) do not recombine quickly because the atom cannot break apart. The result is the sporadic occurrence of long-lived clouds of metallic ion layers in the E region. The fading in and out of distant AM radio signals is due to their reflection to Earth from these sporadic E layers. *See* METEOR.

The E region supports the largest electric currents in the near space regions of the Earth. Currents as high as 10^6 A flow in the polar E region during an intense aurora. In fact, the light associated with the most beautiful of the auroral displays is generated in this region of the ionosphere and atmosphere. The electrical power that is transferred to the atmosphere in such an event is about 100 GW.

F region. The solar extreme-ultraviolet radiations produce the F region, above 85 mi (140 km). This radiation is most strongly absorbed near 95 mi (160 km). Above the region of maximum photoionization, the rate for the loss of electrons and ions decreases with altitude more rapidly than does the source of ionization, so that electron concentrations increase with altitude. This decrease in loss rate occurs because O^+ ions recombine directly with electrons only very slowly in a two-body radiative-recombination process. Another loss process predominates instead, an ion-atom interchange reaction [reaction (1)], followed by a dissociative recombination [reaction (2)]. The loss rate is controlled by the

$$O^+ + N_2 \rightarrow NO^+ + N \qquad (1)$$

$$NO^+ + e^- \rightarrow N + O \qquad (2)$$

ion-atom interchange reaction. Since the N_2 concentration decreases with altitude, the concentration of O^+ increases with altitude above the region where it is actually produced.

The increase in O^+ and electron concentration with altitude finally stops, because the atmospheric density becomes so low that there is a rapid downward flow of ions and electrons in the gravitational field. The region of maximum ionization concentration normally occurs near 180 mi (300 km), and it is known as the F peak. Most of the ionization that occurs above the peak is lost by flow into the dense atmosphere below, where the ionization can be lost by the ion-atom interchange reaction followed by dissociative recombination [reactions (1) and (2)].

The peak daytime plasma concentration in the F region is in the vicinity of 10^{12} ions per cubic meter. Above the peak, the distribution of ionization is in diffusive equilibrium in the gravitational field. The colocated atmospheric region (the thermosphere) has this property as well. Unlike the atmosphere below 60 mi (100 km), which is thoroughly mixed, the various atoms and ions can separate by their mass, the lighter particles extending to very high altitudes. Hydrogen and helium can, in fact, escape the Earth's pull.

The Earth's neutral thermosphere reaches temperatures of well over 1500°F (1188 K) and exchanges the associated energy with the plasma in the ionosphere. This high temperature is due to the absorption of the Sun's deadly x-rays and extreme-ultraviolet light, which also creates the ionized gas. Just as the ozone layer absorbs ultraviolet light in the upper stratosphere and is colocated with a temperature increase, the thermospheric temperature rise marks an important filter of these high-energy solar photons.

The thermosphere is so tenuous that the temperature at a given point responds quickly to solar heat input. The result is a huge temperature difference between night and day, as much as 900°F (500 K). In response, intense winds exceeding 450 mi/h (200 m/s) blow back and forth from high to low temperature zones. As the Earth rotates under this pattern, an observer sees a strong component at a 24-h period. This is called a solar thermal tide in analogy to the 12-h periodicity exhibited by the Earth's oceans responding to the gravitational attraction of the Moon. *See* MOON.

The ionosphere responds to these neutral tidal motions, but with a behavior that is constrained by the Earth's magnetic field and the electric currents that are generated in the conducting fluid. These wind-driven currents also create electric fields. The combined action of electric, gravitational, and magnetic fields, along with the neutral wind, controls the ionospheric material. *See* GEOMAGNETISM; THERMOSPHERE.

Heliosphere and protonosphere. Helium and atomic hydrogen are important constituents of the upper atmosphere, and their ions are also important at levels above 300 mi (500 km) or so. These gases become important, and finally predominant, constituents of the upper atmosphere because of their low mass. In diffusive equilibrium, each neutral gas is distributed

in the gravitational field just as if the other gases were not present, and the lighter ones therefore finally come to predominate over the heavier ones above some sufficiently high altitude. The terms heliosphere and protonosphere are sometimes used to designate the regions in which helium and hydrogen ions respectively are predominant. *See* HELIUM; HYDROGEN.

At midlatitudes these light ions dominate above about 600 mi (1000 km). Also above this height the magnetic field geometry becomes crucial, and a toruslike region called the plasmasphere forms; it is full of dense cool plasma. The torus extends to about four earth radii (11,4000 mi or 19,000 km above the surface) and, following the curved magnetic field lines, this boundary touches the top of the atmosphere at all latitudes below about 60° magnetic latitude (southern Canada in the American Sector). This torus can be considered an extension of the ionosphere, since it is filled from below by the ionosphere during the day and empties back into the ionosphere during the night. The plasmasphere is relatively stable, since it rotates with the planet just as the Earth's gaseous neutral atmosphere does. At the higher latitudes the ionosphere is put into motion by the solar wind and no longer rotates with the planet. *See* SOLAR WIND.

Electrodynamics. In an electrical generator a conducting material is moved across a magnetic field, and the resulting magnetic force drives a current. If there is no external connection, though, the electric charge piles up at the ends of the conductor and an electric field is generated. The electric field attains a magnitude that balances the magnetic forces and cancels the current. If a light bulb is attached to the generator, a small current will flow. In this case an electric field will still exist, but it will be smaller than in the case with no external connection.

The same effect occurs when the Earth's winds blow across its magnetic field. A current is generated by the magnetic force. But when boundaries are encountered, an electric field will be created. These electric fields greatly affect the dynamics of the ionosphere. This wind-driven source of electric fields dominates the electrodynamics of the ionosphere below about 60° latitude.

At higher latitudes the Earth's atmosphere loses control of the plasma, and the plasma is subject to strong electrical forces induced by the flow of the solar wind past the Earth. Up to 200,000 V can be impressed across the Earth's ionosphere when the Sun's magnetic field is aligned opposite to the Earth's. The high-latitude ionosphere is then put into a high-speed circulation characterized by two vortices in each hemisphere. The resulting electrical force can, through collisional forces, create a similar circulation in the upper-atmosphere neutral air. *See* SOLAR MAGNETIC FIELD.

Tremendous variations occur in the ionosphere at high latitudes because of the dynamical effects of these electrical forces and because of the additional sources of plasma production. The most notable is the visual aurora, one of the most spectacular natural sights. Views of the Earth from both scientific and defense satellites have revealed the extent and beauty of the aurora against the background of cities. *See* MAGNETOSPHERE; MILITARY SATELLITES.

The aurora has a poleward and equatorward limit during times of magnetic storms. A resident of the arctic regions of the Northern Hemisphere see the "northern" lights in their southern sky, for example. The aurora forms two rings around the poles of the Earth. The size of the rings waxes and wanes while wavelike disturbances propagate along its extent. *See* AURORA; UPPER-ATMOSPHERE DYNAMICS.

Space plasma physics. The ionosphere is subject to the same disruptions that plague attempts to control nuclear fusion by magnetic confinement. By using the principles of plasma physics, it is possible to explain many of the processes that were known since the early twentieth century to seriously affect radio communication systems. The culprits are generically referred to as plasma instabilities, and they occur when stored energy is rapidly released. Typically such a process is associated with localized turbulent zones and an irregular ionosphere, similar to severe weather in the lower atmosphere.

At low latitudes the most dominant process is the Rayleigh-Taylor instability, in which gravitational energy is released after sunset. Huge convective "thunderstorms" arise that race around the Earth, following the sunset terminator. Intense currents in the lower ionosphere also create large-amplitude plasma sound waves via the Farley-Buneman instability all over the Earth, but with two intense regions: one near the magnetic equator and one at high latitudes where the magnetic field is nearly vertical.

Some of these instabilities are induced by neutral atmospheric waves that are generated in the dense lower atmosphere and propagate upward. These gravity or buoyancy waves have wavelengths of hundreds of kilometers and are created, among other things, by weather fronts, mountain ranges, earthquakes, nuclear explosions, and aurora. Even without their role as a source of plasma instabilities, these waves push the ionosphere around, creating both wavelike horizontal undulations and intense layers in the metallic materials from meteors, the sporadic E layers. *See* ATMOSPHERE.

Michael C. Kelley; Francis S. Johnson

Bibliography. K. Davies, *Ionospheric Radio*, 1990; A. Giraud and M. Petit, *Ionospheric Techniques and Phenomena*, 1978; A. Giraud and M. Petit, *Physics of the Earth's Ionosphere*, 1978; J. K. Hargreaves, *The Solar-Terrestrial Environment*, 1992; M. C. Kelley, *The Earth's Ionosphere: Electrodynamics and Plasma Physics*, 1989; J. A. Ratcliffe, *An Introduction to the Ionosphere and Magnetosphere*, 1972.

Ipecac

A low, perennial shrub or half-shrub of the tropical forest in Brazil and Colombia. Several species are used, but the dried roots and rhizomes of *Cephaelis*

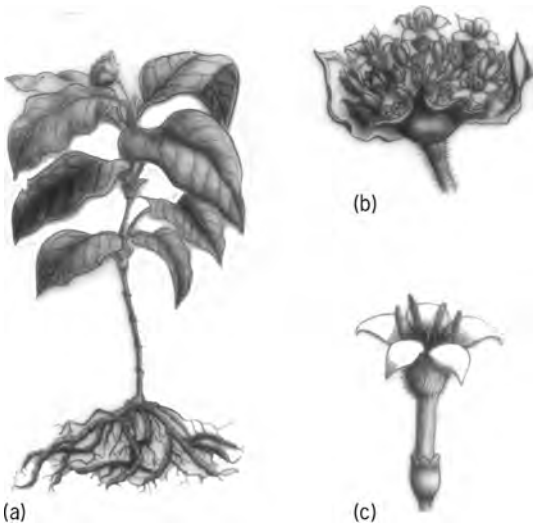

Ipecac (*Cephaelis ipecacuanha*). (*a*) Entire plant. (*b*) Inflorescence. (*c*) Flower.

ipecacuanha (Rubiaceae) constitute the material recognized as the official drug ipecac (see **illus.**). Medically the principal use is as an emetic and an expectorant. *See* RUBIALES.

P. D. Strausbaugh; Earl L. Core

Iridium

A chemical element, Ir, atomic number 77, relative atomic weight 192.22. Iridium is a transition metal and shares similarities with rhodium as well as the other platinum metals, including palladium, platinum, ruthenium, and osmium. The atom in the gas phase has the electronic configuration $1s^2$, $2s^2$, $2p^6$, $3s^2$, $3p^6$, $3d^{10}$, $4s^2$, $4p^6$ $4d^{10}$, $4f^{14}$, $5s^2$, $5p^6$, $5d^7$, $6s^2$. The ionic radius for Ir^{3+} is 0.068 nanometer and its metallic radius is 0.1357 nm. Metallic iridium is slightly less dense than osmium, which is the densest of all the elements. *See* PERIODIC TABLE.

1												13	14	15	16	17	18
1 H	2																2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

The abundance of iridium in the Earth's crust is very low, 0.001 ppm. For mining purposes, it is generally found alloyed with osmium in materials known as osmiridium and iridiosmium, with iridium contents ranging from 25 to 75%.

Solid iridium is a silvery metal with considerable resistance to chemical attack. Upon atmospheric exposure the surface of the metal is covered with a relatively thick layer of iridium dioxide (IrO_2). Important physical properties of metallic iridium are given in the **table**.

Because of its scarcity and high cost, applications of iridium are severely limited. Although iridium metal and many of its complex compounds are good catalysts, no large-scale commercial application for these has been developed. In general, other platinum metals have superior catalytic properties. The high degree of thermal stability of elemental iridium and the stability it imparts to its alloys does give rise to those applications where it has found success. Particularly relevant are its high melting point (2443°C or 4429°F), its oxidation resistance, and the fact that it is the only metal with good mechanical properties that survives atmospheric exposure above 1600°C (2910°F). Iridium is alloyed with platinum to increase tensile strength, hardness, and corrosion resistance. However, the workability of these alloys is decreased. These alloys find use as electrodes for anodic oxidation, for containing and manipulating corrosive chemicals, for electrical contacts that are exposed to corrosive chemicals, and as primary standards for weight and length. Platinum-iridium alloys are used for electrodes in spark plugs that are unusually resistant to fouling by antiknock lead additives. Iridium-rhodium thermocouples are used for high-temperature applications, where they have unique stability. Very pure iridium crucibles are used for growing single crystals of gadolinium gallium garnet for computer memory devices and of yttrium aluminum garnet for solid-state lasers. The radioactive

Physical properties of iridium metal

Property	Value
Crystal structure	Face-centered cubic
Lattice constant *a* at 25°C, nm	0.38394
Thermal neutron capture cross section, barns	440
Density at 25°C, g/cm^3	22.560
Melting point	2443°C (4429°F)
Boiling point	4500°C (8130°F)
Specific heat at 0°C, cal/g	0.0307
Thermal conductivity 0–100°C, cal cm/cm^2 s °C	0.35
Linear coefficient of thermal expansion 20–100°C, μin./in./°C	6.8
Electrical resistivity at 0°C, microhm-cm	4.71
Temperature coefficient of electrical resistance 0–100°C/°C	0.00427
Tensile strength (1000 lb/in.^{2})	
Soft	160–180
Hard	300–360
Young's modulus at 20°C	
lb/in.^{2}, static	75.0×10^6
lb/in.^{2}, dynamic	76.5×10^6
Hardness, diamond pyramid number	
Soft	200–240
Hard	600–700
ΔH_{fusion}, kJ/mol	26.4
$\Delta H_{vaporization}$, kJ/mol	612
ΔH_f monoatomic gas, kJ/mol	669
Electronegativity	2.2

isotope, ^{192}Ir, which is obtained synthetically from ^{191}Ir by irradiation of natural sources, has been used as a portable gamma source for radiographic studies in industry and medicine. *See* HIGH-TEMPERATURE MATERIALS; PLATINUM.

Bibliography. J. D. Atwood (ed.), *Comprehensive Organometallic Chemistry II: Cobalt, Rhodium and Iridium*, 1995; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; A. Earnshaw and N. Greenwood, *Chemistry of the Elements*, 2d ed., Butterworth-Heinemann, 1997.

Iron

A chemical element, Fe, atomic number 26, and atomic weight 55.847. Iron is the fourth most abundant element in the crust of the Earth (5%). It is a malleable, tough, silver-gray, magnetic metal. It melts at 1540°C, boils at 2800°C, and has a density of 7.86 g/cm^3. The four stable, naturally occurring isotopes have masses of 54, 56, 57, and 58. The two main ores are hematite, Fe_2O_3, and limonite, $Fe_2O_3 \cdot 3H_2O$. Pyrites, FeS_2, and chromite, $Fe(CrO_2)_2$, are mined as ores for sulfur and chromium, respectively. Iron is found in many other minerals, and it occurs in groundwaters and in the red hemoglobin of blood. *See* PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

The greatest use of iron is for structural steels; cast iron and wrought iron are made in quantity, also. Magnets, dyes (inks, blueprint paper, rouge pigments), and abrasives (rouge) are among the other uses of iron and iron compounds. *See* CAST IRON; IRON ALLOYS; IRON METALLURGY; STAINLESS STEEL; STEEL MANUFACTURE; WROUGHT IRON.

There are several allotropic forms of iron. Ferrite or α-iron is stable up to 760°C (1400°F). The change of β-iron involves primarily a loss of magnetic permeability because the lattice structure (body-centered cubic) is unchanged. The allotrope called γ-iron has the cubic close-packed arrangements of atoms and is stable from 910 to 1400°C (1670 to 2600°F). Little is known about δ-iron except that it is stable above 1400°C (2600°F) and has a lattice similar to that of α-iron.

The metal is a good reducing agent and, depending on conditions, can be oxidized to the 2+, 3+, or 6+ state. In most iron compounds, the ferrous ion, iron(II), or ferric ion, iron(III), is present as a distinct unit. Ferrous compounds are usually light yellow to dark green-brown in color; the hydrated ion, $Fe(H_2O)_6^{2+}$, which is found in many compounds and in solution, is light green. This ion has little tendency to form coordination complexes except with strong reagents such as cyanide ion, polyamines, and porphyrins. The ferric ion, because of its high charge (3+) and its small size, has a strong tendency to hold anions. The hydrated ion, $Fe(H_2O)_6^{3+}$, which is found in solution, combines with OH^-, F^-, Cl^-, CN^-, SCN^-, N_3^-, $C_2O_4^{2-}$, and other anions to form coordination complexes. *See* COORDINATION CHEMISTRY.

An interesting aspect of iron chemistry is the array of compounds with bonds to carbon. Cementite, Fe_3C, is a component of steel. The cyanide complexes of both ferrous and ferric iron are very stable and are not strongly magnetic in contradistinction to most iron coordination complexes. The cyanide complexes form colored salts. *See* TRANSITION ELEMENTS.

John O. Edwards

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; N. N. Greenwood and A. Earnshaw, *Chemistry of the Elements*, 2d ed., 1997; J. Silver (ed.), *Chemistry of Iron*, 1993.

Iron alloys

Solid solutions of metals, one metal being iron. A great number of commercial alloys have iron as an intentional constituent. Iron is the major constituent of wrought and cast iron and wrought and cast steel. Alloyed with usually large amounts of silicon, manganese, chromium, vanadium, molybdenum, niobium (columbium), selenium, titanium, phosphorus, or other elements, singly or sometimes in combination, iron forms the large group of materials known as ferroalloys that are important as addition agents in steelmaking. Iron is also a major constituent of many special-purpose alloys developed to have exceptional characteristics with respect to magnetic properties, electrical resistance, heat resistance, corrosion resistance, and thermal expansion. **Table 1** lists some of these alloys. *See* ALLOY; FERROALLOY; STEEL.

Because of the enormous number of commercially available materials, this article is limited to the better-known types of alloys. Emphasis is on special-purpose alloys; practically all of these contain relatively large amounts of an alloying element or elements referred to in the classification. Alloys containing less than 50% iron are excluded, with a few exceptions.

Iron-aluminum alloys. Although pure iron has ideal magnetic properties in many ways, its low electrical resistivity makes it unsuitable for use in alternating-current (ac) magnetic circuits. Addition of aluminum in fairly large amounts increases the electrical

TABLE 1. Some typical composition percent ranges of iron alloys classified by important uses*

Type	Fe	C	Mn	Si	Cr	Ni	Co	W	Mo	Al	Cu	Ti
Heat-resistant alloy castings	Bal.	0.30–0.50	—	1–2	8–30	0–7						
	Bal.	0.20–0.75	—	2–2.5	10–30	8–41						
Heat-resistant cast irons	Bal.	1.8–3.0	0.3–1.5	0.5–2.5	15–35	5 max						
	Bal.	1.8–3.0	0.4–1.5	1.0–2.75	1.75–5.5	14–30	—	—	1	—	7	
Corrosion-resistant alloy castings	Bal.	0.15–0.50	1 max	1	11.5–30	0–4	—	0.5 max				
	Bal.	0.03–0.20	1.5 max	1.5–2.0	18–27	8–31						
Corrosion-resistant cast irons	Bal.	1.2–4.0	0.3–1.5	0.5–3.0	12–35	5 max	—	—	4 max		3 max	
	Bal.	1.8–3.0	0.4–1.5	1.0–2.75	1.75–5.5	14–32	—	—	1 max		7 max	
Magnetically soft materials	Bal.	—	—	0.5–4.5								
	Bal.	—	—	—	—	—	—	—	3.5	16		
	Bal.	—	—	—	—	—	—	—	—	16		
	Bal.	—	—	—	—	—	—	—	—	12		
Permanent-magnet materials	Bal.	—	—	—	—	—	12	—	17			
	Bal.	—	—	—	—	—	12	—	20			
	Bal.	—	—	—	—	20	5	—	—	12		
	Bal.	—	—	—	—	17	12.5	—	—	10	6	
	Bal.	—	—	—	—	25	–	—	—	12		
	Bal.	—	—	—	—	28	5	—	—	12		
	Bal.	—	—	—	—	14	24	—	—	8	3	
	Bal.	—	—	—	—	15	24	—	—	8	3	1.25
Low-expansion alloys	Bal.	—	0.15	0.33	—	36						
	Bal.	—	0.24	0.03	—	42						
	61–53	0.5–2.0	0.5–2.0	0.5–2.0	4–5	33–35	—	1–3				

*This table does not include any AISI standard carbon steels, alloy steels, or stainless and heat-resistant steels or plain or alloy cast iron for ordinary engineering uses; it includes only alloys containing at least 50% iron, with a few exceptions. Bal. = balance percent of composition.

resistivity of iron, making the resulting alloys useful in such circuits.

Three commercial iron-aluminum alloys having moderately high permeability at low field strength and high electrical resistance nominally contain 12% aluminum, 16% aluminum, and 16% aluminum with 3.5% molybdenum, respectively. These three alloys are classified as magnetically soft materials; that is, they become magnetized in a magnetic field but are easily demagnetized when the field is removed.

The addition of more than 8% aluminum to iron results in alloys that are too brittle for many uses because of difficulties in fabrication. However, addition of aluminum to iron markedly increases its resistance to oxidation. One steel containing 6% aluminum possesses good oxidation resistance up to 2300°F (1300°C). *See* ALUMINUM ALLOYS.

Iron-carbon alloys. The principal iron-carbon alloys are wrought iron, cast iron, and steel.

Wrought iron of good quality is nearly pure iron; its carbon content seldom exceeds 0.035%. In addition, it contains 0.075–0.15% silicon, 0.10 to less than 0.25% phosphorus, less than 0.02% sulfur, and 0.06–0.10% manganese. Not all of these elements are alloyed with the iron; part of them may be associated with the intermingled slag that is a characteristic of this product. Because of its low carbon content, the properties of wrought iron cannot be altered in any useful way by heat treatment. *See* WROUGHT IRON.

Cast iron may contain 2–4% carbon and varying amounts of silicon, manganese, phosphorus, and sulfur to obtain a wide range of physical and mechanical properties. Alloying elements (silicon, nickel, chromium, molybdenum, copper, titanium, and so on) may be added in amounts varying from a few tenths to 30% or more. Many of the alloyed cast irons have proprietary compositions. *See* CAST IRON; CORROSION.

Steel is a generic name for a large group of iron alloys that include the plain carbon and alloy steels. The plain carbon steels represent the most important group of engineering materials known. Although any iron-carbon alloy containing less than about 2% carbon can be considered a steel, the American Iron and Steel Institute (AISA) standard carbon steels embrace a range of carbon contents from 0.06% maximum to about 1%. In the early days of the American steel industry, hundreds of steels with different chemical compositions were produced to meet individual demands of purchasers. Many of these steels differed only slightly from each other in chemical composition. Studies were undertaken to provide a simplified list of fewer steels that would still serve the varied needs of fabricators and users of steel products. The Society of Automotive Engineers (SAE) and the AISI both were prominent in this effort, and both periodically publish lists of steels, called standard steels, classified by chemical composition. These lists are published in the *SAE Handbook* and the AISI's *Steel Products Manuals*. The lists are altered periodically to accommodate new steels and to provide for changes in consumer requirements. There are minor differences between some of the steels listed by the AISI and SAE. Only the AISI lists will be considered here. The standard steels represent a large percentage of all steel produced and, although considerably fewer in number, have successfully replaced the large number of specialized compositions formerly used.

A numerical system is used to indicate grades of standard steels. Provision also is made to use certain letters of the alphabet to indicate the steel-making

TABLE 2. AISI standard steel designations

Type	Series designation*	Composition
Carbon steels	10xx	Nonresulfurized carbon steel
	11xx	Resulfurized carbon steel
	12xx	Rephosphorized and resulfurized carbon steel
Constructional alloy steels	13xx	Manganese 1.75%
	23xx	Nickel 3.50%
	25xx	Nickel 5.00%
	31xx	Nickel 1.25%, chromium 0.65%
	33xx	Nickel 3.50%, chromium 1.55%
	40xx	Molybdenum 0.25%
	41xx	Chromium 0.50 or 0.95%, molybdenum 0.12 or 0.20%
	43xx	Nickel 1.80%, chromium 0.50 or 0.80%, molybdenum 0.25%
	46xx	Nickel 1.55 or 1.80%, molybdenum 0.20 or 0.25%
	47xx	Nickel 1.05%, chromium 0.45%, molybdenum 0.20%
	48xx	Nickel 3.50%, molybdenum 0.25%
	50xx	Chromium 0.28 or 0.40%
	51xx	Chromium 0.80, 0.90, 0.95, 1.00 or 1.05%
	5xxxx	Carbon 1.00%, chromium 0.50, 1.00, or 1.45%
	61xx	Chromium 0.80 or 0.95%, vanadium 0.10 or 0.15% min
	86xx	Nickel 0.55%, chromium 0.50 or 0.65%, molybdenum 0.20%
	87xx	Nickel 0.55%, chromium 0.50%, molybdenum 0.25%
	92xx	Manganese 0.85%, silicon 2.00%
	93xx	Nickel 3.25%, chromium 1.20%, molybdenum 0.12%
	98xx	Nickel 1.00%, chromium 0.80%, molybdenum 0.25%
Stainless and heat-resisting steels	2xx	Chromium-nickel-manganese steels; nonhardenable, austenitic, and nonmagnetic
	3xx	Chromium-nickel steels; nonhardenable, austenitic, and nonmagnetic
	4xx	Chromium steels of two classes: one class hardenable, martensitic, and magnetic; the other nonhardenable, ferritic, and magnetic
	5xx	Chromium steels; low chromium heat resisting

*The x's are replaced by actual numerals in defining a steel grade, as explained in the text.

process, certain special additions, and steels that are tentatively standard, but these are not pertinent to this discussion. **Table 2** gives the basic numerals for the AISI classification and the corresponding types of steels. In this system the first digit of the series designation indicates the type to which a steel belongs; thus 1 indicates a carbon steel, 2 indicates a nickel steel, and 3 indicates a nickel-chromium steel. In the case of simple alloy steels, the second numeral usually indicates the percentage of the predominating alloying element. Usually, the last two (or three) digits indicate the average carbon content in points, or hundredths of a percent. Thus, 2340 indicates a nickel steel containing about 3% nickel and 0.40% carbon.

All carbon steels contain minor amounts of manganese, silicon, sulfur, phosphorus, and sometimes other elements. At all carbon levels the mechanical properties of carbon steel can be varied to a useful degree by heat treatments that alter its microstructure. Above about 0.25% carbon steel can be hardened by heat treatment. However, most of the carbon steel produced is used without a final heat treatment. *See* HEAT TREATMENT (METALLURGY).

Alloy steels are steels with enhanced properties attributable to the presence of one or more special elements or of larger proportions of manganese or silicon than are present ordinarily in carbon steel. The major classifications of alloy steels are high-strength, low-alloy; AISI alloy; alloy tool; heat-resisting; electrical; and austenitic manganese. Some of these iron alloys are discussed briefly in this article; for more detailed attention *see* STAINLESS STEEL; STEEL.

Iron-chromium alloys. An important class of iron-chromium alloys is exemplified by the wrought stainless and heat-resisting steels of the type 400 series of the AISI standard steels, all of which contain at least 12% chromium, which is about the minimum chromium content that will confer stainlessness. However, considerably less than 12% chromium will improve the oxidation resistance of steel for service up to 1200°F (650°C), as is true of AISI types 501 and 502 steels that nominally contain about 5% chromium and 0.5% molybdenum. A comparable group of heat- and corrosion-resistant alloys, generally similar to the 400 series of the AISI steels, is covered by the Alloy Casting Institute specifications for cast steels.

Corrosion-resistant cast irons alloyed with chromium contain 12–35% of that element and up to 5% nickel. Cast irons classified as heat-resistant contain 15–35% chromium and up to 5% nickel.

During World War I a high-carbon steel used for making permanent magnets contained 1–6% chromium (usually around 3.5%); it was developed to replace the magnet steels containing tungsten that had been formerly but could not then be made because of a shortage of tungsten. *See* MAGNET.

Iron-chromium-nickel alloys. The wrought stainless and heat-resisting steels represented by the type 200 and the type 300 series of the AISI standard steels are an important class of iron-chromium-nickel

alloys. A comparable series of heat- and corrosion-resistant alloys is covered by specifications of the Alloy Casting Institute. Heat- and corrosion-resistant cast irons contain 15–35% chromium and up to 5% nickel.

Iron-chromium-aluminum alloys. Electrical-resistance heating elements are made of several iron alloys of this type. Nominal compositions are 72% iron, 23% chromium, 5% aluminum; and 55% iron, 37.5% chromium, 7.5% aluminum. The iron-chromium-aluminum alloys (with or without 0.5–2% cobalt) have higher electrical resistivity and lower density than nickel-chromium alloys used for the same purpose. When used as heating elements in furnaces, the iron-chromium-aluminum alloys can be operated at temperatures of 4262°F (2350°C) maximum. These alloys are somewhat brittle after elevated temperature use and have a tendency to grow or increase in length while at temperature, so that heating elements made from them should have additional mechanical support. Addition of niobium (columbium) reduces the tendency to grow.

Because of its high electrical resistance, the 72% iron, 23% chromium, 5% aluminum alloy (with 0.5% cobalt) can be used for semiprecision resistors in, for example, potentiometers and rheostats. *See* ELECTRICAL RESISTANCE; RESISTANCE HEATING.

Iron-cobalt alloys. Magnetically soft iron alloys containing up to 65% cobalt have higher saturation values than pure iron. The cost of cobalt limits the use of these alloys to some extent. The alloys also are characterized by low electrical resistivity and high hysteresis loss. Alloys containing more than 30% cobalt are brittle unless modified by additional alloying and special processing. Two commercial alloys with high permeability at high field strengths (in the annealed condition) contain 49% cobalt with 2% vanadium, and 35% cobalt with 1% chromium. The latter alloy can be cold-rolled to a strip that is sufficiently ductile to permit punching and shearing. In the annealed state, these alloys can be used in either ac or dc applications. The alloy of 49% cobalt with 2% vanadium has been used in pole tips, magnet yokes, telephone diaphragms, special transformers, and ultrasonic equipment. The alloy of 35% cobalt with 1% chromium has been used in high-flux-density motors and transformers as well as in some of the applications listed for the higher-cobalt alloy.

Although seldom used now, two high-carbon alloys called cobalt steel were used formerly for making permanent magnets. These were both high-carbon steels. One contained 17% cobalt, 2.5% chromium, 8.25% tungsten; the other contained 36% cobalt, 5.75% chromium, 3.75% tungsten. These are considered magnetically hard materials as compared to the magnetically soft materials discussed in this article.

Iron-manganese alloys. The important commercial alloy in this class is an austenitic manganese steel (sometimes called Hadfield manganese steel after its inventor) that nominally contains 1.2% carbon and 12–13% manganese. This steel is highly resistant to abrasion, impact, and shock.

Iron-nickel alloys. The iron-nickel alloys discussed here exhibit a wide range of properties related to their nickel contents.

Nickel content of a group of magnetically soft materials ranges from 40 to 60%; however, the highest saturation value is obtained at about 50%. Alloys with nickel content of 45–60% are characterized by high permeability and low magnetic losses. They are used in such applications as audio transformers, magnetic amplifiers, magnetic shields, coils, relays, contact rectifiers, and choke coils. The properties of the alloys can be altered to meet specific requirements by special processing techniques involving annealing in hydrogen to minimize the effects of impurities, grain-orientation treatments, and so on.

Another group of iron-nickel alloys, those containing about 30% nickel, is used for compensating changes that occur in magnetic circuits due to temperature changes. The permeability of the alloys decreases predictably with increasing temperature.

Low-expansion alloys are so called because they have low thermal coefficients of linear expansion. Consequently, they are valuable for use as standards of length, surveyors' rods and tapes, compensating pendulums, balance wheels in timepieces, glass-to-metal seals, thermostats, jet-engine parts, electronic devices, and similar applications.

The first alloy of this type contained 36% nickel with small amounts of carbon, silicon, and manganese (totaling less than 1%). Subsequently, a 39% nickel alloy with a coefficient of expansion equal to that of low-expansion glasses and a 46% nickel alloy with a coefficient equal to that of platinum were developed. Another important alloy is one containing 42% nickel that can be used to replace platinum as lead-in wire in light bulbs and vacuum tubes by first coating the alloy with copper. An alloy containing 36% nickel and 12% chromium has a constant modulus of elasticity and low expansivity over a broad range of temperatures. Substitution of 5% cobalt for 5% nickel in the 36% nickel alloy decreases its expansivity. Small amounts of other elements affect the coefficient of linear expansion, as do variations in heat treatment, cold-working, and other processing procedures.

A 9% nickel steel is useful in cryogenic and similar applications because of good mechanical properties at low temperatures. Two steels (one containing 10–12% nickel, 3–5% chromium, about 3% molybdenum, and lesser amounts of titanium and aluminum and another with 17–19% nickel, 8–9% cobalt, 3–3.5% molybdenum, and small amounts of titanium and aluminum) have exceptional strength in the heat-treated (aged) condition. These are known as maraging steels.

Cast irons containing 14–30% nickel and 1.75–5.5% chromium possess good resistance to heat and corrosion. *See* THERMAL EXPANSION.

Iron-silicon alloys. There are two types of iron-silicon alloys that are commercially important: the magnetically soft materials designated silicon or electrical steel, and the corrosion-resistant, high-silicon cast irons.

Most silicon steels used in magnetic circuits contain 0.5–5% silicon. Alloys with these amounts of silicon have high permeability, high electrical resistance, and low hysteresis loss compared with relatively pure iron. Most silicon steel is produced in flat-rolled (sheet) form and is used in transformer cores, stators, and rotors of motors, and so on that are built up in laminated-sheet form to reduce eddy-current losses. Silicon-steel electrical sheets, as they are called commercially, are made in two general classifications: grain-oriented and nonoriented.

The grain-oriented steels are rolled and heat-treated in special ways to cause the edges of most of the unit cubes of the metal lattice to align themselves in the preferred direction of optimum magnetic properties. Magnetic cores are designed with the main flux path in the preferred direction, thereby taking advantage of the directional properties. The grain-oriented steels contain about 3.25% silicon, and they are used in the highest-efficiency distribution and power transformers and in large turbine generators.

The nonoriented steels may be subdivided into low-, intermediate-, and high-silicon classes. Low-silicon steels contain about 0.5–1.5% silicon and are used principally in rotors and stators of motors and generators; steels containing about 1% silicon are also used for reactors, relays, and small intermittent-duty transformers. Intermediate-silicon steels contain about 2.5–3.5% silicon and are used in motors and generators of average to high efficiency and in small- to medium-size intermittent-duty transformers, reactors, and motors. High-silicon steels contain about 3.75–5% silicon and are used in power transformers and communications equipment in highest efficiency motors, generators, and transformers.

High-silicon cast irons containing 14–17% silicon and sometimes up to 3.5% molybdenum possess corrosion resistance that makes them useful for acid-handling equipment and for laboratory drain pipes.

Iron-tungsten alloys. Although tungsten is used in several types of relatively complex alloy (including high-speed steels not discussed here), the only commercial alloy made up principally of iron and tungsten was a tungsten steel containing 0.5% chromium in addition to 6% tungsten that was used up to the time of World War I for making permanent magnets.

Hard-facing alloys. Hard-facing consists of welding a layer of metal of special composition on a metal surface to impart some special property not possessed by the original surface. The deposited metal may be more resistant to abrasion, corrosion, heat, or erosion than the metal to which it is applied. A considerable number of hard-facing alloys are available commercially. Many of these would not be considered iron alloys by the 50% iron content criterion adopted for the iron alloys in this article, and they will not be discussed here. Among the iron alloys are low-alloy facing materials containing chromium as the chief alloying element, with smaller amounts of manganese, silicon, molybdenum, vanadium, tungsten, and in some cases nickel to make a total alloy content of up to 12%, with the balance iron. High-alloy ferrous materials containing a total of 12–25% alloying elements form another group of hard-facing alloys; a third group contains 26–50% alloying elements. Chromium, molybdenum, and manganese are the principal alloying elements in the 12–25% group; smaller amounts of molybdenum, vanadium, nickel, and in some cases titanium are present in various proportions. In the 26–50% alloys, chromium (and in some cases tungsten) is the principal alloying element, with manganese, silicon, nickel, molybdenum, vanadium, niobium (columbium), and boron as the elements from which a selection is made to bring the total alloy content within the 26–50% range.

Permanent-magnet alloys. These are magnetically hard ferrous alloys, many of which are too complex to fit the simple compositional classification used above for other iron alloys. As already mentioned in discussing iron-cobalt and iron-tungsten alloys, the high-carbon steels (with or without alloying elements) are now little used for permanent magnets. These have been supplanted by a group of sometimes complex alloys with much higher retentivities. The ones considered here are all proprietary compositions. Two of the alloys contain 12% cobalt and 17% molybdenum and 12% cobalt and 20% molybdenum.

Members of a group of six related alloys contain iron, nickel, aluminum, and with one exception cobalt; in addition, three of the cobalt-containing alloys contain copper and one has copper and titanium. Unlike magnet steels, these alloys resist demagnetization by shock, vibration, or temperature variations. They are used in magnets for speakers, watt-hour meters, magnetrons, torque motors, panel and switchboard instruments, and so on, where constancy and degree of magnet strength are important. *See* ALLOY STRUCTURES; MAGNETISM.

Harold E. McGannon

Bibliography. American Society for Testing and Materials, *Annual Book of ASTM Standards*, 1992, vol. 01.02: *Ferrous Castings; Ferroalloys*, 1992; F. E. Fletcher, *High-Strength Low-Alloy Steels: Status, Selection, and Physical Metallurgy*, 1982; R. A. Oriani et al. (eds.), *Hydrogen Dehydration of Ferrous Alloys*, 1985; Y. S. Touloukian and C. Y. Ho (eds.), *Properties of Selected Ferrous Alloying Elements*, vol. 3, 1989.

Iron metabolism

A nearly continuous cycle whereby iron from organic and inorganic sources is converted to form iron-porphyrin compounds, which can be utilized by the body. One such compound, termed a heme, is hemoglobin; more than 60% of the iron in the body is used in hemoglobin metabolism. Iron is also essential for other heme compounds, such as myoglobin and cytochromes, and for a wide variety of

nonheme enzymes, including many in the citric acid cycle.

Hemoglobin metabolism involves formation and breakdown. Formation is initiated by the process of absorption. Unlike most elements, iron has no specific mechanism for excretion so that absorption must be closely monitored to control body content and ensure replacement of the daily loss. The normal daily requirements of iron are 1 mg for males, 2 mg for young women (to cover the loss from menstruation), and 3 mg for pregnant women. From an average diet, an absorption rate of 5–10% is possible. Absorption occurs best from animal and plant hemes and less from inorganic ferric salts; different mechanisms are involved for the two types. Many factors influence absorption, especially gastric acid, which can solubilize iron salts and prevent their precipitation in the duodenum.

In the duodenum, iron rapidly enters the mucosal cells of the intestinal villi, where the iron is released from heme by the enzyme heme oxygenase. There, the ferrous iron destined for the formation of hemoglobin in the developing red cells of the marrow (the erythroblasts) is converted to ferric ions by ceruloplasmin and is attached to the transport glycoprotein transferrin. The nonassimilated iron remains in the intestinal cell and is combined with the storage protein apoferritin to form ferritin, which is lost by the body when the mucosal cells are shed after their 3–5-day life cycle. The mechanism by which the mucosal cell knows what to discard and what to assimilate is unknown. Absorption is increased by iron deficiency, anemia of many varieties, hypoxia, and increased red cell production (erythropoiesis).

Once ferric iron is attached to transferrin, it circulates in the blood until it attaches to transferrin receptors on immature red blood cells in the marrow. There may be as many as 300,000 receptors on the surface of an individual cell. Once attached, the transferrin receptor complex is taken in by endocytosis, and the iron is released and reduced to ferrous ions. The transferrin receptor complex returns to the cell surface, and the transferrin reenters the blood plasma. The iron enters the mitochondria and is inserted into protoporphyrin by the enzyme ferrochetalase to form heme, which when combined with the protein globin forms the respiratory pigment hemoglobin.

Mature red blood cells cannot take up iron; at the end of their 120-day life span they are engulfed by the monocyte-macrophage cells in liver and spleen, where the iron is released by the enzyme heme oxygenase. Sixty percent of this iron is rapidly returned to the marrow to produce red blood cells, while the remainder is stored as ferritin in the labile pool for release as needed. Iron excessively absorbed from the gut or released from the labile pool and not destined for the formation of red blood cells may enter the storage compartment as ferritin or hemosiderin. Apoferritin, the iron-free storage protein, exists in most living cells, and ferric ions can be stored in its hollow sphere to form a growing crystal of ferric oxyhydroxide (FeOOH). Small amounts of apoferritin enter the circulatory system at levels directly paralleling those of the stores. Hemosiderin occurs in the monocyte macrophages of the liver and spleen as FeOOH stripped of its apoferritin shell. Both storage forms may be rapidly released on demand for increased red blood cell production. Excessive amounts of iron may be stored as hemosiderin following multiple transfusions of blood in non-iron-deficient anemias, in the disease hemochromatosis, and following undue breakdown of cells. *See* BLOOD; HEMOGLOBIN.

John Murray

Bibliography. W. J. Williams et al. (eds.), *Hematology*, 5th ed., 1995.

Iron metallurgy

Extracting iron from ores and preparing it for use. Extraction involves the conversion of naturally occurring iron-bearing minerals into metallic iron. The term ironmaking is commonly used to include all of the industrial processes that convert raw materials into iron. The major process for the production of iron is the iron blast furnace, which varies widely in size and specific features. However, the growth of alternative direct-reduction processes has been very significant (**Fig. 1**). The principal difference between the blast furnace process and the direct-reduction processes is the temperature of operation. In the blast furnace, high operating temperatures enable the production of molten iron. At the lower operating temperatures of the direct-reduction processes, solid or sponge iron is produced. Most of the iron produced in the world is used in the production of steel. The remainder is converted to iron castings, ferroalloys, and iron powder.

An alternative method for ironmaking is direct smelting. These processes rely on the direct reaction of noncoking coal, oxygen, and iron ore in a high-temperature smelter to produce molten iron. In this sense, they fall between the blast furnace and direct reduction (Fig. 1). The primary impetus for these processes is the elimination of coking in the production of iron from coal. *See* FERROALLOY.

Ore preparation. The major ore deposits contain iron in the form of oxides, such as hematite (Fe_2O_3) and magnetite (Fe_3O_4). These oxides are physically separated from the other constituents in the ore by sequences of operations, including crushing and fine grinding, magnetic and flotation concentration, and dewatering. The separation of iron oxide concentrates in this manner is a branch of a technology called mineral beneficiation or ore dressing. The iron oxide particles in the concentrates are almost as fine in size as talcum powder (−200 mesh). The gangue constituents of the ore are also finely divided, and this has caused major problems in the development of environmentally suitable disposal methods. The iron oxide concentrates must be consolidated into hard pellets, sintered lumps, or briquettes to become suitable charge materials for the iron blast furnace and most of the direct-reduction processes. In general, the ore dressing and consolidation operations

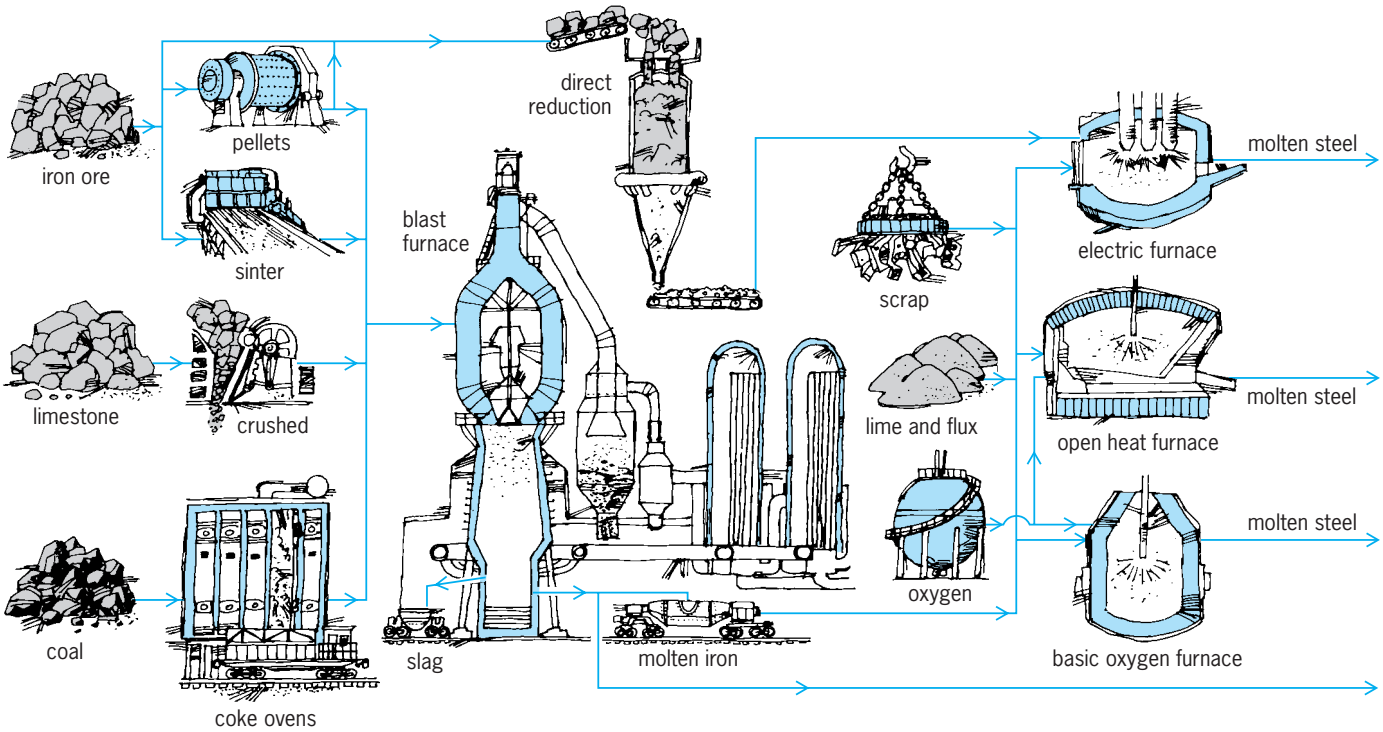

Fig. 1. The position of ironmaking processes (blast furnace and direct reduction) in the steelmaking sequence. (*American Iron and Steel Institute*)

are performed near the mining sites. *See* FLOTATION; ORE DRESSING.

Reduction of oxide concentrates. The conversion of iron oxide to metallic iron is accomplished by the application of heat and reducing agents. In sequence, hematite (Fe_2O_3), magnetite (Fe_3O_4), and wüstite (FeO) are reduced by carbon monoxide (CO) and hydrogen (H_2) as shown in reactions (1)–(3).

$$\tfrac{1}{2}Fe_2O_3 + \tfrac{1}{6}CO \rightarrow \tfrac{1}{3}Fe_3O_4 + \tfrac{1}{6}CO_2 \qquad (1a)$$

$$\tfrac{1}{2}Fe_2O_3 + \tfrac{1}{6}H_2 \rightarrow \tfrac{1}{3}Fe_3O_4 + \tfrac{1}{6}H_2O \qquad (1b)$$

$$\tfrac{1}{3}Fe_3O_4 + \tfrac{1}{3}CO \rightarrow FeO + \tfrac{1}{3}CO_2 \qquad (2a)$$

$$\tfrac{1}{3}Fe_3O_4 + \tfrac{1}{3}H_2 \rightarrow FeO + \tfrac{1}{3}H_2O \qquad (2b)$$

$$FeO + CO \rightarrow Fe + CO_2 \qquad (3a)$$

$$FeO + H_2 \rightarrow Fe + H_2O \qquad (3b)$$

The sum of reactions (1*a*), (2*a*), (3*a*) yields reaction (4), which represents an overall reaction for the

$$\tfrac{1}{2}Fe_2O_3 + \tfrac{3}{2}CO \rightarrow Fe + \tfrac{3}{2}CO_2 \qquad (4)$$

reduction of hematite to metallic iron with carbon monoxide.

The reducing agents, CO and H_2, are derived from coal, natural gas, oil, wood, and other carbonaceous or organic materials. The manner in which the reducing agents are obtained from these raw materials is complex. For example, the iron blast furnace process uses coke (Fig. 1). When coal is heated in coke ovens, a fuel gas is released, leaving a residue of hard carbonaceous material called coke. The released gas is primarily methane (CH_4) with small amounts of other recoverable organic species. After processing, the residual gas is a useful fuel. The hard coke charged at the top of the blast furnace descends slowly to the bosh region (**Fig. 2**), where it is burned with hot air blasted through the tuyeres. In the presence of excess coke the combustion reaction is given by reaction (5). In most blast furnaces, hydrocarbons

$$C + \tfrac{1}{2}O_2 \rightarrow CO \qquad (5)$$

(oil, gas, tar, and so on) are added to the hot blast to provide H_2, which reduces the required amount of coke to produce a unit of iron. For example, where methane is added to the hot blast, hydrogen is generated by reaction (6). In addition, small but significant

$$CH_4 + \tfrac{1}{2}O_2 \rightarrow CO + 2H_2 \qquad (6)$$

amounts of CO and H_2 are regenerated from CO_2 and H_2O at higher levels in the stack by solution loss reactions represented by reactions (7).

$$CO_2 + C \rightarrow 2CO \qquad (7a)$$

$$H_2O + C \rightarrow CO + H_2 \qquad (7b)$$

The sources of the CO_2 and H_2O for these reactions are reactions (1)– (3), which occur simultaneously. As the gases and solids move countercurrently, the descending solid oxides release their oxygen to the ascending gas stream by means of the reduction reactions. The solution loss reactions provide some regeneration of CO and H_2 to maintain the reducing strength of the ascending gases. *See* COKE.

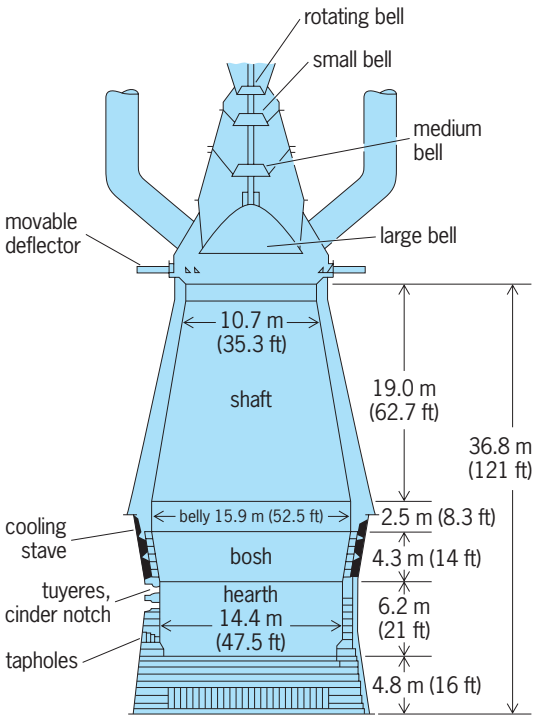

Fig. 2. Iron blast furnace. (*After J. G. Peacey and W. G. Davenport, The Iron Blast Furnace, Pergamon, 1979*)

For the direct-reduction processes, CO and H_2 are produced from the available raw materials in a variety of ways. Several of the rotary kiln processes utilize the direct charging of coal with iron oxide concentrates. The fixed carbon and the volatile matter of the coal are partially burned in place with the iron oxides to generate the reducing gases. Several of the shaft, retort, and fluidized-bed processes utilize the separate generation of reducing gas from natural gas.

In some direct smelting processes, pulverized coal is injected into a molten iron/slag emulsion to provide reductant (fixed carbon and hydrocarbon gases). Simultaneously, iron oxide fines and pellets are fed into the slag, where they dissolve readily. An oxygen lance provides oxidant to burn hydrocarbons and fixed carbon. The overall reaction system in direct smelting is basically the same as that in the blast furnace; however, it is as if the blast furnace were compressed into just the bosh and hearth regions. The process depends upon the fluid dynamic conditions in the iron/slag emulsion, also referred to as a foam.

Material balances. The principal raw materials used in the extraction of iron are oxide concentrates, reducing agents, air, water and steam, auxiliary fuels used directly or converted to electrical energy, and fluxes (**Fig. 3**). The most common flux used in the iron blast furnace is limestone ($CaCO_3$). Large modern furnaces are extremely efficient and values of coke consumption as low as 0.40 kg of coke per kilogram of iron (800 lb/short ton) have been achieved with oil injection. The material balance provides an indication of the relative magnitudes of the major input and output items. Direct-reduction and direct-smelting processes have features in common with the blast furnace; however, the input and output statements may appear to be different to an experienced observer.

The quantity of air required per unit of iron produced in the blast furnace is an impressive figure. At least 1 m^3 of air must be blown into the blast furnace or direct-reduction process for each kilogram of iron produced. The exhaust gases from the blast furnace contain considerable amounts of H_2 and CO.

In modern plants, material balances are monitored by sophisticated computer techniques. It is essential to keep the system operating productively and efficiently. In an integrated steel plant, interruptions in the production of iron from the blast furnace can upset the entire sequence of operations. In this respect, the integrated steel plant is similar to the assembly line for any manufactured product. *See* PRESSURIZED BLAST FURNACE.

Energy balances and temperature profiles. Total energy consumption in steelmaking is much lower than in the production of other common metals (**Table 1**). To some extent the higher thermal efficiency of ironmaking may be attributed to the large scale of iron and steel operations. Another important factor is that recycling and reuse of waste materials and energy sources have been extensively developed (**Table 2**).

Estimated temperature profiles have been established (**Fig. 4**). The temperature of the descending solids rises rapidly to about 1000 K (1340°F) in the top portion of the furnace, rises gradually

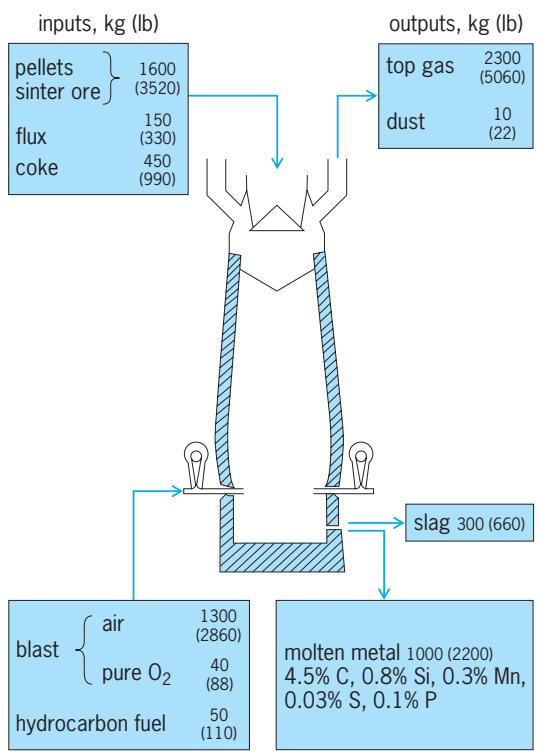

Fig. 3. Simplified material balance for an iron blast furnace. (*After J. G. Peacey and W. G. Davenport, The Iron Blast Furnace, Pergamon, 1979*)

TABLE 1. Approximate energy requirements for the extraction of metals from ores

Metal	Theoretical, 10^2 kJ/kg	Actual, 10^3 kJ/kg
Iron	3.7	19.4
Copper	1.6	54.0
Aluminum	18.3	236.0
Titanium	12.0	501.0

TABLE 2. Approximate simplified energy balance for a blast furnace

Parameter	10^3 kJ/kg
Energy input	
Calorific value of coke	14.87
Sensible heat of hot air blast	1.95
Calorific value of injected fuel	1.12
Total	17.94
Energy output	
Calorific value of recoverable top gas	7.40
Energy consumed	
Heat for reactions, sensible heat of slag and iron, and losses	10.54

to about 1200 K (1700°F) in the central portion, and rises rapidly to temperatures in excess of the fusion point in the lower portion. The radial temperature gradient is very steep at the midheight of the furnace, where the temperature at the vertical axis may be as high as 1900 K (2960°F), while the wall temperature is only 1200 K (1700°F). The residence time of the ascending gases in the furnace is of the order of 1 s, while that of the descending solids is of the order of 10^4 s. Uniformity of gas and solid movement in the furnace is highly desirable. Excessive channeling, erratic flow, and formation of static zones can lead to major losses in efficiency and productivity.

Composition of products. Although meteoric iron containing nickel was used before 4000 B.C., it is believed that iron was not extracted from ore until about 3000 B.C. The earliest extracted form and composition of iron was similar to that produced in modern direct-reduction processes that operate at temperatures below the melting point of iron. This reduced solid iron is called sponge iron because of its porosity. Sponge iron contains significant amounts of trapped carbon and unreduced oxide, and it must be worked severely or melted and refined. Blast furnaces produce liquid iron that is commonly called hot metal or pig iron. Pig iron contains more than 4% carbon and significant amounts of manganese, silicon, and phosphorus. Pig iron must be refined in subsequent steelmaking operations.

Trends. The largest modern furnaces can produce more than 2.2×10^7 lb (10^7 kg) of iron per day, operating with careful charge control, high top pressure, high blast temperature, and fuel injection. It has become common practice to lower the quantity of slag and its basicity. Sulfur control is provided through the addition of a separate process for the desulfurization of hot metal. With the development of advanced sensors and mathematical modeling methods and supercomputers, it has become possible to predict and control the operation of large blast furnaces to an extent far greater than was previously the case. In seconds, analyses of variations can be made, and corrective actions can be taken to maintain high productivity and efficiency. *See* SIMULATION; SUPERCOMPUTER.

The modern iron blast furnace is a highly developed, very efficient, and economical process for large steel plants. However, it requires coke, which must be produced in associated coke ovens, a practice associated with environmental problems; and it is very large and inflexible. Once it becomes operational, the blast furnace goes continuously, and the associated steel plant activities must be matched to take its output. Further, the blast furnace offers little opportunity for incremental increases in steel plant capacity. A major expansion and investment is required for installation of a new blast furnace. A typical large modern integrated steel plant operates with one to four blast furnaces.

Direct-reduction processes have been constructed at many sites throughout the world, particularly in the developing countries. The direct reduction processes offer the alternative of a shippable iron product suitable for steel production in small- and large-scale steel plants. In effect, the direct-reduction iron products act as a substitute for scrap (Fig. 1) with the added advantage that they are cleaner and less contaminated with tramp elements than bundles of commercial scrap. The daily production rates of the direct-reduction processes are less than 2.2×10^6 lb

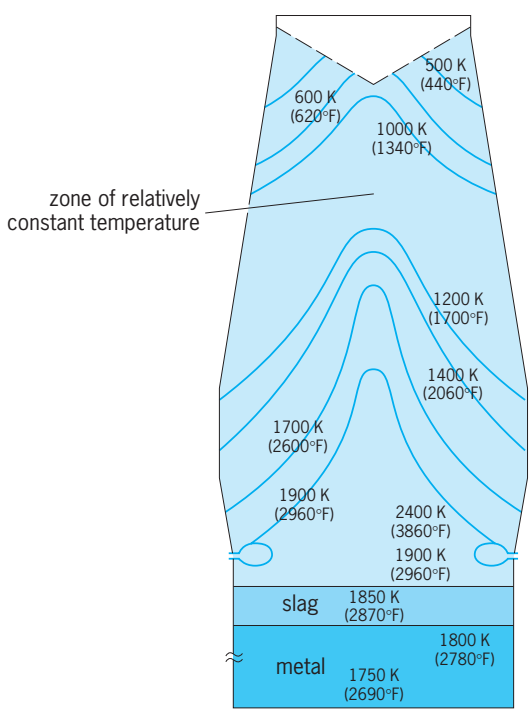

Fig. 4. Temperature profiles in a modern blast furnace. (*After J. G. Peacey and W. G. Davenport, The Iron Blast Furnace, Pergamon, 1979*)

(1×10^6 kg) and often are as low as 4.2×10^5 lb (2×10^5 kg). The direct smelting processes are designed to be flexible and compact with high production intensity (tons of iron produced per unit of working reactor volume) and to start up quickly and adjust production rate to demand. They are adaptable to small-scale steel production facilities. However, molten iron is not shippable and must be transferred to nearby steelmaking furnaces or cast into pigs. *See* CAST IRON; IRON; IRON ALLOYS; PYROMETALLURGY; STEEL MANUFACTURE. George R. St. Pierre

Bibliography. American Institute of Mining, Metallurgical, and Petroleum Engineers (AIME), *Blast Furnace*: *Theory and Practice*, 1969; AIME, *Ironmaking Proceedings*, annually; Gmelin Institute for Inorganic Chemistry of the Max Planck Society for the Advancement of Science, *Metallurgy of Iron*, 4th ed., vols. 1–12, 1992; R. J. Fruehan et al. (eds.), *Making, Shaping, and Treating of Steel*, 11th ed., 1998; W.-K. Lu, A critical review of blast furnace reactions, *Iron Steelmak.*, pp. 51–62, October 1987; T. P. McAloon, Alternate ironmaking update, *Iron Steelmak.*, pp. 37–56, February 1994; J. G. Peacey and W. G. Davenport, *The Iron Blast Furnace*, 1979.

Ironwood

The name given to any of at least 10 kinds of tree in the United States. Because of uncertainty as to just which tree is indicated, the name ironwood has been abandoned in the checklist of the native and naturalized trees of the United States. Probably the best known of the 10 is the American hornbeam (*Carpinus caroliniana*). Some of the others are *Ostrya virginiana*, eastern hophornbeam; *Bumelia lycioides*, buckthorn bumelia; *B. tenax*, tough bumelia; *Cliftonia monophylla*, buckwheat tree; and *Cyrilla racemiflora*, swamp cyrilla or swamp ironwood. All of these species except *Ostrya* are restricted to the southeastern United States. Others commonly called ironwood are *Eugenia confusa*, redberry eugenia of southern Florida and the Florida Keys; *Exothea paniculata*, butterbough of southern Florida; and *Ostrya knowltonii*, Knowlton hophornbeam of southwestern United States. Leadwood (*Krugiodendron ferreum*), a native of southern Florida, has the highest specific gravity of all woods native to the United States and is also known as black ironwood. *See* FOREST AND FORESTRY; HOPHORNBEAM; HORNBEAM; TREE.

Arthur H. Graves; Kenneth P. Davis

Irregularia

The name given by P. A. Latreille in 1825 to what is now recognized as a natural (monophyletic, that is, evolved from a common ancestral form) group of echinoids (sea urchins). Previously, workers such as J. Durham and R. Melville suggested that the group was polyphyletic (evolved from different ancestral lineages), a proposal followed by other influential works. However, this suggestion has been rejected by more recent workers such as M. Jensen, A. Smith, and T. Saucède and colleagues who showed in phylogenetic analyses that all irregular echinoids are derived from eodiadematid-like forms in the Lower Jurassic. Most Irregularia, and all the extant forms, are exocyclic: the periproct (the area surrounding the anus) plus the anus is displaced in the direction of interambulacrum 5 (Lovén's numbering system) to become situated at the edge of or outside the apical system. (The interambulacrum is the area between two ambulacra, which are the radial series of plates along which the tube feet are arranged.) Therefore, these Irregularia exhibit an anterior-posterior axis running down the centers of ambulacrum III and interambulacrum 5, inducing secondary bilateral symmetry. Bilateral symmetry (symmetry along a central axis, with division into equivalent left and right halves) is most pronounced in the spatangoid heart urchins, and in stark contrast to the radial symmetry (symmetry about a central point) of the "regular" urchins. Additional characteristics of the Irregularia include the miniaturization of external appendages [spines, tube feet, pedicellariae (small grasping organs)], the relatively large size of the periproct, the high density of primary tubercles, and in most forms a relatively low test (the internal, limestone skeleton of sea urchins) profile. The Irregularia includes a basal group, the Eodiadematidae, plus holectypoids, pygasteroids, cassiduloid and clypeasteroid neognathostomates, and holasteroid and spatangoid atelostomates. Irregular echinoids left a richer fossil record than the regular urchins, in part because the test of irregulars is less prone to postmortem disintegration and because they tended to live in softer substrata that led to better preservation. Irregular echinoids are largely sediment-swallowing "podial particle pickers" and have developed sophisticated mechanisms for handling fine sediments. *See* ECHINOIDA; ECHINOIDEA; ECHINODERMATA; HOLECTYPOIDA; NEOGNATHOSTOMATA; PYGASTEROIDA. Rich Mooi

Bibliography. J. W. Durham, Phylogeny and evolution, in R. C. Moore (ed.), Part U: *Treatise on Invertebrate Paleontology, Echinodermata-Echinozoa-Echinoidea*, vol. 3, pp. 266–269, 1966; J. W. Durham and R. V. Melville, A classification of echinoids, *J. Paleontol.*, 31:242–272, 1957; M. Jensen, Morphology and classification of the Euechinoidea Bronn, 1860—a cladistic analysis, *Videnskabelige Meddelelser fra Dansk Naturhistorisk Forening i Kjobenhavn*, 143:7–99, 1981; A. B. Smith, *Echinoid Palaeobiology*, Allen & Unwin, London, 1984.

Irrigation (agriculture)

The artificial application of water to the soil to produce plant growth. Irrigation also cools the soil and atmosphere, making the environment favorable for plant growth. The use of some form of irrigation is

TABLE 1. Example of consumption of water by various crops, in inches*

Crop	April	May	June	July	Aug.	Sept.	Oct.	Seasonal total
Alfalfa	3.3	6.7	5.4	7.8	4.2	5.6	4.4	37.4
Beets		1.9	3.3	5.3	6.9	5.8	1.1	24.3
Cotton	1.1	2.0	4.1	5.8	8.6	6.7	2.7	31.0
Peaches	1.0	3.4	6.7	8.4	6.4	3.1	1.1	30.0
Potatoes			0.7	3.4	5.8	4.4		14.0

*1 in. = 25 mm.

well documented throughout the history of civilization.

Use of water by plants. Growing plants use water almost continuously. Growth of crops under irrigation is stimulated by optimum moisture, but retarded by excessive or deficient amounts. Factors influencing the rate of water use by plants include the type of plant and stage of growth, temperature, wind velocity, humidity, sunlight duration and intensity, and available water supply. Plants use the least amount of water upon emergence from the soil and near the end of the growing period. Irrigation and other management practices should be coordinated with the various stages of growth. A vast amount of research has been done on the use of water by plants, and results are available for crops under varying conditions. *See* PLANT GROWTH.

Consumptive use. In planning new or rehabilitating old irrigation projects, consumptive use is the most important factor in determining the amount of water required. It is also used to determine water rights.

Consumptive use, or evapotranspiration, is defined as water entering plant roots to build plant tissues, water retained by the plant, water transpired by leaves into the atmosphere, and water evaporated from plant leaves and from adjacent soil surfaces. Consumptive use of water by various crops under varying conditions has been determined by soil-moisture studies and computed by other well-established methods for many regions of the United States and other countries. Factors which have been shown to influence consumptive use are precipitation, air temperature, humidity, wind movement, the growing season, and latitude, which influences hours of daylight. **Table 1** shows how consumptive use varies at one location during the growing season. When consumptive-use data are computed for an extensive area, such as an irrigation project, the results will be given in acre-feet per acre for each month of the growing season and the entire irrigation period. Peak-use months determine system capacity needs. An acre-foot is the amount of water required to cover 1 acre 1 ft deep (approx. 44 ft^3 or 1214 m^3 of water).

Soil, plant, and water relationships. Soil of root-zone depth is the storage reservoir from which plants obtain moisture to sustain growth. Plants take from the soil not only water but dissolved minerals necessary to build plant cells. How often this reservoir must be filled by irrigation is determined by the storage capacity of the soil, depth of the root zone, water use by the crop, and the amount of depletion allowed before a reduction in yield or quality occurs. **Table 2** shows the approximate amounts of water held by soils of various textures.

Water enters coarse, sandy soils quite readily, but in heavy-textured soils the entry rate is slower. Compaction and surface conditions also affect the rate of entry.

Soil conditions, position of the water table, length of growing season, irrigation frequency, and other factors exert strong influence on root-zone depth. **Table 3** shows typical root-zone depths in well-drained, uniform soils under irrigation. The depth of rooting of annual crops increases during the entire growing period, given a favorable, unrestricted root zone. Plants in deep, uniform soils usually consume water more slowly from the lower root-zone area than from the upper. Thus, the upper portion is the first to be exhausted of moisture. For most crops, the entire root zone should be supplied with moisture when needed.

Maximum production can usually be obtained with most irrigated crops if not more than 50% of the available water in the root zone is exhausted during the critical stages of growth. Many factors influence this safe-removal percentage, including the type of

TABLE 2. Approximate amounts of water in soils available to plants

Soil texture	Water capacity in inches for each foot of depth (in centimeters for each meter of depth)
Coarse sandy soil	0.5–0.75 (4.15–6.25)
Sandy loam	1.25–1.75 (10.40–14.60)
Silt loam	1.75–2.50 (14.60–20.85)
Heavy clay	1.75–2.0 (14.60–16.65) or more

TABLE 3. Approximate effective root-zone depths for various crops

Crop	Root-zone depth, ft (m)
Alfalfa	6 (1.8)
Corn	3 (0.9)
Cotton	4 (1.2)
Potatoes	2 (0.6)
Grasses	2 (0.6)

crop grown and the rate at which water is being removed. Application of irrigation water should not be delayed until plants signal a need for moisture; wilting in the hot parts of the day may reduce crop yields considerably. Determination of the amount of water in the root zone can be done by laboratory methods, which are slow and costly. However, in modern irrigation practice, soil-moisture-sensing devices are used to make rapid determinations directly with enough accuracy for practical use. These devices, placed in selected field locations, permit an operator to schedule periods of water application for best results. Evaporation pans and weather records can be used to estimate plant-water use. Computerizing these data also helps farmers schedule their irrigations. The irrigation system should be designed to supply sufficient water to care for periods of most rapid evapotranspiration. The rate of evapotranspiration may vary from 0 to 0.4 in. per day (10 mm per day) or more. *See* PLANT-WATER RELATIONS.

Water quality. All natural irrigation waters contain salts, but only occasionally are waters too saline for crop production when used properly. When more salt is applied through water and fertilizer than is removed by leaching, a salt buildup can occur. If the salts are mainly calcium and magnesium, the soils become saline, but if salts predominantly are sodium, a sodic condition is possible. These soils are usually found in arid areas, especially in those areas where drainage is poor. Rainfall in humid areas usually carries salts downward to the ground water and eventually to the sea.

Saline soils may reduce yields and can be especially harmful during germination. Some salts are toxic to certain crops, especially when applied by sprinkling and allowed to accumulate on the plants. Salt levels in the soil can be controlled by drainage, by overirrigation, or by maintaining a high moisture level which keeps the salts diluted. *See* PLANTS OF SALINE ENVIRONMENTS.

Sodic soils make tillage and water penetration difficult. Drainage, addition of gypsum or sulfur, and overirrigation usually increase productivity.

Ponding or sprinkling can be used to leach salts. Intermittent application is usually better and, when careful soil-moisture management is practiced, only small amounts of excess irrigation are needed to maintain healthy salt levels.

Diagnoses of both water and soil are necessary for making management decisions. Commercial laboratories and many state universities test both water and soil, and make recommendations.

Methods of application. Water is applied to crops by surface, subsurface, sprinkler, and drip irrigation. Surface irrigation includes furrow and flood methods.

Furrow method. This method is used for row crops (**Fig. 1**). Corrugations or rills are small furrows used on close-growing crops. The flow, carried in furrows, percolates into the soil. Flow to the furrow is usually supplied by siphon tubes, spiles, gated pipe, or valves from buried pipe. Length of furrows and size of stream depend on slope, soil type, and crop; infiltration and erosion must be considered.

Fig. 1. Furrow method of irrigation. Water is supplied by pipes with individual outlets, or by ditches and siphon tubes.

Flood method. Controlled flooding is done with border strips, contour or bench borders, and basins. Border strip irrigation is accomplished by advancing a sheet of water down a long, narrow area between low ridges called borders. Moisture enters the soil as the sheet advances. Strips vary from about 20 to 100 ft (6 to 30 m) in width, depending mainly on slope (both down and across), and amount of water available. The border must be well leveled and the grade uniform; best results are obtained on slopes of 0.5% or less. The flood method is sometimes used on steeper slopes, but maldistribution and erosion make it less effective.

Fig. 2. A side-roll sprinkler system which uses the main supply line (often more than 1000 ft or 300 m long) to carry the sprinkler heads and as the axle for wheels.

Fig. 3. Center-pivot systems are very popular in new irrigation developments.

Bench-border irrigation is sometimes used on moderately gentle, uniform slopes. The border strips, instead of running down the slope, are constructed across it. Since each strip must be level in width, considerable earth moving may be necessary.

Basin irrigation is well adapted to flatlands. It is done by flooding a diked area to a predetermined depth and allowing the water to enter the soil throughout the root zone. Basin irrigation may be utilized for all types of crops, including orchards where soil and topographic conditions permit.

Subirrigation. This type of irrigation is accomplished by raising the water table to the root zone of the crop or by carrying moisture to the root zone by perforated underground pipe. Either method requires special soil conditions for successful operation.

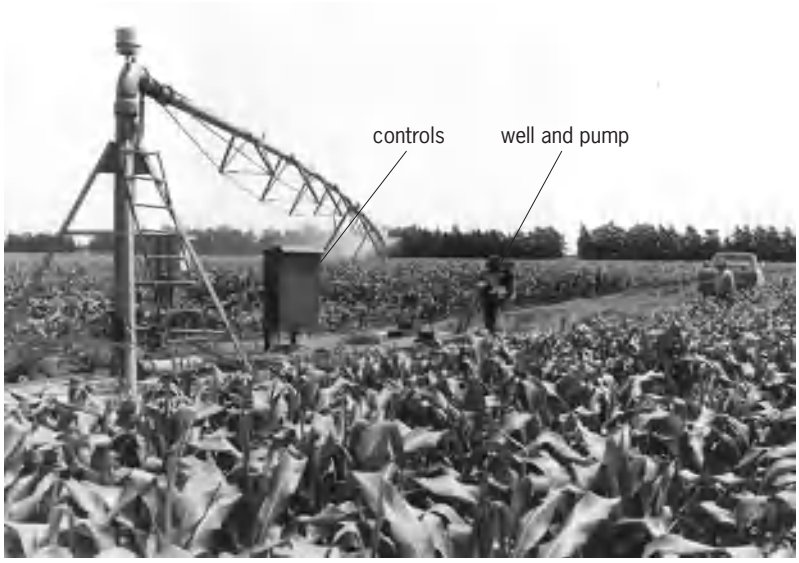

Fig. 4. Center-pivot irrigation system irrigating corn. A single unit can water as much as 500 acres (200 hectares) or more. (*Eastern Iowa Light and Power Cooperative*)

Sprinkler systems. A sprinkler system consists of pipelines which carry water under pressure from a pump or elevated source to lateral lines along which sprinkler heads are spaced at appropriate intervals. Laterals are moved from one location to another by hand or tractor, or they are moved automatically. The side-roll wheel system, which utilizes the lateral as an axle (**Fig. 2**), is very popular as a labor-saving method. The center-pivot sprinkler system (**Fig. 3**) consists of a lateral carrying the sprinkler heads, and is moved by electrical or hydraulic power in a circular course irrigating an area containing up to 135–145 acres (546,200–586,700 m^2).

Extra equipment can be attached in order to irrigate the corners, or solid sets can be used. Solid-set systems are systems with sufficient laterals and sprinklers to irrigate the entire field without being moved. These systems are quite popular for irrigating vegetable crops or other crops requiring light, frequent irrigations and, in orchards, where it is difficult to move the laterals.

Sprinkler irrigation has the advantage of being adaptable to soils too porous for other systems. It can be used on land where soil or topographic conditions are unsuitable for surface methods. It can be used on steep slopes and operates efficiently with a small water supply.

Drip irrigation. This is a method of providing water to plants almost continuously through small-diameter tubes and emitters. It has the advantage of maintaining high moisture levels at relatively low capital costs. It can be used on very steep, sandy, and rocky areas and can utilize saline waters better than most other systems. Clean water, usually filtered, is necessary to prevent blockage of tubes and emitters. The system has been most popular in orchards and vineyards, but is also used for vegetables, ornamentals, and landscape plantings.

Automated systems. Automation is being used with solid-set and continuous-move types of systems, such as the center-pivot (**Fig. 4**) and lateral-move. Surface-irrigated systems are automated with check dams, operated by time clocks or volume meters, which open or close to divert water to other areas. Sprinkler systems, pumps, and check dams can all be activated by radio signals or low-voltage wired systems, which, in turn, can be triggered by soil-moisture-sensing devices or water levels in evaporation pans.

Automatically operated pumpback systems, consisting of a collecting pond and pump, are being used on surface-irrigated farms to better utilize water and prevent silt-laden waters from returning to natural streams.

Multiple uses. With well-designed and -managed irrigation systems, it is possible to apply chemicals and, for short periods of time, to moderate climate. Chemicals which are being used include fertilizers, herbicides, and some fungicides. Effectiveness depends on uniformity of mixing and distribution and on application at the proper times. Chemicals must be registered to be used in this manner.

Solid-set systems are frequently used to prevent

frost damage to plants and trees, since, as water freezes, it releases some heat. A continuous supply of water is needed during the protecting period. However, large volumes of water are required, and ice loads may cause limb breakage. Sequencing of sprinklers for cooling ensures bloom delay in the spring and reduces heat damage in the summer.

Humid and arid regions. The percentage of increase in irrigated land is greater in humid areas than in arid and semiarid areas, although irrigation programs are often more satisfactory where the farmer does not depend on rainfall for crop growth. Good yields are obtained by well-timed irrigation, maintenance of high fertility, keeping the land well cultivated, and using superior crop varieties.

There is little difference in the principles of crop production under irrigation in humid and arid regions. The programming of water application is more difficult in humid areas because natural precipitation cannot be accurately predicted. Most humid areas utilize the sprinkler method.

To be successful, any irrigation system in any location must have careful planning with regard to soil conditions, topography, climate, cropping practices, water quality and supply, as well as engineering requirements. *See* AGRICULTURAL SOIL AND CROP PRACTICES; LAND DRAINAGE (AGRICULTURE); TERRACING (AGRICULTURE); WATER CONSERVATION.

Mel A. Hagood

Bibliography. R. Cuenca, *Irrigation System Design*, 1989; D. R. Hay (ed.), *Planning Now for Irrigation and Drainage in the 21st Century*, 1988; D. Hillel (ed.), *Advances in Irrigation Engineering*, vol. 4, 1987.

Isentropic flow

Compressible flow in which entropy remains constant along streamlines. Generally this implies that entropy is the same everywhere in the flow, in which case the flow is also referred to as homentropic flow. *See* COMPRESSIBLE FLOW; ENTROPY; ISENTROPIC PROCESS.

Because of the second law of thermodynamics, an isentropic flow does not strictly exist. From the definition of entropy, an isentropic flow is both adiabatic and reversible. However, all real flows experience to some extent the irreversible phenomena of friction, thermal conduction, and diffusion. Any nonequilibrium, chemically reacting flow is also irreversible. However, there are a large number of gas dynamic problems with entropy increase negligibly slight, which for the purpose of analysis are assumed to be isentropic. Examples are flow through subsonic and supersonic nozzles, as in wind tunnels and rocket engines; and shock-free flow over a wing, fuselage, or other aerodynamic shape. For these flows, except for the thin boundary-layer region adjacent to the surface where friction and thermal conduction effects can be strong, the outer inviscid flow can be considered isentropic. If shock waves exist in the flow, the entropy increase across these shocks destroys the assumption of isentropic flow, although the flow along streamlines between shocks may be isentropic. *See* ADIABATIC PROCESS; BOUNDARY-LAYER FLOW; NOZZLE; SHOCK WAVE; SUBSONIC FLIGHT; THERMODYNAMIC PRINCIPLES; THERMODYNAMIC PROCESSES.

The assumption of an isentropic flow greatly simplifies the analysis of a flowfield. In many cases, the simple statement of constant entropy replaces a complicated equation such as the momentum or energy equation. In particular, in an isentropic flow of a calorically perfect gas (constant specific heat), the pressure, density, and temperature of the flow are uniquely related to the local Mach number, as shown in Eqs. (1)–(3), where γ is the ratio of specific heat at constant pressure to specific heat at constant volume; p, ρ, and T are the local static pressure, density, and temperature, respectively; M is the local Mach number; and p_0, ρ_0, and T_0 are the total pressure, density, and temperature, respectively. (Total conditions are defined as the conditions that would exist at a point in the flow if the velocity at that point were isentropically reduced to zero.) Equation (3), involving T, holds for any adiabatic flow of a calorically perfect gas; such a flow does not have to be isentropic. Equations (1)–(3) are applicable to isentropic flows of any geometric complexity, including three-dimensional flows. In this manner, isentropic flows are much easier to compute than nonisentropic. *See* FLUID-FLOW PRINCIPLES; GAS DYNAMICS.

$$\frac{p_0}{p} = \left(1 + \frac{\gamma - 1}{2}M^2\right)^{\gamma/(\gamma-1)} \tag{1}$$

$$\frac{\rho_0}{\rho} = \left(1 + \frac{\gamma - 1}{2}M^2\right)^{\gamma/(\gamma-1)} \tag{2}$$

$$\frac{T_0}{T} = 1 + \frac{\gamma - 1}{2}M^2 \tag{3}$$

John D. Anderson, Jr.

Bibliography. J. D. Anderson, Jr., *Fundamentals of Aerodynamics*, 3d ed., 2001; J. D. Anderson, Jr., *Introduction to Flight*, 5th ed., 2005; J. D. Anderson, Jr., *Modern Compressible Flow: With Historical Perspective*, 3d ed., 2003.

Isentropic process

In thermodynamics a change that is accomplished without any increase or decrease of entropy is referred to as isentropic. Since the entropy always increases in a spontaneous process, one must consider reversible or quasistatic processes. During a reversible process the quantity of heat transferred, dQ, is directly proportional to the system's entropy change, dS, as in Eq. (1), where T is the absolute temperature of the system. Systems which are thermally

$$dQ = TdS \tag{1}$$

insulated from their surroundings undergo processes without any heat transfer; such processes are called adiabatic. Thus during an isentropic process there are no dissipative effects and, from Eq. (1), the system neither absorbs nor gives off heat. For this reason the isentropic process is sometimes called the reversible adiabatic process. *See* ADIABATIC PROCESS; ENTROPY; THERMODYNAMIC PROCESSES.

Work done during an isentropic process is produced at the expense of the amount of internal energy stored in the nonflow or closed system. Thus, the useful expansion of a gas is accompanied by a marked decrease in temperature, tangibly demonstrating the decrease of internal energy stored in the system. For ideal gases the isentropic process can be expressed by Eq. (2), where P is the pressure, V is

$$P_1 V_1^{\gamma} = P_2 V_2^{\gamma} = \text{constant} \qquad (2)$$

the volume, and γ is the ratio between the specific heat at constant pressure and the specific heat at constant volume for the given gas. It can be closely approximated by the values of 1.67 and 1.40 for dilute monatomic and diatomic gases, respectively. For a comparison of various processes involving a gas *see* POLYTROPIC PROCESS. Philip E. Bloomfield

Isentropic surfaces

Surfaces along which the entropy and potential temperature of air are constant. Potential temperature, in meteorological usage, is defined by the relationship given below,

$$\Theta = T\left(\frac{1000}{P}\right)^{(c_p - c_v)/c_p}$$

in which T is the air temperature, P is atmospheric pressure expressed in millibars, C_p is the heat capacity of air at constant pressure and C_v is the heat capacity at constant volume. Since the potential temperature of an air parcel does not change if the processes acting on it are adiabatic (no exchange of heat between the parcel and its environment), a surface of constant potential temperature is also a surface of constant entropy. The slope of isentropic surfaces in the atmosphere is of the order of $^1/_{100}$ to $^1/_{1000}$. An advantage of representing meteorological conditions on isentropic surfaces is that there is usually little air motion through such surfaces, since thermodynamic processes in the atmosphere are approximately adiabatic. *See* ADIABATIC PROCESS; ATMOSPHERIC GENERAL CIRCULATION. Frederick Sanders

Bibliography. J. R. Holton, *An Introduction to Dynamic Meteorology*, 4th ed., 2004; J. C. McWilliams, *Fundamentals of Geophysical Fluid Dynamics*, 2006; T. D. Potter and B. R. Colman (eds.), *Handbook of Weather, Climate and Water: Dynamics, Climate, Physical Meteorology, Weather Systems, and Measurements*, 2003.

Ising model

A model which consists of a lattice of "spin" variables with two characteristic properties: (1) each of the spin variables independently takes on either the value +1 or the value −1; and (2) only pairs of nearest neighboring spins can interact. The study of this model (introduced by Ernst Ising in 1925) in two dimensions, where many exact calculations have been carried out explicitly, forms the basis of the modern theory of phase transitions and, more generally, of cooperative phenomena.

The two-dimensional Ising model was shown to have a phase transition by R. E. Peierls in 1936, and the critical temperature or Curie temperature, that is, the temperature at which this phase transition takes place, was calculated by H. A. Kramers and G. H. Wannier and by E. W. Montroll in 1941. Major breakthroughs were accomplished by Lars Onsager in 1944, by Bruria Kaufman and Onsager in 1949, and by Chen Ning Yang in 1952. Onsager first obtained the free energy and showed that the specific heat diverges as $-\ln|1 - T/T_c|$ when the temperature T is near the critical temperature T_c; Kaufman and Onsager computed the short-range order; and Yang calculated the spontaneous magnetization. Since then several other properties have been obtained, and since 1974 connections with relativistic quantum field theory have been made. *See* QUANTUM FIELD THEORY.

Cooperative phenomena. A macroscopic piece of material consists of a large number of atoms, the number being of the order of Avogadro's number (approximately 6×10^{23}). Thermodynamic phenomena all depend on the participation of such a large number of atoms. Even though the fundamental interaction between atoms is short-ranged, the presence of this large number of atoms can, under suitable conditions, lead to an effective interaction between widely separated atoms. Phenomena due to such effective long-range interactions are referred to as cooperative phenomena. The simplest examples of cooperative phenomena are phase transitions. The most familiar phase transition is either the condensation of steam into water or the freezing of water into ice. Only slightly less familiar is the ferromagnetic phase transition that takes place at the Curie temperature, which, for example, is roughly 1043 K for iron.

Of the several models which exhibit a phase transition, the Ising model is the best known. In three dimensions the model is so complicated that no exact computation has ever been made, while in one dimension the Ising model does not undergo a phase transition. However, in two dimensions the Ising model not only has a ferromagnetic phase transition but also has very many physical properties which may be exactly computed. Indeed, despite the restriction on dimensionality, the two-dimensional Ising model exhibits all of the phenomena peculiar to magnetic systems near the Curie temperature. *See* CURIE TEMPERATURE; FERROMAGNETISM; MAGNETISM.

Definition of model. The mutual interaction energy of the pair of spins σ_α and $\sigma_{\alpha'}$ when α and α' are nearest neighbors may be written as $-E(\alpha,\alpha') \cdot \sigma_{\alpha'}$. The meaning of this is that the interaction energy is $-E(\alpha,\alpha')$ when σ_α and $\sigma\alpha'$ are both $+1$ or -1, and is $+E(\alpha,\alpha')$ when $\sigma_\alpha = 1, \sigma\alpha' = -1$, or $\sigma_\alpha = -1$, $\sigma\alpha' = 1$. In addition, a spin may interact with an external magnetic field H with energy $-H\sigma\alpha$. From these two basic interactions the total interaction energy for the square lattice may be written as Eq. (1), where j

$$E = -\sum_j \sum_k [E_1(j,k)\sigma_{j,k}\sigma_{j,k+1} + E_2(j,k)\sigma_{j,k}\sigma_{j+1,k} + H\sigma_{j,k}] \quad (1)$$

specifies the row and k specifies the column of the lattice. In this form the interaction energies $E_1(j,k)$ and $E_2(j,k)$ are allowed to vary arbitrarily throughout the lattice. A special case of great importance is the translationality invariant case (E_1 and E_2 independent of j and k) which was studied by Onsager in 1944. This is the model needed to study a pure ferromagnet without impurities.

Several generalizations of Ising's original model have been considered. For example, σ can be allowed to take on more values than just $\pm$, and interactions other than nearest neighbor can be used. For these generalizations no exact calculations have been performed in two or three dimensions. However, various approximate calculations indicate that the phase transition properties of these models are the same as those of the Onsager lattice.

The extension to the nontranslationally invariant case where $E_1(j, k)$ and $E_2(j, k)$ are treated as independent random variables is important for studying the effects of impurities in ferromagnetics.

Thermodynamic properties. The basic simplification in framing the definition of the Ising model of the preceding section is the choosing of the fundamental variables to be the numbers $\sigma_{j,k}$ which can be only $+1$ or -1. Because of this choice there can be no terms in the interaction energy which refer to kinetic energy or to angular momentum. Consequently, the $\sigma_{j,k}$ do not change with time, and study of the system is, by necessity, confined to those physical properties which depend only on the distribution of energy levels of the system. When the number of energy levels is large, this study requires the use of statistical mechanics.

Statistical mechanics allows the calculation of average macroscopic properties from the microscopic interaction E. If A is some property of the spins σ of the system, then the thermodynamic average of A is given by Eq. (2), where T is the temperature, k is Boltzmann's constant, Z is given by Eq. (3),

$$\langle A \rangle \lim_{N\to\infty} \frac{1}{Z} \sum_{\{\sigma\}} A e^{-E(\sigma)/kT} \quad (2)$$

$$Z = \sum e^{-E(\sigma)/kT} \quad (3)$$

the sums are over all values of $\sigma_{j,k} = \pm 1$, and N is the number of rows and the number of columns. It is mandatory that the thermodynamic limit $N\to\infty$ be taken for these thermodynamic averages to have a precise meaning. *See* BOLTZMANN CONSTANT.

The most important thermodynamic properties of a ferromagnet are the internal energy per site $u = \langle E/N^2 \rangle$ the specific heat $c = \partial u/\partial T$, the magnetization per site $M = \langle \sigma \rangle$ and the magnetic susceptibility $\chi = \partial M \partial H$. These quantities have been computed for the two-dimensional Ising model $E_1 \neq E_2$, but for convenience this discussion is restricted to $E_1 = E_2 = E$.

Onsager studied the two-dimensional square lattice at $H = O$ and computed the specific heat exactly. From that calculation he found that the specific heat was infinite at the critical temperature of Kramers and Wannier given as the solution of Eq. (4). When

$$\sinh \frac{2E}{kT} = 1 \quad (4)$$

T is close to the critical temperature T_c, the specific heat is approximated by Eq. (5). The behavior of

$$c \sim -\frac{8E^2}{kT_c{}^2\pi} \ln \left| 1 - \frac{T}{T_c} \right| \quad (5)$$

the specific heat for any temperature is plotted in **Fig. 1**.

The spontaneous magnetization isdefined as

$$M(0) = \lim_{H\to 0^+} M(H)$$

For $T > T_c$, $M(0) = 0$. For $T < T_c$, Yang found that $M(0)$ is given by Eq. (6). When T is near T_c, $M(0)$ is approximated by Eq. (7). The behavior $M(0)$ as a function of T is plotted in **Fig. 2**.

$$M(0) = \left[1 - \left(\sinh \frac{2E}{kT} \right)^{-4} \right]^{1/8} \quad (6)$$

$$M(0) \sim \left[\frac{8\sqrt{2}}{kT_c} \left(1 - \frac{T}{T_c} \right) \right]^{1/8} \quad (7)$$

The magnetic susceptibility χ at $H = 0$ is much

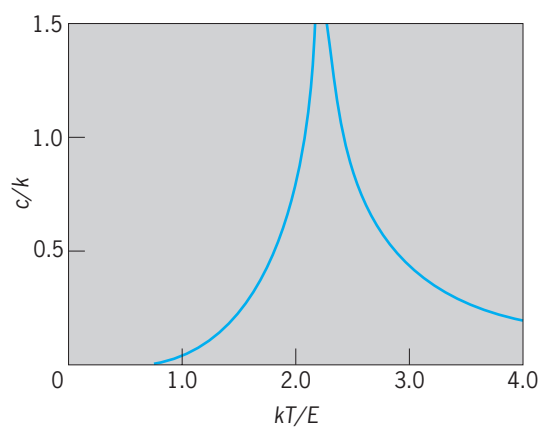

Fig. 1. Specific heat of Onsager's lattice for $E_2/E_1 = 1$.

more difficult to compute than either the specific heat or the spontaneous magnetization. Indeed, no closed form expression is known for χ over the entire range of temperature. However, near T_c it is known that as $T \to T_c^+$, χ is approximated by Eq. (8), and, as $T \to T_c^-$, is approximated by Eq. (9).

$$\chi(T) \sim C_0^+ \left|1 - \frac{T_c}{T}\right|^{-7/4} + C_1^+ \left|1 - \frac{T_c}{T}\right|^{-3/4} + C_2 \quad (8)$$

$$\chi(T) \sim C_0^- \left|1 - \frac{T_c}{T}\right|^{-7/4} + C_1^- \left|1 - \frac{T_c}{T}\right|^{-3/4} + C_2 \quad (9)$$

where

$$C_0^+ = 0.96258\,17322\cdots$$
$$C_0^- = 0.02553\,69719\cdots$$
$$C_1^+ = 0.07498\,81538\cdots$$
$$C_1^- = -0.0198\,94107\cdots$$

See INTERNAL ENERGY; MAGNETIC SUSCEPTIBILITY; MAGNETIZATION; SPECIFIC HEAT OF SOLIDS; STATISTICAL MECHANICS; THERMODYNAMIC PRINCIPLES.

Random impurities. A question which can be very usefully studied in the Ising model is the generalization of statistical mechanics to deal with the experimental situation in which the interaction energy of the system is not completely known because of the presence of impurities. The term "impurity" refers not only to the presence of foreign material in a sample but to any physical property such as defects or isotopic composition which makes lattice sites different from one another. The distribution of these impurities is governed by spin-independent forces. At least two different situations can be distinguished. (1) As the temperature changes, the distribution of impurities may change; such a situation will occur, for example, near the melting point of a lattice. (2) The distribution of impurities may be independent of temperature, at least on the time scale of laboratory measurements; such a distribution will obtain when the temperature of a lattice is well below the melting temperature. Impurities of this sort are said to be frozen in. *See* CRYSTAL DEFECTS.

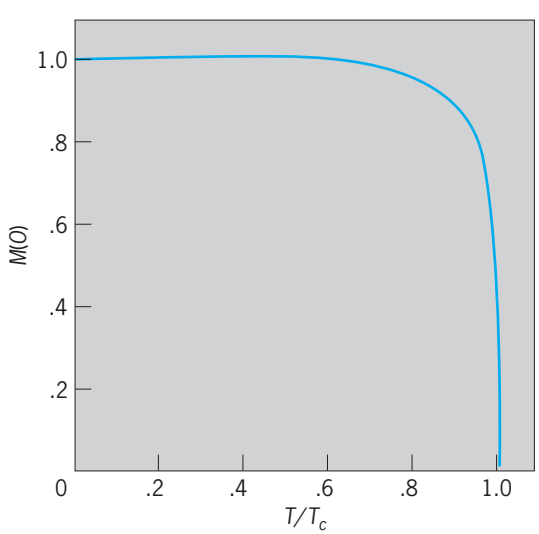

Fig. 2. Spontaneous magnetization $M(0)$ of Onsager's lattice for $E_2/E_1 = 1$.

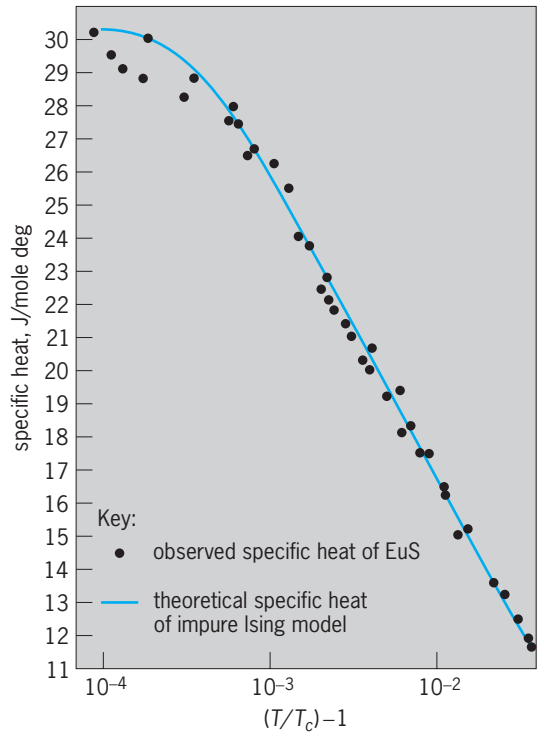

Fig. 3. Comparison of the impure Ising model specific heat with the observed specific heat of EuS for $T > T_c$.

For a study of frozen-in impurities to be realistic, the impurities must be distributed at random throughout the lattice. The translational invariance of the system is now totally destroyed, and it is not at all clear that the phase transition behavior of the pure and random system should be related to each other at all.

These problems were studied for a special case of the Ising model by McCoy and Wu in 1968. They let $E_2(j, k)$ depend on j but not k and kept $E_1(j, k)$ independent of both j and k. Then the variables $E_2(j)$ were chosen with a probability distribution $P(E_2)$. When $P(E_2)$ was of narrow width, they showed that logarithmic divergence of Onsager's specific heat is smoothed out into an infinitely differentiable essential singularity. Such a smoothing out of sharp phase transition behavior may in fact have been observed. The results of one such experiment, carried out by B. J. C. van der Hoeven and colleagues, are compared with the results of the Ising model random specific heat calculation in **Fig. 3**.

Barry M. McCoy; Tai Tsun Wu

Bibliography. C. Domb and M. Green, *Phase Transitions and Critical Phenomena*, 1972; R. Liebmann, *Statistical Mechanics of Periodic Frustrated Ising Systems*, 1986; B. McCoy and T. T. Wu, *The Two Dimensional Ising Model*, 1973.

Island biogeography

The distribution of plants and animals on islands. Islands harbor the greatest number of endemic species. The relative isolation of many islands has allowed populations to evolve in the absence of competitors and predators, leading to the evolution of unique species that can differ dramatically from their mainland ancestors.

Dispersal of species. Plant species produce seeds, spores, and fruits that are carried by wind or water currents, or by the feet, feathers, and digestive tracts of birds and other animals. The dispersal of animal species is more improbable, but animals can also be carried long distances by wind and water currents, or rafted on vegetation and oceanic debris. Long-distance dispersal acts as a selective filter that determines the initial composition of an island community. For example, land snails of many tropical Pacific islands are dominated by small-bodied species that have probably been carried by wind dispersal. Many species of continental origin may never reach islands unless humans accidentally or deliberately introduce them. Consequently, although islands harbor the greatest number of unique species, the density of species on islands (number of species per area) is typically lower than the density of species in mainland areas of comparable habitat. *See* POPULATION DISPERSAL.

Unique morphological traits. Once a species reaches an island and establishes a viable population, it may undergo evolutionary change because of genetic drift, climatic differences between the mainland and the island, or the absence of predators and competitors from the mainland. Consequently, body size, coloration, and morphology of island species often evolve rapidly, producing forms unlike any related species elsewhere. Examples include the giant land tortoises of the Galápagos, and the Komodo dragon, a species of monitor lizard from Indonesia. *See* POLYMORPHISM (GENETICS); POPULATION GENETICS; SQUAMATA.

If enough morphological change occurs, the island population becomes reproductively isolated from its mainland ancestor, and it is recognized as a unique species. Because long-distance dispersal is relatively infrequent, repeated speciation may occur as populations of the same species successively colonize an island and differentiate. The most celebrated example is Darwin's finches, a group of related species that inhabit the Galápagos Islands and were derived from South American ancestors. The island species have evolved different body and bill sizes, and in some cases occupy unique ecological niches that are normally filled by mainland bird species such as warblers and woodpeckers. The morphology of these finches was first studied by Charles Darwin and constituted important evidence for his theory of natural selection. *See* ANIMAL EVOLUTION; SPECIATION.

Theory. For most groups of organisms, the number of species increases in a nonlinear fashion with island area. Mathematically, the relationship can be described by a power function, $S = kA^z$, where S is the number of species on the island, A is the island area, and k and z are fitted constants. On a double logarithmic plot of species versus area, the relationship is linear, with the constant z describing the slope of the line. Although the slope of the species-area curve is steeper for isolated archipelagoes, the number of species per unit area tends to decrease with increasing isolation.

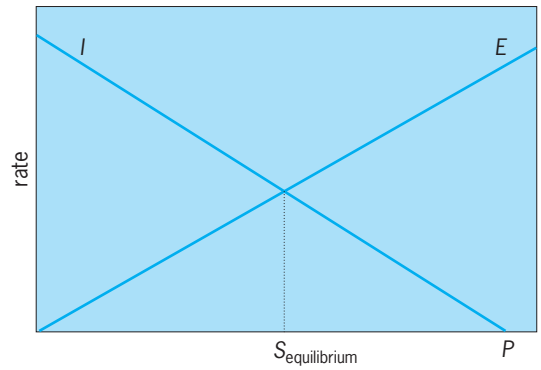

Equilibrium number of species ($S_{equilibrium}$) is determined by the intersection of the immigration (*I*) and extinction (*E*) rate curves. *P* is the maximum number of species in the source pool.

Although descriptions of the effects of area and isolation on species richness began about a century ago, only since the late 1960s have these observations been incorporated into a quantitative model. The MacArthur-Wilson equilibrium theory of island biogeography proposes that the number of species on an island represents an ecological balance between ongoing colonization and extinction of populations (see **illus.**). The model assumes that an island receives its colonists from a permanent source pool (P) of mainland species. The colonization rate (number of new species per unit time) is at a maximum value (I) when the island is empty, and the extinction rate (number of established island populations going extinct per unit time) is at a maximum value (E) when all P species are present on the island. The equilibrium number of species is determined by the intersection of the colonization and extinction curves, which occurs at the point $S_{\text{equilibrium}} = PI/(I + E)$. The model assumes that population size is proportional to island area, and that the risk of population extinction increases at small population size. Consequently, the extinction curve for small islands is steeper than for large islands, producing a trend toward more species at equilibrium as island size increases. The model also assumes that the immigration curve is shallower for more isolated islands, reducing the equilibrium number of species.

The key prediction of the model is that island populations are transient and that colonizations and extinctions are frequent. The most successful test of the equilibrium theory was carried out on tiny mangrove islands in the Florida Keys that were experimentally defaunated and then recolonized by insects over a one-year period. The experiment confirmed

many predictions of the equilibrium model, although much of the turnover may have been from transient species that had not established viable populations.

Nature reserves. The equilibrium theory of island biogeography has been used as a tool for guiding conservation decisions. Its principal prescription is that large, contiguous areas will conserve the maximum number of species. However, specific application of the model has been problematic because many plant and animal communities do not exhibit the turnover predicted by the model. Moreover, the model predicts only the number of species, not their identities. Finally, the model assumes an intact mainland fauna, which may no longer exist in fragmented landscapes. *See* LANDSCAPE ECOLOGY; ZOOGEOGRAPHY.

Advances. Island biogeography theory has recently been extended to describe the persistence of single-species metapopulations. A metapopulation is a set of connected local populations in a fragmented landscape that does not include a persistent source pool region. Instead, the fragments themselves serve as stepping stones for local colonization and extinction. The most successful application of the metapopulation model has been to spotted owl populations of old-growth forest fragments in the northwestern United States. The model prescribes the minimum number, area, and spatial arrangement of forest fragments that are necessary to ensure the long-term persistence of the spotted owl metapopulation. *See* BIOGEOGRAPHY; ECOLOGICAL COMMUNITIES; ECOSYSTEM.

Nicholas J. Gottelli

Bibliography. T. J. Case and M. L. Cody, Testing theories of island biogeography, *Amer. Sci.*, 75:402–411, 1987; R. H. MacArthur and E. O. Wilson, *The Theory of Island Biogeography*, Princeton University Press, Princeton, NJ, 1969; K. A. McGuinness, Equations and explanations in the study of species-area curves, *Biol. Rev.*, 59:423–440, 1984; D. Quammen, *The Song of the Dodo: Island Biogeography in an Age of Extinctions*, Scribner, New York, 1996; M. Williamson, *Island populations*, Oxford University Press, Oxford, 1981.

Isoantigen

An immunologically active protein or polysaccharide present in some but not all individuals in a particular species. These substances initiate the formation of antibodies when introduced into other individuals of the species that genetically lack the isoantigen. Like all antigens, they are also active in stimulating antibody production in heterologous species. The ABO, MN, and Rh blood factors in humans constitute important examples; thus, elaborate precautions for typing are required in blood transfusions. *See* BLOOD GROUPS.

Analogous situations exist for the bloods of most other animal species. Isoantigens are also believed responsible for the ultimate failure of tissue grafts between individuals of the same species, except that of the same genetic constitution or those that have been rendered tolerant. *See* ACQUIRED IMMUNOLOGICAL TOLERANCE; TRANSPLANTATION BIOLOGY.

Isoantigens are to be distinguished from autoantigens, which are antigens active even in the species from which they are derived and in individuals who already possess the antigen. Brain and lens tissue, as well as sperm, constitute examples. These exceptions to the usual rule of nonantigenicity for self-constituents may be more apparent than real, however, since the substances cited are all protected to some degree from contact with the blood, and thus normally do not reach the sites of antibody formation except after experimental manipulation.

Autoantibodies may also be produced in various disease states, perhaps as a result of modification of normal host tissue by the infecting microorganism or by altered host metabolism. Examples are the paroxysmal hemoglobinuria observed in syphilis, acquired hemolytic anemia, or some of the manifestations in rheumatic fever. *See* ANTIBODY; ANTIGEN; AUTOIMMUNITY; POLYSACCHARIDE; PROTEIN.

Margaret J. Polley

Bibliography. E. L. Cooper et al. (eds.), *Developmental and Comparative Immunology*, 1987; D. Male et al., *Advanced Immunology*, 1987; S. B. Mizel and P. Jaret, *The Human Immune System: The New Frontier in Medicine*, 1986; P. N. Plowman, *Hematology and Immunology*, 1987.

Isobar (meteorology)

A curve along which pressure is constant. Leading examples of its uses are in weather forecasting and meteorology. The most common weather maps are charts of weather conditions at the Earth's surface and mean sea level, and they contain isobars as principal information. Areas of bad or unsettled weather are readily defined by roughly circular isobars around low-pressure centers at mean sea level. Likewise, closed isobars around high-pressure centers define areas of generally fair weather (**Fig. 1**). *See* AIR PRESSURE.

Geostrophic wind. A principal use of isobars stems from the so-called geostrophic wind, which approximates the actual wind on a large scale. The direction of the geostrophic wind is parallel to the isobars, in the sense that if an observer stands facing away from the wind, higher pressures are to the person's right if in the Northern Hemisphere and to the left if in the Southern. Thus, in the Northern Hemisphere, flow is counterclockwise about low-pressure centers and clockwise about high-pressure centers, with the direction of the flow reversed in the Southern Hemisphere.

The speed of the geostrophic wind is inversely proportional to distance between isobars drawn at regular close intervals. More precisely, the geostrophic wind is as in Eq. (1).

$$V = \frac{1}{f\rho}\frac{\partial p}{\partial n} \tag{1}$$

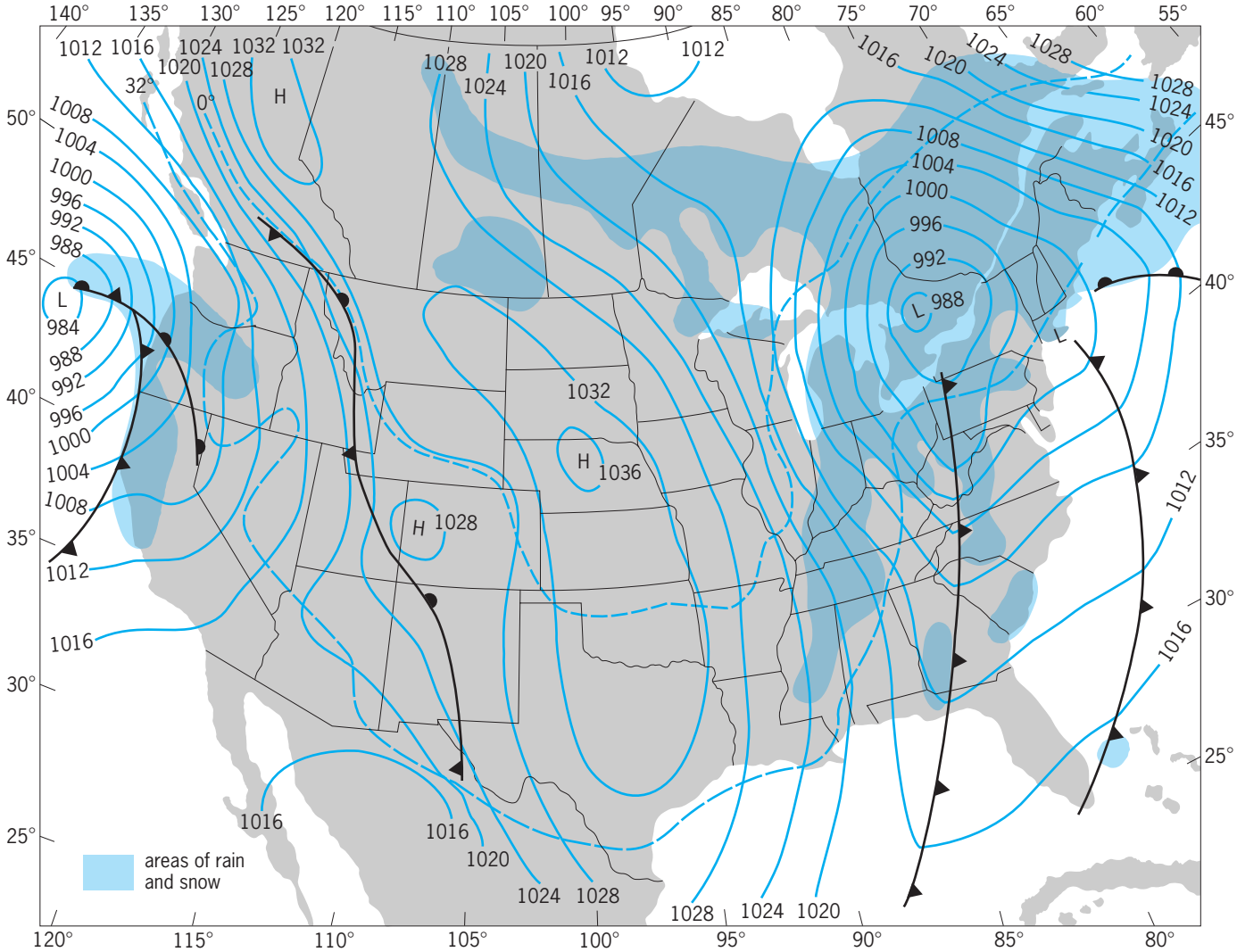

Fig. 1. Weather map at mean sea level for 7:00 A.M., Eastern Standard Time, January 14, 1979. Isobars are labeled in millibars. 1 mb = 10^2 Pa. (*After U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Environmental Data and Information Service, Daily Weather Maps, Weekly Series, January 8–14, 1979*)

V = geostrophic wind speed
$f = 2\omega \sin \phi$, and is called the Coriolis parameter
$\omega = 7.9292 \times 10^{-5}$ radian s^{-1}, the angular velocity of the Earth
ϕ = latitude, positive in the Northern Hemisphere, negative in the Southern
ρ = density of the air, varying about a value of of 1.2×10^{-3} g cm^{-3} at mean sea level
p = pressure, varying about a value of 1013 millibars (101.3 kPa) at mean sea level
n = distance normal to isobar and toward higher pressure

Note that f is positive in the Northern Hemisphere and negative in the Southern. The geostrophic wind law written this way, then, provides a single rule for both Northern and Southern hemispheres, if it is understood that a change in sign of V implies reversal in direction.

When the geostrophic wind is calculated from isobars on a weather map, Eq. (1) is approximated by Eq. (2) [**Fig. 2**], where Δp is the interval of pressure

$$V = \frac{1}{f\rho}\frac{\Delta p}{d} \tag{2}$$

at which the isobars are drawn, and d is the distance between two adjacent isobars. Thus, if isobars are drawn at 4-millibar intervals, then $\Delta p = 4000$ g cm^{-1} s^{-2}; and if $d = 250$ km $= 2.5 \times 10^7$ cm, $f = 10^{-4}$ radian s^{-1} ($\phi = 43.3°$N), and $\rho = 1.2 \times 10^{-3}$ g cm^{-3}, then

$$V = \frac{4000}{10^{-4} \times 1.2 \times 10^{-3} \times 2.5 \times 10^7} \text{ cm s}^{-1}$$
$$= 1333 \text{ cm s}^{-1} = 13 \text{ m s}^{-1} = 25 \text{ knots}$$

Although the geostrophic wind is useful everywhere as an approximation to the wind (except in low latitudes where f approaches zero), at mean sea level it is only a very rough approximation. The departure of the actual wind at mean sea level from

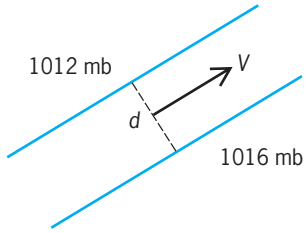

Fig. 2. Schematic of two isobars illustrating the elements in Eq. (2).

the geostrophic is due principally to friction and turbulence, which are highly variable; and also to the impossibility of precisely reducing pressure to mean sea level from plateaus such as the Rocky Mountains and Himalayas. Thus, although the speed of the geostrophic wind varies with air density, at mean sea level it is hardly worthwhile to take into account the variations of air density. A constant, such as $\rho = 1.2 \times 10^{-3}$ g cm^{-3}, is often used instead.

In the free atmosphere, however, where modern aircraft fly, the geostrophic wind is quite accurate as an approximation to the actual wind. Above 1 km (0.6 mi) or so, turbulence and frictional forces are very small, and at even greater heights, above Earth's plateaus, the pressure reduction problem vanishes. In the upper atmosphere, therefore, it is worthwhile to take into account the 15% or so variation of density in the horizontal. *See* GEOSTROPHIC WIND.

Isobaric surfaces. In practice, density variations in the free atmosphere are rather neatly taken into account by constructing weather charts on isobaric surfaces (surfaces of constant pressure), rather than on level surfaces (surfaces of constant height).

Isobars are meaningless on isobaric surfaces, since pressure is constant along any arbitrary curve on such a surface. Instead, curves of constant height are used. These are called isohyets or, more commonly, height contours. The geostrophic wind on isobaric surfaces is as in Eq. (3), where z is height, g is force

$$V = \frac{g}{f}\frac{\partial z}{\partial n} \tag{3}$$

of gravity, and n is distance normal to the contour in the direction of increasing height. *See* WEATHER MAP.

Frederick G. Shuman

Isobar (nuclear physics)

One of two or more atoms which have a common mass number A but which differ in atomic number Z. Thus, although isobars possess approximately equal masses, they differ in chemical properties; they are atoms of different elements. Isobars whose atomic numbers differ by unity cannot both be stable; one will inevitably decay into the other by β^- emission ($Z \rightarrow Z + 1$), β^+ emission ($Z \rightarrow Z - 1$), or electron capture ($Z \rightarrow Z - 1$). There are many examples of stable isobaric pairs, such as ^{50}Ti ($Z = 24$) and ^{50}Cr ($Z = 26$), and four examples of stable isobaric triplets. At most values of A the number of known radioactive isobars exceeds the number of stable ones. *See* ELECTRON CAPTURE; RADIOACTIVITY.

Henry E. Duckworth

Isobaric process

A thermodynamic process during which the pressure remains constant. When heat is transferred to or from a gaseous system, a volume change occurs at constant pressure. This thermodynamic process can be illustrated by the expansion of a substance when it is heated. The system is then capable of doing an amount of work on its surroundings. The maximum work is done when the external pressure P_{ext} of the surroundings on the system is equal to P, the pressure of the system. If V is the volume of the system, the work performed as the system moves from state 1 to 2 during an isobaric thermodynamic process, W_{12}, is the maximum work as given by Eq. (1). For

$$W_{12} = \int_1^2 P\,dV = P\int_1^2 dV = P(V_2 - V_1) \tag{1}$$

an ideal gas the volume increase is a result of a corresponding temperature increase, so that Eq. (1) yields Eq. (2), where n is the number of moles of gas and

$$W_{12} = nR(T_2 - T_1) \tag{2}$$

R is the gas constant. Then W_{12} represents the work done by the system and is positive if $V_2 > V_1$ (that is, $T_2 > T_1$ for an ideal gas). *See* HEAT TRANSFER.

By the first law of thermodynamics, the change of the internal energy, U, in any process is equal to the difference between the heat gained, Q_p, and the work done by the system, W_{12}; thus, Eq. (3) holds. From Eqs. (1) and (3), it follows that Eq. (4) holds for isobaric processes.

$$U_2 - U_1 = U_{12} = Q_p - W_{12} \tag{3}$$

$$Q_p = (U_2 + PV_2) - (U_1 + PV_1) \tag{4}$$

For isobaric processes, it is useful to introduce the enthalpy, H, which is given by the sum of the internal energy U and PV. Then Eq. (5) can be formulated to

$$Q_p = H_2 - H_1 = \int_1^2 C_p\,dT \tag{5}$$

represent Q_p, the transferred heat at constant pressure, where $C_p = (dQ/dT)|_p$ is the heat capacity at constant pressure. For an ideal gas, Eq. (6) holds.

$$Q_p = C_p(T_2 - T_1) \tag{6}$$

See ENTHALPY.

If the isobaric process is also a reversible process, Eq. (7) is obtained, where S is the entropy.

$$Q_p = \int_1^2 T\,dS \tag{7}$$

See ENTROPY.

For an ideal gas the internal energy change can be expressed in terms of the heat capacity at constant

volume, as in Eq. (8). The combination of Eqs. (2), (3), (6), and (8) yields Eq. (9).

$$U_{12} = C_V(T_2 - T_1) \tag{8}$$

$$C_p = C_V + nR \tag{9}$$

From Eq. (8) it follows that, for an isometric process (fixed volume), heat goes into internal energy only. However, from Eqs. (9) and (3) it follows that, for an isobaric process, $C_p > C_V$, and the heat input is converted into internal energy increase as well as work output by the system. *See* ADIABATIC PROCESS; ISOMETRIC PROCESS; THERMODYNAMIC PROCESSES.

For a comparison of the isobaric process with other processes involving a gas *See* POLYTROPIC PROCESS.

Philip E. Bloomfield

Bibliography. D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, 6th ed., 2000; J. R. Howell and R. O. Buckius, *Fundamentals of Engineering Thermodynamics*, 2d ed., 1992; G. J. Van Wylen, R. E. Sonntag, and C. Borgnakke, *Fundamentals of Classical Thermodynamics*, 4th ed., 1994; M. W. Zemansky, M. M. Abbott, and H. C. Van Ness, *Basic Engineering Thermodynamics*, 2d ed., 1975.

Isobryales

An order of the true mosses (subclass Bryidae). This order is difficult to define precisely, but includes plants that generally grow from a creeping primary stem with leaves reduced or essentially lacking and plants that have spreading to ascending secondary stems which may be pinnately branched (see **illus.**). Paraphyllia and pseudoparaphyllia are sometimes present on the stems. The leaves may have single or double and sometimes short costae. The cells may be short or elongate and smooth or papillose, with those at basal angles sometimes differentiated. The sporophytes are lateral, usually with elongate setae and capsules. The double peristome, sometimes reduced, consists of 16 teeth which are papillose on the outer surface, or less often cross-striate at the base, and an endostome with narrow segments and a low basal membrane or none at all. The calyptrae are cucullate and naked, or mitrate and hairy

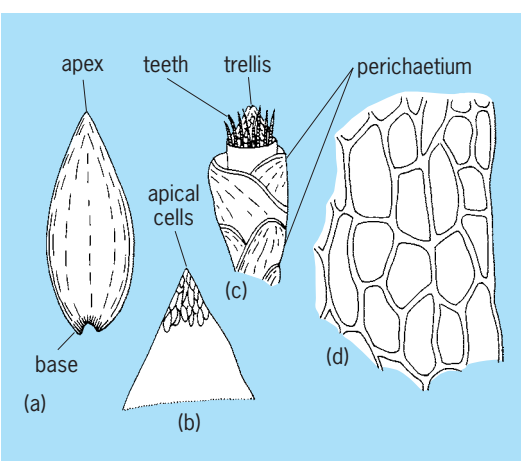

***Fontinalis duriaei*, an example of the order Isobryales. (*a*) Leaf. (*b*) Apex of leaf. (*c*) Trellis and teeth. (*d*) Alar cells. (*After W. H. Welch, Mosses of Indiana, Indiana Department of Conservation, 1957*)**

The order is surely heterogeneous in composition. It consists of about 19 families and 124 genera, some of which may be better placed in the Hypnales. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA; HYPNALES.

Howard Crum

Isocyanate

A derivative of isocyanic acid. Isocyanates are represented by the general formula —N═C═O, where R is predominantly alkyl or aryl; however, stable isocyanates in which the N═C═O group is linked to elements such as sulfur, silicon, phosphorus, nitrogen, or the halogens have also been prepared. Most members of this class of compounds are liquids that are sensitive to hydrolysis and are strong lacrimators. Isocyanates are extremely reactive, especially toward substrates containing active hydrogen. They have found wide use in the commercial manufacture of polyurethanes, which are used in making rigid and flexible foam, elastomers, coatings, and adhesives.

Preparation. Most isocyanates are prepared by reacting excess phosgene ($COCl_2$) with primary amines in high-boiling-point solvents, as in reaction (1). The reaction involves several steps: the

$$\mathrm{R{-}NH_2 + COCl_2 \xrightarrow{-HCL} \underset{\textbf{(1)}}{R{-}NHCOCl} \xrightarrow{\Delta} R{-}N{=}C{=}O + HCl} \tag{1}$$

first, leading to carbamoylchloride (**1**), is exothermic, while the hydrogen chloride (HCl) abstraction to form the isocyanates requires heat. The excess of phosgene prevents or suppresses the formation of side products, such as ureas. The industrial manufacture of isocyanates uses continuous processes in which streams of amine and phosgene, optionally both in solvents, are mixed and the resultant reaction mixtures heated, while the by-product hydrogen chloride is vented off. Products are typically purified by (vacuum) distillation. Other methods of preparation, such as the reaction of salts of isocyanic acid with alkyl halides or the thermal or photolytic degradation of carbonylazides, are less common and are used only in special cases.

The direct conversion of aromatic amines or nitro compounds with carbon monoxide (CO) in the presence of platinum metal catalysts [reaction (2)]

$$\mathrm{R{-}NO_2 + 3CO \longrightarrow R{-}N{=}C{=}O + 2CO_2} \tag{2}$$

has been explored as a viable alternative for the industrial manufacture of isocyanate in place of the

phosgenation technology, which is less desirable in terms of the effect on the environment. However, technical problems have so far prevented the commercialization of any of the various proposed routes.

Characterization. The heterocumulene system N═C═O absorbs strongly in the infrared region at approximately 2250–2275 cm^{-1}; this band is very characteristic and is widely used for the spectroscopic identification of isocyanates. Isocyanates react readily with alcohols to form urethanes (also known as carbamates), which are used for their characterization. *See* INFRARED SPECTROSCOPY.

Reactions. Isocyanates are very reactive compounds. They have been used extensively as building blocks of nitrogen heterocycles, especially in the synthesis of fused-ring systems. Typically, the C═N double bond of the isocyanate group reacts readily with active hydrogen-containing substrates, such as amines, alcohols, or water. Alcohols react with isocyanates to form urethanes [reaction (3)] while pri-

$$R{-}N{=}C{=}O + R'OH \longrightarrow R{-}NACOOR' \qquad (3)$$

mary amines give *N*,*N*′-disubstituted ureas. In the latter reaction, water adds to give initially *N*-substituted carbamic acids, which undergo decarboxylation at room temperature to re-form amine and carbon dioxide [reaction(4)]. Most often, excess isocyanate con-

$$(R{-}N{=}C{=}O + H_2O \longrightarrow$$
$$[R{-}NHCOOH] \longrightarrow R{-}NH_2 + CO_2 \qquad (4)$$

sumes the liberated amine to yield symmetric disubstituted ureas [reaction (5)].

$$RNH_2 + R{-}N{=}C{=}O \longrightarrow RNH{-}CONHR \qquad (5)$$

See AMINE; UREA.

Isocyanates readily undergo cycloaddition reactions with other compounds having activated double bonds, such as C═C, C═N, or C═O. The simplest such adducts are 2 + 2 cycloadducts, in which four-membered ring heterocycles are formed from an enamine [reaction (6)], a compound in which amino

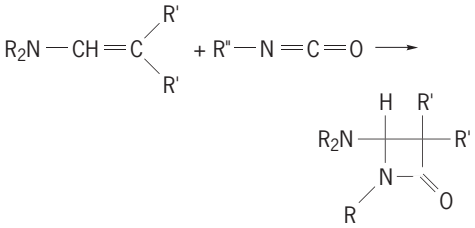

(6)

groups are directly linked to carbon-to-carbon double bonds.

More common, however, are adducts formed from one double-bond component and two molecules of isocyanate, or vice versa. The reactions have been shown to proceed stepwise via dipolar intermediates. An example is the formation of a hexahydro-1,3,5-triazine derivative from an imine and isocyanate [reaction (7)]. The 1,4-dipolar intermediate

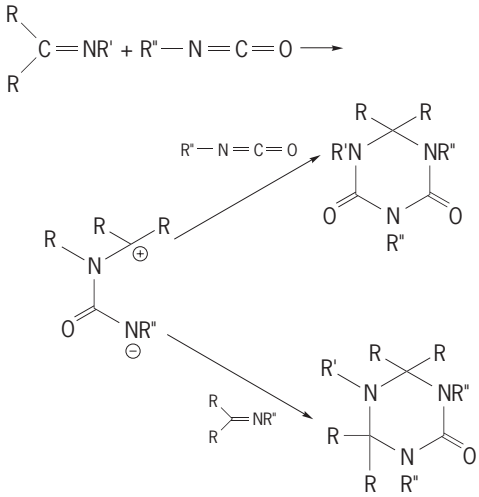

(7)

can be intercepted by either a second molecule of isocyanate or an imine.

Systems containing conjugated double bonds have been shown to react with (predominantly) aryl and sulfonyl isocyanates across both double bonds in a hetero-diene synthesis; an example is shown in reaction (8), the formation of an oxazolopyrimidine

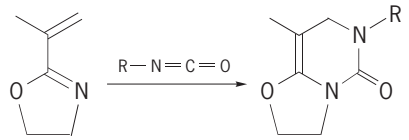

(8)

from 2-vinyloxazole. Isocyanates are also known to react with each other to form low-molecular-weight oligomers or polymers under certain catalyzed conditions. Most commonly they form dimers, 1,3-diazetidinediones or trimers with hexahydro-1,3,5-triazinetrione structure, also known as isocyanurates [reaction (9)]. These cyclizations have also been

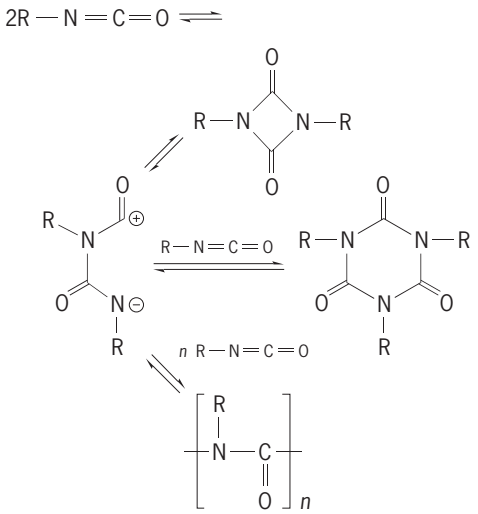

(9)

shown to involve 1,4-dipolar intermediates; they are thermally reversible, especially in the case of the cyclodimers which dissociate back to starting isocyanates on heating to approximately 180–200°C

(350–390°F). An extension of this reaction type is the catalyzed polymerization of alkyl and aryl isocyanates to polyureas or 1-nylons at low temperature. These homopolymers, which are stabilized by addition of base as chain terminators, are thermolabile and usually melt with decomposition at 180–200°C (350–390°F); they have not found industrial application. *See* POLYMERIZATION.

The thermal generation of *N,N'*-disubstituted carbodiimides from isocyanates in the absence of catalysts is likely to involve a labile asymmetric isocyanate dimer in which the C═O group of the heterocumulene is involved in the cycloadduct formation [reaction (10)]. Other heterocumulenes,

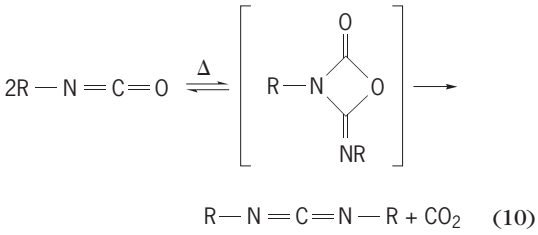

(10)

such as carbodiimides (RN═C═NR), isothiocyanates (RN═C═S), and ketenes (R_2C═C═O), are also known to undergo cycloadditions with isocyanates.

Isocyanates react with organometallic and phosphorus compounds primarily by inserting into the metal-X bond (X = OR, H, halogen, NR_2, and so forth). Several of these additions, especially those involving certain tin (Sn) compounds with SnOR or SnOSn bonds, have been studied in detail, since they exhibit catalytic activity in many isocyanate reactions such as trimerization and urethanization. *See* ORGANOMETALLIC COMPOUND.

Isocyanate-based polymers. Polymers based on the reactions of isocyanates were first developed in 1937. Diisocyanates react with difunctional reagents, such as diols, to form addition polymers with a wide variety of properties. The flexibility in the choice of starting materials (diisocyanate, diol, diamine, diacid, and so forth) and consequently in the multitude of possible adducts make this product group unique in the field of polymeric materials.

Two aromatic diisocyantes, tolylene diisocyanate (TDI; **2**) and di(4-isocyanatophenyl)methane (MDI; **3**) have become the major starting materials for a family of polymeric products, such as flexible

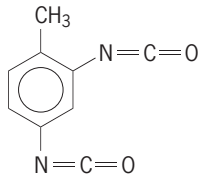

(2)

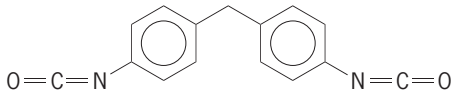

(3)

and rigid polyurethane foams used in construction and appliance insulation, automotive seating, and furniture. Elastomers based on MDI, polyols, and polyamines are widely used in the automotive industry, where reaction injection molding technology is used for the manufacture of exterior parts such as body panels and bumpers. *See* POLYOL.

Thermoplastic polyurethane elastomers (TPU) are used in the molding and extrusion of many industrial and consumer products with superior abrasion resistance and toughness. *See* POLYURETHANE RESINS.

The trimerization to polyisocyanurates and the formation of polyamides from dicarboxylic acids have been utilized to synthesize polymers with excellent thermal properties. Aliphatic diisocyanates, notably 1,6-diisocyanatohexane (HDI), fully hydrogenated MDI (H_{12}MDI), and isophorone diisocyanate (IPDI), have become building blocks for color-stable polyurethane coatings and elastomers with high abrasion resistance. *See* POLYAMIDE RESINS; POLYMER.

Toxicology. Most low-molecular-weight isocyanates are strong lacrimators, and exposure to vapors should be avoided. Inhalation of vapors can cause pulmonary edema. Isocyanates are also known to be strong skin irritants that can cause allergic eczema and bronchial asthma, and they have been known to produce severe skin and lung hypersensitivity. Most isocyanates have a low level of oral toxicity. Amine precursors of several industrial diisocyanates are believed to be carcinogenic. Exposure limits have been set for isocyanate vapors in industrial environments.

Reinhard H. Richter

Bibliography. V. I. Gorbatenko and L. I. Samarai, Synthesis and reactions of haloalkyl isocyanates, *Synthesis*, 1980:85–110, 1980; G. Oertel, *Polyurethane Handbook*, 2d ed., 1993; S. Patai (ed.), *The Chemistry of Cyanates and Their Thioderivatives*, 1977; A. A. R. Sayigh, H. Ulrich, and W. J. Farrissey, Jr., Diisocyanates in J. K. Stille (ed.), *Condensation Polymers*, pp. 369–476, 1972; H. Ulrich, Synthesis of polymers from isocyanates in polar solvents, *J. Polym. Sci.: Macromol. Rev.*, 11:93–133, 1976.

Isoelectric point

The pH of the acidity or alkalinity of the dispersion medium of a colloidal suspension or an ampholyte at which the solute does not move in an electrophoretic field. The term isoelectric point is abbreviated pI.

Ampholytes. These molecules carry acid as well as basic functional groups. When dissolved in a suitable medium, ampholytes may acquire positive and negative charges by dissociation or by accepting or losing protons, thereby becoming bipolar ions (zwitterions). Ampholytes may be as small as glycine and carry just one chargeable group each; or as large as polyampholytes (polyelectrolytes that carry positive charges, negative charges, or both). They may possess molecular weights in the hundreds of thousands like proteins or in the millions like nucleic acids, and

carry many hundreds of chargeable groups. *See* ION; NUCLEIC ACID; PH; PROTEIN.

Determination. An example of establishing the isoelectric point is shown by the course of the pH changes during the titration of alanine [$NH_2CH(CH_3)COOH$], a 1:1 ampholyte, meaning a molecule that carries one positively and one negatively ionizable group. Starting from acid solution (**Fig. 1**), relatively small pH changes (with alkali as the titrant) are observed between pH 2 and 3 (acidic), and again between pH 9.5 and 10.5 (alkaline), caused by the buffering capacity of the carboxyl (—COOH) and amine (—NH_2) groups as weak electrolytes. The pH for $^1/_2$ equivalence corresponds to the pK of the acid function (a value related to the equilibrium constant), where one-half of the alanine molecules still carry only a positive charge ($-NH_3^+$), while the other half are also negatively charged ($-COO^-$). Thus alanine exists in the form of zwitterions. *See* PK.

In the case of alanine, the titration curve is repeated with close symmetry on its alkaline side. After adding 1.5 equivalents of sodium hydroxide (NaOH), the pH is that of the pK of the basic groups, which is also the range of maximum buffering capacity for alkaline pHs. At this point, one-half of the alanine molecules are still in the zwitterionic state, and one-half are entirely negatively charged.

Halfway between the acid and alkaline titration ranges, after adding 1 equivalent of NaOH and thus creating acid-base neutrality, all alanine molecules are in the zwitterionic, net-zero-charge state, and the pH ≡ pI. The value of pI in this example can be calculated as the average of the two [basic (b) and acidic (a)] pK values by the equation

$$pI = \frac{pK_a + pK_b}{2} = 6.1$$

The net charges of alanine vary in an easily recognizable way from the far acid to far alkaline pHs, because the 2 pK values of the two chargeable groups lie far apart. The same is true for the titration curve of

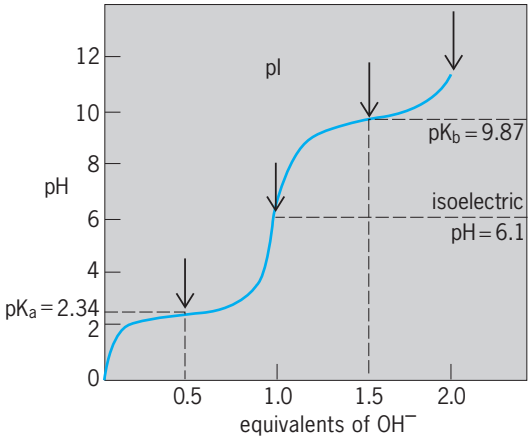

Fig. 1. Titration of alanine with sodium hydroxide (NaOH), showing the course of pH with added fractional equivalents. The four arrows show, from left to right, the pK_a [$^1/_2$ cations: $NH_3^+(CH_3)COOH$, and $^1/_2$ zwitterions: $NH_3^+(CH_3)CHCOO^-$], the pI (all zwitterions), the pK_b ($^1/_2$ zwitterions, $^1/_2$ anions), and the end of titration, when all alanine molecules are in the anionic form: $NH_2(CH_3)CHOO^-$.

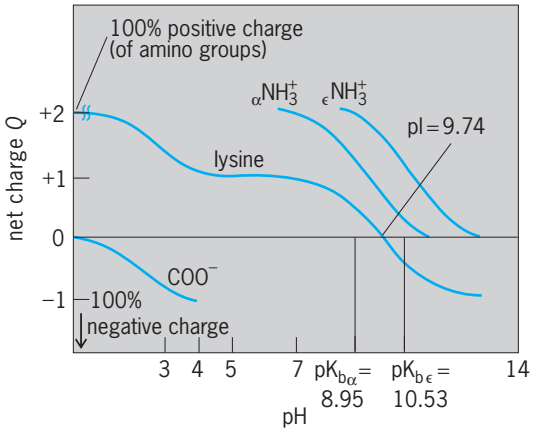

Fig. 2. Course of the charges carried by lysine ($pK_a = 2.18$) molecules in dilute aqueous solution with pH during a titration with alkali. Shown are the net charge, going from +2 at low pH to +1 between pH 4.5 and 6.5, the isoelectric point at pH 9.74, the net charge approaching −1 at about pH 13, and the courses of the individual charges on the three chargeable groups. The break in the curve indicates that in this range the course of the net charges (Q) is not shown completely, but eventually reaches a value of +2.

the amino acid lysine, $NH_2CH[(CH_2)_4NH_2]COOH$, a 2:1ampholyte, meaning a molecule that carries three ionizable (chargeable) groups, two of which are the same and one of opposite sign (**Fig. 2**). The inflection points mark the pK's. The pH at which the curve of overall charge crosses from positive to negative values marks the isoelectric point at pH = 9.74 = pI. Here, lysine is in the purely zwitterionic state.

For molecules that carry four or more chargeable groups, that is, for polyelectrolytes, the courses of the overall titration curves may no longer reflect the individual dissociation steps clearly, as the dissociation areas usually overlap. The isoelectric point then becomes an isoelectric range, such as for pigskin (parent) gelatin, a protein that exhibits an electrically neutral isoelectric range from pH 7 to pH 9.

Since ampholytes in an electric field migrate according to their pI with a specific velocity to the cathode or anode, the blood proteins, for example, can be separated by the techniques of gel or capillary-zone electrophoresis. *See* ELECTROPHORESIS; TITRATION.

Isoelectric focusing. In the method of isoelectric focusing, a pH gradient within a chromatographic column parallels the potential gradient. The components (for example, proteins) of a mixture will each migrate in the electric field until they reach their respective isoelectric zone, at which point they will stop moving. Different components thus become neatly separated along the electrophoretic path.

Influence of change in reactivities. Since the pI depends on the pKs of the individual groups of an ampholyte, anything that affects the reactivities of the groups also affects the pI. This is the case when a variety of cosolutes, especially electrolytes, are present. An example is gelatin, whose midpoint of the isoelectric range will move from 8 to 6 if the concentration of codissolved sodium chloride (NaCl) changes from 10^{-3} to 10^{-1} *N*. Similarly, the solubility of an ampholyte, which depends on the solvation

of its charge-carrying groups, varies with pH, with codissolved electrolyte, or with other cosolutes. At the isoelectric point, many ampholytes show a maximum of solution turbidity, some form of precipitation, or a larger friction factor and a maximum or minimum of viscosity. *See* CONCENTRATION SCALES; ELECTROLYTE.

Ion-exchange chromatography. The important separation technique of ion-exchange chromatography is based on the selective adsorption of ampholytes on the resins with which the column is filled, at a given pH. For example, the larger the net positive charge of an ampholyte, the more strongly will it be bound to a negative ion-exchange resin and the slower will it move through the column. By rinsing with solutions of gradually increasing pH, the ampholytes of a mixture can be eluted and made to emerge separately from the column and be collected. Automated amino acid analyzers are built on this principle.

Significance. The notion that some ampholytes may pass with changing pH through a state of zero charge (zero zeta potential) on their way from the positively to the negatively charged state has become so useful for specifying and handling polyampholytes that it was extended to all kinds of colloids, and to solid surfaces that are chargeable in contact with aqueous solutions. Practically all metal oxides, hydroxides, or hydroxy-oxides become charged by the adsorption of hydrogen ions (H^+) or hydroxide ions (OH^-), while remaining neutral at a specific pH. Strictly speaking, the isoelectric point of electrophoretically moving entities is given by the pH at which the zeta potential at the shear plane of the moving particles becomes zero. The point of zero charge at the particle (solid or surface) is somewhat different but often is not distinguished from the isoelectric point. It is determined by solubility minima or, for solid surfaces, is found by the rate of slowest adsorption of colloids (for example, latexes) of well-defined charge. *See* AMINO ACIDS; COLLOID; ELECTROKINETIC PHENOMENA; ION EXCHANGE; ION-SELECTIVE MEMBRANES AND ELECTRODES. F. R. Eirich

Bibliography. P. C. Hiemenz, *Principles of Colloid and Surface Chemistry*, 3d ed., 1997; J. W. Jorgenson, New directions in electrophoretic methods, *Amer. Chem. Soc. Symp.*, 335:182–198, 1987; N. Kallay et al., Determination of the isoelectric points of several metals by an adhesive method, *J. Phys. Chem.*, 95:7028–7032, 1991; P. G. Rhigetti, *Isoelectric Focusing*, 1983; C. Tanford, *Physical Chemistry of Macromolecules*, 1961.

Isoelectronic sequence

A term used in spectroscopy to designate the set of spectra produced by different chemical elements ionized in such a way that their atoms or ions contain the same number of electrons. An atom or ion with the same number of electrons will generally have spectra with similar features, as the number of electrons determines most atomic properties.

The sequence in the **table** is an example. Since the neutral atoms of these elements each contain Z electrons, removal of one electron from scandium, two from titanium, and so forth, yields a series of ions all of which have 20 electrons. Their spectra are therefore qualitatively similar, but the spectral terms (energy levels) increase approximately in proportion to the square of the core charge, just as they depend on Z^2 in the one-electron sequence H, He^+, Li^{2+}, and so forth. As a result, the successive spectra shift progressively toward shorter wavelengths (higher energies), but maintain similar spectral features. Isoelectronic sequences are useful in predicting unknown spectra of ions belonging to a sequence in which other spectra are known. *See* ATOMIC STRUCTURE AND SPECTRA.

Example of isoelectronic sequence

Designation of spectrum	Emitting atom or ion	Atomic number, Z	Number of electrons
CaI	Ca	20	20
ScII	Sc^+	21	20
TiIII	Ti^{2+}	22	20
VIV	V^{3+}	23	20
CrV	Cr^{4+}	24	20
MnVI	Mn^{5+}	25	20

F. A. Jenkins; W. W. Watson; Alfred S. Schlachter

Isoetales

An order in the class Lycopsida that diverged from the Lepidodendrales in the Late Devonian. These two groups have several characters that are not found in other lycopsids, notably a centralized, shootlike rooting structure (rhizomorph) that allows finite growth, wood production, and tree-sized dimensions. Isoetaleans evolved from trees as an increasingly specialized and reduced lineage; all but the earliest are small-bodied shrubs and pseudoherbs. A reduced morphology characterizes the only living isoetalean genus, *Isoetes*. *See* LEPIDODENDRALES.

Isoetes (see **illus.**) is a globally distributed genus of approximately 150 species. Many are poorly differentiated, as their simple morphology lacks taxonomically useful characters and hence indicates a high probability of evolutionary convergence.

The slow-growing plants resemble a small grass or rush. The basal corm is a highly reduced, woody, bipolar rhizomorph, anchored in the substrate by rootlets and capped by a vestigial stem. It is surrounded by clasping quill-like leaves that function as sporophylls, developing large sporangia. As resources decline during the growing season, the sporangia that are formed on successive sporophylls switch from being "female" megasporangia to "male" microsporangia, and eventually to abortive sporangia. Megasporangia contain several hundred megaspores, whereas microsporangia contain numerous small microspores that germinate to release spermatozoids that are multiflagellate.

Isoetaleans are delimited by the bilateral symmetry and ontogeny of their rhizomorphs, relative to the

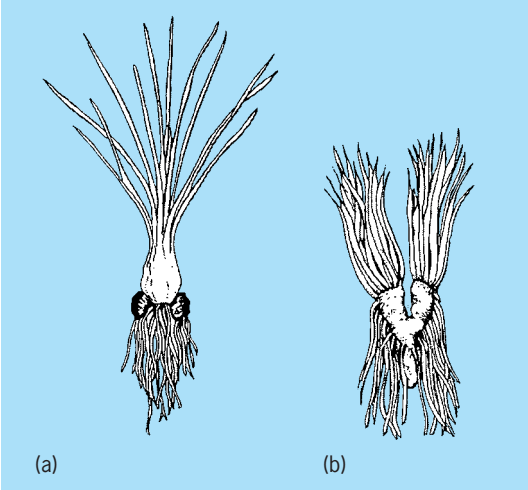

Morphology and anatomy of extant *Isoetes*. (*a*) Compact, cormose rhizomorph and unbranched stem of *I. melanopoda*. (*b*) More elongate, once-branched rhizomorph and stem of *I. andicola*. (*After W. N. Stewart and G. W. Rothwell, Paleobotany and the Evolution of Plants, 2d ed., Cambridge University Press, 1993*)

radial rhizomorph symmetry of the Lepidodendrales. The most likely origin of the Isoetales is from among the relatively primitive lepidodendraleans that possess bisporangiate cones resembling those of the Selaginellales. The group originated in the Late Devonian with poorly known small trees such as *Lepidosigillaria, Cyclostigma*, and later *Protostigmaria-Lepidodendropsis*. The Late Carboniferous *Sporangiostrobus* and *Chaloneria* were small trees and shrubs inhabiting coal swamps. The Early Triassic *Pleuromeia* and Early Cretaceous *Nathorstiana* (Pleuromeiales) resembled *Chaloneria* in growth habit but showed reductions of the rhizomorph. Plants resembling *Isoetes* occur in the Middle to Late Triassic periods, but the earliest definite *Isoetes* species are of Early Cretaceous age. *See* SELAGINELLALES.

Isoetaleans inhabited Paleozoic coal swamps (*Chaloneria*), Mesozoic coastal marshes (*Pleuromeia*), and Cenozoic river and lake systems (*Isoetes*). They typically formed a major ecological component of low-diversity communities that lacked large trees. *See* LYCOPHYTA; LYCOPSIDA.

Richard M. Bateman; William A. DiMichele

Bibliography. K. B. Pigg, Evolution of isoetalean lycopsids, *Ann. Mo. Bot. Gard.*, 79:589–612, 1992; J. E. Skog and C. R. Hill, The Mesozoic herbaceous lycopsids, *Ann. Mo. Bot. Gard.*, 79:648–675, 1992; W. C. Taylor and R. J. Hickey, Habitat, evolution, and speciation in *Isoetes, Ann. Mo. Bot. Gard.*, 79:613–622, 1992.

Isolaimida

An order of nematodes comprising the single superfamily Isolaimoidea. The order consists of one family and one genus. They are rather large for free-living soil nematodes (0.1–0.2 in. or 3–6 mm) and are found in seldom-cultivated sandy soils. Some forms have anterior annulations, while others have posterior transverse rows of punctations. The diagnostic characteristics of this superfamily are the presence of six hollow tubes around the oral opening and two whorls of six circumoral sensilla. Amphids are apparently absent, though some authors speculate that their function is taken over by the dorsolateral papillae of the second whorl. The triradiate stoma is elongate and has thickened walls anteriorly. The esophagus is clavate. The female ovaries are opposed and outstretched, and may have flexures. The gubernaculum, supporting the thickened male spicules, has two dorsal apophyses, and male preanal supplementary organs are papilloid. Paired caudal papillae are large in males, small in females. *See* NEMATA (NEMATODA).

Armand R. Maggenti

Isomerization

Rearrangement of the atoms within hydrocarbon molecules. Isomerization processes of practical significance in petroleum chemistry are (1) migration of alkyl groups, (2) shift of a single-carbon bond in naphthenes, and (3) double-bond shift in olefins.

Migration of alkyl groups. An example of alkyl group migration (skeletal isomerization) is reaction (1). Isomerization to more highly branched con-

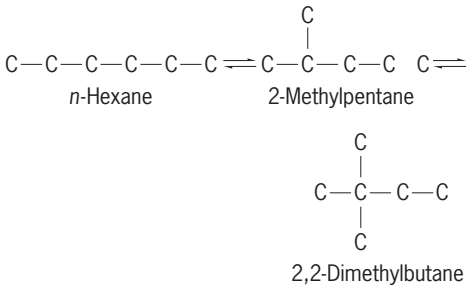

(1)

figurations has commercial importance since it results in improvement in combustion quality in the automobile engine as measured by octane number and increased chemical reactivity because tertiary carbon atoms result. The unleaded, motor-method octane numbers of the hexane isomers shown in reaction (1) are 26.0, 73.5, and 93.4, respectively. Normal butane is converted to isobutane (which has a tertiary carbon atom) to attain chemical reactivity with olefins in alkylation reactions where *n*-butane is inert.

Isomerization of paraffins is a reversible first-order reaction limited by thermodynamic equilibrium which favors increased branching at lower temperatures. Undesirable cracking reactions leading to catalyst deactivation occur at higher temperatures. They are controlled by adding a cracking suppressor such as hydrogen.

Conversion of normal butane to isobutane is the major commercial use of isomerization. Usually, it is carried out in either liquid- or vapor-phase over aluminum chloride catalyst promoted with hydrogen chloride. In the vapor-phase process (250–300°F or 120–150°C), the aluminum chloride is often supported on bauxite. In the liquid-phase processes (180°F or 82°C), it is dissolved in molten antimony

trichloride or used in the form of a liquid complex with hydrocarbon. A second type of catalyst for vapor-phase isomerization (300–850°F or 150–450°C) is a noble metal, usually platinum, supported on a carrier. This may be alumina with halide added to provide an acidic surface. All the processes are selective (95–98% to isobutane). Approximately 60% of the *n*-butane feed is converted per pass to isobutane in the liquid-phase process.

Isopentane, a high-octane component used in aviation gasoline, is made commercially by isomerization of *n*-pentane. Petroleum naphthas containing five- and six-carbon hydrocarbons also are isomerized commercially for improvement in motor-fuel octane numbers. Noble-metal catalyst is normally used with higher-molecular-weight feeds. Isomerization of paraffins above six carbon atoms is of less importance, since octane improvement is limited by predominance of monomethyl branching at equilibrium. Skeletal isomerization is an important secondary reaction in catalytic cracking and catalytic reforming. Aromatics and olefins undergo skeletal isomerization as do paraffins.

Single-carbon bond shift. This process, in the case of naphthenes, is illustrated by reaction (2). Cyclo-

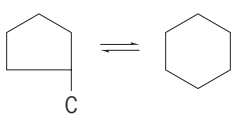

(2)

hexane and methylcyclohexane have been produced commercially by liquid-phase isomerization of the five-carbon ring isomers over aluminum chloride–hydrocarbon-complex catalyst promoted by hydrogen chloride. Conversion per pass is high, selectivity excellent, and reaction conditions mild (190°F or 88°C). Cyclohexane is a raw material for making nylon, and it may be dehydrogenated to benzene. Methylcyclohexane has been used to make synthetic nitration-grade toluene.

Shift of a double-bond. This process is usefully applied when a specific olefin is needed for chemical synthesis, as in reaction (3). Double-bond shift

$$\mathrm{C{-}C{-}C{=}C} \rightleftharpoons \mathrm{C{-}C{=}C{-}C} \qquad (3)$$

occurs selectively over acidic catalysts at temperatures below 450°F (230°C). However, the proportion undergoing skeletal isomerization increases as temperature is increased until, at temperatures in the range of 600–950°F (300–510°C), equilibrium is approached at fairly high space velocities. Equilibrium favors movement of double bonds to the more stable internal positions (85.6% 2-butene at 400°F or 200°C), and octane improvement accompanies this shift; however, the increase of octane number normally is insufficient to justify the cost of processing thermally cracked gasolines solely for this purpose. This type of isomerization occurs as a secondary reaction in the catalytic cracking and catalytic polymerization processes, in part accounting for the high octane numbers of the gasolines. *See* AROMATIZATION; CRACKING; MOLECULAR ISOMERISM; PETROLEUM PROCESSING AND REFINING.

George E. Liedholm

Isometric process

A constant-volume thermodynamic process in which the system is confined by mechanically rigid boundaries. No direct mechanical work can be done on the surroundings by a system with rigid boundaries; therefore the heat transferred into or out of the system equals the change of internal energy stored in the system. This change in the internal energy, in turn, is a function of the specific heat and the temperature change in the system as in Eq. (1), where Q_V

$$Q_V = U_2 - U_1 = \int_1^2 C_V dT \qquad (1)$$

is the heat transferred at constant volume, U is the internal energy, C_V is the heat capacity at constant volume, and T is the absolute temperature. If the process occurs reversibly (the system going through a continuous sequence of equilibrium states), Eq. (2)

$$Q_V = \int_1^2 T dS \qquad (2)$$

holds, where S is the entropy. There is an increase in both the temperature and the pressure of a constant volume of gas as heat is transferred into the system. For a comparison of the isometric process with other processes involving a gas *see* POLYTROPIC PROCESS

Philip E. Bloomfield

Isopoda

An order of invertebrate crustaceans related to crabs, shrimps, and lobsters, all members of the class Malacostraca. The closest relatives of Isopoda are the amphipods, mysids (opposum shrimps), cumaceans, and tanaids, all of which are placed in the superorder Peracarida. Isopods are generally small but very common, highly diversified, and occurring in marine, freshwater, and terrestrial habitats. Sow bugs, pill bugs, and woodlice, as well as their marine relatives (such as gribbles and sea slaters), belong to this group. *See* CRUSTACEA; MALACOSTRACA; PERACARIDA.

Morphology. Isopods are characterized by the lack of a carapace (an outer case or covering made of chitin), having the first thoracic segment fused with the head, one pair of maxillipeds (one of the four sets of paired mouthparts used for feeding), and usually seven pairs of similar legs, although the first leg may be clawlike. Adults of the suborder Gnathiidea possess five pairs of legs.

As in most crustaceans, the body has three parts: the head (cephalon), thorax (pereon), and abdomen (pleon) [see **illustration**]. The cephalon bears sessile eyes and two pairs of antennae, the first of which is chemosensory and the second tactile in function. Each of the seven segments of the pereon

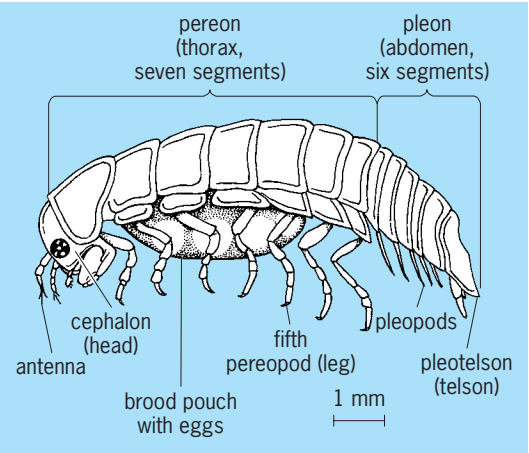

***Limnoria* (Flabellifera), female. (*After R. J. Menzies, The comparative biology of the wood-boring isopod crustacean Limnoria, Museum of Comparative Zoology, Harvard Coll. Bull., 112(5):363–388, 1954*)**

(pereonites) normally bears a pair of legs called pereopods. These are used for locomotion and grabbing prey. The abdomen typically consists of five segments or pleonites (which may be fused in some) plus a sixth segment called the pleotelson. Uropods are the paired appendages at the base of the pleotelson; these are used in orientation and swimming. Each of the pleonites has a ventral pair of foliaceous appendages (leaflike pleopods) which are used for swimming and respiration.

Ecology. At present there are more than 10,400 described species of isopods. Of those, about 4200 species are terrestrial, and these include the familiar sow bugs, pill bugs, roly-polys, and woodlice. They may be found in any habitat on land, including forests, caves, deserts, city sidewalks, and greenhouses.

Terrestrial isopods feed on algae, moss, bark, and decaying plant and animal matter. Like insects, they play a role in the breakdown and recycling of organic nutrients. Cellulose digestion results from the presence of bacteria in the digestive system.

In the continental United States and Canada, there are about 120 known species of terrestrial isopods. The most common are widely distributed and have probably been introduced from Europe. A majority of the most abundant species in North America are concentrated in the northeast and northwest, possibly a result of the transport of their ancestors into harbors by human immigrants.

Marine isopods, numbering more than 5600 species, are found at all latitudes and in all marine habitats, from the intertidal zone to the deep sea, including coral reefs, sponges, algae, and seagrass beds. They range in size from less than a millimeter (<0.04 in.) [one of the parasitic forms] to 1.5 mm (0.06 in.) [*Microcharon*, found living between sand grains] to 44 cm (17.3 in.) [*Bathynomus* in the depths of the Caribbean Sea]. A few, such as *Idothea*, are cosmopolitan (having worldwide distribution), perhaps being distributed by floating on rafts of seaweed. Others travel in ballast water in oceangoing ships from place to place.

Because they have limited swimming ability and move chiefly by crawling on the bottom, most isopods have relatively small ranges. They recycle nutrients by eating detritus, both plant and animal material, and become part of the food chain in the sea as they are eaten by larger animals, particularly fish. Marine isopods may be omnivores, herbivores, or predators; some specialize as parasites, living off fluids and tissues of their hosts, mainly fish and other crustaceans. They fulfill many of the same ecological niches as insects.

The isopod species found in freshwater total about 600. Their habitats include rivers, subterranean streams, caverns, wells, swamps, and hot springs. They tend to have very restricted habitats and thus are the most likely isopods to be designated as threatened or endangered species.

Reproductive biology. As is typical of the peracarids, isopods develop inside the female brood pouch (marsupium) located on the underside of the abdomen. When the young emerge, they typically lack the seventh pair of legs and are known as manca but otherwise look like adults (except in Gnathiidea). Some evidence of parental care has been recorded in the family Sphaeromatidae (suborder Flabellifera); and in *Iais* (suborder Asellota) adults are known to carry the manca about by clasping it to their ventral abdomen. Mate guarding has also been observed. Examples of sequential hermaphroditism or sex change have been found in several groups, both male-to-female and female-to-male. However, sexes are usually separate and can be virtually identical or widely divergent in morphology, depending on the suborder or family.

Evolution. Isopods date from the Paleozoic Era, when malacostracans as a group underwent a rapid evolution. The earliest fossil records of isopods contain members of the present-day Phreatoicidea, dating 325 million years ago, representing a very early offshoot from the original common ancestor. These animals, currently found only in freshwater habitats, are only slightly changed from their fossil ancestors and may be considered as living fossils (living species belonging to an ancient stock otherwise known only as fossils). Other isopod groups probably evolved as benthic (bottom) dwellers in shallow marine environments such as estuaries and embayments before radiating into the deep sea, and they also conquered land as the only truly terrestrial crustaceans (suborder Oniscidea). *Cyclosphaeroma* from the Jurassic, discovered in England, is related to members of the contemporary family Sphaeromatidae.

Classification. The classification system currently in wide use recognizes 10 suborders, listed below.

Anthuridea are generally elongate and cylindrical, living in soft bottoms of estuaries and oceans, especially in tropical reef environments.

Asellota are very diverse, are generally herbivores, and live in freshwater as well as marine habitats. They are particularly diverse in the deep sea. One pair of pleopods forms a cover over the rest of the pleopods.

Calabozoida have been found thus far only in

freshwater wells in Venezuela. They have unique body segmentation and pleopod structure.

Epicaridea are exclusively external parasites on other crustaceans and may be highly modified. The male is much smaller than the female, whose body may be very distorted.

Flabellifera inhabit both fresh and salt water; they can be free-living, predaceous, commensal, or parasitic. A common characteristic is a tail fan comprising the pleotelson and posterior appendages (uropods).

Gnathiidea can be found only in marine environments but at all latitudes. The male is known for large frontal mandibles. The juveniles are parasitic on fish, while the adults probably do not feed.

Microcerberidea may be found in marine, brackish, or freshwater in sand or sediment. They are less than 2 mm (0.08 in.) in length.

Oniscidea, the sow bugs or pill bugs, have successfully conquered almost all terrestrial habitats, as their pleopods are adapted for breathing air.

Phreatoicidea, relict (being a persistent, isolated remnant of a once-abundant species) crustaceans, are found thus far only in Australia, Tasmania, and South Africa. They are laterally compressed with long antennae and long uropods.

Valvifera have uropods which are flexed beneath the abdomen and cover the pleopods. Many live on seaweed or seagrass, from the tropics to the polar regions.

Several authors (A. Brandt and G. C. B. Poore) have suggested that the Flabellifera is an artificial grouping, so further reorganization of this system is certain. These authors have also erected two new suborders, Tainisopodidea, including members of the family Tainisopidae, found only in ground waters in western Australia, and Phoratopidea, comprising a single species found only in Australia. Marilyn Schotte

Bibliography. A. Brandt and G. C. B. Poore, Higher classification of the flabelliferan and related Isopoda based on reappraisal of relationships, *Invert. Systemat.*, 17(6):893–923, 2003; G. Brusca and R. C. Brusca, *Invertebrates*, Sinauer Associates, Sunderland, 1990; B. Kensley and M. Schotte, *Guide to the Marine Isopod Crustaceans of the Caribbean*, Smithsonian Institution Press, Washington, DC, 1989; B. Kensley, M. Schotte, and S. Schilling, *World List of Marine, Freshwater and Terrestrial Crustacea Isopoda*, www.nmnh.si.edu/iz/isopod; M. R. Warburg, *Evolutionary Biology of Land Isopods*, Springer-Verlag, Berlin, 1993.

Isoprostanes

A class of natural products discovered relatively recently. The isoprostanes (iPs) are isomeric with another class of natural products, the prostaglandins. The prostaglandins are the result of enzymatic oxygenation of polyunsaturated fatty acids (PUFAs), in particular arachidonic acid (AA). In contrast, the iPs are formed in vivo by a nonenzymatic, free-radical oxygenation of arachidonic acid. This important distinction in the mode of formation of the iPs is responsible for a more complex mixture of iPs being generated in vivo. For example, whereas the endoperoxide prostaglandin G_2 (PGG_2) is formed by the cyclooxygenase enzymes (COX1 and COX2), four classes of iPs are formed as a result of the free-radical oxygenation of arachidonic acid (**Fig. 1**), with each class containing 16 iPs for a total of 64 individual iP molecules. The discovery of iPs is important for two reasons: (1) group III iPs are incidental ligands for the prostaglandin receptors, hence they possess biological activity; (2) iPs are the product of oxidative stress. Their measurements have been shown to be a predictor of the onset and severity of inflammatory diseases such as Alzheimer's disease and atherosclerosis. *See* ALZHEIMER'S DISEASE; CIRCULATION DISORDERS; EICOSANOIDS; FREE RADICAL.

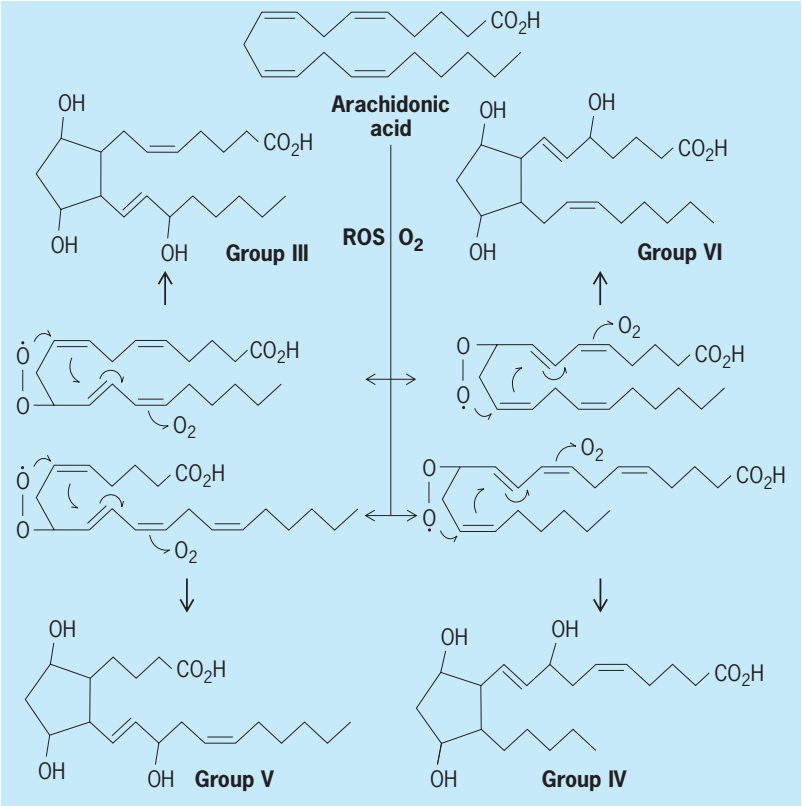

Fig. 1. Endoperoxide mechanism for the formation of isoprostanes. ROS = reactive oxygen species.

Mechanism for formation. Two mechanisms for the formation of iPs have been proposed based on in-depth regiochemical and stereochemical analysis of the peroxidation process. These are the endoperoxide mechanism (Fig. 1) and the dioxetane mechanism. Recently, evidence has been reported showing that the dioxetane mechanism may not be operative.

A more detailed mechanistic view for the formation of group VI iPs by the endoperoxide mechanism is presented in **Fig. 2**. The four iPs shown in the box have been prepared by total synthesis and used in discovery and identification of the same products in human urine (Fig. 4).

Nomenclature. A comprehensive nomenclature for iPs has been introduced utilizing the symbols in the prostaglandin nomenclature. The new classification

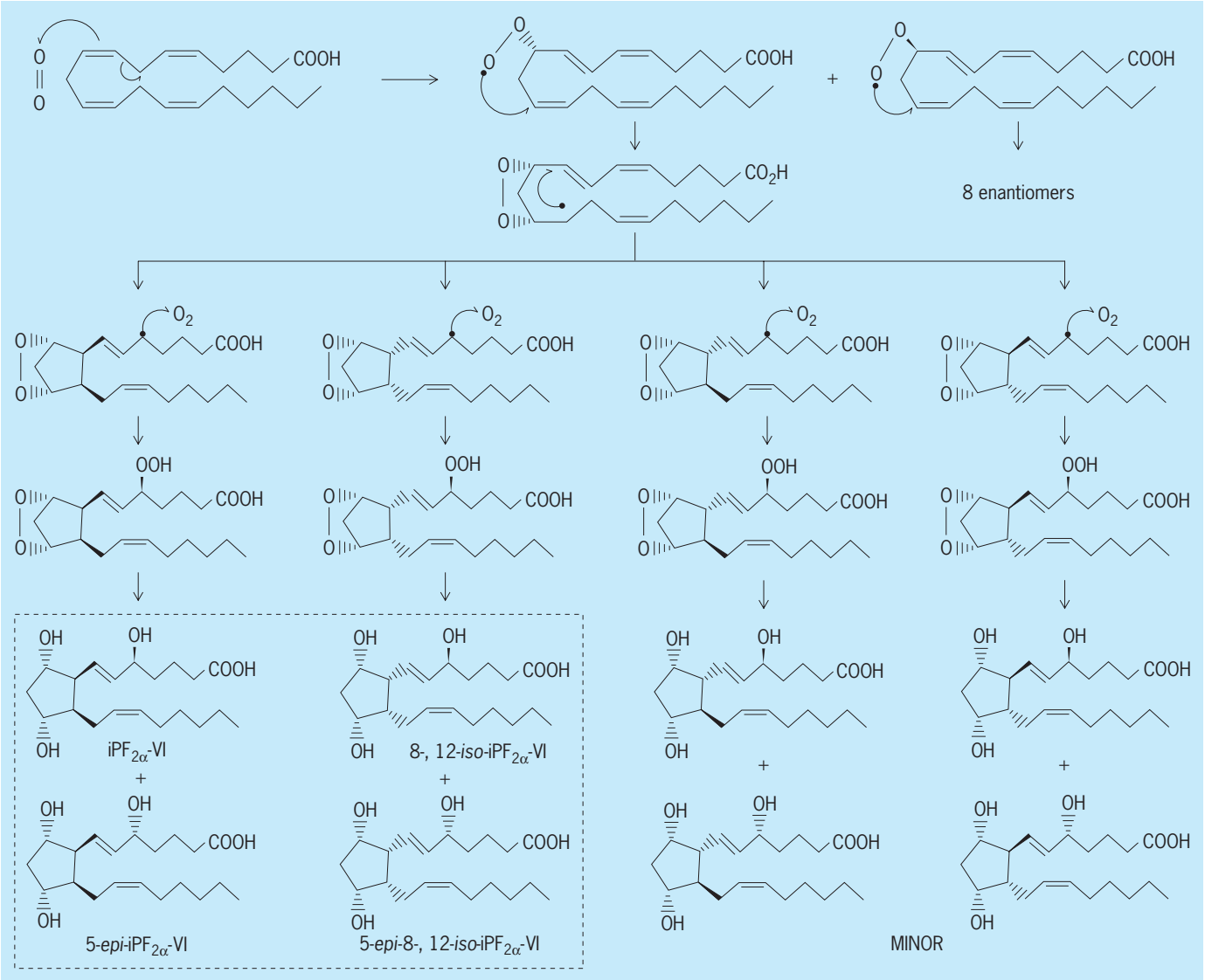

Fig. 2. Mechanism for the formation of group VI isoprostanes. There are 16 products, but for clarity 8 of the enantiomers are not shown.

(roman numerals) reflects the use of the ω end of the molecule instead of the COOH in the double-bond counting system—the same way that double bonds in polyunsaturated fatty acids are referred to as ω-3, ω-6, etc. Another nomenclature is also being used for iPs based on the location of the hydroxyl group in the side chain, the carboxyl end being C-1.

Measurement. Two techniques for measuring iPs are discussed below.

Gas chromatography/mass spectrometry (GC/MS) methods have been developed to quantitate individual iPs, particularly group VI iPs, which appear to be most abundant in urine. Group VI was chosen as the main initial focus because the position of the OH on carbon-5 of the upper side chain allows a lactone to form, which can easily be separated from the rest of the iPs, allowing for a more accurate and reliable determination in biological fluids (**Fig. 3**).

The GC/MS method outlined in Fig. 3 has enabled measurement of in vitro and in vivo elevated levels of $iPF_{2\alpha}$-VI in a number of clinical settings, such as ischemia-reperfusion syndromes, hypercholesterolemia, Alzheimer's disease, and alcohol-induced liver disease. Elevated levels of $iPF_{2\alpha}$-VI have been measured in oxidized low-density lipoprotein (LDL) and in aortic atherosclerotic lesions. In all cases, the discovery and identification of iPs of different groups, including those shown in Figs. 2 and 4, have been made possible by the initial availability of authentic material prepared by total synthesis. *See* GAS CHROMATOGRAPHY; MASS CHROMATOGRAPHY.

Liquid chromatography tandem mass spectrometry (LC/MS/MS) is another promising method for the identification, separation, and quantitation of iPs. Neither derivatization nor prior purification is required. Arachidonic acid–derived iPs readily generate abundant molecular ions (m/z 353) under ESI (electrospray ionization) conditions in the negative-ion mode, which fragments further to ions, some of which are group-specific (m/z 115, 193, 127, 151). **Figure 4** shows the chromatogram of the four groups of arachidonic acid–derived iPs from human

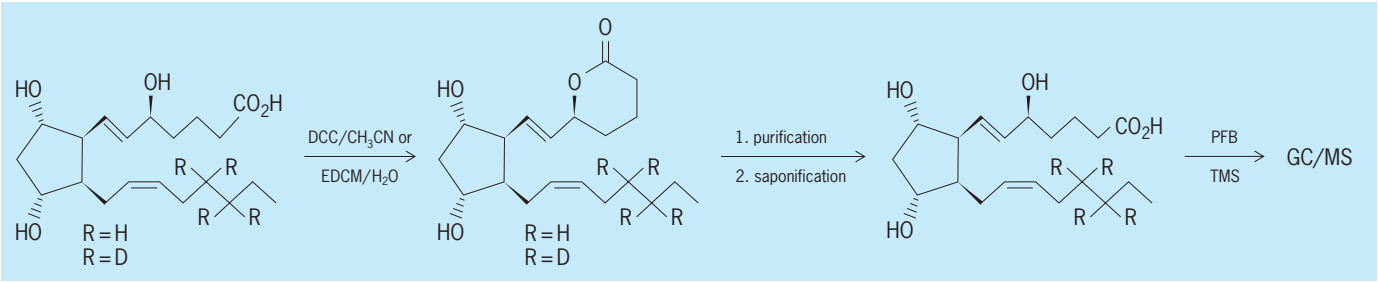

Fig. 3. GC/MS method measurement of iPF$_{2\alpha}$-VI. DCC = dicyclohexylcarbodiimide; EDCM = 1-ethyl-3(3-dimethylaminopropyl)carbodiimide methiodide; PFB = pentafluorobenzyl; TMS = trimethylsilyl; GC/MS = gas chromatography/mass spectrometry; H = hydrogen; D = deuterium.

urine. The chemical structure in the chromatogram identifies the individual iPs. The four most abundant iPs of urinary group VI have the cis configuration of the side chain. *See* LIQUID CHROMATOGRAPHY.

Other novel iPs and isofurans. Reports have shown that oxidation of eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) result in the formation of F_3- and F_4-iPs. Since DHA is extremely rich in the brain, F_4-iPs, collectively called neuroprostanes, could be of great importance in assessing the severity of disease states such as Alzheimer's disease. Chemical syntheses of iPF$_{4\alpha}$-VI and iPF$_{4\alpha}$-III have been reported.

Recently, another series of isomeric compounds called isofurans has been discovered. These compounds, which contain a substituted tetrahydrofuran (THF) ring, are present at detectable levels in normal tissues and biological fluids. Increased amounts of these compounds have been observed in animal models exposed to high oxidant stress.

Biological implications. The importance of free-radical-generated iPs encompasses three general areas.

Damage to cell membranes. One consequence of oxidant injury is alteration in the physicochemical properties and function of cellular membranes. This has been attributed to an accumulation of products of lipid peroxidation in lipid bilayers. Phospholipids are essential constituents of the cell membrane. Radical oxygenation of the polyunsaturated fatty acids, located at the *sn*-2 position of the phospholipids, is bound to create disturbances in the cell membrane. The tightness of the fatty acid portion of the phospholipids is what causes the cell membrane to be hydrophobic and guarantees the integrity of the cell membrane. Polyhydroxylation of the unsaturated fatty acid to yield iPs, which contain three OH groups, is bound to disrupt the tightness of the membrane and have profound effects on the fluidity and integrity of the cellular membrane. For example, a leak could develop in the membrane due to the repulsion by the newly created OH groups and the adjacent lipophilic fatty acids, leading to cell death. *See* CELL MEMBRANES.

Biological activity. Although the mechanism by which iPs are released from the cell membrane remains to be elucidated, phospholipid-containing iPs are subject to cleavage by phospholipases. These iPs are released in the bloodstream, circulate through the liver and kidney, and eventually are excreted in urine, although their metabolic disposition in vivo remains largely unknown. Several iPs have been shown to have biological effects in vitro. For instance, iPF$_{2\alpha}$-III is a vasoconstrictor when infused into the rat renal artery and the rabbit pulmonary artery. In both cases the effects were prevented by pharmacological blockage of thromboxane receptors. Similarly, iPF$_{2\alpha}$-III modulates platelet function. Although it does not cause irreversible aggregation, it facilitates the ability of subthreshold concentrations of conventional agonists to induce this response. These effects require micromolar concentrations of iPF$_{2\alpha}$-III. Interestingly, much lower concentrations of this compound have been reported to modulate platelet adhesion.

Much less information is available about the potential biological activity of other iPs. For example, iPE$_{2\alpha}$-III behaves much like iPF$_{2\alpha}$-III on platelet aggregation. In contrast to these compounds, which appear to act as incidental ligands via the thromboxane receptor, 8,12-*iso*-iPF$_{2\alpha}$-III induces proliferative responses in NIH 3T3 cells via the PGF$_{2\alpha}$ receptor. This compound activates this prostanoid receptor selectively and in a saturable manner, but at concentrations roughly one order of magnitude greater than its natural ligand.

Given the difficulty of interpreting the biological potency of individual iPs when multiple species are likely to be formed coincidentally in vivo, it is premature to decide whether the biological properties of these compounds might be relevant under conditions of oxidant stress in vivo. Nonetheless, the possibility of coordinate cellular activation by multiple species, their delivery to cell membranes in highly concentrated form, and the superiority of thromboxane antagonism to cyclooxygenase inhibition in experimental settings suggest a potential role for these compounds as mediators as well as indices of oxidant stress.

Indices of oxidant stress. Perhaps the most interesting aspect of iPs research at present is the possibility that measurement of these compounds in biological fluids might afford a quantitative index of lipid peroxidation in humans. Increased plasma and urinary concentrations of iPF$_{2\alpha}$-VI and iPF$_{2\alpha}$-III have been reported in several syndromes thought to be associated with excessive generation of free radicals. These include coronary reperfusion after a period of ischemia, poisoning with paraquat and paracetemol,

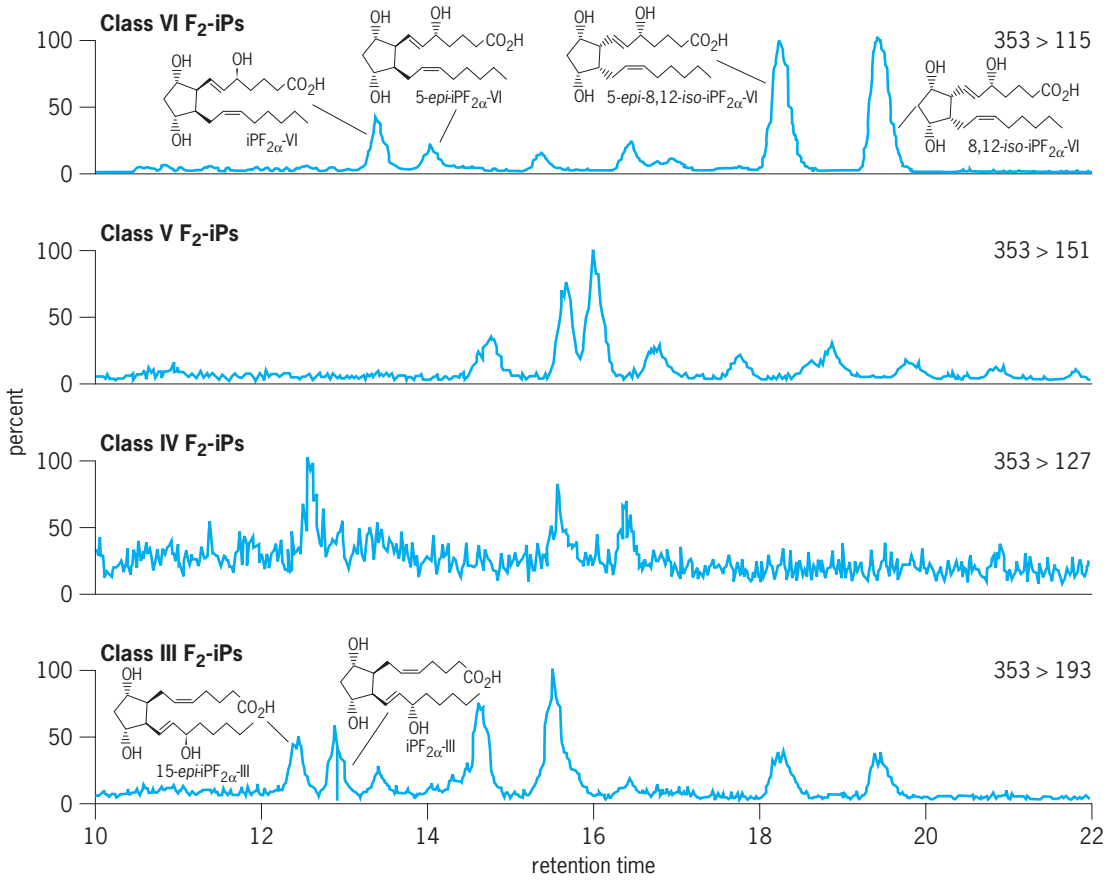

Fig. 4. LC/MS/MS of urinary isoprostanes. Selected peaks were identified by comparison with synthetic standards.

cigarette smoking-related illnesses, and alcohol-induced liver disease. Isoprostanes, which are increased in low density lipoprotein when it is oxidized in vitro, have been found to be increased in the urine of asymptomatic patients with hypercholesterolemia and are present in human atherosclerotic plaque. These observations, and their measurement in urine models of atherosclerosis, raise the likelihood that they may be useful in elucidating the role of LDL and cellular protein oxidation in atherogenesis. In perhaps the most striking evidence to date of the free-radical involvement in atherosclerosis, J. Rokach and G. A. FitzGerald have shown that suppression of elevated iP generation ($iPF_{2\alpha}$-VI) in vivo reduces the progress of atherosclerosis in apolipoprotein E-deficient mice. Significantly, they have also shown in this study that the level of cholesterol remains high and unchanged. It suggests that in cholesteryl linoleate, the entity present in LDL, the cholesterol moiety is the carrier of the linoleic acid which is responsible for the oxidative stress and the subsequent cell damage. Presumably the cholesterol part is left intact. Recently, the most abundant iP in urine, 8,12-*iso*-$iPF_{2\alpha}$-VI (Fig. 4), has been used as an index of oxidant stress and severity of Alzheimer's disease. *See* CHOLESTEROL.

Garret A. FitzGerald; Joshua Rokach

Bibliography. M. Adiyaman et al., Total synthesis of a novel isoprostane $iPF_{2\alpha}$-I and its identification in biological fluids, *Tetrahed. Lett.*, 37:4849–4852, 1996; J. P. Fessel et al., Discovery of lipid peroxidation products formed in vivo with a substituted tetrahydrofuran ring (isofurans) that are favored by increased oxygen tension, *Proc. Nat. Acad. Sci. USA*, 99:16713–16718, 2002; S. Kim et al., The first total synthesis of $iPF_{4\alpha}$-VI and its deuterated analog, *Tetrahed. Lett.*, 43:2801–2805, 2002; H. Li et al., Quantitative high performance liquid chromatography/tandem mass spectrometric analysis of the four classes of F2-isoprostanes in human urine, *Proc. Nat. Acad. Sci. USA*, 96:13381–13386, 1999; J. D. Morrow et al., A series of prostaglandin F2-like compounds are produced *in vivo* in humans by a non-cyclooxygenase, free radical-catalyzed mechanism, *Proc. Nat. Acad. Sci. USA*, 87:9383–9387, 1990; D. Pratico et al., Vitamin E suppresses isoprostance generation in vivo and reduces atherosclerosis in ApoE-deficient mice, *Nature Med.*, 4:1189–1192, 1998; L. G. Quan and J. K. Cha, Preparation of isoprostanes and neuroprostanes, *J. Amer. Chem. Soc.*, 124:12424–12425, 2002; J. Rokach et al., Nomenclature of isoprostanes: A proposal, *Prostaglandins*, 54:853–873, 1997; D. F. Taber, J. D. Morrow, and L. J. Roberts II, A nomenclature system for the isoprostanes, *Prostaglandins*, 53:63–67, 1997; E. A. Weinstein et al., Prothrombinase acceleration by oxidatively damaged phospholipids, *J. Biol. Chem.*, 275:22925–22930, 2000.

Isoptera

An order of the Insecta, commonly called termites, with the general characteristics and stages of other exopterygote insects. They are closely related to the Blattodea. Indeed, the only wood-eating cockroach, *Cryptocercus*, has a similar collection of protozoa in its hindgut to that of the primitive termite *Zootermopsis*. Similarly, the Australian *Mastotermes*, though socially complex, is considered the most primitive extant termite by reason of its having an anal lobe in the adult wing; this character is found throughout the cockroaches.

Approximately 2000 species of termites have been described and these are placed in six or seven families: Mastotermitidae, Hodotermitidae, Termopsidae (sometimes included in the Hodotermitidae), Kalotermitidae, Rhinotermitidae, Serritermitidae and the Termitidae. The latter family represents the higher termites and includes over 80% of all known termite species. *See* EXOPTERYGOTA; ORTHOPTERA.

The termite group is typically a tropical one, but certain genera do occur outside the tropics and may be found as far north as British Columbia and Ontario. The group is distinguished by the fact that all species are eusocial and all feed on cellulose. The castes, apart from the imagos and primary reproductives, are drawn from the larval forms. In this respect they differ from the Hymenoptera where the castes are all variants of adults. Termites live in nests of varying degrees of complexity, ranging from large exigeous mounds to diffuse or temporary galleries in wood or soil. *See* HYMENOPTERA.

Anatomical characteristics. The mature termite (alate or imago) has membranous wings which extend beyond the end of the abdomen. There is a pair of compound eyes, and a pair of ocelli is present in most groups. The wings are superimposed flat on the abdomen when the insect is not in flight. Flight is weak and fluttering and is usually short. When the alate alights, the wings are shed along a basal suture with the base of each wing (the wing scale) being retained. The alates vary in color from yellow, through brown, to coal black. Some species (usually lightly pigmented) fly during the night; others (usually heavily pigmented) fly during the day. The time of flight varies from species to species both with respect to the season of the year and the time of the day or night.

Soldier caste. In almost all termite species a second type of individual is produced in the colony. This is the soldier, which lacks wings, is nonreproductive, and is variously modified for defense. There are four rather distinct types of soldiers: mandibulate, phragmotic, nasutoid, and nasute. In mandibulate soldiers the head and mandibles are greatly enlarged and heavily sclerotized (**Fig. 1***a*). The mandibles may be biting, snapping, or pincherlike and more or less symmetrical or strongly asymmetrical. In phragmotic soldiers the mandibles are not as conspicuously enlarged as in mandibulate forms. The head is high and truncate in front and is used to plug openings in the workings.

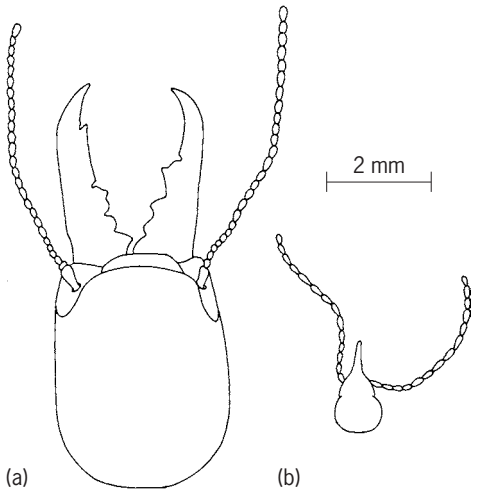

Fig. 1. Head capsule and antennae of (*a*) mandibulate soldier of *Zootermopsis nevadensis* showing enlarged mandibles, and (*b*) nasute soldier of *Lacessititermes palawanensis*.

In the families Rhinotermitidae and Termitidae there is a cephalic gland which opens via a small pore on the dorsal head surface. In some groups this gland and its opening have been variously modified for defense in the soldiers. In some rhinotermitid soldiers the opening of the gland lies at the anterior margin of the head, and the fluid is discharged into an open troughlike structure which extends forward from the head capsule. These have been termed nasutoid soldiers. Certain species have both mandibulate and nasutoid forms. Finally, in the termitid subfamily Nasutitermitinae the cephalic gland opens at the tip of an elongated tube or nasus which projects anteriorly, giving the head the appearance of a pear-shaped syringe (Fig. 1*b*). These are the nasute soldiers.

Soldiers are preceded during development by an intermediate stage (the soldier-nymph or white soldier) which is soldierlike in form but is unsclerotized. In general, mandibulate soldiers are rather large and occur in relatively small numbers in the colonies, whereas nasute soldiers are relatively minute and constitute as much as 30% of the population.

Worker caste. In the more advanced termites there is a third caste, the worker. True workers usually have some pigmentation as opposed to the immature termites, which are generally white. Workers lack wings, are nonreproductive, and have mandibles which resemble those of the imagos; they are usually blind. In many lower termites there is no distinct worker caste, and the work of the colony is performed by young individuals which have undergone at least two stages of development. These are still capable of becoming either alates or soldiers, but may continue to molt without becoming differentiated. Eventually they may undergo stationary molts with no appreciable change in size or structure. These "stabilized" individuals, which function as workers but are still capable of differentiation into other castes, have been termed pseudergates or pseudoworkers.

Replacement reproductives. In addition to the definitive castes (alate, soldier, and worker) another type of individual may occur in the colony under certain circumstances. These individuals are the supplementary or replacement reproductives. Although the original pair (king and queen) may live for two or three decades, the life of the colony itself is not limited by their survival. If one or both are lost, other individuals in the colony become reproductive. If these new reproductives have wing buds, they are termed nymphoid (or brachypterous or second-form) reproductives. If they lack wing buds, they are termed ergatoid (or apterous or third-form) reproductives.

Physiology. Two main areas of physiology which are of particular interest in Isoptera are nutrition and the development of the different castes.

Nutrition. Cellulose is the prime source of nutrition in termites, and may be obtained in nature as wood, humus in soil, or as grass. This consumption of cellulose, while most beneficial in natural areas, especially in the tropics, has caused some of the wood and grass eating forms to be considered pests.

The cellulose is masticated and passed through the gut to the hind intestinal caecum where it is digested by anaerobic protozoa (in the lower termites) and by bacteria (in the higher termites). Most researchers now believe that the cellulose is degraded to glucose and then to acetate which is able to pass through the chitinous lining of the hindgut where the large glucose molecule could not. The older idea that the glucose in the hindgut passes forward through a valve to the midgut for absorption by active transport is not now supported.

Nitrogen is obtained and conserved by both cannibalism and by adventitious consumption of fungi. Recently it has also been demonstrated that some of the symbiotic bacteria are able to fix atmospheric nitrogen. Uric acid is conserved and plays an important role as a nitrogen source.

Caste development. In many families of termites when the functional reproductives are removed from a colony some individuals that remain can transform into egg-laying and sperm-producing forms while retaining larval character; these are the replacement (or neotenic) reproductives. In 1932 it was proposed that the functional reproductives produce a chemical substance (pheromone) that inhibits the development of the gonads of the larvae. The hypothesis was elaborated over the years to include the involvement of twelve pheromones in the caste regulation of *Kalotermes* and *Zootermopsis*. This complex hypothesis, however, was quite speculative, as none of the postulated pheromones has yet been identified nor their origin discovered. Work by other investigators has discounted many of the earlier experiments. It is known, however, that vitellogenesis is affected by the presence of reproductives and that juvenile hormone plays an important part in the regulation; the mandibular glands may be involved. There probably is an inhibitory pheromone that requires contact or ingestion. *See* NEOTENY; PHEROMONE.

Soldier development is more complex. All that is known about it is that injection of a high titer of juvenile hormone into the larvae of some species will cause soldiers to form.

Biology. Eggs laid by a female reproductive hatch into mobile six-legged larvae which are broadly similar in form to the adult. They undergo successive molts over time and the later stages have external wing buds. The wing-budded forms finally molt into winged imagos with dark sclerotized bodies and well developed eyes. These imagos (alates) are produced at certain times of the year, and under certain weather conditions leave the nest all at once in a dispersal flight. The alates fly poorly and eventually settle on the ground or vegetation. Most species then lose their wings by actively twisting their bodies and breaking each wing along a basal suture line. The imagos with short wing stubs, now known as dealates, form into tandems, that is, a male following a female, and move to crevices in soil or wood. There they begin to excavate a small cavity, the nuptial cell or copularium, where mating takes place. The male and female are now referred to as the king and queen, and unlike the hymenopterans the male persists in the colony. Eggs are laid, hatch, and the new larval individuals are the nucleus of a new colony.

Not all larvae in a colony, however, become alate imagos. In most termites larvae may transform into soldiers and workers, while in many species larval forms may become sexually mature without becoming true adults. The main castes found in the Isoptera are these larval forms and, in this order particularly, they can be thought of as end points in individual development: once a termite transforms into a certain caste it loses ability to molt again. The soldier is designed for active defense while the worker is for excavation, feeding, and caring for other castes and individuals. As described above, some genera lack soldiers, some have more than one type of soldier, and some lower termites lack true workers.

While new colonies may be formed by imagos after a flight, in many genera new colonies can also be formed by a process known as sociotomy. This occurs when part of a colony is cut off from the main group by chance, or by the colony having grown so big as to stress the communication system that holds it together, or, as happens in some genera, by the active emigration of part of a colony.

All termites consume cellulose, but in some instances the source is not dead wood. Many termites, such as *Cornitermes*, ingest humus; the harvesting termites (for example, *Hodotermes* in Africa and *Drepanotermes* in Australia) collect and consume grass, while the spectacular *Macrotermes* of Africa construct fungus gardens.

Behavior. As in all social insects, communication is essential to their behavior. Colony cohesion and recognition, defense, construction of galleries and nests, pairing, and mating all are mediated through communication. Visual communication occurs only in adults, where the eyes are well developed; in other castes, where the eyes are rudimentary, communication is by chemical and tactile means. *See* ANIMAL COMMUNICATION; CHEMICAL ECOLOGY; SOCIAL INSECTS.

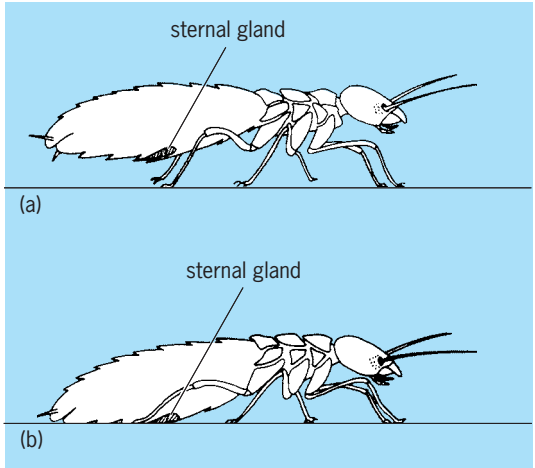

Fig. 2. Posture of a *Zootermopsis* nymph (*a*) during normal activity and (*b*) while laying a trail. (*After K. Krishna and F. Weesner, eds., Biology of Termites, vol. 1, 1969*)

Alarm and defense. Many termites are known to communicate alarm by a combination of characteristic tactile movements coupled with the laying of a chemical trail. The trail has its origin in the secretion of a pheromone by the sternal gland or glands of the abdomen (**Figs. 2** and **3**). A greater or lesser number of termites, depending on the intensity of the initial excitation, are recruited to the spot where the original disturbance occurred by following the trail after having been alerted by physical contact with an excited individual or individuals. This type of communication has also been found in other social insects and has been termed alarm recruitment.

Once a termite has been recruited to a specific locus by this mechanism, the action it then takes is dictated by factors at the site of disturbance. If the cause of alarm is a break in the nest or gallery, then building (in the form of fecal material deposition and the placement of debris) will occur by workers, while soldiers guard the site. If an intruder was the cause, then snapping will be exhibited by workers and mandibulate soldiers; soldiers of the Nasutitermitidae will eject a sticky terpene-containing substance from a gland situated in their heads. When the intruder is immobilized or killed, less excitation is generated and the termites' response is to bury it by depositing fecal material and debris as in a gallery break.

The response of a colony to alarm is further tuned by the phenomenon of polyethism, that is, the differential response of castes, and of individuals of differing ages within a caste, to the same stimulus. This phenomenon is seen most spectacularly in many species of *Nasutitermes* where in a localized disturbance there is an initial recruitment of soldiers as described above, but at the same time the workers retreat. In this case polyethism occurs partly in response to the frontal gland secretion of the soldiers, which has been shown to act as a short-lived recruitment pheromone for other soldiers in addition to its immobilizing function on intruders, and partly in the response to the sternal gland pheromone. The latter pheromone, when combined with the communication of other information such as the presence of food, is responsible for exciting the workers. *See* PHEROMONE.

Colony and kin recognition. Communication of kinship and membership of a colony is known to involve an odor characteristic of the colony. This odor is a composite one with contributions from pheromones, the recognition odors of various castes, the environment, and food. Intruders, including conspecifics from other colonies, are recognized and attacked when their odor is detected and does not match the generalized acquired colony odor. Colony odor is important in identifying aliens, in intraspecific competition, and in enhancing the degree of relationship between colony members by favoring inbreeding. The maintenance of a high degree of inbreeding would in turn favor the evolution of sociality by kin selection in these social insects with normal diploid sex determination.

Pairing and mating. Following the dispersal flight, pairing of male and female imagos takes place without copulation; this occurs after the nuptial cell has been formed. Typically a female de-alate will take up a calling position with the tip of her abdomen pointing vertically, exposing the sternal gland. (In imagos this gland has been shown to produce an excitatory pheromone for males.) A male bumping into a calling female becomes very excited and will tend to follow another moving termite. Upon being bumped, the female lowers her abdomen and begins running and the male follows her in tandem. If the male loses antennal contact with the female she will stop and resume the calling position while he moves randomly; once more contacting the female the process is repeated. Excavation occurs when the pair reaches wood or a crevice. In certain African termites pairing has a distinct visual component: before losing their wings, the females alight on stems of grass and flutter their wings in a typical display which attracts the males.

Mating takes place in the nuptial cell and occurs at intervals throughout the life of the reproductives. Copulation occurs in a rear-to-rear position.

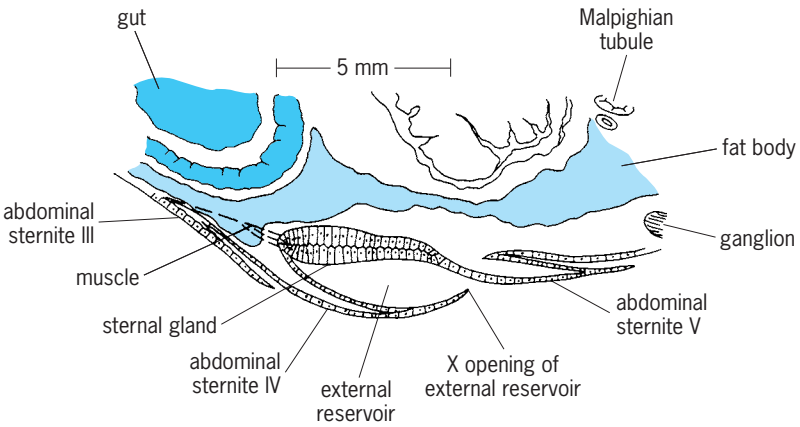

Fig. 3. Section through the abdomen of a nymph of *Zootermopsis nevadensis*, showing the sternal gland. (*After K. Krishna and F. Weesner, eds., Biology of Termites, vol. 1, 1969*)

Construction behavior. Building behavior has many manifestations, from the construction of simple galleries to the construction of the large bisymmetrical "magnetic" mounds made by *Amitermes meridionialis* in northern Australia. These mounds take up a north-south longitudinal axis which may occur through differential building activity in relation to a heat stimulus, as has been shown to occur with another termite species in the laboratory. The shape and orientation have evolved as a temperature regulating system in a termite that is unable to descend into the soil to avoid the heat (as other members of its genus do) because it is found in a habitat that is seasonally inundated.

For all termites, the basic act of building occurs as a response to a specific external stimulus. In some instances this primary building results in construction which, through its particular shape or chemical composition or both, acts as a new stimulus for further construction; this phenomenon is called "stigmergy." The building will stop when the results of building eventually reduce the original stimulus that initiated the building. This behavioral feedback applies in this situation and in alarm recruitment.

Evolution of sociality. Much has been proposed to explain the evolution of social behavior in the insects, and sociobiologists have heavily endorsed the idea that the haplodiploid form of sex determination found in the Hymenoptera is the basis of insect sociality. The termites, however, are diploid in both sexes and so destroy the generality of the hypothesis. Cytological studies, nonetheless, have inferred an exceptionally close genetic relationship between members of a termite colony, and this could partially explain the evolution of sociality and the concomitant altruistic behavior of nonsexual forms.

Probably several factors were involved in the evolution of sociality in termites, including the need to nurture the cellulose-degrading fauna of the hindgut. The similarity of the gut fauna of the wood-eating cockroach *Cryptocercus* to that of *Zootermopsis* supports the importance of diet specificity in the evolution of the social termites from a cockroach ancestor. *See* INSECTA; SOCIOBIOLOGY.

Alastair M. Stuart

Bibliography. K. Krishna and F. M. Weesner (eds.), *Biology of Termites*, vol. 1, 1969, vol. 2, 1970; A. M. Stuart, The role of pheromones in the initiation of foraging, recruitment and defence by soldiers of a tropical termite, *Nasutitermes corniger* (Motschulsky), *Chem. Senses*, 6:409–420, 1981.

Isopycnic

The line of intersection of an atmospheric isopycnic surface with some other surface, for instance, a surface of constant elevation or pressure. An isopycnic surface is a surface in which the density of the air is constant. Since specific volume is the reciprocal of density, isosteric surfaces coincide with isopycnic surfaces. On a surface of constant pressure, isopycnics coincide with isotherms, because on such a surface, density is a function solely of temperature. On a constant-pressure surface, isopycnics lie close together when the field is strongly baroclinic and are absent when the field is barotropic. *See* BAROCLINIC FIELD; BAROTROPIC FIELD; SOLENOID (METEOROLOGY).

Frederick Sanders; Howard B. Bluestein

Bibliography. J. R. Holton, *An Introduction to Dynamic meteorology*, 4th ed., 2004; G. Visconti, *Fundamentals of Physics and Chemistry of the Atmosphere*, 2001.

Isostasy

The application of Archimedes' principle to the layered structure of the Earth. The elevated topography of Earth is roughly equivalent to an iceberg that floats in the surrounding, denser water. Just as an iceberg extends beneath the exposed ice, the concept of isostasy proposes that topography is supported, or compensated, by a deep root. The buoyant outer shell of the Earth, the crust, displaces the denser, viscous mantle in proportion to the surface elevation. Isostasy implies the existence of a level surface of constant pressure within the mantle, the depth of compensation. Above this surface the mass of any vertical column is equal. Equal pressure at depth can also be achieved by varying density structure or by the regional deflection of the lithosphere. *See* ARCHIMEDES' PRINCIPLE; EARTH CRUST.

This theory was independently proposed to explain inconsistencies in geodetic observations made by P. Bouguer (1735) in Peru and G. Everest (1836) in India. Each set of observations depended critically on the determination of vertical from a suspended plumb bob. In the absence of mass anomalies or topography, the bob will orient itself in the local direction of gravity, perpendicular to the surface of the Earth. Early geodesists recognized that mountain ranges such as the Andes and the Himalayas should deflect the plumb bob toward the mountains. Subsequent, independent work by Pratt (1855) demonstrated that the deflection was less than predicted by calculations, indicating that the gravitational attraction of the topography was somehow diminished. One explanation that was advanced (Airy isostasy; **illus.** *a*) for this discrepancy advocated a large iceberglike root of constant density, the mirror image of the surface topography. An alternative view (Pratt isostasy; illus. *b*) suggested that the lower interface was horizontal and that the density of the mountains was less than the density of the surrounding level regions. These theories, advanced in 1855, served to reduce errors in geodetic surveys, but they were unconfirmed until the direct measurement of crustal thickness beneath mountain ranges using seismic refraction techniques began in the late 1920s. These studies of crustal thickness have confirmed that mountains are being underlain by thick roots. *See* EARTH, GRAVITY FIELD OF; GEODESY.

Local isostasy achieves equilibrium directly beneath a load by varying either the density or thickness of that mass column. This model attributes no

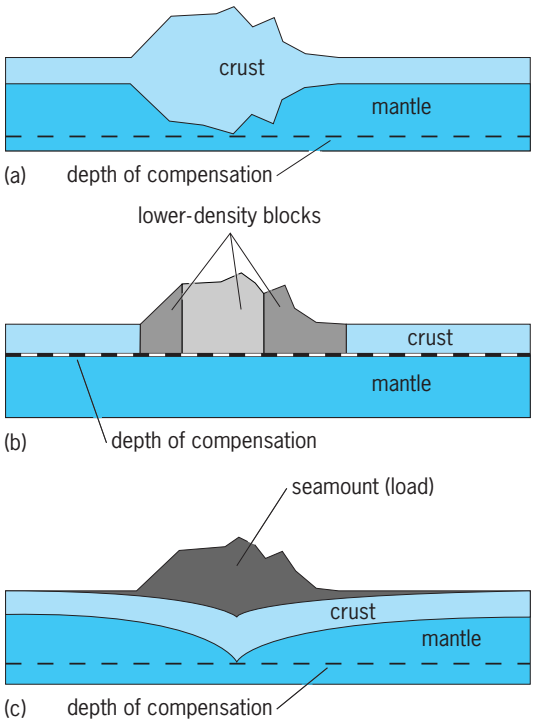

Three major modes of isostatic compensation. (*a*) Airy isostasy, where the crustal density is constant beneath both the elevated topography and the level region; a large root extends beneath the elevated topography, and the depth of compensation is at the base of the crust where the pressure is constant. (*b*) Pratt isostasy, where the density of the crust varies inversely with the height of the topography; the depth of compensation is at the base of the horizontal crust–mantle boundary. (*c*) Flexural or regional isostasy, where the crust has some strength and is deflected beneath the elevated topography; the depth of compensation is a horizontal surface beneath the lowest extent of the crust.

inherent strength to the crust and assumes that the mantle is a simple fluid, redistributing mass to minimize pressure differences at depth. From studies of seamounts, oceanic trenches, foreland basins, and glacial rebound, it has become known that the outer shell of the Earth is rigid, responding to loads over a region broader than the load itself, and that the mantle is a viscous fluid with a time-dependent response to loads.

The simplest method of examining the response of the Earth is to study an area influenced by a discrete load such as a seamount or a continental glacier. If local isostasy (Pratt and Airy) is applicable, the region surrounding the load will be horizontal, unaffected by the load. In contrast, if the lithosphere has finite strength and regional or flexural isostasy (illus. *c*) is applicable, the surrounding regions will be deflected down toward the load. Gravity, bathymetry, and seismic studies of the crust surrounding Hawaii and other seamounts have demonstrated that the crust is downwarped beneath seamounts. The implication of this regional response is that the oceanic lithosphere has some strength and that the Earth's outer shell behaves elastically.

Just as seamounts can be considered loads on oceanic crust, so can continental ice sheets be viewed as loads on continental crust. Since the glacial retreat, Fennoscandia and Canada, which were directly beneath continental ice sheets, have experienced rapid uplift, as recorded by the series of beaches now resting tens to hundreds of meters above sea level. Uplift in this region directly beneath the previously loaded region confirms the prediction that the Earth's surface is in hydrostatic equilibrium, but does not discriminate between the local and regional isostasy. Beyond the furthest extent of the ice sheet, raised beaches are also found, indicating that the crustal deformation associated with the ice loading was regional rather than local. This regional deflection provides compelling evidence for lithospheric strength and regional isostasy. *See* EARTH; LITHOSPHERE. Robin E. Bell; Bernard J. Coakley

Bibliography. C. M. R. Fowler, *The Solid Earth: An Introduction to Global Geophysics*, 2d ed., 2005; R. Sabadini, K. Lambeck, and E. Boschi (eds.), *Glacial Isostasy, Sea-Level and Mantle Rheology*, 1991; B. J. Skinner and S. C. Porter, *The Dynamic Earth: An Introduction to Physical Geology*, 5th ed., 2003; A. B. Watts, *Isostasy & Flexure of the Lithosphere*, 2001; A. B. Watts et al., A multichannel seismic study of lithospheric flexure across the Hawaii-Emperor seamount chain, *Nature*, 313(6015):105–111, 1985.

Isotach

A line along which the speed of the wind is constant. Isotachs are customarily represented on surfaces of constant elevation or atmospheric pressure, or in vertical cross sections. The closeness of spacing of the isotachs is indicative of the intensity of the wind shear on such surfaces. In the region of a jet stream the isotachs are approximately parallel to the streamlines of wind direction and are closely spaced on either side of the core of maximum speed. *See* JET STREAM; WIND. Frederick Sanders; Howard B. Bluestein

Isothermal chart

A map showing the distribution of air temperature (or sometimes sea surface or soil temperature) over a portion of the Earth's surface, or at some level in the atmosphere. On it, isotherms are lines connecting places of equal temperature. The temperatures thus displayed may all refer to the same instant, may be averages for a day, month, season, or year, or may be the hottest or coldest temperatures reported during some interval.

Maps of mean monthly or mean annual temperature for continents, hemispheres, or the world sometimes show values reduced to sea level to eliminate the effect of elevation in decreasing average temperature by about 3.3°F/1000 ft (5.9°C/1000 m; see **illus.**). Such adjusted or sea-level maps represent the effects of latitude, continents, and oceans in modifying temperature; but they conceal the effect of mountains and highlands on temperature

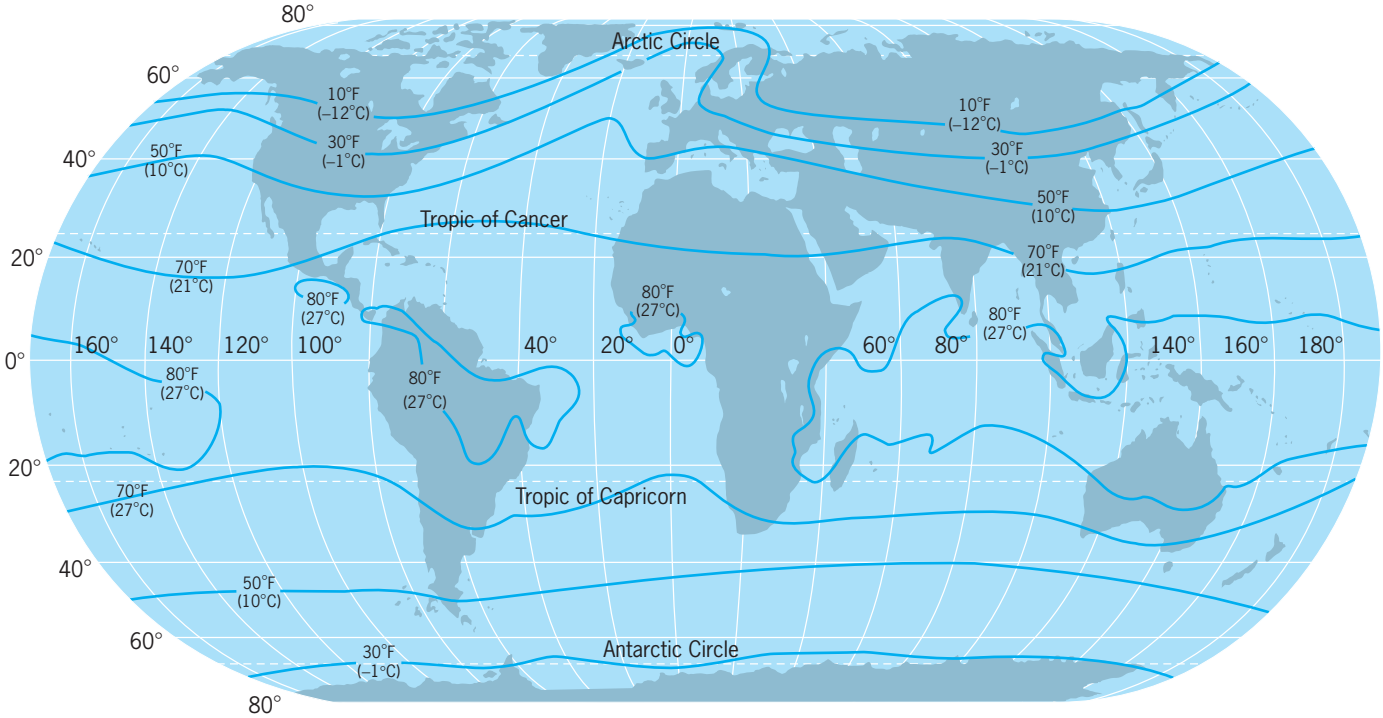

Chart showing selected January average isotherms. (*After D. K. Fellows, Our Environment: An Introduction to Physical Geography, 2d ed., John Wiley and Sons, 1980*)

distributions. The first isothermal chart, prepared by Alexander von Humboldt in 1817 for low and middle latitudes of the Northern Hemisphere, was the first use of isopleth methods to show the geographic distribution of a quantity other than elevation.

These maps are now varied in type and use. Isothermal charts are drawn daily in major weather forecasting centers; 5-day, 2-week, and monthly charts are used regularly in long-range forecasting; mean monthly and mean annual charts are compiled and published by most national weather services, and are presented in standard books on, for example, climate, geography, and agriculture. *See* AIR TEMPERATURE; TEMPERATURE INVERSION; WEATHER FORECASTING AND PREDICTION.

Arnold Court

Isothermal process

A thermodynamic process that occurs with a heat addition or removal rate just adequate to maintain constant temperature. The change in internal energy per mole, U, accompanying a change in volume in an isothermal process is given by Eq. (1), where T

$$U_2 - U_1 = \int_{V_1}^{V_2} \left[T\left(\frac{\partial P}{\partial T}\right)_V - P \right] dV \qquad (1)$$

is the temperature, P is pressure, and V is the volume per mole. The integral in Eq. (1) is zero for an ideal gas (which has the equation of state $PV = nRT$, where R is the gas constant and n the amount, so that the integrand vanishes), and for an incompressible condensed phase (solid or a liquid) for which the volume does not change with pressure. Thus, in both these cases, $U_2 = U_1$. For real gases and compressible liquids, the integral is nonzero, and the internal change is computed using a volumetric equation of state, that is, a relation between pressure, temperature, and volume.

The change in entropy, S, in an isothermal process is given by Eq. (2). For condensed phases with

$$S_2 - S_1 = \int_{V_1}^{V_2} \left(\frac{\partial P}{\partial T}\right)_V dV = -\int_{P_1}^{P_2} \left(\frac{\partial V}{\partial T}\right)_P dP \qquad (2)$$

a small coefficient of thermal expansion, $S_2 \approx S_1$, while for real fluids the equation of state is used in Eq. (2) to obtain the numerical value of the entropy change. Equation (3) is used for the isothermal entropy change of an ideal gas.

$$S_2 - S_1 = nR \ln \frac{V_2}{V_1} = nR \ln \frac{P_2}{P_1} \qquad (3)$$

In the case of a reversible isothermal process, the work done, W, by the fluid and the transferred heat to the fluid, Q, are given by Eqs. (4) and (5). The

$$W = -\int_{V_1}^{V_2} P dV = -Q + (U_2 - U_1) \qquad (4)$$

$$Q = T(S_2 - S_1) \qquad (5)$$

sign convention used here is that heat supplied to the fluid and work done on the fluid are positive. For a reversible isothermal process in an ideal gas, these equations reduce to Eq. (6). The comparable expres-

$$W = -Q = -nRT \ln \frac{V_2}{V_1} = +nRT \ln \frac{P_2}{P_1} \qquad (6)$$

sions for real fluids, which depend on the equation of state, can be found in thermodynamics textbooks.

Phase transitions (that is, first-order phase transitions such as solid to liquid, or liquid to vapor) in a pure material occur at constant temperature and constant pressure according to the Gibbs phase rule. The Gibbs free energy change for a first-order phase transition, ΔG, is identically zero at the transition temperature, the enthalpy and entropy changes, ΔH and ΔS respectively, are related as given in Eq. (7).

$$\Delta G = 0 = \Delta H - T\Delta S$$
$$\text{or} \quad \Delta S = \frac{\Delta H}{T} = \frac{\Delta U + P\Delta V}{T} \tag{7}$$

See CHEMICAL THERMODYNAMICS; GAS; PHASE RULE; PHASE TRANSITIONS; THERMODYNAMIC PROCESSES.

Stanley I. Sandler

Bibliography. S. I. Sandler, *Chemical and Engineering Thermodynamics*, 3d ed., 1999; R. E. Sonntag, C. Borgnakke, and G. J. Van Wylen, *Fundamentals of Thermodynamics*, 5th ed., 1997.

Isotone

One of two or more atoms which display a constant difference $A - Z$ between their mass number A and their atomic number Z. Thus, despite differences in the total number of nuclear constituents, the numbers of neutrons in the nuclei of isotones are the same. The numbers of naturally occurring isotones provide useful evidence concerning the stability of particular neutron configurations. For example, the relatively large number (six and seven, respectively) of naturally occurring 50- and 82-neutron isotones suggests that these nuclear configurations are especially stable. On the other hand, from the fact that most atoms with odd numbers of neutrons are anisotonic, one may conclude that odd-neutron configurations are relatively unstable. *See* NUCLEAR STRUCTURE.

Henry E. Duckworth

Isotope

One member of a (chemical-element) family of atomic species which has two or more nuclides with the same number of protons (Z) but a different number of neutrons (N). Because the atomic mass is determined by the sum of the number of protons and neutrons contained in the nucleus, isotopes differ in mass. Since they contain the same number of protons (and hence electrons), isotopes have the same chemical properties. However, the nuclear and atomic properties of isotopes can be different. The electronic energy levels of an atom depend upon the nuclear mass. Thus, corresponding atomic levels of isotopes are slightly shifted relative to each other. A nucleus can have a magnetic moment which can interact with the magnetic field generated by the electrons and lead to a splitting of the electronic levels. The number of resulting states of nearly the same energy depends upon the spin of the nucleus and the characteristics of the specific electronic level. *See* ATOMIC STRUCTURE AND SPECTRA; HYPERFINE STRUCTURE; ISOTOPE SHIFT.

Of the 108 elements reported thus far, 81 have at least one stable isotope whereas the others exist only in the form of radioactive nuclides. Some radioactive nuclides (for example, ^{115}In, ^{232}Th, ^{235}U, ^{238}U) have survived from the time of formation of the elements. Several thousand radioactive nuclides produced through natural or artificial means have been identified. *See* RADIOISOTOPE.

Of the 83 elements which occur naturally in significant quantities on Earth, 20 are found as a single isotope (mononuclidic), and the others as admixtures containing from 2 to 10 isotopes. Isotopic composition is mainly determined by mass spectroscopy. *See* MASS SPECTROSCOPE.

Nuclides with identical mass number (that is, $A = N + Z$) but differing in the number of protons in the nucleus are called isobars. Nuclides having different mass number but the same number of neutrons are called isotones. *See* ISOBAR (NUCLEAR PHYSICS); ISOTONE.

Nuclear stability. The stability of a nuclide is governed by its total energy E as given by the Einstein relation $E = Mc^2$, where M is the nuclidic mass and c is the velocity of light. If E is less than the combined energies of possible decay products, the nuclide will be stable. A major factor in determining stability is the relative strength of the nuclear force which acts to attract nucleons and the coulombic force (repulsive) which arises from the electric charge on the protons. Nuclides with an even number of protons or neutrons are prevalent in the **table**. Of the 287 nuclides tabulated, 168 are even-even (that is, an even number of both neutrons and protons), 110 are odd, and only 9 are odd-odd. This demonstrates the increased attraction of the nuclear force between pairs of nucleons of the same type (the pairing effect). Nuclides for which the number of either protons or neutrons (or both) comprises so-called magic numbers (for example, 8, 20, 50, 82, etc.) have increased stability. *See* NUCLEAR STRUCTURE.

Isotopic abundance. The term isotopic abundance refers to the isotopic composition of an element found in its natural terrestrial state. The isotopic composition for most elements does not vary much from sample to sample. This is true even for samples of extraterrestrial origin such as meteorites and lunar materials brought back to Earth by United States crewed and Soviet uncrewed missions. However, there are a few exceptional cases for which variations of up to several percent have been observed. There are several phenomena that can account for such variations, the most likely being some type of nuclear process which changes the abundance of one isotope relative to the others. For some of the lighter elements, the processes of distillation or chemical exchange between different chemical compounds could be responsible for isotopic differences. *See* NUCLEAR REACTION; RADIOACTIVITY.

The lead isotopes ^{206}Pb, ^{207}Pb, and ^{208}Pb are stable and are end products of naturally decaying ^{238}U,

Natural isotopic compositions of the elements

Atomic no.	Element symbol	Mass no.	Isotopic abundance, %
1	H*	1	99.985
		2	0.015
2	He*	3	0.000138
		4	99.999862
3	Li*	6	7.5
		7	92.5
4	Be	9	100
5	B*	10	19.9
		11	80.1
6	C*	12	98.90
		13	1.10
7	N*	14	99.634
		15	0.366
8	O	16	99.762
		17	0.038
		18	0.200
9	F*	19	100
10	Ne*†	20	90.51
		21	0.27
		22	9.22
11	Na	23	100
12	Mg	24	78.99
		25	10.00
		26	11.01
13	Al	27	100
14	Si	28	92.23
		29	4.67
		30	3.10
15	P	31	100
16	S*	32	95.02
		33	0.75
		34	4.21
		36	0.02
17	Cl	35	75.77
		37	24.23
18	Ar†	36	0.337
		38	0.063
		40	99.600
19	K	39	93.2581
		40	0.0117
		41	6.7302
20	Ca	40	96.941
		42	0.647
		43	0.135
		44	2.086
		46	0.004
		48	0.187
21	Sc	45	100
22	Ti	46	8.0
		47	7.3
		48	73.8
		49	5.5
		50	5.4
23	V	50	0.250
		51	99.750
24	Cr	50	4.35
		52	83.79
		53	9.50
		54	2.36
25	Mn	55	100
26	Fe	54	5.8
		56	91.72
		57	2.2
		58	0.28
27	Co	59	100
28	Ni	58	68.27
		60	26.10
		61	1.13
		62	3.59
		64	0.91
29	Cu	63	69.17
		65	30.83
30	Zn	64	48.6
		66	27.9
		67	4.1
		68	18.8
		70	0.6
31	Ga	69	60.1
		71	39.9
32	Ge	70	20.5
		72	27.4
		73	7.8
		74	36.5
		76	7.8
33	As	75	100
34	Se	74	0.9
		76	9.0
		77	7.6
		78	23.5
		80	49.6
		82	9.4
35	Br	79	50.69
		81	49.31
36	Kr†	78	0.35
		80	2.25
		82	11.6
		83	11.5
		84	57.0
		86	17.3
37	Rb	85	72.165
		87	27.835
38	Sr†	84	0.56
		86	9.86
		87	7.00
		88	82.58
39	Y	89	100
40	Zr	90	51.45
		91	11.27
		92	17.17
		94	17.33
		96	2.78
41	Nb	93	100
42	Mo	92	14.84
		94	9.25
		95	15.92
		96	16.68
		97	9.55
		98	24.13
		100	9.63
44	Ru	96	5.52
		98	1.88
		99	12.7
		100	12.6
		101	17.0
		102	31.6
		104	18.7
45	Rh	103	100
46	Pd	102	1.02
		104	11.14
		105	22.33
		106	27.33
		108	24.46
		110	11.72
47	Ag	107	51.839
		109	48.161
48	Cd	106	1.25
		108	0.89
		110	12.49
		111	12.80
		112	24.13
		113	12.22
		114	28.73
		116	7.49
49	In	113	4.3
		115	95.7
50	Sn	112	1.0
		114	0.7
		115	0.4
		116	14.7
		117	7.7
		118	24.3
		119	8.6
		120	32.4
		122	4.6
		124	5.6

Natural isotopic compositions of the elements (*cont.*)

Atomic no.	Element symbol	Mass no.	Isotopic abundance, %
51	Sb	121	57.3
		123	42.7
52	Te	71	39.9
		120	0.096
		122	2.60
		123	0.908
		124	4.816
		125	7.14
		126	18.95
		128	31.69
		130	33.80
53	I	127	100
54	Xe†	124	0.10
		126	0.09
		128	1.91
		129	26.4
		130	4.1
		131	21.2
		132	26.9
		134	10.4
		136	8.9
55	Cs	133	100
56	Ba	130	0.106
		132	0.101
		134	2.417
		135	6.592
		136	7.854
		137	11.23
		138	71.70
57	La	138	0.09
		139	99.91
58	Ce	136	0.19
		138	0.25
		140	88.48
		142	11.08
59	Pr	141	100
60	Nd	142	27.13
		143	12.18
		144	23.80
		145	8.30
		146	17.19
		148	5.76
		150	5.64
62	Sm	144	3.1
		147	15.0
		148	11.3
		149	13.8
		150	7.4
		152	26.7
		154	22.7
63	Eu	151	47.8
		153	52.2
64	Gd	152	0.20
		154	2.18
		155	14.80
		156	20.47
		157	15.65
		158	24.84
		160	21.86
65	Tb	159	100
66	Dy	156	0.06
		158	0.10
		160	2.34
		161	18.9
		162	25.5
		163	24.9
		164	28.2
67	Ho	165	100
68	Er	162	0.14
		164	1.61
		166	33.6
		167	22.95
		168	26.8
	Er	170	14.9
69	Ta	169	100
70	Yb	168	0.13
		170	3.05
		171	14.3
		172	21.9
		173	16.12
		174	31.8
		176	12.7
71	Lu	175	97.40
		176	2.60
72	Hf	174	0.16
		176	5.2
		177	18.6
		178	27.1
		179	13.74
		180	35.2
73	Ta	180	0.012
		181	99.988
74	W	180	0.13
		182	26.3
		183	14.3
		184	30.67
		186	28.6
75	Re	185	37.40
		187	62.60
76	Os†	184	0.02
		186	1.58
		187	1.6
		188	13.3
		189	16.1
		190	26.4
		192	41.0
77	Ir	191	37.3
		193	62.7
78	Pt	190	0.01
		192	0.79
		194	32.9
		195	33.8
		196	25.3
		198	7.2
79	Au	197	100
80	Hg	196	0.15
		198	10.1
		199	17.0
		200	23.1
		201	13.2
		202	29.65
		204	6.8
81	Tl	203	29.524
		205	70.467
82	Pb†	204	1.4
		206	24.1
		207	22.1
		208	52.4
83	Bi	209	100
90	Th	232	100
92	U*	234	0.0055
		235	0.7200
		236	99.2745

*Isotopic composition may vary with sample depending upon geological or biological origin.
†Isotopic composition may vary with sample because some of the isotopes may be formed as a result of radioactive decay or nuclear reactions.

^{235}U, and ^{232}Th, respectively, whereas ^{204}Pb is not produced by any long-lived decay chain. Thus, the isotopic composition of lead samples will depend upon their prior contact with thorium and uranium. The potassium isotope ^{40}K has a half-life of 1.28×10^9 years and decays by beta-ray emission to ^{40}Ar and by electron capture to ^{40}Ca, which are both stable. This can cause the argon in potassium-bearing minerals to differ in isotopic abundance from that found in air. It is possible to determine the age of rocks by measuring the ratio of their $^{40}K/^{40}Ar$ content. This ratio technique can also be used for rock samples which bear other long-lived, naturally occurring isotopes such as ^{87}Rb (rubidium), thorium, and uranium. *See* LEAD ISOTOPES (GEOCHEMISTRY); ROCK AGE DETERMINATION.

An interesting example of anomalous isotopic compositions has been observed in the Oklo uranium deposit in Gabon (western Africa). Based upon extensive research, it has been concluded that this is the site of a natural chain reaction that took place about 1.8×10^9 years ago. Much of the uranium in this formation has been depleted of the fissionable isotope ^{235}U. The isotopic composition of some of the other elements found at or near this deposit has also been altered as a result of fission, neutron absorption, and radioactive decay.

Use of separated isotopes. The areas in which separated (or enriched) isotopes are utilized have become fairly extensive, and a partial list includes nuclear research, nuclear power generation, nuclear weapons, nuclear medicine, and agricultural research. Various methods are employed to prepare separated isotopes. Mass spectroscopy is used in the United States and Russia to prepare inventories of separated stable isotopes. Distillative, exchange, and electrolysis processes have been used to produce heavy water (enriched in deuterium, ^2H), which is used as a neutron moderator in some reactors. The uranium enrichment of ^{235}U, which is used as a fuel in nuclear reactors, has mainly been accomplished by using the process of gaseous diffusion in uranium hexafluoride gas in very large plants. This method has the disadvantage of requiring large power consumption. Techniques which can overcome this problem and are finding increasing favor include centrifugal separation and laser isotope separation. *See* ISOTOPE SEPARATION; NUCLEAR REACTOR.

For many applications there is a need for separated radioactive isotopes. These are usually obtained through chemical separations of the desired element following production by means of a suitable nuclear reaction. Separated radioactive isotopes are used for a number of diagnostic studies in nuclear medicine, including the technique of positron tomography. *See* NUCLEAR MEDICINE.

Studies of metabolism, drug utilization, and other reactions in living organisms can be done with stable isotopes such as ^{13}C, ^{15}N, ^{18}O, and ^2H. Molecular compounds are "spiked" with these isotopes, and the metabolized products are analyzed by using a mass spectrometer to measure the altered isotopic ratios. The use of separated isotopes as tracers for determining the content of a particular element in a sample by dilution techniques had broad applicability. *See* ISOTOPE DILUTION TECHNIQUES; RADIOISOTOPE (BIOLOGY).

Atomic mass. Atomic masses are given in terms of an internationally accepted standard which at present defines an atomic mass unit (amu) as exactly equal to one-twelfth the mass of a neutral atom of ^{12}C in its electronic and nuclear ground states. On this scale the mass of a ^{12}C atom is equal to 12.0 amu. Atomic masses can be determined quite accurately by means of mass spectroscopy, and also by the use of nuclear reaction data with the aid of the Einstein mass-energy relation. Atomic weight is the average mass per atom per amu of the natural isotopic composition of an element. Practically all atomic weights are now calculated by taking the sum of the products of the fractional isotopic abundances times their respective masses in amu. The atomic weight of a neutral atom of a nuclide is nearly equal to an integer value A, because the mass of both the neutron and proton is almost identically 1 amu. *See* ATOMIC MASS.

Daniel J. Horen

Bibliography. A. P. Dickin, *Radiogenic Isotope Geology*, 1994, reprint 1997; H. E. Duckworth, *Mass Spectroscopy*, 2d ed., 1990; K. Lajtha and R. Michener, *Stable Isotopes in Ecology*, 1994; D. S. Schimel, *Theory and Application of Tracers*, 1993; H. P. Taylor, J. R. O'Neill, and I. R. Kaplan, *Stable Isotope Geochemistry*, 1991; R. R. Wolfe, *Radioactive and Stable Isotope Tracers in Biomedicine*, 1992.

Isotope dilution techniques

Analytical techniques that involve the addition to a sample of an isotopically labeled compound. Soon after the discovery of the stable heavy isotopes of hydrogen, carbon, nitrogen, and oxygen, their value in analytical chemistry was recognized. Stable isotopes were particularly useful in the analysis of complex mixtures of organic compounds where the isolation of the desired compound with satisfactory purity was difficult and effected only with low or uncertain yields. The addition of a known concentration of an isotopically labeled compound to a sample immediately produces isotope dilution if the particular compound is present in the sample. After thorough mixing of the isotopically labeled compound with the sample, any technique that determines the extent of the isotopic dilution suffices to establish the original concentration of the compound in the mixture. Isotope dilution techniques exploit the difficulty in the separation of isotopes, with the isotopically labeled "spike" following the analytically targeted compound through a variety of separation procedures prior to isotopic analysis. *See* ISOTOPE.

The technique depends on the availability of a stable or radioisotope diluent with isotope abundance ratios differing markedly from those of the naturally occurring elements. With monoisotopic elements, such as sodium or cesium, radioactive elements of sufficiently long life can be used in isotope dilution

techniques. In 1954 M. G. Ingraham identified 68 elements that were analyzable by isotope dilution techniques. The list has since been expanded considerably.

Biochemical applications. The original applications of isotope dilution were by biochemists interested in complex mixtures of organic compounds. In these studies care had to be taken to ensure the stability of the labeled compound and its resistance to isotopic exchange reactions. Nitrogen-15–labeled glycine for example, could be used to determine glycine in a mixture of amino acids obtained from a protein. Deuterium-labeled glycine could not be used reliably if the deuterium isotopes were attached to the glycine amino or carboxyl group, because in these locations deuterium is known to undergo exchange reactions with hydrogens in the solvent or in other amino acids. Deuterium is very useful in elemental isotopic analysis where total hydrogen or exchangeable hydrogen concentrations are desired. *See* BIOCHEMISTRY; DEUTERIUM.

High-sensitivity applications. Applications of isotope dilution techniques have also been found in geology, nuclear science, and materials science. These applications generally focus on the very high sensitivity attainable with these techniques. Isotopes of argon, uranium, lead, thorium, strontium, and rubidium have been used in geologic age determinations of minerals and meteorites. Taking the estimated error as a measurement of sensitivity, isotopic dilution analyses of uranium have been done down to 4 parts in 10^{12} and on thorium to 8 parts in 10^{9}. Studies in geology and nuclear science require the determination of trace amounts of radiogenic products. If the half-life and decay scheme of the parent nuclide is known, then isotopic dilution determinations of parent and daughter isotopes provide a basis for the calculation of the age of the sample. If the age or history of the sample is known, then determination of the trace concentrations of isotopes provides information on pathways of nuclear reactions. *See* RADIOACTIVITY; ROCK AGE DETERMINATION.

An example of the latter type of investigation is the determination of yields of cesium (Cs) from bombardment of gold or uranium with very high-energy protons. The use of radioactive ^{131}Cs with techniques of high-sensitivity mass spectrometry permitted the determination of stable and long-lived Cs isotopes in concentrations of less than 10^{-15} g. The fundamental limitation in the sensitivity of isotope dilution as a trace analytical technique is in contamination of reagents used in the analytical procedure. In the case of the study on cesium cited above, reagent water prepared by triple distillation in quartz columns was found to have approximately 10^{-15} g of natural ^{133}Cs per cubic centimeter.

Studies using isotope dilution techniques have dealt with trace analysis of toxic materials in biological systems. Uranium-233 has been used as an isotopic spike in the determination of uranium in bovine liver and in oyster tissue. Assay of 10^{-10} g of uranium with a precision of better than 0.5% was achieved. Contamination of reagents remains the limiting factor in sensitivity, with water assayed at the 4.5–22-femtogram level by isotope dilution. Uranium is an example of an element that is a potent biotoxin as well as a radiological hazard. Its assay in trace concentrations is of considerable importance in toxicology. *See* RADIATION BIOLOGY; TOXICOLOGY.

Lewis Friedman

Bibliography. J. H. Chen and G. J. Wasserburg, Isotopic determination of uranium in picomole and subpicomole quantities, *Anal. Chem.*, 53:2060–2067, 1981; M. G. Ingraham, Stable isotope dilution as an analytical tool, *Annu. Rev. Nucl. Sci.*, 4:81–92, 1954; J. A. Jonckheere, A. P. DeLeenheer, and H. L. Steyaert, Statistical evaluation of calibration curve nonlinearity in isotope dilution gas chromatography/mass spectrometry, *Anal. Chem.*, 55:153–155, 1983; J. F. Pickup and K. McPherson, Theoretical considerations in stable isotope dilution mass spectrometry for organic analysis, *Anal. Chem.*, 48:1885–1890, 1976; D. C. Reamer and C. Veillon, Determination of selenium in biological materials by stable isotope dilution gas chromatography-mass spectrometry, *Anal. Chem.*, 53:2166–2169, 1981.

Isotope separation

The physical separation of different isotopes of an element from one another. The different isotopes of an element as it occurs in nature may have similar chemical properties but completely different nuclear reaction properties. Therefore, nuclear physics and nuclear energy applications often require that the different isotopes be separated. However, similar physical and chemical properties make isotope separation by conventional techniques unusually difficult. Fortunately, the slight mass difference of isotopes of the same element makes separation possible by using especially developed processes, some of which involve chemical industry distillation concepts. *See* ISOTOPE.

Isotope separation depends on the element involved and its industrial application. Uranium isotope separation has by far the greatest industrial importance, because uranium is used as a fuel for nuclear power reactors, particularly the predominant light-water reactor type. The two main isotopes found in nature are ^{235}U and ^{238}U, which are present in weight percentages (w/o) of 0.711 and 99.283, respectively. Trace amounts of ^{234}U are also present in natural uranium. In order to be useful as a fuel for light-water reactors, the weight percentage of ^{235}U must be increased to between 2 and 5. The process of increasing the ^{235}U content is known as uranium enrichment, and the process of enriching is referred to as performing separative work. *See* NUCLEAR FUELS; NUCLEAR REACTOR; URANIUM.

The production of heavy water is another example of isotope separation. Because of its exceptional neutron slowing-down characteristics, heavy water is used as a moderator in some types of natural and very low enriched uranium-fueled nuclear reactors. Heavy water is obtained by isotope separation of

light hydrogen (^1H) and heavy hydrogen (^2H) in natural water. Heavy hydrogen is usually referred to as deuterium (D). All natural waters contain ^1H and ^2H, in concentrations of 99.985 and 0.015 w/o, respectively, in the form of H_2O and D_2O (deuterium oxide). Isotope separation increases the concentration of the D_2O, and thus the purity of the heavy water. *See* DEUTERIUM; HEAVY WATER.

The development of laser isotope separation technology provided a range of potential applications. For example, enrichment of natural gadolinium in the isotope ^{157}Gd could result in the production of very high quality material for reactivity control of light-water reactors. Other isotope separation applications range from space-flight power sources (^{238}Pu) to medical magnetic resonance imaging (^{13}C) and medical research (^{15}O).

Separation technologies. The isotope separation process that is best suited to a particular application depends on the state of technology development as well as on the mass of the subject element and the quantities of material involved. Processes such as electromagnetic separation which are suited to research quantities of material are generally not suited to industrial separation quantities. However, the industrial processes that are used, gaseous diffusion and gas centrifugation, are not suited to separating small quantities of material.

Uranium isotope separation by either the gaseous diffusion or gas centrifuge process requires that the uranium be in the form of uranium hexafluoride (UF_6) gas. Both separation processes take advantage of the slight difference in the mass of the ^{235}U and ^{238}U atoms. For gaseous diffusion, the separator stage may consist of a single gas diffuser, gas compressor, gas cooler, compressor drive motor, and associated piping and valving, while a gas-centrifuge stage may consist of many centrifuge machines and associated hardware operating in parallel. In either case, the unit or units operating in parallel on material of the same mole fraction form a stage. Each stage has feed, enriched product, and depleted waste streams and a characteristic stage-separation efficiency factor. Stages connected in series to obtain progressive separation are referred to as cascades.

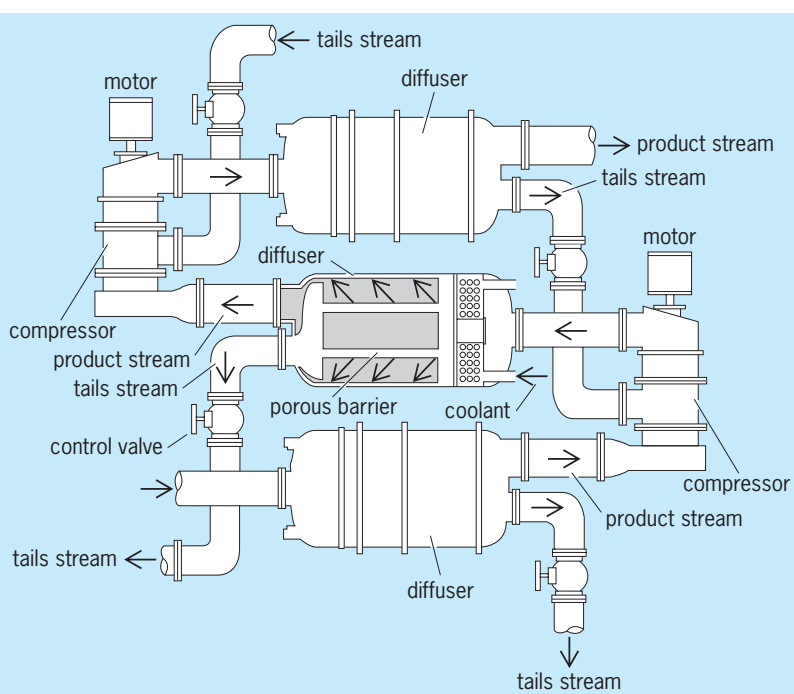

Fig. 1. Gaseous diffusion stage. (*After U.S. Atomic Energy Commission Report ORO-684, January 1972*)

In the course of accomplishing the separation of ^{235}U and ^{238}U in each of the separation stages (either diffusion or centrifuge), large amounts of uranium hexafluoride gas are recirculated. Application of chemical engineering theory using abundance ratios, reflux ratios, and flow-rate considerations, analogous to distillation theory, yields a term for the total flow of uranium in the separation cascade consisting of the product of two terms. One term is a function of the stage-separation factor, which is dependent on the molecular weights of the isotopes being separated, and the other is a function of the flow rates and isotopic composition of the feed, product, and waste. This second term is called separative work (SW), and is usually referred to in the following units: either kgU SWU or SWU in the United States, and either kgSW or tonnes SW in other countries.

Gaseous diffusion. In the gaseous diffusion process, uranium hexafluoride gas is circulated under pressure through specially designed equipment containing porous barriers. The barriers contain hundreds of millions of pores per square inch, with the average diameter of the pores being only a few millionths of an inch. The kinetic energies of the uranium hexafluoride molecules depend only on the temperature, which is the same within each stage. Since the kinetic energy of a molecule is proportional to the product of its mass and the square of its speed, the lighter molecules (^{235}U) have higher speeds causing them to strike the barrier more frequently than the heavy molecules (^{238}U), and therefore they have a higher probability of penetration. Thus, the lighter molecules (^{235}U) diffuse (or more correctly, effuse) through the barriers faster than the heavy molecules (^{238}U), with the result that the downstream side of the barrier is enriched in the lighter gas and the upstream side is depleted in the lighter gas. *See* DIFFUSION; KINETIC THEORY OF MATTER.

The stage-separation factor for the gaseous diffusion process depends only on the molecular weights of $^{235}U^{19}F_6$ and $^{238}U^{19}F_6$, which are 349 and 352, respectively; it is equal to 1.0043. Because the maximum possible value of the separation factor is so close to unity, a gaseous diffusion plant typically needs a very large number of diffuser-compressor-cooler units (**Fig. 1**) connected in series in order to obtain sufficiently enriched product. For example, approximately 1400 stages are needed to produce the 4 w/o ^{235}U that is typical of that used in light-water reactors. The throughput in each stage of a large diffusion plant can require large compressors driven by motors of several thousand horsepower (several megawatts).

Gas centrifugation. The gas centrifuge process has the advantage that its isotope separation capability for uranium is proportional to the absolute mass difference of the uranium hexafluoride molecules, rather than the relative difference, which is the case for gaseous diffusion. The separation effect is achieved by pressure diffusion through the gravitational field generated within the centrifuge rotor. The concentration of the heavy molecule tends to increase toward the rotor wall, while that of the lighter molecule tends to increase at the center of the rotor.

The radial separation factor of a gas centrifuge machine depends on the angular velocity of rotation, the radius of the rotor wall, and the absolute gas temperature. For a rotor radius of 6 cm (2.4 in.), a rotational speed of 40,000 revolutions per minute (corresponding to a peripheral speed of 250 m/s or 800 ft/s), and a gas temperature of 300 K (80°F), the separation factor is 1.0387; the difference of this separation factor from unity is nine times that of the separation factor for the gaseous diffusion process. The separation capability is approximately inversely proportional to the absolute temperature, and operating at a low or at least moderate temperature is therefore desirable. However, an adequate temperature margin must always be maintained to ensure that the uranium hexafluoride process gas does not condense.

In a practical centrifuge, an internal countercurrent circulation flow is introduced (**Fig. 2**). This flow has the effect of generating an axial concentration difference which is a factor larger than the elementary radial separation factor discussed above. Thus, even a centrifuge of very modest speed has a separation factor greatly exceeding that of the gas diffusion process. The countercurrent flow also permits the enriched and depleted fractions to be removed at the end of the rotor. In practice, centrifuge throughput has a strong dependence on the rotor peripheral speed and length. However, rotor material composition limits peripheral speed since the rotor will break if induced stresses exceed the tensile strength of the material. *See* CENTRIFUGATION.

Laser separation. Three experimental laser isotope separation technologies for uranium are the atomic vapor laser isotope separation (AVLIS) process, the uranium hexafluoride molecular laser isotope separation (MLIS) process, and the separation of isotopes by laser excitation (SILEX) process, which was announced in Australia in the early 1990s. The AVLIS process, which is more experimentally advanced than the MLIS and SILEX processes, exploits the fact that the different electron energies of ^{235}U and ^{238}U absorb different colors of light (that is, different wavelengths). In the AVLIS process, lasers are tuned to emit a combination of colors that will be absorbed only by a ^{235}U atom, which then emits an electron. This photoionization leaves the ^{235}U atom with a net positive charge, allowing the atom, now an ion, to be selectively separated by using electromagnetic fields. AVLIS technology is inherently more efficient than either the gaseous diffusion or gas centrifuge processes. It can enrich natural uranium to 4 w/o ^{235}U in a single step. In the United States, the AVLIS process is being developed to eventually replace the gaseous diffusion process for commercially enriching uranium. *See* PHOTOIONIZATION.

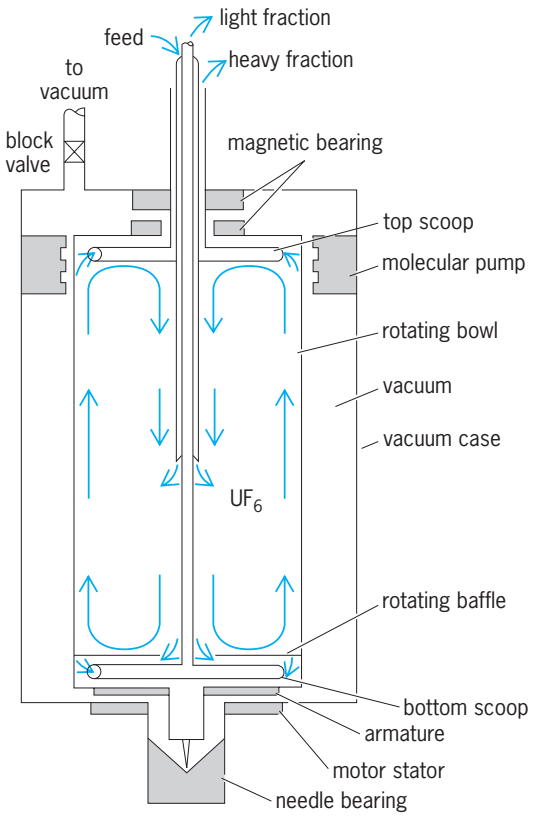

Fig. 2. Gas centrifuge machine. (*After M. Benedict, T. H. Pigford, and H. W. Levi, Nuclear Chemical Engineering, 2d ed., McGraw-Hill, 1981*)

The AVLIS process includes two major component systems: a laser system and a separator system (**Fig. 3**). The laser system uses two types of lasers: dye lasers that generate the light used for photoionization of the uranium, and copper-vapor lasers that energize the dye lasers. Powerful green-yellow light from electrically driven copper-vapor lasers is converted to red-orange light in the dye laser. This red-orange light is tuned to the precise colors that are absorbed by ^{235}U but not by ^{238}U. This laser–pumped-laser scheme is required because dye lasers cannot be powered directly by electricity. *See* LASER.

In the separator system, uranium metal is melted and vaporized by means of an electron beam that

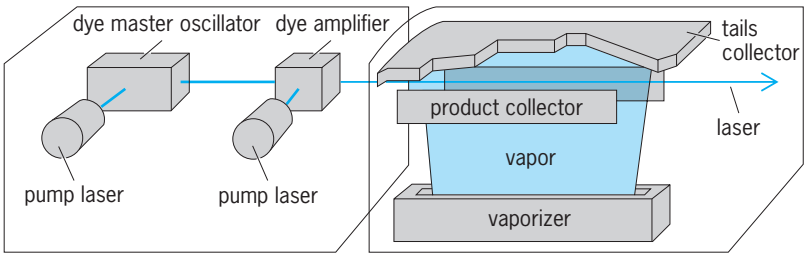

Fig. 3. Atomic vapor laser isotope separation (AVLIS) process. (*After Lawrence Livermore National Laboratory Report LLL-TB-072, 1985*)

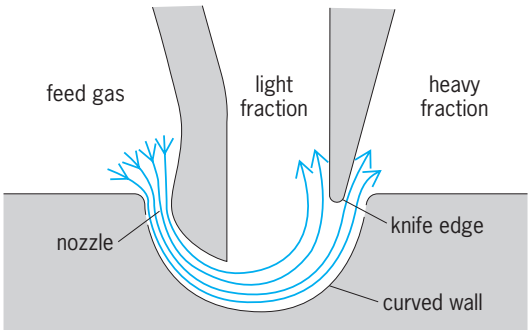

Fig. 4. Cross section of slit-shaped separation nozzle system with streamlines of the gas flow. (*After M. Benedict, T. H. Pigford, and H. W. Levi, Nuclear Chemical Engineering, 2d ed., McGraw-Hill, 1981*)

creates an atomic vapor stream of ^{235}U and ^{238}U. The tuned dye laser beams are passed through the vapor stream, where they photoionize the ^{235}U atoms. An electromagnetic field deflects the selected photoions to the product collector, where they condense. The ^{238}U atoms, which are unaffected by the color-selective laser beams, pass through the product collector to condense on the tails collector. The enriched uranium liquid-metal condensate flows out of the separator to be cast and stored in solid metallic form, or to be converted to uranium hexafluoride or oxide as required.

Nozzle process. In the German jet nozzle or Becker process, which is a laboratory process, uranium hexafluoride gas is diluted in a helium or hydrogen carrier gas, and fed by compressor into the nozzle on one side of a 1–2-m long (3–6-ft) semicircular slot in which there is a deflector blade on the other side to separate enriched and depleted streams (**Fig. 4**). The semicircular wall of the slot provides a partial centrifugal effect such that the heavier isotope tends to stay closer to the wall compared to the lighter isotope. The radius of the semicircular slot is of the order of 0.1 mm and is formed by stacking photoetched foils.

Helikon process. The South African Helikon process is aerodynamically similar to the Becker process from which it was developed. However, it is believed to use a vortex tube instead of a slot and opposing gas flows instead of a deflector blade for flow separation. South Africa's small power-intensive commercial Helikon uranium enrichment plant was shut down in the mid-1990s because of its high cost of operation.

Thermal diffusion. The separation of isotopes by thermal diffusion is based on the fact that when a temperature gradient is established in a mixture of uniform composition, one component will concentrate near the hot region and the other near the cold region. Thermal diffusion is carried out in the annular space between two vertical concentric pipes, the inner one heated and the outer one cooled. Thermal diffusion was used in the United States during World War II for uranium, but it is not efficient for industrial separation. It has been used to separate small quantities of isotopes for research purposes.

Chemical exchange. The chemical exchange process has proved to be the most efficient for separating isotopes of the lighter elements. This process is based on the fact that if equilibrium is established between, for example, a gas and a liquid phase, the composition of the isotopes will be different in the two phases. Thus, if hydrogen gas is brought into equilibrium with water, the ratio of deuterium to hydrogen is several times greater in the water than in the hydrogen gas. By repeating the process in a suitable cascade, it is possible to effect a substantial separation of the isotopes with a relatively small number of stages. The chief use of chemical exchange is in the large-scale production of heavy water. Chemical exchange has also been used for the large-scale separation of other isotopes. For example, the isotopes of boron have been separated by fractional distillation of the boron trifluoride–dimethyl ether complex. *See* CHEMICAL EQUILIBRIUM.

Distillation. The separation of isotopes by distillation is much less efficient than separation by other methods. Distillation was used during World War II to produce heavy water, but the cost was high and the plants are no longer in existence. *See* DISTILLATION.

Electrolysis. Electrolysis of water is the oldest large-scale method of producing heavy water. Under favorable conditions, the ratio of hydrogen to deuterium in the gas leaving a cell in which water is electrolyzed is eight times the ratio of these isotopes in the liquid. Because it is electricity-intensive, the process is very expensive unless low-cost hydroelectric generating capacity is available. Electrolysis is used in the United States only as a finishing step to concentrate heavy water to final-product specifications. *See* ELECTROLYSIS.

Electromagnetic process. Electromagnetic separation was the method which was first used to prove existence of isotopes. The mass spectrometer and mass spectrograph are still widely used as research tools. In the electromagnetic process, vapors of the material to be analyzed are ionized, are accelerated in an electric field, and enter a magnetic field which causes the ions to be bent in a circular path. Since the light ions have less momentum than the heavy ions, they are bent through a circle of smaller radius, and the two isotopes can be separated by placing collectors at the proper location. The large-capacity electromagnetic uranium isotope separators developed in the United States during World War II were called calutrons. *See* CHARGED PARTICLE OPTICS; MASS SPECTROSCOPE.

Julian J. Steyn

Bibliography. M. Benedict, T. H. Pigford, and H. W. Levi, *Nuclear Chemical Engineering*, 2d ed., 1981; Edison Electric Institute, *EEI Enrichment Handbook*, Rep. NFC-90-001, 1990.

Isotope shift

A small difference between the different isotopes of an element in the transition energies corresponding to a given spectral-line transition. For an electronic spectral-line transition between two energy levels *a*

and b in an atom or ion with atomic number Z, the small difference $\Delta E_{ab} = E_{ab}(A') - E_{ab}(A)$ in the transition energy between isotopes with mass numbers A' and A is the isotope shift. It consists largely of the sum of two contributions, the mass shift (MS) and the field shift (FS), also called the volume shift. The mass shift is customarily divided into a normal mass shift (NMS) and a specific mass shift (SMS); each is proportional to the factor $(A' - A)/A'A$. The normal mass shift is a reduced mass correction that is easily calculated for all transitions. The specific mass shift is produced by the correlated motion of different pairs of atomic electrons and is thus absent in one-electron systems.

It is generally difficult to calculate precisely the specific mass shift, which may be 30 times larger than the normal mass shift for some transitions. The field shift is produced by the change in the finite size and shape of the distribution of nuclear charge (protons) when the number of neutrons $N = A - Z$ varies with Z fixed. Since electrons whose orbits penetrate the nucleus are influenced most, *S-P* and *P-S* transitions have the largest field shift. Measurements of isotope shifts, coupled with accurate theoretical calculations of mass shifts and relativistic and other corrections, can be used to determine the change $\delta(r_c^2)$ in the mean-square nuclear charge radius r_c^2 as N varies with Z fixed.

For very light elements, $Z \lesssim 37$, the mass shift dominates the field shift. For $Z = 1$, the 0.13 nanometer shift (0.02%) in the red Balmer line led to the discovery of deuterium, the $A = 2$ isotope of hydrogen. For medium-heavy elements, $38 \lesssim Z \gtrsim 57$, the mass shift and field shift contributions to the isotope shift are comparable. For heavier elements, $Z \gtrsim 58$, the field shift dominates the mass shift. A representative case is shown in the **illustration**. *See* ATOMIC STRUCTURE AND SPECTRA; DEUTERIUM; RYDBERG CONSTANT.

When isotope shift data have been obtained for at least two pairs of isotopes of a given element, a graphical method introduced by W. H. King in 1963 can be used to evaluate quantitatively the separate contributions of the mass shift and the field shift. Experimentally determined field shifts can be used to test theoretical models of nuclear structure, shape, and multipole moments. Experimentally determined specific mass shifts can be used to test detailed theories of atomic structure and relativistic effects. *See* NUCLEAR MOMENTS; NUCLEAR STRUCTURE.

Some atomic transitions exhibit hyperfine structure, which generally is dominated by the interaction of electron orbital and nuclear magnetic dipole moments, the latter existing for isotopes with nonzero nuclear spin. Accurate measurements of magnetic hyperfine structures exhibiting the so-called hyperfine anomaly (Bohr-Weisskopf effect) gives information about how the distribution of nuclear magnetization changes as N varies with Z fixed. *See* HYPERFINE STRUCTURE.

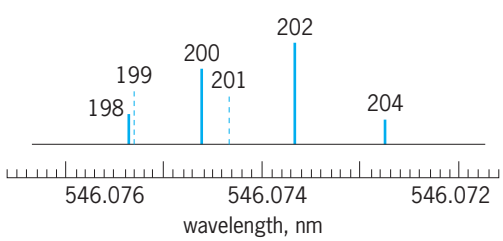

Some isotope shifts in the green line of mercury, $Z = 80$. In this heavy element, the contribution of the field shift is much larger than that of the mass shift.

Experimental techniques that have greatly increased both the amount and the precision of isotope shift data include energetic nuclear reactions for online production of isotopes with half-lives as short as milliseconds, and subsequent high-resolution laser-spectroscopic studies of ions or atoms containing them. Laser spectroscopy of collimated, fast-moving beams of ions or atoms containing the isotopes uses techniques that exploit (rather than are hampered by) the Doppler shift caused by the motion of the ions or atoms through the laser light. Techniques developed for cooling and trapping of neutral atoms or highly charged ions with electromagnetic fields have been extended to short-lived isotopes and allow as few as one trapped atom or ion to be detected. Isotope shifts measured at low Z have tended to focus on precise tests of atomic few-body dynamics and determination of (r_c^2) for short-lived, exotic "halo nuclei" such as ^{6}He and ^{11}Li. Isotope shifts measured with laser or x-ray methods at high Z have tended to focus on determination of $\delta(r_c^2)$ and tests of atomic structure calculations where many-body, relativistic, quantum-electrodynamical (QED), and other interactions are significant and difficult to treat. These must be well understood for atomic parity-non-conservation (PNC) experiments, which provide stringent and useful tests of electroweak and other theories when different isotopes of high-Z atoms are used. *See* ATOMIC THEORY; DOPPLER EFFECT; ELECTROWEAK INTERACTION; LASER; LASER COOLING; LASER SPECTROSCOPY; PARITY (QUANTUM MECHANICS); PARTICLE TRAP; QUANTUM ELECTRODYNAMICS; RELATIVISTIC QUANTUM THEORY.

Isotope shift data have also been obtained for x-ray transitions of electrons in inner atomic shells and of muons in muonic atoms. *See* LEPTON. Peter M. Koch

Bibliography. J. Bauche and R. J. Campeau, Recent advances in the theory of atomic isotope shift, *Advan. At. Mol. Phys.*, 12:39–86, 1976; J. C. Berengut, V. V. Flambaum, and M. G. Kozlov, Calculation of relativistic and isotope shifts in Mg I, *Phys. Rev. A*, 72:044501, 2005; S. R. Elliott, P. Beiersdorfer, and M. H. Chen, Trapped-ion technique for measuring the nuclear charge radii of highly charged radioactive isotopes, *Phys. Rev. Lett.*, 76:1031–1034, 1995; E. Gomez, L. A. Orozco, and G. D. Sprouse, Spectroscopy with trapped francium: Advances and perspectives for weak interaction studies, *Rep. Prog. Phys.*, 69:79–118, 2006; R. Sánchez et al., Nuclear charge radii of ^{9,11}Li: The influence of halo electrons, *Phys. Rev. Lett.*, 96:033002, 2006; L.-B. Wang et al., Laser spectroscopic determination of the ^{6}He nuclear charge radius, *Phys. Rev. Lett.*, 93:142501, 2004.

Isotopic irradiation

The uses of the radiation emitted by radioactive isotopes (radioisotopes), principally in industry and medicine. Although this article deals primarily with gamma irradiation, many other forms of radiation from many different sources are often used for irradiation purposes. The use of radioactive isotopes to irradiate various materials can be divided into a number of major areas, including industrial, medical, and research (**Table 1**). For most applications the radioactive isotope is placed inside a capsule, but for some medical applications it is dispersed in the material to be irradiated. *See* RADIOISOTOPE.

The radiation from radioisotopes produces essentially the same effect as the radiation from electron linear accelerators and other high-voltage particle accelerators, and the choice of which to use is based primarily on convenience and cost. Although the radioisotope radiation source does not require the extensive and complex circuitry necessary for a high-voltage radiation source, its radiation is always present and requires elaborate shielding for health protection and specialized mechanisms for bringing the irradiated objects into and out of the radiation beam. Further, the radiation output decreases with time according to the half-life of the radioisotope, which must therefore be replaced periodically. *See* PARTICLE ACCELERATOR.

Industrial applications. The two main radioisotopes for industrial processing are cobalt-60 with a half-life of 5.271 years and an average gamma-ray energy of 1.25 MeV, and cesium-137 with a half-life of 30.07 years and a gamma-ray energy of 0.662 MeV. *See* CESIUM; COBALT.

The application of industrial irradiation is increasing, with sterilization of medical disposables using cobalt-60 gamma rays being the most common. This includes the irradiation of plastic syringes, surgical gloves, and nonwoven drapes and gowns. In addition to the medical supplies, gamma sterilization of disposable laboratory supplies—petri dishes, pipettes, and test tubes—is also a fast-growing application.

A promising and appealing application is food preservation, including the reduction of postharvest losses by the elimination of pests with irradiation. Other food irradiation effects can be accomplished at distinct irradiation doses (**Table 2**). Regulatory bodies in several countries approve the consumption of over 20 irradiated foods. Extensive research on this subject has been done at the Bhabha Atomic Research Center in Trombay, India, where the need for bulk sterilization of food is more pressing.

TABLE 2. Food irradiation effects

Effect	Irradiation dose
Inhibition of sprouting	6–15 krad (60–150 gray)
Disinfection	20–1000 krad (200–1000 Gy)
Control of parasites	20–200 krad (200–2000 Gy)
Shelf-life extension	200–500 krad (2000–5000 Gy)
Extended refrigerated storage	200–500 krad (2000–5000 Gy)
Elimination of pathogens	300–700 krad (3000–7000 Gy)
Reduction of microbial load	Up to 1 Mrad (10 kGy)
Sterilization	Over 1 Mrad (10 kGy)

The biocidal effect of the gamma-irradiation process is effective for the control of microbiological contamination in many raw materials used by the pharmaceutical and cosmetic industries. Such materials include distilled water, starch powder, rice powder, lactose, talc, cocoa butter, inorganic pigments, petroleum jelly, and enzymes. Other applications include irradiation of male insects for pest control, sterilization of corks, sterilization of laboratory animal bedding, preparation of vaccines, degradation of poly(tetrafluoroethylene) scraps used in lubricants, cross-linking of polyethylene for shrink films, and production of wood-polymer and concrete-polymer composites.

Cesium-137 is also used for sludge processing to produce fertilizer or cattle feed. It is available in large quantities in the reprocessing of fuel elements from nuclear reactors, and the process helps solve the problem of disposing of cesium from this source. *See* NUCLEAR FUELS REPROCESSING.

Medical applications. Radioisotopes are used in the treatment of cancer by radiation. Encapsulated sources are used in two ways: the radioisotopes may be external to the body and the radiation allowed to impinge upon and pass through the patient (teletherapy), or the radiation sources may be placed within the body (brachytherapy).

For teletherapy purposes, cobalt-60 is the most commonly used isotope, with source strengths ranging from several thousand curies up to 10^4 Ci (3.7×10^{14} becquerels). Some cesium-137 irradiators have been built, but cesium-137 radiation is not as penetrating as that from cobalt-60.

TABLE 1. Most commonly used radioisotopes and major applications

Isotope	Half-life	Average γ-ray energy, MeV	Main uses
Radium-226 in equilibrium (^{226}Ra)	1600 years	0.83	Medical (brachytherapy)
Gold-198 (^{198}Au)	2.695 days	0.412	Medical (brachytherapy—permanent implants)
Iridium-192 (^{192}Ir)	73.8 days	0.38	Medical (brachytherapy)
Cesium-137 (^{137}Cs)	30.07 years	0.662	Industrial, medical (brachytherapy), research irradiators
Cobalt-60 (^{60}Co)	5.271 years	1.25	Industrial, medical (teletherapy), research irradiators
Iodine-125 (^{125}I)	59.4 days	0.035	Medical (brachytherapy—permanent implants)
Iodine-131 (^{131}I)	8.021 days	0.187, average β-ray energy	Medical (dispersal technique)

Historically, radium-226 has been used as the brachytherapy encapsulated source, but due to serious problems that can arise from a health physics aspect if the capsule breaks, radium is being replaced with cesium-137 sources. Iridium-172 is also used for brachytherapy applications. *See* HEALTH PHYSICS.

In order to avoid the long half-life of radium, its daughter product radon-222, which is an inert gas, was used in very small glass or gold seeds. Today radon has been replaced by iodine-125 or gold-198 seeds, which are put directly in a tumor and permanently left in place. *See* RADIUM.

For some medical applications the radioisotope is dispersed in the body; the most commonly used is iodine-131. The specific advantage of internal therapy with iodine-131 is that the thyroid gland concentrates the element iodine. When the radioisotope is administered either orally or intravenously in a highly purified form, it goes to the thyroid, where certain forms of thyroid disorders and cancers can be treated by the radiation. Unlike the other radioisotopes discussed here, the iodine-131's therapeutic effectiveness depends upon the beta rays emitted, not the gamma rays. *See* IODINE.

Another medical use for radioisotopes is the requirement that all transfusions of blood be irradiated before being given to certain patients so that most of the lymphocytes present are destroyed. For this purpose, self-contained irradiators are used. Cobalt-60 or cesium-137 encapsulated sources are permanently fixed and sealed off within the radiation shield, and the irradiation chamber is introduced into the radiation field by rotation. Such units are also extensively used for radiobiological and other research purposes. *See* RADIOACTIVITY AND RADIATION APPLICATIONS; RADIOLOGY.

Peter R. Almond

Bibliography. D. A. Bromley, Neutrons in science and technology, *Phys. Today*, 36(12):31–39, December 1983; Institute of Food Technologists, Food irradiation (a collection of articles), *Food Technol.*, 27(2):38–60, February 1983; T. Ouwerkerk, An overview of the most promising industrial applications of gamma processing, *International Symposium on Application and Technology of Ionizing Radiations*, Riyadh, Saudi Arabia, March 1982; M. Satin, *Food Irradiation: A Guidebook*, 2d ed., 1996.

Ixodides

A suborder of the Acarina, class Arachnida, comprising the ticks. Ticks differ from mites, their nearest relatives, in their larger size and in having a pair of breathing pores, or spiracles, behind the third or fourth pair of legs. They have a gnathosoma (or so-called head or capitulum), which consists of a base (basis capituli), a pair of palps, and a rigid, elongated, ventrally toothed hypostome which anchors the parasite to its host. They also have a pair of protrusible cutting organs, or chelicerae, which permit the insertion of the hypostome. The stages in the life cycle are egg, larva, nymph, and adult. The larvae have three pairs of legs; nymphs and adults, four. The 600 or so known species are all bloodsucking, external parasites of vertebrates including amphibians, reptiles, birds, and mammals.

Ticks are divided into three families, Argasidae, Ixodidae, and Nuttalliellidae. The last contains but one exceedingly rare African species, *Nuttalliella namaqua*, morphologically intermediate between the Argasidae and the Ixodidae. It is of no known importance, either medically or economically.

Argasidae. Argasids, or the soft ticks, differ greatly from ixodids in that the sexes are similar; the integument of adults and nymphs is leathery and wrinkled; there is no dorsal plate, or scutum; the gnathosoma is ventral in adults and nymphs but anterior in larvae; and the spiracles are small and anterior to the hindlegs (**Fig. 1**). These ticks frequent nests, dens, and resting places of their hosts. Adults feed intermittently and eggs are laid a few at a time in niches where the females seek shelter. Larvae feed for a few minutes to several days, then detach and transform to nymphs which feed and molt several times before transforming to adults. Nymphs and adults are notably resistant to starvation; some are known to live 10 years or longer without feeding.

The family contains about 85 species, with 20 in the genus *Argas* and 60 in *Ornithodoros*. Several are of medical or veterinary importance. *Argus persicus*, *A. miniatus*, *A. radiatus*, *A. sanchezi*, and *A. neghmei* are serious pests of poultry. Some of these species carry fowl spirochetosis, a disease with a high mortality rate. Larvae and nymphs of *Otobius megnini*, the spinose ear tick, feed deep in the ears of domesticated animals in many semiarid regions. Heavy infestations cause intense irritation which leads to unthriftiness and sometimes to death. The life history is unusual in that the adults do not feed. *Ornithodoros moubata* transmits

Fig. 1. Argasid tick, *Ornithodoros coriaceus*, shown enlarged.

Fig. 2. Ixodid tick, *Dermacentor andersoni*, female, shown enlarged.

relapsing fever in East, Central, and South Africa. It is a highly domestic parasite and humans are probably the chief host. *Ornithodoros turicata*, *O. hermsi*, *O. talaje*, and *O. rudis* are important vectors of relapsing fever in the Western Hemisphere, as is *O. tholozani* in Asia. The bites of most species that attack humans produce local and systemic reactions; the bites of some species, especially *O. coriaceus* of California and Mexico, are extremely venomous. *See* RELAPSING FEVER.

Ixodidae. In contrast to argasids, Ixodidae have a scutum covering most of the dorsal surface of the male but only the anterior portion of females, nymphs, and larvae (**Fig. 2**). They are known as the hard ticks. The sexes are thus markedly dissimilar. The gnathosoma extends anteriorly, and the large spiracles are posterior to the hindlegs. Instead of frequenting nesting places, these ticks are usually more or less randomly distributed throughout their hosts' environment. Larvae, nymphs, and adults feed but once, and several days are required for engorgement. The immature stages of most species drop to the ground for molting but those of the genus *Boophilus* and a few others molt on the host. The female lays a mass containing up to 10,000 or more eggs. The life cycle is usually completed in 1–2 years.

The family consists of about 11 well-defined genera with 500 species. Many species transmit disease agents to humans and animals; included are viruses, rickettsiae, bacteria, protozoa, and toxins. Transmission is by bite or by contact with crushed tick tissues or excrement. Virus diseases of humans include Colorado tick fever of the western United States, transmitted by *Dermacentor andersoni*, and Russian spring-summer encephalitis and related diseases in Europe and Asia, transmitted by *Ixodes persulcatus*. Some important rickettsial diseases of humans are spotted fever, widely distributed in the Western Hemisphere, of which *D. andersoni*, *D. variabilis*, *Rhipicephalus sanguineus*, and *Amblyomma cajennense* are important vectors; boutonneuse fever and related diseases of the Mediterranean region and Africa transmitted by *R. sanguineus* and some other species; and Q fever, the agent of which, although occurring in many ticks throughout the world, is not commonly transmitted directly by them. Tularemia (rabbit fever) is a bacterial disease transmitted by several species of ticks in the Northern Hemisphere. Tick paralysis, which is believed to be caused by toxins, occurs in many parts of the world and is produced by several species of ticks during the feeding process.

The numerous tick-borne diseases of animals cause vast economic losses, especially in tropical and subtropical regions. Examples are babesiasis, a protozoan disease caused by species of *Babesia* including *B. bigemina*, the agent of the widely distributed Texas cattle fever, which is transmitted principally by species of *Boophilus*; East Coast fever of Africa, another protozoan disease, caused by *Theileria parva*, is carried by several species of *Rhipicephalus*. Aside from carrying disease, several species are extremely important pests of humans and animals. Heavy infestations of ticks produce severe anemia in livestock, and even death, from the loss of blood alone. *See* TICK PARALYSIS; TULAREMIA.

Glen M. Kohls

J/psi particle

An elementary particle with an unusually long lifetime or, from the Heisenberg uncertainty principle, with an extremely narrow width $\Gamma = 86.6 \pm 6.0$ keV, and a large mass $m = 3096.93 \pm 0.09$ MeV. It is a bound state containing a charm quark and an anticharm quark. The discovery of the J/psi particle is one of the cornerstones of the standard model.

Discovery. Prior to 1974, there was no theoretical justification, and no predictions existed, for long-lived particles in the mass region 1–10 GeV/c^2.

The J/psi particles are rarely produced in proton-proton (*p-p*) collisions. Statistically, they occur once after hundreds of millions of subnuclear reactions, in which most of the particles are "ordinary" elementary particles, such as the kaon (K), pion (π), or proton (p). Searches for the J/psi particle are conducted by detecting its electron-positron (e^+e^-) decays. A two-particle spectrometer was used to discover this particle. A successful experiment must have: (1) a very high intensity incident proton beam to produce a sufficient amount of J particles for detection; and (2) the ability, in a billionth of a second, to pick out the J/psi $\rightarrow e^-e^+$ pairs amidst billions of other particles through the detection apparatus. *See* PARTICLE ACCELERATOR.

The detector is called a magnetic spectrometer. A positive particle and a negative particle each traversed one of two 70-ft-long (21-m) arms of the spectrometer. The e^+ and e^- were identified by the fact that a special counter, called a Cerenkov counter, measured their speed as being slightly greater than that of all other charged particles. Precisely measured magnetic fields bent them and measured their energy. Finally, as a redundant check, the particles plowed into high-intensity lead-glass and the e^+ and e^- immediately transformed their energy into light. When collected, this light "tagged" these particles as the e^+ and e^-, and not heavier particles such as the π, K, or p. The simultaneous arrival of an e^- and

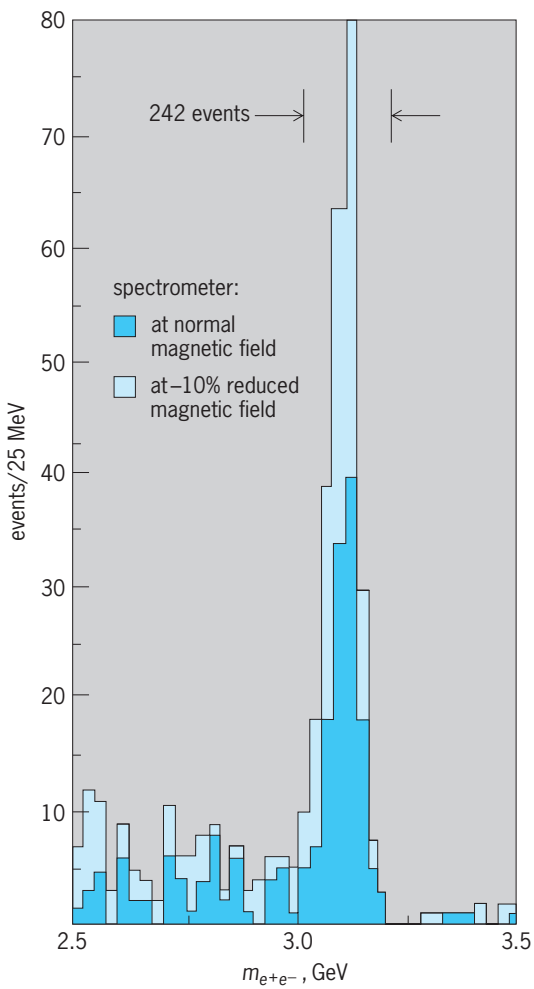

On-line data from August and October 1974, showing existence of J/psi particle. (*After J. J. Aubert et al., Discovery of the new particle, J. Nuc. Phys., B89(1):1–18, 1975*)

e^+ in the two arms indicated the creation of high-energy light quanta from nuclear interactions. The sudden increase in the number of e^+e^- pairs at a given energy (or mass) indicated the existence of a new particle (see **illus.**). *See* CERENKOV RADIATION; PARTICLE DETECTOR.

Properties. Since its discovery in 1974, more than 10^9 J/psi particles have been produced. More than 100 different decay modes and new particles radiating from the J/psi particle have been observed. The J/psi particle has been shown to be a bound state of charm quarks. The long lifetime of the J/psi results from its mass being less than the masses of particles which separately contain a charm and an anticharm quark. This situation permits the J/psi to decay only into noncharm quarks, and empirically this restriction was found to lead to a suppression of the decay rate resulting in a long lifetime and narrow width. The subsequent discovery of the b quark and the intermediate vector bosons Z^0 and $W^{\pm}$, and studies of Z^0 decays into charm, b, and other quarks, show that the theory of the standard model is in complete agreement with experimental data to an accuracy of better than 1%. *See* CHARM; ELEMENTARY PARTICLE; INTERMEDIATE VECTOR BOSON; QUARKS; STANDARD MODEL; UPSILON PARTICLES. Samuel C. C. Ting

Bibliography. O. Adriani et al., Results from the L3 experiment, *Phys. Rep.,* 236:1–146, 1993; J. J. Aubert et al., *Phys. Rev. Lett.*, 33:1404 and 1624, 1974; E. J. Augustin et al., *Phys. Rev. Lett.*, 33:1406, 1974; W. Braunschweig et al., *Phys. Lett.*, 53B:393, 1974; K. Hikasa et al., Review of particle properties, *Phys. Rev. D*, vol. 45, no. 11, pt. 2, 1992.

Jade

A name that may be applied correctly to two distinct minerals. The two true jades are jadeite and nephrite. In addition, a variety of other minerals are incorrectly called jade. Idocrase is called California jade, dyed calcite is called Mexican jade, and green grossularite garnet is called Transvaal or South African jade. The most commonly encountered jade substitute is the mineral serpentine. It is often called new jade or Korean jade. The most widely distributed and earliest known true type is nephrite, the less valuable of the two. Jadeite, the most precious of gemstones to the Chinese, is much rarer and more expensive.

Nephrite. Nephrite is one of the amphibole group of rock-forming minerals, and it occurs as a variety of a combination of the minerals tremolite and actinolite. Tremolite is a calcium-magnesium-aluminum silicate, whereas iron replaces the magnesium in actinolite. Although single crystals of the amphiboles are fragile because of two directions of easy cleavage, the minutely fibrous structure of nephrite makes it exceedingly durable. It occurs in a variety of colors, mostly of low intensity, including medium and dark green, yellow, white, black, and blue-gray. Nephrite has a hardness of 6 to $6^1/_2$ on Mohs scale, a specific gravity near 2.95, and refractive indices of 1.61 to 1.64. On the refractometer, nephrite gemstones show a single index near 1.61. Nephrite occurs throughout the world; important sources include Russia, New Zealand, Alaska, several provinces of China, and a number of states in the western United States. *See* AMPHIBOLE; GEM; TREMOLITE.

Jadeite. Jadeite is the more cherished of the two jade minerals, because of the more intense colors it displays. It is best known in the lovely intense green color resembling that of emerald (caused by a small amount of chromic oxide). In the quality known as imperial jade, the material is at least semitransparent. White, green and white, light reddish violet, bluish violet, brown, and orange colors are also found. Jadeite also has two directions of easy cleavage, but a comparable random fibrous structure creates an exceptional toughness. Although jadeite has been found in California and Guatemala, the only important source of gem-quality material ever discovered is the Mogaung region of upper Burma. The hardness of jadeite is $6^1/_2$ to 7, its specific gravity is approximately 3.34, and its refractive indices are 1.66 to 1.68; on a refractometer, only the value 1.66 is usually seen. *See* JADEITE. Richard T. Liddicoat, Jr.

Jadeite

The monoclinic sodium aluminum pyroxene, $NaAlSi_2O_6$. Free crystals are rare. Jadeite usually occurs as dense, felted masses of elongated blades or as fine-grained granular aggregates. It has a Mohs hardness of 6.5 and a density of 3.25–3.35. It has a vitreous or waxy luster, and is commonly green but may also be white, violet, or brown. Jadeite exhibits the 93° and 87° pyroxene cleavages and a splintery fracture. It is extremely tough.

Jadeite is always found in metamorphic rocks. It is associated with serpentine at Tawmaw, Burma; Kotaki, Japan; and San Benito County, California. It occurs in metasedimentary rocks of the Franciscan group in California and in Celebes. It is also found in Tibet; Yunan Province, China; and Guatemala. In addition, the jadeite "molecule" is present in high concentrations in solid solution in omphacite, a dark green pyroxene which, like garnet, is a major constituent of eclogite, a high-pressure, high-temperature metamorphic rock formed in the deep crust or upper mantle. Eclogites are found in the Bessi district, Japan; Mampong, Ghana; Gertrusk, Austria; Sonoma County, California; and as nodules in South African kimberlite pipes. *See* ECLOGITE; SOLID SOLUTION.

Jadeite's composition is intermediate between that of nepheline and albite, but jadeite does not crystallize in the binary system $NaAlSiO_4$–$NaAlSi_3O_8$, or in the ternary system Na_2O–Al_2O_3–SiO_2, at low pressures. Its high density compared with the above tektosilicates and its presence in eclogite suggest that it is a high-pressure phase. This was confirmed by high pressure experiments on stability relationships between jadeite, nepheline, and albite (see **illus.**). Pure jadeite usually occurs with minerals indicative

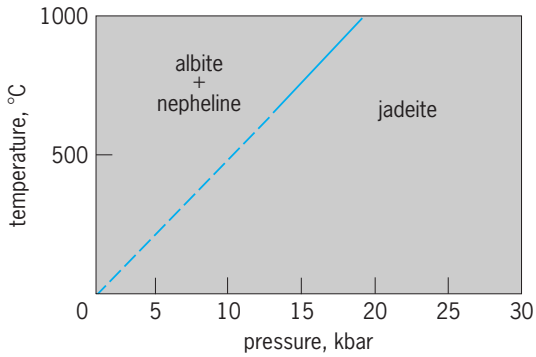

Equilibrium curve for the reaction albite + nepheline → 2 jadeite. 1 kilobar = 10 pascals. °F = (°C × 1.8) + 32.

of relatively low temperatures (390–570°F or 200–300°C) such as analcite, lawsonite, chlorite, glaucophane, stilpnomelane, and zeolites.

Jadeite is valued as a precious stone for carvings. *See* HIGH-PRESSURE MINERAL SYNTHESIS; JADE; PYROXENE.

Lawrence Grossman

Jahn-Teller effect

A distortion of a highly symmetrical molecule, which reduces its symmetry and lowers its energy. The effect occurs for all nonlinear molecules in degenerate electronic states, the degeneracy of the state being removed by the effect. It was first predicted in 1937 by H. A. Jahn and E. Teller. In early experimental work, the effect often "disappeared" or was masked by other molecular interactions. This has surrounded the Jahn-Teller effect with a certain mystery and allure, rarely found in science today. However, there are now a number of clear-cut experimental examples which correlate well with theoretical predictions. These examples range from the excited states of the most simple polyatomic molecule, H_3, through moderate-sized organic molecules, like the ions of substituted benzene, to complex solid state phenomena involving crystals or localized impurity centers. *See* DEGENERACY (QUANTUM MECHANICS); MOLECULAR STRUCTURE AND SPECTRA; QUANTUM MECHANICS.

Square planar molecule example. The fact that electronic degeneracy may destroy the symmetry upon which it is based is most easily demonstrated for molecules with fourfold symmetry. Consider a square planar molecule (with fourfold symmetry, D_{4h}) with an unpaired electron in an orbitally degenerate state (for example, E_u). This state will be characterized by two orthogonal wavefunctions ψ_A and ψ_B and electron density distributions with mutually perpendicular nodal planes, as shown schematically in **Fig. 1***a* and *b*. As long as the nuclei are undisplaced, the molecules in Fig. 1*a* and 1*b* are geometrically congruent and the energy of the system will be the same, regardless of whether the electron is in orbital ψ_A or ψ_B, that is, $E_A(0) = E_B(0)$. Imagine the molecule undergoing an in-plane bending vibration ($B_{2g}\nu_4$) and consider the nuclei displaced an increment $+q$ along the appropriate normal coordinate Q. It is easily seen that the positions of the nuclei with respect to the electronic distributions characterized by ψ_A and ψ_B will not be equivalent, and the energies of the states A and B will differ, as shown on the right side of the potential energy diagram, Fig. 1*c*. The same argument will, of course, hold for the opposite phase of the vibration and displacement $-q$ as shown on the left side of Fig. 1*c*. It is easily seen that the molecule in Fig. 1*a* with displacement $+q$ is geometrically congruent with the molecule in Fig. 1*b* for displacement $-q$. It can therefore be deduced that for the electronic energies $E_A(q) = E_B(-q)$ and similarly $E_A(-q) = E_B(q)$. The establishment of these five points determines the general form of the Jahn-Teller potential curve in Fig. 1*c*. The potential surfaces of the states A and B must cross. The crossing point at zero displacement, with the corresponding energy $E(0)$, is clearly not an absolute minimum; therefore, the totally symmetric square configuration will not be a stable equilibrium for the degenerate electronic state. Therefore, at equilibrium, the molecule will be distorted from a square, and its energy will be lowered. *See* MOLECULAR ORBITAL THEORY.

The above arguments are not restricted to square molecules. With the exception of linear molecules which suffer Renner-Teller effects, all polyatomic molecules of sufficiently high symmetry to possess

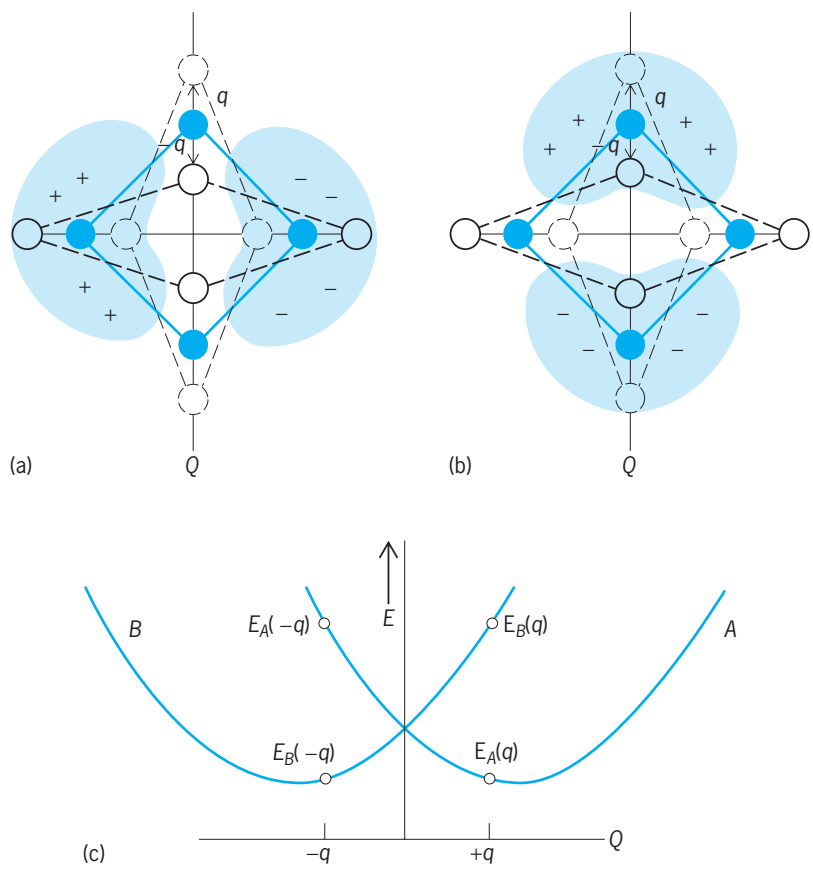

Fig. 1. Jahn-Teller effect for a square planar molecule. Displacement of the nuclei of the molecule in the electronic field provided by (*a*) the wavefunction Ψ_A and (*b*) Ψ_B. (*c*) Energy *E* corresponding to Ψ_A and Ψ_B as a function of the displacement coordinate *Q*.

orbitally degenerate electronic states will be subject to the Jahn-Teller instability. However, in cases other than the fourfold symmetry discussed above, the proof is somewhat more involved and requires the use of the principles of group theory. *See* GROUP THEORY; RENNER-TELLER EFFECT.

Vibrational structure. The distortion of the electronic potential surface of a degenerate electronic state by the Jahn-Teller effect has been considered. The nuclear motion on such a distorted potential surface leads to perturbations in the vibrational structure. These perturbations result in shifts and splittings of the vibrational levels. While group theory, as noted above, can demonstrate the susceptibility of a given molecule to Jahn-Teller instability, it gives no information about the magnitude of the effect; such information must come from experiments. The hexagonal benzene molecule, C_6H_6, and its symmetrically substituted derivatives, represent almost ideal examples of highly symmetrical molecules. Their least tightly bound electrons reside in a pair of degenerate π orbitals. The removal of one of these electrons produces a benzene positive ion whose ground state is electronically degenerate and so fulfills the criteria for a Jahn-Teller effect. These ions have been studied extensively experimentally and represent excellent examples with which to illustrate vibrational effects.

For molecules with the sixfold symmetry (D_{6h}) of a regular hexagon, group theoretical arguments demonstrate that Jahn-Teller distortions of the nuclei occur for motions along four doubly degenerate e_{2g} vibrational normal modes denoted ν_{15}–ν_{18}. These are shown schematically in **Fig. 2**, where the arrows represent the cartesian components of the atomic displacements. The three-dimensional form of the potential energy E with respect to any one of these doubly degenerate coordinates Q_i is shown in

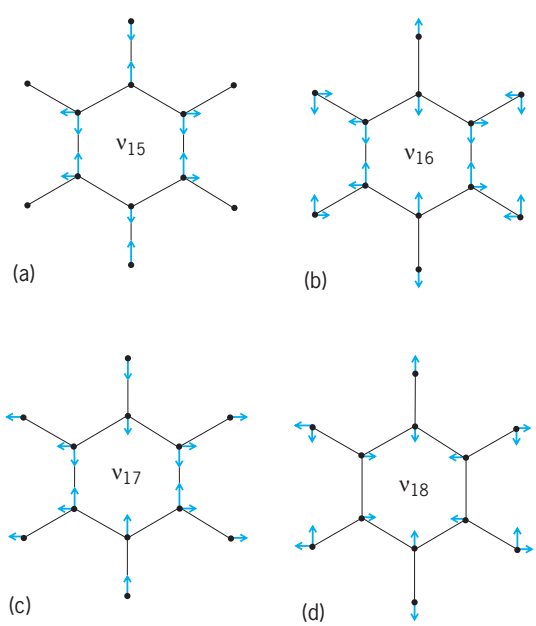

Fig. 2. The four e_{2g} vibrational modes which are Jahn-Teller-active for the positive ion of any substituted benzene with D_{6h} symmetry: (*a*) C-C stretch, (*b*) C-F stretch, (*c*) C-C-C bend, and (*d*) C-F bend.

Fig. 3*a*; Fig. 3*b* shows cuts through this type of surface assuming various magnitudes for the Jahn-Teller distortion and illustrates how the introduction of the Jahn-Teller effect changes the potential surface. On the far left with no Jahn-Teller effect, the doubly degenerate electronic potential has a minimum at the symmetrical configuration (a regular hexagon). As the Jahn-Teller effect increases, the electronic potentials split and lowered minima (Jahn-Teller stabilization) in the potential occur at distorted nuclear configurations.

The Jahn-Teller effect is usually characterized in terms of a linear parameter D defined by $D = E_{JT}/\hbar\omega$, where E_{JT} is the Jahn-Teller stabilization energy (the separation between the central maximum and the distorted minima) and $\hbar\omega$ is the energy of the vibrational mode leading to the instability. As long as only the linear Jahn-Teller effect (wherein only terms linear in the displacement coordinate appear in the hamiltonian) is considered, the distorted potential of the D_{6h} species in Fig. 3*a* exhibits an equipotential moat around the central maximum. The quadratic Jahn-Teller effect (caused by terms quadratic in the displacement coordinate) produces the threefold local minima and saddle points illustrated in Fig. 3*a*, and obviously requires the introduction of an additional coupling parameter, analogous to D, for its characterization. For small distortion when D is much less than 1 and quadratic terms are insignificant, approximate formulas describing the vibronic levels have been derived. For the general case when the distortion is large (D approximately 1 or greater) or when terms quadratic in Q are important, no explicit formulas can usually be given. More sophisticated techniques are needed to calculate the vibrational structure, and usually involve the diagonalization of large matrices.

Such analysis shows that for the substituted benzenoid cations mentioned above the linear Jahn-Teller parameter D is not greater than about 1.0, while quadratic and higher terms are almost negligible. The energy reduction, that is, the Jahn-Teller stabilization energy E_{JT}, is about 2–3 kcal/mole (8–12 kJ/mole) for these ions. This energy reduction is obtained with the benzene ring distorted from a regular hexagon by 2–3° in its internal angles, and with the C-C bond lengths differing by about 4 picometers.

Static and dynamic effects. The terms static and dynamic Jahn-Teller effect are frequently used, but unfortunately their use is not standardized. The term static Jahn-Teller effect is often used to describe the distortion of the electronic potential as seen in Figs. 1*c* and 3, with corresponding use of dynamic Jahn-Teller effect to describe vibrational motion on this surface. However, other uses of the terms static and dynamic involve the magnitude of the Jahn-Teller effect and the time scale of the experiments used to measure its properties. As long as the distortion is small (D less than 1), all vibrational levels ν will be located above the central maximum (Fig. 3*b*). On the time scale of most experiments, a time-averaged, symmetrical configuration will be sampled, and the

term dynamic Jahn-Teller effect is therefore often applied. For a sufficiently large distortion (D greater than 1), one or more vibrational levels may be localized in separate, disjoint minima, with interconversions between them being slow. This situation (D greater than 1) correspondingly would be referred to as the static Jahn-Teller effect. The distortion will therefore manifest itself as a static or dynamic effect depending on the time scale of the experiment. Further complication arises when the Jahn-Teller active mode is itself degenerate, as in the case of the e_{2g} vibrations of Fig. 2. In this case the switching between the potentials U' and U' of Fig. 3 can occur by pseudorotation within the moat, without ever surmounting the symmetric maximum. Then large quadratic terms, resulting in deep minima around the moat, are required to produce a really static effect.

Solids. While the Jahn-Teller effect was originally postulated for highly symmetric discrete molecules, it can also appear in crystals and other highly symmetric extended systems. Searches for the observable consequences of the Jahn-Teller effect in molecular spectra were initially hampered by the fact that typically only excited electronic states of stable molecules or states of transient species—molecular ions and free radicals—possess the orbital degeneracy required for its observation. Most of the early experimental observations therefore involved optical or microwave spectra of point defects and impurity centers in solids. *See* CRYSTAL DEFECTS.

In treating these systems, usually only the impurity or defect and its nearest-neighbor atoms are considered, and the rest of the solid is regarded as a featureless heat bath. As an example, **Fig.** 4*a* shows schematically a simple cluster, consisting of an electron in a p state trapped at a negative ion vacancy in an alkali halide (a so-called F-center), surrounded by a regular octahedron of its nearest neighbor positive ions. In the undistorted lattice, the three possible states, p_x, p_y, and p_z, will be obviously degenerate. Now suppose the cluster undergoes an asymmetric vibration and the ions start moving in the directions indicated by the arrow. It is easily seen that the energy of the p_z state will be initially lowered, since the positive ions will approach more closely the negative lobes of the p_z wave function, while, by a similar argument, those of the p_y and p_x states will be raised (Fig. 4*b*). The surrounding crystal will counterbalance the distortion, with the corresponding energy increase being, by Hooke's law, initially quadratic in displacement. As in the molecular case, it is easy to see that an energy minimum for the F-center electron will not be in the symmetric configuration but at some finite displacement. *See* COLOR CENTERS.

Besides the optical and microwave spectra, the consequences of the distorted geometry are also observable in electron spin resonance and Zeeman spectra. When, for instance, a magnetic field H is applied parallel to the z axis of Fig. 4*a*, the degeneracy of the p state is split into three states with energies of approximately $-\beta H$, corresponding to the three possible angular momentum projection eigenvalues of $L_x = +1, 0, -1$, with β the Bohr magneton. In the Jahn-Teller distorted case, the electron can no longer rotate freely, it has to "drag" its distortion with it. This increases its effective mass and quenches its orbital contribution to the magnetic splitting. Such reductions have been observed in the Zeeman spectra of numerous transition-metal ions in crystals. The magnitude of this reduction is dependent on the Jahn-Teller stabilization energy. At absolute zero (0 K) it was shown to be approximately $\exp(-3E_{JT}/2\hbar\omega)$. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; ZEEMAN EFFECT.

In some ions the electron-lattice interaction is so weak that even though their ground states are orbitally degenerate, Jahn-Teller effects are ordinarily not detected when they result from isolated impurities. In crystals containing large concentrations

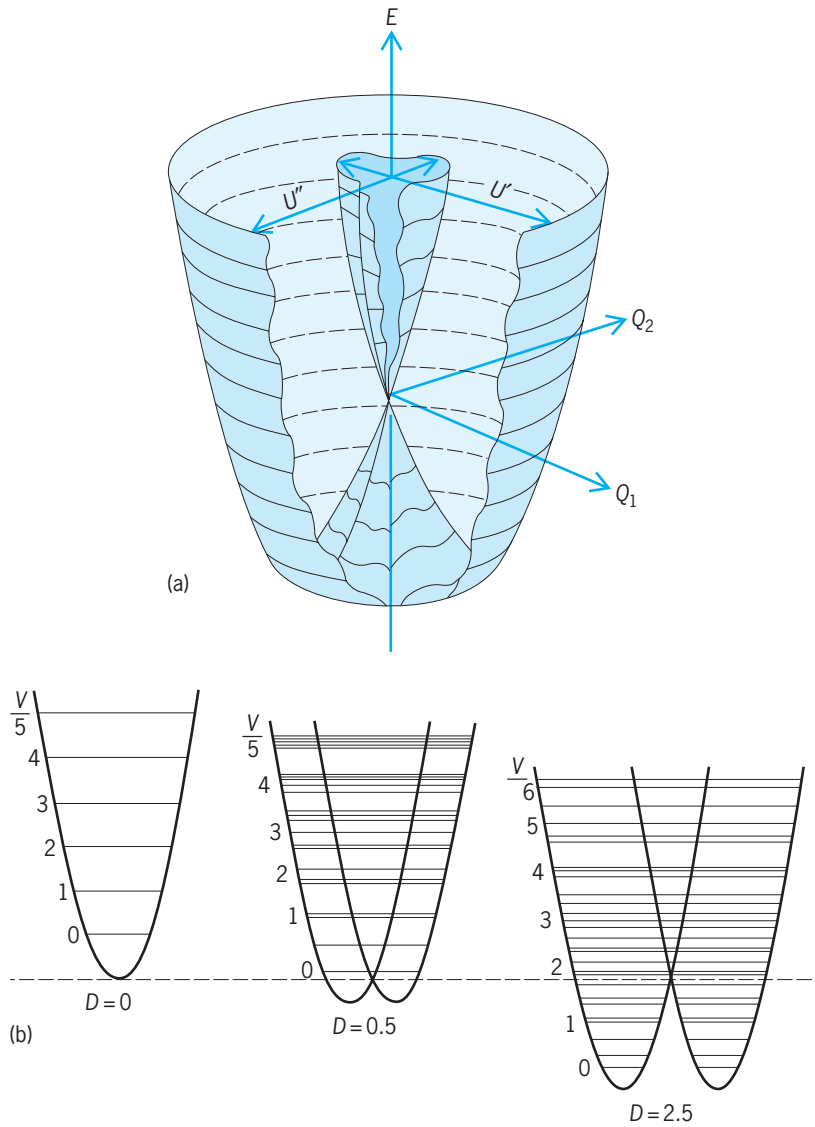

Fig. 3. Potential energy for a doubly degenerate, Jahn-Teller active vibration of a molecule with D_{6h} or D_{3h} symmetry. (*a*) Potential energy surface. (*b*) Cuts through the surface for increasing magnitude of the Jahn-Teller effect going from left to right. Horizontal lines represent vibrational levels, which are numbered with the index *V* for the undistorted molecule.

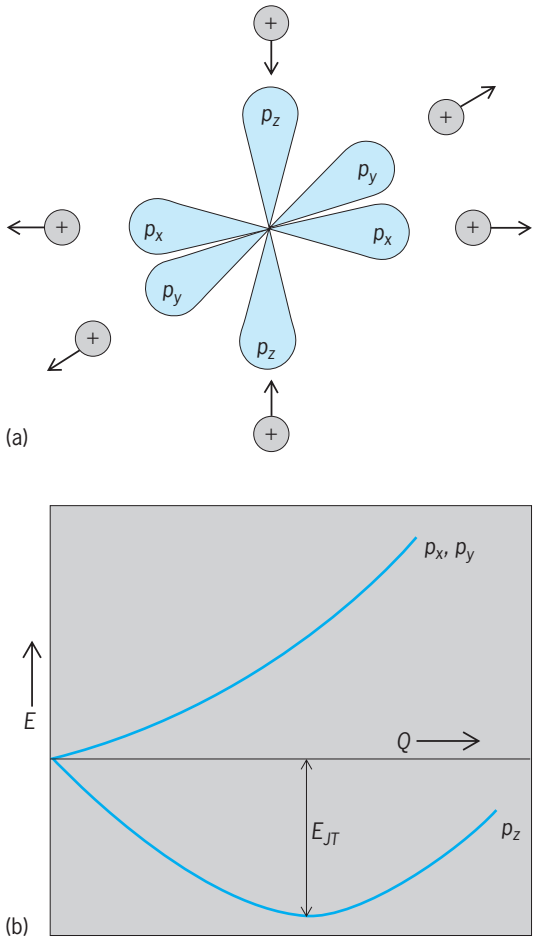

Fig. 4. Jahn-Teller effect in an F-center. (*a*) The *p* orbitals of the F-center surrounded by an octahedron of positive ions. (*b*) The potential energy *E* of the *p* orbitals as the positive ions are displaced, as a function of displacement coordinate *Q*.

of these ions, the ion-ion interactions can greatly enhance these effects. Many rare-earth compounds have been found to undergo second-order phase transitions in which the overall crystal symmetry is reduced and the electronic degeneracy removed in a manner expected for the Jahn-Teller effect. Such distortions, involving cooperation of many ions, are usually called the cooperative Jahn-Teller effect. These minute distortions may cause shifts in phonon modes, which are observable in Raman spectra or neutron diffraction experiments. *See* NEUTRON DIFFRACTION; RAMAN EFFECT.

V. E. Bondybey; Terry A. Miller

Bibliography. I. B. Bersuker, *The Jahn-Teller Effect*, 2006; C. C. Chancey and M. C. M. O'Brien, *The Jahn-Teller Effect in C_{60} and Other Icosahedral Complexes*, 1997; R. Englman, *The Jahn-Teller Effect in Molecules and Crystals*, 1972; T. A. Miller and V. E. Bondybey (eds.), *Molecular Ions: Spectroscopy, Structure and Chemistry*, 1983; Yu. Perlin and M. Wagner, *Modern Problems in Condensed Matter Sciences: The Dynamical Jahn-Teller Effect in Localized Systems*, 1985.

Jamming

Intentional generation of interfering signals in the electromagnetic and infrared spectrum by powerful transmitters as a countermeasure intended to block a communication, radar, or infrared system or to impair its effectiveness appreciably. Radio broadcasts or radio messages can be jammed by beaming a more powerful signal on the same frequency at the area in which reception is to be impaired, using carefully selected noise modulation to give maximum impairment of intelligibility of reception. When stations on many different frequencies are to be jammed or when an enemy is changing frequencies to avoid jamming, the jamming transmitter is correspondingly changed in frequency or swept through a range of frequencies over and over again. Techniques similar to this are also used at radar frequencies to jam early-warning and gunfire-control radar systems. *See* ELECTRONIC WARFARE.

John Markus; Paul J. DeLia

Jasper

An opaque, impure type of massive fine-grained quartz that typically has a tile-red, dark-brownish-red, brown, or brownish-yellow color. The color of the reddish varieties of jasper is caused by admixed, finely divided hematite, and that of the brownish types by finely divided goethite. Jasper has been used since ancient times as an ornamental stone, chiefly of inlay work, and as a semiprecious gem material. Under the microscope, jasper generally has a fine, granular structure, but fairly large amounts of fibrous or spherulitic silica also may be present.

Jasper has a smooth conchoidal fracture with a dull luster. The specific gravity and hardness are variable, depending upon particle size and the nature and amount of the impurities present; both values approach those of quartz. The color of jasper often is variegated in banded, spotted, or orbicular types. Heliotrope is a translucent greenish chalcedony containing spots or streaks of opaque red jasper, and jaspagate contains bands of chalcedonic silica alternating with jasper. Jaspilite is a metamorphic rock composed of alternating layers of jasper with black or red bands of hematite. *See* CHALCEDONY; GEM; QUARTZ.

Clifford Frondel

Jaundice

The yellow staining of the skin and mucous membranes associated with the accumulation of bile pigments in the blood plasma. Bile pigments are the normal result of the metabolism of blood pigments, and are normally excreted from the blood into the bile by the liver. An increase in circulating bile pigments can, therefore, come about through increased breakdown of blood (hemolytic jaundice), through lack of patency of the bile ducts (obstructive jaundice), through inability or failure of the liver to clear

the plasma (parenchymal jaundice), or through combinations of these. *See* GALLBLADDER; LIVER.

Metabolic pathway. Metabolism of the hemoglobin from destroyed red blood cells is carried on in organs of the reticuloendothelial system, such as the spleen, and the resulting bilirubin is liberated to the plasma. The plasma then circulates through the liver, where the bilirubin is conjugated enzymatically with glucuronic acid, and is excreted in the bile. Bile travels through the bile ducts to the small intestine, whence a small amount of altered bilirubin, termed urobilinogen, may be reabsorbed into the plasma. Excessive destruction of red blood cells causes accelerated production of bilirubin, overloads the ability of the liver to remove the pigment from the circulation, and produces jaundice. Blockage of the bile ducts causes elevation of plasma bilirubin glucuronide level because of the inability to dispose of this material in the usual channel. Damage to liver cells may cause elevation of the plasma bilirubin or bilirubin glucuronide or both, depending on the type and severity of liver cell damage.

Although the major portion of circulating bilirubin is derived from the breakdown of red cells, some is also contributed through inefficient or incomplete utilization of precursors of hemoglobin, which spill into the plasma bilirubin pool without having been used for hemoglobin synthesis. In addition, metabolism of cytochromes and similar pigmented compounds found in all cells of the body yields small quantities of bilirubin. *See* HEMOGLOBIN.

Jaundice occurs when the level of these circulating pigments becomes so high that they are visible in the skin and mucous membranes, where they are bound by a reaction which has not been identified. In the normal adult, levels of total bilirubin, that is, the total bilirubin and bilirubin glucuronide, rarely exceed 0.8–1.0 mg/100 ml of plasma, while jaundice usually becomes visible when total bilirubin approaches 1.5 mg. *See* BILIRUBIN.

Hemolytic jaundice. Destruction of red blood cells in the normal human adult proceeds at a rate at which about 0.8% of the circulating hemoglobin is broken down each day. This can be increased in states of excessive hemolysis up to 10- to 15-fold without overtaxing the remarkable ability of the liver to clear bilirubin from the plasma. Even this rate of clearing can be exceeded, however, in certain morbid states which result in hemolytic jaundice; such states include various hemolytic anemias, hemolysis resulting from incompatible blood transfusion, severe thermal or electric injuries, or introduction of hemolytic agents into the bloodstream. Similar jaundice occurs in pulmonary infarction.

In infants, and especially in premature infants, the ability of the liver to conjugate bilirubin with glucuronide is much less than in adults, apparently because of the lack of suitable enzymes. Jaundice appears in many infants shortly after birth, then disappears within a few days, with development of the appropriate enzyme structure. In the uncommon constitutional hepatic dysfunction, this enzyme defect apparently persists into adult life. The infantile jaundice accompanying erythroblastosis fetalis (Rh babies) is due to the inability of the infantile liver to metabolize the bilirubin resulting from markedly accelerated hemolysis. *See* BLOOD GROUPS; RH INCOMPATIBILITY.

A related form of jaundice occurs when an abnormality exists in the hemoglobin-forming cells of the bone marrow in which inadequate utilization of hemoglobin precursors occurs. Moderate degrees of jaundice can result from the accumulation of these substances in the bloodstream.

Obstructive jaundice. The highest levels of total bilirubin are seen in chronic obstructive jaundice, in which plasma levels may reach 50–60 mg/100 ml, and the skin may take on a remarkable deep-yellow hue. This condition may be brought about through a variety of means. In the infant there may be a severe maldevelopment of the bile ducts such that no channel for the flow of bile exists, while in the adult obstructive jaundice is most commonly caused by impaction of a gallstone in the ducts. Benign and malignant tumors of the gallbladder, bile ducts, pancreas, lymph nodes, and other organs may also cause compression of the bile ducts with loss of patency, and bile duct stricture may follow surgery or inflammation in the region. A similar, less severe, reversible picture is seen as a hypersensitivity response to the administration of some drugs, the most common of which are chlorpromazine and related drugs, and methyl testosterone. In the uncommon benign disorder known as idiopathic familial jaundice, there appears to be decreased ability to excrete conjugated bilirubin into the bile ducts, giving rise to a constant, but usually quite slight, elevation in the plasma bilirubin glucuronide. *See* GALLBLADDER DISORDERS.

Parenchymal jaundice. A wide variety of diseases exists in which part of the jaundice can be accounted for by actual damage to liver cells, with consequent decrease in their ability to conjugate bilirubin and excrete the glucuronide, causing an elevation of both fractions in the plasma. This group comprises such conditions as inflammations of the liver, including viral hepatitis, Weil's disease, syphilis, parasitic infestations, and bacterial infections; toxic conditions, including poisoning from a wide variety of organic and inorganic compounds and in a broader sense the toxemias associated with severe systemic diseases; tumorous conditions, including primary hepatic tumors and those metastatic from other organs; and other conditions, the most common of which is congestive heart failure. Some of these conditions have an added component of obstructive or hemolytic jaundice, which confuses the picture for the clinician.

Symptomatology. The appearance and symptomatology of subjects suffering from jaundice vary from case to case and depend on the underlying disease process producing the jaundice. Elevation of bilirubin by itself has very limited deleterious effects on overall physiology, with two exceptions: The brain of very young infants is subject to damage by high levels of circulating bilirubin (a condition termed

kernicterus); and, under certain conditions in the adult, high levels of circulating bilirubin appear to contribute to kidney damage.

Aside from the effects of the jaundice itself, however, individuals with hemolytic jaundice usually have an accompanying anemia. Those with obstructive jaundice commonly note that their stools are not brown, a symptom caused by lack of bile pigments (acholic stools), while bilirubin glucuronide appears in the urine and causes it to turn dark. Persons with malignancies demonstrate the usual signs of weight loss and anemia, while those with inflammatory conditions commonly have fever, chills, and prostration. *See* ANEMIA; PIGMENTATION. Rolla B. Hill

Jawless vertebrates

The common name for members of the agnathans. Jawless vertebrates include the cyclostomes (modern lampreys and hagfishes) as well as extinct armored fishes, known colloquially as ostracoderms ("bony-skinned"), that lived in the Ordovician, Silurian, and Devonian periods (**Fig. 1**). Agnathans have pouchlike gills opening through small pores, rather than slits as in jawed vertebrates (Gnathostomata). Primitively, agnathans lack jaws, they show a persistent notochord, and most have no paired fins. *See* CYCLOSTOMATA (CHORDATA).

Types. The lampreys (Petromyzontiformes; 38 extant species) and hagfishes (Myxiniformes; ~70 species) are scaleless, eel-shaped fishes with round mouths, inside which are keratinized teeth carried upon a complex tongue and oral hood. Parasitic lampreys attach to host fishes by a sucker and use the tongue to rasp away flesh and blood. Hagfishes are exclusively marine, are blind, and live most of their lives buried in mud, emerging to eat polychaete worms, mammal carcasses, or dead or dying fishes. Lampreys and hagfishes show a bipolar distribution (they occur in both Northern and Southern hemispheres, but only 4 species occur in the Southern) and prefer cool waters. *See* MYXINIFORMES; PETROMYZONTIFORMES.

Ostracoderms were very variable in shape and were covered with a superficial bony armor made up of solid shields or scales. One of the best known was the osteostracans (cephalaspids) of Eurasia and North America, with a solid semicircular head shield pierced by dorsally placed eyes and a small circular mouth on the undersurface. The head shield is also marked by sensory fields (lacunae filled with small plates) which were specialized parts of the lateral-line system. Osteostracans probably lived on, or were partly buried in, mud and sand, where they sucked in small food particles. *See* CEPHALASPIDOMORPHA; OSTEOSTRACI; OSTRACODERM.

Galeaspids were superficially like osteostracans and probably occupied a similar ecological niche in China. The head shield in many was extended into an elaborate snout.

Heterostracans (pteraspids) were torpedo-shaped active swimmers that lived alongside osteostracans. The head was encased in large dorsal and ventral shields which surrounded small, laterally directed eyes. The mouth, formed by a lower lip equipped with small bony plates, may have been extended as a scoop to ingest food particles. A few ostracoderms superficially look like heterostracans and sometimes are classified with them because they show a similar type of bone which lacks enclosed osteocytes (bone cells). Heterostracans are the oldest of completely preserved vertebrates and lived in seas 460 million years ago. *See* HETEROSTRACI.

The Thelodonti are a group of small ostracoderms in which the body is completely covered with

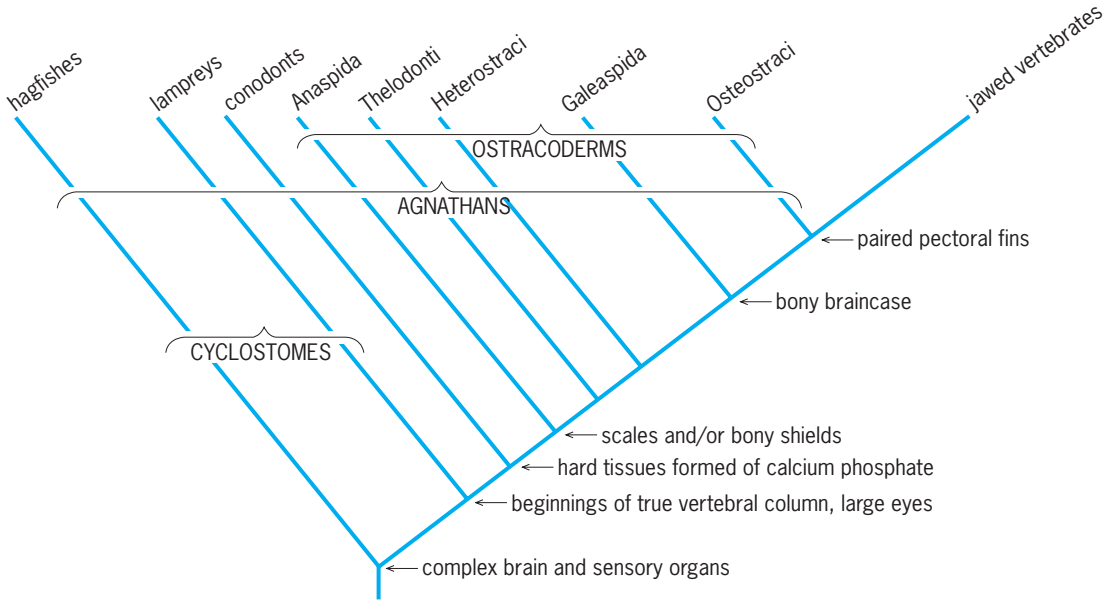

Interrelationships of jawless and jawed vertebrates, showing distribution of some major evolutionary steps.

minute, finely sculptured scales. These scales are found worldwide and in great abundance in Upper Silurian and Lower Devonian rocks. Consequently, they have been used for stratigraphic zonation and correlation. Some scientists think that thelodonts may be closely related to sharks and rays (Chondrichthyes). *See* CHONDRICHTHYES.

Interrelationships. Modern ideas of the interrelationships of agnathan fishes suggest that they are not a natural group; that is, some such as the osteostracans and galeaspids are more closely related to jawed vertebrates than to other agnathans. Similarly, lampreys share many specializations with jawed vertebrates not seen in hagfishes, such as neural arches along the notochord, eye muscles, nervous control of heartbeat, and the capability to osmoregulate and adapt to freshwater. These attributes suggest that lampreys are more closely related to jawed vertebrates than to hagfishes and that the Cyclostomata are not a natural group.

Recently, Conodonta have been classified as vertebrates. Conodonts are preserved as tiny, phosphatic, toothlike elements found in marine deposits from the Cambrian to Triassic. Usually found in great abundance, they are useful for dating and correlating rocks. Rare fossils of soft-bodied, eel-shaped animals show that conodonts are part of complex feeding devices within the mouth. They grew to at least 5 cm (2 in.) (some incompletely known forms must have been considerably longer) and had large eyes, suggesting an active predatory lifestyle. *See* CONODONT.

A classification of jawless vertebrates is given below:

Subphylum Craniata (complex brain and sensory organs)
- Class Hyperotreti
  - Order Myxiniformes
- Class Vertebrata (with neural arches in backbone)
  - Order: Petromyzontiformes
    - Conodonta
    - Anaspida
    - Thelodonti
    - Heterostraci (pteraspids)
    - Galeaspida
    - Osteostraci (cephalaspids)

Peter L. Forey

Bibliography. P. C. J. Donoghue et al., Conodont affinity and chordate phylogeny, *Biol. Rev.*, 75:191–251, 2000; P. L. Forey and P. Janvier, Evolution of the early vertebrates, *Amer. Sci.*, pp. 554–565, November–December 1994; P. Janvier, *Early Vertebrates*, Clarendon Press, Oxford, 1996.

Jerboa

The common name for 33 species of rodents in the family Dipodidae (see **table**). Jerboas inhabit the hot arid desert and semiarid regions of Turkey, North Africa, and Asia. *See* DESERT; RODENTIA.

Description. Jerboas are rodents ranging 2–8 in. (50–200 mm) in head and body length (see **illustration**). The rounded tail is usually about 2 in. (50 mm) longer than the head and body. The ears are large, equaling half the length of the head and body in some species. The eyes are large. Coloration varies but is usually some shade of buff (moderate orange-yellow) mixed with black or russet; the underparts are white.

TABLE 1. Genera and common names of jerboas

Genus	Common name
Cardiocranius	Five-toed dwarf jerboa (1 species)
Salpingotus	Three-toed dwarf jerboas (5 species)
Salpingotulus	Baluchistan pygmy jerboa (1 species)
Paradipus	Comb-toed jerboa (1 species)
Euchoreutes	Long-eared jerboa (1 species)
Dipus	Rough-legged jerboa, or northern three-toed jerboa (1 species)
Eremodipus	Central Asian jerboa (1 species)
Jaculus	Desert jerboas (4 species)
Stylodipus	Thick-tailed three-toed jerboas (3 species)
Allactaga	Four- and five-toed jerboas (11 species)
Allactodipus	Bobrinski's jerboa (1 species)
Alactagulus	Lesser five-toed jerboa, or little earth hare (1 species)
Pygeretmus	Fat-tailed jerboas (2 species)

Of all the rodents, this kangaroolike mammal is perhaps the most highly developed for getting around on two feet. The hindlimbs are large and very long (at least four times the length of the front limbs). In most species the three main foot bones are fused into a single "cannon bone" for greater strength. The front limbs are considerably smaller and are used primarily for grooming and for holding the grass seeds on which many species feed. Jerboas prefer to stand up on their hindlegs and hop, covering 2–6 ft (0.6–1.8 m) in a single jump. When alarmed, a jerboa can travel by leaps and bounds of 5–10 ft (1.5–3 m), considerably faster than a human can run. *Allactaga elater* has been timed at a speed of 48 km/h (30 mi/h). The tail, which ends in a tuft of hair, not only gives the jerboa support when it is standing but also helps the animal maintain its balance as it leaps along. The thick hairs on the feet absorb some of the landing shock and provide traction on the sand.

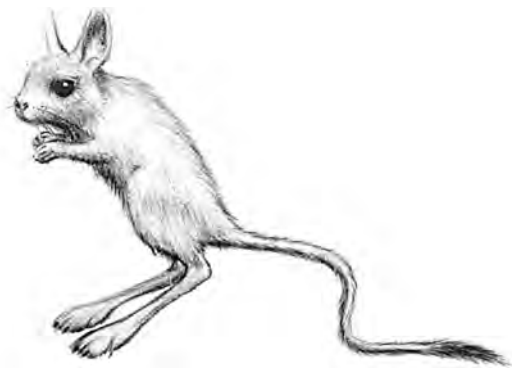

Typical jerboa.

The largest of all jerboas is the four-toed jerboa (*Allactaga major*). Head and body length may range 90–263 mm (3.5–10 in.), tail length is 142–308 mm (5.5–12 in.), and weight is 280–420 grams (9.75–14.75 oz).

Behavior. Jerboas are nocturnal in order to escape the intense heat of the day. They spend the day holed up in their underground den. When they retire for the day, they close up the burrow entrance to keep out the hot sunlight. Four types of burrows are utilized by different species of jerboas: temporary summer day burrows, temporary summer night burrows, permanent summer burrows used as living quarters and for producing young, and permanent winter burrows for hibernation. Temporary shelters during the day may consist of a long but simple tunnel up to 13 m (42 ft) in length and 20–30 cm (8–12 in.) deep. Some species are solitary, while others are more social.

Jerboas in Russia, Ukraine, and Central Asia may hibernate for 6–7 months, whereas jerboas in more southern regions may not hibernate at all. Some species feed primarily on seeds and other plant material such as bulbs, roots, and rhizomes (underground horizontal stems); others are mainly insectivorous. Some fluids are obtained through their food, but jerboas also produce fluid internally as a result of metabolic processes ("metabolic water"). *See* METABOLISM.

Breeding extends from about March to October with females giving birth to two or three litters annually. Most litters contain three to five young. The gestation period is 25–40 days. Females of some species attain sexual maturity in $3^1/_2$ months; therefore, a female born early in the breeding season could give birth to her own litter during the same breeding season.

Several species of jerboas are classified as endangered, threatened, or vulnerable by the International Union for the Conservation of Nature and Natural Resources (IUCN) because of their restricted range and/or habitat disturbance. Donald W. Linzey

Bibliography. D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Jet (gemology)

A black, opaque material that takes a high polish. Jet has been used for many centuries for ornamental purposes. It is a compact variety of lignite coal. It has a refractive index of 1.66 (which accounts for its relatively high luster), a hardness of 3–4 on Mohs scale, and a specific gravity of 1.30–1.35. Jet is compact and durable, and can be carved or even turned on a lathe. The principal source is Whitby, England, where it occurs in hard shales. Although popular from Pliny's day until the nineteenth century, the use of jet for jewelry purposes declined markedly for centuries until a resurgence of interest occurred in the 1960s. *See* GEM; LIGNITE. Richard T. Liddicoat, Jr.

Jet flow

A fluid flow in which a stream of one fluid mixes with a surrounding medium, at rest or in motion. Such flows occur in a wide variety of situations, and the geometries, sizes, and flow conditions cover a large range. To create some order, two dimensionless groupings are used. The first is the Reynolds number, defined in Eq. (1), where ρ is the density, V

$$\mathrm{Re} \equiv \frac{\rho V L}{\mu} \tag{1}$$

is a characteristic velocity (for example, the jet exit velocity), L is a characteristic length (for example, the jet diameter), and μ is the viscosity. The second is the Mach number, defined in Eq. (2), where a is

$$\mathrm{M} \equiv \frac{V}{a} \tag{2}$$

the speed of sound. Jet flows vary greatly, depending on the values of these two numbers. *See* MACH NUMBER; REYNOLDS NUMBER.

Laminar and turbulent flows. Jets exiting into a similar fluid (liquid into liquid, or gas into gas) almost always become turbulent within a short distance of their origin, irrespective of whether the flow in the supply line is laminar or turbulent. For example, a water jet exiting into water at rest with Re $\approx$ 2300 (**Fig. 1**) is initially in the simple laminar state, but at

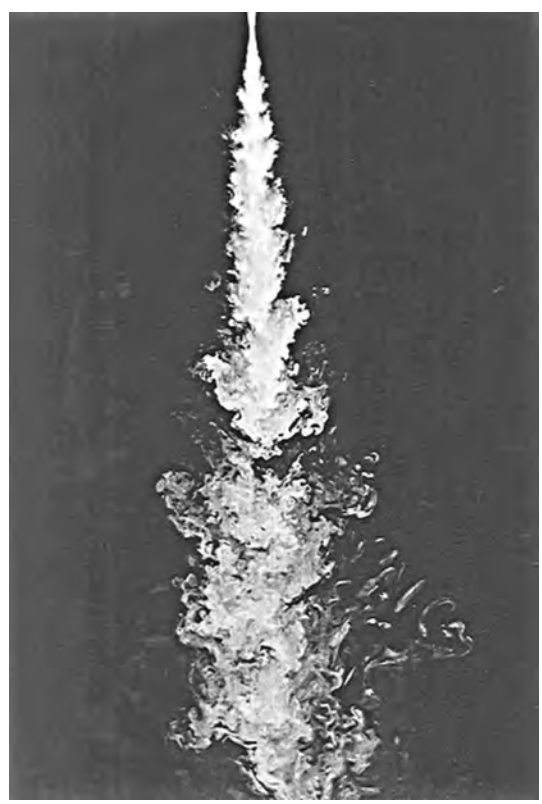

Fig. 1. Laser-induced fluorescence of a round dyed water jet directed downward into water at rest, showing the instantaneous concentration of injected fluid in the plane of the jet axis. The Reynolds number is approximately 2300. (*From P. E. Dimotakis, R. C. Lye, and D. Z. Papantoniou, Proceedings of the 15th International Symposium on Fluid Dynamics, Jachranka, Poland, 1981*)

this Reynolds number that state is unstable and the flow undergoes a transition to the more chaotic turbulent state. Turbulent structures called eddies are formed with a large range of sizes. The large-scale structures are responsible for capturing fluid from the surroundings and entraining it into the jet. However, the jet and external fluids are not thoroughly mixed until diffusion is completed by the small-scale structures. A more stable case, which may remain laminar for a long distance, is where a liquid jet issues into air or other light gas. *See* DIFFUSION; LAMINAR FLOW; TURBULENT FLOW.

High-speed flows. When the velocities in the jet are greater than the speed of sound ($M > 1$) the flow is said to be supersonic, and important qualitative changes in the flow occur. The most prominent change is the occurrence of shock waves. For example, a supersonic air jet exhausting from a nozzle at low pressure into higher-pressure air at rest is said to be overexpanded. Such a flow can be visualized by an optical technique called schlieren photography (**Fig. 2**). As the jet leaves the nozzle, it senses the higher pressure around it and adjusts through oblique shock waves emanating from the edges of the nozzle. As the waves approach the jet axis, they are focused and strengthen to form a vertical Mach disk, visible in the middle of the jet. The pattern repeats itself until dissipated by viscous effects. If the supersonic jet exits the nozzle at higher pressure than the surroundings, it will plume out. Such behavior is observed as the space shuttle rises to high altitudes with the main engines running. *See* SCHLIEREN PHOTOGRAPHY; SHOCK WAVE; SUPERSONIC FLOW.

Plumes. Another class of jet flows is identified by the fact that the motion of the jet is induced primarily by buoyancy forces. A common example is a hot gas exhaust rising in the atmosphere. Such jet flows are called buoyant plumes, or simply plumes, as distinct from the momentum jets, or simply jets, discussed above. These flows become very interesting when the surrounding fluid is stratified in different density layers, as can happen in the atmosphere and ocean by heating from the Sun. A heated plume may then rise rapidly until it reaches a layer of gas at the same density, where the buoyancy force due to temperature difference will drop to zero. The plume will then spread out in that layer without further rise. Such situations can be important for pollution-dispersal problems. *See* AIR POLLUTION; BOUNDARY-LAYER FLOW; FLUID FLOW.

Joseph A. Schetz

Fig. 2. Schlieren photograph of an overexpanded Mach 1.8 round air jet exiting from left to right into ambient air. The adjustment of the lower pressure in the surroundings leads to oblique shock waves and Mach disks in the jet. (*From H. Oertel, Modern Developments in Shock Tube Research, Shock Tube Research Society, Japan, 1975*)

Bibliography. G. N. Abramovich, *Theory of Turbulent Jets*, 1963; J. A. Schetz, *Boundary Layer Analysis*, 1993; J. A. Schetz, *Injection and Mixing in Turbulent Flow*, 1980.

Jet stream

A relatively narrow, fast-moving wind current flanked by more slowly moving currents. Jet streams are observed principally in the zone of prevailing westerlies above the lower troposphere and in most cases reach maximum intensity, with regard both to speed and to concentration, near the tropopause. At a given time, the position and intensity of the jet stream may significantly influence aircraft operations because of the great speed of the wind at the jet core and the rapid spatial variation of wind speed in its vicinity. Lying in the zone of maximum temperature contrast between cold air masses to the north and warm air masses to the south, the position of the jet stream on a given day usually coincides in part with the regions of greatest storminess in the lower troposphere, though portions of the jet stream occur over regions which are entirely devoid of cloud. The jet stream is often called the polar jet, because of the importance of cold, polar air. The subtropical jet is not associated with surface temperature contrasts, like the polar jet. Maxima in wind speed within the jet stream are called jet streaks. *See* CLEAR-AIR TURBULENCE.

Characteristics. The specific characteristics of the jet stream depend upon whether the reference is to a single instantaneous flow pattern or to an averaged circulation pattern, such as one averaged with respect to time, or averaged with respect both to time and to longitude.

If the winter circulation pattern on the Northern Hemisphere is averaged with respect to both time and longitude, a westerly jet stream is found at an elevation of about 8 mi (13 km) near latitude (lat) 25°. The speed of the averaged wind at the jet core is about 80 knots (148 km/h). In summer the jet is displaced poleward to a position near lat 41°. It is found at an elevation of about 7 mi (12 km) with a maximum speed of about 30 knots (56 km/h). In both summer and winter a speed equal to one-half the peak value is found approximately 15° of latitude south, 20° of latitude north, 3–6 mi (5–10 km) above, and 3–6 mi (5–10 km) below the location of the jet core itself.

If the winter circulation is averaged only with respect to time, it is found that both the intensity and the latitude of the westerly jet stream vary from one sector of the Northern Hemisphere to another. The most intense portion, with a maximum speed of about 100 knots (185 km/h), lies over the extreme western portion of the North Pacific Ocean

at about lat 22°. Lesser maxima of about 85 knots (157 km/h) are found at lat 35° over the east coast of North America, and at lat 21° over the eastern Sahara and over the Arabian Sea. In summer, maxima are found at lat 46° over the Great Lakes region, at lat 40° over the western Mediterranean Sea, and at lat 35° over the central North Pacific Ocean. Peak speeds in these regions range between 40 and 45 knots (74 and 83 km/h). The degree of concentration of these jet streams, as measured by the distance from the core to the position at which the speed is one-half the core speed, is only slightly greater than the degree of concentration of the jet stream averaged with respect to time and longitude. At both seasons at all longitudes the elevation of these jet streams varies between 6.5 and 8.5 mi (11 and 14 km).

Variations. On individual days there is a considerable latitudinal variability of the jet stream, particularly in the western North American and western European sectors. It is principally for this reason that the time-averaged jet stream is not well defined in these regions. There is also a great day-to-day variability in the intensity of the jet stream throughout the hemisphere. On a given winter day, speeds in the jet core may exceed 200 knots (370 km/h) for a distance of several hundred miles along the direction of the wind. Lateral wind shears in the direction normal to the jet stream frequently attain values as high as 100 knots per 300 nautical miles (185 km/h per 556 km) to the right of the direction of the jet stream current and as high as 100 knots per 100 nautical miles (185 km/h per 185 km) to the left. Vertical shears below and above the jet core as often as large as 20 knots per 1000 ft (37 km/h per 305 m). Daily jet streams are predominantly westerly, but northerly, southerly, and even easterly jet streams may occur in middle or high latitudes when ridges and troughs in the normal westerly current are particularly pronounced or when unusually intense cyclones and anticyclones occur at upper levels.

Insufficiency of data on the Southern Hemisphere precludes a detailed description of the jet stream, but it appears that the major characteristics resemble quite closely those of the jet stream on the Northern Hemisphere. The day-to-day variability of the jet stream, however, appears to be less on the Southern Hemisphere.

It appears that an intense jet stream occurs at high latitudes on both hemispheres in the winter stratosphere at elevations above 12 mi (20 km). The data available, however, are insufficient to permit the precise location or detailed description of this phenomenon. *See* AIR MASS; ATMOSPHERE; ATMOSPHERIC WAVES, UPPER SYNOPTIC; GEOSTROPHIC WIND; STORM; VORTEX

. Frederick Sanders; Howard B. Bluestein

Jewel bearing

A bearing used in quality timekeeping devices, gyros, and instruments; usually made of synthetic corundum (crystallized Al_2O_3), which is more commonly known as ruby or sapphire. The extensive use of such bearings in the design of precision devices is mainly due to the outstanding qualities of the material. Sapphire's extreme hardness imparts to the bearing excellent wear resistance, as well as the ability to withstand heavy loads without deformation of shape or structure. The crystalline nature of sapphire lends itself to very fine polishing and this, combined with the excellent oil- and lubricant-retention ability of the surface, adds to the natural low-friction characteristics of the material. Sapphire is also nonmagnetic and oxidization-resistant, and has a very high melting point (3685°F or 2029°C). Ruby has the same properties as sapphire; the red coloration is due to the introduction of a small amount of chromium oxide. *See* ANTIFRICTION BEARING; GEM; GYROSCOPE; RUBY; SAPPHIRE; WATCH.

Types. Jewel bearings, classified as either instrument or watch jewels, are also categorized according to their configuration or function. The ring jewel is the most common type. It is basically a journal bearing which supports a cylindrical pivot. The wall of the hole can be either left straight (bar hole) or can be imparted a slight curvature from end to end (olive hole). This last configuration is designed to reduce friction, compensate for slight misalignment, and help lubrication. A large variety of instrument and timing devices are fitted with such bearings, including missile and aircraft guidance systems. *See* GUIDANCE SYSTEMS.

Vee, or V, jewels are used in conjunction with a conical pivot, the bearing surface being a small radius located at the apex of a conical recess. This type of bearing is found primarily in electric measuring instruments.

Cup jewels have a highly polished concave recess mated to a rounded pivot or a steel ball. Typical are compass and electric-meter bearings.

End stone and cap jewels, combined with ring jewels, control the end play of the pivot and support axial thrust. They consist of a disk with highly polished flat or convex ends. Other relatively common jewel bearings are pallet stones and roller pins; both are part of the time- keeping device's escapement.

Fig. 1. Automatic cupping machines for the manufacture of jewel bearings. (*Bulova Watch Co.*)

Dimensions. Minute dimensions are a characteristic of jewel bearings. A typical watch jewel may be 0.040 in. (0.10 cm) in diameter with a 0.004-in. (0.01-cm) hole, but these dimensions may go down to 0.015 and 0.002 in. (0.038 and 0.005 cm), respectively. Jewels with a diameter of more than $^1/_{\sqrt{16}}$ in. (0.16 cm) are considered large. It is usual for critical dimensions, such as hole diameter and roundness, to have a tolerance of 0.0001 in. (0.00025 cm) or less. In some instances these tolerances may be as low as 0.0000020 in. (0.000005 cm).

Manufacturing. Because of its hardness sapphire can only be worked by diamond, which is consequently the main tool for the production of jewel bearings. Both natural and synthetic diamond are used, mostly under the form of slurry, broaches, and grinding wheels. *See* DIAMOND.

The machining of the blanks, small disks, or cylinders of varied diameter and thickness is the first step in the manufacturing process, and is common to most types of jewel bearings. The boules (pear-shaped crystals of synthetic ruby or sapphire) are first oriented according to the optical axis to ensure maximum hardness of the bearing working surface. They are then sliced, diced, and ground flat, and rounded by centerless grinding to the required blank dimensions. From this point on, the process varies considerably, according to the type of bearing.

Ring jewels are drilled with a steel or tungsten wire and coarse diamond slurry, or bored with a small grinding tool. The hole is then enlarged and sized by reciprocating wires of increasingly larger sizes through a string of jewels, until the required hole size is achieved. Fine diamond powder and a very slow rate of material removal permit the respect of strict tolerances and high-finish quality requirements. The jewels, supported by a wire strung through the hole, are then ground in a special centerless-type grinding machine to the desired outside diameter dimension. After the cutting of a concave recess, which functions as an oil reservoir, the top and bottom of the bearing are polished and beveled by lapping and brushing. Finally, the "olive" configuration is obtained by oscillating a diamond charged wire through the hole of the rotating jewel. Between each operation the boiling of the jewels in a bath of sulfuric and nitric acid disposes of remaining slurries and other contaminating products.

Fig. 2. Individual head of an automatic cupping machine. (*Bulova Watch Co.*)

Fig. 3. Technician shown operating an automatic cupping machine. (*Bulova Watch Co.*)

The conical recess in vee jewels is first roughly shaped by a pyramidal diamond tool. The wall of the vee and the radius are then polished and blended with an agglomerated diamond broach, and a succession of brushing operations. Lapping of the top of the jewel and brushing a bevel around the upper outside edge conclude the process.

Most other types of jewel bearings, such as end stones and pallet stones, are shaped by a series of grinding, lapping, and brushing operations.

A full line of automatic and semiautomatic high-precision equipment has been developed to handle and machine jewel bearings efficiently, permitting mass production and relatively low cost (**Figs. 1–3**). Traditionally, a large proportion of the labor involved is devoted to in-process and final quality control.

Robert M. Schultz

Johne's disease

A slowly progressive diarrheal disease that causes major economic loss to the cattle industry; also known as paratuberculosis. It is caused by *Mycobacterium avium ss paratuberculosis (MAP)*, which produces chronic inflammation of the mucosa of the ileocecal valve and adjacent tissues of the gastrointestinal tract of cattle, sheep, goats, and wild ruminants. The organism has been isolated from 2–3% of the adult cows in the United States; in some exotic species (such as mouflon and mountain goats), more than 90% of the animals in a herd may be infected. Animals less than 2 months of age are most susceptible to infection. Therefore, eliminating or minimizing exposure of young animals to *MAP* is important in the control of Johne's disease. It is recommended that calves be removed from dams immediately after birth and fed colostrum from negative animals. *See* MYCOBACTERIAL DISEASES.

Transmission. Transmission of *MAP* is primarily by ingestion of feces from animals shedding the organism. The incubation period varies from 1 to 3 years

or more. Diseased animals in early or subclinical stages have intermittent or persistent diarrhea without fever and in advanced stages become emaciated. Congenital infections have been reported, and the organism has been isolated from colostrum, supramammary lymph nodes, fetal tissues, blood, and semen of cattle with clinical disease.

Diagnosis. Johne's disease is diagnosed by mycobacteriologic examinations conducted on feces or tissues collected by biopsy or at necropsy. Enzyme-linked immunosorbent assays (ELISA) have been developed using a purified protein derivative prepared from the culture filtrate of *MAP* or chemical extracts of the bacterium, including proteins and carbohydrates (antigens) for detecting antibodies in sera of animals exposed to the organism. ELISA reportedly fails to identify three of four contagious cows shedding *MAP* in feces. Moreover, animals in the early stages of disease are frequently negative for humoral antibodies on ELISA, complement fixation, and/or agar gel immunodiffusion tests. Therefore, to confirm a diagnosis of Johne's disease, it is necessary to isolate and identify the etiologic agent by mycobactin dependency and/or polymerase chain reaction (PCR).

Vaccination. A killed whole-cell vaccine available for use in calves 1–35 days of age markedly reduces and often eliminates the occurrence of clinical disease in adult cattle. Live attenuated strains of *MAP* have been used for vaccinating cattle in a few countries, but are not approved for use in the United States or Canada. Therapeutic drugs are not available for routine treatment of animals.

Relation to Crohn's disease. The superficial similarity of Johne's disease to Crohn's disease in humans has led to the hypothesis that Crohn's disease is also caused by *MAP*. Although *MAP* has been isolated from some Crohn's disease patients, no definitive information on association is available. Moreover, published information on Crohn's disease in humans fails to show any association with bovine paratuberculosis.

Available information indicates that the parasitic worm *Trichuris suis* may be useful in the treatment of patients with Crohn's disease. This helminth may play a role in altering host mucosal immunity by inhibiting dysregulated inflammatory responses. *See* INFLAMMATORY BOWEL DISEASE. Charles O. Thoen

Bibliography. O. Chacon, L. E. Bermudez, and R. G. Baretta, Johne's disease, inflammatory bowel disease and *Mycobacterium paratuberculosis*, *Annu. Rev. Microbiol.*, 58:329–363, 2004; P. H. Jones et al., Crohn's disease in people exposed to clinical cases of bovine paratuberculosis, *Epidemiol. Infect.*, pp. 1–8, January 2005; R. W. Summers et al., *Trichuris suis* therapy in Crohn's disease, *Gut*, 54:87–89, 2005; C. O. Thoen and R. G. Barletta, *Mycobacterium*, in *Pathogenesis of Bacterial Infections in Animals*, pp. 70–77, Blackwell, 2004; C. O. Thoen and J. A. Haagsma, Molecular techniques in the diagnosis and control of paratuberculosis in cattle, *J. Amer. Vet. Med. Ass.*, 209:734–737, 1996; J. V. Weinstock, R. W. Summers, and D. E. Elliott, Role of helminths in regulating mucosal inflammation, *Springer Seminars Immunopathol.*, 27:249–271, 2005.

Joint (anatomy)

The structural component of an animal skeleton where two or more skeletal elements meet, including the supporting structures within and surrounding it. The relative range of motion between the skeletal elements of a joint depends on the type of material between these elements, the shapes of the contacting surfaces, and the configuration of the supporting structures.

In bony skeletal systems, there are three general classes of joints: synarthroses, amphiarthroses, and diarthroses. Synarthroses are joints where bony surfaces are directly connected with fibrous tissue, allowing very little if any motion. Synarthroses may be further classified as sutures, syndesmoses, and gomphoses. Sutures are joined with fibrous tissue, as in the coronal suture where the parietal and frontal bones of the human skull meet; in adulthood, this suture becomes an essentially rigid bony union called a synostosis. Syndesmoses are connected with ligaments, as are the shafts of the tibia and fibula. The roots of a tooth that are anchored in the jaw bone with fibrous tissue form a gomphosis. Amphiarthroses are joints where bones are directly connected with fibrocartilage or hyaline cartilage and allow only limited motion. An amphiarthrosis joined with fibrocartilage, as found between the two pubic bones of the pelvis, is known as a symphysis; but when hyaline cartilage joins the bones, a synchondrosis is formed, an example being the first sternocostal joint. The greatest range of motion is found in diarthrodial joints, where the articulating surfaces slide and to varying degrees roll against each other. *See* LIGAMENT.

Diarthrodial joints. The contacting surfaces of the bones of a diarthrodial joint are covered with articular cartilage, an avascular, highly durable hydrated soft tissue that provides shock absorption and lubrication functions to the joint (**Fig. 1**). Articular cartilage is composed mainly of water, proteoglycans, and collagen. The joint is surrounded by a fibrous joint capsule lined with synovium, which produces lubricating synovial fluid and nutrients required by the tissues within the joint. Joint motion is provided by the muscles that are attached to the bone with tendons. Strong flexible ligaments connected across the bones stabilize the joint and may constrain its motion.

Different ranges of motion result from several basic types of diarthrodial joints: pivot, gliding, hinge, saddle, condyloid, and ball-and-socket. *See* COLLAGEN.

Pivot. Pivot (trochoid) joints allow rotation of one bone about another, as does the median atlantoaxial joint (atlas about the second cervical vertebra) at the base of the skull.

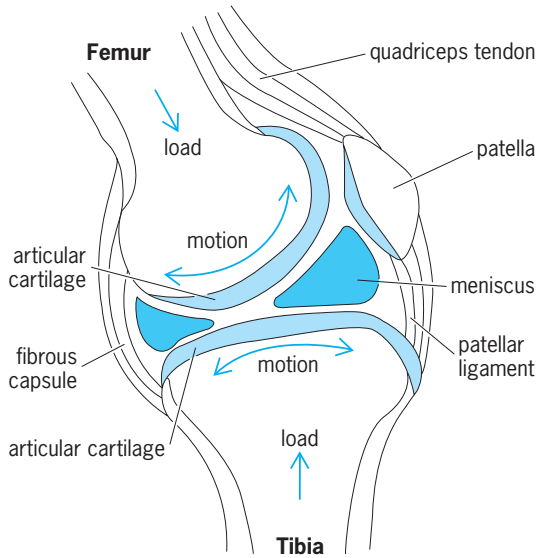

Fig. 1. Cross section of the human knee showing its major components. This diarthrodial joint contains contacting surfaces on the tibia, femur, meniscus, and patella (knee cap). The patella protects the joint and also serves to redirect the force exerted by the quadriceps muscles to the tibia. (*After R. Skalak and S. Chien, eds., Handbook of Bioengineering, McGraw-Hill, 1987*)

Gliding. Some of the joints found in the wrist and ankle are examples of gliding (arthrodia) joints. The surfaces of these joints slide against one another and provide limited motion in multiple directions.

Hinge. Although the hinge (ginglymus) joint generally provides motion in only one plane (flexion and extension), its motion, as in the human knee, can be complex. The articulating surfaces of the knee both roll and slide past each other, and the motion of the tibia relative to the femur displays a small amount of rotation. The knee also contains two fibrocartilaginous (primarily collagenous) crescent-shaped discs between the surfaces of the tibia and femur, called menisci, which provide shock absorption, stability, and lubrication functions to the knee (Fig. 1). The elbow is another example of a hinge joint.

Saddle. This joint is found at the base of the human thumb between the carpal and metacarpal (trapezium). It is characterized by saddle-shaped (concave and convex) articulating surfaces. The saddle joint provides a large range of motion (flexion, extension, abduction and adduction) with limited rotation.

Condyloid. The range of motion of the condyloid joint is similar to that of the saddle joint. Here, the articulating surfaces are ovoid, and one surface is convex, the other concave. This joint is found at the base of the fingers where the metacarpal and proximal phalanx meet. The surfaces of the knee joint are also of condyloid shapes (**Fig. 2**).

Ball-and-socket. The human shoulder and hip are examples of the ball-and-socket (enarthrosis) joint. In addition to flexion, extension, abduction, and adduction, they provide rotation about the long axis of the bone. The low relative-surface-area ratio of the glenoid (socket) to humeral head (ball) of the shoulder gives it the largest range of motion of all human joints. This high degree of mobility is stabilized by a complex rotator cuff consisting of ligaments and muscular attachments. The larger ratio of acetabulum (socket) to femoral head (ball) surface area of the hip constrains the joint to lesser mobility but provides it with greater inherent stability.

Lubrication. The lubrication mechanisms of diarthrodial joints are extremely effective. Although the surface roughness of articular cartilage is higher than that of some engineered bearing surfaces, the coefficient of friction between cartilage surfaces is significantly lower than between most manufactured materials. The rate of wear of the articular surfaces is also extremely low in the normally functioning joint. While many details of diarthrodial joint lubrication mechanisms are not fully understood, these mechanisms fall into two categories: fluid-film lubrication, where the joint surfaces are separated by a layer of lubricating fluid (synovial fluid); and boundary lubrication, where the joint surfaces are in contact with each other. These mechanisms probably occur together to varying extents, depending upon the particular joint and the type of motion.

Some of the major theories involving fluid-film lubrication are hydrodynamic, squeeze film, hydrostatic, and boosted lubrication.

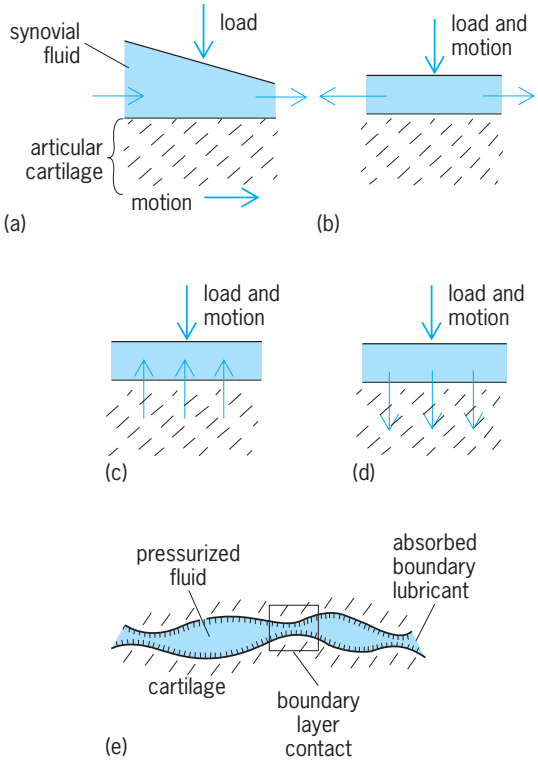

Fig. 2. Possible modes of lubrication in diarthrodial joints. (*a–d*) Different types of fluid-film lubrication showing the directions of loading, motion, and fluid flow: (*a*) hydrodynamic; (*b*) squeeze film; (*c*) hydrostatic weeping; (*d*) boosted. (*e*) A combination of boundary layer lubrication at points of contact and fluid-film lubrication. (*After V. C. Mow and W. C. Hayes, eds., Basic Orthopaedic Biomechanics, Raven Press, 1991*)

Hydrodynamic. If the relative velocity between the joint surfaces is sufficiently high, a pressurized, wedge-shaped layer of the intervening synovial fluid may be formed (Fig. 2*a*). This fluid layer generates lift which forces the joint surfaces apart. Hydrodynamic lubrication probably does not occur to any great extent in diarthrodial joints, except under conditions of low load and high velocity that might occur, for example, during high-speed pitching of a ball.

Squeeze film. When both surfaces of a joint approach each other, pressure in the synovial fluid is generated because this viscous lubricant cannot instantaneously be squeezed from the gap (Fig. 2*b*). The resulting pressurized squeeze film can support large loads which may be generated, for example, in the knee joint when jumping from a height. Since the articular cartilage of the joint surfaces is deformable, the synovial fluid may become trapped in localized depressions. This film may support large loads for many minutes before becoming depleted.

Hydrostatic. During normal joint loading, compression of the articular cartilage surfaces may result in the weeping of fluid from the cartilage into the joint cavity (Fig. 2*c*). Weeping lubrication, a form of hydrostatic lubrication, is facilitated by this self-pressurization of the fluid within the cartilage.

Boosted. As the surfaces of the joint approach each other during normal function, it is possible that the resulting pressure causes some of the water in the synovial fluid to be forced into the cartilage (Fig. 2*d*). Boosted lubrication results from the remaining, highly viscous synovial fluid concentrate between the two articulating surfaces.

Boundary. Boundary layer lubrication takes place when the surfaces of the joints are touching each other. It has been proposed that a monolayer of a polypeptide chain, called lubricin, is adsorbed onto the cartilage surface. This boundary lubricant separates the surfaces during normal joint function and provides the ability to carry weight and reduce friction. Fluid-film lubrication probably occurs simultaneously in the gaps between the joint surfaces that result from undulations in the articular cartilage surfaces (Fig. 2*e*).

The types of lubrication mechanisms active in a joint depend on the contact pattern of the surfaces, the applied load, and the relative velocity of the surfaces. The contact patterns between joint surfaces vary with different joints and also vary within a joint; depending on the relative positioning of the joint surfaces during a particular motion. The applied load and relative velocity of the surfaces depend on the particular action of the joint. Thus, different lubrication modes may occur within the same joint for different motions or even for the same motion, depending on the applied load and the speed with which the motion is accomplished. *See* SKELETAL SYSTEM.

Van C. Mow; Robert J. Foster

Bibliography. V. C. Mow and W. C. Hayes (eds.), *Basic Orthopaedic Biomechanics*, 2d ed., 1997; V. C. Mow, A. Ratcliffe, and S. L. Y. Woo (eds.), *Biomechanics of Diarthrodial Joints*, vols. 1 and 2, 1990; C. Norkin and P. K. Levangie, *Joint Structure and Function*, 3d ed., 2000.

Joint (structures)

The surface at which two or more mechanical or structural components are united. Whenever parts of a machine or structure are brought together and fastened into position, a joint is formed. *See* STRUCTURAL CONNECTIONS.

Mechanical joints can be fabricated by a great variety of methods, but all can be classified into two general types, temporary (screw, snap, or clamp, for example), and permanent (brazed, welded, or riveted, for example). The following list includes many of the more common methods of forming joints. (1) Screw threads: bolt and nut, machine screw and nut, machine screw into tapped hole, threaded parts (rod, pipe), self-tapping screw, lockscrew, studs with nuts, threaded inserts, coiled-wire inserts, drive screws. (2) Rivets: solid, hollow, explosive and other blind side types. (3) Welding. (4) Soldering. (5) Brazing. (6) Adhesive. (7) Friction-held: nails, dowels, pins, clamps, clips, keys, shrink and press fits. (8) Interlocking: twisted tabs, snap ring, twisted wire, crimp. (9) Other: peening, staking, wiring, stapling, retaining rings, magnetic. Also, pipe joints are made with screw threads, couplings, caulking, and by welding or brazing; masonry joints are made with cement mortar. *See* ADHESIVE; ASPHALT AND ASPHALTITE; BOLT; BOLTED JOINT; BRAZING; MUCILAGE; NUT (ENGINEERING); RIVET; SCREW; SCREW FASTENER; SOLDERING; WASHER; WELDED JOINT; WELDING AND CUTTING OF MATERIALS.

William H. Crouse

Joint disorders

Local or generalized conditions that affect the joints and related tissues. About one person in 16 who is over 15 years old has some form of joint disturbance. The most common conditions are forms of arthritis, which cause the inflammation or degeneration of joint structures. These may follow specific infections, injury, generalized disease states, or degenerative changes which largely parallel the aging processes in the body.

Rheumatoid arthritis, one of the most common diseases which involve joints, is of unknown etiology and occurs most often in young or middle-aged adults. Multiple joints are involved, especially those of the small bones in the hands, feet, and wrists or the jaw. Other connective tissue diseases, such as systemic lupus erythematosus and vasculitis, frequently are accompanied by joint symptomatology.

Degenerative joint disease (osteoarthritis) is a ubiquitous joint disease prevalent in older people. This form of arthritis is characterized by deterioration of the articular surfaces of the weight-bearing joints of the knees, hips, and spine. Frequent involvement is seen at the distal interphalangeal joints

(Heberden's nodes). Arthritis produced by trauma may also result in degenerative joint disease.

Infectious arthritis may be caused by gonorrhea, pneumococci, streptococci, and other pathogenic organisms such as tuberculosis and syphilis.

Gout is an example of crystal-deposition arthritis caused by the presence of monosodium urate crystals in synovial fluid. There is an abnormality of purine metabolism that eventuates in the accumulation of urates and uric acid in joint tissues. The physical findings and clinical course are usually distinctive. *See* GOUT.

Spondyloarthropathies constitute a group of arthritis disorders characterized by involvement of the axial (central) skeleton and association with the histocompatibility system.

Miscellaneous forms of arthritis include those systemic diseases of unknown or quite different etiology which may produce arthritis or joint degeneration. Syringomyelia, syphilis (tabes dorsalis), and leprosy are examples in which destructive joint lesions may be secondary to nerve damage of the affected joints (Charcot joints). Hemophilia and sickle-cell anemia (blood disorders) and acromegaly, hypothyroidism, and hyperparathyroidism (endocrine disorders) are also examples of these unusual rheumatic diseases. *See* ARTHRITIS.

Tumors of joint tissues are usually benign chondromas, fibromas, and lipomas, but malignant synovio sarcoma can occur. *See* TUMOR.

Rheumatism is a nonspecific, predominantly lay term which includes local pain and tenderness. Most causes are chronic inflammations or mildly progressive degenerations, and many, when investigated, fall into one of the previously mentioned categories. *See* BURSITIS; RHEUMATISM.

Robert Searles

Jojoba

Simmondsia chinensis, the only plant known to produce and store a liquid wax in its seed. The jojoba plant is native to the southwestern United States and Mexico. It is tolerant of some of the highest temperatures and most arid regions, and is being domesticated as a crop for hot low-rainfall regions around the world. A broadleaf evergreen shrub that is typically 3–10 ft (1–3 m) in height, it can grow as tall as 20 ft (6 m).

Whereas the seed-storage lipid of other oilseed plants such as soybean and sunflower is a branched ester based on the three-carbon glycerol molecule, that of jojoba is a straight-chain ester. A majority of the wax molecules of jojoba are formed from acids and alcohols with 20 or 22 carbon atoms and one double bond. Many modifications can be made at the double bond, which results in the plant's versatility as an ingredient in a wide range of chemical products. Jojoba wax, used in cosmetics and lubricants, has the potential to serve as a basic feedstock if seed production costs are reduced. Jojoba is a relatively new commercial crop that is being developed simultaneously in many places around the world, and cultivation methods are variable and change rapidly. *See* FAT AND OIL (FOOD); WAX, ANIMAL AND VEGETABLE.

Site and variety selection. Jojoba is a perennial, and flower buds that are formed in the spring, summer, and fall of one year must overwinter before producing a crop in the following year. The buds can succumb at temperatures below −3.5°C (25°F), and so sites where such temperatures are expected are unsuitable for planting. Jojoba is sensitive to standing water, and so sites with well-drained soil away from potential flooding should be selected.

Because widely tested cultivars of jojoba are unavailable, if jojoba is being introduced to an area, small test plantings should first be established and locally adapted varieties selected. Jojoba is dioecious, and both male and female plants are needed for seed production.

Propagation. Jojoba is easy to grow from seed. Germination is best in the temperature range of 68–95°F (20–35°C). Seed-propagated plants are extremely variable in yield, size, sex, and other traits, and so use of seed is not recommended for management systems that require high yield or plant uniformity.

Several asexual methods of propagation are effective for jojoba, including stem cuttage, micropropagation, layering, and grafting. These methods allow propagation of selected plants that have a high yield potential or other desired traits. Stem cuttage is the most widely used. Stem tips with four to six pairs of leaves are collected during the warmer months of the year, basally treated with an auxin, and then rooted by using standard mist propagation techniques. Rooting occurs in 4–8 weeks, and plants are ready for field planting in 4–6 months. *See* REPRODUCTION (PLANT).

Irrigation. Although jojoba grows under natural rainfall conditions in several countries, including Paraguay and Australia, and has grown well in some plantings with natural rainfall in Arizona, most growers irrigate with a drip, furrow, or sprinkler system. Water requirements of jojoba for profitable cultivation are not well defined. In nature, jojoba grows where average annual rainfall is as low as 3 in. (76 mm), but growth is confined to washes, where rainfall is concentrated. The most vigorous plants in nature occur where average annual rainfall is 10–20 in. (250–500 mm). In commercial plantings, good yields have been obtained by using about 1.5 ft (460 mm) of water in the central valley of California and about 2 ft (600 mm) in south-central Arizona. Some growers, however, use 4 ft (1200 mm) or more. Restricting irrigation as winter approaches appears to improve frost resistance. *See* IRRIGATION (AGRICULTURE).

Insects and diseases. Several pest and disease problems have been encountered for which control strategies must be developed. Damage has resulted from attacks by spider mites and several types of insects including stinkbugs, inchworms, thrips, grasshoppers, and lepidopterous larvae. In addition, relatively innocuous infestations with aphids, scales, and mealy bugs have been observed.

Several root and foliar fungal pathogens have been isolated from diseased jojoba plants growing in both field and nursery environments. To date, virus diseases have not been found in association with jojoba. *See* PLANT PATHOLOGY.

Harvesting and processing. Although economic conditions in many countries permit hand harvesting, successful cultivation in the United States will probably rely on an efficient mechanical harvester. Growers have evaluated many types of mechanical harvesters, some of which shake the plants and catch the seed as it falls while others vacuum the seed from the ground after it has dropped. Harvesting requires several passes because of nonuniform maturity. Once collected, the seeds are cleaned by using standard technology developed for other oilseeds. The wax is then pressed or solvent-extracted from the seed. After filtering, a transparent, yellowish oil remains that may undergo an additional decolorization process.

David A. Palzkill

Bibliography. H. S. Gentry, The natural history of jojoba (*Simmondsia chinensis*) and its cultural aspects, *Econ. Bot.*, 12:261–295, 1958; G. Robbelen, R. K. Downey, and D. A. Ashri (eds.), *Oil Crops of the World*, 1989; W. C. Sherbrooke and E. F. Haase, *Jojoba: A Wax Producing Shrub of the Sonoran Desert*, 1974; J. Wisniak, *The Chemistry and Technology of Jojoba Oil*, 1987.

Josephson effect

The passage of paired electrons (Cooper pairs) through a weak connection (Josephson junction) between superconductors, as in the tunnel passage of paired electrons through a thin dielectric layer separating two superconductors.

Nature of the effect. Quantum-mechanical tunneling of Cooper pairs through a thin insulating barrier (on the order of a few nanometers thick) between two superconductors was theoretically predicted by Brian D. Josephson in 1962. Josephson calculated the currents that could be expected to flow during such superconductive tunneling, and found that a current of paired electrons (supercurrent) would flow in addition to the usual current that results from the tunneling of single electrons (single or unpaired electrons are present in a superconductor along with bound pairs). Josephson specifically predicted that if the current did not exceed a limiting value (the critical current), there would be no voltage drop across the tunnel barrier. This zero-voltage current flow resulting from the tunneling of Cooper pairs is known as the dc Josephson effect. Josephson also predicted that if a constant nonzero voltage V were maintained across the tunnel barrier, an alternating supercurrent would flow through the barrier in addition to the direct current produced by the tunneling of unpaired electrons. The frequency ν of the alternating supercurrent is given by Eq. (1), where e is

$$\nu = \frac{2eV}{h} \tag{1}$$

the magnitude of the charge of an electron and h is Planck's constant. The oscillating current of Cooper pairs that flows when a steady voltage is maintained across a tunnel barrier is known as the ac Josephson effect. Josephson further predicted that if an alternating voltage at frequency f were superimposed on the steady voltage applied across the barrier, the alternating supercurrent would be frequency-modulated and could have a dc component whenever ν was an integral multiple of f. Depending upon the amplitude and frequency of the ac voltage, the dc current-voltage characteristic would display zero-resistance parts (constant-voltage steps) at voltages V given by Eq. (2), where n is any integer. Finally, Josephson

$$V = \frac{nhf}{2e} \tag{2}$$

predicted that effects similar to the above would also occur for two superconducting metals separated by a thin layer of nonsuperconducting (normal) metal. In 1963 the existence of the dc Josephson effect was experimentally confirmed by P. W. Anderson and J. M. Rowell, and the existence of the ac Josephson effect was experimentally confirmed by S. Shapiro. *See* NONRELATIVISTIC QUANTUM THEORY; TUNNELING IN SOLIDS.

Theory of the effect. The superconducting state has been described as a manifestation of quantum mechanics on a macroscopic scale, and the Josephson effect is best explained in terms of phase, a basic concept in the mathematics of quantum mechanics and wave motion. For example, two sine waves of the same wavelength λ are said to have the same phase if their maxima coincide, and to have a phase difference equal to $2\pi/\lambda$ if their maxima are displaced by a distance δ. An appreciation of the importance that phase can have in physical systems can be gained by considering the radiation from excited atoms in a ruby rod. For a given transition, the atoms emit radiation of the same wavelength; if the atoms also emit the radiation in phase, the result is the ruby laser.

According to the Bardeen-Cooper-Schrieffer (BCS) theory of superconductivity, an electron can be attracted by the deformation of the metal lattice produced by another electron, and thereby be indirectly attracted to the other electron. This indirect attraction tends to unite pairs of electrons having equal and opposite momentum and antiparallel spins into the bound pairs known as Cooper pairs. In the quantum-mechanical description of a superconductor, all Cooper pairs in the superconductor have the same wavelength and phase. It is this phase coherence that is responsible for the remarkable properties of the superconducting state. The common phase of the Cooper pairs in a superconductor is referred to simply as the phase of the superconductor. *See* COHERENCE; PHASE (PERIODIC PHENOMENA); QUANTUM MECHANICS.

The phases of two isolated superconductors are totally unrelated, while two superconductors in perfect contact have the same phase. If the

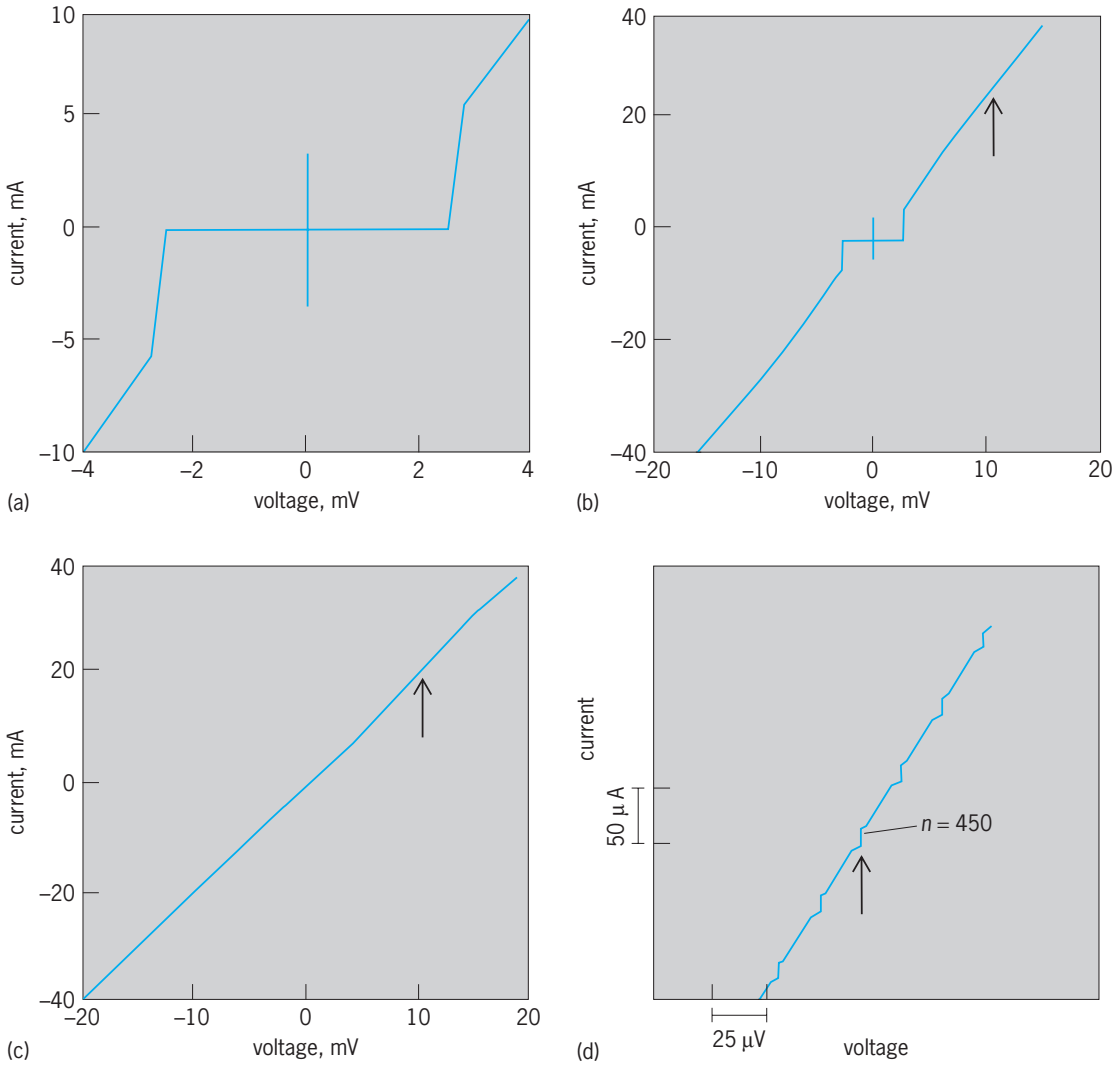

Fig. 1. DC current-voltage characteristics of lead-lead oxide-lead Josephson tunnel junction at 1.2 K. (*a*) Without microwave power. (*b*) Same characteristic with reduced scale. (*c*) 11-GHz microwave power applied. (*d*) Expanded portion of *c*; arrow indicates a constant-voltage step near 10.2 mV corresponding to $n = 450$ in Eq. (2). This voltage is also indicated by arrows in *b* and *c*. (*After T. F. Finnegan, A. Denenstein, and D. N. Langenberg, AC-Josephson-effect determination of e/h: A standard of electrochemical potential based on macroscopic quantum phase coherence in superconductors, Phys. Rev. B, 4:1487–1522, 1971*)

superconductors are weakly connected (as they are when separated by a sufficiently thin tunnel barrier), the phases can be different but not independent. If ϕ is the difference in phase between superconductors on opposite sides of a tunnel barrier, the results of Josephson's calculation of the total current I through the junction can be written as Eq. (3), where I_0 is the

$$I = I_0 + I_1 \sin \phi \tag{3}$$

current due to single electron tunneling, and $I_1 \sin \phi$ is the current due to pair tunneling. The time dependence of the phase is given by Eq. (4). In general, the

$$\frac{\partial \phi}{\partial t} = 2\pi \frac{2eV}{h} \tag{4}$$

currents I, I_0, and I_1 are all functions of the voltage across the junction. For $V = 0$, I_0 is zero and ϕ is constant. The value of I_1 depends on the properties of the tunnel barrier, and the zero-voltage supercurrent is a sinusoidal function of the phase difference between the two superconductors. However, it is not the phase difference that is under the control of the experimenter, but the current through the junction, and the phase difference adjusts to accommodate the current. The maximum value $\sin \phi$ can assume is 1, and so the zero-voltage value of I_1 is the critical current of the junction.

Integration of Eq. (4) shows the phase changes linearly in time for a constant voltage V maintained across the barrier, and the current through the barrier is given by Eq. (5), where ϕ_0 is a constant. The

$$I = I_0 + I_1 \sin\left(2\pi \frac{2eV}{h} t + \phi_0\right) \tag{5}$$

supercurrent is seen to be an alternating current with frequency $2eV/h$. The supercurrent time-averages to

zero, and so the direct current through the barrier is just the single-electron tunneling current I_0.

If the voltage across the junction is $V + v \cos 2\pi ft$, Eqs. (3) and (4) give Eq. (6) for the current. The expression for the supercurrent is a conventional expression in frequency-modulation theory and can be rewritten as expression (7), where J_n is an integer order Bessel function of the first kind. This expression time-averages to zero except when $V = nhf/2e$, in which case the supercurrent has a dc component given by $(-1)^n I_1 J_n(2ev/hf) \sin \phi_0$. As for the zero-voltage direct supercurrent, the phase difference ϕ_0 adjusts to accommodate changes in current at this value of V, and the dc current-voltage characteristic displays a constant-voltage step.

$$I = I_o + I_1 \sin\left(2\pi \frac{2eV}{h} t + \phi_0 + \frac{2ev}{hf} \sin 2\pi ft\right) \quad (6)$$

$$I_1 \sum_{n=-\infty}^{n=\infty} +(-1)^n J_n \frac{2ev}{hf} \times \sin\left(2\pi \frac{2eV}{h} t - 2\pi nft + \phi_0\right) \quad (7)$$

The dc current-voltage characteristic of a Josephson tunnel junction with and without a microwave-frequency ac voltage is shown in **Fig. 1**. The straightening of the current-voltage characteristic in the presence of microwave power displayed in Fig. 1*c* is due to the phenomenon of photon-assisted tunneling, which is essentially identical to classical rectification for the junction and frequency in question. *See* BESSEL FUNCTIONS.

Josephson pointed out that the magnitude of the maximum zero-voltage supercurrent would be reduced by a magnetic field. In fact, the magnetic field dependence of the magnitude of the critical current is one of the more striking features of the Josephson effect. Circulating supercurrents flow through the tunnel barrier to screen an applied magnetic field from the interior of the Josephson junction just as if the tunnel barrier itself were weakly superconducting. The screening effect produces a spatial variation of the transport current, and the critical current goes through a series of maxima and minima as the field is increased. **Figure 2** shows the variation of the critical current with magnetic field for a tunnel junction whose length and width are small in comparison with the characteristic screening length of the junction (the Josephson penetration depth, λ_J). The mathematical function which describes the magnetic field dependence of the critical current for this case is the same function as that which describes the diffraction pattern produced when light passes through a single narrow slit. *See* DIFFRACTION.

Josephson junctions. The weak connections between superconductors through which the Josephson effects are realized are known as Josephson junctions. Historically, superconductor-insulator-superconductor tunnel junctions have been used to study the Josephson effect, primarily because these are physical situations for which detailed calculations can be made. However, the Josephson effect is not necessarily a tunneling phenomenon, and the Josephson effect is indeed observed in other types of junctions, such as the superconductor-normal metal-superconductor junction. A particularly useful Josephson junction, the point contact, is formed by bringing a sharply pointed superconductor into contact with a blunt superconductor. The critical current of a point contact can be adjusted by changing the pressure of the contact. The low capacitance of the device makes it well suited for high-frequency applications. Thin-film microbridges form another group of Josephson junctions. The simplest microbridge is a short narrow constriction (length and width on the order of a few micrometers or smaller) in a superconducting film known as the Anderson-Dayem bridge. If the microbridge region is also thinner than the rest of the superconducting film, the resulting variable-thickness microbridge has better performance in most device applications. If a narrow strip of superconducting film is overcoated along a few micrometers of its length with a normal metal, superconductivity is weakened beneath the normal metal, and the resulting microbridge is known as a proximity-effect or Notarys-Mercereau microbridge. Among the many other types of Josephson junctions are the superconductor-semiconductor-superconductor and other artificial-barrier tunnel junctions, superconductor-oxide-normal metal-superconductor junctions, and the so-called SLUG junction, which consists of a drop of lead-tin solder solidified around a niobium wire. Some different types of Josephson junctions are illustrated in **Fig. 3**.

The dc current-voltage characteristics of different types of Josephson junctions may differ, but all show a zero-voltage supercurrent, and constant-voltage steps can be induced in the dc characteristics

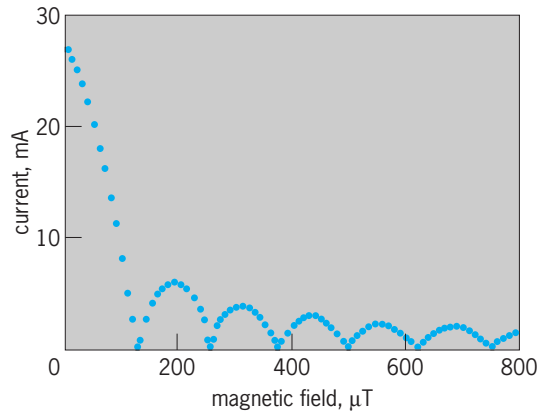

Fig. 2. Magnetic field dependence of the critical current of a Josephson tunnel junction. Data are for a tin-tin oxide-tin junction at 1.2 K, with the magnetic field in the plane of the barrier. (*After D. N. Langenberg, D. J. Scalapino, and B. N. Taylor, The Josephson effects, Sci. Amer., 214(5):30–39, May 1966*)

at voltages given by Eq. (2) when an ac voltage is applied. The dc characteristics of a microbridge and a tunnel junction are compared in **Fig. 4**.

Applications. From July 1972 through December 1989, the United States legal volt, V_{NBS}, was defined by Eq. (1) through the assigned value given by Eq. (8),

$$\frac{2e}{h} = 483,593,420 \text{ MHz/V}_{NBS} \tag{8}$$

and was maintained at the National Institute of Standards and Technology (formerly the National Bureau of Standards) within a few parts in 10^8 by using the ac Josephson effect. By international agreement, a new practical reference standard for the volt, based on the ac Josephson effect, was adopted worldwide on January 1, 1990. This internationally adopted representation of the volt is given by Eq. (9), where K_J

$$K_{J-90} = 483,597.9 \text{ GHz/V} \tag{9}$$

is called the Josephson constant and is theoretically equal to $2e/h$. This developed as a natural consequence of extremely precise measurements of $2e/h$ via the Josephson effect, and the recognition that a Josephson junction is a precise frequency-to-voltage converter and that atomic frequency standards are inherently more stable than electrochemical voltage standards. *See* ELECTRICAL UNITS AND STANDARDS; FUNDAMENTAL CONSTANTS.

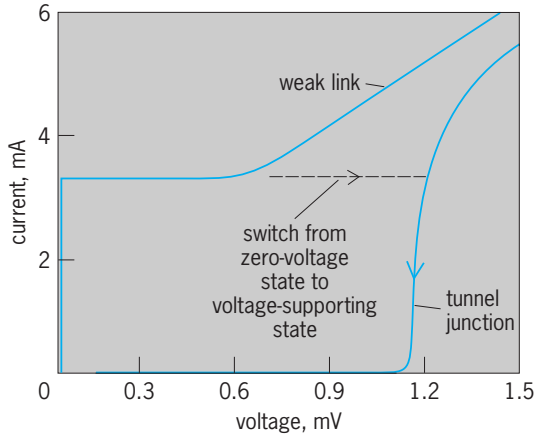

Fig. 4. DC current-voltage characteristics for a weak link and a tunnel junction. (*After E. Burstein and S. Lundqvist, eds., Tunneling Phenomena in Solids, Plenum, 1969*)

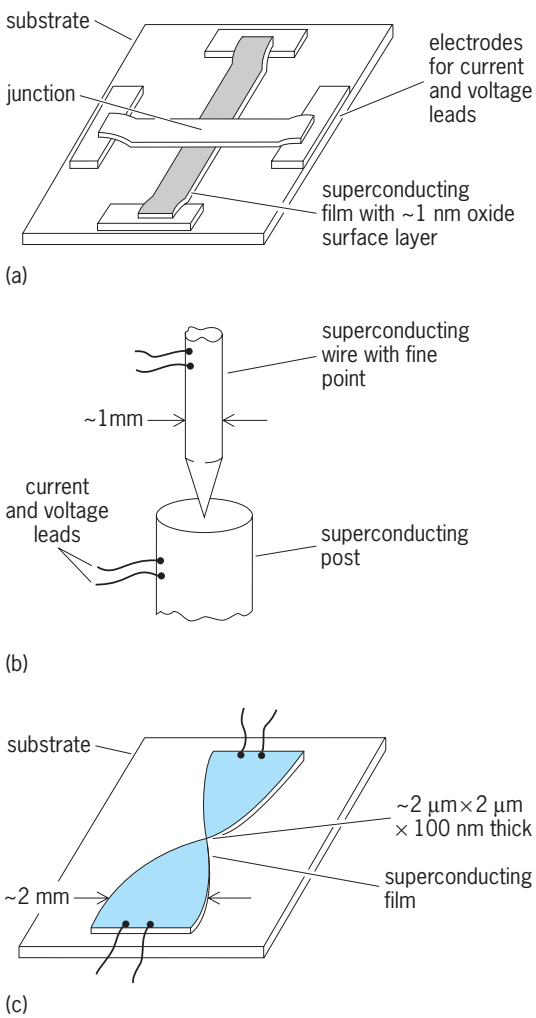

Fig. 3. Some types of Josephson junctions. (*a*) Thin-film tunnel junction. (*b*) Point contact. (*c*) Thin-film weak link. (*After E. Burstein and S. Lundqvist, eds., Tunneling Phenomena in Solids, Plenum, 1969*)

The Josephson effect permits measurement of absolute temperature: a voltage drop across a resistor in parallel with a Josephson junction causes the junction to emit radiation at the frequency given by Eq. (1), but voltage fluctuations resulting from thermal noise produce frequency fluctuations that depend on absolute temperature. *See* LOW-TEMPERATURE THERMOMETRY.

Josephson junctions, and instruments incorporating Josephson junctions, are used in other applications for metrology at dc and microwave frequencies, frequency metrology, magnetometry, detection and amplification of electromagnetic signals, and other superconducting electronics such as high-speed analog-to-digital converters and computers. A Josephson junction, like a vacuum tube or a transistor, is capable of switching signals from one circuit to another; a Josephson tunnel junction is capable of switching states in as little as 6 picoseconds and is the fastest switch known. Josephson junction circuits are capable of storing information. Finally, because a Josephson junction is a superconducting device, its power dissipation is extremely small, so that Josephson junction circuits can be packed together as tightly as fabrication techniques will permit. All the basic circuit elements required for a Josephson junction computer have been developed. *See* SUPERCONDUCTING DEVICES; SUPERCONDUCTIVITY.

Louis B. Holdeman

Bibliography. A. Barone and G. Paterno, *Physics and Applications of the Josephson Effect*, 1982; A. De Chiara and M. Russo, *Tunneling Phenomena in High and Low T_c Superconductors*, 1993.

Joule's law

A quantitative relationship between the quantity of heat produced in a conductor and an electric current flowing through it. As experimentally determined and announced by J. P. Joule, the law states that when a current of voltaic electricity is propagated along a metallic conductor, the heat evolved in a given time is proportional to the resistance of the conductor multiplied by the square of the electric intensity. Today the law would be stated as $H = RI^2$, where H is rate of evolution of heat in watts, the unit of heat being the joule; R is resistance in ohms; and I is current in amperes. This statement is more general than the one sometimes given that specifies that R be independent of I. Also, it is now known that the application of the law is not limited to metallic conductors.

Although Joule's discovery of the law was based on experimental work, it can be deduced rather easily for the special case of steady conditions of current and temperature. As a current flows through a conductor, one would expect the observed heat output to be accompanied by a loss in potential energy of the moving charges that constitute the current. This loss would result in a descending potential gradient along the conductor in the direction of the current flow, as usually defined. If E is the total potential drop, this loss, by definition, is equal to E in joules for every coulomb of charge that traverses the conductor. The loss conceivably might appear as heat, as a change in the internal energy of the conductor, as work done on the environment, or as some combination of these. The second is ruled out, however, because the temperature is constant and no physical or chemical change in a conductor as a result of current flow has ever been detected. The third is ruled out by hypothesis, leaving only the generation of heat. Therefore, $H = EI$ in joules per second, or watts. By definition, $R = E/I$, a ratio which has positive varying values. Elimination of E between these two equations gives the equation below, which is Joule's law as stated

$$H = RI^2$$

above. If I changes to a new steady value I', R to R', and H and $H' = R'I'^2$ as before. The simplest case occurs where R is independent of I. If the current is varying, the resulting variations in temperature and internal energy undoubtedly exist and, strictly speaking, should be allowed for in the theory. Yet, in all but the most exceptional cases, any correction would be negligible.

This phenomenon is irreversible in the sense that a reversal of the current will not reverse the outflow of heat, a feature of paramount importance in many problems in physics and engineering. Thus the heat evolved by an alternating current is found by taking the time average of both sides of the equation. Incidentally, the changes in internal energy, if they were included in the theory, would average out. Hence the equation continues to have a similar form, $\overline{H} = \overline{RI^2}$, for ac applications. *See* ELECTRIC HEATING; OHM'S LAW.

Lewellyn G. Hoxton; John W. Stewart

Juglandales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Hamamelidae of the class Magnoliopsida (dicotyledons). The order consists of two families: the Juglandaceae with a little over 50 species and the Rhoipteleaceae with only one species. Within its subclass the order is sharply set off by its compound leaves. *Juglans* (walnut and butternut) and *Carya* (hickory, including the pecan, *C. illinoensis*) are familiar genera of the Juglandaceae. *See* HAMAMELIDAE; HICKORY; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Juncales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Commelinidae of the class Liliopsida (monocotyledons). The order consists of the family Juncaceae, with about 300 species, and the family Thurniaceae, with only three. Within its subclass the order is marked by its reduced, mostly wind-pollinated flowers and capsular fruits with one to many anatropous ovules per carpel. The flowers have six sepals arranged in two more or less similar whorls, both sets chaffy and usually brown or green. The ovary is tricarpellate, with axile or parietal placentation. The pollen grains are borne in tetrads, and the embryo is surrounded by endosperm. The order is most unusual among higher plants in that, together with at least some members of the Cyperaceae in the related order Cyperales, it has chromosomes with diffuse centromeres. *See* COMMELINIDAE; CYPERALES; LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM.

Arthur Cronquist

Junction detector

A device in which detection of radiation takes place in or near the depletion region of a reverse-biased semiconductor junction. The electrical output pulse is linearly proportional to the energy deposited in the junction depletion layer by the incident ionizing radiation. *See* CRYSTAL COUNTER; IONIZATION CHAMBER.

Introduced into nuclear studies in 1958, the junction detector, or more generally, the nuclear semiconductor detector, revolutionized the field. In the detection of both charged particles and gamma radiation, these devices typically improved experimentally attainable energy resolutions by about two

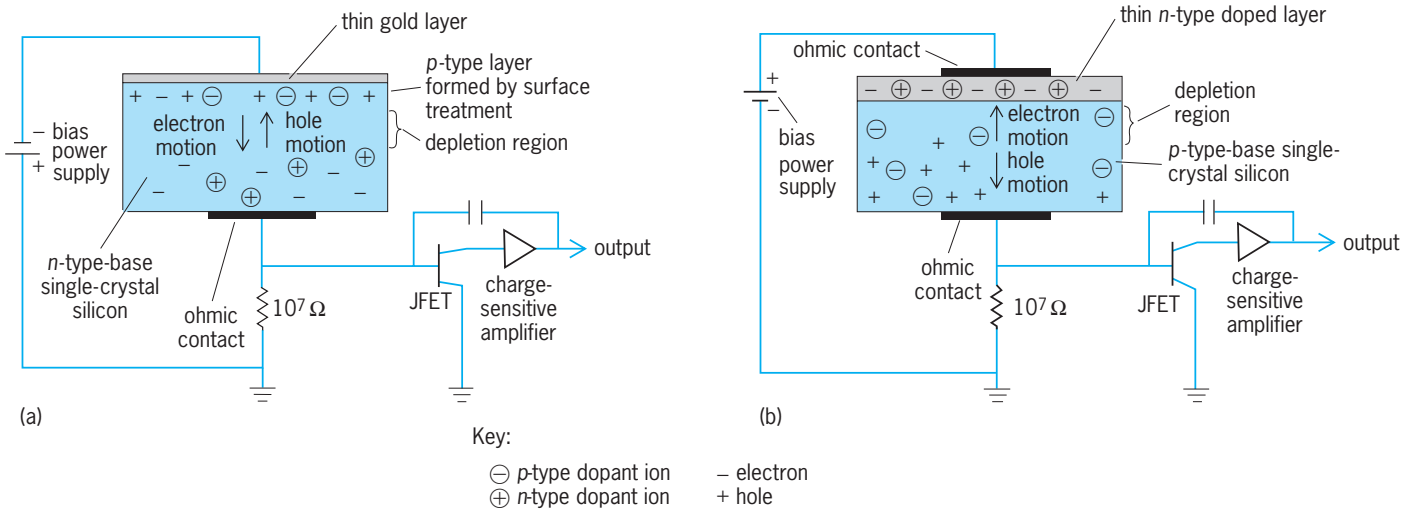

Fig. 1. Silicon junction detectors. (*a*) Surface barrier detector. (*b*) A *pn* junction detector. The *p*-type dopant ions are fixed in the crystal lattice. JFET = junction field-effect transistor.

orders of magnitude over that previously attainable. To this they added unprecedented flexibility of utilization, speed of response, miniaturization, freedom from deleterious effects of extraneous electromagnetic (and often nuclear) radiation fields, low-voltage requirements, and effectively perfect linearity of output response. They are now used for a wide variety of diverse applications, from logging drill holes for uranium to examining the Shroud of Turin. They are used for general analytical applications, giving both qualitative and quantitative analysis in the microprobe and the scanning transmission electron microscopes. They are used in medicine, biology, environmental studies, and the space program. In the last category they play a very fundamental role, ranging from studies of the radiation fields in the solar system to the composition of extraterrestrial surfaces. *See* ELECTRON MICROSCOPE; SECONDARY ION MASS SPECTROMETRY (SIMS).

Fabrication of diodes. The first practical detectors were prepared by evaporating a very thin gold layer on a polished and etched wafer of *n*-type germanium (Ge). To reduce noise these devices were operated at liquid nitrogen temperature (77 K). Silicon (Si), however, with its larger band gap, 1.107 eV compared to 0.67 eV for germanium, offered the possibility of room-temperature operation. Gold-silicon surface barrier detectors and silicon *pn* junction detectors were soon developed.

Surface barrier detectors are made from wafers of *n*-type silicon semiconductor crystals. The etching and surface treatments create a thin *p* layer, and the gold contacts this layer (**Fig. 1***a*). The *pn* junction silicon detectors are usually made by diffusing phosphorus about 2 micrometers into the surface of a *p*-type silicon base (Fig. 1*b*). Both techniques give a *pn* junction. When this junction is reverse-biased, a depletion region, or a region devoid of carriers (electrons and holes), forms mainly in the higher-resistivity base material. A high field now exists in this region, and any carriers born or generated in it are rapidly swept from the region. The requirement for detection is that the ionizing radiation must lose its energy by creating electron-hole pairs (2.96 eV/pair in germanium and 3.66 eV/pair in silicon) in the depletion region or within a carrier diffusion length of this region. Both carriers have to be collected to give an output pulse proportional to the energy of the incident particle. Electrons and holes have similar mobilities in both silicon and germanium, and although carrier trapping occurs it is not as severe as in the 2–16 compounds.

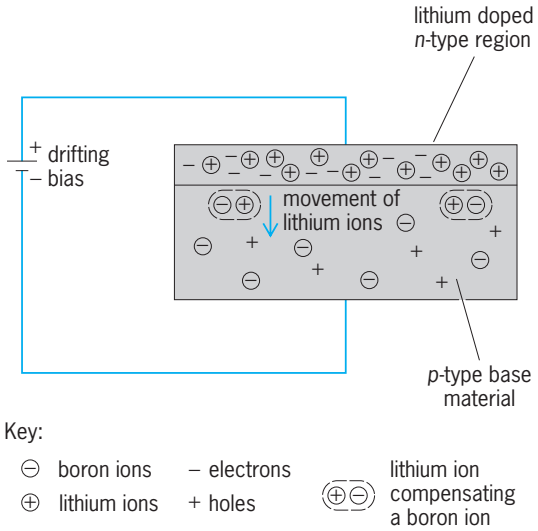

Fig. 2. Compensation of *p*-type semiconductor material with lithium (at 212–392°F or 100–200°C for silicon, 104–140°F or 40–60°C for germanium). Boron ions are fixed in the lattice. Lithium ions are fixed in the lattice, but at elevated temperature can be drifted under an electric field and will compensate boron ions to widen the depletion region.

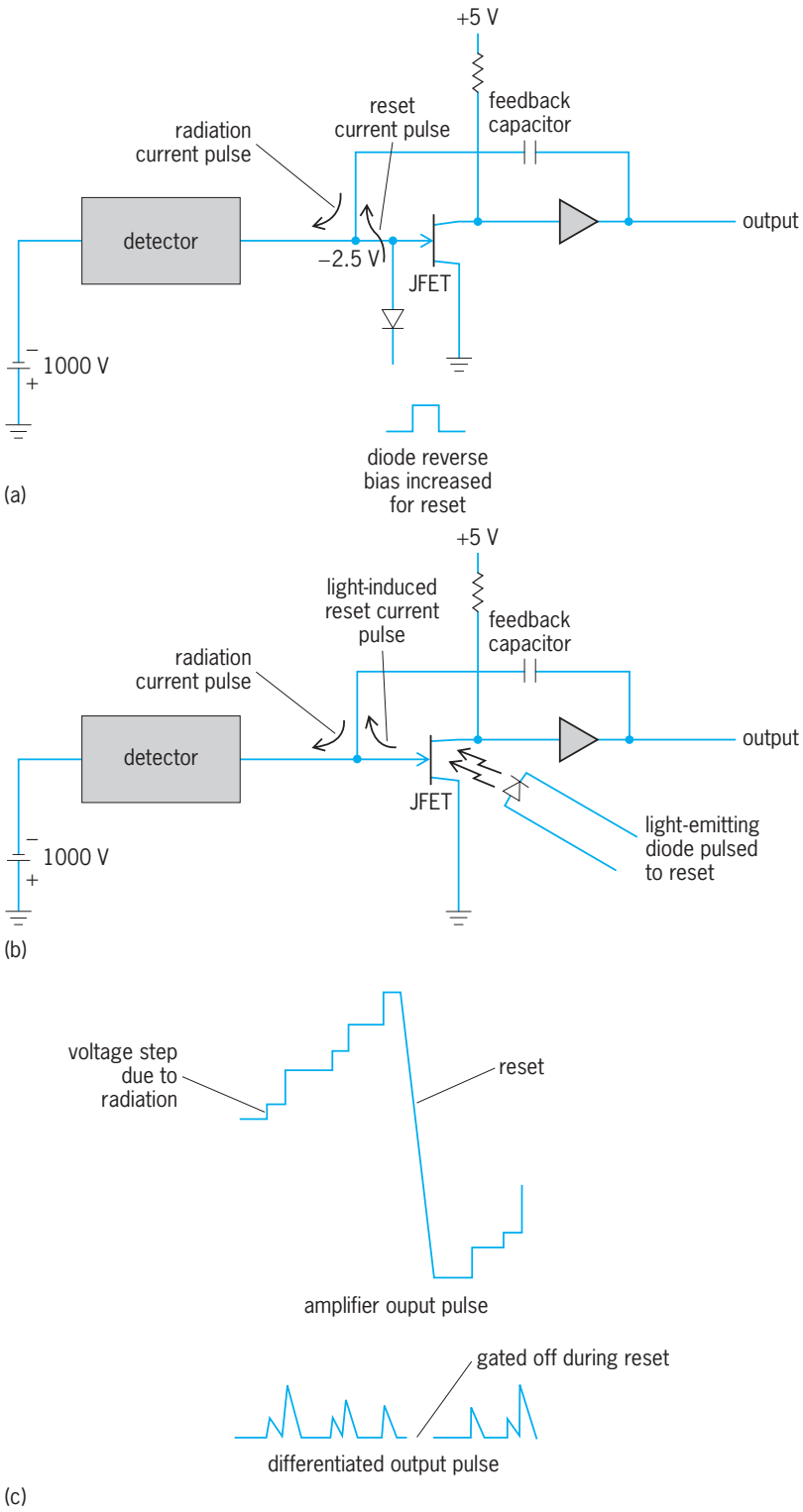

Fig. 3. Reset mechanisms for junction detectors. (*a*) Diode reset. (*b*) Optical reset. (*c*) Amplifier output and differentiated output for the pulse height analyzer.

Control of depletion region width. The detection of charged particles in the presence of gamma rays or higher-energy particles can be optimized by controlling the width of the depletion region. This width is a function of the reverse bias and of the resistivity of the base material. There is a practical limit to the voltage that can be applied to a junction. Thus detectors for higher-energy or lower-mass particles (electrons) requiring wider depletion regions are made from high-resistivity material. This material occurs, by accident, during the growth of some crystals.

Lithium-drifted silicon detectors. Still wider depletion-width detectors can be made from lithium-drifted silicon. Lithium (Li) is a donor in silicon. In addition, at elevated temperatures (392°F or 200°C), the lithium ion is itself mobile. Thus when lithium is diffused in *p*-type silicon, a *pn* junction results. Reverse-biasing this junction at elevated temperatures causes the lithium ion, now appearing as a positive charge, to migrate toward the negative side. On the way it encounters an acceptor ion, negatively charged, which is fixed in the crystal lattice. The lithium ion and the acceptor ion compensate each other, and the lithium ion remains in this location. As more lithium ions drift into and across the depletion region, they compensate the acceptor ions and the region widens (**Fig. 2**). Depletion regions, or compensated regions, up to 0.8 in. (2 cm) wide have been achieved with this technique. Lithium-drifted silicon detectors can be operated at room temperature, but the larger volume gives a greater thermally generated leakage current, which degrades the resolution. The best energy resolution is obtained by operating the detectors at low temperature. However, they may be stored at room temperature.

Lithium-drifted silicon detectors are widely used to detect particle- or photon-induced x-rays. The resolution, when operated at 77 K, is sufficient to resolve the K x-rays for all elements higher in atomic number Z than carbon ($Z = 6$). A resolution of 100 eV has been obtained at 2 keV. At the lower x-ray energies the effects of the detector window thickness and the absorption in the window of the mounting are important, and silicon is preferred for these applications. For x-rays the efficiency of a 5-mm depletion-width lithium-drifted silicon detector is about 50% at 30 keV and 5% at 60 keV. Typically these detectors have capacitances of about 2 picofarads and, to minimize noise, are operated with an optical or diode reset mechanism rather than a feedback resistor (**Fig. 3**). The detector bias is about 1000 V, and the junction field-effect transistor (JFET) gate operates at about −2.5 V. A radiation event causes a pulse of current in the detector. The amplifier drives this current i through the feedback capacitor with capacitance C and in doing so steps a voltage an amount e_{step} proportional to the charge, as given by the equation below. Each subsequent radiation event causes a voltage step. To keep the amplifier within its dynamic range the feedback capacitor must be discharged. The analyzing circuits are first gated off, and in the diode case (Fig. 3*a*) the reverse bias on the diode is momentarily increased to give a picoampere current pulse. The amplifier output

$$e_{\text{step}} = \frac{1}{C} \int i \, dt$$

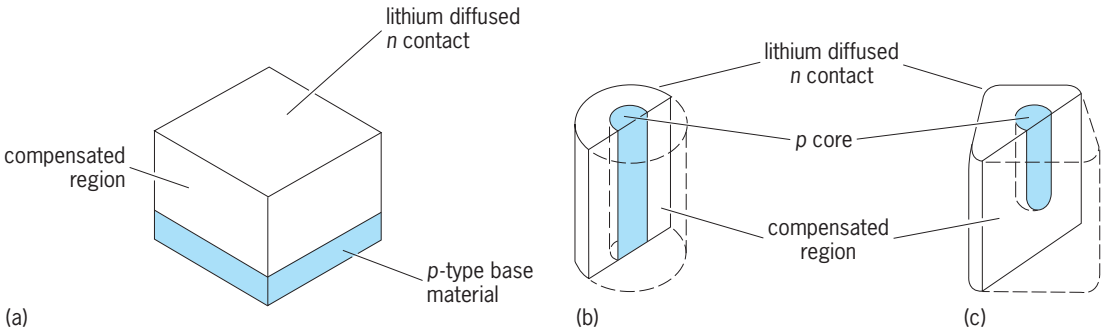

Fig. 4. Lithium-drifted detectors. (*a*) Planar. (*b*) Coaxial. (*c*) Open-one-end coaxial.

voltage changes to allow this current to flow through the feedback capacitor, discharging it. The analyzing circuit is now gated on, and counting can resume. For the optical reset (Fig. 3*b*), a light is flashed on the JFET, momentarily increasing the source-to-gate leakage current and discharging the feedback capacitor. The output from the amplifier and the differentiated output for the analyzer are shown in Fig. 3*c*.

Lithium-drifted germanium detectors. Germanium with its higher atomic number, 32 compared with 14 for silicon, has higher radiation absorption than silicon. Lithium may also be drifted in germanium. But in germanium, lithium is mobile at room temperature and will precipitate or diffuse further if the units, after fabrication, are not kept at liquid nitrogen temperature. Lithium-drifted germanium detectors revolutionized the field of gamma-ray spectroscopy. They may be manufactured in planar, coaxial, or open-one-end coaxial geometry (**Fig. 4**).

Figure 5 compares the gamma-ray spectrum of ^{188}Os taken with a 21-cm^3 lithium-germanium detector with that from a sodium iodide (NaI) scintillator-type spectrometer which is 3 in. (7.5 cm) in diameter by 3 in. deep (27 in.^{3} or 330 cm^3). The counting efficiency of the lithium-germanium detector is lower than the scintillator, but the resolution is at least an order of magnitude better. This higher resolution often reduces the actual counting time to adequately identify a particular energy peak even with an order-of-magnitude less sensitive volume. Also, as shown in Fig. 5, the lithium-germanium detector is able to resolve more energy groups than the scintillator.

Hyperpure germanium detectors. Intrinsic or hyperpure germanium (**Fig. 6**) was grown to overcome the low-temperature-storage and the lithium-drifting problems associated with lithium-germanium. Planar detectors with up to an 0.8-in.-thick (2-cm) depletion region and coaxial detectors with 3-in.^{3} (50-cm^3)

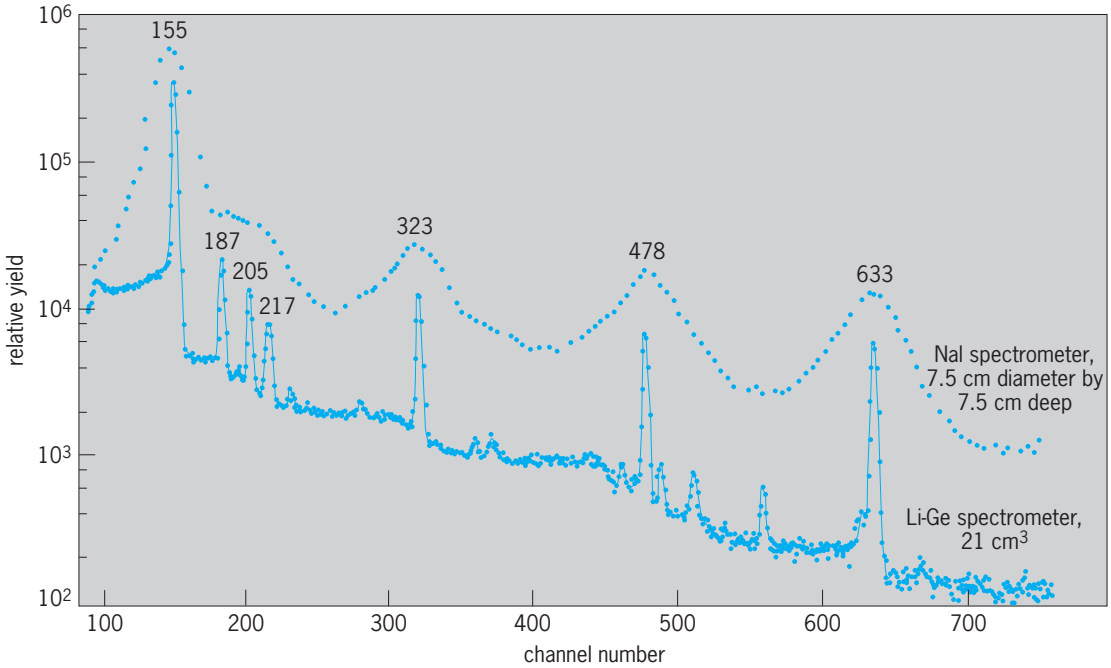

Fig. 5. Gamma-radiation spectra from ^{188}Os as detected in sodium iodide (NaI) and lithium-germanium (Li-Ge) spectrometers. 1 cm = 0.4 in.; 1 cm^3 = 6 × 10^{-2} in.^{3}

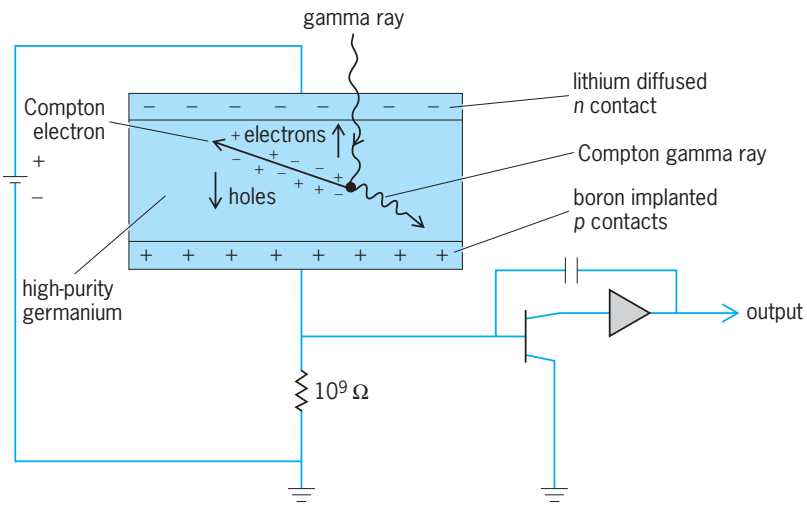

Fig. 6. High-purity germanium gamma-ray detector.

volume have been made with the material. Low-temperature processing is used in the fabrication—usually lithium diffused at 536°F (280°C) for the $n+$ contact and implanted boron for the $p+$ contract. This low-temperature processing is desirable to prevent diffusion of copper, with its subsequent charge trapping, into the germanium. Presently hyperpure germanium detectors cannot be made either as large as, nor with as high a resolution as, lithium-germanium detectors. Both types are operated at liquid nitrogen temperature, 77 K. However, the hyperpure germanium detector is easier to manufacture and can be stored at room temperature when not in use. This is a tremendous practical advantage.

Special detector configurations. Among the many other advantages of semiconductor detectors is the ease with which special detector configurations may be fabricated. One of the simple yet very important examples of this is the annular detector (**Fig. 7**), which is characteristically used to detect nuclear reaction products from a bombarded target in a tight cone around the incident beam. By examining the decay radiation in coincidence with such products, studies may be carried out only on residual nuclei which have had their spins very highly aligned in the nuclear reaction; this has been shown to provide an extremely powerful nuclear spectroscopic probe. The annular detector is extensively used in laboratories worldwide.

Composite detector systems are very readily assembled with the semiconductor devices. For example, it is standard in charged-particle detection to use a very thin detector and a very thick detector (or even two thin and one thick) in series. Multiplication of the resultant signals readily provides a characteristic identification signature for each nuclear particle species in addition to its energy. Three-crystal gamma-ray spectrometers are readily assembled, wherein only the output of the central detector is examined whenever it occurs in time coincidence with two correlated annihilation quanta escaping from the central detector. These systems essentially eliminate background from Compton scattering of other more complex electro-magnetic interactions and yield sharp single peaks for each incident photon energy (**Fig. 8**).

Similarly neutrons may be indirectly detected through examination of recoil protons from a hydrogenous radiator in the case of high-energy neutrons, or through examination of fission fragments resulting from slow neutrons incident on a fissile converter foil mounted with the semiconductor detectors. (It should be noted that the response of the detectors is essentially perfectly linear all the way from electrons and photons to fission fragments.) Neutrons also may be detected and their energy spectra studied through examination of the charged products of the $(n\alpha)$ reaction (where alpha particles are emitted from incident neutrons) induced in the silicon or germanium base material of the detector itself.

Fabrication of triodes. Whereas the detectors thus far discussed are electrically nothing more than diodes, it has been possible to construct equivalent triodes which have extremely important uses in that they provide not only an output which is linearly proportional to the energy deposited in them, but also a second output which in combination with the first establishes the precise location on the detector itself where the ionizing radiation was incident. This has very obvious advantages in the construction of simple systems for the measurement of angular distributions, where such position-sensitive detectors are located about a bombarded target. Their most important impact, however, has been in terms of their on-line use in the focal planes of large nuclear magnetic spectrographs. Simultaneous determination of the energy and location of a particle in the focal plane, together with the momentum

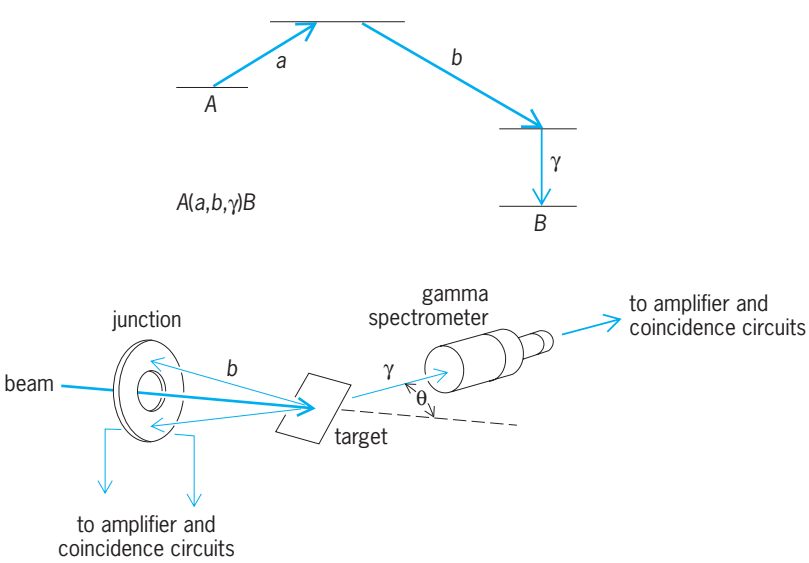

Fig. 7. Schematic of use of annular detector in nuclear reaction studies.

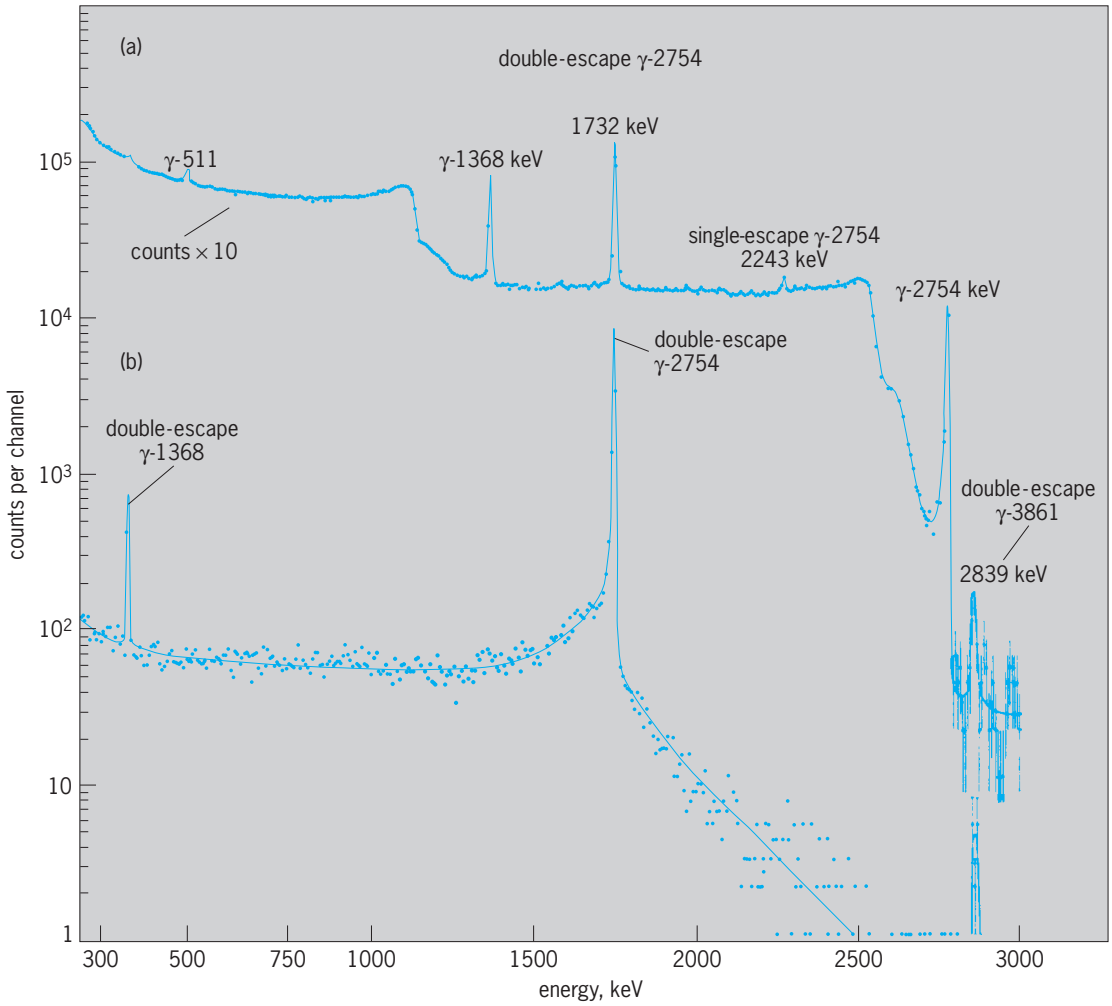

Fig. 8. Comparison of (*a*) direct single detector and (*b*) three-crystal spectrometer spectra from ^{24}Na source.

determination by the magnet itself, establishes unambiguously both the mass and energy of the particle, and does so instantaneously so that additional logical constraints may be imposed through a connected on-line computer—something totally impossible with the earlier photographic plate focal-plane detectors (**Fig. 9**).

A further important utilization of the nuclear triodes has followed their fabrication in an annular geometry similar to that shown in Fig. 7. With radial position sensitivity it becomes possible to correct online, and event by event, for the kinematic variation of particle energy with angle over the aperture of the detector. Without this correction possibility all

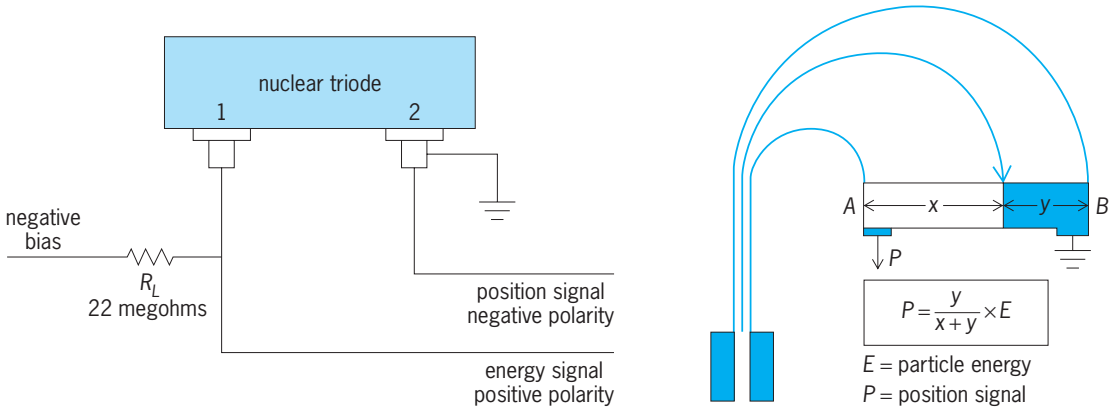

Fig. 9. Schematic of position-sensitive (nuclear triode) detector in focal plane of 180° magnetic spectrograph.

particle group structures in the detector spectrum are smeared beyond recognition. *See* PARTICLE DETECTOR; SEMICONDUCTOR. James M. McKenzie

Bibliography. D. A. Bromley, Detectors in nuclear science, *Nucl. Instrum. Meth.*, 162:1–8, 1979; G. T. Ewan, The solid ionization chamber, *Nucl. Instrum. Meth.*, 162:75–92, 1979; E. Laegsgaard, Position-sensitive semiconductor detectors, *Nucl. Instrum. Meth.*, 162:93–111, 1979; J. M. McKenzie, Development of the semiconductor radiation detector, *Nucl. Instrum. Meth.*, 162:49–73, 1979.

Junction diode

A semiconductor rectifying device in which the barrier between two regions of opposite conductivity type produces the rectification (**Fig. 1**). Junction diodes are used in computers, radio and television, brushless generators, battery chargers, and electrochemical processes requiring high direct current and low voltage. Lower-power units are usually called semiconductor diodes, and the higher-power units are usually called semiconductor rectifiers. For a discussion of conductivity types, carriers, and impurities *see* SEMICONDUCTOR.

Junction diodes are classified by the method of preparation of the junction, the semiconductor material, and the general category of use of the finished device. By far the great majority of modern junction diodes use silicon as the basic semiconductor material. Germanium material was used in the first decade of semiconductor diode technology, but has given way to the all-pervasive silicon technology, which allows wider temperature limits of operation and produces stable characteristics more easily. Other materials are the group 13–15 compounds, the most common being gallium arsenide, which is used where its relatively large band-gap energy is needed. A partial list of silicon types includes the diffused silicon switching diode, alloyed silicon voltage reference diode, epitaxial planar silicon photodiode, and diffused silicon rectifier. Other types include the ion-implanted varactor diode and the gallium arsenide light-emitting diode.

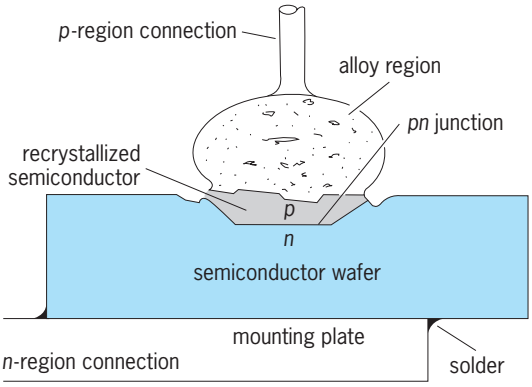

Fig. 1. Section of a bonded or fused junction diode.

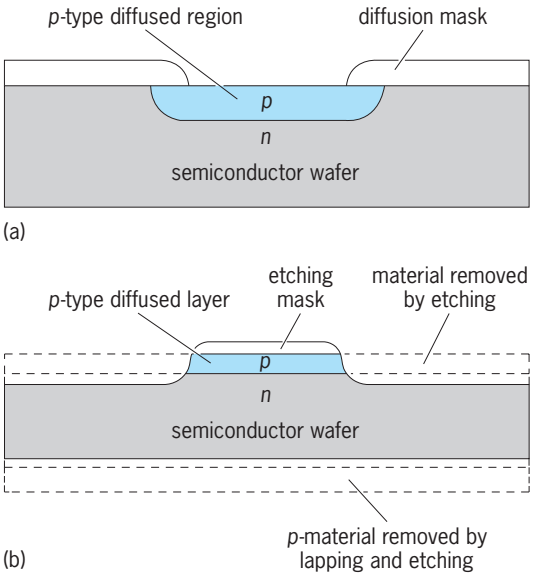

Fig. 2. High-speed diffused silicon diodes. (*a*) Mesaless structure. (*b*) Mesa structure.

In silicon units nearly all categories of diodes are made by self-masked diffusion, as shown in **Fig. 2***a*. Exceptions are diodes where special control of the doping profile is necessary. In such cases, a variety of doping techniques may be used, including ion implantation, alloying with variable recrystallization rate, silicon transmutation by neutron absorption, and variable-impurity epitaxial growth. The mesa structure shown in Fig. 2*b* is used for some varactor and switching diodes if close control of capacitance and voltage breakdown is required. *See* ELECTRONIC SWITCH; RECTIFIER; SEMICONDUCTOR DIODE.

Fabrication methods. The alloy and mesa techniques are largely historical, but were important in the development of junction diodes. The alloy junction section (Fig. 1) is produced by placing a pill of doped alloying material on the clean flat surface of a properly oriented semiconductor wafer and heating it until the molten alloy dissolves a portion of the semiconductor immediately beneath it. Upon cooling, the dissolved semiconductor, now containing the doping impurity, recrystallizes upon the surface of the undissolved semiconductor, reproducing its crystal structure and creating a *pn* junction at the position marking the limit of the solution of the original wafer. If such a junction is held at the peak temperature of its alloying cycle for sufficient time to allow diffusion of the alloy impurity beyond the limit of the dissolved semiconductor into the solid semiconductor, the junction produced is called alloy-diffused.

The planar diffused junction section (Fig. 2*a*) is produced in silicon by first polishing the top surface of a large silicon wafer and then oxidizing the surface by heating the wafer at about 1000°C (1800°F) in the presence of wet oxygen. After about 0.5 micrometer of oxide has grown on the surface, the wafer is cooled, and an array of holes is opened in the

oxide by high-precision etching geometrically controlled by a photoresist technique. A very heavily doped thin oxide layer is chemically deposited in the holes opened in the oxide. This predeposition step is followed by a drive-in diffusion at a higher temperature, causing the deposited impurity to penetrate the substrate, thereby forming diffused *pn* junctions beneath each hole. Subsequently the individual junctions are separated out of the large wafer by scribing and breaking and are encapsulated as individual diodes. Such planar diffused diodes have relatively high breakdown voltages and low leakage currents. The ends of the junction are automatically protected by the oxide mask so that such diodes show long-term stability. This protection by the oxide is often referred to as passivation.

Planar diodes and planar transistors are used in integrated circuits. The diodes in integrated circuits usually consist of the emitter junction or collector junction of a transistor structure rather than being fabricated as a separate diode. Most discrete diodes are power rectifiers, voltage regulators, varactors, or light-emitting diodes. *See* INTEGRATED CIRCUITS; VARACTOR; VOLTAGE REGULATOR.

The mesa structure (Fig. 2*b*) is produced by diffusing the entire surface of the large wafer and then delineating the individual diode areas by a photoresist-controlled etch that removes the entire diffused area except the island or mesa at each diode site.

Still another method of doping control used in modern diodes is through epitaxially deposited material. In this process the polished wafer is subjected at an elevated temperature to a vapor containing a compound of the semiconductor together with a compound containing the appropriate doping element. These compounds decompose upon contact with the surface of the wafer and cause the semiconductor to grow a layer of doped material on its surface. Under proper conditions of cleanliness and growth rate, the underlying crystal structure is propagated into the growing layer, which is then said to be epitaxial in character. In this way either localized areas or entire surfaces of either conductivity type may be produced. In diode fabrication it is typical to use the epitaxially grown material as a lightly doped layer over the original substrate material of the same conductivity type. The junction is then formed in the lightly doped layer by masked diffusion of the opposite-conductivity-type material. By this means the thickness of the web of lightly doped material immediately beneath the diffusion can be controlled to give both a desired reverse breakdown voltage and a relatively constant capacitance. Forward-bias recovery time can be controlled in a trade-off with reverse breakdown voltage in such a structure.

A method of doping control used when special doping concentration profiles are needed, or when localized doping must be accomplished without self-masking oxide, is ion implantation. At present the largest use of this technique in *pn* junction fabrication is to replace the chemical predeposition step in the planar diffusion process. Here ion implantation gives a much more precise control of the sheet resistivity of the diffusion, and it can be accomplished without opening holes in the protective oxide. Crystal damage is automatically healed during the subsequent drive-in diffusion. *See* ION IMPLANTATION.

Junction rectification. Rectification occurs in a semiconductor wherever there is a relatively abrupt change of conductivity type. In any semiconductor the product of the concentrations of the majority and minority current carriers is a temperature-dependent equilibrium constant. The conductivity is proportional to the majority carrier concentration and inversely proportional to the minority-carrier concentration. When a *pn* junction is reverse-biased (*p*-region negative with respect to the *n*-region), the majority carriers are blocked completely by the barrier, and only the minority carriers can flow under the barrier. This minority carrier current is the sum of the individual currents from the *n*- and *p*-regions, and each component is inversely proportional to the conductivity of its region. In addition, there is a thermal regeneration current of minority carriers generated in the depletion region of the reverse-biased junction. In silicon the regeneration current dominates and is about 10^{-3} A/m^2 at room temperature.

When a *pn* junction is forward-biased (*p*-region positive with respect to the *n*-region), the majority hole and electron distributions can flow into the opposite region because the bias has markedly lowered the barrier. Since electrons flowing into a *p*-region or holes flowing into an *n*-region represent a great increase in minority-carrier concentration, the thermodynamic equilibrium of the holes and electrons is disturbed, and the product of their concentrations increases as the junction is approached. The resistivity of both the *n*- and *p*-type regions is considerably lowered by these excess minority carriers, and the forward current is greater than the current through a geometrically equivalent bar of material containing no *pn* junction.

The electrons in an *n*-type semiconductor are given up to the conduction process by donor impurity atoms which remain as fixed, positively charged centers. Similarly, the holes of a *p*-region are created by the capture of electrons by acceptor impurity atoms which remain as fixed, negatively charged centers. In both cases the space charge of the ionized impurity centers is neutralized by the space charge of the majority carriers.

At a *pn* junction the barrier that keeps majority carriers away consists of a dipole layer of charged impurity centers, positive on the *n*-type side and negative on the *p*-type side. When a reverse bias is applied, the barrier height increases and requires more charge in the dipole layer to produce the required step in voltage. To add to the charge, the layer must widen, because ionized impurities are in fixed positions in the crystal. As the layer widens, the capacitance of the junction decreases since the plates of the capacitor are farther apart. Therefore, a *pn* junction acts as a variable capacitance as well as a variable resistance. In this application, it is called a varicap or a varactor diode.

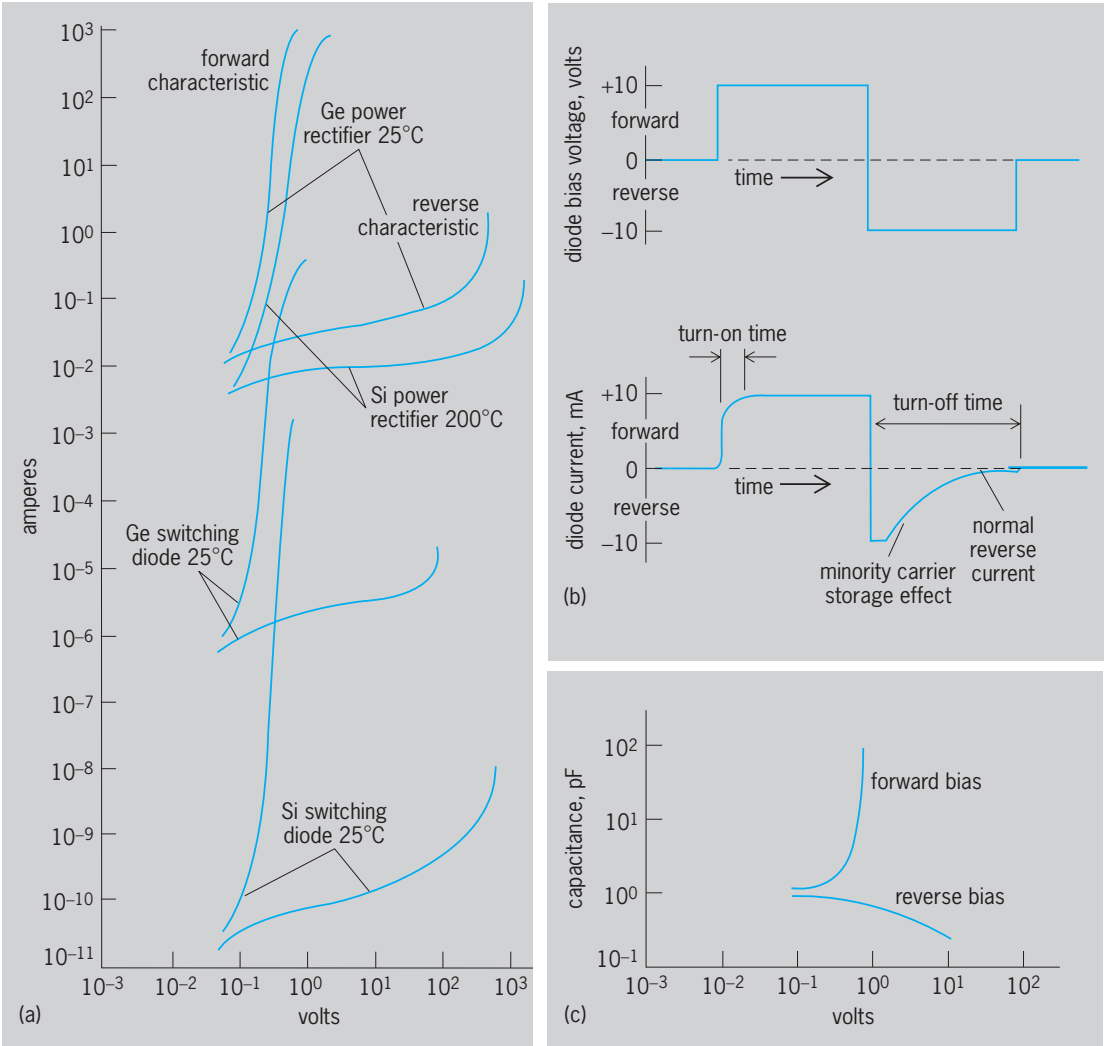

Fig. 3. Junction diode characteristics. (*a*) Rectification. (*b*) Switching. (*c*) Silicon switching diode capacitance. °F = (°C × 1.8) + 32.

Optical properties. When light of sufficient energy is absorbed by a semiconductor, excess minority carriers are created. In a *pn* junction device these excess carriers will increase the reverse-bias leakage current by a large factor if they are within diffusion distance of the junction. If the junction is open-circuited, a forward voltage will develop to oppose the diffusion of the excess carriers generated by the light absorption. This photovoltaic response is the basis of the operation of solar cells and most photographic exposure meters. *See* EXPOSURE METER; PHOTOVOLTAIC CELL; PHOTOVOLTAIC EFFECT; SOLAR CELL.

The inverse of the above effect also exists. When a *pn* junction in a so-called direct-gap semiconductor is forward-biased, the electrically injected excess minority carriers recombine to generate light. This is the basis of light-emitting diodes and injection lasers. Typical direct-gap semiconductors (suitable for light-emitting diodes) are compounds between 13 and 15 group elements of the periodic table such as gallium arsenide. *See* LASER.

Characteristics. Typical rectification, switching, and capacitance characteristics of a junction diode are shown in **Fig. 3**. Rectification characteristics (Fig. 3*a*) show that silicon units provide much lower reverse leakage currents and higher voltage breakdowns and can operate up to 200°C (392°F).

For switching purposes, turn-on and turn-off times are most important (Fig. 3*b*). The turn-on time of a diode is governed by its junction capacitance and is usually short. The turn-off time, usually the critical characteristic, is governed by the time required to remove all of the excess minority carriers injected into the *n*- and *p*-regions while the diode was in the forward-bias state. This is called the minority carrier storage effect, and it is of the order of a few microseconds for good switching diodes. Silicon diodes are usually somewhat superior to germanium units in this respect. The limits of operation of present junction diodes are about 2500 V reverse-standoff voltage and 1500 A forward current in power rectifiers; about 1.0 nanosecond reverse recovery time and 100 picoseconds rise time for fast-switching diodes; a minimum reverse leakage current in a small signal diode is about 0.01 nA.

For further discussion of the properties of *pn* junctions *see* JUNCTION TRANSISTOR; TRANSISTOR.

Lloyd P. Hunter

Bibliography. J. Singh, *Semiconductor Devices: An Introduction*, 1994; S. M. Sze, *Semiconductor Devices*, 1985; E. S. Yang, *Microelectronic Devices*, 1988; M. Zambuto, *Semiconductor Devices*, 1989.

Junction transistor

A transistor in which emitter and collector barriers are formed by *pn* junctions between semiconductor regions of opposite conductivity type. These junctions are separated by a distance considerably less than a minority-carrier diffusion length, so that minority carriers injected at the emitter junction will not recombine before reaching the collector barrier and therefore be effective in modulating the collector-barrier impedance. Junction transistors are widely used both as discrete devices and in integrated circuits. The discrete devices are found in the high-power and high-frequency applications. Junction transistors range in power rating from a few milliwatts to about 300 W, in characteristic frequency from 0.5 to 2000 MHz, and in gain from 10 to 50 dB. Silicon is the most widely used semiconductor material, although germanium is still used for some applications. Junction transistors are applicable to any electronic amplification, detection, or switching problem not requiring operation above 200°C (392°F), 700 V, or 2000 MHz. Not all these limits can be achieved in one device, however. Junction transistors are classified by the number and order of their regions of different conductivity type, by the method of fabricating and structure, and sometimes by the principle of operation. Most modern transistors are fabricated by the silicon self-masked planar double-diffusion technique. The alloy technique and the grown-junction technique are primarily of historical importance. For a general description and definition of terms used here and a description of the mechanism of operation *see* TRANSISTOR.

Alloy-junction transistors. Also called fused-junction transistors, these are made in the *pnp* and *npn* forms. The emitter and collector regions are formed by recrystallization of semiconductor material from a solution of semiconductor material dissolved in some suitably doped metal alloy. The major metal of the alloy serves as the solvent for the semiconductor, while the minor element serves as a source for doping impurity in order to render the recrystallized material opposite in conductivity type to the original wafer.

Alloy junctions are abrupt and allow for bidirectional operation. They usually show a low series resistance, and were therefore used in high-power transistors.

Figure 1 compares several transistor profiles which show how the impurity content varies through the structure. In these profiles C_p is the concentration of the *p*-type impurity; C_n is the concentration of the *n*-type impurity. The net impurity content determines the conductivity type and magnitude. The profile of the alloy transistor shows that there are abrupt changes of impurity concentration at emitter and collector junctions and that the conductivities of emitter and collector regions are therefore high compared to those of the base region. Such a structure shows good emitter-injection efficiency but only moderate collector-voltage rating and relatively high collector capacitance. *See* SEMICONDUCTOR.

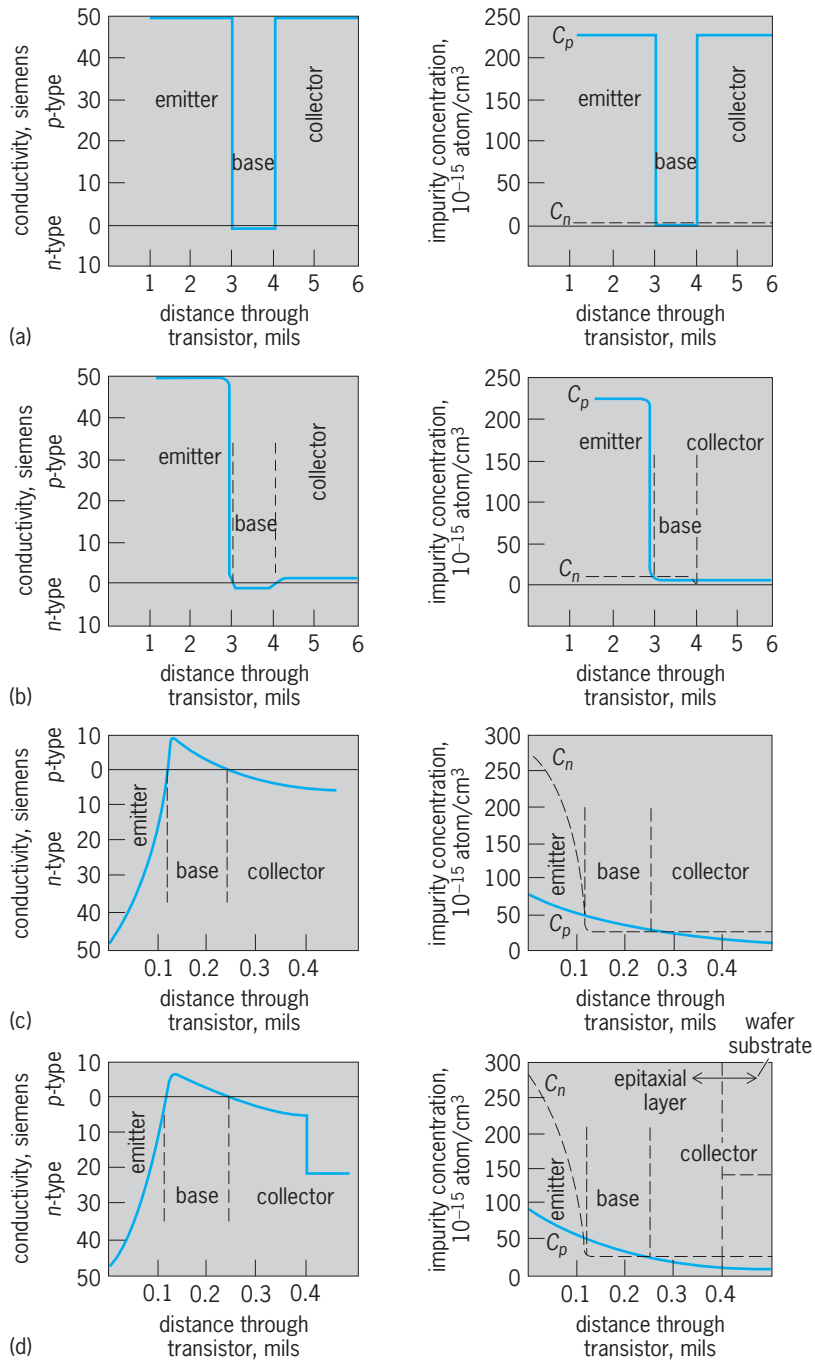

Fig. 1. Conductivity and impurity profiles of typical junction transistors. (*a*) *pnp* alloy-junction type. (*b*) *pnp* grown-junction type. (*c*) *npn* double-diffused-junction type. (*d*) *npn* epitaxial double-diffused-junction type. 1 mil = 25.4 μm.

Grown-junction transistors. These are made in the *pnp* and *npn* forms, as well as in more complicated forms. There are several variations of the grown-junction technique. The simplest consists of successively adding different types of impurities to the melt from which the semiconductor crystal is being grown.

A semiconductor crystal is usually grown by dipping the end of a seed crystal into molten semiconductor and by arranging the thermal gradients so that new semiconductor solidifies on the end of the seed as it is slowly withdrawn. The solid-liquid interface is roughly a plane perpendicular to the axis of withdrawal. A *pnp* structure can be grown by starting with a *p*-type melt; by adding, at one point in the crystal growth, enough *n*-type impurity to give a slight excess over the *p*-type impurity originally present; and, after growth has continued for a few micrometers, by adding an excess of *p*-type impurity. The last-grown region will be the emitter region, and the original *p*-type crystal will be the collector region. The impurity profile of such a structure is shown in Fig. 1*b*.

The high-conductivity emitter region gives a good injection efficiency, and the junction between the base and collector regions is gradual enough so that the unit will show a relatively low collector capacitance and a high breakdown voltage. The one disadvantage of this method is that both the collector and base regions show relatively high series resistances.

Planar diffused epitaxial transistors. The structure of this transistor is shown in section in **Fig. 2**, and the doping profile through the emitter, base, and collector is shown in Fig. 1*d*. In this structure both collector and emitter junctions are formed by diffusion of impurities from the top surface, as shown in Fig. 2. Using silicon, the structure is formed by growing a diffusion mask of native oxide (silicon dioxide) on the previously polished wafer. A hole is opened in the oxide by a photoresist etch technique (Fig. 2*a*) to define the area of the collector buried layer. For a *p*-type substrate a heavy concentration (n^+) of *n*-type impurity such as phosphorus is diffused into the substrate through the opening in the masking oxide. The oxide is etched away, and an epitaxial layer of lightly doped (n^-) silicon is grown over the entire wafer by vapor decomposition at a temperature low enough to prevent significant diffusion of the n^+ material out of the buried layer (Fig. 2*b*). A new oxide layer is grown on the surface of the epitaxial layer, and an opening is etched in it to define the limits of the *p*-type base diffusion (Fig. 2*c*). (This automatically controls the collector junction geometry and capacitance.) The masking oxide is again stripped and regrown for the masking of the n^+ diffusion used to form the emitter and collector contact region (Fig. 2*d*). Next the emitter mask is removed, and an impervious layer of oxide is formed over the surface of the crystal. A layer of glass is bonded to the crystal by means of the oxide layer. The glass must match the expansion coefficient of silicon fairly well, and the oxide must be sufficiently impervious to the glass at the bonding temperature to prevent the diffusion of impurities from the glass into the silicon transistor structure. Finally, holes are etched in the glass-oxide structure so that electrical contact can be made to the various regions of the transistor (Fig. 2*e*).

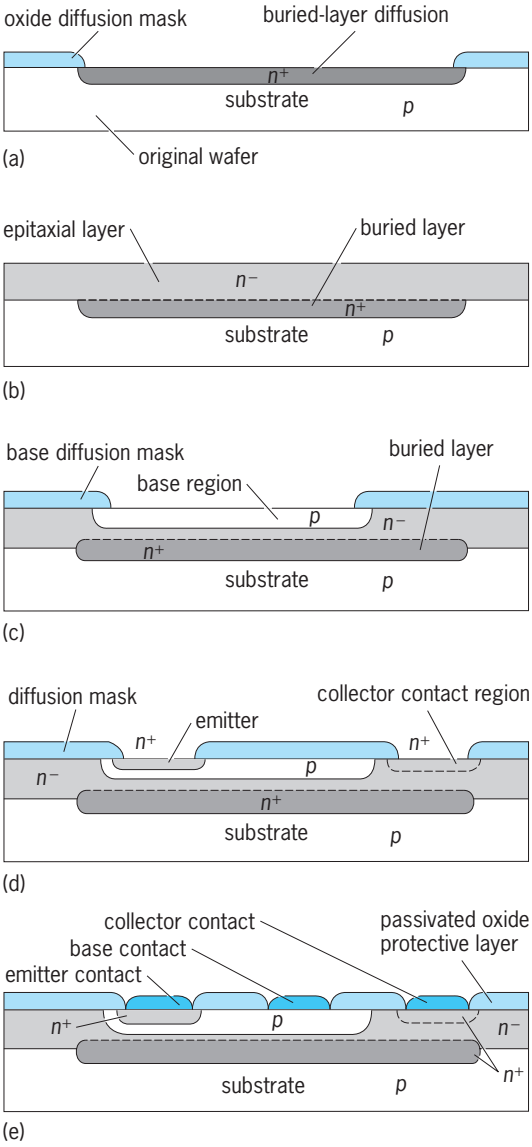

Fig. 2. Double-diffused planar epitaxial transistor structure and method of fabrication. (*a*) Buried layer. (*b*) Epitaxial layer. (*c*) Collector junction formation. (*d*) Emitter junction. (*e*) Contact stripe placement.

In modern technology the above-described base and emitter diffusions are carried out in two steps: a predeposition step, in which a very thin layer of heavily doped oxide is chemically deposited over the open surface of the silicon in the hole opened in the masking oxide; and a drive-in diffusion step, in which the deposited dopant is diffused into the silicon at a higher temperature than that used for the predeposition. This typically controls the sheet resistance of the final diffusions to about ±10% of the design value. The forward current-transfer ratio of the transistor is determined by the ratio of these sheet resistances through the medium of the injection efficiency of minority carriers.

The chemical predeposition step is being replaced by ion implantation directly through the oxide. The masking is accomplished by placing a layer of photoresist on top of the oxide. This eliminates the oxide etching step and allows an accurately metered deposition by controlling the time of bombardment and the current of the ion beam. This modification of the process promises to keep the emitter and base region sheet resistivities within about ±1% of the design value. *See* ION IMPLANTATION.

In this transistor, formation of the base region by diffusion from the emitter side produces a steep doping gradient and thereby a strong electric field in the base region. In the typical alloy-junction transistor (uniform base doping) the minority-carrier transport across the base is achieved by a relatively slow diffusion process. In this diffused base type (sometimes called a drift transistor) the base region shows a high conductivity gradient, decreasing from the emitter to the collector (Fig. 1*d*). This conductivity gradient means that the majority-carrier concentration is much greater near the emitter than near the collector. In order to cancel the natural diffusion of majority carriers from the high- to the low-concentration region, an electric field must exist of such a polarity as to tend to drive majority carriers back toward the emitter. This polarity of field then tends to drive minority carriers from the emitter to the collector; when normal bias is applied to the device, excess injected minority carriers will be accelerated across the base by this field.

The buried layer of n^+-doped material has very low resistance and acts as a shorting bar between the area immediately beneath the emitter and the area immediately beneath the collector contact stripe, thus maintaining a low collector series resistance even if the n^- material of the epitaxial region is of quite high resistivity. The collector breakdown voltage may be maintained at a reasonably high value and the collector capacitance at a low value by controlling the thickness of the n^- material between the base and the buried layer and by keeping the doping level of the n^- material quite low.

Mesa transistors. These transistors minimize the collector capacitance by limiting the collector junction area. This area limitation is achieved by etching away the surrounding material so that the entire transistor structure stands up from the body of the wafer like a small mesa. This structure gave the highest frequency response for many years. It is now replaced by the planar type of structure.

Power transistors. These are used in the output stage of an electronic circuit both as switches and as amplifiers. Depending on the load, a high voltage rating, a high current rating, or a high power rating may be required. With any of these, heat dissipation within the device is a serious limitation. In order to obtain high current capability in a power transistor, a large emitter junction area is required. The base-region recombination current produces a lateral voltage drop between the center of the emitter area and the center of the base contact area in a planar device. This voltage tends to bias the center of the emitter area to off and concentrate the injection current at the periphery of the emitter. Modern silicon power transitors minimize this effect by using emitter junctions with a large perimeter-to-area ratio, usually in the form of a multipronged fork with base contacts interdigitated between the tines. This preserves high forward current-transfer ratio to large values of emitter current. *See* CONTROLLED RECTIFIER.

Unijunction transistor. This device is really a diode in which the resistance of a portion of the base region is modulated by minority carriers injected by forward-biasing its single junction. Its structure typically consists of a lightly doped base region with ohmic contacts at opposite ends. The single junction is formed over a narrow range near the center of the base region by a shallow diffusion of heavily doped material of the opposite conductivity type. If a bias current is set up from end to end in the base, the potential at the junction can be set at a desired reverse bias relative to ground. If a signal is applied to the junction electrode, the device will turn on when the signal exceeds the original reverse-bias potential of the base at that point. Once forward-biased, the junction injects sufficient minority carriers into the base to short the region beneath the junction to the ground end of the base, and the device remains conducting until reset, either by the base bias or by the emitter signal. These devices show a typical negative resistance characteristic and are used for timing, control, and sensing circuits.

Summary. Silicon planar passivated transistors show a wide range of performance with characteristic frequencies up to 2000 MHz, voltage ratings of 12–700 V, and power dissipation ratings of 100 mW–300 W. The highest-frequency devices range up to 4000 MHz. Silicon planar technology is used in fabricating integrated circuit chips. The general form of the transistor structure displayed in Fig. 2 is used in integrated circuits. Such a structure is used for diodes as well as transistors since, for example, it is necessary only to connect the base and collector contacts to use the collector junction as a diode. *See* INTEGRATED CIRCUITS.

Lloyd P Hunter

Bibliography. J. J. Brophy, *Basic Electronics for Scientists*, 5th ed., 1990; G. W. Neudeck, *The Bipolar Junction Transistor*, 2d ed., 1989; S. M. Sze, *Semiconductor Devices*, 1985; E. S. Yang, *Microelectronic Devices*, 1988.

Jungermanniales

The largest order of liverworts, often called the leafy liverworts; it consists of 43 families. The leaves are in three rows, with the underleaves usually reduced or lacking. Other distinctive features include a perianth formed by a fusion of modified leaves, a short-lived seta, and a four-valved capsule. The leaves have an embryonic bilobed phase which may be lost on further development.

The plants of this order are dorsiventrally organized and leafy. They grow by means of an apical cell with three cutting faces, resulting in two

rows of lateral leaves and a third row of underleaves which are generally reduced, and sometimes lacking. The stems lack a central strand. A medulla of delicate cells is generally surrounded by a cortex. Rhizoids are usually present, all smooth. The leaves pass through a primordial two-lobed stage but may become two- to several-lobed, or unlobed (owing to obliteration of one primordial lobe). A midrib is lacking. The cells often have corner thickenings, and oil bodies are usually present in all green cells. Asexual reproduction by gemmae is common. Antheridia occur in leaf axils, sometimes also in axils of underleaves. The stalk is usually 1-2- or less commonly 4-6-seriate. Archegonia are terminal and have a 5-seriate neck. The sporophyte is usually protected by a perianth (in addition to a calyptra) formed by the fusion of leaves. The seta, usually long, consists of delicate, hyaline cells. The capsule is four-valved, with a wall of 2–10 layers of cells. The elaters have spiral-thickened walls. The haploid chromosome number is 8 or 9. *See* BRYOPHYTA; HEPATICOPSIDA; JUNGERMANNIIDAE.

Howard Crum

Bibliography. M. Fulford, Manual of the leafy Hepaticae of Latin America, *Mem. N.Y. Bot. Gard.*, 11:1–172 (1963), 173–276 (1966), 277–392 (1968), 393–535 (1976); R. M. Schuster, *The Hepaticae and Anthocerotae of North America East of the Hundredth Meridian*, vols. 1–4, 1966–1980.

Jungermanniidae

One of the two subclasses of liverworts (class Hepaticopsida). The plants may be thallose, with little or no tissue differentiation, or they may be organized into erect or prostrate stems with leafy appendages. The leaves, generally one cell in thickness, are mostly arranged in three rows, with the third row of underleaves commonly reduced or even lacking. Oil bodies are usually present in all of the cells. The rhizoids are smooth. The capsules, generally dehiscing by four valves, are usually elevated on a long, delicate, short-lived seta. The spore mother cells are deeply lobed.

The subclass consists of the orders Takakiales, Calobryales and Jungermanniales, which are leafy, and the Metzgeriales, which are mostly thallose. *See* BRYOPHYTA; CALOBRYALES; HEPATICOPSIDA; JUNGERMANNIALES; METZGERIALES; TAKAKIALES.

Howard Crum

Bibliography. H. Inoue, *Illustrations of Japanese Hepaticae*, vol. 1, 1974; R. M. Schuster, Phylogenetic and taxonomic studies on Jungermanniidae, *J. Hattori Bot. Lab.*, 36:321–405, 1972.

Jupiter

The largest planet in the solar system, and the fifth in the order of distance from the Sun. It is visible to the unaided eye except for short periods when in near conjunction with the Sun. Usually it is the second brightest planet in the sky; only Mars at its maximum luminosity and Venus appear brighter.

Planet and its orbit. The main orbital elements are a semimajor axis, or mean distance to the Sun, of 484×10^6 mi (778×10^6 km); an eccentricity of 0.0489, causing the distance to the Sun to vary about 47×10^6 mi (75×10^6 km) between perihelion and aphelion; sidereal revolution period of 11.86 years; mean orbital velocity of 8.1 mi/s (13.1 km/s); and inclination of orbital plane to ecliptic of 1.3°. *See* PLANET.

The apparent equatorial diameter of its disk varies from about 47″ at mean opposition to 32″ at conjunction. The polar flattening due to its rapid rotation is considerable and is easily detected by visual inspection; the ellipticity is $(r_e - r_p)/r_e = 0.065$, where r_e is the equatorial radius and r_p is the polar radius. The equatorial diameter is about 88,850 mi (142,984 km), and the polar diameter is 88,086 mi (133,708 km). The volume is about 1321 (Earth = 1), and the mass is about 317.5 (Earth = 1). The mean density is 1.33 g/cm^3, a low value characteristic of the four giant planets. The mean acceleration of gravity at the visible surface is about 85 ft/s^2 (26 m/s^2); however, the centrifugal force at the equator reduces the effective acceleration of gravity to about 78 ft/s^2 (24 m/s^2).

Phases. As an exterior planet, Jupiter shows only gibbous phases and full phase from Earth. Because of the large size of Jupiter's orbit compared with that of the Earth, the maximum phase angle is only 12° at quadratures, and the phase effect shows up only as a slightly increased darkening of the edge at the terminator. The apparent visual magnitude at mean opposition is −2.4, and the corresponding value of the reflectivity (geometrical albedo) is about 0.3; the physical albedo is 0.34. The high value of the albedo, characteristic of the four giant planets, indicates the presence of a dense, cloud-laden atmosphere. *See* ALBEDO.

Telescopic appearance. Through an optical telescope Jupiter appears as an elliptical disk, strongly darkened near the limb and crossed by a series of bands parallel to the equator (**Fig. 1**). Even fairly small telescopes show a great deal of complex structure in the bands and disclose the rapid rotation of the planet. The period of rotation is about 9 h 55 m, the shortest of any planet. The features observed, however, do not correspond to the solid body of a planet but to clouds in its atmosphere, and the rotation period varies markedly with latitude. The rotation period of any given zone is not exactly constant but suffers continual fluctuations about a mean value. Occasionally, short-lived atmospheric phenomena may depart more strongly from the mean rotation period of the zone in which they appear, and thus drift rapidly with respect to other details in the zone. The rotation axis is inclined only 3° to the perpendicular to the orbital plane, so that seasonal effects are practically negligible.

Red Spot. Apart from the constantly changing details of the belts, some permanent or semipermanent markings have been observed to last for decades or even centuries, with some fluctuations in visibility. The most conspicuous and permanent marking is the

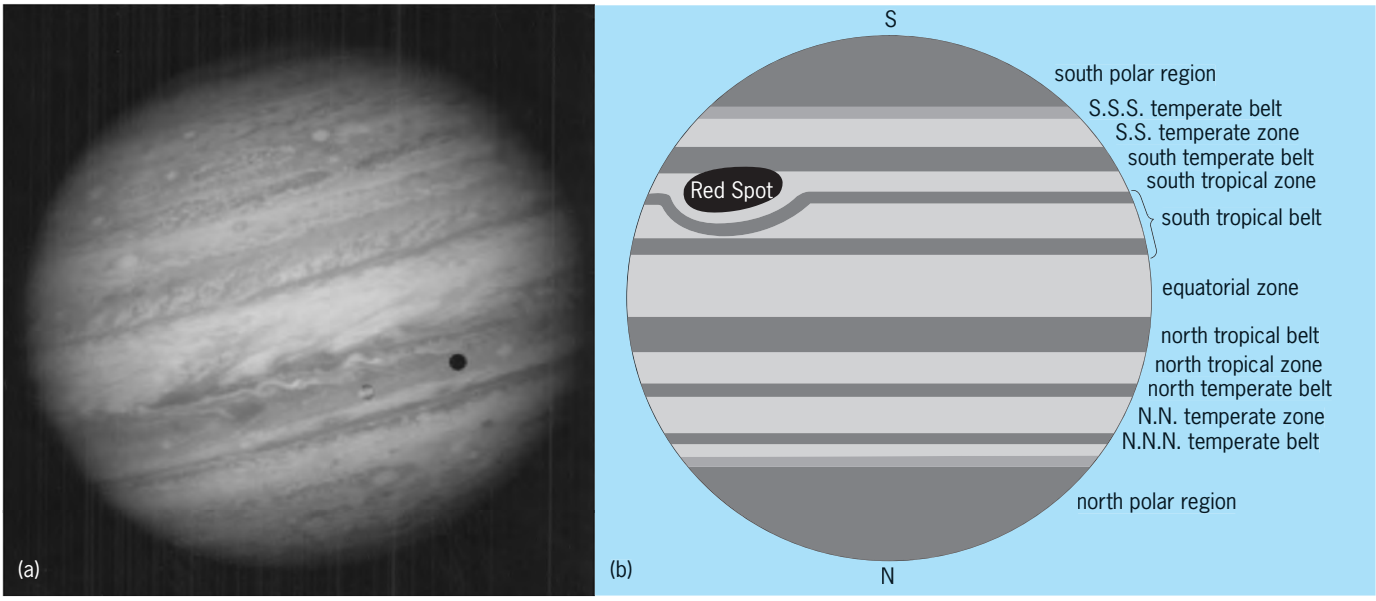

Fig. 1. Jupiter. (*a*) Telescopic appearance from the *Hubble Space Telescope* (*Space Telescope Science Institute: Jet Propulsion Laboratory; NASA*). (*b*) Principal bands.

great Red Spot, discovered by J. D. Cassini in 1665 and observed continually since 1878, when it gained its present red color. At times it has been conspicuous and strongly colored; at other times it has been faint and only slightly colored, and occasionally only its outline or that of the bright "hollow" of the south temperate zone which surrounds it has remained visible.

The current dimensions of the Red Spot are 15,500 mi (25,000 km) in longitude and 7500 mi (12,000 km) in latitude, though the longitudinal dimension varies from 25,000 mi (40,000 km) to 9000 mi (14,000 km). The Spot has been shrinking since 1943 at a rate of 0.19° of longitude per year. It also drifts in longitude by 4.5° per year. It has an essentially calm center, surrounded by a restraining "collar" of high-speed winds which rotate counterclockwise in 6–10 days. This anticyclonic storm (rotation in opposite direction to the planet's) consists of an elevated central region of ammonia gas and ice clouds (**Fig. 2**). Wind speed increases with depth. The Spot is surrounded by towering water-vapor-based thunderstorms, with the deepest clouds in the surrounding collar. The total depth is probably 12–25 mi (20–40 km). These clouds are about 18°F (10°C) colder than the surrounding weather. The Spot's distinctive coloration is probably due to chemical compounds (perhaps containing phosphorus) transported from deep within the atmosphere.

The origin and longevity of the Great Red Spot remain difficult to explain. The vortex is probably not attached to any solid surface feature below the clouds. It is thought to be an eddy of atmospheric gases driven by the strong Coriolis force that stems from the planetary rotation. Atmospheric perturbations or whirlpools coalesce as they pass in the underlying jet stream, forming a vortex of increasing size, which can maintain stability for some time. Several smaller and similarly colored versions of the Spot have been observed, but none has attained the size or lifetime of the original. *Galileo* probe observations suggest that the energy necessary to sustain Red Spot survival comes from the Spot's regular consumption of these eddies several thousand kilometers in diameter, which carry up convective energy from below. This implies a possible lifetime around 100,000 year. *See* CORIOLIS ACCELERATION.

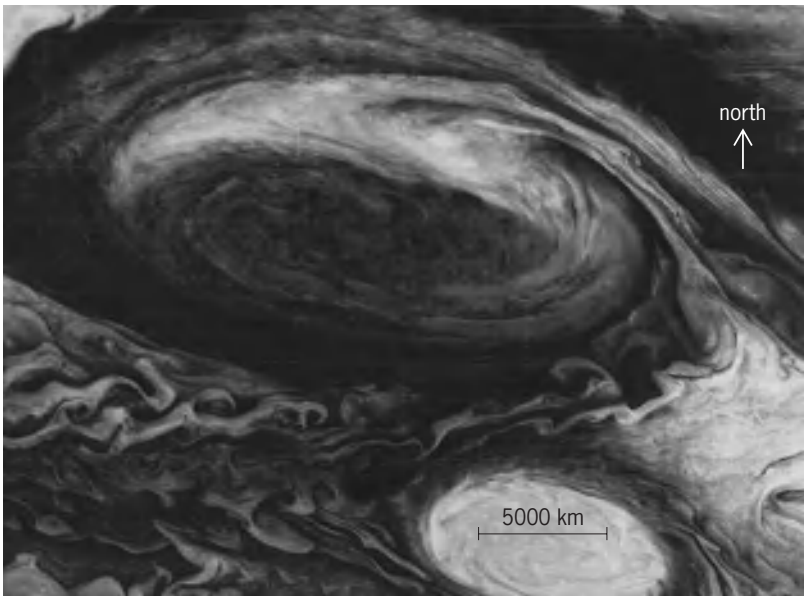

Fig. 2. Details of the Red Spot as seen by the *Voyager 1* fly by. Features include a white oval with a wake of counterrotating vortices, puffy features inside the Red Spot, and reverse-S spirals inside both the Red Spot and the oval. The large white feature extending over the northern part of the Red Spot was observed to revolve about the Red Spot center with a period of 6 days. 5000 km = 3000 mi. (*NASA*)

Atmosphere. Jupiter's visible belts and zones reflect the complicated vertical atmospheric structure. The atmosphere consists primarily of hydrogen (H_2), helium (He), ammonia (NH_3), methane (CH_4), water (H_2O), hydrogen sulfide (H_2S), and other trace compounds such as acetylene ($HC{\equiv}CH$), ethane (CH_3CH_3), hydrogen cyanide (HCN), hydrogen deuteride (HD), carbon monoxide (CO), and phosphine (H_3P). The highest atmospheric features are ammonia crystal haze layers, variable in appearance and density. The visible cloud tops are formed of aerosols of mostly ammonia, with little water vapor, akin to cirrus clouds in Earth's atmosphere. As depth increases, cloud particles become ammonia crystals and ammonium hydrosulfide (NH_4SH) droplets, followed by cumulus clouds of water ice crystals and droplets (**Fig. 3**).

All the compounds previously mentioned are colorless; therefore, the strong yellow, red, blue, and brown colorations seen in the clouds must be due to other agents. Perhaps phosphorus and sulfur compounds create the yellows and reds, while the blue features may be hydrogen made visible by Rayleigh scattering. Infrared observations have found the highest features to be red, followed by white, brown, and blue.

The *Galileo* probe found unexpected results in its plunge into the upper Jovian atmosphere. No distinct water cloud was observed, and less than 10% of the expected abundance of water vapor was seen, although two distinct ammonia and ammonium hydrosulfide cloud levels were seen. It is thought that the probe may have entered an unusually clear and dry area of the atmosphere, a figurative desert. More recent observations give an overall water content similar to solar values. When signal was lost, the probe had penetrated 0.22% of the Jovian atmosphere (about 43 mi or 70 km) where the temperature was 1970°F (1350 K, hotter than expected). This excess heat may have been caused by gravity waves from the lower atmosphere or else heating by soft electron collisions on the hydrogen (H_2) molecules. The ammonia cloud was very tenuous and inhomogeneous near the probe entry point. An ammonia-sulfur cloud was encountered well below its expected height, pointing out the extreme variability of the atmosphere. The probe also found helium raining from upper to lower depths. The *Galileo* probe discovered anomalously high quantities of argon, krypton, and xenon, suggesting that Jupiter had to trap them physically by condensation or freezing. This would require Jovian formation farther out in the solar system than its present location, possibly in the Kuiper Belt. *See* KUIPER BELT.

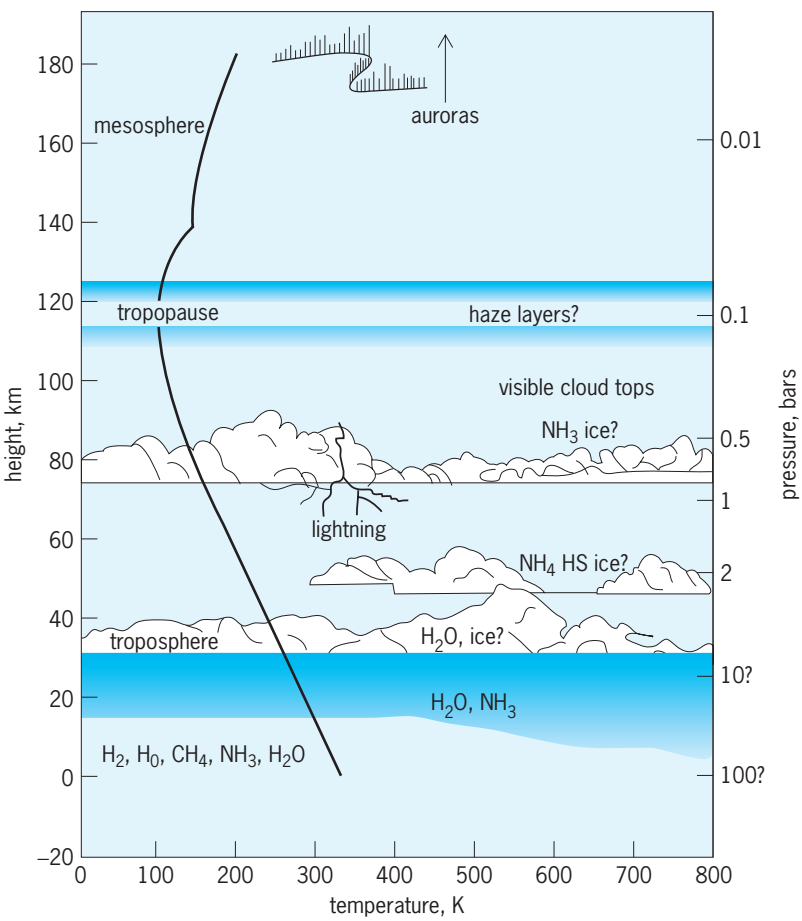

Fig. 3. Diagram showing the locations of the cloud decks of Jupiter. 1 km = 0.6 mi.

The atmospheric structure displays latitudinal variations. Clouds near the equator are lifting and spreading toward higer latitudes, as in the horse latitudes on Earth. Atmospheric aerosols subside at the equator and upwell at high latitudes. The mean temperature of the Jovian disk is −259°F (129 K), but bands of high temperature are found at the poles and mid-latitudes. The equatorial regions alternate between warm and cold over a cycle of 4–5 years.

Zones and belts. The belts and zones represent regions of differing cloud altitudes and compositions. The atmosphere is divided by a series of prograding and retrograding jet streams. Convection is strong since the bright zones are cooler and higher by 9–12 mi (15–20 km) than the dark belts, with their higher albedos arising from solid ammonia crystals. The zones are regions of upward motion and high pressure. *Galileo* found wind speeds of 0.1–0.15 mi/s (160–220 m/s) to extend to unexpected depths, as well as supersonic turbulence in the upper atmosphere on the order of 0.6–6 mi/s (1–10 km/s). This is similar to motions in stellar, not planetary, atmospheres, suggesting particularly stormy upper regions. The energy to maintain such high-velocity winds likely comes from fluid-mechanical convective processes from the interior heat source (discussed below), with some contribution from lightning. The equatorial zonal jet streams have not sensibly changed in over 80 years of observation, despite vigorous turbulence in the defining shear zones.

Belts are aerosol-free and dry; zones are cloudy and ammonia-rich. The motions in the belts are cyclonic, and anticyclonic in the zones. The anticyclones have smooth, nearly elliptical cloud shapes, while the cyclones have large, filamentary, diffuse cloud patterns (**Fig. 4**). The large weather systems

Fig. 4. Fine details of Jovian atmospheric features as photographed by *Voyager 1*. The dark halo surrounding the smaller bright spot to the right of large oval is a hot spot, a warmer region of Jupiter's atmosphere. (*NASA*)

move more slowly than the jets that constrain them.

Individual vortices form within the zones and belts, some persisting for decades. The major white ovals are hot-spot regions of strong infrared emission and are exceptionally dry regions of descending air, while other features such as the plumes observed in the equatorial region may be surface phenomena. The great scale changes in the belts and zones can be gradual (time scales of several months) or cataclysmic and sporadic (several days). For example, in 1990 a new temperate belt was created over a 3-month period as several bright features and white and dark spots coalesced.

Oscillations. Jupiter has global acoustic-mode radial oscillations, discovered in 1991, with periods of 8–17 min. Therefore, the entire planetary atmosphere also has cells of regular vertical motion which may contribute to the complexities of its meteorology.

Comet impact. During July 16–22, 1994, twenty-one observable fragments of Comet Shoemaker-Levy 9 impacted the Jovian atmosphere at nearly 120,000 mi/h (200,000 km/H), the first time such an event had been witnessed. Though occurring on Jupiter's nonvisible side, each strike was clearly visible as a dark area in the south temperate zone as the planet's rapid rotation brought it into view. In addition, ejecta and impact plumes were seen to rise approximately 1860 mi (3000 km) above the planetary limb. The impact of fragment G, thought to be the largest at 2–3 mi (3–4 km) in diameter, released energy equivalent to 6×10^{12} tons of TNT and left a dark spot 1550 mi (2500 km) in diameter surrounded by concentric rings (**Fig. 5**), which expanded at 0.28 mi/s (0.45 km/s). Each strike produced superheated gaseous bubbles, observed at infrared wavelengths, which rose above the stratospheric methane and cooled into aerosols (silicate dust, hyrocarbons, or hydrogen cyanide), as well as increased aurorae near the sites.

The event provided a first probe into the Jovian atmosphere. Traces of ammonia, sulfur (S_2), and hydrogen sulfide in the fireballs showed that they penetrated the upper two cloud layers of ammonia and ammonium hydrosulfide, but the absence of water vapor suggests that the comet itself was low in oxygen-rich ices and volatiles. Surges in Jupiter's decimetric radio emission during the impacts suggested that large numbers of electrons were injected into the Jovian magnetosphere. Given the number of such objects in the solar system, it is thought that Jupiter is impacted once every 500 years by a 0.2-mi-diameter (0.3-km) object and every 6000 years by a 1-mi (1.6-km) comet such as Shoemaker-Levy 9. *See* COMET.

Interior composition and structure. Jupiter primarily consists of liquid and metallic hydrogen. Early measures of the ratios of helium, carbon, and nitrogen to hydrogen gave values resembling those of the Sun, and therefore the primordial composition of the solar system. However, later analyses of the methane spectrum showed a two- to threefold overabundance of carbon as compared to solar values, a result confirmed by gravity analyses of the rocky core. Still in a late phase of its gravitational

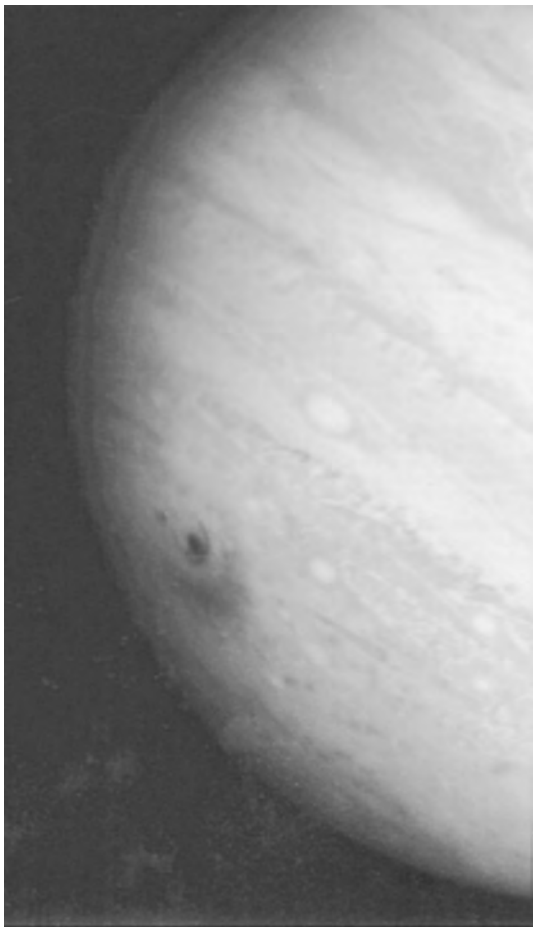

Fig. 5. Image of Jupiter from the *Hubble Space Telescope* showing impact sites of fragments D and G of Comet Shoemaker-Levy 9. The large feature with a ring and crescent-shaped ejecta was created by fragment G. The smaller feature to the left of fragment G impact site was created by fragment D. (*H. Hammel and NASA*)

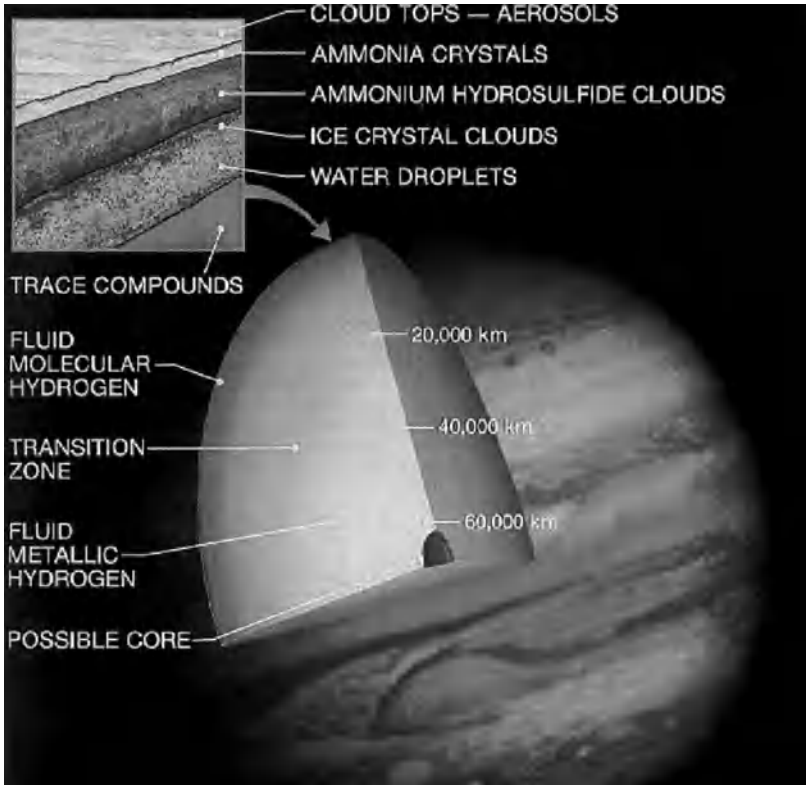

Fig. 6. Schematic diagram of the interior of Jupiter.

contraction, the planet converts the released gravitational energy into heat, emitting 1.668 times as much thermal energy as it receives from the Sun.

Below the clouds is a 13,000-mi-thick (21,000-km) layer of hydrogen and helium, which changes from gas to liquid as the pressure increases. Beneath the liquid hydrogen layer is a 25,000-mi-deep (40,000-km) sea of metallic hydrogen, where the pressure is approximately 3 million times that at Earth's sea level. This layer is electrically conductive and causes Jupiter's intense magnetic field. Acoustic oscillations confirm the presence of a small high-density solid core, perhaps 1.5 times the Earth's diameter, with a density 10–30 times that of Earth's mean density and a temperature of 55,000°F (30,000°C) (**Fig. 6**).

Differentiation occurs. The metallic envelope has a mass fraction of heavy elements that is less than 7.5 times the solar value, while this value ranges from 1 to 7.2 in the molecular envelope. The *Galileo* probe found that the helium-to-hydrogen ratio was 0.156 (near solar), but that carbon, sulfur, and xenon were enhanced and oxygen and neon depressed with respect to solar values. The large abundance of carbon suggests that gases other than those from the solar nebula contributed to the composition of Jovian volatile gases. Probably the entire planet holds 11–45 earth masses of elements other than hydrogen and helium.

Magnetosphere. Jupiter possesses the strongest magnetic field and most complex magnetosphere of any planet in the solar system (**Fig. 7**). The Jovian magnetosphere is the largest object in the solar system, extending well beyond the orbit of Saturn, which sometimes passes through it. If it could be seen visually from Earth, it would subtend several times the diameter of the full moon on the sky. The Jovian magnetic field rotates with the rotational period of the planet and contains an embedded plasma trapped in the field. At the distance of the satellite Io, the field revolves faster than the satellite, and so numerous collisions occur with the atmospheric gas of that body, resulting in the stripping away of 10^{28}–10^{29} ions per second. The energy involved slows the magnetic field, and so, beyond Io, the magnetic field no longer rotates synchronously with the planet. The ions removed from Io spiral around the magnetic lines of force, oscillating above and below the plane of Io's orbit. This ring of ions is known as the Io plasma torus and emits strongly in the ultraviolet. The motion of Io through the magnetosphere creates a 400,000-V, 2×10^{12} W circuit, sufficient to cause pronounced aurorae in both the equatorial and polar regions of the satellite.

Except near the planet, the major component of the Jovian magnetic field is dipolar and opposite in direction from that of the Earth. At a distance of three Jovian radii, the field strength is 0.16 gauss (16 microtesla). Closer to the planet, the field is quadrupolar and octopolar. There the field strength varies from 3 to 14 gauss (0.3 to 1.4 millitesla). The magnetic axis is aligned about 10° with respect to the rotational axis. The inner radiation ring flux intensity peaks at 2.2 Jovian radii, while its innermost edge is at 1.35 Jovian radii, well above the atmospheric cut-off. Radiation is also absorbed near the outer edge of the bright dust ring.

Ion sources for the magnetospheric plasma are the solar wind; the Jovian satellites, particularly Io; and the ionosphere itself. Ion temperatures in the Io plasma torus reach 17,500°F (10,000 K). The dayside Jovian magnetosphere extends to 105 Jovian radii. On the duskside high-latitude region, intense

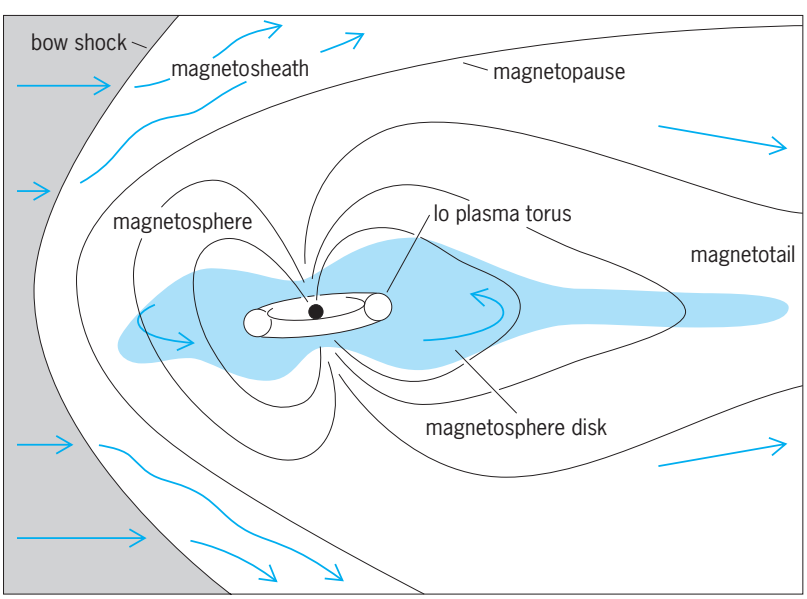

Fig. 7. Jovian magnetic environment. (*From Return to Jupiter, Astronomy, 7(9):6–24, 1979*)

fluxes of counterstreaming ions and electrons from the edge of the plasma torus to the duskside magnetosphere are aligned tightly with the magnetic field, superimposed on a hot plasma background. In addition, a thin plasma sheet (its depth at 15 Jovian radii is 2 radii) has temperatures of 100–400 eV. At polar latitudes, energetically charged particles following the Jovian magnetic lines of force may have access to the interplanetary medium. The Jovian magnetotail extends at least 4.5 astronomical units (4×10^8 mi or 6.5×10^8 km). *See* SOLAR WIND.

Jupiter possesses the most powerful aurorae in the solar system. They are observed at soft x-ray, ultraviolet, visible, and infrared wavelengths. Energies are 1000 times higher than in terrestrial aurorae. Protons, electrons, sulfur, and oxygen ions (from Io) precipitate onto the planet near its magnetic poles. These also form H_3^+ (trihydrogen) ions, the first ever observed. The *Galileo* probes observed visible auroral rings, highly structured and rapidly varying in time, 300–5000 mi (500–8000 km) in diameter at an altitude of 150 mi (240 km) above the 1-bar pressure level. *See* AURORA; MAGNETOSPHERE.

Radio astronomy. Jupiter produces distinct types of radio emission. Thermal radiation from the high stratosphere is detectable at wavelengths below about 10 cm and indicates temperatures in the upper emitting layers of −280 to −225°F (100 to 130 K). Microwave nonthermal emission in the band from about 3 to 70 cm arises from synchrotron radiation from relativistic electrons in extended Jovian Van Allen belts. Decametric radio emission (of frequencies 10–23 MHz) is also formed by synchrotron emission of these electrons and is 100% elliptically polarized. It occasionally produces S-bursts, with durations of tens of milliseconds, caused by instabilities of the electron beam in the Io-Jupiter flux tube. Hectometric emission (around 500 kHz) is probably rooted in the auroral ovals. The *Ulysses* probe detected radio emission at 3–30 kHz. Some radio noise is thought to originate in the lightning storms first observed by *Voyager 2* at depths of 50 mi (80 km) below the ammonia clouds. *See* POLARIZATION OF WAVES; SYNCHROTRON RADIATION; VAN ALLEN RADIATION.

Subsequent observations by the *Galileo* space probe have shown that lightning occurs in the Jovian atmosphere at a rate of approximately 20 bolts per second, approximately one-tenth the Earth's rate. The storms producing this lightning have been observed to be over 600 mi (1000 km) in diameter and over 45 mi (75 km) tall. Most lightning storms are concentrated between north latitudes 47 and 49°.

Periodicities in the radio-noise storms and rocking of the polarization plane of the microwave nonthermal emission led to the well-determined radio rotation period 9h55m29.7s. The difference between the radio and the various other observed Jovian rotation periods suggests that the core of Jupiter is rotating about 13 s faster than the mantle and that angular momentum may be significantly exchanged among the core, mantle, and atmosphere over periods of years. *See* RADIO ASTRONOMY.

Jovian ring. *Voyager 1* and *2* first detected a faint ring encircling Jupiter. Later probes found that Jupiter possesses a tripartite ring system. There is a flat main ring, a lenticular toroidal halo interior to this, and a gossamer ring which is exterior to the main ring. The entire ring system begins at about 55,000 mi (92,000 km) from Jupiter's center and extends to 150,000 mi (250,000 km). The main ring has a radial structure with dips in its apparent brightness due to the perturbations of the closer satellites. It is largely made up of 0.3–1.6-micrometer particles with human-scale lifetimes, whose source is dust and debris knocked off the Jovian moons XV Adrastea and XVI Metis by impacts with meteoroids, comets, and asteroids. Its inner and outer radii are respectively 76,900 mi (123,000 km) and 80,590 mi (128,940 km). The estimated mass of all particles forming it is 1×10^{13} kg. The toroidal halo interior to this ring was probably formed by electromagnetic forces pushing the smallest grains out of the ring plane. The exterior gossamer ring has two components, one associated with V Amalthea and the other with XIV Thebe. The Amalthea ring extends to precisely the orbital distance of the satellite and ends abruptly, while extending to one-half the thickness of the moon's excursion off the Jovian equatorial plane. The external gossamer ring particles orbit in retrograde orbits. The Amalthea ring forms the largest and sparsest ring known in the solar system. The material appears to be a dark reddish soot, consistent with the release of small particles from the two moons.

Satellites. As of October 2003, sixty-four known satellites had been discovered to orbit Jupiter, the largest number for any planet in the solar system. The four largest are I Io, II Europa, III Ganymede, and IV Callisto, discovered by Galileo in 1610 in Italy and independently by Simon Marius in Germany, who named them (though they are collectively known as the Galilean satellites) [see **table**].

The four Galilean satellites are of fifth and sixth stellar magnitudes and would be visible to the naked eye if they were not so close to the much brighter parent planet. They are easily visible in binoculars. The majority of the others are very faint objects only recently discovered with large telescopes.

The planes of the orbits of the major satellites are inclined less than 0.5° to the equatorial plane of Jupiter, so that with the occasional exception of IV, they are eclipsed in Jupiter's shadows and transit in front of its disk near conjunction. The eclipses, transits, and occultations of Jupiter's satellites led to the discovery of the finite velocity of propagation of light by O. Roemer in 1675. Satellites VIII, IX, XI, and XII have retrograde motion. *See* ECLIPSE; OCCULTATION; TRANSIT (ASTRONOMY).

The four Galilean satellites are orbitally connected. It is thought that Io was once much closer to Jupiter. Over the course of 100 million years, rotational energy was transferred to the satellite, causing it to move outward until it locked Europa into a 2:1 synchronous orbit. The two satellites then spiraled out as a pair and changed the eccentricity of Ganymede's orbit. Eventually, the mutual synchronicity will begin

The 16 brightest satellites of Jupiter

Satellite	Diameter, mi (km)	Mean distance from center of planet, 10^3 mi (10^3 km)	Orbital period, days	Magnitude at mean opposition	Year of discovery	Rotation*	Mass, 10^{20} lb (10^{20} kg)
XVI Metis	37 × 21 (60 × 34)	80 (128)	0.294	17.5	1979	Sync.	0.002 (0.001)
XV Adrastea	12 × 9 (20 × 14)	80 (129)	0.297	18.7	1979	Sync.	0.0004 (0.0002)
V Amalthea	155 × 80 (250 × 128)	113 (181)	0.498	14.1	1892	Sync.	0.16 (0.07)
XIV Thebe	72 × 52 (116 × 84)	138 (222)	0.674	16.0	1979	Sync.	0.017 (0.008)
I Io	2256 (3630)	262 (422)	1.769	5.0	1610	Sync.	1971 (894)
II Europa	1950 (3138)	417 (671)	3.551	5.3	1610	Sync.	1058 (480)
III Ganymede	3270 (5262)	665 (1,070)	7.155	4.6	1610	Sync.	3267 (1482)
IV Callisto	2983 (4800)	1170 (1883)	16.689	5.6	1610	Sync.	2372 (1076)
XIII Leda	10 (16)	6893 (11,094)	240	19.5	1974	Unknown	0.00013 (0.00006)
VI Himalia	116 (186)	7133 (11,480)	251	14.6	1904	Nonsync.	0.21 (0.10)
X Lysithea	22 (36)	7282 (11,720)	260	18.3	1938	Nonsync.	0.0017 (0.0008)
VII Elara	47 (76)	7293 (11,737)	260	16.3	1905	Nonsync.	0.017 (0.008)
XII Ananke	19 (30)	13,173 (21,200)	631	18.8	1951	Nonsync.	0.0008 (0.0004)
XI Carme	25 (40)	14,043 22,600	692	17.6	1938	Unknown	0.002 (0.001)
VIII Pasiphae	31 (50)	14,602 23,500	735	17.0	1908	Unknown	0.004 (0.002)
IX Sinope	22 (36)	14,726 (23,700)	758	18.1	1914	Nonsync.	0.0017 (0.0008)

* Synchronous or nonsynchronous.
SOURCE: After Lunar and Planetary Laboratory, University of Arizona; and J. A. Wood, Forging the planets: The orgin of our solar system, *Sky Telesc.*, 97(1):36–48, January 1999.

to affect the orbit of Callisto and lock it in, too. The gravitational effects of proximity of Jupiter and the slight elliptical orbits cause surface flexure and tidal heating strong enough to produce internal melting and major upheaval. All four satellites revolve within the Jovian magnetosphere, causing heavy ion bombardment of their surfaces.

The Jovian satellites fall into eight groups, based on their orbital characteristics. The Small Inner Irregulars and Rings (XVI Metis, XV Adrastea, V Amaltheal, XIV Thebe) are closest to the planet and govern ring structure. Irregular refers to their origin: they are probable captures from the asteroid belt after Jupiter's formation. The Galilean moons (as above) undoubtedly formed independently from the solar nebula or else as agglomerations of nebular material, likely at the same time as Jupiter. The Themisto (XVIII) and Himalia (VI) Prograde Irregular Groups comprise five satellites in total, all with moderately high eccentricites. They, too are likely captures and orbit the planet in prograde orbits. Three other groups (Retrograde Irregular, Carme [XI] Retrograde Irregular, and Pasiphae [VIII] Retrograde Irregular) orbit Jupiter in a retrograde direction and have moderate eccentricities. The newest discoveries have not been assigned to a group, though they share similar orbits and may be fragments of a single large body that broke up. In appearance, they are very dark and quite faint.

Io. The close approaches of the *Voyager* and *Galileo* spacecraft as well as the superior imaging capabilities of the *Hubble Space Telescope* have shown the four Galilean satellites to be very different (**Fig. 6**). I Io is probably the most geologically active body in the solar system. Its surface landforms include active shield volcanoes, calderas, mountains, plateaus, flows, grabens, and scarps. There are strong concentrations of craters near the sub- and anti-Jovian points. High mountains are equally distributed over the surface and are not connected with volcanoes. The tallest mountain known has an altitude of 52,000 ft (16 km). High-temperature (1700–2000 K) magnesium-rich silicate volcanism is common. Volcanism is likely due to a slight eccentricity in the orbit produced by the proximity of II Europa, which causes internal tidal forces to produce heat due to friction. Over 80 active volcanoes have been observed, though there are over 200 calderas larger than 12 mi (20 km) in diameter; by contrast, Earth has but 15. The most active volcanoes are Loki (responsible for half the satellite's total heat output) and Pele.

In 1999 a stupendous eruption was seen at Tvashtar Patera volcano. The lava fountain was observed to be over 0.9 mi (1.5 km) tall and 25 mi (40 km) long. The lava produced was hotter than 2420°F (1600 K) and was visible to Earth-based telescopes observing in the infrared.

Volcanic eruptions generally produce plumes resembling geysers containing fine dust and cold SO_2 (sulfur dioxide) gas and snow with exit velocities of 500–1000 m/s. Other compounds such as diatomic sulfur (S_2) are also ejected. Once on the surface, S_2 changes to S_3 and S_4, which are bright red. Over time, these become S_8 (the common yellow form of sulfur). Therefore, the most active volcanoes are surrounded by crimson rings which fade with time. Other bright colors on the surface are white and gray (SO_2 frost), yellow and brown (sulfurous compounds), red and black (showing the most recent activity), and green (origin unknown). Since Io possesses an ionosphere, there are auroral glows near the volcanic plumes. It is thought that a global surface layer is deposited at a rate of 1 mm to 1 cm every year, thereby explaining the total lack of impact craters. The internal structure of Io is thought to consist of an iron and iron sulfide core, surrounded by a molten silicate interior and a 19-mi-thick-(30-km) silicate crust, able to support the mountains.

Io has been found to have an intrinsic magnetic

field of an amplitude consistent with an internally driven dynamo. Therefore, it is a magnetized, solid planet. However, the field may turn on and off over million-year time scales as the interior recycles.

Despite its small size, Io also possesses a variable atmosphere consisting primarily of sulfur dioxide (with minor components of oxygen, sodium, and sulfur), which has been detected at infrared, millimeter, and ultraviolet wavelengths. It is extremely localized, covering only 10% of the satellite at any given time. Atmospheric pressure is only about 1 microbar. Recently it has been found that the proportion of chlorine in the Io atmosphere is the highest of any object in the solar system and may result from a volcanic eruption component of salt (NaCl), proceeding from underground rivers of salt water. The Jovian magnetosphere sweeps away 1000 kg of volcanic gases and surface material per second. This is the source of material in the Io plasma torus. The sodium cloud thus produced is as large as the Jovian magnetosphere itself, and has been detected not only in the Jovian vicinity but also in interplanetary space.

Europa. Europa possesses the highest albedo of any object in the solar system. The satellite has a varied surface terrain, dominated by tectonic deformation, impact cratering, and the possible emplacement of ice-rich materials and perhaps liquids on the surface. The leading hemisphere is covered with water frost, while the following is darker and redder, possibly caused by the bombardment of sulfur ions from Io. The underlying bright blue plains are the basic surface feature from which the other terrain types are derived. Tectonically deformed bright plains form the mottled regions through local disruptions that allow subsurface salts and clays to well up, akin to sea-floor spreading on Earth. There are striking linear features which range from several miles to 13 mi (20 km) across and which extend for 625 mi (1000 km) or more. These may demonstrate the existence of geyserlike cryovolcanism. Initial eruptions could involve gases or liquids plus rocks. Complex ridges also exist with heights of 330–660 ft (100–200 m), showing that parts of the icy crust have been modified by the intense faulting driven by energy from the planetary interior (**Fig. 9**). These ridges have a reddish coloration at their tops which may be from mineral contaminants in the ice. Finally, there are also impact craters with visible ejecta, though the centers are curiously muted, as though filled immediately with a slushy material. The age of the visible surface has been estimated to be less than 100 million years, due to the tectonic activity. The changing pattern of ridges is consistent with a nonsynchronous rotation of an ice shell with respect to a synchronously rotating interior. Dark patches may show venting of salt-laden water from below the ice. The core of Europa is thought to be iron-rich and is 781 mi (1250 km) in diameter. It is surrounded by a silicate mantle overlain by a subsurface ocean and a thin ice crust. The total ice shell thickness may be 6–94 mi (10–150 km), but it may thin to less than 0.6 mi (1 km) in some places. This is shown by visible "rafting": floating icebergs. That the salt-water ocean of Europa is still liquid below the surface is suggested by the observation that the slight magnetic field observed by *Galileo* changes direction to follow the Jovian magnetosphere. This could occur only in an induced field carried by a salt-rich true liquid layer. Europa also possesses a tenuous oxygen (O_2) atmosphere. It is not inconceivable that simple life could exist in the Europan oceans, much the same way as life-forms congregate near thermal vents in Earth's

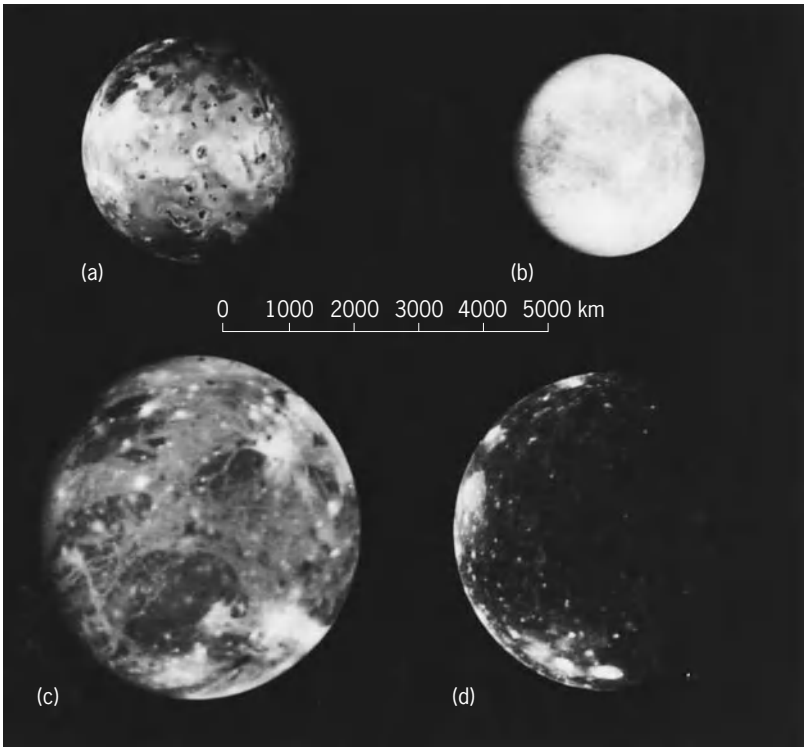

Fig. 8. *Voyager 1* photographs of the Galilean satellites, shown to scale. (*a*) Io. (*b*) Europa. (*c*) Ganymede. (*d*) Callisto. 1 km = 0.6 mi. (*NASA*)

Fig. 9. Intricately textured ridged plains on Europa, photographed by *Galileo* on December 16, 1997, from a distance of 800 mi (1300 km). The region covered is about 12 mi (20 km) on a side, with a resolution of only 85 ft (26 m) per picture element. Multiple generations of overlapping ridges are visible, many of them double. (*Jet Propulsion Laboratory*)

oceans. This possibility led *Galileo* project managers to choose to destroy the probe in 2003 by allowing it to fall into the Jovian atmosphere, rather than to contaminate a possible European ecosystem with Earth microbes.

Ganymede. The density of impact craters on the surface of III Ganymede shows that it is several billion years old. There are two major types of terrain: older, heavily fractured dark and rocky areas and brighter, less densely cratered areas rich in water and dry ice. Crossing the dark terrain are 3–6 mi-wide (5–10 km) furrows which may have formed from impacts which released water ice from below in a form of cryovolcanism. Dark maerial on the surface is likely meteoritic dust. Water and dry ice frost many of the impact craters. Fifty percent of the surface is made up of grooved terrain caused by tectonic stretching of the moon's surface into grabens. Some of the low-lying valleys were flooded with slush about 1 billion years ago, leading to some thought that a subsurface ocean may exist here, also. The partially molten iron core of Ganymede may take up to 50% of its diameter. This is surrounded by a silicate lower mantle, an icy upper mantle, and a thin ice crust. The satellite possesses a magnetic field with a strength 1% that of Earth's. It may possibly form in a salty, electrically conducting layer of water some 110 mi (170 km) below the surface.

Ganymede possesses a slight atomic hydrogen and O_2-O_3 (ozone) atmosphere both above the surface and trapped within the surface ice. In fact, ozone is surprisingly abundant near the poles, while the O_2 is found near the equator.

There is faint fluorescence of oxygen in the ultraviolet at the poles, forming an aurora, while there is a visible light aurora at the equator. A dusting of frost near the poles may derive from the manner in which Ganymede's magnetic field interacts with Jupiter's. Ions and electrons preferentially cascade over the polar regions, which may cause sublimation of water which then accumulates as piles of frost.

Callisto. Since it is similar in radius and density, IV Callisto should have a geological history resembling that of Ganymede, but does not. It is the most globally heavily cratered of the Galilean satellites, with large impact craters (for example Valhalla, with a diameter of 375 mi [600 km]) formed early in its history. The heavy cratering is indicative of its great age, but there is a curious lack of small craters, suggesting that there is ice-rich debris that subsumes them. Sublimation erosion by the Sun is the dominant process, causing crater rims and ridges to become heavily eroded. A long chain of craters is visible which was likely caused by the impact of a comet akin to Shoemaker-Levy 9. The centers of impact craters are often very bright, suggesting frost deposits. There is definite SO_2 frost concentrated on the forward-facing hemisphere. The impact craters can be muted, as if water ice slush filled them after meteoritic impact. There are few signs of internal geology. The crust may be 124 mi (200 km) thick, with a salty ocean 6 mi (10 km) thick beneath. Below this is compressed rock and ice, most forming an undifferentiated core. The liquid ocean is suggested by magnetic irregularities observed by Galileo; there is no intrinsic magnetic field, nor is there an atmosphere. This curious internal structure shows that Callisto formed in a cold environment, unlike the other three Galilean satellites which achieved differentiated cores by forming in a hot region.

Other satellites. The Galileo spacecraft was the first to image the four inner satellites of Jupiter, XVI Metis, XV Adrastea, V Amalthea, and XIV Thebe. All are 25–35% brighter on their leading sides than on their trailing sides. Metis and Adrastea are redder in coloration. Bright spots were seen on Amalthea and Thebe (diameters less than 13 mi [20 km]), as well as what appeared to be higher albedo patches on ridges and crater rims. These probably represent deposition of material from Io. The major geologic features on all four inner satellites are impact craters. The largest is 41 mi (60 km) across on Amalthea. Since its density is too low to allow it to be a solid object, Amalthea is almost certainly a loose agglomeration, held together by gravity. If it ever was solid, it may have been shattered into sand and boulder-sized rubble by impacts.

VI Himalia has been shown from photometry to consist of carbonaceous chondritic material, and probably the remainder of the satellites of Jupiter are constructed of similar material. Satellites VII–XII are asteroidal in nature. They resemble the Trojan asteroids in their colors and albedos, and may share a common origin. They are less elliptical in shape than the other small Jovian satellites, perhaps because they spent more time in the proto-Jupiter nebula. However, their properties are also consistent with a capture origin, and their orbits are subject to large perturbations by the Sun. Thus they may form part of a fluctuating population gained and lost over very long time spans. *See* ASTEROID; PLANETARY PHYSICS; TROJAN ASTEROIDS. Elaine M. Halbedel

Bibliography. J. K. Beatty, *Galileo*: An image gallery II, *Sky Telesc.*, 93(3):28–35, March 1997; J. K. Beatty and D. H. Levy, Crashes to ashes: A comet's demise, *Sky Telesc.*, 90(4):18–26, October 1995; J. K. Beatty, C. C. Petersen, and A. Chaikin (eds.), *The New Solar System*, 4th ed., 1998; M. J. S. Belton (ed.), *In Orbit at Jupiter: Contributions of the Galileo Imaging Science Team*, NASA Galileo Proj. Doc. 625–801, 2000; M. J. S. Belton et al., Galileo's first images of Jupiter and the galilean satellites, *Science*, 274:377–385, 1996; D. Fischer, *Mission Jupiter: The Spectacular Journey of the Galileo Spacecraft*, Copernicus Books, 2001; M. Hanlon, *The Worlds of Galileo: The Inside Story of NASA's Mission to Jupiter*, St. Martin's Press, 2001; D. M. Harland, *Jupiter Odyssey: The Story of NASA's Galileo Mission*, Springer-Praxis, 2000.

Jurassic

The system of rocks deposited during the middle part of the Mesozoic Era, and encompassing an interval of time between about 200 and 142 million

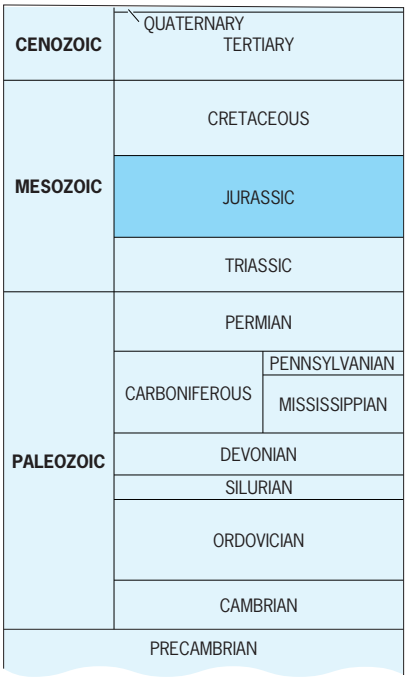

years ago, based on radiometric dating. It takes its name from the Jura Mountains of Switzerland. Its rich marine invertebrate faunas in western Europe have been the subject of intensive study since the pioneering days of geology in the early nineteenth century, and provided the basis for the fundamental stratigraphic concepts of stages and biozones. *See* DATING METHODS.

Subdivisions. The Jurassic System is subdivided into 11 stages which, with the exception of the Tithonian, are named from localities in England, France, and Germany (**Fig. 1**). These, and the much greater number of zones, are based upon ammonites, which are by far the most valuable fossils biostratigraphically because of their high rate of species turnover in time due to rapid evolution and extinction. The most refined stratigraphic subdivisions have been made in the British Isles, with 54 zones and 176 subzones. Because of biogeographic provinciality, with different ammonite taxa inhabiting Boreal and Tethyan realms, difficulties of correlation can occur for younger Jurassic strata; and the youngest stage in the Tethyan Realm, the Tithonian, embracing most of the world, is equivalent to the Volgian stage of the Boreal Realm, extending from northern Eurasia to northern North America. The ammonites of the Volgian are quite different from those of the stratigraphically equivalent Tithonian. In the absence of ammonites, dinoflagellates are the most useful marine fossils for correlation, but in nonmarine strata problems of correlation are considerable, and stratigraphically less satisfactory pollen and spores have to be used. *See* STRATIGRAPHY.

SUBSYSTEM	STAGE	AGE, 10^6 years
		142
UPPER	Tithonian	149
	Kimmeridgian	155
	Oxfordian	157
MIDDLE	Callovian	160
	Bathonian	166
	Bajocian	174
	Aalenian	178
LOWER	Toarcian	183
	Pliensbachian	192
	Sinemurian	197
	Hettangian	200

Fig. 1. Succession of Jurassic stages, with estimated radiometric ages in millions of years. (*After J. Palfy et al., A U-Pb and $^{40}Ar/^{39}Ar$ time scale for the Jurassic, Can. J. Earth Sci., 37:923–944, 2000*)

Paleogeography and sea level. The main continental masses were grouped together as the supercontinent Pangaea, with a northern component, Laurasia, separated from a southern component, Gondwana, by a major seaway, Tethys, which expanded in width eastward (**Fig. 2**). From about Middle Jurassic times onward, this supercontinent began to split up, with a narrow ocean being created between eastern North America and northwestern Africa, corresponding to the central sector of the present Atlantic Ocean. At about the same time, and continuing into the Late Jurassic, separation began between the continents that now surround the Indian Ocean, namely Africa, India, Australia, and Antarctica. As North America moved westward, it collided with a number of oceanic islands in the eastern part of the Paleo-Pacific. Because the impingement was an oblique one, there was a general tendency for these accreted landmasses to be displaced northward along the cordilleran zone of the subcontinent. Other examples of so-called displaced terranes are known on the Asian side of the North Pacific, and some of the accretion of oceanic islands took place in Jurassic times.

There were also important paleogeographic changes later in the period involving the Tethys zone. An older, so-called Palaeotethys was progressively closed as an extensive, narrow continent known as Cimmeria, extending east-west, and collided with the southern margin of Eurasia. The name comes from the Crimean Peninsula of Russia, where there is well-displayed evidence of an intra-Jurassic orogenic disturbance indicative of continental collision. *See* OROGENY; PALEOGEOGRAPHY.

Sea level rose progressively through the period, with a corresponding flooding of the continents by shallow epeiric seas, that is, shallow seas that covered part of the continents but remained connected to the ocean. At the beginning, such seas covered less than 5% of the continents, but near the end, in Oxfordian and Kimmeridgian times, they covered approximately 25% (Fig. 2). The Jurassic sea-level curve also shows a succession of smaller-scale changes, of a duration of a few million years. Some of these, such

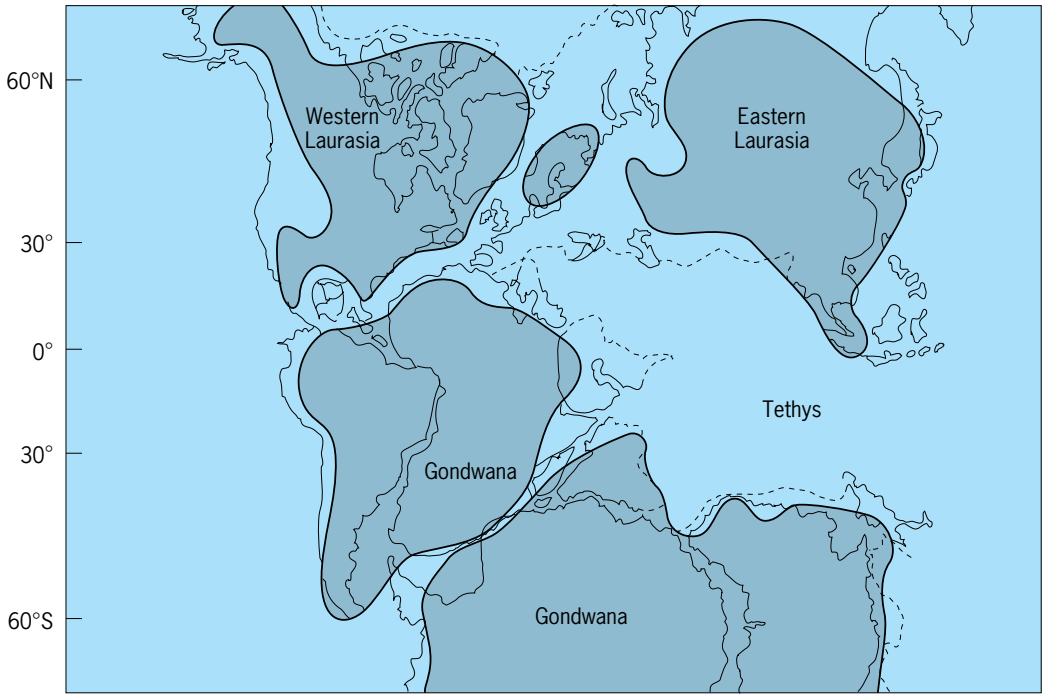

Fig. 2. Approximate distribution of land and sea in the Oxfordian stage. Small islands are excluded, but boundaries of modern continents are included as a reference.

as the Early Toarcian sea-level rise, are clearly global or eustatic, but others are more controversial and may reflect regional tectonic activity rather than truly global phenomena. It is uncertain by how much the sea level rose during the course of the period; but by using a hypsometric method, an estimate of between 330 and 500 ft (100 and 150 m) can be made. *See* PALEOCEANOGRAPHY.

Climate. The climate of Jurassic times was clearly more equable than at present, as indicated by two sets of facts. The first concerns the distribution of fossil organisms. Thus a number of ferns whose living relatives cannot tolerate frost are distributed over a wide range of paleolatitudes, sometimes as far as 60° N and S. Similarly, coral reefs, which are at present confined to the tropics, occur in Jurassic strata in western and central Europe, beyond the paleotropical zone. Many other groups of organisms had wide latitudinal distribution, and there was much less endemism (restriction to a particular area) with respect to latitude than there is today. The second set of facts concerns the lack of evidence for polar icecaps, such as extensive tillites or striated pavements.

However, there must have been strong seasonal contrasts of temperature within the Pangean supercontinent, and climatic modeling suggests winter temperatures at zero Celsius at or close to the paleopoles. A limited amount of evidence from northern Siberia and arctic North America, in the form of apparent glacial dropstones and glendonites, suggests the possibility of some ice, but this ice is likely to have been seasonally transient and small in volume.

There is no evidence of any significant change in the temperature regime through the Jurassic, but there are indications of a change in the humidity-aridity spectrum. Unlike the present, there were no tropical rainforests. Instead, a large area of western Pangea experienced an arid to semiarid climate in low latitudes, especially at some distance from the ocean. Precipitation is likely to have been dominantly monsoonal rather than zonal, a pattern unlike that of today. For most of the period, the continental area represented today by Eurasia had a comparatively humid climate, as indicated in nonmarine sediments by coals and the abundance of the clay mineral kaolinite. Toward the end of the Jurassic, however, there was a change to a more arid climate, indicated by the disappearance of coals and kaolinite and the occurrence of evaporites such as rock salt and gypsum. The reason for this change is unclear, but it may be bound up with a rainshadow effect created by the collision of the Cimmerian continent. *See* PALEOCLIMATOLOGY; SALINE EVAPORITES.

Tectonics and volcanicity. Most of Pangea experienced tensional tectonics as the supercontinent began to break up. This is manifested by graben and half-graben structures, with associated alkaline volcanicity. By far the largest flood basalt province is that of the Karoo in South Africa, most of the basalts and associated igneous rocks being erupted in the Early Jurassic, prior to the breakup of Africa, Madagascar, and India. The Middle Jurassic Ferrar dolerites of Victoria Land, Antarctica, and the contemporaneous Tasmanian dolerites are further manifestations of tensional tectonics, as are earliest Jurassic basalts in eastern North America and Morocco, again signifying tension prior to the Atlantic opening. The North Sea region of western Europe is in effect an aborted oceanic rift, with a major phase of tensional activity and associated volcanicity in the Middle and Late

Jurassic. This did not lead, however, to the creation of true ocean. *See* BASALT; GRABEN.

Compressional tectonics associated with subduction of ocean floor took place in many parts of the Pacific margins, with associated calc-alkaline volcanicity. An excellent example is the Andes. The North Pacific margins were also associated with significant strike-slip faulting bound up with the accretion of displaced terranes. The other important zone of compressional tectonics was along the southern margin of Eurasia, and is involved with the collision of the Cimmerian continent. *See* FAULT AND FAULT STRUCTURES.

Since it is not plausible to invoke the melting and freezing of polar ice caps to account for Jurassic sea-level change, this change must be bound up with tectonic activity. The most plausible mechanism for accounting for long-term sea-level rise is the growth of oceanic ridges, displacing seawater onto the continents, but the cause of short-term sea-level changes is more obscure and remains controversial. *See* CONTINENTS, EVOLUTION OF; GEOSYNCLINE; MID-OCEANIC RIDGE; PLATE TECTONICS; SUBDUCTION ZONES.

Vertebrate fauna. The vertebrate terrestrial life of the Jurassic Period was dominated by the reptiles. The dinosaurs had first appeared late in the Triassic from a thecodont stock, which also gave rise to pterosaurs and, later, birds. From small bipedal animals such as *Coelophysis*, there evolved huge, spectacular creatures. These include the herbivorous *Apatosaurus*, *Brontosaurus*, *Brachiosaurus*, *Diplodocus*, and *Stegosaurus* as well as the carnivorous, bipedal *Allosaurus*. Only two rich dinosaur faunas are known from Jurassic deposits, the Morrison Formation of the United States Western Interior and the approximately contemporary Tendaguru Beds of Tanzania. The two faunas are strikingly similar at family and generic level, which strongly suggests that free land communications existed between western North America and East Africa until quite late in the period, a fact that is not easy to reconcile with some paleogeographic reconstructions. *See* DINOSAUR.

Flying animals include the truly reptilian pterosaurs and the first animals that could be called birds as distinct from reptiles, as represented by the pigeon-sized *Archaeopteryx*. There were two important groups of reptiles that lived in the sea, the dolphinlike ichthyosaurs and the long-necked plesiosaurs. Both of these groups had streamlined bodies and limbs beautifully adapted to marine life. Turtles and crocodiles are also found as fossils in Jurassic deposits. *See* ARCHAEORNITHES; PTEROSAURIA.

Jurassic mammals, known mainly from their teeth alone, were small and obviously did not compete directly with the dinosaurs. They included a number of biologically primitive groups such as the triconodonts, docodonts and multituberculates. The fish faunas were dominated by the holosteans, characterized by heavy rhombic scales. Their evolutionary successors, the teleosts, probably appeared shortly before the end of the period. *See* DOCODONTA; HOLOSTEI; MULTITUBERCULATA; TELEOSTEI; TRICONODONTA.

Invertebrate fauna. Because they are far more abundant, the invertebrate fossil faunas of the sea are of more importance to stratigraphers and paleoecologists than are the vertebrates. By far the most useful for stratigraphic correlation are the ammonites, a group of fossil mollusks related to squids. They were swimmers that lived in the open sea, only rarely braving the fluctuating salinity and temperature of inshore waters. They are characteristically more abundant in marine shales and associated fine-grained limestones. From a solitary family that recovered from near extinction at the close of the Triassic, there radiated an enormous diversity of genera. Many of these were worldwide in distribution, but increasingly throughout the period these was a geographic differentiation into two major realms. The Boreal Realm occupied a northern region embracing the Arctic, northern Europe, and northern North America. The Tethyan Realm, with more diverse faunas, occupied the rest of the world. *See* LIMESTONE; SHALE.

In most facies the bivalves, which flourished in and on shallow, muddy sea bottoms, are the most abundant and diverse of the macrofauna. They included many cemented forms such as *Ostrea*, recliners such as *Gryphaea*, swimmers such as the pectinids and limids, and rock borers such as *Lithophaga*. However, the majority were burrowers: either relatively mobile, shallow burrowers or forms occupying deep permanent burrows and normally still found in their positions of growth. *See* BIVALVIA; FACIES (GEOLOGY).

Brachiopods were much more abundant and diverse than they are today. The range of depths below the sea surface that they occupied is far wider than for the bivalves, and a definite depth zonation can be established in Europe, just as with the ammonites. *See* BRACHIOPODA.

Echinoderms are best represented as fossils by the crinoids and echinoids, and were all inhabitants of shallow seas, unlike some of the modern representatives of this class. The echinoids include both primitive regular forms, such as the cidaroids, and irregular forms, such as *Clypeus* and *Pygaster*. *See* ECHINODERMATA; PYGASTEROIDA.

Corals belonged to the still extant Scleractinia group and included reef builders such as *Isastrea* and *Thamnasteria*. Calcareous and siliceous sponges are also common locally, even forming reefs. It seems likely that the siliceous sponges inhabited somewhat deeper water than the corals. *See* SCLERACTINIA; SCLEROSPONGIAE.

The invertebrate microfaunas are represented by abundant foraminifera, ostracods, and radiolaria. Foraminifera and ostracods are of great value to oil companies in correlation studies. *See* OSTRACODA; RADIOLARIA.

Not all Jurassic invertebrates lived in the sea. Some lived in continental environments such as lakes and rivers; they include a few genera of bivalves, gastropods, and arthropods. These faunas are far less diverse than their marine counterparts. *See* ARTHROPODA; GASTROPODA; PALEONTOLOGY.

Flora. With regard to the plant kingdom, the Jurassic might well be called the age of gymnosperms, the nonflowering "naked seed" plants, forests of which covered much of the land. They included the conifers, gingkos, and their relatives, the cycads. Ferns and horsetails made up much of the remainder of the land flora. These and others of the Jurassic flora are still extant in much the same forms. *See* CYCADALES; GINKGOALES.

Remains of calcareous algae are widely preserved in limestone. Besides the laminated sedimentary structures produced by what have traditionally been regarded as blue-green algae but are actually cyanobacteria, and known as oncolites and stromatolites, there are skeletal secretions of other groups. Some of these are benthic forms, but many pelagic limestones are seen under the electron microscope to be composed largely of tiny plates of calcite, known as coccoliths, which are secreted by certain planktonic algae also called coccoliths. *See* ALGAE; CYANOBACTERIA; STROMATOLITE.

It seems likely that the Late Jurassic saw the emergence of the flowering plants, the angiosperms, since well-developed forms of this group existed in the Early Cretaceous. However, it is not quite understood how they emerged, and a satisfactory direct evolutionary ancestor has yet to be identified with certainty.

Economic geology. Jurassic source rocks in the form of organic-rich marine shale and associated rocks contain a significant proportion of the world's petroleum reserves. A familiar example is the Upper Jurassic Kimmeridge Clay of the North Sea, and its stratigraphic equivalents in western Siberia. Some of the source rocks of the greatest petroleum field of all, in the Middle East, are also of Late Jurassic age. *See* MESOZOIC; PETROLEUM GEOLOGY. A. Hallam

Bibliography. W. J. Arkell, *Jurassic Geology of the World*, 1956; J. W. C. Cope et al. *Jurassic*, pts. 1 and 2, Geol. Soc. Lond. Spec. Rep. 14 and 15, 1980; A. Hallam, Jurassic climates as inferred from the sedimentary and fossil record, *Phil. Trans. Roy. Soc. Lond.*, B341:287–296, 1993; A. Hallam, *Jurassic Environments*, 1975; A. Hallam, A review of the broad pattern of Jurassic sea-level changes and their possible causes in the light of current knowledge, *Palaeogeog., Palaeoclimatol., Palaeoecol.*, 167:23–37, 2001.

Jute

A natural fiber obtained from two Asiatic species, *Corchorus capsularis* and *C. olitorius*, of the plant family Tiliaceae (**Fig. 1**). These are tall, slender, half-shrubby annuals, 8–12 ft (2.5–3.5 m) tall. *See* MALVALES.

When harvested the stems are retted in tanks or pools to rot out the softer tissues. The strands of jute fiber are then loosened by beating the stems on the surface of the water. The fibers are not very strong and deteriorate quickly in the presence of moisture, especially salt water. Despite these weaknesses, jute is much used. It is inexpensive and easily spun and converted into coarse fabrics. It is made into gunny, burlap bags, sacks for wool, potato sacks, covers for cotton bales, twine, carpets, rug cushions, curtains, and a linoleum base. It is also used in making coarse, cheap fabrics, such as novelty dress goods. Most of the commercial supply comes from plants grown in the Ganges and Brahmaputra valleys in Bangladesh and India. *See* FIBER CROPS; NATURAL FIBER. Elton G. Nelson

Fig. 1. Morphological features of *Corchorus capsularis*.

A number of diseases that affect jute cause losses in yield and reduce fiber quality. "Runner" and "specky" fiber are primarily due to disease-producing organisms.

The fungus *Macrophomina phaseolina* is believed to cause the most serious disease of the two species of jute. It is seed-borne and soil-borne, and pycnidiospores from susceptible plants besides jute also serve as sources of infection. The stem, leaves, and roots of both young and older plants are subject to attack. Stem infection usually takes place through a leaf petiole or at a node (**Fig. 2**). Root rot is complicated in that severity is increased when *M. phaseolina* is in combination with other fungi, bacteria, or nematodes, such as *Fusarium solani, Pseudomonas* sp., and *Meloidogyne incognita*, respectively. *See* AGONOMYCETES; LEAF; NEMATA (NEMATODA); ROOT (BOTANY); SEED; STEM.

In contrast to *Macrophomina phaseolina*,

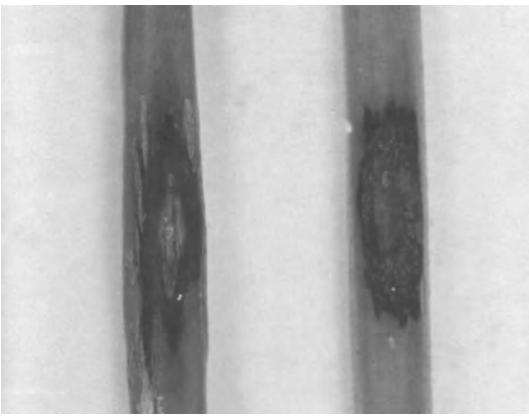

Fig. 2. Disease lesions on jute stems.

Colletotrichum capsici causes lesions on the stem internodes and may also attack seedlings and capsules of *C. capsularis. Macrophoma corchori* and *Diplodia corchori* cause stem diseases. Two species of bacteria, *Xanthomonas makatae* and *X. makatae* var. *olitorii*, attack the stem and leaves of both *C. capsularis* and *C. olitorius. Pythium splendens* causes a root rot and subsequent wilt of *C. capsularis*, and indications are that other species of *Pythium* also are root pathogens of jute. *See* FRUIT.

Other fungi which attack jute are *Sclerotium rolfsii*, *Curvularia subulata*, *Cercospora corchori*, *Rhizoctonia solani*, *Helminthosporium* sp., and *Alternaria* sp.

Seed treatments with approved fungicides are recommended for control of seed-borne pathogens and seedling diseases. Stem rot may be prevented by spraying with fixed copper compounds. The excessive use of nitrogenous fertilizers increases the incidence of stem diseases. Root-rot control requires the use of crop rotation and, in some areas, the use of recently developed varieties of *C. capsularis* which are more tolerant of certain root-rot pathogens than *C. olitorius*. *See* PLANT PATHOLOGY.

Thomas E. Summers

K

Kalahari Desert — Kyanite

Kalahari Desert

A flat, largely waterless, sparsely populated, and sand-covered region in south-central Africa which occupies much of Botswana, plus adjacent parts of South Africa and Namibia. The precise boundaries of the Kalahari are difficult to define since it merges into humid regions to the north and east, and the arid Karoo and Namib further south and west. However, the generally agreed-upon boundaries are the Okavango-Zambezi "swamp zone" in the north and the Orange River in the south, with the western and eastern boundaries (in eastern Namibia and eastern Botswana, respectively) coinciding with the extent of the sandy Kalahari group sediments (see **illustration**). *See* AFRICA.

Climate in the Kalahari varies from arid to semiarid along a southwest–northeast gradient. The driest, southwestern areas (which can be considered true desert) receive as little as 150 mm (6 in.) of rainfall per year, with precipitation levels rising to the north and east to reach around 650 mm (26 in.) in northeast Botswana. However, much of this rain falls during the Southern Hemisphere summer months when high evapotranspiration rates also occur (exceeding 4000 mm or 157 in. per year in the southwest Kalahari), contributing to the desertlike conditions. *See* DESERT.

The Kalahari terrain is relatively flat, broken only by low hills along its periphery and dunefields in the extreme southwest. Permanent drainage is restricted to the north where the Chobe, Zambezi, and Okavango rivers dissect the landscape, the latter ending abruptly in the Okavango Delta swamps. Water is also present on a seasonal basis within the salt lakes of the Makgadikgadi and Etosha basins of central Botswana and northern Namibia, respectively, and within numerous smaller landscape depressions (called pans). In addition to rivers and lake basins, extensive networks of "fossil" dry valleys also cut across the region, suggesting that wetter climates may have occurred in the recent geologic past. *See* DUNE.

Soils developed within the sandy Kalahari sediments are typically skeletal (weakly bound) and have a relatively low nutrient status, although finer-textured soils may be found in association with rivers, pans, dry valleys, and interdune areas. Vegetation patterns broadly follow the climatic gradient, in terms of diversity and biomass, although, again, local variations occur near drainage features. Much of the region is covered by savanna communities, with shrub savanna prevalent in the southwest and tree savanna in the central and northern Kalahari. Exceptions occur in the Okavango swamps, where aquatic grasslands and riparian forests dominate, around saline lake complexes where savanna grasslands are common, and in the relatively wet north and east where dry deciduous forests are the

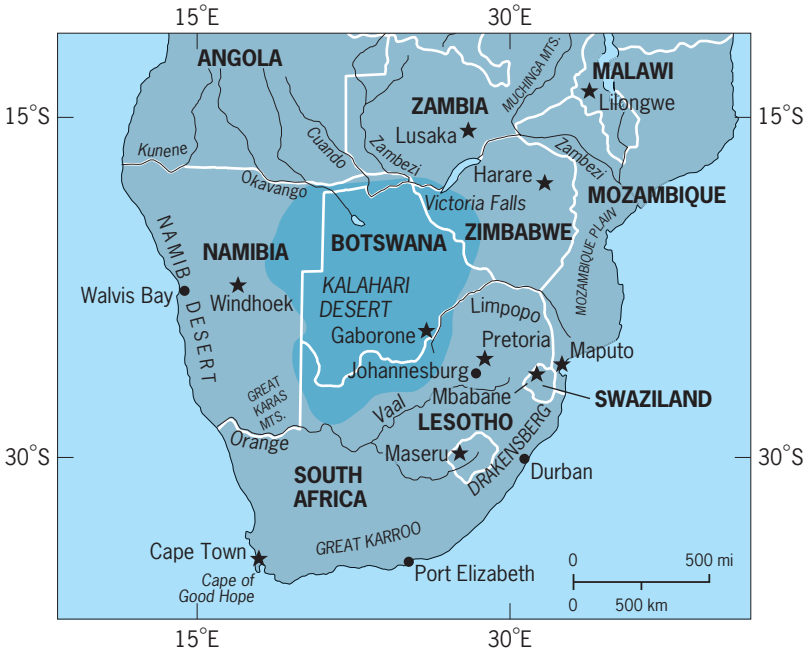

Map of the Kalahari Desert.

dominant vegetation type. *See* PLANT GEOGRAPHY; SAVANNA; SOIL.

The Kalahari has supported, and in some areas still supports, large populations of migratory wild herbivores, including antelope, zebra, and giraffe, along with associated predators, including lion, leopard, cheetah, and hyena. Pressures from hunting and increasing numbers of domestic herbivores, particularly cattle (with the associated fenced-in land), have meant that wildlife populations have been increasingly restricted to protected areas. At the same time, human populations are growing, most notably along the eastern and southern desert periphery (such as Gaborone, Botswana and Upington, South Africa), as well as within the desert core in conjunction with tourism (such as Maun, Botswana) and diamond mining (such as Orapa and Jwaneng, Botswana) developments. David J. Nash

Bibliography. A. S. Goudie, *Great Warm Deserts of the World*, Oxford University Press, 2002; D. Sporton and D. S. G. Thomas (eds.), *Sustainable Livelihoods in Kalahari Environments: Contributions to Global Debates*, Oxford University Press, 2002; D. S. G. Thomas and P. A. Shaw, *The Kalahari Environment*, Cambridge University Press, 1991; L. van der Post, *The Lost World of the Kalahari*, Penguin, London, 1962; E. N. Wilmsen, *Land Filled with Flies: A Political Economy of the Kalahari*, University of Chicago Press, 1989.

Kale

Either of two cool-season biennial crucifers, *Brassica oleracea* var. *acephala* and *B. fimbriata*, of Mediterranean origin. Kale belongs to the plant order Capparales, and the plant family Brassicaceae, commonly known as the mustard family. Cabbage, kohlrabi, broccoli, collards, and Brussels sprouts are all variants of *B. oleracea*

Kale is grown for its nutritious green curled leaves which are cooked as a vegetable (see **illus.**). Distinct varieties (cultivars) are produced in Europe for stock feed. Kale and collards differ only in the form of their leaves; both are minor vegetables in the United States. Cultural practices are similar to those used for cabbage, but kale is more sensitive to high temperatures. Strains of the Scotch and Siberian varieties are most popular. Kale is moderately tolerant of acid soils. Monthly mean temperatures below 70°F (21°C) favor best growth. Harvesting is generally 2–3 months after planting. Virginia is an important producing state. *See* BROCCOLI; BRUSSELS SPROUTS; CABBAGE; CAPPARALES; COLLARD; KOHLRABI. H. John Carew

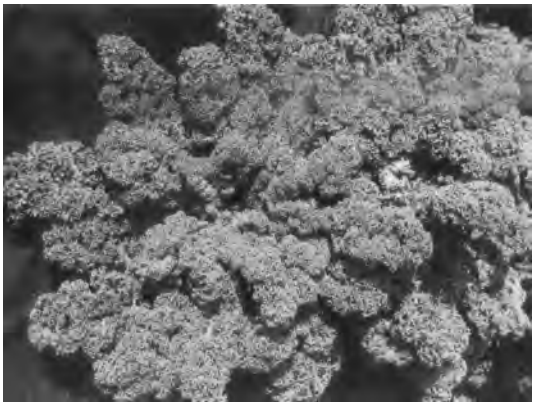

Kale (*Brassica oleracea* var. *acephala*), cultivar Vates. (*Joseph Harris Co., Rochester, New York*)

Kaliophilite

A rare mineral tectosilicate found in volcanic rocks high in potassium and low in silica. Kaliophilite is one of three polymorphic forms of $KAlSiO_4$; the others are the rare mineral kalsilite and an orthorhombic phase formed artificially at about 930°F (500°C). It crystallizes in the hexagonal system in prismatic crystals with poor prismatic and basal cleavage. The hardness is 6 on Mohs scale, and the specific gravity is 2.61. At high temperatures a complete solid-solution series exists between $KAlSiO_4$ and $NaAlSiO_4$, but at low temperatures the series is incomplete. The principal occurrence of kaliophilite is at Monte Somma, Italy. *See* SILICATE MINERALS. Cornelius S. Hurlbut, Jr.

Kalsilite

A rare mineral described in 1942 from volcanic rocks at Mafuru, in southwest Uganda. Kalsilite has since been synthesized. It is one of the three polymorphic forms of $KAlSiO_4$; the others are kaliophilite and an orthorhombic phase formed artificially at about 930°F (500°C). The mineral as shown by x-ray photographs is hexagonal. The specific gravity is 2.59. In index of refraction and general appearance in thin section it resembles nepheline and is difficult to distinguish from it. Structurally the two minerals are similar but belong to different crystal classes. The rock in which kalsilite was found has a dark-colored, fine-grained matrix with large olivine crystals and greenish-yellow patches. These patches are intimate mixtures of diopside, calcite, and kalsilite. *See* FELDSPATHOID. Cornelius S. Hurlbut, Jr.

Kangaroo

Common name for a member of the family Macropodidae, the largest living marsupials. The 61 species in this family also include the wallabies, wallaroos, quokka, and pademelons. In Australia, the term kangaroo is applied only to the largest species of kangaroos, while all small and medium-sized ground-dwelling kangaroos are called wallabies. Except for their size, wallabies are practically the same as kangaroos in appearance as well as habits. *See* AUSTRALIA; MARSUPIALIA.

Description. There are many different kinds of kangaroos, ranging from the familiar great kangaroos

Red kangaroo (*Macropus rufus*). (*Photo by Gerald and Buff Corsi; © California Academy of Sciences*)

(*Macropus*) to the musky rat kangaroo (*Potoroidae*), which is roughly the size of a rabbit but more like a rat in appearance. There are three species of great kangaroos, all of which have extremely long, powerful, highly modified hindlimbs and relatively small, unspecialized forelimbs. The long, thick, muscular tail serves as a balancing organ and also adds impetus to the leaps. Large kangaroos can sustain hopping speeds faster than 55 km/h (35 mi/h). The western gray kangaroo or Forester (*M. fuliginosus*) has a brownish coat with black hands and feet. The eastern gray kangaroo (*M. giganteus*) is the only large kangaroo in Tasmania. It is more reddish-brown than the western gray kangaroo and has coarser and longer fur. The red kangaroo (*M. rufus*) [see **illustration**] is the largest and most gregarious kangaroo. It is richly colored, powerful, and graceful. The male is a brilliant wine-red. The female, or doe, is more slender than the male, is more lightly built, and has great speed. A smoky-blue in color, the female has been named the blue flyer. Adult males, or boomers, may have a head and body length of 1300–1600 mm (50–62 in.) and a tail length of 1000–1200 mm (39–46 in.). They may weigh 20–90 kg (43–197 lb) but seldom exceed 55 kg (120 lb). Females are smaller.

Kangaroos and their relatives primarily inhabit the prairies, open forest, and brush country in Australia but also are found in Tasmania and New Guinea. They feed mostly on plant foods including grasses, forbs (herbs other than grass), leaves, seeds, fruits, and bulbs. The dental formula is I 3/1, C O/O, Pm 2/2, M 4/4 for a total of 32 teeth. Following a gestation of 30–39 days, a single young (joey) is born. A newborn macropod is only 5–15 mm (0.2–0.6 in.) long with undeveloped eyes, hindlimbs, and tail. It uses its strong, well-developed forelimbs to crawl up and into its mother's pouch (marsupium), where it attaches to a teat. Depending on the species, the joey may remain completely in the pouch for 60–120 days. Following the pouch period, it will venture out to find food but will regularly return to the pouch, especially in times of danger. This period, which may last for several weeks to several months, is known as the young-at-foot period. When entering the pouch, the joey enters headfirst and then turns around and pokes its head back out. The joey is forced from the pouch just before the birth of the next young, but it can still put its head back into the pouch and suckle from the teat. The female is capable of producing different qualities of milk from the two teats—a feat achieved by having the mammary glands under separate hormonal control. Longevity in the wild is normally 12–18 years. Although great numbers of red and gray kangaroos are shot each year because they are considered pests of crops and pastures, the chief natural predator is the dingo (*Canis lupus familiaris*), the wild Australian dog.

Tree kangaroos. Tree kangaroos (*Dendrolagus*) are arboreal. They still retain the long hindlimbs for leaping, but the limbs have become shorter and broader, and the foot pads are rough to assist in climbing. They have a head and body length of 520–810 mm (20–32 in.) and a tail length of 408–935 mm (16–36 in.) They weigh 6.5–14.5 kg (14–32 lb). The pelage is usually long and may be blackish, brownish, grayish, or reddish. One species has a bright yellow face, belly, feet, and tail. The forelimbs and hindlimbs are nearly equal in size. The long, well-furred tail serves to help balance and steer the kangaroo as it makes long, flying leaps in the trees. The tail is also used to brace the animal when climbing. The 10 species live primarily in mountainous rainforest in Queensland (Australia) and New Guinea. They are very agile and are active both day and night. Food consists chiefly of tree leaves, ferns, and many kinds of fruits. Some species spend a considerable amount of time on the ground, whereas others spend almost all of their time in the trees. Clearing of rainforest and hunting are reducing their ranges. Habitat fragmentation may prevent dispersal and increase inbreeding. Three species are currently classified as endangered by the International Union for the Conservation of Nature and Natural Resources (IUCN), two are classified as vulnerable, and two as near-threatened.

Donald W. Linzey

Bibliography. B. Grzimek (ed.), *Grzimek's Encyclopedia of Mammals*, McGraw-Hill, 1990; D. Macdonald, *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. N. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Kaolinite

A common hydrous aluminum silicate mineral found in sediments, soils, hydrothermal deposits, and sedimentary rocks. It is a member of a group of clay

minerals called the kaolin group minerals, which include dickite, halloysite (7 Å and 10 Å), nacrite, ordered kaolinite, and disordered kaolinite. These minerals have a theoretical chemical composition of 39.8% alumina, 46.3% silica, and 13.9% water [$Al_2Si_2O_5(OH)_4$], and they generally do not deviate from this ideal composition. They are sheet silicates comprising a single silica tetrahedral layer joined to a single alumina octahedral layer. Although the kaolin group minerals are chemically the same, each is structurally unique as a result of how these layers are stacked on top of one another. Kaolinite is the most common kaolin group mineral and is an important industrial commodity used in ceramics, paper coating and filler, paint, plastics, fiberglass, catalysts, and other specialty applications. *See* CLAY MINERALS; SILICATE MINERALS.

The term "kaolin" is derived from the Chinese word *Kau-ling* (high ridge), which is the name given to a hill in the Jiang Xi Province where white clay was mined for porcelain as early as the seventh century. Deposits of relatively pure kaolinite are called kaolin, which leads to some confusion because the terms are often used interchangeably; however, kaolinite is the name for the mineral whereas kaolin applies specifically to a rock made up of more than one mineral, or to the mineral group. To avoid confusion, the term "kandite" was introduced as the mineral group name, but it has not been universally accepted and consequently kaolin group is preferred.

Structure. The kaolinite unit layer consists of one silica tetrahedral sheet and one alumina octahedral sheet stacked such that apical oxygens from the tetrahedral sheet replace the hydroxyls in the octahedral sheet (**Fig. 1**). This stacked unit of one tetrahedral and one octahedral sheet is known as a 1:1 layer. The tetrahedral sheet is made up of individual silica tetrahedra linked to neighboring tetrahedra through three shared oxygens occupying a basal plane called the siloxane surface. The fourth, or apical, oxygen points upward in a direction normal to the siloxane surface and forms part of the adjacent octahedral sheet. The octahedral sheet is made up of octahedrally coordinated Al^{3+} ions linked together at the octahedral edges. In kaolinite there is little to no cation substitution and therefore essentially no layer charge. In the cases where substitution does occur, the most common are the substitution of Al^{3+} for Si^{4+} in the tetrahedral sheet and Fe^{2+} or Fe^{3+} for Al^{3+} in the octahedral sheet. Kaolinite exhibits varying amounts and types of stacking disorder related to translations between adjacent layers. Ordered kaolinite has very little stacking disorder compared to disordered kaolinite.

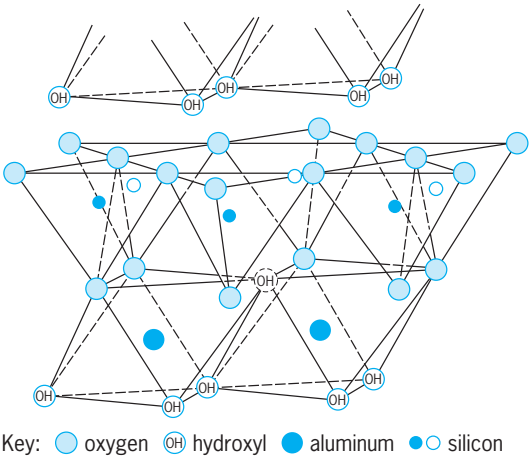

Fig. 1. Diagram of a kaolinite unit layer. (*After R. E. Grim, Applied Clay Mineralogy, McGraw-Hill, 1962*)

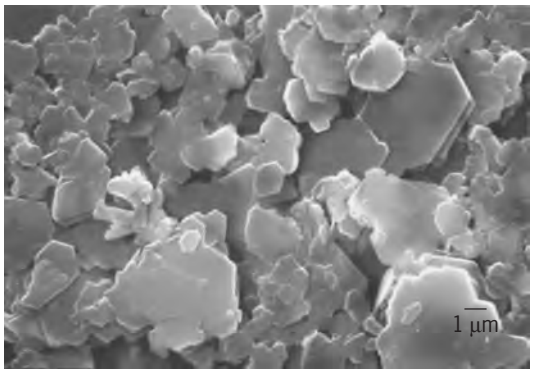

Fig. 2. Scanning electron micrograph showing pseudohexagonal kaolinite particles.

Properties. Kaolinite crystals are typically six-sided or pseudohexagonal in outline and may occur as single thin platelets or large accordionlike vermiform crystals (**Fig. 2**). Individual crystals range from approximately 0.1 to over 10 micrometers in diameter and from 15 to 50 nanometers in thickness. Kaolinite has a relatively high charge density and is therefore naturally hydrophilic, making it easily dispersible in water. It may be chemically modified so that it becomes hydrophobic and organophilic. It does not naturally swell in water unless treated with intercalation compounds such as hydrazine, urea, and ammonium acetate. Kaolinite is intrinsically white, has low plasticity, and has a refractive index of about 1.57. It has a low cation exchange capacity, and is chemically inert over a wide pH range except in some organic systems where it exhibits catalytic activity.

Genesis. Certain geochemical conditions favor the formation of kaolinite, which is most commonly found in surficial environments as a product of chemical weathering. Humid, temperate to tropical environments, where dilute, acid solutions are derived from the leaching of well-drained organic-rich soils, provide ideal conditions for kaolinite formation. The most common parent materials from which kaolinite is derived are feldspars and micas. These minerals, which contain the silica and alumina required for kaolinite formation, readily weather to kaolinite under the conditions listed above. Kaolinite also fills void spaces in sandstone, where it forms as a result of burial diagenesis. *See* DIAGENESIS.

Large accumulations of relatively pure kaolinite are classified as primary or secondary. Primary

deposits are those derived from the in-situ alteration of feldspar-rich igneous, metamorphic, or sedimentary rocks such as granite, gneiss, or arkose. Alteration results from hydrothermal action, surface weathering, ground-water interaction, or a combination of these effects. Secondary deposits, also called sedimentary, are those that have been eroded, transported, and deposited as beds or lenses and are associated with sedimentary rocks. Primary deposits are more common than secondary deposits, which require special geologic conditions for their formation and preservation. For example, the extensive and high-quality paper-coating-grade deposits mined in the southeastern United States were most likely formed as the result of four events: (1) intense chemical weathering of the crystalline source area, (2) reworking and transportation of detrital grains by streams, (3) deposition of kaolinite-rich detritus along a passively subsiding margin, and (4) repeated cycles of oxidative weathering leading to recrystallization and continued kaolinization. *See* WEATHERING PROCESSES.

Uses. Kaolin is mined and beneficiated for a variety of industrial uses ranging from paper coating, to pharmaceuticals, to a source of aluminum for aluminum compounds. It is a versatile mineral that is also used in ceramics, plastics, paints, rubber, adhesives, caulks, catalysts, inks, gaskets, cosmetics, animal feeds, fertilizers, and fiberglass. This diverse range of applications is the result of a unique combination of factors, including particle size, particle shape, color, chemical composition, rheological properties, chemical inertness, thermal properties, and relatively low cost. Functionally kaolin plays many different roles in these products, including adding strength, imparting thixotropy, enhancing chemical resistance and electrical resistivity, and improving opacity. *See* ORE DRESSING; RHEOLOGY.

Because kaolin deposits naturally contain ancillary minerals along with kaolinite, at least some processing is required. Two processing options are available: wet and dry. The quality of the raw material and the desired properties of the final product dictate which process is used. Wet processing results in a purer, more uniform product and may employ a number of processing steps, including particle size separation, selective flocculation, magnetic separation, delamination, flotation, reductive leaching, calcination, and surface treatment. *See* CERAMICS; PARTICULATES.

Jessica Elzea Kogel

Bibliography. D. D. Carr (ed.), *Industrial Minerals and Rocks*, 6th ed., Society of Mining, Metallurgy and Exploration, 1994; R. E. Grim, *Applied Clay Mineralogy*, McGraw-Hill, 1962; D. M. Moore and R. C. Reynolds, *X-ray Diffraction and the Identification and Analysis of Clay Minerals*, Oxford University Press, 1997; H. H. Murray, W. Bundy, and C. Harvey (eds.), Kaolin Genesis and Utilization, *Clay Minerals Soc. Spec. Publ.*, no. 1, 1993; A. C. D. Newman (ed.), Chemistry of Clays and Clay Minerals, *Mineralog. Soc. Monog.*, no. 6, 1987.

Kapitza resistance

A resistance to the flow of heat across the interface between liquid helium and a solid. A temperature difference is required to drive heat from a solid into liquid helium, or vice versa; the temperature discontinuity occurs right at the interface. The Kapitza resistance, discovered by P. L. Kapitza, is defined in the equation below, where T_S and T_H are the solid

$$R_K = \frac{T_S - T_H}{\dot{Q}/A}$$

and helium temperatures and $\dot{Q}/A$ is the heat flow per unit area across the interface. *See* CONDUCTION (HEAT).

In principle, the measured Kapitza resistance should be easily understood. In liquid helium and solids (such as copper), heat is carried by phonons, which are thermal-equilibrium sound waves with frequencies in the gigahertz to terahertz region. The acoustic impedance of helium and solids can differ by up to 1000 times, which means that the phonons mostly reflect at the boundary, like an echo from a cliff face. This property together with the fact that the number of phonons dies away very rapidly at low temperatures means that at about 1 K there are few phonons to carry heat and even fewer get across the interface. The prediction is that the Kapitza resistance at the interface is comparable to the thermal resistance of a 10-m (30-ft) length of copper with the same cross section. *See* ACOUSTIC IMPEDANCE; PHONON; QUANTUM ACOUSTICS.

The reality is that above 0.1 K and below 0.01 K (10 mK) more heat is driven by a temperature difference than is predicted (see **illus.**). Above 0.1 K

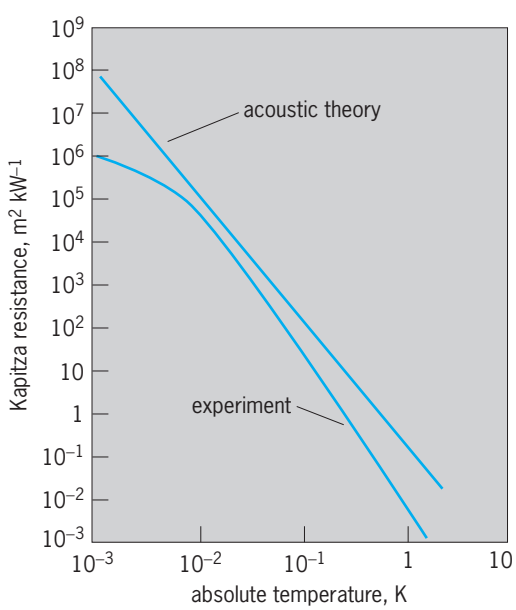

Plot of the Kapitza resistance between copper and liquid helium-3 (^{3}He) as a function of absolute temperature. The acoustic theory is based upon heat transfer by thermal-equilibrium sound waves. The line labeled experiment is representative of many experimental results.

this is now understood to be a result of imperfections such as defects and impurities at the interface, which scatter the phonons and allow greater transmission. *See* CRYSTAL DEFECTS.

The enormous interest in ultralow-temperature (below 10 mK) research generated by the invention of the dilution refrigerator and the discovery of superfluidity in liquid helium-3 (^{3}He) below 0.9 mK also regenerated interest in Kapitza resistance, because heat exchange between liquid helium and solids was important for both the dilution refrigerator and superfluidity research. An ingenious technique was invented to overcome the enormous Kapitza resistance at 1 mK: The solid is powdered, and the powder is packed and sintered to a spongelike structure to enhance the surface area. In this way a 1-cm^3 (0.06-in.^{3}) chamber can contain up to 1 m^2 (10 ft^2) of interface area between the solid and the liquid helium.

It was found that at 1 mK the Kapitza resistance is 100 times smaller than predicted by the phonon model. There have been two explanations for the anomaly, and probably both are relevant. One is that energy is transferred by magnetic coupling between the magnetic ^{3}He atoms and magnetic impurities in the solid or at the surface of the solid; the other is that the spongelike structure has quite different, and many more, phonons than a bulk solid and that these can transfer heat directly to the ^{3}He atoms.

Whatever its cause, this anomaly has had a major impact on ultralow-temperature physics. On the one hand, it has allowed liquid ^{3}He to be cooled by adiabatic demagnetization refrigerators to well below 1 mK (the record is about 10 μK) so that its superfluid properties can be studied over a wide range of temperature. On the other, it has allowed the development of the dilution refrigerator so that it can operate to below 2 mK. In the dilution refrigerator the fluid that refrigerates is liquid ^{3}He. The purpose, however, is to refrigerate a sample or an experiment, and that requires heat exchange to the cold liquid ^{3}He. Such refrigerators are used to investigate the properties of matter at ultralow temperature. *See* ADIABATIC DEMAGNETIZATION; LIQUID HELIUM; LOW-TEMPERATURE PHYSICS; SUPERFLUIDITY.

John P. Harrison

Bibliography. D. F. Brewer (ed.), *Progress in Low Temperature Physics*, vol. 12, 1989; E. T. Swartz and R. O. Pohl, Thermal boundary resistance, *Rev. Mod. Phys.*, 61:605–668, 1989.

Kapok tree

Also called the silk-cotton tree (*Ceiba pentandra*), a member of the bombax family (Bombacaceae). The tree has a bizarre growth habit and produces pods containing seeds covered with silky hairs called silk cotton (see **illus.**). It occurs in the American tropics, and has been introduced into Java, Philippine Islands, and Sri Lanka. The silk cotton is the commercial kapok used for stuffing cushions, mattresses, and pillows. Kapok has a low specific gravity and is impervious to water, making it excellent for filling life preservers. *See* MALVALES.

Pods and leaves of kapok tree (*Ceiba pentandra*).

Perry D. Strausbaugh; Earl L. Core

Kármán vortex street

A double row of line vortices in a fluid. Under certain conditions, a Kármán street is shed in the wake of bluff cylindrical bodies when the relative fluid velocity is perpendicular to the generators of the cylinder (**Fig. 1**). This shedding of eddies occurs first from one side of the body and then from the other, an unusual phenomenon because the oncoming flow may be perfectly steady. Vortex streets can often be seen, for example, in rivers downstream of the columns supporting a bridge. The streets have been studied most completely for circular cylinders at low subsonic flow speeds. Regular, perfectly periodic vortex streets occur in the range of Reynolds number (Re) 50–300, based on cylinder diameter. Above a Reynolds number of 300, a degree of randomness begins to occur in the shedding due to secondary instabilities, which becomes progressively greater as Reynolds number increases, until finally the wake is completely turbulent. The highest Reynolds number at which some slight periodicity is still present in the turbulent wake is about 10^6. *See* REYNOLDS NUMBER.

Vortex streets and their effects are common. For example, they can be created by steady winds blowing past smokestacks, transmission lines, bridges, missiles about to be launched vertically, and pipelines aboveground in the desert. They are also related to the phenomenon that generates cloud

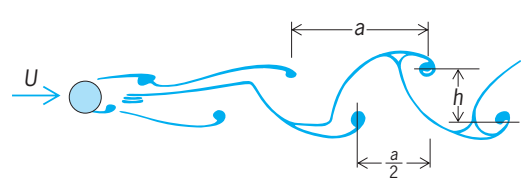

Fig. 1. Diagram of a Kármán vortex street. The streamwise spacing of the vortices, *h*, and the spacing normal to it, *a*, are shown.

Fig. 2. Vortex patterns downstream of Guadelupe Island, imaged by the *SeaWIFS* satellite.

patterns indicating the alternate shedding of vortices downstream of an island, as captured by satellite images (**Fig. 2**). Although the island only forms a surface obstruction to the wind, geophysical effects (Coriolis effects) generate a stagnant column of air in the atmosphere above it that creates an obstruction analogous to a cylinder. *See* CORIOLIS ACCELERATION.

The alternating vortex shedding that coincides with a well-formed vortex street produces oscillation in lateral forces acting on the body. If the vortex shedding frequency is near a natural vibration frequency of the body, the resonant response may cause large-amplitude vibration and in some cases structural damage. The Aeolian tones, or singing of wires in a wind, is an example of forced oscillation due to formation of a vortex street. Such forces can impose on structures unwanted vibrations, often leading to serious damage. An example of unwanted vibrations which led to serious damage is the collapse of the Tacoma Narrows (Washington) suspension bridge in 1940. This large structure made of thousands of tons of steel collapsed after a few hours of ever-growing oscillations induced by a steady wind (**Fig. 3**).

Fig. 3. Tacoma Narrows Bridge in the middle of its wild oscillations, a little before it collapsed.

T. von Kármán showed that an idealized, infinitely long vortex street is stable to small disturbances if the spacing of the vortices is such that $h/a = 0.281$; actual spacings are close to this value. A complete and satisfying explanation of the formation of vortex streets, however, has not yet been given. For $10^3 < \mathrm{Re} < 10^5$, the shedding frequency f for a circular cylinder in low subsonic speed flow is given closely by $fd/U = .21$, where d is the cylinder diameter and U is stream speed; a/d is approximately 5. This means that for a cylinder of a certain diameter immersed in a fluid, the frequency of vortex shedding is proportional to the velocity of the oncoming stream and inversely proportional to the diameter. A. Roshko discovered a spanwise periodicity of vortex shedding on a circular cylinder at $\mathrm{Re} = 80$ of about 18 diameters; thus, it appears that the line vortices are not quite parallel to the cylinder axis. *See* FLUID-FLOW PRINCIPLES; VORTEX.

Arthur E. Bryson, Jr.; Demetri P. Telionis

Bibliography. W. R. Debler, *Fluid Mechanics Fundamentals*, 1990; J. A. Liggett, *Fluid Mechanics*, 1994; R. L. Panton, *Incompressible Flow*, Wiley Interscience, 2d ed., 1996; F. M. White, *Fluid Mechanics*, McGraw-Hill, 6th ed., 2006.

Karst topography

Distinctive associations of third-order, erosional landforms indented into second-order structural forms such as plains and plateaus. They are produced by aqueous dissolution, either acting alone or in conjunction with (and as the trigger for) other erosion processes. Karst is largely restricted to the most soluble rocks, which are salt (360,000 mg/liter), gypsum and anhydrite (2400 mg/liter), and limestone and dolostone (30–400 mg/liter). The numbers denote effective solubility in meteoric waters under standard conditions. Being so soluble, salt is seen only in the driest places (for example, Death Valley, California), where it displays intensive dissolution topography. Surface gypsum karst is also comparatively rare in humid regions, but its dissolution at depth may produce widespread collapse and subsidence landforms (covered karst); it is often well developed in arid and semiarid areas such as the Pecos River valley of New Mexico. Pure limestones ($CaCO_3$) are the principal karst rocks, hosting the greatest extent and variety of features. Dolostone [the double carbonate mineral, $CaMg\ (CO_3)_2$] is significantly less soluble: the type and scale of dissolutional forms developed in it is normally less than in limestone, giving rise to topographies that are transitional to fluvial (stream-derived) landscapes. *See* DOLOMITE ROCK; GYPSUM; LIMESTONE.

Karst rocks outcrop over ~12% of the Earth's continental and island surfaces, but the distinctive topography is limited to about 8%. However, many well-known places are entirely karstic (the Bahamas and the majority of the Greek islands, for example), and 20–25% of the world's population relies on ground water circulating to some extent in karst rocks, so that karst studies are of considerable practical

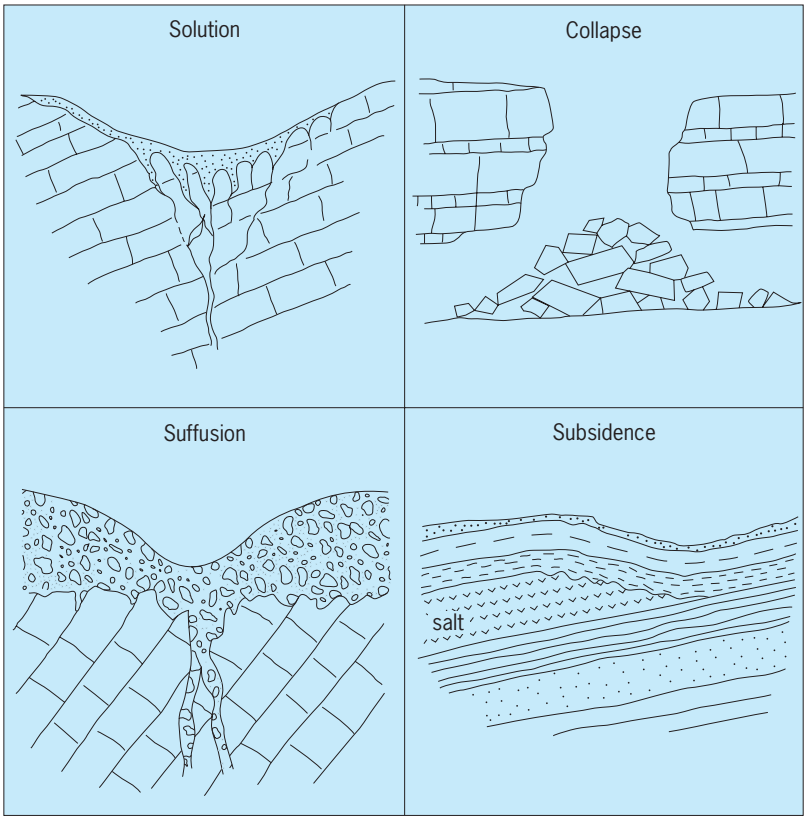

Fig. 1. Principal genetic types of sinkholes. (*After D. C. Ford and P. W. Williams, Karst Geomorphology and Hydrology, 1989; copyright Chapman and Hall; used with kind permission from Kluwer Academic Publishers*)

importance. *See* GROUND-WATER HYDROLOGY.

The essence of the karst dynamic system is that meteoric water (rain or snow) is routed underground, because the rocks are soluble, rather than flowing off in surface river channels. It follows that dissolutional caves develop in fracture systems, resurging as springs at the margins of the soluble rocks or in the lowest places. A consequence is that most karst topography is "swallowing topography," assemblages of landforms created to deliver meteoric water down to the caves. The word "karst" derives from a Slavic regional term, *kras*, meaning stony ground—stony because the former soil cover in western Slovenia and neighboring areas was lost into caves as a consequence of deforestation and overgrazing.

Karst landforms develop at small, intermediate, and large scales.

Karren. This is the general name given to small-scale forms—varieties of dissolutional pits, grooves, and runnels. Individuals are rarely greater than 10 m (30 ft) in length or depth, but assemblages of them can cover hundreds of square kilometers. On bare rock, karren display sharp edges; circular pits or runnels extending downslope predominate. Beneath soil, edges are rounded and forms more varied and intricate.

The largest karren are clefts following major joint sets. Where regular, shallow, and exposed, they create the appearance of a laid paving-limestone pavement, which is common in glaciated regions. Where deeper, they may become ruiniform (corridor karst) or, if the blocks between the clefts are tapered by runnelling dissolution, pinnacled. Dense clusters of pinnacles up to 20 m (65 ft) or more are the most spectacular karren, termed *shilin* (stone forests) in Yunnan, China, where the largest assemblages occur.

Particularly dense and complex pitting patterns occur along seashores. The finest development is on tropical coasts where formation of deep pits and sharp residual pinnacles is enhanced by biological action—phytokarst. Sharp dissolution notching may occur at the high-tide line.

Sinkholes. Also known as dolines or closed depressions, sinkholes are the diagnostic karst (and pseudokarst) landform. They range from shallow, bowllike forms, through steep-sided funnels, to vertical-walled cylinders (**Fig. 1**). Asymmetry is common. Individual sinkholes range from about 1 to 1000 m (3 to 3300 ft) in diameter and are up to 300 m (1000 ft) deep. Many may become partly or largely merged.

Sinkholes can develop entirely by dissolution acting downward into receptor caves, by collapse of cave roofs upward, by washing of unconsolidated fine sediments such as soil cover into underlying cavities, or by slow subsidence. Most karst sinkholes are produced by a combination of the first two or three processes.

On plains, sinkholes may be scattered like shell pits, or packed together with overlapping perimeters (polygonal karst, found in the Sinkhole Plain of Kentucky). Densities greater than 1000/km^2 are known. In mountainous country, deep polygonal karst can create a spectacular "egg carton" topography of steep-sided sinkholes (cockpits) between conical hills, termed *fengcong* (peak cluster) in southern China where the finest examples are found.

Dry valleys and gorges. These large-scale features are carved by normal rivers, but progressively lose their water underground (via sinkholes) as the floors become entrenched into karst strata. Many gradations exist, from valleys that dry up only during dry seasons (initial stage) to those that are without any surface channel flow even in the greatest flood periods (paleo-valleys). They are found in most plateau and mountain karst terrains and are greatest where river water can collect on insoluble rocks before penetrating the karst (allogenic rivers).

Poljes. A Serbo-Croatian term for a field, polje is the generic name adopted for the largest individual karst landform. Ideally, this is a topographically closed depression with a floor of alluvium masking an underlying limestone floor beveled flat by planar corrosion. Rivers may cross the alluvium, flowing from upstream springs to downstream sinks. The polje is flooded seasonally, when dissolutional undercutting takes place around the perimeter of enclosing limestone hills, and further alluvium is added to the insoluble shield protecting the floor. In this manner, the polje expands.

Ideal poljes up to a few kilometers in length are found in many karst regions. Chains of 10 km (6 mi) poljes and individuals up to 60 km (36 mi) in length occur in former Yugoslavia, where they are partly of tectonic origin—karstic adaptations

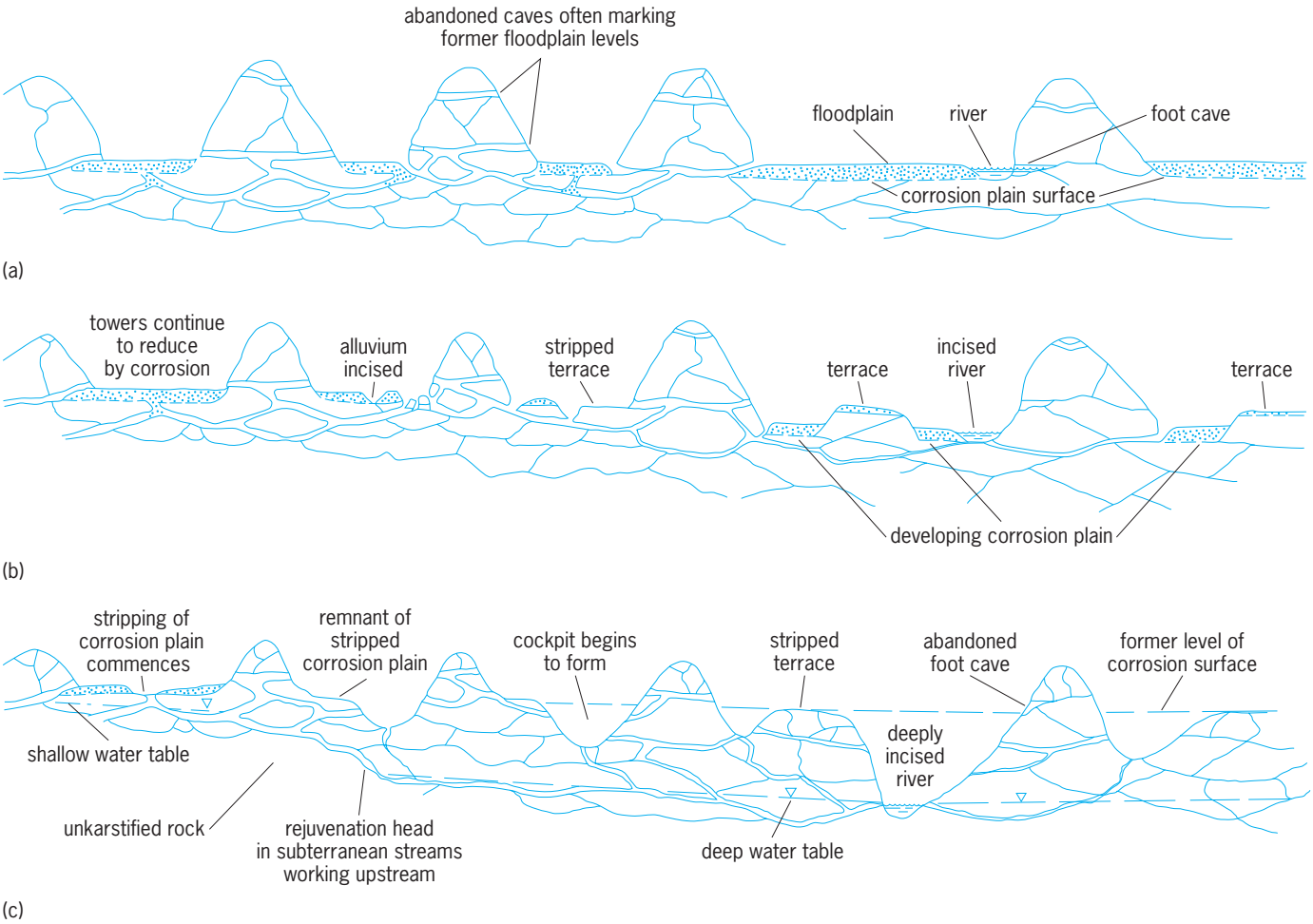

Fig. 2. Development of a multistage, polygenetic karst landscape of towers, cones, relict caves, corrosion plains, and terraces. (*After D. C. Ford and P. W. Williams, Karst Geomorphology and Hydrology, 1989; copyright Chapman and Hall; used with kind permission from Kluwer Academic Publishers*)

of downfaulted blocks. The Yugoslav examples are preeminent and are known as classical karst; they contain most of the cultivatable land in many areas, the intervening hills being unusable because of the great densities of karren and sinkholes. *See* FAULT AND FAULT STRUCTURES.

Karst plains and towers. These are the end stage of karst topographic development in some regions, produced by long-sustained dissolution or by tectonic lowering. The plains are of alluvium, with residual hills (unconsumed intersinkhole limestone) protruding through. Where strata are massively bedded and the hills are vigorously undercut by seasonal floods or allogenic rivers, they may be steepened into vertical towers. Tower karst (*fenglin* in Chinese) is the most accentuated karst topography, with individuals rising to as much as 500 m (1650 ft). The greatest tower karst development is in Guizhou and Guangxi provinces, China. Tower landscapes partly drowned by postglacial rising seas in Haiphong Bay, Vietnam, and Phuket, Thailand, are also famous tourist sights. Where strata are thinner and undercutting is ineffective, the hills tend to be conical. Plains with cone karst are common in Puerto Rico and Cuba.

Tectonic uplift may induce incision by trunk rivers, producing complex multistage karst landscapes of towers, cones, and corrosion plains (**Fig. 2**). *See* CAVE; GEOMORPHOLOGY; WEATHERING PROCESSES.

Derek Ford

Bibliography. W. Dreybrodt, *Processes in Karst Systems*, Springer-Verlag, Berlin, 1988; D. C. Ford and P. W. Williams, *Karst Geomorphology and Hydrology*, Chapman and Hall, London, 1989; J. N. Jennings, *Karst Geomorphology*, Blackwell, Oxford, 1985; P. T. Milanovic, *Karst Hydrogeology*, Water Resources Publications, Littleton, CO, 1981; W. B. White, *Geomorphology and Hydrology of Karst Terrains*, Oxford University Press, 1988; D. Yuan, *Karst of China*, Geological Publishing House, Beijing, 1991.

Kelvin bridge

A specialized version of the Wheatstone bridge network designed to eliminate, or greatly reduce, the effect of lead and contact resistance and thus permit accurate measurement of low resistance. The circuit shown in **illus.** *a* accomplishes this by effectively placing relatively high-resistance-ratio arms in series with the potential leads and contacts of the low-resistance standards and the unknown

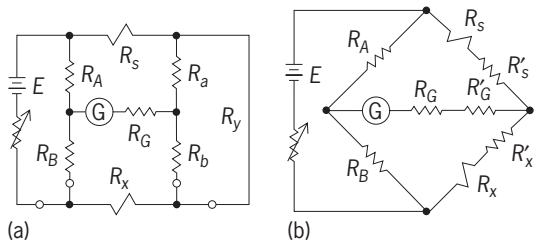

Kelvin bridge. (*a*) Actual circuit. (*b*) Equivalent Wheatstone bridge circuit.

resistance. In this circuit R_A and R_B are the main ratio resistors, R_a and R_b the auxiliary ratio, R_x the unknown, R_s the standard, and R_y a heavy copper yoke of low resistance connected between the unknown and standard resistors.

If the unwanted potential drop along the conductor linking R_s and R_x is divided (by R_a/R_b) in the same ratio (R_s/R_x) as that of the potential drops across R_s and R_x, the zero reading of the galvanometer (G) is unaffected by its actual value R_y.

Therefore, if R_a/R_b is adjustable (for example, by adjusting R_a) so that a null indication of G is obtained no matter what the value of R_y, the basic balance equation for the Wheatstone bridge (1) holds.

$$R_x = \frac{R_B}{R_A} R_s \tag{1}$$

A practical way of attaining this condition is to (1) adjust R_A/R_B with the lowest possible value of R_y (for example, a low-resistance copper link), and then (2) adjust R_a/R_b with R_y infinite (for example, with the link open-circuited). Steps 1 and 2 are carried out alternately as necessary until G is nulled for both conditions without significant further adjustments. The final adjustment of R_A/R_B obtains the bridge balance condition, Eq. (1).

A formal network analysis can be carried out by applying a delta-wye transformation to the network consisting of R_a, R_b, and R_y. The equivalent Wheatstone bridge network shown in illus. *b* is obtained, where Eqs. (2) hold. By an analysis similar to that for

$$\begin{aligned} R'_s &= \frac{R_y R_a}{R_y + R_a + R_b} \\ R'_x &= \frac{R_y R_b}{R_y + R_a + R_b} \\ R'_G &= \frac{R_a R_b}{R_y + R_a + R_b} \end{aligned} \tag{2}$$

the Wheatstone bridge, it can be shown that for a balanced bridge Eq. (3) holds. If Eq. (4) is valid, the

$$R_x = \frac{R_B}{R_A} R_s + R_y \left(\frac{R_b}{R_a + R_b + R_y} \right) \left(\frac{R_B}{R_A} - \frac{R_b}{R_a} \right) \tag{3}$$

$$\frac{R_B}{R_A} = \frac{R_b}{R_a} \tag{4}$$

second term of Eq. (3) is zero, the measurement is independent of R_y, and Eq. (1) is obtained.

As with the Wheatstone bridge, the Kelvin bridge for routine engineering measurements is constructed using both adjustable ratio arms and adjustable standards. However, the ratio is usually continuously adjustable over a short span, and the standard is adjustable in appropriate steps to cover the required range. *See* WHEATSTONE BRIDGE.

Sensitivity. The Kelvin bridge sensitivity can be calculated similarly to the Wheatstone bridge. The open-circuit, unbalance voltage appearing at the detector terminals may be expressed, to a close degree of approximation, as in Eq. (5).

$$e = E \frac{r}{(r+1)^2} \frac{\Delta R_x}{R_x + R_y \left(\dfrac{r}{r+1} \right)} \tag{5}$$

The unbalance detector current for a closed detector circuit may be expressed as in Eq. (6).

$$I_G = \frac{E \left(\dfrac{\Delta R_x}{R_x} \right)}{\dfrac{R_G}{r/(r+1)^2} + R_A + R_B + R_a + R_b} \tag{6}$$

The Kelvin bridge requires a power supply capable of delivering relatively large currents during the time a measurement is being made. The total voltage applied to the bridge is usually limited by the power dissipation capabilities of the standard and unknown resistors.

Errors. Kelvin bridge resistance-measurement errors are caused by the same factors as for the Wheatstone bridge. However, additional sources of error, as implied by the second term of Eq. (2), must be evaluated since these factors will seldom be reduced to zero. For minimum error the yoke resistance should be made as low as possible by physically placing the commonly connected current terminals of the unknown and standard as close together as possible and connecting with a low-resistance lead.

The ratio resistors each include not only the resistance of the resistors but also that of the interconnecting wiring and external leads and the contact resistance of the potential circuit contacts. The external leads are most likely to cause errors, and they should therefore be of the same resistance so that Eq. (3) will be fulfilled as nearly as possible. In addition, they should be relatively short, since the addition of a large resistance (long leads) will introduce an error in the calibrated ratio R_B/R_A. For precise measurements, trimmer adjustments are required in the ratio-arm circuits and provision is made to connect the bridge resistors into two different Wheatstone bridge configurations. By successively balancing first the Kelvin network and then each of the Wheatstone networks, these additive errors are virtually eliminated. *See* BRIDGE CIRCUIT; RESISTANCE MEASUREMENT. Charles E. Applegate; Bryan P. Kibble

Kelvin's circulation theorem

A theorem useful for ideal flows that have negligible viscous forces and a nearly constant density (incompressible flows). Such flows are potential flows and

are applicable to the aerodynamics of wings. One characteristic of potential flows is a lack of vorticity. Vorticity characterizes the solid-body-like rotation of the fluid particles as they translate and deform. If a small element of fluid in a flow were instantaneously frozen, vorticity would indicate its spinning motion. Potential flows have zero vorticity. *See* POTENTIAL FLOW; VORTICITY; WING.

Circulation is a number associated with the spinning motion of a large region of a flow. Consider a plane flow such as the flow around a wing that spans a very big wind tunnel (**Fig. 1**). Imagine that you choose a certain closed path to walk around in a counterclockwise direction. As you walk at a uniform pace, measure and record the wind velocity component that is in the same direction as your path. Ignore the side component. Wind velocity in the same direction as your path is positive, while wind velocity against your path is negative. When you return to the starting point, the average velocity number indicates the circulation Γ for that circuit. Actually, the circulation Γ is the average wind component times the length of the path. Mathematically it can be shown that the circulation is the area integral of the vorticity over the region bounded by the closed-path curve, hence the interpretation as the global spinning character of a large region. In potential flows, the average velocity along the path is always zero for any path within the potential flow region. The average flow against the path is the same as the average flow with the path.

Kelvin's circulation theorem states that the circulation for a circuit of material particles is constant with time in an ideal flow. That is, the average rate of spinning of a piece of fluid is constant as it moves with the flow. This is a direct result of the lack of friction.

As an example, consider a plane vortex flow. This is a swirling potential flow of circular streamlines with zero vorticity everywhere except for the central core, where the vorticity is very high because of viscous forces. Any closed circuit that excludes the core has a circulation of zero. Any closed circuit that includes the core has the same finite circulation. The circulation is a measure of the strength of the vortex.

Consider a wing, helicopter blade, or fan blade starting from rest. The original and current positions are shown in **Fig. 2**. If a fluid is at rest, the circulation is zero for any circuit. A region of potential flow starting from rest will continue to have a circulation of zero for all time. As the wing starts moving, viscous forces are important in a boundary layer near the surface and produce a thin region of vorticity. The boundary layer fluid is not an ideal potential flow. This region is swept off and forms the core of a starting vortex that remains near the starting position. A thin viscous wake connects the starting vortex to the viscous boundary layer region that continues with the wing. A circuit around the wing—outside the boundary layer, the wake, and the starting vortex core—has zero circulation. However, a circuit cutting the viscous wake and including the vortex has a circulation of one sense; and a circuit cutting the wake and including the wing has a circulation

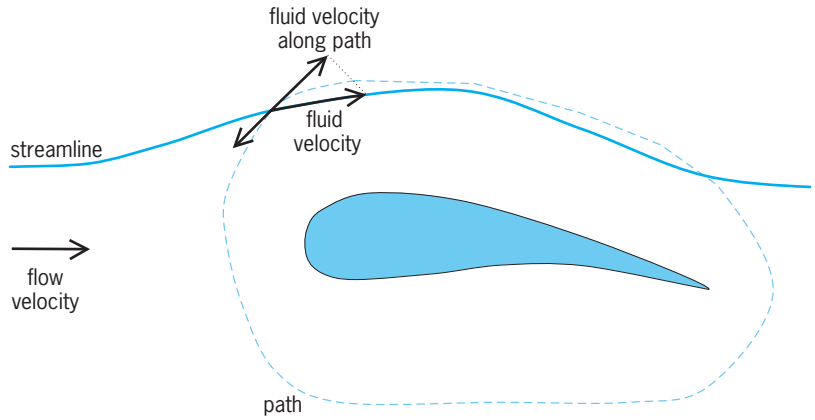

Fig. 1. Cross-section view of a wing in a wind tunnel depicting a circulation path.

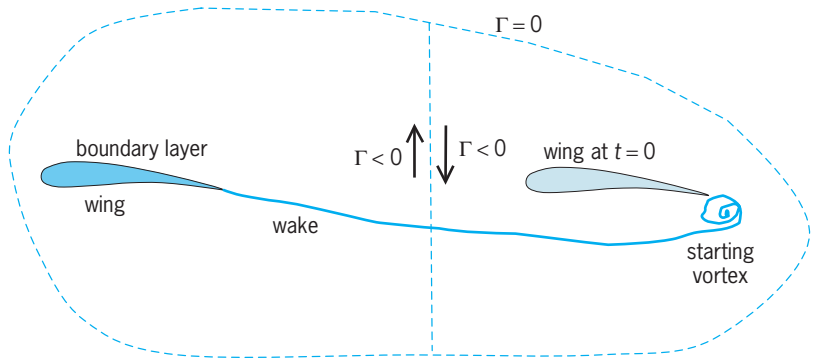

Fig. 2. Cross-section view of a wing starting from rest in a still fluid and leaving a starting vortex.

of equal value but opposite sense. Kelvin's theorem does not apply to the vortex core, wake, or boundary layer particles because they have experienced viscous forces. *See* BOUNDARY LAYER FLOW.

The nonzero circulation actually is directly related to the amount of lift produced by the wing. Circulation implies that the velocities above the wing are higher than those below the wing. The increase in velocity in potential flow is accompanied by a decrease in pressure. Thus, the lower pressures on the top result in a lift force. *See* BERNOULLI'S THEOREM.

A wing has a finite span, and so the disturbed flow does not extend to infinity in the spanwise direction (**Fig. 3**). A circuit path outside the wing/wake/

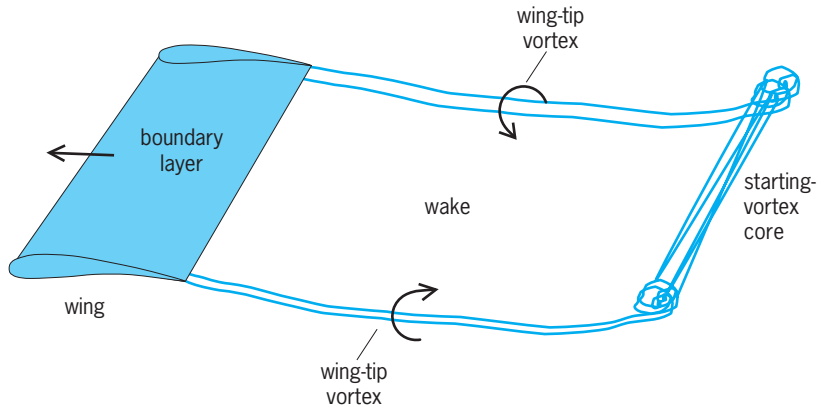

Fig. 3. Wing of finite span showing tip vortices connecting the wing circulation to the starting vortex.

starting vortex has zero circulation, because no viscous forces have touched those particles. However, the starting vortex and the wing are connected by a tip vortex so that a circuit around the tip vortex and through the wake has a finite circulation.

Ronald L. Panton

Bibliography. I. G. Currie, *Fundamental Mechanics of Fluids*, 3d ed., 2003; S. I. Green (ed.), *Fluid Vorticies*, 1995; H. J. Lugt, *Vortex Flows in Nature and Technology*, 1983; R. L. Panton, *Incompressible Flow*, 3d ed., 2005; P. G. Saffman, *Vortex Dynamics*, 1992.

Kelvin's minimum-energy theorem

A theorem in fluid dynamics that pertains to the kinetic energy of an ideal fluid (that is, inviscid, incompressible, and irrotational) and provides uniqueness statements concerning the solution of potential-flow problems.

Much of the value of potential theory lies in the wealth of theorems and corollaries which have been derived in the various fields by using this theory as their basis. Given that the velocity field can be represented by harmonic potential functions, Green's first identity leads to a result known as the energy theorem. This theorem states that the energy of the fluid in a region bounded by a surface S is equal to the product of half the fluid density and an integral over S; the integrand of this integral is the product of the velocity potential and its partial derivative in the direction of the surface normal. (Since the fluid velocity is the gradient of the velocity potential, the normal derivative of this potential equals the normal component of the fluid velocity at the boundary S, and therefore determines the normal motion of this boundary.) Using the energy theorem, Lord Kelvin derived a remarkable theorem which demonstrated that the irrotational motion of a liquid occupying a simply connected region has less kinetic energy than any other motion consistent with the same normal motion of the boundary S. *See* GREEN'S THEOREM; POTENTIALS.

The implications of the minimum-energy of irrotational motion are: (1) irrotational motion is impossible in a simply connected region bounded by fixed walls since in this case the normal derivative of the velocity potential vanishes at all points on the boundary, and therefore, according to the energy theorem, the kinetic energy is zero, or the system is at rest; (2) irrotational motion is impossible in a fluid in which the velocity at infinity vanishes if the internal boundaries are also at rest; (3) if the velocity at infinity vanishes, then the irrotational motion due to prescribed motion of an internal boundary is unique; and (4) if a fluid is in motion with uniform velocity at infinity, then the irrotational motion due to prescribed motion of an internal boundary is unique. *See* FLUID-FLOW PRINCIPLES; LAPLACE'S IRROTATIONAL MOTION.

Eric Paterson; Fred Stern

Bibliography. H. Lamb, *Hydrodynamics*, 6th ed., 1993; H. Rouse, *Advanced Mechanics of Fluids*, 1976.

Kenaf

An annual, short-day, herbaceous plant (*Hibiscus cannabinus*) of the Malvaceae family, cultivated for its stem fibers. The genus *Hibiscus* has approximately 200 species, of which one, roselle (*H. sabdariffa* var. *altissima*), also is occasionally referred to as kenaf.

Kenaf usually grows upright up to 15 ft (5 m) in height, and is cylindrical and either branched or unbranched. The stem is composed of two fibers, bast and core. The bast fibers, located in the bark, are long compared to the core fibers, produced in the stem's interior. The leaves are either entirely heart-shaped or display radiating lobes. Flowers are typically yellow with deep red centers. Wild forms of kenaf are found in east and central Africa, where for several centuries kenaf has been used for both fiber and food. Selection and breeding work have developed varieties with higher fiber yields, improved disease resistance, and reduced branching.

Commercial uses. Kenaf is grown commercially for fiber production in many areas of the world, with the largest producer being China. In the United States, kenaf production is located in Texas, Louisiana, Mississippi, and California. Additional uses of kenaf include the manufacturing of various paper and pulp products, and as poultry litter, potting soil amendments, chemical- and oil-spill absorbents, animal and horse bedding, and packing materials. Other potential uses include the manufacturing of filters and particle and insulation boards. Kenaf leaves, which contain 20–30% crude protein, also may have potential as a livestock feed source.

Cultivation. Kenaf can adapt to a wide range of climates and soils. However, because kenaf cannot tolerate frost, planting should not occur at temperatures below 32°F (0°C). Optimum yields of kenaf are generally obtained on well-drained soils with fertility levels necessary to meet the nutritional requirements. In addition, most varieties of kenaf are photoperiod-sensitive, and vegetative growth increases until the daylight period becomes less than 12 h 30 min. Flowering is then initiated and the vegetative growth rate declines. Early planting maximizes yields by increasing the growing season. Several photoperiod-insensitive varieties have been developed for high-latitude and equatorial areas. Kenaf is propagated by seed and must be replanted annually. Dense plant populations generally produce greater total and bast fiber yields, reduce weed populations, and improve harvesting efficiency.

Harvesting in the United States is accomplished with modified sugarcane or forage chopping equipment. After cutting and drying the stalks, they are processed at fiber separation mills to meet the specific markets. Since the 1970s, there has been an ongoing attempt to focus on kenaf as an alternative

nonwood fiber source for various industrial applications.

Diseases. Kenaf serves as a host to several diseases and nematodes. The southern root-knot nematode (*Meloidogyne incognita*) is probably the greatest constraint to kenaf production throughout the world. Breeding efforts to identify a high level of resistance have not yet been successful. However, several strains which possess significant field tolerance to the root-knot nematode/soil-borne fungi complex have been identified. Crop rotations with nonhost species have provided the best method of reducing yield losses caused by both nematodes and soil-borne pathogens. Breeding and selection work has resulted in varieties and breeding strains with resistance to some of the major diseases of kenaf, including anthracnose (caused by *Collectotrichum gloesporiodes*) and charcoal rot (caused by *Macrophomina phaseolina*). *See* MALVALES; PLANT PATHOLOGY. Charles G. Cook

Bibliography. J. M. Dempsey, Kenaf, *Fiber Crops*, 1975; G. B. Killinger, Kenaf (*Hibiscus cannabinus* L.), A multi-use crop, *Agron. J.*, 61:734–736, 1969; C. S. Taylor and D. E. Kugler, Kenaf: Annual fiber crop products generate a growing response from industry, in *U.S. Department of Agriculture Yearbook*, pt. III: *New Crops, New Uses, New Markets*, 1992; U.S. Department of Agriculture, Agricultural Research Service, Kenaf paper: A forest saving alternative, *Agr. Res.*, 36:6–8, 1988; F. D. Wilson and M. Y. Menzel, Kenaf (*Hibiscus cannabinus*), roselle (*Hibiscus sabdariffa*), *Econ. Bot.*, 18:80–91, 1964.

Kennel cough

A common, highly contagious respiratory disease of dogs, also known as canine infectious tracheobronchitis. Several different bacteria and viruses are usually associated with the disease. Symptoms are generally mild but may vary widely depending on the agent, the host, and environmental factors. The main feature of the disease is sudden onset of violent coughing in dogs that had a recent exposure to other (infected) dogs. The disease is easily transmitted between dogs by droplets in the air or direct contact and often occurs as outbreaks or as a seasonal infection. Most dogs completely recover within 2 weeks; however, chronic and severe forms of the disease sometimes occur.

Infectious agent. Infectious agents commonly associated with the disease are the bacterium *Bordetella bronchiseptica* and canine parainfluenza virus. These agents are widespread in dogs; in fact, few dogs can escape exposure to these agents. *Bordetella bronchiseptica* can, in addition, be found in many other species of animal such as cats, pigs, rats, guinea pigs, rabbits, horses, and rarely humans. Each agent is capable of producing a mild form of the disease; however, most single-agent infections probably show no symptoms of disease. Several species of mycoplasmas have been isolated from the lower respiratory tract of dogs with kennel cough, but always in combination with another agent (for example, bordetella or canine parainfluenza virus). These mycoplasmas are normally found in the upper respiratory tract of healthy dogs. *See* BORDETELLA; MYCOPLASMAS.

Pathogenesis. Replication of canine parainfluenza virus, bordetella, and mycoplasma is usually limited to the epithelium lining the airway. Canine parainfluenza virus ultimately destroys the epithelial cells, while bordetella and mycoplasma remain attached to the outer surface of cells and cause more subtle changes in cell function such as reduced ciliary movement, attraction of inflammatory cells to the tissue, and alteration of phagocytic activities.

Epidemiology. Close contact with other dogs is usually required for transmission of kennel cough. Each of the viral agents of the disease, and possibly some of the mycoplasmas, have host ranges restricted to dogs. Because infections that show no symptoms of disease are also common, it is sometimes difficult to determine the source of the infection. Canine parainfluenza virus and *B. bronchiseptica* do not usually persist longer than a few weeks or a few months, respectively, in an individual dog. Dogs that have recovered from infection develop short-term immunity to reinfection. Therefore, infectious tracheobronchitis is acquired, and the pool of infectious agents is maintained in nature, by situations that bring infected dogs into close contact with susceptible dogs, for example, in closed breeding colonies, pet shops, boarding kennels, and show arenas.

Diagnosis. The most common symptom of kennel cough is often described as a dry hacking or honking cough that may be accompanied by significant retching, as if a bone or other foreign object were lodged in the upper airway. Occasionally, there are discharges from the nose and eyes. Fever is not usually present, and dogs recover spontaneously in a few days to a few weeks. Kennel cough does not appear to have age, breed, geographic, or seasonal limitations; however, other factors, such as air quality, temperature change, stress, and other underlying diseases, may contribute to the onset and severity of disease. Under these circumstances, agents of the disease, particularly bordetella and mycoplasma, may be associated with protracted (chronic) cough or life-threatening pneumonia.

Treatment. Treatment of kennel cough is often unwarranted. However, when indicated, antitussives, bronchodilators, and corticosteroids are used to relieve coughing, and antimicrobials are used to treat or prevent bronchopneumonia. Apart from maintaining dogs in good general health, free of other underlying diseases, the risk of acquiring kennel cough can be reduced by minimizing exposure to infectious agents. This includes isolating infected dogs, disinfecting contaminated areas and objects, avoiding overcrowded conditions, increasing ventilation, and isolating dogs entering a new environment. Regular vaccinations for immunization against canine parainfluenza virus and *B. bronchiseptica* is also important in controlling the disease. David A. Bemis

Bibliography. C. E. Greene (ed.), *Infectious Diseases of the Dog and Cat*, 2d ed., 1998.

Kentucky coffee tree

A large, strikingly distinct tree, *Gymnocladus dioica*, which usually grows 80–90 ft (24–27 m) high, but sometimes attains a height of 110 ft (33 m) and a diameter of 5 ft (1.5 m). The species name, *dioica*, means that the tree is dioecious; that is, male and female flowers are on different individuals. It grows from eastern Nebraska, Kansas, and Oklahoma to southern Ontario, western New York, and Pennsylvania, and southwestward to Louisiana. It can readily be recognized when in fruit by its leguminous pods containing hard, heavy, red-brown seeds, which were used by early settlers as a substitute for coffee; hence the name coffee tree. The branches are stout and thick, and the bark has thin, twisted ridges. The leaves are twice pinnate, and the winter buds, sunken in the bark, are superposed, two or three together (see **illus.**). Never a common tree,

Kentucky coffee tree (*Gymnocladus dioica*). (*a*) Parts of branch showing leaf scars and winter buds. (*b*) Branch.

it is sometimes cultivated in parks and gardens of the eastern United States and northern and central Europe. It is sometimes used as a street tree. The wood is hard and reddish, and is used for construction. It is durable in contact with the soil and is also used for railroad ties and fenceposts. *See* FOREST AND FORESTRY; ROSALES; TREE.

Arthur H. Graves; Kenneth P. Davis

Kepler's equation

The mathematical relationship between two different systems of angular measurement of the position of a body in an ellipse; specifically the relation between the mean anomaly M and eccentric anomaly θ', $M = \theta' - e \sin \theta'$, where e is the eccentricity of the ellipse. *See* ELLIPSE.

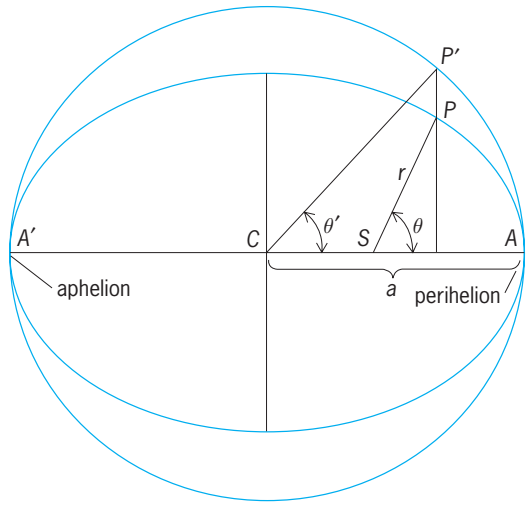

True and eccentric anomaly for a planet *P*.

The true position of a planet P in an elliptical orbit can be represented by the angle θ (true anomaly) measured at the focus S between the line directed to the planet and the line directed to the perihelion A (see **illus.**). The radius vector r from the focus to the planet can be expressed by $r = a(1 - e^2)/(1 + e \cos \theta)$, where a is the semimajor axis. The radius vector may also be expressed in terms of the eccentric anomaly θ' by $r = a(1 - e \cos \theta')$, where θ' is the angle at the center of the ellipse measured from perihelion along the circumscribed circle of radius a to the point whose projection perpendicular to the major axis passes through the planet. The true anomaly θ may be expressed in terms of the eccentric anomaly θ' by $\tan (\theta/2 = [(1 + e)/(1 - e)]_{1/2} \tan (\theta'/2)$. In actual practice, however, it is more convenient to describe the angular position of a planet in an elliptical orbit at any time t by means of its average angular velocity n (called mean motion) and the time T of last perihelion passage. This angle M (mean anomaly) is expressed by $M = n(t - T)$. Therefore, given the orbital elements a, e, n, and T, it is possible by means of Kepler's equation and the intermediary angle θ' to evaluate the true anomaly θ and the actual position of the planet in the orbit for any instant t. *See* PLANETARY PHYSICS.

Several practical methods for the iterative solution of this transcendental equation exist: (1) Starting with the value of M and denoting approximate values of θ' and $\sin \theta'$ by θ'_0 and $\sin M$, solve $\theta'_0 = M + e \sin M$. Denoting a second approximation to θ' by θ'_1, solve $\theta'_1 = M = e \sin \theta'_0$. Indicating the third approximation to θ' by θ'_2, solve $\theta'_2 = M + e \sin \theta'_1$, continuing the iteration until the required convergence of θ'_i is obtained. (2) Starting with M and an approximate value of θ', solve $M_0 = \theta'_0 - e \sin \theta'_0$, where M_0 corresponds to θ'_0. Set $M - M_0 = \Delta M_0$ and let $\Delta\theta'_0$ be a first correction to θ'_0, then $\Delta\theta'_0 = \Delta M_0/1 - e \cos \theta'_0$. Set $\theta'_1 = \Delta\theta'_0 + \Delta\theta'_0$, evaluate

M_1 corresponding to θ_1', form $M - M_1 = \Delta M_1$, and solve $\Delta\theta_1' = \Delta M_1/1 - e \cos \theta_1'$. Then $\theta_2' = \theta_1' + \Delta\theta_1'$, and the process is repeated until the required convergence of $\Delta\theta_i'$ is reached. *See* CELESTIAL MECHANICS. Raynor L. Duncombe

Bibliography. P. Colwell, *Solving Kepler's Equation over Three Centuries*, 1993; J. M. A. Danby, *Fundamentals of Celestial Mechanics*, 2d ed., 1988.

Kepler's laws

The three laws of planetary motion discovered by Johannes Kepler during the early years of the seventeenth century.

First law. The first law of Kepler states that a planet moves in an elliptical orbit around the Sun that is located at one of the two foci of the ellipse. An ellipse is one of the conic curves originally studied by Greek geometers. It is formed when a cone is cut by a plane that is neither parallel nor perpendicular to the axis and not parallel to the side of the cone. As the plane approaches perpendicularity with the axis, the ellipse approaches a circle; Kepler's first law accepts a circular orbit as the limiting case of an ellipse. As the cutting plane approaches a position parallel to the side of the cone, the ellipse approaches a parabola; although a body can move in a parabolic path around another placed at its focus, it would not be a planet in the accepted sense because it would never return. An ellipse is also defined as the locus of points the sum of whose distances from two fixed points (the foci) is a constant. *See* CONIC SECTION; EL-LIPSE.

Kepler discovered the first law by analyzing the observations that Tycho Brahe had made of the planet Mars in the late sixteenth century. An earlier astronomer of that century, Nicolaus Copernicus, had broken with the commonsense notion, apparently supported by daily experience, that the Earth is the center of the universe around which all other bodies move. However, he had accepted another tradition in astronomy, that all heavenly appearances are to be explained by combinations of circles, which were held to be perfect figures. Kepler's ellipses shattered that tradition. They also challenged another premise of earlier astronomy, that the heavens are fundamentally different from the Earth. Although Kepler discovered the ellipse by studying Mars, he generalized the shape to the orbits of all the planets.

In 1610, only one year after Kepler published his first law, Galileo discovered the satellites of Jupiter, and later in the seventeenth century satellites were observed around Saturn. All of these bodies obey Kepler's first law. In 1687, Isaac Newton demonstrated that any body, moving in an orbit around another body that attracts it with a force that varies inversely as the square of the distance between them, must move in a conic section. This path will be an ellipse when the velocity is below a certain limit in relation to the attracting force. Thus the first law is not limited to primary planets around the Sun or to heavenly bodies. It is a general law that applies to all satellites held in orbit by an inverse-square force.

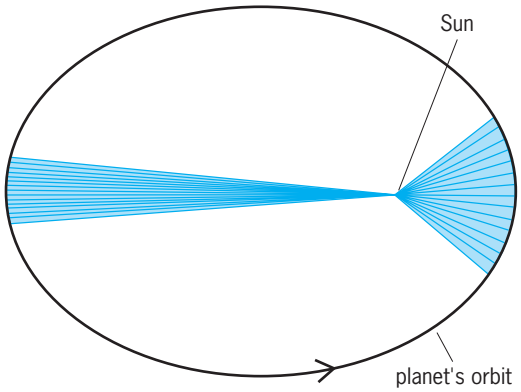

Demonstration of Kepler's first and second laws. The planet moves along an elliptical orbit at a nonuniform rate, so that the radius vector drawn to the Sun, which is located at one focus of the ellipse, sweeps out areas that are proportional to time. Thus, the planet would take equal times (corresponding to the equal areas) to traverse the unequal distances along the ellipse that correspond to the two shaded areas. The diagram greatly exaggerates the eccentricity of any orbital ellipse in the solar system.

Second law. Astronomers needed to predict the location of planets at given times. The system of combining circles, which were assumed to turn uniformly, had made it possible to calculate a position at any time by the vectorial addition of the radii. In rejecting the system of circles, Kepler needed some other method by which to predict locations, and this he provided with his second law. It states that the radius vector of the ellipse (the imaginary line between the planet and the Sun) sweeps out areas that are proportional to time (see **illus.**). A planet does not move along its elliptical path at a uniform velocity; it moves more swiftly when it is closer to the Sun and more slowly when it is farther removed. Behind the nonuniform velocities the second law finds a uniformity in the areas described.

Again Newton demonstrated the dynamic cause behind Kepler's second law. In this case, it is not restricted to forces that vary inversely as the square of the distance; rather it is valid for all forces of attraction between the two bodies, regardless of the law the forces obey. The second law expresses the principle of the conservation of angular momentum. If body B moves in relation to body A in a straight line with a uniform velocity, it has a constant angular momentum in relation to A, and the line joining B to A sweeps out equal areas in equal increments of time. No force of attraction (or repulsion) between the two bodies can alter their angular momentum about each other. The second law then expresses a relation that holds for all pairs of bodies with radial forces between them. *See* ANGULAR MOMENTUM.

Third law. Kepler's first two laws govern the orbits of individual planets around the Sun. His third law defines the relations that hold within the system of planets. It states that the ratio between the square of a planet's period (the time required to complete one orbit) to the cube of the mean radius (the average distance from the Sun during one orbit) is a

constant. The four satellites that Galileo had discovered around Jupiter were found to obey the third law, as did the satellites later found around Saturn. Newton demonstrated once again that the third law is valid for every system of satellites around a central body that attracts them, as the Sun attracts the planets, with a force that varies inversely as the square of the distance.

In fact, none of the three laws precisely describes the motions of the planets around the Sun, for the laws assume the only forces that act are those between the Sun and individual planets. But Newton's law of universal gravitation asserts that all planets also attract each other. In the case of the Moon orbiting the Earth, the attraction that the Sun exerts on the Moon is large enough to introduce perturbations, which astronomers had identified empirically before Kepler stated his three laws. Because the Sun is immensely more massive than the planets, the perturbations caused by the planets' mutual attractions are quite small, and Kepler's laws offer an accurate first approximation of their motions. Kepler's laws of planetary motion were the first of the mathematical laws of modern science; they did much to determine the pattern that science has continued to pursue. *See* CELESTIAL MECHANICS; GRAVITATION; ORBITAL MOTION; PERTURBATION (ASTRONOMY). Richard S. Westfall

Bibliography. G. Holton, Johannes Kepler's universe: Its physics and metaphysics, *Amer. J. Phys.*, 24:340–351, 1956; W. Pauli, *The Influence of Archetypal Ideas on the Scientific Theories of Kepler*, 1955; J. Voekel, *Johannes Kepler and the New Astronomy*, Oxford University Press, 2001.

Kernite

A hydrated borate mineral with chemical composition $Na_2B_4O_6(OH)_2{\cdot}3H_2O$. It is monoclinic with symmetry 2/m. It occurs only very rarely in crystals but is found most commonly in coarse, cleavable masses (see **illus.**) and aggregates. It has two perfect cleavages, {001} and {100}, at about 71° to each other, which together produce cleavage fragments that are elongated parallel to the *b* crystallographic axis. It is colorless to white; colorless and transparent specimens tend to become chalky white on exposure to air with the formation of tincalconite, $Na_2B_4O_5(OH)_4{\cdot}3H_2O$. The structure of kernite, with space group $P\,2/a$, contains infinite chains of composition $[B_4O_6(OH)_2]^{2-}$ which run parallel to the *b* crystallographic axis. The chains are linked to each other by bonding to Na^+ and by hydroxyl-oxygen bonds.

Typical cleavage fragment of kernite, from Boron, Kern County, California. (*Specimen from Department of Geology, Indiana University*)

Kernite was originally discovered at Boron, in the Kramer borate district, in the Mojave Desert of California. It is the second most important ore mineral, after borax, in this borate-producing region. It occurs with borax, $Na_2B_4O_5(OH)_4 \cdot 8H_2O$, and other borates such as ulexite and colemanite in a bedded series of Tertiary clays. Kernite occurs mainly near the bottom of the deposits and is believed to have formed by recrystallization from borax due to increased temperature and pressure. Kernite is also found in borate deposits in Tincalayu, in the province of Salta, Argentina. The three largest producers of crude borates are the United States, Turkey, and Argentina.

Boron compounds are used in the manufacture of glass, especially in glass wool used for insulation purposes. They are also used in soap, in porcelain enamels for coating metal surfaces, and in the preparation of fertilizers and herbicides. *See* BORATE MINERALS; BORON. Cornelis Klein

Bibliography. C. Klein and C. S. Hurlbut, Jr., *Manual of Mineralogy*, 21st ed., rev. 1999.

Kerogen

The complex, disseminated organic matter present in sedimentary rocks that remains undissolved by sequential treatment with common organic solvents (such as a benzene-methanol mixture, toluene, and methylene chloride) followed by treatment with nonoxidizing mineral hydrochloric acid and hydrofluoric acid. *See* SEDIMENTARY ROCKS.

The term kerogen is used to refer to a material derived from an extraction procedure. The term is defined operationally, and does not refer to a well-characterized chemical substance; thus some variability occurs in its definition. Some workers define kerogen as including all sediment organic matter that remains after extraction with organic solvents. However, kerogen isolation also involves recovery of the residual organic matter after treatment of the sediment with hydrochloric and hydrofluoric acids to remove the carbonate and silicate minerals, respectively. Therefore, the definition given here may be more appropriate in terms of the organic fraction actually isolated. The insoluble organic matter initially deposited with unconsolidated sediments often is not considered to be kerogen, but a so-called protokerogen that can be released by more drastic acid treatment (such as boiling several hours with 6 *N* hydrochloric acid) and is gradually converted to kerogen by elimination of carbon dioxide and water during early stages of sediment burial and diagenesis. However, the normal kerogen isolation procedures (using more dilute hydrochloric acid for shorter times) do not provide a clear distinction between kerogen and protokerogen, as defined above.

Kerogen is considered to be the major starting material for most oil and gas generation as sediments are subjected to geothermal heating in the subsurface. It is the most abundant form of organic carbon on Earth—about 1000 times more abundant than coal, which forms primarily from terrigenous remains of higher plants. Kerogen is formed from the remains of marine and lacustrine microorganisms, plants and animals, and variable amounts of terrigenous debris in sediments. The terrestrial portions of kerogen have elemental compositions similar to coal. *See* COAL.

Kerogen is sometimes pictured as a complex material of high molecular weight formed from the random condensation of monomers generated by an initial breakdown of polymeric biological precursor molecules following sediment burial. However, microscopic and chemical evidence is more in accord with a model of substantial incorporation of biological macromolecules that have undergone varying degrees of chemical alteration prior to and after burial.

Kerogens are classified according to their atomic ratios of hydrogen to carbon (H/C) and oxygen to carbon (O/C), with oil-prone kerogens being generally higher in H/C and lower in O/C than the gas-prone kerogens. With increasing length of exposure to subsurface temperatures, all kerogens show decreases in the O/C and H/C ratios as they generate preferentially carbon dioxide and water, then oil, and finally only gas (methane) at progressively higher subsurface depths and temperatures. *See* NATURAL GAS; PETROLEUM.

Jean K. Whelan

Kerosine

A refined petroleum fraction used as a fuel for heating and cooking, jet engines, lamps, and weed burning, and as a base for insecticides. Kerosine, known also as lamp oil, is recovered from crude oil by distillation. It boils in the approximate range of 350–550°F (180–290°C). Most marketed grades, however, have narrower boiling ranges. The specific gravity is about 0.8. Determined by the Abel tester, the flash point is not below 73°F (23°C), but usually a higher flash point is specified. Down to a temperature of −25°F (−32°C), kerosine remains in the liquid phase. Components are mainly paraffinic and naphthenic hydrocarbons which are in the C_{10}–C_{14} range. A low content of aromatics is desirable except when kerosine is used as tractor fuel.

Specifications are established for specific grades of kerosine by government agencies and by refiners. Since these specifications are developed from performance observations, they are adhered to rigidly to assure satisfactory operation. For use in lamps, for example, a highly paraffinic oil is desired because aromatics and naphthenes give a smoky flame; and for satisfactory wick feeding, a viscosity no greater than 2 centipoises is required in this application. Furthermore, the nonvolatile components must be kept low. In order to avoid atmospheric pollution, sulfur content must be low; a minimum flash point of 100°F (38°C) is desirable to reduce explosion hazards. *See* PETROLEUM PRODUCTS.

Harold C. Ries

Kerr effect

Electrically induced birefringence that is proportional to the square of the electric field. When a substance (especially a liquid or a gas) is placed in an electric field, its molecules may become partly oriented. This renders the substance anisotropic and gives it birefringence, that is, the ability to refract light differently in two directions. This effect, which was discovered in 1875 by John Kerr, is called the electrooptical Kerr effect, or simply the Kerr effect. *See* BIREFRINGENCE.

When a liquid is placed in an electric field, it behaves optically like a uniaxial crystal with the optical axis parallel to the electric lines of force. The Kerr effect is usually observed by passing light between two capacitor plates inserted in a glass cell containing the liquid. Such a device is known as a Kerr cell. There are two principal indices of refraction, n_o and n_e (known as the ordinary and extraordinary indices), and the substance is called a positively or negatively birefringent substance, depending on whether $n_e - n_o$ is positive or negative.

Light passing through the medium normal to the electric lines of force (that is, parallel to the capacitor plates) is split into two linearly polarized waves traveling with the velocities c/n_o and c/n_e, respectively, where c is the velocity of light, and with the electric vector vibrating perpendicular and parallel to the lines of force.

The difference in propagation velocity causes a phase difference δ between the two waves, which, for monochromatic light of wavelength λ_0, is $\delta = (n_e - n_o)x/\lambda_0$, where x is the length of the light path in the medium.

Kerr constant. Kerr found empirically that $(n_e - n_o) = \lambda_0 BE^2$, where E is the electric field strength and B a constant characteristic of the material, called the Kerr constant. Havelock's law states that $B\lambda n/(n-1)^2 = k$, where n is the refractive index of the substance in the absence of the field and k is a constant characteristic of the substance but independent of the wavelength λ. Roughly speaking, the Kerr constant is inversely proportional to the absolute temperature. The phase difference δ is determined experimentally by standard optical techniques. If the wavelength λ_0 is expressed in meters, and the field strength E in volts/m, the Kerr constant for carbon disulfide, which has been determined most accurately, is $B = 3.58 \times 10^{-14}$. Values of B range from -25×10^{-14} for paraldehyde to $+384 \times 10^{-14}$ for nitrobenzol.

The theory of the Kerr effect is based on the fact that individual molecules are not electrically isotropic but have permanent or induced electric dipoles. The electric field tends to orient these dipoles, while the normal agitation tends to destroy the orientation. The balance that is struck depends on the size of the dipole moment, the magnitude of

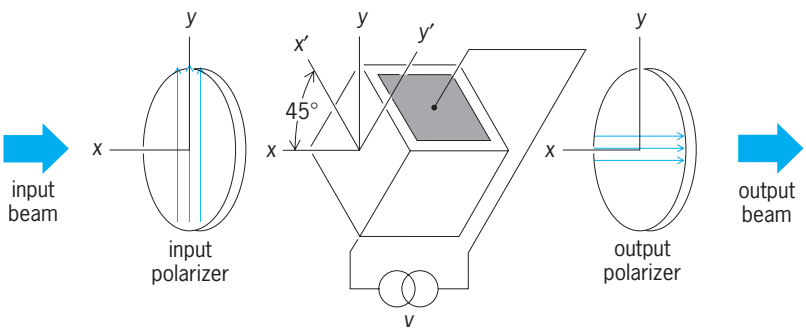

Kerr shutter. The arrows designate a light beam propagating through the cell. Polarizer axes are parallel to *x* and *y* axes; electric field of cell is parallel to *x′* or *y′* axis. (*After A. Yariv, Quantum Electronics, John Wiley and Sons, 1967*)

the electric field, and the temperature. This theory accounts well for the observed properties of the Kerr effect.

In certain crystals there may be an electrically induced birefringence that is proportional to the first power of the electric field. This is called the Pockels effect. In these crystals the Pockels effect usually overshadows the Kerr effect, which is nonetheless present. In crystals of cubic symmetry and in isotropic solids (such as glass) only the Kerr effect is present. In these substances the electrically induced birefringence (Kerr effect) must be carefully distinguished from that due to mechanical strains induced by the same field. *See* ELECTROOPTICS.

Kerr shutter. An optical Kerr shutter or Kerr cell consists of a cell containing a liquid (for example, nitrobenzene) placed between crossed polarizers. As such, its construction very much resembles that of a Pockels cell. With a Kerr cell (see **illus.**), an electric field is applied by means of an electronic driver v and capacitorlike electrodes in contact with the liquid; the field is perpendicular to the axis of light propagation and at 45° to the axis of either polarizer. In the absence of a field, the optical path through the crossed polarizers is opaque. When a field is applied, the liquid becomes birefringent, opening a path through the crossed polarizers. In commercial Kerr cell shutters, the electric field (a typical value is 10 kV/cm) is turned on and off in a matter of several nanoseconds (1 ns = 10^{-9} s). For laser-beam modulation the Pockels cell is preferred because it requires smaller voltage pulses. Kerr cell shutters have the advantage over the Pockels cell of a wider acceptance angle for the incoming light. *See* OPTICAL MODULATORS.

A so-called ac Kerr effect or optical Kerr effect has also been observed and put to use in connection with lasers. When a powerful plane-polarized laser beam propagates through a liquid, it induces a birefringence through a mechanism that is very similar to that of the ordinary, or dc, Kerr effect. In this case, it is the ac electric field of the laser beam (oscillating at a frequency of several hundred terahertz) which lines up the molecules. By using laser pulses with durations of only a few picoseconds (1 ps = 10^{-12} s), extremely fast optical Kerr shutters have been built in the laboratory. *See* LASER.

Michel A. Duguay

Bibliography. A. Yariv, *Optical Electronics*, 5th ed., 1997.

Ketene

A member of a class of organic compounds with the C═C═O group as a common structural element. Ketenes are derivatives of carboxylic acids, from which they are (hypothetically) formed by abstraction of water; they can therefore be considered to be inner anhydrides of acids, as opposed to the common carboxylic acid anhydrides formed from two molecules of a carboxylic acid. An example is the formation of ketene and acetic anhydride from acetic acid [reactions (1) and (2)].

$$CH_3COOH \longrightarrow H_2C{=}C{=}O + H_2O \qquad (1)$$

$$2CH_3COOH \longrightarrow CH_3COOCOCH_3 + H_2O \qquad (2)$$

Infrared spectra of ketenes show very strong bands near 1120 and 2150 cm^{-1}, associated with the asymmetric stretching mode of the C═C and C═O bonds of the heterocumulene (C═C═O); both bonds are used for identification of ketenes. *See* INFRARED SPECTROSCOPY.

Like real anhydrides, ketenes are acylating agents which readily undergo reactions with many compounds containing active hydrogens. They are relatively labile compounds, and only a limited number have been prepared and isolated. Many ketenes have been prepared in place and reacted immediately. *See* ACID ANHYDRIDE.

Preparation. Ketene, the parent compound, is a colorless gas that boils at −41°C (−41.8°F) and is stable for an extent of time only at low temperatures. It is a commercial intermediate for the synthesis of acetic anhydride. Industrially it is produced by dehydration of acetic acid at 700–750°C (1290–1380°F) in the presence of triethyl phosphate as catalyst (90% selectivity). It can also be obtained by thermolysis of acetone at 600–700°C (1110–1290°F) in the presence of catalytic amounts of carbon disulfide [reaction (3)].

$$\begin{array}{ll} CH_3COOH \xrightarrow[-H_2O]{\Delta} & \\ & H_2C{=}C{=}O \\ CH_3COCH_3 \xrightarrow[-CH_4]{\Delta} & \end{array} \qquad (3)$$

Other major laboratory methods, primarily for preparing the higher homologs, involve the thermolysis or radiative rearrangement of diazo carbonyl compounds (Wolff rearrangement) with loss of nitrogen [reaction (4)]; dehalogenation of 2-haloacyl

$$R^1{-}\overset{\overset{\displaystyle O}{\|}}{C}{-}\overset{\overset{\displaystyle N_2}{|}}{C}{-}R^2 \xrightarrow{h\nu,\Delta} \begin{matrix} R^1 \\ R^2 \end{matrix}\!\!>{=}C{=}O + N_2 \qquad (4)$$

halides, predominantly bromides, with zinc [Zn;

reaction (5)]; dehydrohalogenation of acyl halides with tertiary amines [NR_3; reaction (6)].

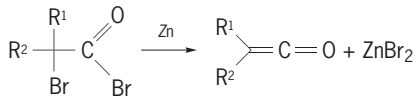

(5)

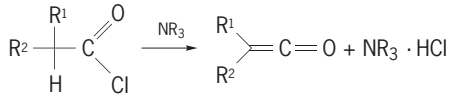

(6)

A variety of ketenes have been synthesized in which R^1 and R^2 in the formulas above represent hydrogen, alkyl, aryl, halogen, alkoxy, amino, carbonyl, thiocarbonyl, imino, cyano, or azido groups. Numerous ketenes in which the heterocumulene group is linked directly to metals such as silicon (Si), germanium (Ge), or tin (Sn) have also been prepared. Many alkyl ketenes have been prepared only in place and converted immediately into stable addition products. Aryl ketenes are more stable and can be prepared and isolated in relatively good yields. Haloketenes are extremely unstable and have most often been identified by their conversion into addition products. Nevertheless, dichloroketene has found wide application in the synthesis of various cycloadducts.

Reactions. Ketenes react readily with nucleophiles such as alcohols to form esters, with amines to form amides, and with acids to yield anhydrides. Reactions of ketenes that are important in chemical synthesis are 2 + 2 cycloadditions with a variety of double-bond systems. Four-membered-ring cycloadducts have been obtained with C═C, C═O, C═N, N═N, and N═S bonds. Primary examples are the dimerization of ketene, which produces an asymmetric adduct having the oxetanone structure [reaction (7)].

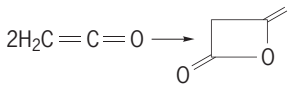

(7)

Mono- and disubstituted ketenes were found to form either symmetric dimers—cyclobutane derivatives, in which cycloaddition takes place only across the C═C double bond of the heterocumulene [reaction (8)]—or asymmetric dimers, or mixtures of both.

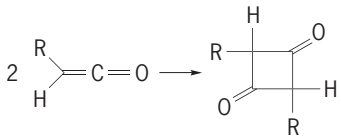

(8)

Halogenated ketenes do not dimerize but undergo cycloadditions with many other double-bond systems. These adducts are very valuable synthetic intermediates, since the halogens can be easily replaced in hydrogenation or substitution reactions. An example is the synthesis of tropolones (I) on the basis of the reaction of dichloroketene with cyclopentadiene [reaction (9)].

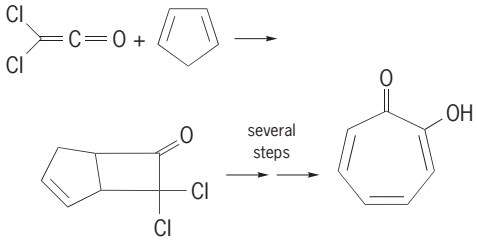

(9)

Spiro compounds have been prepared by reacting methylene cycloalkanes with dichloroketene [reaction (10)].

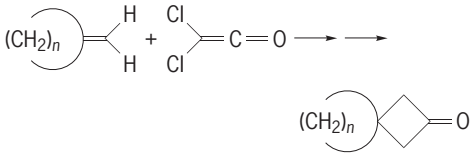

(10)

A large number of azetidinones (β-lactams) have been prepared via the 2 + 2 cycloaddition between ketenes and amidines and azomethines [reaction (11)].

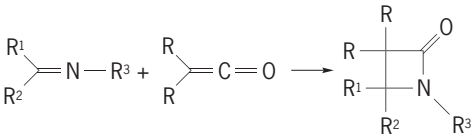

(11)

Similarly, arylketenes, generated from diazoketones as shown in reaction (4), and diarylketenes react with azobenzene to yield a variety of 1,2,3-triaryl or tetraryl-1,2-diazetidinones [reaction (12)].

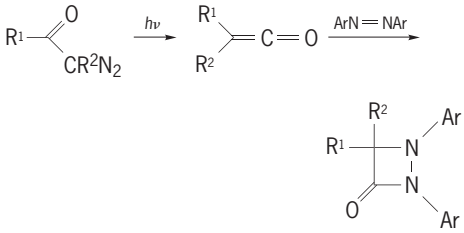

(12)

Ketenes also undergo 1,4-cycloaddition reactions with conjugated double-bond systems, which are used very often for the synthesis of both fused cyclic and heterocyclic systems [reaction (13)].

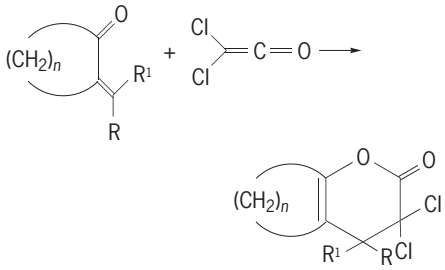

(13)

See STEREOCHEMISTRY. Reinhard H. Richter

Bibliography. S. Patai (ed.), *The Chemistry of Ketenes, Allenes, and Related Compounds*, pt. 1, 1980; M. B. Smith and J. March, *March's Advanced Organic Chemistry: Reactions, Mechanisms, and*

Structure, 5th ed., 2001; T. T. Tidwell, *Ketenes*, 2d ed., 2006.

Ketone

One of a class of chemical compounds of the general formula

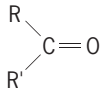

R and R′ are alkyl, aryl, or heterocyclic radicals. The groups R and R′ may be the same or different or incorporated into a ring as in cyclopentanone (**1**).

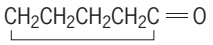

(**1**)

The ketones, acetone and methyl ethyl ketone, are used as solvents. Ketones are important intermediates in the syntheses of organic compounds.

By common nomenclature rules, the R and R′ groups are named, followed by the word ketone, for example, $CH_3CH_2COCH_2CH_3$ (diethyl ketone), $CH_3COCH(CH_3)_2$ (methyl isopropyl ketone), and $C_6H_5COC_6H_5$ (diphenyl ketone). The nomenclature of the International Union of Pure and Applied Chemistry uses the hydrocarbon name corresponding to the maximum number of carbon atoms in a continuous chain in the ketone molecule, followed by "-one," and preceded by a number designating the position of the carbonyl group in the carbon chain. The first two ketones above are named 3-pentanone and 3-methyl-3-butanone.

Properties. The lower-molecular-weight ketones are colorless liquids. Acetone and methyl ethyl ketone are miscible with water; the water solubility of the higher homologs decreases with increasing number of carbon atoms. In the infrared absorption spectrum, the intense carbonyl stretching band near 1715 cm^{-1} is useful for identification and characterization of ketones. Because of their characteristic odors, various ketones are of use in the flavoring and perfumery industry. Biacetyl (**2**) is a principal ingredient in the flavoring of margarine, camphor (**3**) is valued for its medicinal odor, although it appears to have no therapeutic value, and cyclopentadecanone (Exaltone; **4**) has a musk odor and is used in perfumes.

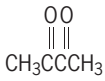

(**2**)

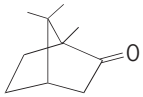

(**3**)

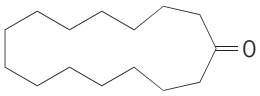

(**4**)

Reactions of ketones. Addition to the carbonyl group is the most important type of ketone reaction. Ketones are generally less reactive than aldehydes in addition reactions. Methyl ketones are more reactive than the higher ketones because of steric group effects.

Hydrogen adds catalytically to the carbonyl group, and lithium aluminum hydride gives the same type of product—the secondary alcohol, as shown in reaction (1). The Grignard reagent adds to the car-

$$RCOR' \xrightarrow[\text{or } LiAlH_4]{H_2(\text{Ni or Pt})} RCHOHR' \qquad (1)$$

bonyl group, and tertiary alcohols are formed by hydrolysis, reaction (2). Hydrogen cyanide and sodium

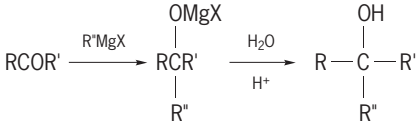

(2)

bisulfite add to methyl ketones, as in reaction (3).

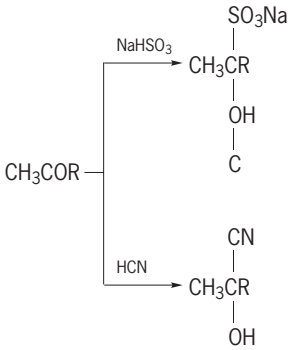

(3)

Alcohols do not add readily to the ketone carbonyl as they do to the aldehyde carbonyl, but ketals may be formed by the action of orthoformates, as in reaction (4). Amine derivatives such as hydroxylamine

$$RCOR' + HC(OC_2H_5)_3 \xrightarrow{H^+} R{-}C(OC_2H_5)_2R' + HCOOC_2H_5 \qquad (4)$$

(NH_2OH), phenylhydrazine ($C_6H_5NHNH_2$), and semicarbazide (NH_2-$CONHNH_2$) add to the carbonyl by breaking an N-H bond and with subsequent loss of water, as in reaction (5). The resulting oximes,

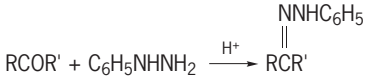

(5)

phenylhydrazones, and semicarbazones are useful derivatives for the identification and characterization of ketones.

Ketones supply alpha hydrogen in aldol-type condensation reactions, as in reaction (6), but can supply

$$H_2CO + CH_3COR \xrightarrow{OH^-} RCOCH_2OH \xrightarrow{H^+} RCOCH{=}CH_2 + H_2O \qquad (6)$$

a carbonyl group only to a limited extent because

of its lower reactivity. An exception is the self-condensation of acetone to diacetone alcohol represented by reaction (7).

$$2CH_3COCH_3 \xrightarrow{Ba(OH)_2} CH_3\overset{\overset{\displaystyle OH}{|}}{\underset{\underset{\displaystyle CH_3}{|}}{C}}CH_2COCH_3 \qquad (7)$$

Ketones are oxidized less readily and less selectively than aldehydes, to give oxidation products such as carboxylic acids, by cleavage of the bonds from the carbonyl carbon atom to the adjacent atom.

Methyl ketones give the haloform reaction with solutions of iodine in aqueous potassium hydroxide, reaction (8).

$$CH_3COR \xrightarrow[KOH]{I_2} CHI_3 + RCOOK \qquad (8)$$

Chlorine or bromine will substitute the alpha hydrogen atoms of ketones, as in reaction (9). The re-

$$RCH_2COR' + Cl_2 \longrightarrow R\overset{\overset{\displaystyle Cl}{|}}{C}HCOR' + HCl \qquad (9)$$

sulting α-haloketones are quite reactive in displacement reactions of the halogen atom.

Preparation. Ketones are formed by the dehydrogenation or oxidation of secondary alcohols, as shown in reaction (10), and this is the method

$$RCH(OH)R' \longrightarrow RCOR' \qquad (10)$$

used industrially for the preparation of acetone and methyl ethyl ketone. In the United States, acetone and methyl ethyl ketone are produced primarily by this reaction.

Aromatic ketones may be prepared by the Friedel-Crafts acylation reaction using either acid halides as shown in reaction (11) for the preparation of ace-

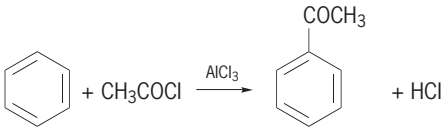

(11)

tophenone, or anhydrides as shown in reaction (12) for the preparation of *o*-benzoylbenzoic acid.

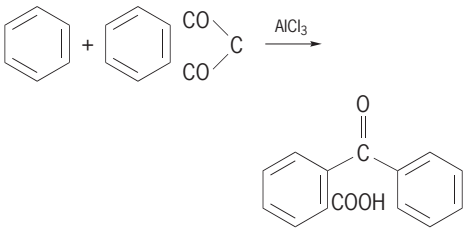

(12)

Organocadmium reagents give ketones on reaction with acid halides, reaction (13).

$$RCOCl + R'CdX \longrightarrow RCOR' + CdXCl \qquad (13)$$

Beta-ketoesters formed in the Claisen condensation are cleaved by aqueous sodium hydroxide solution to form ketones, reactions (14) and (15). The

$$2RCH_2COOEt \xrightarrow{EtONa} RCH_2COCHRCOOEt + EtOH \qquad (14)$$

$$RCH_2COCHRCOOEt + H_2O \xrightarrow{NaOH} RCH_2COCH_2R + CO_2 + EtOH \qquad (15)$$

Dieckmann condensation of esters of dibasic acids leads in similar fashion to cyclic ketones, reactions (16) and (17).

$$EtOOC(CH_2)_4COOEt \xrightarrow{NaOEt} \underbrace{CH_2CH_2CH_2COCH}COOEt \qquad (16)$$

$$\underbrace{CH_2CH_2CH_2COCH}COOEt + H_2O \xrightarrow{NaOH} \underbrace{CH_2CH_2CH_2CH_2CO} + CO_2 + EtOH \qquad (17)$$

Both cyclic and open chain ketones are formed by pyrolysis of calcium or thorium salts of dibasic acids or monobasic acids, reactions (18) and (19).

$$\tfrac{1}{2}CaOOC(CH_2)_4COO\tfrac{1}{2}Ca \xrightarrow{heat} \underbrace{CH_2CH_2CH_2CH_2CO} + CaCO_3 \qquad (18)$$

$$2RCOO^{1}/_{2}Ca \xrightarrow{heat} RCOR + CaCO_3 \qquad (19)$$

See ALDEHYDE; FRIEDEL-CRAFTS REACTION; GRIGNARD REACTION; STERIC EFFECT (CHEMISTRY).

Paul E. Fanta

Bibliography. R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 5th ed., 1994; R. T. Morrison and R. N. Boyd, *Organic Chemistry*, 7th ed., 2001.

Key telephone system

A communications system that allows users to access more than one central office line, answer or access a central office line from more than one telephone, and place a line on hold in order to answer or initiate calls on other lines. The system usually includes an intercom capability that allows users on different telephones in the system to communicate with one

another. *See* INTERCOMMUNICATING SYSTEMS (INTERCOMS).

The main functional parts of a key telephone system are (1) the central service unit, also called the key service unit (KSU), which serves as an interface and switching center between incoming central office lines and system telephones; (2) station cabling to connect each telephone to the key service unit; (3) multibutton telephone sets, or key telephones, on which separate buttons provide access to the central office lines, intercom, and special features; and (4) a power supply, which converts available electrical power to the specifications required by the key telephone system.

The number of station lines and central office lines served by a key telephone system is limited by the size of the key service unit. On a typical key telephone system, each incoming central office line appears on a separate button on every key telephone set that is allowed to make or receive calls on that line. Usually, a visual indicator, such as a small incandescent lamp or light-emitting diode, allows the user to tell if the line is in use. The same lamp flashes at a certain rate when an incoming call is ringing on the line, and flashes at a different rate when a call is on hold. The number of lines that appear on every key telephone set varies with the capacity of buttons on each set. Such capacities vary from a low of 5 buttons to over 30 buttons, depending on the make and model of the key telephone system.

Technology. Key telephone systems are either electromechanical or electronic in design. The electromechanical type is also called a 1A2 system, or industry-standard key telephone system. The key service unit of an electromechanical system contains relays and electrical components, which amplify the voice signal and switch the signal between the central office lines and station lines. Electromechanical systems are highly reliable and relatively low in cost. However, they require station cabling that contains 25 copper wire pairs, which makes installation complex and changes difficult.

Electronic key telephone systems use microprocessors and integrated circuits to accomplish the switching, voice amplification, and feature activation. The key service unit, as well as the telephones, is proprietary and tends to be more costly than the industry-standard system. However, cabling usually consists of only two to four copper wire pairs, installation is simplified, and the features available are much more extensive and varied than with the 1A2 system. Changes to the system are accomplished by simply reprogramming the microprocessors that control the system. *See* INTEGRATED CIRCUITS; MICROPROCESSOR.

Features. As an option, the key telephone system may allow connection of a music source for playback to callers on hold, and access to centralized facilities such as dictation and paging equipment. Other common features are the ability to transfer a call from one telephone in the system to another, and the ability to unite three or more parties into a single conversation called a conference.

In addition, electronic key telephone systems may offer many specialized features that increase user productivity or meet important business needs. Widely used features include (1) last-number redial, which automatically redials the last number entered by the user; (2) speed dialing, which allows users to call a number either by dialing one or two digits or by depressing a programmed button on the key telephone set; (3) station message waiting, which allows one user to request a return call from another user; (4) hands-free operation, using telephones with built-in speakerphones; (5) station message detail recording, which prints a record of every outgoing telephone call made with the system; and (6) the ability to use inexpensive single-line telephones, in addition to the proprietary multibutton key telephones designed for the system.

Most businesses prefer electronic key telephone systems because of their ease of installation and versatility. Nevertheless, the traditional 1A2 system remains a reliable alternative. *See* SWITCHING SYSTEMS (COMMUNICATIONS); TELEPHONE SERVICE.

Sue R. Rubinstein; Vincent F. Rafferty

Bibliography. The Aries Group · MPSG, *Master Glossary of Terminology*, 1989; Datapro Research, *Telecommunications*, 1990.

Kidney

An organ involved with the elimination of water and waste products from the body. In vertebrates the kidneys are paired organs located close to the spine dorsal to the body cavity and are covered ventrally by the coelomic epithelium. They consist of a number of smaller functional units which are called urinary tubules or nephrons. The nephrons open to larger ducts, the collecting ducts, which open into a ureter. The two ureters run backward to open into the cloaca or into a urinary bladder. The shape and the location of the kidneys varies in different animals. In fish, they are extremely elongated and may reach forward almost to the pericardium. In reptiles, the kidneys are smaller and are located in the pelvic region. In mammals, they are bean-shaped and found between the thorax and the pelvis. The number, structure, and function of the nephrons vary with evolution and, in certain significant ways, with the adaptation of the animals to their various habitats.

Nephron. In its most primitive form, found only in invertebrates, the nephron has a funnel opening into the coelomic cavity followed by a urinary tubule leading to an excretory pore. In amphibians, some of the tubules have this funnel, but most of the tubules have a Bowman capsule (**Fig. 1**). In all higher vertebrates, the nephron has the Bowman capsule, which surrounds a tuft of capillary loops, called the glomerulus, constituting the closed end of the nephron. The inner epithelial wall of the Bowman capsule is in intimate contact with the endothelial wall of the capillaries. Electronmicroscopy reveals that the wall of the capillaries, together with

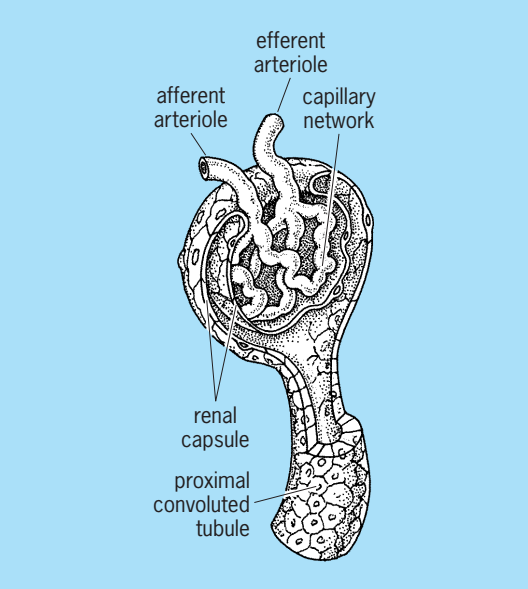

Fig. 1. Nephron, from frog kidney dissected to show glomerulus within Bowman capsule.

the inner wall of the Bowman capsule, forms a membrane ideally suited for filtration of the blood.

The nephron is differentiated into several parts which vary in different groups of vertebrates. Following the glomerulus comes a short neck segment which may be ciliated (this segment is absent in mammals). This is followed by the proximal tubule. Following the proximal tubule in lower vertebrates, there is a short intermediary segment; in mammals, there is the so-called thin segment comprising part of the loop of Henle. This is followed by the distal tubule, which is followed by the collecting duct. In lower vertebrates, all of the nephrons are convoluted and arranged perpendicular to the collecting ducts. In mammals, all of the nephrons have loops of Henle (**Fig. 2**). The glomerulus and convoluted proximal tubule are found in the cortex of the kidney. The loops of Henle run through the outer and inner medulla; some are short, others are long. The thin limb runs into a straight part of the distal tubule, which continues into the distal convoluted tubule found in the cortex. The capillaries are arranged in bundles which form true retia mirabilia, ideally suited for countercurrent exchange. The collecting ducts run from the cortex through the medulla to the tip of the papilla, where they open into the renal pelvis. Through their entire course the collecting ducts are parallel to the loops of Henle and to the capillary loops.

The blood pressure in the capillaries of the glomerulus causes filtering of blood by forcing fluid, small molecules, and ions through the membrane into the lumen of Bowman's capsule. This filtrate contains some of the proteins and all of the smaller molecules in the blood. As the filtrate passes down through the tubule, the walls of the tubule extract those substances not destined for excretion and return them to the blood in adjacent capillaries. Many substances which are toxic to the organism are moved in the opposite direction from the blood into the tubules. The urine thus produced by each nephron is conveyed by the collecting duct and ureter to the cloaca or bladder from which it can be eliminated.

Blood supply. In all classes of vertebrates the renal arteries deliver blood to the glomeruli and through a second capillary net to the tubules. The major blood supply to the kidney tubules comes, however, from the renal portal vein, which is found in all vertebrates except mammals and cyclostomes (Fig. 1). Venous blood from the tail and hindlimbs of the animal drains into the renal portal vein. It then splits up into a second capillary net which surrounds the renal tubules. Waste products from the venous blood can thus be secreted directly into the urinary tubules.

Evolutionary and adaptive aspects. It is assumed that the ancestors of vertebrates were marine forms, with body fluids at osmotic equilibrium with seawater. Migration into fresh water posed the threat of

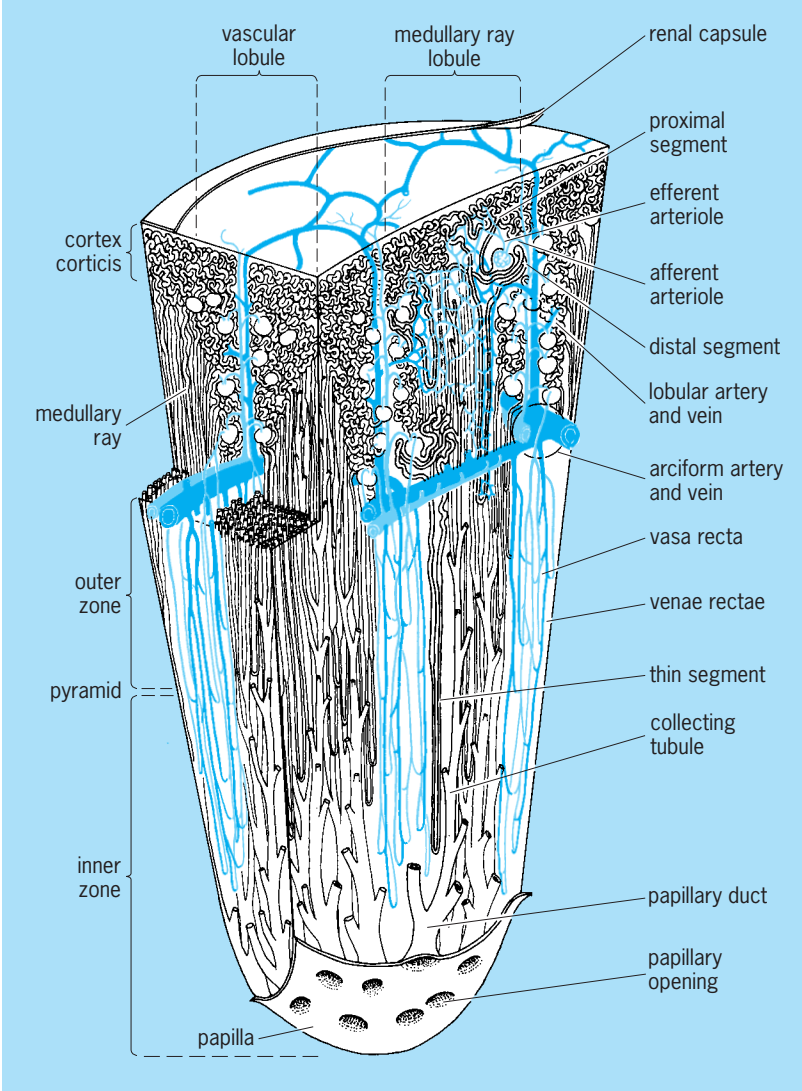

Fig. 2. Diagram of a small section of a human kidney, showing the loops of Henle with their thick and thin segments, and the vasa recta and the collecting ducts as they are arranged parallel to one another in the outer and inner medulla. (*From Homer Smith, Principles of Renal Physiology, Oxford University Press, 1956*)

rapid dilution of body fluids by osmosis. The kidneys in fresh-water forms arose as the major organ for the removal of excess water. A large glomerulus is present in amphibians and fishes found in fresh water today. Their nephrons have a long distal tubule which reabsorbs ions from the filtrate and returns them to the blood. Through this process the urine becomes dilute, and very little salt is lost hereby. However, when fishes reinvaded the ocean, or evolved into land forms, their kidneys had to be modified to their new habitats.

The salt-water teleost fishes maintain an osmotic blood concentration lower than that of the seawater. They are thus in danger of becoming dehydrated, and water conservation becomes a prime objective. Some marine teleosts have lost their glomeruli and the distal tubules entirely, and produce urine by tubular secretion alone. In other marine fishes the glomeruli have been reduced in size.

To maintain osmotic pressure of the blood lower than that of the seawater, the gills secrete salt into the seawater.

In land animals such as reptiles, birds, and mammals, water conservation is also a major objective. The primary function of the kidneys is to excrete salt and nitrogenous waste products. Reptiles and birds save water by excreting their nitrogenous wastes in the form of a nearly dry paste of uric acid crystal. The kidney of reptiles can only make urine that is isoosmotic or hypoosmotic to the blood, but not hyperosmotic. Reptiles faced with a large salt intake, such as marine turtles and lizards, excrete salt by special salt glands in the head. Birds can produce a urine hyperosmotic to the blood because of the countercurrent system formed by the few loops of Henle. They excrete uric acid. Salt-water birds excrete excess salt through glands similar to those of the reptiles.

Mammals are peculiar in that, regardless of habitat, they always excrete almost all of their nitrogenous wastes in the form of urea. It may be for this reason that the mammal has developed its unique kidney, which can produce urine with several times the osmotic concentration of the blood. The ability of the mammalian kidney to concentrate urine is inherent in the spatial arrangement of renal tubules, blood vessels, and collecting ducts which permits the tubules to function together as a countercurrent multiplier system (Fig. 2). *See* EXCRETION; OSMOREGULATORY MECHANISMS.

Embryology. Although nephrons are assembled into kidneys in a variety of ways, the tubules always arise in the intermediate mesoderm along the upper border of the body cavity (**Fig. 3**). The first tubules develop at the anterior end of this mesoderm, followed by others in sequence behind. Variations in this plan permit the following kidney types to be distinguished: holonephros, pronephros, mesonephros, opisthonephros, and metanephros.

Holonephros. There is one nephron beside each somite along the entire length of the coelom. Supposedly, this is the ancestral condition and is thus referred to as an archinephros. It is seen today only in the larvae of a few primitive fishes, the myxinoid cyclostomes.

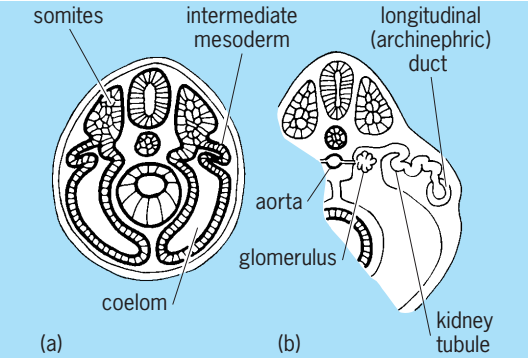

Fig. 3. Diagrammatic cross section of the embryos to show origin of nephron. (*a*) Early stage with mesoderm differentiation. (*b*) Later stage.

Pronephros. Nephrons are formed only in the anterior part of the intermediate mesoderm. The pronephros is rarely found as an adult excretory organ, but is important because in all embryos it initiates the formation of the archinephric duct used by subsequent tubules for drainage. The pronephros is of interest also because it shows features presumed to be primitive, suggesting a series of stages in evolution of the nephron.

Mesonephros. The nephrons form behind the pronephros along the middle zone of the coelom. The tubules usually become involved in sperm transport. Drainage is through the archinephric duct. The mesonephros never forms an entire adult kidney. It is functional in the chick and other amniote embryos.

Opisthonephros. Nephrons develop from the middle and posterior portions of intermediate mesoderm. The middle, or mesonephric, tubules sometimes form only a minor part. The archinephric duct may be supplemented or even replaced by ureterlike accessory drainage ducts. This is the functional adult kidney in amphibians and fishes.

Metanephros. Nephrons form from only the extreme posterior part of the intermediate mesoderm, and are drained by a distinctive new canal, the ureter, which grows forward from the base of the archinephric duct before the tubules have begun to develop. This is the kidney of adult reptiles, birds, and mammals. Their embryos show a distinct sequence of pro-, meso-, and metanephric kidneys. *See* URINARY SYSTEM.

Bodil Schmidt-Neilsen

Kidney disorders

Disorders of the kidney are classified by both their etiology and their anatomic location along the nephron. Disorders other than those arising from tumors become evident when the kidney is unable to regulate the volume and composition of the extracellular fluids, resulting in edema and hypertension. The clinical signs cannot identify the specific segment of the nephron that is affected, since overall

renal function depends on each segment functioning in sequence. The failure of one part, therefore, affects the functioning of the whole unit and makes a precise diagnosis difficult without a kidney biopsy. Knowing the location and type of abnormality is important because it may determine the prognosis or therapy.

Kidney disorders may be divided into the following categories: congenital, inflammatory, hereditary, infection-related, or associated with metabolic diseases (such as diabetes mellitus), toxins or drugs, circulatory collapse, and cancer.

Congenital abnormalities. The absence of kidneys at birth, which is quite rare, is incompatible with life. A slightly more common condition, however, is the presence of only one kidney or the presence of one kidney in an abnormal location.

The most common congenital disorders are abnormalities in the peristaltic passage of urine. Normally, urine flows from the kidneys into the ureters, where it is moved by a rhythmic peristaltic motion to the bladder. It is stored there until voluntary voiding is begun, whereupon it passes through the urethra and out of the body. Any abnormality in such a complex and integrated neuromuscular system responsible for the passage of urine or any mass of tissue that blocks urine flow causes retention of the urine. All such abnormalities are referred to as obstructive lesions, even though some are not actual physical masses. The most common obstructive lesions are tissue masses that act as valves, inhibiting passage of urine in the normal direction but allowing retrograde flow. In many cases, diseases that inhibit the normal peristaltic flow of urine down the ureters because of a neuromuscular abnormality gradually run their course and do not require surgery. On the other hand, "valves" must be promptly removed surgically. *See* CONGENITAL ANOMALIES.

Hereditary disorders. The most common hereditary kidney disorder is adult polycystic kidney disease. It is an autosomal dominant disease that is found in all racial and ethnic groups and is characterized by the formation of cysts along the length of the nephron. As the cysts enlarge, the kidneys likewise enlarge—resulting in kidney failure in midadulthood. Approximately 10% of patients who undergo dialysis have polycystic kidney disease.

Alport's syndrome, a very rare genetic disease of the glomeruli, was first described in Finland and now is known to occur throughout the world. The disease results in glomerular scarring and eventual renal failure within the second or third decade of life that cannot be reversed. *See* HUMAN GENETICS.

Infection-related diseases. The kidneys are subject to infection from both local and remote sources and from a variety of causative organisms.

Pyelonephritis. The most common cause of pyelonephritis, a bacterial infection of the kidney, is obstruction of the bladder or ureters by a tumor or kidney stones that interrupt urine flow. The static urine allows bacteria to ascend to the kidney from the perineum, which accounts for the fact that the organisms most frequently involved are coliform bacteria, which are normal residents of that area. Treatment consists of removal of the obstruction and antibiotic therapy. Bacteria may also reach the kidney through the bloodstream from an infection elsewhere, but that is much less common.

Glomerulonephritis. The kidney may also be subject to infections originating at distant sites. Examples include bacterial products, such as endotoxins, that are released into the bloodstream and antibodies that are formed in response to bacterial invasion. In the latter case, the antibodies form antigen-antibody complexes in the circulation or the kidney. The complexes that are localized in the kidney cause an inflammatory reaction in the glomeruli, a process called acute glomerulonephritis. It is most frequently associated with streptococcal infections of the skin or the upper respiratory tract and is common in children, who frequently develop streptococcal infections. It is usually a short-lived ailment, and nearly all patients recover completely. *See* STREPTOCOCCUS.

Metabolic disorders. The most common metabolic disorder to affect the kidney is diabetes mellitus. Of those persons with the insulin-dependent form of diabetes, 30–40% develop renal disease. Kidney dysfunction is one of the most life-threatening complications of diabetes: between the ages of 25 and 40, the mortality rate of diabetics with kidney diseases is nearly 20 times that of the general population. Diabetes causes progressive glomerular scarring, which leads to kidney failure. No treatment has yet been devised to prevent either the onset or progression of diabetic kidney disease; however, careful control of blood sugar and systemic blood pressure are probably important. *See* DIABETES.

Systemic diseases. The two most common systemic diseases to affect the kidneys are systemic lupus erythematosus and vasculitis. Systemic lupus erythematosus is essentially a disease of young women and is characterized by the deposition of immunoglobulins and other immune substances in the glomeruli. The resulting acute inflammatory reaction may lead to progressive scarring. Systemic lupus erythematosus often responds to treatment with immunosuppressive agents. *See* CONNECTIVE TISSUE DISEASE.

Systemic vasculitis can affect either the medium-sized arteries or the arterioles. Involvement of the former leads to infarcts in various organs but seldom causes kidney failure. When the arterioles are affected, the lesions are diffuse, because the kidney is highly vascular and the arterioles, the smallest of the arteries, are abundant. Kidney failure often occurs. The cause of the disease is unknown, but those who have vasculitis often show systemic signs, including fever. The vascular lesions are characterized by an acute inflammatory infiltrate that involves the blood vessel wall. Treatment with anti-inflammatory medications can help clear the lesions.

Toxins and drugs. The kidney is a major site for the excretion of waste products as well as heavy metals, the acids and alkalies of metabolism, and drugs. Because the urine becomes concentrated (100-fold) as it travels the length of the tubule, the concentration

of otherwise-innocuous, or even therapeutic, substances may reach levels that are toxic to the kidneys. The accumulation of toxins within the tubule often leads to necrosis, or cell death, of the tubular epithelium or an inflammatory reaction in the spaces between the tubules. The result in either case is acute kidney failure. Usually, removing the offending agent and relying on dialysis until renal function returns reverses the condition. Chronic exposure to some phenacetin-containing compounds, however, may lead to progressive kidney fibrosis, which appears slowly and is irreversible. *See* TOXICOLOGY.

Circulatory collapse. If systemic blood pressure drops to levels at which the kidneys are no longer perfused, the tubules become deprived of oxygen and suffer severe injury. The most common causes of underperfusion are blood loss and cardiac failure. If bloodflow to the kidneys is restored, epithelial cell damage is rapidly reversed and renal function returns to normal. *See* VASCULAR DISORDERS.

Cancer. Cancer of the kidney occurs mainly in two age groups—the very young and those over 50 years of age. Tumors found in the young, the most common type being Wilms' tumor, are often bilateral, composed of fetal tissues, and may grow to a large size before detection. Some cancerous tumors of the kidney in children are responsive to radiation or chemotherapy.

Tumors in the adult kidney, adenocarcinomas or hypernephromas, are formed of tubular cells and usually metastasize to the lungs and bones. Blood in the urine is a common early sign, and if all traces of the cancerous growth are removed surgically, a cure is possible. *See* CANCER (MEDICINE); KIDNEY; URINARY SYSTEM.

Gary Striker

Bibliography. S. Massry and R. Glassock (eds.), *Textbook of Nephrology*, 4th ed., 2000; P. Sweny, *The Kidney and Its Disorders*, 1989; P. Sweny and Z. Varghese, *Renal Disease*, 1988; C. Tisher and B. Brenner (eds.), *Renal Pathology*, 2d ed., 1994.

Killing equations

A set of special, linear, first-order partial differential equations involving the coordinate variables x of an N-dimensional riemannian space, the metric tensor g_{ab} of this space, and a vector whose components V_a, V^a ($V^a = g^{ab}V_b$) are functions of x. Specifically, it is given by either Eq. (1) or Eq. (2). The subscript semi-

$$V^c g_{ab;c} + g_{ac}V^c{}_{;b} + g_{bc}V^c{}_{;a} = 0; \qquad (1)$$

$$c \equiv \frac{\partial}{\partial x^c}$$

$$V_{a,b} + V_{b,a} = 0 \qquad (2)$$

colon indicates partial differentiation with respect to a coordinate variable x, while the comma specifies covariant differentiation. The equivalence of Eqs. (1) and (2) may be established as follows. From the definitions of covariant differentiation and the structure of the Christoffel symbols { }, Eq. (3) follows, and one obtains Eq. (4) by virtue of the symmetry of g_{ab}.

$$\begin{aligned} V_{a,b} &= V_{a;b} - V_{d\{^{d}_{ab}\}} \\ &= (V^c g_{ac})_{;b} - V_d g^{cd^{1/2}}[-g_{ab;c} + g_{cb;a} + g_{ac;b}] \end{aligned} \qquad (3)$$

$$\begin{aligned} V_{a,b} + V_{b,a} &= (V^c g_{ac})_{;b} + (V^c g_{bc})_{;a} \\ &\quad -{}^1/_2 V^c[-g_{ab;c} - g_{ba;c} + g_{cba;a} \\ &\quad + g_{ca;b} + g_{ac;b} + g_{bc;a}] \\ &= g_{ac}V^c{}_{;b} + g_{bc}V^c{}_{;a} + V^c g_{ab;c} \end{aligned} \qquad (4)$$

The tensor components $V_a(x)$ are supposedly of the type (0,1, 0). Specifically, transformation equation (5) for V_a is induced by the change of coordi-

$$\overline{V}_r = V_a \frac{\partial x^a}{\partial \overline{x}^r} \qquad (5)$$

nates $x^a = x^a(\overline{x})$, $\overline{x}^r = \overline{x}^r(x)$. The quantities $V_{a,b}$ obtained by covariant differentiation are tensor components of the type (0,2, 0). Specifically, they obey Eq. (6), and therefore Eq. (7) is valid. Hence if for

$$\overline{V}_{r,s} = V_{a,b} \frac{\partial x^a}{\partial \overline{x}^r} \frac{\partial x^b}{\partial x^s} \qquad (6)$$

$$\begin{aligned} \overline{V}_{r,s} + \overline{V}_{s,r} &= V_{a,b} \frac{\partial x^a}{\partial \overline{x}^r} \frac{\partial x^b}{\partial \overline{x}^s} + V_{b,a} \frac{\partial x^b}{\partial \overline{x}^s} \frac{\partial x^a}{\partial \overline{x}^r} \\ &= (V_{a,b} + V_{b,a}) \frac{\partial x^a}{\partial \overline{x}^r} \frac{\partial x^b}{\partial \overline{x}^s} \end{aligned} \qquad (7)$$

a given tensor V_a [type (0,1, 0)] and a certain coordinate system (x) of the infinite collection of tensor analysis the Killing equations are valid, then these equations are valid for each coordinate system of the infinite collection. Briefly, the Killing equations are invariant for class C^l coordinate transformations.

The Killing equations are fundamental with regard to the question of the existence of certain mappings (or, synonymously, transformations) of a space onto itself. These particular transformations are frequently called infinitesimal groups of motions and are related to the ordinary groups of motions, that is, the motions which have the characteristic group properties.

Groups of motions. The one-parameter motions determined by mapping equations of the type in Eq. (8)

$$x^{*a} = f^a(x, t) \qquad (x \equiv \text{set } x^1, x^2, \ldots, x^N) \qquad (8)$$

will be said to constitute a one-parameter group of motions if the following group properties hold: (1) there exists an admissible value of t, t_0, such that for each admissible x, $x^{*a}(t_0) = f^a(x, t_0) = x^a$ (the identity transformation); (2) for each admissible set of x-values x^{*a} and each admissible value for t_1, the equations $x^{*a} = f^a(x, t_1)$ have a unique solution expressible in the form $x^a = f^a(x^*, t_2)$ [existence of an inverse]; (3) for each admissible set of x-values and each admissible pair of t-values t_1, t_2, there exists an admissible parameter value t_3, such that $x^{**a} = f^a(x^*, t_2) = f^a[f(x, t_1)t_2] = f^a(x, t_3)$ [the closure or principal group property]. The term "groups of motions" has been used at times under lighter conditions than

these; for example, condition 1 or 2, or both, has been omitted. *See* GROUP THEORY.

A very simple example of a group of motions for an ordinary euclidean plane E_2 and a rectangular cartesian coordinate system is given by Eq. (9), with t

$$x^{*a} = f^a(x, t) = x^a + t \tag{9}$$

restricted to real numbers. Group property 1 is obviously met by $t = 0$. Property 2 is satisfied since the inverse of $x^{*a} = x^a + t_1$ for each admissible value of t_1 is $x^a = x^{*a} + (-t_1) = f^a(x^*, -t_1)$. Property 3 and Eq. (9) require that for each admissible choice of t_1 and t_2 there exists an admissible t-value t_3 such that $x^{**a} = f^a(x, t_3) = x^a + t_3$ with $x^{**a} = f^a(x^*, t_2) = x^{*a} + t_2 = (x^a + t_1) + t_2$. Obviously the requirement of property 3 is fulfilled by $t_3 = t_1 + t_2$. Thus the group requirements are satisfied. The motion criterion is also met since $g_{ab} = \delta_{ab}$ ($= 1$ for $a = b$, $= 0$ for $a \neq b$) throughout the plane. Thus a change in the evaluation point does not cause a change in the value of g_{ab}, and the metric magnitudes (such as arc lengths and angles between intersecting curves) are preserved. A slight generalization of the motion of Eq. (9) is given by Eq. (10), with k^1 and k^2 given constant real numbers, not both zero.

$$f^a(x, t) = x^a + k^a t \tag{10}$$

As another simple example of a motion, consider a plane E^2 which has a point O deleted and contains a rectangular cartesian coordinate system (x) with center at O; and impose the point transformation in Eqs. (11). The identity transformation is given by

$$x^{*1} = f^1(x, \theta) = \cos\theta\, x^1 - \sin\theta\, x^2 \tag{11a}$$

$$x^{*2} = \sin\theta\, x^1 + \cos\theta\, x^2 \tag{11b}$$

$$0 \leq \theta < 2\pi$$

$\theta = 0$; and for $0 < \theta_1 < 2\pi$, the inverse is $x^a = f^a(x^*, \theta_2)$, $\theta_2 = 2\pi - \theta_1$. The group requirement is satisfied provided there is a functional relation $\theta_3 = \theta_3(\theta_1, \theta_2)$ for $0 \leq \theta_1 < 2\pi$, $0 \leq \theta_2 < 2\pi$, the relation having the following properties (1): $0 \leq \theta_3 < 2\pi$; (2) from $x^{**a} = f^a(x^*, \theta_2)$, $x^{*a} = f^a(x, \theta_1)$, it follows that $x^{**a} = f^a(x, \theta_3)$. From the rotational interpretation of mapping equations (11), it is evident that $\theta_3(\theta_1, \theta_2)$ may be defined by $\theta_3(\theta_1, \theta_2) = \theta_1 + \theta_2$ for $0 \leq \theta_1 + \theta_2 < 2\pi$; and $\theta_3(\theta_1, \theta_2) = \theta_1 + \theta_2 - 2\pi$ for $\theta_1 + \theta_2 \geq 2\pi$. Also, $g^*_{ab}(dx^{*a}du)\,(dx^{*b}d\upsilon) = g_{ab}(dx^a du) \cdot (dx^b d\upsilon)$, and it is evident that the metric entities are preserved in value by the mapping.

Formulation of motions. The definitive requirement for a mapping to be a motion, namely, the preservation of the magnitudes of the metric entities, is of course very stringent. Given a riemannian space R, with a coordinate system (x), one method of attack for searching for a motion is to seek a suitable coordinate transformation which preserves the structure of the metric tensor. If such a coordinate transformation $\bar{x}^r = \bar{x}^r(x)$ exists, then the tensor transformation in Eq. (12) can be altered in structure by replacing

$$g_{ab}(x) = \bar{g}_{rs}[\bar{x}(x)]\frac{\partial \bar{x}^r}{\partial x^a}\frac{\partial \bar{x}^s}{\partial x^b} \tag{12}$$

$\bar{g}_{rs}[\bar{x}(x)]$ with $g_{rs}[\bar{x}(x)]$ since the functional structure of $\bar{g}_{rs}(\bar{x})$ is the same as that of $g_{rs}(x)$ and therefore $\bar{g}_{rs}(\bar{x}) = g_{rs}(\bar{x})$. If this change is made, Eq. (13) results. Conversely, if the coordinate transformation is not known, then this last equation may be regarded tentatively as a set of $N(N + 1)/2$ first-order partial differential equations in N unknown functions $\bar{x}^r(x)$. While motions can be formulated in very special cases, the desirability of replacing the motion requirement with another less demanding one is obvious.

$$g_{ab}(x) = g_{rs}[\bar{x}(x)]\frac{\partial \bar{x}^r}{\partial x^a}\frac{\partial \bar{x}^s}{\partial x^b} \tag{13}$$

Modified motion requirement. If C_1 and C_2 are two class C' arcs given by $x^a = x_1{}^a(u)$, $x^a = x_2{}^a(\upsilon)$ which intersect at a point P for $u = u_0$, $\upsilon = \upsilon_0$ and if, in addition, the motion requirement is met, then $g_{ab}(x^*) \cdot dx_1{}^{*a}/du dx_2{}^{*b}/d\upsilon|u_0, \upsilon_0$ is a constant function of t and therefore its derivative vanishes identically in t. A natural reduction in the severity of this requirement is obtained by imposing the alternate condition that the derivative vanish at t_0, the identity value for t, which is now assumed to be $t_0 = 0$. It follows that $f^s(x, 0) = x^a$ and hence, by the McLaurin expansion for $f^a(x,t)$ that $x^{*a} = x^a + t \cdot [(\partial f^a/\partial t)|t = 0] + \cdots$. These facts suggest consideration of the simpler point transformation, Eq. (14). For each fixed

$$x^{*a} = x^a + tV^a(x) \tag{14}$$

point $P(x)$, Eq. (14) represents a parametrized arc through P, and the tangent vector at P has the components $\partial x^{*a}/\partial t|P$ which are equal to $V^a(x)$ at P. In the literature Eq. (14) is frequently written in the modified form of Eq. (15), and referred to as an infinitesimal transformation.

$$x^{*a} = x^a + \delta t V^a(x) \tag{15}$$

With regard to the motion requirement: (1) for a parametrized arc C given by $x^a = x^a(u)$ and of class C', $dx^{*a}/du = dx^a/du + V^a{}_{;b}(dx^b/du)\, t$ and therefore $\partial/\partial t(dx^{*a}/du) = V^a{}_{;b}dx^b/du$; and (2) $g_{ab}|MP = g_{ab}(x^*) = g_{ab}(x^c + V^c t)$ and hence $\partial g_{ab}/\partial t = \partial g_{ab}/\partial x^{*c}V^c$, and $\partial g_{ab}/\partial t_{t=0} = g_{abc}(x)V^c$. Hence for the case of the two arcs C_1 and C_2, previously introduced, Eq. (16) follows. Thus the Killing equations

$$\begin{aligned}\left.\frac{d\left[g_{ab}(x^*)\dfrac{dx_1{}^{*a}}{du}\dfrac{dx_2{}^{*b}}{dv}\right]}{dt}\right|_{t=0} &= g_{ab;c}(x)V^c\frac{dx_1{}^a}{du}\frac{dx_2{}^b}{dv} + g_{ab}(x)V_{;c}{}^a\frac{dx_1{}^c}{du}\frac{dx_2{}^b}{dv} \\ &\quad + g_{ab}(x)\frac{dx_1{}^a}{du}V_{;c}{}^b\frac{dx_2{}^c}{dv} \\ &= (g_{ab;c}V^c + g_{cb}V_{;a}{}^c + g_{ac}V_{;b}{}^c)\frac{dx_1{}^a}{du}\frac{dx_2{}^b}{dv}\end{aligned} \tag{16}$$

constitute a necessary and sufficient condition for a point transformation, Eq. (14), to be an infinitesimal motion. *See* DIFFERENTIAL EQUATION; RIEMANNIAN GEOMETRY; TENSOR ANALYSIS. Homer V. Craig

Bibliography. I. Chavel, *Riemannian Geometry: A Modern Introduction*, 1995; L. P. Eisenhart, *Riemannian Geometry*, rev. ed., 1950, reprint, 1997; W. P. A. Klingenberg, *Riemannian Geometry*, 2d ed., 1995.

Kiln

A device or enclosure to provide thermal processing of an article or substance in a controlled temperature environment or atmosphere, often by direct firing, but occasionally by convection or radiation heat transfer. Kilns are used in many different industries, and the type of device called a kiln varies with the industry.

"Kiln" usually refers to an oven or furnace which operates at sufficiently high temperature to require that its walls be constructed of refractory materials. The distinction between a kiln and a furnace is often based more on the industry than on the design of the device. For instance, an electrically heated refractory tunnel oven equipped with a stainless mesh conveyor belt to carry the work through is referred to as a tunnel kiln if it is used for sintering small ceramic electronic parts such as ferrite transformer cores. The same device used to sinter small metal parts from powdered aluminum alloys is called a sintering furnace.

Generally the word "kiln" is used when referring to high-temperature treatment of nonmetallic materials such as in the ceramic, the cement, and the lime industries. When melting is involved as in steel manufacture, the term "furnace" is used, as in blast furnace and basic oxygen furnace. In glass manufacture, the melting furnace is often called a glass tank when the process is continuous. *See* FURNACE CONSTRUCTION.

Rotary kiln. This is the largest type of kiln, being used for heating of loose bulk materials such as cement and lime. Rotary kilns are used exclusively for the production of cement and for 85% of the commercial quicklime produced in the United States. Over the years, rotary cement kilns have increased in size to gain greater capacity and productivity from a single unit: kilns having a shell as large as 21 ft (6.3 m) diameter by 700 ft (210 m) long have been constructed. This is probably approaching the practical size limit. *See* CEMENT; LIME (INDUSTRY).

A rotary kiln (**Fig. 1**) is a long, refractory-lined cylinder supported on steel hoops, or "tires," riding on rollers, or "trunnions." The kiln is inclined at a slight slope, often $^1/_2$–$^5/_8$ in. fall per foot (42–52 mm per meter), and rotated at a slow speed (about 0.8–1.0 rpm for cement and lime kilns). For calcining ores, there is a tendency to use flatter slopes and higher rotational speeds because of the improved heat transfer obtained. A burner (usually using pulverized coal, natural gas, fuel, oil, or a combination

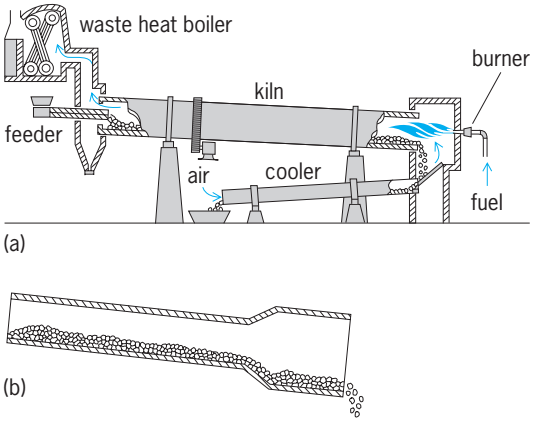

Fig. 1. Rotary kilns: (*a*) single-diameter; (*b*) two-diameter. (*After R. H. Perry and C. H. Chilton, Chemical Engineers' Handbook, 5th ed., McGraw-Hill, 1973*)

of fuels) is placed at the lower end to supply the necessary heat. Raw feed enters the upper end and slowly tumbles through the kiln, being heated by the countercurrent flow of hot combustion gases. Physical and chemical changes take place in the charge, such as drying and dehydration, decomposition of carbonates, and fusing of materials into a clinker.

A rotary kiln tends to have poor thermal efficiency. To conserve fuel, auxiliary devices are usually added. The hot product drops into a cooler which preheats the combustion air. Chain curtains at the feed end and feed-preheating devices are frequently used to extract heat from the flue gas and preheat the kiln feed.

Rotary kilns are also used for calcining or agglomerating phosphate rock; alumina and bauxite; magnesia; clays; carbon and petroleum coke; lightweight aggregates; iron, chromium, lithium, and uranium ores; burning of elemental sulfur; and incineration or pyrolysis of combustible wastes.

Shaft kiln. A shaft kiln (**Fig. 2**) consists of a vertical, stationary, refractory-lined steel tube, frequently tapered to grow slightly larger toward the bottom to permit free downward movement of the charge. These kilns have been used for centuries for calcining limestone and cement and are still used worldwide (except in the United States) for cement production; 15% of the commercial quicklime made in the United States is made in shaft kilns. Shaft kilns are now also used for calcining taconite pellets and as an oil shale retort. The raw charge is added at the top and is discharged through grates at the bottom. Originally, coal or coke was mixed with the charge to supply fuel and was burned out by hot air introduced through the walls above the grates.

Most shaft kilns are now equipped with external burners to supply the heat in order to eliminate product contamination caused by in-place firing of solid fuels. Shaft kilns have much better heat economy than rotary kilns since the gases flowing up through the kiln have more intimate contact with the solid charge. However, shaft kilns generally have lower production rates than rotary kilns, and accurate control of the charge time-temperature relationship is less precise because of gas channeling.

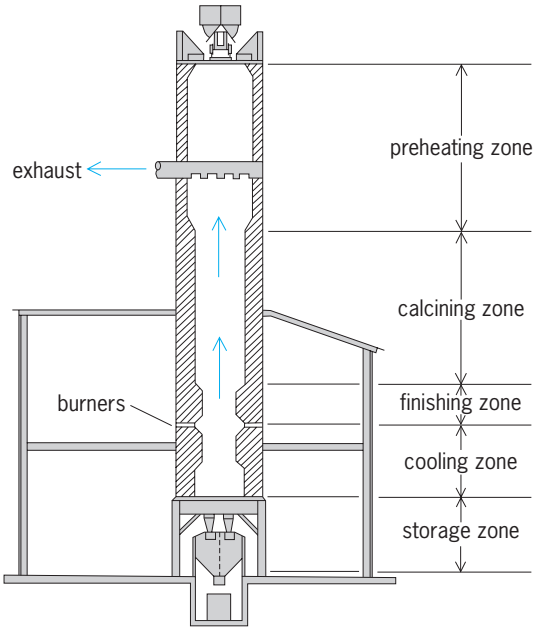

Fig. 2. Vertical shaft kiln. (*After C. J. Lewis and B. B. Crocker, The lime industry's problem of airborne dust, J. APCA, 19:31–39, 1969*)

Newer lime kilns. Two developments for lime calcining are the fluo-solids lime kiln, in which the limestone is calcined in a bed fluidized with hot combustion products, and the Calcimatic kiln, which is a horizontal rotating, circular refractory hearth kiln. The charge is deposited continuously on a pie-shaped element of the hearth as the hearth rotates about its center. The charge advances below radiant burners on the underside of a refractory dome roof. Rotational speed is set so that the charge is calcined when the hearth has rotated about 350°, at which point the charge is scraped from the hearth with a conveyor and transferred to a shaft cooler which preheats combustion air. One advantage of the Calcimatic is the low dust emission in the flue gas.

Kilns for formed ceramic products. Several types of kilns, either batch or continuous, are used for firing formed ceramic products. An early batch kiln is the scove kiln, in which unfired bricks are stacked to provide fireboxes and flues. Thus the basic material to be fired forms the kiln structure, which is covered on the outside with fired brick and a clay wash. Such kilns holding 30,000–200,000 bricks were fired with wood to produce all early American brick. Later, coal or oil fuel was used. *See* CERAMICS.

Periodic kiln. This is the general name given to a batch-type kiln in which the ware passes through a cycle consisting of heat-up, soaking (holding at peak temperature for some time), cooling, and removing or "drawing" the ware at the end of the cycle (**Fig. 3**). Such kilns vary in size from small laboratory or art-pottery kilns to large domed structures 40–50 ft (12–15 m) in size. Periodic kilns have poor thermal efficiency since all the heat used to heat the ware is lost during cooling. Such units are also wasteful of worker-power in loading and unloading.

Another early kiln, the beehive kiln, is a type of periodic kiln. It consists of a permanent, hemispherical chamber surrounded by a number of fuel beds, often burning coal. The hot gases pass up inside the walls and down through the material to be fired to a central flue in the floor.

Chamber kiln. This kiln, widely used in Europe but not generally adopted in the United States, is one of the earliest attempts to improve heat efficiency and approach continuous firing. The kiln is built in the form of a ring or long oval, divided into separate chambers connected with flues. All chambers are loaded with ware. Firing takes several days, the fire being started in one chamber and then moved progressively around the circuit. The combustion gas passes through chambers yet to be fired, where the ware is preheated. After completion of firing in a chamber, combustion air is passed through to cool the ware and preheat itself.

Tunnel kilns. These are long structures in which the ware is moved through the kiln (**Fig. 4**). Combustion air is introduced countercurrent to the unfired ware. Thus the air is preheated by cooling the fired ware. Fuel is burned in the center section, and the hot combustion gases preheat the incoming ware and remove its moisture. Tunnel kilns may be as small as a few inches wide and high by 15–30 ft (4.5–9 m)

Fig. 3. Interior of periodic kiln, with electrical porcelain insulators for firing. (*Swindell-Dressler, Inc.*)

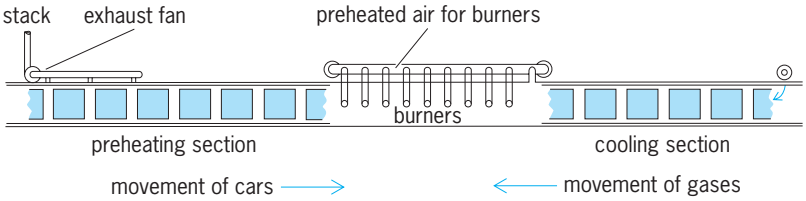

Fig. 4. Section of direct-fired tunnel kiln showing principle of operation. (*After T. J. W. van Thoor, ed., Chemical Technology: An Encyclopedic Treatment, vol. 2, Barnes and Noble, 1971*)

long for small ceramic parts to 6 by 8 ft (1.8 by 2.4 m) in cross section and several hundred feet long for continuous brick processing.

A tunnel kiln is often named after the method of moving the work through the kiln. Brick is loaded on a refractory-topped small car which rolls through the tunnel on a track. The refractory hearth of the car forms the bottom of the kiln. Such kilns are called car tunnel kilns. The car is frequently pushed into the tunnel with a hydraulic ram. Movement of the cars may be either continuous or intermittent at spaced time intervals. Small tunnel kilns may have a traveling mesh belt (mesh belt tunnels), or the ware may be placed in metal or refractory boxes called saggers. The boxes may be pushed through on a stationary hearth (pusher tunnel) or may roll on top of ceramic rollers (roller hearth tunnel).

Muffle kiln. In this type of kiln, which may be either periodic or continuous, the ware is protected from direct contact with the fire by a radiant refractory wall if the kiln is fuel-fired. In many special small muffle tunnels, electric heat is used. The tunnel refractory lining may be wound with a chrome-nickel resistance element, or silicon carbide glowbar units may be suspended from the roof and buried in the hearth. A muffle kiln may be used for firing high-priced dinnerware. Special electrically heated muffle tunnels have been developed for the electronics industry to fire ceramic capacitors, ceramic magnets, and printed ceramic circuits, often with carefully controlled atmospheres such as dry hydrogen.

Lumber and tobacco kilns. Kilns used in these industries are basically low-temperature dryers and do not require refractories for their construction. Lumber-drying kilns are of two basic types: a box or compartment type in which the lumber is loaded in batches for drying, and a progressive kiln in which the lumber travels through on cars much as in a tunnel kiln. The lumber is stacked horizontally with "stickers" (uniform wood spacers) between successive layers. Air of controlled temperature and humidity is passed over and through the lumber, often with circulating fans. Older kilns may have natural circulation produced by rising currents of hot air from steam heating coils. Fan kilns have better thermal efficiency because the air is recirculated more times before venting. Air temperatures vary from 150 to 200°F (66 to 93°C), depending on lumber size and species.

A tobacco-drying kiln is an airtight barn used for flue-curing of tobacco. Leaf strings of tobacco are hung from racks. Wood, coal, oil, or natural gas is burned in a firebox, or "kiln." The hot flue gases circulate through sheet-iron flues near the floor of the barns. The leaf is subjected to a carefully controlled temperature-time cycle (usually 3–5 days long) to produce the desired characteristics in the leaf.

Heating. Gas and oil have been the preferred fuels for all types of kilns and will probably remain so for moderate-priced ceramics. Pulverized coal has been used in the past for firing rotary kilns and will probably become the preferred fuel of the future because of the scarcity of hydrocarbon fuels. Coal has been burned on an external grate to fire shaft kilns when oil or gas is not available. Coal gas and producer gas have also been generated for this purpose. Electric heat is generally limited to smaller kilns processing valuable wares where precise temperature control is needed or where convenience is important, as in laboratory and pilot plant kilns.

Pollution control. Kilns can be sources of air pollution arising either from the combustion of the fuel or from entrainment of dust particles from the charge. The latter source applies only to kilns where loose bulk materials are processed, such as rotary and shaft kilns. The exhaust gases from these kilns are relatively hot, and cyclonic collectors have often been used to remove dust particles down to 5–10 micrometers in size. This collection does not meet today's requirements, and such kilns are being equipped with electrostatic precipitators or bag filters. Wet scrubbing can be used efficiently, but this results in a condensing steam plume in cool weather, which may be esthetically objectionable.

Combustion pollution can take the form of incompletely burned fuel (hydrocarbons, carbon monoxide, soot, and smoke) or oxides of nitrogen. The former is prevented by good combustion practices—adequate excess air, proper mixing of fuel and air, and adequate combustion temperatures. Nitrogen oxides are produced only when combustion is carried out at very high temperatures (usually 2000–2800°F or 1093–1538°C) to produce high peak temperatures in the kiln. NO_x is controlled by reducing flame temperature, using less excess air, two-stage combustion, and recirculation of exhaust gases. *See* AIR POLLUTION; DUST AND MIST COLLECTION.

Burton B. Crocker

Bibliography. L. Culpepper, *High Temperature Drying: Enhancing Kiln Operations*, 1990; F. Olsen, *The Kiln Book*, 3d ed., 2001; K. Peray, *Rotary Cement Kiln*, 2d rev. ed., 1986; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; W. Trinks and W. H. Mawhinney, *Industrial Furnaces*, vol. 1, 5th ed., 1961, vol. 2, 4th ed., 1967.

Kimberlite

A variety of peridotite, an igneous rock containing at least 35% olivine. The full definition is unusually complex for an igneous rock, with many qualifications. Briefly, kimberlite is richer in carbon dioxide than most peridotites, and has crystals larger in diameter than 0.5 mm of olivine, garnet, clinopyroxene, phlogopite, and orthopyroxene. All of these silicate minerals have high Mg/Fe ratios in kimberlites. In addition, there can be 0.5 mm or larger crystals of magnesian ilmenite and titanium-poor chromite. The matrix (crystals less than 0.5 mm) contains olivine plus one or more of monticellite, phlogopite, perovskite, spinel, apatite, serpentine, and carbonate minerals. Melilite, leucite, kalsilite, and alkali feldspar do not crystallize from kimberlite magma. In part, the complexity of this definition is caused by the

complicated processes of generation and emplacement of kimberlites and, in part, a restrictive definition is imposed as a defense against wishful thinking, because kimberlite was long thought to be the only rock that brought diamonds to the Earth's surface. However, in the 1970s other diamond-bearing rocks (lamproites) were found in western Australia and elsewhere, and in the 1980s R. H. Mitchell distinguished orangeite and kimberlite as separate hosts for diamonds in South Africa. Some eclogites have recently been found to contain diamonds also. *See* ECLOGITE; OLIVINE; PERIDOTITE.

Diamonds are the only economically significant mineral extracted from kimberlite. They form deeper than 150 km in the Earth's mantle and are carried upward as "accidental tourists" in kimberlite. Several thousand kimberlite occurrences are known, but less than 10% of these contain any diamonds, and less than 1% can support profitable mines. *See* DIAMOND.

The magmatic liquid that forms kimberlite is generated by the melting of small amounts of the Earth's upper mantle containing water and carbonate. The liquid moves upward, gathering crystals (including diamond) and rock fragments along the way. The high content of carbon dioxide dissolved in kimberlite magma becomes insoluble at lower pressures as the kimberlite ascends; carbon dioxide begins to form a separate gas phase at a depth of about 120 km, but most of the gas comes out at depths of 30 km and less. By the time the kimberlite breaks through the surface of the Earth, it is a suspension of liquid droplets and solid fragments in rapidly expanding gas. No lava flows representing liquid kimberlite that erupted on the surface have been recognized anywhere. This suggests that whenever kimberlite magma breaches the surface it does so explosively.

Kimberlites usually occur in regions of thick and stable continental crust, in southern Africa (including the Kimberley district), India, Siberia, Canada, Colorado-Wyoming, Venezuela, and Brazil. Most kimberlite outcrops appear on the surface as small, roughly circular areas less than 1 km in diameter; they are usually not well exposed because kimberlite weathers rapidly. In three dimensions, kimberlite bodies are dikes or, more commonly, downward-tapering cylinders (pipes). Those that have not been deeply eroded preserve their eruption craters, flaring out at the top of the pipe like the bell of a trumpet, filled with fragmental rock (kimberlite plus anything encountered on the way).

Because kimberlites usually do not crop out boldly, "indicator minerals" are important in prospecting for them. These are much more abundant than diamonds but occur with them in kimberlites and are stable under conditions of surface weathering, erosion, and deposition. They include pyrope garnet, ilmenite, and clinopyroxene, each with a distinctive composition. Using indicator minerals, one can trace debris from weathered kimberlite back to its source, even if that source has been buried.

During the violent injection of kimberlites into the upper crust, some detached fragments from the crust move downward and others from the lower crust and mantle move upward. Therefore kimberlite bodies are important scientifically because they contain fragments of rocks that were once above the present-day erosion surface as well as fragments of the Earth's mantle from 150 km deep or more. *See* IGNEOUS ROCKS. Daniel S. Barker

Bibliography. A. J. A. Janse and P. A. Sheahan, Catalogue of world wide diamond and kimberlite occurrences: A selective and annotative approach, *J. Geochem. Explor.*, 53:73–111, 1995; R. H. Mitchell, *Kimberlites: Mineralogy, Geochemistry, and Petrology*, Plenum, New York, 1986; R. H. Mitchell, *Kimberlites, Orangeites, and Related Rocks*, Plenum, New York, 1995.

Kinematics

That branch of mechanics which deals with the motion of a system of material particles without reference to the forces which act on the system. Kinematics differs from dynamics in that the latter takes these forces into account. *See* DYNAMICS.

For a single particle moving in a straight line (rectilinear motion), the motion is prescribed when the position x of the particle is known as a function of the time. For uniformly accelerated motion, Eq. (1)

$$\frac{d^2x}{dt^2} = a \tag{1}$$

holds, where a is the constant acceleration; x is defined as in Eq. (2), where x_0 and v_0 are the initial po-

$$x = {}^1/_2 at^2 + v_0 t + x_0 \tag{2}$$

sition and velocity of the particle, respectively. *See* ACCELERATION; RECTILINEAR MOTION; VELOCITY.

Plane kinematics of a particle is concerned with the specification of the position of a particle moving in a plane by means of two independent variables, usually the rectangular cartesian coordinates x and y, but often the polar coordinates $r = (x^2 + y^2)^{1/2}$, $\theta = \tan^{-1}(y/x)$, or even other coordinates that may be especially convenient for a particular problem. Here the x and y components of velocity are dx/dt, dy/dt, and of acceleration d^2x/dt^2, d^2y/dt^2, respectively, but the components of velocity in the directions of increasing r and increasing θ are dr/dt and $r(d\theta/dt)$, and are therefore not so simply given in terms of the time derivatives of the coordinates.

The kinematics of a particle in space is concerned with the ways in which three independent coordinates may be chosen to specify the position of the particle at a given time, and with the relations between the first and second time derivatives of these coordinates and the components of velocity and acceleration of the particle. The motion is specified if three such coordinates are given as functions of the time, and the path of the particle in space is then obtained by eliminating the time between these equations.

For describing the simultaneous position in space of N particles, $3N$ coordinates are necessary, and the

configuration of the system may be represented by a point in a $3N$-dimensional space called the configuration space of the system. As the particles move, the corresponding representative point in configuration space traces out a curve called the trajectory of the system. If the particles do not move independently but are subject to m constraints, that is, if m relations between the $3N$ coordinates continue to hold throughout the motion, then it is possible to represent the trajectory in a space of $f = (3N - m)$ dimensions, f being the number of degrees of freedom of the system. For a rigid body, even though N and m are very large, $f = 6$, and the configuration at any time may be represented by six independent variables (usually three to represent the position of the center of mass and three angles to specify the orientation of the body). Kinematics describes the relations between the time derivatives of such angles and the components of the angular velocity of the body. *See* CAYLEY-KLEIN PARAMETERS; CONSTRAINT; DEGREE OF FREEDOM (MECHANICS); EULER ANGLES.

Among the coordinate systems studied in kinematics are those used by observers who are in relative motion. In nonrelativistic kinematics the time coordinate for each such observer is assumed to be the same, but in relativistic kinematics proper account must be taken of the fact that lengths and time intervals appear different to observers moving relative to each other. *See* RELATIVITY.

H. C. Corben; Benard Goodman

Kinetic methods of analysis

The measurement of reaction rates for the analytical determination of the initial concentrations of the species of interest taking part in chemical reactions. This technique can be used since, in most cases, the rates or velocities of chemical reactions are directly proportional to the concentrations of the species taking part in the reactions.

The rate of a chemical reaction is measured by experimentally following the concentration of some reactant or product involved in the reactions as a function of time as the mixture proceeds from a nonequilibrium to an equilibrium or static state (steady state). Thus, kinetic techniques of analysis have the inherent problem of the increased experimental difficulty of making measurements on a dynamic system, as time is now a variable which is not present in making measurements on an equilibrium system. However, kinetic methods often have advantages over equilibrium techniques in spite of the increased experimental difficulty. For example, the equilibrium differentiations or distinctions attainable for the reactions of very closely related compounds are often very small and not sufficiently separated to resolve the individual concentrations of a mixture without prior separation. But the kinetic differentiations or distinctions obtained when such compounds are reacted with a common reagent are often quite large and permit simultaneous analysis. *See* CHEMICAL EQUILIBRIUM.

A further advantage of kinetic methods is that they permit a larger number of chemical reactions to be used analytically. Many reactions, both inorganic and organic, are not sufficiently well behaved to be employed analytically by equilibrium or thermodynamic techniques. Many reactions attain equilibrium too slowly; side reactions occur as the reactions proceed to completion, or the reactions are not sufficiently quantitative (do not go to completion) to be applicable. However, a kinetic-based technique can often be employed in these cases simply by measuring the reaction rate of these reactions during the early or initial portion of the reaction period. Also, the measurement of the rates of catalyzed reactions generally is a considerably more sensitive analytical method for the determination of trace amounts of a large number of species than equilibrium methods. *See* CATALYSIS.

Kinetic methods of analysis can be divided into four basic categories: (1) methods employing uncatalyzed reactions, (2) methods employing catalyzed reactions, (3) enzyme-catalyzed reaction methods, and (4) methods for the simultaneous in-place determination of mixtures.

Uncatalyzed reactions. In order to determine the concentration of a single species in solution, the rate is measured of an irreversible uncatalyzed reaction of the type shown by reaction (1), where the single

$$A + R \xrightarrow{k_A} P \tag{1}$$

species A reacts with the reagent R to form a product or products P; k_A is the rate constant for the reaction. In general, the types of reaction employed for kinetic analysis are second-order irreversible reactions. The rate of the reaction is most conveniently measured by following the rate of formation of the product P as a function of time (although the change in concentration of either A or R could also be followed). Any method of measuring concentrations, such as titrimetry, spectrophotometry, and polarography, can be employed. The rate of formation of the product, $d[P]_t/dt$, is given by differential equation (2).

$$\frac{d[P]_t}{dt} = k_A([R]_0 - [P]_t)([A]_0 - [P]_t) \tag{2}$$

In this relation $[A]_0$ is the initial or original concentration of the species to be determined, $[R]_0$ is the initial concentration of the reagent (added in known concentration), and $[P]_t$ is the concentration of product formed at any time t; thus, the terms $([R_0 - [P]_t)$ and $([A_0 - [P]_t)$ equal the concentration of A and R, respectively, remaining at any time t, and k_A is the second-order rate constant for the reaction of A with R.

If the rate of reaction (1) is measured only during the initial portion of the reaction (during a time period chosen so that the reaction is only 2–3% complete), the concentration of product formed $[P]_t$ is small compared to $[A]_0$ and $[R]_0$. Thus, on rearrangement, Eq. (2) becomes Eq. (3). of Eq. (3) Integration

$$[A]_0 = \frac{d[P]_t/dt(\text{initial})}{k_A[R]_0} \tag{3}$$

from the time interval $t = 0$ to t and for an initial concentration of $P = 0$ at $t = 0$ gives Eq. (4). Both fixed-

$$[A]_0 = \frac{[P]_t/t}{k_A[R]_0} \tag{4}$$

and variable-time methods may be used to calculate $[A]_0$ when employing Eq. (4).

When using the fixed-time method, the concentration of P formed at a chosen fixed time t_f is measured experimentally. As t_f and $[R]_0$ are thus kept constant for all measurements and k_A is a constant for all conditions, the term is the proportionality constant relating $[A]_0$ to $[P]_{tf}$, as shown in **Fig. 1**. For the variable-time method, with $[R]_0$ constant for all solutions, the time $t_{[P]f}$ necessary for the reaction to form a fixed amount of product $[P]_f$ is measured. As $[P]_f$ is now a constant, $[A]_0$ is related to $1/t_{[P]f}$ by the proportionately constant, $P_f/k_A[R]_0$, as shown in **Fig. 2**.

Catalyzed reactions. A catalyst may be broadly defined as an agent that alters the rate of a chemical reaction without shifting the equilibrium of the reaction. Although the catalyst undoubtedly enters into the reaction mechanism at a critical state, it does so in a cyclic manner and hence does not undergo a permanent change. Therefore a catalyst speeds up the rate of attainment of the equilibrium of a system, but it does not change the position of the equilibrium and it is not consumed during the reaction.

In many reactions involving a substance that acts as a catalyst, the concentration of the catalyst is directly proportional to the rate of the reaction. Thus the rates of these reactions can be employed for the analytical methods and are extremely sensitive, since the catalytic agents are not consumed but participate in the mechanism in a cyclic manner. The amount of the catalyst that can be determined employing catalyzed reactions is several orders of magnitude smaller than can be found by most direct equilibrium methods. Also, in many cases—enzyme-catalyzed (biological catalyst) reactions in particular—the catalyzed reaction of a reactant is extremely specific with respect to the chemical nature of the reactant. Such reactions, therefore, can be employed for the in-place analysis of a particular reactant in the presence of either a large excess or a large number of other species which would interfere with conventional equilibrium techniques, unless separation was performed prior to the analysis reaction.

The general mechanism for catalyzed reactions involves the combination of the catalyst C and the reactant, called the substrate S, to form a "complex" X, which then decomposes to form the product. Thus the catalyst is regenerated and combines once more with a substrate molecule. The general mechanism is given by reactions (5) and (6), where R is a reagent

$$C + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} X \tag{5}$$

$$X + R \xrightarrow{k_2} P + C \tag{6}$$

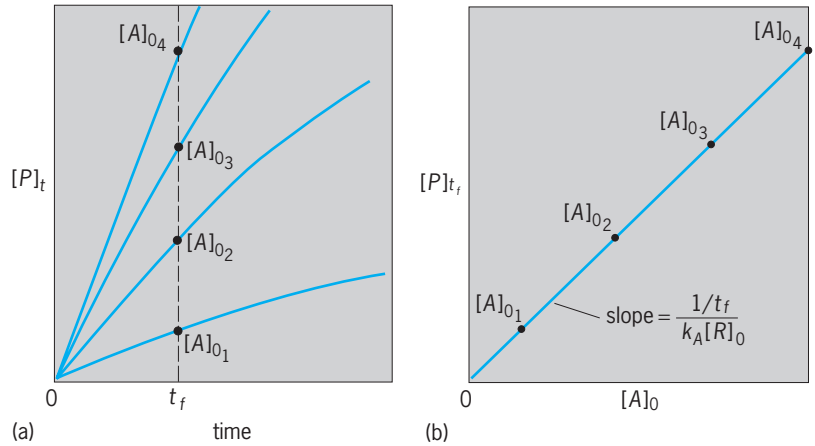

Fig. 1. Fixed-time method for uncatalyzed reactions. (*a*) Variation of [*P*] as a function of time for different initial concentrations of $[A]_0$. (*b*) Variation of amount of *P* formed at fixed time t_f with initial concentration of *A*.

molecule (which in many catalyzed reactions is not needed) that reacts with the complex to give the product P, and k_1 and k_{-1} are the rate constants for the formation of the complex and for the conversion of the complex back to C and S, respectively; the equilibrium constant K for reaction (5) is equal to k_1/k_{-1}, and k_2 is the rate constant for the decomposition of X to form the product and C, as in reaction (6). In many catalyzed reactions the rate of the decomposition reaction (6) is slow compared to the rate of the conversion back to C and $S(k_2 \ll k_{-1})$. Thus, in this case, reaction (6) is the rate-determining step. Under these conditions, the rate of reaction of S in a catalyzed process is given by Eq. (7).

$$-\frac{d[S]_t}{dt} = \frac{k_2[C]_0[S]_0[R]_t}{[S]_0 + 1/K} \tag{7}$$

When the equilibrium constant K is relatively small, so that $[S]_0$ is small compared to $1/K$, and when R is in a nonlimiting excess (or nonexistent) such that its concentration does not change during the initial 2–3% of the reaction, Eq. (7) becomes Eq. (8),

$$-\frac{d[S]_t}{dt} = (K')[C]_0[S]_0 \tag{8}$$

where the constant K' equals $k_2K[R]_0$. Thus the initial

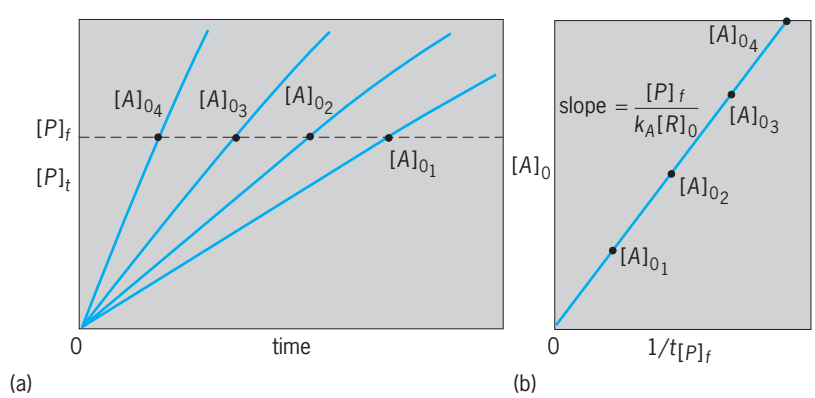

Fig. 2. Variable-time method for uncatalyzed reactions. (*a*) Variation of [*P*] as a function of time for various initial concentrations of $[A]_0$. (*b*) Variation of $1/t_{[P]f}$ required to reach a fixed value of product formed $[P]_f$ with initial concentration of *A*.

rate, $-(d[S]_t)/(dt)$, is directly proportional to the initial concentration of either the catalyst or the substrate, and the initial concentration of either can be determined. If a catalyst is to be determined, the substrate is added to each sample in nonlimiting excess (so that the concentration of S does not change during the initial 2–3% of the reaction). If a substrate is to be measured, the catalyst concentration added is not critical in any way because the catalyst is not consumed during the reaction. The choice of the catalyst concentration is dictated only by the range of initial rates that are conveniently measured. Either the fixed- or variable-time method of analysis can be employed by using catalyzed reactions.

Enzyme-catalyzed reactions. An enzyme is a biological catalyst, and this type of substance exhibits the ultimate in chemical selectivity. Enzyme-catalyzed reactions are used analytically for the determination of enzyme activities as well as selective substrate concentrations in complex biological and other mixtures. Both parameters are obviously of great importance in clinical diagnoses. The number of kinetic-based enzymatic determinations carried out daily in clinical laboratories is responsible for the fact that kinetic-based methods of analysis far outnumber those carried out by equilibrium-based methods or direct instrumental measurement.

In general, the simplified mechanism for enzyme-catalyzed reactions can be as shown in reaction (9),

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES^{+} \xrightarrow{k_2} P + E \qquad (9)$$

where E = enzyme, S = substrate, ES^+ = the activated addition complex, and P = products. Equation (10) is the differential rate equation for reaction (9).

$$-\frac{d[S]_t}{d_t} = \frac{d[P]_t}{d_t} = \frac{k_2[E]_0[S]_t}{K_m + [S]_t} \qquad (10)$$

$$K_m = \frac{k_{-1} + k_2}{k_1}$$

As in catalyzed reactions, experimental conditions can be adjusted for either substrate analysis or enzyme analysis employing the initial portion of the reaction (only 2–3% complete). Again, either the fixed- or variable-time method of analysis can be employed by using enzyme-catalyzed reactions. *See* CLINICAL PATHOLOGY; ENZYME.

Simultaneous determinations. Often the thermodynamic properties of closely related species are such that conventional equilibrium analytical techniques are unable to resolve the analytical concentrations of the components of the mixture without prior separation. However, in general the rates of reaction of such species in a mixture are sufficiently different to enable one to determine the initial concentration of each species without resorting to separation, which is a considerable saving of time and labor. Such techniques are called differential kinetic analysis methods. There are a large number of differential kinetic methods applicable to both first- and second-order reactions. Each of these methods has special conditions under which it is most applicable. The principles of these methods are complex, and an explanation of each is beyond the scope of this section. However, the detailed explanation of the logarithmic extrapolation method given below illustrates the general concepts of differential kinetic methods.

Consider two competing second-order irreversible reactions of the type shown by (11) and (12),

$$A + R \xrightarrow{k_A} P \qquad (11)$$

$$B + R \xrightarrow{k_B} P \qquad (12)$$

where k_A and k_B are the second-order rate constant for the reaction of A and B with R, respectively. If the concentration of R is 50 to 100 times greater than the sum of the concentration of A plus B (the components of interest in the analysis), the reaction is pseudo first-order because $[R]_t$ is constant and is equal to the initial concentration $[R]_0$. The differential rate expressions for the reactions given by (11) and (12) then become Eqs. (13) and (14).

$$-\frac{d[A]_t}{dt} = k_A[R]_0[A]_t = k'_A[A]_t \qquad (13)$$

$$-\frac{d[B]_t}{dt} = k_B[R]_0[B]_t = k'_B[B]_t \qquad (14)$$

The sum of the concentrations of A and B which react competitively to form a common product P is given at any time t by the rate expression, as shown by Eq. (15).

$$\frac{d[P]_t}{dt} = -\left[\frac{d[A]_t}{dt} + \frac{d[B]_t}{dt}\right]$$

$$= k'_A[A]_t + k'_B[B]_t \qquad (15)$$

On integration of Eq. (15) between the limits $t = 0$ and $t = \infty$ Eq. (16) is formed. In the case where

$$[P]_\infty - [P]_t = [A]_t + [B]_t$$
$$= [A]_0 e^{-k'_A t} + [B]_0 e^{-k'_B t} \qquad (16)$$

the rate of reaction of component A is larger than that of B, the term $[A]_0 e^{-k'_A t}$ eventually becomes very small compared to $[B]_0 e^{-k'_B t}$ at some time t after A has reacted essentially to completion ($[A_t] \cong 0$) and can be considered negligible. Thus, by taking the logarithm of both sides of Eq. (16), Eq. (17), is

$$\ln([A]_t + [B]_t) = \ln([P]_\infty - [P]_t)$$
$$= -k'_B t + \ln[B]_0 \qquad (17)$$

obtained, which predicts that a plot $\ln([A]_t + [B]_t)$ or $\ln([P]_\infty - [P]_t)$ versus time t will yield a straight line with a slope of $-k'_B$ and an intercept (at $t = 0$) equal to $\ln [B]_0$. The value of $[A]_0$ is then obtained by subtracting $[B]_0$ from the total initial concentration of the mixture, $[A]_0 + [B]_0$, which must be determined either by independent methods or from $[P]_\infty$. A typical reaction rate curve of this type is illustrated in **Fig. 3**. *See* CHAIN REACTION (CHEMISTRY); CHEMICAL DYNAMICS; ELECTROCHEMICAL TECHNIQUES;

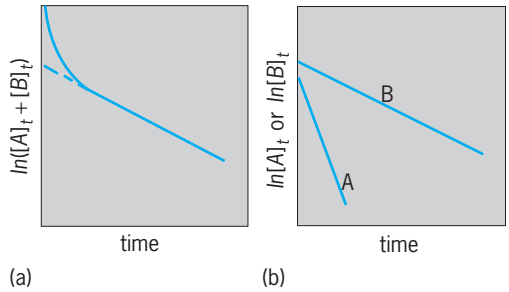

Fig. 3. Logarithmic extrapolation method for a mixture of species *A* and *B* reacting by first-order kinetics. (*a*) Rate data obtained for the mixture. (*b*) Rate data obtained for each component separated.

INHIBITOR (CHEMISTRY); KINETIC THEORY OF MATTER.

Harry B. Mark, Jr.

Bibliography. S. W. Benson, *Foundations of Chemical Kinetics*, 1960, reprint 1982; W. Drenth and H. Kwart, *Kinetics Applied to Organic Reactions* 1980; S. Kotrly and S. Ladislav, *Handbook of Chemical Equilibria in Analytical Chemistry,* 1985; H. B. Mark, Jr., and G. A. Rechnitz, *Kinetics in Analytical Chemistry*, 1968; H. A. Mottola, *Kinetic Aspects of Analytical Chemistry*, 1988; D. Pérez-Bendito and M. Silva, *Kinetic Methods in Analytical Chemistry*, 1988; R. W. Ramette, *Chemical Equilibrium and Analysis*, 1981.

Kinetic theory of matter

A theory which states that the particles of matter in all states of aggregation are in vigorous motion. In computations involving kinetic theory, the methods of statistical mechanics are applied to specific physical systems. The atomistic or molecular structure of the system involved is assumed, and the system is then described in terms of appropriate distribution functions. The main purpose of kinetic theory is to deduce, from the statistical description, results valid for the whole system. The distinction between kinetic theory and statistical mechanics is thus of necessity arbitrary and vague. Historically, kinetic theory is the oldest statistical discipline. Today a kinetic calculation refers to any calculation in which probability methods, models, or distribution functions are involved.

For information which is related to and supplements the present article *see* BOLTZMANN STATISTICS; BOLTZMANN TRANSPORT EQUATION; QUANTUM STATISTICS; STATISTICAL MECHANICS.

Classes of problems. Kinetic calculations are not restricted to gases, but occur in chemical problems, solid-state problems, and problems in radiation theory. Even though the general procedures in these different areas are similar, there are a sufficient number of important differences to make a general classification useful.

Classical ideal equilibrium problems. In these, there are no interactions between the constituents of the system. The system is in equilibrium, and the mechanical laws governing the system are classical. The basic information is contained in the Boltzmann distribution f (also called Maxwell or Maxwell-Boltzmann distribution) which gives the number of particles in a given momentum and positional range ($d^3x = dx\,dy\,dz$, $d^3\upsilon = d\upsilon_x\,d\upsilon_y\,d\upsilon_z$, where x, y, and z are coordinates of position, and υ_x, υ_y, and υ_z are coordinates of velocity). In Eq. (1) ϵ is the energy,

$$f(xyz, \upsilon_x\upsilon_y\upsilon_z) = Ae^{-\beta\epsilon} \tag{1}$$

$\beta = 1/kT$ (where k is the Boltzmann constant and T is the absolute temperature), and A is a constant determined from Eq. (2). The calculations of gas pressure,

$$\int\int d^3x\, d^3\upsilon\, f = N \quad \text{total number of particles} \tag{2}$$

specific heat, and the classical equipartition theorem are all based on these relations.

Classical ideal nonequilibrium problems. Many important physical properties refer not to equilibrium but to nonequilibrium states. Phenomena such as thermal conductivity, viscosity, and electrical conductivity all require a discussion starting from the Boltzmann transport equation, Eq. (3*a*), where $\mathbf{X}$ is the force

$$\frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla f + \mathbf{X}\cdot\nabla_\upsilon f = \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \tag{3a}$$

$$Z_{cl} = e^{-\Psi/kT} = \frac{1}{N!\lambda^{3N}}\int\cdots\int d^3x_1 \cdots d^3x_N \exp[-(1/kT)U_{(x_1,\ldots,z_N)}] \tag{3b}$$

per unit mass, ∇ is the gradient on the space-dependence, ∇_υ is the gradient on the velocity-dependence, and $(\partial f/\partial t)_{\text{coll}}$ is the change due to collisions. If one deals with states that are near equilibrium, the exact Boltzmann equation need not be solved; then it is sufficient to describe the nonstationary situation as a small perturbation superimposed on an equilibrium state. Even though the rigorous discussion of the nonequilibrium processes is difficult, appeal to simple physical pictures frequently leads to quite tractable expressions in terms of the equilibrium distribution function. Of special importance is the example of the so-called electron gas in a metal. The kinetic treatment of this system forms the basis for the classical (Lorentz) conductivity theory. For states far from equilibrium, no general simple theory exists. *See* FREE-ELECTRON THEORY OF METALS.

Classical nonideal equilibrium theory. The basic classical procedure for arbitrary systems (systems with interactions taken into account) that allows the calculation of macroscopic entities is that using the partition function, as shown in Eq. (3*b*), where $\lambda = h/\sqrt{2\pi mkT}$. Here Ψ is the thermodynamic free energy, and h is Planck's constant. Although Eq. (3*b*) is written so that it may be applied to gases, the partition function may also be written for classical spin systems, such as ferromagnetic and paramagnetic solids. The mathematical problems of evaluating the integrals or the sums are difficult. Equation (3*b*), with appropriate modifications, is the

starting point for all these considerations. *See* FREE ENERGY.

Classical nonideal nonequilibrium theory. This is the most general situation that classical statistics can describe. In general, very little is known about such systems. The Liouville equation applies, and has been used as a starting point for these studies, but no spectacular results have yet been obtained. For studies of the liquid state, however, the results have been quite promising.

Quantum problems. There are quantum counterparts to the classifications just described. In a quantum treatment a distribution function is also used for an ideal system in equilibrium to describe its general properties. For systems of particles which must be described by symmetrical wave functions, such as helium atoms and photons, one has the Bose distribution, Eq. (4), where $\beta = 1/kT$, and A is determined by Eq. (2).

$$f(v_x v_y v_z) = \frac{1}{(1/A)e^{\beta\epsilon} - 1} \tag{4}$$

See BOSE-EINSTEIN STATISTICS.

For systems of particles which must be described by antisymmetrical wave functions, such as electrons, protons, and neutrons, one has the Fermi distribution, Eq. (5). Use of these functions in cal-

$$f(v_x v_y v_z) = \frac{1}{(1/A)e^{\beta\epsilon} + 1} \tag{5}$$

culations is actually not very different from the use of the classical distribution function, but the results are quite different, as are the analytical details. As in the classical case, the treatment of the nonequilibrium state can be reduced to a treatment involving the equilibrium distribution only. The application to electrons as an (ideal) Fermi-Dirac gas in a metal is the basis of the Sommerfeld theory of metals. *See* FERMI-DIRAC STATISTICS.

In quantum theory, nonideal systems in equilibrium are described in terms of the quantum partition function, Eq. (6). Here E_n indicates the energy levels

$$Z_g = \sum_n g_n e^{-E_n/kT} = e^{-\Psi/kT} \tag{6}$$

of the system, and g_n indicates the weights of these levels. If one defines a Slater sum by Eq. (7), where

$$S(x_1, \ldots, x_N) = \sum_n e^{-E_n/kT} \, |U_n(x_1, \ldots, x_n)|^2 \tag{7}$$

$U_n(x_1, \ldots, x_N)$ is the wave function of the state n, Z_g may be written as an integral similar to Z, as in Eq. (8). For the applications, the energy levels and

$$Z_g = \int \cdots \int d^3x_1 \cdots d^3x_N S(x_1, \ldots, x_N) \tag{8}$$

the wave functions must be known. In the evaluation of S, given by Eq. (7), the symmetry character of the wave functions must be explicitly introduced. It is sometimes easier to use the grand partition function, Eq. (9). Here μ is the chemical potential, and

$$Z_{q\cdot m\cdot gr} = \sum_N \sum_n e^{(\mu N - EN,n)/kT} \tag{9}$$

$E_{N,n}$ is the nth level of a system having N particles. The current theories of quantum statistics, as for example the hard-sphere Bose gas, are for the most part concerned with questions in this area.

The only technique now available is that of the density matrix. It is possible to express certain entities which characterize transport properties, such as the conductivity tensor, in terms of the unperturbed stationary density matrix. A complete discussion of the validity of the approximations is still lacking. In addition, once the conductivity tensor is obtained in terms of the density matrix of the stationary (but still interacting) system, a problem of the same order of difficulty as the evaluation of the quantum mechanical partition function remains, if explicit expressions for these quantities in terms of the forces between atoms are desired.

Classical examples. Kinetic theory gave the first insight into many of the phenomena that take place in gases as well as in metals, where the free (conduction) electrons can be considered as an ideal gas of electrons. The following examples illustrate some of the more fundamental calculations that have been made.

Ideal-gas pressure. A classical ideal gas is described by the Boltzmann distribution of Eq. (1). The constant A is given by Eq. (10), where m is the mass of an indi-

$$A = \frac{N}{V}\left(\frac{m\beta}{2\pi}\right)^{3/2} \tag{10}$$

vidual molecule and V is the total volume of the gas. A gas exerts a force on the wall by virtue of the fact that the molecules are reflected by it. The component of momentum normal to the wall changes its sign as a consequence of this collision. Hence if the normal velocity of the molecule is υ_n, the momentum given off to the wall is $2m\upsilon_n$. To calculate the total force on a wall, one needs the total momentum transferred to the wall per unit time. Let dS be a small section of the wall, and call θ the angle made by the molecule's velocity vector with the normal to dS (**Fig. 1**). From Eq. (1), the number of molecules per unit volume that have a speed $c = (\upsilon_x^2 + \upsilon_y^2 + \upsilon_z^2)^{1/2}$ and whose velocity vector makes an angle θ with a given axis (call these c, θ molecules) is given by Eq. (11*a*). The

$$f(c, \theta) = 2\pi A e^{(-1/2)\beta mc^2} c^2 \sin\theta \, d\theta dc \tag{11a}$$

$$dS \, c \cos\theta \, dt \, f(c, \theta) \tag{11b}$$

number of such molecules that in time dt will collide with dS (assuming spatial homogeneity) is given by notation (11*b*). From this information an interesting side result may be calculated, namely, that the number of all collisions with a unit area of the wall per second may be written as expression (12*a*). If one introduces the average speed $\bar{c}$, defined in Eq. (12*b*),

$$2\pi A \int_0^{\pi/2} d\theta \, \sin\theta \cos\theta \int_0^\infty dc \, c^3 e^{-(1/2)\beta mc^2} \tag{12a}$$

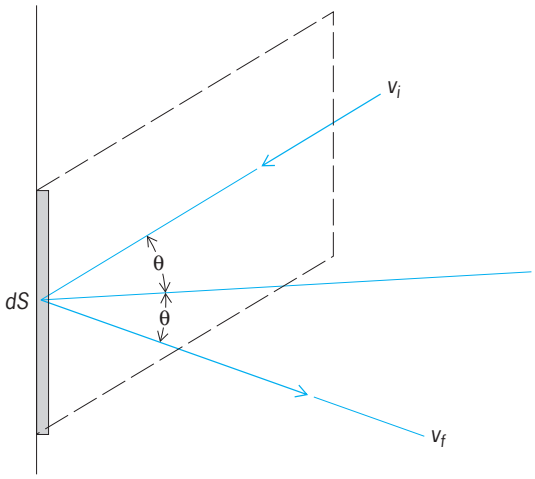

Fig. 1. Change of momentum of gas molecule as it strikes a wall; v_i, initial velocity; v_f, final velocity.

$$\bar{c} = \frac{1}{N} \int \int d^3x \, d^3v \, cf \tag{12b}$$

the total number of collisions may be written as $^1/_4 n\bar{c}$. Here $n = N/V$, the number density. This relation is of importance in calculating the efflux of gases through orifices. Each (c, θ) molecule contributes a momentum of $2mc \cos\theta$ to the wall per collision. Since the pressure p is given by the force per unit area, or the total momentum transferred per unit time per unit area, one obtains for p Eqs. (13) and (14). Since it

$$p = \int_0^\infty \int_0^{\pi/2} 2mc \cos\theta \cdot 2\pi Ac \cos\theta \, e^{-(1/2)\beta mc^2} c^2 \sin\theta \, d\theta \, dc \tag{13}$$

$$p = {}^1/_3 N m\overline{c^2} = (N/V)\beta^{-1} \tag{14}$$

is known experimentally that $pV = NkT$, where k is the Boltzmann constant, $\beta = 1/kT$ is now identified. This pressure calculation is a typical example of a kinetic theory calculation. *See* GAS.

Equipartition theorem and specific heat. In a classical ideal gas, the average kinetic energy associated with any translational degree of freedom is expressed by Eq. (15). It is quite straight forward to show that

$$^1/_2 m\overline{v_x^2} = {}^1/_2 m\overline{v_y^2} = {}^1/_2 m\overline{v_z^2} = \frac{1}{2\beta} = {}^1/_2 kT \tag{15}$$

the average energy associated with any degree of freedom, which occurs as a quadratic term in the expression for the mechanical energy, is given by $^1/_2 kT$. Formally, the result one proves is that, for a Boltzmann distribution written in terms of momenta and generalized coordinates p_i and q_i, one has Eq. (16), which is called the equipartition theorem.

$$\overline{q_i \frac{\partial \epsilon}{\partial q_i}} = \overline{p_i \frac{\partial \epsilon}{\partial p_i}} = kT \tag{16}$$

From the equipartition theorem the specific heat C, defined by Eq. (17), may be immediately obtained.

$$C = \frac{\partial \overline{E}}{\partial T} \tag{17}$$

If one has just the three translational degrees of freedom, the average energy per degree of freedom per molecule is $^1/_2 kT$; hence the average energy $\overline{E} = {}^3/_2 NkT$ and $C = {}^3/_2 R$. The specific heat is thus constant and independent of T. In the case of a diatomic molecule, one usually has three translational and two rotational degrees of freedom. Therefore $\overline{E} = {}^5/_2 NkT$, and $C = {}^5/_2 Nk \equiv {}^5/_2 R$. Suppose there are N atoms in a solid, each bound by elastic forces to a center. In that case, the mechanical energy will be expressed by Eq. (18), where k_F is the elastic force instant.

$$\epsilon = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right) + {}^1/_2 k_F^2 (x^2 + y^2 + z^2) \tag{18}$$

According to the equipartition theorem, these six terms in ϵ will give a specific heat $C = 3Nk$, the Dulong and Petit value for a monatomic solid. *See* SPECIFIC HEAT OF SOLIDS.

The equipartition theorem is especially useful when a gas is at a temperature so high that it becomes necessary to use relativistic mechanics to describe the system properly. This happens at temperatures for which the thermal energy kT is of the same order as mc^2, where c denotes the speed of light. The appropriate relation between energy and momentum is then given by Eq. (19*a*). The equipartition relation (16) may then be written as Eq. (19*b*). From this relation one can obtain the relativistic correction to the average energy of a gas, as shown in Eq. (19*c*).

$$\epsilon = c\left[\left(p_x^2 + p_y^2 + p_z^2\right) + m^2c^2\right]^{1/2} \tag{19a}$$

$$\frac{\overline{c^2 p_x^2}}{\epsilon} = kT \tag{19b}$$

$$\overline{E} = {}^3/_2 NkT\left(1 + \frac{5}{4}\frac{kT}{mc^2} + \cdots\right) \tag{19c}$$

See RELATIVISTIC MECHANICS.

Electrical conductivity of metals. Experimentally, a proportionality is observed between the applied electric field and the current produced in a metallic conductor. Kinetic procedures give an explanation of this general connection. Consider the electrons in a metal as an ideal gas of electrons of mass m and charge e. In equilibrium, the electrons are described by the Boltzmann distribution, Eq. (1), or the Fermi-Dirac distribution, Eq. (5). Call this distribution f. Because f depends on $v_x^2 + v_y^2 + v_z^2$ only, it is clear that $\bar{v}_x = 0$; no net current can flow in an equilibrium state. The application of an electric field therefore results in the destruction of the spherical symmetry in the velocities. This effect is described by the Boltzmann transport equation, Eq. (20).

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \mathbf{X} \cdot \nabla_v f = \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \tag{20}$$

The collision term causes considerable difficulty. In many problems one is justified in introducing a relaxation time τ which may depend on x and v and

which is defined by Eq. (21). Here f_0 is the distri-

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = -\frac{f - f_0}{\tau} \tag{21}$$

bution function at thermal equilibrium. The introduction of such a relaxation time presupposes that a nonequilibrium state will decay exponentially into an equilibrium state as a consequence of the action of the collisions. This is undoubtedly a good approximation near equilibrium. *See* RELAXATION TIME OF ELECTRONS.

Assume the existence of an electric field in the x direction which distorts the initial distribution a small amount. By using Eq. (21), Eq. (20) may be written as Eq. (22). Here E_e is the external electric field. If it

$$\frac{eE_e}{m}\frac{\partial f}{\partial v_x} + v_x\frac{\partial f}{\partial x} = -\frac{f - f_0}{\tau} \tag{22}$$

is now assumed that $(f - f_0)/f_0 \ll 1$, with f very near an equilibrium state, so that quadratic terms may be neglected, and that the parameters A and β in f_0 are independent of x, then an explicit expression for f can be obtained, as shown in Eq. (23).

$$f = f_0 - \tau e E_e v_x \frac{\partial f_0}{\partial \epsilon} \tag{23}$$

The electric current in the x direction is always given by Eq. (24).

$$j_x = \int e v_x f \, d^3v = -\tau e^2 E_e \int v_x^2 \frac{\partial F_0}{\partial \epsilon} d^3v \tag{24}$$

The first term in Eq. (23) does not contribute to the current, as has already been pointed out. This is an example of a formal result in transport theory. The result for j_x is in a form in which only a knowledge of the equilibrium distribution is required to obtain an explicit answer.

In the case of Boltzmann statistics, use of Eqs. (1), (2), (10), and (24) gives Eq. (25*a*). The conductivity

$$j_x = \frac{Ne^2\tau}{m}E_e \tag{25a}$$

$$f_0 = \frac{1}{e^{\beta(\epsilon-\mu)} + 1} \tag{25b}$$

$\sigma = Ne^2\tau/m$ cannot be compared with experiment unless the relaxation time τ is known. In the Fermi case the evaluation of Eq. (24) is facilitated by observing that $\partial f_0/\partial\epsilon$ has a δ-function character. Call $A = e^{\beta\mu}$, so that Eq. (5) reads as Eq. (25*b*).

It is easy to show from Eq. (25*b*) that expression (26) holds for sufficiently low temperatures. This

$$\frac{\partial}{\partial\epsilon}\frac{1}{e^{\beta(\epsilon-\mu)} + 1} \cong -\delta(\epsilon - \mu) \tag{26}$$

allows the immediate calculation of Eq. (24) also for the case in which the relaxation time depends on the velocity.

Viscosity and mean free path. One of the early successes of kinetic theory was the explanation of the viscosity of a gas. Strictly speaking, this is again a transport property, and as such it should be obtained from the Boltzmann transport equation, Eq. (20). It is possible,

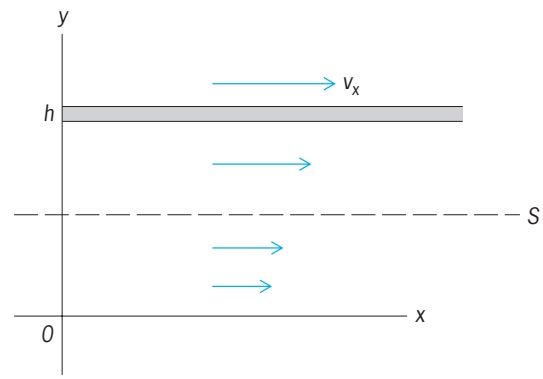

Fig. 2. Explanation of viscosity. The plate at $y = h$ is moving with velocity v_x, and the plate at $y = 0$ is stationary. The horizontal component of velocity of the gas molecules varies from 0 at $y = 0$ to v_x at $y = h$. Momentum therefore must be transferred across surface S by the vertical component of velocity.

however, to give an elementary discussion. Consider a gas that is contained between two walls or plates, the lower one ($y = 0$) at rest and the upper one constrained to move with a given velocity in the x direction (**Fig. 2**). A force is necessary to maintain the constant velocity of the plate. This force is given by Eq. (27). Here dv_x/dy is the velocity gradient, and

$$\mathbf{X}_x = \eta A \frac{dv_x}{dy} \tag{27}$$

$\mathbf{X}$ is the viscous force on the area A, which is perpendicular to the y axis ($\mathbf{X}_x/A$ is sometimes called the shear stress). Equation (27) defines η, the viscosity coefficient. The physical reason for this force stems from the fact that molecules above the surface S have a greater flow velocity than those below this surface. (This will certainly be true on the average.) Thus molecules crossing from above to below will carry a larger amount of momentum in the positive x direction than those crossing from below S upward. Hence the net effect is a transport of momentum in the x direction across the surface. By Newton's second law, this will yield a force. The computation can be carried out in this manner. Consider an area in the xz plane. The number of molecules passing through per second is $f\upsilon_y dS$, where f is the distribution function. The amount of momentum transported in the x direction is (per collision) $f\upsilon_y dS \cdot m\upsilon_x$.

The force per unit area is the sum of these terms, given by expression (28). For the evaluation of this

$$f v_y v_x m \tag{28}$$

sum, the notion of mean free path is useful. The mean free path is the average distance traveled by a molecule between collisions, and is usually designated by λ. To investigate this entity, imagine each molecule to be a hard sphere with radius a. If a molecule moves with average speed $\bar{c}$, it sweeps out a volume $4\pi a^2\bar{c}t$ in time t. If there are $n = N/V$ molecules per unit volume, the number of collisions per second is given by the collision frequency z in Eq. (29).

$$z = n4\pi a^2\bar{c} \tag{29}$$

For a typical gas (oxygen) under standard conditions $n = 3 \times 10^{25}$, $\bar{c} = 4.5 \times 10^4$ cm/s, and $a \cong 1.8 \times 10^{-12}$ cm. Hence, numerically, $z = 5.5 \times 10^9$ collisions/s. The average distance between collisions, that is, the mean free path, is given by Eq. (30*a*). Numerically, $\lambda \cong 8 \times 10^{-6}$ cm. This discussion is, of course, exceedingly crude. Making the calculation on the basis of a Boltzmann distribution gives Eq. (30*b*).

$$\lambda = \frac{\bar{c}t}{n(4\pi a^2)\bar{c}t} = \frac{1}{n(4\pi a^2)} \tag{30a}$$

$$\lambda = \frac{1}{\sqrt{2}n(4\pi a^2)} \tag{30b}$$

Using similar methods, it may be shown that the distribution of the mean free paths (that is, the number of molecules whose mean free path lies between x and $x + dx$) is given by Eq. (31).

$$dN = \frac{N_0}{\lambda} e^{-x/\lambda}\, dx \tag{31}$$

There is an interesting connection between the mean free path and the relaxation time introduced previously. It should be stressed, however, that this connection follows more from a qualitative discussion than from a rigorous calculation. One would guess that Eq. (32) is valid. This means that a relaxa-

$$\tau = \frac{\lambda}{\bar{c}} \tag{32}$$

tion time describes the decay from a state so near an equilibrium state that, when the molecules have traveled (on the average) one mean free path, the equilibrium is reestablished. Stated differently, the nonequilibrium state is, on the average, one collision per molecule removed from the equilibrium state.

Expression (28) may now be evaluated in terms of the mean free path as shown in Eq. (33*a*). From this, the viscosity coefficient follows directly, as shown in Eq. (33*b*). Introducing the mean free path, as given by Eq. (30*a*), one sees that Eq. (33*c*) holds. The re-

$$\text{Force per unit area} = {}^1/_3 nm\lambda\bar{c}\frac{dv_x}{dy} \tag{33a}$$

$$\eta = {}^1/_3 nm\bar{c}\lambda \tag{33b}$$

$$\eta = \frac{1}{3}\frac{m\bar{c}}{4\pi a^2} \tag{33c}$$

markable result is that the viscosity is indeed independent of the pressure. Since Eq. (34) holds, the

$$\bar{c} = \sqrt{8kT/\pi m} \tag{34}$$

viscosity depends on the temperature but not on the pressure. This somewhat unintuitive result was one of the first triumphs of kinetic theory. Both aspects of Eq. (33*c*) are in good agreement with experiment; for not too low densities η is indeed independent of the pressure; η is also proportional to the square root of the temperature. *See* VISCOSITY. Max Dresden

Bibliography. R. L. Liboff, *Kinetic Theory: Classical, Quantum and Relativistic Description*, 2d ed., 1998; P. C. Reidi, *Thermal Physics: An Introduction to Thermodynamics, Statistical Mechanics, and Kinetic Theory*, 2d ed., 1988; P. P. Schram, *Kinetic Theory of Gases and Plasmas*, 1991; L. C. Woods, *An Introduction to the Kinetic Theory of Gases and Magnetoplasmas*, 1993.

Kinetoplastida

An order of the class Zoomastigophorea in the phylum Protozoa, also known as Protomastigida, containing a heterogeneous group of colorless flagellates possessing one or two flagella in some stage of their life cycle. These small organisms (5–89 micrometers in length) typically have pliable bodies. Some species are holozoic and ingest solid particles, while others are saprozoic and obtain their nutrition by absorption. Their life cycles are usually simple but some species have two or more recognizably distinct stages. The species may be either free-living or parasites of vertebrates, invertebrates, and plants. Reproduction is by longitudinal fission, although multiple fission occurs in some species. Several are important disease-causing parasites of humans and domestic and wild animals.

Taxonomy. The order Kinetoplastida includes all of the Protozoa with only one or two flagella that are not markedly ameboid, do not contain chloroplasts, and are not closely related to chloroplast-bearing flagellates. Organisms that possess one or two flagella and are decidedly ameboid belong to the family Mastigamoebidae in the order Rhizomastigida. Those which are colorless but considered to be close relatives of protozoans possessing chloroplasts belong to the family Chlamydomonadidae of the order Volvocida. *See* VOLVOCIDA.

There is disagreement on the division of the order into families. However, the five or more families can be divided into two general groups. The first group contains simple organisms with no distinctive features save one or two flagella of equal or

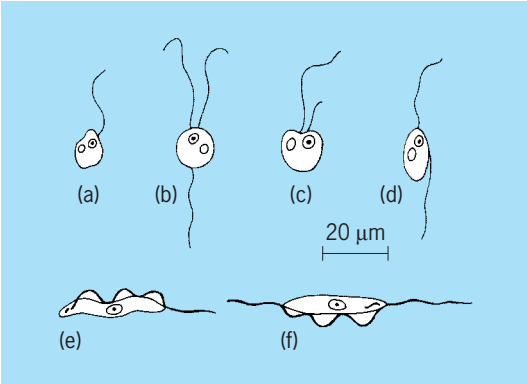

Representative genera of families of order Kinetoplastida. (*a*) *Oikomonas* (family Oikomonadidae), one anterior flagellum. (*b*) *Amphimonas* (family Amphimonadidae), two equally long anterior flagella. (*c*) *Monas* (family Monadidae), two unequally long anterior flagella. (*d*) *Bodo* (family Bodonidae), two unequally long flagella, one of them trailing. (*e*) *Trypanosoma* (family Trypanosomatidae), one flagellum with undulating membrane. (*f*) *Cryptobia* (family Cryptobiidae), two flagella, one free and one with undulating membrane.

unequal length. This includes the families Oikomonadidae, Amphimonadidae, Monadidae, and Bodonidae (**illus.** *a–d*). The second group contains organisms which have an undulating membrane in addition to one or two flagella. The families included in this group are Trypanosomatidae, and Cryptobiidae (illus. *e* and *f*). A third group, including organisms possessing a peculiar collar surrounding a single flagellum, once assigned to this order, now makes up the order Choanoflagellida. *See* CHOANOFLAGELLIDA.

Trypanosomatidae. The most important family of the Kinetoplastida is the Trypanosomatidae, since it includes several species that infect humans and domestic animals with serious diseases, such as African sleeping sickness. The organisms in this family are polymorphic, changing their form in various stages of their development. Their life cycles may involve one or two hosts (invertebrate, vertebrate, or plant). The trypanosome form possesses a single flagellum and an undulating membrane extending the full length of the body. The other related forms are simpler in that they lack one or both of these structures. *See* CILIA AND FLAGELLA; MASTIGOPHORA; PORIFERA; PROTOZOA; TRYPANOSOMATIDAE. M. M. Brooke

Bibliography. S. H. Hutner and A. Lwoff (eds.), *Biochemistry and Physiology of Protozoa,* 3 vols., 1961–1964; T. L. Jahn and F. F. Jahn, *How to Know Protozoa,* 1949; R. R. Kudo, *Protozoology,* 5th ed., 1977; A. Westfall, *Protozoa*, 1976.

Kinorhyncha

A phylum of free-living marine invertebrates less than 1 mm long. They are segmented and lack internal ciliation (see **illustration**). Kinorhyncha, formerly called Echinoderida, are sometimes included in the rejected superphylum Aschelminthes, since their fluid-filled body cavity is probably a pseudocoel, or false coelom. Kinorhynchs are benthonic, generally dwelling in mud or sand from intertidal to deep-sea habitats. Two orders are generally recognized, Cyclorhagida and Homalorhagida.

The body is covered by a transparent cuticle secreted by an underlying epidermis. The cuticle is molted only in the process of juvenile growth. Three body regions are recognized: a head segment, a neck segment, and an 11-segment trunk. The head is completely retractable. When everted, the head extends its five to seven circles of recurved spines called scalids.

The neck consists of plates called placids which function to close the anterior opening of the trunk when the head is retracted. The Cyclorhagida have 14–16 placids arranged around the circular opening of the trunk; but in one suborder, the Conchorhagae, these placids function with the two lateral clam-shell portions of the trunk segment which constitutes the primary closing mechanism. In the Homalorhagida four to eight placids function, with the three ventral plates of the first trunk segment, like a trapdoor—the anterior margin of which is drawn up against a vaulted dorsal plate to close the triangular-shaped opening into which the head is withdrawn.

Most kinorhynchs have a pair of adhesive tubes on the ventral surface of the third or fourth segment. Additional adhesive tubes may occur near some of the lateral spines of more posterior trunk segments.

Circular musculature is present in the head, neck, and pharynx. Trunk musculature is striated and consists of segmentally arranged longitudinal, oblique, and dorsoventral muscles. Anteriorly, the longitudinal muscles form the retractor muscles of the head. Eversion of the head is accomplished by simultaneously relaxing the retractor muscles while contracting paired dorsal ventral muscles, which increases the pressure of the body fluid and everts the head. Retraction of the head is the reverse of this process.

Locomotion is a consequence of repeated eversion and retraction of the head. As the head is everted, mud or sand particles are gripped by the scalids, and the animal is pulled forward.

The nervous system consists of a 10-lobed brain which encircles the anterior end of the pharynx, a poorly developed double ventral nerve cord with segmentally arranged groups of ganglionic cells, and eight other longitudinal nerves situated dorsally, subdorsally, laterally, and lateroventrally.

The mouth is located terminally on the mouth cone, which protrudes from the center of the head. A series of anteriorly directed spines surround the mouth opening. There is a short buccal cavity followed by a muscular pharynx similar to that of the Nematoda and Gastrotricha, except that it is lined with a syncytial epithelium which overlies the cuticle. The pharynx is followed by a short esophagus and stomach-intestine, which is a straight tube with a network of muscle fibers. The end gut is short, lined with cuticle, and separated from the stomach-intestine by a sphincter. The anus is terminal.

A single pair of protonephridial excretory organs

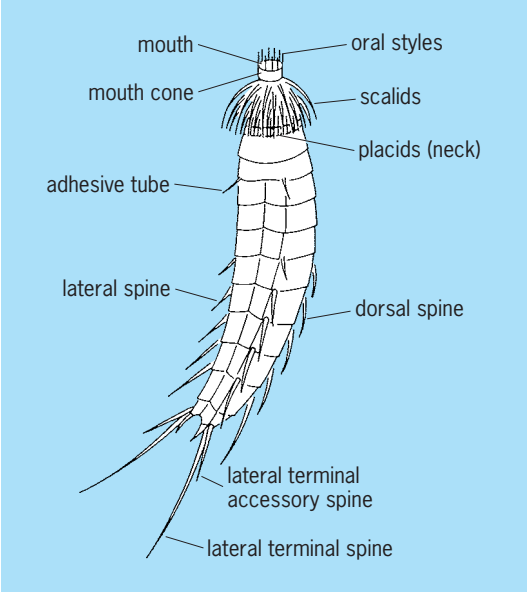

***Echinoderes* sp., a cyclorhagid kinorhynch (ventrolateral view).**

begins in the tenth segment and opens laterodorsally in the eleventh segment. These organs are modified solenocytes, having either a single long driving flagellum as in the Cyclorhagida, or a long flagellum and several shorter driving flagella along the canal in the Homalorhagida.

The Kinorhyncha are dioecious. In each sex there is one pair of gonads. The genital pores are located on the thirteenth segment. The ovary contains both ova and nutritive cells. A short oviduct extends from the ovary to the genital pore. A seminal receptacle may be present. The male genital pores may have several penile spines which may serve as a copulatory apparatus.

The embryology of the Kinorhyncha is unknown. Eggs hatch into minute juveniles which appear to have 11 segments. The juvenile passes through several successive stages by molting, then develops all 13 segments distinctly, and finally becomes an adult. Robert P. Higgins

Bibliography. B. Neuhaus and R. P. Higgins, Ultrastructure, biology, and phylogenetic relationships of Kinorhyncha, *Integr. Comp. Biol.*, 42:619–632, 2002.

Kirchhoff's laws of electric circuits

Fundamental laws dealing with currents at junctions and voltages around any loop within a network. These laws are commonly used in the analysis and solution of networks. They may be used directly to solve circuit problems, and they form the basis for network theorems used with more complex networks.

In the solution of circuit problems, it is necessary to identify the specific physical principles involved in the problem and, on the basis of them, to write equations expressing the relations among the unknowns. Physically, the analysis of networks is based on Ohm's law giving the branch equations, Kirchhoff's voltage law giving the loop voltage equations, and Kirchhoff's current law giving the node current equations. Mathematically, a network may be solved when it is possible to set up a number of independent equations equal to the number of unknowns. *See* CIRCUIT (ELECTRICITY); NETWORK THEORY.

When writing the independent equations, current directions and voltage polarities may be chosen arbitrarily. If the equations are written with due regard for these arbitrary choices, the algebraic signs of current and voltage will take care of themselves.

Kirchhoff's voltage law. One way of stating Kirchhoff's voltage law is: "At each instant of time, the algebraic sum of the voltage rise is equal to the algebraic sum of the voltage drops, both being taken in the same direction around the closed loop."

The application of this law may be illustrated with the circuit in **Fig. 1**. First, it is desirable to consider the significance of a voltage rise and a voltage drop, in relation to the current arrow. The following definitions are illustrated by Fig. 1.

A voltage rise is encountered if, in going from 1 to

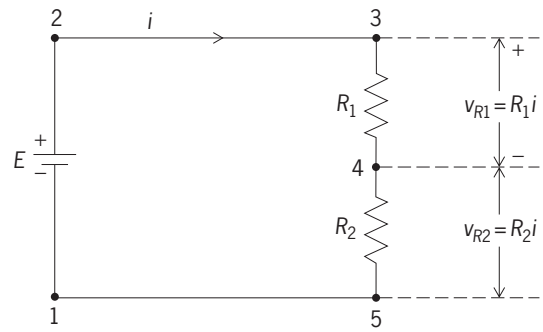

Fig. 1. Simple loop to show Kirchhoff's voltage law.

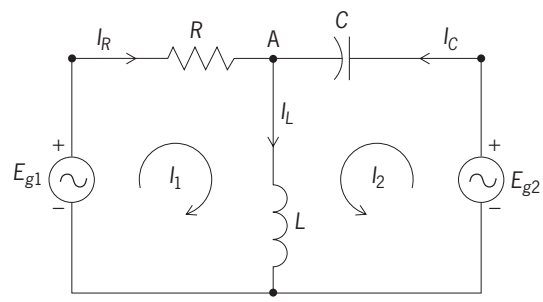

Fig. 2. Two-loop network demonstrating the application of Kirchhoff's voltage law.

2 in the direction of the current arrow, the polarity is from plus to minus. Thus, $v_{R1}; = R_1 i$ is a voltage drop from 3 to 4. The application of Kirchhoff's voltage law gives the loop voltage, Eq. (1).

$$E = v_{R1} + v_{R1} = R_1 i + R_2 i \tag{1}$$

In the network of **Fig. 2** the voltage sources have the same frequency. The positive senses for the branch currents I_R, I_L, and I_C are chosen arbitrarily, as are the loop currents I_1 and I_2. The voltage equations for loops 1 and 2 can be written using instantaneous branch currents, instantaneous loop currents, phasor branch currents, or phasor loop currents.

The loop voltage equations are obtained by applying Kirchhoff's voltage law to each loop as follows.

By using instantaneous branch currents, Eqs. (2) and (3) may be obtained. By using instantaneous loop

$$e_{g1} = Ri_R + L\frac{di_L}{dt} \tag{2}$$

$$e_{g2} = \frac{1}{C}\int i_C dt + L\frac{di_L}{dt} \tag{3}$$

currents, Eqs. (4) and (5) are obtained. Equations (6)

$$e_{g1} = Ri_1 + L\frac{d(i_1 + i_2)}{dt} \tag{4}$$

$$e_{g2} = \frac{1}{C}\int i_2\, dt + L\frac{d(i_2 + i_1)}{dt} \tag{5}$$

and (7) are obtained by using phasor branch

$$\mathbf{E}_{g1} = R\mathbf{I}_R + j\omega L\mathbf{I}_L \tag{6}$$

$$\mathbf{E}_{g2} = -j\frac{1}{\omega C}\mathbf{I}_C + j\omega L\mathbf{I}_L \tag{7}$$

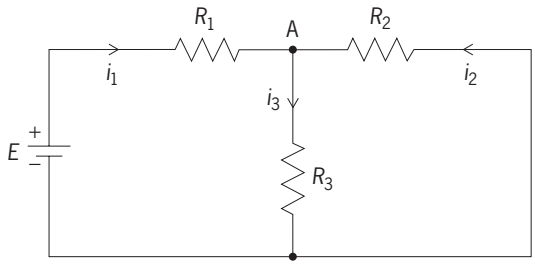

Fig. 3. Circuit demonstrating Kirchhoff's current law.

currents. By using phasor loop currents, Eqs. (8) and (9) may be obtained.

$$\mathbf{E}_{g1} = R\mathbf{I}_1 + j\omega L(\mathbf{I}_1 + \mathbf{I}_2) \tag{8}$$

$$\mathbf{E}_{g2} = -j\frac{1}{\omega C}\mathbf{I}_2 + j\omega L(\mathbf{I}_2 + \mathbf{I}_1) \tag{9}$$

See ALTERNATING-CURRENT CIRCUIT THEORY.

Kirchhoff's current law. Kirchhoff's current law may be expressed as follows: "At any given instant, the sum of the instantaneous values of all the currents flowing toward a point is equal to the sum of the instantaneous values of all the currents flowing away from the point."

The application of this law may be illustrated with the circuit in **Fig. 3**. At node A in the circuit in Fig. 3, the current is given by Eq. (10).

$$i_1 + i_2 = i_3 \tag{10}$$

The current equations at node A in Fig. 2 can be written by using instantaneous branch currents or phasor branch currents.

By using instantaneous branch currents, Eq. (11) is obtained.

$$i_R + i_C = i_L \tag{11}$$

By using phasor branch currents, Eq. (12) is obtained.

$$\mathbf{I}_R + \mathbf{I}_C = \mathbf{I}_L \tag{12}$$

See DIRECT-CURRENT CIRCUIT THEORY.

K. Y. Tang; Robert T. Weil

Kite

A tethered flying device that supports itself and the cable that connects it to the ground by means of the aerodynamic forces created by the relative motion of the wind. This relative wind may arise merely from the natural motions of the air or may be caused by towing the kite through the agency of its connecting cable.

Kites take many forms (see **illus.**); the bow and box kites are common in the United States. In many countries, particularly in Asia, kites are frequently used in rituals and festivals; their bizarre forms and shapes are traditional, some having been developed centuries ago.

The lifting force of all kites is produced by deflecting the air downward, the resulting change in momentum producing an upward force. To be successful, a kite must have an extremely low wing loading (weight/area) so that it can fly even on days when the wind velocity is not high. It must be completely stable, since the only controls available to the operator are the length of cable and the rate at which it is taken in or let out. Efficient design requires that its lift-to-drag ratio be as high as possible. *See* AERODYNAMIC FORCE; AERODYNAMICS.

Experiments on the possible application of efficient aircraft-type lifting surfaces, in which most of the lift arises from the low pressures created by the air flowing over the upper surfaces, have shown them to be too sensitive to changes in wind force and direction. Under normal atmospheric conditions the use of this type of lifting surface results in a kite that behaves in a violent and unpredictable manner. For this reason, the higher drag associated with a surface from which the flow has separated is tolerated, and most of the lifting force is obtained from pressure on the lower surface, because stalled surfaces are much less sensitive to wind changes.

Both the lift-to-drag ratio and the stability of the kite are functions of the length of cable. The more cable released, the more drag created. The increased drag, combined with the increase in weight being supported, causes the kite to sag off downwind, reducing the flight angle, which is the angle formed between the horizontal and a line passing through the kite and the operator.

Most kites with a properly located cable pivot point, generally slightly ahead of the center of gravity, demonstrate longitudinal stability. Lateral and directional instabilities generally couple to produce violent motions. The longer the cable, the more these motions are damped. Lateral and directional stability are improved by the use of effective dihedral (the bow of the bow kite) and a flexible tail, which provides both directional stability and yaw damping.

David C. Hazen

Bibliography. R. Moulton, *Kites*, 1997; C. L. Strong, The lore and aerodynamics of making and flying kites, *Sci. Amer.*, 220(4):130–136, 1969; W. Yolen, *The Complete Book of Kites and Kite Flying*, 1976.

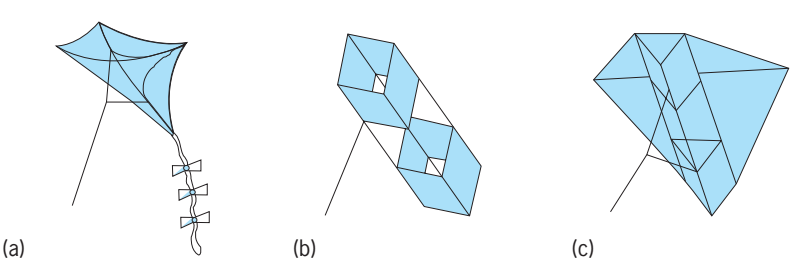

Common forms of kites. (*a*) Bow kite. (*b*) Box kite. (*c*) Modified bow kite.

Kiwifruit

A vigorous deciduous fruiting vine (family Actinidiceae) that is native to central China, where it commonly grows in moist and sheltered areas on the

forest edges. Kiwifruit requires both the female cultivar and a male pollenizer for successful fruit production. The kiwifruit industry depends on a single female cultivar, Hayward, the fruit having a creamy-white central core, black-brown seeds, and a bright translucent green outer flesh surrounded by a light brown fuzzy skin. It is adapted to moderate climates in the temperate zone and requires 600–850 h of winter chilling (temperatures between 32 and 45°F or 0 and 7°C) to ensure uniform budbreak. Kiwifruit wood is susceptible to winter injury at temperatures below 14°F (−10°C), and flower buds can be damaged by frost below 29°F (−1.5°C).

Origin and development. The kiwifruit (*Actinidia deliciosa*) was introduced into cultivation in New Zealand with seed brought from China in 1904, making it one of the most recently domesticated fruiting plants. It appears that all vines now in New Zealand descended from one male and two female plants. The principal kiwifruit-growing countries are Italy, New Zealand, Japan, Chile, France, Greece, United States (California), and Australia.

Kiwifruit is a source of vitamin C, minerals such as potassium, calcium, and phosphorus, and dietary fiber. The primary use is for the fresh market, although culled fruit is processed into canned and frozen fruit slices, wine, jam, juice, and dried products.

Cultivated kiwifruit vines have had very limited genetic selection and are very similar to those found in the wild. There are many species within the genus *Actinidia* with much variation (for example, fruit size and color) that is potentially useful for breeding programs and new introductions.

Propagation and cultivation. The Hayward cultivar is clonally propagated primarily by grafting onto seedling rootstock. Kiwifruit vines require deep, well-drained soils for best yields. Vines are grown on a trellis for support, and two shoots growing in opposite directions from a single straight trunk are trained to form a bilateral cordon structure. Spacing between rows and plants is usually about 15–16 ft (4.5–5 m). Adequate pollination is essential to produce large fruit, as the number of seeds directly correlates to fruit size. Vineyards are commonly planted at a ratio of one male plant to eight female plants. Beehives are used during bloom to increase pollination.

Fruit is borne on 1-year-old canes which arise from the supporting structure and are tied to the trellis wire. Female vines require extensive pruning to maintain strong annual cane growth, regulate crop yield and fruit size, and allow adequate light penetration for bud formation. Dormant pruning consists of removing fruiting canes back to 1-year-old replacement canes which will produce the following year's crop. Windbreaks are often used to shelter vines from the effects of wind damage: broken canes, lost crop, and scarred fruit. Vines have a high moisture requirement. When grown in areas with inadequate summer rainfall, they must have supplemental irrigation. Nitrogen, the most important nutrient, is applied annually during the growing season.

Harvesting and storage. Fruit is hand-picked in the fall once minimum maturity has been achieved. Packed fruit is stored at 32°F (0°C) for several months. During storage, the starch is converted to sugar and the fruit softens.

Diseases. Root and crown rot caused by several species of *Phytophthora* is a fungus disease that thrives in saturated soil conditions. Infected roots turn reddish brown, and the inner wood of infected crowns becomes dark brown. Vines may gradually decline or die within weeks of infection. Control methods include careful soil and water management to prevent the spread of disease from infested soil and irrigation or runoff water; and improved drainage.

Armillaria root rot is mainly a problem in New Zealand. Control methods, none of which are totally satisfactory, include fumigating infested soil, placing barriers in the soil between healthy and infected vines, and removing the remains of felled windbreak trees that harbor the disease.

Bacterial leafspot and blossom blight, caused by *Pseudomonas viridiflava*, is spread to floral buds and young leaves during spring rains. Infected buds may turn brown, fail to open or develop fully, and ultimately drop.

Storage rot caused by *Botrytis cinerea* is the most serious of postharvest diseases. The fungus colonizes stem scars or any damaged area usually after several weeks of storage and may spread to adjacent healthy fruits. Control is based on using preharvest fungicide sprays, avoiding soft fruit through proper cultural practices, and proper handling during harvest and storage. *See* FRUIT; PLANT PATHOLOGY; THEALES.

Janine K. Hasey

Bibliography. K. E. Conn, W. D. Gubler, and J. K. Hasey, Pathogenicity and relative virulence of nine *Phytophthora* spp. from kiwifruit, *Phytopathology*, 81:974–979, 1991; J. K. Hasey and R. S. Johnson (eds.), *Kiwifruit Growing and Handling*, 1994; I. J. Warrington and G. C. Weston (eds.), *Kiwifruit Science and Management*, 1990.

Klebsiella

A genus of gram-negative, nonmotile bacteria whose members ferment lactose, do not decarboxylate ornithine or form hydrogen sulfide, and often produce large mucoid colonies and gas from glucose. Characteristic large mucoid colonies are due to production of a large amount of capsular material. Some isolates, particularly those from urinary tract infections, produce a slow-acting urease enzyme. Species of *Klebsiella* are commonly found in soil and water, on plants, and in animals and humans. Harmless strains of *Klebsiella* are beneficial because they fix nitrogen in soil. Pathogenic species include *K. pneumoniae, K. rhinoscleromatis*, and *K. ozaenae*, also known as *K. pneumoniae* subspecies *pneumoniae, rhinoscleromatis*, and *ozaenae*. Species and strains can be differentiated by biochemical reactions and

typed by bacteriocin susceptibility and serologically by antigen.

Klebsiella pneumoniae is the second most frequently isolated colon-related bacterium in clinical laboratories. The carbohydrate-containing capsule of *Klebsiella* promotes virulence by protecting the encased bacteria from ingestion by leukocytes; nonencapsulated variants of *Klebsiella* do not cause disease. Capsular types 1 and 2 cause pneumonia, accounting for about 3% of all life-threatening acute bacterial pneumonias. Capsular types 8, 9, 10, and 24 are commonly associated with urinary tract infections, second in prevalence only to *Escherichia coli*. *See* ESCHERICHIA; PNEUMONIA.

Klebsiella accounts for a large percentage of hospital-acquired infections, mostly skin infections (in immunocompromised burn patients), bacteremia, and urinary tract infections. It is also the most common contaminant of intravenous fluids such as glucose solutions and other medical devices. Partly, this problem is due to the fact that from 5 to 10% of healthy individuals carry *Klebsiella* in their respiratory tract. *See* HOSPITAL INFECTIONS.

Klebsiella may produce *E. coli*–like enterotoxins and cause acute gastroenteritis in infants and young children. Enteric illnesses due to *Klebsiella* are more predominant where populations are more crowded and conditions less sanitary. Other virulence factors of *Klebsiella* include a relatively high ability to survive and multiply outside the host in a variety of environments, and its relatively simple growth requirements. An important virulence factor of *Klebsiella* is its pyrogenic, or fever-producing, endotoxin, which is a heat-stable lipopolysaccharide component of the organism's cell wall. *Klebsiella* also has a relatively nonspecific virulence factor, which is the so-called common pili produced by almost all enteric bacteria. Common pili are nonflagellar hairlike appendages which function as receptors for mannose, a carbohydrate which is almost universally present on the surface of animal cells. These appendages function as adhesins and are thus important in the colonization phase of infection and in maintaining and spreading infection. *See* ENDOTOXIN.

Klebsiella rhinoscleromatis causes rhinoscleroma, a chronic destructive granulomatous disease of the upper respiratory tract that is most common in eastern Europe, central Africa, and tropical South America. *Klebsiella ozaenae* is one cause of chronic rhinitis (ozena), a destructive atrophy of the nasal mucosa, and is infrequently isolated from urinary tract infections and bacteremia. *See* MEDICAL BACTERIOLOGY.

Doyle J. Evans, Jr.

Bibliography. S. Baron (ed.), *Medical Microbiology*, 4th ed., 1996.

Klystron

An evacuated electron-beam tube in which an initial velocity modulation imparted to electrons in the beam results subsequently in density modulation of the beam. A klystron is used either as an amplifier in the microwave region or as an oscillator. For use as an amplifier, a klystron receives microwave energy at an input cavity through which the electron beam passes. The microwave energy modulates the velocities of electrons in the beam, which then enters a drift space. Here the faster electrons overtake the slower to form bunches. In this manner, the uniform current density of the initial beam is converted to an alternating current. The bunched beam with its significant component of alternating current then passes through an output cavity to which the beam transfers its ac energy. *See* MICROWAVE.

Klystron amplifier. In a typical klystron (**Fig. 1**), a stream of electrons from a concave thermionic cathode is focused into a smaller cylindrical beam by the converging electrostatic fields between the anode, cathode, and focusing electrode. The beam passes through a hole in the anode and enters a magnetic field parallel to the beam axis. The magnetic field holds the beam together, overcoming the electrostatic repulsion between electrons which would otherwise make the beam spread out rapidly. The electron beam goes through the cavities of the klystron in sequence, emerges from the magnetic field, spreads out, and is stopped in a hollow collector where the remaining kinetic energy of the electrons is dissipated as heat. *See* ELECTRON MOTION IN VACUUM.

The signal wave to be amplified is introduced into the first, or buncher, cavity through a coaxial transmission line. This hollow metal cavity is a resonant circuit, analogous to the familiar inductance-capacitance combination, with the electric field largely concentrated in the reentrant noses so that the highest voltage occurs between them. The inductance may be considered as a single-turn conductor formed by the outer metal walls. In Fig. 1 the current in the center conductor of the input transmission line flows through a loop inside the cavity and back to the outer conductor. The magnetic flux generated in the loop links through the cavity inductance, as in a transformer. At the resonant frequency of the cavity, the voltage across the reentrant section through which the electron beam passes is built up by the cavity configuration to 10–100 times the voltage in the input line. *See* CAVITY RESONATOR.

Figure 2 shows the pattern of the electric field in the cavity and how it varies cyclically with time. As electrons pass through the gap in which the cavity field is concentrated (Fig. 2*a*), they are accelerated or decelerated (Fig. 2*b*), depending on the instantaneous direction of the field.

Figure 3 illustrates graphically the effect of these velocity changes. Each slanted line represents the flight of an electron as a function of time measured in electrical degrees as the electron travels from the buncher gap. The slope of a line is thus the velocity of that electron. The velocities leaving the buncher vary sinusoidally with time, as determined by the instantaneous field. The horizontal broken line represents a fixed point beyond the buncher. The flow

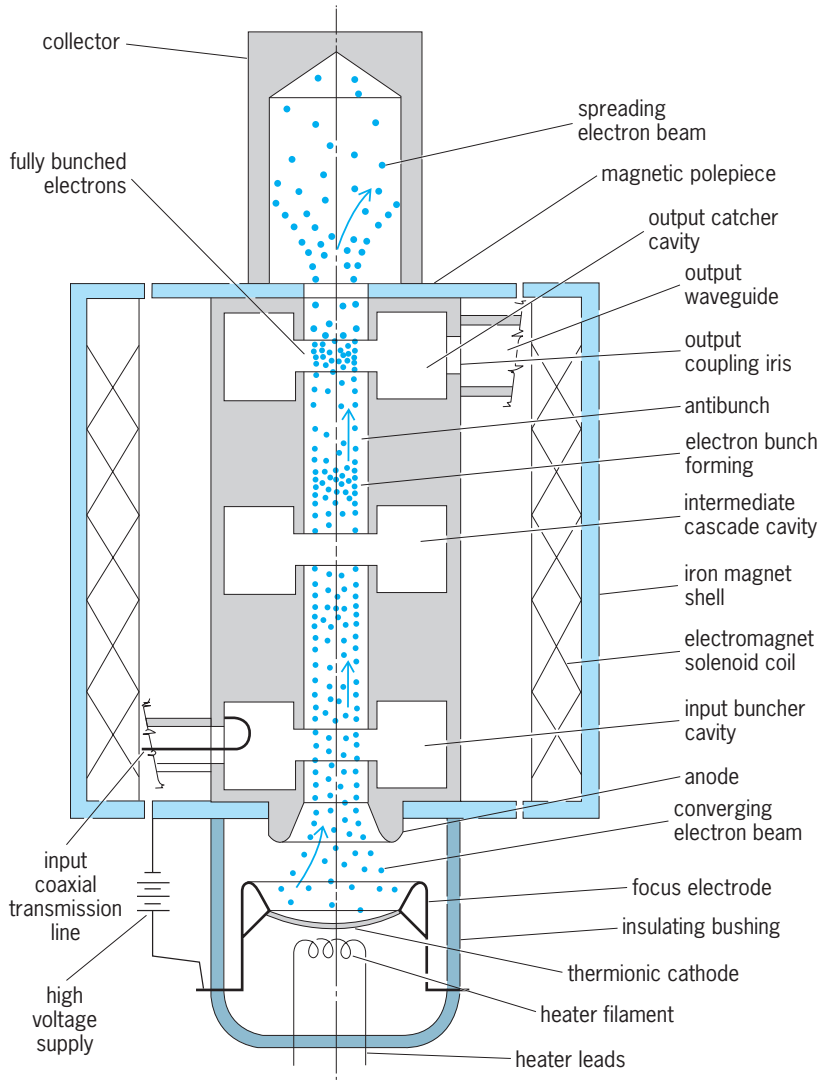

Fig. 1. Cross section of cascade klystron amplifier. (*Varian Associates*)

of electrons past this point is given by the time sequence in which the electron paths cross the broken line. Figure 3 shows how the electrons have gathered into bunches. The rate of current flow is now periodic with time so that the current has an alternating component.

When the bunched beam passes through a second cavity, its space charge induces in the walls of this cavity an alternating current of opposite sign to the electron current. The cavity is tuned to the input frequency so that it has a high resonant impedance; the induced current flowing through this impedance generates voltage in the cavity.

In Fig. 1 the second cavity is not coupled to any outside circuits. Voltage built up here by the beam current produces further velocity modulation in the beam. The resulting alternating current component is about 10 times greater than the initial current. More of these uncoupled cascade cavities can be added for increased amplification.

The final output cavity is coupled into a transmission line (a waveguide in Fig. 1) which carries off the generated power to its useful destination. Because the cavity is tuned to resonance, its reactance is canceled. The induced current flowing through its pure resistive impedance generates in-phase voltage in the direction opposing the current flow. Thus the field in the gap is at its maximum decelerating value at the time a bunch of electrons passes. Most of the electrons therefore are slowed down, and there is a net transfer of kinetic energy of the electrons into electromagnetic energy in the cavity. Klystron amplifiers are used in transmitters for radar and one-way radio communication, for driving particle accelerators, and for dielectric heating. The useful range of frequencies is from 400 MHz to 40 GHz. Power levels range from a few watts up to 400 kW of continuous power or 20 MW for short pulses. Amplification is about 10 dB for a two-cavity tube. With more cavities, gains up to 60 dB are practical.

Figure 4 shows the construction of a four-cavity amplifier rated at 2 MW pulsed output at 2.8 GHz. It operates in a solenoid magnet, as in Fig. 1. The

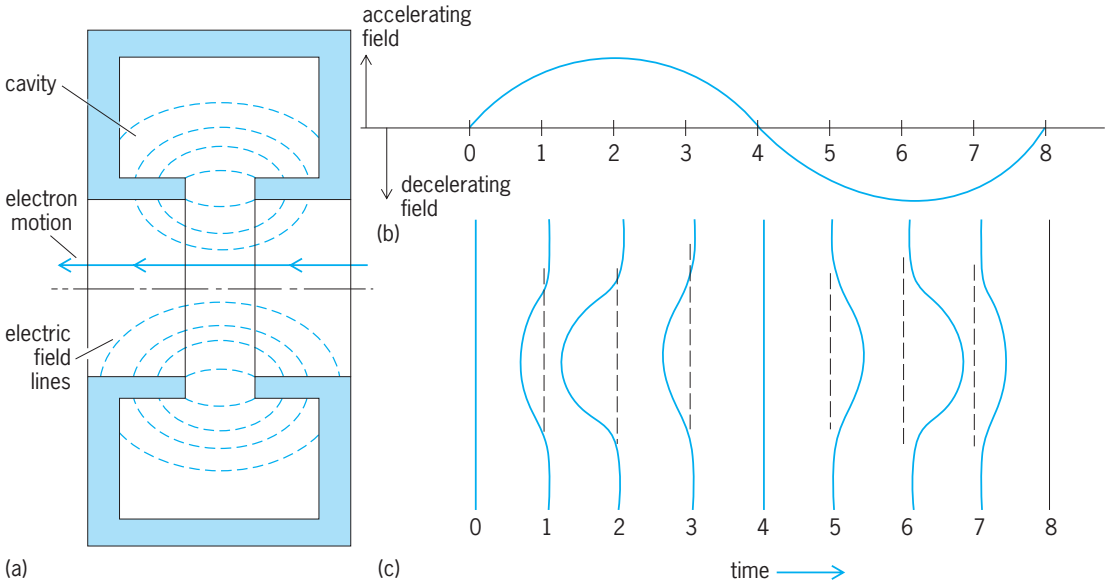

Fig. 2. Cavity concentrates electric field between the reentrant noses. As field varies sinusoidally with time, an electron crossing in the gap between the noses experiences an electric field whose strength and direction depend on the instantaneous phase of the field. (*a*) Map of instantaneous field. (*b*) Cycle variation of field. (*c*) Profile of field on the axis at various times in the cycle. (*Varian Associates*)

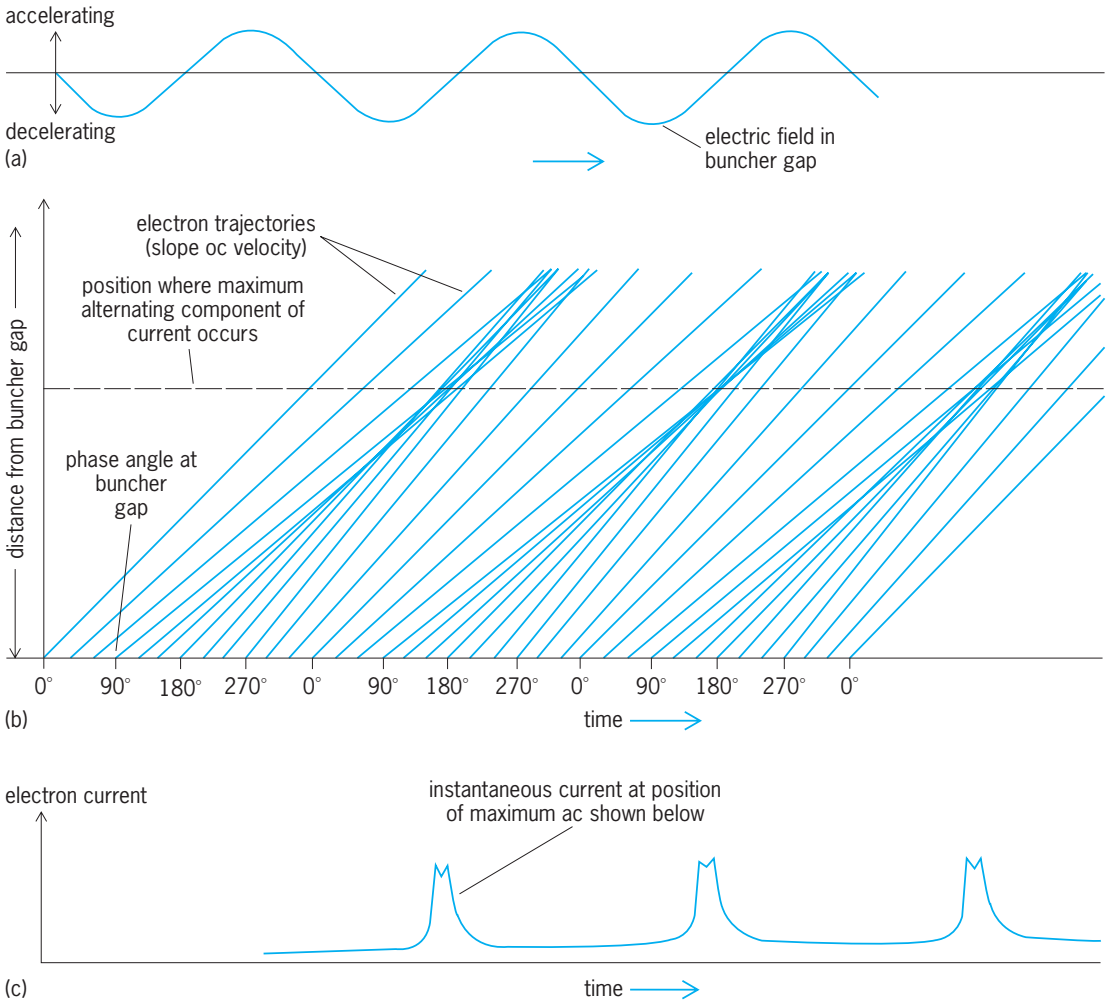

Fig. 3. Bunching the electron stream. (*a*) As the electric field varies periodically, electrons traversing the cavity are speeded up or slowed down. (*b*) Distance-versus-time lines graph the gradual formation of bunches. (*c*) Current passing a fixed point becomes periodic, that is, it becomes alternating. (*Varian Associates*)

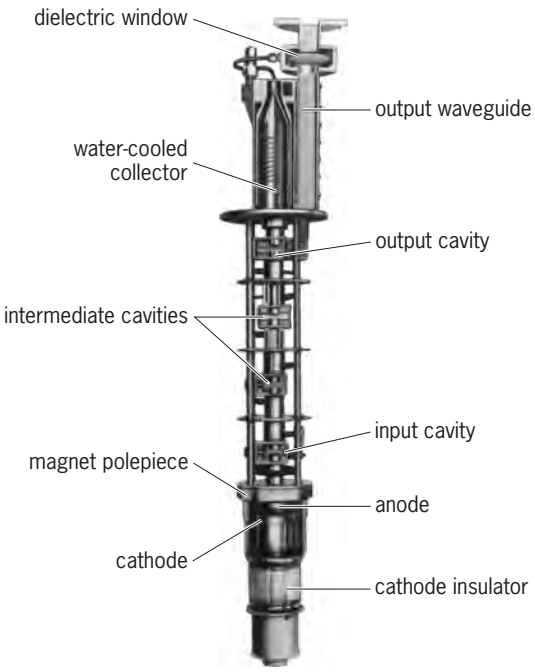

Fig. 4. Cutaway view of four-cavity amplifier. (*Varian Associates*)

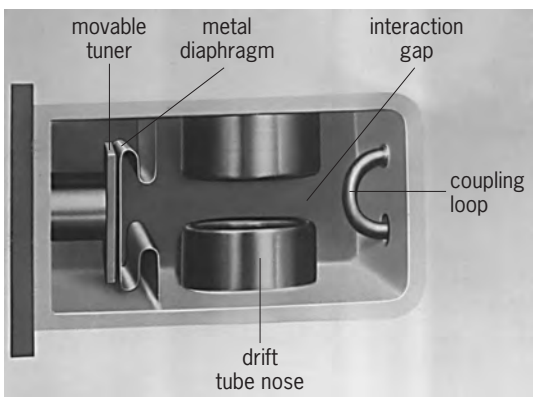

Fig. 5. Detail of input cavity for four-cavity klystron amplifier of Fig. 4. (*Varian Associates*)

cavities are tuned to the operating frequency by moving one flexible inner wall of the box-shaped cavity, changing its volume and its effective inductance. In **Fig. 5** the details of the input cavity are enlarged.

Reflex oscillator. Klystrons may be operated as oscillators by feeding some of the output back into the input circuit. More widely used is the reflex oscillator in which the electron beam itself provides the feedback. **Figure 6** illustrates the operation. The beam is focused through a cavity, as in the amplifier. No magnetic field is needed to keep the beam focused because the total travel distance is short and the amount of natural spreading is tolerable. The cavity usually has grids with open mesh through which the electrons can penetrate. The purpose of the grids is to concentrate the electric field in a short space so that the field can interact with a slow, low-voltage electron beam.

In the cavity the beam is velocity-modulated as in the amplifier. Leaving the cavity, the beam enters a region of dc electric field opposing its motion, produced by a reflector electrode operating at a potential negative with respect to the cathode. The electrons do not have enough energy to reach the electrode, but are reflected in space and return to pass through the cavity again. The points of reflection are determined by electron velocities, the faster electrons going farther against the field and hence taking longer to get back than the slower ones.

A trajectory plot for the reflex oscillator is shown in **Fig. 7**. In a uniform retarding field the space-versus-time curves are parabolas. As in Fig. 3, velocity modulation produces bunches of electrons. If the voltages are adjusted so that the average time to return is $n + {}^3/_4$ cycles (n = integer), the bunches cross the cavity when the alternating field is maximum in the decelerating direction. This transfers beam energy to the cavity.

Because the reflex klystron has only one cavity, it is easy to tune its frequency with a single adjustment. Power output is from 10 mW to a few watts. Reflex oscillators are used as signal sources from 3 to 200 GHz, as the transmitter tubes in line-of-sight radio relay systems, and in low-power radars.

A cutaway of a typical reflex klystron is shown in **Fig. 8**. This tube is tuned by deforming the upper

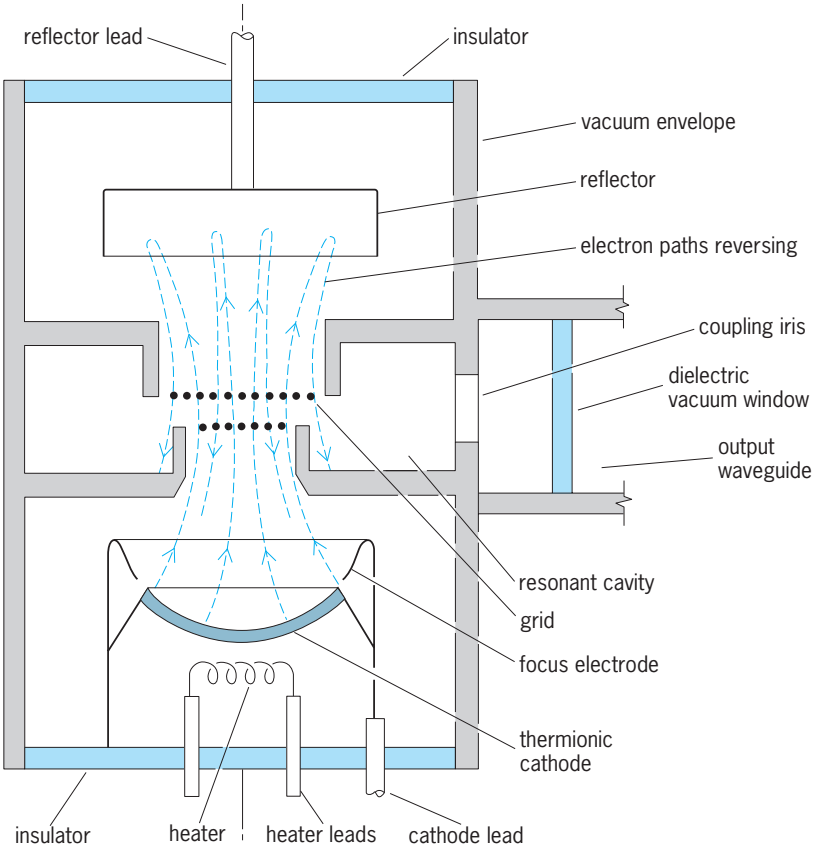

Fig. 6. Schematic cross section of reflex oscillator. (*Varian Associates*)

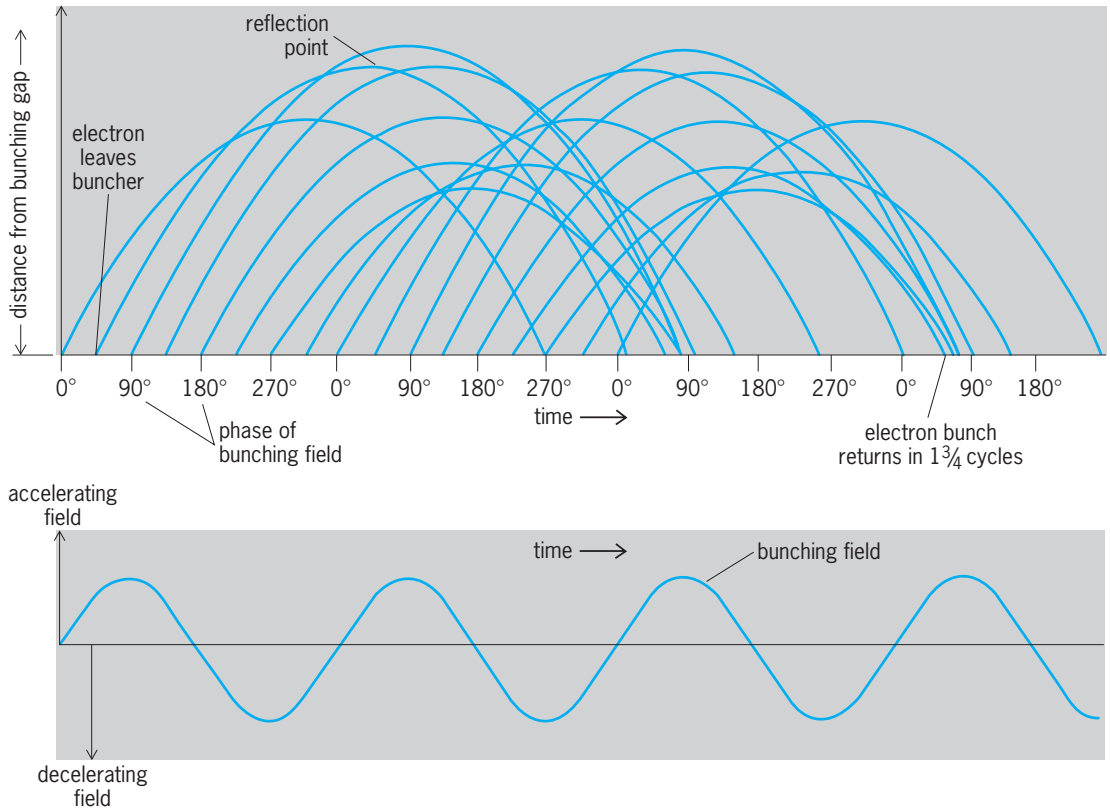

Fig. 7. Trajectories show bunching in a reflex oscillator. Electrons are turned back by a retarding field, faster ones going farther and taking longer. The bunched beam returns through the cavity. (*Varian Associates*)

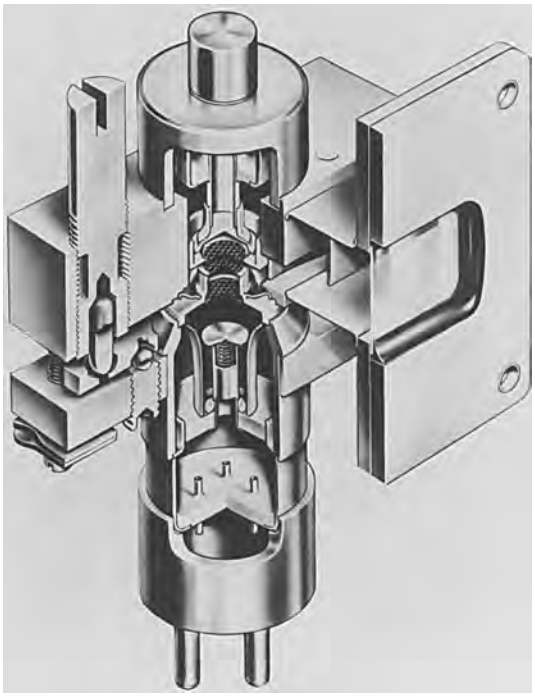

Fig. 8. Cutaway of reflex oscillator shows (at left) the screw that deforms the flexible bottom wall of the cavity to change its resonant frequency. The beam passes through honeycomb grids in the cavity. The cathode leads come out at the bottom socket, and the reflector lead comes out at the top. The output waveguide has a mica vacuum window. (*Varian Associates*)

cavity wall, which varies the spacing between the grids and hence the effective capacitance of the cavity resonator. Richard B. Nelson

Bibliography. J. J. Carr, *Elements of Microwave Technology*, 1989; K. Chang, *Handbook of Microwave and Optical Components*, 4 vols., 1990; A. S. Gilmour, *Microwave Tubes*, 1986; T. K. Ishii, *Practical Microwave Electron Devices*, 1990; S. Y. Liao, *Microwave Devices and Circuits*, 3d ed., 1990; D. M. Pozar, *Microwave Engineering*, 2d ed., 1997.

Knudsen number

In gas dynamics, the ratio of the molecular mean free path λ to some characteristic length L: $\mathrm{Kn} = \lambda/L$. The length chosen will depend on the problem under consideration. It may be, for example, the diameter of a pipe or an object immersed in a flow, or the thickness of a boundary layer or a shock wave. *See* MEAN FREE PATH.

The magnitude of the Knudsen number determines the appropriate gas dynamic regime. When the Knudsen number is small compared to unity, of the order of $\mathrm{Kn} \leq 0.1$, the fluid can be treated as a continuous medium and described in terms of the macroscopic variables: velocity, density, pressure, and temperature. In the transition flow regime, for Knudsen numbers of the order of unity or greater, a microscopic approach is required, wherein the

trajectories of individual representative molecules are considered, and macroscopic variables are obtained from the statistical properties of their motions. In both internal and external flows, for Kn $\geq$ 10, intermolecular collisions in the region of interest are much less frequent than molecular interactions with solid boundaries, and can be ignored. Flows under such conditions are termed collisionless or free molecular. In the range $0.1 \leq \text{Kn} \leq 1.0$, termed the slip flow regime, it is sometimes possible to obtain useful results by treating the gas as a continuum, but allowing for discontinuities in velocity and temperature at solid boundaries.

Since the mean free path is inversely proportional to the gas density, departures from continuum gas dynamics are usually associated with flows in vacuum systems and in high-altitude aerodynamics, often referred to as rarefied gas dynamics. However, such departures can also occur at normal densities owing to the existence of a very small characteristic length, as in the motion of very small particles such as aerosols, and in the flow through very narrow passageways. With the ever-increasing power of computational resources, numerical methods have displaced analytical methods as the predominant approach to the treatment of flows with complicated geometries and complex chemistry, over the entire range of Knudsen number. Monte Carlo methods are particularly useful in the transition flow regime. *See* COMPUTATIONAL FLUID DYNAMICS; GAS DYNAMICS; KINETIC THEORY OF MATTER; MONTE CARLO METHOD; RAREFIED GAS FLOW. Lawrence Talbot

Bibliography. G. A. Bird, *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*, Oxford University Press, 1994.

Koala

A single species, *Phascolarctos cinereus*, which is a member of the family Phalangeridae in the mammalian order Marsupialia (pouch-bearing animals). It is a small, clean, affectionate animal that weighs from 11 to 17 lb (5 to 8 kg) when mature. Koalas are restricted to eastern Australia, where the eucalyptus grows. Not only do they have a specialized diet of eucalyptus leaves, but the leaves must be of a certain age from a specific species of tree, and the tree must grow upon a certain type of soil. Each adult will eat an average of 2–3 lb (0.9–1.5 kg) of leaves each day. While the diet is a definite limiting factor to an increase in the population, millions died from epidemic diseases at the turn of the twentieth century, and later thousands were killed for their fur. These animals are now strictly protected by the Australian government because their numbers are small. Exportation is prohibited.

The koala is well adapted to its arboreal habitat, since both the fore- and hindlimbs have two of the clawed digits opposing the other three, thus constituting an efficient grasping organ (see **illustration**). The koala breeds once each season, and the usual number of offspring is one. It remains in the pouch for 6 months; then it clings to the back of its mother and is carried around in this manner until 1 year old. *See* MAMMALIA; MARSUPIALIA. Charles B. Curtin

Koala (*Phascolarctos cinereus*), a slow-moving arboreal animal. (*Photo by Gerald and Buff Corsi; © 2002 California Academy of Sciences*)

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999.

Kohlrabi

A cool-season biennial crucifer, *Brassica caulorapa* and *B. oleracea* var. *caulo-rapa*, of northern European origin belonging to the plant order Capparales. Kohlrabi is grown for its turniplike enlarged

Kohlrabi (*Brassica caulorapa*), cultivar Early White Vienna. (*Joseph Harris Co., Rochester, New York*)

stem, which is usually eaten as a cooked vegetable (see **illus.**). Kohlrabi is a German word meaning cabbage-turnip and reflects a similarity in taste and appearance to both vegetables. Cultural practices for kohlrabi are similar to those used for turnips. White Vienna and Purple Vienna are popular varieties (cultivars). Harvesting, when the enlarged stems are 2–3 in. (5–8 cm) in diameter, is usually 2 months after planting. A common cooked vegetable in Europe, especially Germany, kohlrabi is of minor importance in the United States. *See* CABBAGE; CAPPARALES; KALE; TURNIP.

H. John Carew

Kondo effect

An unusual, temperature-dependent effect displayed in the thermal, electrical, and magnetic properties of nonmagnetic metals containing very small quantities of magnetic impurities. A striking example is the anomalous, logarithmic increase in the electrical resistivity ρ with decreasing temperature. Other properties, such as heat capacity C_v, magnetic susceptibility χ, and thermoelectric power S, also display anomalous behavior because of the Kondo effect. For these properties, the temperature dependence of a typical dilute magnetic metal (Kondo alloy) differs greatly from the behavior expected of an ordinary metal containing no magnetic impurities (see **illus.**).

The Kondo effect has been observed in a wide variety of dilute magnetic alloys. Usually these alloys are made from a nonmagnetic host such as copper, silver, gold, magnesium, or zinc and a small amount of a magnetic metal impurity such as chromium, manganese, iron, cobalt, nickel, vanadium, or titanium. Typical concentrations range from about one to a few hundred magnetic atoms per million host atoms. At higher concentrations, the dilute magnetic alloys may display spin-glass behavior. *See* SPIN GLASS.

Importance. The Kondo effect is used in thermometry applications, especially thermocouple thermometers at very low temperatures (that is, millikelvin temperatures). In other applications where the properties of pure metals are studied, the Kondo effect serves as a useful indicator of the metal's magnetic-impurity level.

The problem of understanding the Kondo effect is considered important since it is recognized to be a simpler version of the more complex problem of understanding ferromagnetism in magnetic materials, which is one of the great challenges in physics. Basically the Kondo effect is an example of the most simple possible magnetic system—a single magnetic atom in a nonmagnetic environment. (The alloys used are so dilute that the interaction between different magnetic impurities can be safely ignored.) Although this involves a simple physical model, the problem has required some of the most sophisticated mathematical techniques known to advance its understanding.

An important step in this direction was the development of a partial mathematical solution of the Kondo problem using renormalization field theory techniques. Information gained in this step helped with the final development of a mathematically exact solution of the Kondo problem. The exact solution permits a systematic calculation of all properties (resistivity, thermal conductivity, thermopower, specific heat, magnetic susceptibility, neutron scattering behavior, and so forth) and provides a physical understanding of these properties. The theoretical work on the Kondo problem has been connected with new understanding in a variety of other scientific disciplines such as condensed-matter physics, surface physics, critical phenomena, elementary particle physics, magnetism, molecular physics, and chemistry, where parallels and analogs to the Kondo problem can be identified

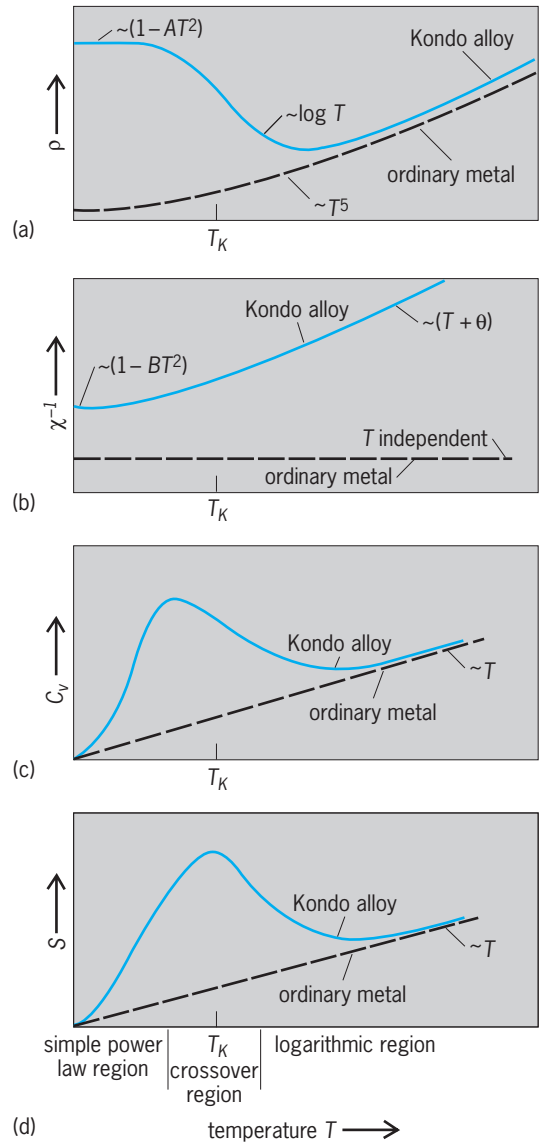

Temperature dependence of a typical Kondo alloy compared with an ordinary metal for four experimentally measurable quantities: (*a*) electrical resistivity ρ, (*b*) inverse magnetic susceptibility χ^{-1}, (*c*) thermodynamic heat capacity C_v, and (*d*) thermoelectric power S. Characteristic dependences on temperature T, the Kondo temperature T_K, and three regions of behavior are identified.

and utilized. *See* CRITICAL PHENOMENA; FERROMAGNETISM; RENORMALIZATION.

Effect of exchange interaction. The first atomic understanding of the phenomenon was achieved in 1964, when Jun Kondo attempted to explain the puzzling minimum in the resistivity of dilute magnetic alloys at low temperatures. Kondo analyzed the simple model of a single magnetic impurity in a nonmagnetic metal by assuming an exchange interaction between the impurity atom and conduction electrons. The exchange interaction J, as it is known in quantum mechanics, is a fundamental consequence of basic symmetry relations that all atomic particles must obey, as a result of their intrinsic spin. By calculating how much the mobile conduction electrons were scattered off the fixed magnetic impurities, Kondo was able to determine the electrical resistivity. He found that the scattering cross section of the electrons showed a remarkable increase (mathematically a logarithmic singularity) at low temperatures when the analysis was carried one step beyond the conventional quantum-mechanical calculation. *See* EXCHANGE INTERACTION.

At a certain characteristic temperature, now referred to as the Kondo temperature T_K, the Kondo effect becomes clearly observable. In Kondo's theory this temperature was identified when the mathematical series of terms in the quantum-mechanical calculation diverged logarithmically. Later work showed that T_K depended on the exchange interaction J, the number of quantum states per unit energy n (often called density of states), and the kinetic energy of the free conduction electrons ϵ_F (more explicitly, the Fermi energy). More refined theories show that a suitable mathematical expression for T_K is

$$T_K \approx \epsilon_F J^{1/2} \exp\left(-\frac{1}{nJ}\right)$$

See FERMI SURFACE; FREE-ELECTRON THEORY OF METALS.

For a given alloy, the values of J, n, and ϵ_F are unique and cannot be adjusted; therefore, at temperatures much higher than T_K the metal atoms experience random thermal agitations that overwhelm the exchange interaction and the Kondo effect becomes insignificant. As temperatures approach T_K, just the reverse occurs. The thermal agitation drops dramatically with falling temperature, and the exchange interaction, which is temperature-independent, dominates; hence the Kondo effect becomes observable (see illus.).

In early experimental investigations (beginning in the 1930s), the Kondo effect was more easily detected in alloys at rather low temperatures (a few hundredths to a few degrees above absolute zero) where the thermal agitation was assuredly small. However, with improved measuring techniques and better alloys, the Kondo effect has been observed at much higher temperatures, corresponding to T_K of several hundred kelvins, well above room temperatures.

Many-body effects. Although the logarithmic rise in the resistivity could be explained, there was still the problem of finding a theory that accounted for the leveling-off of the resistivity at the lowest temperatures (see illus.). This problem was resolved by realizing that a type of many-body effect takes place between the conduction electrons and a magnetic impurity atom, leading to an effective coupling between the impurity and its neighboring electrons. This many-body effect concerns the problem of electrons in a metal experiencing a complex environment where many particles may interact simultaneously around an impurity. In a simple way, the effective coupling may be viewed as a cloud of interactions formed by compounding many simple two-body exchange interactions. At higher temperatures the effective coupling is weak because the thermal agitation rapidly destroys the simple exchange interaction between an electron and a magnetic impurity before another electron, in the vicinity of the impurity, has a chance to react to the effect of the interaction. At lower temperatures the effective coupling becomes more prominent since the interactions are not as rapidly destroyed by thermal agitation. (Technically the increase in coupling is a consequence of an increasing correlation between the spin polarization of the conduction electrons and the impurity spins.) Eventually the coupling gains sufficient strength that the impurity's magnetic spin becomes screened by the magnetic spin of the electron. At this point, there no longer appear to be localized centers of magnetic atoms, and the electron-scattering cross section becomes constant; consequently, the logarithmic rise in the resistivity ceases. *See* SPIN (QUANTUM MECHANICS).

Crossover problem. Bridging the crossover from weak coupling to strong coupling presented a major challenge to many-body physics, and this difficult task was first accomplished in 1973 by Kenneth Wilson through computer calculations based on renormalization group theory. In 1981 this crossover problem was solved by a different method, which yielded an analytical solution. The agreement between these two approaches was considered excellent. There also was the advantage that the analytical method provided a calculated expression for the magnetic susceptibility as observed in the Kondo effect. However, these approaches were still limited by their inability to calculate the temperature dependence of the various properties associated with the Kondo effect.

Exact solution. An important breakthrough in the problem came with the study of dilute impurity cases which have high magnetic impurity concentrations. Especially interesting were alloys containing rare-earth elements such as dilute cerium in a metallic host of lanthanum hexaboron (LaB_6). Such materials are generally classed as heavy fermions or mixed-valence systems. In seeking an explanation of these systems, there arose the realization that the large degeneracy number N of an impurity atom's states (for example, the states of the f-orbital electrons of cerium with the same energy) could be

used to develop mathematical expressions that describe appropriate physical properties. In particular, these mathematical expressions were found to be summations of a convergent series of terms in inverse powers of N, leading subsequently to a formalism known as the $1/N$ expansion method. Coupled with other mathematical formalisms, the $1/N$ expansion method, as applied to the Kondo problem, has been used to develop a consistent set of integral equations that can be solved numerically. Specifically, the computation for one alloy (dilute cerium in lantharum hexaboron) has been carried to completion. In terms of only a single parameter (the peak position of the Kondo resonance T_0 which is proportional to the Kondo temperature T_K), the theory explains the Kondo behavior of this alloy and fits the data of four different Kondo-effect properties, each measured over four decades of temperature. This achievement is considered to be an exact solution of the Kondo problem, and one that represents a universal understanding of various physical properties displaying the Kondo effect.

Many generalizations of the Kondo problem and its relatives, that is, specialized impurity model problems, continue to be a focus of interest. In one special case, namely, the general Kondo problem, success in finding an exact solution has been achieved by using a conformal field theory approach. In this approach the impurity is identified as an immobile quasiparticle while the scattering electron is equated with a mathematical artifice known as a kink, which is a special type of field configuration interpolating between adjacent potential energy wells. *See* ELECTRICAL RESISTIVITY; HEAT CAPACITY; MAGNETIC SUSCEPTIBILITY; SPECIFIC HEAT OF SOLIDS; THERMOELECTRICITY.

Wiley P. Kirk

Bibliography. N. Andrei, K. Furuya, and J. H. Lowenstein, Solution of the Kondo problem, *Rev. Mod. Phys.*, 55:331–402, April 1983; P. Fendley, Kinks in the Kondo problem, *Phys. Rev. Lett.*, 71:2485–2488, 1993; A. C. Hewson, *The Kondo Problem to Heavy Fermions*, 1997; R. M. White, *Quantum Theory of Magnetism*, 1983; J. W. Wilkins, Solution of the Kondo problem (at last), *Phys. Today*, 39(1):22, January 1986; K. G. Wilson, The renormalization group: Critical phenomena and the Kondo problem, *Rev. Mod. Phys.*, 47:773–840, October 1975.

Köppen climate classification system

A system for assigning short codes to the Earth's climate zones. It consists of a hierarchical arrangement of the major climatic groups, subcategories within

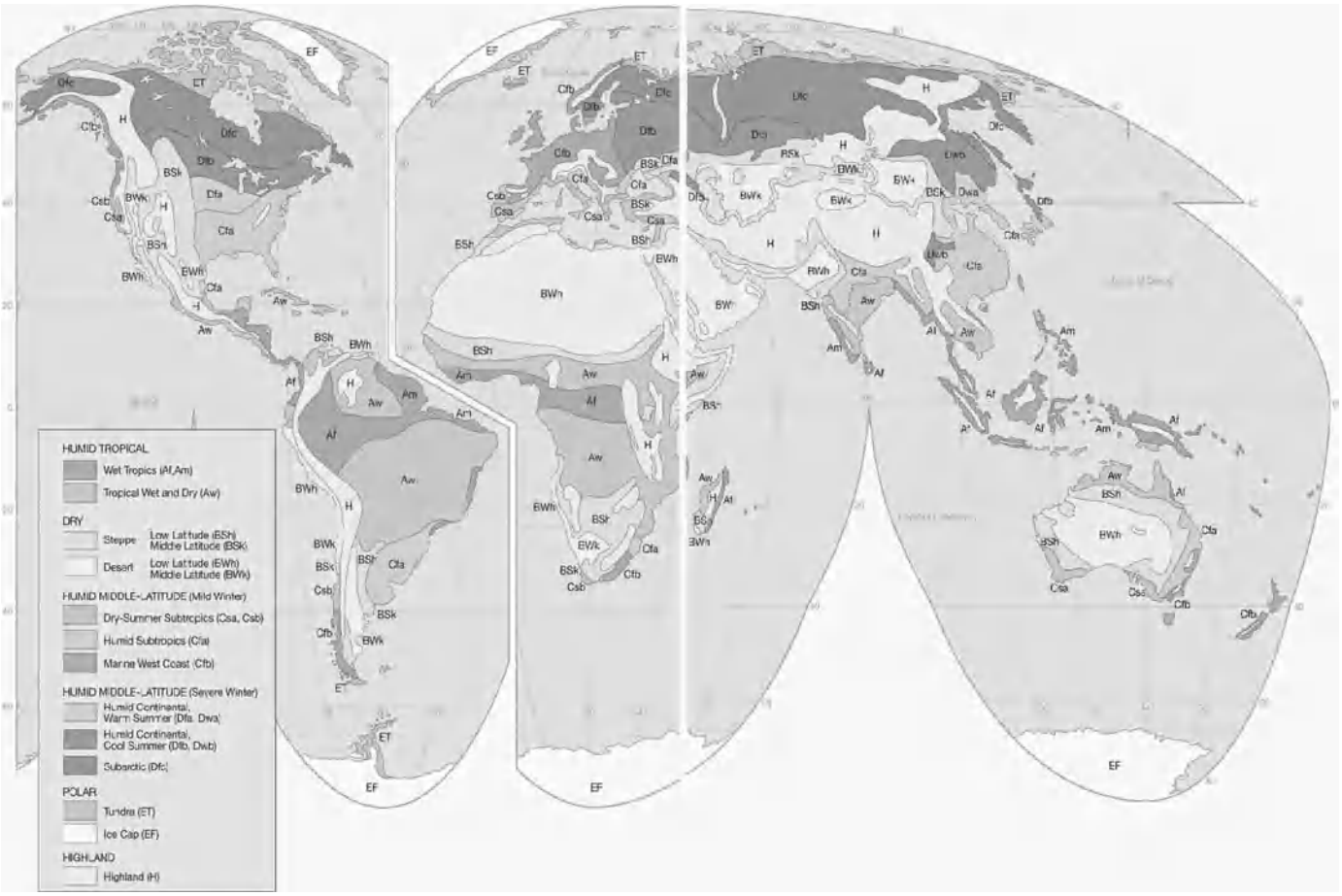

Generalized Köppen climate classification map. (*From E. Tarbuck and F. K. Lutgens, Earth Science, 11th ed., Prentice Hall, 2006*)

Köppen system

First division	Description	Second division	Description	Third division	Description
A	**Tropical** Coolest month's average temperature >64°F/180°C	m	A period of less precipitation in an otherwise very moist climate		
		w	Dry winters, wet summers		
		f	Constantly wet (no less than 2.4 in./60 mm of rain, in the driest month)		
B	**Dry** Potential evaporation and transpiration exceed precipitation	S	More precipitation than deserts, but under 20 in./500 mm per year (semiarid/steppe)	h	Mean annual temperature >64°F/18°C (tropical)
				k	Mean annual temperature <64°F/18°C (temperate)
		W	Less than 10 in./250 mm of precipitation per year (arid/desert)	h	Mean annual temperature >64°F/18°C (tropical)
				k	Mean annual temperature <64°F/18°C (temperate)
C	Midlatitude with mild winters	s	Mild with dry summer (mediterranean)	a	Hot summer season
				b	Warm summer season
		f	Mild with no dry season (humid subtropical)	a	Hot summer season
				b	Warm summer season
				c	Cool summer season
		w	Mild with dry winter	a	Hot summer season
D	Midlatitude with severe winters	f	Humid with severe winters; no dry season (humid-continental)	a	Hot summer season
				b	Warm summer season
				c	Cool summer season
				d	Cool summers and very cold winters
		w	Severe dry winters (subarctic)	a	Hot summer season
				b	Warm summer season
				c	Cool summer season
				d	Cool summers and very cold winters
E	**Polar** Average temperature of the warmest month <50°F/10°C	T	Warmest month's average temperature between 32°F/0°C and 50°F/10C		
		F	Warmest month's average temperature <32°F/0°C		

these groups, and in some cases further subdivisions describing seasonal variations within a climate set.

Climate may be thought of as the average weather in a given location. These general conditions vary over geographic space. Consider the differences between the hot wet tropics close to the Equator and the cold dry conditions typical of the Poles. While the Poles and the Equator illustrate extremes, for every place between them there exists an intermediate climate. Describing an area's average weather is difficult because atmospheric conditions, such as temperature and humidity, vary hourly, daily, seasonally, annually, and through longer periods. However, the climate zones may be identified by generalizations of several variables, based on long-range observations. *See* CLIMATOLOGY; DESERT; EQUATOR; NORTH POLE; SOUTH POLE.

Wladimir Köppen (1846–1940) was a German climatologist. Though weather data were incomplete at the time, Köppen, who was also a trained botanist, recognized that plant distributions were excellent proxies for physical conditions on land. With vegetation as a guide, he and his student/assistant, Rudolph Geiger, developed the first global climate classification map in 1928. The initial draft divided the climates into five major groups, based on inferred temperature and precipitation patterns. The zones were referred to as A (hot and humid), B (dry), C (midlatitude with mild winters), D (midlatitude with cold winters), and E (polar). *See* PLANT GEOGRAPHY; PLANTS, LIFE FORMS OF.

The first divisions in the hierarchy were quite broad and vague, encompassing many possible subclimates. Categories C, D, and E describe their geography more explicitly than they do their associated

climate, since latitude and associated solar intensity are among the most significant controls of a location's climate. In addition, category B describes only the moisture domain without regard to temperature range, as it may refer to hot or cool deserts.

While the major headings are alphabetical, the second and third (if present) subdivisions are less intuitively labeled. Since the subsequent letters are abbreviations for German words, some of them are capitalized and some are not. This has proven confusing to some. Twenty-three originally specified climates are described in the **table**.

The second division denotes further detail, once again considering temperature and/or precipitation characteristics. For example, category B is subdivided into arid regions (deserts), W, and semiarid regions, S, known as steppes. Desert regions are therefore designated BW in the Köppen system. However, some second-order categories denote temperature ranges instead; for example, the polar climate group consists of ice cap (most frigid) and tundra (less cold) subclasses.

Some categories are further divided into subclasses, which are indicated by lowercase letters. For example, deserts (BW) may be further described as warm or cool; the Köppen codes for these climates are BWh and BWk, respectively.

Several changes were made by Köppen, and other geographers have modified the system since his death. A major category for highland, H, was added after the publication of the first matrix, bringing the total number of terrestrial climates to 24. The **illustration** shows a simplified map of several Köppen climate zones.

Classifying climate, or almost any other physical trend, is an attempt to organize natural phenomena. As a result, there are inherent inconsistencies and subjectivity in almost all classification systems, and Köppen's is no exception. Some argue that Köppen's system ignores several significant variables, such as global wind belts, air mass source regions, and frontal zones. Furthermore, the study of global climate change has supported a call for a quantitative approach to climate classification. Digital meteorological data collected globally via satellite can be analyzed by computers, yielding highly accurate, timely, and detailed maps. Even small alterations of a region's climate can be monitored in real time. However, the availability of superior meteorological monitoring equipment, including remote sensors, and nearly 80 additional years of archived weather since Köppen's death have resulted in the development of only a few alternative classification systems. Even so, the descriptive, yet empirical system devised by Köppen remains a model still in use. *See* GLOBAL CLIMATE CHANGE; METEOROLOGICAL SATELLITES; REMOTE SENSING.

Paula Messina

Bibliography. A. Getis, J. Getis, and J. D. Fellman, *Introduction to Geography*, 8th ed., 2001; A. M. Strahler, *Introduction to Physical Geography*, 1965; E. J. Tarbuck and F. K. Lutgens, *Earth Science*, 11th ed., 2005.

Kronig-Penney model

An idealized, one-dimensional model of a crystal which exhibits many of the basic features of the electronic structure of real crystals. Consider the potential energy $V(x)$ of an electron illustrated in **Fig. 1** with an infinite sequence of potential wells of depth $-V_o$ and width a, arranged with a spacing b. The Schrödinger wave equation can be readily solved for such an arrangement to give electron energy as a function of wave number.

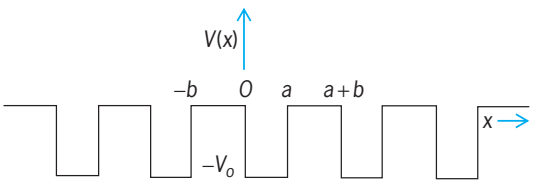

Fig. 1. Potential energy which is assumed for the one-dimensional Kronig-Penney model.

The energy bands that are thus obtained with the choice of constants given in Eq. (1) and (2) are

$$\left(\frac{2ma^2V_o}{b^2}\right)^{1/2} = 12.0 \qquad (1)$$

$$\frac{b}{a} = 0.1 \qquad (2)$$

shown in **Fig. 2**. Notice that the width and the curvatures of the allowed bands increase with energy. The Kronig-Penney model has been extended to include the effects of impurity atoms. A solution can also be derived from the Dirac equation. *See* BAND

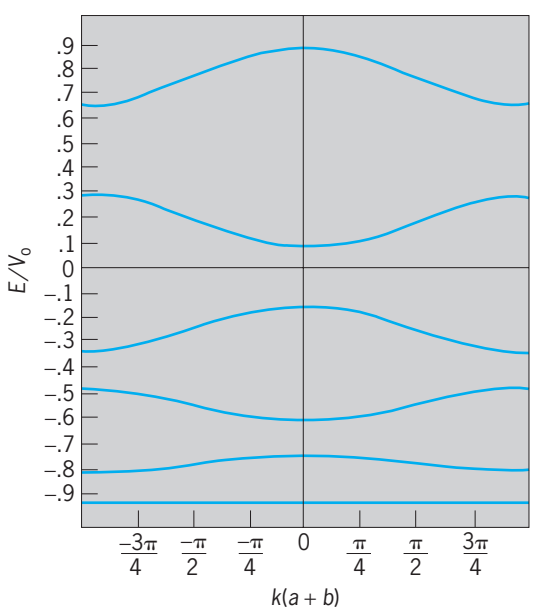

Fig. 2. The ratio of the electron energy E to the magnitude of the well depth V_o shown as a function of electron wave number k times the "lattice constant" $(a + b)$ for the Kronig-Penney model.

THEORY OF SOLIDS; NONRELATIVISTIC QUANTUM THEORY. Joseph Callaway

Bibliography. C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; R. de L. Kronig and W. G. Penney, Quantum mechanics of electrons in crystal lattices, *Proc. Roy. Soc. (London), Ser. A*, 130:499–513, 1931; G. Wannier, *Elements of Solid State Theory*, 1959, reprint 1970.

Krypton

A gaseous chemical element, Kr, atomic number 36, and atomic weight 83.80. Krypton is one of the noble gases in group 18 of the periodic table. Krypton is a colorless, odorless, and tasteless gas. The **table** gives some physical properties of krypton. The principal use for krypton is in filling electric lamps and electronic devices of various types. Krypton-argon mixtures are widely used to fill fluorescent lamps. *See* INERT GASES; PERIODIC TABLE.

1												13	14	15	16	17	18
1 **H**	2																2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

The only commercial source of stable krypton is the air, although traces of krypton are found in minerals and meteorites.

Physical properties of krypton

Property	Value
Atomic number	36
Atomic weight (atmospheric krypton only)	83.80
Melting point, triple point °C	−157.20
Boiling point at 1 atm pressure, °C	−153.35
Gas density at 0°C and 1 atm pressure, g/liter	3.749
Liquid density at its boiling point, g/ml	2.413
Solubility in water at 20°C, ml krypton (STP) per 1000 g water at 1 atm partial pressure krypton	59.4

A mixture of stable and radioactive isotopes of krypton is produced in nuclear reactors by the slow-neutron fission of uranium. It is estimated that about 2×10^{-8}% of the weight of the Earth is krypton. Krypton also occurs outside the Earth. Arthur W. Francis

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; M. Ozima and F. A. Podosek, *Noble Gas Geochemistry*, 2001.

Kudzu

A perennial vine legume, capable of rapid growth in a warm temperate, humid subtropical climate. The name kudzu has a Japanese origin. Kudzu (*Pueraria thunbergiana*) was introduced into the United States in 1876 and used as a shade plant until 1906, when a few enthusiastic growers in the southeastern United States began to use it as a forage crop, a practice that continued for 30 years. It was then promoted as a soil-conserving plant. However, much prejudice developed against its use because of its spread into forest borders, drainage ditches, and other areas.

Kudzu is not adapted to tropical or arid climates or to alkaline soils, and probably requires winter cold for growth rejuvenation. The technique for growing kudzu successfully is to set out a few well-developed plants, use enough commercial fertilizer or equivalent for good growth, and protect the plants from weeds, insects, and grazing animals. The vines will spread over the area to be covered, producing roots at the nodes, which, unless killed by severe winter freezing, become independent plants. Thus, a few plants produce many in 2 or 3 years if adequately protected. *See* FERTILIZING.

Kudzu produces moderate yields of forage, but it must be grazed with care to prevent loss of stand. Since the viney stems are difficult to harvest for hay, adapted grasses and clovers are preferred as forage crops.

A simple, economical method is needed for keeping kudzu from spreading into areas where it is not wanted. It may be stopped by a permanent pasture at the border of a kudzu field, or by repeated cultivation or harrowing the borders of cropland. No acceptable method of spread prevention has been developed for uncultivated or nonpastured areas.

Tropical kudzu (*P. phaseolides*) is one of the most important and widely planted cover and green manure crops of the tropics. It makes rapid vigorous growth, providing quick ground cover and suppressing most other vegetative growth. It is used as a forage crop although careful management is required to prevent complete domination of mixtures with grasses and other species. Its habit of growth is similar to kudzu of subtropical areas. *See* LEGUME FORAGES; ROSALES. Paul Tabor

Kuiper Belt

A vast reservoir of icy bodies in the region of the solar system beyond the orbit of the planet Neptune. The trans-Neptunian region of the solar system was thought to be home only to the small planet Pluto and its satellite Charon until the discovery of the Kuiper

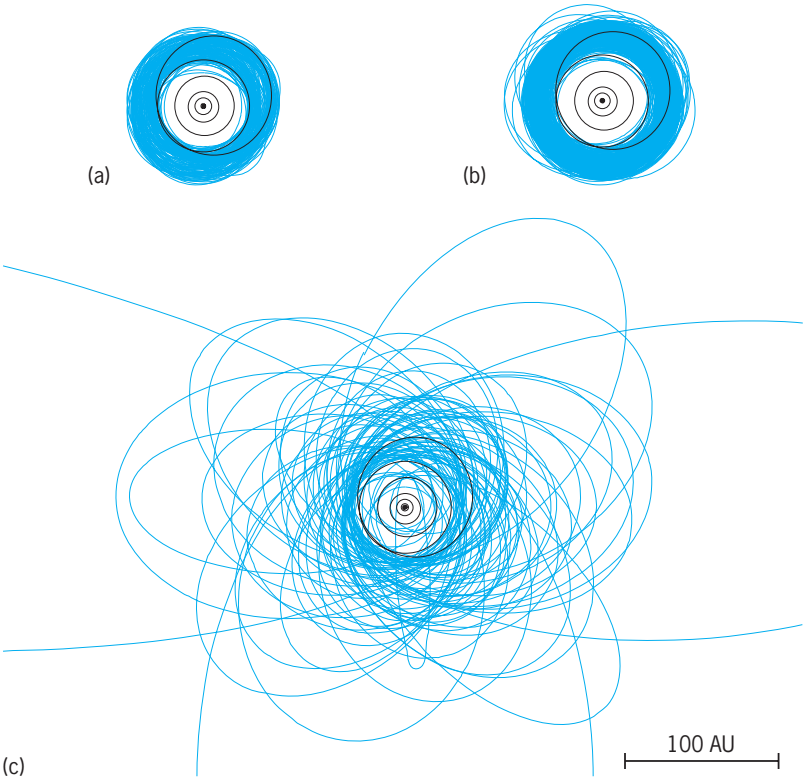

Plane views of the outer solar system showing (*a*) resonant Kuiper Belt objects, (*b*) classical KBOs, and (*c*) scattered KBOs. The types of KBOs are represented by a band of colored orbits in each diagram, while the black concentric circles in the center of each diagram are the orbits of the giant planets, and the black ellipse is the orbit of Pluto. (*Ronnie Hoogerwerf, Smithsonian Astrophysical Observatory*)

Belt object 1992 QB_1 in 1992 by David Jewitt and Jane Luu. Members of the belt (Kuiper Belt objects, or KBOs) are part of the original building blocks (planetesimals) of the outer planets, left behind simply because they were not incorporated into planets or ejected from the solar system. They are relics from the solar nebula, the original disk of gas and dust that gave rise to everything in the solar system. As such, Kuiper Belt objects should preserve records of the early stages of the solar system's evolution. *See* NEPTUNE; PLANET; PLUTO.

Prediction and discovery. The idea that distant planetesimals that did not agglomerate into planets might still exist beyond the planets was proposed in 1951 by Gerard Kuiper. However, Kuiper had a little-known predecessor: the concept of a trans-Neptunian belt of icy bodies was first mentioned by K. E. Edgeworth in 1949. In the 1980s the concept of the Kuiper Belt was revived as a possible source for the short-period comets (those which take less than 200 years to complete one orbit around the Sun). Long-period comets, whose orbital periods are longer than 200 years, have long been recognized to come from the Oort Cloud, a large spherical cloud of comets at roughly 10,000 astronomical units (AU) from the Sun. (An astronomical unit is equal to the distance between the Earth and the Sun, about 1.5×10^8 km or 9.3×10^7 mi.) However, the Oort Cloud could not be the source for the short-period comets, and theoretical arguments suggested that the short-period comets likely originated in a comet belt beyond Neptune. *See* COMET.

Edgeworth and Kuiper had no theoretical basis for their bold idea. They simply thought that there was no a priori reason why the solar system should end abruptly at Neptune or Pluto. Perhaps large planets could not form beyond Neptune, but smaller bodies might be able to. Edgeworth and Kuiper envisioned a population consisting of residual planetesimals left over after planets had finished their formation. Because of the low temperatures in the trans-Neptunian area, these distant objects were likely to be ice-rich, identical to comets. At the time of Edgeworth's and Kuiper's hypothesis, Pluto was the only permanent solar system object known beyond Neptune, and the apparent emptiness of the outer solar system was rather startling compared to the rich abundance of bodies in the inner solar system. In 1987 Jewitt and Luu started a telescopic survey to test this apparent emptiness of the outer solar system. The search finally succeeded in August 1992, when 1992 QB_1 was found.

Population. The Kuiper Belt begins just beyond Neptune and extends to some as yet unknown outer limit. KBOs are not distributed uniformly throughout the belt, but lie in dense bands separated by empty space (see **illustration**). The bands represent stable regions where bodies can survive for the age of the solar system, while empty regions represent unstable areas where previous KBOs have been cleared out by the gravitational influence of Neptune. Based on how their orbits are distributed throughout the region, KBOs are classified as resonant, classical, or scattered.

Resonant KBOs. The Kuiper Belt contains several special areas called mean motion resonances, which are best explained by an example: Located at 39.5 AU, Pluto lies in the 3:2 resonance with Neptune. This means that every time Neptune completes three orbits around the Sun, Pluto completes two such orbits. In the same vein, objects in the 7:4 resonance complete 4 orbits for every 7 orbits completed by Neptune, objects in the 8:5 complete 5 orbits for Neptune's 8 orbits, and so forth. These special relationships with Neptune protect the resonance objects from close encounters with the planet, and the resonance orbits are therefore long-lived. Resonances thus act as stable oases in space where members can survive for the age of the solar system. Just as the 3:2 resonance protects Pluto, it also protects other objects in this resonance (called plutinos to highlight their dynamical similarity with Pluto). The current estimate is that the resonant KBOs (with the plutinos dominating in number) make up roughly 20% of the Kuiper Belt population (illus. *a*), with Pluto being the largest known Kuiper Belt object.

Classical KBOs. The classical KBOs are characterized by near-circular orbits that lie almost entirely in the ecliptic plane between 42 and 48 AU (outside the resonances) and make up about 45% of the total population (illus. *b*). These orbits are what would be expected from the first-generation planetesimals in the solar nebula, suggesting that the classical KBOs

are indeed primordial planetesimals that have managed to preserve their original orbits. Unexpectedly, the classical Kuiper Belt possesses a well-defined edge at approximately 50 AU: as of late 2006, all known KBOs beyond this edge belong to the scattered Kuiper Belt population (see below). The origin for this edge is still not clear: the proposed hypotheses generally postulate that the classical Kuiper Belt beyond 50 AU had been stripped by scattering processes, such as the passage of Earth-sized bodies through the belt or a close encounter with a star. The competing hypotheses make testable predictions, and astronomers hope that as the structure of the Kuiper Belt is better defined, it will help narrow down the competing hypotheses.

Scattered KBOs. The scattered KBOs make up the remaining 35% of the Kuiper Belt, and they easily stand out from the rest of the belt with their very large, very elliptical orbits (illus. *c*). Whereas the known classical and resonant KBOs are clustered inside 50 AU, the object 2000 OO_{67}, the most distant scattered Kuiper Belt object yet discovered, reaches at least to 1053 AU at its farthest point from the Sun. Several theories have been proposed to explain the origin of the scattered KBOs. One theory postulates that the belt is a by-product of the scattering process that produced the Oort Cloud. As Uranus and Neptune reached their full sizes, they scattered nearby comets to very large distances, forming the Oort Cloud; those comets scattered to less extreme distances would have formed the scattered Kuiper Belt population.

Total population. The current total mass of the Kuiper Belt population between 30 and 50 AU is estimated to be roughly 0.1 earth mass, 100 times more massive than the asteroid belt. About 1100 KBOs were known in August 2006 with diameters estimated in the 100–1000-km (60–600-mi) range. (In comparison, Uranus and Neptune are about 50,000 km or 30,000 mi in diameter, while Pluto has a diameter of 2320 km or 1440 mi.) There are currently about 30,000 KBOs with diameters larger than 100 km (60 mi) in the entire Belt, making them roughly 100 times as numerous as similarly sized bodies in the asteroid belt.

Large KBOs. Objects smaller than 100 km are numerous but too small to be detected with conventional telescopes; on the other hand, bodies with diameters of 1000 km (600 mi) or more are easily detectable from the ground but are relatively rare. Surveys have been undertaken that are aimed at both large and small KBOs. By August 2006, roughly 10 KBOs in the 1000-km-diameter range had been discovered by several surveys specifically aimed at the largest bodies. An example of such a survey is the Palomar survey carried out by the California Institute of Technology. Like all other surveys aimed at KBOs, its strategy is to search as much of the sky as possible, at the expense of sensitivity; this strategy works because large KBOs are bright but very rare, so the key to finding them is large sky coverage. These 1000-km objects bridge the gap between the relatively abundant smaller (100 km diameter) bodies and Pluto (2320 km diameter), hitherto the largest Kuiper Belt object known. However, other bodies may have been found that would strip Pluto of such distinction. A case in point is the KBO 2003 UB_{313}, named v36 199 Eris in September 2006, which soon after its discovery was postulated to be even larger than Pluto based on its absolute magnitude. The absolute magnitude is the brightness of an object corrected to a standard distance from the Sun and the Earth, and is an indication of an object's size. However, since the absolute magnitude is equal to the albedo (reflectivity) times the squared diameter, an absolute magnitude brighter than Pluto's is not sufficient to indicate unambiguously a larger size. If the albedo were known, the diameter could be disentangled from the absolute magnitude. The albedo is usually determined by observing an object in the thermal infrared, measuring the object's heat radiation. Such measurements are difficult for KBOs because, due to their great distance from the Sun and cold temperatures, their thermal radiation peaks near 60–70 micrometers, where the Earth's atmosphere is opaque. Until space-based telescopes (such as, NASA's *Spitzer Space Telescope*) can deliver thermal measurements of a large sample of KBOs, the albedos of the largest KBOs can be determined via observations in the submillimeter. As of August 2006, the true size of Eris was still uncertain. Observations by the *Hubble Space Telescope* suggested that Eris was roughly 5% larger than Pluto, but measurement uncertainties prevent a definitive assessment. Whether Eris is larger than Pluto remains to be confirmed by future measurements. *See* ALBEDO; INFRARED ASTRONOMY; SPITZER SPACE TELESCOPE; SUBMILLIMETER ASTRONOMY.

The discovery of each large KBO is often greeted with great excitement as it raises the hope of adding yet another planet to the solar system's inventory, but in truth such large objects are expected if, as is widely believed, KBOs formed basically in the same manner as other bodies in the solar system, that is, by accretion from smaller bodies (planetesimals). According to current accretion models, about 100 such bodies are expected. Besides their appeal to the general public, large KBOs are important for several scientific reasons:

(1) They are bright enough for detailed investigations of their physical properties. The more common 100-km bodies are too faint for anything other than the most rudimentary measurements (such as optical brightness through various filters), and offer little information on their physical makeup. Barring close-up visits with spacecraft in the distant future, astronomers could only hope to extract meaningful physical information such as surface composition, shape, and spin rate from the large bright KBOs.

(2) The number of these large bodies, compared to the number of smaller bodies, is a test of the existing accretion models. These models predict the relative numbers of large and small KBOs, and their confirmation (or the lack thereof) has significant implications for the conditions in the solar nebula when the KBOs were forming.

(3) The largest KBOs seem to be more reflective

than the smaller ones. This suggests that large KBOs may be better at retaining icy material than their smaller counterparts, much like Pluto (albedo = 0.6).

90377 Sedna. In 2004 the Kuiper Belt object 90377 Sedna was discovered in the Palomar survey. Sedna is of great interest for two reasons:

(1) Sedna's perihelion distance (closest approach to the Sun) is 76 AU, well beyond the scattering influence of Neptune. In this sense it resembles the Kuiper Belt object 2000 CT_{105} (perihelion distance 44 AU). It is not clear how these two KBOs attained their orbits beyond Neptune. Their existence seems to suggest an early epoch of strong scattering events during which smaller (less than planet size) solar system bodies were scattered by a configuration of planet-size bodies different from what is known today.

(2) Sedna's large size (1000–1500 km or 600–900 mi) means that it could not have formed in its current orbit far beyond the planets; there is little doubt that Sedna must have formed closer in and was subsequently scattered outward. Sedna's classification is still uncertain. Since it lies between the known Kuiper Belt and the classical Oort Cloud, it has been suggested to be part of the Inner Oort Cloud. This inner part of the Oort Cloud may connect the classical Oort Cloud and the Kuiper Belt.

Small KBOs. Although small KBOs do not tend to generate as much excitement as large ones, there is also much interest in how to detect the small end of the size distribution (KBOs with kilometer-size diameter or smaller). These very small bodies are too faint to be detected by conventional imaging methods and are generally expected to be "seen" only by the occultation method. With this method, the presence of these bodies can be inferred when the bodies pass in front of stars and momentarily block out the stars' light (they occult the stars). The most concerted effort to detect KBOs via occultation is the Taiwan American Occultation Survey (TAOS), which is aimed at detecting KBOs a few kilometers in diameter. The survey's four 0.5-m (20-in.) robotic telescopes will monitor 3000 stars every clear night for several years to search for these occultation events and will send out alerts when events occur. Follow-up observations can then be carried out with larger telescopes. However, there is already a tantalizing result on the smallest end of the Kuiper Belt population: in 2006 a group of Taiwanese astronomers reported occultations in x-rays from Scorpius X-1, the brightest source of x-ray light in the sky, and attributed these occultations to 10–100-m bodies in the Kuiper Belt. However, the occultation rate implies 10^{15} bodies in the 10–100-m range, a population much larger than that predicted by models (10^9–10^{12}). This result remains to be confirmed either by another occultation survey or by space-based missions that can actually monitor the diffraction patterns of the occultation events and establish the distance of the occulting bodies. *See* OCCULTATION.

Population loss. The Kuiper Belt has been much eroded since its formation. The original Kuiper Belt is estimated to have contained roughly 10 earth masses, or 100 times as much mass as the present Kuiper Belt.

Physical properties. The properties of KBOs are of intense interest since these bodies are primordial survivors from the early solar nebula. Unfortunately, most KBOs are so faint that detailed investigations of their properties are extremely difficult. The best determined property is probably their surface colors, which turn out to cover a broad range from neutral to very red. This may be explained in the context of their composition, which is believed to be rich in molecular ices that are abundant in the solar system (water, ammonia, and methane). Laboratory experiments lead to the prediction that after prolonged irradiation by high-energy particles such as cosmic rays and solar-wind particles, KBOs should be covered by an irradiation mantle, a carbon-rich crust that covers the ices buried underneath. Occasional collisions among KBOs may puncture the irradiation crust, revealing pristine unirradiated material which would be of a different color from the crust. The degree to which a Kuiper Belt object is covered by craters and impact debris may explain the wide range of colors.

Given the difficulties in determining the surface properties of KBOs, astronomers look to brighter (larger) analogs of KBOs for clues. A popular counterpart of the KBOs is Triton, a 2700-km-diameter (1680-mi) satellite of Neptune; it orbits Neptune "backwards" (in retrograde orbit) and therefore is suspected to be a captured Kuiper Belt object. Pluto, the largest Kuiper Belt object known, is also often used as a template for KBOs. However, the atmospheres of both Triton and Pluto complicate comparisons with the average Kuiper Belt object with regard to surface properties such as spectra and albedos.

The Kuiper Belt also contains other binary systems besides Pluto and its satellite Charon: several KBOs are now known to have their own moons. Like Charon, these satellites of KBOs are large compared to the primary (as much as one-third the size), but unlike Charon they are also quite far (10,000–100,000 km or 6000–60,000 mi) from the primary. Their formation mechanism, like so many other aspects of the Kuiper Belt, remains a mystery. Binary formation is strongly affected by the dynamical environment in which it occurs, so once understood, binary formation promises to yield important clues to the early Kuiper Belt environment. *See* BINARY ASTEROID; SOLAR SYSTEM.

Jane Luu

Bibliography. D. Jewitt and J. Luu, Discovery of the candidate Kuiper Belt object 1992 QB_1, *Nature*, 362:730–732, 1993; G. P. Kuiper, Planets and satellites, in G. P. Kuiper and B. M. Middlehurst (eds.), *The Solar System*, University of Chicago Press, 1961; J. Luu and D. Jewitt, Kuiper Belt objects: Relics from the accretion disk of the Sun, *Annu. Rev. Astron. Astrophys.*, 40:63–101, 2002.

Kumquat

Shrubs or small trees that are members of the genus *Fortunella*, which is one of the six genera in the group of true citrus fruits. Kumquats are believed to

Fruit cluster and foliage of Nagami kumquat (*Fortunella margarita*). (*McFarland Co.*)

have originated in China and the Malay Peninsula, but are now widely grown in all citrus areas of the world. Of the several species the most common are *F. margarita*, which has oval-shaped fruit (see **illus.**), and *F. japonica*, which has round fruit. *See* SAPINDALES.

The kumquat's stems, leaves, flowers, and fruits resemble those of *Citrus*. Kumquats bear numerous flame- to orange-colored fruits of small size, often less than 1 in. (2.5 cm) in diameter, having three to five locules filled with an acid pulp and a sweet, edible pulpy rind. The trees are among the most cold-hardy of the citrus fruits; they stay dormant even during protracted warm spells in the winter months, which enables them to withstand low temperatures.

Kumquats, with their brilliant orange-colored fruits and dense green foliage, are highly ornamental and are most frequently grown for this reason. Sprays of foliage and fruit are commonly used for decoration, particularly in gift packages of ordinary citrus fruits. Kumquat fruits can be eaten whole without peeling; they are also used in marmalades and preserves and as candied fruits. *See* FRUIT, TREE.

Frank E. Gardner

Kuroshio

A swift, intense current flowing northeastward off the coasts of China and Japan in the upper waters of the North Pacific Ocean. The Kuroshio is the western portion of a giant clockwise, horizontal circulation known as the North Pacific subtropical gyre (**Fig. 1**). This circulation extends from 15° to 45°N across the entire width of the Pacific Ocean. It is driven by the large-scale winds—the trades in the south and the westerlies in the north. As with all other western boundary currents, such as the Gulf Stream, the effect of the Earth's rotation and its spherical shape is to concentrate the Kuroshio flow into a current that is only about 100 km (50 nmi) wide with speeds up to 2 m/s (4 mi/h). *See* CORIOLIS ACCELERATION; GULF STREAM; PACIFIC OCEAN.

The Kuroshio (Japanese, meaning "Black Current") has an apparent blackness resulting from the water clarity, which is a consequence of the low biological productivity of seawater in the area. It originates off the southeast coast of Luzon, the main island of the Philippines. For the first 1000 km (500 nmi), the Kuroshio flows northward along the east coasts of Luzon and Taiwan, until it enters the East China Sea. For the next 1000 km, it flows northeastward near the edge of the continental shelf off eastern China, until it exits the East China Sea through the Tokara Strait. During its final 1000 km, it flows east-northeastward off the southern coast of Japan (where it is sometimes called the Japan Current). Finally, it leaves the Asian coast near Tokyo and travels into the interior of the North Pacific Ocean as a slowly expanding jetlike current known as the Kuroshio Extension (**Fig. 2**). Here it merges with the Oyashio, a cold current with high biological productivity, and becomes the North Pacific Current (Fig. 1).

Like the Gulf Stream in the North Atlantic, the Kuroshio rapidly carries large quantities of warm water from the tropics into midlatitude regions. It is consequently an important agent in redistributing global heat. North of 30°N, where prevailing winds are westerlies, the North American climate is strongly affected by the warmth of these waters. Kuroshio volume transport is huge (more than 100 times that of the Amazon River): typically 2–3 × 10^7 m^3/s (7–10 × 10^8 ft^3/s) until leaving the East China Sea, and then growing to 6 × 10^7 m^3/s (20 × 10^8 ft^3/s) south of Japan. One reason for the increase is the merging of the Kuroshio flow from the East China Sea with northward subtropical gyre flow traveling outside this sea, along the eastern edge of the Ryukyu island chain. Another reason is the contribution of the flow occurring in a local recirculating gyre (Fig. 2).

The Kuroshio and Kuroshio Extension (like the Gulf Stream) are thought to be flanked by a number of such recirculating gyres, which are typically several hundred kilometers across. According to theory, these small gyres may be caused by eddying motions, which are a result of the Kuroshio flow becoming unstable. The instability first appears as side-to-side motions or "meanders" of the Kuroshio current core. The meanders are wavelike in character and usually

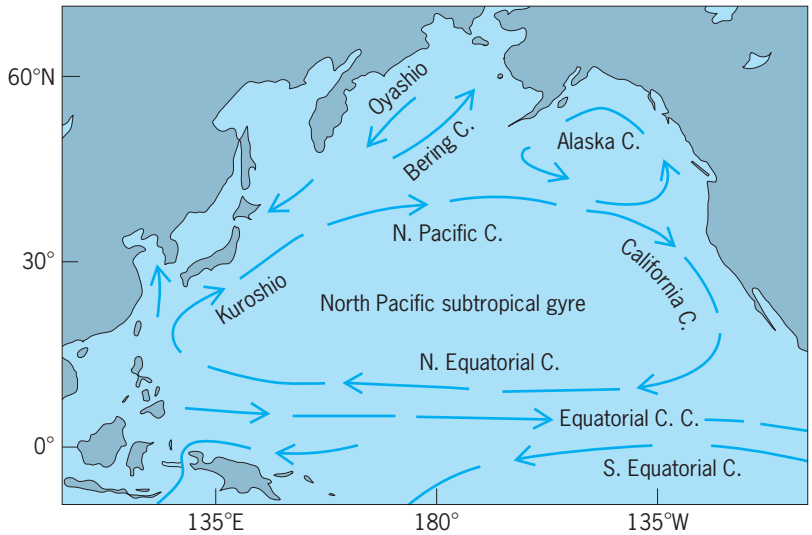

Fig. 1. General surface circulation of the North Pacific Ocean. (*Adapted from Fig. 2.17 in J. R. Apfel, Principles of Ocean Physics, 1987, Academic Press*)

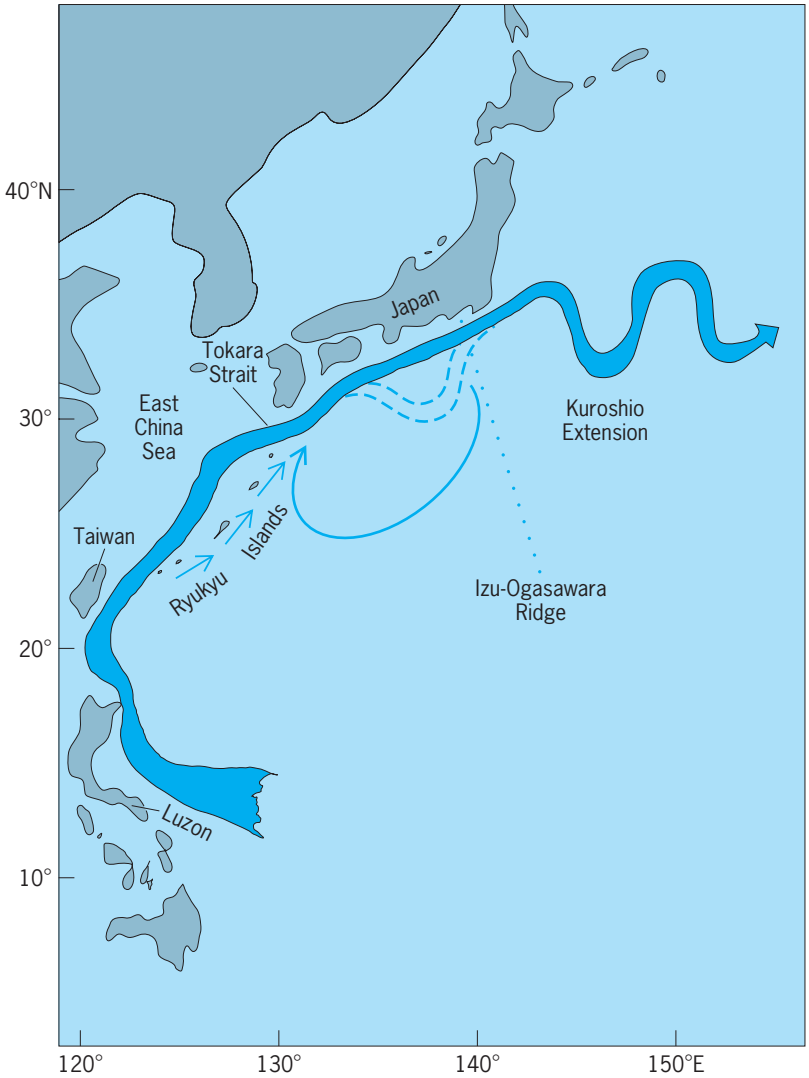

Fig. 2. Flow path of the Kurishio and Kurishio Extension. Broken lines show the alternative "large-meander" path south of Japan. Smaller arrows show other flows described in the text. (*After Y. G. Guo, The Kurishio, Part II, Oceanog. Mar. Bio. Annu. Rev., 29:155–189, 1991*)

propagate slowly (<25 km/day or 0.5 nmi/h) downstream. Occasionally, especially in the Extension region, such a meander can grow until it pinches off into a separate, spinning ring of water about 200 km (100 nmi) in diameter, which can survive for many seasons.

One unique form is the so-called Kuroshio large meander. This stationary meander is typically a 300-km (180-mi) southward loop which forms intermittently off southeast Japan. At present, it occurs about one-third of the time, and once formed lasts 2 to 10 years. It takes place just upstream of where the Kuroshio crosses the Izu-Ogasawara Ridge, and the ridge appears to play a role in causing the meander (Fig. 2). There is evidence that the large meander is associated with weak subtropical gyre flow resulting from weakening of atmospheric high pressure over the subtropical North Pacific. Proxy indicators, in dated sediments from the northern Izu-Ogasawara Ridge, have shown how the tendency to form a large meander there has changed over time, with a pronounced minimum 7000 years ago. These data suggest a strong atmospheric subtropical high at that time, with resulting swift winds and intense gyre circulation. *See* PALEOCLIMATOLOGY.

The Kuroshio is readily apparent in data obtained from satellites. In particular, warm surface water carried by the Kuroshio shows up clearly in sea-surface temperature images. Also, satellite altimeters, which are today able to detect sea-surface height changes of a few centimeters, can readily measure the change associated with the Kuroshio: a phenomenon known as geostrophy, resulting from the rotation of the Earth, makes the offshore side of the Kuroshio about 1 m (3 ft) higher than the inshore side. *See* GEOSTROPHIC WIND; REMOTE SENSING.

Mark Wimbush

Bibliography. Y. J. Guo, The Kuroshio, Part II: Primary productivity and phytoplankton, *Oceanog. Mar. Biol. Annu. Rev.*, 29:155–189, 1991; Y. Hirota, The Kuroshio, Part III: Zooplankton, *Oceanog. Mar. Biol. Annu. Rev.*, 33:151–220, 1995; N. G. Hogg and W. E. Johns, Western boundary currents, *Rev. Geophys.*, 33(suppl.):1311–1334, 1995; H. Kawai, A brief history of recognition of the Kuroshio, *Prog. Oceanog.*, 41:505–578, 1998; K. Sawada and N. Handa, Variability of the path of the Kuroshio ocean current over the past 25,000 years, *Nature*, 392:592–595, 1998; H. Stommel and K. Yoshida (eds.) *Kuroshio: Physical Aspects of the Japan Current*, University of Washington Press, 1972; J. L. Su, B. X. Guan, and J. Z. Hang, The Kuroshio, Part I: Physical features, *Oceanog. Mar. Biol. Annu. Rev.*, 28:11–71, 1990.

Kutorginata

A small extinct class and one of the most primitive types of rhynchonelliformean brachiopods. The kutorginates were widely distributed from the Early Cambrian and became extinct at the end of the Middle Cambrian. *See* BRACHIOPODA; RHYNCHONELLIFORMEA.

Classification. The Kutorginata appear in the classification scheme as follows:

Phylum Brachiopoda
- Subphylum Rhynchonelliformea [Articulata]
  - Class Kutorginata
    - Order Kutorginida
      - Superfamily Kutorginoidea
        - Family Kutorginidae
          - 5 genera (Early to Middle Cambrian)
- Superfamily Nisusioidea
  - Family Nisusiidae
    - 5 genera (Early to Middle Cambrian)

Morphology. Kutorginates have a biconvex, fibrous, impunctuate (lacking holes) shell (see **illustration**). On the posterior margin of the shell, the pedicle (a tough but flexible stalklike appendage used for attachment) protrudes through a large median opening (delthyrium and notothyrium) located between the valves. In the ventral valve

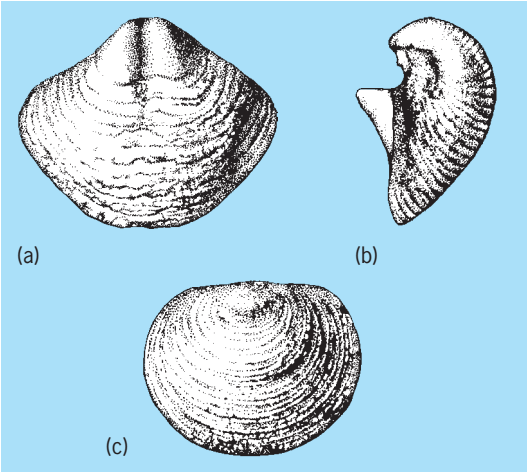

***Kutorgina*, showing (*a*) exterior of pedicle (ventral) valve; (*b*) lateral view of shell; (*c*) exterior of brachial (dorsal) valve. (*After C. D. Walcott, Cambrian Brachiopoda, USGS Monogr., vol. 51, 1912*)**

(formerly named pedicle valve), this opening is partly covered by a triangular convex plate (pseudodeltidium) bounded laterally by furrows, acting in the shell articulation. The pedicle appears at least as strong as that of the coeval (contemporary) lingulids. The putative coelomic cavity is similar anatomically to those of the Cambrian linguliformeans. *See* LINGULIDA.

The articulation of the shell consists of edges of the interareas without teeth and dental sockets: the valves rotate on simple hinge mechanisms that are different from those of other rhynchonelliformean brachiopods. A small apical foramen (opening) may occur on the ventral valve.

The visceral cavity is in a posteromedian position and extends anteriorly over one-third of the shell length. The muscle arrangement is very similar to that of other rhynchonelliformean brachiopods with adductor and diductor muscle systems (that is, muscles used to close and open the shell, respectively). The complete alimentary canal consists of the mouth located at the base of the lophophore (a food-gathering and respiratory organ), a posterodorsal esophagus, an enlarged pouch representing the stomach, a narrow intestine, and a dorsoposterior anus located near the proximal end of the pedicle. The adult lophophore development attains an early spirolophe (a lophophore with spiraled appendages) stage.

The main mantle canals known as vascula lateralia are exclusively disposed radially (pinnate type) in both valves. On the dorsal side, the vascula media are bifurcate.

Members of this group were sessile, but the epifaunal (surface-attaching) position remains questionable.

Christian C. Emig

Bibliography. R. L. Kaesler (ed.), *Treatise on Invertebrate Paleontology*, Part H, revised, *Brachiopoda* Geological Society of America and University of Kansas, 2000; L. E. Popov, A. J. Rowell, and J. S. Peel, Early Cambrian brachiopods from North Greenland, *Palaeontology*, 40:337–354, 1997; A. J. Rowell and N. E. Caruso, The evolutionary significance of *Nisusia sulcata*, an early articulate brachiopod, *J. Paleontol.* 59:1227–1242, 1985; C. B. Skovsted and L. E. Holmer, Early Cambrian brachiopods from Northeast Greenland, *Paleontology* 48(2):325–345, 2005; Z. Zhang et al., Rhynchonelliformean brachiopods with soft-tissue preservation from the Early Cambrian Chengjiang Lagerstätte of South China, *Plaeontology*, in press, 2006.

Labradorite

A plagioclase feldspar with composition $Ab_{50}An_{50}$ to $Ab_{30}An_{70}$ ($Ab = NaAlSi_3O_8$; $An = CaAl_2Si_2O_8$), occurring in basic igneous rocks, that is, basalts, gabbros, and anorthosites. Its hardness is 6–6.5; mean refractive index, 1.56; and specific gravity, 2.70. Labradorite crystals are translucent to dark or cloudy (from inclusions) but often display brilliant colors like those in peristerites and Huttenlocher intergrowths. In nature, most plagioclases throughout the composition range $Ab_{54}An_{46}$ to $Ab_{40}An_{60}$ are exsolved into two e-plagioclases that have compositions $Ab_{56\pm4}An_{44\mp4}$ and $Ab_{42\pm6}An_{58\mp6}$ and are coherently intergrown as somewhat irregular lamellar slabs 50 to 20 nanometers thick and known as Boggild intergrowths (see **illus.**). Intense colors may result from constructive interference of selected wavelengths of light reflected at successive interfaces between the alternating Ab- and An-rich lamellae. In fact, interference colors are so commonly observed in this composition range that the phenomenon has been called labradorescence. Gemologists often use the nonmineralogical term spectrolite in describing them. Iridescent colors in hummingbird feathers and the wings of *Morpho* butterflies are physically similar phenomena. *See* ALBITE; ANDESINE; ANORTHOSITE; BASALT; FELDSPAR; GABBRO; IGNEOUS ROCKS.

Paul H. Ribbe

Labyrinthulia

Protozoa forming a subclass of Rhizopodea with obscure relationships to rest of the class. There is one order, Labyrinthulida (see **illus.**). The mostly marine, ovoid to spindle-shaped, uninucleate organisms secrete a network of filaments (slime tubes) along which they glide, usually singly at rates of 4–150 micrometers/min. This network inspired the name "net slime molds" sometimes applied to them. The mechanism of locomotion is unknown. The unique *Labyrinthula minuta* may move for about 20 μm, then stop, and later move again, either

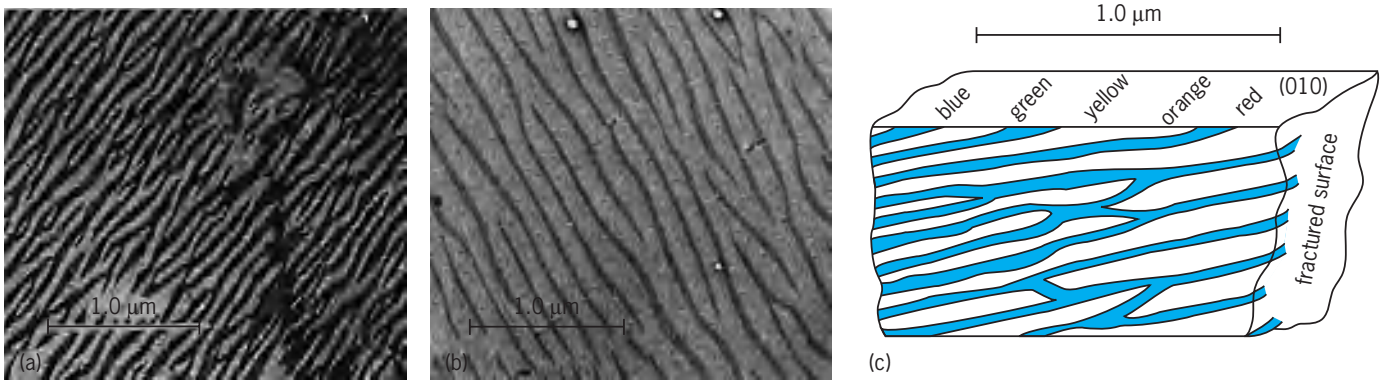

Photoemission electron microscope images of labradorite from an anorthosite in Finland. The more calcic phase is the lighter of the two. The apparent lamellar thickness are greatly exaggerated because the specimens are cut at $\sim 10^\circ$ to the plane of the lamellae: (*a*) "blue" area and (*b*) "red" area (*courtesy of F. Laves*). (*c*) Schematic showing the approximate relations of lamellar thickness and interference color, as observed on the (010) cleavage surface. Bulk composition of the labradorite changes from about $Ab_{50}An_{50}$ (blue) to $Ab_{44}An_{56}$ (red).

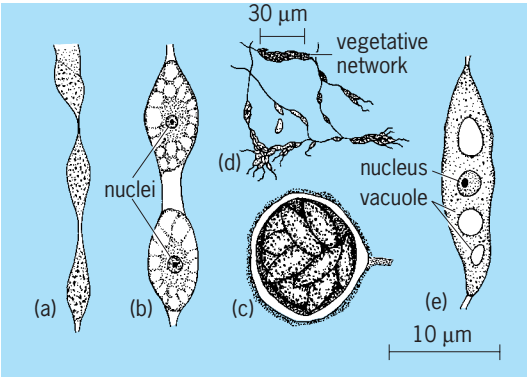

***Labyrinthula zopfi*: (*a*) portion of living network, (*b*) two organisms stained, (*c*) encysted stage. *Labyrinthula macrocystis*: (*d*) vegetative network, (*e*) single organism stained. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)**

forward or backward. The slime filaments, which show an amorphous matrix, may or may not be tubular within a single culture. Individual organisms may move within or along the outside of a slime tube and occasionally may leave the track. A thin pellicle permits changes in diameter, but pseudopodia and phagotrophy have not been seen.

Fission has been described and, in certain species, formation of a sorus by aggregation of individuals. Some produce a pseudosorus by enlargement of one organism; inside, uninucleate pseudospores are produced. The organisms occur on algae and eelgrass, and may cause a disease of the latter. *See* PROTOZOA; RHIZOPODEA; SARCODINA; SARCOMASTIGOPHORA. Richard P. Hall

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; K. L. Pokorny, *J. Protozool.*, 14:697–708, 1967.

Lacquer

A fast-drying, hard, high-gloss surface coating. Lacquers are made by dissolving a cellulose derivative and other modifying materials in a solvent and adding pigment if desired.

The cellulose derivative most commonly used is nitrocellulose, but a cellulose ester such as cellulose acetate or cellulose butyrate, or a cellulose ether such as ethyl cellulose is often used in formulation. Nitrocellulose is not soluble in conventional paint thinners, and so a mixture of solvents with high volatility and extremely fast drying times is used. The mixture usually contains esters (ethylacetate), aromatic hydrocarbons (toluene), and petroleum thinners. Solvent mixtures are formulated carefully in order to produce the combination of solvency and evaporation rate which is required.

Nitrocellulose is extremely hard. Its flexibility is enhanced by addition in the formulation of plasticizers such as vegetable oils (castor or linseed) modified for compatibility with nitrocellulose, or chemical compounds such as dibutyl phthalate or tricresyl phosphate.

Adhesion of lacquers can be improved by the addition of other resins, and these also may reduce the cost by permitting the use of less expensive solvents or, alternatively, by allowing the application of heavier films.

Lacquers dry by evaporation of the solvent. They usually are applied by spray because of their rapid drying properties. The drying time may be extended by altering the formulation to allow application by brushing, but brushing lacquers seldom are used.

Lacquers have been used extensively as fast-drying, weather-resistant finishes for automobiles and as coatings for furniture and other factory-finished items. Their use has diminished, however, as less expensive coating materials with improved properties have been developed. *See* CELLULOSE; SURFACE COATING. C. R. Martinson; C. W. Sisler

Bibliography. *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., 1994.

Lacrimal gland

A tubuloalveolar or acinous skin gland, also known as the tear gland. The lacrimal glands develop from the skin epithelium which folds inward over the developing eye. Two types occur among the vertebrates, the lacrimal proper and the Harderian. These eye glands are first found in the amphibians, associated with the inside of the lower eyelid. In urodele amphibians the eye gland extends along the inner aspect of the lower eyelid. In *Salamandra* it becomes divisible into an anterior Harderian gland associated with lower eyelid structures and a posterior lacrimal gland below the upper eyelid. In frogs and toads only the Harderian gland is present and is associated with the nictitating membrane or third eyelid which develops in relation to the lower lid.

The eye glands drain into the nasal cavity by means of the lacrimal duct. In higher vertebrates, with the exception of snakes, certain lizards, and geckos, Harderian and lacrimal glands are present. In the higher vertebrates the Harderian gland functions in relation to the third eyelid or nictitating membrane, and the lacrimal gland becomes located dorsally above the eyeball near the outer angle of the palpebral fissure. In aquatic mammals, such as seals, whales, and sea cows, the Harderian gland is the more prominent of the two sets of eye glands and it secretes a sebaceous or oily substance. In land mammals the lacrimal gland proper is highly developed as a complex tubuloalveolar structure with several ducts which pour their copious fluid into the outer, upper part of the conjunctival sac or cavity. The tear substance washes across the eyeball to the inner palpebral fissure or commissure. Eventually it passes through two small openings, one on the margin of each lid, into the lacrimal ducts. The latter converge to form the lacrimal sac, from which the nasolacrimal duct leads into the nasal passageway. Tears contain a considerable quantity of the common salt, sodium chloride. *See* EPITHELIUM; EYE (VERTEBRATE); EYE DISORDERS; GLAND. Olin E. Nelsen

Lactam

A cyclic amide that is the nitrogen analog of a lactone. For example, γ-aminobutyric acid readily forms γ-butyrolactam (also known as 2-pyrrolidinone) upon heating, as in reaction (1). The tautomeric enol form

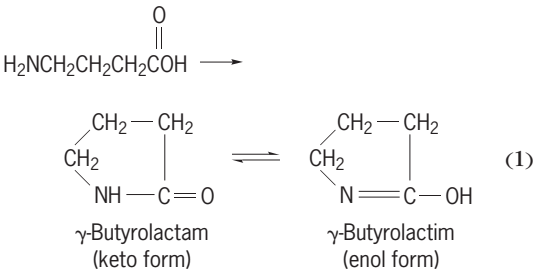

(1)

of a lactam is known as a lactim.

The δ-amino acids similary form δ (six-membered-ring) lactams upon heating, but larger- and smaller-ring lactams must be made by indirect methods, such as addition of ketenes to imines for β-lactams and ring expansion via Beckmann rearrangement of cyclic ketoximes for ϵ- and higher lactams.

Several lactams are of considerable industrial importance. 2-Pyrrolidinone and 1-methyl-2-pyrrolidinone are made by heating γ-butyrolactone with ammonia and methylamine, respectively. They are useful specialty solvents. Vinylation of 2-pyrrolidinone with acetylene gives 1-vinyl-2-pyrrolidinone, which is polymerized to a substance commonly used in aerosol hair sprays, as in reaction (2).

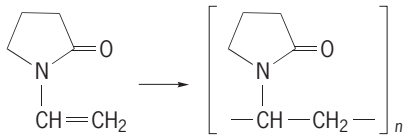

(2)

ϵ-Caprolactam is made commercially by the Beckmann rearrangement of cyclohexanone oxime, and polymerizes when subjected to heat and pressure to form the fibrous polymer nylon-6 [reaction (3)].

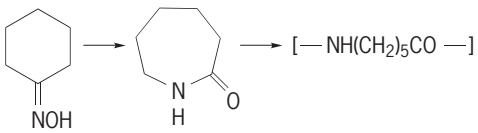

(3)

The β-lactam antibiotics comprise two groups of clinically important therapeutic agents, the penicillins (**1**) and the cephalosporins (**2**). In both cases

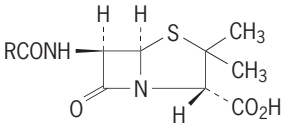

(**1**)

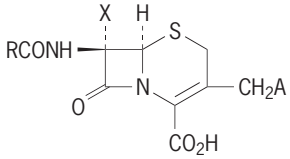

(**2**)

they contain a four-membered or β-lactam ring which has its nitrogen atom and a carbon atom in common with another ring. Such substances are derived commercially from fermentation processes, followed usually by chemical manipulation of the functional groups. *See* AMINO ACIDS; BIOCHEMICAL ENGINEERING; LACTONE; TAUTOMERISM.

Paul E. Fanta

Bibliography. *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., 1994; J. Kroschwitz and M. Winokur, *Chemistry: General, Organic, Biological*, 2d ed., 1990; A. Streitwieser, C. H. Heathcock, and E. M. Kosower, *Introduction to Organic Chemistry*, 4th ed., 1998.

Lactase

An enzyme found in mammals, honeybee larvae, and some plants. It is a β-galactosidase which hydrolyzes lactose to galactose and glucose. In mammals, lactase appears in the intestinal secretion from the intestinal villi, and exerts its effect on lactose in chyme. *See* ENZYME; GLUCOSE.

Daniel N. Lapedes

Lactate

A salt or ester of lactic acid ($CH_3CHOHCOOH$). In lactates, the acidic hydrogen of the carboxyl group has been replaced by a metal or an organic radical. Lactates are optically active, with a chiral center at carbon 2. The two optical isomers of a metal (X) lactate have structures (**1**) and (**2**). Commercial fer-

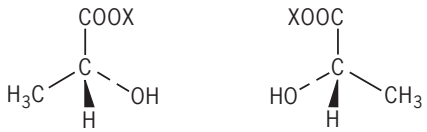

R form (**1**) *S* form (**2**)

mentation produces either the dextrorotatory (*R*) or the levorotatory (*S*) form, depending on the organism involved. *See* OPTICAL ACTIVITY.

The *R* form of lactate occurs in blood and muscle as a product of glycolysis. Lack of sufficient oxygen during strenuous exercise causes enzymatic (lactate dehydrogenase) reduction of pyruvic acid to lactate, which causes tiredness, sore muscles, and even muscle cramps. During renewed oxygen supply (rest) the lactate is reoxidized to pyruvic acid and the fragments enter the Krebs (citric acid) cycle. The plasma membranes of muscle and liver are permeable to pyruvates and lactates, permitting the blood

to transport them to the liver (Cori cycle). Lactates also increase during fasting and in diabetics. *See* BIOLOGICAL OXIDATION; CARBOHYDRATE METABOLISM; CITRIC ACID CYCLE.

Lactates are found in certain foods (sauerkraut), and may be used for flour conditioning and in food emulsification. Alkali-metal salts act as blood coagulants and are used in calcium therapy, while esters are used as plasticizers and as solvents for lacquers. *See* ESTER; SALT (CHEMISTRY). Elbert H. Hadley

Bibliography. J. R. Holum, *Fundamentals of General, Organic, and Biological Chemistry*, 6th ed., 1997; J. Kroschwitz and M. Winokur, *Chemistry: General, Organic, Biological*, 2d ed., 1990; A. Streitweiser, Jr., C. H. Heathcock, and E. M. Kosower, *An Introduction to Organic Chemistry*, 4th ed., 1998; L. Stryer, *Biochemistry*, 4th ed., 1995.

Lactation

The function of the mammary gland providing milk nourishment to the newborn mammal. This process is under the control of the endocrine and nervous systems. It involves transformation of an inactive duct system to a lobuloalveolar glandular structure during pregnancy, cellular production of the components of milk (galactopoiesis), secretion into the ducts, and ejection under the stimulus of milking or suckling.

Galactopoiesis. Formation of the components of milk by the cells composing the lobuloalveolar glandular structure is under the control of a number of hormones. The transformation of the duct structure to a glandular structure during pregnancy initially requires the action of estrogen to induce further duct growth, and of progesterone to induce the development of the lobuloalveolar glandular structure. In addition, somatotropin (STH, growth hormone), corticotropin (ACTH), thyrotropin (TSH), prolactin (PRL) secreted by the anterior lobe of the pituitary (adenohypophysis), and insulin secreted by the pancreas are essential for normal development of the secretory activity of the evolving glandular structure. Cortisone and thyroxine may be substituted for ACTH and TSH. The placenta becomes the source of the estrogen and progesterone and of a mammary-gland hormone which, in humans, is called human placental lactogen (human chorionic somatomammotropin, hCM). Although the gland becomes functional, secretion of milk is minimal until after parturition, when release from the inhibition exerted by the estrogen and progesterone allows secretory activity to rise. Once lactation is initiated after parturition, it is maintained by prolactin with a minimal but necessary contribution from the other pituitary hormones or their target-gland hormones. Milk must be removed by suckling or milking, or the pressure which builds up in the alveoli and ducts will cause retardation of further milk production and involution of the glands.

There is a close correlation between milk secretion and blood levels of prolactin. The initiation of suckling causes a rapid secretion of prolactin and maintenance of the mammary-gland epithelium. The mechanism by which prolactin is thought to effect galactopoiesis involves an interaction of the hormone with a receptor-site protein in the surface of milk secretory cells of the mammary-gland alveoli. The prolactin-receptor site complex is internalized into the cell and, by an undetermined mechanism, enters the nucleus, where it stimulates a genome to transcribe messenger RNA. This nucleic acid then enters the cytoplasm, where the formation it contains is translated into milk protein. The resulting product appears in the lumen of the alveoli as milk (see **illus.**). Between nursings, the blood levels of prolactin are in the range of nonnursing women. Maintenance of the secretory state of the lactating mammary gland requires intermittent stimulation by suckling. Increasing the time between feedings results in less frequent stimulation by prolactin and gradual reduction in milk synthesis and secretion. Secretion of prolactin is held under continuous check by a release-inhibiting hormone produced by the hypothalamus. The suckling stimulus relieves this inhibition and initiates secretion of a releasing hormone.

Lactation makes demands on the maternal regulation of calcium metabolism. Resorption of bone increases in lactating rats and women, and there is a marked increase in the absorption of calcium from the intestine. The elevated need for calcium results in an increased role for parathyroid hormone,

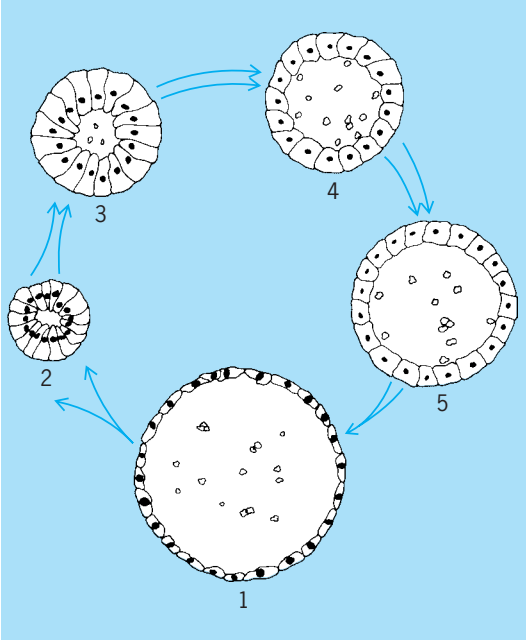

Milk secretory cell activity cycle. (1) The milk-filled alveolus lined with flattened "resting" cells. (2–3) With emptying, as the alveolus contracts and its capillary bed expands, the cells become columnar and resume milk synthesis. (4–5) As milk accumulates in the alveolar lumen, the milk secretory cells regress in height, becoming cuboidal. With further accumulation of milk, the cells reenter their "resting" phase (1). (*After J. M. Nolin and E. M. Bogdanove, Biology of Reproduction, 22:393–416, 1980*)

calcitonin, and vitamin D in the regulation of the absorption and utilization of calcium. In humans a concomitant phenomenon frequently associated with lactation is amenorrhea. Consequently in some societies prolonged nursing is used as a birth control technique.

Milk ejection. Discharge of the milk contained in lactating mammary glands is mediated by a reflex. Suckling stimulates receptors in the nipple that induce impulses in sensory nerves leading into the brain. This stimulus causes the hypothalamus to release oxytocin via its nerves terminating in the neurohypophysis (posterior lobe of the pituitary). The oxytocin is transported by the blood to the basketlike arrangement of myoepithelial cells surrounding the mammary alveoli. The contraction of these cells causes milk to be ejected. Farmers refer to this process as "letdown." Other factors associated with the act of nursing or milking may initiate release of oxytocin. Thus in humans, sexual activity, the approach of the usual nursing time, or preparation of the infant for nursing may become adequate stimuli for release of oxytocin. Analogously, in cows, the sound of buckets, the presence of calves, or washing of the udder may stimulate oxytocin release. Emotional states such as fear or anger which result in secretion of epinephrine may inhibit lactation by inhibiting the action of oxytocin on the myoepithelial basket cells. Nursing mothers should be content, free of stress, and involved primarily with the nursling.

Milk production. The first milk, or colostrum, secreted before parturition, is especially rich in gamma globulin, which contains immune bodies. The digestive tract of the newborn absorbs this protein intact into the blood; this confers a temporary immunity to common diseases. With the regular removal of milk, the amount of milk secreted increases for a period, and then gradually declines until the animal is said to dry up. In dairy animals such as the cow, pregnancy may be induced about the ninetieth day of the lactation period. Lactation and pregnancy with additional growth of the mammary gland occur, so that a second lactation period is initiated in a single year. The total yield to milk increases each year until about the seventh or eighth year, and then declines with old age. *See* ADRENAL GLAND; MAMMARY GLAND; MILK; PITUITARY GLAND; THYROID GLAND. Harry J. Lipner

Bibliography. B. L. Larson, *Lactation*, 1985, reprint 1996; T. B. Mepham, *Physiology of Lactation*, 1987; B. S. Worthington et al., *Nutrition in Pregnancy and Lactation*, 4th ed., 1988.

Lactone

A cyclic, intramolecular ester derived from a hydroxy acid. Simple lactones are designated α, β, γ, δ and so forth according to the position of the hydroxyl group, as shown in (**1**). Five- and six-membered lactones are very readily obtained by cyclization of a hydroxy acid or precursor as shown in reactions (1)–(3). For a γ- or δ-hydroxy acid, the

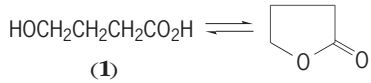

(**1**) (1)

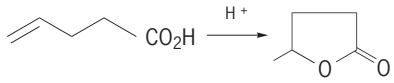

(2)

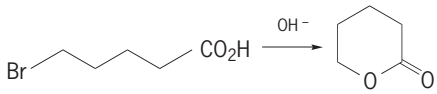

(3)

equilibrium in reaction (1) usually lies far to the right, favoring the lactone. *See* ESTER.

Lactones with three- and four-membered rings are also known. α-Lactones require special preparative methods since, as shown in reaction (4), heating an

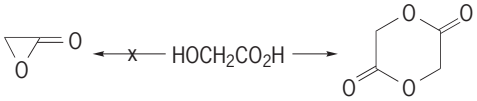

(4)

(**2**)

α-hydroxy acid leads to the dimeric product, known as a lactide, containing two ester groupings in a six-membered ring (**2**). β-Lactones can be prepared from the halo acid or by cycloaddition of a ketene and carbonyl compound [reaction (5)]. The classical method for seven-membered and larger lactones

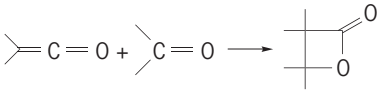

(5)

was oxidation of cyclic ketones or cyclic ethers [reaction (6)], since heating a long-chain hydroxy acid

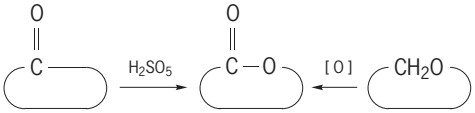

(6)

gives a polymeric product.

Lactones undergo the usual reactions of esters such as nucleophilic attack at the carbonyl group and carbanion reactions of the α-methylene group. The strained α- and β-lactones undergo ring opening at the alkyl-oxygen bond. Although the ring in γ- or δ-lactones is not significantly strained, carbonyl attack is much more rapid than that in open-chain esters (**3**) because the conformation of the ester group in the lactone must adopt the higher-energy Z conformation (**4**).

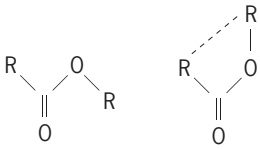

(**3**) (**4**)

Lactones in various forms are found in numerous naturally occuring compounds. Unsaturated γ-lactone rings (**5** and **6**) are present in many components of essential oil. The δ-lactone moiety is present in coumarin (**7**) and other naturally occuring aromatic compounds. Ascorbic acid [vitamin C; (**8**)] is a carbohydrate lactone.

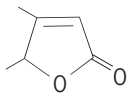

(**5**)

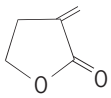

(**6**)

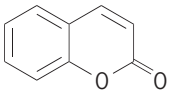

(**7**)

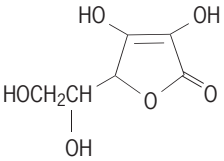

(**8**)

See ASCORBIC ACID; ESSENTIAL OILS.

A major development in the chemistry of naturally occuring compounds has been the isolation from microorganisms of a number of macrocylic lactones with rings containing from 12 to more than 30 atoms. These substances include immunosuppressive agents and antibiotics such as erythromycin. The structures of these macrocylic lactones, or macrolides, are extremely complex, with many chiral centers and often several pendant sugar groups attached as glycosides. Although they represent some of the most complex organic compounds known, several of these macrocylic lactones have been synthesized. The most commonly used approach to systhesis is the assembly of subunits to form a hydroxy acid followed by ring closure by means of a thio ester or other activated carboxyl group. *See* ORGANIC CHEMISTRY; ORGANIC SYNTHESIS.

Paul E. Fanta; James A. Moore

Bibliography. J. W. Apsimon (ed.), *Total Synthesis of Natural Products*, vol. 7, 1988; M. Smith and J. March, *March's Advanced Organic Chemistry*, 5th ed., 2000.

Lactose

Milk sugar or 4-*O*-β-D-galactopyranosyl-D-glucose. This reducing disaccharide is obtained as the β-D-anomer (with the formula shown, where the asterisk indicates a reducing group); the melting point (mp) is 202°C (395.6°F) and the optical activity is $[\alpha]_D^{20}$ + 85.0 → +52.6°. Crystallization at higher temperatures (above 93.5°C or 200.3°F) gives the β-D anomer; mp 252°C (486°F), and $[\alpha]_D^{20}$ + 35 → +55.5°. Lactose is found in the milk of mammals to the extent of approximately 2–8%. It is usually prepared from whey, which is obtained as a by-product in the manufacture of cheese. Upon concentration of the whey, crystalline lactose is deposited. Lactose is not fermentable by ordinary baker's yeast. In the souring of milk, *Lactobacillus acidophilus* and certain other microorganisms bring about lactic acid fermentation by transforming the lactose into lactic acid, $CH_3CHOHCOOH$. *See* CHEESE; MILK; OPTICAL ACTIVITY.

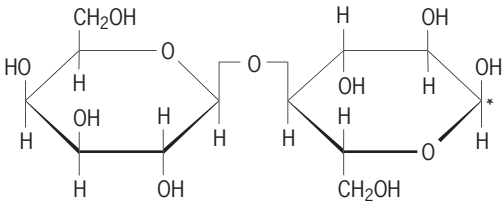

Chemical evidence shows that the glycosidic linkage involves the carbon atom 1 of D-galactose and carbon 4 of D-glucose. Enzymatic studies indicate that the galactosidic linkage has the β configuration. *See* OLIGOSACCHARIDE.

The mammary glands of lactating animals, and their milk, contain an enzyme, lactose synthetase, capable of transferring the D-galactose unit from uridine diphosphate D-galactose to D-glucose, forming lactose according to the scheme: uridine diphosphate D-galactose + D-glucose → lactose + uridine diphosphate.

The lactose synthetase can be resolved into two protein components, A and B, which individually do not exhibit any catalytic activity. Recombination of these fractions, however, restores full lactose synthetase activity. The B fraction was identified as α-lactalbumin. *See* LACTATION; URIDINE DIPHOSPHOGLUCOSE (UDPG).

William Z. Hassid

Bibliography. W. W. Pigman and D. Horton (eds.), *The Carbohydrates*, 2d ed., vols. 1–2, 1970–1980.

Laevicaudata

An order of fresh-water branchiopod crustaceans formerly included in the order Conchostraca. The body is up to about 7 mm (0.3 in.) in length; females have 12 trunk segments, males have 10. The trunk terminates in a feebly developed telson that lacks claws and is covered ventrally by an opercular lamella. The trunk and its appendages are covered by a bivalved carapace that has a true hinge dorsally and that usually lacks lines of growth. The head is free, articulated with the trunk, and can be swung forward. It is covered by a headshield and has paired, sessile eyes. It bears large antennae that are used for swimming. The mandibles, while of the rolling, crushing type, have narrow masticatory surfaces and stout teeth.

The 12 pairs of trunk limbs of the female are foliaceous. The first, and sometimes the second, of the 10 pairs of the male is modified for clasping the female. Clumps of small drought-resistant eggs are carried by the female and shed at the next molt. They

hatch as a late nauplius of distinctive type. The Laevicaudata are inhabitants of temporary water bodies and are almost worldwide in distribution. Although they can swim, they are usually associated with the bottom. *See* BRANCHIOPODA. Geoffrey Fryer

Lagomorpha

An order of mammals composed of two living families: Leporidae (rabbits and hares) and Ochotonidae (pikas). Rabbits and hares have short tails and long ears, whereas pikas lack an external tail and are short-eared. Rabbits and hares occupy a wide variety of habitats and are nearly worldwide in distribution. Pikas are found only in the mountains of western North America, in Siberia, and in the Himalayan region.

Description. Lagomorphs were long considered to be members of the order Rodentia because of their chisel-like incisors, lack of canine teeth, and the presence of a space between the incisors and molariform teeth (known as a diastema). At birth, lagomorphs have three pairs of upper incisors, but the outer on each side is soon lost. In adults, the second pair of incisors is small, is nearly circular, lacks cutting edges, and is set directly behind (not beside) the first functional pair. Each of the anterior upper incisors possesses a longitudinal groove on its anterior face. As in rodents, the incisors grow throughout life, but unlike rodent incisors, they are covered both anteriorly and posteriorly by enamel. In rodents, the enamel is restricted to the anterior face. The distance between the tooth rows of the lower jaws is less than that between the tooth rows of the upper jaws, so only one pair of upper and lower tooth rows is opposable at the same time. The motion of the jaws in mastication is lateral, unlike the rotary jaw movement of rodents. The palate is short and does not extend to the end of the tooth row. It is generally accepted that the resemblances between lagomorphs and rodents are results of convergence rather than of close taxonomic relationship. There is some indication that lagomorphs are more closely related to hooved animals than to rodents. *See* DENTITION; RODENTIA; TOOTH.

The two main lower leg bones, the tibia and fibula, are fused for more than half their length, and the fibula articulates with the calcaneum in the ankle as in artiodactyls. The forelimbs have five digits, while the hindlimbs may have four or five. *See* ARTIODACTYLA.

The diet of lagomorphs is restricted to vegetation. Two types of fecal pellets are expelled: brown pellets and greenish pellets. The reingestion of partially digested vegetation from the greenish pellets is common; this practice is known as coprophagy. It allows the animals to spend relatively little time exposed to predators while in the field actually feeding. They consume green vegetation rapidly and then make optimum use of it in the safety of their brushpile or burrow. This process is somewhat analogous to cud chewing in ruminant animals. These pellets have a high protein content and contain large amounts of B vitamins produced by intestinal bacteria. *See* DIGESTIVE SYSTEM.

Leporidae. This family of 11 living genera and 54 species has a natural range covering most of the world's major landmasses, with the notable exceptions of southern South America, the West Indies, Madagascar, and most islands to the southeast of Asia. Leporids, however, are among the most widely introduced mammals, having been brought by humans to southern South America, Australia, New Zealand, Java, and various oceanic islands. Members of this family have long hindlimbs, oversized feet, furry tails, and long ears. Each limb has five clawed digits, but the first digit is very small. The soles of the feet are covered with hair. A sensory pad, usually hidden by hair-covered folds of skin, is located at the entrance of each nostril, and a Y-shaped naked groove extends from the upper lip to and around the nose. The upper lip is divided. Female leporids are usually larger than males.

The terms "rabbit" and "hare" are often mistakenly interchanged. True rabbits (*Pentalagus, Bunolagus, Pronolagus, Nesolagus, Romerolagus, Brachylagus, Sylvilagus, Oryctolagus, Poelagus*, and *Caprolagus*) prefer habitats with plenty of cover. Young rabbits are born in a well-defined nest. At birth they are altricial; that is, they are blind and naked, have undeveloped ears, and are helpless. Hares (*Lepus*) are bigger than rabbits, have longer hindlegs developed for leaping, and have longer, broader ears. They favor wide-open spaces. Hares, often called jackrabbits in North America, have longer gestation periods than rabbits. The female hare rarely constructs a nest for her young which, at birth, are precocial; that is, they are well developed in that they are furred, have their eyes and ears open, and are able to move about on their own shortly after birth.

Leporids are prolific breeders. Reproductive efficiency is increased by a phenomenon known as induced ovulation. Female leporids normally ovulate only after copulation has occurred, a system that ensures optimum fertilization of ova. *See* REPRODUCTIVE BEHAVIOR; REPRODUCTIVE SYSTEM.

Rabbits and hares are crepuscular, active mainly near dusk and dawn. They are not physically capable of defending themselves against predators, but rely on a keen sense of hearing and smell, the use of burrows, or speed and agility to escape danger. At high speeds, lagomorph limbs function as digitigrade (walking on the digits) appendages, but when moving slowly the entire foot makes contact with the ground (plantigrade).

The Old World, or domestic, rabbit (*Oryctolagus*) has been widely introduced throughout the world. Its spectacular rate of reproduction, especially in areas where no natural predators existed, has proven to be disastrous to other fauna. Nesting areas for petrels and albatrosses on some Pacific islands have been virtually destroyed by the removal of vegetation

Pika (*Ochotona princeps*). (© *California Academy of Sciences*)

by these rabbits and resulting erosion. These rabbits were introduced into Australia by British settlers in 1859 and to New Zealand a few years later. They reproduced until they numbered in the hundreds of millions, causing an ecological disaster in the southern half of Australia. Through their competition and habitat destruction, they have caused the extinction of several native marsupial species and have seriously harmed the sheep-raising industry. In the early 1950s, scientists released myxomatosis, a rabbit-killing virus which proved to be partially successful in controlling the rabbit population but gradually became less effective. In 1995, a calicivirus known as rabbit hemorrhagic disease (RHD) began to be released. By 1997, the wild rabbit population had dropped by 95% in some regions, and native fauna and flora were already staging a comeback. The expectation is that, after the calicivirus kills most of the rabbits, it will remain in the reduced population and act as a long-term regulator of the rabbit population. *See* INFECTIOUS MYXOMATOSIS.

Ochotonidae. Pikas (*Ochotona*), with their stocky bodies and short legs, resemble guinea pigs (see **illustration**). In North America, other common names include cony, rock cony, rock rabbit, mouse hare, whistling hare, little chief hare, and piping hare. While most of the 26 species that occur worldwide inhabit talus, or piles of broken rock, in mountainous regions, several species in Russia, Mongolia, Siberia, and northern China inhabit open plains, deserts, and steppes, and excavate burrows for shelter. Pikas are primarily active during the day, in all seasons. They feed on a variety of vegetable matter. Individuals in most populations spend many hours every day harvesting grasses and herbs and curing them in the sun in exposed locations. These hay piles or haystacks are then stored to be used as food during the winter. As in lagomorphs, coprophagy allows pikas to obtain certain vitamins that would otherwise be lost. Although some species are gregarious and live in large colonies, adult males and females of the two North American species (*O. collaris* and *O. princeps*) are territorial and live separately for most of the year. Most pikas breed twice a year and produce 1–5 young after a gestation of approximately 30 days. Distinctive whistling vocalizations are used during courtship. Rock-dwelling species may live up to 7 years, whereas ground-dwelling species usually live only 1–2 years. Donald W. Linzey

Bibliography. G. A. Feldhamer, B. C. Thompson, and J. A. Chapman, *Wild Mammals of North America: Biology, Management, and Conservation*, 2d ed., Johns Hopkins University Press, 2003; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Lagrange's equations

Equations of motion of a mechanical system for which a classical (nonquantum-mechanical) description is suitable, and which relate the kinetic energy of the system to the generalized coordinates, the generalized forces, and the time. If the configuration of the system is specified by giving the values of f independent quantities $q_1, \ldots, q_f$, there are f such equations of motion. In their usual form, these equations are equivalent to Newton's second law of motion and are differential equations of the second order for the q's as functions of the time t.

Derivation. Let the system consist of N particles. The masses of the particles are m_ρ and their cartesian coordinates are $x_{\rho i}$ ($\rho = 1, 2, \ldots, N$; $i = 1, 2, 3$). These cartesian coordinates are expressible as functions of $f (\leqq 3N)$ generalized coordinates $q_1, \ldots, q_f$ between which there are no constraints. The time does not appear here if the constraints are fixed (as will be assumed), as in Eq. (1). Then, denoting time

$$x_{\rho i} = f_{\rho i}(q_1, \ldots, q_f) \tag{1}$$

differentiation by a dot, Eq. (2) holds. The cartesian

$$\dot{x}_{\rho i} = \sum_{j=1}^{f} \frac{\partial x_{\rho i}}{\partial q_j} \dot{q}_j \tag{2}$$

velocity components are linear functions of the generalized velocities q_j.

Let δq_j represent a small displacement of the system. It is automatically consistent with the constraints. In this displacement, the forces of constraint do no work; their only action is to prevent motion contrary to the constraint, so they have no components in the direction of a displacement consistent with the constraints. If W is the work done during the displacement, it is done entirely by the externally applied forces, as in Eq. (3).

$$W = \sum_{\rho, i} F_{\rho i} \delta x_{\rho i} = \sum_{\rho, i} m_\rho \ddot{x}_{\rho i} \delta x_{\rho i} \tag{3}$$

Each term in the sum of Eq. (3) may receive contributions from forces of constraint, but they cancel when the summation is made. From Eq. (1), Eqs. (4) and (5) are obtained.

$$\delta x_{\rho i} = \sum_{j} \frac{\partial x_{\rho i}}{\partial q_j} \delta q_j \tag{4}$$

$$\frac{\partial \dot{x}_{\rho i}}{\partial \dot{q}_j} = \frac{\partial x_{\rho j}}{\partial q_j} \tag{5}$$

It is readily verified that Eq. (6) is valid. The quan-

$$W = \sum_j \left[\frac{d}{dt}\frac{\partial}{\partial \dot{q}_j}\left(\frac{1}{2}m_\rho \dot{x}_{\rho i}{}^2\right) - \frac{\partial}{\partial q_j}\left(\frac{1}{2}m_\rho \dot{x}_{\rho i}{}^2\right)\right]\delta q_j \tag{6}$$

tity in the brackets is thus, for each j, the noncartesian analog of the carterian $m_\rho \ddot{x}_{\rho i}$.

The only quantity which appears differentiated in Eq. (6) is the total kinetic energy T of the system. This is easily calculated in generalized coordinates because the connection between the cartesian velocities and the generalized velocities is linear and homogeneous. Usually, the kinetic energy can be written by inspection without using Eq. (2) explicitly. Thus Eq. (7) holds.

$$W = \sum_j \left(\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j}\right)\delta q_j \tag{7}$$

Transforming the right-hand side of Newton's equation is simpler, as shown in Eqs. (8) and (9),

$$W = \sum_{\rho,i} F_{\rho i}\delta x_{\rho i} = \sum_j Q_i \delta q_j \tag{8}$$

$$Q_j = \sum_{\rho,i} F_{\rho i}\frac{\partial x_{\rho i}}{\partial q_j} \tag{9}$$

where Eq. (9) is the jth component of the generalized force. By the preceding argument, Q_j depends only on the externally applied forces, the forces of constraint necessarily canceling in the summation.

The displacement δq_j was entirely arbitrary. Thus, it follows from equating expressions (7) and (8) for W that Eq. (10) is valid. Equation (10) sums up

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = Q_j \tag{10}$$

Lagrange's equations of motion. They are valid also when moving constraints are present. *See* CONSTRAINT.

Examples. Use of this form of Lagrange's equations is shown in the two following examples.

Particle in central force field. Here, the force acting on a particle acts always through a fixed point. Choose this point as origin of a spherical coordinate system with coordinates r, θ, ϕ (**Fig. 1**). Only radial displacements involve work, so that only Q_r differs from zero, or Eq. (11) holds.

$$Q_\theta = Q_\phi = 0 \tag{11}$$

The kinetic energy is given by inspection of Fig. 1 and is expressed by Eq. (12).

$$T = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2) \tag{12}$$

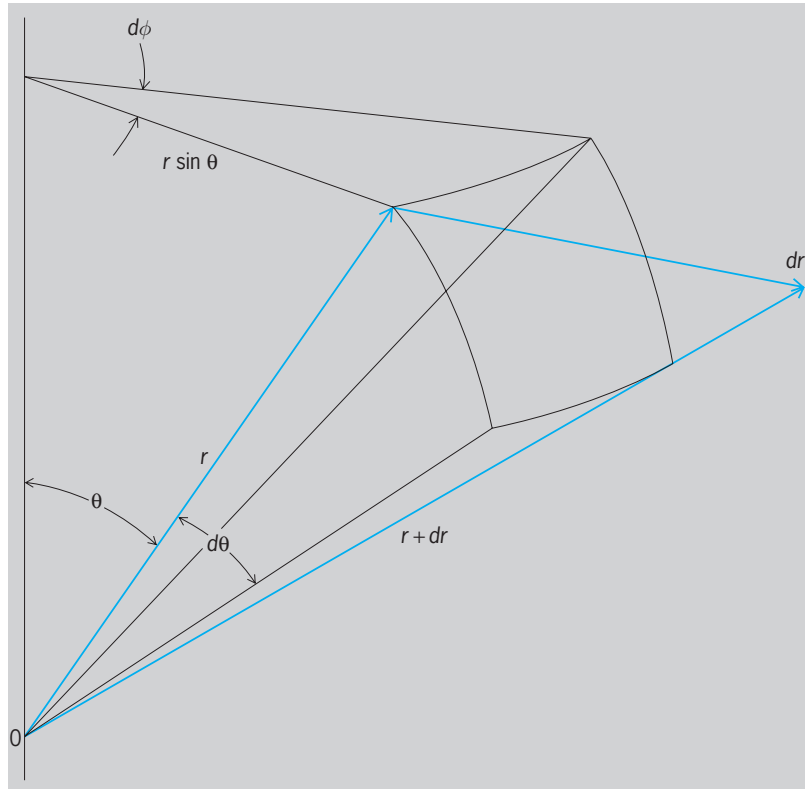

Fig. 1. The vector *dr* is decomposed into three orthogonal vectors of magnitudes *r* sin θ *d*ϕ, *rd*θ, and *dr*, respectively, *r*, θ, and ϕ being spherical coordinates of the terminus of the vector *r*.

Lagrange's equations are Eqs. (13). These may

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{r}} - \frac{\partial T}{\partial r} = m(r - r\dot{\theta}^2 - r\sin^2\theta\,\dot{\phi}^2) = Q_r$$

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{\theta}} - \frac{\partial T}{\partial \theta} = m\left[\frac{d}{dt}(r^2\dot{\theta}) - r^2\sin\theta\cos\theta\,\dot{\phi}^2\right] = 0 \tag{13}$$

be compared with the cartesian equations (14), in

$$\begin{aligned} m\ddot{x} &= \frac{x}{r}F(x,y,z) \\ m\ddot{y} &= \frac{y}{r}F(x,y,z) \\ m\ddot{z} &= \frac{z}{r}F(x,y,z) \end{aligned} \tag{14}$$

which the force function F appears in all three of the equations of motion, while two of the three Lagrange equations are independent of the detailed nature of the force.

Two particles, fixed separation. In this system, there is one constraint, the particles being a constant distance d apart, and so there are five degrees of freedom instead of six. Choose as generalized coordinates the cartesian coordinates X, Y, Z of the center of mass and the polar angles θ, ϕ of the line joining the two particles, as in **Fig. 2**. Then because the kinetic energy is the sum of the kinetic energy of the center of mass and the kinetic energy relative to

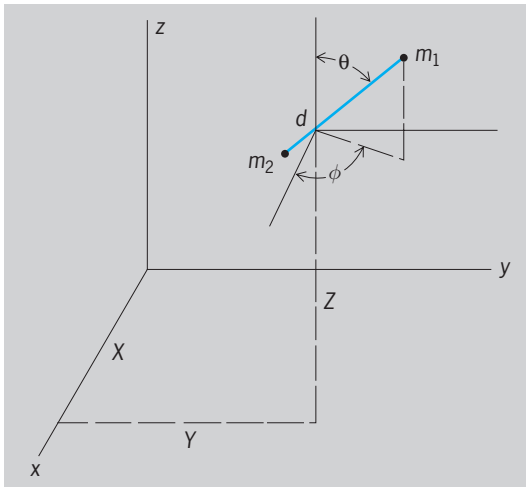

Fig. 2. ***X*, *Y*, and *Z* are cartesian coordinates of the center of mass of dumbbell m_1m_2; θ and ϕ are the polar angles of the dumbbell axis in a set of axes parallel to fixed set of axes and with origin at center of mass.**

the center of mass, Eq. (15) holds. The equations of

$$T = \frac{m_1 + m_2}{2}(\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2) + \frac{1}{2}\left(\frac{m_1 m_2}{m_1 + m_2}\right) d^2(\dot{\theta}^2 + \sin^2\theta\,\dot{\phi}^2) \quad (15)$$

motion may be written down immediately, the generalized forces being evaluated from Eq. (9). There is one equation for each degree of freedom, and the constraint is automatically satisfied.

Conservative systems. In many problems, the forces Q_j are derivable from a potential. Then Eq. (16), and thus Eq. (17), is valid. When this is

$$W = \sum_{j=1}^{f} Q_j dq_j = -dV \quad (16)$$

$$Q_j = -\frac{\partial V}{\partial q_j} \qquad \frac{\partial V}{\partial \dot{q}_j} = 0 \quad (17)$$

so, it is convenient to define a function, called the Lagrangian, by Eq. (18). Then the equations of

$$L(q, \dot{q}t) = T(q, \dot{q}, t) - V(q, t) \quad (18)$$

motion become simply Eq. (19), which is the

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0 \quad (19)$$

most commonly encountered form of Lagrange's equations.

Example of use of L. A simple vibration problem illustrates the use of generalized coordinates and shows the ease of application of Lagrange's equations when T and V are relatively easy to obtain. Consider the one-dimensional system of two equal masses m connected by a spring of restoring constant k. Let the displacements of the masses from their equilibrium positions be x_1 and x_2, respectively.

Now introduce new coordinates, $q_1 = x_2 - x_1$ and $q_2 = x_2 + x_1$. Then Eqs. (20) hold.

$$V = \frac{k{q_1}^2}{2} \qquad T = \frac{m(\dot{q}_1{}^2 + \dot{q}_2{}^2)}{4} \quad (20)$$

Using $L = T - V$, Lagrange's equations for the two variables are written as Eqs. (21).

$$^1/_2 m\dot{q}_1 - kq_1 = 0 \qquad ^1/_2 m\dot{q}_2 = 0 \text{ or } \dot{q}_2 = 0 \quad (21)$$

The solution to the first of these differential equations is $q_1 = A \sin 2k/m\ t$ (setting $q_1 = 0$ at $t = 0$), which is a simple harmonic vibration of the two masses. The second equation merely states that the acceleration of the center of mass of the system is zero.

Nonconservative systems. For (19) is also correct even if the system is not conservative, provided one can write Eq. (22), as in the case of a charged particle in an electromagnetic field. Here V is defined in Eq. (23), where ϕ and $\mathbf{A}$ are the scalar and vector

$$Q_j = \frac{d}{dt}\frac{\partial V}{\partial \dot{q}_j} - \frac{\partial V}{\partial q_j} \quad (22)$$

$$V = e\left(\phi - \frac{\mathbf{v}\cdot\mathbf{A}}{c}\right) \quad (23)$$

potentials of the field, respectively, e is the charge of the particle, $\mathbf{v}$ is the particle's velocity, and c is the velocity of light.

Cyclic coordinate. If L does not depend explicitly on a particular coordinate, say q_k, then Eqs. (24) hold.

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} = 0 \qquad \frac{\partial L}{\partial \dot{q}_k} = \text{constant} \quad (24)$$

Here q_k is an ignorable or cyclic coordinate, and $\partial L/\partial q_k$ is an integral of motion. In the example of the central force field given before, ϕ may be ignored.

Total energy. By use of Eq. (19), Eq. (25) is obtained.

$$\frac{d}{dt}\left(\sum_{j=1}^{f}\frac{\partial L}{\partial \dot{q}_j}\dot{q}_j - L\right) = -\frac{\partial L}{\partial t} \quad (25)$$

Hence, if L does not depend explicitly on the time, expression (26) is an integral of motion. Usually, it

$$\sum_{j=1}^{f}\left(\frac{\partial L}{\partial q_j}q_j - L\right) \quad (26)$$

is the total energy of the system. *See* HAMILTON'S EQUATIONS OF MOTION.

Conjugate momentum. The quantity in Eq. (27) is de-

$$p_k = \frac{\partial L}{\partial \dot{q}_k} \quad (27)$$

fined to be the momentum conjugate to the coordinate q_k. (If a Lagrangian does not exist because of the presence of dissipative forces, the conjugate momentum is sometimes defined as $\partial T/\partial \dot{q}_k$.) This momentum is not necessarily a linear momentum as defined in Newtonian mechanics, since its character depends both on the system and on the nature of the coordinate q_k. If q_k is an ignorable coordinate, its conjugate momentum is a constant of motion.

Kinetic momentum. For a charged particle in an electromagnetic field, the Lagrangian may be written as Eq. (28), and the momentum conjugate to the carte-

$$L = {}^1/_2 m\mathbf{v}^2 - e\phi + \frac{e}{c}\mathbf{v}\cdot\mathbf{A} \tag{28}$$

sian coordinate x is given by Eq. (29). The quantity

$$p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x} + \frac{e}{c}A_x \tag{29}$$

$m\dot{x}$ is called the kinetic momentum. The kinetic momentum is related to p_x much as the kinetic energy is related to the total energy. If Φ and A are independent of x, it is the momentum p_x and not the kinetic momentum $m\dot{x}$ which is a constant of motion.

Relativistic systems. The equations of motion of a relativistic particle may be written in Lagrangian form. The simplest way to do this is to replace the kinetic energy T by another function τ of the mass and velocity so as to get the desired form, namely, Eq. (30). Here m_0 is the rest mass of the particle.

$$\frac{d}{dt}\frac{\partial \tau}{\partial \dot{x}_k} - \frac{\partial V}{\partial x_k} = \frac{d}{dt}\left(\frac{m_0\dot{x}_k}{\sqrt{1-v^2/c^2}}\right) - \frac{\partial V}{\partial x_k} = 0 \tag{30}$$

This is accomplished by setting τ as in Eq. (31),

$$\tau = \left(1 - \sqrt{1-v^2/c^2}\right) m_0 c^2 \tag{31}$$

which reduces to T in the limit $\mathbf{v}/c \to 0$. The equations of motion in this form are still valid in only one reference frame because the time and the coordinates are treated on different bases. *See* RELATIVISTIC MECHANICS; RELATIVITY. Philip M. Stehle

Bibliography. R. Baierlein, *Newtonian Dynamics*, 1983; V. D. Barger and M. G. Olsson, *Classical Mechanics: A Modern Perspective*, 2d ed., 1995; H. C. Corben and P. Stehle, *Classical Mechanics*, 2d ed., 1960, reprint 1994; H. Goldstein, C. P. Poale, and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; L. D. Landau and E. M. Lifshitz, *Mechanics*, 3d ed., 1976; K. R. Symon, *Mechanics*, 3d ed., 1971.

Lagrangian function

A function of the generalized coordinates and velocities of a dynamical system from which the equations of motion in Lagrange's form can be derived. The lagrangian function is denoted by $L(q_1, \ldots, q_f, \dot{q}_1, \ldots, \dot{q}_f, t)$. *See* LAGRANGE'S EQUATIONS.

For systems in which the forcesare derivable from a potential energy V, if the kinetic energy is T, Eq. (1) holds.

$$L = T - V \tag{1}$$

If the system is continuous rather than discrete, the lagrangian function L is the integral of a lagrangian density $\mathscr{L}$, as in Eq. (2), where $\eta(x_1, x_2,$

$$L = \int \mathscr{L}(\eta, \text{grad}\,\eta, x_1, x_2, x_3, t)\, dx_1\, dx_2\, dx_3 \tag{2}$$

$x_3)$ describes the displacement of the medium at the point (x_1, x_2, x_3). The equations of motion are in this case written as Eq. (3).

$$\frac{\partial}{\partial t}\frac{\partial \mathscr{L}}{\partial(\partial\eta/\partial t)} + \sum_{j=1}^{3}\frac{\partial}{\partial x_j}\frac{\partial \mathscr{L}}{\partial(\partial\eta/\partial x_j)} - \frac{\partial \mathscr{L}}{\partial \eta} = 0 \tag{3}$$

This formulation of Lagrange's equations applies to the motion of a gas containing sound waves, to a vibrating jelly, or to any medium where discrete masses are replaced by a continuum. Philip M. Stehle

Lake

An inland body of standing water occupying a depression on the Earth's surface. The surface of a lake is exposed to the atmosphere; in addition, subterranean lakes occur. The term lake includes expanded parts of rivers, reservoirs behind dams, and lake basins intermittently or previously covered with water. Small lakes are usually called ponds, tarns (in mountains), and less frequently pools or meres. Lakes are variable in size, with most lakes being less than 65 ft (20 m) in depth. Collectively, lakes cover approximately 9.6×10^5 mi^2 (2.5×10^6 km^2) of the Earth's surface (see **table**). The largest lake basins together hold about 40% of the Earth's freshwater supplies; about half this volume is contained in a single lake basin, Lake Baikal. *See* HYDROSPHERE.

Most lake basins are elliptic sinusoids in cross section, with mean depths about half their maximum depths. Only a few are considered to be extremely deep, with mean depths that exceed 1300 ft (400 m). Although these generally occur in glaciated mountain regions, the Great Lakes of the United States are included in this category. Lakes Baikal and Tanganyika have maximum depths in excess of 3300 ft (1000 m) and mean water depths of 1650 ft (500 m).

Lakes generally cover relatively small areas of the Earth's surface. The largest of the lakes in terms of area is the Caspian Sea, a salt lake, and the largest fresh-water lake is Lake Superior in North America. Together, the five Great Lakes of North America represent the largest continuous mass of fresh water on the Earth's surface, and account for about 95% of the fresh surface water resources in North America.

Formation. Lake-filled depressions form by a great variety of natural processes, including tectonic activity, volcanic activity, landsliding, glacial activity, bedrock solution, river activity, wind activity, shoreline processes, and biological activity.

Dimensions of some major lakes*

Lake	Area, mi²*	Volume (approx.), 1000 acre-ft	Shoreline, mi	Depth Av. ft	Depth Max. ft
Caspian Sea	168,500	71,300	3,730	675	3,080
Superior	32,200	9,700	1,860	475	1,000
Victoria	26,200	2,180	2,130		
Aral Sea	26,233†	775	—		
Huron	23,010	3,720	1,680		
Michigan	22,400	4,660	—		870
Baikal	13,300†	18,700	—	2,300	5,000
Tanganyika	12,700	8,100	—		4,700
Great Bear	11,490†	—	1,300		
Great Slave	11,170†	—	1,365		
Nyasa	11,000	6,800		900	2,310
Erie	9,940	436			
Winnipeg	9,390†		1,180		
Ontario	7,540	1,390			
Balkash	7,115				
Ladoga	7,000	745			
Chad	6,500†				
Maracaibo	4,000†				
Eyre	3,700†				
Onega	3,764	264			
Rudolf	3,475†				
Nicaragua	3,089	87			
Athabaska	3,085				
Titicaca	3,200	575			
Reindeer	2,445				

*1 mi^2 = 2.6 km^2. 1000 acre-ft = 1.2×10^6 m^3. 1 mi = 1.6 km. 1 ft = 0.3 m.
†Area fluctuates.

Lake basins develop in areas where the Earth's surface is affected by internal geological processes. These processes result in the vertical displacement of the crust and the subsequent development of lakes in grabens or broadly warped depressions. Examples of such tectonic lakes are Lake Baikal and Lake Tanganyika. Some lake basins are tectonically uplifted above sea level, thus isolating the sea floor. Lakes resulting from such uplift include the Caspian and Aral seas. *See* GRABEN; PLATE TECTONICS.

Volcanic lakes form in a variety of ways. Explosive volcanic activity resulting in the formation of either a maar or a caldera is frequently followed by lake formation. In the case of maars, the lake is formed either by lava coming into contact with ground water or from volcanic degassing. In the case of calderas, the lake forms as a result of the seepage of ground water into the subsided crater of the volcano following complete or partial emptying of the magma chamber. Some volcanic lakes form in essentially tectonic depressions formed as a result of accompanying deformation of the ground surface during eruptions. The third principal means by which volcanic lakes form is by modification of drainage patterns due to molten lava blocking stream flow and impounding river water behind the flows. *See* CALDERA; LAVA; MAGMA; VOLCANO.

Temporary or permanent lakes frequently develop in association with landslide occurrences. Landslide events commonly block streams, causing water to accumulate behind the debris. These lakes are commonly ephemeral, surviving only a short period of time before the stream adjusts to the sudden addition of a mass of sediment and erodes through the debris. In other cases, the lake may remain for some time. Substantially smaller and extremely ephemeral lakes may also develop in depressions in the headwall region of the landslide. These lakes commonly drain after a short period of time. *See* LANDSLIDE.

Perhaps the most common cause of lakes is glacial and periglacial activity which occurs widely in the high-latitude areas of the Northern Hemisphere as well as the high altitudes of both hemispheres. Lakes are associated with both the erosional and depositional processes of glaciation. Lakes fill depressions scoured by alpine glaciers in their source areas to form cirque lakes, or tarns. Lakes also commonly fill depressions scoured by alpine glaciers at progressively lower elevations to form a chain of paternoster lakes. Continental glaciers scour large areas of bedrock to form depressions such as those filled by the Great Lakes of North America. Glacial lakes are commonly impounded behind morainal material left both in the terminal position of glaciers and at temporary stillstand positions of a retreating glacier. Many smaller lakes, called kettle lakes, occupy depressions resulting from the melting of blocks of glacial ice in morainal debris. In addition, small lakes may result from the impoundment of water behind lateral moraine features. Thaw lakes commonly form in permafrost-dominated areas, where ground water seeps into and melts permafrost. *See* GLACIATED TERRAIN; MORAINE; PERMAFROST.

A wide variety of other, less catastrophic processes operate to create lakes. These include lakes that form as a result of ground-water and surface-water filling

of solution features in limestone terranes, including dolines and poljes, which are features of karst topography. The erosional and depositional processes of rivers result in a diversity of lake forms along river channels. These include lakes at the base of waterfalls, plunge pools, which may be either active or abandoned. They also include oxbow lakes, which occupy abandoned meander loops of river channels. Coastal-zone lakes commonly form where a coastline embayment becomes blocked by sand transported by longshore drift. In addition, lakes may become entrapped behind coastal dune systems. *See* KARST TOPOGRAPHY; RIVER.

Eolian, or wind-related, processes also create depressions that may be permanently or intermittently filled with water. Interdune corridors may become areas of impoundment. Deflation of desert and semidesert surfaces creates depressions, which often become filled with water to form playas or salinas. *See* EOLIAN LANDFORMS; PLAYA; SALINE EVAPORITES.

Lakes are also created by biological activity. Plant growth due to nutrient enrichment associated with pollution as well as siltation in rivers often leads to lake formation along stream channels. Similarly, beaver activity along streams blocks free flow, and lakes are created. Human activity also leads to the creation of lakes behind a variety of dam structures. *See* EUTROPHICATION; FRESHWATER ECOSYSTEM.

Water characteristics. Lakes are typically subdivided into saline and nonsaline types, with the former including a spectrum of salt contents. Salinity, especially in smaller lakes, is also often variable, fluctuating with seasonal climatic patterns. Typically the salinity of lakes reflects the climatic environment in which they occur. Lakes in arid and semiarid environments, with high potential evaporation, are typically more saline than lakes in humid environments, with low potential evaporations. Examples are the Dead Sea, which has a salinity approximately seven times that of seawater, and the Great Salt Lake in the Western United States, which has a salinity approximately four times that of seawater. Other salt lakes have salinities that are less than that of seawater but are clearly salty to the taste. Freshwater lakes also vary considerably in their salt content, depending on their geologic environment.

Thermal stratification of lakes is a major physical characteristic, reflecting the prevailing climate, lake morphometry, and chemistry; it has profound influences on other physical processes, such as freezing behavior, as well as on the biology of lakes. Most of the world's lakes that are located in temperate latitudes are dimictic; that is, they circulate freely twice a year, with thermal stratification in the summer and inverse stratification in the winter. In the summer the waters of such lakes are warmed from the surface to depth more rapidly than lake-water mixing occurs. In the winter the lake waters cool as heat loss exceeds heat input and inverse stratification occurs. Between these two periods of relative thermal stability, most lakes of moderate depth (>33 ft or 10 m) mix their waters and undergo fall and spring turnover. *See* MEROMICTIC LAKE.

Environmental considerations. Lakes created behind constructions are common features of the Earth's landscape. Such lakes include Lake Mead and Lake Powell on the Colorado River, Lake Roosevelt on the Columbia River, Lake Nasser on the Nile, and Lake Tsimlyanskaya on the Don, which flows into the Black Sea. While major dam construction has largely ceased in the western world, major dam building continues in the developing countries such as Losoto. These lakes, as well as many natural lakes, are significant for the storage and supply of water for irrigation and domestic use, control of floods, generation of electric power, use as commercial waterways, and recreation. However, these diverse uses frequently result in adverse impacts on water quality. Common environmental problems include increases in salinity, siltation of the reservoir, and excessive algal production due to the spillage and inflow of a variety of nutrients. River impoundment also has considerable adverse impact downstream from the dam with respect to flow regimes as well as changes in water temperature. These changes have profound adverse effects on both biota and plant growth. *See* DAM; STREAM TRANSPORT AND DEPOSITION.

While the environmental health of many of the larger lakes has improved considerably with increased global environmental awareness, many are still substantially polluted. For example, there are in excess of 100 toxic substances that can be identified in the waters and sediments of the Great Lakes of North America. While these lakes continue to support a great diversity of life forms, the waters and sediments are contaminated with varying concentrations of chemical elements associated with waste from paper and woodpulp industry as well as from the treatment of sewage. Polychlorinated biphenyls, leaked from electrical capacitors and transformers, continue to be released from sediments on the lake floors and constitute major concerns for human health. Fish production continues to decline in the Great Lakes. *See* WATER POLLUTION. John C. Dixon

Bibliography. A. Lerman (ed.), *Lakes: Chemistry, Geology and Physics*, 2d ed., 1995; J. McManus and R. W. Duck (eds.), *Geomorphology and Sedimentology of Lakes and Reservoirs*, 1993; M. R. Rosen (ed.), *Paleoclimate and Basin Evolution of Playa Systems*, Geol. Soc. Amer. Spec. Pap. 289, 1994; R. G. Wetzel, *Limnology*, 3d ed., 2001.

Lamiales

An order of flowering plants (Magnoliophyta, or angiosperms), in the subclass Asteridae (Eudicotyledons). The order consists of some 22 families with approximately 1100 genera and over 21,000 species. Seven families (Acanthaceae, Gesneriaceae, Lamiaceae, Orobanchaceae, Plantaginaceae, Scrophulariaceae, Verbenaceae) have more than 1000 species. Orobanchaceae, Plantaginaceae, and Scrophulariaceae represent monophyletic groups segregated from the polyphyletic former

Scrophulariaceae *sensu lato*. Members of the Lamiales are distributed worldwide with some families being predominantly tropical (such as Acanthaceae, Bignoniaceae, Gesneriaceae, Verbenaceae) and some being predominantly temperate (such as Oleaceae, Orobanchaceae, Plantaginaceae, Scrophulariaceae). Lamiaceae, the largest family, is well represented in both tropical and temperate floras.

A basal split in the evolutionary relationships within the order sets off the family Oleaceae, with tetramerous flowers, which are radially symmetric, from the rest of the Lamiales. Pentamerous flowers with fewer stamens (two or four) than petals and bilateral symmetry predominate among the other families. Flowers of all families are characterized by petals fused into a tubular corolla, a single whorl of stamens (typically four), and ovaries consisting of two fused carpels. However, the number of ovules per ovary and, therefore, the number of seeds per fruit varies from one to many and helps form the basis for many family distinctions. Iridoid compounds are found in most members of the Lamiales and are thought to function in deterring herbivory by insects. However, iridoids are not universal in the order and, for example, the large subfamily Nepetoideae in family Lamiaceae lacks iridoids but exhibits a diverse array of volatile oils that may serve a similar function.

The volatile oils in the Lamiaceae have made many species valuable as scents and herbs, including lavender (*Lavandula*), mint (*Mentha*), sage (*Salvia*), catnip (*Nepeta*), basil (*Ocimum*), and oregano (*Oreganum*). Additional economically important plants in the Lamiales include those grown for food: sesame (*Sesamum*, Pedaliaceae) and olives (*Olea*, Oleaceae); wood: teak (*Tectona*, Verbenaceae) and ash (*Fraxinus*, Oleaceae); houseplants: African violets and relatives (Gesneriaceae); and many ornamental plants from a wide variety of families. The order also contains some plants with specialized habits, including parasitic plants (Orobanchaceae), insectivorous plants (Byblidaceae, Lentibulariaceae), mangroves (Avicenniaceae), and aquatic plants, such as *Callitriche* and *Hippuris* (Plantaginaceae). *See* ASTERIDAE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM. Richard Olmstead

Bibliography. Angiosperm Phylogeny Group, An ordinal classification of the families of flowering plants, *Ann. Missouri Bot. Gard.*, 85:531–553, 1998; R. M. Harley and T. Reynolds (eds.), *Advances in Labiate Science*, Royal Botanic Gardens Kew, Richmond, 1992; S. J. Wagstaff and R. G. Olmstead, Phylogeny of the Labiatae and Verbenaceae inferred from *rbc*L sequences, *Syst. Bot.*, 22:165–179, 1997.

Laminar flow

A smooth, streamline type of viscous fluid motion characteristic of flow at low-to-moderate deformation rates. The name derives from the fluid's moving in orderly layers or laminae without the formation of small eddies or irregular fluctuations.

The chief criterion for laminar flow is a relatively small value—less than a few thousand—for the Reynolds number, $\mathrm{Re} = \rho VL/\mu$, where ρ is fluid density, V is flow velocity, L is body size, and μ is fluid viscosity. Thus laminar flow may be achieved in many ways: low-density flows as in rarefied gases; low-velocity or "creeping" motions; small-size bodies such as microorganisms swimming in the ocean; or high-viscosity fluids such as lubricating oils. At higher values of the Reynolds number, the flow becomes disorderly or turbulent, with many small eddies, random fluctuations, and streamlines intertwining like spaghetti. *See* CREEPING FLOW; REYNOLDS NUMBER; TURBULENT FLOW; VISCOSITY.

Nearly all of the many known exact solutions of the equations of motion of a viscous fluid are for the case of laminar flow. These mathematically accurate descriptions can be used to give insight into the more complex turbulent and transitional flow patterns for which no exact analyses are known. *See* NAVIER-STOKES EQUATION.

The theory of viscous lubricating fluids in bearings is a highly developed area of laminar flow analysis. Even large Reynolds number flows, such as aircraft in flight, have regions of laminar flow near their leading edges, so that laminar flow analysis can be useful in a variety of practical and scientifically relevant flows. *See* ANTIFRICTION BEARING; BOUNDARY-LAYER FLOW; LUBRICANT. Frank M. White

Bibliography. D. F. Rogers, *Laminar Flow Analysis*, 1992; H. Schlichting, *Boundary Layer Theory*, 8th ed., 2000; F. S. Sherman, *Viscous Flow*, 1990; F. M. White, *Viscous Fluid Flow*, 2d ed., 1991.

Laminariales

An order of large brown algae (Phaeophyceae) commonly called kelps. Definitive features include a life history in which microscopic, filamentous, dioecious gametophytes alternate with a massive, parenchymatous sporophyte; growth of the sporophyte effected by meristems; and production of unilocular zoosporangia in extensive sori on blades of the sporophyte. *See* PHAEOPHYCEAE.

General structure. A mature sporophyte (**Fig. 1***a*) consists typically of a holdfast, stipe, and one or more blades. The holdfast usually comprises a cluster of rootlike structures (haptera); in a few species it is discoid or conical. The stipe, which varies in length from a few millimeters in *Hedophyllum* to more than 100 ft (30 m) in *Macrocystis* (giant kelp), may be branched or unbranched and may bear one or more blades. In some genera, fertile blades are distinct from vegetative blades. In the transition region between the stipe and the blade there is a meristem, which adds material to the upper end of the stipe and the lower end of the blade. The girth of the stipe is increased by activity of a superficial meristem (meristoderm). In some genera the sporophyte is annual; in others the stipe is perennial and develops annual growth rings while the blade is regenerated annually

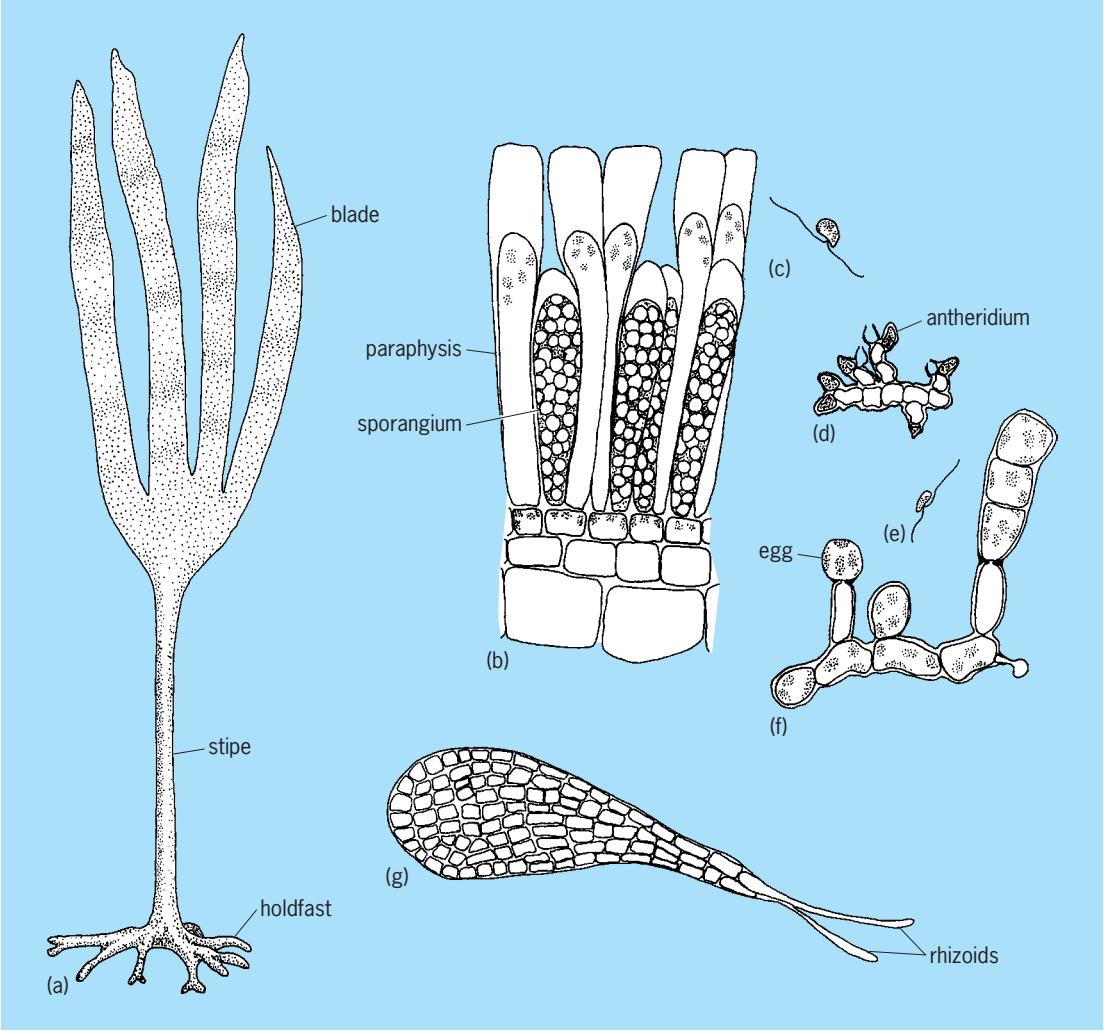

Fig. 1. ***Laminaria*: (*a*) sporophyte; (*b*) cross section of blade, showing sorus of unilocular sporangia interspersed with sterile hairs (paraphyses); (*c*) zoospore; (*d*) male gametophyte with antheridia; (*e*) antherozoid (sperm); (*f*) female gametophyte with egg (left) and three-celled sporophytic germling (right), each perched on an empty oogonium; (*g*) embryonic sporophyte with rhizoids.**

by the transition meristem. A few sporophytes are capable of almost indefinite growth.

Anatomy. Laminariales have a more complex anatomy than any other group of algae, although Fucales (rockweeds) nearly equal them in this regard. A mature stipe comprises three regions: a photosynthetic and meristematic thin outer layer (meristoderm); a parenchymatous cortex; and a filamentous medulla. The cortex is often pierced by mucilage ducts lined with small secretory cells. The medulla comprises both ordinary filaments and sieve tubes, the latter composed of superposed sieve elements with perforated end walls. Both types of filaments are interconnected by transverse filaments. Sieve tubes have been shown to translocate mannitol, a water-soluble sugar alcohol. Translocation is significant in large kelps such as *Macrocystis, Nereocystis,* and *Pelagophycus,* where photosynthate produced by blades floating on the surface is moved long distances to developing stipes and holdfasts at heavily shaded depths. The anatomy of the blade is similar to that of the stipe, except that the meristoderm is thinner and less active in cell division.

Life history and reproduction. The microscopic filamentous gametophytes, except in *Chorda tomentosa*, are dioecious. Female gametophytes (Fig. 1*f*) consist of relatively few stout cells and are sparsely branched. An egg, produced in an intercalary oogonium, is extruded through a rupture in the oogonial wall and remains attached to a mucilaginous cushion. The eggs of some species liberate a chemical that attracts sperm of the same species. Male gametophytes (Fig. 1*d*) have more cells, which are more slender, and are profusely branched. Antheridia, which are produced in clusters, each contain a biflagellate sperm (Fig. 1*e*). The zygote, attached to the empty oogonium, initiates a fairly regular sequence of cell divisions, resulting in an embryo that becomes anchored to the substrate by rhizoids (Fig. 1*g*). Eventually, the gametophyte is overgrown and the rhizoids are replaced by a permanent holdfast. The juvenile plant is clearly differentiated into holdfast, stipe, and blade when it is only 0.8–1.2 in.

(2–3 cm) high. Unilocular sporangia are produced in sori (Fig. 1*b*), accompanied by unicellular hairs (paraphyses), covering large portions of blades. The initial nucleus of a developing sporangium is diploid and undergoes meiosis followed by several mitotic divisions. Mature sporangia commonly contain 64 biflagellate zoospores (Fig. 1*c*), which germinate to produce gametophytes, half of them male and half female.

Classification. Four families of kelps are recognized. The sporophyte of *Chorda,* the only genus in the family Chordaceae, is unique among kelps in being whiplike and not differentiated into stipe and blade. Sporophytes of Laminariaceae display a simple plan of holdfast, stipe, and unbranched (but often laciniate) blade. Genera are distinguished by various patterns of ribs, folds, and perforations of the blade. *Laminaria,* which lacks such modifications, may be considered the basic genus of the family and perhaps of the entire order.

In sporophytes of Alariaceae, there is a differentiation between sterile and fertile blades, the latter (sporophylls) growing from the lateral margins of the blade, transition zone, or stipe. *Alaria* has a long sterile blade, which is ruffled and has a prominent midrib, and two clusters of short winglike sporophylls, one on either side of the transition zone. The long straplike stipes of *Egregia* (feather boa kelp) are fringed on either side with a mixture of sterile blades (some transformed into air bladders) and sporophylls.

In sporophytes of Lessoniaceae, the primary blade splits longitudinally, but unlike Laminariaceae, in which splitting affects only the blade, the splits extend into the transition zone. The meristem is thus dispersed, resulting in branching of the stipe, with each branch terminating in a blade. Sporophytes of some Lessoniaceae are essentially annual, with the thallus either torn from its substrate by winter storms or disintegrating in place; those of other Lessoniaceae are perennial and capable of nearly indefinite growth.

Nereocystis (bull kelp) is a spectacular plant of the Pacific coast of North America from Point Conception, California, northward. Its long, smooth, terete stipe expands into a large bulbous bladder just below the transition zone. Immediately above the bladder, the stipe branches dichotomously and produces numerous blades. The bladder, which buoys the blades at the surface, contains a mixture of gases including carbon monoxide. The sporophyte of the bull kelp appears in the spring and usually disappears in the winter, but it may persist without further growth into the second year. It reaches a length of at least 65 ft (20 m) within a few months, making it one of the world's fastest-growing plants. *Pelagophycus* (elk kelp), with antlerlike branches above the terminal bladder, is the counterpart of the bull kelp in southern California and northern Baja California.

The sporophyte of *Macrocystis* (giant kelp), which is perennial, is anchored by a massive cluster of haptera about 3.5 ft (1 m) broad. The numerous long branches may be regenerated from their base in response to abrasion, grazing, or harvesting. Each ultimate branch ends in a falcate blade, which splits progressively from base to apex to form a series of linear blades, each subtended by a bladder (**Fig. 2**). Sporophylls are distinct structures borne low on the major branches. *Macrocystis* equals or exceeds *Nereocystis* in rapidity of growth.

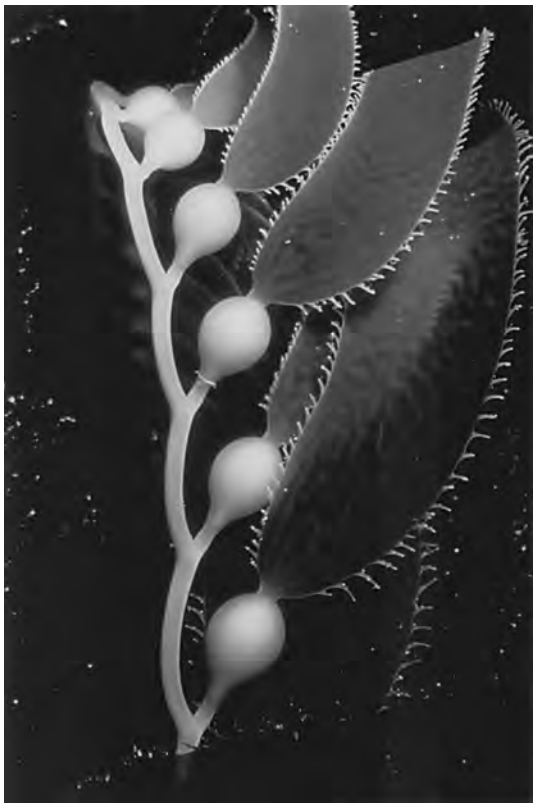

Fig. 2. ***Macrocystis*** **(giant kelp). Part of frond, showing series of blades, each subtended by a bladder. (*Underwater photo by Sylvia A. Earle*)**

Distribution. Kelps are largely confined to cold waters of both hemispheres, displaying a peak of diversity, abundance, and luxuriance all around the North Pacific. Only *Lessonia* and *Ecklonia* (both Lessoniaceae) are characteristic of the Southern Hemisphere, but *Macrocystis* is bipolar, occurring abundantly in subantarctic waters. Kelps are lacking in Antarctica, where their ecological niche is filled by members of the brown algal family Desmarestiaceae.

Some kelps are found in the intertidal zone, usually on rocks exposed to heavy waves, but most are subtidal. *Postelsia* (sea palm) shows amazing tenacity in facing the full force of the surf in the most highly exposed sites. *Macrocystis pyrifera* forms extensive subtidal stands regardless of exposure. The plants are usually anchored at depths of 20–65 ft (6–20 m), with their blades buoyed up toward or on the surface of the water. *Nereocystis* seeks slightly shallower water, often in protected coves.

Economic and ecological importance. Kelps were traditionally used by coastal farmers of the North

Atlantic as a wet manure, or were burned to an ash, which was then used as a fertilizer or in industry. The word kelp, in fact, was originally applied to the ash. During World War I, a potash industry flourished on the Pacific coast of the United States, using *Macrocystis* as the source. The principal use of kelps today is in the alginate industry, with *Macrocystis* the chief source in California, and *Laminaria* together with the rockweed *Ascophyllum* the chief source in the North Atlantic. Various kelps are used extensively for human food in Japan, including *Laminaria* (kombu) and *Undaria* (wakame). In China *Laminaria japonica* has been established as an important maricultural crop plant. Ecologically, such kelps as *Macrocystis* and *Nereocystis* are particularly important because they form protective canopies harboring special communities of smaller seaweeds, invertebrates, fishes, and the sea otter. *See* ALGAE; ALGINATE. Paul C. Silva; Richard L. Moe

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae: Structure and Reproduction*, 2d ed., 1997.

Lamniformes

An order of sharks distinguished from all other sharks by the following combination of characters: trunk cylindrical; head conical to cylindrical, not laterally expanded; two dorsal fins, lacking spines; five gill slits, all anterior to the pectoral fin base or the last two above the pectoral fin base; mouth large, its gape extending well beyond the eyes; spiracles usually present, small and well behind the eyes; eyes without a nictitating membrane; barbels absent; and ring-type intestinal valve. Lamniform sharks are ovoviviparous. Fertilized eggs are enclosed in egg cases in the oviducts (uteri), where they undergo embryonic development, ordinarily from the energy supplied by the yolk. Although the reproductive event may begin with each uterus containing about 20 fertilized eggs, normally only one or two in each uterus reach full term. The gross loss of potential progeny is explained by oophagy or uterine cannibalism, an uncommon phenomenon in which the largest (and presumably oldest) fetuses are free in the uterus to feed on the remaining eggs and developing embryos, thus cannibalizing their potential siblings. This reproductive strategy sacrifices numbers for a few large, strong pups that are like small adults, ready to fend for themselves. *See* ELASMOBRANCHII; SELACHII.

Lamniformes comprises only 15 extant species in seven families (see **illustration**), four of which are monotypic.

Odontaspidae (sand tiger sharks). Sand tiger sharks are distinguished by gill openings that are all in front of the pectoral fin; large teeth with slender cusps and lateral cusplets; a compressed caudal peduncle, without a lateral keel but with an upper precaudal pit; and an asymmetrical caudal fin with a relatively short ventral lobe. Females are the larger of the sexes, with a maximum total length of 370 cm (12 ft).

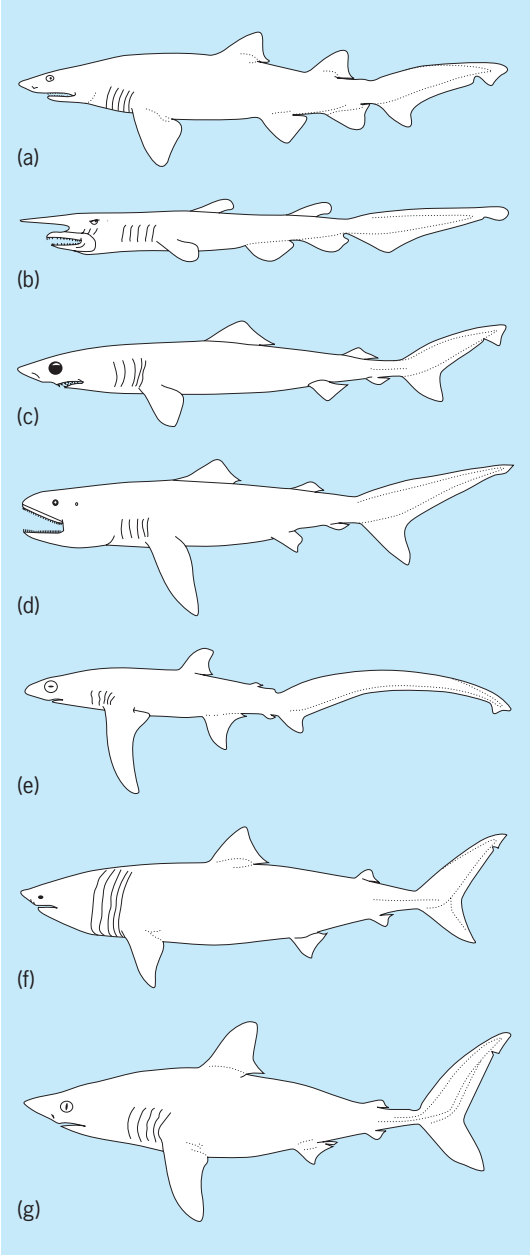

Representatives of seven Lamniformes families. (*a*) Odontaspidae. (*b*) Mitsukurinidae. (*c*) Pseudocarchariidae. (*d*) Megachasmidae. (*e*) Alopiidae. (*f*) Cetorhinidae. (*g*) Lamnidae. (*From J. S. Nelson, Fishes of the World, 4th ed., Wiley, New York, 2006*)

The family comprises two genera, *Carcharias* with one species and *Odontaspia* with two species. The snout is short and flattened in *Carcharias*, and long and conical in *Odontaspia*; dorsal, anal, and pelvic fins are about the same size in *Carcharias*, whereas the first dorsal is larger than the second, and the second is larger than the anal, in *Odontaspia*.

Odontaspids are known from tropical to cool-temperate zones of the Atlantic, Indian, and Pacific oceans, from continental and insular shores to the outer shelves and slopes to about 1600 m (5250 ft). They are voracious feeders on a wide variety of bony fishes, other sharks, rays, squids, and bottom crustaceans. Schools of sand tigers have been observed cooperatively surrounding and bunching schooling prey and then feeding on them. Of the three species, *Carcharias taurus* is by far the best known biologically. The gestation period is 8 to 9 months and the newly born are 95–105 cm (37–42 in.) in total length.

Relatively few attacks by these sharks on people have been recorded, but any large shark should be regarded as potentially dangerous and should not be harassed or provoked.

Mitsukurinidae (goblin sharks). This family consists of only one extant species, *Mitsukurina owstoni*, a bizarre shark that occurs in apparently disjunct populations in the eastern and western Atlantic (including the Gulf of Mexico), off South Africa, in the eastern Pacific (including southern California), and off southern Australia. It is easily recognized by an elongated and flattened snout; small eyes; soft flabby body and fins; very protrusible jaws; teeth with long slender cusps and no cusplets; dorsal fins about equal in size, and smaller than anal and pectoral fins; no precaudal pit; and caudal fin without a ventral lobe. Maximum length is 3.8 m (12.5 ft).

The goblin shark is an uncommon, poorly known, bottom-dwelling shark that inhabits outer continental shelves and beyond to at least 550 m (1800 ft), but rarely occurs near shore. Very little is known of its biology; however, its morphology suggests a relatively inactive, slow-swimming shark that probably uses its sensitive snout to detect prey and its highly protractile jaws to reach out and seize the prey, which may be squid, shrimp, crabs, or a variety of fishes.

Pseudocarchariidae (crocodile sharks). This is a monotypic family, represented by *Pseudocarcharias kamoharai*, which occurs in scattered localities, usually oceanic, in tropical and subtropical waters of the Northern and Southern hemispheres. It has a slender spindle-shaped body; long gill slits that reach dorsally past the axis of the body; highly protrusible jaws that are equipped with long slender teeth; angular dorsal fins; a prominent ventral caudal fin lobe; exceptionally large eyes; upper and lower precaudal pits; and a low lateral keel. This is the smallest lamniform shark, with a maximum adult length of 110 cm (43 in.); it is 41 cm (16 in.) at birth.

Megachasmidae (megamouth shark). This monotypic family, represented by *Megachasma pelagios*, was discovered in 1976 off Oahu, Hawaii. Since then, 35 more specimens (as of May 2006) have been captured in scattered places in warm temperate and tropical shallow coastal waters [depth: 5 m (16 ft)] to oceanic waters, where it inhabits the epipelagic zone (the illuminated surface zone) to depths of 166 m (545 ft). Megamouth is a very distinctive shark in having a large terminal mouth; large head; cylindrical and somewhat compressed trunk; snout that is exceptionally short but broadly rounded; moderately long gill slits, the last two over the pectoral fin base; upper and lower precaudal pits; small teeth in numerous rows; and gill rakers in the form of fingerlike dermal papillae. At a maximum length of 5.5 m (18 ft), megamouth is the smallest of the three giant filter-feeding sharks, the others being the basking shark [10 m (33 ft) or more] and the whale shark [12 m (39 ft) or more].

Alopiidae (thresher sharks). Thresher sharks are easily distinguished by the greatly elongate and curved upper caudal fin lobe, which is about half their total length; long narrow pectoral fins; relatively small gill slits, with the third to fifth over the origin of the pectoral fin; large eyes; a small mouth; precaudal pits; and a caudal peduncle without keels. Maximum length is 5.7 m (18.7 ft).

Thresher sharks occur practically circumglobally in temperate to tropical waters from coastal to oceanic regions. They feed on a variety of schooling fishes, such as mackerels and herrings, as well as squids, octopuses, and pelagic crustaceans. The long caudal fin is used to herd fishes and to slash through a school, killing or stunning prey to be consumed later. There is one genus with three species: *Alopias pelagicus* (pelagic thresher), 347 cm (11 ft) total adult length, 96 cm (38 in.) at birth; *A. superciliosus* (bigeye thresher), 488 cm (16 ft) total adult length, 105 cm (41 in.) at birth; and *A. vulpinus* (thresher shark), 760 cm (25 ft) total adult length, 150 cm (59 in.) at birth.

Cetorhinidae (basking sharks). Cetorhinidae is another monotypic family, which is represented by the basking shark (*Cetorhinus maximus*). It occurs on continental and insular shelves, offshore and often close to land, just free of the surf zone, in boreal (northern) to warm temperate waters of the Atlantic and Pacific oceans. It is easily distinguished from other sharks by the following combination of characters: exceptionally large gill openings practically encircling the head and all in front of the pectoral fin; hairlike gill rakers of modified dermal denticles on the gill arches; very small and hooklike teeth; nearly lunate caudal fin, with a well-developed lower lobe; precaudal pits; and caudal peduncle with strong lateral keels.

The basking shark is a filter feeder, but its method differs from the other two huge filter-feeding sharks (megamouth and whale shark), as it passively strains plankton while swimming with mouth agape and external gill slits open, as opposed to pumping or gulping water into the pharynx and force-expelling it across the gill rakers. The basking shark is the world's second largest living fish, reportedly reaching 15 m (49 ft) in total length, but specimens exceeding 10 m (33 ft) are exceptional. Size at birth is unknown; however, the smallest known free-living basking shark measured 165 cm (65 in.).

Lamnidae (mackerel sharks). Mackerel sharks occur virtually globally from nearshore and the

continental shelf to epipelagic zones of cold-temperate to tropical waters. They are distinguished from other sharks by the following combination of characters: fusiform body moderately slender to robust; snout moderately long, pointed, and conical; large gill openings, extending above the axis of the body (except in *Lamna*) and all in front of the pectoral fin base; no gill rakers on the gill araches; first dorsal fin large, high, erect, and angular or somewhat rounded; second dorsal and anal fins very small; caudal fin lunate or crescent-shaped and caudal peduncle with distinct keels. These are fast-swimming predators with a high level of activity made possible by partial warm-bloodedness (sustained activity is by warm musculature). The family comprises three genera and five living species.

Carcharodon carcharias (white shark or great white shark) inhabits coastal and offshore continental and insular waters and occasionally shallow bays. It is one of the most widely ranging sharks, although uncommon in the tropical waters. Certainly, it is the most notorious and most feared shark. Its large size, large teeth, powerful jaws, and powerful and efficient scombrid-like mode of locomotion that allows it to sustain long cruises make this shark a formidable predator. It feeds on a wide variety of bony fishes, as well as on sharks, rays, seals, dolphins and porpoises, sea birds, squid, octopi, crabs, and carrion. Although its role as a man-eater has been exaggerated, the white shark is still considered very dangerous.

The white shark probably does not exceed 7.5 m (25 ft) in total length. Captured individuals are commonly between 1.4 and 6 m (4.5 and 20 ft). Even at 6 m, it is the world's largest predator. The famous late Pliocene fossil, *Carcharodon megalodon*, known commonly as megatooth shark or megalodon, which reached lengths up to 20 m (66 ft), is a member of this family.

Isurus oxyrinchus (shortfin mako) is a coastal and oceanic inhabitant of the temperate and tropic seas of the world. It is probably the fastest of all sharks, capable of extreme bursts of speed when hooked or in pursuit of prey, and can leap out of the water to a height several times its length. Its maximum total length is about 400 cm (13 ft).

Isurus paucus (longfin mako) is a little-known circumglobal epipelagic shark in the warm-temperate and tropical waters. It is not uncommon in the western North Atlantic and places in the mid-Pacific, but it is rather rare elsewhere. The extremely long pectoral fins suggest a less active shark than its congener, the shortfin mako. Its maximum total length is 417 cm (14 ft).

Lamna ditropis (salmon shark) is the most robust of the lamnid sharks. It is a coastal and oceanic inhabitant of the North Pacific from Japan and Korea to the Bering Sea, and southward to southern California. It is a swift swimmer and voracious predator, and maintains a body temperature several degrees above ambient water temperature. It preys heavily on salmon and attains a maximum total length of about 305 cm (10 ft).

Lamna nasus (porbeagle) is a epipelagic shark that occurs in the western North Atlantic from New Jersey to the Gulf of St. Lawrence, across the Atlantic to the Mediterranean Sea and northward to the western Barents Sea. It also occurs in the cold temperate waters of the Southern Hemisphere, primarily in the southern Indian Ocean, off southern Australia and New Zealand, and off southern South America, tending to avoid water temperatures greater then 18°C (65°F). The porbeagle is most common on offshore fishing banks, where it feeds on a variety of bony fishes, as well as on other sharks and squids. Maximum total length is no less than 300 cm (9.8 ft), possibly 370 cm (12 ft), and it is 60–75 cm (24–30 in.) at birth. Herbert Boschung

Bibliography. H. B. Bigelow and W. C. Schroeder, Sharks, in Fishes of the Western North Atlantic, *Mem. Sears Found. Mar. Res. Mem.*, 1(pt. 1):29–556, 1948; L. J. V. Compagno, Checklist of Chondrichthyes, pp. 503–547, in W. C. Hamlett (ed.), *Reproductive Biology and Phylogeny of Chondrichthyes: Sharks, Batoids and Chimaeras*, Science Publishers, Enfield, NH, 2005; L. J. V. Compagno, Checklist of living elasmobranches, pp. 471–498, in W. C. Hamlett (ed.), *Sharks, Skates, and Rays: The Biology of Elasmobranch Fishes*, Johns Hopkins University Press, Baltimore, 1999; L. J. V. Compagno, *FAO Species Catalogue*, vol. 4: *Sharks of the World: An Annotated and Illustrated Catalogue of Shark Species Known to Date, Part 1: Hexanchiformes to Lamniformes*, FAO Fish. Synop. 125, vol. 4, pt. 1:212–249, 1984; L. J. V. Compagno, *Sharks of the Order Carcharhiniformes*, Princeton University Press, 1988; L. J. V. Compagno, *Sharks of the World: An Annotated and Illustrated Catalogue of Shark Species Known to Date, FAO Species Catalogue for Fishery Purposes* no. 1, vol. 2: *Bullhead, Mackerel and Carpet Sharks (Heterodontiformes, Lamniformes and Orectolobiformes)*, FAO, Rome, 2001; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Lamp

A generic term for a device designed to produce electromagnetic radiation in the form of light, heat, ultraviolet energy, or a combination of the three. The term lamp is applied to the entire range of sources, including flame sources (such as kerosine lamps or gas lamps with Welsbach mantles), incandescent sources, and electric arc discharge sources. Used with a modifier, such as ultraviolet, infrared, or sun, the term lamp is used to indicate sources that radiate energy in the ultraviolet or infrared portions of the electromagnetic spectrum (plus some radiation in the visible part of the spectrum).

Electric lamps are the most common and useful types of light-producing sources. These devices convert electrical energy into light (radiant energy in the visible portion of the spectrum). The most widely used lamps for lighting buildings and

areas, such as parking lots, stadia, and streets, are incandescent lamps, fluorescent lamps, and high-intensity arc discharge lamps (HID lamps). *See* ARC LAMP; FLUORESCENT LAMP; INCANDESCENT LAMP; INFRARED LAMP; SUN; ULTRAVIOLET LAMP; VAPOR LAMP.

G. R. Peirce; Robert Leroy Smith

Bibliography. Illuminating Engineering Society, *IES Lighting Handbook*, 8th ed., 1993.

Lampriformes

An order of fishes also known as Lampridiformes and Allotrignathi. Although the families of Lampriformes appear very dissimilar, they share the following combination of characters that distinguish them from other orders of actinopterygian (ray-finned) fishes: notably compressed body, often ribbonlike; no true spines in fins, rays only; thoracic pelvic fins, if present; swim bladder, if present, physoclistous; and small cycloid scales, commonly absent. In addition, lampriforms have a unique type of protrusible upper jaw, by which the premaxilla excludes the maxilla from the gape, and the maxilla, instead of having a ligamentous attachment to the ethmoid and palatine bones, slides in and out with the highly protractile premaxilla. Most species are large and colorful; some have inspired sea-serpent stories. *See* ACTINOPTERYGII; SWIM BLADDER.

Strong evidence has been presented by several authors supporting the notion that lampriforms are not percomorphs but prepercomorphs and are the primitive sister group to the remaining Acanthomorpha (spiny-rayed fishes). Dating from the Cretaceous Period (135 million years before present), lampriform fishes are presently represented by 21 species in 12 genera and seven families.

Two families (Veliferidae and Lampridae), collectively known as bathysomes, have deep compressed bodies with a symmetrical caudal fin and a well-developed skeleton; the remaining five families (Stylephoridae, Lophotidae, Radiicephalidae, Trachipteridae, and Regalecidae), called taeniosomes, have long ribbonlike bodies with an asymmetrical caudal fin and a weak skeleton.

Veliferidae (velifers). These fish are found in the Indian and western mid-Pacific oceans, usually near the shore. Dorsal and anal fins long and high, with 32–44 and 25–35 rays, respectively; lateral line follows dorsal contour of the body; teeth absent; swim bladder is bifurcate and extends posteriorly far beyond anus; and six branchiostegal rays. Maximum length about 40 cm (16 in.). Two monotypic genera, *Velifer* and *Metavelifer* (**Fig. 1**).

Lampridae (opahs). The opahs are pelagic, living in the open Atlantic, Indian, and Pacific oceans. The dorsal fin is falcate (sickle-shaped) anteriorly; otherwise, it and the anal fins are long and low, with 48–56 and 33–42 rays, respectively; the lateral line is arched high in front; the pectoral fins are falcate; and minute cycloid scales are present. Maximum length is up to 180 cm (71 in.). Two species, *Lampris guttatus* (worldwide) and *L. immaculata* (cold and temperate waters of the Southern Hemisphere), are known.

Fig. 1. Spinyfin velifer (*Metavelifer multiradiatus*). (*Photo by Jack E. Randall*)

Stylephoridae (tube-eyes or thread-tails). These are oceanic fishes, occurring at depths of 300–800 m (980–2620 ft). The body is slender and ribbonlike with up to 200 vertebrae; the dorsal fin extends from the nape to the tail, with 115–124 rays; the anal fin is short, with 15–17 rays; the pelvic fin has only one ray; the caudal fin has seven rays, with the lowermost two being extremely elongate; eyes are large, telescopic, and directed forward or upward; the mouth is small; the snout is tubular; teeth are small; and there is no swim bladder. The only described species, *Stylephorus chordatus*, swims in a vertical position, head uppermost, and feeds by rapidly expanding its buccal cavity, effectively sucking in plankton. Maximum length is 31 cm (12 in.).

Lophotidae (crestfishes). These are oceanic fishes. The dorsal fin originates above or before the tip of the snout and continues to the caudal fin, with 220 to 392 rays; the anal fin is short and posteriorly placed; the body has small deciduous cycloid scales and may appear naked; the caudal fin is symmetrical or nearly so; the pelvic fins are present or absent; the swim bladder is present; and the ink sac opens into the cloaca, with the ink possibly serving to blind predators. Two genera and three species have been identified. Maximum length is about 200 cm (80 in.).

Radiicephalidae (tapertails). These fish are found in the central and eastern Atlantic, off New Guinea, and in the California Current region. The body tapers to a thin caudal filament; the caudal fin has a small upper lobe of four rays and a long slender lower lobe of seven rays; the dorsal fin possesses 152–159 rays; the anal fin is vestigial; scales are found along the lateral line, with the remainder of the body naked; the swim bladder is well developed; the brown ink sac discharges into the cloaca; and there are 114–121 vertebrae. One species, *Radiicephalus elongates*,

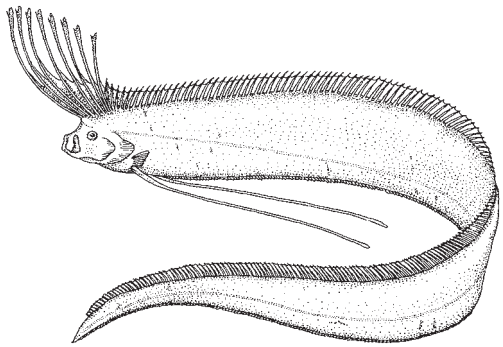

Fig. 2. Oarfish (*Regalecus glesne*). (*After D. S. Jordan and B. W. Evermann, The Fishes of North and Middle America, U.S. Nat. Mus. Bull. 47, 1900*)

is known, although it is rarely captured. Maximum length is about 70 cm (28 in.).

Trachipteridae (ribbonfishes). The ribbonfishes are bathypelagic [living at ocean depths between 600 and 3600 m (2000 and 12,000 ft)] in the Arctic, Atlantic, Indian, and Pacific oceans, as well as the Mediterranean Sea. The body is naked or covered with deciduous cycloid scales or with deciduous modified ctenoid scales; there is no anal fin; the caudal fin is long and at a right angle to the body, usually consisting of an upper lobe only; the dorsal fin is very long, originating distinctly behind the tip of the snout; the eyes are large; teeth are present; the swim bladder is rudimentary or absent; and there are 62–111 vertebrae. Three genera (*Desmodema*, *Trachipterus*, and *Zu*) and 10 species are known. Maximum length is about 170 cm (67 in.).

Regalecidae (oarfishes). Oarfishes are found worldwide in warm-temperate seas. The body is very elongate; scales are absent; the skin is covered with tubercles; the dorsal fin is long (260–412 rays), beginning above the eye and continuing to nearly the end of the tail; the front rays are very long; the pelvic fins are long and slender, with one ray, having a paddlelike expansion near the tip; the anal fin, teeth, and swim bladder are absent; the eyes are small; and there are about 143–170 vertebrae. Two species, *Regalecus glesne* and *Agrostichthys parkeri*, have been identified.

Regalecus glesne (**Fig. 2**) is called oarfish locally, but king-of-the-herring in Europe, from an old tale of this fish announcing the arrival of herrings. The oarfish is a spectacular creature, having a huge ribbonlike body, bright red dorsal fin (the front part crowning the head with long plumelike rays), and long bright-red pelvic fins. Attaining a maximum length of 800 cm (26 ft), it is the longest bony fish and undoubtedly the source of many sea-serpent stories.

Herbert Boschung

Bibliography. S. R. Charter and H. G. Moser, Radicephalidae: Tapertails, pp. 665–667 in H. G. Moser (ed.), The early stages of fishes in the California Current region, *California Cooperative Oceanic Fisheries Investigations (CalCOFI) Atlas*, no. 33, 1996; G. D. Johnson and C. Patterson, Percomorph phylogeny: A survey of acanthomorphs and a new proposal, *Bull. Mar. Sci.*, 52(1):554–626, 1993; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; J. E. Olney, Lampriformes: Development and relationships, in H. G. Moser et al. (eds.), *Ontogeny and Systematics of Fishes*, Spec. Publ. no. 1, American Society of Ichthyology and Herpetology, 368–379, 1984; J. E. Olney, G. D. Johnson, and C. C. Baldwin, Phylogeny of lampridiform fishes, *Bull. Mar. Sci.*, 52(1):137–169, 1993; C. R. Robins and G. C. Ray, *A Field Guide to Atlantic Coast Fishes of North America*, Haughton Mifflin, Boston, 1986; M. M. Smith and P. C. Heemstra (eds.), *Smith's Seafishes*, Macmillan South Africa, Johannesburg, 1986; V. Walters, Synopsis of the lampridiform suborder Veliferoidei, *Copeia*, 1960(3):245–247, 1960; E. O. Wiley, G. D. Johnson, and W. W. Dimmick, The phylogenetic relationships of lampridiform fishes (Teleostei: Acanthomorpha), based on a total evidence analysis of morphological and molecular data, *Mol. Phylogenet. Evol.*, 10(3):417–425, 1998.

Lamprophyre

Any of a heterogeneous group of gray to black, mafic igneous rocks characterized by a distinctive panidiomorphic and porphyritic texture in which abundant euhedral, dark-colored ferromagnesian (femic) minerals (dark mica, amphibole, pyroxene, olivine) occur in two generations—both early as phenocrysts and later in the matrix or groundmass—while felsic minerals (potassium feldspar, plagioclase, analcime, melilite) are restricted to the groundmass (see **illus.**). Compared with the common igneous rocks,

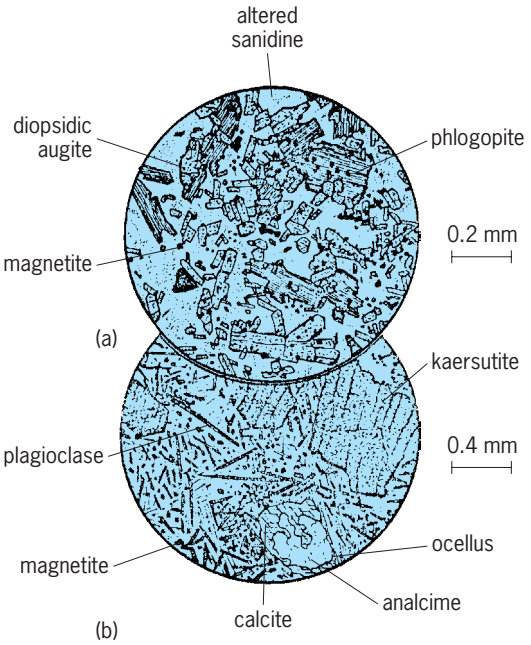

Specimens of lamprophyre. (*a*) Minette, found near Walsenburg, Colorado. The altered sanidine matrix also contains other minor minerals. (*b*) Camptonite, found in Campton Falls, New Hampshire. (*After H. Williams and C. M. Gilbert, Petrography: An Introduction to the Study of Rocks in Thin Sections, W. H. Freeman, 1954*)

Common lamprophyres

Dominant or diagnostic femic mineral	Dominant or diagnostic felsic mineral			
	K-feldspar	Plagioclase	Analcime or glass	Melilite
Phlogopite	Minette	Kersantite	—	Alnoite
Hornblende	Vogesite	Spessartite	—	—
Kaersutite	—	Camptonite	Monchiquite	—

lamprophyres are chemically peculiar. They have low silicon, moderate aluminum, and variable calcium contents, but are rich in alkalies (potassium or sodium or both), magnesium, iron, and volatile constituents (H_2O and CO_2), and contain a wide variety of such minor and trace elements as titanium, phosphorus, barium, strontium, rubidium, zirconium, lanthanum, uranium, thorium, chromium, nickel, and cobalt. Many lamprophyres contain ocelli, small spheroidal bodies consisting of alkali-rich feldspar, analcime, or calcite with minor femic minerals.

Mineralogy and terminology. Many varieties of lamprophyre are known; only the more abundant ones are listed in the **table**. Minettes, kersantites, vogesites, and spessartites are the most common and are sometimes collectively called calcalkaline lamprophyres. Camptonites and monchiquites are less common; alnoites are rare. These varieties, along with some others, are referred to as alkaline lamprophyres. Although lamprophyres are named on the basis of their dominant or diagnostic femic and felsic minerals, their terminology is not straightforward because there is much gradation in mineral assemblages, especially in number and abundance of femic minerals, and felsic minerals are often too altered to permit identification.

All lamprophyres contain essential hydrous minerals, such as dark mica and amphibole, frequently accompanied by significant pyroxene and less olivine. The mica is a magnesium-rich titanium-bearing biotite or phlogopite. Under the microscope the biotite in a thin section of a lamprophyre can be seen as pseudo hexagonal plates that are chemically zoned. The pale magnesium-rich centers of the crystals are surrounded by reddish-brown iron-rich rims; groundmass biotites are chemically similar to phenocryst rims. Amphibole is less common than biotite; it forms prisms or needles of a green or brown hornblende in calcalkaline lamprophyres, and a deep brown kaersutite in alkaline lamprophyres. Pyroxene occurs as stubby monoclinic crystals of diopsidic augite or titanaugite, frequently zoned. Orthorhombic pyroxene is absent. Magnesium-rich olivine is usually the least abundant femic mineral; unlike the biotite, amphibole, and pyroxene, it generally occurs only as phenocrysts and frequently is altered and pseudomorphed by calcite and other secondary minerals such as chlorite and members of the montmorillonite group.

Although in hand specimen the groundmass appears dark, a large proportion of it is composed of colorless plates or laths of potassium feldspar (sanidine or orthoclase) or plagioclase feldspar (albite-andesine in calcalkaline lamprophyres, andesine-labradorite in alkaline). In the rare feldspar-free varieties, monchiquite and alnoite, the groundmass consists predominantly of an alkali-rich commonly altered glass ± analcime, or melilite and a calcium carbonate. Apatite, a magnetitelike spinel, and calcite (primary or secondary or both) are ubiquitous accessory minerals in all lamprophyres; other minerals that may occur include sphene, quartz, analcime, zeolites, nepheline, haüyne, monticellite, pyrite, perovskite, and ilmenite (many of these may be secondary).

Occurrence. Lamprophyres are widespread but volumetrically minor rocks that apparently are restricted to the continents and are the last manifestation of igneous activity in a given area. They usually occur as subparallel or radial swarms of thin (~1.6–160 ft or 0.5–50 m) dikes or, less commonly, sills, volcanic neck fillings, or diatremes, or, rarely, lava flows (lamprophyric lavas are sometimes called lamproites). A swarm may consist of one or more varieties of lamprophyre, but calcalkaline and alkaline lamprophyres usually do not occur together. Calcalkaline varieties frequently are found in the region of large granodiorite or granite bodies, while alkaline lamprophyres tend to be associated with alkaline basalts, nepheline syenites, carbonatites, or kimberlites; many exceptions to this generalization are known.

Alkaline lamprophyres and the rare minette may bear xenocrysts and xenoliths, foreign crystals and rock fragments representing the material that composes the Earth's mantle under the continents at depths greater than approximately 47 mi (75 km). Xenoliths of the continental crust can be found in all lamprophyres. *See* PHENOCRYST; XENOLITH.

Origin. Like all igneous rocks, a lamprophyre is the product of the crystallization of magma (molten rock). The ultimate origin of lamprophyre magmas is the upper portion of the Earth's subcontinental mantle where localized heating causes partial melting (anatexis) of mantle rock. The exact composition of the liquid or magma thus formed, which differs from that of the original unmelted mantle rock (source rock), depends on several factors, including the temperature and pressure at which anatexis occurs, the composition of the source rocks, and the degree of partial melting. Since magma is less dense than solid rock, it rises upward, invades the overlying continental crust, and eventually solidifies. During its ascent from the mantle to its site of final emplacement, a magma's composition may be changed by several processes, two of the most important being

assimilation of foreign material, especially portions of the rocks making up the continental crust; and partial (fractional) crystallization of the magma, resulting in removal of early-formed crystals of femic minerals because of density differences between these crystals and the magma from which they were formed.

The controversy surrounding the origin of lamprophyres centers on the chemical composition of the mantle-derived magma and the extent to which it has been modified prior to its final solidification. Some geologists maintain that a normal basaltic magma (the most common product of anatexis of the mantle) is made lamprophyric by a combination of the magma composition-changing processes. But continuing studies of the major- and trace-element chemistry of lamprophyres is indicating increasingly that many minettes and alkaline lamprophyres result from the crystallization of nearly unmodified mantle-derived magmas. This implies that magmas of lamprophyric composition can be generated within the upper mantle. Although lamprophyric magmas are believed to form by small degrees of partial melting of mantle rock of unusual chemical and mineralogical composition, details of the origin of these magmas are unclear. By the operation of processes such as assimilation of foreign material and partial magma crystallization, lamprophyric magmas may themselves change or evolve during their ascent, and this probably accounts for the gradations between the varieties. The origin of kersantites, vogesites, and spessartites is still a matter of debate; they may represent modified minettelike or other lamprophyric magma. *See* IGNEOUS ROCKS; MAGMA. Sharon W. Bachinski

Land drainage (agriculture)

The removal of water from the surface of the land and the control of the shallow ground water table improves the soil as a medium for plant growth. The sources of excess water may be precipitation, snowmelt, irrigation water, overland flow or underground seepage from adjacent areas, artesian flow from deep aquifers, floodwater from channels, or water applied for such special purposes as leaching salts from the soil or for achieving temperature control.

The purpose of agricultural drainage can be summed up as the improvement of soil water conditions to enhance agricultural use of the land. Such enhancement may come about by direct effects on crop growth, by improving the efficiency of farming operations or, under irrigated conditions, by maintaining or establishing a favorable salt regime. Drainage systems are engineering structures that remove water according to the principles of soil physics and hydraulics. The consequences of drainage, however, may also include a change in the quality of the drainage water.

Agricultural drainage is divided into two broad classes: surface and subsurface. Some installations serve both purposes.

Surface drainage. Poor surface drainage conditions exist over large areas of land in the eastern half of the United States and Canada. The condition is caused by the inability of excessive rainfall to move over the ground surface to an outlet or through the soil to a subsurface drainage system. These poor surface drainage conditions are usually associated with soils that have low hydraulic conductivity. Often the soils are very shallow over a barrier such as rock or a very dense clay pan. The impermeable subsoil prevents the water from moving downward and prevents the proper functioning of a subsurface drainage system. Often the land slope is not sufficient to permit the water to flow across the ground surface. In other cases, the areas lack adequate drainage outlets. In order to correct this problem, something must be done to eliminate the depressions and to provide sufficient slope for overland flow. In addition, it is necessary to provide channels to convey the water from the affected area.

The practice of surface drainage may be defined as the diversion or orderly removal of excess water from the land surface by means of improved natural or constructed channels. The channels may have to be supplemented by shaping and grading the land surface so that the water may flow freely into the channel.

In some instances, a subsurface drainage system of pipes is needed in conjunction with surface drains. The effectiveness of the subsurface drains is increased by the removal of the water from the soil surface. As soon as the surface water is removed, the drain pipes can act to lower the water table and to provide a satisfactory environment for the growth of plants. There are essentially five types of surface field drainage systems in common use today: bedding system (**Fig. 1**); random ditch system; interception system; diversion ditch system; and field ditch system. Combinations of two or more systems may be required by circumstances encountered in the field. The choice of a particular system used for surface drainage depends upon the soil type, topography, crops to be grown, and farmer preference.

Subsurface drainage in humid regions. Subsurface drainage is required where a high water table is present. The main purpose of the drainage is to provide a root environment that is suitable for the maximum growth of plants and to sustain yields over long periods of time. One of the main reasons that poor drainage causes a decrease in crop production is the fact that the plant roots have only a limited amount of soil in which to grow. Not only do the plants lack food, but the plant roots suffer from a deficiency of oxygen which is needed for the respiratory processes, as the water that fills the soil pores displaces the air in the soil. Moreover, the water obstructs the gases which are given off by the roots, and some of these gases inhibit plant growth.

The critical need for drainage occurs in the early spring months when the plants are germinating. Lack of drainage retards the normal rise in the soil temperature, decreases plant resistance to disease, and inhibits root development. Poor drainage discourages the growth of aerobic bacteria which are needed to

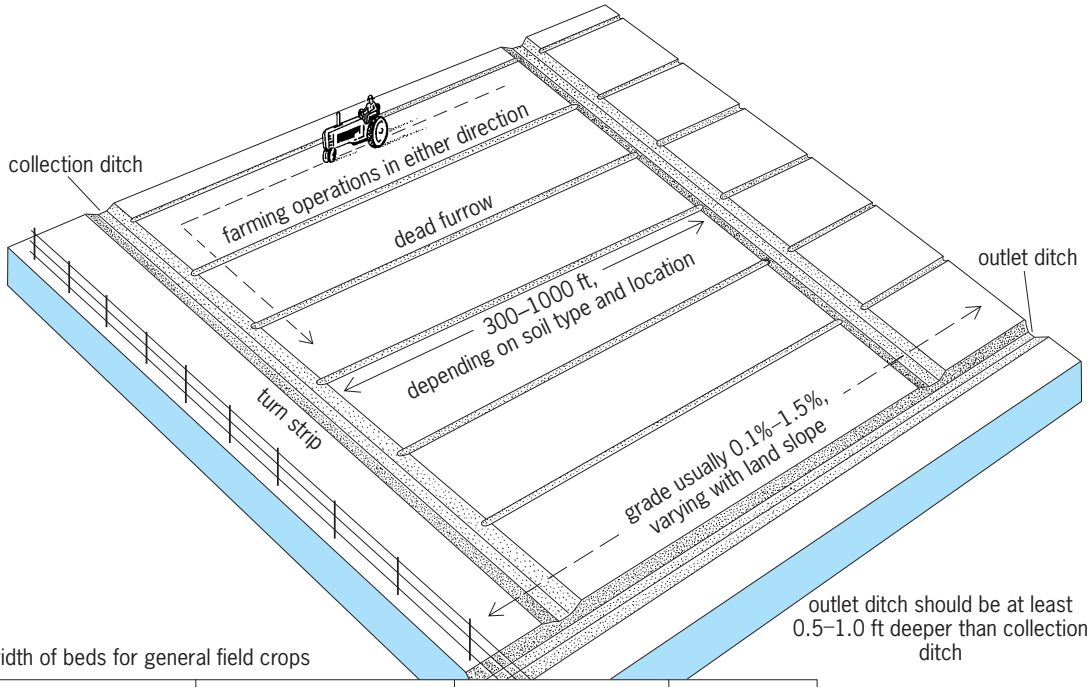

width of beds for general field crops

degree of internal drainage of the soil	width of bed in feet, center to center of dead furrows	number of 3½-ft corn rows with 2 ft allowed per dead furrow	number of rounds using 2–14 in. plows
very slow	23	6	5
	30	8	6½
	37	10	8
slow	44	12	9½
	51	14	11
fair	58	16	12½
	65	18	14
	72	20	15½
	79	22	17
	86	24	18½
	93	26	20

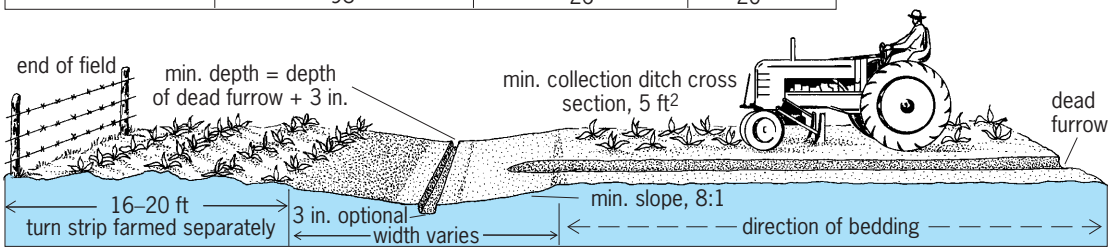

cross section at end of field showing collection ditch and turn strip

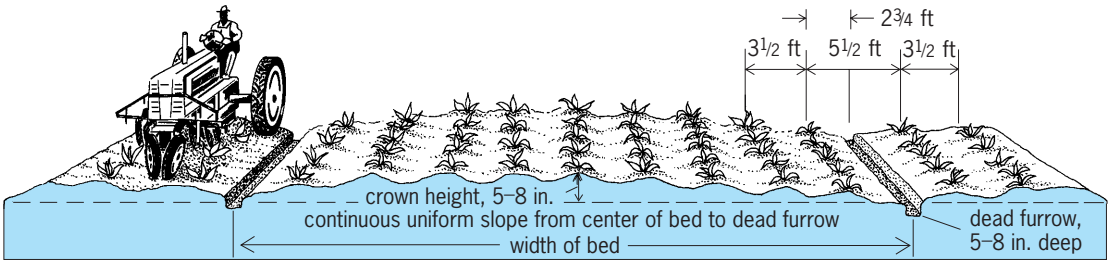

cross section of bed showing crown effect and proper pacing of corn rows

Fig. 1. Surface drainage bedding system. 1 in. = 25 mm; 1 ft = 0.3 m; 1 ft^2 = 0.09 m^2. (*USDA*)

supply nitrogen for crops. Toxic organic and inorganic compounds develop in saturated soils.

The depth of drains in humid regions is largely determined by soil conditions. Drain depths from 24 to 48 in. (60 to 120 cm) are commonly used to control a shallow ground water table. The spacing of the drains depends upon the soil hydraulic conductivity and the amount of the rainfall that must be removed.

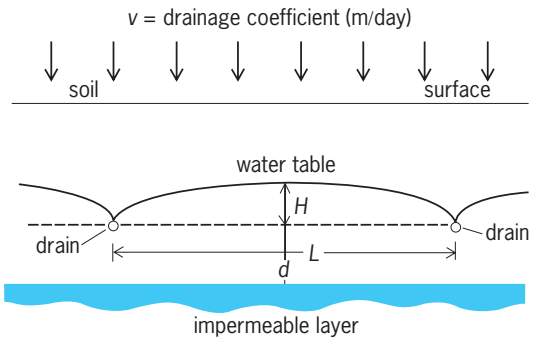

Fig. 2. Hooghoudt's drain spacing formula. Symbols are explained in the text.

Spacings vary between 33 and 100 ft (10 and 30 m).

A number of drain spacing formulas have been proposed and one of the most successful is S. B. Hooghoudt's formula, Eq. (1), where L = drain

$$L^2 = \frac{4K}{\nu}(H^2 + 2dH) \tag{1}$$

spacing (meters), K = hydraulic conductivity (meters/day), ν = drainage coefficient, which is the rate of removal of water from the soil that is necessary to protect the crop from damage (meters/day), H = height of water table above the plant through drains (meters), d = distance from the plane through drains to the impermeable layer. Developed for the drainage of land in the Netherlands, the formula is based upon steady-state replenishment of the ground water. The height of the water table halfway between the drains is calculated as a function of the soil hydraulic conductivity and the rate at which water replenished the ground water table. The significance of the various parameters involved in the formula are presented in **Fig. 2**.

Drainage coefficients are of the order of 5–10 mm/day in Europe. Somewhat higher values (1–4 cm/day) are used in the United States.

Subsurface drainage in arid regions. Irrigation waters contain substantial quantities of salt, from 6 to 240 lb per 1000 ft^3 (0.1 to 4 metric tons per 1000 m^3). Irrigation water is applied at rates of 140,000 to 210,000 ft^3 per acre per year (10,000 to 15,000 m^3 per hectare per year) and hence between 0.05 to 26 tons is added to each acre annually (0.1 to 60 metric tons per hectare). Some of the applied salt precipitates in the soil; a small proportion is used by the plants, and the remainder must be removed from the soil by adding an amount of irrigation water in excess of the crop needs.

The output of salt in the drainage water must equal the input of salt in the irrigation water. If precipitation of salts in the soil and plant uptake of salts are ignored, it can be stated that salt input equals salt output.

The leaching requirement (LR) is defined as the fraction of irrigation water that must be drained in order to maintain the salt balance, as in Eq. (2),

$$\mathrm{LR} = \frac{\nu}{I} \tag{2}$$

where ν represents the drainage coefficient and I the amount of applied irrigation water. The computation of ν depends on the salt tolerance of the crop to be grown.

In arid regions, it is necessary to control the ground water table well below the plant root zone. The plants will extract water from the soil, leaving salt behind. This results in a concentration of salts in the plant root zone; therefore, the water table must be maintained well below the plant root zone so that capillary rise of this salty ground water into the plant root zone is reduced. The depth of the drains then is determined by capillary rise into the plant root zone. Normally drains in irrigated areas are placed about 6 ft (1.8 m) deep. The spacing between the drains may be determined by Hooghoudt's formula.

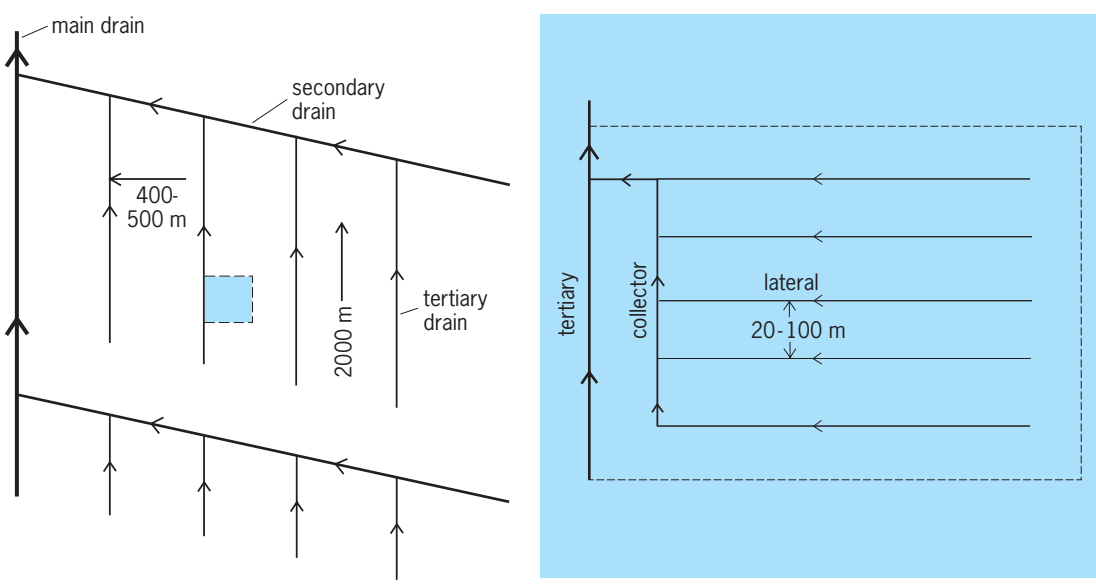

Fig. 3. Typical drainage system for an irrigation project. Shaded box on left is shown in detail on right. 1 m = 3.3 ft.

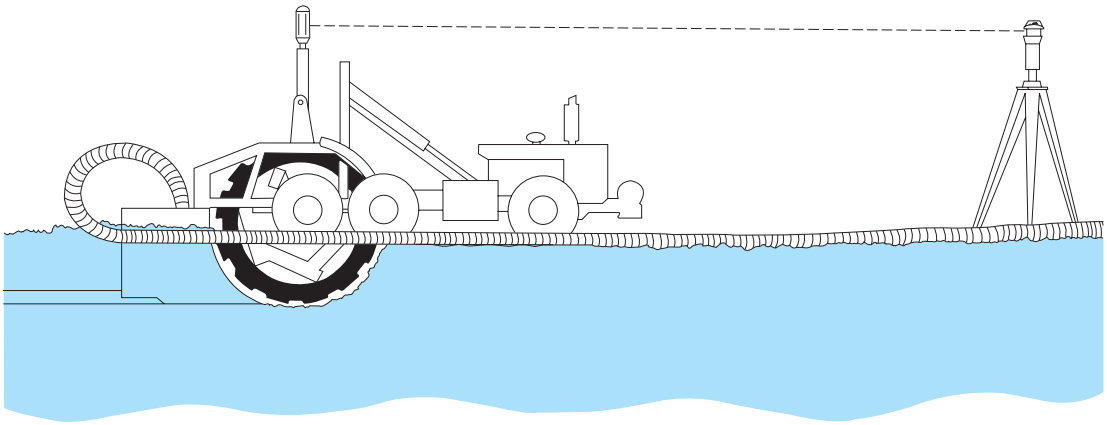

Fig. 4. Grade control by laser beam.

However, the drainage coefficient is determined by a consideration of the leaching requirement, precipitation of salts in soil from irrigation water, amount of leaching due to winter rainfall, and salt tolerance of crops. Typical drain spacings range from 66 to 330 ft or 20 to 100 m (**Fig. 3**).

In addition to the steady-state formula of Hooghoudt, analyses have been made of the transient water table situation. The U.S. Bureau of Reclamation utilizes this transient water table formula for determining the average depth of water table during the growing season. If according to the bureau's method of analysis the water table is rising during the growing season, and continues to rise, the condition indicates a serious drainage problem will occur. The bureau's drainage criteria then are based upon controlling the water table so that it either remains constant or declines over a period. The procedure is called the dynamic equilibrium method.

Construction of drains. The construction of subsurface drains was revolutionized in the years 1976–1979. In the past, a variety of materials such as rocks, clay pipe, concrete pipe, and other materials were used for subsurface drains. With the invention of machines that make perforated corrugated plastic pipe, this pipe largely supplanted the other materials. Corrugated plastic pipe has the advantage of light weight, and it comes in long lengths so the cost of handling is reduced.

Drain pipes either are laid in the bottom of a trench or are pulled into place by a drain plow. The development of laser beams for grade control permits the use of drain plows (called the trenchless method) for installing corrugated plastic tubing (**Fig. 4**).

In unstable soils usually found in arid regions, it is necessary to surround the drain pipe with gravel, sand, synthetic fibers, or organic material in order to prevent fine sands and silts from entering the pipe with the drainage water. *See* AGRICULTURAL SOIL AND CROP PRACTICES; IRRIGATION (AGRICULTURE); PLANT-WATER RELATIONS. J. N. Luthin

Bibliography. G. R. Bumli (ed.), *Principles of Project Formulation for Irrigation and Drainage Projects*, 1982; A. Dinar, *The Economics and Management of Water and Drainage in Agriculture*, 1991; S. C. Harris (ed.), *Irrigation and Drainage*, 1990; E. G. Kruse and C. R. Burdick (eds.), *Environmentally Sound Water and Soil Management*, 1982; J. N. Luthin (ed.), *Drainage Engineering*, rev. ed., 1978.

Land reclamation

The process by which seriously disturbed land surfaces are stabilized against the hazards of water and wind erosion. All seriously disturbed land areas are in need of reclamation and should be stabilized and reclaimed as quickly as possible after disturbance. Disturbance comes from major construction projects such as interstate highway systems, shopping centers, and housing developments, and from surface mining operations for coal, stone, gravel, gold, phosphate, iron, uranium, and clay. Surface mining for coal is responsible for almost one-half of the total land area disturbed in the United States, another one-fourth is from sand and gravel, and the remainder is from mining of other materials and construction. *See* EROSION.

Surface-Mined Areas

The amount of coal recoverable by surface mining is not known since mining technology changes rapidly, and overburden removal is basically a question of economics. The coal required to meet the needs of power production cannot be supplied from deep mining. This means that thousands of acres will be added annually to those already disturbed. In the United States, federal and state laws require that coal exploration and surface coal mining and reclamation operations be conducted in manners that are compatible with the environment and the economic and social needs of each state. The laws also require that reclamation efforts, including but not limited to back filling, grading, top soil replacement, and revegetation of all land disturbed by surface mining activities, shall occur as contemporaneously as practicable with mining operations. Regulations require immediate precautionary measures to be taken to stabilize these severely disturbed areas to avoid serious pollution of the environment.

Surface mining creates many environmental problems. The chemical and physical properties of the resulting spoil is drastically changed and can create a hostile environment for seed germination and subsequent plant growth. However, many of these environmental problems can be overcome or eliminated by proper planning prior to mining and by proper removal and replacement of overburden material during mining operations. Federal and state laws further provide for the control of overland water flow by requiring the construction of temporary diversions as well as permanent diversions to handle peak water flow as needed in order to safeguard surrounding areas from environmental damage. Even with the best mining practices, vegetative cover must be established almost immediately, or the denuded areas will be subject to both wind and water erosion that will pollute surrounding streams with sediment. *See* COAL MINING; MINING; SURFACE MINING; WATER POLLUTION.

In most cases, the spoil material is a conglomerate of rock fragments of various sizes mixed with a small amount of soil. The spoil materials may vary in color from very light to almost black. Because of these color differences alone, temperature and soil moisture may constitute serious problems. Temperature variations created by slope and aspect require use of specific plant species. Mulching is usually essential for germination and seedling establishment. Some spoils are almost sterile and usually require applications of many of the nutrient elements necessary for plant growth. Nitrogen, phosphorus, potassium, calcium, and magnesium deficiencies are major problems on some spoils in the eastern United States, while in other spoils some plant nutrients may be more abundant after mining than before because of weathering of rocks and minerals brought to the surface. Chemical constraints for reclamation in humid areas are low fertility, low pH or high acidity, and toxic concentrations of elements such as aluminum, manganese, and iron. *See* PLANT MINERAL NUTRITION; SOIL CHEMISTRY; WEATHERING PROCESSES.

Arid regions. Under arid conditions of the western United States, excessive amounts of sodium and other elements may constitute serious reclamation problems. High-spodic (heavy, claylike) spoils may require applications of surface soil before good vegetative growth can be accomplished. Under arid and semiarid conditions, lack of plant-available soil water is a factor that ultimately limits plant growth on reclaimed land, just as on unmined land. Practices that increase infiltration, reduce evaporation, and increase potential plant growth generally improve the effectiveness of water conservation and use. Environmental problems of major concern under western arid conditions include salinity levels, exchangeable sodium content, nutrient deficiencies (nitrogen and phosphorus), plant toxicities (magnesium, boron, molybdenum), soil compaction, and steepness of slope.

Alpine and arctic areas. Reclamation and rehabilitation of disturbed land areas in alpine and arctic regions constitute a special problem. These areas are found mainly at high elevations in the western states and the Alaskan arctic tundra area. Examples of land disturbance in arctic and alpine areas involve the Alaskan pipeline and discoveries of large deposits of minerals at high elevations in some of the western states, such as the alpine tundra of the Beartooth Mountain of Montana where chromium and platinum are being surface-mined.

Rehabilitation of these areas is complicated by the short growing season for plants, infertile soils, and the rigorous environments. Some of the greatest problems encountered in reclamation are on areas where total loss of soil has exposed acid-producing pyrites and high concentrations of toxic metals. In some cases, reclamation is aggravated by high winds that remove the exposed soil and sweep away insulating materials, such as snow. *See* TUNDRA.

Topsoil replacement. In the United States the Federal Strip Mine Law requires that topsoil be removed and reapplied on the spoil surface during regrading and reclamation. This practice alone has aided materially in reclamation of surface mine spoil areas throughout the United States. Even when topsoil is reapplied, the surface may contain coarse-textured materials and rock fragments, making it difficult to establish vegetative cover. Many of the eastern mine spoils are derived from sandstone and shales and have a low water-holding capacity. These spoils tend to form crusts and thus create a water-impermeable layer. Practically all of these topsoils have low fertility and thus require extensive fertilization for reclamation and seedling establishment. *See* GROUND-WATER HYDROLOGY.

Placement of overburden. In the eastern United States, many of the surface soils underlined by coal have very low productivity. With proper placement of overburden during mining, use of better regrading techniques, and proper surface management, many surface mine areas can be more productive after mining than before. Some overburden materials contain large amounts of calcium, magnesium, phosphorus, and potassium. If these materials are brought to, or near, the soil surface, weathering may provide a better soil than was there initially.

Use of waste materials. A large number of domestic and industrial waste materials have been tested as amendments on surface mine spoils. These include digested sewage sludges, composted sewage sludge and garbage, tannery waste, bark and fiber mulches, flyash, fluidized-bed combustion waste, and scrubber sludges. Many of these materials contain considerable amounts of essential plant nutrients, but they may also contain some potentially toxic elements. Organic waste materials such as composted sewage and garbage and sewage sludges have been utilized effectively without any detrimental effect. *See* ENVIRONMENTAL ENGINEERING; SOIL.

Wetlands

The term wetlands has been defined in many different ways. In general terms, wetlands are areas that are inundated or saturated by surface or ground

water with sufficient frequency and duration to support vegetation typically adapted for life in saturated soil conditions. Wetlands generally include swamps, marshes, bogs and similar areas. In general, wetlands can be classified as either tidal (including saltwater and brackish water, marshes, intertidal flats, and mangrove swamps) or nontidal (occurring inland and including swamps, bogs, ponds, bottomland, hardwood forests, and areas along streams and lakes).

According to the United States Fish and Wildlife Service, over 50% of the wetlands (approximately 11×10^6 acres or 4.4×10^6 hectares) have been destroyed during the last two centuries. Most of the losses were due to drainage of inland wet areas for agricultural purposes. The development of urban areas accounted for about 8% of the total. Wetlands provide a haven for a multitude of fish and wildlife as well as protection against floods and contaminated water. Coastal wetlands are a vital ecological resource, because they provide shelters and spawning grounds for large numbers of commercially valuable fish. In many land reclamation projects, the value of wetlands is recognized as a vital part of the overall picture. Not only do they provide for spawning and nesting areas for marine and wildlife, but they also provide a natural means of recycling waste materials and cleansing polluted and silt-laden water. A number of communities are taking advantage of the ability of wetlands, both natural and constructed, to treat tertiary wastewater and stormwater.

Preservation. Even though many laws and regulations are in effect to protect wetlands, they still face constant threats from real-estate development and agricultural uses. Developers have practiced converting wetland areas by filling with sand, gravel, soil, and other material to give stable substrate for building sites. This procedure tends to significantly alter adjoining areas by restricting both surface- and ground-water flow. Farmers usually reclaim wetlands for crop production by draining and diking. Despite the laws and programs to protect wetlands, demographic trends, especially along the coastal areas, will continue to increase pressure for development of these areas. Nevertheless, laws and regulations are helping through preservation and a procedure known as mitigation. This term, in general, refers to activities that can avoid or minimize damage to wetlands and, in addition, restore, enhance, or create new wetlands. Thus, mitigation may result in no net loss of wetlands. While the process may reduce wetland losses, mitigation is not without controversy.

Sometimes mitigation involves certain destruction of a wetland site in exchange for uncertain gains. Wetlands involve complex ecosystems of soil, water, and plant and animal life that are difficult to recreate. Nevertheless, mitigation can be successful if appropriate guidelines are followed.

Restoration. Restoration of wetlands, rather than mitigation, is preferred by some states, believing that restoration involves less risk than trying to develop new upland sites. The abused sites can be restored as self-sustaining ecosystems by cutting off pollution discharge, removing fill, and introducing native wetland plants. Restoration can be enhanced by bringing bottom layers of soil and muck from existing wetlands and spreading it on the site to be restored.

Tidal wetlands can be reclaimed by construction canals and opening up access waterways to the main body of water. In many areas this may mean removing some fill dirt to allow free movement of water between high and low tides. This step should be followed by introducing plant species such as salt hay and cord grass in areas along the Atlantic and Gulf coast, and glasswort and Pacific cord grass along the Pacific coast. Many communities along coastal shorelines have established coastal wetlands to control erosion.

When reclaiming nontidal wetlands, a greater variety of plant species can be used, but will depend on type of environmental conditions sought. Freshwater emergent marshes can be seeded with species such as fox sedge, soft rushes, cattail, reed canary grass, and purple loose strife after necessary conditions have been established by removing drainage ways and fill dirt.

Actually, reestablishing wetlands can be accomplished by reversing the process by which they were destroyed. To speed up the reclamation process, it will usually be necessary to remove introduced contaminants and fill. Often marshes, bogs, and swamp areas have been used as dumping grounds in place of approved landfills. Removal of most of this trash is necessary, since it may contain materials and chemicals toxic to both plant and animal life. After cleanup has been accomplished, these areas should be monitored annually by making chemical, physical, and biological evaluations of the soil, water, and plant and animal life. If desirable plant species are not present, they should be introduced as soon as possible by seeding or transplanting once a desirable environment has been established; marine and terrestrial wildlife will follow. Many of the bogs near population centers are in need of cleanup and restoration, especially along the east coast and in Michigan and Wisconsin. Plant species normally adapted to such areas include cranberry, pitcher plant, cattail, Venus' flytrap, evergreen shrubs, and stunted trees. *See* ECOSYSTEM.

Orus L. Bennett

Bibliography. W. H. Armiger, J. N. Jones, Jr., and O. L. Bennett, *Revegetation of Land Disturbed by Strip Mining in Appalachia*, Agricultural Research Service, ARS-Ne-71, 1976; W. Mitsch, *Wetlands*, 2d ed., 1993; D. Salvensen, *Wetlands: Mitigating and Regulating Development Impacts*, 2d ed., 1994; Society of Agricultural Engineers, *Land Reclamation: Advances in Research and Technology*, 1993; R. W. Tiner, *Wetlands Identification, Delineation, and Classification*, 1995; F. W. Schaller and P. Sutton (eds.), *Reclamation of Drastically Disturbed Lands: A Proceedings of a Symposium*, 1978.

Land-use classes

Categories of land resources that provide information about uses and characteristics. In combination with geographic location systems, measurements of the amount of land in use, the amount available for development, and the changes in use over time, decisions about land use can be accurately made. Concerns develop when a certain land quality is required for a specific purpose, such as food production, human living areas, or environmental protection.

Large changes have occurred since the early 1960s concerning land-use classification and natural resource management. Previous concerns were allied to very specific problems such as soil erosion, land availability, desertification, or forest consumption. Technologies and theories now view land-use classification in a larger context and relate it to human use of land-based resources within environmentally acceptable guidelines. Major technological improvements that contributed to these changes include high-resolution films and improved cameras, film processing, lenses, and enlargers. Airplanes, helicopters, and satellites provide platforms for highly developed scanners, return-beam vidicon cameras, radar, lasers, and multispectral retrieval systems. Computerized instrumentation allows merging of numerous sources of data for selected areas of the Earth's surface, including continental or global coverage. Thus, completely new approaches have become feasible, such as the *Land Use Classification System for Use with Remote Sensor Data* prepared by the U.S. Department of the Interior, and previous sources of land-use information have been maintained or enhanced. *See* REMOTE SENSING.

Land is readily classified according to a group of natural characteristics that occur in association. H. E. Conklin developed land classification systems based on soil, climate, topography, and location. Others have preferred to add criteria based on bedrock geology, vegetative cover, subsurface characteristics, landforms, and hydrologic features.

Inventories. Land-use inventories have been encouraged by professional organizations, such as the International Geographical Union. Historically, land-use classification has been based on the concepts of cadastral surveys, which originally included land classification criteria based on quality, quantity, location, and ownership of the land resource. The objective was to provide a basis for equitable taxation. The current objective of land-use classification, however, also includes information for resource management.

Inventories such as the U.S. Department of Agriculture System of Land Capability Classification have centered on the suitability of soils and their limitations for field crop production. The limitations are broadly identified as few, moderate, severe, very severe, impractical to remove, generally too severe for cultivation, very severe limitations for cultivation, or too severe for commercial crop production. Similar major land resource inventories have been developed in Canada, Australia, United States, Netherlands, Denmark, Sweden, England, Switzerland, and other European countries.

Two technologically advanced major inventories, the New York State Land Use and Natural Resource Inventory (LUNR) and the Canada Land Inventory, were developed in the late 1960s for very large areas. The LUNR system provided geographically referenced land-use information on 155 items of information derived primarily from aerial photographs. These can be grouped under the 11 major use headings of agriculture, forest land, water resources, wetlands, residential, commercial and industrial, outdoor recreation, extractive industry, public and semipublic, transportation, and nonproductive uses. Each group is subdivided into 2 to 25 subunits, offering a very detailed classification system for uses of major importance, such as agriculture or residential. Its products included maps, statistical data, computer graphics, and computer printout. The LUNR program sparked activity in land-use classification throughout the United States, and in a few years 32 states developed similar land-use classification systems. The Canada Land Inventory functioned on different concepts, providing highly refined, geographically located in formation on a relatively small number of land-use classification items with computer retrieval capabilities.

Principles. The ideal land-use classification system would provide appropriate identification for all land uses with mutually exclusive units and boundaries clearly defined and easily understood. A thorough description of each unit would describe what the unit contained and why it was placed in its particular classification mode. But natural resources, vegetation, and land patterns influenced by civilization occur in a continuum with transition zones, instead of clearly defined boundaries, between land uses. As a result, generalization of the classification units is required. Because land-based resources are geographically fixed, impurities and mixed units must be included to be comprehensive. The purely theoretical approach may look at vegetative patterns, natural successional systems, and modification of human-impact consequences to maintain the purest possible scientific sequence.

The more common approach is to respond to the request of the sponsor. This usually means forfeiting some of the desired theoretical concepts to maintain cost effectiveness. Steps in this process include: (1) identification of the sponsor's information needs; (2) selection of sources of information; (3) development of a classification system that is comprehensive and provides unique descriptions and discrete assignment; (4) preparation of a geographic referencing system; and (5) selection of appropriate information-retrieval processes.

The classification system should be able to stand alone, independent of other component parts of the inventory project. It should not be dependent on unique talent, special instrumentation, or any one

source of information. Classification and inventory decisions do not require computer or other forms or instrument capability; the use of computers is primarily to store, manipulate, and retrieve information.

Technological development. Development of available technologies has enabled many changes to occur in analytical methodologies. There has been rapid proliferation of methods for analysis of the capability of ecological systems to accommodate particular human land-use needs. These have been accompanied by advances in predictive evaluation of regional landscapes, and have thus carried the process into areas of systems modeling.

Methods for land-use capability analysis have progressed from trial-and-error techniques, which were among the early clearly identifiable regional planning methods of the twentieth century, to inventory methods, characterized by systematic inventories of land uses that combine information from natural sciences with engineering classification of soils and aerial photographic interpretations of cultural land-use patterns. Hand-drawn overlay methods came next to address the complex decisions to be handled at a regional level of planning. Finer distinctions between land uses and resources, and better elucidation of relationships between different types of land use were accomplished by hand-drawn overlay methods than by inventory mapping. Subsequently, data classification systems emerged, as information theory and systems analysis were employed as the organizing elements in data classifications. Computer-based planning methods, including computer-aided assessment procedures, later enabled extensive use of quantitative weighting systems and the sophisticated techniques of derived weights and values.

In the early 1900s in North America, soil and physical resource data were often incomplete and were published in literary form and also at different scales which made comparisons of data very difficult. Later the development of inventory methods was made possible by the considerable increase in knowledge about soils, geology, hydrology, and ecology. Innovations in remote sensing, refinements in soil taxonomy, and enhanced cartographic representations added to the greater availability of natural resource and land-use data. With hand-drawn overlay methods, maps of data developed at a common scale and overlaid are employed in various combinations to create a new level of information: composite maps.

Geographic information systems help serve the interest in finer distinctions between land uses and resources, and the attempts at resolution of conflicts by analysis of relationships between categories of land use. These may be found in the United States' state and regional systems, as well as those in western Europe, Canada, and some developing countries. In the computer era, faster data sources have become more comprehensive, leading to more complex site analysis and land-use planning. Highly advanced landscape planning methods have resulted, which include the use of computer-displayed three-dimensional forms.

In the Netherlands, basic material on the natural environment has been collected over a long period of time. Since 1970, such data has been made suitable for use in behalf of national physical planning policy. A national environmental survey begun in 1972 had as its purpose the elaboration of theoretically demonstrated relationships between community and natural environment for practical situations. This survey and a General Ecological Model (GEM) serve to illustrate the Netherlands' commitment to use of the science of landscape ecology in planning.

A. Lieberman; Ernest E. Hardy

Bibliography. J. R. Anderson et al., *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*, U.S. Government Printing Office, Survey Prof. Pap. 964, 1976; J. G. Fabos, *Planning the Total Landscape: A Guide to Intelligent Land Use*, 1979; E. E. Hardy, *The Design of Natural Resource Inventories*, Resource Information Laboratory, New York State Cooperative Extension, 1979; F. J. Marschner, *Boundaries and Records*, Agricultural Research Service, USDA, 1960; Z. Naveh and A. S. Lieberman, *Landscape Ecology*, 1985.

Land-use planning

The long-term development or conservation of an area and the establishment of a relationship between local objectives and regional goals. Land-use planning is often guided by laws and regulations. The legal basis of land-use regulation in the United States is the police power of a city (or county) to protect the health, safety, and welfare of its residents. The major instrument for current land-use planning is the establishment of zones that divide an area into districts which are subject to specified regulations. Although land-use planning is sometimes done by private property owners, the term usually refers to permitting by government agencies. Land-use planning is conducted at a variety of scales, from plans by local city governments to regulations by federal agencies. The United States has never developed a national land-use plan because land use is considered a local concern.

History. Ancient cities were often designed on rectangular grid patterns and surrounded by protective walls. Modern land-use planning began in the Italian Renaissance during the fifteenth century. Antonio Filarete planned an ideal city, Sforzinda, star-shaped with a radial road pattern. This was an early attempt to organize the land use of an entire city. Early visionaries focused on the form and shape of cities. For example, the inventor Ebenezer Howard introduced the English concept of a "garden city" in the 1890s. In his design, 30,000 people were to live on 1000 acres of land, separated from the city by a greenbelt that was 5000 acres in extent. Homes were laid out in a circular pattern around the city. The design was motivated to overcome the problems of industrial cities. Howard's garden city later gave rise to planned unit developments (PUD), a common tool for developers in suburban areas. Until

the early twentieth century, land-use planning was mostly a physical solution to land-use issues in urban areas.

American land-use planning has an erratic history that draws on a broad variety of disciplines, such as architecture, civic design, parks and open spaces, fire prevention, public health, and housing. An early American planner, Frederick Law Olmsted (1822–1903), was important in the development of land-use planning. His ideas went beyond traditional bounds of shape and form and focused on the functions of different pieces of land, concepts that were a major foundation block in the evolution of zoning. In his plans for the metropolitan area of Boston, he developed some of the first major linear open spaces around the city. He recognized the natural landscape not just as a cultural and esthetic resource but also as an important component in public health. He argued that the closely built walled cities of medieval Europe were responsible for much sickness at that time. Soon after, Charles Eliot planned public open space for Boston around the idea of high-density housing near trolley lines and extensive open spaces in the interstices, along with numerous small squares and parks in the midst of dense populations.

The first comprehensive zoning ordinance in the United States was written by New York City in 1916. It stipulated the regulation of height, area, and use of buildings for public health and safety in various districts. The essence of this ordinance was the designation of different provisions in different areas. In *Village of Euclid, Ohio v. Ambler Realty Co.* in 1926, the U.S. Supreme Court ruled that zoning laws were a valid extension of nuisance laws. The ruling implied that nuisance can be related to the location of a problem rather than just the nature of a problem, such as a pig in a parlor instead of a pig in a pen.

Local-level planning. Most land-use planning occurs at the local level in the United States. American citizens value local autonomy and flexibility in land use. According to one expert, land-use planning in the United States is not a system but 40,000 city and county governments independently administering zoning. Most planning is decided by boards or commissions at the city level, often composed of volunteers or elected officials who are assisted by professional planners. The general public contributes to planning through public hearings and written comments submitted to the planning bodies.

Each local government establishes a general plan which acts as the basic planning document. The general plan indicates the long-term goals for the physical development of the city and any land outside the city's borders that the local government judges to relate to its planning. These plans vary by city and by state. In California, for example, the general plan must have seven elements: a land-use element, a circulation element, a housing element, a conservation element, an open space element, a noise element, and a safety element. The land-use element includes the distribution and intensity of land uses and population density measures. Specific plans are then written to systematically implement the general plan in specific areas. These plans often follow a general process that includes formulation of goals, assembly of data, analysis of data, presentation of results, and implementation.

A major part of local planning is zoning, the division of areas into districts. Zones cover most potential uses, such as residential, commercial, light industry, heavy industry, open space, or transportation infrastructure (such as rail lines or highways). Detailed regulations guide how each zone can be used. For example, residential zoning regulations may cover the height of houses, the distance the houses are placed from the street, or the colors houses can be painted. Open space zoning may include which spaces are designated as picnic areas, as playing fields, as hiking trails, or as strictly wilderness. The placing of different zones relative to each other is increasingly taken into consideration. Zoning has often been used by wealthy communities to promote low-density housing. The use of zoning to exclude certain groups of people has long been criticized.

As a result of pressures from rapid growth, some cities have begun to write growth management plans that limit the pace of growth. Comprehensive city plans aimed to limit the pace of growth have been accepted by the courts. *See* LANDSCAPE ARCHITECTURE.

State-level planning. Very few plans have been undertaken at a statewide scale. In the 1970s, some states developed regional plans to address urgent growth pressures. Each state plan differs by the needs and philosophy of the state. For example, Hawaii introduced statewide zoning to address rapid urbanization of valuable agricultural land; Vermont set up a system of citizen district commissions to administer the state's planning system; and Oregon, in response to large losses of public access to the coast, developed a comprehensive system of local planning in which local plans must conform to specific state goals. The state plans represent a balance of regional structures that address widespread growth with local powers that keep specific decision-making at the local level.

National-level policies. In the United States, very few land-use laws or policies originate at the national level. One of the earliest exceptions was the 1785 Land Ordinance which initiated rectangular surveys in areas west of the Appalachian Mountains. The surveys established 6 mi × 6 mi townships, each with 36 sections of 640 acres (1 mi × 1 mi). The surveys were instrumental in the rapid and orderly land utilization of vast areas. Based on these surveys, the Railway Land Grant gave away over 150 million acres of land to the railroad between 1850 and 1871, mostly in 640-acre sections. These early acts had a major impact on the long-term land use of the United States, and its effects are still emerging in many issues today. For example, the railroad was given every other section in parts of the West, with the remainder staying in public hands. In forested areas, the resulting pattern is that of a checkerboard across vast areas because the railroad's property was sold and logged

by private timber companies, leaving a pattern of clearcuts and mature forest.

Environmental regulations are among the few national-level policies that have direct implications for land-use planning. Four of the major types of environmental laws that impact land-use planning are wetland laws, clean-air laws, clean-water laws, and laws for the protection of endangered species. Wetlands are among the most productive ecosystems, with important roles in flood control, nutrient and pollution removal, and fish and wildlife habitat. Wetlands are defined as those areas inundated or saturated by water at a frequency and duration sufficient to support a prevalence of vegetation typically adapted for life in saturated soil conditions. The U.S. Army Corps of Engineers (ACOE) is responsible for implementing wetland regulations. The ACOE permits development in wetland areas and decides if sites restored to mitigate for wetland losses have been suitably restored. In addition to enforcing wetland laws, the ACOE enforces many regulations that resulted from the Clean Water Act of 1972. The object of this act is to restore and maintain the biological, physical, and chemical integrity of the nation's waters. The Environmental Protection Agency (EPA) oversees standards for clean air that were first established by the Clean Air Act. Although laws to maintain clean air have less direct impact on land use, they have many indirect impacts. For example, the clean-air laws impact what kind of industry can be developed in different places.

The Endangered Species Act of 1974 significantly impacted land-use decisions. It mandates protection of all habitat for endangered and threatened animals and protection of all habitat on public land for endangered and threatened plants. The act has four main components: (1) provisions to list species as threatened or endangered, (2) requirements for federal agencies to consult with the U.S. Fish and Wildlife Service (USFWS) or the National Marine Fisheries Service (NMFS) on federal projects, (3) prohibitions against taking, and (4) provisions for permits to allow incidental take under special circumstances. Species are listed as endangered or threatened if their continued existence is determined to be in jeopardy. Causes for a jeopardy ruling can include threat of habitat destruction, threat from overuse, threat from predation or disease, inadequacy of existing regulation, or other factors that threaten a species' existence. *See* ENDANGERED SPECIES.

In 1982 the act was amended to allow incidental taking of threatened or endangered species. To qualify, the taking must be "incidental to, and not the purpose of, the carrying out of an otherwise lawful activity." The amendment applies to private landowners and nonfederal entities, provided they develop agreements known as habitat conservation plans. These plans agree to minimize and mitigate any taking. These agreements are quickly becoming a prominent mechanism to address threatened and endangered species on private land and had a major impact on land use in the 1990s. Over 200 were approved during the 1990s, spanning scales from private homeowners with less than an acre of land to agencies of state governments that have received permits to use thousands of acres of timberland.

From city to landscape. Land-use planning, in large part, has focused on urban planning. Most texts on the subject dwell on urban plans, zoning (especially residential zoning), and laws implementing zoning at a local level. Increasingly, land-use planning is done at larger scales and involves multiple issues. Awareness of environmental concerns, coupled with the wide availability of technical tools that include digital maps at all scales, has led to new approaches to land-use planning. These approaches often use ideas from landscape ecology in the planning process.

Landscape ecology embraces large, heterogeneous areas. It began in the 1950s as a field of study that combined natural history and physical patterns of large areas with aerial photographs. Four general areas of landscape ecology have direct implications for land-use planning: patches; edges, boundaries, and fragmentation; buffer zones; and corridors and connectivity. *See* ECOLOGY, APPLIED; LANDSCAPE ECOLOGY.

Three features of patches are important in land-use planning: their size, location, and abundance. In general, larger patches are better than smaller patches in maintaining wildlife and vegetation. Similarly, a greater number of intact natural patches is helpful. Development of patches of natural areas may result in loss of rare species and fewer kinds of habitats. It is also better to keep distances between patches short. For the land-use planner, this impacts decisions about which areas to develop. For example, land-use plans for a housing development might consider which areas to develop and which natural areas to leave intact in terms of where natural habitats remain and the size of individual habitat patches, rather than considering only the total acreage of native habitat in the vicinity.

Patch edges often have a different kind of habitat than the middle of a patch. Often this boundary has a negative impact on species that live in the middle of a patch. One reason that larger patches are better for wildlife is that a smaller percentage of the habitat comes in contact with developed areas. Similarly, patches that are more circular in shape maintain more interior habitat than those that have a less regular shape. For this reason, when habitats are fragmented by development into odd shapes and thin strips, little interior habitat area may remain. Also, abrupt edges may have a more negative influence on wildlife than less distinct edges. For example, an abrupt edge may reduce soil erosion and decrease habitat diversity. Thus, land-use plans that impact wildlife areas should consider the shape of patches and boundaries between natural areas and developed areas when designing specific plans.

In natural areas with a high sensitivity to development, a buffer zone provides a protected area between a sensitive area and a developed area. Buffer zones are areas that land-use planners can designate for low-intensity activities, such as parks and recreation.

Many ecological processes are influenced by the connections between patches of natural habitat. Animals may move freely in corridors of native habitat or may be inhibited by large barriers such as interstate highways. In some terrestrial cases, stepping stones of similar habitat may provide enough territory to provide connectivity. Similarly, a stream and its banks must be continuous, not broken, for maintenance of an aquatic community. The connection between sensitive wildlife patches is particularly important in land-use planning because planners must consider how to leave corridors of open space intact and unbroken by development. Cheryl Schultz

Bibliography. J. B. Cullingworth, *The Political Culture of Planning: American Land Use Planning in Comparative Perspective*, 1993; D. J. Curtin, Jr., *California Land Use and Planning Law*, 1994; W. E. Dramstad, J. D. Olson, and R. T. T. Forman, *Landscape Ecology Principles in Landscape Architecture and Land-Use Planning*, 1996; J. G. Fabos, *Land-Use Planning: From Global to Local Challenge*, 1985.

Fig. 1. Landing gear arrangements. (*a*) Tail-wheel arrangement. (*b*) Tricycle arrangement.

Landing gear

That portion of an aircraft consisting of the wheels, tires, brakes, energy absorption mechanism, and drag brace. The landing gear is also referred to as the aircraft undercarriage. Additional components attached to and functioning with the landing gear may include retracting mechanisms, steering devices, shimmy dampers, and door panels.

The landing gear supports the aircraft on the ground and provides a means of moving it. It also serves as the primary means of absorbing the large amounts of energy developed in the transition from flight to ground roll during a landing approach. The brakes, normally located in the main wheels, are used to retard the forward motion of the aircraft on the ground and may provide some control in the steering of the aircraft. In most modern aircraft the landing gear is designed to retract into the aircraft so that it is out of the airstream and drag is thus reduced.

Types. Early aircraft and many small aircraft use a tail-wheel (or skid) in a conventional, or tail-dragger arrangement, in which the main landing gear is located ahead or forward of the center of gravity of the aircraft (**Fig. 1***a*). The popular arrangement on modern aircraft is a tricycle landing gear, with the main gear located behind or aft of the center of gravity, and a nose gear located forward which carries about 20% of the static weight of the aircraft (Fig. 1*b*). Other landing gear arrangements include four wheels, with two wheels on either side of the aircraft in a tandem arrangement; tandem wheel configurations, with the gear mounted on the centerline of the aircraft; single-wheel gear, such as that used for sailplanes; and skid-type gear, often used on helicopters. Skis or floats may be used in lieu of wheels. *See* GLIDER; HELICOPTER; SEAPLANE.

Large aircraft such as the wide-body commercial aircraft and military aircraft like the C-5A employ multiple-wheeled bogies to support their huge weight and, in the case of the C-5A, to provide soft terrain landing and takeoff capability. The bogie landing gear arrangement on the XB-70 aircraft (**Fig. 2**) employs a unique braking system using a small sensing wheel, located between the main tires, to sense impending skid due to braking by measuring the ground speed compared with the wheel speed of the main tires. Several aircraft also employ outrigger landing gear on the wing tips for roll stability. *See* MILITARY AIRCRAFT.

Design considerations. Landing gear design parameters are weight, vertical velocity of the aircraft at gear contact, landing approach speed, braking distance, ground clearance, lift angle requirement, turnover angle (related to the distance between the two main gears and also the distance between the nose gear and main gear), the condition of the landing surface, and whether the landing is on the ground or an aircraft carrier. Once the basic geometric relationships have been defined, the most important design consideration is determining the amount of energy that must be absorbed by the landing gear to prevent overloading the aircraft structure. The ideal situation is to match the loads induced into the structure to those induced by the various flight conditions.

Oleo strut. The most accepted method of absorbing the energy due to landing is an air-oil strut called an oleo. The basic components are an outer cylinder which contains the air–oil mixture and an inner piston that compresses the oil through an orifice. The flow of oil through the orifice is metered by a variable-diameter pin that passes through the orifice as the gear strokes. The flow of oil in effect varies the stiffness of the compression of the gear. Because of the flexibility of the aircraft and gear and the smoothness and strength of the landing surface

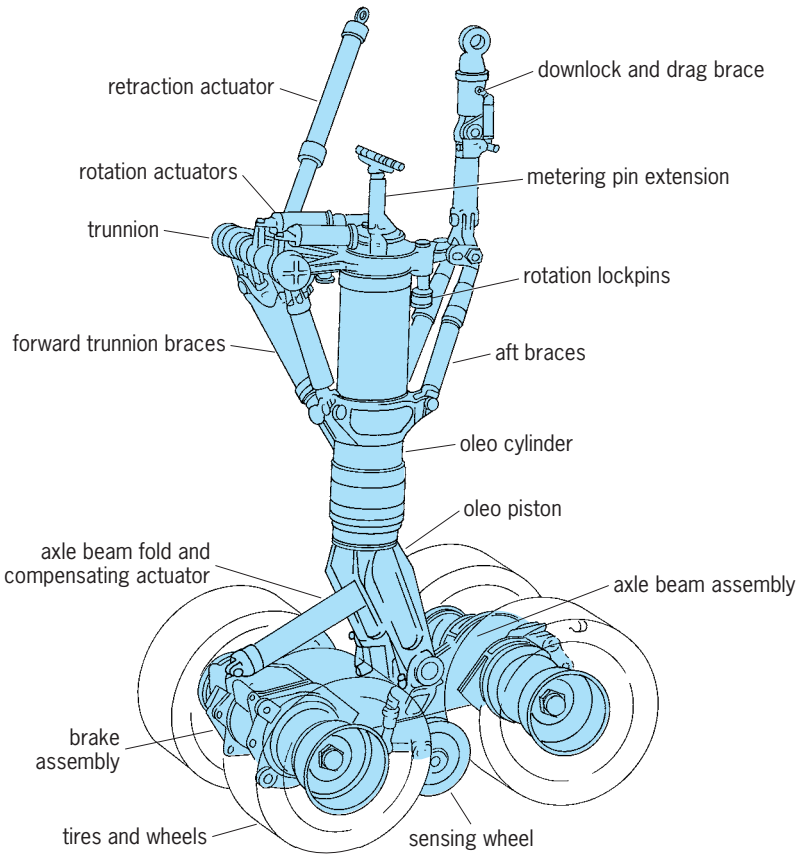

Fig. 2. Main-wheel bogie for the XB-70A aircraft. (*After S. Pace, North American Valkyrie XB-70A, Aero Series vol. 30, Tab Books, 1984*)

metering is necessary to establish the most efficient load–stroke relationship. Thus the overall stroke of the gear is minimized, and the complexity, flexibility, and strength requirements are optimized.

Special requirements. Naval aircraft must be able to land aboard and take off from aircraft carriers. The sink speeds can be as high as 24 ft/s (7.3 m/s), compared with 10 ft/s (3.0 m/s) for Air Force aircraft and 3 ft/s (0.9 m/s) for most commercial aircraft. The high energy generated requires long gear stroking and special metering characteristics. A secondary metering orifice is often required for arrested landings, particularly free-flight engagements. A free-flight engagement occurs when the arresting cable is engaged prior to the aircraft wheels touching down. Very high energy is imparted to the nose gear. The aircraft comes to a stop within approximately 310 ft (100 m). This results in high longitudinal load factors (4–5 *g*) on the aircraft.

The nose gear is specially designed for catapult launches. The design includes a retractable tow bar which drops over the catapult shuttle into a notch. The tow bar is attached to the nose gear strut and transmits the launch force from the catapult to the aircraft through the trunnion and a drag brace. The tow bar also serves to align the aircraft as it approaches the catapult by engaging a track on the carrier deck. A holdback bar trails behind the nose wheel and drops into a slot in the deck. A holdback coupon with a predetermined breaking load based on aircraft weight and catapult characteristics connects the holdback bar to the holdback fitting. This allows the launch load to build, imparting a step acceleration to the aircraft. The longitudinal acceleration during catapulting can exceed 5 *g* with corresponding loads in the launch bar over 250,000 lbf (10^6 newtons). The length of the power stroke of the catapult is approximately 250 ft (80 m), with flying speed being attained in this distance. Dual nose wheel tires are required so that the catapult shuttle can be straddled and maintain symmetric load on the trunnion. *See* NAVAL SURFACE SHIP.

Aircraft that operate from unprepared fields also have special requirements. Large, low-pressure tires are needed to provide a high flotation capability. Since the surface may be rough and contain obstacles such as rocks and logs, the gear is likely to employ additional energy absorption mechanisms. A trailing-arm (articulated) gear is used on the OV-10A aircraft, allowing it to land in soft fields with high sink rates. Active landing gear designs employ computer technology within the gear to sense the rate of onset of load and actively meter the flow of oil to compensate and maintain an optimum load level. *See* AIRPLANE.

Ralph R. Deakins

Bibliography. K. Connes, *Know Your Airplane*, 1986; N. S. Currey, *Aircraft Landing Gear Design: Principles and Practices*, 1988; W. G. Holder, *The B–1 Bomber*, Aero Series vol. 35, 1988; S. Pace, *North American Valkyrie XB70A*, Aero Series vol. 30, 2d ed., 1989; H. Seo, *Boeing 747*, 1984; D. B. Thurston, *Design for Flying*, 2d ed., 1994.

Landscape architecture

The art and profession of designing and planning landscapes. Landscape architects are concerned with improving the ways in which people interact with the landscape, as well as with reducing the negative impacts that human use has upon sensitive landscapes. The history of landscape architecture reaches from the gardens and outdoor environments of ancient civilizations to a broad base in environmental design in the twentieth century. Today, landscape architects are involved in such diverse areas as landscape and urban design, community and regional planning, interior and exterior garden design, appropriate and high technology, agricultural and rural land-use planning, parks and recreation, historic site and natural area preservation, landscape restoration and management, research and academic programs, energy and water conservation, and environmental planning. Landscape architects find employment with private landscape architecture, engineering, and planning firms; public agencies, including city, county, and state design, planning, park, and transportation department; federal agencies, including the U.S. Forest Service, Corps of Engineers, National Park Service, Bureau of Land Management, and Fish and Wildlife Service; and universities and colleges.

The American Society of Landscape Architects

(ASLA) was established in 1899 as the primary informational and lobbying organization for the profession. Over 80% of the states in America have licensure for landscape architects through state consumer, health, and safety regulation boards. There are over 40 accredited college programs in landscape architecture throughout the United States which provide educational training, enabling students to take licensing exams without lengthy apprenticeships. Degrees in landscape architecture are offered at both the bachelor's and master's levels. Many countries have universities with degree programs in landscape architecture, and abundant international employment opportunities exist. There are landscape architects in practice in every country of the world.

Frederick Law Olmsted, the nineteenth-century designer of Central Park in New York City, Yosemite National Park in California, and many gardens and parks throughout the country, is generally considered to be the father of landscape architecture in America. When Olmsted is added to the long list of historical figures associated with the profession, including the eighteenth-century British landscape architect Humphery Repton and the seventeenth-century French landscape architect Andre Le Notre, the form and historical intent of the designed and planned landscape throughout the world can be better understood.

The skills required for landscape architecture are concerned with creativity, ecology, social science, horticulture, engineering, and graphics. Landscape architects are generalists in that their educational and professional experience is very broad. Many environmental and cultural factors affect landscape design and planning, and landscape architects have to know how these factors relate. Design process is the main area of specialization for landscape architects, and decision-making related to design process is the fundamental reason why landscape architects are employed. The ability to utilize the information of specialists and to synthesize the information into a usable format remains the essence of the profession. *See* CIVIL ENGINEERING; ENVIRONMENTAL ENGINEERING; FORESTRY, URBAN; LAND-USE PLANNING.

Kerry J. Dawson

Bibliography. S. Cantor, *Landscape Architecture*, 1992; T. Higuchi, *The Visual and Spatial Structure of Landscapes,* 1988; J. E. Ingels, *Landscaping Principles and Practices*, 5th ed., 1997; T. Turner, *Landscape Planning,* 1987.

Landscape ecology

The study of the distribution and abundance of elements within landscapes, the origins of these elements, and their impacts on organisms and processes. A landscape may be thought of as a heterogeneous assemblage or mosaic of internally uniform elements or patches. Typically, landscape ecologists view the world as one would see it from an airplane (see **illus.**). At this scale, examples of patches include remnant blocks of forest, agricultural fields, and housing subdivisions. Many fields of study focus on large-scale patterns. For example, biogeographers, land-use planners, hydrologists, and ecosystem ecologists are concerned with patterns and processes at large scale. However, landscape ecologists bridge these disciplines in order to understand the interplay between the natural and human factors that influence the development of landscapes, and the impacts of landscape patterns on humans, other organisms, and the flows of materials and energy among patches. Much of landscape ecology is founded on the notion that many observations, such as the persistence of a small mammal population within a forest patch, may be fully understood only by accounting for regional as well as local factors.

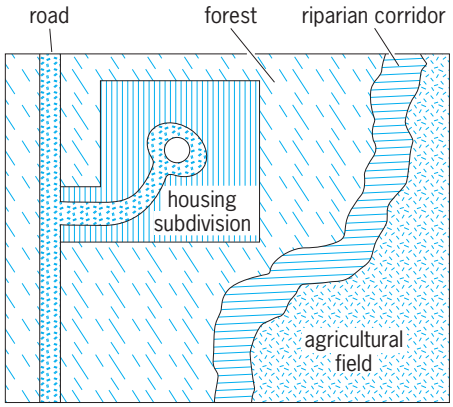

Representative landscape as viewed from above.

Although landscape ecology has become a prominent discipline only recently, it originated over 50 years ago in Europe. Important studies in landscape ecology have shown how the sizes of forest patches have declined in England over the last several centuries, that the fires in Yellowstone National Park in 1988 were probably not unusually intense, and that the dispersal movements of owls can be used to develop better methods of quantifying the distribution of habitat within a landscape. One way of understanding the field of landscape ecology is to divide it into three parts: documenting landscape pattern, understanding the development of landscape pattern, and examining the consequences of landscape pattern.

Documenting landscape patterns. A primary task of landscape ecologists is to describe and measure the attributes of landscapes. Landscape ecologists use a wide variety of measures and often employ high-technology tools such as remotely sensed information from Earth-orbiting satellites to obtain the raw data used when documenting landscape patterns. These large spatial datasets are often analyzed utilizing geographic information systems (GIS) on high-speed computers.

Landscapes are defined in terms of patch types. Deciding how to define patch types is a critical step in any landscape ecological study. Will all vegetation be represented as a single patch type, or will forests

be differentiated from grasslands? Will forests be divided into coniferous (such as pines) and deciduous (such as oaks)? Will forests of different ages be considered as different patch types? The answers to such questions have an overriding effect on the landscape pattern observed and must be closely linked to the researcher's goal. Once patch types have been determined, frequently used measures of landscapes include the number of patches, their size, and their isolation from other patches. Landscape ecologists also have used a number of measures that reflect the shape of patches. The ratio of patch perimeter to patch area increases as patches proceed from globular to sinuous in their shape. The fractal dimension of a landscape is a measure that can quantify the tendency of very complex patches to fill the space within a landscape.

Development of landscapes. Often, landscape attributes have been quantified as part of an effort to understand the factors that led to the development of the landscape pattern. These factors include a combination of human and nonhuman agents that interact to give landscapes their shape. The geology of a region, including the topography and soils along with the regional climate, is strongly linked to the distribution of surface water and the types of vegetation that can exist on a site. These factors influence the pattern of human settlement and the array of past and present uses of land and water. Thus, fertile lowlands once dominated by floodplain forests have often become agricultural bottomlands. Rocky uplands once sparsely covered with drought-resistant plants have often been used for livestock grazing or the extraction of mineral resources.

In many cases, the role of ancient events can still have strong signatures on the shape of landscapes. Throughout the northern regions of Europe, Asia, and North America, glaciers have repeatedly carved the landscape, eradicated vegetation, and forced plants and animals to reinvade from southern refuges. One consequence of such ice ages has been the prevalence of fresh-water lakes and ponds in glaciated regions.

Acknowledging the important role of geology and climate, there can be little doubt that the landscapes of the world have been dramatically and rapidly transformed in recent centuries. A recurrent theme in landscape ecology is the role of humans in shaping the development of the landscape pattern. One prevalent effect is habitat fragmentation, which arises because humans tend to reduce the size and increase the isolation among patches of native habitat. The clearing of land for human settlement, timber, or agriculture tends to replace irregular, contiguous mosaics of native vegetation with geometric fragments surrounded by a matrix of human-dominated patch types. Much of the attention devoted to habitat fragmentation focuses on the initial clearing of land, such as that occurring in many of the world's tropical forests. However, habitat fragmentation can continue for long periods after this initial transition.

Consequences of landscape patterns. Just as many factors act together to lead to the pattern of patches on a landscape, these patterns can have direct effects on many different processes. Human modification of landscapes results in shapes of landscapes which constrain future human uses. The structure and arrangement of patches also can affect the physical movement of materials such as nutrients or pollutants and the fate of populations of plants and animals. Many of these impacts can be traced to two factors, the role of patch edges and the connectedness among patches.

The boundary between two patches can be important by either hindering or promoting movement. In particular, patch boundaries often act as filters or barriers to the transport of biological and physical elements. As an example, leaving buffer strips of native vegetation along stream courses during logging activities can greatly reduce the amount of sediment and nutrients that reach the stream from the logged area.

The part of a patch near a boundary, or its edge, is exposed to materials and organisms that cross that boundary. Edge effects can result when forests are logged and there is a flux of light and wind into areas formerly located in the interior of a forest. In this example, edges can be a less suitable habitat for plants and animals not able to cope with drier, high-light conditions. When habitats are fragmented, patches eventually can become so small that they are all edge. When this happens, forest interior dwellers may become extinct.

The connectivity of landscapes is directly related to the influence of patch boundaries on the flows among patches. Many species that live in a particular patch type are not easily able to move through other patch types. When patch boundaries act as barriers to movement, they can have pronounced effects on the dynamics of populations within and among patches. In the extreme, low connectivity can result in regional extinction even when a suitable habitat remains. This can occur if populations depend on dispersal from neighboring populations. Habitat fragmentation tends to make it harder for animals to successfully disperse. Distances get longer and the matrix between patches is increasingly unsuitable. Eventually, when a population becomes extinct within a patch, there is no way for a colonist to reach the vacant habitat and reestablish the population. This process is repeated until all of the populations within a region disappear. Landscape ecologists have promoted the use of corridors of native habitat between patches to preserve connectivity despite the fragmentation of a landscape.

An overriding theme in landscape ecology is the role of context. The growth rate of a tree or the likelihood that the tree will survive the winter may depend on the size of the forest patch in which it lives and its distance from the edge of the forest. The presence of a salamander species within the same forest patch may depend on the distance to the nearest neighboring forest patch and whether the intervening space is covered with an abandoned agricultural field or a parking lot. Because landscape ecology focuses on large scales and readily incorporates the

role of both humans and natural factors, it has become an important part of conservation efforts: it forces planners to carefully consider their actions across an entire landscape. *See* LAND-USE PLANNING.

David Skelly

Bibliography. R. T. T. Forman, *Land Mosaics: The Ecology of Landscapes and Regions*, Cambridge University Press, Cambridge, 1995; L. Hansson, L. Fahrig, and G. Merriam (eds.), *Mosaic Landscapes and Ecological Processes*, Chapman and Hall, London, 1995.

Landslide

The perceptible downward sliding, falling, or flowing of masses of soil, rock, and debris (mixtures of soil and weathered rock fragments). Landslides range in size from a few cubic yards to over 3.5×10^9 ft^3 (1×10^9 m^3), their velocities range from an inch or so per day to over 300 ft/s (100 m/s), and their displacements may be about an inch to several miles. *See* MASS WASTING.

Types. As with other natural phenomena, difficulties arise in classifying landslides because of the complexity of the processes involved and the limited understanding of these processes. The U.S. Highway Research Board classification is widely accepted, and divides landsliding of rock, soil, and debris, on the basis of the types of movement, into falls, slides, and flows (**Fig. 1**). Other classifications consider flows, along with creep and other kinds of landslides, as general forms of mass wasting.

Falls. This type of landsliding occurs when soil or rock masses free-fall through air. Falls are usually the result of collapse of cliff overhangs which result from undercutting (Fig. 1*a*) by rivers or simply from differential erosion. In temperate and cold regions, ice wedging and roots open cracks parallel to cliff faces, inducing rock slabs, sometimes over 300 ft (100 m) high, to topple and fall (Fig. 1*b*). If these rockfall events are relatively small and piecemeal, the result is an apron of broken rock at the foot of the cliffs known as talus. Talus slopes characteristically assume angles with the horizontal between 30 and 40°.

Slides. Landslides of this type invariably involve shear displacement or failure along one or more narrow zones or planes. Internal deformation of the sliding mass after initial failure depends on the kinetic energy of the moving mass (size and velocity), the distance traveled, and the internal strength of the mass.

Some slides deform very little and move for relatively small distances in comparison with their total size. A slump is a type of these relatively undeformed slides in which the sliding surface or surfaces are spoon-shaped (Fig. 1*c*); in a glide the surface of sliding is planar (Fig. 1*d*). In slumps, the shear plane occurs through intact material; hence, slumps are more common in soil than in rock. In glides, the sliding surfaces are preexisting fractures or weakness planes such as joints, bedding, fault, and foliation planes. Glides are usually associated with rock masses.

Masses of soil, debris, and rock that undergo large deformation during displacement create soil slides, debris slides, and rockslides. Rockslides (Fig. 1*e*) usually involve large ($0.35–35 \times 10^6$ ft or $0.01–1 \times 10^6$ m^3 or more) and extremely large ($0.35–35 \times 10^9$ ft or $0.01–1 \times 10^9$ m^3 or more) rock masses moving for relatively large distances at very high velocities. The most devastating landslides in history

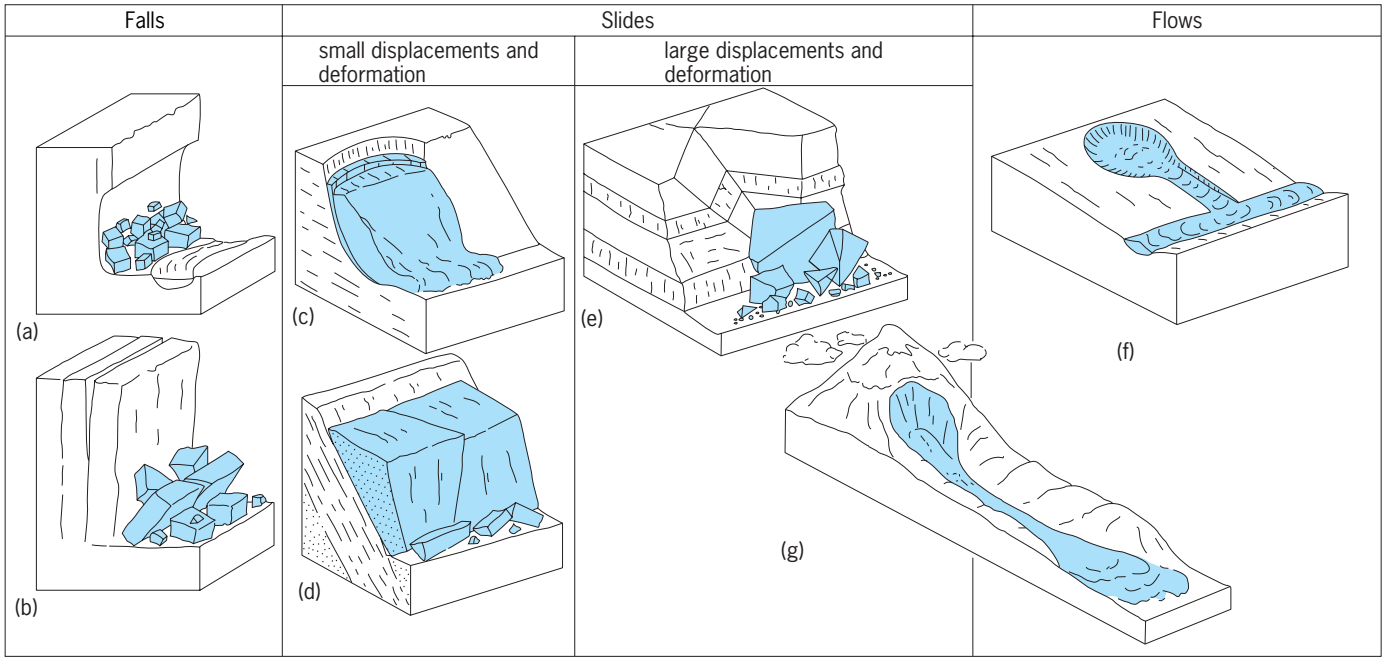

Fig. 1. Types of landslides: (*a*) fall (by undercutting); (*b*) fall (by toppling); (*c*) slump; (*d*) glide; (*e*) rockslide; (*f*) earthflow; and (*g*) rockslide that develops into flow (avalanche). Dimensions may vary considerably from one type to another. The materials involved are rock, soil, and debris.

began as rockslides. A landsliding event may contain more than one type of slide. Thus a rock glide may deform more at its toe and develop a rockslide there. Slumps often exhibit a flowage of material at their toes.

Flows. The internal displacement and the shape of the so-called flows resemble those of viscous fluids,the husaalthough some of them may contain negligible amounts of water. Slip surfaces are not discernible either within the mass or at its base (Fig. 1*f*). Relatively weak and wet masses of shale, weathered rock, and soil may move in the form of debris flows and earthflows; water-soaked soils or weathered rock may displace as mudflows. Some very wet flows may attain large sizes and high velocities and may be able to engulf and displace large rock blocks or human-made structures for miles down their paths.

When extremely large ($0.35–35 \times 10^9$ or $0.01–1 \times 10^9$) masses of rock, ice, and snow tear off the side of a high mountain, the movement is first of the rockslide type (Fig. 1*g*). In some instances, these collossal landslides may be temporarily airborne and hence become rockfalls. In many cases, the gigantic moving masses break down into progressively smaller fragments and engulf snow, ice, and mud to form mixtures that flow very rapidly for long distances and are known as debris avalanches. Some rock avalanches contain very little water. The large runout of relatively dry rock avalanches has been explained as a function of the cushioning effect of a layer of pressurized air trapped at the sole of the mass. The apparent fluidity and the ability of these large rockslides and avalanches to conform to the topographic irregularities of their path, however, may be simply a result of a lack of internal stiffness which results from the breakdown of the original rock mass into progressively smaller blocks until their size is dwarfed by the dimensions of the entire moving mass. Internal collisions allow the transmission of kinetic energy from block to block throughout the mass and create a grain-type of flow in a dispersion, similar to the flow of dry sand. Water vaporization by frictional heat at the base of large rockslides and avalanches can provide a very effective means of reducing base friction and promoting large and violent runouts.

The world's best-known and most devastating landslides have been of the rockslide/avalanche type. In the Frank rockslide of 1903 in southern Alberta, 1×10^9 ft^3 (30×10^6 m^3) of limestone moved along a bedding plane dipping toward the valley, destroyed the southern end of the town of Frank, continued up the opposite side of the valley, and came to rest 20 ft (6 m) above the valley floor. The Gros Ventre rockslide/avalanche of 1925 in northwestern Wyoming was the largest landslide in North America. It involved 1.4×10^9 ft^3 (40×10^6 m^3) of poorly cemented sandstone and was caused by a combination of the progressive leaching of the cement in the sandstone, pore-water pressures caused by heavy rains, and one or more earthquake shocks. The Madison Canyon rockslide (**Fig. 2**) of 1959 in Montana (near Yellowstone Park), and the Hope rockslide of 1965 in the Cascade Mountains of southern British Columbia

Fig. 2. Madison Canyon rockslide showing slide scar on left wall and slide debris damming the Madison River, forming new lake, in foreground. (*J. R. Stacy*, *USGS*)

were also triggered by earthquake shocks. The Huascarán avalanches of 1962 and 1970 in central Peru, were among the most catastrophic natural disasters in historical times. In both events, portions of the side of Mount Huascarán tore off and moved approximately 16 km along a path of destruction. The 1962 event (volume approximately 350×10^6 ft^3 or 10×10^6 m^3; velocity approximately 36 mi/h or 60 km/h) buried the town of Ranrahirca and killed 4000 people. The 1970 event (volume approximately 3.5×10^9 ft^3 or 100×10^6 m^3); velocity approximately 102 mi/h or 170 km/h) was earthquake-triggered, covered the town Yungay under 30 ft (10 m) of mud and rock blocks, and claimed 18,000 lives.

Landslides and engineering. Mining and civil engineering works have induced myriads of landslides, a few of them of a catastrophic nature. Open-pit mines and road cuts create very high and steep slopes, often quite close to their stability limit. Local factors (weak joints, fault planes) or temporary ones (surges of water pressure inside the slopes, earthquake shocks) induce the failure of some of these slopes. The filling of reservoirs submerges the lower portion of natural, marginally stable slopes or old landslides. Water lowers slope stability by softening clays and by buoying the lowermost portion, or toe, of the slope. Slope toes often resist and retain the unstable upper portions of slopes; thus lighter (buoyant) toes can be easily pushed toward the reservoir.

Two well-known landslides that occurred in Europe were triggered by human activities. The Elm avalanche of 1881, which destroyed part of the town of Elm in the Swiss Alps and killed 115 people, was triggered by slate-quarrying activity at the foot of a high mountain side. The Vaiont rockslide of 1963 in northern Italy was even more catastrophic. Filling of the Vaiont dam reservoir induced the violent sliding of 8.5×10^9 ft^3 (240×10^6 m^3) of sedimentary rocks into the reservoir; the reservoir water was displaced by the slide material, overtopped the dam (which withstood the event), and rushed down Vaiont gorge to the Piave valley, where it claimed 2600 lives.

Advances in soil and rock engineering have improved the knowledge of slope stability and the

mechanics of landsliding. Small and medium-sized slopes in soil and rock can be made more stable. Remedial measures include lowering the slope angle, draining the slope, using retaining structures, compressing the slope with rock bolts or steel tendons, and grouting. Engineering geologists often are able to recognize areas with marginally stable slopes; engineering activity in such areas is then avoided, or proper measures are undertaken. One exception to the ability to recognize a potentially disastrous landslide was the mud flow that killed 128 workers at the Guavio hydroelectric project, east of Bogotá, Colombia, on July 28, 1983. The mud flow moved at great speed, involved colluvium and highly weathered rock (saprolite), and was preceded by record rains. Less catastrophic slides have continued to take place in California and other slide-prone areas around the world, chiefly in response to heavy rains. *See* ENGINEERING GEOLOGY; EROSION; SOIL MECHANICS.

Alberto S. Nieto

Bibliography. M. J. Crozier, *Landslides: Causes, Consequences and Environment*, 1986; R. L. Schuster and R. J. Krizek (eds.), *Landslides: Analysis and Control*, NAS-NAC Spec. Rep. 176, 1978; E. F. S. Sharpe, *Landslide and Related Phenomena*, 1938, reprint 1960; J. E. Slosson, A. G. Keene, and J. A. Johnson (eds.), *Landslides: Landslide Migration*, 1993; B. Voight (ed.), *Rockslides and Avalanches*, pt. 1, 1978, pt. 2, 1980.

Langevin function

A mathematical function which is important in the theory of paramagnetism and in the theory of the dielectric properties of insulators. The analytical expression for the Langevin function (see **illus.**) is shown in the equation below. If $x < 1$, $L(x) \cong x/3$.

$$L(x) = \coth x - \frac{1}{x}$$

The paramagnetic susceptibility of a classical (nonquantum-mechanical) collection of magnetic dipoles is given by the Langevin function, as is the polarizability of molecules having a permanent electric dipole moment. In the quantum-mechanical treatment of paramagnetism the Langevin function is replaced by the Brillouin function. For further discussion *see* PARAMAGNETISM; PERMITTIVITY.

Elihu Abrahams; F. Keffer

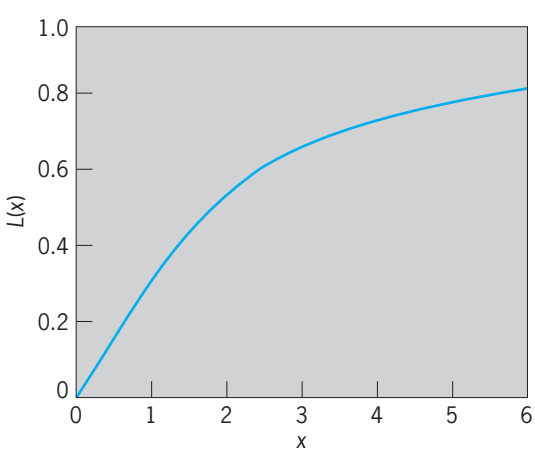

Plot of the Langevin function.

Langmuir circulation

A form of motion found in the near-surface water of lakes and oceans under windy conditions. When the wind is stronger than 5–8 m/s (10–15 knots), streaks of bubbles, seaweed, or flotsam form lines running roughly parallel to the wind, called windrows. Windrows are seen at one time or another on all bodies of water, from ponds to oceans. In the 1920s, Irving Langmuir hypothesized that they are produced by convergences in the water rather than by a direct action of the wind. Langmuir proposed that as the surface water is blown downwind it moves in a spiral fashion, first angling toward the streaks along the surface, next sinking to some depth, then diverging out from under the streaks, and finally rising again in between the streaks (see **illus.**). In a series of observations and experiments conducted in the North Atlantic and on Lake George in New York, he was able to confirm this basic form of the circulation.

In the ocean, the downwelling under windrows can be strong enough to pull bubbles, seaweed, and other buoyant particles down tens to hundreds of meters below the surface. The downward motion is eventually halted by a subtle increase in the water density with depth, associated with colder temperatures and/or higher salinity. The mixed surface layer typically spans most or all of the depth that light penetrates (the euphotic zone). This layer acts like the skin of the sea, through which heat, water vapor, oxygen, carbon dioxide, and all other materials must pass as they enter and leave the air and the sea. The mixing at the bottom of this layer also brings up nutrients and colder water from below.

The existence of an organized pattern of motion invites the notion that marine organisms have evolved to exploit it. For example, A. H. Woodcock noted that most Portuguese man-of-war (*Physalia*) in the North Atlantic are constructed so that they "sail" about 45° to the left of the wind (looking downwind). He speculated that a competitive advantage would be gained if Langmuir circulation were asymmetric (see illus.). Also, H. Stommel argued that particles that sink slowly (such as most phytoplankton) would, upon reaching the lower portion of the rolls, be advected toward the upwelling zones, between streaks, where they would be brought upward by the flow until they diverged away from the upwelling zone. He showed that particles can be trapped in retention zones, which depend on the sinking or rising rate. Sinking particles are trapped toward the upwelling region, while rising ones (such as bubbles) are trapped closer to the downwelling region under the streaks. Thus, the sinking of phytoplankton places them in the upwelling zones where nutrients are more readily available.

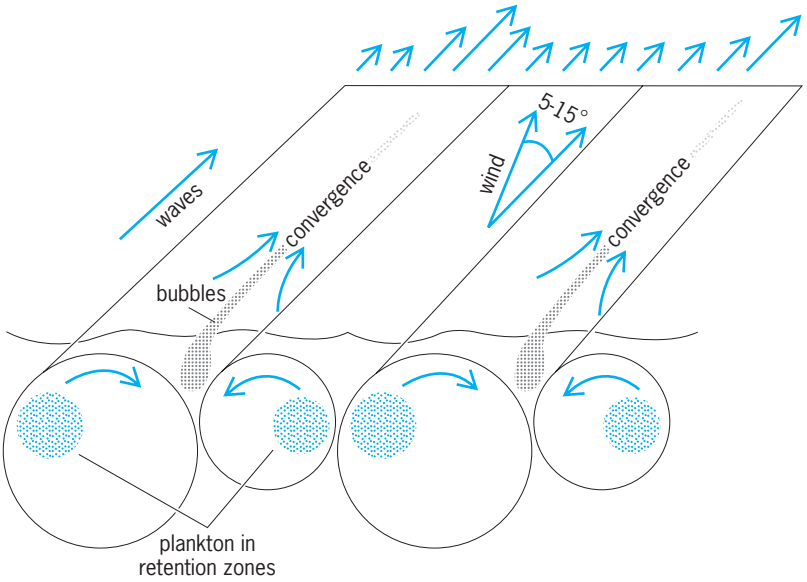

Langmuir circulation consists of a set of alternating rolls of water nearly aligned with the wind. Bubbles, seaweed, and flotsam gather into windrows where the surface water converges. Water beneath these windrows flows downward, drawing these materials to depths as great as 100 m or more. These downwelling zones tend to be narrow and intense compared to the broader, gentler upwelling in between. The rolls may be asymmetric, with stronger flows to the right of the wind at the surface (in the Northern Hemisphere). Details of the lower part of the motion are not yet clear.

In the late 1970s a workable mechanism for Langmuir circulation was deduced. A major clue came from laboratory studies by A. J. Faller and E. A. Caponi that demonstrated the need for both surface waves and wind-driven surface shear (downwind flow that is directly forced by the wind, so it is maximal at the surface and decreases rapidly with depth). Rigorous presentations of this mechanism were given by A. D. D. Craik and S. Leibovich. In brief, given a current pattern that resembles a downwind-directed jet (as might be generated by a large breaking wave), a wave-current interaction forces the surface water to flow from the sides toward the center of that jet. Then, since the surface water is accelerating under the wind stress, by the time it gets to the jet, it is moving downwind faster than the surrounding water. Thus, the jetlike feature becomes stronger and a positive feedback cycle is complete. The initially weak streak grows rapidly until some other influence, such as the development of a large enough density jump at the bottom of the mixed layer, arrests the growth.

The general characteristics of Langmuir circulation as known today are:

1. *Spacing*. In general, larger-scale streaks are associated with stronger current variations, up to some limiting size. The largest scale is related to the mixed layer depth (alluded to above), and the spacing between windrows is two to three times this depth. Since a pair of rolls is contained between each pair of windrows, the rolls themselves are about as deep as they are wide.

2. *Orientation*. Although the streaks are roughly aligned with the wind, there is a systematic bias in open ocean conditions, probably due to the effects of the Earth's rotation. In the Northern Hemisphere, oceanic streaks tend to align some 5–15° to the right of the wind (reversed in the Southern Hemisphere).

3. *Length, degree of organization*. The length of the streaks (in the alongwind direction) seems to vary widely, relative to the distance between them. In open ocean storms, the length may be only a few times the spacing. In contrast, for shallow flat-bottom lakes, the streaks can be quite regular, with the streaks extending alongwind distances many hundreds times the crosswind streak spacing. The streak length also relates to how long the features remain stable. In ocean storms, the streaks evolve rapidly compared to the time it would take for water to make a circuit around the roll (from surface to deep to surface again), while the long regular streaks of a shallow lake hardly vary in time.

4. *Strength*. Leaving aside the mean flow associated with wind (about 3% of the windspeed at the surface), the surface water velocities associated with Langmuir circulation are generally about 1% of the windspeed. However, this estimate can be off by a factor of 3 for reasons that remain unknown. Recent detailed examinations of wind and wave data suggests that the surface currents are more closely tied to the waves than the wind, although these are hard to separate since the waves are typically generated by the wind. But the uncertainty by a factor of 3 remains.

Dispersion of materials at the sea surface, such as oil, seaweed, flotsam, and life rafts, depends on the Langmuir circulation. Materials that are swept into the convergences along the streaks will move more rapidly downwind than they otherwise would. In the cross-wind direction, the spreading sidewise from streak to streak depends sensitively on how erratic versus straight and steady the streaks are.

In spite of decades' worth of data and the development of a reasonable driving mechanism, there remains much to learn about Langmuir circulation. In particular, observations to date have concentrated on the surface circulation, so the bottom portion of the motion has not been well defined.

Jerome A. Smith

Bibliography. A. D. D. Craik, The generation of Langmuir circulation by an instability mechanism, *J. Fluid Mech.*, 81:209–223, 1977; A. J. Faller and E. A. Caponi, Laboratory studies of wind-driven Langmuir circulations, *J. Geophys. Res.*, 83:3617–3633, 1978; I. Langmuir, Surface motion of water induced by wind, *Science*, 87:119–123, 1938; S. Leibovich, Convective instability of stably stratified water in the ocean, *J. Fluid Mech.*, 82:561–581, 1977; J. A. Smith, Evolution of Langmuir circulation during a storm, *J. Geophys. Res.*, 103:12649–12668, 1998; H. Stommel, Trajectories of small bodies sinking slowly through convection cells, *J. Mar. Res.*, 8:24–29, 1949; S. A. Thorpe, On the meandering and dispersion of a plume of floating particles caused by Langmuir circulation and a mean current, *J. Phys. Oceanog.*, 25:685–690, 1995; A. H. Woodcock, A theory of surface water motion deduced from the wind-induced motion of *Physalia*, *J. Mar. Res.*, 5:196–205, 1994.

Language theory

An attempt to formulate the grammar of a language in mathematical terms. Language theory is an important area of linguistics and computer science. Formal language theory was initiated in the mid-1950s in an attempt to develop theories on natural language acquisition. This theory, and in particular context-free grammars, was found to be relevant to the languages used in computers. Interest in the relationship between abstract languages and automata theory began with a seminal paper by S. C. Kleene in 1956, in which he characterized the languages in which membership of a sentence could be decided by a finite-state machine. Formal or abstract languages are based on the mathematical notion of a language as defined by Noam Chomsky around 1956. To understand this concept, we may begin by defining what a language is. The *New Oxford American Dictionary* defines it as "the method of human communication, either spoken or written, consisting of the use of words in a structured and conventional way." However, this definition is too vague to use as a building block of a language theory. To formalize the notion of an abstract language, it is necessary to introduce some preliminary definitions. An alphabet, vocabulary, or the set of terminals, denoted by Σ, is any finite, nonempty set of indivisible symbols. For example, the binary alphabet has only two symbols. This set is generally represented as $\Sigma = \{0, 1\}$. A word or string, over a particular alphabet, is a finite sequence of symbols of the alphabet. In mathematical terms, a typical word, x, can be written as $x = a_1, a_2, a_3, \ldots, a_k$ where $k \geq 0$, $a_i \in \Sigma$ for $1 \leq i \leq k$. Notice that if $k = 0$, the word is called the null word or empty word and is denoted by Λ. For example, using the binary alphabet we can form the words $x = 0010$ and $y = 010$. Given a word, x, the number of occurrences of symbols of a given alphabet in the word is called the length of the word and is denoted by $|x|$. According to this definition, the length of the words $|x| = |0010|$ and $|y| = |010|$ are 4 and 3, respectively. For a particular alphabet, a sentence is a finite sequence of words. *See* AUTOMATA THEORY; LINGUISTICS.

For a given alphabet, an abstract language is the set of all valid sentences of the language. A sentence may be considered valid if it satisfies a finite set of formation or production rules that define the grammar of the language. This definition of language can be applied to any natural or computer language. In a binary language, for instance, if we consider all the sequences of binary symbols of length 4 or less as valid words, then the words $x = 0010$ and $y = 010$ belong to this language. Sentences for a given language can be generated through a grammar. A grammar G can be formally defined as $G = (N, \Sigma, P, S)$, where N is a finite set of nonterminal symbols, also called variables or syntactical categories. Σ is called the alphabet or set of terminal symbols. Σ and N do not have any elements in common. That is, $\Sigma \cap N = \emptyset$. P is a finite set of production or replacement rules. Replacement rules have the form $\alpha \rightarrow \beta$, where α, β are strings. This rule indicates that an instance or occurrence of α can be replaced by an instance or occurrence of β. S is a distinguished element of N called the starting symbol. For example, let us consider a formal minigrammar G that generates four English sentences. This grammar has the following productions or replacement rules:

1. <sentence> → <subject> <verb> <object>
2. <subject> → <article> <noun>
3. <object> → <article> <noun>
4. <verb> → SEES
5. <article> → THE
6. <noun> → BOY
7. <noun> → GIRL

From this set of production rules, we can distinguish the elements that make up the grammar. The set of nonterminal symbols, N, is formed by the entities enclosed in pointed brackets. That is, N = {<sentence>, <subject>, <object>, <verb>, <article>, <noun>}. The alphabet Σ has four elements, namely, Σ = {SEES, THE, BOY, GIRL}. Each of the elements of the alphabet should be considered an indivisible unit. The replacement rules indicate how sequences of words in the language can be put together to generate sentences. For the given set of replacement rules, the interpretation of the first production is that a sentence is formed by a <subject> followed by a <verb> followed by an <object>. Similar interpretations can be given to rules 1 through 3. Rules 4 and 5 indicate that SEES, THE, BOY, and GIRL are instances of a <verb>, <article>, or <noun>, respectively. When there is more than one choice for a particular entity, such as <noun>, the entity can be replaced by the right-hand sides of one of the production rules. In this case, <noun> can be replaced by either BOY or GIRL. The starting symbol of this grammar is, by convention, the element of N that appears at the left of the symbol → of the first production. In this case, <sentence> is the starting nonterminal. One sentence generated by this grammar is THE GIRL SEES THE BOY. The sequence of replacement rules used to generate or derive this sentence is as follows:

<sentence> → <subject> <verb> <object>.
<sentence> → <article> <noun> <verb> <object>.
(using production 2 to replace <subject>)
<sentence> → THE <noun> <verb> <object>.
(using production 5 to replace <article>)
<sentence> → THE GIRL <verb> <object>.
(using production 7 to replace <noun>)
<sentence> → THE GIRL SEES <object>.
(using production 4 to replace <verb>)
<sentence> → THE GIRL SEES <article> <noun>
(using production 2 to replace <subject>)
<sentence> → THE GIRL SEES THE <noun>
(using production 5 to replace <article>)
<sentence> → THE GIRL SEES THE BOY.
(using production 6 to replace <noun>)

Notice that all the elements of the sentence THE

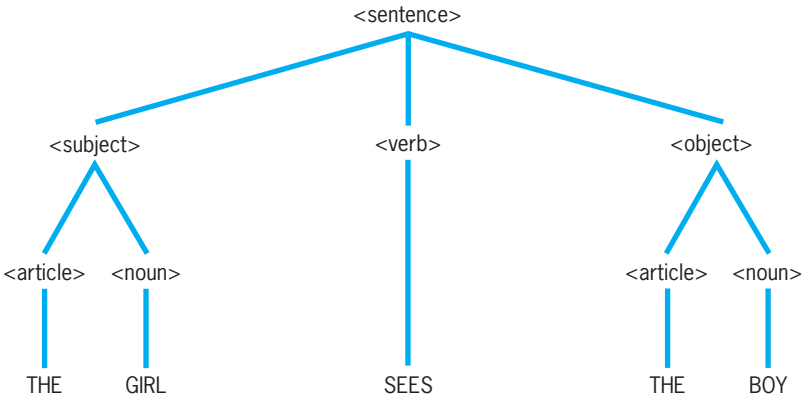

Fig. 1. Derivation tree of the sentence THE GIRL SEES THE BOY.

GIRL SEES THE BOY belong to the set of terminals. At each step of the derivation, a nonterminal on the right-hand side of a production was replaced by another sequence of nonterminal symbols or by a single terminal. The three additional sentences that can be generated using the production rules of this grammar are: THE BOY SEES THE GIRL, THE BOY SEES THE BOY, THE GIRL SEES THE GIRL. These four sentences form the language generated by the grammar. As this example shows, it is not necessary to explicitly indicate the elements of N, Σ, P, and S of the definition of a grammar since all these elements can be derived directly from the set of production rules. **Figure 1** shows a derivation tree of the sentence generated before.

Types of grammars. Grammars can be classified according to the nature of their productions. If $G = (N, \Sigma, P, S)$ and for every production $\alpha \rightarrow \beta$ in P, $|\beta| \geq |\alpha|$, the grammar is said to be context-sensitive or type 1. Some authors require that the productions of this type of grammar be of the form $\alpha_1 A \alpha_2 \rightarrow \alpha_1 \beta \alpha_2$, where α_1, α_2, and β are strings, $\beta \neq \Lambda$, and $A \in N$. This restriction, which does not alter the language generated by the grammar, motivates the term context-sensitive because A can be replaced by β only within the context of α_1 and α_2 according to $\alpha_1 A \alpha_2 \rightarrow \alpha_1 \beta \alpha_2$.

If the productions $\alpha \rightarrow \beta$ of the grammar G are such that $|\alpha| = 1$ and α is a single element of N and β is a string of terminals and/or nonterminals, then the grammar is a context-free grammar or type 2. Context-free grammars and the language that they generate are very important in computer science since most of the computer languages are generated by this type of grammar.

Acceptors, generators, and translators. The relationship between abstract languages and machines can be established through three types of automata: acceptors, generators, and translators. An acceptor is a deterministic finite-state machine which, starting in a predefined initial state, accepts or rejects sequences of input symbols from an alphabet. We say that a finite-state machine operates deterministically if its responses depend solely on its initial state and the input sequence presented. In simple terms, we can picture an acceptor as a machine connected to a lamp which can be turned on and off. As the machine processes each of the symbols of the input sequence, the machine turns the lamp on or off. If the machine turns the lamp on after processing the last symbol of the input sequence, the machine accepts the input sequence; otherwise, the machine rejects the sequence.

A generator is a nondeterministic finite-state machine. When the generator is started from its initial state, it produces a sequence of symbols from a given alphabet known as the output alphabet. It is assumed that the machine will act according to chance producing different sequences each time it is operated. The language of the machine is the set of all sequences that it will ever produce.

A translator is a machine that after processing an input sentence (made up of symbols of an input alphabet) translates the sentence into another sentence (made of symbols of an output alphabet). Translators are important not only in the translation of natural languages (for example, from English

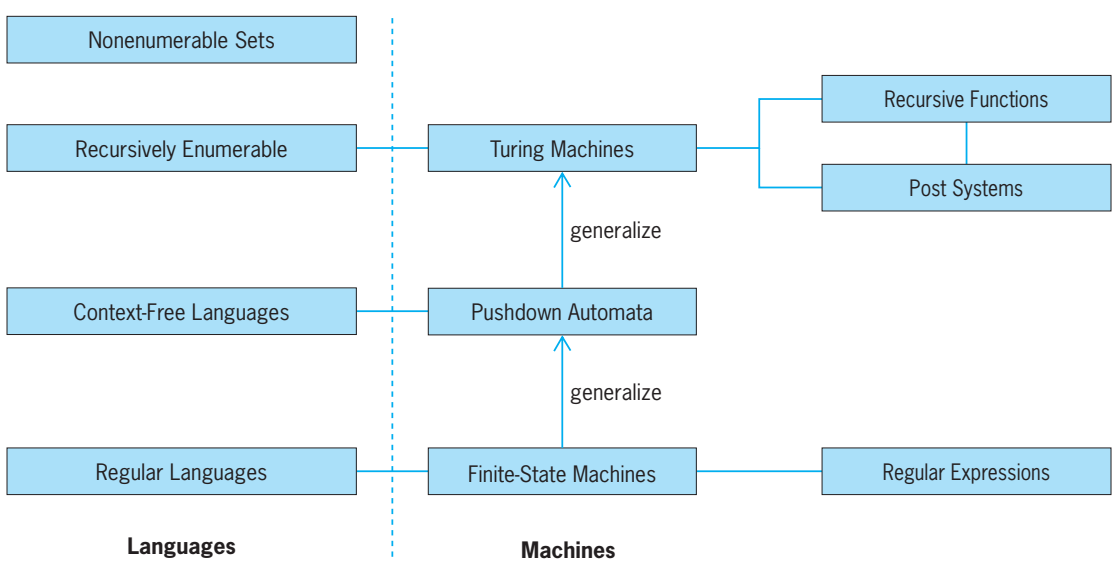

Fig. 2. Hierarchy of abstract machines and languages. (*After P. Denning et al., Machines, Languages, and Computation, 1978*)

to Spanish) but also in the translation of computer languages (for example, from C++ to assembly or machine language). *See* NATURAL LANGUAGE PROCESSING; PROGRAMMING LANGUAGES.

Figure 2 shows a hierarchy of automata and the corresponding hierarchy of languages that they define. Although some of the topics mentioned in this hierarchy are out of the scope of this article, they are included here for completeness.

Ramon A. Mata-Toledo

Bibliography. A. V. Aho et al., *The Theory of Parsing, Translation, and Compiling*, vol. 1, 1972 J. E. Hopcroft et al., *Introduction to Automata Theory, Languages, and Computation*, 2d ed., 2000; J. Martin, *Introduction to Languages and the Theory of Computation*, 2002.

Lanolin

A soft, waxy material derived from the greasy coating on raw wool. In the processing of wool, the fleece is scoured with an aqueous alkaline solution to remove debris, wax, water-soluble material, and free acids. The insoluble fraction, which constitutes about 15–20% of the weight of the original wool, is crude lanolin.

Lanolin is a very complex mixture of esters, similar to the skin lipids of birds and other animals. Analysis is carried out after saponification, which gives equal weights of acids and alcohols. The acids consist of 50–60% saturated fatty acids with chains up to C_{38}–C_{40}, comprising mostly isostructures or anteisostructures with a methyl group that is one or two carbons removed from the end of the chain [$CH_3CH(CH_3)$— or $CH_3CH_2CH(CH_3)$—]. Another group (30–35%) comprises α-hydroxy acids and a small amount of ω-hydroxy acids. *See* ESTER.

The alcohols from lanolin are of two types. About 15% are monohydroxy compounds with the same chain structures as the acids, and a minor amount of 1,2-diols. The major alcohol components are 35% cholesterol and 35–40% lanosterol. The latter is a key intermediate in the biosynthesis of steroids from squalene. *See* CHOLESTEROL; SQUALENE; STEROID.

The supply of lanolin depends directly on the level of wool production; the consumption of lanolin and derived products in the United States is about 1×10^7 lb (4.5×10^6 kg) per year. Crude lanolin from the wool-scouring process can be refined to a colorless, odorless material (USP grades). A particularly useful property of lanolin is the formation of stable water-in-oil emulsions containing up to 25% water (hydrous lanolin). Major uses are as an emollient and skin moisturizer in lotions and cosmetic products, and in medicinal ointments. Unrefined lanolin has some use in inks and as a corrosion and rust preventative. *See* FAT AND OIL; WOOL.

James A. Moore

Lanthanide contraction

The name given to an unusual phenomenon encountered in the rare-earth series of elements. The radii of the atoms of the members of this series decrease slightly as the atomic number increases. Starting with element 58 in the periodic table, the balancing electron fills in an inner incomplete *4f* shell as the charge on the nucleus increases. According to the theory of atomic structure, this shell can hold 14 electrons; so starting with element 58, cerium, there are 14 true rare earths. Lanthanum has no electrons in the *4f* shell, cerium has 1, and lutetium, 14. The *4f* electrons play almost no role in chemical valence; therefore, all rare earths can have three electrons in their valence shell and they all exist as trivalent ions in solution. As the charge on the nucleus increases across the rare-earth series, all electrons are pulled in closer to the nucleus so that the radii of the rare-earth ions decrease slightly as the compounds go across the rare-earth series. Any given compound of the rare earths is very likely to crystallize with the same structure as any other rare earth. However, the lattice parameters become smaller and the crystal denser as the

Atomic and ionic radii of rare-earth metals

Element	Radius of M^{3+}*, nm	Metal crystal structure†	Metallic radii, nm	
Sc		hcp	0.16545‡	0.16280§
Y		hcp	0.18237	0.17780
La	0.1061	hcp	0.18852	0.18694
Ce	0.1034	fcc	0.18248	
Pr	0.1013	hcp	0.18363	0.18201
Nd	0.0995	hcp	0.18290	0.18139
Pm	0.0979			
Sm	0.0964	rhom-hcp	0.18105	0.17943
Eu	0.0950	bcc	0.1994	
Gd	0.0938	hcp	0.18180	0.17865
Tb	0.0923	hcp	0.18005	0.17626
Dy	0.0908	hcp	0.17952	0.17515
Ho	0.0894	hcp	0.17887	0.17428
Er	0.0881	hcp	0.17794	0.17340
Tm	0.0869	hcp	0.17688	0.17237
Yb	0.0858	fcc	0.19397	
Lu	0.0848	hcp	0.17516	0.17171

*Trivalent rare-earth ion.
†hcp, hexagonal close-packed; fcc, face-centered cubic; rhom, rhombic; bcc, body-centered cubic.
‡All values in this column calculated from atoms in basal plane.
§All values in this column are for radii between layers.

compound proceed across the series. This contraction of the lattice parameters is known as the lanthanide contraction. For many compounds the lattice parameters decrease only partway across the series, and when the contraction has progressed to that point, a new crystalline form develops. Frequently, both crystalline forms can be observed for a number of the elements. For this reason, the rare-earth series is of particular interest to scientists because many of the parameters determining the properties of a substance can be kept constant while the lattice spacings can be varied in small increments across the series.

The atomic and ionic radii of atoms are not clearly defined. The atoms can be polarized by the neighboring atoms and there is no clear-cut boundary between the electrons associated with one atom and another. Therefore, the atomic radii will vary somewhat from compound to compound, and the absolute values depend on the method of calculation. However, if most of the parameters are assumed constant, and the difference in lattice parameters in the rare-earth crystalline series is attributed to the rare-earth ion or atom, the lanthanide contraction becomes evident. Although scandium and yttrium are not members of this series, the information is usually wanted at the same time and is given for completeness. The atomic radii of the trivalent ion and the metal atoms are given in the **table**. *See* PERIODIC TABLE; RARE-EARTH ELEMENTS.

Frank H. Spedding

Lanthanum

A chemical element, La, atomic number 57, atomic weight 138.91. Lanthanum, the second most abundant element in the rare-earth group, is a metal. The naturally occurring element is made up of the isotopes ^{138}La, 0.089%, and ^{139}La, 99.91%. ^{138}La is a radioactive positron emitter with a half-life of 1.1×10^{11} years. The element was discovered in 1839 by C. G. Mosander and occurs associated with other rare earths in monazite, bastnasite, and other minerals. It is one of the radioactive products of the fission of uranium, thorium, or plutonium. Lanthanum is the most basic of the rare earths and can be separated rapidly from other members of the rare-earth series by fractional crystallization. Considerable quantities of it are separated commercially, since it is an important ingredient in glass manufacture. Lanthanum imparts a high refractive index to the glass and is used in the manufacture of expensive lenses. The metal is readily attacked in air and is rapidly converted to a white powder. Lanthanum becomes a superconductor below about 6 K (−449°F) in both the hexagonal and face-centered crystal forms. *See* RARE-EARTH ELEMENTS.

Frank J. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

1												13	14	15	16	17	18
1 **H**	2																2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Laplace transform

An integral extensively used by P. S. Laplace in the theory of probability. In simplest form it is expressed as Eq. (1). It is thought of as transforming the de-

$$f(s) = \int_0^\infty e^{-st}\phi(t)\,dt \qquad (1)$$

termining function $\phi(t)$ into the generating function $f(s)$. The variable t is real, the variable s may be real or complex, $s = \sigma + i\tau$. As an example, if $\phi(t) = 1$ the integral converges for $\sigma > 0$, and $f(s) = 1/s$.

The Laplace transform is used for the solution of differential and difference equations, for the evaluation of definite integrals, and in many branches of abstract mathematics (functional analysis, operational calculus, and analytic number theory).

Method. Extensive tables of Laplace transforms exist, and these are used as any table of integrals. To show how a differential equation may be solved, two excerpts (A and B) from such a table can be used.

A. $f(s) = 1/(s-a) \qquad \phi(t) = e^{at}$

B. $f(s) = 1/(s^2+1) \qquad \phi(t) = \sin t$

Suppose it is required to find a solution $y(t)$ of Eq. (2) such that $y(0) = 1$, $y'(0) = 2$. Denote the

$$y''(t) + y(t) = 2e^t y'' = \frac{d^2y}{dt^2}, y' = \frac{dy}{dt} \qquad (2)$$

Laplace transform of the unknown function $y(t)$ by $Y(s)$. Integration by parts gives Eq. (3) on the assump-

$$\int_0^\infty e^{-st}y''(t)\,dt = -y'(0) - y(0)s + s^2\int_0^\infty e^{-st}y(t)\,dt = -2 - s + s^2Y(s) \qquad (3)$$

tion that the integrated part is zero at $t = \infty$. Applying the Laplace transform to Eq. (2) and using A for the right-hand side, one obtains Eq. (4). The differential

$$-2 - s + s^2Y(s) + Y(s) = \frac{2}{s-1} \qquad (4)$$

equation has become an algebraic one, whose

solution is Eq. (5). However, a further use of the table

$$Y(s) = \frac{1}{s-1} + \frac{1}{s^2+1} \tag{5}$$

shows that the Laplace transform of $y(t) = e^t + \sin t$ is precisely the right-hand side of Eq. (5). Assuming uniqueness, one has thus obtained the required solution. Because its properties can be checked directly, the unproved assumptions need not be verified.

This example illustrates the general method. The unknown function is taken as the determining function and the Laplace transform is applied to the differential (or difference) equation. There results an equation with the generating function as unknown, and this must be solved. Finally the determining function must be determined from the generating function, either from tables or by use of an inversion formula. In general, if the original differential equation is partial in any number of independent variables, one application of the Laplace transform reduces the number of these variables by one. If the equation was ordinary (one independent variable), the transformed equation is algebraic, as in the above example.

Properties. Here are the fundamental properties of the Laplace transform:

I. There exists a number σ_c (perhaps $+\infty$ or $-\infty$) called the abscissa of convergence such that the integral in Eq. (1) converges for $\sigma > \sigma_c$, diverges for $\sigma < \sigma_c$. That is, the region of convergence is a half-plane (a half-line if s is real).

II. The generating function is holomorphic for $\sigma > \sigma_c$.

III. The determining function is uniquely determined by the generating function. (Ambiguity is possible only on sets of measure zero.)

IV. The product of two generating functions is in general a generating function. Thus, if Eq. (1) holds for two pairs of functions $f_1(s)$, $\phi_1(t)$ and $f_2(s)$, $\phi_2(t)$, then the product $f_1(s)f_2(s)$ is the transform of the convolution

$$\phi_1(t)^*\phi_2(t) = \int_0^t \phi_1(u)\phi_2(t-u)\,du$$

As was evident in the above example, it is very important to be able to derive the determining function $\phi(t)$ from the generating function $f(s)$. This is especially true when tables are unavailable or inadequate. Any expression of $\phi(t)$ in terms of $f(s)$ is called an inversion formula. Many are known. The classical one is Eq. (6). Here the integration is along any line

$$\phi(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} f(s)e^{st}ds \qquad 0 < t < \infty \tag{6}$$

$\sigma = c$ of the complex s-plane on which the integral in Eq. (1) converges absolutely. Another inversion which employs the real variable only is Eq. (7). Here

$$\phi(t) = \lim_{k\to\infty} \frac{(-1)^k}{k!} f^{(k)}\left(\frac{k}{t}\right)\left(\frac{k}{t}\right)^{k+1} \tag{7}$$

$$0 < t < \infty$$

$f^{(k)}(x)$ means the kth derivative of $f(x)$. Equation (7) can be illustrated by the example A above. For that pair $f^{(k)}(x)$ is easily computed and Eq. (7) becomes Eq. (8), a familiar result of calculus.

$$e^{at} = \lim_{k\to\infty}\left(1 - \frac{at}{k}\right)^{-k-1} \tag{8}$$

See INTEGRAL TRANSFORM.

Generalizations. Certain generalizations of Eq. (1) are in frequent use. The transform shown as Eq. (9)

$$f(s) = \int_{-\infty}^{\infty} e^{-st}\phi(t)\,dt \tag{9}$$

is called the bilateral Laplace transform. An inversion is still provided by Eq. (6), which now holds for $-\infty < t < \infty$. If one sets $s = iy$ in Eq. (9), the result is Eq. (10). This equation defines $g(y)$ as the Fourier

$$g(y) = f(iy) = \int_{-\infty}^{\infty} e^{-iyt}\phi(t)\,dt \tag{10}$$

transform of $\phi(t)$. That is, the Laplace transform (9), if considered along a single line, becomes a Fourier transform. By setting $c = 0$ and $s = iy$ in formula (6) one obtains Eq. (11), the classical inversion of the Fourier transform.

$$\begin{aligned}\phi(t) &= \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} f(s)e^{st}\,ds \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} g(y)e^{iyt}\,dy\end{aligned} \tag{11}$$

Another generalization of Eq. (1) is the Laplace-Stieltjes integral, Eq. (12), where now the integral

$$f(s) = \int_0^{\infty} e^{-st}\,d\alpha(t) \tag{12}$$

is a Stieltjes integral with respect to the "integrator" function $\alpha(t)$. If $\alpha(t)$ has a derivative $\phi(t)$ the integral (12) becomes the integral (1). On the other hand, if $\alpha(t)$ is a step-function, Eq. (12) reduces to a Dirichlet series, Eq. (13), a type of series of great importance

$$f(s) = \sum_{k=1}^{\infty} a_k e^{-\lambda_k s} \tag{13}$$

in analytic number theory.

It must not be supposed that one may choose $\phi(t)$ or $f(s)$ arbitrarily in Eq. (1) and expect its mate to exist. For example if $\phi(t) = e^{t^2}$ the integral (1) diverges for all s and $\sigma_c = +\infty$. Again if $f(s) = s$, no corresponding determining function $\phi(t)$ exists, since it is easily seen that every generating function must approach a limit as $s \to +\infty$ along the real axis. Hence it is clearly important to know what functions $\phi(t)$ and $f(s)$ may be used in Eq. (1). So far as $\phi(t)$ is concerned the problem is completely solved by the formula, with $\sigma_c > 0$, shown as Eq. (14). The other

$$\sigma_c = \overline{\lim_{t\to\infty}} \frac{\log|\alpha(t)|}{t} \tag{14}$$

problem has been partially solved by representation theorems for integral transforms, one striking example of which is presented below.

A function $f(s)$ of the real variable s is said to be completely monotonic on $a < s < \infty$ if and only if relations (15) hold. Examples are $f(s) = 1$, $f(s) =$

$$f(s) \geq 0, f'(s) \leq 0, f''(s) \geq 0, f'''(s) \leq 0, \ldots \qquad a < s < \infty \tag{15}$$

$1/(s - a)$, and $f(s) = e^{-s}$. A theorem of S. Bernstein states that $f(s)$ has a representation (12) converging for $s > a$ and with integrator function $\alpha(t)$ nondecreasing if and only if $f(s)$ is completely monotonic for $a < s < \infty$. For example, if $f(s) = 1$, then $\alpha(t) = 1$, $t > 0$, $\alpha(0) = 0$; if $f(s) = 1/(s - a)$, then $\phi(t) = e^{at}$ and $\alpha(t) = (e^{at} - 1)/a$; if $f(s) = e^{-s}$, $\alpha(t) = 0$ for $0 < t < 1$ and $\alpha(t) = 1$ for $1 < t < \infty$. In each case $\alpha(t)$ is nondecreasing, as predicted by Bernstein's theorem. This result is particularly remarkable because the mere signs of the successive derivatives of a function on the real axis determine not only its holomorphic character (property II above) in a half-plane but also its representation in the form of Eq. (12). David V. Widder

Bibliography. R. E. Bellman and R. S. Roth, *The Laplace Transform*, 1984; A. C. Grove, *Introduction to the Laplace Transform and the Z Transform*, 1991; W. R. LePage, *Complex Variables and the Laplace Transform for Engineers*, 1961, reprint 1980.

Laplace's differential equation

Laplace's equation in two independent variables x and y is given as Eq. (1) and is of central importance

$$\frac{\partial^2 u(x, y)}{\partial x^2} + \frac{\partial^2 u(x, y)}{\partial y^2} = 0 \tag{1}$$

in both pure mathematics and mathematical physics. A function $u(x,y)$ having continuous first and second partial derivatives and satisfying Laplace's equation in a neighborhood of a point is called harmonic at that point. If a plane piece of tinfoil has its edges kept at a temperature which varies from point to point but does not change with time, and if the flow of heat in the tinfoil is steady (that is, independent of the time), the temperature $u(x,y)$ at interior points of the foil is harmonic. Likewise Laplace's equation dominates the flow of electricity (the potential is similarly harmonic) and the flow of any incompressible fluid.

Two-dimensional relations. If $f(z) \equiv u(x, y) + iv(x, y)$ is an analytic function, $u(x, y)$ and $v(x, y)$ are conjugate functions and are harmonic; conversely, if $u(x, y)$ is harmonic in a simply connected region D, one may write Eq. (2), where (x_0, y_0) is fixed in D and

$$v(x, y) \equiv \int_{(x_0, y_0)}^{(x,y)} \left(-\frac{\partial u}{\partial y}\, dx + \frac{\partial u}{\partial x}\, dy \right) \tag{2}$$

(x, y) arbitrary in D. It follows from Green's theorem that the integral over a path in D is independent of the path, so $v(x, y)$ is uniquely defined throughout D; the functions $u(x, y)$ and $u(x, y)$ are conjugate in D, and $f(z) \equiv u + iv$ is analytic there. Under these conditions, let C now be a regular Jordan curve in D; if n denotes the interior normal of C, the equation $\partial u/\partial n = -\partial v/\partial s$ follows from the Cauchy-Riemann equations, whence obtains Eq. (3). The first and last

$$\int_C \frac{\partial u}{\partial n}\, ds = -\int_C \frac{\partial v}{\partial s}\, ds = -v(x, y)\Big|_C = 0 \tag{3}$$

members of this equation form the flux theorem, namely that the total flux (of heat if u is temperature) over C is zero. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

If $u(x, y)$ is harmonic in the closed disk bounded by the circumference γ, and $f(z)$ the corresponding analytic function, one can take the real parts of both members of the equations expressing Cauchy's integral formula, as in Eqs. (4).

$$z - z_0 = \rho(\cos\theta + i\sin\theta) \tag{4a}$$

$$dz = i(z - z_0)\, d\theta \tag{4b}$$

$$f(z_0) = u(x_0, y_0) + iv(x_0, y_0) = \frac{1}{2\pi i}\int_\gamma \frac{f(z)\, dz}{z - z_0}$$

$$= \frac{1}{2\pi}\int_\gamma f(z)\, d\theta \tag{4c}$$

$$u(x_0, y_0) = \frac{1}{2\pi}\int_\gamma u(x, y)\, d\theta \tag{4d}$$

Equation 4*d* expresses Gauss's mean value theorem, that the average of $u(x, y)$ over γ is the value at the center of γ. From this theorem it follows that a function harmonic at a point (x_0, y_0) cannot have a strong local maximum (or minimum) there, and can have a weak local maximum (or minimum) only if identically constant throughout a neighborhood of (x_0, y_0). If $u(x, y)$ is harmonic in a bounded region D, continuous in the corresponding closed region $\overline{D}$, the maximum and minimum of $u(x, y)$ occur on the boundary of D; if a maximum or minimum occurs interior to D, then $u(x, y)$ is identically constant throughout D.

If D is a bounded region with boundary B, and if continuous values $U(x, y)$ are assigned on B, the Dirichlet problem is the problem of determining a function $u(x, y)$ harmonic in D, continuous on $D + B$, equal to $U(x, y)$ on B. If D is a circular region, a Jordan region, or any nonpathological region, the Dirichlet problem has a solution, necessarily (by the absence of nontrivial maxima and minima interior to D) unique. If D is a circular disk of radius a (see **illus.**), the Dirichlet problem for D is solved by Poisson's integral, Eq. (5), using polar coordinates $(r\theta)$

$$u(r, \theta) = \frac{1}{2\pi}\int_0^{2\pi} \frac{(a^2 - r^2)U(\psi)\, d\psi}{a^2 - 2ar\cos(\theta - \psi) + r^2} \tag{5}$$

with pole the center of D. If D is a less elementary region but with smooth boundary B, the Dirichlet problem is solved by Green's formula, Eq. (6), where n indicates the interior normal in this formula.

$$u(x,y) = \frac{1}{2\pi}\int_B U(\xi,\eta)\frac{\partial g}{\partial n}\, ds\,(\xi,\eta) \tag{6}$$

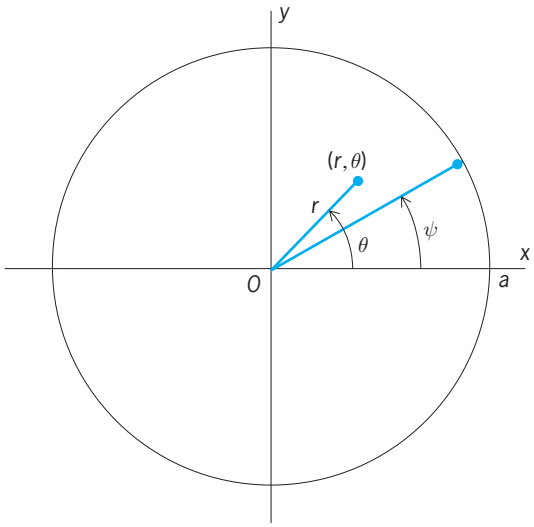

Circular disk *D* of radius *a*.

Green's function $g(x, y; \xi, \eta)$ is harmonic in D except at (x, y), continuous and equal to zero on B, and in the neighborhood of (x, y) has the form $^1/_2 \log [(\xi - x)^2 + (\eta - y)^2] + g_1(\xi, \eta)$, where $g_1(\xi, \eta)$ is harmonic at (x, y). If the boundary B is not smooth, this formula can be expressed in terms of harmonic measure instead of $(\partial g/\partial n)\, ds$.

Numerous series expansions (for example, Fourier's series) can be used for the solution of the Dirichlet problem for various regions.

n-Dimensional relations. The foregoing remarks apply to Laplace's equation with two independent variables; the facts (but not the methods of proof using analytic functions) apply also in three or more dimensions. Thus, in three dimensions, a point distribution of matter of masses m_k at points (x_k, y_k, z_k) has a potential defined by Eq. (7), which is har-

$$u(x, y, z) \equiv \sum m_k[(x - x_k)^2 + (y - y_k)^2 + (z - z_k)^2]^{-1/2} \quad (7)$$

monic except in the points (x_k, y_k, z_k). Except at such points, the force (newtonian law of gravitation) exerted by the distribution on a unit exploratory particle at (x, y, z) has the components $(\partial u/\partial x, \partial u/\partial y, \partial u/\partial z)$, and the component of the force in any direction is the directional derivative of $u(x, y)$ in that direction. *See* POTENTIALS; SPHERICAL HARMONICS.

Joseph L. Walsh

Bibliography. J. Bliedtner, *Potential Theory*, 1986; J. Kral (ed.), *Potential Theory*, 1988.

Laplacian

The differential operator $\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$, in which the symbols x, y, z denote the variables of a rectangular cartesian coordinate system. The laplacian is frequently denoted by the symbol ∇^2 (read del square) in accordance with the fact that the laplacian of a scalar function $S(x, y, z)$ is the divergence of the gradient of S, that is, the equation below applies.

$$\frac{\partial^2 S}{\partial x^2} + \frac{\partial^2 S}{\partial y^2} + \frac{\partial^2 S}{\partial z^2} = \nabla \cdot (\nabla S)$$

The laplacian operator is involved in some of the most fundamental equations of mathematical physics, namely, Laplace's equation ($\nabla^2 u = 0$), Poisson's equation, various wave equations (as those of electricity and magnetism, sound, vibrations, the Schrödinger equation of quantum mechanics), and the heat flow and diffusivity equations. *See* CALCULUS OF VECTORS; GAUSS' THEOREM; GRADIENT OF A SCALAR; GREEN'S THEOREM.

Homer V. Craig

Larch

A genus, *Larix*, of the pine family, with deciduous needles and short spurlike branches, which annually bear a crown of needles. The cones are small and persistent, varying by species in size, number, and form of the cone scales. The tamarack (*L. laricina*), also called hackmatack, is a native species, has erect, narrowly pyramidal habit, and grows in wet and moist soils in the northeastern United States, west to the Lake states, and across Canada to Alaska. The cones are $^1/_2$–$^3/_4$ in. (12–19 mm) long (see **illus.**). The tough resinous wood is durable in contact with the soil and is used for railroad ties, posts, sills, and boats. Other uses include the manufacture of excelsior, cabinet work, interior finish, and utility poles. *See* PINALES.

The western larch (*L. occidentalis*), the most important and largest of all the species, grows in the northwestern United States and southeastern British Columbia. The cones of this species are larger, 1–$1^1/_2$ in. (25–38 mm) long, with bracts growing out beyond the cone scales. The trunk is tall and erect, sometimes attaining a height of 200 ft (60 m) and a diameter of 6–7 ft (1.8–2.1 m). More than one-half of the annual production of larch comes from Montana, the remainder from Idaho, Washington, and Oregon.

The European larch (*L. decidua*) has cones about twice the size of those of the tamarack and 40–50 scales to a cone, whereas tamarack has only

Cone and needles of tamarack (*Larix laricina*).

12–15. The European larch does better in drier soil and is the species usually planted in parks and private grounds.

Golden larch (*Pseudolarix amabilis*), from China, is occasionally cultivated. Its leaves, also deciduous, are golden yellow in the fall. The cone scales fall off one by one, leaving the central axis of the cone on the tree. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Large systems control theory

A branch of control theory concerned with large-scale systems. The three commonly accepted definitions of a large-scale system are based on notions of decomposition, complexity, and centrality. A system is sometimes considered to be large-scale if it can be partitioned or decomposed into small-scale subsystems. Another definition is that a system is large-scale if it is complex; that is, conventional techniques of modeling, analysis, control, design, and computation do not give reasonable solutions with reasonable effort. A third definition is based on the notion of centrality. Until the advent of large-scale systems, almost all control systems analysis and design procedures were limited to components and information grouped in one geographical location or center. Thus, by another definition, a system in which the concept of centrality fails is large-scale. This can be due to a lack of either centralized computing capability or a centralized information structure. Large-scale systems appear in such diversified fields as sociology, management, the economy, the environment, computer networks, power systems, transportation, aerospace, robotics, manufacturing, and navigation. Some examples of large-scale systems are the United States economy, the global telephone communication network, and the electric power generation system for the western United States.

Large-scale control systems have been an active research area for some time, but the field is still in its infancy. Nevertheless, it has set forth foundations for the systematic treatment of a class of very complex physical systems.

Since the 1950s, system theory has evolved from a semiheuristic discipline directed toward the design and analysis of electronic and aerospace systems containing few components into a very sophisticated theory capable of treating large, complex systems with myriad applications. Initially, system engineers attempted to cope with the increasing complexity through the development of sophisticated numerical techniques in order to apply classical theory to large systems. This approach, however, soon reached a point of diminishing returns, and it became apparent that new theoretical techniques would be necessary. Although many such techniques are still being developed, it is generally accepted that the key to the successful treatment of a large-scale system is to exploit fully its structural interconnection. This takes place in two ways: by the full use of sparse matrix techniques or by the decomposition of larger systems into smaller ones. *See* CONTROL SYSTEMS; SYSTEMS ENGINEERING.

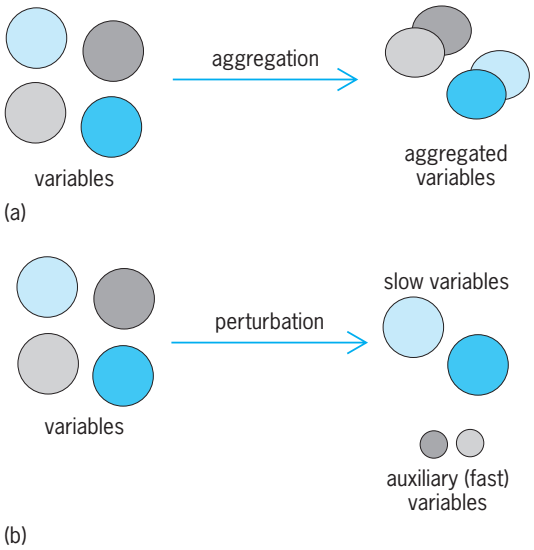

Fig. 1. Model reduction techniques. (*a*) Aggregation. (*b*) Perturbation.

Modeling and model reduction. The first step in any scientific or technological study of a system is to obtain a mathematical model. The major effort in large-scale system modeling has been reduction by aggregation and perturbation. An aggregate model of a system is described by a coarser set of variables. The underlying reason for aggregating a system model is to retain the key qualitative properties of the system, such as stability. Perturbation, on the other hand, is based on ignoring certain interactions of a system. Here again, however, the key properties must not be sacrificed for the sake of reduced computation. The common feature of both aggregation and perturbation is the reduction of system order. **Figure 1** depicts the way in which a fourth-order system is reduced to second order by both schemes. In aggregation, a combination (mixture) of the states (circles) is retained, whereas in perturbation the slow variables are retained and the fast ones are not. *See* PERTURBATION (MATHEMATICS).

In many applications of complex control systems, the accuracy of the model can come into question. To overcome uncertainties and inaccuracies in the model, a larger view of the complex system, encompassing the model, its past history, operational information, and so forth, should be addressed. This view may be designated collectively as knowledge. *See* MODEL THEORY.

Hierarchical control. A large-scale system, as defined above, can be described as a complex system composed of a number of constituents or smaller subsystems serving particular functions, sharing resources, and governed by interrelated goals and constraints. Although interaction among subsystems can take many forms, one of the most common is hierarchical, which appears to be natural in economic management, organizations, and complex industrial

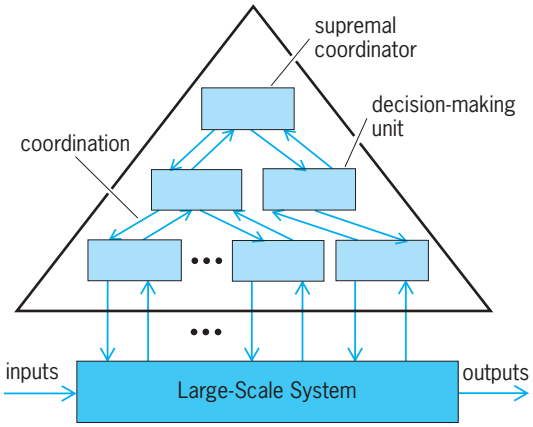

Fig. 2. Pyramidal structure of a hierarchical large-scale system. Each row of boxes represents a hierarchical level. Downward vertical arrows represent control actions, and upward vertical arrows represent output measurements.

systems. Within the hierarchical structure, the subsystems are positioned on different levels. A subsystem at a given level controls or coordinates the units on the level below it and is in turn controlled or coordinated by the unit immediately above.

Figure 2 shows a typical hierarchical (multilevel) system. The highest-level coordinator, sometimes called the supremal coordinator, can be thought of as the board of directors of a corporation, whereas other coordinators may be the president, vice president, and directors. Lower levels are occupied by the plant manager or shop managers; the large-scale system is the corporation itself. In spite of this seemingly natural representation of a hierarchical structure, its exact behavior is not well understood, mainly because relatively little quantitative work has been done on these large-scale systems.

There is no unique or universally accepted set of properties associated with hierarchical systems. However, the following are key properties: (1) a hierarchical system consists of decision-making components structured in a pyramidal shape (Fig. 2); (2) the system has an overall goal that may or may not be that of its components; (3) the various levels in the system exchange information (usually vertically) iteratively; and (4) as the level of hierarchy increases, the time horizon increases; that is, the lower-level components work faster than those at the higher level. *See* MULTILEVEL CONTROL THEORY.

Decentralized control. Most complex systems are characterized by a great multiplicity of measured outputs and inputs. For example, an electric power system has several control substations, each being responsible for the operation of a portion of the overall system. This situation arising in a control system design is often referred to as decentralization. The designer for such systems determines a structure for control which assigns system inputs to a given set of local controllers (stations) that observe only local system outputs. This approach, called decentralized control, attempts to avoid the difficulties associated with data gathering, storage requirements, computer program debugging, and geographical separation of system components.

In hierarchical control, a decomposition in system structure leads to computational efficiency. In decentralized control, simpler controller structures and computational efficiency result from a decomposition in the system's output information. The basic characteristic of any decentralized system is that the transfer of information from one group of sensors or actuators to others is quite restricted. For example, in the two-controller decentralized system of **Fig. 3**, only the output y_1 and external input v_1 are used to find the control u_1, and likewise the control u_2 is obtained through only the output y_2 and external input v_2.

The determination of control signals u_1, and u_2 based on the output signals y_1 and y_2, respectively, is nothing but two independent output feedback problems, which can be used for stabilization or pole-placement purposes. It is therefore clear that the decentralized control scheme is of feedback form, indicating that this method is very useful for complex linear systems. *See* CONTROL SYSTEM STABILITY; LINEAR SYSTEM ANALYSIS.

Attributes of large systems. Although there is no universally accepted set of attributes for large complex systems, the following may be a good starting point.

1. Large complex systems are often controlled by more than one controller or decision maker, involving decentralized computations.
2. The controllers have different but correlated information available to them, possibly at different times.
3. Large complex systems can also be controlled by local controllers at one level whose control actions are being coordinated at another level in a hierarchical (multilevel) structure.
4. Large complex systems are usually represented by imprecise aggregate models.
5. Controllers may operate in a group as a team or in a conflicting manner, with single- or multiple-objective or even conflicting-objective functions.
6. Large complex systems may be satisfactorily optimized by means of suboptimal or near-optimal controls, sometimes termed a satisfying strategy.

Autonomous control systems. The difficulties in developing a complex large-scale system model begin with its availability, and propagate through the

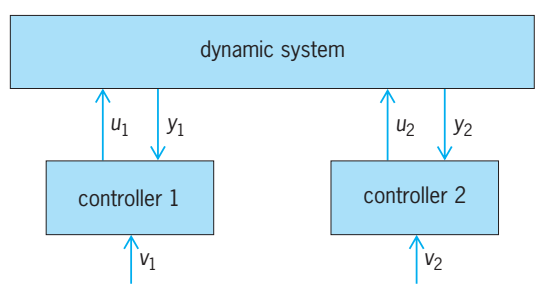

Fig. 3. Two-controller decentralized system.

simulation, analysis, and design of the system. Attempts have been made to introduce a new tool, artificial intelligence, into the modeling and control of systems—large or small in scale. Artificial intelligence can be defined as the science of automating intelligent behavior. The infusion of intelligence into the modeling, analysis, identification, design, and control of the system leads to a new generation of systems whose behavior is autonomous. An excellent example of an autonomous control system is NASA's *Pathfinder* spacecraft, which landed on Mars on July 4, 1997. Of course, such systems must be provided with human supervisory control and, in that sense, are not totally autonomous, but they do have a high degree of autonomy when functioning as intended by their human designers. *See* ARTIFICIAL INTELLIGENCE; SPACE PROBE.

Electric power systems. Another example of a large system is a regional power system. The system consists of many generating substations (subsystems) which are geographically separated but are linked through transmission lines. Here, the system has a number of local controllers, one at each substation (Fig. 3), which collectively work to ensure stability and reliability of the overall system. *See* CONTROL SYSTEMS; ELECTRIC POWER SYSTEMS.

Mohammad Jamshidi

Bibliography. Y. Y. Haimes (ed.), *Large Scale Systems*, North Holland, Amsterdam, 1982; M. Jamshidi, *Large-Scale Systems: Modeling and Control*, Elsevier, 1983; M. Jamshidi, *Large-Scale Systems Modeling, Control and Fuzzy Logic*, Prentice Hall, 1996.

Larmor precession

A precession in a magnetic field of the motion of charged particles or of particles possessing magnetic moments.

Charged particles. The Larmor theorem states that, for electrons moving in a single central field of force, the motion in a uniform magnetic field H is, to first order in H, the same as a possible motion in the absence of H except for the superposition of a common precession of angular frequency given by Eq. (1).

$$\omega_L = \frac{eH}{2mc} \tag{1}$$

Here e/c is the magnitude of the electronic charge in electromagnetic units, and m is the electronic mass. The frequency ω_L is called the Larmor frequency and is numerically equal to 2π times 1.40 MHz per oersted or 2π times 111 MHz per SI unit of magnetic field strength (ampere-turn per meter). *See* PRECESSION.

The Larmor theorem is derived in numerous texts. For the special case of an electron moving in a circular orbit of radius r about a fixed nucleus, with H applied normal to the plane of the orbit, the derivation is as follows: The centripetal force holding the electron in orbit must equal $m\omega^2 r$ and is the sum of the Coulomb force Ze^2/r^2 and the Lorentz force, $(e/c)\omega rH$. Therefore Eq. (2) is valid, where ω_0 is the

$$\omega = \pm\left[\left(\frac{eH}{2mc}\right)^2 + \frac{Ze^2}{mr^3}\right]^{1/2} + \frac{eH}{2mc}$$

$$= \pm(\omega_L{}^2 + \omega_0{}^2)^{1/2} + \omega_L \tag{2}$$

angular frequency in the absence of H. If $\omega_0 > \omega_L$ (bound electron, and first order in H) the approximate angular frequency is given by Eq. (3), which

$$\omega = \pm\omega_0 + \omega_L \tag{3}$$

is the Larmor theorem. For a free or unbound electron (no Coulomb force), the approximation breaks down, but direct solution of the equation involving $m\omega^2 r$ and the Lorentz force yields $\omega = eH/mc$. This is twice the Larmor frequency and is called the cyclotron frequency. *See* PARTICLE ACCELERATOR.

In stating the Larmor theorem, use was made of the phrase "a possible motion." if H is applied sufficiently slowly, it can be proved that the motion is the same as in the absence of H, except for the superposition of the Larmor precession. However, a sudden application of H may change, for example, a circular orbit into an elliptical one. For an important application of the Larmor theorem *see* DIAMAGNETISM; ELECTRON MOTION IN VACUUM.

Magnetic moments. According to elementary electromagnetic theory, a current loop of area A and of current I possesses a magnetic moment $\boldsymbol{\mu}$ of magnitude IA and of direction normal to the loop. Thus an electron moving with a velocity v in a circular orbit of radius r, and hence with current $(-e/c) \cdot (v/2\pi r)$ in emu, has an orbital magnetic moment of magnitude as given by Eq. (4).

$$\mu = \frac{-e}{c}\frac{v}{2\pi r}(\pi r^2) = -\frac{evr}{2c} \tag{4}$$

The electron also has orbital angular momentum mvr, which by quantum theory must equal $\hbar J$, where J is an integer and $\hbar$ is Planck's constant h divided by 2π. The ratio of magnetic moment to angular momentum, Eq. (5), is called the magnetogyric (and

$$\gamma_J \equiv \frac{\mu}{\hbar J} = -\frac{e}{2mc} \tag{5}$$

often the gyromagnetic) factor γ_J. *See* ANGULAR MOMENTUM; GYROMAGNETIC RATIO; NONRELATIVISTIC QUANTUM THEORY.

In terms of the equivalent magnetic moment, Eq. (1) may be written in the form of Eq. (6). In

$$\omega_L = -\gamma_J H = -\frac{\mu}{\hbar J}H \tag{6}$$

this form the Larmor precession is exhibited by any magnetic moment $\boldsymbol{\mu}$ including magnetic moments associated with spin angular momentum as well as those associated with orbital angular momentum. Equation (6) may also be derived from equating the time rate of change of angular momentum $(d/dt)(\hbar\mathbf{J})$

to the magnetic torque $\mu \times \mathbf{H}$, as in Eq. (7). In

$$\frac{d(\hbar\mathbf{J})}{dt} = \mu \times \mathrm{H} = \gamma_J \hbar \mathbf{J} \times \mathbf{H} \qquad (7)$$

this form the Larmor precession applies to experiments in molecular beams, electron paramagnetic resonance (EPR), and nuclear magnetic resonance (NMR). *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; ELECTRON SPIN; MAGNETIC RESONANCE.

Rotating coordinate system. Let $(\partial/\partial t)$ represent differentiation with respect to a coordinate system rotating with angular velocity ω. Then differentiation with respect to a stationary observer (d/dt) is given by Eq. (8). Here $\mathbf{J}$ is measured by the stationary observer. This equation may be combined with Eq. (7) in the form of Eq. (9). Here $\mathbf{H}_r$ is the effective field in the rotating coordinate system as given by Eq. (10).

$$\frac{d\mathbf{J}}{dt} = \frac{\partial \mathbf{J}}{\partial t} + (\omega \times \mathrm{J}) \qquad (8)$$

$$\frac{\partial(\hbar\mathbf{J})}{\partial t} = \gamma_J \hbar \mathbf{J} \times \left(\mathbf{H} + \frac{\omega}{\gamma_J}\right) = \gamma_J \hbar \mathbf{J} \times \mathbf{H}_r \qquad (9)$$

$$\mathbf{H}_r = \mathbf{H} + \frac{\omega}{\gamma_J} \qquad (10)$$

Therefore, in a frame which is rotating at the Larmor frequency, the effect of a constant field $\mathbf{H}$ is reduced to zero.

This result, which is an extension of Larmor's original theorem, also holds in quantum mechanics. It is the basis of simplified analyses of the effects of oscillating magnetic fields on particles with charges and magnetic moments. Elihu Abrahams; Frederic Keffer

Bibliography. R. P. Feynman, *The Feynman Lectures on Physics*, 1964; H. Goldstein, *Classical Mechanics*, 2d ed., 1980; E. M. Purcell, *Electricity and Magnetism*, 2d ed., 1984; K. R. Symon, *Mechanics*, 3d ed., 1971.

Larnite

The alpha polymorph of calcium silicate (Ca_2SiO_4). Larnite is a mineral which crystallizes at high temperature. Its occurrences are practically confined to limestone or chalk zones in contact with semimolten balsalts. The presence of magnesium leads to the formation of other structurally related high-temperature minerals such as merwinite ($Ca_3Mg[SiO_4]_2$), bredigite ($Ca_7Mg[SiO_4]_4$), and alite ($Ca_3O[SiO_4]$). These phases are structurally very dense and have been postulated as possible components in the Earth's mantle. At least four polymorphs of Ca_2SiO_4 are known, and a few others are stabilized by impurities. They often occur in silicate slags and clinkers.

Larnite is stable between 1300 and 2600°F (720 and 1400°C) in the CaO-rich portion of the system $CaO–SiO_2$. At room temperature, larnite is metastable and inverts to its low-temperature polymorph calcio olivine through shock. This leads to fall, or distintegration of slags with time, and presents problems in the cement industry. The mineral is very rare, known from its type locality at Scawt Hill, County Antrim, Ireland, and from Crestmore, near Riverside, California. *See* SILICATE MINERALS.

Paul B. Moore

Larynx

The complex of cartilages and related structures at the opening of the trachea, or windpipe, into the pharynx, or throat.

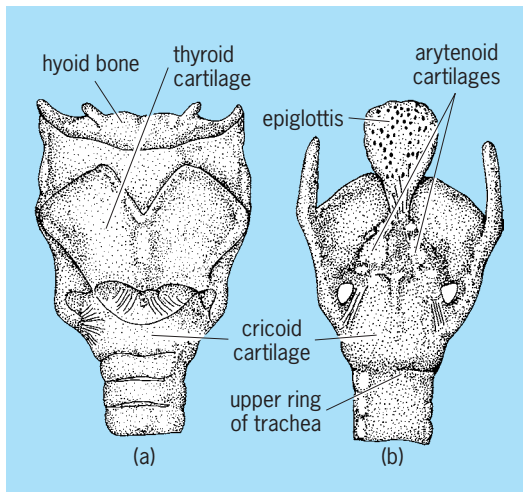

Human laryngeal cartilages and ligaments. (*a*) Front view. (*b*) Back view. Epiglottis acts as lid to larynx.

The cartilages of the larynx (see **illus.**) are derivatives of the primitive gill-bar system. Their original function is to act as a protective sphincter at the entrance of the air passage into the throat, and, when this is the only function, the structure is quite simple, with only one or two pairs of cartilages.

In higher vertebrates the additional function of phonation has been acquired; it is almost limited to mammals since in birds the sound is produced in the syrinx lower in the trachea, but some frogs and reptiles can produce sounds with their larynges.

In humans and most other mammals, the signet-shaped cricoid cartilage forms the base of the larynx and rests upon the trachea. The thyroid cartilage, which forms the prominent Adam's apple ventrally, lies anterior to the cricoid. Dorsally there are paired pivoting cartilages, the arytenoids. Each is pyramid-shaped and acts as the movable posterior attachment for the vocal cords and the laryngeal muscles that regulate the cords. Two other small paired cartilages, the cuneiform and the corniculate, also lie dorsal to the thyroid cartilage. The epiglottis, a leaf-shaped elastic cartilage with its stem inserted into the thyroid notch, forms a lid to the larynx. *See* LARYNX DISORDERS; SPEECH; THYROID GLAND.

Thomas S. Parsons

Larynx disorders

Diseases of the larynx manifest themselves by hoarseness and by stridor, a form of noisy breathing caused by localized narrowing in the larynx or trachea.

Laryngitis is an inflammation of the mucous membrane of the larynx always associated with hoarseness. It frequently occurs with common colds and as a complication of other inflammatory diseases of the upper respiratory system. In diphtheria, the formation of a membrane of fibrin, leukocytes, destroyed tissue, and bacteria can cause severe respiratory difficulties, which may demand tracheotomy. The development of chronic laryngitis is favored by a chronic irritation such as that caused by smoking.

Hoarseness is also a manifestation of paralysis of the recurrent laryngeal nerve of the vagus, which is easily damaged upon surgical removal of a goiter.

Benign tumors of the larynx occur in the younger age group. A tumorlike formation of the vocal cord, known as singer's node, often accounts for the hoarseness of people who abuse the voice. Cancer of the larynx is not uncommon in humans, but is frequently of a slowly growing type and the outlook is often good. The highest incidence is in males over 60.

Ewald R. Weibel

Laser

A device that uses the principle of amplification of electromagnetic waves by stimulated emission of radiation, and operates in the infrared, visible, or ultraviolet region. The term laser is an acronym for light amplification by stimulated emission of radiation, or a light amplifier. However, just as an electronic amplifier can be made into an oscillator by feeding appropriately phased output back into the input, so the laser light amplifier can be made into a laser oscillator, which is really a light source. Laser oscillators are so much more common than laser amplifiers that the unmodified word "laser" has come to mean the oscillator, while the modifier "amplifier" is generally used when the oscillator is not intended. *See* AMPLIFIER; MASER; OSCILLATOR.

The process of stimulated emission can be described as follows: When atoms, ions, or molecules absorb energy, they can emit light spontaneously (as with an incandescent lamp) or they can be stimulated to emit by a light wave. This stimulated emission is the opposite of (stimulated) absorption, where unexcited matter is stimulated into an excited state by a light wave. If a collection of atoms is prepared (pumped) so that more are initially excited than unexcited (population inversion), an incident light wave will stimulate more emission than absorption, and there is net amplification of the incident light beam. This is the way the laser amplifier works.

As with an electronic amplifier, the condition for oscillation is to feed back, in phase with the input, enough of the output to overcome the losses. A laser amplifier can be made into a laser oscillator by arranging suitable mirrors on either end of the amplifier. Together these are called the resonator. Thus the essential parts of a laser oscillator are an amplifying medium, a source of pump power, and a resonator. Radiation that is directed straight along the axis bounces back and forth between the mirrors and can remain in the resonator long enough to build up a strong oscillation. (Waves oriented in other directions soon pass off the edge of the mirrors and are lost before they are much amplified.) Radiation may be coupled out by making one mirror partially transparent so that part of the amplified light can emerge through it (**Fig. 1**). The output wave, like most of the waves being amplified between the mirrors, travels along the axis and is thus very nearly a plane wave. *See* OPTICAL PUMPING.

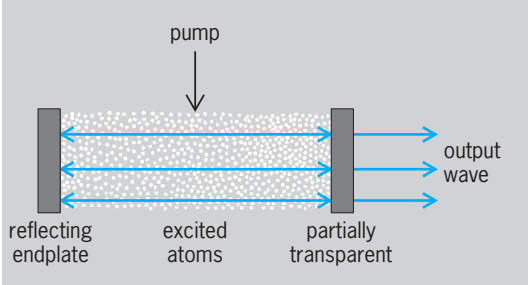

Fig. 1. Structure of a parallel-plate laser.

Comparison with other sources. In contrast to lasers, all conventional light sources are basically hot bodies which radiate by spontaneous emission. The electrons in the tungsten filament of an incandescent lamp are agitated by, and acquire excitation from, the high temperature of the filament. Once excited, they emit light in all directions and revert to a lower energy state. Similarly, in a gas lamp the electric current excites the atoms to high-energy quantum states, and they soon give up this excitation energy by radiating it as light. In all the above, spontaneous emission from each excited electron or atom takes place independently of emission from the others. Thus the overall wave produced by a conventional light source is a jumble of waves from the numerous individual atoms. The phase of the wave emitted by one atom has no relation to the phase emitted by any other atom, so that the overall phase of the light fluctuates randomly from moment to moment and place to place. The lack of correlation is called incoherence.

Hot bodies emit radiation, more or less equally in all directions, whose wavelength distribution is dictated by the Planck blackbody radiation curve. For example, the surface of the Sun radiates like a blackbody at a temperature of about 6000 K, and emits a total of 7 kW/cm^2, spread out over all wavelengths and directions. Light from gas lamps can be more monochromatic (wavelengths radiated are restricted by the quantized energies allowed in atoms), but radiation still occurs in all directions. In contrast, an ideal plane wave would have the same phase all across any wavefront, and the time fluctuations would be highly predictable (coherent). The output of the parallel-plate laser described above is very nearly such a plane

wave and is therefore highly directional. This arises because in the laser oscillator atoms are stimulated to emit in phase with the stimulating wave, rather than independently, and the wave that builds up between the mirrors matches very closely the mirror surfaces. The output is powerful because atoms can be stimulated to emit much faster than they would spontaneously. It is highly monochromatic largely because stimulated emission is a resonance process that occurs most rapidly at the center of the range of wavelengths that would be emitted spontaneously. Since atoms are stimulated to emit in phase with the existing wave, the phase is preserved over many cycles, resulting in the high degree of time coherence of laser radiation.

As a result of this high degree of time coherence, combined with the excellent spatial coherence imposed by the mirror surfaces, lasers can emit vastly more power per unit area and per solid angle than conventional sources. Although such a comparison underscores the spectacular advance in this regard, the difference is not fundamental, but one of magnitude. A truly fundamental difference between laser light and conventional (thermal) light is in the photoelectron statistics when the light is detected. This particular aspect of laser light has yet to be widely utilized, but is crucial for the implementation of quantum computing concepts. *See* COHERENCE; HEAT RADIATION; RADIANCE.

Continuous-wave gas lasers. Perhaps the best-known gas laser is the neutral-atom helium-neon (HeNe) laser, which is an electric-discharge-excited laser involving the noble gases helium and neon. The lasing atom is neon. The wavelength of the transition most used is 632.8 nanometers; however, many helium-neon lasers operate at longer and shorter wavelengths including 3390, 1152, 612, 594, and 543 nm. Output powers are mostly around 1 milliwatt, although at 632.8 nm multimode powers approaching 100 mW can be obtained.

A useful gas laser for the near-ultraviolet region is the helium-cadmium (HeCd) laser, where lasing takes place from singly ionized cadmium. Wavelengths are 325 and 442 nm, with powers up to 150 mW.

The argon ion laser provides continuous-wave (CW) powers up to about 50 W, with principal wavelengths of 514.5 and 488 nm, and a number of weaker transitions at nearby wavelengths. The argon laser is often used to pump other lasers, most importantly tunable dye lasers and titanium: sapphire lasers. Argon lasers have the provision for selecting individual transition wavelengths for oscillation. Single-frequency oscillation can also be produced by inserting an etalon inside the laser resonator. For applications requiring continuous-wave power in the red, the krypton ion laser can provide continuous-wave lasing at 647.1 and 676.4 nm (as well as 521, 568, and other wavelengths), with powers somewhat less than those of the argon ion laser.

The carbon dioxide (CO_2) molecular laser has become the laser of choice for many industrial applications, such as cutting and welding. Oscillation can be obtained on more than 200 vibration-rotation transitions in the 8–18-micrometer wavelength region. The principal transitions are at 10.6 and 9.6 μm, with discrete tuning possible near both. Continuous-wave powers as large as hundreds of kilowatts are obtainable for industrial lasers, making it possible for a factory to use a central, very high power laser and distribute its output as needed. The efficiency of this laser exceeds 10%, which is very desirable for a high-power laser. Terawatts (10^{12} W) of peak power have been obtained in pulsed operation.

Short-pulsed gas lasers. Some lasers can be made to operate only in a pulsed mode. Examples of self-terminating gas lasers are the nitrogen laser (337 nm) and excimer lasers (150 to 400 nm). The nitrogen laser pulse duration is limited because the lower level becomes populated due to stimulated transitions from the upper lasing level, thus introducing absorption at the lasing wavelength. Peak powers as large as 1 MW are possible with pulse durations of 1–10 nanoseconds. Excimer lasers are self-terminating because lasing transitions tear apart the excimer molecules and time is required for fresh molecules to replace them. Other powerful lasers for the vacuum ultraviolet region (100 to 300 nm) use pulsed discharges in high-pressure molecular hydrogen or rare gases, such as xenon. These lasers provide nanosecond pulses with peak powers from tens of kilowatts to a few megawatts.

The nitrogen laser (which is suitable for home workshop laser construction) has proven useful for pumping tunable dye lasers, which lase most readily when pumped with short, intense pulses. Excimer lasers emit at even shorter wavelengths, down to 157 nm. They have proven useful in biological applications because their short wavelengths make possible selective severing of intermolecular bonds. Thus, very fine cutting can be accomplished without excessive heating. The short wavelengths of excimer lasers also allow for high-resolution patterning of photoresist in optical lithography. *See* LASER PHOTOBIOLOGY.

The copper vapor laser is a high-gain neutral-atom laser which operates (pulsed only) at two visible wavelengths, 510.6 and 578.2 nm. Typical pulse durations are 50 ns. A very large copper laser system was built for astronomy applications, which uses a high-repetition-rate laser (5000 pulses per second) followed by three laser amplifiers to achieve an average power of 250 W (peak power of 1 MW). Other applications include pumping large tunable dye lasers for isotope separation and precision machining and cutting, where the short-wavelength diffraction limit is 20 times smaller than for carbon dioxide lasers. *See* ADAPTIVE OPTICS; ISOTOPE SEPARATION.

Solid-state lasers. The term solid-state laser should logically cover all lasers other than gaseous or liquid. Nevertheless, current terminology will be followed, which treats semiconductor (diode) lasers separately from solid-state lasers because the physical mechanisms are somewhat different. With that reservation, virtually all solid-state lasers are optically pumped. Except for some brief remarks about the ruby laser,

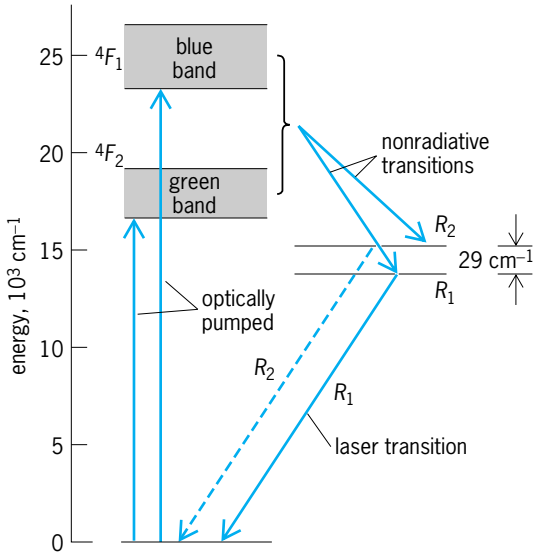

Fig. 2. Energy-level diagram of a ruby laser (three-level system).

the following discussion is limited to the main lasers in current use.

Ruby ($Cr:Al_2O_3$) laser. Historically, the first laser was a single crystal of synthetic ruby, which is aluminum oxide (Al_2O_3 or sapphire), doped with about 0.05% (by weight) chromium oxide (Cr_2O_3). The ruby laser (**Fig. 2**) actually comprises two laser systems with the upper laser levels, R1 and R2. Each is called a three-level system. The major pump bands are blue and green. Unless the ruby laser is pumped by another laser, only pulsed operation is possible. (The first ruby laser was pumped by a flashlamp.) The gain bandwidth is about 300 GHz at room temperature.

Rare-earth lasers and amplifiers. Three important rare-earth laser systems are neodymium:YAG, that is, yttrium aluminum garnet ($Y_3Al_5O_{12}$) doped with neodymium; neodymium:glass; and erbium:glass. Other rare earths and other host materials also find application. In contrast to ruby, the rare-earth ions form four-level systems (**Fig. 3**). This property results in a low pumping threshold for oscillation (about 1000 times lower than for ruby) because it is not necessary to invert the populations of the lower two levels, but only to maintain an inversion between levels 2 and 3. (Level 3 generally empties rapidly.)

1. *Neodymium:YAG laser.* Continuous-wave operation is possible with pumping from flash lamps or efficient diode lasers, typically at a wavelength of 808 nm. The dominant laser transition wavelength is 1064 nm. The gain bandwidth at room temperature is 60 GHz. The overall electrical-to-optical power conversion efficiency with diode laser pumping can exceed 20%. Power scaling of neodynium:YAG lasers is limited to a few kilowatts primarily due to thermal lensing. The laser radiation can be converted efficiently by frequency doubling or frequency tripling to yield green (532 nm) or blue (355 nm) light.

2. *Neodymium:glass laser.* In a typical glass host, neodymium has broader pump bands than in crystalline YAG, which enhance broadband pumping, but also result in a broader gain bandwidth (3000 GHz), which reduces the maximum gain. As a result, neodymium: glass lasers generally have a higher lasing threshold than neodymium:YAG lasers, but prove to be an excellent laser oscillator and amplifier particularly well suited for laser pulses with picosecond to nanosecond duration. Since glass has 40 times poorer thermal conductivity than YAG, it is difficult to remove heat. This severely limits the repetition rate for neodymium:glass lasers. On the positive side, large and uniform glass "logs" can be made by using superb glass technology (**Fig. 4**). Q-switched pulses with peak powers greater than 100 TW are possible.

3. *Erbium:glass laser.* High-bandwidth optical communication systems utilize optical carrier waves in the spectral region around 1550 nm and fiber-optic

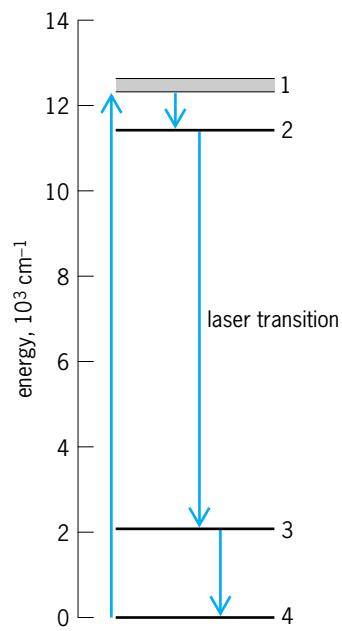

Fig. 3. Energy-level diagram of a rare-earth laser based on the Nd ion in YAG (four-level system). Broadband optical pumping is primarily to the broad, uppermost level. Alternatively, this level can be pumped by efficient semiconductor lasers.

Fig. 4. High-power laser amplifier stage, with a 30-cm (1-ft) aperture, using liquid-cooled slabs of neodymium:glass. (*Lawrence Livermore National Laboratory*)

cable with extremely low transmission losses at these wavelengths. Because of the broad gain band of erbium:glass in this wavelength region, this material is suited for 1550-nm glass lasers and the erbium-doped fiber amplifier (EDFA). Erbium:glass amplifiers and lasers are conveniently pumped by semiconductor diode lasers. The EDFA is a key element in optical communications. It is used to compensate for residual losses of long transmission lines, thereby replacing electronic repeaters which required transforming light to electricity and then back to light. Since signals at many carrier wavelengths (communication channels) can be amplified simultaneously in a broadband EDFA, these devices greatly improve the performance of optical communication systems.

Conventional rare-earth lasers and amplifiers use pencil-sized crystals or glass rods. Slab, thin-disk, and fiber lasers utilize special shapes and sizes of the amplifying medium to improve the laser performance. Slab, thin-disk, and fiber lasers provide much improved heat dissipation compared to traditional laser rods, making them suitable for high-power continuous-wave operation. Thin-disk lasers can generate output powers greater than 1 kW with optical pumping efficiencies up to 50% when pumped by high-power laser diodes.

Fiber lasers use guided-mode propagation for pump and signal light waves to create extremely robust laser structures that do not require any alignment after fabrication. Single-mode signal wave guiding in optical fibers results in diffraction-limited output beams from laser resonators that can be longer than 1 km. Fiber lasers can use all traditional rare-earth laser systems such as neodymium, erbium, or ytterbium. With single-mode output beams of more than 2 kW, continuous-wave power, and efficiencies over 80%, laser-diode-pumped ytterbium:glass fiber lasers are the most prominent high-power fiber lasers. Their emission wavelength of slightly above 1 μm is similar to neodymium:glass lasers. High-power laser diodes are used to pump the ytterbium ions at a wavelength of 980 nm. Erbium:glass fiber lasers that operate around 1550 nm can be used in fiber-optics-based communication and sensing systems that benefit from very long coherence lengths of fiber laser devices. *See* OPTICAL COMMUNICATIONS.

Titanium:sapphire ($Ti:Al_2O_3$) laser. The titanium-doped sapphire laser has become a very important solid-state laser, chiefly by virtue of its large gain bandwidth. Pumped by argon ion or diode lasers, it can be tuned from 660 to 1180 nm, more than 10 times the range of a single-dye laser such as rhodamine 6 G (570 to 610 nm). Commercial models are available with frequency doubling, tripling, and fourth-harmonic generation, respectively, covering the spectral ranges 350–450, 240–300, and 210–225 nm. Titanium:sapphire lasers can be mode-locked to produce femtosecond (10^{-15} s) pulses. Femtosecond titanium:sapphire oscillators typically generate 1 W of average power and, because of their large bandwidth and high saturation fluence, amplified titanium:sapphire laser systems can produce laser pulses with petawatt (10^{15} W) peak power.

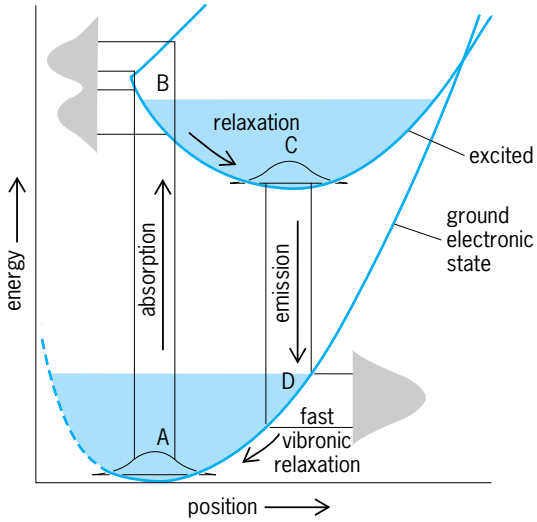

Fig. 5. Energy levels for a titanium-doped sapphire ($Ti:Al_2O_3$) laser. The energy of electronic states is plotted with respect to the position of the titanium ion. Shaded regions indicate continua of vibrational levels.

Since the structure of titanium:sapphire ($Ti:Al_2O_3$) seems similar to that of ruby ($Cr:Al_2O_3$), it might be supposed that titanium:sapphire is just a minor variation on the ruby laser, where one transition element, titanium, is substituted for another, chromium. It turns out that in this case the effect is not minor at all, for it results in a several-thousand-fold increase in the laser gain bandwidth. The reason has to do with the way that titanium ions orient themselves in the aluminum oxide lattice. When titanium ions in this lattice are in their excited states, the overall energy of the system can be lowered if the titanium ion is displaced with respect to the surrounding oxygen atoms (the Jahn-Teller effect). When the titanium ion moves to its new position, it excites lattice vibrations, or phonons (**Fig. 5**). Hence this laser is called a vibronic laser, as is the organic dye laser. When the titanium ion either absorbs or emits light, an atomic electron within the titanium ion reacts more quickly than the heavier titanium nucleus can move (the Franck-Condon principle). Thus, such electronic transitions can be represented as vertical lines in an energy-level diagram (Fig. 5). Gaussian-shaped curves (at points A and C in Fig. 5) can be used to represent the probability of finding the titanium ion at a particular position in the lowest vibrational state of the ground and excited electronic levels, respectively. When this laser is optically pumped in the blue-green region (transition A to B), the excited electron relaxes (to point C) by emitting phonons. Emission (from point C to D) can take place over a wide range of energy differences, which means a broad emission spectrum and wide gain bandwidth. In summary, the huge increase in gain bandwidth of the titanium:sapphire laser over that of the ruby laser is due to the Jahn-Teller effect. *See* FRANCK-CONDON PRINCIPLE; JAHN-TELLER EFFECT.

Semiconductor (diode) lasers. The semiconductor laser is the most important of all lasers, both by

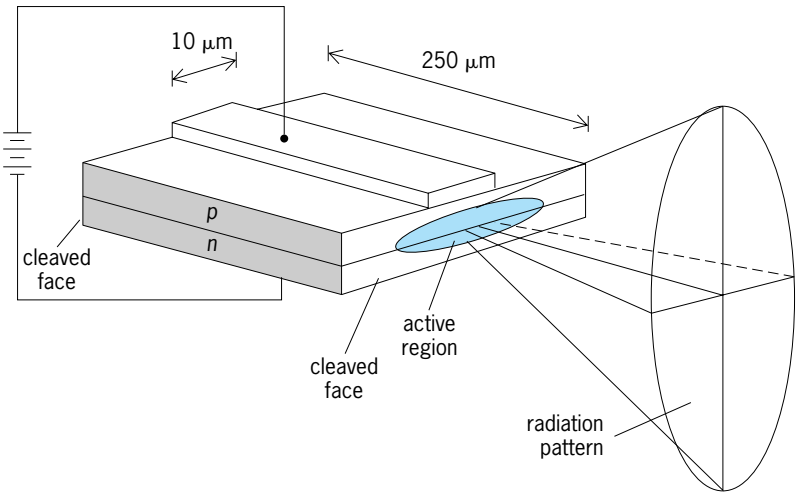

Fig. 6. Semiconductor *pn*-junction laser with horizontal cavity formed by cleaved end facets.

economic standards and by the degree of its applications. Its main features include rugged structure, small size ($250 \times 10 \times 50\ \mu$m), high efficiency (sometimes exceeding 50%), direct pumping by low-power electric current (typically 15 milliamperes at 2 V, which makes it possible to drive it with transistor circuitry), ability to modulate its output by direct modulation of the pumping current at rates exceeding 20 GHz, compatibility of its output beam dimensions with those of optical fibers, feasibility of integrating it monolithically with other semiconductor optoelectronic devices to form integrated circuits, and a manufacturing technology that lends itself to mass production. *See* INTEGRATED OPTICS.

Most semiconductor lasers are based on III-V semiconductors. The laser can be a simple sandwich of *p*- and *n*-type material (**Fig. 6**) such as gallium arsenide (GaAs). The active region is at the junction of the *p* and *n* regions. Electrons and holes are injected through an electrical current into the active region from the *p* and *n* regions, respectively. Light is amplified by stimulating electron-hole recombination. A typical edge-emitting laser has a horizontal laser cavity that can be formed by cleaved-end facets of the chip, distributed Bragg reflectors (DBR), or a distributed feedback (DFB) scheme. *See* ELECTRON-HOLE RECOMBINATION; SEMICONDUCTOR; SEMICONDUCTOR DIODE.

It has been found that a sandwich of *p*-GaAlAs/GaAs/*n*-GaAlAs between the *p*-GaAs and *n*-GaAs layers better confines the light to the gain region and results in lower lasing thresholds. This device is called a double-heterostructure gallium arsenide–gallium aluminum arsenide (GaAs-GaAlAs) laser. GaAs-GaAlAs double-hetrostructure laser diodes reach power levels on the order of 100 mW, and lasing occurs around the wavelength of 800 nm.

Still lower lasing thresholds and higher efficiencies can be achieved using semiconductor nanostructures that provide structural confinement for electrons and holes in one, two, or three-dimensions. Most common are quantum-well structures where the carriers are confined through a potential barrier within nanometer-thin layers called quantum wells. The potential barrier is created through a change in material composition, that is, by including a third (or sometimes fourth) material such as aluminum, either in the barrier or in the well composition. The lowest lasing thresholds have been achieved in quantum-dot lasers, where the carriers are confined in all directions. Such three-dimensional confinement can be obtained through lithographic patterning or special growth modes leading to nanometer-sized semiconductor clusters. *See* ARTIFICIALLY LAYERED STRUCTURES; QUANTIZED ELECTRONIC STRUCTURE (QUEST); SEMICONDUCTOR HETEROSTRUCTURES.

Semiconductor laser diodes can emit at wavelengths from ultraviolet to the infrared with the use of a large variety of materials. Active semiconductor laser structures can include ternary as well as quaternary material compositions. The shortest wavelength ultraviolet diodes use gallium nitride (GaN)–based structures. Indium phosphide (InP)–based compositions are the material of choice at the important range of wavelengths in the optical communications window around 1550 nm. Further in the infrared, even longer wavelengths can be reached with quantum cascade lasers. In quantum cascade lasers, electrons cascade down an energy staircase between neighboring quantum wells, emitting a long-wavelength photon at each step. This is in contrast to all other laser diodes that have stimulating transitions between electron states in conduction bands and hole states in valence bands.

Broad-area semiconductor diode stripes can be fabricated to boost the output power of laser diodes. Lasers with stripes only a few micrometers wide can generate powers up to 100 mW, a 50-μm stripe can generate 500 mW, and more than 1-cm-wide stripes can generate more than 10 W. Monolithic arrays containing dozens of stripes can be manufactured. Power scaling through increasing stripe width, however, results in degradation of the optical beam quality. An alternative approach to generating high-power, high-quality beams is the master oscillator power amplifier (MOPA) structure. Here the high-quality output beam of a master oscillator is coupled into an optical amplifier which retains a virtually unchanged beam quality. Master oscillators and wide-area power amplifiers can be monolithically integrated into the same semiconductor chip, generating several watts of output.

Configurations different from traditional (horizontal resonator) laser diodes (Fig. 6) have been developed in the form of vertically emitting semiconductor lasers. Here the resonator is perpendicular to the plane of the active region, and the laser resonator is formed by mirrors on the top and bottom of the active region. If both mirrors are integrated directly on the chip, such a structure is called a vertical cavity surface-emitting laser (VCSEL), while a laser with one external output coupler is called a vertical external-cavity surface-emitting laser (VECSEL). In both configurations the light passes through

only a short length of active medium (1 μm or less), resulting in low optical gain per cavity round-trip. Therefore, vertical-cavity lasers require mirrors with relatively high reflectivity. Although VCSELs typically have low continuous-wave output powers (a few milliwatts), their beams are more symmetrical than those of edge-emitting diodes. Furthermore, a VCSEL chip can be fabricated as a two-dimensional array of independently addressable laser elements. Although the more complex VECSEL requires one adjustable mirror and is not a completely monolithic laser device, it has the symmetrical beam shape of a surface emitter, it can generate tens of watts of power using a rather large area of the semiconductor chip, and it can provide wavelength tunability.

Low cost, small size, high power, and high efficiency have made semiconductor laser diodes useful in a wide range of applications, including compact-disk players, fiber-optic communications, optical data storage, laser printing, materials processing, and pumping of other (mostly solid-state) lasers. *See* COMPACT DISK; COMPUTER PERIPHERAL DEVICES; OPTICAL FIBERS; OPTICAL RECORDING.

Monochromaticity. When lasers were first developed, they were widely noted for their extreme monochromaticity. They provided far more optical power per spectral range (as well as per angular range) than was previously possible. It has since proven useful to relate laser frequencies to the international time standard (defined by an energy-level difference in the cesium atom), and this was done so precisely, through the use of optical heterodyne techniques, that the standard of length was redefined in such a way that the speed of light is fixed. (One meter is the distance that light travels in vacuum during 1/299,792,458 s.) In addition, extremely stable and monochromatic lasers have been developed, which can be used, for example, for optical communication between remote and moving frames, such as the Moon and the Earth. *See* FREQUENCY MEASUREMENT; HETERODYNE PRINCIPLE; LASER SPECTROSCOPY; LIGHT.

Tunable lasers. Having achieved lasers whose frequencies can be monochromatic, stable, and absolute (traceable to the time standard), the next goal is tunability. Most lasers allow modest tuning over the gain bandwidth of their amplifying medium. However, the laser most widely used for wide tunability has been the (liquid) dye laser. This laser must be optically pumped either by a flash lamp or by another laser, such as the argon ion laser. Considerable engineering has gone into the development of systems to rapidly flow the dye and to provide wavelength tunability. About 20 different dyes are required to cover the region from 270 to 1000 nm.

The dye laser is being challenged by other tunable lasers which do not require flowing dyes. In addition to the titanium:sapphire laser (described above), there is the narrow-band tunable external-cavity diode laser that can be configured as an edge emitter or as a surface emitter (see VECSEL described above). Since the tuning range of diode lasers is much smaller, separate lasers are needed to cover the range from 670 to 1500 nm (in steps of about 10 nm, with linewidths of about 1 MHz). However, compared to the dye laser, the external-cavity diode laser system is much simpler, more compact, and more efficient. Other broadly tunable lasers that are commercially available are the parametric oscillator and the color-center laser (for the region 0.8–4-μm).

Free-electron lasers. The purpose of the free-electron laser is to convert the kinetic energy in an electron beam to electromagnetic radiation. Since it is relatively simple to generate electron beams with peak powers of 10^{10} W, the free-electron laser has the potential for providing high optical power; and since there are no prescribed energy levels, as in the conventional laser, the free-electron laser can operate over a broad spectral range.

To accomplish an energy transfer from electrons to light over an extended distance, it is necessary that the phase velocity of the light wave equals the velocity of the current in the electron beam. This presents a problem, because in vacuum, light waves propagate at the velocity of light c, whereas the electron velocity is less than c. At microwave frequencies, this problem is solved by not having the wave propagate through vacuum but through a waveguide with conducting boundaries that vary periodically in the direction of wave motion. To match this boundary condition requires that a component of the wave propagates with a phase velocity less than c, and it is this component that can absorb energy from an electron beam. A microwave device based upon this principle is known as a traveling-wave tube. Such an approach does not work very well at infrared and optical wavelengths, because the dimensions of the periodic boundary are less than a wavelength, which requires submicrometer tolerances over meter lengths; and because the electromagnetic field decays exponentially from the periodic boundary in the direction normal to the direction of propagation, which limits the transverse interaction area to submicrometer dimensions. *See* TRAVELING-WAVE TUBE; WAVEGUIDE.

The free-electron laser does provide a solution for short wavelengths. Rather than having periodic boundaries for the wave, the electron beam is caused to move in a periodic fashion, which is usually accomplished by passing the beam through a spatially periodic magnetic field generated by an undulator magnet array (**Fig. 7**). Such a beam has a current component with a velocity greater than the electron velocity, and it is this component that delivers energy to the electromagnetic wave. Submicrometer fabrication tolerances are not needed, and the transverse interaction area can be large. Free-electron lasers have operated over the spectral range from microwave to ultraviolet. The far-infrared portion of the spectrum is a promising region of application since there are presently few alternative sources, free-electron lasers operate well in this region, and there are many interesting problems in physics and chemistry to be investigated. *See* SYNCHROTRON RADIATION.

High-power and short-pulse lasers. For high-power applications, it is interesting to compare gas lasers with solid-state lasers. Because they are condensed

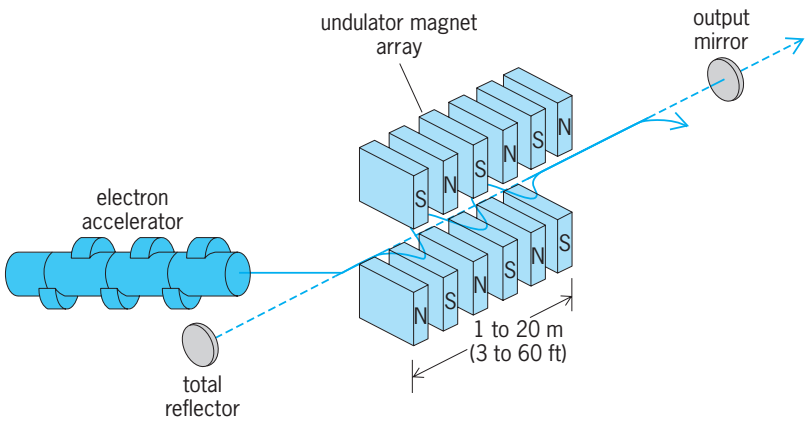

Fig. 7. Free-electron laser. Electrons vibrating in the undulator field amplify the optical beam.

matter, solid-state lasers can have much greater stored energy than gas lasers. However, gas lasers can partially compensate for this through mass transport, that is, replacing used atoms with freshly pumped ones by flowing the gas. The limiting factor in terms of power for gas lasers is the flow rate, which can be supersonic. The limiting factor for solid-state lasers is spontaneous emission lifetime, because flow is not possible, so that stored energy is limited by the time available to pump the same atoms.

For high peak power (pulsed operation), solid-state lasers excel. However, as the repetition rate is increased to increase the average power, the limitation becomes one of laser material damage. For solid-state lasers, the limitation is the heat dissipation rate, while for gas lasers the limitation is the mass transport rate. Gas lasers have the special advantage that if damage occurs, it is self-healing.

Some lasers, notably carbon dioxide and carbon monoxide (CO) lasers, which can emit continuously at modest powers, can also be used to generate intense pulses of microsecond duration when operated at high pressures. At high pressures it is difficult to obtain a uniform electric discharge across the region of lasing. One solution to this problem is the TEA (transverse-electric, atmospheric-pressure) laser, where the electric discharge takes place across the lasing column at pressures of about 1 atm. Still higher powers at higher pressures are obtainable through the ionizing effect of intense bursts of fast electrons from a small accelerator, known as an E-beam.

Short-pulsed operation is achievable from most lasers by a variety of special techniques, including Q-switching and mode-locking. Q-switching involves changing the loss, or Q, of the laser resonator, for example, by effectively misaligning mirrors using a shutter or an optical modulator. The amplifying medium thus becomes excited beyond the ordinary need for lasing, so that additional energy is stored in the medium. At this time, the loss is abruptly removed and the stored energy is released in a giant pulse (1–100 MW, typically lasting 1–30 ns). The Q-switch method works only with lasers whose upper states have long lifetimes. Still shorter and higher-power pulses can be generated by mode-locking or mode generation techniques, sometimes used in conjunction with Q-switching. Modes of a laser oscillator can be synchronized so that the peaks of their light waves occur simultaneously at some instant. The result is a short, intense pulse which quickly ends when the waves of different frequencies get out of step. Mode-locked lasers have generated pulses shorter than 10 femtosecond. Further pulse-shortening techniques have been developed to reach the attosecond (10^{-18} s) regime. *See* OPTICAL PULSES; Q (ELECTRICITY).

The beam may be further intensified by additional stages of amplification, provided the beam diameter is increased by some means such as a telescope, so as to expand (dilute) the beam and thereby prevent damage to laser materials and optics. Sometimes, for solid-state lasers such as neodymium:glass, the amplifying medium is divided into flat slabs separated by cooling liquid (Fig. 4). Here the open faces of the light-amplifying slabs present a large area to receive pumping light and liquid cooling.

Development of very large multistage lasers has been undertaken for research on thermonuclear fusion. In a particularly large laser of this type, the National Ignition Facility at Lawrence Livermore National Laboratory, a single neodymium:glass laser oscillator was designed to drive 192 neodymium:glass amplifier chains made up of together 3072 glass slabs. Each beam will deliver infrared pulses of 10-ns duration and more than 20-kJ energy. After frequency tripling, the combined power of about 500 TW (5×10^{12} W) will be focused onto a millimeter-sized metal target cylinder that contains a deuterium-tritium mixture. The experiment is designed to heat and compress the heavy hydrogen mixture until fusion between hydrogen nuclei is ignited to produce helium, releasing large amounts of nuclear energy. Ultimately this type of controlled laser fusion may become an important source of thermal, electrical, and chemical energy. *See* NUCLEAR FUSION.

Stephen F. Jacobs; Arthur L. Schawlow;
Richard Pantell; Axel Schülzgen;
Nasser Peyghambarian

Bibliography. C. C. Davis, *Lasers and Electro-Optics*, 1996; P. W. Milonni and J. H. Eberly, *Lasers*, 1988; B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics*, 1991; W. T. Silfvast, *Laser Fundamentals*, 2d ed., 2004; J. T. Verdeyen, *Laser Electronics*, 3d ed., 1995; A. Yariv, and P. Yeh, *Photonics: Optical Electronics in Modern Communications*, 6th ed., 2006.

Laser alloying

A material processing method which utilizes the high power density available from focused laser sources to melt metal coatings and a portion of the underlying substrate. Since the melting occurs in a very short time and only at the surface, the bulk of the material remains cool, thus serving as an intimate heat sink. Large temperature gradients exist across

the boundary between the melted surface region and the underlying solid substrate. The result is rapid self-quenching and resolidification.

Transitions. The sequence of schematic cross sections in **Fig. 1** illustrates the transitions occurring during and following an individual laser exposure. In Fig. 1*a*, the metal substrate (B) coated with a thin metal film (A) is irradiated with a laser pulse. A fraction of the incident laser light is absorbed by free carriers within the electromagnetic skin depth of 10^{-7} to 10^{-6} in. (10^{-6} to 10^{-5} cm). For metal surfaces and most laser wavelengths, a significant fraction of the incident light will be specularly or diffusely scattered away (the reflectance process is shown in Fig. 1). The absorbed energy is "instantaneously" (10^{-12} s) transferred to the lattice. The near-surface region very rapidly reaches the melting point, and a liquid-solid interface starts to move through the film (Fig. 1*b*). In Fig. 1*c*, the liquid-solid interface has swept through the original thin-film–substrate interface. Interdiffusion of the film and substrate elements starts. The laser pulse is nearly terminated, and the surface has remained below the vaporization temperature. In Fig. 1*d*, the maximum melt depth has been reached, and interdiffusion continues. The resolidification interface velocity is momentarily zero and then rapidly increases. In Fig. 1*e*, the resolidification interface has moved approximately halfway back to the surface from the melt depth. Interdiffusion in the liquid continues, but the resolidified metal behind the liquid-solid interface cools so rapidly that solid-state diffusion may be neglected. In Fig. 1*f*, the material is completely resolidified, and a "surface alloy" of A in B has been produced. What makes laser surface alloying both attractive and interesting is the wide variety of chemical and microstructural states that can be retained because of the rapid quench from the liquid phase. These include chemical profiles where the "alloyed" element A is highly concentrated near the atomic surface and decreases in concentration over shallow depths (hundreds of nanometers), and uniform profiles where the concentration of A in B is the same throughout the entire melted region. The types of microstructures observed include extended solid solutions (the concentration of A in B greatly exceeds equilibrium values), metastable crystalline phases (high-temperature phases retained because of the rapid return to room temperature), and metallic glasses.

Surface alloying. In general, the output of a laser source must be focused in order to achieve sufficient energy density to induce melting of a metal surface. This focused spot is characterized by its optical spot size. The size of the melted spot (effective spot size) resulting from the laser exposure will generally be smaller than this optical spot size. For a continuous-wave laser source, a melt strip will be produced as the focused beam and metal surface are moved relative to one another. The dwell time, that is, the time that the continuous-wave laser beam irradiates a particular surface point, will strongly influence the depth of the melting. As is schematically shown in **Fig. 2***a*, individual melt strips must be partially

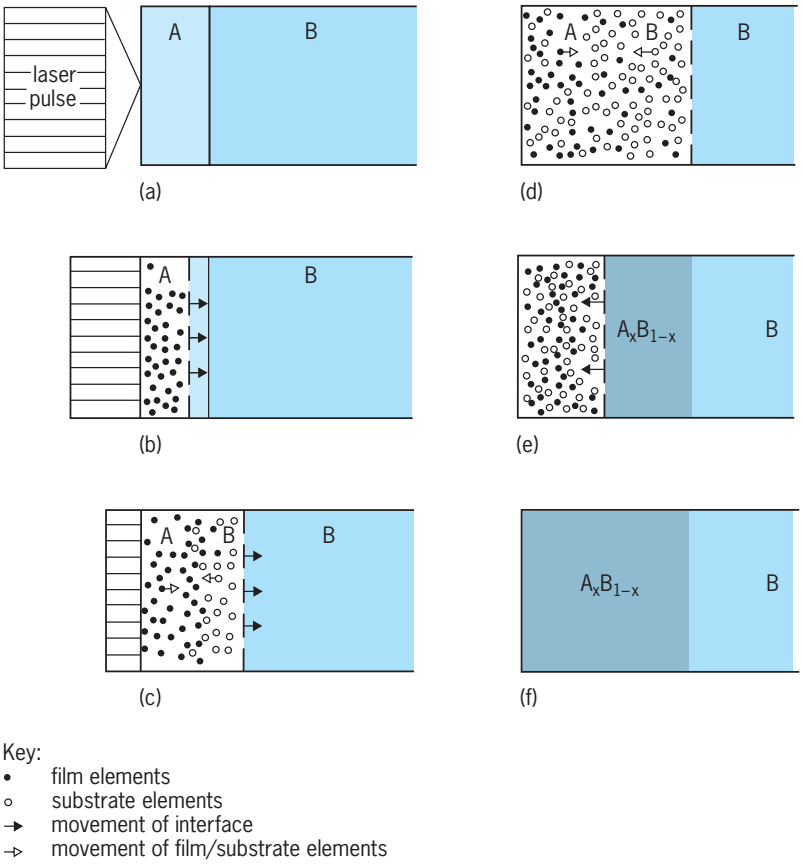

Fig. 1. Sequence of schematic cross sections for laser alloying, with time increasing from *a* to *f*. Ax_xB_{1-x} represents surface alloy (with composition fixed by *x*) of film elements A in substrate elements B.

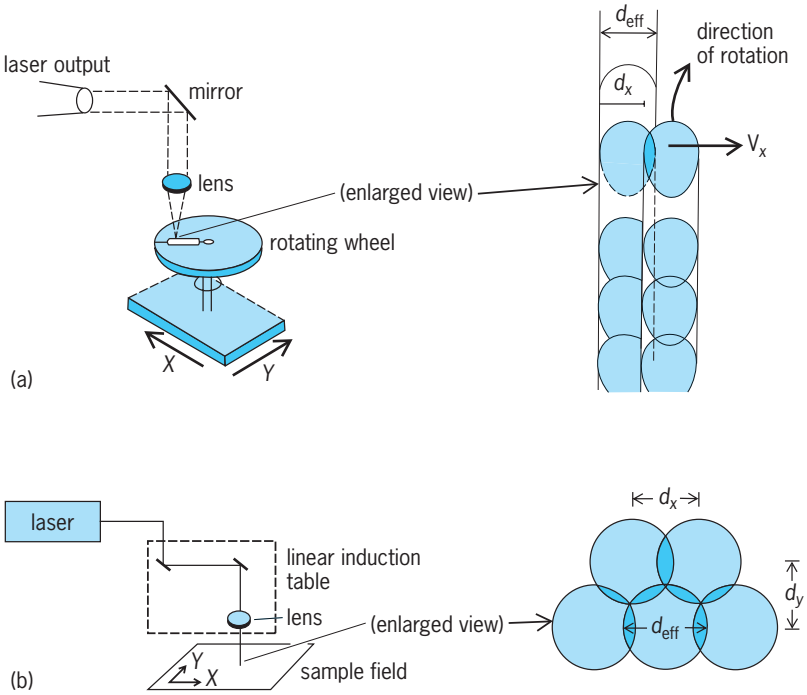

Fig. 2. Schematic drawings and enlarged views of material processing for (*a*) continuous and (*b*) pulsed or Q-switched laser system. d_{eff} = diameter of melt spot or melt strip; V_x = velocity of sample movement in *x* direction; d_x, d_y = displacement between subsequent laser melt strips or spots resulting from relative motion of laser and sample.

overlapped to produce area coverage. The effective spot size, degree of overlap, and relative scanning speed will thus determine the area per unit time which can be produced.

For a pulsed or Q-switched laser (high-power, short-pulse laser) source, there also is an optical and effective spot size, and the "pulse length" is the time of exposure. The laser pulses are emitted in a train of pulses characterized by a repetition rate. As is schematically shown in Fig. 2*b*, this train of pulses is raster-scanned across the metal surface to produce area coverage. For these lasers, the effective spot size, degree of overlap, and repetition rate determine the area per unit time which can be produced.

For all the laser sources (continuous-wave pulsed, and Q-switched), the exposure time (dwell time or pulse length) strongly influences the depth that will be melted. Longer exposure times result in deeper melting. Since deeper melting means a longer total time in the molten state, that means more time available for diffusion of the one or more alloying elements into the molten portion of the substrate. Deeper melting and longer melt times therefore result in more dilute surface alloys, while shallow melting and shorter melt times result in more concentrated surface alloys. It is also evident that in some instances convection, surface tension, and plasma effects can enhance the mixing within the liquid state and drive the melt toward homogenization.

In making laser alloys, many other processing variables need to be considered. In addition to the exposure time just discussed, these include the laser power, the thickness of the film put down prior to laser melting, and in some instances the nature of the gaseous ambient during the laser processing. The processing variables are interrelated, and one variable cannot be freely changed without affecting another. Another consideration is that laser alloying is a liquid state–rapid quenching phenomenon. The near-surface region must be melted and yet vaporization avoided. Different minimum and maximum energy densities are thus defined for each laser exposure time. In addition to these processing constraints, there are certain properties of matter which strongly influence whether or not certain element combinations may be laser-alloyed. For example, it is not possible to laser-alloy a low melting point–high vapor pressure element like zinc (Zn) into a high melting point–low vapor pressure metal substrate such as tungsten (W). The zinc would vaporize before the underlying tungsten could be melted. Since liquid-state intermixing is required, suitable systems must exhibit miscibility in the molten state. Binary systems like silver-nickel (Ag-Ni), iron-lead (Fe-Pb), copper-molybdenum (Cu-Mo), and aluminum-bismuth (Al-Bi) have miscibility gaps in the liquid state spanning nearly all compositions. Such systems cannot be laser-alloyed.

Advantages. Surface alloys, particularly laser surface alloys, have advantages that include tailoring surface properties, conservation of materials, and creation of new metal surfaces.

Surface alloying allows alteration of the metal surface in order to achieve characteristics best suited to the service environment. For example, in an ordinary saw blade the material requirements for the cutting teeth are very different from those of the length of the blade. In many applications requiring corrosion or oxidation resistance, that resistance is needed only on the external surface of the material. It is possible to design the bulk of the material with characteristics most suitable for structural needs or ease of fabrication and to tailor the metal surface for interface requirements. Material conservation is another reason often cited for considering surface alloying. In the case of stainless steels and superalloys, the elements added to impart the special properties are often strategic elements for which there are sparse domestic supplies. For example, chromium (Cr) and nickel (Ni) may not be necessary on the inside of knives, forks, and spoons. Further, precious metals are expensive and in relatively short supply. Thus surface alloying can be very cost-effective.

These first two advantages—tailoring for surface properties and material conservation—are generally important considerations in almost any coating technology. However, laser alloying differs from coating technologies in that the near-surface region is a continuous extension of the interior of the metal. There is no interface between bulk and "coating"; the laser alloy is a mixture of bulk and surface elements. Also, problems in regard to porosity and adherence do not exist. Synergistic benefits may be derived which are not realized by a coating alone. Laser alloying involves very large temperature gradients and quenching from the liquid state. In this way it resembles other rapid-solidification technologies. The thermodynamic constraints which limit the conventional metallurgist do not necessarily apply. It is anticipated that laser alloying will produce many new alloys which could not have been produced by conventional methods. *See* ALLOY; LASER; METAL COATINGS.

Clifton W. Draper

Bibliography. C. W. Draper, Laser surface alloying: A bibliography, *Appl. Opt.,*20(18):3093–3096, September 1981; C. W. Draper, Laser surface alloying: The state of the art, *J. Met.,*34(6):24–32, June 1982; D. S. Gnanamuthu, Laser surface treatment, *Opt. Eng.,* 19(5):783–792, 1980; D. K. Sood, Metastable surface alloys produced by ion implantation, laser of electron beam treatment, *Radiat. Effects,* 63(1):L141–167, January 1982.

Laser cooling

Reducing the thermal motion of atoms with the force exerted by a laser beam. Typically, such cooling is used to reduce the temperature of a gas of atoms, or the velocity spread of atoms in an atomic beam.

Light affects atomic motion when the atoms absorb or emit photons, the particles or quanta that make up light. Photons carry momentum $p = h/\lambda$, where h is Planck's constant and λ is the light's

wavelength. By conservation of momentum, when an atom absorbs or emits a photon, the atom's momentum must change by an amount equal to the photon momentum. Each absorption or emission thus gives the atom a tiny kick, changing its velocity. For most atoms this change is only a few millimeters to a few centimeters per second, while atoms in a gas at room temperature have velocities of a few hundred to a few thousand meters per second. Nevertheless, repeated absorption and emission of photons can have a significant effect on even hot atomic gases or beams. *See* CONSERVATION OF MOMENTUM; LIGHT; MOMENTUM; PHOTON.

Doppler cooling. The keys to using such repeated kicks to reduce the random, thermal motion of a gas of atoms are the monochromatic nature of laser light, the selectivity of absorption of light by atoms, and the Doppler effect. Light is an oscillating electromagnetic wave whose frequency of oscillation determines its color. The energy of each photon is $E = h\nu$, where ν is the frequency. Laser light can have nearly a single frequency or color, so that all the photons have almost identical energies. Atoms absorb only photons whose energy is equal, within a small range, to the difference in energy between two of its quantum states or energy levels. For sodium atoms this resonance frequency is $\nu_0 \equiv 5 \times 10^{14}$ Hz (wavelength $\lambda \equiv 589$ nanometers), but the absorption is efficient only over a range $\Delta\nu = 10^7$ Hz. Moving atoms, however, experience a Doppler shift so that, depending on their speed and whether they are moving along the direction of the laser beam or against it, the light appears to room-temperature atoms to have a frequency shifted up or down by a hundred or more times the natural absorption width $\Delta\nu$. *See* DOPPLER EFFECT; LASER.

If the frequency ν of the laser is tuned to be slightly lower than ν_0, those atoms moving against the laser beam see the laser upshifted, closer to ν_0. These atoms are more likely to absorb photons, receive kicks opposite to the direction of their velocity, and slow down. After absorbing a photon, the atoms are in an excited state and return to the original state by spontaneously emitting a photon. Such photons are radiated in random directions, so the effect of their kicks averages to zero. For atoms held in a trap, as ions generally are, any trapped atom will at some time be traveling against the laser beam and be cooled. Laser cooling was first demonstrated in 1978 with such trapped ions. For free atoms, another, similarly tuned laser beam is added, aimed in the opposite sense, to cool those atoms moving in the opposite direction. More generally, one uses three pairs of mutually perpendicular, counterpropagating laser beams, all tuned below ν_0. Then, no matter the direction of an atom's velocity, there are one or more laser beams that oppose the velocity and slow the atom. Because atoms are slowed as if by friction in a viscous fluid, this configuration of laser beams is called optical molasses.

Cooling limit. Laser cooling's viscosity does not stop atoms. The discrete nature of the kicks imparted by the photons gives the atoms a random motion with a temperature as low as $h\Delta\nu/(2k_B)$, where k_B is Boltzmann's constant. For the example of sodium, this temperature is 240 microkelvins. Surprisingly, in 1988 much lower temperatures were observed in sodium and other atoms. The lower temperatures (as low as 1 μK for cesium atoms) were eventually explained as resulting from additional cooling processes that depend on the atom having multiple energy levels. Such temperatures, the lowest steady-state kinetic temperatures ever achieved for any substance in three dimensions, are close to the recoil limit, or the kinetic energy imparted to an atom from the kick of a single photon. While this might seem to be a fundamental limit, techniques have been developed that go even further, by arranging that when, by chance, an atom achieves a very low velocity it stops absorbing light and remains cold. In this way, atoms have been cooled below the recoil limit.

Applications. Improving atomic clocks, where the thermal motion of atoms reduces the precision and accuracy, was a major motivation to developing laser cooling. Clocks using laser-cooled trapped ions or free neutral atoms rival the performance of the best conventional atomic clocks. Laser cooling is also used in atom optics, where well-collimated, monoenergetic atomic beams are more easily and effectively manipulated. In addition, laser cooling has been used to study collisions between very slow atoms. *See* ATOMIC CLOCK; SCATTERING EXPERIMENTS (ATOMS AND MOLECULES).

Laser cooling is intimately connected with trapping of atoms, because atoms must often be slowed down before they can be held in a trap and because atoms must often be trapped in order to observe laser cooling or its effects. Such effects include cold, trapped ions arranging themselves into a crystal because of the electric repulsion between the charged ions. Neutral atoms can become arrayed on an optical lattice of tiny traps formed by interference between the laser beams used to cool them. In both cases, the spacing between atoms is thousands of times larger than the spacing in solid crystals. Another effect is Bose-Einstein condensation, wherein a gas of atoms whose de Broglie wavelength is comparable to the spacing between atoms has a transition to a state where a significant fraction of the atoms are in the lowest kinetic energy state possible. This phenomenon was first observed for a dilute atomic gas in 1995. *See* BOSE-EINSTEIN STATISTICS; PARTICLE TRAP; QUANTUM MECHANICS.

William D. Phillips

Bibliography. E. Arimondo, W. D. Phillips, and F. Strumia (eds.), *Laser Manipulation of Atoms and Ions*, 1992; S. Chu, Laser trapping of neutral particles, *Sci. Amer.*, 266(2):71–76, February 1992; S. Chu and C. Wieman (eds.), Laser cooling and trapping of atoms (feature issue), *J. Opt. Soc. Amer.*, B11:2019–2278, 1989; C. N. Cohen-Tannoudji and W. D. Phillips, New mechanisms for laser cooling, *Phys. Today*, 43(10):33–40, October 1990; D. J. Wineland and W. M. Itano, Laser cooling, *Phys. Today*, 40(6):34–40, June 1987.

Laser photobiology

The interaction of laser light with biological molecules, and the applications to biology and medicine. The spatial coherence of laser emissions makes it possible to focus beams in small volumes. Consequently, high-power-density irradiations can be obtained, especially with lasers whose energy is delivered in a few nanoseconds or even picoseconds. These two characteristics, combined with the great variety of wavelengths available, have opened up many possibilities of applications to biology and medicine.

Physical aspects. In the visible range, absorption of laser light by biological material is caused only by pigments such as chlorophyll, hemoglobin, cytochromes, or melanin, while ultraviolet laser light is absorbed by nucleic acids and proteins. At low intensity, the absorption of photons produces photochemical effects similar to those obtained with ordinary light sources. By increasing the absorbed energy, heating becomes important, giving rise to induced thermochemical effects (that is, thermal denaturations) that are often useful in biological applications. When the energy density on the target becomes very high, the radiation pressure becomes nonnegligible. Moreover, some nonlinear phenomena may occur, such as optical saturation, two-photon absorption, or dielectric breakdown (at 10^{10} W/cm^2), which lead to ionization and nonselective destruction of the material. *See* CHLOROPHYLL; CYTOCHROME; HEMOGLOBIN; NUCLEIC ACID; PROTEIN.

Laser spectroscopy. This method is used to probe biological processes involving very fast reactions or to study structural changes of complex molecules. The techniques are flash photolysis (in the nanosecond and picosecond range) and Raman spectroscopy.

Many fast reactions can be studied with flash photolysis because lasers provide extremely intense light pulses of very short duration. A large number of excited molecules is therefore produced in a short time. The two main examples of processes studied with this technique are photosynthesis and vision, but the technique has also been applied to study the excited states of many biological molecules. *See* PHOTOCHEMISTRY; PHOTOLYSIS.

Resonance Raman spectroscopy provides information about the vibrational levels of molecules. With lasers used instead of mercury lamps, the quality of the Raman spectrum is improved. *See* RAMAN EFFECT; SPECTROSCOPY; ULTRAFAST MOLECULAR PROCESSES.

Microirradiation of cells. Microirradiation is a useful technique for the study of the photosensitivity of various parts of a single cell. Alteration of a specific organelle after irradiation permits the study of cell functions. This technique has had a wide application in plant cells as well as animal cells. In some cases, animal cells can be stained with supravital dyes to confer absorption at the wavelength of irradiation. Because of the very high power density of the focused laser light, extremely low concentrations of dye can be used, and toxic effects can consequently be avoided. The laser beam is focused through the objective of a microscope onto the cell. The diameter of the beam depends on the divergence of the laser and on the focal length of the objective. In practice, it is easy to obtain spots as small as 1 micrometer in diameter. Ruby, neodymium, and argon lasers are used for this purpose. Ultraviolet excimer lasers are also valuable tools when ultraviolet light and high power density are required.

The organelles which have been the most studied are mitochondria and chromosomes, although other organelles such as mitotic organelles and myofilaments have been irradiated.

Mitochondria. Mitochondria of cells in tissue culture have been irradiated either after staining with Janus green or by using directly the natural absorption of flavins and cytochromes which are present in large amounts in cell types such as heart muscle cells. By using Q-switched lasers, thermal effects are observed in Janus green–stained mitochondria, while with continuous lasers photochemical effects are observed in unstained mitochondria. The possible role of mitochondria in maintaining the beating rhythm of myocardial cells has been investigated in both conditions. *See* MITOCHONDRIA.

Chromosomes. Irradiation of segments of chromosomes is useful for genetic studies. Deletions of small regions of chromosomes are easily obtained after microirradiation with ultraviolet light or visible light plus acridine orange. By this technique, it is possible to assign genes to specific chromosomes, and to generate "mutants" by damaging small selected areas of deoxyribonucleic acid (DNA). Selected parts of chromosomes can be dissected very accurately with an excimer laser into pieces smaller than 0.5 μm in width. Such a technique can be used for microcloning and preparing DNA probes. In addition, location of chromosomes in the interphase nucleus has been performed after ultraviolet or photosensitized damage of chromatin. *See* CHROMOSOME ABERRATION.

Insertion of foreign genetic material. Because self-healing holes in cell walls and membranes can be produced by laser microbeams, micro-irradiation has been applied successfully for the insertion of foreign DNA. In plant cells where conventional techniques can be used only on protoplasts, laser microbeams enable the introduction of DNA without removing the cell wall. The experiments described for mammalian and plant cells clearly indicate that the introduced DNA can be stably incorporated into the genome of the recipient cells.

Optical traps. Optical trapping of biological objects using a laser microbeam allows the manipulation of cells and cellular organelles. This process has been termed optical tweezers. Indeed, significant forces (radiation pressure) are exerted on microscopic objects illuminated by focused beams. By using a 100× objective, cells can be moved over distances greater than 1 mm in the transverse plane, and over about 70 μm in the axial direction.

In addition, optical traps have been used to study

about chromosome motion and cell mobility or to manipulate organelles within cells. They have also been applied to selective cell fusion in combination with laser microsurgery.

Confocal microscopy. Confocal fluorescence microscopy allows cellular three-dimensional imaging and analysis of any fluorophore within a cell. Selective fluorophores of cellular components are excited with laser light, and a complete imaging of the cell structure is built up by scanning all points in the field of view. Quantitative analysis of image parameters is achieved by an image processing computer. Therefore, it is possible to use confocal scanning microscopy to record images for analysis in cell biology. *See* CONFOCAL MICROSCOPY.

Flow cytometry. Mammalian cells, treated with appropriate fluorescent stains, are made to flow in single file through a continuous-wave laser light, such as an argon laser which causes them to fluoresce. By measuring the fluorescence emission, it is possible to analyze heterogeneous cell populations, measure physiological cell response, and isolate cell preparations having different properties.

Surgery. Continuous-wave lasers have been employed as a "light knife," that is, as a surgical cutting and coagulation tool. Generally, high-power CO_2 and argon lasers are used for this purpose. The laser beam is delivered through a mirror system located in an articulated arm (CO_2 laser) or through optical fibers (argon laser) to an endoscope, and can be easily directed into various body cavities. When the laser energy is focused onto a tissue surface, a small volume of tissue is heated, and thus only this area is "cut off." An advantage of this procedure is that small capillaries are coagulated, preventing hemorrhage resulting from cut blood vessels and providing better healing. Lasers have been especially used for otorhinolaryngology, gynecology, urology, and abdominal surgery.

Ophthalmology. Argon laser photocoagulators are commonly used for treating retinopathies, particularly for the prevention of the detachment of the retina and for diabetic retinopathies. Treatment of glaucoma and cataract with pulsed lasers has been attempted.

Dermatology. Laser irradiation is used for removal of foreign pigments in the skin (tattoos), for treatment of vascular disorders ("wine marks"), and for removal of various pigmented skin lesions. Because of the higher absorption of laser light in the pigmented spots than in the skin, thermal effects from the absorbed energy cause preferential coagulation of the pigment and the local constriction of blood vessels with little change in the unpigmented skin.

Cancer photochemotherapy. Red laser light has been used to destroy cancer cells which were previously photosensitized with derivatives of hematoporphyrin, a drug which preferentially localizes in malignant cells. The basis of this technique, also called photodynamic therapy, is an oxygen-dependent reaction produced after light excitation of hematoporphyrin derivatives, which leads to the death of the stained cancer cells. Photodynamic therapy is used for the treatment of certain superficial small tumors, especially in the bladder and the bronchi.

A diagnostic possibility to detect cancers is given by the fact that hematoporphyrin derivatives fluoresce when irradiated with blue or ultraviolet light; because the drug concentrates in cancer cells, tissues which fluoresce are probably malignant.

Low-power density effects. Among controversial laser use and effect, acupuncture and so-called biostimulation can be mentioned. Low-power helium-neon laser beams have been used instead of silver or gold needles for acupuncture. In other respects, laser biostimulation at low fluences in the red has been claimed to increase DNA synthesis, collagen synthesis in fibroblasts, and cell proliferation. Consequently, helium-neon lasers have also been employed to accelerate wound healing. *See* LASER.

Giuliana Moreno; Christian Salet

Bibliography. K. Atsumi (ed.), *New Frontiers in Laser Medicine and Surgery*, 1983; R. V. Bensasson, E. J. Land, and T. G. Truscott, *Excited States and Free Radicals in Biology and Medicine*, 1993; M. W. Berns, W. H. Wright, and R. W. Wiegand Steubing, Laser microbeam as a tool in cell biology, *Int. Rev. Cytol.*, 129:1–44, 1991; B. W. Henderson and T. G. Dougherty (eds.), *Photodynamic Therapy*, 1992; G. Weber and K. O. Greulich, Manipulation of cells, organelles and genomes by laser microbeam and optical trap, *Int. Rev. Cytol.*, 133:1–41, 1992; M. L. Wolbarsht (ed.), *Laser Applications in Medicine and Biology*, vol. 4, 2000.

Laser photochemistry

A branch of chemistry in which reactions are induced or altered by laser light. The initial part of any photochemical reaction involves an optical transition to some excited state of molecule. These excited states could involve electronic, vibrational, and rotational excitation. The particular photochemical product that results from the absorption of light depends on the specific excited state species created during the irradiation. Thus, the properties of the light source often determine the photochemical product. Lasers have had an immense impact on the field of photochemistry by providing scientists with an intense, polarized, and nearly monochromatic source of light. There are lasers that extend from wavelengths of less than 110 nanometers (vacuum ultraviolet) to more than 100,000 nm (far infrared); for comparison, the entire visible spectrum extends from only 400 nm (violet) to 700 nm (red).

Advantages of lasers. Use of lasers in photochemical applications provides three main advantages over conventional light sources such as discharge or arc lamps. First, lasers are generally more powerful than conventional light sources. A continuous-wave argon ion laser can produce 10 W at 514.5 nm. Pulsed lasers, which compress the light energy into very short time periods (10^{-6} to 10^{-12} s), can generate correspondingly higher peak powers, typically from 10^3 to 10^9 W. These very short, intense pulses of light are used by the photochemist to monitor the time

evolution of excited-state species or the appearance rate of products in a photochemical reaction.

Second, laser light can be collimated into a beam with a very small divergence angle, routinely less than 1/100 of a degree. The high degree of collimation of laser light permits efficient illumination at a chosen point within the sample, which could be far from the light source. This collimation is important in photochemistry. It permits the photochemist to confine the photochemical activation to some very small and precisely located area, such as in the fabrication and repair of microelectronic devices. In addition, the photochemical event, which has a very low absorption cross section, can be induced by photolyzing a very small region of the chemical sample with the total laser output power.

Third, laser light is exceptionally pure in color. The spectral purity of light can be described by a band width measured in wavenumbers (cm^{-1}), defined as the frequency width divided by the speed of light. While the full visible spectrum is 10,000 cm^{-1} wide, a typical laser may be as narrow as a few cm^{-1}. This particular advantage of laser light has had its largest impact in photochemistry. It is difficult to obtain spectral purity while maintaining high power with conventional light sources. Consequently, when conventional sources are used to irradiate molecules that have many possible excited-state transitions, a distribution of excited molecules results. These molecules may have been excited to several different electronic states with many possible vibrational or rotational energies. Under such conditions, it is often impossible to identify the excited electronic states that produced the various photochemical products. The high spectral purity of the laser obviates this problem, and so the identity of the excited electronic state is almost always known to the laser photochemist.

The high spectral purity can be a disadvantage when the laser wavelength does not correspond to the required absorption wavelength for the molecule under study. This difficulty has been largely solved in the visible range, where organic dye lasers permit continuous tuning over a significant portion of the spectrum. Nonlinear, pulsed laser techniques, such as frequency doubling, metal-vapor four-wave mixing, stimulated Raman scattering, and optical parametric oscillation, have extended the accessible frequencies to include the vacuum ultraviolet and near-infrared region of the spectrum. *F*-center, diode lasers, and difference-frequency techniques generate continuous-wave radiation over a large portion of the near infrared.

Because of the advantages of lasers, there is a constant demand for the development of lasers that will serve in new regions of the spectrum. In spite of the advantages of lasers, conventional light sources are still widely used in most photochemical investigations and commercial processes because of their simplicity and low cost. *See* COLOR CENTERS; RAMAN EFFECT.

The two general areas of laser-induced photochemistry are vibrational photochemistry and electronic photochemistry. In vibrational photochemistry, the chemical reaction occurs entirely on the ground electronic state of the molecule, whereas electronic photochemistry occurs via some excited electronic state, usually involving the first excited singlet state (all electrons spin-paired) and possibly the lowest-lying triplet state (two electrons spin-unpaired) for most polyatomic molecules. Electronic photochemistry is subject to the selection rules that govern optical electronic transitions. Although electronic photochemistry has been profoundly affected by the advent of lasers, vibrational photochemistry, which requires intense light sources, is not even possible without lasers. *See* ELECTRON SPIN; TRIPLET STATE.

Vibrational photochemistry. In vibrational photochemistry, the chemical reaction occurs on the ground electronic state of both the reactant and product. With this technique, the laser light excites molecules to energetic vibrational states lying above the reaction barrier. By varying the total vibrational energy within the molecule, the photochemist can learn about the energy dependence of the collisional deactivation process and the rate constant for reaction. Laser photochemistry can be used to investigate the interplay between the thermochemical factors and the rate constants of reactions that can also be initiated by thermal activation. This is particularly interesting when there are competitive reaction channels.

Infrared multiphoton photochemistry. Both isomerization and dissociation reactions can be induced with intense infrared lasers. In this photoactivation technique, which has been used to study many different reactions, the first few photons are absorbed more or less resonantly by vibration-rotation states. It is here that it becomes possible to selectively excite one isotope within an isotopic mixture. The excitation scheme for multiphoton and single-photon methods is shown in **Fig. 1**. *See* PHOTON.

As the molecule becomes more energetic, the number of possible vibration-rotation states increases sharply. This is known as the vibrational quasi-continuum. Here the subsequent absorption cross sections are small; but with an intense carbon dioxide (CO_2) laser the molecule will continue to absorb photons and eventually become energetic enough to react. It is often difficult to determine the exact number of photons absorbed by the molecules. Indeed, sometimes the molecules are overexcited beyond the activation barrier so that the reaction rate exceeds the thermal rate. If two dissociation

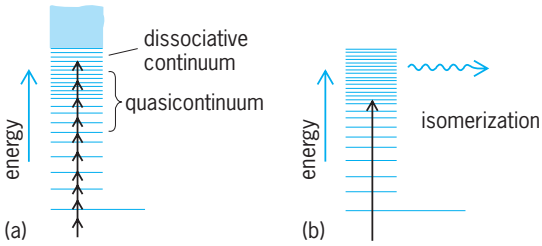

Fig. 1. Excitation scheme using intense infrared lasers. (*a*) Multiphoton. (*b*) Single photon. The black arrows represent the absorption of a photon.

reactions have similar activation barriers, it is possible to change the ratio of the two products by increasing the excitation rate of the laser. For example, there are two dissociation channels for difluorodichloromethane (CF_2Cl_2) that lie close in energy, as shown in reaction (1),

$$CF_2Cl_2 \rightarrow CF_2Cl + Cl \quad \Delta H = 80 \text{ kcal/mole}$$
$$CF_2Cl_2 \rightarrow CF_2 + Cl_2 \quad \Delta H = 71 \text{ kcal/mole} \qquad (1)$$

where ΔH = enthalpy change. When this reaction is initiated by laser multiphoton infrared excitation, the relative yield of $CF_2\cdot$, as measured by resonant fluorescence, depends on the total reaction yield. As the total reaction yield is increased, the amount of $CF_2\cdot$ formed by multiphoton dissociation first increases, then decreases. This change in the relative product yield indicates that the reactant was overexcited above the 71 kcal/mole barrier to a level of excitation where both reaction channels can compete. *See* ENTHALPY.

Single-photon photochemistry (overtone excitation). A method of photoactivation known as laser vibrational overtone activation induces highly excited vibrational states within a molecule by single-photon absorption. In contrast to multiphoton excitation, the exact energy of the activated molecule is known from the wavelength of the excitation laser. Because the molar extinction coefficient ϵ is very small ($\epsilon \simeq 10^{-4}$ or 10^{-5} in contrast to normal allowed electronic transitions, where $\epsilon \simeq 10^4$ to 10^6, or typical infrared transitions, where $\epsilon \simeq 1$ or 10), the gaseous sample is placed within the laser cavity, where the number of photons is a thousand times higher than the laser output. Typically, a tunable dye laser operating in the visible region is tuned to the vibrational overtone transition wavelength, and the photolysis products are detected by gas chromatography after a timed photolysis period. Since each type of carbon-to-hydrogen (C-H) bond (methyl, vinyl, methylenic, and acetylenic) in the vibrational overtone spectra of hydrocarbons absorbs at a characteristic frequency, the specific moieties within the molecule can be excited by tuning the laser wavelength. *See* GAS CHROMATOGRAPHY; PHOTOLYSIS.

Photochemists hoped that laser vibrational overtone activation would provide the long-sought mode-selective chemistry. The idea of mode-selective chemistry involves forcing the molecule to undergo one reaction rather than another by tuning the laser wavelength from one vibrational absorption feature to another. Many unsuccessful attempts have been made to achieve this. One elegant experiment involved 1-cyclopropylcyclobutene (**1**), which isomerizes to 2-cyclopropyl-1,3-butadiene (**2**) as shown in reaction (2).

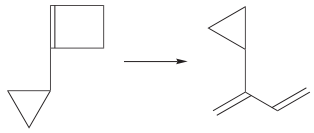

(2)

(**1**) (**2**)

When cyclobutene is excited into the sixth vibrational level of the methylenic C-H bond, the rate constant is 10^8 s^{-1}, but the reaction rate constant for 1-cyclopropylcyclobutene is 10^5 s^{-1}. No difference in rate was observed for this reaction when C-H bonds on the cyclopropyl ring or the cyclobutene ring were excited, indicating that mode selective chemistry was not achieved.

Deanne Snavely

Mode-selective photochemistry. Recently, scientists have become interested in trying to use lasers to cause selective photochemistry. In particular, they would like to break a specific bond in a molecule while leaving the other, weaker bonds unbroken, or to cause a reaction at a specific site in a molecule. This area of research is called mode-selective photochemistry.

Several schemes for mode-selective photochemistry have been tried, and two have shown promise. The first is direct excitation of a molecule with ultraviolet photons to produce an electronic state that correlates to a particular set of reaction products. The second is to vibrationally excite a molecule using visible or infrared photons to energize a particular bond in a small molecule. If the energy can be localized in the bond, it will take less additional energy to break it, and the bond can be thought of as having been weakened. Mode-selective bimolecular photochemistry has also been demonstrated with this second technique.

Laurie Butler, University of Chicago, has used ultraviolet radiation to demonstrate mode-selective photochemistry by using molecules that contain two different halogen atoms bonded to a carbon atom, that is, two chemically similar bonds of differing strengths. Butler has been able to break the stronger C-Br bond of a molecule in the presence of a weaker C-I bond with nearly unit efficiency. **Figure 2** is a schematic of an experiment performed using

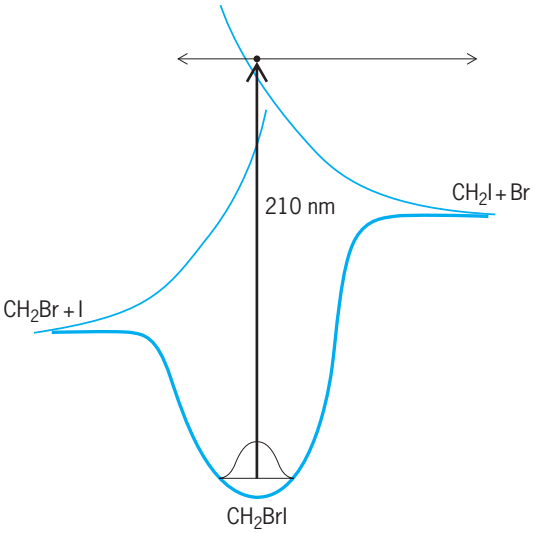

Fig. 2. Schematic representation of a photodissociation experiment where 210-nm light is used to excite bromoiodomethane, causing rupture of the strong C-Br bond in preference to the weaker C-I bond.

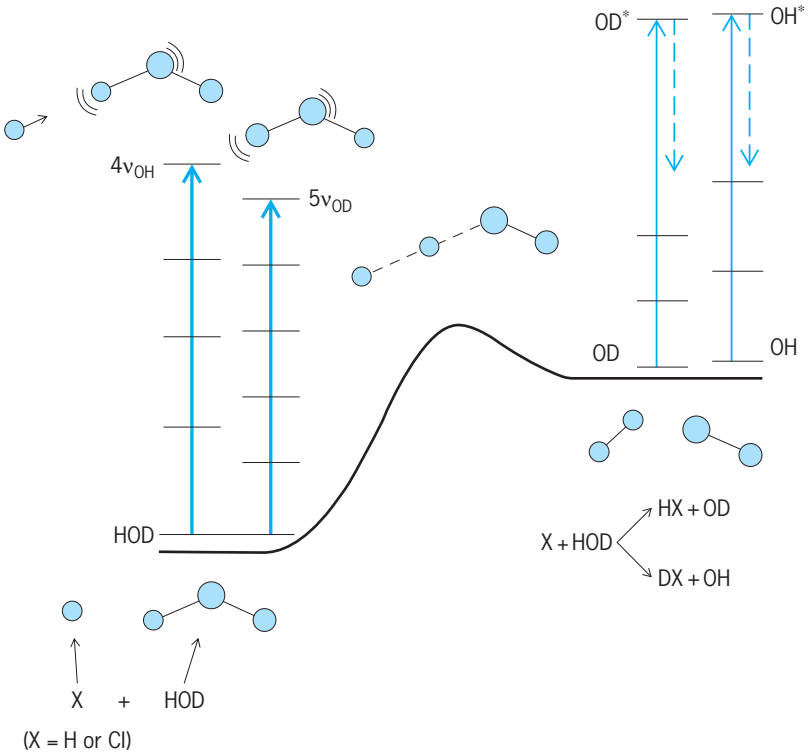

Fig. 3. Schematic drawing of the reaction barrier for breaking either the O-H or the O-D bond in the reaction of H + HOD or Cl + HOD. The thick arrows represent vibrational overtone excitation of either the $4\nu_{OH}$ or the $5\nu_{OD}$ bonds of HOD. The thin arrows represent the use of ultraviolet laser light to detect the OD or OH products.

ultraviolet light near 210 nanometers to excite bromoiodomethane. This wavelength of light will excite an electron to an antibonding orbital that correlates to products CH_2I + Br. At a lower energy, electronic states would couple to the lower-energy-channel products CH_2Br + I. Because the dissociation is very rapid, the energy does not have time to redistribute in the molecule, and only CH_2I + Br products are observed. *See* HALOGEN ELEMENTS.

The idea of using vibrational excitation to selectively weaken a bond has been tried for many years. After the invention, in the 1970s, of intense carbon dioxide lasers, which produced infrared radiation near 10 micrometers, many groups attempted to break particular bonds in molecules by resonantly exciting low-frequency vibrations. This approach was not successful because the energy would rapidly flow throughout the molecule, a process that physical chemists call intramolecular vibrational redistribution, and the weakest bond would predominantly break. With the advent of powerful visible lasers, in the 1980s, scientists began trying to use a single energetic photon to induce vibrational mode-selective photochemistry. They would tune the laser to be resonant with a weak vibrational overtone transition (a transition where the vibrational quantum number changes by more than one) in a molecule possessing a low-energy reaction process, such as isomerization or dissociation. The hope was that the single photon would deposit energy in the molecule with enough specificity to accelerate a particular reaction so that selective chemistry could be observed. Once again, it was found that molecules large enough to have low-lying reaction barriers had a very high density of vibrational states and strong couplings between them, so that the photon energy experienced intramolecular vibrational redistribution, and for the most part statistical behavior was observed.

Fleming Crim, University of Wisconsin, was able to succeed with this approach in small molecules that have a low density of states and therefore little intramolecular vibrational redistribution. In particular, Crim was able to demonstrate a high degree of mode-selective photodissociation using a technique he calls vibrationally mediated photodissociation. In the first step of this process, a small molecule, such as water (HOD) or hydrogen peroxide (HOOH), is excited via an overtone transition. This vibrational excitation does not possess enough energy to cause dissociation, but the further excitation of the molecule from the vibrationally excited state to a dissociative electronic state can be very efficiently and selectively achieved. In this manner, only the initially excited molecules are dissociated, and they dissociate along the initially excited bond. For example, exciting either the OH stretching vibration or the corresponding OD vibration will cleave either the H or the D of the HOD, and in the case of hydrogen peroxide (HOOH) will cause the O-H bond to break despite the presence of the weaker O-O bond.

Bimolecular mode-selective photochemistry has also been demonstrated by Crim (**Fig. 3**). Once again, an overtone vibrational excitation of one of the bonds of HOD is used to energize a bond. An H atom or Cl atom is produced concurrent with the vibrational excitation of the HOD molecules. A collision between a vibrationally excited HOD molecule and an H atom or a Cl atom has enough energy to surmount the barrier to this endothermic reaction. The exciting aspect of this experiment is that the selectivity of the initial vibrational excitation is not lost in the collision that produces the reaction. If the O-H bond is initially excited, then H_2 + OD or HCl + OD are the final products; and if the O-D bond is initially excited, HD + OH or DCl + OH are the final products.

These experiments show that mode-selective photochemistry is not only possible but can be made to be extremely efficient. David W. Chandler

Time-resolved photochemistry. An important advantage of lasers in photochemistry is their ability to time-resolve the chemical processes that occur after the photon is absorbed. Transient absorption and time-resolved laser-induced emission can be used to identify the species involved in the reaction and to determine lifetimes. For applications in organic chemistry, a typical laser flash photolysis experiment employs a nanosecond laser pulse from a nitrogen or excimer laser (perhaps pumping a dye laser) to initiate the chemical reaction and a tunable light source (perhaps a pulsed xenon lamp on another laser) to monitor the transient species. The time resolution of the experiment is determined by the duration of the light source and the response time of the

detector. The overall time resolution for this technique is usually 3–6 nanoseconds.

Laser flash photolysis has been used to measure the lifetimes of many biradical species. In many common organic solvents, these lifetimes range from a few nanoseconds to several hundred nanoseconds. For instance, the biradical lifetimes of structures (**3**)–(**6**)

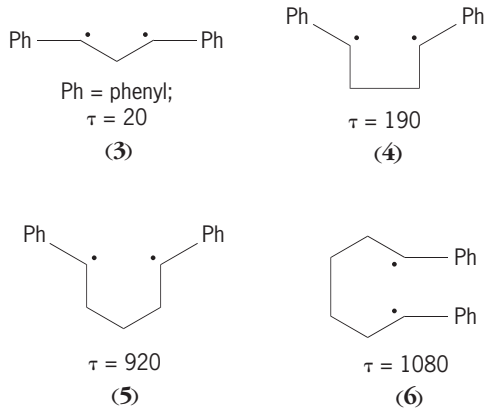

as measured by laser flash photolysis in acetonitrile increases from 20 to 1080 ns. The lifetimes increase as the distance between the radical centers increases. Intersystem crossing occurs when an excited molecule in a singlet state converts to an electronic triplet state. Radicals (**3**)–(**6**) must undergo intersystem crossing before they undergo further chemical reaction. However, for these species this intersystem crossing rate is sufficiently slow to be the rate-determining step. As the distance between the radical centers increases, the rate of intersystem crossing increases, which in turn lengthens the lifetime of the transient diradical. *See* FREE RADICAL; LASER SPECTROSCOPY.

Microelectronics fabrication. There are many advantages of using laser photochemistry to fabricate microelectronic devices. Many of the materials (mercury-cadmium-tellurium, gallium arsenide, and plastics) used in microelectronics are sensitive to thermal decomposition; consequently, it is desirable to maintain lower temperatures during fabrication. With laser processing, no bulk heating occurs, since the chemical activation energy is provided by the laser photon. In addition, the laser beam can be focused to process a very small area on the substrate surface, so that only localized heating is created. The dimensions of the metal lines deposited on the substrate surface must be very narrow (on the order of 1 micrometer); the lower temperatures avoid metal diffusion across the surface, broadening the deposited line. An important advantage of lasers is the generation of patterns of deposited material without the need for masking. In conventional microelectronics fabrication, a patterned mask is usually deposited onto the substrate surface (typically silicon). This mask is then used to transfer the pattern onto the underlying surface, which might be further processed by etching in subsequent steps. With laser deposition and laser-driven doping, the same result can be achieved in a few steps.

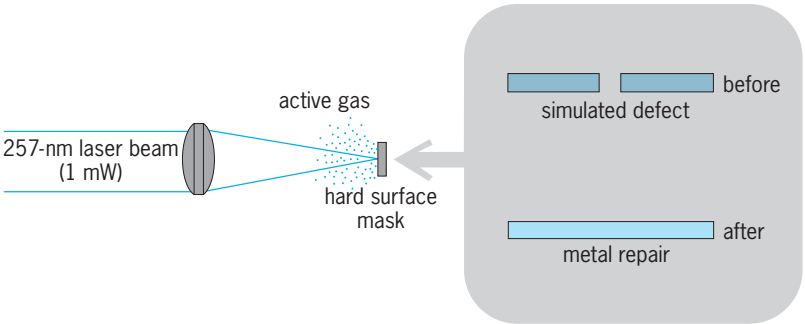

Fig. 4. Mask repair by laser deposition using a low-power laser at 1-μm resolution. (*After K. G. Ibbs and R. M. Osgood, eds., Laser Chemical Processing for Microelectronics, Cambridge University Press, 1989*)

Laser chemical deposition of metal is used in mask repair. This process involves repairing a discrete defect in a metal line on a semiconductor surface (**Fig. 4**). For example, metallic cadmium can be deposited by photodissociating gaseous dimethylcadmium with a weakly focused, frequency-doubled argon ion laser operating at 257 nm. Laser focus can be controlled to cover a 10-μm-diameter area, small enough to patch the defective cadmium line precisely.

The interference pattern generated by two laser beams at a semiconductor surface can generate a diffraction grating by using a technique known as laser-induced wet etching. In this process the semiconductor is typically submerged in an acidic aqueous solution. Light-induced electrochemical reactions involving the metal atoms of the semiconductor render the illuminated portion more soluble in the aqueous solution. Consequently, etching of the surface occurs only in the illuminated area and produces the interference pattern formed by the laser beams. *See* INTEGRATED CIRCUITS; VAPOR DEPOSITION.

Femtosecond photochemistry. Chemical dynamics is an area of chemistry that attempts to determine precisely how an individual molecule undergoes a chemical reaction. This involves preparing the molecular sample in some initial quantum-mechanical state (or some narrow distribution of initial states) and determining the final states that are produced after a chemical reaction. The most widely used technique for producing narrow initial-state populations is the supersonic expansion. A supersonic jet expansion is produced by expanding a mixture of the photochemically active species and a diluent gas such as nitrogen (N_2) or helium (He) from high pressure to low pressure through a small orifice. As the gaseous mixture expands adiabatically into the low-pressure region, it cools substantially. If helium is used, temperatures of a few kelvins are obtained routinely (**Fig. 5**).

Laser photoexcitation techniques such as wavelength selection or polarization alignment further limit the number of initial states for the given photochemical reaction. The initial relative translational energy of reagents in a bimolecular reaction can be controlled also by using two supersonic jets intersecting at a particular angle. Pulsed lasers having

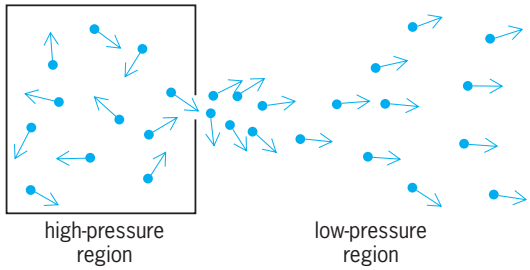

Fig. 5. Diagram of the expansion of a gas in a supersonic jet. The gas cools as it expands into the low-pressure regions.

nanosecond and picosecond pulse duration are used in chemical dynamics to keep the time intervals between the initiation of the reaction and the diagnostics of the final states short to avoid contamination of subsequent chemical reactions or changes of final state by collisions.

As the ability to control the time pulsewidth of lasers has moved into the femtosecond (10^{-15}) range, another area of chemical dynamics has opened—that of observing a chemical reaction while it is occurring. For example, it takes about 200 fs for the dissociation reaction of the triatomic molecule cyanogen iodide (ICN) to occur. If this reaction is initiated by a laser with 6-fs pulses and probed at short intervals after the dissociating pulse, the progress of the reaction through the transition state can be monitored. The time interval of 100 fs between the laser pulse initiating the dissociation reaction and the pulse monitoring the progress of the reaction is generated by sending the monitoring pulse through an extra distance of 3×10^{-5} m on its path to the ICN sample. (Since light travels 3×10^{-7} m/fs, this extra distance introduces a 100-fs delay.) In reaction scheme (3) the excitation laser at 307 nm excites the

$$\mathrm{ICN}^* \longrightarrow [\mathrm{I} \cdots\cdots \mathrm{CN}]^* \longrightarrow \mathrm{I} + \mathrm{CN} \qquad (3)$$

ICN to ICN* on the repulsive excited-state surface labeled V_1 in **Fig. 6**. The delayed probe laser, operating at 388.5 nm, excites the free CN fragment, from which fluorescence is detected. The delay between the excitation laser at 307 nm and the probe laser at 388.5 nm measures the time for the free CN fragment to be generated. Furthermore, the progress of the reaction can be monitored by tuning the probe laser wavelength to longer wavelengths. As can be seen in Fig. 4, the energy difference between V_1 and V_2 depends on the separation of I and CN. This means that if the probe laser is shifted to the red of the free CN absorption wavelength, the CN fragment can be excited as it is being generated. This is the first time that observations have been made of a primary chemical reaction in progress and the time required for the fragments to move along their reaction trajectory has been determined. These techniques are under development for application to more complicated reactions. *See* CHEMICAL DYNAMICS; FLUORESCENCE; LASER; PHOTOCHEMISTRY; ULTRAFAST MOLECULAR PROCESSES.

Deanne Snavely

Bibliography. V. N. Bagratashvili et al., *Multiple Photon Infrared Laser Photophysics and Photochemistry*, 1985; L. J. Butler, Chemical reactions beyond the Born-Oppenheimer approximation, *Annu. Rev. Phys. Chem.*, 49:125–171, 1998; J. Chaiken (ed.), *Laser Chemistry of Organometallics*, 1993; F. F. Crim, Bond-selective chemistry: Vibrational state control of photodissociation and bimolecular reaction, *J. Phys. Chem.*, 100:12725–12734, 1996; K. G. Ibbs and R. M. Osgood (eds.), *Laser Chemical Processing for Microelectronics*, 1989; V. Ramamurthy, *Organic Photochemistry*, 1997; H.-G. Rubahn, *Laser Application in Surface Science and Technology*, 1999; C. E. Wayne and R. P. Wayne, *Photochemistry*, 1996.

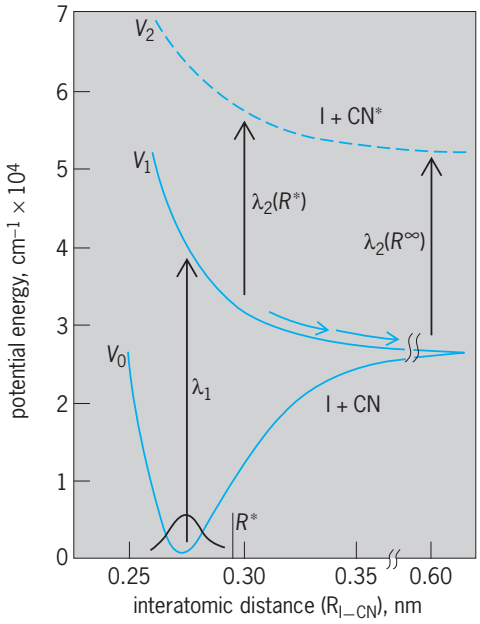

Fig. 6. Potential energy diagram for the dissociation of ICN. (*After A. H. Zewail, Chem. Eng. News, 66(42):24–42, 1988*)

Laser-solid interactions

Interactions of laser light with solids. The term usually refers to the thermal effects of absorption of high-intensity laser beams. For nonthermal laser interactions with matter *see* LASER PHOTOCHEMISTRY; LASER SPECTROSCOPY; NONLINEAR OPTICS.

One of the unique features of lasers is their ability to concentrate the energy of light so that a surface can be exposed to a photon flux of 10^6 to 10^{12} W/cm^2. These extraordinarily high power densities are attainable even with lasers of moderate size and average output of only a few watts, since light can be emitted in pulses of a few nanoseconds or even less than 1 picosecond duration and a coherent beam can be focused into spots as small as 1 micrometer in diameter. The high power densities allow melting and even vaporization of any solid material that is sufficiently opaque at a given wavelength or photon energy. This has led to a number of

applications involving cutting and drilling of ceramics and other brittle materials, even diamonds. Welding of components from the smallest wires to huge steel plates is done commercially with high-power lasers. Metal alloying in surface regions is also a domain of lasers. Modification of semiconductors with intense laser irradiations has emerged as one of the most exciting fields of scientific discovery while at the same time promising substantial practical benefits for the semiconductor device industry. *See* LASER ALLOYING; LASER WELDING; SEMICONDUCTOR.

Photon-electron-phonon interactions. Photons with energy exceeding the fundamental band-gap energy E_g of the semiconductor are absorbed near the surface, exciting electrons from the valence to the conduction band and forming electron-hole pairs. For direct band-gap semiconductors, such as gallium arsenide (GaAs), indium phosphide (InP), and indium antimonide (InSb), light is absorbed within 100 nanometers of the surface. For indirect band-gap semiconductors, such as silicon and germanium, the transition from transparency to absorption is a more gradual function of the light wavelength, because for each transfer of energy from a photon to an electron, a phonon has to be created or annihilated to assure conservation of momentum. *See* BAND THEORY OF SOLIDS; CRYSTAL ABSORPTION SPECTRA.

For photon energies less than E_g, semiconductors are largely transparent, at least at moderate light intensities. Some photons are still absorbed by exciting electrons and holes within the conduction and valence bands to higher energies (free-carrier absorption), but unless the semiconductor is heavily doped with impurities that lead to a large concentration of free carriers, the light beam can penetrate to a depth of 0.04 in. (1 mm) or more before it is fully absorbed. Far-infrared light can interact directly with the lattice vibrations or phonons, that is, the energy of photons can be directly transformed into heat, but this absorption mechanism is quite weak. The low absorption channels are often very important at high photon fluxes, as they can initiate surface heating that tends to make solids more opaque. This feedback mechanism facilitates laser melting and vaporization of almost any solid substance. In semiconductors, free-carrier absorption by electrons and holes generated by optically and thermally stimulated band-to-band transitions contributes a large component to the total absorption at high power densities. *See* LATTICE VIBRATIONS; PHONON.

Excited electrons and holes share energy with one another very rapidly through electron-electron and electron-plasmon interactions. They release energy to other electric carriers and to phonons until they reach the lowest energy possible within their respective energy bands. Then they recombine by one of several mechanisms: radiatively by emitting a photon, nonradiatively by transferring the amount of energy E_g into heat or by an Auger process that transfers the energy to a third particle (electron or hole), exciting it to an energy higher by E_g (**Fig. 1**). Auger interactions are very important for a high density of photoexcited carriers. Although they do not reduce

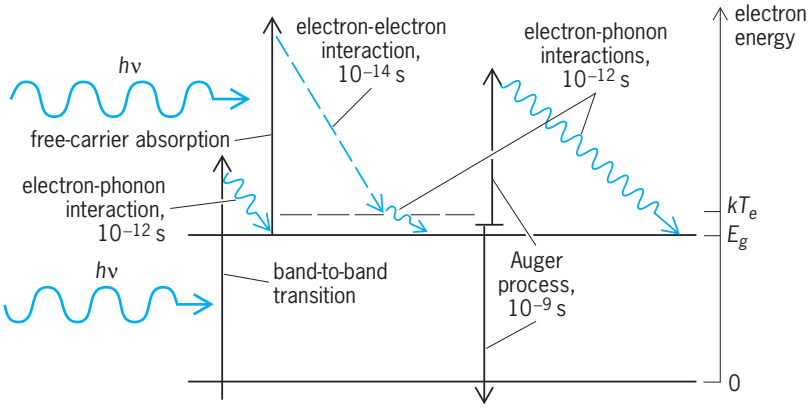

Fig. 1. Absorption of light and energy transfer mechanism for a very intense flux (10^7-10^9 W/cm^2) of photons with energy $h\nu$ exceeding the band-gap energy E_g. Approximate times for processes to take place are given. kT_e is the average thermal energy per electron, the product of Boltzmann's constant k and the electron temperature T_e.

the energy of the electronic system, they convert the band-gap energy into kinetic energy of carriers that can be dissipated rapidly into heat by phonon emission. All the photon energy absorbed by the solid is transformed into heat in about 1 ps, that is, almost instantaneously. *See* AUGER EFFECT.

Annealing of ion-implanted layers. Ion implantation has become a dominant method of introducing controlled quantities of impurities near the surface of silicon and other semiconductors for the purpose of adjusting their electrical conductivity and forming diodes, transistors, and entire integrated circuits. The implanted layers need a heat treatment to repair the displacement damage caused by bombardment with energetic ions and to move the implanted impurity ions into lattice locations where they replace host atoms and become electrically active. Laser heating is particularly suitable for annealing since only the implanted regions are heated. This capability of delivering heat selectively to a very thin layer is the basis of "laser annealing" of semiconductors. *See* ION IMPLANTATION.

Pulsed laser annealing. A short, high-intensity laser pulse melts the layer of silicon damaged by ion implantation. As soon as the pulse is over, the layer freezes, starting at the unmelted single crystal (**Fig. 2**). The single crystal forms a template on which the melt solidifies, preserving the perfection of the template. This process of liquid-phase epitaxy is the same as conventional crystal growth from the melt, except that the solid-liquid interface moves at a velocity of many feet per second, driven by the temperature gradients of about 10^7 degrees per inch. The quality of the crystalline structure is not diminished by the rate of growth, but some new phenomena appear. For example, most impurities that are present in the liquid are incorporated into the solid more completely than during slow crystal growth. Maximum concentrations of some impurities in the crystalline semiconductor can exceed by orders of magnitude the near-equilibrium solid solubilities.

Pulse irradiation allows investigation of the physical limits of crystal growth and the thermodynamic properties of different phases. For example, it was

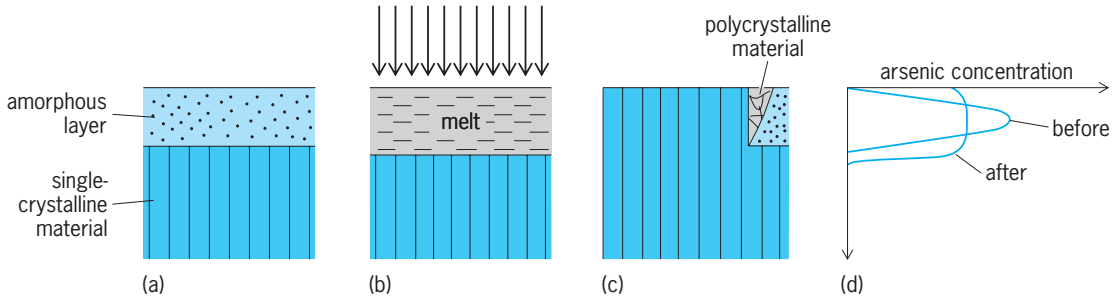

Fig. 2. Pulsed laser annealing by liquid-phase epitaxy. (*a*) Cross section of a crystalline semiconductor with a thin (typically 100 nm thick) amorphous surface layer formed by ion implantation. (*b*) The entire thickness of the amorphous film and some of the underlying crystal melted by a laser pulse. (*c*) Molten layer recrystallized from the crystalline template to the surface. At the edge of the laser-irradiated spot, where the melt has not reached the crystalline substrate, the material is polycrystalline. (*d*) Typical depth profiles of arsenic implanted into silicon, before and after pulsed laser annealing.

determined that amorphous silicon melts at a lower temperature than crystalline silicon, confirming theoretical predictions based on Gibbs free-energy considerations. When the solidification rate of molten semiconductors exceeds about 50 ft/s (15 m/s; this is accomplished with pulses shorter than 2 ns and short-wavelength irradiation), the liquid is supercooled so much that it transforms into an amorphous layer instead of a single crystalline layer. Therefore, depending on irradiation conditions, amorphous → crystalline and crystalline → amorphous phase transitions are achievable, even with the same laser.

Continuous-wave. Laser beams can anneal implantation displacement damage without melting, by solid-phase epitaxy. At elevated temperature the atoms rearrange themselves into an orderly crystalline structure. The process starts at the boundary between crystalline and damaged regions, and this interface moves through the amorphous film (**Fig. 3**), at a velocity determined by the activation energy E_a (for silicon $E_a = 2.7$ eV). Typically about 1 millisecond time is required to regrow ion-implanted surface layers. Pulses of such duration could be used, but better results are obtained with a continuous-wave laser beam scanned across the surface at a velocity that yields about 1 ms dwell time. This type of laser annealing is physically equivalent to furnace annealing, but there are some important differences. Heating is so brief that practically no diffusion of impurities takes place. The impurity profiles after annealing are indistinguishable from as-implanted depth profiles. In contrast, furnace annealing at high temperatures that are necessary for complete defect removal causes significant diffusion of doping impurities. Laser heating is localized both in depth and in a lateral direction. Small areas can be selectively annealed with a tightly focused beam, and micrometer-size patterns can be written in implantation amorphized material. Since high surface temperatures are attained with a laser in less than 1 ms, measurements of solid-phase epitaxial regrowth rates up to the melting temperature became possible, broadening the understanding of solid-phase epitaxy. *See* CRYSTAL GROWTH.

Silicon on insulator. Thin films of single-crystalline silicon over an insulating substrate are very attractive for high-speed integrated circuits. An important approach to the formation of such films is the controlled melting of thin polycrystalline layers deposited over fused silica substrates or over oxidized silicon wafers. Through a careful control of temperature gradients around the molten spot, by shaping the laser beam or patterning the film, single-crystalline regions can be obtained. One particularly successful approach involves growing single-crystalline films from openings in the oxide where the melt is in contact with the single-crystalline template (**Fig. 4**). The formation of silicon-on-insulator structures will lead in the future to three-dimensional circuits, with several levels of transistors on the same chip. *See* INTEGRATED CIRCUITS.

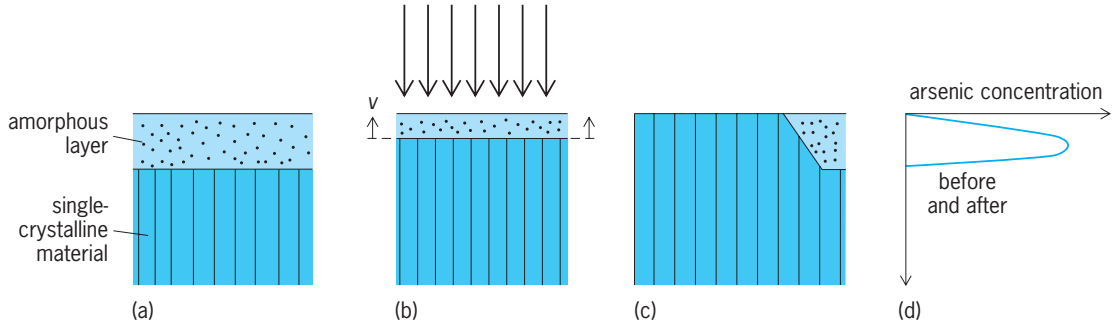

Fig. 3. Continuous-wave laser annealing by solid-phase epitaxy. (*a*) Cross section of the starting structure. (*b*) Crystalline-amorphous interface propagating toward the surface with a velocity *v* that depends on temperature induced by the laser beam. (*c*) The irradiated region fully recrystallized, except at the edge of the laser-heated spot where the interface has a wedge shape. (*d*) Depth profiles of arsenic implanted into silicon, the same before and after continuous-wave laser annealing.

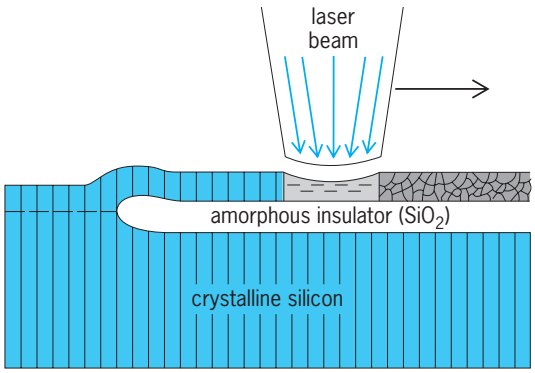

Fig. 4. Melting with a scanning continuous-wave laser beam, allowing lateral epitaxial regrowth of crystalline silicon over a layer of silicon dioxide (SiO_2).

Laser-induced-damage gettering. At high power densities, Q-switched laser pulses can vaporize material from the surface of a solid and form microscopic craters. Shock waves associated with vaporization stress the surface and lead to formation of dislocations. The laser-induced surface damage helps in removing undesirable impurities from the device regions of silicon to an inactive area such as the back side of a wafer. The fast-diffusing impurities become trapped or gettered at the damage sites, from where they do not affect device operation.

Laser deposition. Vaporization of materials with nanosecond laser pulses can be used to grow thin films. Each laser pulse removes a small amount of material from the irradiated surface, with the plume of atoms, ions, and small atomic clusters spreading out and condensing on the surrounding surfaces. To obtain controlled deposition, the target and the object to be coated are placed in a vacuum chamber, with the laser beam entering through a transparent window. Films of amorphous silicon and of refractory materials have been formed in this way. Laser deposition has been successfully applied to formation of superconducting films of $YBa_2Cu_3O_x$ with a critical temperature T_c of 90 K. (This means that the film is superconducting up to 90 K or −298°F.)

Laser generation of x-rays. Very high-intensity, short laser pulses focused on a metallic alloy generate temperatures of many million degrees, vaporizing the surface and forming a plume of highly ionized gas. The ions relax to a lower energy by emitting x-ray radiation. These short bursts of x-rays, with typical wavelengths of 0.6 to 2 nm, find application in x-ray lithography, a process that is necessary for making complex semiconductor chips with millions of transistors, each one smaller than 0.5 micrometer.

Inertial confinement fusion. Gigantic lasers, with multiple beams, are used to implode micrometer-size pellets of deuterium and tritium, producing temperatures and pressures equivalent to those inside the stars. The goal is to overcome repulsion of nuclei and fuse them into helium atoms, releasing controllable amounts of thermonuculear fusion energy. *See* LASER; NUCLEAR FUSION. George K. Celler

Bibliography. G. K. Celler, *Modification of Silicon Properties with Lasers, Electron Beams, and Incoherent Light*, CRC Critical Reviews in Solid State and Material Science, vol. 12, pp. 193–265, 1984; S. S. Charschan, *Lasers in Industry*, 1972; *Materials Research Society Proceedings*, annual volumes since 1979; J. M. Poate and J. W. Mayer, *Laser Annealing of Semiconductors*, 1982.

Laser spectroscopy

Spectroscopy with laser light or, more generally, studies of the interaction between laser radiation and matter. Lasers have led to a revolution of classical spectroscopy, because laser light can far surpass the light from other sources in brightness, spectral purity, and directionality; and if required, laser light can be produced in extremely intense and short pulses, far enhancing the possibility of observing nonlinear phenomena. The use of lasers can greatly increase the resolution and sensitivity of conventional spectroscopic techniques, such as absorption spectroscopy, fluorescence spectroscopy, or Raman spectroscopy. Moreover, interesting phenomena have become observable in the resonant interaction of intense coherent laser light with matter. Some of these effects have become the basis for powerful spectroscopic methods, which offer unprecedented spectral resolution, or which permit the investigation of properties of matter that could not be observed previously. Laser spectroscopy has become a wide and diverse field, with applications in numerous areas of physics, chemistry, and biology.

Tunable sources. Early lasers, such as ruby lasers or helium-neon lasers, worked only at a few discrete wavelengths determined by narrow spectral lines of the active medium. The vigorous advances of laser spectroscopy since about 1970 have been largely due to the development of laser sources which are highly monochromatic and in which the wavelength can be tuned continuously over a wide spectral range.

Tunable coherent sources are available for any wavelength region from the far infrared into the near-vacuum ultraviolet. Tunable infrared lasers include high-pressure molecular gas lasers, semiconductor diode lasers, spin-flip Raman lasers, and color-center lasers. In the visible region, organic dye lasers and later titanium:sapphire lasers have proved particularly versatile and powerful spectroscopic sources. Both pulsed and continuous-wave dye lasers cover the entire visible spectrum, including the bordering near-infrared and near-ultraviolet regions. Titanium:sapphire lasers allow a huge tunable range in the red and near infrared when used as continuous-wave sources. Mode-locked titanium:sapphire systems have allowed the realization of ultrashort optical pulses well down in the femtosecond region. Intense shorter-wave ultraviolet radiation can be generated with excimer lasers over limited regions. Semiconductor diode lasers now cover almost all the wavelengths from ultraviolet to far infrared, allowing great reduction in costs and size of experiments.

The generation of harmonic frequencies or sum frequencies in nonlinear optical crystals or gases

provides a valuable method to produce tunable, coherent ultraviolet radiation from laser light of longer wavelengths. Optical parametric oscillators and difference-frequency crystal mixers offer corresponding alternatives for the generation of coherent infrared radiation. Stimulated Raman scattering can be used to shift the frequency of a tunable laser by some integer multiple of a molecular vibration frequency and thus to extend the tuning range. These nonlinear frequency-mixing and shifting techniques work best with the intense radiation from pulsed lasers. *See* NONLINEAR OPTICS.

New techniques have been developed to take advantage of laser coherence and spectral purity. Considerable engineering efforts have been devoted to the active frequency stabilization of continuous-wave lasers. By electronically locking the frequency of a continuous-wave titanium:sapphire laser to some reference interferometer, for instance, a linewidth of less than 1 MHz is achieved in commercial systems, and linewidths below 1 Hz have been produced with more sophisticated feedback controls or a resolution on the order of 1 part in 10^{15}. Femtosecond pulsed lasers have been used to create a link between microwave frequencies and optical frequencies in order to take full advantage of the narrow linewidth of tunable lasers.

Absorption spectroscopy. Lasers can replace conventional light sources and spectrographs or monochromators for absorption spectroscopy. The spectral purity of laser light can eliminate instrumental resolution limits, the high intensity helps to overcome detector noise, and the good directionality permits long or folded absorption paths. All these factors contribute to improved sensitivity.

Rather than by measuring the attenuation of a laser beam in the sample, the absorption can often be monitored indirectly with still higher sensitivity. For instance, absorption of a modulated laser beam will produce a sound wave in the sample, because some of the light is converted to heat. This sound wave can be picked up by a microphone (optoacoustic detection). Resonant absorption of laser light by atoms or molecules in a gas discharge can change the ionization probability. The resulting changes in discharge current or voltage across the tube are easily measured (optogalvanic detection). The highest sensitivities at visible and ultraviolet wavelengths have been obtained by monitoring the fluorescence or photoionization of laser-excited atoms or molecules in a gas or molecular beam. Single atoms have been selectively detected in this way.

Intracavity absorption. Very high sensitivity has also been obtained by placing an absorbing sample inside the resonator of a broad-band dye laser without any optical tuning elements. Any absorption lines can be detected as dips in the laser emission spectrum when analyzed by a spectrograph. This intracavity spectroscopy can surpass the sensitivity of a conventional single-pass absorption measurement by a large factor (on the order of 10^5), because the sample is effectively traversed by the light many times, and the competition between many simultaneously oscillating modes enhances the suppression of modes with slightly increased losses.

Multiphoton absorption. It has long been known that atoms can be excited to a higher quantum state by simultaneously absorbing two or more photons which together provide the necessary energy. The probability for an N-photon transition grows initially with the N-th power of the intensity. Only the high intensity of laser light has made it possible to observe two-photon and multiphoton absorption in the optical region. Two-photon spectroscopy permits the study of states with the same parity as the absorbing level, which are not normally reached by single-photon transitions, and it requires photons of less energy, which are sometimes more readily produced. *See* SELECTION RULES (PHYSICS).

Fluorescence spectroscopy. Intense laser light is a very effective means to pump a large fraction of an absorbing species to some excited quantum level. Hence, lasers can greatly increase the sensitivity of such classical spectroscopic methods as fluorescence spectroscopy, optical pumping, level-crossing spectroscopy, or double-resonance spectroscopy. Moreover, lasers make it possible to apply these techniques to atomic and molecular transitions at wavelengths where intense spectral lamps are not available, and stepwise excitation permits studies even of highly excited states, including autoionizing levels and very high Rydberg states of atoms.

Studies of the line shape of fluorescent emission at high intensities have permitted interesting tests of the predictions of quantum-electrodynamic theory. *See* QUANTUM ELECTRODYNAMICS.

Raman spectroscopy. Lasers have revolutionized Raman spectroscopy, the observation of scattered light at wavelengths other than that of the exciting light. The high intensity of laser light has greatly increased the sensitivity of this form of two-photon spectroscopy, where the energy difference between incident photon and scattered photon corresponds to a resonant transition between states of equal parity. Tunable lasers have enhanced the sensitivity, resolution, and versatility of Raman spectroscopy even more, by making it possible to excite close to some intermediate resonant transition (resonance Raman spectroscopy), or to observe stimulated Raman scattering with the help of a second, tunable probe laser beam. Polarization anisotropies associated with Raman transitions provide additional information and can further increase the detection sensitivity. *See* RAMAN EFFECT.

Frequency mixing. The high intensity of laser light has also made possible a new class of spectroscopic methods which rely on nonlinear frequency mixing in the sample. For instance, if a sample is irradiated simultaneously by two strong laser beams of frequencies ω_1 and ω_2, the nonlinear response of the driven dipoles leads to the generation of coherent light at a new frequency, $2\omega_1-\omega_2$. This frequency mixing can be described in terms of a third-order nonlinear susceptibility of the sample.

If the frequency ω_2 is tuned to the (redshifted) Stokes line of a Raman transition, excited by ω_1, a resonance is observed in the intensity of the new wave at the (blueshifted) anti-Stokes frequency. (A

Stokes line is a transition close to a Raman transition at a frequency shifted below the Raman resonance by a quantum of vibration of a molecule. On the contrary, if the frequency is shifted above the Raman resonance by the same quantum of vibration, one has the anti-Stokes line.) This type of frequency mixing is sometimes referred to as coherent anti-Stokes Raman spectroscopy (CARS). It can be used to study Raman transitions even in the presence of strong incoherent background radiation, such as in the observation of flames.

High-resolution laser spectroscopy. Lasers have led to particularly noteworthy progress in the field of very high resolution spectroscopy. They have become powerful tools to investigate the structure of atoms, molecules, and ions. They can be used to study fine and hyperfine splittings, Zeeman and Stark splittings, light shifts, collision broadening, collision shifts, and other attributes of spectral lines. Moreover, lasers make it possible to measure the wavelengths of spectral lines with unprecedented accuracy. The laser wavelength can be locked to an atomic or molecular transition, providing an accurate standard length or frequency. Such lasers have become important tools for precision metrology, making possible accurate measurements of fundamental constants and stringent tests of basic physics laws. *See* RYDBERG CONSTANT.

Doppler-free spectroscopy. The sharpest spectral lines are generally those in atoms or molecules which are relatively free and undisturbed. Such molecules, however, are almost inevitably moving with high thermal velocity. Molecules which are moving toward an observer appear to emit or absorb light at higher frequencies than molecules at rest, and molecules moving away appear to absorb at lower frequencies. In a gas, with molecules moving at random in all directions, the lines appear blurred, with typical Doppler widths on the order of 1 part in 1,000,000. To take advantage of the very narrow instrumental linewidth of laser sources, it is necessary to overcome this Doppler broadening of spectral lines. *See* DOPPLER EFFECT.

Saturation spectroscopy. Laser light can easily be intense enough to partly saturate the absorption of a spectral line. That is, those molecules which have absorbed a light quantum are temporarily removed from the initial state. The absorption from that level is reduced or saturated, at least until some relaxation process can replenish the supply of absorbing molecules. In a gas, a monochromatic laser beam will resonantly interact only with those molecules which have the right axial velocity to be Doppler-shifted into resonance. The resulting velocity-selective saturation can be used for Doppler-free spectroscopy.

In 1970 C. Borde and T. W. Hänsch introduced independently a now commonly used form of saturation spectroscopy which achieves good sensitivity with samples outside the laser resonator and which is particularly well suited for use with broadly tunable laser sources. As in **Fig. 1**, the output of a tunable laser is divided by a beam splitter into a stronger saturating beam and a weaker probe beam that are traversing an absorbing gas sample along the same path but in nearly opposite directions. When the saturating beam is on, it bleaches a path through the cell; that is, it depletes those molecules which are Doppler-shifted into resonance, and a stronger probe signal is received at the detector. As the saturating beam is alternately stopped and transmitted by a chopper, the probe signal is modulated. However, that happens only when both beams interact with the same molecules, and those can only be molecules which are standing still or at most moving transversely. Thus the method picks out those molecules which have near-zero component of velocity along the laser beams and ignores others. **Figure 2** illustrates the power of the technique by comparing a Doppler-broadened absorption profile of the well-known Balmer-alpha line of atomic hydrogen with one of the first saturation spectra, recorded with a pulsed dye laser in a glow discharge.

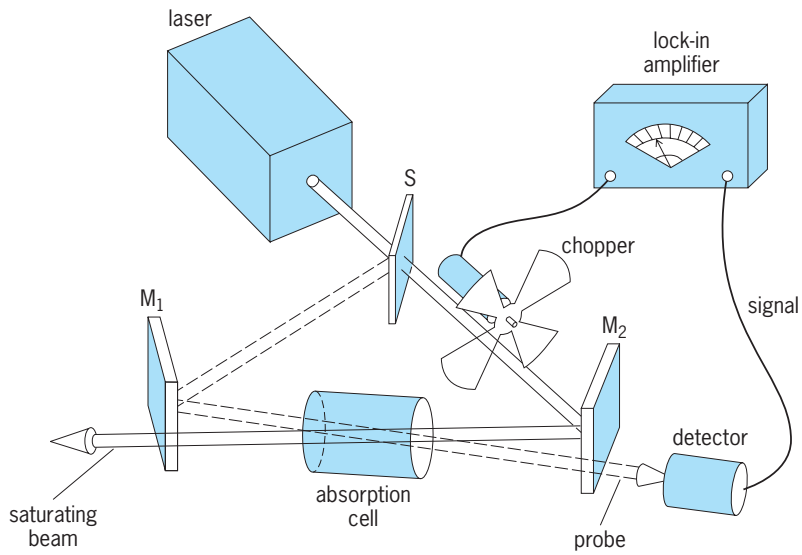

Fig. 1. Scheme of laser saturation spectrometer; S indicates the beam splitter and M_1 and M_2 are mirrors.

The described methods work well only if the sample has noticeable absorption and if the laser is strong enough to excite a substantial fraction of the resonant molecules. Higher sensitivity can sometimes be achieved by observing the absorption of laser light indirectly via the emitted fluorescence or via acoustooptic or optogalvanic detection. In this case, it is advantageous to monitor the nonlinear interaction of two counterpropagating laser beams via intermodulation; that is, the two laser beams are chopped at two frequencies f_1 and f_2, and the spectrum is recorded as a modulation in the signal at the sum or difference frequency, $f_1 + f_2$ or $f_1 - f_2$.

Polarization spectroscopy. There are also purely optical methods, which can be considerably more sensitive than the older saturated absorption method by suppressing the fluctuating background of probe light on which the signal has to be detected. These include, in particular, the technique of polarization spectroscopy introduced in 1976 by C. Wieman and T. W. Hänsch. A polarization spectrometer looks similar to a saturation spectrometer, but takes advantage of the fact that small changes in light

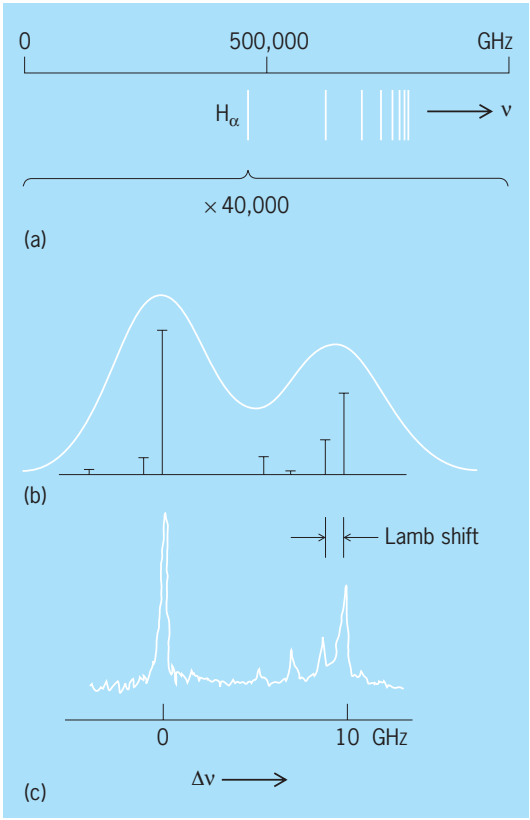

Fig. 2. Laser spectroscopy of atomic hydrogen. (*a*) Balmer series. (*b*) Doppler-broadened absorption profile (300 K) of the Balmer-alpha line with theoretical fine structure. (*c*) Saturation spectrum of Balmer-alpha line with resolved 2*S*-2*P* Lamb shift.

polarization can be detected more easily than changes in light intensity. The probe beam "sees" the sample placed between nearly crossed linear polarizers so that only very little light arrives at the photodetector. The saturating beam is made circularly polarized by a birefringent plate. Alternatively, a linearly polarized beam is used with its polarization axis rotated at 45°. *See* POLARIZED LIGHT.

Normally, in a gas, molecules have their rotation axes distributed at random in all directions, but the probability for absorbing polarized light depends on the molecular orientation. Thus the saturating beam depletes preferentially molecules with a particular orientation, leaving the remaining ones polarized. The polarized ones can then be detected with high sensitivity because they can change the polarization of the probe beam. The probe acquires a component that can pass through the crossed polarizer into the detector, but again this happens only near the center of a Doppler-broadened line where both beams are interacting with the same molecules.

Polarization spectroscopy makes it possible to observe fewer molecules with lower light intensity so that external causes of line broadening and shifts are more easily avoided. A promising related technique, saturated interference spectroscopy, can work even in spectral regions where good polarizers are not available.

Doppler-free two-photon spectroscopy. There is a completely different approach to Doppler-free laser spectroscopy which does not rely on velocity selection. However, this approach works only for two-photon or multiphoton spectroscopy, where the frequencies and directions of the exciting photons can be chosen so that the momenta of the absorbed photons add to zero.

Doppler broadening in two-photon excitation can be eliminated to first order simply by reflecting the output of an intense monochromatic laser back onto itself and placing a gas sample in the resulting standing-wave field, so that the molecules can absorb two photons of equal energy coming from opposite directions. From a moving molecule, one beam will appear Doppler-shifted toward the blue, and the other will appear shifted toward the red by an equal amount. The sum frequency is hence constant, independent of the molecular velocity. If the number of excited molecules is observed during a laser scan, a sharp resonance appears on a low Doppler-broadened background, produced by each traveling wave separately. The Doppler-free signal is strongly enhanced because all molecules, regardless of their velocity, contribute.

Laser cooling of trapped Ions and atoms. Neutral atoms and trapped ions can be cooled to temperatures on the order of 10 microkelvins with laser light. At those temperatures, the atoms have velocity spreads on the order of 1 cm/s. Both the first- and second-order Doppler shifts of the particles are greatly reduced, and the available measurement time is greatly increased. Laser cooling techniques are broadly applicable to spectroscopic measurements where the signal appears as a small frequency shift. Time standards, tests of mass anisotropy, nonlinearity in quantum mechanics, time reversal invariance, and charge neutrality of atoms are examples of work where laser cooling is playing a role. The resolution of any of the described methods is ultimately limited by the natural linewidth of the observed transition. By using almost forbidden transitions in alkali-earth atoms, it should be possible to achieve a resolution of 10^{-17}. *See* ATOMIC CLOCK.

Laser spectroscopy of cooled ions and atoms has also opened up the study of single-ion and few-ion systems, ultracold plasmas, ultracold collisions, and the formation of dilute quantum gases. These new quantum gases can be used to study, with perfect control, many solid-state phenomena such as superconductivity or electron transport in metals. Furthermore, atoms at these temperatures can be manipulated with potentials created with integrated electronic circuits, opening the way to the development of a totally new generation of quantum devices. Thereby, one could exploit the quantum correlations between the ultracold atoms to realize quantum logic circuits or quantum-limited sensors with unprecedented capabilities. *See* LASER COOLING.

Coherent transients. Nonlinear optical processes of a coherent nature can be observed in the time domain, and are capable of yielding detailed dynamic information about relaxation processes as well as Doppler-free spectra. Coherent transient phenomena, such as optical nutation, free induction decay, and photon echoes arise in a sample excited by a sequence of resonant laser pulses. These phenomena

are completely analogous to effects in pulsed NMR spectroscopy. Each transient effect may be used to select and examine a specific dephasing process. The laser field prepares the sample by placing the transition levels in quantum-mechanical superposition, and the resulting dipoles radiate in accordance with the coupled Schrödinger-Maxwell wave equations. The detection of the emitted radiation then yields information about the environment of the oscillators.

Picosecond and femtosecond spectroscopy. Extremely short pulses of light are used to examine events such as vibrational and rotational relaxation of molecules, the thermalization of photo-excited electron plasmas in a semiconductor, and nonthermal processes such as desorption of adsorbates from surfaces. Subpicosecond snapshots of chemical reactions, studies of the primary photochemical behavior of retinal in rhodopsin, and ultrafast electrical pulses are now possible with the short-pulse technology. Most experiments use some variation of the pump-probe technique, where the first optical pulse is used to trigger an event with the absorption of one or more photons. A second pulse optically delayed with respect to the first pulse then probes the system spectroscopically. Coherent transient techniques such as photon echoes or four-wave mixing are also applicable in time domain spectroscopy. The short laser pulses can also be used to generate very short bursts of x-rays, phonons, electrons, and electrical pulses. A very recent development is the achievement of attosecond (10^{-18} s) optical pulses which are capable of directly monitoring electron orbitals in atoms. *See* LASER PHOTOCHEMISTRY.

Optical frequency synthesizers. One of the most recent breakthroughs in laser spectroscopy has been the development of the optical frequency synthesizer, which has completely revolutionized the world of time and frequency standards, allowing the connection of the optical frequency domain with the radio frequency domain. Optical frequency synthsizers are the most precise measurement tool now available, with an inherent stability that has been tested beyond 10^{-18}.

An optical frequency synthesizer is based on a mode-locked pulsed laser capable of emitting pulses that can be made as short as a few femtoseconds. The frequency spectrum of a train of pulses emitted by such a laser is a comb of equally spaced modes (the frequency spacing is just the laser repetition rate) that can span more than one octave. It is possible to compare the two ends of the spectrum by comparing the modes at a given low frequency with the modes at twice that frequency. This can be achieved by frequency-doubling the low-frequency modes in a nonlinear crystal and observing on a fast photodiode the beatnote between the doubled modes and the high-frequency modes. This procedure allows one to know the frequency of each comb mode, but for an offset frequency. The offset frequency can be measured by comparing a suitably chosen mode of the comb with a laser beam stabilized on a stable frequency reference. Indeed, since the modes of the comb can be actively stabilized to the reference, this procedure effectively transfers the stability of the reference to each of the comb modes, which cover the entire visible range, and to the comb frequency spacing, which is in the radio-frequency domain. The intrinsic stability of optical frequency synthesizers is such that at the moment there is a lack of a sufficiently stable reference to take full advantage of the synthesizer. This is one of the main reasons forcing reconsideration of the definition of the second, abandoning the cesium microwave standard in favor of an optical transition either of a laser-cooled atom or of a single trapped ion. *See* FREQUENCY COMB.

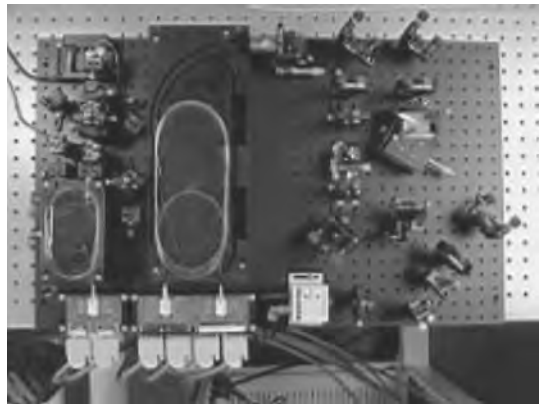

Fig. 3. An optical synthesizer based on a mode-locked fiber laser. The advantage here is that no parts missallign during operation because of slow temperature variations. Thefore this synthesizer can operate continuously for several days.

Other applications. Laser spectroscopy is virtually ubiquitous in natural sciences, but it is also extensively used in many other fields, some as seemingly remote as pollutant monitoring or art restoration. By laser spectroscopy one can readily obtain information on the chemical composition of a given material at a high rate without the need to mechanically touch the material. This is of utter importance in the fabrication of electronic components where spectroscopic techniques can be used to monitor deposition processes while they are being realized.

In analytical chemistry, laser spectroscopy is used for ultrasensitive detection of small concentrations of pollutants, tracer elements, or short-lived intermediate species in chemical reactions. It can be used for the investigation of collision-induced energy-transfer processes or to induce chemical reactions by excitation of one or more of the reactants. Ultrashort laser pulses have led to the birth of real-time femtochemistry where the formation or dissociation of molecules can be observed as it is happening. Laser spectroscopy can also be used for isotope identification and separation.

In biology, laser spectroscopy has been fundamental for the determination of the structure of biological molecules or of energy transfer in DNA, while the extreme time resolution allowed by ultrafast lasers allows the study of fast dynamical processes such as isomerization during photosynthesis.

In medicine, besides the many successful therapeutic applications of lasers, laser spectroscopy is now extensively considered as a diagnostic tool, for

example, in breath analysis or in cancer management.

Geophysics and environmental research benefit from laser spectroscopy in the detection of natural-gas emissions or in pollution monitoring. Of particular interest is the LIDAR (light detection and ranging) technique, in which a laser beam is sent into the atmosphere through an expanding telescope and backscattered off a distant object—for example, a volcanic cloud of ashes or a polluted building facade. By measuring the backscattered radiation with spectral and time resolution, it is possible to gain information on the object distance, extension, and composition.

Another example of a laser spectroscopy application is combustion monitoring, which can be performed in situ and in real time, for example, inside a functioning engine; or the measurement of flow velocities and turbulence.

Raman spectroscopy is widely used for the noninvasive determination of the chemical composition of paint in frescoes and paintings as well as stone analysis in sculpture.

Laser spectroscopy can also provide insights about the purity of material, its composition, and the quality of production processes in the fabrication of electronic circuits. Theo W. Hänsch; Francesco S. Cataliotti

Bibliography. D. L. Andrews (ed.), *Applied Laser Spectroscopy*, 1992; W. Demtröder, *Laser Spectroscopy: Basic Concepts and Instrumentation*, 3d ed., 2003; P. Hannaford (ed.), *Femtosecond Laser Spectroscopy*, 2005; T. W. Hänsch and M. Inguscio (eds.), *Frontiers in Laser Spectroscopy*, 1994; E. R. Menzel, *Laser Spectroscopy: Techniques and Applications*, 1995; S. Svanberg, *Atomic and Molecular Spectroscopy: Basic Aspects and Practical Applications*, 4th ed., 2004.

Laser welding

Welding with a laser beam. The primary apparatus used for both commercial application and industrial research and development is the continuous-wave, convectively cooled CO_2 laser with either oscillator/amplifier (gaussian output beam) or unstable resonator (hollows output beam) optics. These lasers are available in output powers ranging from approximately 1000 to 15,000 W. Substantial advances in laser technology made possible the production of fully automated multikilowatt industrial laser systems which can be operated on a continuous production basis. These systems can be used for a variety of development programs and on-line production applications.

Welding of HY-130 steel. A significant metallurgical phenomenon was reported following the laser welding of HY-130 alloy steel. After welding with a 7-kW continuous-wave coaxial-flow CO_2 laser with oscillator-amplifier optics, concurrent increases in the hardness, tensile strength, and impact energy of the welds were noted when compared with base metal properties. Welding was accomplished in the open atmosphere under gas shielding with no preheating, postheating, or filler metal additions. A decrease in the visible inclusion content was noted both on the impact fracture surface (**Fig. 1**) and in metallographic sections. Chemical analysis revealed significant decreases in the oxygen and nitrogen contents of the weld metal as compared to the base metal. It was concluded that during laser welding of this alloy a purification of the fusion zone with respect to inclusions, and perhaps a change to a more favorable inclusion size distribution, occurs.

Single- vs dual-pass welds. Single- and dual-pass laser welds have been made in an alloy of the 80,000-lb/in.^{2} (5.5×10^8 pascals) yield strength class. **Figure 2** shows cross sections of these welds. The welds were formed with a 12-kW beam from a continuous-wave CO_2 laser. Welding speeds of 25 and 30 in./min (1 and 1.3 cm/s) and 60 and 65 in./min (2.5 and 2.8 cm/s) were used for the single- and dual-pass welds, respectively. The welds were evaluated and exhibited excellent overall mechanical properties. Dual-pass welds exhibited a ductile-to-brittle transition temperature below −60°F (−51.1°C). The increased shelf energy was attributed to a reduction in the visible inclusion content of the fusion zone, while transition temperature was shown to be strongly dependent on grain size.

Titanium alloys. Two in-depth studies involving laser welding of Ti-6Al-4V alloy were conducted. In one study, laser welding was compared with electron-beam and plasma-arc welding. Sound welds were made by all processes, and the tensile properties of these welds exceeded base metal properties. The laser and electron-beam welds, because of their lower specific energies for fabrication, were substantially finer-grained and possessed finer substructures than the plasma-arc welds. Due to the finer microstructure of the beam welds, they exhibited lower K_{IC} (fracture toughness in plane strain) values than the plasma-arc welds, which had a coarser structure. The finer structure of the laser and electron-beam welds showed potential for good fatigue crack initiation resistance.

A comparison study of the fatigue behavior of laser and plasma-arc welds was also conducted. It was found that both techniques were capable of making weld specimens where fatigue failures initiated in the base metal; however, there were a number of specimens welded by each process in which fractures initiated at pores in the weld.

Aluminum alloys. While successful weld penetrations have been made in some aluminum alloys, and reasonable-looking beads have been produced, no extensive weld-quality studies or mechanical-property studies have been reported. Most weld penetrations in aluminum alloys show some porosity, but it appears that the potential of welding aluminum does exist.

Industrial applications. A number of multikilowatt laser welding systems are in use for industrial welding and heat-treating applications, including Ford Motor Company, Western Electric Company, Caterpillar Tractor, and General Motors Corporation.

A system which clearly displays the advantages of laser processing is the underbody welding system at

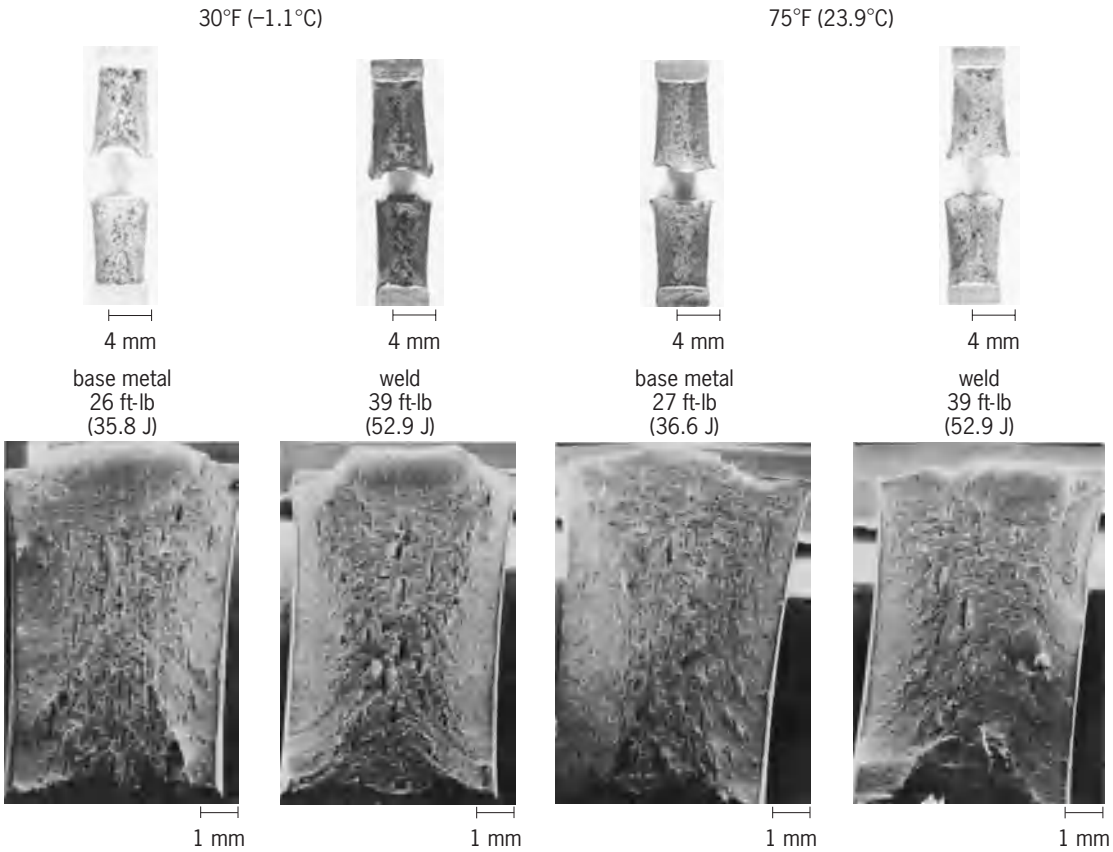

Fig. 1. A change in the shape and distribution of the inclusions on the fracture surfaces of HY-130 weld and base metal impact specimens is shown. The top row comprises low-magnification macrophotographs of both halves of the fractured specimens. The bottom row comprises higher-magnification scanning electron micrographs of one-half of the fracture. Welds were made at 5 kW and a speed of 45 in./min (114 cm/min). (*From E. M. Breinan and C. M. Banas, Fusion zone purification during welding with high power CO_2 lasers, Proceedings of the 2d International Symposium of the Japan Welding Society, Osaka, August 24–28, 1975*)

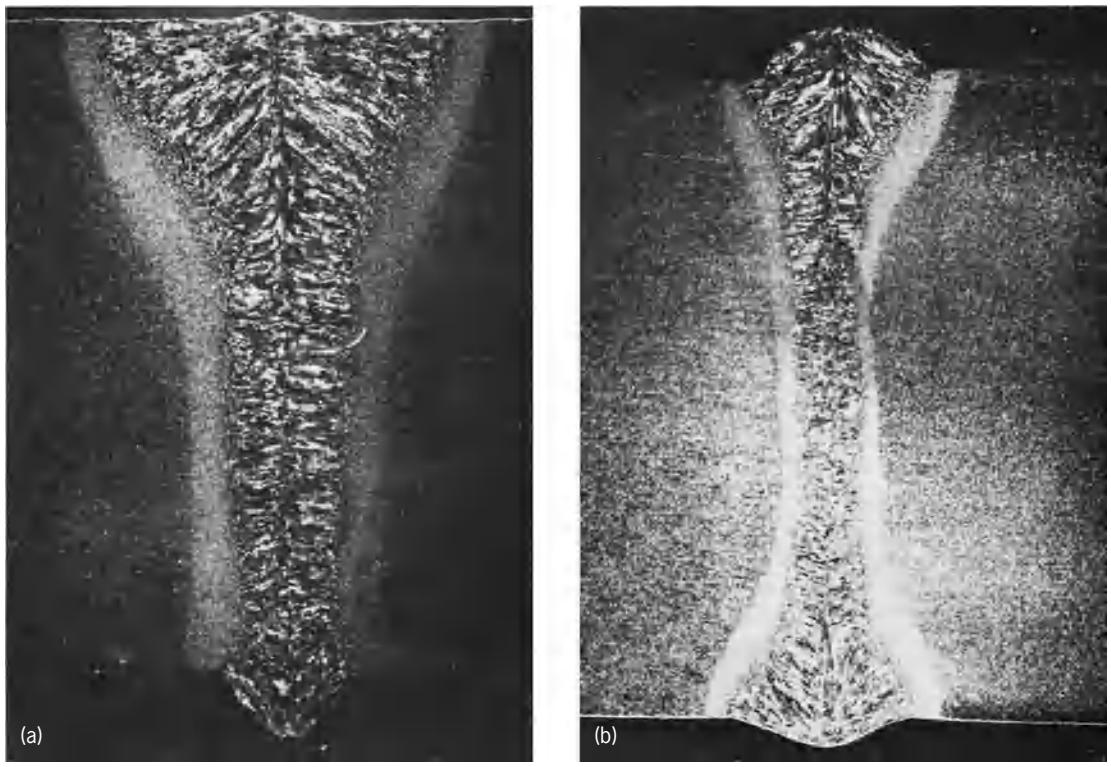

Fig. 2. Cross sections (0.5-in.- or 1.3-cm-thick plate) of laser welds in X-80 arctic pipeline steel. (*a*) Single-pass weld. (*b*) Dual-pass weld. (*From E. M. Breinan and C. M. Banas, Weld. Res. Counc. Bull. 201, pp. 47–57, December 1974*)

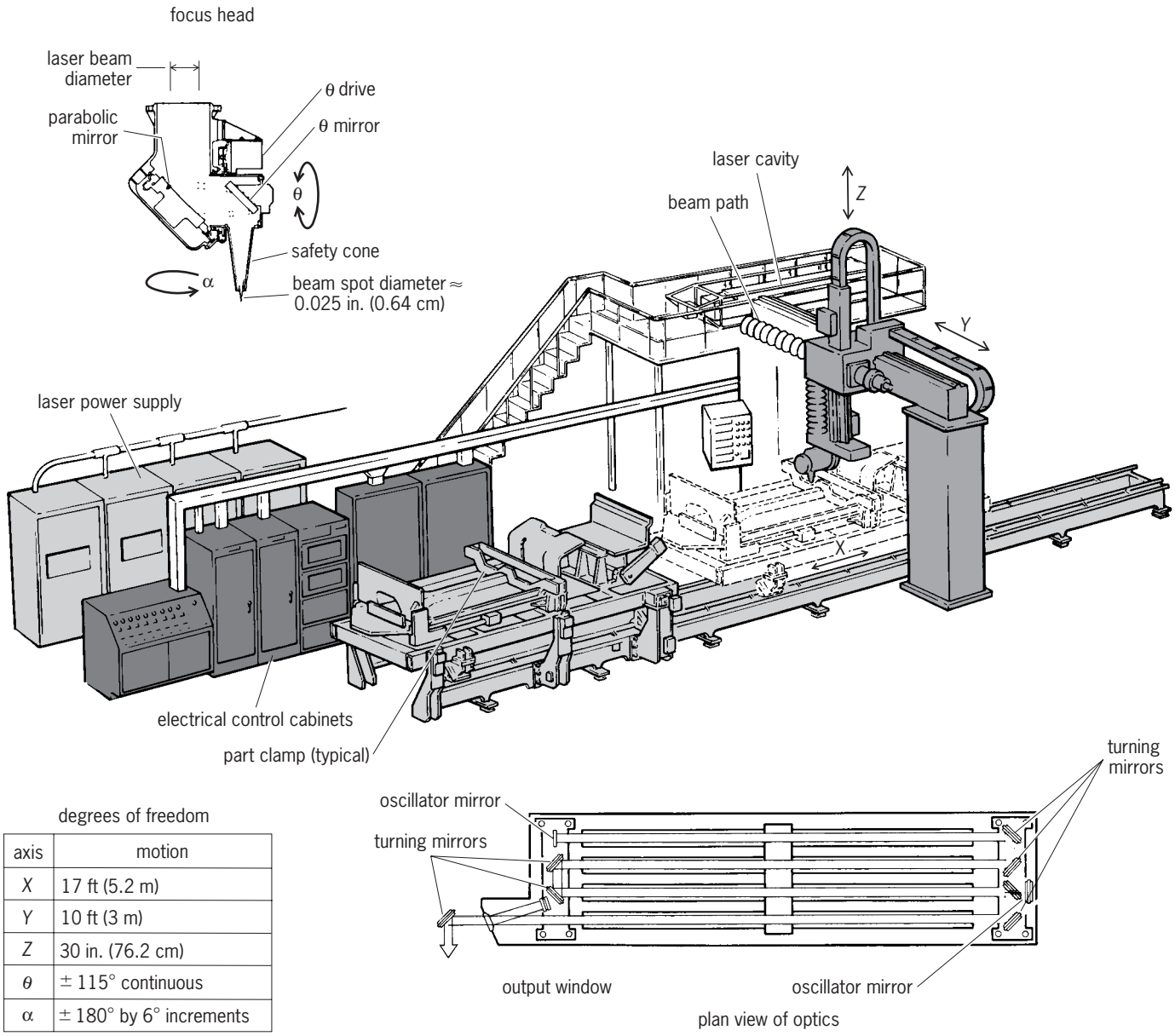

degrees of freedom

axis	motion
X	17 ft (5.2 m)
Y	10 ft (3 m)
Z	30 in. (76.2 cm)
θ	± 115° continuous
α	± 180° by 6° increments

Fig. 3. Schematic diagram of 6-kW multiaxis laser welding system at Ford Motor Company; arrows indicate direction of motion along various axes. This is the first complete high-power laser welding system designed to perform a high-production-rate industrial fabrication operation. (*Hamilton Standard Division*, *United Technologies Corp.*)

Ford (**Fig. 3**). As illustrated, a total of five degrees of freedom are achieved in movement of the focused beam relative to the intricately curved underbody panels. Only one of these, the X direction, is achieved by mechanical actuation of the parts; all the rest are achieved by rotation or translation of water-cooled copper mirrors. A unique 90° off-axis parabolic focusing mirror turns the beam 90° as it focuses the beam, thus allowing easy implementation of the α and θ axes. The focused beam creates a continuous lap weld in the thin steel sheet to a depth of 0.06 in. (0.15 cm) and at a speed of 500 in./min (21 cm/s). The continuous weld achieves much greater strength than spot welding techniques, and assures leak-tightness of the joint. At that welding speed, four large panels are joined together into an underbody in a total cycle time of 1 min.

Future applications. As the power available in commercial lasers increases, the scope of applications is also increasing. In addition to assembly-line operations, several opportunities are seen in the area of welding of heavy sections for industrial purposes including shipbuilding, pipeline welding, and general heavy steel fabrication. The use of added filler metal, which has already proved feasible with laser, promises to extend the thickness capability of the laser welding process.

Laser welding has several potential advantages for deep welding tasks. For one thing, high welding speeds are achievable. In addition, the lack of necessity for filler metal or substantial reduction in the amount of filler required, promises economic advantages for the laser process, particularly when welding arctic pipeline steels and in other applications

normally requiring large quantities of expensive filler metal. Reductions in edge preparation cost may also be realized due to the elimination of the edge chamfer required, although some machining of the butt edges is usually necessary to achieve a joint fitup which is suitable for laser welding. Another important factor is the ability of a single laser installation to supply a beam to several different welding stations, by appropriate indexing of turning mirrors. This capability, along with the high welding speed, allows high productivity on a laser welding installation.

A final advantage to laser welding is in the quality of the joint, apart from the low cost of welding. For quality-critical applications, such as welding of nuclear reactor components or arctic pipeline steel, the joint quality alone may prove to be an important factor in the selection of laser welding. *See* WELDING AND CUTTING OF MATERIALS. Edward M. Breinan

Bibliography. C. Dawes, *Laser Welding*, 1992.

Lateral line system

A primitive vertebrate sensory system that is present in all larval and adult fishes, in larval amphibians (such as tadpoles), and in some adult amphibians that retain an aquatic lifestyle (including the clawed frog). It is absent from all reptiles, birds, and mammals, even those that are aquatic (such as turtles, dolphins, and whales). The lateral line system consists of 100 or more sensory organs (neuromasts) that are typically arranged in lines on or just under the skin of the head and body. Neuromasts are composed of sensory hair cells, which are also found in the auditory system of all vertebrates. The lateral line system responds to water flowing past the skin surface and uses different flow patterns over the body to form hydrodynamic images of the animal's nearby surroundings, just as the visual system forms visual images of the environment using different light patterns on the retina.

Structure and function. Neuromast receptor organs are generally round or oval and can range from 20 micrometers to over 1000 μm (0.001 to 0.04 in.) in length. A neuromast has a central population of up to hundreds of mechanosensory hair cells that are surrounded by support cells. Hair cells function as directional sensors that convey information to the brain about both the strength and direction of water currents. The ciliary bundle of each hair cell contains one long cilium (kinocilium) and a cluster of shorter cilia (stereocilia), which lead up to the kinocilium in a stepwise fashion—the shortest being farthest away from the kinocilium and the longest closest to the kinocilium (**Fig. 1**). The ciliary bundles of the hair cells are embedded in a gelatinous cupula. Hair cells are activated when water flows past the skin surface, causing the cupula to move, thus causing the cilia to bend. The neural response of each hair cell is proportional to both the degree of cilia displacement and the direction in which the stereocilia are displaced relative to the eccentrically placed kinocilium of each hair cell (Fig. 1).

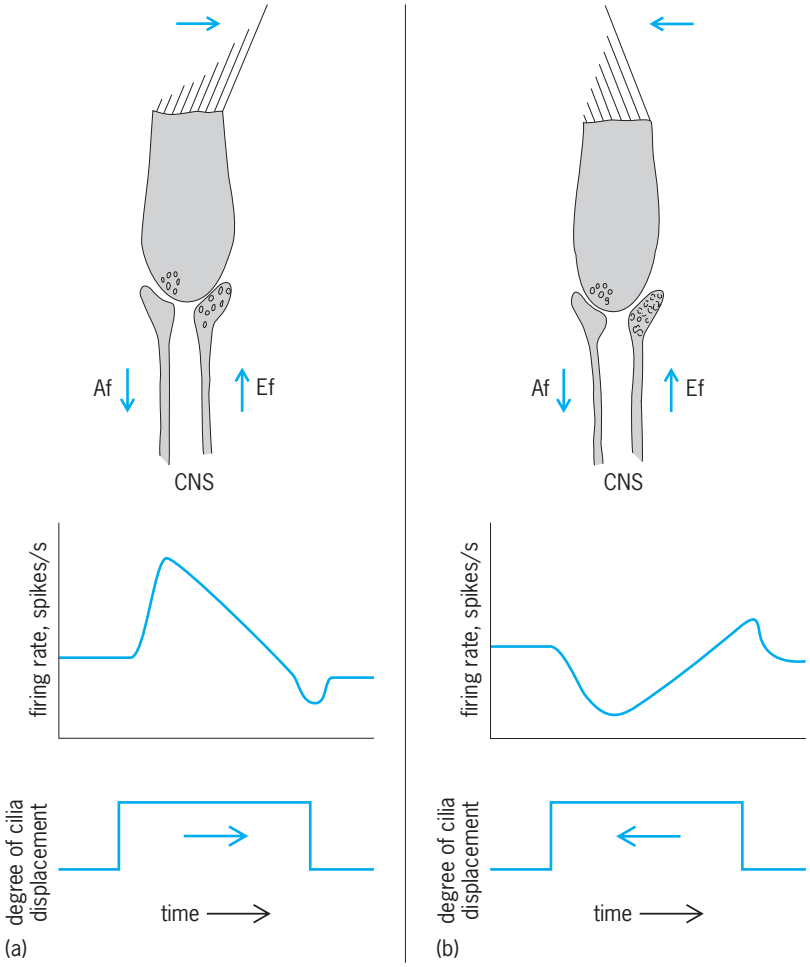

Fig. 1. Hair cells of the lateral line system responding to water flows in two opposing directions. (*a*) Stereocilia displaced toward the elongated kinocilium produce an increase in the firing rate of afferent (Af = afferent, Ef = efferent) fibers carrying information from the hair cell to the central nervous system (CNS). (*b*) Stereocilia displaced away from the kinocilium result in a decrease in neural activity.

Distribution on the head and body. Neuromasts are found just under the skin in fluid-filled canals that communicate with the skin surface through a series of pores (canal neuromasts, found in fishes only), or on the skin surface (superficial neuromasts, found in fishes and amphibians). A prominent canal that often forms a visible line along the trunk of most fishes is probably the origin of the term "lateral line," but in reality the lateral line system includes neuromasts that are distributed all over the head and body of the animal. In fishes, lateral line canals on the head are embedded in bone (bony fishes) or cartilage (cartilaginous sharks, skates and rays), whereas body (trunk) canals are contained in or between scales. Head canals in bony fishes may also vary in width, number of pores, and degree of development. In amphibians, canals are not present. Instead, short rows of up to 10 or so superficial neuromasts form a stitch, and these rows are arranged in lines along the head and trunk of the animal. Canals in fishes and stitch lines in amphibians are distributed in relatively stereotyped patterns (such as above and below the eye, down the cheek, along the lower jaw, across the top of the head, and down the main body) [**Fig. 2**].

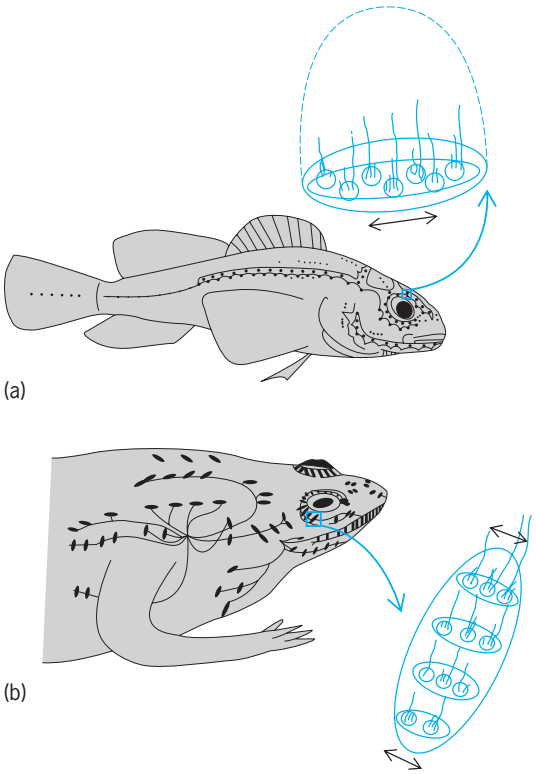

Fig. 2. Lateral line system of a typical bony fish and amphibian. (*a*) In bony fish, neuromasts (black dots) are found on the skin and in pored canals. The hair cells of each neuromast (closeup) are overlain by a dome-shaped gelatinous cupula (dotted line). In amphibians, neuromasts are found only on the skin in stitches (black ovals) that are composed of several neuromasts (closeup).

Superficial neuromasts in fishes are usually smaller (less than 100 μm or 0.004 in. in diameter) and contain fewer hair cells (less than 100) than canal neuromasts. They may occur in rows or singly, sitting flush with the epithelium, in small pits, or on top of papillae or ridges of skin. The number and distribution of superficial neuromasts can also vary considerably among species, ranging from less than 50 distributed in a few (about 10) distinct locations to more than 1000 distributed all over the head and body. *See* TELEOSTEI.

Processing of information. Information from neuromasts is transmitted to the brain by sensory (afferent) nerve fibers, which form five cranial nerves (the lateral line nerves) that terminate in distinct medullary regions of the brainstem (medulla oblongata). Contrary to many earlier published reports, these nerves are distinct from cranial nerves V, VII, IX, and X, so fishes and amphibians have more than the 12 cranial nerves traditionally described in mammals. Distinct regions and pathways in the brain are dedicated to processing information from the lateral line system. These are similar in overall organization, and in proximity to regions of the brain dedicated to processing information from two other closely allied sensory systems, the auditory system and the electrosensory system. Information is also carried from the brain to the sense organs by efferent nerve fibers, which can modulate the sensitivity of the organs to certain stimuli (for example, to reduce sensitivity when water flows are produced by the animal's own movements). *See* AMPHIBIA.

Behavioral significance. The lateral line system is thought to have a function that is separate from, but intermediate between, touch and hearing, and is best described as a sense of touch-at-a-distance. The lateral line system responds to both unidirectional (DC) and oscillatory (AC) water flows that change directions up to 200 times (cycles) per second. In general, large-scale water movements such as oceanic currents, tides, and river flows that are strong enough to carry a fish with them are not by themselves very effective lateral line stimuli. Smaller-scale movements, such as those produced by a slowly moving (less than 8 cm or 0.25 ft/s) stream or by nearby animals (less than one or two body lengths away), can be effective lateral line stimuli. Living organisms can produce water flows in different ways (such as inhalant and exhalant respiratory currents produced by bivalve mollusks, the turbulent wakes left by swimming fishes, and the fine antennal movements of tiny planktonic invertebrates).

Fishes can use different types of water flows to form hydrodynamic images of their surroundings. They can form images passively by remaining still and simply detecting the water currents created by other moving animals, or by detecting current distortions or turbulent wakes created by a stationary obstacle in moving water. Alternatively, fishes can actively form images by swimming past a stationary obstacle and then detecting the distortions in their own self-generated flows due to the presence of the obstacle. In some fishes, such as blind cavefishes, thousands of superficial neuromasts that are over the entire body surface are specialized for active imaging. These fishes are capable of determining very fine spatial details (less than 2 mm or 0.08 in.) of their surroundings.

Fishes and amphibians can also use their lateral line system to orient themselves relative to a water current (rheotaxis), hold a stationary position in a stream, capture prey, avoid predators, and communicate with intraspecifics (**Fig. 3**). Many stream-dwelling fishes (such as trout and salmon) show rheotactic and station-holding behaviors by orienting their bodies upstream and holding positions behind stationary rocks or boulders. These behaviors are important for the upstream spawning migrations of these fishes and for capturing prey that are being carried downstream.

Many different species use their lateral line system to detect prey. Some fishes (such as the topminnow) and amphibians (such as the clawed toad) have specialized canals or receptor organs on their dorsal surface for detecting surface waves produced by struggling insects. Others (such as the Atlantic stingray) appear to have specializations on their ventral surface for detecting benthic prey. Although the functions of the lateral line system in deep-sea fishes have not yet been determined, lateral line morphology in these fishes is frequently very distinctive, with widened canals covering nearly the entire surface

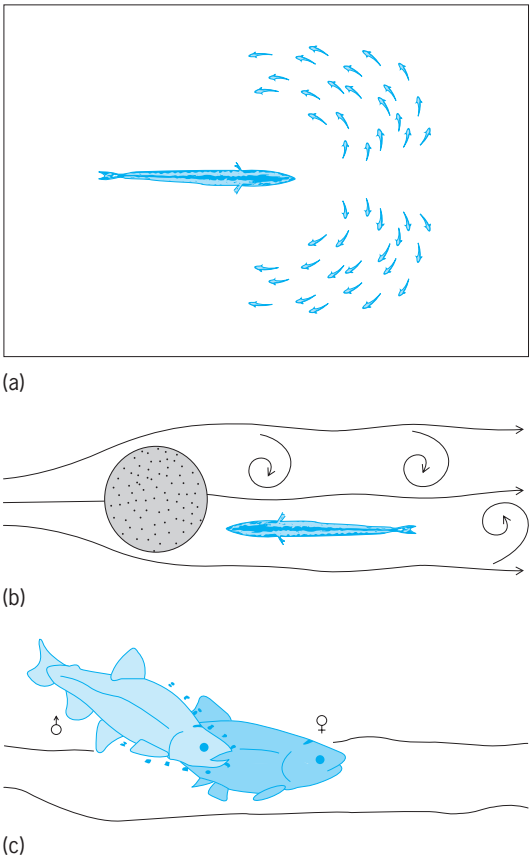

Fig. 3. Uses of lateral line system. (*a*) Rapid schooling maneuver made by a school of herring to escape a rapidly approaching predator. (*b*) Brook trout orienting upstream and holding station behind a large boulder. (*c*) Male salmon quivering next to a female salmon in preparation for mating.

area of the head (as in Poromitra) and prominent superficial neuromasts that sit on the ends of papillae, which elevate the receptors above the skin surface (as seen in the deep-sea anglerfish).

Lateral line specializations are also present in other groups of fishes, including herring, which have a unique connection between a lateral line canal on the head and an air cavity inside the head (near the ear) that functions as an accessory auditory structure. Herring also have extensively branched tubules that connect canals to the outer skin surface and that terminate in many pores. These specializations may allow these fishes to make rapid and coordinated schooling maneuvers to avoid predation. The body movements of nearby fishes can be detected by the lateral line system, and fishes can use this information to help match their swimming velocities and maintain constant distances with adjacent fishes in a school. Body vibrations exchanged between males and females during elaborate courtship dances may serve as another form of intraspecific communication. In land-locked salmon, these vibrational signals are detected by the lateral line system, and they are used to help synchronize the release of gametes (sperm and eggs) so that external fertilization can occur. *See* ELASMOBRANCHII; NERVOUS SYSTEM (VERTEBRATE).

Sheryl Coombs; Jacqueline F. Webb

Bibliography. J. Atema et al. (eds.), *Sensory Biology of Aquatic Animals*, Springer-Verlag, New York, 1988; S. Coombs, P. Görner, and H. Münz (eds.), *The Mechanosensory Lateral Line: Neurobiology and Evolution*, Springer-Verlag, New York, 1989; J. Montgomery, S. Coombs, and M. Halstead, Biology of the mechanosensory lateral line in fishes, *Rev. Fish Biol. Fisheries*, 5:399–416, 1995.

Lateral meristem

Strips or cylinders of dividing cells located parallel to the long axis of the organ in which they occur. Radial enlargement of the cells derived from these meristems increases the diameter of the organ. The lateral meristem is concerned with secondary growth in the sense that its meristematic activity adds cells to the primary body which was derived from the apical meristems. *See* APICAL MERISTEM.

Vascular cambium. The vascular cambium produces secondary xylem on its inner face and secondary phloem on its outer face. In a woody plant the cambium occurs as a continuous sheath, or hollow cylinder, separating xylem and phloem. In stems with separate vascular bundles, fascicular cambium arises within the bundles as strips from the residual procambial cells lying between the primary xylem and primary phloem. In plants with uninterrupted cylinders of secondary vascular tissues, interfascicular cambiums often arise in the ray parenchyma between the bundles in lateral connection with the fascicular cambial strips within the vascular bundles; thus the alternating strips of fascicular and interfascicular cambium form a complete cambial cylinder (**Fig. 1**). The vascular cambium in woody roots arises as strips on the inner faces of the primary phloem groups, and it becomes a laterally continuous sheath when the pericycle cells outside the primary xylem ridges become transformed into cambial initials. *See* PHLOEM; ROOT (BOTANY); XYLEM.

The vascular cambium has two kinds of initial cells: ray and fusiform, or spindle-shaped (**Fig. 2**). Because the immediate derivatives of these cambial cells are quite similar to the initials and may undergo further division, these derivatives and the initials together are said to constitute the cambial zone. The zone may be several cell layers in radial dimension. The composition of the cambium itself, even in a given plant, is not static, for initials may be lost or gained or markedly changed in form. The cambium keeps pace circumferentially with the increasing girth of an organ by self-multiplication of the fusiform initials and by the origin of entirely new groups of ray initials.

Fusiform initials. These are vertically elongate cells, brick-shaped in transectional view, with tapering ends, and with an extensive vacuolar system. These initials give rise to all cells in the vertical system of secondary xylem and phloem. As seen in tangential views, storied cambium has an orderly arrangement in horizontal tiers of uniformly short fusiform initials whose ends slightly overlap those of

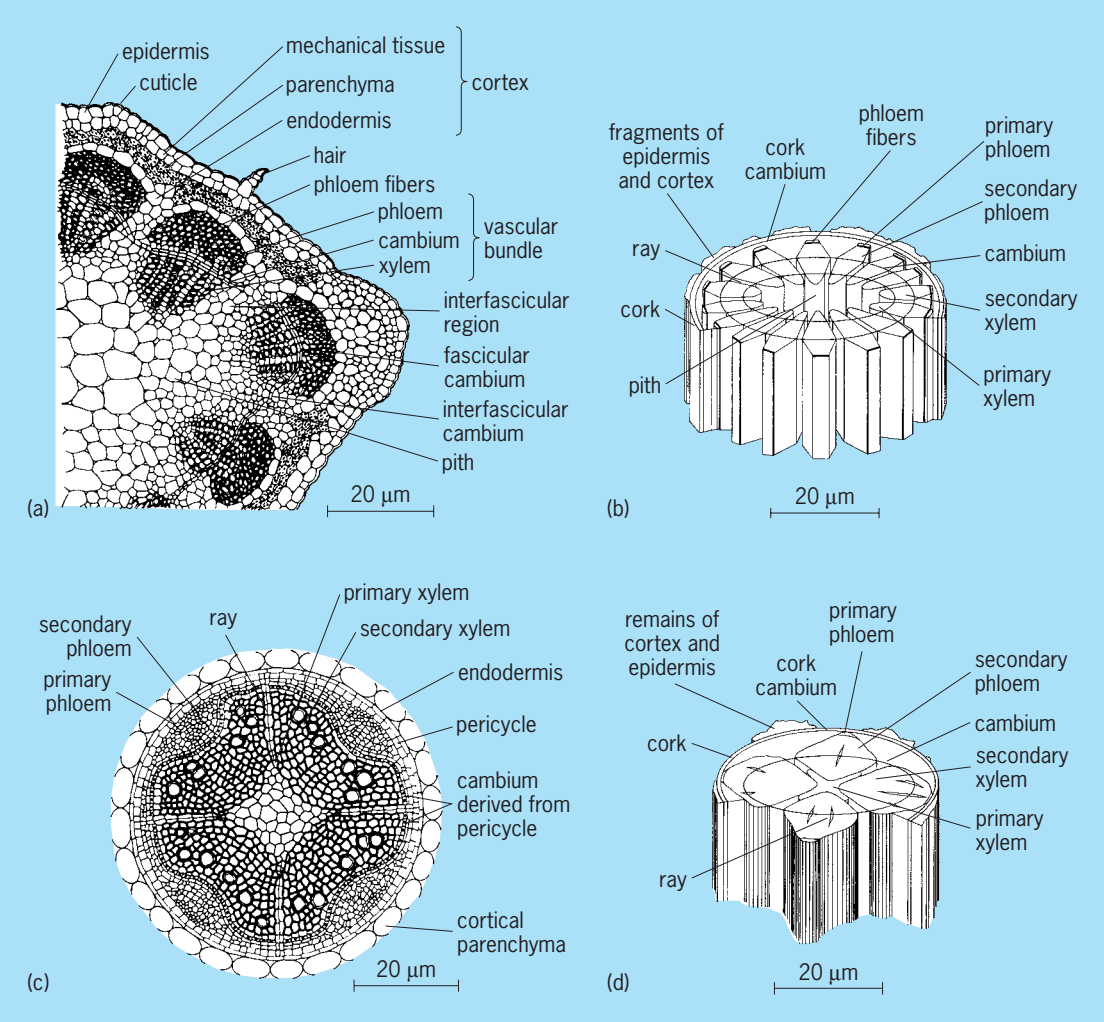

Fig. 1. Diagrams of stem and root. (*a*) Part of a cross section of an alfalfa stem. (*b*) Three-dimensional diagram of a dicotyledonous stem, showing the results of cambial activity. (*c*) Cross section of the vascular cylinder of a dicotyledonous root at level where cambium has become continuous layer and has produced considerable amounts of secondary tissues. (*d*) Three-dimensional diagram of part of root with considerable secondary thickening. (*After G. M. Smith et al.*, *Textbook of General Botany*, *5th ed.*, *Macmillan*, *1953*)

initials in juxtaposed tiers. The more primitive and less orderly nonstoried cambium lacks obvious tiers, and has longer initials of variable length whose ends greatly overlap those of other initials. The radial walls of fusiform initials are relatively thick (with many depressed primary pit fields) in the dormant season, but become much thinner during part of the growing season when the cells are actively dividing. Divisions in the tangential (periclinal) plane produce secondary tissues. Multiplication of initials—which may occur in the dormant season—results from radial (anticlinal) divisions. The separating walls in these divisions may be highly oblique to almost transverse in long initials, or approximately vertical in short initials. The derivatives of oblique divisions elongate by apical intrusive growth.

Ray initials. These cells are nearly isodiametric in form and, though much shorter vertically, often have a wider radial dimension than do fusiform initials. Ray initials divide periclinally in producing new phloem and xylem ray cells. They may also divide

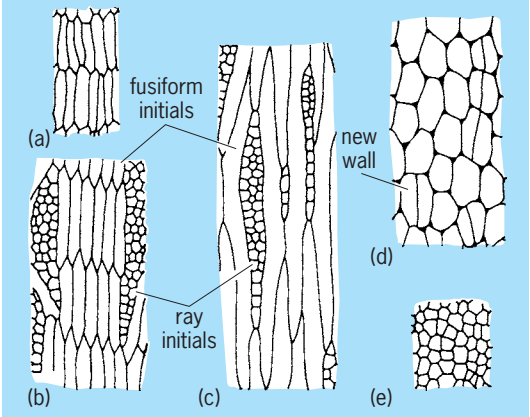

Fig. 2. Tangential views of cambium. (*a*) Root, *Bougainvillea spectabilis*, cork cambium with but one type of initial which is elongate polygonal in shape. (*b*) Stem, *Robinia pseudoacacia*, storied vascular cambium with ray and short fusiform initials. (*c*) Stem, *Juglans regia*, nonstoried vascular cambium with ray and long fusiform initials. (*d*) Stem, *Yucca gloriosa* (a monocotyledon), unusual vascular cambium with but one type of initial. (*e*) Stem, *Lindera benzoin*, cork cambium with but one type of initial, nearly isodiametric in shape.

anticlinally or transversely and thus add to the number of initials in a group. An entirely new group of ray initials may have its origin in a segment cut off from the end of a fusiform initial or through subdivision of an entire fusiform initial.

Cork cambium. This cambium (phellogen) produces cork (phellem) on its outer face and sometimes phelloderm (parenchyma) on its inner face. The initials, unlike those in the vascular cambium, are of a single type. They are short and appear rectangular and flattened in transectional and radial views, and rectangular to polygonal in tangential aspect. Cork and phelloderm are formed by periclinal divisions. The increase in the circumference of the cambium occurs by anticlinal divisions. The phellogen initials in certain monocotyledons divide only a few times and are then transformed into cork cells themselves. *See* PERIDERM.

Other cambiums. The initials responsible for production of secondary tissues in certain woody monocotyledons arise in fundamental parenchyma outside the primary bundles. The initials are short and are variable in shape in tangential aspect. They divide in several vertical planes in producing secondary vascular bundles, but strictly periclinally in forming secondary parenchyma.

Some vascular cambiums differ from the ordinary type chiefly in their unorthodox activity. For example, several concentrically oriented cambiums are simultaneously active in the garden beet. Cambiums in some species may remain active for a time and then become transformed into vascular tissue. New cambium then originates outside the secondary phloem. Some cambiums add both xylem and phloem to the inner side. *See* PLANT TISSUE SYSTEMS.

Vernon I. Cheadle

Laterite

The name given by F. Buchanan in 1807 to the iron-rich weathering product of basalt in southern India. The term is now used in a compositional sense for weathering products composed principally of the oxides and hydrous oxides of iron, aluminum, titanium, and manganese. Iron-rich or ferruginous laterite is largely hematite, Fe_2O_3, and goethite, $HFeO_2$, and may be an ore of iron and nickel (Cuba, New Caledonia). Aluminous laterite is composed of gibbsite and boehmite, and is the principal ore of aluminum. Clay minerals of the kaolin group are typically associated with, and are genetically related to, laterite. Laterites range from soft, earthy, porous material to hard, dense rock.

Concretionary forms of varying size and shape commonly are developed. The color depends on the content of iron oxides and ranges from white to dark red or brown, commonly variegated.

Origin. Laterite is formed by weathering under conditions that lead to the removal of silica, alkalies, and alkaline earths. The resulting concentrations of iron and aluminum oxides sharply differentiate lateritization from temperate-climate weathering in which the end product is largely clay minerals (hydrous aluminum silicates). Early workers from temperate regions considered lateritization as profound leaching (desilication) in a stage beyond ordinary kaolinization. Studies in tropical regions, however, show that weathering of alkaline silicates may yield gibbsite directly without passing through an intermediate clay stage.

Investigations in many parts of the world stress certain genetic factors. A tropical to subtropical climate with high temperature and abundant rainfall, seasonal or at least with periods of marked dryness, is fundamental. Relief sufficient to ensure good drainage is requisite: lateritic soils are permeable and do not erode by sheet wash like clay or shale. Aluminous laterite forms above the water table and may grade to clay in depth. It may be found on hills and on well-drained slopes, but the residual deposits in adjacent valleys usually are kaolin.

Parent materials. The parent material controls or greatly influences the composition of laterites, which may be developed from a variety of igneous, metamorphic, and sedimentary rocks. Iron-rich rocks (peridotite) yield iron ore; aluminous rocks (syenite) produce bauxite; whereas andesite or basalt give intermediate products. Commonly textural and structural features of the parent material are preserved and the more resistant insoluble minerals remain.

Mature lateritic soils lack fertility for most systems of agriculture. Savannas or parklike grasslands are typical on laterite. Clay, not laterite, is found beneath rainforests and jungle vegetation. *See* BAUXITE; CLAY MINERALS; KAOLINITE; WEATHERING PROCESSES.

Samuel S. Goldich

Latitude and longitude

The latitude of a location specifies the angle between an imaginary line directed generally toward the center of the Earth and the Equator. The longitude measures the angle between the meridian (the plane defined by the Earth's axis and this local reference direction) and the plane of the Greenwich meridian (**illus.** *a*).

Astronomical coordinates. Astronomical (or astronomic) latitude and longitude use the direction of gravity for the reference direction. This direction, known as the astronomical vertical, is perpendicular to the equipotential surface of the Earth's gravitational field at the location of the observer. *See* ASTRONOMICAL COORDINATE SYSTEMS.

Geodetic coordinates. A particular geopotential surface approximating mean sea level in the open ocean is called the geoid. A mathematical surface in the form of an oblate ellipsoid may be constructed to approximate the geoid. The direction perpendicular to this reference ellipsoid at the observer's location is used as the reference direction in defining geodetic latitude and longitude. A geodetic datum is defined by its reference ellipsoid, the adopted coordinates of a reference station, and the azimuth of a reference

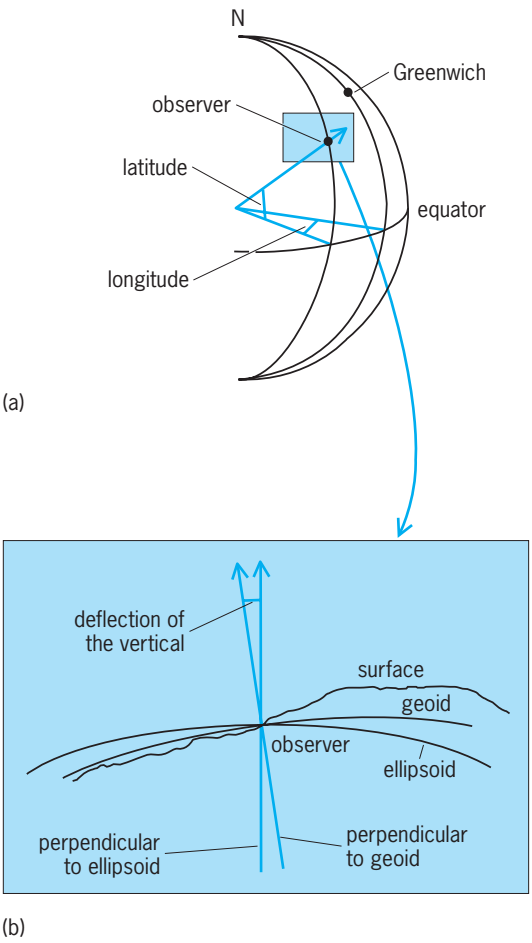

Latitude and longitude. (*a*) Angles formed by reference direction, defining latitude and longitude. (*b*) Detail, showing angle between reference directions in astronomical and geodetic coordinates, the deflection of the vertical.

line. Many geodetic datums are in use. *See* MATHEMATICAL GEOGRAPHY.

Geocentric coordinates. Geocentric latitude and longitude are defined by a reference direction which passes precisely through the center of mass of the Earth. These coordinates are determined mathematically from the geodetic latitude and longitude, assuming a fixed relationship between the center of the geodetic datum and the center of mass and knowing the mathematical shape of the ellipsoid. *See* GEODESY.

Determination. Astronomical latitude is determined by observing the altitude on the meridian of a celestial object whose declination is known. Astronomical longitude is identical with the difference between Universal Time and local mean time. Local mean time is determined by observing the time of meridian transits of celestial objects with known positions. The direction of gravity is established with the aid of a spirit level, a liquid surface, or observation of the horizon corrected for the elevation of the observer. It varies locally, and is generally not directed toward the center of the Earth. *See* TIME.

Geodetic latitude and longitude are determined by referring precise measurements of distance and direction, made with geodetic instruments on the surface of the Earth, to the datum. Geodetic coordinates are used for location of sites, while astronomical latitude and longitude are used to determine the angular orientation of celestial objects with respect to observers on the Earth. The difference between astronomical and geodetic coordinates is called the deflection of the vertical (illus. *b*), and commonly amounts to some seconds of arc, occasionally reaching a minute of arc. *See* EARTH, GRAVITY FIELD OF; GEODESY.

Dennis D. McCarthy

Lattice (mathematics)

Lattice theory deals with properties of order and inclusion, much as group theory treats symmetry. As a generalization of boolean algebra, lattice theory was first applied around 1900 by R. Dedekind to algebraic number theory; however, its recognition as a major branch of mathematics, unifying various aspects of algebra, geometry, and functional analysis, as well as of set theory, logic, and probability (to which boolean algebra had already been applied), dates from the years 1933–1938. *See* BOOLEAN ALGEBRA; GROUP THEORY; SET THEORY.

The most basic concept of lattice theory is that of a partial ordering of a set S of elements $x, y, z, \ldots$. By this is meant a binary relation, usually denoted $\leqq$ (or $\geqq$), with the following properties:

(P1) $x \leqq x$ for all $x \in S$

(P2) If $x \leqq y$ and $y \leqq x$, then $x = y$

(P3) If $x \leqq y$ and $y \leqq z$, then $x \leqq z$

If $\leqq$ is any partial ordering of S, then its converse or dual $\geqq$, defined by statement (1), is also a partial

$$x \leqq y \text{ if and only if } y \leqq x \qquad (1)$$

ordering of S. This easily verified fact provides a fundamental duality principle, which is useful in many connections.

Suppose, for example, the join $x \cup y$ of two elements x and y of a partially ordered set S is defined by conditions (2*a*) and (2*b*). (It is easily shown that

$$x \leqq x \cup y \qquad y \leqq x \cup y \qquad (2a)$$

$$\text{If } x \leqq z \text{ and } y \leqq z\text{, then } x \cup y \leqq z \qquad (2b)$$

there is at most one such $x \cup y$.) Then the duality principle suggests defining the meet $x \cap y$ by Eqs. (3*a*) and (3*b*), and shows that there is at most one such $x \cap y$.

$$x \geqq x \cap y \qquad y \geqq x \cap y \qquad (3a)$$

$$\text{If } x \geqq z \text{ and } y \geqq z\text{, then } x \cap y \geqq z \qquad (3b)$$

A lattice is defined as a partially ordered set (poset, for short) in which any two elements x and y have a

meet $x \cap y$ and a join $x \cup y$. These binary operations satisfy the four basic identities:

(L1) $$x \cap x = x \cup x = x$$

(L2) $$x \cap y = y \cap x \text{ and } x \cup y = y \cup x$$

(L3) $$x \cap (y \cap z) = (x \cap y) \cap z$$

and $$x \cup (y \cup z) = (x \cup y) \cup z$$

(L4) $$x \cap (x \cup y) = x \cup (x \cap y) = x$$

The operations $\cap$ and $\cup$ are connected with the relation $\leqq$ by the condition that $x \leqq y$, $x \cap y = x$, and $x \cup y = y$ are three equivalent statements. Conversely, if L is an algebraic system with operations $\cap$ and $\cup$ satisfying (L1) to (L4) for all x, y, z, then the preceding condition defines $\leqq$ as a partial ordering of L, with respect to which $\cap$ and $\cup$ have the meanings defined above. This principle was discovered in 1880 by C. S. Peirce.

Kinds of lattices. There are many different kinds of lattices. Thus the real numbers form a lattice if $x \leqq y$ is given its usual meaning. This lattice is simply ordered, in the sense that

(P4) Given x and y, either $x \leqq y$ or $y \leqq x$

Any such simply ordered set (or chain) is a lattice, in which $x \cup y$ is simply the larger of x and y, and dually.

Again, the set J of positive integers forms a lattice, if one lets $m \leqq n$ mean "m divides n" (usually denoted $m \mid n$). In this case, $m \cap n = \gcd(m,n)$ and $m \cup n = \operatorname{lcm}(m,n)$. Still again, one can let Σ consist of all subsets $S, T, \ldots$ of a fixed ensemble I, and let $S \leqq T$ mean that every point in S is in T. Then Σ is a lattice, in which $S \cap T$ is the intersection of S and T, whereas $S \cup T$ is their union. Actually, Σ is a boolean algebra.

In all the preceding lattices, the distributive laws hold:

(L5) $$x \cap (y \cup z) = (x \cap y) \cup (x \cap z) \text{ and}$$
$$x \cup (y \cap z) = (x \cup y) \cap (x \cup z), \text{ for all } x, y, z$$

Such lattices are called distributive lattices. Any chain is a distributive lattice; so is any boolean algebra. More generally, a ring of sets is defined as a family Φ of subsets of a fixed set I which contains, with any S and T, also their intersections $S \cap T$ and their union $S \cup T$. Then any ring of sets is a distributive lattice.

It is obvious that each of the two identities of (L5) is dual to the other. It is a curious fact that, in a lattice, each also implies the other.

If G is any group, then its subgroups form a lattice, and its normal subgroups also form a lattice, in both cases under inclusion. The normal subgroups satisfy the (self-dual) modular law:

(L6) If $x \leqq z$, then $x \cup (y \cap z) = (x \cup y) \cap z$

In general, lattices satisfying (L6) are called modular and every distributive lattice is modular.

The lattice of all linear subspaces of the n-dimensional vector space $\mathbf{V}_n(F)$ over any field (or division ring) F is also a modular lattice, usually called

Fig. 1. Ordinal number 4.

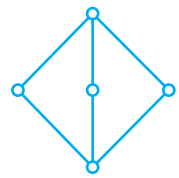

Fig. 2. The projective line over the field Z_2 of integers mod 2.

the $(n-1)$-dimensional projective geometry $P_{n-1}(F)$ over F. This lattice contains special elements $\mathbf{0}$ (the zero vector) and $I = V_n(F)$ (the whole space), such that

(P5) $$\mathbf{0} \leqq x \leqq I \text{ for all } x$$

Such special elements always exist in any lattice whose chains all have finite length, but they need not exist in general; for example, they do not in the simply ordered set of real numbers.

The lattice $P_{n-1}(F)$ is complemented, in the sense that each subspace x has at least one complement x', with the property

(L7) $$x \cap x' = \mathbf{0} \text{ and } x \cup x' = I$$

Thus $P_{n-1}(F)$ is a complemented modular lattice. Similarly, it may be verified that the class of boolean algebras is precisely the class of complemented distributive lattices. This principle enables one to consider boolean algebra as a branch of lattice theory.

Lattices L containing few elements can be conveniently visualized by diagrams. In these diagrams, small circles represent elements of L, a being higher than b whenever $a > b$. A segment is then drawn from a to b whenever $a > b$, but no x exists such that $a > x > b$. Any such diagram defines L up to isomophism: $a > b$ if and only if one can travel from a to b along a descending broken line. **Figures 1–6** are typical of such diagrams.

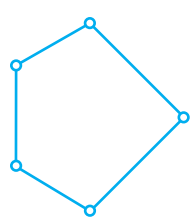

Fig. 3. The simplest nonmodular lattice.

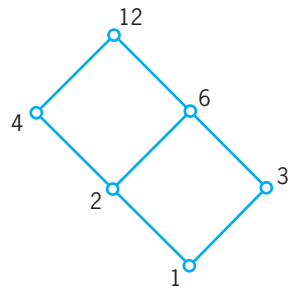

Fig. 4. The lattice of divisors of 12, under divisibility.

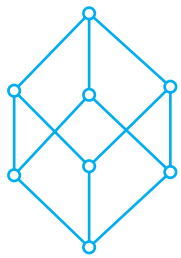

Fig. 5. The boolean algebra of order 8.

Such graphs often give useful information very simply. For example, let a finite lattice be called semimodular if any two elements a and b immediately above (covering) a given element c are also immediately under (covered by) another element $d = a \cup b$. This condition can easily be tested by inspection. Dedekind showed that a finite lattice L was modular if and only if it and its dual were both semimodular. For L to be distributive, the extra condition of containing no subgraph such as that of Fig. 2 is necessary and sufficient.

Applications to algebra and geometry. Lattices, like groups and rings, can be defined as abstract algebras, that is, as systems of elements combined by universally defined operations. These operations may be unary, binary, or ternary. In any such abstract algebra A, define a subset S to be a subalgebra of A when the result of performing any operation of A on elements in S is again in S. Call an equivalence relation $a \equiv b \pmod{\theta}$ on A a congruence relation when, for any n-ary operation f of A, $a_i \equiv b_i \pmod{\theta}$ for $i \equiv 1, \ldots, n$ implies that $f(a_1, \ldots, a_n \equiv f(b_1, \ldots, b_n) \pmod{\theta}$. Then the subalgebras of A form one (complete) lattice, and the congruence relations form another.

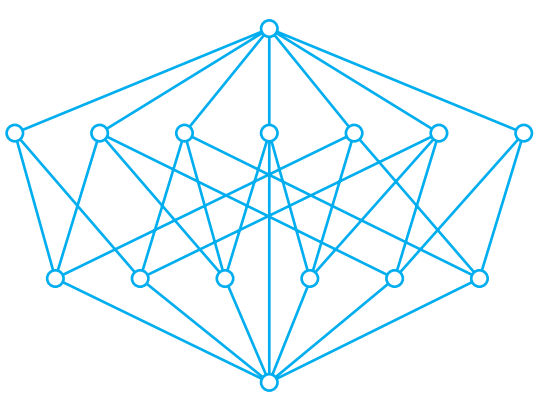

Fig. 6. Lattice of partitions of [*a*,*b*,*c*,*d*].

Such results, which are true of abstract algebras in general, are called theorems of universal algebra. Another such result is the theorem that any algebra can be decomposed into subdirectly irreducible algebras. From this it follows easily that any distributive lattice is isomorphic with a ring of sets.

In groups, rings, and many other algebras, all congruence relations are permutable; it follows that the congruence relations form a modular lattice. This fact, combined with the existence of one-element subalgebras, permits the development of an extensive structure theory that includes a unique factorization theorem under appropriate finiteness conditions. Curiously, lattices themselves satisfy a unique factorization theorem, but for another reason.

It is important to note that the lattice products LM, involved in the statement of the preceding unique factorization theorem, are one of the six basic operations of a general arithmetic of partially ordered sets, which contains the usual cardinal and ordinal arithmetic as a special case. Thus, the most general finite boolean algebra is just $\mathbf{2}^n$, where $\mathbf{2}$ is the ordinal number two and n is a finite cardinal number.

In any lattice with O, a "point" is an element that covers O. In most lattices arising in geometry, every element is a join of a finite number of points; semimodular lattices with this property are called geometric lattices. The lattice of all partitions of a finite set into nonoverlapping subsets is such a geometric lattice. Figure 6 depicts the geometric lattice of all partitions of four objects. Any geometric lattice is complemented; conversely, any complemented modular lattice in which all chains are finite is geometric.

It was mentioned above that the $(n - 1)$-dimensional projective geometry $P_{n-1}(F)$ over a division ring F was always a complemented modular lattice. Many interesting applications of lattice theory to geometry are extensions of this observation. For instance, let P be any abstract projective geometry, defined by incidence relations. Then P is also a complemented modular lattice. Conversely, any finite-dimensional complemented modular lattice is a product of projective geometries and a boolean algebra.

The preceding abstract combinatorial approach to projective geometries led to the construction in 1936, by J. von Neumann, of his continuous-dimensional projective geometries (so-called continuous geometries). O. Frink, in 1946, developed a parallel theory of projective geometries of discretely infinite dimension. By analogy, affine geometries can also be regarded as lattices, and any affine geometry in which all chains are finite is a geometric lattice.

Although important applications of geometric lattices to combinatorial problems have been made by G.-C. Rota and others since 1960, the theory of semimodular lattices in general is quite limited. A negative result is Dilworth's theorem, which states that

every finite lattice is isomorphic with a sublattice of a finite semimodular lattice. Similarly, P. M. Whitman has shown that every lattice is isomorphic with a sublattice of the lattice of all partitions of some (infinite) class. *See* ABSTRACT ALGEBRA; COMBINATORIAL THEORY.

Relation to set theory and logic. In set theory, one is frequently concerned with various special families of sets such as closed sets, open sets, measurable sets, and Borel sets. Some of these families form boolean algebras. But the closed sets (and, dually, the open sets) of a topological space X usually form just an uncomplemented distributive lattice $L(X)$. In $L(X)$, the complemented elements (which necessarily form a boolean algebra) are those which disconnect X.

Both the theory of finite sets and the theory of finite distributive lattices are fairly trivial. Each such lattice L has a unique representation as the cardinal power $L = \mathbf{2}^P$ of the two-element boolean algebra $\mathbf{2}$, with a general finite partially ordered set P as exponent. Moreover, any finite-dimensional distributive lattice is finite.

If L is an infinite-dimensional lattice, distributive or not, then one can define several topologies on L, which may or may not be equivalent. Most of these are suggested by a consideration of the special case that L is the rational number system. Adapting Dedekind cuts, as first shown in 1935 by H. M. MacNeille, one can extend any lattice (or partially ordered set) to a complete lattice, in which any subset A of elements x_α has a least upper bound (join) $\vee\, x_\alpha$ and greatest lower bound $\wedge\, x_\alpha$. In such a complete lattice, one can define lim sup $\{x_n\}$ and lim inf $\{x_x\}$ for any sequence $x_1, x_2, x_3, \ldots$ of elements. One can define a convergence topology in L by letting $x_n \to x$ mean that lim sup $\{x_n\}$ = lim inf $\{x_n\} = x$. Alternatively, one can define an interval topology by taking the closed intervals $[a, b]$, each consisting of all $x \in L$ with $a \leqq x \leqq b$, as subbasis of closed sets.

The concept of convergence, in lattices and other spaces, can be extended by considering directed sets of indices α, this concept itself being lattice-theoretic. To describe convergence in topological spaces not satisfying countability axioms, such directed sets are essential.

Using the concepts of lattice completeness and lattice topology, one can give interesting characterizations of various lattices associated with topological spaces. Thus, this was done for the lattice of regular open sets by S. Ulam and G. Birkhoff. It was done for the (distributive) lattice of continuous functions by Irving Kaplansky. Finally, it was done for closed (and open) sets by J. C. C. McKinsey and Alfred Tarski, all since 1935.

An interesting analogy between set theory and logic is provided by the equivalence between the algebra of open sets and brouwerian logic, which arose entirely from a consideration of formal logic. Abstractly, both deal with relatively pseudocomplemented distributive lattices.

Another type of logic, directly suggested by quantum theory, was developed in 1936 by G. Birkhoff and von Neumann. In this logic, properties form an orthocomplemented modular lattice. The aristotelian logic of classical mechanics appears as the special, distributive (hence boolean) case of permutable observations. The formal properties of the probability $p[x]$ of an event x are shown to be similar to those of dimension in projective geometry.

Lattice theory also gives perspective on other kinds of formal logic, including modal logic and "strict implication." However, since the deepest problems of formal logic concern the foundations of set theory, these applications will not be discussed here.

Lattice-ordered groups. The first study of lattice-ordered groups was made in 1897 by Dedekind, in analyzing the properties of the rational number system R. Letting $a \leqq b$ signify that $ma = b$ for some integer m (that is, that b is a multiple of a), the nonzero rational numbers form a lattice in which $a \cap b$ and $a \cup b$ again mean the lcm and gcd of a and b. Also, with respect to ordinary multiplication, the nonzero rationals form a (commutative) group. Finally, in R, implications (4*a*) hold; that is, every group transla-

$$a \leqq b \text{ implies } ac \leqq bc \tag{4a}$$

$$f(x) \leqq g(x) \text{ implies } f(x) + c(x) \leqq g(x) + c(x) \tag{4b}$$

$$f \geqq g \text{ and } \lambda \geqq 0 \text{ imply } \lambda f \geqq \lambda g \tag{4c}$$

tion is a lattice automorphism. Any lattice L which is also a group, and which shares this property, is called a lattice-ordered group, or *l*-group.

Though many noncommutative *l*-groups exist, it is a striking fact that every complete *l*-group is necessarily commutative.

Lattice-ordered groups arise in function theory, as well as in number theory. If $f \leqq g$ means that $f(x) \leqq g(x)$ for all (or almost all) x, then most function spaces of real functions form lattices. They are also commutative groups under addition. Moreover, implication (4*b*) holds for any $c(x)$; this is simply (4*a*) in additive notation. Hence most (real) function spaces are *l*-groups. In addition, they are vector spaces, in which implication (4*c*) holds. The *l*-groups with these additional properties are vector lattices. Although E. H. Moore and F. Riesz had discussed related ideas earlier, the first systematic analysis of vector lattices as such was made in 1937 by L. V. Kantorovitch.

The application of vector lattice concepts to function theory is still not very far advanced. Using the intrinsic lattice topologies defined earlier, and others related to them, one can avoid the necessity of introducing a distance function, in many function spaces. Thus, the notion of metric boundedness is equivalent to order boundedness for functionals on any Banach lattice, and metric convergence is equivalent to relative uniform star-convergence, which is a purely lattice-theoretic concept.

An additive *l*-group which is also a ring is called a lattice-ordered ring or *l*-ring when its multiplication satisfies the partial analog of implication (4*b*) shown

in formula (5). Such rings have been studied sys-

$$\text{If} \qquad f \geqq 0 \text{ and } g \geqq 0, \text{ then } fg \geq 0 \qquad (5)$$

tematically only since 1955; a typical theorem about them is the following: An *l*-ring is a product of simply ordered rings if and only if it satisfies implication (6).

$$a \cap b = 0 \text{ and } c \geqq 0 \text{ imply } ca \cap b = ac \cap b = 0 \qquad (6)$$

Further applications. It is clear that the concepts of vector lattice and of *l*-ring are essential in various physical applications. This was first apparent in connection with the ergodic theorem, as proved in 1931 by Birkhoff and von Neumann, for the deterministic processes of classical mechanics. A generalization of this theorem to stochastic processes, whose natural formulation is based on the concept of a vector lattice, was proved in 1939–1941 by Shizuo Kakutani and Kosaku Yosida.

A second application is to the theory of Reynolds operators, or averaging operators, arising in turbulent fluid motions. The essential connection with the order relation is simply the obvious principle that any average of nonnegative quantities is nonnegative. Using this principle and the theory of *l*-rings, one can decompose (subdirectly) any vector-averaging operator into scalar components.

A third application is to the concept of criticality in nuclear reactor theory. Neutron chain reactions involve the birth (through fission), migration, and death (through absorption) of neutrons. The laws governing the evolution of the statistical distributions of neutrons (as functions of position x, velocity υ, and time t) evidently carry nonnegative distributions into nonnegative distributions. To deduce the mathematical principle that the neutron distribution must satisfy the asymptotic relation $N(x,\upsilon,t) \sim e^{\lambda t}N_0(x,\upsilon)$, λ^{-1} being called the reactor "period," it is again most convenient to reformulate the problem in lattice-theoretic language. One can then apply results of Oskar Perron, G. Frobenius, and R. Jentzsch on positive linear operators to prove the desired result. *See* LOGIC; RING THEORY. Garrett Birkhoff

Bibliography. M. Anderson and T. Feil (eds.), *Lattice-Ordered Groups: An Introduction*, 1988; G. Birkhoff, *Lattice Theory*, 3d ed., 1967, reprint 1984; S. D. Cower (ed.), *Universal Algebra and Lattice Theory*, 1985; B. A. Davey and H. A. Priestley, *Introduction to Lattices and Order*, 2d ed., 2001; R. S. Freese and O. C. Garcia (eds.), *Universal Algebra and Lattice Theory*, 1983; G. Gierz et al., *A Compendium of Continuous Lattices*, 1980; G. Gratzer, *General Lattice Theory*, 2d ed., 1998; A. Hoffman, *Continuous Lattices and Their Applications*, 1985; S. Mac Lane and G. Birkhoff, *Algebra*, 3d ed., 1987.

Lattice vibrations

The oscillations of atoms in a solid about their equilibrium positions. In a crystal, these positions form a regular lattice. Because the atoms are bound not to their average positions but to the neighboring atoms, vibrations of neighbors are not independent of each other. In a regular lattice with harmonic forces between atoms, the normal modes of vibrations are lattice waves. These are progressive waves, and at low frequencies they are the elastic waves in the corresponding anisotropic continuum. The spectrum of lattice waves ranges from these low frequencies to frequencies of the order of 10^{13} Hz, and sometimes even higher. The wavelengths at these highest frequencies are of the order of interatomic spacings. *See* CRYSTAL STRUCTURE; VIBRATION; WAVE MOTION.

At room temperature and above, most of the thermal energy resides in the waves of highest frequency. Because of the short wavelength, the motion of neighboring atoms is essentially uncorrelated, so that for many purposes the vibrations can be regarded as those of independently vibrating atoms, each moving about its average position in three dimensions with average vibrational energy of $3kT$, where k is the Boltzmann constant and T the absolute temperature. The wave character of the vibrations is needed, however, to describe heat transport by lattice waves. Also, lattice vibrations interact with free electrons in a conducting solid and give rise to electrical resistance. The temperature variation at low temperatures provides evidence that this interaction is with waves. *See* ELECTRICAL RESISTIVITY.

Lattice waves. In a discrete lattice the elastic equations of a continuum are replaced by a set of linear equations in the displacements $\mathbf{u}(\mathbf{x})$ of atoms at position $\mathbf{x}$ and their accelerations. In a uniform continuum the displacements can be expressed as a superposition of progressive waves, which are defined by their wave vectors, angular frequencies, and polarizations. There are three mutually perpendicular directions of polarization, in general not simply related to the direction of the wave vector. Also, the dependence of the frequency of the wave on the wave vector is linear for a given direction and polarization; the wave frequency is equal to the product of the wave vector and the phase velocity of the wave, which is determined by the elastic moduli and the density of the continuum, and depends on the direction of the wave vector but not its magnitude. *See* PHASE VELOCITY.

In a discrete lattice with perfect periodicity, the normal modes are still progressive waves, and at low frequency they are the same as in the corresponding anisotropic continuum. However, the displacements of atoms are defined only at discrete lattice points, and when the wavelength is no longer large compared to the lattice spacing, the dependence of the wave frequency on the wave vector for a given direction and polarization is no longer linear. This is because the square of the wave frequency can be expressed as a sum of contributions from the force constants of linkages between an atom and its different neighbors, and each contribution contains a factor which varies as the square of the wave vector only when the wave vector is small. There are still three mutually perpendicular polarization directions for

every value of the wave vector, but these directions now rotate as the magnitude of the wave vector is changed. This adds complexity to the problem of calculating the frequencies, and numerical methods are required.

The conditions that the lattice be stable and the undisplaced atoms be in positions of equilibrium require that the square of the wave frequency be real and positive for every mode. This implies that for every mode there is a mode of equal frequency and opposite velocity, even in lattices without inversion symmetry. If this were not so, the second law of thermodynamics could be circumvented. *See* THERMODYNAMIC PRINCIPLES.

The dispersion relation, that is, the dependence of the frequency of the wave on the wave vector, determines the density of lattice modes per unit volume and frequency interval.

Zone structure. Like a stroboscopic effect, but in space rather than time, the addition of a reciprocal lattice vector to the wave vector of a particular wave yields the same displacement pattern of the discrete lattice as that of the original wave. Thus, all the lattice waves can be identified with wave vectors in a fundamental zone, bounded by the perpendicular bisector planes of the lowest reciprocal lattice vectors. This fundamental zone contains as many wave-vector values as there are unit cells in the crystal. All real crystals contain at least two atoms per unit cell, and some more. Correspondingly there are, for each wave vector, not only the three polarization modes of the acoustic branches, for which the frequency goes to zero as the wave vector goes to zero (that is, as the wavelength becomes very long), but also the optical branches, which have nonvanishing frequencies everywhere in wave-vector space. When the wave vector equals 0, they describe the relative motion of atoms within a unit cell. In polar crystals they interact strongly with radiation, which accounts for their name. *See* BRILLOUIN ZONE; CRYSTAL ABSORPTION SPECTRA; CRYSTALLOGRAPHY.

Localized vibrations. Dispersion curves of optical modes are frequently flat, that is, the frequency does not vary much over the zone. It thus appears, particularly in Raman spectra, that these modes of vibration have a single frequency, only slightly broadened by interactions with neighboring cells, as if the corresponding vibrations were localized within each cell. In principle, all modes of a harmonic lattice should be wavelike. However, in real crystals there are interactions between modes due to anharmonicities and defects; since these processes can be strong at high frequencies, a model of localized vibrations of atoms relative to each other in a unit cell may sometimes provide a better model. The acoustic modes are always wavelike in crystals. In amorphous solids and highly disordered crystals, they are wavelike at low frequencies but may be localized at the highest frequencies.

Quantization and phonons. A normal mode is like a harmonic oscillator, and according to quantum mechanics its energy is quantized in integral units of hf, where h is the Planck constant. These quanta are termed phonons in progressive lattice or elastic waves. In addition, each mode has a minimum or zero-point energy, $hf/2$; the sum over all normal modes of the zero-point energy is part of the formation energy of the crystal, and not regarded as thermal energy. *See* HARMONIC OSCILLATOR; NONRELATIVISTIC QUANTUM THEORY; PHONON; QUANTUM MECHANICS.

Neutron diffraction. Neutrons are particles, and can also be considered as waves. In a crystal, a neutron can undergo Bragg diffraction, that is, elastic scattering, changing its wave vector by a reciprocal lattice vector, **b**. In addition, there are weaker inelastic scattering processes, involving absorption or emission of a phonon of frequency f and the wave vector, such that the energy of the neutron increases or decreases by hf and the wave vector of the neutron increases or decreases by the sum of the phonon wave vector and a reciprocal lattice vector. Velocity selection of neutrons scattered in a fixed direction allows their energy and momentum to be determined. Hence, it is possible to obtain the frequency of an interacting phonon of a given wave vector. Through a series of such measurements, the dispersion relation can then be obtained over the entire zone, and the forces between atoms and the density of modes, $g(f)$, can be determined. *See* NEUTRON DIFFRACTION; SLOW NEUTRON SPECTROSCOPY.

Heat transport and wave interactions. Each mode contains thermal energy, and carries a heat current equal to the product of its average energy in thermal equilibrium and the group velocity of the wave. A temperature gradient perturbs the thermal equilibrium, increasing the number of phonons flowing down the temperature gradient. Various interaction processes tend to restore equilibrium, resulting in a steady state with a heat current proportional to the temperature gradient. The constant of proportionality, called the thermal conductivity, involves the contribution to the specific heat per unit volume from waves at various frequencies, the phonon mean free paths or wave attenuation lengths, and the phonon velocities. This relationship applies to isotropic solids and can be generalized for anisotropy. *See* GROUP VELOCITY.

Processes which limit the phonon mean free path are anharmonic interactions, scattering by lattice defects, and scattering by external and internal boundaries. The most important anharmonic interactions are three-phonon processes, in which a phonon is removed from a wave and a phonon is created in each of two other waves, or vice versa. Each such process must satisfy energy conservation as well as a wave-vector interference condition, so that, in the first case, for example, the frequency of the absorbed phonon is the sum of the frequencies of the two created phonons, and the wave vector of the absorbed phonons is the sum of the wave vectors of the two created phonons and a reciprocal lattice vector (**b**), which may be 0. Three-phonon processes cause the mean free path to decrease with increasing temperature.

Scattering of lattice waves by defects increases with increasing frequency (f); its variation depends

on the nature of the defect. Scattering by external and internal boundaries is almost independent of frequency, thus dominating at low frequencies and hence at low temperatures. A study of the thermal conductivity of nonmetallic crystals as function of temperature yields information about the defects present, and about the anharmonic nature of the interatomic forces in the crystal lattice. *See* CRYSTAL DEFECTS; THERMAL CONDUCTION IN SOLIDS.

Paul G. Klemens

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; R. Berman, *Thermal Conduction in Solids*, 1976; H. Bilz and W. Kress, *Phonon Dispersion Relations in Insulators*, 1979; G. Grimvall, *Thermophysical Properties of Materials*, 1986, revised 1999; C. Kittel, *Introduction to Solid State Physics*, 8th ed., 2005.

Launch complex

The composite of facilities and support equipment needed to assemble, check out, and launch a rocket-propelled vehicle. The term usually is applied to the facilities and equipment required to launch larger vehicles for which a substantial amount of prelaunch preparation is needed. Small operational rockets may require similar but highly simplified resources on a much smaller scale. For these, the term launcher is usually used. *See* SPACE FLIGHT.

Prelaunch processing. A rocket vehicle consists of one or more stages and a payload. These elements usually are manufactured at different locations and shipped separately to the launch site. The assembly process consists of properly mating these elements in the launch configuration, assuring that all mechanical and electrical interconnections are properly made. Components that have been shipped separately from the main vehicle elements also are installed. These include items requiring special handling or safety precautions, such as batteries and ordnance devices. *See* GUIDED MISSILE; ROCKET STAGING; SPACECRAFT STRUCTURE.

The checkout process consists of detailed testing of all elements of the vehicle to assure they are functioning properly and ready for launch. Checkout of individual components or subsystems may begin prior to or during assembly when easy access for repair or replacement is available. After completion of assembly, the vehicle is given a detailed overall test, usually called combined systems test or integrated systems test, to verify launch readiness.

The launch process includes the final countdown and lift-off. During countdown, propellants are loaded, all vehicle systems are activated and given a final readiness check, ordnance devices are armed, and first-stage motors are ignited.

The term facilities applies to the larger permanent and usually fixed structures on the complex.

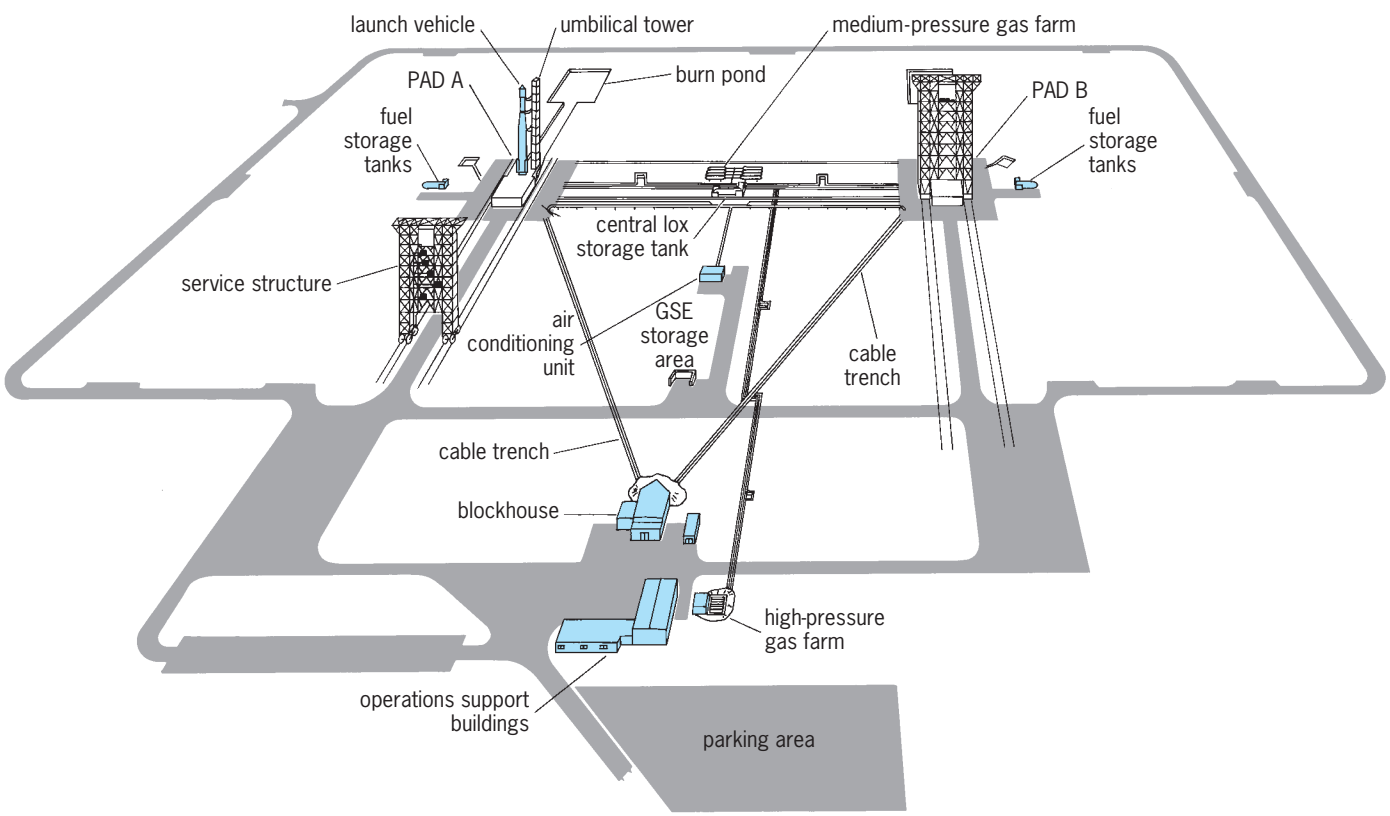

Fig. 1. Overall perspective of a typical launch complex in which one blockhouse serves two launch pads.

The support equipment, often called ground support equipment (GSE), includes all ground equipment necessary to transport, service, test, control, and launch the vehicle. In this grouping are handling equipment, transporters, the service tower and erecting mechanisms, assembly and launch test equipment and consoles, weighing mechanisms, umbilical accessories, control and instrumentation cables, and other external equipment tailored to the missile and required for its preparation and operation.

Launch-complex elements. A typical launch complex (**Fig. 1**) for a vehicle of the Thor-Delta class has two launch pads serviced by a single blockhouse and by common support equipment, providing some economy. The spacing between pads, blockhouse, and other elements of the complex is based on safety considerations. The overall size of a complex therefore depends primarily on the quantities and types of propellants used.

Pad. The launch pad itself is usually a massive concrete structure designed to withstand the heat and pressure of the rocket exhaust. The pad provides physical support for the vehicle prior to launch, but it also has a number of other functions. A propellant-transfer room contains equipment for remote control of propellant loading. An environmental control system supplies conditioned air to the vehicle to maintain normal operating temperatures. A terminal-connection room provides for data and remote-control links between the vehicle and the launch control center or blockhouse.

A flame deflector, located underneath the vehicle, diverts the rocket exhaust in a direction that does not cause harm. A high-pressure, high-volume water supply may be used to cool the flame deflector and otherwise minimize damage to the pad. The water supply also provides fire protection.

Many pads have a hold-down-and-release mechanism that holds the vehicle in place during ignition and releases it when full thrust is built up (**Fig. 2**). The hold-down feature also can be used for test firings of the rocket motors.

The weighing system is a particularly important part of on-stand equipment; it can be used for weighing the missile, determining the amount of propellants loaded, measuring the thrust in static tests, and establishing acceptable thrust performance prior to launching of a missile. Weighing mechanisms are generally load-cell strain gages.

Umbilical connections. During the checkout and countdown phases until just prior to lift-off, the vehicle is serviced from the ground by electrical and mechanical connections. The electrical connections provide electrical power, control signals for remote operation, and data links. The mechanical connections provide for propellant loading, high-pressure gas transfer, and air conditioning. These connections usually are supported by an umbilical tower located on the pad adjacent to the vehicle. Attachment to the vehicle is made by quick-disconnect devices that are pulled away just after rocket motor ignition but prior to lift-off.

Fig. 2. Launch pad with *Soyuz* spacecraft and rocket. Gantry and hold-down mechanisms fold back at proper moments in launch sequence. (*Tass*, *from Sovfoto*)

Service structure. During assembly and checkout, technicians require access to all parts of the vehicle. This is provided by a service structure, or gantry, which can be moved up to the vehicle when needed and pulled away prior to launch. The service

Fig. 3. Minuteman ICBM on the launch pad with its service structure in place. (*Boeing Aircraft*)

Fig. 4. Upper part of a Titan 3 service structure showing the enclosure which provides shelter for the payload. (*TRW Systems*)

structure has movable platforms to provide access to all levels of the vehicle (**Fig. 3**). It may also have means for partially or totally enclosing the vehicle to provide a clean, air-conditioned environment. **Figure 4** shows an environmental enclosure near the top of a Titan 3 service structure intended for sheltering only the payload. **Figure 5** shows the inside of this structure with the hole through which the vehicle will project covered with a safety net.

When the vehicle is assembled on the launch pad, the service structure also provides such necessary mechanical equipment as cranes and hoists, as well as electrical power, gas supplies, and other utilities needed for the process.

Fig. 5. Inside view of the structure of the enclosure shown in Fig. 4. Movable platforms at upper left are temporarily folded away. (*TRW Systems*)

Fueling systems. The launch complex for liquid-fueled vehicles has fuel and oxidizer storage and loading systems. Cryogenic (low-temperature) propellants such as liquid oxygen and liquid hydrogen contribute greatly to the complexity of a launch complex. The cryogens must be stored close enough to the point of use so they will not vaporize in the transfer lines but far enough back from the launch point so that the takeoff blast will not endanger the storage tanks. *See* PROPELLANT; ROCKET PROPULSION.

Extreme cleanliness is necessary in handling liquid oxygen, the most common of oxidizers. Tanks and lines must be thoroughly clean because, under proper conditions, hydrocarbons may combine with an oxidizer to cause a fire explosion. Filters are employed in the system to trap particles which may have been accidentally introduced. Liquid oxygen has a normal temperature of -297°F (-183°C); to minimize losses, the cryogenic tanks are double-jacketed, with a vacuum in the annular space. The entire system is usually made of stainless steel for cleanliness and strength at low temperatures.

Liquid hydrogen, used as a fuel in some vehicles, is even more difficult to handle. It not only has a much lower temperature (-423.2°F or -252.9°C), but hydrogen systems are highly susceptible to leaks. When mixed with air, hydrogen is extremely explosive.

Noncryogenic propellants, such as hydrazine and fuming nitric acid, are used in some vehicles. These are sometimes called storable propellants because they do not boil off at normal ambient temperatures as do cryogens. However, they pose other handling problems because they ignite on contact with each other and are extremely toxic, and fuming nitric acid is highly corrosive.

Fixed fuel storage and pumping areas are ordinarily located on the opposite side of the launch stand from the oxidizer area for safety. For many fueling facilities, the more dangerous and corrosive propellants are handled and transferred by means of specially designed trailers. Even liquid oxygen may be transported in field trailers with their own pumping and pressurization equipment and introduced into the launch vehicle through a loading manifold. Both fuel and oxidizer systems are usually operated remotely from the blockhouse.

A method of dumping propellants is usually provided. Liquid oxygen may be dumped into a concrete evaporation pond, sometimes called a burn pond, on the launch complex. Fuel is rarely dumped on the ground; it is either fed back into its storage tank or piped to a fuel-holding pond where it can be disposed of later. More elaborate closed storage is required for highly toxic or otherwise dangerous propellants.

Control center. The launch operation is controlled from a launch control center. The older control centers, used for uncrewed space missions, are often contained in a blockhouse on the launch complex, a few hundred meters from the launch pad, as in Fig. 1.

Fig. 6. External view of typical blockhouse for uncrewed missions. Massive concrete construction protects workers during launch.

The blockhouse itself is a massive concrete structure (**Fig. 6**) designed to protect personnel from a possible inadvertent impact of the vehicle on the blockhouse with consequent explosion or fire. Thickness of the walls, sometimes 5–7 ft (1.5–2 m) of concrete, is designed to withstand the force of the highest-order explosion that the vehicle propellants could produce. Other precautions, such as baffled underground escape hatches and tunnels and remote air intakes for air conditioning, may be provided.

The launch control center contains both control and monitoring systems. The control systems usually are consoles or panels that enable the operator to remotely operate and control specific functions on the launch pad and in the vehicle. For example, the propellant-transfer console enables the operator to load propellants from the storage tanks to the vehicle. Within the propellant system and the vehicle are sensors that measure flow rates, temperatures, pressures, and other parameters critical to the propellant-loading process. These measurements are transmitted via cable back to the control console, where they are displayed. The operator controls the process and also monitors it to assure satisfactory performance.

The launch control center used for crewed space shuttle missions is located 5 mi (8 km) from the launch pads, far outside the blast area, and therefore does not require a blockhouse. This facility contains four firing rooms, each equipped with consoles which contain controls and displays required for checkout and launch (**Fig. 7**). All four firing rooms are operated by NASA, but two are used to support dedicated Department of Defense missions. The many controls and instruments on the consoles and panels connect to a terminal room on the floor below. Connections are made there to the cables which lead to a similar terminal room at the pad, where in turn connections are made to the vehicle. Closed-circuit television allows direct viewing of many critical operations.

Because many people participate in a launch operation, the communications system that links them together is important to operational success. An operations intercom system may have a dozen or more channels, each channel having a specific assignment. For example, channel 1 may be assigned to stage 1 propulsion, channel 2 to guidance, channel 3 to instrumentation, channel 4 to pad safety, and so on. This scheme of assignment allows each subsystem crew to communicate with each other. The subsystem crew chief also has a link with the system test conductor, for example, with the stage 1 test conductor. The stage 1 test conductor has an additional channel to the chief test conductor, who has overall launch responsibility. The chief test conductor also has direct links to important functions outside the launch complex, such as range control, range safety, and weather forecasting.

Checkout computers. The use of digital computers is important in checkout and launch operations. A typical comprehensive test or launch countdown may require many hundreds of test sequences, each switched on in proper sequence. In addition, many hundreds or even thousands of data measurements

Fig. 7. Typical firing room in the launch control center used for crewed space shuttle missions. (*NASA*)

Fig. 8. Space shuttle Vehicle Assembly Building on Complex 39 at the Kennedy Space Center. The launch control center is at lower left. The space shuttle is on the mobile launch platform in front of the building. (*NASA*)

must be made and evaluated to verify the test results. If this were done manually by human operators, the total time required would be unreasonably long. Computers can control test sequences and data readouts many times faster and with less chance for error. Computers can also promptly evaluate whether a measurement is within a specified tolerance and, if an out-of-tolerance measurement occurs, notify the test conductor. Without the speed and accuracy that computers provide, the checkout and launch of a complex space vehicle, such as a crewed space shuttle, would be virtually impossible. Many rocket vehicle components are deliberately designed for a short life but high performance to achieve weight savings. Such components could wear out faster than they could be checked out in a manual operation. *See* COMPUTER; DIGITAL COMPUTER.

Mobile launcher. Many launch complexes provide for assembly of the rocket vehicle on the launch pad, using the service structure as the assembly facility. Large, complex vehicles such as a space shuttle may require an assembly time of many weeks or months. For them, the launch site is inadequate because the service structure does not usually provide adequate shelter from weather, salt-air corrosion, and other environmental factors; for safety, service shops, stockrooms, engineering offices, and other needed services cannot be located near the launch pad; and the relatively long vehicle preparation time makes the pad unavailable for launchings, thus limiting launch schedules.

To overcome these problems, the mobile-launch concept has been used as the basis for launch-complex design on several programs (Titan 3, Apollo-Saturn 5, and the space shuttle are notable examples). In this procedure, the vehicle is assembled and checked out in specially designed facilities that provide shelter as well as other facilities and support equipment needed for the process. Because highly volatile liquid propellants are not loaded in the assembly building, several vehicles can be safely processed at once, allowing much support equipment to be shared. Usually the vehicle is assembled directly on a transporter which, after the vehicle is assembled and checked out, carries the vehicle to the launch pad and then serves as a launch platform.

The space shuttle Vehicle Assembly Building at the Kennedy Space Center (**Fig. 8**) is 525 ft (160 m) high, 716 ft (213 m) long, and 518 ft (158 m) wide, making it one of the largest (in volume) buildings in the world. A low-bay area is used for overhaul and maintenance of the space shuttle's main engine and serves as a holding area for solid rocket forward assemblies and aft skirts. Solid rocket booster segments and assemblies are processed in adjacent facilities and transferred to the Vehicle Assembly Building for mating with the rest of the complete shuttle vehicle. A high-bay area has four bays, two on each side of the building. Two of the bays are used to store the shuttle's large external fuel tank. The other two bays are used to stack and integrate the complete space shuttle vehicle. Retractable platforms in each bay provide direct access to the vehicle at all work levels. Diagonal to the vehicle assembly building is the launch control center. Two launch pads are located about 4 mi (6 km) distant.

The platform which supports the space shuttle on its mobile launcher (**Fig. 9**) is 160 ft (49 m) long, 135 ft (41 m) wide, and 25 ft (8 m) deep. Two levels within the mobile launcher contain checkout equipment, including a checkout computer, propellant-loading accessories, electrical equipment, and hydraulic servicing units. The mobile launcher serves as the basic support structure for the vehicle during the entire assembly checkout and launch process. The structure, with vehicle aboard, is moved the 4 mi (6 km) from the Vehicle Assembly Building to the launch pad by means of a huge tractorlike device which moves under the mobile launcher, lifts it, and carries it along a specially built roadway.

Space shuttle. Space shuttle operations require two facilities not usually associated with the typical launch complex. The orbiter landing facility is a 3-mi-long (5-km), 300-ft-wide (90-m) runway used for landing operations. The orbiter processing facility

Fig. 9. Fully assembled space shuttle on the mobile launcher leaving the Vehicle Assembly Building. (*NASA*)

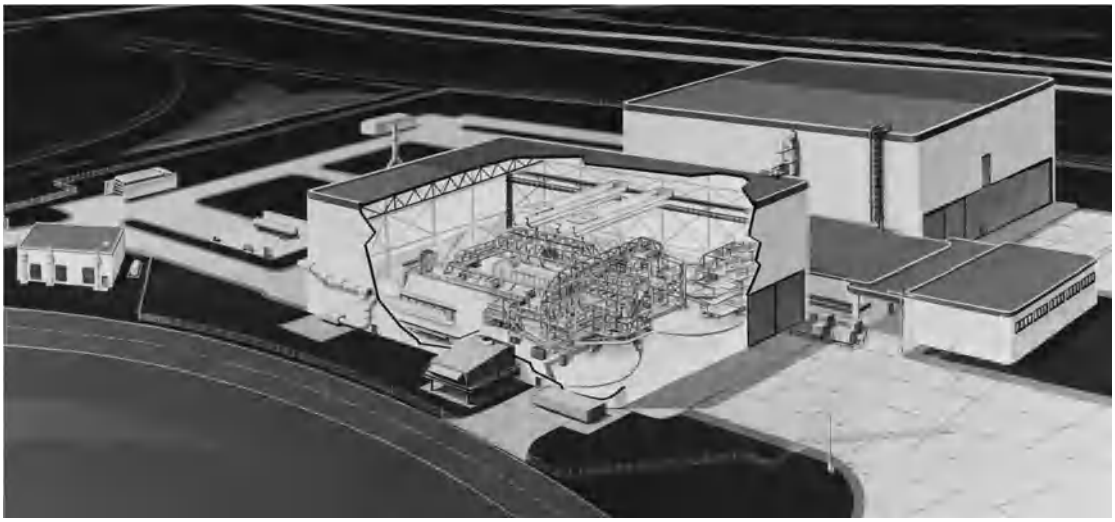

Fig. 10. Orbiter processing facility, used to refurbish orbiter vehicles after return from space and to load and unload heavy cargoes. (*NASA*)

(**Fig. 10**) is used to offload pyrotechnic devices, ordnance items, and toxic fuels, and to refurbish and prepare orbiter vehicles for relaunch after they return from space. The facility is also used to load and unload major payload elements. When vehicle refurbishment and payload operations are completed, the orbiter vehicle is transferred from the processing facility to the assembly building, where it is integrated with the other flight elements, that is, the solid rocket motors and the external fuel tank. *See* SPACE SHUTTLE.

Solid-propellant complex. Solid-propellant rocket vehicles require, in general, a launch complex similar to those just described for liquid-propellant vehicles. The most important difference is the absence of facilities and equipment needed for propellant handling and transfer. This greatly simplifies design of the complex. However, solid-propellant stages are a severe fire or explosion hazard at all times during prelaunch operations. Extreme precautions are taken to eliminate the possibility of accidentally igniting the solid propellants by static electricity, sparks from electrical equipment or machinery, or lightning. Work areas and storage areas are usually separated by revetments.

Operational systems. The launch complex for an operational ballistic missile weapon has essentially the same features as the complexes described above. An operational system is one which is deployed and ready for use by the armed forces. The major differences in the launch complexes are therefore concerned with the operational requirement to maintain a large number of missiles in a launch-ready condition at all times, to have a high probability of successful launches in a short time, and to have a high probability of striking designated targets.

Facilities and equipment for operational systems are designed to have extremely high reliability. Much of the instrumentation used in research-and-development launches is eliminated, and only that which is essential to monitoring launch readiness is retained. A single launch control center may be used for operational control of a large number of missiles, with launch sites for the individual missiles well dispersed around the complex. For weapon systems, solid propellants and storable propellants are much favored over cryogens, which are difficult to store and handle and which cannot be loaded into the vehicle until just prior to launch, greatly increasing the reaction time of the system.

Large weapon-system launch sites may be hardened; that is, they are designed to withstand all but a direct hit from a nuclear weapon in the event of an enemy attack. Hardening usually consists of locating the entire complex underground. The launch pad is a deep hole, called a silo, which contains the vehicle (**Fig. 11**). Tunnels provide limited access for

Fig. 11. Minuteman ICBM emplaced in silo.

umbilicals and at key work levels. As Fig. 11 shows, a technician can be lowered from the surface for additional access, although, on an operational system, access requirements would be minimal.

Perhaps the most unique type of operational launch complex is the Poseidon and Trident-type of submarine. These nuclear-powered vessels can launch 16 Poseidon and 24 Trident intermediate-range ballistic missiles with nuclear warheads while submerged. The relatively short, squat solid-propellant missiles are mounted vertically in injection tubes in a compartment of the submarine. At launch, compressed air ejects the missile to the surface, where the rocket motor is ignited. Complete checkout gear and targeting computers are contained in the submarine. Successful use of the submarine-launched missiles depends on the ability to obtain an accurate geodetic location of the submarine relative to its target. This can be accomplished to a high degree of accuracy anywhere in the world by means of navigation satellites. *See* NAVAL ARMAMENT; SATELLITE NAVIGATION SYSTEMS.

Commercial launch services. With the development of the Space Transportation System, anchored by the space shuttle, NASA began a gradual phaseout of its expendable launch vehicle fleet, consisting of the Atlas-Centaur and the Delta. The last of these vehicles under NASA control were launched in 1989. Under an agreement with the Department of Defense, NASA returned control of Launch Complex 17 (Delta) and Launch Complex 36 (Atlas-Centaur) on Cape Canaveral to the U.S. Air Force. The manufacturers of the Delta and the Atlas-Centaur have agreements with the Air Force for use of these complexes for commercial launch activities involving these vehicles. A similar agreement is in effect between the maker of the Titan vehicle and the Air Force for the use of Complex 40 on Cape Canaveral for commercial launch operations. NASA contracts with the commercial launch organizations for payloads that are not suitable for the space shuttle and that do not require crew-tending during launch, deployment, or orbital operations. Karl E. Kristofferson

Bibliography. *Kennedy Space Center Space Transportation System Ground Operations Plan*, K-STSM-09, Rev. 4, 1979; *Launch Site Accommodations Handbook for STS Payloads*, K-SYSM-14.1, Rev. D, 1989; *Moonport*, SP-4204, 1978; National Aeronautics and Space Administration: *Space Transportation System Facilities and Operations, Kennedy Space Center, Florida*, K-STSM-01, App. A, Rev. A, April 1984; *National Space Transportation System Reference*, vol. 1: *Systems and Facilities*, vol. 2: *Operations*, 1988.

Laurales

An order of flowering plants composed of seven eumagnoliid families of tropical tree species that are important ecologically; some are shrubs. They include in total about 2500 species. They are most closely related to Magnoliales, from which they differ in their partly inferior ovaries and their biaperturate or inaperturate pollen, and then to Winterales and Piperales. Lauraceae (the laurel or cinnamon family) are the best known and largest, but Monimiaceae and its segregates (Atherospermataceae and Siparunaceae) are also important. Nearly all species have aromatic oils, which are important spices, perfumes, and medicines; their flowers are for the most part small and often arranged in distinct whorls, but some such as those of Calycanthaceae are large and much like those of Magnoliales in that parts are arranged spirally and intergrade.

Many species are important as timbers. Cinnamon and camphor come from *Cinnamomum* species, sassafras tea was formerly made from the roots of *Sassafras albida* (now discouraged due to its suspected carcinogenic nature), and avocado comes from *Persea americana*; several genera are cultivated as ornamentals, such as *Calycanthus* (Carolina allspice) and *Chimonanthus* (wintersweet; both Calycanthaceae); and *Laurus* (bay laurel) and *Lindera* (spice bush; both Lauraceae). *See* AVOCADO; CAMPHOR TREE; EUMAGNOLIIDS; MAGNOLIALES; MONOCOTYLEDONS; PIPERALES; SASSAFRAS. Mark Chase

Lava

Molten rock material that is erupted by volcanoes through openings (volcanic vents) in the Earth's surface. Volcanic rock is formed by the cooling and solidification of erupted lava. Beneath the Earth's surface, molten rock material is called magma. All magmas and lavas consist mainly of a liquid, along with much smaller amounts of solid and gaseous matter. The liquid is molten rock that contains some dissolved gases or gas bubbles; the solids are suspended crystals of minerals or incorporated fragments of preexisting rock. Rapid cooling (quenching) of this liquid upon eruption forms a natural volcanic glass, whereas slower cooling allows more minerals to crystallize from the liquid and preexisting minerals to grow in size. The dissolved gases, a large proportion of which are lost on eruption, are mostly water vapor, together with lesser amounts of carbon, sulfur, chlorine, and fluorine gases. With very rare exception, the chemical composition of the liquid part of magmas and lavas is dominated by silicon and oxygen, which form polymers or compounds with other common rock-forming elements, such as aluminum, iron, magnesium, calcium, sodium, potassium, and titanium.

Properties. Viscosity is the principal property which determines the form of erupted lava. It is mainly dependent on chemical composition, temperature, gas content, and the amount of crystals in the magma. Liquid lava with basaltic composition (such as in Hawaii)—relatively low in silicon and aluminum and high in iron, magnesium, and calcium—has higher fluidity (lower viscosity) compared with lava of rhyolitic or dacitic composition (such as at

Mount St. Helens, Washington), with higher abundance of silicon and aluminum but lower amounts of iron, magnesium, and calcium. High temperature and gas content of the liquid lava, combined with low crystal abundance, also contribute to increased lava fluidity. Measured maximum temperatures of basaltic lava (1150–1200°C; 2100–2190°F) are higher than those for andesitic and more silicic lavas (720–850°C; 1330–1560°F). Very fluid basaltic lavas can flow great distances, tens to hundreds of kilometers, from the eruptive vents; in contrast, more silicic lavas travel much shorter distances, forming stubby flows or piling up around the vent to form lava domes. *See* VISCOSITY.

Products. Volcanic products formed by erupted lava vary greatly in size and appearance, depending on volcano type, lava composition, and eruptive style. Most lava products are either lava flows, formed during nonexplosive eruptions by cooling and hardening of flowing lava; or fragmental (pyroclastic) products, formed during explosive eruptions by the shredding apart and ejection into the air of liquid lava.

There are three commonly recognized types of lava flows: pahoehoe, aa, and block lava. Pahoehoe is the Hawaiian name given worldwide to solidified fluid lava whose surface is smooth, gently rolling, and wrinkled to a ropey appearance (see **illus.**); in contrast, aa (also a Hawaiian term) has a very rough, jumbled surface composed of jagged fragments and blocks. Block lava is made up of angular blocks, but bigger and less jagged than those in aa lava.

Fragmental volcanic products differ greatly in size and form, depending on the type of lava ejected and the explosive force of the eruption. All such materials are called pyroclastic. If the explosive force is relatively weak, the lava fragments tend to be larger in size. In contrast, during violently explosive eruptions, the ejected lava is shattered into smaller fragments. However, fragment size in any given explosive eruption can range from room-size blocks to dust-size particles; the largest and heaviest fragments are found closest to the eruptive vent. A great variety of terms have been coined by geologists to describe pyroclastic rocks and deposits, depending on the size, shape, and other characteristics of the fragments. Whether erupted explosively or nonexplosively, basaltic lava generally solidifies to form dark-colored volcanic rocks, whereas more silicic lavas (for example, andesitic, dacitic, and rhyolitic) typically form rocks much lighter in color. *See* ANDESITE; DACITE; IGNEOUS ROCKS; MAGMA; PYROCLASTIC ROCKS; RHYOLITE; TUFF; VISCOSITY; VOLCANIC GLASS; VOLCANO. Robert I. Tilling

Pahoehoe lava (in Kilauea, Hawaii) has a smooth and wavy surface formed by a thin layer of cooler lava at the surface. The surface layer is pushed into folds from the faster-moving, hot fluid lava beneath. (*Photograph by J.D. Griggs, U.S. Geological Survey, June 15, 1989*)

Bibliography. R. W. Decker and B. B. Decker, *Mountains of Fire*, Cambridge University, New York, 1991; R. W. Decker and B. B. Decker, *Volcanoes*, 3d ed., W. H. Freeman, New York, 1998; P. Francis, *Volcanoes: A Planetary Perspective*, Oxford University Press, Oxford, 1993; G. A. Macdonald, *Volcanoes*: Prentice Hall, Englewood Cliffs, NJ, 1972; H. Sigurdsson et al. (eds.), *Encyclopedia of Volcanoes*, Academic Press, San Diego, 2000.

Lawn and turf grasses

Grasses that were once almost exclusively for pasture and meadow and now furnish ground cover for lawns, sportsfields, industrial parks and roadsides. Acreage of such mowed turf in the United States is reckoned in the tens of millions; in many states the turfgrass industry ranks among the top two or three "agricultural" pursuits. The same species important for grazing are also important for lawns. Grass is well adapted to repeated defoliation (grazing or mowing) because its meristem or growing point is basal, and for the most part only foliage tip growth is sacrificed.

Modern turfgrasses. Although much adventive ("natural") grass is still maintained as lawn, the trend has been toward the breeding and selection of cultivars suited specifically for fine turf, not pasture or meadow. These modern turfgrasses, compared with their "common" antecedents, are typically more decumbent (therefore, retain more green foliage after mowing because of this low stature), are denser by virtue of abundant tillering (hence are more resistant to weed invasion and generally wear better), and are more resistant to pest attack (they are especially screened for disease tolerance). Other features of concern to the turfgrass breeder are seed yields (this correlates a bit negatively with attractiveness), rich color and good looks (fine texture is generally preferred), and ability to spread well (abundant production of rhizomes or stolons). Quickness of seed sprouting, early vigor of the seedling, reluctance to thatch, wear quality, compatibility in blends and mixtures, climatic hardiness, and many other traits are also considered.

The production and marketing of improved turfgrass cultivars is not a static endeavor. In time, pests adapted to once-unaffected cultivar evolve, so new selections are constantly sought. Fortunately, the reservoir of turfgrass breeding lines is not as likely

to be exhausted as that of improved agricultural crops, for a diversified germ-plasm bank exists in the many common turfs and the "hedgerows" of suburbia. The parade of new cultivars can be expected to continue, with older varieties dropping out (often because of uneconomical seed yields) and being replaced by still better selections. As an example, representative northern lawn grass cultivars (Lawn Institute Variety Review Board acceptances) are given below:

1. Kentucky bluegrass (*Poa pratensis*); a bit slow from seed, but strong sod, easily cared for.

Adelphi	Fylking	Plush
America	Glade	Ram I
Arboretum	Majestic	Sydsport
Birka	Merion	Touchdown
Bonnieblue	Merit	Vantage
Eclipse	Monopoly	
Emmundi	Nugget	

2. Perennial ryegrass (*Lolium perenne*): quick starting, as elegant but not so widely adapted as bluegrass; most are polycrosses of several parental bloodlines.

Blazer	Fiesta	Pennant
Citation	Manhattan	Pennfine
Derby	NK-200	Regal
Diplomat	Omega	Yorktown II

3. Fine fescue (*Festuca rubra*, in variety); good for shade, drought and low fertility, but often patchy under warm humidity.

Agram	Highlight	Banne
Banne	Koket	Waldorf
Ensylva		

4. Bentgrass (*Agrostis* spp.) and rough bluegrass (*Poa trivialis*): grows in moist shade; mostly for frequently mowed, elegant turfs such as a golf green.

Highland and Exeter colonial bent
Emerald and Prominent creeping bents
Sabre rough bluegrass

5. Tall fescue (*Festuca arundinacea*) "turf-type" polycrosses for difficult border states' "transition zone" climate.

Clemfine, Falcon, and Rebel

Distribution. The **illustration** indicates the general distribution of the major turfgrasses in the United States. The northern species are listed in the table; southern ones are primarily Bermuda grass (*Cynodon*) and manilagrass (*Zoysia*) for the upper South; centipede (*Eremochloa*), St. Augustine (*Stenotaphrum*), and bahia (*Paspalum*) for the deeper South. Tall fescue (*Festuca arundinacea*) is often planted in the "transition zone" where the South meets the North; it is coarse and hardly elegant, but survives well where finer types do not.

As one would suppose, the northern species are at their best during cooler portions of the year. They are especially favored by planting and fertilization in

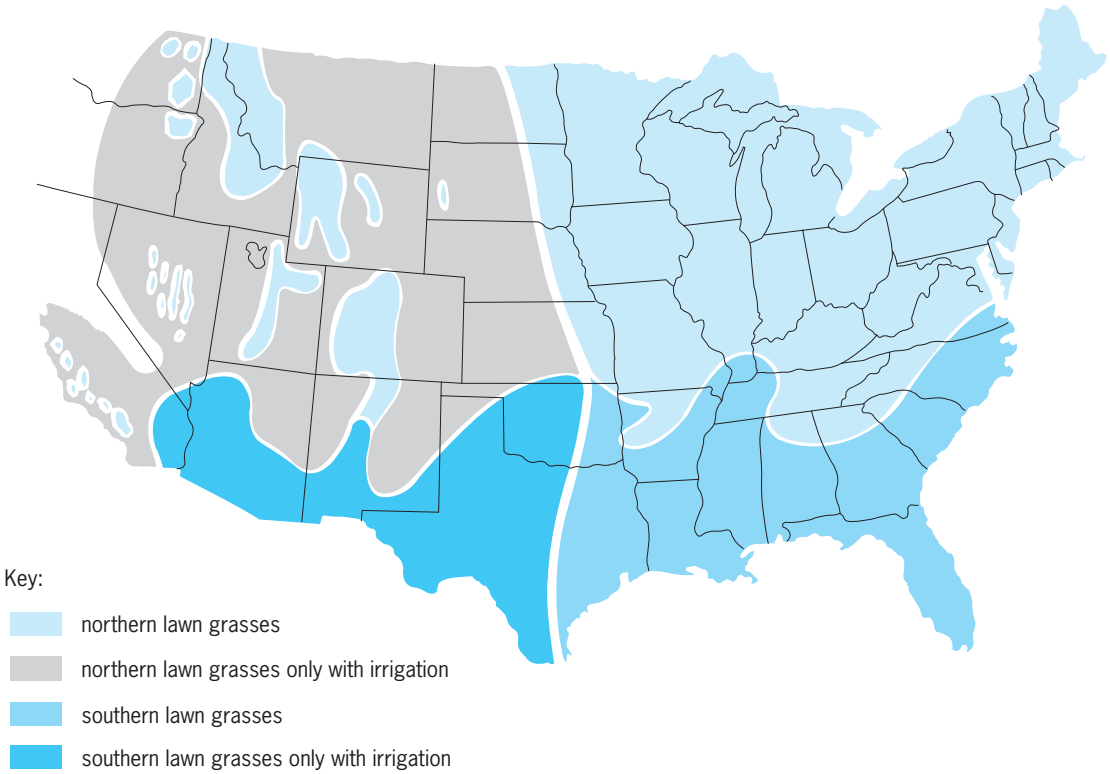

Major lawn grass zones of the United States.

autumn. Southern species are of a subtropical nature and grow best during the spring and summer.

Turfgrass management. Management practices include mowing, timely fertilization, perhaps irrigation (certainly in arid climates), and probably some measure of pest control.

Mowing will vary with the kind of grass: reel mowers are most suitable for stoloniferous species that are mowed low (bentgrasses and most southern species); the more versatile and less complicated rotaries work better for taller turf. High mowing generally favors turfgrass, since more photosynthetic foliage is retained and the grass is more competitive against weeds. Mowing should be frequent enough so that no more than half of the green foliage is sacrificed at any one clipping.

Fertilization has become costly, and self-reliant cultivars adapted to minimal maintenance are increasingly popular. Still, 2 lb of nitrogen per 1000 ft^2 (1 kg/100 m^2) annually (the experts would recommend at least twice this) should be provided for typical lawns and parceled out during seasons appropriate to growth patterns. Fescues and centipede tolerate low fertility well. Sources of nitrogen that feed out the nutrient gradually are now available in a number of forms, and are to be recommended. In most cases a complete fertilizer (containing phosphorus and potassium as well as nitrogen) would be called for, although for turf the chief and most abundant nutrient should be nitrogen. A soil test will show if phosphorus, potassium, or other cations are adequate. A simple pH test indicates the need for correcting acidity by liming (the humid East) or alkalinity with sulfur (the western plains). *See* FERTILIZER.

Irrigation is essential for survival of traditional turfgrasses in many areas, but prairie species such as buffalograss (*Buchloeë*) and grama (*Bouteloua*) often survive without irrigation. Fortunately, the engineering of lawn irrigation devices has made tremendous strides, and automatic systems are available which apply water precisely according to set schedule. Soils vary in their ability to absorb and retain water, and any irrigation system should be adjusted to their capacities. Sandy soils accept insoak rapidly but may retain only an inch or less in the top foot of soil; they should be watered briefly but frequently. Heavy soils (clays) may accept insoak only slowly, but hold up to 3 in. (8 cm) of water in the pore system of the top foot of soil; they must be watered slowly but less frequently. There is often a tendency to overwater. Actually, letting the lawn dry out occasionally almost to the point of wilting can be therapeutic, helping to forestall hydrophilic weeds (such as nutsedge, *Poa annua*) and disease. *See* IRRIGATION (AGRICULTURE).

Pest control. Weeds are perhaps the most evident pests of lawns. An excellent array of simple-to-use herbicides is available that allows one to selectively control most broadleaf weeds (dicotyledons) in grass, annual grasses in grass, but not perennial grasses in turfgrass (unless all vegetation is killed by a general herbicide and the treated spots reseeded). *See* HERBICIDE.

Insect control is more of a problem, because of the upsetting influence of insecticides on the environment. Insect damage tends to be epidemic, from sod webworms, chinch bugs, soil grubs, and so on. Treatment with a registered insecticide, according to directions, can generally rectify the trouble. *See* INSECTICIDE.

Disease inocula are ever present in a lawn, awaiting proper weather, vulnerable physiology of the grass, appropriate season, lawn stress, and so on. At least a low incidence of diseases such as leaf spot, caused by *Helminthosporium* spp., is almost inevitable in cool spring weather, whereas dollar spot, caused by *Sclerotinia homoeocarpa*, and *Fusarium*-induced diseases are generally manifest in hot weather. Innumerable other diseases can also attack turfgrasses. Most are more intense on heavily fertilized turf, but some (such as dollar spot) are less serious under generous fertilization. Turfgrass diseases are difficult for the amateur to recognize, and even the pathologist is often not sure of causality without undertaking laboratory study and reinoculation of the species.

The concept of ecological "balance" among fungi, with serious disease erupting only when conditions get out of balance, is becoming more accepted. Thus *Fusarium* may be beneficial in the normal lawn microcosm, aiding in decomposition of thatch and holding other diseases in check, although capable of damaging a lawn when a susceptible host and the right conditions coincide. Professional grounds managers who must maintain immaculate turf generally undertake preventive fungicidal sprayings at times of year when disease is likely to erupt. Even then fungi may build resistance to a particular fungicide, so that alternating a conventional contact treatment (which kills spores and mycelia on the surfaces) with the newer systemic fungicides such as benomyl is often advocated. *See* FUNGISTAT AND FUNGICIDE.

For a homeowner, it is probably too late to stop the inroads of a disease by the time that symptoms are recognized, and in most cases the disease will have to run its course or it may be terminated by changes in weather without any fungicidal treatment. In any event, a homeowner seldom has either the apparatus or the skill to apply a fungicide accurately and frequently enough so that preventive measures work. The most effective measure is to plant disease-resistant lawn grass cultivars. *See* GRASS CROPS; PLANT PATHOLOGY. Robert W. Schery

Bibliography. J. B. Beard, *Turfgrass Science and Culture*, 1973; W. H. Daniel and R. P. Freeborg, *Turf Manager's Handbook*, 1979; R. W. Schery, *Lawn Keeping*, 1976; A. J. Turgeon, *Turfgrass Management*, 5th ed., 1998.

Lawrencium

A chemical element, symbol Lr, atomic number 103. Lawrencium, named after E. O. Lawrence, is the eleventh transuranium element; it completes the

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

actinide series of elements. *See* ACTINIDE ELEMENTS; PERIODIC TABLE; TRANSURANIUM ELEMENTS.

The nuclear properties of all the isotopes of lawrencium from mass 255 to mass 260 have been established. ^{260}Lr is an alpha emitter with a half-life of 3 min and consequently is the longest-lived isotope known.

Albert Ghiorso

Bibliography. A. Ghiorso et al., New element: Lawrencium, atomic number 103, *Phys. Rev. Lett.*, 6(9):473–475, 1961; S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; G. T. Seaborg and W. D. Loveland, *The Elements Beyond Uranium*, 1990.

Lawson criterion

A necessary but not sufficient condition for the achievement of a net release of energy from nuclear fusion reactions in a fusion reactor. As originally formulated by J. D. Lawson, this condition simply stated that a minimum requirement for net energy release is that the fusion fuel charge must combust for at least enough time for the recovered fusion energy release to equal the sum of energy invested in heating that charge to fusion temperatures, plus other energy losses occurring during combustion.

The result is usually stated in the form of a minimum value of $n\tau$ that must be achieved for energy break-even, where n is the fusion fuel particle density and τ is the confinement time. Lawson considered bremsstrahlung (x-ray) energy losses in his original definition. For many fusion reactor cases, this loss is small enough to be neglected compared to the heating energy. With this simplifying assumption, the basic equation from which the Lawson criterion is derived is obtained by balancing fusion energy release against heat input to the fuel plasma. Assuming hydrogenic isotopes, deuterium and tritium at densities n_D and n_T respectively, with accompanying electrons at density n_e, all at a maxwellian temperature T, one obtains Eq. (1), where the recovered fusion

$$n_D n_T \langle \sigma v \rangle Q \tau \eta_r \geq \left[\frac{3}{2} kt(n_D + n_T + n_e) \right] \frac{1}{\eta_b} \qquad (1)$$

energy release is set equal to or greater than the energy input to heat the fuel. Here $\langle \sigma v \rangle$ is the product of reaction cross section and relative ion velocity, as averaged over the velocity distribution of the ions, Q is the fusion energy release, η_r is the efficiency of recovery of the fusion energy, η_b is the heating efficiency, and k is the Boltzmann constant.

For a fixed mixture of deuterium and tritium ions, Eq. (1) can be rearranged in the general form of Eq. (2). For a 50-50 mixture of deuterium and tritium (see **illus.**), the minimum value of $T/\langle \sigma v \rangle$ occurs at about 25 keV ion kinetic temperature (mean ion energies of about 38 keV). Depending on the assumed efficiencies of the heating and recovery processes, the lower limit values of $n\tau$ range typically between about 10^{14} and 10^{15} cm^{-3} s. These values serve as a handy index of progress toward fusion, although their achievement does not alone guarantee success. Under special circumstances (unequal ion and electron temperatures, unequal deuterium and tritium densities, and nonmaxwellian ion distributions), lower $n\tau$ values may be adequate for nominal break-even.

$$n\tau \geq F(\eta_r, \eta_b, Q) \left[\frac{T}{\langle \sigma v \rangle} \right] \qquad (2)$$

The discussion up to this point has been oriented mainly to situations in which the fusion reactor may be thought of as a driven system, that is, one in which a continuous input of energy from outside the reaction chamber is required to maintain the reaction. Provided the efficiencies of the external heating and energy recovery systems are high, a driven reactor generally would require the lowest $n\tau$ values to produce net power. An important alternative operating mode for a reactor would be an ignition mode, that is, one in which, once the initial heating of the fuel charge is accomplished, energy directly deposited in the plasma by charged reaction products will thereafter sustain the reaction. For example, in the

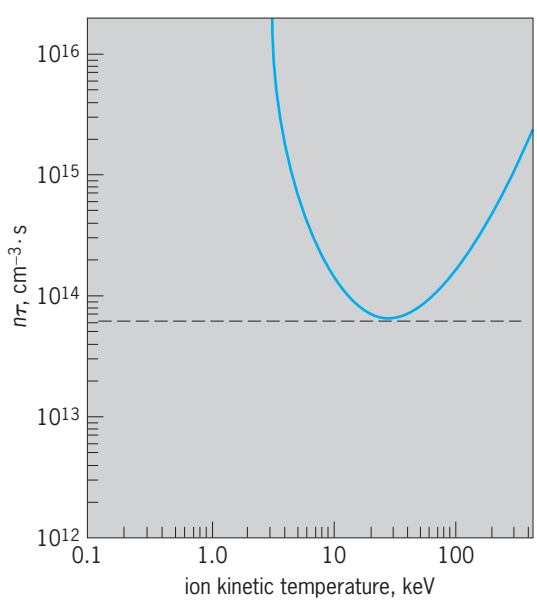

Typical plot of minimum value of $n\tau$ necessary for net release of energy versus ion kinetic temperature in a mixture containing equal amounts of deuterium and tritium. (*After J. M. Hollander, ed., Annu. Rev. Energy, 1:213–255, 1976*)

D-T reaction, approximately 20% of the total energy release is imparted to the alpha particle; in a magnetic confinement system, much of the kinetic energy carried by this charged nucleus may be directly deposited in the plasma, thereby heating it. Thus if the confinement time is adequate, the reaction may become self-sustaining without a further input of energy from external sources. Ignition, however, would generally require $n\tau$ products with a higher range of values, and is thus expected to be more difficult to achieve than the driven type of reaction. However, in all cases the Lawson criterion is to be thought of as only a rule of thumb for measuring fusion progress; detailed evaluation of all energy dissipative and energy recovery processes is required to properly evaluate any specific system. *See* NUCLEAR FUSION; NUCLEAR REACTOR; PLASMA (PHYSICS). Richard F. Post

Lawsonite

A metamorphic silicate mineral related chemically and structurally to the epidote group of minerals. Its composition is $Ca(H_2O)Al_2(OH)_2[Si_2O_7]$, and it is a sorosilicate based on the dimeric $[Si_2O_7]^{6-}$ radical, where two (SiO_4) tetrahedra link at a common vertex to form a bow-tie arrangement. The water molecule is bonded directly to the calcium ion and is tightly held. Although the calcium (Ca^{2+}) and the aluminum (Al^{3+}) cations are octahedrally coordinated by anions (and the water molecule), and the silicon cation (Si^{4+}) is tetrahedrally coordinated by the anions, the structure is not limitingly closest-packed.

Lawsonite possesses two perfect cleavages; crystals are orthorhombic prismatic to tabular, and white to pale blue to colorless; specific gravity is 3.1, and hardness is 6.5 on Mohs scale. Lawsonite has symmetry space group Ccmm, $a = 0.890$, $b = 0.576$, $c = 1.333$ nanometers, and four formula units occur in the cell. *See* CRYSTAL STRUCTURE.

Lawsonite possesses an interesting paragenesis (mineral sequence), occurring mainly along continental margins (subduction zones), and it is a good indicator mineral for high-pressure (6–12 kilobars or 600–1200 megapascals) and low-temperature (300–450°C or 570–840°F) assemblages. It is associated with the glaucophane schists (or blueschists) of the Franciscan Formation in California, the Piedmont metamorphics of Italy, and schists in New Zealand, New Caledonia, and Japan. The premetamorphic assemblage is inferred to be largely pelagic and argillaceous (siliceous) sediments derived from basalt weathering and the like. *See* BLUESCHIST; GLAUCOPHANE; SCHIST.

For metamorphism to lawsonite-glaucophane rocks, the precursor fluid phase must be relatively high in water. One reaction curve has a steep pressure slope with lower limit at about 350°C (660°F) and 6 kb (600 MPa) and an upper limit at about 450°C (840°F) and 12 kb (1200 MPa). *See* AMPHIBOLE; EPIDOTE; METAMORPHISM; SILICATE MINERALS. Paul B. Moore

Layered intrusion

In geology, an igneous rock body of large dimensions, 5–300 mi (8–480 km) across and as much as 23,000 ft (7000 m) thick, within which distinct subhorizontal stratification is apparent and may be continuous over great distances, sometimes more than 60 mi (100 km). Although conspicuous layering may be found in other rocks of syenitic to granitic composition that are richer in silica, the great layered intrusions (complexes) of the world are, in an overall sense, of tholeiitic basaltic composition. (They may be viewed as intrusive analogs to continental flood basalts.) Indeed, their basaltic composition is of paramount significance to their origin. Only basaltic melts, originating in the mantle beneath the Earth's crust, are both voluminous enough to occupy vast magma chambers and fluid enough for mineral layering to develop readily. The relatively low viscosity of basaltic melt is a consequence of its high temperature, 2100–2200°F (1150–1200°C), derived from the mantle source region, and its silica-poor, magnesium- and iron-rich (mafic) composition. *See* BASALT; EARTH; EARTH CRUST; MAGMA.

Cumulus concept. Layered mafic complexes develop upon intrusion of large volumes of basaltic magma (120–24,000 mi^3 or 500–100,000 km^3) into more or less funnel-shaped (smaller intrusions) or dish-shaped (larger complexes) chambers 3–5 mi (5–8 km) beneath the Earth's surface. Although the 1970s and 1980s produced an unusual amount of debate about the development of igneous mineral layering, it is still widely held that such layering is dominantly produced by gravitational settling of early-formed (cumulus) crystals. These crystals begin to grow as the magma cools and, on reaching a critical size, begin to sink because of their greater density relative to that of the hot silicate melt. Although the sequential order of mineral crystallization can vary depending on subtle differences in magma chemistry, a classic sequence of crystallization from basaltic magma is olivine, $(Mg,Fe)SiO_4$; orthopyroxene, $(Mg,Fe)SiO_3$; clinopyroxene, $(Ca,Mg,Fe)SiO_3$; plagioclase $(Ca,Na)(Si,Al)_4O_8$ [**Fig. 1**]. The oxide minerals that are rich in chromium, iron, and titanium, namely chromite, magnetite, and ilmenite, are also common cumulus minerals. Of these silicate and oxide minerals, only plagioclase is light in color. Hence, on various scales, the relative proportion of plagioclase between and within layers is particularly effective in creating the layered aspect of many layered intrusions (for example, **Fig. 2**).

Although layers containing only one cumulus mineral may form under special circumstances, coprecipitation of two or three cumulus minerals—for example, olivine + orthopyroxene; orthopyroxene + plagioclase; or orthopyroxene + clinopyroxene + plagioclase—is more common. Under the influence of gravity and current movements in the cooling, tabular magma chamber, the cumulus minerals accumulate on the ever-rising floor of the chamber. Later recognized as well-formed to rounded grains (**Fig. 3**), these cumulus minerals touch only at occasional

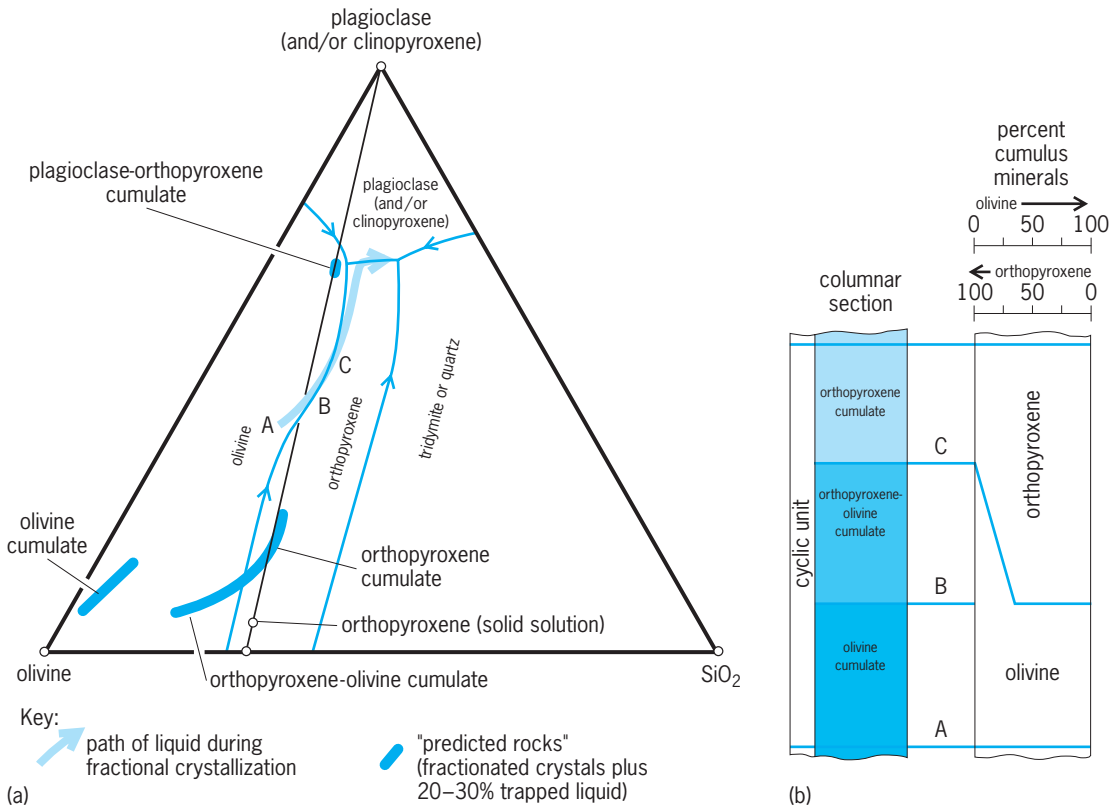

Fig. 1. Sequence of crystallization from basaltic magma. (*a*) Hypothetical phase diagram showing the derivation of the sequence olivine cumulate, orthopyroxene-olivine cumulate, orthopyroxene cumulate, as observed in cyclic units near the base of many layered mafic complexes. The triangular diagram is divided by cotectic lines (in color) into four fields in which melt compositions can be represented. For melt compositions within each field, only the single mineral for which the field is named will crystallize as the melt cools. Thus, because of withdrawal of the constituents of the crystallizing mineral, the melt changes in composition and migrates to a cotectic line. On reaching a cotectic line, the two minerals for which adjacent fields are named will crystallize together as the magma temperature falls (following the arrowheads on the cotectic lines). Plagioclase-orthopyroxene cumulate consists of plagioclase-clinopyroxene-orthopyroxene or plagioclase-clinopyroxene. (*b*) Hypothetical cyclic unit as derived from a liquid that moved from A to B to C. (*After T. N. Irvine, Crystallization sequences in the Muskox intrusion, in D. J. L. Visser and G. von Gruenewaldt, eds., Symposium on the Bushveld Igneous Complex and Other Layered Intrusions, pp. 441–476, 1969*)

points and are immersed in residual, interstitial melt.

As cooling proceeds, the interstitial melt typically solidifies as a mosaic of poorly shaped (anhedral) interlocking grains of one or more minerals that form later in the overall sequence of crystallization (Fig. 3). In other cases, particularly when pyroxene and plagioclase are cumulus phases, interstitial melt also may crystallize as overgrowths on cumulus minerals. The solid rock thus formed is known as a cumulate, a term that emphasizes its mode of origin and predominant content of cumulus minerals.

Fig. 2. A sequence of layers of near-constant thickness and aspect that show a rather abrupt transition from a base rich in pyroxene (pyroxene-plagioclase cumulate) to a top composed nearly entirely of plagioclase (plagioclase cumulate). Ice axe is approximately 3 ft (1 m) long; LaPerouse layered gabbro, Fairweather Range, Alaska.

Cumulate layering. Lithologic layering within a layered mafic intrusion is typically displayed on a variety of scales. On the broadest scale, such an intrusion may contain ultramafic cumulates rich in olivine and orthopyroxene at its base; mafic pyroxene- and plagioclase-rich cumulates at intermediate levels; and more evolved plagioclase-rich cumulates, or even granitic (granophyric) rocks, near its top. In addition to the relatively conspicuous changes in mineral occurrence and relative abundance, there is generally also a so-called cryptic evolution in mineral chemistry. Early-formed olivine and pyroxene are rich in magnesium (Mg), but become richer in iron (Fe) as successively younger cumulates are deposited; similarly, plagioclase evolves from compositions richer in calcium (Ca) to compositions richer in sodium (Na). Within this broad framework, relatively

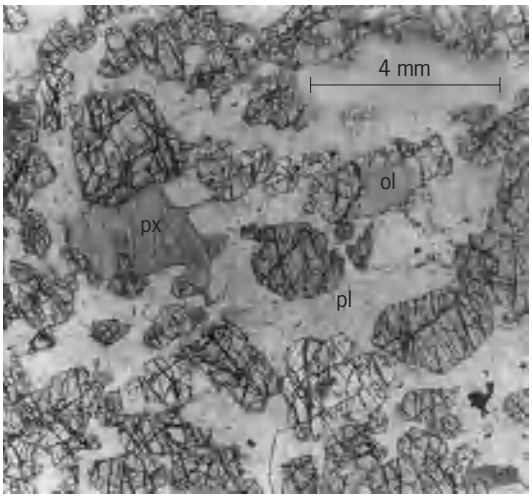

Fig. 3. Photomicrograph of an olivine cumulate from the LaPerouse gabbro. Early-crystallized olivine grains (ol; now fractured and incipiently altered) are enclosed in interstitial plagioclase (pl) and orthopyroxene (px). Mineral proportions suggest that the rock could represent crystallization of a melt with a bulk composition similar to that represented by point A in Fig. 1.

homogeneous cumulate layers range in thickness from 1000 to 2000 ft (300 to 600 m) down to less than 1 in. (2.54 cm). Commonly, layering is rhythmic or repetitious (for example, Fig. 2). The cyclic unit modeled in Fig. 1 may be repeated 15 times in the Ultramafic series of the Stillwater Complex, Montana, where each unit ranges from 80 to 600 ft (24 to 180 m) in stratigraphic thickness.

Mechanics of layering. If a tabular body of basaltic magma were to deposit a cumulate sequence without disturbance, this sequence would not be expected to display many of the observed abrupt changes in cumulus mineral assemblage and chemistry. Thus, commonly observed aberrations in lithologic layering reveal that unusual processes led to their formation. Indeed, many details of these processes have remained speculative.

At least three processes are thought to be commonplace: (1) convective overturn of magma within the chamber; (2) reinjection of new batches of similar magma from the mantle source region; and (3) injection of a chemically distinct, second magma type. In the strictest sense, the first process should produce layered rocks in which mineral proportions and compositions evolve, albeit in a stepwise manner; the second should cause repetition of the initial cumulus assemblage and mineral compositions; while the third should produce such aberrations as monomineralic cumulates, for example, of plagioclase or chromite. One or another of these processes may characterize small layered intrusions, whereas all may have operated to produce the cumulate succession of large complexes. Additional evidence of active current movements within magma chambers are features analogous with sedimentary processes, such as size grading within cumulate layers, cross-bedding, and occasional scour-and-fill structures.

Notable mafic complexes. Geologists have identified several dozen layered mafic complexes, ranging in age from 2700 million years (Stillwater Complex) to 25 million years (LaPerouse, Alaska). By virtue of its enormous size (underlying 25,000 mi^2 or 65,000 km^2) and fabulous mineral wealth, the Bushveld Complex, South Africa, is probably the world's best-known layered intrusion. Quite similar in many respects are the Dufek intrusion of Antarctica (mostly underlying more than 19,000 mi^2 or 49,000 km^2 of glacial ice; **Fig. 4**) and the Stillwater Complex. The Stillwater Complex has been tilted 60° by subsequent tectonic movements, and the 16,000-ft (5000-m) cumulate sequence is now revealed in a 30-mi-long (48-km) cross section. The Great Dyke of Zimbabwe (3.5 by 330 mi or 5.6 by 530 km), the Muskox intrusion of northern Canada, and several complexes in Western Australia are also impressively large. Relatively small complexes such as the Skaergaard intrusion of East Greenland (4.5 by 7 mi or 7 by 11 km), the Rhum intrusion of Scotland, and the Kiglapait intrusion of Labrador (17 by 20 mi or 27 by 32 km) have been studied extensively and have contributed much to the understanding of cumulate processes. The mafic intrusions at Sudbury, Canada, and Duluth, Minnesota, do not qualify as layered complexes in the strictest sense, because of a general absence of well-defined layering.

Economic considerations. Study of layered mafic complexes is of far more than academic interest because many of them host important deposits of chromium (Cr), copper (Cu), nickel (Ni), titanium (Ti), vanadium (V), and the platinum-group elements [platinum (Pt), palladium (Pd), iridium (Ir), osmium (Os), and rhodium (Rh)]. Each of these elements

Fig. 4. Repetitive darker layers of pyroxene cumulates within a thick sequence of relatively monotonous orthopyroxene-plagioclase cumulates, Dufek intrusion, Antartica. Total thickness of cumulates shown is 1750 ft (530 m). (*Art Ford, U.S. Geological Survey*)

has a relatively high initial concentration in mafic magmas, but all must be dramatically concentrated within restricted layers to be recoverable economically. The Bushveld Complex of South Africa has been a dominant source of the world's Cr, Pt, Pd, and V since the 1930s. Significant deposits of Cr in layered complexes are readily recognized because they occur as black, dense, nearly monomineralic cumulate layers of the mineral chromite, $(Cr,Fe)_3O_4$. Such chromite layers may be as thick as 6 ft (2 m) in the Bushveld Complex and have been mapped continuously over distances of tens of miles. Similar, less thick, and continuous layers have been mined from other complexes, including the lower part of the Stillwater Complex, where the chromite is, however, of significantly less value because the proportion of Cr to Fe is substantially lower.

The origin and discovery of platinum-group-element concentrations in layered mafic complexes is more problematic. Indeed, despite intimate knowledge of the Merensky Reef, principal host to the platinoids in the Bushveld Complex, years of intense prospecting preceded discovery of a remarkably similar horizon, the J-M Reef, in the Stillwater Complex. Much of the difficulty experienced in locating the Pt-Pd-enriched horizon in the Stillwater Complex is a reflection of the fact that the Merensky and J-M reefs cannot be readily understood within the framework of the cumulate theory that explains virtually every other feature of these two complexes. It appears that an influx of new magma of undetermined composition, as well as concentration of volatile constituents (such as water and chlorine), trace elements, and platinum-group elements from vast amounts of mafic magma, is essential to formation of a deposit from which Pt and Pd can be recovered profitably. In the Bushveld and Stillwater complexes, the platinum-group elements occur as alloys and in combination with sulfur (S), tellurium (Te), bismuth (Bi), selenium (Se), and other rare elements as minor components of a locally irregular but laterally persistent layer only a few feet thick. *See* IGNEOUS ROCKS.

Gerald K. Czamanske

Bibliography. G. K. Czamanske and M. L. Zientek (eds.), *The Stillwater Complex, Montana: Geology and Guide*, 1985; I. Parsons (ed.), *Origins of Igneous Layering*, 2d ed., 1990; D. J. L. Visser and G. von Gruenewaldt (eds.), *Symposium on the Bushveld Igneous Complex and Other Layered Intrusions*, 1969; L. R. Wager and G. M. Brown, *Layered Igneous Rocks*, 1967.

Layout drawing

A design drawing or graphical statement of the overall form of a component or device, which is usually prepared during the innovative stages of a design. Since it lacks detail and completeness, a layout drawing provides a faithful explanation of the device and its construction only to individuals such as designers and drafters who have been intimately involved in the conceptual stage. In a sense, the layout drawing is a running record of ideas and problems posed as the design evolves. In the layout drawing, for instance, considerations of kinematic design of a mechanical component are explored graphically in incomplete detail, showing only those aspects of the elements and their interrelationships to be considered in the design. In most cases the layout drawing ultimately becomes the primary source of information from which detail drawings and assembly drawings are prepared by other drafters under the guidance of the designer. *See* DRAFTING; ENGINEERING DRAWING.

Robert W. Mann

Lazurite

The chief mineral constituent in the ornamental stone lapis lazuli. It crystallizes in the isometric system, but well-formed crystals, usually dodecahedral, are rare. Most commonly, it is granular or in compact masses. There is imperfect dodecahedral cleavage. The hardness is 5–5.5 on Mohs scale, and the specific gravity is 2.4–2.5. There is vitreous luster and the color is a deep azure, more rarely a greenish-blue. Lazurite is a tectosilicate, the composition of which is expressed by the formula $Na_4Al_3Si_3O_{12}S$, but some S may be replaced by SO_4 or Cl. Lazurite is soluble in HCl with the evolution of hydrogen sulfide.

Lazurite is a feldspathoid but, unlike the other members of that group, is not found in igneous rocks. It occurs exclusively in crystalline limestones as a contact metamorphic mineral. Lapis lazuli is a mixture of lazurite with other silicates and calcite and usually contains disseminated pyrite. Lapis lazuli has long been valued as an ornamental material. Lazurite was formerly used as a blue pigment, ultramarine, in oil painting. Localities of occurrence are in Afghanistan; Lake Baikal, Siberia; Chile; and San Bernardino County, California. *See* FELDSPATHOID; SILICATE MINERALS.

Cornelius S. Hurlbut, Jr.

Le Chatelier's principle

A description of the response of a system in equilibrium to a change in one of the variables determining the equilibrium. Le Chatelier's principle is often stated as the following: For a system in equilibrium, a change in a variable shifts the equilibrium in the direction that opposes the change in that variable. However, this statement is false, since the shift induced by a change in a variable is sometimes in the direction that augments (rather than opposes) the change in the variable. Instead of a single general statement, the response of a system in equilibrium to a change in a variable is best described by a set of particular statements. The following statements are all valid consequences of the laws of thermodynamics. *See* CHEMICAL THERMODYNAMICS.

For any chemical reaction equilibrium or phase equilibrium, an increase in temperature at constant pressure shifts the equilibrium in the direction in which heat is absorbed by the system. For example,

the gas-phase reaction $N_2 + 3H_2 \rightarrow 2NH_3$ is exothermic (heat-producing), so an increase in temperature at constant pressure will shift the equilibrium to the left to produce more N_2 and H_2. *See* CHEMICAL EQUILIBRIUM; PHASE EQUILIBRIUM.

For any reaction equilibrium or phase equilibrium, an increase in pressure at constant temperature shifts the equilibrium in the direction in which the volume of the system decreases. For example, the gas-phase reaction $N_2 + 3H_2 \rightleftharpoons 2NH_3$ has fewer moles of gas on the right than on the left, so the products occupy a smaller volume than the reactants, and a pressure increase at constant temperature will shift the equilibrium to the right to produce more NH_3.

For a reaction equilibrium in a dilute solution, addition of a small amount of a solute species that participates in the reaction will shift the equilibrium in the direction that uses up some of the added solute. For an ideal-gas reaction equilibrium, addition at constant temperature and volume of a species that participates in the reaction shifts the equilibrium in the direction that consumes some of the added species.

For an ideal-gas reaction equilibrium, addition at constant temperature and pressure of a species that participates in the reaction might shift the equilibrium to produce more of the added species or might shift the equilibrium to use up some of the added species; the direction of the shift depends on the reaction, on which species is added, and on the initial composition of the equilibrium mixture. For example, for the gas-phase reaction $N_2 + 3H_2 \rightleftharpoons 2NH_3$, addition of a small amount of N_2 at constant temperature and pressure will shift the equilibrium to the left to produce more N_2 if the initial mole fraction of N_2 is greater than 0.5, and will shift the equilibrium to the right to consume some N_2 if the initial N_2 mole fraction is less than 0.5; addition of H_2 or NH_3 at constant temperature and pressure will always shift the equilibrium to consume some of the added species. General conditions determining which way an ideal-gas equilibrium will shift under constant-temperature-and-pressure addition of a species can be derived by using thermodynamics. Ira N. Levine

Bibliography. G. W. Castellan, *Physical Chemistry,* 3d ed., 1983; J. de Heer, Le Chatelier, scientific principle, or "sacred cow," *J. Chem. Educ.*, 35:133, 135–136, 1958; J. de Heer, The principle of Le Chatelier and Braun, *J. Chem. Educ.,* 34:375–380, 1957; I. N. Levine, *Physical Chemistry*, 4th ed., 1994; R. S. Treptow, Le Chatelier's principle, *J. Chem. Educ.*, 57:417–420, 1980.

Leaching

The removal of a soluble fraction, in the form of a solution, from an insoluble, permeable solid with which it is associated. The separation usually involves selective dissolving, with or without diffusion, but in the extreme case of simple washing it consists merely of the displacement (with some mixing) of one interstitial liquid by another with which it is miscible. The soluble constituent may be solid (as the metal leached from ore) or liquid (as the oil leached from soybeans).

Leaching is closely related to solvent extraction, in which a soluble substance is dissolved from one liquid by a second liquid immiscible with the first. Both leaching and solvent extraction are often called extraction. Because of its variety of applications and its importance to several ancient industries, leaching is known by a number of other names; solid-liquid extraction, lixiviation, percolation, infusion, washing, and decantation-settling. The liquid used to leach away the soluble material (the solute) is termed the solvent. The resulting solution is called the extract or sometimes the miscella.

The mechanism of leaching may involve simple physical solution, or dissolution made possible by chemical reaction. The rate of transportation of solvent into the mass to be leached, or of soluble fraction into the solvent, or of extract solution out of the insoluble material, or some combination of these rates may be significant. A membranous resistance may be involved. A chemical reaction rate may also affect the rate of leaching. The general complication of this simple-appearing process results in design by chiefly empirical methods. Whatever the mechanism, however, it is clear that the leaching process is favored by increased surface per unit volume of solids to be leached and by decreased radial distances that must be traversed within the solids, both of which are favored by decreased particle size. Fine solids, on the other hand, cause mechanical operating problems during leaching, slow filtration and drying rates, and possible poor quality of solid product. The basis for an optimum particle size is established by these characteristics.

Leaching processes fall into two principal classes:

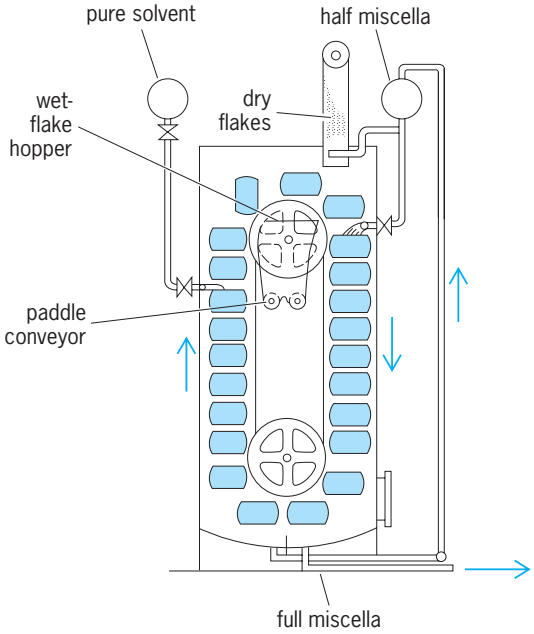

Fig. 1. Bollman extractor. (*After W. L. McCabe and J. C. Smith, Unit Operations of Chemical Engineering, 2d ed., McGraw-Hill, 1967*)

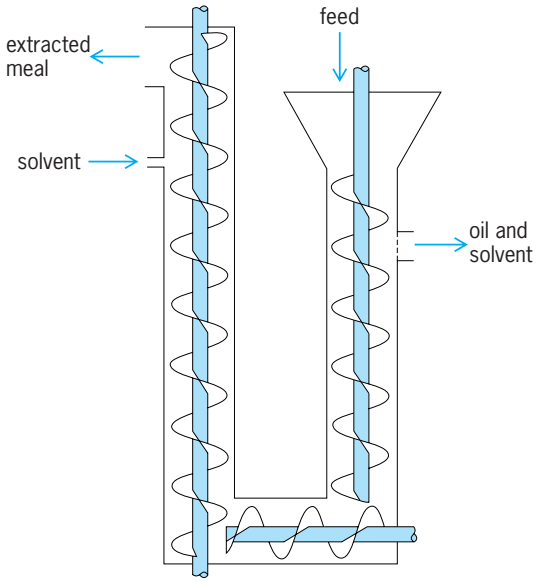

Fig. 2. Hildebrandt screw-conveyor extractor. (*After W. L.McCabe and J. C. Smith, Unit Operations of Chemical Engineering, 2d ed., McGraw-Hill, 1967*)

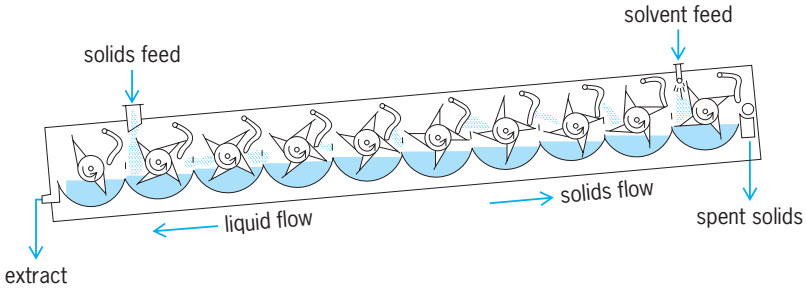

Fig. 3. Kennedy extractor. (*After R. H. Perry and C. H. Chilton, eds., Chemical Engineer's Handbook, 5th ed., McGraw-Hill, 1973*)

those in which the leaching is accomplished by percolation (seeping of solvent through a bed of solids), and those in which particulate solids are dispersed into the extracting liquid and subsequently separated from it. In either case, the operation may be a batch process or continuous. *See* EXTRACTION; SOLVENT EXTRACTION.

Percolation. In addition to being applied to ores and rock in place and by the simple technique of heap leaching, percolation is carried out in batch tanks and in several designs of continuous extractors.

The batch percolator is a large circular or rectangular tank with a false bottom. The solids to be leached are dumped into the tank to a uniform depth. They are sprayed with solvent until their solute content is reduced to an economic minimum and are then excavated. A simple example is the brewing of coffee in a percolator (repeated extraction) or a drip pot (once-through). Countercurrent flow of the solvent through a series of tanks is common, with fresh solvent entering the tank containing the most nearly exhausted material. Some leach tanks operate under pressure, to contain volatile solvents or increase the percolation rate.

Continuous percolators employ the moving-bed principle, implemented by moving baskets that carry the solids past solvent sprays, belt or screw conveyors that move them through streams or showers of solvent, or rakes that transport them along a solvent-filled trough. In a revolving-basket type like the Rotocel extractor, bottomless compartments move in a circular path over a stationary perforated annular disk. They are successively filled with solids, passed under solvent sprays connected by pumps so as to provide countercurrent flow of the extracting liquid, and emptied through a large opening in the disk. Alternating perforated-bottom extraction baskets may be arranged in a bucket-elevator configuration, as in the Bollman extractor (**Fig. 1**). On the up cycle, the partially extracted solids are percolated by fresh solvent sprayed at the top; on the down cycle, fresh solids are sprayed with the extract from the up cycle. A screw-conveyor extractor like the Hildebrandt extractor (**Fig. 2**) moves solids through a V-shaped line in a direction opposite to the flow of solution. The solids may be conveyed instead by rakes or paddles along a horizontal or inclined trough counter to the direction of solvent flow, as in the Kennedy extractor (**Fig. 3**). In the last two types, the action is predominantly percolation but involves some solids dispersal because of agitation by the conveyors. Horizontal continuous vacuum filters of the belt, tray, or table type sometimes are used as leaching equipment. *See* FILTRATION.

Dispersed-solids leaching. Equipment for leaching fine solids by dispersion and separation, a particularly useful technique for solids which disintegrate during leaching, includes batch tanks and continuous extractors.

Inasmuch as the purpose of the dispersion is usually only to permit exposure of the particles to unsaturated solvent, the agitation in a batch-stirred extractor need not be intense. Air agitation is often used. Examples are Pachuca tanks, large cylinders with conical bottoms and an axial air nozzle or air-lift tube. If mechanical agitation is employed, a slow paddle is sufficient. The Dorr agitator (**Fig. 4**) combines a rake with an air lift. In all cases, the mixture of solids and liquid is stirred until the maximum economical degree of leaching has occurred. The solids

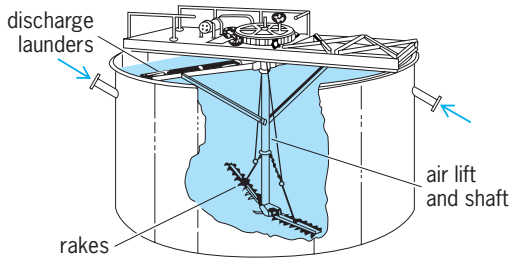

Fig. 4. Dorr agitator for batch washing of precipitates. (*After W. L. Badger and J. T. Banchero, Introduction to Chemical Engineering, McGraw-Hill, 1955*)

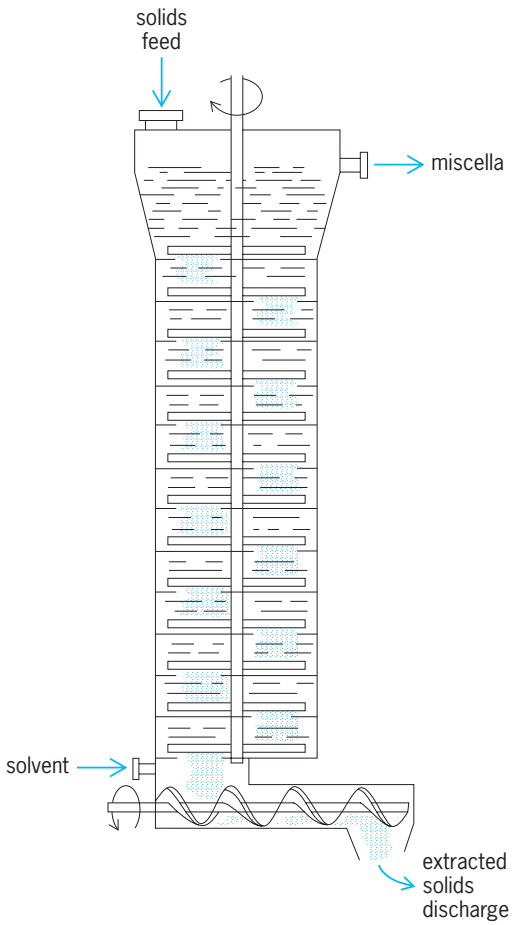

Fig. 5. Bonotto extractor. (*After R. H. Perry and C. H. Chilton, eds., Chemical Engineer's Handbook, 5th ed., McGraw-Hill, 1973*)

are then allowed to settle. The extract is decanted, and the solids, sometimes after successive treatments with fresh solvent, are removed by shoveling or flushing.

Continous dispersed-solids leaching is accomplished in gravity sedimentation tanks or in vertical plate-extractors. An example of the latter is the Bonotto extractor shown in **Fig. 5**. Staggered openings in the plates allow the solids, moved around each plate by a wiping radial blade, to cascade downward from plate to plate through upward-flowing solvent.

Gravity sedimentation thickeners can serve as effective continuous contacting and separating devices for leaching fine solids. A series of such units properly connected provides true continuous countercurrent washing (known as CCD for continuous countercurrent decantation) of the solids. *See* COUNTERCURRENT TRANSFER OPERATIONS; THICKENING.

Shelby A. Miller

Bibliography. C. J. King, *Separation Processes*, 2d ed., 1980; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., 1994; W. L. McCabe, J. C. Smith, and P. Harriot, *Unit Operations and Chemical Engineering*, 6th ed., 2000; R. H. Perry and D. Green (eds.), *Chemical Engineers' Handbook*, 7th ed., 1997; J. R. Richardson and D. G. Peacock (eds.), *Chemical Engineering*, 1979; R. E. Treybal, *Mass-Transfer Operations*: *Chemical Engineering*, 3d ed., 1980.

Lead

A chemical element, Pb, atomic number 82 and atomic weight 207.19. Lead is a heavy metal (specific gravity 11.34 at 16°C or 61°F), of bluish color, which tarnishes to dull gray. It is pliable, inelastic, easily fusible, melts at 327.4°C (621.3°F), and boils at 1740°C (3164°F). The normal chemical valences are 2 and 4. It is relatively resistant to attack by sulfuric and hydrochloric acids but dissolves slowly in nitric acid. Lead is amphoteric, forming lead salts of acids as well as metal salts of plumbic acid. Lead forms many salts, oxides, and organometallic compounds. *See* PERIODIC TABLE.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Industrially, the most important lead compounds are the lead oxides and tetraethyllead. Lead forms alloys with many metals and is generally employed in the form of alloys in most applications. Alloys formed with tin, copper, arsenic, antimony, bismuth, cadmium, and sodium are all of industrial importance. *See* LEAD ALLOYS.

Lead compounds are toxic and have resulted in poisoning of workers from misuse and overexposure. However, lead poisoning is presently rare because of the industrial application of modern hygienic and engineering controls. The greatest hazard arises from the inhalation of vapor or dust. In the case of organolead compounds, absorption through the skin may become significant. Some of the symptoms of lead poisoning are headaches, dizziness, and insomnia. In acute cases there is usually stupor, which progresses to coma and terminates in death. The medical control of employees engaged in lead usage involves precise clinical tests of lead levels in blood and urine. With such control and the proper application of engineering control, industrial lead poisoning may be entirely prevented.

Lead rarely occurs in its elemental state. The most common ore is the sulfide, galena. The other minerals of commercial importance are the carbonate,

cerussite, and the sulfate, anglesite, which are much more rare. Lead also occurs in various uranium and thorium minerals, arising directly from radioactive decay. Commercial lead ores may contain as little as 3% lead, but a lead content of about 10% is most common. The ores are concentrated to 40% or greater lead content before smelting. *See* LEAD METALLURGY.

The largest single use of lead is for the manufacture of storage batteries. Other important applications are for the manufacture of cable covering, construction, pigments, and ammunition.

Organolead compounds are being developed for applications such as catalysts for polyurethane foams, marine antifouling paint toxicants, biocidal agents against gram-positive bacteria, protection of wood against marine borers and fungal attack, preservatives for cotton against rot and mildew, molluscicidal agents, anthelmintic agents, wear-reducing agents in lubricants, and corrosion inhibitors for steel.

Because of its excellent resistance to corrosion, lead finds extensive use in construction, particularly in the chemical industry. It is resistant to attack by many acids because it forms its own protective oxide coating. Because of this advantageous characteristic, lead is used widely in the manufacture and handling of sulfuric acid.

Lead has long been used as protective shielding for x-ray machines. Because of the expanded applications of atomic energy, radiation-shielding applications of lead have become increasingly important.

Lead sheathing for telephone and television cables continues to be a sizable outlet for lead. The unique ductility of lead makes it particularly suitable for this application because it can be extruded in a continuous sheath around the internal conductors.

The use of lead in pigments has been a major outlet for lead but is decreasing in volume. White lead, $2PbCO_3 \cdot Pb(OH)_2$, is the most extensively used lead pigment. Other lead pigments of importance are basic lead sulfate and lead chromates.

A considerable variety of lead compounds, such as silicates, carbonates, and salts of organic acids, are used as heat and light stabilizers for polyvinyl chloride plastics. Lead silicates are used for the manufacture of glass and ceramic frits, which are useful in introducing lead into glass and ceramic finishes. Lead azide, $Pb(N_3)_2$ is the standard detonator for explosives. Lead arsenates are used in large quantities as insecticides for crop protection. Litharge (lead oxide) is widely employed to improve the magnetic properties of barium ferrite ceramic magnets. Also, a calcined mixture of lead zirconate and lead titanate, known as PZT, is finding increasing markets as a piezoelectric material.

Hymin Shapino; James D. Johnston

Bibliography. American Society for Metals, *Metals Handbook*, 2d ed., 1999; P. W. Atkins et al., *Inorganic Chemistry*, 4th ed., 2006; J. Casas and J. Sordo (eds.), *Lead: Chemistry, Analytical Aspects, Environmental Impact and Health Effects*, 2006; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; Z. Rappoport (ed.), *The Chemistry of Organic Germanium, Tin and Lead Compounds*, 2003.

Lead alloys

Substances formed by the addition of one or more elements, usually metals, to lead. Lead alloys may exhibit greatly improved mechanical or chemical properties as compared to pure lead. The major alloying additions to lead are antimony and tin. The solubilities of most other elements in lead are small, but even fractional weight percent additions of some of these elements, notably copper and arsenic, can alter properties appreciably.

Cable-sheathing alloys. Lead is used as a sheath over the electrical components to protect power and telephone cable from moisture. Alloys containing 1% antimony are used for telephone cable, and lead-arsenical alloys, containing 0.15% arsenic, 0.1% tin, and 0.1% bismuth, for example, are used for power cable. Aluminum and plastic cable sheathing have replaced lead alloy sheathing in many applications.

Battery-grid alloys. Lead alloy grids are used in the lead-acid storage battery (the type used in automobiles) to support the active material composing the plates. Lead grid alloys contain 6–12% antimony for strength, small amounts of tin to improve castability, and one or more other minor additions to retard dimensional change in service. No lead alloys capable of replacing the lead-antimony alloys in automobile batteries have been developed. An alloy containing 0.03% calcium for use in large stationary batteries has had success.

Chemical-resistant alloys. Lead alloys are used extensively in many applications requiring resistance to water, atmosphere, or chemical corrosion. They are noted for their resistance to attack by sulfuric acid. Alloys most commonly used contain 0.06% copper, or 1–12% antimony, where greater strength is needed. The presence of antimony lowers corrosion resistance to some degree.

Type metals. Type metals contain $2^1/_2$–12% tin and $2^1/_2$–25% antimony. Antimony increases hardness and reduces shrinkage during solidification. Tin improves fluidity and reproduction of detail. Both elements lower the melting temperature of the alloy. Common type metals melt at 460–475°F (238–246°C).

Bearing metals. Lead bearing metals (babbitt metals) contain 10–15% antimony, 5–10% tin, and for some applications, small amounts of arsenic or copper. Tin and antimony combine to form a compound which provides wear resistance. These alloys find frequent application in cast sleeve bearings, and are used extensively in freight-car journal bearings. In some cast bearing bronzes, the lead content may exceed 25%. *See* ANTIFRICTION BEARING.

Solders. A large number of lead-base solder compositions have been developed. Most contain large amounts of tin with selected minor additions to provide specific benefits, such as improved wetting characteristics. *See* SOLDERING.

Free-machining brasses, bronzes, steels. Lead is added in amounts from 1 to 25% to brasses and bronzes to improve machining characteristics. Lead remains as discrete particles in these alloys. It is also added to some construction steel products to increase machinability. Only about 0.1% is needed, but the tonnage involved is so large that this forms an important use for lead. *See* ALLOY; LEAD; LEAD METALLURGY; TIN ALLOYS. Dean N. Williams

Bibliography. American Society for Testing and Materials, *Nonferrous Metal Products*, vol. 02.04, 1986; W. Hofmann, *Lead and Lead Alloys*: *Properties and Technology*, 1970; T. J. Bertone and J. Neely, *Practical Metallurgy and Materials of Industry*, 5th ed., 1999.

Lead isotopes (geochemistry)

The study of the isotopic composition of stable and radioactive lead in geological and environmental materials to determine their ages or origins. *See* LEAD.

Stable Isotopes

Lead isotope geochemistry provides the principal method for determining the ages of old rocks and the Earth itself, as well as the sources of metals in mineral deposits and the evolution of the mantle.

Geochronology. Lead (Pb) has four stable isotopes of mass 204, 206, 207, and 208. Three are produced by the radioactive decay of uranium (U) and thorium (Th) [reactions (1)–(3), where $t_{1/2}$ is the half-life of

$$^{238}\mathrm{U}(t_{1/2} = 4.5 \times 10^9 \text{ years}) \longrightarrow {}^{206}\mathrm{Pb} + 8\alpha + 6\beta \quad (1)$$

$$^{235}\mathrm{U}(t_{1/2} = 0.71 \times 10^9 \text{ years}) \longrightarrow {}^{207}\mathrm{Pb} + 7\alpha + 4\beta \quad (2)$$

$$^{232}\mathrm{Th}(t_{1/2} = 13.9 \times 10^9 \text{ years}) \longrightarrow {}^{208}\mathrm{Pb} + 6\alpha + 4\beta \quad (3)$$

the isotope and α and β denote alpha and beta particles, respectively]. The lead produced by the decay of uranium and thorium is termed radiogenic. Since ^{204}Pb is not produced by the decay of any naturally occurring radionuclide, it can be used as a monitor of the amount of initial (nonradiogenic) lead in a system. This will include all of the ^{204}Pb and variable amounts of ^{206}Pb, ^{207}Pb, and ^{208}Pb. *See* ALPHA PARTICLES; BETA PARTICLES; RADIOACTIVITY; THORIUM; URANIUM.

Closed systems. It is possible to calculate the isotopic composition of lead at any time t in the past by calculating and deducting the amount of radiogenic lead that will have accumulated, provided a mineral or rock represents a closed system. A closed system is one in which there has been no chemical transfer of uranium, thorium, or lead in or out of the mineral or rock since it formed. All calculations for uranium-lead dating should yield the same age; this is a unique and powerful property. The ratio of radiogenic ^{207}Pb to ^{206}Pb is simply a function of age, not the U/Pb ratio. Certain minerals such as zircon, monazite, and uraninite are particularly well suited for dating because of extremely high concentrations of uranium or thorium relative to initial lead. However, the degree to which they behave as closed systems can vary. For samples having concordant U-Pb ages, data lie along a curved line called the concordia, the line defined by the daughter/parent ratios of each isotopic system that have equal ages (**Fig. 1**). There are several uranium-rich minerals that commonly yield concordant ages, the most useful being the rare-earth phosphate monazite, a common accessory mineral in crustal rocks. However, there are many minerals with appreciable radiogenic lead which have discordant ages, indicating they have not been a closed system throughout their history as a discrete phase. Zircon ($ZrSiO_4$), a common accessory mineral in many types of crustal rocks, has been used more than any other phase for U-Pb dating. However, the data are normally discordant. Data for a series of zircons from the Little Belt Mountains, Montana, lie on a well-defined straight line that intersects the concordia at two points (Fig. 1). It has been shown that phases subject to lead loss (or uranium gain) during a period of time that is short compared with the age of the phase yield daughter/parent ratios defining a straight line termed a discordia. The lower intersection of the discordia with the concordia indicates the time of the episodic bulk lead loss, while the upper intersection represents the age of the phase. Discordance in zircons is more pronounced in uranium-rich varieties and is caused by the severe damage to the lattice produced by recoiling alpha particles. *See* MONAZITE; ROCK AGE DETERMINATION; ZIRCON.

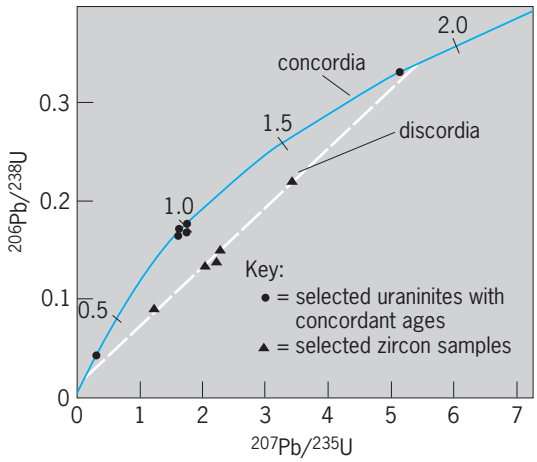

Fig. 1. Systematics of Uranium-lead (U-Pb) dating. For a closed system containing uranium but no primary lead, the ratios of ^{206}Pb/^{238}U and ^{207}Pb/^{235}U will vary with the age of the sample, as shown by the concordia line. Ages indicated by marks along each line are in units of 1 billion years. The data for uraninites are consistent with this. For systems where episodic losses of lead have occurred in the past, values may lie along a discordia, as shown for zircons from the Little Belt Mountains of Montana.

Isochron methods. Even if a rock or mineral contains appreciable initial lead, it may still be dated by using isochron methods. Since the amount of radiogenic lead relative to nonradiogenic lead is a function of the

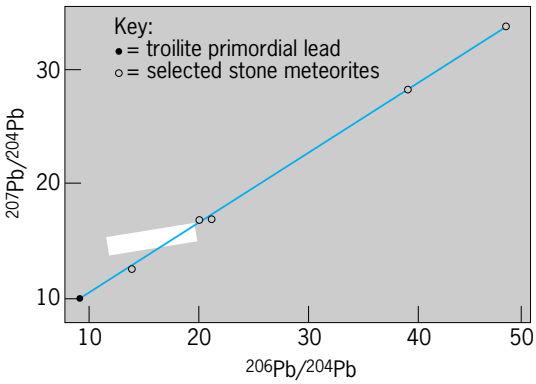

Fig. 2. Plot of $^{207}Pb/^{204}Pb$ versus $^{206}Pb/^{204}Pb$ for troilite primordial lead and selected stone meteorites. The slope of the primary isochron (Geochron) for modern lead indicates an age of 4.55 million years for these materials. The white rectangular area illustrates the range of variation in most terrestrial leads and corresponds approximately to the region detailed in Fig. 3.

U/Pb ratio and time, the slope on a plot of $^{206}Pb/^{204}Pb$ against $^{238}U/^{204}Pb$ is proportional to age. An isochron is a line on a graph defined by data for rocks of the same age with the same initial lead isotopic composition, the slope of which is proportional to the age. In practice, the $^{238}U/^{204}Pb$ ratio may well have been disturbed by recent alteration of the rock because uranium is highly mobile in near-surface environments. For this reason it is more common to combine the two uranium decay schemes and plot $^{207}Pb/^{204}Pb$ against $^{206}Pb/^{204}Pb$; the slope of an isochron on this plot is a function of age.

Isochron dating has been used to determine an age of 4.55 billion years for the Earth and the solar system by dating iron and stony meteorites (**Fig. 2**). The position of data along the isochron is a function of the U/Pb ratio. The iron meteorites are particularly important for defining the initial lead isotopic composition of the solar system since they contain negligible uranium. The meteorite isochron is commonly termed the Geochron. *See* EARTH, AGE OF; GEOCHRONOMETRY; METEORITE.

Geochemistry of Earth. By using the position of data for typical continental crustal rocks and samples of basalt that are derived as magmas from the mantle as shown on the Geochron (Fig. 2), the indication is that the silicate Earth has a U/Pb ratio of about 0.1. This is high relative to chondritic meteorites, commonly considered the best representative of primitive unprocessed preplanetary solar system material. A significant fraction of the Earth's total lead inventory could be in the metallic core. Also, lead is extremely volatile and may have been lost at the temperatures that inner solar system objects may have experienced in their accretionary history.

The Earth's mantle has been depleted by repeated melting during its 4.55-billion-year history, and the loss of such melts should leave the mantle with a low U/Pb ratio. However, close inspection reveals that the lead isotopic compositions of most mantle-derived magmas plot to the right of the Geochron (**Fig. 3**), implying a higher U/Pb ratio since the Earth formed. Originally it was thought that this discrepancy was caused by late accretion of the Earth or late core formation, either of which would displace the mantle to the right of the Geochron. However, there is independent isotopic evidence that the Earth did not accrete late, and there are theoretical reasons why the Earth's core almost certainly formed very early. A more likely explanation is that the mantle has been modified throughout its history by the subduction of ocean-floor basalt enriched in uranium and depleted in lead by low-temperature seawater alteration. The basalt lavas of some ocean islands such as St. Helena have especially radiogenic lead, thought to reflect an extreme example of such reenrichment. *See* SUBDUCTION ZONES.

Tracers. Lead isotopes can serve as tracers in the lithosphere, atmosphere, and hydrosphere. Lead isotopes are commonly used to trace the sources of constituents in continental terranes, granites, ore deposits, and pollutants. For example, the class of low-temperature hydrothermal lead-zinc (Pb-Zn) mineralization known as Mississippi Valley type ore deposits have extremely variable $^{206}Pb/^{204}Pb$ ratios in their galenas, ranging up to 100. These variations reflect the time-integrated U/Pb ratio of the source of the lead, and they can be used to identify specific geological units from which the lead was scavenged. Similarly, some granites such as those of the Isle of Skye in northwest Scotland have very unradiogenic lead, indicating that the magmas were derived by melting portions of the lower continental crust that were depleted in uranium about 3 billion years ago. *See* ORE AND MINERAL DEPOSITS.

The industrialized countries of the world use large tonnages of lead annually, about one-third of which is widely distributed in the air, water, soil, and vegetation of the environment. Isotopic composition of lead in various environmental samples has identified sources and pathways of lead pollution. Most of the lead in the atmosphere originates from the combustion of gasoline containing alkyl lead antiknock compounds. The second-largest emission source of

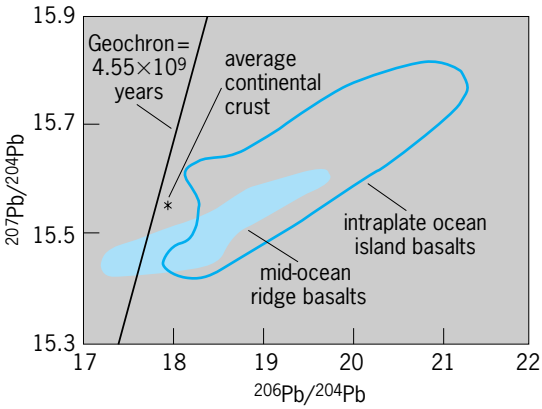

Fig. 3. Lead isotopic compositions of most ocean-floor and ocean-island basalts plot to the right of the Geochron defined by meteorite data (Fig. 2). The composition is the opposite of that predicted from the effects of depletion of the Earth's mantle by partial melting and suggests reenrichment by uranium-enriched subducted ocean floor.

atmospheric lead is coal combustion. Lead aerosols eventually fall to the ground as precipitation or as dust and accumulate in topsoil and in surface water, where they may be incorporated into terrestrial or aquatic life. Lead isotopes have been used to trace contaminant dispersion in the environment. Lead isotope studies, for example, have helped support the contention that high concentrations of lead near roadways are the result of local deposition of large aerosols from automobile exhaust. Similarly, the isotopic composition of lead in natural waters and in sediments has been useful in identifying the extent to which sources are anthropogenic. *See* AIR POLLUTION; WATER POLLUTION.

Radioactive Isotopes

While there are at least 11 known radioactive isotopes of lead, only ^{212}Pb, ^{214}Pb, and especially ^{210}Pb have been of interest geochemically. The usefulness of these isotopes stems from the unique mechanism by which they are separated from parent isotopes in the uranium or thorium decay series.

Atmosphere. Unlike their noble-gas parents, the radioactive lead isotopes as well as other daughter products have a strong affinity for atmospheric aerosols. On formation, the daughter products exist as small positive ions associated with polarized air or water molecules; they form light aggregate particles within periods of tens of seconds. Both ^{212}Pb and ^{214}Pb have been used to study the process of diffusion of ions in gases and the mechanism of attachment of small ions to aerosols. Measurement of the distribution of radon (Rn) daughter product activities (^{212}Pb and ^{214}Pb) with respect to aerosol size has been important in the development of theoretical models of ion-aerosol interactions. The short half-lives of ^{212}Pb and ^{214}Pb also make these isotopes suitable for studies of near-ground atmospheric transport processes.

While the short-lived lead isotopes disappear from the atmosphere primarily by radioactive decay, ^{210}Pb, because of its longer half-life, is removed mainly by precipitation and dry deposition. Its horizontal and vertical distributions are the result of the integrated effects of the distribution and intensity of sources, the large-scale motions of the atmosphere, and the distribution and intensity of removal processes. The inventory of ^{210}Pb in the air is about a thousand times lower than expected, given the amount of its parent ^{222}Rn, a measure of the efficiency with which aerosols are removed from the atmosphere. Numerous measurements of ^{210}Pb as well as other daughter products indicate a tropospheric aerosol residence time of under 1 week. Since the residence time of ^{210}Pb is so short and the oceans are not a significant source, the isotope sometimes can be used to distinguish between air masses originating over the continents and over the oceans. *See* AEROSOL; AIR MASS; ATMOSPHERIC CHEMISTRY; RADON.

Aquatic systems. An important use of ^{210}Pb is as a particle tracer in aquatic systems. Concentrations of ^{210}Pb in surface waves of the oceans generally show the same latitude variations as seen in air and rain. Concentrations in surface waters are roughly 20 times less than expected if there were no removal mechanisms other than radioactive decay. On entering ocean waters, ^{210}Pb is incorporated into microscopic marine organisms, particularly zooplankton, whose remains eventually sink, rapidly conveying ^{210}Pb to underlying waters. Using ^{210}Pb as a tracer has helped explain the mechanisms by which various substances, including pollutants, are removed from the oceans. *See* SEAWATER.

Most of the ^{210}Pb in deep ocean waters is produced from the decay of dissolved ^{226}Ra. The activity of ^{210}Pb is as low as 20% of the activity of radium, indicating the operation of a deep-water scavenging mechanism acting preferentially on ^{210}Pb. Measurements of ^{210}Pb concentrations in deep ocean water indicate a scavenging residence time of around 40 years. The removal process appears to involve horizontal transport of ^{210}Pb to selected areas of intense scavenging by sediments.

One of the most important uses of ^{210}Pb is for dating recent coastal marine and lake sediments. As the isotope is rapidly removed from water to underlying deposits, surface sediments often have a considerable excess of ^{210}Pb. The excess is defined as that present in addition to the amount produced by the decay of radium in the sediments. When the sedimentation rate is constant and the sediments are physically undisturbed, the excess ^{210}Pb decreases exponentially with sediment depth as a result of radioactive decay during burial. The reduction in activity at a given depth, compared with that at the surface, provides a measure of the age of the sediments at that depth. Typically, excess ^{210}Pb can be measured for up to about five half-lives or about 100 years, and it is therefore ideally suited for dating sediments that hold records of human impact on the environment. *See* RADIOISOTOPE (GEOCHEMISTRY); SEDIMENTOLOGY. Alex N. Halliday; John A. Robbins

Bibliography. R. E. Criss, *Principles of Stable Isotope Distribution*, 1999; A. P. Dickin, *Radiogenic Isotope Geology*, 1997; G. Faure and T. M. Mensing, *Isotopes: Principles and Applications*, 3d ed., 2004.

Lead metallurgy

The extraction of lead from ore, its subsequent purification and processing, and its alloying with other metals to achieve desired properties. Possibly the oldest metal known, lead was particularly attractive to alchemists, predecessors of today's metallurgists and chemists, as a substance which they believed could be transformed into gold.

There are several ores of lead; the most important is galena, PbS. Other ores are cerussite, $PbCO_3$, and anglesite, $PbSO_4$. Lead ores are often accompanied by zinc ores, and galena usually contains profitable quantities of silver and gold. *See* CERUSSITE.

Specifications and grades. There are many types and several specifications for lead. In the United States, these include ASTM B29 and Federal Specification QQ-L-171; in Germany, DIN Standard 1719;

Chemical requirements for various grades of lead

	United States ASTM Pig Lead B29				Germany Specification DIN 1719				United Kingdom Specification BS334	
Element	Corroding	Chemical	Acid-copper	Common desilverized	Pure Pb 99.99	Pure Pb 99.985	Smelter lead	Pure copper lead	Chemical type A	Chemical type B
Ag, max %	0.0015	0.020	0.002	0.002	0.001	0.001	0.001	0.0025	0.002	NS*
Ag, min %	—	0.002	—	—	—	—	—	—	—	NS
Cu, max %	0.0015	0.080	0.080	0.0025	0.001	0.001	0.001	0.08	0.003	NS
Cu, min %	—	0.040	0.040	—	—	—	—	0.04	—	NS
Ag + Cu, max %	0.0025	—	—	—	—	—	—	—	—	NS
As, max %	—	—	—	—	0.001	0.001	0.001	0.001	Trace	NS
Sb, max %	—	—	—	—	0.001	0.002	0.002	0.002	0.002	NS
Sn, max %	—	—	—	—	0.001	0.001	0.001	0.001	Trace	NS
As + Sb + Sn, max %	0.002	0.002	0.002	0.005	—	—	—	—	—	NS
Zn, max %	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.002	NS
Fe, max %	0.002	0.002	0.002	0.002	0.001	0.001	0.001	0.001	0.003	NS
Bi, max %	0.050	0.005	0.025	0.150	0.005	0.01	0.05	0.01	0.005	0.005
Cd, max %	—	—	—	—	—	—	—	—	Trace	NS
Co, max %	—	—	—	—	—	—	—	—	0.001	NS
Ni, max %	—	—	—	—	—	—	—	—	0.001	NS
Pb, min %	99.94	99.94	99.90	99.85	99.99	99.985	99.94	99.90	99.99	NS

*NS = not specified; open to negotiation between supplier and customer.

and in the United Kingdom, BS 334. The lead refined to meet these specifications is usually primary lead, that is, lead which is extracted and refined from ore. Secondary lead, recycled mostly from old storage batteries, can be refined to meet the same specifications as the primary metal, but the remelting, drossing, and recasting involved may have a negative impact on composition.

In the United States most lead ore is smelted and refined to a minimum purity of 99.85%. At and above this level of purity, four different grades of lead are recognized by ASTM B29: corroding, chemical, acid-copper, and common desilverized lead. The major differences among these grades are the allowable concentrations of copper, silver, and bismuth. Even trace amounts of these elements can have a significant effect on the properties or cost of the lead and justify having the four grades. Other elements whose concentrations are generally kept low include antimony, arsenic, tin, zinc, and iron. The chemical composition requirements for the four ASTM grades of lead are given in the **table**. The **illustration** shows an abridged flowsheet of lead metallurgy.

Concentrating. The first step in the beneficiation of ores to raise the lead content and to separate the lead from the zinc and iron minerals is concentrating. The standard processing begins with crushing the raw ore, followed by wet grinding, which reduces the product to a particle size of 75% minus 325 mesh. The resulting slurry is conditioned with certain reagents to establish a proper alkalinity, and by further mixing with flotation chemicals that collect the lead minerals in a froth, which is thickened and filtrated. Reagents include sodium carbonate, lime, copper sulfate, pine oil, cresylic acid, xanthate, and sodium cyanide, among others, in amounts ranging from 0.05 to 5.0 lb per ton (0.25 to 2.5 kg per metric ton) of ore. *See* FLOTATION.

During the lead flotation process, zinc minerals, iron compounds, and earth components of the ore are depressed instead of floated by this treatment, and are recovered after further separations. Copper, silver, and gold, if present, normally remain with the lead and are removed and recovered in the refinery. However, copper, if present in sufficient quantity, may be separated from the lead-copper concentrates by special flotation procedures.

The lead concentrate produced in this first processing step has a lead metal content of about 70%.

Smelting. Before smelting, lead concentrates are frequently blended with high-grade raw ores or returned intermediates (such as flue dusts, limerock, and so forth) which are drawn from proportioning bins. These materials are pelletized so that a homogeneous and carefully sized smelter feed is provided. The feed is then sintered to eliminate most of the sulfur and to agglomerate the particles into relatively large, hard lumps that will not be blown out of a blast furnace.

Sinter machines consist of a chain of moving pallets on which the porous feed is ignited. The ore pellets are subjected to an air blast (downdraft or updraft) which burns the sulfur, creating sulfur dioxide, and at the same time oxidizes the metallic elements. The sinter product, with about 9% of its weight of carbon in the form of coke, is charged into the top of a blast furnace. This is a simple vertical shaft that could be as high as 23.3 ft (7.11 m) and with a 5.5-ft (1.68-m) diameter. The coke supplies the fuel for melting the charge, and also the reducing gas which reacts with the lead oxide to form metallic lead. As the charge descends in the furnace, the molten metal flows to the bottom, from where it is

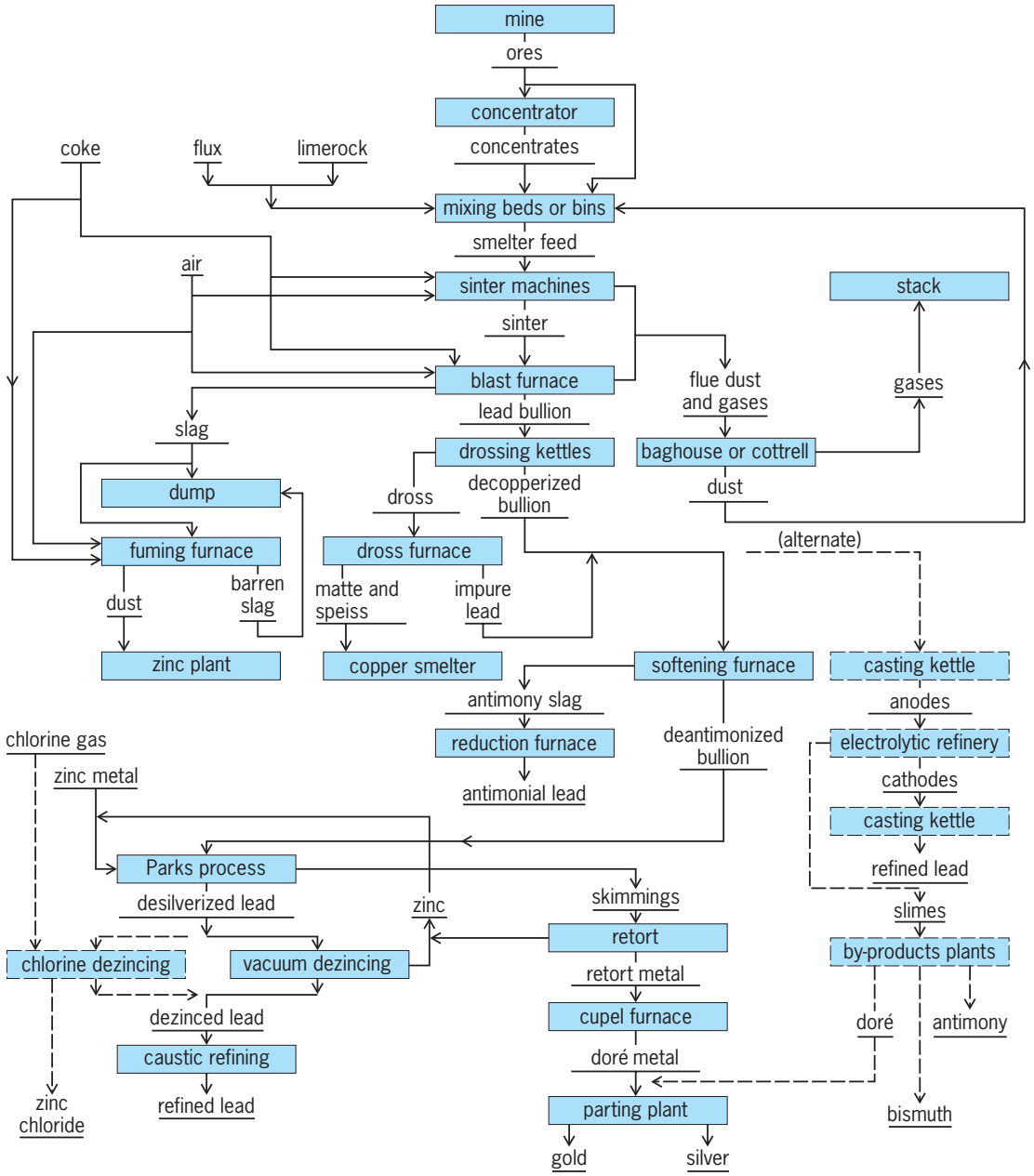

An abridged flowsheet of lead metallurgy.

withdrawn for further treatment. The remainder of the charge forms slag that floats on the lead and is removed from the furnace at a higher level. An improvement is the continuous tapping of slag, which eliminates the high labor requirement for the intermittent operations of the past.

The lead bullion contains the silver, gold, copper, bismuth, antimony, arsenic, tin, and other minor metals in the ore; the slag carries the zinc, iron, silica, lime, and other gangue; the dust contains a variety of elements, including cadmium and indium, plus much lead and zinc. A typical blast furnace takes about 750 tons (680 metric tons) of sinter per day and produces some 250 tons (230 metric tons) of lead bullion. *See* SINTERING.

Refining. The impure bullion is cooled in kettles to about 660°F (350°C), causing a dross to form which carries almost all of the copper as well as major amounts of lead. Most of the copper segregates in the kettle at low temperatures because the solubility of copper in lead, just above the latter's melting point, is very low. After the dross, which has a high copper content, is skimmed off, the rest of the copper is removed by stirring sulfur into the bath. At this stage, the work lead has been decoppered down to 0.01%. If significant quantities of tin are present, the lead bullion may be reheated to 1100°F (600°C), with the introduction of air, and a second dross containing the tin is removed. The lead bullion is then sent on to the refinery.

Slag from the blast furnace is frequently treated by fuming out the zinc and remaining lead. The barren slag is discarded. The dusts from the sinter plant and blast furnace are collected in a baghouse or precipitator, and returned to the mixing bins or proportioning bins.

Decopperized lead bullion, containing significant amounts of silver, gold, and other materials, is refined again by one of two principal processes. Most of the world's output is produced by means of pyrometallurgical techniques, while the remainder—less than 20%—is electrolytically refined. The latter process is used only when the bismuth content of the lead bullion is relatively high.

The first step in the pyrometallurgical operation is to soften the metal by removing the arsenic, antimony, and tin. This can be done with air oxidation of these elements in a small reverberatory furnace at a temperature of 1300–1380°F (700–750°C). The slag, which is normally high in antimony and lead content, is reduced with coke, and the resulting antimonial lead alloy is marketed. Another widely used method of removing these impurities is the Harris process. In this operation, liquid lead bullion is sprayed through molten caustic soda and molten sodium nitrate. The latter reagent oxidizes the arsenic, antimony, and tin, which are converted into sodium salts and skimmed from the bath. The three elements can be recovered from these compounds by wet chemical methods.

After the refining process has been carried to this point, the silver and gold are separated, using the Parkes zinc-desilverizing process. This method is based on the principle that when molten lead is saturated with an excess of zinc the precious metals become insoluble in lead. Intermetallic compounds with zinc are formed, and these float to the surface as a solid crust or dross which can be skimmed off. After two or three such treatments, the lead is free of silver and gold and proceeds to the next operation. The desilverized lead is essentially a eutectic melt of lead and zinc, and the latter is now usually extracted by vacuum dezincing, in which the main part of the zinc is evaporated under vacuum at 1100°F (600°C) and recovered in metallic form. Some plants dezinc the lead by passing chlorine through the metal, forming zinc chloride. The residue is essentially an impure alloy of silver and gold. After an oxidation treatment in a cupel furnace, the now partly refined alloy, termed doré metal, is electrolytically parted into refined silver and gold.

Grades of lead with a high bismuth content are unsuitable for some purposes, and so removal of the bismuth is economical for contents above 0.1%. Minor quantities of bismuth are likely when lead is refined by pyrometallurgical means. The bismuth may be eliminated by the Betterton-Kroll process, which involves treatment of the melt with small amounts of calcium and magnesium. These combine with the bismuth and rise to the surface as a dross which is readily skimmed. There are other methods for reducing the bismuth content to as little as 0.001% by precipitation with sodium. Removal of the bismuth is also possible by electrolytic refining of lead, but economically so only when the bismuth content exceeds 0.5%.

Lead produced by the processes described, or by others, can have a purity of 99.90 to 99.99%. At this level the grades of lead differ only by their bismuth content. Smelting techniques make it possible to produce commercial lead of any desired degree of purity, independent of the nature of the raw material. Lead with so-called six nines purity (99.9999) can be provided for special situations or scientific research. *See* ELECTROLYSIS; PYROMETALLURGY, NONFERROUS.

Alloys. If a user requires certain alloys to achieve desired properties, the elements are added later. In general, addition of small amounts of other metals increases the hardness and tensile strength of lead and decreases its ductility, malleability, and density. Often the tendency for lead to creep, its melting point, and the grain size are also reduced. Subsequent additions of the same metal do not usually provide a linear improvement in the mechanical properties. After a certain amount of alloying metal has been added, further additions often create alloys which are weaker and more brittle.

Another factor which depends on alloy concentrations is the tendency of certain alloys to improve in time due to age hardening. However, alloys that age-harden are also susceptible to overaging, which turns strong alloys into the original soft leads, with the alloying metal ending up in segregated concentrations. Heating can accelerate both age hardening and overaging.

The percentage of alloying metal which will provide a specific amount of improvement in mechanical properties depends in part upon the difference between the atomic radius of lead and that of the alloying metal. It also depends on whether the alloy is heat-treated and whether it is formed by casting, rolling, or extruding. Usually the greater the difference in atomic radii between the lead and the second metal, the greater the improvement in mechanical properties for the same percentage addition.

The effect of adding a third or fourth metal to a binary lead alloy can be too complex to predict. However, one indication of final results is the effect of adding the first alloying metal. In binary, ternary, and higher alloys, there are two other factors, besides a diminishing rate of mechanical improvement and increasing brittleness, which keep the optimal alloy concentrations low. These are cost and the loss of corrosion resistance. Silver, for example, improves corrosion resistance of lead in electrolytic processes, but its cost militates against broad use. Lithium, on the other hand, can make lead more vulnerable to corrosion in electrolytic processes. The amount of lithium that can be added to lead, therefore, can be limited by the unfavorable change in corrosion resistance.

The alloys of lead most frequently used in applications requiring high corrosion resistance are those which contain antimony, silver, tin, calcium, and copper. The addition of 1–13% antimony to lead creates alloys that have much greater tensile strength,

resistance to fatigue, and hardness than pure lead. This is why antimonial lead is often called hard lead and the removal of antimony in lead refining is called softening. Most antimonial lead is used for storage-battery components, and contains 3–8% antimony. Alloys with antimony levels of 0.80–1.15% and 6–8% are used most frequently in nonbattery applications. The low-end alloys are used to produce cable sheathing, while the 6–8% antimony alloys are used to fabricate a wide variety of equipment, such as tank linings, pipe, and anodes for chromium plating. Lead alloys with higher percentages of antimony are used to make castings when hardness is important. However, such alloys are rarely used because they are very brittle and do not have high corrosion resistance.

The addition of small amounts of calcium and tin to lead creates alloys which have significantly increased mechanical strength and which age-harden at room temperatures. The corrosion resistance of these alloys also is higher than that of antimony alloys in many applications. Calcium alloys, with or without tin, are used as battery grids, anodes, and roofing materials.

Tin alloys of lead have much lower densities and melting points than pure lead. These properties continue to decrease as more tin is added until the eutectic composition (approximately 63% Sn, 37% Pb) is reached. The lower melting point of these lead-tin alloys and increased strength have made them especially important as solders. The alloys used primarily for corrosion resistance, however, are low in tin content, with the most important alloy, terne, usually having a composition of 12–20% Sn. Terne is used to coat steel for use in items like roofing and automotive fuel tanks. *See* CORROSION; LEAD; LEAD ALLOYS; METAL, MECHANICAL PROPERTIES OF. A. L. Ponikvar

Bibliography. *ASM Metals Handbook*, vol. 11, pp. 493–522, 9th ed., 1979; D. H. Beilstein, The Herculaneum lead smelter of St. Joe Minerals Corporation, *Proceedings of the AIME World Symposium on Mining and Metallurgy*, *Lead and Zinc*, *St. Louis*, 1970; T. R. A. Davey and W. R. Bull, Process research on lead and zinc extraction, *Proceedings of the AIME World Symposium on Mining and Metallurgy*, *Lead and Zinc*, *St. Louis*, 1970; W. Hofman, *Lead and Lead Alloys*, 1970; M. L. Jaeck (ed.), *International Symposium on Primary and Secondary Lead Processing*, vol. 15, 1989; *Lead for Corrosion Resistant Applications*, Lead Industries Association, Inc., 1974; *A Metal for the Future*, St. Joe Minerals Corp., 1975; D. O. Rausch et al. (eds.), *Lead-Zinc Update*, 1977.

Leaf

A lateral appendage which is borne on a plant stem at a node (joint) and which usually has a bud in its axil. In most plants, leaves are flattened in form, although they may be nearly cylindrical with a sheathing base as in onion. Leaves usually contain chlorophyll and are the principal organs in which the important processes of photosynthesis and transpiration occur.

Morphology

A complete dicotyledon leaf consists of three parts: the expanded portion or blade; the petiole which supports the blades; and the leaf base. Stipules are small appendages that arise as outgrowths of the leaf base and are attached at the base of the petiole (**Fig. 1**). Stipules may be green and bladelike, as in pea; coarse rigid spines, as in black locust; sheaths, as in smart-weed; or more often, temporary structures which function only to protect the leaf while it is folded in the bud. Leaves that have a blade and petiole but no stipules are said to be exstipulate. Some leaves have no apparent petiole and are described as sessile. The leaves of monocotyledons may have a petiole and a blade, or they may be linear in shape without differentiation into these parts; in either case the leaf base usually encircles the stem. The leaves of grasses consist of a linear blade attached to the stem by an encircling sheath. At the junction of the sheath and the blade is a collarlike structure called a ligule.

Arrangement (phyllotaxy). Leaves are borne on a stem in a definite fixed order, or phyllotaxy, according to species (**Fig. 2**). Phyllotaxy is usually helical (spiral), with a single leaf at each node. Leaves may be borne in a compact arrangement as in rosette plants, or if the stem between the nodes (the internode) elongates so that successive leaves are widely separated, the plant is referred to as caulescent. If two leaves are borne on each node, the arrangement is opposite; if successive pairs of opposite leaves are

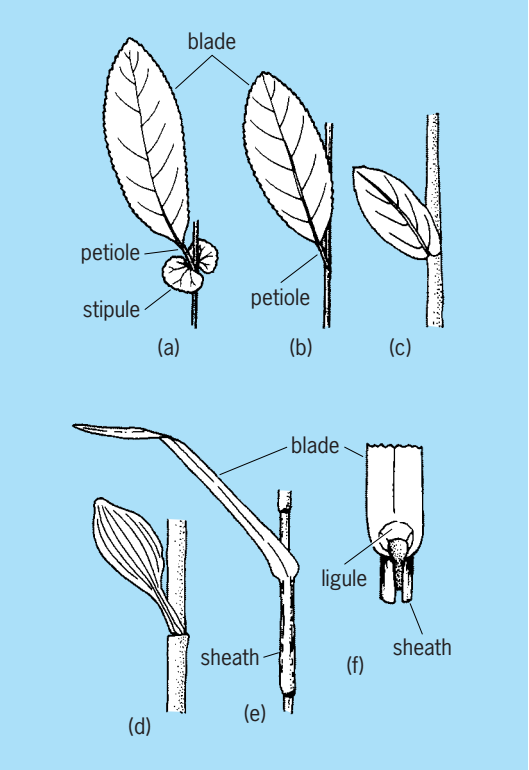

Fig. 1. Leaf parts in different structural patterns. (*a*) Complete (stipulate and petiolate). (*b*) Exstipulate. (*c*) Sessile. (*d*) Expanded monocotyledon leaf with sheathing base. (*e*) Grass leaf. (*f*) Detail of *e*.

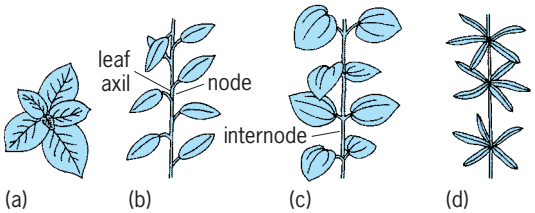

Fig. 2. Leaf arrangement. (*a*) Helical (top view). (*b*) Helical with elongated internodes (alternate). (*c*) Opposite (decussate). (*d*) Whorled (verticillate).

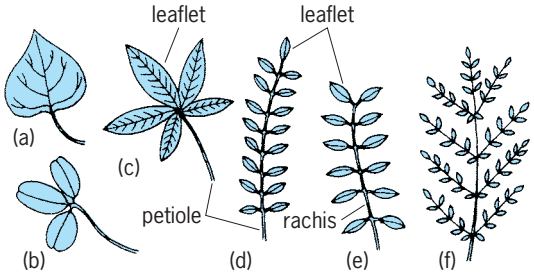

Fig. 3. Leaf types. (*a*) Simple. (*b*) Trifoliate. (*c*) Palmately compound. (*d*) Odd-pinnately compound. (*e*) Even-pinnately compound. (*f*) Decompound.

borne at right angles, it is decussate. When more than two leaves are present at a node, the arrangement is whorled (verticillate).

Types. A leaf is simple when it has only one blade. It is compound when the blade is divided into two or more separate parts called leaflets: trifoliolate if it has three leaflets (**Fig. 3**); palmately compound if leaflets originate from a common point at the end of the petiole; pinnately compound if leaflets are borne on the rachis (continuation of the petiole); odd-pinnate if such a compound leaf is terminated by a leaflet; even-pinnate if it is without a terminal leaflet; decompound (bipinnate) if it is twice compound.

Shapes. A leaf may be linear, long and narrow, with the sides parallel or nearly so (**Fig. 4**); lanceolate, narrow but tapering from base toward apex; oblanceolate, broader at the apex, tapering toward the base; spatulate, broad and obtuse at the apex, tapering to a narrow base; ovate, egg-shaped and broadest toward the base; obovate, the reverse of ovate, broadest toward the apex; elliptic, broadest at the middle and tapering slightly to a broadly rounded base and apex; oblong, somewhat rectangular, with nearly straight sides and a rounded base and apex; deltoid, triangular; reniform, kidney-shaped, broader than it is long; orbicular, circular or nearly so; peltate, shield-shaped, usually a circular leaf with the petiole attached at or near the center of the lower surface; perfoliate, having the stem apparently passing through it; connate, when the bases of two opposite leaves seem to have fused around the stem.

Margins. The margin, or edge, of a leaf may be entire (without indentation, or teeth); serrate (with sharp teeth pointing forward), or serrulate (finely serrate); dentate, with coarse teeth pointing outward, or denticulate (finely dentate); crenate or scalloped, with broad rounded teeth; undulate, with a wavy margin; incised, cut into irregular or jagged teeth or segments (if segments are narrow and pointed, laciniate; if directed backward, runcinate); pinnatifid, deeply pinnately parted (featherlike); or dissected, cut into numerous slender, irregularly branching divisions (**Fig. 5**).

When the blade is deeply cut into fairly large portions, these are called lobes. The degree of such lobing may be designated by the following terms: lobed, with sinuses usually not more than halfway from margin to midrib (midvein) or base, and with lobes and sinuses more or less rounded; cleft, when incisions extend halfway or more from margin to midrib, especially when they are sharp; parted, cut so deeply that the sinuses extend almost to the midrib or base; divided, cut entirely to the midrib, making a leaf compound.

Tips and bases. The tip of a leaf may be acuminate, gradually tapering to a sharp point; acute, tapering more abruptly to a sharp point; obtuse, with a blunt or rounded tip; truncate, seeming to be cut off square or nearly so; emarginate, decidedly notched at the

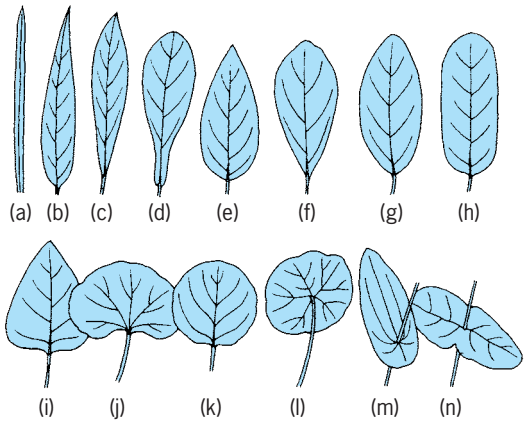

Fig. 4. Leaf shapes. (*a*) Linear. (*b*) Lanceolate. (*c*) Oblanceolate. (*d*) Spatulate. (*e*) Ovate. (*f*) Obovate. (*g*) Elliptic. (*h*) Oblong. (*i*) Deltoid. (*j*) Reniform. (*k*) Orbicular. (*l*) Peltate. (*m*) Perfoliate. (*n*) Connate.

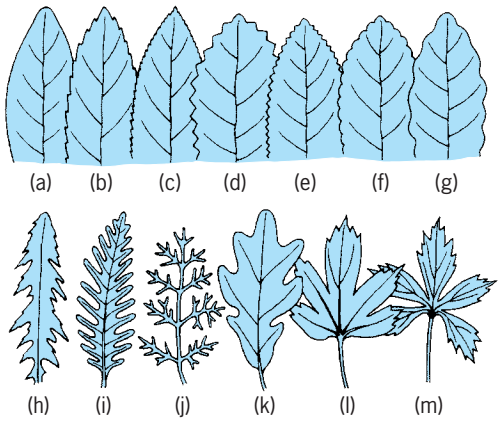

Fig. 5. Leaf margins of various types. (*a*) Entire. (*b*) Serrate. (*c*) Serrulate. (*d*) Dentate. (*e*) Denticulate. (*f*) Crenate. (*g*) Undulate. (*h*) Incised. (*i*) Pinnatifid. (*j*) Dissected. (*k*) Lobed. (*l*) Cleft. (*m*) Parted.

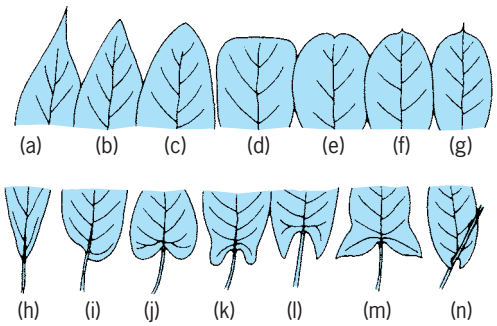

Fig. 6. Leaf tips and bases. (*a*) Acuminate. (*b*) Acute. (*c*) Obtuse. (*d*) Truncate. (*e*) Emarginate. (*f*) Mucronate. (*g*) Cuspidate. (*h*) Cuneate. (*i*) Oblique. (*j*) Cordate. (*k*) Auriculate. (*l*) Sagittate. (*m*) Hastate. (*n*) Clasping.

tip but not lobed; mucronate, abruptly tipped with a small short point; or cuspidate, ending in a sharp rigid point (**Fig. 6**).

The base of the blade may be cuneate, wedge-shaped; oblique, the two sides of the base unequal; cordate, heart-shaped, with a conspicuous sinus; auriculate, with a small earlike lobe on either side of the petiole; sagittate, arrow-shaped, with a pair of basal lobes turned outward; or clasping, sessile and partly investing the stem.

Venation. The arrangement of the veins, or vascular bundles, of a leaf is called venation (**Fig. 7**). As seen with the unaided eye, the venation appears in three basic patterns: dichotomous, reticulate (netted), and parallel. In dichotomous venation the veins are forked, with each vein dividing at intervals into smaller veins of approximately equal size; this pattern is characteristic of most ferns, gingko, and some primitive angiosperms. Reticulate venation may be described as a branching system with successively thinner veins diverging as branches from the thicker veins. The main veins may form a pinnate or palmate pattern, while the small veins form a fine network. In parallel-veined leaves, veins of relatively uniform size are oriented longitudinally, or nearly so, depending on the degree of expansion of the blade. The main longitudinal veins are usually interconnected with small veins. Reticulate venation is most common in dicotyledons, parallel venation in monocotyledons. *See* GINKGOALES; LILIOPSIDA; MAGNOLIOPSIDA; POLYPODIALES.

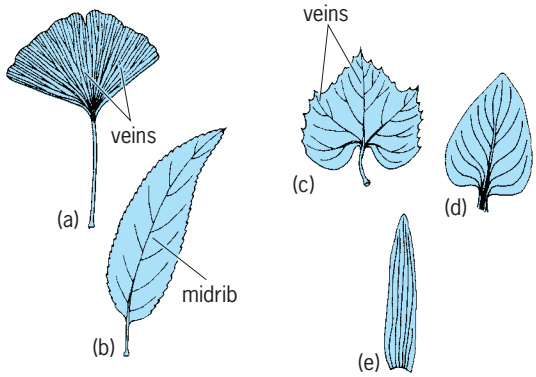

Fig. 7. Leaf venation. (*a*) Dichotomous. (*b*) Pinnate reticulate. (*c*) Palmate reticulate. (*d*) Parallel (expanded leaf). (*e*) Parallel (linear leaf).

Surface. Surfaces of leaves provide many characteristics that are used in identification. A surface is glabrous if it is smooth or free from hairs; glaucous if covered with a whitish, waxy material, or "bloom"; scabrous if rough or harsh to the touch; pubescent, a general term for surfaces that are hairy as opposed to glabrous; puberulent if covered with very fine, downlike hairs; villous if covered with long, soft, shaggy hairs; hirsute if the hairs are short, erect, and stiff; and hispid if they are dense, bristly, and harshly stiff.

Texture. The texture may be described as succulent when the leaf is fleshy and juicy; hyaline if it is thin and almost wholly transparent; chartaceous if papery and opaque but thin; scarious if thin and dry, appearing shriveled; and coriaceous if tough, thickish, and leathery.

Duration. Leaves may be fugacious, falling nearly as soon as formed; deciduous, falling at the end of the growing season; marcescent, withering at the end of the growing season but not falling until toward spring; or persistent, remaining on the stem for more than one season, the plant thus being evergreen. *See* DECIDUOUS PLANTS; EVERGREEN PLANTS.

Anatomy

The foliage leaf is the chief photosynthetic organ of most vascular plants. Although leaves vary greatly in size (from less than 0.04 in. or 1 mm in the duckweeds to over 66 ft or 20 m in some palms) and form (Figs. 1–7), they share the same basic organization of internal tissues and have similar developmental pathways. Like the stem and root, leaves consist of three basic tissue systems: the dermal tissue system, the vascular tissue system, and the ground tissue system. However, unlike stems and roots which usually have radial symmetry, the leaf blade usually shows dorsiventral symmetry, with vascular and other tissues being arranged in a flat plane.

Stems and roots have apical meristems and are thus characterized by indeterminate growth; leaves lack apical meristems, and therefore have determinate growth. Because leaves are more or less ephemeral organs and do not function in the structural support of the plant, they usually lack secondary growth and are composed largely of primary tissue only. *See* APICAL MERISTEM; ROOT (BOTANY); STEM.

The internal organization of the leaf is well adapted for its major functions of photosynthesis, gas exchange, and transpiration. The photosynthetic cells, or chlorenchyma tissue, are normally arranged in horizontal layers, which facilitates maximum interception of the Sun's radiation. The vascular tissues form an extensive network throughout the leaf so that no photosynthetic cell is far from a source of water, and carbohydrates produced by the chlorenchyma cells need travel only a short distance to reach the phloem in order to be transported out of the leaf (**Figs. 8** and **9**). The epidermal tissue forms a continuous covering over the leaf so that undue

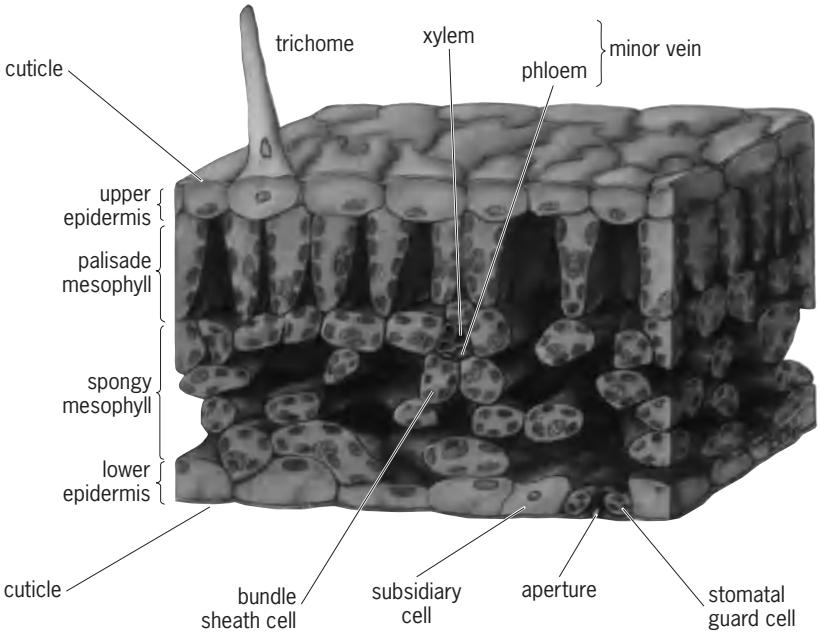

Fig. 8. Three-dimensional diagram of internal structure of a typical dicotyledon leaf.

water loss is reduced, while at the same time the exchange of carbon dioxide and oxygen is controlled.

Epidermis. The epidermis is usually made up of flat tabular cells which may be elongate, as in the linear leaves of grasses, or more or less square or lobed in surface view, as in many broad-leaved dicotyledons (Fig. 8 and **Fig. 10**). Regardless of shape, epidermal cells always fit tightly together without intercellular spaces, and they secrete a layer of hydrophobic substances, such as cutin and waxes, on the outside surface. Both of these adaptations reduce water loss. The layer of cutin is referred to as the cuticle.

The epidermal cells usually contain colorless plastids rather than chloroplasts, so that the epidermis tends to be a clear unpigmented layer of cells which allows light to penetrate to the subjacent photosynthetic tissue. The gas exchange required for photosynthesis takes place through thousands of minute stomates which usually occur at densities of about 0.15–0.75 in.^{2} (100–500/mm^2) [Fig. 10]. Each stomatal apparatus consists of a pore surrounded by two guard cells and associated subsidiary cells, which are sometimes morphologically differentiated from other epidermal cells. In the vast majority of vascular plants, the stomatal pore is closed at night so that evaporation of water vapor from internal air spaces is reduced, while in the light the guard cells are deformed in such a way that the pore is opened and carbon dioxide may enter the leaf. The ingenious mechanism of stomate opening and closure is as yet incompletely understood, but it appears to involve the pumping of potassium ions from the subsidiary cells to the guard cells, raising the osmotic concentration and increasing the hydrostatic pressure within the guard cells. The circumferential arrangement of cellulose microfibrils in the guard cell wall allows longitudinal expansion only; in addition, the ends of the guard cells are fixed, so that with increase in pressure the cells bend away from each other, opening the pore. The guard cells are unique in that they do possess chloroplasts, which provide a ready source of energy for this process. Stomates may occur on both the upper and lower leaf surfaces; or they may be restricted to the lower, shaded epidermis. In some cases such as the floating leaves of the water lily, stomates are present only in the upper epidermis.

The leaf epidermis of many vascular plants bears specialized hairlike structures called trichomes. These range from the simple prickle hairs on the leaves of many grasses, to the multicellular glandular hairs which give plants like geranium and tomato their characteristic odor, and to the complex highly branched hairs of plants like the wooly mullein. Except in obvious cases such as the modified trichomes of Venus' flytrap which secrete digestive

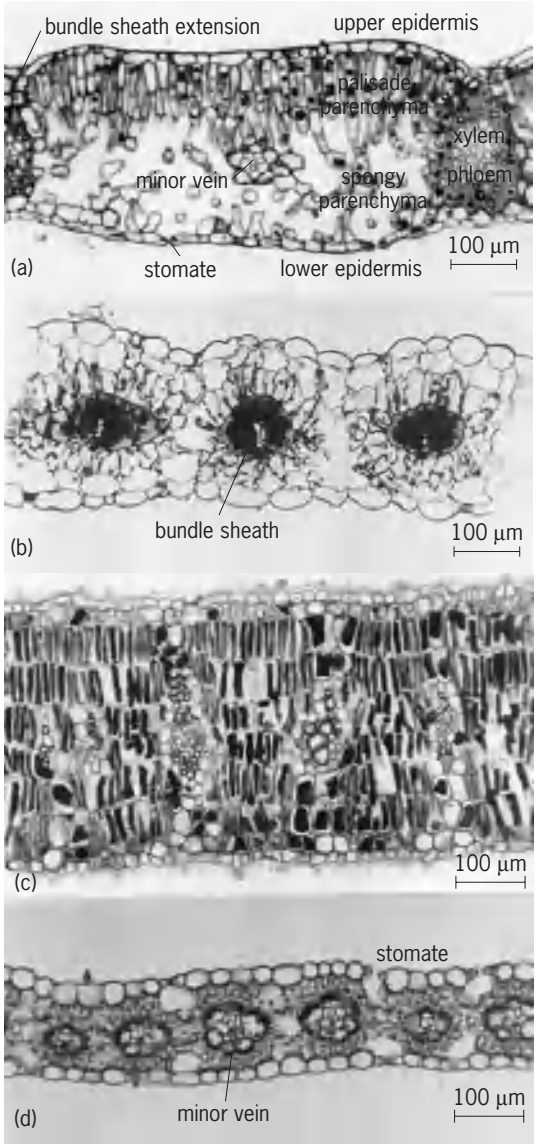

Fig. 9. Leaf cross sections. (*a*) Beech, *Fagus grandifolia*. (*b*) *Atriflex rosea*. (*c*) *Eucalyptus* sp. (*d*) Corn, *Zea mays*.

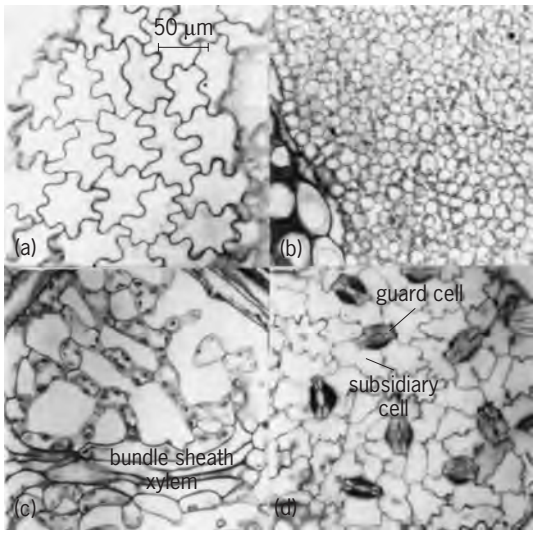

Fig. 10. Sections of beech (*Fagus grandifolia*) cut parallel to the surface (paradermal sections). (*a*) Upper epidermis. (*b*) Palisade parenchyma. (*c*) Spongy parenchyma. (*d*) Lower epidermis. (*From N. G. Dengler, L. B. MacKay, and L. M. Gregory, Cell enlargement and tissue differentiation during leaf expansion in beech, Fagus grandifolia, Can. J. Bot., 53:2846–2865, 1975*)

enzymes, the actual function of these widespread trichomes has not been demonstrated experimentally. Trichomes are believed to reduce water loss through the cuticle and stomates by creating a boundary layer on the leaf surface. A dense layer of trichomes may also reflect solar radiation from the leaf, protecting internal enzymes from damaging high temperature. In some cases trichomes may prevent insects, snails, and slugs from eating the leaves. *See* EPIDERMIS (PLANT).

Mesophyll. The ground tissue of leaves is referred to as mesophyll and consists of loosely packed parenchyma cells separated by intercellular spaces (Figs. 8 and 9). Within these cells are numerous green chlorophyll-containing chloroplasts, which are the site of photosynthesis. The mesophyll is frequently differentiated into palisade parenchyma and spongy parenchyma. The palisade parenchyma consists of one or more layers of columnar cells elongated at right angles to the epidermal layer, while the spongy parenchyma is made up of highly lobed cells with an extensive network of intercellular air spaces. Often, the largest intercellular spaces in the leaf occur in the mesophyll adjacent to a stomate, and are called substomatal chambers. In leaves with a horizontal orientation, the palisade cells usually lie next to the upper epidermis, while in leaves with a vertical orientation, layers of palisade may be adjacent to both epidermal layers, or as in the leaves of many grasses, there may be no differentiation into palisade and spongy parenchyma (Fig. 9). *See* PARENCHYMA.

Vascular tissue. The major veins of the leaf provide the connection between the vascular bundles of the stem and petiole and the xylem and phloem of the minor veins embedded in the mesophyll of the blade. Each major vein is associated with a rib of parenchymatous tissue, which usually projects from the lower surface of the leaf and can be seen with the naked eye. Strengthening tissues such as collenchyma and sclerenchyma are often associated with the major vein and provide the chief means of support for the leaf blade. Regardless of the pattern of the major veins (Fig. 6), the distribution of minor veins is such that every photosynthetic mesophyll cell is usually less than 100 micrometers away from the conducting tissues of the minor veins. Xylem occurs on the adaxial side of the minor veins and consists of both parenchyma and tracheary elements, which often have helical secondary walls (Fig. 10). Phloem occurs on the abaxial side of the vein and consists of sieve tube elements and phloem parenchyma. The conducting tissues at the minor veins are enclosed in a continuous sheath of ground tissue, the bundle sheath. The bundle sheath may be sclerenchymatous, but it more often consists of parenchyma cells with poorly developed chloroplasts. In many leaves the minor veins are connected with the epidermis by panels of bundle sheath cells called bundle sheath extensions. In leaves with freely ending veinlets, these extensions pass over the end of the veinlets so that the vascular tissue is not exposed directly to the intercellular spaces. In some plants, guttation (secretion of water) takes place from specialized vein endings and associated tissues called hydathodes. *See* PHLOEM; XYLEM.

Internal structure in relation to transport. Transpiration water leaving the tracheary elements must cross the bundle sheath parenchyma to reach the photosynthetic mesophyll. The bulk of the water is believed to travel either through the apoplast, that is, outside the living protoplast, or through the walls of the intervening cells. The evaporation of water from the cell walls of the mesophyll into the intercellular spaces provides the driving force for the bulk movement of the continuous column of water that extends from the soil through the xylem tracheary elements of root, stem, and leaf to the mesophyll cell walls. It is the cohesive force of this column which holds all of these together. In addition to moving directly from the xylem to the mesophyll, water moves from the veins through the bundle sheath extensions and then laterally along the epidermis. This pathway is particularly important for the palisade parenchyma cells which have limited lateral contact with one another. The bulk of the water lost from the leaf by transpiration evaporates from the surface of the internal tissues of the leaf, often 10 to 30 times the area of external leaf surface. Water vapor escapes from the leaves through stomatal pores and, to a much lesser extent, through the cuticle. *See* PLANT-WATER RELATIONS.

The photosynthetic mesophyll cells act as the source of carbohydrates for the rest of the plant. Sugars are believed to move in a symplastic pathway, a pathway in which materials move from cell to cell through cytoplasmic connections called plasmodesmata. Materials following this pathway do not leave the living protoplast as they move from the mesophyll, across the bundle sheath, and into the sieve tube elements of the phloem. The process of

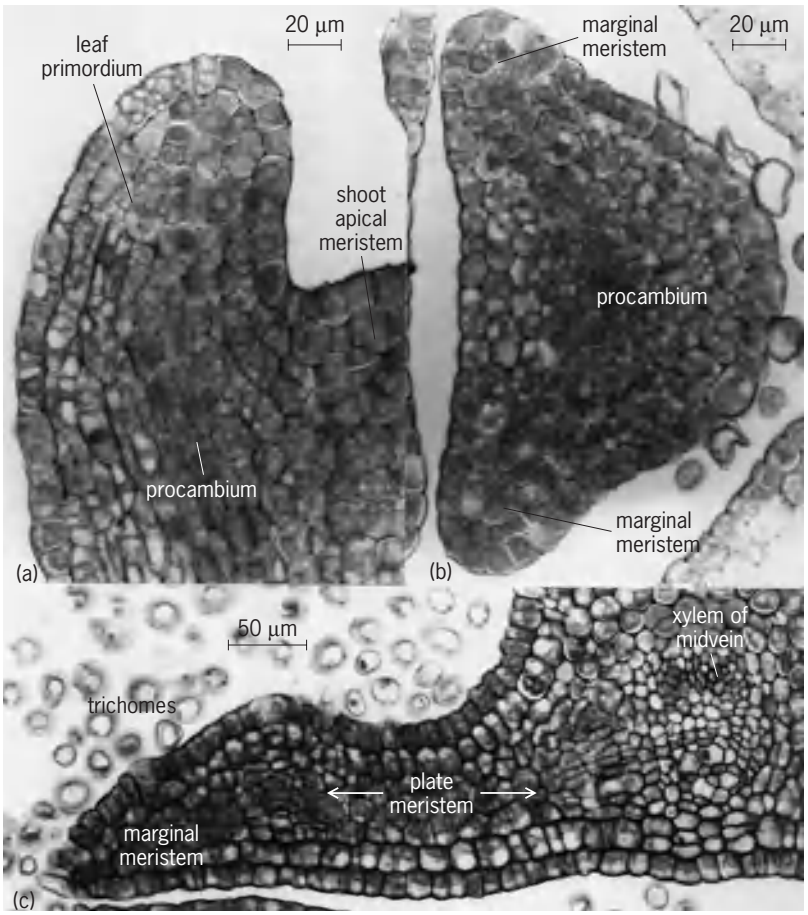

Fig. 11. Sections of developing leaves of *Calycanthus occidentalis*. (*a*) Longitudinal section of leaf primordium. (*b*) Cross section of leaf primordium. (*c*) Cross section of young leaf blade. (*From N. G. Dengler, Ontogeny of the vegetative and floral apex of Calycanthus occidentalis, Can. J. Bot., 50:1349–1356, 1972*)

loading carbohydrates against a concentration gradient into the sieve tube elements requires metabolic energy. The parenchyma cells surrounding the conducting cells often display features of transfer cells—cells with extensive wall invaginations which vastly add to the surface area of the cell membrane. *See* PLANT TRANSPORT OF SOLUTES.

C4 plants. In most monocotyledonous and dicotyledonous species, a three-carbon compound, 3-phosphoglyceric acid, is the first product of photosynthesis. These are called C3 plants. However, since the discovery of a new photosynthetic pathway in the 1960s, a growing number of plants have been shown to initially produce four-carbon organic acids as the first products of photosynthesis. These C4 plants are normally characterized by a distinctive type of leaf anatomy in which the bundle sheath parenchyma is made up of conspicuous dark green cells containing a number of large chloroplasts localized at one end of the cell (Fig. 9*b* and *d*). In addition, the photosynthetic mesophyll cells are often arranged concentrically around the bundle sheath in C4 plants, giving rise to the designation Kranz (wreath) anatomy for these species. Plants with the C4 pathway are more efficient in their incorporation of carbon dioxide than C3 plants, particularly at high light intensities and high temperatures. *See* PHOTOSYNTHESIS.

Development

The foliage leaf, regardless of size and shape at maturity, follows the same basic sequence of developmental events in its formation.

Initiation. The leaf arises as an emergence on the flanks of the apical meristem in the shoot tip. Localized cell divisions and accompanying cell enlargement first give rise to a small mound of cells and then to a fingerlike projection called a leaf primordium (**Fig. 11***a*). The new leaf primordium acts as a growth center, drawing nutrients from surrounding tissues and producing growth regulators such as auxin. Microsurgical experiments have shown that the growth center induces the differentiation of a strand of future vascular tissue called procambium which is in continuity with the vascular system of the stem. The procambium continues to grow toward the tip of the primordium as the primordium itself grows in length. The leaf primordium axis later forms the petiole and midrib of the mature foliage leaf, and the first procambial strand becomes the midvein of the leaf. Initially the leaf primordium elongates through concentrated meristematic activity near the apex; eventually, cells at the primordium apex begin to mature, and further growth in length is the result of cell divisions scattered throughout the young leaf. *See* BUD.

Blade formation. The leaf blade is usually initiated when the leaf primordium is less than a half millimeter in length. The blade arises as a ridge of tissue produced by a zone of cell division and cell enlargement on either side of the primordial axis called marginal meristems (Fig. 11*b*). Analysis of the planes of divisions in the marginal meristems has shown that cell divisions in the surface layer are largely perpendicular to the leaf surface, giving rise to a continuous sheet of cells which will later mature as epidermis. Cell divisions in the internal tissues occur in a variety of planes, adding to the volume of the developing leaf blade, and determining the number of cell layers in the mature leaf. Cell divisions in both the surface and internal layers are most frequent in a region several cells removed from the actual leaf margin. In many plant species, the marginal meristems are active for only a limited period of time, and expansion of the leaf blade is the result of cell division and enlargement throughout the blade between the margin and midrib called plate meristem activity. The plane of cell divisions in the plate meristem is predominantly perpendicular to the leaf surface. This means that the number of cell layers established by the marginal meristem is perpetuated, and that with cell enlargement, the leaf expands in size manyfold in a horizontal plane.

At the time that the leaf expands from the bud, between 50 and 90% of its cells have already been formed. Therefore, the final stages of leaf expansion are mostly the result of enlargement of cells that have been formed by divisions in the plate meristem. Characteristic differences in the rate, duration,

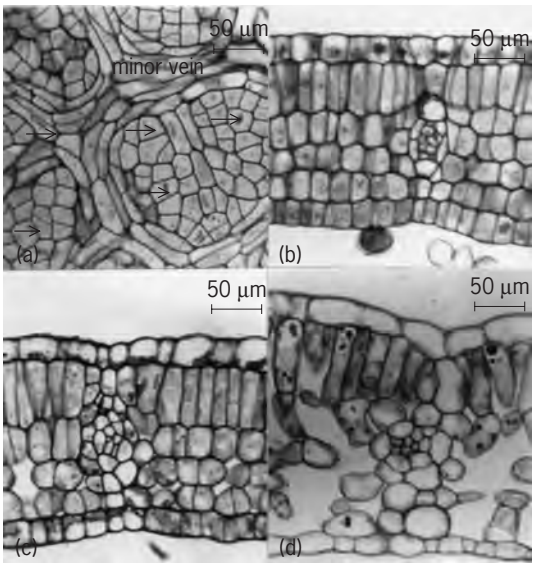

Fig. 12. Sections of beech (*Fagus grandifolia*) leaves during expansion from the winter bud. (*a*) Paradermal section of spongy parenchyma at time of bud swelling. Arrows indicate cell divisions in plate meristem. (*b*) Cross section of leaf at same stage. (*c*) Cross section of leaf 1 week later. (*d*) Cross section of leaf 2 weeks later. (*From N. G. Dengler, L. B. MacKay, and L. M. Gregory, Cell enlargement and tissue differentiation during leaf expansion in beech, Fagus grandifolia, Can. J. Bot., 53:2846–2865, 1975*)

and direction of cell enlargement occur between the various tissues of the leaf blade. For instance, cells that will become the palisade parenchyma elongate greatly in a vertical direction, and very little in a horizontal direction (**Fig. 12**); in addition, cell divisions occur for a longer period in the palisade layers. In contrast, cells of the upper epidermis undergo a great deal of horizontal expansion and very little vertical expansion. The intercellular spaces of the palisade parenchyma layers are formed when the enlargement of the upper epidermal cells draws the attached palisade cells apart. Cell expansion in the spongy parenchyma layers does not occur uniformly over the cell surface, resulting in a lobed cell shape and an extensive network of intercellular spaces. The procambial strands that will differentiate as minor veins are formed from localized cell divisions in the plate meristem (Fig. 12). The minor vein pattern is blocked out early during plate meristem activity, and the areoles, regions of mesophyll surrounded by vascular tissue, are well defined throughout the period of leaf expansion. Maturation of xylem and phloem occurs first in the major veins, starting at the base of the leaf and proceeding toward the apex. This is followed by maturation of the xylem and phloem of the minor veins, which proceeds from the apical portion of the leaf blade toward the base. *See* PLANT GROWTH.

Variations in developmental pattern. The process of leaf development in plants with lobed or compound leaves does not differ qualitatively from the mechanisms described above. Rather, differences in leaf shape arise through differences in the location and direction of marginal and plate meristem activity. The long linear leaves of grasses and some other monocotyledons, however, do possess some unique developmental features. At the earliest stages of development, the leaf primordium is crescent-shaped and encircles the shoot apical meristem. After a short period of apical growth along the rim of the crescent, most of the meristematic activity is localized at the base of the leaf in the intercalary meristem. A majority of cell divisions are in a plane which is perpendicular to the future long axis of the leaf. This means that cells of all tissue layers are formed in long files, with those cells near the apex of the leaf reaching maturity while new cells are being added to the base of each file. Intercalary growth may be prolonged, and may occur sporadically throughout the life of an individual leaf. This unique pattern of growth is often regarded as an adaptation which is associated with the evolution of grazing mammals.

Leaf development in rapidly growing herbaceous plants is a continuous process in which leaves produced on the shoot apical meristem expand, carry on photosynthesis, and eventually senesce when the reproductive stage of the life cycle nears an end. In most perennial plants such as the beech illustrated in Figs. 9, 10, and 12, new leaf primordia are formed late in the growing season and overwinter in the apical bud. In the spring the young leaves can quickly complete plate meristem activity and leaf expansion so as to maximize the period of photosynthesis. Leaf abscission occurs at the end of the growing season in deciduous woody plants. The abscission zone may be structurally weak, but the actual separation of leaf from stem is the result of enzymatic degradation of cell walls. A protective layer is produced by the deposition of protective substances such as suberin and wound gum, and the wound is often further protected by the development of a layer of cork. The resulting leaf scar and the bundle scars which mark the position of the abscission of the leaf midvein are often distinguishing features in the identification of species. *See* ABSCISSION; PERIDERM; PLANT ANATOMY.

Nancy G. Dengler

Lean manufacturing

A unique linked-cell manufacturing system. Initiated in the 1960s by the Toyota Motor Company, it is also known as the Toyota Production System (TPS), the Just-in-Time/Total Quality Control (JIT/TQC) system, or World Class Manufacturing (WCM) system. In 1990, it was given a name that would become universal, "lean production." This term was coined by John Krafcik, an engineer working in the International Motor Vehicle program at MIT with J. P. Womack, D. Roos, and D. T. Jones, who observed that this new system used less of the key resources needed to make goods. What is different about the system is its use of manufacturing cells linked together with a functionally integrated system for inventory and production control. The result is low cost (high efficiency), superior quality, and on-time delivery of unique products from a flexible system. *See* FLEXIBLE MANUFACTURING

SYSTEM; INVENTORY CONTROL; MANUFACTURING ENGINEERING.

Mass versus lean production. The key proprietary aspects of lean production are U-shaped manufacturing and assembly cells, using walking workers and are designed with system requirements in mind. Design decoupling allows the separation (decoupling) of processing times for individual machines from the cycle time for the cell as a whole, permitting the lead time to make a batch of parts to be independent of the processing times for individual machines. This takes all the variation out of the supply chain lead times, so scheduling of the supply system can be greatly simplified. The supply chain is controlled by a pull system of production control called Kanban. Using Kanban, the inventory levels can be dropped, which decreases the throughput time for the manufacturing system. *See* SUPPLY CHAIN MANAGEMENT.

Figure 1 shows an entire mass production factory reconfigured into a lean manufacturing factory. The final assembly lines are converted to mixed model final assembly; this levels the demand for subassemblies and other components. The rate of production is determined by recalculating the monthly demand into a daily demand and trying to make the same product mix every day.

The subassembly lines are reconfigured into U-shaped cells. The daily output from these cells is balanced to match the demand from final assembly.

The traditional job shop generates a variety of unique products in low numbers with its functional design. It is redesigned into U-shaped manufacturing cells that produce families of component parts (**Fig. 2**).

Lean manufacturers focus on sole-sourcing each component or subassembly (that is, they do not have multiple vendors supplying the same components), sharing their knowledge and experience in linked-cell manufacturing with their vendors on a one-to-one basis. For lean automobile manufacturers, the final assembly plant may have only 100 to 400 suppliers, with each supplier becoming a lean or JIT vendor to the company. In the future, the number of vendors supplying a lean manufacturer will decrease even more. The Mercedes-Benz plant in Vance, Alabama, has around 80 suppliers. The subassemblies contain more components as the vendors take on more responsibility for the on-time delivery of a larger portion of the auto.

Lean cell design. In a true lean manufacturing system, manufacturing processes and equipment are designed, built, tested, and implemented into the manufacturing cells. The machine tools and processes, the tooling (workholders, cutting tools), and the material-handling devices (decouplers) are designed specifically for cellular manufacturing. Simple, reliable equipment that can be easily maintained should be specified. In general, flexible, dedicated equipment can be built in-house better than it can be purchased and modified for the needs of the cell.

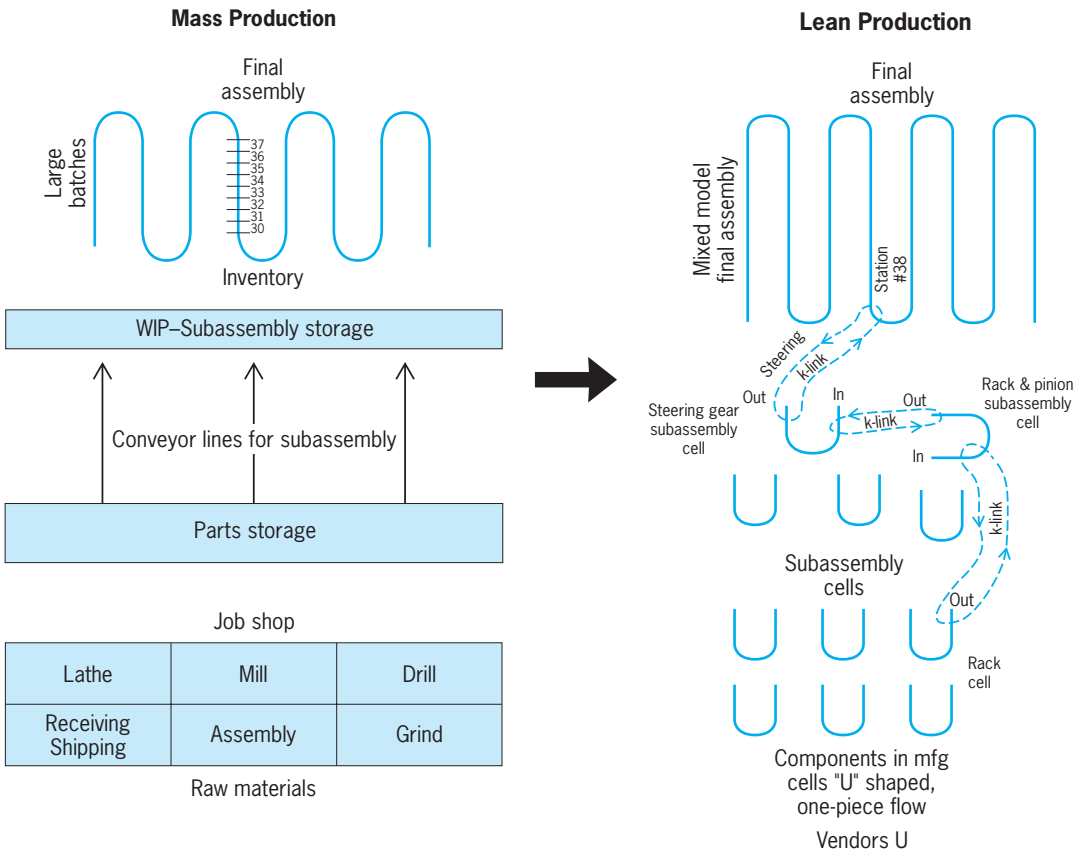

Fig. 1. Restructuring of a job shop/flow shop final assembly design into a linked-cell manufacturing system.

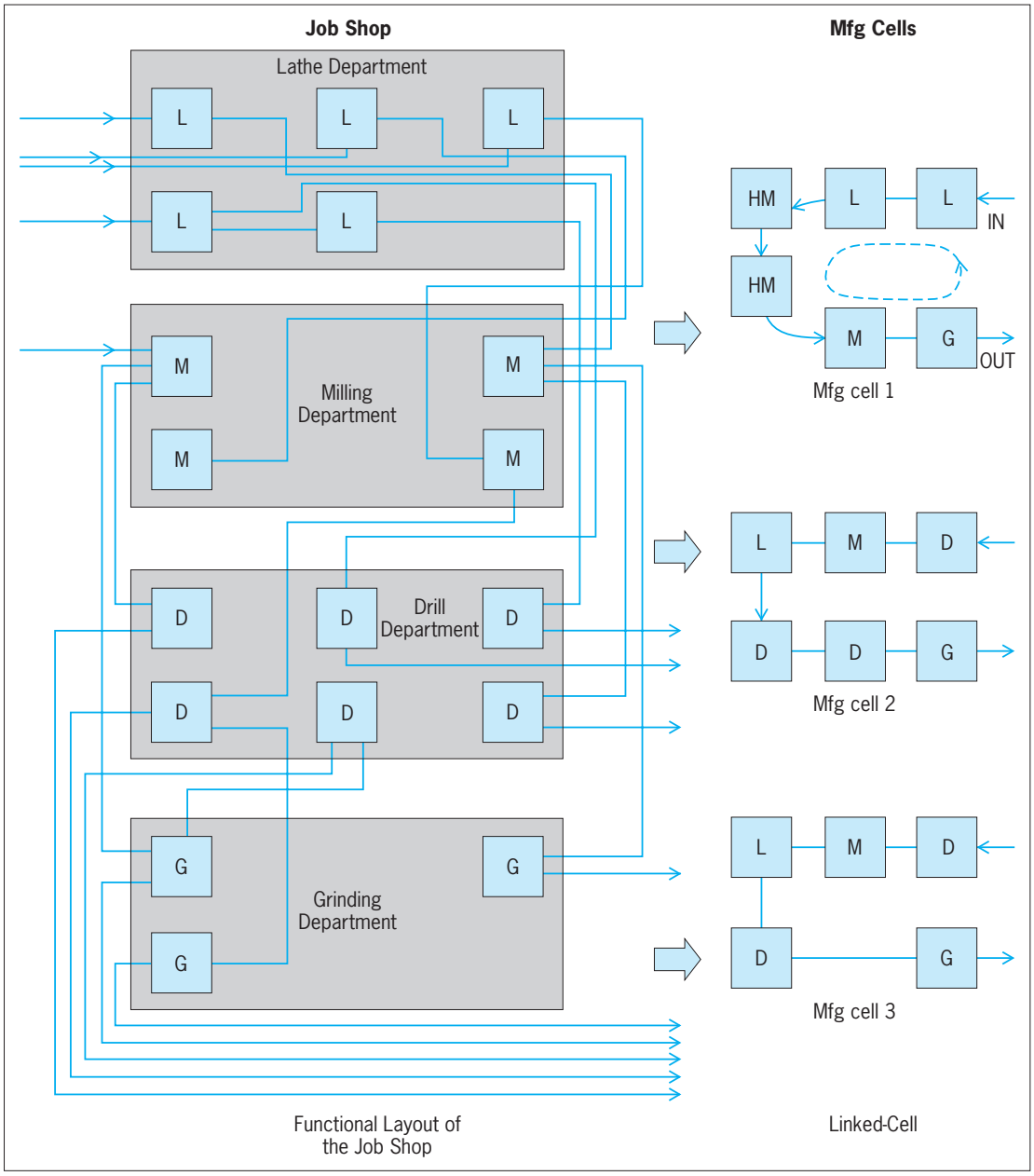

Fig. 2. Job shop portions of the factory require a system level conversion to redesign them into manufacturing cells.

However, many plants that lack the expertise to build machines from scratch do have the expertise to modify equipment to give it unique capabilities, and this interim cell approach can lead a company into true lean manufacturing cells. *See* MATERIALS-HANDLING EQUIPMENT.

Many companies understand that it is not good strategy simply to imitate manufacturing process technology from another company and then expect to make an exceptional product. When process technology is purchased from outside vendors, any unique aspects will be quickly lost. Companies must carry out research and development on manufacturing technologies and systems in order to produce effective and cost-efficient products, but the result makes such investment pay off. There are many advantages to this home-built equipment strategy.

Flexibility (process and tooling adaptable to many types of products). Flexibility requires rapid changeover of jigs, fixtures, and tooling for existing products and rapid modification for new designs. The processes have excess capacity; they can run faster if necessary, but they are designed for less than full-capacity operation.

Building to need. There are three aspects to this. First, there is no unused capability or options. Second, the machine can have unique capabilities that competitors do not have and cannot get access to through equipment vendors. For example, in the lean cell shown in **Fig. 3** is a broaching (cutting) machine for producing the gear teeth on a rack bar. The angle that these teeth make with the bar varies for different types of racks. Job shop broaching machines are not acceptable for this cell because of their

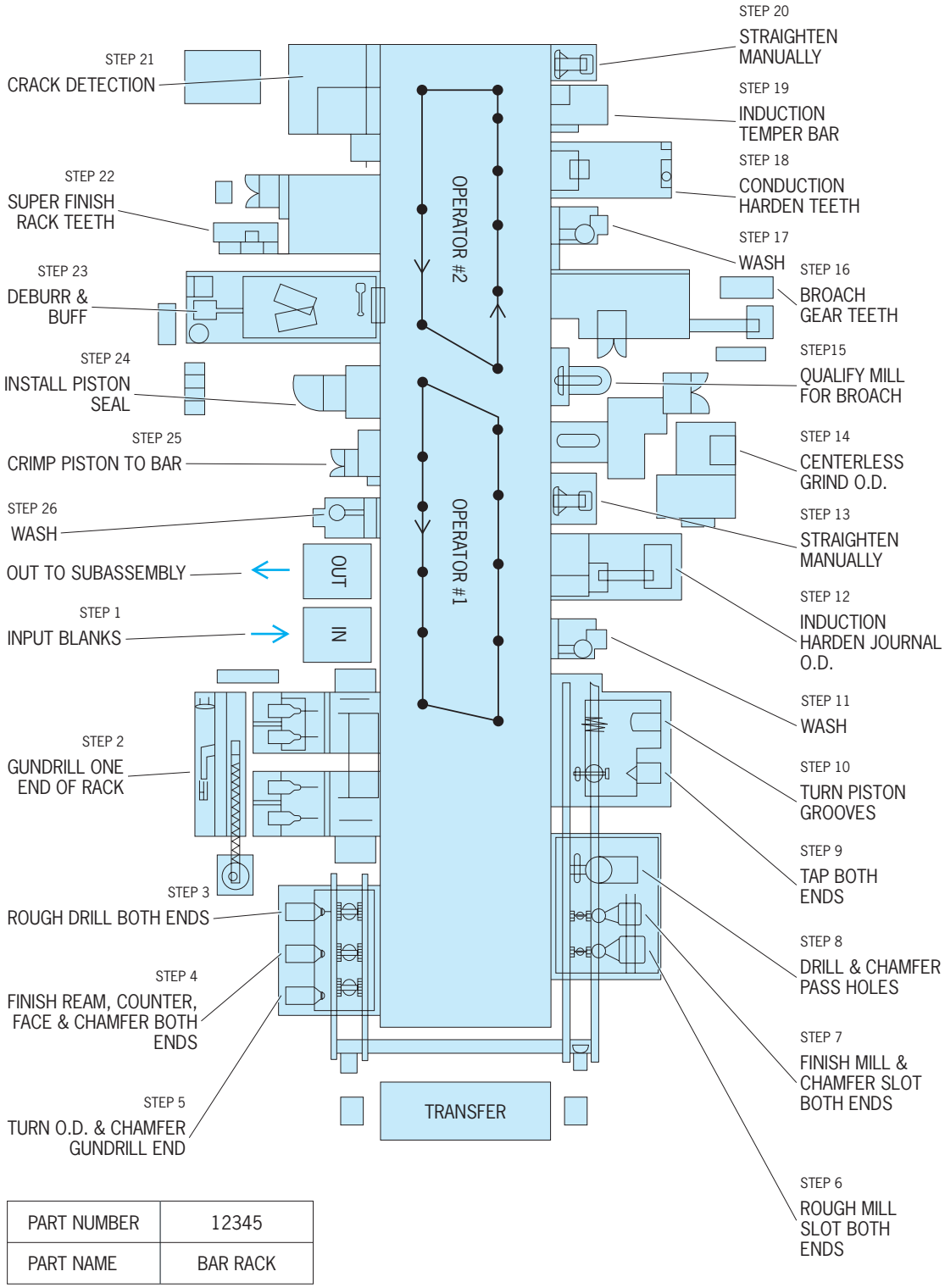

Fig. 3. Lean manufacturing cell for producing rack bars for a rack-and-pinion steering gear.

long changeover times, so a unique machine tool for broaching was designed and built using a proprietary process. Third, the equipment should allow the operator to stand and walk. Equipment should be the appropriate height to allow the operators to easily perform tasks standing up and then move to the next machine in a step or two (that is, the design should have a narrow footprint).

Built-in maintainability/reliability/durability. Equipment should be easy to maintain (to oil, clean, changeover, or replace worn parts, and to use standardized screws). Many of the cells at lean vendors are clones of each other. The vendor company, being the sole source, has the volume and the expertise to get business from many companies, making essentially the same components or subassemblies for many

original equipment manufacturers (OEMs). The equipment can be interchanged from one cell to another in emergencies.

Machines, handling equipment, and tooling built for the needs of the cell and the system. Machines are typically single-cycle automatics but may have capacity for process delay. An example of process delay is a heat treatment process that takes 4 min in a cell with a 1 min cycle/time. If the heat treatment machine can hold four units, each getting 4 min of treatment, one unit is still produced per minute.

Safety. Equipment is designed to prevent accidents (fail-safe).

Equipment designed to be easy to operate, load, and unload. Research has shown that manufacturing cells have ergonomic benefits over the job shop. Toyota ergonomists recommend unloading with the left hand, loading with the right hand, while walking right to left. The wide variety of tasks over the cell cycle time keep the operators from developing cumulative trauma injuries.

Equipment designed to process single units, not batches. Small-area, low-cost equipment is the best. Machining or processing time (MT) should be modified so that it is less than the cycle time (CT), the time in which one unit must be produced. Equipment processing speed should be set in view of the cycle time, such that MT < CT. The MT is related to the machine parameters selected. For example, in a cutting operation this approach often permits the reduction of the cutting speed, thereby increasing the tool life and reducing downtime for tool changes. This approach also reduces equipment stoppages, lengthens the life of the equipment, and may improve quality.

Equipment that contains self-inspection devices (such as sensors and counters) to promote autonomation. Autonomation is the autonomous control of quantity and quality. Often the machine is equipped to count the number of items produced and the number of defects. *See* QUALITY CONTROL.

Equipment that is movable. Machines are equipped with casters or wheels, flexible pipes, and flexible wiring. There are no fixed conveyor lines.

Equipment that is self-cleaning. Equipment disposes of its own chips and trash.

Equipment that is profitable at any production volume. Equipment that needs millions of units to be profitable (R. J. Schonberger calls them supermachines) should be avoided, because if the production volume ever exceeds the maximum capacity that the first supermachine can build, it will be necessary to purchase another supermachine and the new one will not be profitable until it approaches full use. Schedulers of the equipment will typically divide the volume into two machines, so neither will be profitable.

Lean cell operation. The manufacturing cell shown in Fig. 3 typically uses two operators. These standing, walking workers move from machine to machine in counterclockwise (CCW) loops, each completing the loop in about 1 min. Operator 1 typically addresses 10 stations, and operator 2 addresses 11 stations. Most of the steps involve unloading a machine, loading another part into the machine, checking the part unloaded, and dropping the part into the decoupler (handling) elements between the machines. The stock-on-hand (SOH) in the decouplers and the machines helps to maintain the smooth flow of the parts through the machines. The decouplers also can be designed to perform inspections for part quality or necessary process delays while the parts heat up, cool down, cure, and so on. The stock-on-hand is kept as low as possible.

Sometimes the decoupler elements perform the inspection of the part, but mostly they serve to transport parts from one process to the next. The decoupler may also perform a secondary operation such as deburring or degaussing (removal of residual magnetic fields).

By design, one operator controls both the input and output of the cell. One operator always controls the volume of material going through the cell. This keeps the stock-on-hand quantity constant and keeps the cell working in balance with the final or subassembly lines it is feeding.

At the interface between the two operators, either one can perform the necessary operations, depending on when they arrive and when the processes in the machines are finished. That is, the region where the two operators meet is not fixed, but changes or shifts depending upon the way parts are moving about the cell. This is called the relay zone. This added flexibility requires that the workers are cross-trained on all the processes in the cell.

Ergonomics of lean cells. Ergonomics deals with the mental, physical, and social requirements of the job, and how the job is designed (or modified) to accommodate human limitations. For example, the machines in the cell are designed to a common height to minimize lifting of parts, transfer devices are designed for slide on/slide off, and automatic steps equipped with interrupt signaling help workers monitor the process. When the job is primarily loading/unloading, ergonomic concerns regarding lifting and placing parts in machines and operating workholding devices must be addressed. Lean manufacturing cells are relatively free of cumulative trauma syndrome (CTS) problems because the operators' tasks and movements are so varied from machine to machine. Cell designers should try to incorporate ergonomic issues initially rather than trying to implement fixes later.

Human performance in detecting and correcting cell malfunctions will establish utilization and production efficiency. The design of machines for maintainability and diagnostics is critical. In manufacturing and assembly cells where workers operate machines, it is important that all the machines are ergonomically identical. Sewing machines in a cell are a good example: To the operator, all the machines should feel the same in terms of control.

Manufacturing system integration. Many believe that the only way in which manufacturing companies can compete is to automate. This approach, known as computer integrated manufacturing (CIM), was recently renamed agile manufacturing. The concept is to achieve integration through computerization and

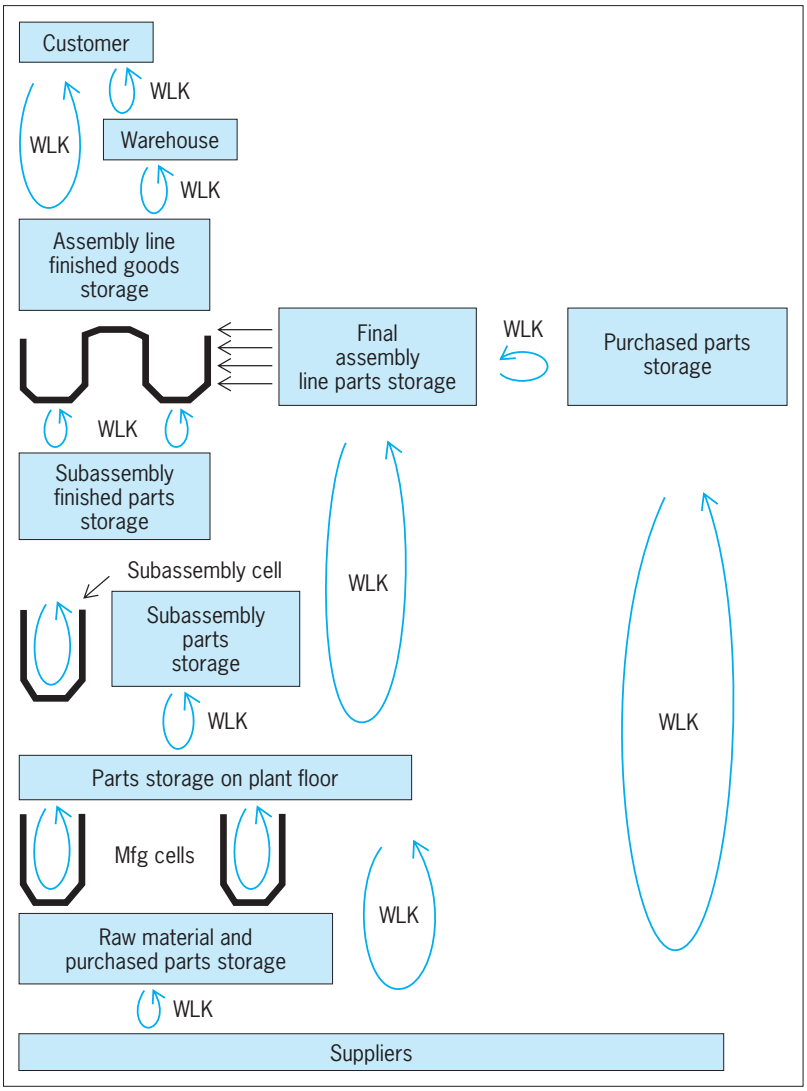

Fig. 4. Linked-cell design. In the lean manufacturing system, the manufacturing and assembly cells are linked to the final assembly area by Kanban (card) inventory links or loops. WLK indicates a withdrawal card.

automation. This often results in trying to computerize, robotize, or automate very complex manufacturing and assembly processes, and it works only when there is little or no variety in the products. Lean manufacturing systems take a different approach—integrate the manufacturing system, then computerize and automate (IM, C). Experts on CIM now agree that lean manufacturing, especially development of manufacturing and assembly cells, must come before efforts to computerize the system. While costs of these systems are difficult to obtain, the early evidence suggests that the lean cell approach is significantly cheaper than the CIM approach. *See* AUTOMATION; COMPUTER-INTEGRATED MANUFACTURING.

Outlook. The lean factory is based on a different design for the manufacturing system in which the sources of variation in time are minimized and delays in the system are systematically removed. In the linked-cell manufacturing system, in which manufacturing and assembly cells are linked together with a pull system for material and information control (**Fig. 4**), downstream processes dictate upstream production rates. The linked-cell manufacturing system strategy simplifies the manufacturing system, integrates the critical control functions before applying technology (automation, robotization, and computerization), avoids risks, and makes automation easier to implement. This is the strategy that will predominate in the next generation of factories worldwide. *See* INDUSTRIAL ENGINEERING; PRODUCTION ENGINEERING; PRODUCTION METHODS. J T. Black

Bibliography. R. U. Ayres and D. C. Butcher, The flexible factory revisited, *Amer. Scientist*, 81:448–459, 1993; J T. Black, Cell design for lean manufacturing, *Trans. NAMRI/SME*, 23:353–358, 2000; J T. Black and B. J. Schroer, Simulation of an apparel assembly cell with walking workers and decouplers, *J. Manuf. Sys.*, 12(2):170–180, 1993; J T. Black, C. C. Jiang, and G. J. Wiens, Design, analysis and control of manufacturing cells, *PED*, vol. 53, ASME, 1991; Y. Monden, *Toyota Production System*, Industrial Engineering and Management Press, IIE, 1983; S. Nakajima, *TPM, Introduction to TPM: Total Productive Maintenance*, Productivity Press, 1988; R. J. Schonberger, *Japanese Manufacturing Techniques: Nine Hidden Lessons in Simplicity*, Free Press, 1982; K. Sekine, *One-Piece Flow: Cell Design for Transforming the Production Process*, Productivity Press, 1990; S. Shingo, *A Study of The Toyota Production System*, Productivity Press, 1989; N. P. Suh, Design axioms and quality control, *Robotics and CIM*, 9(4/A):367, August–October 1992; J. P. Womack, D. T. Jones, and D. Roos, *The Machine That Changed The World*, Harper Perennial, 1991.

Learning mechanisms

Those processes, activities, structural components, and alterations of the nervous system that produce the behavioral manifestations and changes due to experience and event associations (rather than those due to maturation, fatigue, and other developmental or temporary states).

Central issues. The search for a neural basis of learning has long been guided by several central issues, as described below.

Definition of learning. At the root of any search for the neural mechanism of learning is the problem of defining learning. Most simply, learning may be defined as a change in behavior. However, this definition is certainly too broad since any change in behavior could then be called learning whether it resulted from drug intoxication, sleep, or even death. It is now common to use the term behavioral plasticity for behavior changes that seem to involve at least some aspects of learning and to subdivide the definition into specific types of learning, such as motor learning, sensory learning, and associative learning. In this way, the definition of learning has been narrowed to indicate a change in behavior that is relatively permanent, that is not caused by development, maturation, fatigue, or other short-term processes, and that is due to experience or reinforcement. Hence,

what is known as sensitization, a process in which a given behavior increases in intensity simply with repeated occurrences, is an example of behavioral plasticity, whereas learning to play basketball involves the more specific categories of both motor and sensory learning.

Several types of learning have been recognized, perhaps the simplest of which is associative learning, that is, learning by association of one stimulus with another through the repeated simultaneous occurrence of the two stimuli. This type of learning is typified by the dog that learns to associate the sound made when a package of dog meal is opened with the food being placed in its dish. This is an example of Pavlovian conditioning. Motor learning, on the other hand, occurs when an animal or human learns to perform some motor task in response to a given event or stimulus. Sensory learning can be thought of as that learning in which an organism is trained to respond to changes in some stimulus or to differences between stimuli. Included here are such familiar things as learning faces or tunes, learning to distinguish two shades of color, and connecting voices with faces. The last is an example of cross-modal generalization learning in which a stimulus in one sensory modality (the auditory) recalls the image of the face in another (the visual). In the human, there are other categories of learning, such as cognitive learning in which the human learns without any obvious external stimuli, but simply through the internal activity of the brain, commonly referred to as consciousness. This is the most difficult type of learning to investigate, since there is little control over the process. *See* COGNITION; CONDITIONED REFLEX.

It is important to recognize that in the search for neural substrates of learning, the type of learning must be specified, since it is evident that the term learning can apply to many different types of events in the brain. Even though all of these events may have at least some common neural substrates, they involve different areas of the brain—and probably different neural mechanisms as well. The problems encountered when defining learning have become stumbling blocks in determining which neural processes are to be studied, because there is little consensus about what is to be measured.

Relationship between learning and memory. The question of how learning and memory differ is important in designing studies of the underlying processes of learning. Learning is really a hypothetical construct that is presumed to take place when certain conditions are met, and it is the basis for subsequent changes in the behavior of an organism in response to its environment. Memory, on the other hand, is the storage of those alterations for later use. Without memory, there would be no way to determine if learning had occurred since whatever had been learned would be immediately forgotten. Memory, then, is always needed to assess learning and cannot be completely separated from it. In the same way, any assessment of processes underlying learning will also assess memory, although the two can be analyzed somewhat independently by appropriate experimental designs.

The term engram has long been applied to the memory trace of learning. This term denotes that special alteration which occurs as a consequence of learning and which is then translated into altered behavioral output at a later time. Thus, the engram is a stored representation of the learning experience and is presumably open to investigation. *See* MEMORY.

Physiological process. Since the behavior of all animals above the level of the simplest multicellular organisms is the product of a nervous system, the processes underlying those behavioral changes referred to as learning must be the result of some activity or change in the nervous system. The complex human nervous system includes both the central nervous system (the brain and spinal cord) and the peripheral nervous system (the nerves of the body and the autonomic ganglia). It is generally assumed that the changes important in learning occur in the central nervous system, with the peripheral nerves serving a transport function. Within the central nervous system, changes related to learning (the engram) could be limited to one specific region or be dispersed throughout the entire brain and spinal cord. In addition, the learning substrate could be at a subcellular level, within or between individual cells, or among groups of cells at a regional level. *See* NERVOUS SYSTEM (VERTEBRATE).

Only those processes that are known to occur in the brain can be considered candidates for the substrate of learning. For instance, much is known about the mechanics of blood flow and its effects on brain function, but they are unlikely to serve as the basis of learning. However, electrical events, alterations of neural structure, biochemical changes in or around neurons, changes in synapses, and alternations of neurons over large brain regions all seem to be candidates.

Generalizability. Initial observations of the brain's control of learning occurred mainly in humans with various types of brain damage. However, with the advent of new techniques for the study of brain processes, the study of brain function in infrahuman species has become necessary, since many of the techniques are invasive and destructive. Indeed, most current knowledge of brain processes and neural function has come from animal studies, and may or may not apply to humans. When the study subjects are mammals, any processes found there probably will apply to human brain function. However, in studies on animals such as *Aplysia*, an invertebrate sea snail, the parallelism may be more difficult to establish. Such arguments make the assumption that the only goal is to investigate how humans function, and while this is certainly the ultimate goal, it is not the only one. Identifying such processes in these organisms allows for better care of these animals both in captivity and in the wild. In addition, since it is considerably easier to study neural process in simple animals, these studies provide clues about processes that may be important in humans. Obviously, brain processes in humans and other animals differ, but studies of simpler organisms offer clues and information about human learning processes.

Research methods. Attempts to find the neural bases of learning are ultimately aimed at understanding the function of the human central nervous system. However, because many of the techniques used in investigations are inherently destructive to some degree, moral and ethical considerations preclude their use in humans.

Observation and deduction. One of the oldest and still one of the most productive techniques for the early stages of any study is the method of observation and deduction. The subject is put in a learning situation and observed for changes in behavior. Deductions can then be made about the events that take place from what is known about the function of the nervous system. I. Pavlov used that method to develop one of the most widely used learning techniques, classical conditioning, although many of Pavlov's theories about the brain substrates of learning have been shown to be incorrect.

Observation of brain-damaged humans has provided clues about brain areas that are involved in certain learning functions. While this method provides no information about the neural processes underlying various kinds of learning, the damage resulting from illness, closed head and penetrating wounds, and necessary surgical procedures have provided data on which areas of the brain, and especially the cortex, subserve many of the higher human functions. Obviously, it is ethically and morally impossible to undertake controlled lesion or ablation studies on humans, but the uncontrolled brain lesions seen in clinical medicine have been very helpful in guiding controlled studies in animals. With the advent of noninvasive scanning techniques such as positron emission tomography (PET) and computerized tomography (CT) and magnetic resonance imaging (MRI), it is now possible to provide detailed images of the human brain that reveal damaged areas in some detail. *See* MEDICAL IMAGING.

Ablation. The ablation technique involves removal of a part of the brain and observation of the effects after recovery of the subject. The underlying assumption is that removing some part of the nervous system will cause a deficit that accurately reflects the function of the lost part. This technique has been utilized successfully in brain mapping to localize various sensory and motor functions, but localization is only a partial step in characterizing the brain areas responsible for or involved in learning.

Since one brain area may be involved in more than one task, the finding that a cortical area is involved in vision does not preclude its active involvement in other tasks. However, the ablation method assumes that an area is primarily involved in one task and that the function is indeed localized there. Also, when a central brain structure is removed, any subsequent impairment in learning could be due to a loss of sensory or motor ability, a loss of communication links between areas involved in learning, or the destruction of the vital learning area.

The simplest ablations involve removing a brain area. Tissue may also be destroyed by passing an electric current through a needle into the tissue, or by suction, radiation, chemicals, freezing, and stopping blood flow to the region. Reversible ablations, which render the brain tissue temporarily inactive, can be accomplished by using chemicals (such as potassium chloride) or a cold probe.

Following ablation, the animal is allowed to recover and then is given a learning task. The results are compared with those from animals given the same task with no ablation or a different ablation. Performance deficits in the ablated animals are attributed to the loss of the lesioned area. Alternatively, the task may be learned first, the ablation performed, and then the animal tested for loss of learning. Such a design is aimed at localizing the engram or area where the effects of learning are stored, assuming that the storage and processing areas are the same.

Stimulation. The excitation of neural tissue with electric currents or chemical agents is often thought to be the opposite of ablation since one stops activity and the other starts it. However, stimulation rarely produces the opposite results from ablation and rarely has a positive effect on learning. Stimulation of an area can produce some activity that disappears if the area is ablated. The usual result of stimulation is to disrupt or slow learning, so that it can be used as a temporary lesion, to locate sites that lead to disruption of normal learning patterns when stimulated.

Electrical recording. The brain communicates through electrical signals generated by its approximately 20 billion nerve cells. Each cell can generate an electrical impulse and transmit that small electrical charge to other cells, often over substantial distances—3 ft (1 m) in the case of motor neurons that go from the spinal cord to the muscles of the feet in humans. The constant activity of many brain neurons can be recorded by electrodes either as single-cell events, as the activity of small groups of cells, or as the summed activity of millions of cells together. With large electrodes, the activity of cell populations can be recorded directly from the scalp (**Fig. 1**). *See* ELECTROENCEPHALOGRAPHY.

With the development of sensitive amplifiers, single- and multiple-cell recordings became practical. With these techniques, it is possible to implant electrodes in almost any brain area and to record activity from cells for days and often weeks with no discomfort or damage to the animal. To position the electrodes precisely, a stereotaxic instrument (**Fig. 2**) is used to hold the anesthetized animal's head securely and allow the electrode to be lowered into the brain to a precise location. Thus, animals with implanted electrodes can learn tasks while the activity in a brain area is monitored to see if the area participates in the learning. Cellular activity changes that seem to correlate with learning may be observed, but these may be unrelated to the learning of the task itself.

Developmental methods. Increasingly, learning is tested at various stages of human or animal development and the learning possible at that developmental stage is correlated with brain tissue function at the time. As with observational methods, this method gives no

insight into the neural processes underlying learning, but it can give insight into the areas of the brain that subserve various forms of learning. For example, human fetuses late in the last 3 months before birth can learn movements in a Pavlovian motor learning situation through tactile stimuli administered on the mother's stomach over the uterus. Since the fetus can associate two such stimuli, this suggests that such motor learning may not require the active role of the cerebral cortex, which is not developed at that time. Likewise, studying the young of the sea snail *Aplysia*, which can learn some motor tasks, is giving investigators some insight into the parts of the nervous system underlying certain portions of the learning process.

Other methods. Increasingly powerful techniques have been under development for investigating the changes underlying learning. The injection of sensitive dyes and radioactive tracers allows delineation of changes in nerve tracts or chemical alterations. Computer analysis of neural activity or activation of neurons allows the detection of processes underlying the engram formation. Powerful microcomputers can record, stimulate, and even perform ablation studies that were once impossible or too time-consuming. In addition, computers can provide modeling of brain function. However, many of these methods provide clues about which brain areas may be involved in learning rather than the actual process of learning.

Model approaches. Investigators of the neural substrates of learning now use the strategy of operationally defining learning. This means that learning is defined by the operations used to produce it in the experimental situation. Thus, simple associative or motor learning can be defined by the operations of classical conditioning, that is, the presentation of a conditioned stimulus and an unconditioned stimulus in a specified order and time sequence. Because learning is the development of a unique response to two repeated stimuli that are presented at the same time, the behavioral change to be studied can be specified and reproduced precisely.

The use of a simple learning situation is an example of a model approach to the search for neural mechanisms of learning. The investigator can specify the learning task and simplify the complexity of the behavioral changes as well as the neural system to be studied. The human nervous system would be considered the total preparation, with human learning as the total learning system. A nonhuman species for study of a simple learning task and for analysis of the brain processes involved would be a model preparation using a model system.

One of the simplest models for studying the associative learning of Pavlovian conditioning is an invertebrate, such as *Aplysia*. *Aplysia* has no central brain but rather a system of six pairs of ganglia that control its body functions and are interconnected by nerve pathways to allow coordinated activities. In addition, many of the nerve cell bodies are very large (almost visible to the naked eye) and can be identified in each animal. *Aplysia* can learn to associate a light touch to the siphon region with an electrical shock to the tail. The shock causes the animal to quickly and fully withdraw the siphon, whereas the touch does not. Following several paired presentations of the touch (conditioned stimulus) with the shock (unconditioned stimulus), withdrawal to the touch is greatly enhanced. Because that same response does

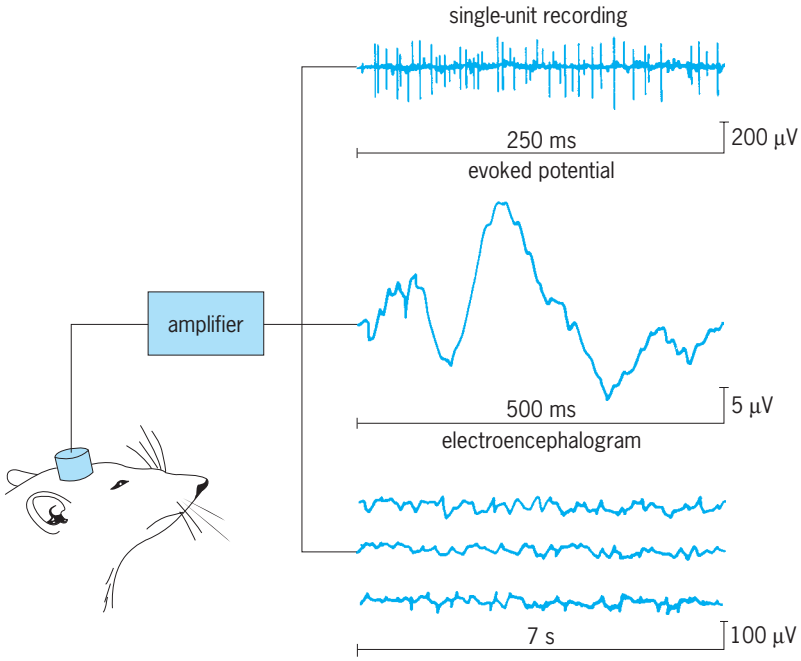

Fig. 1. Brain activity can be recorded in various ways, from single-cell activity (top tracing) to the familiar "brain waves" or electroencephalogram (bottom tracing). (*After C. Cotman and J. McGaugh, Behavioral Neuroscience, Academic Press, 1980*)

Fig. 2. Stereotaxic instrument designed to allow precise placement of electrodes into brain structures. Note the four electrode carriers positioned above the area where the animal's head would be placed. (*Kopf Instrument Co.*)

not take place in animals which receive the same stimuli in other combinations, *Aplysia* can be said to have learned the association. The neural changes underlying its behavior have been located in the abdominal ganglia and seem to involve the activation and alteration of facilitatory neurons that lie between the input sensory neurons and the output motor neurons, and a change in the output of the neurotransmitter from one neuron apparently activates the next neuron in the chain. Such findings cannot automatically be applied to vertebrates and mammals. Mammals have vastly more complex nervous systems and may utilize different neural mechanisms to achieve learning. An understanding of invertebrate learning, however, can offer clues for understanding human learning. *See* NERVOUS SYSTEM (INVERTEBRATE).

Mammalian preparations. To understand learning in the mammalian nervous system, the site at which learning occurs within that system must be found. The simplest mammalian system to be studied is the isolated spinal cord, which is capable of many reflex activities even when separated from the brain. The isolated spinal cord of the cat can alter behaviors when two stimuli are presented to selected sensory inputs in the same way as classical conditioning in the intact animal. Thus, the spinal reflex pathways apparently have the ability to "learn." The neural mechanisms of the change are obscure, but they apparently take place in the small interneurons that connect the sensory inputs with the motor outputs of the spinal cord, and may be alterations in the synaptic junctions between the interneurons. *See* REFLEX.

Classical conditioning has been shown to occur not only in the isolated spinal cord but in cats and rabbits without a cerebrum or cortex as well as in human infants born without a cerebral cortex. The ability of the mammalian nervous system to associate two meaningful stimuli and to change behavior following the association, therefore, appears to be a primitive neural attribute that is probably a basic part of all learning, including human verbal learning. Thus, associative learning appears to occur at various levels of the mammalian nervous system, and the various levels can support different complexities of learning. The total nervous system would, therefore, be necessary for the complexity of human learning. However, the question of the primary location of various learning processes remains.

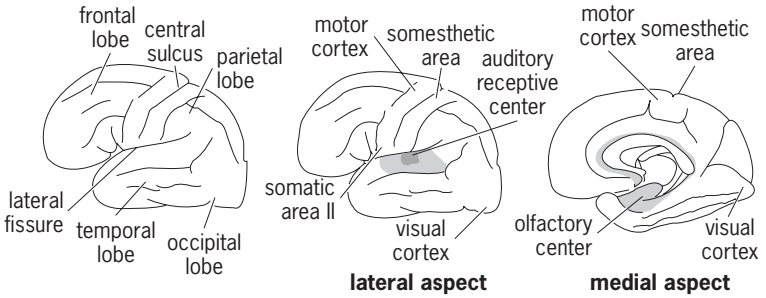

Fig. 3. Diagram of the cortical areas of the human brain. (*After S. W. Ranson and S. L. Clark, Anatomy of the Nervous System, 10th ed., Saunders, 1959*)

Perhaps the most intensive investigation of simple learning in an intact animal has utilized the eye blink response in rabbits. This model system confines the neural system to a specific sensory and motor segment of the very complex rabbit central nervous system. Since the same task is learned in the same way by humans, the learning task can be generalized to humans. By using Pavlovian conditioning, the rabbits learn to blink to a tone sounded just prior to an air puff to the eye. The hippocampus seems to produce a pattern of electrical activity that could be responsible for the learned behavior; unfortunately, the animal can learn the task equally well if the hippocampus is removed. Only when certain complicating factors are added to the simple learning task is the hippocampus found to be necessary for normal learning.

A small area in the cerebellum seems to be necessary for learning the eye blink in the rabbit. When that area is removed, the learned behavior is lost and cannot be regained. The cerebellum has long been considered a strictly motor area involved with coordination, but the learning seen in the eye blink response occurs in the interconnections of two types of neurons in the cerebellum and may involve the alteration of connections between these neurons. Thus, the primary association for learning the eye blink response occurs in the cerebellum. These studies have begun to shed light on the actual mechanisms involved in learning and storage of that learning but have also demonstrated another problem involved in such searches—locating the site of the primary change.

In more complex learning tasks, such as sensory learning, there are various levels of involvement of the higher brain structures. For example, following complete removal of the visual cortex, which subserves vision, a cat or monkey can learn to distinguish lights of different intensities and even speed of movements of visual images. Humans with damage to those visual areas of the cortex can still react to lights in the damaged area of the visual field and can even learn some tasks based on light in that area of the visual field. Some information in the visual sense is, therefore, being processed in other areas of the brain. If the cortex is not completely destroyed, some abilities to learn such things as discrimination will remain but the learning may be impaired, and often learning that had occurred prior to the lesion is lost and must be relearned. When essentially all of the cortex subserving a sense is destroyed, only rudimentary learning can occur, suggesting that some processing of the learning is done at other sites in the brain. **Figure 3** shows some of the important cortical brain areas for sensory information processing and learning in the human.

Observations of humans who have suffered brain damage or who have had brain areas removed surgically have provided information on the primary sites and mechanisms involved in mammalian learning. The most famous, "H. M.," had his hippocampus removed at an early age to control intractable epilepsy. H. M. could remember clearly everything

that happened prior to surgery and had no impairment of intelligence, but after surgery he could no longer remember new facts, demonstrating that the human ability to transfer factual information into long-term memory resides in the hippocampus. However, H. M. could learn certain things. He could learn the rules of a game while he was playing it, but the next day, when asked if he remembered the game he denied ever having seen it. When asked to play, however, he played the game like anyone else who had learned it while still denying ever having seen it. Thus, the learning of the task for verbal or conscious process had not occurred, whereas the learning of motor patterns for playing the game had occurred and been stored. Thus, in humans the learning of a complex task such as a game is divided into at least two parts, one conscious and the other motor. Similar reports of human learning deficits have come from other brain-damaged humans who cannot remember learning anything following the damage but can relearn some sensory or motor tasks. The learning mechanisms are therefore intact for simpler forms of learning, but the results of that learning cannot be brought to consciousness. In addition, evidence suggests that other learning systems in the human are similarly separate. In one case, a brain-damaged accident victim could recall facts about the world and could learn new ones, but he could not remember anything about himself as a participant, even though he remembered who he was. This suggests that in the human, certain brain systems (probably in the frontal lobes of the human brain) may have evolved to support the complex learning of self-awareness and language, which is characteristic of humans.

Theories. There have been several theories formed about the actual mechanisms underlying learning process. These theories have often followed current technology, and all have been simplistic.

Electrical activity. The earliest theories proposed that learning was an electrical circuit set up in some localized area of the brain that remained active, thus explaining forgetting and other aspects of memory. However, studies of cooling the brain to a point at which there was no electrical activity or giving electrical shocks to the brain that disrupted any ongoing activity showed that the memory of learned tasks survived such treatment.

Following this, it was proposed that the initial form of learning was stored as an electrical event that was transformed over several minutes into a more permanent biochemical change in one or more brain neurons. Such notions were modeled after the growing complexity of the telephone system in the 1930s and 1940s. Learning was likened to a switchboard, with changes in individual neurons resulting from the learning activity and producing functional switchings and memories that could be recalled by the appropriate stimuli. Such "one neuron–one memory" theories had some appeal, but they failed to explain how many neurons could be lost from the brain with no apparent loss of memory for learned tasks. In addition, it became increasingly difficult to account for complex learning with the activity of one neuron.

Learning molecules. The deoxyribonucleic acid (DNA) hypotheses of the 1950s and 1960s popularized the theory of learning molecules. If learning was a change in some chemical in the brain or individual neuron, a stable chemical had to be found to account for the stability of the engram. The only sufficiently stable chemical was DNA. Further support for the DNA basis for learning came from studies purporting to show that *Planaria* flatworms that were fed other *Planaria* that had learned a simple task seemed to learn that same task by ingesting the trained flat worms. The hypothesis of learning molecules gained popularity because it suggested the feasibility of learning by pill. However, the effects of such learning molecules, if any, have been shown to be a general one due to enriched diet, and not a specific effect of transferring molecules.

Growth theories. Evidence for a similar theory has come from studies of animals raised in different laboratory environments and given various tasks to learn. Under appropriate conditions, animals in "enriched" environments, that is, with many colors, toys, and textures, have heavier brains than similar animals raised in "poor" environments. Similarly, animals given complex motor tasks to learn may develop larger brain areas associated with motor and learning function. In these studies, however, it is difficult to separate the effects of motor activity from learning itself. The growth theory of learning states that neurons involved in learning grow larger or have more neural processes when compared with noninvolved neurons. How that would occur is not certain, but using the analogy of a muscle, a neuron would enlarge with exercise and shrink with disuse.

Computer analogies. Several theories of learning have drawn analogies between the brain and computer function. These theories postulate that the brain has a central processing unit, or learning center, and that memories are stored in another location to be called up at will. In fact, the human brain does not function like a computer but rather has a vast ability to grasp general concepts and rather limited processing ability.

Activity patterns. A global theory of learning process, drawn from several earlier theories, is based on a probabilistic hypothesis in which an incoming stimulus sets up generalized patterns of activity in some brain area that is capable of causing some specific pattern of motor activity. If the incoming stimulus or a similar stimulus recurs, the same pattern of activity occurs over the same brain area and the learned behavior is executed. The theory depends on the summed performance of millions of neurons. While such a theory is elegant and accounts for many features of learning and memory, it does not specify how a stimulus would create such an activity pattern, what changes would allow that pattern to be stored or recreated by other stimuli, or how the patterns would be translated into motor output. For complex learning, the concept of neural activity pattern changes seems to best account

for human learning. *See* PROBLEM SOLVING (PSYCHOLOGY). Michael M. Patterson

Bibliography. F. E. Bloom and A. Lazerson, *Brain, Mind and Behavior*, 3d ed., 2000; P. M. Groves and G. V. Rebec, *Introduction to Biological Psychology*, 4th ed., 1992; B. Kolb and I. Q. Whishaw, *Fundamentals of Human Neuropsychology*, 3d ed., 1989; C. F. Levinthol, *Introduction to Physiological Psychology*, 4th ed., 1995; R. Ornstein, L. Carstensen, and S. Patnoe, *Psychology, the Study of Human Experience*, 3d ed., 1990; E. Tulving, Remembering and knowing the past, *Amer. Sci.*, 77:361–367, July-August 1989.

Least-action principle

Like Hamilton's principle, the principle of least action is a variational statement that forms a basis from which the equations of motion of a classical dynamical system may be deduced. Consider a mechanical system described by coordinates $q_1, \ldots, q_f$ and their canonically conjugate momenta $p_1, \ldots, p_f$. The action S associated with a segment of the trajectory of the system is defined by Eq. (1), where the inte-

$$S = \int_c \sum_j p_j dq_j \tag{1}$$

gral is evaluated along the given segment c of the trajectory. The action is of interest only when the total energy E is conserved. The principle of least action states that the trajectory of the system is that path which makes the value of S stationary relative to nearby paths between the same configurations and for which the energy has the same constant value. The principle is misnamed, as only the stationary property is required. It is a minimum principle for sufficiently short but finite segments of the trajectory. *See* HAMILTON'S EQUATIONS OF MOTION; HAMILTON'S PRINCIPLE; MINIMAL PRINCIPLES.

Assume that Eq. (2) holds, where $p_j + \delta p_j$ is canon-

$$S + \Delta S = \int_c \sum_j (p_j + \delta p_j) d(q_j + \delta q_j) \tag{2}$$

ically conjugate to $q_j + \delta q_j$. Neglecting second order term in Eq. (2), one obtains Eq. (3), where an inte-

$$\Delta S = \int_c \sum_j (p_j\, d\,\delta q_j + \delta p_j\, dq_j)$$

$$= \int_c \sum_j (\delta p_j\, dq_j - \delta q_j\, dp_j) \tag{3}$$

gration by parts has been made, the integrated parts vanishing.

The vanishing of ΔS requires the integrand to be a perfect differential of a quantity whose end variations vanish. The coefficients of the variations $\delta q_j, \delta p_j$ need not vanish separately because the variations are not independent, the varied q's and p's necessarily being canonically conjugate, as in Eq. (4), with terms defined by Eqs. (5).

$$\sum_j (\delta p_j\, dq_j - \delta q_j\, dp_j) = dU(q, p) \tag{4}$$

$$\delta p_j = \frac{\partial U}{\partial q_j} \qquad \delta q_j = -\frac{\partial U}{\partial p_j} \tag{5}$$

Writing $U = -H\delta t$ leads to Hamilton's equations of motion, Eqs. (6). The quantity $H(q, p)$, known as

$$\dot{p}_j = -\frac{\partial H}{\partial q_j} \qquad \dot{q}_j = \frac{\partial H}{\partial p_j} \tag{6}$$

the hamiltonian function, does not contain the time explicitly because $U(q, p)$ cannot be a function of the time since the end times are not fixed and in general will vary as the path is varied. Thus, the principle is useful only for conservative systems, where H is constant.

If $H(q, p)$ consists of a part H_2 quadratic in the momenta and a part H_0 independent of the momenta, then Eq. (7) holds by Euler's theorem on homoge-

$$S = \int_{t_1}^{t_2} \sum_j p_j \dot{q}_j\, dt$$

$$= \int_{t_1}^{t_2} \sum_j p_j \frac{\partial H}{\partial p_j}\, dt$$

$$= 2\int_{t_1}^{t_2} H_2\, dt \tag{7}$$

neous functions. Usually H_2 is the kinetic energy of the system so that the principle of least action may be written as Eq. (8), where V is the potential energy.

$$\Delta \int_{t_1}^{t_2} 2T\, dt = \Delta \int_{t_1}^{t_2} 2(E - V)\, dt = 0 \tag{8}$$

The principle of least action derives much importance from the fact that it is the action which is quantized in the quantum form of the theory. Planck's constant is the quantum of action. *See* NONRELATIVISTIC QUANTUM THEORY. Philip M. Stehle

Least-squares method

A method, developed originally by A. M. Legendre, of obtaining the best values (the ones with least error) of unknown quantities supposed to satisfy a system of linear equations of the form shown as notation (1),

$$\begin{aligned} M_{11}a_1 + M_{12}a_2 + \cdots + M_{1m}a_m &= b_1 \\ M_{21}a_1 + M_{22}a_2 + \cdots + M_{2m}a_m &= b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ M_{n1}a_1 + M_{n2}a_2 + \cdots + M_{nm}a_m &= b_n \end{aligned} \tag{1}$$

where $n > m$. Since there are more equations than unknowns, the system is said to be overdetermined. Furthermore, the values obtained for the unknowns by solving a given selection, m in number, of the

equations will differ from the values obtained by solving another selection of equations. In the physical situation, the b_i are measured quantities, the M_{ij} are known (or assumed) quantities, and the a_i are to be adjusted to their best values.

Consider a simple example. A quantity y of interest is supposed (perhaps for theoretical reasons) to be a linear function of an independent variable x. For a series of selected values $x_1, x_2, \ldots$ of x the values $y_1, y_2, \ldots$ of y are measured. The expected relation is shown as notation (2), and the problem is to find the

$$\begin{aligned} x_1\alpha + \beta &= y_1 \\ x_2\alpha + \beta &= y_2 \\ x_3\alpha + \beta &= y_3 \\ &\cdots\cdots\cdots \end{aligned} \tag{2}$$

best values of α and β that is, respectively, the slope and intercept of the line which graphically represents the function. The best values of α and β, in the least squares sense, are obtained by writing Eq. (3) and asserting that term (4) shall be minimized with

$$\eta_i = y_i - (x_i\alpha + \beta) \tag{3}$$

$$\sum_{i=1}^{n} \eta_i^2 \tag{4}$$

respect to α and β, that is, Eqs. (5) hold. This leads to Eqs. (6) and (7), which may be solved for α and

$$\begin{aligned} \frac{\partial}{\partial\alpha} \sum_{i=1}^{n} \eta_1^2 &= 0 \\ \frac{\partial}{\partial\beta} \sum_{i=1}^{n} \eta_i^2 &= 0 \end{aligned} \tag{5}$$

$$\alpha \sum_{i=1}^{n} x_i + n\beta - \sum_{i=1}^{n} y_i = 0 \tag{6}$$

$$\alpha \sum_{i=1}^{n} x_i{}^2 + \beta \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i y_i = 0 \tag{7}$$

β. For m, rather than two, unknowns the generalization is obvious in principle, although the labor of solution may be great if m is large unless a high-speed electronic computer is available.

It should be noted that the measurements y_i in the example have all been assumed to be equally good. If it is known that the measurements are of variable quality, a weight may be attached to each value of y_i. The least squares equations are readily modified to take this into account, as in Eqs. (8), where w_i is the

$$\begin{aligned} \alpha \sum_{i=1}^{n} w_i x_i + \beta \sum_{i=1}^{n} w_i - \sum_{i=1}^{n} w_i y_i &= 0 \\ \alpha \sum_{i=1}^{n} w_i x_i{}^2 + \beta \sum_{i=1}^{n} w_i x_i - \sum_{i=1}^{n} w_i x_i y_i &= 0 \end{aligned} \tag{8}$$

weight of measurement y_i.

The least squares equations can be shown to lead to the most probable (in the statistical sense) values of the unknowns under a variety of assumptions about the measurements and their weights. In applications in the physical sciences, however, it is rarely possible to show that one's observations satisfy all, or even any, of the assumptions. However, the conditions may be approximately satisfied in many instances, and the method is widely used because of its convenience. The empirical result is that the unknowns so determined lead to excellent representations of the data in the usual case. *See* CURVE FITTING.

McAllister H. Hull, Jr.

Bibliography. R. W. Farebrother, *Least Squares Computations*, 1988; J. W. Longley, *Least Squares Computations*, 1984.

Leather and fur processing

The technology of processing animal hides and skins. While both leather and fur are derived from animal skins, there are essential differences in the processing techniques.

Leather

Leather processing provides a means of tailoring the unique fibrous architecture of animal skins into materials that possess those attributes for which finished leather is prized: a certain texture (feel) and appearance as well as characteristic physical and chemical properties. The raw materials, primarily skins of cattle, goats, hogs, and sheep, are actually by-products of the many worldwide meat production industries. Articles manufactured from leather include footwear, clothing, gloves, luggage, upholstery, harnesses and saddlery, and mechanical devices.

Raw stock preservation. Since skins and hides putrefy rapidly after slaughter, some immediate treatment is usually necessary. While some fresh skins and hides may undergo the tannage conversion processes immediately, common salt (NaCl) is usually employed as a temporary preservative, either as a dry salt pack (**Fig. 1**) or as a saturated brine solution in an oval circulated vat (raceway) or in rotary vessels.

In the United States, hides from cattle are most frequently treated by the agitated brine curing process.

Fig. 1. Cattlehides being cured with salt in a pack containing about 800 hides.

After the hide is removed from the carcass, it is washed in cool water to lower its temperature and to remove blood and dirt. The flesh that remains on the inner side of the hide along with any manure and other debris is often removed by machine, either before or after brine curing. The brine in the raceway is maintained as close as possible to the saturation point (preferably above 90% saturation) throughout the curing process, which extends over approximately 24 h.

When properly cured, the brined hide will have a water content ranging from 40 to 48%, and the retained water will contain NaCl in excess of the 85% saturation level. Since salt-tolerant (halophilic) bacteria may contaminate recycled brines, a bactericide is added to the brine, particularly when ambient temperatures are relatively high.

Hides that have been cured with either dry salt or brine may be stored for several months, especially in refrigerated storage. This temporary stability facilitates distribution of the cured hides.

The epidermal tissue in the grain area of a typical hide (**Fig. 2**) includes the outer epidermal layers and the associated hair follicles, oil glands, sweat glands, and erector pili muscles. Collagen, a unique triple-helical biopolymer, is the leather-making hide substance; it predominates in the hide corium and is a major proteinaceous component in the grain enamel and fibrous grain layer that surrounds the epidermal tissues. Grain enamel is the extremely thin outer surface cellular layer of a skin or hide; this layer has a unique structural organization and collagen composition. The unique three-dimensional organization of the biopolymers is responsible for the strength, toughness, and flexibility of the final leather. *See* COLLAGEN; SKIN.

The finest leather is produced from skins that possess the proper grain enamel and hair or wool follicle patterns that permit the production of a unique surface appearance in the final product. Skins or hides whose grain quality is inadequate may be finished

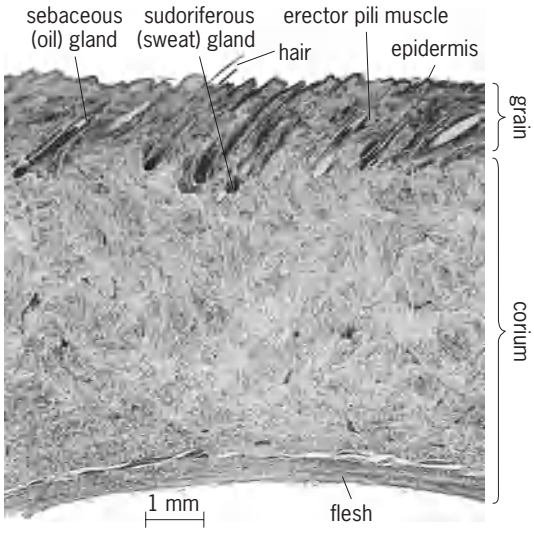

Fig. 2. Cross section through cattlehide. 1 mm = 0.04 in.

into products known as corrected-grain (a leather whose grain surface has been abraded to remove or obscure defects) or sueded leathers. Heavier hides may be mechanically split into two layers; the inner layer or split is usually converted to sueded split leather. These grade/type designations, including species of origin and grain integrity, are protected by fair trade rules.

Beamhouse operations. The preserved dry skins or hides are rehydrated, and the curing salt is removed by soaking in a paddle vat, a rotary drum, or hide processor, a process requiring up to 12 h. Bactericides are added when the soaking process is longer or if warm water is used.

The rehydrated skins or hides are then processed in a saturated lime [$Ca(OH)_2$] solution with caustic sharpeners (alkaline additives), usually sodium sulfide or sodium sulfhydrate. This treatment may either loosen the hair and hair follicles for mechanical removal or pulp the keratinous epidermal tissue so that it may be washed off. Heavy hides may be processed with recovery of salable by-product hair in a 5-day vat process; this is used especially for production of sole leather. Processors containing material with lower ratios of solution to hide by weight and higher concentrations of sharpeners can pulp the hair in a few hours. However, this may require a few hours of additional treatment with lime solution to open up the hide corium to permit proper processing subsequently.

Sheepskins are dewooled by painting a thick chemical suspension for hair removal onto the flesh side. Raw material for leathers that retain the original hair or wool, including sheepskin shearlings, is washed carefully and rewetted to preserve the wool or hair in the follicle.

Chromium tannage. The beamhouse processes remove keratinous tissues from the hide and initiate the preparation of collagenous fibrous substances in the hide for tannage. Further conditioning of the grain (bating) is accomplished with enzymes; these are usually pancreatic enzymes, although sometimes bacterial or fungal enzymes are used. Deliming salts, most frequently ammonium sulfate or ammonium chloride, are employed to adjust the pH of the bath to the range of 8 to 9. *See* PH.

Chromium tannage requires an acidic pickling process. The pickling solution is a sodium chloride brine solution (using water approximately equal in weight to the hide weight and using NaCl equal to 3% of the weight of the water used) and sufficient sulfuric acid to create an acidity of pH 2–3 in the final pickling bath. Pickling, in addition to being a tannage process step, is also used for temporary storage of the skins or hides. This is an important treatment for sheepskins after wool pulling; pickled sheepskins are shipped from New Zealand to worldwide markets.

The tanning material is essentially a complex containing one-third basic chromic sulfate ($CrOHSO_4$). The uptake of the trivalent chromium by the hide substance is moderated by using so-called masking chemicals, often sodium formate. The adsorbed

chromium may be fixed onto the hide collagen by pH adjustment (to approximately pH 3.8–4.0), with sodium bicarbonate being a common basification agent.

A characteristic manifestation of chromium tannage is the increase in the hydrothermal shrinkage temperature of the hide substance from around 149°F (65°C) to temperatures exceeding 212°F (100°C). The amount of the chromium tanning complex that is fixed varies, but it usually requires more than 2% fixation (expressed as the Cr ion) on the protein basis. The tanning complex is fixed by collagenous carboxyl groups, partially forming crosslinks in the collagenous molecule. The entire beamhouse and chromium tannage process for cattlehides may be accomplished in a daily cycle; however, a total time of 48 h for both processes often is preferable.

Since hides may be very greasy, degreasing agents may be utilized in the beamhouse and tannage processes. These are often aqueous solutions; however, pickled sheepskins are commonly degreased with hydrocarbon solvents.

The chromium-tanned stock is blue and is often known as chrome-in-the-blue stock. It is an article of international commerce, and is shipped in the wet stage. While the wet stock is resistant to bacterial deterioration, fungal growth can occur during storage. Hence, fungicides are commonly added to prevent damage. *See* FUNGISTAT AND FUNGICIDE.

Retannage, coloring, and fat-liquoring. The wet chrome-in-the-blue tanned hides are sorted to select the raw material most appropriate for production of the more valuable full-grain leathers. The sorting also groups the stock by hide thickness and surface area to permit the optimal production of the final grain leather thickness and types. The hides are split mechanically, producing a grain layer and a so-called blue drop. After trimming, this inner split layer is usually converted into suede leathers.

Retannage processes may include a second trivalent chromium tannage, a natural vegetable tannage, or a variety of organic synthetic tannages. These combined tannage processes permit the tanner to modify the ionic charges on the tanned collagen fibers facilitating the next two steps in this continuous trio of processes. Proper retannage procedures also facilitate the surface buffing process on the dried leather, which is required for the production of corrected-grain leathers.

Acidic, basic, direct, and reactive dyestuffs produce the colors required for the many fashion markets served by leather. The colorants are selected and applied either to produce surface color effects or to produce leather with complete cross-sectional color penetration. The levels of consumption of the dyes may vary from 1 to 10% by weight of the wet stock, particularly in suede leathers. *See* DYEING.

Fat-liquoring is a lubrication process that uses oil emulsions. It is designed to produce leathers with the suppleness, temper, and strength needed for the varied end uses. The lubrication oils are most frequently natural oils, of animal (especially marine) and vegetable origin; however, synthetic lubricants and mineral oils are also utilized to produce desired effects in the final leather. The emulsions may have either anionic, cationic, or nonionic properties that impart different characteristics to the leather. Sulfated oils, bisulfited oils, and soap emulsions are frequently used in commercial fat-liquoring systems. Levels of lubricant additives vary from a few percent to greater than 10% in oily or so-called stuffed leathers. *See* FAT AND OIL.

Retannage, coloring, and fat-liquoring processes are continuous operations carried out in rotary drums. They are accomplished in a few hours, and they are often carried out in warm water to facilitate the process. A tannery may have available several hundred different formulas for these operations in order to produce a wide enough variety of leathers to meet consumer demand.

Drying and finishing. The development of the final leather characteristics continues in the tannery processes that complete the conversion of skins or hides to leathers for varied commercial end uses. The first step is the removal of the retained process water, so that the dried leather contains about 12% moisture.

Drying processes are designed to produce leathers with the desired feel, surface levelness, and area (these leathers are sold by the square foot). Very soft glove or garment leathers may be dried by slower processes during a hanging period in a drying room with controlled air flow maintained at moderate temperatures. In toggle drying processes, the moist leather is attached to specially designed frames prior to passage through a drying tunnel; this drying process yields dry leather with the desired suppleness and with somewhat increased area yield. Firmer leathers with greater area yield may be affixed upon a drying plate coated with a temporary adhesive in a unit known as a pasted leather tunnel dryer. This process often employs higher drying temperatures and is most often used to produce leather for shoe uppers. Other methods used are vacuum drying processes, radio-frequency drying processes, and dehumidification drying processes.

The selection of the drying process is determined by the desired final leather characteristics and leather area yield. There has been considerable development of processes and machinery involving this step in the tannery process.

Leather suppleness may be further modified after drying by dry milling in a rotary drum or by mechanical flexing processes. Furthermore, dried raw material whose grain quality is inadequate is buffed with selected abrasive papers to modify the grain enamel surface prior to finishing as corrected-grain leathers.

The surface luster of the leather is developed by the application of finishing coats to the grain of the leather. These finishing formulas utilize a film-former binder, a colorant, a carrier, and additives that promote finish coat adhesion to the leather surface to achieve desired leather surface characteristics. The film-forming binder polymers may include proteins, nitrocellulose, acrylics, or polyurethanes.

The colorants are either dyes or pigments, especially on corrected-grain leathers. Water is the basic carrier of the finishing formulation, although other solvents may be used, especially in lacquer topcoats.

The finishing formulations are applied with rotary sprayers or roll coaters, with the skins or hides moving on conveyors. The carrier is removed by evaporation in heated tunnels. These finishing coat applications, which are usually repetitive, are accompanied by mechanical plating operations to facilitate the development of the desired surface luster and appearance. The finished leather is sorted by estimated cutting yield, and the surface area is measured. Packaging and shipment of the finished leather to manufacturers of leather products completes the chromium tannage conversion process.

Vegetable tannage. Vegetable materials are the initial tanning materials in approximately 10% of leather production in the United States. Vegetable tanning materials are polyphenolic extracts of barks and woods. The major commercial extracts are quebracho, wattle (mimosa), and chestnut; these materials are all imported from South America, South Africa, and the Mediterranean countries. Vegetable-tanned leathers in the United States are primarily heavy cattlehide leathers produced for shoe soles and insoles, belting, strap, and specialty leathers. *See* WOOD CHEMICALS.

Vegetable tannery processes are specialized adaptations of the general tannery processes previously discussed. The vegetable-tanned sole leather process utilizes a special hide segmentation process (**Fig. 3**). After the beamhouse operation, the bellies are pickled and used for chrome-tanned leathers for work gloves, and the heads may be used for dog bone production. The shoulders are used for specialty leathers, including waist belt and strap leathers. The bends are converted to shoe sole and insole leathers.

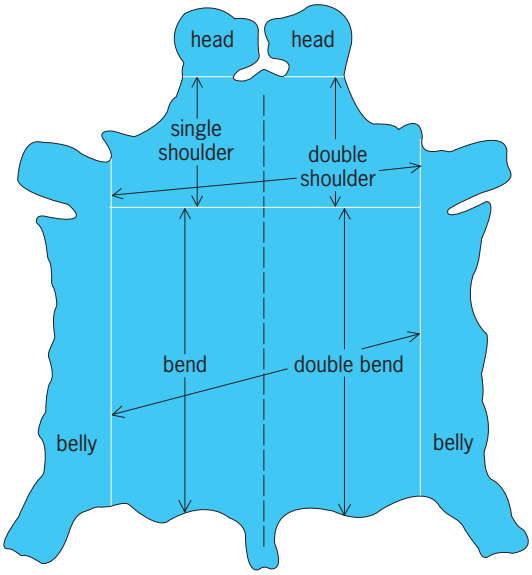

Fig. 3. Sole leather from cattlehide.

The original vegetable tannage process was a 2- to 3-week process in which the hide was sequentially immersed in a series of rocker vats containing progressively increasing tannin strength. This process resulted in incomplete tannin utilization and caused wastewater problems. A more recently developed method uses an initial polymeric phosphate and sulfuric acid pickling bath to create the acidic pH for tannage and to protect the grain of the hide. Vegetable tannage is then carried out in a concentrated tannin solution in a heated rocker-circulator system. This system provides better utilization of the vegetable tannin blend and produces the desired 50–65% combined tannin, expressed on the collagenous protein content of the leather.

The characteristics of solidity and compactness that are necessary to produce good soling leathers are developed further by using a rotary drum known as an oil wheel to apply oils and inorganic salts (for example, epsom salts). The leather is dried very slowly in hanging lofts, and then it is compacted by high-pressure rolling. The leather is not dyed; the grain surface is polished by glazing with a wax emulsion applied to the grain side of the bend.

Fur

The earliest type of clothing for primitive humans was fur, prized for warmth in cold climates and for exotic beauty, color, and texture in warmer climates. Furs possess the unique characteristics of the keratinous epidermal portion of the skin, with the collagenous corium providing the supporting fibrous matrix that enables fur to be used as a clothing material.

Raw materials. The sources of the raw materials are wild and domestic animals. Trappers provide the wild animal skins, which must be preserved under field conditions, sometimes resulting in poor quality. When animals that are bred and raised in captivity, such as mink and rabbit, are produced primarily for their fur skins, better quality control is achieved.

Animals whose skins are used for fur include beaver, fox, mink, muskrat, rabbit, sheep, and seal. Mink represents the highest-dollar-volume fur in worldwide trade. Woolskins from domestic sheep, in addition to being the source of shearlings for garments and footwear, are specially processed to produce mouton fur. The skin of the newborn lamb of the hardy fat-tailed karakul sheep is the source of a major commercial fur. Mink are bred to produce special color styles. Since rabbit skins are numerous and inexpensive, special breeds have also been developed to permit production of rabbit furs with varied colors and textures.

Fur skins possess characteristic colors and color patterns. The natural fur pigmentation is determined by the presence of melanin, an oxidized amino acid derivative that is synthesized by specialized epidermal cells.

The morphology of the epidermal system in fur skins is more complex than that found in raw materials for leather making such as calfskins. The predominant fur of the pelt contains longer hair fibers underlain by a dense, fine, soft undercoat. Some species

also have coarser specialized hairs, including guard hairs.

Processing. The fur dresser rehydrates the fur skins carefully to avoid damage to the expensive raw material. Chemical additives are avoided to protect the integrity of the fur. Mechanical action is minimized, especially on fine, thinner skins. The rehydrated, softened skin is then fleshed, using special machines. Careful fleshing is a necessity, since dressing agents must penetrate from the flesh side of the corium and mechanical action must be avoided.

Tannage, or dressing, of fur skins may employ processes related to those used in leather tannage. Since garments fashioned of fur are seldom exposed to heavy rain and are not laundered, pickled skins containing salt may be merely lubricated prior to drying and finishing. However, alum or chromium tannage processes and aldehyde (formaldehyde or glutaraldehyde) tannages may be employed to stabilize the corium protein further, providing increased resistance to water and heat. Excess fixation of chromium onto the fur protein can lead to color variations in the fur. Chromium tannages may also interact with bleaching processes and dyeing procedures involving oxidation.

Oiling of fur skins is accomplished by applying light oils with low viscosity to the flesh side of the skin. Fat-liquoring operations using oil emulsions in rotary drums would produce fur entanglement and wool felting. Heavier skins may have the oil worked into the skin corium in a device known as a kicker tub; unsaturated oils may be used in this process for lubrication and to produce an oxidized oil tannage. Adjustment of the fur color to produce fashionable colors is a complex process; its use is regulated to assure correctness of labeling.

In fur processing, the hydrothermal stabilization of a skin substance is often low; thus removal of excess water requires care. Centrifugal processes may be used, but the corium retains moisture more tenaciously than the fur fibers. Hanging the fur skins to dry under controlled conditions of temperature and humidity in order to avoid damage completes the process. However, if the corium has been dried to a 15% moisture content, the fur will often be overdried, with low flexibility and excess static electric charge. Agitation in a drum containing sawdust is used to equalize moisture content in the corium and fur and to dryclean the fur. After sawdust removal and adjustment of the corium moisture to about 25%, careful stretching or staking develops skin softness.

Selective removal of guard hairs and shearing are special mechanical processes that may be used. The fur gloss is developed by ironing in rotary ironing machines.

Robert M. Lollar

Bibliography. K. Bienkiewicz, *Physical Chemistry of Leather Making*, 1983; F. O'Flaherty, W. T. Roddy, and R. M. Lollar, *The Chemistry and Technology of Leather Manufacturing*, Amer. Chem. Soc. Monogr. 134, 4 vols., 1956–1965, reprint 1978; G. C. Moseley, *Leather Goods Manufacture: Methods and Processes*, 1986; J. H. Sharphouse, *Leather Technician's Handbook*, 1983; T.C. Thorstensen, *Practical Leather Technology*, 4th ed., 1993.

Lecanicephaloidea

An order of tapeworms of the subclass Cestoda. All species are intestinal parasites of elasmobranch fishes. These tapeworms are distinguished by having a peculiar scolex divided into two portions. The lower portion is collarlike and bears four small suckers; the upper portion may be discoid or tentacle-bearing and is provided with glandular structures (see **illus.**). The scolex is usually buried in the intestinal wall of the host and may produce local pathology. The anatomy of the segments is very similar to that of the Proteocephaloidea. Some authorities place the lecanicephaloids in the order Tetraphyllidea, to which they are obviously closely related. Essentially nothing is known of the life history of lecanicephaloids. *See* EUCESTODA; TETRAPHYLLIDEA.

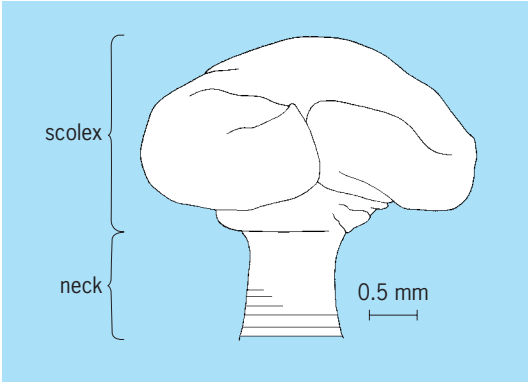

Anterior end of a lecanicephaloid tapeworm.

Clark P. Read

Lecanorales

An order of the Ascolichenes, also known as the Discolichenes. Lecanorales is the largest and most typical order of lichens and parallels closely the fungal order Helotiales. The apothecia are open and discoid, with a typical hymenium and hypothecium. There are four growth forms—crustose, squamulose, foliose, and fruticose—all showing greater variability than any other order of lichens.

Reproduction. Details of sexual reproduction are at best poorly known. Meiotic and mitotic stages leading to spore formation seem normal, but earlier stages of fertilization have not been studied. It is not known if fertilization is somatogamous, or if the microconidia act as spermatia and fertilize the egg in the ascogonium. Mature spores burst the ascal wall and are ejaculated from the apothecial disk to heights of several centimeters. When the spores are

germinated in pure culture, a distinct though slow-growing mycelium is often formed. The mycelium has no resemblance to the growth form of the original parent lichen, nor have apothecia or pycnidia ever been seen in culture. Furthermore, it is doubtful whether anyone has ever synthesized a lichen by combining a fungal culture with an alga. In nature vegetative reproduction is highly developed. Isidia, minute coralloid outgrowths of the cortex, and soredia, powdery excrescences from the medulla, are present in many species. They apparently grow into new lichens when dislodged from the thallus. *See* REPRODUCTION (PLANT).

Taxonomy. The Lecanorales is divided into 25 families, about 160 genera, and 8000–10,000 species. Family divisions are based on growth form of the thallus, structure of the apothecia, the species of symbiotic algae present, and spore characters. Species are separated by such characters as isidia, soredia, rhizines, and pores, and by chemistry. The larger families are described below.

Cladoniaceae. This family includes the reindeer mosses and cup lichens. The main thallus is a hollow structure in the shape of a stalk, cup, or is richly branched. There are about 200 species of *Cladonia*; the most highly developed are found in the boreal and arctic regions.

Lecanoraceae. The thallus is typically crustose, and the apothecia have a distinct thalloid rim. This family is primarily temperate and boreal. The largest genus, *Lecanora*, has over 500 species.

Lecideaceae. This family differs from the Lecanoraceae in lacking a thalloid rim around the apothecia. It is also very common in temperate and boreal zones and includes one of the largest and most difficult lichen genera to diagnose, *Lecidea*, with over 1000 species.

Parmeliaceae. This family, known as the foliose shield lichens, is common on trees throughout the world. *Parmelia* has almost 1000 species; *Cetraria*, a common genus on pine trees, includes the well-known Iceland moss (*Cetraria islandica*).

Umbilicariaceae. The thallus is circular in outline, quite large, and umbilicate. The family occurs on acidic rocks in temperate, boreal, and arctic regions. The members are known as the rock tripes.

Usneaceae. The beard lichens comprise perhaps the outstanding lichen family because of their conspicuous fruticose growth form. The largest genus is *Usnea*, some species of which grow to 5 ft (1.5 m) in length.

Mason E. Hale

Lectins

A class of proteins of nonimmune origin that bind carbohydrates reversibly and noncovalently without inducing any change in the carbohydrate. Since their discovery in 1888, lectins have been labeled agglutinins, antibodylike molecules, hemagglutinins, heteroagglutinins, heterophile agglutinins, natural or normal antibodies, protectins, phytohemagglutinins, and receptor-specific proteins, among other names. The main feature of lectins is their ability to recognize and bind specific carbohydrate structures becoming the translators of the sugar code (the coding of biological information by sugar structures). Lectins bind a variety of cells having cell-surface glycoproteins (carbohydrate-bound proteins) or glycolipids (carbohydrate-bound lipids). The presence of two or more binding sites for each lectin molecule allows the agglutination (clumping) of many cell types, and the agglutination reaction has been used extensively to detect the presence of lectins in extracts from different organisms. However, multivalency (having several sites of attachment) may not be an absolute requirement, even though it is still an important factor for most lectins. Lectins are ubiquitous in nature and may have very different roles according to the organism, tissue, or developmental stage. Their potential ligands, simple or complex carbohydrates, are present in all living cells and in biological fluids, which suggests that protein–carbohydrate interactions constitute basic phenomena common to all organisms. *See* CARBOHYDRATE; PROTEIN.

Distribution. Lectins have been identified in species of virtually all taxa, from viruses and bacteria to vertebrates, and they occur in both prokaryotic and eukaryotic organisms.

Agglutinating activity has been found in more than 1000 plant taxa. At present, around 1000 sequences have been determined for 241 different plants. Moreover, most of the best-characterized lectins come from a single family, Leguminosae, and they constitute about 60% of the well-known lectins. Lectins have been found in almost every plant tissue, but their distribution in different organs varies within the various families already studied. They are abundant in seeds (Leguminosae, Euphorbiaceae, Gramineae), fruits, tubers and bulbs (Solanaceae, Liliaceae), roots (Cucurbitaceae, Leguminosae), and stems and leaves (Cactaceae, Orchidaceae). At the subcellular level, they are present in plant cell walls of almost all tissues, and may have a fundamental role in the assembly of cell-wall polysaccharides and glycoproteins. *See* PLANT METABOLISM.

Recent years have seen an explosion of interest and research in animal lectins. Good examples are galectins (galactose-binding lectins); calcium-dependent (C-type) lectins, which include selectins (membrane-bound lectins specific for leukocyte adhesion) and collectins (mannose-specific lectins); and annexin proteins with affinity for lipids and sugars present in almost all organisms. Vertebrate lectins occur as soluble or integral membrane proteins in embryonic and adult fluids, organs, and tissues. Invertebrate lectins occur in body fluids or secretions, such as fish serum, snake venom, seminal and coelomic fluids, and hemolymph.

Microbial lectins have been isolated mainly from bacteria, but they are also found in viruses, slime molds, protozoa, green algae, and fungi. In many cases the microbial lectins are bound to the bacterial surface and have only one recognition site. The bacteria have the ability to agglutinate other cells, but the isolated lectin does not

and is referred to as lectinlike protein.

Initially, lectins were classified according to their saccharide specificity. However, the introduction of the techniques of molecular biology has changed lectin classification. The homology in the structure of the carbohydrate recognition domain allowing for the drawing of genealogical trees has replaced this criterion. Lectins are classified based on their amino acid sequence homology and relatedness. The crystal structures of more than 500 lectins of different origin have been determined, and a very complete database is available with the known three-dimensional structures.

Molecular properties. The molecular structures of lectins from different organisms show no general pattern. This is not surprising if one considers that plants and animals have undergone long evolutionary divergence. Carbohydrate-binding molecules appear to have arisen at different times and are products of very different evolutionary histories.

Lectins can be monomeric or oligomeric. The latter are usually composed of two to four polypeptide chains and have one binding site per chain. However, they are very different in their chemical and physical properties. Most have covalently linked carbohydrates, which account for 50% in the potato lectin and only 0.5% in the lectin of the sponge *Axinella polypoides*. Some, like the jack bean lectin, concanavalin A, are completely devoid of sugars. The molecular weights of the native oligomers range from 8500 for the lectin of the stinging nettle (*Urtica dioica*) to 600,000 for the lectin from the tunicate *Halocynthia roretzi*. On the other hand, lectins from species that are closely related taxonomically exhibit high homology of the gene sequences encoding for those proteins, as has been shown for animal lectins and legume or cereal lectins.

Some lectins require calcium, as in the animal C-type lectins; others require divalent cations (calcium, magnesium, or manganese) to bind their carbohydrate ligands, as in legume lectins. Other lectins do not require divalent cations for binding activity and instead call for the presence of thiol groups, as in the wheat germ agglutinin or the soluble vertebrate lectins (galectins). The sugar specificity of a lectin is defined by the carbohydrate for which it shows the highest affinity, which means that lectins considered specific for one monosaccharide may also bind other structurally related carbohydrates having a lower affinity. In general, lectins interact with the nonreducing glycosyl groups of polysaccharides and glycoproteins, but some can bind internal sugars or sugars at the reducing end. Some lectins can recognize one particular monosaccharide, which suggests the presence of a small binding site, whereas others bind preferentially to trisaccharides or tetrasaccharides and have an extended binding site. Lectins binding oligosaccharides are thought to be more specific for their ligand than those binding monosaccharides.

Another important property of some lectins is their ability to induce cell proliferation of particular tissues or cells. Such mitogenic lectins are used extensively in biomedical research.

It seems clear that a wide variety of proteins share the ability of lectins to bind carbohydrates. Lectins from closely related organisms probably stem from common ancestors and have high degrees of homologies. That is particularly clear in the relatively less complex structure of plant lectins. However, lectins related to a specific function may or may not have structural homologies, depending on their evolutionary history. In addition, the more complex structure of vertebrate lectins indicates that they are built with different polypeptide domains that present homologies to many unrelated proteins. The carbohydrate recognition domain is the only one offering a certain degree of homology.

Biological functions. From the known examples, it seems clear that lectins are involved in recognition phenomena, and their ability to bind particular carbohydrate structures is the key to their biological functions. These recognition functions include their involvement in interactions with cells or extracellular materials from the same organism (which could be considered self-recognition, or recognition of endogenous ligands) and interactions with foreign particles or cells (recognition of non-self, or recognition of exogenous ligands). *See* IMMUNOLOGICAL PHYLOGENY.

Animal lectins are involved in a variety of cellular processes, including intracellular transport and extracellular assembly, cell–cell and cell–matrix interactions, cell growth control and apoptosis (programmed cell death), and immune functions. Microbial lectins largely function in host cell attachment, tissue colonization, and invasion. Plant lectins have been known for more than 100 years. Although their biological roles are not fully understood, the plant lectins are involved in deposition of storage proteins, maintenance of seed dormancy, defense against pathogen and animal predators, symbiosis, transport of carbohydrates, mitogenic stimulation of embryonic plant cells, assembly and elongation of cell walls, and recognition of pollen.

Lectins as tools. Lectins are very useful reagents for the study of complex carbohydrates and cell surfaces, for the separation and identification of particular cells, and for the stimulation of cell proliferation. Lectins are the main tools for glycan (polysaccharide) profiling aiming at comprehensive elucidation of glycan functions in all organisms. Lectins covalently attached to insoluble matrices are used to separate glycoproteins or glycopeptides that contain different carbohydrates. Labeled lectins are also used in histochemical and cytochemical studies to localize glycoconjugates that carry particular sugars. This technique is particularly interesting, since changes in lectin-binding patterns occur during embryonic differentiation, malignant transformation, aging, and some pathological conditions. The spectrum of uses for lectins as specific reagents continues to expand.

Rafael Pont-Lezica

Bibliography. D. C. Kilpatrick, *Handbook of Animal Lectins: Properties and Biomedical Applications*, Wiley, 2000; N. Sharon and H. Lis, *Lectins*, 2d ed., Kluwer Academic, 2003.

Lecythidales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Dilleniidae of the class Magnoliopsida (dicotyledons). The order consists of the single family Lecythidaceae, with about 400 species. They are tropical, woody plants with alternate, entire leaves, valvate sepals, separate petals, numerous centrifugal stamens, and a syncarpous, inferior ovary with axile placentation. Brazil nuts are the seeds of *Bertholletia excelsa*, a member of the Lecythidaceae. *See* BRAZIL NUT; DILLENIIDAE; MAGNOLIOPHYTA; PLANT KINGDOM. Arthur Cronquist

Legendre functions

Solutions to the differential equation $(1 - x^2)y' - 2xy' + v(v + 1)y = 0$.

Legendre polynomials. The most elementary of the Legendre functions, the Legendre polynomial $P_n(x)$ can be defined by the generating function in Eq. (1). More explicit representations are Eq. (2), and the hypergeometric function, Eq. (3).

$$(1 - 2xr + r^2)^{-1/2} = \sum_{n=0}^{\infty} P_n(x) r^n \qquad (1)$$

$$P_n(x) = \frac{(-1)^n}{2^n n!} \frac{d^n}{dx^n}(1 - x^2)^n \qquad (2)$$

$$P_n(x) = {}_2F_1[-n, n + 1; 1; (1 - x)/2] \qquad (3)$$

See HYPERGEOMETRIC FUNCTIONS.

Generating function (1) implies Eq. (4). The function

$$(a^2 - 2ar\cos\theta + r^2)^{-1/2} = \frac{1}{a}\sum_{n=0}^{\infty} P_n(\cos\theta)\,(r/a)^n \qquad (4)$$
$$0 < r < a$$

$(a^2 - 2ar\cos\theta + r^2)^{-1/2}$ represents the potential in an inverse square field at a point P of a source at A, where r and a are the distances from P and A to a fixed point O, and θ is the angle between the segments PO and OA. These functions were extensively studied by A. M. Legendre and P. S. Laplace because they could be used in the study of the celestial mechanics. They had arisen slightly earlier in probabilistic work of J. Lagrange. *See* POTENTIALS.

Since Legendre polynomials are the zonal spherical harmonics on the surface of the unit sphere in three-space, they arise when studying physical phenomena associated with spherical geometry. They arise in many other applications. One application is to the estimation of the smallest eigenvalue of the truncated Hilbert matrix $(a_{i,j})_0^N$, $a_{i,j} = (i + j + 1)^{-1}$. This matrix is very hard to invert numerically because the condition number, or the ratio of the largest to the smallest eigenvalue, is very large. One of the essential parts of the proof is an asymptotic formula for Legendre polynomials of the interval $[-1,1]$. *See* MATRIX THEORY; SPHERICAL HARMONICS.

Relation to trigonometric functions. When $x = \cos\theta$, Legendre polynomials have a relation with trigonometric functions, given by Eq. (5), where $|R(n,\theta)| \leqq$

$$(\sin\theta)^{1/2}\left(n + \frac{1}{2}\right)^{1/2} P_n(\cos\theta) = \sqrt{\frac{2}{\pi}}\cos\left[\left(n + \frac{1}{2}\right)\theta - \frac{\pi}{4}\right] + R(n,\theta) \qquad (5)$$

$A/(n \sin\theta)$, A being a fixed constant, when $c/n \leqq \theta \leqq \pi - c/n$ for $c > 0$. *See* TRIGONOMETRY.

Properties. A graph of $P_n(x)$ [**Fig. 1**] illustrates a number of properties. $P_n(x)$ is even or odd as n is even or odd, that is, $P_n(-x) = (-1)^n P_n(x)$. All the zeros of $P_n(x)$ are real and lie between -1 and 1. The zeros of $P_{n+1}(x)$ separate the zeros of $P_n(x)$. $P_n(x)$ satisfies the inequality $|P_n(x)| \leqq 1$, $-1 \leqq x \leqq 1$. The successive relative maxima of $|P_n(x)|$ increase in size as x increases over the interval $0 \leqq x \leqq 1$. The closest minimum to $x = 1$ of $P_n(x)$ called $\mu_{1,n}$, satisfies $\mu_{1,n} < \mu_{1,n+1}$. Similar inequalities hold for the first maxima of $P_n(x)$ to the left of $x = 1$, and for the kth relative maxima or minima. All of these results have been used in applications. The last two results about the relative maxima were used to obtain bounds on the phase of scattering amplitudes. From an important limiting relation of F. G. Mehler, Eq. (6), it is easy

$$\lim_{n\to\infty} P_n\left(\cos\frac{\theta}{n}\right) = J_0(\theta) \qquad (6)$$

to show that $\mu_{1,n}$ approaches the minimum value of the Bessel function J_0. *See* BESSEL FUNCTIONS.

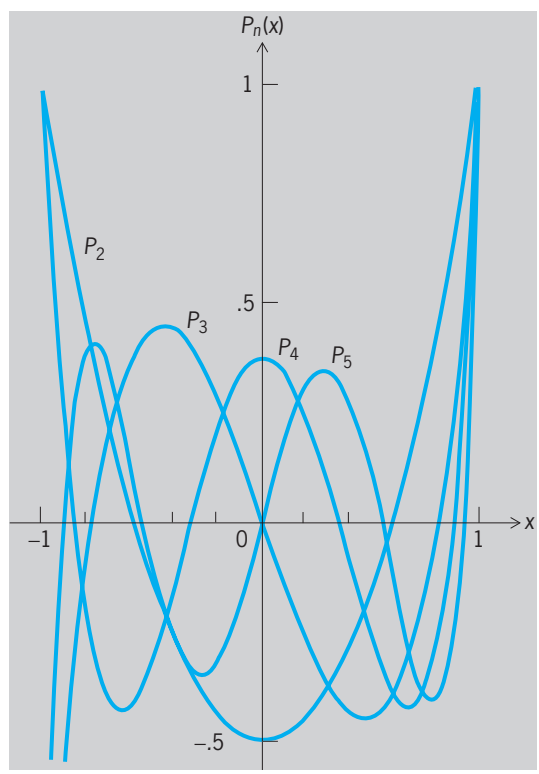

Fig. 1. Legendre polynomials $P_n(x)$, $n = 2, 3, 4, 5$.

Differential equation. One of the main reasons for the occurrence of Legendre polynomials is a differential equation they satisfy, Eq. (7). This equation arises

$$(1-x^2)y'' - 2xy' + n(n+1)y = 0 \qquad (7)$$

$$y = P_n(x)$$

from the solution of Laplace's equation by separation of variables in spherical coordinates. There is a second solution $Q_n(x)$ given by Eq. (8), when x is not in the interval $-1 < x < 1$. For $-1 < x < 1$, $Q_n(x)$ is defined by Eq. (9). $Q_n(x)$ had many of the same properties as $P_n(x)$, illustrated by **Fig. 2**. *See* DIFFERENTIAL EQUATION; LAPLACE'S DIFFERENTIAL EQUATION.

$$Q_n(x) = \frac{1}{2}\int_{-1}^{1} \frac{P_n(t)}{x-t}\,dt \qquad (8)$$

$$Q_n(x) = \lim_{\epsilon\to 0^+} \frac{Q_n(x+i\epsilon) + Q_n(x-i\epsilon)}{2} \qquad (9)$$

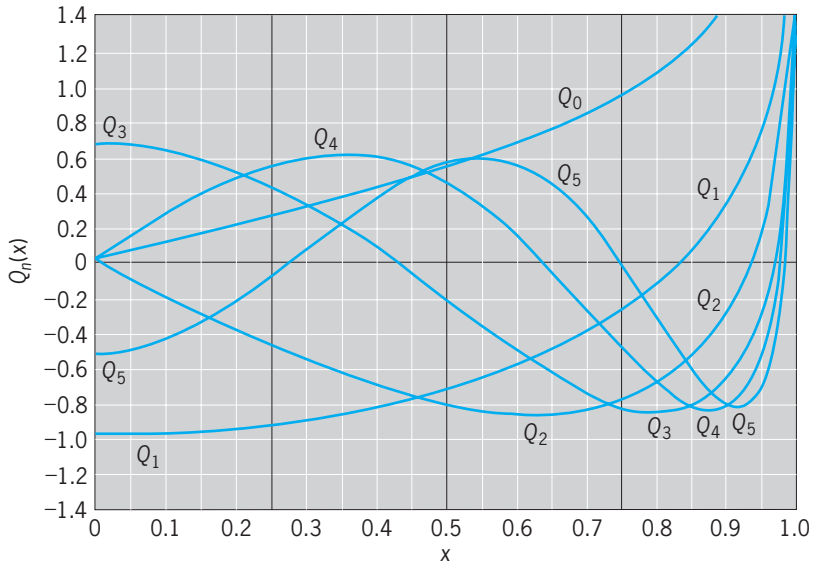

Fig. 2. Legendre functions of the second kind, $Q_n(x)$. (*After E. Jahnke and F. Emde, Tables of Functions with Formulae and Curves, 4th ed., Dover Publications, 1945*)

Addition formula. One of the most important formulas satisfied by Legendre polynomials is the addition formula, Eq. (10). The functions $P_n^m(x)$ are defined by Eq. (11). These functions, often called associated Legendre functions, satisfy differential equation (12).

$$P_n(\cos\theta\cos\varphi + \sin\theta\sin\varphi\cos\chi)$$
$$= P_n(\cos\theta)P_n(\cos\varphi) + 2\sum_{m=1}^{n}\frac{(n-m)!}{(n+m)!}$$
$$\cdot P_n^m(\cos\theta)P_n^m(\cos\varphi)\cos m\psi \qquad (10)$$

$$P_n^m(x) = (-1)^m(1-x^2)^{m/2}\frac{d^m}{dx^m}P_n(x) \qquad (11)$$

$$-1 < x < 1 \quad m = 0, 1, \ldots, n$$

$$(1-x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + \left[n(n+1) - \frac{m^2}{1-x^2}\right]y = 0 \qquad (12)$$

General Legendre functions. In work of F. G. Mehler on electrical distribution with conical symmetry, solutions arose for Eq. (12) with $m = 0$ and $n = -\frac{1}{2} + it$, where t is a parameter that can assume any real value. When Laplace's equation is solved by separation of variables in toroidal coordinates, this equation occurs with $n = \nu - \frac{1}{2}$ and $m = \mu$, where μ and ν are separation parameters. Thus there are reasons for considering solutions to Eq. (12) when n and m are arbitrary complex numbers. In the general context, n and m are usually called ν and μ. $P_\nu^\mu(z)$ is defined by Eq. (13) for $|z-1| < 2$, with the cut $z < 1$ removed. By analytic continuation, this function is extended to the plane cut from 1 to $-\infty$ and then defined for $-1 < z < 1$ by Eq. (14). There is a second solution to differential equation (12) which is called $Q_\nu^\mu(z)$. The easiest way to carry out the analytic extension of Eq. (13) and to define $Q_\nu^\mu(z)$ is to use the general theory of hypergeometric functions. Legendre functions are just appropriate algebraic functions times a hypergeometric function of the type that has a quadratic transformation. The theory of Legendre functions up to but not including the addition formula can be developed very easily in this fashion.

$$P_\nu^\mu(z) = \frac{1}{\Gamma(1-\mu)}\left(\frac{z+1}{z-1}\right)^{\mu/2}$$
$$\cdot\, {}_2F_1[-\nu, \nu+1; 1-\mu; (1-z)/2] \qquad (13)$$

$$P_\nu^\mu(x) = \lim_{\epsilon\to 0^+}\frac{P_\nu^\mu(x+i\epsilon) + P_\nu^\mu(x-i\epsilon)}{2} \qquad (14)$$

Addition formulas are best derived by interpreting the functions as spherical harmonics and using the rotation groups which operate on spheres and then analytically continuing the resulting formulas in the appropriate parameters. Addition formula (10) can be extended not only to the general Legendre function of the first kind, $P_\nu^\mu(z)$, but to the general function of the second kind $Q_\nu^\mu(z)$. A number of results can be obtained from this, such as Eq. (15). This is an extension of the formula $\cos^2 x + \sin^2 x = 1$ to Legendre functions.

$$[P_n(x)]^2 + \frac{4}{\pi}[Q_n(x)]^2$$
$$= \frac{4}{\pi^2}\int_1^\infty Q_n[x^2 + (1-x^2)z](z^2-1)^{-1/2}\,dz \qquad (15)$$

$$-1 < x < 1$$

Gegenbauer functions. There is a related set of functions, called Gegenbauer and associated Gegenbauer functions, which arise as spherical harmonics on higher-dimensional spheres. They are really equivalent to general Legendre functions, since they are

algebraic functions times hypergeometric functions which have a quadratic transformation. The only difference is that different algebraic functions are used as multipliers since the functions satisfy different second-order differential equations. A completely analogous theory has been developed for them. There are two extreme cases of Gegenbauer functions which are very important, and in many respects their theory is easier to develop and is used as a model to suggest further developments. One is connected with the unit sphere in two-space (or the circle $x^2 + y^2 = 1$), and the resulting functions are $\cos \lambda\theta$ and $\sin \lambda\theta$. The other comes from an "infinite dimensional sphere" (actually an appropriate limit of finite dimensional spheres), and the spherical functions are Hermite polynomials and their extensions to Hermite functions. *See* ORTHOGONAL POLYNOMIALS.

Richard Askey

Bibliography. A. Erdéelyi et al., *Higher Transcendental Functions*, vols. 1–3, 1953–1955, reprint 1981; E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics*, 1931, reprint 1975; C. Müller, *Spherical Harmonics*, 1966; I. N. Sneddon, *Special Functions of Mathematical Physics and Chemistry*, 3d ed., 1980; N. Ja. Vilenkin, *Special Functions and the Theory of Group Representations*, 2d ed., 1983, reprint 1988; Z. X. Wang and D. R. Gue, *Special Functions*, 1989.

Legionnaires' disease

A type of pneumonia usually caused by infection with the bacterium *Legionella pneumophila*, but occasionally with a related species (such as *L. micdadei* or *L. dumoffii*). More than 40 species of *Legionella* are known, though not all have been associated with disease. The disease was first observed in an epidemic among those attending an American Legion convention in Philadelphia, Pennsylvania, in 1976. The initial symptoms are headache, fever, muscle aches, and a generalized feeling of discomfort. The fever rises rapidly, reaching 102–105°F (32–41°C), and is usually accompanied by cough, shortness of breath, and chest pains. Abdominal pain and diarrhea are often present. Diagnosis is often difficult, though there are several rapid diagnostic kits, including a urine antigen test that can augment chest x-ray presentation of a patchy interstitial infiltrate (atypical pneumonia in the case of Legionnaires' disease). Proper diagnosis is important since the mortality rate can be as high as 15% in untreated or improperly diagnosed cases. Erythromycin, new-generation fluroquinolones, and rifampicin are considered highly effective medications, whereas the penicillins and cephalosporins are ineffective.

While epidemics of Legionnaires' disease (also referred to as legionellosis) can often be traced to a common source (cooling tower, potable water, or hot tub), most cases seem to occur sporadically. It is estimated that *Legionella* spp. account for approximately 4% of all community- and hospital-acquired pneumonia. Legionnaires' disease is most fequently associated with persons of impaired immune status. Risk factors include smoking, advanced age, diabetes, underlying illness such as cancer, and immunosuppression associated with transplants, in particular, heart transplants. *Legionella* bacteria are commonly found in fresh water and moist soils worldwide and are often spread to humans through inhalation of aerosols containing the bacteria. The legionellae are facultative intracellular parasites of fresh-water protozoa, including most species of amebas. Amebas are also considered as biological amplifiers of *Legionella* bacteria and may be associated with aerosol transmission. Legionnaires' disease is not a communicable disease (no person-to-person spread as occurs with the flu), indicating that human infection is not part of the survival strategy of these bacteria. Therefore, the legionellae are considered opportunistic pathogens of humans. It is technology (air conditioning) and the ability to extend life through medical advances (such as transplantation and treatments for terminal diseases) that have brought these bacteria into proximity with a susceptible population.

Human alveolar macrophages are the primary target cell colonized by *L. pneumophila*, in which the bacteria block early events associated with the fusion of secondary lysosomes with bacteria-laden phagosomes. Instead, the phagosomes associate with and become enveloped by the endoplasmic reticulum, forming a specialized replicative phagosome, permissive for bacterial multiplication. Since this scenario occurs in natural amebic hosts, it is generally believed that the bacteria employ a common infection strategy. The mechanisms by which these bacteria alter host cell biological functions are unknown, and therefore *Legionella* infections of natural and mammalian hosts serve as model systems for study of pathogenesis and of eukaryotic cell biology.

For most humans exposed to *L. pneumophila*, infection is asymptomatic or short-lived. This is attributed to a potent cellular immune response in healthy individuals that includes gamma-interferon production by T lymphocytes as well as quick action by natural killer cells and cytotoxic T cells that kill infected macrophages. Gamma-interferon-activated macrophages restrict the intracellular replication of *L. pneumophila*. In immunocompromised individuals, a lack of gamma-interferon production (suppressed T cells) permits rapid multiplication of the bacteria in alveolar macrophages and development of pneumonia. Recovery from Legionnaires' disease often affords immunity against future infection. However, no vaccine exists at the present time. Most species of *Legionella* can be routinely grown in the laboratory on a special yeast extract medium containing activated charcoal, supplemented with the amino acid cysteine and a source of ferric iron. Bacterial colonies often take 4–7 days to appear, and the bacteria are gram-negative aerobic rods that sometimes display motility. The bacteria can survive in water

for long periods of time following development of a cystlike resting form. *See* MEDICAL BACTERIOLOGY; PNEUMONIA. Paul S. Hoffman

Bibliography. B. S. Fields, The molecular ecology of legionellae, *Trends Microbiol.*, 4:286–290, 1996; P. S. Hoffman, Invasion of eukaryotic cells by *Legionella pneumophila*: A common strategy for all hosts, *Can. J. Infect. Dis.*, 8:139–146, 1997; W. C. Winn, Legionnaires disease: Historical perspective, *Clin. Microbiol. Rev.*, 1:60–81, 1988.

Legume

A member of the plant family Leguminosae, and the name for the fruit produced by members of that family, as idealized in a bean or pea pod. The legume and grass families are by far the world's most important sources of food. Whereas grasses supply starch, legumes, which include beans, soybeans, peas, and alfalfa, supply protein and fats. Through their root nodules, which are inhabited by *Rhizobium* bacteria, the legumes also preserve the nitrogen balance in the soil. *See* FABALES; GRASS CROPS; NITROGEN FIXATION.

Legumes comprise one of the largest plant families in the world, with perhaps 18,000 species. Most species are tropical and include trees, woody vines, and herbaceous plants. All legume plants bear the same type of fruit, the pod or legume, which is diversely modified within the family. Most have compound leaves, and many, particularly tropical legumes, have spectacular flowers, some of which are grown as ornamentals.

The three major groups of Leguminosae differ in the appearance of their flowers, and some botanists consider them to be three separate families. Others, as here, regard them as a single diverse family because all have the same kind of fruit and because the groups intergrade. The first group is the Mimosoids, in which the flowers have reduced petals and usually have conspicuous, long stamens that give color to the flower clusters. The Mimosoid group, which is almost entirely tropical, includes the so-called sensitive plant and the silk tree, or mimosa, a widely planted ornamental in the southern United States. The second group, the Caesalpinioids, is characterized by well-developed petals and includes ornamentals such as the redbud and honey locust in temperate regions and the orchid trees of the tropics. The third and largest group is the Papilionoids, whose members bear a flower that resembles the sweet pea, with a big petal at the top, a wing on either side, and a "keel" of two fused petals that enclose the stamens at the bottom. Many important agricultural plants are members of the Papilionoid group, including peanuts, garden beans, soybeans, garden peas, lentils, chickpeas, cowpeas, clovers, and alfalfa. *Leuceana*, *Prosopis*, and other fast-growing woody genera provide firewood, animal forage, nitrogen fixation, and timber in the tropics. Other species yield natural fibers, flavorings, and medicinal compounds. Many are planted as ornamentals, especially in warm regions and the tropics. At least a dozen tropical species growing in the wild or primitively cultivated, especially in Africa, are locally important for human consumption. *See* ALFALFA; BEAN; CLOVER; COWPEA; LENTIL; PEA; PEANUT; SOYBEAN.

Many economic legumes have not yet been studied scientifically with regard to gene pool exploitation. Considering that presently cultivated wheat and soybeans are far different from the wild plants from which they were derived, these undeveloped legume species may well be a resource for the future. *See* LEGUME FORAGES; NITROGEN-FIXING TREES. Duane Isely

Bibliography. Y. P. Bajaj (ed.), *Legumes and Oilseed Crops I*, 1990; D. Isely, Leguminosae and *Homo sapiens*, *Econ. Bot.*, 36:46–70, 1982; R. H. Mathews, *Legumes: Chemistry*, *Technology*, *and Human Nutrition*, 1989; J. M. Vincent, *Nitrogen Fixation in Legumes*, 1982.

Legume forages

Plants of the legume family used for livestock feed, grazing, hay, or silage. Legume forages are usually richer in protein, calcium, and phosphorus than other kinds of forages, such as grass. The production, preservation, and use of forage legumes require special skills on most soils. One important requirement is a supply of the needed symbiotic nitrogen-fixing bacteria if these are not already in the soil commercial cultures for various strains of these bacteria can be purchased and applied to the legume seed just before planting. Additional lime and commercial fertilizers may be needed on all except fertile soils. Protection from weeds, injurious insects, diseases, and other harmful influences is often required. *See* FERTILIZER; LEGUME; NITROGEN FIXATION.

Legume forages may be preserved for future use as dry hay or silage. To obtain high quality, most legume hay crops must be harvested before the mature stage and thoroughly cured without the loss of leaves; high-moisture hay spoils in warm weather. Legumes are more difficult to preserve as silage than grasses due to their high protein content. However, the addition of a high-carbohydrate material, such as molasses, or, as is usually preferred, the mixing of grass and legume are helpful measures. Legume silage often has an objectionable odor if fermentation is not sufficiently rapid to prevent breakdown of proteins.

Much of the legume forage is grazed. Some crops can be grazed continuously without injury while others require intermittent grazing with rest periods to permit recovery or frequent very light use. A lush growth of a palatable legume crop, such as white clover or alfalfa, often causes bloat in grazing animals, but this may be prevented by restricting intake. The use of drenches to prevent or break up foaming in the stomach act as temporary cures.

Alfalfa is the most important legume forage crop in the United States; it is used mainly for hay but is often grazed. White clover and the annual lespedezas are the most extensively grown legumes for grazing particularly in the southeastern United States. Red clover was an important crop prior to 1930 but is minor now. About a dozen other species of legumes are used for cultivated forage in the United States, and a large number are grown for range grazing. *See* ALFALFA; CLOVER; COVER CROPS; COWPEA; KUDZU; LESPEDEZA; LUPINE. Paul Tabor

Leitneriales

An order of flowering plants, division Magnoliophyta (Angiospermae) in the subclass Hamamelidae of the class Magnoliopsida (dicotyledons). The order consists of a single family, genus, and species (*Leitneria floridana*) of the southeastern United States. The plants are simple-leaved, dioecious shrubs with the flowers in catkins. The ovary is superior and pseudomonomerous, with a single ovule ripening into a small drupe. *See* HAMAMELIDAE; MAGNOLIOPHYTA; PLANT KINGDOM. Arthur Cronquist; T. M. Barkley

Lemming

The name applied to 11 species of rodents in the subfamily Microtinae, family Muridae. These animals (see **illustration**) have a northern circumpolar distribution. Lemmings have been extensively studied from the viewpoint of population dynamics, since their population increases cyclically, resulting in mass migration to the sea.

The lemming, a small burrowing rodent with short legs and stout claws adapted for digging. (*Photo by Robert R. Nack, Minnesota Cooperative Fish and Wildlife Research Unit*)

The Norway lemming (*Lemmus lemmus*) is found in the mountainous wastelands of northern Norway and Lapland. It is usually nocturnal and timid in its habits, except when a population explosion occurs with its resultant migration. Despite many reasons advanced to account for these migrations, the causative factors are not known. Cyclic variations in fertility may be a factor. Usually there are two litters of five offspring each year, but often four litters of two to eight offspring occur. The collared lemmings (*Dicrostonyx*) of Arctic Asia and America and the bog lemmings (*Synaptomys*) of North America have similar migrations. *Lemmus trimucronatus* is found in the Arctic regions of North America. It is a heavy-bodied, short-tailed rodent with short legs, large forefeet with a rudimentary thumb, and short hindfeet. The animal burrows extensively. This species has 16 teeth, with the dental formula I 1/1 C 0/0 Pm 0/0 M 3/3, and eats vegetation. The sexes are identical in appearance. The female may have from two to four litters a year with four to six offspring in each. *See* MUSKRAT; RODENTIA. Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999; D. E. Wilson and S. Ruff, *The Smithsonian Book of North American Mammals*, 1999.

Lemniscate of Bernoulli

A curve shaped like the figure eight (see **illus.**), referred to by Jacques Bernoulli in 1694. Let F_1, F_2 be points of a plane π, with $F_1F_2 = 2a$, $a > 0$. The locus of a point P of π which moves so that $PF_1 \cdot PF_2 = b^2$, where b is a positive constant, is called an

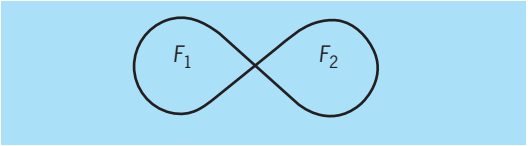

Curve known as a lemniscate.

oval of Cassini. The lemniscate is obtained when $b = a$. Its equation in rectangular coordinates is $(x^2 + y^2)^2 = a^2(x^2 - y^2)$ and in polar coordinates $\rho^2 = a^2 \cos 2\theta$. It is the locus of the point of intersection of a variable tangent to a rectangular hyperbola with the line through the center perpendicular to the tangent. The area enclosed by the lemniscate $\rho^2 = a^2 \cos 2\theta$ is a^2. *See* ANALYTIC GEOMETRY. Leonard M. Blumenthal

Lemon

The fruit *Citrus limon*, commercially the most important of the acid citrus. It probably originated in the eastern Himalayan region of India or adjoining areas, and was introduced into Europe by the Arabs by the twelfth century.

Commercial lemon production developed first in Italy, mainly in Sicily. California is the largest producer, followed closely by Italy. Argentina, Brazil, Spain, Greece, and Turkey are also significant producers, with some commercial production in most Mediterranean countries and in Australia and South Africa. The United States produces about one-third of the total world production.

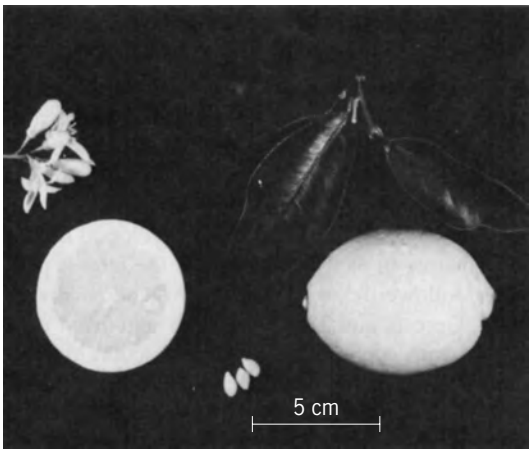

Flowers, foliage, fruit, and seeds of *Citrus limon*.

Culture is generally similar to that for other citrus fruits. Lemons are more sensitive to low temperatures than most citrus, and are susceptible to many problems in warm, humid climates. Therefore, commercial production is restricted to the warmer, drier portions of the citrus areas. In regions with mild winters and cool summers, marketable fruit is available throughout the year.

Lemons are grown on one of several rootstock types. None is entirely satisfactory, and the search for better ones has been ongoing. Sour orange (*C. aurantium*) has been widely used. Sweet orange (*C. sinensis*) is used, but is very susceptible to phytophthora root rot. In the United States, *C. macrophylla* is the most widely used rootstock. *Citrus macrophylla* trees are very productive but short-lived.

In the United States, the principal cultivars are largely classified into two groups, Eureka and Lisbon. Many selections have been made within each of these groups. In Italy, Feminello and selections of it are the most prominent. Berna is the most widely grown cultivar in Spain. Several other cultivars are grown to a limited extent. Most cultivars are vigorous, upright-spreading, and open in growth habit, attaining large size if not controlled by pruning. New shoot growth is purple-tinted, but mature leaves are a light green in contrast to the dark green of oranges. The large, purple-tinted flowers are produced throughout the year. The yellow fruits are medium-sized and elongate with a prominent nipple (see **illus.**).

The principal fungus diseases in the orchard are mal secco (caused by *Phoma tracheiphila*), which is a serious problem in the Mediterranean area, and brown rot (caused by *Phytophthora* sp.), which is a worldwide problem. Several virus diseases are spread through asexual propagation. Exocortis, a viroid, is also spread by pruning and is present in many cultivars. Insect pests are many and varied; several types of scale insects are serious problems, as are various mites. *See* PLANT PATHOLOGY.

The lemon is grown primarily for its acid flavor. For the fresh-fruit market, lemons are picked relatively immature and stored at 55–57°F (13–14°C) and 86–88% relative humidity to provide a constant supply of high-quality fruit. During storage, several fungi can cause serious losses. Blue and green molds (caused by *Penicillium* sp.) and sour rot (caused by *Geotrichum candidum*) are among the most troublesome. Others are *Alternaria* rot, brown rot, stem end rot, and *Trichoderma* rot. Tree-ripe fruits are processed because of their short storage life.

Lemon juice is very high in vitamin C. It has many culinary uses, especially in pies, cakes, ices, candies, jellies, and marmalades, and is also used widely in soft drinks. The principal by-product is lemon oil from the peel. *See* ASCORBIC ACID; FRUIT.

R. K. Soost

Length

A scalar physical quantity indicating a one-dimensional extension in space. Length is one of the three fundamental physical quantities important in mechanics, the other two being mass and time. It can be measured by comparison with an arbitrary standard; the specific one in common usage is the international meter. The most recent definition of the meter was adopted in 1983 at the meeting of the General Conference on Weights and Measures (Conférence Général des Poids et Mesures, or CPGM), when the meter was redefined in terms of time and the speed of light: "The meter is the length of the path traveled by light in a vacuum during a time interval of 1/299,792,458 of a second." The effect was to take the speed of light in a vacuum to be exactly 299,792,458 m/s and then to define the meter in terms of that speed and the most accurately known quantity, the second. *See* LIGHT; TIME.

In theory, anyone with an accurate clock could then measure a meter. Such a time-of-flight method is impractical for most uses, and the meter is measured in terms of the wavelengths of light. The most accurate methods of determining lengths involve the interference of light waves. Distance can be measured by counting the static interference fringes in an interferometer to a precision of a fraction of the wavelength of the light. The wavelength of the light depends on the frequency of the light and on the index of refraction of the medium through which the light travels. The light sources used for metrology have well-known stable wavelengths with uncertainties as small as 1×10^{-10}. For measurements made in air, the index of refraction of the air also must be known to a high degree of precision. *See* FREQUENCY MEASUREMENT; INTERFEROMETRY; LASER; LASER SPECTROSCOPY; REFRACTION OF WAVES.

The meter, originally intended to be 1/10,000,000 of the distance from the Equator to the North Pole along a line of longitude through Paris, was defined by the first CPGM in 1889. It was defined as the distance between two fine lines engraved near the ends of a platinum-iridium bar when the bar was

maintained at 0°C (32°F). The 11th CPGM (1960) replaced the prototype-bar standard with one based on the wavelength of light. Ultimately, measurements of the speed of light were limited by the precision of determining the meter. Finally, the CPGM accepted the best value of the speed of light as a fundamental constant and used that value to define the meter. *See* FUNDAMENTAL CONSTANTS.

Length bars and long gauge blocks are among the commonly used material standards that provide traceability to the definition of the meter. They are used to calibrate a wide range of instruments and other standards. The gauge blocks are calibrated by optical interferometry and may be used in combination to determine larger distances with precision. Longer bars and tapes are calibrated by comparison to these standards. Decimal multiples and submultiples of the meter are used to specify length. In the English system, the inch is defined to be exactly 0.0254 meter. *See* PHYSICAL MEASUREMENT.

Edwin R. Jones; Donald A. Jennings

Bibliography. Documents Concerning the New Definition of the Meter, *Metrologia*, vol. 19, 1984; E. R. Jones and R. L. Childers, *Contemporary College Physics 2001 Update*, 3d ed., McGraw-Hill, New York, 2001; B. N. Taylor, *The International System of Units (SI)*, NIST Spec. Publ., 330, 2001.

Lens (optics)

A curved piece of ground and polished or molded material, usually glass, used for the refraction of light. Its two surfaces have the same axis. Usually this is an axis of rotation symmetry for both surfaces; however, one or both of the surfaces can be toric, cylindrical, or a general surface with double symmetry. The intersection points of the symmetry axis with the two surfaces are called the front and back vertices and their separation is called the thickness of the lens.

Lens Types

There are three lens types, namely, compound, single, and cemented. These are described in the following sections.

Compound lenses. A compound lens is a combination of two or more lenses in which the second surface of one lens has the same radius as the first surface of the following lens and the two lenses are cemented together. Compound lenses are used instead of single lenses for color correction, or to introduce a surface which has no effect on the aperture rays but large effects on the principal rays, or vice versa. Sometimes the term compound lens is applied to any optical system consisting of more than one element, even when they are not in contact.

A group of lenses used together is a lens system. A symmetrical lens is a lens system consisting of two parts, each of which is the mirror image of the other. If one part is a mirror image of the other magnified m times, the system is called hemisymmetric. When $m = 1$, the system is often said to be holosymmetric.

Single lenses. The lens diameter is called the linear aperture, and the ratio of this aperture to the focal length is called the relative aperture. This latter quantity is more often specified by its reciprocal, called the f-number. Thus, if the focal length is 2 in. (50 mm) and the linear aperture 1 in. (25 mm), the relative aperture is 0.5 and the f-number is $f/2$. *See* FOCAL LENGTH.

In precalculation formulas, the lens thicknesses (but not the separations of the lenses) can frequently be neglected. This leads to the convenient fiction of a thin lens.

If ρ_1 and ρ_2 are the front and back curvatures of a lens of refractive index n and thickness d in air, its power is given by Eq. (1). The curvature of the

$$\phi = (n-1)(\rho_1 - \rho_2) + \frac{d(n-1)^2}{n}\rho_1\rho_2 \qquad (1)$$

surface is the reciprocal of its radius.

The distances from the back vertex to the back nodal point and to the back focal point, respectively, are given by Eqs. (2). The last distance is often

$$\begin{aligned} S'_N\phi &= -\frac{d}{n}(n-1)\rho_1 \\ S'_F\phi &= 1 + S'_N\phi \end{aligned} \qquad (2)$$

called the back focus, especially in photographic optics.

The bending of a lens is a change in the curvature of the two surfaces by the same amount. It does not change the power of a thin lens, which is $(n-1)(\rho_1 - \rho_2)$. Bending is an important tool of the designer, for it permits the replacing of one lens by another without changing the data of gaussian optics.

When thick lenses are involved, gaussian optics remains constant only if both the powers of the thick lenses and the distance between the back nodal point of the first lens and the front nodal point of the second remain unchanged. Thus a bending of a thick lens should be accompanied by such an adjustment.

The optical center of a thick lens is the image of the nodal point produced inside the lens. All finite rays through the optical center emerge parallel to their respective directions at their entrance.

An optical center exists also in a hemisymmetric system. It is the point of symmetry which divides the separation of the two parts in the ratio $1/m$. If negative values of m are permitted, any single lens is a hemisymmetric system and the point dividing the thickness of the lens (the separation of the two vertices) in a ratio equal to the ratio of the two radii is the optical center of the lens.

A lens is said to be a collecting lens if $\phi > 0$ and a diverging lens if $\phi < 0$. When $\phi = 0$, the lens is afocal. Several types of collecting and diverging lenses are shown in **Fig. 1**.

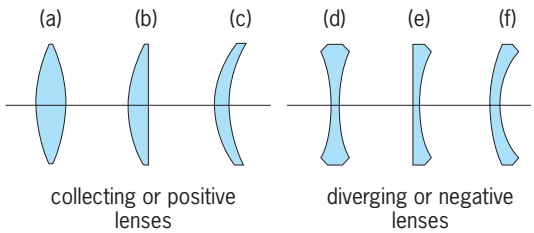

Fig. 1. Common lenses. (*a*) Biconvex. (*b*) Plano-convex. (*c*) Positive meniscus. (*d*) Biconcave. (*e*) Plano-concave. (*f*) Negative meniscus. (*After F. A. Jenkins and H. E. White, Fundamentals of Optics, 3d ed., McGraw-Hill, 1957*)

The surfaces of most lenses are either spherical or planar, but nonspherical surfaces are used on occasion to improve the corrections without changing the power of the lens. *See* OPTICAL SURFACES.

A concentric lens is a lens whose two surfaces have the same center. If the object to be imaged is also at the center, its axis point is sharply imaged upon itself, and since the sine condition is fulfilled, the image is free from asymmetry. Such a lens can be used as an additional system to correct meridional errors.

Another type of lens consists of an aplanatic surface followed by a concentric surface, or vice versa. Such a lens divides the focal length of the original lens to which it is attached by n^2, thus increasing the *f*-number by a factor of n^2 without destroying the axial correction of the preceding system. It does introduce curvature of field which makes a rebalancing of the whole system desirable. *See* ABERRATION (OPTICS).

Cemented lenses. Consider a compound lens made of two or more simple thin lenses cemented together. Let the power of the *k*th simple lens be Φ_k and its Abbe value ν_k. The difference between the powers of the combination for wavelengths corresponding to *C* and *F* is given by Eq. (3) where *N* may be considered

$$\Phi_F - \Phi_C = \frac{\Phi}{N} = \frac{\Sigma \Phi_k}{\nu_k} \tag{3}$$

to be the effective ν-value of the combination. The ν-values of optical glasses vary between 25 and 70, with the ν-value of fluorite being slightly larger ($\nu = 95.1$). By using compound lenses, effective values of *N* can be obtained outside this range. Color correction is achieved as *N* becomes infinite, so that $\Phi_F - \Phi_C = 0$. A lens so corrected is called an achromat. In optical design, it is sometimes desirable to have negative values of *N* to balance the positive values of the rest of the system containing collecting lenses. Such a lens is said to be hyperchromatic. A cemented lens corrected for more than two colors is said to be apochromatic. A lens corrected for all colors of a sizable wavelength range is called a superachromatic lens. *See* CHROMATIC ABERRATION; OPTICAL MATERIALS.

Lens Systems

Optical systems may be divided into four classes: telescopes, oculars (eyepieces), photographic objectives, and enlarging lenses. *See* EYEPIECE; OPTICAL MICROSCOPE.

Telescope systems. A lens system consisting of two systems combined so that the back focal point of the first (the objective) coincides with the front focal point of the second (the ocular) is called a telescope. Parallel entering rays leave the system as parallel rays. The magnification is equal to the ratio of the focal length of the first system to that of the second.

If the second lens has a positive power, the telescope is called a terrestrial or keplerian telescope and the separation of the two parts is equal to the sum of the focal lengths.

If the second lens is negative, the system is called a galilean telescope and the separation of the two parts is the difference of the absolute focal lengths. The galilean telescope has the advantage of shortness (a shorter system enables a larger field to be corrected); the keplerian telescope has a real intermediate image which can be used for introducing a reticle or a scale into the intermediate plane.

Both objective and ocular are in general corrected for certain specific aberrations, while the other abberations are balanced between the two systems. *See* TELESCOPE.

Photographic objectives. A photographic objective images a distant object onto a photographic plate or film. *See* PHOTOGRAPHY.

The amount of light reaching the light-sensitive layer depends on the aperture of the optical system, which is equivalent to the ratio of the lens diameter to the focal length. Its reciprocal is called the *f*-number. The smaller the *f*-number, the more light strikes the film. In a well-corrected lens (corrected for aperture and asymmetry errors), the *f*-number cannot be smaller than 0.5.

The larger the aperture (the smaller the *f*-number), the less adequate may be the scene luminance required to expose the film. Therefore, if pictures of objects in dim light are desired, the *f*-number must be small. On the other hand, for a lens of given focal length, the depth of field is inversely proportional to the aperture.

Since the exposure time is the same for the center as for the edge of the field, it is desirable for the same amount of light to get to the edge as gets to the center, that is, the photographic lens should have little vignetting.

The camera lens can be considered as an eye looking at an object (or its image), with the diaphragm corresponding to the eye pupil. The gaussian image of the diaphragm in the object (image) space is called the entrance (exit) pupil. The angle under which the object (image) is seen from the entrance (exit) pupil is called the object (image) field angle. For most photographic lenses, the entrance and exit pupils are close to the respective nodal points; for such lenses, the object and the image field angles are equal.

In general, photographic objectives with large fields have small apertures: those with large

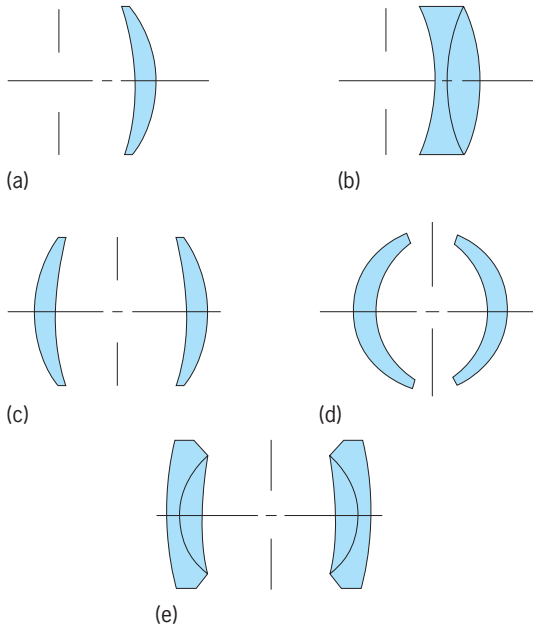

Fig. 2. Older cameralenses. (*a*) Meniscus. (*b*) Simple achromat. (*c*) Periskop. (*d*) Hypergon wide-angle. (*e*) Symmetrical achromat.

apertures have small fields. The construction of the two types of systems is quite different. One can say in general that the larger the aperture, the more complex the lens system must be.

There exist cameras (so-called pinhole cameras) that do not contain any lenses. The image is then produced by optical projection. The aperture in this case should be limited to $f/22$.

Other types of lenses. A single meniscus lens, with its concave side toward the object and with its stop in front at its optical center, gives good definition at $f/16$ over a total field of 50° (**Fig 2***a*). The lens can be a cemented doublet for correcting chromatic errors (Fig. 2*b*). For practical reasons, a reversed meniscus with the stop toward the film is often used.

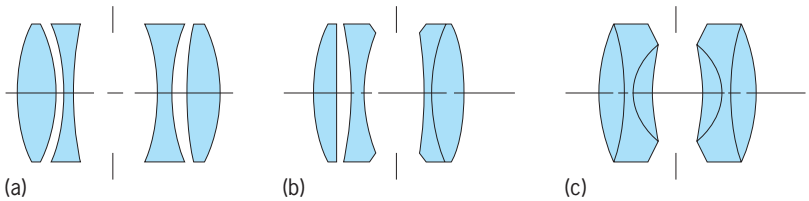

Fig. 3. Types of anastigmats. (*a*) Celor. (*b*) Tessar. (*c*) Dagor.

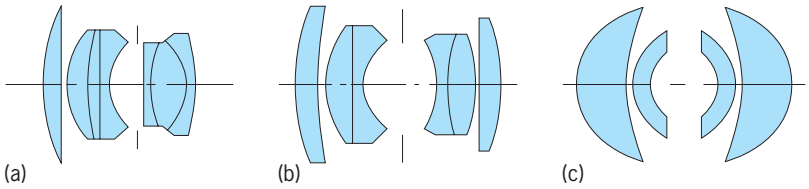

Fig. 4. Modern camera lenses. (*a*) Sonnar. (*b*) Biotar. (*c*) Topogon.

Combining two meniscus lenses to form a symmetrical lens with central stop makes it possible to correct astigmatic and distortion errors for small apertures as well as large field angles (Fig. 2*c*).

The basic type of wide-angle objective is the Hypergon, consisting of two meniscus lenses concentric with regard to the stop (Fig. 2*d*). This type of system can be corrected for astigmatism and field curvature over a total field angle of 180° but it can only be used for a small aperture ($f/12$), since it cannot be corrected for aperture errors. The aperture can be increased to $f/4$ at the expense of field angle by thickening and achromatizing the meniscus lenses and adding symmetrical elements in the center or at the outside of the basic elements.

Two positive achromatic menisci symmetrically arranged around the stop led to the aplanatic type of lens (Fig. 2*e*). This type was spherically and chromatically corrected. Since the field could not be corrected, a compromise was achieved by balancing out sagittal and meridional field curvature so that one image surface lies in front and the other in back of the film.

Anastigmatic lenses. The discovery of the Petzval condition for field correction led to the construction of anastigmatic lenses, for which astigmatism and curvature of field are corrected. Such lenses must contain negative components.

The Celor (Gauss) type consists of two air-spaced achromatic doublets, one on each side of the stop (**Fig. 3***a*). The Cooke triplet combines a negative lens at the aperture stop with two positive lenses, one in front and the other in back. It is called a Tessar (Fig. 3*b*) if the last positive lens is a cemented doublet, or a Heliar if both positive lenses are cemented. The Dagor type consists of two lens systems that are nearly symmetrical with respect to the stop, each system containing three or more lenses (Fig. 3*c*).

Modern lenses. To increase the aperture, the field, or both, it is frequently advantageous to replace one lens by two separated lenses, since the same power is then achieved with larger radii and this means that the single lenses are used with smaller relative apertures. The replacing of a single lens by a cemented lens changes the color balance, and thus the designer may achieve more favorable conditions. Moreover, the introduction of new types of glass (first the glasses containing barium, later the glasses containing rare earths) led to lens elements which for the same power have weaker surfaces and are of great help to the lens designer, since the errors are reduced.

Of modern designs the most successful are the Sonnar, a modified triplet, one form of which is shown in **Fig. 4***a*; the Biotar (Fig. 4*b*), a modified Gauss objective with a large aperture and a field of about 24°; and the Topogon (Fig. 4*c*), a periscopic lens with supplementary thick menisci to permit the correction of aperture aberrations for a moderate aperture and a large field. One or two

plane-parallel plates are sometimes added to correct distortion.

Special objectives. It is frequently desirable to change the focal length of an objective without changing the focus. This can be done by combining a fixed near component behind the stop with an exchangeable set of components in front of the stop. The designer has to be sure that the errors of the two parts are balanced out regardless of which front component is in use. For modern ways to change the magnification *see* ZOOM LENS

The telephoto objective is a specially constructed objective with the rear nodal point in front of the lens, to combine a long focal length with a short back focus. *See* TELEPHOTO LENS.

The Petzval objective is one of the oldest designs (1840) but one of the most ingenious. It consists in general of four lenses ordered in two pairs widely separated from each other. The first pair is cemented and the second usually has a small air space. For a relatively large aperture, it is excellently corrected for aperture and asymmetry errors, as well as for chromatic errors and distortion. It is frequently used as a portrait lens and as a projection lens because of its sharp central definition. Astigmatism can be balanced but not corrected.

Enlarger lenses and magnifiers. The basic type of enlarger lens is a holosymmetric system consisting of two systems of which one is symmetrical with the first system except that all the data are multiplied by the enlarging factor m. When the object is in the focus of the first system, the combination is free from all lateral errors even before correction. A magnifier in optics is a lens that enables an object to be viewed so that it appears larger than its natural size.

The magnifying power is usually given as equal to one-quarter of the power of the lens expressed in diopters. *See* DIOPTER; MAGNIFICATION.

Magnifying lenses of low power are called reading glasses. A simple planoconvex lens in which the principal rays are corrected for astigmatism for a position of the eye at a distance of 10 in. (25 cm) is well suited for this purpose, although low-power magnifiers are often made commercially with biconvex lenses. A system called a verant consists of two lenses corrected for color, astigmatism, and distortion. It is designed for stereoscopic vision at low magnification. *See* STEREOSCOPY.

For higher magnifications, many forms of magnifiers exist. One of the basic designs has the form of a full sphere with a diaphragm at the center, as shown in **Fig. 5***a*. The sphere may be solid or it may be filled with a refracting liquid. When it is solid, the diaphragm may be formed by a deep groove around the equator. Combinations of thin planoconvex lenses as shown in Fig. 5*b* and *c* are much used for moderate powers. Better correction can be attained in the aplanatic magnifier of C. A. Steinheil, in which a biconvex crown lens is cemented between a pair of flint lenses (Fig. 5*d*).

A design by C. Chevalier (Fig. 5*e*) aims for a large object distance. It consists of an achromatic negative lens combined with a distant collecting front lens. A magnifying power of up to 10× with an object distance up to 3 in. (75 mm) can be attained.

Max Herzberger

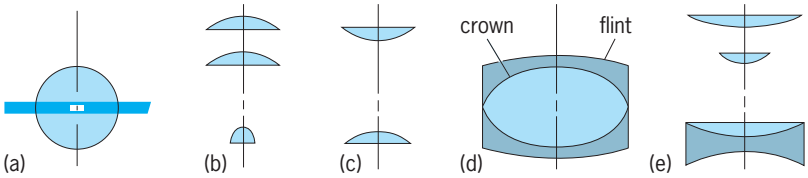

Fig. 5. Typical magnifiers. (*a*) Sphere with equatorial diaphragm. (*b*,*c*) Planoconvex lens combinations. (*d*) Steinheil triple aplanat. (*e*) Chevalier type.

Bibliography. E. Hecht and A. Zajac, *Optics*, 3d ed., 1997; M. Herzberger, *Modern Geometrical Optics*, 1958, reprint 1980; F. A. Jenkins and M. E. White, *Fundamentals of Optics,* 4th ed., 1976; R. Kingslake, R. R. Shannon, and J. C. Wyant (eds.), *Applied Optics and Optical Engineering*, 15 vols., 1965–2000; D. Malacara and Z. Malacara, *Handbook of Lens Design*, 1994; S. F. Ray, *Applied Photographic Optics*, 2d ed., Focal Press, 1996; W. J. Smith, *Modern Lens Design: A Resource Manual*, 1992.

Lentil

A semi-viny annual legume with slender tufted and branched stems 18–22 in. (46–56 cm) long. The lentil plant (*Lens esculenta*) was one of the first plants brought under cultivation. They have been found in the Bronze Age ruins of the ancient lake dwellings of St. Peter's Island, Lake of Bienne, Switzerland. Lentils have been discovered in Poland dating back to the Iron Age. In the Bible the "red pottage" for which Esau gave up his birthright to his brother, Jacob, was probably lentil soup. Large-seeded lentils originated in the Mediterranean region; medium-sized lentils originated in the inner mountains of Asia Minor; and Afghanistan was the original home of the smallest-seeded lentils. *See* LEGUME.

Production. The world's lentil production is centered in Asia, with nearly two-thirds of the production from India, Pakistan, Turkey, and Syria. Whitman and Spokane counties in Washington, and Latah, Benewah, and Nez Perce counties in Idaho grow about 95% of the lentils produced in the United States.

Description. Lentil leaves are pinnately compound and generally resemble vetch. The plant has tendrils similar to those of pea plants.

The seeds grow in short broad pods, each pod producing two or three thin lens-shaped seeds (see **illus.**). Seed color varies from yellow to brown and may be mottled, although mottled seeds are not desirable for marketing. The lentil flowers, ranging in color from white to light purple, are small and delicate and occur at different locations on the stems.

Lens-shaped lentil seeds.

Lentil seed is used primarily for soups but also in salads and casseroles. Lentils are more digestible than meat and are used as a meat substitute in many countries.

Culture. Lentils require a cool growing season; they are injured by severe heat. Therefore, they are planted in April, when soil moisture is adequate and temperatures are cool. A fine firm seedbed is required; the land is usually plowed in the fall and firmed by cultivation in the spring before seeding. Lentils are usually planted in rotation with winter wheat. Seeds are planted in 7–12 in. (21–30 cm) rows at depths of $^1/_2$ to 1 in. (13 to 25 mm), on an average of 60–75 lb/acre (67–84 kg/hectare). Applications of sulfur, molybdenum, and phosphorus are used to increase yields.

Wild oats often infest lentil fields, but chemicals are available to aid weed control. Cowpea and black-bean aphids are the two most important insect pests on lentils. Predators such as ladybird beetles and syrphid-fly larvae usually keep these insects under control, but when their populations are too low to control the aphids, chemical insecticides are used.

Harvesting. Lentils are mowed or swathed when the vines are green and the pods have a golden color. About 10 days later, lentils are ready to combine harvest using the same combines that are used for wheat, oats, and barley. A pick-up attachment picks up the material from the windrows. The combines must be operated at a maximum speed of $1^1/_2$ mi/h (0.7 m/s) to prevent loss of, and damage to, the lentil seed. Average yields are about 900 lb/acre (1000 kg/ha). *See* AGRICULTURAL MACHINERY; AGRICULTURAL SOIL AND CROP PRACTICES.

Kenneth J. Morrison

Bibliography. O. N. Allen and E. K. Allen, *The Leguminosae: A Source Book of Characteristics, Uses, and Nodulation*, 1981; R. J. Summerfield and A. H. Bunting (eds.), *Advances in Legume Science*, vol. 1, 1980.

Lenz's law

A law of electromagnetism which states that, whenever there is an induced electromotive force (emf) in a conductor, it is always in such a direction that the current it would produce would oppose the change which causes the induced emf. If the change is the motion of a conductor through a magnetic field, as in the **illustration**, the induced current must be in such a direction as to produce a force opposing the motion. If the change causing the emf is a change of flux threading a coil, the induced current must produce a flux in such a direction as to oppose the change. That is, if the change is an increase of flux, the flux due to the induced current must be opposite in direction to the increasing flux. If the change is a decrease in flux, the induced current must produce flux in the same direction as the decreasing flux.

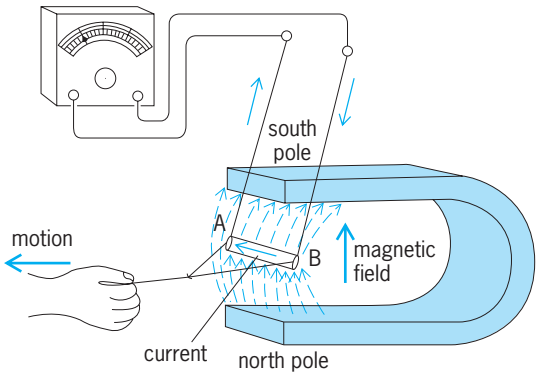

Induced emf in a moving conductor. Direction of current induced in wire AB is indicated by the arrows. (*After M. W. White, K. V. Manning and R. L. Weber, Practical Physics, 2d ed., McGraw-Hill, 1955*)

Lenz's law is a form of the law of conservation of energy, since it states that a change cannot propagate itself. *See* CONSERVATION OF ENERGY; ELECTROMAGNETIC INDUCTION; ELECTROMOTIVE FORCE (EMF).

Kenneth V. Manning

Leo

The Lion, a northern zodiacal constellation (see **illustration**). Its head and front legs are marked by stars tracing out a backward question mark (without a bottom dot), also known as the Sickle. The bright star Regulus is at the bottom of the Sickle. The end of Leo's tail is marked by the bright star Denebola; its name is Arabic for lion's tail. *See* ZODIAC.

Leo is in the evening sky during the spring and early summer. The Leonid meteor shower, which appears to emanate from Leo each November 17, is especially outstanding every 33 years, most recently in 1999–2001, so it will next be prominent starting around 2032. *See* METEOR.

The modern boundaries of the 88 constellations,

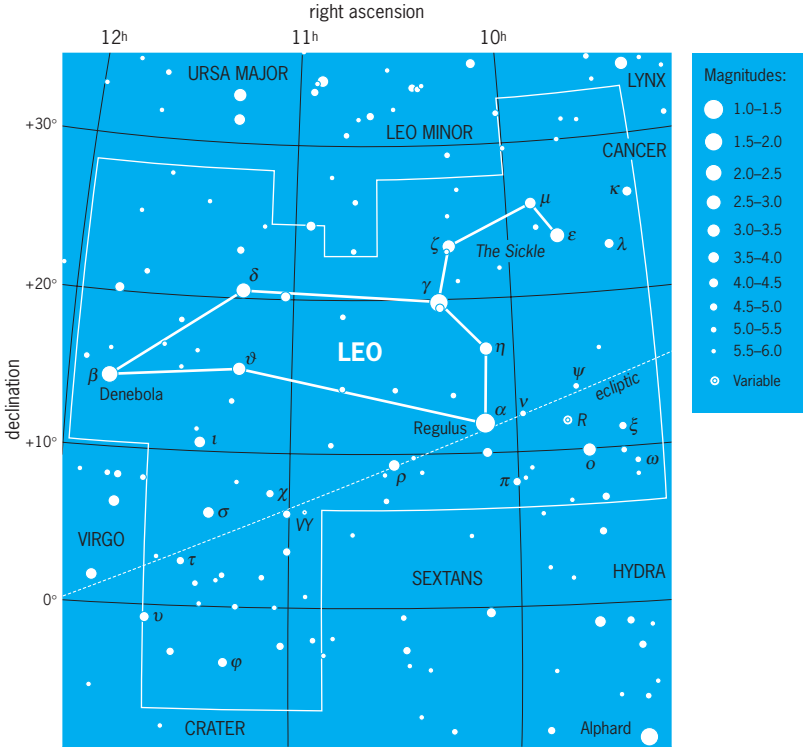

Modern boundaries of the constellation Leo, the Lion. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Lepidodendrales

An extinct order of the class Lycopsida which, together with Isoetales, forms the monophyletic rhizomorphaleans, the most derived and diverse group of clubmosses. The lepidodendraleans are best known as the scale trees that dominate most reconstructions of Upper Carboniferous swamps, where they were the main constituent of many coal-forming peats.

Lepidodendraleans are represented most frequently in the fossil record as characteristic "tire-track" bark fragments, which reflect the regular geometric arrangements of leaves, although taxonomic revisions have focused on the more complete information obtained from fossils preserved three-dimensionally in petrified peats. Reconstruction of the plants from their component organs has revealed that the classic trees (**Fig. 1**) possessed a wide range of growth architectures and shared lowland habitats with other smaller-bodied genera, reflecting a major radiation that occurred in the Late Devonian and perhaps Early Carboniferous.

Rhizomorphic lycopsids are distinguished by their centralized rootstock (rhizomorph), which permitted bipolar growth; both aerial and subterranean axes could branch, unlike the unipolar rhizomes of other, more primitive lycopsids. This trait and the newly acquired ability to produce wood allowed many rhizomorphic lycopsids to form large upright trees, with trunks up to 115 ft (35 m) tall and 3 ft (1 m) in diameter. The tree-lycopsids relied primarily on the external cylinder of bark for support rather than the relatively narrow central cylinder of wood, which was adapted primarily for water transport. Photosynthesis occurred largely in the leaf bases and bark, as well as in the narrow, grasslike leaf laminae.

The lepidodendraleans consisted of four main growth modules: rhizomorph, stem, lateral branches, and crown branches. Only one of the two branch types was well expressed in any one genus. The growth of the tree-lycopsids was probably very rapid. Trees with disposable lateral branches reached reproductive maturity earlier than those that relied only on crown branches. However, unlike most seed plants, branches functioned primarily for reproduction and propagule dispersal rather than light capture; even when mature, they formed open canopies that cast little shade.

The evolution of the group demonstrably entailed progressively increased reproductive sophistication. The more primitive genera possessed bisporangiate cones resembling those of the Selaginellales (**Fig. 2**). Typically, megasporangia containing "female" megaspores developed at the base of the cone where greater resources were available, and

Fig. 1. Reconstructions of tree-sized Carboniferous rhizomorphic tree-lycopsids, illustrating architectural variation. (*After R. M. Bateman et al., Experimental cladistic analysis of anatomically preserved arborescent lycopsids from the Carboniferous of Euramerica, Ann. Mo. Bot. Gard., 79:500–559, 1992*)

microsporangia containing "male" microspores developed closer to the cone apex. The genus *Sigillaria* is characterized by the segregation of the two genders of sporangium in different cones, allowing modifications of the female cone without necessarily modifying the male cone. All subsequent genera underwent selective abortion to leave only one functional megaspore in each megasporangium, now having thinner walls and lacking external ornamentation since it could no longer serve any function. The sporophylls then enclosed the megasporangium to form dispersal units that were the largest of any pteridophyta. Termed aquacarps, they typify *Lepidophlois* and *Miadesemia*. *See* SELAGINELLALES.

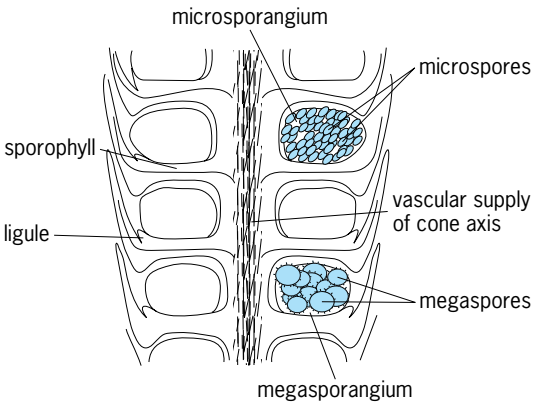

Fig. 2. Primitive bisexual cone of the *Flemingites* type, borne by *Paralycopodites*, with apically concentrated microsporangia and basally concentrated megasporangia containing several megaspores. (*After C. A. Arnold, An Introduction to Paleobotany, McGraw-Hill, 1947*)

These reproductive changes were accompanied by vegetative modifications. Notably, stelar morphology and leaf base anatomy (the basis of a separate organ taxonomy for bark) became more complex, particularly in the larger genera, and ligules became recessed in pits. In contrast, no such progressive increase in evolutionary sophistication occurred in the gross vegetative morphology of the plants. Frequent, profound, and broadly heterochronic changes in body size favored transitions from trees to shrubs and pseudoherbs (plants that were woody and determinate but small bodied and recumbent).

The lepidodendraleans partitioned and occupied most potential ecological niches, at least in the extensive European-American swamps of the Late Carboniferous. The warming and drying of the global climate toward the end of this period greatly reduced the diversity of the rhizomorphic lycopsids, beginning with the most derived and specialized genera such as *Lepidophloios* and eventually leaving only the increasingly reduced and ecologically specialized isoetaleans to survive to present day. *See* CARBONIFEROUS; ISOETALES; LYCOPHYTA; LYCOPSIDA.

Richard M. Bateman; William A. DiMichele

Bibliography. R. M. Bateman, Evolutionary-developmental change in the growth architecture of fossil mizomorphic lycopsids, *Biol. Rev.,* 69:527–597; T. L. Phillips and W. A. DiMichele, Comparative ecology and life-history biology of arborescent lycopsids, *Ann. Mo. Bot. Gard.*, 79:560–588, 1992; W. N. Stewart and G. W. Rothwell, *Paleobotany and the Evolution of Plants*, 2d ed., 1993.

Lepidolite

A mineral of variable composition which is also known as lithium mica and lithionite, $K_2(Li,Al)_{5-6}(Si_{6-7},Al_{2-1})O_{20-21}(F,OH)_{3-4}$. Rubidium (Rb) and cesium (Cs) may replace potassium (K); small amounts of Mn, Mg, Fe(II), and Fe(III) normally are present; and the OH/F ratio varies considerably. Polithionite is a silicon- and lithium-rich, and thus aluminum-poor, variety of lepidolite.

Lepidolite is uncommon, occurring almost exclusively in structurally complex granitic pegmatites, commonly in replacement units. Common associates are quartz, cleavelandite, alkali beryl, and alkali tourmaline. Lepidolite is a commercial source of lithium, commonly used directly in lithium glasses and other ceramic products. Important deposits occur in the Karibib district of South-West Africa and at Bikita, Rhodesia.

The structural modifications show some correlation with lithium content: the six-layer monoclinic form contains 4.0–5.1% Li_2O; the one-layer monoclinic, 5.1–7.26% Li_2O. A three-layer hexagonal form is also found. There is a compositional gradation to muscovite, intermediate types being called lithian

Group of lepidolite crystals found in Zinnwald, Czechoslovakia. (*Specimen from Department of Geology, Bryn Mawr College*)

muscovite, containing 3–4% Li_2O, and having a modified two-layer monoclinic muscovite structure.

Lepidolite usually forms small scales or fine-grained aggregates (see **illus.**). Its colors, pink, lilac, and gray, are a function of the Mn/Fe ratio. It is fusible at 2, yielding the crimson (lithium) flame. It has a perfect basal cleavage. Hardness is 2.5–4.0 on Mohs scale; specific gravity is 2.8–3.0. *See* MICA; SILICATE MINERALS.

E. William Heinrich

Lepidoptera

The order of scaly-winged insects, including the butterflies, skippers, and moths. One of the largest orders in the class Insecta, the Lepidoptera include over 150,000 known species (of which about 10,000 occur in North America) divided among more than 100 families. The adults have a covering of hairs and flattened setae (scales) on the wings, legs, and body, and are often beautifully colored. With minor exceptions, the adults are also characterized by two pairs of membranous wings, and sucking mouthparts that feature a prominent coiled proboscis. This feeding apparatus is formed from a pair of specially elongated and grooved lobes, the galeae (a mouthpart), that are closely joined along their length to make a flexible tube. Adults having a proboscis (the vast majority) can take only liquid food, such as nectar and juices of fruits. Butterflies and skippers usually fly in the daytime, while most moths are nocturnal. The caterpillars are almost always herbivorous and chew their food like a grasshopper or beetle. *See* INSECTA.

Morphology. The head of most adult Lepidoptera species is dominated by the bulbous compound eyes (**Fig. 1***a*), as seen in other insects. The most unusual feature of most moths and butterflies, however, is the form of the mouthparts (Fig. 1*a–c*). The proboscis is thought to be extended by blood pressure created by retraction of special muscles at its base. Other muscles, arranged diagonally within each half of the proboscis, appear to be necessary for the whole proboscis to coil. Liquid is drawn up into the proboscis by the action of a muscular sucking pump inside the insect's head (**Fig. 2***a*). Simple eyes (ocelli) located near the top of the head (Fig. 1*b*) are absent in many groups, such as the Hesperioidea and Papilionoidea. The antennae are variable in form (**Fig. 3**).

Thorax. The prothorax, the first thoracic segment, is well developed in some lower groups, such as the Hepialoidea, but it is considerably smaller than the second and third pterothoracic, or wing-bearing, segments, and is largely membranous in the majority of Lepidoptera. The most prominent feature of the dorsum of the prothorax, in most groups, is a pair of protuberant lobes, the patagia.

The paired basal segments (coxae) of the prothoracic legs are cylindrical and functional, while the two pairs of coxae of the pterothorax (mesothorax and metathorax) are fused with the thoracic capsule and are thus immobile. In each segment of the thorax, the midventral suture, the discrimen, is represented internally by a fine membranous lamella. The form of this lamella and its relationship to the furca (a forked internal projection arising from the base of the thorax), especially in the mesothorax, have proved significant for the classification of some Lepidoptera, including butterflies. The mesothorax is the largest of the three segments and may completely overlap the metathorax.

There is considerable variation throughout the order in the details of the many different sclerites (the hardened plates that make up the body wall) and their connecting sutures. In the butterflies, interesting variation has also been found in the structure of the gut, but this has not been studied extensively in the rest of the order. Many questions remain unanswered concerning the morphology, function, and evolution of this fundamental body system.

Scales. The wing scales are very variable in form (Fig. 1*d*). Generally, they are flat, thin, cuticular sacks with striated upper surfaces. They have a basal stalk, the pedicel, that fits into a socket in the wing membrane. The males of many species have special scales called androconia. These have feathered tips (Fig. 1*e*) or various other forms, and are often gathered together as special androconial organs. Their usual function is to disseminate scents, notably those used by the males during courtship. The vast spectrum of colors seen in the Lepidoptera can be grouped into two sorts, pigmentary and structural. The former result from various chemical pigments, such as melanin (black/brown/orange) or papiliochrome (yellow), deposited in the scales. Structural colors are produced either by fine diffracting ridges on the surface of the scales or, more usually, by layers within the scales that are so close together they interfere with light. These diffraction or interference colors are most often metallic or iridescent greens and blues, but the remarkable cryptic green of some hairstreak butterflies (for example, *Callophrys rubi*) is structural, yet totally noniridescent. In some Lepidoptera, blue bile pigments deposited between the wing membranes also contribute to the overall color pattern, as in species of *Graphium* (Papilionidae).

Wings. In most moths, the fore- and hindwings on each side are coupled together. In females this is achieved by a group of stiff bristles or setae, or

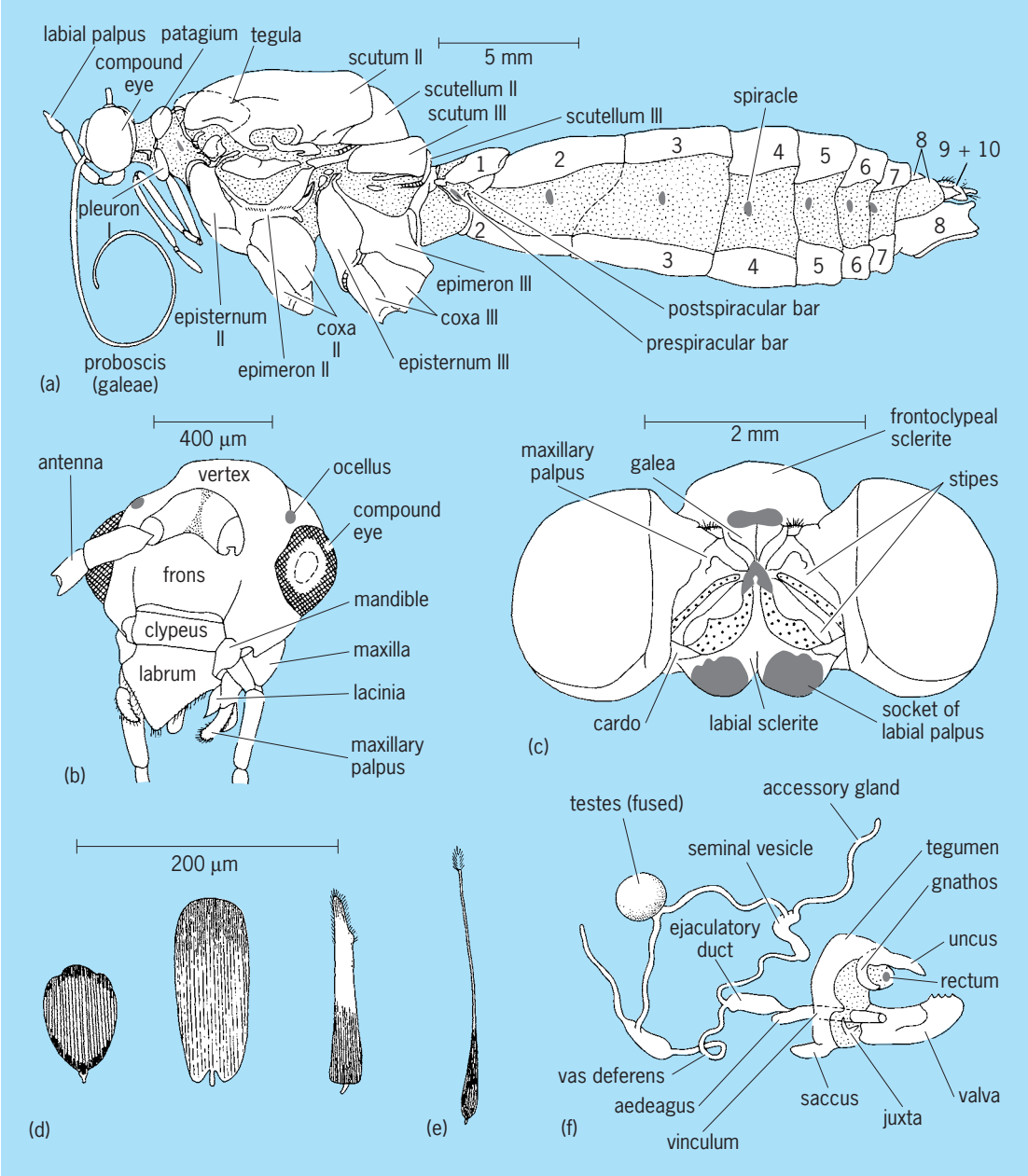

Fig. 1. Adult Lepidoptera anatomy. (*a*) Lateral view of *Danaus plexippus* (Nymphalidae) with hairs, scales, wings, and mid- and hindlegs removed. (*b*) Front view of head of *Epimartyria* (Micropterygidae). (*c*) Ventral view of head of *D. plexippus*. (*d*) Unspecialized scales. (*e*) Androconia (specialized male scales). (*f*) Male genital system.

in males by a large spine formed from several fused setae that projects forward from the base of the hindwing and is held by a clasp on the forewing. The spine or group of setae is known as the frenulum and the clasp as the retinaculum (**Fig. 4***a*). In homoneurous moths (the shape and venation of the fore- and hindwings are similar; see below), there is a lobe, the jugum, at the base of the forewing, which engages with the hindwing or with the frenulum when it is present (Fig. 4*b*). In butterflies, skippers, and some moths, the humeral angle of the hindwing is expanded and strengthened by one or more humeral veins (Fig. 4*c*). In these groups, the frenulum is usually lost, the wings being united functionally by the overlapping lobe.

Genitalia. The external genitalia, together with the anus, usually occupy the last three segments of the abdomen. The form of the external genitalia, especially of the male (Fig. 1*f*), has been used widely in the separation and classification of species. Thus, the valva, uncus, gnathos, and saccus often differ widely between groups and sometimes even closely related species. During copulation, the vulva (ostium bursae) of the female receives the phallus (aedeagus) of the male, through which the sperm is passed in a packet, the spermatophore. The female stores the spermatophore in a special chamber, the corpus bursae, and can then fertilize her eggs as they are laid, sometimes many days later. This is achieved in most Lepidoptera by the sperm passing along a seminal

duct that connects the corpus bursae with the egg duct (oviduct), through which the eggs must pass before they can be laid. In some groups the oviduct terminates in a distinct egg-laying apparatus, the ovipositor.

Developmental Stages

Development starts with the fertilized egg nucleus, which divides and grows, nourished by nutrients stored inside the egg, to give rise to the first larval stage (first instar). This has to bite its way out of the egg shell to commence its independent development. The larvae of Lepidoptera, commonly called caterpillars, are mandibulate and cylindrical, with short articulated thoracic legs and a variable number of fleshy abdominal prolegs. They have one pair of thoracic spiracles and eight pairs of abdominal spiracles (external breathing orifices). After the larva grows and sheds its skin several times, the skin of the fully grown, last larval instar is eventually shed to reveal the immobile pupa. This is the stage during which metamorphosis takes place, the transformation between the wormlike larval stage and winged adult. Inside the pupa, most of the old larval structures are broken down and used as materials and energy supplies to build the adult insect. The pupae of Lepidoptera are variable in form and are often enclosed in a silken cocoon. *See* INSECT PHYSIOLOGY; METAMORPHOSIS.

Egg. There are two general types of eggs, flattened and upright. In both types the surface of the egg shell (the chorion) is usually sculptured, with the sculpturing on the upright type generally more complex (**Fig. 5**). There is a microscopic entrance in the chorion, the micropyle, through which sperm can enter. The eggs are usually laid on the food plant, either singly or in clusters, attached by means of a glue or cement. However, the eggs are sometimes merely scattered in the vicinity of the food plant.

Larva. The caterpillars are variable in color, size, and shape; the presence of warts, hairs, and setae; the arrangement of hooks (or crochets) on the prolegs; and many other superficial details. They are rather constant, however, in basic morphology (**Fig. 6**). There are usually six pairs of ocelli, and compound eyes are absent. At most, eight pairs of prolegs are present, but almost every degree of reduction is found; most species have five pairs of prolegs. *See* CATERPILLAR.

Silk is produced by the labial glands. They are long, coiled, simple tubes which unite anteriorly in a cylindrical spinneret, opening at the front of the labium. *See* SILK.

Pupa. There are three main types, although the vast majority are obtect, with the appendages (legs, antennae, wings) completely sheathed with the rest of the body and immobile, and only the terminal abdominal segments movable. In the Hepialoidea, Cossoidea, and other lower moths, more of the abdominal segments are movable (**Fig. 7**). Although sheathed, the pupal appendages of the Hepialoidea and Cossoidea are partly free, and their pupae are referred to as incomplete. The

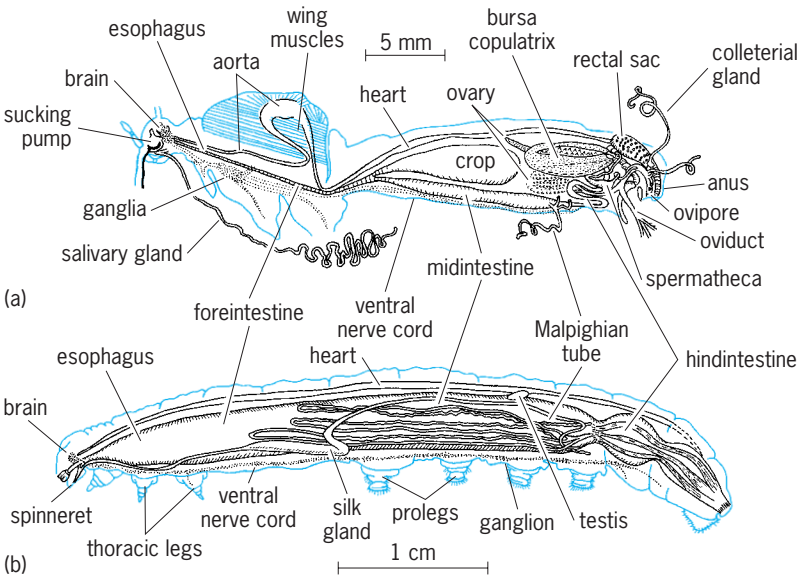

Fig. 2. Internal anatomy of *Danaus plexippus* (Nymphalidae) with tracheal (breathing tube) system and most of musculature omitted. (*a*) Adult female. (*b*) Larval male.

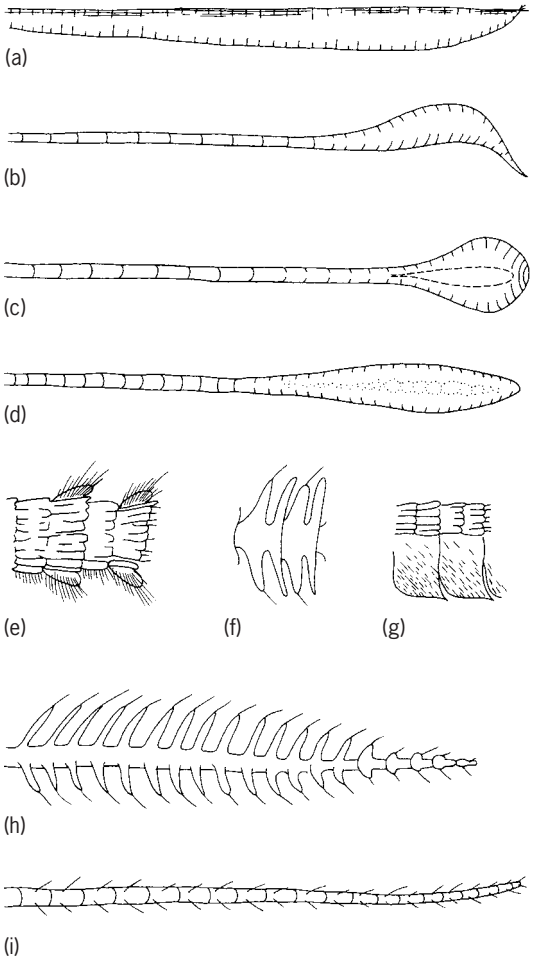

Fig. 3. Types of Lepidoptera antennae. (*a*) Fusiform. (*b*) Clubbed and hooked. (*c*, *d*) Clubbed. (*e*) Serrate and fasciculate, dorsal view. (*f*) Doubly bipectinate (quadripectinate). (*g*) Laminate, lateral view. (*h*) Bipectinate. (*i*) Simple and ciliate.

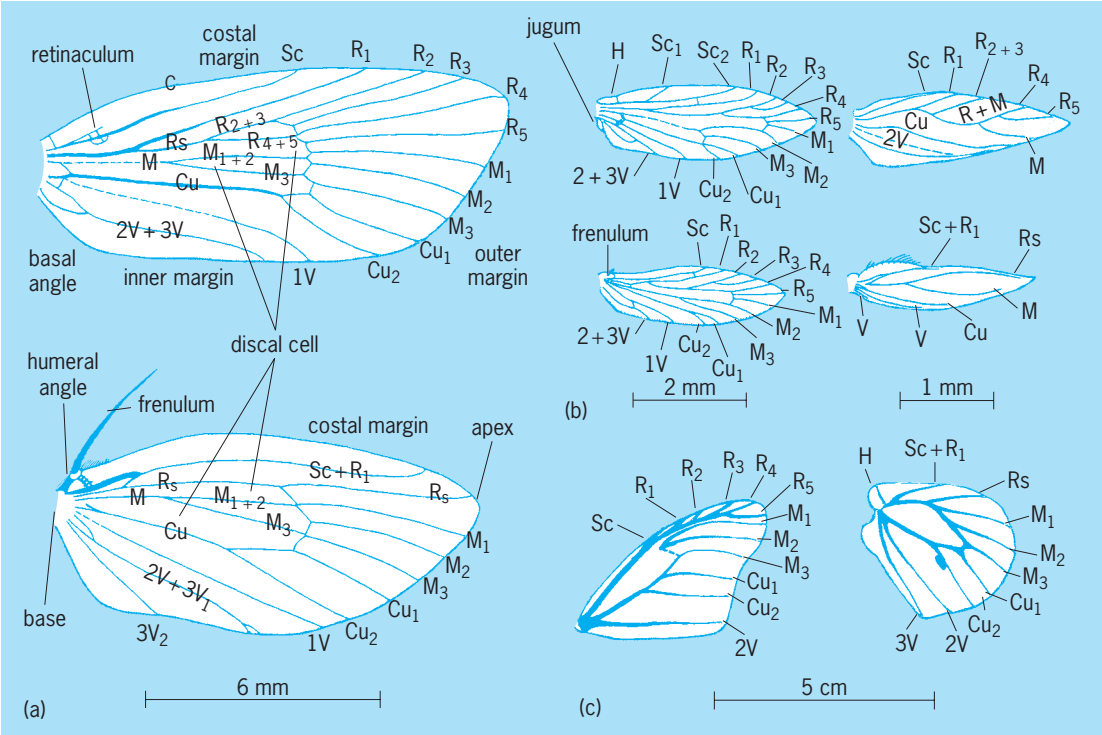

Fig. 4. Wing venation patterns with the veins labeled. (*a*) Male *Acrolophus popeanellus* (Tineidae). (*b*) *Epimartyria* (Micropterygidae) and *Nepticula nyssaefoliella* (Nepticulidae). (*c*) *Danaus plexippus* (Nymphalidae). C = costa; Sc = subcosta; R = radius; R_s = radial sector; M = media; Cu = cubitus; 1V, 2V, 3V = vannal or anal veins; H = humeral vein; subscripts refer to branches (for example, R_2 is second branch of radius).

most primitive moths, such as the Micropterygoidea and Eriocranioidea, have free (exarate) pupae in which the appendages of the head and thorax are not sheathed, the segments of the abdomen are movable, and the outer covering (integument) is soft. These pupae are also decticous, having functional mandibles. All Lepidoptera with obtect pupae, including those that are incomplete, are adecticous, that is, lacking functional mandibles.

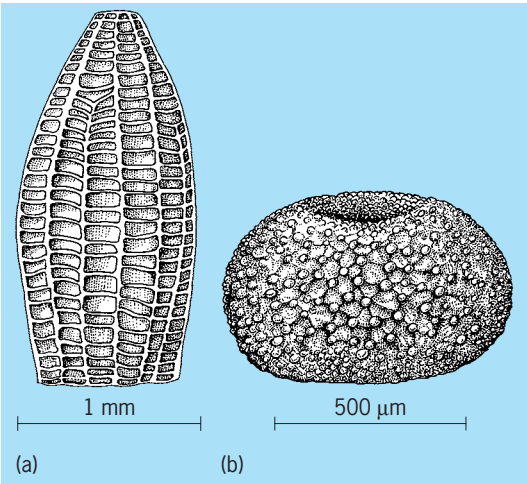

Fig. 5. Butterfly eggs. (*a*) *Satyrium acadica* (Lycaenidae). (*b*) *Anthocharis midea* (Pieridae).

Classification

The Lepidoptera were long divided into two suborders, the Rhopalocera, with clubbed antennae, notably the butterflies and skippers, and the Heterocera, with antennae of other forms, as in most moths. The butterflies and skippers form a major natural unit with over 15,000 species, still often referred to as the Rhopalocera. They are, however, closely related to various higher moths, and the simple division of Lepidoptera into Rhopalocera (butterflies) and Heterocera (moths) is not tenable.

The divisions Macrolepidoptera and Microlepidoptera often appear in older literature. The microlepidoptera comprise the more ancient groups, the most primitive of which have adults with chewing mouthparts and two pairs of similar wings (homoneurous moths), and include all of the families from Micropterygidae to Thyrididae listed in **Table 1**. Generally these are small insects, some with wingspans as small as 3 mm (0.1 in.). The macrolepidoptera include most of the larger moths and the butterflies, with wingspans typically in the 20–80-mm (0.8–3.2-in.) range, with a few reaching as much as 280 mm (11 in.). However, a few microlepidoptera, such as certain cossids and hepialids [wingspans up to 200 mm (7.9 in.) and more], are much larger than many macrolepidoptera, which can have wingspans as small as 8 mm (0.3 in.). In terms of common descent, the Macrolepidoptera do seem to form a natural group, but the microleps definitely do not, and the term Microlepidoptera is not used now, except

informally (for example, micromoth).

Recent comparative research on Lepidoptera indicates that the major groups evolved in a sequential stepwise fashion. This makes natural division into just two or three large groups impossible. The very oldest moths, those with functional adult mandibles, are placed in suborder Zeugloptera, with no more

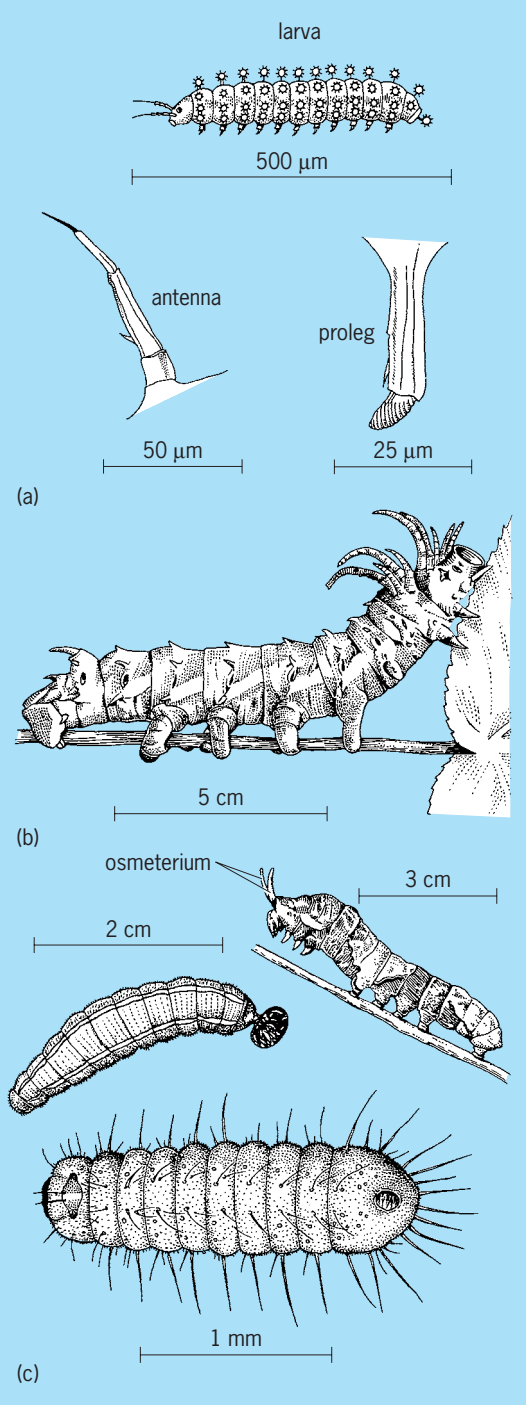

Fig. 6. Larvae. (*a*) Newly hatched larva of *Micropterix calthella* (Micropterygidae). (*b*) Hickory horned devil, larva of *Citheronia regalis* (Saturniidae). (*c*) Skipper and butterfly larvae: mature *Thorybes pylades* (Hesperiidae), showing neck; mature *Papilio cresphontes* (Papilionidae), osmeterium extruded; first instar *Satyrium liparops* (Lycaenidae), about 2 mm (0.08 in.) long.

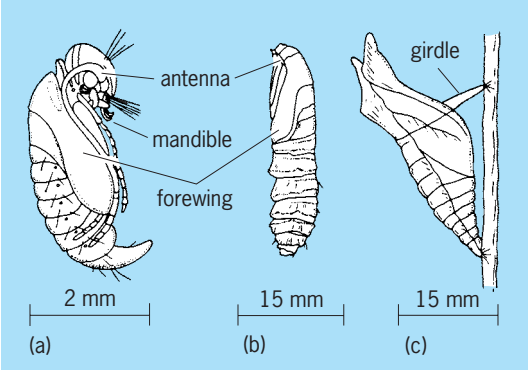

Fig. 7. Pupae. (*a*) *Dyseriocrania auricyanea* (Eriocraniidae). (*b*) *Sthenopis thule* (Hepialidae). (*c*) *Papilio troilus* (Papilionidae).

than about 200 known species. Two small and very obscure groups are considered to be the next oldest known Lepidoptera, the Agathiphagidae and the Heterobathmiidae. All three of these ancient groups lack sucking mouthparts. A major innovation was evolution of the proboscis, giving rise to the so-called Glossata suborder. The living moths that represent the earliest stage of this development are the Eriocraniidae. A further novelty was development of hollow scales, a feature also shared by the great majority of living Lepidoptera. Some of these Coelolepida then evolved muscles that coil the proboscis, and the obtect type of pupa. All Lepidoptera with coilable proboscides, obtect pupae, and larvae with crotchet-bearing abdominal prolegs belong to the Neolepidoptera. When this stage was reached, these Lepidoptera evidently still had two very similar pairs of wings (Homoneura), as seen in ghost moths (Hepialidae). A further step was the evolution of a marked difference between the fore- and hindwings (Heteroneura). A subgroup of these moths then underwent a major change in the female reproductive system, developing separate orifices for copulation (vulva) and egg laying (oviduct) that are connected by a seminal duct (Ditrysia). By this stage it would seem that the Lepidoptera were increasingly successful, evolving step by step many new families. The Apoditrysia evolved with a modification of the thorax; then fusion of the first four abdominal segments of the obtect pupa led to the Obtectomera (although validity of the group based on this character is open to question, as this condition also occurs in some lower Ditrysia, including the Yponomeutoidea, Gelechioidea, and Alucitoidea). Finally, evolution of the Macrolepidoptera occurred, in which all of the many species share a special type of wing base. The familiar butterflies are subordinate members of the Macrolepidoptera, nested inside this huge and highly diverse "crown" group.

It is evident that Lepidoptera classification cannot simply reflect a set of major alternatives, such as butterflies versus moths, or Macrolepidoptera versus Microlepidoptera, but needs to mirror the successive stepwise evolution of one group after another. Thus, not only are the butterflies Macrolepidoptera,

TABLE 1. Size, distribution, and common names of some families of Lepidoptera

Classification	Common name	Distribution	Number of species*
ZEUGLOPTERA			
Micropterygidae	Micropterygids	Cosmopolitan	3 (200)
OTHER NON-GLOSSATA			
Agathiphagidae	Agathiphagids	Queensland, eastern Pacific	(2)
Heterobathmiidae	Heterobathmiids	Southern South America	(10)
GLOSSATA			
Eriocraniidae	Eriocraniids	Holarctic	5 (24)
COELOLEPIDA			
Acanthopteroctetidae		North America, Peru, Crimea	4 (6)
Lophocoronidae	Lophocoronids	Australia	(6)
Neopseustidae	Neopseustids	India to Taiwan, southern South America	(10)
NEOLEPIDOPTERA			
Mnesarchaeidae	Mnesarchaeids	New Zealand	(14)
Hepialidae	Swift or ghost moths	Cosmopolitan	18 (500)
HETERONEURA			
Nepticulidae	Serpentine leaf miners	Cosmopolitan	75 (800)
Incurvariidae	Incurvariids	Mainly Palearctic and Australia	3 (100)
Prodoxidae	Yucca moths and relatives	North America (one species in southern South America)	75
DITRYSIA			
Tineidae	Clothes moths and relatives	Cosmopolitan	135 (3000)
Psychidae	Bagworms	Cosmopolitan	25 (1000)
Gracillariidae	Gracillariids	Cosmopolitan	235 (2000)
Yponomeutidae	Ermine moths	Cosmopolitan	65 (600)
Heliodinidae	Heliodinids	Cosmopolitan	21 (60)
Oecophoridae	Oecophorids	Cosmopolitan but largely Australian	225 (3200)
Coleophoridae	Case bearers	Cosmopolitan	110 (1400)
Gelechiidae	Gelechiids	Cosmopolitan	590 (4600)
Cosmopterigidae	Cosmopterigids	Cosmopolitan	(1650)
APODITRYSIA			
Megalopygidae	Flannel moths	New World, mainly tropical	11 (250)
Limacodidae	Slug moths	Cosmopolitan	50 (1000)
Zygaenidae	Foresters and burnets	Worldwide except New Zealand	20 (1000+)
Cossidae	Goat or carpenter moths	Cosmopolitan	45 (670)
Sesiidae	Clearwing moths	Cosmopolitan	120 (1200)
Castniidae	Castniids	Neotropical and Indo-Australian	(~150)
Tortricidae	Tortricids	Cosmopolitan	925 (5000+)
Alucitidae	Many-plume moths	Cosmopolitan	1 (135)
Pterophoridae	Plume moths	Cosmopolitan	130 (1000)
OBTECTOMERA			
Pyralidae	Pyralids	Cosmopolitan	500 (6,000+)
Crambidae	Snout moths	Cosmopolitan	700 (11,500+)
Thyrididae	Window-winged moths	Tropical	10 (760)
MACROLEPIDOPTERA			
Lasiocampidae	Tent caterpillars, lappet moths	Cosmopolitan except New Zealand; mainly tropical	30 (1500)
Bombycidae	Silkworm and allies	Tropical	1 (introduced) (350)
Saturniidae	Giant silkworms	Cosmopolitan	60 (1480)
Sphingidae	Sphinx, hawk, or hummingbird moths	Cosmopolitan	106 (1250)
Notodontidae	Prominents, puss moths	Cosmopolitan except New Zealand and Pacific	120 (2800+)
Noctuidae	Noctuids, owlets, underwings, millers	Cosmopolitan	2700 (35,000+)
Lymantriidae	Tussock moths	Largely African and Indo-Malayan, but with important Holarctic species	27 (2500+)
Arctiidae	Tiger moths	Cosmopolitan	200 (~6000)
Drepanidae	Hooktips	Holarctic	6 (650)
Geometridae	Measuring worms, loopers, cankerworms, carpets, waves, and pugs	Cosmopolitan	1200 (21,000)
Uraniidae	Uraniids	Tropical	(700)
Hedylidae	Hedylids	Neotropical	(40)
RHOPALOCERA			
Hesperiidae	Skippers, agave worms	Cosmopolitan	240 (3500)
Papilionidae	Swallowtails, bird-wings, parnassians	Cosmopolitan	27 (600)
Pieridae	Whites, sulfurs, orange-tips	Cosmopolitan	61 (1000)
Lycaenidae	Blues, coppers, hairstreaks	Cosmopolitan	138 (5000)
Riodinidae	Metal marks	Mainly Neotropical, but all major regions	1 (1250)
Nymphalidae	Brush-footed butterflies	Cosmopolitan	211 (6000)

*The first figure is the number of described species in North America north of Mexico. In most cases this figure is reasonably accurate. The figure in parentheses is a good estimate of the number of described species in the world. However, it is difficult to derive the total number of living Lepidoptera species from these figures. While in groups such as the Papilionidae more than 90% of existing species have almost certainly been recognized, in others, such as many families of Microlepidoptera, the figure may be well under 25%. Classification is based largely on N. P. Kristensen (ed.), 1999.

but at successively higher levels they are also members of the Obtectomera, Apoditrysia, Ditrysia, Heteroneura, Neolepidoptera, Coelolepida, and Glossata (Table 1). Butterflies are in reality just one of several advanced group of moths. Of all the major formed groups indicated in Table 1, the only group that the butterflies do not belong to is the suborder Zeugloptera, and the same is true for all the other groups listed, insofar as they are included within all the major divisions above their position in Table 1. Some of the most important Lepidoptera families are also listed in this table.

Homoneurous moths. In Table 1, the first nine families listed, up to the Heteroneura, have a primitive homoneurous wing condition, with the fore- and hindwings similar in shape and venation. They do not form a natural group, but it is convenient to deal with them together. Their fore- and hindwings are connected by a jugum and sometimes also by a frenulum. Mouthparts either are mandibulate, with the mandibles functional, vestigial, or absent (**Fig. 8**), or consist of a rudimentary proboscis. The females have a single genital opening, except in the Mnesarchaeoidea and Hepialoidea, which have two openings, one for copulation and the second for egg laying, but they lack the connecting seminal duct found in the Ditrysia. The pupae are exarate (free) or incomplete.

These primitive moths, making up less than 1% of Lepidoptera species, share a number of characteristics with the Trichoptera (caddis flies), to which the Lepidoptera are very closely related. In addition to the three groups described below, six other families of inconspicuous homoneurous moths are listed in Table 1, several of which were discovered only in the 1970s.

Superfamily Micropterygoidea. These moths make up one family, the Micropterygidae. They are small moths [up to 15 mm (0.6 in.) wingspan but usually much less], and the adults have asymmetrically toothed, functional mandibles and lack even a rudimentary proboscis. The galea is short and the adults feed on pollen or fern spores. The larvae, which feed variously on mosses, liverworts, angiosperm leaves, and possibly even fungal hyphae, mostly lack abdominal appendages, although some are very unusual in having eight pairs of simple abdominal prolegs.

It has been suggested that these moths are terrestrial Trichoptera or that they should be placed in a separate order, the Zeugloptera. However, characters of the wing venation, the broad well-developed scales with numerous ridges, and many other features indicate that these insects do belong to the Lepidoptera.

Superfamily Eriocranioidea. In this group of small moths [never more than 15 mm (0.6 in.) wingspan], the adult mandibles are greatly reduced and untoothed, and the galeae form a simple and abbreviated proboscis that cannot be coiled. These moths appear to represent the earliest stage in the evolution of the Glossata. Formerly, three or even more families were placed in this group, but the Neopseustidae, Mnesarchaeidae, and Acanthopteroctetidae are now considered to represent separate, further stages in glossatan evolution (Table 1). The single remaining family, the Eriocraniidae, have legless leaf-mining larvae. The adults fly in sunshine and drink sap from injured leaf buds. The females have a cutting (piercing) ovipositor used to insert their eggs into leaf buds.

Superfamily Hepialoidea. The hepialoids are medium-sized to very large moths possessing short, not

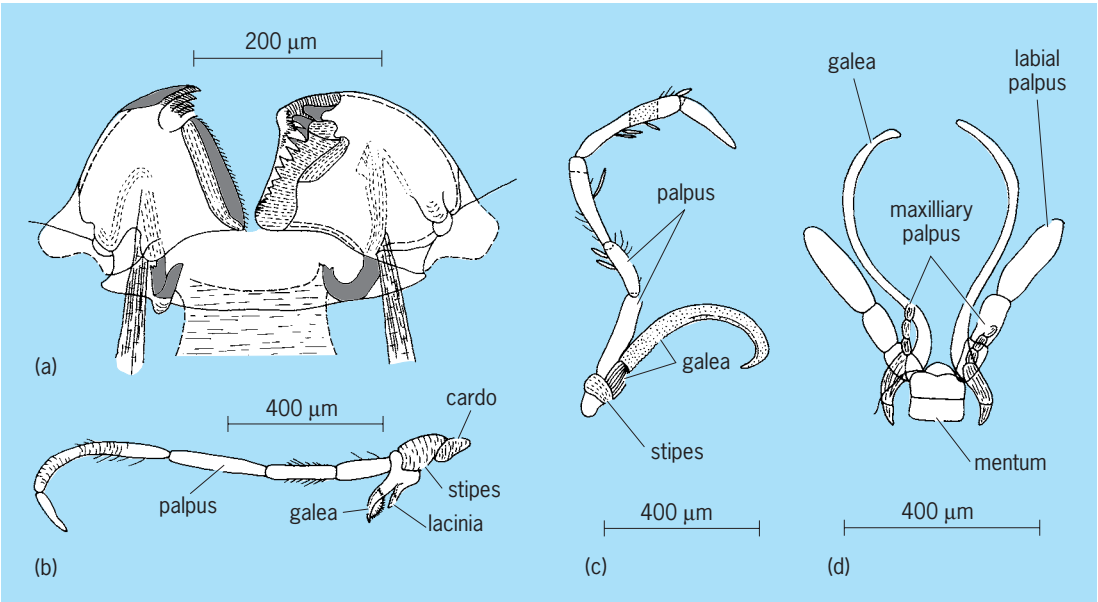

Fig. 8. Mouthparts of adult homoneurous moths. (*a*) Mandibles of *Sabatinca incongruella* (Micropterygidae). (*b*) Maxilla of *Micropteryx aruncella* (Micropterygidae). (*c*) Maxilla of *Eriocrania semipurpurella* (Eriocraniidae). (*d*) Maxillae and labium of *Mnesarchaea paracosma* (Mnesarchaeidae).

fully coilable proboscides that are probably nonfunctional. Together with the Mnesarchaeidae (Mnesarchaeoidea), they form the Exoporia, a group representing the first stage in the evolution of the Neolepidoptera. Hepialoid larvae are borers in stems or roots, either feeding in the tunnels they create or emerging to eat vegetation; some feed on decaying wood and fungi. The adults are rapid flyers and mostly crepuscular (active during twilight or preceding dawn), giving the common name swift moths or ghost moths. There are five families, the best known of which is the Hepialidae with about 500 known species. Some have wingspans approaching 250 mm (9.8 in.).

Heteroneura. Fore- and hindwings in the Heteroneura are markedly different in shape and venation. Usually they are connected by a frenulum and retinaculum. The sucking mouthparts are occasionally vestigial, but the majority of heteroneurans are able to feed on nectar, juices from rotting fruits, and other liquids. The females of all except the Nepticuloidea and Incurvarioidea have two genital openings. The pupae are obtect. The Heteroneura appear to be a major natural group that comprise all the rest of the Lepidoptera, including the Macrolepidoptera. This section deals with the heteroneurous microlepidoptera, leaving the macrolepidoptera to the third section.

Superfamily Nepticuloidea. Two very closely related families are included, Nepticulidae and Opostegidae. These tiny moths have wing spines and females have a single genital opening, but they differ from the Incurvarioidea in having reduced venation, a large eye-cap at the base of the antenna, and a short nonpiercing ovipositor. The larvae, with the exception of some gall-making species (insect galls are enlargements or swellings of plant tissue due to the feeding stimulus of the insects) of the genus *Ectoedemia*, are miners in leaves, bark, and rarely fruits. The pupae have the leg bases free. Many species of the genus *Nepticula* have a wing expanse in the 3–5-mm (0.1–0.2-in.) range, being the smallest insects in the order.

Superfamily Incurvarioidea. Six families are included in this group, with all members having an extensible, piercing ovipositor in the female, and a portable case in the last larval instar. As in the Nepticuloidea and homoneurous moths, the females have a single genital opening. The Incurvariidae have wings covered with microscopic spines, and almost complete venation. The larvae are seed, leaf, stem, or needle miners. In the subfamily Incurvariinae, the larva is first a miner and then a case bearer. With up to 300 species, the fairy moths (Adelidae) are the largest incurvarioid family. Male adelids have spectacularly long antennae, and some species make swarming flights in bright sunshine; the larvae feed on dead leaves and low plants.

The yucca moth, *Tegeticula yuccasella*, belongs to the family Prodoxidae, almost all members of which are North American. This small white moth has an obligatory relationship with yuccas. The females gather pollen with their specialized mouthparts and fertilize the yucca flower. The eggs are laid, using the piercing ovipositor, in the plant ovary, where the larvae grow by eating some of the developing seeds.

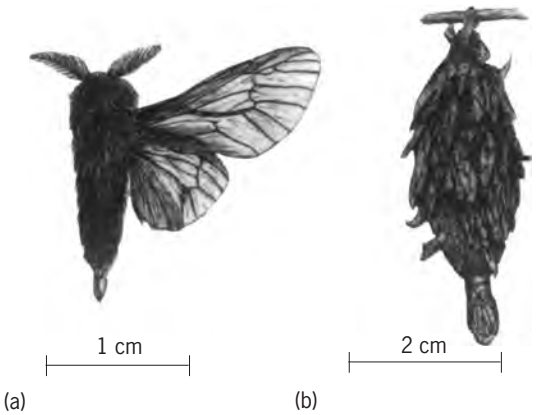

Fig. 9. ***Thyridopteryx ephemeraeformis*** **(Psychidae). (*a*) Male. (*b*) Male case.**

Superfamily Tineoidea. The Tineoidea formerly included a large and heterogeneous assemblage of moths, but only four tineoid families are now accepted, with the others groups variously reassigned. The tineoids can be regarded as the oldest or most primitive members of the Ditrysia.

The Psychidae comprise the remarkable bagworms. The males are hairy, strong-bodied, swift-flying moths with reduced mouthparts. They are relatively large, some with a wing expanse of about 25 mm (1 in.). The females are wingless, legless, and often sluglike. They live concealed in bag-shaped cases made by the caterpillars. The best-known North American representative is *Thyridopteryx ephemeraeformis* (**Fig. 9**). In this species, the larva fastens the bag to a twig and pupates within it. The vermiform (wormlike) female emerges from the pupa and moves to the bottom of the bag, where she is fertilized. This is accomplished from the outside by the highly specialized extrusible genitalia of the male. The female deposits her eggs in the bag, then drops to the ground and dies.

The Tineidae are a family of small, mostly drab moths that often have an erect tuft of yellowish scales between their antennae. The proboscis is reduced or vestigial, but the wing venation is fully developed and only rarely somewhat reduced; the ovipositor is generally long and extensible. The larvae do not feed on green plants but on a variety of substrates, including detritus, fungi, lichens, and notably keratin (wool, skin, etc.). The larvae of many species are case bearers. The best-known species are the clothes moths, which include the case-making clothes moth (*Tinea pellionella*), the webbing clothes moth (*Tineola bisselliella*), and the carpet moth (*Trichophaga tapetzella*); all are important pests whose larvae devour wool and other animal products. These three species have wing expanses of less than 25 mm (1 in.).

Superfamily Gracillarioidea. This group includes four ditrysian families that share a number of specialized

features. The Gracillariidae, by far the largest of the four in terms of included species, have lanceolate (shaped like the head of a lance) and widely fringed pairs of wings, well-developed proboscides, and nonextensible ovipositors; and they comprise the principal group of leaf-mining Lepidoptera. The young larvae are flattened and have bladelike mandibles with which they slash the cells of the leaf, sucking up the exuding juices. When full-grown, the larvae of some species are quite different, being normal in appearance and feeding on leaf parenchyma, either in a mine or externally. In *Phyllocnistis* the final instar does not feed, lacks legs and mouthparts (other than the spinneret), and pupates within its mine (a gallery it creates within the leaf tissue).

Superfamily Yponomeutoidea. This is an assemblage of rather primitive micromoths, with eight families somewhat uncertainly included. The Yponomeutidae represent a heterogeneous assortment of about 600 small, often brightly colored moths that lack ocelli. Most of the 80 or so species of *Yponomeuta* (ermine moths) have larvae that live in communal webs sometimes covering whole trees or shrubs. The Heliodinidae are a small family of tiny moths with ocelli and smooth heads that often have brilliant metallic spots on the forewings. The larvae are leaf miners, borers in stems, buds, or galls, or external leaf feeders; sometimes in communal webs.

Superfamily Gelechioidea. This is a very large group of some 15 families, representing a combined total of over 16,000 described species of micromoths, with wingspans in the range 4–70 mm (0.16–2.7 in.). The superfamily is represented worldwide, but certain families or subfamilies are mainly associated with particular zoogeographical regions. It is anticipated that many thousands more species of these moths await discovery. The group is characterized by the overlapping scales found on the dorsal surface of the proboscis, although this feature does recur in some Ditrysia. *See* ZOOGEOGRAPHY.

Oecophoridae are small to moderately small moths, some of which have a comb of bristles, the pecten, on the antennal scape (shaft). The larvae are varied in habits: some feed in webs or rolled leaves, whereas others are scavengers, notably in Australia where certain species specialize on the tough fallen leaves of *Eucalyptus*. *Stathmopoda* larvae feed on coccids, and some related species feed on spider eggs. *Hofmannophila pseudospretella* is an important pest of stored products, including grain, dried fruits, and furs.

Coleophoridae are small, narrow-winged moths. The young larvae of most species are leaf miners, but from the second instar they become case bearers, carrying shells made out of silk and bits of leaves, to which they add as they grow and within which they eventually pupate. Other coleophorids feed in seed heads, mine pine needles, or are inquilines in galls. *Holcocera pulverea* (subfamily Blastobasinae) is an important predator of the lac insect in India.

The Gelechiidae are a very large family of small to minute moths, usually with rounded forewings or, rarely, pointed ones. The hindwings are trapezoidal and often pointed, with the outer margin excavated or sinuate below the apex. The larvae include seed feeders, miners, borers, gall makers, and foliage feeders, and are known to attack over 80 different plant families. The group includes a number of economically important species, such as the *Angoumois* grain moth (*Sitotroga cerealella*), which infests grain both in the field and in storage, and the pink bollworm (*Pectinophora gossypiella*), an extremely important pest of cotton worldwide. The Cosmopterigidae, with over 1600 species, are the closest relatives of the Gelechiidae. While most are plant or detritus feeders, the larvae of *Euclemensia* are parasitoids, feeding internally in armored scale insects.

Superfamily Zygaenoidea. Currently classified as the first group of the Apoditrysia, the zygaenoids comprise a dozen families apparently linked by having a retractile larval head, and the second abdominal spiracle of the pupa covered by the wings. These moderately small to sometimes quite large moths have complete venation and in some cases a rudimentary proboscis; the wings are usually broad with short fringes. The larvae are stout and sluglike, and are exposed feeders.

The Himantopteridae are a curious group of about 40 species restricted to the Old World tropics. The hindwings of these small or middle-sized moths are very narrow with long ribbonlike tails, and the body and wings are covered with long hairs. The larvae of some species feed on trees (dipterocarps and oaks), but it has been suggested that others live in termite colonies. Supposedly the newly emerged moths escape when attacked by the termites, sacrificing their long hairs that easily pull free and their expendable tails. Recent work has failed to substantiate this, but the potentially fascinating biology of these moths is poorly known.

The Megalopygidae are restricted to the New World and are mainly tropical. The head and thorax are densely scaled, and the abdomen is very hairy. The larvae of some species are pests of palms, ferns, and various other plants, and their penetrating setae (they are often covered by long hairs) can cause dermatitis in humans. The family is considered to be related to the Limacodidae.

The Limacodidae are a family of heavy-bodied, often brightly colored, hairy moths (**Fig. 10**). The

Fig. 10. ***Prolimacodes badia*** **(Limacodidae).**

larvae are short and especially sluglike (slug caterpillars), with a large head concealed by the thorax, and furnished with various spines, tubercles, and even gelatinous warts. The thoracic legs are small or minute, and the abdominal prolegs are replaced by midventral suckers that are used, together with a fluid secreted by the spinneret, to stick the caterpillars tightly to foliage. The best-known North American form is the saddleback caterpillar (*Sabine stimulea*) which, like many limacodid larvae, has irritating (urticating) hairs.

The Zygaenidae are medium-sized, often brightly colored, mainly day-flying moths found in most regions (but not New Zealand), with over 1000 known species. They appear to be chemically protected, and if attacked many produce cyanide, a poison to which they appear to be resistant. Many zygaenids are involved in mimetic associations, notably the Oriental subfamily Chalcosiinae, some of which look very similar to aposematic (pertaining to colors of an organism that give warning of its special means of defense against enemies) butterflies such as Danainae. The larvae are rather stout and feed openly on plants. In some groups, notably the large genus *Zygaena*, the pupa is formed within a characteristic parchmentlike cocoon spun on a grass stem or similar substrate. *See* PROTECTIVE COLORATION.

Superfamily Cossoidea. Two families are included. Because the larvae bore tunnels in wood and smell rather strongly, the Cossidae are commonly called carpenter moths or goat moths. They are heavy-bodied insects, sometimes very large [maximum 235-mm (9.2 in.) wingspan], with the abdomen extending well beyond the hindwings. The mouthparts are largely rudimentary. The median vein (M, Fig. 4) is present and usually forked within the discal cell (a large cell in the central, or disc, part of the wing) of both wings. Although the larvae of some species feed externally on stems or roots, most are tunnelers, even in hardwood trunks. *Prionoxystus robiniae*, the best-known American cossid, is very destructive to a large variety of deciduous trees, and *Zeuzera coffeae* damages numerous species, including coffee. Like the Cossidae, the Dudgeonidae, a group of six stem borers found in the Old World tropics characterized by an abdominal hearing organ, also have a very strongly reduced proboscis.

Superfamily Sesioidea. This group of three families is thought to be very closely related to the Cossoidea. The Sesiidae are called clearwing moths because of the large, transparent, scaleless areas on the wings. Veins Sc and R_s on the hindwing appear to be absent, but are actually concealed in a unique costal fold that acts as an additional wing-coupling mechanism. Diurnal, with contrasting black and transparent patterns, or more brightly colored, many of these moths are excellent wasp mimics. The larvae are mostly borers, and many are economic pests, including the currant borer (*Synanthedon tipuliformis*), peachtree borer (*Sanninoidea exitiosa*), and squash-vine borer (*Melittia satyriniformis*). The South American *Synanthedon coccidivora* is carnivorous on coccids.

The Castniidae are large [wingspan range of 24–190 mm (0.9–7.5 in.)], diurnal, butterflylike moths with broad wings, clubbed antennae, and fusiform eggs. The larvae bore into monocots, and some are minor pests. A proboscis may be present or (rarely) vestigial. These moths have been considered to be related to the butterflies, but the resemblances must be due to convergence, and some are mimetic of butterflies and other moths. They are found in the Neotropical and Indo-Australian regions.

Superfamily Tortricoidea. These small, broad-winged moths comprise just a single very large family (over 5000 species), the Tortricidae, currently divided into three subfamilies. The best character for the group as a whole appears to be the large, flat ovipositor lobes of the females.

The hindwings of the moths belonging to subfamily Olethreutinae usually have a fringe of long hairs on the upper side, running along the basal part of the cubitus. The larvae are generally hidden feeders and live in rolled leaves, in foliage webbed together, or inside fruits. The subfamily contains a number of agriculturally very undesirable species, notably the codling moth (*Cydia pomonella*), a serious pest of apples and other fruits. The genus *Cydia* (formerly called *Laspeyresia*) also contains the curious Mexican-jumping-bean moth, *C. saltitans*. The violent movements of the larvae feeding inside are responsible for the jumping action of the beans.

Moths placed in the subfamily Tortricinae generally lack the long cubital fringe on the hindwing characteristic of the Olethreutinae. The spruce budworm (*Choristoneura fumiferana*), which belongs to the large tortricine tribe Archipini, is probably the most injurious tortricid. In many places, especially eastern Canada, the larvae of this moth have defoliated vast areas of coniferous forest. Members of the third and smallest subfamily, the Chlidanotinae, are largely tropical and relatively poorly known.

Superfamily Alucitoidea. Two small families of rather delicate micromoths are included. The pupae have immovable abdominal segments I–IV. The Alucitidae are the larger of the two families, with larvae that bore into buds, shoots, flowers, and fruits or make galls; the African *Alucita coffeina* is a pest of coffee. The adults of most species have the forewings divided into six featherlike plumes, and the hindwings into six or seven. Although this superfamily is now believed to be very closely related to the plume moths (Pterophoridae), the striking similarity of wing division appears to be a case of convergence.

Superfamily Pterophoroidea. The Pterophoridae are known as the plume moths. The wings of most species are divided into featherlike plumes, of which there are usually two in the forewing and three in the hindwing (**Fig. 11**), similar to the Alucitidae. Although the alucitoids are now thought to be closely related, the evidence is ambiguous; the Pterophoroidea may be closer to the Pyraloidea. The moths lack ocelli and have characteristic slender bodies and long legs, with the wings at rest held outstretched; segments I–IV of the pupae are movable. The larvae feed exposed or are borers. *Buckleria* is remarkable for feeding on insectivorous sundews

Fig. 11. ***Platyptilia carduidactyla*** **(Pterophoridae).**

(Droseraceae); the North American *Platyptilla carduidactyla* (Fig. 11) is a minor pest of artichokes.

Superfamily Pyraloidea. The Pyraloidea comprise one of the largest family-level divisions of the Lepidoptera, with over 16,000 named species, and the expectation is that this will rise to 30,000 or more when the taxonomy is worked out. Currently the group is divided into two families: the Pyralidae with 5 subfamilies and Crambidae with 15. All are linked by possession of a unique type of hearing organ located on the first abdominal segment. They are small to quite large moths, with a wing expanse in the range of 10–150 mm (0.4–6 in.), usually about 20–35 mm (0.8–1.4 in.). The hindwings often have three anal veins, while the legs are usually long and slender.

The Pyralidae differ from the Crambidae in having the case of the hearing organ almost completely closed. The relatively small subfamily Galleriinae includes the bee moth, or wax worm (*Galleria mellonella*), which lives in beehives. The larvae feed on the wax at night and destroy the combs; the species occurs throughout the range of the honeybee. Subfamily Pyralinae reaches its richest development in the Old World tropics. The meal moth (*Pyralis farinalis*) is a cosmopolitan pest of stored products. The Phycitinae comprise a large group of moths (almost 4000 species) in which the frenulum of the female, like that of the male, is a simple spine rather than a bundle of bristles. The larvae have very diverse habits, being leaf rollers, case bearers, borers, and stored-products pests, and some (genus *Laetilia*) are predacious on coccids. The Indian meal moth (*Plodia interpunctella*) is cosmopolitan, feeding on a wide variety of stored products, especially cereals. Another extremely important pest is the Mediterranean flour moth (*Anagasta kuehniella*), which infests cereals throughout the world. In contrast to these harmful species, the subfamily Phycitinae also contains *Cactoblastis cactorum*, imported from South America into Australia, Hawaii, South Africa, and other areas to help control introduced *Opuntia cactus*, which had ruined millions of acres of pasture. This successful program is an outstanding example of biological control. *See* ENTOMOLOGY, ECONOMIC; INSECT CONTROL, BIOLOGICAL.

Crambidae are separated from the Pyralidae by the wide aperture to the hearing organ case and differences in the male genitalia. Subfamily Crambiinae (snout moths) contains small insects common in marshes and grasslands. In this group the labial palpi (Fig. 1*a*) are long and porrect (extended forward), giving the adults a beaked appearance. Species of the small subfamily Schoenobiinae bore in the stems of marsh-living grasses, with *Scirpophaga* and *Rupela* being important rice pests. Members of the subfamily Nymphulinae are notable for being almost entirely aquatic, and in some species the larvae are able to swim from plant to plant. The larvae of some nymphulines have tracheal gills, and the pupae are usually enclosed in a cocoon below the surface, from which the adults emerge through the water. In the genus *Acentria*, most female adults are wingless and reproduce parthenogenetically, never emerging from the water; males and winged sexual females emerge periodically, swarm, mate, and disperse. With about 7400 species, the Pyraustinae are the largest group of pyraloids and include some relatively large moths with wingspans of 30 mm (1.2 in.) or more. This subfamily includes the infamous European corn borer (*Ostrinia nubilalis*), the grape leaf folder (*Desmia funeralis*), and numerous other economically damaging species.

Superfamily Thyridoidea. This mainly tropical group comprises a single family, the Thyrididae, with about 760 described species and many undescribed. They have often been included in the Pyraloidea, but they lack abdominal hearing organs and have a peculiar resting posture in which the body is raised up steeply and the wings held outstretched. Larvae of the Oriental genus *Glanycus* can stridulate (make audible sounds, produced by rubbing various body parts together), possibly a warning signal as these brightly colored moths appear to be chemically protected.

Macrolepidoptera. All of the remaining groups dealt with below are macroleps. Unlike the microlepidoptera, there is some morphological evidence that the macros form a natural group, and they are now formally recognized as such. The Macrolepidoptera are the crown group of the order and include some 100,000 different sorts—more than half of all known Lepidoptera species. The largest moths and butterflies are included, and even the smallest are larger than nepticulids and other micros. Because of their size, the larvae of most macros are external feeders, but there are exceptions.

Superfamily Lasiocampoidea. Two families, somewhat uncertainly united but both apparently related to the Bombycoidea, are included in this group. The Anthelidae are a small family restricted to the Australian region. Male Lasiocampidae always have bipectinate antennae (Fig. 3*b*), right to the tip; the females are usually similar but with shorter rami. This widespread, largely tropical family contains the eggar and lappet moths, and the familiar tent caterpillars of the genus *Malacosoma*. In this genus the larvae live together in large silken nests. Those of *M. americana* are a common sight on trees and shrubs in the eastern United States.

Superfamily Bombycoidea. A group of nine families of often stout-bodied large moths, including some of the most spectacular insects in the whole order. These families appear to be reliably grouped together by a number of characters, including the anterior fusion of the forecoxae in the last instar larvae, and a peculiarity of the forewing radial venation.

The family Eupterotidae represent a group of about 300 species, characterized by the last tarsal segment of the metathoracic leg of the female having a midventral row of spines. On this basis, the sexually dimorphic (having secondary morphological differences between the sexes) African genus *Hibrilides*, in which the females apparently mimic aposematic butterflies, has recently been included in this largely tropical family. *See* SEXUAL DIMORPHISM.

The Bombycidae include the commercial silkworm (*Bombyx mori*). The larvae of this totally domesticated species feed on mulberry leaves, and resemble hawkmoths in that they possess a small caudal horn. The silk from the white or yellow cocoons has been used since prehistoric times. The caterpillars are subject to the ravages of a disease known as pébrine, caused by the protozoon *Nosema bombycis*. They are also attacked by two viral diseases known as grasserie and flacherie. The Bombycidae are recognized as a group based on characters of the thorax and pupa, and worldwide the family includes about 350 species. *See* MICROSPORIDEA.

Saturniidae are the giant silkworms, medium-sized to extremely large insects having a single, often weak humeral vein in the hindwing. The mouthparts are reduced or vestigial, and the antennae are strongly bipectinate. As in the Lasiocampidae, the rami of the male antennae are conspicuously longer than those of the female. The frenulum is absent. The larvae usually bear spiny processes known as scoli. Pupation takes place in large silken cocoons, which are suspended from twigs or formed in leaf litter on the ground. Females of species *Attacus atlas* of the Himalayan foothills and *Coscinoscera hercules* of Australia have wingspans of up to 250 mm (9.8 in.) and possibly more, and these moths are thought to have the largest wing areas in the Insecta. A number of saturniids produce usable silk, which has long been gathered in the Orient. Among these are the Japanese oak silkworm (*Antheraea yamamai*) and the muga silkworm (*A. assamensis*).

The nearly 1500 known species of Saturniidae, which are mostly tropical, are currently divided among nine subfamilies. In North America, the best-known species of the subfamily Saturniinae are the cecropia (*Hyalophera cecropia*), polyphemus (*Antheraea polyphemus*), luna (*Actias luna*), promethea (*Callosamia promethea*), and cynthia (*Samia cynthia*). Likewise in North America, the regal moth, or hickory horned-devil (*Citheronia regalis*), and the imperial moth (*Eacles imperialis*) are the most familiar members of the subfamily Ceratocampinae. In subfamily Hemileucinae, the buck moths and relatives, the io moth (*Automeris io*) is noteworthy because its common larvae, like those of some other members of the subfamily, are equipped with urticating (poisonous) hairs that can cause considerable irritation to the unwary who handle them. The scoli of some tropical species release powerful anticoagulants that can induce bleeding from mucous membranes.

Sphinx, hawk, or hummingbird moths constitute the family Sphingidae, a worldwide group of some 1250 species. Morphological peculiarities of the larvae, pupae, and adults suggest that the family is a natural group. Mainly medium-sized to very large, these heavy-bodied moths have extremely rapid flight. The adults are mostly crepuscular or nocturnal, but a few genera are diurnal, including the hummingbird hawks. They have rather characteristic antennae that are thickened and have a pointed apex, while the proboscis is well developed and often extremely long. The wings are narrow, with the hindwing much shorter than the forewing. A frenulum is present, rarely reduced to a tubular thickening. The larvae are external feeders and usually have a conspicuous caudal horn. The pupa is formed in a cell in the ground or in a loose cocoon at the surface, and in some species the long proboscis is housed in a projecting case resembling a jug handle. One of the most widely distributed species is the white-lined sphinx (*Hyles lineata*). Well known also is the Old World death's-head sphinx (*Acherontia atropos*), the adults of which will enter beehives in search of honey. Of economic importance in the Western Hemisphere are the tomato hornworm (*Manduca quinquemaculatus*) and the tobacco hornworm (*M. sexta*), both of which are pests of solanaceous plants. *Erinnyis ello* is a serious pest of cassava, rubber, and ornamental euphorbs in South America, while *Cephonodes hylas* is a major pest of coffee in many parts of Africa and Asia.

Superfamily Noctuoidea. This huge superfamily contains some 50,000–70,000 species, by far the largest family-level group in the Lepidoptera, and likely to remain so even if the Gelechioidea and Pyraloidea are ever fully worked out. Although rather uniform in basic structure, which has made their classification very difficult, the noctuoids are remarkably varied in size, coloration, and biology. As measured by wingspan, they include the smallest macromoths, down to about 8 mm (0.3 in.), and the largest: the South American *Thysania agrippina* sometimes attains wing spans of 280 mm (11 in.). Currently eight families, some with numerous subfamilies, are included. All members of this great assemblage are linked by possession of a special metathoracic hearing organ that is capable of perceiving ultrasound. This organ not only may be used to detect and escape from bats but may also function in courtship signaling. *See* PHONORECEPTION.

Notodontidae are commonly called prominents, or puss moths. Together with two very small families (Oeonsandridae from Australia and the New World Doidae), they are separable from most other Noctuoidea by the apparently three-branched, or trifid, cubitus. The larvae are external feeders, and the pupa is formed in a cell in the ground or in a loose cocoon on the surface. The Nearctic *Datana*

ministra, the yellow-necked caterpillar, and *Schizura concinna*, the red-humped caterpillar, are pests on apples and other trees.

Noctuidae (once known as the Phalaenidae) are the owlet moths, an extremely large grouping of 35,000 or more mostly dull-colored, medium-sized moths. Evidence that this family, which is currently divided into 28 subfamilies, forms a natural group is unconvincing, and significant changes to the classification should be anticipated. As in most of the remaining Noctuoidea, the noctuid cubitus is usually quadrifid, but this is not a constant feature.

As exemplified by the underwing genus *Catocala*, the forewings of the Noctuidae are almost always dully and cryptically colored. At rest, the forewings cover the hindwings, which may be strikingly colored. Most moths attracted to lights at night belong to this family. The larvae are mostly exposed foliage feeders, but a few such as *Papaipema* are borers. Some *Eublemma* species prey on scale insects, one being an important enemy of the commercially valuable lac insect. Pupation is usually in the ground. The family includes many agricultural pests: the cutworms, *Euxoa* and *Peridroma*, attack a large variety of plants, as does the army worm, *Pseudaletia unipunctata*. The former derive their name from the habit of cutting off shoots at the surface of the soil without consuming them; the latter derive their name from the fact that they often appear in vast numbers. The exceedingly important pest *Helicoverpa armigera* is variously known as the corn earworm, cotton bollworm, and tomato fruitworm.

Lymantriidae are the tussock moths. Evidence that this group of medium-sized moths form a natural group is largely lacking, but the eversible middorsal gland on abdominal segment VI of the larva may be a reliable character. The antennae of the males are broadly pectinate. The end of the female abdomen often has a tuft of detachable hairs which are deposited on top of the eggs. The hairy larvae usually have prominent toothbrush tufts. The most familiar species is the infamous European gypsy moth (*Lymantria dispar*) which, along with the destructive brown-tail moth (*Euproctis chrysorrhoea*), was imported into New England in the latter half of the nineteenth century.

Arctiidae are a relatively large family of often strikingly colored, frequently heavy-bodied insects, the tiger moths. The arctiids are characterized by a pair of eversible pheromone glands on the dorsum of the female abdomen, connected with the anal papillae. The larvae are often very hairy and feed either exposed or in webs. Larval hairs are incorporated into the silken cocoon. In North America, the best-known member of the subfamily Arctiinae is probably the banded wooly bear caterpillar (*Pyrrharctia isabella*). There is a widespread misconception that the banding pattern of this larva predicts the severity of the coming winter. In Europe, the wooly bear of the spectacular garden tiger moth (*Arctia caja*) is equally well known. The larvae of the Lithosiinae are unusual in feeding on algae, lichens, liverworts, or mosses. The third arctiid subfamily currently recognized, the Syntominae, includes many remarkable wasp mimics. Despite their great beauty and fascinating biology, the Arctiidae remain poorly known, and many discoveries can still be made.

Superfamily Drepanoidea. This relatively small superfamily includes just two families, linked by the unusual form of the larval mandibles, which have a large flat lateral area marked off beneath by a distinct ridge. The group is thought to be closely related to the Geometroidea. Epicopeiidae are a small group of brightly colored day-flying species that appear to mimic a variety of other diurnal Lepidoptera; they are known only from the Palearctic and tropical Asia. The more familiar Drepanidae include about 650 species divisible into three subfamilies, of which the Drepaninae are the largest and include the hooktip moths, named for their forewing shape. The Drepanidae are characterized by an abdominal hearing organ located between tergum I and sternum II of a type not found in any other Lepidoptera, including the Geometroidea.

Superfamily Geometroidea. Three families are included, the somewhat butterflylike Sematuridae (about 40 species, which lack abdominal hearing organs), the Uraniidae (about 700 species, also somewhat butterflylike but with abdominal hearing organs), and the Geometridae (a very large group of more than 20,000 species characterized by the unique form of the abdominal hearing organs). The three families are linked together by a character of the larval spinneret.

The tropical Uraniidae include some slender-bodied, brilliantly colored, diurnal insects that lack a frenulum and are often mistaken for butterflies. The multiple-tailed metallic Madagascan sunset moth, *Chrysiridia madagascariensis*, is considered by many to be the most beautiful lepidopteran. It also offers a fascinating example of asymmetry: almost all Lepidoptera have highly symmetrical wing patterns, but this is not the case in *Chrysiridia* (and some other uraniids) in that the black spots on the hindwing orange area, and the black markings on the forewing green areas, invariably differ right to left. Uraniid larvae are diverse in form but have a complete set of prolegs.

The Geometridae include the measuring worms, loopers, and cankerworms, which generally turn into small and medium-sized moths with slender bodies and relatively broad wings. The females are occasionally wingless (apterous). The larvae have reduced or absent anterior prolegs, and usually only those on segments VI and X are well developed. They proceed with a characteristic looping (earth-measuring) motion. The larvae are usually exposed feeders, but sometimes live in folded leaves and often bear a striking resemblance to twigs. This resemblance is enhanced by their habit of resting with the body held rigid at an angle to the branch while holding on with the terminal prolegs. Pupation takes place in the ground or in a weak silken cocoon.

This very large family is currently divided into eight subfamilies. Among the economically

important Nearctic geometrids are the spring and fall cankerworms, *Paleacrita vernata* (subfamily Ennominae) and *Alsophila pometaria* (subfamily Alsophilinae), which attack a large variety of trees. The famous peppered moth (*Biston betularia*), so much studied regarding industrial melanism (below), is another member of the Ennominae. Also of interest is *Idaea bonifata* (subfamily Sterrhinae), the herbarium moth, whose larvae are pests in collections of dried plants. The beautiful emerald moths belong to the subfamily Geometrinae.

Superfamily Hedyloidea. This is a small group of about 40 mothlike species with a curious resting posture, long placed in the Geometridae but recently considered to be the likely closest relatives of the Rhopalocera (skippers and true butterflies). Their early stages are remarkably similar in some ways to those of Papilionoidea; the adults stand on only four legs like Nymphalidae; and research on their DNA confirms a close relationship. The Hedylidae comprise just one genus, *Macrosoma*, found in Central and South America.

Superfamily Hesperioidea. One rather large family, the Hesperiidae, is recognized, with about 3500 known species and probably many more unrecognized ones. The skippers, or skipper butterflies, are small to moderately large, heavy-bodied, mostly diurnal insects. The clubbed antennae are set wide apart on the head and are bent, curved, or reflexed at the tip. Forelegs are fully developed and bear a large spur (epiphysis). The forewings usually have all veins arising separately from the discal cell, and the frenulum is absent except in the male of *Euschemon*. Near the base of the hindwing, a small patch of specialized scales apparently occurs in all species. The larvae have a prominent constriction, or neck, behind the head, and often live in leaves drawn together by silk. However, the caterpillars of the giant skippers (*Megathymus* group), which are borers in yucca and agave in the Americas, lack this neck. The pupa is usually enclosed in a slight cocoon.

The name skipper refers to the typically rapid erratic flight. American species include the silver-spotted skipper (*Epargyreus clarus*) the sachem skipper (*Atalopedes campestris*). In Mexico, agave caterpillars are fried and canned for human consumption. The Australian *Euschemon rafflesia* was once placed in its own subfamily, largely because of the presence of a frenulum in the male, but it is now included in the Pyrginae, a group of some 1000 species and one of six subfamilies into which the skippers are currently divided. New molecular studies based on DNA are expected to bring many improvements to skipper classification.

Superfamily Papilionoidea. These so-called true butterflies, together with the Hesperioidea, make up the Rhopalocera. The two groups share some specialized features in common, are usually regarded as each other's closest relatives, and are together perhaps most closely related to the Hedyloidea. However, it is possible that the hedylids will eventually prove to be more closely related to one of the two Rhopalocera superfamilies. The Papilionoidea are small-to-large diurnal insects [a few are very large, as measured by wingspan, sometimes attaining 250 mm (9.8 in.)], typically with clubbed antennae that are rounded at the tip and not bent or reflexed. The forewings always have two or more stalked veins arising from the discal cell. The frenulum is always absent. The larvae have no constriction behind the head and are usually exposed feeders. The pupa is naked except in the apollos (*Parnassius*) and their relatives and some Lycaenidae, and is most often suspended head-down by caudal hooks (the cremaster) or held semi-upright by a silken girdle. About 14,000 species are currently recognized worldwide, divided among five families.

The Papilionidae are the family that includes the swallowtails and parnassians. These are the only true butterflies with a fully developed foreleg bearing an epiphysis. The hindwings have only one well-developed vannal vein, except in the anomalous Mexican *Baronia brevicornis* (subfamily Baroniinae), while the larvae always have a unique eversible forked organ on the prothorax, the osmeterium. This organ dispenses a disagreeable odor that seems to function as a defense mechanism. The pupa is typically girdled. In the boreal genus *Parnassius* (subfamily Parnassiinae) and some other members of the family, females may be found with a horny pouch, or sphragis, at the tip of the abdomen. This is secreted by the male during copulation and covers the genital opening of the impregnated female to prevent access by rival males. The commonest genera—*Papilio*, *Graphium*, *Battus*, and *Parides* (subfamily Papilioninae)—contain large and attractive species, many of which have the characteristic tails that give the family its name. The birdwing butterflies, *Ornithoptera*, *Trogonoptera*, and *Troides* (Papilioninae), are among the largest and most beautiful species. In North America, the eastern tiger swallowtail (*Papilio glaucus*) has dichromatic females, one form being black-and-yellow-striped like the male, and the other being entirely dark brown or black, apparently mimicking the protected Aristolochia swallowtail (*Battus philenor*). The larva of the orange dog (*Papilio cresphontes*) is sometimes injurious to citrus, and in the Old World there are a number of species that can be serious citrus pests, including the Oriental lime swallowtail (*Papilio demoleus*), recently introduced into the Caribbean.

The Pieridae include the familiar whites, sulfurs, and orange tips. These butterflies appear to have unique types of abdominal base and antennae. The forelegs are completely developed in both sexes but lack an epiphysis. The tarsal claws are bifid, while many species exhibit striking sexual dimorphism. However, most are basically white, yellow, or orange. The larvae feed on plants of the families Brassicaceae (mustards and cabbages), Fabaceae (beans, peas, and other legumes), and Capparidaceae (capers). The pupae are suspended by a girdle, as in the Papilionidae.

Butterflies of this family are often extremely abundant. The giant sulfurs (*Phoebis*) and others often

migrate in huge swarms. The larvae of the Neotropical species *Eucheira socialis* are gregarious, living together in a nest; and in Mexico, where their larvae are harvested for food, there is a population of this butterfly that is semidomesticated. The cabbage worm or small white (*Pieris rapae*) is one of the most economically important butterflies, attacking cultivated crucifers in Europe and North America. The alfalfa butterfly (*Colias eurytheme*) is a major pest in America.

The Lycaenidae include the blues, gossamers, hairstreaks, harvesters, and coppers. This large family, with about 5000 known species and many still to be recognized, is very closely related to the Riodinidae, and the latter have often been included within the lycaenids. All of these butterflies have a unique thorax in that the lamella of the mesodiscrimen is not complete to the furca. In male Lycaenidae the prothoracic legs, which always lack tarsal claws, are functional; but in Riodinidae they are reduced and not used for walking. In both families the eyes are emarginate at the antennal bases or at least contiguous with them. The larvae are mostly flattened, and the pupa is usually girdled. The larvae often have a mutualistic relationship with ants, which tend and protect them in return for honeydew which they secrete; however, a number are predacious on ant larvae or on plant bugs attended by ants. The ant organs of the Riodinidae differ in structure from those of the Lycaenidae, and recent molecular work confirms that the two groups are separate but are each other's closest relatives.

Lycaenidae are small to medium-sized insects that are often metallic blue or green on the upper surface. Species such as *Brephidium exilis* and *Zizula hylax* (subfamily Polyommatinae) have a wingspan of about 15 mm (0.6 in.) and are among the smallest butterflies, all of which are lycaenids. The largest lycaenid is the Indo-Australian *Liphyra brassolis* (subfamily Miletinae), with a wing expanse of about 75 mm (3 in.); its extraordinary armored caterpillars live in weaver ant nests and prey on their larvae. The harvesters (*Miletus, Spalgis, Fenesica*; Miletinae) have predacious larvae that feed on homopterans attended by ants, such as aphids, membracids, and coccids. In the United States, the gray or common hairstreak, *Strymon melinus* (subfamily Theclinae), has at times been an important pest of hops, cotton, and other crops. In various parts of the world, a number of lycaenids are significant pests, notably of beans and other legumes, often feeding on flowers or boring in pods. Species making up the relatively small group of copper butterflies (subfamily Lycaeninae) do not have associations with ants. Widespread in the Holarctic and parts of Asia and Africa, they have a remarkable distribution that includes endemic species in New Zealand, with the closest other members of the group in New Guinea.

The Riodininae, or metal marks, are primarily Neotropical. They are small butterflies, often brilliantly colored. Like the Lycaenidae, many have associations with ants, but it is currently thought that the Riodininae have probably evolved independently. The extraordinary gray and semidiaphanous Peruvian species *Styx infernalis* was formerly placed in its own family or subfamily, but recent discovery of its early stages confirms that it belongs to the Euselasiinae, a group of about 150 species. The 100 or so species that make up the Nemeobiinae form the Old World representatives of the family, and include the European duke of burgundy fritillary (*Hamearis lucina*). The third subfamily, Riodininae, are exclusively American, and are by far the largest group, exhibiting an extraordinary diversity of color and form, including the beautiful *Helicopis* butterflies with six tails on the hindwings and raised underside silvery markings that look like drops of water.

The Nymphalidae are a family of four-footed butterflies, having the prothoracic legs greatly atrophied in both sexes, especially the males, and useless for walking. However, the best character for bringing all 6000 or so included species together is the tricarinate antenna, which has three parallel ridges running along their entire inner-lower surface of the antennal shaft. The female abdomen of those species examined has a von Siebold organ, a structure unknown elsewhere in the butterflies. Although subdivision of this very familiar group of butterflies has been a source of considerable disagreement over the last 200 years, advances in the application of DNA sequence data, coupled with systematic reviews of their morphology and greater knowledge of the early stages, has recently led to greater clarity and an emerging consensus.

The Libytheinae are a small subfamily of less than 20 species divided into two genera, *Libythea* and *Libytheana*. Their long, forwardly directed labial palpi have earned them the name snout butterflies. The eggs and caterpillars of these little butterflies are superficially similar to those of the Pieridae. Formerly often placed as a separate family, the snouts are now widely accepted as the first side branch of the nymphalid family tree.

The next side branch now appears to be the Danainae, a worldwide group of usually warning-colored butterflies which have an acrid taste to humans. Supposedly distasteful to most predators, they appear to be primary models in a great number of mimicry complexes. The weak-flying Neotropical glasswings were formerly placed in their own subfamily, the Ithomiinae, but they are structurally and biologically very similar to the true danaines, and are now included by some in the same subfamily. Many have broad transparent areas in the wings in which the scales are reduced to short hairs; others have bright yellow, black, and orange tiger patterns. The widely distributed and migratory monarch, or milkweed, butterfly (*Danaus plexippus*) is the best-known representative of the subfamily. In North America, it is mimicked by the viceroy (*Limenitus archippus*), a member of another nymphalid subfamily, the Limentidinae. In the Indo-Australian tropics, the largest group is the crow butterflies (*Euploea*), many species of which are black or largely so.

The very large cosmopolitan subfamily Satyrinae, containing the woodnymphs, meadow browns,

graylings, and arctics, is characterized by the typically weak bouncing flight of the imagines (plural form of imago, which is the sexually mature, usually winged stage of an adult insect) and the bladder-like swellings to the basal sections of the forewing veins of most species. The larvae feed primarily on grasses and other monocots, with a few on lower plants such as lycopods. *Erebia* and *Oeneis* penetrate extremely inhospitable arctic and alpine situations, and numerous other satyrines abound in Patagonia, in the high Andes, and in many other montane regions, where many species still await discovery. Subfamily Morphinae is a much smaller but entirely tropical group of large butterflies closely allied to the satyrines. The brilliant metallic-blue species of the genus *Morpho* are among the most beautiful insects. The owl butterflies (*Caligo*) have undersides bearing large and very conspicuous eyespots, and also belong to this group. Two more subfamilies that seem to belong with the satyrine-morphine group are the Sino-Himalayan Calinaginae and the pantropical Charaxinae. This last group includes some of the fastest flying and most beautiful butterflies. Most but by no means all species in this entire assemblage feed as larvae on monocots and are strongly attracted as adults to fallen and fermenting fruit.

The crown group of the Nymphalidae comprises a further series of subfamilies, of which the Heliconiinae and Limenitidinae may form the first major branch. Well-known heliconiines include *Argynnis* and *Boloria* (temperate region fritillaries); *Vindula*, *Cethosia*, and *Acraea* in the Old World tropics; and *Heliconius*, famous for longevity and remarkable mimetic associations, in the Neotropics. The limenitidines comprise *Limenitis* (white admirals) together with numerous tropical genera and species. The final series of subfamilies, which may form a terminal group, include the Cyrestinae (tropical map butterflies), Biblidinae (eighty-eights and numerous tropical species), Apaturinae (purple emperors and hackberry butterflies), and Nymphalinae. This last subfamily includes some of the most familiar butterfly genera, such as *Polygonia* (angle-wings and commas), *Nymphalis* (tortoise-shells), *Vanessa* (red admirals and thistle butterflies), *Junonia* (buckeyes), and *Euphydryas* (checkerspots). The subfamily is also represented in the tropics by numerous genera and species, including the genus *Kallima*, or Indian leaf butterflies, the undersides of which bear a remarkable resemblance to dead leaves. *See* PROTECTIVE COLORATION.

Biological Aspects

The Lepidoptera are a group on which much research remains to be done. A great deal is still unknown about the genetics, developmental biology, physiology, and ecology of these insects despite the fact that butterflies and moths have proved useful as experimental animals in all of these fields. Economically, the larvae of many species are injurious to certain crops (**Table 2**), often causing severe losses; while a few, such as the silkworm (*Bombyx mori*), are the basis of significant industries. In recent years, trade in living tropical butterflies has grown rapidly for the butterfly house industry (insect zoos). A small number of Lepidoptera species cause medical problems, notably larvae with irritating spines and hairs, some of which can cause severe dermatitis.

Ecology and distribution. The variation in larval habits has been discussed under the different families. Lepidoptera of all stages are subject to the attacks of a large number of predators, including birds, mammals, lizards, frogs, and spiders. They also must be wary of rapacious insects, such as dragonflies, preying mantids, pentatomid bugs, asilid flies, and vespoid wasps. Certain wasps (family Sphecidae) paralyze caterpillars, lay eggs in them, and place them in specially constructed cells. Others (braconids and ichneumonids) place their eggs in the caterpillars without paralyzing them. In both cases the caterpillars serve as food for the growing wasp larvae inside them. Many chalcid wasps oviposit (lay or deposit eggs) in the eggs of Lepidoptera. *See* PREDATOR-PREY INTERACTIONS.

One mite, *Myrmonyssus phalaenodectes*, infests the hearing organs of a variety of moths; and another, *Otopheidomenis zalestes*, is restricted, so far as is known, to the genus *Zale* (family Noctuidae). Pseudoscorpions of the genera *Atemnus*, *Stenowithius*, and *Apocheridium* have been found on adults of various species. Lepidoptera are also subject to viral, bacterial, protozoan, and fungal infections. *See* INSECT PATHOLOGY.

The Lepidoptera penetrate almost every place on Earth with the exception of Antarctica. Butterflies of the genus *Boloria* have been taken at Alert on northern Ellesmere Island, about 400 mi (650 km) from the North Pole. Arctic and alpine tundra areas normally support a Lepidoptera fauna that, although relatively poor in species, is rich in numbers. After rains, deserts are often alive with butterflies and moths, but tropical forests are by far the richest in species. One of the strangest habitats occupied by a lepidopteran is the Neotropical three-toed sloth, on which the larvae of the sloth moth (*Bradypodicola hahneli*; Pyralidae) live, feeding on the algae that grow in the sloth's hair. The following list indicates some habitats of larval Lepidoptera.

1. **Plant associations, terrestrial**
   - *a*. Feeding exposed on foliage, flowers, or plant lice
   - *b*. Feeding on foliage in web or case
   - *c*. Leaf and needle mining
   - *d*. Living under bark
   - *e*. Boring in stem, root, fruit, or seeds
   - *f*. In soil, feeding on roots
   - *g*. Boring in or feeding on fungi, mosses, liverworts, club mosses, lichens, and ferns
   - *h*. Living in dried plant products, such as cereals, flour, dried fruit, refuse, and herbarium specimens

2. **Associations with social insects**
   - *a*. In beehives

TABLE 2. Important species of Lepidoptera injurious to crops

Family and species	Damage by larvae
Sesiidae	
Peach-tree borer (*Synanthedon exitiosa*)	Feeds under bark; the most important insect injurious to peach
Squash-vine borer (*Melittia satyriniformis*)	Infests various cucurbits; often very destructive
Gelechiidae	
Angoumois grain moth (*Sitotroga cerealella*)	Feeds on grain in storage and in the field; imported from Europe
Pink bollworm (*Pectinophora gossypiella*)	Feeds on cotton bolls, and causes failure of blossoms to open; most important insect injurious to cotton
Tineidae	
Case-making clothes moth (*Tinea pellionella*); webbing clothes moth (*Tineola bisselliella*)	Clothes moth larvae infest woolen products, upholstery, furs, and other dried animal products
Carpet moth (*Trichophaga tapetzella*)	Damage similar to that by clothes moths but less common
Tortricidae	
Spruce budworm (*Choristoneura fumiferana*)	Defoliator of vast areas of coniferous forest
Ugly-nest tortricids, leaf rollers, etc. (*Cacoecia* spp., *Tortrix* spp.)	Feed on leaves and fruits of apples and other fruits, strawberry, shade trees, ornamental plants, etc.
Codling moth (*Cydia pomonella*)	Larva bores in apples and some other fruits; most important pest of apples
Oriental peach moth (*Cydia molesta*)	Feeds on fruits and twigs of peaches, plums, etc.
Pyralidae	
Bee moth (*Galleria mellonella*)	Destroys combs in neglected beehives
Indian meal moth (*Plodia interpunctella*)	Widespread and very important pest of dried fruit and animal products
Mediterranean flour moth (*Anagasta kuehniella*)	Infests flour, stored grain, cereals
Crambidae	
Oriental rice borer (*Chilo simplex*)	Very destructive to rice in Asia
Grape leaf folder (*Desmia funeralis*)	Sometimes an important defoliator of grapevines
Geometridae	
Spring cankerworm (*Paleacrita vernata*), fall cankerworm (*Alsophila pometaria*)	Defoliate fruit and shade trees in outbreak years; spring cankerworm overwinters as a pupa, fall cankerworm overwinters as an egg
Sphingidae	
Tomato hornworm (*Manduca quinquemaculatus*); tobacco hornworm (*M. sextus*)	Large larvae are most conspicuous pests of tomato and tobacco; also feed on other solanaceous plants
Lasiocampidae	
Tent caterpillars (*Malacosoma* sp.)	Serious defoliators of forest, shade, and orchard trees
Arctiidae	
Fall webworm (*Hyphantria cunea*)	Feeds on wide variety of forest, shade, and fruit trees
Lymantriidae	
European gypsy moth (*Lymantria dispar*)	Important defoliator of forest, shade, and fruit trees in New England
Brown-tail moth (*Euproctis chrysorrhoea*)	Similar to gypsy moth
Noctuidae	
Cutworms (*Euxoa* sp., *Peridroma* sp., etc.)	Numerous species damage many crop plants
Army worm (*Pseudaletia unipunctata*)	Varies tremendously in abundance; in epidemic years it does great damage to corn and other grasses
Corn earworm, cotton bollworm, tomato fruitworm, tobacco budworm (*Helicoverpa armigera*)	Exceedingly important cosmopolitan pest attacking many cultivated plants
Pieridae	
Cabbage butterflies (*Pieris* spp., especially *P. rapae* in United States and Europe)	Pests of various cultivated crucifers
Alfalfa butterfly (*Colias eurytheme*)	Important pest of alfalfa in southwestern United States

b. In ant nests
c. Possibly in termite nests

3. **Associations with other animals**
a. In stored animal products, such as woolens and feathers
b. In hair of three-toed sloths, feeding on algae
c. In plant lice, as parasitoids

4. **Aquatic associations**
a. Feeding on, boring in, or mining in aquatic plants
b. On rocks, feeding on microorganisms

In temperate zones, different species hibernate at every stage of the life cycle. Individual, seasonal, and geographic variations are common and striking in the Lepidoptera. Color and size are characters most frequently affected, but wing shape, venation, genitalia, and other structures can also be involved. The number of broods per season often differs with geographic location.

Behavior and physiology. Migration is the most spectacular behavior occurring in the order. It is most frequent in the butterflies, such as *Phoebis*, *Danaus*, and *Libytheana*, but is also known in moths such as *Chrysiridia* and the bogong moth of Australia (*Agrotis infusa*). Huge migratory swarms of Lepidoptera are frequently reported in many parts of the world; but this phenomenon, as well as the related communal roosting of adults, daily use of flyways, hill topping, and so forth, is poorly understood. Aggregations of butterflies at mud puddles are a frequent sight, and both moths and butterflies have been observed pumping—sipping water steadily from a puddle and ejecting it as a stream of drops or fine spray from the anus.

Evolution and genetics. There is no extinct family in this order, but it must be acknowledged that the known fossil record is quite poor. The family Eosetidae, based on *Eoses triassica* from the Triassic of Australia, was originally placed in the Lepidoptera, but its position is very doubtful. The earliest fossil lepidopteran is a homoneurous micromoth from the Lower Jurassic (*Archeolepis mane*), and so it seems possible that the order did first differentiate in the Triassic. *Archeolepis* was originally placed in its own (extinct) family, but this is no longer accepted, mainly because not enough characters can be appreciated to reliably locate it in, or exclude it from, one of the living families.

The phenomena of mimicry and protective resemblance are widespread in the Lepidoptera. It has been demonstrated that eyespot patterns in certain species elicit escape responses in passerine birds and are thus an effective protective device.

Lepidoptera have advantages for many kinds of genetic investigation, especially for the study of populations. The appearance of dark forms of various moths in heavily industrialized areas that have been polluted and blackened by soot is a widespread phenomenon. While this industrial melanism, most extensively studied in the geometrid *Biston betularia*, is now known to involve complex and shifting selection pressures, it remains one of the best-known examples of evolution in action. *See* POPULATION GENETICS.

Paul R. Ehrlich; R. I. Vane-Wright

Bibliography. C. L. Boggs, W. B. Watt, and P. R. Ehrlich (eds.), *Butterflies: Ecology and Evolution Taking Flight*, 2003; M. M. Douglas, *The Lives of Butterflies*, 1986; J. D. Holloway, G. Kibby, and D. Peggie, *The Families of Malesian Moths and Butterflies*, 2001; N. P. Kristensen (ed.), *Lepidoptera, Moths and Butterflies*, 2 vols., *Handbook of Zoology*, vol. 4 (parts 35, 36), 1999/2003; H. F. Nijhout, *The Development and Evolution of Butterfly Wing Patterns*, 1991; R. Preston-Mafham and K. Preston-Mafham, *Butterflies of the World*, 1988; R. Robinson, *Lepidoptera Genetics*, 1971; M. J. Scoble, *The Lepidoptera*, 1992; N. Wahlberg et al., Synergistic effects of combining morphological and molecular data in resolving the phylogeny of butterflies and skippers, *Proc. Roy. Soc. London B*, 272:1577–1586, 2005; M. Young, *The Natural History of Moths*, 1997.

Lepidosauria

A subclass of living and extinct diapsid reptiles. Lepidosauria and their immediate ancestors constitute the Lepidosauromorpha, one of the two major clades of the Diapsida. By definition, Lepidosauria includes the last common ancestor of the living squamates (lizards, snakes, and the limb-reduced burrowing amphisbaenians) and the New Zealand tuatara (*Sphenodon*) and all of its descendants. This is a narrower definition than that used in older literature, where Lepidosauria was used as a catchall group for all non-archosaurian diapsids. Lepidosaurs differ from archosaurs in that the lower temporal arcade (the inferior border of the lower temporal fenestra) is typically incomplete, there is never an antorbital fenestra in front of the orbit or a mandibular fenestra in the lower jaw, and the teeth are generally fused in position (acrodont or pleurodont) rather than implanted in sockets (thecodont). Other lepidosaurian characters include a specialized skin-shedding mechanism, paired male hemipenes (copulatory organs that characterize lizards and snakes), fracture planes in the caudal vertebrae that allow the tail to be shed if grabbed by a predator, and specialized knee, foot, and ankle joints that improve locomotion. As in mammals, the ends of lepidosaurian long bones develop separate centers of ossification (epiphyses) that fuse to the shaft at the end of skeletal growth. *See* ARCHOSAURIA; DIAPSIDA; REPTILIA.

Characteristics. Living lepidosaurs can be divided into two subequal groups, Rhynchocephalia and Squamata. Abundant and globally diverse during the early Mesozoic, Rhynchocephalia are represented today only by the endangered tuatara (*Sphenodon*) of New Zealand. *Sphenodon* shows some primitive traits in comparison with squamates (such as rudimentary hemipenes, retention of belly ribs, and little or no cranial kinesis), but it is not a "living fossil" as once thought. The skull and dentition are specialized and the eardrum has been lost secondarily. Squamata form the largest and most successful group of living reptiles, with more than 6000 species and a global distribution. Many squamates are small insect eaters, but the group encompasses herbivores and major predators. They range in size from a few millimeters (such as some tiny geckos and chameleons) to several meters (Komodo dragons and anacondas), and in locomotor type they vary from quadrupedal to bipedal, from limbed to limbless, and from terrestrial runners to swimmers, specialized climbers, and gliders. Reproductive strategies range from egg laying to live birth, and at least eight groups have developed one or more all-female (parthenogenetic) species. *See* RHYNCHOCEPHALIA; SQUAMATA.

Fossil record. Most lepidosaurs are small with lightly built skeletons. Such skeletons are preserved and fossilized only under suitable conditions (fine sediments, quiet depositional environment), and their subsequent recovery depends on the use of an appropriate search strategy. Consequently, most lepidosaurian fossils have been recovered either in the search for early mammals (microvertebrate assemblages from caves, fissures, pond deposits) or early birds (fine-grained "lithographic" limestones), and there are major gaps in the record. The ancestors of squamates and rhynchocephalians probably diverged in the Early Triassic, but the first confirmed records are of rhynchocephalians from the Upper Triassic of Britain and Germany. These are already too derived to be an ancestral type, suggesting rhynchocephalians had a significant unrecorded Early to Middle Triassic history. The record of squamates is rather similar. The earliest known lizards are of Early to Middle Jurassic age (India, Central Asia, Britain),

but they include representatives of several major lineages. This implies an earlier, Triassic beginning to the squamate radiation. Susan E. Evans

Bibliography. M. J. Benton, *Vertebrate Palaeontology*, 3d ed., Blackwell, Oxford. 2004; M. J. Benton (ed.), *The Phylogeny and Classification of the Tetrapods*, vol.1: *Amphibians, Reptiles and Birds*, Clarendon Press, Oxford. 1988; T. R. Halliday and K. Adler (eds.), *The New Encyclopedia of Reptiles and Amphibians*, Oxford University Press, 2002; G. R. Zug, L. J. Vitt, and J. P. Caldwell, *Herpetology*, Academic Press, San Diego, 2001.

Lepospondyli

An assemblage of diverse, extinct, small tetrapods found in rocks from the Mississippian to Permian periods (340–245 million years ago). In older Linnean classifications, Lepospondyli is a subclass of Amphibia, along with Labyrinthodontia (other Paleozoic tetrapods) and Lissamphibia (the modern amphibians). Recent analyses of interrelationships of all Paleozoic tetrapods have changed this view. Labyrinthodontia, for instance, has been found not to include all descendant lineages and has been abandoned as a technical name. However, research demonstrates that Lepospondyli includes all descendants of a single common ancestor (a clade), and continues to be used in both Linnean and phylogenetic classifications. Anatomical features uniting lepospondyls include a lack of labyrinthine infolding of tooth dentine, paired palatal fangs and replacement pits, absence of an otic notch in the back of the skull, a small forwardly oriented projection on the first vertebra (reduced in some), and single spool-shaped vertebrae (see **illustration**). However, these features are also characteristic of juvenile or small amphibians. Since all lepospondyls are small, a few specialists still question the group's integrity. *See* AMPHIBIA; LABYRINTHODONTIA; LISSAMPHIBIA.

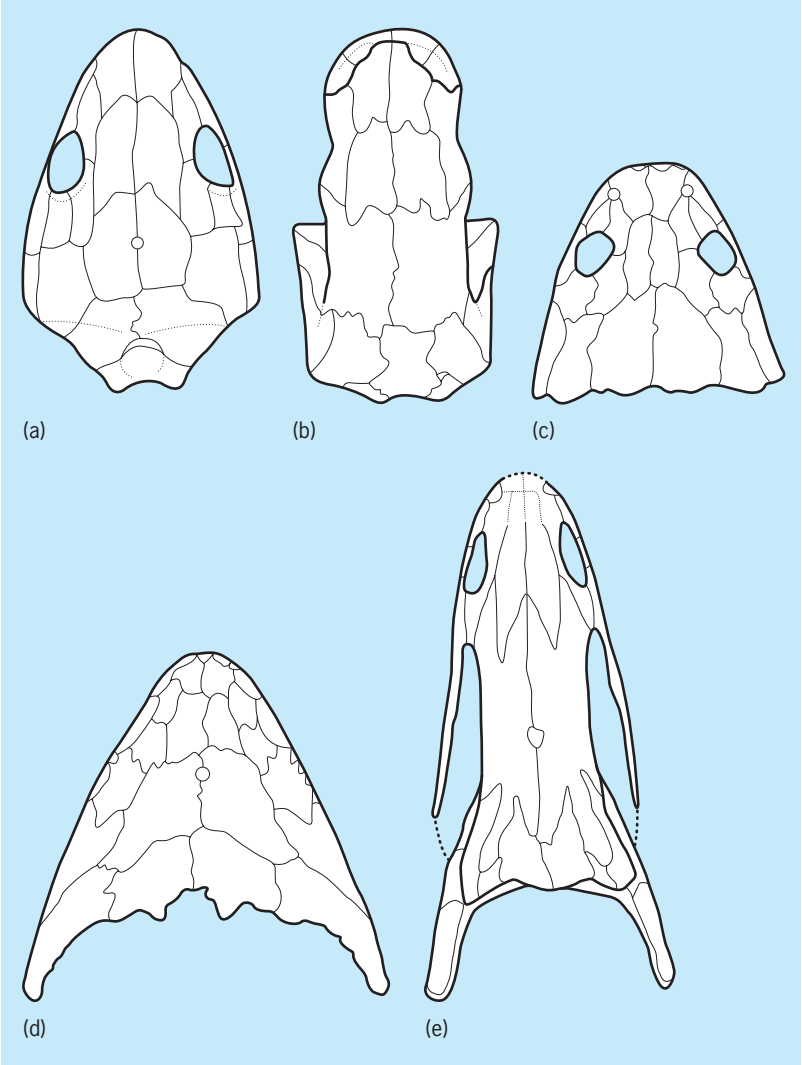

Skulls of lepospondyls. (*a*) *Cardiocephalus*, a microsaur. (*b*) *Brachydectes*, a lysorophian. (*c*) *Scincosaurus*, a nectidean. (*d*) *Batrachiderpeton*, a nectridean. (*e*) *Coloraderpeton*, an aistopod. (*From J. S. Anderson, 2007*)

Taxonomy and characteristics. In Linnean taxonomy, Lepospondyli comprises five diverse orders: Microsauria, Nectridea, Lysorophia, Aistopoda, and Adelospondyli. All are also considered to be clades with the exception of Microsauria, which might include members more closely related to other lepospondyls than other microsaurs. Microsaurs are small, salamander- and lizardlike animals with diverse groupings (11 families) of variously elongate or consolidated body plans and a distinctive cranial articulation. Nectrideans are mostly aquatic, newtlike animals with unique tail vertebrae which support a broad fin for swimming. Adelospondyls, lysorophians, and aistopods are progressively more elongate animals with reduced limbs; aistopods lack limbs entirely, retain only a trace of the shoulder girdle, and have over 200 vertebrae. All are found in North America or Europe (Euramerica), with the exception of one late-occurring nectridean found in Morocco. *See* AISTOPODA.

Paleontology. Lepospondyls are receiving increased attention from paleontologists in recent years for two reasons. First, multiple studies indicate that lepospondyls are the last Paleozoic tetrapod group to share a common ancestor with the amniote lineage, so they are important for understanding the early evolution of amniotes. Second, some suggest that Lepospondyls gave rise to some or all of the modern amphibians. One hypothesis is that lissamphibians share a unique common ancestor closely related to lysorophians instead of another group of Paleozoic tetrapods, the temnospondyls. Another hypothesis is that caecilians are descended from microsaurs, but salamanders and frogs are descended from temnospondyls. A third hypothesis places salamanders and caecilians within microsaurs but leaves the origin of frogs within temnospondyls. These lepospondyl hypotheses for modern amphibian origins are minority views but are driving much new research into the anatomy of amphibians, including, for the first time, comparisons of development patterns between fossil and modern amphibians. *See* TEMNOSPONDYLI. Jason S. Anderson

Bibliography. J. S. Anderson, Incorporating ontogeny into the matrix: A phylogenetic evaluation of developmental evidence for the origin of modern amphibians, in J. S. Anderson and H.-D. Sues (ed.), *Major Transitions in Vertebrate Evolution*, Indiana University Press, Bloomington, in press, 2006; M. J. Benton, *Vertebrate Palaeontology*, 3d ed., Blackwell Press, London and New York, 2004; R. L. Carroll, *Vertebrate Paleontology and Evolution*, W. H. Freeman, New York, 1988.

Leprosy

A chronic infectious disease caused by *Mycobacterium leprae* that primarily affects the skin and peripheral nerves and, to a lesser extent, involves the eyes and mucous membranes. Leprosy, or Hansen's disease, has been known for more than 2000 years. It afflicts at least 3 million people worldwide and is most common in developing countries. There are about 200 new cases yearly in the United States, 85% of them among aliens. The pathogen, *M. leprae*, has never been cultured in artificial media, but it can be grown in the mouse (in the footpad) and the nine-banded armadillo, among other animals. About 2–3% of armadillos in the wild in the southern United States harbor the infection. Its epidemiology is not fully understood, but transmission probably takes place by the respiratory route. The bacillus is very slow-growing, and the incubation period is usually 3–5 years. Less than 5% of any population is susceptible, and these individuals have a deficient cell-mediated immune response specifically to *M. leprae*, which may be genetic in origin. Epidemics have occurred, but are rare. Historically, considerable stigma has been attached to the disease and remains a real, though diminishing, problem. *See* MYCOBACTERIAL DISEASES.

A skin rash and loss of feeling due to nerve damage by *M. leprae* are the hallmarks of leprosy. Usually the nerve damage is mild, but when severe the inability to feel, particularly in the hands and feet, predisposes the individuals to frequent injuries. Nerve involvement may also lead to a loss of muscle function that produces clawing of the fingers and toes as well as other neuromuscular dysfunctions. Manifestations of the disease depend upon the degree of the immune defect. Initially, most patients develop one or several depigmented areas of skin that may have decreased sensation, a stage referred to as indeterminate disease. The condition may self-heal, but if it is diagnosed it is always treated. If treatment or self-healing does not halt the disease, it may progress to one of three advanced types. The mildest of these, tuberculoid disease, is usually manifested as a single large depigmented, scaly, numb area. The most severe type, lepromatous leprosy, usually involves most of the skin to varying degrees, with variously sized nodules or other changes, but some cases have no distinct rash. Between the two extremes immunologically is borderline disease, with skin changes of both types (see **illus.**). The World Health Organization's simplified classification labels indeterminate and tuberculoid patients paucibacillary (few bacilli) and borderline and lepromatous patients multibacillary. Diagnosis is confirmed by a finding of bacteria in material scraped from tiny skin slits or a skin biopsy that shows the presence of bacilli and nerve involvement.

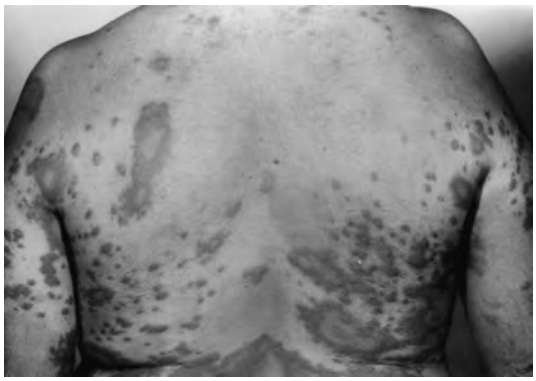

Borderline leprosy showing skin changes ranging from large tuberculoid lesions to lepromatous nodules.

Paucibacillary disease is treated with dapsone plus rifampin for 6–12 months. Clofazimine is added for multibacillary individuals and the treatment continued for 2 or more years. Treatment may be complicated by reactive episodes marked by fever, skin changes, and nerve inflammation. These side effects can usually be managed with drugs such as corticosteroids and thalidomide. Isolation is unnecessary, since patients become noninfectious within days of starting treatment. All can, in effect, be cured. *See* INFECTIOUS DISEASE.

Robert R. Jacobson

Bibliography. R. B. Conn (ed.), *Current Diagnosis 9th ed.*, 1999; H. D. Humes (ed.), *Textbook of Internal Medicine*, 4th ed., 2000; R. E. Rakel (ed.), *Conn's Current Therapy*, 1999.

Leptin

A protein comprising 167 amino acids that is important in the regulation of body weight, metabolism, and reproduction. Leptin was discovered in 1994, when it was identified as the missing protein in mice with a spontaneous single-gene defect that caused obesity. These *ob/ob* mice were very obese, had many of the metabolic abnormalities associated with obesity, and were infertile. However, when leptin levels were examined in other animal models of obesity and in obese humans, the levels were found to be elevated and not low or absent (as in the *ob/ob* mice). Thus, general obesity (not due to leptin deficiency) has been termed a leptin-resistant state. Since these initial findings, the biology of leptin has proven to be more complex than originally thought.

Secretion. Leptin is secreted primarily from adipocytes (fat cells), although other cells in other locations in the body, including the placenta, stomach, and skeletal muscle, also make leptin. Leptin is

secreted in a pulsatile fashion, and the average levels vary over the day (with the lowest levels in midmorning and the highest levels at night). The amount of leptin secreted by adipocytes is in proportion to the size of the cells. In animals, leptin levels increase soon after a meal and decrease with fasting. In humans, leptin levels do not change significantly with normal meals or with reasonable levels of exercise. However, with massive overfeeding, leptin levels will increase by up to 40%, and with fasts longer than 24 hours, leptin levels will decrease as much as 50%. Thus, in humans, leptin appears to play a role in the more chronic regulation of body fat rather than the meal-to-meal changes in energy intake and energy expenditure.

Insulin requirement. Insulin is required for leptin secretion. In animals and in cell culture, insulin increases leptin expression. Likewise, leptin levels decrease when animals are made insulin-deficient. There is also a correlation between insulin levels and leptin levels in humans, but it is difficult to separate the confounding effect of increasing adipose tissue. When humans are deficient in insulin, leptin levels are low, but body fat is also decreased. However, when humans are given excess insulin for several hours (while keeping the blood glucose normal), most studies find no changes in leptin levels. Other hormones that increase leptin secretion include glucocorticoids (such as prednisone) and cytokines (such as tumor necrosis factor alpha and interleukin-1). Hormones that decrease leptin secretion include catecholamines (such as isoproterenol and epinephrine), testosterone, and perhaps thyroid hormone.

Gender difference. The leptin level in women is almost three times higher than the leptin level in men. One reason is that women have a higher percent of body fat than men. Additionally, women have more subcutaneous fat (they tend to be "pear-shaped") than men (they tend to have more abdominal fat and are "apple-shaped"). Subcutaneous fat produces more leptin than abdominal fat, thus contributing to this gender difference. Finally, testosterone, which is present in much greater concentrations in men, decreases leptin and further augments the gender difference.

Receptors. The primary role of leptin appears to be in decreasing food intake by way of its action in the brain. Leptin's actions are exerted by binding to specific receptors. There are at least six different splice variants (isoforms) of the receptor for leptin. All of these receptors contain a leptin-binding area, and all but one receptor span the cell membrane. However, only one receptor (the long-receptor isoform, or Ob-Rb) has complete signaling function and is found predominantly in the hypothalamus (in areas known to control feeding behavior and hunger). The short-receptor isoforms are located mainly in peripheral tissues (such as the lung, kidney, liver, and gonads) and the choroid plexus, but the current role of these short forms of the leptin receptor is still under investigation. They likely play a role in leptin transport in the blood and across the blood-brain barrier. At least part of the resistance to leptin associated with obesity might be due to decreased transport across the blood-brain barrier.

JAK-STAT signaling pathway. When leptin binds to the long form of the receptor, the receptor dimerizes (two receptors come together in combination with leptin) and several cascades are initiated by way of activation of the JAK (janus kinase)–STAT (signal transducers and activators of transcription) pathway (see **illus.**). This activation by JAK results in phosphorylation of tyrosine amino acids on the leptin receptor, which causes changes in gene expression (via STAT pathways that ultimately decrease appetite and increase energy expenditure), downstream phosphorylation and dephosphorylation of enzymes, alterations in ion channels, and other signal modulation (in part via SOCS-3, a suppressor of cytokine signaling). SOCS-3 can actually feed back to the leptin receptor and inhibit leptin signaling, and this may also play a role in the leptin resistance seen in obesity.

It has recently been shown that the effects of leptin on appetite occur through the STAT pathway, but the effects on maintaining reproductive hormones require only the JAK and not the STAT pathway. Additionally, activation of the leptin receptor can also activate pathways analogous to those of the insulin receptor, including the PI-3 kinase pathway (phosphatidyl inositol kinase pathway, which stimulates glucose uptake and other metabolic effects of insulin) and the MAP kinase pathway (mitogen-activated kinase pathway, which stimulates cell growth).

Regulation of body weight. Leptin is believed to regulate body weight through its effects on appetite and energy expenditure. When leptin is given to animals, there is a marked decrease in appetite and an increase in energy expenditure.

Neural pathways of appetite regulation. The decrease in appetite associated with leptin administration is due to the stimulation of neurons that inhibit feeding and the inhibition of neurons that promote feeding. The long form of the leptin receptor is found on two distinct types of neurons, the POMC/CART (proopiomelanocortin and cocaine- and amphetamine-regulated transcript) neurons and the NPY/AGRP (neuropeptide-Y and agouti-related peptide) neurons. The POMC/CART neurons decrease food intake when stimulated by neurotransmitters such as leptin and α-MSH (melanocyte-stimulating hormone, a product of POMC that is made in the hypothalamus and binds to the melanocortin receptors on these neurons). Leptin also binds directly to the NPY/AGRP neurons and decreases food intake by inhibiting the neurons' ability to be stimulated. However, NPY and AGRP stimulate feeding. NPY stimulates food intake by binding directly to receptors on the NPY/AGRP neurons, whereas AGRP stimulates food intake indirectly by antagonizing or counteracting the effect of α-MSH on the melanocortin receptors.

Effects on energy expenditure. The increase in energy expenditure is due to stimulation of the sympathetic

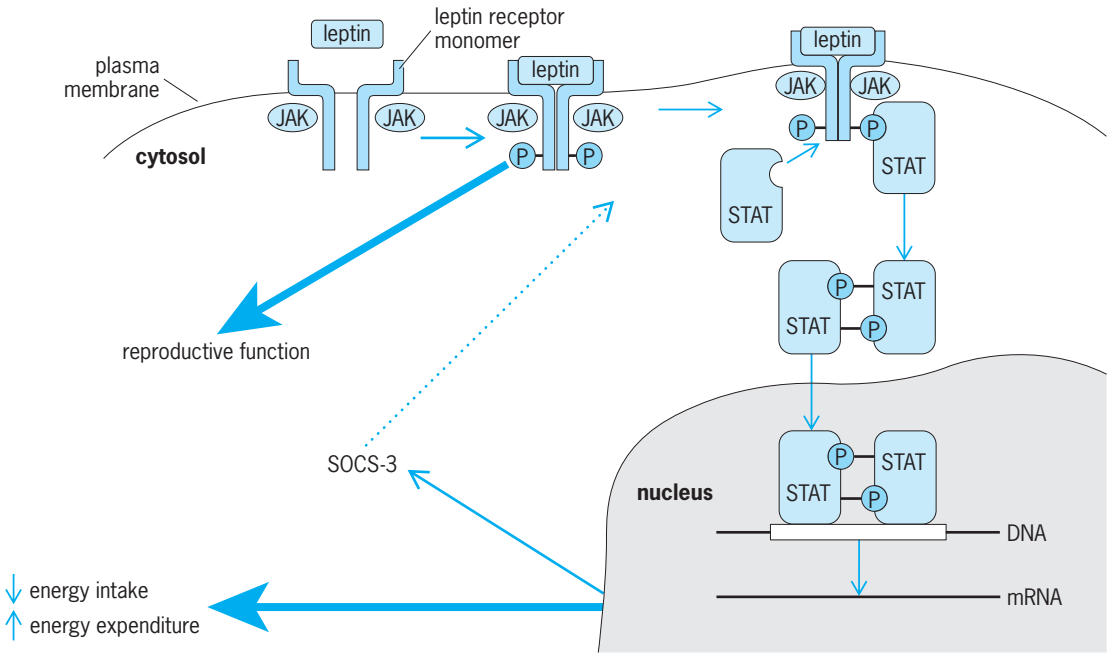

Leptin signal cascade. Leptin binds to the long form of the receptor, resulting in receptor dimerization and initiating several signaling cascades. The primary cascade signals by way of activation of the JAK-STAT pathway. The activation by JAK results in phosphorylation of tyrosine amino acids (shown as P in circle) on the leptin receptor. This in turn results in changes in gene expression via STAT pathways, which ultimately decrease appetite and increase energy expenditure. The maintenance of reproductive function requires only the JAK and not the STAT pathway. Signal modulation occurs in part via SOCS-3, a suppressor of cytokine signaling. SOCS-3 can feed back to the leptin receptor and inhibit leptin signaling. (*Adapted from D. L. Nelson and M. M. Cox, Lehninger Principles of Biochemistry, 3d ed., Worth Publishers, New York, 2000*)

nervous system, which causes increases in oxygen consumption and metabolic rate. In addition, leptin has been shown to induce adipose cell death (apoptosis). As animals become more obese, the response to leptin given peripherally (into the abdomen) decreases, and they become leptin-resistant. Although the ability to respond to central leptin administration (leptin given directly into the brain) remains intact, with continued obesity there is also a decrease in the response to central leptin administration. This adipose cell death, along with an alteration of how the body handles nutrients following leptin administration, might be why there is a preservation of muscle with weight loss (in contrast to weight loss from caloric restriction alone, which is accompanied by a loss in muscle).

Regulation of reproductive hormones. When leptin-deficient mice were first discovered, they were found to be infertile as well as obese. When leptin was given back to these animals, fertility was restored. In addition, when leptin is given to animals prior to puberty, they undergo puberty earlier than control animals. Thus leptin appears to be part of the system that tells the brain there are adequate energy stores (adipose tissue) for reproduction and allows puberty to commence. Leptin has been shown to increase the secretion of the hormones that regulate reproductive function, including gonadotropin releasing hormone, luteinizing hormone, and follicle stimulating hormone.

Other potential roles. Although most of the attention of leptin has centered on its role in decreasing food intake and decreasing weight, it is not certain that this is its primary role. Until very recently, there has been little need for a hormone that inhibits an organism from gaining as much weight as possible, for more efficient weight gain appears to protect against the next famine. Therefore, leptin likely has other roles as a hormone.

Fat accumulation. One role that has received much attention lately is that of a hormone that prevents fat accumulation in nonadipose tissues, including muscle, the liver, and the pancreas. Excess fat accumulation has been shown to be detrimental to these organs and may play a significant role in the development of diabetes. Leptin may traffic fat in such a way that it does not accumulate in these tissues. In addition, leptin increases utilization of fat by producing heat and not energy. Thus leptin might enhance more appropriate and rapid disposal of fat that enters these tissues.

Fasting. Another potential role of leptin is in mediating the endocrine response to fasting. When a human or mouse is fasted, there is a decrease in levels of leptin, reproductive hormones, thyroid hormone, and insulin. However, there is an increase in levels of glucocorticoids, catecholamines, and glucagon. This results in a conservation of energy, an increase in the production of glucose for the brain, and an increase in the release of fatty acids from the adipose tissue for muscle. When mice were fasted but given leptin, there was a blunting in several of these responses, including the changes in levels of glucocorticoids, reproductive hormones, and thyroid hormone. Thus

rather than acting as a signal of excess adipose tissue, leptin's primary role might be in signaling a decrease in energy, and initiating the stress response to this decrease.

Blood cell and bone formation. There is currently much debate on the extent of direct leptin action outside the brain. The gastrointestinal tract does have the long form of the leptin receptor, and signals from this receptor might play a role in nutrient absorption or appetite regulation. In addition, there is evidence that leptin has a direct role in the formation of bone red blood cells, white blood cells, and blood vessels.

Treatment of obesity. There are only five known people (from two families) who are deficient in leptin (like the *ob/ob* mice). They are very obese and seem to have reproductive abnormalities, but they do not appear to have other abnormalities such as increased glucose or insulin. Treatment with leptin has had a dramatic impact on their weight and has allowed puberty to occur. There are also three people (from one family) who do not make normal leptin receptors. These individuals have very high leptin levels, are very obese, do not have normal reproductive function, and have abnormal levels of thyroid hormone and growth hormone. Similar to the patients with leptin deficiency (and unlike the mouse models), they do not have high levels of blood glucose.

Leptin is also under investigation as a treatment for obesity in people with elevated leptin levels. Initial reports suggest that leptin administration results in weight loss in these subjects (with preservation of muscle), with the degree of weight loss similar to other pharmacologic treatments for obesity (about 8%). However, since leptin is a protein, it cannot be taken by mouth and must be injected, which sometimes results in pain, redness, inflammation, and bruising at the injection site. This mode of administration and the injection site reactions limit its use as a treatment. Whether administration of leptin or of a leptin-like compound will ever become a treatment for obesity remains to be seen. However, the discovery of leptin has opened a new area in the understanding of obesity and body weight regulation. *See* DIABETES; ENDOCRINE SYSTEM (VERTEBRATE); HUNGER; INSULIN; LIPID METABOLISM; METABOLIC DISORDERS; OBESITY.

William T. Donahoo; Robert H. Eckel

Bibliography. R. S. Ahima and J. S. Flier, Leptin, *Annu. Rev. Physiol.*, 62:413–437, 2000; C. A. Baile, M. A. Della-Fera, and R. J. Martin, Regulation of metabolism and body fat mass by leptin, *Annu. Rev. Nutrit.*, 20:105–127, 2000; R. B. Harris, Leptin—much more than a satiety signal, *Annu. Rev. Nutr.*, 20:45–75, 2000; B. Jeanrenaud and F. Rohner-Jeanrenaud, Effects of neuropeptides and leptin on nutrient partitioning: Dysregulations in obesity, *Annu. Rev. Med.*, 52:339–351, 2001; D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 3d ed., Worth Publishers, New York, 2000; R. H. Unger, Leptin physiology: A second look, *Regulatory Peptides*, 92(1–3):87–95, 2000.

Lepton

An elementary particle having no internal constituents. There are three known charged leptons: the electron (e^-), the muon (μ^-), and the tau (τ^-). In addition, the corresponding antiparticles e^+, μ^+, τ^+ are known. Each experiences electromagnetic, weak, and gravitational forces, but not the strong (nuclear) force. Associated with each charged lepton is a corresponding neutral lepton, called a neutrino (ν_e, ν_μ, ν_τ) or antineutrino: ($\bar{\nu}_e$, $\bar{\nu}_\mu$, $\bar{\nu}_\tau$); these have only weak and gravitational interactions. A charged lepton and its associated neutrino form a lepton generation; thus there are three known lepton generations (see **table**). *See* ELECTRON; NEUTRINO.

Leptons are very small, less than 10^{-18} m in radius. This is less than 10^{-3} of the radius of a nucleus and less than 10^{-8} of the radius of an atom. Indeed, all existing measurements are consistent with the assumption that leptons are point particles.

The lepton family of particles has very different properties than the quark family of particles. Quarks interact through the strong force as well as through the electromagnetic, weak, and gravitational forces. By means of the strong force, quark-antiquark pairs bind together to form mesons such as the π meson, and three quarks bind together to form baryons such as the proton. Whereas, as far as is known, quarks are always confined inside hadrons (mesons and baryons) and cannot be studied as isolated particles, leptons act as individual particles and can be studied as isolated particles. *See* BARYON; FUNDAMENTAL INTERACTIONS; HADRON; MESON; QUARKS.

More charged leptons or neutrinos may exist. Searches have been conducted up to 45 GeV in mass for additional charged leptons or neutrinos, but none have been found. Existing accelerators are not powerful enough to allow searches above 45 GeV. The reason there are only three lepton generations below 45 GeV in mass (and that there are only three generations of quarks) is unknown.

Decay. The electron is a stable particle: it never decays. The muon and tau are unstable particles, and their average lifetimes have been measured (see table). There is only one known mode of muon decay [reaction (1)].

$$\mu^- \rightarrow \nu_\mu + e^- + \bar{\nu}_e \tag{1}$$

The tau is heavier than the muon and thus decays more rapidly and in many modes, as in reactions (2).

$$\begin{aligned} \tau^- &\rightarrow \nu_\tau + \mu^- + \bar{\nu}_\mu \\ \tau^- &\rightarrow \nu_\tau + e^- + \bar{\nu}_e \\ \tau^- &\rightarrow \nu_\tau + \pi^- \end{aligned} \tag{2}$$

Lepton conservation and lepton mass. The association between charged leptons and their neutrinos has been described by an empirical law called lepton conservation. For example, consider the electron (e^-). There are only two ways in which an e^- can be destroyed: it can be annihilated by combining it with an e^+ (positron); or it can be changed into a ν_e. But an e^- cannot be changed directly into any other

Details of the three lepton generations

	Generation 1	Generation 2	Generation 3
Charged lepton name	Electron	Muon	Tau
Charged lepton symbol	$e^{\pm}$	$\mu^{\pm}$	$\tau^{\pm}$
Dates of discovery	1890s	Late 1930s	1974–1975
Charged lepton mass, MeV	0.51	105.7	~1777
Charged lepton lifetime, s	Stable	2.2×10^{-6}	3.0×10^{-13}
Associated neutrino symbols	ν_e, $\bar{\nu}_e$	ν_μ, $\bar{\nu}_\mu$	ν_τ, $\bar{\nu}_\tau$

lepton or hadron, such as an e^+, μ^-, μ^+, τ^-, τ^+, or π. Furthermore an e^- cannot be changed into a muon-associated neutrino (ν_μ, $\bar{\nu}_\mu$) or into a tau-associated neutrino (ν_τ, $\bar{\nu}_\tau$). Thus it seems that there is a unique property of the electron (e^-) which is preserved or conserved in all reactions; the only other particle carrying this unique property is the electron-associated neutrino (ν_e).

Similar lepton conservation laws appear to hold for the μ and τ. The only other particle carrying the unique property of the μ^- is the muon-associated neutrino (ν_μ). The only other particle carrying the unique property of the τ^- is the tau-associated neutrino (ν_1).

Until 1998 all experimental evidence was consistent with the assumption that lepton conservation is universally valid. However, it is now known that the law breaks down in the phenomenon of neutrino oscillations, which occurs because (1) the neutrinos ν_e, ν_μ, ν_τ do not have definite masses; (2) a quantum-mechanical mixing of neutrino states can occur. It is now known from the results of neutrino oscillation experiments that the quantum-mechanical state of, for example, ν_μ must be expressed as a specific linear combination of three neutrino mass states, at least two of which are associated with definite and distinct nonzero mass (and the same general statement is true of ν_e and ν_τ). Consider a ν_μ that is created at time $t = 0$ in muon decay [reaction (1)]. As this neutrino propagates through space, the states of definite and distinct mass of which it is composed evolve at different rates and thus get out of phase. The result is that after a finite time has elapsed, the neutrino is no longer a pure ν_μ but is in a superposition of ν_e, ν_μ, ν_τ states. Hence there is a finite probability that in a collision with another particle this neutrino could be converted to an electron or (if enough energy is available) even to a tau lepton. This violates lepton conservation.

The masses of the charged leptons (see table) are given in the energy equivalent unit megaelectron-volts (MeV). One MeV is the energy gained by an electron that is accelerated through a voltage of 10^6 V, and is equal to 1.602×10^{-13} joule. The masses of the electron, muon, and neutrinos are smaller than the masses of any of the hadrons. However, the discovery of the tau, whose mass is larger than that of many hadrons, destroyed the concept that leptons had to be very small mass particles.

The mass values m_1, m_2, m_3 of the neutrino states of definite mass are not yet known. At present, results of neutrino oscillation experiments reveal only that Eqs. (3) apply, where c is the velocity of light. These

$$\begin{aligned} m_2^2 - m_1^2 &= 8.1 \cdot 10^{-5} (\mathrm{eV}/c^2) \\ m_3^2 - m_1^2 &= 2.2 \cdot 10^{-5} (\mathrm{eV}/c^2) \end{aligned} \tag{3}$$

results imply that at least two of the three states are associated with nonzero (but certainly very small) mass.

Production. Leptons can be produced in various ways. One way is through the decay of a hadron. For example, a muon and its associated antineutrino are produced when a pion decays, as in reaction (4).

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu \tag{4}$$

Another way to produce leptons is through the annihilation of an electron with a positron; for example, muons can be produced by reaction (5), and taus can be produced by reaction (6).

$$e^+ + e^- \rightarrow \mu^+ + \mu^- \tag{5}$$

$$e^+ + e^- \rightarrow \tau^+ + \tau^- \tag{6}$$

It was through reaction (6) that the tau was discovered. *See* ELEMENTARY PARTICLE.

Eugene D. Commins; Martin L. Perl

Bibliography. E. Commins and P. Bucksbaum, *Weak Interactions of Leptons and Quarks*, Cambridge University Press, 1983; F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, Wiley, New York, 1984; B. Okun, *Particle Physics: The Quest for the Substance of Substance*, Harwood, London, 1985.

Leptospirosis

A bacterial disease caused by a *Leptospira* species that is now recognized as one of the new, emerging infectious diseases. The Centers for Disease Control and Prevention (CDC) has referred to "new, reemerging, or drug-resistant infections whose incidence in humans has increased within the past two decades or whose incidence threatens to increase in the near future." Leptospirosis, presumed to be the most widespread zoonosis (a disease of animals that may be transmitted to humans) in the world, is characterized by a broad spectrum of clinical manifestations, varying from a subclinical infection (showing no symptoms) to a fatal disease.

Etiologic agents. Leptospires are spirochetes (flexible, spiral-shaped bacteria with flagella) that are coiled, thin, and highly motile; have hooked ends;

and contain two flagella, which enable the bacteria to move in a watery environment and burrow into and through tissue. Traditionally, the genus *Leptospira* contains two species: the pathogenic *L. interrogans* and the free-living *L. biflexa*. Leptospires are phylogenetically related to other spirochetes that have a genome approximately 5000 kilobases in size. Recently, a number of leptospiral genes have been cloned and analyzed.

Epidemiology. Leptospirosis is primarily an animal disease with a worldwide distribution that affects at least 160 mammalian species. The incidence is significantly higher in warmer climates than in temperate regions, mainly due to longer survival of leptospires under warm humid conditions. Animals can be divided into maintenance hosts and accidental hosts.

Maintenance hosts are species in which infection is endemic and is usually transferred from animal to animal by direct contact. Infection is usually acquired at an early age, and the prevalence of chronic excretion of the leptospires in the urine increases with the age of the animal. The disease is maintained in nature by chronic infection of the renal tubules of maintenance hosts. Rats, mice, opossums, skunks, raccoons, and foxes are the most important environmental maintenance hosts, with farm animals and dogs being domestic maintenance hosts. All of these animals shed the bacteria in their urine.

Humans are not commonly infected with leptospires and are considered an accidental (incidental) host. Approximately 100 human cases of leptospirosis are reported to the CDC each year, but this probably misrepresents the total number of cases.

Transmission. Transmission of leptospires may follow direct contact with urine, blood, or tissue from an infected animal or exposure to a contaminated environment. Human-to-human transmission is rare. Since leptospires are excreted in the urine, they can survive in water for many months. Leptospires can enter the host through abrasions in the skin or through mucous membranes such as the conjunctiva. Drinking contaminated water may introduce the bacteria through the mouth, throat, or esophagus. The incubation period is usually 1–2 weeks.

Certain occupational groups are at high risk for acquiring leptospires. These include veterinarians, sewage workers, agricultural workers, livestock farmers, slaughterhouse workers, meat inspectors, rodent-control workers, and fish workers. They acquire the bacteria by direct exposure or contact with contaminated soil or water. In western countries, there is a significant risk of acquiring leptospirosis from recreational exposure (canoeing, windsurfing, swimming, water skiing, white-water rafting, freshwater fishing) and domestic animal contact.

Pathogenesis. The mechanisms by which leptospires cause disease are not well understood. It is known that after the leptospires enter the host, they spread to all organs of the body. The bacteria reproduce in the blood and tissues. Although any organ can be infected, the kidneys and liver are most commonly involved. In these organs, the leptospires attach to and destroy the capillary walls. In the liver this leads to jaundice, and in the kidneys it causes inflammation (nephritis) and tubular cell death (necrosis), resulting in renal failure. After antibodies are formed, the leptospires are eliminated from all organ sites except the eyes and kidneys, where they may persist for weeks or months, for unknown reasons.

Clinical manifestations. The clinical manifestations of leptospirosis can be divided into two forms: anicteric (not jaundiced) and icteric (jaundiced).

The majority of infections caused by the leptospires are anicteric (either subclinical or of very mild severity), and patients usually do not seek medical attention. In this mild form, leptospirosis may be similar to an influenza-like illness with headache, fever, chills, nausea, vomiting, and muscle pain. Fever is the most common finding during a physical examination. Most patients become asymptomatic within 1 week.

Between 5 and 15% of all patients with leptospirosis have the icteric (severe) form of the disease, which is characterized by jaundice, renal and liver dysfunction, hemorrhaging, and multiorgan involvement, and can result in death. This severe form is also referred to as Weil's syndrome (after Adolf Weil, a German physician). The mortality rate for is 5–15%. The jaundice occurring in leptospirosis is not associated with hepatocellular necrosis, and liver function returns to normal after recovery. Acute infections in pregnancy have also been reported to cause abortion and fetal death.

The clinical presentation of leptospirosis is biphasic, with the acute or septicemic (characterized by bacteria in the blood) phase lasting about a week. This is followed by the immune phase, characterized by antibody production and excretion of the leptospires in the urine. Most of the complications of leptospirosis are associated with localization of the leptospires within the tissues during the immune phase and thus occur during the second week of the infection.

Laboratory findings. Since the kidneys are invariably involved in leptospirosis, related findings include presence of white blood cells, red blood cells, and granular casts (proteinaceous products of the kidney in the urine); mild proteinuria (protein in the urine); and in severe leptospirosis, renal failure and azotemia (an increase in nitrogenous substances in the blood). The most common x-ray finding in the kidney is a patchy pattern that corresponds to the many hemorrhages due to the destruction of blood vessels.

Diagnosis. A definitive diagnosis of leptospirosis is based either on the isolation of the leptospires from the patient or on immunological tests. Leptospires can be isolated from the blood and/or cerebral spinal fluid during the first 10 days of the illness, and from the urine for several weeks beginning about the first week of the infection. They can be stained using carbol fuchsin, counterstained, and cultured in a simple medium enriched with vitamins.

Treatment. Treatment of leptospirosis differs depending on the severity and duration of symptoms at the time of presentation. Patients with only flu-like symptoms require only symptomatic treatment but should be cautioned to seek further medical help if they develop jaundice. The effect of antimicrobial therapy for the mild form of leptospirosis is controversial. However, such treatment is definitely indicated for the severe form of the disease, and should begin as soon as a severe diagnosis is made. In mild cases, oral treatment with tetracycline, doxycycline, ampicillin, or amoxicillin may be considered. For more severe cases of leptospirosis, intravenous administration of penicillin G, amoxicillin, ampicillin, or erythromycin is recommended. Most patients with leptospirosis recover. Mortality is highest in the elderly and those who have Weil's syndrome. Leptospirosis during pregnancy is associated with high fetal mortality. Long-term follow-up of patients with renal failure and hepatic dysfunction has documented good recovery of both renal and liver function.

Prevention. Individuals who may be exposed to leptospires through either their occupation or involvement in recreational water activities should be informed about the risks. Some measures for controlling leptospirosis include avoiding exposure to urine and tissues from infected animals, vaccination of domestic animals, and rodent control. *See* BACTERIA; MEDICAL BACTERIOLOGY; SPIROCHETE; ZOONOSES.

John P. Harley

Bibliography. R. W. Farrar, Leptospirosis, *Crit. Rev. Clin. Lab. Sci.*, 21(1), 1995; R. W. H. Gillespie, Epidemiology of leptospirosis, *Amer. J. Pub. Health*, vol. 53, 1963; P. N. Levett, Leptospirosis, *Clin. Microbiol. Rev.*, 14(2), 2001; R. van Crevel, Leptospirosis in travelers, *Clin. Infect. Dis.*, 19:132, 2001.

Leptostraca

The only extant order of the crustacean subclass Phyllocarida. The Leptostraca is represented by one fossil and a small number of living genera. These malacostracans (see **illus.**) are unique in having the carapace laterally compressed to such an extent that it forms a bivalvelike shell held together by a strong adductor muscle. The carapace covers only the thorax, leaving exposed the head, with its uniquely movable rostrum, stalked eyes, paired antennules, and antennae. The thorax is composed of eight malacostracan somites, each with a pair of generally similar appendages (thoracopods). Unlike the typical malacostracan abdomen, the leptostracan abdomen consists of seven somites plus a telson with furcal rami. Only the first six abdominal somites carry paired appendages (pleopods). The retention of a seventh somite and caudal furca, in addition to a fully developed telson in the Leptostraca suggested an evolutionarily early offshoot from the malacostracan stem. Although leptostracans are the only extant

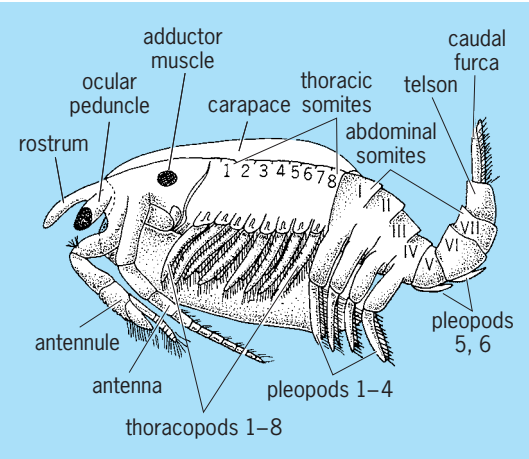

***Paranebalia longipes*. (*After P. A. McLaughlin, Comparative Morphology of Recent Crustacea, W. H. Freeman, 1980*)**

malacostracans to have seven clearly delineated abdominal somites, evidence from embryological and neurological studies indicates that in some members of the Mysidacea a seventh somite, or the ganglia thereof, remains in developmental stages. *See* MALACOSTRACA.

Leptostracans use the thoracopods to produce a feeding current and, in females, to form a brood pouch. That secondary brooding function suggests that egg-bearing females generally do not feed, and evidence confirms this. Locomotion in leptostracans is accomplished by use of the first four pairs of pleopods. Most leptostracans are bottom dwellers, living on or slightly under the substrate, but one species is holopelagic, one inhabits hydrothermal vents, and still another is a marine cave dweller. *See* CRUSTACEA; PHYLLOCARIDA.

Patsy A. McLaughlin

Bibliography. T. E. Bowman, J. Yager, and T. M. Iliffe, *Speonebalia cannoni*, n. gen., n. sp., from the Caicos Islands, the first hypogean leptostracan (Nebaliacea: Nebaliidae), *Proc. Biol. Soc. Wash.*, 98:439–446, 1985; R. R. Hessler, *Dahella caldariensis*, new genus, new species: A leptostracan (Crustacea, Malacostraca) from deep-sea hydrothermal vents, *J. Crust. Biol.*, 4:655–664, 1984; P. A. McLaughlin, *Comparative Morphology of Recent Crustacea*, 1980; F. R. Schram, *Crustacea*, 1986.

Lespedeza

A warm-season legume with trifoliate leaves, small purple pea-shaped blossoms, and one seed per pod. There are 15 American and more than 100 Asiatic species; two annual species, *Lespedeza striata* and *L. stipulacea*, and a perennial, *L. cuneata*, from Asia are grown as field crops in the United States. The American species are small shrubby perennials (see **illus.**) found in open woods and on idle land, rarely in dense stands; they are harmless weeds. Common lespedeza, once known as Japanese clover, is a small

Lespedeza. ***(Soil Conservation Service)***

variety of *L. striata*. This variety, unintentionally introduced in the 1840s and used widely until the late 1920s, has been replaced by a larger variety, Kobe. Korean lespedeza (*L. stipulacea*) has been widely grown since the mid-1920s and is preferred in the northern part of the lespedeza belt, from Missouri eastward, while Kobe is preferred in the lower part of the belt (south to the Gulf of Mexico and across northern Florida). *See* LEGUME.

The annual lespedezas can grow on poorer wetter land than most other forage crops. They give a moderate yield of good quality. The plants reseed well and the young seedlings at the two-leaf stage are very tolerant of cold. Most crop damage is from parasitic nematodes and severe drought, especially in the spring.

The perennial lespedeza (*L. cuneata*), commonly known as sericea, is adapted to the same growing conditions as the annuals but persists better on steep slopes and for this reason has been used widely in soil conservation. The young growth from well-established sericea is a moderate-quality forage, but the older growth is poor. The perennial is difficult to establish because the seed does not germinate readily unless scarified or exposed to cold and moisture for several months.

Several large species of lespedeza, such as the widely used *L. bicolor*, are used in plantings for wild life to provide seed and cover. *See* LEGUME FORAGES; NEMATA (NEMATODA).

Paul Tabor

Lethal dose 50

One special form of the effective dose 50, or median effective dose. The lethal dose 50, or median lethal dose, is also written as LD_{50}. It is used when the response that is being investigated is the death of the experimental animal. The median lethal dose is therefore the dose which is fatal to 50% of the test animals.

The LD_{50} for strychnine given subcutaneously to rabbits is 0.5 mg/kg of body weight. This description shows that in specifying the LD_{50} attention has been given to the species of animal, the route of injection, and the weight of the animals used. For other toxic substances a different set of specifications might be important.

The use of the concept of LD_{50} is not restricted to toxicology, because there are many therapeutic substances, such as digitalis, which can be assayed more their lethal effect than by their therapeutic effect. The same active principle is involved in both effects. *See* BIOASSAY; EFFECTIVE DOSE 50; TOXICOLOGY.

Colin White

Lethal gene

A gene which brings about the death of the organism carrying it. Lethal genes constitute the most common class of gene changes (that is, mutations) and are reflections of the fact that the fundamental function of genes is the control of processes essential to the growth and development of organisms. In higher diploid forms lethals are usually recessive and expressed only in homozygotes. Dominant lethals, expressed in heterozygotes, are rapidly eliminated and thus rarely detected. Recessive zygotic lethals are retained with considerable frequency in natural populations of cross-fertilizing organisms, while gametic lethals (those affecting normal functioning of eggs and sperm among animals, or the pollen and ovules of plants) are subject to stringent selection and are accordingly rare.

Evaluation and analysis. Some lethal mutations have proved to be losses of small or large sections of chromosomal material, rather than gene changes in the strictest sense. The presence of lethal mutants or losses is expressed as a failure of growth, leading to the premature death of the organism carrying them. As most detectable lethal genes are recessive, their effects are observed only in homozygotes. If the processes controlled by a particular gene come early in the developmental sequence, the disturbance usually has fewer ramifications. An analysis of the effects of lethal genes provides a valuable means of investigating the complex processes of embryonic development and cellular differentiation and, in the case of humans, has practical medical implications as well. In microorganisms the study of lethal biochemical mutants opened up a whole new era in the unraveling of biosynthetic processes and led to the establishment of modern biochemical genetics. *See* CHROMOSOME; MOLECULAR BIOLOGY; MUTATION.

Conditional mutants. Although typically lethal genes express themselves independently of internal and external environments, there are cases in which expression is dependent on some particular condition. The temperature-sensitive (*ts*) and amber (*am*) mutants in bacteriophage T4D provide instances where the mutant behaves as a lethal under one set of conditions (restrictive) while under another set of conditions the phenotype of the mutant resembles wild type. Such conditional mutants have been used to analyze gene function in the synthesis and assembly of the macromolecules making up bacteriophage particles. Over 40 conditional lethal mutants have been shown to affect the processes leading to the complete morphogenesis of T4D phage particles. Such conditional lethals are thus invaluable tools for full analysis of gene action in the organization and functioning of biological systems.

Obtaining test materials. One of the difficulties in establishing the mechanism of action of lethal genes in higher organisms is that only one-quarter of the progeny in crosses of heterozygous lethals are homozygous. Analysis at the biochemical level can be greatly simplified if progenies consisting wholly of homozygous lethal individuals are readily obtained. Conditional lethal mutants of the temperature-sensitive type, readily produced in the fly *Drosophila* with the aid of chemical mutagens, which are viable and fertile when raised at 18°C (64°F), and completely lethal at 28°C (82°F), have been the subject of active investigation in several laboratories. By maintaining flies at 18°C and transferring eggs to 28°C, any quantity of lethal tissue becomes available for biochemical study.

Occurrence in natural populations. Lethal genes in natural populations of plants and animals have received considerable attention from geneticists. In certain instances these may form the bases of polymorphisms in such populations. In some organisms they have led to what are known as balanced lethal systems: When two different lethal genes (or losses) are located at different points in different members of a pair of homologous chromosomes in such positions, or circumstances, that recombination by crossing over is prevented, only those individuals heterozygous for both of the two different lethals survive. In this way strains, or even species, which are permanent heterozygotes have arisen as in the evening primroses of the genus *Oenothera*. In human populations the accumulation of lethal mutations have been a matter of some concern in the face of increases in mutagenic agents in the environment and the lowering of selective pressures associated with modern life. The human wastage and misery which accompany high frequencies of lethal genes are considerable. *See* POLYMORPHISM (GENETICS).

Examples. Lethal genes include the following: biochemical lethals in microorganisms such as *Neurospora, Escherichia,* and *Salmonella*; lethal chlorophyll mutants of higher plants; notch and many others in the fruit fly *Drosophila*; creeper in fowl; yellow, brachyury, W-anemia in the mouse; gray-lethal in the rat; Dexter in cattle; thalassemia, sickle-cell anemia, and many others in humans. *See* HUMAN GENETICS.

Donald F. Poulson

Bibliography. E. Hadorn, *Developmental Genetics and Lethal Factors*, 1961; D. L. Hartl, *Genetics*, 4th ed., 1998; M. M. Moore et al. (eds.), *Mammalian Cell Mutagenesis*, 1988; M. W. Strickberger, *Genetics*, 3d ed., 1985.

Lettuce

A cool-season annual, *Lactuca sativa*, of Asian origin and belonging to the tribe Cichorium of the Compositae family. Lettuce is grown for its succulent leaves, which are eaten raw as a salad. Four varieties of this leading salad crop are head lettuce (*L. sativa* var. *capitata*; see **illus.**), leaf or curled lettuce (*L. sativa* var. *longifolia*), and stem or asparagus lettuce (*L. sativa* var. *asparagina*). There are two types of head lettuce: butterhead, and crisphead or iceberg. *See* ASTERALES.

Propagation. The outdoor crop is propagated by seed usually planted directly in the field, but also occasionally planted in greenhouses for later transplanting. Field spacing varies, but usually plants are grown 10–16 in. (25–40 cm) apart in 14–20-in. (35–50-cm) rows. Greenhouse lettuce, predominantly leaf and butterhead varieties, is transplanted to ground beds with plants placed 7–12 in. (18–30 cm) apart.

Uniformly cool weather promotes maximum yields of high-quality lettuce; 55–65°F (13–18°C) is optimum. Heading varieties (cultivars) are particularly sensitive to adverse environment. High temperatures prevent heading, promote seed-stalk development, and result in bitter flavor and tip-burned leaves. Thus, commercial production of lettuce is extensive in several California and Arizona valleys where mild

Head lettuce (*Lactuca sativa* var. *capitata*). (*Asgrow Seed Co., subsidiary of The Upjohn Co.*)

winter and cool summer climates prevail. Production in other areas with hot summer weather is restricted to the cooler spring and fall seasons.

Crisphead or iceberg lettuce is the most widely grown type; cultivars belonging to the Great Lakes variety account for most of the acreage. Popular cultivars of other types of head lettuce are Butterhead, Dark Green Boston, White Boston, and Bibb; leaf lettuce types are Grand Rapids, Prize Head, and Salad Bowl; cos or romaine types are Parris Island, Dark Green, and Valmaine.

Harvesting. The harvesting of heading varieties begins when the heads have become firm enough to satisfy market demands, usually 60–90 days after planting. Most western-grown lettuce is field-packed in paperboard cartons and chilled by vacuum cooling.

Leaf lettuce and cos lettuce are harvested when full-sized, but before the development of seed stalks or a bitter taste. This varies from 40 to 70 days after planting.

California raises more lettuce than any other state; Arizona, Florida, and Texas are next in importance.

Diseases. Lettuce diseases of greatest importance are caused by viruses, fungi, and bacteria which alter the structure and function of the infected parts, causing stunting, malformation, death, discoloration, and breakdown of tissues.

The most serious fungus diseases (and their causal organisms) are sclerotinia rot or drop (*Sclerotinia sclerotiorum* and *S. minor*), gray mold rot (*Botrytis cinerea*), downy mildew (*Bremia lactucae*), anthracnose (*Marssonina panattoniana*), and bottom rot (*Rhizoctonia solani*).

The most serious of the bacteria parasitic on lettuce are *Erwinia carotovora* and *Pseudomonas marginalis*, both of which cause soft rots. Soft rots are especially damaging when lettuce is not cooled quickly after harvest and is not maintained cool (36–40°F or 2–4°C) during storage and transit to market.

The potentially most serious disease of lettuce is common mosaic, caused by a virus that is seed-borne and transmitted from infected to healthy plants by aphids. Fortunately, this disease can be avoided by use of virus-free seed. Lettuce is susceptible to several other aphid-transmitted viruses that are not seed-borne. Diseases caused by these viruses are usually less prevalent than mosaic and of minor importance.

Important diseases due to other infectious agents are aster yellows, caused by a leafhopper-transmitted spiroplasma, and Big Vein, caused by an infectious entity that has not been seen with the electron microscope and thus has not been identified and characterized. The causal agent is soil-borne and transmitted by a chytrid fungus, *Olpidium brassicae*, that infects the roots of lettuce and many other plants without causing evident injury at the site of infection.

Tip burn, a physiological disease, is especially damaging to head lettuce due to death and browning of leaf tissues inside the head. It is induced by high ambient temperatures occurring near the time of harvest that disrupt normal organic acid and calcium metabolism. Cultivars now being used have been selected for increased tolerance to tip burn. *See* PLANT PATHOLOGY.

R. G. Grogan

Leucite rock

Igneous rocks rich in leucite but lacking or poor in alkali feldspar. Those types with essential alkali feldspar are classed as phonolites, feldspathoidal syenite, and feldspathoidal monzonite. The group includes an extremely wide assortment both chemically and mineralogically.

The rocks are generally dark-colored and aphanitic (not visibly crystalline) types of volcanic origin. They consist principally of pyroxene and leucite and may contain calcic plagioclase or olivine. Types with plagioclase in excess of 10% are called leucite basanite (if olivine is present) and leucite tephrite (if olivine is absent). Types with 10% or less plagioclase are called leucitite (if olivine is absent) and olivine leucitite or leucite basalt (if olivine is present).

The texture is usually porphyritic with large crystals (phenocrysts) of augite and leucite in a very fine-grained or partly glassy matrix. If plagioclase occurs as phenocrysts, it is generally labradorite or bytownite and is slightly more calcic than that of the rock matrix. It may be zoned with more calcic cores surrounded by more sodic margins. Leucite appears in two generations. As large phenocrysts it forms slightly rounded to octagonal grains with abundant tiny inclusions of glass or other minerals zonally arranged. Small, round grains of leucite with tiny glass inclusions also occur in the rock matrix. Augite or diopside (sometimes rimmed with aegirine-augite) and aegirine-augite form the mafic phenocrysts. Pyroxene of the matrix is commonly soda-rich. Olivine may occur as well-formed phenocrysts. Other minerals present may include nepheline, sodalite, biotite, hornblende, and melilite. Accessories include sphene, magnetite, ilmenite, apatite, and perovskite.

Leucite rocks are rare. They occur principally as lava flows and small intrusives (dikes and volcanic plugs). Well known are the leucite rocks of the Roman province, in Italy, and the east-central African province. In the Italian area the feldspathoidal lavas are essentially leucite basalts and may have developed by differentiation of basalt magma (rock melt). In the African province the leucitic rocks are associated with ultramafic (periodotite) rocks and may have been derived from periodotitic material. This may have been accomplished by the abstraction of early formed crystals from a peridotitic magma or by the mobilization of peridotitic rocks by emanations from depth. Assimilation of limestone by basaltic magma may help to decrease the silica content and promote the formation of leucite instead of potash feldspar. The crystallization of leucite, however, is in large part a function of temperature and water content of the magma. Conditions of formation, therefore, may strongly influence the formation of leucitic rocks. *See* IGNEOUS ROCKS; LAVA; PHONOLITE.

Carleton A. Chapman

Leukemia

A disease characterized by a progressive and abnormal accumulation of white blood cells, or leukocytes. Leukemic cells are malignant because they have three characteristics common to all cancers: (1) they exhibit uncontrolled growth that is frequently associated with an inability to mature normally; (2) they arise from a single precursor cell; and (3) they disregard anatomic boundaries and metastasize to organs or tissues where leukocytes are not normally found. The expanding clone of leukemic cells infiltrates organs and tissues, particularly the bloodstream and bone marrow, where they disrupt the production of normal cells. The resulting symptoms include fatigue, pallor, infections, bruising and bleeding, and discomfort caused by enlarged organs. In humans, the term leukemia encompasses more than 20 distinct malignancies. *See* BLOOD; HEMATOPOIESIS.

Diagnosis and classification. Normal leukocytes are grouped into two primary types or lineages, myeloid and lymphoid, and virtually any cell of either lineage can become leukemic. Leukemias are also divided into broad categories that are based on the cell involved (myeloid or lymphoid) and disease aggressiveness (either acute or chronic). Subclassifications are based on morphologic, cytochemical, immunologic, cytogenetic, and molecular criteria. Depending on the type of leukemia and reason for the classification (treatment options, research, or reaching a prognosis), only one or all of the criteria may need to be determined. Some characteristics, including morphologic and cytochemical, can be established within hours or days, whereas others, such as molecular and cytogenetic composition, may take weeks. The length of time involved is important because therapy for acute leukemia usually needs to be started promptly.

Morphologic criteria. A microscopic analysis of stained peripheral blood and bone marrow is the first step in diagnosis (**Fig. 1**). In the appropriate clinical setting, that alone may be all that is needed to classify some leukemias, such as chronic myelocytic leukemia (**Fig. 2**). Morphology can differentiate lymphoid from myeloid types by the presence of Auer rods, which are intracellular structures that appear only in the acute myeloid leukemias (**Fig. 3**).

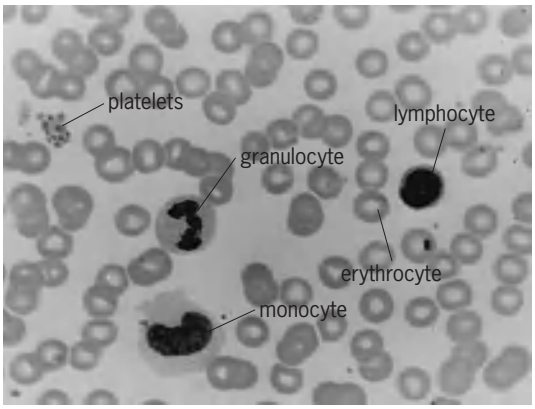

Fig. 1. Normal stained blood smear showing a granulocyte, lymphocyte, monocyte, some erythrocytes, and a clump of platelets.

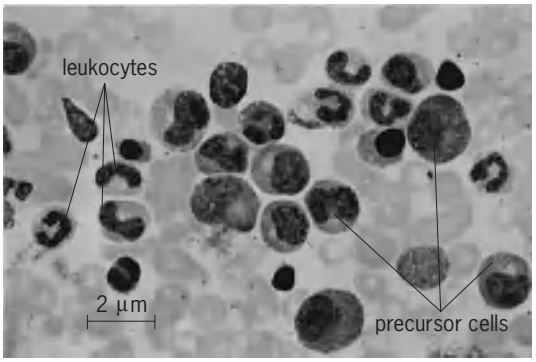

Fig. 2. Stained blood sample taken from a patient with chronic myelocytic leukemia. Many of the cells are bone marrow precursor cells, which are not ordinarily found in the circulating blood. In addition, the total leukocyte count is greatly increased over that found in normal blood.

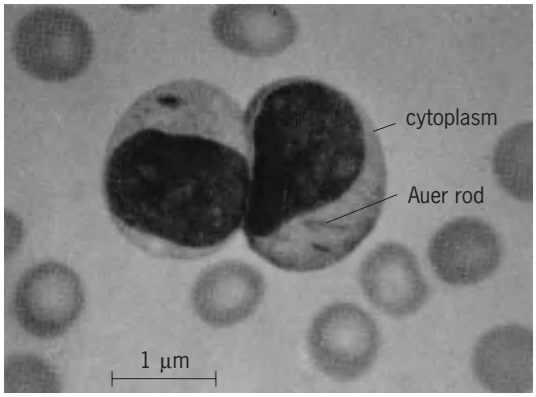

Fig. 3. Two abnormal cells in the peripheral blood of a patient with leukemia, both of which have crystalline structures in their cytoplasm. The crystals, known as Auer rods, are composed of lysosomal enzymes such as acid phosphatase and peroxidase and are virtually always found in patients with acute myeloblastic leukemia.

Cytochemical characteristics. Special stains that impart an easily recognizable color are routinely used to help distinguish cell lineage. One of the most useful stains detects the enzyme peroxidase, which appears only in myeloid cells. Acute myeloblastic leukemias can usually be subclassified by using only morphologic and cytochemical tests.

Immunologic reactivity. Monoclonal antibodies can bind to specific molecules on cells called antigens, and fluorescence or histochemical techniques can be used to reveal antibodies that bind to specific antigens found only on the surface of certain types of cells. Thus, it is possible to assign a leukemia cell to the lymphoid or myeloid categories when the morphologic and cytochemical data are ambiguous or conflicting. Immunodiagnosis has been particularly helpful in subcategorizing lymphoid leukemias and subtyping previously unclassifiable leukemias. Immunologic techniques have also been helpful in establishing clonality (the property of arising from a single cell) and in leading to increased understanding

of basic leukemia biology. *See* CELLULAR IMMUNOLOGY; MONOCLONAL ANTIBODIES.

Cytogenetic analysis. Leukemia diagnosis, prognosis, and etiology have been greatly improved through the analysis of individual chromosomes that have been differentially stained to reveal subchromosomal segments. In most leukemias, distinct chromosomal abnormalities appear nonrandomly with much greater frequency in almost all leukemic cells analyzed. The abnormalities are confined to the leukemic cells and are not found in other cells of the body, and so they are acquired defects. The best-known example is the Philadelphia (Ph) chromosome, a translocation of genetic material between chromosomes 9 and 22 that is detected in more than 90% of chronic myeloid leukemia cases. Cytogenetics also provides information on etiology and prognosis. Deletion of part of the long arm of chromosome 5 (5q-syndrome) or of one of the two chromosomes 7 (monosomy 7) strongly suggests that the leukemia was secondary to prior chemotherapy or arose from a preexisting bone marrow disorder, both of which suggest aggressive disease, poor response to therapy, and a shorter survival time. *See* CHROMOSOME; HUMAN GENETICS.

Molecular studies. Molecular biology identifies and analyzes small fragments of deoxyribonucleic acid (DNA) and has helped to confirm the importance of cytogenetics, the analysis of whole chromosomes. Molecular biology has allowed deeper insight into leukemogenesis and the diagnosis and classification of leukemia. For example, if an increased number of lymphocytes are present (lymphocytosis) but the evidence for lineage or malignancy is not clear, a finding of an abnormal amount of specific DNA that encodes for an antigen present only on the surface of T cells (thymus-derived lymphocytes) is a strong indicator of a T-cell leukemia. Molecular biology has been used following apparently curative therapy to screen blood or bone marrow cells for evidence of abnormal cells not detected by other, less sensitive methods such as morphologic or cytochemical analysis. *See* DEOXYRIBONUCLEIC ACID (DNA); MOLECULAR BIOLOGY.

Differentiating leukemia from a normal reaction to some other disease (secondary leukocytosis) is rarely a problem. However, a persistent yet nonmalignant leukocytosis may require all of the previously mentioned techniques to reach a definite diagnosis. *See* CLINICAL PATHOLOGY.

Epidemiology. In the United States an estimated 55,000 cases of leukemia are diagnosed annually and account for 4–5% of all cancers. These figures, and similar ones for other countries, have remained relatively constant since the mid-1960s. In industrialized countries, leukemia is second only to accidents as the most common cause of death among children under the age of 16, and it is the most common malignancy in these countries. Overall, it is the sixth most common cause of death in the United States. Geographic variations in the incidence of leukemias are less apparent than in any other form of cancer. Although worldwide incidence is variable, there is usually no more than a twofold difference between countries or regions within countries. The exception is chronic lymphocytic leukemia, which is the most common form of adult leukemia in western countries but is unusual in the Far East, Africa, and South America. In Japan, chronic lymphocytic leukemia is virtually nonexistent.

Leukemia can occur at any age and even be present at birth, but its frequency rises sharply after the age of 50. Acute leukemias, primarily the lymphoid type, are most common in children and young adults, whereas the chronic leukemias are rare in children. The majority of adult acute leukemias are myeloid. The incidence of chronic myelocytic leukemia peaks between the ages of 40 and 50. Chronic lymphocytic leukemia is rare under age 40; since 90% of cases are diagnosed in people over 50, it is the most common leukemia in the elderly. A slight male predominance is found in all of the leukemias, especially chronic lymphoid, hairy cell, and childhood acute lymphoid leukemias.

Research into environmental or infectious causes has led to the study of clusters or microepidemics of leukemia that are geographically or chronologically contiguous. There are no convincing arguments that leukemia clusters occur other than randomly, except for the adult T-cell form. Although not genetically transmissible in the classic sense, leukemia has a tendency to occur within families. Most striking is susceptibility in identical twins: the risk of the unaffected twin of a sibling with leukemia developing the disease has been estimated at 200 times that of the general population. Risk falls off rapidly for other family members and is only two to three times higher for nontwin siblings. Even a woman who has leukemia and is pregnant rarely if ever gives birth to a child with leukemia. Certain cell surface antigens are found more often in family members of leukemic patients than in the general population, but the search for a "leukemia susceptibility gene" possibly linked to an inherited human leukocyte antigen (HLA) type has been unsuccessful.

Etiology. An estimated 80% of all human cancers are induced by environmental carcinogens, but only a small fraction of exposed individuals develop cancer. Although many agents are suspected of inducing leukemia, for the great majority of cases the etiology is unknown. It appears that no single factor is causative, but a number of events must take place before leukemia occurs.

Radiation. The evidence for ionizing radiation as a leukemogenic cofactor is virtually irrefutable. A dose- and time-dependent relationship with development principally of myeloid leukemias has been shown following radiation exposure from nuclear explosions, with atomic bomb survivors and atomic test participants; from therapy for ankylosing spondylitis (inflammation of the vertebrae) or cervical cancer; and from work-related contact, such as radiology prior to effective shielding. Radiation appears to have no role in the genesis of chronic lymphoid leukemia, and normal exposure to diagnostic x-rays obtained during a lifetime of routine medical care, to natural sources of radiation, or to nonionizing radiation

does not appreciably increase the risk of leukemia. *See* RADIATION INJURY (BIOLOGY).

Chemicals and drugs. Chronic exposure to high levels of benzene and perhaps related compounds is associated with a tenfold higher risk of developing myeloid leukemia. The risk from exposure to other chemicals is much less clear, however. Agricultural occupations appear to predispose workers to an increased risk of lymphoma and leukemia, possibly as a result of contact with chemicals or animal-borne viruses. *See* LYMPHOMA.

A clear, strong association has been demonstrated between pharmaceuticals (particularly alkylating agents) that are administered as therapy for a primary cancer and the subsequent development of secondary leukemia, virtually always acute myeloblastic leukemia. Compared with chemotherapy that involves alkylators, radiotherapy alone is regarded as weakly leukemogenic, but the combination of both appears to be worse than either alone. Organ transplantation is usually followed by an increased risk of lymphoma, but leukemia risk also rises. The immune compromise induced by drugs that are used to block graft rejection and possibly the chronic reaction to foreign antigens on the graft as well appear to be related to the induction of malignancies.

Viruses. Only in the rare adult T-cell form of leukemia has a virus been implicated. A strong association has been shown between adult T-cell leukemia and a retrovirus [a type of ribonucleic acid (RNA) virus] called human T-cell leukemia virus I, or HTLV-I. The leukemia that it causes occurs endemically in Japan, especially the islands of Kyushu, Shikoku, and Hokkaido; the Caribbean; the southeastern United States; southern Italy; and parts of South America and Africa. Viruses are known to cause lymphomas and leukemias in nonhuman animals. Examples are bovine and feline leukemia viruses and avian leukosis virus, a retrovirus that is the source of a variety of hematologic neoplasias in domestic fowl. *See* AVIAN LEUKOSIS; FELINE LEUKEMIA; TUMOR VIRUSES; VIRUS.

Congenital chromosomal and acquired abnormalities. Persons with Down syndrome are 30 times more likely to develop acute, usually lymphoid leukemia than the rest of the population. Down syndrome patients have an extra chromosome 21, resulting in a total of 47 chromosomes instead of the normal 46, and the extra chromosome is thought to account for the increased risk for developing leukemia. In fact, leukemias in persons without Down syndrome who have abnormalities of chromosome 21 have been reported. Other rare congenital disorders that are characterized by chromosomal abnormalities and increased risk of leukemia include Fanconi's anemia, a potentially fatal blood disorder; ataxia-telangiectasia, a neuromuscular degenerative disease; and Bloom's syndrome, a condition marked by stunted growth and facial flushing. *See* DOWN SYNDROME.

In the primary immunodeficiency states, malignancies develop 10,000 times faster than in unaffected persons, and each of the immunodeficiencies is associated with a distinct leukemia: X-linked agammaglobulinemia (the condition characterized by lack of or extremely low levels of gamma globulin, together with defective antibody production and frequent infections) with lymphoid leukemia, common variable immunodeficiency with chronic lymphoid leukemia, and ataxia-telangiectasia with acute lymphoid leukemia.

Preexisting hematologic disorders. The myelodysplastic syndromes are characterized by ineffective production of normal blood cells, and result in low blood cell counts due to abnormal precursor cells in the bone marrow. The abnormal cells are clonal and manifest a spectrum of morphologic and cytogenetic abnormalities. Most have a tendency to evolve into acute leukemia, with the myeloid type predominating. Other clonal blood disorders that are characterized by increased production of one specific cell type or by fibrotic marrow have a slight tendency to evolve into acute leukemia. That tendency is greatly enhanced as a result of treatment with alkylating agents or radioactive compounds. A related disorder, chronic myelogenous leukemia, almost always transforms into an acute leukemia, a change referred to as a blast crisis. This process is heralded in 70% of cases by increased myeloblasts and in 25% by proliferation of lymphoblasts or other immature cells. The metamorphosis changes chronic myelogenous leukemia from a relatively well-tolerated chronic condition into a highly lethal leukemia that fails to respond to treatment.

Leukemogenesis. The transformation of a normal cell into a malignant one is complex and not fully understood, but it appears to be caused by a series of events that involve a particular cell in a susceptible host. In all likelihood, these events occur with an exact sequence, and timing and must be of sufficient strength and duration, such as exposure to radiation. The transforming events permanently alter or mutate the genome and confer upon the cell and its progeny a survival advantage over normal cells. Whether there is a block in maturation, as in acute leukemias, or an increase in stem cells (formative cells that have the ability to self-replicate and to give rise to specialized cells) with no maturation defect, such as is seen in chronic myelogenous leukemia, the end result is similar—expansion of the malignant clone.

The importance of cytogenetic changes was made clearer with the use of molecular biologic techniques that led to the discovery of oncogenes. Oncogenes are DNA sequences that were discovered in cancer-promoting viruses and that encode for the production of substances important to the normal cell's transfer of growth signals across the cell membrane. Theoretically, when expressed inappropriately, oncogenes malignantly transform cells. Whether these transformations are by themselves sufficient to cause leukemia is debatable. Cytogenetic and molecular discoveries help to further characterize leukemogenesis but do not yet provide a full explanation. *See* ONCOGENES.

Pathogenesis and clinical course. The two major types of leukemia usually differ in signs and symp-

toms. Acute leukemias have a relatively rapid onset, and those with the disease often experience problems immediately. Chronic leukemias have an insidious course and are frequently discovered during an examination for an unrelated problem. For both types, the most consistent symptoms are nonspecific and include weakness, fatigue, mild weight loss, and low-grade fever.

Often the leukemic clone first affects blood-forming cells in the bone marrow. To examine the bone marrow, a sample is removed with a needle from the pelvis and then stained and viewed under the microscope. Depending on the type of leukemia, a partial or near-total replacement of normal cells will be apparent. The great numbers of leukemic cells may obliterate all normal bone marrow architecture, including fat spaces that occupy 20% of the marrow cavity in children and as much as 70% in the elderly. It was once thought that decreased production of blood cells was secondary to "crowding out" by the leukemic cells, but it is now clear that the leukemic cells themselves or their products inhibit normal cell production.

Leukemic cells are found in the bloodstream. Microscopic examination of a drop of stained venous blood may reveal only a rare abnormal cell or a vast sea of leukemic cells that number as many as 50 times the normal total white cell count, as in chronic lymphocytic and acute myeloblastic leukemias. Leukemic cells may infiltrate organs and tissues other than the blood and bone marrow and cause disturbances in normal function. Their spread to the brain, meninges, spinal cord, kidneys, intestine, adrenal glands, lungs, and heart may cause pain, loss of function, or obstruction. However, symptoms referable to organs other than the bone marrow usually do not appear until later stages of the disease. The chronic leukemias are exceptions since the symptoms that they often present are referable to an enlarging spleen or liver, and these organs may infarct or rupture as they become progressively engorged with leukemic cells that compromise the blood supply. Another exception occurs in the myeloid leukemias, where tumors are composed of masses of leukemic cells called chloromas. Chloromas contain an enzyme that imparts an evanescent greenish tinge to their cut surface and can precede a diagnosis or herald a relapse of leukemia before it is detectable elsewhere. Chloromas have a predilection for soft tissues and are usually found in or near the sinuses and orbital areas of the skull, vertebrae, lymph nodes, skin, and gonads.

Because the leukemic cell retains most of the physiologic properties of its normal counterpart, the leukemic clone may appear clinically to be a caricature of a function or attribute of the specific cell type from which it originated. Thus, acute progranulocytic leukemia is invariably accompanied by disturbances in clotting and bleeding due to the release of procoagulant substances from granules in the leukemic cells. Hemorrhage can be fatal, particularly after therapy when leukemic cells die and disgorge their contents. Acute monocytic leukemia, which exaggerates the normal monocyte function of egress into sites of inflammation or infection, is often accompanied by a marked degree of organ and tissue infiltration, especially of the gums, spleen, liver, and lymph nodes.

Therapy. Practical therapeutic goals for the acute and chronic leukemias are distinct. Without prompt, intensive, in-hospital therapy, the acute leukemias usually cause death within a few months. In acute leukemia, the object of therapy is to totally obliterate the leukemic clone and allow normal bone marrow cells to recover. This is achieved through large doses of antileukemic drugs that are potent enough to eliminate all evidence of the leukemia. Unfortunately, such agents lack specificity and destroy normal cells particularly the dividing cells in the bone marrow but the doses are calculated to spare normal stem cells that can eventually repopulate the bone marrow and resume normal blood cell production. After the initial intense treatment, repeated cycles are given for as long as 3 years. Specific treatment of localized collections of leukemic cells, especially in so-called sanctuary sites such as the central nervous system, may be included to prevent relapse.

In the chronic leukemias, standard therapeutic principles are completely different. Many patients who initially require no therapy begin mild forms of outpatient treatment as the disease progresses. The intent is not to cure but to control the disease with minimal toxicity. A certain degree of organ compromise (including the bone marrow) is tolerated as long as it is not life-threatening. When therapy is necessary, it usually consists of milder forms of chemotherapy that do not require hospitalization.

Chemotherapeutic agents can be given singly or in regimens that combine two or more drugs. Radiation therapy is used in the preparative phases before bone marrow transplantation, in preventing central nervous system relapse, and in treating focal deposits of leukemic cells.

Chemotherapy. Many effective antileukemic agents have been synthesized. Combination therapy incorporates drugs that have different modes of action and different toxicities in order to increase cytotoxic potency, account for leukemic cells that may be resistant to a single agent, and lessen cumulative toxicity in any particular organ or tissue. Most antileukemic drugs act by perturbing enzymes or substrates that are related to DNA or RNA synthesis and thus largely affect actively dividing cells. Any treatment must be repeated since the number of leukemic cells may exceed one trillion and a single course of antileukemic drugs will destroy only some of them.

Among the first drugs used to treat leukemia were alkylating agents, including cyclophosphamide and chlorambucil, which act by cross-linking DNA and consequently hindering replication and transcription. Similar effects are produced by ionizing radiation. Paradoxically, it is this effect on genetic material that is thought to account for the increased risk of secondary malignancies after exposure to alkylating agents. Anthracyclines, such as doxorubicin,

daunorubicin, and idarubicin, inhibit replication by insertion between DNA base pairs and are among the most effective antileukemic agents. Plant alkaloids, such as vincristine and vinblastine, which are derived from the periwinkle, prevent cellular division by inactivating the mitotic apparatus. They are most useful when included as part of a combination regimen, since their mode of action and side effects are different from other commonly used agents. Antimetabolites have been synthesized to be molecular mimics of normal substrates required for processes that are integral to cell survival. They become incorporated into the leukemic cell and, because of chemical substitutions at important sites on the molecule, block essential DNA synthetic pathways and ultimately lead to cell death. The nucleosides pentostatin, fludarabine, lymphoma, and cladribine are most effective in chronic lymphocytic leukemia, lymphoma, and hairy cell leukemia. *See* CHEMOTHERAPY.

Stem cell transplantation. Stem cell transplantation has had the most positive impact on the leukemia cure rate. Chemotherapy is administered alone or with radiation therapy in doses much higher than those used in standard antileukemic regimens to abolish the leukemic clone at the expense of the normal stem cells in the bone marrow. Patients would die following such treatment unless "rescued" with cryopreserved stem cells. The thawed stem cells are infused into a vein to circulate in the bloodstream, occupy the marrow spaces, and eventually proliferate and produce all of the various blood cells. The stem cells must come from the patients themselves (autologous stem cell transplant) or from a donor (allogeneic stem cell transplant) whose human leukocyte antigens, commonly called HLAs, match those of the patient's cells as closely as possible (HLAs determine tissue compatibility). In allogeneic stem cell transplants, stem cells are harvested from the peripheral blood of an identical twin (a syngeneic match) or a sibling or unrelated donor (allogeneic). However, the risk of potentially fatal graft-versus-host disease is significant with donor cells that are less than ideally matched. Many patients do not have eligible siblings and require unrelated matched donors. The elderly, who comprise a large segment of those with leukemias, are generally not candidates because of the toxicity involved. Nonetheless, stem cell transplantation is the therapy of choice for eligible patients with acute lymphoid, acute myeloblastic, and other high-risk leukemias who have relapsed after standard chemotherapy. Allogeneic stem cell transplantation is used for patients with chronic myelogenous leukemia who have failed imatinib treatment (see below) and rarely for patients with chronic lymphocytic leukemia and other leukemias. *See* STEM CELLS.

Biologic response modifiers. The search for therapies which are less toxic and more specific for leukemic cells has focused on substances that are derived from natural (biologic) sources or that affect biologic reactions, some of which are thought to be part of the body's natural defense against cancer. Examples include monoclonal antibodies, small molecules that target protein kinases that affect the cell cycle and lead to tumor cell death, cell products manufactured by recombinant DNA technology, and the patient's own killer cells (large, granular lymphocytes that can lyse a variety of target cells) expanded and activated in the laboratory before reinfusion. Alpha-interferon can be used in certain leukemias, including hairy cell and chronic myelogenous leukemia. Recombinant growth factors, such as granulocyte-macrophage colony stimulating factor to promote leukocyte production, show potential as a therapeutic adjunct, shortening the duration of low cell counts after therapy and possibly augmenting the numbers and activity of cells that can destroy leukemic cells. Rituximab, a chimerized (comprising approximately 33% mouse protein and 67% human protein) anti-CD20 monoclonal antibody, is commonly used in the treatment of B-lymphocyte-derived malignancies, including chronic lymphocytic leukemia. Alemtuzumab is a humanized (comprising 5–10% mouse protein and 90–95% human protein) anti-CD52 monoclonal antibody approved for second-line treatment of chronic lymphocytic leukemia. Gemtuzumab ozogamicin is a humanized anti-CD33 monoclonal antibody conjugated to a potent antitumor antibiotic, calicheamicin, and is approved for the treatment of acute myeloblastic leukemia. Imatinib is a small molecule that targets the BCR-ABL tyrosine kinase unique to chronic myelogenous leukemia and sometimes found in acute leukemia. It has revolutionized the treatment of chronic myelogenous leukemia with over 80% of early diagnosed patients responding to therapy. Dasatinib is an inhibitor of multiple tyrosine kinases and has been approved by the U.S. Food and Drug Administration (FDA) for the treatment of imatinib-resistant patients with chronic myelogenous leukemia. *See* CELLULAR IMMUNOLOGY; MONOCLONAL ANTIBODIES.

Complications and prognosis. Each type of leukemia and every therapeutic modality has unique complications. However, because of bone marrow compromise by therapy or the leukemia itself, the most common problems are low numbers of platelets, red blood cells, and neutrophils. Before the advent of blood-product transfusions, most patients—particularly those with acute leukemia—would die of hemorrhage due to low platelet numbers. Decreased production of red blood cells compounded the anemia from blood loss. Now, however, platelet and red cell transfusions can sustain patients with little to no production of their own blood cells for many months. Transfusion of neutrophils, on the other hand, has not been practical and is rarely used. Without neutrophils, patients become susceptible to infections with organisms that normally live as harmless commensals on the skin and mucous membranes and in the gastrointestinal tract. Antimicrobial therapy has succeeded in treating these infections, but the elimination of resident microbes leaves the leukemic patient open to infections caused by unusual and resistant forms of bacteria, fungi, viruses, and protozoa. Thus, infection has replaced hemorrhage as the leading

cause of morbidity and death and overshadows all other complications of therapy. Death from infections usually occurs because therapy fails either to eradicate the leukemia or to promptly restore normal blood cell production. The latter could correct the single most important cause of susceptibility to infections decreased neutrophils. *See* INFECTION.

Despite the differences between the acute and chronic leukemias, and despite the fact that the great majority of patients can be brought into a remission or quiescent phase of the disease, leukemia is one of the most lethal malignancies. If cure is considered to be the absence of disease 3–5 years after cessation of therapy, only a small fraction of all leukemias are curable. An exception is acute lymphoid leukemia in children, where therapeutic advances have resulted in the attainment of a complete remission and cures in the majority of children.

Survival times in other leukemias are different and range from months in some of the acute forms to as long as 20 years in the chronic types. Overall, for adult and childhood acute myeloblastic leukemia, long-term survival is attained in only 20–30% of cases. The figures for adult acute lymphoid leukemia are only slightly better. Stem cell transplantation offers eligible patients a 40–60% chance for a cure. Biologic response modifiers have clearly had an impact on the therapy of many patients with leukemia. *See* CANCER (MEDICINE); ONCOLOGY.

Kenneth A. Foon; Louis Vaickus

Bibliography. D. Catovsky (ed.), *Leukemic Cell*, 2d ed., 1991; J. Fleischer (ed.), *Leukemias*, 1993; A. M. Mauer (ed.), *The Biology of Human Leukemia*, 1990; P. H. Wiernik (ed.), *Neoplastic Diseases of the Blood*, 3d ed., 1995.

Level measurement

The determination of the location of the interface between two fluids, separable by gravity, with respect to a fixed horizontal datum plane. The respective fluids may be any fluids, liquid or gaseous, which do not mix and have specific gravities significantly different from one another. Fluids include granular or particulate solids which are fluidized or handled like fluids. The most common level measurements are, however, between a liquid and a gas or vapour. *See* FLUID; INTERFACE OF PHASES.

Level measurement may be classified in the main categories: direct visual indication of interface location; remotely transmitted indication of interface location; interface location inferred from hydrostatic pressure; interface location inferred from fluid properties. Level measurements may be of an analog or on-off nature.

Direct visual measurement. In many cases, fluid levels may be observed directly and consequently measured to obtain trends or magnitudes in volume.

Graduated scale. Level is measured directly from a vertical graduated scale partially immersed in the liquid.

Glass window. Level is observed through a transparent window in the side of a tank. The window may be graduated with a vertical scale.

Gauge glass. Level is observed in a transparent vertical tube attached to a closed tank. The bottom of the tube is connected to the liquid space and the top of the tube to the gaseous space. The liquid level in the tube corresponds with the level in the tank and may be observed or measured against a graduated scale. Isolating valves usually are fitted in the upper and lower connecting pipes to allow for replacement of the transparent tube (which is usually of glass) without draining the tank (**Fig. 1***a*).

Closed tanks are often under some pressure, hence the use of external tubes able to withstand pressure rather than windows. In pressurized systems the upper and lower connecting tubes are fitted with ball check valves to avoid a dangerous discharge of fluid in the event of a tube rupture. For very high pressure systems, metallic ducts with thick glass windows are used. These windows are sometimes made refractive and artificially illuminated to show more clearly the difference between the two fluids (such as water and steam).

Remote measurement from float. Where levels cannot be observed and hence measured directly, it is common to use a float and to indicate remotely the elevation of this float.

The float must be of an average density between that of the two fluids, the densities of which must be significantly different (such as water and air) to ensure that a sufficient buoyant force is generated on the float, with changing level, to activate the position-sensing mechanism.

The float may generate an analog signal which varies over the whole range of operating level, or may generate an on-off signal as the level rises above or falls below a predetermined elevation. A series of on-off sensors at different elevations can generate a digital type of level indication.

Float with mechanical indicator. Several simple methods allow the float position and hence level to be observed indirectly from outside the vessel containing the fluids. A vertical rod attached to the float and protruding through a hole in the top of the tank can show the level by the amount of protrusion of the rod, which may be graduated. An external weight attached to the float by a rope or tape running over a pulley at the top of the tank can show the level on an inverse scale outside the tank. A rotating shaft passing through a sealed hole in the side of the tank can be connected to the float by a lever so that any rise or fall in the level rotates the shaft appropriately and so moves a pointer on an external graduated scale.

Float with electrical resistance sensor. In a closed vessel a float having a lever connected to a variable-resistance sensor can cause a change in the electrical resistance as the level rises or falls. The change in electrical resistance can be used, via a suitable calibrated electrical instrument, to indicate the level or volume in the tank (Fig. 1*b*).

Float with magnetic switches. In a closed tank a float pivoting about a fixed point can move a magnet close to the wall of the tank down or up as the level rises or falls about a selected level. A similar magnet just outside the tank is flipped by magnetic repulsion to operate electrical contacts to give a corresponding on-off signal. Magnetic repulsion rather than magnetic attraction is employed to create a definite toggle effect (Fig. 1*c*).

Float with buoyancy effect. A float constrained in its vertical movement will exert a varying force on the restraining mechanism. This force in turn can be measured or converted into an analog electrical signal which can be calibrated to indicate the liquid level. In such an application, level can be measured only over the height of the float, so such floats usually have slender dimensions (Fig. 1*d*).

Measurement by hydrostatic pressure. The pressure p at any depth h in a liquid of density ρ is given by the following equation, $p = \rho g h$, where g is acceleration due to gravity. Hence, if the density of liquid is known, the depth of liquid above a selected point can be determined by measuring the pressure in the liquid at that point.

Pressure gauge. In an application with a simple pressure gauge, it may be calibrated to give a direct reading of the depth of liquid. If the tank is closed and there is pressure in the space above the liquid surface, the difference in pressure between this space and the measuring point must be used.

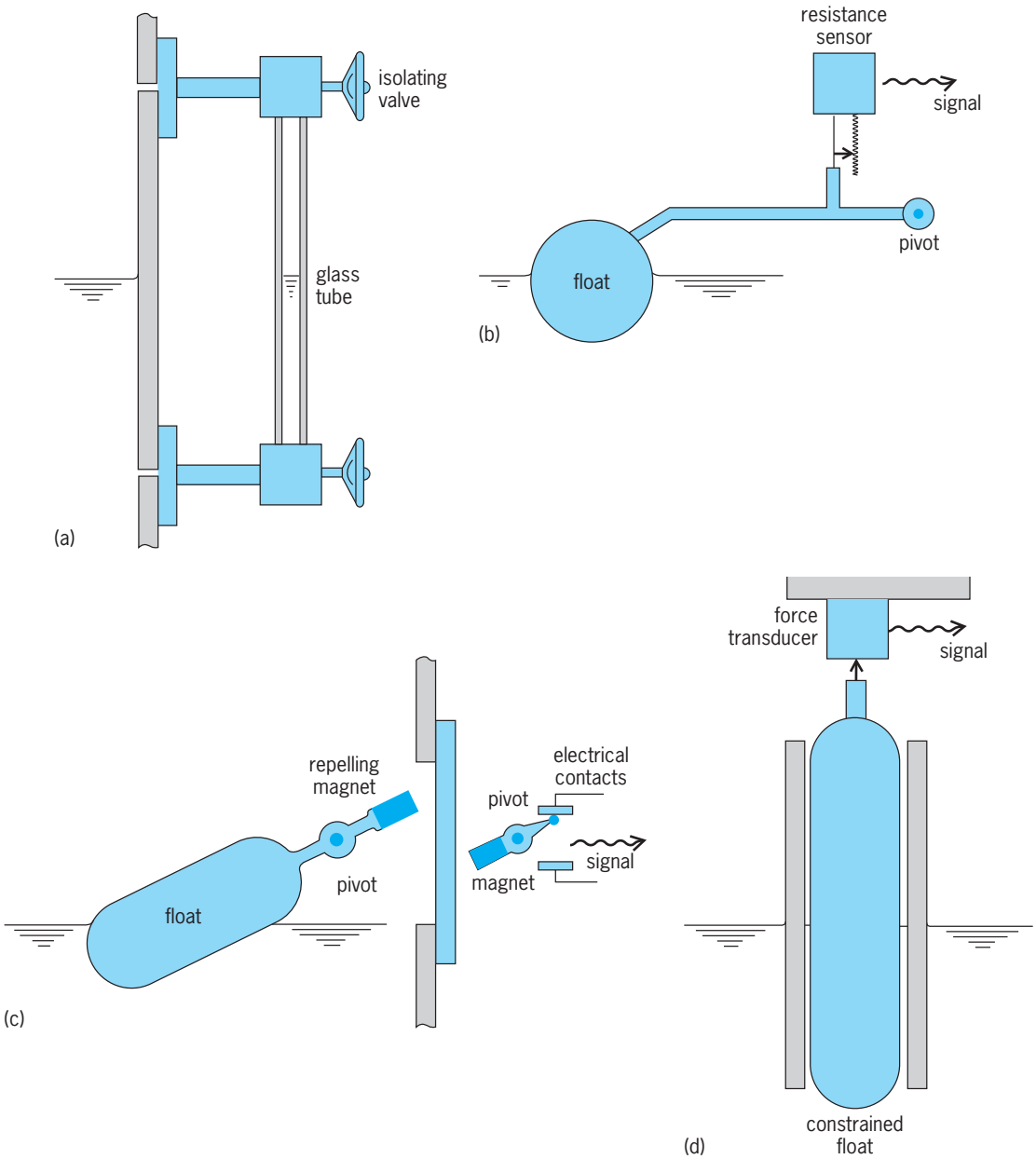

Fig. 1. Direct visual and float-type measurements. (*a*) Gauge glass. (*b*) Float with electrical sensor. (*c*) Float with magnetic switches. (d) Float with buoyancy effect.

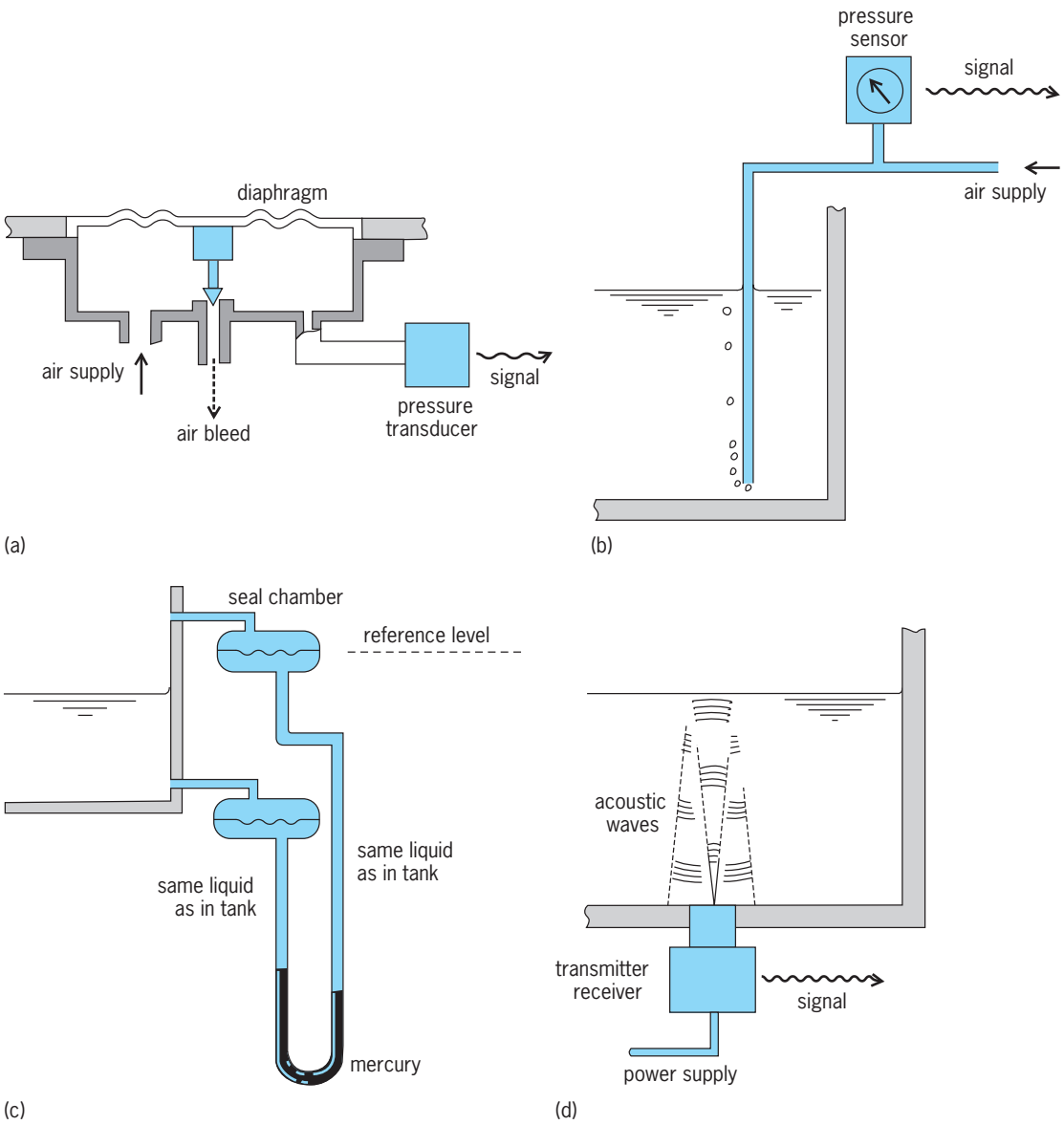

Fig. 2. Hydrostatic pressure and acoustic-wave measurements. (*a*) Pressure diaphragm. (*b*) Bubble tube. (*c*) Manometer. (*d*) Acoustic.

Pressure diaphragm. In applications where the liquid may be contaminated with aggressive impurities or contain solids or sludge, pressure gauges or their pressure-measuring tappings may become blocked and unresponsive. In such cases, pressure diaphragms installed flush with the inside surface of the vessel may be used. Movement of the diaphragm against a spring may be measured directly to give an indication of the pressure or depth of liquid. However, since such movement is limited, it is better to apply a corresponding pressure to the back of the diaphragm so as to restore the diaphragm to its neutral position. Measurement of this external pressure then can be converted into a measure of the depth of liquid (**Fig. 2***a*).

Bubble tube. In order to discharge gas from a pipe or vessel into a liquid at some depth, the pressure of the gas must be at least equal to that of the liquid at that depth. Hence the depth of liquid at the point of discharge can be determined from the gas pressure, provided the gas flow rate is low enough to eliminate dynamic or frictional effects. A pipe supplied with a steady but low flow of gas (such as air) may be inserted into a tank of liquid and the gas pressure in the pipe measured. This pressure measurement can then be converted into a measurement of the depth of liquid above the point of discharge (Fig. 2*b*).

Manometer. A manometer may be used instead of a pressure gauge for measuring pressure (more correctly pressure difference). A manometer using mercury as the reference liquid reduces the level variation by a factor of about 13, making direct measurement more convenient, and is more sensitive than a pressure gauge. Its location relative to that of the vessel in which the liquid level is being measured usually necessitates twin pipes from each side of the manometer extending back to the measuring points and filled with the same liquid as in the vessel (Fig. 2*c*).

Measurement by fluid properties. Certain fluid properties can be readily measured or used to determine the presence or extent of a known fluid. The presence or absence of a fluid can provide an on-off signal indicating whether a certain level has been reached or not, whereas the extent of a fluid can be used to determine the depth.

Conductivity. If a liquid is a conductor of electricity, its presence can be detected by a pair of electrodes subject to a potential difference. When immersed on rising level, they can generate an off-on electrical signal.

Capacitance. If a liquid is a dielectric, probes can be inserted into a tank and the capacitance between them measured. This will vary with the degree of immersion and can be converted to a measurement of level. *See* CAPACITANCE.

Acoustic. Most liquids conduct sound waves readily, and these are reflected from any interfaces, including the liquid surface. If an acoustic transmitter-receiver located at the bottom of a tank directs sound waves vertically upward and senses their reflection from the surface, the depth, and hence level, can be determined by the time taken for the sound wave to travel up and be reflected down (Fig. 2*d*). *See* SOUND.

Nuclear. Since gamma rays are absorbed by many liquids, the presence of liquid can be sensed by the attenuation of gamma rays emitted from a gamma-ray source and measured by a detector a short distance away. *See* GAMMA RAYS.

Thermal. If an electrically heated thermistor is subject to immersion in a liquid, it will be cooled more effectively. The resulting drop in temperature will be reflected as a change in resistance which will indicate the presence of the liquid. *See* THERMISTOR.

Robin Chaplin

Bibliography. E. O. Doebelin, *Measurement Systems: Application and Design*, 5th ed., McGraw-Hill, New York, 2003.

Lever

A pivoted rigid bar used to multiply force or motion, sometimes called the lever and fulcrum (**Fig. 1**). The lever uses one of the two conditions for static equilibrium, which is that the summation of moments about any point equals zero. The other condition is that the summation of forces acting in any direction through a point equals zero. *See* INCLINED PLANE.

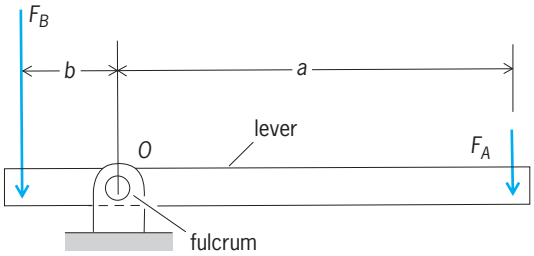

Fig. 1. The lever pivots at the fulcrum.

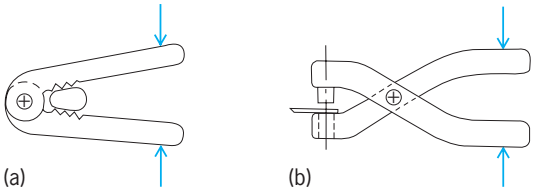

Fig. 2. Two applications of the lever. (*a*) Nutcracker. (*b*) Paper punch.

If moments acting counterclockwise around the fulcrum of a lever are positive, then, for a frictionless lever, $F_B b - F_{BA} a = 0$, which may be rearranged to give Eq. (1).

$$F_B = \frac{a}{b} F_A \tag{1}$$

If F_B represents the output and F_A represents the input, the mechanical advantage, MA, is then given by Eq. (2).

$$\mathrm{MA} = \frac{F_B}{F_A} = \frac{a}{b} \tag{2}$$

Applications of the lever range from the simple nutcracker and paper punch (**Fig. 2**) to complex multiple-lever systems found in scales and in testing machines used in the study of properties of materials. *See* SIMPLE MACHINE.

Richard M. Phelan

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

10 LIB-META

www.MHEST.com

Libra

The Scales, a southern zodiacal constellation (see **illustration**). It has no bright stars. Traditional pictures show a set of scales (a balance) with pans held by chains. Perhaps its name comes from the nearby constellation Virgo, who was the goddess of justice and who would have used scales to balance good and evil deeds. Also, the autumnal equinox, with daytime and nighttime of equal length, used to occur in Libra, though the precession of the equinoxes has moved that point to Virgo. *See* EQUINOX; PRECESSION OF EQUINOXES; VIRGO; ZODIAC.

Libra's brightest stars have the Arabic names Zubenelschamali and Zubenelgenubi—meaning northern and southern claws—that show that they once belonged to Scorpius. *See* SCORPIUS.

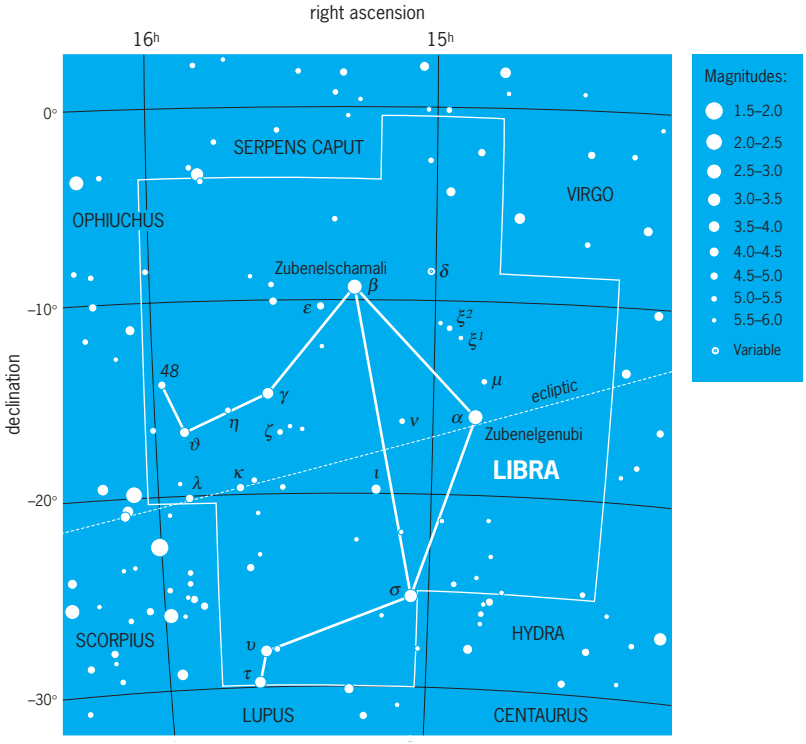

Modern boundaries of the constellation Libra, the scales. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Lichens

Symbiotic associations of fungi (mycobionts) and photosynthetic partners (photobionts). These associations always result in a distinct morphological body termed a thallus that may adhere tightly to the substrate or be leafy, stalked, or hanging. A thallus consists of layers, that is, a cortex and medulla made up of the fungus, and a photosynthetic layer of algal or cyanobacterial cells that are closely associated with fungal hyphae. Rhizoids anchor thalli to their substrates. Some lichens (*Stereocaulon*) with green photobionts have gall-like structures called cephalodia that contain cyanobacteria. *See* CYANOBACTERIA.

Lichens are formed from specialized groups of parasitic fungi; this association is one of a controlled parasitism rather than mutualism. Thus, the photobionts that lichen fungi slowly parasitize should be considered victims and not partners. Lichen-forming fungi share two characteristics with fungi that parasitize plants: concentric bodies and specialized branches of hyphae (haustoria) that penetrate host cells and absorb nutrients from them.

Distribution. Lichens have a worldwide distribution and grow on almost any inanimate object, such as bark, soil, roof tiles, or stone. Some also grow on living leaves of tropical plants and ferns, while a few grow inside sandstone rocks in the dry valleys of Antarctica. Lichens dominate about 8% of the Earth's surface, especially in the northern boreal forests of Canada and Siberia, where reindeer lichens form vast uninterrupted carpets of growth. They are among the hardiest of organisms and thrive in some of the Earth's harshest environments, such as polar regions, deserts, and high mountains.

Classification. The name given to a lichen applies only to the mycobiont, while the photobiont has a separate name. Most of the 15,000 lichen-forming fungi are in the fungal class Ascomycotina (ascolichens); common genera of lichens include *Cladonia, Parmelia, Lecanora, Lecidea*, and *Usnea*. Approximately a dozen species of basidiomycetes form lichens. *Dictyonema pavonium* is a common, tropical basidiolichen with lobed thalli and the blue-green *Scytonema* as photobiont. *Omphalina* forms gilled basidiocarps and two types of associations with the green alga *Coccomyxa* in the form of globular clusters (*Botrydina*-type thallus) or small lobes (*Coriscium*-type thallus). Lichens that do not have sexual reproduction (*Lepraria*) are placed in the Lichenes Imperfecti. *See* ASCOMYCOTA; BASIDIOMYCOTA.

Photobionts of lichens are either green algae (*Trebouxia, Trentepohlia*) or cyanobacteria (*Nostoc, Scytonema*). The most common photobiont is *Trebouxia*. This unicellular green algae has never been found in the free-living state. It is believed that *Trebouxia* is a lichenized and highly modified form of the filamentous alga *Pleurastrum terrestre*.

Growth and reproduction. Lichens grow very slowly. Indeed, the fastest-growing lichen from a temperate zone is *Ramalina menziesii*, which has a maximum annual radial growth of 1 in. (28 mm). Most temperate-zone lichens grow only 0.1–0.2 in. (2–6 mm) per year or less. Lichen growth can be measured in terms of an increase in linear dimension or biomass of thallus. Growth rates of thalli vary greatly according to season, habitat conditions, latitude, and age of the thallus, and even between the lobes of a single thallus. Individual lichens may live for hundreds and even thousands of years.

Fragmentation of thalli is a common way that lichens reproduce. Other asexual reproduction is by means of propagules that are simply outgrowths of the thallus and contain a few photobiont cells closely enclosed by fungal hyphae. Sexual reproduction is presumed to be common among lichen fungi because of their genetic variability and the presence of fungal sexual structures such as spermatia and ascogonial filaments. Ascospores, which are products of sexual recombination, are produced and discharged commonly by many lichens. What happens to the mycobionts after the spores germinate is not clear. Developing hyphae from spores may even steal photobiont cells from other lichens and use them to form their own thalli. *See* REPRODUCTION (PLANT).

Physiology. The basic metabolic processes of lichens are photosynthesis, respiration, and nitrogen fixation. Lichens have adapted these processes to different conditions of light, temperature, day length, and water. The mycobiont causes the photobiont to excrete most of the carbon that it fixes during photosynthesis. Only a single type of compound is excreted. The mycobiont absorbs these compounds and converts them to mannitol, its own storage compound. When dry lichens absorb water, they respire at levels much higher than normal for several hours. The mycobiont respires stored polyols, in the case of chlorolichens, instead of depleting valuable proteins and other structural compounds. *See* PLANT RESPIRATION.

Nitrogen-fixing lichens are common and contribute nitrogen to different ecosystems when they decay. In cyanolichens the mycobiont inhibits the nitrogen-assimilating enzymes of the cyanobiont, causing it to release most of the ammonia it produces. The ammonia is absorbed by the mycobiont and used to make proteins and nucleic acids. *See* NITROGEN FIXATION.

Chemistry. Lichens produce several hundred secondary compounds that accumulate as crystals in the thalli, often at high concentrations. These compounds may protect the slow-growing thalli from harmful bacteria, fungi, and insects and may play a regulatory role in the interactions between bionts. Lichen compounds have been produced by free-living fungi, especially species of *Aspergillus*, and by mycobionts growing alone in culture. Lichen

secondary compounds represent a new class of antibiotics in an age where standard antibiotics such as penicillin are becoming ineffective against antibiotic-resistant microbes. Secondary compounds are used extensively by taxonomists to characterize new taxa of lichens (chemotaxonomy).

Pollution indicators. Lichens have become useful bioindicators of air pollution because of their sensitivity to pollutants such as sulfur dioxide, hydrogen fluoride, ozone, and peroxyacetyl nitrate (PAN) and particulate pollutants such as lead, copper, and zinc. The most widespread and damaging pollutant for lichens is sulfur dioxide, which is most toxic when the pH of the substrate is acidic and the thallus is wet. Nitrogen fixation is the most sensitive to sulfur dioxide, followed by photosynthesis and respiration. Acid rain adversely affects lichen communities but also opens up new niches for acid-loving species. Pollution zones around many major cities have been mapped by using lichens. Lichens have also been useful in measuring radioactive fallout from explosions of nuclear bombs and nuclear reactors. *See* AIR POLLUTION.

Vernon Ahmadjian

Bibliography. V. Ahmadjian, *The Lichen Symbiosis*, 1993; V. Ahmadjian and M. E. Hale (eds.), *The Lichens*, 1973; C. F. Culberson, *Chemical and Botanical Guide to Lichen Products*, 1969; M. Galun (ed.), *CRC Handbook of Lichenology*, 3 vols., 1988; M. E. Hale, *The Biology of Lichens*, 1983; T. H. Nash and V. Wirth (eds.), *Lichens, Bryophytes, and Air Quality*, 1988.

Licorice

A product obtained from the licorice plant (*Glycyrrhiza glabra*) of the legume family (Leguminosae). It is a perennial herb which grows wild and is cultivated in southern Europe and in western and central Asia. The roots are dried for several months and then packaged for shipment. Spain leads in the production of cultivated licorice roots. Licorice is used in medicine to mask objectionable taste and as a laxative; as a flavoring material in the brewing, tobacco, and candy industries; and in the manufacture of shoe polish. *See* ROSALES; SPICE AND FLAVORING.

Perry D. Strausbaugh; Earl L. Core

Lidar

The optical analog of radar. The term lidar is an acronym for light detection and ranging. Lidar systems employ intense pulses of light, typically generated by lasers, and large telescopes and sensitive optical detectors to receive the reflected pulses. They are most commonly used to measure the composition and structure of the atmosphere. The very narrow beamwidth, narrow linewidth, and ultrashort pulses of the laser make it possible to optically probe the atmosphere with exceptional sensitivity and resolution. When used to measure the range and velocity of hard targets, lidars are usually called laser ranging systems or laser radars. *See* LASER; OPTICAL DETECTORS; RADAR.

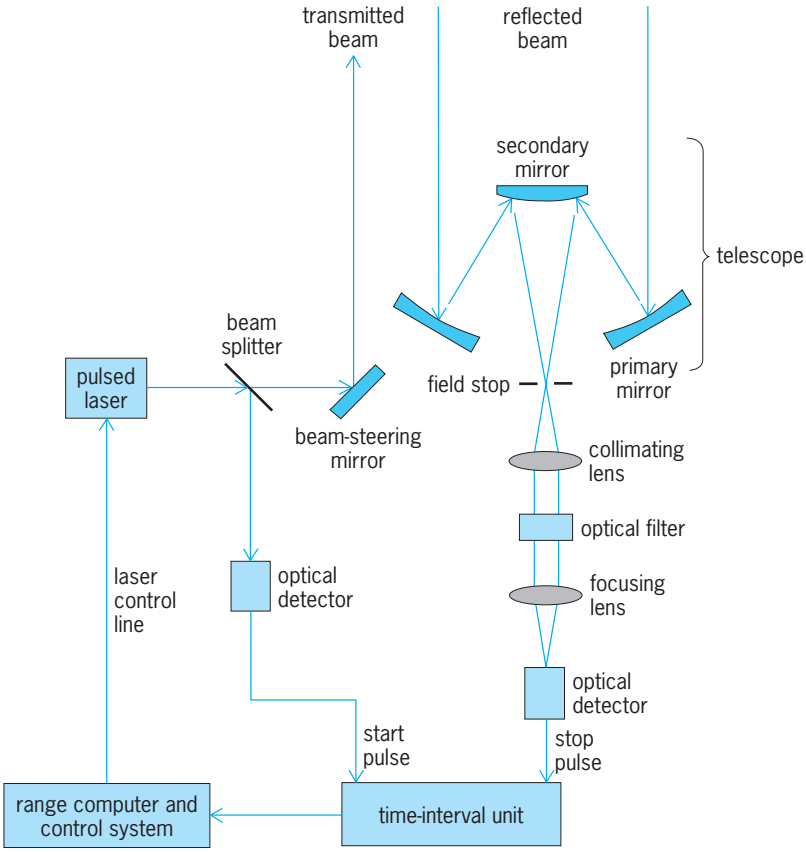

Fig. 1. Block diagram of a typical monostatic system used for ranging or altimetry applications.

Ranging and altimeter systems. The most common lidar configuration is the monostatic system (**Fig. 1**). The laser beam is either projected through the receiving telescope or propagates parallel to the optical axis of the telescope. If the system is designed for ranging or altimetry, the receiver measures the roundtrip propagation time of the laser pulse between the lidar and the target. The distance d to the target is given by the range equation below, where

$$d = \frac{c\tau}{2}$$

c is the speed of light and τ is the roundtrip propagation time of the laser pulse. Precision timing is accomplished electronically by a device called the time interval unit. A beam splitter directs a small fraction (<1%) of the transmitted laser beam to a photodetector. When illuminated by the laser, the photodetector generates an electrical pulse which triggers (starts) a precision clock in the time interval unit. The reflected laser pulse is collected by the receiving telescope and focused onto another, more sensitive photodetector, generating an electrical pulse that stops the clock. The elapsed time is then transmitted to the range computer, which calculates the target distance according to the range equation. The optical filter and field stop in the telescope are

designed to reduce interference and noise caused by other sources of light. The optical filter limits light transmission to the narrow wavelength range (color) of the laser, and the field stop limits the field of view of the telescope to the region illuminated by the laser beam. *See* INTERFERENCE FILTER.

The ranging accuracy depends upon many factors, including the laser pulse length, the received signal strength, and the timing accuracy of the time interval unit. The most sophisticated systems are used for ranging to retroreflector-equipped satellites and to the retroreflector arrays placed on the Moon by the Apollo astronauts. Accuracies of a few centimeters are achieved routinely. Data from these measurements are used to monitor geophysical phenomena such as continental drift, crustal dynamics, and the Earth's rotation rate. Because of the extremely high accuracy required, a more sophisticated version of the range equation must be used which includes the effects of the additional propagation delay introduced by the Earth's atmosphere.

The first successful laser ranging measurements to satellites were conducted in 1964. In the early 1970s, the first space-based laser altimeter was operated in lunar orbit from the *Apollo 15, 16,* and *17* command and service modules. Airborne laser altimeters provide maps of surface topography, coastal water depth, forest canopies, sea ice distribution, volcanic landforms, impact craters, and ocean wave heights. *See* REMOTE SENSING.

Atmospheric lidars. The targets of atmospheric lidars are either suspended dust and aerosols or gas molecules which are continuously distributed in the atmosphere along the propagation path of the laser beam. Atmospheric lidar systems are used to measure density profiles of the scatterers. Profile measurements are accomplished by pulsing the laser and then periodically sampling the detector output (**Fig. 2**). The sampling process is called range gating. The signal level is proportional to the density of scatterers and inversely proportional to the square of the distance to the scattering volume. Because reflected signals from atmospheric aerosols and gas molecules are usually very weak, most atmospheric lidars employ very sensitive photomultiplier tube detectors that operate in the photon-counting mode. The photomultiplier tube signal is processed by a pulse discriminator and range-gated counter. The pulse discriminator converts the low-level (few tenths of a volt) photomultiplier tube pulses generated by the individual detected photons to pulses that can be counted by conventional high-speed digital circuitry. The raw lidar data consists of a series of photon counts corresponding to consecutive range bins. The count in a given range bin is related to the number of scatterers in the corresponding volume illuminated by the laser pulse. *See* ATMOSPHERIC CHEMISTRY; PHOTOMULTIPLIER.

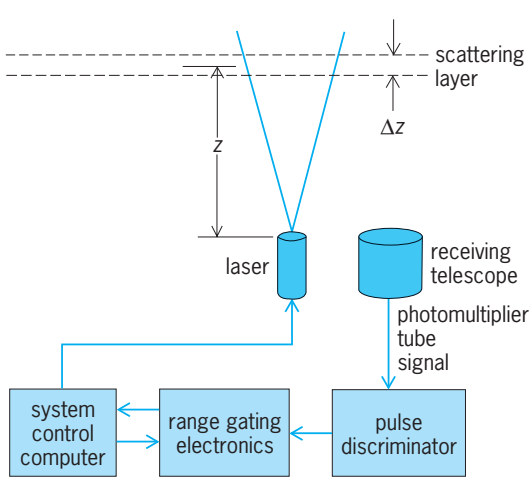

Fig. 2. Block diagram of a typical atmospheric lidar system. Range to scattering volume = *z;* thickness of scattering volume = Δ*z*.

In order to process the photon count profile to yield the density profile of scatterers, other factors, such as the thickness of the scattering volume and the backscatter cross section, must also be taken into account. The thickness of the scattering volume is related to the laser pulse length and the receiver range gate length. The backscatter cross section is the effective cross-sectional area of the scatterer, and depends upon many factors including scattering mechanism, optical wavelength, and optical properties of the scatterer. Only when the scatterer is large compared to the wavelength, which may be the case for aerosol scattering, does the physical size of the scatterer influence the value of the optical backscatter cross section. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

Atmospheric lidars are classified according to the type of scattering mechanism exploited to make the measurement.

Aerosol lidars. These measure scattering from atmospheric dust and aerosols. Observations with ground-based and airborne aerosol lidars have provided much of the data about the dispersion and global distribution of volcanic aerosols and about the life cycle and distribution of polar stratospheric clouds. Both phenomena occur in the altitude range between 10 and 35 km (6 and 22 mi). Volcanic aerosols can have a significant influence on climate, while polar stratospheric clouds play a major role in the springtime depletion of ozone over both polar caps.

Rayleigh lidars. These are designed to measure the molecular scattered signal, which is proportional to atmospheric density. The atmospheric temperature profile can be calculated from the density profile by using the ideal gas law and the hydrostatic equation. Density and temperature profiles at altitudes up to 100 km (60 mi) have been obtained with powerful ground-based systems. *See* ATMOSPHERE.

Resonance fluorescence lidars. These are used to measure the density profiles of specific molecular species such as sodium in the upper atmosphere. The laser wavelength must be tuned to the resonance absorption wavelength of the species of interest. When illuminated at the resonance wavelength by the lidar beam, the atoms fluoresce because some of the photons are resonantly absorbed and then reradiated. The resonant backscatter cross section is typically

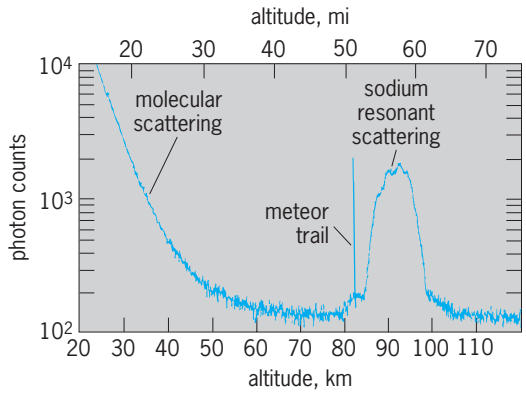

Fig. 3. Photon count profile measured with a sodium resonance fluorescence lidar.

many orders of magnitude larger than the molecular backscatter cross section of the species. **Figure 3** is a photon count profile measured with a sodium resonance fluorescence lidar. The high counts at altitudes below 70 km (45 mi) represent molecular scattering. Sodium resonant scattering is quite strong between 80 and 105 km (50 and 65 mi). The very thin scattering layer at 82 km (51 mi) is a meteor trail. Meteoric ablation is the major source of sodium and other metal layers in the upper atmosphere. By studying the density perturbations of these layers, the characteristics of the background wind perturbations and turbulence can be deduced. *See* METEOR.

Differential absorption lidar (DIAL). This technique is used to measure species concentrations in the lower atmosphere. DIAL systems employ two lasers, one tuned to an absorption line of the species of interest and the other tuned just off the absorption line. The received signal at the more strongly absorbed wavelength is weaker. By comparing the signal levels at the two wavelengths, a density profile of the species can be calculated. A wide variety of important minor constituents in the lower atmosphere, including ozone, are being studied by using DIAL systems.

Raman lidars. These measure the scattered signal at the Raman shifted wavelength. Because the Raman signal is very weak, measurements are usually restricted to the troposphere at altitudes below 10 km (6 mi). However, at these lower altitudes, Raman lidars have provided excellent measurements of atmospheric density, temperature, and water vapor concentration. *See* RAMAN EFFECT.

Doppler lidars. These are used to measure tropospheric winds. The scattered laser pulse is Doppler-shifted in frequency in proportion to the velocity of the scatterer along the propagation path. The frequency shift is measured by combining the reflected signal with the transmitted laser beam, so that the electrical signal generated by the detector is at a frequency equal to the difference between the frequencies of the transmitted and reflected beams. The signal-processing electronics following the photodetector are very similar to those used in Doppler radars. By scanning the lidar beam, the three-dimensional wind vector can be measured. Mobile Doppler lidars have been used to map wind fields associated with a variety of atmospheric phenomena such as frontal passages, downslope flows and gusts, canyon drainage flows, and down bursts caused by severe thunderstorms. *See* DOPPLER EFFECT; DOPPLER RADAR.

Chester S. Gardner

Bibliography. V. A. Banakh and V. L. Mironov, *LIDAR in a Turbulent Atmosphere*, 1988; J. W. Bilbro et al., Airborne Doppler lidar measurements, *Appl. Opt.,* 25:3952–3960, 1986; S. C. Cohen et al., The geoscience laser altimetry/ranging system, *IEEE Trans. Geosci. Remote Sens.,* GE-25:581–592, 1987; C. S. Gardner and J. A. Reagan (eds.), Special section on lidar applications, *Proc. IEEE,* 77(3):408–478, 1989.

Lie detector

A device intended to detect an involuntary physiological response that all persons exhibit when lying but never when telling the truth. Because there is no such specific lie response, the lie detector of popular fancy is mythological. In actual "lie protector" tests, widely used in the United States, breathing movements, blood pressure changes, and electrodermal responses are recorded on a polygraph while the respondent answers "yes" or "no" to a series of 8 to 12 questions. From the polygraph recordings, one can determine whether "relevant" questions had a greater impact on the respondent than did the interpolated "control" questions. In the standard lie test used in specific issue investigations, the relevant questions ask whether the respondent committed the act in question; for example, "On April 12, did you take $2000 from the office safe?" A typical control question might be, "In the first 20 years of your life, did you ever steal anything?" If the examinee reacts more strongly to the relevant than to the control questions, it is inferred that his or her answers to the relevant questions are deceptive. Because an innocent accused also may be disturbed by the relevant questions and react more strongly to them than to the controls, the lie test is biased against the truthful respondent. Research has shown that as many as 50% of innocent criminal suspects may "fail" control-question polygraph tests.

Polygraph tests are also widely used in the United States for preemployment screening. Screening tests normally do not use control questions, but include instead a series of relevant questions that are of interest to the prospective employer: "Have you ever stolen from a previous employer?"; "Have you ever used street drugs?" Strong physiological reactions to any of these questions are interpreted as indicating a deceptive reply. There has been no scientific study to prove the validity of this type of polygraph test, however.

Several devices have been marketed for use in lie detection which are alleged to measure "stress" in the respondent's voice. Because they do not require attachments to the body, such devices can be used covertly and can even be applied to broadcasts or to recordings of the voices of dead people. It has not

yet been shown that these devices can reliably distinguish between different degrees of voice stress. When voice analyzers have been used for lie detection, in parallel with a conventional polygraph, they have done no better than chance in distinguishing truth from deception.

In another type of polygraphic interrogation, the "guilty knowledge" test, the relevant questions name some fact that should be known only by someone who had been present at the crime; the control questions name alternatives that should seem equally plausible to innocent suspects. A differential physiological response to the relevant question suggests that the suspect recognizes its relevance and therefore that he or she possesses guilty knowledge. Although it would seem to hold promise as a tool of criminal investigation, the guilty knowledge test has not yet been adequately studied in real-life applications. *See* ELECTRODERMAL RESPONSE; EMOTION.

David T. Lykken

Bibliography. S. Abrams, *The Complete Polygraph Handbook,* 1989; G. Ben-Shakhar and J. J. Furedy, *Theories and Applications in the Detection of Deception*, 1989.

Lie group

A topological group with only countably many connected components whose identity component is open and is an analytic group. An analytic group or connected Lie group is a topological group with the additional structure of a smooth manifold such that multiplication and inversion are smooth. Many groups that arise naturally as groups of symmetries of physical or mathematical systems are Lie groups. The study of Lie groups has applications to analytic function theory, differential equations, differential geometry, Fourier analysis, algebraic number theory, algebraic geometry, quantum mechanics, relativity, and elementary particle theory. *See* GROUP THEORY; MANIFOLD (MATHEMATICS); TOPOLOGY.

Examples of Lie groups. Euclidean n-dimensional space with the standard differentiable structure and with vector addition as the group operation is an analytic group $\mathbf{R}^n$ of dimension n.

The nonsingular n-by-n real matrices form a Lie group $GL(n,\mathbf{R})$ with the differentiable structure obtained as an open set of n^2-dimensional space and with the group operation given by matrix multiplication. Similarly $GL(n,\mathbf{C})$, the group of nonsingular n-by-n complex matrices under matrix multiplication, forms a Lie group. The latter group is connected.

Each closed subgroup of a Lie group, in its relative topology, admits one differentiable structure on its identity component that makes it a Lie group. With $GL(n,\mathbf{R})$ and $GL(n,\mathbf{C})$, the subgroups of upper triangular matrices, matrices of determinant 1, rotation matrices, orthogonal matrices, and unitary matrices are all examples. *See* MATRIX THEORY.

Lie algebra. Let G be a Lie group. The identity component of G is an open subset of G and is a smooth manifold. Thus it makes sense to speak of tangent spaces. Let $\mathfrak{g}$ be the tangent space to G at the identity element. A multiplication $[X,Y]$ within $\mathfrak{g}$ is defined as follows.

By means of the differential of left translation by p, one can associate a tangent vector X_p to the point p of G. If f is a smooth real-valued function on G, let $\tilde{X}f$ be the function on G with $\tilde{X}f(p) = X_p f$. Then $\tilde{X}f$ is smooth. The tangent vector $[X,Y]$ is defined on the function f by Eq. (1).

$$[X, Y]f = X\left(\tilde{Y}f\right) - Y\tilde{X}f \tag{1}$$

The resulting multiplication satisfies the following properties:

i. $[X,Y]$ is linear in X and in Y.
ii. $[X,X] = 0$.
iii. $[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0$.

The vector space $\mathfrak{g}$ equipped with this multiplication is called the Lie algebra of the Lie group G.

The Lie algebra $\mathfrak{g}$ determines the multiplication in G near the group identity by a result known as the Campbell-Hausdorff formula. Consequently, many properties of the identity component of G can be deduced from the simpler object $\mathfrak{g}$.

Homomorphisms. If φ is a smooth homomorphism from an analytic group G to an analytic group G', the differential $d\varphi$ of φ at the identity is a Lie algebra homomorphism in the sense that it is linear and satisfies $d\varphi([X,Y]) = [d\varphi(X),d\varphi(Y)]$. If φ is one-to-one or onto, so is $d\varphi$. Conversely, if G is simply-connected, each homomorphism of Lie algebras arises as the differential of some unique smooth group homomorphism.

Abstract Lie algebras. An abstract real Lie algebra is a finite-dimensional real vector space with a multiplication satisfying (i), (ii), and (iii). Each abstract real Lie algebra can be identified with the Lie algebra of some analytic group.

Examples. The Lie algebra of $\mathbf{R}^n$ is an n-dimensional vector space with all $[X,Y]$ equal to 0.

For $GL(n,\mathbf{R})$, the Lie algebra can be identified with the vector space of all n-by-n real matrices with multiplication given in terms of matrix multiplication as $[X,Y] = XY - YX$. The identification of a member L of the Lie algebra with a matrix X is done entry by entry. Namely, the i-jth entry of X is the result of applying L to the i-jth entry function, which is a smooth real-valued function on the group.

For $GL(n,\mathbf{C})$, the Lie algebra can be identified with all n-by-n complex matrices with multiplication $[X,Y] = XY - YX$.

For a closed subgroup G of $GL(n,\mathbf{R})$ or $GL(n,\mathbf{C})$, inclusion is a smooth one-to-one homomorphism and allows one to identify the Lie algebra of G with a Lie subalgebra of matrices. The matrices of the Lie subalgebra can be computed conveniently as all entry-by-entry derivatives $c'(0)$ of smooth curves $c(t)$ that lie completely within the subgroup and have $c(0)$ equal to the identity matrix. For unitary matrices, for example, the result is skew-hermitian matrices. (A skew-hermitian matrix is equal to the negative of the transpose of its complex conjugate).

Analytic subgroups. Let G be a Lie group. An analytic subgroup H of G is a subgroup of G that admits the structure of an analytic group in such a way that the inclusion mapping of H into G is smooth. The topology of H as a manifold need not be the relative topology inherited from G.

If H is an analytic subgroup of G, the differential of the inclusion mapping identifies the Lie algebra of H with a Lie subalgebra of the Lie algebra of G. Distinct analytic subgroups give rise to distinct Lie subalgebras. Every Lie subalgebra arises from some analytic subgroup by this construction.

Exponential map. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Let $c(t)$ be a curve in G, that is, a smooth mapping from the real line into G, and suppose that $c(0)$ is the group identity. The image of d/dt under the differential dc of $c(t)$ at 0 is a member of $\mathfrak{g}$. For each X in $\mathfrak{g}$ there is one and only one such curve c_X such that $dc_X(d/dt) = X$ and $c_X(s + t) = c_X(s)c_X(t)$ for all s and t. The exponential mapping is the function from $\mathfrak{g}$ to G defined by $\exp X = c_X(1)$. It is smooth and, in a small open neighborhood of the group identity, has a smooth inverse function that provides local coordinates for G near the identity.

For $GL(n,\mathbf{R})$ or $GL(n,\mathbf{C})$ the exponential mapping is given by the series expansion for the ordinary exponential function e^X if the Lie algebra is identified with the Lie algebra of all matrices. *See* SERIES.

Fourier analysis. In many applications of Lie groups the notion of group representation is a basic tool. A unitary representation of the Lie group G on a Hilbert space H is a function R from G to the set of unitary operators on H such that $R(g_1g_2) = R(g_1)R(g_2)$ for all g_1 and g_2 in G and such that the mapping $g \to R(g)x$ is continuous from G into H for each x in H. [A unitary operator U on H is a linear operator from H onto H that preserves inner products; that is, $(Ux, Uy) = (x,y)$ for each x and y in H.] The representation R is irreducible if there is no proper nonzero closed subspace S such that $R(g)$ maps S into itself for all g in G.

Because of the applications, it is an important mathematical problem to classify the irreducible unitary representations of a given Lie group G and to investigate properties of decompositions of general unitary representations into irreducible ones.

In the analysis of the unitary representations, use of the Lie algebra plays an important role. If X is in the Lie algebra, it is possible to define $R(X)$ as a skew-hermitian operator on a dense subspace by differentiating $R(\exp tX)$ appropriately. Then $iR(X)$ extends to a self-adjoint operator that is easier to deal with than $R(\exp tX)$.

In many applications, H is a space of functions, and some linear operator or set of linear operators is acting in H. Typically, the representation of G is given by a group of transformations that commute with these linear operators. When the representation of G on H is decomposed suitably, one can expect the linear operators to act separately in each of the constituent spaces of H and to be easier to understand on each constituent space than on H.

The prototype for this analysis is the case of classical Fourier series of periodic functions on the line with period 2π. Here G is the circle group (the compact Lie group of real numbers modulo 2π, with addition modulo 2π as group operation), and H is the space $L^2(G)$ of square integrable functions on the interval $-\pi \le x \le \pi$. The unitary representation R is given by $[R(x_0)f](x) = f(x_0 + x)$, with addition taken modulo 2π. For this G, the irreducible unitary representations are all one-dimensional and are parametrized by the integers, the nth one being $R_n(x_0) = e^{inx_0}v$ for v in $\mathbf{C}$. The classical Fourier series expansion given in Eq. (2) corresponds to a de-

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx} \tag{2}$$

composition of R into an infinite orthogonal sum of all R_n, since a single R_n can be realized in the one-dimensional space of multiples of e^{inx} and since Eq. (3) holds. In this context, let L be a bounded

$$\begin{aligned} R(x_0)f(x) = f(x_0 + x) &\sim \sum c_n e^{inx_0} e^{inx} \\ &= \sum R_n(x_0)\left(c_n e^{inx}\right) \end{aligned} \tag{3}$$

linear operator on H that commutes with translations, that is, with all $R(x_0)$. An easy computation shows that L must carry e^{inx} into a multiple of itself. If that multiple is b_n, then L is operating in the space of R_n by multiplication by b_n. Equation (4) is then useful in studying L.

$$Lf(x) \sim \sum_{n=-\infty}^{\infty} b_n c_n e^{inx} \tag{4}$$

See OPERATOR THEORY.

Sample applications. Three examples will illustrate the principle of analyzing a linear operator by decomposing it with respect to a group of operators that commute with it.

Rotations and the Fourier transform. Let H be the Hilbert space of square-integrable functions on $\mathbf{R}^n$, let G be the rotation group, and let the representation of G on H be given by rotation of the coordinates. The Fourier transform is a linear operator on H that commutes with each member of G. The space H decomposes under the representation of G into a sequence of simpler spaces, the kth space being all sums of products of a radial function and a harmonic polynomial homogeneous of degree k. The Fourier transform has a simple form on the kth space, given as a one-dimensional integral involving a Bessel function. *See* BESSEL FUNCTIONS; FOURIER SERIES AND TRANSFORMS.

Solutions of differential equations. Some information about solutions in $\mathbf{R}^n$ to elliptic partial differential equations with constant coefficients can be obtained by using the Fourier transform. Here the differential operator commutes with the representation of $G = \mathbf{R}^n$ given by translation in the space of square-integrable functions on $\mathbf{R}^n$, and the Fourier transform is the device for decomposing the representation. For an operator with nonconstant coefficients, one can apply a freezing principle, replacing a

given equation by one with constant coefficients and obtaining an approximate solution near a point.

For many equations of subelliptic type, which are elliptic except for certain degeneracies, a similar analysis is possible by means of representations of a more complicated Lie group. In the simplest new case, the Lie group is the (non-abelian) group of 3-by-3 real upper-triangle matrices with ones on the diagonal, and the differential equation is Eq. (5) in $\mathbf{R}^3$. Other equations can be analyzed

$$\left(\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}\right)+\left(x\frac{\partial^2 f}{\partial y\partial z}-y\frac{\partial^2 f}{\partial x\partial z}\right)+\frac{1}{4}\left(x^2+y^2\right)\frac{\partial^2 f}{\partial z^2}=0 \qquad (5)$$

by a combination of this method and a freezing principle.

The laplacian $\Delta = \Sigma\partial^2/\partial x_j^2$ in R^n has constant coefficients and can be analyzed by the Fourier transform. But also it commutes with rotations and can be decomposed accordingly. In a ball, for example, the resulting decomposition of solutions of $\Delta u = 0$ corresponds to solving the equation by separation of variables. If the solutions are restricted to a sphere, the analysis amounts to a decomposition of the space of square integrable functions on the sphere into (finite-dimensional) eigenspaces of the spherical laplacian, each of which is the Hilbert space for an irreducible representation of the rotation group. More generally, information about the generalized wave operator given in Eq. (6) can be

$$L=\frac{\partial^2}{\partial x_1^2}+\cdots+\frac{\partial^2}{\partial_{x_m}^2}-\frac{\partial^2}{\partial x_{m+1}^2}-\cdots-\frac{\partial^2}{\partial x_{m+n}^2} \qquad (6)$$

obtained by doing a Fourier analysis of functions on an orbit in R^{m+n} under the action of the noncompact group of linear isometries of a real quadratic form with m plus signs and n minus signs. *See* DIFFERENTIAL EQUATION.

Quantum mechanics. In quantum theory the things one observes from experiments are eigenvalues (or the spectrum, if there are not discrete eigenvalues) of certain self-adjoint operators on Hilbert spaces of wave functions. Conservation laws come from self-adjoint operators whose associated one-parameter groups of unitary operators commute with the hamiltonian. In this way a system of conservation laws leads to a group of symmetries, namely, a group of unitary operators commuting with the hamiltonian. Frequently this group is a Lie group. Analysis of the representation of the Lie group by the unitary operators gives information about the physical system. *See* ELEMENTARY PARTICLE; LORENTZ TRANSFORMATIONS; NONRELATIVISTIC QUANTUM THEORY; RELATIVISTIC QUANTUM THEORY; SYMMETRY LAWS (PHYSICS).

Anthony W. Knapp

Bibliography. M. F. Atiyah et al., *Representation Theory of Lie Groups*, 1980; T. Kawazoe, T. Oshima, and S. Sano, *Representation Theory of Lie Groups and Lie Algebras*, 1992; P. J. Oliver, *Applications of Lie Groups to Differential Equations*, 2d ed., 2000; A. L. Onishchik and E. B. Vinberg, *Lie Groups and Algebraic Groups*, 1990; D. H. Sattinger and O. L. Weaver, *Lie Groups and Lie Algebras with Applications to Physics, Geometry, and Mechanics*, 1993; V. S. Varadarajan, *Lie Groups, Lie Algebras, and Their Representations*, 1974, reprint 1988; N. T. Varopouos, L. Saloff-Coste, and T. Coulhon, *Analysis on Lie Groups*, 1993.

Life zones

Large portions of the Earth's land area which have generally uniform climate and soil and, consequently, a biota showing a high degree of uniformity in species composition and adaptations to environment. Related terms are vegetational formation and biome.

Merriam's zones. Life zones were proposed by A. Humboldt, A. P. DeCandolle, and others who emphasized plants. Around 1900 C. Hart Merriam, then chief of the U.S. Biological Survey, related life zones, as observed in the field, with broad climatic belts across the North American continent designed mainly to order the habitats of America's important animal groups. The first-order differences between the zones, as reflected by their characteristic plants and animals, were related to temperature; moisture and other variables were considered secondary.

Each life zone correlated reasonably well with major crop regions and to some extent with general vegetation types (see **table**). Although later studies led to the development of other, more realistic or detailed systems, Merriam's work provided an important initial stimulus to bioclimatologic work in North America. *See* DENDROLOGY; VEGETATION AND ECOSYSTEM MAPPING.

Work on San Francisco Mountain in Arizona impressed Merriam with the importance of temperature as a cause of biotic zonation in mountains. Isotherms based on sums of effective temperatures correlated with observed distributions of certain animals and plants led to Merriam's first law, that animals and plants are restricted in northward distribution by the sum of the positive temperatures (above 109.4°F or 43°C) during the season of growth and reproduction. The mean temperature for the six hottest weeks of the summer formed the basis for the second law, that plants and animals are restricted in southward distribution by the mean temperature of a brief period covering the hottest part of the year. Merriam's system emphasizes the similarity in biota between arctic and alpine areas and between boreal and montane regions. It was already recognized that latitudinal climatic zones have parallels in altitudinal belts on mountain slopes and that there is some biotic similarity between such areas of similar temperature regime.

In northern North America Merriam's life zones, the Arctic-Alpine, the Hudsonian, and the Canadian,

Characteristics of Merriam's life zones

Zone name	Example	Vegetation	Typical and important plants	Typical and important animals	Typical and important crops
Arctic-alpine	Northern Alaska, Baffin Island	Tundra	Dwarf willow, lichens, heathers	Arctic fox, musk-ox, ptarmigan	None
Hudsonian	Labrador, southern Alaska	Taiga, coniferous forest	Spruce, lichens	Moose, woodland caribou, mountain goat	None
Canadian	Northern Maine, northern Michigan	Coniferous forest	Spruce, fir, aspen, red and jack pine	Lynx, porcupine, Canada jay	Blueberries
Western division					
Humid transition	Northern California coast	Mixed coniferous forest	Redwood, sugar pine, maples	Blacktail deer, Townsend chipmunk, Oregon ruffed grouse	Wheat, oats, apples, pears, Irish potatoes
Arid transition	North Dakota	Conifer, woodland sagebrush	Douglas fir, lodgepole, yellow pine, sage	Mule deer, whitetail, jackrabbit, Columbia ground squirrel	Wheat, oats, corn
Upper Sonoran	Nebraska, southern Idaho	Piñon, savanna, prairie	Junipers, piñons, grama grass, bluestem	Prairie dog, blacktail jackrabbit, sage hen	Wheat, corn, alfalfa, sweet potatoes
Lower Sonoran	Southern Arizona	Desert	Cactus, agave, creosote bush, mesquite	Desert fox, four-toed kangaroo rats, roadrunner	Dates, figs, almonds
Eastern division					
Alleghenian	New England	Mixed conifer and hardwoods	Hemlock, white pine, paper birch	New England cottontail rabbit, wood thrush, bobwhite	Wheat, oats, corn, apples, Irish potatoes
Carolinian	Delaware, Indiana	Deciduous forest	Oaks, hickory, tulip tree, redbud	Opossum, fox, squirrel, cardinal	Corn, grapes, cherries, tobacco, sweet potatoes
Austroriparian	Carolina piedmont, Mississippi	Long-needle conifer forest	Loblolly, slash pine, live oak	Rice rat, woodrat, mocking bird	Tobacco, cotton, peaches, corn
Tropical	Southern Florida	Broadleaf evergreen forest	Palms, mangrove	Armadillo, alligator, roseate spoonbill	Citrus fruit, avocado, banana

are entirely transcontinental (**Fig. 1** and table). Because of climatic and faunistic differences, the eastern and western parts of most life zones in the United States (the Transition, upper Austral, and lower Austral) had to be recognized separately. The western zones had to be further subdivided into humid coastal subzones and arid inland subzones. The Tropical life zone includes the extreme southern edge of the United States, the Mexican lowlands, and Central America.

Although once widely accepted, Merriam's life zones are little used today because they include too much biotic variation and oversimplify the situation. However, much of the terminology he proposed persists, especially in North American zoogeographic literature.

Dice's zones. Another approach to life zones in North America is the biotic province concept of L. R. Dice. Each biotic province covers a large and continuous geographic area and is characterized by the occurrence of at least one important ecological association distinct from those of adjacent provinces. Each biotic province is subdivided into biotic districts which are also continuous, but smaller, areas distinguished by ecological differences of lesser extent or importance than those delimiting provinces. Life belts, or vertical subdivisions, also occur within biotic provinces. These are not necessarily continuous but often recur on widely separated mountains within a province where ecological conditions are appropriate.

Boundaries between biotic provinces were largely subjective, supposedly drawn where the dominant associations of the provinces covered approximately equal areas. In practice, however, too few association data including both plants and animals were available, and vegetation generally offered the most satisfactory guide to boundaries. The Dice system recognizes 29 biotic provinces in North America (**Fig. 2**). The two northernmost

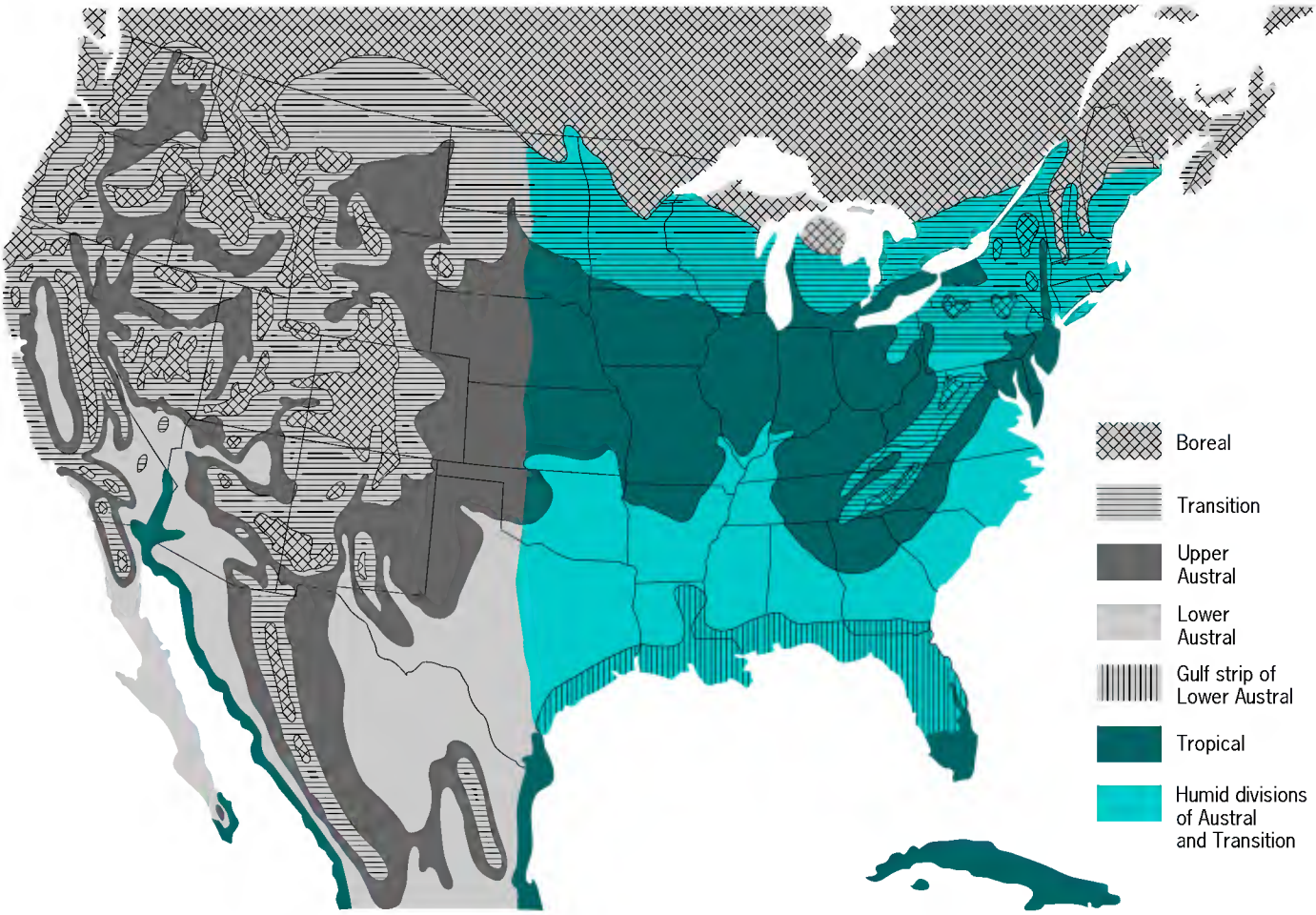

Fig. 1. Life zones of the United States. (*After C. H. Merriam, Life Zones and Crop Zones of the United States, USDA Bull. 10, 1898*)

Fig. 2. Biotic provinces of North America (Veracruzian not shown). 1 mi = 0.6 km. (*After L. R. Dice, The Biotic Provinces of North America, University of Michigan Press, 1943*)

(Eskimoan and Hudsonian) are transcontinental, reflecting broad, high-latitude climatic belts. Those in the eastern United States (Canadian, Carolinian, and Austroriparian) reflect both latitudinal temperature change and physiography, whereas province boundaries in the remainder of North America are strongly influenced by physiography.

Holdridge's zones. A life-zone system with boundaries more explicitly defined is that of L. R. Holdridge. These life zones are defined through the effects of the three weighted climatic indexes: mean annual heat, precipitation, and atmospheric moisture. Each axis of the triangle (**Fig. 3**) represents one climatic component, and the three sets of lines parallel to the axes define the life-zone framework. Climatic values which the lines represent progress geometrically. Mean annual biotemperature is computed by summing the monthly mean temperatures from 32°F or 0°C (setting negative winter months equal to 32°F or 0°C when they occur) to about 86°F or 30°C and dividing by 12. Values increase from top to bottom in Fig. 3 and, by extension to either margin, establish equivalent latitudinal regions and altitudinal belts. Precipitation is computed as mean annual precipitation in millimeters and increases from the apex toward the lower right margin of the diagram. The

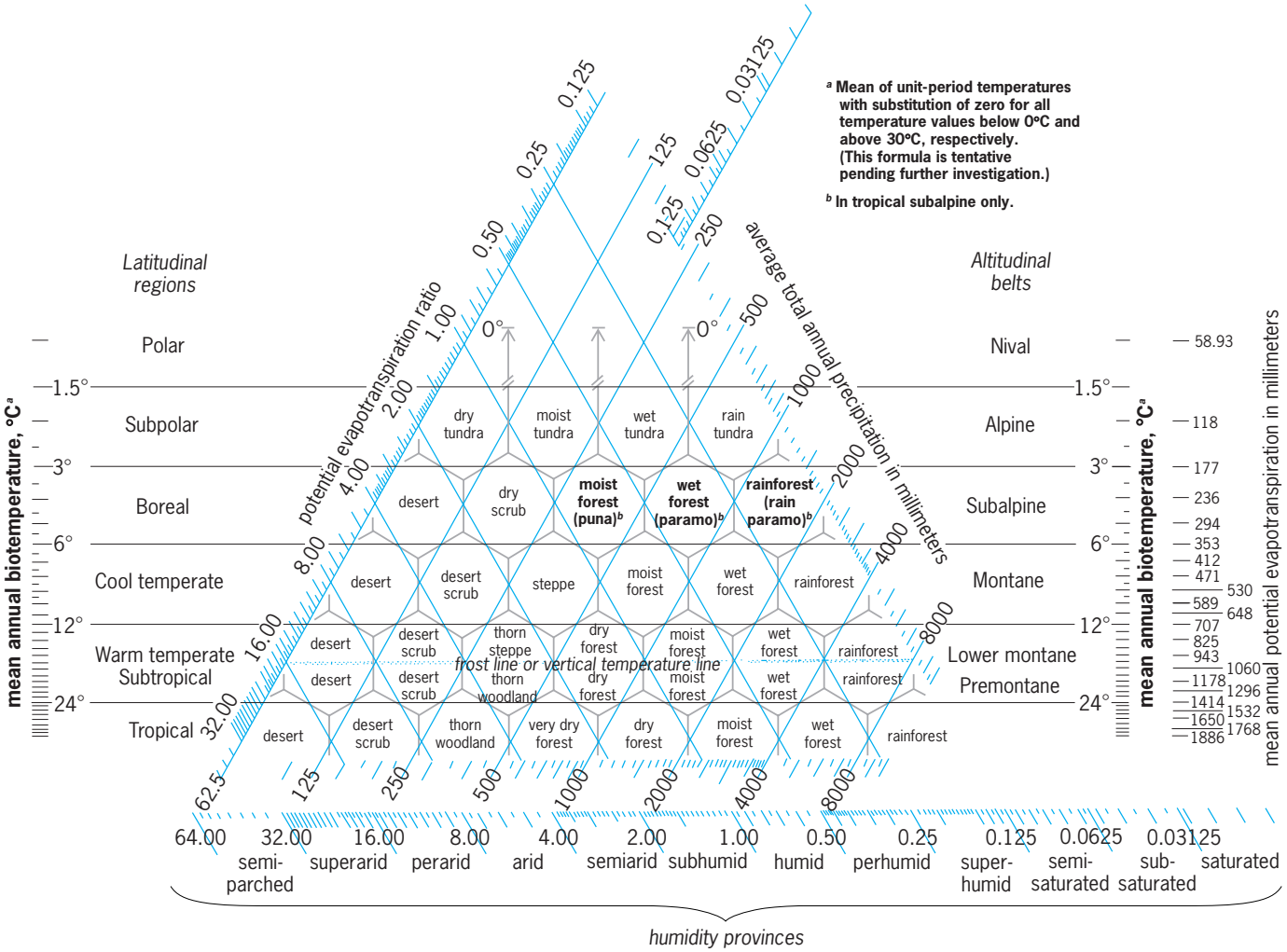

Fig. 3. World life zones. °F = (°C × 1.8) + 32. (*After L. R. Holdridge, Life Zone Ecology, rev. ed., Tropical Science Center, 1967*)

humidity factor is calculated by dividing the value of the mean annual potential evapotranspiration by the mean annual precipitation in millimeters. Values of this ratio increase toward the lower left side of the diagram, signifying a decrease in effective humidity.

Associations of vegetation, animals, climate, physiography, geology, and soils are interrelated in unique combinations, mostly having distinct physiognomy, in Holdridge's climatic grid. Local investigation should show how well the framework represents the most common life zones of Earth, both north and south of the Equator. The system has been applied with considerable success to the mapping of life zones, primarily in the American tropics. Names of Holdridge's life zones are based on the dominant vegetation of each climatic type; animal life of the zone is supposedly distinctive. *See* ALTITUDINAL VEGETATION ZONES; BIOME; ZOOGEOGRAPHY.

Arthur W. Cooper

Bibliography. W. H. Lewis, *Ecology Field Glossary: A Naturalist's Vocabulary*, 1977; C. H. Merriam, *Life Zones and Crop Zones of the United States*, USDA Bull. 10, 1898; H. Walter, *Vegetation of the Earth and Ecological System of the Geobiosphere*, 3d rev. ed., 1994; J. E. Weaver and F. E. Clements, *Plant Ecology*, 2d ed., 1938.

Ligament

A strong, flexible connective tissue band usually found between two bony prominences. Most ligaments are composed of dense fibrous tissue formed by parallel bundles of collagen fibers. They have a shining white appearance and are pliable, strong, and noncompliant. These ligaments are closely associated with all articulations and are largely responsible for the exact movement of the bones. Other ligaments may span larger gaps between bones and may serve as important controlling elements in complex bone-linkage systems.

A second kind of ligament, composed either partly or almost entirely of yellow elastic fibers, is extensible or compliant, thereby allowing the connected bones to move apart. The large ligamenta nuchae between the neural spines of cervical vertebrae and the skull in large ungulates are of this type, and permit lowering and raising the head with less muscular effort. *See* BIOMECHANICS; COLLAGEN; CONNECTIVE TISSUE; JOINT (ANATOMY).

Walter Bock

Ligand

A molecule with an affinity to bind to a second atom or molecule. This affinity can be described in terms of noncovalent interactions, such as the type of binding that occurs in enzymes that are specific for certain substrates; or of a mode of binding where an atom or groups of atoms are covalently bound to a central atom, as in the case of coordination complexes and organometallic compounds. Ligands of the latter type can be further distinguished by the nature of the orbitals used in bond formation. *See* ENZYME.

When a protein binds to another molecule, that molecule may be referred to as a ligand. The site where the ligand is bound is known as the binding or active site of the protein. In order for a molecule to be classified as a ligand for a protein, several weak interactions such as hydrophobic, van der Waals, and hydrogen bonds must take place simultaneously. Therefore, the binding of a ligand by a protein is generally quite specific. For example, the active site of copper-zinc superoxide dismutase (CuZnSOD) has been engineered so that only very small ligands such as superoxide, water, and other small anions are accepted.

Coordination complexes and organometallic compounds have ligands that are a class of ions or neutral molecules that donate an electron pair to a metal atom or ion. Molecules or ions that function as ligands behave as Lewis bases, while a central metal acts as a Lewis acid. The entire molecule, which includes the metal or metal atoms and the attached ligands, is generally known as a coordination complex. An example of a coordination complex having six ligands is hexaammineruthenium (III) (**Fig. 1**). The single pair

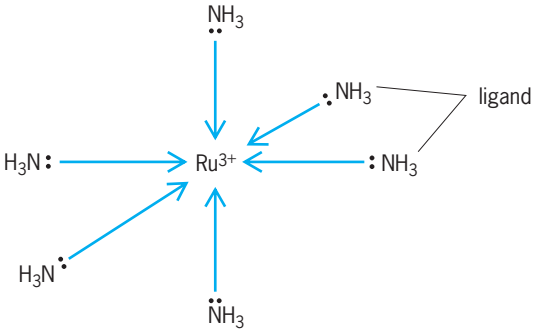

Fig. 1. The complex hexaammineruthenium(III); the six ammine groups (:NH_3) are the ligands.

of electrons present on the ammine (or ammonia) group is donated to a vacant orbital of the ruthenium tripositive ion (Ru^{3+} or Ru^{III}). *See* ACID AND BASE; COORDINATION COMPLEXES; ORGANOMETALLIC COMPOUND.

Mono-, bi-, and multidentate ligands. A monodentate ligand possesses one "tooth" or one pair of electrons available for binding to a metal center. In other words, there is only one site on the ligand where coordination to a metal may occur. Examples of monodentate ligands include the halides, [fluoride (F^-), chloride (Cl^-), bromide (Br^-), and iodide (I^-)] and neutral molecules such as water H_2O and ammonia

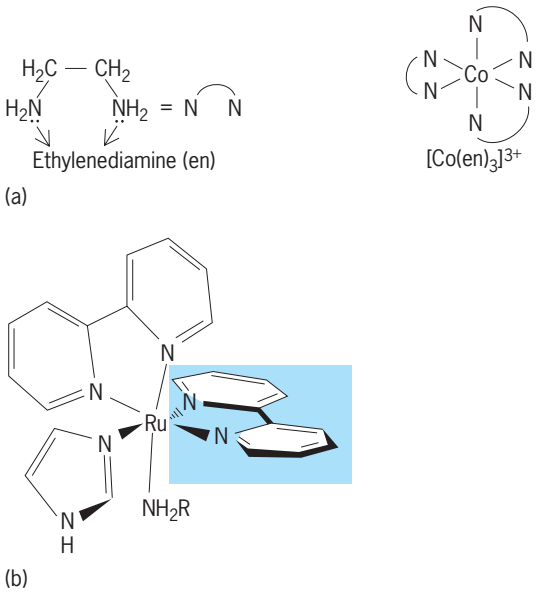

Fig. 2. Bidentate ligands. (*a*) Ethylenediamine (en) and the cobalt coordination complex $[Co(en)_3]^{3+}$ with three en ligands. (*b*) The bidentate aromatic imine 2,2′-bipyridine (bpy) ligand (highlighted) in the complex $Ru(bpy)_2imNH_2R$.

NH_3. Each of these ligands donates a single pair of electrons to the central metal atom.

Bidentate ligands possess two atoms (each with at least one pair of electrons), termed donor atoms, that may coordinate to a metal atom or ion. A well-known example of a bidentate ligand is ethylenediamine (en), and an example of a complex with three en ligands is $[Co(en)_3]^{3+}$ (**Fig. 2***a*). Another example of a bidentate ligand is 2,2′-bipyridine (bpy), an aromatic imine. The single pair of electrons on each of the bipyridine nitrogen (N) atoms is capable of donating to the metal center, as in the coordination complex with two bidentate bpy ligands, the ruthenium complex $Ru(bpy)_2imNH_2R$ (Fig. 2*b*), where R represents an organic group and im = imidazole.

Multidentate ligands possess three or more donor atoms; they are called chelating agents or chelating ligands. An example is diethylenetriamine (**1**) which has three nitrogen donor atoms. Triethylenetetramine derivatives are multidentate ligands that form a species known as tripod complexes (**2**). Ethylenediamine-tetraacetic acid (EDTA, **3**),

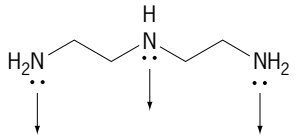

(1)

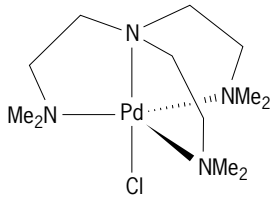

(2)

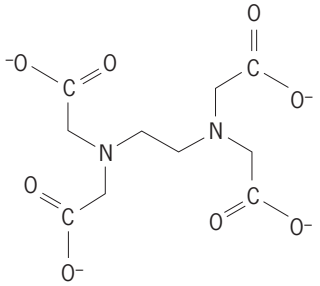

(3)

which has two nitrogen and four oxygen donor atoms, is another example of a multidentate ligand. Chelating ligands like EDTA (six donor atoms per molecule) and diethylenetriamine-pentaacetic acid (DTPA; eight donor atoms per molecule) bind metal ions very tightly because the formation constants of these complexes are very large. Therefore, multidentate ligands are commonly used as scavengers for metal atoms and ions. A Schiff base complex possesses monodentate and multidentate ligands such as the complex in structure (**4**), where the metal is

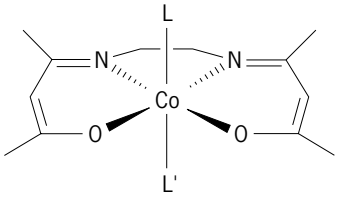

(4)

cobalt (Co); the complex has two nitrogen and two oxygen donor atoms and two monodentate ligands, L and L′. *See* CHELATION.

Bridging ligands. A ligand that is bound to more than one metal center is known as a bridging ligand. These ligands may form two, and in some instances three, bonds to different metal atoms. An example of a unidentate bridging ligand is structure (**5**). Biden-

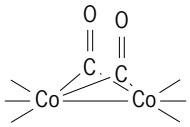

(5)

tate ligands such as the acetate ion (CH_3COO) may behave as both a chelating ligand, where both donor atoms are bound to the same metal, and a bidentate bridging ligand, where binding takes place at more than one metal center as in structure (**6**), where

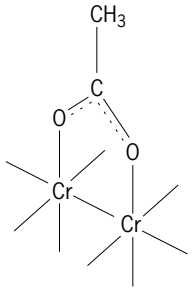

(6)

Cr = chromium and the broken-line bonds indicate that the structure is a resonance hybrid, a configuration in which an electron is delocalized between each carbon-oxygen bond. *See* DELOCALIZATION.

Backbonding ligands. Some ligands can form two or more bonds with a metal atom. These ligands have the capacity to serve as donors and acceptors of electron density that compose the ligand-metal bond. A sigma (σ) bond is formed by donation of a single pair of electrons from the ligand to the metal (M), as in structure (**7**), where the backbonding ligand is car-

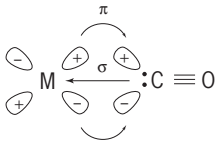

(7)

bon monoxide (CO). The metal atom, in an effort to reduce the electron density, behaves as a π donor by transferring electrons back to the ligand to form a π bond. The backbonding ligand in this case acts as a π acid; that is, it is able to accept electron density into an empty π orbital. As a result of this type of interaction, complexes are produced with metals that have low oxidation states.

The backbonding ligand carbon monoxide (CO) forms complexes known as metal carbonyls. Finally, a prototypic example of metal binding to a π-donor ligand is the complex ferrocene (**8**), with a metal

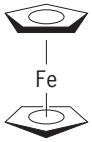

(8)

[in this case iron (Fe)], binding to a π-donor cyclopentadienyl ligand. Each of the cyclopentadienyl rings in this complex donate six electrons to the iron atom. *See* CHEMICAL BONDING; COORDINATION CHEMISTRY; METAL CARBONYL. Thomas J. Meade

Ligand field theory

An essentially ionic approach to chemical bonding which is often used with coordination compounds. These compounds consist of a central transition-metal ion that is surrounded by a regular array of coordinated atoms or ligands. Accordingly, the ligands are assumed to be sources of negative charge which perturb the energy levels of the central metal ion. In this respect the ligands subject the metal ion to an electric field which is analogous to the electric or crystal field produced by the regular distribution of nearest neighbors within an ionic crystalline lattice. For example, the crystal field produced by the Cl ion ligand in octahedral $TiCl_6^{3-}$ is considered to be similar to that produced by the octahedral array of the six Cl ions about each Na ion in NaCl. The Na ion with its rare-gas configuration has an electronic charge distribution which is spherically symmetric

both within and without the crystal field. The paramagnetic Ti(III) ion, which possesses one $3d$ electron (d^1), has a spherically symmetric charge distribution only in the absence of the crystal field produced by the ligands. The presence of the ligands destroys the spherical symmetry and produces a more complex set of energy levels within the central metal ion. The crystal field theory allows the energy levels to be calculated and related to experimental observation. *See* COORDINATION CHEMISTRY.

To illustrate the results of a typical crystal field calculation, assume that the single d electron in the Ti(III) ion will experience a coulombic repulsion with each of the six nearest neighbor Cl ions which are taken as point negative charges. This model for coulombic repulsion may be described mathematically as the summation $eq\Sigma r^{-1}$, where e and q are the electronic and ligand charges, respectively, and r is the distance between the electron and the ligand. The summation extends over all the ligands. A detailed quantum mechanical calculation can then be made. Fortunately, it is possible to arrive at identical results by a very qualitative procedure. This method considers the spatial orientation of d orbitals with relation to the ligands when both are viewed within the same coordinate system (**Fig. 1**).

The d_{z^2} orbital is oriented along the z axis with a ring in the xy plane, while the $d_{x^2-y^2}$ orbital is oriented only along the x and y axes. Although it is not visually obvious, both are equivalent. The d_{xy}, d_{xz}, and d_{yz} orbitals are oriented between the x, y, and

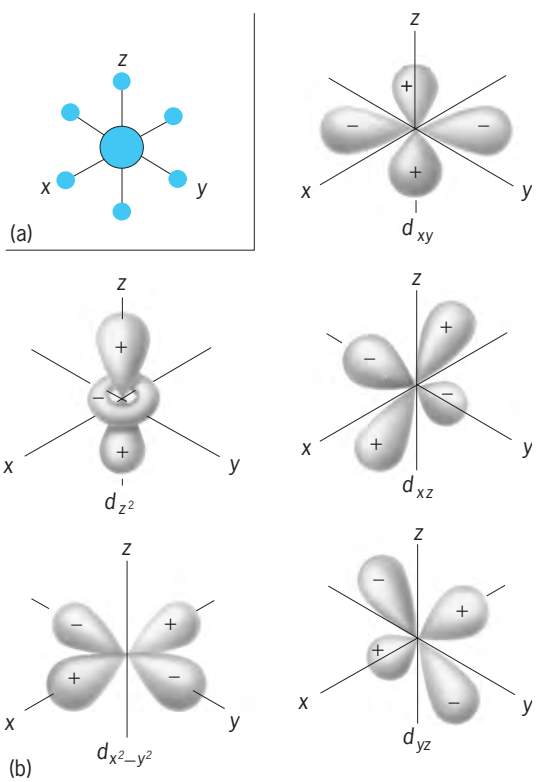

Fig. 1. Spatial orientation of *d* orbitals with relation to ligands in $TiCl_6^{3-}$. (*a*) The coordinate system for the octahedral $TiCl_6^{3-}$ ion. (*b*) The five *d* orbitals. The d_{z^2} orbital is symmetric with respect to the *z* axis.

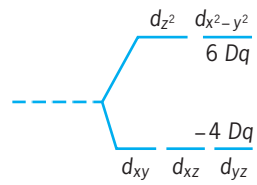

Fig. 2. The splitting of the five *d* orbitals because of an octahedral crystal field.

z axes and are geometrically equivalent. The ligands are placed along the coordinate x, y, and z axes. An electron in any of the five orbitals will experience a coulombic repulsion due to the crystal field of the ligands. However, since the d_{z^2} and $d_{x^2-y^2}$ orbitals are directed toward the ligands, an electron in these orbitals will undergo far more repulsion than one in any of the three other orbitals. The imposition of six point negative charges, then, will not allow the five d orbitals to be equally energetic (degenerate) as they are in the bare Ti(III) ion, but causes three of these orbitals to be more stable than the remaining two. The difference in energy is termed $10Dq$. Thus, in the ground state of $TiCl_6^{3-}$, the single d electron will be found in the lower set of orbitals, whose energy is $-4Dq$. The crystal field stabilization energy (CFSE) is then said to be $4Dq$ (**Fig. 2**).

Coordination compounds are often colored. Crystal field theory suggests that in the case of $TiCl_6^{3-}$ its color is a result of an electronic excitation of the electron from the threefold set of orbitals into the twofold set. In the spectrum of $TiCl_6^{3-}$, an absorption band maximum is found at 13,000 cm^{-1}, which is then the energy associated with the transition, or $10Dq$. *See* MOLECULAR STRUCTURE AND SPECTRA.

Several electrons. When more than one d electron is present, the spectroscopic evaluation of Dq is not as simple, but can generally be accomplished. Nevertheless, the CFSE may be easily and formally obtained for these cases, (see **table**), since this energy is simply the number of electrons occupying the orbital multiplied by the orbital energy. Thus, the CFSE amounts to $8Dq$ in an octahedral d^2 complex, and $(6 \times 4) - (2 \times 6) = 12Dq$ in a similar d^8 complex. With d^4, d^5, d^6, and d^7, however, two possibilities exist. If the crystal field is sufficiently strong to overcome the repulsion energy which will result from pairing the electrons in the lower set of orbitals, the

Crystal field stabilization energy

	Octahedral CFSE (Dq)	
d^n	Weak	Strong
1	4	4
2	8	8
3	12	12
4	6	16
5	0	20
6	4	24
7	8	18
8	12	12
9	6	6

maximum number of electrons will be found in the lower set. This situation is termed the strong-field case. In the weak-field case, the electron-pairing (repulsive) energy is greater than the crystal field (attractive) energy and the maximum number of unpaired electrons will result. Thus, in a strong crystal field a d^5 ion should have only one unpaired electron and a CFSE of $20Dq$. This is found in the $Mn(CN)_6^{4-}$ ion, which has been shown experimentally to possess only one unpaired electron. In most complexes of Mn(II), such as $Mn(H_2O)_6^{2+}$, the crystal field is weak so that five unpaired electrons result and the CFSE is zero.

Simple considerations such as these have enabled inorganic chemists to understand why certain coordination compounds containing a given metal ion may exhibit full paramagnetism, while others containing the metal in exactly the same formal oxidation state may show either a much weaker paramagnetism or none at all. A striking example of this is provided by paramagnetic CoF_6^{3-}, which possesses six unpaired d electrons and the diamagnetic $Co(NH_3)_6^{3+}$, in which none of the d electrons are unpaired. Both contain Co(III).

Tetrahedral array. Tetrahedral complexes may be treated in a similar fashion. The results indicate that the only difference, when compared to the octahedral complex, is in the orbital splitting pattern and the relative magnitude of Dq. The tetrahedral array of ligands causes an inversion of the pattern such that the d_{z^2} and $d_{x^2-y^2}$ orbitals lie lowest. The difference in energy between the two sets of orbitals is now found to be $^4/_9$ of that in an octahedral complex, or Dq (tetrahedral) $= (^4/_9)Dq$ (octahedral). This particularly simple result has led to the understanding of many stereochemical phenomena. An important early example was found in the cation distribution of normal and inverse spinels. The former are double oxides having the general formula $M(II)[M'(III)]_2O_4$ in which the oxygens lie in a close-packed system. The divalent metal ions, M(II), occupy one-eighth of the tetrahedral holes. In the inverse spinel $M'(III)[M(II)M'(III)]O_4$ the divalent metal ions have changed places with one-half of the trivalent ions.

Experimentally, it has been found that $MnCr_2O_4$ [containing Mn(II) with five unpaired electrons] and $NiFe_2O_4$ [containing Fe(III) with five unpaired electrons] have the normal and inverse structures, respectively. A simple application of crystal field theory results in exact agreement with experiment. In $MnCr_2O_4$ the ion at each octahedral site has a stability of $12Dq$, while the tetrahedral site has no CFSE. The total CFSE is then $24Dq$. If the structure were inverse, the total CFSE would be only $^4/_9 \times 12 + 12 = 17.3Dq$. With $NiFe_2O_4$ the CFSE of the normal spinel is only $5.3Dq$, but in the inverse structure the CFSE increases to $12Dq$. In general, agreement between predicted cation distribution and that experimentally observed is good.

The application of these methods to conventional coordination compounds also meets with a fair amount of success. For example, with only one or two exceptions octahedral coordination of Cr(III) and diamagnetic Co(III) prevails in all compounds, as one would predict from crystal field theory. Important exceptions to these rules do exist and point out that other factors, such as ligand-ligand repulsions, sometimes outweigh the CFSE. Octahedral coordination of Ni(II) is favored over the tetrahedral arrangement insofar as CFSE is concerned, yet the tetrahedral $NiCl_4^{2-}$, $NiBr_4^{2-}$, and NiI_4^{2-} ions are well known.

Anomalous effects. Heats of hydration, lattice energies, crystal radii, and oxidation potentials of transition-metal ions and their complexes contain apparent anomalies which are best explained in terms of an effect due to the crystal field. As an example, when the heats of hydration of the ions of the first transition series are plotted with respect to atomic number, a peculiar double-humped curve is obtained (**Fig. 3**). The results for those ions not possessing any CFSE, that is, Ca, Mn, and Zn, lie nearly on a straight line. In the absence of any other effect, it might be expected that with the successive addition of each nuclear charge, a monotonic increase in the heat of hydration would occur. Instead, it was found that the heat of hydration of those metal ions possessing CFSE is far more exothermic than would be predicted by arguments pertaining only to the successive increase in atomic number. In fact, the excess heat is best accounted for in terms of the CFSE. The total heat of hydration may be written as $\Delta H = \Delta H^\circ + \text{CFSE}$, where ΔH° is the heat of hydration that would be expected for a hypothetical metal ion which ignored the crystal field. The change in ΔH° with atomic number would then be expected to be monotonic. When the observed heats of hydration are corrected for the CFSE obtained from spectroscopic data for the resulting $M(H_2O)_6^{2+}$ complexes, the expected monotonic

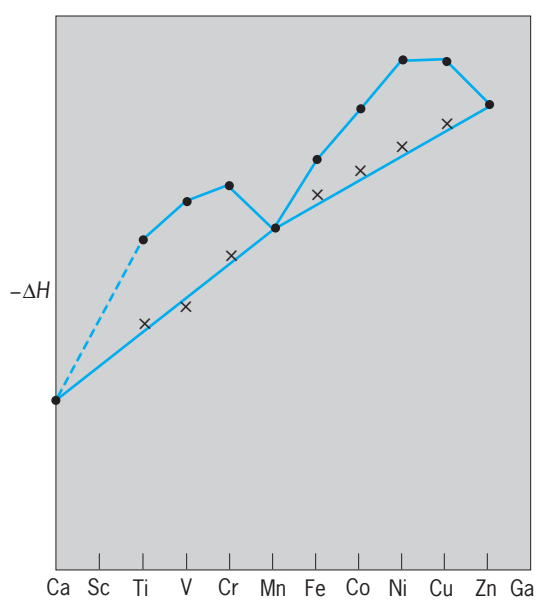

Fig. 3. Variation of heat of hydration $-\Delta H$ of divalent ions of first transition series. Experimental values are given by the dots, and the corresponding values, after correction for CFSE, are given by the x's.

increase is observed. Similar double-humped curves occur with the lattice energies and crystal radii and can be explained by including the effects due to the crystal field.

Stereochemical anomalies can also often be explained through the judicious use of simple arguments. A particularly important example is found in complexes of Cu(II). X-ray crystallography has established that most "octahedral" complexes containing that ion are in fact elongated along one axis. According to a theorem due to H. A. Jahn and E. Teller, this behavior is not unexpected. The theorem states that a system possessing a degenerate ground state will distort in some unspecified manner to remove the degeneracy.

The degeneracy in a regular octahedral complex of Cu(II) is easily illustrated by the possibility of writing the electronic configuration in two distinct, but equally energetic, ways: $(d_{xy})^2(d_{xy})^2(d_{yz})^2(d_{z^2})^2 \times (d_{x^2-y^2})^1$ or $(d_{xy})^2(d_{xz})^2(d_{yz})^2(d_{x^2-y^2})^2(d_{z^2})^1$. For each, the CFSE is $6Dq$. In addition to the twofold degeneracy, there are two separate means by which the degeneracy may be removed. If the ligands along the z axis of the octahedron move away from the metal ion while those in the xy plane move toward the center of the octahedron, then according to simple crystal field arguments, this movement will result in stabilizing the d_{z^2} orbital with respect to the $d_{x^2-y^2}$ orbital. Alternatively, the completely opposite movement will stabilize the $d_{x^2-y^2}$ orbital with respect to the d_{z^2} orbital (**Fig. 4**). In either case, an additional increment of energy is added to the CFSE of the Cu(II) complex: $6Dq + \Delta E$. Thus, the driving force for the distortion is the additional stabilization energy ΔE. Crystal field theory is unable to predict which mode of distortion will occur, but it clearly predicts that a distortion should occur. From structure determinations through the use of x-rays, it is found that elongation along one axis is by far the most predominant mode.

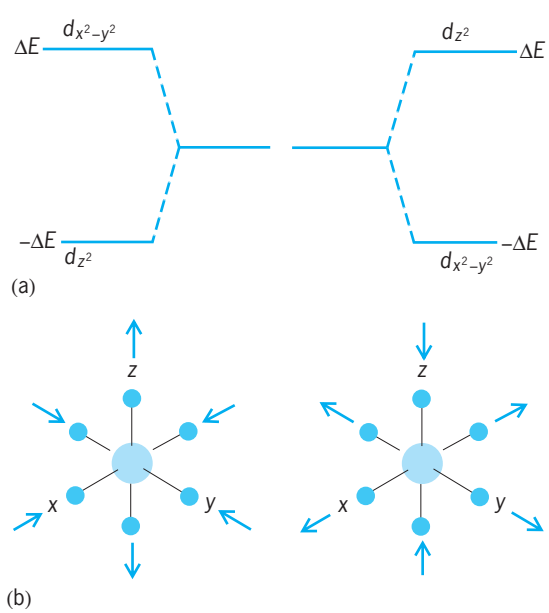

Fig. 4. The splitting of the $d_{x^2-y^2}$ and d_{x^2} orbitals by an axial distortion of the octahedron. (*a*) Alternate possiblities for writing the electronic configuration. (*b*) The two modes of distortion.

Complete theory. The many successes of crystal field theory in the interpretation of the natural phenomena associated with transition-metal compounds should not lead one to the conclusion that the bonding within these compounds can be truthfully represented by a strictly ionic model. Crystal field theory is essentially a very specialized form of the more complete molecular orbital theory. The need for the more complete theory becomes obvious when an attempt is made to rationalize the absolute value of Dq, the observation of ligand-to-metal electronic transitions, and certain observations in both nuclear magnetic resonance and electron spin resonance experiments, which can only be interpreted in terms of some covalent bonding. *See* MOLECULAR ORBITAL THEORY.

R. A. D. Wentworth

Bibliography. R. L. DeKock and H. B. Gray, *Chemical Structure and Bonding*, 2d ed., 1980, reprint 1989; M. Gerloch, *Orbitals, Terms, and States*, 1986; Y. Jean and F. Volatron, *Introduction to Molecular Orbitals*, 1993; A. Rauk, *Orbital Interaction Theory of Organic Chemistry*, 2d ed., 2000.

Light

The term light, as commonly used, refers to the kind of radiant electromagnetic energy that is associated with vision. In a broader sense, light includes the entire range of radiation known as the electromagnetic spectrum. The branch of science dealing with light, its origin and propagation, its effects, and other phenomena associated with it is called optics. Spectroscopy is the branch of optics that pertains to the production and investigation of spectra. *See* OPTICS; SPECTROSCOPY.

Any acceptable theory of the nature of light must stem from observations regarding its behavior under various circumstances. Therefore, this article begins with a brief account of the principal facts known about visible light, including the relation of visible light to the electromagnetic spectrum as a whole, and then describes the apparent dual nature of light. The remainder of the article discusses various experimental and theoretical considerations pertinent to the study of this problem.

Principal Effects

The electromagnetic spectrum is a broad band of radiant energy that extends over a range of wavelengths running from trillionths of inches to hundreds of miles; wavelengths of visible light are measured in hundreds of thousandths of an inch. Arranged in order of increasing wavelength, the radiation making up the electromagnetic spectrum is termed gamma rays, x-rays, ultraviolet rays, visible light, infrared waves, microwaves, radio waves, and very long electromagnetic waves. *See* ELECTROMAGNETIC RADIATION.

Speed of light. The fact that light travels at a finite speed or velocity is well established. In round numbers, the speed of light in vacuum or air may be said to be 186,000 mi/s or 300,000 km/s.

Measurements of the speed of light, c, which had attracted physicists for 308 years, came to an end in 1983 when the new definition of the meter fixed the value of the speed of light. The speed of light is one of the most interesting and important of the fundamental physical constants. It is used to convert light travel times to distance, as in laser and radio measurements of the distance to the Moon and planets. It relates mass m to energy E in Einstein's equation, $E = mc^2$. To fix the value of c in the new definition, highly precise values of c were obtained by extending absolute frequency measurements into a region of the electromagnetic spectrum where wavelengths can be most accurately measured. These advances were facilitated by the use of stabilized lasers and high-speed tungsten-nickel diodes which were used to measure the lasers' frequencies. The measurements of the speed of light and of the frequency of lasers yielded a value of the speed of light limited only by the standard of length which was then in use. This permitted a redefinition of the meter in which the value of the speed of light assumed an exact value, 299,792,458 m/s. The meter is defined as the length of the path traveled by light in vacuum during a time interval of 1/299 792 458 of a second. *See* ASTRONOMICAL UNIT; FUNDAMENTAL CONSTANTS; LASER.

Prior to the observations of O. Roemer in 1675, the speed of light was thought to be infinite. Roemer noted a variation of the orbiting periods of the moons of Jupiter that depended on the annual variation in the distance between Earth and Jupiter. He correctly ascribed the variation to the time it takes light to travel the varying distance between the two planets. The accuracy of Roemer's value of c was limited by a 30% error in the knowledge of the size of the planetary orbits at that time.

The first terrestrial measurement of c was performed by H. L. Fizeau in 1849. His measurement of the time it took light to travel to a distant mirror and return resulted in a value accurate to 15%.

J. C. Maxwell's theory of electromagnetic radiation showed that both light and radio waves were electromagnetic and hence traveled at the same speed in vacuum. This discovery soon led to another method of measuring c: it was the product of the frequency and wavelength of an electromagnetic wave. In 1891 R. Blondlot first used this method to determine a value of c by measuring both the wavelength and the frequency of an electromagnetic, radio-frequency wave. His measurement demonstrated that c was the same for radio and light waves. It is this method which now exhibits the greatest accuracy, and it is used in the most accurate measurements of c using a laser's frequency and wavelength.

In 1958 K. D. Froome reported the speed of light c to be 299,792,500 m/s, with an uncertainty of plus or minus 100 m/s. He measured both the frequency and the wavelength of millimeter waves from klystron oscillators to obtain this result. The major uncertainty lay in measuring the wavelength of the radiation. Since short wavelengths can be measured much more accurately than long wavelengths, a shorter-wavelength source was needed to improve the accuracy further; the stabilized laser soon provided such a source. However, a means of measuring its incredibly high frequency was needed. This problem, too, was soon overcome with the discovery of the tungsten-nickel point-contact diode.

Stabilized lasers. Before the advent of the laser, the most spectrally pure light came from the emission of radiation by atoms in electric discharges. The spectral purity of such radiation was about 1 part per million. Lasers, in contrast, have exhibited short-period spectral purities some 10^8 times greater than this. However, the frequency of this laser radiation was free to wander over the entire emission line, and a means of stabilizing and measuring the frequency was necessary before it could be used in a measurement of c. The technique of sub-Doppler saturated absorption spectroscopy permitted the "locking" of the frequency of the radiation to very narrow spectral features so that the frequency (and, of course, the wavelength) remained fixed. Several different lasers at different wavelengths have been stabilized, and they serve as precise frequency and wavelength sources. Three of the most common are: the helium-neon laser at a wavelength of 3.39 micrometers stabilized with a saturated absorption in methane, the 10-μm carbon dioxide (CO_2) laser stabilized to a saturated fluorescence in carbon dioxide, and the common red helium-neon laser stabilized to an iodine-saturated absorption. The frequencies and wavelengths of these three lasers have been measured to yield values of c. *See* LASER SPECTROSCOPY.

Measurement of wavelength. Precise measurements of wavelengths are commonly made in Fabry-Perot interferometers, in which two wavelengths are compared by the observation of interference fringes of waves reflecting between two mirrors. A bright fringe occurs when the optical path length between the high-reflectivity mirrors is a multiple of a half-wavelength.

With the use of special Fabry-Perot interferometers, wavelength measurements of stabilized lasers were made with accuracies limited by the length standard then in use, the 605.8-nanometer orange radiation from the krypton atom. This limitation affected all speed-of-light measurements, with a resulting uncertainty of about 4 parts in 10^9. *See* INTERFEROMETRY.

Frequency measurement. The measurement of the frequency of an electromagnetic wave in the laser region is performed by a heterodyne technique in which harmonics are generated in high-speed nonlinear devices. For the most accurate measurements of c, the accuracy of the frequency measurements were 10 times more accurate than the wavelength measurements; hence, the uncertainties were dominated by the wavelength measurements. *See* FREQUENCY MEASUREMENT.

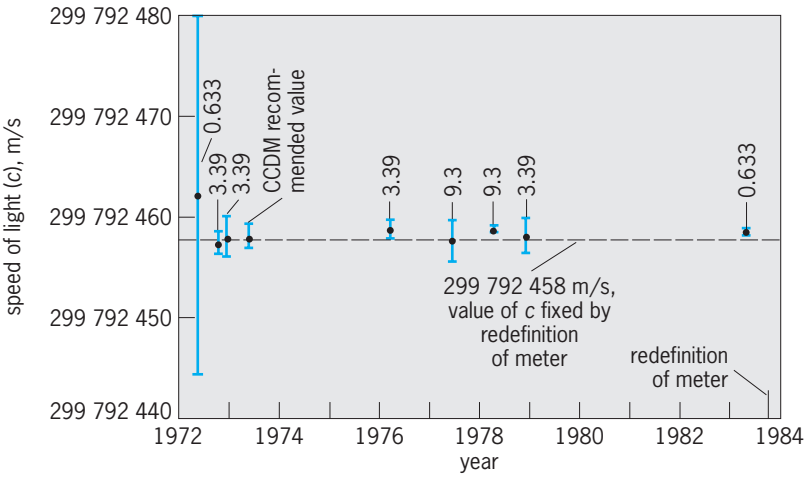

Fig. 1. Laser speed-of-light measurement made from 1972 to 1984. Separate frequency and wavelength measurements were combined to obtain the 1976, 1977, and 1983 measurements. Numbers accompanying data bars indicate the wavelengths in micrometers at which the measurements were carried out.

Results. The values of c obtained from the frequency and wavelength measurements of various lasers are shown in **Fig. 1**. Because of the stability and reproducibility of the stabilized lasers, separate frequency and wavelength measurements were sometimes combined to give independent values of c. The first 1972 measurement did not involve an absolute counting of the laser frequency and was somewhat less accurate. For the first time in history, the various values of the speed of light were in agreement. Prior to 1958, the measured values of c often varied outside the limits of error quoted by the experimenters, and even prompted some observers to think that c might be changing with time. The 1974 meeting of the Consultative Committee for the Definition of the Meter (CCDM) recommended that 299,792,458 m/s be the value of c to be used for converting wavelength to frequency and vice versa, and in all other precise applications involving c. This number was arrived at by the consideration of the first four values of c in Fig. 1. The subsequent measurements of c have shown that this was a good choice, and it is the value used in the redefinition of the meter. The 1983 measurement is the only entry for visible light, and is in excellent agreement with the others. In 1983 the General Conference on Weights and Measures (CCPM) used exactly 299,792,458 m/s in the redefinition of the meter. With this new definition, fixing the value of the speed of light, the era of speed-of-light measurements was at an end. *See* LENGTH.

Kenneth M. Evenson

Diffraction and reflection. One of the most easily observed facts about light is its tendency to travel in straight lines. Careful observation shows, however, that a light ray spreads slightly when passing the edges of an obstacle. This phenomenon is called diffraction. The reflection of light is also well known. The Moon, as well as all other satellites and planets in the solar system, are visible only by reflected light; they are not self-luminous like the Sun. Reflection of light from smooth optical surfaces occurs so that the angle of reflection equals the angle of incidence, a fact that is most readily observed with a plane mirror. When light is reflected irregularly and diffusely, the phenomenon is termed scattering. The scattering of light by gas particles in the atmosphere causes the blue color of the sky. When a change in frequency (or wavelength) of the light occurs during scattering, the scattering is referred to as the Raman effect. *See* DIFFRACTION; RAMAN EFFECT; REFLECTION OF ELECTROMAGNETIC RADIATION; SCATTERING OF ELECTROMAGNETIC RADIATION.

Refraction. The type of bending of light rays called refraction is caused by the fact that light travels at different speeds in different media—faster, for example, in air than in either glass or water. Refraction occurs when light passes from one medium to another in which it moves at a different speed. Familiar examples include the change in direction of light rays in going through a prism, and the bent appearance of a stick partially immersed in water. *See* REFRACTION OF WAVES.

Interference and polarization. In the phenomenon called interference, rays of light emerging from two parallel slits combine on a screen to produce alternating light and dark bands. This effect can be obtained quite easily in the laboratory, and is observed in the colors produced by a thin film of oil on the surface of a pool of water. *See* INTERFERENCE OF WAVES.

Polarization of light is usually shown with Polaroid disks. Such disks are quite transparent individually. When two of them are placed together, however, the degree of transparency of the combination depends upon the relative orientation of the disks. It can be varied from ready transmission of light to almost total opacity, simply by rotating one disk with respect to the other. *See* POLARIZED LIGHT.

Chemical effects. When light is absorbed by certain substances, chemical changes take place. This fact forms the basis for the science of photochemistry. Rapid progress has been made toward an understanding of photosynthesis, the process by which plants produce relatively complex substances such as sugars in the presence of sunlight. This is but one example of the all-important response of plant and animal life to light. *See* PHOTOCHEMISTRY; PHOTOSYNTHESIS.

Historical Approach

Early in the eighteenth century, light was generally believed to consist of tiny particles. Of the phenomena mentioned above, reflection, refraction, and the sharp shadows caused by the straight path of light were well known, and the characteristic of finite velocity was suspected. All of these phenomena except refraction clearly could be expected of streams of particles, and Isaac Newton showed that refraction would occur if the velocity of light increased with the density of the medium through which it traveled.

This theory of the nature of light seemed to be completely upset, however, in the first half of the nineteenth century. During that time, Thomas Young

studied the phenomena of interference, and could see no way to account for them unless light were a wave motion. Diffraction and polarization had also been investigated by that time. Both were easily understandable on the basis of a wave theory of light, and diffraction eliminated the "sharp-shadow" argument for particles. Reflection and finite velocity were consistent with either picture. The final blow to the particle theory seemed to have been struck in 1849, when the speed of light was measured in different media and found to vary in just the opposite manner to that assumed by Newton. Therefore, in 1850, it seemed finally to be settled that light consisted of waves.

Even then, however, there was the problem of the medium in which light waves traveled. All other kinds of waves required a physical medium, but light traveled through a vacuum—faster, in fact, than through air or water. The term ether was proposed by James Clerk Maxwell and his contemporaries as a name for the unknown medium, but this scarcely solved the problem because no ether was ever actually found. Then, near the beginning of the twentieth century, came certain work on the emission and absorption of energy that seemed to be understandable only if light were assumed to have a particle or corpuscular nature. The external photoelectric effect, the emission of electrons from the surfaces of solids when light is incident on the surfaces, was one of these. At that time, then, science found itself in the uncomfortable position of knowing a considerable number of experimental facts about light, of which some were understandable regardless of whether light consisted of waves or particles, others appeared to make sense only if light were wavelike, and still others seemed to require it to have a particle nature. *See* ETHER HYPOTHESIS.

Theory

The study of light deals with some of the most fundamental properties of the physical world and is intimately linked with the study of the properties of submicroscopic particles on the one hand and with the properties of the entire universe on the other. The creation of electromagnetic radiation from matter and the creation of matter from radiation, both of which have been achieved, provide a fascinating insight into the unity of physics. The same is true of the deflection of light beams by strong gravitational fields, such as the bending of starlight passing near the Sun.

A classification of phenomena involving light according to their theoretical interpretation provides the clearest insight into the nature of light. When a detailed accounting of experimental facts is required, two groups of theories appear which, in the majority of cases, account separately for the wave and the corpuscular character of light. The quantum theories seem to obviate questions concerning this dual character of light, and make the classical wave theory and the simple corpuscular theory appear as two very useful limiting theories. It happens that the wave theories of light can cope with a considerable part of the phenomena involving electromagnetic radiation. Geometrical optics, based on the wave theory of light, can solve many of the more common problems of the propagation of light, such as refraction, provided that the limitations of the underlying theory are not disregarded. *See* GEOMETRICAL OPTICS.

Phenomena involving light may be classed into three groups: electromagnetic wave phenomena, corpuscular or quantum phenomena, and relativistic effects. The relativistic effects appear to influence similarly the observation of both corpuscular and wave phenomena. The major developments in the theory of light closely parallel the rise of modern physics. These developments are charted in **Fig. 2**, and are discussed in the remainder of this article.

Wave phenomena. Interference and diffraction, mentioned earlier, are the most striking manifestations of the wave character of light. Their fundamental similarity can be demonstrated in a number of experiments. The wave aspect of the entire spectrum of electromagnetic radiation is most convincingly shown by the similarity of diffraction pictures produced on a photographic plate, placed at some distance behind a diffraction grating, by radiations of different frequencies, such as x-rays and visible light. The interference phenomena of light are, moreover, very similar to interference of electronically produced microwaves and radio waves.

Polarization demonstrates the transverse character of light waves. Further proof of the electromagnetic character of light is found in the possibility of inducing, in a transparent body that is being traversed by a beam of plane-polarized light, the property of rotating the plane of polarization of the beam when the body is placed in a magnetic field. *See* FARADAY EFFECT.

The fact that the velocity of light had been calculated from electric and magnetic parameters (permittivity and permeability) was at the root of Maxwell's conclusion in 1865 that "light, including heat and other radiations if any, is a disturbance in the form of waves propagated . . . according to electromagnetic laws." Finally, the observation that electrons and neutrons can give rise to diffraction patterns quite similar to those produced by visible light has made it necessary to ascribe a wave character to particles. *See* ELECTRON DIFFRACTION; NEUTRON DIFFRACTION.

Electromagnetic-wave propagation. Electromagnetic waves can be propagated through free space, devoid of matter and fields, and with a constant gravitational potential; through space with a varying gravitational potential; and through more or less absorbing material media which may be solids, liquids, or gases. Radiation can be transmitted through waveguides with cylindrical, rectangular, or other boundaries, the insides of which can be either evacuated or filled with a dielectric medium. *See* WAVEGUIDE.

From electromagnetic theory, and especially from the well-known equations formulated by Maxwell, a plane wave disturbance of a single frequency

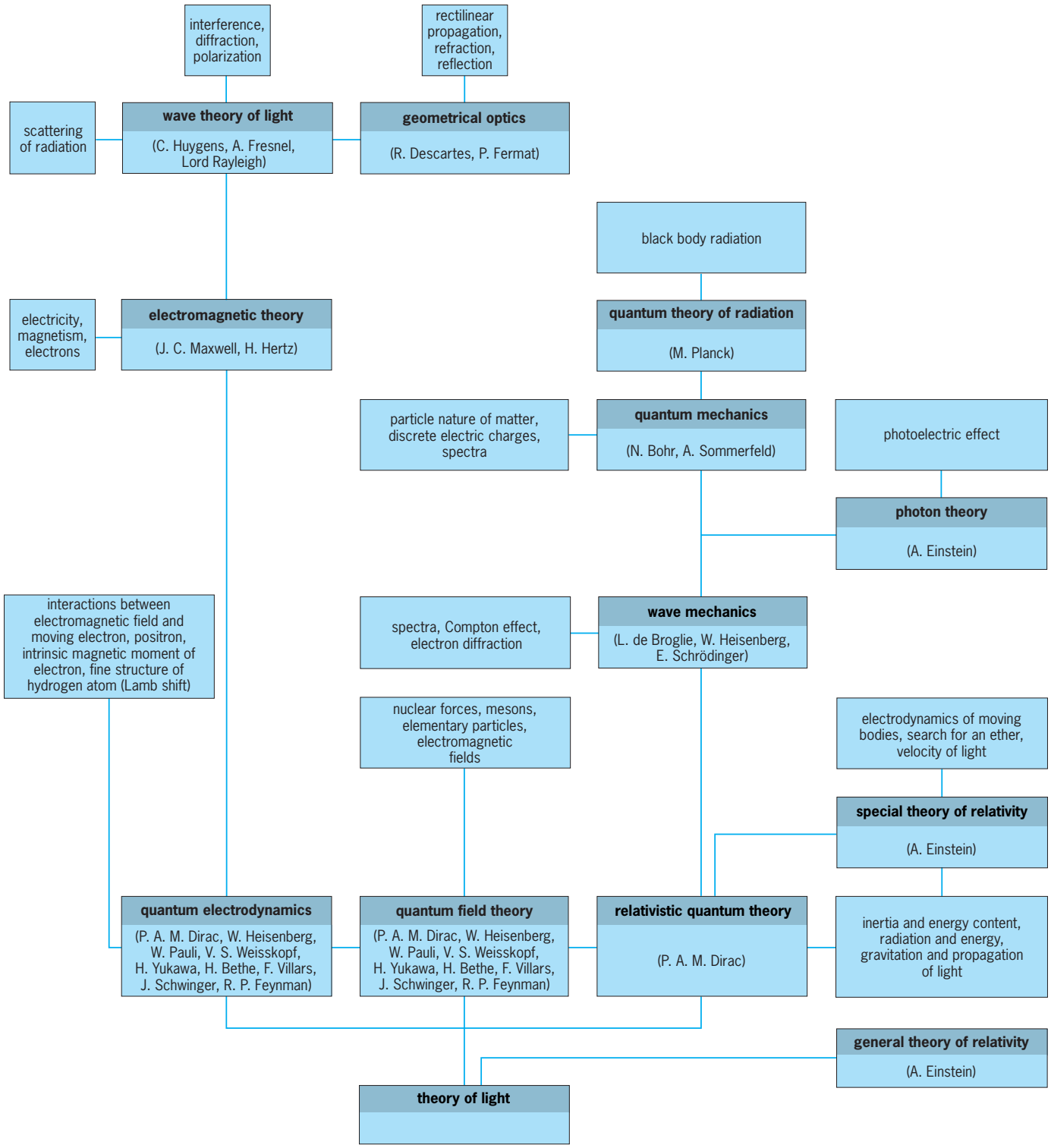

Fig. 2. Chart of important developments in the theory of light.

f is propagated in the x direction with a phase velocity $\upsilon = \lambda f = \lambda\omega/2\pi$, where λ is the free-space wavelength and $\omega = 2\pi f$. The wave can be described by the equation $y = A \cos \omega(t - x/\upsilon)$. Two disturbances of same amplitude A, of respective angular frequencies ω_1 and ω_2 and of velocities υ_1 and υ_2, propagated in the same direction, yield the resulting disturbance y', defined in Eq. (1). Here $\Delta\omega = \omega_1 - \omega_2$

$$y' = y_1 + y_2 = 2A \cos {}^1/_2 \left[(\Delta\omega)\, t - x\Delta\frac{\omega}{\upsilon}\right] \times \cos\left(\omega t - \frac{\omega x}{\upsilon}\right) \quad (1)$$

and $\omega = {}^1/_2(\omega_1 + \omega_2)$. The ratio $u = \Delta\omega/\Delta(\omega/\upsilon)$ is

defined as the group velocity, which is the speed of a light signal, just as the ratio $\omega/(\omega/\upsilon)$ is identical with the phase velocity, which is the speed of a wavefront. In the limit for small $\Delta\omega$, $u = d\omega/d(\omega/\upsilon)$. Noting that $\omega = 2\pi\upsilon/\lambda$, $d\omega = 2\pi(\lambda\, d\upsilon - \upsilon\, d\lambda)/\lambda^2$, and $d(\omega/\upsilon) = -2\pi\ d\lambda/\lambda^2$, an important relation, Eq. (2), between group and phase velocity is ob-

$$u = v - \lambda\frac{dv}{d\lambda} \tag{2}$$

tained. This shows that the group velocity u is different from the phase velocity υ in a medium with dispersion $d\upsilon/d\lambda$. In a vacuum, $u = \upsilon = c$. With the help of Fourier theorems, the preceding expression for u can be shown to apply to the propagation of a wave group of infinite length, but with frequencies extending over a finite small domain. Furthermore, even if the wave train were emitted with an infinite length, modulation or "chopping" would result in a degrading of the monochromaticity by introduction of new frequencies, and hence in the appearance of a group velocity. Such considerations of this nature were not trivial in earlier measurements of the velocity of light, but were quite fundamental to the conversion of instrumental readings to a value of c. Similar considerations apply to the incorporation of the effects of the medium and the boundaries involved in the experiments. Complications arise in the regions of anomalous dispersion (absorption regions), where the phase velocity can exceed c and $d\upsilon/d\lambda$ is positive. *See* DISPERSION (RADIATION); FOURIER SERIES AND TRANSFORMS; GROUP VELOCITY; MAXWELL'S EQUATIONS; PHASE VELOCITY; WAVE EQUATION; WAVE MOTION.

Refractive index. A plane wavefront, in going from a medium in which its phase velocity is υ into a second medium where the velocity is υ', changes direction at the interface. By geometry, it can be shown that $\sin i/\sin i' = \upsilon/\upsilon'$, where i and i' are the angles which the light path forms with a normal to the interface in the two media. It can also be shown that the path between any two points in this system is that which minimizes the time for the light to travel between the points (Fermat's principle). This path would not be a straight line unless $i = 90^\circ$ or $\upsilon = \upsilon'$. Snell's law states that $n \sin i = n' \sin i'$, where n is the index of refraction of the medium. It follows that the refractive index of a medium is $n = c/\upsilon'$, because the refractive index of a vacuum, where $\upsilon = c$, has the value 1.

Dispersion. The dispersion $d\upsilon/d\lambda$ of a medium can easily be obtained from measurements of the refractive index for different wavelengths of light. The weight of experimental evidence, based on astronomical observations, is against a dispersion in a vacuum. Observation of the light reaching the Earth from the eclipsing binary star Algol, 120 light-years distant (1 light-year $\cong 6 \times 10^{12}$ mi or 9.5×10^{12} km is the distance traversed in vacuum by a beam of light in 1 year), shows that the light for all colors is received simultaneously. The eclipsing occurs every 68 h 49 min. Were there a difference of velocity for red light and for blue light in interstellar space as great as 1 part in 10^6, this star would show a measurable time difference in the occurrence of the eclipses in these two colors. *See* BINARY STAR.

Corpuscular phenomena. In its interactions with matter, light exchanges energy only in discrete amounts, called quanta. This fact is difficult to reconcile with the idea that light energy is spread out in a wave, but is easily visualized in terms of corpuscles, or photons, of light. *See* PHOTON.

Blackbody (heat) radiation. The radiation from theoretically perfect heat radiators, called blackbodies, involves the exchange of energy between radiation and matter in an enclosed cavity. The observed frequency distribution of the radiation emitted by the enclosure at a given temperature of the cavity can be correctly described by theory only if it is assumed that light of frequency ν is absorbed in integral multiples of a quantum of energy equal to $h\nu$, where h is a fundamental physical constant called Planck's constant. This startling departure from classical physics was made by Max Planck early in the twentieth century. *See* HEAT RADIATION.

Photoelectric effect. When a monochromatic beam of electromagnetic radiation illuminates the surface of a solid (or less commonly, a liquid), electrons are ejected from the surface in the phenomenon known as photoemission, or the external photoelectric effect. The kinetic energies of the electrons can be measured electronically by means of a collector which is negatively charged with respect to the emitting surface. It is found that the emission of these photoelectrons, as they are called, is immediate, and independent of the intensity of the light beam, even at very low light intensities. This fact excludes the possibility of accumulation of energy from the light beam until an amount corresponding to the kinetic energy of the ejected electron has been reached. The number of electrons is proportional to the intensity of the incident beam. The velocities of the electrons ejected by light at varying frequencies agree with Eq. (3), where m is the mass of the electron, υ_{max}

$$^1/_2 m\upsilon_{max}^2 = h(\nu - \nu_0) \tag{3}$$

the maximum observed velocity, ν the frequency of the illuminating light beam, and ν_0 a threshold frequency characteristic of the emitting solid.

In 1905 Albert Einstein showed that the photoelectric effect could be accounted for by assuming that, if the incident light is composed of photons of energy $h\nu$, part of this energy, $h\nu_0$, is used to overcome the forces binding the electron to the surface. The rest of the energy appears as kinetic energy of the ejected electron.

Compton effect. The scattering of x-rays of frequency ν_0 by the lighter elements is caused by the collision of x-ray photons with electrons. Under such circumstances, both a scattered x-ray photon and a scattered electron are observed, and the scattered x-ray has a lower frequency than the impinging x-ray. The kinetic energy of the impinging x-ray, the scattered

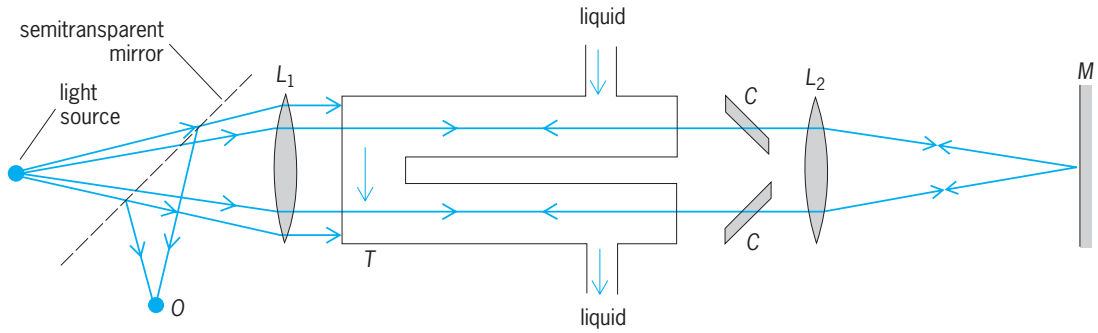

Fig. 3. Fizeau's experiment. *C*, compensator plates; *M*, mirror; L_1, L_2, lenses; *T*, tube; *O*, interference fringes.

x-ray, and the scattered electron, as well as their relative directions, are in agreement with calculations involving the conservation of energy and momentum. *See* COMPTON EFFECT.

Quantum theories. The need for reconciling Maxwell's theory of the electromagnetic field, which describes the electromagnetic wave character of light, with the particle nature of photons, which demonstrates the equally important corpuscular character of light, has resulted in the formulation of several theories which go a long way toward giving a satisfactory unified treatment of the wave and the corpuscular picture. These theories incorporate, on one hand, the theory of quantum electrodynamics, first set forth by P. A. M. Dirac, P. Jordan, W. Heisenberg, and W. Pauli, and on the other, the earlier quantum mechanics of L. de Broglie, Heisenberg, and E. Schrödinger. Unresolved theoretical difficulties persist, however, in the higher-than-first approximations of the interactions between light and elementary particles. The incorporation of a theory of the nucleus into a theory of light is bound to call for additional formulation.

Dirac's synthesis of the wave and corpuscular theories of light is based on rewriting Maxwell's equations in a hamiltonian form resembling the hamiltonian equations of classical mechanics. Using the same formalism involved in the transformation of classical into wave-mechanical equations by the introduction of the quantum of action $h\nu$, Dirac obtained a new equation of the electromagnetic field. The solutions of this equation require quantized waves, corresponding to photons. The superposition of these solutions represents the electromagnetic field. The quantized waves are subject to Heisenberg's uncertainty principle. The quantized description of radiation cannot be taken literally in terms of either photons or waves, but rather is a description of the probability of occurrence in a given region of a given interaction or observation. *See* HAMILTON'S EQUATIONS OF MOTION; NONRELATIVISTIC QUANTUM THEORY; QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; QUANTUM MECHANICS; RELATIVISTIC QUANTUM THEORY; UNCERTAINTY PRINCIPLE.

Relativistic effects. The measured magnitudes of such characteristics as wavelength and frequency, velocity, and the direction of the radiation in a light beam are affected by a relative motion of the source with respect to the observer, occurring during the emission of the signal-carrying electromagnetic wave trains. *See* DOPPLER EFFECT.

A difference in gravitational potential also affects these quantities. Several important observations of this nature are listed in this section, followed by a discussion of several important results of general relativity theory involving light. For extended discussions of both the special theory of relativity and the general theory of relativity *see* RELATIVITY.

Velocity in moving media. In 1818 A. Fresnel suggested that it should be possible to determine the velocity of light in a moving medium, for example, to determine the velocity of a beam of light traversing a column of liquid of length d and of refractive index n, flowing with a velocity υ relative to the observer, by measuring the optical thickness nd. The experiment was carried out by Fizeau in a modified Rayleigh interferometer, shown in **Fig. 3**, by measuring the fringe displacement in O corresponding to the reversing of the direction of flow. If υ' is the phase velocity of light in the medium (deduced from the refractive index by the relation $\upsilon = c/n$), it is found that the measured velocity υ_m in the moving medium can be expressed as $\upsilon_m = \upsilon' + \upsilon(1 - 1/n^2)$ rather than $\upsilon_m = \upsilon' + \upsilon$, as would be the case with a newtonian velocity addition.

Aberration. J. Bradley discovered in 1725 a yearly variation in the angular position of stars, the total variation being 41 seconds of arc. This effect is in addition to the well-known parallax effect, and was properly ascribed to the combination of the velocity of the Earth in its orbit and the speed of light. Bradley used the amplitude of the variation to arrive at a value of the velocity of light. George Airy compared the angle of aberration in a telescope before and after filling it with water, and discovered, contrary to his expectation, that there was no difference in angle. *See* ABERRATION (ASTRONOMY).

Michelson-Morley experiment. The famous Michelson-Morley experiment, one of the most significant experiments of all time, was performed in 1887 to measure the relative velocity of the Earth through inertial space. Inertial space is space in which Newton's laws of motion hold. Dynamically, an inertial frame of reference is one in which the observed accelerations are zero if no forces act. A point in an orbit is the

center of such a frame. *See* FRAME OF REFERENCE.

The rotation of the Earth about its axis, with tangential velocities never exceeding 0.3 mi/s (0.5 km/s) is easily demonstrated mechanically (Foucault's pendulum, precession of gyroscopes) and optically (Michelson's rectangular interferometer). The surface of the Earth is not an inertial frame. In its orbit around the Sun, on the other hand, the Earth has translational velocities of the order of 18 mi/s (30 km/s), but this motion cannot be detected by mechanical experiments because of its orbital nature. The hope existed, however, that optical experiments would permit the detection (and measurement) of the relative motion of the Earth through inertial space by comparing the times of travel of two light beams, one traveling in the direction of the translation through inertial space, and the other at right angles to it. The hope was based on the now disproven notion that the velocity of a light would be equal to the constant c only when measured with respect to the inertial space, but would be measured as smaller $(c - v)$ or greater $(c + v)$ with respect to a reference frame, such as the Earth, moving with a velocity v in inertial space if a light beam were projected respectively in the direction and in the opposite direction of translation of this frame. According to classical velocity addition theorems (which, as is now known, do not apply to light), a velocity difference of $2v$ would be detected under such circumstances. *See* EARTH ROTATION AND ORBITAL MOTION.

Not only does the Earth move in an orbit around the Sun, but it is carried with the Sun in the galactic rotation toward Cygnus with a velocity of several hundreds of miles per second, and the Galaxy itself is moving with a high speed in its local spiral group. Speeds of hundreds and possibly thousands of miles per second should be detectable by measurements on Earth in two orthogonal directions, assuming of course that the Earth motion is itself with respect to inertial space, or indeed that such a space has the physical meaning ascribed to it. The unexpected result of the experiment was that no such velocity difference could be detected, that is, no relative motion could be detected by optical means.

The Michelson-Morley apparatus (**Fig. 4**) consists of a horizontal Michelson interferometer with its two arms at right angles. The mirrors are adjusted so that the central white-light fringe falls on the cross hair of the observing telescope. This indicates equality of optical phase, and therefore an equality of the times taken by the light beams to travel from the beam-splitting surface to each of the two mirrors and back. Rotation of the entire system by 90°, or indeed by any angle, as well as repetition of the experiment at various times of the year all are found to leave the central white-light fringe and associated fringe system undisplaced, indicating no change in the time required by the light to traverse the two arms of the interferometer when their directions relative to the direction of the Earth's motion are varied. Had there been a difference in the velocity of light in the two directions OM and ON, the two arms would be of unequal length in the initial adjustment. For example, if the

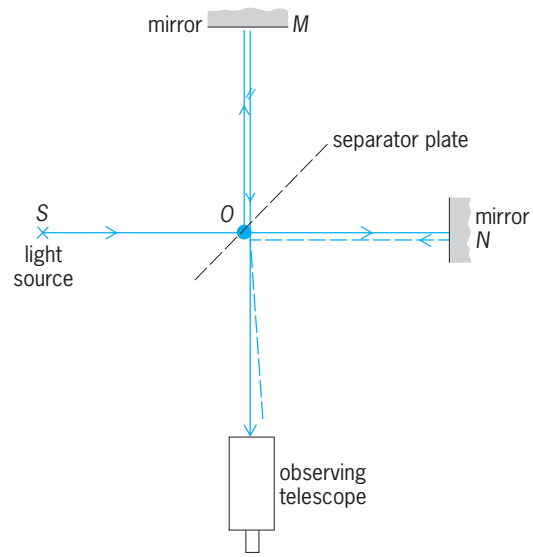

Fig. 4. Michelson-Morley experiment.

light traveled faster in the direction OM (on the average, going back and forth), then the corresponding arm would have to be longer so as to make the time of travel equal in both arms. If the apparatus were turned through 90°, the shorter arm would take the place of the longer arm, and the "faster" light would now travel in the shorter arm, and the "slower" light in the longer arm; a noticeable fringe displacement would, but actually does not, take place.

Einstein's theory accounts for the null result by the simple explanation that no relative motion between the apparatus and the observer exists in the experiment. No change in measured length has occurred in either direction, and because the propagation of light is isotropic, no velocity difference or detection of relative motion should be expected.

Gravitational and cosmological redshifts. Two different kinds of shifts, or displacements of spectral lines toward the red end of the spectrum, are observed in spectrograms taken with starlight or light from nebulae. One is the rare, but extremely significant, gravitational redshift or Einstein shift, which has been measured in some spectra from white dwarf stars. The other, much more widely encountered, is the redshift in spectra from external galaxies usually described as being caused by a radial Doppler effect characterizing an "expansion of the universe."

The most famous example of the gravitational redshift is that observed in the spectrum of the so-called dark companion of Sirius. The existence of this companion was predicted by F. W. Bessel in 1844 on the basis of a perturbation in the motion of Sirius, but because of its weak luminosity (1/360 that of the Sun, 1/11,000 that of Sirius), it was not observed until 1861. This companion is a white dwarf of a mass comparable to that of the Sun (0.95), but having a relatively small radius of 11,000 mi (18,000 km) and the fantastically high density of 61,000 times that of water. This companion shows a shift in its spectral lines relative to the ones emitted by Sirius itself; the shift of 0.03 nm for the Balmer β-line of hydrogen

was reliably determined by W. S. Adams. In 1960, R. V. Pound and G. A. Rebka, Jr., measured the gravitational redshift in a laboratory experiment involving the Mössbauer effect. *See* GRAVITATIONAL REDSHIFT.

The cosmological redshift is a systematic shift observed in the spectra of all galaxies, best measured with the calcium H and K absorption lines. Distances of galaxies are determined photometrically by measurements of their intensities, and it is found that the wavelength shift toward the red increases with the distance of the galaxies from the Earth. (The Earth does not have a privileged position if expansion of the universe is indeed involved. Rather its position is somewhat like that of a person in a crowd which is dispersing—each individual experiencing an ever-increasing distance from every other one.) The change of wavelength with distance d is given by Eq. (4), where $1/H$ is $1\text{–}2 \times 10^{10}$ light-years.

$$\frac{\Delta\lambda}{\lambda} = Hd \tag{4}$$

See REDSHIFT.

Results of general relativity. The propagation of light is influenced by gravitation. This is one of the fundamental results of Einstein's general theory of relativity which has been subjected to experimental tests and found to be verified. Three important results involving light need to be singled out.

1. The velocity of light, measured by the same magnitude c independently of the state of motion of the frame in which the measurement is being carried out, depends on the gravitational potential Φ of the field in which it is being measured according to Eq. (5). Here $\Phi = -GM/R$, where G is the universal

$$c = c_0\left(1 + \frac{\Phi}{c^2}\right) \tag{5}$$

constant of gravitation (6.670×10^{-11} in SI units), M the mass of the celestial body, R the radius of the body, and c_0 the velocity of light in a vacuum devoid of fields.

For example, the absolute value of the term Φ/c^2 is about 3000 times greater on the Sun than on Earth, making the measurements of c smaller by two parts in 10^6 on the Sun as compared to those on Earth.

2. The frequency ν of light emitted from a source in a gravitational field with the gravitational potential Φ is different from the frequency ν_0 emitted by an identical source (atomic, nuclear, molecular) in a field-free region, according to Eq. (6). Spectral lines

$$\nu = \nu_0\left(1 + \frac{\Phi}{c^2}\right) \tag{6}$$

in sunlight should be displaced toward the red by two parts in 10^6 when compared to light from terrestrial sources.

3. Light rays are deflected when passing near a heavenly body according to Eq. (7), where α is the

$$\alpha = \frac{4GM}{c^2R} \tag{7}$$

angular deflection in radians, and R the distance of the beam from the center of the heavenly body of mass M. The deflection is directed so as to increase the apparent angular distance of a star from the center of the Sun when starlight is passing near the edge of the Sun. The deflection according to this equation should be 1.75 seconds of arc, a value which compares favorably with eclipse measurements of the star field around the Sun in 1931. These measurements indicated values up to 2.2 seconds of arc when compared with photographs of the same field 6 months earlier. Measurements of the deflection of radio waves from extremely small diameter celestial radio sources in 1974–1975 agreed with Einstein's theory to within 1%. This prediction of Einstein's theory might seem less surprising today when the corpuscular-photon character of light is widely known, and when a newtonian M/R^2 attraction might be considered to be involved in the motion of a corpuscle with the velocity c past the Sun. However, application of Newton's law predicts a deviation only half as great as the well verified relativistic prediction.

Matter and radiation. The possibility of creating a pair of electrons—a positively charged one (positron) and a negatively charged one (negatron)—by a rapidly varying electromagnetic field (gamma rays of high frequency) was predicted as a consequence of Dirac's wave equation for a free electron and has been experimentally verified. I. Curie and F. Joliot, as well as J. Chadwick, P. M. S. Blackett, G. P. Occhialini, and others have compared the number of positrons and negatrons ejected by gamma rays passing through a thin sheet of lead (and other materials) and have found them to be the same, after accounting for two other groups of electrons also appearing in the experiment (photoelectrons and recoil electrons). Other examples of negatron-positron pair production include the collision of two heavy particles, a fast electron passing through the field of a nucleus, the direct collision of two electrons, the collision of two light quanta in a vacuum, and the action of a nuclear field on a gamma ray emitted by the nucleus involved in the action.

Evidence of the creation of matter from radiation, as well as that of radiation from matter, substantiates Einstein's equation (8), which was first expressed in

$$E = mc^2 \tag{8}$$

the following words: "If a body [of mass m] gives off the energy E in the form of radiation, its mass diminishes by E/c^2."

In regard to exchanges of energy and momentum, electromagnetic waves behave like a group of particles with energy as in Eq. (9) and momentum as in Eq. (10).

$$E = mc^2 = h\nu \tag{9}$$

$$p = \frac{h\nu}{c} = \frac{h}{\lambda} \tag{10}$$

Finally, many experiments with photons show that they also possess an intrinsic angular momentum, as

do particles. Circularly polarized light, for example, carries an experimentally observable angular momentum, and it can be shown that, under certain circumstances, an angular momentum can be imparted to unpolarized or plane-polarized light (plane wave passing through a finite circular aperture). In any case, the angular momentum will be quantized in units of $h/2\pi$.

The inverse process to the creation of electron pairs is the annihilation of a positron and a negatron, resulting in the production of two gamma-ray quanta (two-quantum annihilation). Nuclear chain reactions are known to involve similar processes. *See* CHAIN REACTION (PHYSICS); ELECTRON-POSITRON PAIR PRODUCTION; ELEMENTARY PARTICLE. George W. Stroke

Bibliography. M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; W. Demtröoder, *Laser Spectroscopy: Basic Concepts and Instrumentation*, 2d ed., 1994; E. Hecht and A. Zajac, *Optics*, 3d ed., 1997; J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons*, 2d ed., 1985; F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976; R. Loudon, *The Quantum Theory of Light*, 3d ed., 2000; J. Meyer-Arendt, *Introduction to Classical and Modern Optics*, 4th ed., 1995; W. Rindler, *Essential Relativity*, rev. ed., 1986, paper 1993; F. G. Smith and J. H. Thomson, *Optics*, 2d ed., 1988.

Light amplifier

In the broadest sense, a device which produces an enhanced light output when actuated by incident light. A simple photocell relay–light source combination would satisfy this definition. To make the term more meaningful, common usage has introduced two restrictions: (1) a light amplifier must be a device which, when actuated by a light image, reproduces a similar image of enhanced brightness; and (2) the device must be capable of operating at very low light levels without introducing spurious brightness variations (noise) into the reproduced image. The term is used synonymously with image intensifier. The light amplifier increases the brightness of an image which is below the visual threshold to a level where it can be readily seen with the unaided eye. It is, of course, impossible to see under conditions of complete darkness. Indeed, there is a fundamental lower limit of illumination under which an image of a given quality can be recognized. This limitation arises because of the corpuscular nature of light. *See* PHOTON.

Photons arriving through a lens, or other optical system, onto an image area are random in time. If the number of photons per unit area arriving during the time allotted for image formation (for example, the period of the persistence of vision) is too low, the statistical fluctuation will be greater than the variation in number due to true differences in image brightness. Under these circumstances, image recognition is impossible.

Image-intensifier tubes. Intensifier tubes may consist of a semitransparent photocathode which emits electrons with a density distribution proportional to the distribution of light intensity incident on it. Thus, when a light image falls on one side of the cathode, an electron-current image is emitted from the other side. An electron optical system, which acts on electrons in much the same way as does a glass lens system on light, focuses the electron image onto an intensifier element. This electron optical system may be purely electrostatic or may utilize magnetic focusing. The intensified electron image from the other side of the intensifier element is again focused, by a second electron lens, onto a second intensifier element (if additional amplification is required) or onto a fluorescent viewing screen, where the electron energy is converted to visible light.

Proximity tubes are a class of intensifiers in which the photocathode and the fluorescent viewing screen are parallel and separated by only a short distance. No focusing is required for a proximity tube, since the electrons generated at the photocathode are accelerated along the tube axis by the high electric fields.

The spectral response of the image-intensifier tube depends on the type of photocathode employed, and the tube may serve as a wavelength converter as well as an intensifer. *See* CHARGED PARTICLE OPTICS.

Intensifier elements. Two types of intensifier elements are ordinarily employed. One consists of a transparent support (either a thin film or a fiber optics plate) with a phosphor screen on the side on which the electron image is incident, and a photocathode on the other. An electron striking the phosphor produces a flash of light which causes a release of 50 or more electrons from the photoemitter directly opposite the point of impact. Thus, the intensifier element increases the electron-image current density by a factor of 50. Two such intensifier elements in cascade give a gain of 2500. The **illustration** shows a three-stage electrostatic image-intensifier tube.

A second type of high-gain intensifier element is a thin secondary-emission current amplifier called a microchannel plate, which is placed between the photocathode and screen. It consists of a parallel bundle of small, hollow glass cylinders, where the inside walls of the cylinders are coated with a secondary emitting material. Electrons that are emitted from the photocathode strike the inside walls of the cylinders, causing secondary electron generation. These secondary electrons in turn continue to cascade down the inside walls of the cylinders, resulting in a high total current gain.

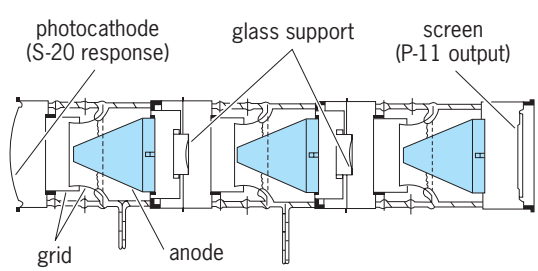

Three-stage electrostatically focused image tube.

Television camera tube. The intensifier image device can be combined with television pickup tubes to make a television camera whose sensitivity is very close to the threshold determined by the photon "shot noise" from the scene being televised. A silicon intensifier target (SIT) camera tube has a wide dynamic range and operates at very low light levels. In the SIT tube, the photoelectrons accelerated from the photocathode are focused onto a thin silicon target consisting of an array of *pn* junctions. The high-energy photoelectrons generate multiple hole-electron pairs in the silicon target, resulting in gain. The signal is read out by scanning the reverse side of the silicon target with an electron beam, similar to the method used in vidicon camera tubes. In other types of electron-bombarded silicon (EBS) tubes, the silicon *pn* junction target is replaced with scanned photodiode arrays or with charge-coupled device (CCD) imagers. In these cases, the electronic readout of the silicon targets is accomplished without the use of a scanned electron beam, and this results in even lower light-level performance. *See* TELEVISION CAMERA TUBE.

Solid-state image intensifiers. A greal deal of exploratory work has been done on solid-state light amplifiers. The form which has been extensively investigated consists of a photoconductive film in contact with an electroluminescent screen. Current flows through the photoconductor at illuminated areas, causing light to be generated in the electroluminescent layer. This type of intensifier gives considerable image enhancement at intermediate light levels, but fails at very low light levels because of extreme time lag. *See* ELECTROLUMINESCENCE.

Another type of solid-state light amplifier which works at intermediate light levels is the liquid-crystal light valve. The light valve consists of an electrically biased multilayer sandwich which has a liquid-crystal layer on the front face and a photoconductive layer on the rear face. In this system, the incident light is focused on the photoconductive side, while a bright, polarized light is projected onto the liquid-crystal side. The incident light from the photoconductive side causes an increase in voltage drop across the liquid crystal. This increased voltage causes a rotation in the orientation of the liquid-crystal molecules, which in turn modulates the polarization of the reflected projection light. Only the modulated reflected light is allowed to be viewed.

Applications. In addition to their application for night vision, light amplifiers have been useful in many fields of science, such as astronomy, nuclear physics, and microbiology. Dean R. Collins

Bibliography. G. Kychakoff (ed.), *Electro-Optic Sensing and Measurement*, 1988; K. Seyrafi and S. A. Hovanessian, *Introduction to Electro-Optical Imaging and Tracking Systems*, 1993.

Light curves

Graphs of the intensity of radiation from astronomical objects as they change with time. Variations may be caused by the changing perspective from the Earth of two stars in orbit around each other, by pulsations that change an individual star's size and surface temperature, by mass ejection or accretion, by explosions, by beams of radiation sweeping across the line of sight from the Earth, or by clouds of very high-energy electrons in powerful magnetic fields. The information contained in the light curve includes the timing of events, such as eclipses or pulses, and the amplitude of changes in the radiation received at Earth.

Each data point in a light curve is a photometric measurement, recorded at a particular time. These points represent measurements of the amount of radiation from the source received at Earth per second per area in a particular bandpass, for example, through a blue filter. In optical light, photometry is affected by the varying transparency of the atmosphere, so that light curves are often obtained as ratios to the intensities of nearby comparison stars. Simultaneous measurements can be made visually, with photometers that have two channels, or with imagers such as charge-coupled devices (CCDs) or photographic emulsions. These relative light curves are put on an absolute scale by means of calibrating measures of the comparison stars taken on clear nights. *See* ASTRONOMICAL PHOTOGRAPHY; CHARGE-COUPLED DEVICES; MAGNITUDE (ASTRONOMY); PHOTOMETRY.

Binary stars. If the orbital plane of a binary star system lies nearly along the line of sight from the Earth, the stars appear to eclipse each other periodically, either partially or totally, depending on the tilt of the orbit (**Fig. 1**). The larger drop in intensity represents the primary eclipse, when the dimmer star passes in front of the brighter; the secondary eclipse is shallower, produced when the light from the dimmer star is blocked. If one face of the cooler star is strongly irradiated by a hot primary or gravity distorts the star into a nonspherical shape, the light curve also shows a hump centered on the secondary eclipse. The timing of the eclipses yields the ratio of the radii of the two stars and the period and tilt of the orbit. Temperature differences between the two stars make the drop in intensity during eclipse differ among color bands. *See* BINARY STAR; ECLIPSING VARIABLE STARS.

Variable stars. Intrinsically variable single stars have distinct signatures in their light curves. Some giant stars have an instability in their outer layers that acts as a heat engine driving radial pulsations of the star's surface. A light curve (**Fig. 2**) reveals the amount of extra flux received at Earth during each pulsation. Spectroscopy gives the temperature and change in radius of the star during the same cycle. When combined, this information yields the distance to the variable star and its intrinsic power output. For example, Cepheid variables can be identified by their light curve signatures and used as distance indicators for nearby galaxies. *See* CEPHEIDS.

Other types of stars also show surface pulsations that appear as variations in their light curves. These types include early B stars on the main sequence, and stellar remnants such as planetary nebula nuclei and white dwarfs. The amplitudes of these

pulsations can be very low (<1% of the white light intensity) and can consist of a complicated blend of many frequencies. These variations are interpreted as nonradial pulsations that act like waves near the star's surface. The central stars of planetary nebulae are evolving so fast that the frequencies of the pulsations show tiny but measurable shifts over the course of 5 to 10 years that are produced by the collapsing cores as they become white dwarf stars. *See* PLANETARY NEBULA; WHITE DWARF STAR.

Many single stars show irregular variations brought about by mass loss. Illuminated ejecta add to the brightness until the material disperses. Newly formed stars may be accreting new gas from disks at their equators and ejecting material in jets at the poles. These T Tauri stars show strong variability. Hot stars of type B with emission lines (Be stars) on the main sequence also have disks or shells that vary in light with irregular mass-loss episodes. Red supergiant stars eject tremendous amounts of matter from their outer envelopes; long-period variables, Miras, and nebular variables all fall in this class. Light curves for these objects in both the optical and infrared provide the time history of mass ejection. *See* STAR; STELLAR EVOLUTION; T TAURI STAR; VARIABLE STAR.

Interacting binaries. The stars in some binary systems are so close together that gas flows between the components. If the primary is a high-density stellar remnant, the gas may spiral in, forming an accretion disk. The accretion flow itself may radiate much of the power from the system, especially in the ultraviolet and x-rays. These high-energy light curves show evidence for flickering and for a hot spot where the stream of material impacts the disk. If the compact star is a white dwarf, the buildup of gas on its surface can cause a nuclear explosion, which is a nova outburst. The light curve shows the object brightening by 10 to 100,000 times its quiescent level. X-ray light curves have shown that some unseen compact companions must be neutron stars or black holes. *See* BLACK HOLE; CATACLYSMIC VARIABLE; NEUTRON STAR; NOVA; X-RAY ASTRONOMY.

Supernovae. The ultimate explosion is when an entire star (or pair of stars) detonates; this explosion is a supernova, which can be brighter than a whole galaxy of stars combined. The shape of the light curves of supernovae demonstrates that after the initial explosion the energy input is dominated by the decay of radioactive elements such as cobalt and nickel isotopes produced by the thermonuclear runaway. The uniformity in the shapes of the light curves of one subgroup of supernovae (type Ia) has suggested that those objects achieve a similar maximum power output and can be used as so-called standard candles for extragalactic distance determinations (**Fig. 3**). The steepness of the light curve immediately after maximum light provides a fine tuning for determining the intrinsic brightness that has an important impact on the distances derived from such explosions. The comparison of the maximum brightness of distant supernovae with that of their local counterparts led to the discovery that the expansion of the universe is accelerating. *See* ACCELERATING UNIVERSE; SUPERNOVA.

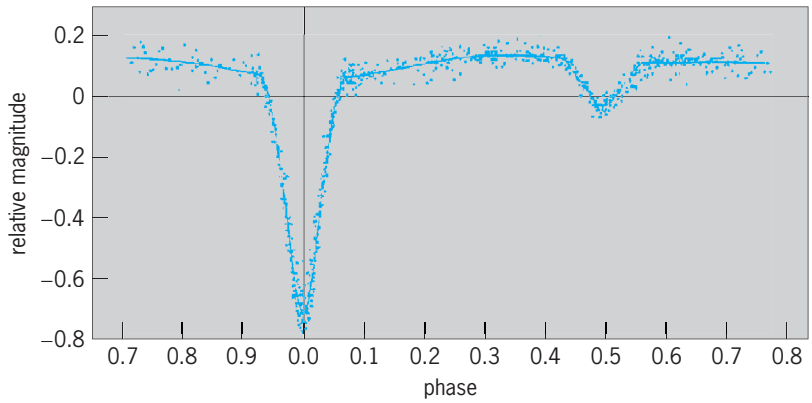

Fig. 1. Light curve of the eclipsing binary Algol. The measured infrared intensity on the vertical axis is given in magnitudes relative to the nearby star κ Persei. The phase is the fraction of the 2.9-day period elapsed relative to the deep primary minimum. Also apparent are the secondary minimum, when the K giant is eclipsed by the B star, and the brightening around secondary minimum caused primarily by heating of one face of the cooler star. (*After K.-Y. Chen and G. Reuning, Astron. J., 71:283–296, 1966*)

Exotic objects. Variable radiation can be emitted by clouds of very high-energy electrons moving in powerful magnetic fields. Beams and jets of radiated energy may result. Pulsars are an example, rapidly rotating neutron stars that sweep the Earth with such beams up to a thousand times per second. Their light curves are commonly observed at radio frequencies, but young pulsars also show pulses in optical light. The timing of the radio pulses for binary pulsars provides a general relativity laboratory for gravitational radiation. The frequency of the pulses from single pulsars can be determined with such accuracy that very small variations have been interpreted as

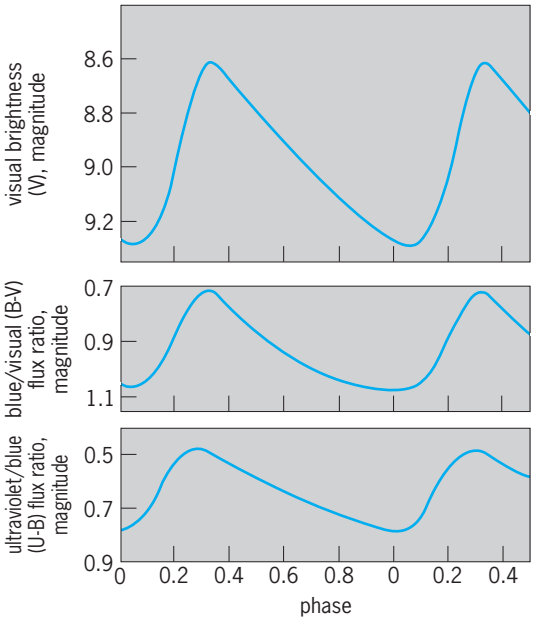

Fig. 2. Light curves of a Cepheid variable star. For the lower two curves, bluer colors are plotted toward the top. The phase is the fraction of the 4.9-day pulsation period, relative to the time of minimum brightness in the cycle. (*After R. I. Mitchell et al., Boletin de los Observatorios Tonantzintla y Tacubaya, 3:153–304, 1964*)

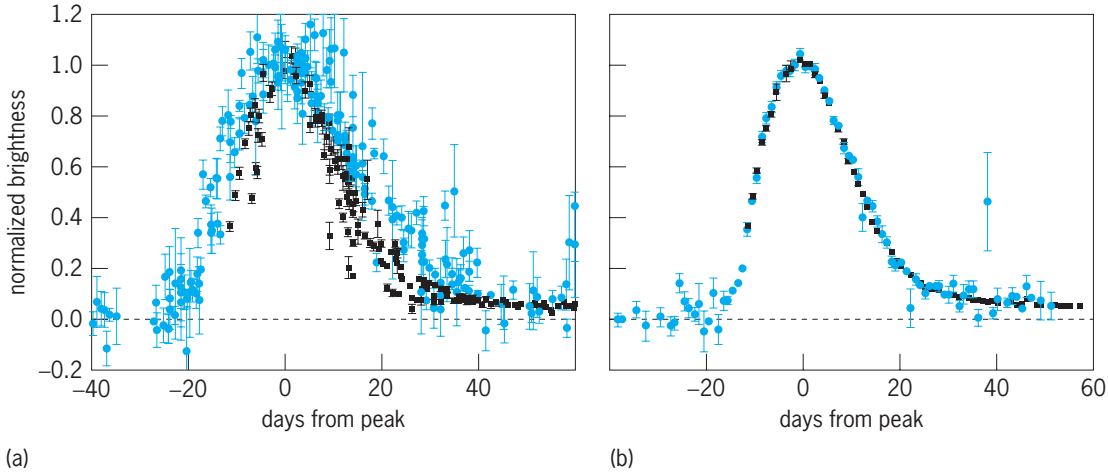

Fig. 3. Light curves for supernovae. (*a*) Composite of blue photoelectric measurements for 18 lower-redshift supernovae (black data points) and red measurements for 35 higher-redshift objects (color data points) normalized to maximum = 1 at time = 0. (*b*) Same data corrected for time dilation and light-curve shape. (*After G. Goldhaber et al., Timescale stretch parameterization of type Ia supernova B-band light curves, Astrophys. J., 558:359–368, 2001*)

arising from the perturbations of orbiting planets. *See* GRAVITATIONAL RADIATION; PULSAR.

Gamma-ray bursts are brief outbursts of high-energy radiation that are random in time and are distributed nearly uniformly over the sky. Most of the energy in a burst is higher than 100,000 eV (more than 7000 times the energy required to strip the electron off a hydrogen atom). Bursts can range in duration from 30 ms to over 1000 s. The shapes of the light curves range from single pulses or spikes through multiple peaks to chaotic or erratic episodes. Observational characterization of bursts has come from the National Aeronautics and Space Administration's (NASA's) *Compton Gamma-Ray Observatory* and more recently NASA's *SWIFT* satellite. Gamma-ray bursts are sometimes followed by afterglows in x-rays and optical light, allowing determination of redshifts for some sources. At least one gamma-ray burst has been unambiguously associated with a supernova. In general, models interpreting gamma-ray bursts are based on mergers of neutron stars or of a neutron star with a black hole, or the core collapse of a massive star in the midst of a merger. *See* GAMMA-RAY BURSTS.

Blazars are actively varying powerful nuclei of galaxies. They may have giant beams and jets produced by supermassive black holes, and present strongly variable light curves in radio, optical, ultraviolet, and x-rays. The time lag of variations in the light curve between the ultraviolet and high-energy x-rays suggests that relativistic beaming enhances the intensity of radiation in the direction of the Earth. The lag also suggests that the mechanism of relativistic electrons giving a boost in energy to optical photons to produce x-rays is not operative in these blazars. Quasars are modeled as having both accretion disks and jets from supermassive black holes. Their light curves show more modest variability than those of blazars. The time delays between variations in the continuous radiation from the central engine and the spectral lines emitted by the surrounding clouds of gas are interpreted to yield the physical size of these active galactic cores. *See* GALAXY, EXTERNAL; QUASAR.

Gravitational microlensing. Einstein's theory of general relativity is based on the underlying principle that the presence of matter determines the local curvature of space-time. A large, concentrated mass can therefore act as a lens and bend the path of light. Radiation from a distant source behind the lensing mass can be focused, thus amplifying the observed intensity. That amplification will vary if the relative position of the lens and background object changes with time. A characteristic gravitational microlensing light curve (roughly symmetrical in time and independent of color) is produced when the lensing object is a star with proper motion that carries it across our line of sight through a critical alignment with a background star. The detailed shape of the light curve can show whether the lensing object is a binary star or if it has an associated planetary system; one such planet was reported in 2006 based on a 1-day microlensing peak superimposed on a year-long microlensing stellar curve. The technique of searching for microlensing signatures in repeated exposures of dense star fields was used to search for the objects in our Galaxy's halo that stabilize its rotation and could comprise the local dark matter, termed massive compact halo objects (MACHOs). The low incidence of microlensing events towards the Large Magellanic Cloud puts stringent limits on the total masses of MACHOs. *See* DARK MATTER; GRAVITATIONAL LENS; MAGELLANIC CLOUDS; MILKY WAY GALAXY.

Richard F. Green

Bibliography. P. Mészáros, Theories of gamma-ray bursts, *Annu. Rev. Astron. Astrophys.*, 40:137–169, 2002; H. R. Miller and P. J. Wiita (eds.), *Variability of Active Galactic Nuclei*, 1991; B. Paczyński, Gravitational microlensing, *Annu. Rev. Astron. Astrophys.*, 34:419–459, 1996; J. R. Percy (ed.), *The Study of Variable Stars Using Small Telescopes*, 1986; E. Valtaoja and M. Valtonen (eds.), *Variability of Blazars*, 1992; S. E. Woosley (ed.), *Supernovae*, 1991.

Light-emitting diode

A rectifying semiconductor device which converts electrical energy into electromagnetic radiation. The wavelength of the emitted radiation ranges from the near-ultraviolet to the near-infrared, that is, from about 400 to over 1500 nanometers.

General considerations. Most commercial light-emitting diodes (LEDs), both visible and infrared, are fabricated from III–V semiconductors. These compounds contain elements such as gallium, indium, and aluminum from column III (or group 13) of the periodic table, as well as arsenic, phosphorus, and nitrogen from column V (or group 15) of the periodic table. There are also LED products made of II–VI (or group 12–16) semiconductors, for example ZnSe and related compounds. Taken together, these semiconductors possess the proper band-gap energies to produce radiation at all wavelengths of interest. Most of these compounds have direct band gaps and, as a consequence, are efficient in the conversion of electrical energy into radiation. With the addition of appropriate chemical impurities, called dopants, both III–V and II–VI compounds can be made *p*- or *n*-type, for the purpose of forming *pn* junctions. All modern-day LEDs contain *pn* junctions. Most of them also have heterostructures, in which the *pn* junctions are surrounded by semiconductor materials with larger band-gap energies. *See* ACCEPTOR ATOM; DONOR ATOM; ELECTROLUMINESCENCE; ELECTRON-HOLE RECOMBINATION; JUNCTION DIODE; JUNCTION TRANSISTOR; LASER; SEMICONDUCTOR; SEMICONDUCTOR DIODE.

Material synthesis. The fabrication of LEDs begins with the preparation of single-crystal substrates, typically 250 to 350 micrometers in thickness, which serve as templates for epitaxial growth and add mechanical strength to the completed device structure. The *p*- and *n*-type minority-carrier injection layers and the photon-generating active regions are formed on top of these substrates by depositing additional material from the liquid or vapor phase (**Fig. 1**). Infrared diodes are made of either gallium indium arsenide layers grown on gallium arsenide substrates or indium gallium arsenide phosphide deposited on indium phosphide. Red-emitting diodes are made of either aluminum gallium arsenide or gallium indium phosphide layers grown on gallium arsenide substrates. Orange- and yellow-emitting diodes are made of aluminum gallium indium phosphide layers deposited on gallium arsenide substrates. Green- and blue-emitting diodes are made of either gallium indium nitride layers grown on sapphire or silicon carbide substrates, or zinc cadmium selenide or zinc tellurium selenide deposited on zinc selenide. LEDs containing aluminum gallium arsenide are made using a process known as liquid-phase epitaxy. All the rest are made using one of three different vapor-based processes; namely, vapor-phase epitaxy, metallorganic chemical vapor deposition, or molecular-beam epitaxy. In the more common III–V LEDs, the *n*-type dopant of choice is either the column IV (or group 14) element silicon, or one of two column VI elements selenium or tellurium. The *p*-type dopant is usually a column II element, for example, zinc or magnesium; however, the column IV element carbon is now a viable alternative. For II–VI LEDs, the *n*-type dopant is the column VII (or group 17) element chlorine, and the *p*-type dopant is the column V element nitrogen. To facilitate electrical access, ohmic contacts are made by evaporating thin metallic films onto both *n*- and *p*-type layers. *See* BAND THEORY OF SOLIDS; CRYSTAL GROWTH; CRYSTAL STRUCTURE; MOLECULAR BEAMS; SEMICONDUCTOR HETEROSTRUCTURES.

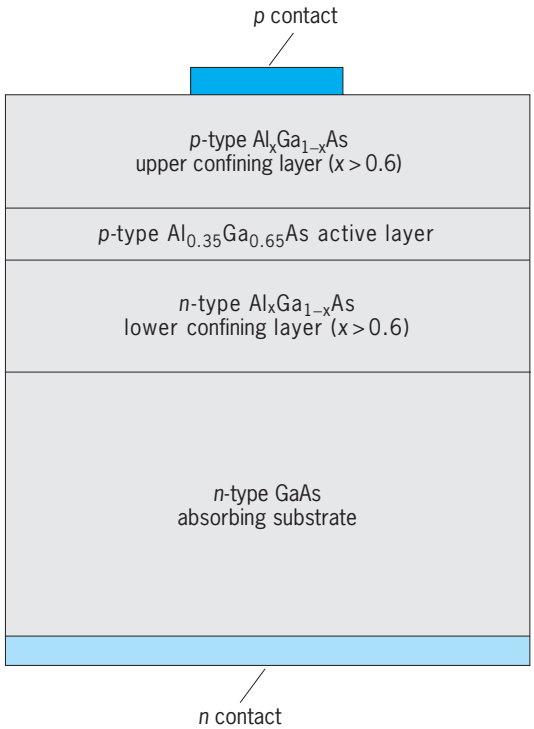

Fig. 1. Cross-sectional view of a typical light-emitting diode structure. The 250–350-μm-thick GaAs substrate serves as a template for epitaxial growth of the active layers and provides mechanical strength to the completed device. The *p*-type and *n*-type $Al_xGa_{1-x}As$ upper and lower confining layers form the *pn* junction that provides the means to inject excess electrons and holes into the active region. These layers are usually 1–2 μm thick. The $Al_{0.35}Ga_{0.65}As$ active region is where the excess electrons and holes recombine to generate photons. This layer is usually 0.1–0.5 μm thick. The *p* and *n* contacts are metal alloys or multilayer composites that are chosen specifically to make low-resistance electrical contacts to the adjacent semiconductor materials. (*Copyright 1993 Hewlett-Packard Company; reproduced with permission*)

Device design. One of the primary objectives of LED design is to maximize external quantum efficiency, defined as the radiant flux (in watts) emitted from the device divided by the product of the photon energy (in electronvolts) of the peak emission and the current (in amperes) flowing through the device. For discussion purposes, it is useful to separate this quantity into two distinct parameters: the internal quantum efficiency and the light extraction coefficient. The internal quantum efficiency (the number of useful photons created per electron-hole pair injected) is limited by the transfer of excess carriers to indirect energy band minima, their leakage

outside the active region, and the trapping of carriers at heterointerfaces—all of which eventually lead to either nonradiative recombination or radiative recombination at undesirable wavelengths. The influence of these excess carrier loss mechanisms can be mitigated by the proper choice of active region materials and layer thicknesses. In the ideal case of a well-developed semiconductor material, combined with an optimized structure, the internal quantum efficiency may approach unity. The light extraction coefficient (or the fraction of photons that escape from the device) is constrained by total internal reflection at the semiconductor/air interface. Light generated by spontaneous emission at any point within the active region is uniformly distributed in all directions. Given the large difference in the refractive indices between semiconductor and air, only a small fraction of light impinges upon the interface at angles less than the angle for total internal reflection. This component, which escapes by simple refraction on its first pass through the device, is of the order of 4% for each of the six escape cones inscribed within the semiconductor parallelepiped (**Fig. 2**). Hence, even without any additional obstructions, the light extraction coefficient has a maximum value of only 24%. With normal epitaxial layer thicknesses and device cross-sectional area, however, almost all of the light emitted toward the four side walls is reflected off the top surface or absorbed in the substrate. Most of the light emitted directly upward toward the top surface is reflected back into the substrate or absorbed within the opaque metal contact pad. All of the light emitted downward toward the substrate is lost to absorption before reaching the back surface. Thus, in practice, a typical red-light-emitting diode with near-unity internal quantum efficiency has an external quantum efficiency of less than 5%. *See* LUMINESCENCE; REFLECTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES.

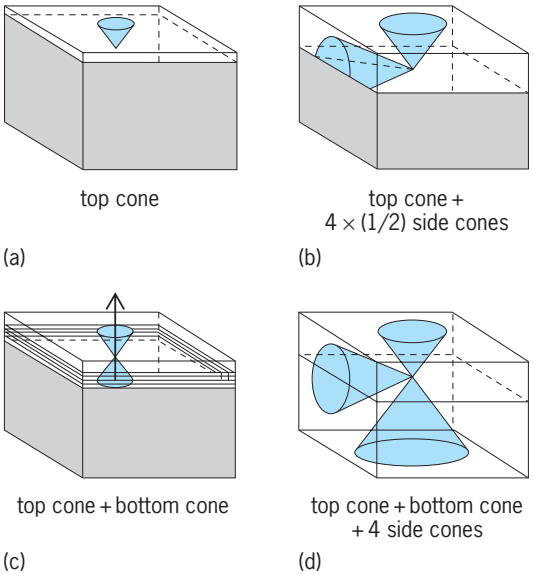

Fig. 2. Schematic diagram showing the means by which photons created at a point source inside a semiconductor parallelepiped escape to the outside environment. Photons incident on the semiconductor/air interface at angles larger than the critical angle are trapped by total internal reflection. The cones represent the solid angles over which photons can escape from the semiconductor after suffering only small Fresnel reflection losses. (*a*) Light escapes only from the top surface because the substrate is highly absorbing at the photon wavelength and the active region is too thin to allow escape from the sample edges. (*b*) Light escapes from both the top and bottom cones. A distributed Bragg mirror was inserted between the active region and the absorbing substrate so that photons emitted into the bottom cone would be reflected back toward the top surface. (*c*) Light escapes from the top surface and the upper halves of the four side cones because the active region was made thick enough that photons could reach the side walls of the parallelepiped before being trapped by total internal reflection off the top surface. (*d*) Light escapes from all six cones. The absorbing substrate was replaced by a transparent material, and thus photons emitted into the bottom cone and into the bottom halves of the four side cones can escape.

Applications. Conventional low-power, visible LEDs are used as solid-state indicator lights in instrument panels, telephone dials, cameras, appliances, dashboards, and computer terminals, and as light sources for numeric and alphanumeric displays. Modern high-brightness, visible LED lamps are used in outdoor applications such as traffic signals, changeable message signs, large-area video displays, and automotive exterior lighting. General-purpose white lighting and multielement array printers are applications in which high-power visible LEDs may soon displace present-day technology. Infrared LEDs, when combined in a hybrid package with solid-state photodetectors, provide a unique electrically isolated optical interface in electronic circuits. Infrared LEDs are also used in optical-fiber communication systems as a low-cost, high-reliability alternative to semiconductor lasers.

Indicators and displays. The advantages of using LEDs to produce visible electromagnetic radiation are their high efficiency, small size, ruggedness, long life, and compatibility with silicon integrated circuits. The most appropriate measure of electrical-to-optical power conversion for display applications is luminous efficiency, defined as the luminous flux (in lumens) divided by the product of applied voltage and drive current (in watts). The luminous flux emitted from an LED is determined by measuring its corresponding radiometric quantity, that is, the radiant flux, and correcting it to account for the wavelength-dependent response of the eye. LEDs with power conversion efficiencies greater than 10 lumens per watt are now commercially available at discrete wavelengths covering the entire visible spectrum. Individual red-, green-, and blue-light-emitting diodes can be combined to produce an enormous variety of intermediate hues, and specially designed ultraviolet diodes can be used to excite an inorganic fluorescent coating for the purpose of generating white light (**Fig. 3**). Within the spectral window from 600 to 640 nm, the luminous efficiency of LEDs exceeds that of tungsten and halogen lamps and, in some laboratory demonstrations, approaches that of fluorescent lighting. This high level of performance is the consequence of many improvements in material quality and device design, particularly in the area of light extraction. Special current spreading layers, distributed

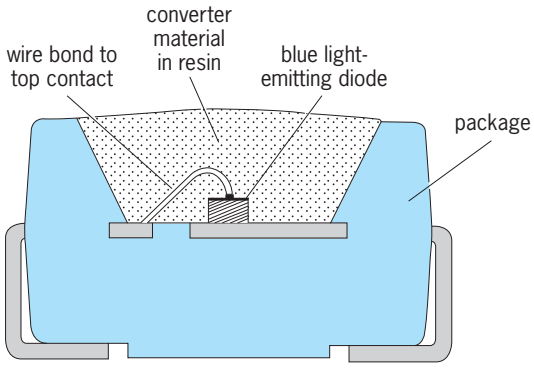

Fig. 3. Cross-sectional view of a light-emitting diode package designed to emit white light by combining blue, green, and red photons. The source of blue photons is the direct light emission from a AlGaN-GaInN double heterostructure LED chip. The sources of green and red photons are two distinct inorganic fluorescent materials embedded in the epoxy resin that encapsulates the LED chip. The green and red photons are generated by an energy down-conversion process; that is, blue photons emitted by the semiconductor LED are absorbed by the fluorescent materials, the excited electronic carriers lose energy through various internal mechanisms and drop down to an intermediate energy state, and then their electronic energy is converted back into photons by the process of flourescence. The intensity of the blue photon source is controlled by the choice of drive current passing through the semiconductor LED chip. The intensity of the green and red photon sources is set by adjusting the composition of the fluorescent materials in the epoxy resin.

Bragg reflectors, and wafer-bonded transparent substrates have been developed to move the light extraction coefficient closer to its theoretical limit (**Fig. 4**). Another important operating parameter is maximum light output power. The simplest way to increase the luminous flux from an LED is to raise its drive current. However, this is accompanied by an increase in Joule heating in the passive regions of the device, leading to a higher *pn* junction operating temperature. Thus, the output power of an LED lamp is limited by heat dissipation within the semiconductor chip, since an increase in junction temperature causes a reduction in internal quantum efficiency and shifts the emitted radiation to longer wavelengths. A luminous output as high as 20 lumens at 619 nm is sustainable by using a special high-power lamp package to improve thermal management. *See* ILLUMINATION; LUMINOUS EFFICIENCY; LUMINOUS FLUX; PHOTOMETRY.

Optoisolators. In the simplest optical interface, the optoisolator, an LED, and a photodetector are optically coupled, but electrically isolated, in a small package. This device can be used, for example, at the interface between two different circuits, such as the switching equipment in a telephone central office and the connecting loop circuit which carries the signals to the telephone sets. The electric signal from the central office is converted to radiation by the LED, which in turn is converted back into an electric signal by the photodetector before it enters the loop circuit. This type of interface is traditionally provided by electromechanical relays or isolation transformers. The electrical isolation resulting from the optical path protects the central office from electromagnetic interferences such as lightning which hits telephone wires or surge currents from electromechanical relays. LEDs are ideal for this application because they are very reliable and can be modulated to carry high-frequency signals. In a typical optoisolator structure, the LED and the phototransistor are mounted on separate metal lead frames, and the two components are coupled optically through a transparent plastic encapsulant (**Fig. 5**). This plastic is also the source of electrical isolation, typically on the order of 2500 V. Final encapsulation is completed with a black, opaque overmold which also provides mechanical stability. Optoisolators are compatible with silicon integrated circuits in size, reliability, and

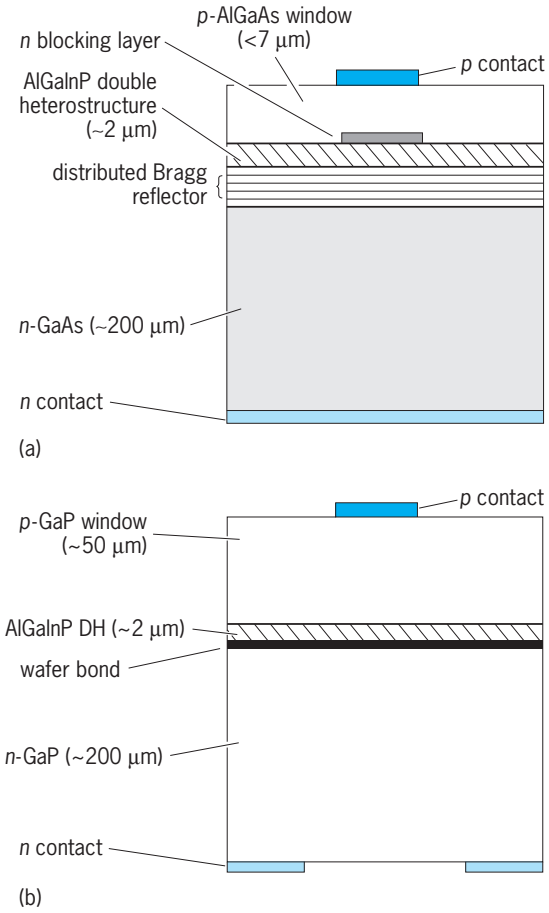

Fig. 4. Cross-sectional view of two different designs for maximizing the luminous efficiency of AlGaInP light-emitting diode structures. (*a*) A thick *p*-type AlGaAs layer between the AlGaInP active region and the top-side *p*-contact serves to spread the current out well beyond the cross-sectional area of the metal pad, combined with an *n*-type semiconductor blocking layer located directly adjacent to the active region that prevents the flow of current through the cross-sectional area directly shadowed by the opaque metal pad. A distributed Bragg mirror was inserted between the AlGaInP active region and the absorbing GaAs substrate to reflect light back up toward the top surface. (*b*) An extremely thick GaP window serves as both a current-spreading layer and a light extraction layer. The GaP window is thick enough that photons emitted within the four side cones can escape before reaching the top surface and being trapped by total internal reflection. The absorbing GaAs substrate that served as the original growth template for the AlGaInP active region was removed, and a transparent GaP substrate was attached to the bottom side of the active region using a wafer-bonding method.

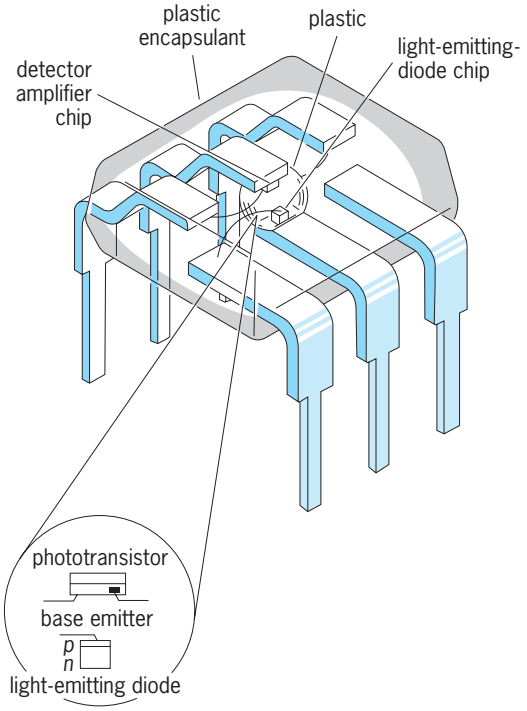

Fig. 5. Typical opto-isolator in a plastic-encapsulated dual in-line package.

performance parameters, giving them a prominent role in modern solid-state circuits. *See* OPTICAL ISOLATOR.

Optical fiber transmission. Another rapidly evolving application of infrared LEDs is in optical fiber communication. An optical signal is fed into a 50–100-μm-diameter glass fiber and transmitted over long distances without the need for regeneration. At the other end of the fiber, a photodetector converts the optical signal back to an electrical signal. This silica-glass fiber, which replaces coaxial cable, is smaller in volume, less expensive, and immune to electromagnetic interference. It also transmits higher data rates and provides longer repeater spacings than metal conductors. Since doped silica fibers exhibit both low loss and minimum material dispersion near 1.3 μm, LED sources at this wavelength can achieve a repeater spacing of 10 km (6 mi) at a data rate of 250 megabits per second. The key operating parameters for this application are fiber-coupled power and response time. Coupling infrared radiation into an optical fiber is difficult because the fiber has a very narrow acceptance angle (10–150°). This task is accomplished by extracting light through a well etched into the bottom side of the substrate with the fiber held in place by a special epoxy resin. A dielectric layer is deposited onto the top side of the structure to provide electrical isolation between the ohmic contact metal and the heavily doped semiconductor surface, thereby limiting the active area of the junction to the size of the optical fiber. It is also necessary to shorten the response time of the LED for high data-rate transmission. This goal can be realized by increasing the doping level in the active region with the intention of decreasing the radiative recombination lifetime. The challenge is to achieve the highest doping level possible without inadvertently introducing nonradiative recombination centers that act to reduce the internal quantum efficiency of the device. *See* OPTICAL COMMUNICATIONS; OPTICAL FIBERS. Jerry M. Woodall; Louis J. Guido

Bibliography. A. A. Bergh and P. J. Dean, *Light Emitting Diodes*, 1976; G. Bogner et al., *Compound Semiconductor*, 5(4):28, 1999; S. E. Miller and I. P. Kaminow (eds.), *Optical Fiber Telecommunications*, vol. 2, 1988; G. B. Stringfellow and M. George Craford (eds.), *Semiconductors and Semimetals*, vol. 48: *High Brightness Light Emitting Diodes*, 1997.

Light panel

A surface-area light source that employs the principle of electroluminescence to produce light. Light panels are composed of two sheets of electrically conductive material, one a thin conducting backing and the other a transparent conductive film, placed on opposite sides of a plastic or ceramic sheet impregnated with a phosphor, such as zinc sulfide, and small amounts of compounds of copper or manganese. When an alternating voltage is applied to the conductive sheets, an electric field is applied to the phosphor. Each time the electric field changes, it dislodges electrons from the edges of the phosphor crystals. As these electrons fall back to their normal atomic state, they affect the atoms of the slight "impurities" of copper or manganese, and radiation of the wavelength of light is emitted. *See* ELECTROLUMINESCENCE.

In contrast to incandescent, vapor-discharge, and fluorescent lamps, which are essentially point or line sources of light, the electroluminescent light panel is essentially a surface source of light. Complete freedom of size and shape is a fascinating aspect of luminescent cells (see **illus.**).

Brightness of the panel depends upon the voltage applied to the phosphor layer and upon the electrical frequency. In general, higher voltage and higher frequency both result in a brighter panel. Blue, green, red, or yellow light can be produced by the choice of phosphors, and the proper blend of these colors

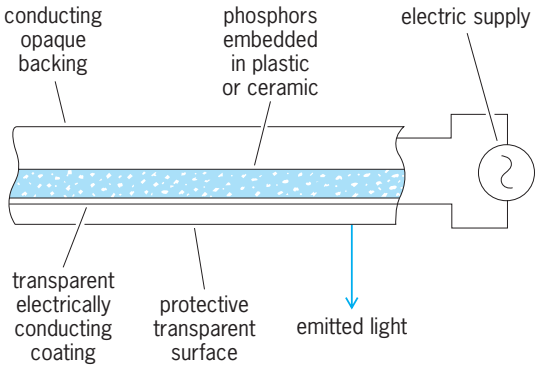

Simplified diagram of an electroluminescent cell; the sketch is not drawn to scale.

produces white light. Color can be varied for a particular phosphor by changing the frequency of the applied voltage. Increasing the frequency shifts the color toward the blue end of the spectrum.

The efficiency of these light panels is only a fraction of that of the most efficient fluorescent lamps. Theoretical limits indicate, however, that the efficiency can be further improved, probably to exceed that for fluorescent lamps. Because panel lights employ no filaments and no evacuated or gas-filled bulbs, replacement of units is virtually eliminated. Glareless uniform distribution of light from large-area sources is possible without shades, louvers, or other control devices. *See* FLUORESCENT LAMP; ILLUMINATION.

Warren B. Boast

Light-scattering techniques

A wide variety of analytic techniques in which the distribution and precise nature of the scattered light carry information about the object from which it was scattered. Light-scattering techniques are used for studying and characterizing particles ranging from nanometer to millimeter sizes.

Whenever light traveling through a medium encounters a change in optical properties (for example, meeting a particle suspended in air), some of the incident light energy is redirected from its original path. This is light scattering. If the wavelength (color) of the scattered light is the same as that of the incident light, the phenomenon is referred to as elastic light scattering; if there is a change in wavelength caused by the particle's motion relative to the incident light or its material properties, so-called quasi-elastic or inelastic light scattering may result. Light-scattering techniques are now found in virtually every branch of science and engineering, from polymer chemistry, pharmaceutical biochemistry, and clinical science to food processing, meteorology, and environmental pollution monitoring. Some of the principal light-scattering techniques used are described below.

Light is that part of the electromagnetic spectrum to which our eyes are sensitive, approximately 400–700 nm (blue to deep red) in wavelength (λ). Light-scattering techniques normally use wavelengths within this range, although the term is often loosely applied when shorter (ultraviolet) or longer (infrared) wavelengths are used. In such cases, the principles of the underlying science remain essentially the same. *See* ELECTROMAGNETIC RADIATION; LIGHT.

Scattering. Classically, the interaction of light with a particle may be considered in terms of the electromagnetic wave nature of light and the electrical charges bound within the particle. The oscillating electric field of the light wave (in the plane perpendicular to the wave's direction) causes the bound electrical charges to oscillate too, and so reradiate secondary electromagnetic waves in all directions (**Fig. 1**). Two arbitrary bound charges at positions A and B within the particle will radiate secondary waves as shown. In certain directions, these waves will be in phase and will constructively interfere, producing a higher intensity; in other directions, they will be out of phase and will tend to cancel out each other (destructive interference). The directions of constructive and destructive interference depend on both the relative positions of the charges and how they are bound within the molecules of the particle. When all bound charges are considered, the unique pattern of scattered light around the particle will depend on the size, shape, and material nature of the particle, as well as the wavelength and polarization state of the incident light (that is, how the electric field oscillations vary in time). It is this unique optical "signature" that is exploited in particle light-scattering techniques. In some cases, light energy may be absorbed by the particle and reemitted at a different wavelength some time later. These processes give rise to phenomena such as fluorescence and phosphorescence, and a more complex quantum-mechanical description is required for full understanding. *See* INTERFERENCE OF WAVES.

If the particle is much smaller than the incident wavelength, all of the bound charges will oscillate essentially in phase and the scattered energy will be virtually the same in forward and backward directions. This is referred to as Rayleigh scattering. When the particle size is the same order as the wavelength of the light, the pattern of scattered light becomes more complex and the energy scattered in the forward direction dominates. This is referred to as Mie scattering. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

Light-scattering techniques can be broadly divided into two categories: those that investigate individual particles in an attempt to characterize or classify them, and those that examine a large number of particles simultaneously in order to determine properties of the whole population. There are variants of both techniques, each of which fulfils an important analytic role. In the majority of cases, the preferred light source is a laser since this provides intense, monochromatic, and coherent (required in some applications) radiation. *See* LASER.

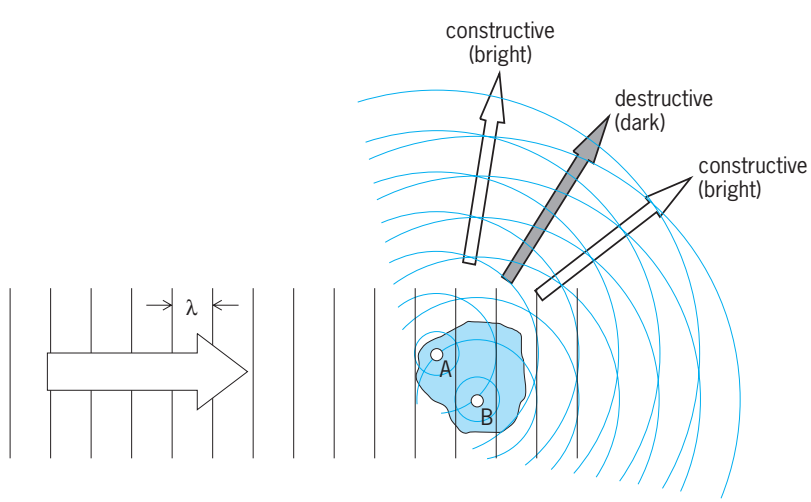

Fig. 1. Simplified interpretation of light scattering by a particle.

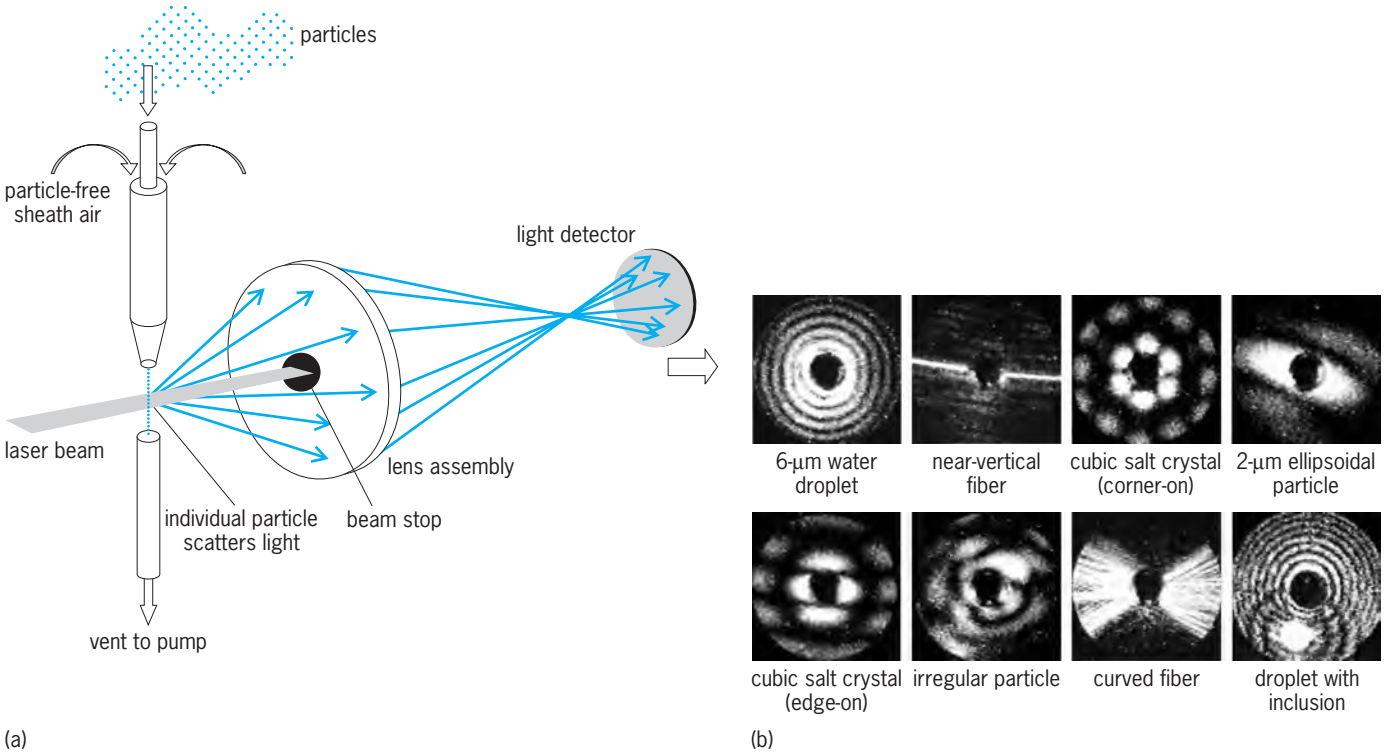

Fig. 2. Single-particle light scattering. (*a*) Apparatus. (*b*) Typical scattering patterns from various particle types.

Single-particle scattering. To achieve single-particle scattering from particles in air or liquid suspension, it is necessary to control the movement of the particles so that they pass rapidly through the illuminating light beam in single file and produce pulses of scattered light that can be detected and recorded. Often, an arrangement is used in which particles are constrained to flow through a laser beam by ensheathing the particle-laden air column with a layer of particle-free air (**Fig. 2***a*).

Optical particle counting. In the simplest case, light scattered by each particle may be recorded by a single optical detector that integrates the light signal across a large, solid angle of scattering. Such optical particle counters (OPCs) are an important class of instrument used in areas such as environmental or industrial monitoring, where knowledge of the number concentration and size distribution of airborne particles is important from a health or process-control point of view. Particles from ~0.3 to 100 μm may be measured, although sizing data must be treated with caution since the magnitude of scattered light falling on the detector will be a function of particle shape and material (refractive index) as well as absolute size. *See* OPTICAL DETECTORS.

Spatial light scattering. In more sophisticated instruments, the single detector is replaced by a detector array to record more detail of the complex pattern of scattered light (Fig. 2*b*). Such spatial light-scattering techniques can allow assessment of individual particle shape and structure as well as size, and are especially useful in areas where the classification of particles into different categories is required. For example, aircraft-borne instruments allow the classification of cloud ice crystals and water droplets (data important for meteorological climate-change research), while in clinical fields instruments can be used to sort various types of liquid-borne cells. Spatial light-scattering techniques are also of importance in monitoring applications where the detection of potentially low concentrations of hazardous particles is required in an ambient atmosphere. Examples include the detection of airborne asbestos fibers around demolition sites and the detection of airborne biological particles in environments susceptible to

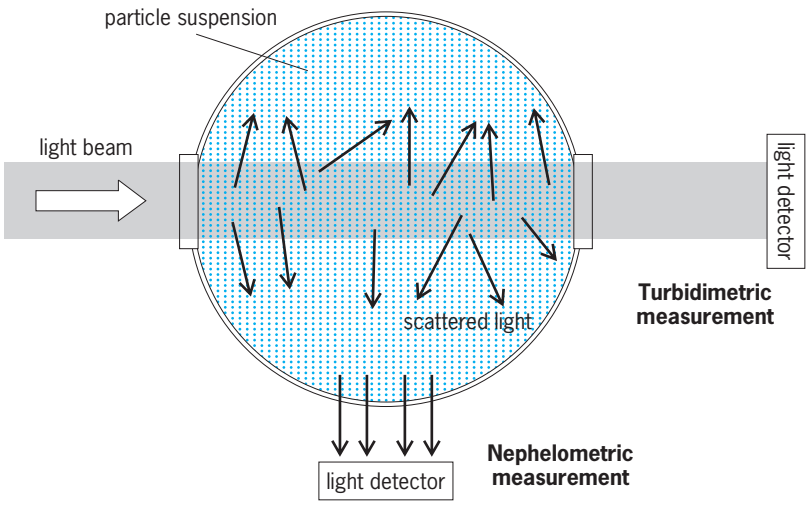

Fig. 3. Light-scatter measurement from particle suspensions.

terrorist attacks. In the latter case, measurements of other optical phenomena, such as particle fluorescence, are often used in parallel with light scattering to improve the confidence of particle classification.

Ensemble particle scattering. The most widely used light-scattering techniques, primarily for reasons of simplicity of implementation and ease of use, are those based on measurement of the scattering properties of large collections or ensembles of particles. In such methods, one or more detectors record the light intensity scattered simultaneously from possibly millions of individual particles, giving rise to average properties of the particle suspension (**Fig. 3**). While such techniques are used in fields such as meteorology and pollution monitoring for visibility assessment, they are most frequently applied to the quantitative measurement of powders, particle-liquid suspensions, and macromolecular suspensions associated with numerous industrial, chemical, or biochemical processes.

Turbidimetry and nephelometry. Measurement of the decrease in intensity of a beam of light resulting from its passage through a suspension of particles is referred to as turbidimetry. This simple technique provides a measure of the size and concentration of the particles, and is widely used in clinical chemistry and polymer science. In a related technique, nephelometry, the detector is placed at an angle, usually 90°, to the beam axis to measure the scattered light from the suspension. For very dilute suspensions or those containing particles much smaller than the wavelength of the light, even down to macromolecular sizes, the relative ease of measuring low scattered-light intensities against a dark background makes nephelometry an extremely sensitive technique. Measuring the change of light-scatter intensity over time, so-called rate nephelometry, allows chemical or biochemical reactions to be followed through changes in particle concentration or size. *See* IMMUNONEPHELOMETRY; POLYMER.

Laser diffraction. In the technique of laser diffraction, the forward-scattered light from a suspension of airborne or liquid-borne particles is measured using a multielement detector, usually comprising concentric annular detector rings. Since proportionally more of the scattered light from the smallest particles is scattered at higher angles while larger particles scatter predominantly at lower forward angles, such a detector arrangement can be used to assess the distribution of particle sizes within the suspension. The technique is applicable to particles from below ~100 nm up to several millimeters in size, depending on the range of scattering angles covered by the instrument and the optical geometry used. Being a noninvasive and essentially real-time technique, it has found application in powder and liquid-particle suspension measuring/monitoring applications in the production of foodstuffs, paints, ceramics, and pharmaceuticals.

Dynamic light scattering. The techniques described above measure the time-averaged intensity of scattered light and are often referred to as total-intensity scattering or static light scattering. For suspensions of particles or macromolecules that are extremely small compared with the wavelength of the light, the time-dependent fluctuations in the scattered light arising from the random Brownian motion of the particles can be exploited to assess particle size. This technique of dynamic light scattering (also known as quasi-elastic light scattering or photon correlation spectroscopy) requires the incident light to be coherent and monochromatic, and has benefited hugely from the development of lasers as well as from advances in computer power. By measuring their diffusion behavior, it allows routine sizing of the particles or macromolecules down to a few nanometers in size, and is a valued technique in numerous research fields and manufacturing processes.

Paul Kaye

Bibliography. B. J. Berne and R. Pecora, *Dynamic Light Scattering with Applications to Chemistry, Biology, and Physics*, Courier Dover Publications, Toronto, 2000; C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, Wiley Interscience, 1998; W. Brown, *Light Scattering: Principles and Development*, Oxford University Press, 1996; R. Xu, *Particle Characterization: Light Scattering Methods*, Kluwer Academic, 2000.

Light-year

A unit of measurement of astronomical distance. A light-year is the distance light travels in 1 sidereal year. One light-year is equivalent to 9.461×10^{12} km, or 5.879×10^{12} mi. Distances to some of the nearer celestial objects, measured in units of light time, are shown in the **table**.

Distances from the Earth to some celestial objects

Object	Distance from Earth (in light time)
Moon (mean)	1.3 s
Sun (mean)	8.3 min
Mars (closest)	3.1 min
Jupiter (closest)	33 min
Pluto (closest)	5.3 h
Nearest star (Proxima Centauri)	4.3 years
Andromeda Galaxy (M31)	2.3×10^6 years
Typical quasar	3×10^9 years
Epoch of recombination	13.7×10^9 years

This unit, while useful for its graphic presentation of the enormous scale of stellar distances, is seldom used technically except in cosmology. *See* ASTRONOMICAL UNIT; COSMOLOGY; PARALLAX (ASTRONOMY); PARSEC.

Jesse L. Greenstein; Jay M. Pasachoff

Bibliography. K. Ferguson, *Measuring the Universe*, Walker, New York, 1999; M. Rowan-Robinson, *Cosmology*, 4th ed., Oxford University Press, 2004.

Lighthouse

A distinctive structure, built on or near a shore, which exhibits a light of distinctive characteristics to serve as an aid to navigation (see **illus.**). Lesser lights may be displayed from fixed structures called beacons or from floating buoys or lightships.

Lighthouse overlooking sea from a coastal headland. (*U.S. Coast Guard*)

The characteristics of the lights displayed by lighthouses are given in light lists available to mariners and, in abbreviated form, on charts. Some lights have one or more sectors in which the light appears red, usually to warn of some danger in this sector. In other sectors most lights are white.

Lighthouses have been diverse in structure and type of light. Towers up to 200 ft (60 m) were constructed along the Mediterranean coast of Egypt many centuries before Christ, with beacon fires maintained by priests. Logs, coal, oil, gas, and electricity have been used to provide lights. An attendant is continuously on duty at many lighthouses, but some are unattended, and some modern installations are controlled remotely from a convenient location. *See* PILOTING. Alton B Moody

Bibliography. N. Bowditch, *American Practical Navigator, An Epitome of Navigation, 1995 edition,*, 1995; U.S. Coast Guard, *Aids to Navigation Manual*, CG-222, 1982.

Lightning

An abrupt, high-current electric discharge that occurs in the atmospheres of the Earth and other planets and that has a path length ranging from hundreds of feet to tens of miles. Lightning occurs in thunderstorms because vertical air motions and interactions between cloud particles cause a separation of positive and negative charges. *See* ATMOSPHERIC ELECTRICITY.

Lightning is both a fascinating weather phenomenon and a severe hazard. The vast majority of flashes are produced by thunderclouds, and most discharges remain within the cloud or are from cloud to air or from cloud to cloud. About one-third of all discharges are from cloud to ground, and this type of lightning is the primary hazard to people or objects on the ground. Lightning can also be initiated artificially (triggered) by tall structures when a thunderstorm is nearby, or by aircraft or rockets when they fly into a highly electrified environment. Triggered lightning can occur in clouds that otherwise do not produce natural lightning; hence this type of discharge is frequently unexpected.

Lightning between cloud and ground. The vast majority of lightning flashes between cloud and ground begin in the cloud with a process known as the preliminary breakdown. After perhaps a tenth of a second, a highly branched discharge, the stepped leader, appears below the cloud base and propagates downward in a succession of intermittent steps. The leader channel is usually negatively charged, and when the tip of a branch of the leader gets to within about 30 m (100 ft) of the ground, the electric field becomes large enough to initiate one or more upward connecting discharges, usually from the tallest objects in the local vicinity of the leader. When contact occurs between an upward discharge and the stepped leader, the first return stroke begins. The return stroke is basically a very intense, positive wave of ionization that propagates up the partially ionized leader channel into the cloud at a speed close to the speed of light. After a pause of 40–80 milliseconds, another leader, the dart leader, forms in the cloud and propagates down the previous return-stroke channel without stepping. When the dart leader makes contact with the ground, a subsequent return stroke propagates back to the cloud. A typical cloud-to-ground flash lasts 0.2–0.3 s and contains about four return strokes; lightning often appears to flicker because the human eye is capable of just resolving the interval between these strokes. In roughly one-third of all flashes to ground, the dart leader propagates down just a portion of the previous return-stroke channel and then forges a different path to ground. In these cases, the discharge actually strikes the ground in two places, and the channel has a characteristic forked appearance that has been frequently photographed (**Fig. 1**).

Lightning between cloud and ground is usually classified according to the direction of propagation and polarity of the initial leader. For example, in the most frequent type of cloud-to-ground lightning a negative discharge is initiated by a downward propagating leader as described above (**Fig. 2***a*). In this case, the total discharge will effectively lower negative charge to ground or, equivalently, will deposit positive charge in the cloud. In a photograph, the direction of the branches is the same as the direction of leader propagation.

A discharge can be initiated by a downward-propagating positive leader (Fig. 2*b*). Positive discharges occur less frequently than negative ones, usually just a few percent of the total number of

Fig. 1. Natural cloud-to-ground lightning near Tucson, Arizona. Note that each discharge strikes the ground in more than one place. (*Courtesy of Michael J. Leuthold*)

flashes, but positive discharges are often quite deleterious. Another type of lightning is a ground-to-cloud discharge that begins with a positive leader propagating upward (Fig. 2*c*); this type is relatively rare and is usually initiated by a tall structure or a mountain peak. The rarest form of lightning is a discharge that begins with a negative leader propagating upward (Fig. 2*d*).

The electric currents that flow in return strokes have been measured during direct strikes to instrumented towers. The peak current in a negative first stroke (Fig. 2*a*) is typically 30 kiloamperes, with a zero-to-peak rise time of just a few microseconds. This current decreases to half-peak value in about 50 microseconds, and then low-level currents of hundreds of amperes may flow for a few to hundreds of milliseconds. The long-continuing currents produce charge transfers on the order of tens of coulombs and are frequently the cause of fires. Subsequent return strokes have peak currents that are typically 10–15 kA, and somewhat faster current rise times. The maximum rate of change of current during the initial rise of both first and subsequent return strokes is on the order of 100 kA/μs (10^{11} A/s). Five percent of the negative discharges to ground generate peak currents that exceed 80 kA, and 5% of the positive discharges (Fig. 2*b*) exceed 250 kA. Positive flashes frequently produce very large charge transfers, with 50% exceeding 80 coulombs and 5% exceeding 350 coulombs.

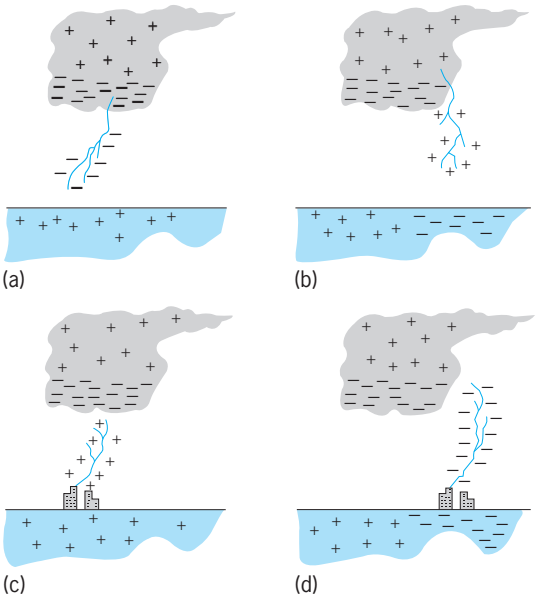

Fig. 2. Sketches of the different types of lightning between an idealized cloud and the ground: (*a*) type 1, (*b*) type 2, (*c*) type 3, and (*d*) type 4. Channel development within the cloud is not shown. Type 1 is the most common form of cloud-to-ground lightning, and type 4 is very rare. (*After M. A. Uman, The Lightning Discharge, Academic Press, 1987*)

The rapidly rising return-stroke current heats the channel to a peak temperature near 55,000°F (30,000°C). This rapid heating produces high pressures that cause sections of the channel to expand behind shock waves that eventually become thunder. The peak light output from a 1-m (3.3-ft) vertical section of channel is over 10^6 W, and the peak light output from the entire stroke is typically 10^9 W or more. Occasionally, flashes produce a light output that exceeds 10^{11} W, and these are sometimes known as superbolts. *See* CLOUD PHYSICS; SHOCK WAVE.

Electromagnetic fields. Electric and magnetic fields that are radiated by different lightning processes have different but characteristic signatures that are reproduced from flash to flash. A typical cloud-to-ground flash at a distance of about 60 km (37 mi) can be characterized by a cloud pulse that was radiated during the preliminary breakdown, the waveform that was produced by the first return stroke, and the signature from a subsequent return stroke.

A peak field of 8 V/m at 100 km (60 mi) means that the peak electromagnetic power that is radiated by the stroke exceeds 10^{10} W; therefore, lightning is easily detected at very large distances. It also means that if such a pulse propagates upward into the middle atmosphere, it will be large enough to heat the electrons and cause regions of enhanced ionization in the lower ionosphere. Such phenomena have occurred in the United States and New Zealand. *See* ELECTROMAGNETIC RADIATION; GEOELECTRICITY; IONOSPHERE.

Triggered lightning. Mountains and tall buildings can sometimes initiate upward, ground-to-cloud lightning. Aircraft experiments have also shown that the majority of lightning strikes to airborne vehicles actually begin at, and hence are triggered by, the airframe.

When a thunderstorm is overhead and the electrical conditions are favorable, lightning can be triggered artificially by using small rockets to carry a grounded wire aloft. If the rocket is fired when the surface electric field is high (3–5 kV/m), about 70%

of all launches will trigger a discharge. Most triggers occur when the rocket is at an altitude of only 100–300 m (330–990 ft), and usually the first stroke in the flash propagates upward into the cloud and is not like natural lightning. The majority of the subsequent strokes, however, follow the first stroke to ground and are almost identical to their natural counterparts.

Triggered lightning is important for research because it provides the capability of studying both the physics of the discharge and its interactions with structures in a partially controlled environment. It is being used to investigate the luminous development of lightning channels, the waveforms of lightning currents, the velocities of return strokes, the relation between currents and fields, the mechanisms of lightning damage, the performance of lightning protection systems, and many other problems. Among the more important findings have been the discovery of submicrosecond fields and currents during return strokes and the general validity of the transmission line model equation. Submicrosecond current variations produce large voltages when lightning strikes an inductive load or a complex system, such as an airplane, that is sensitive to high-frequency disturbances. Submicrosecond currents also place more stringent requirements on the performance of surge arresters and other devices that are used in lightning protection, and also on the engineering tests and test standards that are used to verify that performance. *See* STORM ELECTRICITY; THUNDERSTORM.

E. Philip Krider

Detection. There are several techniques for detecting lightning: electric field amplitude, magnetic radiation field direction finding, electric and magnetic radiation field time of arrival and interferometers, radar, visible light detection, and thunder time of arrival. The electric field amplitude (flash counter) techniques are limited to a range of a few tens of kilometers; they are used primarily in lightning research for studies in small geographical regions. Magnetic field direction finding is the basis for one of two methods of detecting lightning flashes to ground. If two azimuths or directions are known to a lightning flash, the location of the flash to ground is determined by triangulation.

Radiation field time-of-arrival systems are operating over the United States and are effective in locating lightning when the precise relative time is known at four or more stations. The combination of magnetic direction finding and radiation time-of-arrival systems makes up the National Lightning Detection Network. Radar techniques depend on detecting the reflected energy from a lightning channel at wavelengths of 10 cm or longer, and the resulting pattern can produce a pattern of the channel in a cloud. Techniques using visible light for lightning location depend on instruments such as video cameras to record the image of lightning. The thunder time-of-arrival technique depends on the accurate recording of the arrival time of thunder following the time of the visible channel. Since the average speed of sound in the atmosphere is known, the distance to the channel is estimated by allowing 3 s for every kilometer (5 s for every mile) between the observer and the lightning. Of the various lightning-detection methods, the combined use of the magnetic direction finding and time-of-arrival methods provides the most effective technique for locating flashes to ground.

The U.S. National Lightning Detection Network is composed of over one hundred sensors strategically placed throughout the contiguous United States; it provides lightning locations to over 400 users, including universities, electrical power companies, telephone corporations, airlines, and federal agencies including the National Weather Service and Defense Agencies. The National Network was originally established as a large regional lightning detection network in the west to provide early warning of forest fires. The sensors record only lightning flashes to ground, and exclude the more frequent and usually harmless (for ground-based objects) intracloud lightning. Subsequent regional lightning detection networks were established in the Midwest and the East to study the meteorological and physical characteristics of lightning to ground.

A satellite detector launched in April 1995 is in a nearly circular orbit of 740 km (460 mi) with an inclination of 70°. From this altitude, the satellite has a field of view of 1300 by 1300 km (780 × 780 mi). Results from summer observations indicate that the lightning flash rate on Earth is approximately 40 flashes per second. A second satellite with a lightning sensor was launched in November 1997; it also contains a radar to measure rainfall from space. The lightning sensor is a small solid-state camera with special filters that transmit only the peak optical radiation emitted by lightning at 777.4 nanometers. The camera view is 600 km (370 mi) on a side, and each pixel in the focal plane covers 5–10 km (3–6 mi) on the ground, matching the smallest storm cells that might be producing lightning. Researchers are now overlaying these lightning locations onto visible, thermal, and radar images to learn more about storms around the globe. *See* METEOROLOGICAL SATELLITES.

National Lightning Detection Network operation. The National Lightning Detection Network contains sensors that detect the electric and magnetic radiation field and uses the magnetic direction finding method. The magnetic direction finders consist of two orthogonal magnetic-loop antennas, an electric field antenna, associated electronics to process the incoming signals, and an antenna to communicate via satellite with the National Lightning Detection Network operations center in Tucson, Arizona. The bandwidths of the antenna systems receiving the lightning signals are wide (approximately 1–380 kHz), so that the shapes and polarities of the radiation field waveforms are preserved. The orthogonal magnetic-loop antennas sense the magnetic field from lightning, and a voltage is induced in each loop that is proportional to the magnetic field derivative multiplied by the cosine of the angle between the plane of the loop and the direction of propagation of the incoming field. The ratio of the integrated voltages in the orthogonal

loops at the time of the peak voltage field provides the direction to the lightning flash. A 180° ambiguity in direction is removed by the electric field antenna that senses the polarity of the field (and the polarity of the lightning current). The processing electronics is designed to respond to only those waveforms that are characteristic of return strokes in cloud-to-ground flashes.

Information from all the sensors detecting the lightning is sent via satellite to the operations center. When a lightning flash is detected by a sensor, the time, angle, signal amplitude, and polarity of all strokes are stored until the information is electronically requested by a computer in the operations center, typically every 10 s. When two or more sensors detect a flash within a programmed time interval, the location of the flash is computed and plotted using the time-of-arrival and direction finding information. All the sensor information is used to plot the optimal location of the flash.

The sensors in the National Lightning Detection Network use magnetic direction finding and time-of-arrival techniques. The detection efficiency of the network is estimated to be approximately 80–90%, with an accuracy of approximately 500 m (over 1640 ft).

Positive and negative flashes. Although lightning to ground normally lowers negative charge (negative flashes), a small but significant number of flashes involve the lowering of positive charge to ground (positive flashes). Consequently, study of ground flashes must separate the characteristics of these two kinds of lightning. Negative flashes are known to have a median peak current in first return strokes of 30 kA, but positive flashes have a median peak current that is approximately 30% higher. Fortunately, the occurrence of positive lightning to ground is a small percentage of the total number of flashes, typically 10% for the entire year. In some regions of the United States, however, the percentage of positive lightning may exceed 25%.

An annual summary masks some of the interesting polarity variations that have been discovered. For example, the percentage of positive lightning flashes has been shown to vary with season, latitude, and storm. At higher latitudes, characteristic of New England and the upper Midwest, the percentage of positive lightning flashes in winter storms normally exceeds 50% and may be 100% in periods of a storm. At lower latitudes, characterized by Florida, the percentage of positive flashes rarely exceeds 20%. *See* LIGHTNING AND SURGE PROTECTION; STORM DETECTION.

Sprites, elves, and jets. Red sprites, elves, and blue jets are upper atmospheric optical phenomena associated with thunderstorms and have only recently been documented using low-light-level television technology. Since the first sprite images were obtained accidentally in 1989, thousands of images have been obtained, including over 20 from the space shuttle. They show massive but weak luminous flashes appearing directly above active thunderstorms coincident with cloud-to-ground or intracloud lightning. They extend from the cloud tops to about 95 km (59 mi) and are predominantly red. High-speed photometer measurements show that the duration of sprites is only a few milliseconds. Their brightness is comparable to a moderately bright auroral arc. The optical energy is roughly 10–50 kJ per event, with a corresponding optical power of 5–25 MW. Elves are associated with sprites and are a recent discovery. They are optical emissions of approximately 1 millisecond, with a fast lateral, horizontal expansion that emits more red than blue light. They occur at altitudes of 75–95 km (47–59 mi). *See* SPACE SHUTTLE.

Numerous images have been obtained from aircraft of blue jets. They are a second high-altitude optical phenomenon, distinct from sprites and observed above thunderstorms using low-light television systems. Blue jets are optical ejections from the top of the electrically active core regions of thunderstorms. Following the emergence from the top of the thundercloud, they typically propagate upward in narrow cones of about 15° full width at vertical speeds of roughly 100 km/s (60 mi/s), fanning out and disappearing at heights of about 40–50 km (25–30 mi). Richard E. Orville

Bibliography. *The Earth's Electrical Environment*, National Research Council Studies in Geophysics, 1986; D. R. MacGorman and W. D. Rust, *The Electrical Nature of Storms*, 1998; M. A. Uman, *All About Lightning*, 1986; M. A. Uman, *The Lightning Discharge*, 1987; M. A. Uman and E. P. Krider, Natural and artificially initiated lightning, *Science*, 246:457–464, 1989.

Lightning and surge protection

Means of protecting electrical systems, buildings, and other property from lightning and other high-voltage surges.

The destructive effects of natural lightning are well known. Studies of lightning and means of either preventing its striking an object or passing the stroke harmlessly to ground have been going on since the days when Franklin first established that lightning is electrical in nature. From these studies, two conclusions emerge: (1) Lightning will not strike an object if it is placed in a grounded metal cage. (2) Lightning tends to strike, in general, the highest objects on the horizon. *See* ATMOSPHERIC ELECTRICITY; LIGHTNING.

One practical approximation of the grounded metal cage is the well-known lightning rod or mast (**Fig. 1***a*). The effectiveness of this device is evaluated on the cone-of-protection principle. The protected area is the space enclosed by a cone having the mast top as the apex of the cone and tapering out to the base. Laboratory tests and field experience have shown that if the radius of the base of the cone is equal to the height of the mast, equipment inside this cone will rarely be struck. A radius equal to twice the height of the mast gives a cone of shielding within which an object will be struck

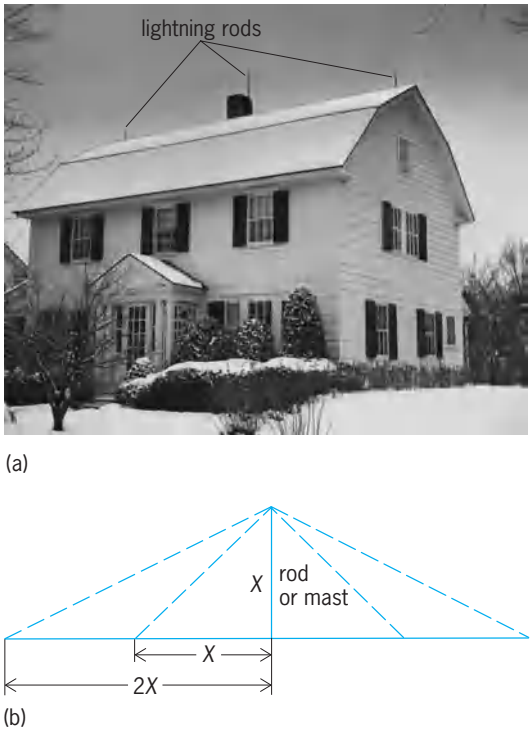

Fig. 1. Lightning rod cone of protection. (*a*) Configuration of rods on a house. (*b*) Geometry of the principle.

occasionally. The cone-of-protection principle is illustrated in Fig. 1*b*.

Another criterion for cone of protection of a lightning mast is based on the distance over which the final lightning stroke breakdown occurs. As the stroke can contact any point within striking distance of this final breakdown, the configuration of a sphere with a radius of the striking distance, conservatively taken as 100 ft (30 m), defines the vulnerable area, and locations outside the sphere would be protected. A different visualization of this would be a ball having a 100-ft radius, rolled on the ground toward a mast; when it contacts the mast, all points between the mast and the point where the ball contacts the ground, outside and below the ball, would be protected.

A building which stands alone, like the Empire State in New York City, is struck many times by lightning during a season. It is protected with a mast, and the strokes are passed harmlessly to ground. However, lightning has been observed to strike part way down the side of this building (**Fig. 2**). This shows that lightning does not always strike the highest object but rather chooses the path having the lowest electrical breakdown.

The probability that an object will be struck by lightning is considerably less if it is located in a valley. Therefore, electric transmission lines which must cross mountain ranges often will be routed through the gaps to avoid the direct exposure of the ridges. *See* ELECTRIC DISTRIBUTION SYSTEMS; ELECTRIC POWER SYSTEMS.

Overhead lines of electric power companies are vulnerable to lightning. Lightning appears on these lines as a transient voltage, which, if of sufficient magnitude, will either flash over or puncture the weakest point in the system insulation.

Many of the troubles that cause service interruptions on electrical systems are the results of flashovers of insulation wherein no permanent damage is done at the point of fault, and service can be restored as soon as the cause of the trouble has disappeared. A puncture or failure of the insulation, on the other hand, requires repair work, and damaged apparatus must be removed from service.

There are a number of protective devices to limit or prevent lightning damage to electric power systems and equipment. The word protective is used to connote either one or two functions: the prevention of trouble, or its elimination after it occurs. Various protective means have been devised either to prevent lightning from entering the system or to dissipate it harmlessly if it does. *See* ELECTRIC PROTECTIVE DEVICES.

Overhead ground wires and lightning rods. These devices are used to prevent lightning from striking the electrical system.

The grounded-metal-cage principle is approached by overhead ground wires, preferably two, installed over the transmission phase conductors and grounded at each tower (**Fig. 3**). The ground wires must be properly located with respect to the phase conductors to provide a cone of protection and have adequate clearance from them, both at the towers and throughout the span. If the resistance to ground of the tower is high, the passing of high lightning currents through it may sufficiently elevate the tower in potential from the transmission line conductors so that electrical flashover can occur. Since the magnitude of the lightning current may be defined in terms of probability, the expected frequency of line flashover may be predicted from expected storm and lightning stroke frequency, the current magnitude, and the tower footing resistance. Where

Fig. 2. Multiple lightning stroke A to Empire State and nearby buildings, August 23, 1936. Single, continuing stroke B to Empire State Building, August 24, 1936.

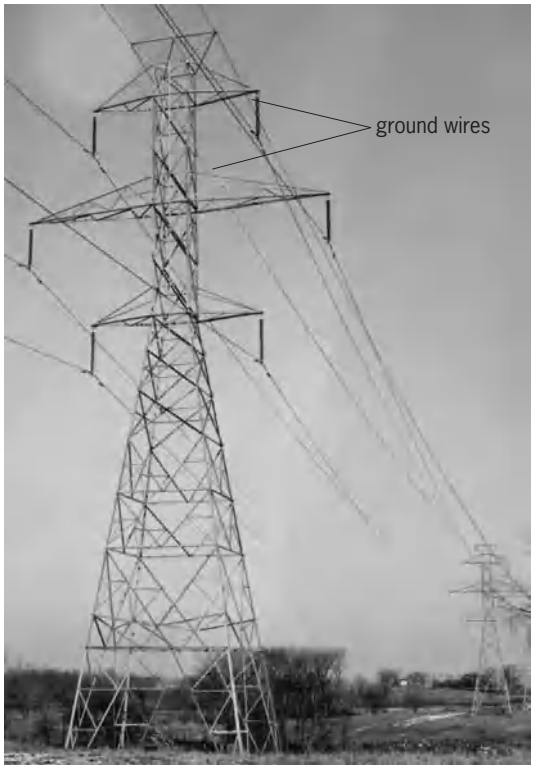

Fig. 3. Double-circuit suspension tower on the Olive–Gooding Grove 345-kV transmission line, with two circuits strung and with two ground wires. (*Indiana and Michigan Electric Co., American Electric Power Co.*)

the expected flashover rate is too high, means to lower the tower footing resistance are employed such as driven ground rods or buried wires connected to the base of the tower. Applications of zinc oxide surge arresters across line insulator strings in areas of high tower footing resistance have been effective in lowering line flashover frequency.

The ground wires are often brought in over the terminal substations. For additional shielding of the substations, lightning rods or masts are installed. Several rods are usually used to obtain the desired protection. *See* ELECTRIC POWER SUBSTATION.

Lightning arresters. These are protective devices for reducing the transient system overvoltages to levels compatible with the terminal-apparatus insulation. They are connected in parallel with the apparatus to be protected. One end of the arrester is grounded and also connected to the case of the equipment being protected; the other end is connected to the electric conductor (**Figs. 4** and **5**).

The nonlinearity of current with respect to voltage of a lightning arrester provides a relatively low discharge path to ground for the transient overvoltages, and a relatively high resistance to the power system follow current, so that their operation does not cause a system short circuit.

In selecting an arrester to protect a transformer, for example, the voltage levels that can be maintained by the arrester both on lightning surges and surges resulting from system switching must be coordinated with the withstand strength of the transformers to these surges. Arrester discharge voltages on various magnitudes and wave shapes of discharge currents, the inductance of connections to the arrester, voltage wave shape, and other factors must all be evaluated. *See* SURGE ARRESTER.

Rod gaps. These also are devices for limiting the magnitude of the transient overvoltages. They usually are formed of two $^1/_2$-in.^{2} (3-cm^2) rods, one of which is grounded and the other connected to the line conductor but may also have the shape of rings or horns. They have no inherent arc-quenching ability, and once conducting, they continue to arc until the system voltage is removed, resulting in a system outage.

These devices are applied on the principle that, if an occasional flashover is to occur in a station, it is best to predetermine the point of flashover so that it will be away from any apparatus that might otherwise be damaged by the short-circuit current and the associated heat.

The flashover characteristics of rod gaps are such that they turn up (increase of breakdown voltages with decreasing time of wavefront) much faster on steep-wavefront surges than the withstand-voltage characteristics of apparatus, with the result that if a gap is set to give a reasonable margin of protection on slow-wavefront surges, there may be little or no protection for steep-wavefront surges. In addition, the gap characteristics may be adversely affected by weather conditions and may result in undesired flashovers.

Immediate reclosure. This is a practice for restoring service after a fault occurs by immediately reclosing automatically the line power circuit breakers that have been tripped. The protective devices involved are the power circuit breaker and the fault-detecting and reclosing relays.

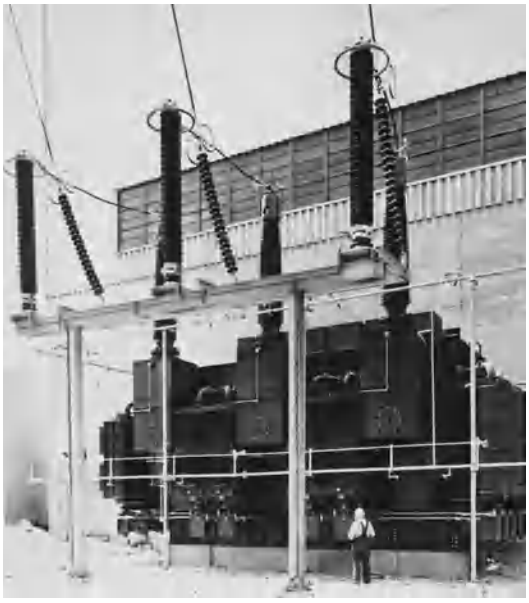

Fig. 4. Installation of three single-pole station lightning arresters rated 276 kV for lightning and surge protection of large 345-kV power transformer. (*General Electric Co.*)

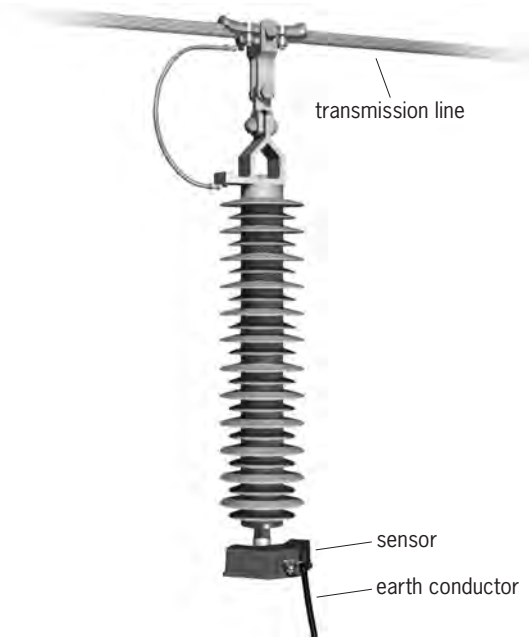

Fig. 5. PEXLINK® high-voltage surge arrester in silicone housing, equipped with special link arrangement for installation on a transmission line. The arrester protects the line from overvoltages such as lightning. The EXCOUNT-II sensor monitors the health of both the transmission line and the surge arrester. Data from the sensor can be read with a wireless transceiver and then transferred to a computer for statistical analysis. (*ABB Ltd.; photograph by Mats Findell*)

This practice is successful because the majority of the short circuits on overhead lines are the result of flashovers of insulators and there is no permanent damage at the point of fault. The fault may be either line-to-ground or between phases. Reclosing relays are available to reenergize the line several times with adjustable time intervals between reclosures.

If the relays go through the full sequence of reclosing and the fault has not cleared, they lock out. If the fault has cleared after a reclosure, the relays return to normal.

Permanent faults must always be removed from a system and the accepted electric protective devices are power circuit breakers and suitable protective relays. Glenn D. Breuer

Bibliography. Electric Power Research Institute, *EPRI AC Transmission Line Reference Book: 200kV and Above*, 3d ed., 2005; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; M. M. Frydenlund, *Lightning Protection for People and Property*, 1993; T. Hovrath, *Computation of Lightning Protection*, 1991; National Fire Protection Association, *NFPA 780, Standard for the Installation of Lightning Protection Systems, 2004 Edition*, 2004.

Lignin

A polymer found extensively in the cell walls of all woody plants; is one of the most abundant natural polymers. Lignin, which constitutes one-fourth to one-third of the total dry weight of trees, combines with hemicellulose materials to help bind the cells together and direct water flow. *See* CELL WALLS (PLANT); HEMICELLULOSE.

Isolation. Several methods have been devised for isolating lignin from wood; isolated lignins are frequently named for the persons who developed the procedures. Brauns lignin is obtained by exhaustively extracting wood meal with 95% ethanol and then diluting the extract with water; the resulting precipitate is taken up in dioxane and reprecipitated with ether. Freudenburg lignin is obtained by a similar process, except the initial extraction is done with 90% acetone and the final purification with ethyl acetate.

Some isolation methods are based on acid treatments in which the carbohydrate components (cellulose and hemicelluloses) are hydrolyzed to water-soluble materials. Klason lignin (72% sulfuric acid) and dioxane lignin (hydrochloric acid) are produced in this way. A much milder and more popular isolation method gives Björkman or "milled wood" lignin. Here, wood meal is ground in a ball mill to break down the cell structures and then extracted with dioxane-water. This procedure, like most of the others, suffers from a poor yield (50%) of lignin. With all these procedures, serious doubts exist as to whether the isolated lignin is representative of the "native" lignin.

Enzymatic digestion of the carbohydrate in wood meal is a lengthy, tedious procedure but offers the greatest promise of leaving lignin unaltered during isolation. The isolated lignin contains some chemically bonded hemicellulose materials known as lignin-carbohydrate complexes.

Structure. Structural studies on lignin have been hampered by the random, cross-linked nature of the polymer. The relative proportions of the monomers which make up lignin [structures (**1**) and (**2**)] vary

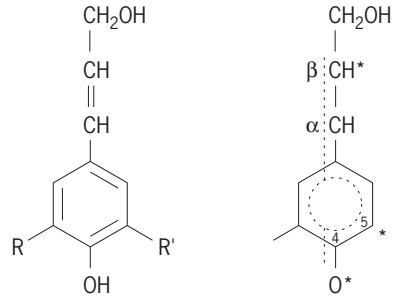

(**1**) (**2**)

with the plant species. In general, lignins in softwoods (pine, spruce, fir) are derived principally from coniferyl alcohol ($R = OCH_3$, $R' = H$), with a few percent sinapyl alcohol ($R = R' = OCH_3$) and trace amounts of *p*-coumaryl alcohol ($R = R' = H$) structures. The precursors of lignins in hardwoods (oak, maple, birch) are coniferyl alcohol and sinapyl alcohol in roughly equal amounts.

Lignin is formed in the plant by an enzymatic dehydrogenation of the monomers. The * designations shown on the structure (**2**) indicate the principal sites for coupling monomers. The **illustration** presents a possible partial lignin structure. There are

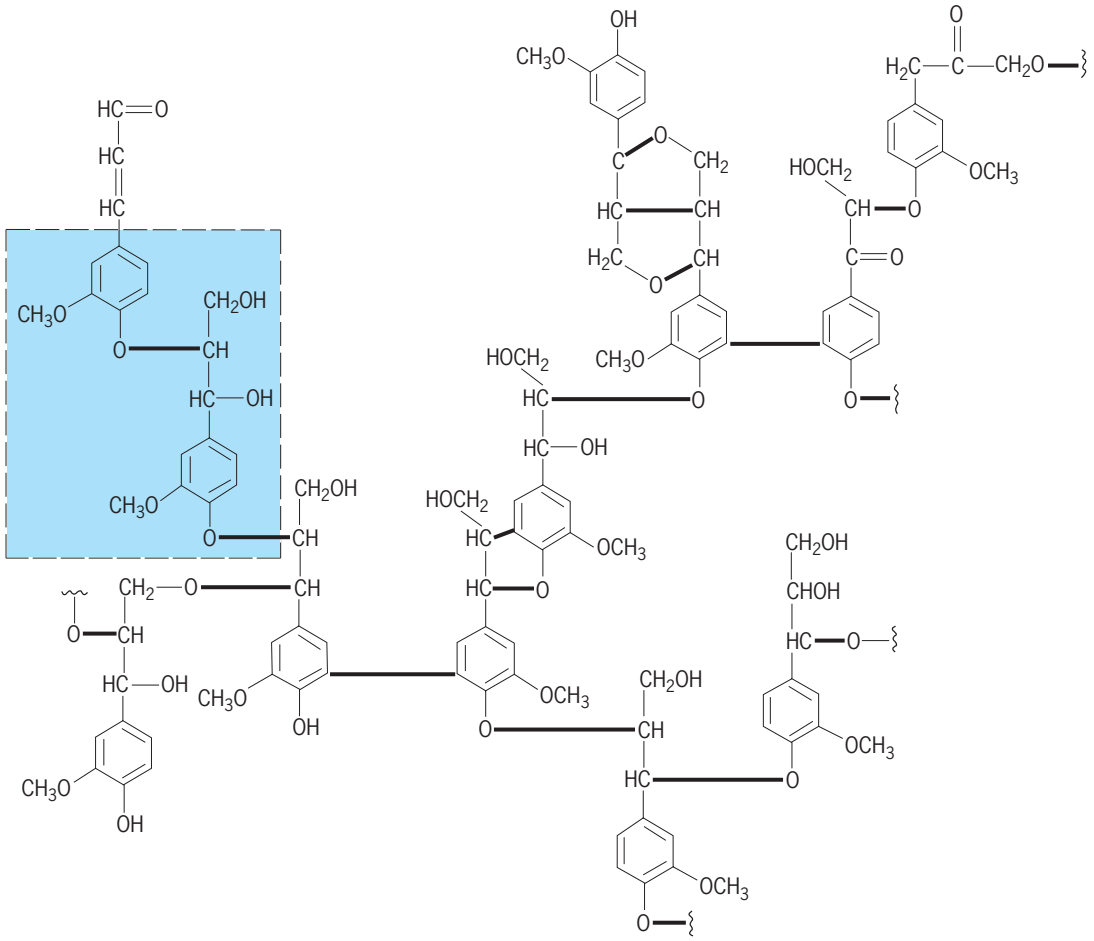

Chemical formula representation of lignin, in which bold lines represent linkages between individual phenylpropane lignin monomer units, and shaded area represents the structure of a typical lignin model compound.

occasional multiple (or secondary) linkages between units, especially at the α-carbon position. Research on lignin chemistry is often done with models of lignin. These models are typically dimers, joined by a β-aryl ether bond (see illus.). Unlike lignin, the model and its reaction products are easily characterized. Chemical reactions of lignin are assumed to parallel those of the models. The goal of chemical pulping is to break down and solubilize the lignin (delignify the wood and, thus, liberate its cellulose fibers for re-formation into paper products). *See* PAPER.

Uses. Very little lignin is isolated from pulp liquors and sold as such. In general, the markets for lignin products are not large or attractive enough to compensate for the cost of isolation and the energy derived from its burning. An exception is lignosulfonate, which is obtained either directly from sulfite pulping liquors or by sulfonation of acid-precipitated kraft lignin; its markets include a dispersant in carbon black slurries, clay products, dyes, cement, oil drilling muds, an asphalt emulsifier, a binder for animal feed pellets, a conditioner for boiler water or cooling water, and an additive to lead-acid storage battery plate expanders. *See* WOOD CHEMICALS.

Donald R. Dimmel

Bibliography. J. P. Casey (ed.), *Pulp and Paper Chemistry and Technology,* 3d ed., vol. 1, 1980; R. L. Crawford, *Lignin Biodegradation and Transformation*, 1981; D. Fengel and G. Wegener, *Wood: Chemistry, Ultrastructure, Reactions*, 1989; W. G. Glasser and S. Sarkanen (eds.), *Lignin: Properties and Materials*, 1989; I. A. Pearl, *The Chemistry of Lignin*, 1967; E. Sjöström, *Wood Chemistry: Fundamentals and Applications*, 2d ed., 1993.

Lignite

A brownish-black, low-rank coal defined in the engineering sense by the American Society for Testing and Materials as having a heating value of less than 8300 Btu/lb (4611 kcal/kg), determined on a moist, mineral-matter-free basis. Lignite occurs in two subclasses: lignite A [8300–6300 Btu/lb (4611–3500 kcal/kg)] and lignite B [less than 6300 Btu/lb (3500 kcal/kg)]. Vitrinite reflectance of lignite is 0.25–0.38%; however, vitrinite reflectance may be more suitably applied to higher-rank coals. Outside North America, low-rank coal is classified as brown coal, which includes lignite, subbituminous, and most high-volatile C bituminous coal of the North American classification system. Brown coal is divided into soft and hard coal; hard coal is subdivided into dull and bright coal. *See* COAL.

Generally, the fixed-carbon content of lignite is 60–68%, and volatile-matter content exceeds 50% (on dry, ash-free basis). The oxygen-to-carbon and hydrogen-to-carbon ratios (calculated on an atomic basis) are greater in lignite than in higher-rank coals. Porosity of lignite is about 25–30%, highest of all coals. Moisture ranges 30–75%; water occurs in pores and capillaries and is bound to organic molecules. Porosity, and thus moisture-holding capacity, decreases with increasing rank in lignite.

In the United States, lignite contains more alkali and alkaline-earth metals (especially sodium in the Northern Great Plains Region) than higher-rank coals. Relative to higher-rank coals, lignite is more reactive and commonly lower in sulfur content; and the contained inorganic matter is more likely to be organically bound (complexed) than in discrete minerals. *See* ALKALI METALS; ALKALINE-EARTH METALS.

Because lignite has undergone less coalification than higher-rank coals, the organic precursor constituents are more easily recognized than those in high-rank coals, and lignite serves as an invaluable link between peat and high-rank coals in studying the coal origin. *See* COAL PALEOBOTANY; PEAT.

Occurrences. World resources of lignite are difficult to assess because of the different classification systems for low-rank coal in North America and elsewhere. However, the world's greatest in-place resources of low-rank coal (lignite and subbituminous coal in the United States and brown coal elsewhere) occur in the United States, Russia, Australia, Germany, and the former Yugoslavia.

The lignite resource in the United States, primarily lignite A, is 5.19×10^{11} short tons (4.71×10^{11} metric tons), which is about 30% of the total identified United States coal resource. United States lignite occurs primarily in the Northern Great Plains Region, which has 4.65×10^{11} short tons (4.22×10^{11} metric tons) [90%], and the Gulf Coast Basin, which has 5.3×10^{10} short tons (4.8×10^{10} metric tons) [10%] the lignite resource. The major lignite-producing state is Texas, followed by North Dakota. Nearly all lignite in the United States is produced by surface mining. *See* SURFACE MINING.

Transportation. Lignite is difficult to store and transport. Its high moisture content adds to cost of transportation, and its high reactivity causes spontaneous combustion. Experiments with drying, spraying with oil, briquetting, and use of a slurry pipeline as alternatives or improvements in the transport and storage performance have shown that these techniques are not cost-effective.

Uses. Therefore, production is usually for use at short distances. In the United States, as elsewhere, lignite is used primarily to generate electricity at mine-mouth power plants. Lignite has been successfully used as a feedstock for gasification, liquefaction, and pyrolysis. Minor uses of lignite are montan wax, activated carbon, firing kilns, and home heating. *See* ACTIVATED CARBON; COAL CHEMICALS.

Lignite has certain utilization advantages over higher-rank coal. Because of its high reactivity, which is due in part to high oxygen content, lignite does not have to be ground as finely as higher-rank coal to ensure complete combustion in pulverized coal systems. Advantages for gasification and liquefaction are high reactivity, low sulfur content, and noncaking properties. Alkali and alkaline-earth elements have catalytic properties in gasification and possibly liquefaction. *See* COAL GASIFICATION; COAL LIQUEFACTION.

Disadvantages of lignite use in combustion, relative to higher-rank coals, are low heating value and the fact that fly ash from lignite may have high electrical resistivity, making it difficult to collect in electrostatic precipitators. Mineral matter in lignite is largely organically bound and inseparable by standard washing techniques; the high sodium content of some lignite contributes to boiler fouling and slagging problems; and the quartz content, where high, accelerates erosion of coal-feed systems and furnace burners. In gasification, moisture acts as a diluent. Walter B. Ayers

Bibliography. American Society for Testing and Materials, *1991 Annual Book of ASTM Standards*, pt. 12: *Gaseous Fuels; Coal and Coke; Atmospheric Analysis*, 1991; N. Berkowitz, *The Chemistry of Coal: Coal Science and Technology,* 7, 1985; D. J. Trantolo and D. L. Wise (eds.), *Energy Recovery from Lignin, Peat, and Lower Rank Coals*, 1989.

Lignumvitae

A tree, *Guaiacum sanctum*, also known as holywood lignumvitae, which is cultivated to some extent in southern California and tropical Florida. Lignumvitae is native in the Florida Keys, Bahamas, West Indies, and Central and South America. It is an evergreen tree of medium size with abruptly pinnate leaves. The tree yields a resin or gum known as gum guaiac or resin of guaiac which is used in medicine. The very heavy black heartwood is used in bowling balls, blocks and pulleys, and parts of instruments. *See* FOREST AND FORESTRY; SAPINDALES; TREE. Arthur H. Graves; Kenneth P. Davis

LIGO (Laser Interferometer Gravitational-wave Observatory)

A physics research facility developed to detect cosmic gravitational waves and to measure these waves for scientific research. It consists of two installations within the continental United States, located in the states of Washington and Louisiana, operated in unison as a single observatory. Funded by the National Science Foundation (NSF), LIGO was designed and constructed by a team of scientists from the California Institute of Technology and the Massachusetts Institute of Technology, and by industrial contractors. Construction of the facilities was completed in 1999, and operation of the detectors began

in 2001. LIGO is the leading member in a developing global network of gravitational-wave observatories.

Operating principles. Gravitational waves are emitted by accelerating masses much as electromagnetic waves are produced by accelerating charges. Unlike electromagnetic waves, astrophysical-sized objects, such as massive stars, are required to produce gravitational waves that are strong enough to be detectable. Such a source of waves may be in our galaxy, the Milky Way, or in a relatively nearby galaxy. As the wave propagates away from its source, it produces small oscillations in the space-time nature of gravity. To measure these oscillations, widely separated test masses have been set up, each hung by fibers to form a pendulum, as probes of the oscillating space-time (**Fig. 1**). The test masses are constructed as mirrors and arranged along the two arms of an L-shape, so that their separations can be monitored using a specialized type of Michelson interferometer. The interferometer is illuminated with a powerful laser, and at the vertex of the L-shape the laser beam is split into two beams, so that half the laser light travels down one arm of the L and half down the other. At the ends of the arms, each beam is reflected on itself by the end test masses, such that the two beams recombine at the vertex beam splitter. At the output port of the beam splitter, a photodetector monitors how much light exits the instrument. Nominally, when the two arms are of equal length, the two beams destructively interfere at the beam splitter, and no light reaches the photodetector. If a gravitational wave passes through the interferometer, the light travel time down and back one arm will be longer than that of the other, and the beams will no longer completely destructively interfere when they recombine at the beam splitter. Thus some light will exit the output port and be detected by the photodetector, indicating the presence of a gravitational wave. *See* ELECTROMAGNETIC RADIATION; INTERFEROMETRY.

The challenge is to make the instrument as sensitive as possible because the scale of the distance changes is incredibly small. A gravitational wave produces an oscillating strain in space, implying that the magnitude of the arm length change, dL, is proportional to the length of each arm. This fact provides the motivation for making the arms as long as possible, specifically 4 km (2.5 mi) long in LIGO. Even at this scale, the motion that must be measured is extremely small: gravitational waves passing through Earth are expected to carry strain amplitudes of order 10^{-21} or smaller, and the resulting arm length change, dL, is of order 10^{-18} m, 1000 times smaller than the diameter of a nucleus.

Refinements. Many refinements have been made to the basic interferometer to increase the sensitivity. One of these is to increase the interaction time of the light with the gravitational wave beyond the round-trip time down and back an arm. This is done by effectively folding the beams multiple times within each arm, by having the two test masses in each arm form a resonant optical cavity. Laser photons entering an arm cavity make, on average, 50 round trips

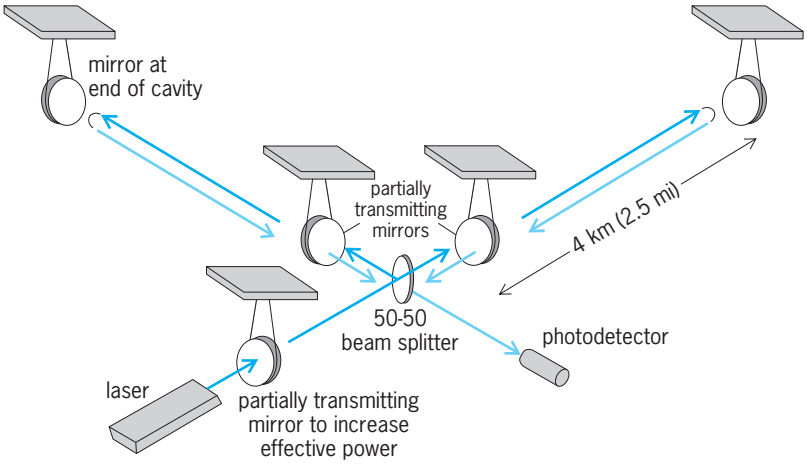

Fig. 1. Experimental setup of LIGO. The interferometric gravitational-wave detector is an equal-arm Michelson laser interferometer whose hanging mirrors serve as gravitational test masses. Beam paths and all optics are enclosed in a high-vacuum system. Each 4-km-long (2.5-mi) arm contains a resonant optical cavity in which each photon, on average, makes about 50 round trips before recombining with photons from the other arm of the beam splitter. The interference is measured with a photodetector, which is sensitive enough to detect phase shifts to a few times 10^{-10} of an interference fringe, such as might be produced by a passing gravitational wave.

in the arm before recombining with photons from the other arm cavity. This increases the interaction time of the light with the gravitational wave to about 1 millisecond, increasing the phase delay difference between the two arms and ultimately increasing the light power at the output photodetector.

Another special technique increases the effective light power by making the entire interferometer a resonant optical cavity. Because the two arms beams interfere destructively at the output port, most of the incident laser light interferes constructively on the input side of the beam splitter and actually returns toward the laser. This makes it possible to achieve a significant gain by placing a partially transmitting mirror between the laser and the beam splitter. By properly positioning this mirror, one can arrange that no light is reflected back toward the laser, so that all the light is resonantly built up in the interferometer. This technique, called power recycling, increases the laser power at the beam splitter by a factor of about 40. *See* CAVITY RESONATOR.

Increasing the interaction time and the effective power serves to mitigate a type of detector noise that is due to the quantum nature of light, often referred to as shot noise. Another type of detector noise that must be addressed is fluctuating random forces applied to the mirrors, which could be from a variety of sources. Isolation from motions of the ground, for example, is provided by cascaded stages of vibration isolation, both passive and active. A more fundamental type of noise is Brownian motion driven by thermal excitations; this is controlled largely by careful choice of materials and engineering. *See* BROWNIAN MOVEMENT; ELECTRICAL NOISE; VIBRATION ISOLATION.

Siting and status. LIGO consists of three of these interferometers, two located at the Hanford, Washington, site (**Fig. 2**) and one at the Livingston,

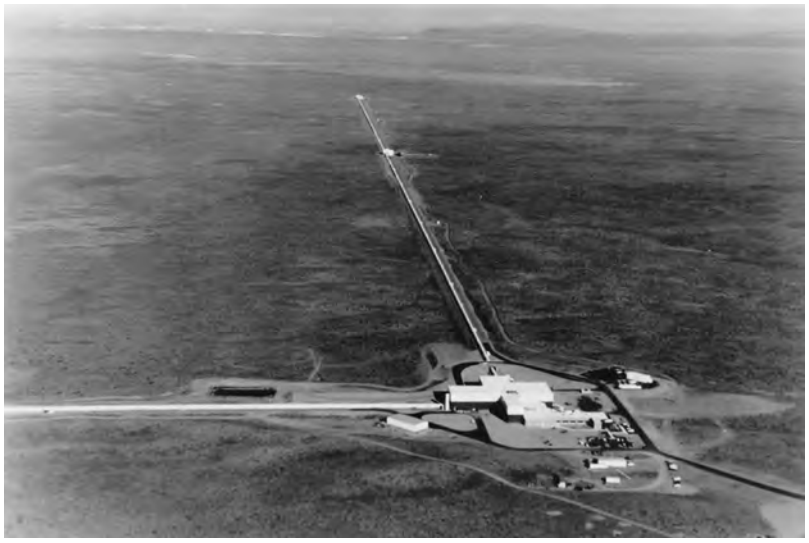

Fig. 2. Aerial photograph of the LIGO at Hanford, Washington. This facility houses two interferometers (a nearly identical facility, with a single interferometer, is in Livingston, Louisiana). The lasers and optics are contained in the white buildings; the 4-km-long (2.5-mi) vacuum tubes are covered by the arched concrete enclosure.

Louisiana, site. At least two detectors located at widely separated sites are essential for the certain detection of gravitational waves. Regional phenomena such as micro-earthquakes, acoustic noise, and laser fluctuations can cause disturbances that simulate a gravitational wave event. Such perceived events may happen locally at one site, but such disturbances are unlikely to happen simultaneously at two widely separated sites.

The LIGO detectors achieved their design sensitivity to gravitational waves near the end of 2005 and began an extended data collection phase. The next generation of LIGO interferometers is being designed and planned for implementation after 2010; these will provide an order of magnitude (10 times) greater sensitivity to gravitational waves. *See* GRAVITATIONAL RADIATION.

Peter K. Fritschel

Bibliography. B. Barish and R. Weiss, LIGO and the detection of gravitational waves, *Phys. Today*, 52(10):44–50, October 1999; M. Bartusiak, Catch a gravity wave, *Astronomy*, 28(10):54–59, October 2000; M. Bartusiak, *Einstein's Unfinished Symphony: Listening to the Sounds of Space-Time*, Joseph Henry Press, Washington, DC, 2000; G. H. Sanders and D. Beckett, LIGO: An antenna tuned to the songs of gravity, *Sky Telesc.*, 100(4):40–48, October 2000.

Liliales

An order of monocotyledons, the well-known lilies or family Liliaceae of many previous botanists, actually consisting of 9 families and about 1600 species. Liliales are clearly circumscribed in deoxyribonucleic acid (DNA) sequence analyses, but are difficult to define on the basis of morphological characters, resulting in varying family composition in different classifications. Separation of Liliales from the lilylike Asparagales has proved particularly problematic, but they differ in nectary position (septal in Asparagales) and absence of phytomelan, a dark seedcoat pigment present in most families of Asparagales. Nearly all features marking the families of Liliales are micromorphological (for example, perigonal nectaries, nuclear endosperm formation).

Many members of the order are herbaceous perennials, but there are also some vines, including the greenbriar (*Smilax*, Smilaceae) and *Lapageria* (the national flower of Chile, Philesiaceae). Liliaceae include species with some of the largest genomes recorded in the angiosperms, the record being held by *Fritillaria syriaca*. Many taxa are extremely poisonous. Most Colchicaceae, for example, possess colchicine-type alkaloids. *Lilium* (lilies), *Tulipa* (tulips), and *Fritillaria* (fritillaries) in Liliaceae, *Colchicum* (autumn crocuses, Colchicaceae), and *Alstroemeria* (Peruvian lilies, Alstroemeriaceae) are well-known horticultural plants. *Trillium* (wakerobin, Melanthiaceae) and *Erythronium* (troutlilies, Liliaceae) are familiar United States wildflowers. *See* ASPARAGUS; COLCHICINE; GARLIC; LILIIDAE; LILIOPSIDA; MAGNOLIOPHYTA; ONION; ORCHIDALES; ORNAMENTAL PLANTS; PLANT KINGDOM; SISAL.

Michael F. Fay; Mark W. Chase

Liliidae

A subclass of the class Liliopsida (monocotyledons) of the division Magnoliophyta (Angiospermae), the flowering plants, consisting of 2 orders (Liliales and Orchidales), 19 families, and about 25,000 species. The Liliidae are syncarpous monocotyledons with both the sepals and the petals usually petaloid. The seeds are either nonendospermous or have an endosperm with various sorts of reserve foods such as cellulose, fats or protein, and only seldom starch. The stomates are without subsidiary cells or, in a few small families, have two subsidiary cells. The pollen is binucleate. The vast majority of the species have the vessels chiefly or wholly confined to the roots. The flowers generally have well-developed nectaries, and pollination is usually by insects or other animals. *See* LILIALES; LILIOPSIDA; MAGNOLIOPHYTA; ORCHIDALES; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Liliopsida

One of the two classes which collectively make up the division Magnoliophyta (Angiospermae), the flowering plants. The Liliopsida, often known as Monocotyledoneae, or monocotyledons, embrace 5 subclasses (Alismatidae, Commelinidae, Arecidae, Zingiberidae, and Liliidae), 18 orders, 61 families, and about 55,000 species.

All of the characters which collectively distinguish the Liliopsida from the Magnoliopsida (dicotyledons) are subject to exception, but most of the Liliopsida

have parallel-veined leaves, and when the embryo is differentiated into recognizable parts, there is only a single cotyledon. The vascular bundles are generally scattered or borne in two or more rings, so the stems and roots do not have a well-defined pith and cortex.

Monocotyledons never have an intrafascicular cambium, and most of them have no secondary growth at all. Some few of them have an unusual type of cambium which produces vascular bundles toward the inside and usually only a small amount of parenchyma toward the outside. Woody monocots such as palms become woody by lignification of the ground tissues in which the vascular bundles of the stem are embedded. The mature root system of monocots is wholly adventitious. The floral parts of monocots, when of definite number, are most often borne in sets of 3, seldom 4, never 5. The pollen is uniaperturate or of uniaperturate-derived type.

It is generally agreed that the monocotyledons are derived from primitive dicotyledons. The solitary cotyledon, the parallel-veined leaves, the absence of a cambium, the dissected stele, and the adventitious root system of monocotyledons are all regarded as secondary rather than primitive characters in the angiosperms as a whole, and any plant which was more primitive than the monocotyledons in these several respects would certainly be a dicotyledon. The monocotyledons are more primitive than the bulk of the dicotyledons in having uniaperturate pollen, but several of the archaic families of dicotyledons also have uniaperturate pollen. *See* ALISMATIDAE; ARECIDAE; COMMELINIDAE; LILIIDAE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM; ZINGIBERIDAE. Arthur Cronquist; T. M. Barkley

Lime (botany)

An acid citrus fruit, *Citrus aurantifolia*, usually grown in tropical or subtropical regions because of its low resistance to cold. The two principal groups of limes are the West Indian or Mexican and the Tahiti or Bearss. The West Indian lime (see **illus.**) is a medium-sized, spreading tree with numerous willowy branches densely armed with short, stiff spines. Flowers are small and flowering occurs throughout the year, but mainly in the spring. The fruit is very small (walnut size) and strongly acid, and drops when fully colored. The West Indian lime is more sensitive to cold than the Tahiti lime, which is a more vigorous tree, bearing fruits of lemon size. The Tahiti lime is seedless and its aroma is less pronounced.

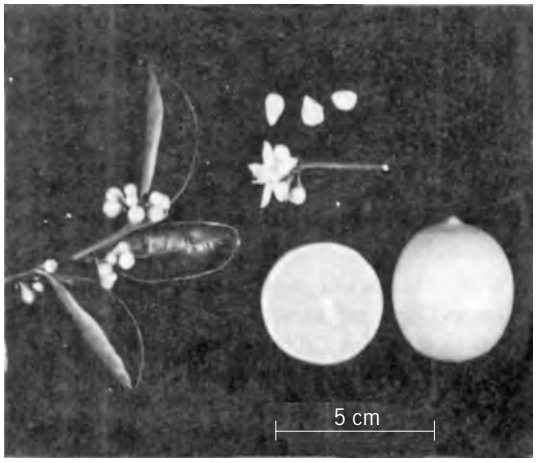

Foliage, seeds, flower, and fruit of the West Indian or Mexican lime.

The limes are believed to have originated in northeastern India or adjoining portions of Burma or northern Malaysia. It probably was introduced into Europe by the Arabs and was brought to the Americas by the Spanish and Portuguese explorers in the sixteenth century. It escaped cultivation and became feral in parts of the West Indies, some Caribbean countries, and southern Florida.

Except in the United States, the commercial lime industry is restricted to the West Indian group, which has a high total heat requirement for good-sized fruit. The major producing areas are India, Mexico, Egypt, and the West Indies. Plantings are scattered and production statistics uncertain.

Commercial production of the Tahiti lime is more recent and largely confined to the United States. It is grown mainly in Florida, with some plantings in the warmer areas of southern California. A little over half of the crop is processed, with the remainder utilized fresh. Fruits are harvested before turning yellow because they have maximum aroma and storage life. *See* FRUIT; FRUIT, TREE. R. K. Soost

Lime (industry)

A general term for burned (or calcined) limestone, also known as quicklime, hydrated lime, and unslaked or slaked lime. Use of lime as a building material in mortar and plaster coincides with earliest recorded history. However, since development of the chemical process industries, its predominant usage is as a basic industrial chemical, where it ranks second to sulfuric acid in tonnage. It still enjoys its traditional building uses, but more than 90% is used as a chemical. Uses in order of decreasing size are steel fluxing, water treatment, nonferrous metals (alumina, magnesium, copper, and others), pulp and paper, refractories, soil stabilization, sewage and trade waste treatment, chemicals, and glass manufacture.

Lime is not a mineral; it is manufactured from a mineral—limestone, coral, oystershell, all being sources of calcium carbonate. Dolomite, a calcium-magnesium carbonate, is used to produce dolomitic (magnesium) lime. Only the purest types of stone or shell are used for lime. *See* LIMESTONE.

The quarried or mined limestone raw material is crushed and screened to produce either pebble size ($^1/_2$–2 in. or 1.3–5 cm) for charging into rotary kilns or lump size (3–8 in. or 7.5–20 cm) for vertical

kilns. At high temperatures of 2000–2500°F (1100–1400°C), the limestone is decomposed, evolving carbon dioxide gas that constitutes 44% of the weight of the limestone, with calcium oxide remaining. *See* KILN.

The resulting lime, depending upon degree of burn, is chemically reactive with water and acids. Soft-burned limes literally explode on contact with water. If the lime is hard-burned, its reactivity is retarded. One lime product is dead-burned dolomite that has no reactivity with water, making it useful as a refractory material. Unlike conventional limes, hydraulic lime is made from impure limestone that contains considerable silica, alumina, and iron. Calcination causes the lime to combine chemically with these impurities, forming silicates and aluminates. Such limes when hydrated will harden under water, somewhat like cement.

When lime hydrates in water, considerable heat is generated, enough for 1 lb of quicklime to heat 2 lb of water from 0°F (−18°C) to boiling. The lime manufacturer produces hydrated lime (calcium hydroxide), an ultrafine, fluffy, dry powder by controlled addition of water to ground quicklime. Hydrated lime is 25% chemically combined water by weight. Lime users also hydrate or slake quicklime with excess water into slurries or "milk of lime" of various concentrations. Dolomitic lime does not slake as readily as high-calcium types. Under normal conditions, only the calcium oxide component hydrates, yielding $Ca(OH)_2 \cdot MgO$. If slaking occurs under pressure, a highly hydrated dolomitic lime is produced.

Hydrated lime, as a base, is a strong alkali and neutralizes the strongest acids, forming neutral salts, such as the sulfates and chlorides of calcium and magnesium. Robert S. Boynton

Bibliography. R. Boynton, *Chemistry and Technology of Lime and Limestone*, 2d ed., 1980; D. D. Carr (ed.), *Industrial Minerals and Rocks*, 6th ed., 1994; K. A. Gutschick (ed.), *Lime for Environmental Uses*, 1987; S. J. Lefond (ed.), *Industrial Minerals and Rocks*, 5th ed., AIME, 1983; U.S. Bureau of Mines, Lime, *Minerals Yearbook*, 1988.

Limestone

A common sedimentary rock composed predominantly of carbonates of calcium and magnesium. Limestones are the most voluminous of the nonsiliciclastic sedimentary rocks. In the strict sense, limestones refer to sedimentary rocks composed of the calcium carbonate mineral calcite ($CaCO_3$). Those rocks, dominated by the magnesium-calcium carbonate mineral dolomite [$CaMg(CO_3)_2$], are known as dolomites or dolostones. Although most limestones are similar in chemical and mineralogical composition, the complex organic and chemical origins of carbonate sediments lead to a wide range of textures and fabrics in the resulting limestones. These textures and fabrics share significant parallels with those found in siliciclastic rocks, and they are quite useful for the classification and determination of depositional environments for limestones. Limestones and dolomites are used commercially as building materials (**Fig. 1**) and as a source for industrial and agricultural lime. In addition, limestones and dolomites are important reservoirs for oil and gas and are the hosts for important mineral deposits, including lead, zinc, silver, and fluorite. *See* LIME (INDUSTRY); ORE AND MINERAL DEPOSITS; PETROLEUM GEOLOGY; STONE AND STONE PRODUCTS.

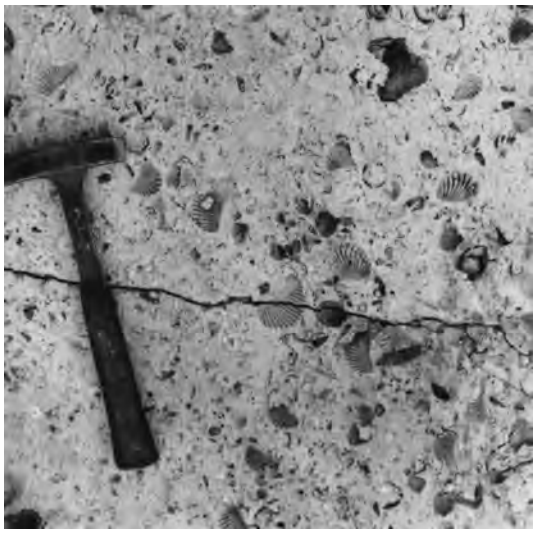

Fig. 1. Surface of a Lower Cretaceous limestone exposed in a quarry just north of Austin, Texas. The limestone is composed predominantly of fossil fragments. The large ribbed fossils are *Trigonia* sp., a common Cretaceous mollusk. This limestone is used extensively as a decorative building stone.

Chemical composition. Limestones are composed principally of calcium oxide (CaO) and carbon dioxide (CO_2). They may contain magnesium oxide (MgO) of up to several weight percent, in which case they are termed magnesian limestones. Limestones constitute the single largest carbon reservoir. Iron (Fe) may be present as an oxide or may (in its reduced state) substitute for magnesium (Mg) or calcium (Ca) in the calcite lattice. Iron commonly substitutes for magnesium in the dolomite lattice. Iron may also occur in association with limestones as a carbonate (siderite) or as a sulfide (pyrite). Strontium (Sr) is a common, important trace element in limestones, occurring at a level of 100–200 ppm. Strontium is a major trace element in aragonite skeletons of shallow marine, limestone-forming organisms (6000–10,000 ppm). Strontium is released into diagenetic waters during the conversion of aragonite into calcite, and is used extensively as a tracer in diagenetic studies of limestones. Isotopic composition of limestones and limestone components, such as fossil fragments, varies widely in response to rock–water interaction, during the extensive diagenetic changes suffered by most limestones as a consequence of their high chemical reactivity. Stable isotopes of oxygen, carbon, and sulfur, and the radiogenic isotopes of strontium are commonly used in diagenetic

studies of limestones. *See* ARAGONITE; CALCITE; DIAGENESIS; PYRITE; SIDERITE.

Mineralogy. Calcite and dolomite are the chief minerals found in ancient limestones and dolomitic limestones. Modern shallow marine carbonate sediments, however, are composed of aragonite, a more soluble polymorph of calcite, and magnesian calcite, a highly soluble form of calcite containing 4–16 mol % $MgCO_3$. This dramatic difference between the mineralogy of carbonate sediments and the mineralogy of ancient limestones suggests that during lithification processes leading to the formation of limestone and dolomitic limestone, a major mineralogical conversion, from relatively soluble aragonite and magnesian calcite to more stable calcite and dolomite, is accomplished. Studies have indicated that aragonite and magnesian calcite sediments were present throughout the Phanerozoic.

This stabilization process takes place relatively early in the burial history of carbonate sediments and generally is completed during the first million years of the limestone's history, as observed by the progressive loss of aragonic and magnesian calcite in Pleistocene limestones. The mineralogical stabilization process is a major factor in early freshwater limestone diagenesis and is extremely important in the porosity evolution of ancient limestone oil and gas reservoirs.

Deep-water carbonate sediments, however, consist of the skeletal remains of planktonic organisms composed of calcite, and they do not undergo the massive mineralogical conversion of their shallow marine counterparts.

Ancillary minerals commonly found in association with limestones include dolomite [$Ca(CO_3)_2$]; celestite ($SrSO_4$); strontianite ($SrCO_3$); anhydrite ($CaSO_4$); fluorite (CaF_2); pyrite (FeS_2); and quartz (SiO_2). *See* ANHYDRITE; CARBONATE MINERALS; CELESTITE; DOLOMITE; FLUORITE; MINERAL; QUARTZ; STRONTIANITE.

Origin. Most marine limestones (perhaps 90% or more) originate as calcium carbonate skeletal elements of various organisms, including both plants (marine algae such as *Lithothamnion* and phytoplankton such as coccoliths) and animals (such as corals, clams, snails, and oysters; Fig. 1). The larger organisms are broken down into cobble-to-silt-sized sediments by biological processes, such as boring, browsing, and grazing, in the environment. Many of the calcareous algae, such as *Halimeda*, and the smaller animals, such as juvenile formanifera, break down into mud-sized carbonate sediments upon death. Once formed, these sediments react to environmental processes as do their siliciclastic counterparts. High-energy conditions in carbonate depositional environments such as a tidal channel are marked by cross-bedded carbonate sands. Low-energy carbonate environments, such as a sheltered lagoon, are characterized by burrowed mud-dominated sediments. Calcareous organisms of pelagic environments are dominated by tiny phytoplankton (coccoliths) and zooplankton (foraminifera). Deep marine limestones, therefore, are very fine-grained and are generally termed chalks. Coccoliths and calcareous pelagic foraminifera became dominant in the Mesozoic. In the Paleozoic, the plankton were dominated by siliceous organisms, so that Paleozoic pelagic sequences are composed of massive cherts. *See* ALGAE; CHALK; CHERT; COCCOLITHOPHORIDA; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; FORAMINIFERIDA; MARINE SEDIMENTS; MARL; STRATIGRAPHY.

Some limestones and limestone components are formed by direct chemical precipitation from marine and meteoric waters. Most modern, tropical, marine surface water is supersaturated with respect to calcium carbonate. If CO_2 is removed from this water by warming, agitation, or photosynthesis, there is a tendency for calcium carbonate to be precipitated. This precipitation can take several forms: an aragonite or magnesian calcite cement, which lithifies carbonate sediment, such as the beach rock commonly found along tropical beaches; an aragonite precipitate on a moving nucleus in a high-energy environment, forming highly polished, round, sand-sized particles termed ooids; or clouds of spontaneously precipitated, clay-sized aragonite, forming on shallow carbonate platforms or in restricted bays. *See* OOLITE.

Finally, some limestones are formed in freshwater environments associated with caves (speleothems, such as stalactites and stalagmites), springs (tufa and travertine), and lakes (almost always chemically precipitated fine muds of calcite, dolomite, or alkali-carbonates). *See* CAVE; STALACTITES AND STALAGMITES; TRAVERTINE; TUFA.

Consequences of biological origin. Most carbonate sediments are forming in place in the depositional environment and generally suffer little transport. This fact, coupled with very high rates of sediment production, generally results in vertical sedimentation and in the common stacking of shoaling-upward cycles on carbonate-dominated marine shelves. This pattern of sedimentation is in direct contrast with siliciclastics, where the source of sediments is outside the basin of deposition, sediment transport to the basin is extremely important, and shoreline progradation is the major pattern of siliciclastic sedimentation on the shelf.

Textures and fabrics in limestones are much more difficult to interpret than in siliciclastics, because of the organic origin of most carbonate grains. While grain size distribution in siliciclastics is controlled by the flow velocity at the site of deposition, grain size distribution in carbonates may be controlled by the types of organisms present in the environment that furnishes the grains. As an example, an environment dominated by large mollusks will tend to produce a sediment characterized by coarse grain sizes, whereas a benthic foraminiferal community will tend to produce grain sizes that are much finer. Roundness in siliciclastic deposits may be used to infer transport and depositional processes. Roundness in the individual grains of a limestone, however, may reflect only the original shape of the organism (that is, a foraminifer or gastropod) or the architecture of its

Classification of carbonate rocks according to depositional texture*

Depositional texture recognizable					Depositional texure not recognizable
Original components not bound together during deposition				Original components were bound together during deposition as shown by intergrown skeletal matter, lamination contrary to gravity, or sediment-floored cavities that are roofed over by organic or questionably organic matter and are too large to be interstices	
Contains mud (particles of clay and fine silt size)			Lacks mud and is grain-supported		Crystalline carbonate
Mud-supported		Grain-supported			
Less than 10% grains	More than 10% grains				(Subdivide according to classifications designed to bear on physical texture or diagenesis)
Mudstone	Wackestone	Packstone	Grainstone	Boundstone	

*After R. J. Dunham, Classification of carbonate rocks according to depositional texture, in W. E. Ham (ed.), *Classification of Carbonate Rocks,* Amer. Ass. Petrol. Geol. Mem. 1, pp. 108–121, 1962.

skeleton. For example, coral septa are composed of spherical bodies. Coral skeletons tend to break down into these spheres during biological erosion after the death of the coral.

Finally, carbonate environments of deposition are characterized by unique faunal and floral communities. Because these organisms give rise to the sediments that ultimately form limestones, the grain composition of a detrital limestone composed of recognizable fossil fragments may directly reflect its environment of deposition. *See* FOSSIL.

Reefs. Certain carbonate-secreting organisms display ability to modify their environment by encrusting, binding, and building a solid organic framework that possesses the strength to resist wave attack and may act as a trap for sediments being transported across the shelf. These important solid limestone structures are known as reefs. **Figure 2** illustrates a schematic reef structure and its major components. Major reef builders today consist of corals and coralline algae. In the past, important reef formers included mollusks, sponges, and algae. Reef masses are important reservoirs for oil and gas. *See* CORALLINALES; MOLLUSCA; REEF.

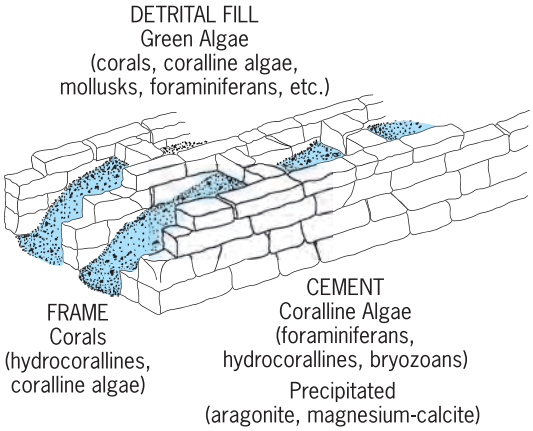

Fig. 2. Schematic rendering of a reef illustrating the complex interaction of biologic framework growth, sedimentation, and chemical precipitation in the formation of reef limestones. (*After R. N. Ginsberg and H. A. Lowenstam, The influence of marine bottom communities on the depositional environment of sediments, J. Geol., 66:310–318, 1958*)

Consequences of high chemical reactivity. Limestones are considerably more soluble and hence maintain a much higher chemical reactivity than their siliciclastic counterparts during burial and lithification. This higher chemical reactivity is reflected in the propensity of limestones to exhibit dramatic, early, surface-related, diagenetic overprints, including extensive dissolution, dolomitization, and massive, pore-filling calcite cements. In contrast, siliciclastics generally exhibit little early diagenetic activity and are buried into the subsurface relatively intact. During subsequent burial, pore space is lost much more rapidly in limestones than in siliciclastics because of pressure solution resulting from steadily increasing lithostatic pressures. Limestones dissolve more readily under pressure because of the high relative solubility of calcite as compared to quartz. As a result, successful exploration for oil and gas in limestone reservoirs has a severe depth constraint not shared by their siliciclastic counterparts.

Classification. The first practical, widely used limestone classification was developed by R. L. Folk in 1959. This classification encompasses the common textural classifications used in siliciclastics, as well as a strong grain compositional component. Folk's classification is complex, is difficult to use in detail, and generally is best used with a petrographic microscope. The most widely used classification is the simple, field-oriented, textural classification published by R. Dunham in 1962 (see **table**). Dunham's major rock classes are based on the presence or absence of organic binding, presence or absence of mud, and the concept of grain versus mud support. The first four major rock types, mudstone, wackestone, packstone, and grainstone, directly parallel common siliciclastic textural divisions and represent a gradient from low energy (mudstone) to high energy (grainstone). The boundstone category emphasizes the unique characteristics of organic

binding and framework construction. *See* DOLOMITE ROCK; SEDIMENTARY ROCKS; SEDIMENTOLOGY.

Clyde H. Moore

Bibliography. R. G. C. Bathurst, *Carbonate Sediments and Their Diagenesis*, 2d ed., 1975, *Developments in Sedimentology*, vol. 12, 1975; L. Glaister, *Limestone and Clay*, 1994; P. A. Scholle, D. G. Bebout, and C. H. Moore (eds.), *Carbonate Depositional Environments*, Amer. Ass. Petrol. Geol. Mem. 33, 1983; J. L. Wilson, *Carbonate Facies in Geologic History*, 1975.

Limiter circuit

A device whose purpose is to ensure that the amplitude of a sensed variable (referred to as a signal) is constrained or limited to lie within prescribed maximum and minimum values. It is more properly termed an amplitude limiter, although convention usually dispenses with the important modifier. Also, unless specified otherwise, the limiting is applied to voltage.

Function. Strictly speaking, the limiter should behave as a perfectly linear device (such as an ideal amplifier or even a piece of wire) until the amplitude of the input signal reaches the upper or lower limit values. In other words, the output of an ideal limiter matches the input perfectly when the input is between the limit values. If the input is greater than the upper limit, the output equals the upper limit; if the input is less than the lower limit, the output equals the lower limit. If the upper and lower limits are equal in magnitude, the limiting is symmetrical. In the past, one or two stages of high-gain amplification would often precede a limiter in order to obtain clean limiting. *See* LINEARITY.

The term limiter is often used to mean sign detector, so that when the input is greater than zero, the output equals some fixed positive value; when the input is less than zero, the output is the negative of this value. This usage is sometimes employed in discussing frequency-modulation (FM) radio receivers.

Mechanization. Discrete-component limiters have become rare. The cost/performance ratio of integrated circuits makes them the only rational choice for commercial applications. Most often, integrated circuits now contain limit operations that are both functionally and physically integrated with companion operations to form a higher-level function. The isolated need for a limit-only function is almost never encountered. Limiting within an integrated circuit is usually obtained by designing an amplifier to be heavily overdriven and therefore to quickly enter the saturation (limiting) range. *See* AMPLIFIER; INTEGRATED CIRCUITS.

The conceptual nonlinear mechanics of limiting are very simple. An ideal situation (**illus.** *a*) can be approximated with readily available components. The signal voltage is generated by a source, which has a resistance. For the sake of clarity, reactance will be ignored; that is, it will be assumed that all impedances are resistive. Two diodes are each back-biased by voltage sources. The backward resistance of each diode is much greater than the source resistance; and in turn, the source resistance is much greater than the forward resistance of each diode. *See* DIODE.

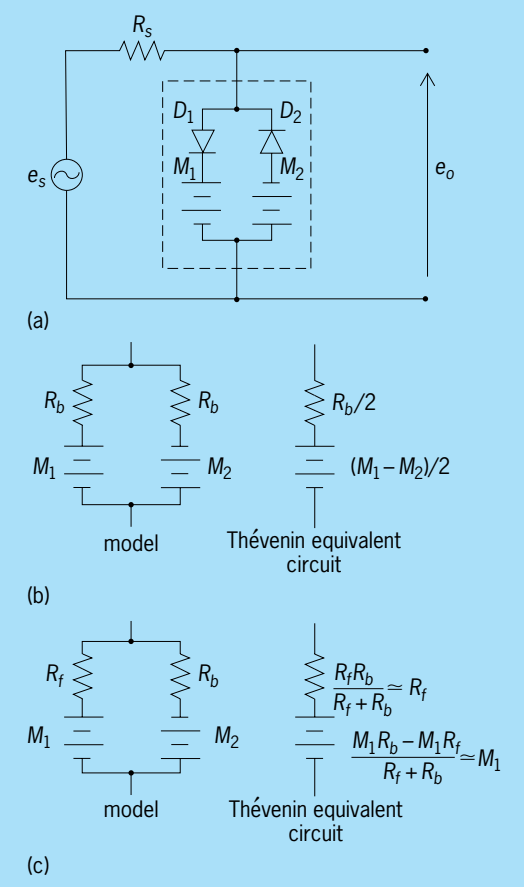

Limiter circuits. (*a*) Generic limiter. The broken line encloses the part of the circuit described by models. e_s = signal voltage; R_s = source resistance; e_o = output voltage. Diodes D_1 and D_2 are back-biased by voltage sources M_1 and M_2, respectively. (*b*) Model and Thévenin equivalent circuit for diodes and their biasing voltage sources when both diodes are back-biased. R_b is the back resistance of both diodes. (*c*) Model and Thévenin equivalent circuit for diodes and their biasing voltage sources when diode D_1 is forward-biased and D_2 is back-biased. R_f is the forward resistance of both diodes.

In order to easily and simply determine the operation of the circuit, the two diodes and two bias sources are combined into an equivalent circuit (consisting of the series combination of one resistor and one voltage source, as given by Thévenin's theorem) for the case when both diodes are back-biased (illus. *b*), and for the case when the first diode is forward-biased and the second diode is back-biased (illus. *c*). If the second diode is forward-biased and the first diode is back-biased, the roles of its corresponding voltage source and the negative of the other voltage source are exchanged. Never can both diodes be forward-biased simultaneously. *See* THÉVENIN'S THEOREM (ELECTRIC NETWORKS).

When both diodes are back-biased (illus. *b*), the output voltage is approximately equal to the signal voltage. When the first diode is forward-biased (illus. *c*), the output voltage is very nearly equal to its corresponding voltage source. Similarly, when the second diode is forward-biased, the output voltage is very nearly equal to the negative of its corresponding voltage source.

Virtually all limiter circuits, upon final analysis, reduce to this simple configuration. Digital limiting can be performed by a comparator and a logic statement.

Limiting for FM reception. The operation of FM detectors (discriminators) requires that the amplitude envelope of the frequency-modulated information signal be held nearly constant. The purpose of this limitation is to remove (or at least seriously reduce) additive noise and amplitude modulation from the signal. These unwanted envelope effects produce undesirable noise components at the output of the FM receiver. Some FM detectors provide internal limiting, but others require that the signal envelope be separately amplitude-limited.

If the FM detector is of the slope-detector type, additional filtering of the amplitude-limited signal is necessary. Because the limiting action modifies the signal wave shape, the frequency content of the signal is also changed by the creation and addition of high-frequency distortion components. Because the frequency content of the signal before limiting consists of the information energy tightly clustered around the carrier frequency, adequate waveform correction and high-frequency removal can be performed by applying the amplitude-limited signal across the terminals of a simple resonant circuit which has been tuned to the carrier frequency. *See* DISTORTION (ELECTRONIC CIRCUITS).

The effectiveness of the limiting is measured by a quieting criterion. When the input signal to the receiver is low, the limiting action does not occur and the receiver output is noticeably noisy. As the strength of the signal increases and limiting action begins to occur, the reduction in the noise level is immediately noticeable. As the strength of the signal continues to grow, the output signal-to-noise ratio continues to grow correspondingly. The level of input signal necessary to achieve a 30-decibel signal-to-noise ratio is called the receiver's sensitivity. Of course, extremely large input signals can also create problems. In practice, the limiter provides automatic-gain-control effects; however, separate automatic-gain-control circuits in the early (radio-frequency) stages of the receiver often are used to regulate the input signal to the limiter by boosting the amplitude of the signal so that amplitude lies in the most effective range for the limiter. *See* AUTOMATIC GAIN CONTROL (AGC); FREQUENCY-MODULATION DETECTOR; SIGNAL-TO-NOISE RATIO.

Stanley A. White

Bibliography. A. B. Carlson, *Communications Systems*, 3d ed., 1986; R. C. Dorf (ed.), *The Electrical Engineering Handbook*, 2d ed., 1997; D. Roddy and J. Coolen, *Electronic Communications*, 4th ed., 1995; R. A. Williams, *Communication System Analysis and Design*, 1987.

Limits and fits

The extreme permissible values of a dimension are known as limits. The degree of tightness or looseness between two mating parts that are intended to act together is known as the fit of the parts. The character of the fit depends upon the use of the parts. Thus, the fit between members that move or rotate relative to each other, such as a shaft rotating in a bearing, is considerably different from the fit that is designed to prevent any relative motion between two parts, such as a wheel attached to an axle.

In selecting and specifying limits and fits for various applications, it is essential in the interests of interchangeable manufacturing that (1) standard definitions of terms relating to limits and fits be used; (2) preferred basic sizes be selected wherever possible to reduce material and tool costs; (3) limits be based upon a series of preferred tolerances and allowances; and (4) a uniform system of applying tolerances (bilateral or unilateral) be used.

Nominal size. The basic size or dimension of a screw thread or machine part is the theoretical nominal standard size from which variations are made. For example, a shaft may have a basic diameter of 2.000 in. with a maximum variation of ±0.010 in. (5.060 ± 0.025 cm). The nominal size is generally the size used for purposes of general identification (see **illus.**)

Tolerance. A tolerance is the amount of variation permitted on dimensions or surfaces of machined parts. The tolerance is equal to the difference between the maximum and minimum limits of any specified dimension. For example, if the maximum limit for the diameter of a shaft is 2.000 in. (5.060 cm) and the minimum limit is 1.990 in. (5.055 cm), the tolerance for this diameter is 0.010 in. (0.025 cm). Any component-part tolerance is mainly decided by the functional requirement of the part as well as the capability of the manufacturing process.

A unilateral tolerance means that the total tolerance, as related to a basic dimension, is in one direction only. For example, if the basic dimension is 1.000 in. (2.540 cm) and the tolerance is expressed as 1.000 − 0.002, or as 1.000 + 0.002, these would be unilateral tolerances, because the total tolerance in each case is in one direction. If the tolerance is divided into plus and minus, for example, 1.000 ± 0.001, it is classified as a bilateral tolerance because the total tolerance of 0.002 is given in two directions. The limits (maximum and minimum) are obtained by adding or subtracting the tolerance from the basic dimension.

Allowance. An allowance is a prescribed difference between the maximum material limits of two mating parts. It is a minimum clearance (positive allowance) or a maximum interference (negative

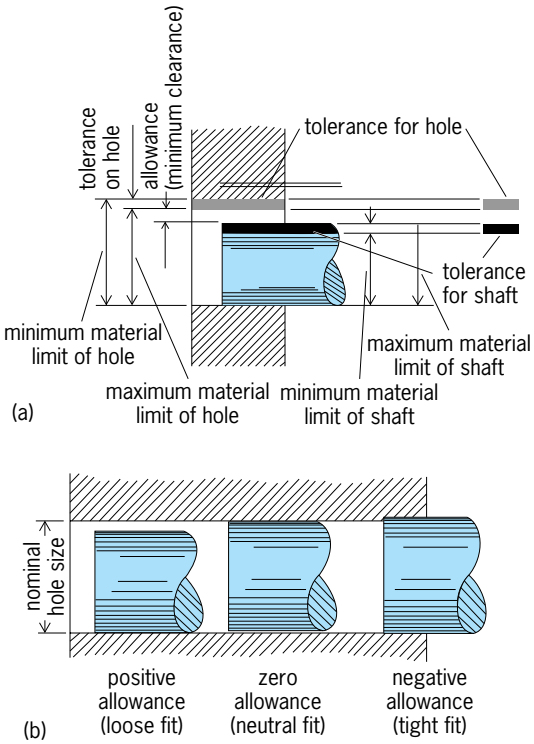

Limits and fits in machine parts. (*a*) Limits, tolerances, and allowance for a shaft. (*b*) Three types of allowances.

allowance) between the parts. Fit is the general term used to signify the range of tightness that may result from the application of a specific combination of allowances and tolerances in the design of mating parts. The actual fit between two mating parts is the relationship existing between them with respect to the amount of clearance or interference that is present when they are assembled.

Type of fit. In selecting limits of size for any application, the type of fit is determined based upon the use or service of the parts being designed. Then the limits of size of the mating parts are established to ensure that the desired fit will be produced. In general, fits may be considered as positive, negative, or zero allowances (illus. *b*).

The classes of fits are arranged in three general groups: running and sliding fits, locational fits, and force fits. Running and sliding fits are intended to provide a similar running performance with suitable lubrication allowances throughout the range of sizes. They range from close sliding fits (for the accurate location of parts that must be assembled without perceptible play between them) to loose running fits, where wide commercial tolerances are acceptable. Locational fits are intended to determine only the location of the mating parts. They may provide rigid or accurate location as in the case of interference fits, or some freedom of location as with clearance fits. Accordingly, they are divided into three groups: clearance fits, transition fits, and interference fits. Force or shrink fits constitute a special type of interference fit normally characterized by maintenance of constant bore pressures throughout the range of sizes. The interference varies almost directly with the diameter of the part and the difference between its minimum and maximum value in order to maintain the resulting pressure between the two mating parts within reasonable limits. The range is from light drive fits that require light pressure, to force fits that are more suitable for parts that can be highly stressed by a given force or load applied to the two parts. *See* DESIGN STANDARDS; ENGINEERING DESIGN; MACHINE DESIGN; MACHINE ELEMENTS; MACHINING. Joseph Elgomayel

Bibliography. E. A. Avallone and T. Baumeister III (ed.), *Marks' Standard Handbook for Mechanical Engineering*, 10th ed., 1996; A. Parrish (ed.), *Mechanical Engineer's Reference Book*, 12th ed., 1994; H. A. Rothbart, *Mechanical Systems Reference Guide*, 1989; R. A. Walsh, *McGraw-Hill Machining and Metalworking Handbook*, 2d ed., 1998.

Limnology

The study of lakes, ponds, rivers, streams, swamps, and reservoirs that make up inland water systems. Each of these inland aquatic environments is physically and chemically connected with its surroundings by meteorologic and hydrogeologic processes (**Fig. 1**). Precipitation and runoff, combined with the gradient and watershed characteristics of a river or stream, provide the physicochemical environment

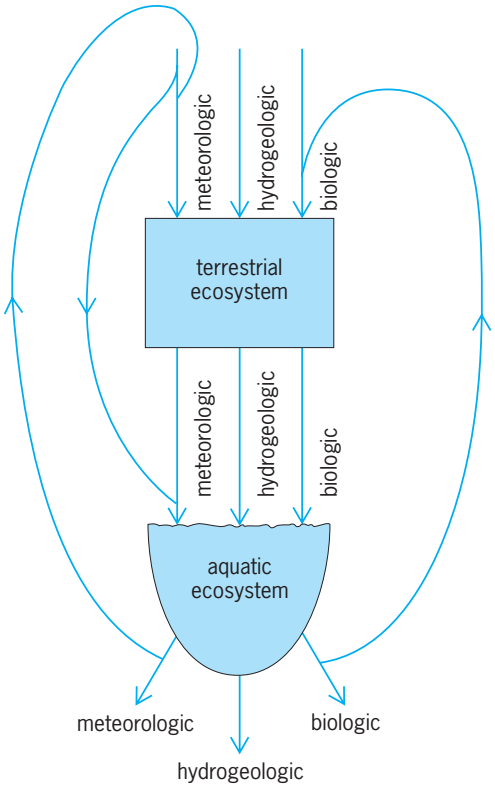

Fig. 1. Diagrammatic model of the functional linkages between terrestrial and aquatic ecosystems. Vectors may be meteorologic, hydrogeologic, or biologic components moving nutrients or energy along the pathway shown.

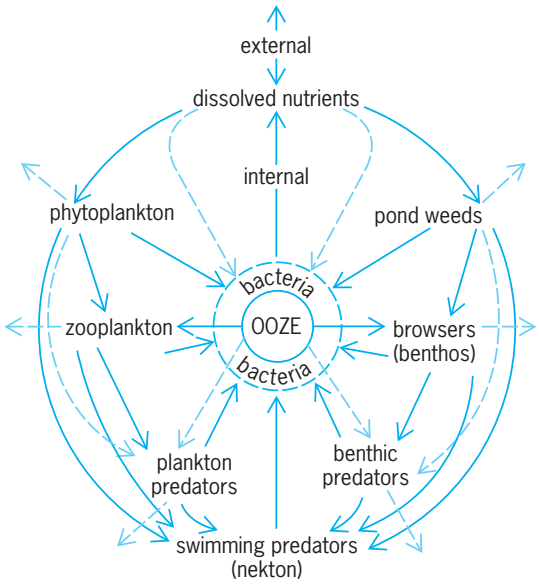

Fig. 2. Diagram of food webs in a lake, showing the interdependence of organisms. Broken lines depict the energy pathways and solid lines the food source.

for the organisms of flowing water (lotic) systems. The frequency and stability of thermal stratification and mixing in standing water (lentic) systems such as lakes, ponds, and reservoirs are determined by the seasonal balance of heating and cooling, the unique temperature-density characteristics of pure water, and the energy of the wind.

Solar energy supply and distribution, physical mixing processes, and nutrients supplied by the surrounding watershed and lake sediments determine how much organic matter is produced by photosynthesis in aquatic environments. That autotrophic production, combined with leaves and other organic matter produced outside the system, supports heterotrophic activities of aquatic communities.

Biogeochemical processes. Food webs are used to illustrate interactions between plants and animals in aquatic communities. The trophic-dynamic approach (**Fig. 2**) traces the energy fixed by photosynthesis through trophic levels from herbivores to top carnivores. The flow of energy and the return of minerals to inorganic nutrient pools maintain internal biogeochemical cycles of aquatic environments. *See* BIOGEOCHEMISTRY; FOOD WEB.

General habitats of aquatic environments. Aquatic organisms are grouped into categories according to where they live within a particular system. Habitats of aquatic systems are defined by unique physical conditions. The open water (limnetic zone) of lentic environments is inhabited by microscopic plants (phytoplankton) and animals (zooplankton). Shallows around the edge of the lake that are limited by light penetration (littoral zone) are occupied by rooted plants (aquatic macrophytes). These plants provide habitats for other plants and animals. The deepest area (profundal zone) is bounded by the limnetic zone on the top and the littoral zone on the sides. Benthic organisms such as insects, oligochaetes, and microorganisms occupy the profundal zone. Insects and free-swimming animals such as fish (nekton) are not restricted to any particular zone of a lentic environment. The feeding behavior of fish and other carnivores influences many food web interactions by removing large herbivores from the food web. The loss of large herbivores alters feeding relationships at the base of the food web, because small herbivores cannot eat large phytoplankton. This cascade of feeding interactions alters the structure of aquatic communities by permitting the accumulation of large phytoplankton and small herbivores. *See* PHYTOPLANKTON; ZOOPLANKTON.

Lotic environments also are divided into zones created by water velocity, bottom conditions, and stream gradients. In mountainous areas, rapids and turbulent water zones (riffles) are contrasted with sluggish water zones (pools); each of these habitats has characteristic organisms. Streams and rivers in valleys and on plains have many meanders, oxbows, and braids developed by the balance of erosion and

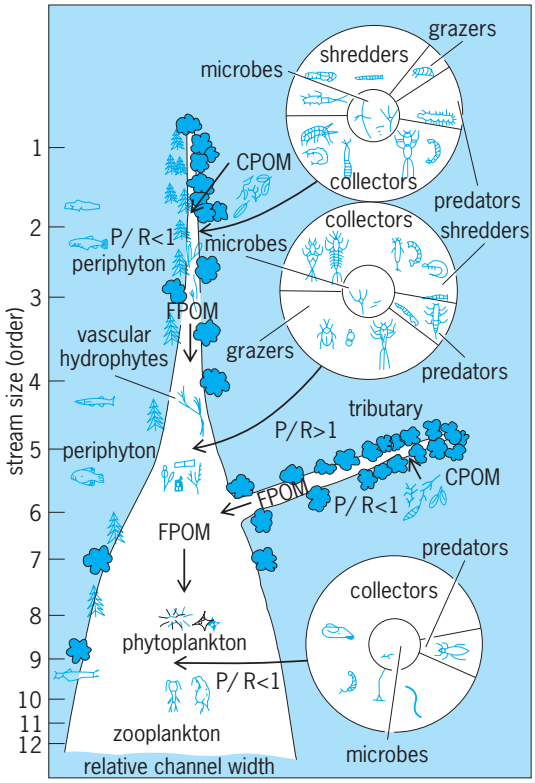

Fig. 3. A proposed relationship between stream size and the progressive shift in structural and functional attributes of lotic communities. Coarse particulate organic material (CPOM) that enters a stream from the watershed is shredded by organisms to form fine particulate organic matter (FPOM). Some organisms collect the fine material, and others graze on autotrophic organisms to sustain the food web of the stream. Production (P) and respiration (R) responses are used to detect changes in autotrophic and heterotrophic processes in the biotic community. (*After R. L. Vannote et al., The river continuum concept, Can. J. Fish. Aquat. Sci., 37:130–137, 1980*)

deposition of the riverbed. When studying the linkage between streams and their terrestrial setting, scientists consider the gradient of physical factors formed by drainage networks. Fluvial geomorphic processes are then integrated with community ecology principles to develop an understanding of functional processes that contribute to mineral cycling and organic matter utilization (**Fig. 3**). *See* FRESHWATER ECOSYSTEM.

Eutrophication and toxic materials. Aquatic systems with excellent physical conditions for production of organisms and high nutrient levels often show signs of eutrophication. Eutrophic lakes are generally identified by large numbers of phytoplankton and aquatic macrophytes and by low oxygen concentrations in the profundal zone. Eutrophication is accelerated by poor watershed management, by runoff from overfertilized lands, and by poor septic facilities. *See* EUTROPHICATION; WATER CONSERVATION.

Organisms of aquatic environments are also affected by metals such as lead and mercury, or pesticides that enter the aquatic environment from poor watershed management practices. These toxic materials alter food webs of aquatic systems by destroying sensitive organisms. Tolerant organisms concentrate toxic elements by feeding and absorption. Toxic elements are concentrated in tolerant organisms by a process called biomagnification. These organisms acquire levels of toxicity that make them dangerous for human consumption. Lakes and rivers are also influenced by changes in the chemical composition of rain. Acid runoffs cause stress to the biota of aquatic systems in many ways, including direct toxicity of hydrogen ions, disruption of normal food web relations, alterations of behavioral patterns of animals, and modification of biogeochemical cycles. *See* ACID RAIN; ECOLOGY; HYDROLOGY; LAKE; RIVER; WATER POLLUTION. James E. Schindler

Bibliography. S. R. Carpenter (ed.), *Complex Interactions in Lake Communities*, 1988; A. J. Horne and C. R. Goldman, *Limnology*, 2d ed., 1994; D. W. Schindler et al., Long-term ecosystem stress: The effects of years of experimental acidification on a small lake, *Science*, 228:1395–1401, 1985; J. V. Ward, *Aquatic Insect Ecology*, 1992; R. G. Wetzel and G. E. Likens, *Limnological Analysis*, 3d ed., 2000.

Limonite

A field or generic term for natural hydrous iron oxides, the most common phase being the mineral goethite, α-FeO(OH). Limonite occurs as a low-temperature mixture of phases and includes the so-called bog iron ores. It is the characteristic brown stain which coats rocks containing sulfide ores, such as pyrite and pyrrhotite, in the zone of weathering of these ores referred to as a gossan. A vast list of pseudonyms, such as limnite, hypoxanthite, and xanthosiderite, plague the earlier literature. Limonite is most commonly a mixture of greater or lesser fractions of goethite, hematite, ferric oxyhydroxide colloids, clays, and manganese oxides. It is formed by biogenic or inorganic precipitation in bog, spring, lacustrine, or marine deposits. *See* GOETHITE; IRON METALLURGY; ORE AND MINERAL DEPOSITS. Paul B. Moore

Limpet

Name given to a variety of species of aquatic gastropod mollusks, all with a characteristic conical shell (see **illus.**) and a suckerlike foot. Limpets (like chitons) are well adapted for life on rocky surfaces exposed to wave action and, in the higher levels of the littoral zone, to alternating tidal submergence and aerrial drying. Apparently the coiled gastropod shell has evolved in many different stocks into the conical shell with a wide opening (the limpet form) by a rapid enlargement of the mantle-shell edge so that the snail is completely accommodated in one open half-whorl. In many cases the initial coiling (recognizable in larval stages) is entirely lost in the adult shell, which is a wide cone. In a few others, initial coiling (termed a protoconch) remains obvious at the apex of the shell throughout adult life. The gastropod crawling foot is greatly enlarged and flattened into a round muscular sucker. All limpets move relatively slowly over rock or other hard surfaces, protecting themselves (against wave action or desiccation or predation) by clamping down the shell opening against the substrate by contraction of the enlarged shell muscles. There is never an operculum on the foot, and limpets are defenseless once detached.

Although not all limpets retain a recognizable anterior mantle cavity, all have a pallial groove between the circular outer edge of the foot and the concentric mantle edge lining the shell. This mantle edge may bear obvious sensory tentacles, but it and the pedal edge always contain tactile receptors and chemoreceptors (used by several different stocks of limpets in "homing" behavior). Limpets usually graze slowly and continously by radular scraping of attached algae and diatoms from rock surfaces. The structural and functional adaptations of the limpet form have arisen in many distinct groups of gastropods.

Limpets among the most primitive snails (those Archaeogastropoda with two aspidobranch ctenidia) include slit shells (*Emarginula*) and keyhole limpets

***Diodora aspera*, the rough-keyhole limpet of the Pacific coast of North America. All limpets have this shell shape as a widely open cone, but only a few genera have an apical "keyhole."**

(*Fissurella*). Other limpets, belonging to archaeogastropod stocks with marked asymmetry of gills and internal organs but with no openings or slits in the conical shell, include *Acmaea* (the most common limpet of the American Atlantic coast), which has a single aspidobranch ctenidium in a small anterior mantle cavity, and *Patella* (the dominant European limpet genus), which has no true ctenidium but a series of respiratory outgrowths (secondary or neomorphic gills) along each pallial groove. The closely related giant owl limpet (*Lottia*) of Californian shores, which can reach a diameter of 4 in. (10 cm), has both a ctenidium and neomorphic gills. Such acmaeids and patellids of the intertidal are often referred to as "true limpets." The more distantly related archaeogastropod family Lepetidae comprises rarer species of tiny limpets (up to 0.3 in. or 8 mm) without any gills, found only offshore in depths of more than 20 ft (6 m).

In the Mesogastropoda (with a pectinibranch ctenidium, or comblike "half-gill" with its axis fused to the mantle-cavity wall), limpets include the filter-feeding slipper limpets (*Crepidula*) and the fool's-cap limpets or capulids, some genera of which have become ectoparasites on echinoderms and on other mollusks. Among the higher gastropods, the subclass Opisthobranchia has a few limpet-shelled forms (*Tylodina*, *Umbraculum*), and the subclass Pulmonata has at least five distinct lines of limpets. Two of these are intertidal: in the Trimusculidae (=Gadiniidae) and Siphonariidae or false limpets, common on Gulf of Mexico and Caribbean shores, the air-breathing lung has been supplemented by secondary gills developed from its roof. Finally there are at least three stocks of fresh-water pulmonate limpets. These include the stream limpets, such as *Ancylus fluviatilis* in Europe and *Ferrissia rivularis* in North America, found abundantly clinging to rocks in swiftly flowing streams (even in waterfalls), and the pond limpets, such as *Acroloxus lacustris* in Europe and *Laevapex fuscus* in North America, which live on reed stalks. All such fresh-water limpets are derived from air-breathing "diving" snails which have developed, along with the limpet form, secondary (neomorphic) gills derived, in part, from ciliated rectal tissues. Outside gastropod stocks, there are the untorted limpets of the class Monoplacophora, and in the class Bivalvia the anomiid saddle oyster (*Enigmonia*), which inhabits mangrove swamps, has resumed a crawling habit on an enlarged foot as a bivalve turned "limpet," although it still filter-feeds as a typical lamellibranch.

This acquisition, occurring independently in so many diverse stocks of gastropods, of the structural and functional adaptations which make up the limpet form provides, as does the similar polyphyletic occurrence of certain structural adaptations to specific habitats in bivalves, a striking example of evolutionary convergence. *See* BIVALVIA; GASTROPODA; MOLLUSCA; MONOPLACOPHORA; OPISTHOBRANCHIA; PROSOBRANCHIA; PULMONATA.

W. D. Russell-Hunter

Line integral

The line integral of a vector function **F** of position over a path C is represented by Eq. (1), where F_x, F_y,

$$\int_C \mathbf{F}\cdot d\mathbf{r} = \int_C F_x(x,y,z)\,dx + \int_C F_y(x,y,z)\,dy + \int_C F_z(x,y,z)\,dz \qquad (1)$$

F_z are the scalar components of **F** along the coordinate axes. The path C is supposed to be a curve, smooth at least in part, defined parametrically by equations of form (2) for each smooth portion. The

$$x = x(p) \qquad y = y(p) \qquad z = z(p) \qquad (2)$$

functions $F_x(x,y,z)$, etc., must be defined at all points of C. When this is so, the line integral can be evaluated by writing Eq. (3), where the prime means dif-

$$\int_c \mathbf{F}\cdot d\mathbf{r} = \int_{p_1}^{p_2} F_x(p)x'(p)dp + \int_{p_1}^{p_2} F_y(p)y'(p)dp + \int_{p_1}^{p_2} F_z(p)z'(p)\,dp \qquad (3)$$

ferentiation with respect to the parameter, and p_1, p_2 are values of the parameter at the end points of the path C or of a smooth piece. The integral has been converted into an ordinary definite integral.

When C is a closed curve, the line integral is called a circuit integral, and is written as notation (4).

$$\oint \mathbf{F}\cdot d\mathbf{r} \qquad (4)$$

Some physical applications of the line integral follow. If **F** is a force, the line integral is the work done in moving a mass along the curve C. If **F** is the velocity of flow of a fluid, the line integral is the circulation of the fluid along the curve. If **F** is the electrostatic field strength, the integral is the electric potential difference between the end points of the curve. If **F** is the electric field strength of an electromagnetic field, the circuit integral is the electromotive force of the circuit. In each example $d\mathbf{r}$ is physically a length. *See* INTEGRATION.

McAllister H. Hull, Jr.

Line spectrum

A discontinuous spectrum characteristic of excited atoms, ions, and certain molecules in the gaseous phase at low pressures, to be distinguished from band spectra, emitted by most free molecules, and continuous spectra, emitted by matter in the solid, liquid, and sometimes gaseous phase. If an electric arc or spark between metallic electrodes, or an electric discharge through a low-pressure gas, is viewed through a spectroscope, images of the spectroscope slit are seen in the characteristic colors emitted by the atoms or ions present.

To avoid the overlapping of close spectral images, the slit illuminated by the light source is made very

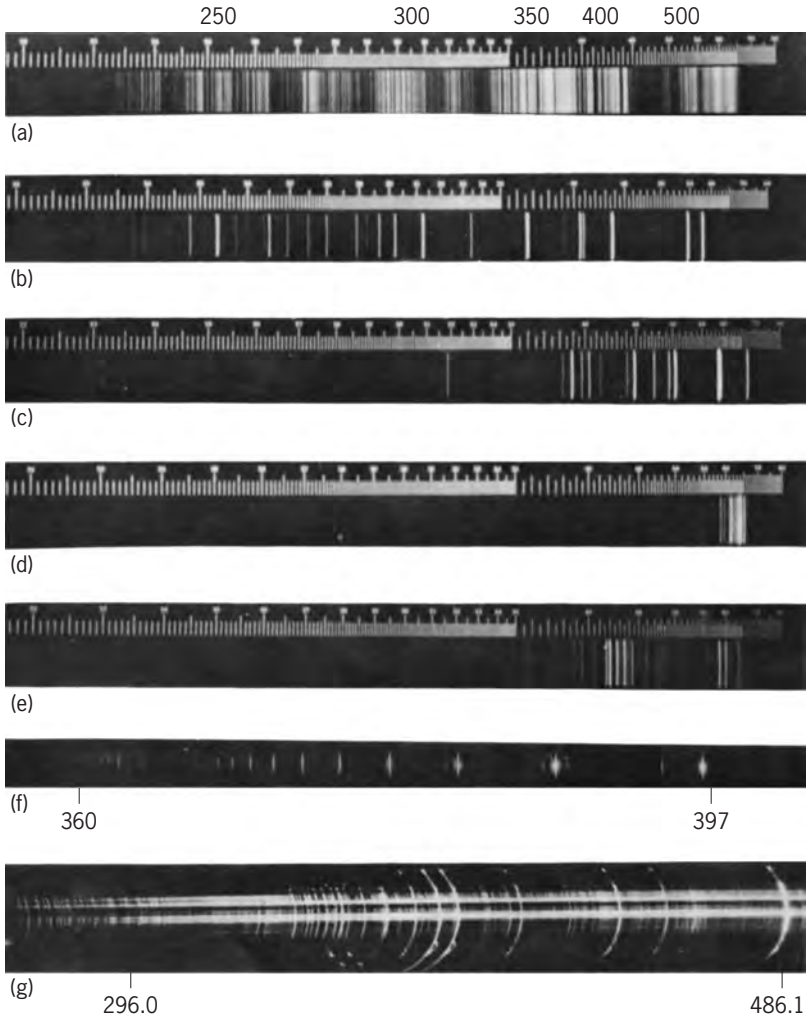

Photographs of common line spectra, wavelengths given in angstroms; emission spectra (*a–e*) all taken with the same quartz spectrograph. (*a*) Spectrum of iron arc. (*b*) Mercury spectrum from an arc enclosed in quartz. (*c*) Helium in a glass discharge tube. (*d*) Neon in a glass discharge tube. (*e*) Argon in a glass discharge tube. (*f*) Balmer series of hydrogen in the ultraviolet, photographed with a grating spectrograph. (*g*) Emission spectrum from gaseous chromosphere of the Sun, a grating spectrum taken without a slit at the instant immediately preceding a total eclipse, when the rest of the Sun is covered by the Moon's disk. Two strongest images, H and K lines of calcium, show marked prominences, or clouds, of calcium vapor. Other strong lines are caused by hydrogen and helium. 1 angstrom unit = 0.1 nanometer. (*From F. A. Jenkins and H. E. White, Fundamentals of Optics, 4th ed., McGraw-Hill, 1976*)

narrow. The spectrum then appears as an array of bright line slit images on a dark background (see **illus.**). Under certain conditions spectra show dark absorption lines against a bright background. *See* ATOMIC STRUCTURE AND SPECTRA; SPECTROSCOPY.

George R. Harrison

Linear algebra

That branch of mathematics which deals with solutions of systems of linear equations and the related geometric notions of vector spaces and linear transformations. It is fundamental in the theory of the calculus of functions of several variables and hence is of great importance in the application of mathematics to physical and biological sciences, economics, and so on.

The word linear is derived from the fact that the equation of a line in two-dimensional analytic geometry has the form shown in Eq. (1) and a system of linear equations has the corresponding form shown in Eq. (2), where $i = 1, 2, \ldots, m$. The a_{ij} and b_i are fixed

$$ax + by = c \tag{1}$$

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \tag{2}$$

quantities belonging to a specified field, for example, the field of real numbers, and solutions $(x_1, x_2, \ldots, x_n)$ are sought in the same field. If $m = n = 2$, the problem of solving the equations is equivalent to the geometric one of finding the points common to two lines (which may be coincident).

Basic concepts. Let R denote the field of real numbers and let R^2 be the set of pairs (x,y) of real numbers x, y. Such a pair can be regarded as the

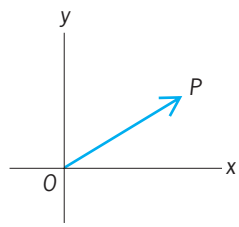

Geometric representation of a vector OP.

coordinates of a point P in a two-dimensional euclidean space. The point P determines a vector, or directed segment **OP** with initial point the origin O and terminal point P, as shown in the **illustration**.

Such vectors are basic in representing certain physical entities, for example, acceleration, velocity, force. Addition of these physical entities corresponds to the addition of vectors. If $\mathbf{v} = \mathbf{OP}$ and $\mathbf{v}' = \mathbf{OP}'$ where P' is the point (x', y'), then $\mathbf{v} + \mathbf{v}' = \mathbf{OQ}$, where $Q = (x + x', y + y')$. If t is a real number, a stretching of the vector $\mathbf{v}$ in the ratio t: 1 gives rise to the vector $t\mathbf{v} = \mathbf{OM}$, where $M = (tx, ty)$.

On the algebraic side these considerations generalize from pairs of real numbers (x,y) to n-tuples of real numbers $(x_1, x_2, \ldots, x_n)$, and one can replace the field R of real numbers by any field F, for example, the field of rational numbers.

Accordingly, let F^n denote the set of n-tuples $\mathbf{x} = (x_1, x_2, \ldots, x_n)$; where the x_i are in a given field F. Such an n-tuple is called a vector. If $\mathbf{y} = (y_1, y_2, \ldots, y_n$; then one defines the sum of x and y by Eq. (3).

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \tag{3}$$

See ABSTRACT ALGEBRA.

The elements of the field F are called scalars. If t is such an element, one defines the multiplication of the vector x by the scalar t by means of Eq. (4).

$$t\mathbf{x} = (tx_1, tx_2, \ldots, tx_n) \tag{4}$$

The basic properties of addition of vectors and multiplication of vectors by scalars are listed below.

A_1. $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
A_2. $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
A_3. There exists a vector $\mathbf{O}$ such that $\mathbf{x} + \mathbf{O} = \mathbf{x} = \mathbf{O} + \mathbf{x}$ for all $\mathbf{x}$.
A_4. For any vector $\mathbf{x}$ there exists a vector $-\mathbf{x}$ such that $\mathbf{x} + (-\mathbf{x}) = \mathbf{O} = (-\mathbf{x}) + \mathbf{x}$.
S_1. $t(\mathbf{x} + \mathbf{y}) = t\mathbf{x} + t\mathbf{y}$
S_2. $(s + t)\mathbf{x} = s\mathbf{x} + t\mathbf{x}$
S_3. $(st)\mathbf{x} = s(t\mathbf{x})$
S_4. $1\mathbf{x} = \mathbf{x}$

Here $\mathbf{x}, \mathbf{y}, \mathbf{z}$, are vectors, s, t are scalars, and 1 is the unit element of F. These properties are immediate consequences of the definitions and the properties of the field F. If one defines unit vectors $\mathbf{e}$ by means of Eq. (5), where the 1 occurs in the ith place, then

$$\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0) \tag{5}$$

a consequence of the definition of vector addition is that any vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ can be expressed as the sum shown in Eq. (6). In addition, this way of

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n \tag{6}$$

presenting the vector $\mathbf{x}$ as a linear combination of the vectors $\mathbf{e}_i$ is unique.

Axiomatic formulation. It is possible to obtain better insight and power in dealing with vectors if one axiomatizes the situation described above. Accordingly, one defines the notion of a vector space V over a field F as a set V with an addition composition $\mathbf{x} + \mathbf{y}$, and a multiplication $t\mathbf{x}$ of elements $\mathbf{x}$ of V by scalars t(elements of F). These are required to satisfy the axioms A_2 through A_4 and S_1 through S_4. It is worth noting that conditions A_1 through A_4 state that V and its addition composition constitute a commutative group. A subset B of V is called a base if every vector $\mathbf{x}$ can be written in one and only one way as a linear combination $\mathbf{x} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_m\mathbf{b}_m$, where the $\mathbf{b}_i$ are in B and the x_1 are in F. Every vector space has a base and any two bases have the same cardinal number which is called the dimension of V. The vector space F^n of n-tuples of elements of the field F has dimension n with base $B =$ base $B = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ defined above. *See* GROUP THEORY.

A subset U of V is called a subspace if the element $s\mathbf{u} + t\mathbf{v}$ is in U whenever $\mathbf{u}$ and $\mathbf{v}$ are in U and s and t are in F. The factor space V/U is the set of cosets $\mathbf{x} + \mathbf{u}$ where $\mathbf{x}$ is fixed and $\mathbf{u}$ ranges over U. Addition and scalar multiplication in V/U are defined by $(\mathbf{x} + U) + (y + U) = \mathbf{x} + \mathbf{y} + U$, $t(\mathbf{x} + U) = t\mathbf{x} + U$, respectively.

Linear transformations. If U and V are vector spaces over the same field F, then one defines a linear transformation $\mathbf{A}$ of U into V as a correspondence $x \to \mathbf{Ax}$ (also written as $\mathbf{xA}$) of U into V such that Eqs. (7)

$$\begin{aligned} \mathbf{A}(\mathbf{x} + \mathbf{y}) &= \mathbf{Ax} + \mathbf{Ay} \\ \mathbf{A}(t\mathbf{x}) &= t(\mathbf{Ax}) \end{aligned} \tag{7}$$

is satisfied for any $\mathbf{x}, \mathbf{y}$ in U and any t in F. The image $\mathbf{A}U$ of U under the linear transformation $\mathbf{A}$ is the set of all vectors $\mathbf{Ax}$ where $\mathbf{x}$ ranges over all of U. $\mathbf{A}U$ is a subspace of V. The kernel of $\mathbf{A}$ is the set of vectors $\mathbf{z}$ in U such that $\mathbf{Az} = 0$. This is also a subspace of U. If U and V are finite dimensional and b is vector in V, then the determination of the vectors x of U such that $\mathbf{Ax} = \mathbf{b}$ is equivalent to the determination of the set of solutions of a system of linear equations.

If $\mathbf{A}$ and $\mathbf{B}$ are linear transformations of a vector space U into another vector space V and if t is in F, then one defines $\mathbf{A} + \mathbf{B}$ and $t\mathbf{A}$ by $(\mathbf{A} + \mathbf{B})\mathbf{x} = \mathbf{Ax} + \mathbf{Bx}$, $(t\mathbf{A})\mathbf{x} = t(\mathbf{Ax})$, respectively. If $\mathbf{C}$ is a linear transformation of V into a third vector space W, then one defines $\mathbf{AC}$ by $(\mathbf{AC})\mathbf{x} = \mathbf{A}(\mathbf{Cx})$. The mappings $\mathbf{A} + \mathbf{B}$, $t\mathbf{A}$, $\mathbf{AC}$ are all linear transformations. One denotes the set of linear transformations of U into V by hom (U,V). This is itself a vector space relative to the addition $\mathbf{A} + \mathbf{B}$ and the multiplication $t\mathbf{A}$ defined above. In the special case in which $V = F$, the vector space $V^* =$ hom (V,F) is called the conjugate space of V or the space of the linear functions on V. If $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$

are in hom(U,U), then $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$, $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$, $(\mathbf{B} + \mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$, $t(\mathbf{AB}) = (t\mathbf{A})\mathbf{B} = \mathbf{A}(t\mathbf{B})$. These properties imply that hom (U,U) is an algebraic system of a type called an associative algebra.

If U and V are vector spaces over F with bases $\{\mathbf{u}_1,\mathbf{u}_2, \ldots, \mathbf{u}_m\}$ and $\{\mathbf{v}_1,\mathbf{v}_2, \ldots, \mathbf{v}_n\}$, respectively, and $\mathbf{A}$ is a linear transformation of U into V, then according to the definition of a base, one can write

$$\mathbf{A}\mathbf{u}_i = \sum_{j=1}^{i} a_{ji}\mathbf{v}_j$$

In this way one can associate with $\mathbf{A}$ the n by m matrix (a_{ji}). The correspondence between hom (U,V) and the set of n by m matrices with entries in F is 1-1 and preserves addition and multiplication by scalars. That is, to each transformation in hom (U,V) there corresponds one and only one matrix, and the matrix which corresponds to the sum of two transformations is the sum of the matrices corresponding to the separate transformations. Also if $V = U$ then the correspondence preserves multiplication. In this way one sees that the theory of linear transformations between finite-dimensional vector spaces is equivalent to matrix theory. *See* MATRIX THEORY.

Multilinear algebra and tensors. Linear algebra is generally regarded as including also multilinear algebra, the basic notion of which is that of a bilinear mapping of the pair of vector spaces U, V into a third vector space W (over the same field F). This is a mapping B of the set of pairs $(\mathbf{x},\mathbf{y})$ with $\mathbf{x}$ in U and $\mathbf{y}$ in V into W such that Eqs. (8) are satisfied for $\mathbf{x}$, $\mathbf{x}'$ in U;

$$\begin{aligned} B(\mathbf{x} + \mathbf{x}', \mathbf{y}) &= B(\mathbf{x}, \mathbf{y}) + B(\mathbf{x}', \mathbf{y}) \\ B(\mathbf{x}, \mathbf{y} + \mathbf{y}') &= B(\mathbf{x}, \mathbf{y}) + B(\mathbf{x}, \mathbf{y}') \\ B(t\mathbf{x}, \mathbf{y}) &= tB(\mathbf{x}, \mathbf{y}) = B(\mathbf{x}, t\,\mathbf{y}) \end{aligned} \tag{8}$$

$\mathbf{y}$, $\mathbf{y}'$ in V; and t in F. This leads to the notion of the tensor product $U \otimes V$, which is a vector space together with a bilinear mapping of U and V into this space which is maximal in a certain sense. One can iterate the construction of $U \otimes V$ to obtain tensor spaces of higher degree. Such spaces are fundamental for applications of algebra to differential geometry and relativity theory.

Quadratic forms. An important part of linear algebra is concerned with the study of a vector space relative to a given quadratic form. This is defined to be a correspondence Q: $\mathbf{x} \to Q(\mathbf{x})$ of V into F such that $Q(t\mathbf{x}) = t^2Q(\mathbf{x})$ and such that $B(\mathbf{x},\mathbf{y}) = Q(\mathbf{x} + \mathbf{y}) - Q(\mathbf{x}) - Q(\mathbf{y})$ is bilinear from the pair V, V to F. If $F = R$, the field of real numbers, and V is complete in the topology defined by Q as a norm (an absolute value), then a real Hilbert space is obtained. Replacement of Q by a hermitian form defines a (complex) Hilbert space. The theory of finite-dimensional real Hilbert spaces is equivalent to the study of n-dimensional euclidean geometry. Infinite-dimensional Hilbert spaces and their generalization called Banach spaces play an important role in modern analysis.

Nathan Jacobson

Bibliography. H. Anton, *Elementary Linear Algebra*, 8th ed., 2000; J. B. Fraleigh and R. A. Beauregard, *Linear Algebra*, 3d ed., 1995; P. R. Halmos, *Finite Dimensional Vector Spaces*, 2d ed., 1958, reprint 1993; S. A. Lang, *Linear Algebra*, 3d ed., 1994; S. J. Leon, *Linear Algebra with Applications*, 5th ed., 1997; W. K. Nicholson, *Linear Algebra with Applications*, 3d ed., 1994.

Linear energy transfer (biology)

A physical factor which is significant in biology because, in many biological actions of high-energy radiations, linear energy transfer (LET) strongly and characteristically modifies the quantitative relations between dose and effect. *See* RADIATION BIOLOGY.

When a moving charged particle (for example, an electron or proton) traverses matter, it progressively loses kinetic energy which is gained, in diverse amounts, by individual molecules along its path. The amount of energy thus transferred, per unit length of path, is the LET. It is roughly proportional, at least in gases, to ionization per unit path (specific ionization; ion density). LET increases as the velocity of the particle decreases; accordingly it varies greatly along any individual path and is maximal near the end. Also, LET increases as the square of the charge on the particle; for example, the LET of an alpha particle is four times that of an electron or proton of the same instantaneous velocity.

For experimental purposes, different values of LET can be obtained in two general ways. In the track-segment method, a beam is used made up of particles (protons) that have uniform initial velocity and produce substantially straight tracks. The experiment is so arranged that they all traverse a sample of biological objects with the same linear segment of their paths. Different LET values are obtained by using suitably different path segments. In the track-average method, either entire paths are spent within the biological object or random linear segments of different paths traverse it. Different average values of LET are obtained by using radiations that produce suitably different spectra of LET. Although the average usually used is the arithmetic mean, other averages will probably be found to be more meaningful as greater understanding of radiobiological mechanisms is achieved.

In some radiobiological actions the shape of the dose-effect curve is changed as LET is varied. However, these cases are rare in comparison with those in which only the slope is changed. In the latter case the effect of LET can be expressed by a single parameter (relative biological effectiveness, RBE) which varies as the slope (or as the reciprocal of the dose, in rads, required to produce a given degree of effect).

For some radiobiological actions, RBE decreases with increase in LET; this is exemplified by inactivation of macromolecules, viruses, and certain bacteria. For other actions, RBE increases with LET; in some of these cases, RBE has been shown to pass through a maximum; there are theoretical reasons

to believe that maxima would be observed for all of them if larger values of LET were available.

The distribution of individual energy transfers along a path corresponds to the distribution of ionized and excited molecules, production of which is believed to initiate the radiobiological mechanism. Because of diffusion, this distribution in space lasts only 1 microsecond or less. Thus the observed variation of RBE with LET demonstrates that some essential part of the radiobiological mechanism is completed in a very short time.

Among other interesting facts are the following: (1) In at least some radiobiological actions, the influence of LET is modified by other factors, such as concentration of molecular oxygen. (2) Influence of LET is also encountered in radiation chemistry. For instance, the decomposition of liquid water depends strongly on LET, and this dependence is modified by oxygen. Facts such as these make LET of prime interest in attempts to understand radiobiological mechanisms. *See* RADIATION; RADIATION INJURY (BIOLOGY).

Raymond E. Zirkle

Bibliography. T. Alper, *Cellular Radiology*, 1979; W. Harm, *Biological Effects of Ultraviolet Radiation*, 1980; D. J. Pizzarello and R. L. Witcofski, *Basic Radiation Biology*, 2d ed., 1975; C. Von Sonntag, *The Chemical Basis for Radiation Biology*, 1987.

Linear momentum

Isaac Newton's "quantity of motion," defined as the product of a particle's mass and velocity, whose time rate of change equals the applied force; in modern physics, the rate of change of energy with respect to velocity. The law of conservation of linear momentum states that the total linear momentum of an isolated system remains constant. *See* ANGULAR MOMENTUM; CONSERVATION OF MOMENTUM; ENERGY; MOMENTUM.

Role in classical mechanics. The linear momentum of a particle is the vector quantity $\mathbf{p}$ given by Eq. (1),

$$\mathbf{p} = m\mathbf{v} \tag{1}$$

where m is the particle's mass and $\mathbf{v}$ is its velocity. *See* MASS; VELOCITY.

Newton's second law of motion states that the rate of change of linear momentum is proportional to the applied force $\mathbf{F}$, Eq. (2).

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} \tag{2}$$

See FORCE.

This is actually the definition of force in terms of the fundamental quantities of physics: mass, length (here, position), and time. If more than one force acts on the particle, then $\mathbf{F}$ is the (vector) sum of the forces. Newton's first law of motion is implied because if $\mathbf{F} = \mathbf{0}$, then $\mathbf{p}$ is constant. *See* DIMENSIONS (MECHANICS); DYNAMICS; EQUILIBRIUM OF FORCES; INERTIA; RESULTANT OF FORCES.

Newton based his laws on a number of assumptions he took for granted. These are (1) total mass is constant (this is sometimes called Newton's zeroth law of motion); (2) masses (which are scalar quantities) add linearly; (3) velocities (which are vector quantities) and forces (also vectors) add linearly, implying that momenta must add linearly; (4) forces do not depend on velocity; (5) forces between particles are directed along the line that connects them; (6) all physical interactions are completely deterministic; and (7) mass, energy, and momentum are infinitely subdivisible. *See* CONSERVATION OF MASS; DETERMINISM; RELATIVE MOTION.

In addition to the above assumptions are others regarding the reference frames: the coordinate systems for specifying positions and velocities. By definition, inertial reference frames are those in which Newton's laws are true; this implies that the frames are neither rotating nor accelerating, but are only moving at a constant velocity with respect to the observer. Newton's laws also assume that all the reference frames are Newtonian, meaning that measurements of position and time are completely independent; there is one absolute universal time measurement; and influences act across all distances instantaneously (the speed of light is infinite). *See* FRAME OF REFERENCE.

The various branches of modern physics are characterized by the extent to which one or more of the above assumptions are dropped. For example, assumptions (1), (2), and (3) are not true for relativistic velocities; assumptions (4) and (5) are not true for magnetic forces between particles; (6) is not true for quantum-mechanical systems; and (7) is not true for atomic, nuclear, and subnuclear physics. General relativity describes physics in accelerating frames of reference.

The total linear momentum of a multiparticle system equals the sum of the linear momenta of its constituent particles. If the system's configuration is not important, then it can be modeled as a single particle, having the same mass as the system, located at the system's center of mass, and subjected to the sum of the external forces. If there are no external forces (or if they always sum to zero), then by Newton's first law, the center of mass will move at a constant velocity. *See* CENTER OF MASS; COLLISION (PHYSICS).

Expanding Eq. (3) by the usual rules of differentiation shows that, in general, change in linear momentum can be caused by (nonrelativistic) change in mass. An example would be in the analysis of rocket flight wherein the total mass of the rocket decreases as its propellants are ejected.

$$\frac{d\mathbf{p}}{dt} = \frac{d}{dt}(m\mathbf{v}) = m\frac{d\mathbf{v}}{dt} + \mathbf{v}\frac{dm}{dt} \tag{3}$$

When mass is constant, Eq. (3) simplifies to Eq. (4),

$$\mathbf{F} = m\frac{d\mathbf{v}}{dt} = m\mathbf{a} \tag{4}$$

where $\mathbf{a}$ is the acceleration of the particle. *See* ACCELERATION.

The units of measure for linear momentum are the Newton second (in SI units) and pound-force second (in U.S Customary units). *See* UNITS OF MEASUREMENT.

Linear momentum in continuous media. Continuous media include deformable solids, fluids (liquids and gases), plasmas, and "empty" space. For such media, it is more appropriate to use momentum density, defined as momentum per unit volume of the medium. Two examples are given here.

Fluid mechanics. The Navier-Stokes equations describe the dynamics of fluids in terms of conservation of momentum. Linear momentum appears as the density (mass per unit volume) times velocity at a given point in the fluid, that is, the momentum density at that point. *See* FLUID-FLOW PRINCIPLES; NAVIER-STOKES EQUATIONS.

Classical electromagnetic theory. Electromagnetic waves transport energy according to the Poynting vector $\mathbf{S} = \mathbf{E} \times \mathbf{H}$, where $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic field vectors, respectively. The energy propagates in the direction of $\mathbf{S}$ at the speed of light, suffusing the (lossless) medium with electromagnetic momentum $\mathbf{g} = (1/c^2)\,\mathbf{S}$, measured in N-s/m^3, that will exert radiation pressure on matter. The pressure is very small: for sunlight, it is about 0.44 kg-force/(km)^2, or 2.5 lb-force/(mi)^2, on the Earth's surface. *See* POYNTING'S VECTOR; RADIATION PRESSURE.

Role in relativistic mechanics. Relativistic effects become important when velocities approach the speed of light c. In particular, a particle moving with speed v appears more massive by the factor $\gamma(v)$, Eq. (5). However, an isolated system will retain the

$$\gamma(v) = (1 - v^2/c^2)^{-\frac{1}{2}} \tag{5}$$

same form of law of conservation of linear momentum provided we use the relativistic momentum, Eq. (6), of each constituent particle.

$$\mathbf{p} = \gamma(v) m_0 \mathbf{v} = m(v)\mathbf{v} \tag{6}$$

Relativistic momentum is sometimes called the 3-momentum because its three components plus an energy term constitute the momentum-energy 4-vector, or 4-momentum, much used in high-energy particle physics. Although different observers will obtain different measurements of a particle's energy and momentum, they will all agree on the value of its rest mass. Mathematically stated, the magnitude of the 4-momentum $c^2 m_0$ (the momentum-energy, or momenergy), given by Eq. (7), is invariant and so

$$c^2 m_0 = \sqrt{E^2 - |\mathbf{p}|^2 c^2} \tag{7}$$

relates any observer's own measurements of the particle's energy E and momentum $\mathbf{p}$. For a particle at rest in the observer's frame, Eq. (7) yields Einstein's famous equation for rest energy: $E_{\text{rest}} = m_0 c^2$. *See* RELATIVISTIC MECHANICS; RELATIVITY.

Linear momentum in quantum theory. Quantum mechanics differs radically from classical mechanics in several ways. With specific regard to linear momentum in nonrelativistic quantum theory:

1. A particle's linear momentum and position cannot both be precisely measured at the same time. Heisenberg's uncertainty principle specifies the lower limit on the product of the uncertainties (probable range of errors) of the two measurements by Eq. (8) where $\hbar = h/(2\pi)$ and h is Planck's constant. *See* PLANCK'S CONSTANT; UNCERTAINTY PRINCIPLE.

$$\Delta p_x \Delta x \geq \frac{\hbar}{2} \tag{8}$$

2. Electromagnetic radiation is quantized into photons whose linear momentum is given by Eq. (9),

$$p = \frac{h}{\lambda} \tag{9}$$

where λ is the electromagnetic wavelength. When a photon loses some of its momentum to a subatomic particle, its wavelength increases correspondingly. *See* COMPTON EFFECT; QUANTUM (PHYSICS).

3. Particles have wavelike properties and their momentum is also given by Eq. (9), where λ is now the particle's de Broglie wavelength. *See* DE BROGLIE WAVELENGTH; WAVE MECHANICS.

4. The Lagrangians and Hamiltonians of classical mechanics are invalid since they are premised on the observer being able to accurately measure momentum and position simultaneously. However, the same form of Hamiltonian is used, but with the dynamic variables replaced by their analogous Hamiltonian operators. For example, the momentum operator $\hat{p}$ in the one-dimensional form of Schrödinger's equation is given by Eq. (10), where i is the "imaginary" unit $\sqrt{-1}$. *See* HAMILTON'S EQUATIONS OF MOTION; LAGRANGE'S EQUATIONS OF MOTION.

$$\hat{p} = \frac{\hbar}{i}\frac{\partial}{\partial x} \tag{10}$$

By applying the momentum operator in the appropriate way to the wave function, one obtains the probability of obtaining any given value of measured momentum. Although repetition of the experiment will yield different results in detail, nevertheless the total linear momentum will always be found to be conserved. *See* NONRELATIVISTIIC QUANTUM THEORY; QUANTUM MECHANICS; SCHRÖDINGER'S WAVE EQUATION.

Role in relativistic quantum theory. When relativistic effects become important, the total momentum-energy of a system is still conserved. However, tools such as Schrödinger's equation become invalid because they use time independently of space and one must instead use the methods of quantum field theories. Among these methods are the diagrammatic momentum-space Feynman rules of quantum electrodynamics. *See* FEYNMAN DIAGRAM; QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; QUANTUM GRAVITATION; RELATIVISTIC QUANTUM THEORY.

Andrej Tenne-Sens

Bibliography. R. P. Feynman, R. B. Leighton, and M. L. Sands, *The Feynman Lectures on Physics: The Definitive and Extended Edition*, 3 vols., Addison-Wesley, Reading, MA, 2005; H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., Pearson Education, San Francisco, 2002; J. Griffiths, *Introduction to Quantum Mechanics*, Prentice Hall,

Englewood Cliffs, NJ, 1994; J. D. Jackson, *Classical Electrodynamics*, 3d ed., Wiley, 1999; R. Penrose, *The Road to Reality: A Complete Guide to the Laws of the Universe*, Jonathan Cape, London, 2004; M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Westview Press, 1995; A. Sommerfeld, *Lectures on Theoretical Physics*, vol. 1: *Mechanics*, Academic Press, New York, 1964; E. F. Taylor and J. A. Wheeler, *Spacetime Physics: Introduction to Special Relativity*, 2d ed., Freeman, New York, 1992.

Linear programming

An area of mathematics concerned with the minimization (or maximization) of a linear function of several variables subject to linear equations and inequalities. Linear programming developed from three main areas: transportation problems, game theory, and input-output models. Work on all these areas took place before and during World War II, with major contributions by L. V. Kantorovich, J. von Neumann, and W. Leontief. The subject in its present form was created in 1947, when G. B. Dantzig defined the general model and proposed the first, and still the most widely used, method for its solution: the simplex method.

Although the linearity assumptions are restrictive, many algorithms for extensions of linear programming, such as problems with nonlinear or integer restrictions, involve successively solving linear programming problems. With a result in 1979 giving a polynomially bounded ellipsoid method, an alternative to the simplex method, linear programming became the focus of work by computer scientists, and nonlinear methods have been refocused on solving the linear programming problem. Work by N. K. Karmarkar announced in 1984 attracted much attention because of claims of vastly improved performance of a new interior method. The relative merits of Karmarkar's method and the simplex method remain to be determined, but there seems to be a place for both methods. Karmarkar's work stimulated considerable activity in linear programming methodology. *See* NONLINEAR PROGRAMMING; OPTIMIZATION.

General problem. The linear programming problem is to minimize linear objective function (1) subject to restrictions (2). The variables $x_1 \ldots, x_n$ are required to take on real values, and the coefficients a_{ij}, c_j, and b_i are real constants. The objective could be to maximize rather than minimize, and among constraints (2) the equations could be replaced by inequalities of the form less-than-or-equal-to or greater-than-or-equal-to. The set of x_j's satisfying constraints (2) form a convex polyhedron, and the optimum value of the objective function will always be assumed at a vertex of the polyhedron unless the objective function is unbounded. The simplex method works by moving from vertex to vertex until the vertex yielding the optimum value of the objective function is reached, while interior methods stay inside the polyhedron. *See* LINEAR SYSTEMS OF EQUATIONS.

$$c_1x_1 + \cdots + c_nx_n \tag{1}$$
$$x_1 \geq 0, \ldots, x_n \geq 0$$

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ \vdots \qquad\qquad\qquad &\quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{2}$$

Transportation problem. The transportation problem is an example of a linear programming model. In this model, there are supply amounts $S_1 \ldots, S_K$ of a given product at supply points $1, \ldots, K$ and demand amounts $D_1, \ldots, D_L$ at demand points $1, \ldots, L$. The variables x_{ij} are associated with a supply point i and demand point j and represent the amount to be shipped from i to j. The constraints are $x_{ij} \geq 0$ and inequalities (3), and the objective is to minimize

$$\begin{array}{llll} x_{11} + \cdots + x_{1L} & & & \leq S_1 \\ \cdots & & & \\ & x_{K1} + \cdots + x_{KL} & & \leq S_K \\ x_{11} + \cdots & + x_{K1} & & \geq D_1 \\ \cdots & & & \\ x_{1L} & + \cdots & + x_{KL} & \geq D_L \end{array} \tag{3}$$

$\Sigma c_{ij}x_{ij}$, the total cost of shipping, where c_{ij} is the per unit cost of shipping from supply point i to demand point j. Variants include bounds and costs on storage at the supply points and penalties for shortages at the demand points.

Dual problem. For every linear programming problem with objective function (1) and constraints (2), there is a dual linear program in variables $y_1, \ldots, y_m$, one variable for each equation in (2). The objective is to maximize function (4), subject to linear inequalities (5). In this form, the y_i's are not constrained to

$$y_1b_1 + \cdots + y_mb_m \tag{4}$$

$$\begin{aligned} y_1a_{11} + \cdots + y_ma_{m1} &\leq c_1 \\ y_1a_{1n} + \cdots + y_ma_{mn} &\leq c_n \end{aligned} \tag{5}$$

be nonnegative, although the problem can be stated so that the dual is more symmetric by changing the equations in (2) to inequalities. The duality theorem states that the maximum value of function (3) is equal to the minimum value of function (1), provided both constraints (2) and (5) have solutions. The dual variables are related to Lagrange multipliers, and can be interpreted as prices on resources constrained by restrictions (2). The duality theorem is mathematically equivalent to the result that any point not in the convex polyhedron defined by constraints (2) can be separated from it by a hyperplane. The original problem with constraints (2) is referred to as the primal problem, and the x_j's are called primal variables. *See* CALCULUS OF VARIATIONS.

Simplex method. The simplex method in its usual form is a primal feasible method; that is, it maintains constraints (2) while working toward optimality. The method must find a solution satisfying

constraints (2), and one way to do this is to initially solve the problem of finding a solution to these constraints while ignoring the original objective function. Such a problem is itself a linear programming problem and is called the phase-one problem. Then, the minimization of the original objective function becomes phase two. The algorithm in phase two maintains primal feasibility, that is, constraints (2), while working toward satisfying inequalities (5) in such a way that once these inequalities are satisfied the solutions x and y to the primal and dual problems will be optimal with regard to these respective problems.

Interior method. Interior methods were proposed at the same time as the simplex method, but were not pursued because they did not prove to be faster than the simplex method. The first promise of success for nonlinear methods was the proof of a polynomial upper bound for the ellipsoid method of L. G. Khatchian in 1979; that is, this method was guaranteed to find a solution in a time bounded by a polynomial function of the data. Despite the theoretical interest in Khatchian's breakthrough, the method proved to be much slower, on the average, than the simplex method, at least for problems of the size encountered in practice. Karmarkar's method in 1984 proved to be competitive with the simplex method and perhaps better on large problems and on certain classes of problems. Karmarkar also gave a smaller polynomial bound. The interior methods led to renewed research. In particular, bounds have been improved, variants developed, prior work (in particular, by Russian mathematicians) rediscovered, and computational experimentation carried out. The three basic interior methods are Karmarkar's method, affine scaling, and barrier methods.

Integer programming. An important extension in practice is to require some of the x_j's to take on integral values. The most common case in practice is where the integer x_j must be 0 or 1, representing decision choices such as to whether to switch from production of one product to another or whether to expand a warehouse to allow for larger throughput. Whereas linear programming solution times tend to be less than an hour, adding the constraint that some or all of the x_j's must be integral may cause the running time to be very long. Work in the 1950s established the usefulness of linear programming in solving the so-called traveling salesman problem. Around the same time, cutting-plane methods were shown to be a convergent process for solving general integer programs, in which all of the variables x_j are required to take on integer values. Cutting-plane methods attempt to adjoin new inequality constraints to constraints (2) so that all integral solutions remain feasible and the optimum solution is integral. Some linear programs have the property that the optimum vertex is already an integral solution, but in practice the main class of such problems is the transportation problem mentioned above. *See* DECISION THEORY.

Although cutting-plane methods offered the hope of converting integer programs to larger linear programs by adjoining more constraints, the simpler branch-and-bound method became more common and was incorporated into commercial codes, perhaps largely due to the fact that the important mixed 0–1 case, where some of the variables must be 0 or 1, is not yet adequately treated by cutting-plane methods. The branch-and-bound method is simple in its original form. The linear programming relaxation (in which the constraints are relaxed by not requiring the x_j's to be integers) is solved, and a branching procedure is carried out with regard to some 0–1 variable that is at a fractional value. Two problems are created: In one the variable is forced to be 0 and in the other it is forced to be 1. The difficulty with the method is its explosive growth in the number of problems to be solved. The "bound" part of branch-and-bound comes from using the objective value of the linear programming relaxation to terminate further exploration of a problem when its objective value is worse than the objective value of some integer solution previously found or given. In its more refined variations, branch-and-bound may employ many heuristics to guide the search and elaborate improvements to the linear programming relaxation in order to improve the bound given by the value of the linear programming objective function.

Computation. Early work on computer programs was done in the 1950s. Commercial computer codes implementing the simplex method have been used in industry since the mid-1960s. Efficient methods for handling the structures encountered have been developed. In particular, the matrices tend to be very sparse, that is, most (usually over 99%) of the a_{ij}'s are zeroes. Efficient methods for basis inversion and update have been the main numerical focus of code developers. Up until the early 1980s, most problems solved by commercial codes were in the range of several hundred to several thousand equations and variables. A typical time on a mainframe computer might have been several seconds to an hour. In the 1980s, intense development in software was begun because of changed hardware and new algorithmic developments. The hardware changes included both larger supercomputers and more powerful workstations. The algorithmic developments were mainly the interior methods as well as improvements in the simplex method. In integer programming, the main algorithmic focus has been on improving the linear programming relaxation to give better bounds in branch-and-bound. *See* MATRIX THEORY; SUPERCOMPUTER.

Applications. Following the early work on codes, the petroleum industry quickly became the major user of linear programming, and still is an important user, especially for blending models in petroleum refining. Commercial codes are used in industry and government for a variety of applications involving planning, scheduling, distribution, manufacturing, and so forth. In universities, linear programming is taught in most business schools, industrial engineering departments, and operations research departments, as well as some mathematics departments. The model is general enough to be useful in the physical and social sciences. The improved

computational efficiency achieved in the 1980s has gone hand-in-hand with expanded applications, particularly in manufacturing, transportation, and finance. *See* OPERATIONS RESEARCH. Ellis L. Johnson

Bibliography. V. Chvatal, *Linear Programming*, 1983; G. B. Dantzig and M. N. Thapa, *Linear Programming*, 2 vols., 1997, 2003; J. P. Ignizio and T. M. Cavalier, *Introduction to Linear Programming*, 1994; G. L. Nemhauser and L. Wolsey, *Integer and Combinatorial Programming*, 1988; A. Sultan, *Linear Programming: An Introduction with Applications*, 1993; R. J. Vanderbei, *Linear Programming: Foundations and Extensions*, 2d ed., 2001; L. N. Vaserstein, *Introduction to Linear Programming*, 2002.

Linear system analysis

The study of properties and behavior of a system using a body of mathematical techniques based on linear system theory. A system can be defined as a set or arrangement of things related in such a way as to form a whole. Linear system analysis is concerned with the study of equilibrium and change in dynamical systems, that is, in systems that contain variables that may change with time. These variables include system inputs (external causes of change or excitation), outputs (measurable results or effects of the behavior, response, or dynamics of the system), as well as variables describing internal states of the system. To perform the analysis, relationships between these variables are described by a set of equations known as the model. In order for linear system analysis to be applicable, the model must possess the linearity property: it must be a linear model. Linearity simplifies the analysis of systems significantly, and hence there is a large body of mathematical techniques and results, referred to as linear system theory, that can be used to study linear systems. *See* LINEARITY; SYSTEMS ENGINEERING.

Examples of systems studied include both natural (biological or environmental) systems and artificial or engineered systems, such as spacecraft, aircraft, and electronic circuits. Models describing such systems can be derived from the laws of physics or fitted empirically by using input and output data. A simple mathematical model of a linear dynamical system can be obtained by applying Newton's second law to the motion, in a single dimension, of a system (**Fig. 1**), which consists of a mass M that may represent a car or a train, as given in Eq. (1). Here $u(t)$ is

$$u(t) = M\frac{d^2y}{dt^2} \tag{1}$$

the net force applied in the direction of movement, and $y(t)$ is the displacement at time t. If the only forces considered are that of the engine $f(t)$ and a viscous friction opposed to the motion whose magnitude is proportional to the velocity of the mass, then Eq. (2) holds, where K is a constant. The model

$$u(t) = f(t) - K\frac{dy}{dt} \tag{2}$$

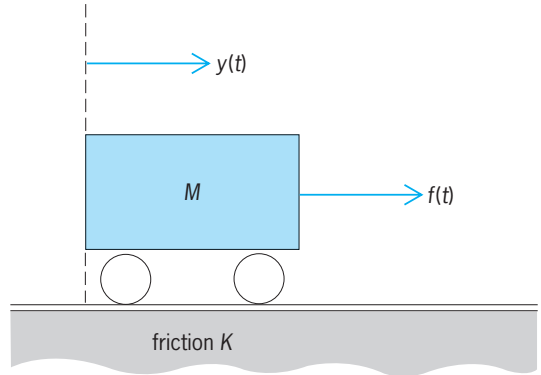

Fig. 1. A simple mass–friction dynamical system.

is then described by the second-order ordinary differential equation (3). In spite of the obvious simplicity

$$M\frac{d^2y}{dt^2} + K\frac{dy}{dt} = f(t) \tag{3}$$

of this model, it is useful in studying many practical situations.

The linearity property implies that the response of a system to the sum of several inputs is equal to the sum of its responses due to each of the individual inputs, and that the response of a system to an input multiplied by a scale factor is equal to the same scale factor times the response of the system to that input. In mathematical terms, linearity can be expressed through the use of Eq. (4), where the

$$y = F(u) \tag{4}$$

operator $F(.)$ describes the relationship between the system input and output. Here, y is the system output, u the system input, and F a mapping from input to output. In general, y and u are functions of time. The operator F is said to be linear if Eq. (5) is valid,

$$F(a_1u_1 + a_2u_2) = a_1F(u_1) + a_2F(u_2) \tag{5}$$

where u_1 and u_2 are two arbitrary inputs, and a_1 and a_2 are arbitrary real or complex numbers.

Although most systems in reality are nonlinear, it is possible to study their behavior in a limited range of values of inputs and outputs around an operating point with the help of linear approximations. By using Taylor's expansion, a nonlinear system $y = F(u)$ can be expressed in the neighborhood of an operating point u_0 as in Eq. (6), or as in Eq. (7), where F' is

$$y = F(u_0) + F'(u_0)(u - u_0) + R \tag{6}$$

$$\delta y = A\delta u + R \tag{7}$$

the first derivative of F with respect to u, $\delta y = y - F(u_0)$, $\delta u = u - u_0$, $A = F'(u_0)$ is a constant, and R is a residue which contains higher-order terms in δu that can be ignored if δu is small to achieve a linear approximation. An electronic amplifier, for example, can be approximated by a linear model for analysis of response to small signals around an operating point defined by dc bias voltages. *See* AMPLIFIER; DIFFERENTIATION.

Mathematical models. Mathematical models of dynamical systems can be divided into linear and nonlinear, time-invariant and time-varying, lumped-parameter and distributed-parameter, continuous-time and discrete-time, and so forth. The behavior of linear, continuous-time systems can be described by a set of differential equations. If the system parameters are time-invariant (that is, do not change value with time) and are lumped, then the model consists of a set or ordinary linear differential equations for continuous-time systems and difference equations for discrete-time systems. These equations can be solved, and the values of all variables at any moment in time t starting from any given initial time t_0 can be found, if all input functions are known from t_0 to t and if the initial conditions (the values at time t_0) of a minimal set of variables, referred to as the state variables of the system, are known. The state variables contain all the information about the history of the system needed to solve its equations. No knowledge of input or state values at time prior to t_0 is needed. Hence the state variables completely characterize the state of the system at any time t. *See* DIFFERENCE EQUATION; DIFFERENTIAL EQUATION.

If the state variable vector is a column vector comprising all the state variables of the system, the state vector for a second-order system is expressed by Eq. (8).

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{8}$$

The state vector is a point in a linear algebraic space referred to as the state space. It is possible to describe the behavior of any linear system by using a specific form of mathematical models that expresses relationships between system inputs, state variables, and system outputs. This form is called the state-variable model or the state-space description. For continuous time it can be expressed as in Eqs. (9) and (10), where $x(t)$ is an n-component state vec-

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{9}$$

$$y(t) = C(t)x(t) + D(t)u(t) \tag{10}$$

tor, $\dot{x}(t)$ the derivative of this vector with respect to t, $u(t)$ a p-component input vector, $y(t)$ an m-component output vector, and $A(t)$, $B(t)$, $C(t)$, and $D(t)$ matrices of appropriate dimensions. The corresponding model for discrete-time systems is given by Eqs. (11) and (12), where $t = kT$, $k = 0, 1, 2, \ldots$,

$$x(k+1) = A(k)x(k) + B(k)u(k) \tag{11}$$

$$y(k) = C(k)x(k) + D(k)u(k) \tag{12}$$

and T is the sampling period. *See* LINEAR ALGEBRA; LINEAR SYSTEMS OF EQUATIONS; MATRIX THEORY.

If the system is linear, the coefficients of these matrices do not depend on the values of the states. If the system is time-invariant, elements of the matrices are constants and the matrices are expressed simply as A, B, C, and D. Only continuous-time systems will be treated in this section; results for discrete-time systems are discussed below.

A set of state variables is defined in Eqs. (13) and (14) for the mass motion example given above. A

$$x_1(t) = y(t) \tag{13}$$

$$x_2(t) = \frac{dy}{dt} = \dot{x}_1(t) \tag{14}$$

state-space model can be formulated using these definitions as in Eqs. (15) and (16).

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & \dfrac{-K}{M} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \dfrac{1}{M} \end{bmatrix} f(t) \tag{15}$$

$$y(t) = [\,1 \;\; 0\,] \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \tag{16}$$

It is possible to describe relationships between inputs and outputs of a linear system by using another type of mathematical model known as a transfer-function model. For systems with single input and single output (SISO systems), the transfer function is the ratio between the Laplace transform of the output and the input. For multiinput–multioutput systems (multivariable or MIMO), the transfer function may be put in matrix form. The element in the ith row and jth column is a scalar transfer function describing how the jth input influences the ith output.

The system transfer function can be obtained from the state-space model by taking the Laplace transform, setting the initial value of $x(t)$ to zero, and solving for $Y(s)$ as in Eq. (17), where $Y(s)$ and $U(s)$

$$Y(s) = [C(sI - A)^{-1}B + D]U(s) \tag{17}$$

are the Laplace transforms of $y(t)$ and $u(t)$, and I is an $n \times n$ identity matrix. *See* LAPLACE TRANSFORM; MODEL THEORY.

Behavior of a linear system. In order to understand the behavior of a linear system, its response to a specified input is often calculated. The output can be calculated by solving the differential equations describing the model. Solution of the state equations provides a systematic way of solving the model equations to calculate the output as a function of time. The state equations can be solved analytically or by computer simulation techniques involving numerical integration. The closed-form solution of the state equations is given by Eq. (18), and the output by Eq. (19), where the $n \times n$ state-transition matrix $\Phi(t,\tau)$ satisfies differential equation (20).

$$x(t) = \Phi(t, t_0)x(t_0) + \int_{t_0}^{t} \Phi(t, \tau)B(\tau)u(\tau)d\tau \tag{18}$$

$$y(t) = C(t)x(t) + D(t)u(t) \tag{19}$$

$$\frac{d\Phi(t, t_0)}{dt} = A(t)\Phi(t, t_0) \qquad \Phi(t_0, t_0) = I \tag{20}$$

See INTEGRAL EQUATION.

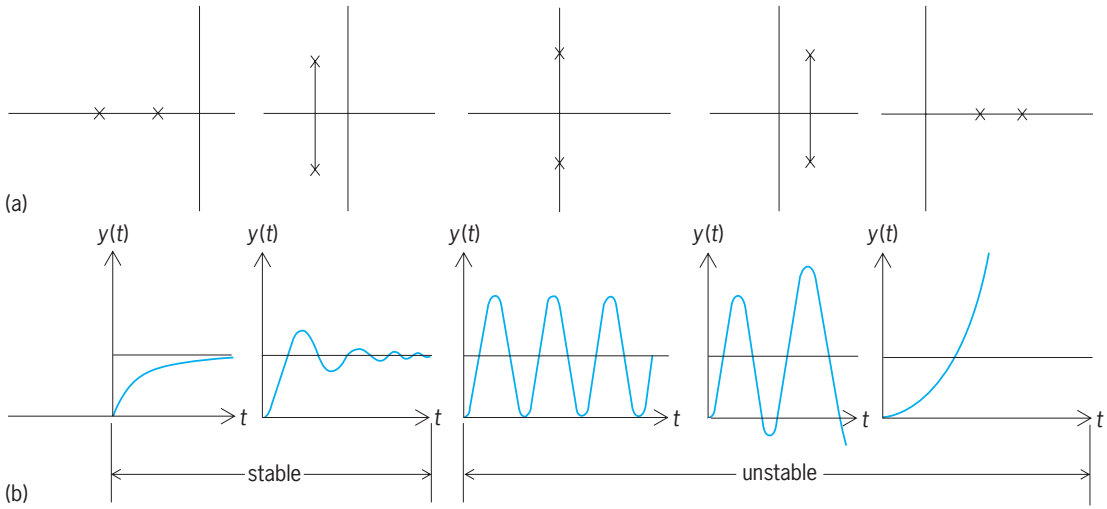

Fig. 2. Dependence of response of linear, second-order, single-input single-ouput system on location of eigenvalues of *A* matrix. (*a*) Locations of eigenvalues in complex plane. (*b*) Corresponding system response to a unit step input (a sudden change from zero to one at $t = 0$).

For time-invariant systems the state-transition matrix is found in closed form as in Eq. (21). The matrix

$$\Phi(t, t_0) = e^{A(t-t_0)} \tag{21}$$

e^{At} can be calculated by one of many possible methods, as shown in Eqs. (22) and (23), where $\mathscr{L}^{-1}$ is the inverse Laplace transform.

$$e^{At} = I + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + \cdots \tag{22}$$

$$e^{At} = \mathscr{L}^{-1}[sI - A]^{-1} \tag{23}$$

If the eigenvalues of A, λ_i ($i = 1, \ldots, n$), are distinct and $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ there is a similarity transformation Q that relates A to Λ by Eq. (24). In this case Eq. (25) holds.

$$A = Q\Lambda Q^{-1} \tag{24}$$

$$e^{At} = Qe^{\Lambda t}Q^{-1} \tag{25}$$

See EIGENFUNCTION.

Several commercial software packages and mathematical software libraries provide efficient numerical algorithms for calculation of e^{At} and for solution of the model using simulation involving numerical integration. *See* NUMERICAL ANALYSIS; SIMULATION.

Characteristics. By using the mathematical model of a linear system, a great deal of information about its behavior can be obtained. Stability, controllability, and observability can be deduced from a system model. There are many mathematical definitions and tests for stability. The bounded-input bounded-output definition is commonly used, which states simply that a system is stable if its output remains bounded for any bounded input applied to it. A continuous-time system is stable if all eigenvalues of the A matrix of its state-space model are in the left side of the complex plane, in other words if all eigenvalues have negative real parts. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

The shape of the response of a linear single-input single-output system as a function of time is determined by the location of the eigenvalues of the A matrix in the complex plane (**Fig. 2**). This is due to the fact that all elements of the matrix e^{At} consist of linear combinations of terms of the form $t^k e^{\lambda_i t}$ called the modes of the system, where the λ_i ($i = 1, \ldots, n$) are eigenvalues of A. For a multivariable system the eigenvalues and eigenvectors of A determine the shape of the response.

A system is controllable if it is possible to cause its state to move from an initial value $\boldsymbol{x}(t_0)$ at time t_0 to a specified final value $x(t_1)$ at a finite time t_1 by using an appropriate control input. A linear, time-invariant system is controllable if the $n \times pn$ controllability matrix of Eq. (26) has rank n.

$$U = [B|AB|A^2B|\cdots|A^{n-1}B] \tag{26}$$

Only the input and output of a system are measurable. The states are, in general, not accessible unless some of them are also outputs. It is sometimes necessary to generate an estimate of the state vector by using the input and output measurement values. This is possible if the system is observable. A linear, time-invariant system is observable if the $n \times mn$ observability matrix of Eq. (27) has rank n, where the

$$V = \left[C^T|A^TC^T|(A^2)^TC^T|\cdots|(A^{n-1})^TC^T\right] \tag{27}$$

superscript T stands for transpose of the matrix.

State feedback control. It is often desired to cause the system response to meet given specifications. This can be achieved by using state feedback of the form of Eq. (28) to place, or assign, the poles of the

$$u(t) = r(t) - Kx(t) \tag{28}$$

resulting (closed-loop) system (which are the same as the eigenvalues of the closed-loop A matrix) to desired locations in the complex plane, where $r(t)$ is a reference input and K a matrix of feedback gains.

It is possible to assign the poles of the closed-loop system (eigenvalues of the matrix A–BK) arbitrarily in the complex plane if the system is controllable. Furthermore, it is possible to design a state estimator of the form of Eq. (29), with arbitrarily chosen poles

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + L[y(t) - C\hat{x}(t)] \qquad (29)$$

(eigenvalues of A–LC), if the system is observable. It can be shown that, under appropriate conditions, the design of a system combining state-feedback control and state estimation (**Fig. 3**) can be separated into two problems, one for control design and one for estimator design. This is known as the principle of separation between estimation and control. *See* ESTIMATION THEORY.

Computer assistance. There are many public-domain and commercial software packages and subroutine (or procedure) libraries that assist in implementing linear system analysis on digital computers. These typically include many collections of matrix and polynomial manipulation functions, such as finding matrix eigenvalues or exponential and polynomial roots. They usually also contain simulation routines based on numerical integration and graphics routines for plotting the outcome of the simulation. Numerical robustness and stability of the algorithms used become very important in computer-assisted analysis. The effect of round-off error resulting from the finite number of bits used to represent numbers in a computer can be significant. How round-off error, stability, and robustness are dealt with determines the quality of the software used. *See* ALGORITHM; CONTROL SYSTEMS.

Hany K. Eldeib

Discrete-time linear systems. A discrete-time system is one whose input and output responses evolve at discrete time instants rather than in continuous time. A discrete-time linear system is a discrete-time system that obeys the superposition property; that is, its response to a linear combination of two inputs is the same linear combination of the responses to the inputs applied separately. It is usually assumed that a discrete-time system is also a quantized system (not explicitly included in the definition), which means that the quantities, or signals, processed can take on only values from a finite set. In that case, a discrete-time system can be implemented by means of either finite-state processing components or a general-purpose or specialized signal-processing computer. One reason for the importance and increasing use of discrete-time (quantized) systems is the ease with which they can be realized and modified simply by programming a signal-processing computer. Other reasons are the great generality of functions that can be implemented, and the lower cost relative to implementation of continuous-time (or analog) processors for comparable signal-processing functions. Applications of discrete-time linear systems include control of a process, data manipulation, and signal processing such as filtering.

Numerical solutions. A discrete-time algorithm can also be used for the numerical solution of a continuous-time model for a system. An example is the simple

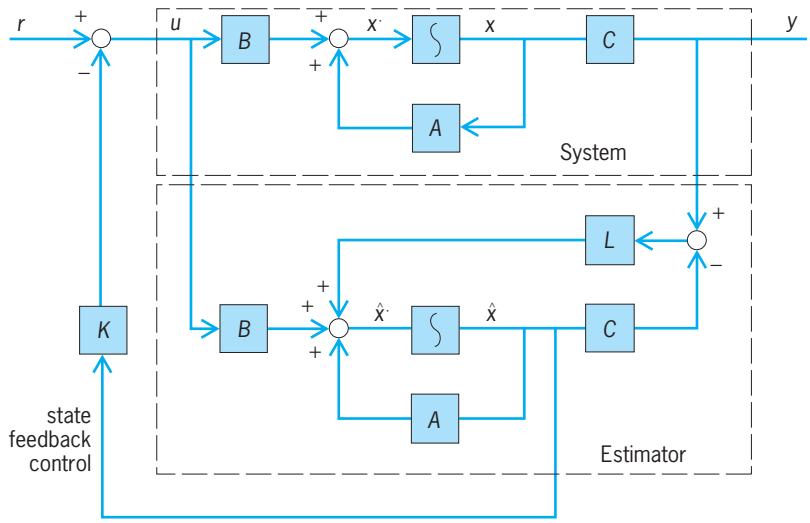

Fig. 3. State-space model of a system that combines state feedback control and state estimation.

first-order linear differential equation (30) with con-

$$\frac{dy(t)}{dt} + ay(t) = bu(t) \qquad (30)$$

stant coefficients a and b, which might relate the response of a system, $y(t)$, to a forcing function or input, $u(t)$. This equation can be put into standard state-variable form (although that would be unnecessary in this simple case) as Eq. (31), where $x(t)$ is

$$\frac{dx(t)}{dt} = -ax(t) + bu(t) \qquad y(t) = x(t) \qquad (31)$$

a state variable. The state equation [the first equation of Eq. (31)] can be approximated by a discrete-time equivalent equation by using a rectangular approximation for integration at the nth step given by Eq. (32). In this equation, the overdot denotes differ-

$$x_{n+1} = x_n + T\dot{x}_n \qquad (32)$$

entiation, T is the step size, and the subscript n denotes evaluation at the nth time instant, nT. Equation (32) is known as the Euler integration formula. The state equation (31) provides a means for eliminating $\dot{x}_n$ in Eq. (32). Rewriting it for the nth time instant gives Eq. (33). This equation can be substituted into Eq. (32) to yield Eq. (34). This first-order difference

$$\dot{x}_n = -ax_n + bu_n \qquad (33)$$

$$\left.\begin{aligned} x_{n+1} &= (1 - aT)x_n + bu_n \\ y_n &= x_n \end{aligned}\right\}, \quad n = 0, 1, 2, \ldots \qquad (34)$$

equation provides an algorithm for a numerical approximation to the solution of Eq. (30). Under the restriction that $aT \ll 1$, the sequence of numbers computed from Eq. (34) for a given input sequence $\{u_n\}$ provides a close approximation to the values of $y(t)$ at the sequence of time instants $\{nT\}$.

The advantage of recasting the original differential equation (30) in state-variable form is that the same form is obtained no matter what the order of the

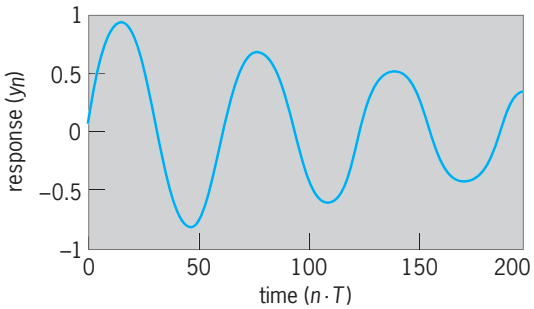

Fig. 4. Response to a unit step of a discrete-time approximation to a second-order system given by Eq. (37) with $T = 1$, $a = 0.02$, $b = 0.01$, $c = 0.1$.

system (or describing differential equation). For example, a continuous-time second-order linear system represented by the differential equation (35), where

$$\frac{d^2y(t)}{dt^2} + a\frac{dy(t)}{dt} + by(t) = c\frac{du(t)}{dt} \tag{35}$$

a, b, and c are constants, can be represented as the state-variable equations (36). The same approach to

$$\begin{aligned} \frac{dx_1(t)}{dt} &= -ax_1(t) - x_2(t) + cu(t) \\ \frac{dx_2(t)}{dt} &= bx_1(t) \\ y(t) &= x_1(t) \end{aligned} \tag{36}$$

numerical integration of these equations [or, equivalently, the second-order differential equation (35)] can be used as for the first-order example of Eq. (30), except that the numerical integration formulas are now matrix equations. The discrete-time equivalents to Eqs. (36) are given by Eqs. (37), where

$$\begin{aligned} x_{1,n+1} &= (1 - aT)x_{1,n} - Tx_{2,n} + cTu_n \\ x_{2,n+1} &= bTx_{1,n} + x_{2,n} \\ y_n &= x_{1,n} \end{aligned} \tag{37}$$

$x_{1,n} = x_1(nT)$, and so forth (**Fig. 4**). As these examples suggest, state-space methods for the description of nth-order systems of difference equations are closely related to those used for similar systems of linear differential equations.

Digital filters. Discrete-time linear systems find application in filtering, that is, the passing of certain frequency components of a signal and the attenuation or blockage of others. In general, a digital filter can be represented by the difference equation (38),

$$y(nT) = \sum_{i=0}^{r} L_i u[(n-i)T] - \sum_{j=1}^{m} K_j y[(n-j)T] \tag{38}$$

where the independent variable is now shown explicitly as integer multiples of the sampling interval to emphasize the fact that this application often concerns sampled continuous-time signals. The second term on the right-hand side of Eq. (38) involves past samples of the output, y, which will generate an infinitely long response for a finite-length input. Thus, with this term nonzero, Eq. (38) represents an infinite impulse response (IIR) filter. If this term is zero, the filter is known as a finite impulse response (FIR) filter, for the response will die out at a time rT after the input goes away.

Taking the z transform of Eq. (38) and solving for the ratio of $Y(z)$ to $X(z)$ gives the transfer function in Eq. (39). The substitution $z = e^{j\omega T}$ gives the frequency response. The magnitude of the frequency response function shows how the frequency components of the input signal are attenuated or enhanced, and it is this function that is often used for design purposes.

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{i=0}^{r} L_i z^{-i}}{1 + \sum_{j=1}^{m} K_j z^{-j}} \tag{39}$$

An often used approach for designing IIR filters is to choose an analog (linear, continuous-time) prototype filter that has the frequency response properties desired for the application at hand, and transform the transfer function of the prototype filter to the discrete-time domain. One such technique, called invariance design, holds a time-domain response function the same between the analog and digital domains. For example, with an impulse-invariant design, the impulse responses of analog and

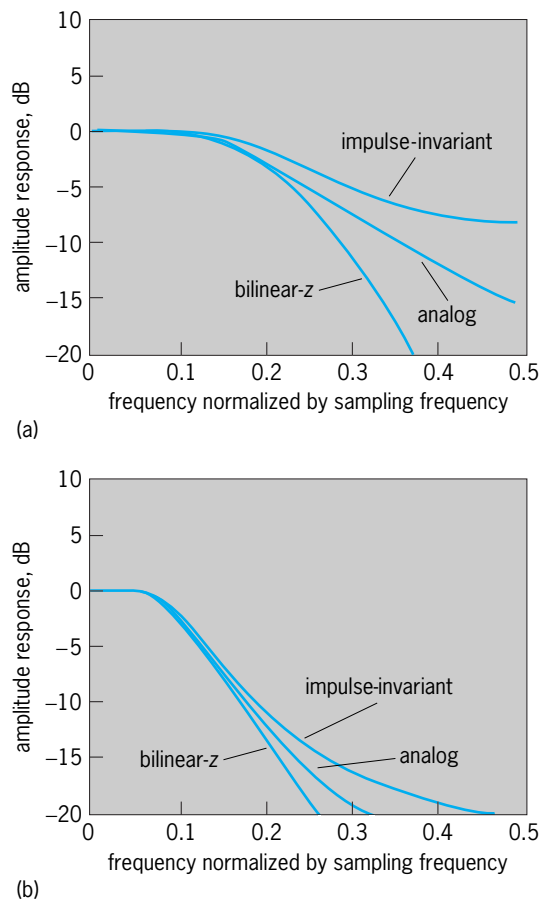

Fig. 5. Comparison of the amplitude response (that is, magnitude of frequency response) of a second-order Butterworth low-pass analog filter with responses of impulse-invariant and bilinear-z discrete-time filter approximations. (*a*) Cutoff frequency/sampling frequency = 0.2. (*b*) Cutoff frequency/sampling frequency = 0.1.

digital filters are proportional. Because invariant filter designs are prone to aliasing distortion, a bilinear-z designed filter is often favored. This filter has a zero in its frequency response at half the sampling frequency, which guarantees no aliasing, at the expense of distortion of the frequency-response function. However, for a sufficiently large ratio of sampling frequency to cutoff frequency of the filter, this distortion is small. (**Fig. 5**). *See* DIGITAL FILTER; DISTORTION (ELECTRONIC CIRCUITS); ELECTRIC FILTER.

Rodger E. Ziemer

Bibliography. C. T. Chen, *Linear System Theory and Design*, 3d ed., 1998; E. P. Cunningham, *Digital Filtering: An Introduction*, 1994; J. J. D'Azzo and C. H. Houpis, *Linear Control System Analysis and Design*, 4th ed., 1995; T. Kailath, *Linear Systems*, 1980; J. G. Proakis and D. G. Manolakkis, *Introduction to Digital Signal Processing*, 1992; R. E. Ziemer, W. H. Tranter, and D. R. Fannin, *Signals and Systems: Continuous and Discrete*, 4th ed., 1998.

Linear systems of equations

Systems of mathematical equations of the form of system (1), where the a_{ij}, $i = 1, 2, \ldots, m$, $j =$

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \tag{1}$$

$1, 2, \ldots, n$, and the b_1, $i = 1, 2, \ldots, m$, are constants, or fixed numbers, and the x_i, $i = 1, 2, \ldots, n$ are called unknowns. System (1) is referred to as m linear equations in n unknowns. A solution of system (1) is a set of numbers $c_1, c_2, \ldots, c_n$ such that when x_1 is replaced by c_1, x_2 by $c_2, \ldots, x_n$ by c_n every equation of system (1) becomes a true equality. The problem posed by such a system of equations is to find criteria for the existence of a solution and, when solutions exist, to obtain systematic methods for finding the solutions.

An example of such a linear system of equations is system (2). System (2) has infinitely many solu-

$$\begin{array}{l} 2x_1 - x_2 \qquad\quad - x_4 + 2x_5 = 2 \\ x_1 + 3x_2 + 4x_3 - 3x_4 - x_5 = 4 \\ 3x_1 + 2x_2 + 4x_3 + 2x_4 + 3x_5 = 2 \end{array} \tag{2}$$

tions which can be given in the form $x_1 = -c_5 - {}^4/_7c_3 - {}^6/_7$, $x_2 = {}^1/_3c_5 - {}^8/_7c_3 + {}^8/_{21}$, $x_3 = c_3$, $x_4 = -{}^1/_3c_5 - {}^2/_3$, $x_5 = c_5$, where c_3 and c_5 are arbitrarily chosen numbers.

Two linear systems of equations are equivalent if every solution of one of the systems is a solution of the other and vice versa. System (1) is replaced by an equivalent system, for which the solutions, if they exist, are easily found by performing certain elementary operations on the equations. These operations are interchanging two equations, adding to one equation another equation multiplied by a constant, or multiplying an equation by a nonzero constant. By a sequence of such operations, system (1) can be replaced by equivalent system (3). If $e_{r+1} =$

$$\begin{array}{r} x_{n_1} + d_{1,n_1+1}x_{n_1+1} + d_{1,n_1+2}x_{n_1+2} + \cdots \\ + d_{1,n}x_n = e_1 \\ x_{n_2} + d_{2,n_2+1}x_{n2_1+1} + \cdots + d_{2,n}x_n = e_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_{n_r} + d_{r,n_r+1}x_{n1_r+1} + \cdots\cdots + d_{r,n}x_n = e_r \\ 0 = e_{r+1} \\ \cdots\cdots\cdots \\ 0 = e_m \end{array} \tag{3}$$

$e_{r+2} = \cdots = e_m = 0$, then the solutions of the system can be obtained by choosing arbitrary values for the unknowns other than

$$x_{n_1}, x_{n_2}, \ldots, x_{n_r}$$

and solving the equations in turn, beginning with the last equation, for

$$x_{n_r}, x_{n_{r-1}}, \ldots, x_{n_2}, x_{n_1}$$

When the system has solutions it is said to be consistent. If any one of $e_{r+1}, e_{r+2}, \ldots, e_m$ is not zero, then the equations do not have a solution and are inconsistent. In example (2), the system is equivalent to the system

$$\begin{array}{r} x_1 + 3x_2 + 4x_3 - 3x_4 - x_5 = 4 \\ x_2 + {}^8/_7x_3 - {}^5/_7x_4 - {}^4/_7x_5 = {}^6/_7 \\ x_4 + {}^1/_3x_5 = -{}^2/_3 \end{array}$$

Arbitrary values c_5 and c_3 can be chosen for x_5 and x_3, respectively, and the equations can be solved successively for x_4, x_2, and x_1.

The operations used in reducing system (1) to system (3) can be most effectively performed on the rows of the matrix of system (4), and they are pre-

$$A* = \begin{pmatrix} a_{11}\ a_{12} \cdots a_{1n}\ b_1 \\ a_{21}\ a_{22} \cdots a_{2n}\ b_1 \\ \cdots\cdots\cdots\cdots\cdots \\ a_{m1}\ a_{m2} \cdots a_{mn}\ b_m \end{pmatrix} \tag{4}$$

cisely the elementary row transformations of a matrix. *See* MATRIX THEORY.

Matrix (4) is called the augmented matrix of the system, and matrix (5) is called the matrix of the co-

$$A = \begin{pmatrix} a_{11}\ a_{12} \cdots a_{1n} \\ a_{21}\ a_{22} \cdots a_{2n} \\ \cdots\cdots\cdots\cdots \\ a_{m1}\ a_{m2} \cdots a_{mn} \end{pmatrix} \tag{5}$$

efficients. A necessary and sufficient condition for the existence of a solution of system (1) is that the rank of the augmented matrix A^* be the same as the rank of the matrix of coefficients A. (In general, the rank of a matrix A may be defined as the largest order r of a nonvanishing minor of A, where a minor of order r is the determinant of an $r \times r$ matrix that

consists of the elements common to r rows and r columns selected from A.) In example (2), the rank of the augmented matrix

$$A^* = \begin{pmatrix} 2 & -1 & 0 & -1 & 2 & 2 \\ 1 & 3 & 4 & -3 & -1 & 4 \\ 3 & 2 & 4 & 2 & 3 & 2 \end{pmatrix}$$

as well as the rank of

$$A = \begin{pmatrix} 2 & -1 & 0 & -1 & 2 \\ 1 & 3 & 4 & -3 & -1 \\ 3 & 2 & 4 & 2 & 3 \end{pmatrix}$$

is equal to 3. Therefore the system is consistent and has the solutions given earlier.

If $m = n$ in system (1) and if the determinant $|A|$ of the matrix of the coefficients A is not zero, then system (1) has a unique solution which can be obtained by Cramer's rule, which gives the values of the unknowns $x_1, x_2, \ldots, x_n$ as the ratio of two determinants. *See* DETERMINANT.

Specifically, Cramer's rule gives

$$x_1 = \begin{vmatrix} b_1 & a_{12} & a_{1n} \\ b_2 & a_{22} & a_{2n} \\ \cdots & \cdots & \cdots \\ b_n & a_{n2} & a_{nn} \end{vmatrix} \Big/ |A|$$

$$x_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} & \cdots & a_{1n} \\ a_{21} & b_2 & a_{23} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \\ a_{n1} & b_n & a_{n3} & \cdots & a_{nn} \end{vmatrix} \Big/ |A|$$

.....

$$x_n = \begin{vmatrix} a_{11} & \cdots & a_{1n-1} & b_1 \\ a_{21} & \cdots & a_{2n-1} & b_2 \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & \cdots & a_{nn-1} & b_n \end{vmatrix} \Big/ |A|$$

In every case the determinant in the denominator is $|A|$, and for the value of the unknown x_i, the determinant in the numerator is the determinant of the matrix obtained from A by replacing the ith column of A by $b_1, b_2, \ldots, b_n$. If $|A| = 0$ in this case of n equations in n unknowns, the system may either be inconsistent or it may have infinitely many solutions. In the example

$$\begin{aligned} x_1 - 4x_2 + 5x_3 &= 6 \\ 2x_1 - 11x_2 + 10x_3 &= -1 \\ 4x_1 - x_2 + 5x_3 &= 0 \end{aligned}$$

$$|A| = \begin{vmatrix} 1 & -4 & 5 \\ 2 & -11 & 10 \\ 4 & -1 & 5 \end{vmatrix} = 45$$

and the unique solution is given by

$$x_1 = \begin{vmatrix} 6 & -4 & 5 \\ -1 & 11 & 10 \\ 0 & -1 & 5 \end{vmatrix} \Big/ 45 = -{}^{285}/_{45} = {}^{19}/_{3}$$

$$x_2 = \begin{vmatrix} 1 & 6 & 5 \\ 2 & -1 & 10 \\ 4 & 0 & \end{vmatrix} \Big/ 45 = {}^{194}/_{45} = {}^{13}/_{3}$$

$$x_3 = \begin{vmatrix} 1 & -4 & 6 \\ 2 & -11 & -1 \\ 4 & -1 & 0 \end{vmatrix} \Big/ 45 = {}^{267}/_{45} = {}^{89}/_{15}$$

If in system (1), $b_1 = b_2 = \cdots = b_m = 0$, system (6)

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned} \tag{6}$$

is called homogeneous. Since the augmented matrix of system (6) differs from the matrix of the coefficients by an additional column of zeros, these matrices always have the same rank, so that a homogeneous system is always consistent. It has the obvious solution $x_1 = 0, x_2 = 0, \ldots, x_n = 0$, which is called the trivial solution. Any solution with at least one $x_i \neq 0$ is called nontrivial. A homogeneous system of m equations in n unknowns has a nontrivial solution if and only if the rank of the matrix of coefficients A is less than n. In particular, if m is less than n, the rank of A is at most m, and the system has a nontrivial solution. Further, if $m = n$, the system has a nontrivial solution if and only if the determinant $|A| = 0$. In system (7)

$$\begin{aligned} 3x_1 + 4x_2 - 2x_3 &= 0 \\ -2x_1 + 3x_2 - 4x_3 &= 0 \\ 5x_1 + x_2 + 2x_3 &= 0 \\ -9x_1 + 5x_2 - 10x_3 &= 0 \end{aligned} \tag{7}$$

the rank of A is 2, so that system (7) has a nontrivial solution:

$$A = \begin{pmatrix} 3 & 4 & -2 \\ -2 & 3 & -4 \\ 5 & 1 & 2 \\ -9 & 5 & -10 \end{pmatrix}$$

The rank of A is seen to be 2 by performing elementary row transformations on A to reduce it to the form

$$\begin{pmatrix} 1 & {}^{4}/_{3} & -{}^{2}/_{3} \\ 0 & 1 & -{}^{16}/_{17} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

System (7) is equivalent to the system

$$\begin{aligned} x_1 + {}^{4}/_{3}x_2 - {}^{2}/_{3}x_3 &= 0 \\ x_2 - {}^{16}/_{17}x_3 &= 0 \end{aligned}$$

The solutions are given by $x_1 = -{}^{10}/_{\sqrt{17}}c_3$, $x_2 = {}^{16}/_{\sqrt{17}}c_3$, $x_3 = c_3$, where c_3 is an arbitrarily chosen number.

In the event that a linear system (1) has infinitely many solutions, they can be given as in example (2) or, alternatively, a finite number of solutions can be listed in terms of which all solutions can be conveniently obtained. Any one fixed solution of system (1) is called a particular solution. In example (2), $x_1 = {}^{6}/_{7}$, $x_2 = {}^{8}/_{\sqrt{21}}$, $x_3 = 0$, $x_4 = -{}^{2}/_{3}$, $x_5 = 0$ is a particular solution. If system (1) is consistent, every solution is of the form $x_1 = c_1 + c_1^{(0)}$, $x_2 = c_2 + c_2^{(0)}$, $\ldots$, $x_n = c_n + c_n^{(0)}$, where $c_1, c_2, \ldots, c_n$ is a fixed particular solution of system (1) and $c_1^{(0)}, c_2^{(0)}, \ldots, c_n^{(0)}$ is any solution of the corresponding homogeneous system (6). By virtue of this result, it is sufficient to investigate the solutions of the homogeneous system (6).

Let $C^{(1)} = (c_1^{(1)}, c_2^{(1)}, \ldots, c_n^{(1)})$ and $C^{(2)} = (c_1^{(2)}, c_2^{(2)}, \ldots, c_n^{(2)})$ be solutions of the homogeneous system (6), written as n-tuples. Then $C^{(1)} + C^{(2)} = (c_1^{(1)} + c_1^{(2)}, c_2^{(2)} + c_2^{(2)}, \ldots, c_n^{(1)} + c_n^{(2)})$ is the sum of the solutions, and $eC^{(1)} = (ec_1^{(1)}, ec_2^{(1)}, \ldots, ec_n^{(1)})$ where e is a fixed number and is a scalar multiple of the solution $C^{(1)}$. It is evident that any linear combination $e_1C^{(1)} + e_2C^{(2)} + \cdots + e_sC^{(s)}$ of solutions $C^{(1)}, C^{(2)}, \ldots, C^{(s)}$, of system (6) is again a solution of system (6). A set of solutions $C^{(1)}, C^{(2)}, \ldots, C^{(s)}$, of the homogeneous system (6) is called a linearly independent set if $e_1C^{(1)} + e_2C^{(2)} + \cdots + e_sC^{(s)} = (0, 0, \ldots, 0)$ implies that $e_1 = e_2 = \cdots = e_s = 0$. Otherwise the solutions form a linearly dependent set, that is, there exist numbers $e_1, e_2, \cdots, e_s$, not all zero such that $e_1C^{(1)} + e_2C^{(2)} + \cdots + e_sC^{(s)} = (0, 0, \ldots, 0)$. A set of solutions is a maximal linearly independent set if the set is linearly independent and if a set formed by adding any other solution of system (6) is linearly dependent. If the rank of the matrix A of system (6) is r, then the number of solutions in a maximal independent set is $n - r$. Every solution of system (6) is a linear combination $e_1C^{(1)} + e_2C^{(2)} + \cdots + e_{n-r}C^{(n-r)}$ of the solutions in a maximal independent set. In example (2), the corresponding homogeneous system is system (8). The rank of the matrix A of system (8)

$$\begin{aligned} 2x_1 - x_2 \qquad\quad - x_4 + 2x_5 &= 0 \\ x_1 + 3x_2 + 4x_3 - 3x_4 - x_5 &= 0 \\ 3x_1 + 2x_2 + 4x_3 + 2x_4 + 3x_5 &= 0 \end{aligned} \tag{8}$$

is 3. Thus there are $5 - 3 = 2$ solutions in a maximal linearly independent set. They can be given as $C^{(1)} = (-4, -8, 7, 0, 0)$ and $C^{(2)} = (-3, 1, 0, -1, 3)$. Every solution of system (8) is given by $e_1C^{(1)} + e_2C^{(2)} = (-4e_1 - 3e_2, -8e_1 + e_2, 7e_1, -e_2, 3e_2)$ for arbitrary e_1 and e_2. Every solution of the original system is given by $({}^{6}/_{7} - 4e_1 - 3e_2, {}^{8}/_{\sqrt{21}} - 8e_1 + e_2, 7e_1, -{}^{2}/_{3} - e_2, 3e_2)$.

In example (7), the rank of the matrix is 2, and there is $3 - 2 = 1$ solution in a maximal linearly independent set. It can be given as $C^{(1)} = (-10, 16, 17)$, and every solution of (7) is a multiple of $C^{(1)}$.

If the system of equations (1) has a solution, then it has a rational solution. This means that there is a solution which is obtained by employing only the rational operations of addition, subtraction, multiplication, and division on the coefficients of the system. The particular solution as well as the linearly independent solutions of the corresponding homogeneous system can always be selected rationally. This is evident since the operations performed on the system to obtain a solution are rational operations. *See* EQUATIONS, THEORY OF; POLYNOMIAL SYSTEMS OF EQUATIONS. Ross A. Beaumont

Bibliography. J. B. Fraleigh and R. A. Beauregard, *Linear Algebra*, 3d ed., 1995; P. K. Rees et al., *College Algebra*, 10th ed., 1990.

Linearity

A relationship between two or more quantities which can be expressed in terms of linear algebraic, differential, or integral equations. A system in which all quantities (or variables) can be described in terms of such equations is said to be a linear system. By definition, linear systems satisfy the principle of superposition. By this principle, the response of a linear system to multiple inputs is given simply by the sum of the responses due to each individual input. In addition, if all inputs are multiplied by a common constant factor, the resulting response is multiplied by the same factor. *See* DIFFERENTIAL EQUATION; INTEGRAL EQUATION; LINEAR ALGEBRA; LINEAR SYSTEMS OF EQUATIONS.

As an example of a simple linear relationship, the voltage V across an ideal (ohmic) resistor is directly proportional to the current I through it, as given by Eq. (1), where R is the constant of proportionality.

$$V(I) = IR \tag{1}$$

Here, the customary notation $V(I)$ is used to denote V as a function of I. The voltage corresponding to a scaled input current, kI, is obtained by applying Eq. (2). A sum of input currents, $I_1 + I_2 + I_3$, gives

$$V(kI) = kIR = kV(I) \tag{2}$$

Eq. (3). Clearly, the principle of superposition is observed in this case. *See* OHM'S LAW; SUPERPOSITION THEOREM (ELECTRIC NETWORKS).

$$\begin{aligned} V(I_1 + I_2 + I_3) &= (I_1 + I_2 + I_3)R \\ &= I_1R + I_2 + I_3R \\ &= V(I_1) + V(I_2) + V(I_3) \end{aligned} \tag{3}$$

In contrast, a semiconductor diode has a nonlinear I–V relationship of the form of Eq. (4), where V_0 and I_0 are constants. In this case, inequalities (5) and (6)

$$V(I) = V_0 \ln \frac{I}{I_0} \tag{4}$$

$$V(kI) \neq kV(I) \tag{5}$$

$$V(I_1 + I_2 + I_3) \neq V(I_1) + V(I_2) + V(I_3) \tag{6}$$

hold, and the superposition principle does not apply. *See* SEMICONDUCTOR DIODE.

Linearity is a desirable characteristic of all systems where an output response is required to be a faithful reproduction (except for a constant scale factor) of one or more inputs. For example, electronic amplifiers used in measurement and signal transmission and reproduction systems are designed with linearity as a primary goal. Although physical systems are generally nonlinear to some degree, in practice the objective is to realize a good approximation to an ideal linear system by minimizing nonlinearities as far as possible. Any departure from linearity in these systems causes unwanted distortion of the original signal and results in a degraded and erroneous response. *See* AMPLIFIER; DISTORTION (ELECTRONIC CIRCUITS).

Ashok P. Nedungadi

Bibliography. J. D. Bell, *Mathematics of Linear and Nonlinear Systems for Engineers and Applied Scientists*, 1990; N. K. Sinha, *Linear Systems*, 1991.

Linen

A widely used cloth made from flax fibers. Linen is noted for its evenness of thread, fineness, and density. Its uses include garments, tablecloths, sheeting, towels, and thread.

The flax plant which supplies the fiber needs a short, cool growing season with plenty of rainfall. Russia leads the world in growing flax for fiber. The manufacture of the fiber into fabric requires unusual care during each process to retain the strength and beauty of the fiber. *See* FLAX.

Processing the fiber. Seeds and leaves are removed from the stems of the flax plant by passing the stalks through coarse combs, a process called rippling. Bundles of stems weighted down with heavy stones to ensure complete immersion are then steeped in water. The wetting allows the bark surrounding the hairlike flax fibers to decompose, thus loosening the gums that hold the fibers. This decomposing, or fermentation, is called retting.

After removal from water the stems may be spread out on the ground for a week or more. After the woody tissue has become dry, it is crushed by fluted iron rollers. Then there remain only small pieces of wood called shives. The shives are removed by means of rotating beaters, thus finally releasing the flax fibers. This can be done by hand or with machinery. The simple combing process, known as hackling, straightens the flax fibers, separates the short from the long staple, and leaves the longer fibers in parallel formation. For coarse linen, hackling is usually done by hand. Fine linen is also hackled by machine, a finer comb being used with each hackling treatment.

Properties of flax fiber. Under a microscope the flax fiber shows hairlike cylindrical filaments with fine pointed ends. The filaments are cemented together by gummy pectin. The long flax filaments contain a lumen, or central canal. The fiber resembles a straight, smooth bamboo stick, with joints producing a slight natural unevenness that cannot be eliminated. Chemically the flax fiber is about 70% cellulose, 25% pectin, and the remainder of woody tissue and ash. From such fibers linen yarns are produced with a smooth, straight, compact, and lustrous appearance. *See* CELLULOSE; NATURAL FIBER; PECTIN; TEXTILE.

Elton G. Nelson

Lines of force

Imaginary lines in fields of force **F** whose tangents at any point *P* give the direction of the field at *P* and whose number through unit area perpendicular to this direction represents the intensity of the field at *P* (see **illustration**). The concept of lines of force is perhaps most common when dealing with electric or magnetic fields.

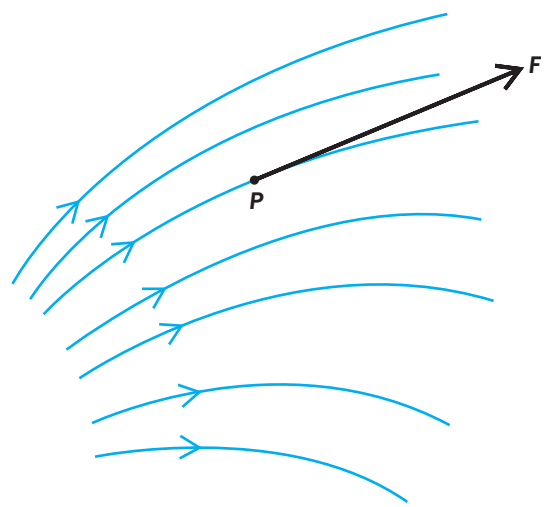

Definition of lines of force.

Electric lines of force. These are lines drawn to represent, or map, an electric field graphically in the space around a charged body. They are of great help in visualizing an electric field and in quantitative thinking about such a field. The direction of an electric line of force at any point is drawn parallel to the direction of the electric field intensity **E** at that point. The quantitative convention is that the number of electric lines of force drawn through an imaginary unit area of surface perpendicular to the field shall be numerically equal to **E** at that area. From this quantitative convention, and using International System (SI) units, in free space q/ϵ_0 lines of force originate on a positive charge q (in coulombs) and a like number terminate on a negative charge $-q$ of the same magnitude. Here $\epsilon_0 = 8.85 \times 10^{-12}$ farad/m is the permittivity of empty space. *See* ELECTRIC FIELD.

Magnetic lines of force. A magnetic field of strength **H** may also be represented by lines called lines of force. The number of lines of force per unit area of a surface perpendicular to **H** is numerically equal to the value of **H**. The lines of magnetic force due to currents are closed curves, as are the lines of magnetic induction. Magnetic lines of force due to magnets originate on north poles and terminate on south poles, both inside and outside the magnet. *See* MAGNETIC FIELD.

Ralph P. Winch; Bryan P. Kibble

Bibliography. D. J. Griffiths, *Introduction to Electrodynamics*, 3d ed., 1999; A. F. Kip, *Fundamentals of Electricity and Magnetism*, 2d ed., 1969; E. M. Purcell, *Electricity and Magnetism*, Berkeley Physics Course, vol. 2, 2d ed., 1985; W. T. Scott, *The Physics of Electricity and Magnetism*, 2d ed., 1966, reprint 1976.

Linewidth

A measure of the width of the band of frequencies of radiation emitted or absorbed in an atomic or molecular transition. One of the dominant sources of electromagnetic radiation of all frequencies is transitions between two energy levels of an atomic or molecular system. The frequency of the radiation is related to the difference in the energy of the two levels by the Bohr relation, Eq. (1), where ν_0 is the frequency

$$\nu_0 = \frac{E_1 - E_2}{h} \tag{1}$$

of the radiation, h is Planck's constant, and E_1 and E_2 are the energies of the levels. This radiation is not monochromatic, but consists of a band of frequencies centered about ν_0 whose intensity $I(\nu)$ can be characterized by the linewidth. The linewidth is the full width at half height of the distribution function $I(\nu)$. The simplest case is for a transition from an excited state to the ground state for an atom or molecule at rest. For this case, the normalized distribution function is the lorentzian line profile given by Eq. (2). Here $\Delta\nu$ is the full width at half maximum

$$I(\nu) = \frac{1}{\pi}\frac{\Delta\nu/2}{(\nu-\nu_0)^2 + (\Delta\nu/2)^2} \tag{2}$$

(FWHM). The FWHM is related to the lifetime τ of the excited level through Eq. (3). This is a manifesta-

$$(\Delta\nu)(\tau) = \frac{1}{2\pi} \tag{3}$$

tion of the quantum-mechanical uncertainty principle, and the linewidth $\Delta\nu$ is referred to as the natural linewidth. For a transition between two unstable levels with lifetimes τ_1 and τ_2, the FWHM is given by Eq. (4).

$$\Delta\nu = \frac{1}{2\pi}\left(\frac{1}{\tau_1} + \frac{1}{\tau_2}\right) \tag{4}$$

See ENERGY LEVEL (QUANTUM MECHANICS); QUANTUM MECHANICS; UNCERTAINTY PRINCIPLE.

Another major source of line broadening for atomic and molecular transitions is the Doppler shift due to thermal motion. To first-order in v/c, the frequency of the radiation depends upon whether the radiating object is moving away from or toward the observer according to Eq. (5). Here v is the com-

$$\nu = \nu_0\left(1 - \frac{v}{c}\right) \tag{5}$$

ponent of the velocity along the line joining the observer and the radiating object taken as positive when the separation is increasing, and c is the speed of light. For a gaseous radiator whose velocity distribution is a maxwellian one characterized by a most probable velocity given by Eq. (6), where k is the

$$v_p = \left(\frac{2kT}{M}\right)^{1/2} \tag{6}$$

Boltzmann constant, T the absolute temperature, and M the mass of the radiating molecule, the distribution of radiation is given by Eq. (7). This is a gaussian

$$I(\nu) = I_0 \exp\{-[c(\nu-\nu_0)/\nu_0 v_p]^2\} \tag{7}$$

line profile with FWHM given by Eq. (8). For most sit-

$$\begin{aligned}(\Delta\nu)_D &= \frac{(2\sqrt{\ln 2})\nu_0 v_p}{c} \\ &= \left(\frac{v_0}{c}\right)\left(\frac{8kT\ln 2}{M}\right)^{1/2}\end{aligned} \tag{8}$$

uations the Doppler width is greater than the natural linewidth. A more accurate line profile is the Voigt line profile, a convolution of the lorentzian and gaussian profiles given by Eq. (9).

$$I(\nu) = (\text{const})\int_0^\infty \frac{\exp\{-[c(\nu'-\nu_0)/\nu_0 v_p]^2\}}{(\nu-\nu')^2 + (\Delta\nu/2)^2}\,d\nu' \tag{9}$$

See BOLTZMANN STATISTICS; DISTRIBUTION (PROBABILITY); DOPPLER EFFECT.

A third major source of line broadening is collisions of the radiating molecule with other molecules. This broadens the line, shifts the center of the line, and shortens the lifetime of the radiating state. The net result is a lorentzian line profile of the form of Eq. (10), where Γ is given by Eq. (11). Here $(\Delta\nu)_N$ is the

$$I(\nu) = I_0\frac{\Gamma^2}{[\nu-\nu_0-(\delta\nu)_{EC}]^2 + \Gamma^2} \tag{10}$$

$$\Gamma = \frac{(\Delta\nu)_N + (\Delta\nu)_{IC}}{2} + (\Delta\nu)_{EC} \tag{11}$$

natural linewidth, $(\Delta\nu)_{IC}$ the width due to inelastic collisions, $(\Delta\nu)_{EC}$ the width due to elastic collisions, and $(\delta\nu)_{EC}$ the shift of the center due to elastic collisions. The combination of the collision width, the natural width, and the Doppler width yields a generalized Voigt profile. This form is used to analyze data from plasmas and stellar atmospheres to determine the temperature of the radiator and the density of the perturbers. *See* PLASMA DIAGNOSTICS.

In many measurements the linewidth limits the precision with which the center of the line can be determined. A number of techniques have been invented to circumvent the limitation due to the Doppler width. One technique, called Dicke narrowing, is to confine the radiating atoms to a volume whose characteristic dimensions are much less than the wavelength of the radiation. *See* LASER SPECTROSCOPY.

For radiating atoms in a liquid or solid the width is usually dominated by the strong interaction of the radiator with the surrounding molecules. This usually results in a spectrum with a wide distribution which cannot be characterized by a simple lorentzian or

gaussian line profile. The inelastic interactions of the excited molecule with its surroundings produce radiationless transitions and shorten the lifetime of the excited state. There are often a variety of radiation sites with different characteristic frequencies. The net result is a broad line profile with a complex structure. *See* BAND THEORY OF SOLIDS; CRYSTAL ABSORPTION SPECTRA.

The above discussion deals principally with the situations where radiation is emitted by an excited molecule. The considerations are similar for the inverse reaction where a ground-state molecule absorbs radiation and makes a transition to an excited state. *See* ATOMIC STRUCTURE AND SPECTRA; MOLECULAR STRUCTURE AND SPECTRA. Francis M. Pipkin

Bibliography. R. C. Breen, Jr., *Theories of Spectral Line Shape*, 1981; W. Demtröder, *Laser Spectroscopy*, 2d ed., 1994.

Linguistics

The science, that is, the general and universal properties, of language. Interest in language seems to have arisen in the earliest history of the human species. Every culture has left records that reveal philosophical or practical concerns for this unique human characteristic, and these concerns have varied over time. The middle of the twentieth century saw a shift in the principal direction of linguistic inquiry from one of data collection and classification to the formulation of a theory of generative grammar, which focuses on the biological basis for the acquisition and use of human language and the universal principles that constrain the class of all languages.

Generative grammar distinguishes between the knowledge of language (linguistic competence), which is represented by mental grammar, and the production and comprehension of speech (linguistic performance). Theories of generative grammar include Government and Binding or Principles and Parameters Theory, Generalized Phrase Structure, and Lexical Functional Grammar, among others.

Grammar. If grammar is defined as the mental representation of linguistic knowledge, then a general theory of language is a theory of grammar. A grammar includes everything one knows about a language; its phonetics and phonology (the sounds and the sound system), its morphology (the structure of words), its lexicon (the words or vocabulary), its syntax (the structure of sentences and the constraints on well-formed sentences), and its semantics (the meaning of words and sentences).

Phonetics and phonology. Different languages use different inventories of discrete speech sounds. Speech is, however, a continuous acoustic signal, not a string of separate sounds. Yet speakers know that a word like "cat" is composed of three sounds represented by the letters c, a, and t. This illustrates in a simple way the difference between knowledge of language and linguistic performance. Phonetics is the study of the production or articulation and perception of speech as well as the acoustic characteristics of the sounds produced. *See* PSYCHOACOUSTICS; SPEECH; SPEECH PERCEPTION.

Speech reflects the more abstract system of phonological knowledge that determines, for example, why the letter g is pronounced in "signature" but not in "sign" or why stress is placed on the first word in the "White House," the official residence of the President of the United States, and on the second in "white house" a designation for any residence that is white.

Morphology. Morphology determines how words are constructed. The word "undesirability," for example, is a complex structure composed of the prefix "un," the root "desire," the suffix "able," and the final suffix "ity."

Syntax. Syntax, an important component of grammar and linguistic theory, concerns the structure of sentences, that is, the order and ways in which words, phrases, and clauses can be combined. It accounts for the ability of speakers to produce and comprehend an infinite number of sentences. Knowledge of syntax accounts for an English speaker knowing that a sentence such as "Linguistics is interesting" is acceptable but "interesting linguistics is" is not; it also explains why and how a sentence can be lengthened indefinitely; for example, "I know that linguistics is interesting," "You know that I know that linguistics is interesting," and so on. Syntactic principles account for the fact that in a sentence such as "Mary thinks that she is pretty," the word "she" may refer to Mary or to someone else, but in the sentence "She thinks that Mary is pretty," the word "she" cannot refer to Mary. Syntax also explains why the sentence "George put the book on the shelf," is grammatical but "George put the book" is not. The words in the mental lexicon contain syntactic information in addition to information about word meanings and pronunciation. Thus, a word like "put" must be followed by both an object and a prepositional phrase.

Semantics. Semantics is the study of the meaning of words and sentences and the relationship between logic and language, a subject of great interest in the Middle Ages and the seventeenth century. Semantics cannot be independent of syntax because meaning often depends on syntactic structure. Thus, a sentence like "John likes Mary more than Joan" has two meanings because of its structure and not because any of the words has two meanings. The semantic component of grammar accounts for the fact that sentences may be paraphrases of each other (that is, are semantic equivalents), as in "Look up the word in the dictionary" and "Look the word up in the dictionary"; or may be anomalous, as in "My sister is an only child"; or may entail other sentences (for example, "It is a red flower" entails the sentence "It is a flower").

Branches. Liquistics is not limited to grammatical theory. Descriptive linguistics analyzes the grammars of individual languages; for example, Indo-European linguistics, Romance linguistics, and African linguistics refer to the studies of particular languages and

language families, both from historical and synchronic points of view. In a more general sense, this area of linguistics investigates why and how languages develop.

Anthropological linguistics, or ethnolinguistics, and sociolinguistics focus on languages in relation to culture, social class, race, and gender. Dialectologists investigate how these factors fragment one language into many. In addition, sociolinguists and applied linguists examine language planning, literacy, bilingualism, and second-language acquisition. The study of Creoles and pidgins, languages that develop out of the need for speakers of different languages to communicate, has proved to provide important insights into language structure and change.

Computational linguistics encompasses automatic parsing, machine processing, and computer simulation of grammatical models for the generation and parsing of sentences. If viewed as a branch of artificial intelligence, computational linguistics has as its goal the modeling of human language as a cognitive system. *See* ARTIFICIAL INTELLIGENCE.

Historically, language has been considered a window into the mind. Whether language ability is distinct or arises in the human species out of more general cognitive capacities has yet to be determined. A branch of linguistics concerned with that question and with the biological basis of language development is neurolinguistics. Research on aphasia (language disorders following brain injury) reveals that damage to specific brain areas results in the disruption of distinct cognitive functions, as well as different motor and perceptual abilities, and that language may be disordered without disruption of other cognitive abilities. Furthermore, different parts of the language system (or grammar) may themselves be uniquely disturbed, a phenomenon that supports a modular concept of the brain, the mind, and the grammar itself in which the components interact but are independent. *See* INFORMATION PROCESSING (PSYCHOLOGY).

The form of language representation in the mind, that is, linguistic competence and the structure and components of the mental grammar, is the concern of theoretical linguistics. The branch of linguistics concerned with linguistic performance, that is, the production and comprehension of speech (or of sign language by the deaf), is called psycholinguistics. Psycholinguists also investigates how children acquire the complex grammar that underlies language use. *See* PSYCHOLINGUISTICS.

Children are not taught a language in the same way they are taught, for example, how to add. Nor are they taught the rules of grammar; those are acquired unconsciously. A child who incorrectly says "mouses" instead of "mice" or "bringed" instead of "brought" has learned the rules and is applying them to a case that is an exception to these rules. The fact that the complex grammar is acquired at a very early age, without being taught and before other seemingly simpler cognitive abilities are developed, supports the view that a child is genetically programmed for language acquisition. *See* COGNITION; LEARNING MECHANISMS.

Victoria A. Fromkin

Bibliography. N. Chomsky, *Knowledge of Language: Its Nature, Origin, and Use*, 1986; N. Chomsky, *Syntactic Structures*, 1957; V. Fromkin and R. Rodman, *An Introduction to Language*, 6th ed., 1998; F. J. Newmeyer, *Linguistic Theory in America*, 2d ed., 1986; F. J. Newmeyer (ed.), *Linguistics: The Cambridge Survey*, vols. 1–4, 1988.

Lingulida

An order of inarticulated brachiopods consisting of an exclusively marine group of lophophorate animals (animals with a tentacular food-gathering, respiratory, and protective organ): sessile benthic suspension-feeders possessing a chitinophosphatic shell. Representatives occur throughout the Phanerozoic Era (from the Early Cambrian to the present). Three superfamilies are recognized: Linguloidea, with the family Lingulidae, extends from the Middle Ordovician to the present; Discinoidea, with the family Discinidae, extends from the Early Ordovician to the present; and Acrotheloidea is known from the Early Cambrian to the Early Ordovician. *See* BRACHIOPODA; INARTICULATA.

General morphology. Shells are tongue-shaped (linguliform) to circular in outline. The valves are separate and lack articulation but possess a complex arrangement of muscles formed by paired anterior adductor muscles and single or paired posterior adductor muscles, and three or four pairs of oblique muscles. The pedicle emerges posteriorly between valves or through a foramen.

Lingulids are suspension feeders. The tentacular feeding organ (lophophore) occupies the mantle cavity. The digestive system consists of a mouth, pharynx, esophagus, stomach, digestive diverticula, pylorus, and intestine, and terminates in a functional anus. Four digestive diverticula open through ducts to the stomach. The excretory system consists of a pair of ciliated funnels (metanephridia) that act as gonoducts during spawning and allow the discharge of gametes from the coelom into the lophophoral cavity. Solid waste, enmeshed in mucus, is also ejected through the nephridiopores, while the excretory product ammonia is voided through the tissues of the mantle and lophophore.

Lingulids possess an open circulatory system of blood vessels and coelomic canals containing a faintly pigmented fluid, which is coagulable and contains free cellular inclusions consisting of blood cells and coelomocytes. In *Lingula*, the blood cells are somewhat round in shape and contain a respiratory protein that consists of hemoerythrins.

Lingulids have a central nervous system containing unsheathed nerves with no obviously differentiated sense organs, although statocysts have been reported in Lingulidae.

All extant members of this order are dioecious (that is, male and female reproductive organs are on separate individuals) and their larval development is

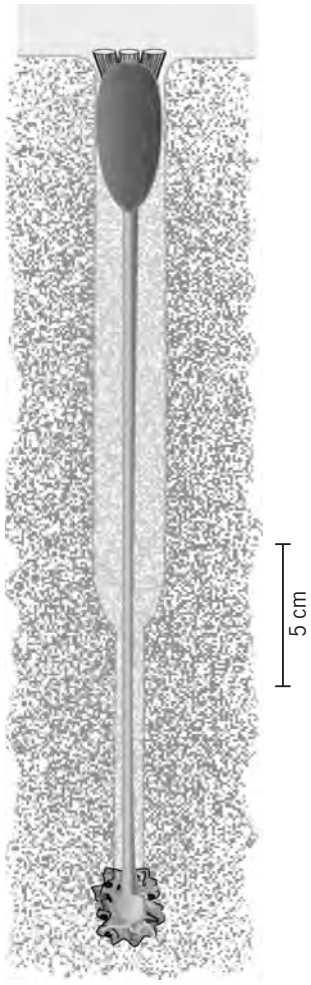

View of living *Lingula anatina* in its burrow. (*Courtesy of C. C. Emig*)

planktotrophic. They extend generally from inshore waters, mainly subtidal, to the bathyal zone.

Linguloidea. Shells are linguliform to circular in outline and are generally biconvex. The two extant genera belong to the family Lingulidae: *Lingula*, wrongly considered as "living fossils" from the Paleozoic, and *Glottidia*. Both genera originated at the beginning of the Tertiary. They have a long, fleshy extensible pedicle and burrow in soft sediments (see **illustration**). Their complex arrangement of muscles consists of two pairs of anterior and one single posterior adductor muscles, unpaired oblique muscles (three pairs and an single additional one), and a longitudinal lateral muscle, which help with movement of the valves.

Discinoidea and Acrotheloidea. Most shells are circular or subcircular in outline. The dorsal valve is usually larger than the ventral valve, and is convex and often conical. The ventral valve is flattened and disklike. Paired anterior and posterior adductor muscles and three pairs of oblique muscles make up the complex arrangement of muscles. A short pedicle in the postlarval stage emerges through a foramen in the ventral valve, but the adults lack a pedicle notch and pedicle. Christian C. Emig; Mark A. James

Bibliography. C. C. Emig, Ecology of inarticulated brachiopods. Biogeography of inarticulated brachiopods, in R. L. Kaesler (ed.), *Treatise on Invertebrate Paleontology*, pt. H revised: *Brachiopoda*, vol. 1, pp. 473–502, Geological Society of America, Boulder, and University of Kansas, Lawrence, 1997; C. C. Emig, Tools for linguloid taxonomy: The genus *Obolus* (Brachiopoda) as an example. *Carnets de Géologie-Notebooks on Geology*, Article 2002/01 (CG2002_A01_CCE), 2002; C. C. Emig, Proof that *Lingula* (Brachiopoda) is not a living-fossil, and emended diagnoses of the Family Lingulidae, *Carnets de Géologie-Notebooks on Geology*, Letter 2003/01 (CG2003_L01_CCE), 2003; R. L. Kaesler (ed.), *Treatise on Invertebrate Paleontology*, pt. H revised: *Brachiopoda*, vols. 1 and 2, Geological Society of America, Boulder, and University of Kansas, Lawrence, 1997 and 2000.

Link

An element of a mechanical linkage. A link may be a straight bar or a disk, or it may have any other shape, simple or complex. It is assumed, for simple analysis, to be made of unyielding material; that is, its shape does not change. The frame, or fixed member, of a linkage is one of the links, whether the frame is fixed relative to the Earth or relative to a movable body such as the chassis of an automobile.

Two links of a kinematic chain meet in a joint, or pair, by which these links are held together. Just as each joint is a pair having two elements, one from each link, so a link in a kinematic chain is the rigid connector of two or more elements belonging to different pairs.

Each free link has three degrees of freedom in the plane (two translational and one rotational). If the tail of one link is pin-joined to the head of its predecessor two degrees of freedom are lost. If several links are joined so as to form a closed loop, $2g$ degrees of freedom are lost, where g is the number of lower pair joints. If one member of an n-member chain is fixed, all three degrees of freedom of the fixed member are eliminated, the remaining degrees of freedom F being given by Eq. (1). Hence for a tri-

$$F = 3(n - 1) - 2g \tag{1}$$

angle $F = 0$ and for a chain of four links called simply a four bar linkage, $F = 1$. *See* DEGREE OF FREEDOM (MECHANICS).

Assuming that $F = 1$ is stipulated for a mechanism, and investigating n in Eq. (1), it is found that Eq. (2)

$$3n = 4 + 2g \tag{2}$$

results. Because 4 is even and $2g$ is even, n must always be even. There can be no pin-joint mechanism with five links; the next highest must have six links. *See* FOUR-BAR LINKAGE; LINKAGE (MECHANISM); MECHANISM. Douglas P. Adams

Bibliography. J. Bickford, *Mechanism for Intermittent Motion,* 1972; R. S. Hartenberg and J. Denavit,

Kinematic Synthesis of Linkages, 1964; G. D. Hiscox, *Mechanical Movements, Powers, and Devices,* 1983.

Linkage (genetics)

Failure of two or more genes to recombine at random as a result of their location on the same chromosome pair. Among the haploid products of a cell which has gone through meiosis, two genes located in the same chromosome pair remain in their two original combinations of alleles (parental) unless an odd number of exchanges of homologous segments occurred within the interval bounded by their loci. The incidence of exchanges of homologous segments at meiosis is roughly proportional to the length of the chromosome segment between two loci. The percentage of recombinants thus provides an estimate of this length and a basis for constructing gene maps on which linked loci are arranged in linear order and spaced out in proportion to the recombination percentages between them. The analogy is with a linear map of the stations along a railway line spaced out according to the time of travel between any two or more of them. By extension of the meaning of linkage to bacterial transduction or transformation, two bacterial genes are said to be linked when both are incorporated together in the chromosome of recipient cells more often than would be expected from mere coincidence of two separate incorporations. This is the consequence of their being frequently carried on the same transducing or transforming DNA fragment. *See* DEOXYRIBONUCLEIC ACID (DNA); MEIOSIS; TRANSDUCTION (BACTERIA); TRANSFORMATION (BACTERIA).

If at meiosis in an individual heterozygous at two loci on the same chromosome pair—for example, *AB*—less than one exchange occurs on the average between them, the proportions of haploid recombinant meiotic products—*aB* and *Ab*—will be less than 50%, and that of the parental meiotic products—*AB* and *ab*—correspondingly more than 50%. This defines the two loci as linked. If, however, two loci on the same chromosome pair are so far apart that on the average one or more exchanges occur between them at meiosis, the proportions of parental and recombinant meiotic products will be equal, that is, 50%. This is the same as for two loci on different chromosome pairs. Thus for two loci between which recombination is 50% there are no means of deciding, on the basis of the recombination percentage alone, whether they are on one and the same chromosome pair but sufficiently far apart, or on different pairs. For an explanation of why 50% is the upper limit of recombination *see* RECOMBINATION (GENETICS).

Thus, while evidence of linkage—less than 50% recombination—implies that two loci are on the same chromosome pair, absence of linkage—50% recombination—does not imply that they are not. Further evidence may reveal that two loci showing no linkage to each other are both linked to a third locus, and so on. This additional evidence permits the construction of linkage maps of all the loci belonging to one and the same linkage group. Linkage groups can then be homologized to chromosome pairs.

In human genetics, since the introduction of the techniques of somatic cell and molecular genetics, very often loci can be assigned to one and the same chromosome pair well before knowing anything about the recombination percentages between any two of them at gametogenesis, this is, before knowing whether or not they are linked in the precise arithmetical sense of the word. To overcome this ambiguity, the term syntenic, proposed by J. Renwick, is now in general use to define two loci carried on the same chromosome pair, irrespective of what is the recombination percentage between them at meiosis. *See* SOMATIC CELL GENETICS.

Genetic linkage is a structural relationship between two genes, that is, a restriction on their recombination at meiosis. It does not necessarily imply a functional relationship. However, in bacteria some of the structural and regulatory genes controlling enzymes sequentially involved in a metabolic pathway often turn out to be closely linked. With very few exceptions, such as homeobox gene clusters, this is not so in eukaryotes.

The biochemical and evolutionary significance of linkage is poorly understood. In the cases just mentioned in bacteria, linkage permits coordinated regulation of enzyme synthesis. In eukaryotes several linkages have been preserved over long periods of evolution, but it cannot be excluded that for some at least this is simply a matter of historical accident. The fact that more and more cases of linkage disequilibrium are coming to light, on the other hand, reflects selective advantages, possibly related to the preservation of genetic variation in natural populations.

G. Pontecorvo

Linkage (mechanism)

A set of rigid bodies, called links, joined together at pivots by means of pins or equivalent devices. A body is considered to be rigid if, for practical purposes, the distances between points on the body do not change. Linkages are used to transmit power and information. They may be employed to make a point on the linkage follow a prescribed curve, regardless of the input motions to the linkage. They are used to produce an angular or linear displacement $f(x)$, where $f(x)$ is a given function of a displacement x. *See* LINK.

If the links are bars the linkage is termed a bar linkage. In first approximations a bar may be treated as a straight line segment or a portion of a curve, and a pivot may be treated as a common point on the bars connected at the pivot. A common form of bar linkage is then one for which the bars are restricted to a given plane, such as a four-bar linkage (**Fig. 1**). A body pivoted to a fixed base and to one or more

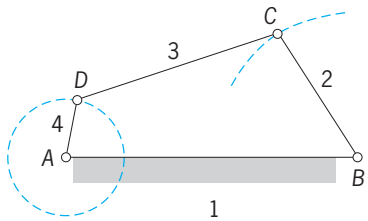

Fig. 1. Four-bar linkage.

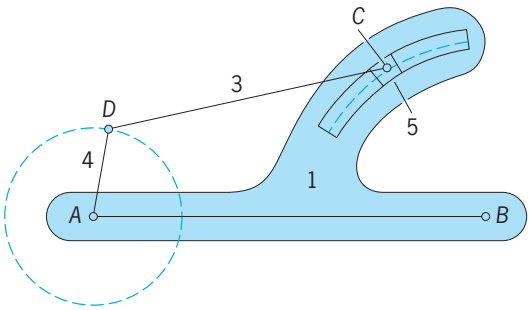

Fig. 2. Equivalent of four-bar linkage.

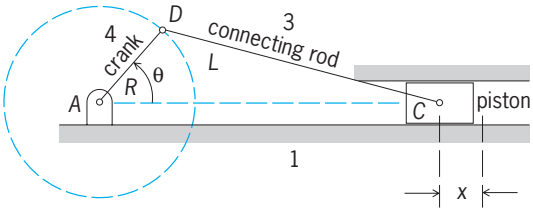

Fig. 3. Slider crank mechanism.

other links is a crank. As crank 4 in Fig. 1 rotates, link 2 (also a crank) oscillates back and forth. This four-bar linkage thus transforms a rotary motion into an oscillatory one, or vice versa. Link *AB*, marked 1, is fixed. Link 2 may be replaced by a slider 5 (**Fig. 2**). As end *D* of crank 4 rotates, pivot *C* describes the same curve as before. The slider moves in a fixed groove. *See* FOUR-BAR LINKAGE.

A commonly occurring variation of the four-bar linkage is the linkage used in reciprocating engines (**Fig. 3**). Slider *C* is the piston in a cylinder, link 3 is the connecting rod, and link 4 is the crank. This mechanism transforms a linear into a circular motion, or vice versa. The straight slider in line with the crank center is equivalent to a pivot at the end of an infinitely long link. Let *R* be the length of the crank, *L* the length of the connecting rod, while θ denotes the angle of the crank as shown in Fig. 3. Also, let x denote the coordinate of the pivot *C* measured so that $x = 0$ when $\theta = 0$. Length *L* normally dominates radius *R*, whence the approximate relation in Eq. (1) holds.

$$x = R(1 - \cos\theta) + \frac{R^2}{2L}\sin^2\theta \qquad (1)$$

A Scotch yoke is employed to convert a steady rotation into a simple harmonic motion (**Fig. 4**). The angle of the crank 2 with respect to the horizontal is denoted by θ, and the coordinate of the slider 4 by x. If *R* again denotes the length of the crank, Eq. (2) holds. In computer mechanisms, slider 3 can

$$x = R\cos\theta \qquad (2)$$

be moved along the crank so that the component of a vector can be obtained mechanically. *See* SLIDER-CRANK MECHANISM.

A lever is normally understood to be a bar connected to three other links. A lever is often used for addition (**Fig. 5**). Point *B* is at the center of lever *AC*. It is assumed that x and y are numerically small compared to the length of the lever. Levers with a fixed pivot (fulcrum) are very useful for amplifying or attenuating linear displacements and forces. *See* LEVER.

The pantograph shown in **Fig. 6** is a five-link mechanism employed in drafting to magnify or reduce diagrams. As point *Q* on link 5 describes a curve, point *P* on link 4 will describe a similar but enlarged curve. Here *ABCD* is a parallelogram. Pivot *A* is fixed. *See* PANTOGRAPH.

A mechanism composed of six links, introduced by Joseph Whitworth (1803–1887), makes the return stroke of a body in oscillatory linear motion faster than the forward stroke. Universal joints for

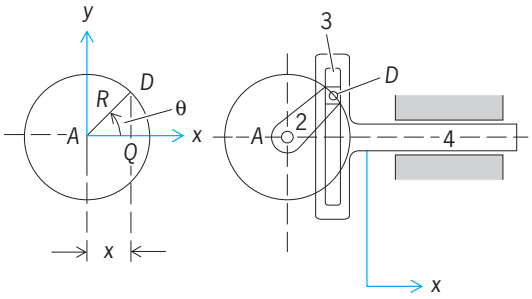

Fig. 4. Scotch yoke.

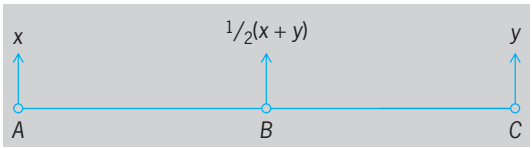

Fig. 5. Addition with a lever.

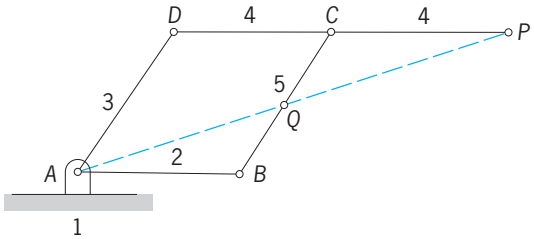

Fig. 6. Pantograph.

transmitting motion between intersecting axes and escapements permitting rotation about an axis in one direction only are examples of the many kinds of linkages used in machines. However, complicated linkages are becoming obsolete because of the ease with which information and power can be transmitted by electrical, hydraulic, and pneumatic means. *See* MECHANISM. Rufus Oldenburger

Bibliography. N. P. Chironis, *Mechanisms and Mechanical Devices Sourcebook*, 3d ed., 2001; H. H. Mabie and C. F. Reinh, *Mechanisms and Dynamics of Machinery*, 4th ed., 1987; J. E. Shigley and C. R. Mischke, *Mechanisms: A Mechanical Designer's Workbook*, 1990; J. E. Shigley and J. J. Viker, *Theory of Machines and Mechanisms*, 2d ed., 1994; C. H. Suh and C. W. Radcliffe, *Kinematics and Mechanisms Design*, 1978, reprint 1983.

Lipid

One of a class of compounds which contains long-chain aliphatic hydrocarbons (cyclic or acyclic) and their derivatives, such as acids (fatty acids), alcohols, amines, amino alcohols, and aldehydes. The presence of the long aliphatic chain as the characteristic component of lipids confers distinct solubility properties on the simpler members of this class of naturally occurring compounds. This led to the traditional definition of lipids as substances which are insoluble in water but soluble in fat solvents such as ether, chloroform, and benzene. However, those lipids (particularly glycolipids and phospholipids) which contain polar components in the molecule may be insoluble in these solvents, and some are even soluble in water. Most of the phosphatides are good emulsifying agents. *See* EMULSION; SOLVENT.

Classification. The lipids are generally classified into the following groups:

A. Simple lipids
   1. Triglycerides or fats and oils are fatty acid esters of glycerol. Examples are lard, corn oil, cottonseed oil, and butter.
   2. Waxes are fatty acid esters of long-chain alcohols. Examples are beeswax, spermaceti, and carnauba wax.
   3. Steroids are lipids derived from partially or completely hydrogenated phenanthrene. Examples are cholesterol and ergosterol.

B. Complex lipids
   1. Phosphatides or phospholipids are lipids which contain phosphorus and, in many instances, nitrogen. Examples are lecithin, cephalin, and phosphatidyl inositol.
   2. Glycolipids are lipids which contain carbohydrate residues. Examples are sterol glycosides, cerebrosides, and plant phytoglycolipids.
   3. Sphingolipids are lipids containing the long-chain amino alcohol sphingosine and its derivatives. Examples are sphingomyelins, ceramides, and cerebrosides.

The classification scheme is not rigid since sphingolipids are found which contain phosphorus and carbohydrate, and glycolipids which contain phosphorus. Those lipids not included in the above groupings are of the chemically simpler type and include the fatty acids, alcohols, ethers such as batyl and chimyl alcohols, an hydrocarbons such as the terpenes and carotenes.

The fatty acids found in lipids may be saturated or unsaturated, cyclic or acyclic, and contain substituents such as hydroxyl or keto groups. They are combined in ester or amide linkage. The alcohols, likewise, may be saturated or unsaturated and are combined in ester or ether linkages; in the case of α-β unsaturated ethers, these are best considered as aldehyde derivatives.The amino alcohols found in lipids are linked as amides, glycosides, or phosphate esters. Sphingosine is the predominant amino alcohol found in animal tissue lipids, and phytosphingosine is its equivalent in the plant kingdom.

Occurrence. Lipids are present in all living cells, but the proportion varies from tissue to tissue. The triglycerides accumulate in certain areas, such as adipose tissue in the human being and in the seeds of plants, where they represent a form of energy storage. The more complex lipids occur closely linked with protein in the membranes of cell and of subcellular particles. More active tissues generally have a higher complex lipid content; for example, the brain, liver, kidney, lung, and blood contain the highest concentration of phosphatides in the mammal. Fish oils are important sources of vitamins A and D, and seed-germ oils contain quantities of vitamin E. The vitamins K occur chiefly in plants and microorganisms, and carotenoids are widely distributed in both the plant and animal kingdoms. Waxes are found in insects and as protective agents on plant leaves and the cuticles of fruits and vegetables (for instance, the bloom on grapes and plums). Soybeans are used for the commercial preparation of phosphatides. *See* CELL MEMBRANES.

Extraction. Many of the lipids, particularly glycolipids, phosphatides, and steroid esters, are present in the tissue in association with other cell components, such as protein. The nature of the linkages in these complexes is not fully understood although they are most certainly not covalent. Combinations of lipid and protein, like lipoproteins and proteolipids, have been isolated from sources such as blood serum and egg yolk. For the efficient extraction of lipids from tissues, these complexes must first be disrupted, and this is usually accomplished by dehydration procedures such as freeze-drying or acetone extraction, or by denaturation with alcohol before extracting the tissue with suitable lipid solvents. The use of mixed solvents, such as ethanol and ether, ethanol and benzene, or chloroform and methanol, allows disruption, dehydration, and extraction in one operation. Partial fractionation of lipids may

be effected by successive extractions with different solvents.

Separation. Prior to 1948, solvent fractionations or the use of specific complexing agents, such as cadmium chloride for the separation of lecithin, were the only methods available for the separation of lipids, and progress in the chemical identification of individual lipids was slow. Since then, the use of column chromatography (particularly on alumina, silicic acid, and diethylaminoethyl cellulose), thin-layer chromatography, countercurrent distribution, and gas-phase chromatography has allowed rapid progress in separation, and many new classes of lipid have been discovered as a result of the application of these techniques. Other methods which allow selective degradation of certain lipids either chemically (for example, the use of dilute alkali to cause preferential hydrolysis of esters) or enzymatically are also being used successfully for the separation of certain lipids. *See* CHROMATOGRAPHY.

Identification. Most of the conventional analytical procedures are used in the identification of lipid samples. Physical methods, such as infrared and ultraviolet spectroscopy, optical rotation, proton magnetic resonance, and mass spectrometry, have been used. The advent of gas-phase chromatography allowed a major advance in lipid chemistry by providing a complete analysis of the fatty acid content of lipid mixtures. The high resolving power of thin-layer chromatography has accelerated the routine analysis of the mixture of lipids from many natural sources; qualitative and quantitative methods for the determination of esters, aldehydes, carbohydrates, long-chain bases, and the other important components of lipids such as glycerol, glycerophosphate, ethanolamine, choline, serine, amino sugars, and inositol; and elementary analyses for nitrogen and phosphorus. The nitrogen-to-phosphorus ratio of a phosphatide sample was once used as a criterion of purity because carbon and hydrogen analyses on a single phosphatide species are not highly revealing, owing to the spectrum of fatty acids that may be present. The degree of unsaturation is determined by hydrogenation or iodine number determination, and free hydroxyl groups are expressed by the acetyl. Lipid samples also may be hydrolyzed under a variety of acid or alkaline conditions and the products investigated by paper chromatography. *See* CARBOXYLIC ACID; FAT AND OIL; FAT AND OIL (FOOD); GLYCOLIPID; LIPID METABOLISM; PHOSPHOLIPID; SPHINGOLIPID; STEROID; TERPENE; TRIGLYCERIDE; VITAMIN; WAX, ANIMAL AND VEGETABLE.

Roy H. Gigg

Bibliography. W. W. Christie, *High Performance Liquid Chromatography and Lipids: A Practical Guide*, 1987; F. D. Gunstone et al., *The Lipid Handbook*, 2d ed., 1994; N. F. Hadley, *The Adaptive Role of Lipids in Biological Systems*, 1985; J. L. Harwood and N. J. Russell, *Lipids in Plants and Microbes*, 1984; M. Horisberger and V. Bracco (eds.), *Lipids in Modern Nutrition*, 1987; J. J. Kabara (ed.), *The Pharmacological Effects of Lipids II*, 1985; J. F. Mead et al., *Lipids: Chemistry, Biochemistry, and Nutrition*, 1986; R. Paoletti et al. (eds.), *Drugs Affecting Lipid Metabolism*, 1988.

Lipid metabolism

Lipids form a class of compounds composed of neutral fats (triglycerides), phospholipids, glycolipids, sterols, and fat-soluble vitamins (A, D, E, and K). As a group they share the common physical property of insolubility in water. This article describes the assimilation of dietary lipids and the synthesis and degradation of lipids by the mammalian organism.

Fat digestion and absorption. The principal dietary fat is triglyceride, an ester of three fatty acids and glycerol. This substance is not digested in the stomach and passes into the duodenum, where it causes the release of enterogastrone, a hormone which inhibits stomach motility. The amount of fat in the diet, therefore, regulates the rate at which enterogastrone is released into the intestinal tract. Fat, together with other partially digested foodstuffs, causes the release of hormones, secretin, pancreozymin, and cholecystokinin from the wall of the duodenum into the bloodstream.

Secretin causes the secretion of an alkaline pancreatic juice rich in bicarbonate ions, while pancreozymin causes secretion of pancreatic enzymes. One of these enzymes, important in the digestion of fat, is lipase. Cholecystokinin, which is a protein substance chemically inseparable from pancreozymin, stimulates the gallbladder to release bile into the duodenum. Bile is secreted by the liver and concentrated in the gallbladder and contains two bile salts, both derived from cholesterol: taurocholic and glycocholic acids. These act as detergents by emulsifying the triglycerides in the intestinal tract, thus making the fats more susceptible to attack by pancreatic lipase. In this reaction, which works best in the alkaline medium provided by the pancreatic juice, each triglyceride is split into three fatty acid chains, forming monoglycerides. The fatty acids pass across the membranes of the intestinal mucosal (lining) cells. Enzymes in the membranes split monoglyceride to glycerol and fatty acid, but triglycerides are reformed within the mucosal cells from glycerol and those fatty acids with a chain length greater than eight carbons: Short- and medium-chain fatty acids are absorbed directly into the bloodstream once they pass through the intestinal mucosa.

The triglycerides formed in the mucosal cells are associated with proteins and phospholipids and pass into the intestinal lymphatics as the lipoproteins called chylomicrons. The thoracic duct, the main lymphatic channel, carries the chylomicrons to the great veins, where they enter the circulatory system and can be taken up and metabolized by most tissues. *See* CHOLESTEROL; DIGESTIVE SYSTEM; GALLBLADDER; LIVER; PANCREAS.

The bile salts are absorbed from the ileum, together with dietary cholesterol. They enter the portal circulation, which carries them to the liver, where they can be reexcreted into the duodenum. The

bile also contains pigments which are derived from hemoglobin metabolism. *See* BILIRUBIN; HEMOGLOBIN.

Abnormalities in the pancreatic secretions, bile secretion, or intestinal wall can lead to either partial or complete blockage of fat absorption and an abnormal excretion of fats in the stool.

Adipose tissue. Accumulations of fat cells in the body are called adipose tissue. This is found beneath the skin, between muscle fibers, around abdominal organs and their supporting structures called mesenteries, and around joints. The fat cell contains a central fat vacuole so large that it fills the cell, pushing the nucleus and cytoplasm to the periphery. The adipose tissue of the body is of considerable magnitude and is quite active metabolically, functioning as a buffer in energy metabolism. *See* ADIPOSE TISSUE.

During meals carbohydrates, amino acids, and fats which are absorbed in excess of immediate energy requirements are converted and stored in the fat-cell vacuoles as triglycerides. These fat depots provide an economical storage form for energy requirements and serve as the source of energy between meals and during periods of fasting.

Triglycerides cannot traverse the cell membrane until they are broken down into glycerol and fatty acids. Lipoprotein lipase, an enzyme which catalyzes the hydrolysis of lipoprotein triglycerides, controls entry of triglycerides into the cell; it is induced by feeding and disappears during periods of fasting. The lipase which breaks down triglycerides before they leave the cell is controlled by hormones. Norepinephrine, a hormone released from sympathetic nerve endings, is probably the most important factor controlling fat mobilization. Epinephrine from the adrenal medulla also activates lipase, as do adrenocorticotropic hormone (ACTH) and glucagon. Growth hormone and a separate fat-mobilizing factor released from the pituitary gland of some species also stimulate lipolysis and release of fatty acids from adipose tissue. While the fat-mobilizing factor may activate the same enzyme as the hormones cited above, growth hormone appears to function through a different mechanism, probably requiring protein synthesis. *See* ADENOHYPOPHYSIS HORMONE; ENDOCRINE MECHANISMS; GLUCAGON.

Obesity is a condition in which excessive fat accumulates in the adipose tissue. One factor responsible for this condition is excessive caloric intake. The metabolic and psychological factors are under investigation. In starvation, uncontrolled diabetes, and many generalized illnesses the opposite occurs and the adipose tissue becomes markedly depleted of lipid. *See* DIABETES; METABOLIC DISORDERS.

Blood lipids. Lipids are present in the blood in the form of lipoproteins and as free fatty acids (FFAs), which are bound to albumin. The FFAs originate in adipose tissue and are mobilized for oxidation in other tissues of the body. The lipoproteins are small droplets of lipids complexed with proteins dispersed in the blood. They originate in the liver, with the exception of the small droplets, called chylomicrons, derived from the intestinal tract. The lipoproteins have been divided into arbitrary classes based upon their physical properties. The two methods of classification separate them according to density by ultracentrifugation and according to electrical charge by electrophoresis. The classes of lipoproteins defined by either of these methods differ in their content of triglycerides, cholesterol, and phospholipids. Some disorders of lipid metabolism are manifested in the blood by an abnormality in the pattern of lipoprotein classes. *See* CENTRIFUGATION; CLINICAL PATHOLOGY; ELECTROPHORESIS.

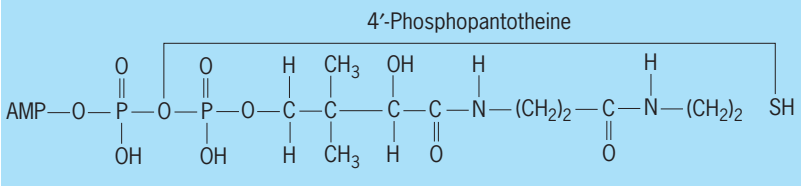

Fig. 1. Structural formula for coenzyme A.

Fatty acid oxidation. Fatty acids are oxidized to meet the energy requirements of the body by a β-oxidation pathway, that is, the second or β-carbon from the carboxyl group is oxidized. Before oxidation can proceed, the fatty acid must be coupled with coenzyme A (CoA), a compound containing adenosine monophosphate (AMP), two molecules of phosphate, and the complex nitrogen derivative pantotheine (**Fig. 1**).

In reaction (1), showing activation, the fatty acid is as shown in **Fig. 2**, in which only the terminal 4-carbon atoms are given. The initial activation proceeds in two steps. First, as shown in reaction (1*a*), the fatty acid reacts with adenosine triphosphate (ATP) to form fatty acid adenylate (fatty acid–AMP). This in turn reacts with CoA to form acyl CoA, as shown in reaction (1*b*). The enzyme that catalyzes reactions (1*a*) and (1*b*) is called the activating enzyme or thiokinase. The fatty acid–CoA cannot penetrate the mitochondrial membrane to gain access to the oxidative enzymes. To traverse the membrane, fatty acid must be esterified with the hydroxyl group of carnitine (**Fig. 3**), a reaction catalyzed by carnitine acyl transferases. The carnitine–fatty acid

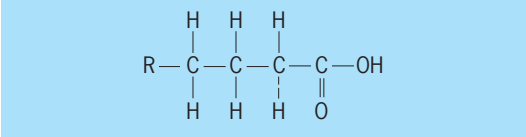

Fig. 2. Structural formula for a fatty acid. R represents chain of up to 14 carbon atoms.

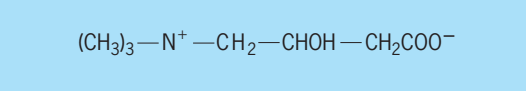

Fig. 3. Structural formula for carnitine.

Reaction *Enzyme*

Activation:

Fatty acid + ATP ⟶ Fatty acid—AMP + PP (1*a*)

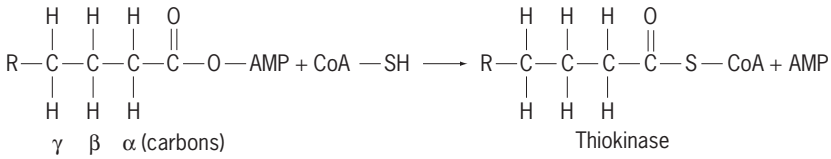

(1*b*)

Oxidation:

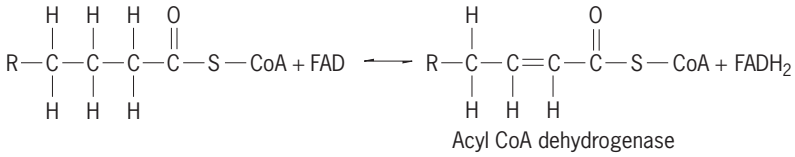

(2)

Hydration:

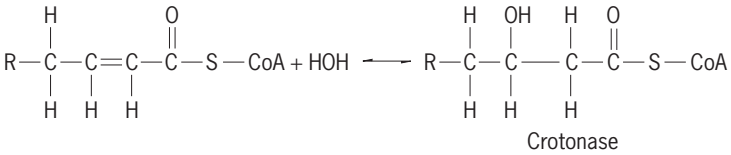

(3)

Oxidation:

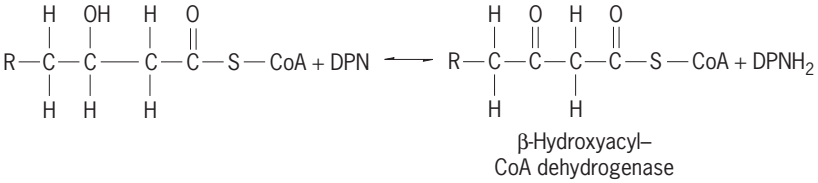

(4)

Cleavage:

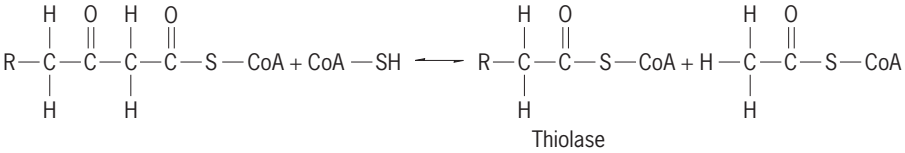

(5)

ester traverses the membrane, and the fatty acid is again coupled to CoA by a transferase on the other side.

In reaction (2), the first of two oxidation reactions, one hydrogen is removed from the α-carbon and one is removed from the β-carbon, producing an unsaturated fatty acid analogous to crotonic acid. The enzyme is called acyl CoA dehydrogenase, and the cofactor which acts as the oxidizing agent and accepts the hydrogen is flavin adenine dinucleotide (FAD).

In reaction (3), hydration, hydrogen and the hydroxyl group (from water) are introduced at the α- and β-carbons, respectively; a β-hydroxy acid is formed; the enzyme is crotonase.

In reaction (4), the second oxidation, two hydrogens are removed from the α and β carbons of the β-hydroxy acid to form a keto acid. The oxidizing agent is nicotinamide adenine dinucleotide (NAD), also called diphosphopyridine nucleotide (DPN). *See* NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD).

In reaction (5), showing cleavage, acetyl CoA is split off. Thus a fatty acid two carbons shorter than the original one is formed. Simultaneously, the shortened fatty acid combines with another molecule of CoA at the newly formed carboxyl group to yield a new activated fatty acid. The enzyme is β-ketothidase. The new activated fatty acid undergoes the same series of reactions, and this process is repeated until the entire fatty acid molecule has been reduced completely to acetyl CoA. Oxidation of the higher fatty acids, such as palmitic acid with 16 carbon atoms, yields eight molecules of acetyl CoA. Two fates of acetyl CoA formed in the liver are of immediate interest (**Fig. 4**): (1) oxidation through the Krebs cycle to carbon dioxide and (2) condensation of two molecules of acetyl CoA with the splitting off of the CoA forming so-called ketone bodies. *See* BIOLOGICAL OXIDATION; CITRIC ACID CYCLE.

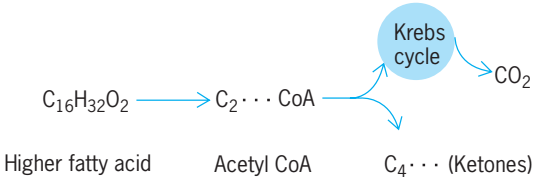

Fig. 4. Fates of acetyl CoA from fatty acid oxidation.

Ketone body formation. If excessive amounts of fatty acids are mobilized from the depots and metabolized by the liver, as occurs in starvation or diabetes, all the acetyl CoA formed cannot be oxidized. Instead, two molecules of acetyl CoA may form acetoacetic acid, which cannot be further oxidized by the liver but is transformed into β-hydroxybutyric acid or acetone according to reactions (6) and (7).

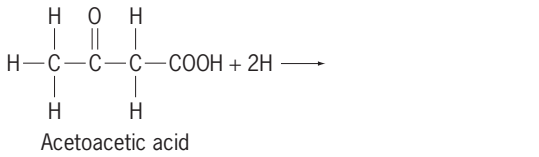

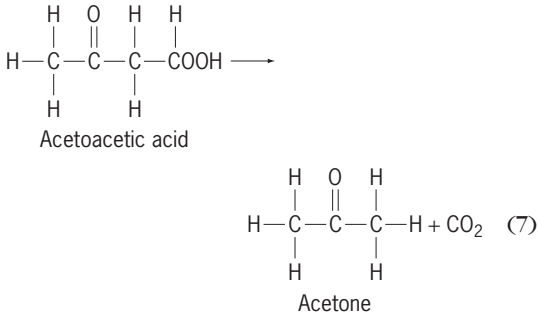

These three compounds, acetoacetic acid, acetone, and β-hydroxybutyric acid, called the ketone bodies, are formed in excessive amounts in the liver not only because of the excess acetyl CoA present but also because the metabolic disturbances caused by such conditions reduce the capacity of the Krebs cycle for oxidation by diverting oxaloacetate, a key intermediate, to gluconeogenesis (synthesis of glucose from noncarbohydrate precursors).

The excessive formation of ketone bodies leads to a high concentration in the blood—a condition called ketonemia. This causes ketone body excretion in the urine, together with water, sodium, and potassium. These losses, along with the excessive acidity of the blood caused by the high concentration of ketone bodies (ketoacidosis), can lead to coma and death if uncorrected. Such fatal acidosis was quite common in diabetics before the advent of insulin. During starvation or diabetic ketosis, muscles are capable of deriving the major portion, if not all, of their energy from oxidizing the ketone bodies. However, in the uncontrolled diabetic subject, the formation of ketone bodies may greatly exceed the utilization of them by the muscle or other tissues, so that ketoacidosis results.

Lipotropic factors. A diet high in fat results in fat accumulation in the liver. This can be prevented by feeding choline, a phosphatide base; the effect has been termed a lipotropic effect. If methionine is present in the diet, it can substitute for choline, since enzymes are present in the liver which can transfer methyl groups from methionine to ethanolamine to form choline, as shown in reaction (8).

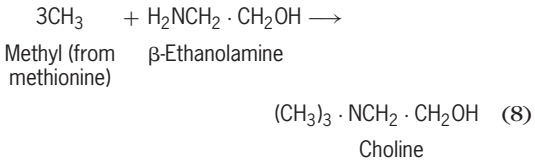

Fatty livers can also occur when large amounts of depot fat are mobilized in starvation and diabetes. This type of fatty liver is not affected by choline. Since choline is a constituent of the phospholipid lecithin, it has been suggested that the lipotropic effect of choline is exerted through lecithin. Evidence to support this suggestion has not been forthcoming.

Fatty acid synthesis. Certain carbohydrates, such as glucose and fructose, yield simpler substances, such as lactic acid, pyruvic acid, and acetic acid, as do proteins, which are potential precursors of acetyl CoA. The acetyl CoA molecules thus formed are used to synthesize higher fatty acids.

The initial reaction in fatty acid synthesis is the carboxylation of acetyl CoA to form malonyl CoA. This reaction is catalyzed by acetyl CoA carboxylase, a biotin-containing enzyme (enz-biotin) and probably proceeds in two steps. In the first, reaction (9*a*), biotin is carboxylated in the 1′-nitrogen position in an ATP-dependent reaction to form 1′-*N*-carboxamide biotin. In the second step, reaction (9*b*), the carboxyl group is transferred to acetyl CoA to form malonyl CoA.

For the synthesis to proceed, the malonyl group must be transferred from CoA to a specific carrier protein called the acyl carrier protein (ACP). The compound with which the malonyl group is to be condensed is similarly transferred from CoA to ACP to form an acyl ACP. The acyl group can be acetate or any fatty acid less than 16 carbons in length, depending upon the species. The enzyme catalyzing the transfer from CoA to ACP is specific for each compound and is called a CoA-ACP transacylase. The active group of both CoA and ACP is 4′-phosphopantotheine (**Fig. 5**). This group is linked to the protein through the hydroxyl group of serine by a phosphodiester bond. The sulfhydryl group is the functional group of both CoA and ACP linkage occurring through sulfur (thiol ester).

The joining of malonyl ACP with an acyl ACP, as shown in reaction (10), is catalyzed by the acylmalonyl ACP condensing enzyme. Since carbon dioxide

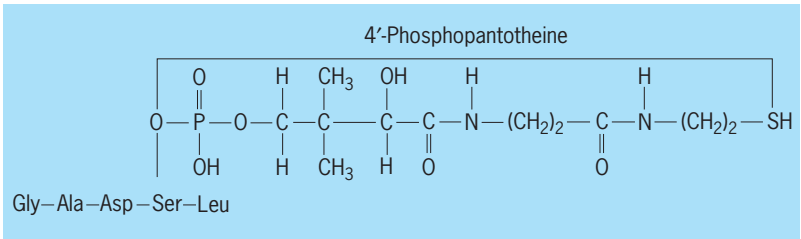

Fig. 5. Active site of acyl carrier protein (ACP). Gly, glycine; Ala, alanine; Asp, aspartic acid; Ser, serine; and Leu, leucine.

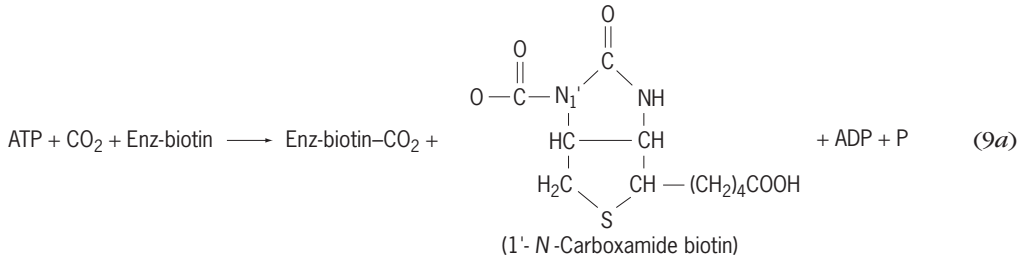

(9a)

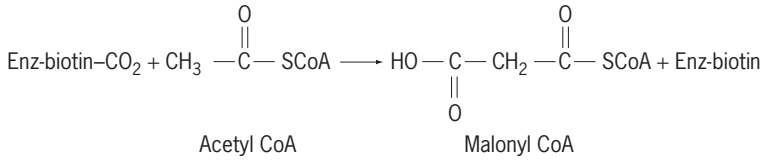

(9b)

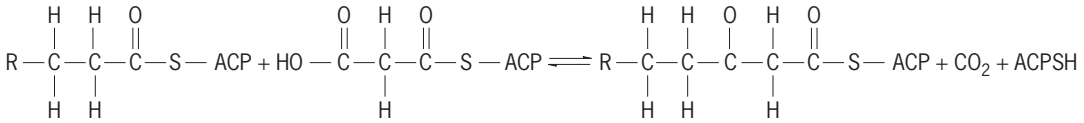

(10)

is released in this reaction, the acyl carbon chain is lengthened by two, resulting in a β-keto acid linked to ACP. This compound is converted to a fatty acid by the reverse of reactions (4), (3), and (2). These reactions are catalyzed by a series of enzymes which are associated in a complex protein called fatty acid syntase in yeast and pigeon liver. The individual enzymes have been isolated from bacterial systems and are specific for ACP compounds, in contrast to the specificity of oxidative enzymes for CoA compounds. The reverse reaction (4) is catalyzed by β-keto ACP reductase and the cofactor is nicotinamide adenine dinucleotide phosphate (NADP), also known as triphosphopyridine nucleotide (TPN), rather than NAD. Reverse reaction (3) is catalyzed by enoyl ACP dehydrase and reverse reaction (2) by enoyl ACP reductase. *See* NICOTINAMIDE ADENINE DINUCLEOTIDE PHOSPHATE (NADP).

Repetition of this series of reactions leads to a building up of the fatty acid molecule to a chain containing 16 or 18 carbons. When the fatty acid reaches this length, it is transferred from ACP to CoA. The CoA is split off by an enzyme, deacylase, yielding the free fatty acid, which in turn combines with glycerol to form phospholipids and the triglyceride that is stored in depots and circulates in lipoproteins.

Triglyceride synthesis. Triglycerides, which consist of three fatty acids joined to glycerol by ester bonds (**Fig. 6**), are synthesized in the liver from glycerol phosphate and CoA derivatives of fatty acids. A series of enzymes called transferases catalyze the joining of the first two fatty acids with glycerol phosphate to form a phosphatidic acid (**Fig. 7**). The phosphate group must be removed to form a 1,2-diglyceride before the third fatty acid can be transferred to form a triglyceride. The removal of phosphate is catalyzed by a specific phosphatase. In the intestinal mucosa, triglycerides can be synthesized from monoglycerides without passing through phosphatidic acid as an intermediate.

Modification of dietary lipids. The fatty acids in the depots of each mammalian species have a characteristic composition. The chain length of ingested fat is either lengthened or shortened by the same pathways used for the synthesis and degradation of fatty acids so that the composition of the depots remains constant. However, when large amounts of a fat of markedly different composition from the animal's depot fat is ingested, the depots slowly change to reflect the characteristics of the ingested fat. The ability to introduce double bonds or reduce them in order to form an unsaturated or a saturated fatty acid is limited; highly unsaturated fatty acids can only be derived from dietary sources.

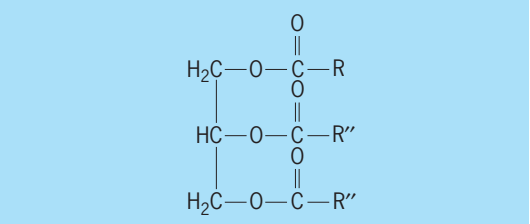

Fig. 6. General structural formula for a triglyceride.

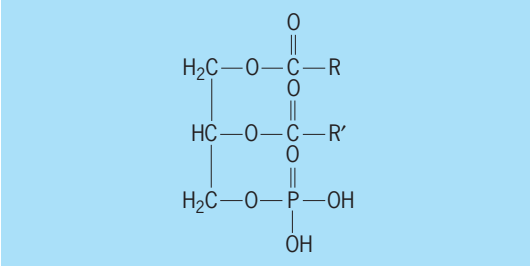

Fig. 7. Structural formula for a phosphatidic acid.

Essential fatty acids. Most mammals, including humans, probably require arachidonic acid for normal growth and for maintenance of the normal condition of the skin, although the mechanism of these effects has not been elucidated. Linoleic acid, a precursor of arachadonic acid in mammals, is the usual dietary source of unsaturated fatty acid. Highly unsaturated fatty acids cannot be synthesized by mammals from either saturated fatty acids or their precursors. Hence, the so-called essential fatty acids must be present in the diet.

Role of phosphatides. The phosphatides are compounds of glycerol, fatty acids, phosphate, and certain bases containing nitrogen, namely, ethanolamine, serine, and choline, the so-called phosphatide bases. When combined with fatty acids, they tend to make the insoluble fatty acids soluble in water. They are important constituents of both outer cell membranes (plasma membranes) and the network of cytoplasmic membranes (endoplasmic reticulum). *See* CELL MEMBRANES.

Phospholipid synthesis. In the liver, phospholipid synthesis follows the same pathway as triglyceride synthesis to the point of diglyceride formation. The nitrogenous base to be coupled at the 3 position of glycerol must first be phosphorylated by ATP and then reacted with cytidine triphosphate to form a cytidine diphosphoryl derivative. This derivative can then be coupled with the diglyceride in a reaction catalyzed by a specific transferase to form a phospholipid. *See* LIPID. Martin A. Rizack

Bibliography. M. J. Halpern, *Lipid Metabolism and Its Pathology*, 1986; M. Horisberger and V. Bracco (eds.), *Lipids in Modern Nutrition*, 1987; J. F. Mead et al., *Lipids, Chemistry, Biochemistry, and Nutrition*, 1986; R. Paoletti et al., *Drugs Affecting Lipid Metabolism*, 1988; R. Salvayre et al. (eds.), *Lipid Storage Disorders: Biological and Medical Aspects*, 1988.

Lipid rafts (membranes)

Microdomains of the cell surface that play key roles in promoting information transfer across the membrane bilayer. The plasma membrane is composed of a lipid bilayer into which are inserted various essential proteins, such as transmembrane ion-channel and transporter proteins (which allow for the movement of small molecules into and out of the cell) and signaling proteins. Many of these proteins are found associated with domains on the surface of the membrane which are rich in free cholesterol (FC) and sphingolipids (lipids containing the amino acid sphingosine or its derivatives); some lipid components of transmembrane signaling pathways are also found here. It seems likely that both leaflets of the bilayer contribute to the structure of the microdomains. It is now evident that these rafts consist of a mixture of two quite different types of FC- and sphingolipid-rich domains: bowl-shaped caveolae and flat lipid rafts. Both caveolae and lipid rafts act as assembly sites (scaffolds) for signaling complexes, and may have important roles in signal transduction and other crucial intracellular functions. *See* CELL MEMBRANES.

Caveolae and lipid raft morphology. The general structures of caveolae and lipid rafts are shown in the **illustration**. Caveolae are invaginated membrane pockets 60–90 nanometers in diameter. They are stabilized by low-molecular-weight structural proteins, called caveolins, which are essential for their formation. In contrast to lipid rafts, caveolae are deficient in GPI-anchored proteins. (GPI stands for glycosylated derivatives of phosphatidylinositol; these are lipid molecules that insert into the membrane lipid bilayer, so that proteins attached to them are held at the membrane surface.) Caveolae and caveolins are abundant in differentiated cells but are relatively deficient (or absent) in cells transformed by cancer. Caveolae are relatively immobile, semipermanent features of the cell surface with a half-life of many hours.

Flat lipid rafts are planar, microdomains rich in FC, sphingolipids, and GPI-anchored proteins, and lacking caveolins. Mean diameters of 10–250 nm have been reported. The rafts are abundant in nonadherent cells, such as lymphocytes, but are probably present to some extent in all vertebrate cells. Lipid rafts are mobile and short-lived, with a half-life estimated by different biophysical techniques to be between a few seconds and several minutes. Microdomains with a composition similar to that of natural lipid rafts are formed spontaneously in synthetic lipid membranes containing FC and sphingolipids. *See* CHOLESTEROL; SPHINGOLIPID.

Presence in cells. There is no evidence that caveolae and lipid rafts are interconvertible, though most cells probably contain both. Depletion of lipid rafts in living cells is associated with an upregulation of the expression of caveolin and caveolae, and vice versa. In caveolin-deficient cells, signaling complexes normally associated with caveolae are found associated with lipid rafts. This may explain why, despite well-defined physiological roles for caveolae, cells from mice in which the caveolin gene has been deleted seem relatively normal. Conversely, in the absence of lipid rafts, signaling proteins that are generally present in the rafts are recruited to caveolae in the course of signal transduction.

Targeted protein assembly. One way in which proteins can be targeted to membrane domains is by covalent binding of long-chain fatty acids (acylation). These increase the association of proteins with membrane lipids. Lipid rafts support the assembly of many cytokine- and chemokine-dependent signaling complexes containing acylated proteins. Caveolae support the assembly of protein complexes that are active in signal transduction and those involved in cell attachment. In addition to acylation, these signaling proteins are bound to caveolin by a motif (local pattern of amino acids) that is rich in aromatic amino acids.

The high FC content of both flat lipid rafts and caveolae appears to be another major factor in targeting proteins to these microdomains. Depleting

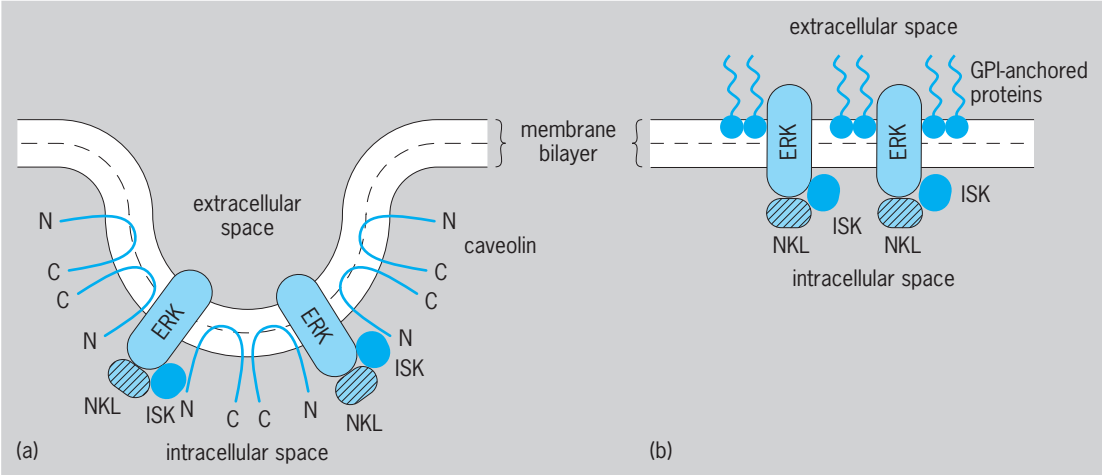

Cholesterol-rich membrane microdomains. (*a*) Caveola. (*b*) Lipid raft. ERK (extracellular receptor kinase), ISK (intracellular signaling kinase), and NKL (nonkinase linker protein) are microdomain-associated proteins that are important in signal transduction and protein complex assembly. N and C refer to the NH_2 and COOH termini of the caveolin protein primary sequence.

membrane FC by using extracellular lipid acceptors such as cyclodextrins (soluble synthetic FC acceptors) favors the transfer of signaling proteins to other parts of the membrane.

Signal transduction. In addition to their activity in assembling of signaling complexes, FC-rich microdomains can play more complex roles. For example, in the case of signaling initiated by the extracellular signaling protein called platelet-derived growth factor (PDGF), which stimulates cell growth and division and is essential to wound healing, FC is required for the initial assembly of the PDGF receptor complex in caveolae, but the FC must be expelled for subsequent steps in the pathway to proceed effectively. Linkage between loss of FC and signal transduction is also shown by the fact that the presence of extracellular FC-binding proteins stimulates both FC efflux and kinase activities in response to PDGF. Far from being inert, FC in caveolae can be dynamically regulated.

Endothelial nitric oxide synthase (e-nos) is an enzyme which regulates production of vascular nitric oxide production, an important mediator of blood vessel contraction. E-nos is localized within a signaling module (a complex of signaling molecules working together) in endothelial cell caveolae. Depletion of FC displaces e-nos from caveolae, rendering it inactive, an effect opposed by the delivery of FC to the caveola.

The signaling kinase Lyn is part of the B-cell antigen receptor (BCR) pathway which mediates the inflammatory response in lymphocytes, where it is mainly localized to lipid rafts. Raft FC mediates Lyn activity by protecting certain Lyn protein phosphoamino acids within rafts from hydrolysis. *See* SIGNAL TRANSDUCTION.

Caveolae, rafts, and cellular FC homeostasis. In addition to activities related to protein assembly and signal transduction, other roles for caveolae and rafts in FC homeostasis have been proposed.

FC efflux. The accessibility of caveolar FC is borne out by its rapid transfer from the cell surface to extracellular lipid acceptors such as plasma lipoproteins, particularly high-density lipoprotein (HDL), and cyclodextrin. The caveolar FC pool has been shown to contribute preferentially to such lipid transfer from caveola-rich primary smooth muscle cells and fibroblasts. This effect was not found in virus-transformed or other continuously dividing cell lines, probably because of the much smaller numbers of caveolae present.

Owing to their short life, a possible contribution of lipid rafts to FC efflux has not been quantified, though FC from rafts, like that from caveolae, is easily transferred to cyclodextrin.

Intracellular FC. An additional role has been suggested for caveolae and caveolin as intracellular mediators of FC transfer between membrane-bound organelles and storage areas. Caveolin has been identified in the perinuclear trans-Golgi fraction, in lipid storage droplets, in a weakly acidic recycling endosome fraction, and in a chaperone complex proposed as a carrier for the newly synthesized FC originating from the endoplasmic reticulum. Many transformed cells transfected with caveolin complementary DNA (cDNA) retain the protein within such intracellular pools.

Lipid rafts and viral infectivity. A recent development is the recognition that signaling proteins in lipid rafts are the binding sites by which many viruses gain entry to cells. The same structures can contribute coat lipids to newly synthesized virions. Lymphocyte raft signaling proteins CD4 and CCR5 are among the contributors to human immunodeficiency virus (HIV-1) binding. Depletion of FC in these microdomains via cyclodextrin significantly reduces infectivity in vitro. Other viruses shown to utilize lipid raft pathways for infection include Echovirus and Ebola virus. Details of the linkage between virus binding and internalization have not yet

been worked out. Preliminary data suggest that caveolae may contribute to the internalization of other microorganisms.

Answers to the many questions remaining about caveolae, lipid rafts, and their functions may be gained from their tertiary structure, when this information becomes available. Present data suggest that structural differences might contribute significantly to heterogeneity within plasma membranes, generating the unique lipid and protein structures that play important roles in cellular transport and communication. *See* VIRUS. Christopher J. Fielding

Bibliography. C. J. Fielding (ed.), *Lipid Rafts and Caveolae: From Membrane Biophysics to Cell Biology*, Wiley-VCH, Weinheim, Germany, 2006; R. G. Parton, Caveolae—from ultrastructure to molecular mechanisms, *Nat. Rev. Mol. Cell Biol.*, 4:162–167, 2003; L. J. Pike, Lipid rafts: Bringing order to chaos, *J. Lipid Res.*, 44:655–667, 2003.

Lipoprotein

Classes of conjugated proteins consisting of a protein combined with a lipid. The normal functioning of higher organisms requires movement of insoluble lipids between tissues. These lipids include cholesterol, used for cell membranes, steroid hormones, and bile; and triglyceride, a major form of dietary and stored energy. To accomplish this movement, lipids are incorporated into macromolecular complexes called lipoproteins.

Structure and metabolism. All major types of lipoproteins share a general structure (see **illus.**). The core of these spherical particles contains primarily cholesteryl ester and triglyceride. These insoluble molecules are surrounded by a coating of proteins and phospholipids that are amphipathic; that is, they have both polar and nonpolar regions. Lipoproteins vary by size and density. The largest lipoproteins, chylomicrons, are up to 500 nanometers in diameter, and since they contain primarily triglyceride they are so buoyant that they float in plasma. Very low density lipoproteins (VLDL) also primarily transport triglyceride; low-density lipoproteins (LDL) and the smallest, most dense lipoproteins, high-density lipoproteins (HDL), transport cholesterol. The interactions of these particles with cell surface receptors and with metabolic enzymes are mediated by the protein components of the particles that are termed apolipoproteins. *See* CHOLESTEROL; TRIGLYCERIDE.

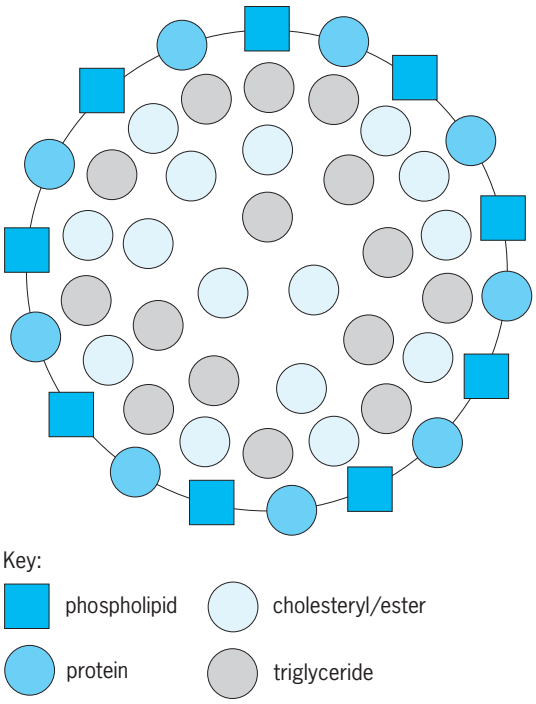

Structural model of a lipoprotein. PL, phospholipid; P, protein; CE, cholesteryl ester; TG, triglyceride.

Chylomicrons and VLDL. These two lipoproteins transport triglyceride. Chylomicrons contain triglyceride (fat) from the diet. In addition, they are the lipoproteins that carry fat-soluble vitamins, such as vitamin A and E, into the circulation. Chylomicrons are produced in the intestine, enter the body via the lymphatic system, and then enter the bloodstream. Very low density lipoproteins are made in the liver and contain triglyceride that is synthesized either from excess carbohydrate sources of calories or from fatty acids that enter the liver and are reassembled into triglyceride. Lipoprotein lipase (LpL) is an enzyme found on the surface of blood vessels that is responsible for the breakdown of triglyceride in lipoproteins. This enzyme is made in muscle and fat. It requires an activator protein, apoCII, that is carried on very low density lipoproteins and chylomicrons. Through this system, energy is supplied to muscle for motion and to fat for storage.

The partially degraded lipoproteins are termed remnants. They are ultimately removed from the circulation by the liver. Two receptors on liver cells, the LDL receptor and the LDL receptor–related protein, recognize apoE on the surface of the remnant lipoproteins.

LDL and HDL. Low-density lipoproteins constitute a metabolic product that results after triglyceride is removed from very low density lipoproteins. This leaves a smaller, denser particle that primarily contains cholesteryl ester as its core lipid and a single protein called apoB. Cells throughout the body contain an LDL receptor that recognizes apoB. This allows the uptake of low-density lipoproteins into cells, supplying them with cholesterol. When sufficient low-density lipoproteins and cholesterol are available, cells use them in preference to synthesizing new cholesterol from precursors.

In contrast, high-density lipoproteins both deliver and remove cholesterol from tissues. In animals that have very little circulating low-density lipoproteins, such as rodents, high-density lipoproteins deliver cholesterol to steroid-producing tissues. Its major protein, apoAI, is made in both the intestine and the liver and is secreted with a small amount of lipid. These small HDL precursors acquire additional cholesterol from other tissues. Via the actions of lecithin cholesterol acyl transferase (LCAT),

the cholesterol is converted to cholesteryl ester—a process that makes it more hydrophobic and allows it to create a lipoprotein with a cholesteryl ester core. High-density lipoproteins have at least one receptor, scavenger receptor B-1 (SRB-1). High-density lipoproteins bind to SRB-1 in the liver and in adrenal cells; and via a poorly understood process, cholesteryl ester, but not HDL protein, is transferred to cells. This process can supply cholesterol for steroid hormones and for bile. A plasma lipid transport protein called cholesteryl ester transfer protein (CETP) exchanges cholesteryl ester in high-density lipoproteins for triglyceride in very low density lipoproteins, thus reducing the size and cholesterol content of high-density lipoproteins. CETP is not produced by rodents.

Lipoprotein(a). Apo(a) is a large protein with homology to the blood-clotting protein plasminogen. Apo(a) complexes with low-density lipoproteins via a disulfide linkage and forms a lipoprotein called Lp(a). Lp(a) contains added protein and is denser than low-density lipoproteins.

Medical significance. Blood levels of lipoproteins are major factors regulating risk for development of coronary artery atherosclerosis, the most common form of heart disease in the western world. Via unknown mechanisms, low-density lipoproteins and remnant lipoproteins infiltrate and then become attached to extracellular matrix molecules within the artery. Some of the lipoproteins are internalized by macrophages and smooth muscle cells. This might first require chemical modification such as oxidation of the lipids. The resulting pathological findings are deposition of cholesterol in cells and matrix within the vessel wall, leading to a decrease in the diameter of the artery.

LDL cholesterol is a major predictor of the development of atherosclerosis in humans and in animals. Moreover, dietary and genetic modifications of animals have conclusively shown that elevated cholesterol in low-density lipoproteins or remnant lipoproteins is the only alteration required to convert a nonatherogenic animal into one that gets this disease.

In contrast, high-density lipoproteins appear to prevent atherosclerosis formation. The reasons are not entirely understood. Most likely, high-density lipoproteins remove excess cholesterol that accumulates in the artery, or high-density lipoproteins prevent the oxidation of low-density lipoproteins.

Lp(a) levels are also correlated with the development of atherosclerosis. This relationship is most evident in Caucasian populations. Blood Lp(a) concentrations are primarily genetically determined; individuals with smaller forms of apo(a) produce a greater number of these molecules. Lp(a) particles may alter blood clotting or have alternative molecular interactions that make them more likely to accumulate in arteries. *See* ARTERIOSCLEROSIS.

Regulation of levels. Diet, exercise, and genetic factors control the levels of lipoproteins in the blood. Diets containing more cholesterol and saturated fats increase the amounts of LDL cholesterol. Lipoproteins containing primarily triglyceride are increased with obesity, diabetes, and intake of foods such as simple sugars and alcohol that are readily converted to triglyceride in the liver. Exercise lowers triglyceride.

HDL levels are increased by high cholesterol–high saturated fat diets, but low-density lipoproteins are increased at the same time. In addition, alcohol, oral estrogen, and exercise raise high-density lipoproteins.

Lp(a) is reduced by estrogen.

Disorders. In most people the concentrations of blood lipoproteins are modulated by lifestyle and by poorly defined polygenetic factors. ApoB is produced in greater amounts in people with familial combined hyperlipoproteinemia. These people have increased very low density lipoproteins and low-density lipoproteins in their blood. Familial hypercholesterolemia is a disease due to a defect in removal of low-density lipoproteins from the bloodstream. This is either due to a defect in the LDL receptor or to a mutation in apoB that prevents low-density lipoproteins from being recognized by the receptor. Defects in lipoprotein lipase or its activator apoCII lead to very high blood levels of triglyceride. Defects in almost all the other apos and lipoprotein-metabolizing enzymes have been described and lead to diseases in humans.

Treatments. A number of medical therapies have been developed to reduce blood concentrations of lipoproteins. Medications that reduce LDL cholesterol are effective in reducing atherosclerosis complications such as heart attacks and strokes. Bile acid–binding resins reduce low-density lipoproteins by absorbing cholesterol in the gastrointestinal tract and augmenting its removal. The vitamin niacin decreases liver production of very low density lipoproteins and low-density lipoproteins. HMG CoA reductase inhibitors, also known as statins, are inhibitors of the cholesterol synthesis in the liver. In turn, the liver increases the number of LDL receptors, leading to greater removal of low-density lipoproteins from the bloodstream. Ira J. Goldberg

Bibliography. H. N. Ginsberg and I. J. Goldberg, Lipoproteins disorders, in *Harrison's Principles of Internal Medicine,* 13th ed., edited by Fauci et al., pp. 2138–2149, 1997; M. Rosseneu (ed.), *Structure and Function of Apolipoproteins*, CRC Press, Boca Raton, FL, 1992.

Liposomes

Aqueous compartments enclosed by lipid bilayer membranes; liposomes are also known as lipid vesicles. Phospholipid molecules consist of an elongated nonpolar (hydrophobic) structure with a polar (hydrophilic) structure at one end. When dispersed in water, they spontaneously form bilayer membranes, also called lamellae, which are composed of two monolayer layer sheets of lipid molecules with their nonpolar (hydrophobic) surfaces facing each other and their polar (hydrophilic) surfaces facing the

aqueous medium. The membranes enclose a portion of the aqueous phase much like the cell membrane which encloses the cell; in fact, the bilayer membrane is essentially a cell membrane without its protein components.

Liposomes are often used to study the characteristics of the lipid bilayer. Properties of liposomes have been characterized by a variety of techniques: molecular organization by x-ray diffraction, nuclear magnetic resonance, electron paramagnetic resonance, and Raman spectroscopy; melting behavior (that is, crystal to liquid-crystal transition) by calorimetry; net electric surface charge by microelectrophoresis; size by light scattering and electron microscopy. *See* LIPID.

Liposomes are most easily formed by hand-shaking a solution of anhydrous lipids with water for seconds to hours (depending on the natures of the lipid or aqueous phases). This simple procedure spontaneously yields large, multilamellar liposomes or vesicles with diameters of 1–10 micrometers, which are composed of a few to hundreds of concentric lipid bilayers alternating with layers of aqueous phase. The thickness of the aqueous layers and, hence, the total amount of aqueous phase trapped by the vesicle, depends on the balance of electrostatic repulsion forces between charged lipids and van der Waals attractive forces between bilayers as a whole. Thus, aqueous layer spacing (and hence, volume trapped) increases with increasing proportion of charged lipids in the membrane and with decreasing concentrations of electrolyte (charged ions) in the aqueous phase. The smallest unilamellar vesicles of 20-nanometer diameter can be produced by subjecting multilamellar vesicles to ultrasound. Unilamellar vesicles of somewhat larger (0.1–1 μm) diameter result when lipid is solubilized in organic solvent or in detergent, and the solubilizing agent removed by evaporation or by dialysis, respectively. Fusion of small unilamellar vesicles by methods requiring particular lipids or stringent dehydration-hydration conditions can yield unilamellar vesicles as large as or larger than cells.

Although hydration of anhydrous phospholipids was long known to generate sheets of lipid bilayer, it was not established until the 1960s that the aqueous volume trapped by the hydrated lipid is totally enclosed. Since the vesicles are continent and can be loaded with almost any water-soluble molecule, it was possible to measure bilayer permeability to ions and metabolites present in cells. It was soon demonstrated that liposome membranes have low permeability to polar metabolites; this is also an essential material property of the similarly constructed matrix of the cell membrane, which means that liposomes can be used to reproduce various cell membrane functions. For example, the criterion for the isolation of membrane transport proteins is usually their incorporation into liposomes in a functional form.

Liposomes have numerous uses as biochemical and biophysical tools: (1) as vehicles for the delivery of both water- and oil-soluble materials to the cell, (2) as immunological adjuvants, (3) as substrates for the study of membrane properties such as rotational or translational diffusion in the plane of the membrane, and (4) as intermediates in the construction of bilayers large enough for the study of electrical properties of membranes. The great interest in liposomes as therapeutic agents was stimulated by the administration to animals of liposomes containing anticancer drugs, which suggested that such an innovation could decrease a person's exposure to toxic compounds while delivering a high dose of the compound to the diseased site. The targeting of drug-containing vesicles to specific sites by modification of the liposome surface—for example, by attaching site-specific antibodies—has met with limited success, however. A notable exception has been the treatment of parasitic diseases of the liver, in which most unmodified liposomes naturally accumulate, since most of the scavenging of particulate matter occurs in the liver. *See* LIVER.

Other medical applications that have been studied include the enhancement of x-ray imaging by radiopaque liposomes; heavy-metal toxicity therapy with liposomes containing chelating agents; protection of otherwise digestible dietary supplements and drugs to enable their oral administration; and preparation of artificial blood composed of hemoglobin-loaded liposomes. Technological applications of liposomes have also been the subject of study. Notable among these are (1) diagnostic assays featuring liposomes constructed to signal formation of antibody-antigen complexes; and (2) artificial photosynthesis by appropriately constructed liposomes, in which a light-driven reaction leads to the release of hydrogen from water. *See* CELL MEMBRANES.

Robert C. MacDonald; Ruby I. MacDonald

Bibliography. A. D. Bangham (ed.), *The Liposome Letters*, 1983; G. Gregoriadis (ed.), *Liposome Technology*, vols. 1–3, 1984; E. D. Korn (ed.), *Methods in Membrane Biology*, vol. 1, 1974.

Liquefaction of gases

The process of refrigerating a gas to a temperature below its critical temperature so that liquid can be formed at some suitable pressure, also below the critical pressure.

Thus gas liquefaction is a special case of gas refrigeration and cannot be separated from it. In both cases, the gas is first compressed to an elevated pressure in an ambient-temperature compressor. This high-pressure gas is passed through a countercurrent heat exchanger to a throttling valve or expansion engine. Upon expanding to the lower pressure, cooling may take place, and some liquid may be formed. The cool, low-pressure gas returns to the compressor inlet to repeat the cycle. The purpose of the countercurrent heat exchanger is to warm the low-pressure gas prior to recompression, and simultaneously to cool the high-pressure gas to the lowest temperature possible prior to expansion. Both refrigerators and liquefiers operate on this same basic principle.

See COMPRESSOR; CRITICAL PHENOMENA; GAS; HEAT EXCHANGER; REFRIGERATION.

There is nonetheless an important distinction between refrigerators and liquefiers: In a continuous refrigeration process, there is no accumulation of refrigerant in any part of the system. This contrasts with a gas-liquefying system, where liquid accumulates and is withdrawn. Thus, in a liquefying system, the total mass of gas that is warmed in the countercurrent heat exchanger is less than the gas to be cooled by the amount that is liquefied, creating an unbalanced flow in the heat exchanger. In a refrigerator, the warm and cool gas flows are equal in the heat exchanger. This results in balanced flow condition. The thermodynamic principles of refrigeration and liquefaction are identical. However, the analysis and design of the two systems are quite different due to the condition of balanced flow in the refrigerator, and unbalanced flow in liquefier systems.

Liquefaction principles. The prerequisite refrigeration for gas liquefaction is accomplished in a thermodynamic process when the process gas absorbs heat at temperatures below that of the environment. As mentioned above, a process for producing refrigeration at liquefied gas temperatures usually involves equipment at ambient temperature in which the gas is compressed and heat is rejected to a coolant. During the ambient-temperature compression process, the enthalpy and entropy, but usually not the temperature of the gas, are decreased. The reduction in temperature of the gas is usually accomplished by heat exchange between the cooling and warming gas streams followed by an expansion of the high-pressure stream. This expansion may take place either through a throttling device (isenthalpic expansion) where there is a reduction in temperature only (when the Joule-Thomson coefficient is positive) or in a work-producing device (isentropic expansion) where both temperature and enthalpy are decreased. *See* ENTHALPY; ENTROPY; ISENTROPIC PROCESS; THERMODYNAMIC PRINCIPLES; THERMODYNAMIC PROCESSES.

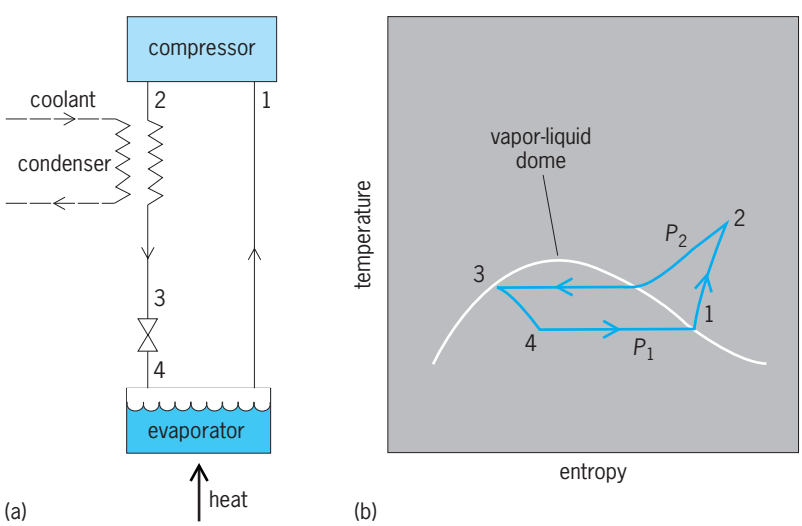

Fig. 1. Compressed-vapor refrigerator. (*a*) Arrangement of equipment. (*b*) Temperature-entropy diagram of process. Numbers indicate corresponding points in the cycle.

Maximum inversion temperatures of frequently liquefied gases*

Fluid	Maximum inversion temperature, °R (K)
Oxygen	1370 (761)
Argon	1300 (722)
Nitrogen	1120 (622)
Air	1085 (603)
Neon	450 (250)
Hydrogen	364 (202)
Helium	72 (40)

*After R. Barron, *Cryogenic Systems*, 1966.

For reasons of efficiency alone, isentropic expansion devices might be preferred for gas liquefaction. However, expansion engines have yet to be developed which can operate with more than a few percent of liquid present, due to erosion. Therefore, simple throttling devices are employed at some point in nearly all gas-liquefaction systems. Throttling is an isenthalpic process, and the slope at any point on an isenthalpic curve (constant enthalpy h) in a pressure-temperature diagram is a quantitative measure of how the temperature T will change with pressure P. This slope is called the Joule-Thomson coefficient and is usually denoted by μ given mathematically by Eq. (1).

$$\mu = \left(\frac{\partial T}{\partial P}\right)_h \tag{1}$$

The Joule-Thomson coefficient is a property of each specific gas, is a function of temperature and pressure, and may be positive, negative, or zero. For instance, hydrogen, helium, and neon have negative Joule-Thomson coefficients at ambient temperature. Consequently, to be used as refrigerants in a throttling process, they must first be cooled either by a separate precoolant fluid or by a work-producing expansion engine to a temperature (called inversion temperature) below that at which the Joule-Thomson coefficient is positive. Only then will throttling cause a further cooling rather than a heating of these gases.

Nitrogen, methane, and other gases, on the other hand, have positive Joule-Thomson coefficients at ambient temperatures and hence produce cooling when expanded through a valve. Accordingly, they may be liquefied directly in a throttling process without the necessity of a precooling step or expansion through a work-extracting device. Maximum inversion temperatures for some of the more frequently liquefied gases are given in the **table**.

Liquefaction methods. General principles of the more useful liquefaction methods are given below.

Vapor compression. In a typical single-stage compressed-vapor refrigerator (**Fig. 1***a*), heat is absorbed by evaporation of a suitable liquid at temperature T_1 and pressure P_1 in an evaporator. In a dry

compression cycle, saturated vapor is compressed to pressure P_2, and superheated vapor at the compressor outlet is condensed in a condenser at temperature T_2 by exchanging heat with a suitable coolant. A throttling valve allows the saturated liquid formed in the condenser to expand approximately isenthalpically from 3 to 4 to the evaporator, completing the cycle. Figure 1*b* shows the states of the refrigerant in this compressed-vapor process on a temperature-entropy diagram.

Certain criteria can be established to aid in choosing a gas for this type of liquefier.

1. The temperature of the evaporator is usually near the normal boiling point of the gas so that the pressure of the evaporator may be approximately atmospheric. Temperatures lower than the normal boiling temperature can be used; however, care must be taken to prevent air leaking into the subatmospheric parts of the cycle.

2. The gas must be chosen with a small compression ratio. This is necessary because the work of compression increases rapidly with the compression ratio. High discharge pressures require massive equipment for proper containment of the refrigerant and result in costly and heavy refrigerators.

3. The ambient temperature must be less than the critical temperature of the gas to effect condensation while using the environment as a coolant.

The efficiency of the compressed-vapor refrigerator is generally higher than that of other methods discussed below because heat is transferred nearly isothermally, and the compression process can be made nearly reversible. Thus, the main deviation of this cycle from a Carnot cycle is the isenthalpic expansion at the throttling valve.

Cascade vapor compression. It is possible to cascade several vapor-compression systems, each containing different refrigerants, to reach lower temperatures (**Fig. 2**). Cascading is necessary because the gases with the lower boiling points have critical temperatures well below the ambient temperature. Thus, it is necessary to provide condenser temperatures for these vapor-compression cycles by the evaporators of other vapor-compression cycles, which reject heat at higher temperatures. The efficiency of cascade systems is greatly improved by the use of heat exchangers which cool the incoming warm gases. The exchangers warm the cool gases before they reach the compressor intake.

Isenthalpic expansion. In a thermodynamic process to obtain cryogenic temperatures, known as the simple Linde cycle (**Fig. 3**), the gas is compressed at ambient temperature approximately isothermally from 1 to 2, rejecting heat to a coolant. The compressed gas is cooled in a heat exchanger by the stream returning to the compressor intake until it reaches the throttling valve. Joule-Thomson cooling upon expansion from 3 to 4 further reduces the temperature until, in the steady state, a portion of the gas is liquefied.

For the simple Linde liquefier shown, the liquefied portion is continuously withdrawn from the reservoir and only the unliquefied portion of the fluid is

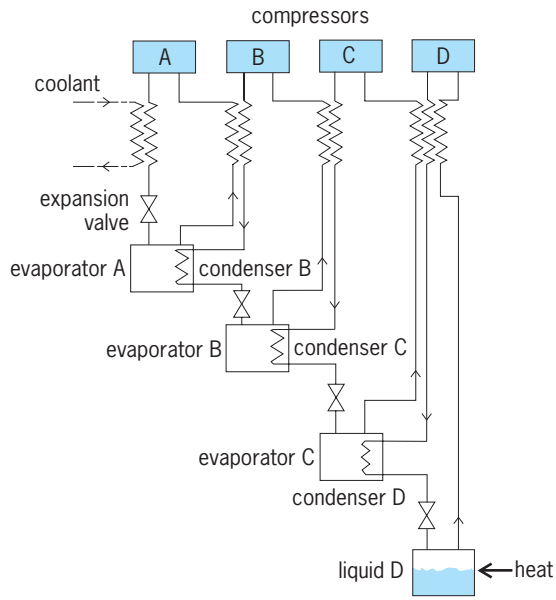

Fig. 2. Cascade compressed-vapor refrigerator.

warmed in the countercurrent heat exchanger and returned to the compressor. The fraction y that is liquefied is obtained by applying the first law of thermodynamics to the heat exchanger, Joule-Thomson valve, and liquid reservoir. This results in Eq. (2),

$$y = \frac{h_1 - h_2}{h_1 - h_f} \tag{2}$$

where h_f is the specific enthalpy of the liquid being withdrawn. Maximum liquefaction occurs when h_1 and h_2 refer to the same temperature.

Gases used in this process have a critical temperature well below the ambient temperature; consequently, liquefaction by direct compression is not possible. In addition, the inversion temperature must be above the ambient temperature to provide

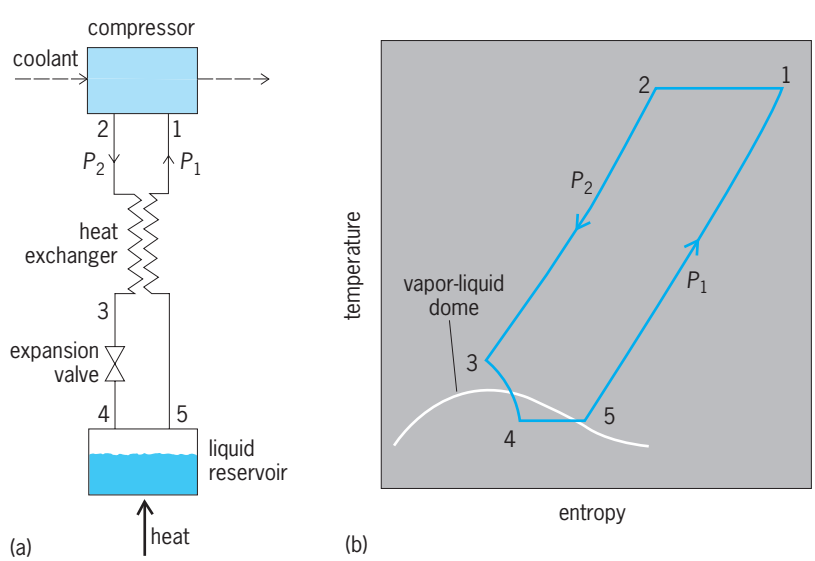

Fig. 3. Refrigerator using simple Linde cycle. (*a*) Arrangement of equipment. (*b*) Temperature-entropy diagram of process.

cooling as the process is started. Auxiliary refrigeration is required if the simple Linde cycle is to be used to liquefy gases whose inversion temperature is below the ambient temperature. Liquid nitrogen is the optimum refrigerant for hydrogen and neon liquefaction systems, while liquid hydrogen is the normal refrigerant for helium liquefaction systems.

Dual-pressure process. To reduce the work of compression, a two-stage or dual-pressure process may be used whereby the pressure is reduced by two successive isenthalpic expansions. Part of the gas is returned to the compressor after the first expansion, providing some precooling for the full high-pressure stream. The remainder of the gas proceeds to the second isenthalpic expansion. Since the work of compression is approximately proportional to the logarithm of the pressure ratio, and the Joule-Thomson cooling is roughly proportional to the total pressure difference, there is a much greater reduction in compressor work than in refrigerating performance. Hence, the dual-pressure process produces a given amount of liquefaction with less energy input than the simple Linde process.

Isentropic expansion. Refrigeration can always be produced by expanding the process gas in an engine and causing it to do work. In a simple refrigerator using this principle (**Fig. 4**), gas compressed isothermally from 1 to 2 at ambient temperature is cooled in a heat exchanger by gas being warmed on its way to the compressor intake. Further cooling takes place during the engine expansion from 3 to 4. In practice, this expansion is never truly isentropic, and this is reflected by the curved path 3–4 in Fig. 4*b*. The net cooling effect is shown between T_4 and T_5 in Fig. 4*b*.

In a work-producing expansion, the temperature of the gas is always reduced; hence, cooling does not depend on being below the inversion temperature prior to expansion. In large machines the work produced during expansion is conserved. In small liquefiers, the energy from the expansion is usually expended in a gas or hydraulic pump, or other suitable device.

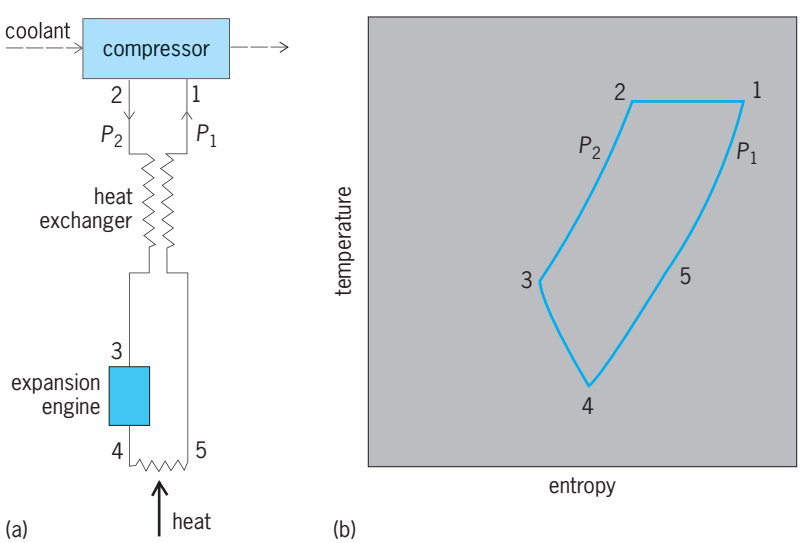

Fig. 4. Isentropic-expansion refrigerator. (*a*) Arrangement of equipment. (*b*) Temperature-entropy diagram of process.

Path 3–4 seldom dips very far into the liquid-vapor region because of technical problems associated with forming liquid in the expansion engine. Since a liquid product is the objective, an isentropic expansion for precooling is combined with isenthalpic expansion for forming liquid, as discussed below.

Combination of isenthalpic and isentropic expansion. It is not uncommon to combine the isentropic and isenthalpic expansions to allow the formation of liquid in the refrigerator. The Claude cycle is an example of a combination of these methods. All of the high-pressure gas is first expanded isentropically to produce cooling. A portion of this cooled gas is bypassed to the low-pressure stream to help precool the incoming high-pressure stream. The remainder of the cooled, high-pressure stream is expanded isenthalpically to produce liquid.

Modifications of Claude cycle. One modification of the Claude cycle that has been used extensively in high-pressure liquefaction plants for air is the Heylandt cycle. In this cycle, the first warm heat exchanger is eliminated, allowing the inlet of the expander to operate at ambient temperatures and thereby minimizing lubrication problems that are often encountered at lower temperatures.

Another modification of the basic Claude cycle is the dual-pressure Claude cycle, similar in principle to the dual-pressure Linde system discussed above. In the dual-pressure Claude cycle, only the gas which is sent through the expansion valve is compressed to the high pressure; this reduces the work requirement per unit mass of gas liquefied. When the Claude dual-pressure cycle is selected in the liquefaction of air, the liquid yield can be doubled while the work per unit mass liquefied can be halved in comparison to the dual-pressure Linde cycle.

Still another extension of the Claude cycle is the Collins helium liquefier. Depending upon the helium inlet pressure, from two to five expansion engines are used in this system. The addition of a liquid-nitrogen precooling bath to this system results in a two- to threefold increase in liquefaction performance. *See* LIQUID HELIUM.

Mixed refrigerant cycle. The mixed refrigerant cycle is a variation of the cascade cycle described above and involves the circulation of a single refrigerant stream which might be a mixture of refrigerants. The simplification of the compression and heat exchange services in such a cycle may, under certain circumstances, offer potential for reduced capital expenditure over the conventional cascade cycle. This cycle is widely used to liquefy natural gas. *See* AIR SEPARATION; CRYOGENICS; LOW-TEMPERATURE PHYSICS.

Thomas M. Flynn

Bibliography. *Advances in Cryogenic Engineering*, vols. 1–40, 1954–1994; R. Barron, *Cryogenic Systems*, 2d ed., 1985; T. M. Flynn and K. D. Timmerhaus, *Cryogenic Process Engineering*, 1989.

Liquefied natural gas (LNG)

Natural gas, usually consisting primarily of methane, that has been processed into liquid form. It is increasingly used for transporting large quantities of natural gas across oceans with specially designed ships. The properties of LNG are those of liquid methane, slightly modified by minor constituents. Natural gas must be cooled to cryogenic temperatures to be liquefied and then stored in well-insulated containers. More than 600 ft^3 (17 m^3) of natural gas condense into 1 ft^3 (0.028 m^3) of liquid at about −258°F (−161°C). In this form, the gas can be conveniently stored and transported. At the receiving end, LNG is revaporized and then compressed and put into the natural gas distribution system. *See* CRYOGENICS; NATURAL GAS.

Contribution to gas supply. By 2005, natural gas became the world's number two source of energy supply (behind oil but surpassing coal) with 23% of the total. Estimates vary but point toward an increasing market share for natural gas over the next several decades. While world energy demand is slated to increase by at least 50%, natural gas will provide at least 28% and as much as 45%.

For the United States, estimates suggest that there will be a need to import 8 trillion cubic feet (Tcf; 2.2×10^{11} m^3) by 2015 and 15 Tcf (4.2×10^{11} m^3) by 2015.

The earliest commercial use for LNG was to supply winter peak demands for home heating fuel. Such a peak management system was constructed in Cleveland, Ohio, in the early 1940s. However, several years of successful operation ended in disaster when a storage tank developed a leak and escaping liquid spread through city streets, producing a very destructive fire. Nearly two decades passed before another LNG plant was built in the United States.

In modern design, dikes separate storage tanks and equipment to ensure that liquid leaks will be confined. An accident or fire in one unit would not be likely to damage any other area of the plant or its surroundings.

Insulated storage tanks constitute the main cost element in LNG plants. Novel approaches have attempted to reduce this cost. Frozen earth can form the lining of an excavated reservoir. As tank size increases, the ratio of wall surface area to container volume declines. With large tanks, the heat conduction rate of natural soil results in tolerable evaporation. However, water flowing through the earth in contact with the frozen soil lining increases heat flow to intolerable levels.

Modern LNG schemes involve four distinct parts: supply of natural gas (likely from a country in South America, Africa, or the Middle East), liquefaction, transportation, and regasification. The relative costs for the four phases, respectively, for 1 Bcf/day (2.86×10^7 m^3/d) are \$1.5 billion, \$2 billion, \$1–2 billion (depending on the distance), and \$0.5 billion. These figures lead to a price per Mcf of \$2–2.45 from South America, \$2.25–3.25 from Africa, and \$2.85–3.60 from the Middle East. Such prices make LNG affordable.

Specialized tanker ships can transport LNG over long distances. Modern LNG tankers transport about 3 Bcf of gas. Storage and regasification facilities are required to reconvert to a gas that can then be fed into the gas distribution system. Worldwide, there are about 17 natural gas liquefaction plants and 40 regasification plants in operation today, and several new liquefaction and regasification plants are under construction. Regasification facilities in the United States are expected to increase from 2 Bcf/d capacity in 2004 to at least 12 Bcf/d by 2010. *See* OIL AND GAS STORAGE.

Use. In 2004, about 27% of the global natural gas trading was done through LNG technology. This constitutes roughly 7% of total world production. Most of the current LNG transactions occur from Asia to Japan. Early on, the Japanese government realized the immense potential of natural gas in their energy mix and began construction of LNG facilities in the 1970s. Today Japan imports 47% of the world's LNG production.

Natural gas demand is fast increasing worldwide. As the United States increases its natural gas import capacity, Canada, its main outside source, will no longer be able to satisfy demand. Thus, LNG terminals are an increasingly viable alternative.

Ocean transport. Gas-producing areas remote from demand centers can ship LNG overseas—Japan is supplied by plants in Abu Dhabi, Alaska, Borneo, Brunei, Indonesia, and the Middle East. Western Europe and the United States receive LNG from the Middle East, Algeria, and Libya. Many other countries, such as Russia, Trinidad, and Australia, have embarked upon ambitious LNG projects. It is likely that China will soon become a major market for LNG imports.

In early 2005, the Atlantic Basin had 14.5 Bcf/d proposed and committed liquefaction plants for shipment to consumers. The numbers for the Middle East and Asia-Pacific were 12.5 Bcf/d and 14.2 Bcf/d, respectively. These projects would increase the total installed capacity in the world from 17 Bcf/d in 2003 to 30 Bcf/d by 2008 and 41 Bcf/d by 2010. Moreover, the world fleet for LNG tankers increased from 90 in 1995, to 128 in 2001, to 174 in 2004. By 2009, this number is estimated to exceed 300.

LNG ships are equipped with specially designed insulated storage tanks of three general types. Balsa wood has been used extensively as an effective insulation. Membrane, or integral, tanks contain LNG within a thin metallic liner, which is supported by load-bearing insulation, which in turn receives its support from the ship's structure. Stand-alone prismatic tanks conform to the hull contour but are self-supporting. Both membrane and prismatic tanks require special design at tank edges and corners to accommodate contraction during cool-down to LNG temperature. Spherical stand-alone cargo tanks minimize thermal problems. However, with spherical tanks the hull volume is poorly utilized. Ship design is influenced by the height of tanks above the

water line. Both LNG and its insulating material have densities significantly less than seawater. Therefore, all LNG vessels carry most of their cargo above the waterline. This design necessity is complemented by spherical tanks. Stability and wind load factors become more important in design and operation. *See* MARINE REFRIGERATION; MERCHANT SHIP.

Challenges. Normally, the composition of LNG is over 97% methane. Heat leak into storage tanks may vaporize essentially pure methane, causing enrichment of the remaining liquid in ethane and propane. At the same time, density increases and boiling temperature rises. Composition, density, and temperature of liquid added to the tank may be significantly different. Mixing large quantities of liquids of differing composition may cause very rapid vaporization. If the rate exceeds the capacity of relief devices, tank facilities may be damaged. *See* ETHANE; METHANE; PROPANE.

Large-scale spills of LNG present unique hazards. As very cold methane vapor mixes with air, the density of the mixture increases. Absorption of heat from contact with surfaces decreases mixture density. As a result, a cloud of cold gas may form and hover near the ground. Since dispersion is slow in relatively still air, combustible mixtures may travel considerable distances before achieving safe dilution.

Michael J. Economides; Arthur W. Francis

Bibliography. M. J. Economides, K. Sun, and G. Subero, *Compressed Natural Gas (CNG): An Alternative to Liquid Natural Gas (LNG)*, Pap. SPE 92047, 2005; C. H. Gatton, *Liquefied Natural Gas Technology and Economics*, 1967; LNG scorecard, *Pipeline Gas J.*, November 1982; National Fire Protection Association, *Liquefied Natural Gas: Production, Storage and Handling*, 1972; G. Subero et al., *A Comparative Study of Sea-Going Natural Gas Transport*, Pap. SPE 90243, 2004; U.S. Department of Energy, *An Approach to LNG Safety*, DOE/EV 002, February 1978.

Liquefied petroleum gas (LPG)

A product of petroleum gases, principally propane and butane, which must be stored under pressure to keep it in a liquid state. At atmospheric pressure and above freezing temperature, these substances would be gases. Large quantities of propane and butane are now available from the gas and petroleum industries. These are often employed as fuel for tractors, trucks, and buses and mainly as a domestic fuel in remote areas. Because of the low boiling point (−47 to 32°F or −44 to 0°C) and high vapor pressure of these gases, their handling as liquids in pressure cylinders is necessary. Owing to demand from industry for butane derivations, LPG sold as fuel is made up largely of propane.

Operating figures for gasoline, diesel, and LPG fuels show that LPG compares favorably in cost per mile. LPG has a high octane rating, making it useful in engines having compression ratios above 10:1.

Another factor of importance in internal combustion engines is that LPG leaves little or no engine deposit in the cylinders when it burns. Also, since it enters the engine as a vapor, it cannot wash down the cylinder walls, remove lubricant, and increase cylinder-wall, piston, and piston-ring wear. Nor does it cause crankcase dilution. All these factors reduce engine wear, increase engine life, and keep maintenance costs low. However, allowances must be made for the extra cost of LPG-handling equipment, including relatively heavy pressurized storage tanks, and special equipment to fill fuel tanks on the vehicles. *See* GAS TURBINE; INTERNAL COMBUSTION ENGINE; PETROLEUM PRODUCTS.

Mott Souders

Liquid

A state of matter intermediate between that of crystalline solids and gases. Macroscopically, liquids are distinguished from crystalline solids in their capacity to flow under the action of extremely small shear stresses and to conform to the shape of a confining vessel. Liquids differ from gases in possessing a free surface and in lacking the capacity to expand without limit. On the scale of molecular dimensions liquids lack the long-range order that characterizes the crystalline state, but nevertheless they possess a degree of structural regularity that extends over distances of a few molecular diameters. In this respect, liquids are wholly unlike gases, whose molecular organization is completely random.

Thermodynamic relations. The thermodynamic conditions under which a substance may exist indefinitely in the liquid state are described by its phase diagram, shown schematically in the **illustration**. The area designated by *L* depicts those pressures and temperatures for which the liquid state is energetically the lowest and therefore the stable state. The areas denoted by *S* and *V* similarly indicate those pressures and temperatures

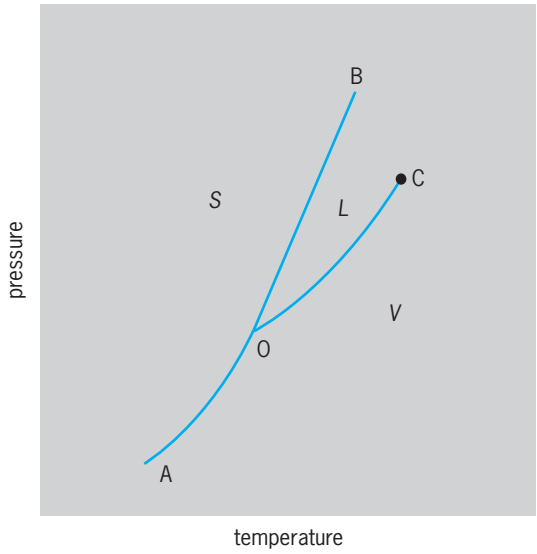

Phase diagram of a pure substance.

for which only the solid or vapor phase may exist. The connecting lines OC, OB, and OA define pressures and temperatures for which the liquid and its vapor, the solid and its liquid, and the solid and its vapor, respectively, may coexist in equilibrium. They are usually termed phase boundary or phase coexistence lines. The intersection of the three lines at O defines a triple point which, for the three states of matter under discussion, is the unique pressure and temperature at which they may coexist at equilibrium. Other triple points exist in the phase diagram of a substance that possesses two or more crystalline modifications, but the one depicted in the figure is the only triple point for the coexistence of the vapor, liquid, and solid. Line OA has its origin at the absolute zero of temperature and OB, the melting line, has no upper limit. The liquid-vapor pressure line OC is different from OB, however, in that it terminates at a precisely reproducible point C, called the critical point. Above the critical temperature no pressure, however large, will liquefy a gas. *See* TRIPLE POINT.

Along any of the coexistence curves the relationship between pressure and temperature is given by the Clausius-Clapeyron equation: $dP/dT = \Delta H/T\Delta V$, where ΔV is the difference in molar volume of the corresponding phases (gas-liquid, gas-solid, or liquid-solid) and ΔH is the molar heat of transition at the temperature in question. By means of this equation the change in the melting point of the solid or the boiling point of the liquid as a function of pressure may be calculated. When a liquid in equilibrium with its vapor is heated in a closed vessel, its vapor pressure and temperature increase along the line OC. ΔH and ΔV both decrease and become zero at the critical point, where all distinction between the two phases vanishes. *See* PHASE EQUILIBRIUM.

Transport properties. Liquids possess important transport properties, notably their capacity to transmit heat (thermal conductivity), to transfer momentum under shear stresses (viscosity), and to attain a state of homogeneous composition when mixed with other miscible liquids (diffusion). These nonequilibrium properties of liquids are well understood in macroscopic terms and are exploited in large-scale engineering and chemical-process operations. Thus, the rate of flow of heat across a layer of liquid is given by $\dot{Q} = \kappa\ dT/dx$, where $\dot{Q}$ is the heat flux, dT/dx is the thermal gradient, and κ is the coefficient of thermal conductivity. Similarly, the shearing of one liquid layer against another is resisted by a force equal to the momentum transfer: $F = \dot{P} = \eta\ dv/dx$, where dv/dx is the velocity gradient and η is the coefficient of viscosity. Likewise, the rate of transport of matter under nonconvective conditions is governed by the gradient of concentration of the diffusing species: $J = -D\ dC/dx$, where J is the matter flux and D is the coefficient of diffusion. Each of these transport coefficients depends upon temperature, pressure, and composition and may be determined experimentally. An a priori calculation of κ, η, and D is a very difficult problem, however, and only approximate theories exist.

Theoretical explanations. In fact, although a great deal of effort has been expended, there still exists no satisfactory theory of the liquid state. Even so commonplace a phenomenon as the melting of ice has no adequate theoretical explanation. The reason for this state of affairs lies in the tremendous structural and dynamical complexity of the liquid state. To understand this, it is useful to compare the structural and kinetic properties of liquids with those of crystalline solids on the one hand and with those of gases on the other.

In crystals, atoms or molecules occupy well-defined positions on a three-dimensional lattice, oscillating about them with small amplitudes; their kinetic energy is entirely distributed among these quantized vibrational states up to the melting point. This nearly perfect spatial order is revealed by diffraction techniques, which utilize the coherent scattering of x-rays or particles have wavelengths comparable with interatomic spacings. The structural and dynamical properties are sufficiently tractable mathematically so that the theory of solids is quite well understood.

The theory of gases is also simple, but for quite a different reason. No vestige of positional regularity of atoms remains in gases, and their energy resides entirely in high-speed translational motion. Except for collisions, which deflect their motions into new straight-line trajectories, atoms in gases do not interact with one another; vibrational modes in monatomic gases are absent.

Liquids, by contrast, lie intermediate between gases and crystals from both a structural and dynamic point of view. Kinetic energy is partitioned among translational and vibrational modes, and diffraction studies reveal a degree of short-range order that extends over several molecular diameters. Moreover, this "structure" is continually changing under the influence of vibrational and translational displacements of atoms. Physical reality may be attributed to this short-range structure, nevertheless, in the sense that a time average over the huge number of possible configurations of atoms may show that a fairly definite number of neighboring atoms lie close to any arbitrary atom. At a somewhat greater distance from this reference atom, the density of neighbors oscillates above and below the average density of atoms in the liquid as a whole.

Information about the degree of local order is contained in the radial distribution function, a mathematical property which may be deduced from diffraction measurements. This is the starting point for a theory of the liquid state, and although research efforts have yielded partial successes, prodigious mathematical difficulties lie in the path of an entirely satisfactory solution. *See* BOILING POINT; KINETIC THEORY OF MATTER; LIQUEFACTION OF GASES; MELTING POINT; VISCOSITY; X-RAY DIFFRACTION.

Norman H. Nachtrieb

Bibliography. J. P. Hansen and I. R. McDonald, *The Theory of Simple Liquids*, 2d ed., 1987, reprint 1990; P. Kruus, *Liquids and Solutions: Structure and Dynamics*, 1977; G. W. Rothschild, *Dynamics*

of Molecular Liquids, 1984; D. Tabor, *Gases, Liquids, and Solids: And Other States of Matter*, 3d ed., 1991.

Liquid chromatography

A method of chemical separation that involves passage of a liquid phase through a solid phase and relies on subtle chemical interactions to resolve complex mixtures into pure compounds. A small amount of the sample to be separated is injected onto the top of a column that is densely packed with spherical particles of small diameter, that is, the stationary phase. A liquid solvent, the mobile phase, flows through the column continuously to carry the sample from the top to the bottom of the column. During passage through the column, the components of the sample are transferred back and forth continuously between the two phases, and small thermodynamic differences in the chemical interactions of the various sample components with the mobile and stationary phases slow the passage of some solutes more than others and lead to their separation. The technique can be performed on very small scales for chemical analysis, dealing with micrograms or even nanograms of sample, or it can be performed on an industrial scale for purification of commercial products. The technique has great resolving power (see **illus.**).

Liquid chromatography as practiced today differs only in detail from the technique as first described by its inventor, M. Tswett, a Russian botanist, in 1906. Many advances in instrumentation and understanding of the chemistry of the interactions involved have led to improvements in the separations, but the fundamental idea of the technique has not changed. This method of chemical analysis is widely used in the pharmaceutical, biomedical, biotechnology, chemical, oil, and polymer industries and to some extent in virtually every other chemical-based industry. *See* CHEMICAL SEPARATION TECHNIQUES.

Instrumentation. In the late 1960s, workers realized that to achieve maximum performance they needed to make the stationary-phase particles very small. As the stationary-phase particles became smaller, they packed too densely to permit gravity-driven flow of the mobile phase. It was necessary to force the mobile phase through the column under high pressure, and the technique was named high-pressure liquid chromatography (HPLC). The meaning of the acronym has been changed to high-performance liquid chromatography, for the elegant separations that are possible. Typical stationary-phase particles are monodisperse, macroporous silica particles either 3 or 5 micrometers in diameter, and the column lengths for analytical scale separations are on the order of 2–10 in. (5–25 cm), with inside diameters of about 0.16 in. (4 mm). The columns are made of stainless steel, which is relatively inert chemically and able to withstand the high pressures applied to the top of the column. Since these columns require pressures of a few hundred to a few thousand pounds per square inch, depending on the mobile-phase flow velocity desired, a high-pressure metering pump is an integral part of a modern liquid chromatograph. Most instruments include a means of performing gradient elution, that is, making a continuous change in the composition of the liquid mobile phase during the separation process. Gradient elution can be performed by using a separate pump for each solvent and changing the relative proportions during the separation, or by using a proportioning valve between the solvent reservoirs and the pump. An injector valve is used to introduce a small volume of sample, typically 5–100 microliters, onto the top of the column without interrupting the mobile-phase flow. These valves can be operated manually, or they can be programmed to perform injections from a tray of samples for routine analyses. After the sample traverses the column, a flow-through detector is employed to generate the chromatogram, which is the visual representation of the separation. Detectors can provide both quantitative and qualitative information about the separated components. Temperature control of the column and detector is important; they are generally operated at or near room temperature, and temperature fluctuations can adversely affect the reproducibility of the separation and detection steps.

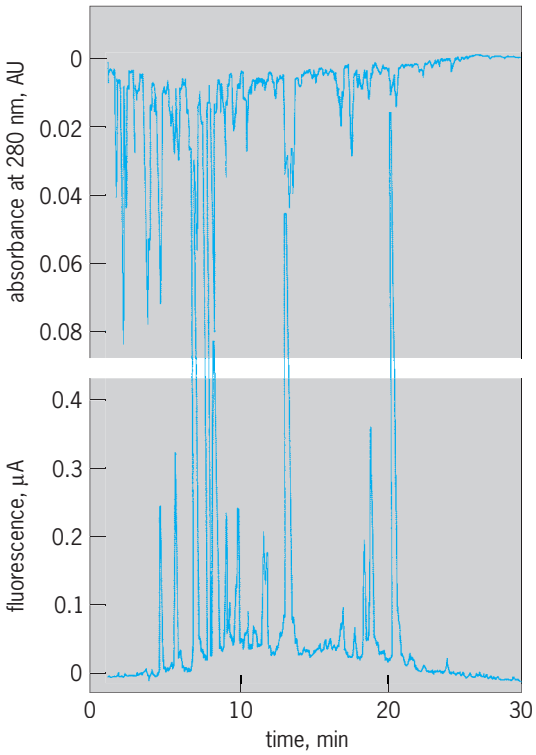

Chromatogram of a urine sample with added aromatic acids. Ten microliters of sample was separated on a 10-in. (25-cm) reversed-phase column using gradient elution. The separation was monitored with both fluorescence and absorbance detectors, and 119 compounds were separated in an analysis time of just under 30 min. (*After I. Molnar, C. Horvath, and P. Jatlow, Separation of urinary aromatic acids and measurement of vanilmandelic acid by high performance liquid chromatography, Chromatographia, 11:260–265, 1978*)

Separation modes. Liquid chromatography very much depends on the highly selective chemical interactions that occur in both the mobile and stationary phases. Rapid separations have become possible for compounds whose difference in free energy of transfer between the two phases is only a few calories per mole. Columns exhibiting virtually every type of possible selectivity exist, including selectivity by shape, charge, size, and optical activity. Additional selectivity can be generated through manipulations of the mobile phase; additives that interact with the solute in the mobile phase can create unique selectivities in columns that do not normally show that type of selectivity.

Ion exchange. This is a popular mechanism for the separation of charged solutes. Ion-exchange chromatography involves the reversible exchange of ions between mobile and stationary phases. The stationary phase most often consists of beads of a polymer with charge-bearing functional groups, and the sample's ions compete with a mobile-phase ion for the ionic sites on the ion exchanger. This separation mode is frequently used in the electronics industry, for environmental analysis, and for the separation of proteins and biopolymers that have charged groups. *See* ION EXCHANGE.

Size exclusion. This type of chromatography involves separation based on the molecular size of the solutes. Exclusion columns contain porous stationary-phase materials, usually silica gel or a porous polymer, with different pore diameters. A solute injected onto such a column will diffuse into those pores that have a diameter greater than the effective diameter of the solute's molecules (under solvent conditions at that instant), and the effective diameters of the various solutes control elution order. The solute whose molecules have the largest effective diameter will elute first, since there are fewer pores that it can fit into. A smaller solute will diffuse into many more of the stationary-phase pores and will require a much larger volume of solvent to achieve elution. This separation mode is heavily used in the petrochemical and polymer industries for characterization of molecular weights and sizes of synthetic and natural polymers.

Bonded phases. In the early 1970s, technology was developed to chemically bond functional groups onto spherical silica particles, providing the efficient and highly reproducible stationary phases needed for liquid chromatography to gain acceptance as a routine method of chemical analysis. This bonding technology has the advantage of allowing researchers to tailor the surface chemistry of small-diameter silica particles to give virtually any desired interaction. There are reports of almost every conceivable functional group's being bonded to the surface, both alone and in mixtures to yield mixed interaction phases. *See* SURFACE AND INTERFACIAL CHEMISTRY.

Normal and reversed phase. Polar groups can be bonded to the surface of silica beads and used with nonpolar mobile phases for a separation mechanism known as normal-phase chromatography. The latter is best suited for separating polar compounds and geometrical isomers and for performing class separations. These separations are similar to those obtained on bare silica. The most popular mode of liquid chromatography is reversed-phase. Here a molecule with nonpolar functionality, typically a fully saturated 18-carbon hydrocarbon chain, is bonded to the silica surface, giving the stationary phase a highly nonpolar character. Aqueous-based mobile phases are used, and it is estimated that this mode of separation is used in 75% or more of analytical scale chromatography. The reversed-phase method has a number of advantages. The aqueous-based mobile phases provide good compatibility with many detection techniques and are also compatible with biological samples. The interactions of the solute with the stationary phase are weak. This provides rapid mass transfer and sharp chromatographic peaks, as well as rapid equilibrium with solvent changes, making the technique ideal for gradient elution methods. A very powerful advantage of aqueous-based mobile phases is the ability to control pH and other mobile-phase-based equilibria for additional separational selectivity. Weak acids and bases are retained very differently in their charged and uncharged forms, so pH control is a powerful means of modulating selectivity and of moving ionizable compounds relative to uncharged species. Complexation equilibria can also be advantageously used, such as with cyclodextrins for the separation of optical isomers.

One of the most popular types of chromatography that uses secondary equilibria is ion-pairing chromatography. This involves the addition of a hydrophobic counter-ion to the mobile phase, which is adsorbed onto the surface of the stationary phase and makes the reversed-phase stationary phase into a dynamic ion exchanger. Ion-pairing chromatography is popular because it allows the separation of charged and neutral species in the same sample. It is frequently used in pharmaceutical analysis.

Hydrophobic interaction. A variation on the reversed-phase mode is hydrophobic interaction chromatography. Nonpolar groups are bonded onto the silica surface at very low densities (number of groups per unit surface area). This technique is used mostly for protein separations, with mobile-phase conditions generally of a descending salt gradient. Retention is controlled by interactions between the hydrophobic ligands of the stationary phase and the hydrophobic entities of the protein. While interactions of small molecules are relatively well understood, the mechanism of protein separations is still largely unknown.

Affinity chromatography. This is the most highly selective chromatographic method. The technique closely mimics the types of binding processes found in nature, including antigen-antibody, enzyme-substrate, and hormone-receptor. Whereas most chromatographic methods are aimed at separating all solutes present in the sample, affinity chromatography attempts to retain only one solute on the column, while all undesired components are swept through. The mobile-phase conditions are then changed, and the desired analyte is eluted in a single peak. Affinity chromatography, really a form of digital chromatography, is useful for both analysis and purification of

biological materials. In affinity chromatography the solute is retained completely; the mobile phase is changed and the solute is instantly eluted. This is analogous to the "on" or "off" in digital electronics.

Other applications. In addition to facilitating chemical analysis, liquid chromatography can be used to obtain physicochemical information. Diffusion coefficients, kinetic parameters, critical micelle concentrations of surfactants, and other information have been estimated from chromatographic data. The most common application is the estimation of hydrophobic parameters, especially as models of biological or environmental partitioning processes (most frequently, of octanol-water partitioning). Bioavailability, bioaccumulation, soil sorption, and various other factors are estimated based on linear free-energy relationships. *See* CHROMATOGRAPHY.

John G. Dorsey

Bibliography. U. D. Neue, *HPLC Columns: Theory, Technology, and Practice*, 1997; P. C. Sadek, *The HPLC Solvent Guide*, 1996; L. R. Snyder, J. J. Kirkland, and J. L. Glajch, *Practical HPLC Method Development*, 2d ed., 1997.

Liquid crystals

A state of matter with an orientational order of building units—individual molecules or their aggregates—and complete or partial absence of the long-range positional order. Liquid crystals, discovered more than 100 years ago, are one of the best-studied classes of soft matter, along with colloids, polymer solutions and melts, gels, and foams.

The liquid-crystalline state, also known as the mesophase, is in between that of a regular solid having long-range positional order of atoms (or molecules which are also orientationally ordered in crystals) and an isotropic fluid in which the molecules show neither positional nor orientational order. Liquid crystals can flow, adopt the shape of a container, and form drops as regular isotropic fluids do; however, the molecules in the liquid-crystal sample are ordered. The direction of average orientation in liquid crystals is called the director **n**.

Molecular interactions responsible for the orientation order in liquid crystals are relatively weak, as most liquid crystals melt into the isotropic phase below 100–150°C (212–300°F). As a result, the structural organization of liquid crystals, most importantly the spatial configuration of the director and thus the optical properties, are very sensitive to the external factors such as electromagnetic field, shear deformations, and type of molecular orientation on bounding walls. This sensitivity allows for numerous applications of liquid crystals, including liquid-crystal displays (LCDs). In these devices, weak electric voltage pulses reorient the director and change the optical appearance of the liquid-crystal cell. *See* OPTICAL MATERIALS.

Classification. Depending on the way the liquid-crystalline state is produced, one distinguishes thermotropic and lyotropic liquid crystals.

Thermotropic (solvent-free) liquid crystals form either by heating a solid crystal or by cooling an isotropic fluid. They exist in a certain temperature range for the materials made of strongly anisometric (elongated or disklike) molecules. An individual substance forming a thermotropic liquid crystal does not require any solvent to exhibit the mesophase. In practical applications, a number of liquid-crystal materials are often mixed together to improve their functional properties (such as their useful temperature range) or doped with additives that are not necessarily mesomorphic (such as dyes, chiral molecules, and polymer inclusions).

Lyotropic liquid crystals form only in the presence of a solvent, such as water or oil. Most commonly, lyotropic mesophases are formed by solutions of amphiphilic molecules (such as soaps, phospholipids, and surfactants). Amphiphilic molecules have two distinct parts: a hydrophilic (polar) head and a hydrophobic (nonpolar) tail, which is generally an aliphatic chain. As a result, amphiphilic molecules in solvents give rise to "self-organization," as manifested by the formation of micelles and bilayers. Mesomorphic states (between liquid and solid) may also form in solutions of certain polymers; polymers may also form thermotropic liquid crystals. *See* MICELLE.

Phases. There are four basic types of liquid-crystalline phases, classified according to the dimensionality of the positional order of building units: nematic (only orientational order and no positional order), smectic (orientational and 1D positional order, caused by a periodic variation of the molecular densities), columnar (orientational and 2D positional order), and various 3D-correlated structures (such as cubic phases and blue phases).

Upon heating, many thermotropic substances, such as octylcyanobiphenyl (8CB) [**Fig. 1**], yield the following phase sequence: solid crystal > smectic (or columnar phase) > nematic > isotropic fluid. On cooling, the sequence is generally reversed, although it often includes phases that are absent during heating.

Nematics. There are three different variations of the nematic phase: uniaxial nematic (UN), biaxial nematic (BN), and cholesteric (Ch). The UN is formed by the molecules of a rodlike or disklike shape that can easily rotate around one of the axes. The average orientation of these axes of rotation coincides with the director **n** (**Fig. 2***a*). Even when the molecules are polar, as the 8CB molecules are, their

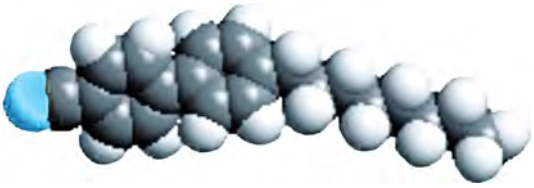

Fig. 1. A rodlike molecule of octylcyanobiphenyl (8CB), where the phases formed with increasing temperature are solid crystal > smectic A (24°C) > nematic (34°C) > isotropic fluid (42.6°C).

head-to-head overlapping and flip-flops establish centrosymmetric arrangement in the nematic bulk. The UN is an optically anisotropic medium; the optic axis coincides with the director (Fig. 2*a*). The speed of light propagation in the nematic medium depends on how the polarization of light is directed with respect to the optic axis, giving rise to the phenomenon of birefringence. *See* BIREFRINGENCE.

In BN, the shape of building units significantly deviates from axially symmetric and can be more appropriately described as boxlike. These materials exhibit the orientational order of all three axes and have three directors (Fig. 2*b*).

When the building unit (molecule or aggregate) is not equal to its mirror image (chiral), the director **n** might show a twisted structure. Most often, the twist is unidirectional and the director forms a helicoid by rotating around a single axis in space, which is the case of the Ch phase (Fig. 2*c*). The UN, BN, Ch, and blue phases are liquid phases (no correlations in molecular positions over large distances). However, Ch shows a periodic structure associated with the spatial twist of molecular orientation (Fig. 2*c*). The pitch (period) of this twist is much larger than the molecular size, as molecular interactions responsible for the twist are weak as compared to the interactions responsible for the mesomorphic state itself. With a typical molecular size of a liquid crystal being 1–10 nanometers, the pitch is often in the range of 0.1–1 micrometer, which overlaps with the range of the visible part of optical spectrum.

Smectics. Smectics are layered phases with a quasi-long-range 1D translational order of centers of molecules in a direction normal to the layers. This positional order is not exactly the long-range order as in a regular 3D crystal. As shown by L. D. Landau and R. E. Peierls, the fluctuative displacements of layers in 1D lattice diverge with the linear size of the sample. However, for typical smectic materials with a period of the order of 1 nm, the effect is noticeable only on scales of 1 mm and larger; it is thus not noticeable in samples that are normally 1–100 μm thick. In smectic A (SmA), the molecules within the layers show fluidlike arrangement, with no long-range in-plane positional order; it is a uniaxial medium with the optic axis and the director perpendicular to the layers (Fig. 2*d*).

In the lyotropic version of SmA, the lamellar L_α phase, the amphiphilic molecules arrange into bilayers. If the solvent is water, the exterior surfaces of the bilayer are formed by polar heads; the hydrophobic tails are hidden in the middle of the bilayer. (The membranes of many biological cells are organized in the similar way.) The periodic structure of alternating surfactant and water layers gives rise to the L_α phase (Fig. 2*e*). The lamellar structure may retain its smectic order, even when strongly diluted, as it is stabilized by thermal fluctuations of bilayers.

Other types of smectics show in-plane order, caused, for example, by a collective tilt of the rodlike molecules with respect to the normals to the layers (the so-called SmC) [Fig. 2*f*]. In chiral materials, the tilt of the molecules might lead to the helicoidal structure. The resulting chiral SmC phase is of considerable interest for applications in fast-switching optical devices (Fig. 2*g*).

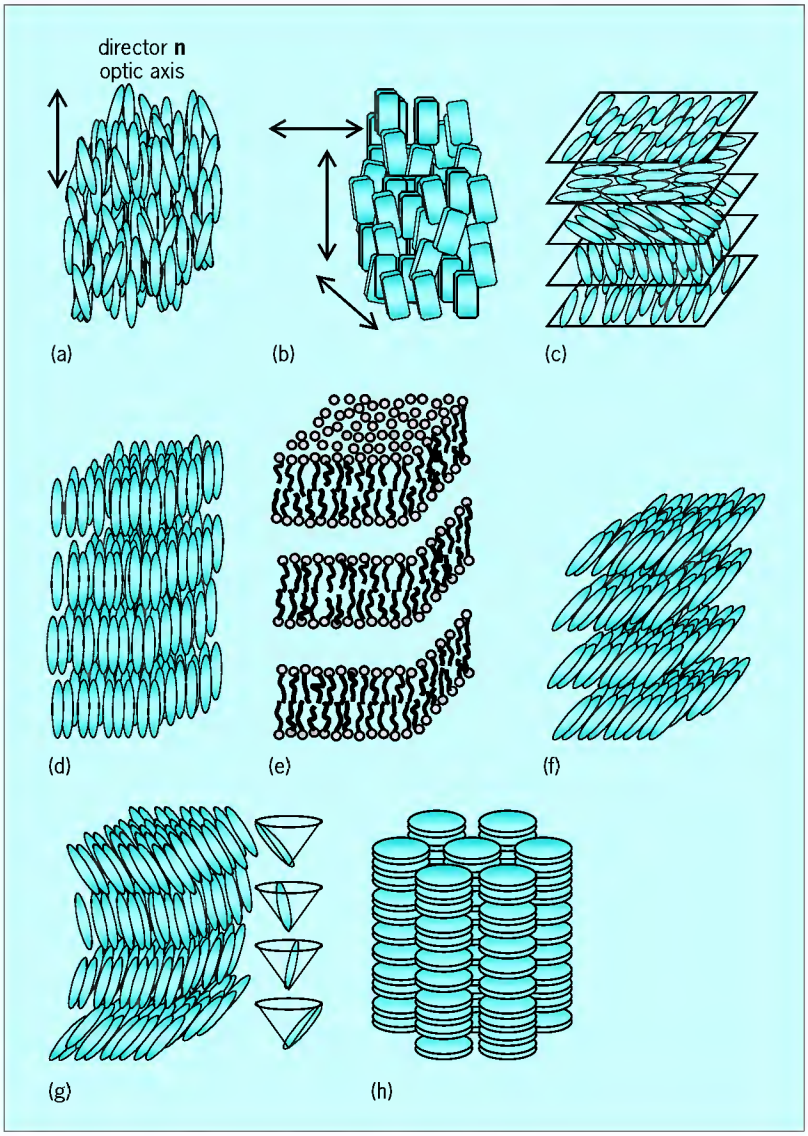

Fig. 2. Liquid-crystalline phases. (a) Uniaxial nematic; the molecular long axes are on average aligned along the director n which is also the optic axis. (*b*) Biaxial nematic; three molecular axes are on average aligned along three directors. (c) Cholesteric: the local nematic director rotates in space forming a helicoid. (*d*) Smectic A type of ordering in the thermotropic liquid crystal; the molecular axes are perpendicular to the plane of layers. (e) Smectic A type of ordering in the lyotropic L_α phase formed by equidistant arrangement of amphiphilic bilayers in water. (*f*) Smectic C; the molecular axes within the layers are tilted. (*g*) Chiral smectic C; the molecular axes are tilted and the direction of tilt forms a helicoid. (*h*) Columnar liquid crystal.

Columnar. Columnar phases are most frequently formed by hexagonal packing of cylindrical aggregates, as in the case of thermotropic materials formed by disklike molecules. The positional order is 2D only, as the intermolecular distances along the axes of the aggregates are not regular (Fig. 2*h*).

3D-correlated structures. These demonstrate a periodic structure along all three coordinates, but they are still different from the 3D crystals, as the periodicity is associated with repetition of the local director patterns rather than with a regular positional order of the molecular centers of mass. The molecules are

free to move around, adjusting to the local director orientation. For example, in cubic lyotropic phases the 3D network is formed by periodically curved layers of amphiphilic molecules. Another example is the blue phases that occur in the narrow temperature range between the isotropic fluid phase and Ch phase, in which the director forms double-twisted structures. The double twist cannot expand through the entire volume of the material and must be accompanied by a 3D periodic lattice of defects.

Order parameter. The director specifies only the direction of average orientation, not the degree of orientational order. For the latter, one can chose the temperature-dependent quantity $s(T) = \frac{1}{2}\langle 3\cos^2\theta - 1\rangle$, where θ is the angle between the axis of an individual molecule and the director, and $\langle \ldots \rangle$ means an average over molecular orientations. The term s is often called the scalar order parameter, where $s = 1$ corresponds to an ideal order with all the molecules rigidly aligned along one direction and $s = 0$ corresponds to absence of orientational order (in which case one deals with regular isotropic fluids).

The order parameter of the liquid crystal can be related to the anisotropy of macroscopic properties such as diamagnetic or dielectric susceptibility. Measuring these anisotropies allows one to determine the degree of orientational order. The magnetic measurements are especially convenient compared with their electric counterparts, as in this case the local field acting on the molecules differs very little from the external field. In uniaxial nematics (UNs), the values of the magnetic susceptibility measured in the direction parallel to the director and perpendicular to it are different. This difference $\chi_a = \chi_\parallel - \chi_\perp$ is called the anisotropy of the magnetic susceptibility. In most thermotropic UNs, $\chi_\parallel < 0$ and $\chi_\perp < 0$ (diamagnetism), and $\chi_a > 0$, so that the director orients along the applied magnetic field. In the isotropic phase, $\chi_a = 0$. For practical applications of UNs, of prime importance is the anisotropy of two other properties: the anisotropy of dielectric permittivity $\varepsilon_a = \varepsilon_\parallel - \varepsilon_\perp$ and optical birefringence—that is, the difference in the refractive indices for light polarized parallel to the director (the extraordinary index of refraction $n_e = n_\parallel$) and for light polarized perpendicularly to the director (the ordinary index of refraction $n_o = n_\perp$), i.e., $\Delta n = n_e - n_o$. When $\varepsilon_a > 0$, the director reorients along the applied electric field; and when $\varepsilon_a < 0$, the director realigns perpendicularly to the field. Because the director is simultaneously the optical axis, and because of the nonzero birefringence, $\Delta n \neq 0$, these field-induced reorientations change the optical properties of the sample, such as the effective value of the refractive index. The latter effect is at the core of numerous devices that use the dielectric response of the UN to display information or to control optical properties of the liquid crystal cell. *See* ELECTRIC SUSCEPTIBILITY; MAGNETIC SUSCEPTIBILITY; REFRACTION OF WAVES.

Elasticity and the free energy of the nematic phase. The molecular orientation in a liquid-crystal sample might change from point to point because of the external fields, boundary conditions, presence of foreign particles, and so on. The order parameter becomes spatially nonuniform. In most problems of practical interest, the typical scale of distortions is much larger than the molecular scale. The deformations are thus weak in the sense that the scalar part of the order parameter, $s(T)$, remains constant despite the spatial gradients of the director $\mathbf{n}(\mathbf{r})$.

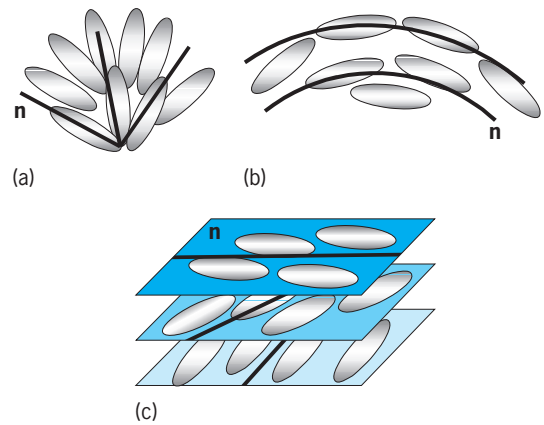

Fig. 3. Basic types of director distortions in the bulk of the uniaxial nematic: (*a*) splay, (*b*) bend, and (*c*) twist.

There are three basic types of director distortions in uniaxial nematics, called splay, twist, and bend (**Fig. 3**). The free-energy density associated with these three types of (small) deformation is written in terms of the spatial director gradients

$$n_{i,j} = \frac{\partial n_j}{\partial x_j}$$

as Eq. (1) and is known as the Frank-Oseen energy

$$f_{\mathrm{FO}} = \frac{1}{2}K_1(\operatorname{div}\mathbf{n})^2 + \frac{1}{2}K_2(\mathbf{n}\cdot\operatorname{curl}\mathbf{n})^2 + \frac{1}{2}K_3(\mathbf{n}\times\operatorname{curl}\mathbf{n})^2 \quad (1)$$

density, with Frank elastic constants of splay (K_1), twist (K_2), and bend (K_3); all three are necessarily positive-definite and the dimensionality is that of a force. The elastic constants can be estimated as the typical energy of molecular interactions responsible for the orientational order divided by the characteristic length, which is the molecular size

$$K \sim U/l \sim k_B T/l \sim 4\times 10^{-21}\mathrm{J}/10^{-9} \sim 4\ \mathrm{pN}$$

which is a very good estimate for many thermotropic uniaxial nematic materials, as the experimental values are between 1 and 10 piconewtons (pN). The energy density [Eq. (1)] is often supplemented with the so-called divergence terms that can influence the equilibrium director through boundary conditions at the surface. In the presence of an external field, the free-energy density acquires additional terms; for example, a diamagnetic term $f_{\mathrm{field}} = -\frac{1}{2}\mu_0^{-1}\chi_a(\mathbf{B}\cdot\mathbf{n})^2$ in the magnetic field $\mathbf{B}$; here $\mu_0 = 4\pi\times 10^{-7}$ henry/m is the magnetic permeability of free space.

Another important contribution to the free energy is related to the phenomenon of director anchoring at the bounding surfaces. Any actual liquid-crystal cell is confined, say, by a pair of parallel glass plates.

The molecular interactions between the liquid crystal and the boundary substrates are anisotropic. This anisotropy establishes one (sometimes more) preferred orientation of **n** at the boundary, the "easy axis." The phenomenon is called surface anchoring. Orienting action of the substrates usually keeps the director uniform if the external field is absent. To change this orientation, one needs to perform some work.

The equilibrium of a uniaxial nematic sample of a given shape in an external field is determined by the minimum of the functrional, Eq. (2),

$$F = \int_v (f_{\mathrm{FO}} + f_{\mathrm{field}})dV + \int_S F_{\mathrm{anch}} dS \qquad (2)$$

where f_{anch} is the surface-anchoring energy density, often represented by the simple potential $f_{\mathrm{anch}} = \frac{1}{2}W_a\alpha^2$, with W_a being the surface-anchoring coefficient and α the angle between the easy axis and the actual director orientation at the surface.

Elasticity of the smectic A phase. For the smectic A (SmA) phase, the elastic free-energy density, Eq. (3),

$$f = \tfrac{1}{2}K_1(\mathrm{div}\,\mathbf{n})^2 + \tfrac{1}{2}B\gamma^2 \qquad (3)$$

is different from Eq. (1) because of certain restrictions that the layered structure imposes onto the director distortions and the elastic cost of changes in the thickness of the layers. Here B is the Young modulus (layers compressibility modulus) and $\gamma = (d - d_0)/d_0$, the relative difference between the equilibrium period d_0 and the actual layer thickness measured along the director **n**. The ratio of K_1 to B defines an important length scale, Eq. (4),

$$\gamma = \sqrt{K_1/B} \qquad (4)$$

called the penetration length γ, and is of the order of the layer separation but diverges when the system approaches the SmA-nematic transition. One expects that a SmA would have K_1 of the same order as in a nematic phase stable at higher temperatures. With $\lambda \approx d_0 \approx (1-3)$ nm, one finds $B \sim 10^6 \div 10^7$ N/m^2, a value that is 10^3 to 10^4 times smaller than the compressibility modulus in a solid.

Dynamics. Liquid crystals are fluids. They can flow while preserving the orientational order. Flow imposes an orientational torque on **n**. Most often, **n** tends to realign along the direction of flow. There is also a reverse effect by which director distortions can cause the flow. This "backflow" effect is of importance in liquid-crystal displays. In the approximation of a constant scalar order parameter, the hydrodynamics of liquid crystals is described in terms of seven variables: (1) mass density, (2) three components of the velocity field, (3) energy density, and (4) two components of the director field $\mathbf{n}(\mathbf{r},t)$. In contrast to an isotropic fluid, the stress tensor depends not only on the gradients of the velocity but also on the director components. The uniaxial nematic phase should be characterized by five different viscosity constants. The number of viscosities reduces to three when the director distortions are small. These three can be chosen as the effective viscosities for three idealized geometries of flow, also known as Miezowicz geometries, in which one assumes that the director is fixed (for example, by a strong magnetic field): (a) When $\mathbf{n} = (1,0,0)$ is perpendicular to both the flow direction and the velocity gradient, the UN behaves as an isotropic fluid with a viscosity η_a; however, director fluctuations coupled with the certain values of the viscosity coefficients might destabilize the initial orientation. (b) When **n** is parallel to the flow or (c) when **n** is parallel to the velocity gradient (**Fig. 4**), the corresponding viscosities η_b and η_c are generally different from η_a and from each other; $\eta_b < \eta_a < \eta_c$ for a typical uniaxial nematic material composed of rodlike molecules; the result $\eta_b < \eta_c$ can be explained by assuming that the friction correlates with the cross section of the molecules seen by the flow.

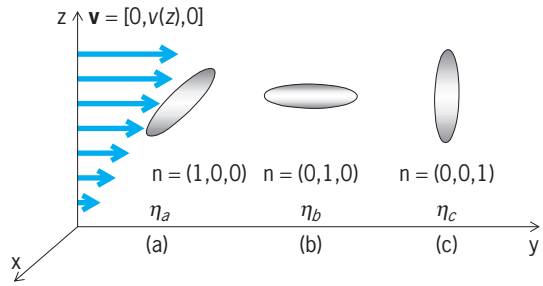

Fig. 4. Miezowicz geometries for effective viscosities of the uniaxial nematic.

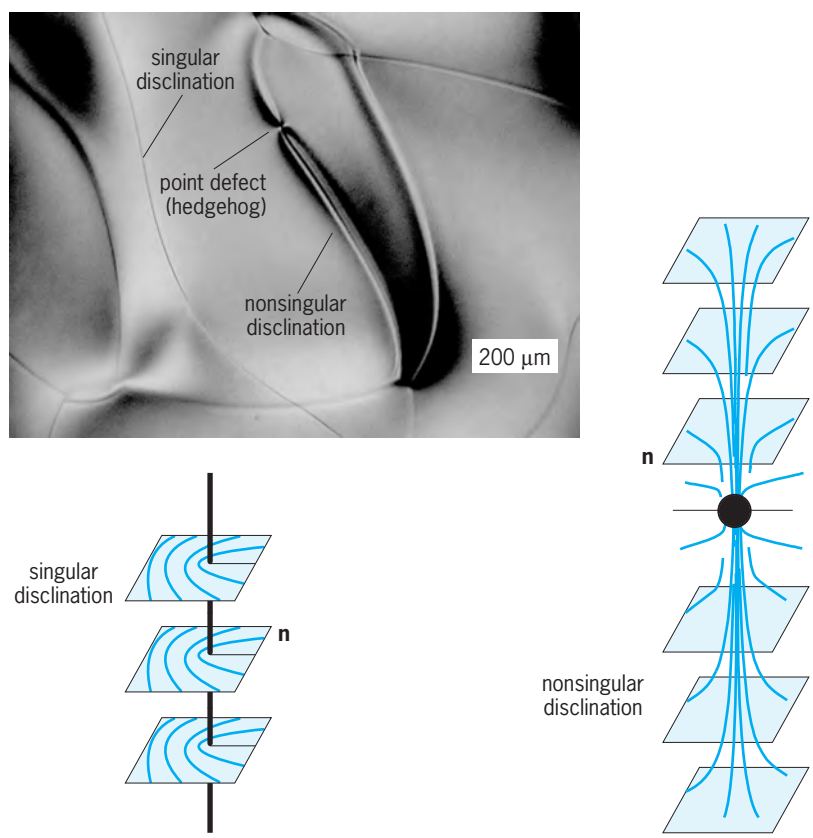

Fig. 5. Thin singular disclinations and thick nonsingular threads in the nematic [*n*-pentylcyanobiphenyl (5CB)] bulk, as observed under the microscope with crossed polarizers, and the corresponding director configurations.

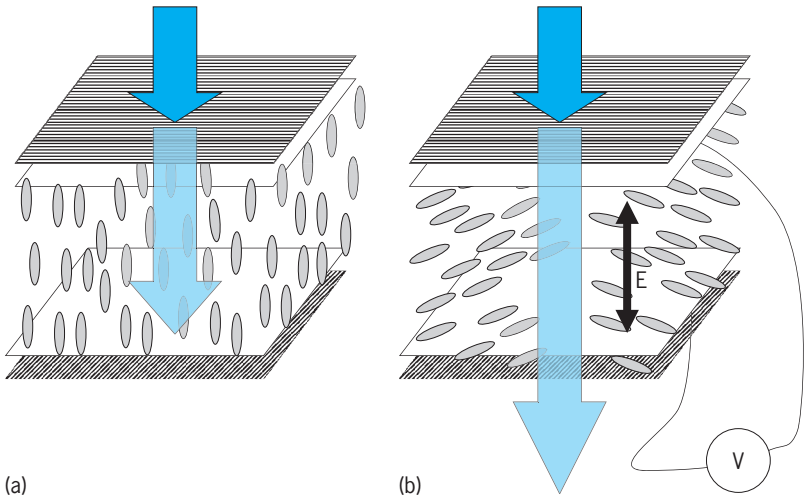

Fig. 6. Electrically controlled uniaxial nematic cell between a pair of crossed linear polarizers that either (*a*) blocks the transmitted light field off state or (*b*) allows the light to pass through the field on state.

Topological defects. When a thick (100-μm) nematic sample with no special aligning layers is viewed with a microscope, one usually observes a number of mobile flexible lines, called disclinations. The disclinations are seen as thin and thick threads (**Fig. 5**). Thin threads strongly scatter light and show up as sharp lines. These are truly topologically stable defect lines, along which the nematic symmetry of rotation is broken. The disclinations are topologically stable in the sense that no continuous deformation can transform them into a state with the uniform director field, $\mathbf{n}(\mathbf{r}) = \text{const}$. Thin disclinations are singular in the sense that the director is not defined along the line. Thick threads are line defects only in appearance; they are not singular disclinations. The director is smoothly curved and well defined everywhere, except for a number of point defects in the nematic bulk, called hedgehogs.

Applications of liquid crystals. The possibility to orient the director by an applied electric field leads to numerous practical applications. In a flat cell, the orienting action of the homogeneously treated substrate usually keeps the director uniform if the external field is absent. However, if the liquid crystal is dielectrically anisotropic, a sufficiently high electric field can overcome both the anchoring effect and the elasticity of the nematic bulk and reorient the director, either parallel ($\varepsilon_a > 0$) to the field or perpendicular to it ($\varepsilon_a < 0$). This is the Frederiks effect, first discovered for the magnetic case. When the field is removed, the surface anchoring and bulk elasticity restore the original director structure. Thus, one can use the electric field and surface anchoring to switch the elastic liquid crystal back and forth.

Various implementations of the Frederiks effect in uniaxial nematic (UN) and cholesteric (Ch) are widely used in electrooptic devices, including displays. The liquid crystal is usually sandwiched between two transparent electroconductive plates (for example, glass covered with indium tin oxide) coated with a suitable alignment layer, most often chosen from the polyimide family. The voltage across the cell controls the director configuration and thus the optical properties of the cell. In one of the most recent and commercially successful embodiments, the initial field-free state is the one in which $\mathbf{n}$ is perpendicular to the bounding plates with electrodes (**Fig. 6**). The cell is sandwiched between two crossed light-polarizing films. The light beam becomes linearly polarized after passing the first polarizer. This state of polarization is not changed by the liquid crystal, and the light beam is thus blocked by the second polarizer. The UN material is of a negative dielectric anisotropy, $\varepsilon_a < 0$. When the electric field is applied across the cell, the director deviates from the vertical axis. In this "field-on" state, because of the finite birefringence, the cell allows the light beam to pass through the second polarizer. The direction of director tilt in the plane of the UN cell can be controlled by using specially patterned electrodes; thus the mode is known as the patterned vertical alignment (PVA) mode. The field can also be applied in the plane of the sample, comprising a layer of a UN with $\varepsilon_a > 0$, thus forming the in-plane switching (IPS) mode. Finally, the polarization of light can be controlled by using the so-called twisted nematic (TN) cell, in which the electric field switched the twisted (chiral) nematic structure into a vertically aligned director configuration, thus altering the appearance of the UN cell between two polarizers, either crossed or parallel. To create a full-color display using the PVA, IPS, or TN techniques, one adds a pixelized color filter with red, blue, and green pixels to the glass plate facing the viewer. The panel that controls the voltage across the liquid crystal also is pixelized, so that the intensity of light transmitted by each pixel and the corresponding color filter is controlled independently of other pixels. The most advanced scheme of electric-addressing the liquid crystal display is the active matrix scheme, based on the array of thin-film transistors (TFTs). The operating voltage needed in liquid-crystal displays is relatively low, usually between 1 and 10 V. *See* ELECTRONIC DISPLAY; ELECTROOPTICS; FLAT-PANEL DISPLAY DEVICE.

The flexibility of molecular architectures of liquid crystals allows many other types of applications, including the information displays based on the effects of light scattering rather than the changes in

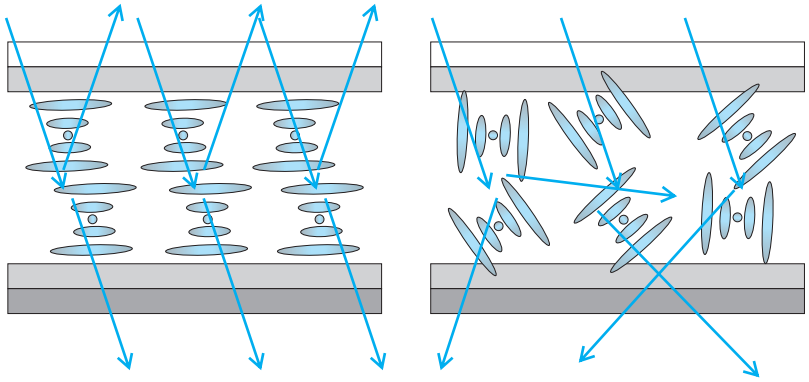

Fig. 7. Ch bistable display based on electrically controlled light scattering.

polarization state. Historically, one of the first liquid-crystal devices was the cholesteric thermometer that changed colors as the function of temperature. In Ch, the helicoidal pitch is often in the range of 0.1–1 μm (the molecular interactions responsible for the chiral structure are relatively weak, thus the pitch is much larger than the size of an individual molecule). The well-aligned periodic director structure is thus capable of selective reflection in the visible part of the spectrum, which brings a "colored" appearance to the sample viewed in the white light. The spectral range of this selective reflection is related to the value of the cholesteric pitch and can be changed by changing the pitch. Ch materials with electrically controlled orientation of the helicoids axis are used in bistable light-reflective liquid-crystals displays (**Fig. 7**). The Ch is switched by the electric voltage pulses between a well-aligned (planar) light-reflecting state and a disoriented light-scattering state with numerous defects of the director configuration. By adding a black absorbing layer to the bottom of the Ch film, one creates a bistable Ch display with a bright colored appearance in the planar state and black appearance in the distorted state. The field is needed only to switch the Ch sample from one state to another, which greatly reduces energy consumption. Although only the planar state is truly stable, the distorted state is relatively stable, as its relaxation into the planar state is hindered by a large energy barriers associated with the helix distortions. As the display works in the light-reflective mode, it does not need a special illuminating panel beneath the liquid-crystal layer, as most other liquid crystals displays do. The Ch reflective displays are used in "electronic books" and in development of flexible displays in which the traditional glass panels are replaced with plastic films.

Some applications use composite materials in which the liquid crystal is a component. Polymer-dispersed liquid crystals are one example. By dispersing the UN droplets in a polymer matrix (through a phase separation process), one can create an effective light-scattering medium, provided the effective index of the liquid-crystal droplet is different from the refractive index of the polymer. By applying the electric field, the droplets and the director inside them reorient along the field. If the materials are selected in such a way that the ordinary index of refraction of the UN matches the refractive index of the polymer, the composite film becomes transparent. The composite film thus represents an electrically switchable light-scattering panel and can be used in large-area (square meters) "privacy windows."

Oleg D. Lavrentovich

Bibliography. L. M. Blinov and V. G. Chigrinov, *Electrooptic Effects in Liquid Crystal Materials*, Springer, 1996; S. Chandrasekhar, *Liquid Crystals*, Cambridge University Press, 1992; P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals*, Oxford Science, 1993; M. Kleman and O. D. Lavrentovich, *Soft Matter Physics: An Introduction*, Springer, 2003; R. G. Larsen, *The Structure and Rheology of Complex Fluids*, Oxford University Press, 1999; E. Lueder, *Liquid Crystal Displays: Addressing Schemes and Electro-optical Effects*, Wiley, 2001; S. T. Wu and D. K. Yang, *Reflective Liquid Crystal Displays*, Wiley, 2001.

Liquid helium

Helium boils at a substantially lower temperature, 4.2 K (−452°F or −269°C), than any other substance; and below 2.172 K (−455.76°F) the liquid exhibits the extraordinary properties of superfluidity, notably the ability to flow through narrow channels with complete absence of friction. In addition to the common isotope of atomic weight 4, helium has a rare isotope of atomic weight 3 with a normal boiling point of 3.2 K (−454°F) and a superfluid transition at a very much lower temperature near 0.001 K. Both forms of helium remain in a liquid state at absolute zero. All of these characteristics are due to the weakness of the attractive force between two helium atoms and to the small atomic mass, which according to the laws of quantum mechanics makes the atoms difficult to localize.

At 4.2 K (−452°F) liquid ^{4}He is colorless and of low refractive index ($n = 1.024$), with a density of 0.125 g/cm^3 (0.125 times that of water). The latent heat of vaporization, 5 cal/g (21 J/g), is very small, and so care must be taken to reduce the heat input by conduction and radiation into the storage container. The classical container consists of two vacuum-insulated vessels of silvered glass (Dewar flask) or metal, with the inner vessel containing the liquid helium immersed in a larger outer vessel filled with liquid nitrogen. Modern superinsulated Dewars are able to dispense with the liquid nitrogen.

The phase diagram of ^{4}He (**Fig. 1**) shows several remarkable characteristics. Helium remains a liquid down to absolute zero unless a pressure greater than 2.53 megapascals (25.0 atm or 367 lb/in.^{2}) is applied. A more subtle feature is a transition between two different liquid phases. This λ-transition is so named

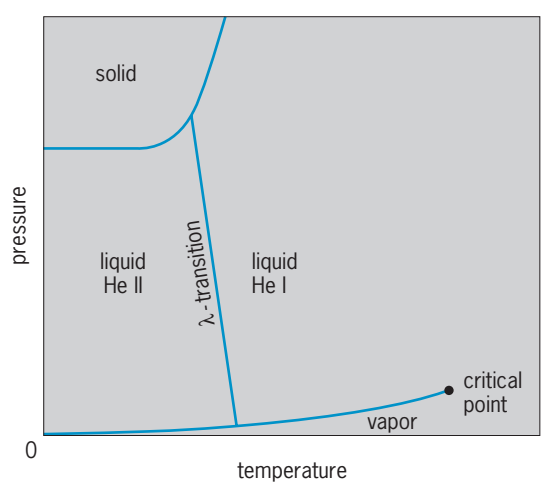

Fig. 1. Phase diagram for ^{4}He. The critical point is at $T_c = 5.20$ K (−450.3°F), $P_c = 229$ kPa (2.26 atm or 33.2 lb/in.^{2}).

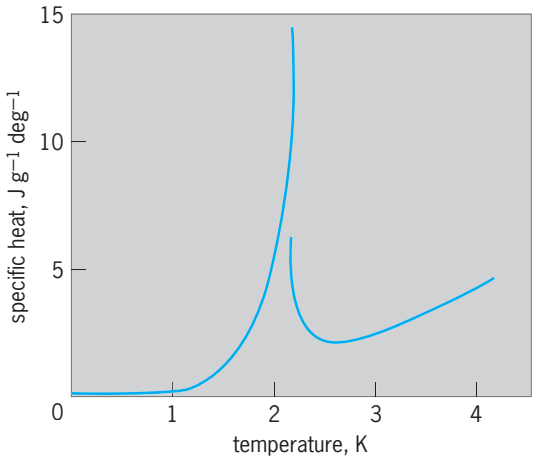

Fig. 2. Specific heat of liquid helium at saturated vapor pressure as a function of temperature °F = (K × 1.8) − 459.67. (*After K. Atkins, Liquid Helium, Cambridge University Press, 1959*)

because the specific heat (**Fig. 2**) has a singularity resembling the Greek letter lambda. There is no latent heat; such a transition is called second-order. The high-temperature liquid phase, called helium I, is a rather ordinary liquid. The λ-transition at 2.172 K or −455.76°F (at vapor pressure) marks the onset of superfluidity, which is the characteristic property of the low-temperature phase, helium II.

Liquid Helium II

When liquid helium is cooled by pumping, it boils vigorously until the λ-temperature is reached at a vapor pressure of 5.0 kilopascals (0.05 atm or 0.7 lb/in.^{2}). The λ-transition is manifest by a sudden cessation of bubbling, and the liquid becomes completely quiescent. Helium II has an enormous effective thermal conductivity, hundreds of times larger than that of the best metallic conductors. This explains the absence of bubbling, as evaporation need occur only at the surface; but it raises the problem of discovering the heat transport mechanism.

Superfluidity. Helium II flows thousands of times faster than He I at low pressure, and behaves as though turbulent. Thus, He II should be considered a superfluid state analogous to the superconducting state which appears in certain metals at low temperatures. *See* SUPERFLUIDITY.

When currents of helium are made to circulate around annular containers packed with fine powder or other porous material (provided that the initial speed of the helium is not too great), the helium continues to circulate for hours with no measurable loss of angular momentum. From these measurements an upper limit can be placed of the order of 10^{-19} cgs unit (1 cgs unit = 0.1 Pa · s) on the viscosity of the circulating superfluid at low speeds (about 10^{-15} that of air). *See* VISCOSITY.

However, a paradox exists because a different type of viscosity measurement, based on the drag torque on a cylinder rotating in the liquid, shows that the viscosity of He II is only about three times smaller than that of He I.

Thermomechanical effect. If two vessels partially filled with He II are connected by narrow channels and heat is supplied to one vessel, helium will flow into the warmer vessel and cause the level to rise against gravity. There is a converse mechanocaloric effect: if He II is forced by pressure through narrow channels, such as a porous plug, the upstream liquid becomes warmer and the downstream liquid becomes colder.

Ions in liquid helium. Electrons or positive ions injected into liquid helium form rather large complexes. An electron, which interacts repulsively with helium atoms, reduces its zero point energy by forming a cavity or bubble of 1.7-nm radius, within which it is confined. Thus the negative ion bubble displaces about 480 helium atoms. A He^+ ion attracts neutral helium atoms to form a "snowball" with an effective radius of 0.6 nm and an effective mass of about 44 helium atoms.

Critical velocities. Superfluid flow is frictionless only if the velocity is not too large. Above some "critical" velocity a pressure drop (or more precisely, a chemical potential difference) appears along the flow channel. The magnitude of the critical velocity may be less than 0.04 in./s (1 mm/s) in a wide channel (near 1-mm diameter) or as large as 70 ft/s (21 m/s) in a very short, narrow channel at very low temperatures. The value may depend on temperature; on approaching the λ-temperature the critical velocity decreases to zero. *See* CRITICAL PHENOMENA.

Film flow. Liquid helium was long thought to wet all solid surfaces, because a helium atom is more strongly attracted to other substances, which have molecules that are more polarizable than helium atoms. Consequently the surface of any solid in contact with liquid helium or its saturated vapor is covered, in thermodynamic equilibrium, by a helium film a few tens of nanometers thick. What is unique in the case of He II is that the film, like the bulk liquid, is superfluid and can flow as fast as about 1 ft/s (30 cm/s) without friction. It has been found that helium does not wet metallic cesium, and a cesium layer can interrupt superfluid film flow.

Thinner helium films are formed on solid surfaces exposed to helium vapor at less than the saturated vapor pressure. As the film becomes thinner, the onset of superfluidity is depressed progressively farther below the λ-temperature of the bulk liquid. Superfluid flow has been observed in films as thin as two atomic layers, at temperatures well below 1 K (−458°F). *See* PHASE TRANSITIONS.

Liquid Helium-3

The stable isotope ^{3}He occurs naturally as about 1 part per million (ppm) of atmospheric helium and 0.1 ppm of well helium. Commercial ^{3}He is obtained from beta decay of tritium.

Liquid ^{3}He-^{4}He mixtures. The addition of ^{3}He to liquid ^{4}He depresses the λ-transition to lower temperatures (**Fig. 3**). Below 0.87 K (−458.1°F) the

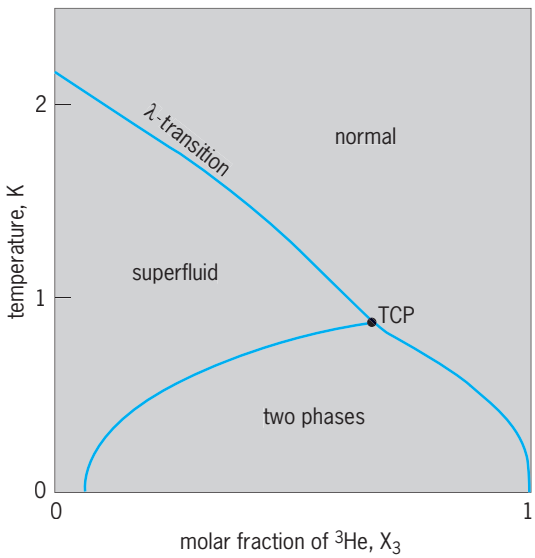

Fig. 3. Phase diagram for ^{3}He-^{4}He mixtures at saturated vapor pressure in the temperature-^{3}He concentration plane. The tricritical point (TCP) is at $X_3 = 0.67$, $T = 0.87$ K (−458.1°F). °F = (K × 1.8) − 459.67.

homogeneous solution is unstable and separates into a nonsuperfluid ^{3}He-rich phase floating on top of a superfluid ^{4}He-rich phase. In the superfluid phase, ^{3}He acts as part of the normal fluid.

Usually when a two-component system exhibits a liquid-liquid phase separation, the top of the phase-separation curve is rounded and the highest point is a critical point. By contrast in the helium isotope case, the two branches of the phase-separation curve meet in an angular contact with the λ-line. The junction, called a tricritical point, marks the change of the second-order λ-transition into a first-order phase-separation transition. Singularities in various thermodynamic functions at the tricritical point (in addition to those associated with the λ-line) have drawn considerable theoretical interest.

As the temperature is lowered toward absolute zero, the ^{3}He-rich phase rapidly becomes pure ^{3}He, but ^{3}He remains soluble up to a concentration of about 6% in the ^{4}He-rich phase. An insignificant number of phonons and rotons remain in the ^{4}He at temperatures well below 0.5 K (−458.8°F), and it is essentially in its ground state. This pure superfluid ^{4}He acts as an inert background for a gaslike solute of ^{3}He atoms, except that it increases the effective mass of the ^{3}He atoms and weakens the ^{3}He-^{3}He interaction. This has a very important practical consequence: It is possible to "evaporate" ^{3}He into superfluid ^{4}He with absorption of latent heat far below the temperature at which the true vapor pressure of ^{3}He becomes insignificant. Helium-3 dilution refrigerators based on this principle can attain temperatures less than 3 mK in continuous, recirculating operation.

Pure liquid helium-3. Liquid ^{3}He solidifies at absolute zero only at a pressure of 3.4 MPa (34 atm or 500 lb/in.^{2}). There is no λ-transition, because the exclusion principle prevents condensation of particles with half-integer spin such as ^{3}He atoms into a macroscopic quantum state, as occurs in superfluid ^{4}He. Instead, below about 0.5 K (−458.8°F) the liquid gradually becomes "degenerate" in the same sense as the electrons in a normal metal are degenerate. That is, a ^{3}He atom with each of the two allowed nuclear spin directions occupies each of the lowest momentum quantum states. Practically all the quantum states with energies up to the so-called Fermi level are occupied, and practically all above this energy are empty. The system is described in terms of the coordinates of weakly interacting quasiparticles, each moving in the mean field of the rest of the system. This Fermi liquid theory works well below about 100 mK. The mean field for nuclear spin interaction is such that the liquid is nearly ferromagnetic.

One consequence of the Fermi degeneracy is that the liquid has lower entropy than the solid, in which the nuclear spins are randomly oriented down to about 1 mK, where nuclear antiferromagnetic ordering occurs. In accord with the Clausius-Clapeyron equation of thermodynamics, the melting curve then has negative slope. This has an application known as Pomeranchuk cooling: if liquid ^{3}He at the solidification pressure in the region of negative slope is further compressed adiabatically, it begins to solidify, and this is accompanied by cooling. The amount of cooling can be very substantial if the initial temperature is sufficiently low. *See* FERMI-DIRAC STATISTICS; FREE-ELECTRON THEORY OF METALS; VAPOR PRESSURE.

Superfluid ^{3}He. The BCS theory explains how superconductivity occurs in certain metals, by a mechanism of electron pairing. Electrons of opposite spin and momentum become bound in pairs, of zero net spin, which condense into a macroscopic quantum state responsible for superconductivity. It became apparent that a similar pairing mechanism might result in a superfluid transition in liquid ^{3}He, which was studied at progressively lower temperatures, using an ^{3}He-dilution refrigerator and an additional cooling stage involving either Pomeranchuk cooling or adiabatic demagnetization. *See* ADIABATIC DEMAGNETIZATION; SUPERCONDUCTIVITY.

Sound propagation. Like ^{4}He-II, the superfluid phases of ^{3}He support first, second, third, and fourth sound waves. In the low-temperature limit, there are no collisions between the particles and the first sound mode is highly damped. However, a new mode called zero arises in the normal Fermi liquid because of the strong effective interactions between the ^{3}He quasiparticles. This mode is present in the superfluid phases also and, through its coupling to other excitations in the superfluid, demonstrates the triplet nature of the Cooper pairs in many dramatic ways.

Bellave S. Shivaram

Bibliography. K. H. Benneman and J. B. Ketterson (eds.), *The Physics of Solid and Liquid Helium*, 2 vols., 1976, 1978; K. Mendelssohn, *The Quest for Absolute Zero*, 2d ed., 1977; D. R. Tilley and J. Tilley, *Superfluidity and Superconductivity*, 3d ed., 1990; G. E. Volovik, *Exotic Properties of Superfluid Helium Three*, 1992; J. C. Wheatley, Helium three, *Phys. Today*, 29(2):32–42, 1976; J. Wilks and D. S. Betts, *An Introduction to Liquid Helium*, 2d ed., 1987.

Lissajous figures

Plane curves traced by a point which executes two independent harmonic motions in perpendicular direction, the frequencies of the motion being in the ratio of two integers. Such figures, produced on the screen of a cathode-ray tube, are widely used in frequency and phase measurements. *See* HARMONIC MOTION.

Methods of constructing. Mechanical, optical, and electronic methods may be used to construct Lissajous figures. Blackburn's pendulum consists of one ordinary pendulum hung from another with the two axes of rotation at right angles. The Lissajous figure is traced by the bob. An example of an optical method is afforded by a beam of light reflected successively from two small mirrors, each mounted on the prong of a separate tuning fork; the forks vibrate in perpendicular directions, and the image of the light source, focused on a screen, traces out the figure.

The cathode-ray oscilloscope furnishes the most important and practical means for the generation of the figures. The x-deflection plates of the tube are supplied with one alternating voltage, and the y-deflection plates with another. If the frequencies are incommensurable, the figure is not a closed curve and, except for very low frequencies, will appear as a patch of light because of the persistence of the screen. On the other hand, if the frequencies are commensurable, the figure is closed and strictly periodic; it is a true Lissajous figure, stationary on the screen and, if the persistence is sufficient, visible continuously as a complete pattern. *See* OSCILLOSCOPE.

Frequency measurements. Measurement of frequency in sinusoidal wave motion (for example, in alternating electric current) may be accomplished with an oscilloscope if one of the alternating voltages is derived from a variable-frequency calibrated oscillator. The standard is adjusted until a stationary figure results; if f_v is the frequency of the signal on the vertical deflection plates, and f_h that on the horizontal plates, then f_v/f_h equals the number of times a side of the figure is tangent to a vertical line divided by the number of times the top or bottom is tangent to a horizontal line. The shape of the figure depends also on the relative phase of the signals (**Fig. 1**), and in the case where one-half of the figure coincides with the other, the rule just stated does not hold. This situation is not troublesome in practice because unavoidable slow-frequency drifts of the signals prevent the figure from being exactly stationary; instead, the shape of the figure changes as though the relative phase of the signals were slowly changing.

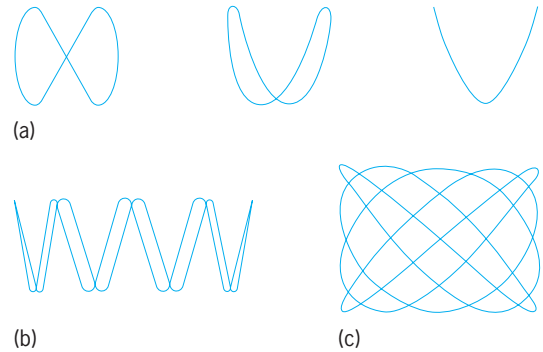

Fig. 1. Typical Lissajous figures for ratios of vertical frequency to horizontal frequency. (*a*) 2:1, with various phase relations. (*b*) 8:1. (*c*) 5:4. (*After F. E. Terman, Radio Engineers' Handbook, McGraw-Hill, 1943*)

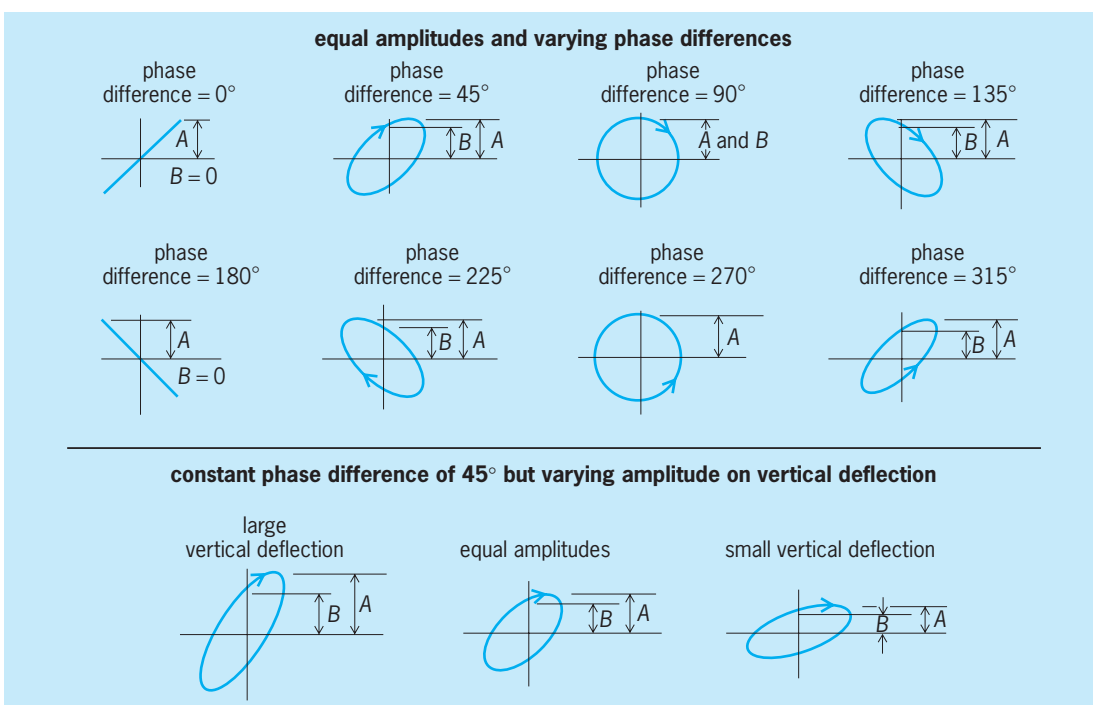

Fig. 2. Typical Lissajous figures of 1:1 frequency ratio. (*After F. E. Terman, Radio Engineers' Handbook, McGraw-Hill, 1943*)

In some cases, the standard is of fixed frequency and the unknown is nearly equal to the standard (or to some multiple or submultiple thereof). The figure will drift slowly through the various aspects it can assume for the nominal ratio involved. The actual ratio can be obtained from the nominal ratio, and a correction determined by the rate and direction of drift of the figure. This method is particularly valuable for the intercomparison of standards.

Phase measurements. These usually involve two signals of the same frequency. In this case, the Lissajous figure is an ellipse of which the shape and orientation depend on the relative phase and relative amplitudes of the signals, as shown in **Fig. 2**. The relative phase θ is given by the equation below,

$$\sin\theta = \pm\frac{B}{A}$$

in which A is the maximum half-height and B is the intercept on the y axis. Figure 2 shows how the quadrant in which θ lies is determined by the orientation of the major axis of the ellipse and the direction of motion of the spot. If the latter is unknown, the resulting ambiguity amounts to that in the sign of $\sin\theta$, once it is decided which signal is the reference. The ambiguity can be resolved by shifting the phase of either signal in a known direction and noting whether the phase shift indicated by the figure increases or decreases.

The most precise comparison can be made when the ellipse degenerates into a straight line. This condition is therefore used as a null in measurements made with a calibrated phase shifter. The unknown phase shift disturbs the null, and the calibrated phase shifter restores it.

More complicated figures are employed in the calibration of phase shifters. A signal at the desired frequency is connected through the phase shifter to the y plates of an oscilloscope and also through a frequency multiplier to the x plates. If the frequency multiplication is 8, for example, the Lissajous figure will appear as in Fig. 1*b*. The phase shifter is then successively set to the various positions for which one-half of the figure coincides with the other; these positions will correspond to phase differences of, in this case, $\pi/8$. *See* PHASE-ANGLE MEASUREMENT.

Martin Greenspan

Bibliography. A. D. Helfrick and W. D. Cooper, *Modern Electronic Instrumentation and Measurement Techniques*, 1990; J. D. Lenk, *Handbook of Oscilloscopes*, 1982; Lord Rayleigh, *Theory of Sound*, vol. 1, reprint 1945.

Lissamphibia

The subclass of Amphibia including all living amphibians (frogs, toads, salamanders, and apodans). The other two subclasses are the Labyrinthodontia and the Lepospondyli, predominantly Paleozoic fossil groups lacking recent representatives. *See* LABYRINTHODONTIA; LEPOSPONDYLI.

Morphology. Living amphibians are grouped together by possession of a unique series of characters, the most important of which are (1) pedicellate teeth; (2) an operculum-plectrum complex of the middle ear; (3) the papillar amphibiorum, a special sensory area of the inner ear; (4) green rods in the retina of the eye; (5) similar skin glands; and (6) a highly vascular skin used in respiration (cutaneous respiration). Only the first two characters can be used to separate the Lissamphibia from the other two subclasses, because the other characters are soft anatomy features not ordinarily preservable as fossils. *See* EAR (VERTEBRATE); EYE (VERTEBRATE); RESPIRATION.

Dentition. The pedicellate condition of the lissamphibian tooth is formed by a separation of the tooth into two segments, a crown and a pedicel (**Fig. 1**). The pedicel is usually fused to the jaw, but crown and pedicel are separated by uncalcified or fibrous material, which in effect functions as a hinge in the tooth. In at least some lissamphibians the tooth bends inward more easily than outward. This indicates that it may be an adaptation for allowing prey to enter the mouth easily, but tends to prevent its escape. Some teleost fishes also have such a hinged tooth that bends inward toward the throat and is relatively rigid when bent back; however, there the hinge is between tooth and jaw, without an intervening pedicel, and is not comparable with the

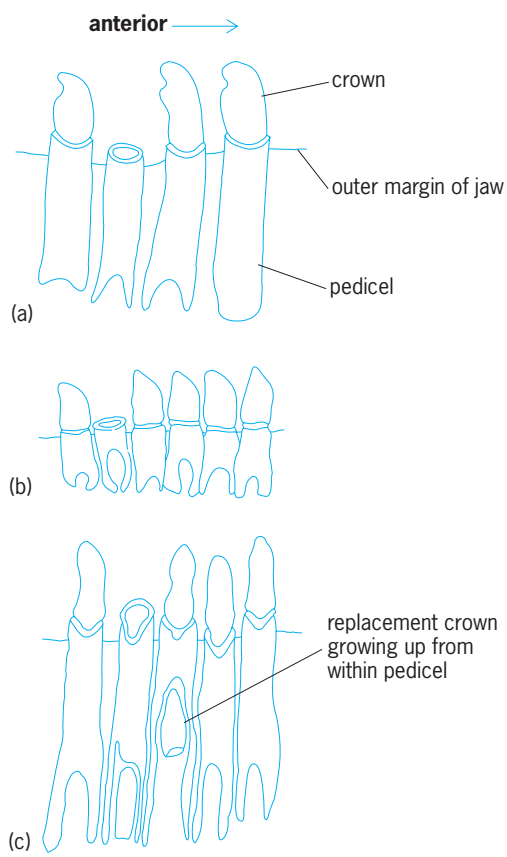

Fig. 1. Pedicellate teeth in Lissamphibia. (*a*) Salamander. (*b*) Apodan. (*c*) Frog. Teeth are from the lower jaw and are viewed from inside the mouth. (*After T. Parsons and E. Williams, J. Morphol., 110:375–390, 1962*)

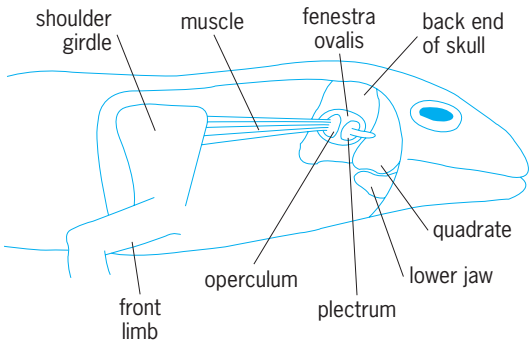

Fig. 2. Operculum-plectrum complex of a primitive salamander (skin removed to show internal structures).

lissamphibian condition except perhaps in function. *See* TOOTH.

Sound perception. The operculum-plectrum complex of the middle ear is an unusual feature unknown in other tetrapods. In the characteristic tetrapod condition a single sound-transmitting bone, the stapes, fills the opening (fenestra ovalis) of the middle ear and connects to the inner ear. In lissamphibians two structures occur in the fenestra ovalis: plectrum (stapes) and operculum (**Fig. 2**). The plectrum has the usual tetrapod connections to tympanic membrane or to quadrate. The operculum is connected to the shoulder girdle by a specialized shoulder muscle that may provide passage to the ear for vibrations traveling through the front limb from the substrate. Thus the animal may receive both airborne and ground-borne sounds.

Phylogeny and fossil record. Apodans lack a fossil record; frogs and salamanders go back to the Jurassic. There is a Triassic fossil from Madagascar that may be near the ancestry of frogs, but it does not indicate the group of origin. Except for this Triassic fossil, the ancient amphibians and the lissamphibians do not overlap in their geological ranges. On the basis of vertebral similarities (solid, spool-shaped centrum), lissamphibians have often been thought to be lepospondyl descendants, but in the skull (especially the palate) there are similarities to temnospondyl labyrinthodonts.

Pedicellate teeth and the operculum-plectrum complex are characters that form the best evidence for close relationship of modern amphibians to each other. Pedicellate teeth and the possible presence of an operculum have been found in one family of primitive rhachitome labyrinthodonts. These two occurrences may indicate that lissamphibians are derived from the latter group, although new fossil finds are necessary to clarify lissamphibian ancestry. *See* AMPHIBIA; ANURA; APODA; URODELA.

Richard Estes

Bibliography. E. Jarvik, *Basic Structure and Evolution of the Vertebrates*, 2 vols., 1981; T. Monath, The opercular apparatus of salamanders, *J. Morphol.*, 116:149–170, 1965; T. Parsons and E. Williams, The relationships of the modern Amphibia, *Quart. Rev. Biol.*, 38:26–53, 1963; A. Romer, *Vertebrate Paleontology*, 3d ed., 1966.

Listeriosis

An infectious disease of humans and animals caused by the bacteria *Listeria monocytogenes* and *L. ivanovii*. Both humans and animals can be carriers, which excrete the bacterium in feces. Sheep, goats, and cattle can excrete the bacteria in milk, without clinical symptoms of mastitis. The most important pathway of infection is probably through food. *Listeria monocytogenes* has frequently been isolated from grass silage (fermented fodder), especially from silage of inferior quality, which is an important source of infection in ruminants. Relatively few animals develop clinical disease, but a high proportion can be latent carriers. Farm leys often are fertilized with manure, and the bacteria in infected manure may be distributed on plants and in soil. During forage harvesting, soil and dung may be brought into the silo together with the grass, thus introducing the bacteria into the silage. *Listeria monocytogenes* is able to multiply in silage with a high pH, which probably is the main source of infection in ruminants.

Humans most likely ingest the bacteria with contaminated food, such as meat and meat products, raw milk and milk products, and unwashed vegetables. *See* FOOD MICROBIOLOGY.

Virulence and pathogenesis. Virulence depends upon attachment of the bacterium to a tissue cell, entrance into and multiplication in the cell, resistance to host defense, and damage to host mechanisms. Cells of *L. monocytogenes* are hydrophobic, which can be an attachment factor. They produce phospholipases which can disrupt cell surfaces, and catalase and superoxide dismutase to inactivate lethal enzymes in the macrophages. They also produce a hemolysin which disrupts the phagolysosomal membrane and releases the bacterium into the cytoplasm, where it can grow and multiply.

The *L. monocytogenes* organism usually enters the body by way of the alimentary route. The number of bacteria is substantially reduced after passing through the gastric acid, but bacteria that reach the intestine may attach to a portion of the epithelium. They rapidly infect the lymphoid tissue and may extend to other internal organs.

In ruminants, another mechanism seems more important. The bacteria pass the buccal mucosa and reach one or several of the brain nerves. They proceed to the brain through an intracellular route, and during this process they are shielded from the host's defense mechanisms. After about 2 weeks, they reach the brainstem and cause lesions within a fairly local area, with damage of neurons.

Clinical symptoms. Encephalitis is the most common form of disease in ruminants, and septicemia with involvement of several organs, including the pregnant uterus, occurs most commonly in monogastric animals, including very young sheep, goats, and calves.

Symptoms of encephalitis are dullness, and facial paralysis with dropping ear and lowered eyelid on the affected side. The head is often tilted to one side,

with slight paralysis of the tongue and loss of ability to salivate or swallow. The animal may walk in circles, become recumbent after a while and, if not recovering, pass into coma and die. About 40% of sheep and goats will recover if treated with antibiotics. The prognosis is better for cattle.

In cases of septicemia the animal looks depressed, with reduced appetite, high fever, and often diarrhea. Pregnant animals may abort, especially in the last trimester. They usually respond well to treatment with antibiotics.

Most human cases are sporadic, but food-borne epidemics occur. Abortions, perinatal disease, and disease in immunosuppressed individuals are most common. Perinatal disease is dominated by septicemia, widespread microscopic abscesses, and meningitis. In adults, meningitis is by far the most common manifestation, but in immunosuppressed individuals encephalitis occurs, possibly with a pathogenesis similar to encephalitis in ruminants.

Immunity. Cell-mediated immunity is the main defense against *L. monocytogenes*, and activated macrophages are the main effector cells. Neutrophils play an important part in defense before cell-mediated immunity has been developed. They destroy infected nonphagocytic cells in various organs, leaving the bacteria open for inactivation by other defense mechanisms. Immunity can be strengthened by use of a vaccine that is available for sheep in several countries. *See* IMMUNITY.

Diagnosis. Clinical diagnosis is based on symptoms, on isolation of *L. monocytogenes*, and on histopathological examination of affected tissue, especially brain tissue. Food samples are examined for the bacteria. The bacteria can be further identified in biochemical tests, by sero- and phagetyping, and by techniques such as gene probes, deoxyribonucleic acid (DNA) fingerprinting, isoenzyme typing, and ribosomal ribonucleic acid (rRNA) typing. *See* MEDICAL BACTERIOLOGY. Hallstein Gronstol

Bibliography. J. M. Farber and P. I. Peterkin, *Listeria monocytogenes*: A food-borne pathogen, *Microbiol. Rev.*, 55:476–511, 1991; A. L. Miller, J. L. Smith, and G. A. Somkuti (eds.), *Foodborne Listeriosis*, 1990.

Literature of science and technology

The accumulated body of scientific-technical writings published to serve the informational needs of, primarily, scientists, engineers, and research workers, and, to the extent that it can be understood, the general public. This vast literature is worldwide in origin, international in language, diverse in subject content, varied in form, uneven in quality, and expensive. An ever-increasing portion of it is being produced, stored, and retrieved or published electronically, and is being made available on film, tape, CD-ROM or magnetic disk, or through remote terminals as well as in print-on-paper format. *See* COMPACT DISK; COMPUTER STORAGE TECHNOLOGY.

The primary sources of scientific-technical information are the first (and often the only) published records of original research and development, and accounts of new applications of science and engineering to technology and industry. These unorganized and usually not closely related contributions appear almost exclusively in periodicals, topical compilations such as Festschriften, and independently published proceedings of conferences, separate research reports or monographs, patents, standards, dissertations, or manufacturers' literature.

Organized works and compilations that derive from or refer to the primary-source literature make up the secondary sources of scientific-technical information: the handbooks, encyclopedias, treatises, bibliographies, reviews, abstracting and indexing serials, machine-readable bibliographic databases and databanks, other reference works, and, of course, all translations.

Guides to the literature, directories (of persons, organizations or products, for example), and textbooks are often looked upon as tertiary sources. Popularizations, biographies, histories, and similar relatively nontechnical publications are thought of as being about science rather than of science.

Primary Sources

These are the records of original research, patents, standards, and manufacturers' literature.

Periodicals. These publications traditionally make up the bulk of the primary-source literature of science and technology. Periodicals include journals, bulletins, transactions, proceedings, or other serial publications that appear regularly and continuously in numbered sequence, but not newspapers or annuals. More than 20,000 scientific and technical periodicals are published worldwide.

At least 150 countries regularly publish research journals in science-technology, and these may appear in many different languages. The growth of English as the major language of science has not displaced the considerable body of material in other languages, but the importance of language varies from subject to subject, and the importance of subject varies from country to country.

Current issues of some periodicals are available on black-and-white or colored microfiche as well as in print-on-paper format, and complete volumes of many periodicals are available on microfilm for space-saving storage. Some periodicals are available online in electronic form, although most have printed counterparts. Many authors are still reluctant to publish in nonprint formats alone, but electronic peer-reviewed journals with graphics and searchable text are an alternative. *See* PHOTOCOPYING PROCESSES; PRINTING.

The contents of periodicals vary considerably in the kind of material included as well as in the technical level. Professional, scientific, and technical societies, for example, tend to emphasize basic research and the more technical aspects to a subject, while industrial and trade associations lean toward the practical, personal, and popular side. Societies, too, almost uniformly make use of highly qualified subject experts, called referees, to evaluate critically all

original contributions before they are published, thus imparting reliability, authoritativeness, and prestige to their publications. Other organizations, such as research institutes, university experiment stations, and government agencies, publish periodicals comparable with those published by the societies, but such organizations are more likely to emphasize their own work. There are, of course, noteworthy exceptions.

Conference papers. Papers presented at national and international scientific conferences, symposia, colloquia, seminars, and other technical meetings often are important contributions to the primary-source literature of science and technology but are also often uneven and elusive. At widely varying intervals following their presentation they may be published as complete (or partial) collections by the convening bodies themselves, by the parent scientific societies, or by private publishers; they may appear as articles in, or supplements to, regular journals; or they may only be noted or at most abstracted. Audiotapes are sometimes available as supplements to, or substitutes for, printed proceedings. Sometimes papers are preprinted and not later published at all, thus being made available only to those who attend the meetings or who write for them in time. Announcements of forthcoming scientific meetings as well as of resulting technical publications are now published regularly. *World Meetings*, for example, is a 2-year registry of such meetings, and *Index to Scientific and Technical Proceedings* is one of several indexes to the resulting publications.

Research monographs. These are separately published reports on original research that are too long, too specialized, or otherwise unsuitable for publication in one of the standard journals. Each monograph is self-contained, frequently summarizes existing theory or practice before presenting the author's original and previously unpublished work, and is likely to be one of a series in the same field. Research monographs are also being published on microfiche with paper copies available on demand, a practice which may escalate as costs of print-on-paper publishing continue to rise.

Research reports. Reports from research and development projects make up another important part of the primary-source literature of science and technology. Reports are a more primitive form of literature because they are produced earlier in the research program, often as progress reports. While a progress report is only a temporary reference, it is important because it may include negative results and incidental data which may be missing from a final report.

Research reports originate in all research organizations doing work under contract to the United States government and are in turn distributed to all installations (including designated depository libraries) that have an interest in them. Classified reports, those containing data having some bearing on national security, are distributed only to properly cleared individuals or organizations. Since distribution of even the unclassified reports is somewhat limited, few of them are indexed, abstracted, or even listed by the standard services which cover other parts of the primary literature.

The National Technical Information Service (NTIS) of the U.S. Department of Commerce (Springfield, Virginia) is the agency responsible for the announcement, sales, and distribution of unclassified research, development, and engineering reports, translations, and other analyses prepared by United States and foreign government agencies, by their contractors or grantees, or by special technology groups, including many information analysis centers. NTIS also is the source for federally generated machine-processable data files and software. Since 1981, it has had the additional responsibility of operating an indexing service covering ongoing federally sponsored research projects, a service previously provided by the Smithsonian Science Information Exchange (SSIE). Established in 1970, NTIS superseded the Clearinghouse for Federal Scientific and Technical Information.

Summaries of unclassified government research reports and of available translations of foreign research reports are published in the comprehensive biweekly *Government Reports Announcements & Index (GRA&I)* and in the 27 *NTIS Alerts*, each covering a different subject area. Copies of the full reports are available from NTIS in paper copy or microform, and the NTIS database can be searched online through major vendors. Other more specific current awareness services are also available.

While the *GRA&I* and the 27 *NTIS Alerts* cover most government unclassified research reports, several government agencies also publish their own more specialized announcements and abstracts. Examples include *Scientific and Technical Aerospace Reports (STAR)*, published by the National Aeronautics and Space Administration (NASA), and *Energy Research Abstracts (ERA)*, published by the U.S. Department of Energy.

Preprints. Preprints are research-in-progress reports to the small but often international group of fellow scientists working in the same relatively narrow research specialty, somewhat after the fashion of seventeenth-century "invisible colleges." Like government research reports, the preprint is a primitive form of technical literature, unrefereed and unannounced but historical. In the past, preprints were far more elusive than conference papers since only those on particular "preprint circuits" knew of their existence, even though they may have been cited as references (for example, "in preparation"). However, now that many scientists are conferring with each other via computer conferencing and electronic mail, paper preprints are becoming obsolete except for possible record purposes. Preprint databases are being developed with abstracts distributed via electronic mail to announce the availability of new titles and the full text available for downloading via a computer network. To scientists on the research frontiers, paper preprints or preprint electronic reports are vital (and quite conventional); to others they may be stimulating or perhaps just interesting. Yet even mere identification of specific areas of ongoing research can help research and

development managers and policy makers to avoid unwarranted duplication of research effort and expense as well as to suggest other areas where research could be more profitable. *See* ELECTRONIC MAIL.

Patents. Patents are considered part of the primary-source literature, since an invention must be new in order to be patented and more often than not there will be no other published description either of the idea or of its application. Patents are sources of information about new products, processes, plants, microorganisms, and chemical compositions. Announcement publications, abstracts journals, and the individual patents are available from the various national patent offices and in selected depository libraries, but extensive patent searches require knowledge of both the subject field and patent law, as well as much time. There are more than 130 patenting authorities in the world. *See* PATENT.

Standards. Often a standard is a pamphlet describing acceptable levels of dimension, quality, performance, or other attributes of materials, products, or processes. Standards are established by general agreement among representatives of concerned groups, such as manufacturers and consumers, and may be issued by companies, trade associations, technical societies, government agencies, and national or international organizations. Technical societies also often issue statements of recommended practice as followed by competent practitioners.

Dissertations. Doctoral dissertations and master's theses are reports of original research. Although the research may later be reported in one or more periodical articles, dissertations and theses may still be consulted for their comprehensive literature reviews and for details not included in subsequent publications.

Festschriften. Compilations of scholarly papers memorializing a distinguished scientist, scientific society, or important scientific event often contain accounts or results of original research, the significance of which may well be comparable with that of the person, society, or event being commemorated. Each paper usually is related to some aspect of the subject's work or professional interests.

Manufacturers' literature. These publications are often the only source of specific information about particular products or their development; if the information is not published elsewhere, this source becomes a primary source. Technical service bulletins, data sheets, price lists, catalogs, and other specialized publications describing a manufacturer's equipment or processes are typical examples.

Secondary Sources

Reference works, which make up the bulk of the secondary literature, are of three types:

1. Those which in effect index selected portions of the primary literature and which thus aid in finding what has been published on a given subject (and where) generally or specifically, recently or retrospectively; for example, indexes themselves, bibliographies, indexing serials, and under certain conditions, abstracting serials, all of which, if computer-compiled, may become machine-readable bibliographic databases.

2. Those which survey selected portions of the primary literature and which thus aid in acquiring state-of-the-art, recent background, or comprehensive and definitive information on a given subject; for example, reviews, treatises, monographs, and again under certain conditions, abstracting serials.

3. Those which themselves contain the desired information (but rather concisely presented), such as established facts, formulas, procedures, equations, meanings, theories, history, and biography. All this information is compiled selectively from the primary literature and arranged in some definite manner for convenience in finding; for example, dictionaries, encyclopedias, handbooks, and compilations of tables. An increasing amount of the statistical, numerical, and structural data found in these works is also available from machine-readable data banks.

4. Technical translations are another important (and expensive) part of the secondary literature of science and technology. For all practical purposes these take on the characteristics of the particular primary-, secondary-, or tertiary-source literature that was translated, and they are often treated as if they were in fact the original publications. Translations, of course, are not new to science and technology, but since 1957 a crash program to disseminate scientific-technical information has encouraged many more government, society, academic, and private organizations throughout the world to produce English translations of works published originally in Russian, Chinese, Japanese, and other languages; translation from English into these languages is also occurring.

Some 500 periodicals translated into English are now available in cover-to-cover, selected articles, or abstract form, and additional thousands of separate articles, monographs, and other works translated into English are available from commercial sources, from NTIS, from the National Translations Center (U.S. Library of Congress), and from similar agencies in Canada and elsewhere. Many authors of research papers still prefer to cite original sources, so translations tend to serve largely as a kind of monitoring system with an inherent time delay for non-Western science and technology. *World Translations Index*, published by the International Translations Centre, announces translations in all fields of science and technology.

Index types. The following subsections discuss the various indexing services.

Indexes. An index is a detailed alphabetical list of the names, terms, topics, places, formulas, numbers, or other significant items in a completed work (such as a book, set, or bound journal) with exact page references to material discussed in that work. It may be part of the work or it may be published separately. Rather than alphabetically, it may be arranged chronologically, geographically, numerically, or in some other fashion. No reference work except, of course, a dictionary and possibly an

encyclopedia is any better than its index; a properly compiled index is the information hunter's most useful tool.

Bibliographies. A bibliography is a list of references to primary or other sources relating to some given subject or person; it is usually arranged chronologically or alphabetically by authors. It may be comprehensive or selective and annotated or otherwise evaluative. It may be published separately or as a part of a larger work to acknowledge information sources, recommend additional reading, or direct attention to more detailed treatment elsewhere. Almost always the time span and the subject coverage are well defined, so that a good bibliography represents a definitive coverage of the literature within the defined limits and makes retrospective literature searching that much easier.

Indexing serials. An indexing serial, sometimes called a current bibliography, is a regularly issued compilation of titles of articles (or of subsequent references to them) that appear in current primary-source journals; titles of new books and of other separately published material are often also included. In each issue the titles usually are listed alphabetically by subject; but they may be listed by both subject and author in one alphabet, or the authors may be indexed separately. Issues are published with a specific frequency, weekly or monthly, for example, and they may cumulate. Some indexing serials aim to cover only selected journals in their fields but to do it comprehensively, listing all articles of any possible interest; others aim to cover all journals in their fields but to do it selectively, listing only articles of real substantive interest. Current issues of typical indexing serials aid in finding recent articles, while bound sets are helpful in searching older literature, especially that published before the advent of machine-readable bibliographic databases in the 1970s.

Express services (such as *Current Contents*), which reproduce, often translate, and index the tables of contents of selected current primary-source journals and then distribute the usually weekly compilations by air mail or computer networks, are important current awareness aids as well as pertinent leads to reprint sources.

Most indexing serials are now being compiled and printed with the aid of computers, and the tapes become available for custom literature searches by the producer or for online searches through vendors of information services which lease them. In some cases the tapes may also be purchased.

Abstracting serials. An abstracting serial, also at times called a current bibliography, is a regularly issued compilation of concise summaries of (1) significant articles (often in a very limited subject field) that appear in current primary-source journals and (2) important new research monographs, reports, patents, and other primary-source publications in that field. In each issue the summaries usually are arranged by rather broad subjects followed by an author index, but some abstracting serials also include report number, corporate author, and more exact subject indexes as well. Most abstracting serials aim at selective substantive coverage of all known and available journals in their fields. Issues may be published with almost any specific frequency, but usually they do not cumulate. Instead, exhaustive annual or multiple-year indexes according to author, subject, patent, formula, or report number are published. Most abstracting serials are now being compiled and printed with the aid of computers, and the tapes become available for custom or online searches or for purchase. Long runs of many abstracting serials are available in microform.

Abstracts, however published, are likely to be further behind the current primary literature than title compilations are, but they also cover much more primary literature and give much more subject information per reference. Serious research workers find lengthy runs of abstracts, properly indexed, particularly suitable for in-depth searching of the older subject literature, especially that published before the advent of the applicable subject machine-readable bibliographic database. They find the current issues to be particularly suitable for keeping abreast of worldwide scientific progress in their chosen fields. Abstracts thus serve both as indexes to specific information and as surveys of particular subject areas.

Bibliographic databases. A bibliographic database is essentially a current bibliography or an indexing or abstracting serial in machine-readable form. It may be used to print paper copies in journal format, or it may be used to provide online (user-interactive) or batch (custom once-only) search service by the producer or by an information service vendor. Bibliographic databases are also distributed on CD-ROM, searchable by using a microcomputer. Some databases cover the literature of a subject, while others cover particular types of primary literature (dissertations, standards, patents, conference papers). References may be retrieved online or from CD-ROMs by using sophisticated search strategies, including terms from title and abstract as well as index terms. *See* DATABASE MANAGEMENT SYSTEM.

Survey types. The following subsections discuss the various survey types of reference works.

Reviews. A review is a subject survey of the primary literature usually covering the material that was published since the most recent review on the same subject. A good review accumulates, digests, and correlates the current literature and indicates the direction future research might take. A critical or evaluative review by a competent author is an important aid in keeping up with the work being done in a particular field and in spotting the outstanding developments. A review may appear as one of a collection of more or less related papers in an annual, a quarterly, or a monthly series, or it may appear as an article in a regularly issued primary-source subject journal.

Review reference lists are often extensive and, in a way, form a definitive bibliography of the subject for the period covered. One shortcoming of a review, though, is the inevitable delay before comprehensive and critical coverage of the primary literature can be obtained; one hindrance in getting a good

review is the difficulty in finding a qualified specialist with both the time and the literature-searching competence to do the job. Nevertheless, information experts believe that reviews are and will continue to be an important part of the literature of science and technology.

Treatises. A treatise is a comprehensive, authoritative, systematic, well-documented compilation or summary of known information on a subject, covering it so completely that the work becomes definitive at the time of writing. Thus a treatise is an aid in acquiring a foundation in a subject adequate to enable a trained individual to carry on advanced research; at the same time, it is a source of facts, procedures, theories, and other important data, presented in such a way as to show their development, correlation, and probable reliability. A treatise combines characteristics of an advanced textbook and a handbook; in fact, certain multivolume German *Handbücher* and certain multivolume British textbooks are good examples of treatises. One drawback to a treatise is that it is soon out of date. By the time the last part of a multipart treatise is published, the first parts are already quite old. Here, too, the time lag is unavoidable because of the scholarly inventory nature of the work itself. Periodic supplements and occasional new editions of the whole or of consecutive parts help overcome this lack of recency of the basic work, and so long as the dates of publication and the probable value of the original sources are clearly brought out, the usefulness of a treatise will not be impaired.

Monographs. A monograph, in effect, is a short treatise on a particular subject, a single division or part of a larger branch of knowledge. Hence it can be more up to date when it is published and can be brought up to date more easily for a revised edition; it frequently is published as part of a series. Almost any book on a special topic may be called a monograph unless obviously it is something else, for example, a research monograph, part of the primary literature already discussed, or a textbook, discussed later.

Reference tools. Five major types of reference tools are encyclopedias, dictionaries, handbooks, critical tables, and machine-readable data banks.

Encyclopedias. An encyclopedia is an idea book. It deals with concepts and is usually arranged alphabetically by separate subjects, those parts of all knowledge or of a branch of knowledge which are considered to be significant by the compilers. Treatment is as thorough as space will permit and may be descriptive, explanatory, statistical, historical, or evaluative, but it should be impartial and should summarize the information thought to be of the most general value at the time of writing. Illustrations, tables, drawings, and photographs are often used, lavishly on occasion. Source references may or may not be given, so authoritativeness depends largely on the reputations of the publisher and of the authors of the individual articles. The large-subject approach usual in a multivolume set requires a competent index, while the small-subject approach usual in a one-volume work does not; however, if the subject arrangement is not alphabetic, an index is indispensable.

An encyclopedia is designed to give a summary of background knowledge and a discussion of concepts in a particular field, useful for orientation in a strange subject area and for determining key words for more specific searching. One-volume encyclopedias give something less than this and tend to be considered and used as expanded dictionaries.

Dictionaries. A dictionary is a word book. In science and technology its purpose is to define commonly used terms as simply as possible, preferably without much recourse to other technical terms and, in addition, with some indication of the specific subject field in which a given definition is the accepted one. Etymology and pronunciation of terms may also be included. Terms may be illustrated to show shape or use, and the treatment of terms may be expanded beyond simple definitions to include information bordering on encyclopedic or handbook coverage. A good technical dictionary always gives the generally accepted correct spellings of the words it defines, and it often notes variants. For acronyms and abbreviations, the fully spelled out terms are given. Bilingual and polylingual dictionaries normally give just the direct equivalent word in the other language with no definition other than identifying the subject field in which it is used.

Handbooks. A handbook (or manual) in science and engineering usage is a compact, fairly up-to-date, relatively complete, authoritative compilation of specific data, procedures, and professional principles of a subject field. Much of the information is given in tables, graphs, and diagrams, and illustrations are freely used. Symbols, equations, formulas, abbreviations, and concise technical language all help condense much practical information into a handbook, but they also require that the reader already have rather broad knowledge of the field in order to use the work effectively. Hallmarks of a good handbook are an exhaustive index, up-to-date references, expert editorial staff, easy-to-read printing, and convenient format. The term handbook implies ready reference in a technical subject field. The term manual in common usage implies instruction in how to do something with the aid of very explicit step-by-step directions.

Critical tables. A collection of standard reference data (for example, physical and chemical properties of materials) that has been compiled, documented, evaluated, published, and kept up-to-date by reputable and responsible specialists is an essential reference tool to the working scientist or engineer. The compilation of a directory of critical (that is, critically evaluated) tables of standard reference data from worldwide sources is the responsibility of the World Data Referral Centre in Paris, established by the international Committee on Data for Science and Technology (CODATA). The National Standard Reference Data System of the National Institute of Standards and Technology is the United States affiliate.

Data banks. A data bank is a compilation of numerical, statistical, or structural (for example, chemical)

data in machine-readable form. Advances in instrumentation in all fields of science and technology have led to dramatic increases in the amount of data available, from gene sequences to space-based Earth measurements. When data banks are accessed online, numerical data may be retrieved and manipulated using available computer programs.

Tertiary Sources

These sources include textbooks, directories, and literature guides.

Textbooks. A textbook is a standard work used for instruction, arranged so as to develop an understanding of a branch of knowledge rather than to impart information for its own sake. Textbooks characteristically are graded in difficulty and are, therefore, suitable for either introductory or advanced study at almost any level, depending on the background of the reader. The most elementary textbooks assume no prior subject knowledge or experience at all; they work with and build on long-established (and sometimes already superseded) theories and concepts and proceed logically to newer and more advanced ideas as the student learns to master the material. The most advanced textbooks, on the other hand, take on the aspects of monographs or treatises in the thorough coverage, systematic development, and rigorous treatment of their subjects.

Directories. A directory is an alphabetical listing of at least the names and addresses of persons, organizations, manufacturers, or periodicals belonging to a particular class or group. The directory may include indexes or supplementary lists by subject field, geographical location, product, or some other desirable classification. It may be published separately in one or more volumes, or it may appear as a part in, an issue of, or a supplement to another work such as a journal or the proceedings of a society; it may be revised from time to time or not at all.

Biographical directories vary in coverage from annual membership lists with only the briefest information about the persons listed to the specialized compendiums of the *Who's Who* type. Trade and product directories, often appearing as annual buyers' guide issues of certain industrial and trade periodicals, frequently contain an unsuspected variety of information, such as sources, producers, dealers, properties, hazards, shipping instructions, trade names, and trademarks. Industrial directories—city, state, national, and international—often contain considerable data on manufacturers and their officials, subsidiaries, plants, dealers, capitalization, trade names, and products.

Organization directories of learned societies, professional associations, research establishments, trade groups, educational institutions, and government agencies, for example, usually give at least the officers, the address, the functions, and the date of establishment of each organization listed; coverage varies from community-size to worldwide. Periodicals directories include full bibliographical information, frequency, publisher, and price of each title listed; some also include index coverage, advertising arrangements, and availability in libraries. The more complete periodicals directories and lists are often considered to be national or subject bibliographies rather than directories.

Literature guides. A literature guide is a reference manual designed to aid research workers or other prospective users in finding their way into and through the literature of a specific subject field. Its purpose is to acquaint its readers with the important types of information sources available to satisfy their needs and to help them make fruitful current or retrospective literature searches. Library resources and services are discussed, and all important specific and related titles of works in the field are cited and many are described quite fully. Guides also exist, in both print and audiovisual forms, to assist users of specific reference tools, searching aids, or information services.

Information Science

Collecting, organizing, and providing easy access to the literature are generally performed by libraries with the aid of the tools produced by professional, governmental, academic, trade, and other members of the international information industry.

Role of libraries. The library has long been the institution most concerned with the methods, skills, and systems for the acquisition, storage, preservation, retrieval, and use of literature. The depth of a library's concern has always been determined primarily by the needs of those it serves and by the fields of literature it handles.

A subject-specialized research library acquires, classifies, catalogs, stores, and supplies both whole works and bits of information as precisely as possible, combining library practice with the techniques of documentation or information science.

But science-technology literature is vast, polylingual, expensive, and only imperfectly distributed or indexed (although better than most other subject literatures). No large research library can have everything published in even a relatively narrow subject field, to say nothing of the broader subject disciplines.

To the serious research worker, to industry, and to science itself, this present inability of libraries (however large, rich, dedicated, or specialized) to cope with the literature of science and technology is a very critical matter. In an attempt at solving this problem, local, regional, national, and international networks have been developed among libraries of various types, with members associated electronically in order to share resources, services, or internal technical procedures. Thus anyone in need of specialized information can get help from the total resources of a particular network or, if need be, any other network, by inquiring first at the local public, academic, or special library.

Role of information technology. Of even greater help to the serious research worker is the rapid development of electronic information and telecommunication systems combined with improvements in input and output devices and storage media. The *Gale*

Directory of Databases lists over 2200 databases of all types in science and technology. In addition to the bibliographic databases and data banks listed as separate categories above, many other databases are machine-readable counterparts of various primary, secondary, or tertiary sources. Examples include the full text of several periodicals dealing with chemistry and medicine, an encyclopedia of chemical technology, chemical dictionaries, biographical and organization directories, textbooks in emergency medicine, and the full text of technical reports. At present, most databases are equivalent to or supplements of their printed counterparts rather than total replacements. In the case of full text in electronic form, text can be searched for combinations of key words, and retrieved text can be displayed or printed on demand at remote terminals, or it may be produced in microform (computer output microfiche, or COM).

The kinds of information needed by research workers and now found in the present varied literature of science and technology probably will not change much over the years. But the form and manner in which information is produced, published, stored, retrieved, and analyzed certainly will change as information science and technology and the information industry (1) find out more about how information really is used by research workers, and (2) find ways to solve problems that seem to beset the new electronic information systems: reliability, security, access vocabulary, copyright, archiving, and user visual fatigue, among others. Some solutions to these problems are emerging. For example, the Copyright Clearance Center operates a centralized photocopy authorization and payment system on behalf of copyright owners, both foreign and domestic, providing a practical means for organizations to comply with copyright law in the use of scientific-technical information.

New technologies continue to be introduced. CD-ROMs offer the possibility of high-volume, low-cost storage of databases, including multimedia. Microcomputers are proving to be versatile devices, used by individual scientists and engineers for both document creation and retrieval of information from databases. The Internet, a worldwide system of interconnected computer networks, allows access to information resources all over the world from a desktop computer workstation. Software tools for navigation and retrieval improve access to distributed resources, including data, images, sounds, text, and computer programs. *See* MICROCOMPUTER; MULTIMEDIA TECHNOLOGY; OPTICAL RECORDING.

Documenting progress. Research, development, and state-of-the-art reports and current news items covering all aspects of the science information field may be found through *Library and Information Science Abstracts* (an abstracting serial) and *Library Literature* (an indexing serial). Machine-readable database material is featured in *Online*, *Online & CD-ROM Review,* and *Database*.

George S. Bonn; Linda C. Smith

Bibliography. C. D. Hurt, *Information Sources in Science and Technology*, 3d ed., 1998; H. R. Malinowsky, *Reference Sources in Science, Engineering, Medicine, and Agriculture*, 1994; E. Mount and B. Kovacs, *Using Science and Technology Information Sources*, 1995; National Academy of Sciences, *Realizing the Information Future: The Internet and Beyond*, 1994; G. Stix, The speed of write, *Sci. Amer.*, 271(6):72–77, December 1994; R. D. Walker and C. D. Hurt, *Scientific and Technical Literature: An Introduction to Forms of Communication*, 1990.

Lithium

A chemical element, Li, atomic number 3, and atomic weight 6.939. Lithium heads the alkali metal family in the periodic table. In nature it is a mixture of the isotopes ^{6}Li and ^{7}Li. Lithium, the lightest solid element, is a soft, low-melting, reactive metal. In many physical and chemical properties it resembles the alkaline-earth metals as much as, or more than, it does the alkali metals. *See* ALKALINE-EARTH METALS; PERIODIC TABLE.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

The major industrial use of lithium is in the form of lithium stearate as a thickener for lubricating greases. Other important uses of lithium compounds are in ceramics, specifically in porcelain enamel formulation; as an additive to give longer life and higher output in alkaline storage batteries; and in welding and brazing fluxes.

Lithium is a moderately abundant element and is present in the Earth's crust to the extent of 65 parts per million (ppm). This places lithium a little below nickel, copper, and tungsten, and a little above cerium and tin in abundance.

Noteworthy among lithium's physical properties are the high specific heat (heat capacity), large temperature range of the liquid phase, high thermal conductivity, low viscosity, and very low density. Lithium metal is soluble in liquid ammonia and is slightly soluble in the lower aliphatic amines, such as ethylamine. It is insoluble in hydrocarbons. The physical properties of lithium metal are summarized in the **table**.

Lithium undergoes a large number of reactions with both organic, and inorganic, reagents. It

Physical properties of lithium metal

Property	Temperature, °C (°F)	Metric (scientific) units	British (engineering) units
Density	20 (68)	0.534 g/cm^3	33.7 lb/ft^3
	400 (752)	0.490 g/cm^3	30.6 lb/ft^3
	800 (1472)	0.457 g/cm^3	28.6 lb/ft^3
Melting point	179 (354)		
Boiling point	1317 (2403)		
Heat of fusion	179 (354)	103.2 cal/g	186 Btu/lb
Heat of vaporization	1317 (2403)	4680 cal/g	8420 Btu/lb
Viscosity	200 (392)	5.62 millipoises	11.1 kinetic units
	400 (752)	4.02 millipoises	8.2 kinetic units
	600 (1112)	3.17 millipoises	6.7 kinetic units
Vapor pressure	727	0.78 mm	0.015 lb/in.^{2}
	1077	91.0 mm	1.76 lb/in.^{2}
Thermal conductivity	216 (420)	0.109 cal/(s)(cm^2)(cm)(°C)	26.5 Btu/(h)(ft^2)(°F)
	539 (1002)	0.073 cal/(s)(cm^2)(cm)(°C)	17.6 Btu/(h)(ft^2)(°F)
Heat capacity	100 (212)	0.90 cal/(g)(°C)	0.90 Btu/(lb)(°F)
	300 (573)	1.02 cal/(g)(°C)	1.02 Btu/(lb)(°F)
	800 (1472)	0.99 cal/(g)(°C)	0.99 Btu/(lb)(°F)
Electrical resistivity	0 (32)	1.34 microhm-cm	
	100 (212)	12.7 microhm-cm	
Surface tension	200–500 (392–932)	About 400 dynes/cm	

reacts with oxygen to form the monoxide, Li_2O, and the peroxide, Li_2O_2. Lithium is the only alkali metal that reacts with nitrogen at room temperature to form a nitride, Li_3N, which is black. Lithium reacts readily with hydrogen at about 930°F (500°C) to form lithium hydride, LiH. The reaction of lithium metal with water is exceedingly vigorous. Lithium reacts directly with carbon to form the carbide, Li_2C_2. Lithium combines readily with the halogens, forming halides with the emission of light. While lithium does not react with paraffin hydrocarbons, it does undergo addition reactions with arylated alkenes and with dienes. Lithium also reacts with acetylenic compounds, forming lithium acetylides, which are important in the synthesis of vitamin A.

The most important lithium compound is lithium hydroxide. It is a white powder, and the material of commerce is actually lithium hydroxide monohydrate, $LiOH \cdot H_2O$. Lithium carbonate, $LiCO_3$, finds application in the ceramic industries and in medicine as an antidepressant. Both lithium halides, lithium chloride and lithium bromide, form concentrated brines with ability to absorb moisture over a wide temperature range; these brines are used in commercial air conditioning systems.

Marshall Sittig

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; G. N. Schrauzer and K. F. Klippel (eds.), *Lithium in Biology and Medicine*, 1991.

Lithosphere

The rigid or mechanically strong outer layer of the Earth that can support stress. The lithosphere is divided into 12 major plates, the boundaries of which are zones of intense activity that produce many of the large-scale geological features that characterize the Earth. These plates move as coherent units with velocities of up to several centimeters per year, and their relative movement and interaction form the foundation for the theory of plate tectonics.

The lithosphere comprises the crust (either continental or oceanic) and a portion of the upper mantle that together overlie a zone of relative weakness termed the asthenosphere (**Fig. 1**). The boundary between the crust and the mantle is known as the Mohorovičić discontinuity (Moho), and is compositional in origin—that is, the crust and mantle are distinguished by fundamental differences in rock chemistry. In contrast, the boundary between the lithosphere and the asthenosphere represents an isotherm that separates a conductively cooling lithosphere from a quasi-isothermal convecting asthenosphere. The asthenosphere differs from the overlying lithosphere principally in its ability to flow on geological time scales. These differences arise from the fact that temperature (and thus the fluid or flow properties of rocks) increases as a function of depth in the Earth. Whereas the lithosphere tends to be resistant to deformation, the asthenosphere deforms by flowing. A complication arises when the geothermal temperature is sufficiently high to

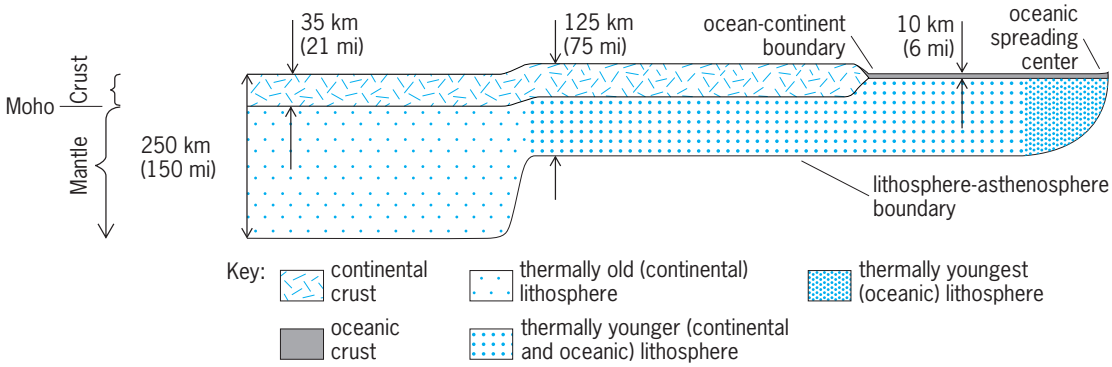

Fig. 1. Representation of oceanic and continental lithosphere as defined by thermal structure. Also shown are representative crustal and lithospheric thicknesses.

induce the lower crust to flow while the upper mantle is still relatively strong (for example, the Basin and Range Province of the western United States). Even in this case, a lithosphere-asthenosphere boundary still exists at depth. The lithosphere is either oceanic or continental, each type being fundamentally different in terms of the formation and composition of the rocks that constitute the crust and upper mantle. While the most common definition of the lithosphere is in terms of its temperature structure, there exist a whole range of alternative definitions that consider the seismic, mechanical, rheological, and chemical characteristics of the crust and mantle. *See* ASTHENOSPHERE; EARTH CRUST; MOHO (MOHOROVICIC DISCONTINUITY); PLATE TECTONICS; RHEOLOGY.

Oceanic lithosphere. The oceanic lithosphere ranges in thickness from near zero (young lithosphere at an oceanic spreading center) to 125 km (75 mi) within the oldest existing ocean basins ($\sim 2 \times 10^8$ years). Ages for oceanic lithosphere indicate not only their absolute age but their temperature state as well. Thus, young oceanic lithosphere refers to the new, hot lithosphere created at a spreading center as part of the plate tectonic cycle. In turn, the spreading center is the region where mantle (asthenosphere) material rises, melts, and is emplaced to produce new lithosphere. The thickness of oceanic crust typically ranges 5–10 km (3–6 mi), although thickness can approach 20 km (12 mi) when the normal oceanic crustal formation process is augmented thermally by the presence of hot spots (for example, Iceland). The oceanic lithosphere consists of a basaltic crust created at spreading centers by the passive emplacement of magma and an underlying mantle component. The new crustal material (oceanic crust) is extracted from the decompressive melting of the underlying asthenosphere. The mantle component consists of the residuum of this melting process (harzburgite rocks) and pristine upper-mantle material (lherzolite rocks) created by the transformation of asthenosphere to lithosphere by conductive cooling. Lherzolite is considered pristine because it is chemically indistinguishable from the composition of the underlying asthenosphere. The thickness of the oceanic crust is a function of the temperature of the underlying asthenosphere: the hotter the asthenosphere, the thicker the oceanic crust. *See* HOT SPOTS (GEOLOGY); MAGMA.

The oceanic lithosphere cools and thickens as it moves away from the spreading center, and ultimately it is consumed or recycled as it descends back into the asthenosphere in collisions with other plates to form subduction zones. The relatively young age of oceanic lithosphere ($\leq 0.2 \times 10^9$ years) compared to that of continental lithosphere ($\leq 3.8 \times 10^9$ years) is a direct consequence of this recycling process. *See* SUBDUCTION ZONES.

Continental lithosphere. The continental lithosphere ranges in thickness from approximately 50 to 250 km (30 to 150 mi), with crustal thicknesses ranging 20–70 km (12–40 mi). It consists of a crust created by the accretion of older, relatively low-density continental fragments augmented by the ascent and emplacement of new igneous material or even slivers of oceanic crust at subduction zones, and a mantle component that has been created by a range of processes. These processes include (1) the fracturing and emplacement of older continental mantle during plate collisions, (2) the fracturing and emplacement of older oceanic mantle during subduction, (3) the accretion of low-density depleted mantle derived from melting of the mantle that overlies subducted oceanic lithosphere, and (4) the transformation of asthenosphere to pristine lithosphere by long-term conductive cooling.

New continental crust is derived by partial melting of both the mantle overlying subducted oceanic lithosphere and the sediments that are transported along with the subducted plate during the formation of island arcs and continental arcs. In turn, island arcs are accreted to the continents during plate collisions. The decompressive melting of mantle material associated with hot spots and plumes is an important intraplate process for incorporating spatially extensive and thick basalts into the continental crust. *See* BASALT; OCEANIC ISLANDS.

Continental crust generally contains relatively light, granitic and gneissic rocks, and it tends to be thicker than its oceanic counterpart. While the upper continental crust is lithologically and chemically heterogeneous, it has an average composition

equivalent to granodiorite. The lower continental crust is significantly more mafic (containing magnesium and iron), with an average composition similar to that of oceanic crust (that is, basaltic). Because of their buoyancy, continental crust and strongly depleted continental mantle are difficult to subduct, and thus they are perpetually reincorporated into the structure of the interacting plates. It is this inability to subduct continental crust that helps explain the apparent logevity of the continents with respect to the oceans. *See* GNEISS; GRANITE; GRANODIORITE.

The continental crust is characterized by relatively abundant concentrations of long-lived radioactive isotopes of uranium (^{235}U and ^{238}U), potassium (^{40}K), and thorium (^{232}Th). Subduction recycling, surficial weathering, and magmatic and hydrothermal differentiation processes all help to concentrate these elements into the upper crust. The observed linear relationship between surface heat flow and radioactive heat content of surface rocks is satisfactorily explained by a distribution of radioactivity that decreases exponentially with depth. The contribution of radioactive heating to the mantle heat flux in the continents can be as high as 50%. *See* RADIOISOTOPE (GEOCHEMISTRY).

Other definitions. Although the most common and important definition of the lithosphere is in terms of its deformation properties, other definitions exist; they depend on identifying the differences between the lithosphere and the asthenosphere.

Thermal lithosphere. This is a thermal boundary layer between the surface of the Earth, which cools by conduction, and the asthenosphere, which cools by convection. In contrast, the underlying asthenosphere is characterized by convective stirring such that it is approximately isothermal, with a temperature of 1200–1400°C (2200–2600°F). The lithosphere is related to the asthenosphere by a physical change of state in that the boundary between them demarcates the onset of minor amounts of mantle melting (partial melting). Thus the asthenosphere has the ability to flow. It is precisely for this reason that the lithosphere-asthenosphere boundary represents the boundary between convective and conductive thermal processes.

The thickness of the oceanic thermal boundary layer (that is, oceanic lithosphere) is estimated by using the observed subsidence and heat flow as a function of crustal age. Typically, the thermal boundary layer thickness is 125 ± 10 km (75 ± 6 mi). In contrast, it is significantly more difficult to estimate the thickness of the continental thermal boundary layer (that is, the continental lithosphere) because of the effects of radioactive heating. The usual method for determining the thickness of the continental thermal boundary layer is to measure the length of time over which a sedimentary basin forms (known as the thermal time constant of the lithosphere). The thermal time constant estimated from continental sedimentary basins spans $6.5–26 \times 10^7$ years, implying that the continental lithosphere thickness ranges 125–250 km (75–150 mi). The thicker lithospheres tend to be associated with old, stable continental cratons. Their increased depth to the 1200–1400°C (2200–2600°F) isotherm (that is, the base of the lithosphere) is consistent with the temperature and pressure conditions obtained from studies on diamonds and their inclusions. *See* DIAMOND.

Two competing models have been used to explain the temperature behavior of the lithosphere: a cooling half-space model and a cooling plate model. For the cooling half-space model, it is assumed that conductive cooling progressively thickens the lithospheric mantle without limit; that is, the longer the interval of cooling, the thicker the lithosphere. By contrast, the cooling plate model assumes that there is a maximum thickness that limits conductive cooling. While observations of basement subsidence and heat-flow history favor the cooling plate model for the oceanic lithosphere (and by inference, the continental lithosphere), the subsidence data used for the model calibration may be biased by the effects of thermal plumes in the asthenosphere. This issue has not been resolved. It is probably fair to say that thermal steady state is never achieved in either the oceanic or continental lithosphere. In the continents, the complicating effects of erosion, radioactive heating, crustal thickening by the stacking of thrust sheets during broad continent-continent collisions (for example, the Himalayas and the Tibetan Plateau; the Hercynian orogeny of Europe), and the perpetual recycling of continental crust all mitigate against a simple application of the same techniques to estimate continental lithosphere thickness. Nevertheless, the general correlation between surface heat flow and the age of the last thermal event that affected the continental crust suggests that the temperature structure of the continents is primarily conductive, although it is not possible to discriminate between the applicability of the half-space or plate cooling models. *See* EARTH, CONVECTION IN; EARTH, HEAT FLOW IN.

Seismic lithosphere. The seismic lithosphere is that region between the surface of the Earth and the asthenosphere that is characterized by the efficient propagation of seismic waves. Since seismic velocity and attenuation are sensitive to temperature and melt content, the base of the lithosphere can be mapped using the variations in seismic P (longitudinal) waves and S (shear) waves. The lithosphere-asthenosphere boundary is marked by an abrupt decrease in shear-wave velocity. This low-velocity zone is 50–100 km (30–60 mi) thick and represents a region of incipient partial melting (approximately 1%). In effect, the velocities and efficient propagation of seismic body waves define a high-velocity "lid" that lies above the low-velocity asthenosphere. Surface wave studies also indicate that the depth to the low-velocity zone associated with relatively old oceanic lithosphere is approximately 100–125 km (60–75 mi). It has proved difficult to recognize the low-velocity zone beneath continental regions.

A subregion of the seismic lithosphere in which intraplate earthquakes are generated is termed the seismogenic zone. The seismogenic zone represents

those layers of the crust and mantle that are sufficiently cool that they can rupture by brittle fracture. *See* EARTHQUAKE; SEISMOLOGY.

Mechanical lithosphere. The mechanical lithosphere, also known as the elastic lithosphere, is a mechanical analog used to simulate the flexural deformation or response of the Earth to applied geological loads such as seamounts, subducted slabs, mountain ranges and thrust sheets, and deltas. Similarly, the Earth can be unloaded by processes such as lithosphere extension and erosion. The mechanical analog is a thin elastic plate overlying a fluid substratum. The flexural properties of the plate are characterized by its flexural rigidity, or equivalently by its thickness. When applied to the lithosphere, this thickness is termed the effective elastic thickness (T_e) of the lithosphere because it reflects the mechanical strength of the lithosphere in a depth-averaged sense. Even though the effective elastic thickness is a parametrization of real lithosphere behavior, its use does not imply that the lithosphere is necessarily elastic or that a geological or rheological boundary occurs at a specific depth in the lithosphere. In general, the effective elastic thickness should be thought of as an empirical parameter representing the integrated rheological strength of the entire lithosphere, with all its complexity, at a given time and location. Thus, the deformed shape of an elastic plate of thickness T_e is the same as the deformation of the lithosphere for the same load.

Three classes of studies have been used to determine the effective elastic thickness for the oceans and continents: (1) inverse gravity admittance techniques, which exploit the statistical relationship between topography and gravity anomalies; (2) forward modeling of the crustal shape and stratigraphy of extensional basins and their respective free-air gravity anomalies; and (3) determining the relationship between the basement relief and the Moho topography. Critical in these analyses is the fact that the wavelength of flexure indicates the strength of the lithosphere—a strong lithosphere produces a long-wavelength deformation of small amplitude, while a weaker lithosphere results in a shorter-wavelength deformation but of larger amplitude. *See* MODEL THEORY; SIMULATION.

In the oceans, a simple straightforward application of the admittance between bathymetry and free-air gravity (that is, the ratio of the gravity and bathymetry power spectra) has proved valuable in defining both the spatial and temporal variations of the effective elastic thickness. Most importantly, the effective elastic thickness tends to be a function of the age (and hence temperature structure) of the lithosphere at the time that it was loaded. The flexural strength of the oceanic lithosphere, as expressed by the effective elastic thickness, ranges 5–40 km (3–24 mi) and correlates with the depth to the 300–600°C (570–1110°F) isotherm as determined from thermal models for the lithosphere. The continents are somewhat more problematic. An implicit assumption in the interpretation of admittance functions is that the topography is the load responsible for deforming the crust and lithosphere. When the topography (or surface load) is augmented with subsurface loads, the effective elastic thickness estimates are suspect and are biased toward low values. This problem of subsurface loading seems most prevalent in the continents. When allowance is made for subsurface loads, the admittance technique predicts a range of the values of the effective elastic thickness for the Australian and North American continents [15 km (9 mi) $< T_e <$ 135 km (81 mi) and 4 km (2.4 mi) $\leq T_e \leq$ 128 km (79 mi), respectively] that correlates with the age range of orogenic events across each continent. Extremely low values of T_e are invariably related to regions that have recently undergone or are presently undergoing tectonic deformation involving large amounts of heat (for example, the Basin and Range Province of the western United States).

The flexural strength of the lithosphere can be determined by forward-modeling the topographic relief across regions of continental extension and compression. For example, in extensional settings the topography flanking rift basins is produced by the flexural rebound of the lithosphere in response to its tectonic unloading (that is, lateral removal of crustal and mantle material) during extension. Lithospheric flexural strength is maintained despite localized but intense fracturing of the crust by normal faults that penetrate to lower crustal levels.

Within compressional settings, studies of the gravity anomaly over mountain belts (that is, the load) and their associated foreland basins (that is, the flexural response or deformation of the lithosphere) indicate that the continental lithosphere in these tectonic settings has significant flexural strength. Estimated effective elastic thicknesses range 20–150 km (12–90 mi) and suggest that, in a broad way, the continents behave in a manner similar to the flexural deformation of oceanic lithosphere. The relationship between the flexural strength and the thermal age for the continental lithosphere appears similar to that found in the oceans as evidenced by the widths of foreland basins, given that foreland basin width is primarily controlled by the flexural strength of the lithosphere. Foreland basins on relatively young crust (for example, Molasse basin/Hercynian crust with an age of $2.45–2.6 \times 10^8$ years is 200 km or 120 mi wide) are narrower than foreland basins on relatively old crust (for example, Ganges basin/Archean crust with an age of 2.5×10^9 years is 400 km or 240 mi wide). The Molasse basin of Switzerland, southern Germany, and Austria represents the sediment-filled flexural deflection of the European crust, the deflection being produced in response to the weight and subsequent loading of the Alpine mountain thrust load. Similarly, the Ganges basin of northern India represents the sediment-filled flexural deflection of the Indian shield produced by the lateral emplacement of the thrust sheets that constitute, in part, the Himalayan Mountains. The thermal states of the European and northern Indian lithospheres are very different. Whereas the Indian lithosphere consists of extremely old, cold, crustal

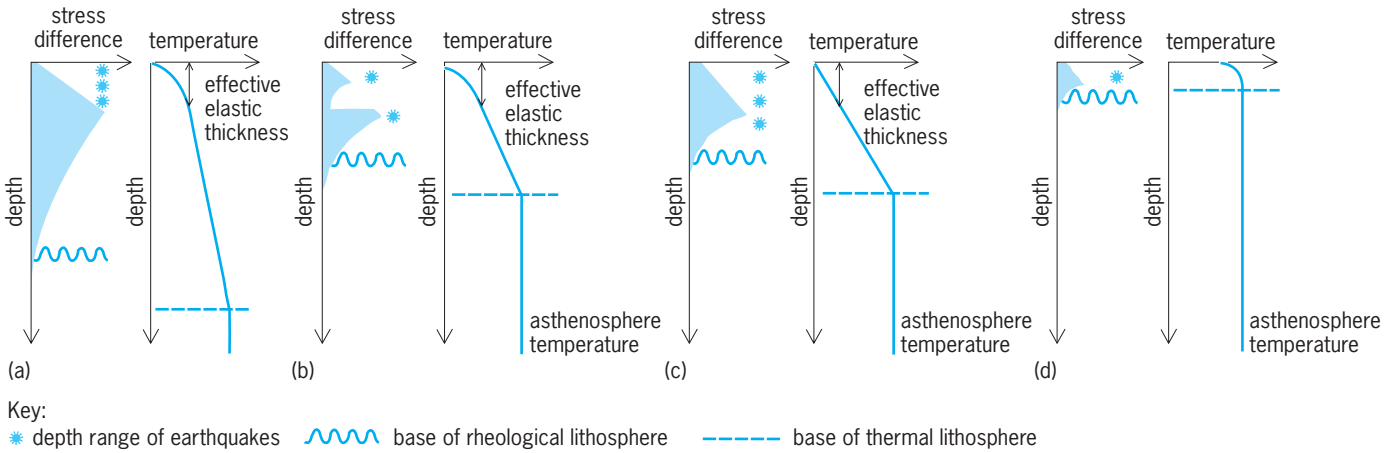

Fig. 2. Representation of oceanic and continental lithosphere as defined by thermal structure. Also shown are representative crustal and lithospheric thicknesses.

and mantle material, the European crust and mantle were severely deformed and heated during a major period of mountain building known as the Hercynian orogeny. The flexural response of the two lithospheres to thrust sheet loading (that is, the widths of their foreland basins) reflects the difference in their thermal states. *See* OROGENY.

While the importance of the free-air gravity cannot be overstated, there are some precautions that should be taken when using gravity to assess the flexural strength of the lithosphere. For example, a common misconception is that extensional basins are characterized by low-amplitude free-air gravity anomalies and thus the extended lithosphere has low flexural strength. A comparison of the similarly dimensioned Central Graben of the North Sea (−10 milligals) with the Tucano basin of Brazil (−140 milligals) indicates that the statement on gravity amplitude is incorrect. Another problem is that the gravity anomaly does not necessarily reflect details of the basement geometry because of the effects of compaction. As sediments compact, the density contrast with respect to the basement decreases, resulting in a decrease in the amplitude of the gravity anomaly. Thus, low-amplitude gravity anomalies over extensional basins do not necessarily reflect the flexural strength of the lithosphere, which is in marked contrast to basins that exhibit large negative gravity anomalies. Such gravity anomalies are indicative of high flexural strength.

Rheological lithosphere. This is the region of the crust and mantle over which the rheological or flow properties of rocks are insufficient to relieve stresses (primarily elastic). In general terms, stress is the pressure being sustained or supported within a material, whereas strain is the degree of deformation suffered by that material. Since by definition the lithosphere represents that region of the crust and mantle that can support significant stress, the more fluidlike the rocks become with increasing depth and temperature, the less shear stress they can support. A level of 10–50 megapascals is arbitrarily assigned to define the base of the rheological lithosphere.

Rock strength varies with depth in the lithosphere in response to changes in ambient temperature, confining pressure, and compositional variations. At shallow depths in the lithosphere, rock strength depends mainly on confining pressure, and yielding occurs as frictional sliding on fracture surfaces (that is, brittle deformation). It is this episodic stick-and-slip nature of frictional sliding that produces earthquakes. With increasing depth in the lithosphere, rock strength is expected to decrease as increasing temperatures and confining pressure promote thermally activated creep processes (that is, the rocks deform plastically). A particularly useful representation of the rheological lithosphere that incorporates these processes is in terms of a yield-stress envelope (**Fig. 2**). The yield-stress envelope defines the depth dependence of the maximum supportable stress difference in the lithosphere as a function of depth and composition. The continental yield-stress envelope has an additional region of low strength within the lower crust caused by the plastic deformation of the quartz-rich rocks that characterize this type of crust. The integral of the yield-stress envelope as a function of depth is a proxy for the total strength of the lithosphere, and it represents the force (per unit length) that opposes the breaching or breaking of the lithosphere due to horizontal forces.

According to the concept of the yield-stress envelope, even small levels of applied force are expected to produce yielding at the top and bottom of the lithosphere by brittle faulting and ductile creep, respectively (Fig. 2). The intervening core region can support large stress differences elastically without yielding on geological time scales, and correlates in a general way with the effective elastic thickness of the lithosphere. As with effective elastic thickness, it is this core region that controls the wavelength of flexural deformation in response to applied forces and loads, this wavelength being on the order of tens to hundreds of kilometers. In contrast, the brittle deformation associated with yielding in the top portion of the lithosphere might not have any preferential length scale, depending instead on the distribution

of existing faults and fractures in the crust and their orientation in relation to the applied force. *See* RHEOLOGY; STRESS AND STRAIN.

Chemical lithosphere. This lithosphere is defined as a chemical boundary layer between the surface of the Earth and the asthenosphere that cools by conduction and contains both the material differentiated or extracted from the mantle (for example, oceanic and continental crust) and mantle material modified by various degrees of depletion.

In general, the continental chemical lithosphere represents (1) the accretion of low-density depleted mantle originally associated with the formation of oceanic crust but remobilized by mantle convection systems, (2) the accretion of low-density depleted mantle derived from melting of the mantle overlying subducted oceanic lithosphere, and (3) the fracture and emplacement of older continental crust and mantle during plate collisions. As such, these magmatic differentiation processes explicitly result in chemical heterogeneities within the lithosphere. Perhaps more importantly, the low-density accreted mantle material is crucial in offsetting the subsidence engendered by the continued cooling and contraction of the lithosphere, thus for example helping to maintain cratons at relatively low but broadly distributed elevations (200–800 m or 660–2640 ft). The degree to which the lithospheric mantle is chemically distinct from the asthenosphere mantle is controversial. Seismic tomographic imaging of the mantle has revealed large velocity anomalies that correlate with the past location of subduction zones. Presumably, the subduction recycling process is conducive to a heterogeneous mantle asthenosphere with some affinity to the chemical composition of the lithosphere. *See* COMPUTERIZED TOMOGRAPHY.

Estimates for the thickness of old continental lithosphere are constrained primarily from the existence and composition of carbonatitic intrusions, particularly kimberlites, and the diamonds they may contain. Kimberlites are of special significance because they contain xenoliths of ultramafic rocks and eclogite from the upper mantle. Kimberlitic magmas result from the low partial melting of upper-mantle rocks (garnet lherzolites) at depths of 50–300 km (30–180 mi) and are rapidly emplaced. Kimberlites range in age from 20 to 1750 million years and occur in all major Precambrian cratons, suggesting that they are generated within the lower continental lithosphere. This implies that the lithospheric mantle of old continental terranes is chemically distinct from either younger continental terrane or the oceanic lithosphere. *See* ECLOGITE; KIMBERLITE; PRECAMBRIAN; XENOLITH.

Lithosphere interrelationships. A useful way to view the relationships between the various definitions for the lithosphere is by considering the depth distribution of earthquakes, temperature, and stress differences within old and young oceanic and continental lithosphere: old, cold continental lithosphere; relatively young continental lithosphere; relatively old, cold oceanic lithosphere; and young, hot oceanic lithosphere (Fig. 2).

The base of the thermal lithosphere is defined by T_m, which is the temperature of the asthenosphere at which olivine-bearing rock begins to flow easily. This boundary thus seperates material that remains in place and cools by heat conduction from material that moves as it is stirred by convection. The change in the rheological properties between old and young continental lithosphere is a function of the increased temperature gradient associated with the younger continental lithosphere and the accompanying creep activation of the quartz-rich rocks that compose the continental crust. As the lithosphere cools, this creep component becomes subordinate to the brittle deformation process, causing the yield-stress envelope bite within the lower crust to systemically diminish. It is the absence of this quartz-rich component in oceanic crust that results in the brittle-dominated segment of the yield-stress envelope observed for both old and young oceanic lithosphere.

Because of the pivotal role played by temperature in controlling the flow of rocks, there is a general temporal relationship between thermal lithospheric thickness, temperature gradients, brittle failure within the crust and upper mantle (defining the seismogenic layer), and the effective elastic thickness of the lithosphere. *See* EARTH. Garry D. Karner

Bibliography. M. H. P. Bott, *The Interior of the Earth: Its Structure, Constitution and Evolution*, 2d ed., 1982; K. C. Condie, *Plate Tectonics and Crustal Evolution*, 3d ed., 1989; C. M. R. Fowler, *The Solid Earth: An Introduction to Global Geophysics*, 1990; H. L. Levin, *The Earth Through Time*, 3d ed., 1988.

Litopterna

Hoofed herbivores confined to the Cenozoic of South America. The order was well represented from the Paleocene to the Pleistocene, and apparently arose on that continent from a condylarth ancestry. By later Paleocene time two main lines of descent were clearly demarcated. The Proterotheriidae (**Fig. 1**) displayed a remarkable evolutionary convergence with the horses in their dentition and in reduction of the lateral digits of their feet. In one group the foot was reduced to a single median toe by early Miocene time. The members of the Macraucheniidae (**Fig. 2**) were proportioned much as in the camels and by late

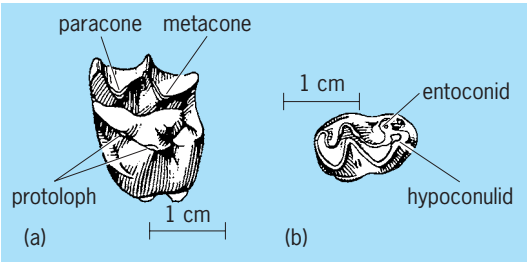

Fig. 1. Molars, (*a*) upper and (*b*) lower, of *Proterotherium*, an early Miocene proterotheriid litoptern from the Santa Cruz Formation of Patagonia, Argentina.

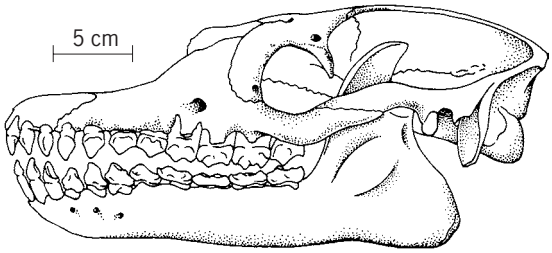

Fig. 2. Skull and jaw of *Theosodon garrettorum,* an early Miocene macraucheniid litoptern, Santa Cruz Formation, Patagonia, Argentina.

Tertiary time had similarly lost the vertebral arterial canal of the cervical vertebrae. They early developed a tendency toward retraction of the nasal bones; by late Tertiary time, the nasal opening was situated above the eyes on top of the skull. This specialization has been thought to indicate the presence of a trunk or perhaps dorsally situated nostrils adapted for a submerged amphibious existence. Three functional digits remain in the feet of even the most advanced macraucheniids. *See* ARCHAIC UNGULATE.

Both groups share the following characters held typical for the order. The skull was without expansion of the temporal or squamosal sinuses, as in the notoungulates; a postorbital bar was present. Dentition was primitive, with full eutherian formula and a tendency to reduce incisors and upper canine teeth; cheek teeth were unreduced, low- to high-crowned, the posterior premolars becoming molariform. The upper molars were quadrate, the paracone and metacone were crescentic, the protoloph was present and the metaloph was absent. The lower molars were bicrescentic; the entoconid was not isolated but joined to the posterior end of the tooth. The feet were functionally three-toed or reduced to a single digit, the weight borne by digit three. *See* DENTITION.

The earliest known mammalian faunas in South America (later Paleocene) contained representatives of both groups of litopterns. By Oligocene and Miocene time the litopterns reached their greatest recorded abundance. The proterotheres did not survive the Pliocene, and of the macraucheniids only the large *Macrauchenia* survived until Pleistocene time. *See* MAMMALIA. Richard H. Tedford

Bibliography. M. J. Benton, *Vertebrate Paleontology*, 1991; R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988.

Liver

A large gland found in all vertebrates. It consists of a continuous parenchymal mass forming a system of walls (muralium) through which venous blood from the gut must pass. This strategic localization between nutrient-laden capillary beds and the general circulation is associated with hepatic regulation of metabolite levels in the blood through storage and mobilization mechanisms controlled by liver enzymes.

The liver is the sole source of such necessary constituents of the blood as fibrinogen, serum albumin, and cholinesterase (an enzyme which may prevent harmful buildup of the neuronal transmitter agent acetylcholine). In the embryonic stage of most vertebrates the liver serves as the major manufacturing site of erythrocytes, a process known as erythropoiesis. The liver also removes toxins from the systemic circulation and degrades them, as well as excess hormones, by oxidation, hydroxylation or conjugation. Particulate material may be removed through a phagocytic action of specialized cells (Kupffer cells) lining the lumen of the hepatic "capillary spaces," or sinusoids. In addition to the products which the liver delivers directly to the general circulation (endocrine function), it secretes bile through a duct system which, involving the gallbladder as a storage chamber, eventually passes into the duodenum (exocrine function). Bile functions as an emulsifier of fats to facilitate their digestion by fat-splitting lipases, and may also activate the lipase directly.

The liver does not exist in invertebrates except as the smaller analogous hepatopancreas of the crustaceans. In the vertebrate embryo the liver is a rapidly growing organ, but it rarely undergoes cellular division in the adult unless a portion is removed. Excision leads to such rapid restoration that regenerating mammalian liver has been used as a model in growth studies. As the largest discrete internal organ, it may make up more than 5% of total body weight in rodents and up to 3% in humans. Though the liver is highly variable in shape in different vertebrate groups, the basic features of circulation are probably similar in all livers.

The major afferent blood supply (70%) is from the hepatic portal vein, carrying blood from the spleen and digestive organs, with a smaller contribution from the hepatic artery. The smaller branches of these vessels join with the bile ductules and, as portal canals, reach the sinusoids. The major efferent vessel is the hepatic vein, whose smaller processes are known as central veins. Drainage is into the inferior vena cava.

Within the liver there are small branches of both the hepatic artery and portal vein supplying an elaborate network of capillaries apposed to the bile canaliculi (minute channels coursing in a continuous network between the parenchymal cells of the liver). These capillaries drain into the portal vein to form a minute portal system. This subsidiary system, supplying the liver cells directly with nutrient and oxygen, may represent a significant site for hepatic artery function in some species. In a number of mammalian species, for example, the pig, a discrete connective tissue boundary demarcates the liver mass into hexagonal bundles or lobules with portal canals arranged peripherally and central veins occupying the center. Even those forms lacking patent lobular arrangements demonstrate lobular distribution of key metabolites.

Comparative Anatomy

The liver is found only in vertebrates and is ubiquitous in this group. Despite the variations in embryonic hepatic development encountered in the

various vertebrate prototypes, all adult livers are essentially similar. This phenomenon is at sharp variance with Haeckel's biogenetic law, which stresses the similarity of embryonic structures and the divergence of adult structures. It also attests to the evolutionary success of the basic pattern of hepatic architecture: a continuous convoluted wall of functional parenchymal cells tunneled by passageways which contain a network of blood vessels and biliary tracts in homogeneous fashion within the organ.

In *Branchiostoma* (*Amphioxus*) a hollow outpocketing of the foregut, called a hepatic cecum, receives veins from the intestine. This may presage the portal system and liver of higher forms, and many consider this structure the first true homolog of the vertebrate liver.

Humans. Knowledge of the precise connections between liver and intestine and of liver and systemic circulation provides the mechanical basis for understanding hepatic function (**Fig. 1**). Although complex biochemically, the human liver is extremely simple structurally; it will be used here as a model for the vertebrates in terms of gross anatomic features, blood supply, and intrahepatic morphology. The liver is a remarkably plastic structure, and its general contour varies greatly from one class to another; however, the underlying features that determine function are sufficiently similar to justify selection of a single prototype.

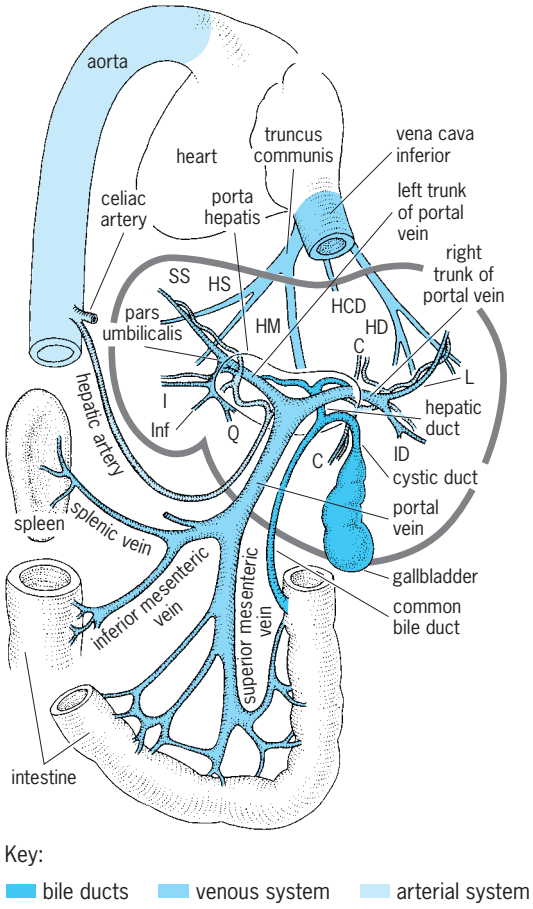

Fig. 1. Human liver and its major dorsal connections. Rami venae portae: C, centralis; I, intermedius; ID, inferior dexter; Inf, inferior sinister; L, lateralis; Q, quadratus; and SS, superior sinister. Venae hepaticae: HCD, dorsocaudalis; HD, dextra; HM, media; and HS, sinister.

General features. The human liver is a massive wedge-shaped organ divided into a large right lobe and a smaller left lobe. Its anterior surface underlies the diaphragm. The upper portion of the liver is partially covered ventrally by the lungs, whereas the lower portion overhangs the stomach and intestine. Both liver and gallbladder are innervated by the right phrenic nerve and by branches of the sympathetic fibers arising from the celiac ganglion. The vagus constitutes the parasympathetic source. The entire liver is covered by Glisson's capsule, an adherent membranous sheet of collagenous and elastic fibers.

Blood supply. Venous blood from the intestine, and to a lesser extent from spleen and stomach, converges upon a short broad vessel, called the hepatic portal vein, which enters the liver through a depression in the dorsocaudal surface termed the porta hepatis. There the hepatic portal vein divides (Fig. 1) into a short right branch (truncus dexter) and a longer left branch (truncus sinister). These vessels then ramify into the small branches which actually penetrate the functional parenchymal mass as the inner tubes of the portal canals (**Fig. 2**).

The hepatic artery, which is a branch of the larger celiac artery, also enters at the porta hepatis and ramifies unto smaller branches, which flank the portal venules within the portal canals (Figs. 1 and 2). Anastomoses of these arterial pairs result in the formation of a loose arterial plexus around each portal-vein branch. Arterial blood is also brought to the bile ducts and the lining cells of the portal canals by other sets of arterioles branching from tributaries of the hepatic artery.

The branches of the portal vein and hepatic artery then empty into the sinusoids (Fig. 2), the specialized capillaries lying within the spaces (lacunae hepatis) of the hepatic muralium. The sinusoids, which are major regions of hepatovascular exchange, communicate with small branches of the hepatic veins and, through the hepatic vein, the blood is returned to the heart by way of the vena cava. *See* CARDIOVASCULAR SYSTEM.

The hepatic canals house the hepatic venous branches in a manner similar to the containment of portal-vein and hepatic-artery branches within the portal canals. These respective canal systems are independent and are separated by the parenchymal cells (**Fig. 3**). Sinusoidal elements provide a functional continuum between afferent and efferent vascular systems. The absence of anastomoses within the portal canal system, as well as within the hepatic canal system, leads to a highly discrete pattern of vascular supply and drainage. This is of particular importance in surgical manipulations of the liver, since collateral circulation is of little significance in providing connections to those lobes whose vessels have been ligated.

Bile ducts. The tiny bile canaliculi, which lie between grooves in adjacent parenchymal cells, communicate with tiny intralobular bile ducts

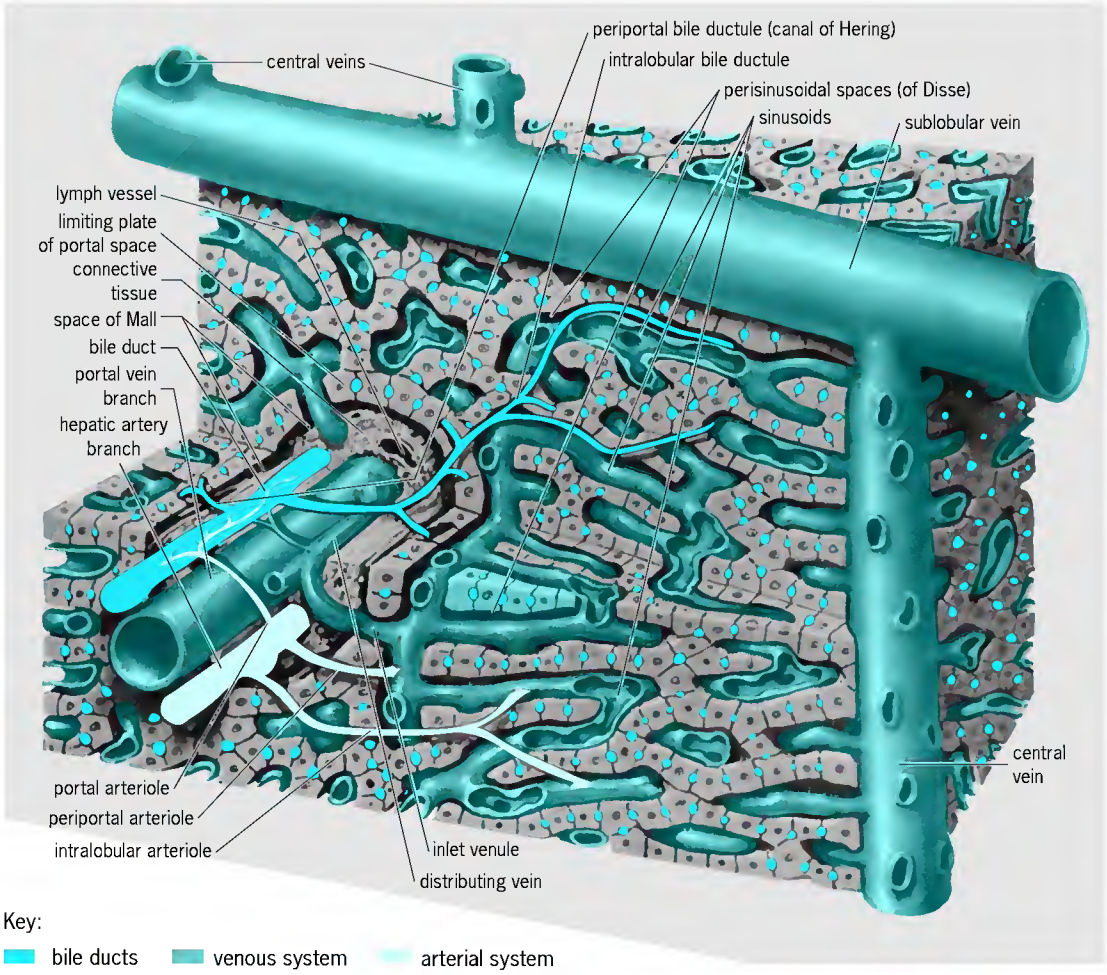

Fig. 2. Liver structure. ***(After E. Oppenheimer, ed., The Ciba Collection of Medical Illustrations, vol. 3: Digestive System, pt. 3: Liver, Biliary Tract and Pancreas, Ciba, 1957)***

(cholangioles or ductules), termed the canals of Hering (Fig. 2 and **Fig. 4**). These intralobular bile ducts empty into increasingly larger interlobular bile ducts which lie within the portal canals and make up the third element of the so-called portal triad (**Fig. 5**).

Although the biliary system is intimately associated with the portal system anatomically, even in terms of egress through the porta hepatis, it should be stressed that bile flow within this biliary system is in the opposite direction from that of blood flow. The hepatic duct arises in the region of the porta from the confluence of the largest bile ducts.

Gallbladder. A large outpouching of the hepatic duct forms the gallbladder (vesica fellea) and its cystic duct. Beyond the bifurcation the common bile duct enters the duodenum, thus serving as a conduit for bile from both the liver and its bile storage organ. Several mammals (for example, the horse and rat) do not possess gallbladders, and large numbers of humans have had their gallbladders removed surgically, thus indicating that the gallbladder is not essential for life. Electron microscopy studies indicate that the tall columnar cells which line the lumen of the gallbladder contain many microvilli. These may function in the concentration of bile, a significant function of gallbladder epithelium involving the active transport of sodium and potassium.

In addition to acting as a reservoir for the bile synthesized by the liver, the gallbladder plays a significant role in controlling the flow of bile to the intestine, when needed, because of its exquisite sensitivity to ingested fat or meat. Following a high-fat meal, up to three-fourths of gallbladder bile is expelled in less than 1 h as a result of contraction of the musculature of that organ and the concomitant relaxation of the sphincter choledochus (a ring of muscle) at the terminus of the common bile duct. Spasms of the sphincter may lead to abdominal pain and respiratory distress, symptoms which may be confused with those of heart seizures. *See* GALLBLADDER.

Vertebrata. The primitive vertebrate group Agnatha is represented by the extant cyclostomes: hagfishes and lampreys. The cyclostome liver is extremely small and consists of only a single lobe in the adult lamprey. The bilobed structure of the hagfish is more typical of the vertebrate archetype. The livers of higher fish are relatively large and lobated.

In some forms the gallbladder and even the bile duct may degenerate during metamorphosis and thus may be absent in the adult. A gallbladder is generally

present, although a few species of cartilaginous fish (shark) may lack this structure. In the teleosts (bony fish), the gallbladder is often quite large.

The economic importance of fish liver (oils and vitamins) stems in part from the capacity of this organ to synthesize and store lipids. Researchers hypothesized that in lower vertebrates the liver carries out synthetic functions, particularly of a lipid nature, which are delegated to the increasingly important adipose-tissue mass in birds and mammals. Shape and pancreatic relationship are quite variable for teleost liver.

Amphibians. The large liver of the amphibians, which is lobed and possesses a gallbladder, receives venous blood from the abdominal vein as well as the hepatic portal vein. Some variation in lobation patterns are encountered in the members of this class. *See* AMPHIBIA.

Reptiles. Reptiles demonstrate typical patterns of hepatic topography. In snakes (Squamata) one elongated lobe is found, whereas turtles (Chelonia) and crocodiles (Crocodylia) have two lateral masses connected by a thin isthmus. The gallbladder is invariably present. *See* REPTILIA.

Birds. The liver of birds is invariably lobed, but quite compact. The gallbladder is often absent but may exist as primordium in the embryo. In the pigeon two separate bile ducts lead from each lobe of the liver to the duodenum. *See* AVES.

Mammals. A great deal of variation characterizes the lobation of mammals (**Fig. 6**). Two major lobes may be subdivided in a variety of ways. If the lobations are superficial, the liver is classed as unified, as in the case of humans, whales, and many of the ruminants. In these livers the portal and hepatic canals tend to run at right angles to each other, whereas markedly lobated livers show a parallel arrangement of the large branches of these canals. Removal of a portion of a unified liver is a more difficult procedure than similar extirpation from a lobated liver.

Many mammals lack a gallbladder (rat, whale, coney, some of the artiodactyla, and all hoofed animals). In humans, those individuals who carry typhoid fever usually harbor the typhoid bacilli within the gallbladder and discharge the organisms in their feces. *See* ARTIODACTYLA; MAMMALIA.

Histology

In all vertebrates, with the exception of the adult lamprey, the liver is essentially a three-dimensional lattice of interconnected walls or plates made up of epithelial cells (hepatic parenchyma) [**Fig. 7**]. The cylindrical spaces (lacunae hepatis) which tunnel within this parenchymal mass interconnect to form a continuous maze of corridors known as the labyrinthus hepatus. Suspended in the lacunae hepatis are the sinusoids, or liver capillaries. These are relatively large, irregular, anastomosing vessels lined discontinuously with a specialized endothelial layer. Each functional unit of the liver is in contact with terminal branches of the two pervasive tunnel systems traversing the liver: the portal and hepatic canals. The branches of these two systems never

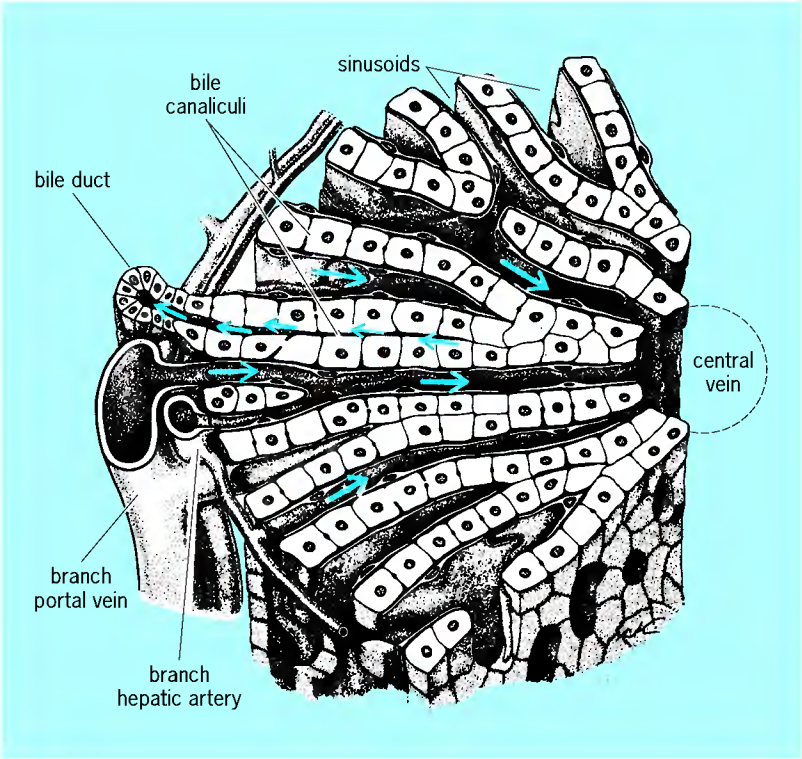

Fig. 3. Radial disposition of the liver-cell plates and sinusoids around the terminal hepatic venule or central vein, showing centripetal flow of blood from branches of the hepatic artery and portal vein and centrifugal flow of bile (small arrows) to the small bile duct in the portal space. (*After W. Bloom and D. W. Fawcett, A Textbook of Histology, 9th ed., Saunders, 1968*)

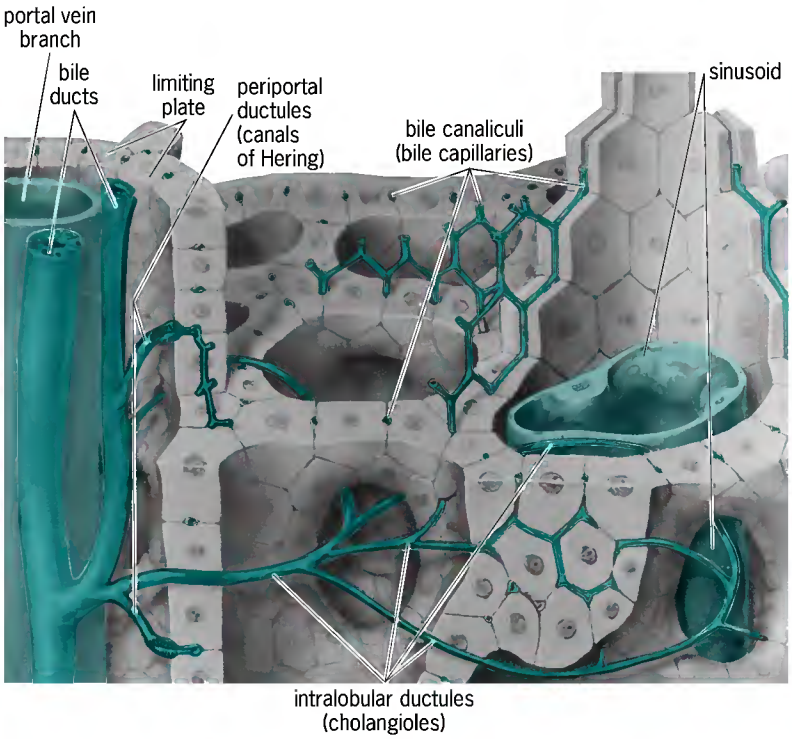

Fig. 4. Intrahepatic biliary system showing the arrangement of the bile ducts and the venous system. (*After E. Oppenheimer, ed., The Ciba Collection of Medical Illustrations, vol. 3: Digestive System, pt. 3: Liver, Biliary Tract and Pancreas, Ciba, 1957*)

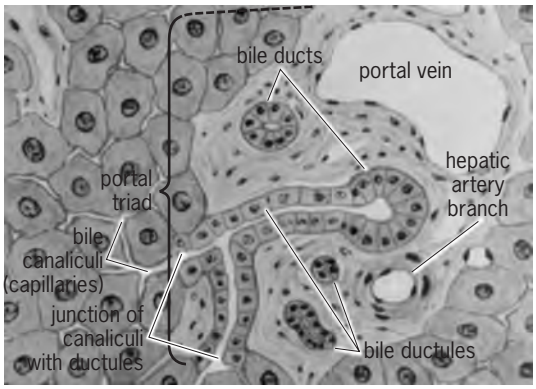

Fig. 5. Liver section, with components of portal triad. (*After E. Oppenheimer, ed., The Ciba Collection of Medical Illustrations, vol. 3: Digestive System, pt. 3: Liver, Biliary Tract and Pancreas, Ciba, 1957*)

communicate directly, but each is dovetailed into the sinusoid, the site of functional linkage between the two arborescent systems.

Limiting plate. The external surface of the liver is covered by a single layer of thin parenchymal cells lying immediately beneath the fibrous Glisson's capsule. This limiting plate is continuous with the hepatic muralium. It is reflected inward at the porta hepatis to form a fenestrated wall around the portal canals. In a similar manner the hepatic canals are also lined with a limiting plate (Figs. 2 and 4). Thus the external and internal limiting plates make up a continuum which is in intimate association with the internal muralium. The fibrous capsule is also reflected into the interior of both canal systems and provides a matrix for bringing blood vessels, lymph channels, nerves, and ducts into the interior of the liver.

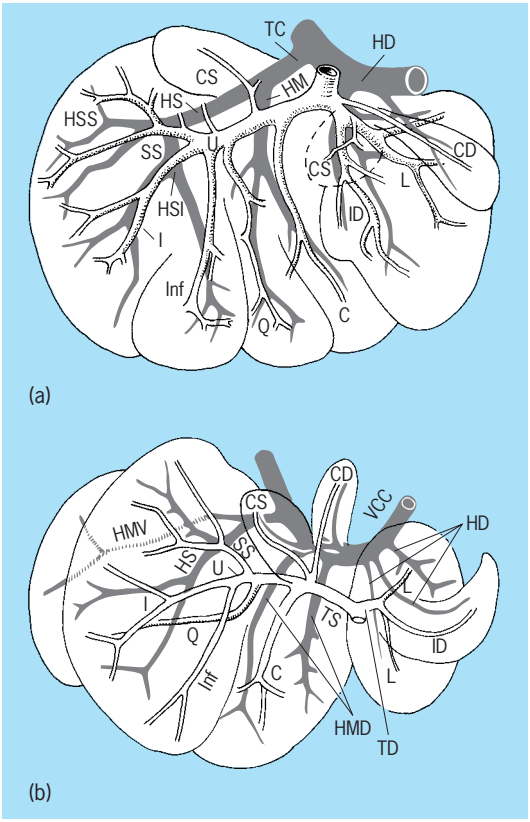

Fig. 6. Dorsal aspect of (*a*) liver of dog, (*b*) liver of rat. Rami venae portae: C, centralis; I, intermedius; ID, inferior dexter; Inf, inferior sinister; L, lateralis; Q, quadratus; SS, superior sinister. Venae hepaticae: HD, dextra; HM, media; HMV, ventral division of media; HMD, dorsal division of media; HS, sinister; HSI, sinistrae radix inferior; HSS, sinistrae radix superior. Ramuli: CD, caudatus dexter; CS, caudatus sinister; TC, truncus communis. Trunci venae portae: TD, dexter; TS, sinister; U, pars umbilicalis; VCC, vena cava inferior.

Sinusoids. The lining of these specialized capillaries has been the subject of considerable debate. Older theories postulated a syncytium (absence of individual cell boundaries), but electron micrographs show separate cells present. These lining cells were once hypothesized to be of only one type, the Kupffer cells, with a great capacity for engulfing foreign material (phagocytosis). It is now thought that there may be as many as three distinct types making up the lining, although transformation of one type to another is possible. The large macrophages (Kupffer cells) engulf degenerative erythrocytes as well as foreign or unassimilated material, thus "straining" the blood which moves through the liver sinusoids from the vessels of a portal canal to the central veins. Once again electron microscopy has clearly demonstrated that, although some of these lining cells tend to overlap one another, others have margins which are separated by as much as 0.5 μm. Some animals, for example, the calf, do not demonstrate this discontinuity; this raises the interesting problem of possible species differences. The regulation of blood flow within the hepatic lobule is achieved in part through a swelling of the macrophages lining the sinusoid.

A patent space, called the space of Disse, stands between sinusoid cells and the liver parenchymal cells (Fig. 2). Irregularly arranged projections, or microvilli, extend from the parenchymal cell and lie within the space of Disse. No limiting basement membrane blocks the free interchange between sinusoid gaps and the parenchymal cell. Futhermore, the villi of the parenchymal cell lying within the perisinusoidal space increase parenchymal cell surface across which exchanges may occur with the lymph-like fluid emanating abundantly from the sinusoid. This increase in the area of exchange is particularly important in terms of the rapid transfer of metabolites that must occur between the liver and portal blood.

Biliary system. The bile canaliculi are found between parenchymal cells throughout the liver. They are composed of a continuous network of polyhedral links associated with the convoluted muralium. If a single parenchymal cell is isolated, a hexagonal mesh of bile canaliculi appears to surround it. Although the specialized membranes of adjacent liver cells appear to extend into the lumen of the bile canaliculus in the form of microvilli, it is quite clear that no intracellular channels extend from the canaliculus into the parenchymal cytoplasm. Older studies suggested that the canalicular apparatus contained discrete walls which reinforced the histoarchitecture of the parenchymal mass. However, it is now

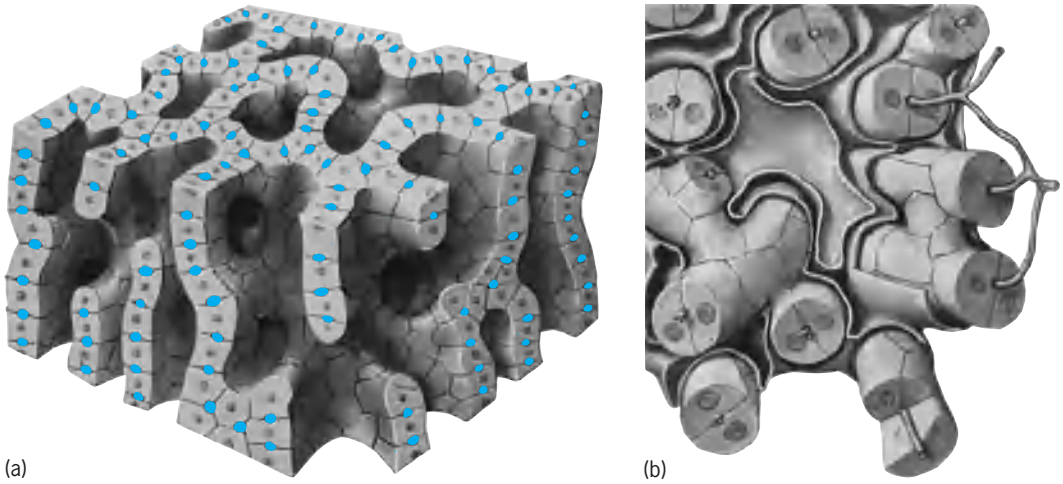

Fig. 7. Liver structure. (*a*) Stereogram of liver-cell plates after removal of ducts, vessels, and connective tissue. (*b*) A former concept of liver structure which was thought to be composed of cell cords. (*After E. Oppenheimer, ed., Ciba Collection of Medical Illustrations, vol. 3: Digestive System, pt. 3: Liver, Biliary Tract and Pancreas, Ciba, 1957*)

believed that the lumen of the canaliculus is merely an expansion of the parenchymal intercellular space, with fibrous reinforcement of the cytoplasm immediately surrounding the lumen. In the normal liver no communication exists between the canaliculi and the Disse space; under pathological conditions, however, the microvilli may swell and obstructive pressure may then lead to passage of bile into the Disse spaces. The high levels of the enzyme adenosinetriphosphatase (ATPase) found around the lumen of the bile canaliculi serve as a histochemical marker for these tiny structures.

Comparative histology. In early radiations of the vertebrate group, a parenchymal muralium two cells in thickness developed. In mammals, songbirds, and several primitive bird groups, a single-cell-thick muralium (simplex) is found. This mammalian pattern is more efficient in increasing the blood–parenchymal cell interface area and affords a more rapid exchange between blood and liver, which is appropriate to the metabolically more active homeotherms. Exchange is also facilitated by the microvilli which extend from parenchymal cell surfaces apposed to sinusoidal spaces.

Cyclostomes. This group is sufficiently different from other vertebrate groups to justify a separate categorization of their cellular arrangement: large bile canaliculi lined by parenchymal cells demonstrating polarity (structural and functional specialization in relation to the cell axis).

Fishes. The elasmobranchs (sharks) show bile canaliculi tubules of smaller bore lined by large, fat-laden cells. Some degree of polarity persists in the hepatocytes of teleosts. The lobular arrangement of higher forms, involving centripetal flow of blood through the sinusoid from peripheral vessels of portal triad to central veins, is absent. During vitellogenesis (egg development) or pregnancy in many teleosts, a marked reduction of hepatic lipid and hepatocyte volume is found, which suggests a calorie storage function for the liver in lower vertebrates.

Great variations in liver cell nuclei size have been reported in several teleosts, with some indications of a rhythmic alternation. This alternation of size may be related to the mitotic periodicities of higher vertebrate forms, such as the circadian (24-h) rhythms of the cortical liver cells in certain amphibians; namely, salamanders and newts.

Amphibians and reptiles. The capillaries between portal and hepatic channels are not arranged in typical radial fashion in the frog, so that a lobular pattern is not discernible. In addition, there are marked seasonal variations in the appearance of the liver cells and in their content of glycogen and fat. Large lymphatic channels are a characteristic of many amphibian livers, and cells within these ducts, as well as the hepatocytes themselves, are often pigmented. Similar pigmentation is encountered in the reptiles, whose livers also lack the clear-cut lobular pattern of birds and mammals.

Birds and mammals. In birds and mammals the classical lobular pattern is usually a distinguishing feature, although the absence of connective tissue boundaries may obscure this basic recurrent form of peripheral portal triads, radial sinusoids, and central veins. Some investigators maintain that the classical lobule is merely an artifact of blood flow patterns within the liver, and attempts to construct a functional unit in which the portal triad is the center of a smaller "acinar" lobule have received considerable support. In a number of mammalian species vinyl-injection techniques have provided histological evidence for such a pattern; this would also accord with the more usual glandular arrangement of secretory drainage ducts standing in the center of the lobule.

Cytology

The significant functional cell of the liver is the hepatic parenchymal cell (**Fig. 8**). The sinusoids and a small amount of connective tissue also contribute distinctive cell types.

Parenchymal cells. These are polyhedral units of at least six sides; varying in volume from less than

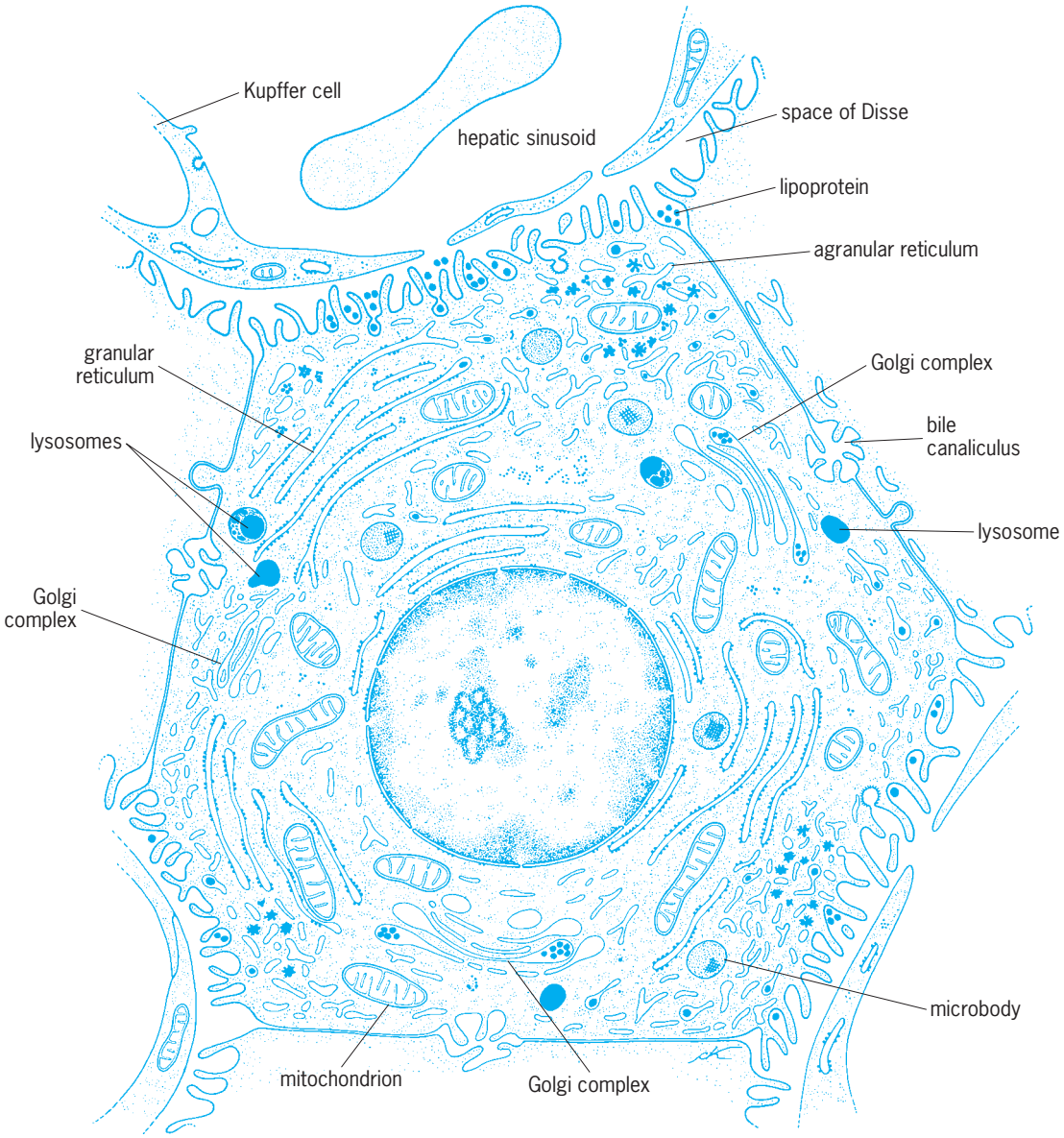

Fig. 8. Relationship of the liver cells to each other and to the sinusoids, with principal components of the hepatic cell as seen in electron micrographs. (*After W. Bloom and D. W. Fawcett, A Textbook of Histology, 9th ed., Saunders, 1968*)

10,000 to 35,000 cubic micrometers. The surfaces may be in contact with the perisinusoidal space; the lumen of a bile canaliculus; or an adjacent parenchymal cell.

The round nucleus may be quite large and variable, a consequence of the polyploidy (containing more than two sets of chromosomes) which characterizes many liver cells. Occasionally two nuclei are found. Nucleoli are prominent, and scattered chromatin clumps may be seen within the smooth nucleus. The cytoplasm varies with the nutritional state and general health of the animal, particularly in terms of glycogen space and fat vacuoles. Numerous mitochondria (filamentous) are present. A compact series of vesicles, the Golgi complex, is discernible near the bile canaliculus, where it probably acts to concentrate and "package" the bile. *See* CELL (BIOLOGY).

Electron microscopy. Electron micrographs have extended the understanding of hepatic ultrastructure. Thus both smooth- and rough-surfaced types of endoplasmic reticulum (tubular cytoskeleton) are present, with the rough-surfaced type being a major site of protein synthesis. The endoplasmic reticulum is continuous with the external membrane of the double nuclear envelope and possibly with the cytoplasmic membrane, so that a continuous system of channels may completely traverse the cell. Glycogen appears as single granules 40 nanometers thick, or in conglomerates up to 0.2 micrometer across, in close association with the smooth endoplasmic reticulum. *See* ELECTRON MICROSCOPE.

Distributed in the region of the bile canaliculus are dense rounded bodies, called lysosomes, containing several hydrolytic enzymes. The peroxisomes are similar in structure but contain crystalline inclusions rich in the enzyme uricase, which splits uric acid. They are also rich in enzymes which split hydrogen peroxide or use it to effect oxidations.

Adjacent parenchymal cells are reciprocally attached through tiny projections in one that fit snugly into recesses of the other in a manner analogous to snap fasteners.

Fractionation techniques. The region of cytoplasm immediately surrounding the bile canaliculus is fibrillar and probably gelated and may persist as tubular fragments even when the liver is macerated or homogenized. The homogenizing process involves the breakdown and dispersal of the cellular elements of a tissue to yield a homogeneous mixture for assay of enzymes, metabolites, and so forth. Differential centrifugation may then yield fractions which are particularly rich in mitochondria, other cellular organelles, or the nonparticulate cell sap. These fractionation techniques were originally developed with mammalian liver as the test material and sharply increased understanding of the biochemical patterns of liver organelles. *See* CENTRIFUGATION.

Sinusoidal lining cells. A great deal of variation is characteristic of these "littoral" cells, and even the shape and texture of the nucleus differs from one preparation to the next. Mitochondria have been noted, but not in all cells. Lipid has been demonstrated in these cells, but the existence of glycogen is in doubt. Lysosome activity has been indirectly demonstrated. The phagocytosing cells are believed to contain lysosomes and are characterized by long cytoplasmic extensions which prompted earlier investigators to call them the stellate (starlike) cells of Kupffer. Some of the living endothelial cells are vacuolated and are believed to serve as fat-storing cells, whereas others are highly granular. About 15% of the human liver cell population is made up of littoral cells.

Connective tissue. Although the liver is the largest internal organ in the body, it contains only slight amounts of supporting connective tissue, averaging less than 1%. Connective tissue is found in the capsular material, the sheath which surrounds the vessels and ducts of the portal canal system, and in lesser amounts within the hepatic canal system. *See* CONNECTIVE TISSUE.

Collagenous fibers are found within the canal systems, whereas delicate reticular fibers extend between the parenchymal cells of the muralium and the endothelial cells within the space of Disse. These fibers, appearing early in embryonic development, probably arise from the sinusoidal living cells. The connective tissue in the perisinusoidal space, which is continuous with the dense connective material of the portal sheath, is believed to function in channeling the large amount of lymph which arises as tissue fluid in the liver.

Embryology

The liver arises initially as a thickening and then an outpocketing of the ventral surface of that portion of the gut destined to become the duodenum. Two branches are quickly formed: an anterior portion from which the liver parenchyma develops; and a posterior portion giving rise to the gallbladder, its cystic duct, and the common bile duct.

Morphogenesis. The liver parenchyma develops as a series of cords or strands which are essentially tubular. At first these proliferating cords lie alongside the vitelline vessels of the yolk sac; then they are found ramified within these vessels to form an intermingled mass of epithelial cords and tiny channels, or sinusoids, representing subdivisions of the original larger venules. *See* FETAL MEMBRANE; YOLK SAC.

Alternative explanations for sinusoid development have been offered. Branching of the cords takes place in a manner characteristic of each vertebrate group and results in the varied hepatic architectures found in the adult. There may be as many as 12 basic patterns of hepatic development among vertebrates and the proliferation of hollow cords may be characteristic of only a few species.

The parenchymal cells, which are arranged in a continuous muralium (lamina) either two cells thick (lower vertebrates) or one cell thick (mammals and songbirds), are mainly derived from the endoderm. In early states the muralium may be several cells thick. The sinusoidal elements are chiefly derived from mesenchyme, but the origin of the blood-forming cells that characterize the embryonic liver is not fully known.

The bile canaliculi, which also form a continuous network in the liver, arise as a result of parenchymal cell differentiation. They are probably formed even before the liver is capable of producing bile and ramify into ductules and ducts which join the hollow posterior pouch, or future gallbladder. One contribution of electron microscopy has been the finding that the walls of the bile canaliculi do not exist as separate structures but simply represent differentiated surfaces of apposed hepatic cells.

Metabolic aspects. Although the liver plays a key role in assuring carbohydrate homeostasis (dynamic steady-state conditions) through the storage of sugar as glycogen, the embryonic livers of most species are incapable of synthesizing glycogen until a precise point is reached, rather late in development, which varies with different species. In most mammals a maximum reserve is achieved at birth, with a rapid falling off soon after. In amphibians glycogen build up is associated with metamorphosis, whereas in birds several cycles are apparent. Control is probably achieved through hormonal action of the pituitary-adrenal axis, although the role of enzyme activation has also been explored.

Much less is known about the handling of lipids, although it is clear that hepatic fat appears quite early in most mammalian embryos and is extremely high in the newborn. High lipid reserves exist in chick livers at hatching but diminish sharply thereafter. Protein incorporation, as measured by the uptake of radioactive-labeled amino acids, is extremely high in the livers of rat embryos despite the generally lower protein levels in comparison to adults. These livers also demonstrate an early competence for the synthesis of histamine.

Many of the enzymes which play a key role in the metabolic transformations of the adult mammalian liver are either low or absent at birth. The synthesis of

these enzymes at term may explain the sharp changes at birth in the levels of such metabolites as glycogen. Iron reserves also tend to increase late in embryonic life, and the storage in the liver of histologically identifiable iron is associated with the gradual diminution of the erythropoietic function. Electron microscopy studies have received an interest in the nature of the hemopoietic (blood-forming) function of the nature and origin of the blood-forming island cells which proliferate within the parenchymal elements early in embryonic life.

Physiology

The large size of the liver is matched by its functional complexity and involvement in a diverse array of regulatory mechanisms. Although said to perform over 100 specific functions, the liver's key physiological role may be delineated under seven major headings.

Control of carbohydrate metabolism. Following the ingestion of food, blood entering the liver from the hepatic portal vein is rich in glucose and other monosaccharides (simple sugars). The parenchymal cells, which are freely permeable to these sugars, remove them from the general circulation and store them as a complex insoluble polymer, glycogen. Since glycogen is a polysaccharide made up of glucose subunits, the liver must first convert other dietary monosaccharides (fructose, galactose) to glucose prior to glycogen synthesis (glycogenesis). As a result of this storage mechanism, the blood leaving the liver through the hepatic veins contains normal levels of glucose. In the intervals between ingestion of food, liver glycogen is broken down. This process, known as glycogenolysis, tends to maintain blood sugar levels between 80 and 100 mg per 100 ml of blood. *See* CARBOHYDRATE METABOLISM; GLUCOSE; GLYCOGEN; MONOSACCHARIDE.

Blood sugar level. Since the brain and other vital central nervous system structures are almost entirely dependent on the glucose fuel of the blood, hepatic regulation of circulating carbohydrate is necessary to life. Under conditions of prolonged fast, where glycogen stores are exhausted, the liver is capable of converting noncarbohydrate metabolites into glucose to maintain blood sugar levels. This process, by which amino acids, fats, glycerol, and other metabolites are converted enzymatically to glucose, is called gluconeogenesis. It represents a significant metabolic pathway for the regulation of carbohydrate metabolism by the liver, although the kidney is also capable of carrying out gluconeogenesis in mammals.

All of the processes involved in the manipulation of carbohydrate by the liver are active constantly, but the relative rates of individual pathways vary in response to food intake, tissue uptake, hormone levels, and so on. The maintenance of constancy through the dynamic balance of opposing metabolic forces is called homeostasis, a concept which was first developed a century ago by Claude Bernard in his studies of the biochemical competency of the liver. *See* HOMEOSTASIS.

Endocrine regulation. It has long been recognized that the pancreatic hormone insulin is necessary for the utilization and storage of glucose, while hormones of the adrenal cortex are involved in the gluconeogenic process. The actions of epinephrine and glucagon in enhancing hepatic glycogenolysis have been elucidated. Thus the complex steps involved in maintaining carbohydrate homeostasis are subject to endocrine control, with the liver serving as a particularly sensitive target organ for the hormone regulators. *See* INSULIN.

Enzyme systems. The precise nature of these hepatic responses are being explored, and emphasis is placed upon the role of specific enzyme systems which are rate-limiting in some overall pathway or which occupy a branch position serving to divert metabolites from one major pathway to another. *See* ENZYME.

1. *Gene activation*. George Weber showed that insulin, which is known to enhance uptake of glucose by muscle and adipose tissue, tends to increase the formation of three "bottleneck" enzymes in the liver which act to degrade glucose (glycolysis) and prepare its carbon skeletons for eventual oxidation by the mitochondria. On the other hand, insulin inhibits formation of a quartet of hepatic enzymes which act to bring the degradation products of glucose back to free monosaccharide. Hormones of the adrenal cortex have a completely opposite effect on these same target enzymes, thus explaining the opposing actions of cortisone (which increases blood sugar) and insulin (which decreases blood sugar) on carbohydrate balance. Presumably these hormones act on a specific gene which is responsible for the synthesis of a particular group of enzymes influencing a single effect. *See* GENE ACTION.

2. *Enzyme targets*. Direct action on an enzyme already present in the liver cell is also utilized. Thus epinephrine and glucagon (a pancreatic hormone that is antagonistic to insulin) each act to produce increased quantities of cyclic adenosine monophosphate (cAMP), which activates the hepatic enzyme involved in the cleavage of glycogen bonds to yield glucose. The rise in blood sugar associated with increased circulating levels of epinephrine is effected through this mechanism. Thyroxine also influences blood sugar levels, but its mode of action is unclear.

Negative feedback mechanisms. These provide a fine-control system for the hormones themselves in carbohydrate regulation. Thus insulin release is enhanced by high blood sugar levels (hyperglycemia), but its release brings into play mechanisms which tend to lower blood sugar levels and hence dampen the evoking stimulus. Long-term heavy ingestion of carbohydrate may exhaust the insulin-producing tissue of the pancreas and lead to diabetes—a complex disease usually involving insulin deficiency with concomitant disturbances of carbohydrate and fat metabolism. *See* DIABETES; ENDOCRINE MECHANISMS; LIPID METABOLISM.

In genetic obesity in mice, characterized by a seemingly paradoxical situation of high insulin levels and hyperglycemia, elevations in hepatic glycolytic

enzymes are found. This is consistent with Weber's findings of a similar hepatic glycolytic enzyme increase in animals injected with insulin. The relatively high fasting levels of liver glycogen also associated with this syndrome may be attributed to the induced synthesis of glycogen synthetase (glycogen-forming enzyme). This synthesis has been shown to occur with insulin and is related to the action of insulin in promoting carbohydrate storage in the liver. The liver possesses an enzyme, called insulinase, which degrades insulin. Its importance under physiological conditions is not clear; however, with overproduction of insulin, insulinase may prevent excessive buildup of the hormone in the bloodstream.

Interconversion of metabolites. The liver is a major site of the production of both fatty acids and neutral fats (triglycerides—three fatty acids joined to glycerol). Under conditions of adequate feeding and high carbohydrate intake, a significant proportion of the food entering the liver through the portal vein, or coming in from the systemic cirulation through the hepatic artery, will be converted to fatty acids and neutral fat. The liver possesses enzyme systems which activate both fatty acids and glycerol to facilitate their conversion to the less soluble triglycerides. In addition, the manipulation of carbohydrate in the liver leads to the formation of substrates and coenzymes (accessory molecules for enzyme action) necessary for fat synthesis, although an increasingly important role in fat metabolism is being assigned to the adipose tissue mass in mammals.

The link between the pathways of carbohydrate and fat metabolism is the probable basis for the disruption of fat synthesis and storage which occurs when carbohydrate metabolism is impaired. Synthesis of the phospholipids, which are readily transportable lipids, also occurs in the liver. The production of fatty infiltrations in the livers of animals deficient in choline or methionine may be associated with an impairment in phospholipid metabolism, since these substances are necessary for their formation. Inability to form lipoproteins has also been suggested as a factor. Agents which increase fatty acid transport to the liver or decrease hepatic fatty acid oxidation may also bring about fatty degeneration. Chronic accumulation of lipid in the liver leads to fibrotic changes and eventual cirrhosis, a condition in which the liver loses its resiliency and gradually ceases to function as the parenchymal mass is replaced by connective tissue. A variety of hepatotoxins (alcohol, chloroform, carbon tetrachloride, lead, and others) also produce fatty livers. *See* COENZYME; LIPID; LIVER DISORDERS.

Ketone bodies. Representing partial oxidation products of fatty acid degradation, these are produced by the liver and utilized as a rapid short-term energy source by extrahepatic tissues. In starvation and advanced cases of diabetes, the excess production of acidic ketone bodies poses a serious threat to life. It has been shown that hepatic glutamine production is markedly augmented in these circumstances, presumably to meet the increased needs of extrahepatic tissues. *See* KETONE.

Sterols and steroids. Cholesterol is also synthesized by the liver, although small amounts are ingested. High levels of cholesterol are believed to be associated with atherosclerosis (hardening of the arteries) by some investigators, but the relationship is a complex one. Cholesterol is excreted in the feces following its conversion to bile salts, but considerable recycling of these sterol compounds may occur between intestine and liver. Steroid hormones, which resemble cholesterol in their chemical structure, are degraded in the liver and may then be excreted in the bile. *See* CHOLESTEROL; STEROID; STEROL.

Amino acids. From the gut amino acids reach the liver through the portal circulation, and a small portion may be incorporated into hepatic structural protein to serve as a labile reserve. Protein is not generally regarded as being stored in animal tissues. During gluconeogenesis, which is enhanced by glucagon as well as the adrenal steroids, amino acids are stripped of their nitrogen and the remainder of the carbon skeleton is utilized for the production of glucose. The nitrogen arising from this deamination process may be secreted into the blood as ammonia, or the ammonia may be converted by the liver to urea or uric acid. The latter two nitrogenous end products are associated with vertebrates inhabiting marine or terrestrial environments. Nitrogen, or more specifically amino groups, may also be shifted from one organic molecule to another through the process of transamination. High levels of the enzymes mediating these transfers (transaminases) are found in the liver, and they may pass into the bloodstream during hepatic degeneration, where they serve as a clinical sign of liver damage. *See* AMINO ACIDS; CLINICAL PATHOLOGY; EXCRETION; PROTEIN METABOLISM; TRANSAMINATION.

Blood protein manufacture. The liver is the only source of plasma cholinesterase, albumin, prothrombin, and fibrinogen. Most of the globulins also arise in the liver, but gamma globulin is produced in plasma cells found principally in the spleen with only a minor contribution coming from the lymphoid tissue of the liver. On the basis of the hepatic contribution to the pool of blood proteins, a significant role may be averred for the liver in maintaining the osmotic pressure of the blood, preventing hemorrhage, providing a reserve of body protein (7 g per 100 ml of blood), promoting the transport of lipids, and preventing excess buildup of acetylcholine. *See* ACETYLCHOLINE; BLOOD; GLOBULIN; HEMORRHAGE.

Erythropoietic function. In almost all vertebrates the liver is the predominant embryonic source of red blood cells. Thrombocytes (platelets) are also produced in this organ, and in embryonic forms other than man the liver may be an important region for lymphocyte and granulocyte production as well. In urodeles the cortical region of the liver continues its erythropoietic function throughout the life cycle. *See* HEMATOPOIESIS.

Detoxification. Many poisons are rendered harmless by degradation reactions in the liver. These reactions involve conjugation, methylation, oxidation, reduction, and complex syntheses. The

microsomes (endoplasmic reticulum and ribosomes) are particularly active in neutralizing potentially harmful drugs, for example, barbiturates, tranquilizers, and hallucinogens such as lysergic acid diethylamide (LSD). Detoxification mechanisms are also involved in the formation of bile pigments, the esterification of cholesterol, and the conjugation or oxidation of steroid hormones. A diseased liver cannot carry out these functions, which lowers the body threshold to most toxic materials.

Bile formation. Bile is formed at the rate of about a quart per day. It consists of bile pigments, bile salts, protein, cholesterol, and inorganic salts. The pigments, bilirubin and biliverdin, arise from the breakdown of heme in those tissues (spleen, bone marrow) where erythrocytes are usually broken down, and then are conjugated with glucuronic acid by the liver. Bile salts and cholesterol are formed in the liver. The significance of bile release into the intestine lies in the ability of the bile salts to emulsify fats, activate lipase, and promote the uptake of the fat-soluble vitamins by the intestine. *See* BILIRUBIN.

Vitamin storage. The liver stores vitamins A and D and is apparently a reservoir for iron as well. Some of the B vitamins are also at high levels in liver tissue, suggesting a possible storage mechanism. Carotene in the liver is converted to vitamin A. *See* DIGESTIVE SYSTEM; VITAMIN A.

George H. Fried

Bibliography. D. Fawcett, *Bloom and Fawcett: A Textbook of Histology*, 12th ed., 2001; A. Francavilla et al. (eds.), *Liver and Hormones*, 1987; M. F. Heyworth and A. L. Jones (eds.), *Immunology of the Gastrointestinal Tract and Liver*, 1988; H. A. Harper et al., *Review of Physiological Chemistry*, 17th rev. ed., 1979; L. Weiss (ed.), *Cell and Tissue Biology: A Textbook of Histology*, 6th ed., 1988.

Liver disorders

A heterogeneous group of diseases that affect the liver and are of particular importance because of the many essential functions of the liver. It is a large organ (50–63 oz or 1400–1800 g), and it is involved in the metabolism of proteins, fats, and carbohydrates; in the production of bile and blood clotting substances; and in the metabolism and detoxification of many hormones and drugs. Because of the numerous functions of the liver, persons with liver disorders can develop a number of different signs and symptoms. One common symptom of liver disease is fatigue, although some people are entirely asymptomatic. A striking sign of liver disease is jaundice (yellowing of the skin and eyes) due to abnormal production or transport of bile. The presence of liver disease can be confirmed by performing liver function tests on the blood. Such tests can sometimes indicate the severity of the disease and occasionally even the exact type of abnormality. Usually, however, liver function tests can indicate only the presence of a liver abnormality. But a needle biopsy, in which a small amount of tissue is removed through a needle puncture for microscopic examination, can be done safely with minimal pain and inconvenience to the patient. A needle biopsy can help determine the type of liver disorder and the extent of organ damage. *See* JAUNDICE.

Fatty metamorphosis. This common condition is characterized by accumulation of fat (lipid) within the liver cells. The individual cells become engorged with fat, and overall liver size can increase to two or three times normal size. Persons who develop fatty metamorphosis may have no signs or symptoms except for a vague feeling of increased "fullness" in the right upper quadrant of the abdominal area, which is the location of the liver. An enlarged liver may also be diagnosed during a routine physical examination. In the United States the most common cause of a fatty liver is excessive alcohol intake, which appears to cause a change in the metabolism of the liver cells. In the developing areas of the world, malnutrition is the major cause of a fatty liver. The cause in this case is not a lack of calories but a deficiency in specific amino acids that are necessary for normal metabolism. Some obese persons also develop a fatty liver, perhaps as a manifestation of their overall excess body fat. Fatty liver caused by excess alcohol may progress to cirrhosis. *See* ADIPOSE TISSUE; ALCOHOLISM; MALNUTRITION; OBESITY.

Hepatitis. Hepatitis, which is an inflammation of the liver, has a variety of causes. The specific type of hepatitis present is usually determined by studies of serum, which identify the type of antibody directed against a component of the causative virus, or by liver biopsy. The prognosis for otherwise healthy persons who develop hepatitis is excellent: even those with hepatitis B have a greater than 85% survival rate. *See* HEPATITIS.

Cirrhosis. Cirrhosis is a disease in the liver that is characterized by scarring, which produces a marked nodularity of the liver. Cirrhosis has a variety of causes, although in the United States the most common by far is excessive alcohol intake. Cirrhosis caused by alcohol has been referred to as nutritional cirrhosis because it was originally believed to be related to a deficiency of certain dietary substances relating to the alcoholic condition. However, studies of the disease suggest that alcohol is a direct liver toxin and is itself responsible for the changes that eventually lead to cirrhosis. Cirrhosis may also result from almost any type of injury to the liver that does not heal but instead leads to progressive inflammation and scarring. Some cases of viral hepatitis progress to cirrhosis, as can obstruction of the bile duct. When the duct cannot drain bile from the liver, the result is a form of the disease that is referred to as secondary biliary cirrhosis. In general, cirrhosis is a devastating disease and has many serious complications due to increased pressure in the portal venous system, with shunting of blood through the systemic venous circulation. This shunting of blood frequently results in rupture of blood vessels near surfaces of various organs and can cause, for example, bleeding into the intestine. *See* CIRRHOSIS.

Neoplasms. Most neoplasms in the liver are malignant and are metastatic from a primary site. Because the liver receives the entire venous drainage from the intestine and thus is a common site for the spread

of tumors from the intestine. Lung and breast cancer also frequently metastasize to the liver. Primary tumors of the liver are rare in the United States, although in some parts of Africa they are the most common type of tumor. Usually their cause is unknown, but in the United States primary malignant tumors are most frequently associated with cirrhosis. Primary liver cancers are difficult to treat because they usually grow rapidly and involve other structures, preventing total surgical removal. *See* CANCER (MEDICINE); ONCOLOGY.

Uncommon or rare disorders. Reye's syndrome is a liver disorder that affects primarily infants and young children; it usually occurs during or after an episode of viral influenza and it has been causally related to the ingestion of aspirin for treatment of influenza. The effects of Reye's syndrome include a finely vacuolated fatty change in the liver. Wilson's disease, also called hepatolenticular degeneration, is a rare inherited disorder of copper metabolism in which cirrhosis of the liver is associated with degeneration of certain regions of the brain. Hemochromatosis, another hereditary disease, is marked by excessive deposition of iron in the liver because of faulty iron metabolism. It causes cirrhosis and is frequently associated with primary cancer of the liver. In alpha-1 antitrypsin deficiency, a disease characterized by an absence of an antienzyme in the blood, the liver can show extensive fibrosis and even cirrhosis. Some cases of primary cancer of the liver have been associated with an alpha-1 antitrypsin deficiency. *See* LIVER.

Samuel P. Hammar

Bibliography. J. H. Alter (ed.), Viral hepatitis, *Semin. Liver Dis.*, 6:1–95, 1986; J. R. Hodges et al., Chronic active hepatitis: The spectrum of disease, *Lancet*, 1:550–553, 1982; J. Ludwig and R. Axelsen, Drug effects on the liver: An updated tabular compilation of drugs and drug-related hepatic diseases, *Dig. Dis. Sci.*, 28:651–666, 1983; S. Sherlock, *Diseases of the Liver and Biliary System*, 10th ed., 1997; S. N. Zama et al., Risk factors in development of carcinoma in cirrhosis: Prospective study of 613 patients, *Lancet*, 1:1357–1359, 1985.

Living fossils

Living species very closely resembling fossil relatives in most anatomical details. The term is a relative one and, applied loosely, could embrace nearly all extant animals and plants. In its more restricted usage, the term applies to living species with four additional characteristics: (1) truly close anatomical similarity to (2) an ancient fossil species—generally at least 1×10^8 years old; (3) living members of the group are represented by only a single or at best a few species, which are (4) often found in a very limited geographic area. Examples are horseshoe crabs (found on the eastern shores of the two Northern Hemisphere continental landmasses, with closely similar relatives over 2×10^8 years old); ginkgo trees, dating from the Mesozoic era, which, until their recent introduction over the world, were restricted to Asia; and coelacanth fish, which have not changed much since the Devonian and which today are found only between eastern Africa and Madagascar and the Comoro Islands.

Living fossils are important because they retain many primitive traits, hence their study helps elucidate the anatomy of their fossil relatives. Living fossils also raise the issue of the lack of evolutionary change. Survival of primitive organisms in an unchanged state over many millions of years implies a low rate of evolution. Evidence suggests that slowly evolving groups speciate at a lower rate than faster-evolving groups. Such slowly evolving groups are frequently ecologically generalized, that is, able to exploit a wide variety of food resources and withstand large-scale fluctuations in the physical and chemical characteristics of their habitats. Organisms so adapted are unable to withstand competition from closely related species with similar habits. This factor reduces the number of available habitats and hence reduces the chance of new species evolving. Without speciation there can be little evolution, and as long as the habitat persists, so will the organisms, surviving virtually unchanged for millions of years. *See* SPECIATION.

Niles Eldredge

Bibliography. N. Eldredge, Survivors from the good old, old, old days, *Nat. Hist.*, 74(2):60–69, 1975; J. Martin, *Living Fossils*, 1997.

Llama

A member of the camel family (Camelidae) found only in South America. The llama is an artiodactyl, or even-toed ungulate, with two toes on each foot. The upper lip is cleft and prehensile. The animal has a

The llama of South America.

Hybrids produced by crossing different llamas

Male	Female	Hybrid
Llama	Alpaca	Huarizo
Alpaca	Llama	Misti; machurga
Vicuna	Llama	Llamavicuna
Guanaco	Llama	Llamahuanaco
Vicuna	Alpaca	Vacovicuna

long neck, and attains a maximum length of less than 8 ft (2.4 m) and a maximum weight of nearly 300 lb (135 kg; see **illus.**). A single young is born after a gestation period of about 11 months. The maximum life-span is about 20 years. Like other members of the family, the llama is herbivorous. It has 36 teeth with the dental formula I 1/3 C 1/1 Pm 3/3 M 3/3. These animals lack a gallbladder. Many interesting crosses have occurred among the different breeds in South America (see **table**). *See* ARTIODACTYLA.

Lama glama has been domesticated since the Inca civilization by the Peruvian Indians, who still keep large herds. While the fur provides material for rugs, rope, and cloth, the main use of this animal is as a beast of burden, especially in mining areas. The males are used as pack animals and the females are kept for breeding and wool production. The alpaca (*L. pacos*) is more restricted in its distribution and is specialized in wool production.

There are two species of wild llama, the guanaco or huanaco (*L. guanicoe*) and the vicuna (*L. vicugna*). The most widespread is the guanaco, which stands about 4 ft (1.2 m) at the shoulder and just over 5 ft (1.5 m) at the top of the head. The guanacos live in small herds composed of a male, several females, and the young. It was formerly thought that this stock represented the forerunner of the domestic breeds. The vicuna is becoming rare because of overshooting for its pelt of soft wool. *See* ALPACA; VICUNA.

Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, 6th ed., 1999.

Loads, dynamic

Forces which are derived from moving loads such as wind, earthquakes, machinery, vehicles, trains, cranes, and hoists. Analysis techniques which take into account the vibrations of the structures are required for loads which are repeated many times, such as machinery in motion, and produce harmonic motions of equal amplitude and constant frequency (cyclic loading); loads such as the motion produced by earthquakes (random motion); and varied loads, such as that of the wind, which produce gusts or short-duration impulses.

Repeated loads applied to a structural member can cause failure by fracture of the material. This fracture can occur at various stress levels depending upon the amplitude or acceleration of the motion, frequency, and duration. Often the stress level is below that of the design level for statically applied loads, and is referred to as the fatigue strength of the material for that application. The fatigue strength may vary for different applications of dynamic loadings and configurations of the structures or components. The fatigue strength will be lowered in the presence of localized stress concentrations.

Cycle of stress. The variation of stress S during one typical fluctuation when the load is repeatedly applied and removed or when the load fluctuates continuously from tension to compression is a stress cycle (see **illus.**). A fluctuating or pulsating stress varies between maximum and minimum values of the same sign; this condition is called repeated if the stress reduces to zero in each cycle. Reversed stress varies between values of opposite signs and is partly reversed when the opposite stresses are unequal, and completely reversed when the opposite stresses are equal.

Stress variation is described either by the maximum stress, its type (tension or compression), and the range of fluctuations, or by the mean or steady stress (S_m in illus.) together with the magnitude of the superimposed alternating stress producing the cyclic variation. Repeated loads superimposed on a constant load, such as dead weight of the structure, may produce fluctuation without reversal. Engine crankshafts and parts, rails, and axles are subjected to the reversed-type cyclic loading, which may involve billions of repetitions in the life of the member.

Stress raisers. Dimensional changes, internal discontinuities, or other conditions causing localized disturbance of the normal stress distribution, with increase of stress over that normally produced, are termed stress raisers. The maximum stress is called a stress concentration. High local stresses exist at bolt threads, abrupt change of shaft diameter, notches, holes, keyways, or fillet wells. Internal discontinuities, such as blowholes, inclusions, seams or cracks, ducts and knots in wood, voids in concrete, and variable stiffness of component constituents, cause stress concentrations. At points of contact of ball or roller bearings, gear teeth, or other local load applications, the stress may be greatly increased. Maximum stress is obtained by multiplying the nominal stress, computed without regard to the modifying effect of the stress raiser, by a stress concentration factor. This factor, found experimentally, depends on the type

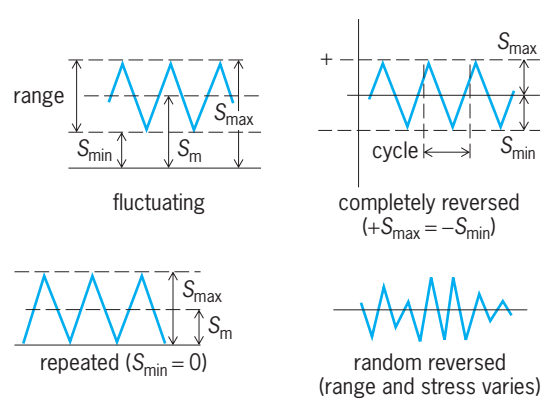

Types of repeated stress.

of discontinuity and the properties of the material under static load. Ductile material relieves most of the stress concentration by plastic yielding. Stress raisers contribute to failure under high stress rates and low temperatures, which tend to inhibit plastic flow, and are of great importance under repeated loads, which produce progressive fracture, called fatigue failure. *See* STRESS AND STRAIN. John B. Scalzi

Loads, transverse

Forces applied perpendicularly to the longitudinal axis of a member. Transverse loading causes the member to bend and deflect from its original position, with internal tensile and compressive strains accompanying change in curvature.

Concentrated loads are applied over areas or lengths which are relatively small compared with the dimensions of the supporting member. Often a single resultant force is used to analyze effects on the member. Examples are a heavy machine occupying limited floor area, wheel loads on a rail, or a tie rod attached locally. Loads may be stationary or they may be moving, as with the carriage of a crane hoist or with truck wheels.

Distributed loads are forces applied completely over large areas with uniform or nonuniform intensity. Closely stacked contents on warehouse floors, snow, or wind pressures are considered to be uniform loads. An equivalent uniform load may represent multiple closely spaced, concentrated loads. Variably distributed load intensities include foundation soil pressures and hydrostatic pressures.

Bending and shear. Transverse forces produce bending moments and transverse (shearing) forces at every section which must be balanced by an internal couple or resisting moment M and an internal shear force V (**Fig. 1***a*). These component forces are evaluated by free-body analysis. The bending moment and resultant shear force can be expressed analytically as functions of the loads and the distance locating the section from a reference origin. The magnitude and distribution of internal stresses associated with the external force system are calculated by the theory of flexure.

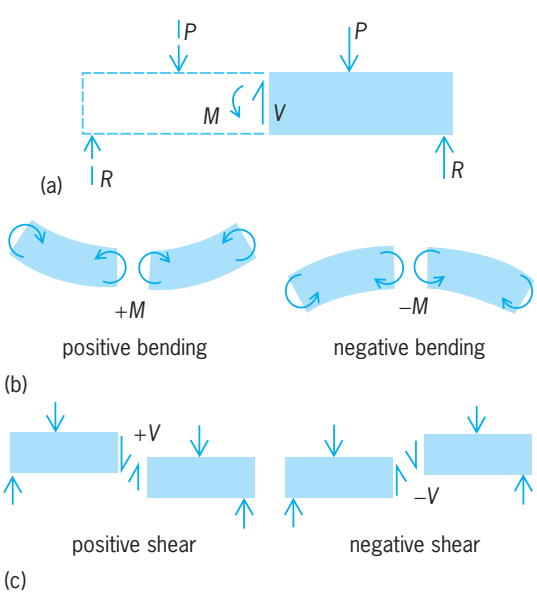

Fig. 1. Beam loading analysis. (*a*) By free-body analysis of beam. (*b*) Conventions of sign for bending moment. (*c*) Conventions of sign for shear.

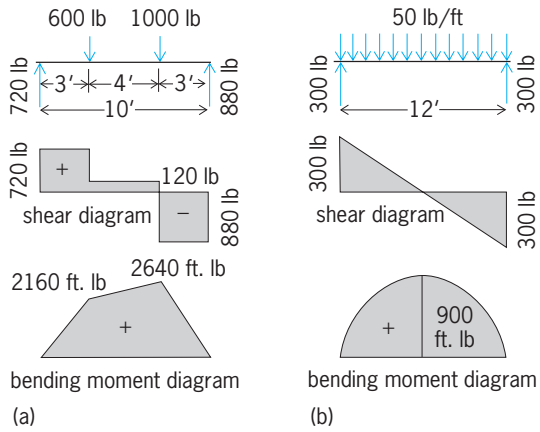

Fig. 2. Typical shear and bending moment diagrams for (*a*) concentrated and (*b*) uniformly distributed loads. 1 ft = 0.3 m; 1 lb = 4.45 N; 1 ft-lb = 1.36 N-m.

Bending moment diagram. A graphical representation indicates the bending moment produced at all sections of a member by a specified loading. The bending moment is the sum of the moments of external forces acting to the left or right of the section, taken about the section. The sign of the curvature is related to the sign of the bending moment (Fig. 1*b*).

Typical bending moment and shear diagrams are shown in **Fig. 2** for concentrated and uniformly distributed loads on a simply supported beam. The reactions and moments are calculated for equilibrium. *See* STATICS.

Shear diagram. A graphic representation of the transverse (shearing) force at all sections of a beam produced by specified loading is called a shear diagram. The shear at any section is equal to the sum of the forces on a segment to the left or right of the section considered (Fig. 1*c*). Positive shears are produced by upward forces on the left section (Fig. 2).

Relationships between shear and bending moment that assist in construction of diagrams are the following: (1) maximum moment occurs where shear is zero; (2) area in the shear diagram equals change of bending moment between sections; (3) ordinates in the shear diagram equal slope of moment diagram; (4) shear is constant between concentrated loads and has constant slope for uniformly distributed loads. Combined loading can be represented by conventional composite moment diagrams or by separate diagrams, called diagrams by parts.

Beams. Members subjected to bending by transverse loads are classed as beams. The span is the unsupported length. Beams may have single or multiple spans and are classified according to type of support, which may permit freedom of rotation or furnish restraint (**Fig. 3**).

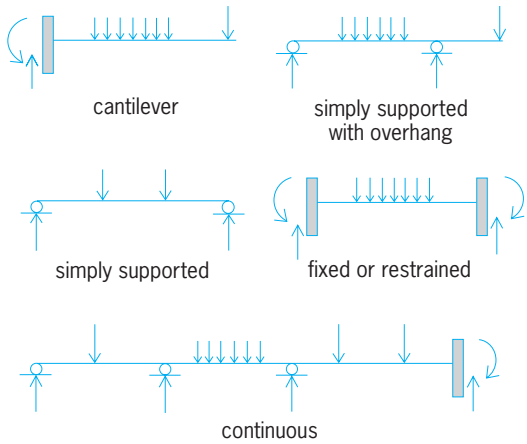

Fig. 3. Diagrams showing types of beams.

The degree of restraint at supports determines the stresses, curvature, and deflection. A beam is statically determinate when all reaction components can be evaluated by the equations of statics. Two equations are available for transverse loads and only two reaction components can be found. Fixed-end and continuous beams are statically indeterminate, and additional load deformation relationships are required to supplement statics.

Bending stresses are the internal tensile and compressive longitudinal stresses developed in response to curvature induced by external loads. Their magnitudes depend on the bending moment and the properties of the section. The theory of flexure assumes elastic behavior. No stresses act along a longitudinal plane surface within the beam called the neutral surface. A segment subjected to constant bending moment, as when bent by end couples, is in pure bending, and stresses vary linearly across the section which remains plane. Simultaneous shear causes warping and nonlinear stress distribution. The flexure stress is maximum at boundary surfaces of sections where bending moment M is greatest. Maximum stress σ at distance c from the neutral axis to the extreme element of a section having a moment of inertia I is $\sigma = Mc/I = M/S$, where $S = I/c$ is called the section modulus, a measure of the section depending upon shape and dimensions. Greatest economy results when the section provides the required section modulus with least area. The common theory of flexure applies only to elastic stresses. Stresses exceeding the elastic limit involve plastic strains producing permanent deflections upon load removal.

Inelastic stresses are first produced at surface elements of sections resisting maximum moment. A distinct yield point in materials, such as structural grade steel, causes stresses near the surface to remain constant, while interior stresses increase with increasing load. Redistribution of stress continues until the entire section behaves plastically. The fully plastic resisting moment of a standard rolled, wide-flange section is about 12% greater than the moment just producing yield point at the surface. Where small increases in deflection can be tolerated, this plastic-hinge moment is used in design.

Shearing stresses are intensities of the shear force distributed over a beam cross section. At any point in the beam, the vertical and longitudinal shear stresses are equal. They are zero at the boundary surfaces and maximum at the neutral axis. The shear stress according to the common theory of flexure is $S_s = VQ/Ib$, where V is total shear force, Q is statical moment about the neutral axis of that area of the beam section between the level of investigation and the top or bottom of the section, I is moment of inertia, and b is width. In rectangular sections the distribution is parabolic, and maximum S_s is one and one-half times the average stress.

Shear flow is the shear force per unit length acting along component areas of the section. The concept of flow is derived from the distribution of unit shear forces represented by vectors pointing in the same direction along the median line of a thin element, and the similarity of the expression for unit shear to the equation for quantity of liquid flowing in a pipe or channel. The direction of unit shear forces acting in typical structural shapes is shown in **Fig. 4**. *See* TORSION.

Deflection. A beam is said to undergo deflection when the displacement is normal to the original unloaded position and is due to curvature produced by loads. Deflections may be limited to less than structurally safe values in buildings to avoid cracking the ceiling or in machines because of necessary clearances.

Elastic curve is the curved shape of the longitudinal centroidal surface when loads produce elastic stresses only. Deflection at any point is calculated from the equation of this curve. Shear forces contribute additional deflection.

Radius of curvature depends on the moment and stiffness of the beam and is constant, with bending to a circular arc, when the moment is constant. This is simple bending. Curvature varies under variable moment.

Statically indeterminate beam. In the presence of more reaction components than can be determined by the equations of statics alone, a beam is statically indeterminate. The first step in analysis is to find the reactions. Equations of statics are supplemented by additional equations relating external forces to slope and deflection.

The curvature and accompanying stresses depend on the loads and the restraint imposed by the

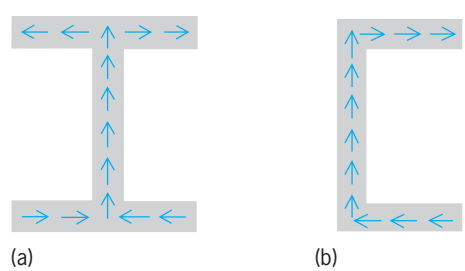

Fig. 4. Shear flow in (*a*) a wide-flange section and (*b*) a channel section.

supports. In a simply supported beam, only vertical displacement is prevented at the supports, whereas when the beam is fixed, rotation is also prevented at the ends. Similarly, beams with overhanging ends and beams continuous over multiple supports are partially restrained. Introduction of the boundary conditions describing the conditions imposed by the supports in the solution of the basic differential equation for the elastic curve furnishes relations between the loads, reaction components, and properties of the beam. For elastic behavior these relations, together with the equations of statics, evaluate the indeterminate reactions. Inelastic behavior is caused by plastic deformations at critical sections of high flexural stress when the yield point of the material is exceeded. When a fully plastic condition is reached, the section acts as a plastic hinge, with constant resisting moment during unrestrained rotation. These known moments at supports or intermediate points permit determination of the reactions by the equations of statics. Advantage is taken of this inelastic action in plastic design.

Unsymmetrical bending. Loads that produce bending moments in planes oblique to the principal axes of a section cause unsymmetrical bending. Components of the applied moment produce curvature and deflection in planes perpendicular to these bending axes; the stresses produced by these components acting separately can be superimposed. To avoid torsional twist, the oblique forces must pass through either the geometric center of symmetrical sections or a particular point called the shear center for unsymmetrical sections.

Lateral buckling of beams is caused by instability of the compression flange, which results in lateral deflection and twist of the section. The action corresponds to buckling of long columns. The critical compressive stress producing buckling depends on the sectional dimensions and type of loading, which in turn determine the flange stresses and torsional rigidity of the section. Beams having small moments of inertia about the weak axis with low torsional resistance are susceptible to buckling. For rolled beams, buckling resistance is related approximately to the Ld/A_f ratio, where L is span, d is beam depth, and A_f is the area of the flange.

Buckling of beam webs is caused by excessive diagonal compression induced by the shear stresses in the web or by the combined shear and bending stresses in the web acting as an edge-supported plate. Because the web is relatively thin, it tends to buckle laterally in a series of waves. For deep beams the shear stress must be reduced by increased web thickness or reinforcement in the form of stiffener angles or plates added to increase resistance to lateral buckling.

Variable cross sections make more efficient use of material in beams by providing resisting moment at all sections more nearly equal to the applied moment. If the section modulus varies directly as the bending moment, stress M/S will be constant at all sections, and the beam is said to have constant strength. Dimensions of the section can be varied by tapering the width and depth or by adding cover plates to a built-up section to vary the section modulus. Common examples of beams with variable section are tapered cantilevers, plate girders with multiple cover plates, forgings, and other machine elements. A leaf spring is a tapered beam in which uniform stress increases energy absorption. *See* PLATE GIRDER; SHEAR CENTER.

Beams on elastic foundation. These members are subjected to transverse loads while resting either on a continuous supporting foundation or on closely spaced supports which behave elastically. The curvature, deflection, and bending moment depend upon the relative stiffness of the beam and supporting foundation. Beams of this type include railroad rails on ties, a timber resting on level ground, a long pipe supported by closely spaced hanger rods or springs, and a concrete footing on a soil foundation. The unknown reactive force per unit length of beam is assumed to be proportional to the deflection, and the solution of this statically indeterminate problem requires derivation of the equation of the elastic curve whose differential equation is $EI(d^4y/dx^4) = -ky$, where k is the spring constant per unit length of support and ky is the elastic force exerted per unit length of beam. Solution of this equation determines the distribution of reactive forces from which the shear and moments can be calculated. *See* STRENGTH OF MATERIALS.

John B. Scalzi

Bibliography. S. Berkowitz, *Introduction to Structural Analysis*, 2d ed., 2000; R. N. White, P. Gergely, and R. G. Sexsmith, *Structural Engineering*, vol. 1: *Introduction to Design Concepts and Analysis*, 2d ed., 1976.

Lobata

An order of the phylum Ctenophora (comb jellies) comprising the families Bathocyroidae, Bolinopsidae, Eurhamphaeidae, Kiyohimeidae, Leucotheidae, and Ocyropsidae. Lobate ctenophores are characterized by large winglike lobes on the oral end of the body that are used for capturing food. All species are predators on zooplankton. Lobates are among the largest ctenophores; some epipelagic species attain an oral-aboral height of 8 in. (20 cm), and some deep-sea species have lobes nearly 39 in. (1 m) across. Most species are transparent, but some have conspicuous brown, blue, or purple spots on the lobes. Deep-sea forms often have red pigment surrounding the stomodeum. Lobates are bioluminescent, like other ctenophores. Bright blue-green light is produced in the meridional canals, and some species also release a cloud of luminous material into the water when disturbed, probably as part of an escape response.

The upper body is usually compressed in the tentacular plane. The comb (ctene) rows in this plane are short, but the substomodeal comb rows extend onto the outer surface of the lobes. Meridional canals are long and interconnected; the substomodeal canals often make complex loops in the

lobes that are characteristic for different species. Pendant structures called auricles extend from the oral end of the body into the space enclosed by the oral lobes. Auricles may be short and stubby or long and sinuous in different species. Locomotion is produced by the comb rows, but some genera (*Ocyropsis*, *Bathocyroe*) can also swim by flapping the oral lobes.

Except for the genus *Ocyropsis*, all lobate ctenophores are simultaneous hermaphrodites, and have ovaries and testes situated along the meridional canals. The gonads open to the outside by way of numerous gonopores along the comb rows, and many species are thought to be self-fertilizing. Coastal species can produce thousands of eggs per day, but oceanic and deep-sea species probably produce fewer. Eggs develop into the cydippid larva typical of most ctenophores. After a variable length of time of swimming and feeding as a cydippid, the larva metamorphoses into the adult form, passing through intermediate larval stages in some species (for example, *Leucothea*, *Mnemiopsis*). Metamorphosis from the cydippid form involves loss of the primary tentacles, development of the oral lobes and auricles, and growth of secondary tentacles near the mouth.

In most lobate ctenophores, feeding is accomplished by swimming slowly forward with the lobes spread open around the mouth. The secondary tentacles give off numerous fine side branches that cover the inner surface of the lobes. Movement of the auricles may startle the crustacean copepods and other small zooplankton so that they try to escape but instead swim into an outspread oral lobe. On contact, the lobe closes over the prey, which becomes tangled in the tentilla and is then transported to the mouth. In some species both lobes contract to enclose the prey, and in others the lobes operate independently. Species of *Ocyropsis* lose their tentacles entirely during metamorphosis and rely on their muscular lobes alone to catch and hold prey before transferring it to the mouth. This mechanism enables them to catch larger and more active prey than other species. Many deep-sea species have exceptionally large, flimsy oral lobes, presumably to increase the chance of capturing sparsely distributed prey.

Lobate ctenophores are found in both coastal and oceanic waters at all latitudes and depths. In coastal waters and estuaries, their high fecundity, efficient feeding, and rapid growth rates can lead to extremely dense seasonal populations that have a significant impact as predators on smaller zooplankton and competitors for young fish. Lobate ctenophores are parasitized by hyperiid amphipods and preyed upon by medusae, the ctenophore *Beröe*, and various fishes. *See* CTENOPHORA; CYDIPPIDA. Laurence P. Madin

Bibliography. G. R. Harbison, L. P. Madin, and N. R. Swanberg, On the natural history and distribution of oceanic ctenophores, *Deep-sea Res.*, 25:233–256, 1978; F. W. Harrison and J. A. Westfall (eds.), *Microscopic Anatomy of Invertebrates*, vol. 2, 1991; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 1982.

Lobosia

A subclass of Rhizopodea characterized by lobopodia predominantly, although certain of these protozoan species also may form slender pseudopodia, or even develop several different kinds. Lobosia are divided into two orders which differ in presence or absence of a test. The Amoebida include the amebas, which do not produce a test. The Arcellinida possess a one-chambered test ranging from a predominantly chitinous to a mainly siliceous covering. Pseudopodia emerge through the aperture of the test. *See* AMOEBIDA; ARCELLINIDA; PROTOZOA; SARCODINA; SARCOMASTIGOPHORA. Richard P. Hall

Local-area networks

Computer networks that usually cover a limited range, say, within the boundary of a building. A computer network is two or more computers that communicate with each other through some physical medium. The primary usage of local-area networks (LANs) is the sharing of hardware, software, or information, such as data files, multimedia files, or electronic mail. Resource sharing provided by local-area networks improves efficiency and reduces overhead. *See* DIGITAL COMPUTER; ELECTRONIC MAIL; MULTIMEDIA TECHNOLOGY.

A local-area network can be thought of as similar to a telephone network which uses phone wires installed in a star topology. A local-area network must also use some transmission medium installed with a selected topology.

Transmission media. Four basic types of media are used in local-area networks: coaxial cable, twisted-pair wires, fiber-optic cable, and wireless. Each medium has its advantages and disadvantages relative to cost, speed, and expandability. Originally, most local-area networks used coaxial cables. A coaxial cable consists of one or two conducting wires encapsulated by several layers of insulation and shielding (**Fig. 1***a*). Coaxial cables provide high speed and low error rates. Twisted-pair wires (Fig. 1*b*) are cheaper than coaxial cables, can sustain the speeds common to most personal computers, and are easy to install. Fiber-optic cable (Fig. 1*c*) is the medium of choice for high-speed local-area networks, operating at speeds of 100 megabits per second or higher. A fiber-optic cable uses light pulses to represent data. Because light signals are not distorted by electric or magnetic fields, fiber-optic cables provide excellent error characteristics. The disadvantages of fiber-optic cables include their high cost and the difficulty of adding or removing stations. Messages in wireless local-area networks are transferred through the air as radio waves rather than through a conductive cable or wire. Hence, wireless local-area networks have the advantage of expandability. *See* COAXIAL CABLE; COMMUNICATIONS CABLE; FIBER-OPTIC CIRCUIT; OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

Topology. The topology of a local-area networks is the physical layout of the network. For wired

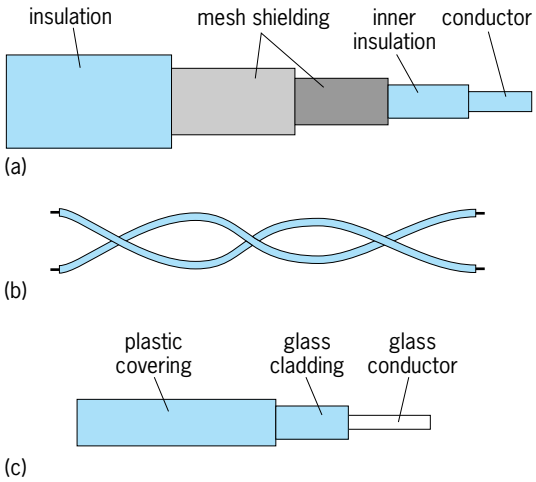

Fig. 1. Media used in local-area networks. (*a*) Coaxial cable. (*b*) Twisted-pair wires. (*c*) Fiber-optic cable.

local-area networks, there are four basic topologies: bus, ring, star, and mesh (**Fig. 2**). The most widely used local-area network topology is the bus (Fig. 2*a*), where the medium consists of a single wire or cable to which nodes are attached. A message transmitted over a bus propagates in both directions along the bus, passing each tap until it is finally absorbed at the ends. In a ring topology (Fig. 2*b*), the medium forms a closed loop, and all stations are connected to the loop or ring. Data propagate in one direction. The star topology (Fig. 2*c*) has a central node to which the computers and devices are separately connected. All the network traffic passes through the central node. In a mesh topology (Fig. 2*d*), messages are transmitted through several switches that connect nodes. The asynchronous-transfer-mode (ATM) technology favors the use of mesh topology. Modifying and combining some of the characteristics of these basic network topologies may result in hybrid topologies, which often provide greater network efficiency.

Circuit versus packet switching. There are a number of ways in which nodes can communicate over a network. The simplest is to establish a dedicated link between the transmitting and receiving stations. This technique is known as circuit switching. In the topologies outlined above, this would involve dedicating the medium to communication between the nodes. However, data communications over a local-area network tend to be bursty; that is, they are characterized by periods of intensive data transfer, followed by lulls. During this period of inactivity, precious bandwidth capability is being wasted.

A better way of communicating is to use a technique known as packet switching, in which a dedicated path is not reserved between the source and the destination. Data are wrapped up in a packet and launched into the network. In this way, a node only has exclusive access to the medium while it is sending a packet. During its inactive period, other nodes can transmit. Thus, the available bandwidth can be better utilized. The problem of sharing access to the network is reduced to establishing some rules that will allow each node on the network to launch its packet in a fair manner. This is achieved by access protocols, discussed below. *See* PACKET SWITCHING.

Packet format. A typical packet is divided into preamble, address, control, data, and error-check fields. The field of the preamble contains some bit sequence that never occurs in normal data. It serves to inform all other nodes on the network that transmission of a packet has begun. The address field contains addresses of both the sender node and receiver

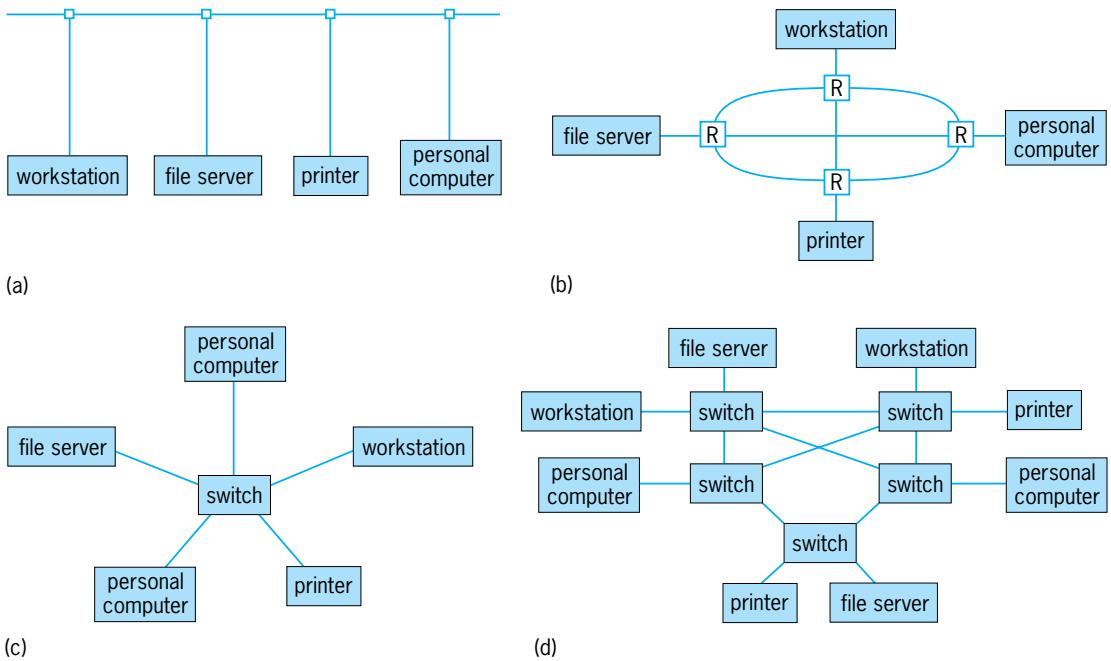

Fig. 2. Local-area network topologies. (*a*) Bus. (*b*) Ring; R = ring repeater. (*c*) Star. (*d*) Mesh.

node. This allows the nodes on the network to recognize where the packet comes from and who should receive it. The control field identifies the purpose of the packet as normal data transmission or a special management purpose. Although a packet usually contains normal data, it can also be sent for a network management purpose (for example, to discover the status of nodes and medium). The data field contains the actual data to be transferred. Although the error rate in a local-area network is typically very low, a packet usually has an error-check field that allows the receiver node to detect or recover transmission errors. The sizes of all these fields are usually fixed except for the data field, which may have a variable length, depending on the amount of data to be transferred.

Access protocols. An access protocol is a set of rules observed by all the nodes in a local-area network so that one node can get the attention of another and its data packet can be transferred. Two common protocols are carrier sense multiple access with collision detection (CSMA/CD) and token passing.

With the CSMA/CD protocol, a node that wants to transmit its data must first listen to the medium to hear if any other node is using the medium. If not, the node may transmit immediately. However, while the transmission is taking place, the transmitting node must continue listening to ascertain if anyone else has begun transmitting. This is possible because, with the propagation delay over the medium, two or more nodes could simultaneously decide that the medium is idle and start transmission virtually at the same time. Then, data transmitted by the nodes collide and get garbled. Hence, if the transmitting node detects that someone else is also transmitting, the node aborts its own transmission, waits for a random amount of time, and then restarts the process until its data transmission succeeds.

With the token-passing protocol, the right to transmit is granted by a token, a predefined bit pattern that is recognized by each node. The token is passed for one node to another in a pre-determined order. For example, in a ring network the token passes from one node to its adjacent node. When a node obtains the token, it has two options: it can transmit a message, if any, or it can pass the token to the next node. If the node transmits a message, then at the end of its transmission it releases the token to the next node. When a token network starts up, an initial token must be generated, and when the token is damaged because of node and transmission errors, a token must be regenerated. This function is usually carried out by a designated monitor node. If the token does not arrive at the monitor node within a certain time, it generates a new token. this technique works because the token traveling time around the network is usually predictable.

Software. The success of the local-area network depends on how application and system software interacts with the hardware in setting up the communication capability. Application software is designed to solve users' problems. It is assisted in this goal by supporting system software. Local-area network system software is essentially an extension of the operating system. It insulates applications from hardware details and carries out hardware-oriented local-area network tasks, such as interfacing to the medium, directing printing jobs, and sending and receiving service requests. That is, the applications can make requests for services, and system software contains the logic to carry out these requests for a specific type of hardware. *See* OPERATING SYSTEM; SOFTWARE.

Network services. The primary usage of local-area networks is resource sharing. The resources can be information, such as data files, multimedia files, electronic mail, or software. They can also be hardware devices such as printers, plotters, scanners, or storage devices. Resource sharing is realized by services provided by the local-area network. Usually some nodes in the network function as file servers. File services provide the central repository of data or application programs, and offer consistent and easy-to-use file access to all the users and applications. Management services provide uniform interfaces so that a small central staff can monitor and control end-user activities, network changes, security, and system performance across an entire corporate network. Security services aim at establishing consistent security across a variety of network resources. Usually these services are integrated with other network services and are enforced transparently with every network activity. Time services synchronize enterprise-wide services and applications, making network events temporally consistent. *See* COMPUTER PERIPHERAL DEVICES; COMPUTER SECURITY.

Connection-oriented communications. While the packet-switching method is cost-effective and acceptable for many common applications such as electronic mail and file transfer, it does not provide a prior guarantee regarding the quality of service that an overlying application may receive. Hence, it is insufficient for applications such as multimedia and teleconferencing. *See* TELECONFERENCING.

Connection-oriented communication technology has been developed to improve the packet-switching method. A connection is an abstract communication service provided by the network to a particular application. It can be considered as a two-way contract: The communicating applications specify the characteristics of traffic which they may generate (say, the maximum rates), and the network agrees to provide the requested quality of service to the applications. The network may disallow a connection if it cannot provide the requested quality-of-service guarantees. Thus, a connection may be viewed as a virtual link that is dedicated for the use of an application and that has a certain traffic-carrying capacity. Connection-oriented communication technology exploits the advantages of packet switching while overcoming its weaknesses.

Development of networks. A representative of the first generation of local-area networks is the Ethernet, a CSMA/CD bus. It uses coaxial cable with a maximum data rate of 10 megabits per second. The best-known second-generation local-area network is

the Fiber Distributed Data Interface (FDDI), a token-ring network using optical fibers with a data rate of 100 megabits per second.

Another local-area network technology, which uses the asynchronous transmission mode, is ATM LAN. An ATM LAN is a mesh-based network using the optical-fiber medium with data rates of the order of 150 megabits to 10 gigabits per second. With the high bandwidth provided by ATM LANs, many applications such as integrated real-time data and multimedia services become possible.

The basic premise of virtual local-area networks is to separate logical work groups from physical constraints. This enables devices that are connected to different components throughout the network to be logically grouped into communities of interest. Thus, virtual local-area networks are built by administrative and functional needs rather than attachment. Virtual local-area network technology provides scalability, flexibility, and policy-based management. *See* DATA COMMUNICATIONS. Wei Zhao

Bibliography. N. Malcolm and W. Zhao, Hard real-time communication in multi-access networks, *J. Real-time Syst.*, 8(1):35–77, January 1995; W. Stallings, *Data and Computer Communications*, 6th ed., 1999; A. Tanenbaum, *Computer Networks*, 3d ed., 1996.

Local Group

The small cluster of galaxies that contains the Milky Way Galaxy. Galaxies exhibit a pronounced tendency to clump together on a variety of scales. It is assumed that gravitational attraction draws galaxies together. On a very large scale, this attractive process may still be at an early stage, resulting in filamentary structures which are only mildly perturbed in their motions from the general expansion of the universe. On a smaller scale, the process has led to collapse. Galaxies have fallen together, though with enough angular momentum that they usually orbit each other rather than collide.

Collapsed structures that contain a few dozen to a few thousand substantial galaxies are called clusters; collapsed structures that contain a few but less than 10 or 20 big galaxies are called groups. The Milky Way Galaxy is a large but not exceptional spiral galaxy and one of two dominant members of a small assemblage referred to as the Local Group.

Environment. On successively larger scales, the Local Group is a member of the Coma-Sculptor Cloud and the Local Supercluster. The Coma-Sculptor Cloud contains several hundred galaxies in an extremely flattened distribution that extends across 4×10^7 light-years (1 light-year equals 9.46×10^{12} km or 5.48×10^{12} mi), in the general direction toward the Virgo cluster at the center of the Local Supercluster. The dominant members of the nearest-neighboring groups (see **illus.**) are New General Catalogue (NGC) 253, Maffei 1, Index Catalogue (IC) 342, Messier (M) 81, and the prominent source of radio emission, Centaurus A. The closest galaxies

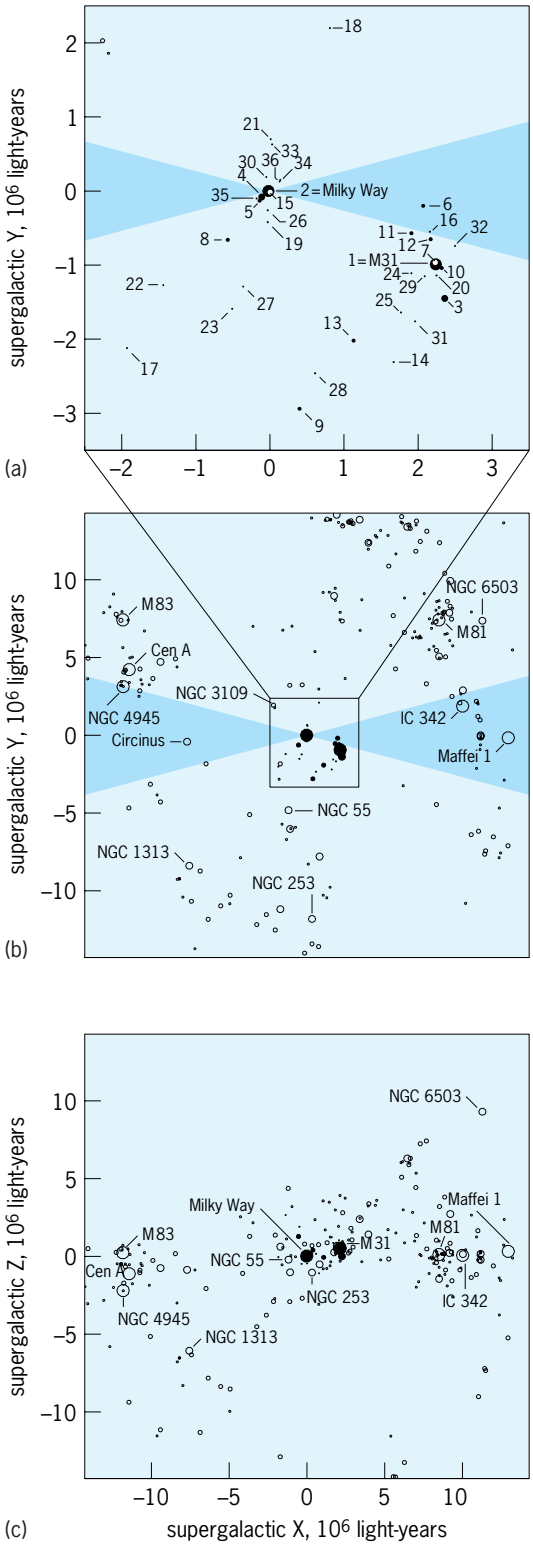

Neighborhood of the Local Group. Shaded wedges indicate the zone partially obscured by the Milky Way Galaxy. Filled circles denote members of the Local Group. Open circles locate galaxies beyond the Local Group. Members of the Local Group are identified by the running number given in the table. Several prominent galaxies beyond the Local Group are identified by name. (*a*) Blowup of the Local Group. (*b*) View down onto supergalactic equatorial plane ($Z = 0$) of the region around the Local Group. (*c*) View edge-on to the supergalactic plane, showing that most nearby galaxies lie near this plane.

Local Group Galaxies*

Name	Type	Distance, 10^6 light-years	Luminosity (X Sun)
1. M31 = Andromeda	Spiral	2.5	3×10^{10}
2. Milky Way Galaxy	Spiral	—	2×10^{10}
3. M33 = Triangulum	Spiral	2.8	6×10^{9}
4. Large Magellanic Cloud	Irregular	0.16	3×10^{9}
5. Small Magellanic Cloud	Irregular	0.19	7×10^{8}
6. IC 10	Irregular	2.2	3×10^{8}
7. NGC 205	Elliptical	2.5	3×10^{8}
8. NGC 6822	Irregular	1.6	2×10^{8}
9. WLM	Irregular	3.0	2×10^{8}
10. M32	Elliptical	2.6	2×10^{8}
11. NGC 185	Elliptical	2.0	2×10^{8}
12. NGC 147	Elliptical	2.3	1×10^{8}
13. IC 1613	Irregular	2.3	9×10^{7}
14. Pegasus	Dwarf irregular	3.1	3×10^{7}
15. Sagittarius	Dwarf spheroidal	0.08	2×10^{7}
16. Cassiopeia	Dwarf spheroidal	2.5	1×10^{7}
17. Tucana	Dwarf spheroidal	2.6	9×10^{6}
18. Leo A	Dwarf irregular	2.6	8×10^{6}
19. Fornax	Dwarf spheroidal	0.48	7×10^{6}
20. Andromeda I	Dwarf spheroidal	2.6	6×10^{6}
21. Leo I	Dwarf spheroidal	0.83	6×10^{6}
22. SagDIG	Dwarf irregular	3.4	5×10^{6}
23. DDO 210	Dwarf irregular	2.6	4×10^{6}
24. Andromeda II	Dwarf spheroidal	2.2	4×10^{6}
25. Andromeda VI	Dwarf spheroidal	2.6	3×10^{6}
26. Sculptor	Dwarf spheroidal	0.26	2×10^{6}
27. Phoenix	Dwarf irregular	1.4	2×10^{6}
28. Cetus	Dwarf spheroidal	2.5	2×10^{6}
29. Andromeda III	Dwarf spheroidal	2.4	1×10^{6}
30. Sextans	Dwarf spheroidal	0.26	1×10^{6}
31. LGS 3	Dwarf irregular	2.6	8×10^{5}
32. Andromeda V	Dwarf spheroidal	2.6	7×10^{5}
33. Leo II	Dwarf spheroidal	0.65	6×10^{5}
34. Ursa Minor	Dwarf spheroidal	0.22	4×10^{5}
35. Carina	Dwarf spheroidal	0.35	4×10^{5}
36. Draco	Dwarf spheroidal	0.26	3×10^{5}

*Ordered from brightest to faintest.

beyond the Local Group are at a distance of 7×10^6 light-years.

Members. The Local Group has 36 known or suspected members (see **table**). The only system larger than the Milky Way Galaxy is M31, the Andromeda Nebula. These two giant galaxies generate 80% of the light of the group. There are also two intermediate-scale galaxies: M33, the Triangulum Nebula near Andromeda; and the Large Magellanic Cloud, a close companion of the Milky Way and a conspicuous feature of the night sky in the Southern Hemisphere. Fainter than these are 32 small systems. The smallest among them are identified only because they are so close. Many would not be detected in even the nearest adjacent groups. The census of Local Group members may be very incomplete at the faint end and in the zones obscured by the plane of the Milky Way. *See* ANDROMEDA GALAXY; MAGELLANIC CLOUDS; MILKY WAY GALAXY.

Subclustering. The two giant galaxies have entourages of small systems. The Large and Small Magellanic clouds are so close that they may be in the process of merging with the Milky Way Galaxy. The Milky Way also has the nine dwarf satellites: Fornax, Sculptor, Leo I, Leo II, Sextans, Draco, Carina, Ursa Minor, and Sagittarius. These are gas-depleted spheroidal systems with essentially no ongoing star formation. The Sagittarius dwarf is in the process of being absorbed by the Milky Way Galaxy.

In any photograph of M31 the high surface-brightness companions M32 and NGC 205 are clearly visible. In addition, M33 is within 6.5×10^5 light-years of M31, and in proximity there are NGC 147, NGC 185, IC 10, and the very small systems LGS 3 and Andromeda I, II, III, V, and VI.

In another part of the sky, NGC 6822, David Dunlap Observatory (DDO) 210, SagDIG (Sagittarius Dwarf Irregular Galaxy), and Phoenix are close together. The nearest neighbors to the Local Group are four other dwarf irregular systems close enough to each other that they could be thought of as a distinct entity: NGC 3109, Sextans A, Sextans B, and Antlia. A rigorous criterion for group membership might require total or continuing collapse toward the common potential of the group. However, the distances and three-dimensional motions of each of the systems or of the underlying mass distribution are not well enough known to identify with certainty which systems are bound and which are separating.

Types of galaxies. Galaxies within the Local Group, as with most others, can be assigned to one of two families. Some have substantial reservoirs of hydrogen and are actively forming new stars; others are

gas-depleted and contain only old stars. The three spirals and the Magellanic or dwarf irregulars have young stars and gas as well as old constituents. The elliptical, lenticular, and dwarf spheroidal systems are all deficient in interstellar gas and have the smooth textures of old, dynamically relaxed galaxies.

All the known gas-poor systems in the Local Group lie close to one of the giant galaxies and to each other. In contrast, the gas-rich systems tend to be farther apart, even though still identified with the group. Tidal or direct collisional stripping has probably affected the systems in crowded regions and deprived them of the gas reserves necessary for continued star formation.

Distances. The galaxies in the Local Group are so close that individual stars can be resolved with modern detectors on large telescopes. More massive young stars are easily observed in the spiral and irregular galaxies. Among the populations of old stars, only the brightest of those in a certain phase of their evolution called red giants are visible in the more distant parts of the Local Group. In each case, all well-studied systems contain stars with familiar properties that give an indication of their distances. The most useful are the Cepheid and RR Lyrae stars because they pulsate with frequencies that are related to their intrinsic luminosities. With the Hubble Space Telescope, red giant stars are easily observed, and accurate distances can be determined from the properties of the brightest of these. It has been established that the distances measured in the first half of the twentieth century were badly underestimated. The current estimates (see table) are fundamental for the calibration of the extragalactic distance scale. *See* CEPHEIDS; GIANT STAR; VARIABLE STAR.

Mass. Assuming that the galaxies within 3×10^6 light-years of either the Milky Way or M31 are part of a dynamically collapsed system, the mass of the unit can be estimated with the virial theorem. This theorem specifies the relationship between the mass, dimensions, and motions in an isolated, relaxed system of gravitating particles. In practice, there are problems because not all the velocity and position information is available, and the assumption that the system has collapsed and relaxed is probably poorly satisfied. Nevertheless, the results of such an analysis are consistent with similar analyses of other groups and can be taken to be valid in a statistical sense. It is estimated that the mass of the Local Group is roughly 10^{12} times the mass of the Sun (2×10^{42} kg or 4×10^{42} lb). Hence, the ratio of the mass to light of the Local Group is 20 in solar units. This value is lower than a more typical value of 100 for galaxy groups, but much larger than the value of about 5 for the luminous parts of individual galaxies. It is supposed that in the Local Group, as in many other environments, the potential well is dominated by a component of invisible matter, so-called dark matter. The mass contributed by the stars that produce the light and by the interstellar hydrogen are relatively insignificant. *See* COSMOLOGY; GALAXY, EXTERNAL. R. Brent Tully

Bibliography. P. Hodge, Populations in Local Group galaxies, *Annu. Rev. Astron. Astrophys.*, 27:139–159, 1989; A. Layden, R. C. Smith, and J. Storm (eds.), *The Local Group: Comparative and Global Properties*, 3d CTIO/ESO Workshop, La Serena, Chile, Jan. 25–28, 1994; M. L. Mateo, Dwarf galaxies of the Local Group, *Annu. Rev. Astron. Astrophys.*, 36:435–506, 1998; A. R. Sandage and J. Bedke, *The Carnegie Atlas of Galaxies*, 2 vols., 1994; R. B. Tully and J. R. Fisher, *Nearby Galaxies Atlas*, 1987; S. van den Bergh, *The Galaxies of the Local Group*, Cambridge Astrophysics Series 35, 2000.

Location fit

Mechanical sizes of mating parts such that, when assembled, they are accurately positioned in relation to each other. Locational fits are intended only to determine the orientation of the parts. For normally stationary parts that require ease of assembly or disassembly, for parts that fit snugly, and for parts that move yet fit closely as in spigots, slight clearance is provided between parts. Where accuracy of location is important, transition fits are used. In these fits the holes and shafts are normally nearly the same diameter. For greater accuracy of location, the shafts are made slightly larger than the holes; such fits are termed location interference fits. *See* ALLOWANCE.

Paul H. Black

Locomotive

A machine mounted on flanged wheels which converts some form of potential energy into the mechanical work of propelling itself and other, nonpowered vehicles over a railroad track. The dimensions and weight of a locomotive are restricted by the presence of stationary structures adjacent to and above the track, such as overhead bridges and tunnels, and by the strength of the track and bridges that support it.

Economics demands that locomotives use an available and plentiful form of potential energy; that they convert it to useful work efficiently; that they be capable of applying sufficient force to overcome the resistance of the loaded vehicles they are to move, and they do so in the time allotted; and that they be durable, reliable, and relatively easy to maintain.

Characteristics. Locomotives are described in terms of the type of energy they consume; their size and weight; the tractive effort, that is, the force they are capable of exerting; and their power, that is, the rate at which they convert energy to work. Locomotive tractive effort varies with speed, so it is often described by a curve that shows the relationship between speed and the force applied to the leading coupler of a train. The force at zero speed is known as the starting tractive effort and is related to the locomotive's weight on driving wheels by the factor of adhesion. The factor of adhesion can exceed the nominal coefficient of friction by employing sophisticated antislip technologies.

The coefficient of friction between a steel wheel

and a steel rail depends upon the surface conditions of both and varies considerably even over short distances and at different times. The nominal coefficient of friction is therefore only an average value. Railroads have used a number of techniques to enhance friction levels. Spreading sand ahead of driving wheels has been used for years. Antislip controls that detect slipping wheels and momentarily reduce power to those wheels have been developed. The use of alternating-current (ac) synchronous traction motors where speed is strictly controlled by current frequency has further raised the effective coefficient of friction. In such cases the locomotive wheels improve the surface conditions of wheels and rails by instantaneous slippage or stick-slip phenomena, which burn off contaminants that would otherwise lower the friction coefficient by lubricating the contacting surfaces. Thus, nominal coefficients of friction of the order of 20–30% can be raised to factors of adhesion well above 40%. *See* FRICTION.

Steam models. The first locomotives (around 1830) were steam propelled. The first successful steam locomotives employed a multiple fire-tube boiler in which the hot gases from the combustion of coal or wood were passed through submerged pipes (flues) to heat water to or above its boiling point. The resulting steam was expanded against pistons connected to driving wheels, the linear reciprocating motion of the pistons being converted to rotation by a system of cranks. Over the first 100 years or so of railroad history, steam locomotives were the primary source of power for railroads throughout the world. The basic concept of firebox, boiler, and smoke box arranged longitudinally from rear to front and cylinders under the smoke box lasted throughout the history of steam with some minor exceptions. Within this concept, however, many changes and improvements were made. In particular, as rail and track strength were increased, locomotive size and weight did as well. The first steam locomotives were restricted to weights of about 10 tons (9 metric tons), but by the 1940s huge machines weighing over 300 tons (270 metric tons) were being produced. These locomotives were equipped with superheaters and feedwater heaters to improve their efficiency, sanders to reduce wheel slippage in situations requiring high tractive effort, mechanical stokers and lubricators, and electric generators for lighting; they used specially treated water; at speed, they could produce upward of 6000 horsepower. Because of the many advantages of other forms of motive power, steam locomotives began to be replaced shortly after World War II and had virtually disappeared in the United States by the late 1950s. However, steam power is still used in China on some railways preserved as attractions for tourists and railroad enthusiasts. *See* STEAM ENGINE.

Electric models. Electric locomotives draw electricity from a third rail set alongside the running rails of track through a sliding shoe that transmits electric power to the locomotive, or from a configuration of overhead wires known as a catenary by means of a pantograph that raises or lowers to compensate for wire height variations. The cost of these electric distribution systems has restricted the use of electric motive power to only very high traffic routes such as the Northeast Corridor of Amtrak in the United States, commuter lines, and some routes in Japan, Russia, England, and continental Europe.

Electric locomotives can be designed and constructed to use a wide variety or even combinations of voltages and either direct or alternating currents. Low-voltage direct current (dc) is often distributed over relatively short distances by a third rail. High-voltage direct or alternating current (1500 dc, 3000 dc, 11,000 ac, 15,000 ac, 25,000 ac) is distributed by overhead catenary.

Electric locomotives can be overloaded for short periods of time so that a nominal 4000–5000–hp locomotive can produce 7000–8000 hp during acceleration or when climbing grades.

Therefore, since locomotives do not require their maximum power at all times, electrification permits a railroad to purchase a smaller quantity of total locomotive horsepower than would be required for other types of power. Besides the high cost of electric distribution systems, the disadvantages of electrification include reduced overhead clearance where a catenary is used, danger to employees working on or around the third rail or catenary, and generally lower utilization factors for locomotives. Utilization factors are the measures used by railroads to describe the percentage of time that a locomotive (or other piece of equipment) is actually in service. The utilization of electric locomotives is restricted by the limits of the power distribution system. Electric locomotives can go only where the electrification system goes. This differs from availability factors (which are high for electric locomotives) that measure the percentage of time that a locomotive is ready for service. Thus an electric locomotive may be available (that is, may be functional and serviced) but cannot be assigned since there are no trains destined to locations on the electrified lines. *See* ALTERNATING-CURRENT MOTOR.

Diesel models. Individual diesel-electric locomotive units are produced in a number of sizes, with corresponding power (**Fig. 1**). Switching or shunting locomotives are of approximately 1000 hp and

Fig. 1. Modern 4000-hp diesel-electric ac locomotive. (*Electro-Motive Division of General Motors Corp.*)

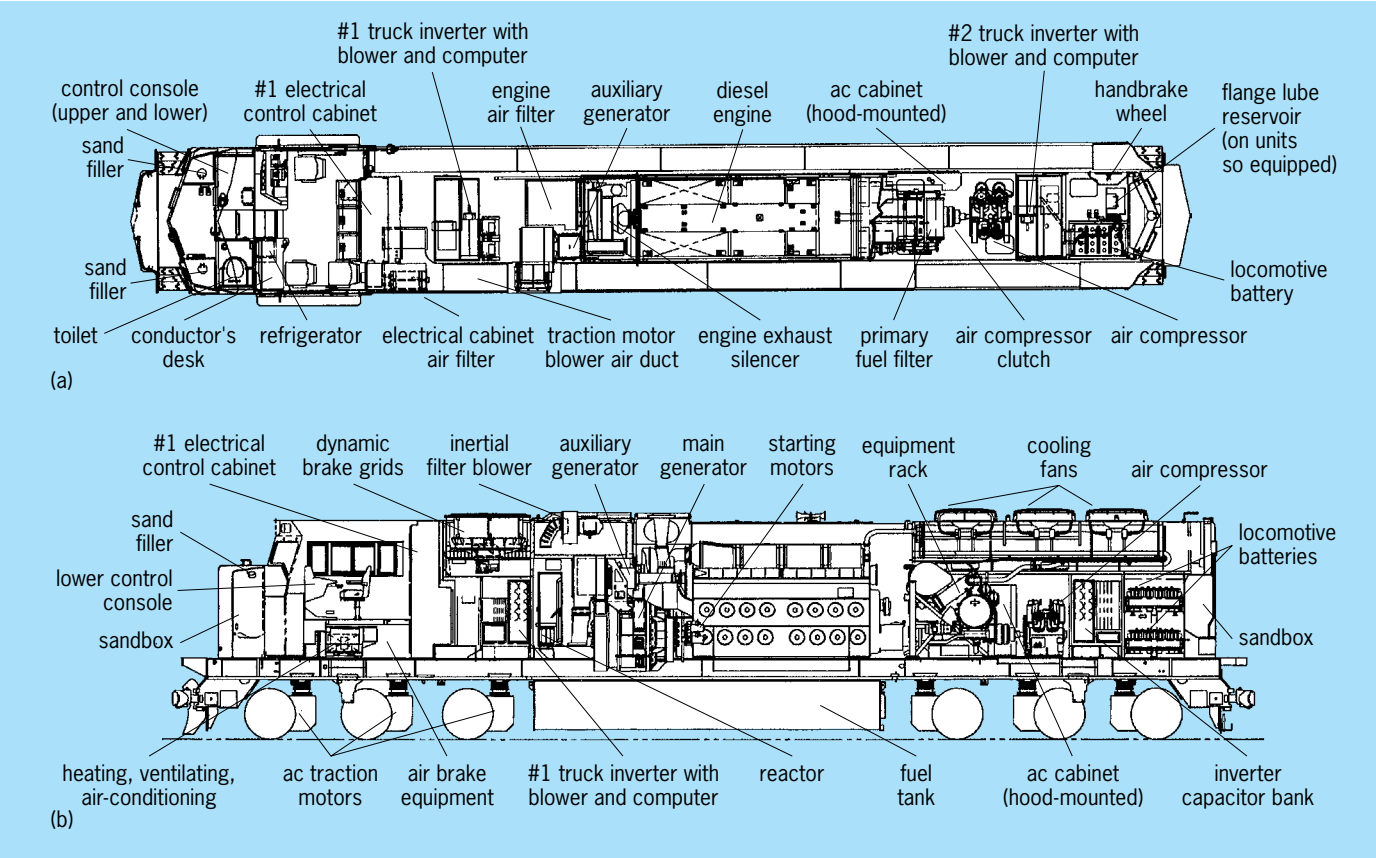

Fig. 2. Major components of a modern diesel-electric locomotive. (*a*) Plan view. (*b*) Cross section. (*Electro-Motive Division of General Motors Corp.*)

low maximum speed, while mainline passenger and freight units are being designed to produce 5000–6000 hp. Several diesel-electric units are most often coupled together to produce a single locomotive of as much as 15,000 hp or more. In such cases the entire locomotive assembly is controlled from the leading unit.

The diesel engines are designed either as two- or four-cycle (**Fig. 2**). They are most often of the vee configuration of 12–20 cylinders 9–11 in. (23–28 cm) in diameter. Most are turbo-supercharged and operate at approximately 1000 rpm. The lubricating oil system on a diesel unit ensures a supply of clean, cool oil to the engine, and it provides for extracting carbon deposits left by combustion along with some heat from the engine. The oil is circulated by engine-driven pressure pumps through removable filter elements, strainers, and a water-cooled heat exchanger. The engine cooling system also uses engine-driven centrifugal pumps to circulate treated water through cylinder heads, liners, and the oil cooler to banks of radiator units. The large fans that force air through the radiators are driven mechanically from the engine or by electric motors. *See* TURBOCHARGER.

The fuel system includes a motor-driven pump to deliver fuel oil from a storage tank, usually mounted under the carbody between the trucks, with a capacity of about 4000 gal (15,000 liters). The oil is forced through a strainer and filter to the injectors, where it is sprayed into the cylinders by high-pressure pumps and nozzles. The quantity of fuel delivered for each stroke of the engine piston is regulated by a mechanical linkage connected to the engine governor. This device keeps engine speed constant for each throttle setting, regardless of load. *See* GOVERNOR; LINKAGE (MECHANISM).

These engines turn an alternator, which produces the electricity to power the traction motors. The traction motors are mounted in the truck assemblies with their armature shafts parallel to the locomotive's driving axles. A pinion gear on the armature shaft turns a bull gear on the axle to propel the locomotive. Traction motors are cooled by fan-driven air through ducts. There may be four or six such motor-driven axles per locomotive unit, and the gear ratio between pinion and bull gear teeth is varied according to the locomotive's intended use.

At one time, all traction motors were dc, series-wound machines, and motor speed had to be limited to prevent damage to the commutator and armature from centrifugal force. However, low motor speed for extended periods is associated with high current flow and damage from excessive heat. Thus, gear ratio has been critical, and it determined the speed range of the locomotive. With the advent of ac squirrel-cage motors, gear ratio is less critical, and locomotive assignments can be more flexible as these ac machines are equally suited to both high-speed

passenger or merchandise service and lower-speed heavy-tonnage assignments. *See* DIESEL ENGINE.

Braking. Both diesel-electric and straight electric locomotives can be equipped with dynamic brakes where traction motors act as generators resisting rotation, thus slowing the train. In the straight electric locomotive the current produced is fed back into the catenary, while in the diesel-electric the generated current is fed to high-resistance grids, which must be cooled by fan-driven air. Dynamic braking can be advantageous in reducing wear and tear on train brakes and wheels, but it must be used judiciously to avoid excessive longitudinal train forces. *See* BRAKE; DYNAMIC BRAKING; GAS TURBINE; RAILROAD ENGINEERING; TURBINE PROPULSION. George H. Way

Bibliography. H. I. Andrews, *Railway Traction: The Principles of Mechanical and Electrical Railway Traction*, 1986; J. H. Armstrong, *The Railroad, What It Is, What It Does*, 4th ed., 1990; W. W. Hay, *Railroad Engineering*, 2d ed., 1982; C. J. Riley, *The Encyclopedia of Trains and Locomotives*, 1994.

Locust (forestry)

A name commonly applied to two trees, the black locust (*Robinia pseudoacacia*) and the honey locust (*Gleditsia triacanthos*). Both of these commercially important trees have podlike fruits similar to those of the pea or bean. *See* LEGUME.

The black locust, which may attain a height of 75 ft (22 m), is native in the Appalachian and the Ozark regions and is now widely naturalized in the eastern United States, southern Canada, and Europe. It is characterized by odd-pinnate leaves, fragrant white blossoms like those of pea plants in long clusters, flat brown pods 2–4 in. (5–10 cm) long, and a pair of stipular thorns at the base of each leaf. In winter these paired thorns and the superposed buds hidden under the leaf scar are important distinguishing features (see **illus.**).

The wood is hard and durable in the soil and is used for fence posts, mine timbers, poles, and railroad ties. It is the principal wood for insulator pins on utility poles. Because of its wide-spreading, fibrous root system which supports nitrogen-fixing nodules, the black locust is well adapted to erosion control and soil reclamation. Many varieties are planted as shade or ornamental trees. *See* SOIL CONSERVATION.

The honey locust, which reaches a height of 135 ft (41 m), is native in the Appalachian and the Mississippi Valley regions, but is also widely naturalized in the eastern United States and southern Canada. The flowers are inconspicuous, but the large brown fruit pods (12–18 in. or 30–45 cm long) are striking. The tree may be easily recognized at any season of the year by the branched thorns growing on the trunk or branches. The leaves are either pinnate or bipinnate. The reddish wood is hard, strong, and coarse-grained and takes a high polish. Because it is durable in contact with soil, it is used for fence posts and railroad ties; it is also used for construction, furniture, and interior finish. Although it belongs to the legume family, the roots do not possess nitrogen-fixing bacteria. The tree is often used for hedges, as an ornamental, and for shade. *See* FOREST AND FORESTRY; TREE. Arthur H. Graves; Kenneth P. Davis

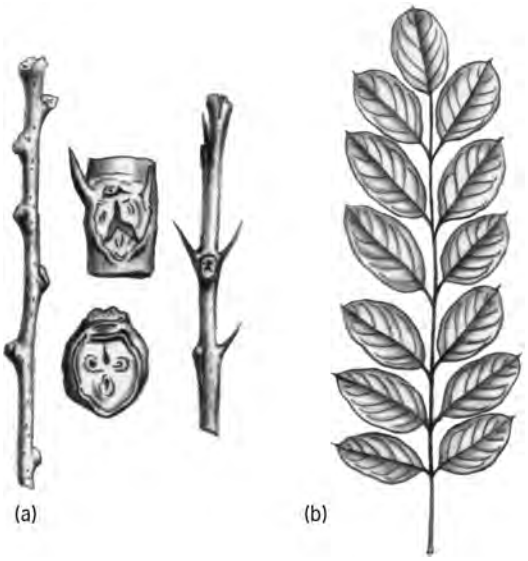

Black locust (*Robinia pseudoacacia*). (*a*) Branches, with and without thorns, and enlarged views of leaf scars. (*b*) Odd-pinnate leaves.

Loess

Deposits of windblown fine-grained angular particles that are produced by the breakdown of larger particles at desert or glacial margins. Suspension and transport by the wind produces particularly homogenous and well-sorted loess upon deposition. Loess is dominated by silt-sized (2–63 micrometers in diameter) particles of quartz, with smaller amounts of feldspars, calcium carbonate, and clay minerals, and can form successive, blanketing layers of thickness from centimeters to hundreds of meters, as in the Loess Plateau region of north-central China (**Fig. 1**). Loess occurs globally, encompassing the pampas of Argentina, the central and east European and Russian steppe, and the midcontinent United States, to name some locations. It is porous and unconsolidated, making it a valuable agricultural resource. Loess is highly erodible in wet conditions, although often mechanically strong when dry because calcium carbonate and clay aggregates bind the silt particles together. *See* CALCITE; CLAY MINERALS; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; EOLIAN LANDFORMS; FELDSPAR; QUARTZ.

Paleoclimate and the Chinese Loess Plateau. In addition to its agricultural importance, loess is of key geological significance. Thick sequences of loess and their interbedded buried soils (Fig. 1) are the subject of intense scientific attention because they contain the longest, most detailed record of Quaternary climatic changes yet found on land. In China, these sediments span at least the last 2.5 million years of the Earth's history, and cover some 500,000 km^2 (193,000 mi^2) of land (**Fig. 2**). The vertical loess

Fig. 1. Loess/paleosol sequence in the central Loess Plateau of China. The paler layers of loess are interbedded with the darker paleosol units, two of which are indicated by arrows.

thickness is ~350 m (1150 ft) near the nothern and western Plateau edges, bordering Mongolia and Tibet, ~150 m (490 ft) in the central area, and thins out towards the southern and eastern cities of Xi'an and Beijing.

Sediment source. Because the Chinese loess is thickest and its particle size coarsest toward the northwest (**Fig. 3**), it was thought that the northwestern deserts were the source of the loess, blown by winter monsoon winds. There is a good match of the geochemical "fingerprints," including strontium and neodymium isotopic compositions and rare-earth and trace-element concentrations, between the loess and desert sands. However, the loess is often much

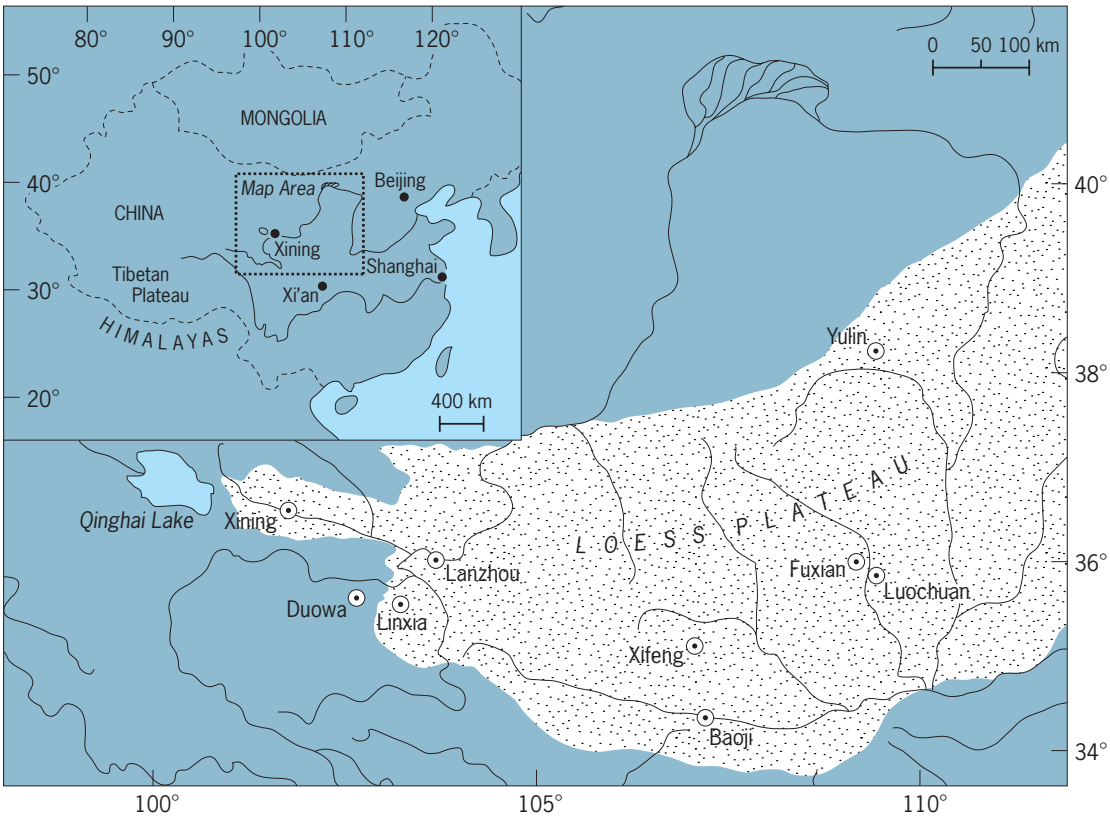

Fig. 2. Loess Plateau, north-central China.

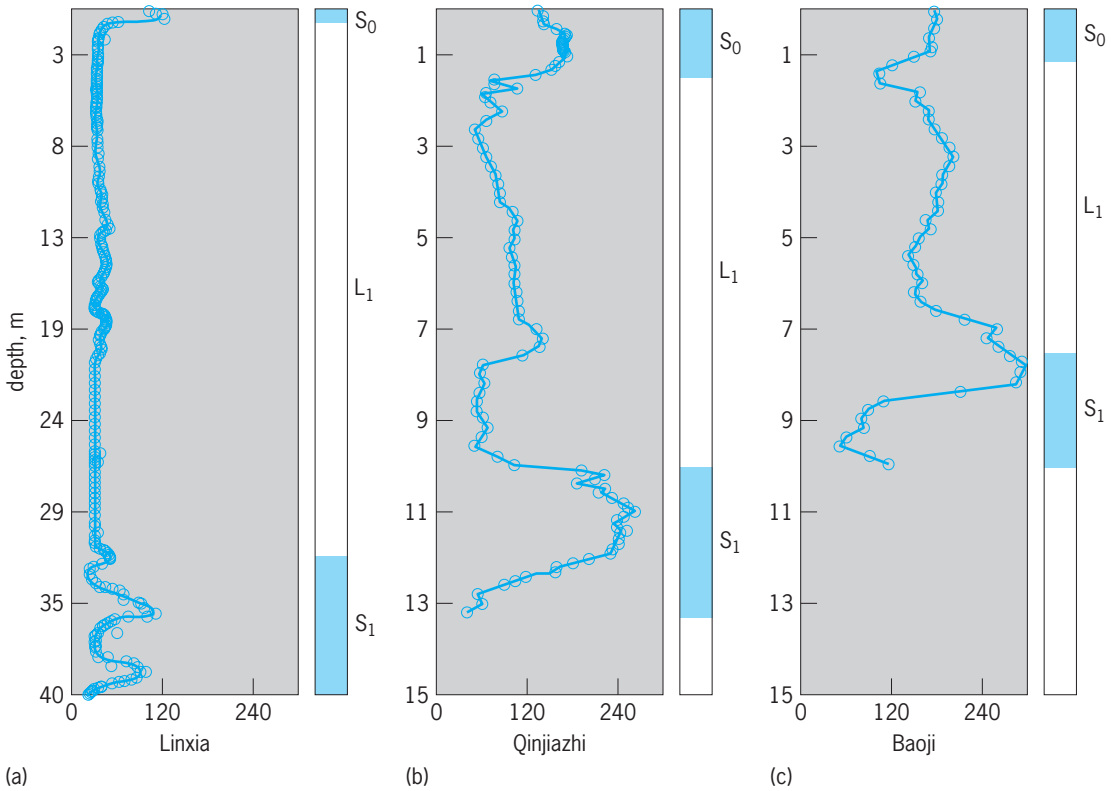

Fig. 3. Loess/paleosol stratigraphy (modern soil—S_0; last glacial loess unit—L_1; and last interglacial paleosol—S_1) and magnetic susceptibility (10^{-8} m^3 kg^{-1}). (*a*) A western Loess Plateau site (note 40 m or 131 ft of sediment spanning the soil/loess/paleosol units, and a maximum paleosol susceptibility value of 120 × 10^{-8} m^3 kg^{-1}). (*b*) A central site (note 13 m or 43 ft of sediment for the same time span as *a*, and maximum paleosol susceptibility value of 245 × 10^{-8} m^3 kg^{-1}). (*c*) A southern site (10 m or 33 ft of sediment for the same time span as *a* and *b*, and maximum paleosol susceptibility value of 300 10^{-8} m^3 kg^{-1}). The higher the susceptibility, the higher the paleo-rainfall.

more magnetic than modern desert sand. It also has variable but often very high values of beryllium-10, which forms in the atmosphere through nuclear reactions caused by cosmic rays and is brought down to the Earth's service by precipitation, indicating a source area with moderate to high rainfall such as the Himalayas or Tibetan Plateau. The loess transport path might originate with winter surface winds, blowing from the western Himalayas to the northeast, and supplying Himalayan dust to the deserts to the north and west of the Loess Plateau. During the spring months of the winter monsoon, the dust is blown from these deserts to the Loess Plateau to the southeast. *See* ASIA; BERYLLIUM; COSMIC RAYS; COSMOGENIC NUCLIDE; MAGNETIC SUSCEPTIBILITY; MONSOON METEOROLOGY; QUATERNARY; RADIOISOTOPE (GEOCHEMISTRY); SEDIMENTOLOGY.

Paleosols and climate. Interlayered within the "buff"-colored loess are buried soils (paleosols), traceable over hundreds of kilometers as near-horizontal, reddened layers (Fig. 1). These ancient soils formed when the climate became warmer and wetter so that at successive intervals the loess surface became vegetated and weathered. More than 30 alternations between cooler, drier periods (during which loess was transported and deposited) and warmer, wetter, soil-forming periods have occurred. These climatic variations have occurred over the past 2.5 million years. The basal age of the Chinese sequences has been defined paleomagnetically, and the central area of the Chinese Loess Plateau has recorded the last six major reversals of the Earth's magnetic field. *See* CLIMATE HISTORY; MAGNETIC FIELD; PALEOMAGNETISM; PALEOSOL; ROCK MAGNETISM; SOIL.

Magnetic variations. It may be possible to quantify these monsoon rainfall variations from the mineral magnetic properties of the loess and paleosol units. The paleosols are much more magnetic than the loess layers. In terms of their response to an applied magnetic field (that is, their magnetic susceptibility), paleosol layers have values 2–5 times higher than the loess layers (Fig. 3). In modern soils, strongly magnetic minerals (magnetite and maghemite) are formed in situ in well-drained, moderately acidic soils (pH ~5.5–7) on iron-bearing, weatherable parent materials. Magnetic measurements of modern soils across the Chinese Loess Plateau and Russian steppe show that the amount of magnetic material formed is strongly correlated with annual rainfall. The numerical relationship between magnetic susceptibility and rainfall defined from modern soils can be used to calculate paleo-rainfall from the magnetic susceptibility values of the buried soils in the Chinese loess/soil sequences. This approach identifies high rainfall (more intense summer monsoons) in the early Holocene (8000 to 6000 years ago) and late Holocene (5000 to 3000 years ago) and the last interglacial, especially in the presently semiarid western Plateau. During the last glacial stage, rainfall was

reduced across the whole region. Controversially, G. Kukla and colleagues believe that the susceptibility variations are mostly controlled not by rainfall but by the rate of deposition of the magnetically poor dust. *See* GLACIOLOGY; MAGNETITE.

Multiproxy climate indicators. In addition to geochemical and paleomagnetic indicators, other proxy climate indicators retrieved from global loess/soil sequences include fossil snails, plant phytoliths (silicate exudates from plant roots), loess particle size, and carbonate content. It is possible to extract preserved fossil snails from loess/paleosol sequences, with different ecological groups reflecting different temperature and moisture requirements of each taxon. Phytolith analysis provides information on vegetation changes, with trees and grasses forming differently shaped phytoliths, but phytoliths can also record isotopic information relating to changes in temperature and rainfall. As with the fossil snails, the phytoliths have to be extracted from the loess and both are subject to dissolution in some sequences. Particle-size variations in the loess can record the strength of the winter monsoon winds. Larger detrital grain sizes may indicate periods of increased wind strength (or increased proximity to dust sources), and smaller sizes may indicate periods of decreased winter monsoon intensity (or increased distance from dust sources). Carbonate content may reflect variations in summer monsoon moisture, causing dissolution of detrital carbonate and its reprecipitation in lower soil horizons. High-resolution studies of the loess/soil sections in the western Chinese Loess Plateau, where the sequences are at their thickest, show that the intensities of the summer and winter monsoons are not directly coupled. From about 12,000 to 8000 years ago the summer monsoon was generally strong and the winter monsoon was weak, while about 8000 years ago both the winter and summer monsoons were strong.

Magnetic susceptibility/climate relationships. Thinner, less complete loess/soil sequences occur around the globe. Those in Alaska, Siberia, and Poland show the opposite pattern of magnetism, with low susceptibility values associated with the paleosols and higher values for the loess layers. One reason for this is that the loess in these regions is very strongly magnetic (much more magnetic than the Chinese loess). Also, modern soils show that arid conditions prevent formation of new magnetic material, while waterlogging and acidification of soils not only inhibit magnetic mineral formation but also destroy detrital magnetic minerals. Thus, we can infer that the climate at soil-forming intervals in these locations was not conducive to in situ magnetic mineral production.

Barbara A. Maher

Bibliography. F. Heller and M. E. Evans, Loess magnetism, *Rev. Geophys.*, 33:211–240, 1995; H. Y. Lu et al., Seasonal climatic variation recorded by phytolith assemblages from the Baoji loess sequence in central China over the last 150 ka, *Past Global Changes Newslett*, 6:4–5, 1998; B. A. Maher, Magnetic properties of modern soils and loessic paleosols: Implications for paleoclimate, *Palaeogeog. Palaeoclimatol. Palaeoecol.*, 137:25–54, 1998; B. A. Maher and R. Thompson, Palaeomonsoons I: The magnetic record of palaeoclimate in the terrestrial loess and palaeosol sequences, in B. A. Maher and R. Thompson, *Quaternary Climates, Environments and Magnetism*, pp. 81–125, Cambridge University Press, 1999; B. A. Maher, A. Alekseev, and T. Alekseeva, Variation of soil magnetism across the Russian steppe: Its significance for use of soil magnetism as a palaeorainfall proxy, *Quat. Sci. Rev.*, 21:1571–1576, 2002.

Logarithm

An exponent of a suitably chosen positive number (base) larger than unity. Logarithms are of value in mathematical computation and in the equations and formulas used in expressing natural phenomena.

The invention of natural logarithms (base $e = 2.718\ldots$) is usually attributed to John Napier, who published a table of logarithms in Edinburgh in 1614 under the title *Mirifici Logarithmorum Canonis Descriptio*, even though the functions tabulated by Napier are merely related to, and not identical with, natural logarithms.

The invention of common logarithms (base 10) is generally attributed to Henri Briggs; Napier's share is acknowledged, however, by Briggs himself in his *Arithmetica Logarithmica* in 1624.

Theory. If $b_l = n$, where b is a positive number larger than unity, then l is called the logarithm of n to the base b and is written $l = \log_b n$. From this definition it follows at once that all positive numbers larger than unity have positive logarithms, all positive numbers smaller than unity have negative logarithms, and the logarithm of unity is equal to 0. Since $b^0 = 1$ irrespective of the value of b, it follows that the logarithm of unity is equal to 0. From the known properties of exponentials expressed by relations (1), it follows immediately that (a) the log-

$$\begin{aligned} b^{l_1} \cdot b^{l_2} &= b^{l_1+l_2} \qquad & b^{l_1} \div b^{l_2} &= b^{l_1-l_2} \\ \left(b^l\right)^m &= b^{lm} \qquad & \sqrt[m]{b^l} &= b^{l/m} \end{aligned} \tag{1}$$

arithm of a product of two or more factors is equal to the sum of the logarithms of the factors; (b) the logarithm of the ratio of two numbers is equal to the difference between the logarithm of the numerator and the logarithm of the denominator; (c) the logarithm of the mth power of a number is the product of m and the logarithm of the number; and (d) the logarithm of the mth root of a number is equal to the logarithm of the number divided by m. These properties are responsible for the great simplification in the task of carrying out numerical computations involving multiplication, division, raising numbers to certain powers, or extracting the roots of a certain order of given numbers. It also follows from these properties that in constructing a table of logarithms to a certain base from the beginning, the computer need only compute the logarithms of prime numbers, because the logarithm of any number which is

not prime can be obtained from the logarithms of its prime factors by simple operations of addition and multiplication. *See* EXPONENT.

Although any positive number b larger than unity might have been chosen as the base of a system of logarithms, actually two numbers have been chosen in the construction of tables of logarithms, namely, the number $b = 10$ and the number e defined as the sum of infinite series (2). The system of logarithms

$$1 + \frac{1}{1} + \frac{1}{1 \cdot 2} + \frac{1}{1 \cdot 2 \cdot 3} + \cdots + \frac{1}{1 \cdot 2 \cdot 3 \cdots n} + \cdots \quad (2)$$

to the base 10 is usually referred to as common logarithms; the system of logarithms to the base e is called natural logarithms.

Common logarithms have certain obvious advantages not shared by natural logarithms. All numbers between 1 and 10 have logarithms between 0 and 1, all numbers between 10 and 100 have logarithms between 1 and 2, and so on. Thus the integral part of the common logarithm of a number larger than unity is one unit less than the number of digits before the decimal point. This integral part is called the characteristic; the decimal part is called the mantissa. Similarly, because the common logarithms of 0.1, 0.01, and 0.001 are -1, -2, and -3, it follows that the logarithm of a number smaller than unity having p zeros after the decimal point may be expressed as the sum of a negative characteristic equal to $-(p + 1)$ and a positive mantissa.

Natural logarithms. Because the number $e = 2.718 \ldots$, the base of the system of natural logarithms is irrational; that is, it cannot be expressed as the ratio of two integers (and therefore when expressed as a decimal number, it involves an infinite number of decimals with no repeating groups of decimals). It might seem odd, therefore, that it has been chosen as the base of a system of logarithms. The primary motivation for this choice lies in the fact that the solutions of numerous problems in applied mathematics are most naturally expressed in terms of powers of e. Thus, for instance, the solutions of the problems of equilibrium of a flexible cable, the transient flow of electric current in a circuit, and the disintegration of radioactive elements are expressed in terms of e^x, where x is either positive or negative and depends on the physical parameters of the problem in question. Thus, the tabulation of the function e^x for both positive and negative values of x was an indispensable aid in obtaining the solutions of many physical problems. Because the logarithmic function is the inverse of the exponential function, a table of natural logarithms can be constructed with relative ease from a table of e^x by the process of inverse interpolation. Another motivation for constructing tables of natural logarithms is the fact that natural logarithms arise in the integration of quotients of polynomials and of functions which may be approximated by quotients of polynomials. Indeed, it is a well-known fact that if the n roots of a polynomial $Pn(x)$ are known so that

Table of logarithms

Number	Natural logarithm	Common logarithm
0.01	5.586–10	−2 (8 − 10)
1.0	0	0
$e = 2.718 \ldots$	1	0.4343
10	2.303	1
1000	6.909	3

$Pn(x) = a_0(x - x_1)(x - x_3) \cdots (x - x_n)$ and if $Q(x)$ is a polynomial of degree smaller than n, then the quotient $Q(x)/Pn(x)$ may be expressed in the form of expression (3) when the A_ks may be easily determined. Accordingly Eq. (4) may be written.

$$\sum_{k=1}^{n} \frac{A_k}{x - x_k} \quad (3)$$

$$\int \frac{Q(x)}{Pn(x)}\, dx = \sum_{k=1}^{n} A_k \int \frac{dx}{x - x_k} = \sum_{k=1}^{n} A_k \log_e\, (x - x_k) \quad (4)$$

Relation between logarithm bases. Let l_1 be the known natural logarithm of a number N. What is the common logarithm of N? Clearly $N = e_1^l$, and therefore $\log_{10} N = l_l \log_{10} e$. Since $\log_{10} e = 0.43429$ (to five decimals) it can be concluded that natural logarithms are converted into common logarithms by multiplying the natural logarithm by the factor 0.43429, which is called the modulus. Similarly the common logarithm is converted into a natural logarithm by division by the modulus 0.43429 or multiplication by 2.303. The **table** shows natural and common logarithms for some numbers from 0.01 to 1000.

Logarithms of complex numbers. Since b^l is positive when b is positive, negative numbers have no real logarithms. Nevertheless, Euler's famous formula $e^{i\theta} = \cos\theta + i \sin\theta$, where $i = \sqrt{-1}$, makes it possible to define not only the logarithm of a negative number but also the logarithm of a complex number. Indeed, any complex number $a + ib$ may be written in the form $\rho e^{i\theta}$, where $\rho = \sqrt{a^2 + b^2}$ and $\theta = \arctan(b/a)$. It follows that the logarithm of a complex number $a + ib$ is a complex number whose real part is the logarithm of its modulus ρ and whose imaginary part is its phase (or argument) ρ. Because $e^{i\theta} = e^{i(\theta + 2n\pi)}$, it follows that a complex number has an infinite number of logarithms given by Eq. (5), where

$$\log(a + ib) = \log\sqrt{a^2 + b^2} + i(\theta + 2n\pi) \quad (5)$$

n is a positive or negative integer.

As previously mentioned, the evaluation of natural logarithms may be based on inverse interpolation in a table of e^x. For small values of x, positive or negative, it is possible to use expansion (6). This expan-

$$\log_e(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n+1}\frac{x^n}{n} + \cdots \quad (6)$$

sion may be derived from Taylor's series, notation (7),

$$f(x+a) = f(a) + \frac{x}{1}f'(a) + \frac{x^2}{1\cdot 2}f''(a) + \cdots + \frac{x^n}{1\cdot 2\cdot 3\cdot m}f^{(n)}(a) + \cdots \quad (7)$$

by writing $f(x) = \log_e x$. After a sufficiently accurate table of $\log(1 \pm x)$, for $x = k10^{-n}$, $k = 1, 2, 3, \ldots, 9$, and $n = 1, 2, 3, \ldots, N$, where N is sufficiently large (perhaps 15) has been computed, this table may be used to compute the logarithm of any number. From the expansion (6) for $\log_e(1 + x)$ it is easy to derive expansion (8), where Eq. (9) holds; this has the

$$\log_e p = {}^1/_2\{\log_e(p-1) + \log_e(p+1)\} + S(p) \quad (8)$$

$$S(p) = \frac{1}{2p^2-1} + \frac{1}{3(2p^2-1)^3} + \cdots + \frac{1}{(2n+1)(2p^2-1)^{2n+1}} + \cdots \quad (9)$$

virtue that $S(p)$ converges quite rapidly. Thus for $p >$ 10,000 the error resulting from neglecting all but the first term in $S(p)$ is considerably less than 10^{-20}. *See* E (MATHEMATICS); NUMBER THEORY; NUMERICAL ANALYSIS. Arnold N. Lowan; Solomon Bochner

Logging

Those processes required to bring all or a portion of a tree from the stump to the mill facilities. Logging (tree harvesting) processes are clustered into tree conversion, woods transport (off-road transportation), landing operations (wood transfer), transport from landing to mill facility (truck, water, or rail), and unloading at the mill facility (wood transfer).

Logging systems are networks that connect harvesting processes so that the whole operation is physically feasible, economically efficient, environmentally sound, and institutionally feasible. They are usually characterized by the degree of tree conversion before off-road transportation: cut-to-log-length, tree-length, or full-tree. Off-road transportation options are the decisive factors for the use of equipment, and restrict the possible configuration of harvesting systems. The major off-road transportation options include ground-based transportation (vehicles operating on the surface using unpaved "skid trails" through the forest), cable systems, and airships (such as helicopters).

Logging operations around the world vary with the level of industrial development of the local forest industry. Logging in some countries is done by hand tools and human or animal power (biomechanical). In other countries, machines with on-board computers and microprocessors carry out the harvesting activities without workers setting foot on the forest floor (mechanized logging).

Logging is both a transportation process and an allocation process. The allocation process relates how the various parts of the tree can meet a product or environmental need. Depending upon the forest management plan, some trees or parts of trees are removed to make some of the many wood products such as fuelwood, materials for housing, and furniture. Other trees are left in the forest for wildlife, forest renewal, or soil nutrient reasons. The transportation process moves the logs (woody stems), chips, chunks, flakes, or biomass (whole or parts of the tree destined to be used for energy) from the forest to mill facilities. *See* FOREST TIMBER RESOURCES; PAPER; PLYWOOD; WOOD PRODUCTS.

Felling (tree conversion). The start of harvesting is the cutting down of trees with hand tools, chain saws, or mechanized felling machines (using chain saws, circular saws, or hydraulic shears to cut the tree at the base). The tree may be further cut into suitable lengths (bucking), or it may be transported whole or in tree-lengths. Tree products may be allocated during bucking with the aid of a computer on the felling and bucking machine or by a faller using a log order list or a hand-held computer to help decide the log products to make.

Felling is the most dangerous part of the harvesting operation. Larger trees generally must be felled manually with a chainsaw. Manual fallers often work under difficult conditions and must possess stamina, dexterity, skill with hand tools, mechanical ability, tree identification skills, and a thorough understanding of the rules involved in bucking. Mechanized felling operators must be adequately literate; have good dexterity, mechanical ability, tree identification skills, and decision-making skills; and be able to follow complex plans while working independently.

The objective of the tree falling operation is to fell the tree with minimum damage, to avoid damaging surrounding trees, to minimize soil and water impacts, and to position the tree or logs for the next phase of harvesting (**Fig. 1**). The goal of bucking is to produce the most valuable assortment of logs from the tree while considering the physical

Fig. 1. Feller-buncher cutting tree. (*John Deere*)

capability of the skidding (log-dragging), yarding (moving of logs to a landing), or forwarding (log-carrying) equipment.

Woods transport (off-road transport). Logs in lengths from about 1 to 10 m (3 to 33 ft) or other products must be transported from the stump to a place where they are further processed (often called a landing). In some cases, entire trees are pulled to the landing. Humans, animals (for example, oxen, horses, and elephants), crawler tractors and wheeled skidders (machines that drag the logs), forwarders (machines that carry loads of logs), farm tractors with winches or trailers, cable logging systems, balloons, or even helicopters transport logs and tree products to landings.

Normally, logs are hooked to the skidding animal or vehicle using wire rope (choker) or chains, or by a grapple arm. Forwarders use a hydraulic grapple to load the forwarder in the woods and unload it at the landing (**Fig. 2**).

The choice of the transport system depends on topography, roads and other transport options, size of products, economics, environmental considerations, soil types, weather, and the development stage of transport systems for the region or country. It may also depend upon the subsequent silvicultural activities. In some situations, soil disturbance may be encouraged to provide beds for seedlings; while in other situations, soil disturbance may be minimized to prevent erosion.

Where terrain (soil and slopes) is unsuitable for machines to operate on the ground surface for environmental or economic reasons, cable systems may be used to lift and drag or fly logs or felled trees to a landing area. Skyline systems may use wire rope 10–40 mm (0.4–1.6 in.) in diameter suspended between a steel tower 9–35 m (30–115 ft) high and a stump or tree to carry a load of logs connected to a carriage (device that rides on the skyline on pulley blocks) from where the tree was cut to the landing. The distance may be as short as 200 m (660 ft) or as long as 1700 m (5600 ft) depending on roads and terrain. Intermediate supports may be needed to provide necessary skyline clearance over the terrain. Alternately, helicopters used for heavy lifting may carry logs or trees by flying them to landings up to 2 km (1.2 mi) away.

Fig. 2. Forwarder loading logs for transport to landing. *(John Deere)*

Yarding productivity depends upon the system, equipment, yarding distance, and other circumstances. Equipment factors include the pulling power, load capacity, and speed. Environmental factors affecting productivity include tree characteristics; stand characteristics of size, shape, and location; the volume of timber to be harvested; yarding distance and direction; terrain conditions; and weather. Other productivity factors include the harvesting method, type of cutting (manual or mechanized), and skill and organization of the crew.

Landing operations. At the landing, the logs or trees may be stored or directly processed for transport. They may be loaded onto trucks, trains, barges, or ships, or prepared for water transport. Whole trees brought to a landing may have limbs and bark removed, and then be chipped and loaded into chip vans for transport to a pulp mill. Tree-length segments may be delimbed and bucked into logs for different market destinations. Trees may be shredded, chunked, or processed through machines for use as fuel. The allocation process may include measurement by volume or weight of the products. In some cases, logs can be offloaded from a woods transport vehicle, such as a forwarder, directly to a trailer, minimizing rehandling.

Loading. Because logs are heavy, they are normally loaded mechanically, although some regions still use manual or animal methods involving ramps. There are two general types of mechanical loaders used at roadside: swingboom loaders with grapples, and front-end loaders fitted with a log fork or grapple. Both are mobile, mounted on tracked or rubber-tired carriers. Forwarders usually unload themselves either into log decks for storage or onto setout trailers. For log products, loads are usually secured by using wire rope wrappers and binders to hold logs on the trailer during transport. Loader operators must be highly skilled. They are responsible for sorting logs on the landing, loading trucks heading for different destinations, carefully building stable loads for transport, without overloading the truck or injuring landing personnel.

Transportation (landing to mill facility). Trucks are most commonly used to transport log products to mill facilities (**Fig. 3**). They vary from small vehicles hauling 5–8 tons on straight beds to large specialized off-highway vehicles hauling 50 tons or more. A variety of truck trailers are used depending upon the type of product. Trees or long logs may be loaded onto pole trailers. Short logs, 2.5 m (8 ft) and less, are often stacked sideways on flat-bed trailers with bunks. Chips or flakes are hauled in specially designed chip vans. Many pole-trailers permit the trailer to be loaded onto the truck tractor for the return trip to improve traction and reduce tire wear. Logging trucks need to be stronger than normal highway trucks because they travel on lower-quality roads. Truck components must be suitably matched, including engines, brakes, transmissions, differentials, and tires. Drivers must be especially skilled in

Fig. 3. Heavy-duty log truck-trailer being loaded with a grapple. Part of a steel yarding tower is visible at the left and is rigged with a skyline cable system. (*Kenworth Truck Co., Division of PACCAR*)

maneuvering under adverse forest conditions as well as on public highways.

Water transport in barges, as log rafts, and as free-floating logs is used in some areas. Where water transport is consistent with environmental concerns, it is often the most economical alternative. It also may be used in conjunction with truck transport. Logs can be banded on the truck trailer before they are lifted from the truck and placed in the water. Bundles of logs form rafts which are then towed to processing destinations or to sites for loading onto seagoing vessels.

At one time, railroads were one of the principal methods of moving timber from the forest to a river landing or to the sawmill. With the development of modern road construction, improved road maintenance, and more powerful and reliable trucks, rail transport has declined.

Unloading. Log products can be unloaded from truck trailers by lifting, rolling, or dumping over the side or end, depending upon the type of trailer. Trees and long logs are usually lifted from trailers by a grapple on mobile wheel loaders or overhead cranes (which can unload the entire truck in one pass). Shorter logs are often unloaded using slings. In some cases, short wood for pulp is swept directly off the trailer and fed into a debarking machine to eliminate rehandling. Chip trailers are often tilted and end-dumped on large hydraulic ramps.

Not all logs are processed in the countries where they were harvested. Sometimes logs are shipped from countries with greater forest resources to those needing logs for their processing facilities.

Loggers. The people who conduct tree harvesting operations are called loggers. In some countries, loggers are farmers for part of the year who then use their tractors for logging in the forests adjacent to their farms. In others, loggers are highly skilled technicians operating from computerized cabs. Loggers may be self-employed and own their own equipment, or they may be employed by logging contractors or companies. Logging work is characterized as difficult and dangerous. In most countries, logging has one of the highest accident rates of any industrial activity. Most of the accidents occur from the use of chain saws. When mechanized harvesting operations replace manual harvesting, accidents decrease dramatically. Loggers implement the silvicultural prescriptions of forestry professionals. Logging contractors provide logging services to governmental and large and small forest owners, often using capital equipment worth millions of dollars.

Logging is a means of resource development that can be abused if modern forestry practices are not regulated or enforced. Throughout the world, many communities and cultures are dependent on the forest resource and its utilization that starts with the process of logging. Developed countries have forest practices regulations that are strongly enforced to protect environmental and societal values.

Environmental effects. Potential risks from logging within the natural environment include loss, destruction, and compaction of soils; damage to residual trees; disturbance of flora and fauna; and pollution of water. Risks in the social setting are to human health and safety, disturbance of landscape esthetics, and disturbance of cultural heritage. At present, a major emphasis of the logging industry worldwide is on reducing environmental impacts. Environmental as well as safety and health standards are prescribed in codes of practices to help control and reduce risks.

When logging is done poorly and without regard for environmental concerns, resource damage to trees, soil, water, and wildlife can occur. When environmental concerns are accommodated, such as careful selection of roads and landings, control over the season of operation, minimizing the area where machines travel, minimizing damage to surrounding trees, careful activities around streams, and controlled waste disposal from operations, logging can be conducted as part of a sustainable basis of development.

Planning for logging operations. An adequate plan is essential for a successful logging operation. Logging operations are expensive and if done incorrectly can be costly to correct. The logging plan often specifies the type of equipment used, the location and size of roads and landings, the location of skid trails, the felling pattern, and the season of operation. Logging planning must consider subsequent forestry operations such as site preparation.

The tools to plan and organize a logging operation include on-the-ground reconnaissance and computer programs that assess logging feasibility, equipment performance, environmental impacts, and economic performance. Planning is often conducted by professional forest engineers.

Training and supervision. Good plans and technical solutions require effective training and on-site supervision. On-site supervision and communication with loggers are vital in controlling soil and water impacts and minimizing resource problems. In many cases, small adjustments by equipment operators can produce significant benefits. For example, skidder

operators can keep the front end of the log fully suspended and not use the skidder blade as a brake to avoid disturbing the soil. Careful supervision is especially important when operating during wet weather or on weak soils to avoid soil and water damage. Trained loggers blend harvesting and environmental knowledge into safe and effective forest operations.

[The authors thank the Forest Engineering Department, Oregon State University, for cooperation in the preparation of this article.]

John Sessions; John Garland

Bibliography. S. Conway, *Logging Practices*, Miller Freeman, 1982; *Model Code of Forest Harvesting Practice*, Food and Agriculture organization of the United Nations, 1996; L. Pancel (ed.), *Tropical Forestry Handbook*, Springer-Verlag, 1993; Planning Forest Roads and Harvesting Systems, *FAO Forestry Pap.*, no. 2, Food and Agriculture Organization of the United Nations, 1977; U. Sundberg and C. Silversides (eds.), *Operational Efficiency in Forestry*, Kluwer Academic, 1988; Workers Compensation Board of British Columbia, *Handbooks for Falling, Skidding, Yarding, Loading, Helicopter Logging*.

Logic

The subject that investigates, formulates, and establishes principles of valid reasoning.

The first attempt to investigate systematically acceptable modes of reasoning was made by Aristotle, in whose *Organon* reasoning was recognized as the subject of a special science. He formulated three basic "laws of thought": (1) the law of contradiction (no proposition is both true and false), (2) the law of excluded middle (each proposition is either true or false), and (3) the law of identity (each proposition implies itself). Advanced as his views of logic were, it seems doubtful that the idea of axiomatizing it ever occurred to him, despite the success of his contemporary Euclid in organizing geometry. That aspect of logic which enunciates or establishes valid reasoning (rather than merely investigating it) began with the publication in 1854 of George Boole's *An Investigation of the Laws of Thought*. The partly abstract treatment of logic presented in this work initiated the completely abstract developments that were to follow. In it the laws of thought are regarded as mere conventions which, like the postulates of euclidean geometry, might be modified or even rejected to create new logics. The only requirement that a "new" logic must satisfy is the one demanded of every deductive system—consistency. Just as different geometries are useful for different purposes, a logic that is appropriate in one environment might not be so in another. There are a growing number of competent individuals, for example, who consider that any logic containing Aristotle's law of excluded middle is not suitable for mathematics. *See* BOOLEAN ALGEBRA; EUCLIDEAN GEOMETRY; GEOMETRY.

This article discusses some of the principles for testing inferences, the use of formal procedures in clarifying arguments, and how important parts of logic have been developed axiomatically. The classical theory of the syllogism will be examined, and the deductive theory of truth functions (the calculus of propositions) will be presented in some detail. Finally, mathematical logic, which arose from work on the foundations of mathematics and encompasses proof theory, set theory, model theory, and recursion theory, will be discussed. For the most part, the discussion will be limited to what is called symbolic logic, for a special set of symbols is employed to exhibit in the clearest way the (at times) complicated logical relations involved. The advantages of a thoroughgoing use of a good symbolization are enormous. It is by such means that modern logic is able to solve problems that were quite beyond the power of logicians of an earlier time, and to render some of the achievements of those logicians almost trivial by comparison.

The syllogism. The formulation of the syllogism is contained in Aristotle's *Organon*. It had a great fascination for medieval logicians, for almost all their work centered about ascertaining its valid moods. The three characteristic properties of a syllogism are as follows:

1. It consists of three statements, the third one (conclusion) being a logical consequence of the first two (the premises).

2. Each of the three sentences has one of the following four forms.

(A) All X is Y (universal affirmative judgment).
(E) No X is Y (universal negative judgment).
(I) Some X is Y (particular affirmative judgment).
(O) Some X is not Y (particular negative judgment).

3. The three sentences contain altogether three terms, S (the subject of the conclusion), P (the predicate of the conclusion), and M (the middle term), that occur in each of the premises. P is called the major term, S the minor term, and a premise is called major or minor depending on whether it contains P or S, respectively. It is by virtue of the middle term, M, which occurs in both premises, that a relation between S and P can be asserted. Thus the three statements

No professors are rich.
All handsome men are professors.
No handsome men are rich.

form a syllogism in which M stands for professors, P for rich, and S for handsome men. This particular syllogism is valid (that is, the conclusion is a consequence of the premises), and hence the validity of a syllogism is seen to be independent of the truth or falsity of its conclusion. Notice that the major premise, minor premise, and conclusion of the above syllogism are statements of type (E), (A), and (E), respectively.

The four figures of the syllogism are represented

by notation (1). Each symbol *MP*, *SM*, and so on

$$\begin{array}{cccc} \text{(I)} & \text{(II)} & \text{(III)} & \text{(IV)} \\ \\ MP & PM & MP & PM \\ SM & SM & MS & MS \\ SP & SP & SP & SP \end{array} \qquad (1)$$

stands for one of the four statements (A), (E), (I), and (O); for example, *MP* might symbolize the statement "No *M* is *P*." Hence each figure has $4^3 = 64$ possible forms, and consequently there are $4 \cdot 4^3 = 256$ possible forms for a syllogism. The great attention given the syllogism by logicians of the Middle Ages resulted in establishing that only 19 of the possible 256 forms that a syllogism might take are admissible (that is, yield valid conclusions). It should be noted that the result they obtained is correct provided one accepts the Aristotelian principle that the statement "All *X* is *Y*" implies that "There are *X*'s." It was also maintained that "No *X* is *Y*" implies that "Some *X* is not *Y*." Thus, according to early logicians, universals have existential implications. Modern logicians do not take this point of view. This results in eliminating 4 of the 19 admissible forms (moods). Three-syllable mnemonic words were coined in which the vowels indicated, in order, those combinations of the statements (A), (E), (I), and (O) that resulted in valid syllogisms. For figure (I) they are *Barbara*, *Celarent*, *Darii*, *Ferio*; figure (II), *Cesare*, *Camestres*, *Festino*, *Baroco*; figure (III), *Datisi*, *Feriso*, *Disamis*, *Bocardo*, *Darapti*, *Felapton;* and figure (IV), *Calemes*, *Fresison*, *Dimatis*, *Bamalip*, and *Fesapo*. The illustrative syllogism given above is Celarent, figure (I). The four moods that depend for their validity on the classical assumption that universals imply existence (and which must be eliminated if that assumption is not made) are *Darapti*, *Bamalip*, *Fesapo*, and *Felapton*.

The syllogism is now of little more than historical interest. Its limitations were recognized quite early, for it was clear that few of the inferences needed in mathematics could be put in syllogistic form. As remarked by Hans Reichenbach, "It is only the combination of a logic of functions of several variables with the logic of propositions that leads to the full understanding of conversational language as well as of mathematics."

Propositional calculus. The unit of the propositional calculus is the proposition, which may be defined as the meaning of an indicative or declarative sentence. Each proposition has a truth-value; it is either true or false in the classical logic, but may assume other truth-values (for example, uncertain) in some of the new logics. Thus, "To err is human" is a proposition, whereas "Lead us not into temptation" is not a proposition. It is convenient to use letters p, q, r, and so on to denote propositions. Propositions are combined by logical connectives to form other propositions. Chief among these are (1) negation (the negation of a proposition p, symbolized by p', may be formed verbally by writing "It is not the case that" before p); (2) conjunction (a binary operation symbolized by $\cdot$ or &, for example, $p \cdot q$, read "p and q"); (3) disjunction (a binary operation symbolized by $+$ or $\vee$; for example, $p \vee q$, read "p or q"); (4) implication (a binary operation symbolized by $\supset$ or $\rightarrow$, for example, $p \rightarrow q$, read "p implies q"); and (5) equivalence (a binary operation denoted by $\equiv$ for example, $p \equiv q$, read "p is equivalent to q," means that each proposition implies the other). The proposition $p \cdot q$ asserts that both p and q are true. It is important to note that the disjunctive "or" is always used in the inclusive sense; that is, $p \vee q$ asserts that at least one, possibly both, of the propositions p and q is true. In addition, the meaning of "implication" in logic differs markedly from the ordinary sense of that term. The proposition $p \rightarrow q$ asserts merely that it is not the case that p is true and q is false; that is, in classical logic $p \rightarrow q$ means that either p is false or q is true. Thus any false proposition (for example, Napoleon Bonaparte was born in the year 1600) implies any proposition whatever. It implies, New York is a large city, or Cleopatra was queen of France. Similarly, any true proposition is implied by any proposition whatever. In an implication $p \rightarrow q$, p is called the antecedent and q the consequent. The logical connectives introduced above are examples of logical constants. They are not independent notions. Implication may be defined in terms of negation and conjunction, and this is true for all of the connectives. By use of the symbol $\equiv_D$ to mean "is defined," notation (2) follows, for example.

$$\begin{gathered} p \cdot q \equiv_D (p' \vee q')' \qquad p \supset q \equiv_D p' \vee q \\ p \equiv q \equiv_D [(p' \vee q)' \vee (q' \vee p)']' \end{gathered} \qquad (2)$$

In fact, all the connectives can be defined in terms of one binary operation, the so-called Scheffer stroke function $p|q$, which may be read "p is false or q is false." Therefore, notation (3) follows.

$$\begin{gathered} p' \equiv_D p|p \qquad p \cdot q \equiv_D (p|q)|(p|q) \\ p \vee q \equiv_D (p|p)|(q|q) \\ p \supset q \equiv_D [(p|p)|(p|p)]|(q|q) \end{gathered} \qquad (3)$$

The logical connectives, the letters p, q, r, and so on, that stand for propositions (propositional variables), and the parentheses, brackets, and braces needed for punctuation are formal concepts. A propositional function is a combination of concepts, formal or nonformal, involving at least one variable, which is not a proposition but becomes one when all the variables are given specified values. Examples are (1) p and q, (2) $x^2 + y^2 = 1$, and (3) x is president of the United States. Unlike a proposition, a propositional function has no truth-value. On the other hand, the statement "for every integer x, $x^2 + 2 = 0$" is a proposition (not a propositional function) even though it contains a (bound) variable. A formal function is a propositional function that contains only formal concepts—for example $[(p \vee q) \cdot r] \supset s$. The form of a given proposition is the formal function obtained from it by substituting formal concepts for all nonformal ones. Thus $p \cdot q$ is the form of the proposition "Leonard is happy and the moon is blue." Notice that in formalizing the above proposition, different propositions were replaced by

different propositional variables. A truth function is a propositional function in which only propositional variables occur, and such that the truth value of each proposition obtained from the function by substituting specific propositions for the variables depends only on the truth values of those propositions. Examples are (1) $(p \cdot q) \vee r'$, (2) Napoleon was a great general and p (where p is a propositional variable). On the other hand, "Jones stated that p" is not a truth function.

This article is concerned principally with the class of formal truth functions (that is, those containing only formal concepts) and with an important subclass called tautologies [that is, those formal truth functions such that every proposition obtained from them by substituting specific values (propositions) for the variables they contain has the truth value T (true)]. A truth function is contingent provided it assumes both truth values T and F (false)—here only the classical two-valued (Aristotelian logic) truth functions are considered—and self-inconsistent provided it has only the value F. The character of a truth function is determined when it is shown to be self-inconsistent, contingent, or a tautology.

Character of a truth function. A truth function may be analyzed by means of a truth table, or matrix. Such tables are made first for the elementary truth functions p', $p \vee q$, $p \cdot q$, $p \supset q$ as follows:

(I)

p	p'
T	F
F	T

(II)

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

(III)

p	q	$p \cdot q$
T	T	T
T	F	F
F	T	F
F	F	F

(IV)

p	q	$p \supset q$
T	T	T
T	F	F
F	T	T
F	F	T

In every table each row of the last column gives that truth value of the function which is determined by those truth values (assigned to the variables) that are entered in the preceding columns of that row. For example, the second row of table (IV) states that the truth function $p \supset q$ has value F. Since all logical connectives are used in the logic of propositions are expressible in terms of negation and disjunction, only tables (I) and (II) are really essential, but tables (III) and (IV) are convenient when analyzing complicated truth functions. Although the character of any truth function may be determined by constructing its truth table, the method is tedious when the function contains a large number of variables, since the truth table of a truth function of n variables has 2^n rows. As an example, the character of the truth function $(p \supset q) \supset (q' \supset p')$ may be determined. Using tables (I) and (IV), the truth table of the function is found to be

p	q	p'	q'	$p \supset q$	$q' \supset p'$	$(p \supset q) \supset (q' \supset p')$
T	T	F	F	T	T	T
T	F	F	T	F	F	T
F	T	T	F	T	T	T
F	F	T	T	T	T	T

Since every entry in the last column is T, the function $(p \supset q) \supset (q' \supset p')$ is seen to be a tautology. It may also be observed that in each row the entries in columns 5 and 6 are the same; that is, for any given values of p and q, the functions $p \supset q$ and $q' \supset p'$ assume the same truth value. Hence $(q' \supset p') \supset (p \supset q)$ and the two functions are logically equivalent. The principle embodied in this equivalence (the law of contraposition) is often used in mathematics.

Another device for ascertaining the character of a truth function is the method of negation. It sometimes happens that the character of the function obtained by negating a given function can be easily established, and since a function is contingent if and only if its negation is, self-inconsistent if and only if its negation is a tautology, and a tautology provided its negation is self-inconsistent, a knowledge of the character of the negated function solves the problem. The method makes use of the so-called De Morgan formulas $(p \vee q)' = p' \cdot q'$, $(p \cdot q)' = (p' \vee q')$, in which the sign $=$ signifies that an expression appearing on one side of it may be replaced by the expression appearing on the other side, and the boolean expansion of a truth function. This is an expansion that exhibits all of the conjunctive possibilities for truth value T; It is a disjunction of mutually exclusive (incompatible) conjuncts. For example, $p \vee q$ assumes truth value T if and only if the function $(p \cdot q') \vee (p \cdot q) \vee (p' \cdot q)$ does. The latter function is the boolean expansion of $p \vee q$. The method of negation can be illustrated first by considering the same truth function $(p \supset q) \supset (q' \supset p')$ analyzed by the matrix method. Writing Eq. (4) and noting that

$$\begin{aligned}[(p \supset q) \supset (q' \supset p')]' &= (p \supset q) \cdot (q' \supset p')' \\ &= (p' \vee q) \cdot (q \vee p')' \\ &= (p' \vee q) \cdot (p' \vee q)' \qquad (4)\end{aligned}$$

$(p' \vee q) \cdot (p' \vee q)'$, as the conjunction of a proposition and its negation, is self-inconsistent, it follows that $[(p \supset q) \supset (q' \supset p')]'$ is self-inconsistent, and so $(p \supset q) \supset (q' \supset p')$ is a tautology. A second example, using the boolean expansion of $p \vee q$, is Eq. (5). Clearly, the truth function $p \cdot q'$ is

$$\begin{aligned}[(p \vee q) \supset q]' &= (p \vee q) \cdot q' \\ &= [(p \cdot q) \vee (p \cdot q') \vee (p' \cdot q)] \cdot q' \\ &= p \cdot q' \qquad (5)\end{aligned}$$

contingent, and consequently so is the function $(p \vee q) \supset q$.

Tautologies and forms of argumentation. One truth function T_1 necessarily implies another T_2 provided the truth function $T_1 \supset T_2$ is a tautology. For example, $(p \supset q) \supset (q' \supset p')$ is a tautology, and so the truth function $p \supset q$ necessarily implies the truth function $q' \supset p'$. On the other hand, neither of those two truth functions is a tautology, and so the implications they exhibit are not necessary implications. Valid forms of argumentation are based on the notion of necessary implication. Two such forms were known to classical logic as *modus ponendo ponens* and *modus tollendo tollens*. The first states that if an implication

$p \supset q$ and its antecedent p are asserted, then the consequent q may be asserted; that is, the truth function $(p \supset q) \cdot p$ necessarily implies q. This is established by proving that the truth function $[(p \supset q) \cdot p] \supset q$ is a tautology. The *modus tollendo tollens* form of argument permits the denial of the antecedent of an asserted implication when its consequent is denied. The validity of the argument rests upon the tautology $[(p \supset q) \cdot q'] \supset p'$. A common fallacy (fallacy of the consequent) is to assert the antecedent of an implication whose consequent has been asserted. This is not permissible since the truth function $[(p \supset q) \cdot q] \supset p$ is not a tautology. Although such an argument is logically unwarranted, it is frequently invoked in testing experimentally a hypothesis p, for if q follows from p, numerous instances of the validity of q are sometimes taken to furnish strong evidence of the validity of p.

A traditional form of argumentation known as the dilemma is an application of *modus ponendo ponens* or *modus tollendo tollens*. Consider an example:

1. If I speak the truth, men will hate me; if I lie, the gods will hate me.
2. I must either speak the truth or lie.
3. Therefore, either men will hate me, or the gods will hate me.

The form of this dilemma is as follows:

1. $(p \supset q) \cdot (r \supset s)$
2. $p \vee r$
3. $q \vee s$

The argument is valid by *modus ponendo ponens,* for since premise (2) affirms the antecedent of at least one of the implications of premise (1), at least one of the consequents may be asserted, which is what the conclusion (3) does.

Deductive theory. The intent of this section is to sketch the procedure by which all tautologies are deducible from a set of just four of them. This is analogous to what is done in an axiomatic geometry in which all the theorems of the geometry are deduced from a selected set of axioms or postulates that are assumed to be true. The nature of the geometry is fully determined by the set of statements forming the postulates.

Let P denote a nonempty abstact set whose elements, denoted by $p, q, r,$ and so on, are called propositions (for suggestiveness), and suppose P is closed with respect to a unary operation, denoted by $(')$, and a binary operation, denoted by $\vee$; that is, if p and q are any elemets of P, then p' and $p \vee q$ must be unique elemets of P. An additional binary operation, denoted by $\supset$ and read "implies" is defined in P:

If p, q are elements of P, $p \supset q \equiv_D p' \vee q$.

An expression formed by elements of P and the operations $(')$, $\vee$ is a truth function if and only if it conforms to the following syntactical rules: (1) $p, q, r,$ and so on are truth functions; (2) if A and B are truth functions, then so are A' and $\vee B$; (3) no sequence of symbols denotes a truth function unless it is constructed according to (1) or (2), or unless it is replaceable by definition by a truth function. A subclass of the class of truth functions constitutes the class of tautologies. It is assumed that the following four truth functions are tautologies (these are the postulates):

P1 $(p \vee p) \supset p$
P2 $q \supset (p \vee q)$
P3 $(p \vee q) \supset (q \vee p)$
P4 $(q \supset r) \supset [(p \vee q) \supset (p \vee r)]$

These simple tautologies were selected by Alfred North Whitehead and Bertrand Russell in their monumental work *Principia Mathematica* as the basic set from which all tautologies can be derived. They had an additional postulate

P5 $[p \vee (q \vee r)] \supset [q \vee (p \vee r)]$

which was later shown by Paul Bernays to be deducible from the first four postulates, and hence may be deleted as a postulate. None of the three Aristotelian principles (laws of identity, contradiction, and excluded middle) are in the basic set. They are all deduced as theorems.

Before any theorems can be deduced there must be available rules that make it possible to go from one step in a proof to another. Among such rules are those of formal substitution, according to which any truth function may be substituted for any of the variables occurring in a postulate or theorem, provided the substitution is carried out completely, and any expression may be replaced by one definitionally identical. Among the rules of formal assertion are the principle of inference (that is, if P and $P \supset Q$ are postulates or proved theorems, then Q is a theorem) and the principle of adjunction (that is, if P, Q are postulates or theorems then $P \cdot Q$ is a theorem).

It was proved by E. L. Post that postulates P1–P4 are consistent; that is, it is impossible to deduce from them truth functions X and X'. He also proved that the four postulates form an independent set; that is, no one of the postulates can be deduced from the others.

Non-Aristotelian logics. In these logics the principle of the excluded middle (that is, each proposition is either true or false) is not valid. There are, then, at least truth values possible for a proposition, and for this reason such logics are known as many-valued. In a many-valued logic the *reductio ad absurdum* procedure that is so frequently used in mathematics is much less effective than it is in two-valued classical logic, for if it is not possible to conclude (as is permissible in Aristotelian logic) that the proposition is true. The Aristotelian tautology $(p')' \equiv p$ is not a tautology in a logic with at least three truth values. Doubts that the classical logic is an appropriate one to use in mathematics stem from the modern view that the concepts and the relationships dealt

with in mathematics have no real existence in the external world. It would seem to follow that the a priori nature of "true" or "false" then disappears and with it the law of excluded middle! If "true" is interpreted in mathematics to mean "provable," and a proposition is called false provided its negation is provable, then the law of excluded middle is demonstably false, since it has been shown that mathematics contains propositions p such that neither p nor its negation p' is provable. The intuitionistic school of the Dutch mathematician L. E. J. Brouwer rejected the excluded middle law, and it forms no part of the logic of that school, formulated by A. Heyting. *See* FUZZY SETS AND SYSTEMS. Leonard M. Blumenthal

Mathematical logic. Mathematical logic is an area of research that has emerged from the study of formal systems. It has its roots in work on the foundations of mathematics in the nineteenth and early twentieth centuries, showing that the customary forms of mathematical reasoning can be adequately expressed in formal (purely symbolic) systems. Mathematical logic contains four areas of research: proof theory, set theory, model theory, and recursion theory.

Formal systems. The original objective of work in the foundations of mathematics was the analysis and precise formulation of mathematical concepts, such as "limit," "function," and "differential manifold," as well as a complete and explicit determination of the valid forms of mathematical argument. This goal was achieved by the formalization of mathematics.

A formal system consists of a symbolic notation (or language) with rules for constructing expressions using the specified notation, axioms (expressions taken as given), and rules of inference used to derive additional expressions from a given set of expressions. To formalize an area of mathematics means to give a procedure for representing proofs in the specified area within a definite formal system. The process of formalization requires the construction of a formal language in which all the fundamental concepts of the theory can be adequately represented, as well as the specification of appropriate axioms and rules of inference. One formal system that is adequate for the formalization of all existing mathematics is the language of set theory, which can be set up in a variety of ways. There are various standard systems of axioms for set theory, notably the axiom system ZFC, consisting of axioms proposed by E. Zermelo and A. Fränkel, with the addition of the axiom of choice. A typical formalized axiom, the power set axiom of ZFC, is displayed in notation (6).

$$(\forall u)(\exists v)(\forall w)((w \in v) \Longleftrightarrow (\forall x)(x \in w \Longrightarrow x \in u)) \tag{6}$$

Proof theory. Proof theoretic research continues the study of formal systems that are adequate for the formalization of various parts of mathematics. Two basic results, due to K. Gödel, relate to the adequacy of formal systems in general: the so-called completeness and incompleteness theorems.

The completeness theorem essentially states that the laws of inference are adequately mirrored in the usual formal systems in the sense that any law of inference that is in fact valid in all mathematical contexts is derivable in any of these systems.

The incompleteness theorem, on the other hand, can be interpreted as stating that it is impossible to give a completely satisfactory axiom system for any theory that can serve as a formal foundation for mathematics. More precisely, for any axiom system that is adequate for the development of number theory and that is explicitly presented, it is possible to write down a true sentence of number theory that is not derivable from the given axioms. As a result, although the logic actually used by mathematicians in their proofs can be fully captured by a single formal system, at a certain theoretical level any formalization is necessarily incomplete. One way to prove the incompleteness theorem is to show that, for any explicitly given formal system T capable of formalizing number theory, the statement "T is a consistent system" can be formalized in T, and this formalized consistency statement is not provable within the system T.

Modern proof theory studies the strength of a variety of formal systems, generally using a convenient formalism due to G. Gentzen.

Axiomatic set theory. Axiomatic set theory was developed in part as a way of obtaining a formalism adequate for all of mathematics, and also for concrete mathematical reasons connected with the development of general topology and measure theory around 1900. Set theory is probably the most active research area within mathematical logic. It has been shown that a variety of problems in analysis, general topology, and algebra (especially homological algebra) cannot be settled within the framework of any of the standard set theoretic systems like ZFC. Thus such questions cannot be settled rigorously unless additional axioms of set theory (or equivalent methods of proof) are accepted. Many such axioms have been proposed, and their relationships intensively investigated, but no consensus has emerged as to which additional axioms should be accepted. *See* SET THEORY.

Model theory. The development of formal systems in logic is related to the development of axiomatic systems in modern algebra, although the motivation is very different. Model theory combines the algebraic and logical points of view. In applied model theory, the tools of mathematical logic are applied to problems arising in algebra. In pure model theory, axiomatic systems are studied as if they were algebraic systems, and the class of all models of a given axiomatic system is investigated.

The basic result in model theory is the compactness theorem, stating that an infinite axiom system will have a model whenever all of its finite subsystems have models. This follows from the completeness theorem. One curious consequence of the compactness theorem is the existence of mathematical structures that have all the formal properties

of the real number system but contain so-called infinitesimal elements as well; this gives a rigorous foundation for differential and integral calculus that appears closer in spirit to the original seventeenth-century viewpoint than the so-called (ϵ, δ) formulation of A. Cauchy and K. Weierstrass. *See* CALCULUS; DIFFERENTIATION; INTEGRATION.

Recursion theory. The notion of a computable function was defined rigorously in connection with problems in formal logic prior to the development of electronic computers. Such a definition is needed, for example, in order to state precisely what is meant by an explicitly given axiom system. A wide variety of noncomputable functions arise in mathematics; many of them correspond to concrete problems that cannot be solved algorithmically.

For example, the Diophantine problem, associated with Diophantos of Alexandria in the third century and reformulated by D. Hilbert in 1900, asks for a procedure to determine whether an arbitrary polynomial equation in several variables with integer coefficients has an integer solution. It was shown in 1970 that this problem has no algorithmic solution; in other words, it is not possible to write a computer program that solves all instances of this problem correctly. This result depends on a substantial refinement of the encoding technique used by Gödel in the proof of his incompleteness theorem, and the proof also requires a precise notion of computability. The theoretical analysis of the notion of computability was also of some importance in developing the electronic computer, particularly in connection with the work of A. Turing and J. von Neumann during World War II.

Given two functions, f and g, from the natural numbers to the natural numbers, f is said to be reducible to g if there is an algorithm that will compute f by using information about g. For example, any function is reducible to itself, and a computable function is reducible to any other function. Two functions are equivalent if each can be reduced to the other. The equivalence classes are called degrees; they can be thought of as degrees of complexity. Modern recursion theory has produced a very detailed picture of the structure of the degrees and the reducibility relation. Modified notions of reducibility are used also in theoretical computer science, in which all algorithms are assumed to satisfy additional restrictions in the amount of time or space used in the computation. *See* DIGITAL COMPUTER; MATHEMATICS; SCIENCE; SCIENTIFIC METHODS. Gregory Cherlin

Bibliography. S. Barker, *Elements of Logic*, 5th ed., 1989; I. M. Copi and C. Cohen, *Introduction to Logic*, 10th ed., 1998; H. J. Gensler, *Introduction to Logic*, 2002; D. L. Johnson, *Elements of Logic via Numbers and Sets*, 2001; E. Lepore, *Meaning and Argument: An Introduction to Logic Through Language*, 2d ed., 2003; E. Mendelson, *Introduction to Mathematical Logic*, 4th ed., 1997; E. Nagel, J. R. Newman, and D. R. Hofstadter, *Gödel's Proof*, 2d ed., 2002; G. Takeuti, *Proof Theory*, 2d ed., 1987; J. van Heijenoort, *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, 1967.

Logic circuits

The basic building blocks used to realize consumer and industrial products that incorporate digital electronics. Such products include digital computers, video games, voice synthesizers, pocket calculators, and robot controls. *See* CALCULATORS; DIGITAL COMPUTER; ROBOTICS; VIDEO GAMES; VOICE RESPONSE.

The change that has enabled widespread economical use of digital logic once found only in very expensive, room-sized computers has been a dramatic evolution in device technology. Logic circuits which comprise several basic electronic devices (typically transistors, resistors, and diodes) were once designed with each device as a separate physical entity. Now, very large scale integration delivers millions of equivalent basic devices on one piece of silicon typically rectangular with maximum dimensions of a few tenths of an inch (0.1 in. = 2.5 mm) per side. *See* INTEGRATED CIRCUITS.

This dramatic reduction in size and concomitant increase in integrated circuit capabilities has been accompanied by a number of effects. Both the cost and power consumption per logic device have been greatly reduced, while increasingly powerful computing capability has been incorporated into ever smaller packages. The modestly priced, multicapable, digital watch that can run for over a year on a tiny battery exemplifies this progression.

Logic circuit operation. Logic circuits process information encoded as voltage or current levels. The adjective "digital" derives from the fact that symbols are encoded as one of a limited set of specific values. While it would be feasible to encode information using many distinct voltage levels (for example, each of 10 voltage levels might correspond to a distinct decimal digit value), there are several reasons for using binary, or two-valued, signals to represent information: the complexity of the sending and receiving circuit is reduced to a simple on-off-switch type of operation corresponding to the two-valued signals; the speed of operation is far greater since it is not necessary to wait until a changing signal has had time to converge to a final value before interpreting its binary value; and the on-off operation makes the circuitry very tolerant of both changing characteristics of devices due to aging or temperature–humidity environment and electromagnetically induced noise added to the voltage levels.

Since signals are interpreted to be one of only two values (denoted as 0 and 1), these binary digits, or bits, are used in a number system with digit weights that change by multiples of 2 rather than multiples of 10 as in the decimal system. In this system, successive digits or bits have weights of 1, 2, 4, 8, . . . , in contrast to the decimal system where successive digits have weights of 1, 10, 100, ... *See* NUMBER SYSTEMS.

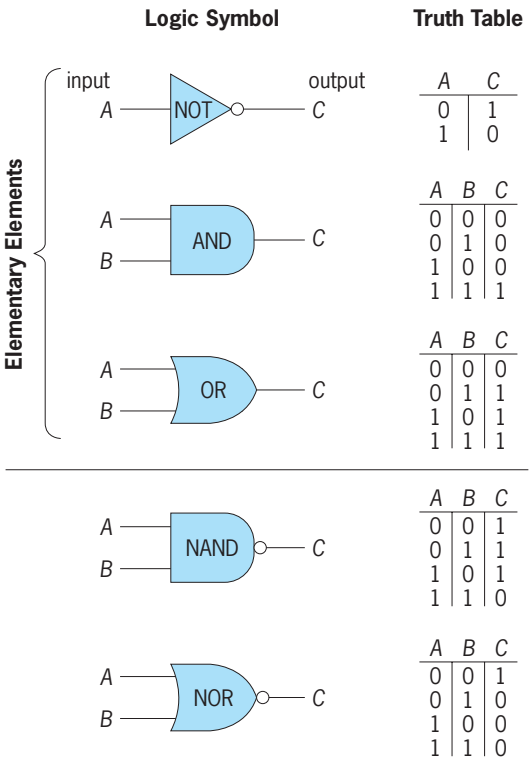

Fig. 1. Logic elements.

Bit patterns also are used in computers to represent alphanumeric characters. A very useful character-encoding convention, called Unicode®, uses a standard collection of 8, 16, or 32 of these binary signals to represent all keyboard-derived symbols in any of the world's languages.

Types of logic functions. All logic circuits may be described in terms of three fundamental elements (**Fig. 1**). The NOT element has one input and one output; the output generated is the opposite of the input in binary. In other words, a 0 input value causes a 1 to appear at the output; a 1 input results in a 0 output. Two NOT elements in series, one output driving the input to the other, simply reproduce a copy of the input signal.

The AND element has an arbitrary number of inputs and a single output. (Electrical characteristics of devices used or physical packaging of devices result in some practical limit to the number of inputs allowed.) The output becomes 1 if, and only if, all of the inputs are 1; otherwise the output is 0. The AND together with the NOT circuit therefore enables searching for a particular combination of binary signals. If, for example, a signal is needed that goes to 1 only when signals *A*, *B*, and *C* are 1, 0, and 1, respectively, then the circuit indicated in **Fig. 2** will generate the desired output.

The third element is the OR function. As with the AND, an arbitrary number of inputs may exist and one output is generated. The OR output is 1 if one or more inputs are 1.

The operations of AND and OR have some analogies to the arithmetic operations of multiplication and addition, respectively. The collection of mathematical rules and properties of these operations is called Boolean algebra. *See* BOOLEAN ALGEBRA.

While the NOT, AND, and OR functions have been designed as individual circuits in many circuit families, by far the most common functions realized as individual circuits are the NAND and NOR circuits of Fig. 1. A NAND may be described as equivalent to an AND element driving a NOT element. Similarly, a NOR is equivalent to an OR element driving a NOT element. (The reason for this strong bias favoring inverting outputs is that the transistor, and the vacuum tube which preceded it, are by nature inverters, or NOT-type circuit devices, when used as signal amplifiers.) An interesting property of the NAND or NOR circuits is that all logic functions may be accessed using either type of circuit. A NOT element, for example, is realized as a one-input NAND. An AND element is realized as NAND followed by a NOT element. An OR element is realized by applying a NOT to each input individually and then applying the resulting outputs to the NAND as inputs.

Combinational and sequential logic. Logic circuits respond to combinations of input signals. In an arithmetic adder, for example, a network of logic elements is interconnected to generate the sum digit as an output by monitoring combinations of input digits; the network generates a 1 for each output only in response to those combinations in the addition table that call for it. Logic networks which are interconnected so that the current set of output signals is responsive only to the current set of input signals are appropriately termed combinational logic.

A further capability for processing information is memory, or the ability to store information. In digital systems, the memory function has been provided for by a variety of technologies, including magnetic cores, stored charge, magnetic bubbles, and magnetic tape. The logic circuits themselves must provide a memory function if information is to be manipulated at the speeds the logic is capable of. The logic elements defined above may be interconnected to provide this memory by the use of feedback. *See* COMPUTER STORAGE TECHNOLOGY.

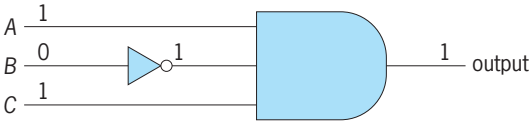

Fig. 2. Logic circuit whose output is 1 only when input signals *A*, *B*, and *C* are 1, 0, and 1, respectively.

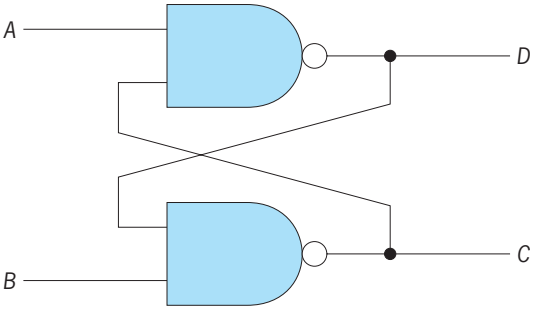

Fig. 3. A memory element.

The circuit in **Fig. 3** illustrates a basic and perhaps the most commonly used form of memory circuit. Normally both inputs A and B to the cross-coupled NAND gates are at the voltage corresponding to logic 1; if a 0 is applied to signal A, outputs C and D assume the values 0 and 1, respectively, based on the definition of NAND operation; the outputs retain these values after input A is returned to the quiescent value of 1. Conversely, if B is set to 0 while A is 1, the outputs C and D become 1 and 0, respectively, and again retain these values after B is returned to a value of 1. In the jargon of logic circuit design, this circuit is called a flip-flop, toggle, or trigger. Like the light switch on the wall, the circuit retains one of two stable states after the force initiating change to that state is removed. This circuit is therefore well matched to the storage of binary, or two-valued, signals. Logic circuit networks that include feedback paths to retain information are termed sequential logic networks, since outputs are in part dependent on the prior input signals and in particular on the sequence in which the signals were applied. *See* SWITCHING CIRCUIT.

Logic circuit embodiment. Alternatives exist for the digital designer to create a digital system, and seven are summarized below.

1. Ready-made catalog-order devices can be combined as building blocks. In this case, manufacturers have attempted to provide a repertoire of logic networks that will find common usage and generality. Some examples of the types of circuits in digital circuit manuals or catalogs are arrays of individual logic gates (typically four to six) with outputs and inputs available as external connections to the device; arrays of memory elements (flip-flops) used to hold a collection of binary signals or digits; arithmetic adders capable of adding four corresponding binary digits from each of two binary numbers and of being cascaded to produce an adder for any multiple of four digits; and multipliers capable of forming the product of two 16-bit numbers in a single step (that is, a combinational multiplier).

The significance of these examples is that the number of gates or logic elements realized on individual devices ranges from a few to several thousand. Clearly, a design that can use the devices with larger gate counts is likely to be more compact and more economical.

2. An increasingly common realization of logic gates is custom-designed devices. Two factors which make this practical are computer-aided-design (CAD) tools and libraries of subnetworks that can be used as building blocks. Custom design of a logic network implies the generation of highly complicated artwork patterns, called masks, for use in photographic-development-like steps used to produce the integrated circuit networks. The availability of a library of subnetworks for which this artwork has already been generated reduces the additional artwork required to interconnection patterns among these library elements to set up the desired logic network. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

3. Another option is gate-array devices. In this case, the device vendor manufactures devices comprising a two-dimensional array of logic cells. Each cell is equivalent to one or a few logic gates. The final layers of metallization that determine the exact function of each cell and interconnect the cells to form a specific network are deferred until the customer orders such a device. This procedure uses mass production for the majority of processing steps necessary to manufacture a device, including most of the necessary artwork. Since the interconnecting metallization layers are a relatively small and simple part of the total device fabrication, the customization cost and time can be reduced significantly by this method relative to a totally custom fabrication. The sacrifice in this approach is that the packing density of the two-dimensional array will tend to be less than a custom-designed layout, since routing channels must allow for reasonably general interconnection patterns.

4. An additional option in realizing logic functions is programmable logic arrays (PLAs). The manufactured part has the potential for realizing any of a large number of different sets of logic functions; the particular functions produced are determined by actually blowing microscopic fuses built into the device, in effect removing unwanted connections. In one device, each of 16 device-input signals is an input to each of 48 AND gates; the 48 AND outputs are input signals to 8 OR gates that generate the device outputs. Selective removal of inputs to the AND gates, and between the AND and OR gates, enables implementation of a wide variety of functions.

5. Another embodiment of logic network functions involves the use of table look-up. A logic network has several binary input signals and one or more binary outputs. The transformation of input-signal combinations to output-signal values need not be realized by AND-OR-type logic elements; instead, the collection of input signals can be grouped arbitrarily as address digits to a memory device. In response to any particular combination of inputs, the memory device location that is addressed becomes the output. For example, a logic network function involving 10 binary inputs and 8 binary output signals could be realized with a single memory device that holds 1024 memory cells, with 8 bits stored in each cell. (Ten bits address 2^{10} or 1024 locations.)

The form of memory device used for table look-up is usually read-only; that is, memory contents are defined as part of the manufacturing process by means of the last layer of metallization, or one of a number of other means is used to fix the memory cell contents so that it is not necessary to reload the memory after each turn-off of power. Typically, this approach may lead to slower operation than the use of logic circuits, but in some cases can lead to very significant economies.

6. The field-programmable gate array (FPGA) is another embodiment of logic functions. The logic network on an FPGA is configured by loading gate interconnection data over certain FPGA pins after power is applied. The FPGA retains this configuration until its power source is disconnected or until it is reprogrammed. Powerful high-end FPGAs can include onboard microcomputers in addition to their

usual arrays of reconfigurable logic gates. *See* FIELD-PROGRAMMABLE GATE ARRAYS.

7. The last form of logic network embodiment that will be discussed is the microcontroller. A microcontroller is a single device that includes a read-only memory to hold a program, a processor capable of reading and executing that program, and a small read-write memory for scratch working space. Just as a memory device provides an efficient realization of a combinational network, the equivalent of a highly complex sequential network can be had with a microcontroller. Two advantages of this approach are that little or no custom-device fabrication is required and that the programmability permits utilization of complicated and modifiable equivalent networks. As with table look-up, the principal disadvantage is that the speed of operation may be slower than if an actual network of high-speed logic circuits was used. *See* MICROCOMPUTER; MICROPROCESSOR.

Technology. There are basically two logic circuit families in use: bipolar and metal-oxide-semiconductor (MOS), with MOS technology the dominant one.

The basic MOS device is formed by using a silicon substrate which has been doped in such a way as to greatly increase the hole/electron relative density, forming a *p*-type substrate. In addition to the substrate, two diffused *n*-type regions are formed. A layer of metal, insulated from the substrate by a deposit of oxide, is situated between the two *n* regions. This metal connection is known as the gate, and the two *n* regions are called the source and drain; since the device is symmetrical in every respect, source and drain are functionally interchangeable. When a sufficient voltage exists between the metal gate and the substrate, electrons are conducted between the drain and source connections, essentially shorting the two together. Otherwise, a high resistance exists between the two. Thus, the gate input is analogous to the control lever of a switch. In like manner, an *n*-type doping of the substrate, along with a *p*-type source and drain, may be used to implement the switch. This will result in hole migration between source and drain when the gate is turned on. *See* SEMICONDUCTOR; TRANSISTOR.

The simplest MOS logic structure is the transmission gate (also known as pass transistor and steering logic). Only one transistor is used to implement the function, where source and gate leads represent the two input variables, and the drain becomes the output value. The transistor is turned on when the gate voltage is above the threshold voltage, assuring sufficient current flow through the channel; in other words, it acts as a short between source and drain. This is analogous to a mechanical switch in the "on" position. When the gate is below the threshold potential, the switch is opened, thus performing an AND-like function on the inputs.

The NAND logic function is also easily derived with MOS circuits. **Figure 4** shows a sample NAND circuit in which the transistors *A*, *B*, and *C* are configured in series and must therefore all be in the "on" state for a conduction path from output *Z* to ground to be established. If *A*, *B*, and *C* are all above threshold, output *Z* is shorted to ground or is forced to logical zero. On the other hand, if any of *A*, *B*, or *C* is off, an open circuit essentially exists between *Z* and ground. In this case, output node *Z* exhibits much less resistance to the power source ($R \ll R_a + R_b + R_c$) and therefore will be at the voltage potential of the power source. (The resistance *R* is commonly formed by a type of MOS device known as a depletion-mode transistor.)

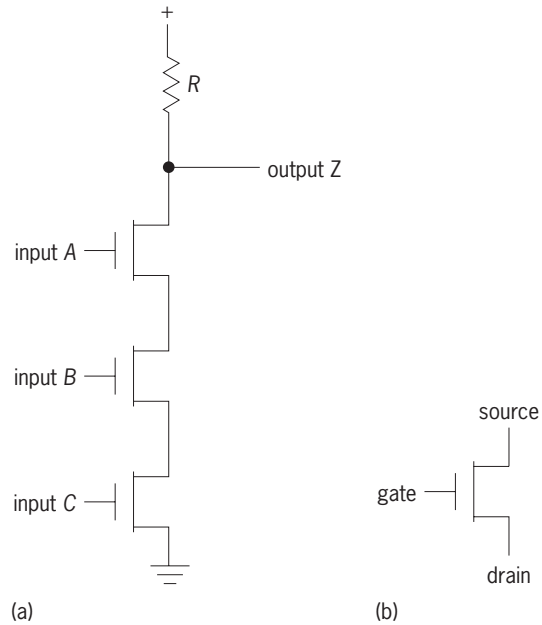

Fig. 4. MOS NAND gate. (*a*) Circuit configuration. (*b*) MOS transistor (*n* or *p*), the basic circuit device.

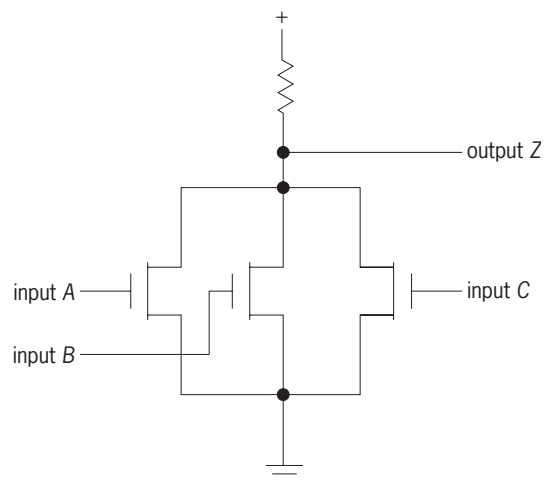

Fig. 5. MOS NOR gate.

This NAND circuit can be reconfigured into the NOR circuit of **Fig. 5**. Here, the transistors are in parallel instead of in series, and when any of the inputs are turned on, the output will be grounded. Otherwise, *Z* will be pulled up to the power supply voltage.

Another useful MOS logic structure is the AND-OR-INVERT, which performs the logic functions exactly as stated in the name. The A-O-I is shown in **Fig. 6** to be a combination of the NAND and NOR configurations.

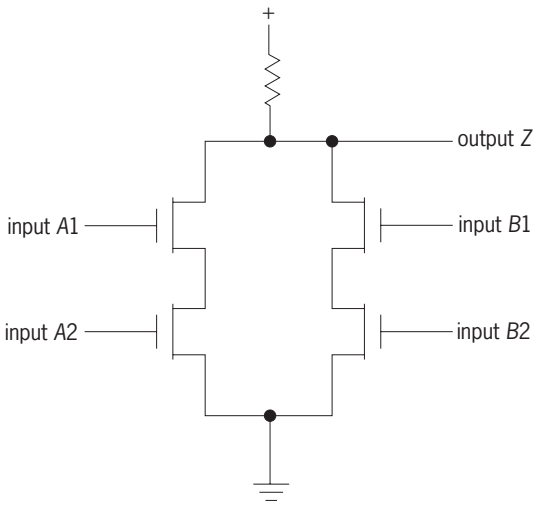

Fig. 6. MOS A-O-I (AND-OR-INVERT) gate.

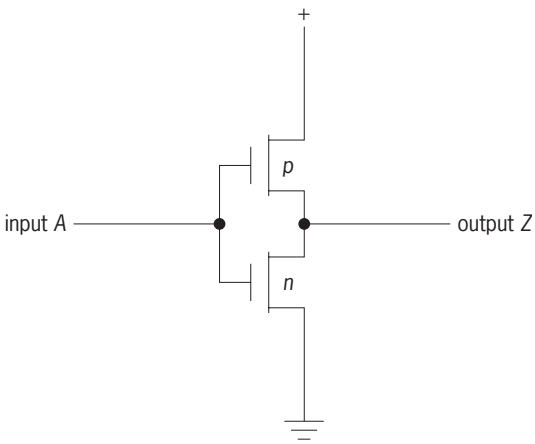

Fig. 7. CMOS inverter.

Output Z is low only if the function ($A1$ AND $A2$) OR ($B1$ AND $B2$) is high.

As described earlier, either an n-channel substrate (NMOS) or a p-channel substrate (PMOS) may be used to implement the logic structures. Additionally, both types may be used concurrently in the logic circuit to fabricate a more sophisticated device. **Figure 7** shows an inverter built with both NMOS and PMOS devices. When the input is low, the NMOS transistor is cut off and the PMOS transistor conducts, shorting the output to the power source. On the other hand, when the input is high, the PMOS gate is shut off and the NMOS gate shorts the output to ground, thus performing the inverter function. The advantage of such a complementary configuration, known as CMOS, is primarily power dissipation. Owing to the fact that only one complementary transistor conducts at any time (other than the short overlap during switching), there is always a very high resistance from the power source to ground, and therefore very little current flows through the pair. The primary disadvantage of CMOS over the common MOS circuit is the increased fabrication complexity due to the n and p substrates occupying the same chip, and the secondary disadvantage is the increased area needed to build the CMOS structure because of the use of complementary pairs.

The charge carriers in an MOS transistor are either free electrons or holes. For this reason, MOS is known as a unipolar device. In contrast, devices that utilize both free electrons and hole migration, such as the junction transistor, are known as bipolar devices. Although general comparisons are hard to make, bipolar devices typically exhibit much less input resistance and demand a more sophisticated fabrication technique than do MOS devices. In terms of logic realization, the bipolar family encompasses many common circuit types. Perhaps the best-known and most widely used implementation of bipolar logic switches is transistor-transistor logic (TTL). The basic TTL NAND gate (**Fig. 8**) is formed by a multiemitter transistor (turned on only if every input is high) followed by an output transistor that acts as a pullup/buffer. Thus, the first transistor performs an AND operation on the inputs and the second transistor

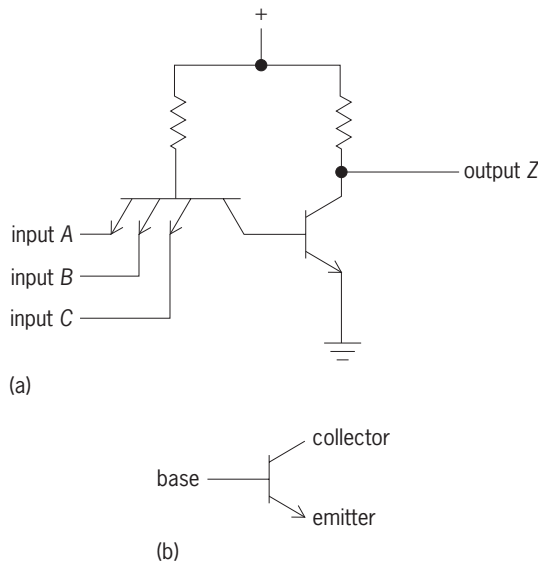

Fig. 8. Simplified TTL (transistor-transistor-logic) NAND gate. (*a*) Circuit configuration. (*b*) Bipolar junction transistor (*npn*), the basic circuit device.

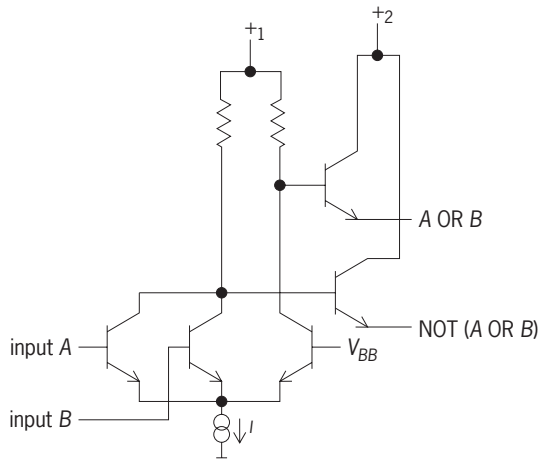

Fig. 9. Simplified ECL (emitter-coupled-logic) OR/NOR gate.

completes the NAND by performing an inversion. *See* JUNCTION TRANSISTOR.

TTL transistors are operated in the saturation mode; in other words, the transistors are driven hard to either the cutoff or the saturation limits. This overdriving introduces a time delay that does not exist if the transistors are operated in the nonsaturated mode. Such nonsaturating logic, while inherently faster, is more susceptible to noise since it is biased in the linear region.

Current-mode logic (CML) is a popular form of nonsaturation logic, and most often takes the configuration of emitter-coupled logic (ECL). The basic ECL gate (**Fig. 9**) is composed of current-steering transistors that perform an OR operation on the inputs. Typically, the gate output is amplified by an emitter-follower transistor, and both the true and complement signals are available simultaneously. *See* EMITTER FOLLOWER.

Speed and power. Historically, the operational speeds achievable with MOS, TTL, and ECL logic have been well ordered, that is, ECL is the fastest, MOS the slowest, and TTL in between. Power levels vary inversely with gate delays, so that the compensating advantage of MOS is low power. CMOS in particular has the advantage that virtually no power is dissipated (typically microwatts per device) during periods when no switching is taking place, and power consumption in general is directly proportional to the frequency of operation.

With the advent of very large scale integration (VLSI) and corresponding fine-line lithography with feature sizes on digital chips of 90 nanometers or less, the speed categorization described above is less well defined. As a result of the smaller feature sizes, gate delays achievable in all three technologies typically are below 1 nanosecond. MOS technology is now the choice for even very high speed digital systems because its reduced power levels and simpler fabrication procedure enable placing more devices in the same area on individual chips.

R. R. Shively; W. V. Robinson; Philip V. Lopresti

Bibliography. R. J. Baker, *CMOS: Circuit Design, Layout, and Simulation*, 2d ed., 2004; R. H. Katz and G. Borriello, *Contemporary Logic Design*, 2d ed., 2005; J. M. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital Integrated Circuits*, 2d ed., 2003; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 5th ed., 2004; J. F. Wakerly, *Digital Design: Principles and Practices*, 4th ed., 2006.

Lophiiformes

The anglerfishes and their relatives. This highly modified order of antinopterygian fishes, also known as the Pediculati, is thought to have been derived from a stock close to the Batrachoidiformes, or toadfishes, at least as early as Eocene time. The first dorsal fin is reduced to a few flexible rays, of which the first is placed on top of the head and bears a terminal bulb or tassel and functions as a fishing lure (see **illus.**). The epiotics are large and meet behind the supraoccipital. There are no ribs or epipleurals, and the gill opening is small and placed far back. Pelvic fins, if present, are thoracic, and the pectoral girdle is notably modified. *See* BATRACHOIDIFORMES.

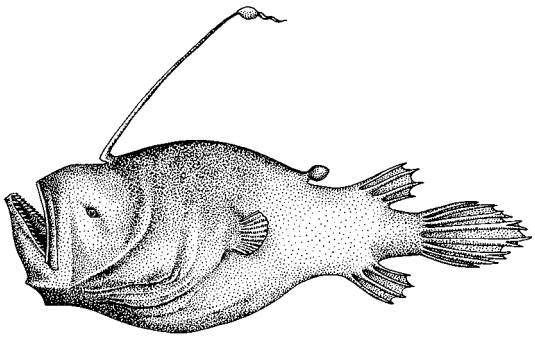

Anglerfish (*Cryptopsaras couesi*). (*After G. B. Goode and T. H. Bean, Oceanic Ichthyology, U.S. Nat. Mus. Spec. Bull. 2, 1985*)

There are 3 suborders, 15 families, nearly 60 genera, and about 195 known species. All are marine. Most species inhabit deep seas or tropical shore waters; the shore forms are commonly sedentary in habit. Luminescent organs may be present, and in some deep-sea species the males are dwarfed and attached as ectoparasites on the females. *See* ACTINOPTERYGII.

Reeve M. Bailey

Bibliography. E. Bertelsen, *The Ceratioid Fishes: Ontogeny, Taxonomy, Distribution and Biology,* Dana Rep. 39, 1951.

Lophophore

The crown of tentacles which surrounds the mouth in the Bryozoa, Phoronida, and Brachiopoda. Other groups of invertebrates also have tentacle crowns used in feeding, but the term lophophore is limited to the structure common to these three phyla in which the numerous ciliated tentacles arise from a circular or horseshoe-shaped fold of the body wall (see **illus.**). The tentacles are hollow outgrowths of the body wall, each containing fluid-filled extensions of the body cavity and extended hydraulically.

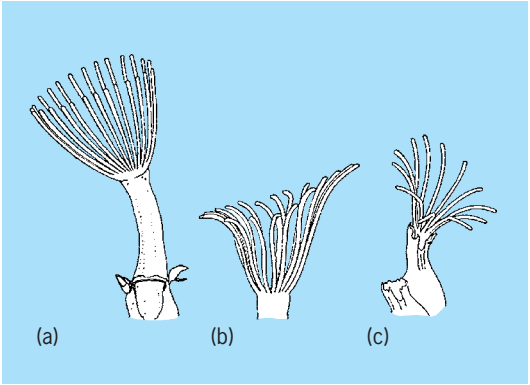

Examples of lophophores. (*a*) Equitentacled (*Beania intermedia*). (*b*) Scalloped (*Trematooecia turrita*). (*c*) Bent-tentacled (*Reteporellina evelinae*). (*After J. E. Winston, Polypide morphology and feeding behavior in marine ectoprocts, Bull. Mar. Sci., 28(1):1–31, 1978*)

The lophophore surrounds the mouth but not the anus of the animals. The primary function of the lophophore is to gather food. On the tentacles are ciliary tracts which drive a current of food-particle-bearing water through the lophophore. The action of the cilia, the flow of the current, and the movements of the flexible tentacles (which are provided with muscles under control of the nervous system) combine to collect food and deliver it to the mouth, where it is ingested. While the lophophore is primarily a feeding organ, it may also play a role in reproduction, respiration, and larval locomotion.

Bryozoa. Bryozoans are microscopic colonial animals. Marine bryozoans (gymnolaemates and stenolaemates) have the simplest lophophores: circular funnels which have the tentacles either equal in length or obliquely truncate, have their edges straight or curving, and are usually less than a millimeter in size. The lophophore of bryozoans can be protruded from the body and opened like an umbrella for feeding, or the tentacles may be furled and the whole lophophore drawn back into the zooid. In the gymnolaemates the lophophore is also important in reproduction. Sperm are liberated through pores in the tentacle tips. Some species have special male zooids with lophophores bearing only a few long tentacles. Species which do not brood their eggs have a vase-shaped structure between two tentacles. The eggs pass from the body cavity of the zooid into this intertentacular organ to be shed into the sea.

One fresh-water (phylactolaemate) bryozoan has a round lophophore, similar to that of marine bryozoans, but most have large horseshoe-shaped lophophores with a double series of tentacles arranged in a crescent and the mouth in the bend of the U. Their lophophores can also be retracted. *See* BRYOZOA.

Phoronida. Phoronids are larger tube-dwelling animals. The lophophore of phoronids is similar to that of the phylactolaemate bryozoans, with the tentacles arranged around the perimeter of the crescent. In some species the ends of the lophophore may be spirally coiled. The cilia of the tentacles create a current directed toward the groove between the two ridges of tentacles. This current moves toward the mouth and exits dorsally over the anus. Food particles carried by the current are entangled with mucus from the tentacles and carried by cilia on the basal groove to the mouth. The lophophores of phoronids cannot be introverted, but the animals can draw them back into their tubes. Phoronids can orient their expanded lophophores in various directions to take advantage of the water currents where they live. Phoronids also utilize the lophophore in reproduction. In *Phoronis viridis* sperm packages freed into seawater by a male dissolve a pore in a tentacle of a female, and the sperm travel down the tentacle and penetrate into the main body cavity to fertilize the eggs. The eggs of some species of phoronids are brooded between the two arms of the lophophore. *See* PHORONIDA.

Brachiopoda. Within their bivalved shells the brachiopods have the largest, most complicated lophophores, coiled and drawn out into two complex arms or lobes called brachia. Often this lophophore is provided with a rigid coiled support or brachidium. The tentacles of brachiopods are aligned in rows along the brachia. The lophophore divides the "mantle cavity" into incurrent and excurrent areas. The mouth is located medially in the incurrent area, in a groove which extends along both brachia below the tentacle-bearing ridges. Food is transported by cilia to the mouth. In the shelled planktonic larval stage of some inarticulate brachiopods, the tentacles of the lophophore may be used temporarily for swimming as well as for feeding. *See* BRACHIOPODA; FEEDING MECHANISMS (INVERTEBRATE).

Judith E. Winston

Loran

A navigation system from which hyperbolic lines of position are determined by measuring the difference in times of arrival of pulses from widely spaced, synchronized transmitting stations. Since radio waves travel at the speed of light, a line of position represents a constant range difference from two transmitters. Loran is used by commercial and military ships and aircraft. The name is derived from "long-range navigation." *See* HYPERBOLIC NAVIGATION SYSTEM.

In **Fig. 1**, it is assumed that the master station M transmits a pulse signal at time $t = 0$, and that the secondary station S transmits a similar pulse signal Δ microseconds after receiving the master pulse signal. The secondary station receives the master signal at time $t = \beta$ microseconds, and therefore transmits its pulse at time $t = \beta + \Delta$ microseconds. A receiver at R measuring the difference in the times of arrival of the signals from the secondary and master stations measures the time difference $\mathrm{TD} = (\beta + \Delta) + t_{SR} - t_{MR}$. The locus of all points with this common time difference is a hyperbola through the receiver

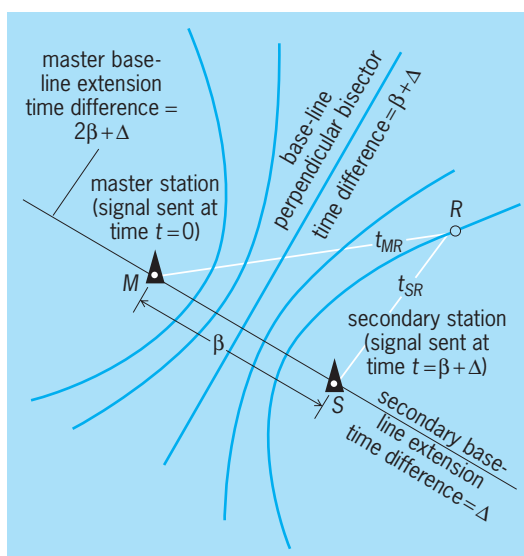

Fig. 1. Principle of position fixing by hyperbolic navigation systems. (*After J. P. Van Etten, Navigation systems: Fundamentals of low- and very-low-frequency hyperbolic techniques, Electr. Commun., 45(3):192–212, 1970*)

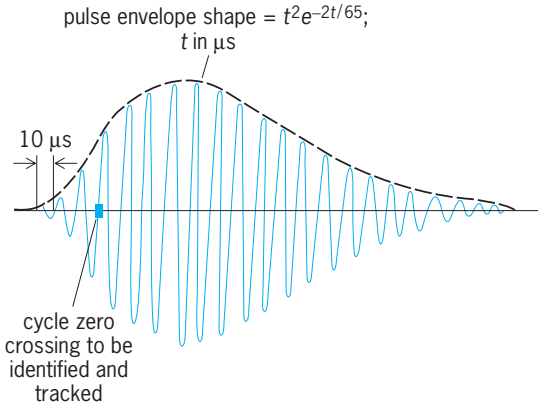

Fig. 2. Loran C pulse. (*After U.S. Coast Guard, Department of Transportation, Loran-C User Handbook, COMDTPUB P16562.6, 1992*)

position R. Similarly, every hyperbolic line of position is uniquely defined by a time difference. The intersection of two such lines of position gives a navigational fix, or location. Historically, maritime charts were printed with these lines of constant time difference and fixes were manually plotted. Since it is common for a Loran receiver to be able to receive the signals from four or more transmitters, there will exist more than two hyperbolic lines of position and they will intersect at many different locations. More recently, techniques have been developed to both automatically convert receiver observations into position and to utilize this redundant information to improve the accuracy and integrity of the fix obtained.

Development. The Loran position-fixing system was originally developed during World War II as an aid to the navigation of Allied aircraft and the North Atlantic convoys. Following the war, its use was extended by the U.S. Coast Guard to aid marine navigation. Operating at a frequency of about 2 MHz, this system was subsequently known as Loran A. The Loran B system, a high-accuracy cycle-matching version of Loran A, was not implemented because propagation disturbances rendered cycle selection unreliable. The Loran C system evolved from Loran A to provide greater range and more precise, repeatable, and accurate navigation. By the end of 1980, all Loran A service operated by the U.S. government had been terminated. Loran D, which evolved as a tactical (short-range transportable) version of Loran C that could be more quickly deployed, is no longer in use.

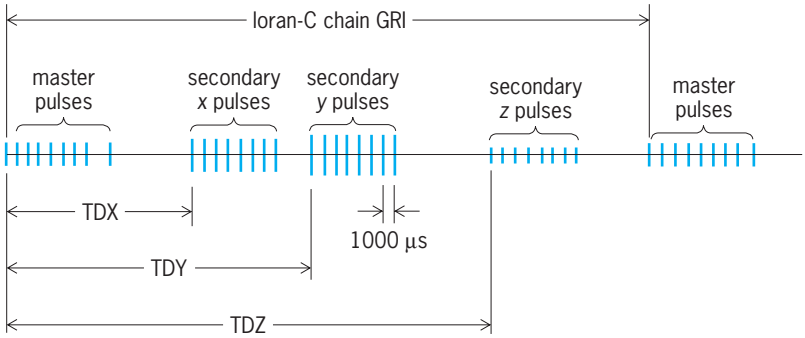

Fig. 3. Loran C signal format. (*After U.S. Coast Guard, Department of Transportation, Loran-C User Handbook, COMDTPUB P16562.6, 1992*)

Loran C. Loran C uses ground-wave transmissions at low frequencies to give an operating range in excess of 1000 nautical miles (1800 km). Sky-wave contamination is avoided by using pulse techniques. High accuracy is obtained by carrier-phase comparison, and cycle identification is accomplished by measurement of the pulse envelope. Operation is possible in poor signal-to-noise conditions by using correlation techniques and time-domain filtering. Loran C has been conservatively evaluated as providing an absolute accuracy of better than 0.25 nmi (500 m) at least 95% of the time with better than 165 ft (50 m) repeatable accuracy from day to day throughout 80% of the fishing grounds, where repeatable accuracy is desired for economic reasons. In the United States, differential Loran is being developed such that the 95% accuracy will be improved to approximately 65 ft (20 m) in areas that have been surveyed and are near a differential monitor site. In the United States the Loran C system is operated and maintained by the Coast Guard.

Transmission. Loran C chains are composed of a master transmitting station and two or more secondary transmitting stations. Master and secondary stations are generally separated by about 600 mi (970 km). Peak power of up to 2 MW is generated by the transmitter. Most stations use top-loaded monopole antennas up to 720 ft (220 m) in height. Six sites have multitower antenna arrays. *See* ANTENNA (ELECTROMAGNETISM).

All Loran C transmitters operate at a fixed frequency of 100 kHz and confine 99% of their radiated energy within the 90–110-kHz band. Each radiated pulse is designed, therefore, to build up quickly and decay slowly (**Fig. 2**). The radio frequency within each pulse is coherent with the repetition frequency.

Each Loran C transmitting station transmits a group of these pulses at a specified group repetition interval (GRI; **Fig. 3**). The master pulse group consists of eight pulses spaced 1000 microseconds apart, and a ninth pulse 2000 μs after the eighth. Secondary pulse groups contain eight pulses spaced 1000 μs apart.

Multiple pulses are used so that more signal energy is available at the receiver, improving significantly the signal-to-noise ratio without an increase in the peak transmitted power capability of the transmitters. The master station transmits a ninth pulse that can be used for visual identification. In the event that a baseline is unusable for navigation, the secondary transmitter of the affected baseline notifies the navigator by blinking the first two pulses off and on.

Each pulse within a group may have its radio-frequency cycles in phase or 180° out of phase with an established reference. The phase coding identifies master or secondary transmissions so that automatic signal acquisition can be accomplished unambiguously; coding is chosen so that the effects of long sky-wave pulse trains (over 1000 μs) cancel, thus permitting ground-wave accuracy to be retained under all conditions.

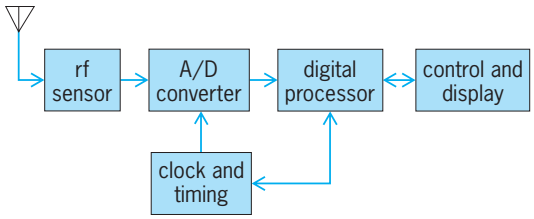

Fig. 4. Loran C navigation receiver.

Reception. High accuracy is obtained in the receiver by measuring the crossover time of individual radio-frequency (rf) cycles on the leading edge of the pulse envelope. Because of the longer propagation path of the sky wave, sky-wave contamination does not start until at least 30 μs after the beginning of the ground-wave pulse. Therefore, the first three cycles of the received signal are always stable ground waves. Tracking is accomplished on the third rf cycle of each pulse in the group to obtain the greatest measurement precision and stability. Cycle identification requires integration or filtering of measurements of the pulse envelope to permit a reliable decision as to whether the correct third cycle or an earlier or later cycle is being tracked. Once this decision is made, the correct cycle can be selected. Continuous monitoring of this cycle selection process is generally implemented, with an alarm if correction is required. *See* RADIO-WAVE PROPAGATION.

There are five major sections of a modern Loran C automatic acquisition and tracking, cycle-matching receiver (**Fig. 4**): the rf sensor, the analog-to-digital (A/D) converter, a Loran clock and timing circuits, a digital processor, and appropriate operator controls and displays. *See* ANALOG-TO-DIGITAL CONVERTER.

The rf sensor receives and processes the signal to provide adequate overall signal-to-interference protection, to "notch out" severe continuous-wave interferences, and, if necessary, to provide amplification to a suitable level for digital sampling. The analog-to-digital converter is used to convert the analog rf signal information from the rf sensor to digital samples associated with each Loran pulse or pulse group. The digital samples so obtained are then processed in the digital-processor. *See* SIGNAL PROCESSING.

Technology that emerged during 1972–1975 made it possible to incorporate tracking-loop algorithms, search algorithms, cycle identification algorithms, and so on, in digital software using inexpensive read-only memory for control of an inexpensive microprocessor. As a result, fully automatic microprocessor-based Loran C receivers were produced at a user cost comparable to that of the manual, and comparatively crude, predecessor Loran A equipment. *See* MICROPROCESSOR.

Operator controls and displays can take many forms, depending on the application. The navigation display, in its simplest form, is time-difference line-of-position numbers in microseconds for two signal pairs. More useful navigation outputs in terms of latitude-longitude coordinates, course-to-steer and distance-to-go to a selected destination, along-track and cross-track error, and so on, are generally provided in avionics equipment configurations and are generally offered as options in marine equipment configurations. Map displays with automatic Loran C position plot are also available.

Enhancements in the United States. Considerable effort and funding has been expended since the mid-1990s, to recapitalize the infrastructure and enhance the performance of Loran in the United States. All of the cesium atomic clocks have been replaced with newer versions. New timing and control equipment has been installed to enable precise synchronization to Universal Time Coordinated (UTC). Transmitter uninterruptible power supplies (UPS) have been installed to eliminate temporary loss of signal when commercial power is lost. New solid-state transmitters have replaced all tube transmitters in the contiguous 48 states and will replace those in Alaska as additional funding becomes available. *See* ATOMIC CLOCK; TIME; UNINTERRUPTIBLE POWER SUPPLY.

The Coast Guard has begun implementation of a Loran data channel to improve the accuracy, integrity, and availability for aviation, maritime, and timing users. The data channel consists of an additional ninth pulse in each group for data purposes only. This pulse is pulse position modulated in 32 possible states to transmit five bits per group. A message consists of 24 words: 15 words are Reed-Solomon parity and nine words or 45 bits are data. Several types of messages have been designed.

One type of message transmits absolute time, station identification, and leap-second information. Historically, since Loran signals are periodic and synchronized to UTC, they have been used as a source of frequency and of time modulo the period of the signal. They can now be used as a source of absolute time as well. Since the periodicity of all Loran signals in North America is in multiples of 200 μs, an inherent 200-μs lane ambiguity can hinder attempts to combine observations from more than one Loran chain. In addition, since all secondary signals are of the same format, if a receiver can successfully track a Loran secondary but not the corresponding master, there can be uncertainty as to which secondary is being tracked. Transmitting absolute time to resolve the interchain lane ambiguities, as well as station identification, enables master independent, multichain navigation with high integrity.

A second type of message transmits differential Loran corrections. Local monitors measure the times of arrival (TOAs) of the various Loran signals and offsets from nominal TOAs are broadcast. These corrections remove the effects of diurnal and seasonal variations in phase propagation velocity for maritime and timing users. For maritime users it is also necessary to conduct precise surveys of a harbor and its approaches to account for spatial variations in propagation as well. By removing errors due to both spatial and temporal variations in phase it is possible to obtain 2DRMS or 95% accuracies of 65 ft (20 m) or better.

Another message warns users of anomalous propagation caused by solar disturbances resulting in abnormally low ionospheric layers and early sky-wave

interference. These solar disturbances are usually referred to as polar-cap disturbances (PCDs), which affect only those paths in the auroral zone, and sudden ionospheric disturbances (SIDs), which affect paths at all latitudes. These disturbances will be detected by the monitor network. The warning message will transmit the geomagnetic latitude of the lower limit of the auroral zone. This will be set to 90° during times of no warning, to 0° during SIDs, and to the geomagnetic latitude of the lower limit of the auroral zone during PCDs. Receivers should not use signals when the midpoint of the path exceeds this geomagnetic latitude, the propagation distance exceeds 500 mi (800 km), and the midpoint of the propagation path is illuminated by the Sun. *See* GEOMAGNETIC VARIATIONS; IONOSPHERE; SOLAR WIND.

Other messages will be used to transmit system status, monitor locations, and other data channel information. In early 2006 the Loran data channel was being transmitted only on secondary signals. Whether to, and if so how to, implement it on master signals was under discussion.

EUROFIX. Beginning in the late 1980s, at the Delft University of Technology in the Netherlands, a group led by Durk van Willigen developed a Loran communications scheme called EUROFIX. In EUROFIX, the last six navigational pulses in a group are pulse-position modulated, resulting in one seven-bit word per group. A complete message contains 30 words: 20 words are Reed-Solomon parity and 10 words or 70 bits are data. The system was originally designed to transmit differential Global Positioning System (DGPS) corrections, and a message type has been added to transmit absolute time as well. As of January 2006, there were EUROFIX broadcasts of both DGPS and time from four European stations (Lessay, France; Sylt, Germany; and Vaerlandet and Boe, Norway) and DGPS-only broadcasts from three Saudi stations. *See* SATELLITE NAVIGATION SYSTEMS.

Status. The 2005 edition of the (U.S.) Federal Radionavigation Plan states "The Government continues to operate the Loran-C system in the short term while evaluating the long-term need for the system. If a decision is made to discontinue Loran as a result of these evaluations, then at least 6 months notice will be provided to the public prior to the termination of service." In 2002, a team was chartered by the Federal Aviation Administration's Loran evaluation program to determine if a modernized Loran system could meet aviation nonprecision approach (NPA), maritime harbor entrance and approach (HEA), and time and frequency dissemination requirements in the event of a GPS outage. This team's report,

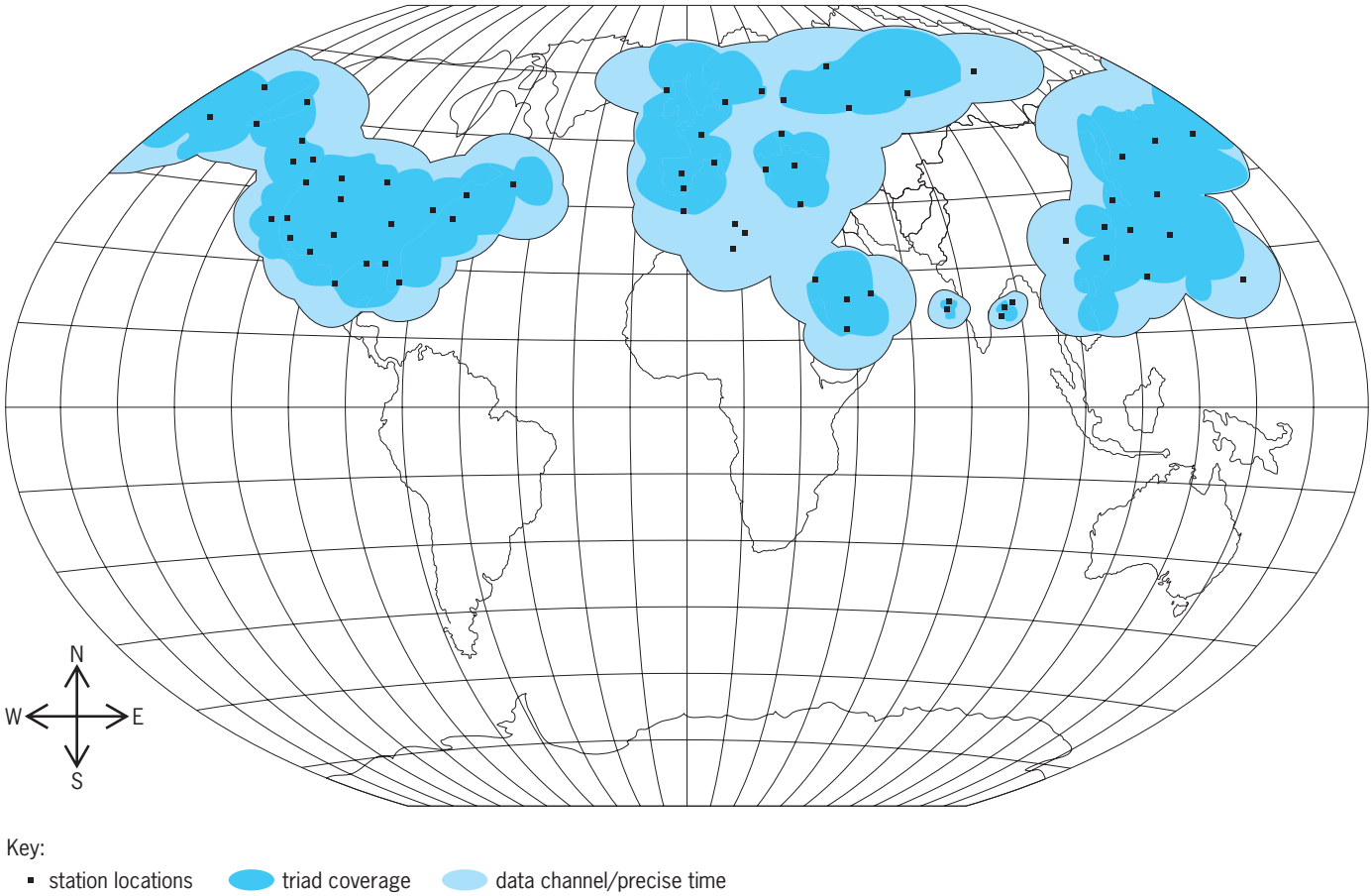

Fig. 5. Locations of Loran C transmitting stations and predicted coverage throughout the world as of 2006. (*Produced by and used with the permission of Megapulse, Inc.*)

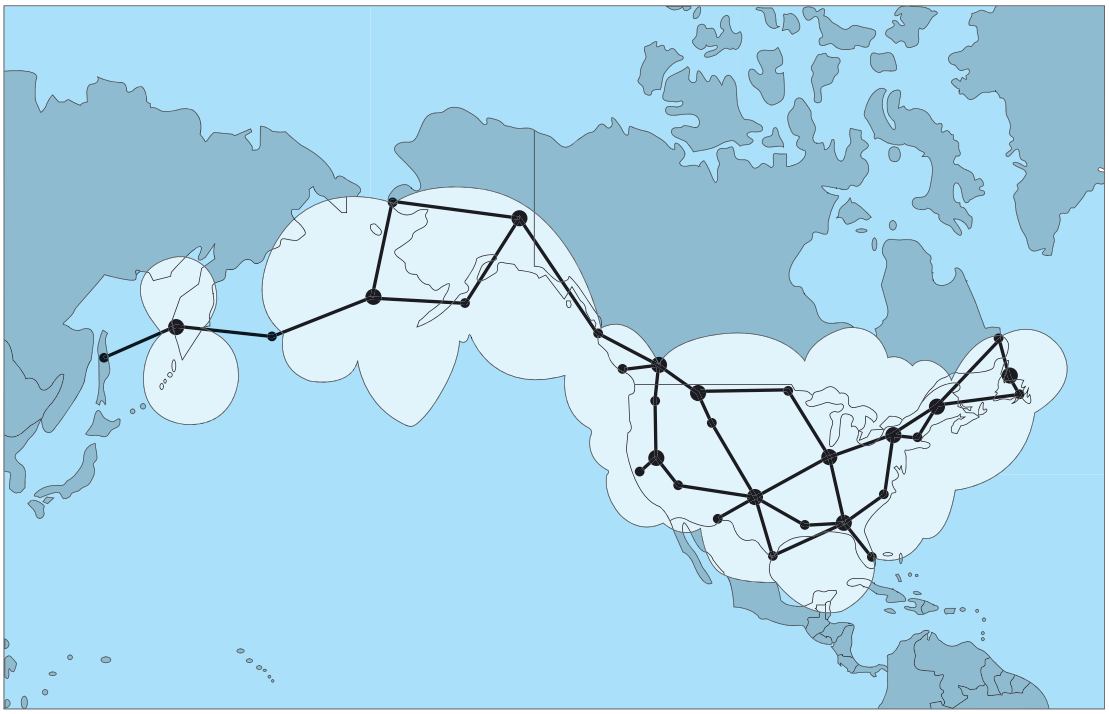

Fig. 6. Coverage provided by Loran C stations operated or supported by the United States as of 2006. Master stations are larger black dots. (*After 1999 Federal Radionavigation Plan, DOT-VNTSC-RSPA-98-1, U.S. Department of Transportation and U.S. Department of Defense, 1999*)

released in 2004, concluded that an enhanced Loran can meet all of these requirements. However, in early 2006 there was still no U.S. decision on the long-term future of Loran.

For a 10-year period ending January 1, 2006, the Northwest European Loran C System (NELS) was the operator of Loran in Western Europe. The NELS member nations included Germany, France, the Netherlands, Norway, Denmark, and Ireland. In 2005, the French Navy committed to operate the two French stations (Lessay and Soustons) for its own needs until 2015 and to enter into bilateral agreements with the goal of maintaining service of as many of the existing stations as possible. Also in 2005, the United Kingdom established a station at Rugby for testing purposes. However, Germany said it would terminate transmissions from Sylt, and Norway announced that it would demolish its stations, although both governments said that they would delay implementation of these plans. The Danish station at Ejde in the Faroe Islands probably will continue operations with financial support from the French government.

Coverage. **Figure 5** illustrates the worldwide distribution of Loran C transmitters as of January 2006. The inner area labeled "triad coverage" indicates where navigation fixes from Loran are available. In the outer area labeled "data channel/precise time," at least one signal can be received, and Loran is a source of frequency. It would potentially be the area of coverage for a data channel and absolute time if all signals were modulated, which at present they are not. **Figure 6** illustrates Loran C coverage provided or supported by the United States as of 2006. This includes Loran C sites operated in cooperation with Canada and Russia. Additional stations are operated by India, China, Japan, Korea, and Saudi Arabia. Russia with partner nations operates Chayka stations, which broadcast navigation signals that are very similar to Loran C. Transmitters for the Chayka "European" chain are located within Russia, Ukraine, and Belarus; transmitters for the "Eastern" chain are located within Russia and include two stations that operate cooperatively with the Loran C station operated by the United States at Attu, Alaska.

Applications. Although originally implemented as an electronic aid to marine navigation, the Loran C system can serve diverse airborne and land position location and navigation applications as well.

In the late twentieth century, technological advances rapidly lowered receiver costs, and coastal coverage limitations were eliminated or reduced by system improvements and expansion; consequently, there was a high degree of user acceptance by fishing crews and other commercial marine users, by the marine recreational community, and by civil aviation. In 1990, there were over 400,000 civil maritime users and approximately 75,000 civil aviation users of Loran C worldwide.

The determination of atmospheric winds using retransmission with navigational aids is accepted as economically and technically viable. For this application, meteorological probes, in the form of balloon-borne radiosondes, retransmit narrow-band Loran C signals to be processed at the tracking facility. Accurate velocity (not accurate position) is the sought-after parameter; hence narrow-banding of the Loran signals prior to retransmission is permissible; this has the advantage of greatly reduced

out-of-band interference and an improved signal-to-noise ratio. The improved Loran C worldwide coverage makes extensive use of this technique. *See* WIND MEASUREMENT.

There is also potential terrestrial use of Loran C for radiolocation, such as automatic vehicle monitoring in both rural and urban environments, and for precise determination of time and time intervals.

However, since the 1990s, actual use of Loran has diminished due to the better performance of GPS. The major argument for retaining Loran has become the vulnerability of GPS and other satellite-based navigation systems to jamming or interference, both intentional and unintentional. The major issue regarding Loran facing the United States and other governments are whether these vulnerabilities justify the costs of a terrestrial backup to satellite-based navigation systems; and, if so, if Loran is the appropriate backup. *See* AIR NAVIGATION; ELECTRONIC NAVIGATION SYSTEMS; MARINE NAVIGATION.

Benjamin B. Peterson; James P. Van Etten

Bibliography. *Loran-C and the Northwest European Loran-C System*, NELS Coordinating Agency, Oslo, 1998; *LORAN-C User Handbook*, COMDTPUB P16562.6, U.S. Coast Guard, Department of Transportation, 1992; L. Melton, *The Complete Loran-C Handbook*, International Marine Publishing Co., McGraw-Hill, 1987; *2005 Federal Radionavigation Plan*, DOT-VNTSC-RITA-05-12, U.S. Department of Transportation and U.S. Department of Defense, 2005; U.S. Federal Aviation Administration, *Loran's Capability to Mitigate the Impact of a GPS Outage on GPS Position, Navigation, and Time Applications*, March 2004.

Lorentz transformations

The relationship in the special theory of relativity between the sets of coordinates (t, x, y, z) and (t', x', y', z') used to label events in spacetime by two inertial observers, O and O', who are moving with respect to each other. Many of the effects predicted by special relativity can be derived in a direct manner from the Lorentz transformation formulas.

By definition, an event in spacetime is a point of space at an instant of time. It is an empirical fact that the collection of events in spacetime constitutes a four-dimensional continuum. This means that it takes four numbers to specify a particular event: one number to specify its "time" and three numbers to specify its spatial position.

Construction of coordinate systems. In both prerelativity physics and in special relativity, an inertial observer, O, is one who is not acted upon by any external forces and thus undergoes straight-line motion. It is assumed in prerelativity physics and in special relativity that any inertial observer, O, can use the following procedure to assign four numbers (t, x, y, z) to each event in spacetime: First, O builds a rigid cartesian grid of meter sticks, all of which intersect each other at right angles. Each point on this meter-stick grid may then be labeled by the three numbers (x, y, z) representing the distance of this point from O along each of the three orthogonal directions of the cartesian grid. The observer O may then carry the grid without rotating it, so that each point on the grid undergoes inertial motion. Next, O may arrange to have a clock placed at each point of this grid. Finally, O may synchronize all of the clocks placed on the grid. In prerelativity physics, this clock synchronization can, in principle, be achieved in a very straightforward manner because there is no limit to the speed at which a material body can move. Thus, at time $t = 0$ on O's clock, O could send off an assistant to visit and start each of the clocks on the grid and return to O before any appreciable time has elapsed on O's clock. *See* COORDINATE SYSTEMS.

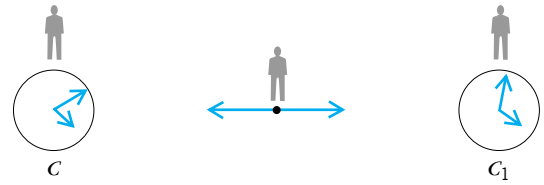

Fig. 1. Procedure that can be used in special relativity to synchronize clocks that are at rest with respect to each other.

However, in special relativity this straightforward (in principle) synchronization procedure cannot be used because the velocity of the assistant (or any other material body) is limited to be less than the velocity of light, c. Nevertheless, any clock, C_1, on the grid can be synchronized with O's clock C by stationing an assistant half-way between C_1 and C, and having the assistant send a signal (such as a light ray) simultaneously toward C_1 and C in a symmetrical manner (**Fig. 1**). Another assistant who is stationed at C_1 is instructed to record the reading of C_1 at the event at which the signal is received. Finally, yet another assistant may travel from O to C_1 (at a speed less than c) carrying the information of the reading on O's clock, C, when O received the signal. The assistant stationed at C_1 is then instructed to correct the "zero" of the clock reading there (by setting it forward or backward, but not changing its rate) by the amount of the difference of the readings of C_1 and C at the time the signal was received. By performing the same procedure with all the clocks on his grid, an observer O in special relativity can achieve the desired clock synchronization.

After performing this clock synchronization, an inertial observer in both prerelativity physics and in special relativity can label any event by the four numbers (t, x, y, z) as follows: The numbers (x, y, z) assigned to the event are the spatial coordinates of the meter-stick gridpoint at which the event occurred, while t is the time of the event as determined by the (synchronized) clock located at that gridpoint. This correspondence of the numbers (t, x, y, z) to events in spacetime is called the global inertial coordinate system associated with observer O.

Relationship of coordinate systems. Another inertial observer, O', who moves with respect to O at velocity v in the x direction, may also construct a global inertial coordinate system by using exactly the same procedure as O. The observer O' can thereby label events in spacetime by the numbers (t', x', y', z').

It is of interest to compare the way the same events are labeled by O and O'. For simplicity, it is assumed that O and O' meet at an event A, and that they adjust their clocks so that $t = t' = 0$ at this event. It is also assumed for simplicity that the axes of the grid of meter sticks carried by O' are aligned (that is, not rotated) with respect to the axes of O when they meet.

In prerelativistic physics, the relationship between the global inertial coordinate systems of O and O' is given by a galilean transformation defined by Eqs. (1).

$$t' = t \tag{1a}$$

$$x' = x - vt \tag{1b}$$

$$y' = y \tag{1c}$$

$$z' = z \tag{1d}$$

Equation (1a) states that O and O' agree on the time labeling of any two events. In particular, if two events are judged to be simultaneous by O (that is, if $t_1 = t_2$ for the two events), they also will be judged to be simultaneous by O' (that is, $t_1' = t_2'$). This reflects the fact that the notion of simultaneity is an absolute one in prerelativistic physics. The observers O and O' disagree only about the x-labeling of events, and this disagreement is readily understood as resulting from their relative motion.

In special relativity the corresponding relationship between the global inertial coordinates of O and O' is a Lorentz transformation, defined by Eqs. (2). The

$$t' = \frac{t - \dfrac{xv}{c^2}}{\sqrt{1 - \dfrac{v^2}{c^2}}} \tag{2a}$$

$$x' = \frac{x - vt}{\sqrt{1 - \dfrac{v^2}{c^2}}} \tag{2b}$$

$$y' = y \tag{2c}$$

$$z' = z \tag{2d}$$

Lorentz transformation differs dramatically from the galilean transformation in that O and O' disagree over the time labeling of events (**Fig. 2**). In particular, two events judged by O to be simultaneous (that is, $t_1 = t_2$) will not, in general, be simultaneous according to O'. This reflects the fact that simultaneity is not an absolute notion in special relativity.

General transformation. Equation (2) gives the Lorentz transformation for the simple case where O' moves in the x direction with respect to O, and O and O' have their axes aligned. The general Lorentz transformation allows for an arbitrary velocity, $\vec{v}$, of O and an arbitrary rotation of the axes of O and O'. For a general Lorentz transformation, Eq. (3) is satisfied. [It is readily verified that this relationship holds for the particular Lorentz transformation given in Eq. (2).] Indeed, the collection of all possible Lorentz transformations is composed precisely by the linear

$$x'^2 + y'^2 + z'^2 - c^2t'^2 = x^2 + y^2 + z^2 - c^2t^2 \tag{3}$$

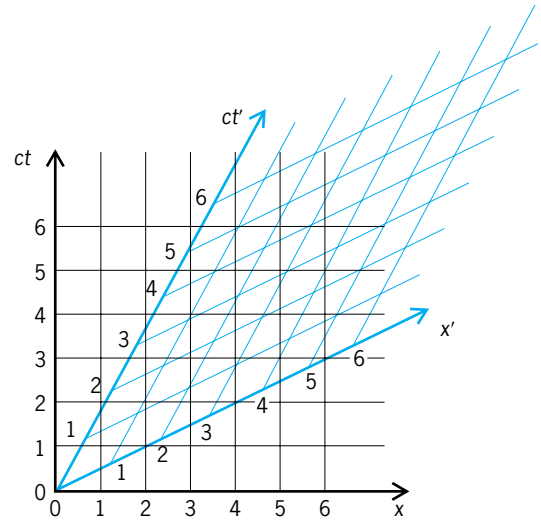

Fig. 2. Spacetime diagram illustrating the Lorentz transformation for the case $v = c/2$. The black lines indicate the (t, x) labels given to events in spacetime by inertial observer O, whereas the color lines indicate the corresponding (t', x') labeling given by inertial observer O'.

transformations of coordinates which satisfy Eq. (3). The mathematical role played by Lorentz transformations in special relativity thus is closely analogous to the mathematical role played by rotations in ordinary euclidean geometry in four dimensions, since a rotation may be defined to be a linear transformation of the coordinates (x_1, x_2, x_3, x_4) for which $x_1^2 + x_2'^2 + x_3'^2 + x_4'^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2$. *See* EUCLIDEAN GEOMETRY.

Transformation of physical quantities. All quantities of interest in special relativity which have an intrinsic, observer-independent meaning have well-defined transformation properties under Lorentz transformations. A good example is provided by the energy and momentum of a particle. In special relativity the energy, E, and 3-momentum, $\vec{p}$, of a particle are combined into a spacetime 4-vector, $(E/c, p_x, p_y, p_z)$. In order to get simple Lorentz transformation properties it is necessary to choose the "zero" of energy so that this vector is tangent to the worldline of the particle in spacetime. This implies that a particle of mass m which is at rest with respect to a particular global inertial coordinate system is assigned the energy $E = mc^2$ in that system. Under a Lorentz transformation of the form given by Eq. (2), these quantities transform in the same way as the spacetime position 4-vector (ct, x, y, z). Then Eqs. (4) are obtained. In particular, if the particle is

$$E' = \frac{E - p_x v}{\sqrt{1 - \dfrac{v^2}{c^2}}} \tag{4a}$$

$$p_x' = \frac{p_x - E\dfrac{v}{c^2}}{\sqrt{1 - \dfrac{v^2}{c^2}}} \tag{4b}$$

$$p_y' = p_y \tag{4c}$$

$$p_z' = p_z \tag{4d}$$

at rest with respect to O, the energy and magnitude of momentum of the particle as seen by O' are given by Eqs. (5).

$$E' = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad (5a)$$

$$|\vec{p}| = \frac{mv}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad (5b)$$

See RELATIVISTIC MECHANICS.

Another interesting example of the behavior of quantities under Lorentz transformations is provided by the electromagnetic field. In special relativity, the electric field vector $\vec{E}$ and magnetic field vector $\vec{B}$ are combined into a single spacetime tensor with simple properties under Lorentz transformation. Under the Lorentz transformation (2), the electric and magnetic fields at the event where O and O' meet transform as Eqs. (6). Thus, for example, if according to O only an

$$E'_x = E_x \qquad B'_x = B_x \qquad (6a)$$

$$E'_y = \frac{E_y - \frac{v}{c}B_z}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad B'_y = \frac{B_y + \frac{v}{c}E_z}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad (6b)$$

$$E'_z = \frac{E_z + \frac{v}{c}B_y}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad B'_z = \frac{B_z - \frac{v}{c}E_y}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad (6c)$$

electric field is present, then, in general, according to observer O' a magnetic field also will be present. However, it can be verified from Eq. (6) that the quantities $\vec{E} \cdot \vec{B}$ and $(|\vec{E}|^2 - |\vec{B}|^2)$ remain unchanged under Lorentz transformations, so all observers agree on the values of these quantities. *See* CALCULUS OF VECTORS; RELATIVISTIC ELECTRODYNAMICS; RELATIVITY.

Robert M. Wald

Bibliography. H. Bondi, *Relativity and Common Sense*, 1964, reprint 1980; R. M. Wald, *Space, Time, and Gravity*, 2d ed., 1992.

Loricifera

A phylum of multicellular invertebrates. These marine organisms are entirely meiobenthic; that is, they never exceed a maximum dimension of 400 micrometers and live in sediments ranging from deep-sea red clay to coarse sand or shell hash. They have some of the smallest known cells in the animal kingdom. Although they have been found throughout the world, only 10 species representing three genera and two families have been described: the first described species (*Nannaloricus mysticus*) from shallow coastal waters near Roscof, France; eight species (including the first loriciferan seen, *Pliciloricus enigmaticus*) off North Carolina, and a single species

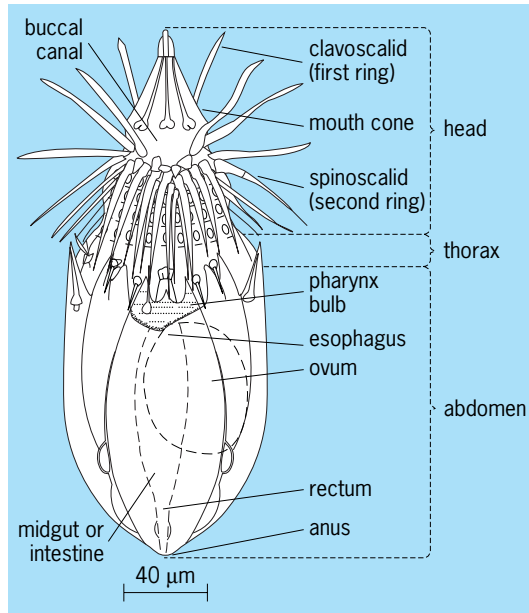

Ventral view of adult female loriciferan (*Nannaloricus mysticus*). (*After R. M. Kristensen, Loricifera, a new phylum with Aschelminthes characters from the meiobenthus, Z. Zool. Syst. Evolutionsforsch., 21(3):163–180, 1983*)

from the western Pacific. Well over 50 species thought to represent several additional genera are known but remain undescribed.

Adult loriciferans are bilaterally symmetrical. The body is regionated into a mouth cone which telescopically protrudes anteriorly from the center of a spherical, appendage-ringed (up to 400 rings may be present) head; a short, usually plated neck region; and a distinctively loricated (thick-cuticled) thorax-abdomen (see **illus.**). The head, along with its withdrawn mouth cone, may be inverted into the thorax-abdomen, which is then closed anteriorly by the cuticle of the neck region.

The first ring of appendages of the head comprises clavoscalids. They usually consist of eight club-shaped structures that extend anteriorly, similar to the arrangement of less prominent scalids in the Priapulida. Up to eight additional rings of posteriorly directed scalids follow in posterior progression. When the head is inverted into the lorica, the scalids fold together like the ribs of an umbrella, closing the wrong way in a manner similar to the scalids of the Kinorhyncha, a postulated close relative.

Internally, an eight-lobed brain surrounds the anterior part of the alimentary tract. Ten nerve cords extend posteriorly. At various places on the cuticle are special receptor organs called flosculi. The digestive system consists of a terminal mouth, followed by a cuticular buccal tube, muscular pharynx, and a short cuticle-lined esophagus. The esophagus opens into a large midgut, which is followed by a cuticle-lined hindgut opening dorso- or ventrocaudally. Sexes are separate in the Loricifera; clavoscalids of some may exhibit dimorphic characters. Females have dorsal ovaries and laterodorsal oviducts that open caudally. Males have two saccate testes, each with large seminal vesicles. The excretory system consists of a single

pair of protonephridia uniquely located within the gonads.

Two to five immature stages (called Higgins larvae) precede the adult. Higgins larvae have the same body regions as adults and range from 50 to 385 μm long. The structure of the mouth cone, the morphology and number of rings of head appendages, a thinner and slightly modified cuticular pattern of the remaining body, and presence of two posterior toes are characters that identify the immature stages. The toes may be used for locomotion or possibly for adhesion to the substrate.

Loriciferans appear to be closely related to the Priapulida and perhaps less so to the Kinorhyncha. Certain structures of the mouth cone suggest affinities to the Tardigrada. Their presence may support the idea that proarthropods could have been developed from aschelminth ancestral stock. *See* ANIMAL KINGDOM; ARTHROPODA; KINORHYNCHA; PRIAPULIDA; TARDIGRADA. Robert P. Higgins

Bibliography. R. P. Higgins and R. M. Kristensen, New Loricifera from southeastern United States coastal waters, *Smithson. Contrib. Zool.*, 438:1–70, 1986; R. P. Higgins and H. Thiel (eds.), *Introduction to the Study of Meiofauna*, 1988; E. E. Ruppert and R. D. Barnes, *Invertebrate Zoology*, 6th ed., 1994.

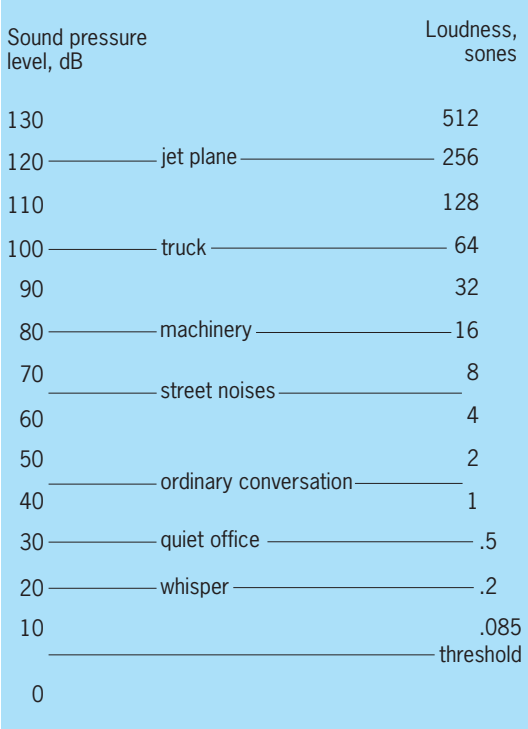

Fig. 1. Decibel and loudness scales of common sounds.

Loudness

The perceptual intensity of sound. Loudness depends importantly on the physical intensity of sound, increasing when physical intensity increases and decreasing when physical intensity decreases. But loudness also depends on other physical properties of sound, such as frequency and duration. Sound waves with frequencies between 1000 and 5000 Hz are louder than sound waves that have the same intensity but lower or higher frequencies. Very brief bursts of sound are less loud than are longer bursts, loudness increasing regularly as duration increases up to about 0.1–0.2 s, beyond which point in time, loudness no longer increases with increasing duration.

Decibels. In most work, the physical level of sound is expressed in decibels (dB). Decibels are units that denote relative intensity or pressure of a sound. One bel is defined as the logarithm to the base 10 of the ratio of a given sound intensity to a reference intensity. A decibel is one-tenth of a bel. Hence, a decibel level of 10 dB means an intensity 10 times that of the reference; a level of 20 dB means an intensity 100 times the reference.

Although decibels are defined primarily in terms of relative intensity, sound levels are usually measured in units of sound pressure. Under most circumstances, the intensity of a sound wave is proportional, not to its pressure, but to the square of its pressure. A sound whose pressure is 10 times that of the reference, therefore, has an intensity 100 times the reference, and decibel level of 20 dB. The most commonly used reference level for decibel scales is a sound pressure of 0.00002 newton/m^2. When expressed in terms of this reference a decibel level is called sound pressure level (SPL). Typical SPLs of some ordinary sounds are: a whisper, 20 dB; a quiet office, 30 dB; street noises, 60–70 dB; a truck, 100 dB; a jet plane, 120 dB (**Fig. 1**).

Phons. Because loudness depends on sound frequency as well as sound intensity or pressure, different stimuli with the same SPL may not be equally loud. One type of decibel scale, called the phon scale, overcomes this deficiency. The level of a sound in phons is the SPL in decibels of an equally loud 1000-Hz tone. Thus a 1000 Hz-tone at 40-dB SPL has a level of 40 phons, as do all other sounds that equal its loudness, even though these other sounds may have SPLs much greater than 40 dB. **Figure 2** shows the relationship between decibels SPL and phons in terms of several equal-loudness contours. The lowest curve is the audibility function, the absolute threshold for sound. Note that the equal-loudness curves become flatter and flatter at higher and higher levels of loudness. This flattening is most evident at low frequencies. In other words, the higher the intensity, the more closely phon levels correspond with SPL throughout the range of audible frequencies.

Loud sounds (70–100 phons) with the same frequency spectrum can be discriminated on the basis of their loudness if they differ by about 0.5 dB. Soft sounds (10–40 phons) must differ by 1–1.5 dB to be discriminated.

Sones. The phon scale is often designated as a scale of loudness level. It is a scale of equal loudness, in that all equally loud sounds take the same level in phons, regardless of the SPL. Nevertheless, the phon scale is not a true scale of loudness, because it is a physical (decibel) scale. That is, a sound of 80 phons is not necessarily twice as loud as a

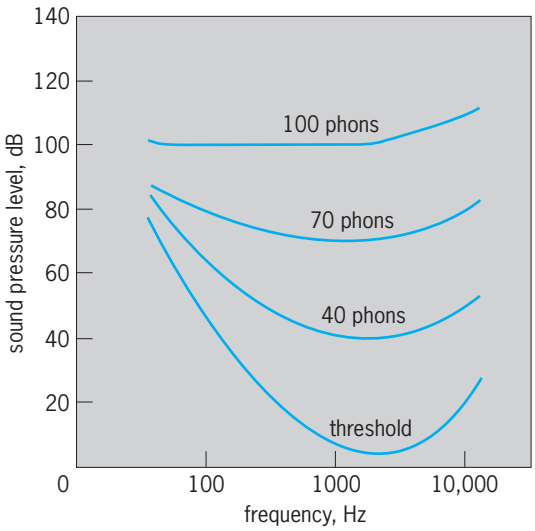

Fig. 2. Relationship between loudness level in phons and sound pressure level for audible frequencies.

sound of 40 phons. In order to determine a scale of perceived loudness, observers were asked to set the level of one sound to make it appear twice or half as loud as standard sounds. Observers were also asked to make numerical judgments of the degree of loudness. From such judgments a scale of loudness in sones was established. One sone is defined as the loudness of a 1000-Hz tone at 40-dB SPL heard with both ears. (When heard with one ear, loudness is half as great.) Above 40 dB, loudness in sones doubles with every 10-dB increase in the level of sound. Below 40 dB, loudness falls by half with decrements smaller than 10 dB (Fig. 1).

The loudness of complex sounds depends on how the sound energy is distributed across the frequency range. If components are sufficiently separated in frequency, the loudness of the total complex in sones equals the linear sum of the loudnesses of the components presented separately. If the components are not widely separated, the total loudness in sones is smaller than the sum of the component loudnesses.

Several methods exist to calculate the loudness in sones or the loudness level in phons of any complex sound. Sound-level meters essentially use the ear's equal-loudness curves (Fig. 2) to weigh the spectrum and produce an estimate of the overall loudness level in phons. Most often used is the A scale, the weighting that corresponds to the 40-phon curve. More accurate methods for computing loudness and loudness level require the individual evaluation of partial loudnesses in many frequency bands across the spectrum, together with complex rules for combining the partial loudnesses.

Noise. The problem of unwanted noise in the environment often reduces largely to a question of loudness. Judgments of noisiness, unpleasantness, and annoyance of sound correspond closely, in most instances, with judgments of perceived loudness, though other acoustical properties, such as impulsiveness, as well as nonacoustical factors, such as predictability, are also important. *See* ACOUSTIC NOISE; HEARING (HUMAN).

Lawrence E. Marks

Bibliography. K. K. Charan, *Fundamentals of Hearing: A Textbook on Acoustics and Psychoacoustics*, 1989; T. N. Decker (ed.), *Introduction to Audiology and Hearing Science*, 1990; G. M. Edelman et al., *Auditory Function: Neurobiological Bases of Hearing*, 1988; J. Katz (ed.), *Handbook of Clinical Audiology*, 4th ed., 1994; K. D. Kryter, *The Handbook of Hearing and the Effects of Noise*, 1994; J. O. Pickles, *An Introduction to the Physiology of Hearing*, 2d ed., 1988.

Loudspeaker

A device that converts an electrical signal from an amplifier into sound. A loudspeaker driver is an electromechanical-acoustic device with two electrical input terminals, to which an electrical signal is applied, and a diaphragm which vibrates and radiates sound. An electromechanical motor mechanism exerts a force on the diaphragm to cause it to vibrate. By far the most common type of motor mechanism, the electromagnetic-mechanical transducer, employs a coil of wire immersed in a magnetic field; electric current flowing through the coil causes a force to be exerted on the coil which is mechanically coupled to the diaphragm. A less common type of motor mechanism is the electrostatic-mechanical transducer. It uses a capacitor to which an electric voltage is applied, causing a force to be exerted between the capacitor plates, which is mechanically coupled to the diaphragm. Two types of electrostatic transducers are used: the piezoelectric transducer and the condenser transducer. The piezoelectric transducer uses a piezoelectric crystal between the capacitor plates. The condenser transducer uses an air dielectric. One plate of the capacitor is a flexible membrane which serves as the diaphragm. *See* AUDIO AMPLIFIER.

A loudspeaker system employs one or more loudspeaker drivers in a common enclosure. A one-way loudspeaker system employs a full-range driver to cover the full audio spectrum. A two-way system employs a low-frequency driver called a woofer and a high-frequency driver called a tweeter. An electrical low-pass filter is used in series with the woofer and a high-pass filter is used in series with the tweeter. These filters are commonly referred to as the crossover networks. A three-way system adds a mid-frequency driver called the midrange or squawker. The crossover network for this driver is a band-pass filter. In some systems, a driver called a supertweeter is used to reproduce audio frequencies into the ultrasonic range. A driver called a subwoofer is used to reproduce audio frequencies into the infrasonic range. *See* ELECTRIC FILTER; NETWORK THEORY.

Crossover networks can be either passive or active. Passive filters contain inductors, capacitors, and resistors, which are commonly mounted inside the loudspeaker enclosure. Active filters contain operational amplifier circuits, and must precede the power amplifiers that drive the loudspeakers.

A biamplified system uses two power amplifiers, one for the woofer and another for the midrange and tweeter. Passive networks are used between the midrange and tweeter. A triamplified system is one in which three power amplifiers are used, one for the woofer, one for the midrange, and one for the tweeter. *See* OPERATIONAL AMPLIFIER; POWER AMPLIFIER.

Loudspeaker drivers can also be operated either as direct-radiator or horn-loaded. A direct-radiator driver is one whose diaphragm radiates directly into the external air load. A horn-loaded driver has an acoustic horn between the diaphragm and the air load. A horn can be used to improve the efficiency and to control the directional pattern of the radiated sound at the expense of frequency response. *See* DIRECTIVITY; SOUND-REINFORCEMENT SYSTEM.

Dynamic drivers. A driver that uses an electromagnetic-mechanical transducer is called a dynamic or moving-coil driver. (**Fig. 1**). The moving surface which radiates sound is called the diaphragm. It is also called the cone because of its shape. Some drivers are designed with flat diaphragms. However, the cone shape exhibits better rigidity and is preferred.

The diaphragm is usually made of paper or paper-felt. Where deterioration due to weather or direct sunlight is a problem, a plastic is used. Paper-felt usually exhibits better damping of mechanical resonant modes that can cause peaks and dips in the response in the upper frequency range. The diaphragm is supported by an inner suspension, called the spider, and an outer suspension. These permit diaphragm motion in the axial direction when a signal is applied, while providing the restoring force to keep it centered in the absence of a signal. *See* DAMPING.

There are two basic types of outer suspensions, rolled and accordion. A rolled suspension has an arc-shaped cross section and is very compliant, allowing the diaphragm to move with relatively little restoring force. The suspension is usually made of thin foam rubber or butyl rubber. Drivers with rolled suspensions are better suited for closed-box systems where the spring constant of the air inside the enclosure dominates over the spring constant of the driver suspension. An accordion suspension has a cross section that resembles the folds of an accordion. It is stiffer than the rolled suspension and is used primarily for drivers designed for open-back and vented-box systems. The suspension is made of paper-felt or lacquered cloth.

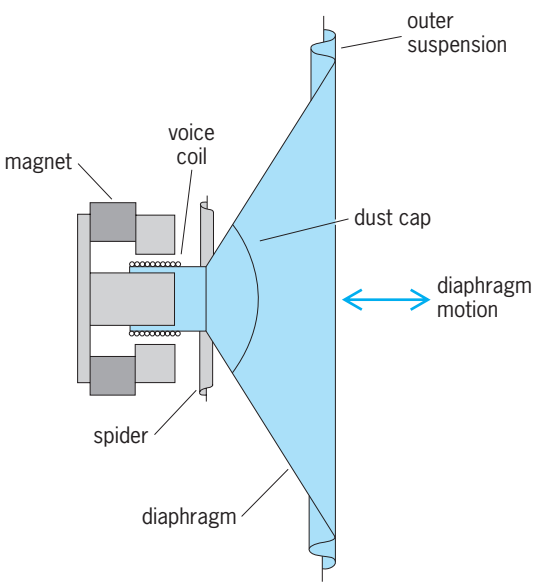

Fig. 1. Cross section of a typical dynamic loudspeaker driver.

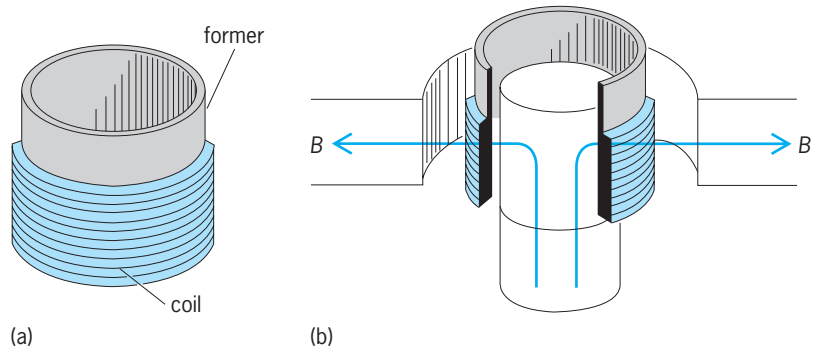

Fig. 2. Voice coil. (*a*) Coil wound on a former. (*b*) Magnetic flux, *B*, through an air gap in the magnet assembly where the voice coil is located. The force generated is in the axial direction when a current flows in the voice coil.

The electrical input signal is applied to the voice coil, a coil of wire wound on a cylindrical form, called the voice-coil former (**Fig. 2***a*), which is glued to the diaphragm. The voice-coil former is usually made of paper in low-power drivers. In high-power drivers, it may be made of aluminum or some other heat resistant material. A dome-shaped cover prevents dust from entering the driver through the voice coil.

When current flows through the voice coil, a force is generated, which causes the diaphragm to move and radiate sound. The interaction of the current and the magnetic flux through the voice coil produces the force. For maximum force, the magnetic field must be perpendicular to the direction of the current flow. The field is set up by a permanent magnet that is part of a magnet assembly which focuses the field through the voice coil (Fig. 2*b*). *See* MAGNETISM.

The magnet assembly consists of a permanent magnet and several pole pieces. The magnet is normally made of either an alnico (aluminum-nickel-cobalt) alloy or ceramic. The pole pieces, made of a high-permeability metal, form a low-reluctance path to focus the magnetic field through the air gap (**Fig. 3**). Ceramic magnets are normally donut-shaped (Fig. 3), while alnico magnets have a slug shape and occupy the position of the center pole piece. *See* MAGNETIC MATERIALS; RELUCTANCE.

Figure 4 illustrates how the force is generated through the voice coil. The direction of the force is that of the vector cross product of the direction of current flow and the magnetic flux B ($\vec{n} \times \vec{B}$, where $\vec{n}$ is in the direction of positive current flow). Because an audio signal is an ac signal, the force reverses direction each time the current reverses, so that the

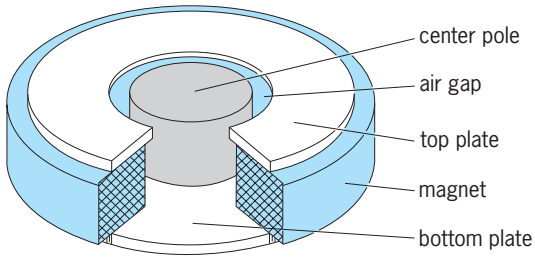

Fig. 3. Typical magnet assembly. The cylindrical, donut-shaped magnet is sandwiched between top and bottom pole plates. The air gap is the space between the center pole and the top pole plate.

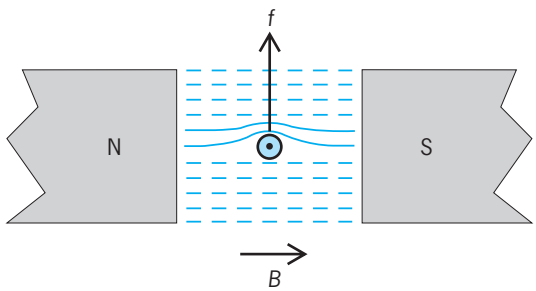

Fig. 4. Cross section of one turn of voice-coil wire with the current flowing, for illustration, in the direction out of the page. $\vec{B}$ is the direction of the magnetic flux. For current flowing out of the page, the force *f* generated on the voice coil is upward.

force causes the diaphragm to move in and out with the signal.

The force on the diaphragm is linearly related to the current in the voice coil only if the total number of turns of wire in the magnetic field does not change when the diaphragm moves. There are two methods by which this may be achieved. The long voice coil (**Fig. 5***a*) is wound so that there is an overhang on each side of the air gap. When the voice coil moves, turns of wire leave the gap on one side while simultaneously the same number of turns enter the gap on the other side so that the total number of turns in the gap remains constant. The long voice coil is preferred in woofers and higher-power drivers because a greater diaphragm excursion can be obtained and its heat dissipation is better than for the short voice coil.

In the short voice coil (Fig. 5*b*), all turns of wire are located in the central portion of the air gap, so that no turns leave the gap when the diaphragm moves. Because all the turns are in the magnetic field, the efficiency of the driver is improved. However, heat dissipation is poor because of overlapping turns. Thus, the short voice coil is usually not used in high-power drivers.

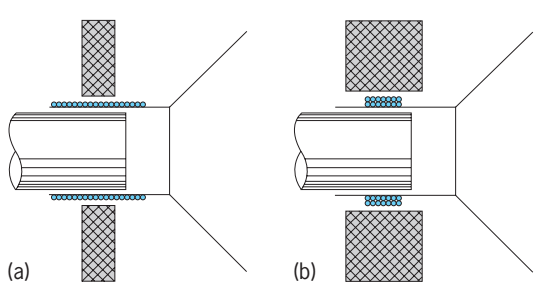

Fig. 5. Voice-coil configurations. (*a*) Long voice coil. (*b*) Short voice coil.

There is a maximum distance that the voice coil can move for either geometry before the number of turns in the magnetic field no longer remains constant. When this occurs, the driver becomes nonlinear and produces a distorted output. The maximum distance that the diaphragm can move from its rest position (zero), either inward or outward, before it becomes nonlinear is usually denoted by the symbol x_{max}. *See* DISTORTION (ELECTRONIC CIRCUITS).

Another nonlinear effect is caused by the driver's inner (spider) and outer suspensions (Fig. 1). The spring constant of these materials is constant only for small diaphragm displacements. As the displacement increases, the suspension become nonlinear. When the suspension bottoms out, the diaphragm cannot move further. In a well-designed driver, the voice coil displacement limit x_{max} is reached before the suspension bottoms out.

In modeling the acoustic radiation from a driver, it is common to assume that the diaphragm is a flat piston. The piston radius of the driver is measured by taking one-half the diameter of the diaphragm aperture, measured half-way into the outer suspension on each side. As a rule, a driver's piston radius in centimeters is numerically approximately equal to its advertised diameter in inches.

The flat piston approximation fails at high frequencies. The frequency at which this occurs is a function of the piston radius, the depth of the diaphragm, and the diaphragm material. Although it is impossible to specify an exact relation for the upper-frequency limit, it is normal to take it to be the frequency at which the circumference of the equivalent flat piston is somewhere between $\lambda/2$ and λ, where λ is the wavelength. The depth of the diaphragm should not exceed about $\lambda/10$ at this frequency. Unless specified otherwise, the upper frequency limit is usually taken to be the frequency at which the circumference is λ. Above this frequency, the radiation from the diaphragm can become directional and exhibit undesirable peaks and dips as the frequency is increased.

Small-signal parameters. The basic parameters commonly specified for dynamic drivers (see **table**) are called small-signal parameters because they assume the diaphragm is not driven into nonlinear operation. Knowledge of these parameters is necessary to design system enclosures, particularly woofer enclosures because the lower cutoff frequency of a system is determined by the woofer. A typical value for the voice-coil resistance is in the range from 4 to 8 ohms. This resistance value is commonly used in specifying the rated impedance of a system. The other parameters determine the optimum enclosure volume for the driver and its lower cutoff frequency. For woofer drivers, the resonance frequency is typically in the range of 20 to 40 Hz. For woofer drivers made for musical instruments, it is typically 40 to 80 Hz. For tweeter drivers, it is typically in the range

Loudspeaker parameters

Parameter	Symbol
Voice-coil resistance	R_E
Resonance frequency	f_s
Electrical quality factor	Q_{ES}
Mechanical quality factor	Q_{MS}
Total quality factor	Q_{TS}
Volume compliance	V_{AS}

from 1 to 2 kHz. Midrange drivers have intermediate values. *See* MUSICAL INSTRUMENTS.

The total quality factor Q_{TS} is related to the mechanical and electrical quality factors by Eq. (1). Q_{TS}

$$Q_{TS} = \frac{Q_{MS}Q_{ES}}{Q_{TS} + Q_{ES}} \tag{1}$$

has a typical value of 0.2 to 0.5 for woofers intended for use in enclosures and a value of 0.7 to 1.0 for woofers intended for use in walls or open-back enclosures. The electrical quality factor is typically 5 to 20% greater than the total quality factor. The mechanical quality factor is typically 5 to 20 times the electrical quality factor. If a resistor R is inserted in series with the driver voice coil, the electrical quality factor is increased to an effective value Q'_{ES} given by $Q'_{ES} = Q_{ES}(1 + R/R_E)$, where R is the value of the added resistor. For example, R might be the dc resistance of any passive crossover network elements in series with the driver. *See* Q (ELECTRICITY).

Driver sensitivity. The sensitivity of a driver is usually defined as the on-axis sound pressure level (SPL) 1 m (3.3 ft) in front of the driver with an electrical input power of 1 W. The frequency at which the sensitivity is measured should be stated as part of the driver's specification. The voice-coil voltage is usually calculated assuming the voice coil is a resistor having a value equal to the dc resistance of the driver. For 1 W of input power, the root-mean-square (rms) voice coil voltage has the value $\sqrt{R_E}$ V. For example, a driver with $R_E = 8$ ohms would require 2.83 V rms. Typical sensitivity values are in the range from 85 to 95 dB. When comparing drivers with different voice-coil resistances, the sensitivities must be referenced to the same voice-coil voltage. For example, the sensitivity for $\sqrt{R_E}$ V rms across the voice coil can be converted to a sensitivity for 1 V rms by subtracting $10 \log_{10}(R_E)$ from the 1-W specification. *See* SOUND PRESSURE.

Closed-box systems. A closed-box system is one in which the woofer is mounted in a sealed box (**Fig. 6***a*). When the diaphragm vibrates, it compresses and rarefies (expands) the air in the box, causing an increase in the effective spring constant of the driver suspension. This causes an increase in the driver resonance frequency and quality factors. If the spring constant of the air in the enclosure dominates over the spring constant of the driver suspension, the system is called an acoustic-suspension or air-suspension system. If the spring constant of the diaphragm suspension dominates over that of the air in the enclosure, the system is called an infinite-baffle system. Closed-box systems are usually loosely filled with an absorbing material such as fiberglass. This has the beneficial effect of damping standing waves inside the box. In addition, it increases the effective internal volume of the box. This occurs because the filling acts as a heat reservoir to maintain the air at constant temperature in the box as it is compressed and rarefied, thus decreasing the change in pressure as the diaphragm vibrates.

Vented-box systems. A vented-box or ported system is one in which the woofer is mounted in a box with a vent connecting the air inside the box to the outside. The vent can be a rectangular or cylindrical tube installed in the front panel of the enclosure (Fig. 6*b*). Sometimes it is formed by a shelf across the width of the box below the driver so that the volume of air in the vent is the air between the shelf and the bottom of the box. Vented-box systems are

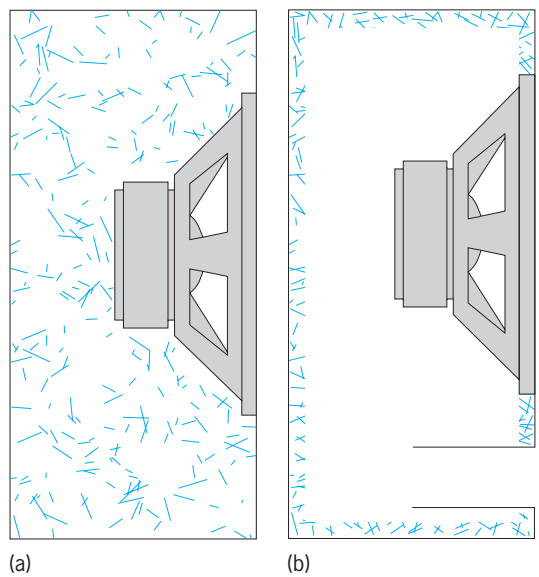

Fig. 6. Driver mountings. (*a*) Mounting in a filled closed box. (*b*) Mounting in a lined vented box in which the vent is a tube installed in the front panel of the enclosure.

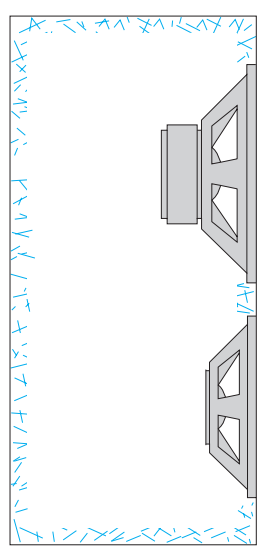

Fig. 7. Passive radiator (passive lower driver) system.

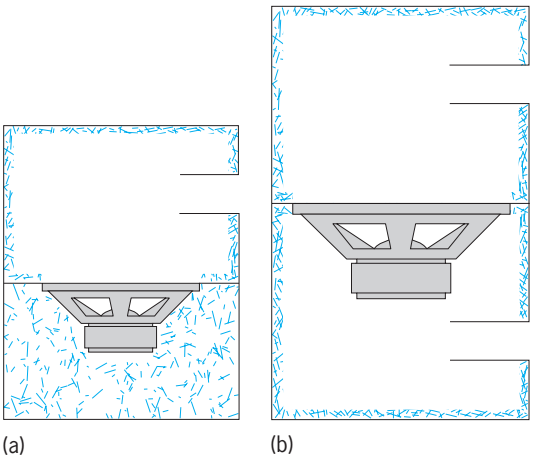

Fig. 8. Band-pass systems. (*a*) Fourth-order system. (*b*) Sixth-order asymmetrical system.

usually lined with a 1-in. (2.5-cm) thickness of filling material such as fiberglass. This helps absorb standing waves inside the box but has little effect on the effective volume of air in the box.

The air in the vent and the air trapped in the box form an acoustical resonator called a Helmholtz resonator that is used to extend the response of the system to lower frequencies. At the Helmholtz resonance frequency, the mass of the air in the vent resonates with the spring constant of the air in the box to cause strong radiation from the vent. At the same time, the pressure of the air in the box tends to oppose motion of the driver diaphragm so that little radiation comes from the driver. If the system is driven hard at the Helmholtz frequency, the airflow in the vent can become so great that noise can be generated by turbulent airflow. To minimize this noise, the vent area should be as large as possible. For a given Helmholtz frequency, the vent length must be made longer if its area is increased. However, increasing the vent's area may require a vent longer than the depth of the box, which can be accomplished using a vent tube with a 90° angle inside the box. *See* ACOUSTIC RESONATOR.

An aperiodic vented-box system has some type of acoustical damping material, such as fiberglass, placed inside the vent. This tends to damp the radiation from the vent at the Helmholtz frequency, thus decreasing the output from the system at low frequencies. A properly designed vented-box system does not require such damping. It is useful only if the box is too small, thus causing an undesirable peak in the frequency response at the Helmholtz frequency.

Passive-radiator systems. A passive-radiator system replaces the vent in a vented-box system with a passive radiator, that is, a driver having no voice-coil and magnet assemblies (**Fig. 7**). The advantage is that the Helmholtz frequency of the system can be controlled by varying the mass of the passive radiator diaphragm instead of the length of a vent tube.

Like the vented-box system, the passive-radiator system exhibits a Helmholtz resonance frequency at which the mass of the passive radiator diaphragm resonates with the spring constant of the air in the enclosure. This resonance can be used to improve the low-frequency response of the system. However, there is a second resonance frequency at which the radiation from the passive radiator tends to cancel that from the driver. At this frequency, the system exhibits a null in its frequency response. In a properly designed system, this resonance should be well below the Helmholtz resonance so that is has a minimal effect on the lower half-power or −3-dB cutoff frequency of the system. *See* RESONANCE (ACOUSTICS AND MECHANICS).

Band-pass systems. A band-pass enclosure is often used for subwoofers. In such a system, the loudspeaker driver is mounted on a shelf inside the enclosure, which divides the enclosure into two parts. For the system of **Fig. 8***a*, one of the two enclosures is coupled to the outside of the box through a vent. The frequency response of this system is that of a fourth-order band-pass filter which rolls off at a rate of 12 dB per octave or 40 dB per decade below its lower cutoff frequency and above its upper cutoff frequency. In the system of Fig. 8*b*, the two enclosures are coupled to the outside of the box through separate vents. The frequency response of this system is that of a sixth-order asymmetrical band-pass filter which rolls off at a rate of 24 dB per octave or 80 dB per decade below its lower cutoff frequency and a rate of 12 dB per octave or 40 dB per decade above its upper cutoff frequency. **Figure 9** compares the responses of the two systems, and illustrates why the sixth-order system is said to have an asymmetrical response. Both of the curves exhibit a maximally flat response in the midband range that is typical of a Butterworth response. (A Chebyschev response would exhibit peaks at the edges of the midband range.)

Horn-loaded drivers. A driver may be connected to an acoustic horn (**Fig. 10**). The sound radiated by the driver enters the throat of the horn and propagates to the mouth, where it is radiated. The acoustic impedance seen by the diaphragm of the driver is much higher with the horn than for a direct-radiator driver. This improves the efficiency so that less power is required to drive the system. In addition, the radiation from the mouth is more directional than for a direct-radiator driver. This is advantageous in public address systems in large rooms, where it is desirable to direct the sound to

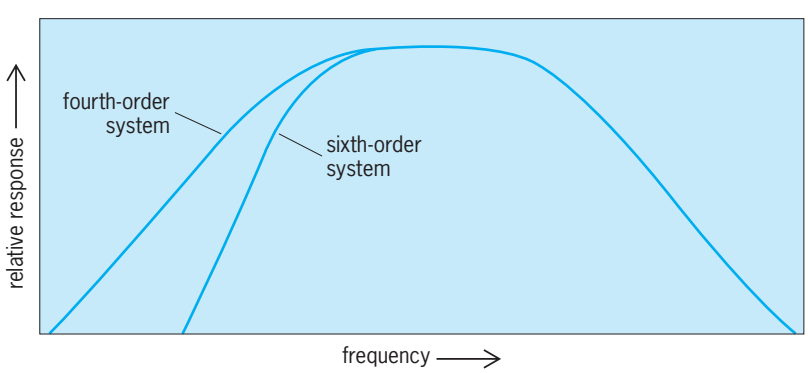

Fig. 9. Comparison of the low-frequency responses of fourth-order and sixth-order band-pass systems.

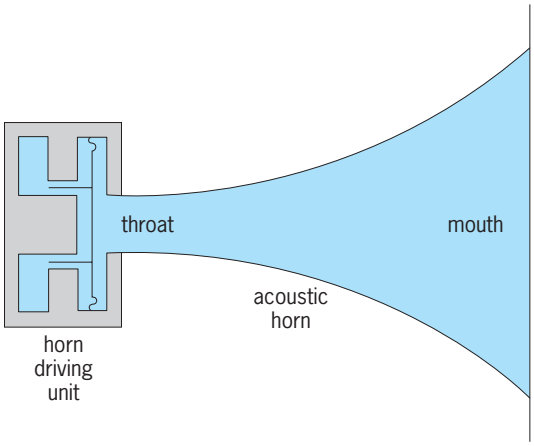

Fig. 10. Horn-loaded driver.

listeners at the back of the room without reflections from the side walls. Horn-loaded systems become physically very large when designed for low-frequency applications. For this reason, horns are normally used with midrange and tweeter drivers. *See* ACOUSTIC IMPEDANCE. W. Marshall Leach, Jr.

Bibliography. F. V. Hunt, *Electroacoustics*, American Institute of Physics for the Acoustical Society of America, 1982; W. M. Leach, Jr., *Introduction to Electroacoustics and Audio Amplifier Design*, 2d ed., Kendall/Hunt, 1999; *Loudspeakers*, vols. 1–4, Audio Engineering Society Anthology Series; L. L. Beranek, *Acoustics*, rev. ed., Acoustical Society of America, 1986; H. F. Olson, *Acoustical Engineering*, Van Nostrand, New York, 1957.

Louping ill

A viral disease of sheep, capable of producing central nervous system manifestations. It occurs chiefly in the British Isles. Infections have been reported, although rarely, among persons working with sheep.

In sheep the disease is usually biphasic with a systemic influenzalike phase, which is followed by encephalitic signs. In infected humans the first phase is generally the extent of the illness.

The virus is a member of the Russian tick-borne complex of the group B arboviruses. Characteristics, diagnosis, and epidemiology are similar to those of other viruses of this complex. *See* ANIMAL VIRUS; ARBOVIRAL ENCEPHALITIDES. Joseph L. Melnick

Low-level counting

The measurement of very small amounts of radioactivity. This can be achieved by measuring radioactivity in large samples or modifying conventional radiation-detection instruments for greater detection sensitivity, or both. All approaches must take into account the characteristics of the radiations, notably the very short range and intense energy deposition of alpha particles, intermediate range and energy deposition of beta particles, and high penetration of matter with low linear energy transfer by gamma rays. These techniques are important for measuring radionuclides from naturally occurring terrestrial and cosmic-ray-produced sources, and anthropogenic radionuclides to characterize the decay schemes of newly formed isotopes and analyze environmental and biological samples. *See* ALPHA PARTICLES; BETA PARTICLES; GAMMA RAYS.

Statistical error. Accurate measurements are defined by a small statistical error E, in percent, for the net sample count rate R, in inverse seconds (s^{-1}). Because the statistical error depends both on R and on the radiation background count rate R_b, in s^{-1}, both measured for the counting period t, in seconds (s), the value of R_b must be kept as low as possible. When R_b is much greater than R, E can be estimated from Eq. (1). For example, a sample counted for 5000 s

$$E = 100(2R_b t^{-1} R^{-2})^{0.5} \tag{1}$$

at a net count rate of 0.001 s^{-1} in a low-background beta-particle detector with a background count rate of 0.02 s^{-1} has an unacceptably high error of 283%. To reduce this error tenfold, Eq. (1) indicates that the background count rate has to be reduced by a factor of 100, or the time increased by a factor of 100, or the sample net count rate increased by a factor of 10. If R and R_b are of the same magnitude, the statistical uncertainty of both the gross count rate and the background must be considered. *See* BACKGROUND COUNT; STATISTICS.

Concentration. The concentration C, in units of s^{-1} kg^{-1}, of a sample is related to R by Eq. (2). Here e is

$$C = R(efdm)^{-1} \tag{2}$$

the overall fractional counting efficiency of the detector for the sample, f is the fraction of the counted radiation type relative to the emitted radiation, d is the fraction of the radionuclide remaining after radioactive decay during the interval between sample collection and counting, and m is the sample mass (in kg). For measuring very low concentrations, the values of e, f, and d should be as near unity as possible, and m should be large.

Radiation detectors. These concepts apply to a variety of radiation detectors for counting solid, liquid, and gaseous samples. Commonly used radiation detectors are gas-ionization systems operating in the ionization, proportional, or Geiger-Müller regions; scintillation systems with scintillators that are inorganic solids, organic solids, or organic liquids; and solid-state systems with germanium or silicon detectors. If possible, the various factors cited above are combined to improve the detection sensitivity, although some circumstances require a compromise between factors—for example, balancing the advantages of large sample mass and a counting efficiency near unity—if the two cannot be achieved simultaneously. *See* GEIGER-MÜLLER COUNTER; IONIZATION CHAMBER; JUNCTION DETECTOR; PARTICLE DETECTOR; SCINTILLATION COUNTER.

Sample preparation. Radionuclides in very large samples may be measured directly by counting the

energetic gamma rays they emit, for example in whole-body counting. If such samples emit only alpha particles, beta particles, or weak gamma rays (x-rays), then the radionuclides must be extracted and concentrated by chemical or physical processes. Radionuclides can be concentrated from water on ion-exchange resins or by precipitation, and from air by sorption on materials such as charcoal or molecular sieves. Filtering will collect radioactive solids from both media. Other processes include condensation of gases from air and volatilization from liquids or solids. The radionuclides must then be prepared for counting as thin solids, as solutes in a liquid scintillation cocktail, or as gases that are part of an ionization-detector filling. *See* ADSORPTION; CHARCOAL; FILTRATION; ION EXCHANGE; MOLECULAR SIEVE; PRECIPITATION (CHEMISTRY); VOLATILIZATION.

Detection efficiency. To maximize the detection efficiency, the intrinsic efficiency for recording the radiation should be near unity and the detector should view the sample from as large a solid angle as possible. Even relatively thin detectors have intrinsic efficiencies near unity for alpha particles, beta particles, and x-rays, but energetic gamma rays are detected with much lower intrinsic efficiency, except by very thick detectors. In general, the larger the sample, the smaller the angle subtended by the detector, although large detectors have been made for large samples. Relatively small samples can be surrounded more easily by a detector.

Radiation background. Achieving a low and stable radiation background is one basic aim in low-level counting. Stability is necessary because the detector background generally is determined by counting blanks between samples and subtracting these blank counts from the sample count on the assumption that the background is invariant. The counting facility must be situated to avoid fluctuating external radiation fields at nuclear facilities or near radiation sources. Airborne radon and progeny may have to be controlled by air-cleaning processes to avoid radiation background fluctuations. *See* RADON.

Shielding. The detector background from ambient radiation is reduced by shielding and by anticoincidence systems. A typical shield consists of 2 in. (5 cm) of a heavy metal such as lead or mercury, or thicker iron, to surround a beta-particle detector, or at least twice that thickness for a gamma-ray detector. This shielding absorbs all beta particles and most gamma rays. Water or paraffin is added to shield against neutrons (**Fig. 1**). For ultralow-level counting, the entire system can be shielded with massive amounts of iron and buried below ground to reduce the cosmic-ray background. *See* NEUTRON; RADIATION SHIELDING.

Radioactive contamination. The shielding materials and the detector plus ancillary equipment such as supports and cryostat within the shield must be as free as possible of radioactive contamination. Small amounts of natural and anthropogenic radionuclides may be in the mineral from which the metal is refined, may be added during production of detector-system components, or may be deposited on a

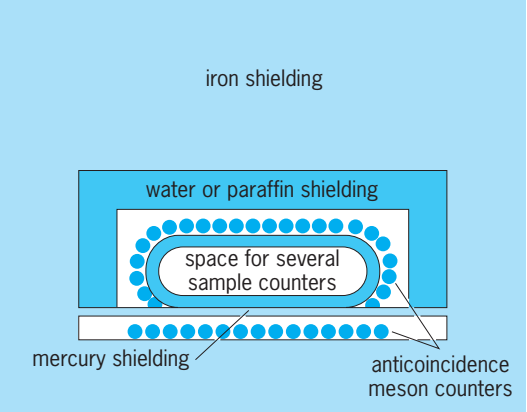

Fig. 1. Cross section of the shielding used for large-volume ultralow-level gas counters.

surface within the shield. Contaminants that have been found include thorium-232 in aluminum, cobalt-60 and ruthenium-106 in steel, and lead-210 in lead. These radionuclides may be at low levels that are difficult to detect until they are used in the low-level counter.

Interfering radionuclides. Radionuclides in the sample that interfere with measuring the radionuclide of interest can be considered another component of the background. These interfering radionuclides may be removed during the chemical and physical separations for extracting the radionuclide of interest from a large sample and for preparing the radionuclide in dimensions that are suitably thin and have been calibrated for counting. In addition, or as an alternative, electronic discrimination may be used to prevent or reduce such interference.

Electronic discrimination. An anticoincidence detector system, either surrounding the detector (Fig. 1) or part of it (**Fig. 2**), is used to reduce the count rate due to external beta particles and gamma rays. The electronic system cancels a pulse from the detector when a simultaneous pulse is recorded in an anticoincidence detector. It must be realized that the system also cancels pulses due to radiations emitted by the source that are counted both in the detector

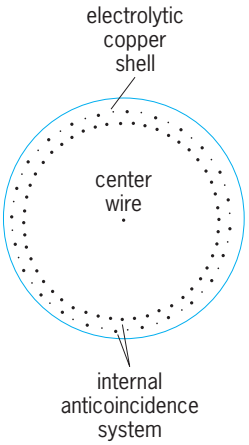

Fig. 2. Cross section of a proportional gas counter with an internal anticoincidence system.

and the anticoincidence detector.

Coincidence systems are used to detect selectively any radionuclide that emits without delay two or more successive radiations. They reverse the anticoincidence mode, accepting only pulses that are recorded simultaneously in two detectors. The two radiations may be an alpha particle, a beta particle, or a gamma ray followed by the same or a different radiation, each counted by a detector that responds preferentially to the radiation of interest. In a different context, liquid scintillation counters couple two photomultiplier tubes and only accept pulses due to scintillations recorded by both tubes in coincidence, to eliminate noise pulses from the individual tubes.

Pulse-height discrimination removes pulses due to radiations that have energies other than those of the radionuclide of interest. This approach is most effective for discriminating among alpha particles, which are emitted typically at one to three discrete energies per radionuclide. It is also effective for gamma rays, which deposit in the detector by photoelectric and multiple interactions discrete energies that correspond to those of emitted gamma rays, in addition to other energies deposited by Compton scattering and pair production. Various detector systems include spectrometers for pulse-height analysis. The pulse height is calibrated in terms of gamma-ray energy, and the area beneath the peak is calibrated for counting efficiency. The peak energies are used to identify the radionuclide, while the areas beneath the peaks give the amount present.

Simple pulse-height analysis is not feasible for beta particles because they are emitted in an energy continuum, with a characteristic maximum energy. It is possible to distinguish among beta-particle groups that have widely different maximum energies and to exclude beta particles beyond the maximum energy of the radionuclide of interest. Pulse-height discrimination also is used to avoid recording low-energy noise, and to separate different types of radiations, for example, alpha particles from beta particles, when the energy deposited in a detector by the different radiations differs widely.

Pulse-shape discrimination is also used to distinguish among types of radiations in a single detector or among pulses detected in different detectors within a counting system. This process can reduce the background by removing noise or pulses from interfering radiations.

Track-detection systems. These can be particularly useful for low-level measurements because they permit detection over periods of months with materials that respond only to specific types of radiation. Selected photographic films are used to record alpha-particle tracks. Alpha-particle tracks are also formed in certain plastics and some natural substances, and can be made visible under a microscope by etching with a strong base. *See* PARTICLE TRACK ETCHING.

Mass measurement. Special mass-measurement techniques can reach levels of activity per atom (that is, per unit mass) of the isotope too low for low-level counting. The activity A, in disintegrations per unit time, is related to the number of atoms N and the half-life of the radionuclide T, in the same units of time, according to Eq. (3). When the half-life

$$A = 0.693NT^{-1} \tag{3}$$

of the radionuclide is much longer than a day, the number of its atoms present is much greater than the number of radiations emitted in a day. By separating the atoms of such a long-lived radionuclide from interfering atoms, ionizing them, and counting them in an ionization detector, the detection sensitivity is lowered by orders of magnitude compared to counting the emitted radiations. Atoms have been separated for this purpose by mass spectrometer, accelerator, and laser excitation. A technique for the chemical analysis of elements at the picogram level that can be applied to measuring low levels of radioisotopes that have half lives in excess of about 10,000 years is inductively-coupled plasma mass spectrometry. The activity can be calculated from the mass according to Eq. (3). *See* ACCELERATOR MASS SPECTROMETRY; ISOTOPE SEPARATION; MASS SPECTROMETRY; RADIOACTIVITY. Bernd Kahn

Bibliography. National Council on Radiation Protection and Measurements, *A Handbook of Radioactivity Measurement Procedures*, 2d ed., 1985.

Low-temperature acoustics

The application of acoustics to research on the properties of condensed matter at low temperatures. Acoustic techniques are readily adaptable to the cryogenic environment and make possible many measurements of the structural and thermodynamic properties of materials at temperatures approaching absolute zero (0 K, which is −460°F or −273°C). The study of sound propagation has also yielded major insights into the low-temperature phenomena of superconductivity in metals and superfluidity in liquid helium.

Low-Temperature Properties of Solid Materials

Acoustic measurements have been used to characterize the properties of a wide variety of solid-state materials, such as metals, dielectric crystals, amorphous solids, and magnetic materials. A measurement of the velocity of sound in a substance gives information on its elastic properties, while the attenuation of the sound characterizes the interaction of the lattice vibrations with the electronic and structural properties of the material.

Experimental methods. Ultrasonic frequencies, in the range from 20 kHz to 100 MHz and above, are commonly employed in these measurements because of the ease of generating and detecting the sound with piezoelectric quartz crystals. In a typical experimental arrangement (**Fig. 1**), the opposite sides of the material being investigated are cut flat and the surfaces polished. Two quartz transducers are cemented to the flat sides. The sound is generated by exciting the quartz with a short voltage pulse at its resonant frequency. The pulse of sound propagates

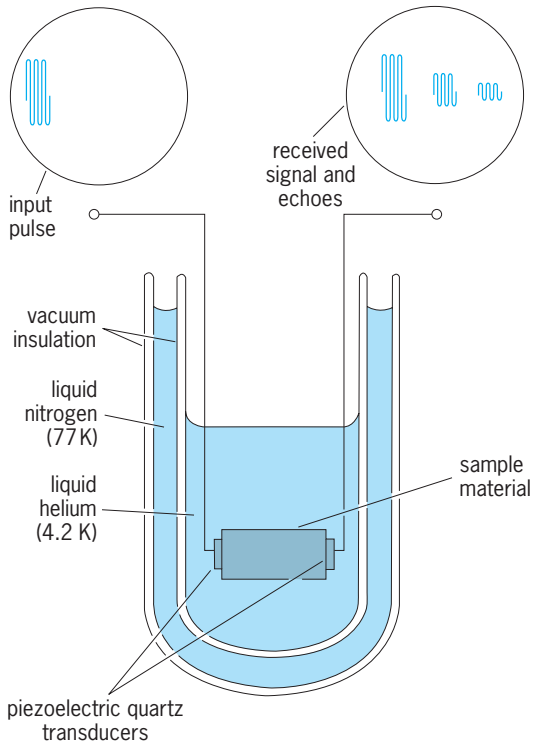

Fig. 1. Experimental arrangement to measure the velocity and attenuation of sound in a sample cooled to 4.2 K (−452°F) using liquid helium.

through the sample and is detected by the receiver-transducer. By measuring the time of flight across the sample, the sound velocity can be obtained. As the sound then reflects back and forth across the sample, the attentuation of the wave is found by measuring the rate of decay of the echo pulses. The temperature of the sample can be varied by bringing it into contact with cryogenic fluids such as liquid nitrogen (77 K or −321°F) or liquid helium (1–4.2 K or −458 to −452°F). Various refrigeration techniques using these liquids allow measurements over the entire range from room temperature (300 K or 80°F) down to well below 1 K (−458°F). *See* ULTRASONICS.

Metals and crystals. Because the sound velocity effectively measures elastic constants, such measurements are used to characterize phase transitions in crystals where the structure of the lattice changes. For example, in solid metallic sodium there is a transition near 36 K (−395°F), where the crystal lattice changes from body-centered cubic to hexagonal close-packed. Both the sound velocity and the attenuation change abruptly when the transition occurs, and provide a means of monitoring the nature of the transition. *See* CRYSTAL STRUCTURE; PHASE TRANSITIONS.

The attenuation of sound in many crystals is due to defects and impurities in the crystal lattice and provides information on such structures. If there is a line of missing atoms (a line dislocation), the stress from the sound field causes it to move back and forth. However, because of pinning forces, the dislocation cannot freely follow the sound oscillation and hence absorbs energy from it. In a metal at very low temperatures, the dominant source of attenuation is the interaction of the sound with the conduction electrons. Oscillations of the positively charged ions making up the crystal lattice couple to the negatively charged electrons. When the mean free path of the electrons becomes longer than the acoustic wavelength (which occurs at low temperatures), the sound wave loses energy to the electrons. Studies of this process are important in determining the strength of the electron-phonon coupling in a material (phonons are thermally excited quantized lattice vibrations). The coupling between the electrons and phonons is responsible for superconductivity. *See* CRYSTAL DEFECTS; LATTICE VIBRATIONS; PHONON; SOUND ABSORPTION.

A large variety of magnetoacoustic effects are observed in metals and crystals. In these measurements, changes in the sound attenuation occur as the strength of a magnetic field applied to the sample is increased. One example is the phenomenon of nuclear acoustic resonance. The nuclear spins in a crystal interact with vibrations of the lattice. When the frequency of the sound matches the Larmor frequency of the spins (which is determined by the magnetic field), the sound can induce transitions in the spin states to higher energies. This results in absorption of the sound, and a resonance is observed in the graph of attenuation versus field. Other magnetoacoustic effects in metals are useful in determining the orbits followed by the conduction electrons in the metal. *See* DE HAAS-VAN ALPHEN EFFECT; LARMOR PRECESSION.

Amorphous materials. Sound propagation is useful for studying amorphous materials. In materials such as silica glass (amorphous silicon dioxide, SiO_2), only two quantum energy levels are found to be important at low temperatures. These levels correspond to two nearly equivalent arrangements of the atoms, with one arrangement having slightly higher energy. An imposed sound field can cause a transition from one arrangement to the other. If the relaxation rate back to the original configuration is comparable to the sound frequency, there will be a net absorption of energy from the sound wave. A peak in the attenuation in silica glass near 50 K (−370°F) has been identified as being due to this process, and measurements as a function of frequency allow a determination of the relaxation rate. *See* AMORPHOUS SOLID.

Superconductors. When a metal is cooled below its superconducting transition temperature, there are striking changes in the attenuation of sound. At the transition some of the electrons near the Fermi surface begin to pair together, because of the attractive electron-phonon coupling. Once this occurs, the electrons can no longer exchange momentum with the lattice, and hence have zero resistance. This also means that the paired electrons no longer absorb energy from the sound wave, and the attenuation is from the remaining unpaired normal electrons. As the temperature is lowered well below the transition, the density of the unpaired electrons drops rapidly, and the attenuation becomes very small. The

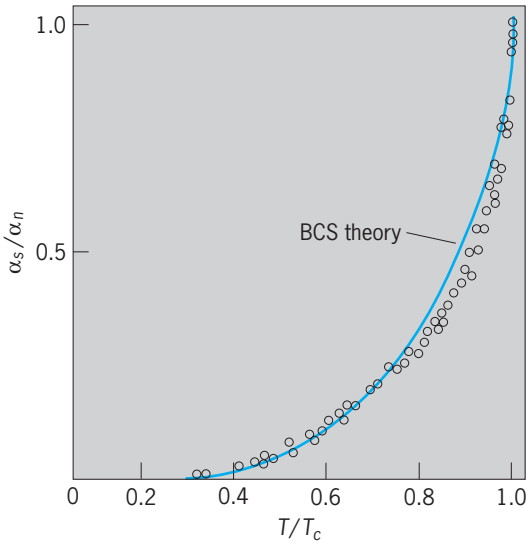

Fig. 2. Ratio of the sound attenuation α_s in the superconducting state to that in the normal state α_n for metallic tin, as a function of temperature T. The superconducting transition temperature T_c is 3.71 K (−453.0°F) for tin.

exponential drop in attenuation is well predicted by the Bardeen-Cooper-Schrieffer (BCS) theory of superconductivity, in substances such as metallic tin (**Fig. 2**). From attenuation measurements, values of the energy gap of the electrons (Δ) can be obtained, and this is the fundamental parameter characterizing the superconducting state.

The sound attenuation is also used to study the characteristics of type II superconductors. In these superconductors the magnetic flux applied to a sample is expelled until a critical field H_{c1} is reached. Above that field the flux enters the sample by threading the materials with an array of flux tubes known as quantized vortex lines. The motion of these pinned vortex lines contributes to the sound attenuation. As the magnetic field is increased, there is a further sharp increase in the attenuation at the second critical field H_{c2}, where the superconductivity is destroyed.

Acoustic techniques play a major role in understanding the unconventional properties of two material systems, the heavy-fermion superconductors and the high-T_c superconductors (where T_c is the superconducting transition temperature). The heavy-fermion metals are transition-metal compounds such as UBe_{13}, $CeCuSi_2$, and UPt_3, in which the conduction electrons are found to have effective masses as much as a thousand times greater than ordinary electrons. The discovery of superconductivity in these systems was quite unexpected, and a number of observations have shown that the predictions of the BCS theory are not followed for these materials. Ultrasonic attenuation measurements show an unusual peak in the attenuation just below the transition temperature, and at lower temperatures the attenuation follows a power-law behavior versus temperature rather than the exponential behavior discussed above. This indicates that the electrons may be paired in an unusual state where their spins point in the same direction (a spin triplet state), rather than in opposite directions (spin singlet state) as assumed by the BCS theory.

The discovery in 1986 of the high-T_c ceramic superconductors [such as $YBa_2Cu_3O_7$ with $T_c = 94$ K (−290°F)] generated intense research activity into the possible mechanisms that could account for such high transition temperatures. By using acoustic measurements of velocity and attenuation, features have been observed at and below the transistion temperature that are connected with the superconducting transition, but a clear interpretation of the data has not yet emerged. *See* SUPERCONDUCTIVITY.

Superfluid Helium

Sound propagation has been extensively used to probe many of the unusual properties of superfluid helium. When liquid ^{4}He is cooled below $T_\lambda = 2.17$ K (−455.76°F), it undergoes a phase transition into the superfluid state. In the two-fluid model the liquid helium below T_λ is treated as being a mixture of two fluids, a superfluid component with density ρ_s and a normal fluid component with density ρ_n. Just below the transition at T_λ, the superfluid density ρ_s is nearly zero, but it increases rapidly as the liquid is cooled, and below 1 K (−458°F) ρ_s is nearly equal to ρ.

The superfluid component has the unusual property that its viscosity is identically zero, and it is able to flow through the tiniest of capillaries. The normal fluid component, on the other hand, has the properties of an ordinary viscous liquid and is completely immobilized in a small capillary. A second characteristic of the superfluid is that it has zero entropy because it is an ordered quantum system, and all of the liquid's entropy is carried by the normal fluid component. This leads to an unusual flow property of the superfluid, namely, that it flows in response to both pressure gradients and temperature gradients in the liquid. An ordinary fluid is accelerated only by pressure gradients. These novel features of the superfluid (zero viscosity and entropy) give rise to a rich variety of different types of sound which can propagate in the superfluid helium. Five distinct sound modes have been identified and observed experimentally (**Fig. 3**).

First sound. First sound is a pressure wave which propagates in the bulk liquid. It is quite similar to sound in ordinary fluids. The velocity C_1 is given by Eq. (1), where B is the adiabatic bulk modulus of

$$C_1^2 = \frac{B}{\rho} \tag{1}$$

the helium. The velocity is relatively temperature-independent at a value of about 778 ft/s (237 m/s), except for the region near T_λ. At that point the velocity dips in a sharp cusp. High-resolution measurements of the sound velocity and attenuation in this region yield valuable information on the critical properties of the superfluid transition.

An important application of first sound is the acoustic microscope using liquid helium at 0.1 K (−459.5°F) as the coupling fluid for imaging.

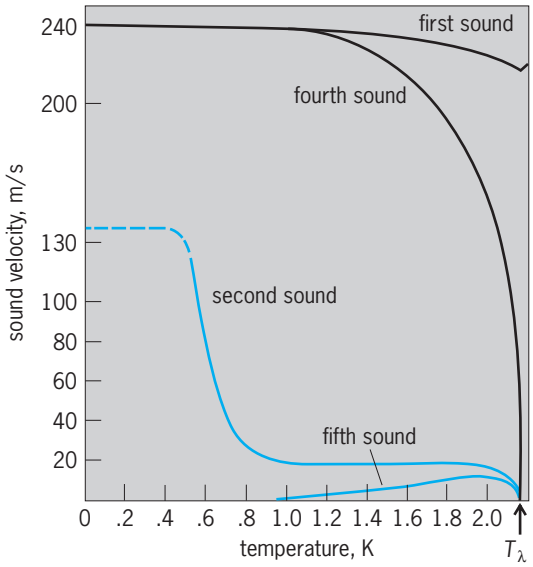

Fig. 3. Velocity of the various types of sound in superfluid ^{4}He as a function of temperature. 1 m/s = 3.28 ft/s. °F = (K × 1.8) − 459.67.

Because of the very low attenuation of first sound at this temperature, it was possible to use higher ultrasonic frequencies (up to 10 GHz) than can be achieved with similar microscopes that operate with water. The higher frequencies resulted in an image resolution of less than 20 nanometers, more than an order of magnitude better than can be obtained with light microscopes. *See* ACOUSTIC MICROSCOPE.

Second sound. Second sound is an unusual type of wave: it is a temperature wave in the bulk superfluid. In this mode the normal fluid and superfluid move in opposite directions. This keeps the density constant, and hence there are no pressure oscillations in the wave (as in first sound); but because only the normal fluid carries entropy, there are oscillations in the entropy and thus in the temperature of the liquid. The second sound can be generated with a heater wire and detected with a carbon resistance thermometer. According to the theory, the sound velocity is given by Eq. (2), where T is the temperature, S the specific entropy, and C the specific heat. The velocity is zero at T_λ (since $\rho_s/\rho = 0$ there), and then increases at lower temperatures to a value of about 65 ft/s (20 m/s). Below 1 K (−458°F) there is a further rapid increase to a value $C_2 = C_1/\sqrt{3} = 460$ ft/s = 140 m/s, but in this region the mean free path of the normal fluid phonon excitations becomes longer than the sound wavelength, and the second sound can no longer be observed below about 0.6 K (−458.6°F). *See* SECOND SOUND.

$$C_2^2 = \frac{\rho_s}{\rho_n}\frac{TS^2}{C} \tag{2}$$

Third sound. Third sound is a wave which propagates in very thin films of helium. If helium gas is introduced into a chamber at low temperatures, some of the helium atoms are attracted to solid surfaces within the chamber by van der Waals forces, forming a thin layer of liquid on all of the surfaces. The thickness of the film can be varied between about 0.5 and 30 nm by adjusting the pressure of the gas in the chamber. The third sound is a wave in which the thickness of the film varies, somewhat like waves in a tank of water. Because the films are so thin, only the superfluid can move, the normal fluid being immobilized by its viscosity. The van der Waals force is the restoring force for the wave, and the velocity is given by Eq. (3), where α is the van der Waals coefficient and d is the film thickness. Experiments have verified this formula. *See* INTERMOLECULAR FORCES.

$$C_3^2 = \frac{\rho_s}{\rho}\frac{3\alpha}{d^3} \tag{3}$$

Third sound has been used to investigate the superfluid transition in thin films, which occurs at temperatures lower than T_λ and is found to depend on the exact thickness of the helium film. These observations have served to verify a general theory of two-dimensional matter, which predicts that ρ_s should drop discontinuously from a finite value to zero at the transition in the thin films.

Fourth sound. Fourth sound is a pressure wave which propagates in superfluid helium when it is confined in a porous material such as a tightly packed powder. In such a situation the normal fluid is immobilized, and only the superfluid can flow freely because of its zero viscosity (the porous materials are often called superleaks for this reason). The fourth sound is analogous to first sound because it involves density and pressure oscillations. Because only the superfluid component can move in the superleak, the velocity is modified from that of first sound and is given by Eq. (4). Because C_1 is known quite accurately, measurements of C_4 can be used with Eq. (4) to obtain highly precise values of ρ_s/ρ as a function of temperature.

$$C_4^2 = \frac{\rho_s}{\rho}C_1^2 \tag{4}$$

Fifth sound. Fifth sound is a temperature wave which can propagate in helium confined in a superleak. It is analogous to second sound, except that again only the superfluid component can flow. This modifies the velocity from that of second sound to Eq. (5). This velocity is zero both at T_λ (where C_2 is zero) and $T = 0$ (where $\rho_n/\rho = 0$). The fifth sound is quite a low-velocity mode, reaching a maximum value of about 40 ft/s (12 m/s) at 1.9 K (−456.2°F).

$$C_5^2 = \frac{\rho_n}{\rho}C_2^2 \tag{5}$$

Detailed measurements of both fourth and fifth sound in a variety of different porous materials have shown that these sound modes can be used to characterize the hydrodynamic flow in the pores. One application of the sound modes has been to study the properties of the interconnected pores in ceramic materials. It is found that the fourth sound velocity of helium in these materials decreases substantially as they are made denser by stronger sintering, allowing a study of the structural changes in the ceramic

as complete solidification is approached. These techniques also have possible applications to the oil industry, where fluid flow in porous rocks is an important technical problem. Fourth sound measurements in porous sandstones can yield parameters that help to interpret data from the logging of the size of oilfield reservoirs.

Superfluid ^{3}He. Acoustic methods have also been very important in determining the superfluid properties of the isotope ^{3}He. A phase transition was discovered in liquid ^{3}He in 1973 at ultralow temperatures, below 0.003 K. The first conclusive evidence that this was a transition to a superfluid state came when it was discovered that fourth sound can propagate in a superleak filled with the ^{3}He. The fourth sound exists only if there is a frictionless superfluid component, and the measurements of ρ_s derived from C_4 were in general agreement with theoretical predictions.

When the sound frequency is increased to a value greater than the rate of collisions between the ^{3}He atoms, the characteristics of the wave change from first sound to what is known as zero sound, where the velocity is slightly higher and the attenuation is reduced. At the superfluid transition there is a sharp peak in the attenuation as a function of temperature, and further variation is found at lower temperatures. In the magnetic A phase, the zero sound attenuation is found to depend on the angle between the direction of the sound beam and the direction of a magnetic field applied on the liquid. This anisotropy is an unusual feature, and the zero sound measurements are able to probe the microscopic structure of the paired ^{3}He atoms. *See* LIQUID HELIUM.

Gary A. Williams

Bibliography. W. P. Mason (ed.), *Physical Acoustics*, vols. 1–5, 1964–1969; W. P. Mason and R. N. Thurston (eds.), *Physical Acoustics*, vols. 6–18, 1970–1988; S. J. Putterman, *Superfluid Hydrodynamics*, 1974; R. N. Thurston and A. D. Pierce (eds.), *Physical Acoustics*, vols. 19–22, 1990–1992.

Low-temperature physics

A branch of physics dealing with physical properties of matter at temperatures such that thermal fluctuations are greatly reduced and effects of interactions at the quantum-mechanical level can be observed. As the temperature is lowered, order sets in (either in space or in motion), and quantum-mechanical phenomena can be observed on a macroscopic scale. As an example, consider a magnetic system. In a sample at high temperatures, there is a high degree of disorder (**Fig. 1***a*) corresponding to a large entropy because of the random orientations of its magnetic moments, the sample being in a paramagnetic state. At low temperatures (Fig. 1*b*), the sample is in a highly ordered state, the ferromagnetic state (lower entropy), and the details of the interactions responsible for this state can be studied. *See* ENTROPY; MAGNETISM.

Some of the most interesting manifestations of low temperatures have been investigated in the temperature range from 4 K (−452°F) down to less than a nanokelvin above absolute zero. (1 K is equal to 1.8°F above absolute zero, or −459.67°F.) Certain metals become superconducting, losing their electrical resistance entirely; hence persistent currents can flow indefinitely in a superconducting ring or coil, displaying quantum-mechanical coherence over large distances. The liquids helium-3 (^{3}He) and helium-4 (^{4}He) remain liquid down to absolute zero under their own vapor pressure due to the large zero-point energy of these light atoms. (To overcome the large zero-point energy in liquid ^{3}He and liquid ^{4}He, a large pressure, approximately 30 atm or 3 megapascals, must be applied to cause these systems to solidify.) Liquid ^{4}He becomes superfluid, exhibiting no resistance to flow under certain conditions; when set in circulation, the fluid current persists indefinitely. Liquid ^{3}He also becomes superfluid at a much lower temperature with interesting magnetic and orbital effects. At sufficiently low temperatures, nuclear magnetic ordering has been observed in solid ^{3}He, in magnetic insulators, and in metallic systems. Silver becomes a nuclear antiferromagnet in the nanokelvin range as a result of quantum-mechanical exchange interactions. Considerable attention has been addressed to the general problem of ordering in disordered systems leading to studies of spin glasses, localization, and lower dimensionality. Quantum statistics are investigated in atomic hydrogen and deuterium, stabilized in states known as spin-polarized hydrogen (H↓) and spin-polarized deuterium (D↓). Because of its light mass and weak interactions, spin-polarized hydrogen is expected to remain gaseous down to absolute zero, whereas spin-polarized deuterium might liquefy at low temperatures. *See* HYDROGEN; LIQUID HELIUM; SUPERFLUIDITY.

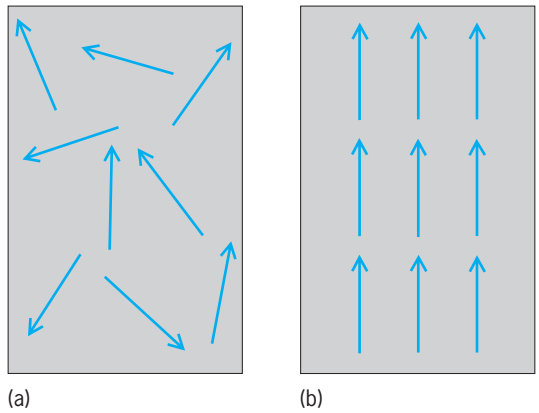

Fig. 1. Ordering of a magnetic system at (*a*) high temperatures and (*b*) low temperatures.

Low-temperature research also deals with problems of thermometry and heat transfer between systems and within systems. Many practical applications have emerged, including the use of superconductivity for large magnets, ultrafast electronics for computers, and low-noise and high-sensitivity instrumentation. This type of instrumentation has

opened new areas of research in biophysics, and in fundamental problems such as the search for magnetic monopoles, gravity waves, and quarks. *See* LOW-TEMPERATURE THERMOMETRY; SUPERCONDUCTING DEVICES.

Production of low temperatures. Because of the need for small quantities of liquid ^{4}He in superconducting instrumentation and electronics, small cyclic heat engines are needed for continuous refrigeration. For lower temperatures, the refrigeration is produced by a variety of methods (depending on the experiment) in the research apparatus itself, called the cryostat. Some of the most important ones are described below.

Liquefaction of helium. Liquid helium at its normal boiling point of 4.2 K (−452.1°F) is a basic working substance used as a platform for many experiments and for attaining lower temperatures. Liquid is produced from the gas phase by a combination of adiabatic work on a piston and expansion through a Joule-Thompson valve; the cycle is maintained by use of a compressor, liquid-nitrogen precooling, and heat exchangers.

Since ^{3}He is rare and hence expensive, small quantities are liquefied in the cryostat in a closed cycle using liquid ^{4}He whose temperature is reduced by adiabatic evaporation. *See* LIQUEFACTION OF GASES.

Adiabatic evaporation. Evaporation is one of the most powerful methods of producing cooling, and it is used for reducing the temperature of liquid ^{4}He or liquid ^{3}He. **Figure 2** shows a typical arrangement where the liquid is in a container, isolated by a vacuum. Cooling results when the vapor pressure in the container is reduced by a pump. The rate of energy extraction by this refrigerator depends on the latent heat of the liquid and the pumping speed. The lowest temperatures achieved are typically 0.8 K (−458.2°F) for liquid ^{4}He and 0.2 K (−459.3°F) for liquid ^{3}He. At these temperatures, the vapor pressure is so low that the slow rate at which atoms can be removed from the liquid phase sets the limit.

^{3}He–^{4}He dilution refrigerator. This is an important method for producing low temperatures continuously by dissolving ^{3}He in liquid ^{4}He. This process is analogous to the above evaporation process but with an inverted geometry (**Fig. 3**), and relies on the fact that a mixture of ^{3}He–^{4}He separates into two phases below 0.87 K (−458.10°F). In a container, the mixing chamber, liquid ^{3}He is diluted in ^{4}He, since it has a finite solubility in ^{4}He, being 6.4% at 0 K (−459.67°F). Because liquid ^{4}He is a superfluid, it acts like a mechanical vacuum wherein ^{3}He atoms behave like an ideal gas. On passing from the ^{3}He-rich phase to the ^{4}He-rich phase, ^{3}He atoms produce cooling. To maintain continuous cooling, ^{3}He is extracted from the ^{4}He-rich phase at a still by means of a pump; it is then recondensed at a 1 K liquid-^{4}He evaporator and dissolved again in the mixing chamber, the ^{4}He being stationary in the cycle. Efficiency is improved by the use of heat exchangers. Temperatures as low as 2 millikelvins (3.6×10^{-3}°F above absolute zero) can be achieved continuously. The limiting temperature is determined by the decreasing entropy available and the intrinsic limitations due to viscous heating of the circulating ^{3}He and other heat influxes.

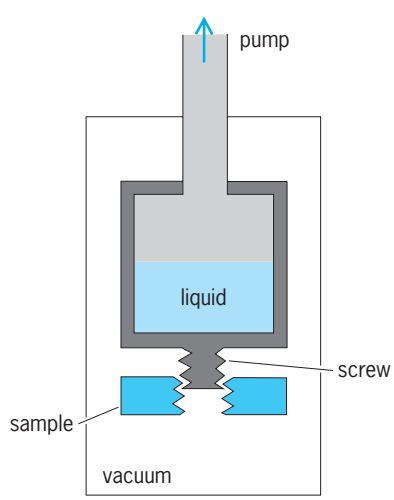

Fig. 2. Liquid-helium evaporator with screw for attaching sample.

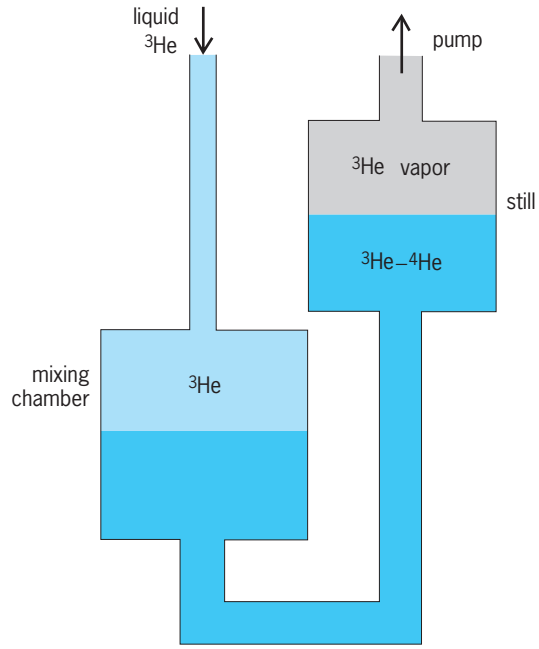

Fig. 3. Basic ^{3}He–^{4}He dilution refrigerator.

Adiabatic compression of ^{3}He. This method of refrigeration is based on the fact that below 0.32 K (−459.09°F) the entropy of solid ^{3}He is larger than that of the liquid. Liquid ^{3}He is a Fermi liquid with a high degree of motional order (because of quantum statistics); the solid has a high degree of disorder due to the random orientations of the ^{3}He nuclear magnetic moments. Changing the liquid to the solid under adiabatic conditions below 0.32 K (−459.09°F) leads to cooling of the solid-liquid mixture to temperatures as low as 1 mK (1.8×10^{-3} °F above absolute zero). The conversion to solid is achieved by applying pressure at the melting curve of ^{3}He. In accord with the Clausius-Clapeyron equation, the slope of the melting curve is negative below 0.32 K (−459.09°F), and hence solid is produced.

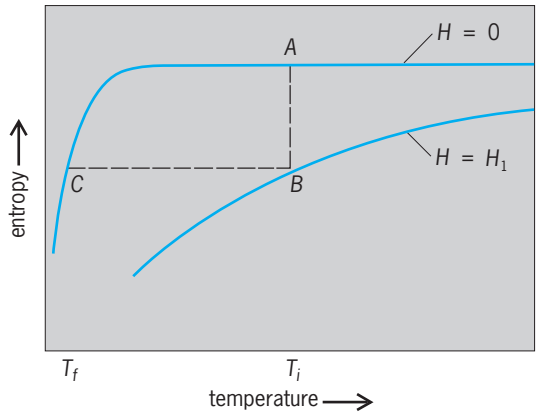

Fig. 4. Magnetic cooling process on entropy-temperature diagram.

This method is particularly interesting for cooling solid ^{3}He and even liquid ^{3}He.

Adiabatic demagnetization. This is a method of producing temperatures much below 1 K (−458°F) using a paramagnetic system whose entropy can be reduced by a magnetic field. **Figure 4** shows the entropy due to random orientations of magnetic moments in a paramagnet for zero magnetic field, $H = 0$. In a magnetic field and at low temperatures, the entropy is reduced and it is shown for some value of the field, H_1. As the system follows the path from A to B at constant temperature T_i, heat is given off during magnetization. Then, by an adiabatic (isentropic) demagnetization from B to C, the system cools to T_f. The limiting temperature is determined by the internal magnetic fields in the system. With an electronic paramagnet such as cerium magnesium nitrate, temperatures of approximately 1 mK (1.8×10^{-3} °F above absolute zero) can be reached from 1 K (−458°F). With a nuclear paramagnet, much lower temperatures can be achieved since the internal fields are much weaker than in electronic systems. Usually, metallic copper is used, and submicrodegree temperatures have been achieved in the spin system. Starting conditions require a large magnetic field and low temperatures. For work just below 1 mK (1.8×10^{-3} °F above absolute zero), it is easier to use hyperfine-enhanced nuclear cooling with a system such as $PrNi_5$. The strong coupling between electronic and nuclear magnetic moments enhances the entropy reduction with modest magnetic fields, and hence the initial conditions are less stringent. *See* ADIABATIC DEMAGNETIZATION.

Low-temperature phenomena. At a temperature of 2.17 K (−455.76°F), liquid ^{4}He undergoes a phase transition to a state characterized by flow without any resistance under certain conditions. The liquid then has very large heat conductivity (much better than many metals), peculiar film-flow behavior, and interesting and unusual sound and thermal propagation properties. Quantized vortices can exist in the rotating fluid at certain speeds, and above a certain critical velocity dissipation sets in. Rotating superfluid ^{4}He provides the best testing ground for theories of vortex dynamics. Numerous experiments demonstrate the unusual behavior of a superfluid, including photographic recordings of vortices. *See* QUANTIZED VORTICES.

The explanation for this behavior is that liquid ^{4}He obeys Bose-Einstein statistics and that a condensation to the ground state is expected. Liquid ^{3}He obeys Fermi-Dirac statistics, and its behavior is different than ^{4}He. After much theoretical and experimental work, it was discovered that liquid ^{3}He also undergoes a superfluid transition at 0.93 mK (1.67×10^{-3} °F above absolute zero) under zero pressure. The behavior is comparable to superconductivity, but instead of having pairing in a nonmagnetic singlet state, the ^{3}He spins pair in triplet states. Unusual behavior occurs in liquid ^{3}He as a result of the magnetic and nonmagnetic triplet pairs. The pairing explains the three known superfluid phases: A_1, A, and B near the melting curve. Many experimental and theoretical investigations of these ordered phases of the magnetic superfluid have discovered unexpected behavior in their magnetic properties, flow characteristics, and sound propagation, where collective excitations contribute from zero sound to fourth sound. *See* BOSE-EINSTEIN STATISTICS; FERMI-DIRAC STATISTICS; LOW-TEMPERATURE ACOUSTICS.

Superconductivity. This phenomenon is another manifestation of superfluidity, the fluid here being the conduction electrons in certain metals and compounds. Below a critical temperature T_c ranging from 133 K (−220°F) for Hg-Ba-Ca-Cu-O to 0.325 mK (5.85×10^{-4} °F above absolute zero) for rhodium, the materials show no electrical resistance as a result of quantum-mechanical phase coherence over macroscopic distances. The impetus provided by the BCS (Bardeen-Cooper-Schrieffer) theory of superconductivity led to investigations of a variety of phenomena such as tunneling, flux quantization, very sensitive devices called SQUID (superconducting quantum interference devices), new materials, amorphous superconductors, vortex flow and dissipation, and magnetic superconductors. The BCS theory led to the suggestion that electron pairing could occur in media of high electrical polarizability. This suggestion of possible superconductivity in organic molecules has stimulated a great deal of research in the field of conducting organic molecular solids. Consequently, materials such as $(TMTSF)_2ClO_4$ (TMTSF = tetramethyltetraselenafulvalene) have been found to be superconducting below 1 K (−458°F). The discovery of the Josephson effect and flux quantization in superconductors has led to the development of devices of extremely high sensitivity. *See* JOSEPHSON EFFECT; SQUID.

Because the cost of refrigeration is an important factor in the applications of superconductivity, there has always been interest in trying to find superconductors with high transition temperatures. For many years, Nb_3Ge had the highest transition temperature, 23.3 K (−417.7°F). In 1986 a new class of materials was discovered to have high critical temperatures ranging from 40 to 93 K (−388 to −292°F). These materials are oxides, usually in ceramic form, $YBa_2CU_3O_7$ being a typical example with

temperature of 92 K (−294°F). Transition temperatures as high as 133 K (−220°F) have been observed. Such unexpected behavior attracted much interest with the hope of achieving even higher transition temperatures and hence using cheaper refrigeration. The mechanism responsible for superconductivity in these materials is not clear; based on experimental facts, the applicability of the BCS theory to the high-T_c materials has been questioned. Although at the present there are problems of material preparation and of physical understanding of their behavior, these materials have great potential for a variety of applications ranging from high magnetic fields to electronics. The field of high-T_c superconductors is still part of low-temperature physics since it deals with a quantum-mechanical system on a macroscopic scale. *See* SUPERCONDUCTIVITY.

Quantum crystals. The amplitude of quantum-mechanical zero-point vibrations is especially large in crystals of light elements. It can become comparable to the lattice parameters. As a consequence of these vibrations, there is a high quantum-mechanical probability of particles tunneling into the neighboring lattice sites. This leads to quantum-mechanically delocalized particles in the crystal, similar to a quantum fluid. The largest effects occur for ^{3}He, followed by ^{4}He, hydrogen, and neon. Such effects are responsible for the direct-exchange interactions which influence the nuclear magnetism of ^{3}He. Nuclear-magnetic-resonance measurements of spin diffusion and relaxation, as well as susceptibility measurements, substantiate such a model. The antiferromagnetic phase structure found in solid ^{3}He around 1 mK (1.8×10^{-3}°F above absolute zero) is attributed to such many-particle exchange. *See* EXCHANGE INTERACTION; NUCLEAR MAGNETIC RESONANCE (NMR); QUANTUM SOLIDS.

Lower dimensionality and localization. These phenomena involve electrical conduction in thin films behaving as two-dimensional systems and in one-dimensional-like long narrow channels. This subject is of great theoretical importance, and is also of great importance in the development of very small circuits and devices in the submicrometer limit. Localization theory argues that a wave cannot propagate through a disordered one-dimensional system. In a disordered system, electrons can be localized in regions of disorder, and hence electrical conductivity will be strongly affected at very low temperatures. Quantum-mechanical effects govern the behavior, one being the exponential increase of resistance with the length of a very thin wire. Fundamental questions of electrical conduction are studied, experimentally and theoretically, with benefits for the design of very small circuits.

Lower-dimensionality magnetism. In order to understand three-dimensional systems, models in one dimension (for example, the Ising model) have been proposed and studied. The results are extremely interesting, and they have been applied to real systems of long chains and two-dimensional arrays of magnetic moments. The absence of phase transitions in one dimension has led to a more fundamental understanding of magnetic ordering at low temperatures. *See* ISING MODEL.

Dilute magnetic alloys. The study of dilute alloys was originally proposed to help solve the problem of magnetism in metals, but it became a fundamental problem in itself. The basic question deals with the behavior of a magnetic impurity in a nonmagnetic metallic host. A wide spectrum of behavior is observed from the nonmagnetic behavior of the impurity in one host to strong magnetic behavior in others. The electrical resistance, after going through a minimum, shows a logarithmic rise as the temperature is lowered; this is the Kondo effect. At higher magnetic impurity levels, the exchange coupling between the impurities, being of random strength and sign, leads to spin-glass freezing below a certain temperature, and to irreversible behavior. A question of fundamental interest is the nature of the freezing and the type of order that sets in at lower temperatures. This problem has been extended to semiconductors and insulators where randomness in the impurity interactions has also led to spin-glass behavior. *See* KONDO EFFECT; SPIN GLASS.

Spin-polarized hydrogen and deuterium. When atomic hydrogen and deuterium are stabilized (so as not to form molecules), interesting quantum gas behavior is observed. Stabilization can be achieved by large magnetic fields, and the systems are then called spin-polarized hydrogen and spin-polarized deuterium. Spin-polarized hydrogen provides a gas for which Bose-Einstein condensation is expected, while spin-polarized deuterium is expected to produce superfluidity due to fermion pairing, similar to liquid ^{3}He. Working with a gas provides researchers with the possibility of studying weakly interacting systems for which theories exist. Actually it has never been shown that liquid ^{4}He is really a Bose-Einstein system since the interactions are so strong. Research in this area, which is usually at temperatures below 1 K (−458°F), deals with fundamental questions of quantum statistics.

Magnetic superconductors. The observation of the coexistence of superconductivity and magnetism, made possible by the discovery of superconducting stoichiometric rare-earth compounds, has in turn opened the possibility of observing a wide range of phenomena. The occurrence of antiferromagnetic order in the superconducting state is deduced from the presence of a lambda-type anomaly in the specific heat and a cusp in the magnetic susceptibility. In some ternary rare-earth compounds, superconductivity is destroyed at a second lower critical temperature by the onset of long-range ferromagnetic ordering.

Quantum Hall effect. This is a form of the Hall effect observed in semiconductors with thin-film structure. At low temperatures and in high magnetic fields, the Hall conductance is quantized in integral multiples of e^2/h, where e is the electronic charge and h is Planck's constant, varying as the number of electrons is changed and dependent on localized states. Experiments have shown that the Hall conductance below 2 K (−456°F) exhibits also odd-denominator

and even-denominator simple fractions of e^2/h, attributed to two-dimensional quantum fluid behavior of electrons. This fractional quantum Hall effect is a many-particle effect that is based on electron-electron correlations. *See* HALL EFFECT.

Tunneling. According to quantum mechanics, particles trapped between potential barriers can escape by tunneling due to the wave nature of the particles. At low temperatures, thermal processes are reduced, making it possible to observe such tunneling. For very small structures, such as a junction formed by two electrodes separated by an insulator, single-electron tunneling effects can be observed when the electron electrostatic energy is larger than its thermal energy. This effect has been applied to single-electron transistors and quantum dots. On a larger scale, the applicability of quantum mechanics to macroscopic objects is being studied in macroscopic quantum tunneling experiments on two systems, Josephson junctions and small magnetic particles. *See* QUANTIZED ELECTRONIC STRUCTURE (QUEST); TRANSISTOR.

When the dissipation is reduced, tunneling of a macroscopic variable can be observed. At very low dissipation, there is a chance of seeing macroscopic quantum coherence in the tunneling experiments. *See* NONRELATIVISTIC QUANTUM THEORY; TUNNELING IN SOLIDS.

Other phenomena. The development of low-temperature techniques has revealed a wide range of other phenomena. The behavior of oriented nuclei is studied by observing the distribution of gamma-ray emission of radioactive nuclei oriented in a magnetic field. Other areas of study include surfaces of liquid ^{3}He and liquid ^{4}He, ^{3}He–^{4}He mixtures, cryogenics, acoustic microscopy, phonon spectroscopy, monolayer helium films, molecular hydrogen, determination of the voltage standard, and phase transitions. *See* ACOUSTIC MICROSCOPE; CRYOGENICS; ELECTRICAL UNITS AND STANDARDS; NUCLEAR ORIENTATION; PHASE TRANSITIONS.

Orest G. Symko

Bibliography. D. S. Betts, *Introduction to Millikelvin Technology*, 1989; D. F. Brewer (ed.), *Progress in Low Temperature Physics*, vols. 7–12, 1982–1989; W. G. Clark, Proceedings of the 16th International Conference on Low Temperature Physics, *Physica*, vols. 109, 110 B and C, 1982; R. P. Hudson, *Principles and Application of Magnetic Cooling*, 1972; O. Lounasmaa, *Experimental Principles and Methods Below 1 K*, 1974; K. Mendelssohn, *The Quest for Absolute Zero*, 1977; F. Pobell, *Matter and Methods at Low Temperatures*, 2d ed., 1996; G. K. White, *Experimental Techniques in Low-Temperature Physics*, 3d ed., 1987.

Low-temperature thermometry

The measurement of temperature below 0°C (32°F). Very few thermometers are truly wide-range, and hence most of the conventional methods of thermometry tend to fail the further the temperature drops below room temperature (see **table**). The defining instrument for a great part of the International Temperature Scale, the platinum resistance thermometer, rapidly loses sensitivity below 30 K, and its official limit is set at 13.81 K, the triple point of equilibrium hydrogen. This scale is based upon measurements of thermodynamic temperature made with the gas thermometer; the gas thermometer may be used down to about 2 K, for which purpose it is usually filled at, and "piggybacked" upon, a temperature in the region of the normal boiling point of equilibrium hydrogen, about 20.28 K. Liquid-in-glass thermometers may be used down to about −39°C or −38°F (mercury), about −100°C or −148°F (ethyl alcohol), and almost down to 70 K with various organic fluids. A liquid becomes sluggish at the lower end of its range, and also, with the exception of mercury, it presents problems of wetting the capillary wall and draining. Thermocouples are quite commonly encountered as temperature sensors down to 77 K and below, and thermistors and semiconductor diode sensors are used down to the liquid helium region, all three being available commercially in conjunction with linearizing electronics. Certain thermocouples have been developed especially for use in the 1–20 K region, and others with enough sensitivity for use in the millikelvin region. The electronic thermometer based upon the thermal expansion of a quartz frequency-controlling element may be used down to −40°C (−40°F). *See* GAS THERMOMETRY; THERMISTOR; THERMOCOUPLE.

In the liquid cryogen region, the saturation vapor pressure is used as a convenient and sensitive indicator of temperature. The usable range from 10 pascals pressure, for example, up to the critical point for the "permanently" liquid heliums corresponds to a temperature range from 0.45 to 3.32 K (at 1.15 atm or 116.5 kilopascals) for ^{3}He and 0.95 to 5.20 K (at 2.26 atm or 229.0 kPa) for ^{4}He. The method is also quite commonly used for equilibrium hydrogen between the triple point and the normal boiling point or above, but not much in modern times for liquid nitrogen (triple point of 63.15 K, normal boiling point of 77.35 K). *See* LIQUID HELIUM; VAPOR PRESSURE.

Thermometry below 1 K. The region below 1 K is special, not only because quantum phenomena are present on a macroscopic scale, but also because of the intrusion of a thermodynamic uncertainty principle: the small heat capacities of many systems studied preclude measurement techniques which gratuitously introduce energy and modify the state of the subject. In addition, relaxation times are long, and a system, once perturbed, may take minutes, hours, or even days to return to a state of equilibrium. Thus it is difficult to find an accurate, reproducible, and "well-behaved" thermometer that is very sensitive, small, and of low heat capacity, and the investigator is inclined to discount cost and complexity. The magnitude of a temperature, or a temperature change, is significant only in relation to the quantity E/k, or $\Delta E/k$, where k is Boltzmann's constant, and E and ΔE are energies which characterize the phenomenon under investigation. Thus, to a nuclear spin system

Ranges and sensitivities of low-temperature thermometers

Thermometer	Temperature range, K	Typical sensitivity or precision at low-temperature limit
Thermocouples		
300 to 700 ppm Fe in Au/Ag + 0.37 at. % Au	1–25	10 $\mu V \cdot K^{-1}$
Chromel/300 to 700 ppm Fe in Au	1–300	10 $\mu V \cdot K^{-1}$
Chromel/Constantan	20–1100	8.5 $\mu V \cdot K^{-1}$
Resistance thermometers		
Platinum (capsule)	4–500	±3 mK [3.3 $\mu V \cdot K^{-1}$ at 10 mA]
Rhodium + 0.5 at. % Fe	0.5–300	±30 μK [350 $\mu V \cdot K^{-1}$ at 0.5 mA]
Carbon	0.01–300	~5 μK
Germanium	0.01–30	~5 μK
Saturation vapor pressure thermometers		
Hydrogen	14–21	1 mK [dT/dp ~30 mK $(mmHg)^{-1}$~0.23 mK $\cdot Pa^{-1}$]
Helium-4	1.0–5.2	20 mK [dT/dp ~0.85 K $(mmHg)^{-1}$ ~6.4 mK $\cdot Pa^{-1}$]
Helium-3	0.5–3.3	10 mK [dT/dp ~0.44 K $(mmHg)^{-1}$ ~3.3 mK $\cdot Pa^{-1}$]
Noise thermometer*	0.002–0.1	1–5%
Magnetic thermometers		
Gadolinium metaphosphate, $Gd(PO_3)_3$	2–100	100 nK
Cerous magnesium nitrate (CMN), $Ce_2Mg_3(NO_3)_{12} \cdot 24H_2O$; single crystal	0.003–4	5 nK
CMN powder sphere or cylinder	0.002–4	1 nK
Copper (and other nuclear paramagnets)	0.001–0.01	100 nK
Gamma-ray anisotropy thermometers*	0.002–0.05	1–10%
^{60}Co in hexagonal close-packed cobalt single crystal	0.002–0.04	
^{54}Mn in iron	0.003–0.03	
^{54}Mn in nickel	0.004–0.045	
^{3}He melting-curve thermometer	0.001–1	1 μK
Nuclear resonance thermometer	310 nK–2 K	1–20%

*Primary thermometer.

which might become ferromagnetic at 10^{-6} K, a temperature change of 1 microkelvin is very large, while room temperature would be much as the Sun's interior is to a human being. *See* ADIABATIC DEMAGNETIZATION.

For many years, the traditional method for measuring temperature below 1 K was that of magnetic thermometry, wherein the susceptibility χ of paramagnetic salts which closely obey the Curie law, $\chi \propto T^{-1}$, is utilized. A sample is calibrated at a known temperature in the 1–4 K region, and then, for lower temperatures, a "magnetic temperature" T^* is extracted from the inversion of Curie's law. The establishment of the exact relationship between T^* and T, which is of greatest importance at the low end of the range, where the simple law begins to break down, was a lengthy and difficult process. The magnetic thermometer is still used because of its basic simplicity, and also as a verifier and "smoother" of temperature scales derived by other methods. Enhanced detection sensitivity is available through superconducting quantum interference device (SQUID) technology, so that very much smaller sensors can be employed, always an important advantage for low-temperature research. The magnetic-thermometry technique has also been adapted to nuclear spin systems for measurements in the microkelvin region (discussed below). *See* CURIE-WEISS LAW; LOW-TEMPERATURE PHYSICS; MAGNETIC THERMOMETER; PARAMAGNETISM; SQUID; SUPERCONDUCTING DEVICES.

Although primary in principle, the ^{3}He melting-curve thermometer seems to deliver the highest accuracy as a secondary instrument, making use of a pressure-temperature relation obtained by temperature calibration. It is used over the range 1 mK to 1 K and offers very high precision (better than 1 μK at the low-temperature end). The accuracy remains somewhat uncertain, especially at the lowest temperatures, and has been the subject of intensive research. *See* QUANTUM SOLIDS.

As at room temperature and above, resistance thermometers tend to be accorded highest favor as far as practicality is concerned, but very few types having adequate sensitivity exist. Doped germanium semiconductors and carbon resistors are much in use down to temperatures as low as a few millikelvins, the germanium being somewhat more reproducible, if much more costly. A normally resistive rhodium-iron alloy (0.5% Fe) is also useful for the region above 0.1 K (and up to 300°C), and serve, in fact, as an extremely important transfer standard for use in thermometry calibration service.

Superconducting transitions have been developed for temperature reference points between 15 mK (tungsten) and 9.3 K (niobium).

Primary thermometers. The low-temperature region is unique in having available several different types of primary thermometers, all of which are quite practical. The least practical, perhaps, is the acoustic thermometer, which uses the property that, extrapolated to zero pressure, the speed of sound in a gas is proportional to $T^{1/2}$. This has been used in the range 2–20 K as an alternative to, and check upon, the gas thermometer. The Johnson noise in a resistor can be used with particular advantage at low temperatures when allied with SQUID detector technology. This yields a value of temperature through a restricted form of the Nyquist relation $\overline{P(t)} = 4kTR\Delta f$, where $\overline{P(t)}$ is the mean noise power (averaged over

infinite time, to be rigorous), R the resistance, k the Boltzmann constant, and Δf the frequency bandwidth. *See* ELECTRICAL NOISE; SOUND.

In suitable systems it is possible to spatially orient atomic nuclei at very low temperatures, and if these nuclei are emitters of gamma rays, the emission pattern is anisotropic to a degree which is a measure of the thermodynamic temperature. Finally the magnetic susceptibility of suitable atomic nuclei may also be employed via the Curie law. This magnetism is very weak, which gave it the disadvantage that it was difficult to measure accurately, prior to the advent of modern technologies, and the advantage that the Curie law validity persists down to the microkelvin region. Nuclear magnetic resonance or static SQUID-based techniques may be employed, but the former is preferable in being unaffected by magnetic impurities, to which the second method falls hostage.

The nuclear magnetic resonance thermometer needs one calibration point, and since this calibration cannot be performed at the kelvin-defining triple point of water, but must be carried out at 1 K or lower, some deny it primary thermometer status. *See* NUCLEAR MAGNETIC RESONANCE (NMR); NUCLEAR ORIENTATION.

Revision of IPTS-68. In 1990, the International Committee of Weights and Measures promulgated an extensively revised version (ITS-90) of the scale that it recommended 22 years earlier, the IPTS-68. Two significant features are an extension down to 0.65 K (far below the earlier limit of 13.81 K) and improved approximation to thermodynamic temperature in all ranges. In this effort, input data came from many sources. A magnetic thermometer was used between 1 and 83 K to detect irregularities in the 1968 scale. Values for defining fixed points were improved by using acoustic, constant-volume gas, and gas-isotherm thermometry over the range 1–30 K. The role of the high-temperature thermocouple has been eliminated, and the platinum resistance thermometer has been adopted as the recommended interpolation instrument for the entire temperature region between the freezing point of silver (961.78°C) and the triple point of hydrogen (13.81 K). The most striking change is that, for the first time, alternative realizations of the scale are permitted, for operational convenience and when no significant discrepancies arise from those alternatives. Thus, for example, the lowest region of the scale, 0.65 K–3.2 K, utilizes the saturation vapor pressure of ^{3}He, while the next higher range, 1.25 K–5.0 K, is realized in terms of the saturation vapor pressure of ^{4}He, with evident overlap between 1.25 K and 3.2 K. *See* TEMPERATURE; TEMPERATURE MEASUREMENT; THERMOMETER.

Ralph P. Hudson

Bibliography. R. P. Hudson et al., Recent advances in thermometry below 300 mK, *J. Low Temp. Phys.*, 20:1, 1975; B. W. Mangum and G. T. Furukawa, *Guidelines for Realizing the International Temperature Scale of 1990 (ITS-90)*, NIST Tech. Note 1265, 1990; F. Pobell, *Matter and Methods at Low Temperatures*, 2d ed., 1996.

Lubricant

A gas, liquid, or solid used to prevent contact of parts in relative motion, and thereby reduce friction and wear. In many machines, cooling by the lubricant is equally important. The lubricant may also be called upon to prevent rusting and the deposition of solids on close-fitting parts.

Liquid hydrocarbons are the most commonly used lubricants because they are inexpensive, easily applied, and good coolants. In most cases, petroleum fractions are applicable, but for special conditions such as extremes of temperature, select synthetic liquids may be used. At very high temperatures, or in places where renewal of liquid lubricants is impossible, solid lubricants (graphite or molybdenum disulfide) may be used.

Petroleum lubricants. Crude petroleum is an excellent source of lubricants because a very wide range of suitable liquids, varying in molecular weight from 150 to over 1000 and in viscosity from light machine oils to heavy gear oils, can be produced by various refining processes. Crude petroleum fractions consist of saturates (normal, iso-, and cyclo- paraffins); monoaromatics which may contain saturated rings as well as saturated side chains; substituted polyaromatics; hetero compounds containing sulfur, nitrogen, and oxygen; and asphaltic material made up of polycondensed aromatics and hetero compounds. Of these, the wax-free saturates and monoaromatics are desired in the finished oil in order to obtain the desired viscometric properties and oxidation and thermal stability. So far as possible, the other types of compounds are generally removed. Classically, only those crudes (so-called paraffinic crudes) relatively rich in the desired type of hydrocarbons were used in lubricating oil manufacture. By atmospheric distillation, neutrals were produced as distillates and the residual fractions were treated with activated clays, which removed asphaltic and hetero material to produce bright stock.

In modern refining, vacuum distillation removes the desired hydrocarbons from the asphaltic constituents (which may be present in amounts up to 40% of the total) and gives oils in the required viscosity and boiling range. Extraction of the distilled fractions with solvents such as liquid sulfur dioxide, furfural, and phenol permits the removal of the polyaromatic and hetero compounds to improve viscosity-temperature characteristics (viscosity index, VI) and stability of the oil. Dewaxing removes the high-melting paraffins. *See* DEWAXING OF PETROLEUM.

If the viscosity index of the oil is not important for a particular application, the most reactive polyaromatics and hetero compounds may be removed by treatment with concentrated sulfuric acid. In either case, the oil is treated finally with an active clay to remove trace amounts of residual acids and resins (heterocyclics).

Small amounts of heterocyclics, such as substituted benzthiophenes, quinolines, and indoles, may remain in the finished oil. The sulfur-containing

Viscosity of oils for various applications

Application	Viscosity in centistokes at 77°F (25°C)	Primary function
Engine oils		
SAE 10W	60–90	Lubricate piston rings, cylinders, valve gear, bearings; cool piston; prevent deposition on metal surfaces
SAE 20	90–180	
SAE 30	180–280	
SAE 40	280–450	
SAE 50	450–800	
Gear oils		
SAE 80	100–400	Prevent metal contact and wear of spur gears, hypold gears, worm gears; cool gear cases
SAE 90	400–1000	
SAE 140	1000–2200	
Aviation engine oils	200–700	Same as engine oils
Torque converter fluid	80–140	Lubricate, transmit power
Hydraulic brake fluid	35	Transmit power
Refrigerator oils	30–260	Lubricate compressor pump
Steam-turbine oil	55–300	Lubricate reduction gearing, cool
Steam cylinder oil	1500–3300	Lubricate in presence of steam at high temperatures

compounds serve the useful purpose of acting as natural antioxidants. The nitrogen compounds may be harmful in certain applications because of their propensity to form deposits on hot surfaces and are generally reduced to low concentrations.

The **table** lists typical specification data for lubricants in several applications. Viscosity is a determining factor in lubricant selection. Machines are generally designed to operate on the lowest practicable viscosity, since the lighter fluids give lower friction and better heat-transfer rates. However, if loads or temperatures are high, more viscous and less volatile lubricants are required. Change of viscosity with temperature is often of considerable practical importance and is customarily expressed in terms of viscosity index, an arbitrarily chosen scale on which an oil from a Pennsylvania crude, high in saturate and monoaromatic content, is assigned a value of 100, and those containing a relatively large amount of cyclohydrocarbons, both paraffinic and aromatic (from naphthenic crudes) are in the 0–50 viscosity-index range. As already described, the refiner can produce oils of high viscosity index from naphthenic stocks, but yields may be low.

Multigrade oils. In order to standardize on nomenclature for oils of differing viscosity, the Society of Automotive Engineers (SAE) has established viscosity ranges for the various SAE designations (see table). By the use of relatively large amounts of additives for improving viscosity index, it is possible to formulate one oil which will fall within the range of more than one SAE viscosity grade, the so-called multigrade oils, illustrated by **Fig. 1**. Since all viscosity-index improvers also increase viscosity, it is necessary to use a base oil of low viscosity to formulate such oils. Since the viscosity-increasing effect of the additive decreases with increase in shear rate, an artificially thickened oil of this type behaves as a light oil in engine parts under high shear, and thereby friction is low, but acts as a heavier oil in low-shear regions or at higher temperatures. However, volatility and ability to protect high shear parts from rubbing set limits to the use of light oils in these formulations.

Additives for lubricating oils. It is often desirable to add various chemicals to lubricating oils to improve their physical properties or to obtain some needed improvement in performance.

Viscosity-index improvers. The fall in viscosity with increase in temperature of oils of a given grade, or viscosity level, can be made less steep by thickening lighter oils with polymeric substances such as polybutenes and copolymers of polymethacrylates. The polymer may increase in solubility as the temperature increases, and correspondingly, the molecules uncoil and thicken the base oil more at high than at low temperatures and thus counteract, to some extent, the natural decrease in viscosity the base oil would undergo (Fig. 1).

Pour-point depressants. The dewaxing process removes the higher melting hydrocarbons, but some components remain which may solidify and gel the oil, and thus reduce its fluidity at low temperatures. Small amounts of chemicals, such as metallic soaps, condensation products of chlorinated wax and alkyl naphthalenes or phenols, polymethacrylates, and a host of others, increase fluidity at low temperatures. The mechanism of their action is still uncertain, but

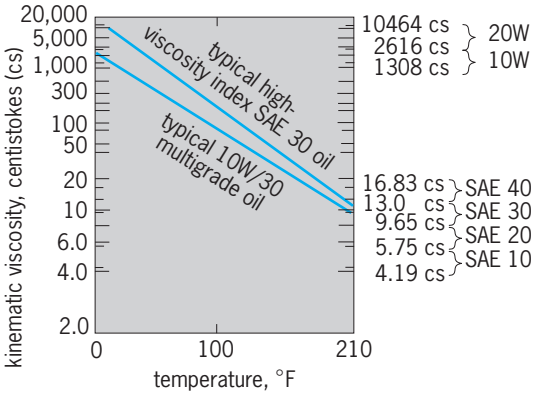

Fig. 1. Viscosity-temperature relationships indicated for additive-thickened multigrade oils. °C = (°F − 32)/1.8.

it probably involves adsorption of additive molecules on the surface of the growing wax crystals. The adsorbed layer either reduces intercrystal forces or modifies crystal growth.

Antioxidants. Lubricants are exposed to oxidation by atmospheric oxygen in practically all of their applications. This results in formation of acids and sludges which interfere with the primary function of the lubricant. Substantial increases in service life can be obtained by using small amounts (0.1–1.0%) of antioxidants. For lighter, more highly refined oils such as steam turbine oils, hindered phenols such as dibutyl-*p*-cresol are very effective. In heavy lubricants, such as engine oils, amines (phenyl-α-naphthylamine), metal phenates (alkali-earth salts of phenol disulfides), and zinc salts of thiophosphates and carbamates are used.

Antiwear and friction-reducing additives. If sliding surfaces can be completely separated by an oil film, friction and wear will be at a minimum. There are many systems, however, in which the combination of component geometry and operating conditions is such that a continuous oil film cannot be maintained. If no separating film of any kind were interposed, however, complete seizure would result. If the pressures and temperatures between such contacting surfaces are moderate, the provision of a boundary lubricating film will suffice, whereas if, as in some gears, conditions of both temperature and pressure are severe, some form of extreme pressure (EP) lubrication may be necessary.

The form in which wear manifests itself in machine components varies widely with the conditions, from the catastrophic welding together of gear-tooth surfaces to the slow continuous removal of material from an engine cylinder; even within the same mechanical system, it may change profoundly with changes in operating and environmental conditions.

In internal-combustion engines operating at low cylinder temperatures, for instance, condensation of acids from the gaseous combustion products on the cylinder and ring surfaces results in corrosive wear; in such cases, the addition of alkaline-earth phenates to the lubricating oil to neutralize the acids has succeeded in reducing wear rates.

Extreme pressure additives. Certain types of gears, particularly the hypoids used in automotive rear-axle transmissions, operate under such severe conditions of load and sliding speed, with resulting high temperature and pressure, that ordinary lubricants cannot provide complete protection against metal contact; this leads to welding, transfer of surface metal, and ultimate destruction of the gears. Also, in certain machining operations, it is necessary to prevent the chip from welding to the cutting tool. For such applications, lubricants containing sulfur and chlorine compounds are used. At the temperatures developed in the contact, these react chemically with the metal surfaces; the resulting sulfide and chloride films provide penetration-resistant, low-shear-strength films which prevent damage to the surfaces. Care in formulating these lubricants must, however, be exercised to ensure that corrosion of metal at normal temperatures does not occur. **Figure 2** shows the damage which may occur in the absence of protective additives.

Fig. 2. Two examples of gear damage.

Dispersants. In internal combustion engines, some of the products of combustion, which include carbonaceous particles, partially burnt fuel, sulfur acids, and water, enter the oil film on the cylinder walls. These materials may react to form lacquers and sludges which are deposited on the working parts of the engine. The lubricant is called upon to keep engine parts (oil screens, ring grooves) clean, and this is accomplished by the use of so-called detergent oils (**Fig. 3**).

The most commonly used dispersants are alkaline-earth salts. Of these, the most successful have been the salts of petroleum sulfonic acids, phenates, salicylates, thiophosphates, and oxidized olefinphosphorus pentasulfide reaction products. These materials owe their effectiveness partly to their surface activity, which ensures that foreign particles are kept in suspension in the lubricating oil, and partly to their alkalinity, which enables them to neutralize combustion acids that would otherwise catalyze the formation of lacquers.

Polymeric additives have been introduced for dispersancy. They are copolymers of a long-chain methacrylate, for example, lauryl methacrylate, and a nitrogen-containing olefin, such as vinyl pyrrolidone. These nonash dispersants are superior to the metal-containing additives for low-temperature operation and can be used very effectively in motor oils. Because they are less alkaline than metallic additives, the polymeric additives are less effective in diesel engines whose fuels often lead to combustion products which are much more acid than those from gasoline.

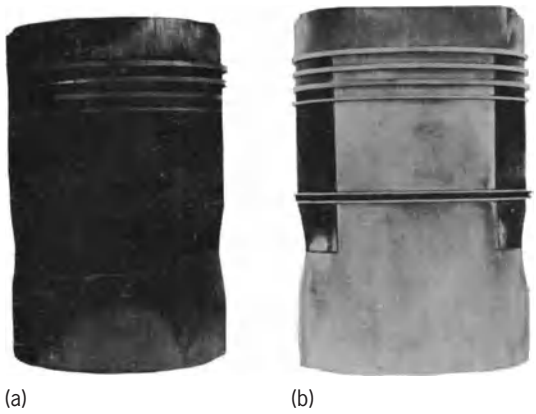

Fig. 3. Effect of dispersant on cleanliness. (*a*) Ordinary oil. (*b*) Oil containing dispersant additive.

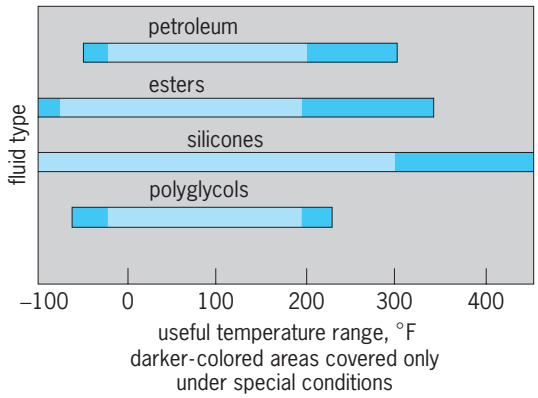

Fig. 4. Useful temperature range of synthetic oils. °C = (°F − 32)/1.8.

Boundary lubricants. Boundary conditions are encountered in many metal-forming processes in which the pressures required to deform the metal are too high to allow an oil film to form. In such applications, fatty oils, such as palm oil, or lubricants containing fatty materials are employed to reduce the friction and wear; the fatty acids react with the metal surface to form a tenacious soap film which provides lubrication up to temperatures near the melting point of the soap, usually about 250°F (120°C). *See* MACHINABILITY OF METALS.

Synthetic oils. During World War II, synthetic lubricants were extensively employed by Germany as substitutes for mineral lubricants which were in short supply. Some of these, such as the ester oils, were in fact superior to mineral lubricants in some applications. After the war, the use of synthetics for special applications, where their performance justifies the higher costs, steadily increased. Their main advantage is that they have a greater operating range than a mineral oil (**Fig. 4**). Esters such as 2-ethylhexyl sebacate, containing oxidation inhibitors and sometimes mild extreme pressure additives, are used as lubricants for aircraft jet engines. Silicones, although ideal with regard to thermal stability and viscosity characteristics, are poor lubricants for steel on steel, and this has restricted their use, although they are invaluable in some applications. Another class of widely used synthetic oils are the polyglycols, such as polypropylene and ethylene oxides. These polymers are available in a wide molecular-weight range, and they vary considerably in solubility in water and hydrocarbons. Thus, water-base lubricants may be formed which are used when a fire hazard exists.

Solid lubricants. The most useful solid lubricants are those with a layer structure in which the molecular platelets will readily slide over each other. Graphite, molybdenum disulfide, talc, and boron nitride possess this property.

The principal difficulty encountered with the use of solid lubricants is that of maintaining an adequate lubricant layer between the sliding metal surfaces. If the solid lubricant is applied as a suspension in a fluid, there is a tendency for the particles to settle out, and they may not reach the region where they are required. If they are applied as a thick paste to overcome the tendency to settle out, it is frequently difficult to force the paste through the narrow clearances between the sliding surfaces. A third method is to pretreat the surfaces with a relatively thick film of solid lubricant suspended in a resin and bake it on. Difficulties may then arise through progressively increasing clearances as the film wears, and there is no method for its continuous renewal. Materials such as graphite and molybdenum disulfide oxidize quite rapidly in air at 750°F (400°C). Unless air can be excluded, therefore, alternative solid lubricants will be required at temperatures greatly in excess of this.

A unique type of solid lubricant is provided by the plastic polytetrafluoroethylene (PTFE). At low speeds, PTFE slides on itself or on metals with a coefficient of friction in the lowest range observed for boundary lubrication (about 0.05). The friction rises to values common for other plastics after the sliding speed passes a critical value, and this high friction tends to persist when returning to low speeds.

It is likely that this is a thermal effect, the incidence of which is promoted by the extremely low thermal conductivity of PTFE, and it can be combated to a considerable extent by using the PTFE in a copper matrix. The use of such a matrix also reduces the wear which would occur with PTFE alone. This type of structure is used as a self-lubricated bearing material.

Greases. A lubricating grease is a solid or semifluid lubricant comprising a thickening (or gelling) agent in a liquid lubricant. Other ingredients imparting special properties may be included. An important property of a grease is its solid nature; it has a yield value. This enables grease to retain itself in a bearing assembly without the aid of expensive seals, to provide its own seal against the ingress of moisture and dirt, and to remain on vertical surfaces and protect against moisture corrosion, especially during shutdown periods.

The gelling agents used in most greases are fatty acid soaps of lithium, sodium, calcium, and aluminum. Potassium, barium, and lead soaps are used occasionally. The fatty acids employed are oleic, palmitic, stearic, and other carboxylic acids derived from animal, fish, and vegetable oils. The use of soaps in greases is limited by phase changes (or melting points) that occur in the range 120–390°F (50–200°C). Because many applications require operation above this temperature, greases have been introduced which are thickened by high-melting solids such as clays, silica, organic dyes, aromatic amides, and urea derivatives. In these cases, the upper limit of application is set by the volatility or by the oxidation and thermal stability of the base oil. To extend the operating range to extremes of high and low temperatures, synthetic oils may be used as the substrate. Silicones, esters, and fluorocarbons have all been employed for this purpose. As in the case of liquid lubricants, various additives such as oxidation inhibitors, EP agents, and antirust additives are frequently used in greases. Volatility of the base oil is

particularly important in greases, because it is constantly exposed as a thin film on the bearing surfaces.

Structure. The gelling agent in most soap-base greases is present as crystalline fibers having lengths of 1–100 micrometers and diameters $^1/_{10}$ to $^1/_{100}$ of their lengths. The fibers thicken the oil by forming a network or brush-heap structure in which the oil is held by capillary forces. For a given concentration of soap, the larger the length-to-diameter ratio of the fibers, the greater is the probability of network junction formation and the harder is the grease. However, not all soaps form such crystalline fibers; aluminum soap is a notable exception. There is evidence that aluminum soap molecules form a network akin to that formed by polymer molecules.

Nonsoap thickeners are present as small isometric particles much smaller than 1 μm in linear size. Silica particles are round, and some clay particles are plates. Forces of interaction between silica, clay particles, and oil are easily destroyed by water, so it is necessary to waterproof the primary particles by special treatments. Electron micrographs of various types of greases are shown in **Fig. 5**.

Mechanical properties. When the applied stress exceeds the yield stress, the grease flows and the viscosity falls rapidly with further increase of stress until it reaches a value only a little higher than that of the base oil. This fall in viscosity is largely reversible, since it is caused by the rupture of network junctions which, following the release of stress, can reform. However, continued vigorous shearing usually results in the rupture of fibers and permanent softening of the grease. Here, certain greases such as those derived from lithium salts of hydroxy acids are superior to other types, although the manner of grease preparation exercises considerable influence.

In a major grease application, for example, ball and roller bearings, shearing between the races and rolling elements is extremely severe. The reason that a good bearing grease does not soften and run out lies in the fact that only a very small fraction of the grease put in the bearing is actually subjected to shearing. As soon as a freshly packed bearing is set in motion, most of the grease is redistributed to places where it can remain static in the cover plate recesses and attached to the bore of the cage between the balls or rollers. The small amount of grease remaining on the working surfaces of the bearing is quickly broken down to a soft oily material which lubricates these parts, and because it has only a low viscosity, it develops very little friction. *See* ANTIFRICTION BEARING.

A number of methods exist for the measurement of yield stress and viscosity at various temperatures. One of the oldest of grease tests, still universally used, is a consistency (hardness) test in which the depth of penetration of a cone of standard weight and dimensions is determined. However, few of the many tests used are helpful in predicting the performance of a grease. Consequently, in development work, recourse is made to actual performance data and to the use of rigs that simulate field conditions. *See* GEL; PETROLEUM PROCESSING AND REFINING.

Robert G. Larsen

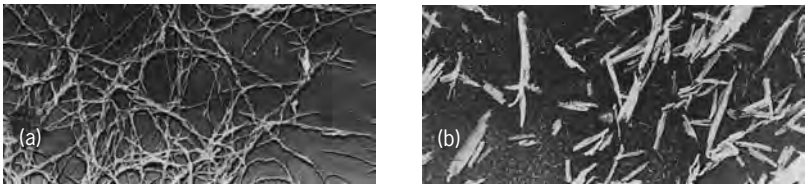

Fig. 5. Electron micrographs showing (*a*) soap-thickened grease and (*b*) dye-thickened grease.

Bibliography. J. Bruce, *Synthetic Functional Fluids and Lubricants*, 1994; Business Trend Analysts, *Lubricating Oils and Greases*, 1992.

Lubrication

The use of lubricants to reduce friction and wear. Whenever two bodies in contact are made to slide relative to one another, a resistance to the motion is experienced. This resistance, called friction, is present in all machinery. Approximately 30% of the power of an automobile engine is consumed by friction. Friction and wear can be significantly reduced, and thus relative motion of machine parts made possible, by interposing a lubricant at the interface of the contacting surfaces; the machine elements designed to accomplish this are called bearings. Bearings can be lubricated by solids such as graphite or, more commonly, by liquids and gases. *See* ANTIFRICTION BEARING; FRICTION; GRAPHITE; LUBRICANT; SURFACE AND INTERFACIAL CHEMISTRY; WEAR.

Conventionally, lubrication has been divided into (1) fluid-film lubrication (hydrostatic, hydrodynamic, and elastohydrodynamic), where the sliding surfaces are separated by a relatively thick, continuous film of lubricant; and (2) boundary lubrication, where contact surface separation is but a few molecular layers and asperity contact is unavoidable. Under conditions of boundary lubrication, solid lubricants are often used as films, or in particulate form as dispersions in oils, greases, and gases, to provide relatively stress-free surface conditions. Solid lubricants are also employed in very high temperature or very low pressure environments. The solid lubricants most commonly employed in particulate form are graphite and molybdenum disulfide (MoS_2), but polytetrafluoroethylene (Teflon) and lead compounds are also used. Another method of application of solid film lubricants is through bonded films; these are solid lubricants, such as molybdenum disulfide, graphite, Teflon, and tungsten disulfide, (WS_2), attached to the metal surface by means of a binder material such as a resin, a silicate, or a ceramic.

Lubricant films. When the lubricant film in a bearing is discontinuous, partial or boundary lubrication results. In this case, the coefficient of friction $f = F/N$, where F represents tangential (frictional) and N normal components of the contact force, depends on the characteristics of both the lubricant and the

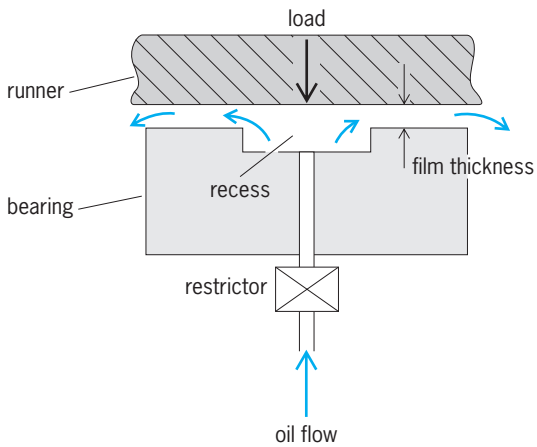

Fig. 1. Hydrostatic bearing pad.

contacting surfaces. The coefficient of friction can be analyzed by the methods of surface chemistry and physics. When the lubricant film is continuous, such as in the case of fluid film lubrication, frictional characteristics are defined solely by the material properties of the lubricant but not of the bounding surfaces. Continuous lubricant films can be generated in either of two modes, termed hydrostatic (externally pressurized) and hydrodynamic (self-acting).

Hydrostatic bearings. Hydrostatic films are created when a high-pressure lubricant is injected between opposing (parallel) surfaces (pad and runner), thereby separating them and preventing their coming into direct contact. Hydrostatic bearings require external pressurization. The film is 5–50 micrometers thick, depending on application. Though hydrostatic lubrication does not rely on relative motion of the surfaces, relative motion is permitted and can even be discontinuous. **Figure 1** is a schematic of a hydrostatic bearing pad. To handle asymmetric loads, hydrostatic systems generally employ several evenly spaced pads. The pads are usually supplied from a common reservoir and are equipped with flow restrictors. Flow restrictors are flow control devices, such as orifices and capillaries, across which minute changes in flow rate cause large changes in pressure drop. Thus, if variation in the load vector causes even a small variation in film thickness over a particular pad, a large change in the recess pressure of that pad will result, yielding a righting moment on the runner. Hydrostatic bearings find application where relative positioning is of extreme importance; machine tools using this type of bearing can grind parts round within 5 μm, with a 2.5-μm surface finish.

Hydrostatic bearings are also applied where a low coefficient of friction ($\sim 10^{-6}$ to 10^{-4}) at vanishing relative velocity is required. The 200-in. (5-m) optical telescope at Mount Palomar, which weighs 500 tons (454 metric tons), is supported on hydrostatic pads and kept synchronized with the aid of a $^1/_4$-hp (190-W) clock motor. Hydrostatic lifts are often built into the hydrodynamic bearings of a large rotating apparatus to facilitate starting and stopping, when a continuous hydrodynamic film is absent. The advantages of hydrostatic bearings are low friction, exact relative positioning of the bearing surfaces, and insensitivity to magnitude and continuity of relative motion. Among their disadvantages is the need for auxiliary equipment, the high risk of failure in the absence of a back-up system, and the high capital cost. Because of these disadvantages, hydrostatic bearings are used less frequently than are hydrodynamic bearings.

Hydrodynamic bearings. Hydrodynamic bearings are self-acting. To create and maintain a load-carrying hydrodynamic film, it is necessary only that the bearing surfaces move relative to one another and ample lubricant is available. The surfaces must be inclined to form a clearance space in the shape of a wedge, which converges in the direction of relative motion. The lubricant film is then created as the lubricant is dragged into the clearance by the relative motion. This viscous action results in a pressure build-up within the film (**Fig. 2**). The fact that hydrodynamic bearings are self-generating and do not rely on auxiliary equipment makes these bearings very reliable. Hydrodynamic lubrication can be found to operate even when highly undesirable; on the rain-covered pavement, an automobile tire rides on a hydrodynamic film, thereby reducing friction between road and tire by orders of magnitude compared to dry conditions.

Hydrodynamic journal bearings (**Fig. 3**) and thrust bearings (**Fig. 4**) are designed to support radial and axial loads, respectively, on a rotating shaft. The film thickness (h) in these bearings is in the $\sim$10–200-μm range, and the coefficient of friction $\sim 10^{-2}$ to 10^{-3}. Because of their relatively large contact area, the pressures generated in hydrodynamic bearings are low, in the range of 1–10 megapascals (10–100 atm). They operate with the bearing surfaces remaining rigid, although at higher speeds and loads both thermal and elastic deformations can be considerable. Hydrodynamic bearings are usually lubricated with liquids (such as mineral oil), but for light loads and high speeds, as in sophisticated navigational equipment, gas lubrication is preferred.

Intense development of gas lubrication technology was spurred, beginning in the 1950s, by the demands of sophisticated navigation systems, by the operation of gas-cooled nuclear energy sources, and by the quest for exotic machinery in aerospace applications. The magnetic recording industry has

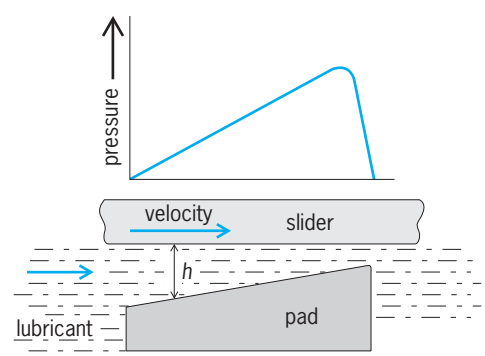

Fig. 2. Hydrodynamic film formation.

further emphasized the importance of this technology; the pursuit of ever-increasing densities of information storage has led to the design of disk drives with slider bearings, or "sliders," that carry the read/write head, allowing it to skim over disk surfaces at minuscule distances, a fraction of the mean free path (the average distance traveled by gas molecules between collisions, about 60 nm for air). *See* COMPUTER STORAGE TECHNOLOGY; MAGNETIC RECORDING.

Shaped bearings. Journal and thrust bearings are conformal bearings; that is the opposing bearing surfaces conform in shape. Ball and roller bearings, also known as rolling contact bearings, are counterformal. Counterformal bearings always operate in the hydrodynamic mode, but because the contact area in these bearings is small the pressure attains high values, in the range of 1–3 gigapascals (10,000–30,000 atm). In consequence, the surfaces deform elastically and the lubricant viscosity increases by several orders of magnitude—the variation of lubricant viscosity, μ, with pressure, P, approximately follows the law $\mu = \mu_0 \exp(\alpha p)$, where μ_0 represents viscosity at atmospheric pressure and α is a material constant. Conformal bearings are said to operate in the elastohydrodynamic (EHD) mode. The schematic of a ball bearing is shown in the **Figure 5**. Rolling contact bearings can be designed to support either radial or axial loads or any combination thereof. The greatest disadvantage of rolling contact bearings is that they posses little or no damping, necessary to dissipate the vibrational energy when approaching system resonance. The friction coefficient of rolling contact bearings is about the same as journal bearings, but the film thickness is two orders or magnitude smaller; compared to this film thickness, the roughness of the bearing surfaces can no longer be neglected, unless the surfaces are made particularly smooth. *See* VIBRATION.

Lubricants. Lubricants that are liquids at normal temperatures played a minor role until the industrial revolution; before then greases and pastes of animal and vegetable origin were the principal lubricants. Today, mineral oils manufactured from petroleum are the most common liquid lubricants. The manufacturer of petroleum lubricants can choose from a wide variety of crude oils, and the choice is of great importance because the lubricating oil fraction of crude oils varies widely. Petroleum oils are excellent lubricants because they wet metal surfaces, are available in a wide range of viscosities, and have viscosity that increases with pressure faster than that of almost all other fluids (of great importance in rolling contact bearings). The next important property of petroleum oils is their ability to resist oxidation. For most applications of petroleum oil lubricants, however, certain materials must be added to provide a new and desirable property that was not originally present, to enhance a desirable property already possessed in some degree, or to overcome some natural dificiency. Additives are used to reduce thermal and oxidative degradation, lessen the accumulation of harmful deposits, change the viscosity characteristics, minimize rust and corrosion, control frictional behavior, reduce wear, prevent bacterial growth, prevent destructive metal to metal contact, and control foaming. *See* PETROLEUM; VISCOSITY.

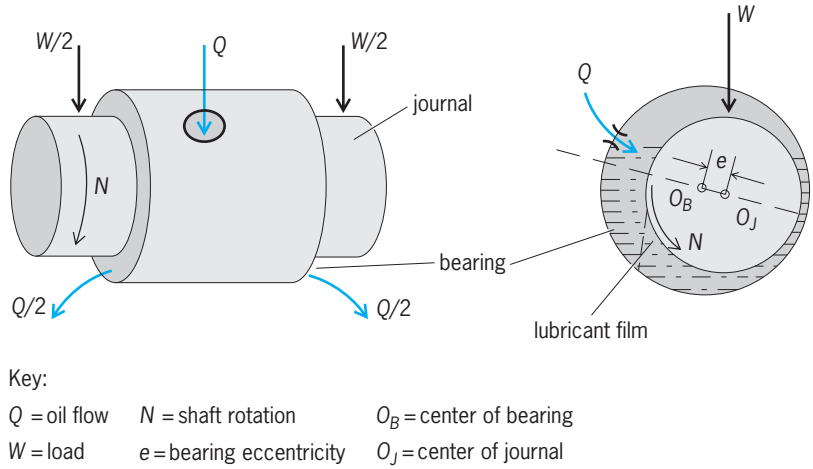

Fig. 3. Journal bearing.

Bearing materials. The selection of the bearing material for a particular application depends on the type of bearing, the type of lubricant, and environmental conditions. For hydrodynamic bearings, babbitts are among the most widely used materials. They are either tin- or lead-base alloys, having excellent embeddability (the ability to embed dirt or other foreign material so that damage to the shaft is minimized) and conformability (to misalignment)

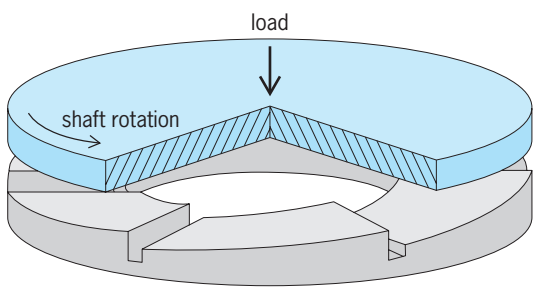

Fig. 4. Thrust bearing.

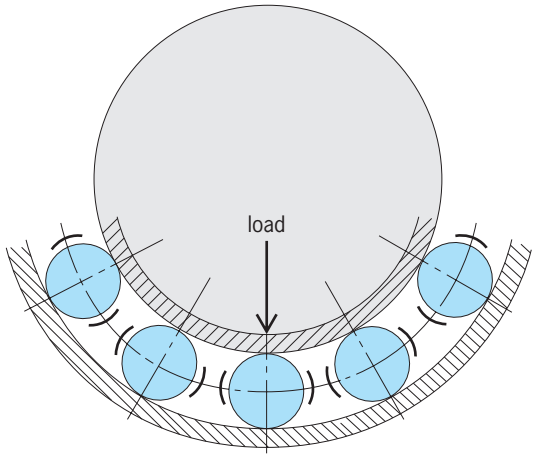

Fig. 5. Ball bearing.

characteristics. Babbitts, however, have relatively low load-carrying capacity. To increase their load capability, these alloys are metallurgically bonded to stronger backing materials such as steel, cast iron, and bronze. Rolling contact bearings operate under conditions that impose high compressive stress for millions of stress cycles as the balls or rollers rotate through the loaded zone of the bearing. For such applications, the materials of the races and balls should be hard and have high fatigue resistance. The most common material used is hardened chrome steel. In high-temperature applications, high-speed tool steels with better hardness retention are used. In any high-temperature application, it is normally the thermal stability of lubricants, rather than the bearing materials, that limits the maximum operating temperature. *See* ALLOY; BRONZE; IRON; LEAD; METALLURGY; STEEL; TIN. Andras Z. Szeri

Bibliography. K. E. Bannister, *Lubrication for Industry*, Industrial Press, 1996; B. Bhuushan (ed.), *Modern Tribology Handbook*, CRC Press, 2000; B. Bhushan, *Principles and Applications of Tribology*, Wiley, 1999; B. J. Hamrock, *Fundamentals of Fluid Film Lubrication*, McGraw-Hill, 1993; K. C. Ludema, *Friction, Wear, Lubrication: A Textbook in Tribology*, CRC Press, 1996; M. J. Neale (ed.), *The Tribology Handbook*, 2d ed., Butterworth-Heinemann, 1995; B. N. J. Persson, *Sliding Friction: Physical Principles and Applications (Nanoscience and Technology)*, 2d ed., Springer-Verlag, 2000: A. Z. Szeri, *Fluid Film Lubrication: Theory & Design*, Cambridge University Press, 1998.

Lumber

Timber sawed or split into planks, boards, and similar products. Lumber can come in many forms, species, and types from a wide variety of commercial sources. Because most lumber is manufactured similarly and graded by standardized rules, it is fairly uniform throughout the United States. This article describes the manufacturing, processing, and grading of commercial lumber and the effects of moisture content. *See* WOOD PRODUCTS.

Manufacturing and processing. Lumber is manufactured from round logs primarily in rectangular shapes of different dimensions. Lumber length is recorded in actual dimensions. Width and thickness are traditionally recorded in nominal dimensions, which are somewhat more than actual dimensions. Softwood lumber is manufactured in length multiples of 305 mm (1 ft) as specified in various grading rules. In practice, 610-mm (2-ft) multiples are most common. Lumber width commonly varies from 38 to 387 mm (nominal 2 to 16 in.). Lumber is classified by thickness into three categories: (1) board, lumber less than 38 mm (nominally 2 in.) thick; (2) dimension, lumber from 38 mm to, but not including, 114 mm (nominally 5 in.) thick; and (3) timber, lumber 114 mm (nominally 5 in.) or more in thickness in the least dimension. *See* LOGGING.

Lumber can be produced with either a rough or surfaced (dressed) finish. Rough-sawn lumber has surface imperfections caused by the primary sawing operations. Surfaced lumber is smoothed on either one or both sides and one or both edges.

Grades. In general, the grade of a piece of lumber is based on the number, character, and location of features that may lower the strength, durability, or utility value of the wood. Lumber grading can be divided into two main categories: remanufacture "shop grade" and structural "stress grade."

Shop grade. Sorting of lumber for remanufacture is based on visual inspection. The wood is designated shop grade on the proportion of defect-free or clear cuttings of a certain size that can be made from a piece of lumber. The larger volume and more frequent number of clear cuttings, the higher the grade. The rules adopted by the National Hardwood Lumber Association are considered standard in grading hardwood lumber intended for cutting into smaller pieces to make furniture or other fabricated products. Some shop grades commonly used for this material are Select, No. 1 Common, and No. 2A Common.

Stress grade. Pieces of lumber graded for structural uses are put into classes with similar mechanical properties called stress grades. Stress grades are characterized by (1) one or more sorting criteria, (2) a set of allowable properties for engineering design, and (3) a unique grade name. The allowable properties depend on the particular sorting criteria and additional factors that are independent of sorting criteria. Allowable properties are usually much lower than the properties of clear, straight-grained wood. The allowable properties are inferred through visual grading criteria or are determined nondestructively by machine-grading criteria.

Visually graded. Visual grading is the oldest stress-grading method. It is based on the premise that mechanical properties of lumber differ from mechanical properties of clear wood. Growth characteristics, which affect properties and can be seen and judged by eye, are used to sort the lumber into stress grades. Typical visual sorting criteria include density, decay, proportion of heartwood and sapwood, slope of grain, knots, shake, checks and splits, wane, and pitch pockets. Some grades commonly used for this material are Select Structural, No. 1, No. 2, and Construction. Two species with the same grade look alike but may have different individual properties.

Machine graded. Machine-graded lumber is evaluated by a machine using a nondestructive test followed by visual grading to evaluate certain characteristics that the machine cannot or may not properly evaluate. Machine-stress-rated (MSR), machine-evaluated (MEL), and E-rated lumber are three types of machine-graded lumber. Machine-graded lumber allows for better sorting of material for specific applications in engineered structures. The basic components of a machine-grading system are as follows: (1) sorting and prediction of strength through machine-measured nondestructive testing coupled with visual assessment of growth characteristics; (2) assignment of design properties based on

strength prediction; and (3) quality control to ensure that assigned properties are being obtained. The MSR and MEL systems differ in grade names, quality control, and coefficient of variation (COV) for modulus of elasticity (E) values. Grade names for MSR lumber are a combination of the design bending stress and average modulus of elasticity (such as 1650F 1.5E), whereas grade names for MEL lumber start with an M designation (such as M-14). Machine grading is species-independent because grading is based on properties.

Moisture content. Clear, straight-grained lumber can be about 50% stronger when dry than when wet. For lumber containing knots, the increase in strength with decreasing moisture content is dependent on lumber quality. As the frequency and size of knots increase, the reduction in strength resulting from the knots begins to negate the increase in strength in the clear wood portion of the lumber. Very low quality lumber, which has many large knots, may be insensitive to changes in moisture content. Some strength values may decrease with decreasing moisture content (less than about 8%), and care should be exercised in these situations. For timber, often no adjustment for moisture content is made because properties are assigned on the basis of wood in the green condition. Timber is usually put into service without drying, and it is assumed that drying degrade offsets the increase in strength normally associated with loss in moisture. *See* WOOD PRODUCTS; WOOD PROPERTIES. David E. Kretschmann

Bibliography. *Rules for the Measurement and Inspection of Hardwood and Cypress*, National Hardwood Lumber Association, Memphis, TN, 1990; *Wood Handbook: Wood as an Engineering Material*, Gen. Tech. Rep. FPL-GTR-113, U.S. Department of Agriculture, Forest Service, Forest Products Laboratory, Madison, WI, 1999.

Luminance

The luminous intensity of any surface in a given direction per unit of projected area of the surface viewed from that direction. The International Commission on Illumination defines it as the quotient of the luminous intensity in the given direction of an infinitesimal element of the surface containing the point under consideration, by the orthogonally projected area of the element on a plane perpendicular to the given direction. Simply, it is the luminous intensity per unit area. Luminance is also called photometric brightness.

Since the candela is the unit of luminous intensity, the luminance, or photometric brightness, of a surface may be expressed in candelas/cm^2, candelas/in.^{2}, and so forth.

Mathematically, luminance L may be found from the equation below, where θ is the angle between

$$L = \frac{dI}{dA\ \cos\theta}$$

the line of sight and the normal to the surface area A considered, and I is the luminous intensity.

The stilb is a unit of luminance (photometric brightness) equal to 1 candela/cm^2. It is often used in Europe, but the practice in America is to use the term candela/cm^2 in its place.

The apostilb is another unit of luminance sometimes used in Europe. It is equal to the luminance of a perfectly diffusing surface emitting or diffusing light at the rate of 1 lumen/m^2. *See* LUMINOUS INTENSITY; PHOTOMETRY. Russell C. Putnam

Luminescence

Light emission that cannot be attributed merely to the temperature of the emitting body. Various types of luminescence are often distinguished according to the source of the energy which excites the emission. When the light energy emitted results from a chemical reaction, such as in the slow oxidation of phosphorus at ordinary temperatures, the emission is called chemiluminescence. When the luminescent chemical reaction occurs in a living system, such as in the glow of the firefly, the emission is called bioluminescence. In the foregoing two examples part of the energy of a chemical reaction is converted into light. There are also types of luminescence that are initiated by the flow of some form of energy into the body from the outside. According to the source of the exciting energy, these luminescences are designated as cathodoluminescence if the energy comes from electron bombardment; radioluminescence or roentgenoluminescence if the energy comes from x-rays or from gamma rays; photoluminescence if the energy comes from ultraviolet, visible, or infrared radiation; and electroluminescence if the energy comes from the application of an electric field. By attaching a suitable prefix to the word luminescence, similar designations may be coined to characterize luminescence excited by other agents. Since a given substance can frequently be made to luminesce by a number of different external exciting agents, and since the atomic and electronic phenomena that cause luminescence are basically the same regardless of the mode of excitation, the classification of luminescence phenomena into the foregoing categories is essentially only a matter of convenience, not of fundamental distinction.

When a luminescent system provided with a special configuration is excited, or "pumped," with sufficient intensity of excitation to cause an excess of excited atoms over unexcited atoms (a so-called population inversion), it can produce laser action. (Laser is an acronym for light amplification by stimulated emission of radiation.) This laser emission is a coherent stimulated luminescence, in contrast to the incoherent spontaneous emission from most luminescent systems as they are ordinarily excited and used. *See* LASER; OPTICAL PUMPING.

Fluorescence and phosphorescence. A second basis frequently used for characterizing luminescence is its persistence after the source of exciting energy is removed. Many substances continue to luminesce

for extended periods after the exciting energy is shut off. The delayed light emission (afterglow) is generally called phosphorescence; the light emitted during the period of excitation is generally called fluorescence. In an exact sense, this classification, based on persistence of the afterglow, is not meaningful because it depends on the properties of the detector used to observe the luminescence. With appropriate instruments one can detect afterglows lasting on the order of a few thousandths of a microsecond, which would be imperceptible to the human eye. The characterization of such a luminescence, based on its persistence, as either fluorescence or phosphorescence would therefore depend upon whether the observation was made by eye or by instrumental means. These terms are nevertheless commonly used in the approximate sense defined here, and are convenient for many practical purposes. However, they can be given a more precise meaning. For example, fluorescence may be defined as a luminescence emission having an afterglow duration which is temperature-independent, while phosphorescence may be defined as a luminescence with an afterglow duration which becomes shorter with increasing temperature. *See* BIOLUMINESCENCE; CATHODOLUMINESCENCE; CHEMILUMINESCENCE; ELECTROLUMINESCENCE; FLUORESCENCE; PHOSPHORESCENCE; PHOTOLUMINESCENCE; THERMOLUMINESCENCE.

Type of radiation emitted. Because of their many practical applications, materials that give a visible luminescence have been studied and developed more intensively than those which emit in other spectral regions. Luminescence, however, may consist of radiation in any region of the electromagnetic spectrum. The production of x-radiation by the bombardment of a metal target by a fast electron beam is an example of luminescence. Certain fluorescent lamps, called black-light lamps, are coated with a luminescent powder chosen for its ability to emit ultraviolet light of approximately 360 nanometers rather than visible light. A number of luminescent solids have been developed which luminesce in the near infrared under excitation by visible light, by a cathode-ray beam, or by electric fields. *See* ULTRAVIOLET LAMP.

Luminescent substances. The ability to luminesce is not confined to any particular state of matter. Luminescence is observed in gases, liquids, and amorphous and crystalline solids. The passage of an electrical discharge through a gas will excite the gaseous atoms or molecules to luminesce under certain conditions. An example of such a gaseous luminescence is the mercury-vapor lamp, in which an electrical discharge excites the mercury vapor to emit both visible and ultraviolet light. Many liquids, such as oil or solutions of certain dyestuffs in various solvents, luminesce very strongly under ultraviolet light. A large number of solids, including many natural minerals as well as thousands of synthetic inorganic and organic compounds, luminesce under various types of excitation. Applications of gas luminescence, formerly confined to advertising signs (neon signs) and fluorescent lamps, have multiplied greatly with the invention of the laser, many atomic and molecular gases providing the laser-active materials. The same is true of dye solutions and inorganic glasses. The major nonlaser applications of luminescence involve solid luminescent materials. *See* MERCURY-VAPOR LAMP.

The term phosphor, originally applied to certain solids that exhibit long afterglows, has been extended to include any luminescent solid regardless of its afterglow properties. Other terms sometimes used synonymously with phosphor are luminophor, fluor, or fluorphor. The term luminophor is preferable, since it carries no connotation that afterglow times are long or short. In conformity with current usage, however, the term phosphor is used in the succeeding discussion of solid materials.

Comparatively few pure solids luminesce efficiently, at least at normal temperatures. In the category of organic solids, the pure aromatic hydrocarbons, such as naphthalene and anthracene, which consist exclusively of condensed phenyl rings, are luminescent, as are many heterocyclic closed-ring compounds. However, the closed phenyl ring structure is not in itself a guarantee of efficient luminescence, since certain substituents for hydrogen in the structure, particularly halogens, tend to reduce luminescence efficiency. This quenching of efficiency is called internal conversion. Among the pure inorganic solids that luminesce efficiently at room temperature are the tungstates, uranyl salts, platinocyanids, and a number of salts of the rare-earth elements.

Activators and poisons. The development of a large number of inorganic phosphors is due to the discovery that certain impurities, called activators, when present in amounts ranging from a few parts per million to several percent, can confer luminescent properties on the compounds (host or matrix compounds) in which they are incorporated. The activator and its nearby atoms are often referred to as the luminescent center.

By the same token, small amounts of other impurities or imperfections, called poisons or quenchers, can inhibit or destroy the luminescence, evidently by providing alternative mechanisms for the radiationless dissipation of the energy imparted to the phosphor or by emitting radiation characteristic of the poison itself which may not be in a spectral region of interest for the purpose at hand.

Manganese is a particularly effective activator in a wide variety of matrices when incorporated in amounts ranging from a small trace up to the order of several percent. The emission spectrum of these manganese-activated phosphors generally lies in the green, yellow, or orange spectral regions. Other frequently used activators are copper, silver, thallium, lead, cerium, chromium, titanium, antimony, tin, and the rare-earth elements. Poisons in inorganic phosphors generally are from the iron-nickel-cobalt group. The most important host materials are silicates, phosphates, aluminates, sulfides, selenides, the alkali halides, and oxides of calcium, magnesium, barium, and zinc. Preparation of phosphors

and the incorporation of the activators is generally done by high-temperature reaction of well-mixed, finely ground powders of the components. A number of semiconducting compounds such as gallium phosphide (GaP) and gallium arsenide (GaAs) can be made to luminesce and display laser action when prepared so that charges can be injected into the semiconductor. *See* JUNCTION DIODE; SEMICONDUCTOR; SEMICONDUCTOR DIODE.

In order for ultraviolet light to provoke luminescence in a substance, the substance must be able to absorb it. Because of variations in its ability to absorb ultraviolet light of different wavelengths, a material may be nonluminescent under ultraviolet light of one wavelength and strongly luminescent under a different wavelength in the ultraviolet region. For a similar reason a substance may not luminesce under ultraviolet light at all and yet be strongly luminescent under x-ray irradiation.

Luminescence in Atomic Gases

The processes that occur in luminescence may be most simply illustrated in the case of an atom in a gas. The atom can exist only in certain specific states of energy, some of which are shown schematically in **Fig. 1**. The lowest energy level of the atom, E_1, corresponds to the atom in its unexcited, or ground, state, and the higher energy levels, E_2 and E_3, represent electronically excited states of the atom. The excitation of the atom from state E_1 to state E_2 requires the absorption of an amount of energy $\Delta E = E_2 - E_1$. If this excitation is to be produced by the absorption of light, the energy of the acting light photon, E_{absorbed}, must equal ΔE. The frequency of the exciting light must therefore be $\upsilon = E_{\text{absorbed}}/h = (E_2 - E_1)/h$, and the wavelength of the exciting light must be $\lambda_{\text{absorbed}} = hc/(E_2 - E_1)$, where h is Planck's constant and c is the velocity of light. In an isolated atom this extra energy cannot be dissipated and is emitted as radiation when the atom eventually returns to its ground state. The emitted light will therefore again correspond in energy to ΔE and it will have a wavelength $\lambda_{\text{emitted}} = \lambda_{\text{absorbed}}$. When $\lambda_{\text{emitted}} = \lambda_{\text{absorbed}}$, the emitted light is sometimes referred to as resonance luminescence or resonance radiation. *See* ATOMIC STRUCTURE AND SPECTRA.

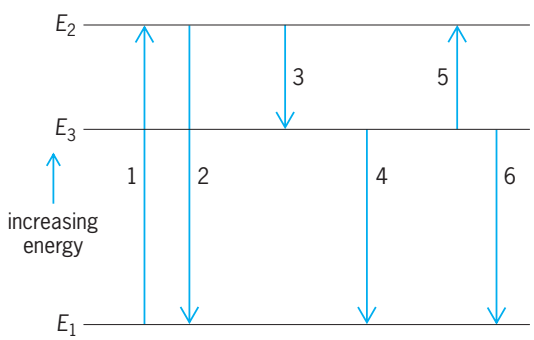

Fig. 1. Schematic representation of energy levels and electronic transitions in an atomic gas. E_1, ground state; E_2 and E_3, excited states; 1, excitation; 2, emission of resonance luminescence; 3, radiationless transition to lower excited state; 4, luminescence emission, if transition $E_3 \rightarrow E_1$ is allowed (if it is not allowed, 4 does not occur and E_3 is called a metastable state); 5, stimulation of atom back to emitting state; 6, radiationless transition of atom back to ground state (quenching).

If a large number of atoms are excited and the excitation then removed, the luminescence intensity will decrease with time exponentially according to the equation $I_t = I_0 e^{-t/\tau_l}$, where I_t is the intensity of luminescence at a time t after removal of the excitation, I_0 is the intensity at $t = 0$, and τ_l is the average time required by an atom to make a spontaneous luminescent transition. The quantity τ_l, called the radiative lifetime, is independent of temperature, and it is this temperature independence that is emphasized in the more precise definition of fluorescence given earlier. If the transition between the energy levels E_1 and E_2 is highly probable (a so-called permitted or allowed transition), τ_l is very small, of the order of 10^{-8} s for transitions involving visible light.

The excited atom can also lose a certain amount of energy and fall to an energy state E_3, of intermediate energy between E_1 and E_2. This can happen, for example, if the excited atom collides with another atom. If the transition from state E_3 to the ground state E_1 can occur with a reasonably high probability, fluorescence will occur starting from this intermediate excited state. In this case the fluorescent wavelength $\lambda_{\text{emitted}} = hc/(E_3 - E_1)$. Since $(E_3 - E_1)$ is smaller than $(E_2 - E_1)$, the fluorescence in this case will be of longer wavelength than the resonance radiation. Although the quantum efficiency of luminescence is unity (one photon of emitted light per photon of absorbed light), the energy efficiency is less than unity.

If, however, a transition between state E_3 and state E_1 is highly improbable (a so-called forbidden transition), state E_3 is known as a metastable state. The atom can remain in this state for long periods of time and cannot return to the ground state with the emission of radiation. Luminescence can occur under these circumstances only if the atom regains the energy $(E_2 - E_3)$ by a collision with another atom or by some other process. Once the atom has been brought back to state E_2, a transition to the ground state is again allowed, and luminescence corresponding to $(E_2 - E_1)$ will be emitted. The existence of metastable states like E_3 explains the delayed emission termed phosphorescence. The atom may spend a considerable amount of time in such a state before some external influence causes it to return to an emitting state such as E_2, in which case the luminescence is correspondingly delayed and appears as an afterglow. In order for the atom to get from E_3 to E_2 it must absorb energy somehow. The rate of return to the emitting state, and hence the duration of the afterglow, will therefore depend to a very large extent on temperature. At high temperatures the atoms will be excited back to the emitting state rapidly, and there will be a bright afterglow of short duration. At lower temperatures the atoms will be raised back to the emitting state very slowly, and the afterglow will be of long duration but of low intensity. This temperature dependence is the basis for the more precise definition of phosphorescence that was given earlier.

As an alternative to regaining the energy ($E_2 - E_3$), the atom may lose the energy ($E_3 - E_1$) by a competing radiationless process, for example, by collision with another atom. The energy of excitation is thus dissipated without luminescence, and the quantum yield of luminescence is zero (quenching).

The principles illustrated in the foregoing discussion may be extended with little modification to the case where the primary absorption or excitation act completely removes an electron from an atom (that is, ionizes the atom) instead of merely raising the atom to an excited state. Under these conditions the electron can be trapped temporarily by other atoms, and its return to the parent atom can also be delayed by this mechanism.

The same principles are also operative in the case of complex configurations of atoms, such as in organic molecules or solids. However, the forbiddenness of radiative transitions can be modified in these cases, and efficient luminescent emissions can consequently be observed due to radiative transitions from metastable states, albeit with longer afterglows.

When nonlinear optics conditions exist, luminescence can be excited in certain systems by multiphoton absorption, the resulting luminescence being of higher frequency than the exciting light because the energy of two or more photons of the latter are combined to give one photon of emitted light. There are also cases where a single photon is absorbed by a pair of activator atoms or ions, each of which then emits a luminescence photon, leading to a quantum efficiency of 2. However, the preceding discussion and the exposition of those principles which follow deal primarily with the more usual case of single-photon absorption and emission. *See* NONLINEAR OPTICS.

Configuration Coordinate Curve Model

Luminescence in atomic gases is adequately described by the concepts of atomic spectroscopy, but luminescence in molecular gases, in liquids, and in solids introduces two major new effects which need special explanation. One is that the emission band appears on the long wavelength (low-energy) side of the absorption band; the other is that emission and absorption often show as bands tens of nanometers wide instead of as the lines found in atomic gases.

Both of these effects may be explained by using the concept of configuration coordinate curves illustrated in **Fig. 2**. As in the case of atomic gases, the ground and excited states represent different electronic states of the luminescent center, that is, the region containing the atoms, or electrons, or both, involved in the luminescent transition. On these curves the energy of the ground and excited state is shown to vary parabolically as some configuration coordinate, usually the distance from the activator to its nearest neighbors, is changed. There is a value of the coordinate for which the energy is a minimum, but this value is different for the ground and excited states because of the different interactions of the activator with its neighbors. Absorption of light gives rise to the transition from A to B. This transition occurs so rapidly that the ions in the luminescent center do not have time to rearrange. Once the system is at B it gives up heat energy to its surroundings by means of lattice vibrations and reaches the new equilibrium position at C. Emission occurs when the system makes the transition from C to D, and once again heat energy is given up when the system goes from D back down to A. This loss of energy in the form of heat causes the energy associated with the emission $C \rightarrow D$ to be less than that associated with the absorption $A \rightarrow B$. *See* FRANCK-CONDON PRINCIPLE; LATTICE VIBRATIONS.

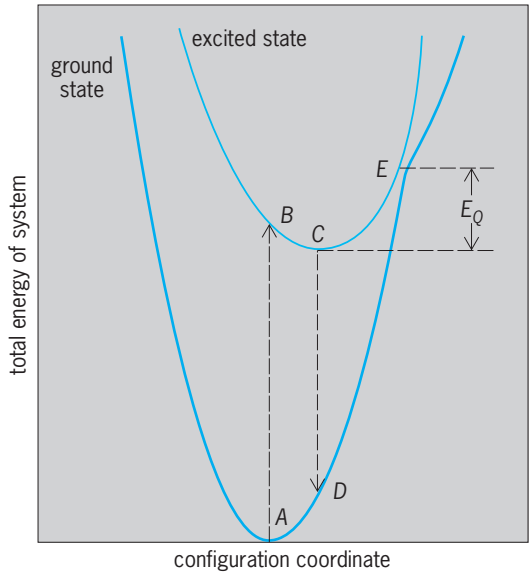

Fig. 2. Configuration coordinate curves for a simple luminescence center. Transition $A \rightarrow B$ shows absorption, and transition $C \rightarrow D$ shows emission. Energy E_Q is that necessary for the excited state to reach point E, from which a transition to the ground state can be made without luminescence.

When the system is at an equilibrium position, such as C of the excited-state curve, it is not at rest but migrates over a small region around C because of the thermal energy of the system. At higher temperatures these fluctuations cover a wider range of the configuration coordinate. As a result, the emission transition is not just to point D on the ground state curve but covers a region around D. In the vicinity of D the ground state curve shows a rapid change of energy, so that even a small range of values for the configuration coordinate leads to a large range of energies in the optical transition. This explains the broad emission and absorption bands that are observed. An analysis of this sort predicts that the widths of the band (usually measured in energy units between the points at which the emission or absorption is half its maximum value) should vary as the square root of the temperature. For many systems this relationship is valid for temperatures near and above room temperature. At low temperatures, quantum-mechanical effects, described below, become dominant.

Two other phenomena which can be explained on the basis of the model described in Fig. 2 are temperature quenching of luminescence and the variation of the decay time of luminescence

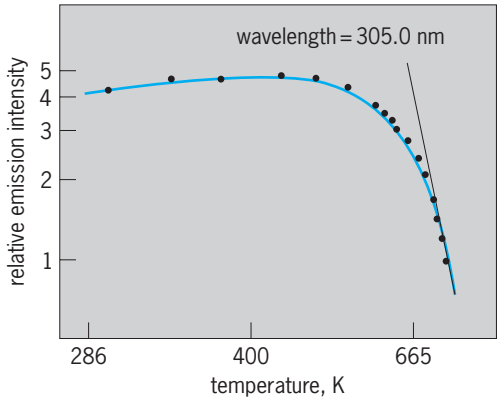

Fig. 3. Variation of the brightness of a thallium-activated potassium chloride phosphor with temperature. °F = (K × 1.8) − 459.67.

with temperature. In **Fig. 3** a curve of temperature quenching for the emission in thallium-activated potassium chloride is shown. At low temperatures there is very little change in brightness with temperature, but at elevated temperatures the luminescence efficiency decreases rapidly (so-called thermal quenching). On the scheme of Fig. 2 this is interpreted as meaning that the thermal vibrations become sufficiently intense to raise the system to point E. From point E the system can fall to the ground state by emitting a small amount of heat, or infrared radiation. If point E is at an energy E_Q above the minimum of the excited state curve, it may be shown that the quantum efficiency, η, of luminescence is given by Eq. (1), where C is a constant, k is Boltzmann's constant, and T is the temperature on the Kelvin scale. By fitting an expression of this form to the data of Fig. 3, a value of 0.60 eV is obtained for E_Q.

$$\eta = 1 + C \exp\left(-\frac{E_Q}{kT}\right)^{-1} \tag{1}$$

Another result of the onset of thermal quenching is that the luminescence decays faster, since the excited state is now depopulated by two processes simultaneously—a dissipative process in parallel with the luminescent process.

Quantum-mechanical corrections. Although the configuration coordinate curve model of Fig. 2 is successful in describing many aspects of luminescence in solids, it predicts that emission and absorption bands should become narrow lines as the temperature is reduced to absolute zero (0 K). This is not the case, as is shown in **Fig. 4**, which gives the width of the absorption band of the F-center in potassium chloride as a function of the square root of the temperature. (An F-center is the simplest type of color center, a color center being a lattice defect which can absorb light.) At high temperatures the previously quoted results are valid, but at low temperatures the width of the band is constant. *See* COLOR CENTERS.

This phenomenon may be explained by treating the configuration coordinate curves quantum mechanically. The curves have the energy versus displacement characteristics of the simple harmonic oscillator. For this case the quantum-mechanical analysis shows that there is a series of equally spaced energy levels separated by an energy of $h\nu$, where h is Planck's constant and ν is the frequency of vibration. The lowest of these levels occurs at a value of $^1/_2\, h\nu$ above the minimum of the classical curve, and this energy is called the zero point energy. Its importance is that even at absolute zero the system is not at rest but varies over a range of configuration coordinates characteristic of this lowest vibrational level. Analysis shows that under these conditions the widths of the bands, ΔE, vary as in Eq. (2), where A is a constant. A curve of this form is drawn through the experimentally obtained points of Fig. 4 and shows satisfactory agreement.

$$\Delta E = A\left[\coth\left(\frac{h\nu}{2kT}\right)\right]^{1/2} \tag{2}$$

One other result of the introduction of quantum mechanics is that this simple model predicts that both absorption and emission bands should be gaussian in shape at all temperatures; that is, they should be of the form of Eq. (3), where I is the emission intensity or absorption strength for light of energy E, and I_0 and E_0 refer to corresponding quantities at the maximum of the curve. **Figure 5** shows the emission spectrum of thallium-activated potassium chloride plotted in such a way that the expression of Eq. (3) would give two straight lines making equal angles with the horizontal. Although there is some disagreement in the wings of the emission spectrum, and although the lines are not quite at the same angle, there still is fairly good agreement with the predictions of Eq. (3).

$$I = I_0 \exp\left[-A(E - E_0)^2\right] \tag{3}$$

High dielectric constant materials. The use of configuration coordinate curves is justifiable only when the electron taking part in an optical transition is tightly bound to a luminescent center and interacts primarily with its nearest neighbors. This

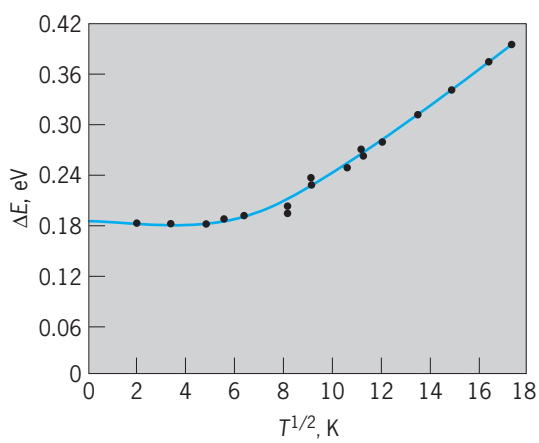

Fig. 4. Variation of the width of the *F*-center absorption band in potassium chloride at its half-maximum points, *FE*, as a function of the square root of the temperature. °F = (K × 1.8) − 459.67.

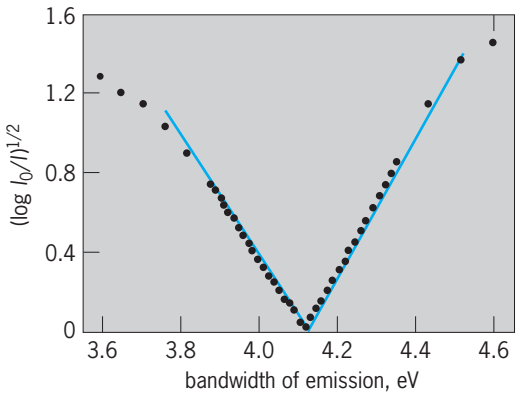

Fig. 5. Emission of thallium-activated potassium chloride at 4 K (−452°F). In a plot of this particular form, a gaussian curve would consist of two straight lines which make equal angles with the abscissa.

appears to be generally the case in materials with low dielectric constants such as the alkali halides. For high dielectric constant materials the situation is very different. It has been estimated that for boron in silicon the electron is spread over about 500 atoms so that its interaction with the nearest neighbors is small. In these cases the center interacts with the lattice during an optical transition through the creation of many vibrational phonons (sound quanta) at relatively large distances from the center. A case of this sort is illustrated in **Fig. 6**, which shows edge emission in cadmium sulfide near the absorption edge of the material. The major peaks correspond to an optical transition with simultaneous emission of 0, 1, 2, and 3 phonons, respectively, starting with the highest peak. The peaks are equidistant in energy, and the spacing is just that to be expected for the creation of phonons. Although both short- and long-range interactions of a center with its environment occur in all cases, the short-range interaction dominates for tightly bound electrons in low dielectric constant materials, while the long-range interaction dominates for loosely bound electrons in high dielectric constant materials. *See* DIELECTRIC MATERIALS; PHONON.

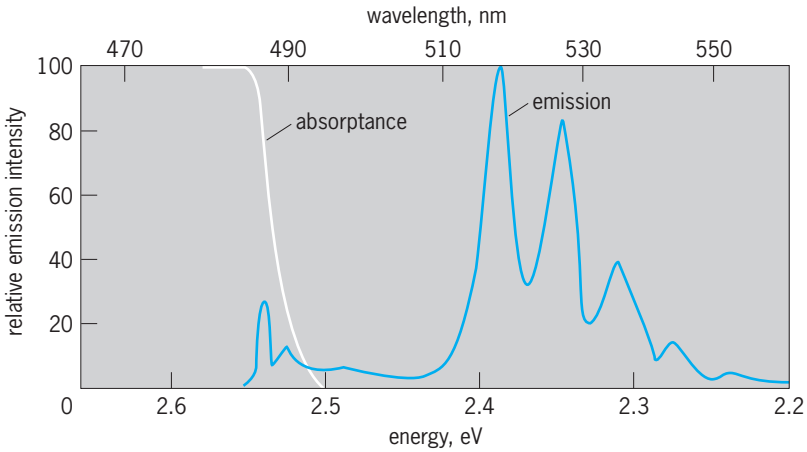

Fig. 6. Graph showing the absorptance (100 minus percent of transmission) and edge emission of cadmium sulfide at 4 K (−452°F).

Sensitized Luminescence

A process of considerable interest occurs in systems where one type of activator absorbs the exciting light and transfers its energy to a second type of activator which then emits. The transfer does not involve motion of electrons. This process is often called sensitized luminescence, and it is widespread in both inorganic and organic luminescent systems. The principal phosphors used in fluorescent lamps are of this type. The results on calcium carbonate ($CaCO_3$) shown in **Fig. 7** illustrate this kind of

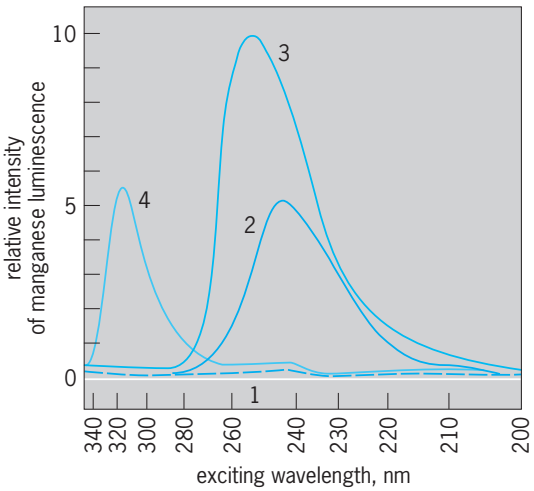

Fig. 7. Excitation spectra for emission of calcium carbonate phosphors, using manganese as an activator, with different impurity sensitizers: 1, no sensitizer; 2, thallium; 3, lead; and 4, cerium.

system. If divalent manganese (Mn^{2+}) alone is incorporated as an activator into $CaCO_3$, the system gives only a very weak Mn^{2+} emission for all wavelengths of exciting ultraviolet light. This is because the Mn^{2+} ion has only forbidden transitions in this spectral region, as shown by its very low absorption strength and slow decay of luminescence. In other words, the Mn^{2+} does not absorb very much ultraviolet light. Bright phosphors may be prepared by incorporating divalent lead (Pb^{2+}), monovalent thallium (Tl^{+}), or trivalent cerium (Ce^{3+}) along with Mn^{2+} in $CaCO_3$. In each case the Mn^{2+} emission is the same, but the wavelengths which excite these phosphors vary as the other activator (called the sensitizer) is changed. All these absorptions occur for energies well below the energy at which free electrons are formed in the solid. The ultraviolet light is absorbed by the Pb^{2+}, Tl^{+}, or Ce^{3+} sensitizers, which have allowed transitions leading to strong absorption of the light; this absorbed energy is then passed on to nearby Mn^{2+} activators and excite the Mn^{2+} emission. It is important to know over what distances the energy may be transferred from absorber to emitter. Experiments on calcium silicate with Pb^{2+} and Mn^{2+} as added impurities have shown that if a Mn^{2+} ion is in any one of the 28 nearest lattice sites around the Pb^{2+} ion, the energy transfer may take place.

Resonant transfer. The transfer of energy has been treated theoretically using quantum mechanics and

assuming a model of resonance between coupled systems. A number of cases have been examined; in each case the sensitizer has an allowed transition since only in this case will it absorb the exciting light appreciably.

The first case is one in which the activator undergoes an allowed transition. The probability of energy transfer, P (defined as the reciprocal of the time required for a transition to occur), varies as in expression (4). Here R is the distance between sensitizer

$$P \propto \frac{1}{R^6 n^4} \int f_s F_A dE \qquad (4)$$

and activator; n is the index of refraction of the host material, which may be liquid or solid; f_s is a function describing the emission band of the sensitizer as a function of energy, and is normalized so that the area under the curve is unity; and F_A is a function describing the absorption band of the activator, also normalized to unity. The integral of expression (4) measures the overlap of these bands and determines the resonance transfer. In typical cases the sensitizer will transfer energy to an activator if the activator occupies any one of the 1000–10,000 nearest available sites around the sensitizer. Another case is that of a forbidden quadrupole transition in the activator; in this case the number of sites for transfer would be about 100. If the transition in the activator is even more strongly forbidden, quantum-mechanical "exchange" effects predominate and the number of available sites for transfer should be reduced to 40 or less. In both of the cases just mentioned the integral measuring the overlap enters in the same way as it does in expression (4). From this theoretical treatment it appears that phosphors with Mn^{2+} as the activator probably receive their energy by exchange interactions.

Concentration quenching. Another phenomenon related to the resonant transfer of energy is that of concentration quenching. If phosphors are prepared with increasing concentrations of activators, the brightness will first increase but eventually will be quenched at high concentrations. It is believed that at high concentrations the absorbed energy is able to move from one activator to a nearby one by resonant transfer and thus migrate through the solid or liquid. If there are "poisons," or quenching sites, distributed in the material, the migrating energy may reach one of these, be transferred to it, and dissipate without luminescence. Impurity atoms, vacancies, jogs at dislocations, normal lattice ions near dislocations, and even a small fraction of the activator ions themselves, when associated in pairs or higher aggregates, can act as poisons. As the concentration is increased, the speed of migration is increased, and the quenching process becomes increasingly important.

Luminescence Involving Electron Motion

In an important group of luminescent materials the transfer of energy to the luminescent center is brought about by the motion of electrons. Many oxides, sulfides, selenides, tellurides, arsenides, and phosphides are of this type, and of these, zinc sulfide has been most widely studied. It is frequently used as the luminescent material in cathode-ray tubes and electroluminescent lamps.

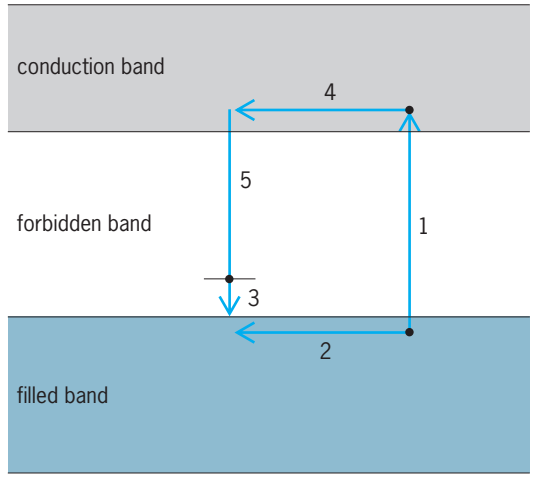

Fig. 8. Energy-band model for luminescence processes in zinc sulfide. Electron energy increases vertically.

Electronic processes in insulating materials are described by a band model such as that illustrated in **Fig. 8**. The electron energy increases vertically and the horizontal dimension shows position in a crystal. In the shaded areas, called the filled band, or the valence band, all the energy levels are filled with electrons. No electron can be accelerated or moved to higher energies within this band since the higher levels are already filled. Thus the material is an insulator. Above the valence band is an energy region called the forbidden band, which has no energy levels in it for pure materials. However, imperfections may introduce a local energy level in this region, as illustrated by the short line above 3 in Fig. 8. Above the forbidden band is an energy region called the conduction band. Here there are energy levels but, in an insulator, no electrons. If an electron in the filled band absorbs light of sufficiently high energy, it may jump up into the conduction band, as shown in 1 of Fig. 8. The empty position left behind in the filled band, called a hole, has properties which allow it to be described as an electron except that it has a positive charge. Both the electron in the conduction band and the hole in the filled band are free to move and gain energy. As a result, current can flow when a voltage is applied externally. The electron and hole can also diffuse far from their origin and can thus transport energy to a distant luminescent center. *See* HOLE STATES IN SOLIDS.

A simple luminescent transition is illustrated in Fig. 8. Assume that the impurity level (black dot in the forbidden band) is due to a luminescent center and that there is an electron in the level at the beginning of the process. Transition (1) shows the creation of a free electron and hole due to the absorption of light. The hole migrates to the center (2), and the electron in the impurity level falls into the hole (3),

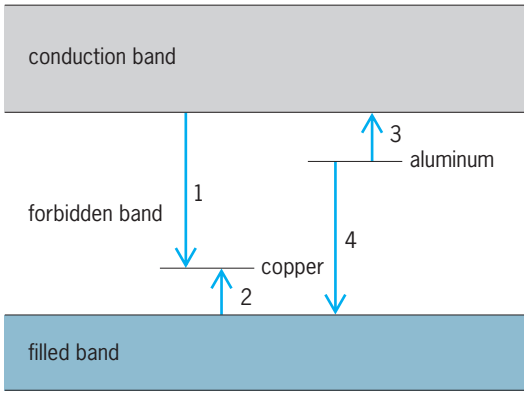

Fig. 9. Energy-band model for luminescence processes in zinc sulfide which uses copper as the activator and aluminum as the coactivator.

thus destroying it. The free electron now migrates toward the center (4) and falls into it (5), giving off luminescence. The cycle is complete and the center once again has an electron. It is important to note that in this process both electrons and holes must be assumed to move.

In zinc sulfide the normal activators are monovalent metals such as Ag^+ or Cu^+. They replace Zn^{2+} ions. To maintain electrical neutrality it is necessary also to incorporate into the lattice ions such as Cl^- for S^{2-} or Al^{3+} for Zn^{2+}. These additional ions are called coactivators. The situation is now quite complex, and is illustrated for ZnS with Cu and Al as added impurities in **Fig. 9**. The various arrows in this figure show some of the electronic transitions that give rise to observable effects. Arrows 1 and 4 show two different luminescent transitions. Arrow 2 shows the excitation of an electron into the luminescent center, which may occur by absorption of infrared light or by thermal fluctuations. In either case the center is prevented from luminescing and thus is quenched. Transition 3 is the excitation of an electron into the conduction band. The electron may be excited by thermal energy and its appearance may be detected by the appearance of transition 1. This process is called thermoluminescence.

The wide variety of luminescence phenomena, the insight they give into the constitution of very different classes of materials, and their many important applications to light production, imaging devices, and radiation detectors make luminescence a perennially challenging field of study.

Clifford C. Klick; James H. Schulman

Bibliography. H. J. Cantow et al. (eds.), *Luminescence*, 1981; J. D. Dunitz et al. (eds.), *Luminescence and Energy Transfer*, 1981; A. H. Kitai (ed.), *Solid State Luminescence: Theory, Materials, and Devices*, 1993; R. P. Rao (ed.), *Luminiscence: Phenomena, Materials and Devices*, 1991; D. Rendell, *Fluorescence and Phosphorescence Spectroscopy*, 1987; R. C. Ropp, *Luminescence and the Solid State*, 1991; S. G. Schulman (ed.), *Molecular Luminescence Spectroscopy*, 3 vols., 1985, 1988, 1993.

Luminescence analysis

Methods of chemical analysis in which analyte concentration is related to luminescence intensity or some other property of luminescence. Photoluminescence, particularly fluorescence, is the most widely used type of luminescence for chemical analysis. However, there are also important analytical applications of both chemiluminescence and bioluminescence.

Principles. Luminescence is light that accompanies the transition from an electronically excited atom or molecule to a lower energy state. The forms of luminescence are distinguished by the method used to produce the electronically excited species. When produced by absorption of incident radiation, the light emission is known as photoluminescence. Photoluminescence that is short-lived (10^{-8} s or less between excitation and emission) is known as fluorescence. Photoluminescence that is longer-lived (from 10^{-6} s all the way up to seconds) is known as phosphorescence. The reason for the difference in lifetime is that fluorescence involves an allowed, high-probability transition while phosphorescence involves a forbidden, low-probability transition. *See* FLUORESCENCE; LUMINESCENCE; PHOSPHORESCENCE; PHOTOLUMINESCENCE; SELECTION RULES (PHYSICS).

Chemiluminescence is observed when the electronically excited atom or molecule is formed as the product of a chemical reaction. If the light-producing reaction occurs in nature, such as the light emitted by fireflies, it is known as bioluminescence. *See* BIOLUMINESCENCE; CHEMILUMINESCENCE.

Spectra. Photoluminescence excitation spectra are determined by measuring emission intensity at a fixed wavelength while varying the wavelength of the incident light used to produce the electronically excited species responsible for emission. The excitation spectrum is a measure of the efficiency of electronic excitation as a function of excitation wavelength. Photoluminescence emission spectra are determined by exciting at a fixed wavelength and varying the wavelength at which emission is observed. **Figure 1** shows typical excitation and fluorescence and phosphorescence emission spectra for a molecule in solution. Between excitation and emission, electronically excited molecules normally lose some of their energy because of relaxation processes. As a consequence, the emission spectrum is at longer wavelengths, that is, at lower energy, than the excitation spectrum. Because the magnitude of the energy loss due to relaxation processes is greater for phosphorescence than for fluorescence, phosphorescence occurs at longer wavelengths than fluorescence. The spectra extend over a range of wavelengths, because in solution, molecules exist in a continuous distribution of vibrational and rotational energy levels. *See* ENERGY LEVEL (QUANTUM MECHANICS).

Chemiluminescence and bioluminescence spectra are equivalent to photoluminescence emission spectra for the electronically excited reaction products. The difference between these forms of luminescence

lies only in the method of production of the excited state.

Methods based on chemiluminescence and bioluminescence are restricted in scope because relatively few reactions produce light.

Intensities. Observed luminescence intensities depend on three factors: (1) the number of electronically excited molecules or atoms produced by the excitation process; (2) the fraction of electronically excited molecules that emit light as they relax to a lower energy state; and (3) the fraction of the emitted luminescence that impinges on the detector and is measured. In the case of photoluminescence, the number of excited molecules is proportional to incident excitation intensity, the concentration of the luminescent species, and the efficiency with which the luminescence species absorbs the incident radiation. In the case of chemiluminescence and bioluminescence, the number of excited molecules depends on reactant concentrations and the efficiency with which the reaction pathway leads to production of the excited state. The dependence of intensity on concentration is the basis for chemical analysis based on luminescence.

The fraction of excited species emitting light depends on the rate of the emission process relative to the rates of competing nonemissive processes and is expressed as the quantum yield of fluorescence or phosphorescence.

The fraction of the emitted light that is observed by the detector is a function of the optics of the instrumentation.

Instrumentation. Measurements of photoluminescence require a source of radiation to excite the sample and a detection system to measure the intensity of the resulting emission (**Fig. 2**). The most common radiation source is a xenon-arc lamp, which emits strongly throughout the ultraviolet and visible regions of the electromagnetic spectrum. Lasers provide high performance for specialized applications, but are not widely used for routine work because they are expensive to purchase and maintain and they emit at a limited number of wavelengths. Mercury-arc lamps provide high intensity at mercury emission lines and are relatively inexpensive. A filter or monochromator is used to transmit the excitation radiation of the desired wavelength selectively and to reject all other wavelengths. The light is focused on the sample. *See* SPECTROSCOPY.

As shown in Fig. 2, the detection system is usually placed at right angles to the incident light beam. This arrangement is employed to minimize the amount of excitation radiation that enters the detection system. A second filter or monochromator is used in the detection system to reject scattered excitation radiation selectively and transmit the desired photoluminescence that is at a longer wavelength. A photomultiplier tube detects and amplifies the luminescence signal. *See* PHOTOMULTIPLIER.

Instrumentation designed for phosphorescence measurements also includes provisions for exciting the sample sequentially and then measuring photoluminescence intensity after the source radiation is no longer incident on the sample. This arrangement discriminates temporally against scattered excitation radiation, making it easier to measure small luminescence signals. The same approach can also be applied to fluorescence measurements; but because fluorescence is so short-lived, the required instrumentation is too complex and expensive to be practical for most analytical applications.

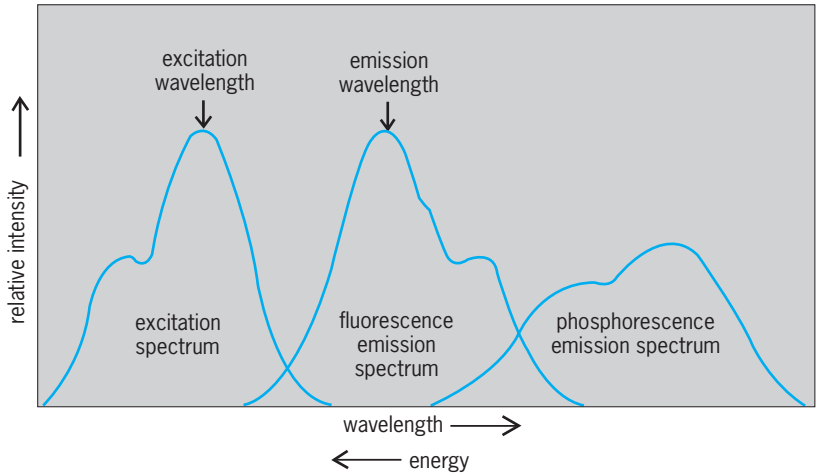

Fig. 1. Typical photoluminescence excitation and emission spectra for molecules in solution.

Chemiluminescence and bioluminescence measurements do not require excitation radiation or a means of resolving the emission of interest from scattered excitation radiation. Instead they require a system for initiating the reaction that gives rise to light production. Most frequently, this is accomplished by placing one component of the reaction in a sample container positioned in front of a photomultiplier tube that measures light intensity. The reaction is initiated by injecting a known volume of solution containing the other components required for light production.

Analytical variables. To maximize the observed intensity, most analytical photoluminescence

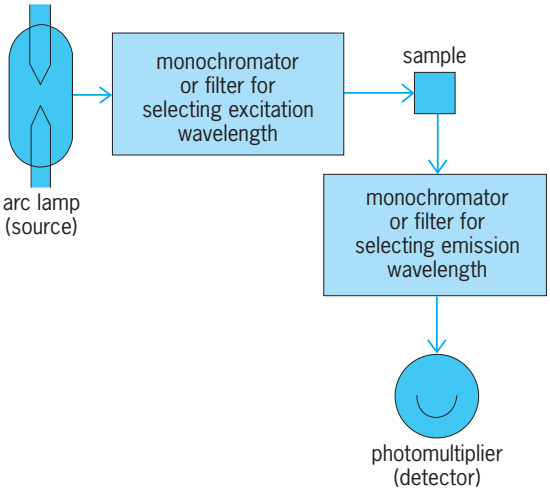

Fig. 2. Components of instrumentation for luminescence measurements.

measurements are made by exciting at the maximum point in the excitation spectrum and observing emission at the maximum point in the emission spectrum. These points are indicated in Fig. 1. However, if the sample contains multiple photoluminescent species, wavelengths may be chosen on the basis of selectivity. The goal is to maximize the signal for the photoluminescent species of interest relative to the signals for other interfering species.

Samples that are colored attenuate the excitation radiation as it passes through, thereby affecting the observed intensity. Therefore, they are avoided where possible for analytical photoluminescence measurements. This is most easily accomplished by diluting the sample with solvent until there is no longer any observable color. Dilution with solvent also reduces the likelihood that interaction between the luminescent species and other components of the sample will lead to quenching of fluorescence.

Colloidal samples scatter excitation radiation, and can be used for analytical photoluminescence measurements only with instrumentation that includes an emission monochromator of sufficiently high quality to resolve a relatively weak luminescence signal from a much stronger signal due to scattered excitation. *See* COLLOID.

Because solutions are clear and homogeneous, they are the preferred media for most analytical photoluminescence measurements. However, they are also subject to some drawbacks. The motion of molecules in liquids causes them to collide with each other, which can lead to quenching. This is rarely a serious problem for fluorescence, because it occurs on such a rapid time scale. However, phosphorescence is rarely observed from solution, because quenching processes occur more rapidly than light emission. Dissolved oxygen, in particular, is extremely effective at quenching phosphorescence. As a consequence, it is usually necessary to work with deoxygenated solid-phase samples to observe phosphorescence. The sample can be prepared by freezing a solution of analyte or by spotting analyte solution on an appropriate solid substrate and allowing the solvent to evaporate. Because of the inconvenience of working with solid-phase samples, phosphorescence is much less widely used for analysis than fluorescence.

Luminescent molecules in a solid phase can exist only in a limited number of unique environments rather than in a continuum of environments as is the case in solution. This leads to excitation and emission spectra that consist of a series of narrow lines rather than the broad bands illustrated in Fig. 1. By using solid-phase measurements to get narrow-line spectra, it is possible to identify individual components of mixtures of complex molecules that fluoresce (fluorophors), for example, polynuclear aromatic hydrocarbons in fossil fuels.

Chemiluminescence and bioluminescence measurements are usually made in solution or the gas phase. After the reaction is initiated, intensity decreases with time as reactants are consumed. The measured parameter can be the initial intensity or the intensity integrated over a specified time interval after the reaction has been initiated. An alternative approach is to mix the reactants required for light emission continuously in a flow system and to measure steady-state light output. In all cases, reactant concentrations need to be chosen so that the observed intensities depend on the concentration of analyte.

Applications. By far the most important advantage of luminescence methods of chemical analysis is their ability to measure extremely low concentrations. Analyte levels in the range of 0.001 to 10 micrograms per liter are readily measured by luminescence techniques. This advantage arises because luminescence is measured relative to weak background signals. In contrast, methods based on the absorption of light require that a small difference between two large signals be measured. Detection limits for luminescence methods are established by the magnitude of the luminescence relative to the background signal, which can arise from several sources including fluorescence from impurities in the solvent, Raman scattered light, detector dark current, and stray excitation light that is not eliminated by the emission filter or monochromator. *See* RAMAN EFFECT.

Compounds that are intrinsically luminescent may be determined directly. Polynuclear aromatic hydrocarbons are an important example. While most intrinsically luminescent compounds are aromatic organic substances, some inorganic compounds also luminesce. Uranium salts are an important example.

Luminescence methods may be extended to nonluminescent analytes by using reagents that react with the analyte to form a luminescent derivative. There is a vast number of such reactions. Some are specific for a particular analyte. Others are generic for a particular type of compound, such as amines or aldehydes. The generic reactions are often applied in combination with chemical separations. An important example is the analysis of amino acids by ion-exchange chromatography coupled to fluorescence derivatization and detection. *See* ION EXCHANGE.

Luminescence methods are particularly important in the life sciences. For example, sensitive measurements of enzyme activities may be made by using substrates that are converted to fluorescent products by the action of the enzyme. Immunological assays, which are based on the highly specific interaction between an antibody and an antigen, are frequently coupled to fluorescence measurements by chemical bonding of a high-efficiency fluorophor to either the antibody or the antigen. *See* ENZYME; IMMUNOASSAY.

W. Rudolf Seitz

Bibliography. M. A. DeLuca and W. D. McElroy (eds.), *Bioluminescence and Chemiluminescence, Basic Chemistry and Analytical Applications*, 1981; P. J. Elving (ed.), *Treatise on Analytical Chemistry*, 2d ed., pt. 1, vol. 7, 1981; G. G. Guilbault, *Practical Fluorescence: Theory, Methods and Techniques*, 1973; D. M. Hercules, *Fluorescence and Phosphorescence Analysis*, 1966; J. R. Lackowicz, *Principles of Fluorescence Spectroscopy*, 2d ed., 1999; C. S. Parker, *Photoluminescence of Solutions*, 1968; S. G. Schulman (ed.), *Molecular Luminescence Spectroscopy, Methods and Applications*, pt. 2, 1988.

Luminous efficacy

There are three ways this term can be used: (1) The luminous efficacy of a source of light is the quotient of the total luminous flux emitted divided by the total lamp power input. Light is visually evaluated radiant energy. Luminous flux is the time rate of flow of light. Luminous efficacy is expressed in lumens per watt. (2) The luminous efficacy of radiant power is the quotient of the total luminous flux emitted divided by the total radiant power emitted. This is always somewhat larger for a particular lamp than the previous measure, since not all the input power is transformed into radiant power. (3) The spectral luminous efficacy of radiant power is the quotient of the luminous flux at a given wavelength of light divided by the radiant power at that wavelength. A plot of this quotient versus wavelength displays the spectral response of the human visual system. It is, of course, zero for all wavelengths outside the range from 380 to 760 nanometers. It rises to a maximum near the center of this range. Both the value and the wavelength of this maximum depend on the degree of dark adaptation present. However, an accepted value of 683 lumens per watt maximum at 555 nm represents a standard observer in a light-adapted condition. The reciprocal of this maximum spectral luminous efficacy of radiant power is sometimes referred to as the mechanical equivalent of light, with a probable value of 0.00147 W/lm.

If the spectral luminous efficacy values at all wavelengths are each divided by the value at the maximum (683 lm/W), the spectral luminous efficiency of radiant power is obtained. It is dimensionless.

For purposes of illuminating engineering, light is radiant energy which is evaluated in terms of what is now simply and descriptively the spectral luminous efficiency curve, V_λ. In this case it is the intent that an illuminating engineer have a measure for the capacity or capability for the production of luminous flux from radiant flux, or that radiant flux has efficacy for the production of luminous flux.

Since the human eye is sensitive only to radiations of wavelengths between about 400 and 700 nm, it follows that the sensation of sight evoked by a radiator is due only to the radiation within this limited wavelength band. If all the wavelengths within this band were equally effective for the purpose of vision, then the area under the graph of radiant power distribution would be a measure of the light output.

The **illustration** shows the spectral distribution curve for an electric tungsten filament lamp. The total area under the curve between the wavelength limits $\lambda = 0$ and $\lambda = \infty$ is proportional to the power of the total radiation. In illus. *a* the shaded area between the limits of $\lambda = 400$ and $\lambda = 700$ nm is a measure of the power of visible radiation. However, this is not the same thing as luminous efficacy because of the variable sensitivity of the eye with respect to wavelength. If each ordinate, say at 10-nm wavelength intervals, is multiplied by the appropriate value of V_λ a new distribution is obtained, as shown in the shaded area of illus. *b*. This area is a measure of the light output of the source. The ratio of the shaded area of illus. *b* to the area under the larger curve is the luminous efficacy in terms of lumens per watt, and this is the method of determination that is generally used. The name luminous efficacy expresses exactly what is meant by the effectiveness of 1 watt of radiant power in producing luminous flux. *See* ILLUMINATION; LUMINOUS EFFICIENCY; LUMINOUS FLUX; PHOTOMETRY.

G. A. Horton

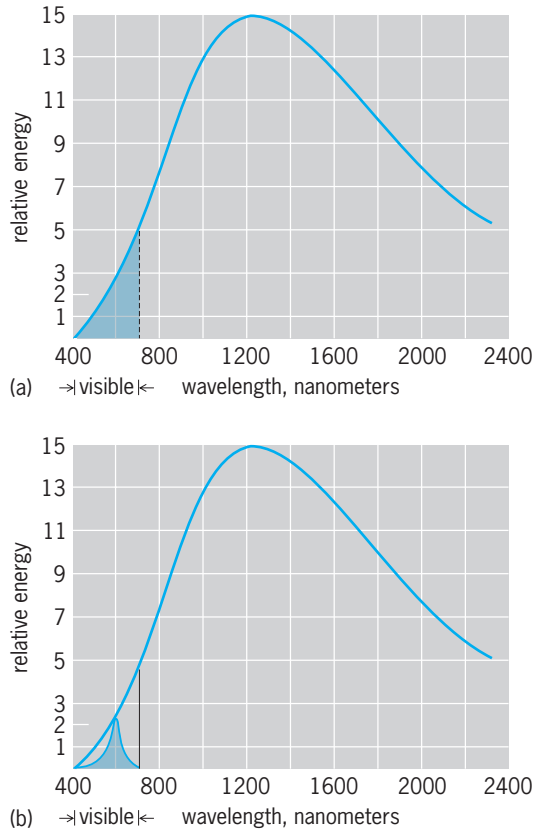

Spectral energy distribution of electric filament lamp. (*a*) Colored area is measure of power of visible radiation. (*b*) Colored area is measure of light output.

Luminous efficiency

Visual efficacy of visible radiation, a function of the spectral distribution of the source radiation in accordance with the "spectral luminous efficiency curve," usually for the light-adapted eye or photopic vision, or in some instances for the dark-adapted eye or scotopic vision. Account is taken of this fact in photometry, or the measurement of light. Actually, instead of indicating directly the physical radiant power of light, photometers are designed to indicate the corresponding luminous flux of the visually effective value of the radiation.

The spectral luminous efficiency of radiant flux is the ratio of luminous efficacy for a given wavelength to the value of maximum luminous efficacy. It is a dimensionless ratio. Values of spectral luminous efficiency for photopic vision V_λ at 10-nanometer-wavelength intervals were provisionally adopted by the International Commission on Illumination in

1924 and by the International Committee on Weights and Measures in 1933, as a basis for the establishment of photometric standards of types of sources differing from the primary standard in spectral distribution of radiant flux. These standard values of spectral luminous efficiency were determined by visual photometric observations with a 2° photometric field having a moderately high luminance (photometric brightness); photometric evaluations based upon them, consequently, do not agree exactly with other conditions of observations, such as in physical radiometry. Watts weighted in accordance with these standard values are often referred to as light-watts.

Values of spectral luminous efficiency for scotopic vision V'_λ at 10-nm intervals were provisionally adopted by the International Commission on Illumination in 1951. These values of spectral luminous efficiency were determined by visual observations by young, dark-adapted observers, using extra-foveal vision at near-threshold luminance.

The main difference between the terms luminous efficacy and luminous efficiency is that the latter came into use a century or so before the former. It should also be noted that in engineering, general values of efficiency never exceed 1.0, and illuminating engineers were the only ones who had values of efficiency up to 683, but this has now been changed by the usage of the term luminous efficacy. In view of the above discussion, the term luminous efficiency is evidently a misnomer. *See* ILLUMINATION; LUMINOUS EFFICACY; PHOTOMETRY. G. A. Horton

Bibliography. Illuminating Engineering Soc., *American National Standard Nomenclature and Definitions for Illuminating Engineering*, ANSI/IES R16–1986.

Luminous energy

The radiant energy in the visible region or quantity of light. It is in the form of electromagnetic waves, and since the visible region is commonly taken as extending 380–760 nanometers in wavelength, the luminous energy is contained within that region. It is equal to the time integral of the production of the luminous flux. *See* PHOTOMETRY.

Russell C. Putnam

Luminous flux

The time rate of flow of light. It is radiant flux in the form of electromagnetic waves which affects the eye or, more strictly, the time rate of flow of radiant energy evaluated according to its capacity to produce visual sensation. The visible spectrum is ordinarily considered to extend from 380 to 760 nanometers in wavelength; luminous flux is radiant flux in that region of the electromagnetic spectrum. The unit of measure of luminous flux is the lumen. *See* PHOTOMETRY. Russell C. Putnam

Luminous intensity

The solid angular luminous flux density in a given direction from a light source. It may be considered as the luminous flux on a small surface normal to the given direction, divided by the solid angle (in steradians) which the surface subtends at the source of light. Since the apex of a solid angle is a point, this concept applies exactly only to a point source. The size of the source, however, is often extremely small when compared with the distance from which it is observed, so in practice the luminous flux coming from such a source may be taken as coming from a point. For accuracy, the ratio of the diameter of the light source to the measuring distance should be about 1:10, although in practice ratios as large as 1:5 have been used without excessive error.

Mathematically, luminous intensity I is given by the equation below, where ω the solid angle is through

$$I = \frac{dF}{d\omega}$$

which the flux from the point source is radiated, and F is the luminous flux. The luminous intensity is often expressed as candlepower. *See* CANDLEPOWER; PHOTOMETRY. Russell C. Putnam

Lung

Paired, air-filled respiratory sacs, usually in the anterior or anteroventral part of the trunk of most tetrapods. Many elongated tetrapods, such as snakes, possess a simple lung, and one group of salamanders have lost their lungs completely with gas exchange occurring through the skin of the back. Lungs originated, however, early in fish evolution, presumably in the placoderms, as an adaptation to obtain atmospheric oxygen required for life in stagnant, oxygen-depleted waters. Lungs are still found in a few living species of freshwater fishes (such as lungfish) in which they may fuse to form a single organ and may come to lie well dorsally in the body; in many fishes the lungs have evolved into the swim bladder, a dorsal hydrostatic organ. *See* DIPNOI; SWIM BLADDER.

Lungs develop as outgrowths of the anterior end of the digestive tract, from the pharynx or throat, and remain connected to it by the trachea or windpipe. They lie within the coelom and are covered by peritoneum. In mammals they are within special chambers of the coelom known as pleural cavities and the peritoneum is termed pleura.

Amphibian lungs are often simple sacs, with only small ridges on the internal walls. In higher forms the lungs become more and more subdivided internally, thus increasing greatly the surface areas across which the respiratory exchange takes place. However, even in many reptiles the lungs may be quite simple. Birds have especially complex lungs with a highly differentiated system of tubes leading into and through them to the air sacs which are contained in many parts of the bird's body. Mammalian lungs are simpler, but in them the internal subdivision into tiny sacs or alveoli

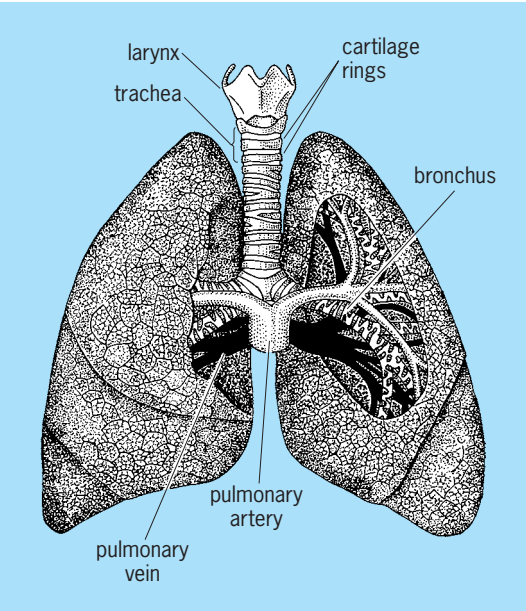

Human lungs. (*After T. I. Storer and R. L. Usinger, General Zoology, 4th ed., McGraw-Hill, 1965*)

is extreme; there may be over 350 million of them in one human lung.

In humans the two lungs lie within the chest, separated by the heart and mediastinum. The right lung has three lobes and the left lung two. A bronchus, an artery, and a vein enter each lung medially at the hilum; each branches again and again as it enters the lobules and smaller divisions of the lungs (see **illus.**). The terminal airways or bronchioles expand into small clusters of grapelike air cells, the alveoli. The alveolar walls consist of a single layer of epithelium and collectively present a huge surface. A small network of blood capillaries in the walls of the alveoli affords surfaces for the actual exchange of gases. *See* RESPIRATION; RESPIRATORY SYSTEM.

Thomas S. Parsons

Lupine

A cool-season legume with an upright stem, leaves divided into several digitate leaflets, and terminal racemes of pea-shaped blossoms. Three species, yellow, blue, and white, each named for the color of its blossoms, are cultivated as field crops; several hybrids are grown as ornamentals; and many species occur as wild plants. The yellow crop varieties are usually the earliest and smallest, the whites latest and largest. Field crop lupines have been grown in Europe since early Roman times as a soil-improving crop. The older varieties could not be used as forage since the plants contained a bitter, water-soluble, toxic alkaloid. However, since 1912, plant breeders have developed "sweet" varieties with only traces of alkaloid. *See* LEGUME.

Blue lupine was grown extensively as a winter cover crop in the southeastern United States during the 1940s. From the 1950s, however, lupine culture declined, partly because cheap nitrogen fertilizer competes with the nitrogen-fixing value of the cover crop, and also because diseases and winter killing of lupines increased.

The ornamental lupines are perennials unsuited to areas with hot summers. The many kinds of wild lupines which occur in the western part of the United States vary greatly in length of life, size, and site requirement. In the East a few small perennial wild lupines are found on deep sands. *See* BREEDING (PLANT); COVER CROPS; LEGUME FORAGES. Paul Tabor

Lutetium

A chemical element, Lu, atomic number 71, atomic weight 174.97, a very rare metal and the heaviest member of the rare-earth group. The naturally occurring element is made up of the stable isotope ^{175}Lu, 97.41%, and the long-life β-emitter ^{176}Lu with a half-life of 2.1×10^{10} years. *See* PERIODIC TABLE.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Lutetium, along with yttrium and lanthanum, is of interest to scientists studying magnetism. All of these elements form trivalent ions with only subshells which have been completed, so they have no unpaired electrons to contribute to the magnetism. Their radii with regard to the other rare-earth ions or metals are very similar so they form at almost all compositions either solid solutions or mixed crystals with the strongly magnetic rare-earth elements. Therefore, the scientist can dilute the magnetically active rare earths in a continuous manner without changing appreciably the crystal environment. *See* MAGNETOCHEMISTRY; RARE-EARTH ELEMENTS. Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; R. J. Elliott (ed.), *Magnetic Properties of Rare Earth Metals*, 1972.

Lychee

The plant *Litchi chinensis*, also called litchi, a member of the soapberry family (Sapindaceae). It is a native of southern China, where it has been cultivated for more than 2000 years. It is now being grown successfully in India, Republic of South Africa, Hawaii,

Burma, Madagascar, West Indies, Brazil, Honduras, Japan, Australia, and the southern United States. *See* SAPINDALES.

The fruit is a one-seeded berry. The thin, leathery, rough shell or pericarp of the ripe fruit is bright red in most varieties. Beneath the shell, completely surrounding the seed, is the edible aril or pulp. The fruit may be eaten fresh, canned, or dried. In China the fresh fruit is considered a great delicacy. *See* FRUIT.

Perry D. Strausbaugh; Earl L. Core

Lycophyta

A division (formerly Lycopodiophyta) of the subkingdom Embryobionta which represents one of two main lineages that evolved during the Early Devonian from the earliest land plants to develop vascular tissue and multiple sporangia. Integration of living and fossil taxa has allowed tentative reconstruction of the evolutionary history of the lycophyte group (**Fig. 1**), which is generally regarded as monophyletic (includes all the descendants of a single putative ancestor). The division Lycophyta contains two classes: the extinct Zosterophyllopsida (zosterophylls; **Fig. 2**) and the Lycopsida (club-mosses; **Fig. 3**), which contains 10 extant genera. *See* LYCOPSIDA; ZOSTEROPHYLLOPSIDA.

Living and fossil taxa. Living lycophytes are confined to the Lycopsida and are of far less phenotypic diversity and ecological significance than the fossils; most are small-bodied rhizomatous herbs or tuberous pseudoherbs occupying moist niches of low interspecific competition. Their economic significance is confined to various industrial applications of the abundant, small, and unusually uniformly sized spores of *Lycopodium*.

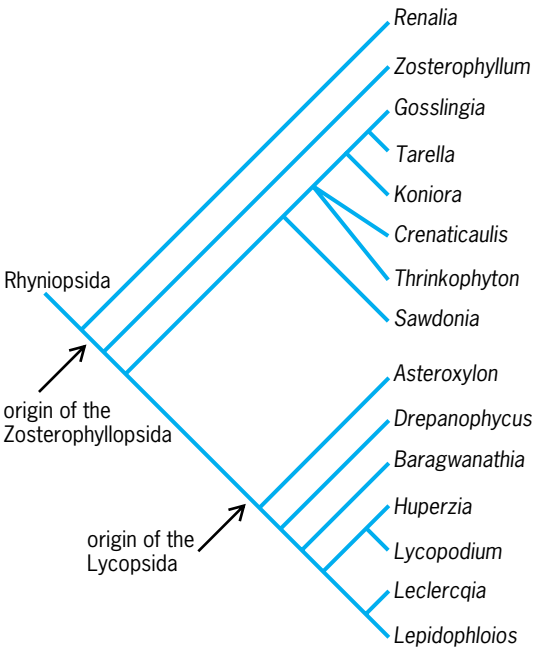

Fig. 1. Cladistic phylogeny of the Lycophyta. Only *Huperzia* and *Lycopodium* include extant species. (*After P. G. Gensel, Phylogenetic relationships of the zosterophylls and lycopsids, Ann. Mo. Bot. Gard., 79:450–473, 1992*)

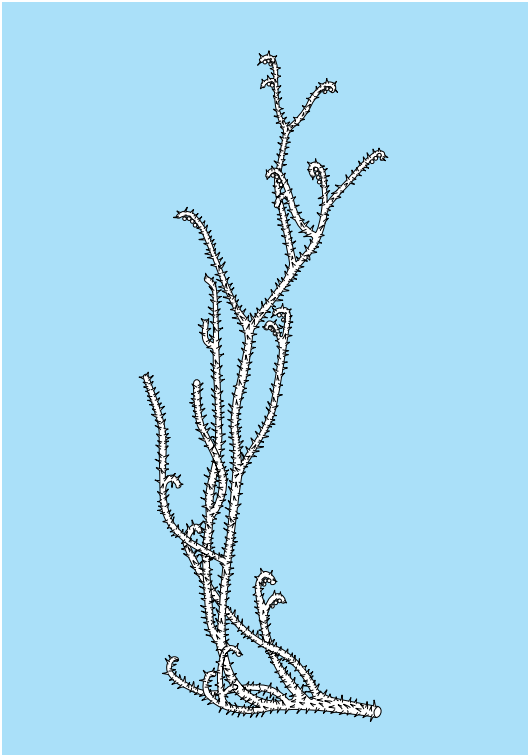

Fig. 2. Reconstruction of *Sawdonia ornata*. (*After Ananiev and Stopanov, Finds of sporogenous organs of Psilophyton princeps, Treat. Tomsk Lab. St. Univ. Geol. Ser., 202:30–46, 1968*)

The extant genera represent perhaps half of the major lycopsid groups known in the fossil record. Although living genera such as the lycopsid *Lycopodiella* (Fig. 3) and fully reconstructed fossils such as the zosterophyllopsid *Sawdonia* (Fig. 2) provide a strong phylogenetic framework, some key inferences of relationship depend on poorly known fossils, such as the putative rhyniopsid *Renalia*, which has been implicated as a potential ancestor of the Lycophyta. *See* RHYNIOPSIDA.

Characteristics. The lycophytes have a small gametophyte that is morphologically distinct from, and physically independent of, the larger sporophyte. Gametophytes are either supraterranean and photosynthetic or subterranean and saprophytic. The primitive growth habit is unipolar; rhizomes periodically generate adventitious roots and aerial branches. Vegetative branches are small, nonwoody, dichotomously branched, and either naked or, more commonly, scattered with cortical projections. The latter are termed emergences if unvascularized and microphyllous if supplied by a typically small and unbranched vascular strand. The primary xylem undergoes external maturation. Other key features of the Lycophyta are unvascularized, kidney-shaped eusporangia that are positioned laterally rather than terminating axes. The spores are released through distal lateral dehiscence slits.

Evolutionary pattern. The evolutionary history of the lycophytes generally is documented by increases

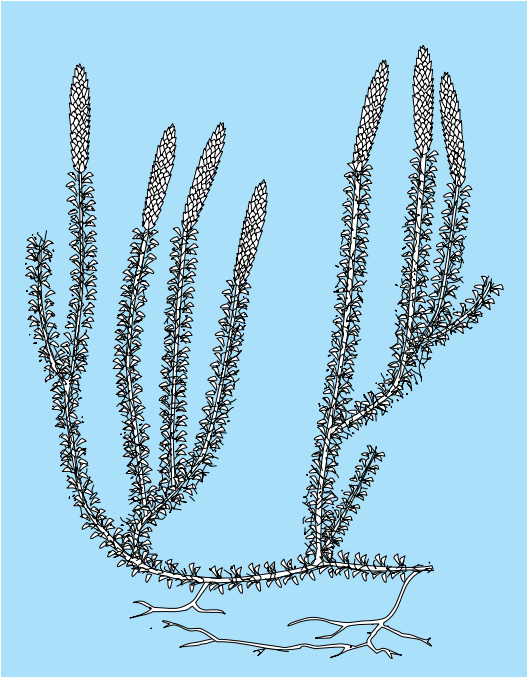

Fig. 3. ***Lycopodiella annotina*, an extant club moss. (*After W. N. Stewart and G. W. Rothwell, Paleobotany and the Evolution of Plants, 2d ed., Cambridge University Press, 1993*)**

in body size, vegetative complexity, and reproductive sophistication. Sporangia were borne on naked axes in zosterophyllopsids, but in the lycopsids were produced on the upper surface of the sporophylls. The sporophylls increasingly diverged in morphology from the leaves and became terminally concentrated as cones.

Primitive lycophytes produced only one size of spore. However, the most advanced lycopsid orders (selaginellaleans, lepidodendraleans, and isoetaleans) are uniformly heterosporous, producing microspores that form male gametophytes and megaspores that form female gametophytes. Further evolution in the rhizomorphic lycopsids led to retention of the megaspore in the megasporangium, in a structure analogous to a true seed. The rhizomorphic groups were also characterized by a centralized rootstock (rhizomorph), bipolar growth, large medullated protosteles, and wood production. These characters allowed growth to tree-sized dimensions, producing the most sophisticated of all pteridophytes. Richard M. Bateman; William A. DiMichele

Bibliography. D. W. Bierhorst, *Morphology of Vascular Plants*, 1971; W. N. Stewart and G. W. Rothwell, *Paleobotany and the Evolution of Plants*, 2d ed., 1993.

Lycopodiales

An order of the class Lycopsida (clubmosses) that is evolutionarily positioned between two extinct orders. The Lycopodiales are more advanced than the Asteroxylales because their sporangia are unequivocally positioned on the upper surface of modified leaves, but less advanced than the Protolepidodendrales because they lack a uniquely lycopsid feature termed the ligule. All have steles with metaxylem toward the center of the axis, but otherwise show considerable variation. All are perennial nonwoody herbs. *See* ASTEROXYLALES; PROTOLEPIDODENDRALES.

The fossil record of lycopodialeans begins with *Baragwanathia*. Its earliest occurrence, in Australia, has been controversially dated as Late Silurian or Early Devonian. Unlike most other lycopsid orders, the range of variation observed among the fossils is well reflected by the extant genera. Eight genera (of approximately 350 extant species) are recognized: the widespread *Huperzia, Phlegmariurus, Lycopodium, Diphasiastrum, Pseudolycopodiella, Lycopodiella*, and *Palhinhaea*, together with the enigmatic antipodean endemic *Phylloglossum* (see **illus.**).

Phylloglossum drummondii is one of the great curiosities of modern botany. Found in shrubland habitats in Australia and New Zealand, it does not resemble other lycopodialeans because of its apparently highly reduced morphology and centralized tuberous growth. Each growing season, the short underground stem branches to generate a new fleshy tuberous storage organ that in turn produces up to 10 quill-like leaves at the soil surface. From the center of this winter-green pseudowhorl emerges a single naked cone axis up to 2 in. (5 cm) long, with an unusual stelar morphology (mesarch siphonostele), and bears at its apex a small dense cone of up to 20 sporophylls. Each sporophyll bears a single kidney-shaped sporangium on its upper surface. Like all lycopodialeans, the plant produces only one size of spore. These spores generate gametophytes that initially are subterranean, but later grow to the soil surface and become photosynthetic. The highly specialized

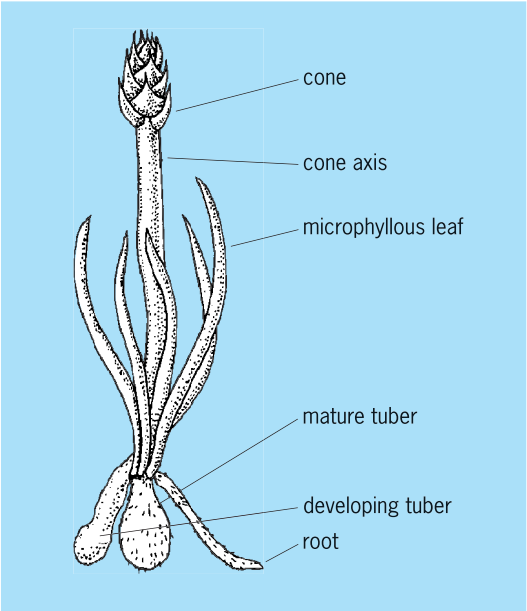

Fertile sporophyte of *Phylloglossum drummondii*. (*After A. W. Haupt, Plant Morphology, McGraw-Hill, 1953*)

morphology of *Phylloglossum* makes a comparison with other lycopsids difficult. Thus, its evolutionary position within the Lycopodiales, and even its inclusion in the Lycopodiales, remains questionable.

Phylogeny of extant genera. *Huperzia* and *Phlegmariurus* are the most primitive of the living genera. They have non-rhizomatous aerial axes with undivided, star-shaped actinosteles. Sporangia are simple in construction and positioned on sporophylls that resemble leaves; they do not form cones. In contrast, the remainder of the genera have rhizomatous growth and possess thin-walled sporangia and sporophylls that are aggregated into cones. *Lycopodium* and *Diphasiastrum* gametophytes share an annular meristem, and the stem vascular system is partially dissected into a plectostele. In *Pseudolycopodiella, Lycopodiella*, and *Palhinhaea*, the vascular system is further dissected into an annulus of apparently discrete but actually interconnected strands, and axial branching is more complex. Because gametophytes are lobed and occur on the soil surface, they are able to photosynthesize, thus reducing their dependence for nutrition on symbiotic mycorrhizal fungi. Also, the deciduous sporophytes of *Lycopodiella* and *Palhinhaea* develop from a protocorm and produce globose sporangia.

Habitats and life history. *Phlegmariurus* species grow hanging from other plants; *Pseudolycopodiella* and *Lycopodiella* are semiaquatics. The remaining genera are dominantly terrestrial. Lycopodialeans tend to prefer habitats where interspecific competition is low; few clubmoss species form a major proportion of the biomass in their preferred communities. Nonetheless, together they have a global distribution. Species-level diversity is greatest in the tropics, although a few species reach the Arctic Circle.

Lycopodialeans have a large and long-lived diploid sporophyte, which alternates with the wholly physiologically independent, smaller, and generally shorter-lived haploid gametophyte. The sporophyte produces spores that are released from the sporangium and germinate to produce gametophytes, which in most genera are subterranean and rely for nutrition on mycorrhizal fungi. The large gamete-producing sex organs are borne terminally. Antheridia produce biflagellate spermatozoids, which swim to archegonia and fuse with the egg to form a diploid zygote that develops into a juvenile sporophyte.

Fossil record. Many fossils of a wide age range have been assigned to the Lycopodiales. However, most are poorly understood vegetative fragments that often prove to belong to other orders of the Lycopsida, notably the Asteroxylales for Devonian material, Lepidodendrales for the upper Carboniferous, and Selaginellales for the Mesozoic. Identification as lycopodialean rests primarily on demonstrating the absence of wood, ligules, and megaspores—features of the more evolutionarily derived orders Protolepidodendrales, Selaginellales, Lepidodendrales, and Isoetales. The generic radiation within the Lycopodiales may have occurred in the Late Jurassic and Cretaceous periods, contemporaneously with the major radiations of the polypodiaceous ferns and angiosperms. *See* LYCOPHYTA; LYCOPSIDA. Richard M. Bateman; William A. DiMichele

Bibliography. B. A. Thomas, Paleozoic herbaceous lycopsids and the beginnings of extant *Lycopodium* and *Selaginella, Ann. Mo. Bot. Gard.*, 79:623–631, 1992; W. H. Wagner, Jr., and J. M. Beitel, Generic classification of modern North American Lycopodiaceae, *Ann. Mo. Bot. Gard.*, 79:676–686, 1992; D. P. Whittier and J. E. Braggins, The young gametophyte of *Phylloglossum* (Lycopodiaceae), *Ann. Mo. Bot. Gard.*, 79:730–736, 1992.

Lycopsida

A class (formerly Lycopodiopsida) of the division Lycophyta. It probably evolved in the Early Devonian from a zosterophyllopsid ancester; all six constituent taxonomic orders had probably evolved by the end of the Devonian. *See* ZOSTEROPHYLLOPSIDA.

Since this early and rapid radiation, lycopsids have always been globally distributed, with peak diversity in the tropics. In terms of both species diversity and biomass, lycophytes dominated many Devonian and Carboniferous plant communities, including the classic coal-swamp forests. Living lycopsids are of far less phenotypic diversity and ecological significance than the fossils; most are small-bodied rhizomatous perennial herbs (see **illus.**) or tuberous pseudoherbs

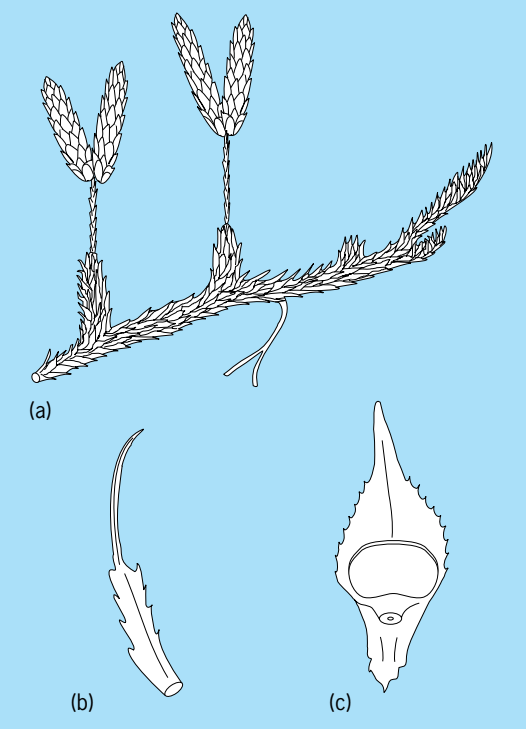

***Lycopodium clavatum*, a representative homosporous lycopsid. (*a*) Sporophyte showing typical rhizomatous growth and paired cones. (*b*) Microphyllous leaf. (*c*) Sporophyll bearing a sporangium on the upper surface. (*From K. R. Sporne, The Morphology of Pteridophytes, 2d ed., Hutchinson, 1975*)**

occupying moist niches of low interspecific competition. Many are polyploids.

Extant genera are relatively evenly distributed from a phylogenetic viewpoint. Fourteen representative genera of lycopsids belong to six orders, which either lack living species (Asteroxylales, Protolepidodendrales, Lepidodendrales) or possess them (Lycopodiales, Selaginellales, Isoetales). *See* ASTEROXYLALES; ISOETALES; LEPIDODENDRALES; LYCOPODIALES; PROTOLEPIDODENDRALES; SELAGINELLALES.

Because the prelycopsids of the Asteroxylales are treated as the most primitive lycopsids rather than as the most advanced zosterophyllopsids, characters delimiting the lycopsids concern vasculature. The inception of the order Lycopodiales was marked by vascularization of the microphylls and by the unequivocal sharing of a single vascular trace by the sporangium and the subtending microphyll. The Lycopodiales achieved considerable phenotypic diversity that is well represented in the extant flora although all remained rhizomatous herbs. Sporangia of the Protolepidodendrales remained adaxial but were situated farther from the axis; they were also radially elongate and had a longitudinal dehiscence slit. This group also showed the earliest evidence of a uniquely lycopsid feature, the ligule.

All the preceding orders are homosporous; that is, they generate only one spore type. However, the Selaginellales are heterosporous and produce large megaspores and small microspores that generate male and female gametophytes. Megasporangia and microsporangia are borne together in bisporangiate cones.

The Lepidodendrales and the Isoetales are characterized by finite growth from a centralized axial rooting structure termed the rhizomorph. All rhizomorphic species undergo at least some secondary thickening to generate wood and bark. These characters probably first appeared in the Lepidodendrales, allowing the development of large trees as well as shrubs and pseudoherbs which together could occupy a wide range of habitats.

The lycopsids illustrate the value of integrating evolutionary data from living and fossil plants. The fossils provide characters and, more importantly, combinations of characters that do not occur in the living genera, thereby bridging apparent evolutionary discontinuities. However, many fossils provide tantalizingly limited ranges of characters and so are problematic, both phylogenetically and taxonomically. Although some lineages remain geographically widespread and speciose, most species are only trivially distinct in morphological terms; many are polyploid. Also, they are biased toward morphologically simple but ecologically specialized species that occupy low-stress habitats, where they are interstitial elements in communities dominated by other plant groups. *See* LYCOPHYTA.

Richard M. Bateman; William A. DiMichele

Bibliography. W. N. Stewart and G. W. Rothwell, *Paleobotany and the Evolution of Plants*, 2d ed., 1993.

Lyme disease

A multisystem illness caused by the tick-borne spirochete *Borrelia burgdorferi*. The disease, also known as Lyme borreliosis, generally begins with a unique expanding skin lesion, erythema migrans, which is often accompanied by symptoms resembling those of influenza or meningitis. During the weeks or months following the tick bite, some individuals may develop cardiac and neurological abnormalities, particularly meningitis or inflammation of the cranial or peripheral nerves. If the disease is untreated, intermittent or chronic arthritis and progressive encephalomyelitis may develop months or years after primary infection. *See* NERVOUS SYSTEM DISORDERS.

Human. Erythema migrans, which occurs in about 60% of persons with the disease, is a unique clinical marker. It can be characterized by a small red papule or macule, often considered the site of the tick bite, that 3 to 23 days later spreads centrifugally, with induration up to 0.8 in. (2 cm) wide and usually flat borders. The centers of early lesions are usually erythematous and indurated but may also be vesicular or necrotic. As the lesions expand, the centers clear. Annular secondary lesions appear in almost half the individuals within several days of the initial lesion.

Causative agent. The causative agent, *B. burgdorferi*, is a helically shaped bacterium with dimensions of 0.18–0.25 by 4–30 micrometers (see **illus.**). An outer membrane surrounds the protoplasmic cylinder, which consists of a peptidoglycan layer, a cytoplasmic membrane, and enclosed cytoplasmic contents. Seven to eleven periplasmic flagella are attached subterminally and overlap at the central region of the cell; they are responsible for the organism's rotational and translational movements. *Borrelia burgdorferi* is gram-negative and uses carbohydrates such as glucose as energy and carbon sources.

Epidemiology. Once thought to be limited to the European continent, Lyme borreliosis and related disorders are now known to occur also in North

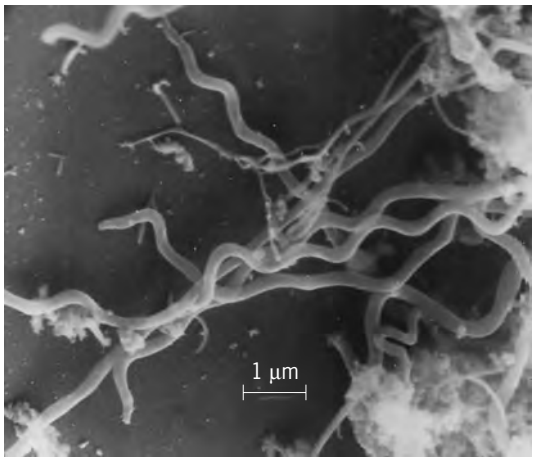

Scanning electron micrograph of *Borrelia burgdorferi*.

America, Russia, Japan, China, Australia, and Africa, where *B. burgdorferi* is maintained and transmitted by ticks of the genus *Ixodes*, namely *I. dammini, I. pacificus*, and possibly *I. scapularis* in the United States, *I. ricinus* in Europe, and *I. persulcatus* in Asia. *See* IXODIDES.

In the United States, Lyme borreliosis is the most prevalent arthropod-borne illness, affecting thousands of children and adults. Because of variations in case definition, no accurate data on its incidence are available.

Reports of Lyme disease in areas where neither *I. dammini* nor *I. pacificus* is present suggest that other species of ticks or possibly other bloodsucking arthropods such as biting flies or fleas may be involved in maintaining and transmitting the spirochetes.

Diagnosis and treatment. Diagnosis of Lyme borreliosis is based chiefly on clinical and epidemiologic evidence, particularly the presence of the unique skin lesion, erythema migrans. However, in as many as 40% of patients, that pathognomonic sign does not develop, rendering an early diagnosis only suggestive. Serologic tests such as indirect fluorescent antibody staining and enzyme-linked immunosorbent assay have been applied successfully in differentiating Lyme disease in its later stages from aseptic meningitis, rheumatoid arthritis, and other nervous system disorders such as multiple sclerosis. Both methods, however, lack sensitivity and specificity during the early course of the illness, and they have been shown to produce false-positive reactions. Because the procedures for both assays have not yet been standardized, differences in test results (serovariability) are common. The use of pathogen-specific antigens of *B. burgdorferi* should facilitate serologic diagnosis of Lyme disease.

All stages of Lyme borreliosis may respond to antibiotic therapy. Early treatment with oral tetracycline, doxycycline, penicillin, amoxicillin, or erythromycin can shorten the duration of symptoms and prevent later disease. Optimal treatment for arthritis, carditis, and neurological involvement is not yet clear, but intravenous penicillin and the cephalosporin, ceftriaxone, are highly promising. *See* ANTIBIOTIC.

Prevention and control. Prevention and control of Lyme borreliosis must be directed toward reduction of the tick population. This can be accomplished through reducing the population of animals that serve as hosts for the adult ticks, elimination of rodents that are not only the preferred hosts but also the source for infecting immature ticks with *B. burgdorferi*, and application of tick-killing agents to vegetation in infested areas. The Damminix method, a commercially available regimen that uses permethrin-impregnated cotton fibers, involves application directly to tick-infested rodents. Personal protection methods, including the use of effective tick repellents and toxins such as Permanone, are also recommended. *See* INFECTIOUS DISEASE; INSECTICIDE.

Animal. Lyme disease affects not only humans but also domestic animals such as dogs, horses, and cattle that serve as hosts for the tick vectors. Animals affected show migratory, intermittent arthritis in some joints similar to that observed in humans. In horses and cattle the most frequently observed clinical signs are lameness and swollen joints, although stiffness, inflammation of the hoof, abortion, and fever may occur. Spirochetes in the urine of such animals are said to be responsible for transmission of the disease by direct contact of an infected animal with an uninfected one. Birds also have been shown susceptible to *B. burgdorferi*, but their potential role in the epidemiology of the agent is still a matter of conjecture.

The disease is called canine Lyme borreliosis when it affects dogs. The transmission of *B. burgdorferi* from the tick to its host is minimal during the first 48 h of attachment; therefore, daily inspection and removal of ticks should be performed on dogs that have been in a tick habitat. A vaccine for the prevention of canine Lyme borreliosis is available, which is meant to be administered prior to *B. burgdorferi* exposure. In endemic areas, minimizing a dog's exposure to potentially infected ticks though the use of tick collars, sprays, and dips can be of value. Tick repellents containing DEET (*N,N*-diethyl-*meta*-toluamide) or permethrin are approved for veterinary use, and dog owners are advised to consult with their veterinarians before using these repellents and acaricides.

Because of the highly regional nature of canine Lyme borreliosis, dog owners considering either the vaccine or chemical tick-control measures should learn about the incidence of the disease in those areas that will be frequented by the dog. State health departments can provide information on the incidence of indigenously acquired Lyme disease and the presence of vector ticks in the geographic area of interest. Willy Burgdorfer; James J. Kasmierczak

Bibliography. M. J. Appel et al., Experimental Lyme disease in dogs produces arthritis and persistent infection, *J. Infect. Dis.*, 167:651–664, 1993; J. L. Benach and E. M. Bosler (eds.), Lyme disease and related disorders, *Ann. N.Y. Acad. Sci.*, 539:180–191, 1988; R. C. Johnson et al., *Borrelia burgdorferi* sp. nov.: Etiologic agent of Lyme disease, *Int. J. Syst. Bacteriol.*, 34:496–497, 1984; J. J. Kazmierczak and F. E. Sorhage, Current understanding of *Borrelia burgdorferi* infection, with emphasis on its prevention in dogs, *J. Amer. Vet. Med. Ass.*, 203:1524–1528, 1993; A. C. Steere, Lyme disease, *N. Engl. J. Med.*, 321:586–596, 1989.

Lymphatic system

A system of vessels in the vertebrate body, beginning in a network of exceedingly thin-walled capillaries in almost all the organs and tissues except the brain and bones. This network is drained by larger channels, mostly coursing along the veins and eventually joining to form a large vessel, the thoracic duct, which runs beside the spinal column to enter the left subclavian vein at the base of the neck. The lymph fluid originates in the tissue spaces by filtration from

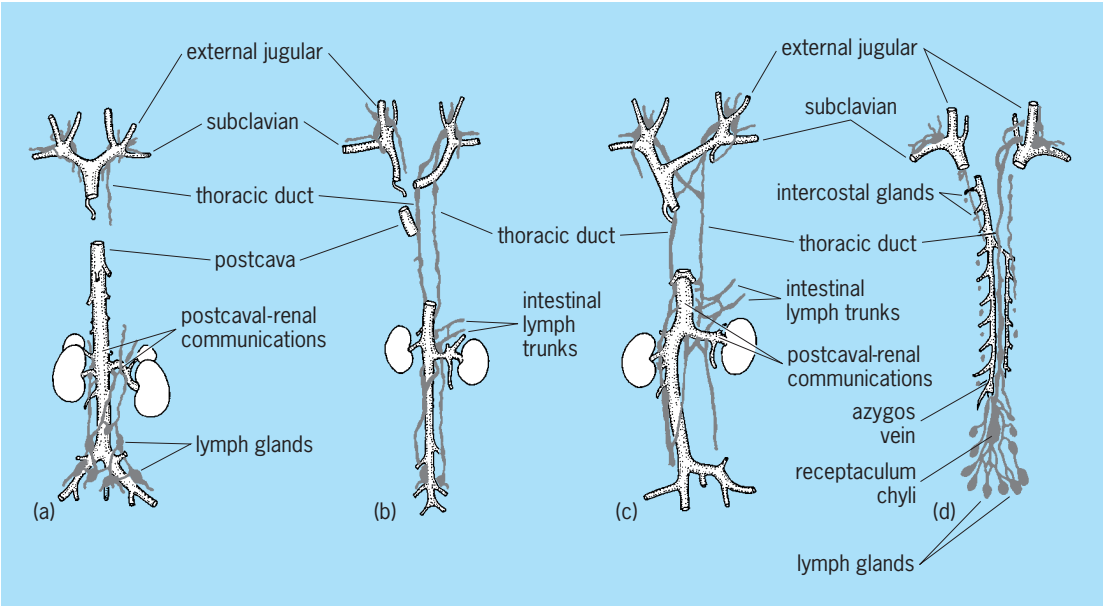

Fig. 1. Chief mammalian lymphatic trunks and their relations to the veins; ventral views. (*a*) South American monkeys. (*b*) Mammals (*Lepus*) in which postcaval-renal communications are wanting. (*c*) Mammals in general. (*d*) Humans (postcava and kidneys not shown). In all mammals the lymph enters the veins at the point of junction between the jugular and the subclavian veins. In most mammals there is also communication between the lymph vessels and the postcaval and renal veins. In humans the right thoracic duct degenerates in part and the only communication with veins is at the root of the jugulars. (*After H. V. Neal and H. W. Rand, Chordate Anatomy, Blakiston–McGraw-Hill, 1939*)

the blood capillaries. While in the lymphatic capillaries, it is clear and watery. However, at intervals along the larger lymphatic vessels, the lymph passes through spongelike lymph nodes, where it receives great numbers of cells, the lymphocytes, and becomes turbid. The lymphatic vessels other than capillaries contain numerous valves preventing backflow of the lymph. The functions of the lymphatic system are to remove particulate materials such as molecular proteins and bacteria from the tissues, to transport fat from the intestine to the blood, and to supply the blood with lymphocytes.

This article contains sections on the comparative anatomy, histology, embryology, and physiology of the lymphatic system.

Comparative Anatomy

It is disputable how far down in the animal series lymphoid tissue is found. Some researchers speak of lymphocytic cells in annelid worms, in mollusks, and in other invertebrates.

In general, the more primitive and less differentiated blood cells resemble more closely the lymphocytes of vertebrates. This is true also of the lower chordates in which so-called lymphocytes of the tunicates are said to differentiate into any of the other cell types of the body. Lymph nodules are described in some species of tunicates in the body wall and in the gut wall.

In the nonchordates it is difficult or impossible to make a distinction between blood vessels and lymphatic vessels. Probably the two systems had a common type of origin.

In *Branchiostoma* (*Amphioxus*), however, among the lower chordates, lymphatic vessels can be identified surrounding the blood vessels, and are seen also around the central nervous system, in the dorsal fin, and in the metapleural folds.

The lymphatics of vertebrates are distributed in somewhat the same manner as are the veins (**Fig. 1**). Thus they are divided into a deep and a superficial set. The deep set develops in relation to the cardinal veins and then makes connection with them and with the superficial set.

Fishes. In fishes the lymphatic system varies in degree of development in different species. In some, the veins seem to carry on the function of lymphatic drainage, but in a number of species of ganoids, teleosts, and elasmobranchs a well-developed lymphatic system is seen. The lymphatic vessels surround the veins in the elasmobranchs; in other fishes they are not so closely related to the blood vessels. In the elasmobranchs, also, there may be one or two large trunks which run parallel with the aorta and are comparable to the thoracic duct or ducts of humans and other mammals. Pulsating organs called lymph hearts, which assist in driving the lymph along the vessels, occur in some fishes (**Fig. 2**). The lymph vessels lack valves. Lymph nodes are absent, although there are collections of lymphoid tissue, chiefly close to mucosal surfaces. Nodules are found in these collections but they lack germinal centers.

The organ of Leydig of the selachian fishes consists of two large accumulations of lymphoid tissue which run longitudinally the whole length of the esophagus and even extend into the stomach. Although the cells of this organ are primarily lymphocytic, they apparently serve as stem cells for all the types of white blood corpuscles in these fishes, and hence the tissue perhaps should be spoken of as lymphomyeloid,

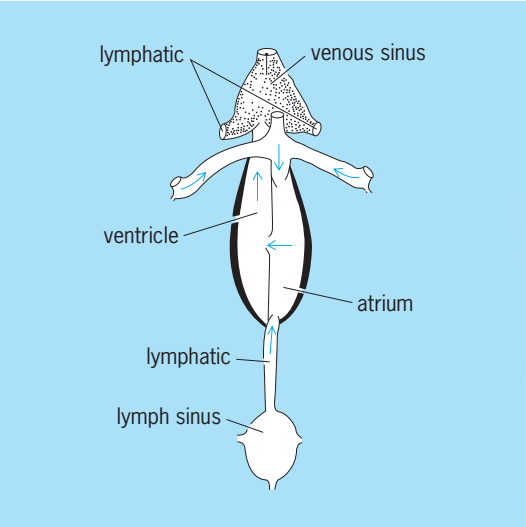

Fig. 2. Lymph heart in the tail of a teleost. (*After G. C. Kent, Comparative Anatomy of the Vertebrates, 2d ed., Blakiston–McGraw-Hill, 1965*)

indicating its hematopoietic as well as lymphoid character.

In the sturgeon a mass of lymphoid tissue has been described which occurs near the heart and resembles a lymph node, but its homologies are doubtful.

Amphibia. In the Amphibia the lymphatic system often includes very large subcutaneous sinuses, probably to protect against the danger of loss of moisture. Dilatations of the lymphatics, or lymph sacs, occur frequently. Lymph hearts are common in various parts of the body and a thoracic duct is present. Accumulations of lymphoid tissue are often present adjacent to the lymph sacs, but are not comparable in architecture to the lymph nodes of higher vertebrates. They may, however, represent an early stage in the development of such nodes.

Reptiles. Reptiles have a well-developed lymphatic system but lack the special features seen in amphibians. In the crocodiles true lymph nodes have been described in the mesentery.

Birds. The birds present a special condition in relation to the development of a lymphatic system. Here, as in other vertebrates, the more fluid portions of the blood readily leave the capillaries to diffuse among the tissue cells. This exudate is collected into lymphatic capillaries which join to form larger lymphatic vessels. The entire system converges toward the two superior (or anterior) venae cavae into which the main lymphatic trunks (thoracic ducts) empty. These veins bring the blood from the head, neck, and wings of the bird. The description of the lymphatic system of birds resembles that of mammals, even to the union of lymphatic with venous system.

The lymphatic vessels of birds are similar in structure to those of mammals and are provided with valves to help ensure one-way flow. However, the number of vessels is less than in mammals.

Lymph nodes are absent in vertebrates lower than the birds; that is, they are lacking in cold-blooded animals. Lymph nodes are not present in the common fowl (*Gallus domesticus*), but plexuses of lymphatic vessels do occur in the places where nodes are found in some other birds. Nodes are developed only in certain families of birds. This is not to say, however, that lymphatic tissue is lacking in birds without lymph nodes, for such tissue occurs in both diffuse and nodular form in the walls of the alimentary and respiratory systems: "Tonsillar" structures occur in the upper parts of these tracts; a large "esophageal tonsil" is found in the common domestic fowl; the bursa of Fabricius is a lymphoid organ in the wall of the posterior part of the alimentary tract; and the thymus appears to be an important producer of lymphocytes.

Early studies demonstrated that the quantity of lymphatic tissue increases, in general, toward the anal region in all birds studied. The avian cecum, in such mammals as the rabbit, is particularly rich in lymphoid tissue and may be said to constitute an "intestinal tonsil."

According to other findings, true lymph nodes probably occur only in water, shore, and swamp birds. Even when present, they are few, two pairs being found for the cervical-thoracic and two for the lumbar region. The nodes generally are elongated and spindle-shaped and lack a hilum. Capsule and trabeculae are lacking or very poorly developed.

Lymph nodes arise in birds about midway in embryonic life. The wall of the lymphatic vessel in the region of development becomes surrounded by a thick mass of mesenchyme which sends cordlike growths into the lumen of the vessel and eventually leaves only the central portion free, to serve as a central sinus into which the peripheral spaces open. Infiltrating lymphocytes transform the mesenchyme into truly lymphoid tissue. Nodules are formed, and in the centers of these nodules the germinal, or reaction, center develops, which is not found in fishes, amphibians, or reptiles.

These centers contain many phagocytic cells, or macrophages, as well as lymphocytes, some of which are undergoing mitotic division. These new centers in birds are found not only in the lymph nodes but also in individual nodules in other locations.

The lymph hearts, which are prominent features in cold-blooded vertebrates, were described in a number of species of birds, including the ostrich, cassowary, goose, swan, stork, and dove. The lymph hearts are paired structures in close relation to the kidneys and are true pulsatile organs provided with striated muscle. Two additional lymph hearts have been described in the goose; these have delicate, bladderlike walls which lie close under the hypogastric veins and are places of convergence for lymphatic vessels coming from several directions. From each of these hearts, a lymphatic vessel passes forward to the lumbar lymph nodes, and in its course receives tributaries from the pelvis and lower extremities.

In the fowl (*Gallus*) genuine lymph hearts with striated muscle are found only in the embryo. They

do not degenerate completely, however, until after hatching, and traces may be seen up to the thirty-fifth day after hatching.

Lymph hearts are said to occur in the embryos of all birds and are found on either side of the last sacral vertebrae. After hatching they disappear in some birds, as in the fowl, but are retained to some degree in the majority. They seem to be most prominent in the ostrich.

There does not appear to be a regular pulsation in the lymph hearts of birds but rather a contraction initiated by stretching of the wall due to filling. In the birds as a class the lymph hearts appear to be in a regressive condition when compared with such structures in amphibians and reptiles.

Mammals. The lymph nodes of mammals vary in number, size, form, and structure in different species. There appear to be some rather surprising differences in numbers of lymph nodes in different kinds of mammals. For example, the dog is said to have about 60, the cow 300, human 465, and the horse 8000. In the dog and cow, however, the lymph nodes are relatively large; in humans and the horse, they are relatively small. Their shape also varies. In the cow, for instance, the mesenteric nodes are long and narrow, attaining a length of several decimeters. In the rabbit the mesenteric nodes seem to have fused into one huge node, the pancreas aselli, weighing several times as much as all the other nodes combined. There is a similar arrangement for the mesenteric nodes in some other rodents and in some carnivores. These species include the mole, the dolphin, the seal, and the narwhal.

The amount of connective tissue of the lymph nodes, that is, the degree of development of the capsule and trabeculae, varies in different mammals. It is weakly developed in the pig, dog, sheep, rabbit, guinea pig, and rat. In humans and in the horse the trabeculae of the medullary part are more strongly developed than those of the cortex. Many mammals show a strong development of smooth muscle in the capsule and trabeculae, for instance, the horse, sheep, cow, and dog. The cow has an especially thick capsule and trabeculae.

In some lymph nodes the cortex, with its nodules, is very sharply set off from the medulla, with its more diffuse lymphoid tissue. This is true of the cow. In other animals the nodes seem to vary in this respect in different parts of the body or even in a rather haphazard manner, whereas in still others, such as the pig, there may seem to be almost a reversal of arrangement, with the nodular tissue situated centrally and the diffuse tissue peripherally.

Other lymphoid organs. These include the tonsils, thymus gland, and spleen, and in certain classes and groups of animals, structures which are confined to such groups, for instance, the bursa of Fabricius in the birds, a diverticulum from the lower end of the alimentary canal. *See* SPLEEN; THYMUS GLAND; TONSIL.

Tonsils. The tonsils seem to be represented in the frog and toad by two large fossae palatinae, invaginations of epithelium surrounded by dense masses of lymphoid tissue. Smaller structures of simple construction also are present.

In the posterior part of the buccal cavity of *Proteus* and *Salamandra* there are areas in which the epithelium is thickly infiltrated by lymphocytes.

Among the reptiles, *Lacerta agilis* shows definite tonsils in the pharynx, with elongated lymph nodules occurring along epithelial invaginations and with an apparent lack of basement membrane and a complete intermingling of epithelial and lymphoid tissue near the surface. Tonsils are seen also in the crocodiles.

Birds not only have large aggregations of lymphoid tissue in the pharynx but often also a large esophageal tonsil. The light-staining germinal centers are seen for the first time in birds.

A ring of lymphoid tissue, known as Waldeyer's ring, consisting of palatine, pharyngeal, and lingual tonsils, is present in many mammals, for instance, humans. Other mammals, however, may lack one or more of its component parts. In the rat and mouse, there are no palatine tonsils. In the hare and rabbit, pharyngeal tonsils are lacking. The rabbit, hare, dog, cat, and sheep apparently lack lingual tonsils.

Tonsils of the horse are peculiar in that they contain structures apparently identical with the Hassall's corpuscles of the thymus gland.

Spleen. The spleen is an organ of dual nature, belonging to both the blood-vascular and the lymphatic system. In the lower vertebrates it is actually within the wall of the alimentary canal. In cyclostomes it lies in the submucosa of stomach and intestine. In the lamprey it is within the spiral valve. Among the Dipnoi it is found within the stomach wall, but in elasmobranchs and ganoids it is a separate structure attached to the mesentery. In fishes and amphibians it is active as an erythrocytopoietic and lymphocytopoietic organ. Its function of forming lymphocytes is important also in higher vertebrates (including humans and other mammals), but its blood-cell-forming function is lost and replaced by one of destruction of "old" or worn-out red corpuscles and, to some degree, of blood storage. *See* HEMATOPOIESIS.

The lymphoid tissue of the spleen in higher vertebrates contains many nodules known as Malpighian bodies. These usually contain active germinal centers.

Histology

Lymphoid tissue is a tissue in which the predominant cell type, under normal conditions, is the lymphocyte. Lymphocytes are associated with reticular fibroblasts which form a delicate framework of fibers which can be demonstrated by their affinity for silver. There usually are cells of other types present also, including macrophages (fixed and free), plasma cells, and occasional eosinophile granulocytes.

Function. The chief function of the lymphoid tissue is to produce lymphocytes. In efferent lymphatic vessels, those leaving the lymph nodes, lymphocytes are considerably more abundant than in afferent lymphatics. Lymphocytes are added to the

lymph as it flows through the lymph nodes. The lymphocyte content of veins coming from the nodes also is greater than that of arteries going to them. Apparently newly formed lymphocytes also migrate into the capillaries and small veins. A second function of lymphoid tissue is protection against infection, which it carries out by means of its phagocytic and antibody-producing activity.

Cellular composition. Lymphocytes are of different sizes, small, medium, and large. The range in size of these three kinds, however, varies with different classes and even species of animals. The large lymphocyte resembles very closely the stem cell of blood-forming tissue; the hemocytoblast. In fact, lymphoid tissue itself bears a considerable resemblance to hematopoietic tissue, and in lower vertebrates it is not always possible to distinguish one from the other.

The function of the lymphocytes themselves is imperfectly known. The study of comparative histology, however, seems to throw light on their probable role in higher forms. In many invertebrates, the regenerative powers are much greater than in higher vertebrates. It has been shown that primitive mesenchymatous and free lymphoid cells often aggregate to form the various organ primordia and eventually even a completely differentiated animal. In some flatworms it has been found that not only do new parts arise from such elements (cells with conspicuous nucleoli and basophilic cytoplasm comparable to the large lymphocyte) but that the relative power of regeneration in different species may be correlated with the relative numbers of such cells. Thus, where they are numerous, as in *Planaria*, regenerative power is high. When they are not so common, as in *Procotyla*, regenerative power is low. In flatworms these cells will form such varied structures as epithelium of the gut, germ cells, and brain.

Although the lymphocytes of mammals do not have such remarkable potentialities, many workers have shown that they can develop into various other types of cells, including macrophages, which are large phagocytic cells; some types of granular white cells of the blood; red blood cells; and fibroblasts, or fiber-forming cells. The suggestion has been made also that they may metamorphose into epithelial cells in some locations. Thus one of their important roles in the higher vertebrates is still a regenerative or cell-replacing one.

Lymph vessels. The lymph capillaries, the smallest lymph vessels, form a dense network in almost every portion of the body. They appear to end blindly and consist only of an endothelial layer. As they unite to form larger vessels, connective tissue appears in the wall. The blind, rounded ends of lymphatic capillaries are easily seen in the lacteals in the villi, which receive the fatty foods from the intestinal lumen and hence contain chyle, a fluid made milky by the presence of multitudes of minute droplets of fat (**Fig. 3**).

In larger lymphatic vessels there are more collagenous fibers and also some elastic and smooth muscle fibers. In vessels with a diameter of more than 0.008 in. (0.2 mm), the wall generally shows three layers which bear the same names as those in the blood vessels, namely, intima, media, and adventitia. In lymph vessels the adventitia is usually the thickest layer and often contains smooth muscle cells as well as collagenous and elastic connective tissue fibers. In the largest lymphatic vessels, the right lymphatic duct and the thoracic duct, the media is more fully developed and contains considerable smooth muscle.

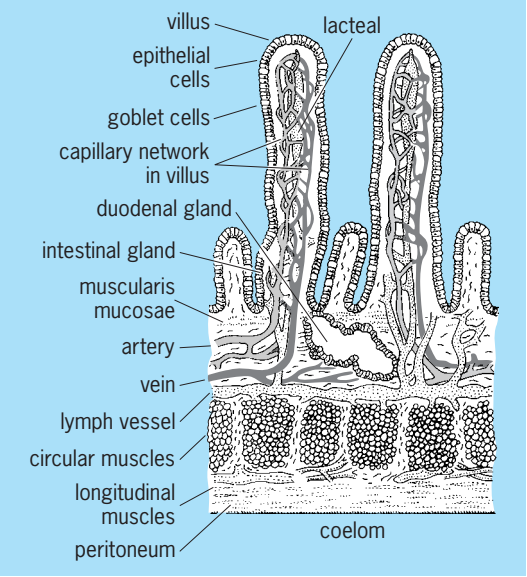

Fig. 3. Cross section of duodenum. (*After T. I. Storer and R. L. Usinger, General Zoology, 4th ed., McGraw-Hill, 1965*)

Lymph nodes. The lymph nodes are masses of lymphoid tissue with a special type of architecture. They are interposed as way stations in the course of the lymphatic stream. The vessels entering the node are called afferent; those leaving it are the efferent vessels.

The cortex usually consists of rather dense lymphoid tissue composed of rounded nodules, made up chiefly of small lymphocytes. In the centers of the nodules are lighter-staining portions, the germinal centers with macrophages and lymphocytes of medium and large size, frequently seen in mitosis.

The medulla of the lymph node consists of anastomosing channels or sinusoids, through which the lymph is coursing toward the efferent vessels and of cords of lymphoid tissue. It is more spongy than the cortex.

The lymph node is surrounded by a capsule consisting of collagenous and elastic connective tissue fibers and often containing some smooth muscle. At the hilum, where the blood vessels of the node enter and leave and from which the efferent lymphatics emerge, the connective tissue of the capsule penetrates well into the node.

Warren Andrew

Embryology

The mode of development of the lymphatic system, in contradistinction to the arterial and venous

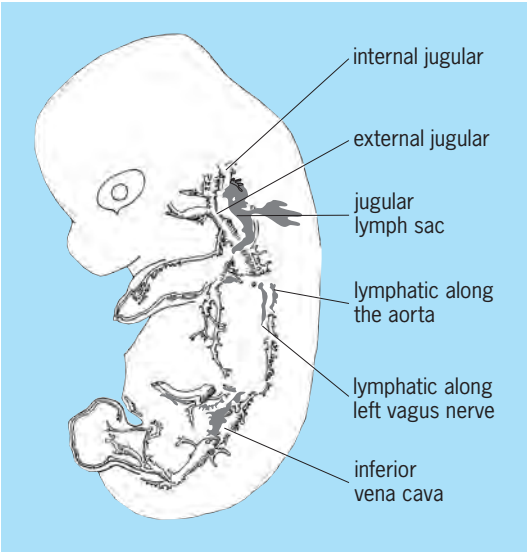

Fig. 4. Lymphatic vessels and veins in a 14.5-mm rabbit embryo. (*After R. O. Greep, ed., Histology, 2d ed., Blakiston-McGraw-Hill, 1966*)

systems, has been elucidated only since about 1910. Embryos of representative species from all classes of vertebrates have been studied, and as yet there is no agreement on fundamental questions pertaining to the early embryology of this elusive system. The unsolved status appears to rest more on differences of interpretations than on sufficient observations and facts.

A sharp distinction must be made between the processes of the earliest formation of discrete lymph vessels and those leading merely to an extension of once-established channels. The former entail actual genesis of endothelium (vasculogenesis); the latter merely represent growth. Growth by budding and branching of newly established vessels occurs throughout the vascular system, whether hemal (arterial and venous) or lymphatic. The controversial aspects mainly concern the earliest beginnings of lymph vessels.

Theories. There are two conflicting concepts. One is that all lymph channels develop by centrifugal outgrowth from embryonic veins, specifically from their endothelial lining. In the neck and loin regions, as well as in the posterior abdominal wall and at the root of the intestinal mesentery, such outgrowths form saclike expansions (**Fig. 4**; this figure, as well as **Fig. 5** and Fig. 7, shows general topography and structural plan in schematic form, but in detail these situations are not precisely as shown in the diagrams). These are, respectively, the jugular and iliac (lumbar) lymph sacs, the future cisterna chyli (pool of chyle near the base of the abdominal aorta), and the mesenteric lymph sac. All the body-pervading lymphatics are said to be formed by sprouting from these sacs and thus to be derived indirectly from the lining epithelium of blood vessels. The original points of budding from certain veins (cardinals, for instance) either could be retained or lost, and secondary evacuation taps for the lymph could be established.

The opposing view holds that the lymphatics arise directly from mesenchymal spaces, either adjacent to decadent venules or unassociated with veins. During early embryonic development, some rich and temporary venous plexuses become reorganized. Certain venules atrophy and disappear, particularly in regions of the developing lymph sacs. The drainage function is taken over by a lymphatic network which temporarily may carry blood cells received either from degenerative venules or from blood islands in the surrounding mesenchyme. The ultimate connection with major stem veins is always secondary. The system arises in a discontinuous manner and independently of the endothelium of veins. The overall direction of its development is centripetal. The isolated and discrete spaces converge and fuse to form continuous lymphatic trunks and networks. The main connection of the jugular lymph sacs with the venous system develops early, before the more distal lymphatic primordia have interconnected or joined the sacs. This link between two drainage systems, lymphatic and venous, occurs near the base of the neck (Fig. 4) at the jugulosubclavian junction.

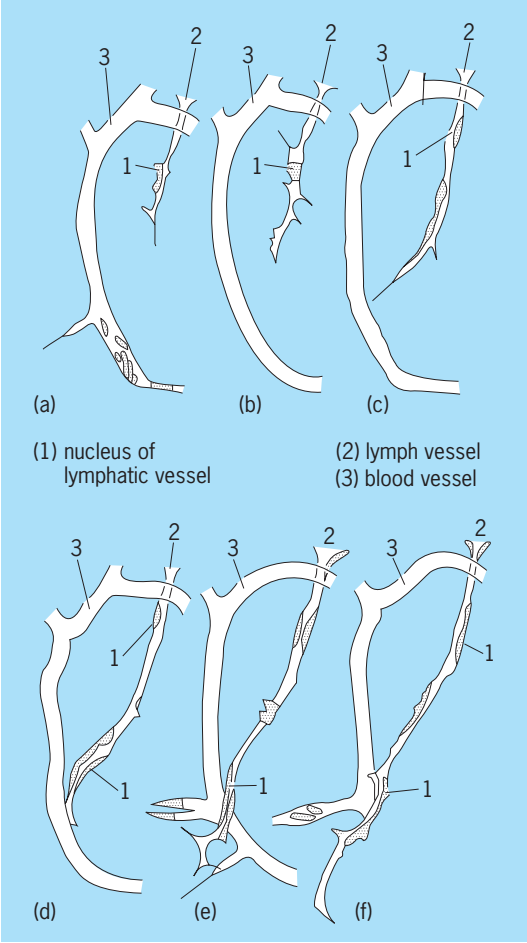

Fig. 5. Successive stages *a-f* in growth of lymphatic vessel in tail of tadpole during 72 h. (*After R. O. Greep, ed., Histology, 2d ed., Blakiston-McGraw-Hill, 1966*)

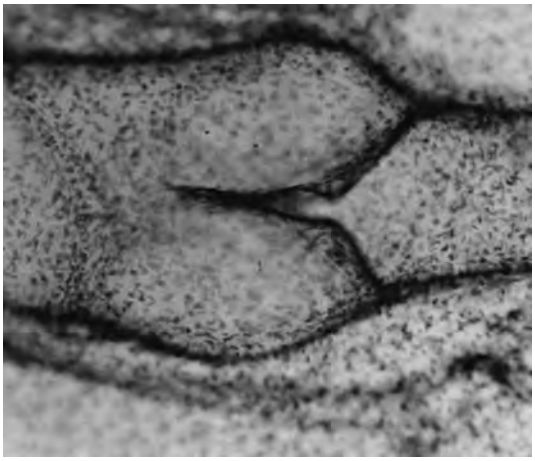

Fig. 6. Valve in lymph vessel to prevent backflow of blood. (*General Biology Supply House, Inc.*)

Investigative methods. There are two general methods for the investigation of lymphatic genesis; each has serious limitations. They consist either of the study of serial sections of well-preserved embryos or of injections of the developing lymphatics (Fig. 5). Crucial events in that development take place very early, in human embryos, for instance, when they measure only about 0.5 in. (10 mm). Obviously it is impossible to detect discontinuous segments of developing lymphatics by the injection method. Serial sections, on the other hand, can be obtained of uninjected as well as of injected embryos. This method in its turn presents difficulties in regard to fixation effects, shrinkage of the delicate embryonic tissues, and interpretive errors in reconstructions.

Evidence strongly supports the centripetal view of the origin of the lymphatic system. According to this view, the very first beginnings are identical with those of the arterial and venous components of the vascular system. All vasculogenesis appears to occur through the same fundamental transformation of mesenchymal cells into endothelial cells. The lymphatic precursor stages, succeeding the development of veins, as a rule are less distinct, evanescent, and therefore difficult to recognize. This has caused differing interpretations.

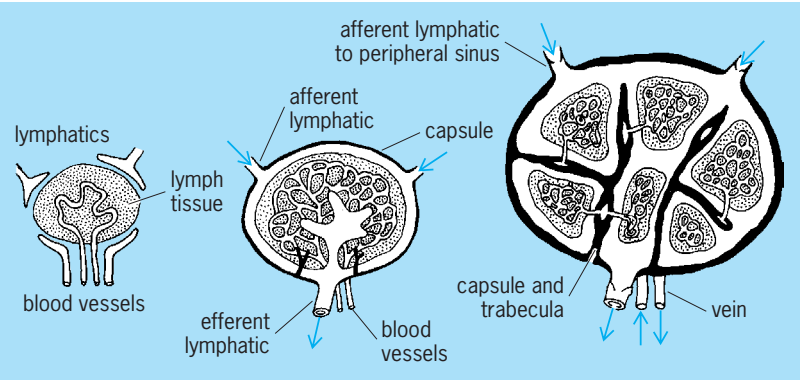

Fig. 7. Development of a lymphatic gland, with ingrowth of connective tissue strands. (*After L. B. Arey, Developmental Anatomy, 7th ed., Saunders, 1965*)

Embryonic phase. In human embryos of about 1 in. (20–30 mm) a connected primitive lymph system is already established. The main lymph channel ascending through the thorax in front of the vertebral column (thoracic duct) has joined the jugular lymph sac. At its distal end, it is continuous with the cisterna chyli, into which converge the lymphatics from the intestinal tract and those of the lower trunk and limbs. The jugular lymph sacs also receive lymphatic drainage from the developing upper limbs through the primitive ulnar lymphatics. Other lymph trunks develop in the head and neck regions, so that the jugular lymph sacs soon receive returning tissue fluids from all body regions. Lymph thereby is returned to the economy of the blood vascular system.

In early stages, the thoracic ducts are paired and symmetrically developed. Some animals retain double thoracic ducts. In many mammals and in humans the thoracic duct system becomes asymmetrical through the formation of interconnecting links and atrophy of certain segments. Henceforth, lymph drainage from the lower body regions occurs through a single main trunk, which terminates at the left jugulosubclavian junction.

Fetal phase. During early fetal months, the main lymph trunks, as well as the peripheral lymph vessels, become provided with valves which arise earlier in lymphatic vessels than in veins. Their structural arrangement enhances centripetal flow of lymph. By the fifth month, valves are present throughout the system (**Fig. 6**).

By ingrowth of connective tissue strands, the lymph sacs become transformed into a lacy plexus or fenestrated structure. In the meshes between the remaining lymph spaces the mesenchymal cells proliferate rapidly. Blood capillaries pervade the tissue bridges and immature lymphocytes appear in great numbers. The whole represents the beginning of a lymph node (**Fig. 7**). True lymphocytes further infiltrate the structure and produce compact cellular masses surrounded by lymph vessels. The complex may become subdivided to form groups of smaller lymph nodes. By the middle of fetal life the surrounding lymph vessels form a characteristic peripheral sinus, and an investing capsule is formed by the adjacent connective tissue. The full differentiation of lymph nodes is attained only after birth.

Arnold A. Zimmermann

Physiology

In any given region the fluid which enters a lymphatic capillary may be that which is close to the arteriolar or to the venous end of a blood capillary, or it may be fluid relatively distant from either. There exists a gradient of pressure from the tissue fluid to the lumen of the lymphatic capillary. In edema such as can be brought about experimentally by treatment with xylol or by heat, the gradient is increased. This pressure gradient is known to be an important factor in the formation of lymph. *See* EDEMA.

TABLE 1. Average values for the concentrations of electrolytes in lymph and plasma

Substance or property	Source of lymph	Subject*	Plasma	Lymph	Anesthetic
Sodium, meq/liter	Thoracic duct	Human (5)	127	127	
	Cervical duct	Dog	163	157	Nembutal
Potassium, meq/liter	Thoracic duct	Human (5)	5.0	4.7	
	Right lymph duct	Dog (5)		5.4	Nembutal
Calcium, meq/liter	Thoracic duct	Human (5)	5.0	4.2	
	Thoracic duct	Dog (1)	5.2	4.6	Amytal
	Cervical duct	Dog (11)	5.8	4.9	Nembutal
Chloride, meq/liter	Thoracic duct	Human (5)	96	98	
	Thoracic duct	Dog (1)	110	116	Amytal
	Cervical duct	Dog (7)	116	122	Nembutal
Inorganic phosphorus, mg %	Thoracic duct	Human (5)	4.5	4.4	
	Thoracic duct	Dog (1)	4.3	3.6	Amytal
	Cervical duct	Dog (3)	5.6	5.9	Nembutal
Carbon dioxide, ml %	Cervical duct	Dog (7)	56.8	58.8	Nembutal
Carbon dioxide tension, mmHg	Cervical duct	Dog (7)	46.4	40.3	Nembutal
pH	Cervical duct	Dog (7)	7.34	7.41	Nembutal

*Number of subjects studied is shown.

Any fluid in which filtration exceeds absorption through the wall of the blood capillary is normally drained by the lymphatics. Experimentally, this drainage is demonstrated where the filtration rate is increased by intravenous transfusions or by warming of parts.

Lymph collection. The function of the lymphatics is to remove such excess fluid and also to return to the bloodstream the protein which has escaped from the blood capillaries. Lymph collected from lymphatics of the kidney, leg, and cervical region, as well as from the thoracic duct, has been shown to contain this protein which has been "leaked" from the blood-vascular system. These proteins include serum globulin, serum albumin, and fibrinogen. The filtrate from the blood capillaries consists of water and of sugar and salts in concentrations found within the bloodstream, plus these proteins in low concentration. The salts and much of the water are reabsorbed by blood vessels, whereas the protein, together with some water and salt, enters the lymphatics. When the normal drainage of protein from the tissue spaces is prevented, elephantiasis may result, as when the lymphatics are obstructed in filariasis. The result is massive edema of a part, with dilatation and thickening of the lymphatic capillaries and an overgrowth of connective tissue.

In addition to water, ions, small molecules, and the large molecules of extravascular protein, particles such as injected graphite or extravasated red cells also enter the lymphatic capillaries. Such relatively huge structures pass between the cells.

Composition and properties of lymph. Basically, the composition of lymph closely resembles that of the plasma. Chemical and electrophoretic studies show that lymph contains all of the types of protein found in plasma, but in lower concentration. The composition of lymph varies to some extent from one part of the body to another. Thus, the lymph from the liver contains more protein than that from the skin.

The ability of lymph to coagulate depends largely on its content of fibrinogen and prothrombin, and this varies rather markedly. In dogs the mean prothrombin level is 7.6 for leg lymph, 51.2 for thoracic duct lymph, and 93.2 for liver lymph. There are no blood platelets in lymph, and it is thought that thromboplastin, generally considered necessary in clotting, is formed at the time of clotting from precursors present in the lymph. Calcium, also needed in clotting, is present in the lymph in slightly lower concentration than in the blood.

The ionic content of lymph does not vary much from that of plasma **(Table 1)**. Chloride and bicarbonate levels are higher in the lymph than in the plasma, whereas total base is slightly lower.

The glucose content usually is essentially the same in the lymph as in the blood. Curiously, the lymph of the kidney has a lower concentration of glucose. When carbohydrate is fed or when an intravenous injection of glucose is made, the concentration of this sugar in the lymph lags behind that in the plasma, but eventually catches up when an equilibrium between plasma and lymph is attained. Concentrations of amino acid, creatinine, and urea are approximately equal in lymph and plasma, but in the kidney the urea of the lymph is slightly higher than that of the plasma. It seems probable that some of the renal lymph comes from fluid reabsorbed from the distal and collecting tubules.

The lipids consist of neutral fat forming the chylomicrons, and of cholesterol and phospholipid, the latter two kinds of compounds being associated chiefly with proteins to form lipoproteins. The quantity of neutral fat depends on the fat absorption by the lacteals of the intestine, and the amounts of cholesterol and phospholipid vary with the amounts of protein present.

Lymph flow. Contractility of the walls of lymphatic vessels in humans and other mammals appears to be only slight and to play a minor role in lymph flow. This is in marked contrast to the condition in some lower vertebrates, particularly the

TABLE 2. Average flow and protein content of lymph from various lymph ducts in dogs under nembutal anesthesia*

Lymph duct†	Lymph flow, ml/h (in.^{3}/m)	g %	Lymph protein g/day (oz/day)	Lymph protein % of total intravascular plasma protein
Thoracic duct (4)	26.0 (1.6)	4.1	26.1 (0.92)	47.5
Right lymph duct (21)	2.3 (0.14)	3.7	2.0 (0.07)	3.6
Cervical ducts (18)	2.2 (0.13)	2.8	1.3 (0.05)	2.4
Foreleg ducts (5)	3.8 (0.23)	1.3	1.2 (0.04)	2.2
Hindleg ducts (17)	2.2 (0.13)	1.8	1.0 (0.04)	1.8

*After C. K. Drinker and J. M. Yoffey, *Lymphatics, Lymph and Lymphoid Tissue,* 2d ed., Harvard University Press, 1956. Average weight of each animal in each group was approximately 46 lb (21 kg) and average intravascular plasma protein 1.9 oz (55 g). Total lymph protein from leg and cervical ducts includes both sides.

† In parentheses is the number of animals in each group.

Amphibia, in which well-developed pulsatile lymph hearts occur along the course of the lymphatic channels (Fig. 2). The lymph flow in mammals appears to be very sluggish compared with that in the frog. The rather slow rate of flow of the lymph in dogs is seen in **Table 2**. It is calculated that extravascular circulation involves 50–100% of the plasma proteins each day.

The pressure of the lymph flow is variable, even in any given lymphatic vessel. Because lymph flows centripetally, it is clear that the peripheral lymph is under higher pressure than that farther along the vessels. Muscle contractions, such as those of the legs in walking, may greatly increase the pressure of the lymph in an individual region. During contraction of the muscle the lymphatics tend to empty, whereas during relaxation they fill again. In one experiment no pressure was recorded from a lymphatic just below the popliteal lymph node of a dog when the animal was at rest; but during rapid flexion and extension of the leg, the pressure rose to 27 in. (68 cm) of water.

Measurements of lymph pressure usually are made under standard conditions of rest. Pressure in a major division of the thoracic duct of the dog has been recorded as 2.6 in. (6.4 cm) of water, compared with a pressure of 1 in. (2.4 cm) in the internal jugular vein of the same dog. Such a pressure difference promotes flow from the thoracic duct into the venous system. Pressures have been recorded from a number of different lymphatic vessels. That of the lymphatic capillaries is rather high, having been recorded as 17 in. (42.6 cm) of water or 1.25 in. (31.3 mm) of mercury in the capillary of the villus.

Immunity. The lymph nodes serve as filtering-out places for foreign particles, including microorganisms, because the lymph comes into intimate contact with the many phagocytic cells of the sinusoids. These macrophages are of both the fixed and free wandering types. *See* PHAGOCYTOSIS.

In addition to the phagocytic function, experimental studies have demonstrated that lymphoid tissue produces antibodies. If antigens such as typhoid bacilli or sheep erythrocytes are injected into the pad of the foot of an experimental animal, antibodies begin to appear in the efferent lymph vessels of the popliteal node a few days later and their titer rapidly increases to a maximum and then declines. Some investigators have found that most of the antibody is within the lymphocytes rather than in the fluid. Germinal centers of the lymph node increase in size, and there is an increased output of lymphocytes. *See* ANTIBODY.

The actual process of antibody formation in lymphoid tissue is not well understood. A considerable body of evidence would assign the chief role to the plasma cell, generally abundant in lymphoid tissue, rather than to the lymphocyte. Antigen injection has been shown to bring about an intense plasma cell response in many instances.

The use of fluorescent antigens and antibodies offers evidence that antigen is fixed by plasma cells in lymph nodes which are forming antibodies, but is not fixed by lymphocytes. The increased numbers of lymphocytes in infection and their passage into the bloodstream may thus have some significance not connected with antibody formation.

Lymphocyte transformations. A large body of evidence supports the view that lymphocytes undergo transformations. Tracer studies combined with tissue culture methods for peripheral blood seem to show conclusively that the large proliferating cells in such cultures originate from small lymphocytes. Phytohemagglutinin acting on cultures of peripheral blood induces a transformation of small lymphocytes into large stem-type cells, which then become active mitotically.

A profound modification occurs in the ultrastructure of human lymphocytes cultured with phytohemagglutinin: New organelles increase in number in the cytoplasm of such cells; mitochondria become more numerous; the Golgi apparatus enlarges; and a variety of vesicles is seen. A transformation of small lymphocytes (parental strain) into larger cells of pyroninophilic type has been shown in F_1 hybrid rats. Cytochemical studies on lymphocytes and large blast cells in short-term culture show the dehydrogenase system highly active in glycolysis.

Cerebrospinal fluid. The cerebrospinal fluid fills the ventricles of the brain and the space between the arachnoid mater and the pia mater. Cerebrospinal fluid is a clear liquid derived from the capillary

vessels of the choroid plexuses which permit the passage of water, electrolytes, and small molecules, but hold back the proteins of the blood. Fluid passing through the walls of the capillaries must also go through a layer of epithelium covering the capillaries. Traces of protein are found in the normal cerebrospinal fluid. *See* NERVOUS SYSTEM (VERTEBRATE).

By far the greater part of the cerebrospinal fluid is absorbed into the veins of the arachnoid villi. The relationship between the cerebrospinal fluid and the lymphatic system is found in certain regions in which the fluid escapes from the meninges and enters lymphatic capillaries. Thus it may pass by way of the arachnoid sheaths of the olfactory nerves into the submucosa of the nasal cavity, and so reach the cervical lymphatic vessels. Arachnoid coverings of the sensory ganglia also are locations where cerebrospinal fluid may reach the lymphatics in the normal animal.

Clinically, such connections may serve as routes by which bacteria or viruses can enter the central nervous system. They may also, in cases of meningitis or of hemorrhage, help to clear the cerebrospinal fluid of cells and of plasma proteins.

Cerebrospinal fluid differs from other lymph in its low protein content, low cell content, and lower specific gravity (1.004–1.008). It contains no fibrinogen.

The cerebrospinal fluid has at least three functions: (1) It acts as a fluid buffer against shocks and jars of the central nervous system. (2) It serves as a reservoir in regulation of the cranial contents. (3) It acts as a mechanism for exchange of gases and nutrients in the nervous system. In relation to these functions, its properties are particular, but it may be considered as a specialized type of lymph. Warren Andrew

Bibliography. J. A. Barrowman, *Physiology of the Gastro-Intestinal Lymphatic System,* 1978; G. F. Bastian, *Lymphatic and Immune System,* 1993; M. Battezzati and I. Donini, *The Lymphatic System,* 1974; R. C. Myall and M. H. White, *Progress in Lymphology,* 1977; M. Nishi, S. Uchino, and S. Yabuki (eds.), *Progress in Lymphology,* no. 12, Proceedings of the 12th International Congress on Lymphology, 1990.

Lymphoma

Any of a group of malignant neoplasms derived from cells endogenous to lymphoid tissue. Lymphomas are grouped into two major categories: Hodgkin's disease and non-Hodgkin's lymphomas. Lymphomas usually originate in the lymph nodes located throughout the body, but they can arise from lymphoid tissue that does not form distinct nodes, such as that in the gastrointestinal tract or lung. Determination of the specific variety of Hodgkin's disease or non-Hodgkin's lymphoma was formerly based on the appearance of the cells when examined under a light microscope. Identification now relies on the nature of the cells with respect to certain substances (antigens) that they have on their surface or within their cytoplasm. For clinical purposes, lymphomas are categorized into three grades, low, intermediate, and high, with low-grade lymphomas having the best prognosis. *See* HODGKIN'S DISEASE.

Etiology. The etiology of most lymphomas is unknown. In experimental and domestic animals, viruses can cause lymphomas. Burkitt's lymphoma, a type of lymphoma that is rare in the United States but relatively common in children of central Africa, is thought to be caused by Epstein-Barr virus, a member of the herpes virus group. Electron microscopic studies of Burkitt's tumor cells have shown intracellular viral particles; Burkitt's lymphoma produces elevated serum antibody levels against Epstein-Barr virus. A form of T-cell lymphoma that has been identified in southern Japan has been attributed to a retrovirus referred to as human T-cell lymphoma-leukemia virus type 1 (HTLV-1). *See* EPSTEIN-BARR VIRUS; RETROVIRUS.

Clinical features of lymphomas. Patients with lymphomas may have painless swelling of various lymph nodes, such as those in the neck or near the armpit. Some patients, especially those with Hodgkin's disease, are referred to as B symptoms (fever, malaise, and weight loss). If the lymphoma originates in lymphoid tissue outside the lymph nodes, abdominal pain will signal lymphoma of the gastrointestinal tract and a cough will point to lymphoma of the lung.

Pathologic features. Lymph nodes involved by lymphoma are characteristically enlarged. They may be firm and have a consistency resembling fish flesh. In rare cases they are rock hard and they may show areas of cellular death (necrosis).

The appearance of lymphomas under a light microscope is highly variable. They can range from neoplasms composed of cells that very closely resemble normal lymphocytes, to tumors that are composed of large neoplastic cells. Hodgkin's disease has four microscopic subtypes, although the malignant cell in all of these subtypes is what is referred to as a Reed-Sternberg cell. The Reed-Sternberg cell is a tumor giant cell that frequently measures as large as 40 micrometers. It can have different forms, depending on the histologic subtype of Hodgkin's disease.

Cells of origin. Advances in the understanding of immunology have led to new concepts concerning lymphomas. It has long been known that there are two major classes of lymphocytes, referred to as B lymphocytes and T lymphocytes (B cells and T cells), both of which are formed in bone marrow. B lymphocytes differentiate into cells called plasma cells, which produce immunoglobulin. T lymphocytes are involved in cell-mediated immunity; before entering lymph nodes and other lymphoid organs, they are processed by the thymus gland to become functional, mature lymphocytes. Macrophages, sometimes referred to as histiocytes or M cells, are also of bone marrow origin, and are involved in a

number of functions, such as antigen presentation to the lymphocytes. By using modern immunologic techniques, most lymphomas can be identified as B-cell, T-cell, or M-cell type; about 90% of lymphomas are of B-lymphocyte origin. A T-cell lymphoma that occurs in the skin, referred to as mycosis fungoides, is a lymphoma of a specific subtype of lymphocyte labeled a T-helper/inducer lymphocyte. It is that subtype that is depleted in patients with acquired immune deficiency syndrome (AIDS). *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); CELLULAR IMMUNOLOGY; IMMUNOLOGY.

It is not always clear whether a lymph node shows the microscopic changes of a lymphoma or is a benign, reactive type of lymphoid cell proliferation. By using immunologic techniques with antibodies directed against various types of lymphocytes, the malignancy of the cells can often be determined. Other available techniques can ascertain whether the cells are of clonal origin, suggesting they are neoplastic, by determining that specific changes have taken place in their genes.

Evaluation of patients. Most diagnoses of lymphoma are made by surgical removal of a lymph node. The pathologist slices a thin section of the node, stains it with dyes, and examines it under a light microscope. If the lymph node shows a malignancy and is a lymphoma, then the type of lymphoma is determined by the appearance of the cells. Once a diagnosis of lymphoma is established, the patient usually must undergo a series of staging studies. These include a liver-spleen scan to determine if those organs are involved as well as a bone marrow biopsy to check for the presence of malignant cells. Some patients with Hodgkin's disease undergo exploratory abdominal surgery, at which time their spleen is removed for examination, and their liver and abdominal lymph nodes are biopsied and examined for evidence of the disease. *See* ONCOLOGY.

Treatment and prognosis. The treatment of lymphomas depends on the type of lymphoma diagnosed. Lymphomas in the low-grade group are usually not treated, since treatment does not increase life expectancy. Patients who have Hodgkin's disease or who have intermediate or high-grade non-Hodgkin's lymphoma are usually treated with chemotherapy with or without concurrent radiation.

Most cases of Hodgkin's disease and more than half the cases of intermediate and high-grade non-Hodgkin's lymphomas are potentially curable. Persons with low-grade non-Hodgkin's lymphomas usually have a life expectancy of 7–10 years, although some live considerably longer. Treatment protocols are now being developed for the low-grade lymphomas in the hope of increasing life expectancy. Among the experimental therapies that have been under investigation are treatment with interferon and with antibodies directed against tumor cells. *See* CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; LYMPHATIC SYSTEM.

Samuel P. Hammar

Bibliography. A. R. Moossa et al. (eds.), *Comprehensive Textbook of Oncology*, 2d ed., 1991; P. K. Pattengale, R. J. Lukes, and C. R. Taylor (eds.), *Lymphoproliferation Diseases*: *Pathogenesis, Diagnoses, Therapy*, 1985.

Lyophilization

Solvent removal from the frozen state by sublimation; commonly referred to as freeze-drying. The dried material, a porous solid, is sealed under vacuum or inert-gas atmosphere and retains its physical and biological characteristics indefinitely. When reconstituted for use, the porous product readily readmits solvent; hence the term lyophile, or solvent-loving.

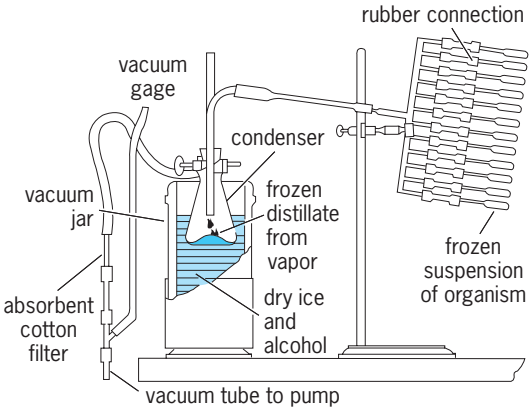

Apparatus for rapid high-vacuum desiccation of frozen suspensions of biologic materials. (*After A. J. Salle, Fundamental Principles of Bacteriology, 4th ed., McGraw-Hill, 1954*)

Lyophilization is accomplished by freezing the material to be dried below its eutectic point and then providing the latent heat of sublimation. Precise control of heat input permits drying from the frozen state without product melt-back. In practical application, the process is accelerated and more precisely controlled under reduced pressure conditions.

Several forms of apparatus are in use, ranging from the small laboratory equipment in the **illustration** to large chamber-type installations for preservation of serums, vaccines, blood plasma, soluble beverages, and foodstuffs.

Samuel C. Tease

Lyra

The Lyre (a stringed musical instrument), a small northern constellation (see **illustration**). It is marked by the fifth brightest star in the sky, Vega. In Greek mythology, the god Mercury made the first lyre by placing strings across a turtleshell. The constellation commemorates the lyre that the god Apollo gave to Orpheus, a gifted musician who used it to induce Pluto to release Orpheus's wife, Eurydice, from the underworld. Orpheus, warned not to look backward while liberating Eurydice, could

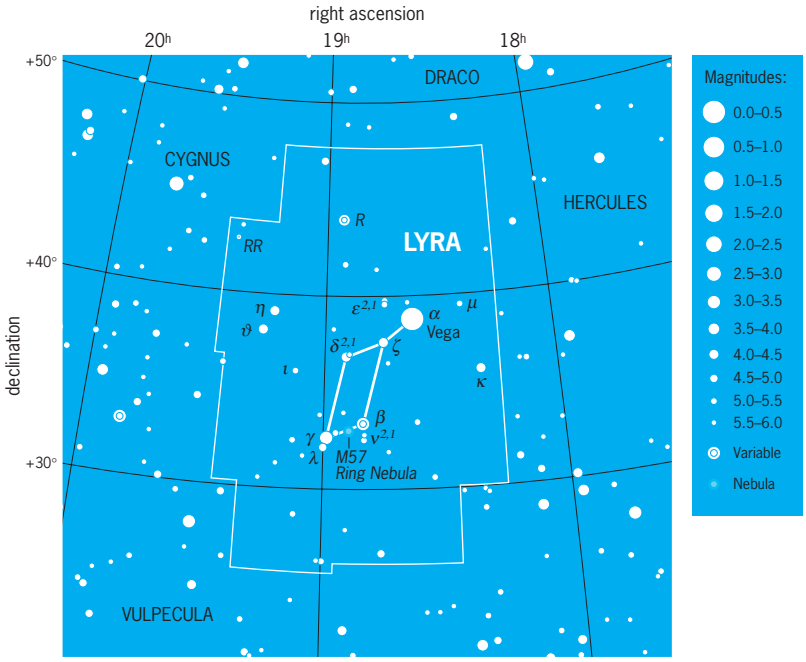

Modern boundaries of the constellation Lyra, the Lyre. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

not help himself from doing so, and lost her again. *See* VEGA.

Lyra contains the Ring Nebula, the most famous example of a planetary nebula, the gas given off by dying sunlike stars. Deneb in Cygnus, Vega in Lyra, and Altair in Aquila make the "summer triangle." Epsilon Lyrae, the "double-double" (that is, a multiple star system with two widely spaced components, each of which has two closely spaced stars), is a popular object for amateur astronomers. *See* BINARY STAR; CYGNUS; PLANETARY NEBULA.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Lysin

A term used in biology to describe substances that will disrupt a cell, with the release of some of its constituents. Unless the damage is minor, this action leads to the death of the cell. Lysins vary in the range of host species whose cells they will attack and in their requirements for accessory factors for lysis; the immune lysins are strictest in their requirements. Erythrocytes are lysed by a wide variety of chemicals, including water and hypertonic salt solutions, which displace the osmotic pressure from that of isotonicity. They are also susceptible to surface-active substances, such as saponin. Many bacteria, such as the staphylococcus and the streptococcus, elaborate one or more hemolysins that will lyse erythrocytes from certain, although not all, species of animals. These patterns may be used in the taxonomic classification of microorganisms. There are a number of other lytic substances, such as lysozyme, an enzyme in the eye aqueous humor and to some extent the bacteriophages, which attack diverse species of bacteria. These actions are not entirely nonspecific since they depend on the presence of common cell constituents or receptors which need not, however, coincide with those involved in immune lysis. *See* BACTERIOPHAGE; LYSOGENY; STAPHYLOCOCCUS; STREPTOCOCCUS.

Immune lysin is a term used in serology to designate an antibody that, in the presence of complement and cofactors, such as Mg^{2+} (magnesium ion) and Ca^{2+} (calcium ion), will disrupt a particular type of cell, with the release of some of its constituents. This action is in accord with the principles of serologic specificity and is to be distinguished from the more general actions of other lytic chemicals. Immune lysins are also classified according to the general class of cell attacked: Hemolysins are antibodies lysing erythrocytes, while bacteriolysins cause the lysis of bacterial cells. Hemolysins against various foreign erythrocytes occur normally in many sera; the Forssman and other heterophile antibodies and the blood group isoantibodies constitute special instances. High titers of hemolysins may also be produced by immunization with erythrocytes, and rabbit antibody against sheep erythrocytes is an important laboratory reagent, the so-called amboceptor, in complement-fixation tests. For example, sheep erythrocytes injected into a rabbit produce antibodies in the rabbit against the sheep red blood cells. The lysis of erythrocytes releases hemoglobin,

which may be quantitated by eye or by a colorimeter. *See* ANTIBODY; ANTIGEN; COMPLEMENT; COMPLEMENT-FIXATION TEST; HETEROPHILE ANTIGEN; LYTIC REACTION; SEROLOGY. Henry P. Treffers

Lysogeny

Almost all strains of bacteria are lysogenic; that is, they have the capacity on rare occasions to lyse with the liberation of particles of bacteriophage (see **illus.**). Such particles can be detected by their ability to form plaques (colonies of bacteriophage) on lawns of sensitive (indicator) bacteria. The genetic determinant of the capacity of lysogenic bacteria to produce bacteriophage is a repressed phage genome (provirus) which exists in the bacterium in one of two states: (1) integrated into the bacterial chromosome (most cases), or (2) occupying some extrachromosomal location (rare cases).

Bacteriophages which have the potential to exist as provirus are called temperate phages. When the provirus is integrated into the bacterial genome, it is called prophage. When the germinal substance (deoxyribonucleic acid, DNA, or deoxyribonucleoprotein) of certain temperate phages (for example, wild type of coliphage λ) enters a sensitive bacterium, the outcome may be death (lysis) for the bacterium as a result of phage multiplication, or it may result in the integration of the phage nucleic acid into the host genome (as a prophage), with the formation of a stable lysogenic bacterium. The lysogenic strain is designated by the name of the sensitive strain, followed, in parentheses, by the strain of lysogenizing phage, for example, *Escherichia coli* (λ). Such a bacterium differs from its nonlysogenic ancestor in one special way: It is immune to lysis by phage homologous to its carried prophage. *See* DEOXYRIBONUCLEIC ACID (DNA).

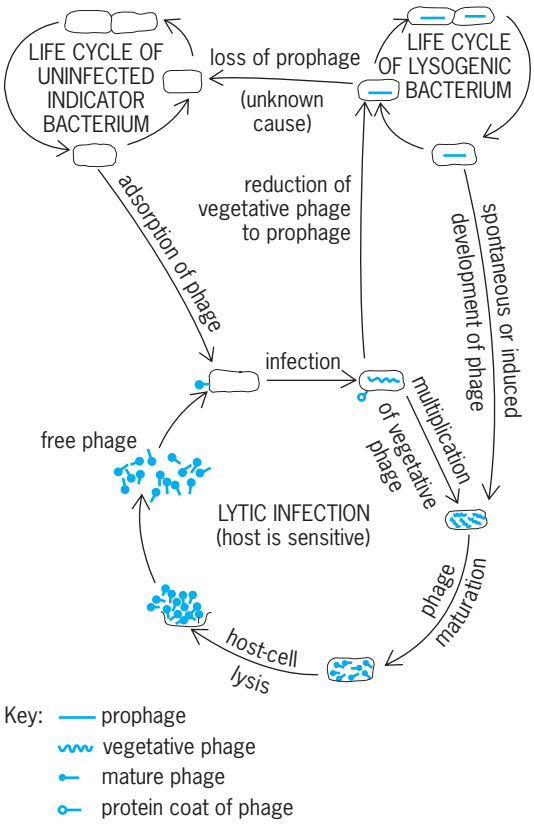

Life cycles of phage and bacterial host. (*After E. Jawetz, J. L. Melnick, and E. A. Adelberg, Review of Medical Microbiology, 2d ed., Lange, 1956*)

Repression. The conversion of a nonlysogenic cell to an immune lysogenic state by its resident prophage is the essential element of lysogeny, and is related to the regulatory mechanism by which the lysogenic state is maintained stably. When a prophage becomes part of the bacterial genome, most of the prophage genes, especially those having to do with DNA replication and the synthesis of viral protein, are repressed. This repression is attributed to the synthesis, under the direction of the prophage genome, of a repressor substance; thus at least one gene of prophage is expressed. This process is responsible for the immunity of lysogenic bacteria to lytic superinfection.

Immunity. The immunity conferred by prophage explains the fact that in most lysogenic cultures, while the host bacteria multiply normally, mature phage particles of the kind for which the host is lysogenic can be found in the surrounding medium. In a very few cells in any lysogenic population, the repressive mechanism fails to function, and prophage initiates replication and the host cell undergoes productive lysis. Since the population as a whole is immune, free virus merely persists in the environment. Under certain conditions, of which irradiation and the introduction of certain chemicals into the medium have been the best studied, a lysogenic culture may be induced to shift from the prophage to the lytic cycle; that is, the regulatory mechanism in the majority of cells is caused to break down and mass lysis of the culture occurs, with the production of many infectious phage particles. *See* LYTIC INFECTION.

Not all genes of the prophage are repressed, and a number of properties of lysogenic bacteria, assumed to be solely under the genetic control of the host cell, have been shown to be the expression of genes of the harbored prophage. Thus, the synthesis of diphtherial toxin by the diphtheria bacillus is controlled by the tox^+ gene of several corynebacteriophages, which render nontoxinogenic diphtheria bacilli both lysogenic and toxinogenic. Certain strains of *Salmonella* synthesize one kind of antigen when they are nonlysogenic, while when they harbor a specific prophage, antigen synthesis is altered by an expressed prophage enzyme, giving rise to another kind of antigen.

In one well-studied case, the origin of viral genes expressed in the lysogenic host cell is clear. Coliphage λ, liberated from galactose-fermenting *E. coli* (λ) has, sometimes, in the process of prophage integration, exchanged λ genes having to do with capsid production with host cell genes having to do with the enzymes of galactose utilization. When the repression mechanism for prophage fails, copies of λDNA

now contain genes for galactose fermentation. The DNA of this mutant, λdg, converts *E. coli* cells unable to utilize galactose into *E. coli* (λdg), the cells of which are lysogenic for λdg and able to utilize galactose. This galactose-positive state results from the expression of bacterial genes that have become an integral part of a viral DNA. *See* BACTERIOPHAGE; COLIPHAGE. Lane Barksdale

Bibliography. E. A. Birge, *Bacterial and Bacteriophage Genetics*, 4th ed., 2000; R. Dulbecco and H. Ginsberg, *Virology*, 2d ed., 1988; M. Dworkin, *Developmental Biology of the Bacteria*, 1985; W. Hayes, *The Genetics of Bacteria and Their Viruses*, 2d ed., 1976; W. K. Jokklik et al., *Zinsser Microbiology*, 20th ed., 1995.

Lysorophia

An order and clade of elongate lepospondyls with highly reduced limbs and a lightly built, open skull known from the Middle Pennsylvanian to Lower Permian (305–270 million years ago) of North America and the British Isles. Unlike other lepospondyls, no lysorophians have been found in continental Europe. The vast majority of specimens come from two localities: the coal deposits of Linton, Ohio, and the Clear Fork redbeds of north central Texas. They were aquatic in their lifestyle, and are frequently found in large numbers curled in burrows. These burrows are believed to be estivation chambers, structures constructed by animals for periods of dormancy associated with suboptimal living conditions. *See* LEPOSPONDYLI.

Features. Three species from one family are currently recognized: *Brachydectes newberryi, B. elongatus*, and *Pleuropteryx clavatus*, ranging in size from a few centimeters to about 1.5 m (4.5 ft). All possess the same unique anatomical features. The skull of lysorophians has a large opening inside the skull that extends the orbit back to the bones of the jaw articulation. The jaw articulation is inclined in a forward direction, and the lower jaw is short. Lysorophians have highly ossified remnants of the gill support skeleton, unique in lepospondyls, that could have supported three sets of gills, but it is unclear whether gills were present or if these bones performed other functions such as anchoring the tongue muscles. Vertebrae are distinctive, having neural arches that do not meet dorsally and are sutured, not fused as in other lepospondyls, to the single, spool-shaped vertebral body (or centrum). There are between 69 and 97 presacral vertebrae, and tails were probably relatively short. Limbs are short, about as long as four vertebrae, and probably had four fingers and five toes, but they are very poorly preserved so certainty is impossible.

Phylogeny. Relationships of lysorophians are very unclear. Traditionally they have been placed within the microsaurs, a group of salamander-like lepospondyls. In recent analyses of relationship they are frequently placed next to other elongate, limbless lepospondyls such as aistopods and adelospondylids, but many of the uniting features can easily be associated with being long and limbless and not recent shared common ancestry. Recently a small number of scientists have hypothesized that lysorophians are connected to the origins of modern amphibians (frogs, salamanders, and caecilians). Critics of this idea point out that most of the features that link lysorophians and modern amphibians are shared absences, which is not strong evidence. However, this idea is driving much new research on the question of the origins of modern amphibians. *See* AISTOPODS; AMPHIBIA; MICROSAURIA. Jason S. Anderson

Bibliography. M. J. Benton, *Vertebrate Palaeontology*, 3d ed., Blackwell Press, London, 2004; R. L. Carroll, Lepospondyls, in H. Heatwole and R. L. Carroll (eds.), *Amphibian Biology*, vol. 4: *Palaeontology: The Evolutionary History of Amphibians*, pp. 1198–1269, Surrey Beatty, Chipping Norton, NSW, Australia, 2000; R. L. Carroll, *Vertebrate Paleontology and Evolution*, Freeman, New York, 1988; C. F. Wellstead, Taxonomic revision of the Lysorophia, Permo-Carboniferous lepospondyl amphibians, *Bull. Amer. Mus. Nat. Hist.*, 209:1–90, 1991; C. F. Wellstead, Order Lysorophia, in P. Wellnhofer and R. L. Carroll (eds.), *Encyclopedia of Paleoherpetology*, vol. 1: *Lepospondyli*, pp. 133–148, Verlag Dr. Friedrich Pfeil, München, 1998.

Lysosome

A digestive structure found within virtually all types of animal cells. Lysosome sizes, microscopic appearances, and other properties vary among different cell types and circumstances owing, in part, to differences in their functions and states. Typical lysosomes (**Figs. 1** and **2**) are roughly spherical or elongate

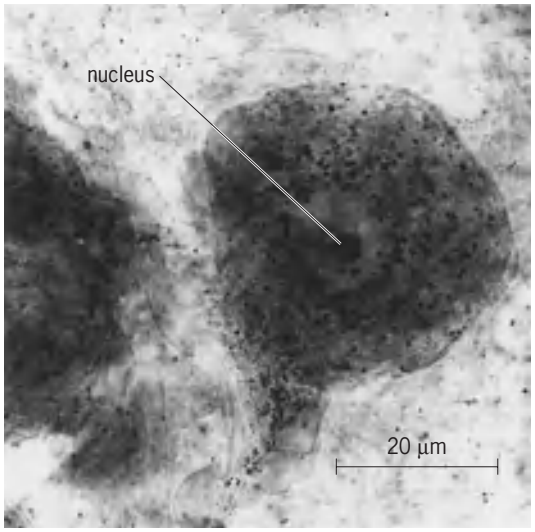

Fig. 1. Light microscope image of a nerve cell treated to show the locations of the lysosomal enzyme aryl sulfatase. The lysosomes show as small dark granules in the cell's cytoplasm. (*From E. Holtzman, Lysosomes in the physiology and pathology of neurons, in J. T. Dingle and H. B. Fell, eds., Lysosomes in Biology and Pathology, vol. 1, North Holland, 1969*)

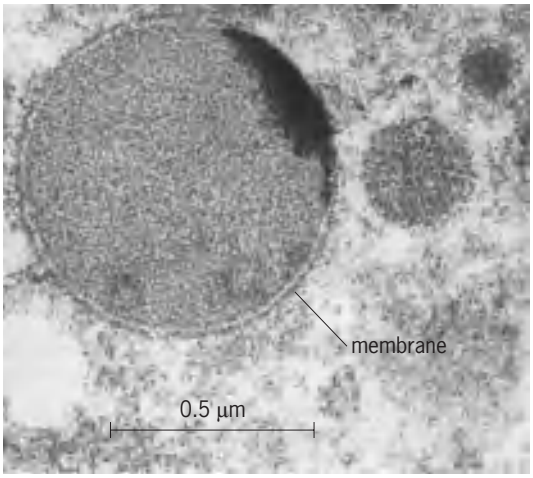

Fig. 2. Electron microscope image of a lysosome from the adrenal gland after treatment to reveal the lysosomal enzyme aryl sulfatase. The membrane bounding a mature secondary lysosome surrounds contents that appear moderately dark and heterogeneous. The very dark patch results from the cytochemical reaction that demonstrates the presence of the sulfatase enzyme. (*Reproduced, with permission, from E. Holtzman and R. Dominitz, Cytochemical studies of lysosomes, Golgi apparatus and endoplasmic reticulum in secretion and protein uptake by adrenal medulla cells of the rat, J. Histochem. Cytochem., 16:320–326, 1968*)

bodies with largest dimensions of 0.1–1 micrometer or greater; tens to hundreds are present in a single cell.

Enzymes. Each lysosome is bounded by a membrane and contains several dozen different species of digestive enzymes, each of which can sever (by hydrolysis) particular chemical bonds found in natural materials. Collectively, the enzymes are capable of splitting most of the large molecules of living systems—proteins, nucleic acids, complex lipids, polysaccharides—into their component parts, such as amino acids. The enzymes are made by the ribosomes of the rough endoplasmic reticulum and pass to the Golgi apparatus before being sorted to the lysosomes. In cells of higher animals, this sorting depends partly on the enzyme molecules being tagged with the sugar derivative mannose-6-phosphate. Specific membrane-associated proteins, known as mannose phosphate receptors, recognize and bind to molecules carrying the mannose-6-phosphate tag and segregate them into membrane-bounded transport vehicles, such as vesicles, for eventual delivery to digestive bodies. The transport vehicles for newly made hydrolases are sometimes called primary lysosomes to distinguish them from late endosomes and other secondary vacuoles, in which the enzymes have been mixed with materials undergoing digestion or with indigestible residues of earlier digestion. *See* CELL MEMBRANES; ENDOPLASMIC RETICULUM; ENZYME; GOLGI APPARATUS.

Most lysosomal enzymes function best in an acid environment, and the interior of the endosomal-lysosomal system displays a pH gradient (pH 5–6 or less) with the lysosome itself exhibiting the greatest acidity. This acidification is accomplished by an energy-dependent mechanism, a proton pump, which is built into the membrane surrounding the lysosome and which effects the transport of hydrogen ions into the lysosomes. *See* ION TRANSPORT.

Principal roles. Lysosomes digest materials taken into the cell from the outside (a process known as heterophagy) as well as other materials that originate in the cell's own cytoplasm (autophagy) [**Fig. 3**]. The materials to be digested are ultimately incorporated into the same membrane-bounded compartments as the lysosomal enzymes. Selective degradative products can pass out of the lysosome by crossing the membrane, but the enzymes cannot. This sequestration, which protects the cell, persists because the admixture of the enzymes and the materials to digest takes place through fusion of membrane-bounded compartments.

In heterophagy, the cell takes up particles or molecules by the process of endocytosis, engulfing them in membrane-bounded vesicles or vacuoles that are formed at the cell surface. The endocytosed material enters lysosomes via intermediate compartments known as endosomes. These are membrane-bounded bodies with moderately acidic interiors that possess limited digestive capacities at some stages of their life histories. One role of endosomes is the separation of endocytosed materials destined for lysosomal degradation from membrane-associated molecules which are sent back to the cell surface to be reused. Trafficking of molecules through the endosomal-lysosomal system may have significant influence over receptor-mediated signaling and signal transduction. *See* ENDOCYTOSIS.

Heterophagy is employed by many protozoa to obtain nutrients. In multicellular organisms, it is used to process some nutrients. For example, cholesterol can enter mammalian cells by uptake and lysosomal hydrolysis of particulate complexes of lipids and proteins that circulate in the blood. In higher animals, however, heterophagy is most prominently used by leukocytes and macrophages. These specialized cells endocytose invasive microorganisms and use endocytosis in clearing debris and disposing of dead or senescent cells that result from injury or from natural processes such as the regression of tadpole tails or the aging of red blood cells. *See* CELL SENESCENCE AND DEATH; PHAGOCYTOSIS.

Other cell types call upon elements of the endocytic and heterophagic systems for other specialized functions. The thyroid gland produces hormones by using heterophagy to release them from large proteins. Cells of the immune system employ endocytosis and endosome- or lysosome-related structures to fragment some of the foreign materials (antigens) that the immune system is designed to eliminate; the molecular fragments are then moved to the cell surface, where they take part in processes of antibody production. *See* CELLULAR IMMUNOLOGY.

In autophagy (Fig. 3), cells segregate regions of their own cytoplasm within compartments that come to be bounded by single membranes and to receive lysosomal enzymes. Autophagic lysosomes take part in the remodeling of cells as part of

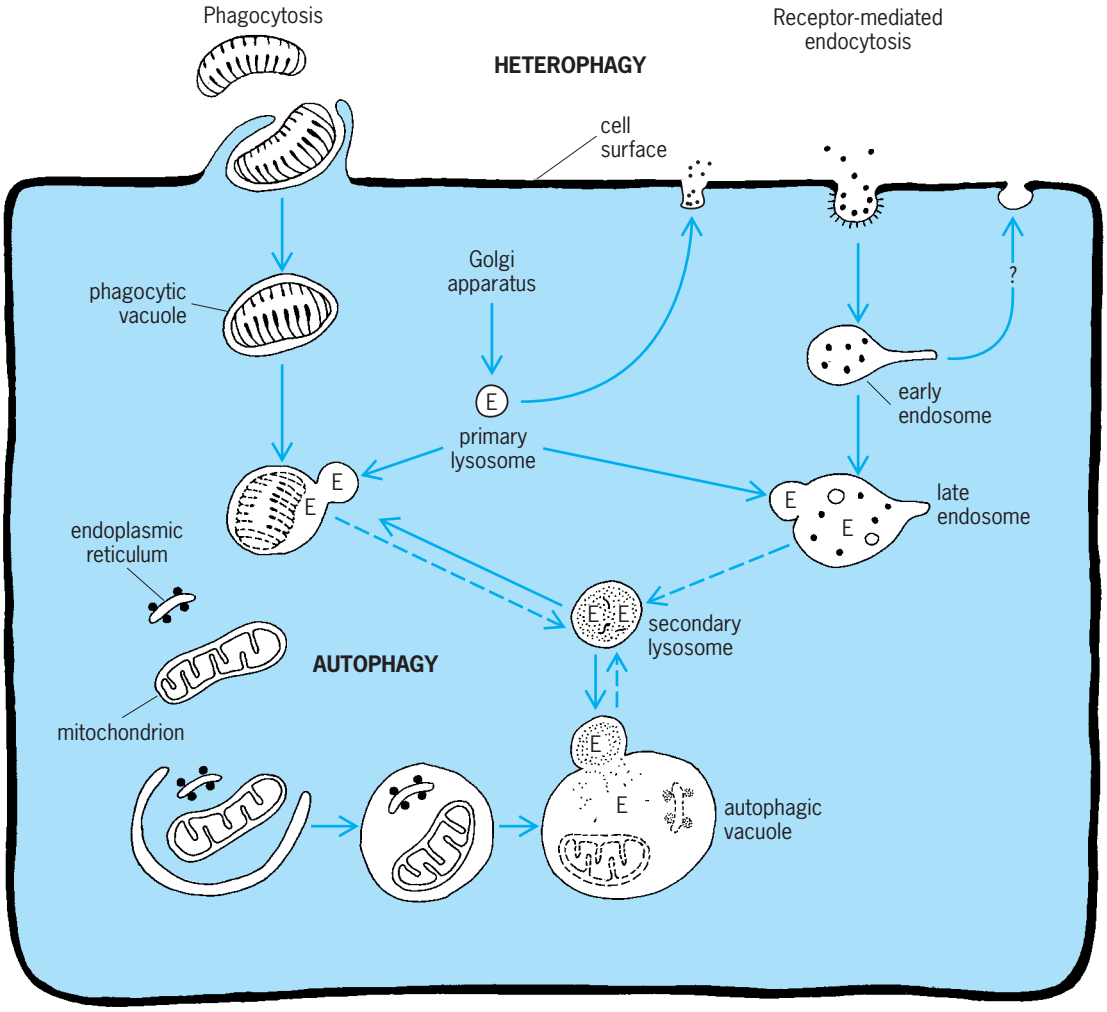

Fig. 3. Intracellular digestive phenomena. In one form of heterophagy, phagocyte cells take up particles into vacuoles arising from the plasma membrane. In receptor-mediated endocytosis, another form of heterophagy, molecules bind to receptors in the plasma membrane. After internalization by the cell, the molecules are delivered to an endosome where the receptors may be freed to return to the cell surface while other endocytosed molecules undergo digestion by lysosomal enzymes (E). In autophagy, the cell sequesters and digests portions of its own cytoplasm. Both autophagic and heterophagic lysosomes may evolve (broken arrows) into bodies containing lysosomal enzymes and indigestible residues.

the processes of development and during stressful circumstances. They also participate, along with nonlysosomal enzymes and heterophagic lysosomes, in normal turnover of the body's constituents—the balanced synthesis and destruction through which most molecules of most cells are replaced by new molecules over periods of hours to weeks.

Older lysosomes can recycle by fusing with newer digestive structures (Fig. 3). Lysosomes also sometimes fuse with the cell surface, releasing their contents to the extracellular space, and differential lysosomal enzyme secretion may be a common feature in many cell types. Cells known as osteoclasts use such fusion to help create extracellular digestive chambers for sculpting and remodeling of bone during growth and development. *See* SKELETAL SYSTEM.

Involvement in disease. Genetic defects in lysosomal enzymes and related proteins are known to be associated with a large number of rare disorders in humans and animals (such as Tay-Sachs disease and Niemann-Pick disease type C). Defective lysosomal function leads to storage of particular classes of molecules that cannot be degraded and, in long-lived cells such as neurons, to complex pathogenic cascades with widespread impact on endosomal-lysosomal function, membrane trafficking, and signal transduction. Such disorders are most often fatal. Lysosomes or prelysosomal structures also have been "adopted" as intracellular homes by certain pathogenic microorganisms that avoid or survive the attacks of the lysosomal system. Some strains of viruses, and toxins such as the one responsible for diphtheria, may use endosomes as their route of entry into the cell, penetrating through the endosomal membrane into the surrounding cytoplasm.

Plants, yeast, and fungi. In plant cells, yeast, and fungi, many of the structures that are known as vacuoles are closely related to lysosomes in that they are acidified and contain degradative enzymes. In some plant cells, a single vacuole is the largest intracellular compartment, occupying as much as 90%

of the cell's volume. Less is known about the digestive functions of vacuoles than of the lysosomes of animal cells, but vacuoles can be used for extensive storage of amino acids, proteins, pigments, and other molecules. In addition, they play a central role in the osmotic regulation of water balances, which is crucial to plant cell architecture and growth. *See* CELL (BIOLOGY); VACUOLE.

Eric Holtzman; S. U. Walkley

Bibliography. E. Holtzman, *Lysosomes*, 1989; J. B. Lloyd and R. W. Mason (eds.), Biology of the Lysosome, *Subcellular Biochemistry*, vol. 27, 1996; B. Storrie and R. F. Murphy (eds.), *Endosomes and Lysosomes*: *A Dynamic Relationship*, 1993.

Lysozyme

An enyme that was first identified and named by Alexander Fleming, who recognized its bacteriolytic properties. It has been designated muramidase, since it is known to facilitate the hydrolysis of a β-1-4-glycosidic bond between *N*-acetylglucosamine and *N*-acetylmuramic acid in bacterial cell walls; it also hydrolyzes similar glycosidic bonds in fragments of chitin. In recognition of this enzyme function, the International Enzyme Commission assigned the official chemical name of *N*-acetylmuramide glycanohydrolyase, coded as 3.2.1.17.

The most detailed studies have been performed on hen egg-white lysozyme, because this product is readily available. However, enzymes possessing lysozyme activity have been found in bacteria, bacteriophages, and plants and in human leukocytes, nasal secretions, saliva, and tears. People with a certain type of monocytic leukemia excrete large quantities of lysozyme in their urine. This excretion is presumed to be secondary to the synthesis of the enzyme by the large mass of leukemic cells.

The three-dimensional structure of the protein has been defined by x-ray crystallography. Additional data are available for the amino acid sequence of human lysozyme and also for a bacteriophage lysozyme. These results have given rise to speculation concerning the origin of the lysozyme gene during evolution.

Amino acid sequence. The complete amino acid sequence of hen egg lysozyme was determined independently by P. Jollès and by R. E. Canfield with their respective coworkers. The molecule consists of a single polypeptide chain containing 129 amino acid residues. The complete structure is illustrated in **Fig. 1**. This was determined by a technique which involved splitting the reduced alkylated lysozyme molecule into small peptide fragments with the proteolytic enzymes trypsin, chymotrypsin, and pepsin. The precise sequence of amino acids in each of the isolated fragments was then determined by the Edman degradation technique.

There are eight half-cystine residues in hen egg lysozyme, arranged in the native molecule as four cystine bridges. M. Inouye and A. Tsugita reported the amino acid sequence of a lysozyme synthesized by T4 bacteriophage that facilitates the penetration of phage deoxyribonucleic acid (DNA) through the bacterial cell wall. The protein is a single polypeptide chain containing 160 amino acid residues and no cystine bridges.

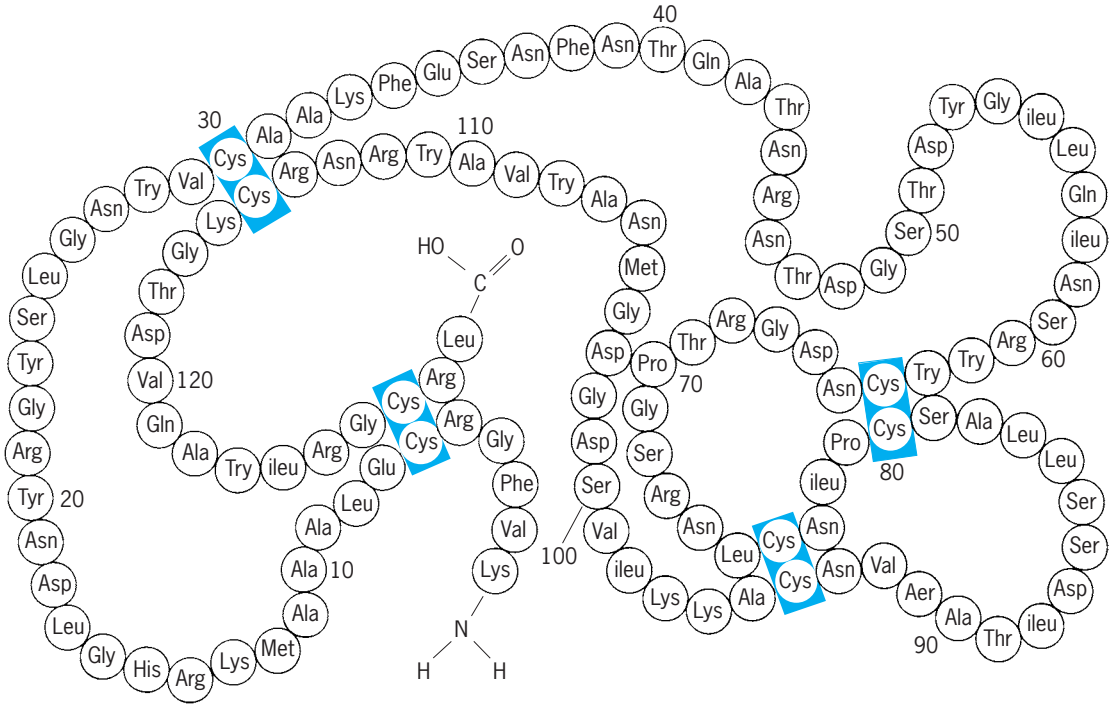

Fig. 1. Amino acid sequence of hen egg lysozyme. Four rectangles indicate cystine bridges. Ala, alanine; Arg, arginine; Asn, asparagine; Asp, aspartic acid; Cys, half-cystine; Gln, glutamine, Glu, glutamic acid; Gly, glycine; His, histidine; Ilen, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Phi, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Try, tryptophan; Tyr, tyrosine; Val, valine. (*After R. E. Canfield and A. K. Liu, J. Biol. Chem., 240:1997, 1965*)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
(a)	Lys	Val	Phe	Gly	Arg	Cys	Glu	Leu	Ala	Ala	Ala	Met	Lys	Arg	His	Gly	Leu	Asp	Asn	Tyr	Arg	Gly	Tyr	Ser	Leu
(b)	Lys	Val	Phe	Glu	Arg	Cys	Glu	Leu	Ala	Arg	Thr	Leu	Lys	Arg	Leu	Gly	Met	Asp	Gly	Tyr	Arg	Gly	Ilu	Ser	Leu
(c)	Glu	Gln	Leu	Thr	Lys	Cys	Glu	Val	Phe	Arg	Glu	Leu	Lys	.....		Asp	Leu	Lys	Gly	Tyr	Gly	Gly	Val	Ser	Leu

Fig. 2. Amino acid terminal sequences for (*a*) hen egg lysozyme, (*b*) human lysozyme, and (*c*) α-lactalbumin from cow milk. Rectangles identify common structural features. Dotted line for milk α-lactalbumin at positions 14-15 indicates gene deletion. Abbreviations are as in Fig. 1.

A portion of the structure of human lysozyme, described by Canfield, is illustrated in **Fig. 2**.

X-ray crystallography. D. C. Phillips and coworkers constructed a model of hen egg lysozyme based on a detailed x-ray crystallographic study of the three-dimensional structure of the enzyme at a resolution that permitted discrimination of structural features that were separated by a distance of 2 angstroms (0.2 nanometer). This model agreed with the primary structure shown in Fig. 1, including the location of the disulfide bonds. The lysozyme molecule appears to be folded in a fashion that accommodates most of the hydrophobic amino acid side chains inside the molecule, where they are removed from contact with the aqueous solvent, while most of the polar groups are distributed at the surface.

A unique aspect of the hen egg lysozyme model is the presence of a long cleft in the surface of the molecule. A trisaccharide inhibitor of lysozyme (tri-*N*-acetylglucosamine) can be bound in the cleft. This binding occurred at the same location when the inhibitor was diffused into crystals of lysozyme and when lysozyme was crystallized in the presence of the inhibitor. This unique binding of the inhibitor in the cleft may serve to localize the region of the active site of lysozyme. Of the various ways in which a model of the enzyme's substrate can be fitted into the region of the active site, the most favorable arrangement places the β-1-4-glycosidic bond that is hydrolyzed by lysozyme directly between the carboxyl group of the aspartic acid in position 52 and the carboxyl group of the glutamic acid in position 35 (Fig. 1). Structural observations indicate that these two amino acids have an important function in the mechanism of action of lysozyme.

Evolution. Certain enzyme functions appear to be widely distributed in nature. The amino acid sequences of proteins possessing these functions reflect changes that have occurred in the course of evolution. The structures of lysozymes from three sources, distant in evolution, have been carefully examined. Hen egg lysozyme (Fig. 1) has no structural elements in common with bacteriophage lysozyme. Thus it must be concluded that these two enzymes emerged in evolution completely independent of each other. Preliminary studies of the structure of human lysozyme reveal considerable similarity to the structure of hen egg lysozyme. In fact, the resemblance is so great, including the strategically placed aspartic and glutamic acids in the active site, that it can be concluded that these proteins evolved from the same gene and have an essentially identical mechanism of action. Figure 2 illustrates this similarity, or homology, for the first 25 amino acid residues of these two lysozymes, where 18 of the first 25 are identical.

The amino acid composition of α-lactalbumin, a protein in cow's milk, is quite similar to that of hen egg lysozyme; nearly half of the amino acid positions in these two proteins are identical (Fig. 2). This is surprising since hen egg lysozyme operates to hydrolyze a β-1-4-glycosidic bond between amino sugars in bacterial cell walls, while α-lactalbumin facilitates the synthesis of β-1-4-glycosidic bond between glucose and galactose to form lactose. One possible explanation is that the lysozyme gene was accidentally duplicated at some point in evolution and that subsequently one of the genes acquired lactose synthetase function. This new function seems to have appeared late in evolution, and is coincident with the development of breast feeding.

It is postulated from a comparison of the amino acid sequences of hen egg lysozyme, human lysozyme, and α-lactalbumin (Fig. 2) that a "deletion" occurred during evolution in the α-lactalbumin gene with a resulting loss of information for two amino acids near position 13. In addition, positions 10, 12, and 19 in human lysozyme and α-lactalbumin are identical, so it is possible to see remnants of a common ancestral gene in all three proteins. These data illustrate the manner in which amino acid sequence information is being used as a molecular reflection of the paths of evolution. *See* PROTEINS, EVOLUTION OF.

Experimental approach. One approach to the problem of determining the mechanism of action of the enzyme is to determine the structure of other lysozymes that derive their catalytic activity from amino acid sequences different from that of hen egg lysozyme. An examination of the structure of the active sites in these other lysozymes will permit the exploration of the common features that are essential for enzyme activity. If the present hypothesis based on the location of the active site for the mechanism of lysozyme action is correct, then the finding of two carboxyl groups immediately adjacent to the substrate bond that is hydrolyzed is to be expected. The data for human lysozyme appear to follow this pattern.

The synthesis of various substrate analogs will also be important in future work, because as these are fitted into the model of the lysozyme cleft, it should be possible to predict their chemical behavior and

verify the predictions by experiments. *See* BACTERIA; CRYSTALLOGRAPHY; ENZYME; LYSIN; LYTIC REACTION; POLYSACCHARIDE. Robert E. Canfield

Bibliography. K. Hayashi and N. Sakamoto (eds.), *Dynamic Analysis of Enzyme Systems*, 1985; C. P. Hollenberg and H. Sahm (eds.), *Microbial Genetical Engineering and Enzyme Technology*, 1987; S. Smith-Gill and E. Sercarz (eds.), *The Immune Response to Structurally Defined Proteins*: *The Lysozyme Model*, 1989.

Lytic infection

Infection of a bacterium by a bacteriophage with subsequent production of more phage particles and lysis, or dissolution, of the cell. The viruses responsible are commonly called virulent phages. Lytic infection is one of the two major bacteriophage–bacterium relationships, the other being lysogenic infection. *See* BACTERIOPHAGE; COLIPHAGE; LYSOGENY. Philip B. Cowles

Lytic reaction

A term used in serology to describe a reaction that leads to the disruption or lysis of a cell. The best example is the lysis of sheep red blood cells by specific antibody and complement in the presence of Ca^{2+} (calcium ion) and Mg^{2+} (magnesium ion), a reaction that forms the indicator system of the standard Wassermann test for syphilis, as well as other complement-fixation reactions. In this example, lysis results in the release of cellular hemoglobin into the medium; the reaction may be followed by visual or instrumental estimation of the decreased cell turbidity or the increased color of the medium due to the free hemoglobin. The initiation of lysis by complement can apparently proceed after the attachment of only one molecule of IgM or two molecules of IgG antibody to the red blood cell. IgM and IgG are both immunoglobulins. *See* IMMUNOGLOBULIN.

Pfeiffer reaction. This is an example of a bacteriolytic reaction, carried out with gram-negative bacteria, such as cholera vibrios, which are especially susceptible to the lytic action of specific antibody, complement, and Mg^{2+} and Ca^{2+}. This reaction may be carried out in the living subject (Pfeiffer's system) or in the test tube, and is a valuable aid in diagnosis. An analogous reaction is given by the spirochetes of relapsing fever. Although demonstrable instances of bacteriolysis occur, it is not clear whether all the bactericidal actions of antibody and complement are necessarily preceded by lysis.

Phagocytic reaction. The engulfment of microorganisms by phagocytic cells is often accompanied by their digestion and lysis. The initial stages of this reaction are aided by specific antibody and by complement. It has been estimated that no more than 8 molecules of IgM or 2200 molecules of IgG are needed for the phagocytosis of a typical enteric bacterium. *See* PHAGOCYTOSIS.

Cytotoxic assays. A variety of reactions have come into use because of the necessity of detecting tissue incompatibilities before transplantation. The cell damage, including permeability changes resulting from the actions of antibody and complement, can be estimated from the differential staining by dyes or by assay of the release of tracers. *See* COMPLEMENT; COMPLEMENT-FIXATION TEST; SEROLOGY. Henry P. Treffers

Bibliography. P. Agre and J. C. Parker, *Red Blood Cell Membranes*: *Structure*, *Function*, *Clinical Implications*, 1989; J. Bryant, *Laboratory Immunology and Serology*, 2d ed., 1986; M. L. Turgeon, *Immunology and Serology in Laboratory Medicine*, 2d ed., 1996.

M

Macadamia nut — Metazoa

Macadamia nut

The fruit of a tropical evergreen tree, *Macadamia ternifolia*, native to Queensland and New South Wales and now grown commercially in Australia and Hawaii. The trees grow to 50 ft (15 m) in height and have dense foliage of glossy, leathery leaves. They bear many small white or pinkish flowers in drooping racemes, each of which may mature from 1 to 20 fruits. These consist of a leathery outer husk (pericarp) which splits along one side at maturity, freeing the very hard-shelled, nearly round seed or nut about an inch in diameter. Two types of nuts are recognized: the most important commercially having a smooth shell and the other having a rough shell and sometimes referred to another species, *M. integrifolia* (see **illus.**).

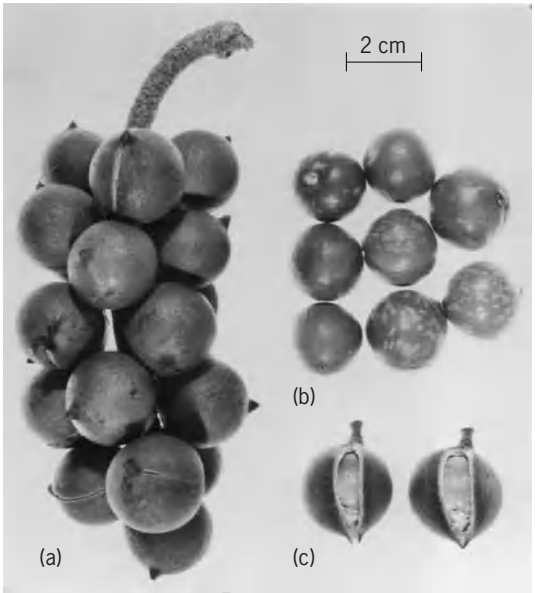

***Macadamia integrifolia*. (*a*) Mature nuts. (*b*) Nuts without husks. (*c*) Nuts in husk showing method of dehiscence. (*From R. A. Jaynes, ed., Handbook of North American Nut Trees, Humphrey Press, 1969*)**

Successful culture of the trees as an orchard crop requires a tropical climate, good well-drained soil, abundant rainfall (50 in. or 1.3 m or more), and soil management to control competing vegetation. Protection from strong winds is essential. Trees have been grown successfully in Florida and California, but the nuts are commercially important only in Hawaii.

Macadamia nuts ripen over a period of several months: As immature nuts they are of little value; the nuts are allowed to mature on the trees and fall to the ground, where they are picked up by hand, machine-hulled, dried, and stored for processing. After machine-cracking, the kernels are graded by floating in water, where the high-quality kernels float and are skimmed off. They are then roasted, usually in oil, salted, and packed, with glass jars being used for best-grade nuts. The processed nuts are of high quality and find a ready market. *See* NUT CROP CULTURE.

Laurence H. MacDaniels

Mach number

In the flow of a fluid, the ratio of the flow velocity, V, at a given point in the flow to the local speed of sound, a, at that same point. That is, the Mach number, M, is defined as V/a. In a flowfield where the properties vary in time and/or space, the local value of M will also vary in time and/or space. In aeronautics, Mach number is frequently used to denote the ratio of the airspeed of an aircraft to the speed of sound in the freestream far ahead of the aircraft; this is called the freestream Mach number. The Mach number is a convenient index used to define the following flow regimes: (1) subsonic, where M is less than 1 everywhere throughout the flow; (2) supersonic, where M is greater than 1 everywhere throughout the flow; (3) transonic, where the flow is composed of mixed regions of locally subsonic and supersonic flows, all with local Mach numbers near 1, typically between 0.8 and 1.2; and

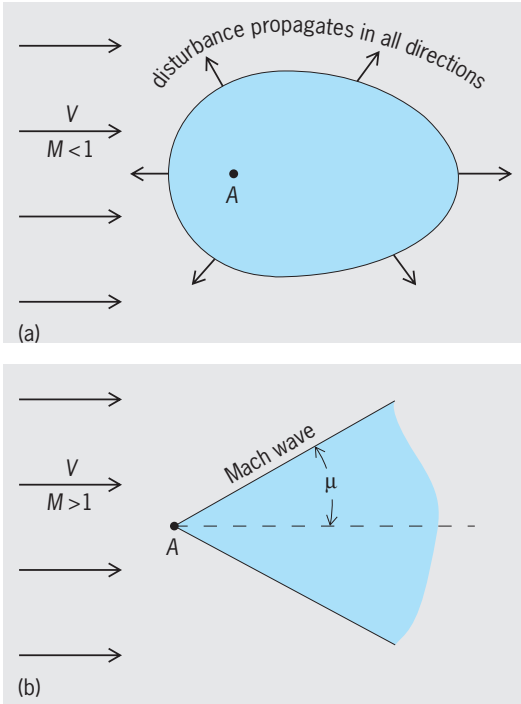

Propagation of disturbances from a stationary source in (*a*) subsonic flow and (*b*) supersonic flow.

(4) hypersonic, where (by arbitrary definition) M is 5 or greater.

Perhaps the most important physical aspect of Mach number is in the completely different ways that disturbances propagate in subsonic flow compared to that in a supersonic flow. Consider an initially uniform flow from left to right in the plane of this page, as shown in the **illustration**. First consider a subsonic flow. Imagine that you jab the subsonic flow in illus. *a* with a needle perpendicular to the page at point A. The weak disturbance created will propagate in all directions at the local speed of sound relative to the flow. Because the initially uniform flow is moving at less than the speed of sound, the disturbance will work its way upstream, eventually covering the entire extent of the flow. The shaded region in illus. *a* is a qualitative sketch of the subsonic flow region affected by the disturbance at some instant after you have jabbed it at point A. In illus. *b*, when you jab the supersonic flow with the needle at point A, once again the disturbance will want to propagate in all directions at the local speed of sound. However, because the initially uniform flow is already moving at a velocity greater than the speed of sound, the disturbance will not be able to work its way upstream; rather, it will be swept downstream, affecting only the conical shaded region. This region is bounded by a Mach wave at an angle μ to the freestream direction. The angle μ is defined as the Mach angle, given by $\mu = \sin^{-1}(1/M)$. If the disturbance at point A is strong (not weak, as in the example), the disturbed region is bounded by an oblique shock wave which is at an angle larger than μ. Severe, almost discontinuous changes take place across shock waves; for example, the pressure, density, and temperature increase across a shock wave, sometimes dramatically. Shock waves are a ubiquitous aspect of supersonic flows. Because of the phenomena in the illustration, the physical aspects of subsonic and supersonic flows are quite different.

The physical importance of the ratio of flow velocity to the speed of sound was recognized and studied as early as 1887 by the scientist and philosopher Ernst Mach in Prague. The term "Mach number" for this ratio was coined in 1929 by Jakob Ackeret in Zurich, and did not find its way into the English literature until the late 1930s. *See* COMPRESSIBLE FLOW; SHOCK WAVE; SONIC BOOM; SUPERSONIC FLIGHT.

John D. Anderson, Jr.

Bibliography. J. D. Anderson, Jr., *Fundamentals of Aerodynamics*, 2d ed., McGraw-Hill, 1991; J. D. Anderson, Jr., *A History of Aerodynamics*, Cambridge University Press, 1997; J. D. Anderson, Jr., *Modern Compressible Flow, with Historical Perspective*, 2d ed., McGraw-Hill, 1990; L. W. Reithmaier, *Mach 1 and Beyond*, McGraw-Hill, 1994.

Machinability of metals

The ease and economy with which a metal may be cut under average conditions. Frequently no truly quantitative assessment is made, but rather a rating or an index is established vis-à-vis a reference material. More quantitative comparisons are based on tool life. For example, maximum cutting speeds for a given tool life may be used as a rating of machinability. Alternatively, tool wear rate may be the basis for a machinability rating. Surface finish is sometimes used for assessing machinability.

Machining process. The wide range of metal-cutting processes may be represented, with some oversimplification, by the orthogonal cutting process (**Fig. 1**). A rectangular metal workpiece is machined by a tool with a face at a rake angle γ_c measured from the normal of the surface to be machined. A clearance angle Θ exists between the flank of the tool and the machined surface of the workpiece. The

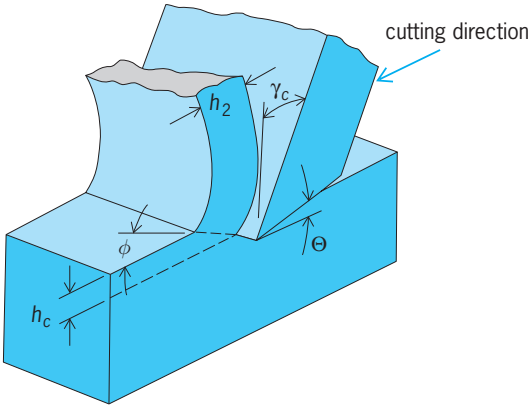

Fig. 1. Orthogonal cutting process: a simplified view of machining. (*After J. A. Schey, Introduction to Manufacturing Processes, McGraw-Hill, 1977*)

tool cuts into the workpiece to a depth h_c. However, the emerging chip is thicker than h_c. The chip thickness h_2 reflects intense shearing that takes place at an angle ϕ to the workpiece surface.

Cutting speed. The machining response of ductile metals is very sensitive to cutting speed (**Fig. 2**). Below about 0.06 ft/s (0.02 m/s), chips form discontinuously, metal chunks are lifted out of the surface, and the surface is scalloped or pockmarked (Fig. 2*b*). When the speed is in the 0.1 m/s range, chips are formed continuously, the shear zone is narrow, and the chip slides on the tool face (Fig. 2*b*). Under these conditions a cutting fluid can lubricate both the rake and flank faces of the tool. If the speed is increased to 1.0 ft/s (0.3 m/s) or so, sticking starts to occur at the tool-workpiece interface due to the increased heat generation. The metal begins to shear along a built-up edge of metal which is stuck to the tool face (Fig. 2*c*). The appearance of the built-up edge has several practical consequences. The effective rake angle becomes quite large, and energy consumption drops. The surface finish is poor, however, owing to the ill-defined tool tip and sporadic breakdowns in the builtup edge. At much higher speeds (above 3.2 ft/s or 1.0 m/s), heat generation is even more intense, and temperature increases of several hundred degrees Celsius may occur. Under these conditions the built-up edge disappears, and the chip makes full, sticking contact with the tool face (Fig. 2*d*). The chip moves by shearing along the so-called secondary shear zone. With the high temperature, workpiece strength may be so low that power consumption is lower in spite of the extensive shearing. The most serious problem associated with the high temperature may be the possibility of diffusional bonding occurring between the tool and workpiece. If alloying elements of the tool material diffuse into the workpiece, very rapid wear (called crater wear) develops.

Cutting fluid. The interaction of the tool and the workpiece is considerably affected by the presence of cutting fluids. Cutting fluid has two primary functions. First, as long as the cutting speed is slow, the cutting fluid can act as a lubricant between the chip and the tool face. Even at higher speeds some lubricating effect at the flank face may be present. Second, and perhaps more importantly, the cutting fluid serves as a coolant. In most instances the cutting fluid will be an emulsion of a lubricating phase (oil, graphite, and so on) in water, since water is the best heat-transfer medium readily available. Beyond lubrication and cooling, the cutting fluid can be used to flush out the cutting zone.

Surface quality and tool wear. Cutting speed thus affects the all-important machinability considerations of surface quality and tool wear. As implied by Fig. 2, surface finish is best with well-lubricated, moderately low-speed operation or high-speed cutting with no built-up edge. The high pressures and temperatures of operation, abetted by shock loading and vibrations, can lead to rapid tool wear. Tool wear is often sufficiently rapid to make tool replacement a major factor in machining economics. Tools must

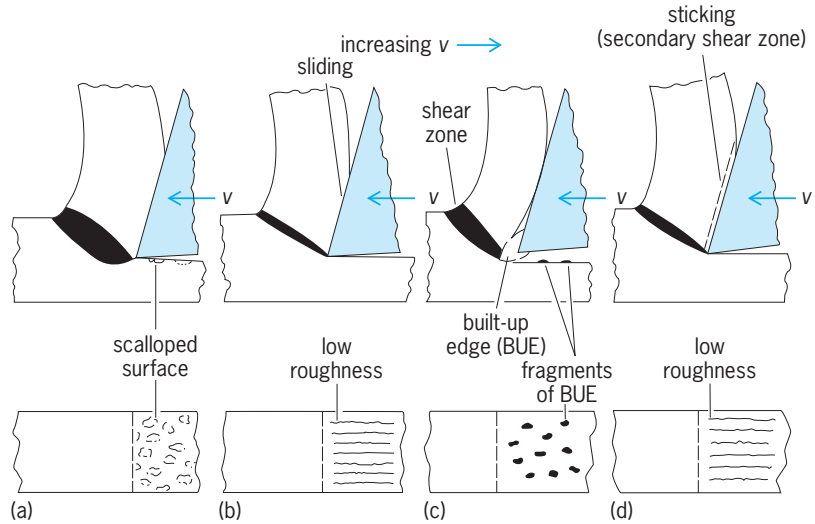

Fig. 2. Changes in chip formation and surface finish with increasing cutting speed: (*a*) 0.02 m/s; (*b*) 0.1 m/s; (*c*) 0.3 m/s; and (*d*) above 1.0 m/s. (*After J. A. Schey, Introduction to Manufacturing Processes, McGraw-Hill, 1977*)

be replaced when they break or when they have worn to the point of producing an unacceptable surface finish or an unacceptable degree of surface heating. Whereas cracking and chipping occur suddenly, other forms of wear occur gradually. Gradual wear is related to rubbing distance and temperature, and thus to speed v. It often follows the Taylor power law, $C = vt^n$, where C is the cutting speed leading specifically to 1 minute of tool life, t is actual tool life in minutes, and n is the Taylor exponent. Both C and n reflect the tool and workpiece combination.

Metal properties. A machining operation can be optimized to effect metal removal with the least energy, or the best surface, or the longest tool life, or a reasonable compromise among these factors. Even so, it remains that the optimum ease and economy of machining some alloys is vastly different from that of others, and some basic attributes of easily machined alloys can be set forth.

Toughness. To ensure that chip separation occurs after minimum sliding, low ductility is required. To minimize cutting force, low strength and low ductility combine to mean low toughness. Toughness is generally defined as energy per unit volume consumed en route to fracture. Ironically, materials of maximized toughness are desirable for most engineering applications and, thus, some of the most attractive alloys, such as austenitic stainless steels, are difficult to machine.

Adhesion. The degree to which the metal adheres to the tool material is important to its machinability. Actually this attribute can work to advantage or disadvantage. If diffusion results, the tool can be weakened, and rapid wear occurs. Otherwise, high adhesion will stabilize the secondary shear zone.

Workpiece second phases. Small particles or inclusions in the metal can have a marked effect on machinability. Hard sharp oxides, carbides, and certain intermetallic compounds abrade the tool and accelerate

tool wear. On the other hand, soft second phases are beneficial because they promote localized shear and chip breakage.

Thermal conductivity. In some cases, workpiece thermal conductivity can be important to machinability. A low thermal conductivity generally results in high shear-zone temperature. This can be advantageous in reducing the strength of the metal or in softening second-phase particles. Of course, if adhesion and diffusional depletion of tool alloy content result, the high temperature is a problem. The workpiece temperature can be managed by cutting-speed and cutting-fluid manipulation.

Alloy systems. Commercial alloys can be grouped into two categories, namely those designed for ease of machining (so-called free machining grades), and the vast majority which are of widely varying but generally less than optimum machinability. Considering these latter, ordinary alloys, it can be shown that their machinability may be considerably improved by metallurgical operations which limit strength or ductility or both. Of course, it is not often possible to simultaneously reduce both strength and ductility. Even so, machinability often can be improved by grossly reducing one or the other property.

Consider the range of plain carbon steels represented in **Fig. 3**. In each case, the steel can be produced in three relevant forms: fully annealed (pearlitic), spheroidized, or cold-worked. For steel in the 0.20% C range, the annealed and spheroidized conditions involve too much ductility for optimum machinability. Better machinability can be had by reducing the ductility through cold work, even if strength is increased. In the 0.45% C range, the cold-worked material is too strong, and the intermediate strength and ductility of the annealed condition is optimum. In the 0.70% C range, the large quantities of lamellar carbide in the annealed pearlitic form are too abrasive and promote tool wear. Better machinability is possible with the globular carbides and lower strength of the spheroidized condition despite its high ductility. *See* STEEL.

Similar cases may be cited for other alloy systems, and the following principles are fairly general: (1) soft, ductile materials are more machinable when work-hardened; (2) hard materials with hard second phases are best machined in the well-annealed or overaged condition where globular particle shapes are developed; (3) moderate-strength age-hardenable alloys (such as age-hardenable Al alloys) are best machined in the aged and strengthened condition. *See* ALLOY.

Cast iron. Cast iron is a "natural" free-machining material. Gray cast iron with its large graphite flakes produces short chips without extra alloying or further process control. Moreover, the graphite itself serves as a lubricant. However, the machined surface is somewhat roughened by the breaking out of graphite particles. The machinability of gray iron may be further enhanced with a subcritical anneal (1 h at about 1380°F or 750°C) to graphitize the carbide that exists in the pearlite lamellae. The resultant structure is totally iron and flake graphite. *See* CAST IRON.

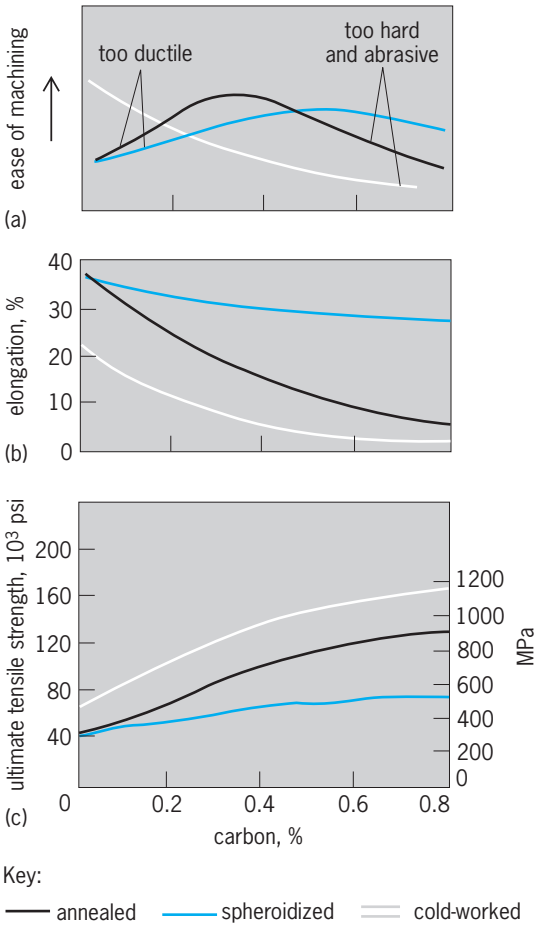

Fig. 3. The ease of machining carbon steels as a function of their metallurgical condition. (*a*) Ease of machining versus carbon content. (*b*) Elongation versus carbon content. (*c*) Ultimate tensile strength versus carbon content. (*After J. A. Schey, Introduction to Manufacturing Processes, McGraw-Hill, 1977*)

Lead additive. One of the most universal free-machining additives is lead. Lead is insoluble in iron, copper, aluminum, and alloys thereof. It can be finely dispersed as a soft second phase, which helps to break up the chips and, furthermore, lubricates the tool-workpiece interface. *See* LEAD ALLOYS; METAL, MECHANICAL PROPERTIES OF; PLASTIC DEFORMATION OF METAL.

Roger N. Wright

Bibliography. American Society for Metals, *Metals Handbook*, 9th ed., vol. 16: *Machining*, 1989; R. Baril, *Modern Machining Technology*, 1987; J. N. Harris, *Mechanical Working of Metals: Theory and Practice*, 1983; J. A. Schey, *Introduction to Manufacturing Processes*, 2d ed., 1987, rev. 2000; R. Stevenson and D. A. Stephenson (eds.), *Materials Issues and Physics of Machining Processes*, 1992.

Machine

A combination of rigid or resistant bodies having definite motions and capable of performing useful work. The term mechanism is closely related but applies only to the physical arrangement that provides for the definite motions of the parts of a machine. For

example, a wristwatch is a mechanism, but it does no useful work and thus is not a machine. Machines vary widely in appearance, function, and complexity from the simple hand-operated paper punch to the ocean liner, which is itself composed of many simple and complex machines. No matter how complicated in appearance, every machine may be broken down into smaller and smaller assemblies, until an analysis of the operation becomes dependent upon an understanding of a few basic concepts, most of which come from elementary physics. *See* MECHANISM; SIMPLE MACHINE. Richard M. Phelan

Bibliography. K. Sherwin, *Engineering Design for Performance*, 1982; J. E. Shigley and L. Mitchell, *Mechanical Engineering Design*, 6th ed., 2000.

Machine design

Application of science and invention to the development and construction of machines. An understanding of the basic laws of nature is essential to a proper perspective in the approach to machine design. Knowledge of the past development of machine elements makes possible their effective application. Inventiveness consists of producing new combinations of old elements or, where extreme need arises, of exercising genius either in breaking the bounds of convention, or in evolving new principles not hitherto applied or known.

In machine design, accomplishment takes on two forms: one is the drawings and blueprints, which completely describe the machine, and the other is the assembled product. In addition, most machines go through periods of evolution, and later models may show little outward similarity to the original design. *See* ENGINEERING DRAWING.

Machine design consists of the conception of a machine that will meet a specific need. Before constructing a machine to fulfill the need, the designer must thoroughly understand the application, and mentally modify an old machine or devise a new machine as required. A certain cost for the machine and a probable time for its construction are estimated. These estimates consider the materials required, the equipment necessary for its manufacture and testing, and the final operation in meeting the original need. If the machine is desirable, construction of the unit follows. In time the machine may become obsolete due to advances in the technology; it may then be rebuilt or replaced, possibly under the direction of the original designer.

The working tools in machine design are an understanding of the basic elements of machines that have been developed in the past and a thorough knowledge of the mechanical fields of science including mathematics, physics, statics and dynamics, strength of materials, kinematics, mechanisms, and the laboratories associated with them. *See* DYNAMICS; KINEMATICS; MATHEMATICS; MECHANISM; PHYSICS; STATICS; STRENGTH OF MATERIALS. James J. Ryan

Bibliography. R. H. Creamer, *Machine Design*: *Mechanical Engineering Technology*, 3d ed., 1983.

Machine elements

Elementary mechanical parts used as building blocks for the construction of most devices, apparatus, and machinery. The gradual development of these building blocks, following the invention of the roller or wheel and the arm or lever in ancient times, brought about the Industrial Revolution, starting with the assembly of James Watt's engine for harnessing the force of steam and proceeding into the advanced mechanization of present automatic control.

The most common example of a machine element is a gear, which, fundamentally, is a combination of the wheel and the lever to form a toothed wheel (see **illus.**). The rotation of this gear on a hub or shaft drives other gears which may rotate faster or slower, depending upon the number of teeth on the basic wheels. The material from which the gear is made establishes its strength, and the hardness of its surface determines its resistance to wear. Knowledge of the forces on the gear makes possible the determination of its size. Changes in its shape allow modifications in its use. These applications, as in most machine elements, have developed into many standard forms, such as spur, bevel, helical, and worm gears. Each of these forms has required the development of a special technology for its production and use. *See* GEAR.

Other fundamental machine elements have evolved from wheels and levers. A wheel must have a shaft on which it may rotate. The wheel is fastened to the shaft with a key, and the shaft is joined to other shafts with couplings. The shaft must rest in bearings, such as journal bearings, ball bearings, or roller bearings. The shaft may be started by a clutch or stopped with a brake. It may be turned by a pulley with a flat belt, a V belt, or a chain or a rope connecting it to a pulley on a second shaft. The supporting structure may be assembled with bolts or rivets or by welding. Proper application of these machine elements depends upon a knowledge of the forces on the structure and the strength of the materials employed. In the design, calculations must accommodate the forces to the materials in the simplest construction. Other machine elements have been evolved whose applications are more specific in construction. *See* ANTIFRICTION BEARING; BRAKE; CAM

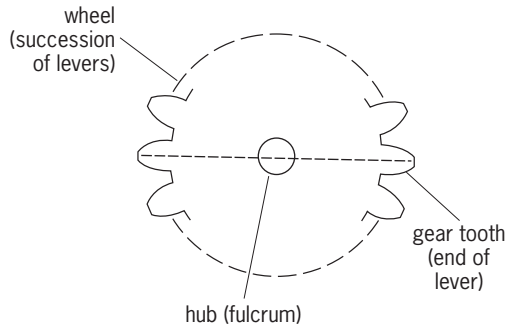

The gear, a machine element which combines features of the wheel and the lever.

MECHANISM; CLUTCH; COUPLING; FLYWHEEL; FOUR-BAR LINKAGE; GOVERNOR; MACHINE KEY; SCREW FASTENER; SHAFTING; SPRING (MACHINES); WHEEL AND AXLE.

Machine parts which are commonly used have been developed into standardized designs. Manufacturing specialists have concentrated upon the development of standard elements and have mass-produced these parts with a high degree of perfection at reduced cost. Standard elements, as applied in machine design, may be modified as desired, although certain ones, through the hazards to safety under improper use, must be modified only in line with the requirements of codes established by regulating bodies. *See* DESIGN STANDARDS. James J. Ryan

Bibliography. E. Oberg et al., *Machinery's Handbook*, 26th ed., 2000.

Machine key

The most common function of a key is to prevent relative rotation of a shaft and the member to which it is connected, such as the hub of a gear, pulley, or crank. Many types of keys are available, and the choice in any installation depends on such factors as power requirements, tightness of fit, stability of connection, and cost. For light power requirements a setscrew may be tightened against the round shaft or against a flat spot on the shaft. For most requirements a positive connection, such as by a key, is necessary. A setscrew is frequently used to seat the key and to prevent axial motion.

Square keys are common in general industrial machinery (**Fig. 1***a*). Flat keys are used where added stability of the connection is desired, as in machine tools (Fig. 1*b*). Square or flat keys may be of uniform cross section or they may be tapered. In tapered keys the width is uniform and the height of the key tapers. Tapered keys may have gib heads to facilitate removal (Fig. 1*c*).

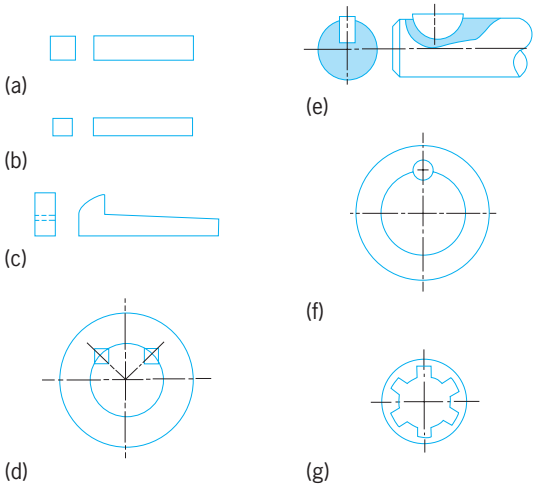

Fig. 1. Types of keys. (*a*) Square key. (*b*) Flat key. (*c*) Taper key with gib head. (*d*) Kennedy key. (*e*) Woodruff key. (*f*) Round key. (*g*) Spline fitting. (*After P. H. Black and O. E. Adams, Jr., Machine Design, 3d ed., McGraw-Hill, 1968*)

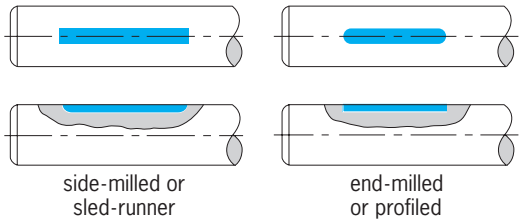

Fig. 2. Types of keyways. (*After P. H. Black and O. E. Adams, Jr., Machine Design, 3d ed., McGraw-Hill, 1968*)

The Kennedy key is used for heavy duty and consists of two keys driven 90° apart (Fig. 1*d*). The hub is bored to fit the shaft and is then rebored slightly off center. The keys force the shaft and hub into concentric position.

The Woodruff key requires a key seat formed by a special side-milling cutter (Fig. 1*e*). This key will align itself in the key seat. It has the disadvantage of weakening the shaft more than a straight key.

The round key, or pin, introduces less stress concentration at the key seat in the shaft and is satisfactory except for the necessity of drilling the hole to accommodate the pin after assembly of the hub and shaft (Fig. 1*f*). This may be a disadvantage in production and prevents interchangeability.

A spline fitting is composed of a splined shaft formed by milling and a mating hub with internal splines formed by broaching (Fig. 1*g*). The splines in reality are a number of keys integral with the shaft. Splined fittings are adaptable to mass production and are used where radial space must be conserved.

Straight-side splines are being replaced at an increasing rate by stub involute splines. These splines have the advantages of greater strength, a self-centering feature, and production economy.

Keyways for straight keys are formed either by a side-milling cutter which forms a sled-runner keyway, or by an end-milling cutter which forms profiled keyways (**Fig. 2**). The sled-runner keyway requires a longer space between the end of the key and the end of the keyway than does an end-milled keyway. This favors the end-milled keyway in locations near a shoulder. However, the end-milled keyway reduces the endurance limit of a shaft more than does the sled-runner type.

Feather keys are used where it is necessary to slide a keyed gear or pulley along the shaft. The key is generally tight in the shaft and with clearance between key and hub keyway. When a gear or pulley must be moved axially along the shaft while power is being transmitted, it is desirable to provide for a minimum of force necessary to move the hub along the shaft. The use of two feather keys equally spaced reduces the necessary axial force to half that for one key. *See* SHAFTING. Paul H. Black

Bibliography. P. H. Black and O. E. Adams, Jr., *Machine Design*, 3d ed., 1968; A. D. Deutschman et al., *Machine Design: Theory and Practice*, 1975; E. Oberg et al., *Machinery's Handbook*, 26th ed., 2000; J. Shigley, *Machine Design*, 1989.

Machinery

A group of parts arranged to perform a useful function. Normally some of the parts are capable of motion; others are stationary and provide a frame for the moving parts. The terms machine and machinery are so closely related as to be almost synonymous; however, machinery has a plural implication, suggesting more than one machine. Common examples of machinery include automobiles, clothes washers, and airplanes; machinery differs greatly in number of parts and complexity.

Some machinery simply provides a mechanical advantage for human effort. Other machinery performs functions that no human being can do for long-sustained periods. A jackscrew does nothing until a person pulls on a lever; then they are able to move objects many times their weight. Conversely, an internal combustion engine can run unattended for hours, requiring only the press of a button to start it.

The need for machinery usually stems from a desire to do a job at less cost. Evolution of machinery for a certain function may be gradual or rapid. If only small quantities of a product are needed, it is likely that machinery used in making the product will not change rapidly or possess the highest degree of automation. On the other hand, machinery used in making automobiles has evolved into some of the most complex automatic machines in existence. *See* AUTOMATION; MACHINE; MACHINE DESIGN; MECHANICAL ENGINEERING; SIMPLE MACHINE.

Robert S. Sherwood

Machining

Any one of a group of operations that change the shape, surface finish, or mechanical properties of a material by the application of special tools and equipment. Machining almost always is a process where a cutting tool removes material to effect the desired change in the workpiece. Typically, powered machinery is required to operate the cutting tools. *See* PRODUCTION METHODS.

Although various machining operations may appear to be very different, most are very similar: they make chips. These chips vary in size from the long continuous ribbons produced on a lathe to the microfine sludge produced by lapping or grinding. These chips are formed by shearing away the workpiece material by the action of a cutting tool. Cylindrical holes can be produced in a workpiece by drilling, milling, reaming, turning, and electric discharge machining. Rectangular (or nonround) holes and slots may be produced by broaching, electric discharge machining, milling, grinding, and nibbling; and cylinders may be produced on lathes and grinders. Special geometries, such as threads and gears, are produced with special tooling and equipment utilizing the same turning and grinding processes mentioned above. Polishing, lapping, and buffing are variants of grinding where a very small amount of stock is removed from the workpiece to produce a high-quality surface.

In almost every case, machining accuracy, economics, and production rates are controlled by the careful evaluation and selection of tooling and equipment. Speed of cut, depth of cut, cutting-tool material selection, and machine-tool selection have a tremendous impact on machining. In general, the more rigid and vibration-free a machining tool is, the better it will perform. Jigs and fixtures are often used to support the workpiece. Since it relies on the plastic deformation and shearing of the workpiece by the cutting tool, machining generates heat that must be dissipated before it damages the workpiece or tooling. Coolants, which also act as lubricants, are often used.

Boring. This machining operation increases the size of an existing hole in a workpiece. The usual purpose of boring is to produce a hole with an accurate diameter and good surface finish. Boring can be performed on a special machine or a lathe, with either the workpiece or the boring tool being on a movable table. A rotating spindle, holding either a single-point cutting tool or the workpiece, is fed into the work. As the spindle rotates, the cutting tool engages the interior of the existing hole, and chips are formed as the tool cuts into the workpiece. The actual cutting action of a boring tool is very similar to a lathe turning tool. Both the dimensional quality and the surface finish of the hole are determined by the rigidity and accuracy of the machine spindle and tool holder. The more rigid the machine is, and the more precisely the boring tool and spindle are controlled, the more precise the hole will be, both as to dimension and surface finish.

Broaching. This is the removal of material to produce a slot (or other formed shape) in a workpiece by moving a multiple-tooth, barlike tool across the workpiece. The cutting action results from the configuration of each tooth being progressively higher than the preceding one. Each tooth of the broach removes a small, predetermined amount of stock, the chip (**Fig. 1**). The peripheries or cutting edge of the broach are shaped to give a desired surface contour. Often a slot, it may also be almost any geometric shape desired. Broaching is usually performed on special machines that pull or push the broach across the workpiece. Internal broaching utilizes an existing hole to provide clearance for the broaching tool. External broaches do not require existing slots,

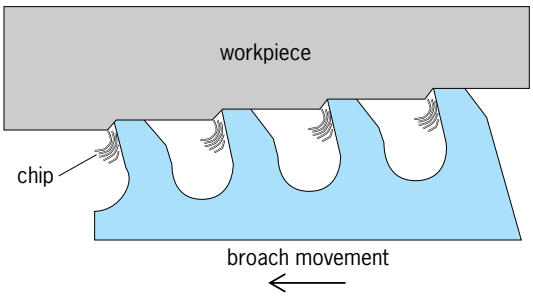

Fig. 1. Broach cutting a small chip from a workpiece.

but they may have guide bushings to support the broaching tool and a fixture to hold the workpiece. Broaching is a very economical machining operation, although tool costs can be high; accordingly, it is applied most often to high-volume production.

Drilling. One of the most common machining operations, drilling is a method of producing a cylindrical hole in a workpiece. A typical twist drill consists of a helically grooved steel rod with two cutting edges on the end. The helical flutes or grooves in the drill allow the chips to be removed from the cutting edge, conduct coolant to the cutting lips, and form part of the cutting edge geometry. Although drills may appear to be simple, their geometries are carefully controlled. Chisel angle, point angle, and web thickness, for example, are optimized to speed drilling and maximize tool life (**Fig. 2**). If this geometry is significantly altered during resharpening, drill life and hole quality may be compromised.

Drilling can be performed on special machines, which range in size and complexity from small sensitive drill presses through upright and multiple-spindle models designed for mass production. Jigs are often used to guide and support the drill. The use of a bushing to guide the drill into the workpiece can significantly improve drilling accuracy and tool life by improving the rigidity of the drill and assisting in properly locating it in relation to the workpiece. Of course, drilling can also be performed with simple hand drills. Drilling is a very fast and economical process, but it usually does not produce a very accurate hole diameter or a fine surface finish.

Turning. This type of machining is performed on a lathe. The process involves the removal of material from a workpiece by rotating the workpiece under power against a cutting tool. The cutting tool is held in a tool post that is supported on a cross slide and carriage. The tool may be moved radially or longitudinally in relation to the turning axis of the workpiece (**Fig. 3**). Forms such as cones, spheres, and related workpieces of concentric shape as well as true cylinders can be turned on a lathe. Machining operations such as facing and boring are variations of the turning process, and they can also be performed on a lathe.

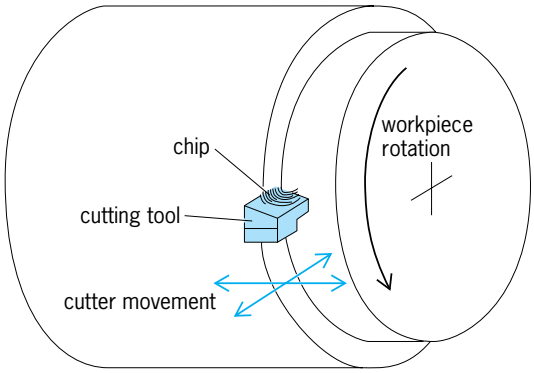

Fig. 3. Formation of turning chips as a single-point tool contacts the exterior of the spinning workpiece.

The most common lathe is an engine lathe, where the workpiece may be rotated and held between tapered centers or by means of a collet or chuck. A turret lathe is an engine lathe that has a multisided, indexing tool holder, or turret, instead of a tailstock center. This adds to the versatility of the machine by allowing a greater variety of cutting tools to be applied to the rotating workpiece. In addition to the turning tools used on an engine lathe, the turret lathe has tool holders that can be set up to drill, tap, and bore the workpiece.

Automatic screw machines. Automatic screw machines are sophisticated lathes that have been designed to perform several turning operations automatically in rapid succession without removing the workpiece from the machine. They are used when the volume and complexity of the required workpieces justifies the expense of setting up and operating these very versatile but complex machines. A screw machine (named for the screw manufacturing role for which is was created) is cam controlled or driven by computer numerical control (CNC). Several individually controlled slides are provided to hold turning, grooving, or cutoff tools. The machines also have turrets that hold drills, reamers, hollow mills, and counterboring tools. Special attachments permit performance of operations such as milling, index drilling, or thread chasing. *See* COMPUTER NUMERICAL CONTROL.

Milling. This process removes material by feeding a workpiece through the periphery of a rotating circular cutter. Each tooth of the rotating multitoothed milling cutter removes a portion of material from the passing workpiece.

Milling cutters are designed for particular operations and are classified as either shell type or end type. Shell mills are disk shaped and usually produce continuous slots. When well supported in the

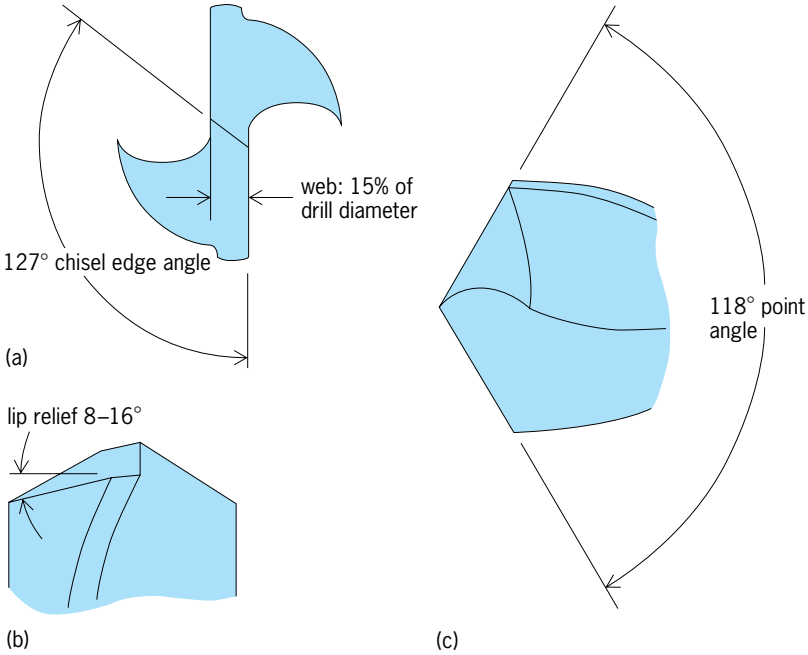

Fig. 2. Critical items on a drill point and their geometry. (*a*) End view. (*b*) Side view. (*c*) Side view, with drill rotated 90° about its axis from the position shown in *b*.

machine tool to minimize vibration, they perform well and economically. The more versatile end mill is held by its shank only. End mills can be used for slotting just like a shell mill, but they also can cut pockets, contours, and even cylindrical holes. End mills with four or six flutes are stronger and more rigid than two-flute mills, and so are better able to machine tougher materials. They have less flute room to allow chips to be removed, however, so they do not perform well on soft material at high speeds. Two-flute end mills are typically used for high production rates on soft materials such as aluminum. Ball-nose end mills, which have round ends, can be used to machine complex three-dimensional contours.

Reaming. This machining operation enlarges an existing hole by a few thousandths of an inch and produces a hole whose diameter is very accurately controlled. The cutting edges of a reamer may be ground on the apexes between longitudinal flutes or grooves, or cutting may take place on chamfered edges at the end of the reamer. Reaming is performed either manually or by machine, often as a finishing operation after drilling. Reamed holes have good surface finish. Hand reamers have a slight end taper to aid in starting. Fluted machine reamers for finishing work remove material by both end and side cutting, while the heavier tools known as rose reamers cut only on their chamfered ends.

Sawing. This is the parting of material by using blades, bands, or abrasive disks as the cutting tools. In the most common type of saw, a toothed blade is passed across the workpiece in either a reciprocating or continuous motion. The teeth can be mounted on continuous bands, short steel blades, or the periphery of a disk. Friction sawing is a rapid process used to cut steel as well as certain plastics. A very high speed blade softens the workpiece material with frictional heat. The material is then wiped away from the workpiece by the cutting blade. Since many teeth engage the workpiece, no single tooth overheats. Abrasive sawing looks similar to friction sawing, except that a thin rubber or bakelite bonded abrasive disk grinds the material away instead of simply softening and wiping it.

Shaping. This process cuts flat or contoured surfaces by reciprocating a single-point tool across the workpiece. The tool is mounted on a hinged unit known as a clapper box, which lifts up to disengage the tool from the surface on the return stroke. The cutting action of a shaper is actually similar to turning, except that the single-point cutting tool moves straight across a workpiece, instead of having the workpiece rotate against the cutting tool. Because shapers have a limited stroke, shaping is usually performed on small pieces requiring a short cutter stroke. Larger pieces requiring longer strokes are usually milled or planed.

By and large, shaping is obsolete except in small shops; shapers are no longer being produced. Planing, when employed to substitute for shaping, is done by reversing the relative motion between tool and workpiece; that is, the tool is stationary during the cut, and is fed incrementally between cuts, while the workpiece is reciprocated past the stationary tool.

Nibbling. Nibbling is the operation that cuts away small pieces of material by the action of a reciprocating punch. A nibbler takes repeated small bites from the workpiece (which is usually a thin sheet of material) utilizing a quickly reciprocating punch system. As the work is passed beneath the punch, a small nibble of material is removed by the punch during each punch cycle. After each quick nibble, the workpiece is advanced under the punch to allow another small bite to be taken away. In this manner, a great deal of material can be removed from the workpiece after several tens or hundreds of individual punching cycles. By controlling the movement of the workpiece under the reciprocating punch, slots, profiles, and radii may be cut into the workpiece. Additionally, since there is a great deal of versatility in workpiece movement, complex or irregular shapes may easily be produced. Duplicate pieces may be made simultaneously by stacking the sheet workpieces and allowing the punch to cut several sheets at a time. Templates are also used as guides to direct the material movement beneath the punch to produce large predetermined shapes.

Nibbling machines are constructed with considerable distance, or throat, between the punch and its drive and support mechanism. This distance allows relatively large workpieces to be machined. The punches themselves are frequently round or square, but also may have complex shapes if special detailed shapes are desired on the workpiece.

Grinding. This process removes material by the cutting action of a solid, rotating, grinding wheel. The abrasive grains of the wheel perform a multitude of minute machining cuts on the workpiece. Although grinding is sometimes used as the sole machining operation on a surface, it is generally considered a finishing process used to obtain a fine surface and extremely accurate dimensions.

Grinding is used to machine a wide range of metals, carbide materials, stone, and ceramics. The grinding process may be used on metals too hard to machine otherwise because commercial grinding abrasives are many times harder than the metals to be machined.

Grinding wheels are composed of abrasive grains plus a bonding material (**Fig. 4**). The wheels are often very porous, with homogeneous open areas between the grains. The abrasives most commonly used are silicon carbide (SiC) and aluminum oxide (Al_2O_3). Coarse-grained wheels are used for rapid removal of stock; wheels with fine grains cut more slowly but give smoother finishes. Coolants are applied to the grinding point to dissipate the heat generated and to flush away the fine chips. *See* ABRASIVE.

Cylindrical grinding is performed on the peripheries or shoulders of workpieces such as shafts, cylinders, rolls, and axles. They rotate the workpiece from a power headstock while it is held between centers, gripped in a jawed chuck, or fastened to a faceplate. Centerless grinding is similar to cylindrical grinding, except that the work is supported on

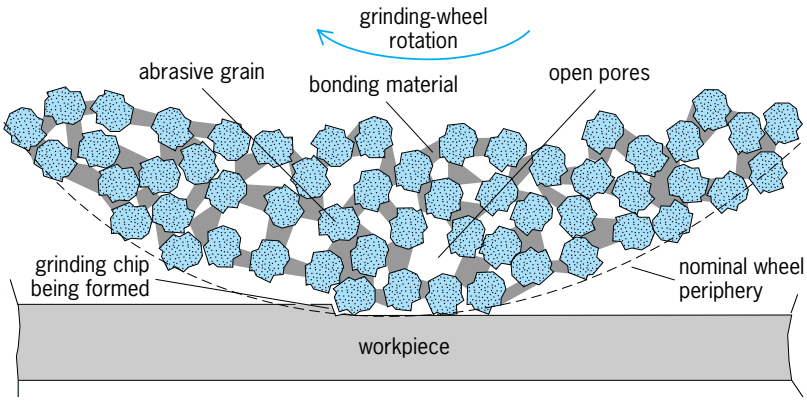

Fig. 4. Action of a grinding wheel, showing how individual grains of the wheel produce chips from a workpiece.

a blade between two abrasive wheels. One wheel is a regular grinding wheel rotating at normal grinding speed, while the second is a rubber-bonded regulating wheel that causes the workpiece to rotate against the grinding wheel (**Fig. 5**). Centerless grinders are faster than cylindrical grinders, but they do not provide the accurate concentricity that cylindrical grinders provide. Surface grinding is accomplished by holding one or several workpieces on a reciprocating or rotating horizontal table and feeding them through the cutting path of a rotating grinding wheel.

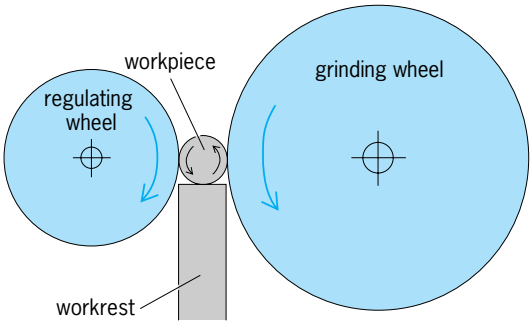

Fig. 5. Configuration of a centerless grinding wheel.

Creep-feed grinding is an abrasive machine process capable of removal of heavy stock in working difficult-to-machine materials. The process has the ability to hold both size and form to very close tolerances (**Fig. 6**). Specialized equipment and grinding wheels are required to properly perform creep-feed grinding. Creep-feed grinding is analogous to shell milling except that a grinding wheel is used in place of the milling cutter. The multitude of abrasive grains acts as the cutting teeth of a shell mill. The cutting forces are very high in the creep-feed grinding process; therefore the machine tool must be designed for rigidity and must be capable of closely controlled slow table speeds. Creep-feed grinding also requires special wheels. These grinding wheels are very soft, and have pores in the wheel that allow the cutting fluid to be taken up and into the arc of the cut. Superabrasive wheels are sometimes used, because their high initial cost is justified by their long life and high grinding productivity.

Cubic boron nitride and polycrystalline diamond are classified as superabrasives. Cubic boron nitride was discovered as the product of a failed attempt to make diamonds in the laboratory. It is almost as hard as diamond. Since cubic boron nitride is a synthetic material, the structure and size of the grains can be closely controlled and optimized for specific applications. Polycrystalline diamond is a synthetic diamond that cleaves as it fractures under load, thus exposing a new, sharp cutting edge instead of simply dulling as other abrasives do. Therefore polycrystalline diamond provides longer cutting life compared to other abrasives.

Honing. Honing is a grinding process that removes a small amount of material from a workpiece by means of abrasive stones. It is able to produce extremely close dimensional tolerances and very fine surface finishes. The abrading action of the fine grit stones occurs on a wide surface area rather than on a line of contact as in grinding. A hone floats as it moves across the workpiece. The hone removes high spots and smooth surface inaccuracies. It is well suited to smoothing and improving the roundness of a hole, but it will not correct the position of a hole or establish its alignment.

Lapping. This is precision abrading process used to finish a surface to a desired state of refinement or dimensional accuracy by removing an extremely small amount of material. Lapping is accomplished by abrading a surface with a fine abrasive grit rubbed about it in a random manner. A loose unbonded grit is used. It is traversed about on a lap, made of a somewhat softer material than the workpiece. The unbonded grit is mixed with a vehicle such as oil, grease, or soap and water. When a bonded grit is used, it may be in the form of a bonded abrasive lap or a charged lap such as cast iron or copper with the lapping compound embedded in it. In some cases, abrasive-covered paper laps are used.

Polishing. This is the smoothing of a surface by the cutting action of an abrasive grit that is either glued to or impregnated in a flexible wheel or belt. Polishing is not a precision process; it removes stock until the desired surface condition is obtained. Polishing

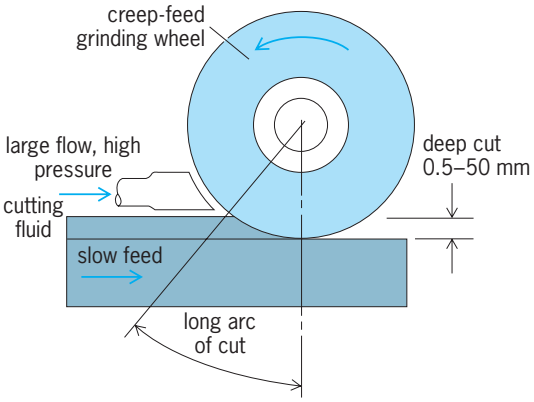

Fig. 6. Creep-feed grinding process.

wheels are built up from layers of soft material such as wood or leather.

Superfinishing. For mirrorlike surface, a precision abrading process known as superfinishing is used. Superfinishing removes minute flaws or inequalities because it is performed with an extremely fine-grit abrasive stone, shaped to match and cover a large portion of work surface. As the workpiece turns, the stone reciprocates across the work under a flood of lubrication. The process is usually performed on cylindrical pieces.

Buffing. This is the smoothing and brightening of a surface by rubbing it with a fine abrasive compound carried in a soft wheel or belt. The abrasive is a fine powder or flour mixed with tallow or wax to form a smooth paste. Buffing is performed on the same types of machines as polishing; frequently both buffing and polishing wheels are included on the same machine. Buffing differs from polishing in that a finer grit is used and less material is removed from the workpiece. Although some material is removed, the amount is almost undetectable.

Ultrasonic machining. In this process, material is removed by abrasive bombardment and crushing in which a vibrating tool drives an abrasive grit against a workpiece. In ultrasonic machining, the tool never directly contacts the workpiece; rather, the vibrations (typically 20,000 Hz) drive the abrasive, which is suspended in a liquid, against the workpiece. Impressions may be economically sunk in glass, ceramics, carbides, and hard brittle metals by this method. The shape of the tool basically determines the shape of the impression. Often, several cuts will be required, with progressively finer abrasive grit sizes, to obtain the desired dimensional accuracy and surface finish. *See* ULTRASONICS.

Electric-discharge machining. This is a machining process in which electrically conductive materials can be removed by repeated electric sparks. Electric-discharge machining is used to form holes of varied shape in materials of poor machinability. Unlike other machining operations, it does not rely on a cutting tool to shear away the workpiece. Instead, it uses electrical energy to melt or vaporize small areas on the workpiece. The sparks created by an electric-discharge machining unit (at a rate greater than 20,000 per second) are discharged through the space between the tool (cathode) and the workpiece (anode; **Fig. 7**). The small gap between the tool and workpiece is filled with a circulating dielectric hydrocarbon oil, which serves as a cooling medium and flushes away metal particles. Although not as fast as other machining processes, electric-discharge machining has the unique ability to remove hard materials that otherwise would not be machinable. Electric-discharge machining was originally developed as a method of removing broken taps or drills from holes.

Gear cutting. This operation results in a uniform series of toothed projections on the surface of a workpiece, the teeth being designed to mesh with a mating tooth series in order to transmit power or motion. Gear-cutting machines may use either conventional cutting tools or grinding wheels to produce the desired gear shape.

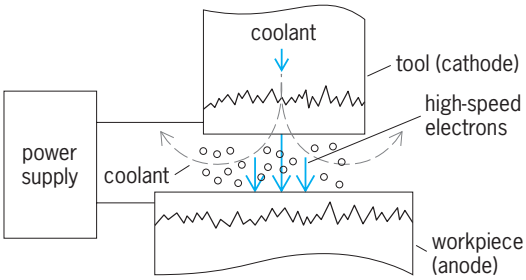

Fig. 7. Electric-discharge, or electric-spark, machining process. (*Elox Corporation of Michigan*)

Gear forms may be generated by the total cutting action of the tool plus the rotation of the workpiece, or they may be formed by a tool whose cutting profile matches the tooth form. Frequently, gear teeth may be rough-cut by a generating method and finished to size with a form tool.

Gear hobbing machines revolve the gear blank while the rotating hob, or cutter, moves across the face of the gear to generate the teeth (**Fig. 8**). A system of index gearing maintains the relative positions of the blank and the hob and controls the number

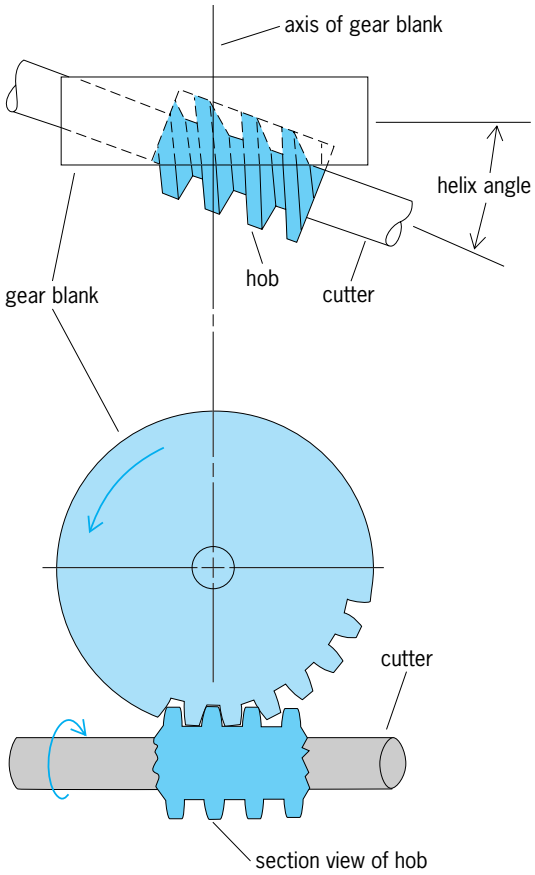

Fig. 8. Gear hobbing operation; rotation of the hob and gear blank is synchronized and continuous. (*After J. G. Bralla, ed., Handbook of Product Design for Manufacturing, McGraw-Hill, 1986*)

of teeth to be cut. In addition to use of specialized hobbing machines, gears can be formed by milling and broaching.

Gear-grinding machines produce gear teeth by using preformed grinding wheels. Although some fine-pitch gear teeth are finish-ground from solid stock, gear grinding is primarily a finishing operation. Grinding is the only way to finish gears that have been brought to a high degree of hardness. A grinding wheel, which has an involute tooth form dressed into both sides, grinds both sides of the space between the two gear teeth as it moves across the face of the gear. *See* GEAR.

Threading. Threading is the forming of a ridge and valley of uniform cross section that spiral about the inner or outer diameter of a cylinder or cone in an even, continuing manner. The work must be produced with sufficient uniformity and accuracy that the resulting threaded part will accomplish its intended purpose of fastening, transmitting motion or power, or measuring.

Threads are produced by several machining methods. External threads may be chased by preformed cutting dies, turned with single-point turning tools on a lathe, milled, hobbed, ground, or rolled. Most internal threading is cut with a preformed tap although turning, hobbing, grinding, and milling are also used. *See* SCREW THREADS.

Peening. This is a metal-finishing operation, also known as shot peening. During a peening operation, small steel shot is driven against a workpiece to plastically deform its surface to a depth of a few thousandths of an inch, producing residual compressive stress. The material is thus made more resistant to fatigue failure. Surface hardness of the material is also increased slightly by the cold working produced by shot peening. The shot is driven at high velocity against the work by centrifugal force or by an air blast.

Cutting-tool coatings. To increase the life and speed of cutting tools, they are often coated with a thin layer of extremely hard material such as titanium nitride or zirconium nitride. These materials, which are applied over the cutting edges, provide excellent wear resistance. They are also relatively brittle, so they rely on the toughness of the underlying cutting tool to support them. Coated tools are more expensive than conventional tools, but they can often cut at much higher rates and last significantly longer. When used properly on sufficiently rigid machine tools, they are far more economical than conventional tooling. *See* METAL COATINGS.

Jig and fixture design. Successful machining usually requires that the workpiece be securely presented to the cutting tool. In machines such as lathes or cylindrical grinders, the holding device is typically a chuck, collet, or center driver. Often, however, special devices must be constructed to position and support the workpiece, especially for drilling, milling, welding, and assembly operations. Jigs and fixtures provide this function. Jigs support the workpiece and provide locating guides for the cutting tool. Fixtures support the workpiece but do not provide locating guides. The workpiece is held against one or more reference surfaces, from which all machining cuts are measured. Some type of clamping device, such as a toggle clamp or locking screw, is also frequently used. The most common locating guide is a drill or reamer bushing, which is a hardened steel tube whose inside diameter is just slightly larger than the cutting-tool diameter. The bushing guides the drill or reamer to the workpiece and helps to prevent deflection of the cutting tool under load, thereby improving cutting accuracy and extending cutting tool life.

J. R. Casey Bralla

Bibliography. W. E. Boyes (ed.), *Jigs and Fixtures*, 3d ed., Society of Manufacturing Engineers, 1989; J. G. Bralla (ed.), *Product Design for Manufacturing*, 1986; R. J. Christoforo, *Jigs, Fixtures, and Shop Accessories*, 1993; S. F. Krar, *Grinding Technology*, 2d ed., 1994; S. F. Krar, *Machine Tool Operations*, 1983; R. L. Little, *Metalworking Technology*, 1976; Society of Manufacturing Engineers, *SME Tool and Manufacturing Engineers Handbook*, vols. 1–5, 1983–1987; R. A. Walsh, *McGraw-Hill Machining and Metalworking Handbook*, 2d ed., 1998.

Mackerel

A fish which is a member of the order Perciformes, family Scombridae. There are about 50 carnivorous species found in the middle layer or near the surface of tropical and temperate seas.

Mackerel are characterized by a long slender body, pointed head, and large mouth. The dorsal fin is spinous, the anal fin is soft-rayed, and the tail fins are widely forked and powerful. The teeth are feeble, and the maxillary bones are fused to the premaxillaries.

Scomber scombrus, the common mackerel (see **illus.**), is an important fish commercially. It does not have a swim bladder and can move from the depths to the surface rapidly. It is a migratory species that is found on both sides of the North Atlantic. It swims in large schools, which approach the coast in the early summer to spawn, and then return to the ocean. This mackerel is about 2 ft (0.6 m) long when mature.

Common mackerel (*Scomber scombrus*).

The Pacific mackerel (*Pneumatophorous diego*) is also an important commercial fish but differs from the common mackerel in having a swim bladder. The American Spanish mackerel (*Scomberomorus maculatus*) is a choice food fish that may attain a weight of 10 lb (4.5 kg). *See* PERCIFORMES; SWIM BLADDER.

Charles B. Curtin

Macrocyclic compound

An organic compound that contains a large ring. In the organic chemistry of alicyclic compounds, a closed chain of 12 carbon (C) atoms is usually regarded as the minimum size for a large ring; crown ethers are similarly defined. Macrocyclic compounds may be a single, continuous thread of atoms, as in cyclododecane [$(CH_2)_{12}$], or they may incorporate more than one strand or other ring systems (subcyclic units) within the macrocycle or macroring. In addition, macrocycles may be composed of aromatic rings that confer considerable rigidity upon the cyclic system. These aromatic rings may be joined together or coupled by spacer units consisting of one or more carbon atoms. The compounds generally thought of as macrocycles include crown ethers, cryptands, spherands, carcerands, cyclodextrins, cyclophanes, and calixarenes. *See* AROMATIC HYDROCARBON.

Classes of macrocyclic polyethers. Crown ethers are generally composed of repeating ethylene (CH_2CH_2) units separated by noncarbon atoms such as oxygen (O), nitrogen (N), sulfur (S), phosphorus (P), or silicon (Si). Cryptands are similar to crown ethers but have a third organic chain attached at two points on the crown ring, usually trivalent nitrogen. Although other carbon-containing subunits such as propylene ($CH_2CH_2CH_2$) or methylene (CH_2) may be included, these subunits are less common. The conformations of the ethylene unit are more favorable than those of propylene, and the ethylene unit is more stable to hydrolysis than is the O—CH_2—O linkage (an acetal). Organic units sterically equivalent to the ethylene unit, such as 1,2-benzo, have been incorporated, as have units sterically equivalent to two ethyleneoxy units, such as 2,6-bis(hydroxymethyl)pyridine. Replacement of an OCH_2CH_2O unit by the 1,2-dihydroxyethyl unit of tartaric acid (HOOC—CHOH—CHOH—COOH) or by the adjacent hydroxyl groups of carbohydrates (sugars) alters hydrophobicity and introduces chirality. By far, the most common heteroatom present in the macrorings of crowns [X in $(XCH_2CH_2)_n$] is oxygen; but as more intricate structures are prepared, nitrogen, sulfur, phosphorus, silicon, or siloxy residues are becoming much more common. *See* STEREOCHEMISTRY.

The name crown ether is informal, developed to bypass the complex systematic names that these compounds are given under traditional nomenclature rules. A simple formal system of nomenclature for compounds of this type has yet to be developed, although several systems have been suggested. Of these, the names coronand for a cycle and coronates for its complex have been most widely adopted. By tradition, however, crown ethers are named according to the total number of atoms composing the macrocyclic ring and the number of heteroatoms contained within it (**Fig. 1***a*). A crown composed of five ethyleneoxy [CH_2CH_2O] units would be called 15-crown-5. Substitution of one ethyleneoxy unit by a 1,2-benzo unit would lead to benzo-15-crown-5.

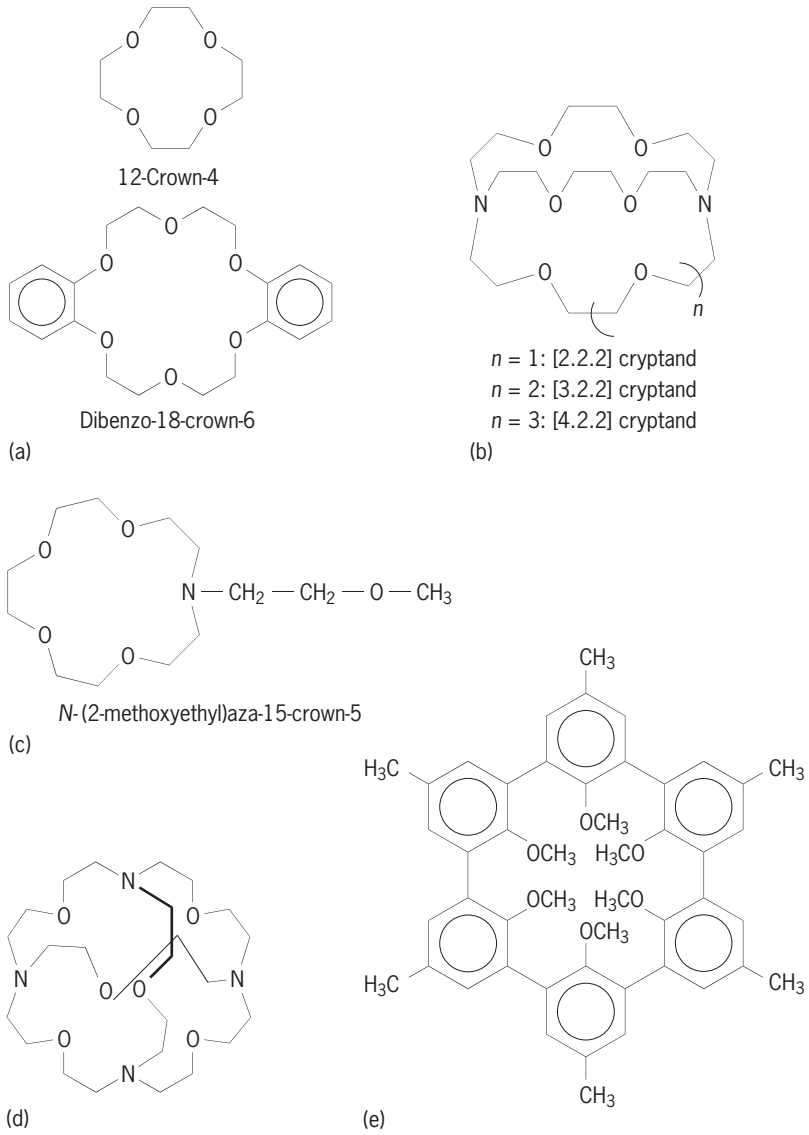

Fig. 1. Macrocyclic structures. (*a*) Two crown ethers. (*b*) A basic structure for three cryptands. (*c*) A lariat ether. (*d*) All-aliphatic spherand. (*e*) All-aromatic spherand.

Substitution of a nitrogen atom for an oxygen would transform 15-crown-5 into aza-15-crown-5.

Crown ether variants. Numerous structural features have been incorporated into the crown ether ring system. The simplest variation is substitution of nitrogen, sulfur, phosphorus, silicon, or other heteroatom for oxygen. Such substitution leads, respectively, to azacrowns, thiacrowns, phosphacrowns, and silacrowns. Multiple substitution of heteroatoms is also possible. A 14-membered ring containing 4 oxygen heteroatoms is called 14-crown-4. The all-nitrogen analog may be called tetraaza-14-crown-4, or 14-ane-4 and has the trivial name cyclam. Such compounds are of special interest to inorganic chemists, as these analogs of crown ethers bind transition and coinage metals. When the —CH_2—O—(ether) link is replaced by a —CO—O—(ester) unit, the crown ether is transformed into a crown ester. Likewise the interchange (ether for amide) of reaction (1) affords macrocyclic

$$—CH_2—O— \longrightarrow —CO—NH— \qquad (1)$$

lactams. Crown ethers have also been prepared that incorporate subcyclic units such as benzene, furan, and pyridine. *See* ESTER; ETHER; LACTAM; LACTONE.

Cryptands. By adding a third strand to the simple macrocyclic polyethers, three-dimensional compounds based on the crown framework are formed. Typically, two of the oxygen atoms across the ring from each other are replaced by nitrogens, and a third ethyleneoxy chain is attached to them. Known as cryptands (Fig. 1*b*), these structures completely encapsulate cations smaller than their internal cavities and strongly bind the most similar in size. Systematic nomenclature of cryptands is complicated. A cryptand is usually presumed to be composed of ethyleneoxy chains attached to a nitrogen atom at either end. A typical cryptand would be $N[(CH_2CH_2O_2)\ CH_2CH_2]_3N$. Since each chain contains two oxygen heteroatoms, it is referred to as [2.2.2]-cryptand, or simply [2.2.2]. A cryptand composed of chains having two, three, and four ethyleneoxy units tied to nitrogen atoms at either end has one, two, and three oxygens respectively in the chains and is called [3.2.1] cryptand.

Lariat ethers. The name lariat ether (Fig. 1*c*) has been given to compounds that, like the cryptands, have a ring and a third strand, but the additional chain is attached at only one point. This permits three-dimensional encapsulation of a cation while retaining a high degree of flexibility. Two-armed lariat ethers are referred to as bibracchial lariat ethers or BiBLEs. Three-armed systems are TriBLEs, and so forth.

Spherands. Compounds other than simple cryptands have been developed that are capable of either partially (cavitands) or completely (spherands) enveloping cations. The donor atoms (O, N, S) are arranged so that they provide a solvation sphere to the encapsulated cation. The earliest example of a spherand is Lehn's all-aliphatic spherand (Fig. 1*d*). An all-aromatic spherand (Fig. 1*e*) is a compound that can envelop a guest three-dimensionally to an even greater extent than can a cryptand. When a crown ether ring is attached to three of the six aromatic rings shown for the all-aromatic spherand, a cryptaspherand results. If only an ethyleneoxy chain, rather than a crown ether, is appended to the aromatic unit, a hemispherand results.

Calixarenes, cavitands, and carcerands. Calixarenes are cyclic oligomers of phenol and formaldehyde. Calix[*n*]arenes from *p-tert*-butylphenol have been developed in which *n* is 4, 6, or 8, corresponding to that number of aromatic rings and CH_2 units (Fig. 2*b*). Less common but also available are the five- and seven-ring calix[*n*]arenes. All these compounds have aromatic walls and a hydrophobic interior. The larger molecules are more flexible than the smaller ones, but both are capable of including a variety of organic guest molecules.

An extension of calixarenes to a more rigid, basket-shaped system can be accomplished by the reaction of resorcinol and acetaldehyde. Subsequent reaction with formaldehyde gives the enforced-cavity structure called a cavitand (Fig. 2*c*), a structure that possesses a molecular void (cavity) that is invariably filled by solvent when the molecule crystallizes.

An even greater level of encapsulation (incarceration) of a molecular guest is possible when two cavitands are combined as top and bottom shells of a carcerand (Fig. 2*d*). These compounds can effectively isolate the bound guest from the chemical environment (solvent) in which the host·guest (solvent·carcerand) complex exists. In keeping with nomenclature traditions, the complexes of these two structures are referred to as cavitates and carcerates. *See* SUPRAMOLECULAR CHEMISTRY.

Cyclophane is the name given to macrocyclic compounds that contain organic (usually aromatic) rings as part of a cavity-containing structure (Fig. 2*e*). The first such compound was [2.2]-paracyclophane. In it, two benzene rings are joined by ethylene (CH_2CH_2) chains in their para positions. *See* CYCLOPHANE.

Bicyclic cryptands. Two crown ether rings may be held together by a crown-ether-like strand to give a bicyclic cryptand (**Fig. 2***a*). These have sometimes been referred to as ditopic receptors because they possess two distinct binding sites. The cryptands and spherands differ from the ditopic or bicyclic cryptands in that they normally possess a three-dimensional array of binding sites focused for a single guest.

Host-guest chemistry. The terms host and guest designate a pair of molecular species, comprising one molecule that can bind another. In principle, either molecule of the pair may be the host or the guest. The distinction between host (receptor) and guest (substrate) is arbitrarily made based upon size. The larger (higher molecular weight) molecule is the host (receptor) and is the compound that is deliberately prepared by chemical synthesis or that has evolved to complexity in nature for example, (proteins, DNA). In chemistry the host or receptor molecule (especially for a cation) is called a ligand, but in biology it is the substrate or guest that is often called a ligand.

Calixarene is representative of a growing family of diverse molecular receptors in which four, six, and eight aromatic rings are common in otherwise analogous structures (**Fig. 3**). As with crown ethers, assorted molecular species may be complexed by these diverse receptors. In many cases, the focus will be on a single receptor molecule and many potential substrates in order to define the efficacy and selectivity of the host. Natural receptors such as proteins have been studied in order to understand which guest molecules are bound and transformed (enzymatic catalysis) by them. Other families of receptor molecules have been devised which are based upon the sugar oligomers called cyclodextrins containing five to seven glucose molecules in a cyclic array. These natural and semisynthetic hosts bind a variety of organic substrates (guests) in water solution.

Complexation phenomena. It is the ability of these macrocyclic host compounds to complex a variety of guest species that makes these structures interesting.

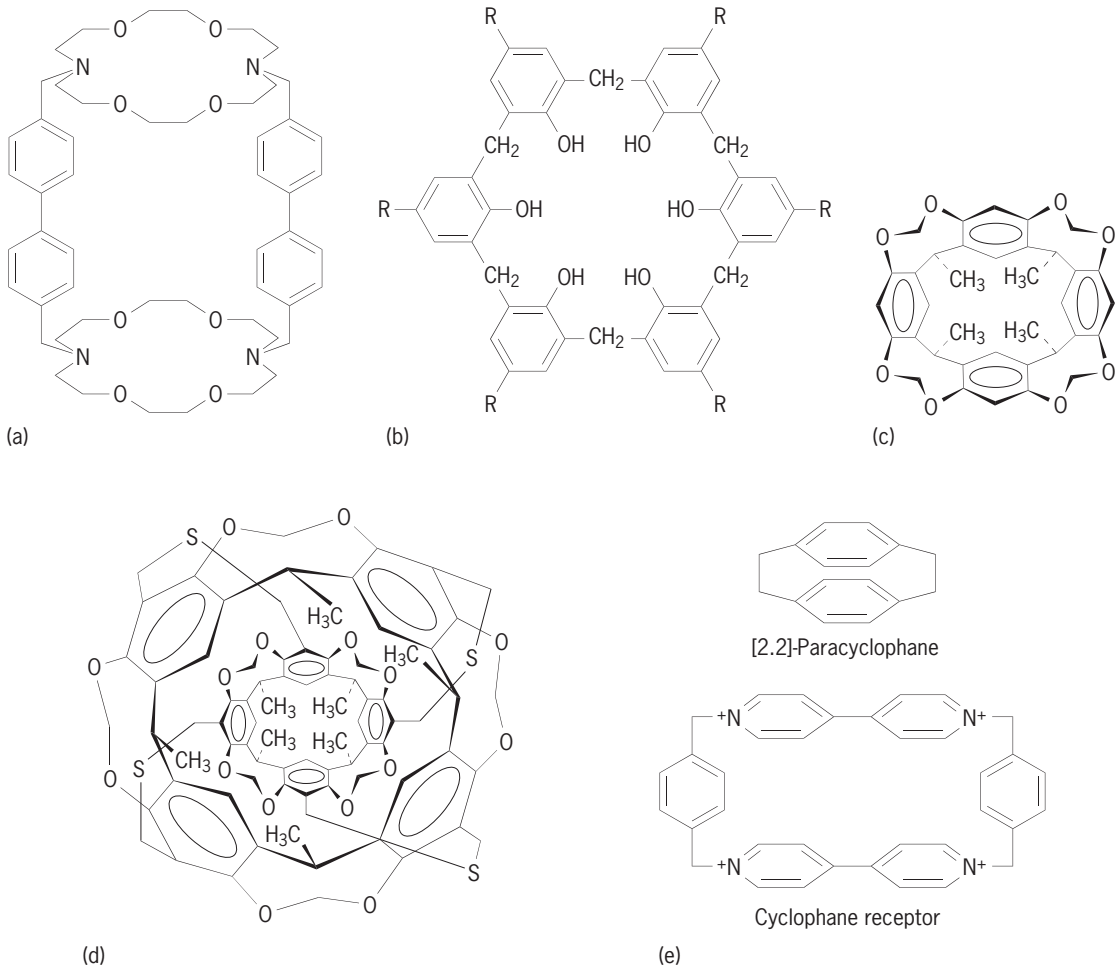

Fig. 2. Macrocycles with various structural features. (*a*) Bicyclic cryptand (ditopic receptor). (*b*) Calix[6]arene; R represents a functional group. (*c*) Cavitand. (*d*) Carcerand. (*e*) Cyclophanes.

A crown ether can be described as a doughnut with an electron-rich and polar hole and a greasy or lipophilic (hydrophobic) exterior. As a result, these compounds are usually quite soluble in organic solvents but accommodate positively charged species in their holes. The complexation process is usually characterized by a binding or stability constant (K_s) of the form $K_s = k_c/k_d$, where k_c represents the rate constant for complex formation (k_{complex}) and k_d represents the rate of decomplexation ($k_{\text{decomplex}}$ or k_{release}).

The two important components of complexation are the magnitude of the equilibrium constant and the kinetics of the process. For simple crown ethers, k_c is usually very large and k_d is also large. These rate constants are solvent-dependent, but k_d decreases more in nonpolar solvents than does k_c. Thus, K_s is smaller in polar solvents and larger in nonpolar solvents. Both k_c and k_d are lower for cryptands and spherands than for crowns, but k_d is lowered more. As a result, binding constants (K_s) for the three-dimensional hosts are generally higher (they bind cations more strongly) than for simple two-dimensional crowns.

Binding constants. Binding or stability constants can be assessed by several methods. One approach is to dissolve in water a cation such as Li^+ or Na^+ paired with a colored organic anion such as picrate. Shaking with an equal volume of organic solvent does not extract the salt. When a crown ether is added, the cation is complexed and transported into the organic phase. The crown cation–complex is accompanied by the colored anion, so the extent of extraction can be determined by colorimetric analysis. The extraction constants are expressed numerically as the percent of total salt extracted.

Homogeneous binding constants are determined for a ligand and a cation present simultaneously in solution. These K_s values may be determined by a variety of techniques (nuclear magnetic resonance, conductometric methods, calorimetry), but are most often obtained from either calorimetric or ion-selective electrode studies.

The interior cavity (hole) of a crown ether or cryptand varies in size depending upon the compound's structure. The series of compounds 12-crown-4, 15-crown-5, 18-crown-6, and so forth, increase in size. Thus, the hole of 15-crown-5 is about 0.20 nanometer, almost the same size as the ionic diameter of the sodium ion (Na^+). Likewise, the hole of 18-crown-6 and the potassium ion (K^+) are similar at about 0.27 nm. This similarity has given rise to

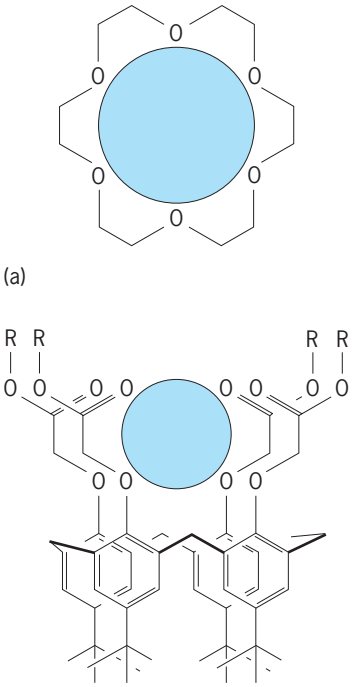

Fig. 3. The metal cation is the great and the organic structure is the host in (*a*) 18-crown-6 complex of potassium cation (K^+) and (*b*) sodium (Na^+) complexed by a calix[4]arene.

the hole/size relationship, which holds that binding is at a maximum when the cation and crown ether's hole are similarly sized. This principle is valid for the cryptands, which select cations on the basis of size. The correlation is tenuous for crown ethers in solution.

Complexing species. A variety of organic cations have been found to complex with crown ethers and related hosts. It has been suggested that for a host-guest interaction to occur, the host must have convergent binding sites and the guest must have divergent sites. This is illustrated by the interaction between optically active dibinaphtho-22-crown-6 (**Fig. 4*a***) and optically active phenethylammonium chloride. The crown ether oxygen atoms converge to the center of a hole and the ammonium hydrogens diverge from nitrogen. Three complementary O—H—N hydrogen bonds stabilize the complex (Fig. 4*b*). In this particular case, different steric interactions between the optically active crown and the enantiomers of the complex permit resolution of the salt.

Other organic cations have also been complexed, either by insertion of the charged function in the crown's polar hole or by less distinct interactions observed in the solid state. *See* COORDINATION CHEMISTRY; COORDINATION COMPLEXES.

Applications. Macrocyclic compounds have numerous interesting properties that have led to many applications.

Cation complexation. The striking ability of neutral macrocyclic polyethers to complex with alkali and alkaline-earth cations as well as a variety of other species has proved of considerable interest to the chemistry community. Crown ethers may complex the cation associated with an organic salt and cause separation of the ions. In the absence of cations to neutralize them, many anions show considerably enhanced reactivity. *See* ORGANIC REACTION MECHANISM.

By far, the most common effect of crown or cryptand compounds is to enhance the reactivity of counteranions by complexation of cations. There are two notable exceptions. First, when complexed by crowns, the normally unstable or explosive arenediazonium cations exhibit remarkably enhanced stability. More dramatic is the complete isolation of sodium cation, which allowed the preparation of the sodium anion (Na^-) for the first time. The crystals of cryptated Na^+ in the presence of Na^- are deep gold in color. This ability of cryptands to surround and protect the cation also allowed the preparation of the first electride, a cation salt of a lone electron. *See* ALKALI METALS; ALKALINE-EARTH METALS.

Phase-transfer catalysis. One of the important modern developments in synthetic chemistry was the use of the phase-transfer technique. Nucleophiles such as cyanide are often insoluble in media that dissolve organic compounds with which they react. Thus 1-bromooctane may be heated in the presence of sodium cyanide for days with no product formation. When a crown ether is added, two things change. First, solubility is enhanced because the crown wraps about the cation, making it more lipophilic. This, in turn, makes the entire salt more lipophilic. Second, by solvating the cation, the association between cation and anion and the interactions with solvent are weakened, thus activating the anion for reaction. This approach has been used to assist the

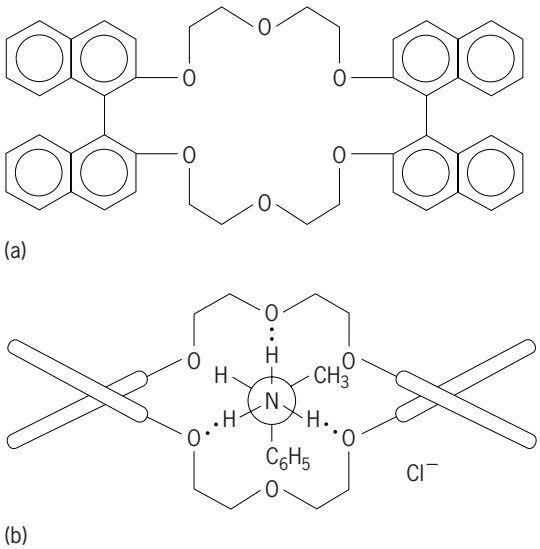

Fig. 4. Complexing species. (*a*) Dibinaphtho-22-crown-6. (*b*) R,R-dibinaphtho-22-crown-6 complex of phenethylammonium chloride.

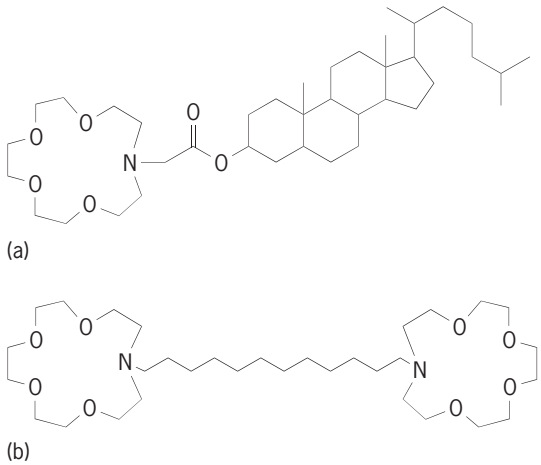

Fig. 5. Amphiphiles. (*a*) Crown ether-steroid amphiphile. (*b*) Crown ether bolaamphiphile.

dissolution of potassium permanganate ($KMnO_4$) in benzene (to give so-called purple benzene) in which solvent permanganate is a powerful oxidizing agent. One striking example of solubilization is the displacement of chloride (Cl^-) by fluoride (F^-) in dimethyl 2-chloroethylene-1,1-dicarboxylate by using the KF complex of dicyclohexano-18-crown-6, reaction (2).

$$ClHC{=}C(COOCH_3)_2 \longrightarrow FHC{=}C(COOCH_3)_2 \qquad (2)$$

In reaction (2), a crown provides solubility for an otherwise insoluble or marginally soluble salt. Use of crowns to transfer a salt from the solid phase into an organic phase is often referred to as solid-liquid phase-transfer catalysis. *See* CATALYSIS; PHASE-TRANSFER CATALYSIS.

Indicators and sensors. Since crown ethers and related species complex cations selectively, they can be used as sensors. Crowns have been incorporated into electrodes for this purpose, and crowns having various appended chromophores have been prepared. When a cation is bound within the macroring, a change in electron density is felt in the chromophore. The chromophores are often nitroaromatic residues and therefore highly colored. The color change that accompanies complexation can be easily detected and quantitated. *See* ION-SELECTIVE MEMBRANES AND ELECTRODES.

Membrane formation. Lipid bilayer membranes form from compounds that have polar head groups and nonpolar tails. Such compounds are known as amphiphiles. Crown ethers are capable of serving as the head group in a family of compounds that form membranes. The crown ether-steroid compound (**Fig. 5***a*) forms traditional bilayers, but the bis(crown) compound (Fig. 5*b*), which has two head groups (a bolaamphiphile), forms ultrathin monolayer membranes in which a single covalent chain spans the entire membrane. *See* MICELLE; MONOMOLECULAR FILM.

Toxicity. Some crowns and cryptands, like many other organic compounds, are irritants or toxic. 15-Crown-5 has produced a slight redness in skin tests and is reported to be toxic in mice at a level of about 0.01 oz/lb (800 mg/kg) of body weight. This means that a 150-lb (70-kg) human would have to ingest about 2 oz (56 g) of this compound to be in danger. The toxicity of simple aspirin is about half this (that is, twice as much aspirin, 0.0267 oz/lb or 1.75 g/kg, is required for the same toxic effect), so the toxicity level of this crown is not high. Of course, not all crowns have been tested, and some may be dangerous. *See* TOXICOLOGY. George W. Gokel

Bibliography. V. Balzani and F. Scandola, *Supramolecular Photochemistry*, 1991; D. J. Cram and J. M. Cram, *Container Molecules and Their Guests*, 1994; F. Diederich, *Cyclophanes*, 1991; B. Dietrich, P. Viout, and J.-M. Lehn, *Macrocyclic Chemistry*, 1992; G. W. Gokel, *Crown Ethers and Cryptands*, 1991; C. D. Gutsche, *Calixarenes*, 1989; Y. Inoue and G. W. Gokel (eds.), *Cation Binding by Macrocycles*, 1990; C. M. Starks, C. L. Liotta, and M. Halpern, *Phase Transfer Catalysis*, 1993; J. W. Steed and J. L. Atwood, *Supramolecular Chemistry*, 2000.

Macrodasyida

An order of the phylum Gastrotricha. Members all have pharyngeal lumina that are inverted-Y–shaped, and most have pharyngeal pores. They inhabit marine or brackish waters, seldom fresh waters. Some do not exceed 0.5 mm (0.02 in.) in length, and most are not more than 1–1.5 mm (0.04–0.06 in.);

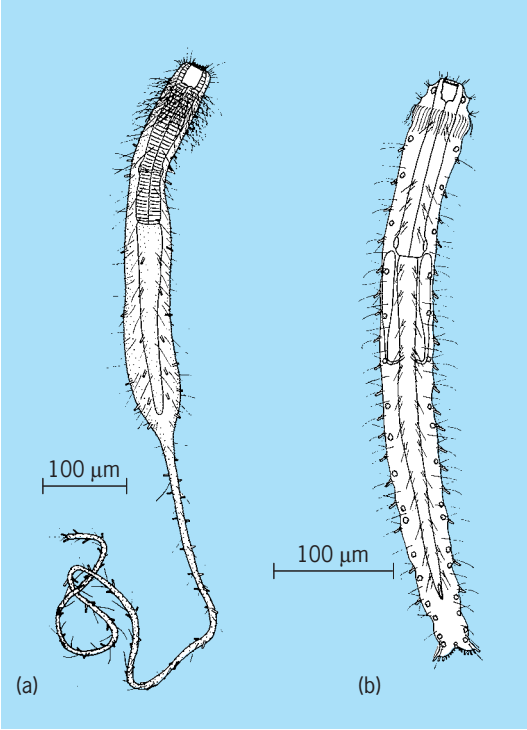

Marine gastrotrichs of the order Macrodasyida; dorsal views. (*a*) *Urodasys (courtesy of W. D. Hummon)*. (*b*) *Turbanella (after W. D. Hummon, Intertidal marine Gastrotricha from Colombia, Bull. Mar. Sci., 24:396–408, 1974)*.

these live in clean to detritus-rich marine sands of littoral or sublittoral areas. Both *Dolichodasys*, which reaches 2.5 mm (0.1 in.), and *Megadasys*, which reaches 4 mm (0.16 in.), inhabit black sulfide zones beneath quiet waters. All have front and rear groups of adhesive tubes; most also have tubes along their sides. Some have cuticular thickenings (*Lepidodasys* in one family) or scales or hooks (*Acanthodasys, Diplodasys, Tetranchyroderma*, and *Thaumastoderma* in another family). *Macrodasys* has a pointed rear, *Urodasys* has a tail (**illus.** *a*). Dactylopodolids, planodasyids, and turbanellids all have bilobed rear ends; lepidodasyids have rounded ones. *Turbanella* (illus. *b*) and *Tetranchyroderma* are the most common and most abundant of macrodasyids, with numbers of 50–100 per cubic centimeter of sand (800–1600 per cubic inch) not being unusual. *See* GASTROTRICHA. William D. Hummon

Macroevolution

Large-scale patterns and processes in the history of life, including the origins of novel organismal designs, evolutionary trends, adaptive radiations, and extinctions. Macroevolutionary research is based on phylogeny, the history of common descent among species. The formation of species and branching of evolutionary lineages mark the interface between macroevolution and microevolution, which addresses the dynamics of genetic variation within populations. The term macroevolution was used by the geneticist Richard Goldschmidt around 1940 to challenge the then prevailing notion that major features of evolutionary history could be understood as simple extrapolations of population genetic principles. The term still often implies this challenge and the expansion of Darwinian evolutionary theory to include evolutionary processes emerging above the species level on time scales encompassing multiple millions of years. Phylogenetic reconstruction, the developmental basis of evolutionary change, and long-term trends in patterns of speciation and extinction among lineages constitute major foci of macroevolutionary studies.

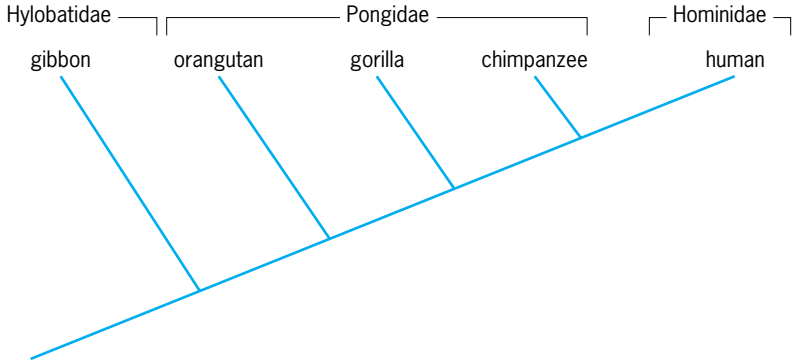

Fig. 1. Phylogenetic relationships of anthropoid primates showing traditional family-level taxa. All apes and humans together form a monophyletic group. The family Pongidae is paraphyletic, and therefore considered invalid by cladistic taxonomists. (*After C. P. Hickman, Jr., L. S. Roberts, and A. Larson, Integrated Principles of Zoology, 9th ed., 1993*)

Systematics and phylogeny. The goals of systematics are to reconstruct the phylogeny of species and to produce a taxonomy of species that reflects their phylogenetic relationships. Phylogenetic relationships are revealed by the sharing of evolutionarily derived characteristics among species, which provides evidence for common ancestry. Shared derived characteristics are termed synapomorphies, and are equated by many systematists with the older concept of homology. Characteristics of different organisms are homologous if they descend, with some modification, from an equivalent characteristic of their most recent common ancestor. Closely related species share more homologous characteristics than do species whose common ancestry is more distant. Species are grouped into clades according to patterns of shared homologies. The clades form a nested hierarchy in which large clades are subdivided into smaller, less inclusive ones, and are depicted by a branching diagram called a cladogram. A phylogenetic tree is a branching diagram, congruent with the cladogram, that represents real lineages of past evolutionary history; that is, it has a time division.

A cladogram or phylogenetic tree is necessary for constructing a taxonomy, but the principles by which higher taxa are recognized remain controversial. The traditional evolutionary taxonomy of G. G. Simpson recognizes higher taxa as units of adaptive evolution called adaptive zones. Species of an adaptive zone share common ancestry, and distinctive morphological or behavioral characteristics associated with use of environmental resources. Higher taxa receive Linnean categorical ranks (genus, family, order, and so forth) reflecting the breadth and distinctness of their adaptive zones. All taxa must have a single evolutionary origin, which means that the taxon must include the most recent common ancestor of all included species. A taxon is monophyletic if it contains all descendants of the group's most recent common ancestor, or paraphyletic if some descendants of the group's most recent common ancestor are excluded because they have evolved a new adaptive zone. For example, evolutionary taxonomy of the anthropoid primates groups the orangutan, gorilla, and chimpanzee in the paraphyletic family Pongidae and the humans in the monophyletic family Hominidae. Although the humans and chimpanzees share more recent common ancestry than either does with the gorilla or orangutan, the chimpanzees are grouped with the latter species at the family level and the humans are placed in a different family because they are considered to have evolved a new adaptive zone. The Hominidae and Pongidae together form a monophyletic group at a higher level (**Fig. 1**). *See* ANIMAL SYSTEMATICS.

Cladistic taxonomy or phylogenetic systematics accepts only monophyletic taxa because these alone are considered natural units of common descent. Such taxa form a nested hierarchy of groups within groups corresponding to the structure of the cladogram or phylogenetic tree. Linnean rankings are

considered unimportant. Taxa recognized using both the Simpsonian and cladistic taxonomies are standardly used in macroevolutionary analyses of extinction and patterns of diversity through time. The Simpsonian versus cladistic taxonomies often lead to fundamentally different interpretations, however. For example, extinction of a paraphyletic group, such as dinosaurs, would be considered pseudoextinction by cladists because some descendants of the group's most recent common ancestor survive. Birds are living descendants of the most recent common ancestor of all dinosaurs. The dinosaurs as traditionally recognized, therefore, do not form a valid cladistic taxon. *See* AVES; DINOSAURIA; PHYLOGENY.

Origins of evolutionary novelty. Comparative studies of organismal ontogeny are used to find where in development the key features of higher taxa appear and how developmental processes differ between taxa. Evolutionary developmental biologists denote the characteristic body plans of taxa by the term Bauplan, and the developmental stage during which key features of the Bauplan are formed by the term phylotypic stage. Developmental stages preceding and following phylotypic stages are subject to profound modification among species, but the phylotypic stages are evolutionarily stable. Specific patterns of gene expression that characterize some phylotypic stages in animal development have been identified through molecular genetics. The major characteristics of animal phyla and their developmental and molecular attributes appear to have arisen and stabilized early in the history of life, during the Cambrian Period. Subsequent evolutionary diversification builds upon the Baupläne established early in animal evolution. *See* CAMBRIAN.

Heterochrony. Particularly important to the evolutionary diversification of life are historical processes that generate change by altering the timing of organismal development, a phenomenon called heterochrony. Heterochronic changes may alter the ontogeny and life history of the organism as a whole or may be limited to certain structures or organs. The different heterochronic processes are described by using an ontogenetic trajectory that describes changes in form occurring throughout the organism's life. Ontogenetic changes are measured relative to the timing of reproductive maturation.

Heterochronic changes can produce either paedomorphic or peramorphic results. Paedomorphosis denotes the retention of preadult characteristics of ancestors in the adult stages of descendants; peramorphosis is the opposite outcome, in which the descendant ontogeny transcends that of the ancestor, adding new features at the final stages. Heterochronic changes can be produced by changing the rates of developmental processes or the times of their onset or termination. Axolotls of the salamander genus *Ambystoma* illustrate paedomorphosis produced by the heterochronic process neoteny, in which the rate of morphological development is decreased relative to reproductive maturation. The ancestral life history of *Ambystoma* is biphasic, possessing an aquatic larval stage followed by a metamorphosed, terrestrial adult. In the axolotl, reproductive maturity proceeds normally, but metamorphosis does not occur and the adult retains the morphological characteristics of the aquatic larva, including gills and tail fin.

Developmental dissociation occurs when different kinds of heterochronic change alter the development of different parts of the organism independently. Extensive dissociation can fundamentally restructure organismal ontogeny, producing ontogenetic repatterning. However, it is rare that a single heterochronic transformation affects all parts of the organism simultaneously, as seen in the axolotl example above. Dissociation and paedomorphosis together refute the statement that ontogeny recapitulates phylogeny, a claim that preadult stages of organismal development correspond to adult morphologies of ancestors. Ontogeny recapitulates phylogeny only in the unusual event that evolutionary change produces peramorphosis affecting all parts of the organism simultaneously. Ontogeny recapitulates phylogeny for individual structures or organs experiencing peramorphic evolution, whereas reverse recapitulation characterizes paedomorphic transformations. For most taxa, novel morphologies are produced by a mosaic of different heterochronic processes and by changes in the physical location of developmental events within the organism. *See* HETEROCHRONY.

Key innovation. A novel morphological or behavioral feature that alters the functional relationship between organisms and their environments is called a key innovation. A key innovation changes the selective factors acting on populational variation and leads to new directions of adaptive evolution. The trophically diverse cichlid fishes illustrate the concept of a key innovation and its evolutionary consequences. The key innovation in this group constitutes changes in the structure of the pharyngeal jaws that permit them to process food. Without the modified pharyngeal jaws, the maxillary and mandibular jaws must both obtain and process food items, preventing the jaws from becoming specialized for obtaining food. The cichlid pharyngeal jaws permit the maxillary and mandibular jaws to specialize for obtaining food in diverse ways, and this has produced an enormous diversity in feeding morphology and ecology among cichlid species. A key innovation or Bauplan of a taxon is evolutionarily stable because many subsequently evolved features are built upon it, causing it to become developmentally and functionally burdened; any major alteration would disrupt the adaptive system and have disastrous consequences. A trophically specialized cichlid fish would be unable to survive, for example, if its pharyngeal jaws lost their ability to process food.

Hierarchical expansion. Traditional Darwinian theory emphasizes natural selection acting on varying organisms within populations as the main causal factor of evolutionary change. Over many generations, the accumulation of favorable variants by natural selection produces new adaptations and new species. Macroevolutionary theory postulates two additional

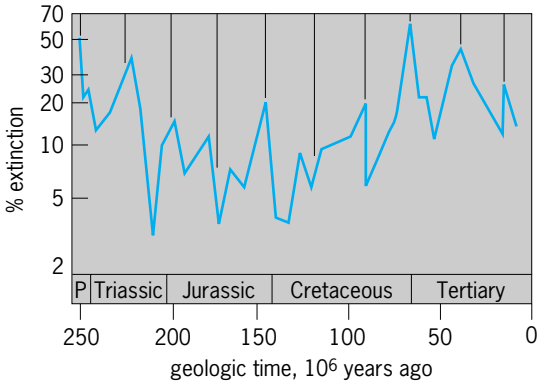

Fig. 2. Simpsonian patterns of mass extinction in the post-Permian fossil record of marine invertebrates. Mass extinctions (peaks) correspond approximately to 26-million-year intervals (vertical bars). (*After D. M. Raup and J. J. Sepkoski, Jr., Periodicity of extinctions in the geologic past, Proceedings of the National Academy of Sciences USA, 81:801–805, 1984*)

processes analogous to natural selection that act above the species level and on much longer time scales. An evolving lineage ultimately experiences one of two fates, branching speciation or extinction. Lineages that have a high propensity to produce new species and an ability to withstand extinction will dominate evolutionary history to the detriment of lineages that are prone to extinction and that have low rates of speciation. Biological characteristics are transmitted from ancestral to descendant species, and characteristics associated with lineages having high rates of speciation and low rates of extinction will prevail, even if they do not experience variation and selection within populations.

The higher-level process of differential speciation and extinction caused by the varying characteristics of species or lineages has been called species selection. Because the precise meaning of the term species is controversial, the more neutral terms lineage selection and clade selection are sometimes substituted for species selection. Most species show an evolutionary duration from a few million to approximately 10 million years in the fossil record between geologically instantaneous events of branching speciation. Species selection therefore generally occurs on a time scale of millions of years, rather than the generational time scale of natural selection. Species selection may be the primary factor underlying morphological evolutionary trends at this scale if lineages evolve by punctuated equilibrium, in which most morphological evolutionary change accompanies branching speciation, and species remain morphologically stable between speciational events. *See* SPECIATION.

The fossil record reveals mass extinctions in which enormous numbers of species from many different taxa are lost within a relatively short interval of geological time. For example, approximately 90% of the species and half of the families of shallow-water marine invertebrates were lost within a few million years during the Permian mass extinction (225 million years ago). A major recent discovery is that mass extinctions may occur at a periodicity of approximately 26 million years in the post-Permian fossil record of marine invertebrates (**Fig. 2**). This periodicity may represent periodic bombardment of the Earth by asteroids, causing massive environmental destruction. Some lineages may be better able to survive mass extinction events than others, and the characteristics that make a lineage prone to survive mass extinction may be very different from those that influence species selection between events of mass extinction. Catastrophic species selection denotes differential survival and extinction of lineages during events of mass extinction as determined by character variation among lineages. Prior to the Cretaceous mass extinction, dinosaur taxa dominated mammalian taxa, whereas mammals survived the mass extinction and then diversified extensively. The characteristics of the ancestral mammals may have permitted them to survive environmental challenges to which dinosaurs were susceptible. *See* EXTINCTION (BIOLOGY); FOSSIL; MAMMALIA; PALEONTOLOGY; PERMIAN.

Because natural selection, species selection, and catastrophic species selection can differ in the biological characteristics they promote, higher-level processes may undo or reverse evolutionary trends arising from lower-level processes. For this reason, macroevolutionists often resist attempts to explain evolutionary trends strictly from the occurrence of natural selection acting within populations. Species selection and catastrophic species selection are being investigated by paleontological research where processes occurring on geological time scales of multiple millions of years can be investigated. *See* ORGANIC EVOLUTION.

Allan Larson

Bibliography. W. Alvarez and F. Asaro, An extraterrestrial impact, *Sci. Amer.*, 263(4):78–84, 1990; N. Eldredge and J. Cracraft, *Phylogenetic Patterns and the Evolutionary Process*, 1980; S. J. Gould, Darwinism and the expansion of evolutionary theory, *Science*, 216:380–387, 1982; S. J. Gould, The paradox of the first tier: An agenda for paleobiology, *Paleobiology*, 11:2–12, 1985; S. J. Gould and N. Eldredge, Punctuated equilibrium comes of age, *Nature*, 366:223–227, 1993; S. J. Gould, N. L. Gilinsky, and R. Z. German, Asymmetry of lineages and the direction of evolutionary time, *Science*, 236:1437–1441, 1987; B. K. Hall, *Evolutionary Developmental Biology*, 2d ed., 1998; J. Levinton, *Genetics, Paleontology, and Macroevolution*, 2d ed., 2001; M. H. Nitecki, *Evolutionary Innovations*, 1990; R. M. Ross and W. D. Allmon, *Causes of Evolution*, 1990; A. Somit and S. A. Peterson, *The Dynamics of Evolution*, 1989.

Macromolecular engineering

The process of designing and synthesizing well-defined complex macromolecular architectures. This process allows for the control of molecular parameters such as molecular-weight/molecular-weight distribution, microstructure/structure, topology,

and the nature and number of functional groups. In addition, macromolecular engineering is the key to establishing the relationships between the precise molecular architectures and their properties. The understanding of the structure–property interplay is critical for the successful use of these elegantly tailored structures in the design of novel polymeric materials for applications such as tissue engineering, drug delivery, molecular filtration, micro- and optoelectronics, and polymer conductivity. Complex architectures, including star-shaped, branched, grafted, and dendritic-like polymers, have been prepared using living polymerization methods (for which there is no termination step to stop chain growth) such as anionic, cationic, living radical, metal-catalyzed polymerization, or combinations of these methods. *See* POLYMER; POLYMERIZATION.

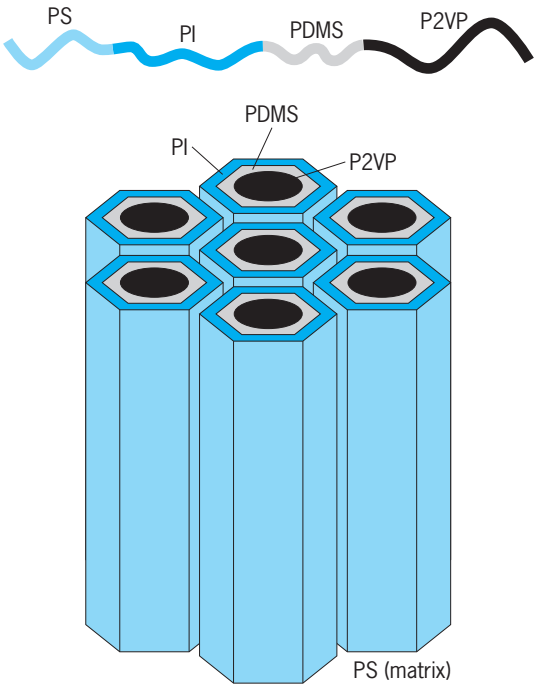

Fig. 1. Hexagonal triple coaxial cylindrical structure observed in PS-*b*-PI-*b*-PDMS-*b*-P2VP found by transmission electron microscopy. (*Reprinted with permission from Macromolecules, 35:4859–4861, 2002. Copyright 2002 American Chemical Society.*)

Anionic polymerization. The first organometallic miktoarm (mikto is from a Greek word meaning mixed) star copolymer of the AB_3 type, where A is polyferrocenyldimethylsilane (PFS) and B is polyisoprene-3,4 (PI), was synthesized by I. Manners and coworkers by anionic polymerization. The synthetic approach involved the anionic ring-opening polymerization of ferrocenophane, followed by reaction with excess silicon tetrachloride. After the removal of the excess of the volatile linking agent, the resulting macromolecular linking agent PFS-$SiCl_3$ was reacted with excess of living polyisoprenyllithium. The desired product (PFS)Si$(PI)_3$ was separated from

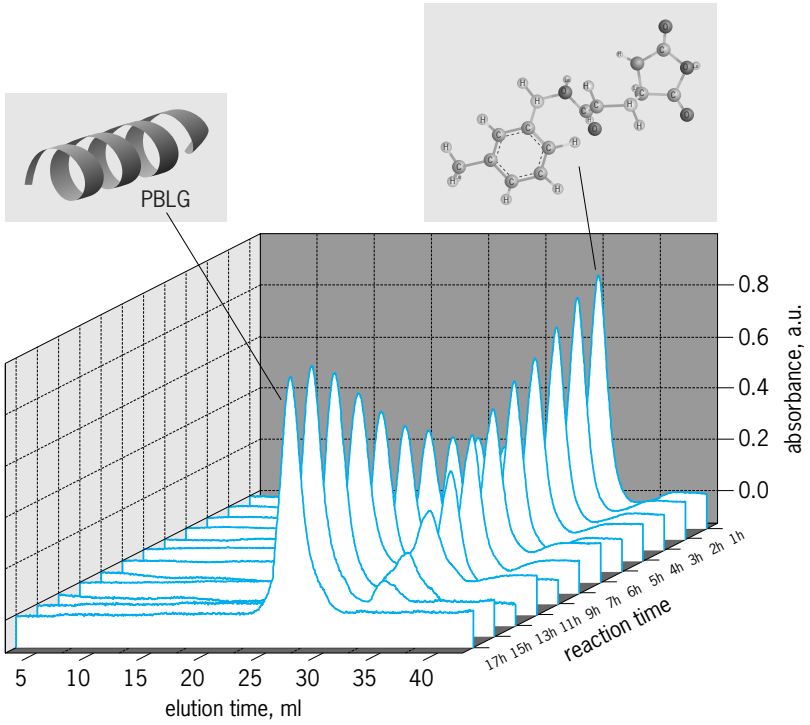

Fig. 2. Monitoring the living polymerization of γ-benzyl-L-glutamate NCA with *n*-hexylamine by size exclusion chromatography with a UV detector. The monomer is transformed 100% to polypeptide. (*Reprinted with permission from Biomacromolecules, 5(5):1653–1656, 2004. Copyright 2004 American Chemical Society.*)

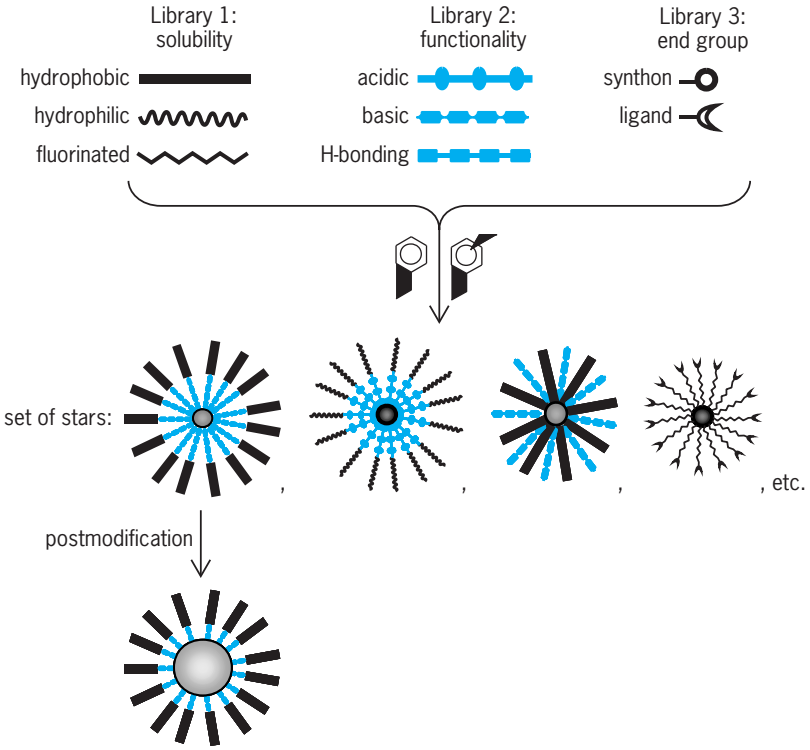

Fig. 3. Schematic representation of the modular approach to star polymers by nitroxide-mediated living radical polymerization. (*Reprinted with permission from J. Amer. Chem. Soc., 125:715–728, 2003. Copyright 2003 American Chemical Society.*)

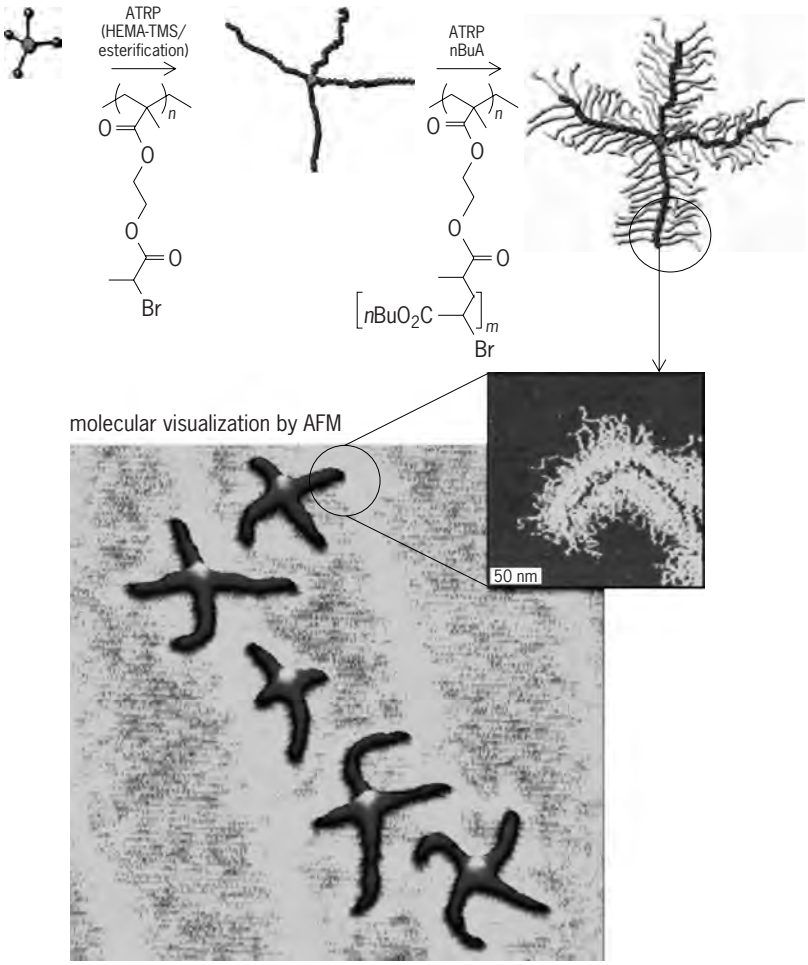

Fig. 4. Reaction sequence for the synthesis of four-arm star brushes, and AFM images. (*Reprinted with permission from Macromolecules, 36:1843–1849, 2003. Copyright 2003 American Chemical Society.*)

the excess PI arm by preparative size exclusion chromatography to yield a near-monodisperse miktoarm star copolymer. The presence of a metallopolymeric block can give redox activity, as well as semiconducting and preceramic properties, to these materials. *See* BRANCHED POLYMER; COPOLYMER; RING-OPENING POLYMERIZATION.

The first tetrablock quaterpolymers of polystyrene (PS), polyisoprene (PI), polydimethylsiloxane (PDMS), and poly(2-vinylpyridine) [P2VP] having different total molecular weights and compositions were synthesized by N. Hadjichristidis and coworkers. The synthetic strategy relied on recent advances in the living anionic polymerization of hexamethylcyclotrisiloxane and the use of a heterofunctional linking agent. This heterofunctional linking agent contains two functional groups, one of which reacts selectively with the polydimethylsiloxanyl lithium. The synthetic approach involved the sequential polymerization of styrene, isoprene, and hexamethylcyclotrisiloxane, followed by reaction with a stoichiometric amount of the heterofunctional linking agent 2-(chloromethylphenyl)ethyldimethyl chlorosilane. A slight excess of living poly(2-vinylpyridine) lithium was then added to react with the functional group of the resulting macromolecular linking agent PS-PI-PDMS-Ph-CH_2Cl. The excess P2VP was removed by fractional precipitation to give the pure tetrablock quaterpolymers. Due to the incompatibility of the four different blocks, these new quaterpolymers gave a unique four-phase triple coaxial cylindrical microdomain morphology never obtained before (**Fig. 1**). These structures could be used, for example, as multifunctional sensors or multiselective catalysts for sequential or simultaneous chemical reactions.

New ways of living ring-opening polymerization of α-amino acid-*N*-carboxy anhydrides (NCA) to yield polypeptides have been presented, thereby resolving a problem which had existed for more than 50 years. H. Schlaad and coworkers synthesized polystyrene-*b*-poly(γ-benzy-L-glutamate) block copolymers by using the hydrochloric salt of an ω-functionalized —NH_2 polystyrene as the macroinitiator of the ring-opening polymerization of γ-benzy-L-glutamate NCA. The use of the hydrochloric acid, instead of the free amine, suppressed the termination reactions. Hadjichristidis and coworkers synthesized well-defined di- and triblock copolypeptides, as well as star-shaped homo- and copolypeptides, by extensively purifying the solvent (dimethylformamide), the monomers (NCAs), and initiator (primary amines) in order to maintain the conditions necessary for the living polymerization. A variety of NCAs, such as γ-benzy-L-glutamate NCA, ϵ-carbobenzoxy-L-lysine NCA, glycine NCA, and *O*-benzyl-L-tyrosine NCA, were used to synthesize several combinations of block copolypeptides, showing that this is a general methodology (**Fig. 2**). The combination of the self-assembly of block copolymers and the highly ordered three-dimensional structures of proteins is expected to give novel supramolecular structures. *See* SUPRAMOLECULAR CHEMISTRY.

Living radical polymerization. A modular strategy for preparing functional multiarm star polymers by using nitroxide-mediated, living radical polymerization was proposed by J. M. Fréchet, C. Hawker, and coworkers. The approach used a variety of alkoxyamine-functional initiators for the polymerization of several vinyl monomers. These linear chains, containing a dormant chain end, were coupled with a crosslinkable monomer, such as divinylbenzene, resulting in a star polymer. The ability of this initiator to polymerize a great variety of vinyl monomers, along with the great diversity of the block sequence, led to the synthesis of a myriad of functionalized three-dimensional star polymers. A few examples are given in **Fig. 3**. The diversity of the method becomes even higher when performing postpolymerization reactions on specific blocks of the stars. These unique structures are useful in a range of applications such as supramolecular hosts, catalytic scaffolds, and substrates for nanoparticle formation.

K. Matyjaszewski and coworkers synthesized a series of three- and four-arm stars of poly(bromo-

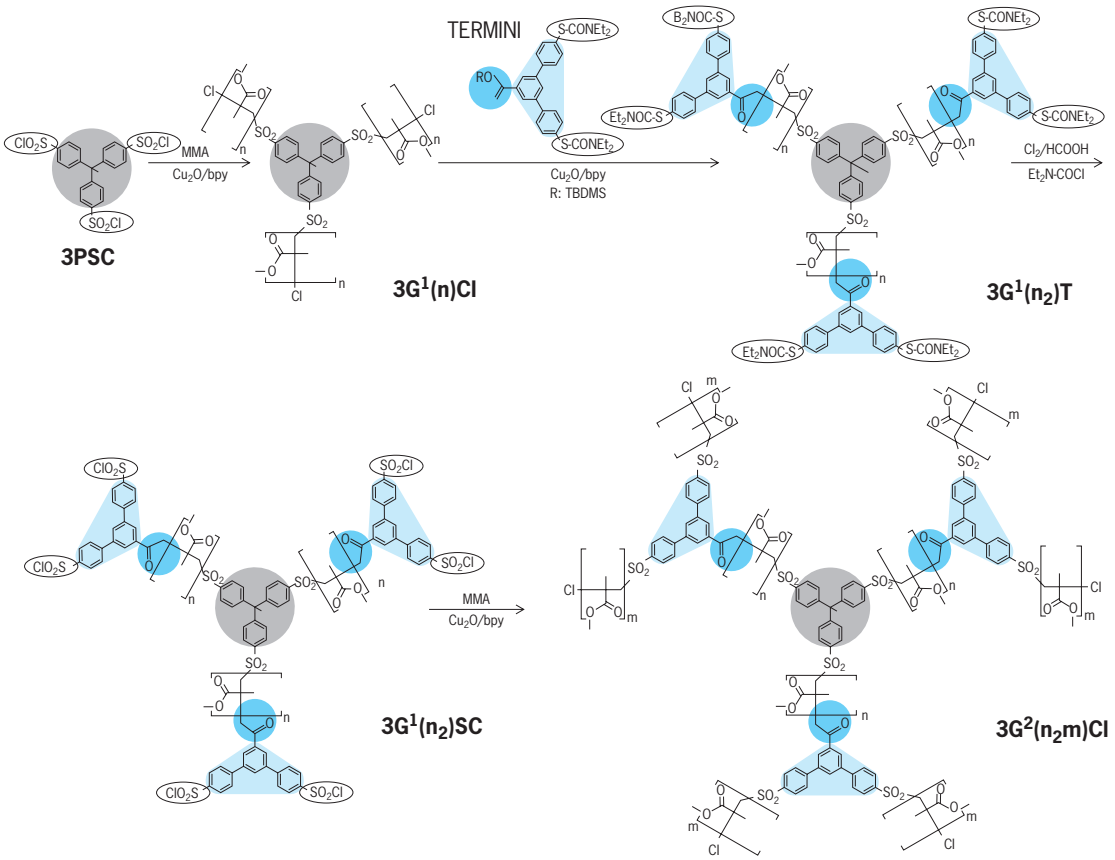

Fig. 5. Divergent iterative synthetic strategy for synthesis of dendritic PMMA. (*Reprinted with permission from J. Amer. Chem. Soc., 125:6503–6516, 2003. Copyright 2003 American Chemical Society.*)

propionylethyl-methacrylate-*g*-*n*-butyl acrylate) by using atom transfer radical polymerization (ATRP). The synthetic approach involved the use of tri- and tetrafunctional initiators for the ATRP of a 2-(trimethylsiloxyethyl)ethyl methacrylate (HEMA-TMS), which is a monomer containing protected hydroxyl groups. After the polymerization, the hydroxyl groups were deprotected and transformed to new bromine-containing groups, which are initiating sites of ATRP. The three- and four-arm star multifunctional macroinitiators were subsequently transformed to the densely grafted molecular brushes of poly(bromopropionylethyl-methacrylate-*g*-*n*-butyl acrylate), after the ATRP of poly(*n*-butyl acrylate) [Fig. 4]. The researchers were able to see the actual stars by atomic force microscopy (AFM). The synthesis of four-arm star brushes is given in **Fig. 4**.

A universal iterative strategy for the divergent synthesis of dendritic macromolecules was presented by V. Percec and coworkers. They used living radical polymerization of conventional monomers, such as methyl methacrylate (MMA), to synthesize the polymeric blocks of the dendrimers, followed by reaction with a specially designed compound, called TERMINI (from TERminator Multifunctional INItiator). TERMINI acts both as a terminator of the living polymer chains and as a masked difunctional initiator. Consequently, after terminating the living polymer and demasking, TERMINI is able to reinitiate the polymerization of MMA by ATRP, generating two new PMMA branches per TERMINI unit with nearly 100% efficiency. Using this methodology, dendrimers with up to four generations were synthesized (**Fig. 5**). The dendrimers were extensively characterized, showing that they exhibited high-molecular-weight homogeneity. *See* DENDRITIC MACROMOLECULE.

Catalytic polymerizations. Living polymerization of propylene was achieved by T. Fujita and coworkers

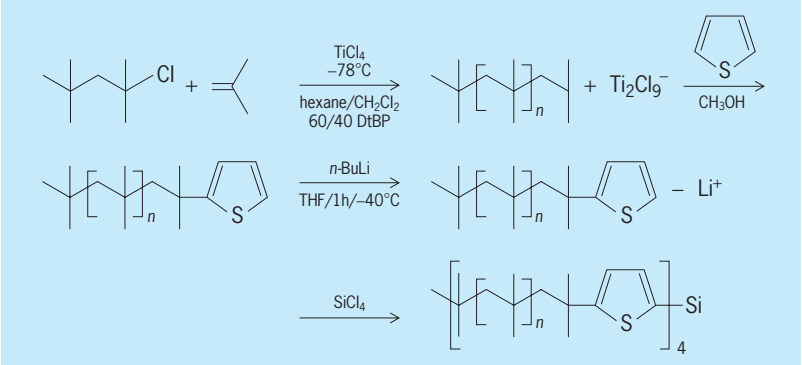

Fig. 6. Synthesis of four-arm PIB stars by site transformation of thiophene end-capped PIB chain ends.

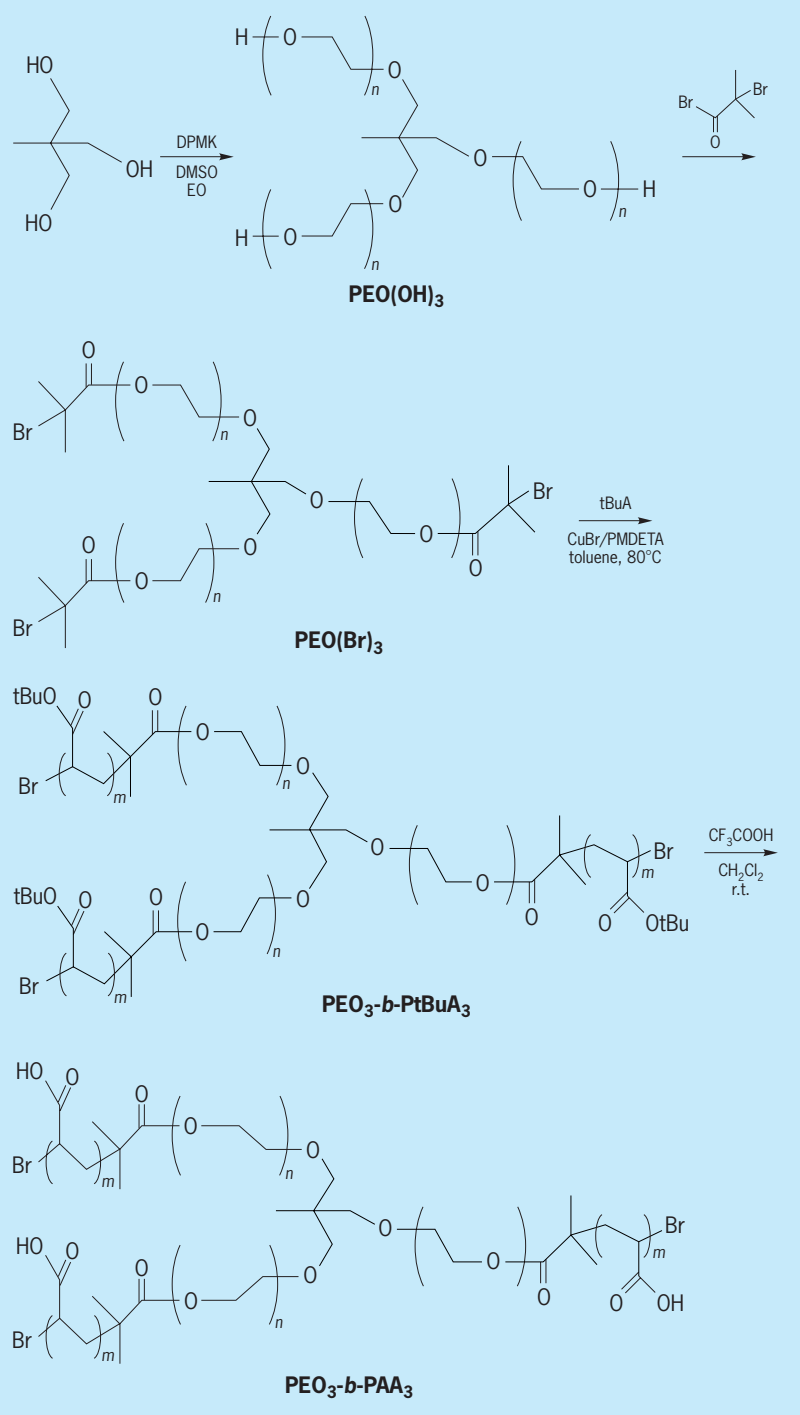

Fig. 7. Synthesis of PEO_3-*b*-PAA_3 star block copolymers.

using a series of titanium (Ti) complexes featuring fluorine-containing phenoxy-imine chelate ligands in the presence of methylaluminoxane (MAO). Despite the fact that the Ti catalyst possesses C_2 symmetry, the polypropylene produced is syndiospecific, through a chain-end control mechanism. Highly syndiotactic and monodisperse polypropylenes were synthesized at the temperature range between 0 and 50°C (32 and 122°F).

Well-defined poly(*n*-butyl methacrylate-*b*-methyl methacrylate) block copolymers were prepared by Hadjichristidis and coworkers, using zirconocene catalysts, fluorophenyl borate cocatalysts, and sequential addition of monomers, starting from the polymerization of *n*-butyl methacrylate. The success of the procedure was based on previous kinetic investigations that have revealed the exact experimental conditions (such as the temperature, nature of catalytic system, catalyst and cocatalyst concentration, and time for the completion of polymerization) under which the polymerization is controlled and the termination reactions are minimized.

Combinations of techniques. Specific polymerization techniques are applicable to a limited number of monomers. Therefore, the combination of different polymerization methods is expected to lead to new and unique macromolecular architectures. A new and efficient procedure for the synthesis of polyisobutylene (PIB) stars and PIB-poly(*t*-butyl methacrylate) block copolymers was developed by A. Müller, R. Faust, and collaborators by combining living carbocationic and anionic polymerizations. Monoaddition of thiophene to living PIB leads to the formation of 2-polyisobutylene-thiophene macroinitiator, which was subsequently metalated with *n*-butyllithium to transform the living cations to the respective anions. The stable carbanionic species were then used to initiate the polymerization of *t*-butyl methacrylate. PIB four-arm stars were also prepared via coupling of the living macrocarbanions with $SiCl_4$ as a coupling agent (**Fig. 6**).

Y. Gnanou and coworkers prepared double hydrophilic star-block (PEO_3-*b*-PAA_3) and dendrimer-like copolymers (PEO_3-*b*-PAA_6) consisting of three inner poly(ethylene oxide) [PEO] arms and either three or six outer poly(acrylic acid) [PAA] arms. Three arm PEO stars, having terminal —OH groups, were prepared by anionic polymerization using a suitable trifunctional initiator. The hydroxyl functions were subsequently transformed to either three or six bromo-ester groups, which were used to initiate the polymerization of *t*-butyl acrylate by ATRP. Subsequent hydrolysis of the *t*-butyl groups yielded the desired products (**Fig. 7**). These double hydrophilic block copolymers exhibit stimuli-responsive properties with potential biotech applications.

Graft copolymers having poly(methyl methacrylate) [PMMA] backbone and polystyrene [PS], polyisoprene [PI], poly(ethylene oxide) [PEO], poly(2-methyl-1,3-pentadiene) [P2MP], and PS-*b*-PI branches were prepared by Hadjichristidis and coworkers using the macromonomer methodology. The methacrylic macromonomers were synthesized by anionic polymerization, whereas their homopolymerization and copolymerization with MMA were performed by metallocene catalysts. Relatively high macromonomer conversions were obtained in all cases. Parameters such as the flexibility, structure, and molecular weight of the macromonomer as well

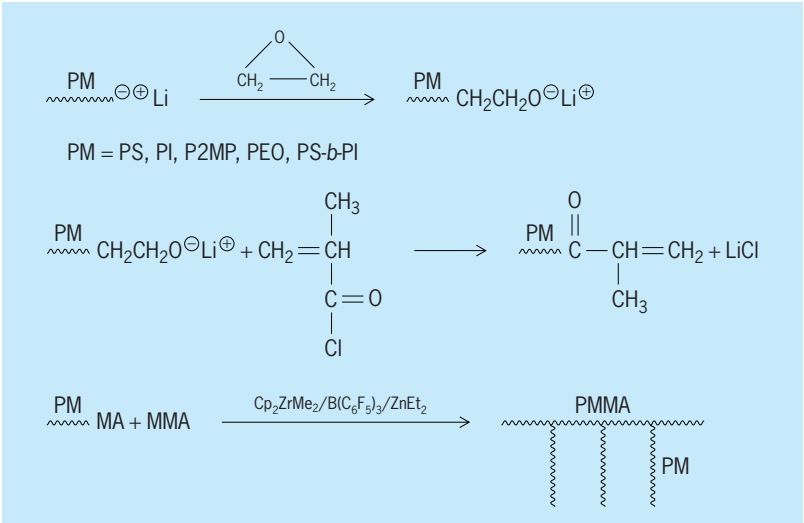

Fig. 8. Synthesis of graft copolymers by combination of anionic and metallocene-catalyzed polymerizations.

as the nature of the catalytic system were found to affect the copolymerization behavior (**Fig. 8**).

Nikos Hadjichristidis; Hermis Iatrou; Marinos Pitsikalis

Bibliography. J. M. Fréchet et al., A modular approach toward functionalized three-dimensional macromolecules: From synthetic concepts to practical applications, *J. Amer. Chem. Soc.*, 125:715–728, 2003; T. Fujita et al., Syndiospecific living propylene polymerization catalyzed by titanium complexes having fluorine-containing phenoxy-imine chelate ligands, *J. Amer. Chem. Soc.*, 125:4293–4305, 2003; Y. Gnanou et al., Synthesis and surface properties of amphiphilic star-shaped and dendrimer-like copolymers based on polystyrene core and poly(ethylene oxide) corona, *Macromolecules*, 36:8253–8259, 2003; Y. Gnanou et al., Synthesis of water-soluble star-block and dendrimer-like copolymers based on poly(ethylene oxide) and poly(acrylic acid), *Macromolecules*, 36:3874–3881, 2003; N. Hadjichristidis et al., *Block Copolymers: Synthetic Strategies, Physical Properties and Applications*, Wiley-Interscience, 2003; N. Hadjichristidis et al., Complex macromolecular architectures utilizing metallocene catalysts, *Macromolecules*, 36:9763–9774, 2003; N. Hadjichristidis et al., Four-phase triple coaxial cylindrical microdomain morphology in a linear tetrablock quarterpolymer of styrene, isoprene, dimethylsiloxane and 2-vinylpyridine, *Macromolecules*, 35:4859–4861, 2002; N. Hadjichristidis et al., Living polypeptides, *Biomacromolecules*, 5(5):1653–1656, 2004; N. Hadjichristidis et al., Synthesis and characterization of linear tetrablock quaterpolymers of styrene, isoprene, dimethylsiloxane and 2-vinylpyridine, *J. Polym. Sci., Polym. Chem.*, 42:514–519, 2004; I. Manners et al., Synthesis of the first organometallic miktoarm star polymer, *Macromol. Rapid Commun.*, 24:403–407, 2003; K. Matyjaszewski et al., Effect of initiation conditions on the uniformity of three-arm star molecular brushes, *Macromolecules*, 36:1843–1849, 2003; A. Müller et al., Polyisobutylene stars and polyisobutylene-block-poly(*tert*-butyl methacrylate) block copolymers by site transformation of thiophene end-capped polyisobutylene chain ends, *Macromolecules*, 36:6985–6994, 2003; V. Percec et al., Universal iterative strategy for the divergent synthesis of dendritic macromolecules from conventional monomers by a combination of living radical polymerization and irreversible terminator multifunctional initiator (TERMIND), *J. Amer. Chem. Soc.*, 125:6503–6516, 2003; H. Schlaad et al., Synthesis of nearly monodisperse polystyrene-polypeptide block copolymers *via* polymerisation of *N*-carboxyanhydrides, *Chem. Commun.*, 23:2944–2945, 2003.

Macroscelidea

An order of mammals containing the single family Macroscelididae (elephant shrews). Although formerly thought to be most closely associated with shrews, ungulates, and rabbits and hares, they are now thought to be most closely associated with elephants, aardvarks, golden moles, hyraxes, and sea cows on the basis of DNA investigations. *See* MAMMALIA.

General morphology. The 15 species of elephant shrews have chubby, soft-furred bodies, long hindlegs, and a long, slender, ratlike tail usually covered with knobbed bristles (see **illustration**). The fur is brownish or grayish. A gland on the underside of the tail produces a musky odor, especially in males. The hindlegs are much longer than the front limbs. Four or five toes are present on each foot. They may walk on all four feet, or when alarmed or pursued they hop on their hindlegs like a kangaroo. The eyes are large, the ears are well developed, and the elongated, trunklike snout is flexible and sensitive. The nostrils are located at the end of the snout. All

Elephant shrew, *Elephantulus*. (*Photo by Dr. Lloyd Glenn Ingles; © 2001 California Academy of Sciences*)

elephant shrews have long tongues that extend beyond the tips of their noses and are used to flick small food items into their mouths. The dental formula is I 1-3/3, C 1/1, PM 4/4, M 2/2-3 × 2 for a total of 36 to 42 teeth. Females have two or three pairs of mammae. The head and body length ranges 95–315 mm (3.5–12 in.) and the tail 80–265 mm (3–10 in.) The largest species is the checkered elephant shrew (*Rhynchocyon*) whose back is marked like a checkerboard, with four or five rows of evenly spaced light and dark square spots.

Behavior and habitats. Elephant shrews are found throughout Africa in a variety of habitats from the deserts of the Sahara to grassy plains, to rocky outcrops and boulder fields, to the well-forested regions of the Congo. They are primarily diurnal, but may be nocturnal during hot weather and on moonlit nights. They often bask and stretch out in the warm sunshine. While not a typical burrowing animal, as judged by its claws, the elephant shrew can quickly dig its way out of sight in reasonably soft soil or sand. The den may be a crevice in the rocks, a depression beneath a log, or a burrow it has dug in the ground. All burrows have a perpendicular emergency exit situated on the surface a few feet from the main entrance.

Food consists primarily of ants, termites, beetles, centipedes, spiders, and earthworms. The large checkered elephant shrew may also consume small mammals, birds, eggs, and mollusks. Some, like the short-eared elephant shrew (*Macroscelides*), may include shoots, roots, and berries in their diet. The elephant shrew, like most insect-eating animals, is not a particularly sociable creature: most species live either singly or in pairs. Several species including the golden-rumped elephant shrew (*Rhynchocyon chrysopygus*), the four-toed elephant shrew (*Petrodomus tetradactylus*), the short-eared elephant shrew (*Macroscelides proboscideus*), the rufous elephant shrew (*Elephantulus rufescens*), and the western rock elephant shrew (*Elephanteus rupestris*) are monogamous.

Reproduction. Breeding in some species may be continuous throughout the year, whereas in others it is seasonal. Following a gestation period of about 2 months, a litter of one or two young is born. Newborn elephant shrews are well developed, covered with fur, and usually have their eyes open. They can move about shortly after birth, and they nurse for a relatively short period of time. Sexual maturity is attained at 5 to 6 weeks.

Threats. Longevity in some species in the wild may be up to 4 years. Some of the larger species, such as the golden-rumped elephant shrew and the four-toed elephant shrew, may be taken for food in Kenya. Several species are classified as vulnerable or endangered by the International Union for the Conservation of Nature and Natural Resources (IUCN) primarily due to restricted areas of suitable habitat that are expected to decline due to human activity for subsistence farming, exotic tree plantations, and urban developments. Donald W. Linzey

Bibliography. D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford Limited, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Macrostomorpha

More than 300 species of small (mostly 1 mm size) flatworms usually with a ventral mouth situated about one quarter distant from the rostral tip of the body. A ciliated tubular foregut leads into a ciliated gut. Except for the Haplopharyngida, an anus is lacking. A hindgut is never differentiated. Macrostomorpha are covered by a single-layered, ciliated epidermis; they move by ciliary gliding and have special duo-gland adhesive papillae, at least in the tail region, which they use for temporary adhesion to and quick release from the substrate (only the marine, pelagic microstomid *Alaurina* does not possess papillae). The space between epidermis and gut is filled with parenchymal cells and muscle cells among others. The animals are hermaphroditic; sexual reproduction is typical. Larval stages are not known. Many species are found in marine sand and mud, and in various freshwater habitats. These flatworms were formerly members of the Turbellaria-Rhabdocoela, but are now included in the Rhabditophora, which is one of three monophyletic taxa constituting the Platyhelminthes (Acoelomorpha, Catenulida, and Rhabditophora). *See* PLATYHELMINTHES.

Taxonomy. Macrostomorpha are subdivided into two taxonomic groups of equal rank:

The Haplopharyngida are more than 5 mm (0.2 in.) long. They have a particularly pronounced glandular (sometimes highly muscular) proboscis-like organ in front of the mouth. Only two marine species in one family (Haplopharyngidae) are known.

The Macrostomida include all other Macrostomorpha. Three families are distinguished: (1) Microstomidae (three genera with 35 species) is known from many marine and freshwater habitats worldwide. The genus *Microstomum* (with 29 species) is characterized by asexual reproduction through transverse fission in addition to sexual reproduction. (2) Dolichomacrostomidae (35 species in three

subfamilies) is known only from marine sediments. The genus *Paromalostomum* shows worldwide distribution. These animals exhibit very complicated reproductive organs and an unpaired, common gonopore through which eggs are laid and sperm is transferred during copulation. (3) Macrostomidae includes most of the 11 genera of widely recognized species. Marine, brackish-water, and freshwater forms are found in all regions of the world. The best-known genus of Macrostomorpha is *Macrostomum*. The species *M. lignano* from the northern Adriatic can be cultured in large numbers and is a new model organism for developmental and evolutionary studies (see **illustration**). Another species with worldwide distribution, *M. orthostylum*, is cultured as food for fish larvae in India.

In the illustration, the general body plan of the Macrostomorpha is shown in *Microstomum lignano*, with hermaphroditic reproductive structures.

Based on body plan and molecular data, the Macrostomorpha are considered to be a basal-most branch of the Rhabditophora. Above Macrostomorpha, the following branches form a sequence from basal to most derived: Polycladida, Lecithoepitheliata, Prolecithophora, Proseriata, Tricladida, Rhabdocoela, and the parasitic Neodermata (with the Trematoda, Monogenea, and Cestoda).

Reinhard Rieger

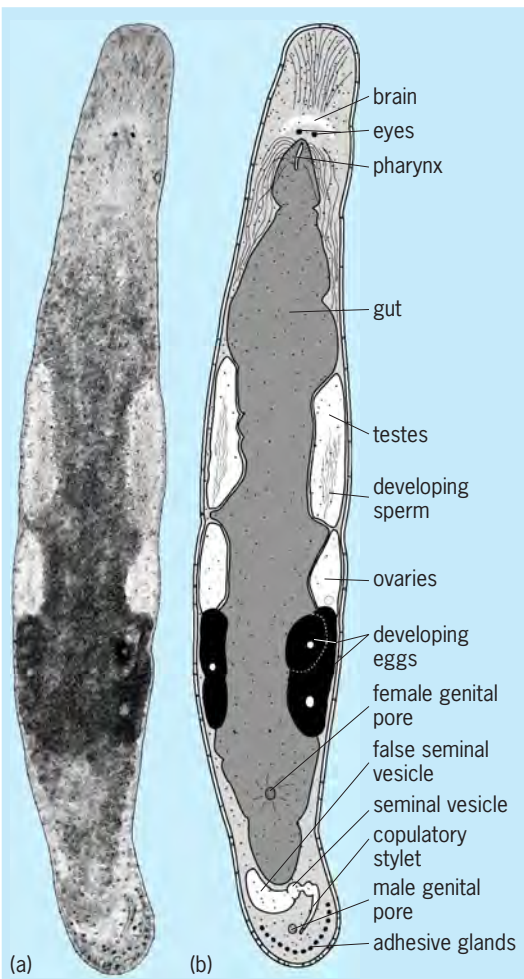

***Macrostomum lignano*, a new species from the northern Adriatic near Lignano, Italy. (*a*) Photomicrograph of a living, slightly squeezed worm. (*b*) Diagram showing the major organ systems. (*From L. Schärer, G. Joss, and P. Sandner, Mating behavior of the marine turbellarian Macrostomum sp., Mar. Biol., 145:373–380, 2004; with kind permission of Springer Science and Business Media*)**

Bibliography. L. R. G. Cannon, *Turbellaria of the World: A Guide to Families & Genera*, Queensland Museum Cultural Centre, Brisbane, 1986; M. Crezee, Turbellaria, in S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, pp. 718–740, McGraw-Hill, New York, 1982; R. Rieger, Phylogentic systematics of the Macrostomorpha, in D. T. J. Littlewood and R. A. Bray (eds.), *Interrelationship of the Platyhelminthes*, pp. 28–38, Taylor and Francis, New York, 2001.

Madelung constant

A numerical constant α_M in terms of which the electrostatic energy U of a three-dimensional periodic crystal lattice of positive and negative point charges q_+, q_-, N in number, is given by Eq. (1), where d is

$$U = -\frac{1}{2}\frac{Nq_+q_-}{d}\alpha_M \tag{1}$$

the nearest-neighbor distance between positive and negative charges and N is large. Knowledge of such electrostatic energies as given by the Madelung constant is of importance in the calculation of the cohesive energies of ionic crystals and in many other problems in the physics of solids. *See* IONIC CRYSTALS.

Designate the lattice sites by indices i or j, and let the distance between sites i and j be given by r_{ij}. Then U is given by the sum of the Coulomb interaction energies of all pairs of charges as shown in Eq. (2). In the summation, i and j range over all sites

$$U = \frac{1}{2}\sum_{i,j}{}' \frac{q_iq_j}{r_{ij}} \tag{2}$$

in the lattice, and the prime on the summation sign indicates the exclusion of terms for which $i = j$. The factor $^1/_2$ avoids counting pairs of charges twice. In Eq. (2) $q_i = q_+$ or q_-, depending upon whether the lattice site i is occupied by a positive or a negative ion. From this expression for U it can be seen by comparison of Eq. (2) with Eq. (1) that Eq. (3) holds.

$$\alpha_M = -\frac{1}{2}\sum{}'\left(\frac{q_i}{q_+}\right)\left(\frac{q_j}{q_-}\right)\left(\frac{d}{r_{ij}}\right) \tag{3}$$

This number is called the Madelung constant after E. Madelung, who first calculated the sum in connection with the cohesive energies of ionic crystals. It is characteristic of the lattice structure but independent of the dimensions of the lattice. Its numerical value does, however, depend on the

Madelung constants for some common ionic crystals

Name and formula	Madelung constant, α_M
Sodium chloride, NaCl	1.7476
Cesium chloride, CsCl	1.7627
Zinc blende, α-ZnS	1.6381
Wurtzite, β-ZnS	1.641
Fluorite, CaF_2	5.0388
Cuprite, CuO_2	4.1155
Rutile, TiO_2	4.816
Anatase, TiO_2	4.800
Corundum, Al_2O_3	25.0312

choice of the unit distance within the crystal, d. In Eq. (1), d has been chosen to be the nearest-neighbor positive-negative charge separation. Occasionally other choices of d are used instead, such as the cube root of the molecular volume or the unit cube edge in cubic crystals.

The calculation of Madelung constants requires some care since they are given by slowly and conditionally convergent series. Attempts at direct summation can involve thousands of terms without impressive accuracy. Various ingenious schemes for calculation have been used, but the most general and powerful is Ewald's method, also called the theta function method.

The Madelung constants for a number of common ionic crystal structures are given in the **table**. For these cases d is chosen as the nearest-neighbor distance. *See* CRYSTAL STRUCTURE. B. Gale Dick

Magellanic Clouds

Two small, irregular galaxies that are close companions of the Milky Way Galaxy. Both are nearby galaxies that are located in the southern sky, not far from the south celestial pole. When viewed without a telescope, they resemble small sections of the Milky Way that might have drifted away from the main arc of that soft, glowing belt of light. On a very dark clear night the Magellanic Clouds can also be mistaken for small atmospheric clouds because of their diffuse appearance. *See* MILKY WAY GALAXY.

The Large Magellanic Cloud (LMC) subtends an angular extent of about 5° in the sky (**Fig. 1**), and the Small Magellanic Cloud (SMC) is about 3° across. Telescopic studies show, however, that each is really much larger than it appears. Outer stars are detected over 10° and 6°, respectively, and there is even a faint bridge of stars that joins them tenuously together.

Because they are not visible from the Northern Hemisphere, the Magellanic Clouds were not known to European astronomers until Portuguese explorers brought back word of them from their travels down the coasts of Africa. They were given their names when Antonio Pigafetta, Ferdinand Magellan's navigator, published his description of the circumnavigation of the world, suggesting that these celestial objects be named for Magellan as a memorial. (Magellan did not complete his journey, having been killed in the Philippines.)

Distances. The Magellanic Clouds are now known to be rather small irregular-type galaxies. The Large Magellanic Cloud is at a distance of 160,000 light-years (1.5×10^{18} km or 9×10^{17} mi), and the Small Cloud is about 10% farther away. Their distances have been measured by many methods, including observations of Cepheid and RR Lyrae variables and (for the Large Magellanic Cloud) Supernova 1987A. The explosion of the supernova was seen in February 1987, and then several days later the effects were detected of its burst of light on a circumstellar ring of material that had been formed around the star before its final explosion. The time difference, combined with the angular size of the ring, provided a measurement of the distance to the Large Magellanic Cloud. *See* CEPHEIDS; SUPERNOVA; VARIABLE STAR.

Motions. The motions of the Magellanic Clouds across the sky have been measured by precise astrometry over many years. Combined with radial velocities measured spectroscopically, they provide knowledge of the motion of the Clouds in space. In both cases it is clear that these galaxies are satellites of the Milky Way Galaxy, with orbits that bring them close enough to the Milky Way and to each other to provide strong tidal effects. Indications are that such close encounters have led to periods of enhanced star formation.

Star clusters. In addition to some 5×10^9 ordinary stars, the Large Magellanic Cloud has a suite of about 10 massive old globular star clusters and several thousand smaller, younger star clusters. It was shown in 1998, using Hubble Space Telescope measurements, that the oldest globular clusters in the Large Magellanic Cloud have the same ages as the oldest such clusters in the Milky Way Galaxy, indicating that both galaxies must have formed at nearly the same time.

The Small Magellanic Cloud may have formed more recently, perhaps 10×10^9 rather than

Fig. 1. Large Magellanic Cloud. (*Harvard Observatory photo by P. Hodge*)

Fig. 2. Giant gas cloud centered on 30 Doradus in the Large Magellanic Cloud. (*Cerro Tololo Interamerican Observatory photo by P. Hodge*)

15×10^9 years ago, as no very old clusters have yet been identified among its several hundred star clusters. It is less massive than the Large Magellanic Cloud, with a total content of about 5×10^8 stars. Both Magellanic Clouds have large numbers of very young stars, most of which are located in stellar associations, each of which contains several hundred recently formed stars in loose aggregates about 200 light-years (2×10^{15} km or 1.2×10^{15} mi) across. *See* STAR CLUSTERS.

Gas and nebulae. Both objects also are rich in gas, mostly neutral hydrogen gas, detected by radio telescopes, such as the giant Australia Telescope array. In certain areas the gas is heated by nearby bright stars, producing brilliant glowing nebulae. The brightest and biggest is the 30 Doradus nebula (see **Fig. 2**), one of the most remarkable objects of the nearby universe. It extends over 500 light-years (5×10^{15} km or 3×10^{15} mi), and its primary source of energy is the ultraviolet light from many very young, massive stars, including about 10 extreme stars that are clustered together so compactly that ordinary ground-based telescopes mistakenly led astronomers to think that only one immense star was involved. Special interferometry techniques and the Hubble Space Telescope resolved this remarkable object in the 1990s, providing an example of a rare phenomenon: a "nursery" for the formation of supergiant stars. *See* INTERSTELLAR MATTER; NEBULA; SUPERGIANT STAR; SUPERMASSIVE STARS.

The gas clouds of both galaxies can be used to measure their chemical compositions, which are different from that of the Sun and the nearby stars. The Small Magellanic Cloud is depleted in heavy elements by a factor of about 6, and the Large Magellanic Cloud by a factor of about 2. *See* GALAXY, EXTERNAL.

Paul Hodge

Bibliography. P. Hodge and F. Wright, The Large Magellanic Cloud, *Smithson. Publ.*, no. 4699, 1967; P. Hodge and F. Wright, *The Small Magellanic Cloud*, University of Washington Press, 1977; B. Westerlund, *The Magellanic Clouds*, Cambridge University Press, 1998.

Magic numbers

The number of neutrons or protons in nuclei which are required to fill major quantum shells. They occur at particle numbers 2, 8, 20, 50, and 82.

In atoms, the electrons that orbit the nucleus fill quantum electron shells at atomic numbers of 2 (helium), 10 (neon), 18 (argon), 36 (krypton), and 54 (xenon). These elements are chemically inert and difficult to ionize because the energies of orbits are grouped in bunches or shells with large gaps between them. In nuclei, an analogous behavior is found; quantum orbits completely filled with neutrons or protons result in extra stability. The neutrons and protons fill their quantum states independently, so that both full neutron and full proton shells can occur as magic nuclei. In a few cases, for example oxygen-16 ($^{16}_{8}O_8$) and calcium-40 ($^{40}_{20}Ca_{20}$), doubly magic nuclei have full neutron and proton shells. Between the major shell gaps, smaller subshell gaps cause some extra stabilization and semimagic behavior is found at particle numbers 14, 28, 40, and 64. *See* ATOMIC STRUCTURE AND SPECTRA; ELECTRON CONFIGURATION.

In very heavy nuclei the Coulomb repulsion between the protons results in a different sequence of states for neutrons and protons and different major shell gaps. For neutrons the magic sequence continues at $N = 126$; the next shell gap is predicted at $N = 184$. For protons the next major shell gap is anticipated at $Z = 114$. The latter shell gaps lie beyond the heaviest nuclei known, but calculations indicate that the extra stability gained by producing nuclei with these particle numbers may result in an island of long-lived superheavy nuclei.

The closing of nuclear quantum shells has many observable consequences. The nuclei are more tightly bound than average, and the extra stability leads to anomalously high abundances of magic nuclei in nature. The full shells require unusually high energies to remove the least bound neutron or proton, and the probability of capturing extra particles is lower than expected. Furthermore, the full shells are spherically symmetric, and the nuclei have very small electric quadrupole moments. Many of these properties were known before the nuclear shell model was developed to account for quantum-level ordering and gaps between major shells. The different shell closures for atomic and nuclear systems reflect the differences between the Coulomb force that binds electrons to nuclei and the strong force that holds the nucleus together. An important component of the strong force in nuclei is the spin-orbit term, which makes the energy of a state strongly dependent on the relative orientation of spin and orbital angular momentum. *See* ANGULAR MOMENTUM; ELEMENTS, COSMIC ABUNDANCE OF; ISOTOPE; NUCLEAR MOMENTS; STRONG NUCLEAR INTERACTIONS.

Unlike atoms, nuclei that have both neutrons and protons in excess of magic closings are not confined to sphericity and can be polarized to exotic shapes. The bunching of quantum states into shells is

removed by a small deformation and the influences of shell gaps disappear. However, if the deformation becomes very large and nuclei become polarized to ellipsoidal shapes with axis ratios 2:1, 3:1, and so on, new symmetry is then imposed and the quantum states bunch into new shell groups with large gaps between them. Hence, deformed gaps and magic numbers appear which help stabilize these highly eccentric shapes. The production of nuclei in these highly elongated shapes is difficult and requires the fusion of heavy ions at high rotational frequency. The population of these superdeformed nuclei, the position of the deformed gaps and magic numbers, and the degree of stabilization they cause are of great interest. *See* EXOTIC NUCLEI; NUCLEAR STRUCTURE.

C. J. Lister

Bibliography. W. E. Burcham and M. Jobes, *Nuclear and Particle Physics*, 1994; A. De Shalit, *Theoretical Nuclear Physics*, 1990; K. S. Krane, *Introduction to Nuclear Physics*, 1987.

Magma

The hot material, partly or wholly liquid, from which igneous rocks form. Besides liquids, solids and gas may be present in magma. Most observed magmas are silicate melts with associated crystals and gas, but some inferred magmas are carbonate, phosphate, oxide, sulfide, and sulfur melts.

Strictly, any natural material which contains a finite proportion of melt (hot liquid) is a magma. However, magmas which contain more than about 60% by volume of solids generally have finite strength and fracture like solids.

Hypothetical, wholly liquid magmas which develop by partial melting of previously solid rock and segregation of the liquid into a volume free of suspended solids and gas are called primary magmas. Hypothetical, wholly liquid magmas which develop by crystallization of a primary magma and isolation of rest liquid free of suspended solids are called parental (or secondary) magmas. Although no unquestioned natural examples of either primary or parental magmas are known, the concepts implied by the definitions are useful in discussing the origins of magmas.

Bodies of flowing lava and natural volcanic glass prove the existence of magmas. Such proven magmas include the silicate magmas corresponding to such rocks as basalt, and andesite, dacite, and rhyolite as well as rare carbonate-rich magmas and sulfur melts. Oxide-rich and sulfide-rich magmas are inferred from textural and structural evidence of fluidity as well as mineralogical evidence of high temperature, together with the results of experiments on the equilibrium relations of melts and crystals. *See* LAVA.

Temperatures. The temperatures observed for silicate magmas flowing on the surface (lavas) range from slightly greater than 2200°F (1200°C) for basalts to about 1500°F (800°C) for rhyolites. The lower temperatures reflect both the nonequilibrium persistence of supercooled liquid and a lower temperature of melting at equilibrium for rhyolite than for basalt. Temperatures inferred for intrusive magmas vary widely and are both higher and lower than those observed for lavas. Both relations are possible, because some very hot magmas may cool before reaching the surface and because pressure enhances the solubility in melts of H_2O and other volatile substances which lower the equilibrium temperature of freezing. The volatile substances may largely escape during crystallization, and the minerals in a rock may interact during slow cooling; therefore, it is generally impossible to infer, with certainty, the temperatures of intrusive magmas.

Density. The mass density of silicate magmas (about 2.7 to 2.4 g/cm^3 or 1.6 to 1.4 oz/in^3) depends upon composition and temperature. In general, the density of the melt is less than that of crystals forming from it. Thus, buoyance forces lead to gravitative settling of crystals. In hot fluid basaltic magmas, millimeter-sized crystals may settle at rates of centimeters per hour. Indeed, some pillows of basalts 8 in. (20 cm) in diameter are known to contain accumulations of olivine crystals in their lower parts. In some iron-rich melts, feldspar crystals may float. Flotation of feldspar may occur on Earth rarely and probably has occurred on the Moon.

Many basaltic magmas are denser than continental crustal material and are necessarily forced to the surface either by hydrostatic head in the mantle, where the encasing rock is denser than the melt, or by gas propulsion. Most other magmas are less dense than most rock types and are buoyant everywhere within the crust and upper mantle. *See* BASALT.

Viscosity. Silicate magmas are variably viscous, ranging from about 100 poise (10 N·s/m^2) to more than 100,000 poise (10,000 N·s/m^2). Viscosity of magmas increases with decrease in temperature, with decrease in dissolved H_2O, with increase in crystal fraction, and with increase in fraction of SiO_2. In natural magmas the above variables are not independent; the viscosity of silicate magmas generally increases with fraction of SiO_2 in the melt. Thus, basaltic magmas produce flat-lying lava fields, whereas rhyolitic magmas commonly yield mounds and domes. Magmas with more than a few percent of crystals may have a finite yield strength.

Structure. Silicate melts are ionic solutions and have significant electrical conductivity (100 ohm^{-1}·cm^{-1}). Melts rich in silica are highly polymerized, which fact accounts for their high viscosity. Following the practice used in studies of slags, siliceous melts are commonly termed acidic (low activity of oxygen anion), and melts rich in basic oxides such as MgO and CaO are termed basic (high activity of oxygen anion).

Crystallization. Magmas crystallize as a consequence of cooling or effervescence, or both. Minerals grow in crystallizing magmas according to mass-action principles whereby the most abundant, least soluble constituents crystallize first, followed by less abundant, more soluble substances. With crystallization, most magmas display progressive enrichment to both crystals and residual melts in soda relative to lime and ferrous iron relative to magnesia. In many systems an immiscible sulfide melt develops

at an early stage. Minerals (such as apatite and zircon) which contain high concentrations of minor elements generally crystallize late and tend to have comparatively high ratios of surface area to volume. However, minerals which crystallize early in one magma may form later or not at all in other magmas. *See* PETROLOGY.

Magmas which cool and crystallize quickly (in a few decades or less) develop small crystals (a few millimeters in diameter or smaller). Magmas which cool and crystallize slowly (hundreds of thousands of years) develop large crystals (up to a few centimeters, rarely many tens of centimeters).

Occurrence. Magma is presumed to underlie regions of active volcanism and to occupy volumes comparable in size and shape to plutons of eroded igneous rocks. However, it is not certain that individual plutons existed wholly as magma at one time. Some catastrophic eruptions have yielded up to about 250 mi^3 (1000 km^3) of material on a void-free basis in a single event of short duration (few weeks or years), it is certain that some large bodies of magma have existed at one time, and presumably exist today. Magma may underlie some regions where no volcanic activity exists, because many plutons appear not to have vented to the surface.

Remote detection of bodies of magma is difficult because the properties of magma are closely approximated by hot fractured rock filled with water or water vapor. Seismic, gravimetric, and magnetic data are consistent with the existence of major bodies of magma beneath Yellowstone Park; Katmai-Trident, Alaska; several regions in Kamchatka; Japan; and elsewhere.

Origin. Diverse origins are probable for various magmas. Basaltic magmas because of their high temperatures probably originate within the mantle several tens of kilometers beneath the surface of the Earth. Experiments at high pressure have led many scientists to conclude that most basaltic magmas are modified by crystallization and separation of residual liquid from crystals between the site of melting in the mantle and the place of crystallization within the crust. Rhyolitic magmas may originate through crystallization differentiation of basaltic magmas or by melting of crustal rock. Intermediate magmas may originate within the mantle or by crystallization of basaltic magmas, by melting of appropriate crustal rock, and also by mixing of magmas or by assimilation of an appropriate rock by an appropriate magma. *See* IGNEOUS ROCKS; VOLCANO.

Alfred T. Anderson

Bibliography. M. G. Best, *Igneous and Metamorphic Petrology,* 1982; L. Civetta et al., *Physical Volcanology,* 1974; C. J. Hughes, *Igneous Petrology,* 1982; A. R. McBirney, *Igneous Petrology*, 2d ed., 1992.

Magnesite

A member of the calcite-type car-bonates having the formula $MgCO_3$. It forms dolomite [$CaMg(CO_3)_2$] with calcite ($CaCO_3$) in the system $CaCO_3$—$MgCO_3$. Pure magnesite is not common in nature because there exists a complete series of solid solutions between $MgCO_3$ and $FeCO_3$, which is constantly present in magnesite in its natural occurrence. *See* CARBONATE MINERALS; MAGNESIUM.

The calcite-type structure of magnesite has $a = 0.4632$ nm, $c = 1.5032$ nm, $z = 6$, and space group $R\bar{3}c$. The interatomic distances are C—O, 0.12851 nm, and Mg—O, 0.21018 nm, which produce a MgO octahedron with less distortion than that of CaO_6 in calcite. Magnesite decomposes to MgO and CO_2 in the temperature range of 625–643°C as determined by differential thermal analysis. Crystallization of magnesite in aqueous solution is inhibited by species of magnesium in solution and strong positive catalysis by ionic strength and CO_2 pressure. Therefore, highly saline conditions provide a favorable environment for magnesite crystallization.

Magnesite is uniaxial and optically negative with extreme birefringence (0.191). Refractive indices are $\epsilon = 1.509$ (extraordinary index) and $\omega = 1.700$ (ordinary index). Magnesite is usually white, but it may be light to dark brown if iron-bearing. The hardness of magnesite is $3\frac{1}{2}$–$4\frac{1}{2}$ on the Mohs scale, and the specific gravity is 3.00. *See* BIREFRINGENCE; CRYSTAL OPTICS; HARDNESS SCALES.

Magnesite deposits are of two general types: massive and crystalline. Massive magnesite is an alteration product of serpentine which has been subjected to the action of carbonate waters. Crystalline magnesite is usually found in association with dolomite. It is generally thought to be a secondary replacement of magnesite in preexisting dolomite by magnesium-rich fluids. Massive magnesite deposits are found in Greece, Turkey, India, China, North Korea, and Australia, while those of crystalline magnesite occur in the United States, Austria, Spain, Russia, and the Czech Republic.

Magnesite is an important industrial mineral. Various types of magnesite or magnesia (MgO) are produced by different thermal treatment. The caustic-calcined magnesite or magnesia is used in the chemical industry for the production of magnesium compounds, while dead-burned or sintered magnesite or magnesia is used in refractory materials. Fused magnesia is used as an insulating material in the electrical industry because of its high electrical resistance and high thermal conductivity. Luke L. Y. Chang

Bibliography. L. Chang, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, vol. 5B: *Non-Silicates*, 2d ed., Geological Society Publishing House, 1998.

Magnesium

A metallic chemical element, Mg, in group 2 of the periodic system, atomic number 12, atomic weight 24.312. Magnesium is silvery white and extremely light in weight. The specific gravity is 1.74, and the density is 1740 kg/m^3 (0.063 $lb/in.^3$ or 108.6 lb/ft^3). Because of this lightness combined with alloy strength suitable for many structural

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

uses, magnesium has long been known as industry's lightest structural metal. *See* PERIODIC TABLE.

With a density only two-thirds that of aluminum, magnesium is used in countless applications where weight saving is an important consideration. The metal also has, however, many desirable chemical and metallurgical properties which account for its extensive use in a variety of nonstructural applications.

Magnesium is very abundant in nature, occurring in substantial amounts in many rock-forming minerals such as dolomite, magnesite, olivine, and serpentine. In addition, magnesium is also found in sea water, subterranean brines, and salt beds. It is the third most abundant structural metal in the Earth's crust, exceeded only by aluminum and iron.

Some of the properties of magnesium in metallic form are listed in **Table 1**. Magnesium is very active chemically. It will actually displace hydrogen from boiling water, and a large number of metals can be prepared by thermal reduction of their salts and oxides with magnesium. The metal will combine with most nonmetals and with practically all acids. Magnesium reacts only slightly or not at all with most alkalies and many organic chemicals, including hydrocarbons, aldehydes, alcohols, phenols, amines, esters, and most oils. As a catalyst, magnesium is useful for promoting organic condensation, reduction, addition, and dehalogenation reactions. It has long been used for the synthesis of complex and special organic compounds by the well-known Grignard reaction. Principal alloying ingredients include aluminum, manganese, zirconium, zinc, rare-earth metals, and thorium. *See* MAGNESIUM ALLOYS.

Magnesium compounds are used extensively in industry and agriculture. **Table 2** lists the major magnesium compounds and indicates some of their more significant applications.

William H. Gross; Stephen C. Encleson

TABLE 1. Physical properties of primary magnesium (99.9% pure)

Property	Value
Atomic number	12
Atomic weight	24.312
Atomic volume, cm^3/g-atom	14.0
Crystal structure	Close-packed hexagonal
Electron arrangement in free atoms	(2) (8) 2
Mass numbers of the isotopes	24, 25, 26
Percent relative abundances of ^{24}Mg, ^{25}Mg, ^{26}Mg	77, 11.5, 11.5
Density, g/cm^3 at 20°C	1.738
Specific heat, cal/g/°C at 20°C (1 cal = 4.2 joules)	0.245
Melting point, °C	650
Boiling point, °C	1110 ± 10

TABLE 2. Principal magnesium compounds and uses

Compound	Uses
Magnesium carbonte	Refractories, production of other magnesium compounds, water treatment, fertilizers
Magnesium chloride	Cell feed for production of metallic magnesium, oxychloride cement, refrigerating brines, catalyst in organic chemistry, production of other magnesium compounds, flocculating agent, treatment of foliage to prevent fire and resist fire, magnesium melting and welding fluxes
Magnesium hydroxide	Chemical intermediate, alkali, medicinal
Magnesium oxide	Insulation, refractories, oxychloride and oxysulfate cements, fertilizers, rayon-textile processing, water treatment, papermaking, household cleaners, alkali, pharmaceuticals, rubber filler catalyst
Magnesium sulfate	Leather tanning, paper sizing, oxychloride and oxysulfate cements, rayon delustrant, textile dyeing and printing, medicinal, fertilizer ingredient, livestock-food additive, ceramics, explosives, match manufacture

The magnesium ion (Mg^{2+}) is an essential component of all living systems; it is the second most abundant positive ion within cells and the fourth most common cation in the body. Magnesium plays a structural role in bones and teeth. More importantly, magnesium ion is involved in a variety of cellular biochemical processes. Intracellular magnesium ion either is present in its free form or is bound to biomolecules such as adenosine triphosphate (ATP), nucleic acids, chlorophyll, and Mg^{2+}-activated enzymes. The bound form of intracellular magnesium ion predominates over its free from because of the high charge density of this metal ion. The partitioning of intracellular magnesium ion between bound and free forms regulates the biological activity of many Mg^{2+}-dependent enzymes, which participate in glycolysis, muscle contraction, ion transport, protein synthesis, respiration, photosynthesis, and signal transduction. The free intracellular magnesium ion concentration, unlike that of the calcium ion (Ca^{2+}), is well maintained within the cell; it has therefore been suggested that magnesium ion is not a trigger of cell activity but a static regulator. *See* BIOINORGANIC CHEMISTRY.

Duarte Mota de Freitas

Bibliography. B. Alberts et al., *Molecular Biology of the Cell*, 4th ed., Garland Publishing, 2002; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999.

Magnesium alloys

Solid solutions of magnesium and aluminum, zinc, manganese, silicon, zirconium, rare-earth metals, thorium, or yttrium. The specific gravity of magnesium alloys ranges from 1.74 to 1.83. This low specific gravity has led to a great many structural applications in the aircraft, transportation, materials-handling, and portable-tool and -equipment industries. *See* MAGNESIUM.

The magnesium-aluminum-zinc (Mg-Al-Zn) alloys are the most important commercial alloys. These alloys are produced in the form of sand, permanent-mold, and die castings; extrusions; rolled sheet and plate; and forgings. They are also cast by some of the less common methods, including plaster-molding, centrifugal-molding, shell-molding, and investment-molding processes. Their properties may be modified by appropriate heat treatments. Sand and permanent-mold castings are used primarily in the solution heat-treated condition and in the solution heat-treated plus artificially aged condition. The alloys are used in the solution heat-treated condition, where maximum ductility and toughness with adequate strength are required, and in the artificially aged condition, where maximum strength is needed and somewhat lower ductility can be tolerated. Die castings are used in the as-cast condition. Extrusions and forgings are used either as fabricated or after an artificial aging treatment. Sheet and plate are produced primarily in two tempers, annealed and strain-hardened plus partially annealed.

Magnesium-zinc-zircon alloys. The magnesium-zinc-zircon alloys are high-strength alloys with good ductility and toughness. They are generally used in the artificially aged condition, but for maximum strength in forgings they are solution-heat-treated followed by artificial aging.

Some magnesium-zinc-zircon alloys can also be used as sand castings. These alloys are high-strength, having an excellent combination of mechanical strength, ductility, and toughness. However, they are poor in overall castability, having a high tendency toward hot cracking and microshrinkage porosity. They are suitable only for producing relatively simple castings which do not have pressure-tightness requirements and which are not intended for use at temperatures above 250°F (121°C). The addition of misch metal and thorium to these alloys greatly improves their castability and weldability. *See* ZINC; ZIRCONIUM.

Rare-earth and thorium alloy ingredients. The alloys described thus far have the disadvantage of losing their strength rapidly with increasing temperature, especially above 300°F (149°C). The need for magnesium alloys with improved strength and creep resistance at elevated temperatures has been met by the development of alloys containing rare-earth metals or thorium, or both, as the principal alloying constituents. The temperature range over which magnesium alloys exhibit structurally useful mechanical strength has been markedly extended by the use of rare-earth metals and thorium. Magnesium-thorium alloys, for example, are used up to temperatures in the range of 700–900°F (371–482°C) and even higher, depending on the time duration of exposure at the elevated temperature.

The most effective rare earth is neodymium, and it is economically available in the form of didymium, which is also referred to as neodymium-rich rare-earth metal. Didymium is the rare-earth mixture remaining when both cerium and lanthanum are separated from misch metal. It consists of 80–85% neodymium, and the balance is substantially praseodymium with small amounts of samarium, gadolinium, cerium, and lanthanum. Didymium is used in two commercial casting alloys. Silver is a major alloying constituent, and zirconium is added for grain refinement and improved castability and weldability. Both alloys have exceptionally high strength from ambient temperature up to 600°F (316°C). There are no rare-earth alloys used for wrought products.

Several magnesium alloys containing thorium as the principal alloying ingredient can be used for wrought products and for cast products. The wrought and magnesium-thorium alloys are used for all applications involving exposure to temperatures above 250–300°F (121–149°C). An important consideration for an elevated temperature application, especially for long time exposure, is resistance to creep. Another characteristic important in structural design is modulus of elasticity. Magnesium alloys containing thorium and rare-earth metals retain a high modulus over a wide temperature range. The combination of light weight, high strength and modulus, and good creep resistance permits the design of structures having the highest stiffness-to-weight ratio over a wide temperature range. The use of rare earths and thorium has expanded markedly the applicability of magnesium alloys in the rather limited field of missiles and spacecraft and has retained a position for magnesium in aircraft engine applications.

The need for die-casting alloys with improved creep resistance at temperatures in the range of 275–375°F (135–190°C) has produced a magnesium-aluminum-silicon (Mg-Al-Si) alloy, that requires the use of higher casting temperatures; but at the same time the presence of silicon improves fluidity and results in overall better castability. It is especially suitable for automotive engines and power-train housings.

A little-known advantage of magnesium over other common structural materials is its inherently high damping capacity, that is, its capacity to absorb mechanical vibrations. In general, addition of alloying elements decreases this property. *See* THORIUM.

Magnesium-yttrium alloys. A family of magnesium alloys based on the addition of yttrium has surpassed the performance of all previous magnesium-based casting alloys, particularly in the areas of temperature tolerance and corrosion resistance. In fact, the general corrosion resistance of one member of this family is more comparable to aluminum alloys than to the less corrosion-resistant magnesium alloys. In

addition, the room-temperature tensile strength of this particular alloy exceeds proposed specification minimums.

Alloy preparation. Metal for the production of either shape castings or ingots for subsequent fabrication is prepared by melting and alloying in either stationary or tilting iron pots or crucibles. The capacity of the pots may vary from a few hundred pounds to 6000 lb (2700 kg). Induction furnaces or ceramic-lined reverberatory or muffle furnaces may also be used.

The surface of the molten metal is protected from atmospheric oxidation by a cover of flux consisting of a mixture of alkali and alkaline-earth halides. A technique for protecting the surface of molten magnesium involves the use of an atmosphere of air containing a few tenths of one percent sulfur hexafluoride. Elimination of the use of halide fluxes greatly alleviates environmental control problems and reduces metal loss and processing costs. Aluminum, zinc, thorium, misch metal, and didymium are usually added in the elemental form. Manganese and zirconium are alloyed by the addition of a proprietary fused salt containing either the chloride or fluoride of the desired metal which is reduced by the molten magnesium. Zirconium and thorium may be added in the form of magnesium hardeners containing 20–35% of the alloying element. Beryllium in amounts of 0.0003–0.001% is added particularly to magnesium die-casting alloys in order to minimize oxidation in the molten state during melting and casting operations.

Mg-Al-Zn alloys are grain refined either by superheating or by the addition of a carbonaceous material. Alloys containing zirconium are inherently fine grained. Molten magnesium alloys are transferred from one pot to another or from pot to mold by air-driven mechanical pumps through steel pipes.

Fabricability. In addition to being adaptable to all the primary working operations already mentioned, magnesium alloys can be fabricated by all the common metalworking processes such as stamping, deep and shallow drawing, blanking, coining, spinning, impact extrusion, and forging. For forging operations both press and hammer equipment are used, but the former is more commonly used because the physical structure of magnesium alloys makes the metals better adapted to the squeezing action of the forging press. *See* FORGING.

Magnesium alloys exhibit excellent machinability. They can be machined at higher speeds and with larger feeds and depths of cut than is possible with most other commonly used metals. Power requirements needed in machining are the lowest for magnesium among all structural metals. *See* MACHINABILITY OF METALS.

Magnesium alloy parts can be joined by any of the common methods. Inert-gas, shielded-arc, and electrical-resistance welding, adhesive bonding, and riveting are in daily production use. Brazing and gas welding, although not as frequently used as the other methods, are also suitable for joining magnesium alloys. *See* ADHESIVE BONDING; BRAZING; WELDING AND CUTTING OF MATERIALS.

A wide variety of protective and decorative surface-finishing systems can be applied to magnesium alloys. They can be treated chemically and electrochemically to produce a protective and paint-adherent surface. In addition they can be painted, electroplated, anodized, and clad with plastic sheathing. *See* METAL COATINGS.

Uses. Reviewing the nonstructural uses of magnesium gives an accurate perspective of the entire industry. Nonstructural uses are those in which magnesium is used for its particular chemical, electrochemical, and metallurgical properties. The principal nonstructural uses are as alloying constituent in other metals (principally aluminum); nodularizing agent for graphite in cast iron; reducing agent in the production of other metals, such as uranium, titanium, zirconium, beryllium, and hafnium; desulfurizing agent in iron and steel production; sacrificial anode in the protection of other metals against corrosion; and anode material in reserve batteries and dry cells.

As structural materials, magnesium alloys are best known for their light weight and high strength-to-weight ratio. Accordingly, they are used generally in applications where weight is a critical factor and where high mechanical integrity is needed. Magnesium alloys are most commonly used in the form of die casting. Of all the metalworking and fabricating processes, die casting is the most readily adaptable to the specific characteristics of magnesium. Magnesium has a relatively low melting point, low heat content per unit volume, and low reactivity toward ferrous materials. Coupling these characteristics with the high productivity of the die-casting process results in the most economical technique for the mass production of a large number of complex parts of the same configuration. Magnesium die castings are produced on hot-chamber die-casting machines. The hot-chamber machine is even more amenable to automation than the cold-chamber machine and is capable of higher production rates and overall lower operating costs.

The hot-chamber machine has greatly enhanced the potential for expansion of the use of magnesium die castings in the automotive and appliance industries. Some automobiles contain magnesium parts such as the distributor diaphragm, steering column brackets, and a lever cover plate. Magnesium-alloy die castings are also used on chain saws, portable power tools, cameras and projectors, office and business machines, tape reels, sporting goods, luggage frames, and many other products.

Uses of sand and permanent mold castings are confined largely to aircraft engine and airframe components. Engine parts include gearboxes, compressor housings, diffusers, fan thrust reversers, and miscellaneous brackets. On the airframe, magnesium sand castings are used for leading edge flaps, control pulleys and brackets, entry door gates, and various

cockpit components. This is a small market in volume, but one requiring highly technical skills in production.

The use of magnesium wrought products has become rather limited. Magnesium sheet, extrusions, and forgings continue to be used in a few limited air-frame applications. The high-temperature alloys are used in at least 20 different missiles (including ICBMs) and on various spacecraft. *See* ALLOY; HEAT TREATMENT (METALLURGY); METAL CASTING; METAL FORMING.

Thomas E. Leontis

Bibliography. American Foundrymen's Society, *Conference on Recent Advances in Magnesium Technology*, 1987; *Annual Book of ASTM Standards*, vol. 2.02: *Die-Aluminum and Magnesium Alloys*, 1993; National Fire Protection Association, *Storage, Handling, and Processing of Magnesium*, 1987; H. Paris and W. H. Hunt (eds.), *Advances in Magnesium Alloys and Composites*, 1989.

Magnet

An object or device that produces a magnetic field. Magnets are essential for the generation of electric power and are used in motors, generators, labor-saving electromechanical devices, information storage, recording, and numerous specialized applications, for example, seals of refrigerator doors. The magnetic fields produced by magnets apply a force at a distance on other magnets, charged particles, electric currents, and magnetic materials. *See* GENERATOR; MAGNETIC RECORDING; MOTOR.

Magnets may be classified as either permanent or excited. Permanent magnets are composed of so-called hard magnetic material, which retains an alignment of the magnetization in the presence of ambient fields. Excited magnets use controllable energizing currents to generate magnetic fields in either electromagnets or air-cored magnets. Most electromagnets use iron or its alloys, allowing efficient generation of fields up to 2 teslas. Air-cored magnets often are used to generate fields above the saturation of ferromagnetic materials. The power requirements of excited magnets can be reduced by using superconducting wire to carry the current. *See* ELECTROMAGNET; FERROMAGNETISM; SUPERCONDUCTIVITY.

Theory and design. Flux density B (in teslas), field strength H (in amperes/meter), and magnetization M are related by Eq. (1), where μ_0 is the permeability

$$\vec{B} = \mu_0(\vec{H} + \vec{M}) \tag{1}$$

of free space ($4\pi \times 10^{-7}$ henry/meter). In ferromagnetic materials, M is a nonlinear and hysteretic function of B.

Field strength H is related to current I by Ampère's law, Eq. (2). Here the integral of the scalar product is

$$\int \vec{H} \cdot \vec{ds} = I \tag{2}$$

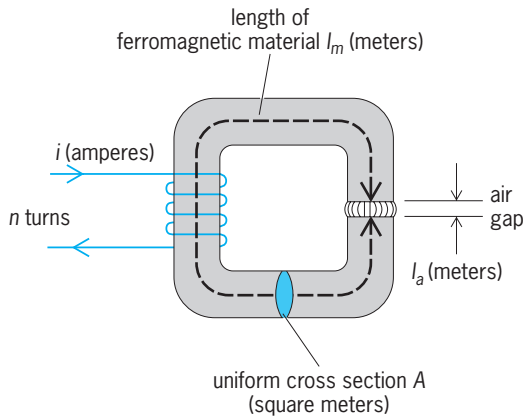

Fig. 1. Simple electromagnetic circuit.

taken around any path s encompassing the currents whose sum is I.

The relation between B and H for a ferromagnetic material is a property of the material expressed graphically and known variously as a magnetization curve, demagnetization curve, or hysteresis loop.

The procedure for calculating B in electromagnets is illustrated by an electromagnetic circuit containing an air gap (**Fig. 1**). A simplified analysis reveals the following features:

1. The relation between the field strength in the ferromagnetic material H_m and the flux density in the air gap B is defined simultaneously by the properties of the material and by Eq. (3), where l_m is the length

$$H_m = \left(ni - \frac{Bl_a}{\mu_0}\right) l_m \tag{3}$$

of the ferromagnetic material (in meters), l_a is the distance across the air gap (in meters), i is the current through the coil (in amperes), and n is the number of turns in the coil.

2. The energy density in the air gap is given by Eq. (4). This is also the force of attraction per unit area in pascals acting on the magnet pole faces.

$$W = \frac{B^2}{2\mu_0} \qquad \text{joules/m}^3 \tag{4}$$

If the ferromagnetic material is magnetically hard and i is zero, then the electromagnet becomes a permanent magnet and Ampère's law shows that H_m must be negative. It follows that the volume of ferromagnetic material equals $-2E/BH_m$, where E is the air gap energy. Hence for economic reasons, permanent-magnet designs should maximize the product $(-BH_m)$. Moreover, the maximum attainable value of this product, $(-BH_m)_{\max}$, is an important property of a material.

In air-cored magnets, flux density may be computed by using the Biot-Savart law or one of its many derivatives.

For discussion of the theory underlying the functioning and design of magnets *see* AMPÈRE'S LAW; BIOT-SAVART LAW; MAGNETISM; MAGNETIZATION.

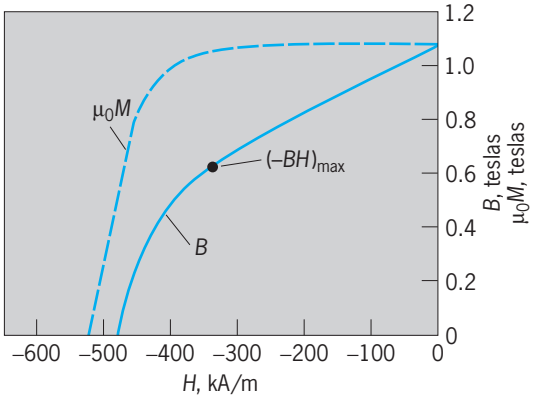

Fig. 2. The variation of *B* (the demagnetization curve) and $\mu_0 M$ as *H* is increased in the negative direction (opposite to *B* and *M*) for the permanent-magnet material $Sm_2(Co, Cu,Fe,Zr)_{17}$. B_r is the remanence, H_c the coercive force, and $(-BH)_{max}$ the optimum operating point. (*After Z. A. Abdelnour, H. F. Mildrum, and K. J. Strnat, Properties of various sintered rare earth–cobalt permanent magnets between −60° and +200°C, IEEE Trans. Magnet., MAG-16:944–996, 1980*)

Materials for magnets. The essential characteristic of permanent-magnet materials is an inherent resistance to change in magnetization over a wide range of field strength H (**Fig. 2**). Resistance to change in M in this type of material and in a great many of its predecessors is due to two factors: (1) the material consists of particles smaller than the size of a domain, a circumstance which prevents the gradual change in M which would otherwise take place through the movement of domain wall boundaries; and (2) the particles exhibit a marked magnetocrystalline anisotropy. During manufacture the particles are aligned in a magnetic field before being sintered or bonded in a soft metal or polyester resin. Compounds of neodymium, iron, and boron have the highest known values of $(-BH_m)_{max}$. *See* IRON ALLOYS.

Electromagnets rely on magnetically soft or permeable materials which are well annealed and homogeneous so as to allow easy motion of domain wall boundaries. Ideally the coercive force H_c should be zero, permeability should be high, and the flux density saturation level should be high. Coincidentally the hysteresis energy loss represented by the area of the hysteresis curve is small. This property and high electrical resistance (for the reduction of eddy currents) are required where the magnetic field is to vary rapidly. This is accomplished by laminating the core and using iron alloyed with a few percent silicon that increases the resistivity.

Electromagnets. Electromagnets usually have an energizing winding made of copper and a permeable iron core. Applications include relays, motors, generators, magnetic clutches, switches, scanning magnets for electron beams (for example, in television receivers), lifting magnets for handling scrap, and magnetic recording heads. *See* CATHODE-RAY TUBE; CLUTCH; ELECTRIC SWITCH; RELAY.

Special iron-cored electromagnets designed with highly homogeneous fields are used for special analytical applications in, for example, electron or nuclear magnetic resonance, or as bending magnets for particle accelerators. *See* MAGNETIC RESONANCE; PARTICLE ACCELERATOR.

Normal-conductor air-cored magnets. Air-cored electromagnets are usually employed above the saturation flux density of iron (about 2 T); at lower fields, iron-cored magnets require much less power because the excitation currents needed then are required only to generate a small field to magnetize the iron. The air-cored magnets are usually in the form of a solenoid with an axial hole allowing access to the high field in the center. The conductor, usually copper or a copper alloy, must be cooled to dissipate the heat generated by resistive losses. In addition, the conductor and supporting structure must be sufficiently strong to support the forces generated in the magnet. The effective pressure in the magnet bore is approximately $B^2/2\mu_0$, or 40 megapascals (6000 lbf/in.^{2}) for a flux density of 10 T. *See* SOLENOID (ELECTRICITY).

Bitter design. A design for high-field solenoids developed by F. Bitter, which addresses these problems, employs a stack of mass-produced and interchangeable slit disks (**Fig. 3***a*). The Bitter stack is interleaved with insulation and compressed axially so as to form a compact helical conductor structurally bonded by friction. Hole patterns punched through the conducting plates and insulators are carefully aligned to allow axial flow of the cooling water. A modification of this helical structure (Figs. 3*b* and *c*) employs two

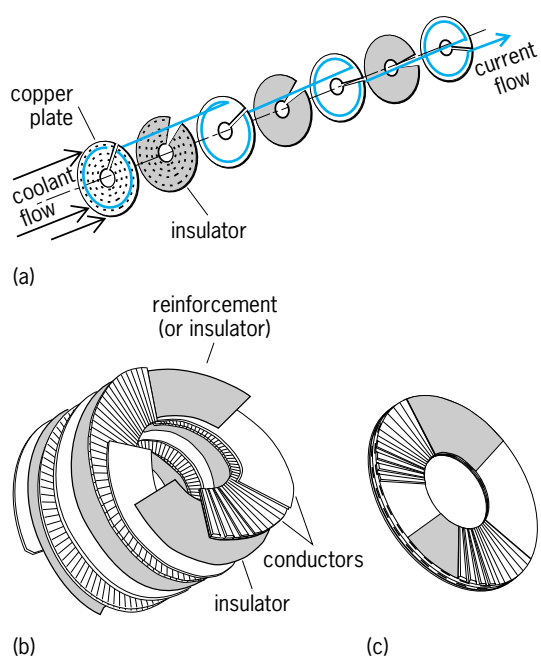

Fig. 3. Bitter magnet stacks. (*a*) Exploded view of helical current path through slotted plates in a stack. (*b*) Expanded view of a stack that has a similar current path shared by the two conductors in a helical plate structure, but has radial grooves in one of them for cooling. (*c*) Assembled view of a few plates. (*After D. B. Montgomery, The Magnetic and Mechanical Aspects of Resistive and Superconducting Systems, Robert E. Krieger Publishing Co., 1980*)

parallel slit copper-plate structures as conductors (one with etched radial cooling channels). A reinforcing plate structure may also be included in either design. Radial cooling is used when higher power dissipation is required. A Bitter magnet produced a sustained field of 30 T and required 30 MW of power. A slightly more efficient but more complex magnet is composed of parallel-connected nested coils with water-cooling passages between them.

Pulsed magnets. Higher fields can be generated by limiting the excitation to short pulses (usually furnished by the energy stored in a capacitor bank) and cooling the magnet between pulses. The pulse length is limited by the heat capacity of the magnet and the resistive losses in the conductor. The highest fields are generally achieved in small volumes. The fundamental limit for an efficient magnet generating pulsed fields repeatedly is set by the conductivity and strength of the conductor. Often additional structures are required to support the large forces generated. A field of 75 T has been generated for 120 microseconds. Special high-strength, high-conductivity conductors for such magnets include a copper-niobium microcomposite and a copper-silver alloy. Pulse durations of 100 ms or more require very large volume magnets and large energies.

The highest fields are generated in magnets which fail during the pulse. Over 200 T is generated for a few microseconds in a single-turn magnet which explodes radially outward. Even higher fields (over 1500 T) are generated by explosive compression of magnetic flux trapped in a single-turn coil.

Superconductor air-cored magnets. Large-volume or high-field magnets are often fabricated with superconducting wire in order to avoid the large resistive power losses of normal conductors. The two commercially available superconducting wire materials are (1) alloys of niobium-titanium, a ductile material which is used for generating fields up to about 9 T; and (2) a brittle alloy of niobium and tin (Nb_3Sn) for fields above 9 T. Practical superconducting wires use complex structures of fine filaments of superconductor that are twisted together and embedded in a copper matrix. Practical superconducting wires can carry current densities of 10^8 A/m^2 with no resistive power loss. However, at sufficiently high fields the superconductor becomes resistive. The conductors are supported against the electromagnetic forces and cooled by liquid helium at 4.2 K (−452°F). A surrounding thermal insulating enclosure such as a dewar minimizes the heat flow from the surroundings.

Superconducting magnets operating over 20 T have been made with niobium-titanium outer sections and niobium-tin inner sections. Niobium-titanium is used in whole-body nuclear magnetic resonance imaging magnets for medical diagnostics. Other applications of superconducting magnets include their use in nuclear magnetic resonance for chemical analysis, particle accelerators, containment of plasma in fusion reactors, magnetic separation, and magnetic levitation. The nuclear magnetic resonance magnets are usually operated as permanent magnets. Once the excitation current reaches the appropriate value, it is switched through a superconducting joint between the magnet leads so that current flows continuously with no resistive loss as long as the conductor is cooled to maintain superconductivity. *See* MAGNETIC LEVITATION; MAGNETIC SEPARATION METHODS; MEDICAL IMAGING; NUCLEAR FUSION; NUCLEAR MAGNETIC RESONANCE (NMR); SUPERCONDUCTING DEVICES.

Hybrid air-cored magnets. The highest continuous fields are generated by hybrid magnets. A large-volume (lower-field) superconducting magnet that has no resistive power losses surrounds a water-cooled inner magnet that operates at the highest field. The fields of the two magnets add. Over 35 T has been generated continuously.

There are active world-wide efforts to increase the high field capabilities of the various air-cooled magnets.

Simon Foner

Bibliography. B. I. Bleaney and B. Bleaney, *Electricity and Magnetism*, 3d ed., 1976, paper 1989; J. E. Evetts (ed.), *Encyclopedia of Magnetic and Superconducting Materials*, 1992; D. C. Jiles, *Introduction to Magnetism and Magnetic Materials*, 2d ed., 1998; D. B. Montgomery, *Solenoid Magnet Design*, 1969, reprint 1980; S. G. Sankar, *Permanent Magnet Materials and Devices*, 2000; E. P. Wohlfarth et al. (eds.), *Ferromagnetic Materials*, vols. 1–5, 1980–1990.

Magnet wire

The insulated copper or aluminum wire used in the coils of all types of electromagnetic machines and devices. It is single-strand wire insulated with enamel, varnish, cotton, glass, asbestos, or combinations of these. To meet the immense variety of uses and to gain competitive advantage, a great number of kinds of enamel and of fiber insulations have been developed and are widely available. And to ensure that the many synthetic materials and combinations are properly applied, many elaborate test procedures have been devised to evaluate heat shock, solubility, blister and abrasion resistance, flexibility, and so on as covered in the National Electrical Manufacturers Association (NEMA) Standards. Most important, the thermal endurance is determined by repeated exposure to high temperature, moisture, and electrical stress, as specified in the American Society for Testing and Materials (ASTM).

The Institute of Electrical and Electronics Engineers (IEEE) thermal classes of insulation, defined by upper temperature limits at which the untreated insulation will have a life expectancy of at least 20,000 h, are O (176°F or 80°C), A (221°F or 105°C), B (266°F or 130°C), F (311°F or 155°C), and H (356°F or 180°C). In general, materials such as cotton, paper, and silk are class O. Organic materials, such as oleoresinous and Formvar enamels, varnish-treated cotton, paper, and silk are class A, and asbestos, mica, silicone varnishes, and polyimide are class H, while

Data on film- and fiber-insulated magnet wire

Type	Shape	Size range, mils*	NEMA spec.	Thermal class	Applications	Qualities
Formvar coated	Round Rectangular	1–204 2100–65,000†	MW15 MW18	A 105°C	Universal use for 105°C limit	Tough, high dielectric strength, chemical stability, but crazes in polar solvents
Alkanex coated	Round Rectangular	1–204 2100–65,000†	MW5	F 155°C	Motors, transformers, control coils for 130° and 155°C limits	Same as Formex, but higher thermal endurance
Polyurethane coated	Round	1–102	MW2	A 105°C	Automatic coil winding	Solderable, low dielectric losses, chemical stability, but low abrasion resistance
Nylon coated	Round	1–102	MW6	A 105°C	Automatic winding with varnish treatment	Solvent resistant, solderable, but high electrical losses and moisture-absorbent
Fluoropolymer resin coated	Round	2–57	MW10	H 180°C	High-temperature and corrosive ambients	Thermal and chemical stability, but low abrasion resistance, and cold flow
Butvar coated	Round	2–64	MW19	A 105°C	Self-supporting coils	Self-bonding, no varnish required for physical support
Imidex E or Isomid coated	Round	2–64	—	180–200°C	High-temperature ambients	No extractibles and highly Freon-resistant
Glass fiber, varnish-treated	Round Rectangular	36–258 2400–93,000†	MW41 MW42	B 130°C	Traction motors and high-temperature windings	Thermal endurance, good dielectric, positive spacing
Dacron and glass fiber, varnish-treated	Round Rectangular	36–258 2400–93,000†	MW45 MW46	B, F, or H, depending on varnish	Machine windings and control coils	Good flexibility and temperature endurance, tough, smooth surface
Asbestos with varnish impregnation	Round Rectangular	18–258 2400–166,000†	MIL-W-583A	B, F, or H, depending on varnish	Large machine windings and traction motors	Good resistance to forming stresses and temperature endurance

*1 mil = 25 μm.
†In mils^2: 1 mil^2 = 645 μm^2.

various synthetic enamels fall in the B and F classes. The polyesterimide enamels, however, are capable of withstanding temperatures of 356–392°F (180–200°C).

The **table** provides a summary of the most widely used types of film- and fiber-insulated magnet wire.

Almost all magnet wire is insulated soft-drawn electrolytic copper, but aluminum is being used more, especially in times when copper is scarce or high-priced and where space is not limiting. Round aluminum wire, being soft, flattens under pressure, giving a higher space factor in coils than might be expected. At temperatures above 392°F (200°C) copper oxidizes rapidly; it also becomes brittle when under stress at such temperatures. Its high-temperature strength is increased considerably by adding a small amount of silver (about 30 oz/ton or 1 kg/metric ton). Anodized aluminum is preferred for higher temperatures, up to 572°F (300°C) or more, while copper nickel wire with an anodized aluminum coating may be operated at still higher temperatures.

The increase in wire diameter due to a single coat of enamel varies from about 1 mil (25 micrometers) for a wire diameter of 80 mils (2 mm) to about 5 mils (15 μm) for a wire diameter of 200 mils (5 mm). The double coating that is frequently called for is not quite twice as thick. The fiber coatings are about three times as thick as the enamel, and they too can be obtained in double thickness. *See* ELECTRICAL INSULATION; MAGNET. Philip L. Alger; C. J. Herman

Bibliography. American Society for Testing and Materials, *Test Method for Relative Thermal Endurance of Film-Insulated Round Magnet Wire*, Stand. ASTM-D2307-94, 1994; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; National Electrical Manufacturers Association, *Magnet Wire*, Stand. MW 1000-1993.

Magnetic compass

A compass depending for its directive force upon the attraction of the Earth's magnetism for a magnet free to turn in any horizontal direction. A compass is an instrument used for determining horizontal direction.

Principle of operation. The magnetic compass operates on the principle that like magnetic poles repel each other whereas unlike poles attract each other. The Earth has internal magnetism similar to that which would result from a short, powerful bar magnet at the center of the Earth. The lines of force connecting the two poles are vertical at two points on the surface of the Earth, called magnetic poles,

and horizontal along the magnetic equator, a line approximating a great circle nearly midway between the magnetic poles. The horizontal component of the Earth's magnetic field varies from a maximum on or near the magnetic equator to zero at the magnetic poles. *See* GEOMAGNETISM; MAGNET; MAGNETIC FIELD; MAGNETISM.

A simple magnetic compass consists of a magnetized needle mounted so as to be free to align itself with the horizontal component of the Earth's field. A magnetic compass is unreliable or inoperative in the vicinities of the magnetic poles. Other sources of error are described below.

Uses. Compasses are used to determine the direction in which a craft is pointed, the direction of travel on land, the direction or bearing of an object, and for other purposes. Mariners, aviators, explorers, guides, scouts, and surveyors have need for a compass. *See* NAVIGATION; SURVEYING.

Direction. The primary function of a compass is to indicate a reference direction, from which all other directions are derived. South is the usual reference direction for astronomers and surveyors; north is the reference direction for others.

Compasses are generally graduated so that, when the north and south graduations on the compass readout are aligned with the north-seeking and south-seeking ends of the compass needle, respectively, other directions are shown in their correct positions as seen from the center of the compass. Most modern navigational compasses mounted permanently in ships, aircraft, and land vehicles are graduated in degrees, from 0° at north clockwise through 360°; the graduations are found on a lightweight compass card attached to the directive element. Pocket or wrist compasses are usually graduated in compass points; the graduations in this instance are found in the case.

Errors. In addition to such errors as faulty alignment or inaccurate graduation, magnetic compasses are subject to two principal errors if the instruments are to be used to determine geographical north and south: variation and deviation.

Variation. This is the angle between geographic and magnetic meridians. The magnetic needle of a compass tends to align itself with the magnetic meridian, which is the horizontal component of a line of force of the magnetic field of the Earth. Magnetic meridians do not, in general, coincide with the geographic meridians, which converge at the geographical poles. The magnetic meridians converge at the magnetic poles, some distance away. The magnetic meridians, unlike the geographic meridians, are not great circles. This is true because the magnetic field is not uniform; it is distorted considerably by irregularities in the magnetic properties of the materials forming the Earth's crust.

Variation of the compass, called declination by magneticians, depends upon the location of the compass. It is determined by magnetic surveys and is published on charts and elsewhere. It varies somewhat with time, both diurnally (daily) and secularly (long-term).

Deviation. The angle between a magnetic meridian and the axis of a compass needle is termed deviation. It is caused by local magnetic effects, such as magnetism of a ship, the magnetic field around a nearby conductor of direct-current electricity, or induced magnetism in a nearby tool or pocketknife.

This source of error generally can be eliminated in a portable compass by using it in a position free from such disturbing influences. For permanent installations in ships, aircraft, and land vehicles, provision is generally made for setting up equal and opposite forces near the compass to neutralize the effects of magnetic influences at the compass. The process is called compass adjustment or compass compensation. In case the deviation cannot be eliminated completely, a deviation table is made up listing the residual deviation on each of a number of headings of the craft. Each deviation table applies only to the compass installation for which it was compiled, since the deviations are caused by local magnetic effects.

Marine compass. The magnetic compass has been largely replaced by the gyrocompass, which is not subject to variation or deviation, as the primary source of directional information at sea. However, the magnetic compass, which does not require an external source of electric power and does not have a mechanism subject to mechanical failure, is still standard equipment aboard ship (**illus.** *a*), and small

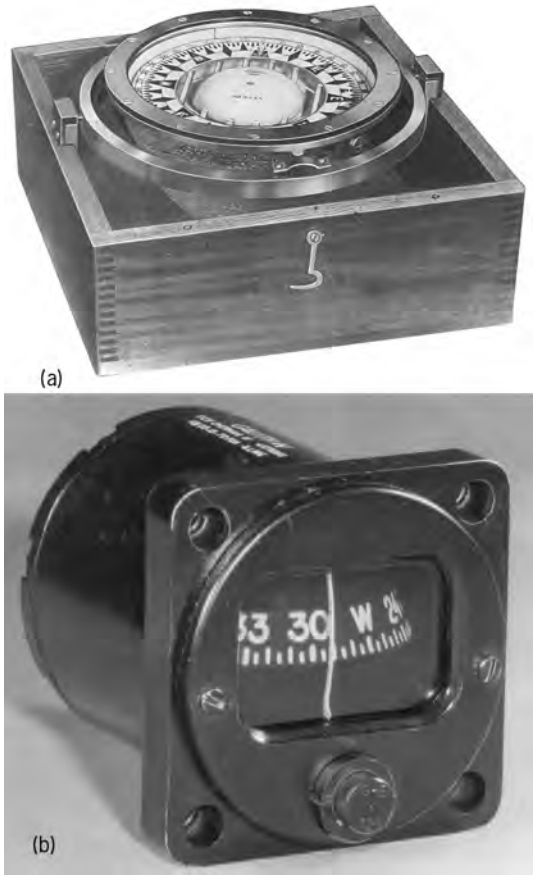

Typical magnetic compasses. (*a*) Marine. (*b*) Aircraft (*Aviation Electric, Ltd.*).

craft almost universally use the less expensive magnetic compass exclusively. The directional element of a marine compass is mounted in a bowl generally filled with liquid (usually a petroleum distillate) that reduces friction on the pivot, thus making the directional force more effective, and damps oscillations of the card. The entire compass is mounted in a stand called a binnacle (or box for small craft) that has provision for magnets and perhaps electric currents to largely neutralize deviation. It may also have a device to facilitate use of the compass for observing bearings of external objects. *See* GYROCOMPASS.

Aircraft compass. Aircraft compasses (illus. *b*) are generally smaller than those used aboard ship and may be mounted in the instrument panel, with graduations on the after (rear) side of the compass card. However, because of the adverse effects of inertial errors related to motions of the aircraft and magnetic disturbances from electric currents and other instruments in the instrument panel, most aircraft compasses are installed in a more favorable location such as a wing or the tail, with remote indicators where needed. A widely used type of remote-indicating aircraft compass is known as a gyro fluxgate compass, or simply fluxgate compass. The entire system consists of a magnetic compass stabilized by means of a gyroscope to reduce the inertial and magnetic errors (which are particularly severe when the aircraft tilts or turns), a transmitter, an amplifier, and one or more remote indicators. The fluxgate itself consists of a fixed soft-iron-core artificial magnet energized by an alternating electric current of suitable strength to saturate the coil with magnetism at each maximum current. The addition of the horizontal component of the Earth's magnetism distorts the core response in such a way that the Earth's horizontal lines of force can be duplicated, making possible the rotation of a moving coil to align itself with the magnetic meridian of the Earth. This alignment can be transmitted to remote indicators throughout the aircraft. *See* AIRCRAFT INSTRUMENTATION; GYROSCOPE.

Alton B Moody

Bibliography. W. Denne, *Magnetic Compass Deviation and Correction*, 3d ed., 1979; J. Klinkert, *Compass-wise or Getting to Know Your Compass*, 1976.

Magnetic ferroelectrics

Materials which display both magnetic order and spontaneous electric polarization. Research on these materials has enabled considerable advances to be made in understanding the interplay between magnetism and ferroelectricity. The existence of both linear and higher-order coupling terms has been confirmed, and their consequences studied. They have given rise, in particular, to a number of magnetically induced polar anomalies and have even provided an example of a ferromagnet whose magnetic moment per unit volume is totally induced by its coupling via linear terms to a spontaneous electric dipole moment.

Most known ferromagnetic materials are metals or alloys. Ferroelectric materials, on the other hand, are nonmetals by definition since they are materials which can maintain a spontaneous electric moment per unit volume (called the polarization) which can be reversed by the application of an external electric field. It therefore comes as no surprise to find that there are no known room-temperature ferromagnetic ferroelectrics. In fact, there are no well-characterized materials which are known to be both strongly ferromagnetic and ferroelectric at any temperature. This is unfortunate since not only would a study of the interactions between ferromagnetism and ferroelectricity be valuable as basic research but such interplay could well give rise to important device applications.

Most antiferromagnetic materials, however, are nonmetals, so there is no apparent reason why ferroelectricity and antiferromagnetism should not coexist. A study of antiferromagnetic ferroelectrics would also provide much information concerning the interplay of magnetic and ferroelectric characteristics even if the device potential were very much reduced. Somewhat unaccountably, antiferromagnetic ferroelectrics are also comparative rarities in nature. Nevertheless, a few are known, and among them the barium–transition-metal fluorides are virtually unique in providing a complete series of isostructural examples. First characterized in the late 1960s, they have the chemical composition $BaXF_4$ in which X is a divalent ion of one of the 3*d* transition metals, manganese, iron, cobalt, or nickel. Nonmagnetic, but still ferroelectric, isomorphs (that is, isostructural equivalents) also exist in the series in which X is magnesium or zinc.

The antiferromagnetic ferroelectrics $BaMnF_4$, $BaFeF_4$, $BaNiF_4$, and $BaCoF_4$ and their nonmagnetic magnesium and zinc counterparts are orthorhombic and all spontaneously polar (that is, pyroelectric) at room temperature. For all except the iron and manganese materials, which have a higher electrical conductivity than the others, the polarization has been reversed by the application of an electric field, so that they are correctly classified as ferroelectric, although their ferroelectric transition (or Curie) temperatures are in general higher than their melting points. The elementary magnetic moments of the manganese, iron, cobalt, and nickel structures are thermally disordered at room temperature, and long-range antiferromagnetic ordering sets in at temperatures somewhat below 100 K (-280°F). Structurally the materials consist of XF_6 octahedra which share corners to form puckered *xy* sheets which are linked in the third dimension *z* by the barium atoms. The magnetic interactions are very dominantly within the *xy* layers, leading to two-dimensional magnetic characteristics, with the magnetic spins eventually aligning themselves antiferromagnetically within the planes. The axis of the spontaneous polarization is also contained within the planes. *See* CRYSTAL STRUCTURE.

The importance of these magnetic ferroelectrics is the opportunity they provide to study and to

separate the effects of a variety of magnetic and nonmagnetic excitations upon the ferroelectric properties and particularly upon the spontaneous polarization. Measurements are often made via the pyroelectric effect p, which is the variation of polarization P with temperature T; that is, $p = dP/dT$. This effect is an extremely sensitive indicator of electronic and ionic charge perturbations in polar materials. Through these perturbations the effects of propagating lattice vibrations (phonons), magnetic excitations (magnons), electronic excitations (excitons), and even subtle structural transitions can all be probed with precision. *See* PYROELECTRICITY.

Phonons. At sufficiently low temperatures the dominant contribution to the pyroelectric effect p comes from thermal excitation of acoustic-phonons, with p varying as T^3. At higher temperatures optic-phonon-mode excitations contribute additional components of exponential form and tend to dominate the acoustic terms. In the magnetic ferroelectrics these contributions must compete with the components arising from the thermal disruption of the ordered antiferromagnetic state. These low-temperature magnetic components take different forms depending on the character of the low-temperature magnetic ordering (in particular, depending on whether the magnetic anisotropy is large, as in the iron and cobalt systems, or small as in the manganese and nickel analogs). The two types of phonon terms are present, of course, in both the magnetic and the nonmagnetic materials, with the additional magnon contributions present only in the magnetic systems. Although the basic contributions are readily detectable in the pyroelectric response, a full separation of all three contributions has not yet been convincingly obtained at low temperatures. *See* LATTICE VIBRATIONS; PHONON.

Magnons. The magnetic contributions to p are seen most easily near the magnetic transitions (that is, the Néel points). Since the magnetic systems undergo very drastic changes, here the perturbation of p is large and easily recognizable. The Néel temperature anomalies take on a cusplike form and arise physically because both the polarization and the magnetic interactions couple strongly to crystal strain, the former via the piezoelectric effect and the latter via magnetostriction. The details of the magnetically produced cusps differ for the different magnetic materials depending on the details of the magnetic anisotropy, but they are now very well understood. Detailed theories have been given, and the quality of agreement between theory and experiment indicates that the magnetic to ferroelectric coupling mechanisms are now known in considerable detail. Since the magnetism is basically of a two-dimensional nature, a significant vestige of the Néel point p-anomalies persists to temperatures well in excess of the transition. This reflects the well-known persistence of short-range spin correlations in two dimensions. *See* MAGNETOSTRICTION; MAGNON; PHASE TRANSITIONS; PIEZOELECTRICITY.

Excitons. The optical absorption spectrum in the magnetic ferroelectrics $BaXF_4$ is dominated throughout the visible region by the d-state-to-d-state localized electronic excitations on the magnetic ions. The approximately cubic crystal field environment splits the magnetic d states into two subgroups which are typically separated in energy by values corresponding to the visible or near-visible range. The excited states on the individual magnetic ions tend to be long-lived at low temperatures, and the excitation can propagate through the crystal via magnetic (exchange) interactions with its magnetic neighbors. These excitations are known as Frenkel excitons, and they couple to the spontaneous polarization via the elementary electric moment perturbation associated with each exciton due to the different charge distributions in the ground and excited d states. The electrical response of the magnetic crystals to short-duration optical pulses clearly shows the exciton contributions to p and their relaxation as the excitations gradually decay back to their ground states. In $BaMnF_4$, for which the most detailed exciton work has been performed, the excited-state polarization is of opposite sign to the spontaneous polarization, enabling a particularly convincing verification of the exciton origin of the effect to be made. *See* CRYSTAL ABSORPTION SPECTRA; EXCHANGE INTERACTION; EXCITON.

Structural transitions. Of all the X ions present in the series $BaXF_4$, the largest is Mn^{2+}. As the temperature is reduced from room temperature, the fluorine cages contract and eventually the divalent manganous ion becomes too big for its cage, precipitating a structural transition at 250 K (−10°F). For the rest of the X ions this overcrowding effect does not occur. The resulting transition in $BaMnF_4$ is quite a complicated one, and produces a cell doubling in the xz plane and a repeat distance in the y direction which is incommensurate with the high-temperature unit cell. The transition involves a complex series of rigid rotations of the fluorine octahedral cages. The subtle atomic rearrangements which take place perturb the local charge environment in such a way that the response is again easily picked up by the pyroelectric effect. A pronounced but rather diffuse peak is seen in p, although the structural transition is known to be sharp from the findings of other probes such as neutron or light scattering, and fluorine nuclear magnetic resonance. A detailed interpretation of this pyroelectric response has not yet been forthcoming.

Magnetoelectric effect. One of the more interesting effects of the incommensurate phase transition at 250 K (−10°F) in $BaMnF_4$ is that it produces a lower-temperature phase with a crystal symmetry low enough to support the existence of the linear magnetoelectric effect. The couplings between magnetism and ferroelectricity discussed above all involve quadratic spin terms. The magnetoelectric effect in contrast is a linear coupling between magnetization and polarization. It is forbidden by symmetry in most magnetic systems, but its presence when allowed always raises the possibility of pyroelectrically inducing a ferromagnetic moment in an antiferromagnetically ordered system. Below the

antiferromagnetic transition at 26 K in $BaMnF_4$ this linear coupling does indeed produce a canting of the antiferromagnetic sublattices through a very small angle (of order 0.2 degree of arc). The result is a spontaneous, polarization-induced magnetic moment. At low temperatures $BaMnF_4$ is therefore technically a weak ferromagnet, although the resultant magnetic moment is extremely small, and it is more usually referred to as a canted antiferromagnet. Pyroelectrically driven ferromagnetism is extremely rare, and this is the only well-categorized example. The magnetoelectric effect also produces additional perturbations of the dielectric susceptibility near the magnetic Néel temperature for the manganese material which are not seen in the other magnetic ferroelectrics of the series for which the linear coupling term is absent for symmetry reasons. *See* ANTIFERROMAGNETISM; FERROELECTRICS; FERROMAGNETISM; MAGNETISM.

Malcolm E. Lines

Bibliography. M. E. Lines and A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials*, 1977; X. Yuhuan, *Ferroelectric Materials and Their Applications*, 1991.

Magnetic instruments

Instruments designed for the measurement of magnetic field strength or magnetic flux density, depending on their principle of operation.

Hall-effect instruments. Often called gaussmeters, these instruments measure magnetic field strength. They have a useful working range from 10 A/m to 8 MA/m (0.125 oersted to 100 kOe). When a magnetic field, H_z, is applied in a direction at right angles to the current flowing in a conductor (or semiconductor), a voltage proportional to H_z is produced across the conductor in a direction mutually perpendicular to the current and the applied magnetic field (**Fig. 1**). This phenomenon is called the Hall effect. The output voltage of the Hall probe is proportional to the Hall coefficient, which is a characteristic of the Hall-element material, and is inversely proportional to the thickness of this material. *See* HALL EFFECT.

For a sensitive Hall probe, the material is thin with a large Hall coefficient. The semiconducting materials indium arsenide and indium antimonide are particularly suitable. Because of the small size of the Hall element (usually of the order of 2 mm square), it is difficult to align the voltage contacts exactly opposite each other. A compensating network comprising a resistive divider (with the addition of a variable capacitor if ac excitation is used) is connected (Fig. 1) to counteract the voltage drop produced by the current flowing in the region between the voltage contacts. The temperature coefficient (about 0.05% per Celsius degree for indium arsenide) can be compensated by the addition of a suitable thermistor. The Hall probe is connected by a flexible cable to a control unit housing the current supply and Hall voltage detection unit. At high magnetic field strengths (above 1.2 MA/m or 15 kOe), the linearity of the Hall-effect device falls off because of the deflection of the electrons within the sensor by the strong magnetic field. By calibrating the instrument, its useful range can be extended. *See* THERMISTOR.

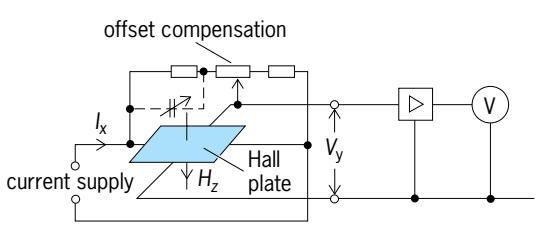

Fig. 1. Principle of the Hall-effect gaussmeter.

Hall-effect gaussmeters are calibrated at low magnetic field strengths by mounting the probe at the center of a long solenoid or pair of Helmholtz coils for which the field-to-current ratio has been determined. At higher magnetic field strengths, the probe is mounted in the gap of a large electromagnet having a uniform magnetic field which is measured by a magnetic resonance magnetometer. *See* HELMHOLTZ COILS; NUCLEAR MAGNETIC RESONANCE (NMR); SOLENOID (ELECTRICITY).

Fluxgate magnetometer. This instrument is used to measure low magnetic field strengths. It is usually calibrated as a gaussmeter with a useful range of 0.2 millitesla to 0.1 nanotesla (2 gauss to 1 microgauss).

Fluxmeter. This instrument is designed to measure magnetic flux. A fluxmeter is a form of galvanometer in which the torsional control is very small and heavy damping is produced by currents induced in the coil by its motion. This enables a fluxmeter to accurately integrate an emf produced in a search coil when the latter is withdrawn from a magnetic field, almost independently of the time taken for the search coil to be moved. A fluxmeter comprises a rectangular coil of wire held between the poles of a permanent magnet. To minimize the torsional control, the suspension is either by an unspun silk fiber or, in the case of more robust instruments, jeweled bearings. Connection between the coil and fluxmeter terminals is made through thin spirals of strip silver, copper, or gold. Assuming the torsional control is negligible and no air damping exists, the fluxmeter receives an impulse and the coil moves through an angle independent of the rate of change of magnetic flux. In practice, there is a small torsional control and some air damping; these cause the instrument to drift slowly toward zero after the maximum deflection has been attained. In the case of Grassof fluxmeters, the resistance of the external circuit should be less than 10 ohms. Some types of light-spot fluxmeters require an external circuit of resistance exactly 30 ohms to maintain the marked calibration. Fluxmeters can be calibrated in a similar way to electronic charge integrators. *See* GALVANOMETER.

Electronic charge integrators. Often termed an integrator or gaussmeter, an electronic charge integrator, in conjunction with a search coil of known effective area, is used for the measurement of magnetic flux density. Integrators have almost exclusively replaced

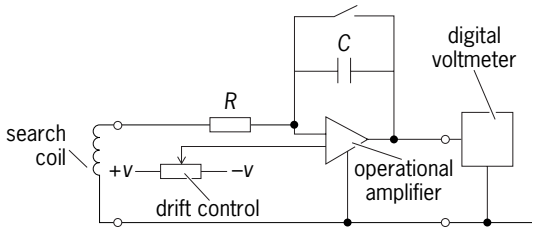

Fig. 2. Arrangement of an electronic charge integrator.

fluxmeters because of their independence of level and vibration. The instrument (**Fig. 2**) consists of a high-open-loop-gain (10^7 or more) operational amplifier with capacitive feedback and resistive input. The value of the input resistor is high (10 kΩ) compared with that of the search coil. The feedback capacitor is chosen to give a convenient time constant and is typically 0.1 or 1 microfarad. The output voltage is proportional to the product of the effective area of the search coil and the change in the magnetic flux density. In practice, consideration must be given to the insulation resistance of the feedback capacitor and any parallel circuits, to thermal and stray pick-up voltages arising in the input circuit, and to changes in the input offset current of the operational amplifier. *See* OPERATIONAL AMPLIFIER.

Electronic charge integrators may be calibrated in one of several ways, as follows:

1. The input can be connected to the secondary of a mutual inductor of known value, and a measured change made in the primary current. The integrator will be direct-reading if the product of the mutual inductance in henrys and the primary current in amperes is equal to the product of the change in the magnetic flux density in teslas and the effective area of the search coil in square meters.

2. A capacitor of known value can be charged up to a known voltage and discharged into the integrator.

3. A linear ramp voltage generator can be connected to the input of the integrator for a measured period of time, t.

If the search coil is wound onto a specimen of ferromagnetic material, it will measure the magnetic flux induced in that material by a known magnetic field strength. Using a Hall-effect instrument in conjunction with an electronic integrator and search coil enables the magnetization curve and hysteresis loop of the magnetic material to be measured. *See* MAGNETIZATION.

Rotating-coil gaussmeter. This instrument measures low magnetic field strengths and flux densities. It comprises a coil mounted on a nonmagnetic shaft remote from a motor mounted at the other end (**Fig. 3**). The motor causes the coil to rotate at a constant speed δ and, in the presence of a magnetic field or magnetic flux density, a voltage is induced in the search coil. The magnitude of the voltage is proportional to the effective area of the search coil and the speed of rotation.

The output of the search coil is taken via slip rings on the driving shaft, rectified by using a synchronous cam-driven switch and measured by using a dc voltmeter. The linearity is exceptionally good over the whole working range from below the Earth's magnetic field strength up to very high magnetic field strengths, limited only by the design of the search coil. Calibration of the instrument is achieved by inserting the rotating coil in a uniform field, for example in the gap of a large electromagnet, the value of which is determined by using a magnetic resonance magnetometer.

Alternating magnetic field meters. These instruments essentially comprise a search coil connected to a root-mean-square (rms) sensing voltmeter and frequency meter. The search coil may be a single turn of calculable area, or a coil of many turns in one or more layers. Alternating magnetic field meters are used extensively in measuring magnetic field strengths from power frequencies up to very high radio frequencies (in the megahertz region), in particular where the effects of stray electromagnetic radiation are of concern. They can be calibrated as an entity over a wide frequency range by mounting the search coil in a specially constructed high-frequency Helmholtz coil system for which the field-to-current and the frequency characteristics have been determined. For an alternating magnetic field of sine-wave form, the magnetic field strength, H (A/m), is given by

$$H = \frac{2\sqrt{2}V}{\pi \mu_o f(AN)}$$

where V is the rms value of the induced voltage in the search coil (V), μ_o is the magnetic constant ($= 4\pi \times 10^{-7}$ H/m), f is the frequency of the alternating magnetic field (Hz), and AN is the effective area-turns product of the search coil (m^2). The measurement of complex waveforms can be achieved by the use of a search coil connected to an oscilloscope with the facility for fast Fourier transform analysis. *See* HELMHOLTZ COILS.

Permeability indicators. These are instruments used for the measurement of the relative magnetic permeability, μ_r, of low-magnetic-permeability materials, for instance stainless steel, in order to assess the effect of the material on the ambient magnetic field. The range of permeability covered is typically 1.001–2. The simplest form of permeability indicator

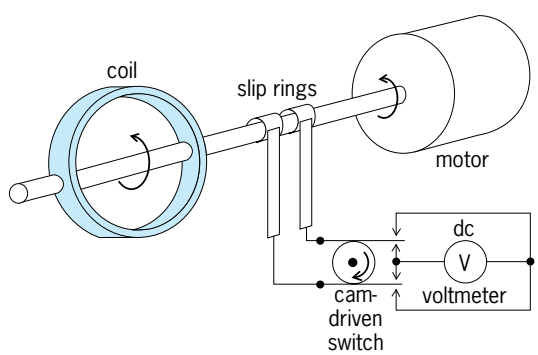

Fig. 3. Arrangement of a rotating-coil gaussmeter. (*M. Lush*)

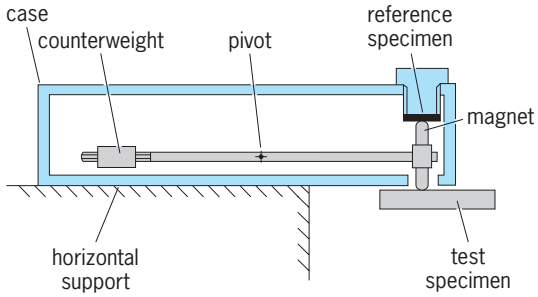

Fig. 4. A typical magnetic balance, side view.

is the magnetic balance (**Fig. 4**). The balance comprises a rod pivoted at its center with a hemispherically ended cylindrical permanent magnet at one end and a counterweight at the other. A number of reference specimens are supplied in the form of discs which can be screwed into the balance case above the upper end of the permanent magnet. Starting with the reference disc of lowest relative magnetic permeability in place, a test specimen is brought into contact with the lower end of the permanent magnet. If the magnet is attracted to the test specimen, then higher values of reference disc are sequentially inserted into the balance until the magnet is attracted to the reference disc. The relative magnetic permeability of the test specimen is then in the range between the values of this disc and the previous reference disc of lower value.

Another type of instrument is based on a fluxgate magnetometer. The permeability-measuring probe includes a permanent magnet, one end of which is placed in contact with the test specimen (**Fig. 5**). The change in the magnetic field around this magnet due to the presence of the test specimen is detected by a pair of fluxgate magnetometer elements. The magnetometer is scaled in terms of $(\mu_r - 1)$, and its advantage over the magnetic-balance type of instrument is that it has a continuous scale.

Both instruments can be calibrated using specially prepared certified reference materials available from some National Standards Laboratories. *See* MAGNETISM; MAGNETOMETER. A. E. Drake

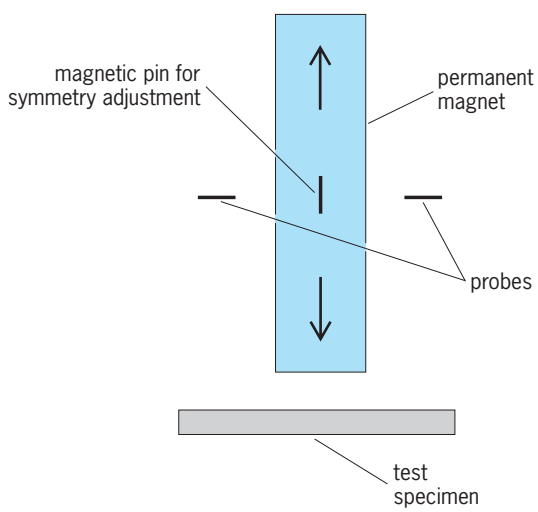

Fig. 5. Arrangement of the permeability meter probe.

Bibliography. E. W. Golding and F. C. Widdis, *Electrical Measurements and Measuring Instruments*, 1963; M. J. Hall et al., Low magnetic permeability reference standards with improved high magnetic field strength performance, *IEE Proc. Sci. Meas. Technol.*, vol. 145, no. 4, July 1998; F. K. Harris, *Electrical Measurements*, 1952, reprint 1975; M. Lush, *Rotating Coil Gaussmeters*, 1964; D. A. Rollett and A. E. Drake, Traceable magnetic field strength measurements to 120 kHz, *IEE Proc. Sci. Meas. Technol.*, vol. 143, no. 4, July 1996.

Magnetic lens

A magnetic field with axial symmetry capable of converging beams of charged particles of uniform velocity and of forming images of objects placed in the path of such beams. Magnetic lenses are employed as condensers, objectives, and projection lenses in magnetic electron microscopes, as final focusing lenses in the electron guns of cathode-ray tubes, and for the selection of groups of charged particles of specific velocity in velocity spectrographs.

As shown in the **illustration**, magnetic lenses may be formed by solenoids or helical coils of wire traversed by electric current, by axially symmetric pole pieces excited by a coil encased in a high-permeability material such as soft iron, or by similar pole pieces excited by permanent magnets. In the last two instances the armatures and pole pieces serve to concentrate the magnetic field in a narrow region about the axis.

Magnetic lenses are always converging lenses. Their action differs from that of electrostatic lenses and glass lenses in that they produce a rotation of the image in addition to the focusing action. For the simple uniform magnetic field within a long solenoid the image rotation is exactly 180°. Thus a uniform magnetic field forms an erect real image of an object on its axis. This image has unity magnification and is at a distance from the object equal to $(8\pi^2 m\phi/eB^2)^{1/2} = 21.08\phi^{1/2}/B$, where m is the mass of the particles, e their charge, ϕ the accelerating potential of the particles, and B the magnetic induction on an axis of symmetry of the field. The numerical coefficient 21.08 applies for electrons, with ϕ in volts and B in gauss.

For short magnetic lenses, or lens fields which are short compared to the focal length, both the magnification and image position depend on the position of the object. The focal length f is given by Eq. (1). The

$$\frac{1}{f} = \frac{e}{8m\phi}\int B^2\,dz = \frac{0.022}{\phi}\int B^2 \cdot dz \quad \text{cm}^{-1} \qquad (1)$$

integration is carried out over the extent of the lens field along the axis of symmetry (the z axis). The numerical coefficient 0.022 again applies for electrons. At the same time the field produces an image

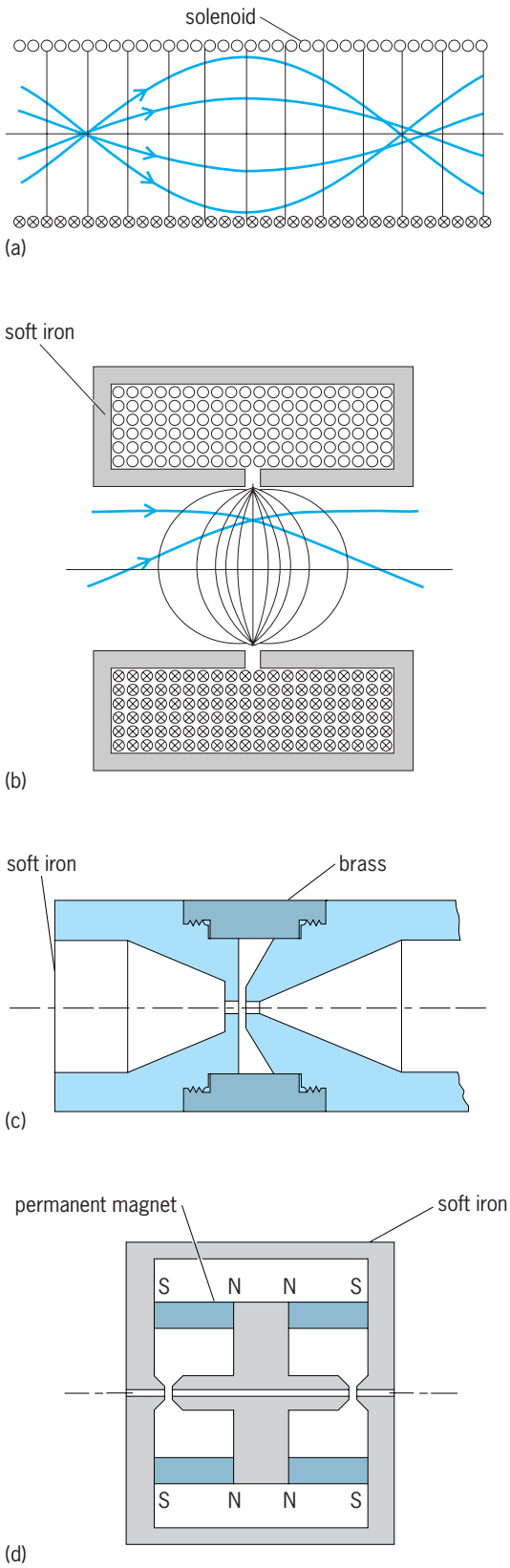

Magnetic lenses. (*a*) Uniform magnetic field. (*b*) Short magnetic lens formed at gap in soft-iron casing about coil. (*c*) Pole pieces for magnetic electron microscope objective. (*d*) Double magnetic lens excited by permanent magnets. (*After E. G. Ramberg and G. A. Morton, J. Appl. Phys., vol. 10, 1939, and J. Hillier and E. G. Ramberg, J. Appl. Phys., vol. 18, 1947*)

rotation through an angle ϕ, as given in Eq. (2). Thus

$$\theta = \left(\frac{e}{8m\phi}\right)^{1/2} \int B\,dz = \frac{0.147}{\phi^{1/2}} \int B\,dz \qquad (2)$$

the magnetic lens field formed by two identical coils in tandem, traversed by oppositely directed currents, is rotation free. As a specific example of a short lens, a single circular loop of wire of radius r traversed by current I produces a lens with focal length $f = 96.8r\phi/I^2$ cm. Here ϕ is in volts and I in amperes. The image rotation for this lens is $\theta = 0.185I/\phi^{1/2}$ radian. *See* ELECTRON MOTION IN VACUUM. E. G. Ramberg

Bibliography. D. A. DeWolf, *Basis of Electron Optics*, 1990; P. W. Hawkes and E. Kaspar (eds.), *Principles of Electron Optics*, 2 vols., 1989; M. Szilagyi, *Electron and Ion Optics*, 1988.

Magnetic levitation

A method of supporting and transporting objects or vehicles which is based on the physical property that the force between two magnetized bodies is inversely proportional to their distance. By using this magnetic force to counterbalance the gravitational pull, a stable and contactless suspension between a magnet (magnetic body) and a fixed guideway (magnetized body) may be obtained. In magnetic levitation (maglev), also known as magnetic suspension, this basic principle is used to suspend (or levitate) vehicles weighing 40 tons or more by generating a controlled magnetic force. By removing friction, these vehicles can travel at speeds higher than wheeled trains, with considerably improved propulsion efficiency (thrust energy/input energy) and reduced noise. In maglev vehicles, chassis-mounted magnets are either suspended underneath a ferromagnetic guideway (track), or levitated above an aluminum track. *See* MAGNET; MAGNETISM.

Attraction-type suspension system. In the attraction-type system, a magnet-guideway geometry, (**Fig. 1***a*) is used to attract a direct-current electromagnet toward the track. This system, also known as the electromagnetic suspension (EMS) system, is suitable for low- and high-speed passenger-carrying vehicles and a wide range of magnetic bearings. The electromagnetic suspension system is inherently non-linear and unstable, requiring an active feedback to maintain an upward lift force equal to the weight of the suspended magnet and its payload (vehicle). The upward life force, $F(i,z,t)$, is given by Eq. (1), where i is the current and z is the

$$F(i,z,t) = \frac{B^2A}{\mu_0} = \frac{\mu_0 N^2 A}{4}\left(\frac{i}{z}\right)^2 \qquad (1)$$

airgap, which both depend on time t, B is the airgap flux density, N is the number of turns in the magnet winding, A is the pole-face area of the magnet, and μ_0 is the permeability of free space. A basic form of state feedback controller which maintains a constant clearance (airgap) between the guideway and the magnet is shown in Fig. 1*b*. The position feedback

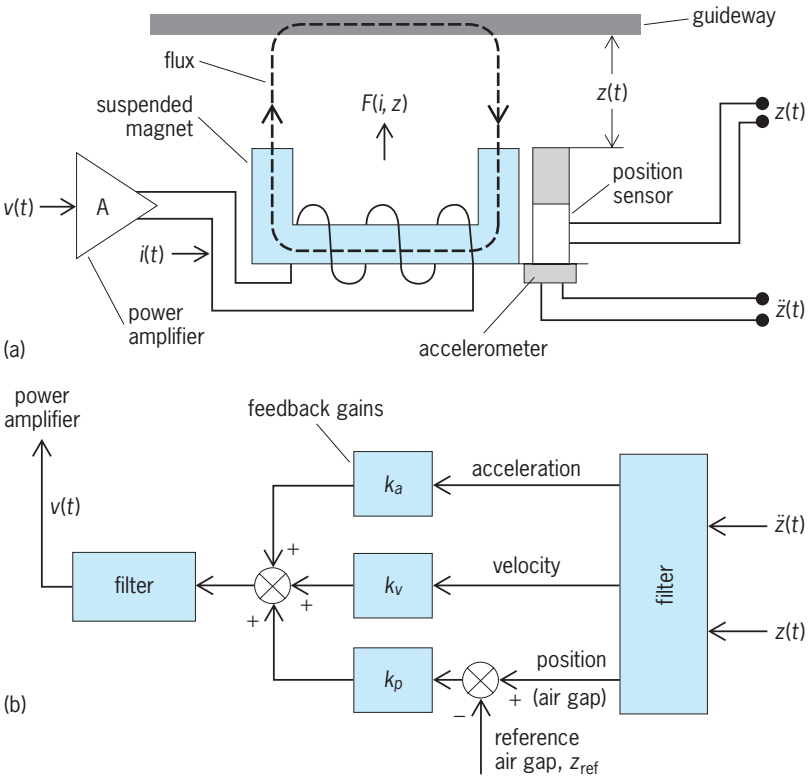

Fig. 1. Attraction-type suspension system. (*a*) Electromagnet-guideway (track) geometry with flux linkages between the magnet and the track. (*b*) Basic state feedback controller. Here, *v*(*t*) is the voltage input to the power amplifier, and *z*(*t*) is the second time derivative (acceleration) of the airgap or clearance, *z*(*t*). Other symbols are explained in the text. (*After P. K. Sinha, Electromagnetic Suspension: Dynamics and Control, Institution of Electrical Engineers, London, 1987*)

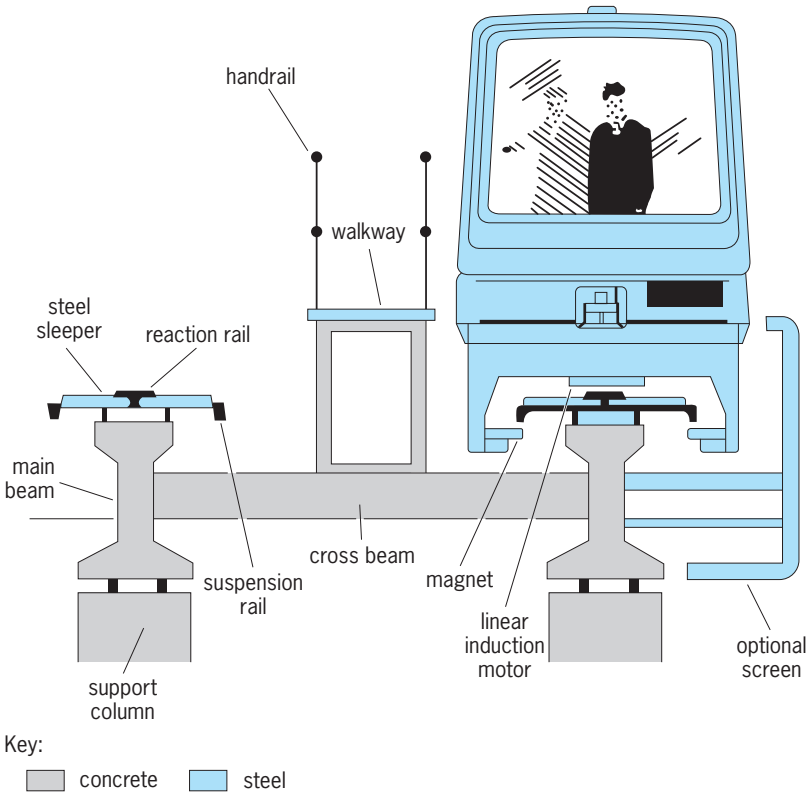

Fig. 2. Sectional view of the Birmingham system, with key components marked. (*After P. K. Sinha, Electromagnetic Suspension: Dynamics and Control, Institution of Electrical Engineers, London, 1987*)

maintains a constant airgap (z_{ref}) and suspension stiffness, the velocity feedback introduces damping, while the acceleration feedback controls damping and ride quality. High-performance microprocessors and digital signal processors are used to implement this and other forms of advanced control laws to provide a very high degree of suspension stability and ride characteristics. For vehicle suspension applications, the airgap is kept in the range of 12–15 mm (0.5–0.6 in.), with airgap flux densities varying from 0.8–1 tesla. For other applications, such as magnetic bearings, the airgaps can be as small as 1 mm (0.04 in.) with considerably lower flux densities. *See* CONTROL SYSTEMS; ELECTROMAGNET; MICROPROCESSOR.

Development of electromagnetic suspension technology for vehicle suspension has been actively pursued since the 1960s, with several experimental and demonstration vehicles running at various locations. The first public system was installed in 1984 at the Birmingham International airport in England to connect the inter-city railway station to the expanded airport. During its operation over several years, this kilometer-long (0.6-mi) system established the electromagnetic suspension technology and its role in public transport. The Birmingham system (**Fig. 2**) used a set of staggered magnets to generate vertical suspension and lateral guidance forces, and linear induction motors for propulsion.

Since the Birmingham system, several full-scale prototype vehicles have been exhibited. Germany leads the technology with the Transrapid series of vehicles (**Fig. 3**) and the construction of a 500-km/h (300-mi/h) maglev link between Hamburg and Berlin. Development of a 100-km/h (60-mi/h) system has been undertaken in Japan for use in urban transit systems.

Repulsion-type levitation system. Most conductors exhibit a reduction in resistance with decreasing temperature; some "lose" their resistance completely below the critical temperature (T_c) to become ideal conductors (superconductors). An electric current set up in a superconducting coil will thus continue to flow and produce a magnetic field as long as the coil remains below T_c (around 20 K or −424°F for copper). This low-temperature condition is achieved by immersing the coil in a coolant (typically, liquid helium) stored in a specially constructed chamber (cryostat). This persistent current makes the superconducting coil behave like a high-field permanent magnet. *See* SUPERCONDUCTIVITY.

In the repulsion-type system, also known as the electrodynamic levitation system (EDS or EDL), a superconducting coil operating in persistent-current mode is moved longitudinally along a conducting surface (an aluminum plate fixed on the ground and acting as the guideway) to induce circulating eddy currents in the aluminum plate. These eddy currents create a magnetic field which, by Lenz's law, oppose the magnetic field generated by the travelling coil. This interaction produces a repulsion force on the moving coil. At lower speeds, this vertical force is not sufficient to lift the coil (and its payload), so

Fig. 3. Magnetic levitation systems and vehicles. (*a*) German Transrapid TR07 vehicle (*Thyssen Henschel, Munich*). (*b*) Japanese HSST system. (*c*) Japanese superconducting test vehicle MLU-002 (*HSST Corporation, Tokyo*).

supporting auxiliary wheels are needed until the net repulsion force is positive. The speed at which the net upward lift force is positive (critical speed) is dependent on the magnetic field in the airgap and payload, and is typically around 80 km/h (50 mi/h). To produce high flux from the travelling coils, hard superconductors (type II) with relatively high values of the critical field (the magnetic field strength of the coil at 0 K) are used to yield airgap flux densities of over 4 tesla. With this choice, the strong eddy-current induced magnetic field is rejected by the superconducting field, giving a self-stabilizing levitation force at high speeds (though additional control circuitry is required for adequate damping and ride quality). For mathematical analysis, the interaction between the two fields is expressed as the repulsive force between the moving superconducting coil above the aluminum track and an imaginary image coil of opposite polarity located at an equal distance below the track. *See* EDDY CURRENT.

Only direct-current superconducting magnets are suitable for this system because of the relatively high hysteresis loss in hard superconductors with alternating-current excitation. In practice, there is some residual resistance in the direct-current superconducting loop, which, although very small, eventually reduces the amplitude of the persistent current. In addition, the coolant in the cryostat needs replacing to keep the coil temperature below T_c. Thus some charging-up current as well as coolant topping up are required from time to time to sustain the levitation force on the moving magnet. Also, there is a certain amount of ohmic loss in the aluminum track, which generates heat and a magnetic drag force as the magnet travels along. In the simplest form, the lift and the drag forces due to a single-turn coil, at any values of the longitudinal velocity v and the airgap z [100 mm (4 in.) or more], are given by Eqs. (2), with $w = 2/(\mu_0 \sigma h)$, where h is

$$F_{\text{drag}}(v,z,t) = \frac{w}{v} F_{\text{lift}}(v,z,t) \qquad (2a)$$

$$F_{\text{lift}}(v,z,t) = \left[\frac{v^2}{v^2 + w^2}\right] F_{\text{coil}}(I,z,t) \qquad (2b)$$

the thickness of the aluminum plate and σ is the conductivity of the superconducting coil material. Here, $F_{\text{coil}} = \mu_0 I^2/(4\pi z)$ is the repulsive force between the levitated coil with a persistent current of I amperes and its image at a distance of $2z$.

The first fully operational supeconducting maglev vehicle was demonstrated by the Japanese National Railways in 1972. This was followed by systems from Germany and the United States. Speeds of up to 517 km/h (321 mi/h) were recorded on the Myazaki test track, in Japan in 1980, and in 1998 a three-car 80-m-long (260-ft) long test vehicle operating on the Yamanashi test track near Tokyo achieved a speed of 550 km/h (342 mi/h).

Linear motors. Due to their contactless operation, linear motors are used to propel maglev vehicles: linear induction motors for low-speed vehicles and linear synchronous motors for high-speed systems. Although their performance parameters are different, operationally they are the unrolled versions of the conventional rotary motors. For linear synchronous motors, three-phase windings are mounted on the guideway, and propulsion force is generated by the interaction between these long stator windings and vehicle-borne direct-current electromagnets (field coils). In the electrodynamic repulsion system, field coils are made of a separate set of superconducting loops. In high-speed electromagnetic attraction systems, such as the Transrapid vehicles, the lift magnets are constructed to provide

the vertical suspension force and act as the field coils of the linear synchronous motor. *See* INDUCTION MOTOR; SYNCHRONOUS MOTOR.

While high-speed vehicles have a restricted magnet-guideway configuration, several types of low-speed systems with different magnet-guideway geometries have been developed. The Berlin Magnetbahn uses a combination of small wheels and permanent magnets for suspension and a linear synchronous motor for propulsion, while the Tokyo continuous transit magnet (CTM) vehicles are supported by rubber wheels and propelled by the coupling force between the vehicle-mounted permanent magnets and a series of linear magnetic belts.

Magnetic bearings. Suspending the rotating part of a machine in a magnetic field may eliminate the contact friction present in conventional mechanical bearings. Magnetic bearings may be based on either attractive or repulsive forces. While the attraction type offers the advantage of controlling the suspension dynamics (stiffness in particular), it is inherently more complex and is used in high-performance and heavy-duty applications. A repulsion system may be very compact and is better suited where physical size is more important than flexibility in the choice of suspension characteristics. Although well developed, radial magnetic bearings are relatively expensive and complex, and are used in specialized areas such as vibration dampers for large drive shafts for marine propellers. In contrast, the axial versions of magnetic bearings are in common use in heavy-duty applications, such as large pump shafts and industrial drums. *See* ANTIFRICTION BEARING.

Maglev technology has evolved from tabletop demonstrations, through high-speed passenger trains traveling at over 550 km/h (342 mi/h), to contactless bearings to support a 3-m-long (10-ft) pump shafts in hazardous environments. As the demand for precise, quiet, and efficient drives grows, the concepts of magnetic suspension and linear motors are likely to find new applications such as the clean-room wafer carrying mechanism, reusable space rocket launchers, and even household appliances. P. K. Sinha

Bibliography. IEEE Magnetics Society, *Intermag, 2000*, 2000; P. K. Sinha, *Electromagnetic Suspension: Dynamics and Control*, Institution of Electrical Engineers, London, 1987.

Magnetic materials

Materials exhibiting ferromagnetism. The magnetic properties of all materials make them respond in some way to a magnetic field, but most materials are diamagnetic or paramagnetic and show almost no response. The materials that are most important to magnetic technology are ferromagnetic and ferrimagnetic materials. Their response to a field H is to create an internal contribution to the magnetic induction B proportional to H, expressed as $B = \mu H$, where μ the permeability, varies with H for ferromagnetic materials. Ferromagnetic materials are the elements iron, cobalt, nickel, and their alloys, some manganese compounds, and some rare earths. Ferrimagnetic materials are spinels of the general composition MFe_2O_4, and garnets, $M_3Fe_5O_{12}$, where M represents a metal. *See* FERRIMAGNETISM; FERROMAGNETISM; MAGNETISM; MAGNETIZATION.

Ferromagnetic materials are characterized by a Curie temperature, above which thermal agitation destroys the magnetic coupling giving rise to the alignment of the elementary magnets (electron spins) of adjacent atoms in a crystal lattice. Below the Curie temperature, ferromagnetism appears spontaneously in small volumes called domains. In the absence of a magnetic field, the domain arrangement minimizes the external energy, and the bulk material appears unmagnetized. *See* CURIE TEMPERATURE.

Magnetic materials are further classified as soft or hard according to the ease of magnetization. Soft materials are used in devices in which change in the magnetization during operation is desirable, sometimes rapidly, as in ac generators and transformers. Hard materials are used to supply a fixed field either to act alone, as in a magnetic separator, or to interact with others, as in loudspeakers and instruments. Both soft and hard materials are characterized by their magnetic hysteresis curve (see **illus.**). *See* ELECTRIC ROTATING MACHINERY; ELECTRICAL MEASUREMENTS; GENERATOR; INDUCTOR; LOUDSPEAKER; MAGNETIC SEPARATION METHODS; MICROPHONE; TRANSFORMER.

Soft magnetic materials. These materials are characterized by their low loss and high permeability. There are a variety of alloys used with various combinations of magnetic properties, mechanical properties, and cost (**Table 1**). There are seven major groups of commercially important materials: iron and low-carbon steels, iron-silicon alloys, iron-aluminum-silicon alloys, nickel-iron alloys, iron-cobalt alloys, ferrites, and amorphous alloys.

The behavior of soft materials is controlled by the pinning of domain walls at heterogeneities such as grain boundaries and inclusions. Thus the common goal in their production is to minimize such heterogeneities. In addition, eddy-current loss is minimized through alloying additions which increase the electrical resistivity. Initial permeability, important in electronic transformers and inductors, is improved by minimizing all sources of magnetic anisotropy, for example, by using amorphous metallic alloys and by using zero magnetostrictive alloys. A high maximum permeability, which is necessary for motors and power transformers, is increased by the alignment of the anisotropy, for example, through development of crystal texture or magnetically induced anisotropy. *See* EDDY CURRENT; MAGNETOSTRICTION.

The class of alloys used in largest volume is by far iron and 1–3.5% silicon-iron for applications in motors and large transformers. In these applications the cost of the material is often the dominant factor, with losses and excitation power secondary

but still important. Thus, the improvement over the years in these alloys has been in developing lower losses without increases in cost. A major improvement in these alloys occurred around 1940 with the development of a $\{110\}\langle 001\rangle$ crystal texture in 3.2% silicon-iron which greatly reduced losses and increased permeability. Since 1880 the losses have decreased steadily through alloying, texture improvements, decreasing the thickness of the strip, and application of stressed insulating coatings. *See* IRON ALLOYS.

Many special alloys find their use in special devices designed to exploit unusual properties. These include materials with high-saturation induction B_s (the magnetic induction at very large values of H), usually cobalt-iron alloys; and alloys with high permeability μ, most often nickel-iron alloys. These alloys are the mainstay of the telecommunications industry. Some of the devices require high initial permeability μ_0 (the permeability at very low fields), and others may depend on high values of μ_{max}, the maximum permeability.

In power equipment the major consideration is the power lost in the magnetic circuit under operating conditions. The total core loss P_c consists of the hysteresis loss P_b and eddy-current loss P_e. The hysteresis loss represents the energy lost at each cycle in the magnetization process. It is proportional to the area of the hysteresis loop and is directly dependent on the frequency f. The loss due to eddy currents depends on the square of the elecromotive force (emf) developed in the conducting magnetic material, that is, B^2f^2, and inversely on the electrical resistivity ρ of the material. By combining geometric factors into constants, the equation below is obtained.

$$P_{c(B,f)} = P_b + P_e = \text{const} \times f + \text{const} \times \frac{B^2f^2}{\rho}$$

The hysteresis loss may be kept low by using a material with a narrow hysteresis loop (one with low coercive force H_c, the value of H required to reduce B to zero). The low H_c is obtained by choosing an alloy with zero stress sensitivity or by minimizing mechanical strain after the final stress-relief anneal. Eddy-current loss may be reduced by breaking up eddy-current paths, for example, by using laminations rather than solid cores.

Audio-frequency devices require very thin laminations or resin-bonded alloy powder cores. The very high electrical resistivity (1–10^6 ohm-cm) qualifies many ferrites and garnets for very high-frequency applications, for example, in cores, transformers, phase shifters, and isolators. *See* FERRITE.

Hard magnetic materials. Permanent magnets, or hard magnetic materials, strongly resist demagnetization once magnetized (**Table 2**). They are used, for example, in motors, loudspeakers, meters, and holding devices, and have coercivities H_c from several hundred to many thousands of oersteds (10 to over 100 kA/m). The bulk of commercial permanent magnets are of the ceramic type, followed by the

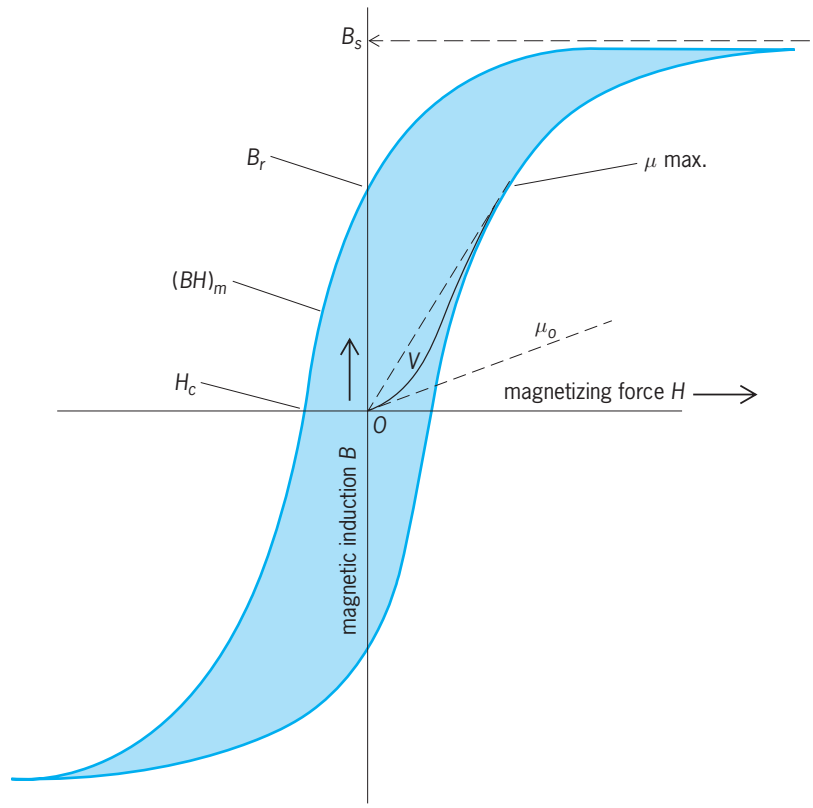

Identification of properties related to the magnetic hysteresis loop. Colored area is the magnetization energy lost in every cycle (hysteresis loss).

Alnicos and the cobalt-samarium, iron-neodymium, iron-chromium-cobalt, and elongated single-domain (ESD) types in decreasing sequence of usage. The overall quality of a permanent magnet is represented by the highest-energy product $(BH)_m$; but depending on the design considerations, high H_c, high residual induction B_r (the magnetic induction when H is reduced to zero), and reversibility of permeability may also be controlling factors. *See* MAGNET.

To understand the relation between the resistance to demagnetization, that is, the coercivity, and the metallurgical microstructure, it is necessary to understand the mechanisms of magnetization reversal. The two major mechanisms are reversal against a shape anisotropy and reversal through nucleation and growth of reverse magnetic domains against crystal anisotropy. The Alnicos, the iron-chromium-cobalt alloys, and the elongated single-domain Lodex alloys are examples of materials of the shape anisotropy structure, while barium ferrites, the cobalt-samarium alloys, and the iron-neodymium-boron alloys are examples of the crystal anisotropy-controlled materials.

Recording materials. Recording tapes and disks for many years were mostly coated with δ–Fe_2O_3, but a variety of new materials have been introduced, for example, metallic particles, chrome oxide, metallic alloy tapes, and cobalt-doped Fe_2O_3. The forming of uniform, and often oriented, dispersions of these materials on flexible (tapes) or rigid (disks)

TABLE 1. Some properties of selected soft magnetic materials

Material or trade name	Composition % by weight, remainder Fe	Relative permeability Initial μ_0	Relative permeability Max. μ_{max}	Coercive force H_c, A/m (oersteds)	Saturation induction B_s, teslas (gauss)	Curie temperature T_c, °F (°C)	Electrical resistivity ρ, microhm-cm	Core loss at 60 Hz, 1.5 teslas (15 kilogauss), W/lb (W/kg)	Sample thickness, in. (mm)
Iron, high-purity	0.05 impurity	10,000	200,000	88 (1.1)	2.15 (21,500)	1420 (770)	10	5.9 (13)	0.025 (0.64)
Iron, commercial-purity	0.2 impurity	250	9,000	72 (0.9)	2.14 (21,400)	1420 (770)	12	3.7 (8.1)	0.019 (0.47)
Carbonyl iron powder	—	60	150	—	—	—	10×10^6		
Armature M-43	0.95Si	—	4,100	75 (0.94)	2.13 (21,300)	1400 (760)	24	2.3 (5.1)	0.019 (0.47)
Electric M-36	2Si	—	7,500	40 (0.5)	2.11 (21,100)	1390 (755)	41	1.9 (4.2)	0.014 (0.36)
Dyname M-22	2.8Si	—	9,400	32 (0.4)	2.03 (20,300)	1365 (740)	49	1.7 (3.7)	0.014 (0.36)
Transformer, M-15	2.2Si	1,500	7,000	28 (0.35)	1.95 (19,500)	1345 (730)	53	1.4 (3.0)	0.014 (0.36)
Oriented cube-on-edge texture, M-4	3.2Si	7,500	55,000	8 (0.1)	2.01 (20,100)	1365 (740)	48	0.55 (1.2)	0.011 (0.27)
High-permeability G.O.	3.2Si	—	—	—	2.01 (20,100)	1365 (740)	45	0.35 (0.77)	0.008 (0.20)
Low-aluminum iron (3.5%)	3.5Al	500	19,000	24 (0.3)	1.90 (19,000)	1380 (750)	47		
16% Al-Fe	16.0Al	3,900	110,000	1.2 (0.015)	0.50 (5,000)	840 (450)	153		
Sendust	5.0Al, 10.0Sl	30,000	120,000	4 (0.05)	1.00 (10,000)	930 (500)	45		
Thermoperm	30Ni	—	—	—	0.203 (2,030)	120 (50)	—		
45 Permalloy	45Ni	2,500	25,000	8 (0.10)	1.60 (16,000)	750 (400)	45		
50–50 Ni-Fe	50Ni	4,000	100,000	4 (0.05)	1.60 (16,000)	930 (500)	45		
Mumetal	77Ni,5Cu,2Cr	20,000	100,000	4 (0.05)	0.65 (6,500)	— 62			
78 Permalloy	78.5Ni	8,000	100,000	4 (0.05)	1.08 (10,800)	1110 (600)	16		
Supermalloy	79Ni,5Mo	100,000	1,000,000	0.32 (0.004)	0.78 (7,800)	750 (400)	60		
2–81 Moly permalloy powder	81 Ni,2Mo	125	130	—	—	—	16×10^6		
27% Co-Fe	27Co, 1Cr	650	10,000	56 (0.70)	2.42 (24,200)	— 28			
50% Co-Fe	49Co,2V, or 2Cr	800	5,000	160 (2.0)	2.45 (24,500)	1800 (980)	7		
Supermendur	49Co,2V	800	70,000	24 (0.3)	2.40 (24,000)	1800 (980)	40		
45–25 Perminvar	45Ni,25Co	400	2,000	95 (1.2)	1.55 (15,500)	1320 (715)	19		
Mn-Zn Ferrite	Mn-δZn-δFe_2O_4	1,500	2,500	16 (0.2)	0.34 (3,400)	265 (130)	20×10^6		
Ni-Zn Ferrite	—	2,500	5,000	8 (0.1)	0.32 (3,200)	285 (140)	10^{11}		
Amorphous Fe-B-Sl, METGLAS 2605S-2	4B,3Sl	15,000	300,000	1.6 (0.02)	1.56 (15,600)	780 (415)	130	0.1 (0.2) at 1.4 T (14 kG)	0.001 (0.025)

TABLE 2. Representative permanent-magnet properties*

Material	Composition % by weight	Curie temperature T_c °F	Curie temperature T_c °C	Coercive force H_c kA/m	Coercive force H_c Oe	Residual induction B_c T	Residual induction B_c G	Max. energy product $(BH)_m$ kJ/m^3	Max. energy product $(BH)_m$ MGOe	Preparation	Mechanical properties
Ba ferrite	$BaO \cdot 6Fe_2O_3$	840	450	170	2,100	0.43	4,300	36	4.5	Press, sinter	Brittle
Sr ferrite	$SrO \cdot 6Fe_2O_3$	860	460	250	3,100	0.42	4,200	36	4.5		
Alnico 5	50Fe,24Co,15Ni, 8Al,3Cu	1650	900	58	620	1.25	12,500	42	5.3	Cast, anneal	Hard, brittle
Alnico 8	34Fe,35Co,15Ni, 7Al,5Ti,4 Cu	1580	860	130	1,600	0.83	8,300	40	5.0		
Alnico 9	34Fe,35Co,15Ni, 7Al,5ti,4Cu			120	1,450	1.05	10,500	68	8.5		
Fe-Cr-Co	63Fe,22Cr,15Co	1165	630	51	640	1.56	15,600	66	8.3	Cast, anneal	Hard
Fe-Cr-Co-Cu	42Fe,33Cr,23Co,2Cu			86	1080	1.30	13,000	78	9.8	Roll, anneal	Hard
Co_5Sm	66Co,34Sm	1290	700	665	8,300	0.91	9,050	160	20	Press, sinter,	Brittle
Co_{17} Sm_2	77Co,23Sm	1470	800	670	8,400	1.08	10,800	223	28	Press, sinter	Brittle
Elongated single domain (ESD) Fe-Co(Lodex)	9.9Fe,5.5Co,77Pb, 8.6Sn	1795	980	70	870	0.8	8,000	25	3.2	Electroplate, distill, press	Soft
Mn-Al-C	70Mn,29Al0.5Ni,0.5C	570	300	220	2,700	0.61	6,100	56	7.0	Cast, extrude, anneal	
Co-Pt	77Pt,23Co	895	480	360	4,500	0.65	6,500	73	9.2	Cast, anneal	Hard, strong
Fe-Nd-B	66Fe,33Nd,1B	570	300	905	11,300	1.21	12,100	280	35	Press, sinter, anneal	Brittle

*Oe = oersteds, T = teslas, G = gauss, MGOe = megagauss-oersteds.

substrates is a challenging manufacturing problem. *See* ALLOY; AMORPHOUS SOLID; MAGNETIC RECORDING.

F. E. Luborsky

Bibliography. American Society for Metals, *Metals Handbook, Desk Edition*, 1985; K. M. Buschow et al. (eds.), *Ferromagnetic Materials*, vols. 1–6, 1980–1991; P. Campbell, *Permanent Magnetic Materials and Their Application*, 1994; C. W. Chen, *Magnetism and Metallurgy of Soft Magnetic Materials*, 1977, reprint 1986; J. Crangle, *Solid State Magnetism*, 1991; J. E. Evetts (ed.), *Concise Encyclopedia of Magnetic and Superconducting Materials*, 1992; J. P. Jakubovics, *Magnetism and Magnetic Materials*, 2d ed., 1994; D. C. Jiles, *Introduction to Magnetism and Magnetic Materials*, 1991; C. Kittel, *Introduction to Solid State Physics*, 6th ed., 1986; P. Robert, *Electrical and Magnetic Properties of Materials*, 1988; S. G. Sankar, *Permanent Magnet Materials and Devices*, 1995.

Magnetic monopoles

Magnetically charged particles. Such particles are predicted by various physical theories, but so far all experimental searches have failed to demonstrate their existence.

Theoretical basis. The fundamental laws governing electricity and magnetism become symmetric if particles exist that carry magnetic charge. Current understanding of electromagnetic physical phenomena is based on the existence of electric monopoles, which are sources or sinks of electric field lines (**Fig. 1***a*), and which when set into motion generate magnetic fields. The magnetic field lines produced by such a current have no beginning or end and form closed loops (Fig. 1*a*). All magnetic fields occurring in nature can be explained as arising from currents. However, theories of electromagnetism become symmetric if magnetic charges also exist. These would be sources or sinks of magnetic field and when set into motion would generate electric fields whose lines would be closed without ends (Fig. 1*b*). *See* ELECTRIC FIELD.

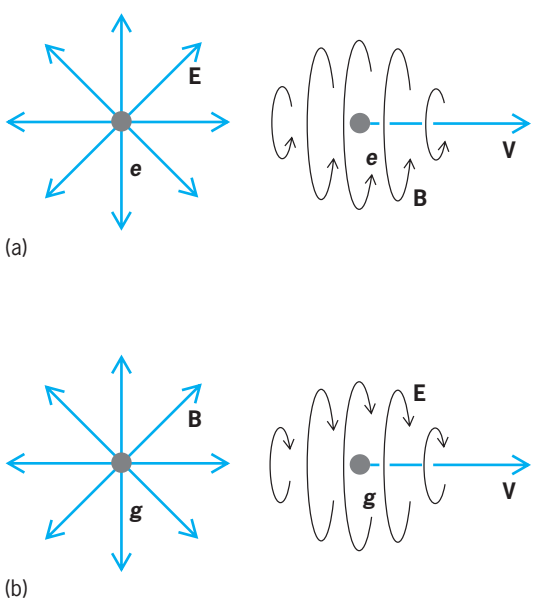

Fig. 1. Electric (*E*) and magnetic (*B*) field lines generated by monopoles and by their motion with velocity *v*. (*a*) Electric monopole with electric charge *e*. (*b*) Magnetic monopole with magnetic charge *g*.

J. C. Maxwell's equations describing electromagnetism can be generalized by adding the magnetic charge density ρ_m and the magnetic current density $\mathbf{j}_m$. The generalized Maxwell equations are Eqs. (1)

$$\begin{aligned} \nabla \times \mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} &= \frac{4\pi}{c}\mathbf{j}_e \\ \nabla \cdot \mathbf{E} &= 4\pi\rho_e \\ \nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} &= -\frac{4\pi}{c}\mathbf{j}_m \\ \nabla \cdot \mathbf{B} &= 4\pi\rho_m \end{aligned} \tag{1}$$

in cgs units, where c is the speed of light. The first equation (Ampère's law) says that electric currents ($\mathbf{j}_e$) and time-varying electric fields ($\mathbf{E}$) produce magnetic fields ($\mathbf{B}$); the second (Gauss's law), that electric charges (with density ρ_e) produce electric fields with strength inversely proportional to the square of the distance from the charge; the third (Faraday's law), that electric fields are produced by time-varying magnetic fields and by magnetic charge currents (if they exist); and the fourth (Gauss's law for magnetic monopoles), that magnetic charges would produce magnetic fields with strength inversely proportional to the square of the distance from the charge. *See* AMPÈRE'S LAW; ELECTROMAGNETIC INDUCTION; GAUSS' THEOREM; MAXWELL'S EQUATIONS.

However, this symmetrization of Maxwell's equations is not a strong reason for suggesting the existence of magnetic charges. New electric and magnetic fields can be defined as linear combinations of the present fields in such a way that Maxwell's equations become symmetric but no new physics is introduced.

Dirac's theory. In 1931 P. A. M. Dirac found a more fundamental reason for hypothesizing magnetic charges, when he showed that this would explain the observed quantization of electric charge. He showed that all electric and magnetic charges e and g must obey Eq. (2), where k must be an integer and is Planck's constant divided by 2π. Equation (2) can be satisfied only if all electric and magnetic charges are integer multiples of an elementary electric charge e_0 and an elementary magnetic charge g_0. Since the size of the elementary electric charge, the charge carried by an electron or proton, is known experimentally, Dirac's equation predicts the size of the elementary magnetic charge to be given by Eq. (3). Since the fine structure constant α is given by Eq. (4), the elementary magnetic charge g_0 is about 68.5 times larger than the elementary electric

$$eg = k(^1/_2\hbar c) \tag{2}$$

$$g_0 = \frac{1}{2}\frac{\hbar c}{e_0} \tag{3}$$

$$\alpha = \frac{e_0^2}{\hbar c} \approx \frac{1}{137} \tag{4}$$

charge e_0. *See* FINE STRUCTURE (SPECTRAL LINES); FUNDAMENTAL CONSTANTS.

Grand unification theories. In 1983 a successful theoretical unification of the electromagnetic and weak forces culminated in the detection of the W^+, the W^-, and the Z^0 particles predicted by the theory. This success has encouraged the search for a grand unification theory that would include the electroweak force and the nuclear or color force under one consistent description. In 1974 G. 't Hooft and independently A. M. Polyakov showed that magnetically charged particles are necessarily present in all true unification theories (those based on simple or semisimple compact groups). These theories predict the same long-range field and thus the same charge g_0 as the Dirac solution; now, however, the near field is also specified, leading to a calculable mass. The SU(5) model predicts a monopole mass of 10^{16} GeV/c^2, while theories based on supersymmetry or Kaluza-Klein models yield even higher masses up to the Planck mass of 10^{19} GeV/c^2. *See* ELECTROWEAK INTERACTION; FUNDAMENTAL INTERACTIONS; GRAND UNIFICATION THEORIES; QUANTUM GRAVITATION; SUPERGRAVITY; SUPERSYMMETRY.

Astrophysical limits. Although grand unification theories clearly predict the existence of monopoles, cosmological models based on these theories lead to impossibly high or unobservably low predictions for monopole fluxes. However, the low predictions are based on the inflationary model, involving an exponential expansion of the early universe followed by a reheating. In this model, the number of monopoles depends exponentially on the reheating energy, and thus small changes in this energy, which cannot be calculated in detail, can result in observable monopole fluxes. *See* INFLATIONARY UNIVERSE COSMOLOGY.

An upper bound (the Parker bound) of 10^{-15} cm^{-2} sr^{-1} s^{-1} is obtained, assuming an isotropic flux, from arguments based on the existence of the 3-microgauss (3×10^{-10} tesla) galactic magnetic field. The Parker bound becomes less severe linearly with monopole mass for masses above 10^{17} GeV/c^2. In addition, models incorporating monopole plasma oscillations would allow a much larger particle flux. All of these bounds assume particle velocities in gravitational virial equilibrium, that is, very near 10^{-3} c.

It has been shown theoretically that the supermassive monopoles arising from many grand unification theories would catalyze nucleon decay processes. Arguments based on x-ray flux limits from galactic neutron stars, which assume a strong-interaction cross section for proton decay catalysis, lead to an upper bound for magnetic particle flux of about 10^{-21} cm^{-2} sr^{-1} s^{-1}, which would make direct detection of monopoles virtually impossible. However, the incomplete understanding of both the catalysis cross section and neutron stars, and the failure to observe proton decay, make this bound highly uncertain. *See* NEUTRON STAR; PROTON.

Detectors. There are two classes of magnetic monopole detectors, superconducting and conventional.

Superconducting detectors. These are based on Faraday's generalized law of induction [the third of Maxwell's equations in (1)], where a magnetic charge passing through a conducting coil will induce a current, in the same way that passing one end of a solenoid through a conducting coil induces a current. If the coil is superconducting, the current never dies away so that the passage of a magnetic charge would show itself as a step in the persistent current of the coil. This type of detector directly measures the magnetic charge of a particle, independent of particle velocity and other properties. Because of their velocity-independent response, these detectors are a natural choice in searches for supermassive (and therefore slow) magnetically charged particles. *See* SUPERCONDUCTING DEVICES; SUPERCONDUCTIVITY.

In addition, the detector response is based on simple, fundamental, and extremely convincing theoretical arguments.

The remarkable theoretical similarities between flux quantization in superconductors and Dirac magnetic monopoles make superconducting systems natural detectors for these particles. The magnetic flux emanating from a Dirac charge g_0 is exactly twice the flux quantum of superconductivity. Since no magnetic field line can pass through a superconductor, when a magnetic charge passes through a superconducting ring (**Fig. 2**), every flux line emanating from the pole must leave a closed loop around the ring wire as the particle moves through. Thus, the magnetic flux change through the ring must exactly equal

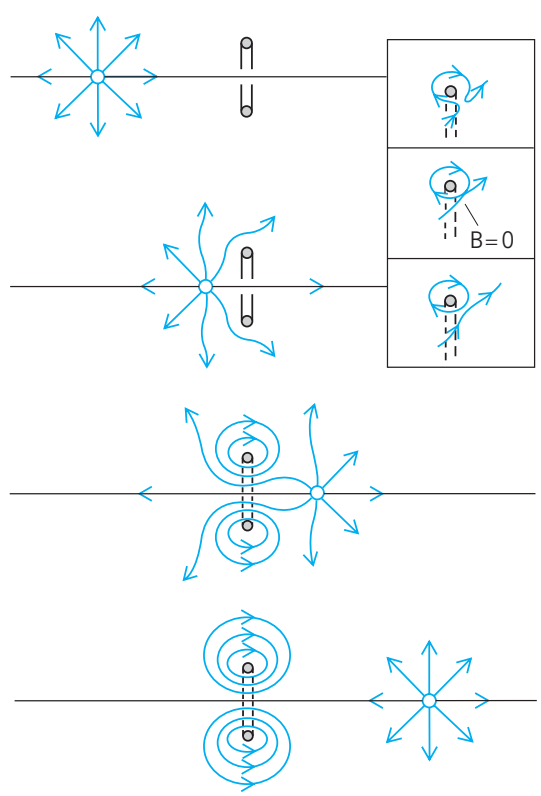

Fig. 2. Successive configurations of magnetic field lines as a monopole passes through a superconducting ring. Inset shows detail of a single magnetic field line leaving a closed loop around the ring; *B* is the magnetic flux density.

the total flux from the pole: two flux quanta for a monopole of unit Dirac charge. No net flux change results form trajectories that miss the ring.

On February 14, 1982, a prototype superconducting detector operating at Stanford University observed a single candidate event. Since then a number of groups have operated larger second- and third-generation detectors, and their combined data have placed a limit on the monopole flux more than 3000 times lower than the value from the data set that included the original event. Thus the possibility that this event was caused by the passage of a magnetic monopole has been largely discounted.

Another single event was observed in a second-generation detector at Imperial College, London, on August 11, 1985. The possibility of the event having been caused by the passage of a magnetic monopole also has been largely discounted because it was a noncoincident signal observed by only one SQUID, and thus more susceptible to a spurious cause. All other second-generation detectors have used fully coincident multiple-loop detection schemes. *See* SQUID.

Several groups have operated larger third-generation detectors. The detector shown in **Fig. 3**, with a sensing area greater than 1.5 m^2 (16 ft^2) times 4π sr, is composed of eight planar superconducting detection coils arranged around a cylinder with an octagonal cross section. Each coil is a gradiometer connected to a high-sensitivity radio-frequency SQUID current sensor.

The gradiometer winding pattern is an important design feature for superconducting monopole detectors. For example, a figure-eight coil couples no net magnetic flux for any change in a uniform applied field. The sensitivity of a gradiometer to external magnetic field changes is substantially reduced over that from a simple coil, whereas the sensitivity to the passage of a magnetic charge remains high since the particle passes through only one element of the gradiometer pattern. A further improvement was achieved by breaking the loop up into a number of separate elements that are connected to one SQUID in parallel, thereby reducing coupling losses to the SQUID.

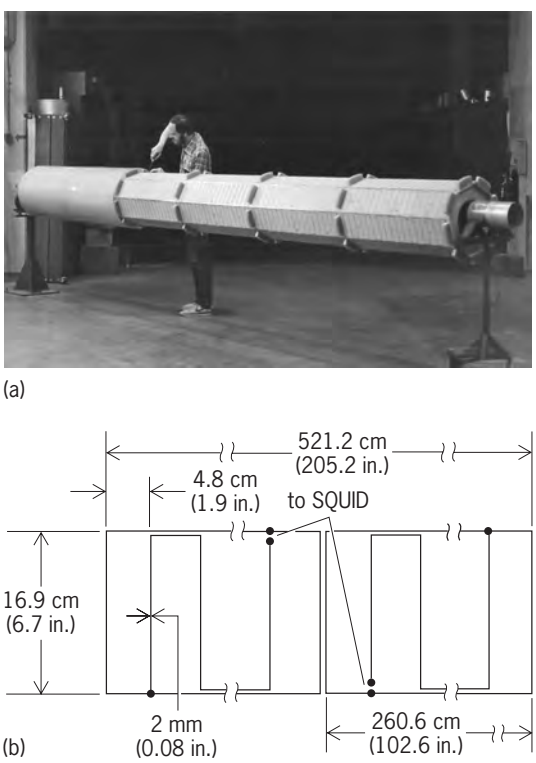

Fig. 3. Eight-SQUID 1.5-m^2 (16-ft^2) octagonal detector in operation at Stanford University. (*a*) Detector being adjusted. (*b*) Schematic diagram of detector.

The monopole flux limit from the combined data of all detectors is below 3×10^{-13} cm^{-2} sr^{-1} s^{-1} (90% confidence limit). Each of the third-generation detectors is capable of surpassing the peak of the mass-dependent Parker bound in 3 years of operation, or convincingly discovering magnetic monopoles in cosmic rays. If magnetic monopoles exist in cosmic rays, they are rare and very large detectors are required to observe them.

Conventional detectors. These measure the weak ionization of matter expected along the trajectory of a magnetic charge. Particle detectors such as scintillators, which collect the fluorescent light from the ionization, and proportional counters, which collect the electrons produced by the ionization, can be used. The signature of a monopole in these conventional detectors would be that of a massive particle moving a thousand times slower than the speed of light. Such a signature would not directly measure a magnetic charge, but the detection of any massive ionizing particle, whether electrically or magnetically charged, would be highly significant. The primary advantage of conventional detectors over superconducting ones is that sensing areas roughly 10 times larger can be instrumented for the same costs. In addition, the conventional detectors are also sensitive to known particles such as cosmic-ray neutrinos, allowing other experiments to be run simultaneously. *See* IONIZATION CHAMBER; PARTICLE DETECTOR; SCINTILLATION COUNTER. Blas Cabrera

Bibliography. A. S. Goldhaber and W. P. Trower (eds.), *Magnetic Monopoles: Selected Reprints*, 1990; P. A. Horvathy, *Introduction to Monopoles*, 1988.

Magnetic reception (biology)

Sensitivity to magnetic stimuli. For more than a century, biologists have speculated whether any living organism can detect magnetic stimuli, especially the very weak ones occurring naturally in the environment (the intensity of the Earth's magnetic field, the geomagnetic field, is roughly 0.5 gauss). A great variety of biological effects resulting from exposure to fields many thousands of times more intense than Earth's have been reported. Among these are changes in plant growth rates, retardation of embryo development, changes in enzyme activity, alterations of tumor growth, and other indications of stress. While the evidence for some of the reported

effects is not very convincing, it does seem likely that, under certain conditions, such intense fields can indeed produce stress effects in living tissues similar to the effects of factors such as extremes of heat, cold, or starvation. This article is primarily concerned with sensory detection, not with stress effects, and focuses primarily on fields of geomagnetic intensities.

Invertebrates. Most of the evidence for magnetic detection comes from experiments performed during the 1960s and 1970s. During the first half of this period most of the experiments were performed by F. A. Brown and colleagues, who reported turning or orientational responses to weak magnetic fields in a variety of invertebrates, including protozoa, flatworms, and snails.

In the late 1960s and early 1970s a number of other laboratories began to find evidence of responses to weak magnetic fields. Attracting special attention were experiments on insects and on birds (and later, on fish and on bacteria). In 1968 Martin Lindauer and Herman Martin first published extensive data showing that the geomagnetic field influences the orientation of the waggle-run dance by which a scout honeybee communicates the distance and direction of a food source to the forager bees. Later, Lindauer and Martin showed that honeybees are so sensitive to magnetic stimuli that fluctuations of less than 10^{-4} gauss (roughly 1/10,000 of the Earth's field) can influence their behavior. Other investigators found evidence of magnetic detection in other kinds of insects, including termites, beetles, and fruit flies (*Drosophila*).

Birds. Most of the evidence for magnetic detection by birds has come from studies of their migratory and homing behavior. Experiments on birds exhibiting oriented migratory restlessness in circular test cages and experiments on the initial orientation of homing pigeons have yielded results strongly suggesting that birds possess a magnetic compass, that is, they can determine compass bearings from the geomagnetic field. Evidence indicates that birds' sensitivity to magnetic stimuli is roughly similar to the honeybees'; they too can probably detect fluctuations of less than 10^{-4} gauss. It appears that the tiny fluctuations in the Earth's magnetic field caused by solar flares and other solar disturbances have a detectable effect on birds' navigation.

Although behavioral effects of magnetic stimuli have been found in many kinds of animals, no one has yet succeeded in conditioning an animal to a magnetic stimulus in the laboratory. There is abundant evidence that the detection process is not quick, usually taking 15 min or more; hence, the flash stimuli presented in most classical conditioning attempts may be undetectable.

It appears that birds do not read the magnetic compass from its horizontal component the way people do. Rather, they probably rely on the angle between gravity and the magnetic total vector (which points north and down in the Northern Hemisphere); for a bird in the Northern Hemisphere, north is apparently that direction in which the gravity and magnetic vectors form the most acute angle. Only the magnetic vector itself, not its polarity, is important. Moreover, the detection system probably has a narrow range of sensitivity; magnetic fields much stronger or weaker than the Earth's probably cannot be detected. *See* MIGRATORY BEHAVIOR.

Mechanisms. The physical mechanism for magnetic detection by living organisms is unknown, though a variety of possibilities have been put forward, such as an induced electromotive force (emf) as a result of motion through the magnetic field; deflection of moving charges by means of the Hall effect; distortions of molecular bond angles; nuclear magnetic resonance effects; direct deflection of ferromagnetic particles; and many others. But the evidence is still so scanty that any choice between them is more theoretical than empirical. What does seem very likely, in view of the sensitivity of detection already demonstrated, which is at a level where thermal noise becomes a major problem, is that the process must involve extensive integration across sites or across time, or both. William T. Keeton

Bibliography. M. F. Barnothy (ed.), *Biological Effects of Magnetic Fields*, 1964, vol. 2, 1969; W. T. Keeton, Magnets interfere with pigeon homing, *Proc. Nat. Acad. Sci.*, 68:102–106, 1971; C. Walcott and R. P. Green, Orientation of homing pigeons altered by a change in the direction of an applied magnetic field, *Science*, 184:180–182, 1974; W. Wiltschko, in *Orientation and Navigation: Birds, Humans, and Other Animals*, Paper No. 12, 1993; W. Wiltschko and R. Wiltschko, Magnetic compass of European robins, *Science*, 176:62–64, 1972.

Magnetic recording

The technique of storing information as a magnetic pattern on a moving magnetic medium. The medium may be a disk, either flexible (floppy) or rigid, or a tape. There are various techniques for "writing" the patterns, including the use of the magnetic field generated by an electric current in a miniature coil, or the use of a laser beam to heat the material. There are several techniques for "reading" the pattern, such as measuring the voltage induced in the same or different coil by the magnetic fields associated with the pattern, measuring the change in the resistance of a metal stripe due to this field, or measuring the polarization of a laser beam upon reflection from the medium.

This combination of formats and techniques has given rise to a variety of products with applications ranging from audio recorders to large optical data files.

Magnetic medium. All materials respond in some way to an applied magnetic field, but the term "magnetic material" generally means one that maintains a magnetic polarization in the absence of an applied field. This remanence depends upon the magnetic field history to which the material was exposed. This history can be plotted on a graph of magnetization

versus magnetic field, giving rise to hysteresis loops (**Fig. 1**). If the material is initially completely demagnetized (A in Fig. 1, sometimes referred to as the ac erase state), and the applied field is increased to some intermediate value and then reversed, a minor loop (BCDEB) is obtained. At zero magnetic field there is a remanence. If the field is increased to the point that further increases in the field result in no further increases in magnetization, the material is said to be saturated. The remanence, at zero magnetic field, is a function of the maximum magnetic field to which the material has been exposed (Fig. 1). Also, it takes a finite field in the reverse direction to drive the magnetization to zero. This field is called the coercivity of the material, and is an important parameter in magnetic recording. *See* MAGNETIC MATERIALS; MAGNETIZATION.

Audio recording. The hysteretic behavior of magnetic materials, and in particular, the field dependence of the remanence (Fig. 1), is the basis for recording sound. The basic idea is to use the electric current from a microphone to generate a magnetic field (Ampère's law) that magnetizes portions of a magnetic medium (originally, segments of a steel wire) in proportion to this current. The resulting magnetic pattern along the medium can then be read back as a voltage induced in a pick-up coil (Faraday's induction law) as the fringing fields from the magnetic medium pass by. *See* AMPÈRE'S LAW; FARADAY'S LAW OF INDUCTION.

The writing field associated with a coil can be enhanced by filling the coil with a magnetic material. The reason is that the magnetic flux density generated by an electric current is proportional to the current through a constant of proportionality called the permeability. In free space, this quantity is usually denoted by μ_0. The permeability of a magnetic material, μ, however, is much larger. For example, the nickel-iron alloy permalloy has a relative permeability, μ/μ_0, of the order of 10,000. Therefore, the magnetic flux density inside a permalloy core would be 10,000 times that of an air core.

A high-permeability core also serves to confine the flux density. The field in a very narrow gap in the magnetic material can therefore be relatively large. Thus, recording generally employs an electromagnet with a narrow gap (**Fig. 2**). The writing and reading element is referred to as the head.

The relationship between the remanence and the field is very nonlinear (Fig. 1), which led to a good deal of distortion in early recorders. However, if an alternating current (ac) is added to the signal current, the resulting remanence, called the anhysteretic remanence, becomes linear at low fields. The amplitude of this bias must be sufficient to produce a field greater than the coercivity, and the bias frequency must be higher than the highest signal frequency. This ac-biasing technique, which makes magnetic recording practical, was discovered in 1927 but was not widely applied to recording until the 1940s. The first magnetic tape, also introduced in 1927, consisted of paper coated with a magnetic powder. *See* SOUND RECORDING.

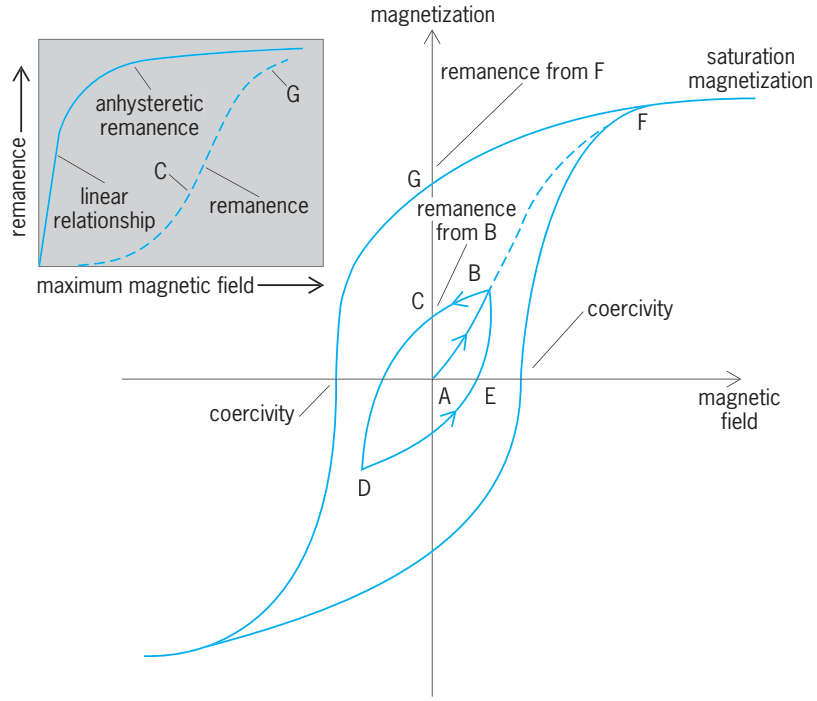

Fig. 1. Typical magnetization curves. At A, material is completely demagnetized. If the magnetic field is increased to B and then reversed, the minor loop BCDE is obtained. The inset shows remanence as a function of maximum field; points C and G correspond to the hysteresis loops shown.

Video and data recording. To extend magnetic recording to video recording with signals as high as 5 MHz requires increasing the speed between the tape and the recording head. Video recording is based on an approach in which a number of heads are mounted on the face of a drum that rotates rapidly in a direction transverse to the direction of the tape motion. The first commercial broadcast product, a video tape recorder (VTR), introduced in 1956, used a rotating drum mounted with four heads; this became known as the quadruplex head. *See* TELEVISION.

In order to reduce flicker, television systems interlace two scans down the screen. Each scan is called a field. To record one field continuously on a tape, the video track is recorded diagonally. This is accomplished by a helical-scan approach, in which the tape is wrapped around the rotating-head drum in a helical path.

In order to adapt this commercial technology for home recording (that is, video cassette recorders or VCRs), it is necessary to provide a longer playing time, and this requires a better utilization of the tape area. This is accomplished by azimuthal recording, in which two heads are used whose gaps are inclined at different angles to the direction of the track. This gives a herringbone recording pattern with very low crosstalk between the two tracks, and eliminates the need for guard bands between them.

Magnetic recording was applied to the storage of data in the early 1950s. Data generally means information represented in a digital form, that is, a sequence of 0's and 1's. In the first tape system for data storage, the data were recorded longitudinally along seven tracks. In a disk system, data are stored along

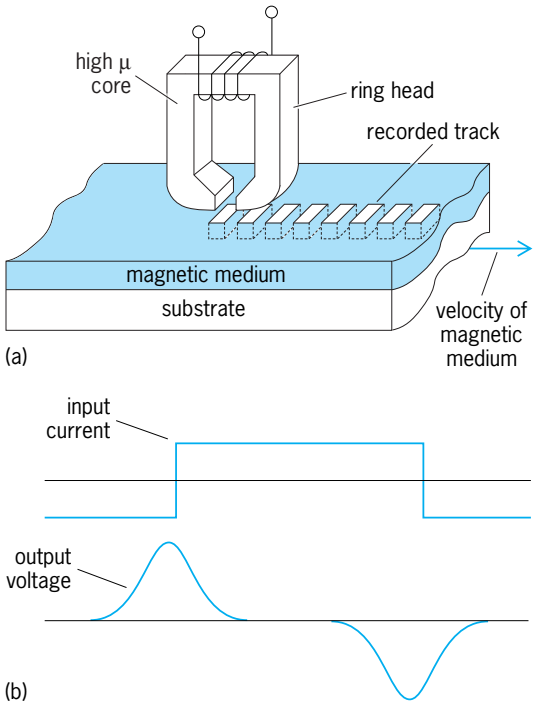

Fig. 2. Writing and reading process. (*a*) Motion of the magnetic medium past the electromagnet in the form of a ring head. (*b*) Variation with time of input current and output voltage.

concentric tracks. The figure of merit is the areal bit density, which is the product of the linear density along a track and the track density. *See* COMPUTER STORAGE TECHNOLOGY.

Digital recording. The three basic types of signals for magnetic recording are audio, video, and digital (**Fig. 3**). The audio signal contains a high-frequency ac bias, and the video signal is frequency-modulated. In the video and digital cases the maximum current is sufficient to produce a field that takes the medium to saturation magnetization (Fig. 1). The video signal is analog in that the location of the zero-crossings is continuous. In the digital signal the zero-crossings occur, or do not occur, in synchronization with a clock signal. *See* FREQUENCY MODULATION.

If an analog signal is sampled at short time intervals and the measured amplitude expressed digitally, it is possible to take advantage of digital processing techniques to eliminate many of the sources of degradation of the analog signal. This is the reason for the widespread use of compact-disk read-only memories (CD-ROMs) and digital audio tapes (DATs). Professional video recorders have also made the transition to a digital format. Consumer digital VCRs are technically feasible. The remainder of this article discusses digital systems. *See* ANALOG-TO-DIGITAL CONVERTER; COMPACT DISK.

The digital input signal, consisting of 1's and 0's, must be represented in the magnetic medium. A straightforward method would be to associate a 1 with one direction of the magnetization and a 0 with the opposite direction. This is referred to as non-return-to-zero (NRZ) recording. A major disadvantage of this approach is that if there is an error in one bit it will propagate indefinitely. Also, large strings of 1's or 0's require very accurate clocking to assure that a bit is not lost. The approach that is now commonly used is to associate a 1 with a transition, and a 0 with the absence of a transition. To avoid the clocking problem, the incoming data are precoded such that after a certain number of 0's, denoted by k, a 1 is inserted into the data stream. Likewise, a string of 1's is not desirable since the associated transitions interfere with one another. Therefore, some number of 0's, denoted d, is inserted between consecutive 1's. The result is that n bits of coded data are used to represent m bits of original data. The ratio m/n is called the code rate, and the code is indicated by m/n (d,k). For finite k, the code is said to be run-length-limited (RLL). Thus, the fundamental physical object in magnetic recording is the magnetic transition.

Writing. When current passes through the coil (Fig. 2), a magnetic field is created in the vicinity of the gap. In magnetic recording systems the magnetization of the medium lies in the plane of the medium. Therefore, this writing is carried out by the longitudinal component of the magnetic field. The contours of constant longitudinal field ($H_{\parallel}$) form nested circles.

The region of the medium that experiences fields exceeding the coercivity (H_c) will be reoriented in the direction of the field. As the medium passes under the head, this reversed region grows out behind the trailing edge of the head. The coercivity

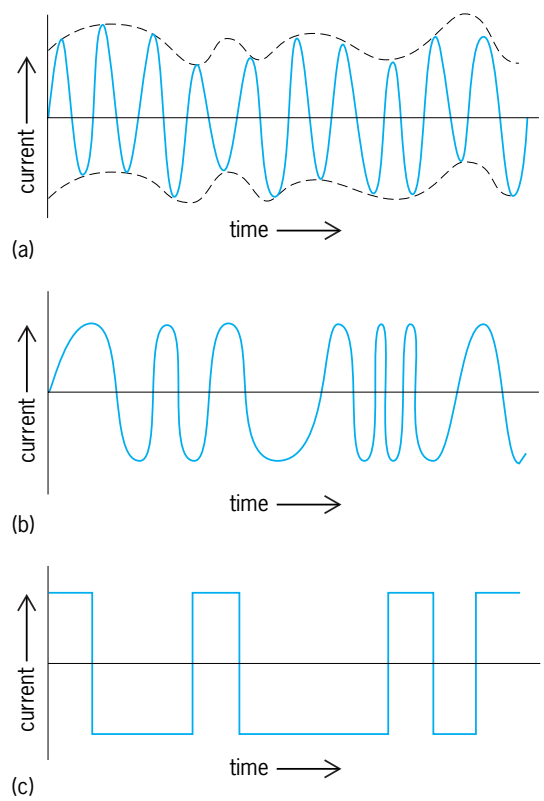

Fig. 3. Three types of input current signals. (*a*) Analog signal (audio with ac bias). (*b*) Frequency-modulated (FM) video signal. (*c*) Digital signal.

cannot be too high or the head will not be able to switch the medium.

Head materials are also characterized by a hysteresis loop (Fig. 1), but for a head material a small remanance is desirable (so the head does not write when the signal current is zero), as well as a large saturation magnetization. Once the core saturates, a larger signal current has little effect on the field produced. Thus, the maximum head field is of the order of half the saturation magnetization of the head material. Earlier heads were commonly of the form called ring heads (Fig. 2), where the electromagnet was machined out of a ceramic made from a class of magnetic oxides known as ferrites. Today the electromagnet is fabricated as a thin-film coil using photolithographic techniques (**Fig. 4**). *See* FERRITE.

Since heads saturate at the pole tips where the field is strongest, one way of extending the maximum field of a ring head is to line the gap with a material having a higher saturation magnetization. This is usually a metallic material, and thus the head is called a metal-in-gap (MIG) head.

Even though the current in the head can be switched virtually instantaneously, the resulting transition in the magnetic medium is not correspondingly sharp. The reason is that like magnetic poles repel. As a result, the transition spreads out over a distance, a, that depends upon the characteristics of this hysteresis loop (Fig. 1), the saturation magnetization, M_s, and the coercivity, H_c, as given by the equation below. Here δ is the thickness of the medium.

$$a = \frac{M_s \delta}{4\pi H_c}$$

The parameter a is important because it determines how closely transitions can be placed and therefore limits the recording density.

Media. The equation for a defines the strategy for achieving higher recording densities. To reduce a, thinner media with higher coercivity are needed. Reducing the saturation magnetization also reduces the readout, so this is not an option. In fact, a larger magnetization is generally desired to give a larger signal.

Particulate media. Early tapes and disks utilized particulate media in which the magnetic ingredient consisted of microscopic particles of a magnetic oxide. These particles were immersed in a polymeric binder that served to separate the particles from one another and bind them to the substrate. The particles first used were the gamma form of iron oxide. These particles are needle-shaped, with a length-to-width ratio of 5-to-1 and a length less than 1 micrometer, and have a coercivity of the order of 200 oersteds (16 kA/m). If the particles are randomly oriented in the tape, the coercivity of the tape is half that of an individual particle. To recover this factor of 2, the particles are physically oriented by exposing them to a magnetic field during the coating process. The search for higher coercivity led to the discovery that impregnating the surfaces of iron oxide particles with cobalt gave coercivities in the range 550–750 Oe (44–60 kA/m). These cobalt-modified particles are widely used in high-performance audio and video tapes.

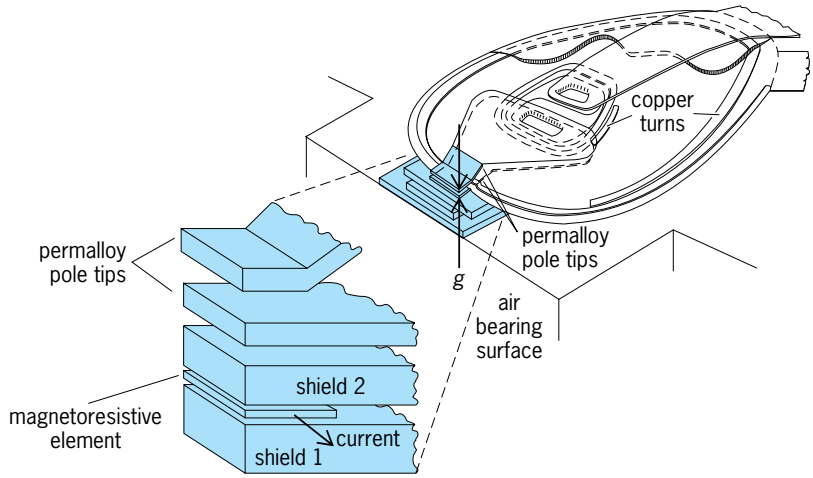

Fig. 4. Thin-film head, showing incorporation of a magnetoresistive read element. Copper turns are a two-dimensional version of the coil in a conventional head. (Not shown are the leads attached to the magnetoresistive element.)

Higher levels of magnetization and coercivity can be obtained in particles made of the ferromagnetic elements, iron and cobalt, and their alloys than is possible with oxide particles. Offsetting these advantages are some serious disadvantages. Metal particles tend to corrode in the atmosphere and to react with binders and so must be passivated at some cost in saturation magnetization. The particles are also difficult to disperse and are much more expensive than particles of iron oxide. Metal particles having coercivities in the range 700–1150 Oe (56–92 kA/m) are used in premium audio tapes, and particles with coercivities of 1350–1550 Oe (107–123 kA/m) are used in 8-mm video tapes.

Metallic films. These are used on magnetic disks to reduce the thickness of the magnetic medium and retain a large magnetization. Among the magnetic elements (iron, cobalt, and nickel), cobalt has a hexagonal crystalline structure that leads to a large coercivity. Therefore, many metallic media consist of cobalt with additional elements to stabilize the hexagonal phase, and also, for longitudinal recording, to ensure that the hexagonal axis lies in the plane of the film. Cobalt-rhenium (Co-Re) is an example. The addition of 10% rhenium stabilizes the hexagonal phase. Another widely used alloy is cobalt-chromium-tantalum (CoCrTa). In many ways, metallic films resemble particulate media. They consist of grains typically 30 nanometers in size. The magnetic properties are strongly dependent upon the size of these grains and the strength of the magnetic coupling between them.

Two techniques are used to deposit magnetic films, sputtering and electron-beam evaporation. Sputtering is primarily used for disks. In this process, the material to be deposited, identified as the target, is placed in a vacuum chamber opposite the disk substrate to be coated. The chamber is filled with argon gas, and a voltage is applied between the target and anode. This forms a plasma of argon ions.

These ions strike the target, knocking out atoms that travel through the plasma and deposit on the substrate. This technique is particularly well suited for the deposition of alloys. *See* SPUTTERING.

A magnetic recording disk actually consists of many layers. The aluminum substrate is first coated with a nickel-phosphorus sublayer to provide hardness. Then a 100-nm chromium underlayer is deposited in order to provide the proper crystallographic texture for the 30-nm magnetic layer, which is deposited next. A 25-nm carbon overcoat is then added for protection.

Because sputtering is a relatively high energy process, this technique does not work well for organic substrates such as the polyethyleneterephthalate (PET) used in most tapes. Therefore, the cobalt-nickel (Co-Ni) films on 8-mm tape, for example, are obtained by electron-beam evaporation of the alloy. The geometry is such that the metal vapor strikes the tape at an oblique angle. As a result, the crystallites making up the film grow as columns at an angle to the film surface. The angle of incidence and the amount of oxygen present during deposition govern the magnetic properties of the tape.

Inductive readout. According to a basic law of physics, any spatial change in the magnetization becomes a source of magnetic fields. Thus, each transition in the medium produces a fountain of magnetic field lines emanating from the surface of the medium. One way of detecting these fields is to use the same transducer that is used to write these data, but in an open-circuit configuration. As the transition sweeps under the head, this magnetic field is attracted into the core because of its large permeability. As the field lines thread through the coils at the back of the head, they induce a voltage in accordance with Faraday's law of induction. *See* MAGNETISM.

As a function of time, this voltage appears as a bell-shaped pulse (Fig. 2). The height of this pulse is proportional, among other factors, to the relative velocity between the head and the medium, the track width, and the number of turns of the coil. The width of this pulse, at half its maximum, depends upon the various geometrical parameters such as the gap between the pole tips, the magnetic spacing between the head and the surface of the magnetic medium, the transition length, and the thickness of the medium.

The linear recording density along a track is limited by the width of this pulse. The interdependence of these geometrical parameters means that to achieve higher densities all these dimensions must be reduced together. Rigid-disk systems use media with a thickness of 30 nm and heads with a gap of 0.2 μm. The magnetic spacing is of the order of 20 nm.

In the early 1980s, most heads consisted of ferrite cores with coils that were hand-wound. In 1989 a thin-film version of an inductive head was introduced. This thin-film structure is deposited on the trailing edge of a slider that is mounted at the end of a swing arm. This slider is designed to create an air bearing between the air-bearing surface (Fig. 4) and the surface of the recording medium. One of the attractive features of a thin-film head is that its pole material, permalloy, has a higher saturation magnetization than ferrite, and therefore can write on higher-coercivity media, which translates to sharper transitions and higher densities.

Magnetoresistive readout. The thin-film head was later modified by the addition of a strip of magnetoresistive material. Magnetoresistance is the phenomenon in which the resistance of a material depends upon the direction in which the current flows relative to the direction of the magnetization. If, in zero magnetic field, the magnetization makes some angle with respect to the current, then in the presence of the field from the media the magnetization will, in general, be tipped away from this direction, producing a change in the resistance (Fig. 4). The magnetoresistive strip is bounded by two layers of soft magnetic material which, like the pole tips themselves, draw in magnetic flux. This has the effect of shielding the magnetoresistive strip so that it senses a magnetic transition only when the transition is directly under the strip; that is, the shields enhance the resolution of the magnetoresistive element. The magnetoresistive effect provides higher sensitivity than inductive readout, and therefore quickly became the preferred readout approach for high-density drives. The separation of the writing and reading functions also means that each can be optimized independently.

In 1988 it was discovered that the resistance of a sandwich of two magnetic films separated by a thin (1-nm) copper layer is dramatically different depending upon the relative orientation of the magnetization in the two films. Thus, if the magnetization in one of the films is pinned while that of the other is free to be oriented by the magnetic field from a transition, the resistance of the sandwich may change by as much as 10% in a field as small as 10 Oe (0.8 kA/m). This giant magnetoresistance (GMR) effect can provide heads with much greater sensitivity than single-strip magnetoresistive heads. *See* MAGNETORESISTANCE.

Recording systems. The major applications for magnetic recording are audio, video, and data storage. A widely used audio system was the compact cassette. This has been largely replaced by the digital compact disk. Digital audio type systems are another viable technology, for example, the rotary digital audio tape (R-DAT) system.

R-DAT. This system uses the helical scan technology developed for video systems (**Fig. 5**). The rotating drum contains two MIG heads. The tape is a metal-particle medium, 0.16 in. (4 mm) wide. R-DAT uses an 8/10 code, which suppresses the low-frequency crosstalk that survives the azimuthal recording.

In the write mode, R-DAT uses a 48 kilosample-per-second sampling rate with a 16-bit analog-to-digital conversion. This leads to a digital signal at 61,000 bits/in. (2400 bits/mm). With the 8/10 code, this means 76,000 magnetic flux transitions per inch (3000 per millimeter). The cassette has dimensions

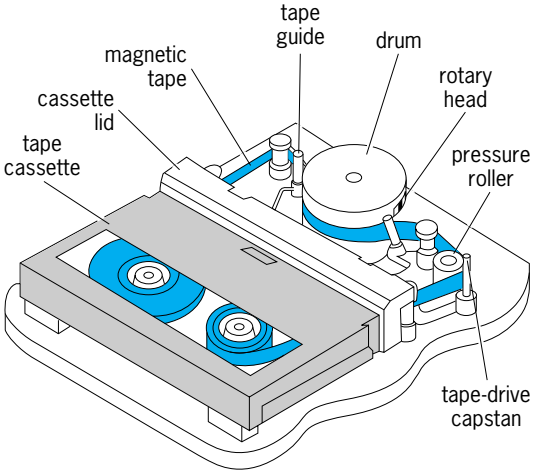

Fig. 5. R-DAT recorder. (*After J. Watkinson, An Introduction to Digital Audio, Focal Press, Butterworth-Heinemann Ltd., 1994*)

of 2.9 × 2.1 × 0.4 in. (74 × 53 × 10 mm) and a volume storage density of 2×10^{11} bits/in.^{3} (1.2×10^7 bits/mm^3). Because of this very high packaging density, this technology is used for data storage in the form of disk backup. Cassettes holding 12 gigabytes are available.

Home video recorders have long been based on a 0.5-in.-wide (13-mm) tape cartridge format. Improvements in heads and media over the years contributed to better picture quality, and in 1984 an 8-mm (0.3-in.) video technology was introduced. The Hi-Band 8-mm VCR uses MIG ferrite heads with metal-evaporated media. Being incompatible with the 0.5-in. format, this technology found success in video camcorders. As with the R-DAT technology, the 8-mm technology was also adapted for digital data storage backup. Cartridges capable of holding 20 gigabytes of data are available.

Rigid-disk drives. Tape systems are inherently sequential and therefore appropriate for audio and video applications. For data applications, however, it is necessary to access data from any location in the file. For this purpose, a disk format is preferable (**Fig. 6**). Data are recorded on concentric tracks on the surfaces of the disks. The heads are mounted at the end of arms that are positioned over the disks by an electromagnetic actuator. Since rapid access to data is important, the disks rotate at high speeds, up to 10,000 revolutions per minute in some drives. In order to reduce wear and prevent contamination, the heads ride above the surface of the disk on a cushion of air, or air bearing; flying heights are of the order of 10 nm. Heights less than this require ultrasmooth surfaces. The typical roughness on a nickel-phosphorus/aluminum (NiP/Al) disk might be 25 nm, while ultrasmooth surfaces require roughnesses less than 0.5 nm. The requirement for extreme cleanliness means the head-disk assembly must be enclosed in its own clean environment. Thus, the disks in high-performance rigid-disk systems are not removable, unlike the diskettes in lower-performance flexible-disk drives.

Track widths are of the order of 3 or 4 μm. To follow such narrow tracks, one approach has been to dedicate one of the surfaces in the stack of disks (Fig. 6) to positional information which is fed back to the actuator. It is then assumed that the other heads on the actuator will be aligned with this servo head. However, as track widths decrease, uneven thermal expansion of the actuator comb can lead to track misregistration. One solution is to dedicate certain sectors on each disk to positional information, thereby embedding the servo information within the data itself.

Various technologies have increased the areal density of rigid-disk systems a millionfold since the first disk system in 1954. This increased areal density has meant that larger capacities can be achieved with smaller disk sizes. The dominant disk size is 3.5 in. (90 mm). One 3.5-in. system, for example, stores 36.7 gigabytes on 6 disks and uses magnetoresistive heads and thin-film media to achieve an areal density of 6 gigabits/in.^{2} (9.3 megabits/mm^2).

At sufficiently low recording densities, the precoding leads to separated output pulses (Fig. 2) that may be detected by looking for peaks in the signal. Most high-performance disk systems use a 1/2 (2,7) or 2/3 (1,7) code. However, the system discussed above has a bit density of 344,000 bits per inch (13,150 bits per millimeter), and at this density the output pulses overlap, resulting in intersymbol interference. The addition of noise further complicates the output signal. To recover the data from such a signal, techniques that were originally developed for communication channels have been adopted.

Magnetic recording channel. The magnetic recording channel behaves almost like an ideal channel studied in signal processing called a class IV partial response (PR4) channel. The advantage of this behavior is that if the isolated pulse shape of a PR4 channel is sampled at time intervals equal to the channel bit period T, all values are zero except those at $t = 0$, the time when the transition was written, and $t = T$. These two nonzero values are equal and assigned the value 1. Thus, the samples of an isolated pulse are – – – 0 0 1 1 0 0 0 – – –. If an opposite transition is written at time $t = T$, the linear superposition of the two transitions (called a dibit) gives the samples – – – 0 0 1 0–1 0 0 0 – – –. The samples of an arbitrary stream of data will consist of 1's, 0's, and −1's.

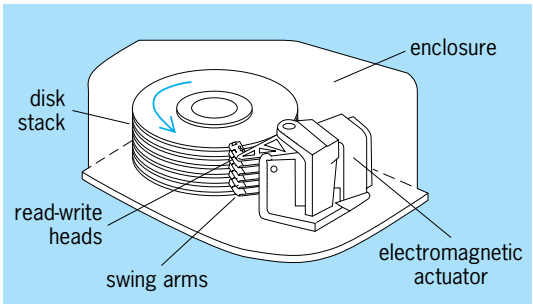

Fig. 6. Rigid-disk drive. (*After C. D. Mee and E. D. Daniel, eds., Magnetic Recording Handbook: Technology and Applications, McGraw-Hill, 1990*)

Since the magnetic recording channel is not an ideal PR4 channel, and since there is obviously noise present, there will be some uncertainty. Rather than trying to decide whether one sample is a 1, 0, or −1, a sequence of samples is examined and the sequence with the maximum likelihood is determined. For example, the sequence – – – 1 1 1 – – – cannot exist since transitions must alternate. The algorithm used to establish the most likely sequence was developed by Andrew Viterbi, and is implemented as a Viterbi detector. The use of such a partial-response maximum-likelihood (PRML) technique provides a 1.5-dB signal-to-noise gain over peak detection. *See* ELECTRICAL NOISE; SIGNAL-TO-NOISE RATIO.

Limits to recording density. Magnetic recording has been a remarkably resilient technology, and hard-disk areal densities have increased over a million times since the first product. The nature of the ultimate limits to recording density has occupied engineers since the technology was introduced. Today, the major challenge is thought to arise from thermal effects. Thin-film media are composed of crystalline grains. As the recording density increases, the number of grains contributing to a bit must at least stay the same in order to maintain a given signal-to-noise ratio. This entails the preparation of media with smaller grains. However, as the grains become smaller, they eventually reach a size where their magnetic properties decrease with time due to thermal fluctuations. Such fluctuating magnetic particles are said to be superparamagnetic. The result is that the data decay with time, a very undesirable situation. Various approaches have been proposed to stabilize media at high densities. One is to place the recording medium on top of an antiferromagnetic underlayer, which serves to increase the volume of the bit cell. Another approach is to record perpendicular to the plane of the media, which also allows somewhat thicker media. It is believed that these approaches may extend the superparamagnetic limit well beyond 100 gigabits/in.^{2} (150 megabits/mm^2).

Perpendicular recording. As discussed above, the transition length arises from the demagnetizing fields associated with the opposing magnetic alignments. If, however, the alternating magnetic regions are perpendicular to the plane of the medium, the north and south poles of these regions occur next to one another. Since this is a favorable situation, the transition length in this case actually becomes smaller with increasing recording density. This was the basis for early interest in perpendicular recording. However, a small transition length by itself is not advantageous. A small gap, thin media, and a small head-to-medium spacing are also necessary. While the fringing field from a ring head (Fig. 2) has a vertical component that can be used to write on perpendicular media, it is more efficient to use a pole head in connection with a magnetically soft (that is, low-coercivity) underlayer.

Optical storage. Optical storage encompasses read-only optical systems such as the CD-ROM, write-once systems, and erasable optical systems. In the CD-ROM systems, the data are impressed on a disk by a stamping process and read out by means of a solid-state laser. This medium is widely used for distributing system documentation and software. In write-once systems the user is provided with a medium whose state may be irreversibly changed by the application of local heating by a focused laser beam. This change may be from an amorphous to a crystalline state, or it may involve actual deformation of the material by local melting. The medium is typically an alloy of tellurium. Once created, the data may be read many times by a lower-power laser.

Phase-change recording. The tellurium-based materials used for write-once read-many-time (WORM) systems may also be modified for erasable applications. As an alloy such as germanium-antimony-tellurium (Ge-Sb-Te) of suitable composition is rapidly cooled from above its melting point, it solidifies into a glassy, or amorphous, state in which the atoms are randomly arranged. If this amorphous material is heated, it crystallizes before it melts. Thus, by controlling the power level and the pulse length of a laser beam, it is possible to reversibly alter the state of such materials. The reflectivities of these two states are generally quite different. Therefore, the readback signal is much larger than that obtained, for example, in a magnetooptic system. The challenge is to find materials that can be crystallized quickly with a laser but are thermally stable at room temperature.

Magnetooptic effect. Another widely used technology for erasable optical storage is based on the magnetooptic effect. Since light consists of oscillating electric and magnetic fields, when light passes through a magnetic material it undergoes change. One of the features of light that is affected is polarization. When a laser beam is emitted from a solid-state laser, its electric field oscillations lie in a plane governed by the geometry of the laser, and the beam is said to be linearly polarized. When such a beam passes through a magnetic material in which the direction of the magnetization is parallel (or antiparallel) to the beam, the plane of the polarization of the light is rotated clockwise (or counterclockwise). This magnetooptic effect is called the Faraday effect. A similar change that occurs when light is reflected from the surface of a metallic magnetic material (**Fig. 7**) is the magnetoopic Kerr effect. For typical materials, this rotation is of the order of 1°. *See* FARADAY EFFECT; MAGNETOOPTICS.

Perpendicular magnetization. To utilize this magnetoopic effect on a storage system requires a medium with a perpendicular magnetization. Amorphous films containing the rare-earth elements, such as gadolinium and terbium, together with the transition-metal elements iron, cobalt, and nickel, can be prepared with a perpendicular magnetic orientation. Furthermore, the coercivities of these materials are very temperature dependent. These properties are the basis for the use of these materials in optical storage.

A composition is found for which the coercivity at room temperature is very high but decreases with increasing temperature. Typical materials are terbium-iron-cobalt (TbFeCo) and gadolinium-terbium-iron

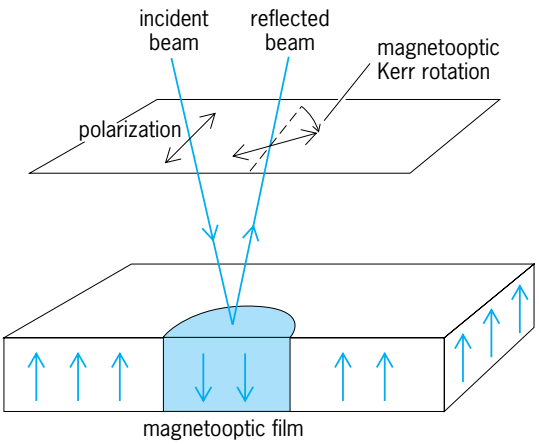

Fig. 7. Cross section of a magnetooptic film showing rotation of polarized light upon reflection.

(GdTbFe). If the material is heated with a laser beam in the presence of a magnetic field, then as the coercivity decreases, the heated region within the laser spot will align with the applied field. When the heating is removed, the coercivity again increases, freezing in the cylindrical domain. Assuming that the disk was initially all polarized, say, "down," an "up" region would be identified as a 1, in analogy with the magnetic transition on a rigid disk. Magnetooptic systems therefore also precode the signals with a run-length-limited code such as the (2,7) code.

Areal densities. Since the power in the laser beam has a bell-shaped profile, the size of the recorded spot will depend on the corresponding temperature profile in the medium and can therefore be made arbitrarily small. However, the detection of spots is limited by the diffraction that takes place during readout. Assuming that the diffraction spots just touch, their separation becomes the limiting diameter of a bit cell. Since current solid-state lasers for recording have infrared wavelengths, spot sizes are of the order of 1 μm. This gives areal densities in the range of 540 megabits/in.^{2} (0.8 megabit/mm^2) and total capacities on a 5.25-in. (130-mm) removable cartridge of 1.3 gigabytes. Optical storage was once

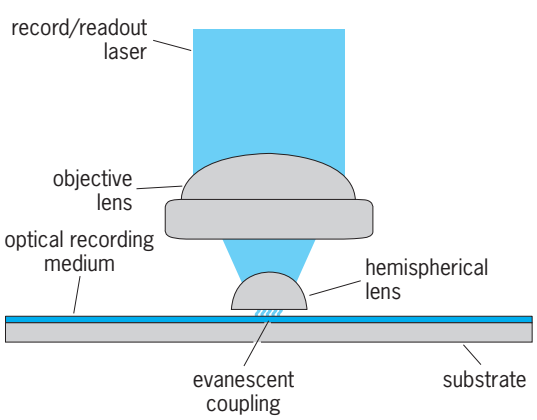

Fig. 8. Near-field optical recording system.

favored for its high areal densities, but magnetoresistive heads and thin-film media enabled conventional magnetic recording densities to surpass optical densities. However, optical densities will increase with higher-frequency lasers. Magnetooptical disks also have the attractive feature of being removable. *See* LASER.

Biasing and direct overwrite. One of the challenges in a magnetooptic system is to generate the biasing magnetic field. Early designs used an electromagnet. However, since it is difficult to switch such a magnet at recording data rates, these systems required two revolutions to rewrite data, a distinct disadvantage. Of course, it is possible to use a conventional recording head flying close to the disk, but this compromises the advantage of optical systems, which is their beam-addressable nature, the fact that the head is a focal length away from the medium. This is what makes it possible to encapsulate the medium, so that it is removable.

Approaches to achieving direct overwrite in magnetooptic systems involve additional layers of material. For example, if a memory layer, with a relatively high coercivity that is sensitive to temperature, is placed on top of a reference layer, with a low coercivity that is relatively insensitive to temperature, then as the bilayer passes through an initializing field, the reference layer is polarized by this field. However, the data in the memory layer are unaffected. If the bilayer now moves into a region with a reverse-bias field and is simultaneously exposed to a low-power laser pulse that reduces the coercivity of the memory layer, this layer will be oriented by the strong field from the reference layer. If, however, the laser pulse has enough power to reduce the coercivities of both layers, both will be aligned by the bias field. Thus, by modulating the laser power, direct overwrite can be achieved.

Near-field recording. It is possible to improve on the diffraction limit if a very small head-to-disk spacing is acceptable. The idea is to focus the optical beam through a hemispherical lens to a point at the back (flat) face of the lens (**Fig. 8**). This arrangement reduces the diffraction limit by the index of refraction of the lens (typically a factor close to 2). The beam is "immersed" in the solid lens, which is therefore called a solid immersion lens (SIL). The disadvantage of this scheme is that in focusing a broad beam within the lens some of the light is internally reflected. It is a property of internally reflected light that the field of the light actually extends beyond the boundary; this extension is called an evanescent wave, or near-field. If the medium is placed within this near-field, it interacts with the beam and utilizes the higher resolution. *See* OPTICAL RECORDING.

Robert M. White

Bibliography. R. L. Comstock, *Introduction to Magnetism and Magnetic Recording*, Wiley, 1999; E. D. Daniel, C. D. Mee, and M. H. Clark (eds.), *Magnetic Recording: The First 100 Years*, IEEE Press, 1999; C. D. Mee and E. D. Daniel (eds.), *Magnetic Recording Technology*, 2d ed., McGraw-Hill, 1996; S. X. Wang and A. M. Taratorin, *Magnetic Information Storage Technology*, Academic Press, 1999.

Magnetic relaxation

The relaxation or approach of a magnetic system to an equilibrium or steady-state condition as the magnetic field is changed. This relaxation is not instantaneous but requires time. The characteristic times involved in magnetic relaxation are known as relaxation times.

Magnetism is associated with angular momentum called spin, because it usually arises from spin of nuclei or electrons. The spins may interact with applied magnetic fields, the so-called Zeeman energy; with electric fields, usually atomic in origin; and with one another through magnetic dipole or exchange coupling, the so-called spin-spin energy. Relaxation which changes the total energy of these interactions is called spin-lattice relaxation; that which does not is called spin-spin relaxation. (As used here, the term lattice does not refer to an ordered crystal but rather signifies degrees of freedom other than spin orientation, for example, translational motion of molecules in a liquid.) Spin-lattice relaxation is associated with the approach of the spin system to thermal equilibrium with the host material; spin-spin relaxation is associated with an internal equilibrium of the spins among themselves. *See* MAGNETISM; SPIN (QUANTUM MECHANICS).

The measurement of relaxation times is of interest in the study of molecular motions and interactions. An important practical application of magnetic relaxation is the attainment of very low temperatures. For example, a sample of a paramagnetic salt can be cooled by liquid helium to about 1 K (2°F above absolute zero, −459.67°F), then subjected to a magnetic field. Heat, evolved in the magnetization, will be dissipated into the liquid helium. If thermal contact to the helium is broken, and the magnetic field is then removed, the relaxation of the sample to more random orientation will be accompanied by a further decrease in temperature, in some cases to 0.01 K (0.02°F above absolute zero). A similar adiabatic demagnetization of copper, in this case involving nuclear spin, has cooled the sample to 0.00002 K (0.00004°F above absolute zero). *See* ADIABATIC DEMAGNETIZATION.

Relaxation process. In an unmagnetized sample (ferromagnets are excluded for the moment) the spin orientations are random. In the presence of a static magnetic field and at thermal equilibrium, an excess of the spin vectors points in the lower energy orientation along the field. The buildup of magnetization following application of the field requires that some of the spins turn from antiparallel to parallel orientations. The energy given up in this turning process may go to either or both of two reservoirs. If the applied field is weak, the spin-spin energy constitutes an effective reservoir. If the applied field is strong, the spin-spin (free) energy is negligible and magnetization requires an interchange with the lattice. At low fields the transfer of energy into the spin-spin reservoir may be thought of as corresponding to a heating of the spin system. The spin-lattice coupling will eventually bring the spins into thermal equilibrium with the lattice (by changing either the spin "temperature" or the lattice temperature, or both). Thus for low fields the process of magnetization would exhibit two stages: (1) the redistribution of energy in the spin system, or spin-spin relaxation, and (2) equilibration of the spin system with the lattice, or spin-lattice relaxation. For high fields, only spin-lattice relaxation plays a role in establishing the magnetization.

Although relaxation exhibits itself in a macroscopic manner, an explanation of the rates of relaxation requires consideration of atomic processes. Two requirements must be met for spin-lattice relaxation: (1) There must be a coupling of the spin to the lattice, and (2) the coupling must be time-dependent, its frequency spectrum containing natural frequencies of the spin system (for example, the precession frequency of the spins in the applied field).

Types of relaxation. Relaxation has been studied for nuclear magnetism, electron paramagnetism, and ferromagnetism.

Nuclear relaxation. For nuclei, the Zeeman and spin-spin interactions are always important, and in some cases, the nonspherical charge distribution (electric quadrupole moment) of a nucleus affects the relaxation time. When the quadrupolar energy may be neglected, the magnetization usually approaches thermal equilibrium exponentially after application of a strong static field, the time constant being the spin-lattice relaxation time. Nuclear relaxation is almost invariably studied by magnetic resonance; the weak nuclear paramagnetism cannot be detected by other methods. *See* MAGNETIC RESONANCE.

If bulk nuclear magnetization **M** is tilted away from the direction of the static field **H**, the **M** vector precesses about **H**. The component of **M** parallel to **H** returns to its thermal equilibrium value by means of spin-lattice relaxation. The spin-spin coupling gives a spread in precession frequencies, causing the nuclei to get out of step with one another, as manifested by the decay of the components of **M** perpendicular to **H** without energy exchange with the lattice (spin-spin relaxation).

Rapid motion of a nucleus relative to its neighbors lengthens the spin-spin relaxation. The times range from tens of microseconds, in almost all solids, to several seconds, in cases of rapid nuclear motion. Measurement of spin-spin relaxation has been extensively used to study atomic motion.

Spin-lattice relaxation arises from magnetic dipole coupling with electrons in the case of metals, paramagnetic substances, or diamagnetic insulators possessing paramagnetic impurities; with other nuclear moments when the relative nuclear positions undergo large changes due to sufficiently rapid self-diffusion or molecular rotation in liquids, gases, or even some solids; or from the nuclear electric quadrupole coupling to the electric fields of the host material. Spin-lattice relaxation times range from less than 1 ms to several hours.

Paramagnetic relaxation. Resonance methods can be used for the study of paramagnetic relaxation. However, the observation of the ability of the

magnetization to follow a changing field and even the observation of the change in the lattice temperature brought about by spin-lattice relaxation have also been used. The first interest in magnetic relaxation, in fact, stemmed from the desire to cool the lattice by adiabatic demagnetization of the spins of paramagnetic ions. Large electrostatic interactions between the ions and their surroundings make major contributions to the total spin energy, and it is rarely possible to neglect this effect. Other important contributions to the spin energy arise from dipolar coupling to applied fields, spin-spin coupling by both magnetic dipole and exchange mechanisms, and magnetic coupling to nuclei, particularly the nuclei of the paramagnetic ions.

The principal spin-lattice relaxation mechanisms involve modulation of the spin-orbit coupling by lattice motion, modulation of the magnetic coupling to nuclei by lattice motion, and coupling with paramagnetic impurity atoms. Typical spin-lattice relaxation times vary from less than 10^{-8}s at room temperature to tens of seconds at liquid helium temperature. Although the study of paramagnetic relaxation was begun in the 1930s, it is not as thoroughly understood as the more recently studied nuclear relaxation, primarily because of the complexity of the electronic energy levels of the paramagnetic ions.

Ferromagnetic relaxation. In ferromagnetic substances, strong exchange coupling between electron spins, causing them to prefer an alignment parallel to one another, and demagnetizing effects cause ferromagnetic relaxation to differ from paramagnetic relaxation.

The strong tendency of ferromagnetic spins to remain parallel makes it convenient to distinguish between two types of relaxation: (1) that in which the total magnetization changes direction but not magnitude, thereby keeping the exchange energy constant, and (2) that in which the magnitude of the magnetization changes, producing a change in the exchange and also in the demagnetizing energies. The second mechanism depends on the sample shape and may arise from irregularities in the ferromagnetic lattice (impurities and nonstochiometric composition, as in ferrites) as well as from thermal vibrations. Ferromagnetic relaxation has been studied primarily by magnetic resonance. Charles P. Slichter

Magnetic resonance

A phenomenon exhibited by the magnetic spin systems of certain atoms whereby the spin systems absorb energy at specific (resonant) frequencies when subjected to alternating magnetic fields. The magnetic fields must alternate in synchronism with natural frequencies of the magnetic system. In most cases the natural frequency is that of precession of the bulk magnetic moment **M** of constituent atoms or nuclei about some magnetic field **H**. Because the natural frequencies are highly specific as to their origin (nuclear magnetism, electron spin magnetism, and so on), the resonant method makes possible the selective study of particular features of interest. For example, it is possible to study weak nuclear magnetism unmasked by the much larger electronic paramagnetism or diamagnetism which usually accompanies it.

Nuclear magnetic resonance (that is, resonance exhibited by nuclei) reveals not only the presence of a nucleus such as hydrogen, which possesses a magnetic moment, but also its interaction with nearby nuclei. It has therefore become a most powerful method of determining molecular structure. The detection of resonance displayed by unpaired electrons, called electron paramagnetic resonance, is also an important application. These two phenomena, as well as other related resonance phenomena, are discussed in this article. *See* MAGNETISM.

Origin. Because **M** has its origin in circulating currents or intrinsic spins, there is always an angular momentum **J** associated with it. The vector quantities **M** and **J** are related by Eq. (1), where γ is called the gyromagnetic ratio.

$$\mathbf{M} = \gamma \mathbf{J} \tag{1}$$

For a system to exhibit a magnetic resonance, it must possess a magnetic moment, possess angular momentum, and experience torques. The magnetic moment may arise from nuclei of atoms, from orbital electronic motion, from electronic spins, or from moving nuclear charges during molecular rotation. The angular momentum arises from the same sources. The torques may arise from externally applied magnetic fields; from magnetic dipole fields exerted by neighboring nuclei, atoms, or molecules; from electric fields acting on, for example, a nuclear electric quadrupole moment or the nonspherical electron cloud of an atom; or from electron exchange coupling. In any given case it is necessary to decide which interactions are large and which are small. Thus, for a paramagnetic atom possessing a nuclear magnetic moment, a distinction might be made between small applied static fields, in which the coupled nucleus and electron angular momenta act as a unit, and large magnetic fields, which decouple them so that they act independently. An effective angular momentum and magnetic moment can often be defined, giving an effective γ (in analogy to the Landé g factor of optical spectroscopy).

Several types of resonances have been observed; these differ in one or more of the three basic requirements. However, the principal features can be understood by assuming that the torques arise from an effective static applied magnetic field **H**. The torque $\mathbf{M} \times \mathbf{H}$ causes the angular momentum to change with time according to Eq. (2). The resultant motion

$$\frac{d\mathbf{J}}{dt} = \mathbf{M} \times \mathbf{H} \qquad \text{or} \qquad \frac{d\mathbf{M}}{dt} = \gamma \mathbf{M} \times \mathbf{H} \tag{2}$$

of **M** is a precession at angular frequency γH about the direction of **H**. At thermal equilibrium, **M** is parallel to **H**, and no precession occurs. Application of an alternating magnetic field $H_x \cos \omega t$ perpendicular to **H** causes **M** to tilt away from **H** with a consequent absorption of energy, provided the resonant condition

$\omega = \gamma H$ is satisfied. In practice, the absorption takes place over a narrow range of frequency on both sides of γH. The magnetization M_x parallel to H_x obeys Eq. (3), where $\chi'(\omega)$ and $\chi''(\omega)$ are the real and imag-

$$M_x = H_x \chi'(\omega) \cos \omega t + H_x \chi''(\omega) \sin \omega t \qquad (3)$$

inary parts of the complex magnetic susceptibility $\chi = \chi'(\omega) - j\chi''(\omega)$ ($j = \sqrt{-1}$), and characterize dispersion and absorption respectively. For a typical resonance, $\chi''(\omega)$ attains maximum value for a region of frequencies near $\omega = \gamma H$.

For many cases, it is necessary to analyze magnetic resonance by quantum theory. Consider a system composed of many (weakly interacting) identical parts (atoms or nuclei), each with angular momentum quantum number F (total angular momentum $= \sqrt{F(F+1)}\,\hbar$, where $\hbar$ is Planck's constant h divided by 2π). The spatial quantization in the magnetic field H gives $2F + 1$ equally spaced energy levels labeled by the quantum number $M_F = (F, F-1, \ldots, -F)$ and energy spacing between adjacent levels of $\gamma\hbar H$. The field $H_x \cos \omega t$ produces transitions with the selection rule $\Delta M_F = \pm 1$. To satisfy the Bohr frequency condition, $\hbar w = \gamma \hbar H$, in agreement with the classical result, $\omega = \gamma H$.

According to quantum theory, the probability of transition from any energy level A to any other B is the same as that from B to A; thus a net absorption of energy requires that the population of the lower energy states be greater than that of the upper. The reverse situation in which the upper states are more populated leads to an induced emission and is the basis for the operation of the solid-state maser. At thermal equilibrium, as a result of spin-lattice relaxation, the lower energy states are more heavily populated in accordance with the classical Maxwell-Boltzmann statistics (ordinarily it is unnecessary to use either Fermi-Dirac or Bose-Einstein statistics). When H_x becomes sufficiently large, the level populations become disturbed from thermal equilibrium. The population difference between states joined by H_x decreases, a phenomenon known as saturation, because it causes χ' and χ'' to diminish with increasing H_x. Population differences between pairs of states other than A and B may simultaneously be increased, or even inverted, as in the three-level maser. The intensity of the alternating field necessary to produce saturation depends on the width of the absorption line and on the spin-lattice relaxation time (wider lines or shorter times require larger H_x). *See* MAGNETIC RELAXATION; MASER; NONRELATIVISTIC QUANTUM THEORY.

Observation. Experimentally it is possible to detect magnetic resonance by measuring the absorption of magnetic energy of a circuit containing the magnetic material or by measuring the change in inductance or resonant frequency of the circuit. The two methods measure $\chi''(\omega)$ or $\chi''(\omega)$, respectively. The resonant condition $\omega = \gamma H$ may be produced by varying ω, or more customarily, by changing H. In some experiments, one tilts **M** away from the direction of **H** by alternating fields of short duration and then observes voltages induced by the subsequent free precession of **M**. This method is particularly useful for studying relaxation times.

Nuclear magnetic resonance (NMR). The nuclei of many atoms possess angular momentum (spin) and nonvanishing magnetic moments. The former may be characterized by an angular momentum quantum number I (integer or half integer) of the nuclear particles. As far as is known, stable nuclei with an even number of neutrons and even number of protons have zero spin and magnetic moment, and hence are incapable of exhibiting magnetic resonance. *See* NUCLEAR MOMENTS.

Nuclear resonances have been observed in insulators, metals, paramagnetic salts, antiferromagnetic substances, and other solids, and in gases and liquids. Often, to observe NMR a sample is placed between the poles of an electromagnet (**Fig. 1**) which in addition to the main winding carries a small auxiliary winding or sweep. A coil connected to an oscillator surrounds the sample, as does a second coil at right angles to both the oscillator coil and the sweep winding, to avoid direct coupling. The oscillator frequency is fixed and the sweep circuit is used to vary the magnetic field strength continuously. When a resonance frequency of the sample is reached, a signal induced in the second coil is detected and amplified. Typical resonance frequencies in a field of 10,000 gauss lie in the radio-frequency region (1–45 MHz). For example, the H^1 nucleus shows a resonance frequency of 42.6 MHz at this field strength. ^{13}C nuclei give a much weaker signal, further decreased in a sample containing this isotope in its natural abundance of 1.1%.

Nuclei with quadrupole moments. If a nucleus has a spin $\geqq 1$, it generally has a nonvanishing electric quadrupole moment, expressing the deviation of its spatial change distribution from spherical symmetry (there are good grounds for believing that all nuclei have zero electric dipole moments). The electrical interaction between the nucleus and electric potentials $V(x,y,z)$ from other charges depends on the nuclear orientation (specified by the direction of nuclear spin) and on the spatial second derivatives of

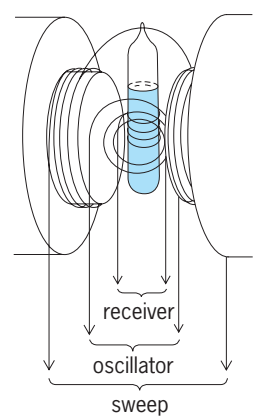

Fig. 1. Arrangement of a sample and coils in a nuclear induction apparatus. (*After J. D. Roberts, Nuclear Magnetic Resonance, McGraw-Hill, 1959*)

the potential $\partial^2 V/\partial x^2$, $\partial^2 V/\partial x\ \partial y$, and so on, at the position of the nucleus. For potentials of spherical, tetrahedral, or cubic symmetry, the interaction energy is independent of orientation and may be disregarded. When the electric quadrupole interaction is nonzero but nevertheless much weaker than the static magnetic interaction, the unique resonance condition $\omega = \gamma H$ is changed, and the resonance line splits into $2I$ components centered around γH. A convenient method of determining the nuclear spin is thereby provided.

The name "nuclear quadrupole resonance" is used when one dispenses with the static field H and observes the reorientation of the nucleus among its various quantized orientations with respect to the electric potential alone. Classically, the nonspherical nuclear charge experiences a torque which causes precession. This technique is particularly useful when the electric quadrupole coupling is large and the sample is available only in powder form. For such a sample, the NMR lines tend to be broad, owing to the distribution of the orientations of the powder granules with respect to the direction of the applied magnetic field. No such broadening arises in the absence of a magnetic field since the frequency of the nuclear quadrupole resonance does not depend on spatial orientation.

Applications. Nuclear magnetic resonance has been used widely to measure nuclear magnetic moments, electric quadrupole moments, and spins. Because the resonance lines may be on occasion very sharp (1 cycle wide at 40 MHz), nuclear resonance is frequently employed to measure magnetic fields with great precision. For an example *see* MAGNETOMETER.

The extensive use of NMR in molecular structure determinations arises from the slight shift in the resonance frequency of an atom—commonly that of a proton—due to the environment of neighboring atoms. Because the magnitude of this shift depends on the type of environment, it is called the chemical shift. **Figure 2** shows the NMR spectrum of ethyl alcohol, CH_3CH_2OH. The three main resonance frequencies are due to protons in the OH, CH_2, and CH_3 groups, respectively, and the spacing between them (which varies with the field strength) shows the chemical shift characteristic of the protons in these three typical structural groups.

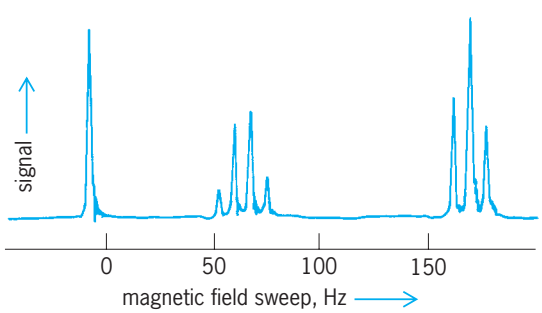

Fig. 2. Proton resonance spectra of ethyl alcohol at 40 MHz. (*After J. D. Roberts, Nuclear Magnetic Resonance, McGraw-Hill, 1959*)

The separate peaks at each frequency are due to spin-spin splitting. Often, the presence of n protons will split the frequency of a given, structurally different proton into $n + 1$ peaks, in direct analogy to ordinary spectral lines. The triplet in the NMR spectrum of ethyl alcohol results from the splitting of the frequency of the CH_3 protons by the two adjacent CH_2 protons; the quadruplet is at the typical CH_2 frequency and is split into four peaks by the three protons of the CH_3 group. *See* MOLECULAR STRUCTURE AND SPECTRA.

Because the time required for nuclear transitions is relatively long, substances undergoing fast reactions show altered NMR spectra, and rates of such rapid processes as ionization or intramolecular rotation can be measured. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

Magnetic resonance imagining (MRI). Ordinarily in magnetic resonance, it is desirable to have the static field, H_0, as uniform as possible in order that the NMR lines reveal the detailed structure originating in the chemical shifts and spin-spin coupling. In 1973, Paul Lauterbur and Peter Mansfield independently realized that by making the field nonuniform in a controlled manner one could form spatial images of the number density and other properties of nuclei in a sample. This method is the basis for MRI.

The general principles of MRI are as follows. In order to image a sample, it is placed in a strong uni-form laboratory magnetic field oriented in the z direction, of strength H_0. Added to that field is a second one, called the gradient field, whose z component, H_G, at spatial position specified by coordinates (x,y,z), is a linear function of those coordinates [Eq. (4), where a, b, and c are constants controlled

$$H_G(x, y, z) = ax + by + cz \tag{4}$$

by the experimenter]. Such a gradient is called a linear gradient. The resonance angular frequency, ω, now becomes a function of position, obeying the law shown in Eq. (5). Choice of the constants a, b, and

$$\omega(x, y, z) = \gamma\,[H_0 + H_G(x, y, z)] \tag{5}$$

c defines planes in space over which the resonance frequency has the same value, and how rapidly the magnetic field changes when one moves from one plane to another. For example, if only a is nonzero, all nuclei with the same coordinate x have the same resonance frequency, independent of their y or z coordinates. The nuclei between two planes $x = x_1$ and $x = x_2$ would produce a resonance extending from frequency $\omega(x_1) = g[H_0 + ax_1]$ to frequency $\omega(x_2) = g[H_0 + ax_2]$ whose total intensity is proportional to the number of nuclei in the sample between these two planes. The size of a determines how big a change in frequency accompanies a change in x. Since a is controlled by the experimenter, it is a known quantity, and thus the relationship between frequency and x is also known.

The NMR absorption is recorded as a function of frequency. A plot of NMR signal strength versus ω can

thus be converted to a plot of NMR signal strength versus x. It gives one the number of nuclei per unit distance in the x direction versus x. By varying the quantities a, b, and c, one can change the spatial orientation of the planes over which the magnetic field values are uniform (that is, one can change the direction of the field gradient). One proceeds to gather data for a variety of directions of gradient. These data can then be processed in a computer using well-known mathematical methods (such as Fourier transforms) to obtain n (x,y,z), the number of nuclei per unit volume at position x,y,z in the sample. *See* FOURIER SERIES AND TRANSFORMS.

Such results for planes in space (called slices) are often displayed as images whose darkness at a particular spatial position corresponds to the numerical value of n at that position. The resulting pattern of varying lightness and darkness gives an "image" of n (x,y,z). A given set of experimental data can be used to display slices of any orientation.

A unique feature of NMR images is that the strength of the NMR signal may be adjusted to tell more than just the spatial distribution of the number density of nuclei n (x,y,z). For example, a so-called T_1 weighted image shows how the spin-lattice relaxation time, T_1, varies with position. One may also give a T_2-weighted image to show spatial variation of the transverse relaxation time, T_2. Or one can pick a given region of space to display the chemical shift data. Using these methods, it is possible to use NMR images to display flow patterns of fluids, such as blood, or to show the existence of spatial diffusion of nuclei. These latter quantities are useful in diagnosis of stroke.

Two quantities of great interest are the spatial resolving power of the imaging method and apparatus and the length of time required to form an image. In general, the higher the spatial resolution, the longer the time to form the image. In practice, the sequence of gradients which are employed (through the sequence of values of a, b, and c employed to form the image) and the method of recording the NMR spectra control how quickly an image can be formed and how high its spatial resolution will be. The method of choice depends on the purpose for which the image is being obtained. For some medical applications, for instance, one may wish to emphasize speed of data acquisition (reducing data acquisition times to minutes rather than hours) at the expense of spatial resolving power. A research area of growing importance is "functional imaging," in which the goal is to locate where in the brain the response occurs to an external stimulus.

Electron-nuclear double resonance (ENDOR). In this technique the magnetic resonance of a nucleus is detected by observing that of a nearby electron. The magnetic coupling between the nuclear and electronic magnetic moments gives rise to a back reaction on the electron resonance when the nuclei are brought into resonance. The sample under study is placed in a conventional electron resonance apparatus. In addition, an oscillator drives a coil to produce alternating magnetic fields at the sample under study, the frequency of alternation being in an appropriate range to produce nuclear transitions. Typically, the static magnetic field is adjusted to produce electron resonance. With the static magnetic field and the frequency of the electron resonance apparatus held fixed, the frequency of the nuclear resonance oscillator is then swept. As it passes through the resonant frequency of any nucleus that is coupled to the electron, a change in the electron absorption occurs.

The electron resonance and the nuclear resonance both represent transitions between energy levels of the combined system of electron and nucleus, each resonance being between a pair of levels. The strength of the electron resonance depends on the difference in population between its two levels. Thus, if one of these levels takes part in both resonances, the nuclear resonance may influence the electron resonance by changing the population of the common energy level.

Since the individual quanta absorbed in an electron resonance are much larger in energy than those absorbed in a nuclear resonance, the double resonance often permits the detection of the resonant absorption of a smaller number of nuclei than it would be possible to detect by a direct observation of the nuclear resonance. Since ENDOR requires both electron and nuclear resonances, it is applicable only to systems possessing both. It has been used, for example, to study the nuclear resonance associated with impurity atoms in semiconductors and that associated with F-centers in alkali halides, as well as to measure nuclear moments of rare elements. Important results have included the determination of the electronic structure of point imperfections and the measurement of nuclear magnetic moments and hyperfine anomalies.

Paramagnetic resonance. Magnetic resonance arising from electrons in paramagnetic substances or from electrons in paramagnetic centers in diamagnetic substances is called paramagnetic resonance. For applied fields of several thousand gauss, the electron paramagnetic resonance (EPR) experiments are done at microwave frequencies, commonly at 3-cm (1.2-in.) or at 1-cm (0.4-in.) wavelengths. In some instances, nuclear resonance apparatus has been used with correspondingly lower applied fields. The most sensitive apparatus detects approximately 10^{12} electron spins for a line 1 gauss broad, a sensitivity far greater than that obtained by nonresonant methods, such as those utilizing paramagnetic susceptibility.

Resonances have been observed in atoms of the iron group, rare earths, and other transition elements; in paramagnetic gases; in organic free radicals; in color centers in crystals (such as F- and V-centers); in metals (conduction electron spin); and in semiconductors (both conduction electron and impurity center spins).

When two paramagnetic ions or molecules approach one another, the spins become coupled via the exchange interaction. Much weaker couplings than the usual exchange interactions within atoms, in chemical bonds, or in ferromagnets produce pronounced effects. For additional information on this

phenomenon and its applications . *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY.

Ferromagnetic resonance. In the case of both nuclear and paramagnetic resonance, the spins of neighboring atoms are nearly randomly oriented with respect to one another. In contrast, the electron spins in one domain of a ferromagnet are nearly all parallel for temperatures sufficiently below the Curie point. The alignment may be described in terms of the exchange coupling between neighboring spins, or equivalently in terms of the Weiss molecular magnetic field $\mathbf{H}_w$.

It is simplest to consider the case where magnetization is uniform throughout the sample. Neglecting relaxation effects, the equation of motion is still Eq. (2), but an effective field is substituted for $\mathbf{H}$, consisting of the applied static and alternating fields, the demagnetizing corrections (from the electron magnetic dipolar fields), and the effects of crystalline anisotropy. Because the Weiss molecular field is always parallel to $\mathbf{M}$, $\mathbf{H}_w$ exerts no torque and plays no role as long as the magnetization is uniform throughout the sample. (The exchange energy between spins does not change as long as their relative orientation does not change.)

The crystalline anisotropy can be shown to be equivalent to a magnetic field $\mathbf{H}_A$ along the direction of easy magnetization as long as $\mathbf{M}$ points nearly in that direction.

Because of the demagnetizing effects, the resonant frequency depends on sample geometry. For an infinite plane perpendicular to the applied field, the resonant angular frequency ω is given by $\omega = \gamma\sqrt{BH}$, but for a sphere, $\omega = \gamma H$.

The large demagnetizing and exchange fields are the principal difference between ferromagnetic and paramagnetic resonance. The demagnetizing field has components $-N_x M_x$, $-N_y M_y$, and $-N_z M_z$, where N_x, N_y, and N_z are the demagnetizing coefficients. Thus, suppose the static field and $\mathbf{M}$ lie along the z direction. Application of the alternating field tilts $\mathbf{M}$ away from z, changing the effective z field. If the change brings the spins closer to resonance, $\mathbf{M}$ may tilt out more. It is possible for such a nonlinear effect to be unstable for sufficiently large alternating fields. This instability is utilized in the Suhl ferromagnetic amplifier.

Antiferromagnetic resonance. The two sublattices of spins in an antiferromagnet are strongly coupled together by exchange forces. If both magnetizations ($\mathbf{M}_1$ and $\mathbf{M}_2$) are tilted together, away from the normal direction of magnetization in the crystal (call this the z direction), the only change in energy results from the anisotropy fields. However, because the anisotropy fields are reversed in direction for the two lattices, the magnetizations $\mathbf{M}_1$ and $\mathbf{M}_2$ tend to precess in opposite directions, bringing about a change in the exchange energy. An external field along the z direction aids one anisotropy field but opposes the other. The resonant angular frequency ω for a sphere is given by Eq. (6), where H is the applied field, H_A

$$\omega = \gamma\left[H \pm \sqrt{H_A(H_A + 2H_E)}\right] \tag{6}$$

the equivalent anisotropy field, and H_E the equivalent exchange field. The plus and minus signs refer to two opposite directions of rotating magnetic fields which may be used to observe the resonance. If H_E is 10^6 oersteds, and H_A is 10^4 oersteds, the corresponding frequency is 3×10^{11} Hz.

Ferrimagnetic resonance. Magnetic resonance in ferrites is called ferrimagnetic resonance. Ferrites are the natural generalization of antiferromagnets, containing two or more sublattices which may differ in magnetization. The basic coupling terms are still anisotropy fields, exchange fields, and the applied fields. The resonant angular frequency ω for the case of two sublattices is given by Eq. (7), where η is

$$\omega = \gamma\left[H - \frac{\eta H_E}{2} \pm \sqrt{\left(\frac{\eta H_E}{2}\right)^2 + H_E H_A(2-\eta) + H_A{}^2}\right] \tag{7}$$

a parameter measuring the relative sizes of the two magnetization vectors. Taking $\mathbf{M}_1$ to be the smaller magnetization, η is defined by Eq. (8). This equation

$$\mathbf{M}_1 = (1-\eta)\mathbf{M}_2 \tag{8}$$

assumes the magnetizations to be at saturation and the two sublattices to have the same γ (deviations might differ from spin-orbit coupling). *See* MOLECULAR BEAMS.

Charles P. Slichter

Bibliography. R. J. Abraham, J. P. Fisher, and P. Loftus, *Introduction to NMR Spectroscopy*, 1992; H. Friebolin, *Basic One- and Two-Dimensional NMR Spectroscopy*, 2d ed., 1993; J. W. Hennel and J. Klinowski, *Fundamentals of Nuclear Magnetic Resonance*, 1993; C. P. Poole and H. A. Farach, *Theory of Magnetic Resonance*, 2d ed., 1987; C. P. Slichter, The golden anniversary of nuclear magnetic resonance, NMR—Fifty years of surprises, *Proc. Amer. Phil. Soc.*, 142:533–556, 1998; C. P. Slichter, *Principles of Magnetic Resonance*, 3d ed., 1990, reprint 1992; W. S. Warren (ed.), *Advances in Magnetic Resonance*, vols. 1–14, 1965–1990.

Magnetic reversals

The Earth's magnetic field has reversed polarity hundreds of times. That is, at different times in Earth's past, a compass would have pointed south instead of north. Recognition that the geomagnetic field has repeatedly reversed polarity played a key role in the revolution that transformed the geological sciences in the 1960s—the acceptance of the theory of plate tectonics. It is generally accepted that the geomagnetic field is generated by motion of electrically conducting molten metal in Earth's outer core. However, the mechanism by which the field decays and reverses polarity remains one of the great unknowns in geophysics. *See* GEOMAGNETISM; PLATE TECTONICS.

The last magnetic field reversal occurred long before humans were aware of the geomagnetic field

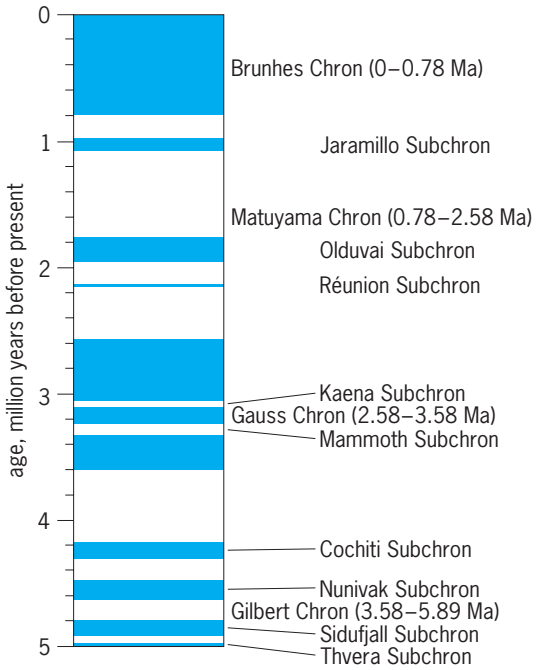

Fig. 1. Time scale of reversals of Earth's magnetic field for the last 5 million years. Color intervals represent periods when Earth's magnetic field was in the same state as the present field (called the normal polarity state) when a compass would point north. White intervals represent periods when Earth's magnetic field was in the opposite state as the present field (called the reversed polarity state) when a compass would point south. The boundaries between these polarity intervals represent short periods of geological time. Long intervals of nearly constant polarity are termed chrons, while shorter periods of opposite polarity within a chron are termed subchrons. The last four chrons are named after famous contributors to geomagnetism (Brunhes, Matuyama, Gauss, and Gilbert), while the subchrons in these intervals are named after the locations at which they were first discovered. (*From S. C. Cande and D. V. Kent, Revised calibration of the geomagnetic polarity time scale for the Late Cretaceous and Cenozoic, J. Geophys. Res., 100:6093–6095, 1995*)

(780,000 years ago; **Fig. 1**), so it is necessary to study geological records to understand the process by which the field reverses. The ability of rocks to act as fossilized compasses, which record a permanent "memory" of Earth's magnetic field at the time of rock formation, makes them suitable for detailed studies of ancient geomagnetic field behavior. *See* PALEOMAGNETISM; ROCK MAGNETISM.

In order to understand the mechanism of field reversal, it is necessary to develop a realistic view of the geometry of the field during polarity transitions. The interval of time required for the field to switch polarity is between about 5000 and 10,000 years. In a typical sedimentary rock unit, a polarity transition is recorded across a thin interval of strata, generally less than a meter thick. Many detailed studies have been undertaken from closely spaced samples through such intervals. During the early 1990s, a compilation of existing records indicated a distinct pattern in the way the field behaves during reversals. In particular, if the position of the magnetic pole was plotted as it progressed through a reversal (**Fig. 2**), it seemed that a large proportion of the poles tracked across a longitudinal band through the Americas. A second set of poles followed a path on the other side of the Earth (180° from the first) through western Australia and east Asia. This compilation suggested that the transitional field was dominantly dipolar and that biased field configurations may have persisted for several reversals during the last 10 million years. The time constants of convective motion in Earth's fluid outer core are a great deal shorter than in the solid overlying mantle. The persistence of these patterns over such long time intervals therefore suggested that fluid motion in the core is strongly affected by Earth's mantle and that the proposed bias may reflect long-term characteristics of deep-Earth dynamics. The coincidence in the location of transitional geomagnetic poles and other geophysical observations, such as anomalously cool lower-mantle temperatures, fluid flow patterns at the top of Earth's outer core, and significant features in the present geomagnetic field, excited a great deal of interest. As a result, the nature of transitional geomagnetic field behavior became one of the most hotly debated issues in geophysics in the 1990s. *See* EARTH, CONVECTION IN.

When observations concerning transitional geomagnetic field behavior were subjected to more rigorous statistical analysis, it was concluded that it is premature to accept the hypothesis of lower-mantle control of the geodynamo. Despite the compilation of many polarity transition records, there is an uneven distribution of sites around the world from which polarity transition records are available. A much wider distribution of sites is required for adequate testing of current hypotheses concerning

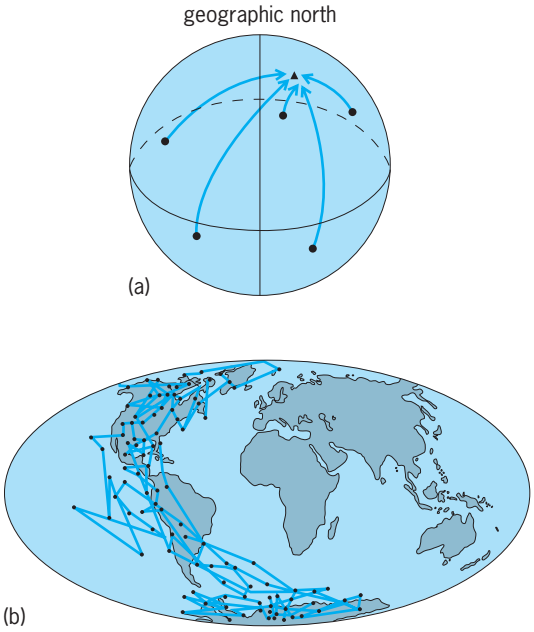

Fig. 2. Position of the magnetic pole plotted as it progresses through a reversal. (*a*) In a purely dipole field, such as that produced by a bar magnet, paleomagnetic data from any site on Earth's surface will describe a single pole (Fuller et al., 1996). (*b*) An example of a series of poles determined from a detailed paleomagnetic study of a polarity transition (Tric et al., 1991). When a number of such records were plotted together, the majority of poles plotted over a longitudinal band across the Americas (Laj et al., 1991), which indicated that the transitional field was dominantly dipolar.

transitional field behavior. In addition, there is considerable concern about the ability of some geological materials, particularly sediments, to record the geomagnetic field during periods when the field is extremely weak (such as during a reversal when the field decays to about one-tenth of its nontransitional value). Thus, rigorous methods must also be developed for assessing the reliability of the records used to make inferences about transitional field behavior. *See* GEODYNAMO.

Major advances in understanding geomagnetic field behavior have recently been made using numerical models that require thousands of hours of super-computer processing. The models provide fairly realistic simulations for how the geomagnetic field is generated, and a polarity reversal has been simulated. This simulation displays many features in common with paleomagnetic records of reversals, but much more work is needed to get a more realistic match. Paleomagnetic observations of polarity transitions are still required from widely distributed sites around the world. Debate concerning geomagnetic field behavior during polarity transitions is therefore far from settled and is likely to rage for some time to come. Andrew P. Roberts

Bibliography. S. C. Cande and D. V. Kent, Revised calibration of the geomagnetic polarity time scale for the Late Cretaceous and Cenozoic, *J. Geophys. Res.*, 100:6093–6095, 1995; M. Fuller, C. Laj, and E. Herrero-Bervera, The reversal of the Earth's magnetic field, *Amer. Sci.*, 84:552–561, 1996; G. A. Glatzmaier and P. H. Roberts, A three-dimensional self-consistent computer simulation of a geomagnetic field reversal, *Nature*, 377:203–209, 1995; J. A. Jacobs, *Reversals of the Earth's Magnetic Field*, Cambridge University Press, 1994; C. Laj et al., Geomagnetic reversal paths, *Nature*, 351:447, 1991; P. L. McFadden, C. E. Barton, and R. T. Merrill, Do virtual geomagnetic poles follow preferred paths during geomagnetic reversals?, *Nature*, 361:342–344, 1993; A. P. Roberts, Polarity transitions and excursions of the geomagnetic field, *Rev. Geophys. Suppl.*, 35:153–160, 1995; E. Tric et al., High-resolution record of the Upper Olduvai transition from Po Valley (Italy) sediments: Support for dipolar transition geometry?, *Phys. Earth Planet. Inter.*, 65:319–336, 1991.

Magnetic separation methods

All materials possess magnetic properties. Substances that have a greater permeability than air are classified as paramagnetic; those with a lower permeability are called diamagnetic. Paramagnetic materials are attracted to a magnet; diamagnetic substances are repelled. Very strongly paramagnetic materials are classified as ferromagnetic and include such metals as iron, nickel, and cobalt, and such minerals as magnetite, pyrrhotite, and ilmenite. Such substances can be separated from weakly or nonmagnetic materials by the use of low-intensity magnetic separators. Minerals such as hematite, limonite, and garnet are weakly magnetic and can be separated from nonmagnetics by the use of high-intensity separators.

Magnetic separators are widely used to remove tramp iron from ores being crushed, thereby protecting the crushers; to remove contaminating magnetics from food and industrial products; to recover magnetite and ferrosilicon in the float-sink methods of ore concentration; and to upgrade or concentrate ores. Tramp iron may be removed by a stationary magnet suspended over a conveyor belt or the material may be passed over a magnetic pulley; in both cases, magnetics are attracted by the magnetic field and separated.

Magnetic separators are extensively used to concentrate ores, particularly iron ores, when one of the principal constituents is magnetic. When the chief economic mineral is magnetite, iron ores can be cheaply and effectively separated by low-intensity separators. Such separators may be dry or wet. If an ore can be crushed to give substantial liberation of minerals at sizes coarser than $^1/_4$ in. (0.6 cm), separations can be made on the $+^1/_4$-in. (0.6-cm) sizes on a belt-type magnetic cobber (see **illus.**). Such a machine is usually used to cob out or reject waste materials. If clean mineral occurs at this size, it can be recovered by splitting the discharge or by retreatment on a lower strength machine. Wet magnetic separators are usually used to treat ore finer than $^1/_8$ in. (0.3 cm). These separators may be of the belt type or of the more common rotating-drum type. Drum-type separators consist of one or more rotatable drums having inner nonrotatable magnet elements with 3–7 poles. The magnets may be either electromagnets or permanent magnets. After the feed enters the machine as a slurry, the magnetics are attracted to the pole pieces and are carried to a discharge point on the surface of the drum. Many types of box designs are in use. The concurrent type is frequently used on relatively coarse material to reject a clean waste product. The countercurrent type is employed on fine ore to give a clean concentrate. Magnets may be of either the electromagnetic or permanent type. Electromagnets were formerly used almost exclusively but are now used mainly when exceptionally high

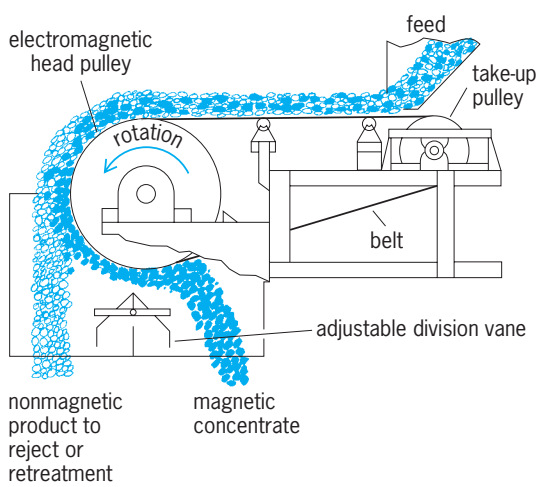

Belt-type magnetic cobber. (*Stearns Magnetic Products*)

field strengths are required or when it is desirable to vary magnet strength. Permanent magnets are now widely used since modern materials permit charging and retaining high field strengths permanently. Most permanent magnets are of the alnico type but the ceramic types containing barium ferrites are coming into use. Several types of magnetic separators employing alternating current have been devised but have found little commercial use.

High-intensity separators for the separation of weakly magnetic minerals are usually of the dry type. Surface-tension effects usually rule out wet separations. Since magnetic attraction varies inversely as the square of the distance, weakly magnetic minerals must be brought close to the magnets if they are to be separated. Both belt-type and induced-roll machines are used. The ore must be fully dried and closely sized for best results. *See* CHEMICAL SEPARATION TECHNIQUES; MAGNETISM; MECHANICAL SEPARATION TECHNIQUES; ORE DRESSING. Fred D. DeVaney

Bibliography. A. Grandison and M. Lewis, *Separation Processes*, 1993; C. J. King, *Separation Processes*, 2d ed., 1980; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; J. A. Wesselingh and R. Krishna, *Mass Transfer*, 1991.

Magnetic susceptibility

The magnetization of a material per unit applied field. It describes the magnetic response of a substance to an applied magnetic field. If M is the magnetization and H the applied magnetic field, then the magnetic susceptibility, denoted by χ, is given by Eq. (1). In the case that M is not parallel to H,

$$\chi = \frac{M}{H} \tag{1}$$

χ is a tensor. Otherwise, it is a simple number. For a crystalline material, χ may depend upon the direction of H with respect to the axes of the crystal because of anisotropic effects. For an elementary discussion of M and H *see* MAGNETISM; MAGNETIZATION.

The magnetic susceptibility is expressed in a variety of ways: per gram, per atom, per unit volume, and per mole. In this article, electromagnetic units are used. In Eq. (1), the units of χ are ergs per oersted per unit volume. **Figure 1** shows the atomic susceptibilities χ_A (ergs per oersted per atom) of the elements. The static susceptibility is measured in constant applied magnetic fields. The frequency-dependent susceptibility is measured in alternating magnetic fields. It is usually a complex quantity in which both the real and imaginary parts are functions of frequency.

Ferromagnetic susceptibility. The general behavior of the susceptibility of ferromagnetic materials above the Curie temperature follows the Curie-Weiss law, Eq. (2). This behavior is followed in the region well

$$\chi = \frac{C}{T - \theta} \tag{2}$$

above the ferromagnetic Curie temperature T_c. The paramagnetic Curie temperature θ is usually slightly greater than the temperature of transition T_c. Comparison of θ and T_c for three ferromagnetic metals is given in the **table**.

Comparison of θ and T_c

Parameter	Fe	Co	Ni
θ, K	1093	1428	650
T, K	1043	1393	631

All ferromagnetic materials exhibit paramagnetic behavior above their ferromagnetic Curie points. The magnitude of the paramagnetic susceptibility is determined by the Curie constant C, as in Eq. (2). A typical value of C is 0.2 K/cm^3 for iron. For the theory of ferromagnetic susceptibility *see* CURIE TEMPERATURE; CURIE-WEISS LAW; FERROMAGNETISM.

In the region just a fraction of a degree above the "critical point," or Curie temperature T_c, the susceptibility is found to approximate Eq. (3),with γ generally very close to 1.33. The theory is extremely complicated and not entirely satisfactory.

$$\chi = \frac{C'}{(T - T_c)^{\gamma}} \tag{3}$$

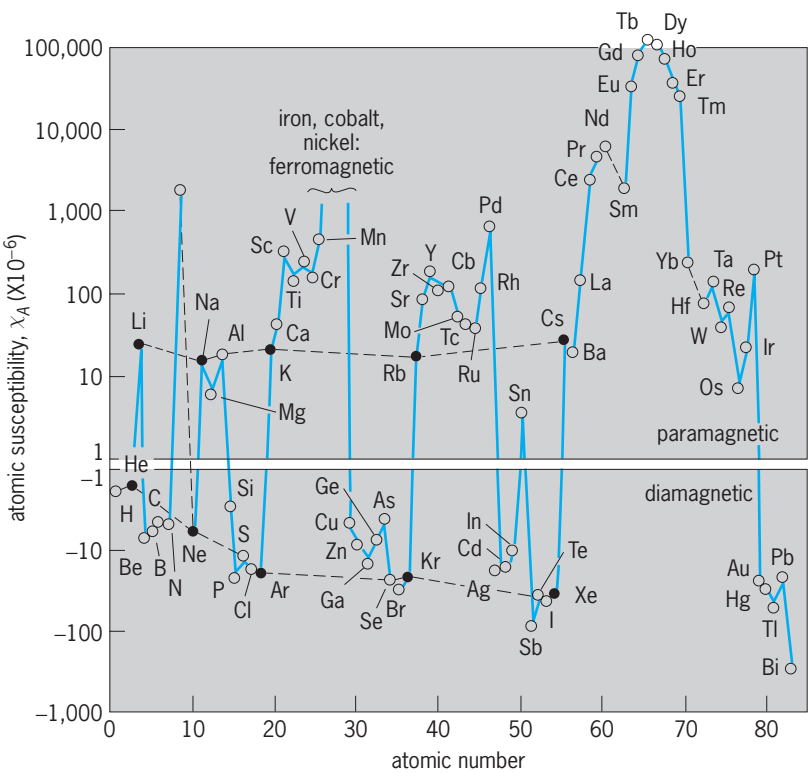

Fig. 1. Atomic susceptibilities of the elements at room temperature. The broken lines connect the alkali metals (paramagnetic) and the rare gases (diamagnetic).

Below the Curie point, the static susceptibility is not usually defined for a ferromagnetic substance, since the ferromagnet may have a finite magnetization in zero applied field.

The initial permeability is the slope of the magnetization curve (B plotted against H) for magnetic field strength $H = 0$. Initial permeabilities vary from almost 300 (platinum-cobalt) to 100,000 (supermalloy).

The frequency-dependent permeability is that measured in an alternating magnetic field. The experiments are usually carried out in small magnetic fields so that it is the frequency dependence of the initial permeability which is measured. Such experiments give information on the structure of ferromagnetic domains and the motion of domain walls.

Paramagnetic susceptibility. Most paramagnetic substances at room temperature have a static susceptibility which follows a Langevin-Debye law, Eq. (4),

$$\chi = \frac{Np^2\mu_B^2}{3kT} + N\alpha \tag{4}$$

where N is the number of magnetic dipoles per unit volume, p is the effective magneton number, μ_B is the Bohr magneton, k is Boltzmann's constant, T is the absolute temperature, and α is the temperature-independent contribution of Van Vleck paramagnetism. The first term of Eq. (4) is referred to as the Curie law because of its $1/T$ dependence. Experimental results are often expressed in terms of what the effective magneton numbers must be in order to account for the variation of χ with $1/T$. The interpretation of the experimental data reveals the nature of the energy levels of the paramagnetic ions, the symmetry of paramagnetic crystals, the effects of crystalline electric fields on the energy levels, and the influence of the paramagnetic ions on one another.

Saturation of the paramagnetic susceptibility occurs when a further increase of the applied magnetic field fails to increase the magnetization, because practically all the magnetic dipoles are already oriented parallel to the field. Saturation will occur either at very strong fields or at very low temperatures; that is, in cases when the approximation leading to the Langevin-Debye formula fails. Some paramagnetic solids have susceptibilities which follow a Curie-Weiss law rather than a Curie law. *See* PARAMAGNETISM.

Diamagnetic susceptibility. The susceptibility of diamagnetic materials is negative, since a diamagnetic substance is magnetized in a direction opposite to that of the applied magnetic field. The diamagnetic susceptibility is independent of temperature.

Diamagnetic susceptibility depends upon the distribution of electronic charge in an atom and upon the energy levels. Interpretation of the experimental data reveals the nature of the atomic wave functions of the atom in question. For the theory of diamagnetic susceptibility and a listing of values of the molar diamagnetic susceptibilities of several rare gases and ions in crystals *see* DIAMAGNETISM.

Antiferromagnetic susceptibility. The susceptibility of antiferromagnetic materials above the Néel point, which marks the transition from antiferromagnetic to paramagnetic behavior, follows a Curie-Weiss law with a negative paramagnetic Curie temperature $-\theta$, as in Eq. (5). The Néel temperature is always somewhat less than θ.

$$\chi = \frac{C}{T + \theta} \tag{5}$$

The antiferromagnetic susceptibility is a maximum at the Néel temperature. Below the Néel temperature, it behaves in a way determined by the angle between the field direction and the crystal axes—it usually decreases as the temperature lowers. For the theory of antiferromagnetic susceptibility *see* ANTIFERROMAGNETISM.

Measurement of susceptibility. Magnetic susceptibilities may be measured by several methods, depending upon the quantity sought. Ferromagnetic static permeabilities may be measured by the Rowland ballistic method. In this technique, the specimen is cut in the shape of a ring, and a magnetic field is applied by means of a primary winding on the ring. A secondary winding is connected to a ballistic galvanometer. The current through the primary is changed suddenly. The resulting abrupt change in magnetic field H causes a change in magnetic induction B ($B = H + 4\pi M$) in the specimen, which in turn induces a voltage in the secondary. Thus, the galvanometer suffers a deflection proportional to the change in B. One starts with a known value of B (usually zero) and plots the magnetization curve in this way.

Paramagnetic and diamagnetic susceptibilities may be measured by the balance method. A magnetized substance in an inhomogeneous magnetic field experiences a force given by Eq. (6).

$$F = {}^1/_2 \operatorname{grad} \int M \cdot H \, dV \tag{6}$$

The integral is over the volume of the specimen. The susceptibility is $\chi = M/H$. Therefore Eq. (7) holds.

$$F = {}^1/_2 \chi \operatorname{grad} \int H^2 \, dV \tag{7}$$

If the magnetic field now varies only in one direction, say the z direction, then Eq. (8) is valid,

$$F = {}^1/_2 \chi A \int \frac{d}{dz}(H^2)\, dz = {}^1/_2 \chi A (H_1^2 - H_2^2) \tag{8}$$

where the specimen is in the shape of a rod of cross-sectional area A and is suspended in the z direction in the inhomogeneous field (**Fig. 2**). In Fig. 2, H_1 and H_2 are the values of the magnetic field strength at the two ends of the rod. The force is measured and, when H_1 and H_2 (or H and dH/dz) are known, the susceptibility can be determined. This is known as the Gouy balance method. Note that the total magnetic susceptibility is measured, both para- and diamagnetic contributions.

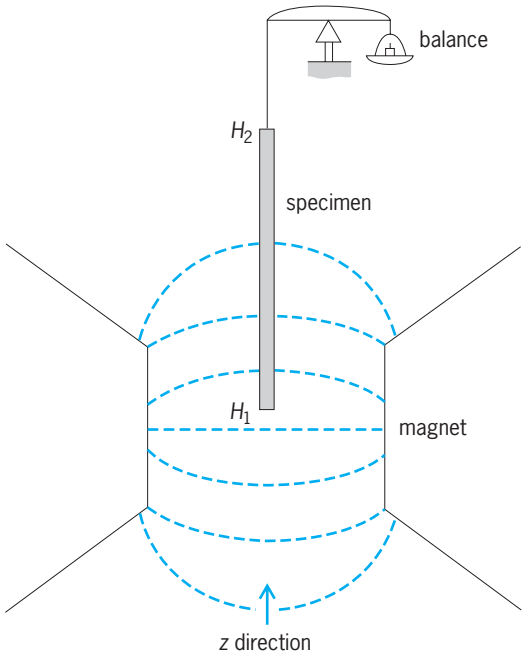

Fig. 2. Gouy balance method for measurement of magnetic susceptibilities. (*After C. Kittel, Introduction to Solid State Physics, 3d ed., John Wiley and Sons, 1966*)

The paramagnetic susceptibility may sometimes be measured separately by spin-resonance techniques. This is especially true of the Pauli paramagnetic susceptibility of conduction electrons in alkali metals. *See* MAGNETIC RESONANCE.

At low temperatures, relative measurements of paramagnetic susceptibility are very simple in comparison with absolute measurements and yet can provide much of the ionic and crystal-field information referred to above. Such measurements are most commonly carried out using an audio-frequency mutual-inductance technique. Exactly the same procedure can be employed for another important purpose—namely, low-temperature thermometry—making use of the Curie law (with or without refinements). *See* LOW-TEMPERATURE THERMOMETRY; MAGNETIC THERMOMETER; SQUID. Elihu Abrahams; Frederic Keffer

Bibliography. R. A. Hein, T. L. Francavilla, and D.H. Liebenberg (eds.), *Magnetic Susceptibility of Superconductors and Other Spin Systems*, 1992; J. R. Hook and H. E. Hall, *Solid State Physics*, 2d ed., 1995; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996.

Magnetic thermometer

A thermometer whose operation is based on Curie's law, which states that the magnetic susceptibility of noninteracting (that is, paramagnetic) dipole moments is inversely proportional to absolute temperature. Magnetic thermometers are typically used at temperatures below 1 K (−458°F). The magnetic moments in the thermometric material may be of either electronic or nuclear origin. Generally the magnetic thermometer must be calibrated at one (or more) reference temperatures. *See* ELECTRON; NUCLEAR MOMENTS; PARAMAGNETISM.

Principle. In a magnetic field H, paramagnetic moments are polarized so that the vectorial sum of the individual dipole moments μ produces a magnetization M. At absolute temperatures T that are relatively high compared to the magnetic energy of the dipole (that is, at temperatures at which k_BT is much greater than μH, where k_B is Boltzmann's constant), the magnetization obeys the Curie-Weiss law, Eq. (1). Here χ is the magnetic susceptibility of the

$$M = \chi H = \frac{C}{T - \Delta} H \qquad (1)$$

thermometric material, C its Curie constant, and Δ a small correction which becomes important only at the low-temperature end of the thermometer's working range. *See* CURIE-WEISS LAW.

For an actual device, C in Eq. (1) becomes the so-called thermometer constant C', comprising the Curie constant multiplied by detection-sensitivity parameters. By measuring the magnetization M in a constant polarizing field H, the absolute temperature T can be determined if the thermometer constant C is known from a calibration with some other thermometer.

Toward high temperatures, the useful temperature range is limited by the decreasing sensitivity (since the increment in M for a unit change in T is proportional to $1/T^2$). Toward low temperatures, the measurement sensitivity increases, but at the lowest temperatures the range is restricted, first by the need to calibrate the correction term Δ, and finally by the breakdown of the Curie-Weiss behavior. Even at the expense of a reduced signal amplitude, it is therefore preferable to use both a small polarizing field H and a low measuring frequency, well below the equivalent of the spin-lattice relaxation time (to avoid loss of thermal contact between the magnetic moments and their environment in the crystalline lattice). *See* MAGNETIC SUSCEPTIBILITY; MAGNETIZATION.

Electronic magnetic thermometer. At temperatures from a few millikelvins upward, the thermometric material is preferably an electronic paramagnet, typically a nonconducting hydrous rare-earth salt. For higher temperatures, an ion is selected with a large magnetic moment in a crystalline environment with a high density of magnetic ions. In contrast, for low temperature use the magnetic exchange interactions between the magnetic ions should be small, which is accomplished by selecting an ion with a well-localized moment and by maintaining a large separation between the magnetic ions by means of diamagnetic atoms. This is the case in cerium magnesium nitrate (CMN) $[2Ce(NO_3)_3 \cdot 3Mg(NO_3)_3 \cdot 24H_2O]$. Here, the Ce^{3+} ion is responsible for the magnetic moment, which is well localized within the incompletely filled $4f$ shell relatively deep below the outer valence electrons. To reduce the magnetic interactions between the Ce^{3+} ions further, Ce^{3+} may be partly substituted with diamagnetic La^{3+} ions. Lanthanum-diluted CMN has been used for

thermometry to below 1 mK. *See* EXCHANGE INTERACTION.

A mutual-inductance bridge, originally known as the Hartshorn bridge, has been the most widely employed measuring circuit for precision thermometry (**Fig. 1**). The bridge is driven by a low-frequency alternating-current source. The inductance at low temperatures consists of two coils, which are as identical as possible. The voltages induced across them by the drive current are compared by means of a high-input-impedance ratio transformer. The output level of this voltage divider is adjusted to equal that of the midpoint between the two coils, using as null indicator a narrow-band preamplifier and a phase-sensitive (lock-in) detector. Thus, without a paramagnetic specimen, the bridge is balanced with the decade divider adjusted at its midpoint, while with the specimen inside one of the coils the change in the divider reading at bridge balance is proportional to the sample magnetization. For high-resolution thermometry it has become standard practice to replace the room-temperature zero detector with a SQUID magnetometer circuit. This also allows the mass of the sample to be reduced from several grams to the 1-mg level. *See* INDUCTANCE MEASUREMENT; SQUID.

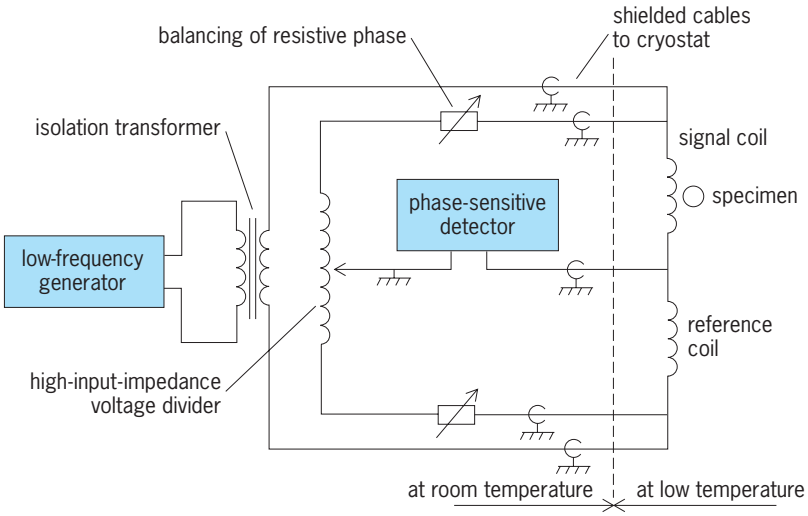

Fig. 1. Low-frequency (10–10^4 Hz) inductance bridge for the measurement of the susceptibility of an electronic paramagnet. The two identical solenoidal coils with three shielded cables are located at low temperatures, with the thermometric specimen in good thermal contact with the body whose temperature is measured.

Nuclear magnetic thermometer. Nuclear magnetic moments are smaller by a factor of 10^3 and are used for thermometry only in the ultralow-temperature region. For this the Curie-law behavior with $\Delta = 0$ in Eq. (1) is generally sufficient down to the lowest temperatures. The nuclear paramagnetic thermometer loses adequate sensitivity for calibration purposes above 50–100 millikelvins, unless it is operated in a high polarizing field (H greater than 0.1 tesla). It can be utilized as a self-calibrating primary thermometer if the spin-lattice relaxation time is measured in parallel with the nuclear Curie susceptibility. In metals where the temperature-independent Pauli paramagnetism of the conduction electrons is predominant and which are not in the superconducting state, the nuclear spin-lattice relaxation time τ_1 is controlled by the conduction electrons according to Korringa's relation, Eq. (2). Here, κ is the Korringa constant,

$$\tau_1 T = \kappa \tag{2}$$

which for platinum is 30 ms·K and for copper 1.2 s·K. The thermometric specimen is in metallic form in order to secure fast spin-lattice relaxation and adequate thermal homogeneity through electronic

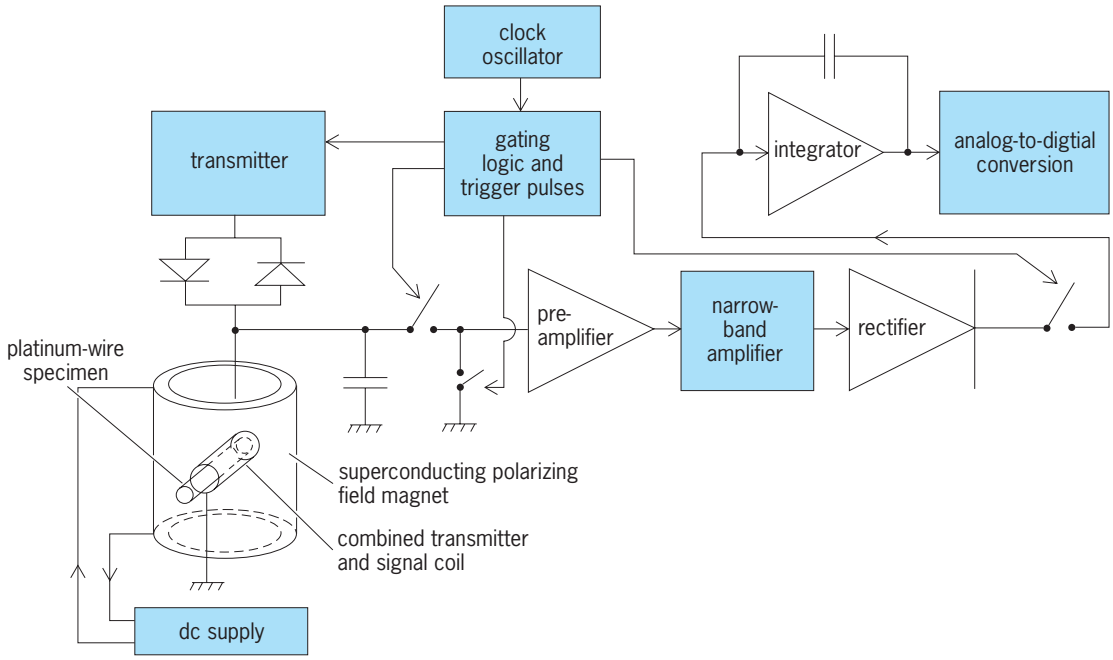

Fig. 2. Pulsed nuclear magnetic resonance (NMR) thermometer for the measurement of the nuclear paramagnetic susceptibility and spin-lattice relaxation time of a metallic wire or powder sample at temperatures below 100 mK and magnetic fields less than 50 millitesla. (*PicoWatt, Inc., Finland*)

thermal conduction. Extreme purity with respect to electronic magnetic impurities is important. With nuclear magnetic resonance (NMR) techniques the measurement can be restricted to probe only the magnetism of one nuclear spin species in the specimen. Pulsed NMR measurement on the ^{195}Pt isotope in natural platinum metal provides presently the most widely used thermometry at temperatures below 1 mK. In the Curie-susceptibility measuring mode, it has been extended down to 10 μK. *See* MAGNETIC RELAXATION.

NMR measurement is selectively sensitive only to those paramagnetic nuclei for which the Larmor frequency in the externally applied polarizing field H coincides with the frequency of the rf excitation field (oriented transverse to the steady field H), which is also the frequency of the tuned amplifier chain (**Fig. 2**). This property is important for reducing the influence from electronic-impurity magnetism. Measurement in the pulsed mode reduces problems associated with heating from the resonance excitation field and makes it possible to measure the spin-lattice relaxation time. The rectified and integrated value of the NMR free-induction-decay (FID) signal is usually treated without any corrections as inversely proportional to absolute temperature. *See* LOW-TEMPERATURE THERMOMETRY; NUCLEAR MAGNETIC RESONANCE (NMR).

Matti Krusius

Bibliography. D. S. Betts, *Refrigeration and Thermometry Below 1 K*, 1976; R. P. Hudson, *Principles and Applications of Magnetic Cooling*, 1972; O. V. Lounasmaa, *Experimental Principles and Methods Below 1 K*, 1974; F. Pobell, *Matter and Methods at Low Temperatures*, 2d ed., 1996; R. C. Richardson and E. N. Smith, *Experimental Techniques in Condensed Matter Physics at Low Temperatures*, 1988.

Magnetic thin films

Sheets of magnetic material with thicknesses of a few micrometers or less, used in the electronics industry. Magnetic films can be single-crystal, polycrystalline, amorphous, or multilayered in the arrangement of their atoms. Applications include magnetooptic storage, inductive recording media, magnetoresist sensors, and thin-film heads. *See* COMPUTER STORAGE TECHNOLOGY; MAGNETIC RECORDING; MAGNETOOPTICS; MAGNETORESISTANCE.

Fabrication. Thin films are usually directly deposited on a substrate. The techniques of deposition vary from high-temperature liquid-phase epitaxy such as that used in the fabrication of garnet films, through electroplating, to computer-controlled vapor deposition and sputtering. Permalloy, for example, can be both electroplated and sputtered or vapor-deposited. Magnetooptic films for storage applications are usually sputtered. *See* CRYSTAL GROWTH; ELECTROPLATING OF METALS; SPUTTERING.

Magnetic order. Both ferro- and ferrimagnetic films are used. The ferromagnetic films are usually transition-metal-based alloys. For example, permalloy is a nickel-iron alloy. The ferrimagnetic films, such as garnets or the amorphous films, contain transition metals such as iron or cobalt and rare earths. The ferrimagnetic properties are advantageous in magnetooptic applications where a low overall magnetic moment can be achieved without a significant change in the Curie temperature. *See* FERRIMAGNETISM; FERROMAGNETISM.

Magnetic anisotropy. In an isotropic bulk magnetic sample the overall magnetization does not prefer a particular direction. However, the same material in the form of a thin film has a lower energy when the magnetization lies in the plane of the film. This shape-induced anisotropy is useful in applications such as permalloy drive or sensor patterns used in magnetoresist-head technology. However, there are other situations where it is advantageous to have perpendicular rather than in-plane anisotropy. For example, magnetooptic memory materials have perpendicular anisotropy. This is obtained by using anisotropic crystals, by growing films in a magnetic field or directional deposition, and by annealing in a magnetic field or applying an anisotropic stress. A preferred direction of magnetic bias in multilayer films can also be achieved through exchange coupling.

Domain structure. Depending upon the nature of anisotropy (perpendicular to planar), a variety of domain patterns can be obtained. For example, when the perpendicular anisotropy is large relative to shape-induced anisotropy, stripe domains are observed. These form cylindrical domains if an external magnetic field is applied perpendicular to the plane of the films. Such cylindrical domains are used to represent bits of information in magnetooptical storage. The domain structure can be observed in an optical microscope with polarized light by using either the Faraday or Kerr effect. This effect is also used to sense a bit of information in magnetooptic storage. Domains can also be observed by using ferrofluids. Domains and domain wall structure are also observable by electron microscope techniques. *See* DOMAIN (ELECTRICITY AND MAGNETISM); FARADAY EFFECT; KERR EFFECT.

Coercivity. The coercivity of thin film varies with applications. For most applications a low coercive force is desired. Coercivity is generally associated with inhomogeneities such as stress fields, compositional fluctuations, defects, and surface roughness. *See* MAGNETISM; MAGNETIZATION.

Magnetotransport. The change in electrical properties, such as the electrical resistance, with a magnetic field is used in sensor elements. The most notable of these in semiconductor technology is the magnetoresist head used in disk storage technology. Very large magnetoresist signals (called giant magnetoresistance) are observed in magnetic multilayers and composites containing a magnetic and nonmagnetic material.

Praveen Chaudhari

Bibliography. J. A. Bland and H. Bretislav (eds.), *Ultrathin Magnetic Structures I: Introduction to Electronic, Magnetic, and Structural Properties*,

1994; P. Chaudhari and D. Turnbull, Structure and properties of metallic glasses, *Science*, 199:11–21, 1978; L. M. Falicov, Metallic magnetic superlattices, *Phys. Today*, 45(10):46–51, October 1992; I. S. Jacobs, Magnetic materials and applications: Quarter century overview, *J. Appl. Phys.*, 50:7294–7306, 1979; D. L. Smith, *Thin Film Deposition and Applications*, 1995.

Magnetism

The branch of science that describes the effects of the interactions between charges due to their motion and spin. These may appear in various forms, including electric currents and permanent magnets. The interactions are described in terms of the magnetic field, although the field hypothesis cannot be tested independently of the electrokinetic effects by which it is defined. The magnetic field complements the concept of the electrostatic field used to describe the potential energy between charges due to their relative positions. Special relativity theory relates the two, showing that magnetism is a relativistic modification of the electrostatic forces. The two together form the electromagnetic interactions which are propagated as electromagnetic waves, including light. They control the structure of materials at distances between the long-range gravitational actions and the short-range "strong" and "weak" forces most evident within the atomic nucleus. *See* ELECTROMAGNETIC RADIATION; RELATIVITY.

The term "magnetism" originates in the material magnetite, an iron ore, which produces weak natural magnets in the form of lodestones, exerting forces on each other and on pieces of iron. Peregrinus showed in 1269 that the behavior can be described in terms of magnetic poles on opposite end surfaces. The analogy between magnetic poles and electric charge greatly influenced the later development of the subject. The Earth provides an example of the subsequent explanation of magnetic behavior in terms of the flow of current and movement of charge, including the quantum state described as spin. This leaves open the question of whether or not isolated magnetic poles, or monopoles, exist as separate physical entities. *See* MAGNETIC MONOPOLES; MAGNETITE.

The magnetic field can be visualized as a set of lines (**Fig. 1**) illustrated by iron filings scattered on a suitable surface. The intensity of the field is indicated by the line spacing, and the direction by arrows pointing along the lines. The sign convention is chosen so that the Earth's magnetic field is directed from the north magnetic pole toward the south magnetic pole. The field can be defined and measured in various ways, including the forces on the equivalent magnetic poles, and on currents or moving charges. Bringing a coil of wire into the field, or removing it, induces an electromotive force (emf) which depends on the rate at which the number of field lines, referred to as lines of magnetic flux, linking the coil changes in time. This provides a definition of flux,

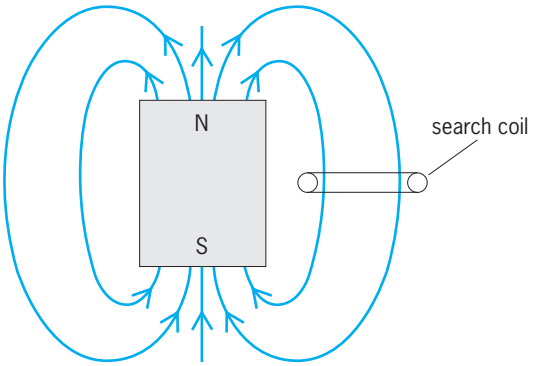

Fig. 1. Magnetic lines of a bar magnet.

Φ, in terms of the emf, e, given by Eq. (1) for a coil

$$e = -N\, d\Phi/dt \qquad \text{volts} \tag{1}$$

of N turns wound sufficiently closely to make the number of lines linking each the same. The International System (SI) unit of Φ, the weber (Wb), is defined accordingly as the volt-second. The symbol B is used to denote the flux, or line, density, as in Eq. (2), when the area of the coil is sufficiently small

$$B = \Phi/\text{area} \tag{2}$$

to sample conditions at a point, and the coil is oriented so that the induced emf is a maximum. The SI unit of B, the tesla (T), is the Wb/m^2. The sign of the emf, e, is measured positively in the direction of a right-hand screw pointing in the direction of the flux lines. It is often convenient, particularly when calculating induced emfs, to describe the field in terms of a magnetic vector potential function instead of flux.

Magnetic circuits. The magnetic circuit provides a useful method of analyzing devices with ferromagnetic parts, and introduces various quantities used in magnetism. It describes the use of ferromagnetic materials to control the flux paths in a manner analogous to the role of conductors in carrying currents around electrical circuits. For example, pieces of iron may be used to guide the flux which is produced by a magnet along a path which includes an air gap (**Fig. 2**), giving an increase in the flux density, B, if the cross-sectional area of the gap is less than that of the magnet. *See* MAGNET; MAGNETIC MATERIALS.

The magnet may be replaced by a coil of N turns carrying a current, i, wound over a piece of iron, or ferromagnetic material, in the form of a ring of uniform cross section (**Fig. 3**). The flux linking each turn of the coil, and each turn of a secondary coil wound

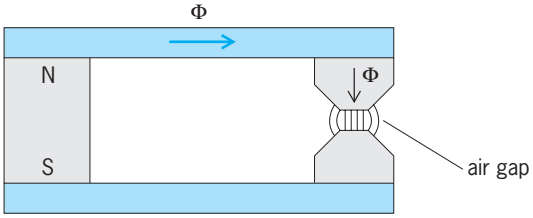

Fig. 2. Magnetic circuit with an air gap.

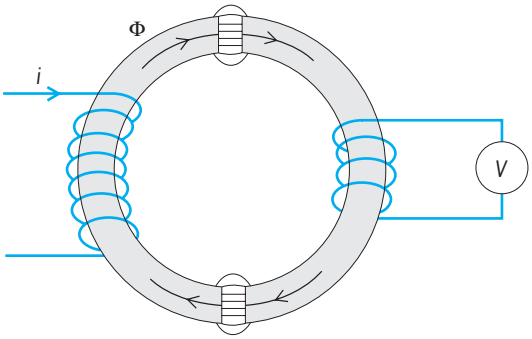

Fig. 3. Coils coupled by a ring of ferromagnetic material of uniform cross section.

separately from the first, is then approximately the same, giving the same induced emf per turn [according to Eq. (1)] when the supply current, i, and hence the flux, Φ, changes in time. The arrangement is typical of many different devices. It provides, for example, an electrical transformer whose input and output voltages are directly proportional to the numbers of turns in the windings. Emf's also appear within the iron, and tend to produce circulating currents and losses. These are commonly reduced by dividing the material into thin laminations. *See* EDDY CURRENT; TRANSFORMER.

The amount of flux produced by a given supply current is reduced by the presence of any air gaps which may be introduced to contribute constructional convenience or to allow a part to move. The effects of the gaps, and of different magnetic materials, can be predicted by utilizing the analogy between flux, Φ, and the flow of electric current through a circuit consisting of resistors connected in series (**Fig. 4**). Since Φ depends on the product, iN, of the winding current and number of turns, as in Eq. (3),

$$iN = \Phi\Re \tag{3}$$

the ratio between them, termed the reluctance, $\Re$, is the analog of electrical resistance. It may be constant, or may vary with Φ. The quantity iN is the magnetomotive force (mmf), analogous to voltage or emf in the equivalent electrical circuit. The relationship between the two exchanges the potental and flow quantities, since the magnetic mmf depends on current, i, and the electrical emf on $d\Phi/dt$. Electric and magnetic equivalent circuits are referred to as duals. *See* RELUCTANCE.

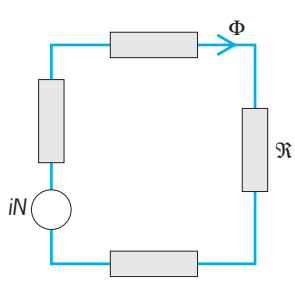

Fig. 4. Circuit analogy. Components of a magnetic circuit carrying a flux Φ, analogous to current i. The reluctances $\Re$ of the components are analogous to resistance, R.

Any part of the magnetic circuit of length l, in which the cross section, a, and flux density, B, are uniform has a reluctance given by Eq. (4). This equation parallels Eq. (5) for the resistance, R, of a con-

$$\Re = l/(a\mu) \tag{4}$$

$$R = l/(a\sigma) \tag{5}$$

duct or of the same dimensions. The permeability, μ, is the magnetic equivalent of the conductivity, σ, of the conducting material. Using a magnet as a flux source (Fig. 2) gives an mmf which varies with the air gap reluctance. In the absence of any magnetizable materials, as in the air gaps, the permeability is given by Eq. (6) in SI units (Wb/A-m). The quantity μ_0

$$\mu = \mu_0 = 4\pi \times 10^{-7} \tag{6}$$

is sometimes referred to as the permeability of free space. Material properties are described by the relative permeability, μ_r in accordance with Eq. (7), The

$$\mu = \mu_r\mu_0 \tag{7}$$

materials which are important in magnetic circuits are the ferromagnetics and ferrites characterized by large value of μ_r, sometimes in excess of 10,000 at low flux densities.

The various reluctances, $\Re_1, \Re_2, \ldots$, forming the different parts of the magnetic circuit can be added, yielding Eq. (8), when each part carries the same

$$\Phi = iN/(\Re_1 + \Re_2 + \Re_3 + \cdots) \tag{8}$$

flux, Φ. The largest reluctances are often those of the air gaps. In iron the nonlinear relationship between mmf and flux may produce large variations in $\Re$, and it is then convenient to specify B, and use the magnetization surve, relating B to the mmf per meter, to obtain the mmf, iN. Fringe fields in air may cause significant differences in fluxes in the different parts of the circuit, referred to as flux leakage, and these usually require corresponding adjustments to B. Greater accuracy is obtained by using numerical methods to compute the field.

The analogy between magnetic and electric circuits fails in that energy is dissipated in resistance but not in reluctance. Equations (1) and (3) show that a coil of N turns magnetizing a magnetic circuit of reluctance $\Re$ absorbs an energy given by Eq. (9)

$$ei\,dt = iN\,d\Phi = \Re\Phi\,d\Phi \qquad \text{joules} \tag{9}$$

in time dt, where $d\Phi$ is the flux change [Eq. (1)]. Summing shows that, if $\Re$ is constant, then the total energy required to increase the flux from zero to Φ is given by Eq. (10), and this energy, measured in

$$\int ei\,dt = \Re\Phi^2/2 \qquad \text{joules} \tag{10}$$

joules, is stored, not dissipated, since it is recovered when the flux is driven back to zero.

The storage of energy is a key property of a magnetic circuit or, more generally, of a magnetic field.

Since the flux $\delta\Phi$ in a volume element of cross-sectional area δa is $B\ \delta a$, and the reluctance $\delta\Re$ is proportional to the length δl [Eq. (4)], the energy required to magnetize the element is proportional to its volume. Substituting in Eq. (10) shows that the energy per unit volume is given by Eq. (11).

$$\mathfrak{E} = B^2/2\mu \qquad \mathrm{J/m^3} \tag{11}$$

The high permeability, μ, of the iron causes a high concentration of energy, $\mathfrak{E}$, in the air gaps of a magnetic circuit in which B is uniform. This is recovered if the iron surfaces are allowed to come together, while B is kept constant. Hence, the force per unit area is given by Eq. (12) when the flux density, B,

$$f = B^2/2\mu_0 \tag{12}$$

in the air is in the direction normal to the surface. The force of attraction between the two halves of the magnetic circuit in Fig. 3 is obtained by summing f over the areas of both air gaps. The attraction is important in many magnetic devices. An example is the use of electromagnets to lift steel scrap. *See* ELECTROMAGNET.

The highest fields obtainable in air, using superconducting coils, give B values in excess of 1 tesla, at which the energy density is approximately 4×10^5 J/m^3, or 0.4 J/cm^3. The recovery of the energy takes the form of an arc if any attempt is made to reduce the current to zero by opening a switch. Some arcing occurs across the contacts of any switch which energizes a magnetic circuit, or field, and protective measures may be necessary to prevent damage.

Comparing energy properties makes $\Re$ the magnetic equivalent of an electrical capacitor. The magnetic "flow" quantity, or "current," becomes $d\Phi/dt$ instead of Φ and describes the movement of poles, corresponding to the flow of charge in the electrical equivalent. Although this analogy is helpful, and more systematic (including aspects such as the modeling of eddy currents), the comparison of $\Re$ with electrical resistance is more usual.

Magnetic field strength. It is convenient to introduce two different measures of the magnetic field: the flux density, B, and the mmf per meter, referred to as the field strength, or field intensity, H. The field strength, H, provides a measure of the currents and other magnetic field sources, excluding those representing polarizable materials. It may also be defined in terms of the force on a unit pole.

A straight wire carrying a current I sets up a field (**Fig. 5**) whose intensity at a point at distance r is given by Eq. (13). The field strength, H, like B, is a

$$H = \frac{I}{2\pi r} \tag{13}$$

vector quantity pointing in the direction of rotation of a right-hand screw advancing in the direction of current flow. The intensity of the field is shown by the number of field lines intersecting a unit area. The straight wire provides one example of the circuital

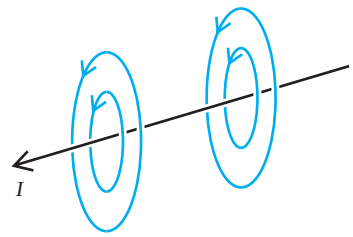

Fig. 5. Magnetic field of a straight wire.

law, known as Ampère's law, given in vector notation by Eq. (14). Here, θ is the angle between the vector

$$\oint H \cdot dl = \oint H \cos\theta\, dl = I \tag{14}$$

H and the element **dl** of any closed path of summation, or integration, and H and d denote magnitudes. I is the current which links this path. Choosing a circular path, centered on a straight wire, reduces the integral to $H\,(2\pi r)$. Equation (4) cannot be applied unless a path can be found along which H is uniform, or the way in which H varies with position is known.

A long, straight, uniformly wound coil (**Fig. 6**), for example, produces a field which is uniform in the interior and zero outside. The interior magnetic field, H, points in the direction parallel to the coil axis. Applying Eq. (14) to the rectangle *pqrs* of unit length in the axial direction shows that the only contribution is from *pq*, giving Eq. (15), where $\boldsymbol{n}$ is the number

$$H = In \tag{15}$$

of turns, per unit length, carrying the current, I. The magnetic field strength, H, remains the same, by definition, whether the interior of the coil is empty or is filled with ferromagnetic material of uniform properties. The interior forms part of a magnetic circuit in which In is the mmf per unit length. H is the analog of the electric field strength E, in the conductors of an electric circuit. The flux density, B, describes the effect of the field, in the sense of the voltage which is induced in a search coil by changes in time [Eq. (1)]. The ratio of H to B is the reluctance of a volume element of unit length and unit cross section in which

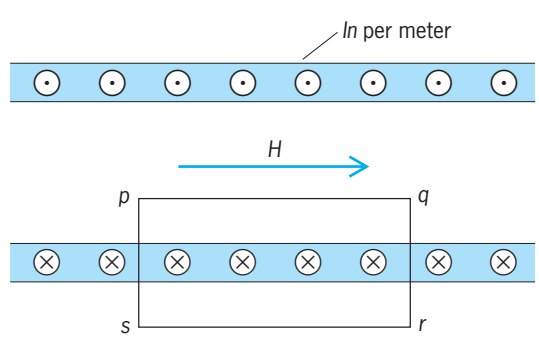

Fig. 6. Cross section of part of a long, straight uniformly wound coil. Black circles indicate current flowing out of the plane of the page, and crosses indicate flow in the opposite direction. Rectangle *pqrs* is used to calculate magnetic field strength, *H*, within the coil.

the field is uniform. Equation (4), shows that the two quantities are related by Eq. (16). The energy density in the field [Eq. (11)] is given by Eq. (17).

$$B = \mu H \tag{16}$$

$$\mathfrak{E} = HB/2 \qquad \text{J/m}^3 \tag{17}$$

The permeability, μ, is defined by Eq. (16). The relative permeability, μ_r, of polarizable materials [Eq. (7)] is measured accordingly by subjecting a sample to a uniform field inside a long coil such as that shown in Fig. 6 and using the emf induced in a search coil wound around the specimen to observe the flux in it.

The magnetic field due to current-carrying conductors in air may also be found by dividing the circuit into elements and summing the contributions from each. Provided that I is constant, or varies only slowly in time, the field, dH, at a point, P, due to a short length, dl, of a thin wire carrying a current, I, is given by the Biot-Savart law, Eq. (18), where θ is

$$dH = \left(\frac{I\,dl}{4\pi r^2}\right)\sin\theta \tag{18}$$

the angle between the axis of the wire, in the current flow direction, and the line of length r joining P to the current element. The field lines form circles centered on the current element axis, directed according to the right-handed screw rule. The field due to a point charge, q, moving at velocity u is likewise given by Eq. (19) at points close enough to make the

$$H = \left(\frac{qu}{4\pi r^2}\right)\sin\theta \tag{19}$$

propagation delay negligible.

Magnetic flux and flux density. Magnetic flux is defined in terms of the forces exerted by the magnetic field on electric charge. The forces can be described in terms of changes in flux with time [Eq. (1)], caused either by motion relative to the source or by changes in the source current, describing the effect of charge acceleration.

Since the magnetic, or electrokinetic, energy of current flowing in parallel wires depends on their spacing, the wires are subject to forces tending to change the configuration. The force, dF, on an element of wire carrying a current, i, is given by Eq. (20), and this provides a definition of the flux

$$dF = Bi\,dl \qquad \text{newtons} \tag{20}$$

density, B, due to the wires which exert the force. The quantity B is sometimes referred to as magnetic induction. The SI unit, called the tesla, or Wb/m^2, is the N/A-m. The flux density, B, equals $\mu_0 H$ in empty space, or in any material which is not magnetizable [Eq. (16)]. An example is the force which is exerted by a long straight wire on another which is parallel to it, at distance r. Equation (13) shows that the force, F, per meter length, between wires carrying currents I and i is given by Eq. (21). when the wires carry cur-

$$F = \frac{\mu_0 Ii}{2\pi r} = 2\times 10^{-7}\,Ii/r \qquad \text{newtons} \tag{21}$$

rents I and i. F is accounted for by the electrokinetic interactions between the conduction charges, and describes the relativistic modification of the electric forces between them due to their motion. In consequence, F is small when the currents are of the order of 1 A, but is sufficient to cause damage in high-power devices, particularly when the currents are of an impulsive nature, due to fault conditions or lightning strokes. *See* LIGHTNING AND SURGE PROTECTION.

In general, any charge, q, moving at velocity u is subject to a force given by Eq. (22), where $\mathbf{u}\times\mathbf{B}$

$$\mathbf{f} = q\,\mathbf{u}\times\mathbf{B} \qquad \text{newtons} \tag{22}$$

denotes the cross-product between vector quantities. That is, the magnitude of $\mathbf{f}$ depends on the sine of the angle θ between the vectors $\mathbf{u}$ and $\mathbf{B}$, of magnitudes u and B, according to Eq. (23). The force

$$f = quB\sin\theta \tag{23}$$

on a positive charge is at right angles to the plane containing $\mathbf{u}$ and $\mathbf{B}$ and points in the direction of a right-handed screw turned from $\mathbf{u}$ to $\mathbf{B}$.

The same force also acts in the axial direction on the conduction electrons in a wire moving in a magnetic field, and this force generates an emf in the wire. The emf in an element of wire of length dl is greatest when the wire is at right angles to the $\mathbf{B}$ vector, and the motion is at right angles to both. The emf is then given by Eq. (24). More generally, u is

$$\text{emf} = uB\,dl \tag{24}$$

the component of velocity normal to $\mathbf{B}$, and the emf depends on the sine of the angle between $\mathbf{dl}$ and the plane containing the velocity and the $\mathbf{B}$ vectors. The sign is given by the right-handed screw rule, as applied to Eq. (23).

Since the wire moves through a distance $u\,dt$ in time dt, and B denotes flux density, the element dl "cuts through" an amount of flux given by Eq. (25).

$$d\Phi = uB\,dl\,dt \tag{25}$$

It follows that the motion of any closed loop generates a net emf given by the rate of change of flux linkage, $d\Phi/dt$, with the loop, in accordance with Eq. (1). The concept of motion relative to the field which is implied by "flux cutting" has caused much debate in applications such as homopolar devices in which B is uniform. The difficulty originates in the interpretation of velocity in Eq. (22). In general, the Lorentz force, Eq. (26), depends on the definition of

$$\mathbf{f} = q(\mathbf{E} + \mathbf{u}\times\mathbf{B}) \tag{26}$$

$\mathbf{u}$ as relative to any specified inertial reference frame in which the electric field is $\mathbf{E}$. The magnetic field

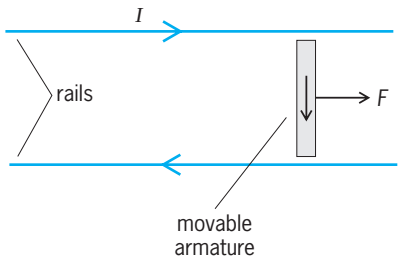

Fig. 7. Electromagnetic railgun. Passing a large current, *I*, along the rails and through the movable armature produces force, *F*, on the armature.

sources are not necessarily stationary in this reference. They may contribute to **E**, so that the electric field is not, in general, confined to electrostatics. **E** depends on the velocity of the observer, as does **B**, since the magnetic field of a moving charge is zero in the reference in which the charge is stationary [Eq. (19)].

A small search coil provides a convenient way of measuring B [Eq. (2)]. Either removing or inserting it, or changing the source current, induces an emf in accordance with Eq. (1). The change in Φ is obtained by integrating the induced voltage in time as, for example, with a ballistic galvanometer, with an integrating amplifier, or by digital processing. The transverse Hall voltage which appears in a stationary current-carrying conductor as a consequence of the Lorentz forces on the moving electrons is also commonly used to measure B. *See* GALVANOMETER; HALL EFFECT.

Since the forces on current elements are in the direction at right angles to the element, the action and reaction forces between two elements are not in opposite directions. A force balance is obtained only between total forces on closed circuits. The electromagnetic railgun (**Fig. 7**) provides an example. Passing a large current, I, from a source to the left along the rails and through a movable armature propels the armature to the right. The source of the action is the B field of the rail currents, but these are not subject to any recoil since the forces on them are at right angles to the rails. The equilibrium is accounted for locally by attributing a state of stress to the magnetic field in the gun interior, producing an outward pressure on each of the conductors. The concept of stress, due to James Clerk Maxwell, describes the flux lines as in tension, producing the force per unit area given by Eq. (12), together with a transverse pressure of the same magnitude. The stress is commonly used to predict forces in magnetic devices, particularly when the field is calculated by numerical methods.

Equation (22) shows that any charge which moves freely in a magnetic field is subjected to a force at right angles to its motion, tending to drive it into a closed circular path. This force, quB, is balanced by the centrifugal force mu^2/r on a particle of mass m, causing it to travel in a circle whose radius is given by Eq. (27) [with a transit time which is independent of u], unless there is a component of the velocity, u, in the direction of B. Since this component causes no interaction, the charges then tend to spiral along the flux lines. The behavior is characteristic of charged particles entering the Earth's magnetic field from outer space, and is utilized in many applications requiring the control of moving electrons, such as the magnetic focusing of electron beams. *See* ELECTRON MOTION IN VACUUM; VAN ALLEN RADIATION.

$$r = \frac{mu}{qB} \qquad \text{meters} \qquad (27)$$

Magnetic flux linkage. The magnetic flux linking any closed path is obtained by counting the number of flux lines passing through any surface, s, which is bounded by the path. Stated more formally, the linkage depends on the sum given in Eq. (28), where B_n

$$\Phi = \iint B_n \, ds \qquad \text{webers} \qquad (28)$$

denotes the component of B in the direction normal to the area element, ds. The rate of change of linkage gives the emf induced in any conducting wire which follows the path [Eq. (1)].

The flux linkage with a coil (**Fig. 8**) is usually calculated by assuming that each turn of the coil closes on itself, giving a flux pattern which likewise consists of a large number of separate closed loops. Each links some of the turns, so that the two cannot be separated without breaking, or "tearing," either the loop or the turn. The total linkage with the coil is then obtained by adding the contributions from each turn.

Linkage can be defined only when both the flux lines and the conductors form closed paths, so that its use depends on the absence of magnetic monopoles, since these terminate flux lines. The linkage concept cannot be applied to circuits which are not closed, such as some forms of wire antenna, and it may be difficult to apply to conductors of complex shape, as in coils that are not closely wound but form an open spiral (Fig. 8), producing a flux pattern in which the flux lines likewise form spirals. The concept of linkage attributes the induced emf to flux which may be remote from the wire in which the emf is induced, and this action-at-a-distance effect has been the source of much discussion in some applications. One such application is the magnetic Aharonov-Bohm effect, which describes the effect of the flux linkage on quantum phase. All of the actions attributed to flux linkage can be described more directly in terms of the magnetic vector potential. *See* AHARONOV-BOHM EFFECT.

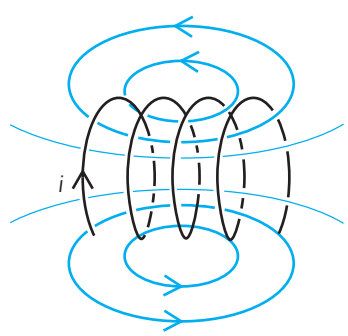

Fig. 8. *B* (magnetic flux) field of a short coil.

The inductance, L, is a property of a circuit defined by the emf which is induced by changes of current in time, as in Eq. (29). The SI unit of inductance is

$$e = -L\,di/dt \tag{29}$$

the henry (H), or V-s/A. The negative sign shows that e opposes an increase in current (Lenz's law). From Eq. (1) the inductance of a coil of N turns, each linking the same flux, Φ, is given by Eq. (30), so that

$$L = N\,\Phi/i \qquad \text{henrys} \tag{30}$$

the henry is also the Wb/A. When different turns, or different parts of a circuit, do not link the same flux, the product $N\Phi$ is replaced by the total flux linkage, Φ, with the circuit as a whole. Applying the result to a air-cored coil of N turns whose length, d, is sufficient to give an approximately uniform field, H, in the interior [Eq. (15)] gives Eq. (31), if all of the

$$L = \mu_0 N^2\,\pi r^2/d \qquad \text{henrys} \tag{31}$$

turns (Fig. 6) are circular in shape with a common radius, r. Doubling the number of turns produces a fourfold increase in L because the same current produces twice as much flux, linking twice as many turns. *See* INDUCTANCE.

Reducing the length brings the turns closer together and increases the inductance, since this consists of the sum of the mutual interactions between the various parts. The mutual inductance, M, between any two coils, or circuit parts, is defined by emf which is induced in one by a change of current in the other. Using 1 and 2 to distinguish between them, the emf induced in coil 1 is given by Eq. (32*a*),

$$e_1 = -M_{12}\,di_2/dt \tag{32a}$$

$$e_2 = -M_{21}\,di_1/dt \tag{32b}$$

where the sign convention is consistent with that used for L when the coils are connected in the series. The quantity L is referred to as the self-inductance. Likewise, the emf induced in coil 2 when the roles of the windings are reversed is given by Eq. (32*b*). The interaction satisfies the reciprocity condition of Eq. (33), so that the suffixes may be omitted. If coil

$$M_{21} = M_{12} \tag{33}$$

1 consists of N_1 turns wound sufficiently closely together to link the same flux, then the mutual inductance is given by Eq. (34).

$$M = N_1\,\Phi_1/i_2 \qquad \text{henrys} \tag{34}$$

The choice of winding role may sometimes simplify the calculation. For example, a short coil of N_2 turns, each in the form of a circle of radius r_2, will link a flux which is proportional to the circles area when placed inside, and coaxial with, the coil whose self-inductance is given by Eq. (31). Applying Eq. (15) shows that the mutual inductance between the two is given by Eq. (35) if the larger coil consists of N_1 turns wound over a length d_1. Reversing the roles makes it more difficult to calculate the mutual flux linkage when the smaller coil is too short to produce a uniform field. The mutual inductance between two coils which have approximately the same length and radius, and are sufficiently close, is given by Eq. (36). This equation applies more generally to

$$M = \mu_0 N_1 N_2\,\pi r_2^2/d_1 \qquad \text{henrys} \tag{35}$$

$$M = L_1\,N_2/N_1 \tag{36}$$

any two coils whose turns share the same flux, usually because they are linked by a common flux path of a sufficiently high permeability (Fig. 3).

Parallel wires whose length is sufficient to ignore axial variations in the field provide another important example. Current flowing in any one of them produces a magnetic field strength, H, varying inversely with the distance, r, from it [Eq. (13)]. Hence, the total flux per meter length linking two other wires, at distances r_1 and r_2 from the source, is given by Eq. (37). Dividing by the current, i, gives the mutual inductance in H/m. Any assembly of wires can be treated by superposition.

$$\Phi = \int B\,dr = (\mu_0/2\pi)i\,\log_e\,[r_2/r_1] \qquad \text{Wb/m} \tag{37}$$

The field conditions are similar between the core and sheath of a coaxial cable consisting of a core conductor of circular cross section, with radius r_1, inside a sheath of radius r_2 carrying the return current. Equation (14) shows that the field, H, varies inversely with radius in the annular space between the two. Hence Eq. (37) gives the flux, and inductance per meter, when the frequency is sufficiently high to confine the current to the surface layers of the conductors. *See* COAXIAL CABLE.

Current flow paths inside the core add to the flux, giving these paths a higher inductance than those on the surface. This increases their inductive impedance, defined as the ratio of the induced emf, e, to the current. The increase is typical of internal paths, in all problems of alternating currents flowing in solid conductors, and tends to confine the currents to a surface layer whose thickness diminishes as the frequency rises. The same effect may also restrict the current to a more limited path within the layer, as when a wire carries a current, I, at high frequency close to a conducting plate (**Fig. 9**). The requirement that the inductance, and hence the magnetic flux linkage, is a minimum produces a concentration of induced current in a path immediately beneath the wire, known as the shadowing effect. The local heating may raise the temperature of the strip sufficiently to make it glow, and provides a method of controlled welding. *See* SKIN EFFECT (ELECTRICITY).

Magnetostatics. The term "magnetostatics" is usually interpreted as the magnet equivalent of the electrostatic interactions between electric charges. The

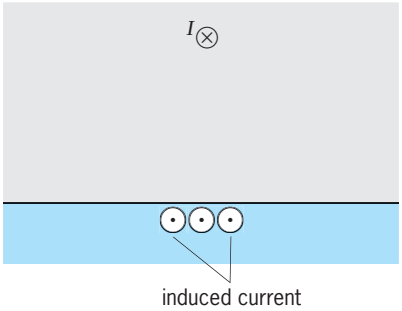

Fig. 9. Current shadowing.

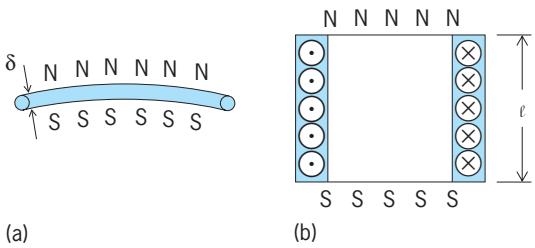

Fig. 10. Magnetic shell equivalents for (*a*) wire loop and (*b*) multiturn coil.

equivalence is described most directly in terms of the magnetic pole, since the forces between poles, like those between charges, vary inversely with the square of the separation distance. Although no isolated poles, or monopoles, have yet been observed, the forces which act on both magnets and on coils are consistent with the assumption that the end surfaces are equivalent to magnetic poles.

This follows from the magnetic shell principle. Separating two surfaces of magnetic charge of opposite sign by a small distance, δ (**Fig. 10***a*), creates a constant difference in potential, or mmf, if the charges, or poles, are uniformly distributed. A current flowing in a loop which bounds the edge of the sheets produces the same magnetic field, other than in the space between the sheets. Superposing turns shows that a uniformly wound coil of N turns of wire carrying a current, i (Fig. 10*b*), can be replaced by two sheets of magnetic charge of opposite sign, each uniformly distributed over surfaces bounded by the end turns of the coil. Both the turns and the pole sheets are shown plane in Fig. 10*b*, but in general the magnetic shell of a single turn can have any shape which is bounded by the turn, and this applies by superposition to larger coils.

The pole density, per unit area, is given by Eq. (38) for a single turn, or by Eq. (39) for a coil of N turns and length l. The total force, or torque, acting on the shell is the same as that on the loop or coil, where the force on a pole of strength q_m is given by Eq. (40). Figure 10*b* also illustrates the two alterna-

$$\rho_m = \mu_0\, i/\delta \tag{38}$$

$$\rho_m = \mu_0\, iN/l \tag{39}$$

$$f = q_m H \tag{40}$$

tive models of a uniformly polarized magnet, one consisting of surface currents representing the sum of the "Ampèrean currents," and the other the surface poles. Equations (14) and (40) show that the work done in carrying a unit pole around any closed path is equal to the current I, or iN, which is linked by the path.

Replacing currents by poles is useful for computational purposes since it replaces the mmf by a magnetic scalar potential, analogous to the electrostatic potential. Both are conservative; that is, the work required to move a pole, or charge, between any two points is independent of the path taken.

Magnetic moment. The magnetic moment of a small current loop, or magnet, can be defined in terms of the torque which acts on it when placed in a magnetic flux density, B, which is sufficiently uniform in the region of the loop. Equal but opposite forces then act on opposite sides of length a of a rectangular loop of N turns (**Fig. 11**) carrying a current, i. The force is $iNBa$ [Eq. (20)], and the torque, given by Eq. (41), depends on the effective

$$T = iNBa\,b \sin\theta \qquad \text{N-m} \tag{41}$$

distance, $b \sin\theta$, between the wires. It is proportional to the area ab, and is a maximum when the angle θ between B and the axis of the loop is 90°. A current loop of any other shape can be replaced by a set of smaller rectangles placed edge to edge, and the torques of these added to give the total on the loop. The magnetic moment, m, of any loop of area s is defined as the ratio of the maximum torque to the flux density, so that m is given by Eq. (42). This

$$m = iNs \tag{42}$$

is the magnitude of a vector quantity $\mathbf{m}$ pointing in the direction of a right-handed screw turned in the direction of current flow. It gives the torque $\mathbf{T}$ in Eq. (43) in vector cross-product notation.

$$\mathbf{T} = \mathbf{m} \times \mathbf{B} \tag{43}$$

See TORQUE.

The coil, or magnet, may also replaced by its magnetic pole, or magnetic shell, equivalent. The torque is then accounted for by the forces on poles of strength q_m at opposite ends distance l apart, where q_m is the pole density times the area, s. Hence, from

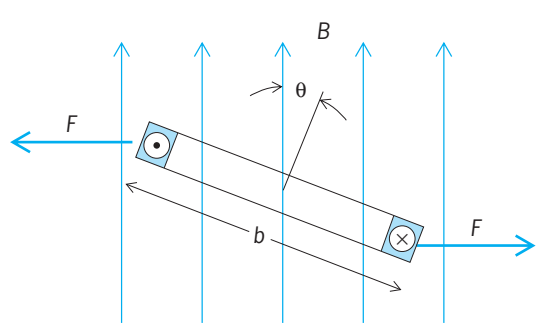

Fig. 11. Forces acting on rectangular current loop, shown in cross section, due to the magnetic flux density *B*. Current flow direction is shown in the same way as in Fig. 6.

Eqs. (39) and (40), the torque is given by Eq. (44),

$$T = \mu_0 iNHs \sin\theta \tag{44}$$

and this comparison demonstrates the equivalence. The moment, m^*, may be defined by describing the torque in terms of H instead of B, giving Eq. (45).

$$m^* = \mu_0 m \tag{45}$$

An electron of charge of $q_e$orbiting at frequency f is the equivalent of a current $i = q_e f$, giving Eq. (46)

$$m_0 = q_e f s \tag{46}$$

for the moment, where s is the area of the orbit. The permissible values are determined by the quantum energy levels. The electron spin is a quantum state which can likewise be visualized as a small current loop. Atomic nuclei also possess magnetic moments. *See* ELECTRON SPIN; MAGNETON; NUCLEAR MOMENTS.

Magnetic polarization. Materials are described as magnetic when their response to the magnetic field controls the ratio of B to H. The behavior is accounted for by the magnetic moments produced mainly by the electron spins and orbital motions. These respond to the field and contribute to it in a process referred to as magnetic polarization. The effects are greatest in ferromagnetics and in ferrites, in which the action is described as ferrimagnetic. *See* FERRIMAGNETISM; FERRITE; FERROMAGNETISM.

Current-based model. The sources are the equivalent of miniature "Ampèrean currents" whose sum, in any volume element, is equivalent to a loop of current flowing along the surface of the element. The flux density, B, depends on the field intensity, H, which is defined so that its value inside a long ferromagnetic rod of uniform cross section placed inside a long coil (Fig. 6) is not affected by the rod. It is given by Eqs. (14) and (15) and is the same inside the material as in the annular gap between the rod and the coil, in accordance with Eq. (15). If the field is not sufficiently uniform, H can be measured by using a search coil to observe the flux density, $\mu_0 H$, in the gap. The flux density inside the rod is given by Eq. (47), where

$$B = \mu_r B_0 \tag{47}$$

B_0 denotes $\mu_0 H$, and μ_r is the relative permeability [Eqs. (7) and (16)]. The same flux, B, is obtained by replacing the material by a coil in which the current in amperes per unit length is given by Eq. (48).

$$\begin{aligned} J_s &= (B - B_0)/\mu_0 \\ &= (1 - 1/\mu_r)B/\mu_0 \end{aligned} \tag{48}$$

The magnetic moment, [Eq. (42)], of a volume element of length dz, due to the current J_s flowing over the surface enclosing the area, $dx\,dy$, is given by Eq. (49). The moment per unit volume defines

$$dm = (J_s\,dz)\,dx\,dy \tag{49}$$

the magnetic polarization, as in Eq. (50). The polarization, $\mathbf{M}$, is a vector pointing in the direction of $\mathbf{dm}$ with magnitude J_s. The surface current accounts for $\mathbf{B}$ but produces an $\mathbf{H}$-like, or $\mathbf{B}/\mu_0$, field which is entirely different from $\mathbf{H}$ in the material. Substituting from Eq. (48) gives Eq. (51).

$$M = dm/dx\,dy\,dz \tag{50}$$

$$\mathbf{H} = \frac{\mathbf{B}}{\mu_0} - \mathbf{M} \tag{51}$$

This model of the material gives the flux field, $\mathbf{B}$, as observed by the voltage induced in a search coil wound around the specimen, and $\mathbf{H}$, becomes an auxiliary quantity representing the difference between $\mathbf{B}/\mu_0$ and the polarization, $\mathbf{M}$. $\mathbf{B}$ is the magnetizing field to which $\mathbf{M}$ responds. The polarization, $\mathbf{M}$, contributes to that field, when B is greater than B_0 [Eq. (47)], since the equivalent surface current is then in the same direction as the current in the magnetizing coil.

Pole-based model. The rod may also be replaced by equivalent magnetic poles distributed over the two end surfaces (Fig. 10*b*), and this alternative likewise applies to its volume elements. Since the flux lines are continuous, by definition, the component of $\mathbf{B}$ normal to the surface is the same on both sides, and the relationship $B = \mu H$ requires a change in H. The change, from H_0 on the outside to H, given by Eq. (52), in the interior, is accounted for by the

$$H = H_0/\mu_r \tag{52}$$

magnetic poles. The pole density, ρ_s^*, per unit surface area is given by Eq. (53).

$$\begin{aligned} \rho_s^* &= \mu_0(H_0 - H) \\ &= (\mu_r - 1)\mu_0 H \end{aligned} \tag{53}$$

Each volume element $dx\,dy\,dz$ acquires a magnetic moment given by Eq. (54), and a corresponding mag-

$$dm^* = (\rho_s\,dx\,dy)\,dz \tag{54}$$

netic moment, $\mathbf{M}^*$, per unit volume. The equivalent poles are defined so as to produce the same field, $\mathbf{H}$, as in the material, and the corresponding $\mu_0\mathbf{H}$ flux field is entirely different from $\mathbf{B}$, which is obtained by adding $\mathbf{M}^*$ to $\mu_0\mathbf{H}$, as in Eq. (55). It is B which be-

$$\mathbf{B} = \mu_0\mathbf{H} + \mathbf{M}^* \tag{55}$$

comes the auxiliary vector in the pole-based model, whose polarization vector, $\mathbf{M}^*$, is $\mu_0\mathbf{M}$, in accordance with Eq. (45). If a sample of the material is placed in an air gap, the surface poles acquire an opposite sign to those on the iron surfaces adjacent to them.

Additional equivalent sources appear inside the material when the changes in permeability are not confined to the surfaces. The customary use of $\mathbf{H}$ as the "magnetizing field" follows from the historical role of the pole as the magnetic analog of electric charge, and the corresponding comparison between the treatment of magnetic materials, in terms of displaced poles, with dielectrics, whose behavior depends on the displacement of charge. Both $\mathbf{B}$ and $\mathbf{H}$ represent averages of the complex field conditions

which are produced by arrays of dipoles, and are difficult to observe since local measurements require cavities whose shape determines the result. The potential functions provide more meaningful measures, in terms of the array energy, and provide a definition of the vector **B** as the differential of the magnetic vector potential, **A**, discussed below.

Magnetic hysteresis. The relationship between the flux density, B, and the field intensity, H, in ferromagnetic materials depends on the past history of magnetization. The effect is known as hysteresis. It is demonstrated by subjecting the material to a symmetrical cycle of change during which H is varied continuously between the positive and negative limits $+H_m$ and $-H_m$ (**Fig. 12**). The path that is traced by repeating the cycle a sufficient number of times is the hysteresis loop. The sequence is counterclockwise, so that B is larger when H is diminishing than when it is increasing, in the region of positive H. The flux density, B_r, which is left when H falls to zero is called the remanence, or retentivity. The magnetically "hard" materials used for permanent magnets are characterised by a high B_r, together with a high value of the field strength, $-H_c$, which is needed to reduce B to zero. The field strength, H_c, is known as the coercive force, or coercivity. Cycling the material over a reduced range in H gives the path in Fig. 12 traced by the broken line, lying inside the larger loop. The locus of the tips of such loops is known as the normal magnetization curve. The initial magnetization curve is the B-H relationship which is followed when H is progressively increased in one direction after the material has first been demagnetized ($B = H = 0$).

A magnet which forms a part of a magnetic circuit (Fig. 2) is driven to the point, P, in the part of the hysteresis loop in which B and H are in opposite directions. The mmf of the magnet of length l is Hl if the field intensity, H, is the same through the volume of the material. The reluctance $\Re$ of the rest of the magnetic circuit produces an equal but opposite mmf in response to the flux, Φ, obtained by multiplying B by the magnet cross-sectional area, a. Hence, Eq. (56) holds.

$$Hl = \Re Ba \tag{56}$$

This approximation to the position of P can be expressed graphically by observing that the slope, B/H, of the line OP is inversely proportional to $\Re$. Reducing the reluctance does not, in general, return the magnet to the point B_r, but establishes a path, PQ, known as a recoil curve, forming a part of a minor hysteresis loop.

Materials that are easily demagnetized are referred to as magnetically soft. When these are used in applications in which the supply current alternates in time, the hysteresis action changes the time phase between B and H, producing losses in the form of heat. The field strength, H, is a measure of the current, i, flowing in a single turn of suitable shape which is required to magnetize a volume element of material of unit size. Since the flux linking the coil, of unit

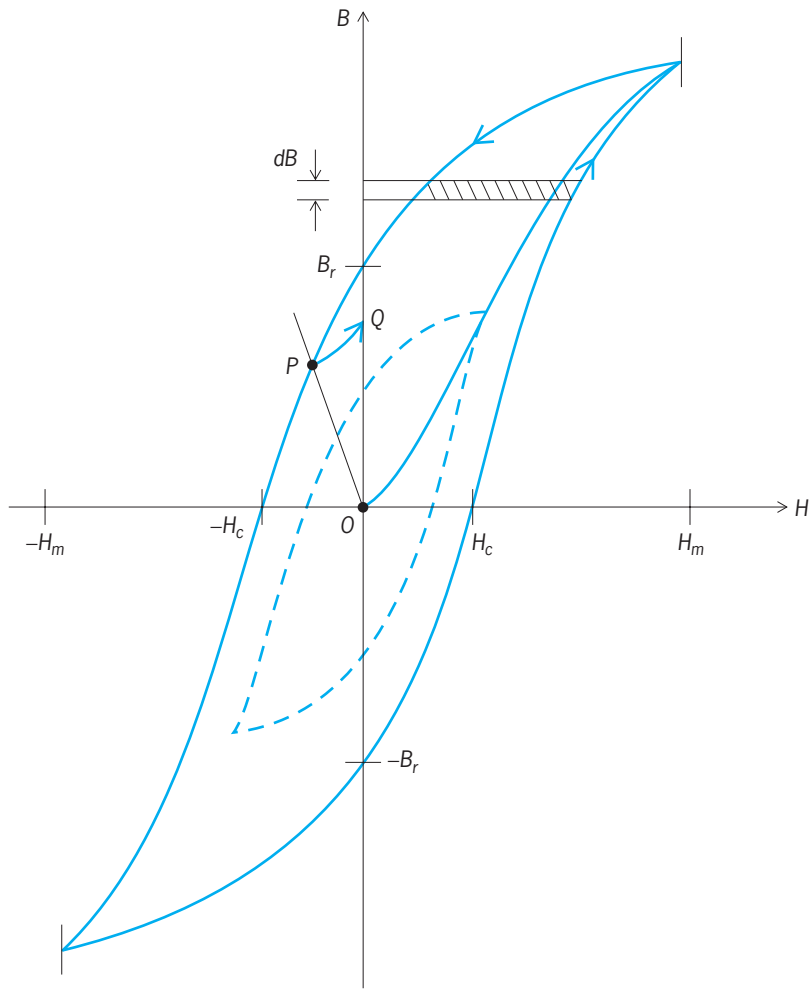

Fig. 12. Hysteresis behavior of a ferromagnetic material.

area, is B, the voltage, e, induced in the coil [Eq. (1)] is given by dB/dt, and the energy supplied to the element during any interval of time, t, is given by Eq. (57). In the last integral, $H\,dB$ is the area of a

$$\int ei\,dt = \int (dB/dt)H\,dt = \int H\,dB \tag{57}$$

strip of width dB drawn between the B-H curve and the $H = 0$ axis (Fig. 12), so that the energy injected is the area over the curve. When B varies linearly with H, up to the values B_m and H_m, the area forms a triangle, and the input, given by Eq. (58), is recov-

$$\int ei\,dt = H_m B_m/2 \tag{58}$$

ered when the line retraced is in accordance with the stored energy density in Eq. (17). Retracing the magnetization curve along a different path results in a hysteresis loop whose area is the loss, W, per unit volume per cycle. The Steinmetz coefficient, η, gives the net loss in terms of the empirical relationship of Eq. (59). The exponent n may vary between 1.5 and

$$W = \eta(B_m)^n \tag{59}$$

2.5. The Steinmetz value of 1.6 is appropriate for many materials.

Vector potential. The use of magnetic poles provides a treatment of magnets and current-carrying coils in terms of a magnetic scalar potential, but the application is limited by the need to replace current sources by magnetic pole equivalents. The more direct description of moving charges, or currents, as field sources requires the magnetic vector potential, **A**, whose significance is illustrated by its application to the calculation of induced emf's. A magnetic field which changes in time induces an emf in an element of wire of length dl given by Eq. (60), where θ is the

$$de = -(dA/dt)\,dl\,\cos\theta \tag{60}$$

angle between the vectors **dl** and **A**. The total, e, for any wire, or conducting path, is obtained by summation. The quantity $N\Phi$ in Eq. (1) can therefore be replaced by the symbol Ψ, defined by Eq. (61), with-

$$\Psi = \int A\,\cos\theta\,dl \tag{61}$$

out reference to flux or flux linkage, as in Eq. (28). The integral may extend around the N turns of a coil by adding the contributions of each, but does not require that the path, or any part of it, closes on itself. Other advantages include the description of emf in terms of the local values of **A** at the wire, in place of flux linkages requiring the summation of remote values of B over surfaces bounded by the wire [Eq. (28)].

The relationship between the magnetic field and the source currents is described most directly by treating an element, dl, of a wire carrying a current, i, as the magnetic analog of electric charge. The magnitude of **A**, like the electric potential, φ, of a point charge, varies inversely with the distance, r, from the source in accordance with Eq. (62). **dA** points

$$\mathbf{dA} = \frac{\mu_0 i\,\mathbf{dl}}{4\pi r} \tag{62}$$

in the direction of current flow in the element. This is an example of the more general relationship given by Eq. (63) between the vector potential, at point

$$\mathbf{A} = \mu\varphi/c^2 \tag{63}$$

P, due to any group of charges moving at velocity u and the potential, φ, of the same group at P. Here, c is the velocity at which changes in both potentials propagate in empty space (the velocity of light). It follows that the constant μ_0 is given by Eq. (64) in

$$\mu_0 = 1/\varepsilon_0 c^2 \tag{64}$$

terms of ε_0, sometimes known as the permittivity of free space.

The mutual inductance between any two wires, as defined by Eqs. (32) and (60), is given by Eq. (65)

$$M = \Psi_2/i_1 \tag{65}$$

in terms of the summation along wire 2 of the contributions to **A** from the current, i_1, flowing in wire 1.

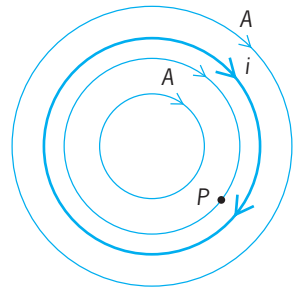

Fig. 13. Magnetic vector potential, *A*, due to a loop of wire.

The calculation cannot be expressed in terms of flux linkage if wire 2 does not form a closed circuit. The two summations can be combined to give Eq. (66)

$$M = \mu_0 \int_1 \int_2 (1/r)\,dl_1\,dl_2\,\cos\theta \tag{66}$$

in terms of elements dl_1 and dl_2 at distance r and angle θ to each other. The result, due to Franz Neumann, is limited to circuits in air formed from wires, or filaments, in which the current is everywhere the same. In general, it is necessary to calculate **A** explicitly, particularly when considering the effects of induced currents and of ferromagnetic materials.

The magnetic field due to currents flowing in parallel straight wires is described by an **A** vector which points everywhere in the axial direction. Equation (63) shows that an equipotential map of the surfaces on which the magnitude **A** is constant is identical to the map of the φ equipotentials when the same wires carry charge instead of current. Hence the inductances between various conductors, when the currents are confined to the surface layers, are proportional to the reciprocal of the capacitances between the same conductors. Closing up any current-carrying wire into a loop (**Fig. 13**) likewise "curls up" its **A** vector, since Eq. (62) requires that the direction of **A** at any point, P, close to the wire is that of the nearest current. If the frequency is sufficiently low to give the same current at all points in the wire, then **A** likewise forms closed loops. Equations (28) and (61) can be expressed in the differential form of Eq. (67), and this provides an alternative definition

$$\mathbf{B} = \operatorname{curl}\mathbf{A} \tag{67}$$

of magnetic flux density. The "general equations of the electromagnetic field" set out by Maxwell in his *Treatise* depend on this use of **A** as the principal measure of the magnetic field, and **B** as a symbol for its differential. Maxwell's approach is reversed in modern practice, in which **B** is treated as the source and **A** derived from it, leaving open the choice of the vector property of divergence, div **A**, or gauge. Equation (62) describes the magnetic field of a current element in terms of an **A** vector which does not necessarily close on itself, but satisfies the Lorentz condition, given by Eq. (68).

$$\operatorname{div}\mathbf{A} = -(1/c^2)\,\partial\varphi/\partial t \tag{68}$$

The emf, *de*, in Eq. (60) is an example of the force given by Eq. (69), which acts on any charge,

$$\mathbf{f} = -q d\mathbf{A}/dt \qquad (69)$$

q, due to a change of **A**. This provides a description of charge behavior in terms of an electrokinetic momentum, $q\mathbf{A}$, predicting forces which tend to oppose any change. Interrupting the movement of the conduction charges in a current-carrying winding, for example, by opening a switch, causes an arc to form across the contacts due to a rise in electrical potential which is analogous to the rise in pressure when the flow of a fluid through a pipeline is suddenly blocked. The comparison illustrates the physical significance of the circuit inductance. The quantity $q\mathbf{A}$ describes the sum of the interactions between all of the moving electrons and q, and is thus a property of the system as a whole. The magnetic, or electrokinetic, energy is likewise obtained by summing the energy density, $\mathbf{J} \cdot \mathbf{A}/2$, over all of the currents, of density **J**, and it is this from which the alternative description of energy as a property of the field [Eq. (17)] is derived. The vector **A** is sufficient to describe many magnetic, or electrokinetic, effects without reference to the vector **B**. *See* POTENTIALS. C. J. Carpenter

Bibliography. W. Berkson, *Fields of Force: Development of a World View from Faraday to Einstein*, 1974; B. I. Bleaney and B. Bleaney, *Electricity and Magnetism*, 3d ed., 1976; J. D. Buchwald, *From Maxwell to Microphysics*, 1985; H. E. Burke, *Handbook of Magnetic Phenomena*, 1986; W. J. Duffin, *Electricity and Magnetism*, 4th ed., 1990; I. S. Grant and W. R. Phillips, *Electromagnetism*, 2d ed., 1990; D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, 6th ed., Wiley, 2000; P. Hammond, *Electromagnetism for Engineers: An Introductory Course*, 4th ed., Oxford University Press, 1997; Y. Ishikawa and Miura (eds.), *Physics and Engineering Applications of Magnetism*, 1991; J. D. Jackson, *Classical Electrodynamics*, 3d ed., 1998; J. D. Krauss, *Electromagnetism*, 4th ed., McGraw-Hill, 1991; P. F. Mottelay, *Biographical History of Electricity and Magnetism*, 1975, reprint 1991; R. G. Powell, *Electromagnetism*, 1990; J. R. Reitz, F. J. Milford, and R. W. Christy, *Foundations of Electromagnetic Theory*, 4th ed., Addison-Wesley, 1993; M. N. O. Sadiku, *Elements of Electromagnetism*, 1995; M. Sibley, *Introduction to Electromagnetics*, 1996; J. Vanderlinde, *Classical Electromagnetic Theory*, 1993; E. Whittaker, *History of Theories of Aether and Electricity*, reprinted 1987; H. D. Young and R. A. Freedman, *University Physics*, 9th ed., 1998.

Magnetite

A cubic mineral and member of the spinel structure type, space group Fd3m, with composition $[Fe^{3+}]^{IV}[Fe^{2+}Fe^{3+}]^{VI}O_4$. It possesses the inverse spinel structure, in which half the ferric iron is tetrahedrally coordinated and the remaining half as well as all ferrous iron are octahedrally coordinated by the cubic close-packed oxygen atoms. The color is opaque iron-black and streak black, the hardness is 6 (Mohs scale), and the specific gravity is 5.20. The habit is octahedral (see **illus.**), but the mineral usually occurs in granular to massive form, sometimes of enormous dimensions. Though possessing no cleavage, some specimens show an octahedral parting. Limited Al^{3+} may substitute for Fe^{3+}; and Ca^{2+}, Mn^{2+}, and Mg^{2+} for Fe^{2+}.

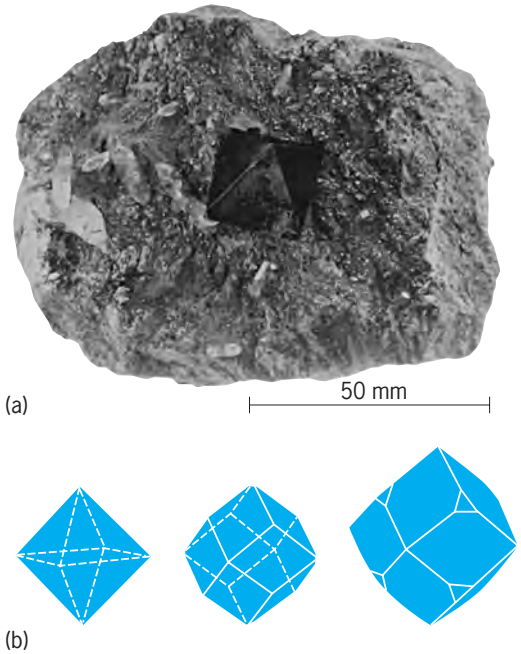

Magnetite. (*a*) Crystal on block of mica schist with small quartz crystals on left, Binnenthal, Switzerland (*American Museum of Natural History specimens*). (*b*) Crystal habits (*after C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*).

Magnetite is a natural ferrimagnet, but heated above 578°C (1072°F; the Curie temperature) it becomes paramagnetic. It can be prepared by heating hematite (α-Fe_2O_3) in a reducing atmosphere, by heating hematite in air above 1400°C (2552°F), or by the oxidation of iron at high temperatures. Oxidation of magnetite leads to maghemite (γ-Fe_2O_3), which gradually inverts to hematite.

The major magnetic ore of iron, magnetite may be economically important if it occurs in sufficient quantities. It often occurs as an early magmatic segregation in basic rocks, such as anorthosites and norites where it has accumulated in bands by gravity settling. It is also a contact metamorphic product, occurring with limestones, leptites, and so forth. Moderately resistant to weathering, magnetite sometimes accumulates in detrital sands. The most spectacular ore body occurs at Kiruna in northern Sweden, where the magnetite fluid differentiate was injected into syenitic rocks. Other important occurrences are in Norway, Russia, and Canada. Magnetite is common as an accessory mineral in igneous

rocks throughout the world. *See* EMERY; HEMATITE; ILMENITE; IRON METALLURGY; MAGNETIZATION; ORE AND MINERAL DEPOSITS; SPINEL. Paul B. Moore

Magnetization

The process of becoming magnetized; also the property and in particular the extent of being magnetized. Magnetization has an effect on many of the physical properties of a substance. Among these are electrical resistance, specific heat, and elastic strain. *See* MAGNETOCALORIC EFFECT; MAGNETORESISTANCE; MAGNETOSTRICTION.

The magnetization **M** of a body is caused by circulating electric currents or by elementary atomic magnetic moments, and is defined as the magnetic moment per unit volume of such currents or moments. In the electromagnetic system of units (emu), **M** is measured in gauss; in the mks system, **M** is measured in webers per square meter. For **M**, 1 weber/m^2 = $10^4/4\pi$ gauss.

The magnetic induction or magnetic flux density **B** is given by Eq. (1), where **B** and **M** are in gauss,

$$\mathbf{B} = \mathbf{H} + 4\pi\mathbf{M} \text{ (emu)} \tag{1}$$

and **H**, the applied magnetic field, is in oersteds (equivalent to gauss); or by Eq. (2), where **B** and

$$\mathbf{B} = \mu_0\mathbf{H} + \mathbf{M} \text{ (mks)} \tag{2}$$

M are in webers/m^2, **H** is in ampere-turns/m, and μ_0, the permeability of free space, is defined as $4\pi \times 10^{-7}$ henry/m, that is, webers/(ampere-turn)(m). *See* ELECTRICAL UNITS AND STANDARDS.

The permeability μ of a substance is defined as the ratio **B**/**H** (in emu) or $\mathbf{B}/\mu_0\mathbf{H}$ (in mks). The magnetic susceptibility χ is defined as the ratio **M**/**H** (in emu) or $\mathbf{M}/\mu_0\mathbf{H}$ (in mks). From Eqs. (1) and (2) Eq. (3) is

$$\mu = 1 + (4\pi)\chi \tag{3}$$

obtained, where the 4π is to be used only in the emu system. It is to be noted that the magnitude of μ is the same in both systems of units, but the magnitude of χ differs by a factor of 4π. Both μ and χ may be tensors, although usually they are simple numbers. *See* MAGNETIC SUSCEPTIBILITY.

The topic of magnetization is generally restricted to materials exhibiting spontaneous magnetization, that is, magnetization in the absence of **H**. All such materials will be referred to as ferromagnets, including the special category of ferrimagnets. A ferromagnet is composed of an assemblage of spontaneously magnetized regions called domains. Within each domain, the elementary atomic magnetic moments are essentially aligned, that is, each domain may be envisioned as a small magnet. An unmagnetized ferromagnet is composed of numerous domains, oriented in some fashion as shown in **Fig. 1**, so that the total magnetization is zero.

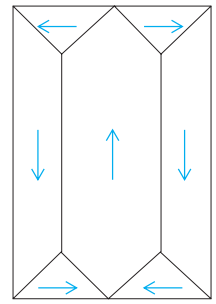

Fig. 1. Possible arrangement of domains in an unmagnetized single crystal. The arrangement is much more haphazard in polycrystals.

The process of magnetization in an applied field **H** consists of growth of those domains oriented most nearly in the direction of **H** at the expense of others (**Fig. 2***a*), followed by rotation of the direction of magnetization against anisotropy forces (Fig. 2*b*). For discussions of anisotropy and of domains *see* FERROMAGNETISM

On removal of the field **H**, some magnetization will remain, called the remanence $\mathbf{M}_r$.

Magnetization curves. These curves, sometimes called *B-H* curves, are used to describe magnetic materials. They are plotted with **H** as abscissa and with either **M** or **B** as ordinate. In **Fig. 3**, $\mathbf{B}_r$ is the remanent induction ($\mathbf{B}_r = 4\pi\mathbf{M}_r$); $\mathbf{H}_c$ is the coercive force, or reverse field required to bring the induction **B** back to zero; and $\mathbf{M}_s$ is the saturation magnetization, or magnetization when all domains are aligned. The saturation magnetization is equal to the spontaneous magnetization of a single domain, except that it is possible to increase this magnetization slightly by application of an extremely large field. Saturation magnetization is temperature dependent, and disappears completely above the Curie temperature T_c where a ferromagnet changes into a paramagnet. The **table** lists $\mathbf{M}_s$ and T_c for a few ferromagnetic materials. *See* CURIE TEMPERATURE.

The initial permeability μ_i of a substance is the slope of the magnetization curve at $\mathbf{H} = 0$. There is a definite correlation between initial permeability and coercive force; that is, materials of large μ_i have small $\mathbf{H}_c$ and vice versa.

The relationship between **B** and **H** may be studied

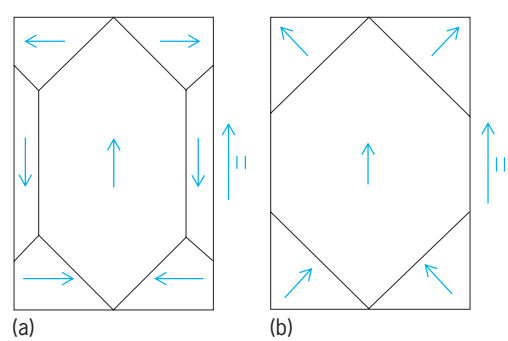

Fig. 2. Process of magnetization. (*a*) Domain growth. (*b*) Rotation.

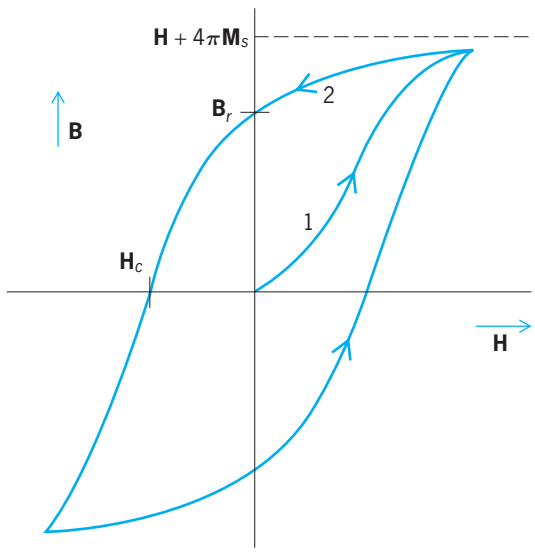

Fig. 3. Magnetization or *B-H* curves.

Values of saturation magnetization $\mathbf{M}_s$ and ferromagnetic Curie temperature T_c

Substance	$\mathbf{M}_s$ at 293 K (68°F), gauss	$\mathbf{M}_s$ at 0 K (−460°F), gauss	T_c, K (°F)
Fe	1707	1752	1043 (1418)
Co	1400	1446	1394 (2050)
Ni	485	510	631 (676)
Gd	0	1980	293 (68)
MnBi	600	675	670 (746)
Fe_3O_4	480	510	858 (1085)
EuO	0	1910	77 (−321)
$Y_3Fe_5O_{12}$	135	195	560 (548)

by means of a Rowland ring or toroid in which the core is the material to be studied. For this arrangement, the numerical value of **H** can be computed from the relation in Eq. (4), where N is the number

$$H = \frac{NI}{l} \tag{4}$$

of turns of the coil, I is the current in the coil, and l is the length of the mean circumference of the coil. If the radius of the ring is large compared to the radial distance across the core, the flux density will be nearly uniform.

The flux density **B** is not a linear function of **H**, and furthermore, its value depends not only upon **H** but upon the previous magnetic history of the sample used.

The ferromagnetic core of the Rowland ring may be demagnetized by successive reversals of the current as the current is gradually reduced from a sufficiently high initial current to a current that will produce a magnetizing force less than any value to be used subsequently. *See* DEMAGNETIZATION.

Changes in **B**, but not **B** itself, can be measured by wrapping a search coil on the toroid and observing the throw of a ballistic galvanometer.

Normal magnetization curve. When the core of the Rowland ring is initially demagnetized, $\mathbf{B} = 0$ and $\mathbf{H} = 0$. If the current is quickly increased to give a predetermined value $\mathbf{H}_1$, a flux $\Phi_1 - 0$ will be measured by the throw of the ballistic galvanometer. From the flux and the area A of cross section of the core, the value of $\mathbf{B} = \Phi/A$ corresponding to the selected **H** is computed. Then, without reducing the current, a second increase in current is made to reach $\mathbf{H}_2$. Again, the throw of the galvanometer measures the change in flux $\Phi_2 - \Phi_1$, and hence Φ_2 is found and $\mathbf{B}_2$ is computed corresponding to $\mathbf{H}_2$. By continuing this process for as many pairs of **B** and **H** as are necessary, points on the magnetization curve are obtained. The curve obtained starting with the core demagnetized is called the normal magnetization curve (**Fig. 4**). If the specimen were not initially demagnetized, the process here described would yield a magnetization curve that would differ in shape from the normal magnetization curve and, in general, would not pass through the origin.

In the curve of Fig. 4, **B** increases slowly at first as **H** rises, then increases rapidly until the "knee" of the curve is reached. Here, the rate of change of **B** with respect to **H** decreases and becomes small as saturation is approached.

Single crystals. Magnetization curves of single crystals depend upon the direction of **H** with respect to the crystallographic axes. If a weak field is applied to a crystal, those domains whose magnetic moments are most nearly parallel to **H** increase in size at the expense of those in other directions. There is then a small net contribution of the domains to the flux. As the field increases, the boundaries of the domains continue to change, and **B** increases rapidly along the steep part of the curve. Near the end of this change, all the magnetic moments in the domains rotate in a direction parallel to those that have been increasing. The rotation within the domains requires greater change in **H**, and thus the slope changes at the knee. As **H** is further increased, the domains rotate until their magnetic moments are all parallel to the field

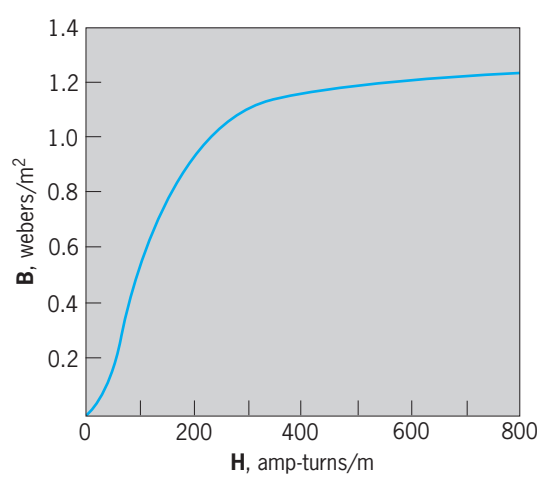

Fig. 4. Normal magnetization curve.

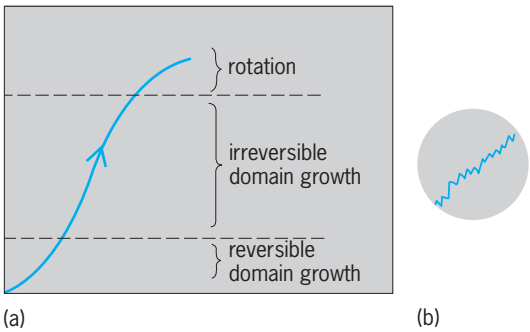

Fig. 5. Process of magnetization. (*a*) Portion of a magnetization curve, showing magnetization processes. (*b*) Enlarged section of path, showing Barkhausen effect.

and saturation has been reached. *See* SINGLE CRYSTAL.

The ferromagnetic materials in common use are polycrystalline; that is, a piece of the material consists of a tremendous number of single crystals of random orientation. In each single crystal, there are many domains. The magnetization of the whole body consists of the magnetization of the various single crystals within that body.

Domain growth. The process of magnetization is shown in **Fig. 5***a*. Domain growth proceeds by movement of the so-called Bloch wall between domains. This takes place reversibly at first, then irreversibly. Irreversible motion causes sudden almost discontinuous changes in magnetization, called the Barkhausen effect (Fig. 5*b*). *See* BARKHAUSEN EFFECT.

Hysteresis. The irreversible nature of magnetization is shown most strikingly by the fact that the path of demagnetization does not retrace the path of magnetization—path 2 of Fig. 3 does not retrace path 1. There is a tendency for the magnetization to show hysteresis, that is, to lag behind the applied field, and the loop of Fig. 3 is called a hysteresis loop. The area of the loop is a measure of the energy lost to heat, per cycle, in going around the loop, and therefore is related to the energy lost by a magnet in an alternating field. An additional energy loss is that from eddy currents; this is absent in the nonconducting ferrites and may be minimized in metals by lamination.

Demagnetizing fields. The effect of an applied field **H** on a sample is always reduced by a demagnetizing field $-N\mathbf{M}$ coming from the surface "poles," where N, the demagnetizing factor, depends upon the sample shape and orientation with respect to **H**. It varies from nearly 0 in a sample long and thin in the direction of **H**, to nearly 4π (in emu) in a sample short and fat, like a disk of revolution about **H**. Magnetization curves are generally obtained from long, thin samples to avoid demagnetization. The demagnetization effects are extremely important in ferromagnetic resonance. *See* MAGNETIC RESONANCE.

Elihu Abrahams; Frederick Keffer; Kenneth V. Manning

Bibliography. B. I. Bleaney and B. Bleaney, *Electricity and Magnetism*, 3d ed., 1976, paper 1989; D. Craik, *Magnetism*, 1995; W. J. Duffin, *Electricity and Magnetism*, 4th ed., 1990; J. D. Jakubovics, *Magnetism and Magnetic Materials*, 2d ed., 1994.

Magneto

A type of permanent-magnet alternating-current generator frequently used as a source of ignition energy on tractor, marine, industrial, and aviation engines. The higher cost of magneto ignition is not warranted in modern automobiles, where storage batteries are required for other electrically operated equipment. Hand-operated magneto generators were once widely used for signaling from local battery telephone sets. *See* GENERATOR.

Modern induction-type magnetos consist of a permanent-magnet rotor and stationary low- and high-tension windings, also called the primary and secondary windings. The **illustration** shows two induction-type magnetos in a dual-ignition circuit for a nine-cylinder aircraft engine. The two magnetos are identical, but only magneto no. 1 is shown schematically. The ignition system may receive energy from either or both magnetos, according to the position of the switch.

The energy output of a magneto is obtained as a result of a rapid rate of change of flux through the stationary windings. The primary winding has comparatively few turns and the secondary winding has many thousand turns of fine wire. One end of the secondary winding is connected to an end of the primary winding and grounded to the frame of the magneto. The primary winding is closed on itself through a breaker mechanism actuated by a cam on the magneto shaft. The breaker is mechanically set to interrupt the primary circuit each time the flux

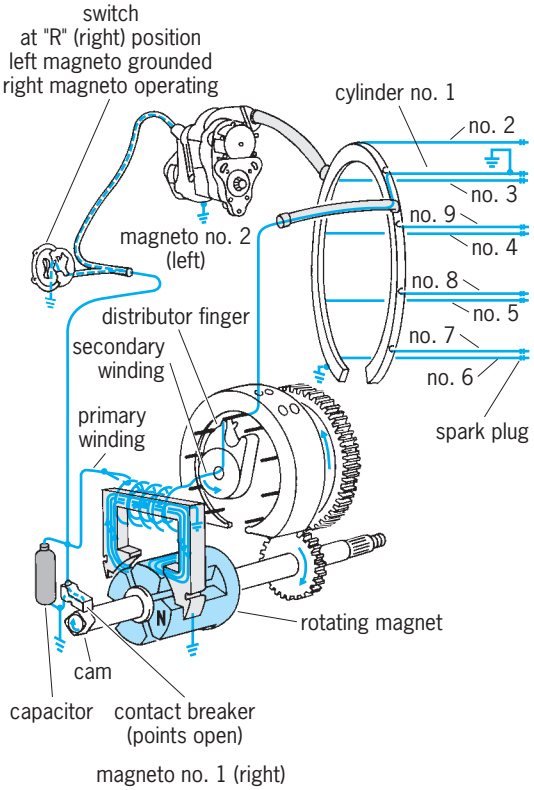

Multipolar aviation magneto in dual-ignition circuit for nine-cylinder radial engine. (*Bendix Aviation Corp.*)

through the winding is changing at its greatest rate. The sudden collapse of the primary current induces a very high voltage in the secondary winding.

Magnetos are always geared to the engine shaft and timed to open the breaker at the proper instant. Magneto speed depends on the number of poles of the magneto and the number of engine cylinders. The distributor finger speed is always one-half engine speed. *See* IGNITION SYSTEM; INTERNAL COMBUSTION ENGINE.

Robert T. Weil, Jr.

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers,* 10th ed., 1996; D. Fink and H. W. Beatty (eds.), *Standard Handbook for Electrical Engineers,* 14th ed., 1999.

Magnetocaloric effect

The reversible change of temperature accompanying the change of magnetization of a ferromagnetic or paramagnetic material. Thermodynamics theory shows that, for an adiabatic change of field ΔH, the change of temperature ΔT is given by the equation below, where c_H is the specific heat per unit volume

$$\frac{\Delta T}{\Delta H} = -\frac{T}{c_H}\left(\frac{\partial M}{\partial T}\right)_H$$

at constant H, and M is the magnetization.

This change in temperature may be of the order of 2°F (1°C), and is not to be confused with the much smaller hysteresis heating effect, which is irreversible. *See* THERMAL HYSTERESIS.

Except in antiferromagnets, $(\partial M/\partial T)_H$ is negative and an adiabatic decrease in H causes T to drop. This is the basis of the Giauque-Debye adiabatic demagnetization of paramagnetic salts, a technique which has achieved extremely low temperatures. *See* ADIABATIC DEMAGNETIZATION.

Elihu Abrahams; Frederick Keffer

Magnetochemistry

The study of magnetic properties of materials by analyzing the interactions between a magnetic field and chemical substances, which include both molecular and extended structures. All substances respond to a magnetic field. Some materials, called permanent magnets, generate magnetic fields.

Magnetic properties. When a substance is placed in a magnetic field **H**, its response **B**, called magnetic induction, is given by Eq. (1). Here **M** is the

$$\mathbf{B} = \mu\mathbf{H} = (1 + 4\pi\kappa)\mathbf{H} = \mathbf{H} + 4\pi\mathbf{M} \qquad (1)$$

magnetization, which is total magnetic moment per volume in the material, **H**, **B**, and **M** are vector quantities; their corresponding magnitudes (intensities or strengths) are symbolized as H, B, and M. The permeability μ is the tensor relationship between the external field and induction; the volume susceptibility κ is the tensor relationship between the external field and magnetization ($\mathbf{M} = \kappa\mathbf{H}$). When magnetic measurements are made on single crystals with well-defined directions, the tensor properties of κ can be obtained. Susceptibilities measured along different crystal axes are called principal susceptibilities, and the difference between principal susceptibilities is called the magnetic anisotropy of the substance. However, measurements are usually made on powders, so κ is typically reported as a single number (a scalar quantity), which is the average of the three principal susceptibilities and represents the ratio M/H. More frequently, the molar susceptibility is reported, which is $\chi_M = \kappa \cdot \mathrm{FW}/d$, where FW is the formula weight and d is the density of the substance.

The susceptibility can be negative, zero, or positive: when negative, $B < H$ and the substance is repelled by a magnet (diamagnetic); when zero, $B = H$ (nonmagnetic and highly unusual); when positive, $B > H$ and the substance is attracted by a magnet (paramagnetic). *See* MAGNETIC SUSCEPTIBILITY.

Magnetism originates in matter from the intrinsic spin of nuclei (protons and neutrons) and electrons, as well as from the orbital motion of electrons. The spins of protons, neutrons, and electrons create magnetic moments. The magnitudes of these moments are related inversely to their masses, so the electronic magnetic moment is about 2000 times larger than that of a proton or neutron. Therefore, macroscopic magnetic behavior arises essentially from electronic properties. The magnetism associated with nuclei is used in nuclear magnetic resonance (NMR) and is very useful for structural information and analysis. *See* ATOMIC STRUCTURE AND SPECTRA; MAGNETIC RESONANCE; NUCLEAR MAGNETIC RESONANCE (NMR); NUCLEAR MOMENTS.

The orbital motion of electrons can also create a magnetic moment if the orbital angular momentum is not quenched. The combination of electron spin and orbital momenta produces a local magnetic moment assigned to each atom in a chemical structure. These local moments can interact with one another and with an external magnetic field, leading to the substance's magnetic behavior and response. If the orbital angular momentum for an atom is quenched, which is often the case, then the local moment comes entirely from the total electron spin. If all electron spins are paired, an atom or ion has no local magnetic moment. Therefore, local magnetic moments primarily arise from unpaired electrons. *See* ELECTRON SPIN; MAGNETISM; MOLECULAR ORBITAL THEORY.

Units. Magnetic moments are given in units of Bohr magnetons, $\mu_B = (e/2mc) = 9.274 \times 10^{-21}$ erg/gauss $= 9.274 \times 10^{-24}$ joule/tesla. The Bohr magneton is not an SI unit but corresponds to the magnetic moment of a 1s electron in hydrogen. Magnetic field strengths and inductions are given in units of gauss, Oersted, or tesla (1 gauss = 1 Oersted = 10^4 tesla). Volume magnetic susceptibility is dimensionless; molar susceptibility has the units emu/mole = 4π cm^3/mole.

Measurements. Magnetic susceptibilities are commonly determined by measuring a material's response to an external magnetic field. The Gouy and Faraday methods rely on measuring the weight difference for a sample in and out of a magnetic field. A SQUID (superconducting quantum interference device) uses changes in magnetic fluxes in a superconducting ring and is a sensitive instrument for measuring magnetic responses. NMR studies of solutions can also be used to approximate susceptibilities. Magnetic susceptibilities are typically measured over a range of temperature, typically from about 4 to 300 K. In some cases, magnetization is recorded as a function of external magnetic field. *See* SQUID.

Magnetic states of matter. The nature of a material's response to an external magnetic field essentially depends on the number of unpaired electrons. In an external magnetic field, all electrons, as charged particles, will develop trajectories that create an induced magnetic field opposing the applied field, according to Lenz's law. The presence of unpaired electron spins will create local moments that counteract this induced field. If these local moments interact strongly with one another within the material in the absence of the external field, there can be cooperative behavior. The commonly observed magnetic states are diamagnetism, paramagnetism, ferromagnetism, antiferromagnetism, ferrimagnetism, and superconductivity. *See* LENZ'S LAW.

Diamagnetism. This is a property of a material which is repelled by an external magnetic field due to the induced motion of electrons. All electron spins are paired giving a net spin of zero. This is the most common behavior for main-group molecules and solids. The susceptibility is independent of temperature and field strength and is usually of the order of magnitude 10^{-6} to 10^{-5} emu/mole. *See* DIAMAGNETISM.

Paramagnetism. This is the property of a material which is attracted by an external magnetic field. There are unpaired electrons, which are commonly observed for transition-metal and rare-earth metal complexes. The susceptibility is inversely proportional to temperature but independent of field strength, and is usually of the order of magnitude 10^{-4} to 10^{-3} emu/mole. *See* PARAMAGNETISM.

Ferromagnetism. This property is exhibited by materials that are strongly attracted by an external magnetic field and can produce a permanent magnet. There are unpaired electron spins that are held in parallel alignment between adjacent magnetic atoms (**Fig. 1**). Ferromagnets show paramagnetic behavior above a critical temperature called the Curie temperature. *See* CURIE TEMPERATURE; FERROMAGNETISM.

Antiferromagnetism. This property is exhibited by materials that are weakly attracted by an external magnetic field because there are unpaired electrons that are held in antiparallel alignment between different atoms, and there is no net magnetization (**Fig. 2**). If the external field is oriented parallel to the direction of these local moments (called the easy axis of magnetization), the material would be repelled by the field. Antiferromagnets show paramagnetic behavior above a critical temperature called the Neél temperature. *See* ANTIFERROMAGNETISM.

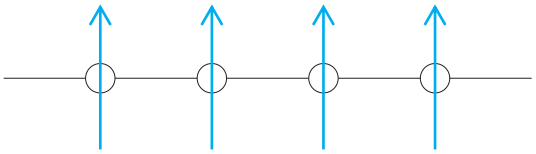

Fig. 1. Diagram of ferromagnetism.

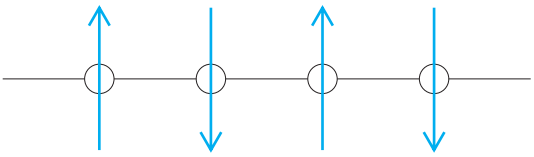

Fig. 2. Diagram of antiferromagnetism.

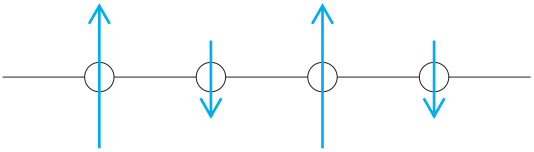

Fig. 3. Diagram of ferrimagnetism.

Ferrimagnetism. This is a property of materials that are attracted by an external magnetic field as there are unpaired electrons with antiparallel alignment, but there is an overall net magnetization (**Fig. 3**). Ferrimagnets show paramagnetic behavior above a critical temperature. *See* FERRIMAGNETISM.

Superconductors. These are "perfect" diamagnets with an immeasurably small electrical resistivity. Electrons are coupled into Cooper pairs, and the material completely repels an external magnetic field. Superconductors adopt "normal" behavior above a critical temperature. *See* SUPERCONDUCTIVITY.

Understanding magnetic behavior. One appeal of magnetochemistry is the ability to construct accurate mathematical models to understand magnetic responses. These models rely on constructing expressions for the spectrum of energy levels of the chemical system in the presence of a magnetic field. From a thermodynamic perspective, an external magnetic field does work on the system of electrons that is at thermal equilibrium. Therefore, the magnetic field affects the spectrum of electronic energy states, while the temperature controls how these states are occupied by electrons. Quantum mechanics provides the physical and mathematical machinery for interpreting measurements of magnetic susceptibilities with the Van Vleck equation. If a substance has unpaired spins (total spin S), the ground electronic state is split into $2S + 1$ levels by a magnetic field according to the Zeeman energy term in Eq. (2), and g is about 2. The occupation of the excited energy levels at temperature T are determined by the factor $\exp(-g\mu_B H/kT) = \exp(-1.34H/T)$ [H in tesla, T in K]. Since H is typically less than 10 tesla, magnetic fields place a small perturbation on

$$E(H) = -g\mu_B H M_S \qquad (2)$$
$$M_S = S, S-1, \ldots, -S$$

the occupation of electronic states until very low temperatures are reached. *See* MOLECULAR STRUCTURE AND SPECTRA.

Noninteracting magnetic atoms. If local magnetic moments are widely separated, they do not interact with one another, and the molar susceptibility has three components $\chi_M = \chi_{dia} + \chi_{TIP} + \chi_{Curie}$. The first, χ_{dia}, is the diamagnetic (Langevin) susceptibility due to the induced magnetic field from the circulating electrons. This term is negative. It increases with the size of the atom/ion and is additive over all atoms making up the chemical substance. For a molecule or unit cell, the Pascal relation approximates the diamagnetic contribution $\chi_{dia} = \Sigma n_A \chi_{dia}(A) + \Sigma n_B \chi_{dia}(B)$, where A = atoms, B = bonds, n_A = number of A atoms, n_B = number of B bonds. $\chi_{dia}(A)$ and $\chi_{dia}(B)$ are empirical constants derived from measured susceptibilities. The second, χ_{TIP}, is Van Vleck (temperature-independent) paramagnetism and occurs when the external magnetic field causes excited electronic states to mix with the ground state. Molecular ions, such as $(CrO_4)^{2-}$ and $(MnO_4)^-$, show temperature-independent paramagnetism. The third, χ_{Curie}, is the typical temperature-dependent paramagnetism. As temperature increases, the magnetic susceptibility decreases; thus $\chi_{Curie} = 0.125\mu_{eff}^2/T$, where μ_{eff} is the effective magnetic moment per atom/ion (χ_{Curie} in cgs units emu/mol; T in K). When orbital angular momentum is quenched, $\mu_{eff} = g[S(S+1)]^{1/2} \cong [n(n+2)]^{1/2}$ [$n = 2S$ = number of unpaired electrons] in units of Bohr magnetons, called spin-only magnetic moments. For rare-earth ions and certain transition-metal complexes, the total orbital angular momentum is not quenched, and through spin-orbit coupling we must consider the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$, so that $\mu_{eff} = g[J(J+1)]^{1/2}$, where g = Landé g-factor. One of the early major successes of quantum mechanics, this formula successfully explained the observed magnetic behavior of rare-earth salts. At low temperatures or high magnetic fields (that is, large H/T), the Curie expression breaks down because only the very lowest electronic energy levels are occupied. In this regime, the molar susceptibility follows the approximate relation $\chi_M = M_{sat}/H$, where M_{sat} is saturation magnetization—the largest possible magnetization for the substance. *See* RARE-EARTH ELEMENTS.

Interacting magnetic atoms. Magnetic centers are never truly isolated from one another in any substance. The interaction between two centers of unpaired spins, $\mathbf{S}_1$ and $\mathbf{S}_2$, is called exchange and the potential energy is $-J_{12}\,\mathbf{S}_1 \cdot \mathbf{S}_2$, where J_{12} is the exchange constant (given in units of K), and called the Heisenberg exchange model. If $J_{12} > 0$, the preferred interaction is ferromagnetic; if $J_{12} < 0$, the interaction is antiferromagnetic. When exchange is included in the Van Vleck equation, the temperature-dependent molar susceptibility becomes $\chi_{Curie\text{-}Weiss} = 0.125\mu_{eff}^2/(T - \Theta)$, called the Curie-Weiss law, where Θ, the Weiss constant, is proportional to J_{12}. Plots of $1/\chi_{Curie\text{-}Weiss}$ versus T are linear, and the slopes allow determination of μ_{eff} and the intercepts along the T axis tell the sign and magnitude of the exchange constant (**Fig. 4**). *See* CURIE-WEISS LAW.

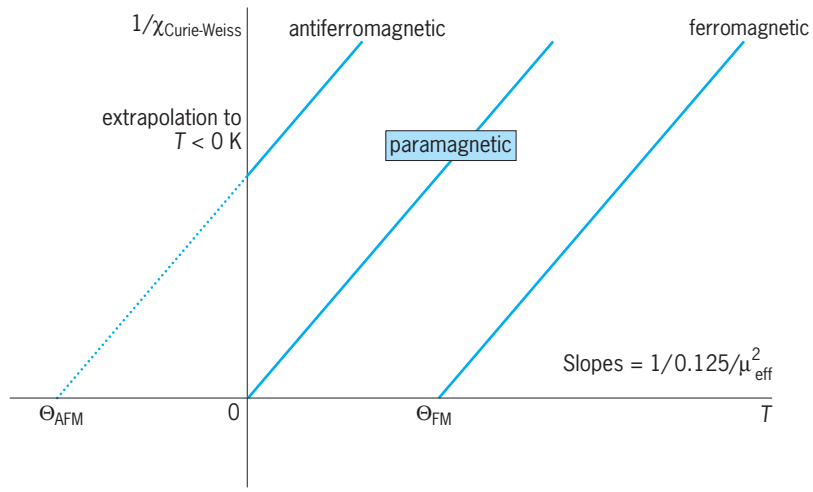

Fig. 4. Qualitative plots of $1/\chi_{Curie\text{-}Weiss}$ versus T.

Exchange is a purely quantum-mechanical phenomenon and arises as a consequence of the Pauli exclusion principle. Intra-atomic exchange (within the atom) is responsible for Hund's rule, which creates special stability to half-filled orbital subshells. Interatomic exchange (between atoms) can be characterized as direct or indirect. Direct exchange involves overlapping electronic wavefunctions, and its magnitude decreases rapidly with distance. The sign of the exchange interaction is related to the strength of the orbital overlap [strong overlap, negative exchange (antiferromagnetic); weak overlap, positive exchange (ferromagnetic)]. Indirect exchange produces an interaction between magnetic moments through an intermediary. In molecules and magnetic insulators, superexchange is the coupling of magnetic moments via bridging ligands. J. B. Goodenough and J. Kanamori developed a series of rules for transition-metal oxides based on the electronic configuration at the metal atoms and the metal-oxide-metal (M-O-M) angle for whether the exchange is ferromagnetic (parallel) or antiferromagnetic (antiparallel). In metals, magnetic moments associated with atoms, such as rare-earth elements, can interact via the spins of the conduction electrons. This mechanism is called RKKY exchange (for Ruderman, Kittel, Kasuya, and Yosida, who developed the model). *See* EXCHANGE INTERACTION.

Examples of magnetism for chemical substances. Molecular oxygen (O_2) has two unpaired electrons in its ground state and is paramagnetic. This is clearly observed when blue, liquid oxygen is poured between the poles of a magnet. Its molar susceptibility over a wide range of pressures and temperatures is $\chi_{Curie} = 0.993/T$. Other paramagnetic molecules include NO, NO_2, ClO_2, and ClO_3.

Most organic compounds are diamagnetic, but there are some, called free radicals, with one or two unpaired electrons. One example is hexaphenylethane, which dissociates in benzene

solution to give the triphenylmethyl radical. Aromaticity assists to stabilize the unpaired electrons. *See* FREE RADICAL.

Numerous monatomic transition-metal complexes, especially those involving the 3*d* metal ions, show paramagnetic behavior. Often the orbital angular momentum is quenched, and these ions exhibit spin-only magnetic moments. The number of unpaired electrons depends on the strength of the ligand field created by the coordination environment, where strong ligand fields generate low-spin configurations, and weak ligand fields generate high-spin configurations. Most tetrahedral complexes show high-spin behavior. The nature of ligand field strength is given by the spectrochemical series: cyanide (CN^-), carbonyl (CO), and nitrosyl (NO^+) are strong-field ligands; iodide (I^-) and bromide (Br^-) are weak-field ligands. In certain cases of octahedral or tetrahedral coordination, there can be some orbital effects via spin-orbit coupling. The 4*d* and 5*d* transition metals often show low-spin configurations and greater effects of ligand orbitals mixing into the valence orbitals, which complicate the magnetic response. Rare-earth ions show significant magnetic behavior which utilizes their total angular momentum. **Figure 5** illustrates the variation in theoretical and experimental magnetic moments for 3*d* transition-metal ions and 4*f* rare-earth ions to show how successful these two interpretations are. *See* LIGAND FIELD THEORY.

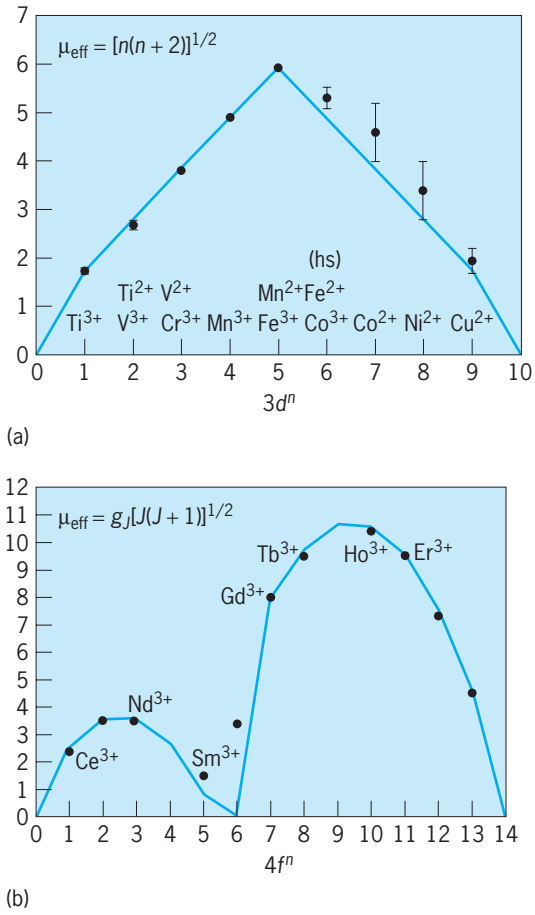

Fig. 5. Plots of effective magnetic moments versus numbers of valence electrons for (*a*) 3*d* ions, where theory uses the spin-only model; and (*b*) 4*f* ions, where theory uses the total angular momentum for each electronic configuration. Lines correspond to theoretically predicted values; dots with error bars indicate experimentally measured values.

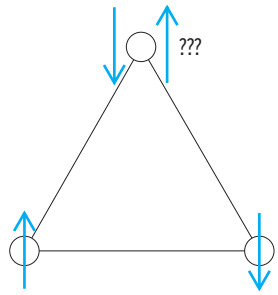

Fig. 6. Diagram of antiferromagnetic frustration.

Clusters of magnetic ions can lead to interesting magnetic behavior because the magnetic exchange interaction can vary with temperature to show high-spin to low-spin transitions. If the exchange interaction is antiferromagnetic and the cluster contains triangles, then frustration arises, as complete antiferromagnetic coupling between magnetic centers cannot be sustained by the structure (**Fig. 6**). Large clusters can lead to superparamagnetism, in which molecules can behave as classical magnets with large magnetic moments.

Metals typically show Pauli paramagnetism. In the presence of an external magnetic field, the energy bands associated with spin-up electrons are separated from the energy bands of the spin-down electrons (Zeeman interaction), resulting in more spin-up than spin-down electrons and an induced, temperature-independent magnetization. This effect is observed for sodium, magnesium, and aluminum (Al becomes superconducting at low temperatures). In poor metals (for example, graphite) and semimetals (such as antimony and bismuth), the induced magnetic field from the trajectories of the conduction electrons in the external magnetic field create a net diamagnetic response called Landau diamagnetism. In anisotropic materials, such as graphite, this diamagnetism shows that the principal susceptibility perpendicular to the hexagonal axis of a single crystal of graphite at room temperature is about -0.5×10^{-6} emu/mol, but parallel with the hexagonal axis, it is -21.5×10^{-6} emu/mol. The enhanced diamagnetism when the field is parallel to the hexagonal axis is thought to arise from the trajectories of the conduction electrons in the graphite planes. Some metallic elements, such as iron, cobalt, nickel, gadolinium, and dysprosium, are ferromagnetic, which means they show net magnetization in the absence of a magnetic field. A few elements are antiferromagnetic, such as manganese and chromium. Recent investigations with possible applications to magnetic storage and retrieval involve half-metallic ferromagnets, in which only the spin-up electrons contribute to electronic conduction.

Gordon J. Miller

Bibliography. R. L. Carlin, *Magnetochemistry*, Springer-Verlag, Berlin, 1986; P. Day (ed.), *Electronic Structure and Magnetism of Inorganic Compounds*, vols. 1–7, Chemical Society, London, 1970–1982; J. B. Goodenough, *Magnetism and the Chemical Bond*, R.E. Krieger Publishing, 1976; A. F. Orchard, *Magnetochemistry*, Oxford University Press, 2003; G. T. Rado and H. Suhl (eds.), *Magnetism: A Treatise on Modern Theory and Materials*, 5 vols., Academic Press, 1963–1973; É. du Trémolet de Lacheisserie, D. Gignoix, M. Schlenker (eds.), *Magnetism: I, Fundamentals, and II, Materials and Applications*, Kluwer Academic, Dordrecht, 2002.

Magnetohydrodynamic power generator

A system for the generation of electrical power through the interaction of a flowing, electrically conducting fluid with a magnetic field. As in a conventional electrical generator, the Faraday principle of motional induction is employed, but solid conductors are replaced by an electrically conducting fluid. The interactions between this conducting fluid and the electromagnetic field system through which power is delivered to a circuit are determined by the magnetohydrodynamic (MHD) equations, while the properties of electrically conducting gases or plasmas are established from the appropriate relationships of plasma physics. Major emphasis has been placed on MHD systems utilizing an ionized gas, but an electrically conducting liquid or a two-phase flow can also be employed. *See* ELECTROMAGNETIC INDUCTION; GENERATOR; MAGNETOHYDRODYNAMICS; PLASMA (PHYSICS).

Principle. Electrical conductivity in an MHD generator can be achieved in a number of ways. At the heat-source operating temperatures of MHD systems (1300–5000°F or 1000–3000 K), the working fluids usually considered are gases derived from combustion, noble gases, and alkali metal vapors. In the case of combustion gases, a seed material such as potassium carbonate is added in small amounts, typically about 1% of the total mass flow. The seed material is thermally ionized and yields the electron number density required for adequate electrical conductivity above about 4000°F (2500 K). With monatomic gases, operation at temperatures down to about 2200°F (1500 K) is possible through the use of cesium as a seed material. In plasmas of this type, the electron temperature can be elevated above that of the gas (nonequilibrium ionization) to provide adequate electrical conductivity at lower temperatures than with thermal ionization. In so-called liquid metal MHD, electrical conductivity is obtained by injecting a liquid metal into a vapor or gas stream to obtain a continuous liquid phase.

The conversion process in the MHD generator itself occurs in a channel or duct in which a plasma flows usually above the speed of sound through a magnetic field. The power output per unit volume W_e is given by the equation below, where σ is the

$$W_e = \sigma v^2 B^2 k(1 - k)$$

electrical conductivity of the gas, v the velocity of the working fluid, B the magnetic flux density, and k the electrical loading factor (terminal voltage/induced emf).

With a conductivity of 10 siemens, a value typical for fossil fuel combustion, $B = 6$ teslas, $v = 600$ m/s = 1970 ft/s, and $k = 0.75$, the power density is 24 MW/m^3 (680 kW/ft^3). Increasing the plasma values to $\sigma = 120$ S and $v = 1000$ m/s = 3280 ft/s and setting $k = 0.5$ yields a power density of 1.08 GW/m^3 (30.6 kW/ft^3). High power densities are one of the attractive features of MHD power generators.

Under the magnetic field strengths required for MHD generators, the plasma displays a pronounced Hall effect. To permit the basic Faraday motional induction interaction and simultaneously support the resulting Hall potential in the flow direction, a linear channel requires segmented walls comprising alternately electrodes (anode or cathode) and insulators. Various connections are possible to combine the outputs of the individual anode-cathode pairs thus formed. These are (1) Faraday, with individual loading of each anode-cathode pair or cell; (2) diagonal loading, with cells series-connected along diagonal equipotentials; and (3) Hall, involving direct loading of the Hall potential with short-circuiting of the individual Faraday anode-cathode pairs. From an electrical machine viewpoint, both individual cells and the complete generator may be regarded as a gyrator. Indeed, the optimum loading of the MHD channel is achieved by extracting power from both the Faraday and Hall terminals, and this is most readily accomplished through consolidation of the dc outputs of individual electrode pairs using power electronics (**Fig. 1**), rather than by the direct physical connections just described. *See* GYRATOR; HALL EFFECT.

For most applications, a superconducting magnet system is needed to provide the 4–6-T field, which is at least twice the value utilized in conventional machines. *See* MAGNET; SUPERCONDUCTING DEVICES.

System considerations. The MHD generator is a heat engine or electromagnetic turbine which

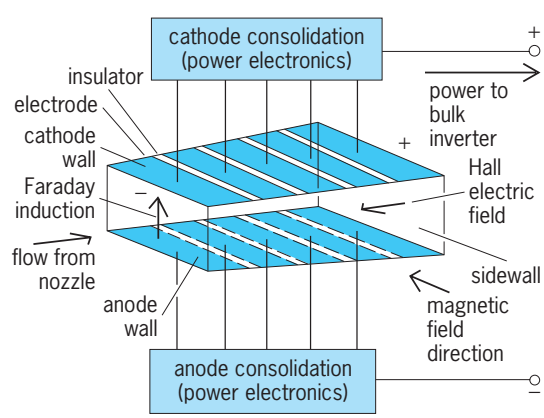

Fig. 1. MHD generator with five electrode segments.

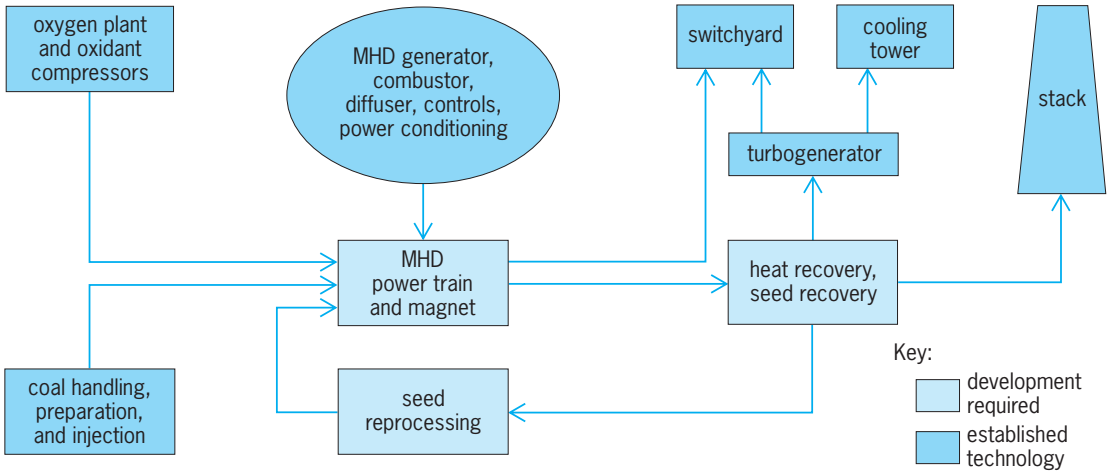

Fig. 2. Major subsystems of an MHD-steam combined-cycle plant.

converts thermal energy to a direct electrical output via the intermediate step of the energy of the flowing working fluid. As such, it must be embedded in a complete power train comprising a heat source, a nozzle for fluid acceleration, the channel itself, a diffuser for pressure recovery, a magnet subsystem, and the power conditioning subsystem to combine the many direct-current outputs resulting from segmentation and match them to the voltage and frequency requirements of the load or system to which the generator is connected. Fuel, oxidizer, and ionizing seed supplies complete the basic MHD power generating system (**Fig. 2**). *See* ENGINE; THERMODYNAMIC PROCESSES.

The gas exhausted from an MHD system generally contains significant recoverable energy. A combined cycle, generally with a steam turbine, is advantageous when the objective, as is the case for electrical utility application, is to take advantage of the higher thermodynamic efficiency available with a 4600°F (2800 K) inlet temperature to the conversion system (Fig. 2).

Improvement of the overall thermal efficiency of central station power plants has been the continuing objective of power engineers. Conventional plants based on steam turbine technology are limited to about 40% efficiency, imposed by a combination of working-fluid properties and limits on the operating temperatures of materials. Application of the MHD interaction to electrical power generation removes the restrictions imposed by the blade structure of turbines and enables the working-fluid temperature to be increased substantially. This enables the working fluid to be rendered electrically conducting and yields a conversion process based on a body force of electromagnetic origin.

When combined with a steam turbine system to serve as the high-temperature or topping stage of a binary cycle, an MHD generator has the potential for increasing the overall plant thermal efficiency to around 50%, and values higher than 60% have been predicted for advanced systems. *See* ELECTRIC POWER GENERATION; STEAM TURBINE.

Types. When this generator is embedded in an overall electrical power generation system, a number of alternatives are possible, depending on the heat source and working fluid selected. The temperature range required by MHD can be achieved through combustion with oxygen or compressed preheated air. The association of MHD with nuclear heat sources has also been considered, but in this case limitations on the temperature of nuclear fission heat sources with solid fuel elements has thus far precluded any practical scheme being developed where a plasma serves the working fluid. The possibility of coupling MHD to a fusion reactor has been explored, and it is possible that twenty-first-century central station power systems will comprise a fusion reactor and an MHD energy conversion system. *See* NUCLEAR FUSION; NUCLEAR REACTOR.

MHD power systems are classified into open- or closed-cycle systems, depending respectively on whether the working fluid is utilized on a once-through basis or recirculated via a compressor. For fossil fuels, the open-cycle system offers the inherent advantage of interposing no solid heat-exchange surface between the combustor and the MHD generator, thus avoiding any limitation being placed on the cycle by the temperature attainable over a long period of operation by construction materials in the heat exchanger. Closed-cycle systems were originally proposed for nuclear heat sources, and the working fluid can be either a seeded noble gas or a liquid metal-vapor mixture.

Features and applications. The greatest development effort in MHD power generation has been applied to fossil-fired open-cycle systems, but sufficient progress has been made in closed-cycle systems to establish their potential and to identify the engineering problems which must be solved before they can be considered practical.

In addition to offering increased power plant efficiencies, MHD power generation also has important potential environmental advantages. These are of special significance when coal is the primary fuel, for it appears that MHD systems can utilize coal directly

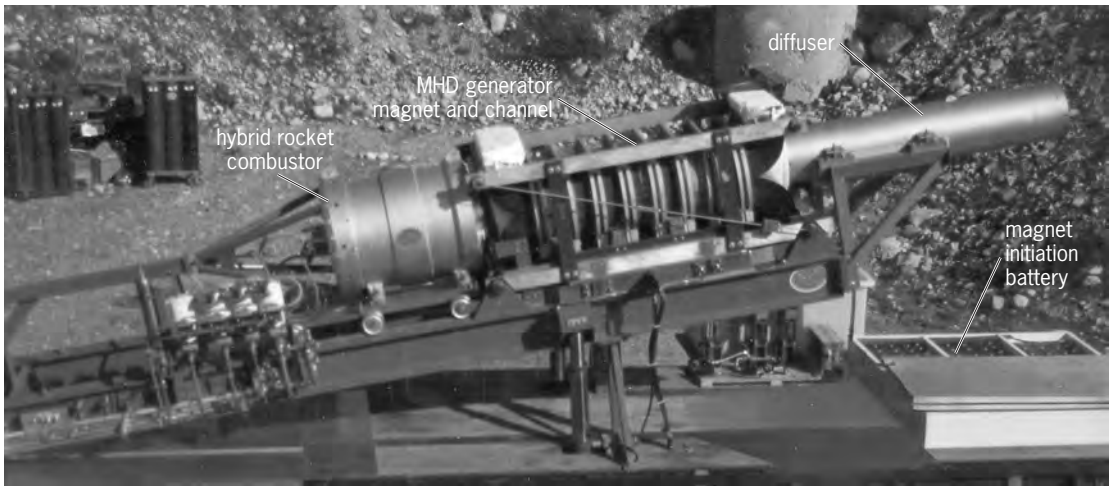

Fig. 3. Self-excited, rocket-motor–driven, pulsed MHD generator. (*STD Research Corp.*)

without the cost and loss of efficiency resulting from the processing of coal into a clean fuel required by competing systems. The use of a seed material to obtain electrical conductivity in the working fluid also places the requirement on the MHD system that a high level of recovery be attained to avoid adverse environmental impact and also to ensure acceptable plant economics. The seed recovery system required by an MHD plant also serves to recover all particulate material in the plant effluent. A further consequence of the use of seed material is its demonstrated ability to remove sulfur from coal combustion products. This occurs because the seed material is completely dissociated in the combustor, and the recombination phenomena downstream of the MHD generator favor formation of potassium sulfate in the presence of sulfur. Accordingly, seed material acts as a built-in vehicle for removal of sulfur. Laboratory experiments have shown that the sulfur dioxide emissions can be reduced to levels experienced with natural-gas–fired plants. A further important consideration is the reduction of the emission of oxides of nitrogen through control of combustion and the design of component operating conditions. Tests at the 30-MW (thermal) Coal-Fired Flow Facility in Tennessee have shown that these emissions can be held to about 10% of the values allowed by the New Source Performance Standards of the Environmental Protection Administration. *See* AIR POLLUTION.

While not a property of the MHD system in itself, the potential of MHD to operate at higher thermal plant efficiencies has the consequence of substantial reduction in thermal waste discharge, following the relationship that the heat rejected per unit of electricity generated is given by $(1 - \eta)/\eta$, where η is the plant efficiency. As the technology of MHD is developed along with that of advanced gas turbines, it is also possible that MHD systems can dispense entirely with the need for large amounts of cooling water for steam condensation through the coupling of MHD generators with closed-cycle gas turbines. *See* GAS TURBINE; WATER POLLUTION.

The high-power-density feature has led to MHD being considered for situations where compact electrical sources, usually of the pulse power type, are required. Energy storage in high-energy fuel, together with the rapid start-up capability of a rocket engine–type combustor, leads to the possibility of pulsed systems (1 s or less) at powers up to at least 1 GW. *See* PLASMA PROPULSION.

An experimental pulse MHD generator of 10-MW electrical output, driven by a solid-fuel rocket motor, is shown in **Fig. 3**. It is self-excited, and the combustion gas plasma has an electrical conductivity greater than 100 S/m.

One application of MHD results from the advantages inherent in using a liquid metal as the heat-transfer medium in a solar central receiver system. The liquid-metal type of MHD is readily coupled to this type of power tower. Liquid-metal systems have also been considered for long-duration space power applications. *See* SOLAR ENERGY; SPACE POWER SYSTEMS.

State of development. Development efforts on MHD have been mainly focused on fossil-fired systems, the fossil fuel selected being determined by national energy considerations. Coal is particularly attractive as a fuel for MHD plants because the basic electromagnetic interaction process is not compromised by ash, slag, or combustion products. Efforts in the United States, China, India, Australia, and the European community have centered on open-cycle systems utilizing this fuel. In Russia, the large reserves of natural gas have led to its extensive use in MHD research and development, although emphasis on coal-fired systems is increasing. Work in Japan was originally based on petroleum-based fuels, but now is also oriented toward coal firing with closed-cycle systems.

The most complete MHD plant constructed to date is the open-cycle U-25 installation near Moscow. As first constructed, it operated with natural gas and achieved a maximum power output of 25 MW and continuous operation for 250 h at low power levels. In both cases, the electrical output was delivered to the Moscow grid. This plant has been extensively

modified and, as the U-25M, resumed operation in 1989.

In the United States, important engineering progress has been achieved on coal-fired systems. A 50-MW coal-fired combustor has been developed, and over 1000 h of operation of an open-cycle channel have been completed. A large experimental generator in Tennessee reached an electrical power above 30 MW. A 6-T superconducting magnet has been constructed and operated by the Argonne National Laboratory. Design studies have dealt with the feasibility of retrofitting existing plants with an MHD generator of about 25-MW electrical output. Electrical power conditioning has been addressed, and a 3.5-MW inverter has been successfully coupled to an MHD generator at a Department of Energy development facility.

Pulsed MHD has also been applied to geophysical research in Russia. Up to 100 MW peak power is delivered in 1–3 s through suitably placed electrodes to the Earth's crust, and electrical conductivity in the surrounding region is recorded by stations which pick up the transmitted pulse. Detection of impending earthquakes and the location of minerals have been achieved by this method. *See* GEOPHYSICAL EXPLORATION; ROCK, ELECTRICAL PROPERTIES OF.

William D. Jackson

Bibliography. H. K. Messerle, *Magnetohydrodynamic Electrical Power Generation*, 1995; R. J. Rosa, *Magnetohydrodynamic Energy Conversion*, 1968, reprint 1987; P. E. Sens (ed.), *Coal-Fired Magnetohydrodynamic (MHD) Electric Power Generation*, 1992.

Magnetohydrodynamics

The interaction of electrically conducting fluids with magnetic fields. The fluids can be ionized gases (commonly called plasmas) or liquid metals. Magnetohydrodynamic (MHD) phenomena occur naturally in the Earth's interior, constituting the dynamo that produces the Earth's magnetic field; in the magnetosphere that surrounds the Earth; and in the Sun and throughout astrophysics. In the laboratory, magnetohydrodynamics is important in the magnetic confinement of plasmas in experiments on controlled thermonuclear fusion. Magnetohydrodynamic principles are also used in plasma accelerators for ion thrusters for spacecraft propulsion, for light–ion-beam powered inertial confinement, and for magnetohydrodynamic power generation. Magnetic fields can also be spontaneously generated in inertial confinement experiments, making magnetohydrodynamics relevant in this area. *See* COSMIC RAYS; GEOMAGNETISM; ION PROPULSION; MAGNETOHYDRODYNAMIC POWER GENERATOR; MAGNETOSPHERE; NUCLEAR FUSION; PLASMA (PHYSICS); SOLAR WIND; SUN.

The conducting fluid and magnetic field interact through electric currents that flow in the fluid. The currents are induced as the conducting fluid moves across the magnetic field lines. In turn, the currents influence both the magnetic field and the motion of the fluid. Qualitatively, the magnetohydrodynamic interactions tend to link the fluid and the field lines so as to make them move together. *See* ELECTRIC CURRENT.

The generation of the currents and their subsequent effects are governed by the familiar laws of electricity and magnetism. The motion of a conductor across magnetic lines of force causes a voltage drop or electric field at right angles to the direction of the motion and the field lines; the induced voltage drop causes a current to flow as in the armature of a generator.

The currents themselves create magnetic fields which tend to loop around each current element. The currents heat the conductor and also give rise to mechanical ponderomotive forces when flowing across a magnetic field. (These are the forces which cause the armature of an electric motor to turn.) In a fluid, the ponderomotive forces combine with the pressure forces to determine the fluid motion. *See* ELECTRICITY; GENERATOR; MAGNETISM; MOTOR.

Magnetohydrodynamic phenomena involve two well-known branches of physics, electrodynamics and hydrodynamics, with some modifications to account for their interplay. The basic laws of electrodynamics as formulated by J. C. Maxwell apply without any change. However, Ohm's law, which relates the current flow to the induced voltage, has to be modified for a moving conductor. *See* ELECTRODYNAMICS; HYDRODYNAMICS; MAXWELL'S EQUATIONS; OHM'S LAW.

It is useful to consider first the extreme case of a fluid with a very large electrical conductivity. Maxwell's equations predict, according to H. Alfvén, that for a fluid of this kind the lines of the magnetic field **B** move with the material. The picture of moving lines of force is convenient but must be used with care because such a motion is not observable. It may be defined, however, in terms of observable consequences by either of the following statements: (1) a line moving with the fluid, which is initially a line of force, will remain one; or (2) the magnetic flux through a closed loop moving with the fluid remains unchanged.

If the conductivity is low, this is not true and the fluid and the field lines slip across each other. This is similar to a diffusion of two gases across one another and is governed by similar mathematical laws.

As in ordinary hydrodynamics, the dynamics of the fluid obeys theorems expressing the conservation of mass, momentum, and energy. These theorems treat the fluid as a continuum. This is justified if the mean free path of the individual particles is much shorter than the distances that characterize the structure of the flow. Although this assumption does not generally hold for plasmas, one can gain much insight into magnetohydrodynamics from the continuum approximation. The ordinary laws of hydrodynamics can then easily be extended to cover the effect of magnetic and electric fields on the fluid by adding a magnetic force to the momentum-conservation equation and electric heating and work to the energy-conservation equation.

Ideal magnetohydrodynamics. Ideal magnetohydrodynamics refers to the study of magnetohydrodynamics for a perfectly conducting fluid, that is, one obeying the form of Ohm's law, given by Eq. (1), where

$$0 = \mathbf{E} + \mathbf{v} \times \mathbf{B} \tag{1}$$

E is the electric field, **B** the magnetic field, and **v** the center-of-mass velocity of the fluid. This equation means that the electric field in the rest frame of the fluid is zero. *See* ELECTRIC FIELD; RELATIVISTIC ELECTRODYNAMICS.

Frozen flow. Equation (1) can be combined with Faraday's law, Eq. (2), to give Eq. (3), which shows

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2}$$

$$\begin{aligned} \frac{d}{dt}\int \mathbf{B}\cdot d\mathbf{S} &= \int \frac{\partial \mathbf{B}}{\partial t}\cdot d\mathbf{S} + \int \mathbf{B}\cdot\frac{\partial}{\partial t}d\mathbf{S} \\ &= -\int \nabla\times\mathbf{E}\cdot d\mathbf{S} + \oint \mathbf{B}\cdot\mathbf{v}\times d\mathbf{l} \\ &= -\oint(\mathbf{E}+\mathbf{v}\times\mathbf{B})\cdot d\mathbf{l} = 0 \end{aligned} \tag{3}$$

that the magnetic flux, $\int \mathbf{B}\cdot d\mathbf{S}$, where $d\mathbf{S}$ is an element of area in the fluid, is conserved in the frame of reference of the fluid. *See* CALCULUS OF VECTORS; FARADAY'S LAW OF INDUCTION.

It can thus be said that the magnetic field and the fluid are frozen together. This result also implies that magnetic field lines immersed in a perfectly conducting fluid cannot change their topology.

Alfvén waves. If a uniform plasma of density ρ_0 is immersed in a uniform magnetic field $\mathbf{B}_0$, then if there is a periodic transverse displacement of the frozen-in plasma and magnetic field (**Fig. 1**), it can be shown that the displacement propagates as a wave with a velocity $c_A = (B^2{}_0/\mu_0\rho_0)^{1/2}$, where μ_0 is the permeability of empty space. This is called an Alfvén wave. Physically this is analogous to a wave on a loaded string, the magnetic field lines being the string with tension $B^2{}_0/\mu_0$, and the mass loading of the string being ρ_0 per unit volume [Eq. (16)]. This result can easily be derived by using perturbation techniques.

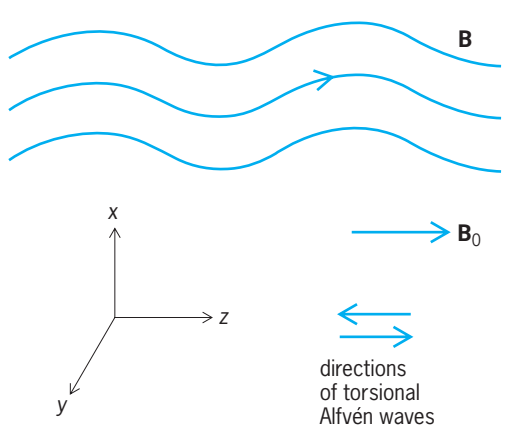

Fig. 1. Periodic transverse displacement of a uniform magnetic field B_0 and a uniform, frozen-in plasma.

The equation of motion for the plasma is Eq. (4),

$$\rho\frac{d\mathbf{v}}{dt} = -\nabla p + \mathbf{J}\times\mathbf{B} \tag{4}$$

where p is the pressure and **J** is the current density. For the situation shown in Fig. 1 this becomes Eq. (5),

$$\rho_0\frac{\partial v_{x1}}{\partial t} = J_{y1}B_{z0} \tag{5}$$

where J_{y1} is the perturbed current density and v_{x1} is the x-component of the perturbed velocity. The plasma pressure p does not contribute to Eq. (5) for purely transverse motion. Ampère's law (neglecting the displacement current which is negligible at such low frequencies) is Eq. (6), and relates the per-

$$\nabla\times\mathbf{B} = \mu_0\mathbf{J} \tag{6}$$

turbation current of the wave with the perturbed magnetic field through Eq. (7). Combining Eqs. (1)

$$\frac{\partial B_{x1}}{\partial z} = \mu_0 J_{y1} \tag{7}$$

and (2) yields Eq. (8), and in perturbation form this becomes Eq. (9). Differentiating Eq. (9) with respect to time t and employing Eqs. (5) and (7) gives Eq. (10), which is the wave equation for a wave prop-

$$\frac{\partial\mathbf{B}}{\partial t} = \nabla\times(\mathbf{v}\times\mathbf{B}) \tag{8}$$

$$\frac{\partial B_{x1}}{\partial t} = B_{z0}\frac{\partial v_{x1}}{\partial z} \tag{9}$$

$$\frac{\partial^2 B_{x1}}{\partial t^2} - \frac{B_{z0}^2}{\mu_0\rho_0}\frac{\partial^2 B_{x1}}{\partial z^2} = 0 \tag{10}$$

agating in the $\pm z$ direction with a wave velocity c_A, the Alfvén speed. *See* ALFVÉN WAVES; AMPÈRE'S LAW; WAVE MOTION.

Magnetosonic waves. For the more general case when the wave vector **k** is at some arbitrary angle θ to the equilibrium magnetic field $\mathbf{B}_0$, the plasma pressure also contributes to the restoring force of the waves because, in general, compression of the plasma now occurs. Therefore, in addition to the Alfvén wave there are two magnetosonic waves with phase velocity given by Eq. (11), where $c_s = (\gamma p_0/\rho_0)^{1/2}$ is

$$v_\pm = \left\{{}^1/_2\left[c_A^2 + c_s^2 \pm\sqrt{\left(c_A^2+c_s^2\right)^2 - 4c_A^2c_s^2\cos^2\theta}\right]\right\}^{1/2} \tag{11}$$

the sound speed, γ being the ratio of specific heats. The plus sign refers to the fast magnetosonic wave in which the plasma pressure perturbations and the magnetic force perturbations act in phase, while the minus sign refers to the slow magnetosonic wave in which the two forces are out of phase. The first (+) magnetosonic wave is the only one that can propagate perpendicular to $\mathbf{B}_0$, and then both the magnetic pressure $B_0^2 2\mu_0$ [Eq. (16)] and the plasma pressure p_0 contribute to the restoring force of this compressional longitudinal wave, and it propagates with a velocity $(c_A^2 + c_S^2)^{1/2}$. *See* SOUND.

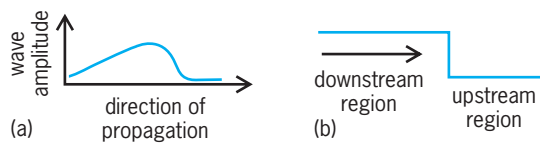

Fig. 2. Evolution of a large-amplitude wave. (*a*) Steepening of wave. (*b*) Formation of shock wave.

Shock waves. When the amplitude of a wave is large, it tends to steepen in the direction of propagation (**Fig. 2***a*). The reason is that the local phase velocity increases where the temperature and magnetic field are higher. The steepening process increases with wave amplitude until a sharp shock (Fig. 2*b*) occurs. The shock thickness and structure are determined by dissipative processes such as viscosity, thermal conduction, and electrical resistivity, discussed below. Conservation laws for mass, momentum, energy, and magnetic flux can be applied in the frame of reference of the shock front in which a steady state can be assumed. For a shock in which the magnetic field is transverse to the direction of propagation, the conservation relations between the upstream region (indicated by subscript 1) and downstream region (indicated by subscript 2) are given by Eqs. (12). The

$$\begin{aligned} \rho_1 v_1 &= \rho_2 v_2 \\ \rho_1 v_1^2 + p_1 + \frac{B_1^2}{2\mu_0} &= \rho_2 v_2^2 + p_2 + \frac{B_2^2}{2\mu_0} \\ {}^1\!/_2 \rho_1 v_1^3 + \frac{\gamma}{\gamma-1} p_1 v_1 + \frac{B_1^2 v_1}{\mu_0} &= {}^1\!/_2 \rho_2 v_2^3 + \frac{\gamma}{\gamma-1} p_2 v_2 + \frac{B_2^2 v_2}{\mu_0} \\ v_1 B_1 &= v_2 B_2 \end{aligned} \tag{12}$$

first three of these equations describe conservation of mass, momentum, and energy in that order, the fourth relation being conservation of magnetic flux, derived from Eq. (8). *See* CONSERVATION OF ENERGY; CONSERVATION OF MASS; CONSERVATION OF MOMENTUM.

The shock velocity v_1 in the laboratory frame can be found from Eqs. (12) in terms of the density ratio ρ_1/ρ_2 from Eq. (13). An important result from this is

$$v_1^2 = \frac{2}{\gamma+1} \cdot \frac{c_s^2 + c_A^2\left[1 + \left(1 - {}^1\!/_2\gamma\right)(\rho_2/\rho_1 - 1)\right]}{\rho_1/\rho_2 - (\gamma-1)/(\gamma+1)} \tag{13}$$

that the maximum density ratio that can occur in a single shock compression is a factor of 4 for $\gamma = {}^5\!/_3$. For this reason the high compression that is required in inertial confinement fusion cannot be attained by a single shock wave, but requires a tailored compression wave that is equivalent to many weaker shocks following each other.

More generally, it is possible to drive a current in the shock front which causes other components of the magnetic field to be generated or removed. Such shocks are called switch-on or switch-off shocks, respectively. *See* SHOCK WAVE.

Collisionless shock. At low density and high temperature the mean free path of electrons and ions is too large for collisional processes to determine the shock structure. Instead, within the shock structure microscopic kinetic instabilities occur because the relative drift velocity of electrons and ions (associated with the net current density in the shock) exceeds some critical value such as the ion sound speed, c_S [Eq. (11)]. These instabilities grow to give a turbulent state yielding corresponding anomalous transport coefficients. *See* TRANSPORT PROCESSES; TURBULENT FLOW.

Magnetostatic equilibria. If the center-of-mass velocity **v** is very small, a stationary equilibrium can exist in which the pressure gradient in Eq. (4) is balanced by the magnetomotive force, $\mathbf{J} \times \mathbf{B}$; that is, Eq. (14)

$$\nabla p = \mathbf{J} \times \mathbf{B} \tag{14}$$

is satisfied. This nonlinear equation together with Eq. (6) represents the set of equations describing magnetostatic equilibria.

Two important results follow from Eq. (14). Taking the scalar product first with **B** and then with **J** yields Eqs. (15), which shows that the pressure is constant

$$\mathbf{B} \cdot \nabla p = 0 \qquad \mathbf{J} \cdot \nabla p = 0 \tag{15}$$

along both a magnetic field line and the current density direction. Applying this to the case of a toroidal axisymmetric equilibrium (**Fig. 3**), it is found that magnetic field lines with toroidal B_ϕ and poloidal B_p components form nested surfaces on each of which the pressure is constant. The current density **J** lies in the plane of each magnetic surface and is related to **B** through Eq. (6). Thus Eq. (14) can be written as Eq. (16), where $\perp$ signifies the component per-

$$\begin{aligned} \nabla p &= \frac{(\nabla \times \mathbf{B}) \times \mathbf{B}}{\mu_0} \\ &= -\nabla_\perp \left(\frac{B^2}{2\mu_0}\right) + \frac{B^2}{\mu_0}(\mathbf{b} \cdot \nabla)\mathbf{b} \end{aligned} \tag{16}$$

pendicular to the magnetic field, and $\mathbf{b} = \mathbf{B}/|B|$ is the unit vector in the direction of the magnetic field. Then it is possible to identify the magnetic pressure, $B^2/(2\mu_0)$, and the tension in the magnetic field lines, B^2/μ_0, which causes a force through the curvature of the lines of force, $(\mathbf{b} \cdot \nabla)\mathbf{b}$. The quantity $2\mu_0 p/B^2$ defines β, the ratio of plasma pressure to magnetic pressure.

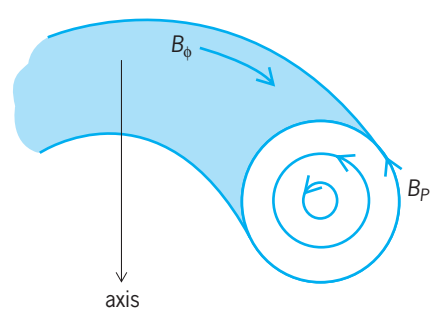

Fig. 3. Perfectly conducting fluid in toroidal axisymmetric equilibrium.

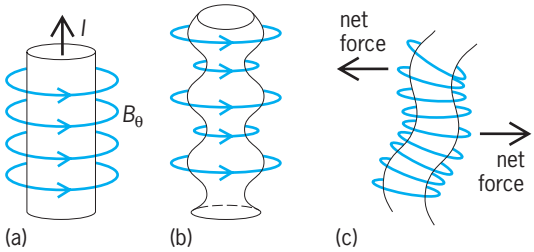

Fig. 4. Instabilities of a Z-pinch configuration. (*a*) Initial configuration of magnetically confined plasma. (*b*) Sausage instability. (*c*) Kink instability.

The solution of Eq. (16) for the axisymmetric large aspect ratio toroidal equilibrium shown in Fig. 3 gives nested circular magnetic surfaces that are not concentric; the inner ones are shifted out, away from the major axis of the torus, so that the surfaces are closer together on the outside of the torus. The physical reason for this is that the toroidal current I_θ which produces the poloidal magnetic field B_p flows as in a hoop, leading to an outward force on the plasma. (Each current segment experiences a repulsive force from the oppositely flowing current in the segment opposite it.) This toroidal configuration is the basis of the tokamak approach to magnetic fusion energy.

Stability of ideal plasmas. An equilibrium state as described above is not necessarily stable. The most elementary configuration of a magnetically confined plasma is a Z-pinch (**Fig. 4***a*) in which a cylindrical column carrying an axial current I is confined by the self-magnetic field produced by the current I itself. This was the earliest concept for confinement of plasma for controlled fusion. Consider an azimuthally symmetric perturbation (Fig. 4*b*) that makes the configuration look rather like a string of sausages. The magnetic field B_θ at the surface of radius a is $\mu_0 I/2\pi a$, so that the magnetic field is stronger where the radius a is narrower, and vice versa. Hence the plasma column (unless it is very diffuse) can be unstable to this perturbation because of the enhanced $\mathbf{J} \times \mathbf{B}$ force where a is small. *See* PINCH EFFECT.

Another type of instability is the kink mode (Fig. 4*c*) where the column is bent one way and the other in an alternating way. Again the driving force for the instability can be seen from an examination of the magnetic lines of force. In the kink mode they are concentrated together on the concave side and are further apart on the convex side of each kink. This leads to a net force (as shown in Fig. 4*c*) which increases the amplitude of the kink.

Such instabilities have been seen experimentally, and led to the introduction of another magnetic field component, an axial magnetic field B_z, in order to stabilize the configuration. With both B_z and B_θ components present, the magnetic field lines are now helices. If the pitch of the helices varies across the radius of the plasma column, then this shear of the field can stabilize the discharge against localized modes.

The necessary stability criterion is given by Eq. (17), where r is the distance from the axis of the column and μ, the pitch, is defined by Eq. (18).

$$\left(\frac{1}{\mu}\frac{d\mu}{dr}\right)^2 + \frac{8\mu_0}{rB_z{}^2}\frac{dp}{dr} \geq 0 \tag{17}$$

$$\mu = \frac{B_\theta}{rB_z} \tag{18}$$

For a confined plasma the second term involving the pressure gradient, dp/dr, is always a negative contribution, and so this condition determines the shear $d\mu/dr$ required for stability. When the configuration is bent into a torus, the second term is multiplied by a factor $(1 - q^2)$, where q, the safety factor, is rB_ϕ/RB_θ; this leads to stability of these internal modes for $q > 1$. Here R is the major radius of the torus, and B_ϕ is the toroidal component of magnetic field.

A purely hydrodynamic instability, found also in dynamic Z-pinches and exploding supernovae, but also of vital concern in the concept of inertial confinement, is the Rayleigh-Taylor instability. The original problem concerned the stability of a heavy fluid above a light fluid under gravitational acceleration g. A small perturbation of the interface (**Fig. 5**) is clearly unstable, and one with a wave number k (where the wavelength $\lambda = 2\pi/k$) can be shown to have a growth rate $\gamma = (kg)^{1/2}$. Short wavelengths thus have the highest growth rates, and in inertial confinement the shell thickness is likely to determine the most dangerous mode. The reason why the Rayleigh-Taylor instability applies here is that the effect of g is the same as a fluid acceleration in the opposite direction.

Energy principle. Two theoretical methods have been developed for examining the linear stability of an equilibrium configuration; one is a normal mode analysis which directly determines the growth rate, and the other uses an energy principle. When more complex physical processes such as the nonvanishing ion Larmor radius or the Hall effect are included, the former method is applicable; but where the complexity is entirely in the geometrical configuration of the magnetically confined plasma, the energy principle is the more appropriate technique. In ideal magnetohydrodynamics, Eqs. (1), (2), (4), and (6) can be combined to linearly relate the acceleration $\ddot{\boldsymbol{\xi}}$ of a fluid element to the force $\mathbf{F}(\boldsymbol{\xi})$, where $\boldsymbol{\xi}$ is the local fluid displacement and the two dots above $\boldsymbol{\xi}$ means the second derivation with respect to time. Because of the adjointness of $\mathbf{F}$, the stability is ensured if the change in potential energy associated with the

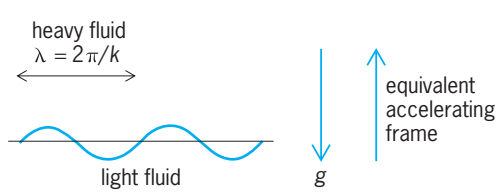

Fig. 5. Rayleigh-Taylor instability, involving a perturbation in the interface between a light fluid and a heavy fluid above it.

displacement, given by Eq. (19), is positive, where

$$\delta W = -\frac{1}{2}\int \ddot{\xi} \cdot \mathbf{F} d^3 r \tag{19}$$

the integral is taken over the entire volume of the configuration. Equation (17) was derived from the energy principle. [By adjointness is meant that **F** satisfies Eq. (20), where η obeys the same boundary conditions as ξ.]

$$\int \xi \cdot F(\eta)\, d^3 r = \int \eta \cdot \mathbf{F}(\xi) d^3 r \tag{20}$$

Thus the rate of increase of kinetic energy (that is, the rate of doing work) is $\int \xi \cdot \mathbf{F}(\xi) d^3 r$. By virtue of conservation of total energy and the adjointness of $\mathbf{F}(\xi)$, it follows that the corresponding change in potential energy δW is given by Eq. (19). If the kinetic energy is increasing through growth of the instability, there is a corresponding decrease in potential energy. Conversely, if δW is positive, the system is stable. More strictly, if δW is found to be positive for all allowable displacements ξ, the plasma is stable; but if δW is negative for any allowable displacement ξ, the plasma is unstable.

The most simple example of the use of the energy principle is the interchange mode. **Figure 6** illustrates two flux tubes each having a magnetic flux $\delta\phi = B\delta A$, where δA is the local area of cross section. The volume of a flux tube, δV, is the integral $\int \delta A dl$, where dl is an element of length along the tube, and therefore the plasma thermal energy in a flux tube is given by Eq. (21). When the plasma on the two

$$\frac{p}{\gamma - 1} dV = \frac{p}{\gamma - 1}\delta\phi \int \frac{dl}{B} \tag{21}$$

flux tubes is interchanged, δW is found to be positive if $\int dl/B$ decreases with the spatially decreasing plasma pressure as one moves toward the walls of the confining vessel. This leads to the concept of a minimum-B confinement configuration, that is, one in which the magnitude of the magnetic field is a minimum at the center of the plasma and increases in all directions away from the center.

Resistive magnetohydrodynamics. An electrically conducting fluid acquires new properties when its resistivity is taken into account.

Stability of resistive plasmas. While the resistivity η of thermonuclear plasmas at a temperature of 10 keV is very low (approximately 10^{-9} ohm·m), it is nevertheless sufficient to permit the magnetic field configuration to change its topology in a thin layer inside which the wave number **k** of the perturbation mode

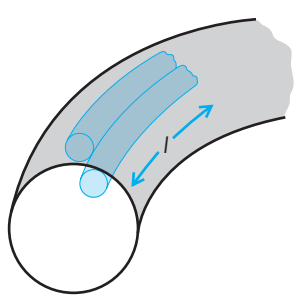

Fig. 6. Two flux tubes in a magnetic field configuration.

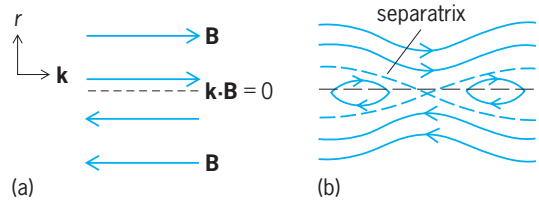

Fig. 7. Effect of nonvanishing resistivity on plasma stability. (*a*) Location of resonant surface k · B = 0. (*b*) Breaking and reconnection of magnetic field lines about the resonant surface in a resistive plasma.

and the magnetic field vector **B** satisfy the condition $\mathbf{k} \cdot \mathbf{B} = 0$. The surface that satisfies this condition is called a resonant surface since the magnetic field lies in a plane of constant phase of the plasma displacement. In a torus, such resonant surfaces occur at rational values of q, the safety factor.

Figure 7*a* illustrates the plane containing **k** and the coordinate r perpendicular to the magnetic surface, surrounding the resonant surface, $\mathbf{k} \cdot \mathbf{B} = 0$. With nonvanishing resistivity η, Ohm's law is given by Eq. (22), replacing Eq. (1). With the presence of

$$\eta \mathbf{J} = \mathbf{E} + \mathbf{v} \times \mathbf{B} \tag{22}$$

the $\eta\mathbf{J}$ term, the magnetic field lines can break and reconnect to form a different configuration consisting of magnetic islands centered on the $\mathbf{k} \cdot \mathbf{B} = 0$ surface. There are corresponding X-type and O-type neutral points (Fig. 7*b*), the former lying at the intersection of the separatrix and the $\mathbf{k} \cdot \mathbf{B} = 0$ surface. The growth of such a magnetic island structure to a finite size is of major importance in understanding the magnetohydrodynamic behavior of toroidal magnetically confined plasma, particularly with regard to disruptions in tokamaks. The phenomenon of magnetic reconnection is also important in the evolution of solar flares. The nonlinear evolution of the new closed magnetic lines topology, called magnetic islands, is the subject of much research. For example, the Sweet-Parker model suggests that the open field lines drift slowly in toward the $\mathbf{k} \cdot \mathbf{B} = 0$ plane, while the plasma flows in the $\pm\mathbf{k}$ direction within the reconnection layer have an amplitude of the order of the Alfvén velocity. It is also possible that the high current density induced in the reconnection layer will, as in collisionless shocks, lead to the triggering of microinstabilities with an associated anomalous resistivity, resulting in a fast reconnection. This is particularly applicable to astrophysical reconnection processes.

Helicity. The possibility of magnetic reconnection even when the resistivity is small but nonzero allows the overriding thermodynamic tendencies to dominate. Hence, for example, in the toroidal reversed-field pinch, the plasma and magnetic field can relax to a state of minimum energy. In doing so, a three-dimensional dynamo occurs as a result of the resistive MHD instabilities and the associated plasma flows, causing the toroidal magnetic field to peak on the minor axis and to reverse in direction at the plasma edge. During this process, it can be reasonably postulated that the magnetic helicity. $\int \mathbf{A} \cdot \mathbf{B}\, dV$ is approximately conserved. Here **A** is the magnetic

vector potential defined by $\mathbf{B} = \nabla \times \mathbf{A}$. The reason for this conservation is that the local helicity density $\mathbf{A} \cdot \mathbf{B}$ obeys Eq. (23), where Φ is the electrostatic

$$\frac{\partial}{\partial t}(\mathbf{A} \cdot \mathbf{B}) + \nabla(\Phi\mathbf{B} + \mathbf{E} \times \mathbf{A}) = -2\mathbf{E} \cdot \mathbf{B} \qquad (23)$$

potential. If the flux of helicity, $(\Phi\ \mathbf{B} + \mathbf{E} \times \mathbf{A})$, is zero, then the source or sink of helicity is the term $-2\mathbf{E} \cdot \mathbf{B}$, which is small everywhere for a plasma of high conductivity, and is only significant in the reconnection layer associated with resistive MHD instabilities. (In evaluating helicity, care must be taken with the choice of gauge for $\mathbf{A}$.)

Diffusion. Finite resistivity also allows the plasma to diffuse across magnetic field lines. Classical diffusion of a low β plasma across an axial magnetic field B_z can be described in terms of the pressure balance of Eq. (14), which here becomes Eq. (24). In this

$$\frac{\partial p}{\partial r} = J_\theta B_z \qquad (24)$$

the azimuthal current density is driven by the radial diffusion velocity υ_r through Ohm's law, Eq. (22), which here becomes Eq. (25). Thus, the flux of par-

$$\eta J_\theta = -\upsilon_r B_z \qquad (25)$$

ticles of density n in an isothermal plasma is given by Eq. (26). This is Fick's law of diffusion, with a dif-

$$n\upsilon_r = -\frac{\eta p}{B^2}\frac{\partial n}{\partial r} \qquad (26)$$

fusion coefficient $D = \eta p/B^2$. The B^{-2} dependence is characteristic of classical diffusion. Plasmas that are microscopically turbulent tend to obey Bohm diffusion where the diffusion coefficient $k_B T/(16eB)$ [where k_B is Boltzmann's constant, T the absolute temperature, and e the electronic charge] is larger by a factor of $\omega_e\tau_e/16$; ω_e is the electron cyclotron frequency, τ_e is the electron-ion collision time, and the product, $\omega_e\tau_e$, is called the Hall parameter, being equal to $B/(ne\eta)$. *See* DIFFUSION; ELECTRON MOTION IN VACUUM; HALL EFFECT; PARTICLE ACCELERATOR; RELAXATION TIME OF ELECTRONS.

In a toroidal configuration with high safety factor q, greater than 1 (usually called a tokamak), classical diffusion is enhanced by two effects. The first is connected with the magnetohydrodynamic equilibrium itself, as defined by Eq. (14), in which, to balance the outward hoop force of the toroidal current itself, a vertical magnetic field component is required which causes the magnetic surfaces to be shifted outward. Consistent with this, the toroidal current must increase with increasing distance R from the major axis. However, the applied electric field E_ϕ (equal to $-\partial A_\phi/\partial t$, where A_ϕ is the toroidal component of the electromagnetic vector potential) varies as R^{-1}. Therefore an outward diffusion velocity υ_R occurs, so that the current density J_ϕ in Ohm's law, here given by Eq. (27), has the required spatial depen-

$$\eta J_\phi = E_\phi + (\mathbf{v} \times \mathbf{B})_\phi \qquad (27)$$

dence. This diffusion is $2q^2$ times the classical value. *See* POTENTIALS.

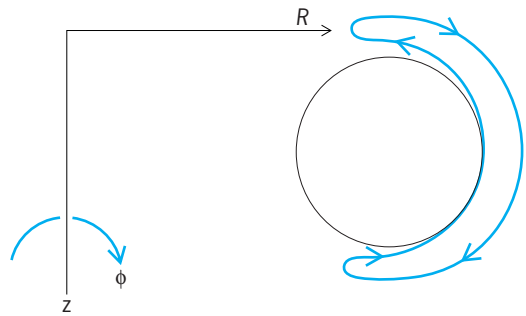

Fig. 8. Projection of a guiding center orbit onto a poloidal plane of a tokamak.

The second effect is important at high temperature when the plasma is relatively collisionless, and when the trajectories of particles have large excursions from magnetic surfaces. Because the toroidal magnetic field varies as $1/R$, the magnitude of the total field is greater near the inside of the torus. A guiding center for the motion of a particle can therefore be reflected near the inside of the torus and, in continuing its drift motion, executes a trajectory which is banana-shaped when projected onto a poloidal plane (**Fig. 8**). The width of the "banana" is $2m\upsilon_{||}/(ZeB_p)$, where B_p is the poloidal component of the magnetic field, m and Z are the particle's mass and charge number, and $\upsilon_{||}$ is the component of its velocity parallel to the magnetic field. This width is much greater than the Larmor radius $m\upsilon_\perp/(ZeB_\phi)$, and when such particles have even only weak collisions, an enhanced diffusion occurs; here $\upsilon_\perp$ is the component of velocity perpendicular to the magnetic field. *See* CHARGED PARTICLE OPTICS.

Magnetic Reynolds number. The relative strength of the $\mathbf{v} \times \mathbf{B}$ and $\eta\mathbf{J}$ terms in Ohm's law can be written with the aid of Eq. (6) as the ratio R_M given by Eq. (28),

$$R_M = \frac{\mu_0 \upsilon L}{\eta} \qquad (28)$$

the magnetic Reynolds number, where L is a characteristic length. When R_M is much greater than 1, the magnetic field is frozen to the conducting fluid, while for R_M much less than 1 the fluid very easily penetrates an applied magnetic field (as in magnetohydrodynamic power generators). When resistive damping is included in the theory of Alfvén waves and magnetohydrodynamic stability theory, a corresponding dimensionless parameter, the Lundquist parameter, given by Eq. (29), occurs, where λ is the wavelength.

$$S = \frac{\mu_0 c_A \lambda}{\eta} \qquad (29)$$

When S is much greater than 1 the resistive damping or thickness of the tearing layer is small, while for S less than 1 no Alfvén wave propagation occurs. *See* DIMENSIONLESS GROUPS.

Electrically conducting fluids. Electrically conducting fluids are either liquid metals or plasmas.

Liquid metals. Mercury and liquid sodium are two familiar examples of electrically conducting, incompressible fluids. The core of the Earth is also

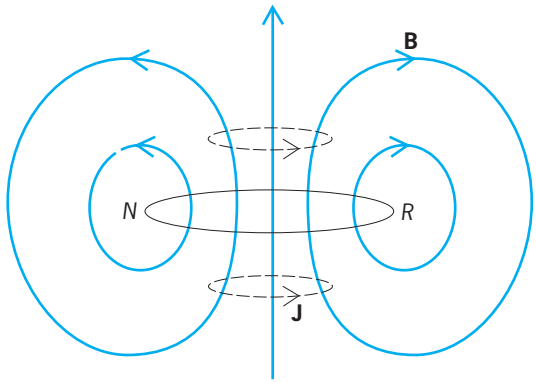

Fig. 9. Axisymmetric, two-dimensional, liquid-metal system in which only an azimuthal current and poloidal magnetic field components exist.

interpreted to be in the liquid-metal state and able to generate currents and magnetic fields. In particular, in a three-dimensional interaction involving the Coriolis force and the $\mathbf{v} \times \mathbf{B}$ force associated with the Earth's rotation, a dynamo action is postulated to account for the Earth's magnetic field. A similar mechanism could generate the solar magnetic field. *See* CORIOLIS ACCELERATION.

The necessity of involving a three-dimensional model for dynamo action can be seen from consideration of an axisymmetric, two-dimensional system (**Fig. 9**), in which only an azimuthal current and poloidal magnetic field components exist. It has been shown that such a system cannot be sustained by fluid motion alone in a steady state. The argument is as follows: the azimuthal electric field must vanish in a steady state so that $\eta\mathbf{J} = \mathbf{v} \times \mathbf{B}$; therefore the neutral ring NR where $\mathbf{B}$ is zero will also have $\mathbf{J} = 0$ and hence $\nabla \times \mathbf{B} = 0$. If this condition is satisfied—for example, if, in the vicinity of the neutral ring, $|B| \propto r^{\alpha}$ for $\alpha > 1$, where r is the distance from the neutral ring—then Ampère's law gives $J \propto r^{\alpha-1}$, but Ohm's law then yields $\upsilon \propto 1/r$ which is an impossible singularity.

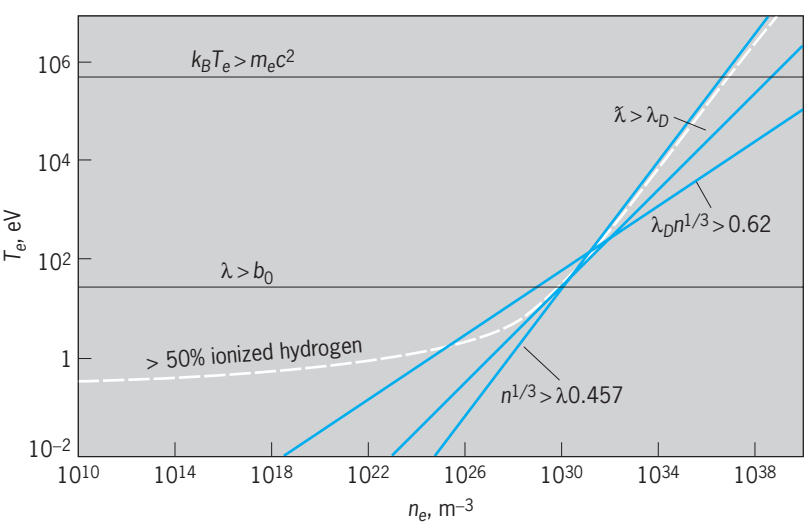

Fig. 10. Plot of electron temperature T_e and density n_e in a plasma, showing values of these parameters where various types of behavior occur.

Another theorem for the same axisymmetric system is as follows: Because the current is purely azimuthal, the curl of $\mathbf{J}$ has no azimuthal component. For uniform resistivity, this means that the azimuthal component of curl $(\mathbf{v} \times \mathbf{B})$ vanishes, which, with the Maxwell equation (30), gives $\mathbf{B} \cdot \nabla\Omega = 0$, where

$$\nabla \cdot \mathbf{B} = 0 \tag{30}$$

$r\Omega = \upsilon_\theta$. This means that the fluid at all points on a magnetic field line rotates about the axis with a uniform angular velocity Ω, r being the distance from the axis.

Plasmas. Most work in magnetohydrodynamics has concerned ionized gases or plasmas. Plasma physics concerns the collective interaction of ions and electrons in a usually quasi-neutral system. In such a plasma there are many particles in a Debye sphere. This follows from the condition that for the gas to remain fully ionized, the mean kinetic energy k_BT_e (where k_B is Boltzmann's constant, and T_e is the electron temperature) must be larger than the mean potential energy $n_e^{1/3}e^2/\epsilon_0$, where ϵ_0 is the permittivity of empty space, and n_e is the number of electrons per cubic meter, so that the average distance between particles is $n^{-1/3}{}_e$. Raising this condition to the power $^3/_2$ yields the condition $n_e\lambda^3{}_D \gg 1$, where λ_D (the Debye length) is $[\epsilon_0 k_B T_e/(n_e e^2)]^{1/2}$. In contrast to this condition, there is increasing interest in "strongly coupled" plasmas where $n_e\lambda^3{}_D \lesssim 1$, as these are relevant to the early stages of inertial confinement when the plasma is very dense and relatively cold. **Figure 10** illustrates the regimes in $n_e - T_e$ parameter space, and shows three quantum-mechanical effects: (1) The classical minimum impact parameter b_0 (the Landau parameter), given by Eq. (31), is less than the de Broglie wavelength, given by Eq. (32), where $\hbar$ is Planck's constant h divided

$$b_0 = \frac{Ze^2}{4\pi\epsilon_0 k_B T_e} \tag{31}$$

$$\lambda\!\!\!^{-} = \frac{\hbar}{(m_e k_B T_e)^{1/2}} \tag{32}$$

by 2π, and m_e is the electron mass, above a temperature of 27 eV (1 eV $= 1.16 \times 10^4$ K). [The effective minimum impact parameter is then the de Broglie wavelength, though strictly the reason for the departure from classical Rutherford scattering is that the presence of Debye shielding modifies the r^{-1} dependence of the Coulomb potential to $r^{-1}\exp(-r/\lambda_D)$, which quantum-mechanically leads to the maximum impact parameter being reduced from λ_D to $\lambda_D\, b_0/\lambda\!\!\!^{-}$. Furthermore, electrons colliding with ions with an impact parameter of $\lambda\!\!\!^{-}$ have a maximum cross section for radiating their energy due to their large acceleration during the collision. This radiation loss is hence called bremsstrahlung, literally, braking radiation. The radiation is characterized by photon energy $h\nu$ of order k_BT.] (2) Below 1 eV a typical gas is only weakly ionized. (3) For dense plasmas when $n_e^{-1/3}$ is less than $\lambda\!\!\!^{-}$, the electron gas becomes degenerate. Above 0.5 MeV the electron energy is relativistic ($k_BT_e > m_ec^2$, where m_e is the electron mass and c is

the speed of light). *See* BREMSSTRAHLUNG; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; RELATIVISTIC QUANTUM THEORY.

The classical collision cross section for a single 90° deflection of an electron by an ion is πb_0^2. However, since there are many particles in a Debye sphere, each particle simultaneously interacts with $n\lambda^3{}_D$ particles. This leads to an increase in collision cross section by a factor 8 ln Λ, where Λ is λ_D/b_0 (or λ_D/λ for $T_e > 10^5$ K). Apart from the weak logarithmic dependence, the mean free path is then found to be proportional to T^2/n and the mean collision frequency $\bar{v}$ to $nT^{-3/2}$; for an electron of velocity υ, the mean free path is proportional to υ^4/n, and its collision frequency (with stationary ions) ν is proportional to $n\upsilon^{-3}$. This velocity dependence of the collision frequency is responsible for the penetration and preheating by more energetic electrons in laser-induced compression experiments, for so-called runaway electrons in ohmically heated toroidal plasmas, and in linear transport theory for the existence of large thermoelectric and Nernst terms in Ohm's law, discussed below. When the plasma is only partially ionized (for example, in a magnetohydrodynamic power generator), the effective collision frequency is the sum of the electron-ion and electron-neutral collision frequencies. *See* KINETIC THEORY OF MATTER.

For a fully ionized plasma the electrical conductivity σ (equal to η^{-1}) is $ne^2/(m_e\upsilon)$ and therefore is independent of electron density and proportional to $T^{3/2}$. Similarly the thermal conductivity and the viscosity are proportional to $T^{5/2}$ in the absence of a magnetic field. *See* CONDUCTION (ELECTRICITY); CONDUCTION (HEAT); VISCOSITY.

Fluid equations for a plasma. The full set of fluid equations for a plasma in a magnetic field displays a large number of terms which give many interesting effects in magnetohydrodynamics. This set can be derived from the Boltzmann kinetic equation. *See* BOLTZMANN TRANSPORT EQUATION.

Conservation equations. The first of these equations, Eq. (33), describes the conservation of mass, where

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{33}$$

ρ is the mass density and $\mathbf{v}$ is the fluid center-of-mass velocity. That this does indeed give conservation, which means no creation or destruction of mass, can be shown by integrating Eq. (33) over a volume V and employing Gauss' theorem, Eq. (34), so that the mathematical result of Eq. (35) means that the rate

$$\int \nabla \cdot (\rho \mathbf{v})\, dV = \oint \rho \mathbf{v} \cdot d\mathbf{S} \tag{34}$$

$$\frac{d}{dt} \int_V \rho\, dV = -\oint \rho \mathbf{v} \cdot d\mathbf{S} \tag{35}$$

of change of mass in the volume V is equal to the net inward flux, $-\rho\mathbf{v}$, integrated over the closed surface $d\mathbf{S}$ enclosing the volume V. *See* EQUATION OF CONTINUITY; GAUSS' THEOREM.

The conservation of total momentum is of similar form, given by Eq. (36). Here $\rho\mathbf{v}$ is the momentum of

$$\frac{d}{dt}(\rho \mathbf{v} + \mathbf{D} \times \mathbf{B}) + \nabla \cdot (\rho \mathbf{v}\mathbf{v} + \mathsf{P} + \mathsf{T}) = 0 \tag{36}$$

the fluid per unit volume, and to this must be added the momentum of the electromagnetic fields themselves per unit volume, $\mathbf{D} \times \mathbf{B}$. Since flux is a vector and momentum is also a vector, the flux of momentum is in general a tensor. For example, there can be a momentum component $\rho\upsilon_x$ transported with a velocity component υ_y. Thus the terms in the divergence operator are tensors: $\rho\mathbf{v}\mathbf{v}$ is the flux of momentum associated with the center-of-mass velocity; P, called the stress tensor, is the flux of momentum associated with the random or thermal motion of the fluid particles; and T is the flux of electromagnetic momentum usually called the Maxwell stress tensor. In most cases, Eq. (36) can be reduced to a more simple form that does not involve tensors, as given in Eq. (40). *See* TENSOR ANALYSIS.

The conservation of total energy also obeys a similar equation, Eq. (37), where $^1/_2\rho\upsilon^2$ and $^1/_2\rho\upsilon^2\mathbf{v}$ are

$$\frac{\partial}{\partial t}(^1/_2\rho v^2 + U + {}^1/_2\mathbf{E} \cdot \mathbf{D} + {}^1/_2\mathbf{H} \cdot \mathbf{B}) + \nabla \cdot (^1/_2\rho v^2\mathbf{v} + U\mathbf{v} + \mathsf{P} \cdot \mathbf{v} + \mathbf{q} + \mathbf{E} \times \mathbf{H}) = 0 \tag{37}$$

the kinetic energy per unit volume and flux of kinetic energy, respectively; $U = {}^3/_2 nk_BT$ is the internal energy density, n being the particle number density, k_B Boltzmann's constant, and T the temperature; and $\mathbf{q}$ is the heat flux. The terms $^1/_2\mathbf{E} \cdot \mathbf{D}$ and $^1/_2\mathbf{H} \cdot \mathbf{B}$ are the energy per unit volume in electric and magnetic fields, respectively, and $\mathbf{E} \times \mathbf{H}$ is the electromagnetic energy flux, commonly called Poynting's vector, and equal to $c^2\mathbf{D} \times \mathbf{B}$ for vacuum fields. *See* POYNTING'S VECTOR.

Application of Maxwell's equations. Maxwell's equations are Eqs. (38), (2), (39), and (30), where q_υ is the

$$\nabla \cdot \mathbf{D} = q_\upsilon \tag{38}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \tag{39}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{30}$$

charge density. From these equations it can be shown that the electromagnetic force per unit volume is given by Eq. (40), where T is given by Eq. (41) and

$$q_\upsilon\mathbf{E} + \mathbf{J} \times \mathbf{B} = -\frac{\partial}{\partial t}(\mathbf{D} \times \mathbf{B}) - \nabla \cdot \mathsf{T} \tag{40}$$

$$\mathsf{T} = -\mathbf{E}\,\mathbf{D} + {}^1/_2\mathbf{E} \cdot \mathbf{D}\mathbf{1} - \mathbf{H}\,\mathbf{B} + {}^1/_2\mathbf{H} \cdot \mathbf{B}\,\mathbf{1} \tag{41}$$

$\mathbf{1}$ is the unit tensor. The left-hand side of Eq. (40) can also be considered as the rate at which electromagnetic momentum per unit volume is given to the fluid or plasma, or minus the rate at which fluid momentum density is given to the electromagnetic fields. Through the effects of magnetic field tension

and pressure, the electromagnetic stress **T** provides the means, for example, of transferring the equal and opposite momentum onto a field coil when a plasma is being accelerated.

Using Eq. (40) and subtracting **v** times Eq. (33), Eq. (36) can now be written simply as Eq. (42),

$$\rho\frac{\partial \mathbf{v}}{\partial t} + \rho(\mathbf{v}\cdot\nabla)\mathbf{v} = \rho\frac{d\mathbf{v}}{dt} = -\nabla p + \mathbf{J}\times\mathbf{B} \qquad (42)$$

where the stress tensor has been approximated by the isotropic scalar pressure $p = nk_BT$. It has also been assumed here that the plasma is quasi-neutral, that is, that there are many particles in a Debye sphere so that the residual charge density q_v is small, causing $q_v\mathbf{E}$ to be a negligible force per unit volume. For this reason Eq. (38) is employed only for describing high-frequency phenomena. The operator d/dt refers to the absolute rate of change with time as seen by a moving fluid element.

Maxwell's equations can also be used to derive Eq. (43) for $\mathbf{J}\cdot\mathbf{E}$, the rate of transfer of energy per

$$\mathbf{J}\cdot\mathbf{E} = -\frac{\partial}{\partial t}({}^1/_2\mathbf{E}\cdot\mathbf{D} + {}^1/_2\mathbf{H}\cdot\mathbf{B}) - \nabla\cdot(\mathbf{E}\times\mathbf{H}) \qquad (43)$$

unit volume from electromagnetic to plasma energy. For a dynamo or a generator (for example, a magnetohydrodynamic generator) $\mathbf{J}\cdot\mathbf{E}$ must be negative, while for dissipation in or Joule heating of a stationary plasma, $\mathbf{J}\cdot\mathbf{E}$ will be positive.

Viscous effects. When the stress tensor is retained in the equation of motion, kinetic theory shows that in addition to the isotropic pressure there are contributions proportional to gradients or shear in velocity **v**. For an isotropic plasma (that is, one without large magnetic field effects) the equation of motion (42) is modified to give Eq. (44), where μ is the ion

$$\rho\frac{\partial \mathbf{v}}{\partial t} + \rho(\mathbf{v}\cdot\nabla)\mathbf{v} = -\nabla p + \mathbf{J}\times\mathbf{B} + \mu\nabla^2\mathbf{v} + {}^1/_3\mu\nabla(\nabla\cdot\mathbf{v}) \qquad (44)$$

viscosity, equal to the ion pressure divided by the collision frequency for ion-ion collisions. The viscous terms in Eq. (44) are the characteristic Navier-Stokes terms, the second one vanishing for an incompressible fluid. *See* FLUID-FLOW PRINCIPLES; NAVIER-STOKES EQUATION; VISCOSITY.

Other dimensionless numbers. The ratio of $\rho(\mathbf{v}\cdot\nabla)\mathbf{v}$ to $\mu\nabla^2\mathbf{v}$ in Eq. (44) gives Eq. (45) for the (vis-

$$R = \frac{\rho v L}{\mu} \qquad (45)$$

cous) Reynolds number, where L is the characteristic length over which the velocity shear exists. R describes the ratio of convection of vorticity, $\nabla\times\mathbf{v}$, to diffusion. A large value of R of the order of 10^3 leads to hydrodynamic turbulence. *See* REYNOLDS NUMBER.

The Peclet or thermal Reynolds number R_T is the ratio of the enthalpy flow $(U + p)\mathbf{v}$ to the heat flow **q**. If **q** is given by Fourier's law, Eq. (46), where κ is the thermal conductivity, the thermal Reynolds number R_T describes in Eq. (47) the ratio of convection of temperature to diffusion.

$$\mathbf{q} = -\kappa\nabla T \qquad (46)$$

$$R_T = \frac{5}{2}\frac{nk_BvL}{\kappa} \qquad (47)$$

The combined effect of convection and diffusion can lead to boundary layers being established. The relative thickness of the thermal viscous and magnetic (or skin current) boundary layers is described by Prandtl numbers which are the ratio of the various Reynolds numbers or diffusivities. The original Prandtl number is R_T/R, while the magnetic Prandtl number is R_M/R, where R_M is the magnetic Reynolds number defined in Eq. (28). *See* BOUNDARY-LAYER FLOW; SKIN EFFECT (ELECTRICITY).

The relative strength of $\mathbf{J}\times\mathbf{B}$ and $\mu\nabla^2\mathbf{v}$ in Eq. (44) leads to the definition of the Hartmann number H by Eq. (48), where **J** is taken to be of order $\mathbf{v}\times$

$$H^2 = \frac{B^2L^2}{\mu\eta} \qquad (48)$$

$\mathbf{B}/\eta$. Setting H approximately equal to 1 defines the scale length over which magnetic and viscous forces are comparable in magnitude.

Diamagnetism. The mean center-of-mass velocity **v** for a species should not be confused with the mean guiding-center velocity $\mathbf{v}_{gc}$ associated with the motion of individual particles of the species. Ignoring inertial effects and collisions, **v** in Eq. (34) can be shown to be the sum of $\mathbf{v}_{gc}$ and the diamagnetic velocity $1/Zen\ \mathrm{curl}_\perp\ (p_\perp\mathbf{B}/B^2)$, where $\mathrm{curl}_\perp$ is the component of the curl perpendicular to the magnetic field, and $p_\perp$ is the component of the anisotropic pressure perpendicular to the magnetic field. The physical meaning of the diamagnetic velocity can be understood from **Fig. 11**, which shows particle orbits in the plane perpendicular to the magnetic field B_z in a frame of reference in which the mean guiding-center motion is zero. It can be seen that if the density or temperature increases in the x direction, or if B_z decreases in the x direction, there will be a net contribution to momentum in the y direction in the shaded element due to the time-averaged statistical contribution from particles whose orbits intersect the element.

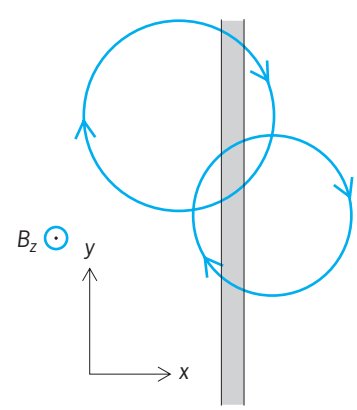

Fig. 11. Particles orbits in the plane perpendicular to a magnetic field B_z in a frame of reference in which the mean guiding center motion is zero.

Transport equations for a two-fluid plasma in a magnetic field. By solving the Boltzmann equation for the case when the electric field **E** and gradients of pressure, temperature, and velocity are small, transport equations can be derived relating the current density **J** and the heat flux **q** to these quantities. Thermodynamically the fluxes **J** and **q** are driven by the thermodynamic forces, **E** and ∇T, so that the irreversible generation of entropy is always positive locally. *See* ENTROPY.

In general, cross-phenomena occur; for example, current can flow as a result of a temperature gradient. The physical reason for this is that the collision frequency is velocity-dependent. Thus, when a temperature gradient exists, higher-velocity particles in the hotter region will flow with fewer collisions, so that there will be not only a heat flow but also a current. If a current is prevented from flowing, as in the symmetric radially inward heat flow in laser fusion targets, an electric field will develop instead to retard the hotter electrons and also to drive an equal and opposite cold electron current. The presence of the electric field will also affect the heat flow. *See* THERMOELECTRICITY.

If there is also a magnetic field which is orthogonal to an electric field or to a temperature gradient, the electrons will cause a current or heat flow not only parallel to **E** or ∇T but also orthogonal to both **E** or ∇T and the magnetic field. This is because of the cycloidal motion of charged particles in a magnetic field. *See* GALVANOMAGNETIC EFFECTS; THERMOMAGNETIC EFFECTS.

When the center-of-mass velocities and temperature of the electrons and ions are considered to be different, separate equations for the two fluids have to be employed. Using the notation of Braginskii, the transport equations for the electron fluid (subscript e) can be written as Eqs. (49) and (50), where $\mathbf{E}^*$ is a generalized electric field given by Eq. (51); **b** is

$$n_e e\mathbf{E}^* = \frac{\alpha_\parallel \mathbf{J}_\parallel + \alpha_\perp \mathbf{J}_\perp - \alpha_\wedge \mathbf{b}\times\mathbf{J}}{n_e e} - n_e\left(\beta_\parallel \nabla_\parallel T_e + \beta_\perp \nabla_\perp T_e + \beta_\wedge \mathbf{b}\times\nabla T_e\right) \quad (49)$$

$$\mathbf{q}_e = -\left(\kappa_\parallel \nabla_\parallel T_e + \kappa_\perp \nabla_\perp T_e + \kappa_\wedge \mathbf{b}\times\nabla T_e\right) - \frac{\left(\beta_\parallel \mathbf{J}_\parallel + \beta_\perp \mathbf{J}_\perp + \beta_\wedge \mathbf{b}\times\mathbf{J}\right) T_e}{e} \quad (50)$$

$$\mathbf{E}^* = \mathbf{E} + \mathbf{v}\times\mathbf{B} - \frac{\mathbf{J}\times\mathbf{B}}{n_e e} + \frac{\nabla p_e}{n_e e} \quad (51)$$

the unit vector in the direction of the magnetic field; the coefficients $\alpha_\perp$, $\alpha_\wedge$, and so forth are functions of $(\omega_e\tau_e)$, the Hall parameter; $\alpha_{||}$ is the value of $\alpha_\perp$ when $\omega_e\tau_e = 0$, and so forth; and $\mathbf{J}_{||}$, $\mathbf{J}_\perp$, $\nabla_{||}T_e$, and $\nabla_\perp T_e$ are given by Eqs. (52)–(55).

$$\begin{aligned} \mathbf{J}_\parallel &= \mathbf{b}(\mathbf{J}\cdot\mathbf{b}) \\ \mathbf{J}_\perp &= \mathbf{b}\times(\mathbf{J}\times\mathbf{b}) \\ \nabla_\parallel T_e &= \mathbf{b}(\nabla T_e\cdot\mathbf{b}) \\ \nabla_\perp T_e &= \mathbf{b}\times(\nabla T_e\times\mathbf{b}) \end{aligned} \quad (52)$$

In Eq. (49) the Hall electric field is principally included in the $\mathbf{J}\times\mathbf{B}$ term in the definition of $\mathbf{E}^*$ [Eq. (51)], while the correction $\alpha_\wedge$ has been shown to vary as $(\omega_e\tau_e)^{-2/3}$ for values of $(\omega_e\tau_e)$ much greater than 1. The main property of the Hall effect comes from a comparison of the $\alpha_\perp$ term and the $\mathbf{J}\times\mathbf{B}$ term in $\mathbf{E}^*$: This is that the components of current density **J** and the general electric field $\mathbf{E}+\mathbf{v}\times\mathbf{B}+\nabla p_e/(n_e e)$ in the plane perpendicular to the magnetic field are at an angle $\theta \simeq \tan^{-1}(\omega_e\tau_e)$ to each other. Thus in a magnetically confined plasma there is an electric field component orthogonal to the magnetic surface. In magnetohydrodynamic power generators the electrodes have to be segmented in order that the current flow is orthogonal to the gas flow, so that the $\mathbf{J}\times\mathbf{B}$ force is directed to oppose this flow, and the generator impedance is minimized. The segmentation of the electrodes permits an electric field to be set up between the segments.

When an electric field **E** is applied to a plasma in a direction orthogonal to a magnetic field so that the current can flow freely at the angle $\tan^{-1}(\omega_e\,\tau_e)$ to **E**, the component of current density orthogonal to **E** is called the Hall current. This can be used to give a $\mathbf{J}\times\mathbf{B}$ acceleration of the ions in a suitable magnetic field configuration such as in magnetically insulated diodes (as used for light ion-beam production). The Hall term modifies the ideal magnetohydrodynamic frozen-flow motion so that magnetic fields are now frozen to the electron center-of-mass velocity.

The last term in Eq. (49) describes the thermoelectric term, and through this it is possible to generate a current parallel to ∇T_e (Seebeck effect) or perpendicular to ∇T_e (Nernst effect). The Nernst effect can approximately be represented as an electic field equal to $\mathbf{q}_e\times\mathbf{B}/(^3/_2 n_e k_B T_e)$ showing by analogy with Eqs. (1) and (3) that the magnetic field is convected in the direction of heat flow with a velocity given by Eq. (56). In the case of laser fusion, this

$$\mathbf{v}_T = \frac{\mathbf{q}_e}{^3/_2 n k_B T_e} \quad (56)$$

can lead to a magnetic field being convected from the critical surface to the ablation surface, and, furthermore, amplified because of the increase in number density n which reduces the velocity $\mathbf{v}_T$ causing the transverse magnetic field lines to bunch up. This further modification of frozen flow implies that the magnetic field is frozen to the hotter (and less collisional) electrons that carry the heat. Indeed, it is a special feature of inertial confinement fusion that megagauss magnetic fields can spontaneously be generated in nonuniformly irradiated targets through the $\nabla p_e/(n_e e)$ term in $\mathbf{E}^*$ [Eq. (50)] from Eq. (2), so that Eq. (57) holds. (In laser fusion the ideal situ-

$$\frac{\partial\mathbf{B}}{\partial t} = -\nabla\times\mathbf{E} = k_B\nabla T_e\times\frac{\nabla n_e}{n_e e} \quad (57)$$

ation is one in which laser radiation is uniformly irradiated onto and absorbed by a perfectly symmetrical spherical target. The above magnetic fields are generated only through a lack of symmetry arising from nonuniform irradiation or through instabilities such as the Rayleigh-Taylor instability discussed above.)

In Eq. (50), it is found that while $\kappa_{||}$ is proportional to $T_e^{5/2}$, $\kappa_\perp$ varies as $\kappa_{||}(\omega_e\tau_e)^{-2}$ for values of $\omega_e\tau_e$ much greater than 1. Therefore, in magnetically confined plasmas the electron thermal conduction across magnetic surfaces is considerably reduced, and indeed ion thermal conduction across field lines is $[m_i/(Zm_e)]^{1/2}$ times larger. The heat flow parallel to a current flow essentially represents the flux of enthalpy of the current carriers and is called the Peltier effect. The heat flux defined by Eq. (50), however, is defined in the center-of-mass frame of the electrons, and so the $\beta_{||}$ and $\beta_\perp$ terms refer to the corrections to the enthalpy flow caused by the velocity dependence of the collision frequency. The Ettingshausen effect is the heat flow in the direction of the Hall electric field, while the Righi-Leduc effect is the heat flow orthogonal to ∇T_e and $\mathbf{B}$. $\kappa_\wedge$ and $\beta_\wedge$ are proportional to $\omega_e\tau_e$ for $\omega_e\tau_e$ much less than 1, and to $(\omega_e\tau_e)^{-1}$ for $\omega_e\tau_e$ much greater than 1, while $\beta_\perp$ varies as $(\omega_e\tau_e)^{-5/3}$ for $\omega_e\tau_e$ much greater than 1. It follows that $\alpha_\wedge$, $\beta_\wedge$, and $\kappa_\wedge$ all vanish as the magnetic field tends to 0. The peculiar fractional powers in $\alpha_\wedge$ and $\beta_\perp$ occur because for a value of $\omega_e\tau_e$ much larger than 1 it is found that a fraction of electrons with a velocity less than $(\omega_e\tau_e)^{-1/3}$ times the mean thermal speed $\sqrt{(2k_BT_e/m_e)}$ have in fact unmagnetized orbits. This is because they have a low velocity υ so that their collision frequency which is proportional to υ^{-3} exceeds their cyclotron frequency.

The ion stress tensor is also modified by a strong magnetic field, and for $\omega_i\tau_i \gg 1$ leads to the persistence of terms in velocity gradients with a coefficient p_i/ω_i, where p_i is the ion pressure and ω_i the ion cyclotron frequency. These residual terms describe finite ion Larmor radius effects, which can be a beneficial correction to ideal magnetohydrodynamic stability theory. Malcolm G. Haines

Bibliography. G. Bateman, *MHD Instabilities*, 1980; D. Biskamp, *Nonlinear Magnetohydrodynamics*, 1993; F. F. Chen, *Introduction to Plasma Physics and Controlled Fusion*, vol 1: *Plasma Physics*, 2d ed., 1984; E. M. Epperlein and M. G. Haines, *Physics of Fluids*, 29:1029, 1986; N. A. Krall and W. Trivelpiece, *Principles of Plasma Physics*, 1973, reprint 1986; K. Miyamoto, *Plasma Physics for Nuclear Fusion*, 1976, rev. ed.1989; J. A. Wesson, *Tokamaks* 1987, 2d ed., 1997.

Magnetometer

An instrument that measures the magnitude and/or direction of a magnetic field. A magnetometer can be either a scalar instrument that measures the magnitude of the total field or a vector instrument that measures one or more vector components of the field. Some magnetometers are relative devices that are calibrated with respect to a known field. Others are absolute devices that yield magnetic-field values without calibration. Arrays of magnetometers can be configured as gradiometers that suppress or enhance far-field magnetic sources while enhancing or suppressing near-field sources. Magnetometer technology is available for detecting a range of magnetic fields (**Fig. 1**).

Induction. Perhaps the simplest magnetometer is an induction coil that employs the Faraday effect, a voltage induced in a conductor by a time-varying magnetic flux. Induction magnetometers are generally multiturn coils that are either air-wound or wound around a ferromagnetic core to increase flux through the coil. Fixed coils can measure only time-varying fields, but static fields can be measured using moving induction coils that rotate, vibrate, or deform. Their sensitivity is governed by the time rate of change of the magnetic flux change in the coil, the number of turns in the coil, the magnetic permeability of the core, the electrical resistance and parasitic capacitance of the coil, and thermal noise. Induction coils are usually vector sensors that require external calibration. *See* FARADAY EFFECT.

Hall effect. In the presence of a magnetic field, moving charge carriers in a current-carrying solid conductor or semiconductor are subjected to a Lorentz force that depends upon their charge, velocity, and magnitude and direction of the magnetic field. This force deflects the charge carriers and produces a voltage difference on opposite sides of the solid known as the Hall effect. *See* HALL EFFECT; MAGNETIC FIELD.

Modern Hall-effect sensors are made from InSb or InAs alloy thin films produced by vacuum deposition or molecular-beam epitaxy. These solid-state vector devices can be fabricated as integrated circuit packages that can measure alternating-current and direct-current magnetic fields. By combining multiple calibrated sensors into a single package, the total magnetic field and magnetic-field gradients can be measured. Hall-effect magnetometers are usually used to measure relatively strong magnetic fields.

Magnetoresistance and SDT. Magnetoresistance is the dependence of the electrical resistance of a solid upon an applied magnetic field. A wide variety of magnetoresistive sensors have been designed and constructed for medium field strengths. They can be manufactured in integrated circuit packages in small sizes at relatively low cost and used for magnetic detection and navigation. *See* MAGNETORESISTANCE.

Thin-film alloys of iron, nickel, and/or cobalt are used to measure magnetic fields using anisotropic magnetoresistance (AMR). The more recently discovered giant magnetoresistive (GMR) effect sensors are made of sandwiched layers of ferromagnetic iron-nickel-cobalt alloy and a nonferromagnetic material. A recent variant of this type of sensor is the spin-dependent tunneling (SDT) device (**Fig. 2**), where the nonferromagnetic layer is replaced by an insulator (commonly aluminum oxide) yielding up to 20 times the sensitivity of a GMR device and perhaps 200 times the sensitivity of an AMR sensor. *See* MAGNETIC THIN FILMS.

Fluxgate. The fluxgate (saturable-core) sensor is constructed from a pair of high-magnetic-permeability cores. Wound around each core in opposite directions is a primary field core driven to

magnetic saturation by an audio-frequency current. The primary coils are connected in series. A secondary field core wound collectively around both cores is connected to a detector circuit.

When a core reaches magnetic saturation, the driving current waveform is distorted by a change in self-inductance in the primary coil. In the absence of an external magnetic field, both cores saturate at the same instant and no voltage is developed in the secondary coil. In the presence of a biasing external magnetic field, one core comes to saturation before the other, producing a pair of voltage pulses in the secondary coil at twice the driving frequency. For small fields, the amplitude of this pulse is proportional to the strength of the ambient field, and thus the device is a sensitive magnetometer capable of measurement precision better than 0.1 nanotesla over a frequency range from direct current to low audio frequencies.

All fluxgate magnetometers are relative vector instruments that require calibration in a known magnetic field to produce accurate results. Orthogonal sets of fluxgate sensors can be used to measure all three field components and thereby the total field vector.

Resonance. Two general classes of resonance magnetometers are proton precession magnetometers and optically pumped magnetometers. Both are absolute instruments that measure total field strength without the need for calibration using a known magnetic field.

Proton precession magnetometer. The proton precession magnetometer employs nuclear magnetic resonance to measure magnetic-field strength. The sensor contains a reservoir of a proton-rich fluid (water or a low-density hydrocarbon like kerosene or decane) surrounded by a magnetic-field coil. The coil initially produces a direct-current magnetic field that aligns the magnetic moments of the protons. The direct-current field is abruptly turned off, whereupon the proton magnetic moments precess about the ambient magnetic-field direction for a few seconds, and produce an alternating-current signal in the coil whose frequency is directly proportional to the intensity of the ambient magnetic field. The Overhauser magnetometer is a variant of this design that aligns the protons by radio-frequency electron spin resonance coupling. This technique requires far less power and is well suited for spacecraft.

The precession signal is essentially independent of the orientation of the sensor, and accuracy is limited by the magnitude of the precession signal and the accuracy of the frequency counter connected to the coil. Typical precision is on the order of 0.01 nT. Using a larger Helmholtz coil to null one component of the ambient field allows the proton precession magnetometer to be converted to a vector magnetometer that can measure components of a magnetic field. The measurement bandwidth of the proton magnetometer is limited to low frequencies because the polarization/frequency counting sequence can take several seconds. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

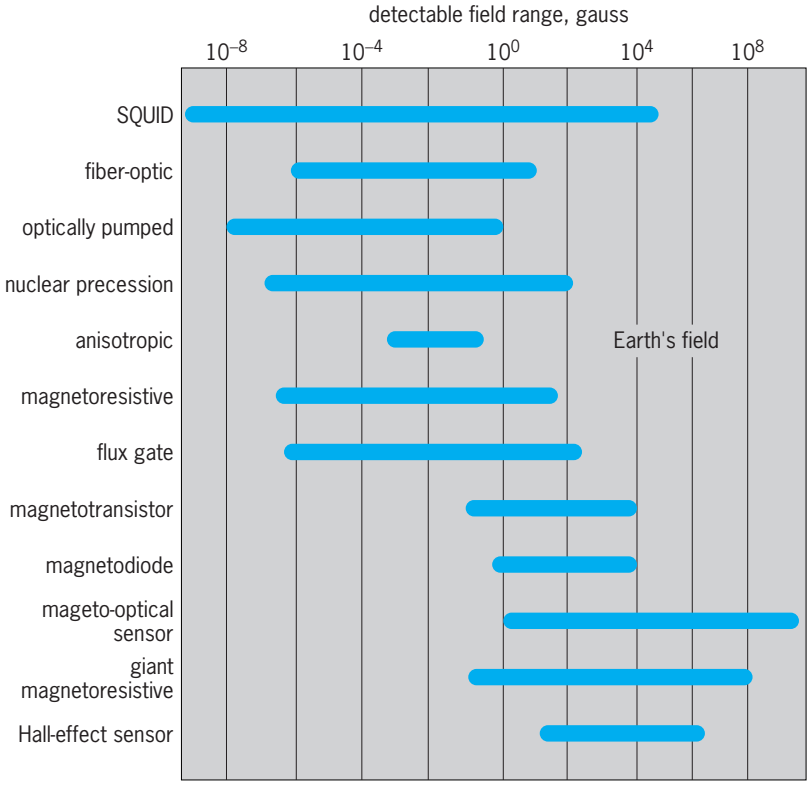

Fig. 1. Magnetic sensor technologies and ranges. 1 gauss = 10^{-4} tesla = 10^{5} gamma.

Optically pumped magnetometer. This device is analogous to the proton magnetometer in that it employs optical pumping to cause atomic or electron spin precession. A sealed glass cell containing an alkali vapor (typically cesium, rubidium, or potassium) is placed next to a lamp containing the same alkali vapor, and produces light at a characteristic frequency (**Fig. 3**). Opposite the lamp is a photodetector whose output is connected through an amplifier to a coil surrounding the cell. Resonance absorption and reradiation of optical energy from the cell produce oscillation in a feedback loop whose frequency is proportional to the intensity of the ambient magnetic field, as in the case of the proton magnetometer. A variant is the metastable ^{4}He magnetometer, where the alkali vapor cell is replaced by

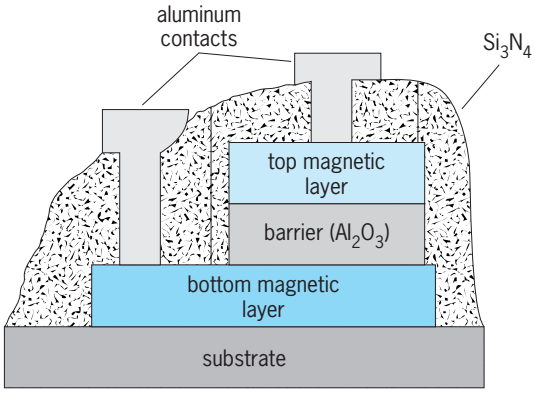

Fig. 2. Schematic cross section of the STD structure.

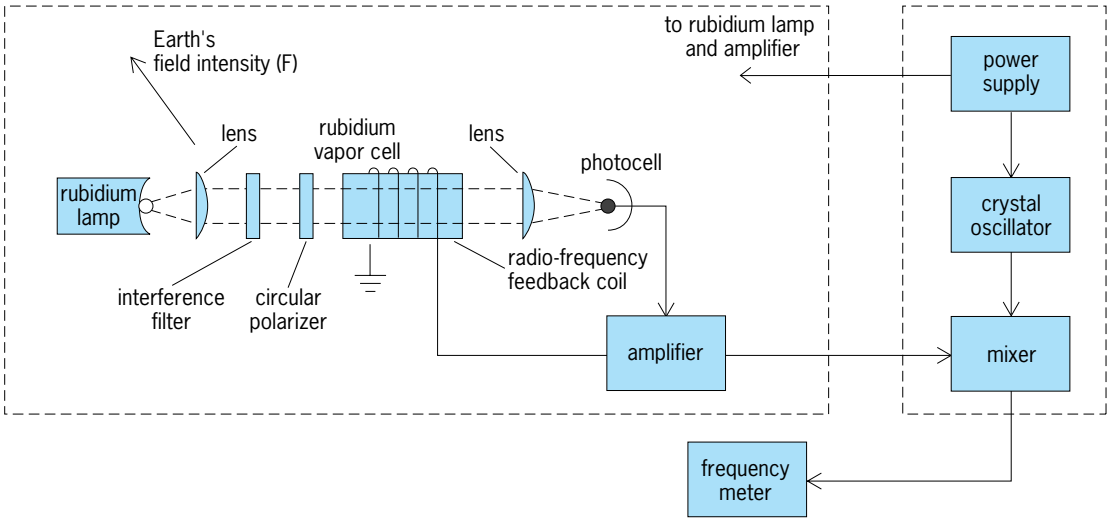

Fig. 3. Rubidium vapor magnetometer. The rubidium vapor cell, photocell, amplifier, and feedback coil constitute an oscillator whose frequency is proportional to the ambient magnetic field.

a helium gas cell. In the most modern designs, the lamps are replaced by lasers and the cells contain an alkali metal vapor and helium, providing vastly increased sensitivity and spatial resolution.

The sensitivity of most optically pumped magnetometers is superior to that of proton precession magnetometers. A typical cesium vapor magnetometer has a measurement precision better than 0.001 nT, but the absolute accuracy is on the order of 1 nT because of field-dependent complexities in the energy transitions in the cell. In contrast to the proton magnetometer, the measurement bandwidth of optically pumped magnetometers is quite broad (from direct current up to several megahertz) because the pumped energy transitions occur over very short time intervals. *See* OPTICAL PUMPING.

Superconducting. The cryogenic or SQUID magnetometer uses one or more Josephson junctions as a magnetic-field sensor. A Josephson junction is a zone of weak magnetic coupling (a weak link) between two regions of superconducting material in which current will flow without resistance. A change in the magnetic field applied to the weak link produces a proportional change in magnetic flux within the Josephson junction (**Fig. 4**). The SQUID is the most sensitive magnetometer in general use, capable of measuring flux changes only a small fraction of a flux quantum Φ (2.07×10^{-15} weber), but its capabilities are now rivaled by modern spin-polarized alkali vapor magnetometers. *See* JOSEPHSON EFFECT; SUPERCONDUCTIVITY.

Two types of SQUID are in common use. The radio-frequency SQUID employs a single weak link, whereas the direct-current SQUID uses a pair of Josephson junctions. Flux changes can be introduced directly to the SQUID, but more commonly a flux transformer is used, where external sensing coils are inductively linked to the SQUID. The external coils are made of superconducting material and can be configured as a single-turn or multiturn coil, a Helmholtz pair, or even a first- or second-order gradiometer. The practical noise limits of conventional niobium SQUID devices are on the order of 10^{-6} nT, orders of magnitude better than other devices. *See* HELMHOLTZ COILS.

All SQUID magnetometers are relative, vector instruments. The principal advantages of the SQUID magnetometer over proton, optically pumped, and fluxgate magnetometers are sensitivity and frequency response. The principal disadvantage of the SQUID magnetometer is that it must be kept in a

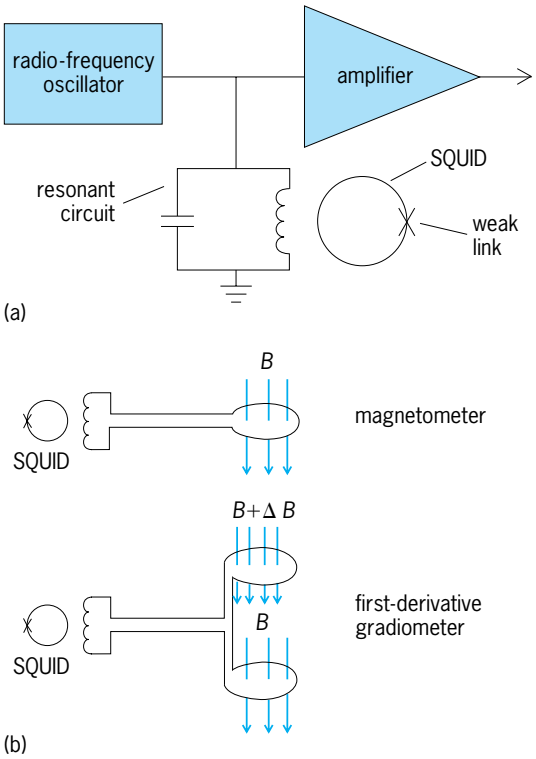

Fig. 4. SQUID magnetometer. (*a*) Conventional radio-frequency SQUID electronics. (*b*) Sensing-coil configurations.

superconducting state. At one time, this meant that the sensor had to be maintained at 4.2 K (−452°F) by using liquid helium as a cryogen. Now inexpensive direct-current SQUID devices can be made from a thin-film substrate of YBCO ($Y_1Ba_2Cu_3/SO_7$), which use liquid nitrogen (77 K or −321°F) as a cryogen. The performance of the high-temperature SQUID device does not yet match that of the conventional liquid-helium-cooled niobium SQUID. *See* SQUID.

Fiber-optic. Some ferromagnetic materials exhibit magnetostriction, the dependence of the state of strain of a solid (and hence its dimensions) on magnetic-field strength. By bonding a magnetostrictive material to an optical fiber, its optical path length can be made to change in response to a magnetic field. Sensitive magnetometers have been constructed that employ a Mach-Zehnder interferometer to measure the optical path length difference between the sensor fiber and a reference fiber at mechanical resonance. *See* FIBER-OPTIC SENSOR; MAGNETOSTRICTION.

Other magnetometers. Classical mechanical magnetometers employ small magnets suspended on balances or torsion fibers. The first absolute device was probably constructed by K. F. Gauss in 1832 and consisted of a magnet suspended by a torsion fiber and free to rotate about a vertical axis. The absolute value of the horizontal magnetic-field component is estimated by measuring the period of oscillation of the magnet. The field inclination is measured with a dip needle, resulting in a complete specification of the vector field.

Other devices that employ suspended or balanced magnets are the dip needle, sine galvanometer, Schmidt vertical balance, quartz horizontal magnetometer, astatic magnetometer, and magnetometer zero balance. Compared to modern instruments, their main disadvantages include the time required to make a measurement, temperature dependence of the sensor, and reduced sensitivity. *See* GEOPHYSICS; MAGNET. Michael McWilliams

Bibliography. D. Grundler, Highly sensitive YBCO DC SQUID magnetometer with thin-film flux transformer, *Appl. Phys. Lett.*, 63:2700–2702, 1993; D. A. Konotop, Development of high T_c superconductor SQUID-based magnetometer, *Cryogenics*, 33:632–636, 1993; P. Ripka, *Magnetic Sensors and Magnetometers*, Artech House, 2001; S. Tanaka, Four-channel YBCO DC SQUID magnetometer for biomagnetic measurements, *Appl. Phys. Lett.*, 64:514–516, 1994; W. M. Telford, L. P. Geldart, and R. E. Sheriff, *Applied Geophysics*, 2d ed., 1990.

Magneton

A unit of magnetic moment used to describe atomic, molecular, or nuclear magnets. More precisely, one unit, the Bohr magneton, is used at the atomic and molecular levels, and another unit, the nuclear magneton, is used at the nuclear level. Still another unit (which might be called the muon magneton, but is usually not named) is used to describe the magnetic moment of the muon.

Bohr magneton. The Bohr magneton μ_B is defined and its value given in Eq. (1), where $-e$ and m are the

$$\mu_B = \frac{e\hbar}{2m} = (9.274\,009\,49 \pm 0.000\,000\,80) \times 10^{-24}\ \text{joule/tesla} \quad (1)$$

charge and mass of the electron respectively and $\hbar$ is Planck's constant divided by 2π. In Dirac's theory the magnetic moment of the electron is exactly $-\mu_B$, but according to the theory of quantum electrodynamics the electron has a small anomalous magnetic moment. (The anomalous magnetic moment of a particle of spin $^1/_2$ is the deviation of its magnetic moment from the value given by the Dirac equation.) The experimental value of the electron magnetic moment μ_e is given by Eq. (2), in agreement with the pre-

$$\mu_e = -(1.001\,159\,652\,185\,9 \pm 0.000\,000\,000\,003\,8)\mu_B \quad (2)$$

diction of quantum electrodynamics within the errors. The minus sign in Eq. (2), which arises from the fact that the electron has negative charge, is usually omitted, and just the magnitude of the electron magnetic moment is given. According to quantum electrodynamics, the magnetic moment of the positron is equal in magnitude to the magnetic moment of the electron but positive in sign. More generally, the magnetic moments of any particle and its antiparticle are expected to be equal in magnitude and opposite in sign. *See* ELECTRON SPIN; QUANTUM ELECTRODYNAMICS; RELATIVISTIC QUANTUM THEORY.

The magnetic moment of a composite particle results from contributions from both orbital angular momentum and spin of its constituents. In the case of atoms and molecules, by far the major contribution comes from the electrons, the contribution from the atomic nuclei usually being neglected. If an atom is regarded as having an infinitely heavy nucleus, the contribution from the orbital angular momentum of an electron is an integral number of Bohr magnetons.

Muon magnetic moment. The unit of magnetic moment to describe the muon is obtained from the Bohr magneton by replacing m in Eq. (1) by the muon mass m_μ. The value of the muon magnetic moment is given in Eq. (3). The deviation of the muon magnetic

$$\mu_\mu = (1.001\,165\,919\,81 \pm 0.000\,000\,000\,62)\frac{e\hbar}{2m_\mu} \quad (3)$$

moment from its Dirac value can also be accounted for by the theory of quantum electrodynamics. *See* LEPTON.

Nuclear magneton. The nuclear magneton is obtained from the Bohr magneton by replacing m by the proton mass m_p. The value of the nuclear magneton is given in Eq. (4). The nuclear magneton is used

$$\mu_N = (5.050\,783\,43 \pm 0.000\,000\,43) \times 10^{-27}\ \text{joule/tesla} \quad (4)$$

not only as the unit for the magnetic moment of the proton but also for the neutron and other hadrons and for atomic nuclei. If the proton and neutron were Dirac particles, the proton's magnetic moment would be one nuclear magneton (except for a small correction arising from quantum electrodynamics) and the neutron's magnetic moment would be zero (because the neutron is uncharged). However, the proton and neutron have large anomalous magnetic moments. The magnetic moments of the proton and neutron are given in Eqs. (5).

$$\begin{aligned} \mu_p &= (2.792\,847\,351 \pm 0.000\,000\,028)\mu_N \\ \mu_n &= (-1.913\,042\,73 \pm 0.000\,000\,45)\mu_N \end{aligned} \tag{5}$$

See NEUTRON; NUCLEAR MOMENTS; PROTON.

Origin of anomalous moments. According to present theory, the proton, neutron, and other hadrons have large anomalous magnetic moments because these particles are not elementary but composite. In the theory of quantum chromodynamics, the principal constituents of a baryon, such as the proton or neutron, are three quarks. In a simple model that neglects the contributions from quark orbital angular momenta, from relativistic effects, and from higher-order effects of quantum chromodynamics, the magnetic moment of a baryon can be calculated in terms of the magnetic moments of its three constituent quarks. These calculations give the remarkably good prediction that $\mu_p/\mu_n = -1.5$, a value only 3% from the measured value. In the case of the higher-mass baryons (Λ, Σ, Ξ), the predicted magnetic moments are all of the correct sign, but the disagreement with experiment is up to 20%. Inclusion of contributions from quark orbital angular momentum and other higher-order effects improves the agreement with experiment, but a residual disagreement remains. *See* BARYON; ELEMENTARY PARTICLE; FUNDAMENTAL CONSTANTS; QUANTUM CHROMODYNAMICS; QUARKS.

Don B. Lichtenberg

Bibliography. S. Eidelman et al., Review of particle physics, *Phys. Lett. B*, 592:1–1109, 2004; K. J. Heller (ed.), *High-Energy Spin Physics 8th International Symposium*, Amer. Inst. Phys. Conf. Proc. 187, 1989; P. J. Mohr and B. N. Taylor, CODATA recommended values of the fundamental physical constants: 2002, *Rev. Mod. Phys.*, 77:1–107, 2005.

Magnetooptics

That branch of physics which deals with the influence of a magnetic field on optical phenomena. Considering the fact that light is electromagnetic radiation, an interaction between light and a magnetic field would seem quite plausible. It is, however, not the direct interaction of the magnetic field and light that produces the known magnetooptic effects, but the influence of the magnetic field upon matter which is in the process of emitting or absorbing light.

Zeeman effect. This produces a splitting of spectrum lines when the emitting light source is placed in a magnetic field. The inverse Zeeman effect refers to a similar splitting of absorption lines when the absorbing substance is in a magnetic field. The Zeeman effect for spectrum lines originating from closely spaced levels is called the Paschen-Back effect. The explanation of most other magnetooptical phenomena is based on the Zeeman effect, which therefore may be regarded as the basic magnetooptic effect. *See* PASCHEN-BACK EFFECT; ZEEMAN EFFECT.

Faraday effect. This is the rotation of the plane of polarization of light when light traverses certain substances in a magnetic field. *See* FARADAY EFFECT.

Voigt effect. An anisotropic substance placed in a magnetic field becomes birefringent (doubly refracting), and its optical properties are similar to those of a uniaxial crystal. The Faraday effect is the result of this birefringence when observations are made parallel to the magnetic field. The analogous observations perpendicular to the magnetic lines of force are more difficult and were not successfully carried out until 1898 because of the smallness of the effect. The transverse magnetooptic birefringence is called the Voigt effect after its discoverer, W. Voigt.

The Voigt effect (also called magnetic double refraction) can easily be calculated for substances having a normal Zeeman effect. For more complicated Zeeman effects the results can also be theoretically predicted but are less simple quantitatively, though not essentially different from those in the simpler cases.

The Voigt effect depends on the fact that the indices n_s and n_p for light polarized perpendicular or parallel to the magnetic lines of force respectively are different from one another in a magnetic field where the absorption line shows a Zeeman effect. The value of n_p is independent of the magnetic field because the central component does not change while $n_s = {}^1/_2(n^+ + n^-)$. The appropriately labeled curve in the **illustration** gives $n_s - n_p$, on which the observed effects of the transverse magnetic double refraction depend.

The theoretical formulas which represent the double refraction, though readily derived, are quite complicated. When the wavelength is considerably

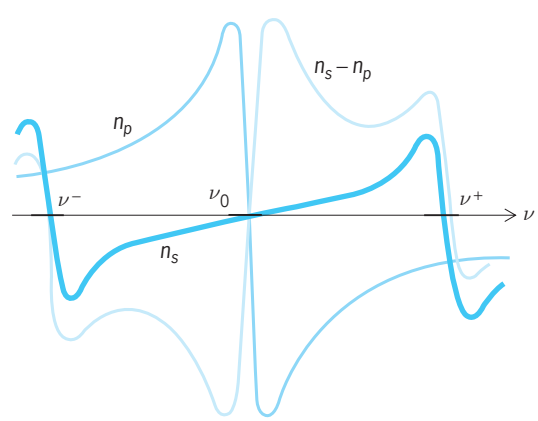

Index of refraction of light polarized parallel (n_p) and perpendicular (n_s) to the magnetic field in the vicinity of a Zeeman triplet. The Voigt effect is proportional to $n_s - n_p$ (shown as a curve).

removed from the Zeeman triplet, the phase difference is as given by Eq. (1), where ν_0 is absorption

$$\delta = \frac{2\pi x}{c}(n_p - n_s) = \frac{e^4 f x}{32\pi^2 c^3 n_0 (\nu - \nu_0)^3} N H^2 \quad (1)$$

line frequency, ν is frequency of transmitted light, e is charge of electron, x is path length, c is velocity of light, n_0 is index of refraction without field, N is number of absorbing atoms per unit volume, H is magnetic field strength, and f is so-called oscillator strength, measure of strength of absorption.

Contrary to the situation existing in the Faraday effect, first-order effects of magnetic double refraction are canceled out because of the presence of two perpendicular, symmetrically placed Zeeman components of combined strength equal to that of the parallel component. Because of this the Voigt effect can be observed only in the vicinity of sharply defined absorption lines, that is, in gases and in certain crystals having sharp lines, such as rare-earth salts. In the case of the rare-earth salts a linear Voigt effect is also possible at extremely low temperatures.

Since the formula for phase difference contains the oscillator strength f, the Voigt effect may be used to measure this important quantity.

Cotton-Mouton effect. This effect is concerned with the double refraction of light in a liquid when the liquid is placed in a transverse magnetic field. It is analogous to the electrooptical Kerr effect and is observed in liquids with complicated molecular structure. If the molecule has a magnetic moment, the field tries to orient the molecule, but the thermal motion tends to oppose this action. There is thus a degree of orientation, which is dependent on the temperature. If the molecule itself is optically anisotropic, the liquid also will be anisotropic and will exhibit double refraction. *See* KERR EFFECT.

The Cotton-Mouton effect is observed chiefly in nitrobenzene and aromatic organic liquids. Aliphatic compounds have a considerably smaller effect.

The phase difference of the Cotton-Mouton effect is expressed by Eq. (2), where x is the path length

$$\delta = C_m x H^2 \quad (2)$$

and C_m is called the Cotton-Mouton constant. For nitrobenzene at a temperature of 61.3°F (16.3°C) and a wavelength λ of 578 nanometers, $C_m = 2.53 \times 10^{-2}/(\mathrm{m} \cdot \mathrm{T}^2) = 2.53 \times 10^{-12}/(\mathrm{cm} \cdot \mathrm{gauss}^2)$. With large magnets ($H = 4.65$ teslas $= 46{,}500$ gauss) under the most favorable circumstances, rotations of the plane of polarization up to 27° have been observed.

The dispersion of the Cotton-Mouton effect is given by Havelock's law as in the Kerr effect.

Magnetooptic Kerr effect. This deals with the changes that are produced in the optical properties of a reflecting surface of a ferromagnetic substance when the substance is magnetized. In a typical case this will result in elliptically polarized light appearing in reflection, when the ordinary rules of metallic reflection would given only plane-polarized light. The component produced by the magnetic field when the magnetization is close to saturation is only of the order 10^{-3} of that normally present. The explanation must be sought in the fact that the conduction electrons made to vibrate by the incident light will have a curved path in the magnetic field.

Majorana effect. This deals with optical anisotropy of colloidal solutions. The effect probably is caused by the orientation of the particles in the magnetic field.

Magnetooptic effects have played an increasingly important role in microwave spectroscopy, where transitions between the Zeeman components of a single level can be observed directly. *See* MICROWAVE SPECTROSCOPY.

G. H. Dieke; William W. Watson

Magnetoresistance

The change of electrical resistance produced in a current-carrying conductor or semiconductor on application of a magnetic field H. Magnetoresistance is one of the galvanomagnetic effects. It is observed with H both parallel to and transverse to the current flow. The change of resistance usually is proportional to H^2 for small fields, but at high fields it can rise faster than H^2, increase linearly with H, or tend to a constant (that is, saturate), depending on the material. In most nonmagnetic solids the magnetoresistance is positive. *See* GALVANOMAGNETIC EFFECTS.

In semiconductors, the magnetoresistance is unusually large (especially in indium antimonide) and is highly anisotropic with respect to the angle between the field direction and the current flow in single crystals. When the magnetoresistance is measured as a function of field, it is the basis for the Shubnikov–de Haas effect, much as the field dependence of the magnetization gives rise to the de Haas–van Alphen effect. Measurement of either effect as the field direction changes with respect to the crystal axes serves as a powerful probe of the Fermi surface. Magnetoresistance measurements also yield information about current carrier mobilities. Important to practical applications is the fact that the geometry of a semiconductor sample can generate very large magnetoresistance, as in the Corbino disk. *See* CORBINO DISK; DE HAAS-VAN ALPHEN EFFECT; FERMI SURFACE; SEMICONDUCTOR.

In some effectively two-dimensional systems, such as GaAs-AlGaAs field-effect transistors, the high-field magnetoresistance goes to zero in fields at which the Hall resistance exhibits plateaus. The Hall resistance is then independent of H and its value is $h/(e^2 i)$, where h is Planck's constant, e is the electronic charge, and i is an integer or, in some cases, a simple fraction. This behavior is known as the quantum Hall effect. *See* HALL EFFECT.

Multilayered structures composed of alternating layers of magnetic and nonmagnetic metals, such as iron/chromium or cobalt/copper, can feature very large, negative values of magnetoresistance. This effect, called giant magnetoresistance, arises from the spin dependence of the electron scattering which causes resistance. When consecutive magnetic layers have their magnetizations antiparallel

(antiferromagnetic alignment), the resistance of the structure is larger than when they are parallel (ferromagnetic alignment). Since the magnetic alignment can be changed with an applied magnetic field, the resistance of the structure is sensitive to the field. Giant magnetoresistance can also be observed in a simpler structure known as a spin valve, which consists of a nonmagnetic layer (for example, copper) sandwiched between two ferromagnetic layers (for example, cobalt). The magnetization direction in one of the ferromagnetic layers is fixed by an antiferromagnetic coating on the outside, while the magnetization direction in the other layer, and hence the resistance of the structure, can be changed by an external magnetic field. Films of nonmagnetic metals containing ferromagnetic granules, such as cobalt precipitates in copper, have been found to exhibit giant magnetoresistance as well. *See* ANTIFERROMAGNETISM; FERROMAGNETISM; MAGNETIC THIN FILMS; MAGNETIZATION.

Magnetoresistors, especially those consisting of semiconductors such as indium antimonide or ferromagnets such as permalloy, are important to a variety of devices which detect magnetic fields. These include magnetic recording heads and position and speed sensors. *See* MAGNETIC RECORDING.

J. F. Herbst

Magnetosphere

A comet-shaped cavity or bubble around the Earth, carved in the solar wind. This cavity is formed because the Earth's magnetic field represents an obstacle to the solar wind, which is a supersonic flow of plasma blowing away from the Sun. As a result, the solar wind flows around the Earth, confining the Earth and its magnetic field into a long cylindrical cavity with a blunt nose (**Fig. 1**). Since the solar wind is a supersonic flow, it also forms a bow shock a few earth radii away from the front of the cavity. The boundary of the cavity is called the magnetopause. The region between the bow shock and the magnetopause is called the magnetosheath. The Earth is located about 10 earth radii from the blunt-nosed front of the magnetopause. The long cylindrical section of the cavity is called the magnetotail, which is on the order of a few thousand earth radii in length, extending approximately radially away from the Sun. *See* SOLAR WIND.

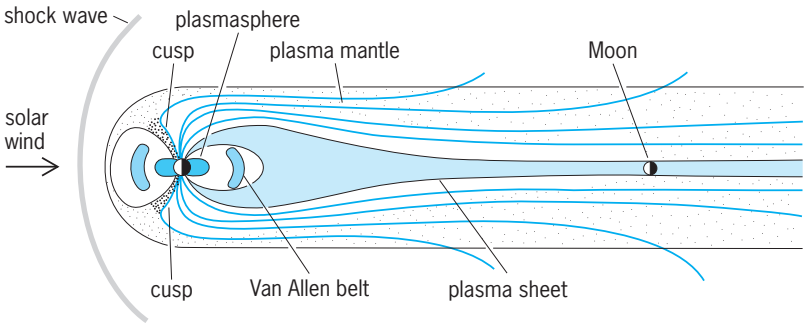

Fig. 1. Noon-midnight cross section of the magnetosphere and the distribution of various plasmas inside it.

The concept of the magnetosphere was first formulated by S. Chapman and V. C. A. Ferraro in 1931, but the term magnetosphere was introduced in 1959 by T. Gold, who defined it as "the region above the ionosphere in which the magnetic field of the Earth has a dominant control over the motions of gas and fast charged particles." It was, however, only during the 1970s that the magnetosphere was extensively explored by a number of satellites carrying sophisticated instruments. This exploration was a part of the effort to understand that region of the Earth's environment above the level where most meteorological phenomena take place, namely, the space between the Sun and the stratosphere/mesosphere. The space which surrounds the magnetosphere is often referred to as geospace. *See* IONOSPHERE; SCIENTIFIC AND APPLICATIONS SATELLITES.

Plasma. The satellite observations have indicated that the cavity is not an empty one, but is filled with plasmas of different characteristics (Fig. 1). Just inside the magnetopause some solar wind plasma blows; this particular plasma region is called the plasma mantle. There are also two funnel-shaped regions (one in each hemisphere), extending from the front magnetopause to the polar ionosphere, which are also filled with solar wind plasma; these regions are called the cusp. The extensive tail region of the magnetosphere is divided into the northern and southern halves by a thin sheet of plasma, called the plasma sheet. The distant tail region (beyond the lunar distance) has been infrequently explored when it was traversed by space probes on their way to the planets. The inner magnetosphere is occupied by the Van Allen belts, the ring current belt, and the plasmasphere. All of them are doughnut shaped and surround the Earth. The ring current belt is colocated with the Van Allen belts and is so named because it carries a large amount of a westward-directed electric current around the Earth.

It is generally believed that the main part of the plasma sheet is of solar origin; the mantle plasma may be an important source of it. The plasmasphere is formed mainly by ionospheric plasma which diffuses outward along magnetic field lines. At times the ring current belt contains heavy ions of ionospheric origin, such as helium and oxgen ions, as well as protons of solar origin. The heavy ions are found also in the plasma sheet.

The Earth's dipolar magnetic field is considerably deformed by these plasmas and the electric currents generated by them. The dipolar field is compressed in the dayside part of the magnetosphere, mainly by the action of the impacting solar wind, whereas it is considerably stretched along the equatorial plane in the magnetotail (Fig. 1). *See* VAN ALLEN RADIATION.

Solar wind–magnetosphere generator. If there were no magnetic field in the solar wind, the Earth's magnetic field would simply be confined by the solar wind into a teardrop-shaped cavity (rather than a comet-shaped cavity), and the cavity would be

almost empty. It is the magnetic field of the solar wind which introduces considerable complexity in the interaction between the solar wind and the magnetosphere. Some of the magnetic field lines originating from the polar region are almost always connected to the solar wind magnetic field lines across the magnetopause (**Fig. 2**). In fact, because the solar wind magnetic field lines originate from the Sun, it may be said that the Sun and the Earth are connected by magnetic field lines. More specifically, the connected field lines originate from an oval area in the polar region. This area is called the polar cap and is bounded by a belt of the aurora called the auroral oval. The magnetic field lines originating from the northern polar cap are stretched along the northern half of the magnetotail as a bundle of magnetic field lines and are connected to the solar wind magnetic field lines across the magnetopause. Similarly, the magnetic flux from the southern polar cap occupies the southern half of the magnetotail.

This connection between the solar wind magnetic field lines and the Earth's magnetic field lines provides the most important aspect of the solar wind–magnetosphere interaction. Since the solar wind (a fully ionized plasma and thus a conductor) must blow across the connected magnetic field lines as it flows along the magnetopause (Fig. 2), an electromotive force is generated. This force powers the solar wind–magnetosphere generator and thus provides practically all the electric power for various magnetospheric processes, such as auroral phenomena and magnetospheric disturbances, including geomagnetic storms. It is precisely in this way that the solar wind (a flow of magnetized plasma) couples its energy to the magnetosphere. Therefore, if there were no connection of the field lines, there would be no energy transfer from the solar wind to the magnetosphere, and thus there would be no auroral phenomena, which are one of the most spectacular wonders of nature.

Because the solar wind–magnetosphere generator is operated on the magnetopause, where the solar wind blows across the connected field lines, the whole cylindrical surface of the magnetotail can be considered to be the generator in which the current flows from the duskside magnetopause to the dawnside magnetopause along the cylindrical surface. Much of the current thus generated flows in the plasma sheet across the magnetotail; this part of the circuit may be called the cross-tail circuit (**Fig. 3**). Thus, the magnetotail may be viewed as two long solenoids, producing two bundles of magnetic field lines separated by the plasma sheet. As a result, the Earth's dipolar field is stretched along the magnetotail in the antisolar direction. The generator also powers a large-scale motion of plasmas in the magnetosphere. In the main part of the plasma sheet, the plasma flows toward the Earth. *See* GEOMAGNETIC VARIATIONS; GEOMAGNETISM; SOLAR MAGNETIC FIELD.

Auroral phenomena. Another important circuit that is connected to the generator consists of a current flow along magnetic field lines from the morning side of the magnetotail to the morning half of the auroral oval, a flow along the auroral ionosphere (the auroral oval), and a flow along magnetic field lines from the evening half of the auroral oval to the evening side of the magnetotail. This discharge circuit may be called the auroral circuit because it is responsible for causing acceleration of current-carrying charged particles (Fig. 3). When the accelerated particles collide with upper atmospheric particles, they excite or ionize them. The auroral lights are emitted by those excited or ionized particles when they return to a lower or the ground state.

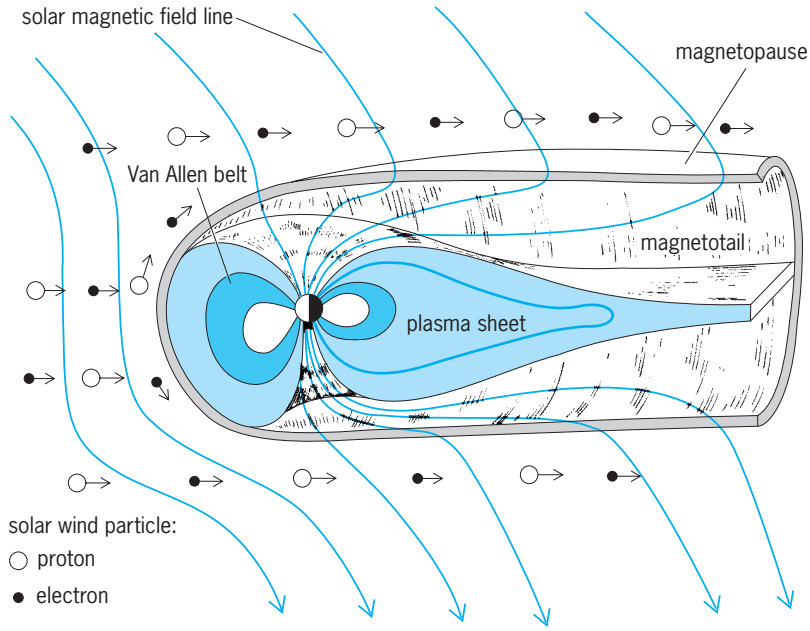

Fig. 2. Internal structure of the magnetosphere (noon-midnight cut). Motions of solar wind protons and electrons around the magnetopause and the linkage of the solar wind magnetic field lines and Earth's magnetic field lines are also shown.

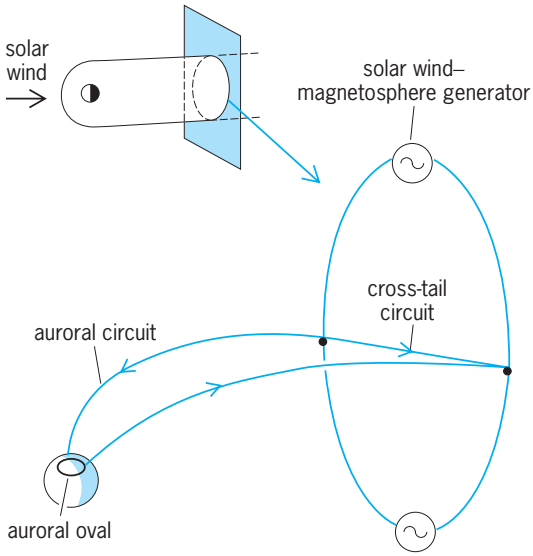

Fig. 3. The solar wind–magnetosphere generator located on the magnetopause (one cross section is shown) and its two main circuits: the cross-tail circuit and the auroral circuit.

The total power generated by the solar wind–magnetosphere generator depends on the magnitude and direction of the solar wind magnetic field, as well as on the speed of the solar wind. In particular, when the solar wind magnetic field is directed southward, the generated power becomes maximum. On the other hand, when the solar wind magnetic field is directed northward, the power becomes minimum. The power is also greater if the magnitude of the solar wind magnetic field B and the speed of the solar wind V are higher. The power P is given by the equation below, where θ is the polar angle of the

$$P = VB^2 \sin^4\left(\frac{\theta}{2}\right) l_0^2$$

solar wind vector projected to a plane perpendicular to the Sun-Earth line and l_0 is a constant of about 7 earth radii.

When the solar wind magnetic field vector turns gradually from a northern to a southern direction in the vicinity of the magnetosphere, the power generated by the dynamo begins to increase. As a result, the cross-tail electric current in the plasma sheet is increased. As the generated power is very high, the consumption of the power by the magnetosphere also becomes high. An indication of this high consumption rate is the phenomenon called the magnetospheric substorm which manifests itself in various magnetospheric and polar upper-atmospheric disturbances. As a result of this process, the earthward edge of the plasma sheet is suddenly pushed toward the Earth, and the plasma thus injected into the inner magnetosphere forms the ring current belt.

At the same time, the electric current along the auroral circuit is also intensified. As a result, the aurora becomes bright and active. The enhanced ionospheric portion of the auroral circuit is called the auroral electrojet. Intense magnetic field variations are experienced under and near the auroral electrojet. In this way, effects of the enhanced power of the solar wind–magnetosphere generator are manifested most conspicuously in the polar region. An intense auroral activity associated with this phenomenon is called the auroral substorm. The associated intense magnetic disturbances are called polar magnetic substorms. The ionosphere is also greatly disturbed in this period, and this phenomenon is called the ionospheric substorm. The ionospheric disturbances greatly affect high-frequency and very high-frequency radio communication in high latitudes. Even satellite microwave communication is at times disturbed by auroral activity.

Because the Earth is a good conductor, the auroral electrojets induce a potential difference on the Earth's surface. Therefore, if two points, along which the potential difference is present, are connected by a conductor, an electric current flows along it. It is for this reason that electric currents are induced in oil and gas pipelines and in power transmission lines in high latitudes during auroral activity. In the power lines there are severe power fluctuations when the electrojets are particularly intense.

Various solar activities are responsible for the time variations of the three solar wind quantities which determine the power of the generator. In particular, an intense explosion on the solar surface, a solar flare, causes a gusty solar wind and a large-scale deformation of the solar wind magnetic field. If this deformation happens to produce a large southward-directed magnetic field around the magnetosphere for an extended period (approximately 6 to 24 h), the generated power is considerably enhanced, and intense magnetospheric substorms occur. As a result, an unusually intense ring current grows, carrying a westward current surrounding the Earth and producing a southward-directed magnetic field around the Earth. Since the Earth's magnetic field is directed northward, this addition of the southward field is observed as a decrease of the Earth's magnetic field, one typical aspect of geomagnetic storms.

The Sun has also a few regions from which a high-speed stream is generated, particularly during the declining epoch of the 11-year sunspot cycle. Such a stream lasts for an extended period, from several months to more than 1 year. The solar wind magnetic field is markedly variable in the high-speed stream. As the Sun and the stream rotate with a period of about 27 days, the Earth is immersed in the stream once each period. As a result, the power of the solar wind–magnetosphere dynamo is increased and fluctuates considerably, causing magnetospheric substorms and auroral activity. *See* AURORA; SUN.

Planets. All other magnetic planets, such as Mercury, Jupiter, and Saturn, have magnetospheres which are similar in many respects to the magnetosphere of the Earth. The magnetosphere of Mercury is much smaller than that of the Earth. Since the solar wind pressure is higher at Mercury's than at the Earth's distance and since the magnetic field of Mercury is weaker than that of the Earth, the blunt-nosed front of the magnetosphere of Mercury is located at a distance of only 0.5 Mercury radius from its surface. On the other hand, because of a weak solar wind, a very strong magnetic field, and a fast rotation, the magnetosphere of Jupiter is gigantic and has an extensive current disk which lies along its equator. The distance to the blunt-nosed front is about 100 Jupiter radii. The *Voyager 2* space probe found that the magnetosphere of Jupiter is filled with sulfuric ions, presumably injected from one of its satellites, Io. The space probe also traversed its gigantic magnetotail. The magnetosphere of Saturn is a little smaller than that of Jupiter, but extends well beyond its rings. *See* JUPITER; MERCURY (PLANET); PLANETARY PHYSICS; SATURN.

A great variety of plasma processes, plasma waves, and plasma instabilities are involved in all magnetospheric phenomena. Some of them occur commonly in a thermonuclear machine. The magnetosphere provides a natural laboratory in which these plasma behaviors can be studied without the so-called wall effects. Among a variety of astrophysical conditions, the magnetospheres of the Earth and the planets are the only regions in which plasma behaviors can be studied in natural conditions. Therefore, a study of

the magnetosphere provides an important foundation upon which astrophysical plasma theories can be tested and constructed. It has also been speculated that many magnetic stars, pulsars, and even some galaxies have structures similar to the Earth's magnetosphere. Since a magnetosphere results from an interaction of magnetized plasma flow and a magnetized celestial body, it is likely that it is of common occurrence in galaxies. *See* MILKY WAY GALAXY; PLASMA (PHYSICS); PULSAR. S.-I. Akasofu

Bibliography. S.-I. Akasofu, *Physics of Magnetospheric Substorms*, 1977; S.-I. Akasofu and S. Chapman, *Solar-Terrestrial Physics*, 1972.

Magnetostriction

The change of length of a ferromagnetic substance when it is magnetized. More generally, it is the phenomenon that the state of strain of a ferromagnetic sample depends on the direction and extent of magnetization. The phenomenon has an important application in magnetostriction transducers.

Physical cause. Magnetostriction results from the dependence of the crystalline anisotropy energy upon the state of strain of the crystalline lattice. If the crystal deforms (for example, suffers a change in length), the anisotropy energy may be lowered more than the elastic energy is raised. Thus, a strained state will be favored. For a discussion of crystalline anisotropy energy *see* FERROMAGNETISM.

The total energy of a ferromagnetic substance depends upon the state of strain and the direction of magnetization through three contributions. The first two consist of the crystalline anisotropy energy of the unstrained lattice plus a correction which takes into account the dependence of the anisotropy energy on the state of strain. The third contribution is that of the elastic energy, which is independent of magnetization direction and is a minimum in the unstrained state. The state of strain of the crystal will be that which makes the sum of the three contributions to the energy a minimum. The result is that, when magnetized, the lattice is always distorted from the unstrained state, unless there is no anisotropy.

Since spontaneous magnetization occurs below the Curie temperature, there will always be a spontaneous lattice distortion which depends on magnetization direction in the ferromagnetic state. In nickel, the lattice spacing parallel to the magnetization is always smaller than the lattice spacing perpendicular to the magnetization. *See* CURIE TEMPERATURE.

Magnetoelastic coupling constants. These determine the magnitude of the correction arising from strains to the anisotropy energy. In cubic crystals, there are two magnetoelastic coupling constants B_1 and B_2. The constant B_1 determines the change in anisotropy due to a diagonal component of strain, and B_2 that due to a mixed component. The values of the strain components which lead to a minimum in the total magnetoelastic energy are given in terms of the magnetoelastic coupling constants, the elastic constants, and the direction cosines of the magnetization with respect to the crystal axes, α_1, α_2, α_3. In the strain state of lowest magnetoelastic energy, that part of the energy which depends upon magnetization direction is given by the equation below, where

$$U = (K_1 + \Delta K) \times (\alpha_1{}^2\alpha_2{}^2 + \alpha_2{}^2\alpha_3{}^2 + \alpha_3{}^2\alpha_1{}^2) + \cdots$$

ΔK is a correction to the first-order anisotropy constant K_1. The quantity ΔK depends only upon the magnetoelastic constants and the elastic constants.

When a high permeability (soft) magnetic material is required, the magnetostriction should be small in order that anisotropy not be induced by lattice distortions. Magnetostriction appears to be a major source of transformer hum. As the silicon content of soft magnetic steel is increased toward 6.5%, the magnetostriction disappears. Unfortunately, the metal becomes excessively brittle.

Applications. The magnetostrictive effect is exploited in transducers used for the reception and transmission of high-frequency sound vibrations. Nickel is often used for this application. *See* ELASTICITY; SONAR; ULTRASONICS.

Elihu Abrahams; Frederick Keffer

Bibliography. D. Craik, *Magnetism*, 1994; E. Du Tremolet de Lacheisseri, *Magnetostriction: Theory and Applications—Magnetoelasticity*, 1993.

Magnetron

The oldest of a family of crossed-field microwave electron tubes wherein electrons, generated from a heated cathode, move under the combined force of a radial electric field and an axial magnetic field. By its structure a magnetron causes moving electrons to interact synchronously with traveling-wave components of a microwave standing-wave pattern in such a manner that electron potential energy is converted

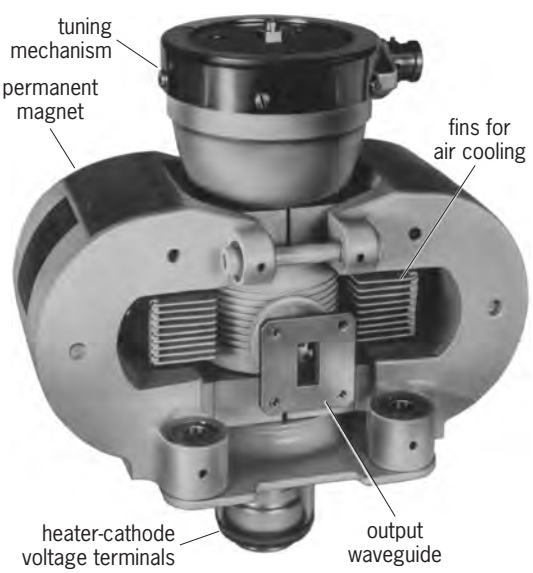

Fig. 1. Coaxial cavity magnetron with horseshoe-shaped magnets. (*After J. W. Gewartowski and H. A. Watson, Principles of Electron Tubes, Van Nostrand, 1965*)

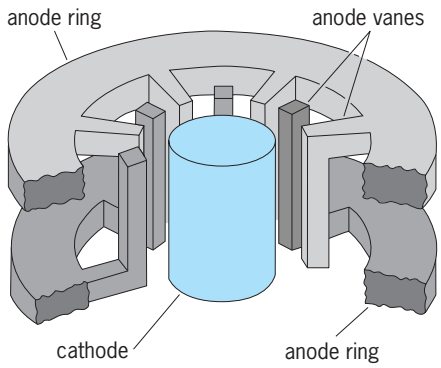

Fig. 2. Interdigital-vane anode circuit and cathode, indicating the basic cylindrical geometry. (*After G. D. Sims and I. M. Stephenson, Microwave Tubes and Semiconductor Devices, Blackie and Son, London, 1963*)

to microwave energy with high efficiency. Magnetrons have been used since the 1940s as pulsed microwave radiation sources for radar tracking. Because of their compactness and the high efficiency with which they can emit short bursts of megawatt peak output power, they have proved excellent for installation in aircraft as well as in ground radar stations. In continuous operation, a magnetron can produce a kilowatt of microwave power, which is appropriate for rapid microwave cooking.

The magnetron in an oscillator; it generates radio-frequency energy, usually over limited portions of the microwave frequency range of 1–40 gigahertz. It can produce microsecond pulses of peak power as high as 10 megawatts but with low duty-cycle ratios (ratio of pulse length of pulse repetition period). Mechanical tuning of the magnetron microwave circuit can change the center frequency of the output radiation by about ±5%.

Configuration. Raw materials of a magnetron are chiefly copper, iron, and ceramic welded and brazed together to form a vacuum tube. A slow-wave microwave circuit, made of copper to reduce ohmic losses, serves to provide a traveling wave that can interact with the electrons. Iron pole pieces guide the magnetic flux from externally attached horseshoe-shaped magnets (**Fig. 1**). Pole pieces shape the kilogauss magnetic field in the electron-microwave interaction region to be perpendicular to the electric field and to constrain the electrons to the interaction region. Ceramic (or sometimes glass) serves as a standoff insulator for the tens of kilovolts applied between cathode and anode. Ceramic is also used as the microwave window to couple the power generated inside the magnetron through the vacuum enclosure to externally attached waveguides. The breakdown limit of this window may be the chief factor in setting the peak power output obtainable from a magnetron.

The magnetron is a device of essentially cylindrical symmetry (**Fig. 2**). On the central axis is a hollow cylindrical cathode. The outer surface of the cathode carries electron-emitting materials, primarily barium and strontium oxides in a nickel matrix. Such a matrix is capable of emitting electrons when current flows through the heater inside the cathode cylinder. *See* VACUUM TUBE.

At a radius somewhat larger than the outer radius of the cathode is a concentric cylindrical anode. The anode serves two functions: (1) to collect electrons emitted by the cathode and (2) to store and guide microwave energy. The anode consists of a series of quarter-wavelength cavity resonators symmetrically arranged around the cathode. *See* CAVITY RESONATOR.

A radial dc electric field (perpendicular to the cathode) is applied between cathode and anode. This electric field and the axial magnetic field (parallel and coaxial with the cathode) introduced by pole pieces at either end of the cathode, as described above, provide the required crossed-field configuration.

Microwave generation. Provided that the magnitudes of the electric field and the magnetic field are properly adjusted, an electron emerging from the heated cathode proceeds to orbit the cylindrical cathode, moving on the average in a direction perpendicular to both applied fields (**Fig. 3**). The totality of electrons emitted by the cathode form a swarm or hub of negative charge rotating about the cathode axis. The hub thickness extends from the cathode surface outward to a radius intermediate between cathode and anode radii. Motion of the rotating swarm of electrons past the surface of the concentric-anode microwave circuit induces a noise current in the copper circuit. This noise shock excites the resonators so that microwave fields build up at the resonant frequency. The velocity of the electrons, as determined by the applied voltage and magnetic field, is made to be close to or synchronous with the slow-wave phase velocity characteristic of the microwave circuit. Consequently, as the electron swarm rotates, it concentrates into bunches that deliver microwave energy to the resonators. Thus, as in most oscillators, the process begins with noise; then electromagnetic fields build up at the resonant

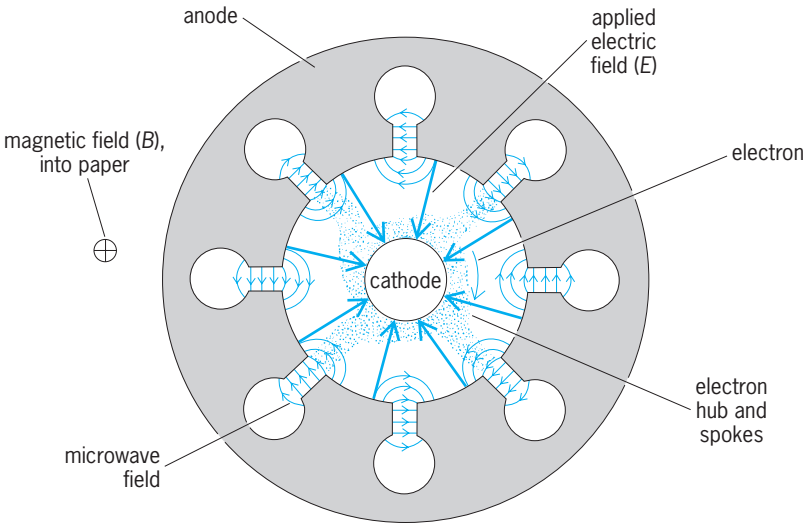

Fig. 3. Electron motion in magnetron interaction region under action of applied electric and magnetic fields and generated microwave field. (*After C. H. Dix and W. H. Aldous, Microwave Valves, Illiffe Books, London, 1966*)

frequency. These fields react, in turn, upon the electrons, remove some of their potential energy, and cause them to approach the circuit more closely, whereupon the process intensifies. Buildup of the microwave field is limited by internal ohmic losses in the circuit, external loading of the output coupling, and the electron energy available for conversion to microwave energy.

Electron bunching. Bunching of electrons plays a prominent role in the generation of microwave power in the magnetron, as it does in most microwave tubes. The applied radial electric field E and the axial field B constrain the average motion of the hub electrons to a direction perpendicular to both fields and hence circumferential about the cathode. (There is cycloidal motion superimposed on the average, as well.) The average circumferential electron velocity, called drift velocity, has a magnitude $u = E/B$. The microwave circuit is designed to support a traveling wave whose phase velocity υ is equal to u. If $u = \upsilon$, the electron hub rotation is said to be synchronous with the rotating traveling wave. Under these conditions, one can consider the interaction of electrons with microwave fields from the anode in a frame of reference which rotates at the synchronous velocity.

The quarter-wave cavity resonators, which support the rotating microwaves and surround the electron hub, present at the hub a rotating microwave field configuration consisting of electric fields having strong circumferential components. These components are 180° out of phase (oppositely directed) for neighboring cavities, as shown in Fig. 3. This set of alternating fields can be considered as composed of two sets of uniformly directed fields (each internally inphase), one set 180° out of phase with the other and the two interlaced. (Such a microwave field configuration is called a π mode, the two sets of fields being phased π radians from each other.) In the frame of reference rotating in synchronism with these two sets of microwave fields, the circumferential components of the microwave electric fields can be regarded as steady or static fields crossed with the applied axial magnetic field.

Those hub electrons rotating in phase and under the influence of one of the sets of microwave fields, because of the crossed-field situation, are constrained to move radially toward the anode (average electron motion is always perpendicular to both electric and magnetic crossed-field directions). In doing so they contribute to the strength of the microwave field, converting their potential energy in the applied electric field to microwave energy. The process continues until the electrons finally fall through the entire hub-to-anode potential and are collected on the anode. As a consequence, the distribution of electrons in the interaction region between cathode and anode looks like a hub with spokes, all rotating together at drift velocity E/B (Fig. 3). In this sense the electrons are bunched. Throughout the interaction the electrons continue to rotate at close to synchronous velocity, remaining in phase with the wave and continuing to convert dc potential energy into microwave energy until striking the anode. This staying-in-synchronism results in high energy conversion efficiency (values may be as high as 65–75%) of the magnetron in contrast to those microwave tubes that depend on differential velocities of electrons for bunching. It also means that the anode must be rugged enough to absorb the full kinetic energy of the electrons ($mu^2/2$, where m is the electron mass). *See* KLYSTRON.

Those electrons in phase with the remaining set of microwave fields (the set 180° out of phase with the first) will be constrained to move radially back through the hub toward the cathode. This motion results in an increase of electron potential energy at the expense of the microwave field. However, the loss in microwave energy is small because such electrons move out of high microwave field regions toward weaker ones and so move to minimize their interaction. On striking the cathode these electrons heat up the cathode and reduce the required heater power. The phenomenon of electron bunching, wherein certain in-phase electrons give up their potential energy to the microwave field while other out-of-phase electrons are forced out of the microwave field region to prevent their extracting microwave energy, is further abetted by the shape of the microwave fields near the anode. The field shape is such as to cause electrons approaching the anode to become even more closely in phase so as to deliver maximum power to the circuit.

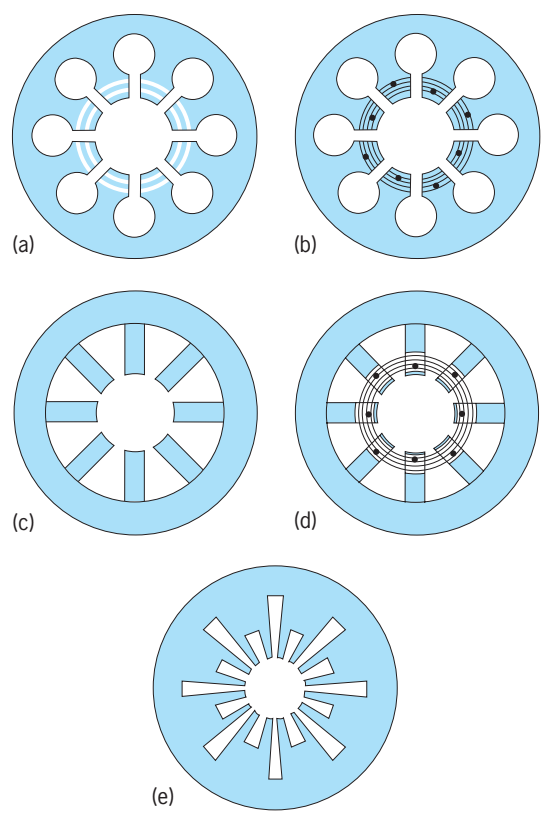

Fig. 4. Magnetron slow-wave circuits. (*a*) Unstrapped cavity. (*b*) Strapped cavity. (*c*) Unstrapped vane block. (*d*) Strapped vane block. (*e*) Rising-sun block. (*After C. H. Dix and W. H. Aldous, Microwave Valves, Illiffe Books, London, 1966*)

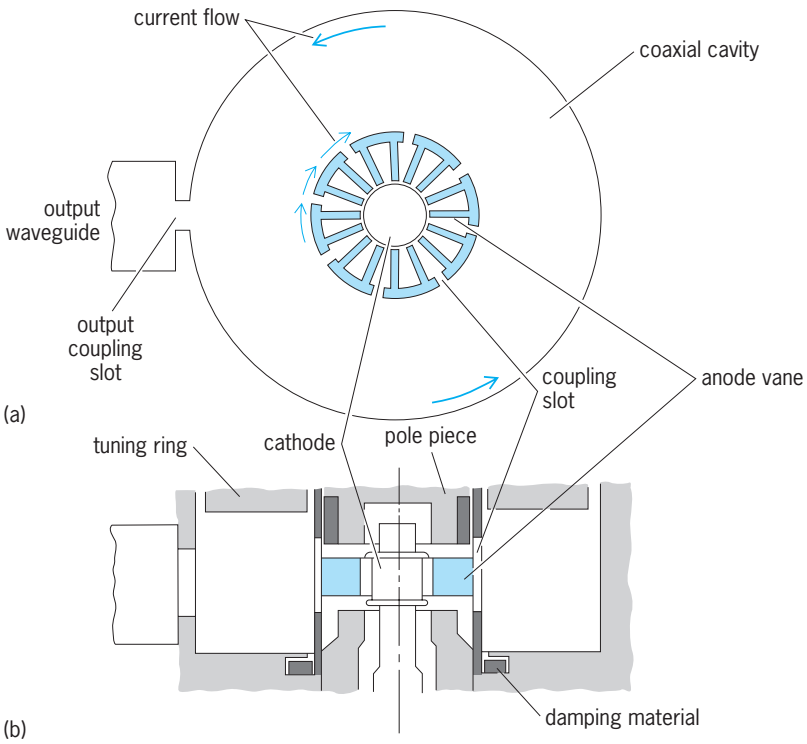

Fig. 5. Coaxial cavity magnetron. (*a*) End view. (*b*) Cross-sectional view. (*After J. W. Gewartowski and H. A. Watson, Principles of Electron Tubes, Van Nostrand, 1965*)

Microwave circuit. The only aspect of the magnetron that is not symmetric is the output coupler. Being an oscillator and generating its own signal starting from noise-induced fields, the magnetron must maintain enough energy in its resonant circuit to keep the generation process going. On the other hand, only the microwave energy coupled out of the magnetron is useful to the radar operator. Thus some compromise must determine the fraction of the generated power which is coupled out through the output coupling slot, the ceramic window, and output waveguide.

The π-mode electromagnetic field configuration is desirable for interaction with the electrons as described above, and the output coupling is generally designed for such a resonant condition. However, the microwave circuit can sustain other resonant modes which can interact with the electrons and which are generally less efficient. Some of these undesirable modes can so orient themselves as to present a null in the microwave field pattern at the output. All the energy converted from interaction with electrons is then stored inside the magnetron in such modes, and none is usefully coupled out. The fields in the interaction region can become high enough to block operation in the π mode. Such undesirable modes can be close in frequency to the π mode, and unless special circuit design is incorporated, they are difficult to discriminate against. Much of the design of magnetron microwave structures has to do with maximizing efficiency of π-mode interaction while inhibiting electron interaction with other undesirable modes.

A variety of shapes have been used in the design of the quarter-wavelength cavities whose open-end, capacitative section provides the electric fields for interaction with the electrons. There are vane-type cavities and hole-and-slot versions (**Fig. 4**). These may be strapped (alternate cavities tied together with a copper strap to promote π-mode operation), or alternate cavities can be elongated (rising-sun configuration) to separate the undesirable modes from the π mode. Such circuits are generally tuned in frequency by inserting copper pins into the induction sections of the quarter-wave cavities. Voltage tuning has been achieved in certain low-power magnetrons.

Coaxial cavity magnetron. A more recent magnetron design, the coaxial cavity magnetron (**Fig. 5**), achieves mode separation, high efficiency, stability, and ease of mechanical tuning by coupling a coaxial high-Q cavity to a normal set of quarter-wavelength vane cavities. *See* CAVITY RESONATOR; Q (ELECTRICITY).

The high-Q cavity is resonant in the TE_{011} mode, a mode in which current flow on the inner conductor is circumferential and phase is not a function of circumferential position. To couple microwave energy from the slow-wave circuit into the coaxial cavity, slots are cut into the backs of alternate quarter-wavelength cavities. If the rotation of the electron hub induces current flow on the vanes, so that a π-mode resonant buildup of microwave fields commences there, the slot coupling geometry is correct to also stimulate the buildup of the TE_{011} mode in the outer coaxial cavity. Because alternate vane cavities are in phase in π-mode operation, the current coupled into the high-Q cavity through slots cut only in alternate cavities will also be in phase. This is the condition required to induce the TE_{011} mode in the coaxial cavity. The output frequency can be tuned merely by altering the axial length of the outer coaxial cavity (a tuning piston is employed). Most stored microwave energy is in the coaxial cavity, and its high Q ensures efficient storage. The coaxial cavity magnetron has replaced many earlier designs for airborne radar.

Some characteristics typical for the coaxial cavity magnetron are as follows.

Frequency (tunable)	15.5–17.5 GHz
Peak output power	125 kW
Peak applied voltage	17.5 kV
Peak anode current	19 A
Pulse duration	3μs
Duty cycle	0.001s
Magnetic field	0.74 weber/in.^{2} (1150 teslas)
Overall efficiency	43

Other crossed-field tubes. The magnetron is a member of the crossed-field family of devices. Other members are platinotrons, amplitrons, forward-wave crossed-field amplifiers, non-reentrant-beam crossed-field amplifiers, and carcinotrons. Some of these devices are oscillators and others amplifiers (**Fig. 6**).

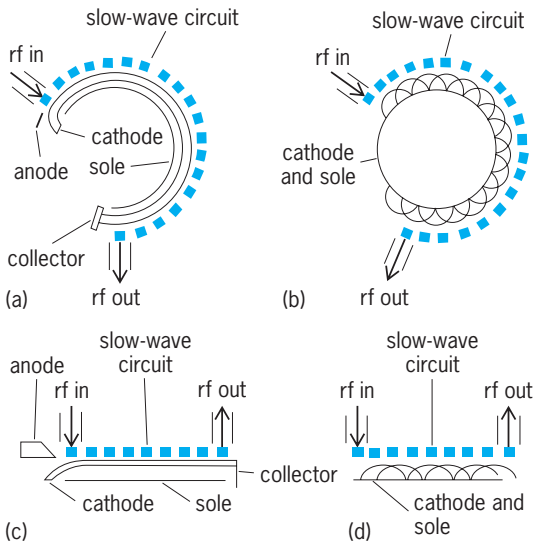

Fig. 6. Magnetron amplifiers. (*a*) Circular type with beam injection. (*b*) Circular type with continuously emitting sole. (*c*) Linear type with beam injection. (*d*) Linear type with continuously emitting sole.

Some interact with backward slow waves, others with forward waves. Some employ an electron gun, others an emitting sole. The magnetron oscillator itself has both a reentrant electron beam and a reentrant microwave circuit. (The electron flow is continuous in the circumferential direction, orbiting the axial cathode, while the quarter-wavelength cavities are symmetrically disposed around the cathode with no beginning or end.)

Power generation in an oscillator begins with noise, and the resonant nature of the circuit produces the high electric fields required to further react with the electrons. Amplifiers, on the other hand, have non-reentrant circuits with a definite input and output. The input signal provides the initial microwave field to begin interaction between electrons and circuit. The electron beam may be reentrant or not, depending upon the geometry of the circuit, that is, whether the cavities are arranged with cylindrical or linear symmetry.

The magnetron and related crossed-field amplifiers and oscillators are rather noisy devices and operate best in the high-output-power region. As amplifiers, they have generally low gain, but their output power is high. Compactness and high efficiency are the most attractive features of crossed-field electron tubes. Conversion of electron potential energy to microwave energy in characteristic of all crossed-field devices (M-type carcinotrons) in contrast to conversion of kinetic energy in the case of traveling-wave tubes (O-type devices). *See* TRAVELING-WAVE TUBE.

R. J. Collier

Bibliography. R. E. Collin, *Foundations for Microwave Engineering*, 2d ed., 2000; A. S. Gilmour, *Microwave Tubes*, 1986; J. Helszajn, *Microwave Engineering*, 1992; C. A. Lee and G. C. Delman, *Microwave Devices, Circuits and Their Interaction*, 1994; S. Liao, *Microwave Devices and Circuits*, 3d ed., 1996.

Magnification

A measure of the effectiveness of an optical system in enlarging or reducing an image. For an optical system that forms a real image, such a measure is the lateral magnification m, which is the ratio of the size of the image to the size of the object. If the magnification is greater than unity, it is an enlargement; if less than unity, it is a reduction.

The ratio of the longitudinal (with respect to the optical axis) dimensions of the image to the corresponding dimensions of the object is known as longitudinal magnification, which in first order equals the square of the lateral magnification.

The angular magnification γ is the ratio of the angles formed by the image and the object at the eye. The relation $n'\gamma m = n$ relates angular to lateral magnification. Here n and n' are the refractive indices of the media containing the object and image, respectively. In telescopes the angular magnification (or, better, the ratio of the tangents of the angles under which the object is seen with and without the lens, respectively) can be taken as a measure of the effectiveness of the instrument.

A small off-axis element is imaged with different magnification in both the meridional and sagittal directions. This may be called differential magnification.

Magnifying power is the measure of the effectiveness of an optical system used in connection with the eye. The magnifying power of a spectacle lens is the ratio of the tangents of the angles under which the object is seen with and without the lens, respectively. The magnifying power of a magnifier or an ocular is the ratio of the size under which an object would appear seen through the instrument at a distance of 10 in. (25 cm; the distance of distinct vision) divided by the object size.

E. Abbe suggested defining the magnifying power of an optical system as the ratio of the tangent of the visual angle under which the object appears to the object size. This quantity is approximately equal to the power of the system, which is the reciprocal of the focal length, $1/f$. *See* DIOPTER; LENS (OPTICS); OPTICAL IMAGE.

Max Herzberger

Bibliography. M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; M. Herzberger, *Modern Geometrical Optics*, 1958, reprint 1980; F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976; J. R. Meyer-Arendt, *Introduction to Classical and Modern Optics*, 4th ed., 1995.

Magnitude (astronomy)

The brightness of an astronomical object, expressed on a unique numerical scale. The stellar magnitude scale is logarithmic and is inverted in that fainter objects have numerically larger magnitudes. The term "magnitude" in this context refers to brightness only and has nothing to do with the size or other quantities. Although used primarily for stars, the stellar magnitude scale can also be used to

express the brightness of the Sun, planets, asteroids, comets, nebulae, galaxies, and even background radiation.

Since the brightness of any object varies with wavelength, many different magnitude scales have been defined corresponding to different spectral regions, bandwidths, and methods of observation. Visual magnitudes, corresponding to the sensitivity of the human eye centered in the yellow part of the spectrum, are usually implied if the type is unspecified. A distinction is made between the apparent magnitude of an object viewed from the Earth and its absolute magnitude, which measures the object's intrinsic luminosity by indicating its apparent magnitude as seen from a standard distance.

Development. The star catalog of Hipparchus (about 150 B.C.) is thought to have contained approximately 850 naked-eye stars classified according to brightness. The 15 or so brightest stars were referred to as stars of the first magnitude, while second-magnitude stars were on the average two or three times fainter, and so on. The scale is logarithmic because intervals that are perceived as equal intervals are, in fact, equal brightness ratios.

Measurements of brightness ratios in the nineteenth century showed that, on average, stars of the sixth magnitude (near the limit of naked-eye vision) were about 100 times fainter than those of the first. On the scale introduced by N. R. Pogson in 1856 and universally adopted, an interval of 5 magnitudes corresponds to a factor of exactly 100, so that each magnitude corresponds to a factor of $\sqrt[5]{100} \approx 2.512\ldots$. The zero point of the Pogson scale was set so that most stars retained their customary magnitudes.

Mathematical properties. The stellar magnitude scale has many convenient properties. Its logarithmic character means that brightness ratios can be calculated by simple subtraction, and that the entire range of astronomical brightnesses can be represented in terms of small numbers, ranging from -27 (the apparent magnitude of the Sun) to approximately $+31$ (the faintest objects detected with confidence on deep-field images taken with the *Hubble Space Telescope*). Its inverted nature seems natural to observers, who are apt to think in terms of a star's faintness, that is, how hard it is to make a particular observation. It is particularly convenient in that the most familiar stars have magnitudes near zero.

The correspondence between magnitude interval and brightness ratio is shown in **Table 1**. In general, an interval of n magnitudes corresponds to a factor of 2.512^n in brightness. A factor of 10 corresponds to exactly 2.5 magnitudes, a factor of 100 to exactly 5 magnitudes. Any linear measure of brightness X can be expressed as a magnitude m through Eq. (1), where the negative sign is needed to produce

$$m = -2.5 \log_{10} X + \text{ constant} \qquad (1)$$

an inverted scale and the constant can be evaluated by measuring X for any star of known magnitude. *See* LOGARITHM.

TABLE 1. Magnitude scale

Magnitude difference	Brightness ratio
0.00	1
0.01	1.01
0.05	1.05
0.30	1.32
0.753	2
1.000	$2.512\ldots$ ($=\sqrt[5]{100}$)
1.5	~4
2	$6.31\ldots$ ($=2.512^2$)
2.5	10 exactly
3	~16 ($=2.512^3$)
3.5	$25.12\ldots$
4	~40
5	100 exactly
10	10^4
15	10^6
50	10^{20}

An attractive feature of the magnitude scale is the ease with which fractional magnitudes can be interpreted. Each change of 1% in the brightness of an object corresponds to a change of 0.01 in the magnitude, and this numerical correspondence holds to good accuracy for changes up to about 30%.

Absolute magnitude. The apparent magnitude m of a star observed from the Earth depends upon both its absolute magnitude M (an intrinsic property of the star) and its distance d from Earth. The difference $m - M$ between the apparent and absolute magnitude is thus a function of the distance, given by Eq. (2), and is sometimes called the distance modulus.

$$m - M = 5 \log_{10} d - 5 \qquad (2)$$

Here the distance d is expressed in parsecs (1 pc = 3.26 light-years = 1.92×10^{13} mi = 3.09×10^{13} km). The constants in the formula come from the inverse-square law of light, the definition of a magnitude, and the choice of $d = 10$ pc as the standard distance at which the apparent and absolute magnitudes are defined to be equal. Thus the absolute magnitude may be defined as the apparent magnitude an object would have if viewed from a distance of 10 pc. *See* PARSEC.

The absolute magnitude can thus be found for any astronomical object whose distance has been measured (**Table 2**). The Sun has an absolute visual magnitude of $+4.79$; that is, it would appear as about a fifth-magnitude star to an observer 10 pc away. The brightest stable stars in the Milky Way Galaxy and in external galaxies have absolute magnitudes about -9, while the faintest are near $+20$ or even fainter. Supernova 1987A in the Large Magellanic Cloud ($m - M \approx 18.5$) reached an absolute magnitude of -15.5 at its maximum of May 1987; more typical supernovae attain even higher luminosities. *See* STAR; SUPERNOVA.

It is often possible to assign an absolute magnitude to a star on the basis of its spectral classification, in which case its distance can be calculated from Eq. (2). The problem of determining absolute

TABLE 2. Apparent and absolute visual magnitudes and distances of familiar stars

Name	m	M	d, pc	Comment
Sun	−26.78	4.79	$\frac{1}{206{,}265}$	
Barnard's star	9.54	13.25	1.81	Nearby red dwarf
Sirius	−1.46	1.42	2.65	The brightest star
Arcturus	−0.06	−0.3	11	Red giant
Betelgeuse	0.5 var.	−5.2	130	Cool supergiant
Deneb	1.25	−7.2	500	Very bright supergiant
SN 1987A (maximum)	3.0	−15.5	50,000	Supernova in Large Magellanic Cloud

magnitudes is thus equivalent to that of determining distances.

Absorption by interstellar dust affects the apparent magnitudes of distant stars, particularly those lying in the plane of the Milky Way. If the absorption A is expressed in magnitudes, Eq. (2) may be rewritten as Eq. (3).

$$m - M = 5 \log_{10} d - 5 + A \tag{3}$$

See INTERSTELLAR EXTINCTION.

Visual magnitudes. The scale of visual magnitudes was originally established with the dark-adapted eye as the detector. Visual magnitudes can be measured photographically (in which case they are called photovisual) by using an emulsion sensitive to yellow light and a filter to block the shorter wavelengths. However, most modern measurements of visual magnitudes are done photoelectrically, with either a photomultiplier or a charge-coupled device (CCD), in combination with a suitable filter; the symbol V is reserved for photoelectrically determined visual magnitudes. Eye estimates are often employed to monitor the visual magnitudes of variable stars. *See* CHARGE-COUPLED DEVICES; PHOTOELECTRIC DEVICES; PHOTOMETRY; VARIABLE STAR.

Photographic magnitudes. Magnitudes of stars can be determined from image sizes on astronomical photographs. Magnitudes can be found by interpolation if they are known for a sequence of stars appearing on the same photograph. The advantage of photographic photometry is that many thousands of stellar images can often be recorded on a single exposure.

The basic photographic emulsion is sensitive to blue light, and the term photographic magnitude (designated m_{pg}) specifically means a wideband blue magnitude that includes the near-ultraviolet but excludes the yellow region and all longer wavelengths. Star catalogs usually list either m_{pg} or the visual magnitude m_v. When both are known, the difference can be used as a quantitative measure of the color of the star, given by Eq. (4). The zero point of the m_{pg} scale

$$\text{Color index} = m_{pg} - m_v \tag{4}$$

has been set so that $m_{pg} - m_v = 0.0$ for stars similar to Vega, of spectral type A. The color index is positive for stars redder than Vega, negative for bluer stars. Although colors can be affected by interstellar absorption, they serve primarily as indicators of surface temperature; cooler stars have redder (numerically larger) colors. *See* ASTRONOMICAL PHOTOGRAPHY.

Multicolor photoelectric photometry. Magnitudes can be measured accurately and conveniently with a photomultiplier tube and a filter to limit the bandpass. Aperture photometry, the traditional method by which stars are measured one at a time through a small circular aperture, is now used primarily when the number of filters is large; by measuring a star through n filters, a magnitude and $n - 1$ color indices are obtained. The various color indices of a star are normally well correlated, but they often provide additional information because of the presence of absorption features in the stellar spectrum. An accuracy of 0.01 magnitude or better is routinely achieved in this work. *See* PHOTOMULTIPLIER; RADIOMETRY.

Most photometry today is done with CCD detectors. In this case, an image of the field is recorded through each filter at the telescope, and the photometry is done digitally on the computer. The advantages of CCDs are particularly great for faint stars in crowded fields.

Many multicolor photometric systems have been defined with wide, intermediate, and narrow bandpass filters. One of the first and still the most widely used is the wideband UBV (ultraviolet, blue, visual) system introduced by H. L. Johnson about 1950. The B and V filters are the photoelectric equivalents of m_{pg} and m_v, respectively (the most significant difference being that the B filter excludes the ultraviolet light that contributes to the photographic magnitude), and B − V has replaced $m_{pg} - m_v$ as the general-purpose color index most commonly employed. The addition of the U filter, placed in a spectral region strongly affected by absorption by neutral hydrogen in hot stars, makes it possible to determine the intrinsic properties of stars from UBV photometry alone, even in the presence of interstellar absorption.

Measurements in red and near-infrared (R,I) filters are often added to UBV photometry, especially when the object is to determine stellar temperatures without relying on the U and B filters, which are strongly affected by absorption lines in the stellar spectrum.

Infrared magnitudes. With the improvement of infrared detectors and the development of telescopes

TABLE 3. Filter characteristics for standard wideband multicolor photometric system

Filter	Central wavelength, μm	Band-width μm	Flux density* for Magnitude = 0.00	Thermal sources peaking in filter: Temperature, K†	Thermal sources peaking in filter: Example
U	0.36	0.07	4350	8500	Fairly hot star
B	0.44	0.10	7200	7000	
V	0.55	0.08	3920	5500	Solar-type star
R	0.70	0.21	1760	4300	
I	0.90	0.22	830	3300	Typical red giant
J	1.25	0.4	340	2400	
H	1.6	0.5	120	1900	Very cool star
K	2.2	0.6	39	1400	
L	3.4	0.7	8.1	900	Circumstellar dust
M	5.0	1.2	2.2	600	
N	10	5	0.12	300	Room temperature

*Unit = 10^{-15} W·cm^{-2}·μm^{-1}.

optimized for infrared work, Johnson was able to extend the wideband UBVRI photometric system to much longer wavelengths. Filters were chosen to match the windows between bands of atmospheric absorption. The 11 filters of the standard multicolor system, as originally defined by Johnson, have the central wavelengths and bandwidths given in **Table 3**. (These bandwidths are the full widths at one-half the peak transmission.) Subsequent observers have often used somewhat narrower bands. The magnitude scale for each filter has an arbitrary zero point, and these have been set by requiring a typical star of type A0 (such as Vega) to have the same magnitude in each filter.

Applications of multicolor infrared photometry have included refinement of the stellar temperature scale, especially for cool stars, and the determination of "bolometric magnitudes." Unlike other magnitudes, which refer to particular wavelength regions, the bolometric magnitude is an integrated quantity referring to the total radiative energy output of the star. That is, it represents, on the magnitude scale, the area under the flux curve. The absolute bolometric magnitude (referred to a distance of 10 pc) represents the intrinsic luminosity of the star. *See* INFRARED ASTRONOMY.

Robert F. Wing

Bibliography. E. Budding, *Introduction to Astronomical Photometry*, 2d ed., 2006; A. A. Henden and R. H. Kaitchuck, *Astronomical Photometry*, 1982, reprint 1990; S. B. Howell, *Handbook of CCD Astronomy*, 2d ed., 2006; J. R. Percy (ed.), *The Study of Variable Stars Using Small Telescopes*, 1986; C. Sterken and J. Manfroid, *Astronomical Photometry, a Guide*, 1992.

Magnolia

A genus of trees with large, chiefly white flowers, and simple, entire, usually large alternate leaves. In the winter the twigs may be recognized by their aromatic odor when bruised, by the encircling stipule scars at the nodes, and by the chambered pith (see **illus.**).

The most important species industrially is *Magnolia acuminata*, commonly called cucumber tree, which grows in the Appalachian and Ozark mountains and may reach a height of 80 ft (20 m), rarely 100 ft (30 m). The fruit is red when ripe and resembles a small cucumber in shape.

The wood of the magnolia is similar to that of the tulip tree and is rather soft, but it is of such wide

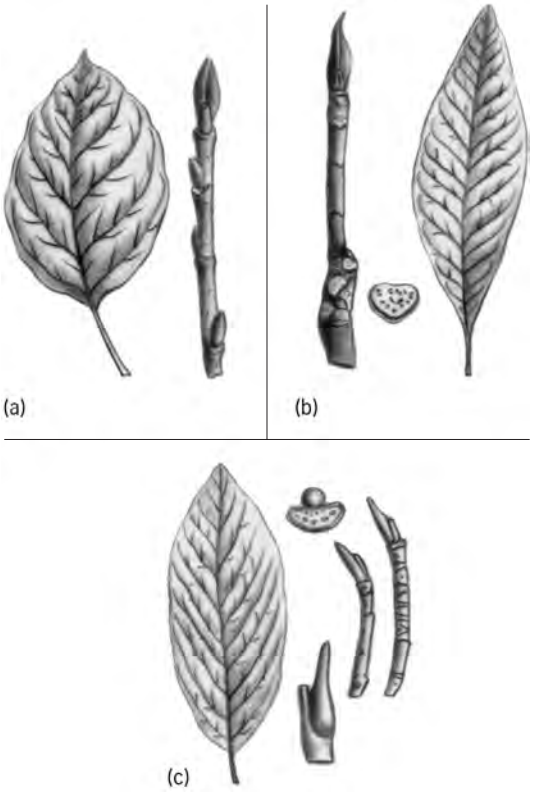

Magnolia leaves and twigs. (*a*) Cucumber tree (*Magnolia acuminata*). (*b*) Umbrella magnolia (*M. tripetala*). (*c*) Sweetbay (*M. virginiana*). In *b* and *c*, note the chambered pith of twig cross sections.

natural dimensions that it is valued for furniture, cabinetwork, flooring, and interior finish. The annual cut amounts to several million board feet. *See* MAGNOLIALES; RANUNCULALES; TULIP TREE.

Magnolia species occur naturally in a broad belt in the eastern United States and Central America, with a similar region in eastern Asia and the Himalayas. With a few exceptions the Asiatic species bloom early, before the leaves appear. These are valued as ornamental trees in the United States. *Magnolia stellata* is shrubby with white or pinkish flowers; *M. yulan* is a small tree with cream-colored flowers; *M. kobus* is similar to *M. yulan*; and *M. soulangeana*, the saucercup magnolia, has pink flowers and is supposedly a hybrid. The species of the eastern United States, all important ornamentals, bloom after the leaves appear and include *M. tripetala*, the umbrella magnolia; *M. virginiana*, the sweetbay; *M. macrophylla*, the big-leaf magnolia; and *M. grandiflora*, the southern magnolia. *See* FOREST AND FORESTRY; TREE. Arthur H. Graves; Kenneth P. Davis

Magnoliales

An order of flowering plants consisting of six families, the best known of which are Magnoliaceae (220 species) and Annonaceae (2200 species). The others contain some botanically interesting and peculiar plants (305 species), but none besides Myristicaceae (300 species) are commonly encountered. Previously, many authors have considered these families as among the most primitive of the flowering plants, but in all cases the plants exhibit some highly derived traits. Although their fruits are often cited as the most primitive in the flowering plants, their carpels are not fused but the edges are sealed, unlike those of Nymphaeaceae and Chloranthaceae, which are not sealed at maturity. Their "primitiveness" has been overemphasized; studies of sequences of deoxyribonucleic acid (DNA) demonstrate that Magnoliales are closely related to Laurales, Piperales, and Winterales, and then more distantly to the monocotyledons.

Some species are of minor economic importance in various parts of the tropics: Annonaceae contain the custard apple, soursop, and sweetsop (*Annona* species), ylang-ylang (an aromatic oil is produced by the flowers of *Canaga odorata*), and *Mkilua fragrans* is the source of a perfume. *Myristica fragrans* (Myristicaeae) is the source of nutmeg and mace, and seeds of several species in this family are used locally to produce consumable oils that are sometimes also used for candle making. Of Magnoliaceae, many species of *Magnolia* (tulip trees) are used as ornamental trees in the temperate and tropical zones, and *Liriodendron* (yellow poplar or green tulip tree) is a commonly planted shade tree throughout the north temperate region. The latter is also a valuable timber-producing genus. *See* LAURALES; PIPERALES; PLANT KINGDOM; TULIP TREE. Mark Chase

Magnoliidae

A subclass of the class Magnoliopsida (dicotyledons) in the division Magnoliophyta (Angiospermae), the flowering plants. The subclass consists of 8 orders, 39 families, and more than 12,000 species. The Magnoliidae are the most primitive subclass of flowering plants, being essentially coextensive with what has often been called the Ranalian complex. In general, they have a well-developed perianth, which may or may not be numerous, centripetal stamens, and they are apocarpous. Their ovules are bitegmic and crassinucellate, and their seeds usually have a small embryo and copious endosperm. All of these characters are subject to exception, however, and there is no one character or easy set of characters by which the group can be formally defined. Various other putatively primitive characters, such as vesselless xylem, laminar stamens, laminar placentation, and the presence of more than two cotyledons, are largely or wholly restricted to this subclass, but are not standard features. *See* ARISTOLOCHIALES; ILLICIALES; LAURALES; MAGNOLIALES; MAGNOLIOPHYTA; MAGNOLIOPSIDA; NYMPHAEALES; PAPAVERALES; PIPERALES; PLANT KINGDOM; RANUNCULALES. Arthur Cronquist; T. M. Barkley

Magnoliophyta

A division of seed plants consisting of about 250,000 species, which form the bulk and most conspicuous element of the land plants. Often called flowering plants or angiosperms, they have several unique characteristics, the most prominent of which are their reproductive structure, flowers, and covered seeds. The other obvious woody land plants are the gymnosperms, which have cones instead of flowers and have naked seeds. Another trait distinguishing the angiosperms is the presence of double fertilization, which results in the production of stored food (starch or oils) within their seeds.

Angiosperms range from some of the smallest plants known to large forest trees, and they occur in all habitats, including the oceans, where they are only a minor element in most marine ecosystems. Some are capable of growing directly on rock surfaces as well as on the limbs of trees (*Tillandsia*, the so-called Spanish moss of the family Bromeliaceae, even grows on telephone wires). *Borya* (of the family Boryaceae from Australia) grows in rock crevices; during the dry season, it dries down to a fraction of its fresh weight and, like many mosses, is a "resurrection plant" that quickly resumes photosynthesis once rain occurs. The flowering plants thus have become masters of the world in many ways. The success of humankind is completely dependent on them, and if human activities alter their success the results will be catastrophic. Mark W. Chase

The angiosperms may be characterized as vascular plants with roots, stems, and leaves, usually with well-developed vessels in the xylem and with

companion cells in the phloem. The central cylinder has leaf gaps or scattered vascular bundles; the ovules are enclosed in an ovary; and the female gametophyte is reduced to a few-nucleate embryo sac without an archegonium. The male gametophyte is reduced to a tiny pollen grain that gives rise to a pollen tube containing a tube nucleus and two sperms; one sperm fuses with the egg in the embryo sac to form a zygote, and the other fuses with two nuclei of the embryo sac to form a triple fusion nucleus that is typically the forerunner of the endosperm of the seed. *See* LEAF; PHLOEM; POLLEN; ROOT (BOTANY); SEED; STEM; XYLEM.

Life cycle. The angiosperms are often called flowering plants because the flower is a characteristic feature not found in other groups of plants. A flower is a special sort of short shoot with modified leaves, some of which produce sexual reproductive structures, leading ultimately to the formation of seeds. A flower typically has four kinds of modified leaves—the sepals, petals, stamens, and carpels—attached to a stem tip called the receptacle. The sepals are on the outside, the petals next, and then the stamens, and the carpels are in the center. One or more of these kinds of parts may be missing, but the order (in normal specimens) is invariable. Typically the sepals are greenish and protective in function, whereas the petals are larger, brightly colored, and serve to attract insects or other pollinators. The sepals collectively constitute the calyx, the petals collectively constitute the corolla, and the calyx and corolla together constitute the perianth. *See* FLOWER.

Sporophyte generation. The stamens, collectively constituting the androecium, are essentially microsporophylls which have been modified to the point that they do not resemble leaves in structure. Typically a stamen has a slender stalk (filament) and a thicker, terminal anther with two pollen sacs (microsporangia), in which the pollen grains (young male gametophytes) are produced.

The carpels, collectively constituting the gynoecium, are essentially megasporophylls. The carpels may be separate from each other, each one folded and joined at the margins to form a simple pistil, or they may all be united to form a compound pistil. A pistil is usually differentiated into three externally visible parts: the ovary, style or styles, and stigma or stigmas. The ovary, at the base, contains the ovules (essentially megasporangia, with some additional outer coats). The stigma, which receives the pollen, is usually elevated above the ovary on a slender style. Each ovule typically contains a single embryo sac, representing the female gametophyte.

Gametophyte generation. The pollen grains are transported to the stigma (often of a different flower) by one means or another, often incidentally by insects seeking nectar. Birds, bats, wind, and even water may also serve as pollinating agents for some species. Some kinds of flowers are self-pollinated, the pollen being deposited directly on the stigma as the flower develops. The pollen grain germinates on the stigma, producing a pollen tube which grows down through the style into the ovary, eventually delivering its two sperms into the embryo sac of an ovule. The ovule develops into the seed, the fertilized egg in the embryo sac develops into the embryo (young sporophyte) of the seed, the triple fusion nucleus (resulting from the fusion of the second sperm with two nuclei of the embryo sac) is the forerunner of the endosperm in the seed, and the integuments (outer coatings) of the ovule develop into the seed coat.

The endosperm typically serves as a food storage tissue for the seed, but in some kinds of seeds the food is absorbed by the embryo during early development so that the ripe seed has no endosperm. Food may also be stored in the nucellus, which is the megasporangium wall that encloses the embryo sac and is in turn largely enclosed by the integuments of the ovule. A small opening through the integuments to the nucellus in the young ovule is called the micropyle. The pollen tube may enter the ovule through the micropyle, or it may digest its way through at some other point. *See* REPRODUCTION (PLANT).

Evolutionary position. Among plants with alternation of sporophyte and gametophyte generations, the angiosperms represent the most extreme stage in reduction of the gametophyte, which in effect is reduced to a mere stage in the reproduction of the sporophyte. The pollen grain, with its associated pollen tube, and the embryo sac represent the male and female gametophyte generations; the endosperm is a new structure not referable to either generation; and the remainder of the plant throughout its life cycle is the sporophyte. Many angiosperms can also propagate asexually by means of creeping stems or roots or by other specialized vegetative structures such as bulbils.

The angiosperms are usually considered to be the most highly evolved division of the subkingdom Embryobionta. Their highly specialized and relatively efficient conducting tissues, combined with the protection of their ovules in an ovary, give them a competitive advantage over most other groups of land plants in most regions. This general advantage is reflected in their abundance and in their great diversification, with different species adapted to a wide range of ecological niches. *See* EMBRYOBIONTA.

It is obvious to biologists that the angiosperms must have evolved from gymnosperms, but beyond this the facts are obscure. They appear in the fossil record early in the Cretaceous Period as obvious angiosperms, without any hint of a connection to any particular group of gymnosperms. Many believe that among the gymnosperms the seed ferns provide the most likely ancestors. *See* PALEOBOTANY; PINOPHYTA. Arthur Cronquist; T. M. Barkley

The Magnoliophyta consist of two large groups that have not been formally named: eudicots (formerly considered dicotyledons with triaperturate pollen, and the magnoliids (including the monocotyledons) with uniaperturate pollen. Although the difference in pollen formation has been known for

a long time, the significance was not appreciated until deoxyribonucleic acid (DNA) studies produced results that clearly reflect this difference. The eudicots are characterized by flowers that are highly organized in terms of the number and orientation of parts whereas the magnoliids have many parts without any particular fixed patterns among the parts—except for the monocots, in which the most developed groups, like the eudicots, exhibit developed flowes with highly organized patterns. *See* LILIOPSIDA; MAGNOLIOPSIDA. Mark W. Chase

Economic importance. The vast majority of conspicuous land plants with green leaves—other than the ferns and the evergreen, needle-leaved trees—are angiosperms. Their significance to humans is more than proportional to their abundance. All the important staples of the human diet are either angiosperms or are derived from animals (except for fish) which feed largely on angiosperms. In addition to food, they provide clothing and shelter and furnish raw materials for countless industrial processes. Even the air has its oxygen supply constantly replenished by angiosperms (and other plants) as a result of their food-making processes. Angiosperms are so essential to human life that they may properly be said to be the foundation of human existence. *See* PHOTOSYNTHESIS; PLANT KINGDOM. Arthur Cronquist; T. M. Barkley

Bibliography. Angiosperm Phylogeny Group (K. Bremer, M. W. Chase, and P. Stevens, eds.), An ordinal classification for the families of flowering plants, *Ann. Missouri Bot. Garden*, 85:531–553, 1998; A. Cronquist, *The Evolution and Classification of Flowering Plants*, 1968; A. Cronquist, *Integrated System of Classification of Flowering Plants*, 1981; A. Cronquist, *Introductory Botany*, 1961; A. Eames, *Morphology of Angiosperms*, 1961; P. Maheshwari, *An Introduction to the Embryology of Angiosperms*, 1950.

Magnoliopsida

One of the two classes of flowering plants which collectively make up the division Magnoliophyta (Angiospermae). The Magnoliopsida, often known as Dicotyledoneae or dicotyledons, embrace 6 subclasses, 64 orders, 318 families, and about 165,000 species.

All of the characters which collectively distinguish the Magnoliopsida from the Liliopsida (monocotyledons) are subject to exception, but in general the Magnoliopsida have two cotyledons and net-veined leaves. The vascular bundles are typically borne in a ring (or cylinder) enclosing a pith. Increase in thickness of stems and roots, after the primary tissues have matured, results from meristematic activity of a cambial layer which passes through the vascular bundles. In about half of the species of the group, the cambium of the stem forms a continuous cylinder which produces a new layer of wood (secondary xylem) and bark (secondary phloem) each growing season for year after year. Such plants become trees or shrubs.

Some other species have an active cambium and are potentially woody, but the stem or the whole plant dies at the end of the year. In still other species the cambium is confined to the vascular bundles and produces relatively little secondary tissue, or there may be no cambium at all. The root system in many species is derived wholly from the primary root; in others it is partly or wholly adventitious.

The floral parts of dicotyledons, when of definite number, are most often borne in sets of five, less often four, and seldom only three (except that there are often only two carpels, or even only one). The pollen is typically triaperturate or of triaperturate-derived type, except in a few of the more primitive families in which it is uniaperturate.

It is widely agreed that the most existing angiosperms belong to the dicotyledons (especially the order Magnoliales) and that most of the characters which distinguish the monocotyledons as a group are derived rather than primitive. *See* ASTERIDAE; CARYOPHYLLIDAE; DILLENIIDAE; HAMAMELIDAE; LILIOPSIDA; MAGNOLIALES; MAGNOLIIDAE; MAGNOLIOPHYTA; PLANT KINGDOM; ROSIDAE.

Arthur Cronquist; T. M. Barkley

Magnon

A quantum of a spin wave; an elementary excitation of a magnetic system which is usually long-range-ordered, such as a ferromagnet. *See* ANTIFERROMAGNETISM; FERRIMAGNETISM; FERROMAGNETISM.

In the lowest energy state of a simple ferromagnet, all the magnetic moments of the individual atoms are parallel (say, to the z axis). Each atomic moment derives mainly from the electron spin angular momentum of the atom. In the next-to-lowest energy state (first excited state), the total z component of spin angular momentum, S_z, is reduced by one unit of $\hbar = h/2\pi$, where h is Planck's constant. In the case of a crystalline material, this unit is shared equally by all the spins, each of which lies on a cone (see **illus.**), precessing at an angular rate ω. These spins form a wave, known as a spin wave, having a repeat distance or wavelength, λ. The wave amplitude (that is, the cone angle) is extremely small, because of the sharing among all the spins whose number N is very large, roughly 10^{23}. Thus, each atom's share of the reduction in S_z, labeled Δ, is only $\hbar/N$, whereas the z component of the atomic spin in the fully aligned state is typically 1–10 times $\hbar$. It follows from simple geometry that the cone half-angle is of order 10^{-11} to 10^{-12} radian. The state with this value of the amplitude is said to be a one-magnon state with wave number $k = 1/\lambda$. If Δ is doubled to $2\hbar/N$, the state is a two-magnon state, and so forth. The integer values of $N\Delta/\hbar$ correspond to the possible changes in S_z being integral multiples of $\hbar$. *See* ELECTRON SPIN; WAVE MOTION.

While the spin waves associated with energy states, that is, stationary states, in crystals vary

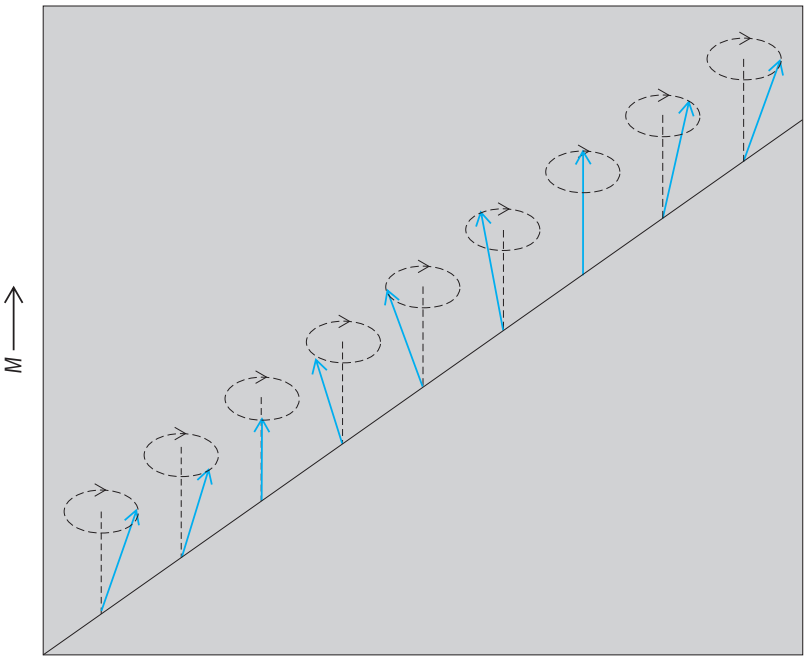

Spin wave in a linear ferromagnetic array of precessing atomic spins of equal magnitude, represented as arrows (vectors) in perspective. The axis of precession is along the vertical direction of the total magnetization, *M*.

sinusoidally in space (see illus.), magnons can be associated, instead, with nonstationary states (wave packets) in some situations. Closely analogous to magnons are phonons and photons, quanta of mass-density waves and electromagnetic waves, respectively. *See* ELECTROMAGNETIC WAVE; PHONON; PHOTON; QUANTUM MECHANICS.

Energy of magnon states in ferromagnets. For a ferromagnet the energy is minimum when the spins are parallel; in the spin-wave state shown (see illus.), the energy is raised by an amount ΔE that depends both on the wavelength and on the amplitude. In the Heisenberg model where the energy (called exchange energy) is proportional to the cosine of the angle between neighboring spins, ΔE for one-magnon states is proportional to k^2 when the wave number (k) is small (that is, for long wavelengths). To determine this energy more generally, it is usually necessary to resort to approximations. A rather broadly applicable one is the semiclassical spin-wave approximation, which assumes that the x and y components of the spins are much smaller than the z components. Then ΔE for an n-magnon state with wave number k is $n\epsilon_k$, where ϵ_k is the energy increase for a single-magnon state. The linear dependence on n follows from the energy being proportional to the cosine of the angle between neighboring spins, $\mathbf{S}_1$ and $\mathbf{S}_2$, or to $\mathbf{S}_1 \cdot \mathbf{S}_2$, with the two spins lying on the surface of the cone, and the cone angle being extremely small. Then (see illus.) it is found that $\mathbf{S}_1 \cdot \mathbf{S}_2 - S^2$ is proportional to the square of the cone angle, which is linear in n. Also, the energy, ϵ_k, of a single-magnon state is equal to $\hbar\omega$, where ω is the precession frequency. When there are magnons with different wave numbers (k) simultaneously present, it is found that they contribute additively to the total increase in energy.

Energy in other magnets. For magnons in antiferromagnets and ferrimagnets, the properties observed in ferromagnets still hold, with two principal exceptions, as follows:

1. The excitation energy ϵ_k can be very different. For example, in antiferromagnets, ϵ_k is proportional to k when k is small, rather than k^2; for ferrimagnets the lowest values of ϵ_k go as k^2.

2. Whereas in the ground state of a Heisenberg ferromagnet each atomic spin has its maximum allowed z component (S), in an antiferromagnet or ferrimagnet this full value cannot be achieved, the expectation value being reduced from S in general. In the latter ground states, the number of magnons present is still zero, but there is a nonzero cone angle φ_0 for each spin. Excited states have nonzero values of the magnon numbers, n_k, as with ferromagnets, which now increase the cone angles beyond the zero-point value φ_0. The ground-state reduction of the z component of each spin is said to be due to zero-point quantum spin fluctuations, in close analogy to zero-point lattice vibrations and vacuum fluctuations of the electromagnetic field (related to phonons and photons, respectively). *See* LATTICE VIBRATIONS; QUANTUM ELECTRODYNAMICS.

These concepts of magnon or spin-wave theory have been applied with remarkable success to many experimentally observed properties.

Thermal properties. The fact that magnons with different wave numbers contribute additively to the total energy increase, plus the fact that the numbers n_k of magnons with wave number k take on the values 0, 1, 2, . . . , imply that the average values of the magnon numbers, denoted $\langle n_k \rangle$, in thermal equilibrium at absolute temperature T, are given by the Bose-Einstein distribution. For this reason, magnons can be considered particles of the type called bosons. From the values of $\langle n_k \rangle$ are found the internal energy and therefore the specific heat. Also, for a ferromagnet, for example, the average spin, $\langle S_z \rangle$, can be found. These results imply particular behaviors with T at low temperature, depending on the form of ϵ_k at small k, and the dimensionality of the structure. For a typical ferromagnet or ferrimagnet with essentially three-dimensional bonding, for example, the cubic materials europium oxide (EuO) and magnetite (Fe_3O_4), the decrease in $\langle S_z \rangle$ from S is proportional to $T^{3/2}$. The magnetic specific heat for these cases is proportional to $T^{3/2}$, whereas for cubic antiferromagnets the magnetic specific heat goes as T^3. *See* BOSE-EINSTEIN STATISTICS; SPECIFIC HEAT OF SOLIDS; STATISTICAL MECHANICS.

Neutron scattering. Elastic or Bragg scattering of neutrons via their magnetic moment helps determine the type of ordering of the atomic moments, while inelastic scattering measures the magnon energy ϵ_k. The particle nature of magnons is reinforced by the fact that the assignment of momentum a $\hbar$ (related to/$\vec{k}$) to a magnon leads to the conservation of energy and momentum of the system of neutron

and magnons in the scattering process. The intensities of the Bragg peaks are related to the strength of the ordered moments; they consequently give information about the zero-point spin fluctuations. These are particularly large in the antiferromagnets La_2CuO_4 and $YBa_2Cu_3O_4$ (the insulating parents of high-temperature superconductors) due to the essentially two-dimensional structure of these layered materials and the fact that for the magnetic ion (Cu^{2+}), $S = {}^1/_2$, the smallest possible nonzero value. *See* NEUTRON DIFFRACTION; SLOW NEUTRON SPECTROSCOPY; SUPERCONDUCTIVITY. Thomas A. Kaplan

Bibliography. C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; D. C. Mattis, *The Theory of Magnetism*, 2d ed., 1987.

Mahogany

A hard, red or yellow-brown wood which takes a high polish and is extensively used for furniture and cabinetwork. The West Indies mahogany tree (*Swietenia mahagoni*), a native of tropical regions in North and South America, is a large evergreen tree with smooth pinnate leaves (see **illus.**). Together with other species it yields the world's most valuable cabinet wood. Varieties with ornamental figures in the grain are highly prized. In the United States it occurs naturally only in the extreme southern tip of Florida, but it is planted elsewhere in the state as an ornamental and shade tree. The cigarbox or West Indian cedar (*Cedrela odorata*) belongs to the same family. *See* FOREST AND FORESTRY; GERANIALES; TREE.

Leaf and fruit of the West Indies mahogany (*Swietenia mahagoni*).

Arthur H. Graves; Kenneth P. Davis

Maillard reaction

A nonenzymatic chemical reaction involving condensation of an amino group and a reducing group, resulting in the formation of intermediates which ultimately polymerize to form brown pigments (melanoidins). It was named for the French biochemist Louis-Camille Maillard, who published initial studies of the reaction between 1912 and 1916. The reaction is of extreme importance to food chemistry, especially because of its ramifications in terms of food quality. *See* AMINE; REACTIVE INTERMEDIATES.

Mechanism. There are three major stages of the reaction. The first comprises glycosylamine formation and rearrangement to *N*-substituted-1-amino-1-deoxy-2-ketose (Amadori compound). The second phase involves loss of the amine to form carbonyl intermediates, which upon dehydration or fission form highly reactive carbonyl compounds through several pathways. Three of these pathways are given in the **illustration**.

Pathway (1) leads to the formation of fission product reductones such as pyruvaldehyde, diacetyl, and hydroxyacetone; pathway (2) forms 2-furfural or its derivatives by dehydration; and pathway (3) is responsible for forming flavor compounds such as maltol, isomaltol, and 4-hydroxy-5-methyl-3(2H)-furanone by dehydration.

Another degrading pathway is the Strecker degradation of amino acids by α-dicarbonyls produced in the above pathways, which by decarboxylation and oxidation forms aldehydes with one less carbon atom than the original amino acid.

The third phase of the Maillard reaction occurring upon subsequent heating involves the interaction of the carbonyl flavor compounds (see illus.) with other constituents to form brown nitrogen-containing pigments (melanoidins) by aldol condensation, carbonyl-amine polymerization, and the production of *N*-, *O*-, and *S*-heterocyclic compounds such as furans, furanones, pyrazines, thiazoles, thiophenes, oxazolines, and polysulfides responsible for heated-food flavors (corny, nutty, bready, and roasted). These are highly desirous compounds in certain foods browned by heating in the presence of oxygen.

Antioxidant properties. Melanoidins or pre-melanoidins resulting from heating various sugar-amine mixtures have strong antioxidant properties in certain lipid mixtures. The use of these high-molecular-weight compounds and their intermediates have tremendous potential as antioxidants in food systems. *See* ANTIOXIDANT.

Undesirable aspects. The Maillard reaction is considered undesirable in some biological and food systems. The interaction of carbonyl and amine compounds might damage the nutritional quality of proteins by reducing the availability of lysine and other essential amino acids and by forming inhibitory or antinutritional compounds. The reaction is also associated with undesirable flavors and colors in some foods, particularly dehydrated foods. Some of the Maillard interaction products are also possibly toxic to animals. Maillard-type carbonyl–amine reactions in living organisms have specific physiological effects that may influence human health.

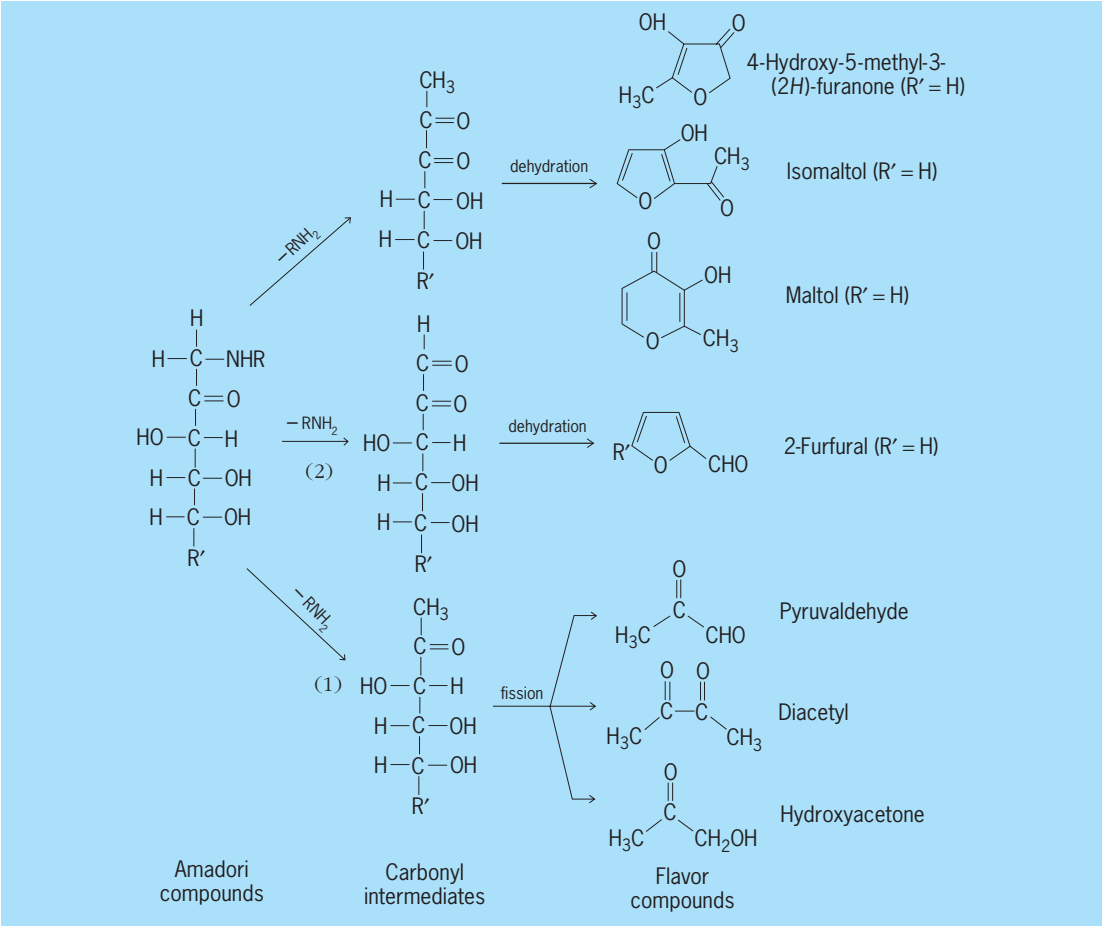

Degradation of Amadori compounds to carbonyl intermediates in Maillard browning reactions.

See AMINO ACIDS; CARBONYL; FOOD MANUFACTURING.

Milton E. Bailey

Bibliography. C. E. Eriksson, Maillard reactions in food, *Prog. Food Nutrit. Sci.*, vol. 5, 1981; P. A. Finot et al. (eds.), *The Maillard Reaction in Food Processing, Human Nutrition, and Physiology*, 1990; T. M. Reynolds, Chemistry of nonenzymic browning II, *Adv. Food Res.*, 14:167–283, 1965; R. Waller and M. S. Feather, *The Maillard Reaction in Foods and Nutrition*, ACS Symp. Ser. 215, 1982.

Maintenance, industrial and production

The actions taken to preserve the operation of devices, particularly of electromechanical equipment, to ensure that the devices can perform their intended functions when needed. The field of maintenance science is an interdisciplinary research area that employs techniques from physics, engineering, and decision analysis to develop methods for the safe and reliable performance of devices and equipment. Traditionally, the focus of maintenance has been on equipment availability—the ratio of operating time less downtime to total available time. Modern maintenance practices focus on increasing equipment effectiveness, that is, making sure that the equipment is both available and capable of producing superior-quality products. *See* SYSTEMS ENGINEERING.

Life-cycle considerations. It has been estimated that up to 50% of all life-cycle equipment costs are attributable to operation and maintenance. As customers recognize that their options are to "pay now" (higher purchase price) or "pay later" (increased maintenance costs), maintenance considerations are increasingly important at the design phase. Equipment buyers are now requiring better information on time to failure and repair. Suppliers have responded by including such things as failure mode and effects analysis, statistical information on failure times, cost-effective maintenance procedures, and better customer training with their products. Additionally, many companies emphasize design enhancements to improve maintainability, such as built-in diagnostics, greater standardization and modularity, and improved component accessibility. While these enhancements may add to the purchase price of equipment, in most cases they more than offset this price in reduced operating and maintenance costs over the equipment's life cycle. *See* ENGINEERING DESIGN; MACHINE DESIGN.

Categories of maintenance activities. Maintenance activities can be classified into several broad

categories, depending on whether they respond to failures that have occurred or whether they attempt to prevent failures.

Reactive maintenance. The simplest and least sophisticated maintenance strategy still used by many companies is reactive maintenance or breakdown maintenance. Equipment is operated until it fails, then repaired or replaced. No effort is expended on activities that monitor the ongoing "health" of the equipment, and maintenance is focused on quick repairs that return the equipment to production as soon as possible. This type of maintenance has two significant drawbacks: (1) repairs are unpredictable and often result in significant unplanned downtime, and (2) equipment often experiences accelerated degradation as a result of interactions at the subsystem or component level within the equipment.

Preventive maintenance. A slightly more sophisticated maintenance strategy is preventive maintenance, also known as calendar-based maintenance. This system involves detailed, planned maintenance activities on a periodic basis, usually monthly, quarterly, semiannually, or annually. As in reactive maintenance, preventive maintenance does not monitor information on equipment status. Rather, it attempts to avoid unplanned failures through planned repairs or replacements. The advantage to this strategy is that one knows in advance when repairs are to be made, but a big drawback is that properly working components may actually be replaced with defective ones. Moreover, operating equipment is taken offline while such repairs are made, and a large spare parts inventory may be required.

Predictive maintenance. Also known as condition-based maintenance, predictive maintenance is based on an ongoing (continuous or periodic) assessment of the actual operating condition of equipment. The equipment is monitored while in operation, and repair or replacement is scheduled only when measurements indicate that it is required. Predictive maintenance programs seek to control maintenance activities to avoid both unplanned equipment outages and unnecessary maintenance and overhauls. An important aspect of predictive maintenance strategies is the use of sophisticated sensoring technologies in combination with data-driven models to characterize equipment degradation. Among the new sensoring technologies for monitoring critical parameters are acoustic sensors for vibration analysis, microelectronic devices for gas analysis, sensors for oil wear analysis, infrared thermography to locate aberrantly high temperatures indicative of excessive wear, ultrasound for detection of incipient cracks, fiber optics for strain gauge measurements, high-speed video, and laser alignment. *See* FIBER-OPTIC SENSOR; MICROSENSOR.

Industry-driven maintenance initiatives. The importance of equipment maintenance for product quality and plant safety has long been recognized. Recently, industry has begun to understand the impact of maintenance practice on the broader elements of manufacturing productivity, such as factory output and product cycle time, as well as on product yield. In equipment-intensive industries, such as the semiconductor industry and the electric utility industry, even small improvements in equipment use can result in an immediate competitive advantage.

Reliability-centered maintenance. Developed in the late 1960s by the aerospace industry, and since then widely adopted in the passenger airline industry and in the nuclear power industry, reliability-centered maintenance (RCM) involves the use of reliability methods, such as failure mode, effects and criticality analysis (FMECA), to identify and prioritize component failures in a systematic and structured way. Key equipment components are then selected and monitored, and appropriate maintenance actions are determined for all components based on cost, safety, environmental, and operational requirements. The rationale behind reliability-centered maintenance is that most vendor-supplied recommendations for fixed, time-in-service maintenance intervals are overly conservative, and are applied to individual components without consideration of their overall system functions. The focus of reliability-centered maintenance is on failure "triage," that is, on identifying and avoiding machine conditions that are likely to be the most serious and lead to expensive equipment failures. For critical equipment, reliability-centered maintenance replaces fixed maintenance intervals with intervals that depend on use and condition, as determined by analysis of past performance and machine diagnostics. For noncritical equipment, cost analysis is performed to determine an appropriate balance between reactive and preventive maintenance. Maintenance tasks are not performed unless they affect system availability, safety, or overall cost. Companies that have introduced reliability-centered maintenance report significant savings in maintenance costs and improved equipment effectiveness. Moreover, the data collected as part of the reliability-centered maintenance program are valuable in guiding ongoing maintenance improvements.

Total productive maintenance. Total productive maintenance (TPM), developed in Japan in the 1980s, has been adopted by many industries, notably by semiconductor manufacturers. The development of total productive maintenance stems from the recognition that production-level cost, inventory, safety, and product quality depend directly on equipment performance. The goals of total productive maintenance—zero unplanned downtime, zero equipment-caused product defects, and zero loss of equipment speed—are approached via a continuous improvement strategy that involves all workers. Managers, engineers, production operators, and maintenance personnel are organized into small autonomous groups and charged with maintenance responsibility for a machine or set of machines. Machine operators are assigned certain routine maintenance tasks, such as cleaning, lubrication, calibration, and bolt tightening, while maintenance personnel have responsibility for larger, less frequent tasks, such as overhauls and component replacements. Their activities are directed toward five

main categories of equipment losses that reduce equipment effectiveness: set-ups and adjustments, equipment failures, idling and minor stoppages, reductions in equipment speed, and process defects. Total productive maintenance prescribes a data collection strategy that ensures a consistent set of metrics and allows groups to prioritize improvements.

Integrated production operations. As the impact of equipment maintenance on plant productivity has become better understood, recent interest has focused on integrating production and maintenance decisions. While the initiatives described above have been successful in reducing processing time variability due to unplanned equipment downtime, maintenance activities are still largely scheduled independently of the production status (the amount and location of work) of the plant. As a result, planned maintenance activities have the potential for significant disruptions in product flow. Many companies now employ strategies to execute maintenance activities "opportunistically," that is, when production status allows, subject to achieving certain performance objectives. The performance objective may involve meeting targeted equipment availability or minimizing total expected production and maintenance costs. The long-term goal of this approach is to develop controllers that schedule appropriate production and maintenance actions based on the combined input of a shop-floor control system and condition-monitoring sensors. Georgia-Ann Klutke

Bibliography. B. S. Blanchard, D. Verma, and E. L. Peterson, *Maintainability: A Key to Effective Serviceability and Maintenance Management*, Wiley, New York, 1995; J. Douglas, The maintenance revolution, *EPRI J.*, January–February 1995; R. K. Mobley, *An Introduction to Predictive Maintenance*, Van Nostrand Reinhold, New York, 1990; S. Nakajima, *Introduction to TPM*, Productivity Press, Cambridge, MA, 1988; A. M. Smith, *Reliability-Centered Maintenance*, McGraw-Hill, New York, 1993.

Malachite

A bright-green, basic carbonate of copper [$Cu_2CO_3(OH)_2$]. Malachite is the most stable copper mineral in natural environments in contact with the atmosphere and hydrosphere. It occurs as an ore mineral in oxidized copper sulfide deposits; as a stain on fractures in rock outcrops; as a corrosion product of copper and its alloys (except in industrial-urban environments, where the basic copper sulfate dominates); as suspended particles in streams and in alluvial sediments; and as encrustations on bronze artifacts in seawater and on coccoliths floating in the oceans. It can be distinguished from other green copper minerals by its effervescence in acid. The combination of hardness (3.5–4 on Mohs scale) ideal for carving, color variation in concentric layers, and adamantine-to-silky luster has made malachite a highly prized ornamental stone. Its rare blocky-tabular crystals up to 5 mm (0.2 in.), its pseudomorphs after azurite crystals to 2 cm (0.8 in.), and its more common felty tufts perched on bright blue azurite are eagerly sought by mineral collectors.

Malachite is an important copper ore mineral in supergene copper oxide deposits formed by weathering of primary copper sulfide deposits. The high acidity necessary for dissolution of sulfide minerals and solution transport of copper (pH = 2–3) occurs with the oxidation of pyrite followed by generation of sulfuric acid in the unsaturated zone above the water table. Malachite, stable in oxidized ground waters at pH > 5, readily precipitates in large amounts in places where copper-laden, acidic waters from a nearby sulfide deposit encounter a neutralizer that also can supply carbonate ion, such as limestone wall rocks or calcite gangue. If cupric solutions are neutralized in the absence of carbonate, chrysocolla ($CuSiO_3 \cdot 2H_2O$) or tenorite (CuO) may form; in arid climates, the soluble basic copper sulfates are stable (chalcanthite, antlerite, brochantite). Azurite may form rather than malachite if the concentration of carbon dioxide (CO_2) gas is high (as in soils producing CO_2 by oxidation of organic matter) or if the relative humidity is low (as in arid climates), as shown by the reaction below.

$$\underset{\text{Malachite}}{3Cu_2CO_3(OH)_2} + \underset{\text{Carbon dioxide}}{CO_2} \longrightarrow \underset{\text{Azurite}}{2Cu_3(CO_3)_2(OH)_2} + \underset{\text{Water}}{H_2O}$$

Copper oxide deposits with malachite as an important ore mineral include N'Changa and Chambishi in the Zambian copperbelt, Tsumeb in Namibia, Mount Isa in Australia, Bisbee in Arizona, and oxidized portions of copper-bearing skarn and vein deposits in limestone throughout the world. *See* AZURITE; CARBONATE MINERALS; COPPER. Marco T. Einaudi

Bibliography. L. L. Y. Chang, R. A. Howie, and J. Zussman, *Rock Forming Minerals,* vol. 5B: *Nonsilicates: Sulfates, Carbonates, Phosphates, Halides*, 1995; C. Palache, H. Berman, and C. Frondel, *The System of Mineralogy of James Dwight Dana and Edward Salisbury Dana*, 7th ed., vol. 2, 1951.

Malacostraca

The largest and most diversified class of the Crustacea; includes the shrimps, lobsters, crabs, sow bugs, beach hoppers, and their allies. The shell or carapace may be large, small, vestigial, or absent; the tail or abdomen is long or short; the eyes are generally set on movable stalks but may be sessile or even coalesced. Despite this diversity, the unity of the group is demonstrated by the following characteristics which all share. The maximum number of appendages is 19 pairs, even where there is an additional body segment such as is found in the Phyllocarida and embryonic Mysidacea. The trunk limbs are sharply differentiated into a thoracic series

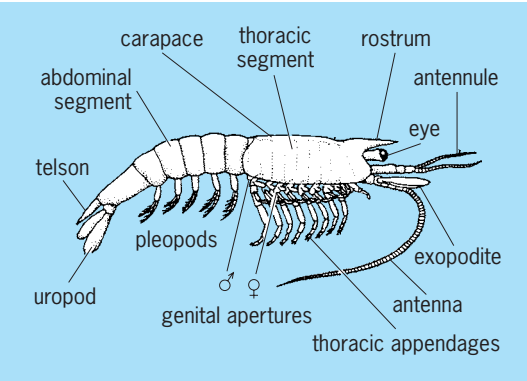

Fig. 1. Caridoid facies. (*After E. R. Lankester, ed., A Treatise on Zoology, pt. 7, fasc. 3, A. and C. Black, 1909*)

of eight pairs and an abdominal series of six pairs. The female genital duct always opens at the level of the sixth thoracic segment, whereas those of the male open at the level of the eighth. Additionally, the frontal organs of malacostracans differ significantly from nonmalacostracans. Even greater, more fundamental differences between malacostracans and nonmalacostracans have been found in their compound eyes.

Classification. Malacostraca are divided into three subclasses, the Phyllocarida, Hoplocarida, and Eumalacostraca. Central to the classification of the Eumalacostraca has been the concept of the "caridoid facies" (**Fig. 1**). This term refers to a series of morphological attributes generally common to the three orders, Syncarida, Peracarida, and Eucarida. Those attributes are a well-developed carapace covering the cephalothorax; stalked and movable eyes; biramous antennules; the exopod of the antenna developed as a blade or scale; natatory exopods on the thoracic appendages; a long, ventrally flexing abdomen; and a tail fan formed by the uropods and telson. Other shared characteristics include a complex set of intertwining abdominal muscles that provide the flexing mechanism seen in escape reactions, internal organs located primarily in the thorax, and a similarity in pleopod structure and function. *See* EUCARIDA; PERACARIDA; SYNCARIDA.

Although they share some of the more primitive aspects of the caridoid facies, the Phyllocarida are distinguished by their retaining the primitive condition of a seventh abdominal somite, the presence of caudal rami on the telson, and a bivalve shell-like carapace held together by an adductor muscle. Similarly, aspects of Hoplocarida morphology have led some researchers to suggest the Hoplocarida represents another lineage within the Malacostraca that has evolved independently of the Eumalacostraca.

Evolution. The first unequivocal malacostracan fossil is that of a phyllocarid. Comparative functional morphology of extant malacostracans, however, suggests that the phyllocarids, and the hoplocarids as well, were early offshoots from the basically eumalacostracan stem. Although conflicting theories regarding a eumalacostracan precursor are numerous, it is probable that possession of a carapace represents the primitive condition and its loss in such groups as the Syncarida and some Peracarida is a derived condition. Other caridoid features, such as the antennal scale, tail fan, and abdominal musculature, appear to have played important roles in eumalacostracan evolution, but radiation also seems to have been dependent upon a change in feeding strategies from primitive thoracopodan phyllopod filter feeding to maxillary feeding and concurrent development of the stenopodial limb for locomotion. The more anterior (thoracic) and more posterior (abdominal) appendages have become specialized, whereas the central pairs, which are also thoracic, have developed as locomotory limbs. Walking legs, for example, have evolved through the reduction or loss of the exopod.

Terminology. The terminology applied to the thoracic appendages and their segments in the various orders differs widely. From one to three of the anterior pairs may be called maxillipeds and the other seven to five pairs pereopods. The second and third (rarely also the fourth) pairs in some orders are known as gnathopods. In Stomatopoda the first five pairs are often termed maxillipeds, although gnathopods would be more appropriate. The abdominal appendages are called pleopods, unless they are modified for reproductive purposes. In the latter, they are referred to as gonopods.

The basis malacostracan appendage is made up of a basal protopod and two terminal rami. The protopod is usually divided into a coxa and a basis, one or both of which may have mesial (endites) or

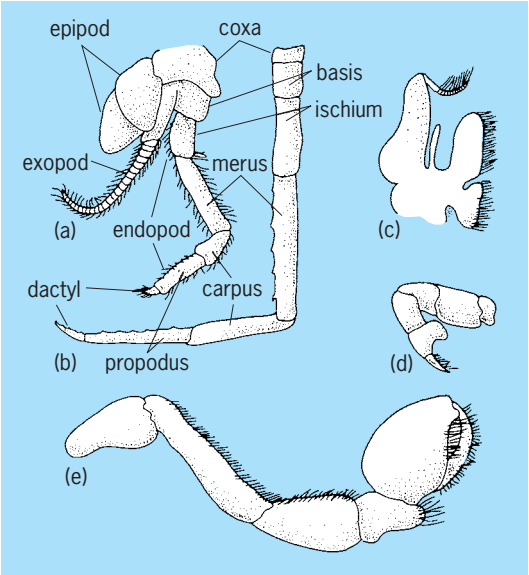

Fig. 2. Representative malacostracan thoracic appendages: (*a*) biramous ambulatory thoracic appendage of an Anaspidacea; (*b*) uniramous thoracic walking leg of a Stenopodidea; (*c*) first maxilliped of a Paguroidea; (*d*) third maxilliped of a Thalassinoidea; (*e*) raptorial appendage ("maxilliped") of a Stomatopoda. (*After P. A. McLaughlin, Comparative morphology of crustacean appendages, in L. G. Abele, ed., The Biology of Crustacea, vol. 2: Embryology, Morphology, and Genetics, Academic Press, 1982*)

lateral (epipods) protrusions. An exception exists in the Stomatopoda, where a precoxal segment may also present. Endites function primarily in food transport or in mastication; epipods, on the other hand, most often are specialized as respiratory structures, or gills. The exterior ramus (exopod), if not reduced or absent, consists of two segments, with the distal segment often multiarticulate or annulated. Exopods may be modified for feeding or for facilitating movement in water, but more often they function as swimming appendages. The inner ramus (endopod) is usually the best developed and consists of five segments: ischium, merus, carpus, propodus, and dactyl (**Fig. 2**). The endopods may be unspecialized locomotive limbs or may be transformed into highly specialized appendages used in sensory detection, feeding, grasping, walking, or burrowing. *See* CRUSTACEA; EUMALACOSTRACA; HOPLOCARIDA; PHYLLOCARIDA. Patsy A. McLaughlin

Bibliography. L. G. Abele (ed.), *The Biology of Crustacea*, vol. 2: *Embryology, Morphology, and Genetics*, 1982; F. R. Schram (ed.), *Crustacean Phylogeny, Crustacean Issues*, vol. 1, 1983.

Malaria

A disease caused by members of the protozoan genus *Plasmodium*, a widespread group of sporozoans that parasitize the human liver and red blood cells. Four species can infect humans: *P. vivax*, causing vivax or benign tertian malaria; *P. ovale,* a very similar form found chiefly in central Africa that causes ovale malaria; *P. malariae,* which causes malariae or quartan malaria; and *P. falciparum,* the highly pathogenic causative organism of falciparum or malignant tertian malaria. Malaria is characterized by periodic chills, fever, and sweats, often leading to severe anemia, an enlarged spleen, and other complications that may result in loss of life, especially among infants whose deaths are almost always attributed to falciparum malaria. The infective agents are inoculated into the human bloodstream by the bite of an infected female *Anopheles* mosquito, more than 60 species of which can carry the infection to humans. The disease is found in all tropical and some temperate regions, but it has been eradicated in North America, Europe, and Russia. Despite control efforts, malaria has probably been the greatest single killer disease throughout human history and continues to be a major infectious disease. An estimated 1 million children die of malaria in Africa each year, and estimates of the world's malaria death toll vary from 1.5 to 2 million cases annually. *See* EPIDEMIC.

Life cycle. The vast reproductive capacity of *Plasmodium* parasites is illustrated by their life cycle, which begins as a series of asexual divisions in human liver and then red blood cells. Transfer of the parasites to the mosquito host depends on the rate of sexual multiplication that begins in the infected human red blood cells and is completed in the mosquito stomach, followed by asexual multiple division of the product of sexual fusion. During the developmental stage, enormous numbers of the sporozoites are produced in cysts massed on the stomach wall of the mosquito. These elongated microscopic forms break out of their cysts and concentrate in the mosquito salivary glands, from which they are pumped into another human when the mosquito next takes a blood meal. In a matter of minutes, the sporozoites enter human liver cells, where they divide many times in each infected cell. However, this does not initiate any host recognition or disease response. The progeny, called merozoites, burst out of the infected liver cells and may invade more liver cells for another generation or two, but most leave the liver and invade red cells in the bloodstream. Only then does the body respond with fever and other symptoms associated with the disease.

Clinical malaria usually begins 7–18 days after infection with sporozoites. Progression to that phase is most rapid with *P. falciparum* and slowest with *P. malariae*. In *P. vivax* and *P. ovale,* some of the invading sporozoites from the infective mosquitoes enter liver cells but do not develop or divide. These "sleeping" parasites (hypnozoites) can remain unchanged for months or years. They eventually grow, divide, and infect the blood cells to initiate bouts of clinical malaria up to 3 years after the initial infection (relapse malaria). Red cell infections tend to follow a remarkably synchronous division cycle. The parasite progresses from merozoite to a vegetative phase (trophozoite) to a division stage (schizont), ending with the new generation of merozoites ready to break out in a burst of parasite releases and initiate the chills–fever–sweat phase of the disease. This symptomatic period takes place in a 48-h pattern, at the start of the third day (hence tertian), with *P. vivax* and *P. ovale.* With *P. falciparum,* an extended fever phase occurs during the 48-h cycle, with the fever lasting 1–8 h rather than 30 min to 1 h, as in the others. *Plasmodium malariae* has a 72-h, or quartan (at the start of fourth day), periodicity. Synchronous schizogony continues for 3 to 6 weeks, up to about 3 months, depending upon the host's acquired or genetic resistance. *See* HAEMOSPORINA; SPOROZOA.

Clinical malaria. Clinical malaria begins with the erythrocytic cycle, and the sequence of chills, fever, and sweats is the result of simultaneous red cell destruction at 48- or 72-h intervals. Bouts of chills and fever often begin irregularly but usually revert to the synchronous pattern. *Plasmodium falciparum* is the most uncertain and variable of the four species, and it is by far the most dangerous. The period from infective mosquito bite to onset of symptoms (incubation period) varies from about 12 days for falciparum malaria (range 6–25 days), to 14 days (range 8–27 days) for vivax and ovale malaria, to 4 weeks (range 16 days–8 weeks) for quartan malaria. The incubation period may be extended to several months with resistant or partially resistant falciparum malaria when prophylactic drugs have been used. Early symptoms, usually 1 or 2 days before the start of an acute attack, include headache, low fever

with chills, and muscle aches and pains. The full attack generally begins with shaking chills, followed by intermittent fever bouts up to 105.8°F (41°C), followed by the 1–2-h sweating stage. The entire attack usually lasts 4–8 h, the longer period usually due to *P. falciparum*. Associated signs and symptoms may include severe headache, abdominal pain with nausea and sometimes forceful vomiting, joint pain, extreme fatigue, and various signs and symptoms that suggest a number of other conditions. Hence, malaria is frequently called the great mimic. High fever with chills and headache are the most consistent symptoms. After about a week, a soft, enlarged, and sometimes tender spleen can usually be detected. The liver may also become enlarged and tender. Between malarial bouts, the patient is fatigued but usually asymptomatic.

A thick blood smear is preferred for diagnosing malaria, but a thin blood smear is essential for reaching a definitive or species diagnosis, particularly for identifying a life-threatening case of falciparum malaria. Frequently, many smears must be made at various periods of the erythrocytic cycle to confirm the diagnosis.

Pathogenesis. The destruction of red cells results in anemia, which is the most characteristic cause of malarial pathology. An enlarged spleen develops because the spleen has the burden of filtering out the many red cells that are burst or damaged because of the disease, as well as removing the characteristic pigment clusters that are present in all red cells infected with *Plasmodium*. Fatalities from falciparum malaria can be directly traced to any or all of three factors: (1) destruction of stem cells in the bone marrow; (2) infection of large numbers of red cells (50,000–2,000,000 or more infected cells per cubic millimeter of blood, rather than the 10,000–35,000 that is indicative of the other three species); and (3) induction of "knobby" or sticky infected red cells, which then adhere to endothelial linings of blood vessels. Should the stickiness lead to vessel blockage in the brain, cerebral malaria, the primary cause of death, can rapidly result from brain anoxia. Similar blockage of essential circulatory vessels in the intestine, lungs, liver, and kidneys can also cause rapid death from falciparum malaria. Death may follow this severe form of attack in 6–12 h. Hence, falciparum malaria should be considered a medical emergency.

Chemotherapy. Chloroquine remains the drug of choice for prevention as well as treatment of vivax, ovale, and malariae malaria. However, most strains of falciparum malaria have become strongly chloroquine-resistant, with the sole remaining chloroquine-sensitive areas being the Caribbean region and the Arabian peninsula.

For prevention of chloroquine-resistant falciparum malaria (and in some areas vivax malaria is now chloroquine-resistant as well) a weekly dose of mefloquine beginning a week before, then during, and for 4 weeks after leaving the endemic area is recommended. Resistance to mefloquine is a major concern in Southeast Asia and other countries. The tetracycline doxycycline can be used in mefloquine-resistant areas, except by pregnant women and children under 8. Possible relapse from latent liver infections with vivax or ovale malaria in travelers can be prevented by administration of primaquine upon return.

Chloroquine-resistant malaria is chiefly treated with the oldest known malaricide, quinine, in the form of quinine sulfate, plus pyrimethamine-sulfadoxine. If sulfa-based compounds cannot be used (in cases such as pregnancy), tetracycline and clindamycin (though even here pregnancy may rule out their use) may be used with quinine. Mefloquine, halofantrine, a phenanthrine-methanol compound, or a herbal drug used for many years in China, qinghaosu (the active agent is artemesin, from the plant *Artemisia*), are effective against chloroquine-resistant strains of *Plasmodium*. *See* DRUG RESISTANCE; QUININE.

Eradication efforts. Malaria is perhaps the most urgent problem in tropical medicine. Failure of earlier efforts to eradicate malaria and the rapid spread of resistant strains of both parasites and their mosquito vectors necessitated renewed interest in prevention of exposure by avoidance of mosquito bites using pyrethrin-treated bednets, coverage of exposed skin during active mosquito periods (usually dawn, dusk, and evening hours), and use of insect-repellent lotions. A balance between epidemiological and immunological approaches to prevention, and the continued development of new drugs for prophylaxis and treatment are recognized as the most effective means to combat one of the most dangerous and widespread threats to humankind from an infectious agent. *See* MEDICAL PARASITOLOGY.

Donald Heyneman

Bibliography. P. Gardner (ed.), *Health Issues of International Travelers*, 1992; H. M. Gilles, *Management of Severe and Complicated Malaria: A Practical Handbook*, 1991; F. J. López-Antuñano and G. Schmunis (eds.), *Diagnosis of Malaria*, 1990; S. C. Oaks, Jr., et al. (eds.), *Malaria: Obstacles and Opportunities*, 1991; R. S. Phillips, Vaccination against malaria, *Immunobiology*, 184:240–262, 1992; F. O. Ter Kuile et al., Halofantrine versus mefloquine in treatment of multidrug-resistant falciparum malaria, *Lancet*, 341:1044–1049, 1993; M. V. Valero et al., Vaccination with SPf66, a chemically synthesized vaccine, against *Plasmodium falciparum* malaria in Colombia, *Lancet*, 341(8847):705–710, 1993; D. J. Wyler, Malaria chemoprophylaxis for the traveler, *New Engl. J. Med.*, 329:31–37, 1993.

Malnutrition

Impaired health caused by a dietary deficiency, excess, or imbalance. To support human life, energy (from fat, carbohydrate, and protein), water, and more than 40 different food substances must be obtained from the diet in appropriate amounts. Malnutrition can result from the chronic intake of any

of these substances at levels above, as well as below, ranges that are adequate and safe, but commonly, the term refers only to deficient intake.

Primary (diet-induced) malnutrition. The number of people throughout the world who suffer from nutritional deficiencies as a result of inadequate dietary intake is uncertain, but even the most conservative estimates place that figure at hundreds of millions; many experts consider the actual number to approach 1 billion. Most malnourished people live in developing countries where income, education, and housing are inadequate to buy, transport, store, and prepare food and where nutritional deficiencies are almost always related to poverty. All but the very poorest nations produce enough food to meet the energy requirements of their populations, but the food that is produced is either inadequately distributed or too expensive.

In industrialized countries, chronic conditions of deficient dietary intake occur far less frequently but are reported occasionally among people who are dieting to lose weight, fasting, or on an unusually restrictive ("fad") diet. The widening gap in income between rich and poor has focused attention on inadequate food intake among low-income populations. Hunger—when defined as inadequate access to food assistance—has been reported to affect 20 million children and adults in the United States, and many surveys have identified substandard levels of dietary intake (below the levels generally recommended for prevention of nutritional deficiencies) among participants in assistance programs. *See* OBESITY.

Pregnant women, infants, and children are most at risk for inadequate dietary intake because their nutritional requirements are relatively high. Surveys also have identified dietary deficiencies among older adults who do not consume enough food to meet nutritional requirements. Their inadequate food intake often is associated with unemployment, reduced access to health care, low levels of education, and other factors that adversely affect health.

Secondary (illness-induced) malnutrition. Nutritional deficiencies also occur as a result of illness, injury, or alcohol or drug abuse that interferes with appetite; the inability to eat; defective digestion, absorption, or metabolism of food molecules; or disease states that increase nutrient losses. Secondary malnutrition has been observed frequently among medical and surgical patients who are treated in hospitals for prolonged periods of time. Hospital malnutrition can be prevented or corrected through appropriate nutritional support, which is a major reason for efforts to improve the nutrition training of physicians and other health professionals. Regardless of cause, the effects of malnutrition can range from minor symptoms to severe syndromes of starvation, protein-calorie malnutrition, or single-nutrient deficiencies. *See* METABOLIC DISORDERS.

Starvation. The chronic intake of energy below the level of expenditure induces rapid losses in body weight and muscle mass accompanied by profound changes in physiology and behavior. Together, these effects cause a starving person to become weak, apathetic, depressed, and unable to work productively and to do whatever is necessary to reverse the malnutrition. The consequences of nutritional deficiencies are seen first in tissues that are growing rapidly. These changes are most evident in the gastrointestinal tract, skin, blood cells, and nervous system as indigestion, malabsorption, skin lesions, anemia, or neurologic and behavior changes.

Of special concern is the loss of immune function that accompanies severe malnutrition. Malnourished individuals lose cellular immune competence and are less resistant to infectious disease. Because infections interfere with the body's ability to digest, absorb, and use food, they increase nutritional requirements. This cyclical interaction of malnutrition and infection produces the greatest damage among people in developing countries, where water supplies are contaminated with disease-causing microorganisms and where that interaction is the principal cause of death in young children and a chief cause of illness in adults.

Protein-calorie malnutrition. The combined effects of malnutrition and infection in young children are referred to as protein-calorie malnutrition. Because breast milk contains immune substances that fight infection, the children most likely to become malnourished are those who are not breast-fed. Not only do infant formulas lack the immune factors, but they also must be diluted in clean water and stored under refrigeration—conditions often unavailable in developing countries. For these reasons, the World Health Organization has developed guidelines to regulate the promotion of infant formulas in these countries by their manufacturers. Children who have just been weaned are also at high risk, because traditional weaning foods may lack essential nutrients and sufficient protein or calories.

Protein-calorie malnutrition usually is classified into two entities, marasmus and kwashiorkor, on the basis of physical appearance and the relative proportions of protein and calories in the diet. Children with the marasmus form appear generally wasted as a result of diets that are chronically deficient in calories as well as protein and other nutrients. Children with kwashiorkor are also very thin but have characteristically bloated bellies due to fluid retention and accumulation of fat in the liver, symptoms attributed to diets relatively deficient in protein. These distinctions are not precise, however; undernourished children raised on similar diets may exhibit signs characteristic of either form of the disease. In any case, children who survive protein-calorie malnutrition display the typical effects of starvation: depression, apathy, irritability, and growth retardation. *See* ADIPOSE TISSUE; PROTEIN METABOLISM.

Deficiencies of single nutrients. Deficiency conditions due to lack of a single vitamin or mineral occur rarely and usually reflect the lack of the most limiting nutrient in a generally deficient diet. In industrialized countries, single-nutrient deficiencies are most evident in individuals who abuse alcohol or drugs.

Classic conditions of deficiency of niacin (pellagra), thiamine (beriberi), vitamin C (scurvy), and vitamin D (rickets) have virtually disappeared as a result of food fortification programs and the development of food distribution systems that provide fresh fruits and vegetables throughout the year. Iron-deficiency anemia also has declined in prevalence, although children in low-income families remain at risk. *See* ANEMIA.

In developing countries, however, such conditions are still observed among people whose diets depend on one staple food as the major source of calories.

Beriberi, for example, was common among Japanese sailors whose diets were based mainly on white rice that had been milled to remove the outer thiamine-containing bran layer. The sailors developed disorders of the heart muscle and of neurologic function. The disease declined in prevalence when the diet became more varied and when it was discovered that par-boiled rice retains thiamine. However, it still occurs in parts of the world where milled rice is the principal source of calories.

Pellagra is caused by a deficiency of the vitamin niacin. Because some niacin can be synthesized from the amino acid tryptophan, pellagra is also considered to be a disease of protein deficiency. Pellagra was common through the 1930s in parts of the southern United States where the diet consisted mainly of corn. Corn is low in tryptophan, and its niacin occurs in an unavailable form. Increasing the variety of foods in the diet and fortifying corn meal and flour with niacin have effectively eliminated pellagra from industrialized countries.

Scurvy was observed often among sailors who were unable to obtain vitamin C from fresh fruits and vegetables during long sea voyages. Providing limes on board ship prevented this condition. Scurvy is reported occasionally in the United States among alcoholics and among infants born to mothers who consume unusually restricted diets. *See* ALCOHOLISM.

The characteristic bowed legs of children with rickets are rarely seen in countries where milk is fortified with vitamin D. This vitamin is synthesized in the body by the action of sunlight on skin, and rickets continues to occur among children who live in parts of the world where sunlight is limited and among those confined indoors.

A condition of substantial current public health importance is vitamin A deficiency, which is the principal cause of blindness and a major contributor to illness and death among children in developing countries. The carotene precursors of vitamin A are readily available in the local fruits and vegetables, but custom dictates that those foods are inappropriate for young children. Education and various forms of fortification and supplementation can serve as preventive measures, but because children with vitamin A deficiency are generally undernourished, the most effective form of prevention is improvement of the overall diet. *See* VITAMIN.

Marion Nestle

Bibliography. J. L. Brown, Hunger in the U.S., *Sci. Amer.*, 256(2):37–41, 1987; L. C. Chen and N. S. Scrimshaw (eds.), *Diarrhea and Malnutrition: Interactions, Mechanisms, and Interventions*, 1983; W. A. Coward and P. G. Lunn, The biochemistry and physiology of kwashiorkor and marasmus, *Brit. Med. Bull.*, 37(1):19–24, 1981; P. R. Kerndt et al., Fasting: The history, pathophysiology, and complications, *West. J. Med.*, 137:379–399, 1982; F. M. Lappe and J. Collins, *World Hunger: Twelve Myths*, 2d ed., 1998; R. Roubenoff et al., Malnutrition among hospitalized patients: A problem of physician awareness. *Arch. Int. Med.*, 147:1462–1465, 1987; N. S. Scrimshaw, The politics of starvation, *Technol. Rev.*, 87:19–27,50, 1984; R. R. Watson (ed.), *Nutrition, Disease Resistance, and Immune Function*, 1984; World Bank staff, *Poverty and Hunger: Issues and Options for Food Security in Developing Countries*, 1986.

Malpighiales

One of the largest orders of the rosid eudicotyledons, comprising more than 30 families. Distributed worldwide, this order has never before encompassed this composition. Recent analyses of deoxyribonucleic acid (DNA) sequences, both plastid and nuclear, led to its recognition, even though the group is highly heterogeneous and difficult to characterize. The largest families are Euphorbiaceae (8000 species), Clusiaceae (1400), Malpighiaceae (1100), Flacourtiaceae (900), and Violaceae (850). Most of the order is composed of woody species, many of regional importance as timber and medicines. Several of the smaller families are significant as well, including Salicaceae (used as coppice, and the original source of aspirin) and Rhizophoraceae (the ecologically significant mangroves). *See* MAGNOLIOPHYTA.

Mark W. Chase

Malt beverage

A fermented beverage produced from grain. Beer is a generic term used to describe alcoholic beverages made from cereal grains, especially barley, in the form of malt. Ale, lager, porter, and stout are different kinds of beer made by recognizably similar processes. The United States is the largest producer of beer in the world.

The manufacture of beer is a complex natural process of three general parts: the preparation of barley by germination, or the malting process; the actual digestion of barley (now malt) starch to produce a solution of sugars (called wort) and the adjustment of flavor with hops, which are the brewhouse processes (see **illus.**); and the fermentation of these sugars by yeast to yield alcohol, carbon dioxide gas, and flavor compounds to produce beer. *See* HOP.

Malting process. Barleys for malting in the United States are special varieties grown in the Middle West (six-rowed types) or the West (two-rowed types). Barley is a hard, flavorless grain with an incomplete

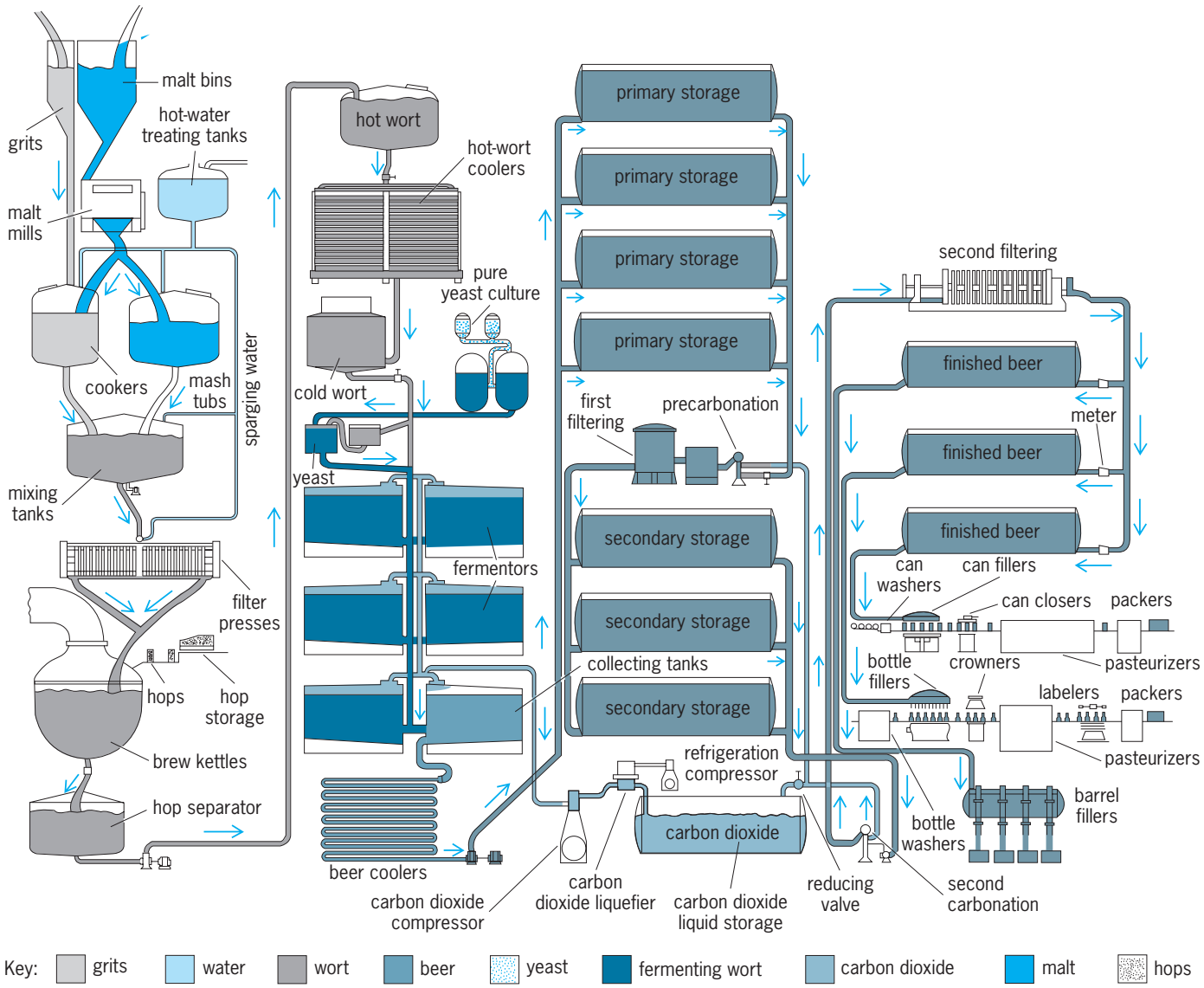

Flow chart of brewing process. (*Pabst Breweries*)

enzyme content (from the brewer's viewpoint), but malt is friable, flavorful, and has an enzyme complement necessary for brewing. The malting process contains three interrelated parts: steeping or soaking the grain, germination, and kilning. *See* BARLEY.

Steeping. When barley is harvested, it contains about 12% moisture and therefore cannot germinate. In large conical-bottomed vessels, barley and water at about 59°F (15°C) are mixed, and the grain rapidly absorbs water. The first steep water quickly becomes brown as the barley is washed. This cleaning is improved by forcing air through the grain and by changing the water often. The air also provides oxygen for germination which begins during steeping. After about 40 h or so, the barley contains about 45% moisture, and the root sheath begins to show.

Germination. The barley is now transferred to large vessels through which cooled humidified air is passed, and screwlike ploughs keep the rootlets from matting. The airflow and moisture are required for the grain to germinate optimally and allow the maltster to control temperature at about 68°F (20°C) and moisture at about 45%. Germination continues 3 to 5 days, depending on the variety of barley and the kind of malt being made.

Kilning. When germination is complete, as judged by the soft consistency of the contents of the kernels and by the growth of the shoot (about three-fourths the length of the kernel beneath the husk), germination is halted by a programmed process of heat and draft (kilning). There are three stages: At first, low heat and high draft reduce malt moisture to about 8 to 10%, then higher heat reduces moisture to about 5%. Finally, intense heat for a short time (called curing) drives the level of moisture to about 4% or less and induces good malt flavor. This heating schedule helps to preserve the enzymes present in the malt (which are heat-sensitive when wet) and to produce a flavorful, friable product (malt), the starch of which has been modified to render it more readily soluble in water. In addition, during the germination

process the starch-splitting enzyme α-amylase is created (β-amylase exists in barley), and substantial protein breakdown to amino acids is achieved.

Brewing. Malt is the main raw material of brewing. Corn (or cornflakes or a variety of corn syrups) or rice are used as adjuncts to provide 20 to 50% of the total raw materials. They provide cheaper starch than does malt, and (because they are low in nitrogenous compounds, color, and flavor) they permit pale-colored beers to be made which have a light and fresh flavor.

In a brewery two main processes occur: the transformation of the malt and other raw materials into a fermentable liquid in the brewhouse and the cellar operations of fermentation and finishing.

The objective of brewhouse operations is to extract the raw materials and to recover the desirable liquid (wort) from the insoluble solids (spent grain). These are the processes of mashing, lautering, and wort boiling.

Mashing. In the United States the adjunct (corn or rice containing a little malt) is gradually heated with water to boiling to gelatinize the starch of the adjunct. For the separate malt mash the malt is first milled, mixed with hot water, and stirred at about 100°F (38°C) for about 10 to 20 min. The boiling adjunct mash is then mixed into the malt mash, and the temperature of this combined mash is 158 to 160°F (70 to 71°C). The temperature is critical because it is here that the α- and β-amylases of the malt intensively digest the starch of the malt and adjunct to produce simple fermentable sugars (such as maltose) and unfermentable dextrins. This conversion takes 15 to 30 min. Many other compounds essential to fermentation and beer quality, including amino acids, vitamins, and minerals, are dissolved in this process.

Lautering. The converted mash is transferred to the lauter vessel—a wide, flat vessel with a slotted false bottom—where the wort can be filtered from the spent grains. The bed of spent grains is then sparged (rinsed) with hot water to recover the maximum amount of sugar.

Boiling. The sweet wort is collected in the kettle and boiled for up to 1 h with 0.25 to 0.35 lb/barrel (0.97 to 1.35 kg/m^3) of hops or extract of hops. The amount of hops used, the blend of foreign and domestic hops, and the time of their addition to the kettle are selected by the brewer. The hop resins, called α-acids, by boiling are extracted and isomerized to iso-α-acids, which impart bitterness to beer. Boiling also sterilizes the wort and precipitates some protein which otherwise would cause haze in the finished beer. After separation in a hop strainer or whirlpool, the wort is cooled, yeast is added, and the fermentation stage begins.

Sugar syrups of various kinds (made from cornstarch) do not have to be digested in brewing and are added directly to the boiling stage. This allows the brewer to produce a more concentrated wort (about 16 to 18% sugar) than is economically feasible with traditional processes (about 10 to 12% sugar). When fermented, this beer is high in alcohol and needs to be diluted to normal strength with deaerated, carbonated water. The advantage of this "high gravity brewing" is that the brewer can increase production without additional equipment.

Fermentation and finishing. Yeasts are unicellular fungi which are classified in 37 genera, but all brewers' yeasts are selected strains of only two species of one genus, namely *Saccharomyces carlsbergensis* (the lager yeast, presently called *S. uvarum*) and *S. cerevisiae* (the ale yeast; other strains of this yeast are used for winemaking and baking). A brewing strain may be unique to one company, though not necessarily, and is chosen because it ferments at the right speed, can be removed easily at the end of the process, and produces beer of the desired flavor.

Yeast is added to wort immediately after it is cooled to about 55–60°F (13–16°C) and as it is being run to a fermenter. A traditional fermenter is a horizontal cylinder of about 1000 barrels (117 m^3) capacity, but some major breweries now have additional vertical tanks of six to eight times this capacity. Fermentation takes approximately 1 week. During this time the yeast grows to about three to five times its original amount, using the sugar and other materials in wort as fuel and food for this growth. Beer flavor characteristics are largely created during fermentation and are the waste products of anaerobic yeast growth. Alcohol (ethanol) and carbon dioxide gas are obvious products of fermentation, but the yeast produces a host of additional materials, including many other alcohols, esters, aldehydes and ketones, sulfur-containing materials, and organic acids. The precise array of materials produced strongly affects beer flavor and is a function of wort composition, conditions of fermentation, and yeast characteristics.

Toward the end of fermentation, when the sugars are almost exhausted, the fermentation is cooled. This aids the yeast's natural tendency to settle out (flocculate). The cooled beer is then transferred off the settled yeast to tanks in the lager (resting) cellar which is kept at 35°F (2°C) or so. Here the beer is allowed to mature and clarify quietly for about 3 weeks in the presence of some yeast transferred from the primary fermentation. *See* FERMENTATION; YEAST.

After lagering, beer is filtered at 29°F (−2°C), usually through diatomaceous earth, carbonated, packaged, and pasteurized to complete the process. However, other treatments may variously be used in finishing. For example, many beers are treated with papain (a protease enzyme from papaya) or an insoluble absorbent (such as silica gel) to remove some protein which otherwise could cause haze in the beer when it is cooled. A fairly common practice at this stage is krausening, which is the addition of some freshly fermenting wort to the beer while it is in large tanks in cold storage to achieve natural carbonation. Beers that have been brewed at high gravity are adjusted with deaerated carbonated water en route to the bottle shop to achieve the required final level of alcohol and flavor components. Hops extract may be added at this point to adjust bitterness to a standard

value, and caramel (made from roasted sugar) may be added to produce dark beer.

Bottling or canning of beer is a highly developed technology. Speed and exclusion of air are the keys to successful packaging of beer, and great improvements have been made in recent years. Because beer does not improve in the package, it should be drunk fresh. Brewers therefore make every effort to get the beer to the consumer as rapidly and as cool as possible. About 80% of production is sold in the retail package; only 15% is marketed in kegs for draft sale.

Light (low-calorie) beer. The calories in beer (about 150 per 10-oz can or 630 joules per 283-g can) arise from the alcohol present (two-thirds) and the residual carbohydrates (one-third), which are flavorless unfermentable dextrins produced from starch during mashing. Low-calorie beer is made by eliminating these dextrins. This is achieved by controlling the amount of dextrins entering the wort, since in making light beer the adjunct used is fully fermentable (dextrin-free) corn syrup; and by converting the dextrins to fermentable sugar, which is done by adding to the cooled wort an enzyme of microbial origin (called amyloglucosidase or glucamylase). Fermentation of such wort converts all the carbohydrate to alcohol. This yields a beer of nearly 5% alcohol which (because the conversion of sugar to alcohol removes few calories) gives little saving of calories. Therefore the beer is diluted to a normal alcohol content (about 3.4 to 3.8% by weight) with pure water. The end product contains no dextrins, and almost all the calories (about 100 or 420 J) derive from an essentially normal content of alcohol. Any further savings in calories, to produce beers of 60 to 90 calories (250–380 J), can be made only by dilution below the normal alcohol content.

M. J. Lewis

Malvales

An order of flowering plants in the core eudicots. The order consists of 10 families and more than 5500 species. The circumscription of the order has been altered greatly in recent years, largely on the basis of deoxyribonucleic acid (DNA) sequence data. Of the five families in the traditional concept (Bombacaceae, Malvaceae, Sterculiaceae, Tiliaceae, and Elaeocarpaceae), the first four have been combined as Malvaceae, due to lack of monophyly of three of these families in their traditional circumscription; and Elaeocarpaceae are included in Oxalidales. In addition, other families have been transferred to Malvales, including Dipterocarpaceae and Thymelaeaceae. The expanded order is characterized by the presence of mucilage in epidermal cells and cavities; and of palmate leaves, stellate hairs, and numerous stamens with partially fused filaments occur frequently.

Malvaceae are cosmopolitan and include economic crops such as cotton (*Gossypium*), cocoa (*Theobroma*), and durian (*Durio*) and horticultural plants such as hollyhocks (*Althaea*) and rose of Sharon (*Hibiscus*). Dipterocarpaceae are important elements of tropical forests, especially in Southeast Asia, and provide hardwood timbers. Other horticultural genera in the order include *Cistus* and *Helianthemum* (Cistaceae) and *Daphne* (Thymelaeaceae). Several genera of Thymelaeaceae provide fibers used for making paper, and *Bixa* (Bixaceae) is the source of the orange dyestuff anatto. *See* CACAO; COTTON; EUDICOTYLEDONS.

Mark Chase

Mammalia

The class Mammalia is the dominant group of vertebrates today. They have ruled the planet since the extinction of the dinosaurs 65 million years ago. There are over 4200 living species of mammals, classified into over 1000 genera, 140 families, and 18 orders. However, the number of extinct mammals is at least five times that. Most living mammals are terrestrial, including such large beasts as elephants, rhinos, hippos, and giraffes, as well as a great diversity of smaller land animals. The largest known land mammal was the extinct 20-ton hornless rhino *Paraceratherium*. Many groups of mammals moved to the water from land-dwelling ancestors. These included manatees and dugongs (which are distantly related to elephants), otters (which are related to weasels), seals, sea lions, and walruses (which are distantly related to bears), and whales (which are distantly related to even-toed hoofed mammals), as well as numerous extinct groups. The living blue whale [reaching 30 m (100 ft) in length and 150 tons] is by far the largest animal that has ever lived. Mammals have also taken to the air, with over 920 living species of bats, as well as numerous gliding forms such as the flying squirrels, phalangerid marsupials, and flying lemurs or colugos. Mammals are even more successful at small body sizes, with hundreds of small species of rodents, rabbits, and insectivores. The smallest living mammal, the 1.5-g Kitti's hognosed bat, is at the lower limit of body size for mammals, since physiology and anatomy prevent them from thriving in the tiny-body-size niche inhabited by insects and other arthropods.

Anatomy and physiology. Mammals are distinguished from all other animals by a number of unique characteristics. These include a body covered with hair or fur (secondarily reduced in some mammals, particularly aquatic forms); mammary glands in the female for nursing the young; a jaw composed of a single bone, the dentary; and three middle ear bones, the incus, malleus, and stapes. All mammals maintain a constant body temperature through metabolic heat. Their four-chambered heart (two ventricles and two atria) keeps the circulation of the lungs separate from that of the rest of the body, resulting in more efficient oxygen transport to the body tissues. They have many other adaptations for their active life-style, including specialized teeth (incisors, canines, molars, and premolars) for biting, tearing, and grinding up food for more efficient digestion. These teeth

are replaced only once in the lifetime of the animal (rather than continuous replacement, found in other toothed vertebrates). Mammals have a unique set of muscles that allow the jaw to move in many directions for chewing and for stronger bite force. Their secondary palate encloses the internal nasal passage and allows breathing while they have food in the mouth. Ribs (found only in the thoracic region) are firmly attached to the breastbone (sternum), so that expansion of the lung cavity is accomplished by a muscular wall in the abdominal cavity called the diaphragm. *See* CARDIOVASCULAR SYSTEM; DENTITION; GHOST IMAGE (OPTICS); HAIR; LACTATION; MAMMARY GLAND; THERMOREGULATION; TOOTH.

All mammals have large brains relative to their body size. Most mammals have excellent senses, and some have extraordinary senses of sight, smell, and hearing. To accommodate their large brains and more sophisticated development, mammals are born alive (except for the platypus and echidnas, which lay eggs), and may require considerable parental care before they are ready to fend for themselves. Juvenile mammals have separate bony caps (epiphyses) on the long bones, separated from the shaft of the bone by a layer of cartilage. This allows the long bones to grow rapidly while still having a strong, bony articulation at the end. When a mammal reaches maturity, these epiphyses fuse to the shaft, and the mammal stops growing (in contrast to other vertebrates, which grow continuously through their lives). *See* BRAIN; NERVOUS SYSTEM (VERTEBRATE); SKELETAL SYSTEM.

Reproduction and classification. The living mammals are divided into three major groups: the monotremes (platypus and echidnas), which still lay eggs, retain a number of reptilian bones in their skeletons, and have other primitive features of their anatomy and physiology; the marsupials (opossums, kangaroos, koalas, wombats, and their relatives), which give birth to an immature embryo that must crawl into the mother's pouch (marsupium), where it finishes development; and the placentals (the rest of the living mammals), which carry the young through a long gestation until they give birth to relatively well-developed progeny. In addition to these three living groups, there were many other major groups, such as the rodentlike multituberculates, now extinct. The most recent classification of the mammals can be summarized as follows:

Class Mammalia
- Subclass Prototheria (monotremes)
- Subclass Theriiformes
  - Infraclass Holotheria
    - Cohort Marsupialia (marsupials or pouched mammals)
    - Cohort Placentalia (placentals)
      - Magnorder Xenarthra (sloths, anteaters, armadillos)
      - Magnorder Epitheria
        - Grandorder Anagalida (= Glires) (rodents, rabbits, elephant shrews)
        - Grandorder Ferae (carnivores, pangolins, many extinct groups)
        - Grandorder Lipotyphla (hedgehogs, shrews, moles, tenreces, and kin)
        - Grandorder Archonta
          - Order Chiroptera (bats)
          - Order Primates (lemurs, monkeys, apes, humans)
          - Order Scandentia (tree shrews)
        - Grandorder Ungulata (hoofed mammals)
          - Order Tubulidentata (aardvarks)
          - Order Artiodactyla (even-toed hoofed mammals: pigs, hippos, camels, deer, antelopes, cattle, giraffes, pronghorns, and relatives)
          - Order Cete (whales and their extinct land relatives)
          - Order Perissodactyla (odd-toed hoofed mammals: horses, rhinos, tapirs, and extinct relatives)
          - Order Hyracoidea (hyraxes)
          - Order Tethytheria (elephants, manatees, and extinct relatives)

This classification does not list all the extinct groups, which include at least a dozen more ordinal-level taxa. It is a considerable improvement over previous mammalian classifications, which used to list over 30 different orders with no indication of how they were related to one another. Although there are more ranks in this classification than is traditional, this nesting of groups within groups accurately reflects the evolutionary branching sequence (see **illus.**). *See* REPRODUCTIVE SYSTEM.

Evolution. Mammals evolved from the Synapsida, an early branch of the terrestrial amniotes that has been erroneously called the mammallike reptiles. This name is inappropriate because synapsids were never related to reptiles. Synapsids and reptiles originated independently from one another about 320 million years ago, and evolved separately ever since. Early synapsids, such as the finback *Dimetrodon*, show relatively few mammalian characteristics, but as their evolution progressed through the late Paleozoic, synapsids became progressively more mammallike.

The first undoubted mammals appeared in the Late Triassic (about 210 million years ago), and were tiny insectivorous forms much like living shrews. Through the rest of the age of dinosaurs, a number of different groups evolved over the next 145 million years of the Jurassic and Cretaceous. Most remained tiny, shrewlike animals, hiding from the dinosaurs in the underbrush and coming out mostly at night. The first two-thirds of mammalian history had passed before the dinosaurs became extinct 65 million years ago, and this allowed mammals to emerge from their shadow.

Opossumlike marsupials are known from the Early Cretaceous (about 110 million years ago), and they were more common than placentals just before the

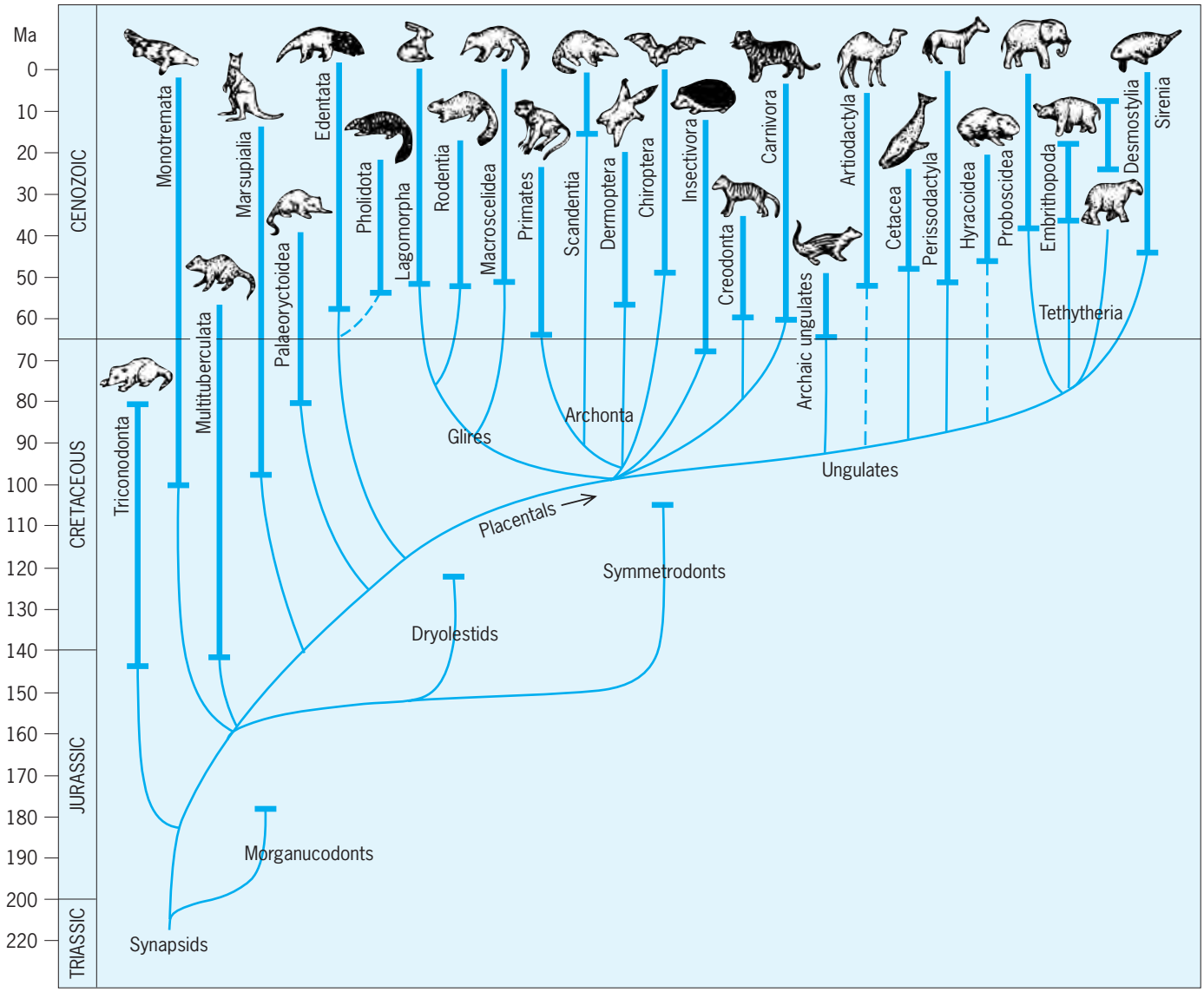

Evolutionary relationships of the major groups of mammals.

end of the age of dinosaurs. Although opossums were found on most continents during the Cenozoic, other marsupials diversified primarily on the southern continents of Australia or South America, which were isolated from the main developments in placental evolution. On both of these continents, marsupials evolved in parallel with placentals, so that Australia has marsupials shaped like woodchucks (wombats), rabbits (bandicoots), flying squirrels (phalangers), monkeys (cuscuses), dogs (Tasmanian wolf), wolverines (Tasmanian devil), anteaters (myrmecobiids), cats (dasyurines), moles (notoryctines), mice (dasycercines), as well as extinct marsupial lions and rhino-sized diprotodonts. South America was once the home of marsupial sabertooths and hyaenalike marsupials known as borhyaenids, although they apparently became extinct about 3 million years ago when placental carnivores came from North America via the Panamanian land bridge.

Between 65 and 55 million years ago, a rapid adaptive radiation (see illus.) yielded all the living orders of placental mammals and many extinct forms as well. The Xenarthra, or edentates, were the first group of placentals to branch off. Although "edentate" means toothless, this is true only of anteaters; sloths and armadillos have simple peglike teeth lacking enamel. The archaic nature of edentates among the placentals is shown by a variety of characters, including a uterus simplex (divided by a septum and lacking a cervix), a slower, less well regulated metabolism, retention of several reptilian bones lost in all other placentals, and a primitive rodlike stapes in the middle ear. Like the marsupial carnivores, edentates evolved in isolation in South America throughout most of their history, developing into a variety of sloths (both tree sloths and huge ground sloths weighing up to 3 tons), anteaters, and armadillos [including the giant glyptodonts, which were

2 m (6.5 ft) long, had 400 kg (880 lb) of body armor, and a spiked club on the tip of the tail]. The long period of isolation in South America ended about 3 million years ago, and ground sloths and glyptodonts migrated to central America and parts of North America, but they disappeared at the end of the last ice age.

The remaining (nonedentate) placentals, or epitheres, diversified primarily in Eurasia and North America and spread throughout the world in the early Cenozoic. The true lipotyphlan insectivorans (represented by shrews, moles, and hedgehogs) continued to diversify all over the northern continents throughout the Cenozoic. Most remained small, and they ate insects, worms, and other small animals, although the extinct giant hedgehog *Dinogalerix* was as big as a large dog and killed sizable prey.

The archontan radiation began with an enormous expansion of primitive lemurlike primates in the early Cenozoic, when the world had dense jungle vegetation all the way to the Poles. Primates declined in the Oligocene, when their forest habitats disappeared, and they became restricted to Africa (Old World monkeys and apes) and South America (New World monkeys). The earliest bats (known from about 50 million years ago) already had fully developed wings. They are the second most diverse group of mammals, after the rodents.

The radiation of the Glires began in the Paleocene of Asia, where numerous primitive relatives of rodents and rabbits are found. In the Eocene, both groups migrated to Europe and North America, where they soon took over the niche of small-bodied fruit, seed, and nut eaters that had been occupied by multituberculates and primitive primates. Although rodents and rabbits are separate orders, they are both characterized by chisellike, ever-growing front incisors that are used in gnawing. However, rodents have only a single pair of incisors, while rabbits have two. The enormous diversification of the Rodentia since the Eocene has given rise to over 1700 species (about 40% of the Mammalia), with forms as large as the pig-sized capybara down to the tiny mice and voles.

Predatory mammals (the Ferae) include the extinct creodonts (an archaic group that were the dominant predators and scavengers of the early Cenozoic), and the living order Carnivora (cats, hyenas, mongooses, civets, dogs, bears, weasels and their kins, seals and sea lions, raccoons, and many extinct groups). All Carnivora are distinguished by distinctive shearing teeth, the carnassials, developed between the last upper premolar and first lower molar. True Carnivora began as weasellike forms in the early Eocene, but by the Oligocene they had taken over most of the predatory niches from the creodonts. By the Miocene, the ancestors of seals and sea lions had evolved from bearlike ancestors. Carnivorous mammals show remarkable examples of convergence on a limited number of body forms. For example, sabertoothed forms evolved four times: once in the creodonts, once in the true cats, and once in the extinct catlike nimravids, which are related to dogs (plus the extinct sabertoothed marsupial). In North America, the borophagine dogs converged on hyenas, with similar bone-crushing teeth.

The hoofed mammals, or ungulates, are first known from about 85 million years ago in central Asia. In the latest Cretaceous and Paleocene, archaic hoofed mammals ("condylarths") were among the most common forms in North America and Asia. From these roots, numerous orders evolved. The first to branch off were the even-toed Artiodactyla, which have two or four toes on each foot and a distinctive ankle structure. First appearing in Pakistan in the earliest Eocene, artiodactyls quickly diversified into a number of different groups, including suoids (pigs, peccaries, hippos), tylopods (camels and their extinct relatives), and ruminants (deer, giraffes, pronghorns, cattle, sheep, goats, antelopes). With their four-chambered stomachs for more efficient digestion, the ruminants became the dominant group of large herbivorous mammals as global climates became drier and grasslands expanded in the later Cenozoic.

The ancestors of whales were large hoofed predators known as mesonychids. Recently, transitional forms between mesonychids and primitive whales have been found in the Eocene of Pakistan. By the Oligocene, whales had diversified into the predatory toothed whales (dolphins, orcas, sperm whales) and filter-feeding baleen whales (blue, right, humpback, gray, and many other whales).

The Perissodactyla, or odd-toed ungulates, have one or three toes on each foot. Today they include horses, rhinos, and tapirs, but they were much more diverse in the past, with huge slingshot-horned brontotheres and bizarre clawed chalicotheres. Closely related to perissodactyls are the hyraxes or conies, and the tethytheres (elephants, sea cows, and their relatives). Tethytheres originated from late Paleocene ancestors that once lived along the Tethys Seaway (which stretched from Gibraltar to Indonesia). Although sea cows spread around the world by the Eocene, the early evolution of hyraxes, and of elephants and their kin was restricted to Africa until the middle Miocene, when both groups spread to Eurasia, and mastodonts even reached North America.

Donald R. Prothero

Bibliography. J. F. Eisenberg, *The Mammalian Radiations*, 1981; J. H. Honacki, K. E. Kinman, and J. W. Koeppl, *Mammal Species of the World*, 1982; D. Macdonald, *The Encyclopedia of Mammals*, 1984; M. C. McKenna and S. K. Bell, *Classification of Mammals above the Species Level*, 1997; R. M. Nowak, *Walker's Mammals of the World*, 1999; D. R. Prothero and R. M. Schoch, *Horns, Hooves, and Flippers: The Evolution of Hoofed Mammals and Their Relatives*, 2000; R. J. G. Savage and M. R. Long, *Mammal Evolution: An Illustrated Guide*, 1986; F. S. Szalay, M. J. Novacek, and M. C. McKenna (eds.), *Mammal Phylogeny*, 1993; J. Z. Young, *The Life of Mammals*, 1975.

Mammary gland

A unique anatomical structure of mammals that secretes milk for the nourishment of the newborn. The mammary gland contains thousands of milk-producing units called alveoli, each of which consists of a unicellular layer of epithelial cells arranged in a spheroid structure (**Fig. 1**). The alveolar epithelial cells take up a variety of nutrients from the blood that perfuses the outer surface of the alveolar structures. Some of the nutrients are then secreted directly into the alveolar lumen; other nutrients are used to synthesize the unique constituents of milk which are then secreted. Each alveolus is connected to a duct through which milk flows. The ducts from many alveoli are connected via a converging ductal system which opens externally by way of the lactiferous pore (**Fig. 2**).

Surrounding each alveolus and its associated small ducts are smooth muscle cells called myoepithelial cells. These cells contract in response to the posterior pituitary hormone oxytocin; milk is thus forced out of the alveoli, through the ductal system, and out the lactiferous pore for the nourishment of the newborn. The release of oxytocin is a neuroendocrine reflex triggered by the stimulation of sensory receptors by the suckling of the newborn. *See* ENDOCRINE MECHANISMS.

Embryology. Mammary glands are basically highly modified and specialized sebaceous glands which derive from ectoderm. In the embryo, mammary lines, formed on both sides of the midventral line, mark the location of future mammary glands. In some species (such as elephants, sirenians, and primates) mammary glands develop in the thoracic region, in others (ungulates, cetaceans) in the inguinal region; in litter-bearing animals they extend from the thoracic to the inguinal region. Along the mammary lines discrete ectodermal ingrowths, called mammary buds, produce a rudimentary branched system of ducts at birth. For each mammary gland, the ingrowths begin as primary sprouts which number from 1, as is the case in the cow, to as many as 20, as in humans. In all species (except the monotremes) a nipple or teat develops in concert with the mammary buds. The most common type of nipple proliferates outward as the primary sprouts grow downward. A second type, found in some marsupials, grows inward into a pocket and everts after sexual maturity. A third type, found in rats and mice, is formed by an epithelial ingrowth in the gland area, and at the time of puberty the nipple emerges as the mammary gland develops. Each teat or nipple has lactiferous pores from 1 to 20 excurrent ducts; the number is determined by the number of primary sprouts that develop in embryonic life. In the most primitive mammal (the duckbill or platypus), which lacks nipples or teats, milk simply oozes out of the two mammary gland areas and is lapped up by the young.

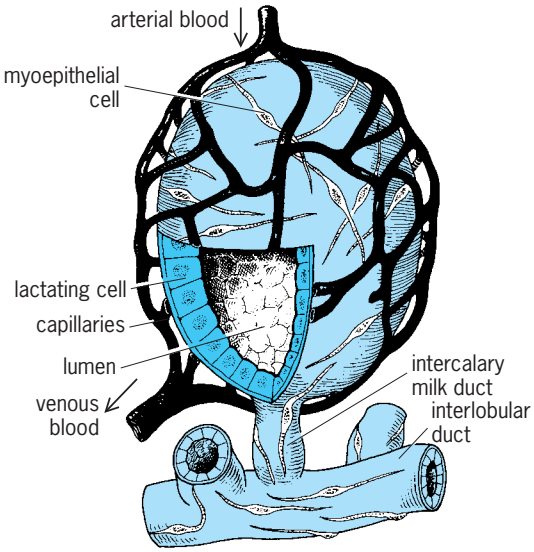

Fig. 1. An alveolus, the milk-producing unit of the mammary gland.

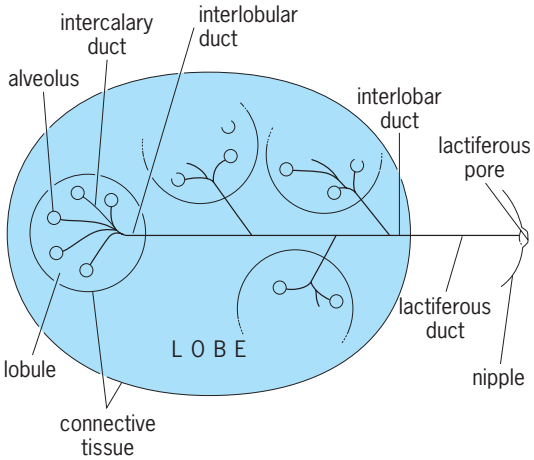

Fig. 2. Structural components of the human mammary gland.

Puberty. From birth to sexual maturity the mammary gland consists of a nipple and a rudimentary ductal system in both males and females. At the onset of puberty in the female, the enhanced secretion of estrogen causes a further development of the mammary ductal system and an accumulation of lipids in fat cells. After puberty in women, the mammary gland consists of about 85% fat cells and a partially developed ductal system. *See* ESTROGEN.

Pregnancy. During pregnancy the mammary gland comes under the influence of estrogen and progesterone which are derived from both the ovary and placenta. These hormones cause a further branching of the ductal system and the development of milk-secreting structures, the alveoli. In humans, approximately 200 alveoli are surrounded by a connective tissue sheath forming a structure called a lobule (Fig. 2). About 26 lobules are packaged via another connective tissue sheath into a larger structure called a lobe. Each of 15–20 lobes is exteriorized into the nipple via separate lactiferous pores. *See* PROGESTERONE.

A complement of hormones maximizes the development of the ductal and lobuloalveolar elements in the mammary gland. Optimal ductal growth is

attained with estrogen, a glucocorticoid, prolactin, and insulin. Maximal lobuloalveolar growth is obtained with estrogen, progesterone, growth hormone, prolactin, a glucocorticoid, and insulin. During pregnancy estrogen and progesterone stimulate mammary development but inhibit milk production.

During the final third of pregnancy, the alveolar epithelial cells begin secreting a fluid called colostrum. This fluid fills the alveoli and causes a gradual enlargement of the breast or udder. At parturition, the inhibitory influence of estrogen and progesterone is removed, and the gland can secrete milk under the influence of a further complement of hormones including prolactin, a glucocorticoid, insulin, and the thyroid hormones.

At birth the mammary glands of about 80% of human newborns are found to secrete a watery substance called witch's milk. During pregnancy the estrogen and progesterone in the mother pass into the fetal circulation and influence the mammary cells of the fetus. *See* GLAND; LACTATION; MAMMALIA; MILK; PREGNANCY. James A. Rillema

Bibliography. T. C. Jones, U. Mohr, and R. D. Hun (eds.), *Integument and Mammary Glands*, 1989; M. C. Neville and C. W. Daniel (eds.), *The Mammary Gland: Development, Regulation, and Function*, 1987; T. B. Mepham, *Physiology of Lactation*, 1991.

Mammography

The radiological imaging of breast tissue. This procedure is used to identify cancer, preferably when still impalpable. Because of improved resolution in high-contrast film, only minor exposure of skin to x-rays is required, so that national screening programs have been set up in many western industrialized countries.

Mammography depends on the tumor being reflected as a dense focus in contrast to surrounding tissue or less dense glandular and ductal parts of the structure. Discrimination between benign and malignant lesions may sometimes be difficult, but may be aided by increased magnification of the image and by spot compression to confirm otherwise equivocally benign lesions. In addition, ultrasonography may be used to distinguish between cystic (mainly benign) and solid (possibly malignant) masses. Many women with impalpable cancers discovered through mammography have microcalcifications as the diagnostic feature; others are diagnosed by the presence of a distinct mass, architectural asymmetry of the glandular and ductal tissue, or tissue distortion.

Mammography has recognized limitations: the possibility exists that in some postmenopausal women it will fail to reveal cancer that is present. Similarly, in premenopausal women in whom the breast tissue is often dense, mammography can be falsely negative in some women with subsequently proven cancers.

National screening mammography programs in Europe have shown some reduction of mortality from breast cancer in screened women between 50 and 70 years of age. The value of screening mammography in women between 40 and 50 who have no family history of breast cancer is unknown.

Western experience suggests that of every 1000 women between the age of 40 and 70 initially screened, approximately 70 will be recalled for more meticulous diagnostic mammography, 15 will have surgical biopsy, and 6 will be found to have cancer. The confirmation necessary for the diagnosis may be made by surgical excision preceded by insertion of a wire under mammographic guidance for localization of the impalpable abnormal shadow or microcalcifications. Ultrasonographic or stereotactic guided core biopsy has also been investigated as a technique to prevent a surgical operation solely for diagnostic purposes.

By revealing unsuspected shadows due to early cancer in asymptomatic women, mammography is estimated to account for much of the apparent increase in incidence of breast cancer: A number of these lesions are still confined to the mammary duct and may be as small as 4 mm. As this type of lesion does not necessarily progress to invasive cancer, it is not certain whether mammography will improve outcome in this group of very small lesions. The ability of mammography to reflect breast changes over time and its value in the surveillance of women with breast cancer treated by conservation techniques (lumpectomy and radiation) has made mammography an invaluable clinical tool. Efforts to refine the technique still further include the exploration of digital mammography using computer-acquired data to provide more accurate stereotactic guidance to the biopsy of impalpable suspicious lesions. *See* BREAST DISORDERS; CANCER (MEDICINE); MEDICAL ULTRASONIC TOMOGRAPHY; RADIOGRAPHY. Alexander J. Walt

Bibliography. D. B. Kopans, *Breast Imaging*, 2d ed., 1998; A. K. Tucker (ed.), *Textbook of Mammography*, 1993.

Mandarin

A name used to designate a large group of citrus fruits in the species *Citrus reticulata* and some of its hybrids. This group is variable in the character of trees and fruits since the term is used in a general sense to include many different forms, such as tangerines, King oranges, Temple oranges, tangelos (hybrids between grapefruit and tangerine), Satsuma oranges, and Calamondin, presumably a hybrid between a mandarin and a kumquat. *See* KUMQUAT; ORANGE; TANGERINE.

Many varieties of citrus falling in the mandarin classification are compact trees with willowy twigs and rather small, narrow, pointed leaves. The fruits are usually yellow, orange, or red and have loose skins which peel easily. A distinguishing feature common to the true mandarins and their hybrids is the chlorophyll in the cotyledons (seed leaves), giving them a pistachio-green color.

Although tangerines are the most extensively planted of the mandarin group, others, particularly the Temple orange, the Murcott orange, and the tangelos, of which there are several varieties, are important commercial fruits in the United States. *See* FRUIT; FRUIT, TREE; SAPINDALES. Frank E. Gardner

Mandibulata

A subphylum of the phylum Arthropoda in the long-established classification scheme which comprises three groups at subphylum level; it is almost certainly not a natural assemblage. Each of the other two subphyla (Trilobitomorpha and Chelicerata) is more likely to be monophyletic. *See* CHELICERATA; TRILOBITA.

The subphylum Mandibulata, defined as those arthropods possessing mandibles (lateral jaws) and antennae, includes the six classes Insecta, Chilopoda, Diplopoda, Symphyla, Pauropoda, and Crustacea. The subphylum could also be defined, more correctly, as comprising all classes of living arthropods that lack chelicerae and pedipalps. In fact, neither in their development nor in their functional morphology are the mandibles of crustaceans homologous with the mandibles of insects. The available fossil evidence for primitive crustaceans also supports a different phyletic origin for these appendages. In addition, insects have a single pair of antennae, while crustaceans have two pairs (borne on the second and third preoral segments of the head, with the second pair of antennae typically biramous). In contrast, most authorities agree on a closer relationship between the insects and the four myriapodous classes, and among these five classes the mandibles are almost certainly homologous. *See* ARTHROPODA; CHILOPODA; CRUSTACEA; DIPLOPODA; INSECTA; PAUROPODA; SYMPHYLA. W. D. Russell-Hunter

Manganese

A metallic element, Mn, atomic number 25, and atomic weight 54.9380 g/mole. Manganese is one of the transition elements of the first long period of the periodic table, falling between chromium and iron. The principal properties of manganese are given in the table. It is the twelfth most abundant element in the Earth's crust (approximately 0.1%) and occurs naturally in several forms, primarily as the silicate ($MnSiO_3$) but also as the carbonate ($MnCO_3$) and a variety of oxides, including pyrolusite (MnO_2) and hausmannite (Mn_3O_4). Weathering of land deposits has led to large amounts of the oxide being washed out to sea, where they have aggregated into the so-called manganese nodules containing 15–30% Mn. Vast deposits, estimated at over 10^{12} metric tons, have been detected on the seabed, and a further 10^7 metric tons is deposited every year. The nodules also contain smaller amounts of the oxides of other metals such as iron (Fe), cobalt (Co), nickel (Ni), and copper (Cu). The economic importance of the nodules as a source of these important metals is enormous. *See* HAUSMANNITE; MANGANESE NODULES; PERIODIC TABLE; PYROLUSITE.

1												13	14	15	16	17	18
1 H	2																2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Manganese is more electropositive than its near neighbors in the periodic table, and consequently more reactive. The bulk metal undergoes only surface oxidation when exposed to atmospheric oxygen, but finely divided metal is pyrophoric.

Manganese is a trace element essential to a variety of living systems, including bacteria, plants, and animals. In contrast to iron (Fe), its neighbor in the periodic table, the exact function of the manganese in many of these systems was determined only recently. The manganese superoxide dismutases have been isolated from bacteria, plants, and animals, and are relatively small enzymes with molecular weights of approximately 20,000. The function of the enzyme is believed to be protection of living tissue from the harmful effects of the superoxide ion (O_2^-), a radical formed from partial reduction of O_2 in the cells of respiring (O_2-utilizing) cells.

The most important biological role yet recognized for manganese is in the enzyme responsible for photosynthetic water oxidation to oxygen in plants and certain photosynthetic bacteria. This reaction represents the source of oxygen gas on the Earth and is therefore responsible for the development of the most common forms of life.

All steels contain some manganese, the major advantage being an increase in hardness, although it also serves as a scavenger of oxygen and sulfur impurities that would induce defects and consequent

Properties of manganese

Property	Value
Atomic number	25
Atomic weight, g/mole	54.9380
Naturally occurring isotope	^{55}Mn (100%)
Electronic configuration	$[Ar]3d^5 4s^2$
Electronegativity	1.5
Metal radius, picometers	127
Melting point, °C (°F)	1244 ± 3 (2271 ± 5.4)
Boiling point, °C (°F)	1962 (3563)
Density (25 °C or 77°F), g/cm^3 (oz/in.^{3})	7.43 (4.30)
Electrical resistivity, ohm·cm	185×10^{-6}

brittleness in the steel. Manganese even has some use in the electronics industry, where manganese dioxide, either natural or synthetic, is employed to produce manganese compounds possessing high electrical resistivity; among other applications, these are utilized as components in every television set. *See* ELECTROLYSIS; GLASS; METAL; TRANSITION ELEMENTS.

George Christou

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; N. N. Greenwood and A. Earnshaw, *Chemistry of the Elements*, 2d ed., 1997; A. G. Sykes (ed.), *Advances in Inorganic Chemistry*, vol. 49, 1999.

Manganese nodules

Concentrations of manganese (Mn) and iron oxides found on the floors of many oceans. The complex growth histories of manganese nodules are revealed by the textures of nodule interiors (**Fig. 1**).

Nodules from certain regions are significantly enriched in nickel, copper, cobalt, zinc, molybdenum, and other elements so as to make them important reserves for these strategic metals. Modern oceanographic surveys have delineated areas of the world's sea floors where nodule abundances and metal concentrations are highest.

Regions of metal-rich nodules. Although manganiferous nodules and crusts have been sampled or observed on most sea floors, attention has focused on the nickel-copper–rich nodules (2–3 wt %) from the north equatorial Pacific belt stretching from southeast Hawaii to Baja California, as well as the high-cobalt nodules and crusts from seamounts in the central Pacific. Manganese nodules from the Atlantic Ocean and from higher latitudes in the Pacific Ocean have significantly lower concentrations of the minor strategic metals. Surveys of the Indian Ocean have revealed metal-enrichment trends comparable to those found in the Pacific Ocean nodules; high nickel-cobalt-copper–bearing nodules are found near the Equator. The ferromanganese nodules and crusts associated with submarine hydrothermal deposits have extremely low concentrations of nickel-cobalt-copper.

Fig. 1. Reflected-light photograph of the polished surface of a sectioned manganese nodule showing the complex growth history of the concretionary deposit (diameter 1.6 in. or 4 cm).

The metal-rich nodules of the north equatorial Pacific are underlain by siliceous ooze sediments composed primarily of skeletons of radiolarians and diatoms. The sea floor here is generally deeper than the carbonate compensation depth, below which the calcium carbonate ($CaCO_3$) in tests of foraminifera, coccoliths, and other calcareous organisms is dissolved at the high pressures (approximately 500 atm or 50 megapascals) and low temperatures (approximately 36°F or 2°C) of the deep ocean. South of the Equator, calcareous ooze sediments predominate and abundances of manganese nodules are considerably lower. At higher latitudes, away from the high biological productivity surface waters that occur near the Equator, pelagic red clays underlie the manganese nodules, which generally have lower Ni + Cu, but higher iron (Fe), contents than equatorial nodules.

The growth rate of marine manganese nodules, determined by decay measurements of traces of radioactive isotopes such as thorium-230 (^{230}Th), protactinium-231 (^{231}Pa), beryllium-10 (^{10}Be), and potassium-40 (^{40}K) contained in the nodules is very slow, averaging a few millimeters each million years. Since the sedimentation rates of accompanying siliceous ooze or pelagic red clay sediments are almost a thousand times faster than the nodule growth rates, an enigma in manganese nodule geochemistry is how they avoid being buried by sediment raining down on them. One explanation is the high populations of benthic organisms living in the inhospitable, dark, high-hydrostatic-pressure, low-temperature environment on deep sea floors which could periodically roll over the nodules and keep them at the sediment-seawater interface. Other possible mechanisms are winnowing of sediments by bottom currents, less dense nodules floating on top of denser sediments, and diagenetic remobilization of manganese in the sediment column, dissolving buried nodules and redepositing the manganese on surface nodules. *See* DEEP-SEA FAUNA.

Chemistry of nodules. Marine manganese nodules are usually classified by mode of formation into hydrogenetic, diagenetic, and hydrothermal types. A fourth, mixed type also exists. Hydrogenetic nodules form by direct precipitation of manganese–iron oxide phases onto an existing nucleus at the sediment-water interface; diagenetic nodules are believed to be biologically driven; and hydrothermal nodules form by submarine hydrothermal activity. Hydrothermal nodules are rarely reported; in submarine hydrothermal environments, manganese–iron oxides occur mainly as crusts. Cobalt is the only strategic metal reported in hydrogenetic nodules that have relatively low manganese/iron ratios. Nodules

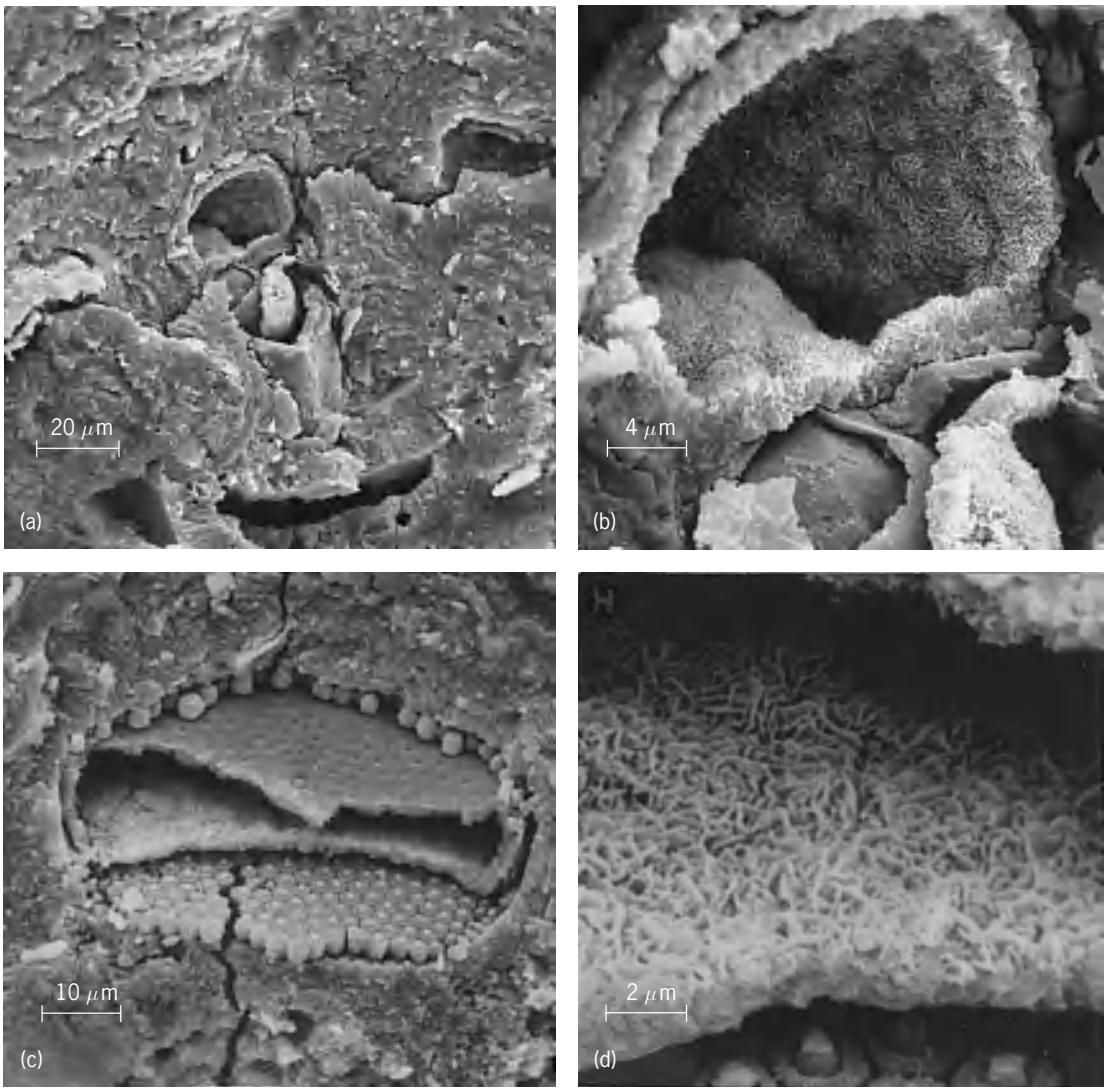

Fig. 2. Scanning electron microscopy photographs showing the growth of Ni + Cu–rich todorokite inside a manganese nodule from the north equatorial Pacific. The meshworks of tiny todorokite crystals (*a, b*) line a cavity and (*c, d*) replace biogenic debris in a fossil cast inside the nodule. (*From R. G. Burns and V. M. Burns, Manganese oxides, in R. G. Burns, ed., Marine Minerals, Mineralogical Society of America Publication Series Reviews in Mineralogy, vol. 6, pp. 1–46, 1979*)

with high nickel, cobalt, and copper contents are diagenetic and have relatively high manganese/iron ratios. Chemical differences exist between the outermost top layer (exposed to seawater) and bottom layer (immersed in sediment) of manganese nodules. Surfaces buried in underlying sediments are generally higher in manganese, nickel, and copper contents, compared to the more iron-cobalt–rich surfaces exposed to seawater. Episodic rolling-over of a nodule would account for fluctuating concentrations of manganese, iron, nickel, cobalt, copper, and other metals across sectioned manganese nodules.

There is a biological source of the metals and diagenetic remobilization of them in the underlying sediment (**Fig. 2**). Very fine grained, highly dispersed clay minerals and oxides of iron and manganese, derived from various sources, are buried with biogenic debris from deceased organisms living in the surface waters of the equatorial photic zone. Traces of the strategic metals, nickel, copper, cobalt, and so on, either are adsorbed onto the surfaces of clays or manganese and iron oxides or are organically complexed in biological matter settling to the sea floor. A reducing environment is set up at depths in the buried sediments, and organic debris is oxidized by the manganese(IV) and iron(III) oxides, leading to remobilization of soluble cations such as manganese, iron, nickel, and copper in pore waters of the sediments. Oxidation by aerated pore water or ocean bottom water near the sediment-seawater interface, first of iron and then of manganese, produces insoluble authigenic hydrated manganese(IV) and iron(III) oxide phases that accrete onto surfaces of manganese nodules lying in the sediments. Ions of nickel, copper, and other elements are bound as essential constituents in the crystal structures of these diagenetically formed manganese oxides. *See* AUTHIGENIC MINERALS; DIAGENESIS.

Minerals in manganese nodules. Manganese nodules contain extremely small, poorly crystalline

phases of minerals, often finely intermixed. In cobalt-rich nodules from seamounts and relatively shallow sea floors, vernadite (a manganese oxide; δ-MnO_2) accommodates cobalt(III) ions, which have ionic radii identical to those of manganese(IV) ions. Vernadite, which can incorporate iron into its structure, is the characteristic mineral phase in hydrogenetic nodules. The characteristic mineral in diagenetic nodules is todorokite, which concentrates nickel and copper ions in the metal-rich nodules. Todorokite has a framework structure consisting of long chains and tunnels that accommodates both divalent transition metals, and larger cations and water molecules. Ions such as nickel and copper stabilize the todorokite structure. In hydrothermal manganese nodules and crusts near hot-water sources, relatively well-crystallized todorokite occurs along with birnessite, a layer mineral. Iron occurs in nodules as feroxyhyte, which is usually so fine-grained as to be x-ray amorphous and may be epitaxially intergrown with vernadite, especially in hydrogenetic nodules from pelagic red clays and seamounts. Other minerals in low or trace amounts in manganese nodules include smectite, phillipsite, plagioclase, and quartz. *See* HYDROTHERMAL VENT; MARINE MINING; MARINE SEDIMENTS; ORE AND MINERAL DEPOSITS; TODOROKITE.

Roger G. Burns; Virginia Mee-Burns

Bibliography. R. G. Burns and V. M. Burns, The mineralogy and crystal chemistry of deep-sea manganese nodules, a polymetallic resource of the twenty-first century, *Phil. Trans. Roy. Soc. London*, A285:249–258, 1977; D. S. Cronan, *Marine Minerals in Exclusive Economic Zones*, 1991; A. Martin-Barajas, E. Lallier-Verges, and L. Leclaire, Characteristics of manganese nodules from the Central Indian Basin: Relationship with the sedimentary environment, *Mar. Geol.*, 101:249–265, 1991; P. A. Rona and S. D. Scott, A special issue on sea-floor hydrothermal mineralization: New perspectives, *Econ. Geol.*, 88:1935–1976, 1993.

Mango

A tree (*Mangifera indica*) of the family Anacardiaceae that originated in the Indo-Burma region and is now grown throughout the tropical and subtropical areas of the world. The mango is a medium to large evergreen tree; it produces a dense, round canopy, with leaves up to 15 in. (38 cm) long by 2.5 in. (10 cm) wide, which are reddish brown when young and dark green when mature. The tree puts down a deep taproot and has many fibrous roots near the surface. Small, pinkish-white flowers, numbering 600 to 2000 and mostly staminate (male), grow on large terminal panicles. *See* SAPINDALES.

In the United States, mangoes are grown only in Florida and on a small scale in Hawaii, usually as backyard trees. Their greatest importance is in India, which contains about 75% of the world's area of mango production. The ripe fruit is eaten raw as a dessert or used in the manufacture of juice, jams, jellies, and preserves. Unripe fruit can be made into pickles or chutneys. Mangoes are a good source of vitamins A and C.

Mango is grown from sea level to 4000 ft (1220 m) in the tropics. Optimum growth temperature is 75–80°F (24–27°C). Dry weather during flowering and fruiting is important. A high carbon/nitrogen ratio is required for flower initiation. Mangoes tend to bear fruit every two years.

Pollination, which is necessary for setting the fruit, is usually accomplished by flies. Rain and high humidity reduce the fruit set, especially if the anthracnose fungus (*Colletotrichum gloeosporioides*) is present. Benlate or captan sprays give control.

There are three groups of mangoes grown: (1) seedling races in the Western Hemisphere (such as Turpentine, Number 11, and Hawaiian), producing fruit of inferior quality and poor color; (2) Indochinese or Saigon, including Philippine, types which are also rather colorless but have flesh with practically no fiber and a delicate flavor (these are polyembryonic and seedlings come true to type); and (3) the Indian race which is better flavored, has brilliant red, purple, and yellow skins, and is good for commercial production. Practically all of these are monoembryonic and must be propagated asexually. Usually either a side wedge or a veneer graft is used.

Florida developed a large number of mango cultivars in a very short time. The Haden originated in Miami as a seedling of Mulgaba, an Indian cultivar introduced to Florida in 1883. Haden was the standard in commerce for more than 40 years. In Hawaii, Haden and Pirie are popular. Other cultivars include Irwin, Edward, Smith, Tommy Atkins, Kent, Keitt, Palmer, Sensation, Pope, Momi-k, Gouveia, and Ah Ping. *See* FRUIT, TREE.

Robert M. Warner

Bibliography. S. L. Kochhar, *Tropical Crops: A Textbook of Economic Botany*, 1989; S. E. Malo, Mango and avocado cultivars, *Fla. State Hort. Soc.*, 83:357–362, 1970.

Mangrove

A taxonomically diverse assemblage of trees and shrubs that form the dominant plant communities in tidal, saline wetlands along sheltered tropical and subtropical coasts. The development and composition of mangrove communities depend largely on temperature, soil type and salinity, duration and frequency of inundation, accretion of silt, tidal and wave energy, and such fortuitous factors as cyclone or flood frequencies (see **illus.**). Extensive mangrove communities seem to correlate with areas in which the water temperature of the warmest month exceeds 75°F (24°C), and they are absent from waters that never exceed 75°F (24°C) during the year. Intertidal, sheltered, low-energy, muddy sediments are the most suitable habitats for mangrove communities, and under optimal conditions, forests up to 148 ft (45 m) in height can develop, such as those in Ecuador, Thailand, and Malaysia. Where less

Mangroves, Great Barrier Reef, off northeast Australia. (*a*) Mangroves at northwestern corner of King Island. (*b*) Interior of mangrove swamp, Newton Island. (*c*) Mangroves on north side of Howick Island. (*From J. A. Steers, Salt marshes, Endeavour, 18(70):75–82, 1959*)

favorable conditions are found, mangrove communities may reach maturity at heights of only 3 ft (1 m). *See* ECOSYSTEM.

Approximately 58,000 mi^2 (150,000 km^2) of the land in the Americas, Africa, Asia, and Australia is occupied by mangrove communities, which show a great similarity in structure, flora, and function. Plants of the mangrove community belong to many different genera and families, many of which are not closely related to one another phylogenetically. However, they do share a variety of morphological, physiological, and reproductive adaptations that enable them to grow in an unstable, harsh, and salty environment. On the basis of the commonality of those various adaptations, approximately 80 species of plants belonging to about 30 genera in over 20 families are recognized throughout the world as being indigenous to mangroves. About 60 species occur on the east coasts of Africa and Australasia, whereas about 20 species are found in the Western Hemisphere.

At the generic level, *Avicennia* and *Rhizophora* are the dominant plants of mangrove communities throughout the world, with each genus having several closely related species in both hemispheres. At the species level, however, only a few species, such as the portia tree (*Thespesia populnea*), the mangrove fern (*Acrostichum aureum*), and the swamp hibiscus (*Hibiscus tiliaceus*), occur in both hemispheres. Such a present-day distribution can be satisfactorily explained only if the ancestors of the mangroves evolved in the Lower Cretaceous, were dispersed outward from their center of origin in the remnants of the Tethys Sea, and subsequently developed further as two isolated groups following the closure of the Mediterranean Sea as a dispersal route between the Eastern and Western hemispheres. *See* CRETACEOUS.

The mangrove community is often strikingly zoned parallel to the shoreline, with a sequence of different species dominating from open water to the landward margins. These zones are the response of individual species to gradients of inundation frequency, waterlogging, nutrient availability, and soil salt concentrations across the intertidal area, rather than a reflection of ecological succession, as earlier studies had suggested. Not surprisingly, the zonation patterns show some similarity between the Eastern and Western hemispheres. In the Americas and western Africa, *Rhizophora* forms the outermost (seaward) zone, followed by *Avicennia*, then *Laguncularia*, with a sporadic landward fringe of *Conocarpus*. In the Eastern Hemisphere, *Rhizophora* forms the outermost zone, although at some localities *Sonneratia* and *Avicennia* may also be present. *Avicennia* generally forms monospecific stands behind the outer zone, followed by mixed stands of *Bruguiera*, *Heritiera*, and *Xylocarpus*, with a landward zone of *Ceriops* mixed with *Lumnitzera* and *Avicennia*. *See* ECOLOGICAL SUCCESSION.

Most plants of the mangrove community are halophytes, well adapted to salt water and fluctuations of tide level. Many species show modified root structures such as stilt or prop roots, which offer support on the semiliquid or shifting sediments, whereas others have erect root structures (pneumatophores) that facilitate oxygen penetration to the roots in a hypoxic environment. Salt glands, which allow excess salt to be extruded through the leaves, occur in several species; others show a range of physiological mechanisms that either exclude salt from the plants or minimize the damage excess salts can cause by separating the salt from the sensitive enzyme systems of the plant. Several species have well-developed vivipary of their seeds, whereby the hypocotyl develops while the fruit is still attached to the tree. The seedlings are generally buoyant, able to float over long distances in the sea and rapidly establish themselves once stranded in a suitable habitat. *See* PLANTS, SALINE ENVIRONMENTS OF.

A mangrove may be considered either a sheltered, muddy, intertidal habitat or a forest community; consequently, the mangrove fauna comprises elements dependent on either of these habitats. The sediment surface of mangrove communities abounds with species that have marine affinities, including brightly colored fiddler crabs, mound-building mud lobsters (*Thalassina anomala*), and a variety of mollusks and worms, as well as specialized gobiid fish, whose ability to travel over the mud surface by using their fins has earned them the common name mudskippers. The waterways among the mangroves are important feeding and nursery areas for a variety of juvenile finfish as well as crustaceans such as portunid crabs and penaeid shrimps, all of which are important sources of protein for indigenous populations. Animals with forest affinities that are associated with mangroves include snakes, lizards, deer, tigers, crab-eating monkeys, bats, and many species of birds, including flamingos, ibis, sea eagles, herons, pelicans, and several species of kingfisher.

Economically, mangroves are a major source of timber, poles, thatch, and fuel. The bark of some trees is used for tanning materials, whereas other species have food or medicinal value. In Malaysia, Bangladesh, India, and Thailand, managed forest operations produce commercial timber, fuelwood, charcoal, or woodpulp. Subsistence harvesting of mangrove resources is widespread in Central America and parts of Asia and Africa. Elsewhere, mangroves are gathering areas for honey and fodder. Despite their obvious economic value, mangroves have often been considered wastelands and have been converted to other forms of land use. Gradually, however, opinion has changed, and mangrove communities are now recognized not merely for their economic value but also as living systems of major intrinsic scientific interest. *See* ECOLOGICAL COMMUNITIES; FOREST MANAGEMENT. Peter Saenger

Bibliography. P. Hutchings and P. Saenger, *Ecology of Mangroves*, 1987; V. Jaccarini and E. Martens (eds.), *The Ecology of Mangrove and Related Ecosystems*, 1993; P. B. Tomlinson, *The Botany of Mangroves*, 1994.

Manifold (mathematics)

A Hausdorff topological space with an n-dimensional atlas of charts for some integer n. The integer n is called the dimension of the manifold. If M is a Hausdorff space, an n-chart on M is a pair (U, φ) with U an open subset of M and φ a homeomorphism of U onto an open subset $\varphi(U)$ of n-dimensional euclidean space $\mathbf{R}^n$. An n-dimensional atlas for M is a system of n-charts (U, φ) such that the union of the sets U is all of M. *See* TOPOLOGY.

The terminology "charts" and "atlases" comes from the geographer's way of viewing the surface of the Earth. The manifold described in this case is the two-dimensional sphere $M = S^2$.

Many mathematicians include an additional restriction in the definition of manifold—sometimes the underlying topological space is connected, sometimes the topology is given by a metric, sometimes both.

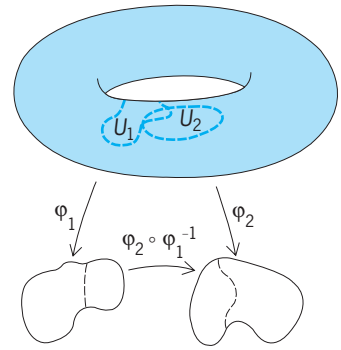

Sets and mappings defining compatible n-charts (U_1, φ_1) and (U_2, φ_2).

Differentiable structure. With an appropriate additional structure on a manifold, it is possible to speak of differentiability or smoothness of real-valued functions. In this setting, the notion of a smooth mapping between smooth manifolds can be defined, and the differential of a smooth mapping, which is a generalization of the derivative, can be introduced.

Two charts (U_1, φ_1) and (U_2, φ_2) in an atlas for an n-dimensional manifold M are compatible (for defining smoothness) if the mapping $\varphi_2 \circ \varphi_1^{-1}$ from the open set $\varphi_1(U_1 \cap U_2)$ in $\mathbf{R}^n$ to the open set $\varphi_2(U_1 \cap U_2)$ is infinitely differentiable and has an infinitely differentiable inverse function (see **illus.**). If all the charts in an atlas are compatible with one another, the atlas is said to determine a differentiable structure on M, and M with its atlas is then called a smooth manifold. Two atlases determine the same differentiable structure if all the charts in question are compatible with one another. Smooth manifolds are called also differentiable manifolds, differential manifolds, and C^∞ manifolds.

Examples. If M is an open subset of $\mathbf{R}^n$, then M and the identity mapping form a single chart, that is, an atlas of compatible charts for M. The resulting differentiable structure is the standard differentiable structure on this open subset of $\mathbf{R}^n$.

The unit sphere S^n in $\mathbf{R}^{n+1}$ can be made into a smooth manifold of dimension n by using the two charts given in Eqs. (1) and (2). The function φ_1 is defined on $U_1 = S^n - \{(0, \ldots, 0, -1)\}$, and the function φ_2 is defined on $U_2 = S^n - \{(0, \ldots, 0, -1)\}$.

$$\varphi_1(x_1, \ldots, x_{n+1}) = \left(\frac{x_1}{1 - x_{n+1}}, \ldots, \frac{x_n}{1 - x_{n+1}}\right) \tag{1}$$

$$\varphi_2(x_1, \ldots, x_{n+1}) = \left(\frac{x_1}{1 + x_{n+1}}, \ldots, \frac{x_n}{1 + x_{n+1}}\right) \tag{2}$$

The projective space $\mathbf{R}P^n$ is another example of a smooth manifold. A point of $\mathbf{R}P^n$ is all nonzero points on a line through the origin of $\mathbf{R}^{n+1}$ and is written $[(x_1, \ldots, x_{n+1})]$. The space $\mathbf{R}P^n$ is given the quotient topology. An atlas (U_j, φ_j), $1 \le j \le n + 1$, is defined

as follows: U_j is the set of members of $\mathbf{R}P^n$ whose jth coordinate is nonzero, and φ_j is given on U_j by Eq. (3). The space $\mathbf{R}P^2$ is the underlying set in which

$$\varphi_j[(x_1, \ldots, x_{n+1})] = x_j^{-1}(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_{n+1}) \quad (3)$$

two-dimensional projective geometry is studied. *See* GEOMETRY; PROJECTIVE GEOMETRY.

The product of two smooth manifolds is a smooth manifold. The torus $S^1 \times S^1$ (surface of a doughnut) is an example. *See* TORUS.

Nonsingular plane curves give examples of one-dimensional smooth manifolds, and nonsingular surfaces in $\mathbf{R}^3$ give examples of two-dimensional smooth manifolds. These facts can be proved from the implicit function theorem. *See* DIFFERENTIAL GEOMETRY; PARTIAL DIFFERENTIATION.

Smooth functions and mappings. Functions on a smooth manifold can be referred to functions on euclidean space by means of the charts, which can be regarded as supplying local coordinate systems for the manifold. A real-valued function on a smooth manifold or a mapping between smooth manifolds is said to be smooth if, when referred to local coordinates, it is infinitely differentiable.

Tangent space. The real-valued smooth functions on M form a real vector space. A linear functional L on this space is called a tangent vector at a point p of M if Eq. (4) is satisfied for all smooth f and g. In local coordinates, L is a linear combination, with numerical coefficients, of first partial derivatives evaluated at p. The vector space of all tangent vectors at p is called the tangent space to M at p. If M has dimension n, its tangent space at p has dimension n. *See* LINEAR ALGEBRA; OPERATOR THEORY.

$$L(fg) = f(p)L(g) + g(p)L(f) \quad (4)$$

Differential. If $\varphi: M \to N$ is a smooth mapping between smooth manifolds, p is a point of M, and X is a tangent vector at p, then a tangent vector L to N at $\varphi(p)$ is defined by Eq. (5), where f is any smooth real-valued function on N; $d\varphi_p(X)$ is written for L. The mapping $d\varphi_p$ defined in this way and carrying the tangent space to M at p to the tangent space to N at $\varphi(p)$ is linear and is called the differential of φ at p.

$$Lf = X(f \circ \varphi) \quad (5)$$

If M and N are both the real line, the differential of φ at p is a multiple of the differential of the identity mapping at p, and that multiple is exactly the derivative of φ at p. *See* DIFFERENTIATION.

Other structures. Alternate definitions of compatibility of charts lead to other classes of manifolds. Thus a manifold with its atlas is said to be piecewise linear (PL) or real analytic if all the mappings $\varphi_2 \circ \varphi_1^{-1}$ have the corresponding property.

If the dimension n is even, say $2m$, then the open sets in $\mathbf{R}^n$ can be regarded as open sets in the complex space $\mathbf{C}^m$ The manifold with its atlas is called a complex manifold of dimension m if the mappings $\varphi_2 \circ \varphi_1^{-1}$ are analytic functions of several variables. Complex manifolds are used in studying varieties in algebraic geometry. *See* ALGEBRAIC GEOMETRY; COMPLEX NUMBERS AND COMPLEX VARIABLES; SERIES.

Anthony W. Knapp

Bibliography. L. Conlon, *Differentiable Manifolds: A First Course*, 2d ed., 2001; G. de Rham, *Differentiable Manifolds: Forms, Currents, Harmonic Forms*, 1984; A. A. Kosinski, *Differential Manifolds*, 1997; S. Lang, *Differential and Riemannian Manifolds*, 3d ed., 1995; K. Shiohama (ed.), *Geometry of Manifolds*, 1989; F. W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, 1996.

Mantophasmatodea

An order of Insecta, known as heel-walkers or gladiators, occurring in Africa south of the Equator. Members of the 13 known species are 1–3 cm (0.4–1.2 in.) long, slender, moderately long-legged, and secondarily wingless. The habitus (general appearance) resembles a mixture of grasshopper, stick-insect, and praying mantis (see **illustration**).

While amber fossils were mentioned in the literature in 1997 (nymphs only), the first species was not described until 2001 (*Raptophasma*, adult specimens). The two first extant specimens, collected in Namibia and Tanzania, were found in the natural history museums of Berlin and London in 2001 and described in 2002 (*Mantophasma* and *Tanzaniophasma*). During 2002 and 2003 living specimens were reported from many places in Namibia and South Africa.

Morphology. The biting mouthparts are directed downward. The antennae are long and thin; the forelegs have a strong femur and two rows of spines on the tibia. The females possess a short ovipositor (egg-laying device) on the posterior abdominal segments. The cerci (appendages of last abdominal segment) are short and one-segmented, forming claspers in the male. Derived characters distinguishing Mantophasmatodea from other insects are

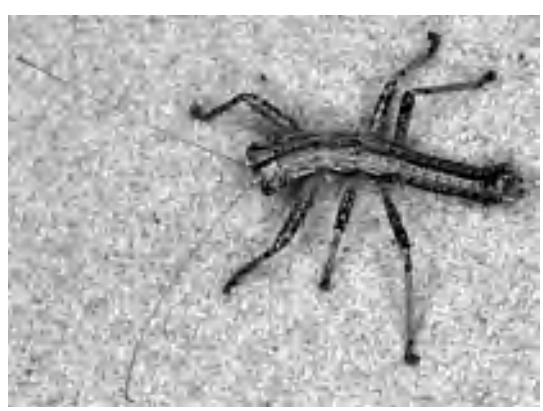

Male of *Karoophasma biedouwensis* from Succulent Karoo biome of western South Africa. (*Photo by M. Picker, University of Cape Town*)

(1) anterior tentorial pits (invagination points of the skeleton inside the head) located far above the base of the mandibles; (2) antennae differentiated into a sparsely setose, multijointed basal part and a wider, more strongly setose distal part with seven joints showing specific length ratios; (3) a huge lobe (arolium) between the claws of the legs. The coloration is patchy, mainly brown, yellow, grey, and green, varying among and within species.

Diversity and distribution. Four species occur in Namibia (genera *Mantophasma, Sclerophasma, Praedatophasma*, and *Tyrannophasma*), eight in western South Africa (*Austrophasma, Karoophasma, Lobophasma, Hemilobophasma*, and *Namaquaphasma*), and one in Tanzania (*Tanzaniophasma*). Species differ mainly in the structure of the male and female genitalia; *Praedatophasma* and *Tyrannophasma* are conspicuously spinose (hence the common name "gladiator"). The single species usually exhibit limited distribution. Mantophasmatodea inhabit several major biomes of southwestern Africa: the dry Nama Karoo and Succulent Karoo, and the more humid Fynbos. Fossils from the Lower Tertiary were found in Europe (*Raptophasma* and *Adicophasma* in Baltic amber).

Life history. The heel-walkers live singly and prey on insects, which they grasp using their strong, spiny forelegs. The terminal tarsomere (foot segment), with the claws and arolium, is usually held elevated (hence the common name). The males use a median projection of their subgenital plate to knock on the ground—probably for communication. To copulate, the male jumps onto the female and bends its abdomen aside and downward to bring the copulatory organs into contact. The female produces sausage-shaped egg pods, each containing 10–20 eggs, enclosed in a hard envelope made from gland secretions and sand. The nymphs, which resemble the adults (hemimetabolism), hatch in the early humid season and reach maturity near its end.

Phylogenetic relationships. Anatomical details of the spiracles, musculature, and ovipositor show that despite their lacking wings the Mantophasmatodea belong to the subclass Pterygota of the Insecta ("winged insects"). DNA studies suggest the wingless Notoptera (ice-crawlers) to be their closest relatives. The morphological evidence is conjectural: different characters point to a close relationship with Dictyoptera (cockroach-mantis-termite group) based on the structure of the male genitalia and spermatozoa, with Phasmatodea (stick-insects) based on the sclerotized process above the male genitalia, with Plecoptera (stoneflies) based on the median process on the male subgenital plate, or Notoptera based on the foregut structure. Klaus-Dieter Klass

Bibliography. K.-D. Klass, Mantophasmatodea, pp. 182–184 in *2004 McGraw-Hill Yearbook of Science & Technology*, McGraw-Hill, New York, 2004; K.-D. Klass, Mantophasmatodea, pp. 677–678 in V. H. Resh and R. T. Cardé (eds.), *Encyclopedia of Insects*, Academic Press, San Diego, 2003; K.-D. Klass et al., Mantophasmatodea: A new insect order with extant members in the Afrotropics, *Science*, 296:1456–1459, 2002; K.-D. Klass et al., The taxonomy, genitalic morphology, and phylogenetic relationships of southern African Mantophasmatodea (Insecta), *Entomol. Abhandlungen*, 61:3–67, 2003.

Manufactured fiber

Any of a number of textile fibers produced from chemical substances of natural origin or synthetic origin; the latter are also known as synthetic fibers. Among the natural sources of manufactured fibers are plant cellulose and protein, rubber, metals, and nonmetallic inorganics. The synthetic fibers are produced from organic intermediates derived from petroleum, coal, and natural gas.

In addition to names based on their chemical composition, manufactured fibers may be known by the manufacturer's name or trademark; terms referring to characteristic properties, for example, thermoplastic fibers; and method of manufacture, for example, wet-spun fiber. In the United States, the Textile Fiber Products Identification Act (TFPIA) standardized the nomenclature. In the following discussions, each fiber will be defined according to the TFPIA. Fiber structures are shown in **Fig. 1**.

Manufacture

With the exceptions of glass and metal fibers, the manufactured fibers are made from very long chainlike molecules called linear polymers (**Fig. 2**). These polymers may be naturally occurring (cellulose from cotton or wood pulp) or may be synthetic (polyester). When the polymer is of biological origin, the problem is one of converting it from its natural form into a fiber. The synthetic polymers may be formed either by addition polymerization or by condensation polymerization. In the former process, small molecules (monomer) containing unsaturated carbon-carbon bonds are built into polymers by using catalysts to open the double bonds joining the carbon atoms of one molecule. This reactive molecule subsequently "attacks" a second molecule, joins with it, and propagates the reactive entity. Thus, a number of monomer units may be joined together in a single chain by the addition of links of monomer. The growing chains may be terminated by the joining of two growing chains or by the addition of substances which will combine with the reactive species to form inactive chain ends.

In condensation polymerization, molecules containing reactive groups at each end are combined, with the elimination of water as a by-product, to yield larger molecules which still contain reactive groups at each end. The reactive groups can continue to combine to build, step by step, a large polymer. The reaction may be terminated by the addition of a stopping agent, which will inactivate one or both of the reactive species. *See* POLYMERIZATION.

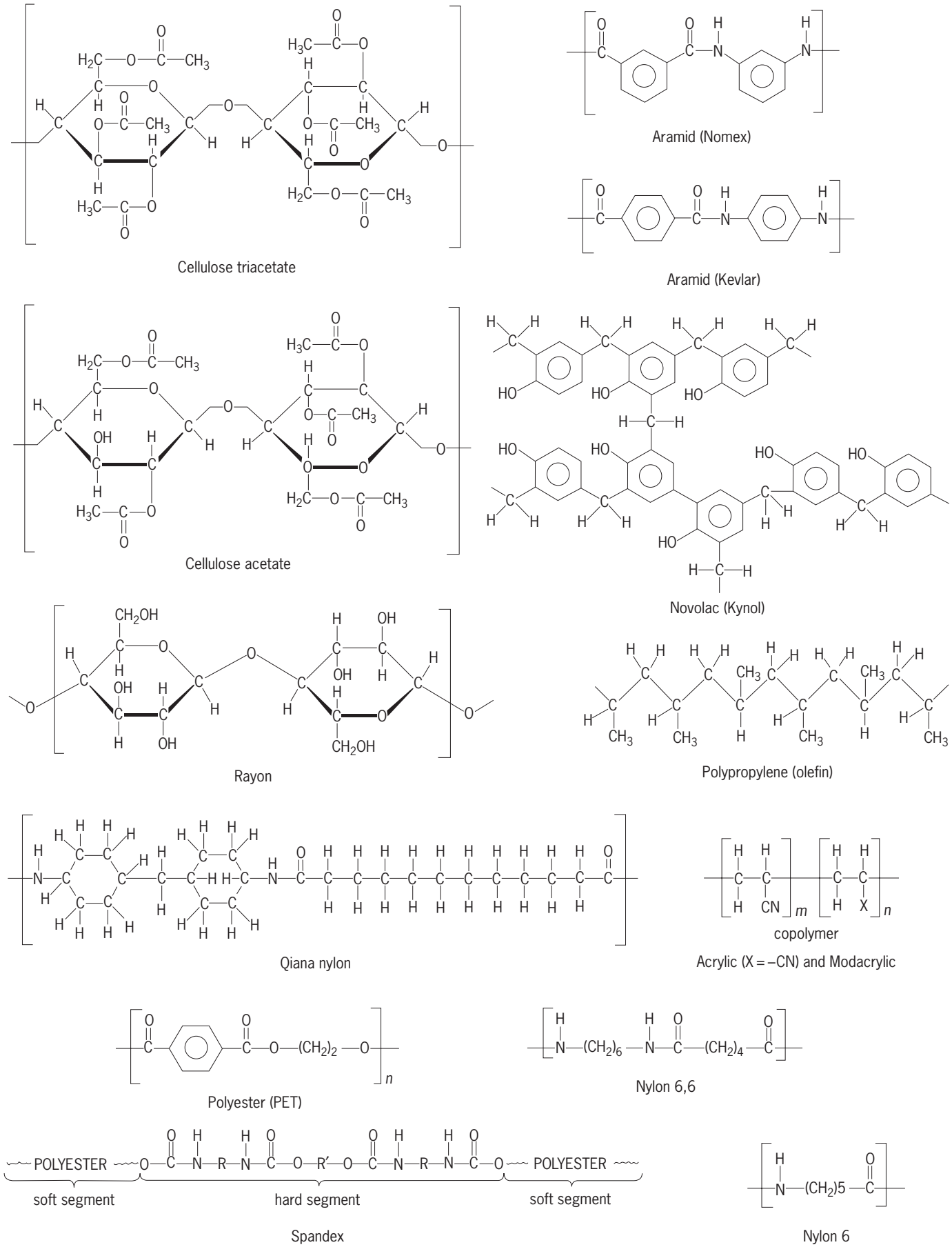

Fig. 1. Structures of the manufactured fibers. (*After B. F. Smith and I. Block, Textiles in Perspective, Prentice-Hall, 1982*)

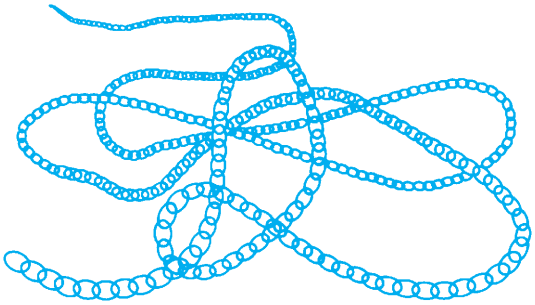

Fig. 2. Idealized polymer modeled as a chain of many identical links. The molecule tends to curl rather than stretch out. (*After B. F. Smith and I. Block, Textiles in Perspective, Prentice-Hall, 1982*)

Polymer characteristics. Irrespective of their chemical nature, fiber-forming polymers must possess the following characteristics: (1) great length—at least 200 monomer units must be joined in a chain; (2) a high degree of intramolecular and intermolecular attraction, whether through primary chemical bonds or other attractive forces; (3) the ability to be oriented along the axis of the fiber; and (4) the ability to form well-ordered crystals or pseudocrystals. All of these parameters are sensitive to the chemical nature of the polymer and the processes of manufacture of the fiber. In turn, they establish the properties of the fiber, such as strength, flexibility, resilience, and abrasion resistance, which contribute to their usefulness in various end uses for apparel, home furnishings, and commercial and industrial applications.

Only a fraction of those substances capable of forming fibers prove to have all of the characteristics necessary for commercial success. The fiber types of major importance in the United States are classified by composition as follows: cellulosic (composed of regenerated cellulose, cellulose diacetate, and cellulose triacetate); synthetic (composed of polyamide, polyester, polyacrylic, polyvinyl, and polyolefin resins); and inorganic (composed of glass and metal).

All of the manufactured fibers are produced according to the same principles: (1) the fiber-forming material must first be made fluid; (2) the fluid is forced under pressure (extruded) through tiny holes into a medium which causes it to solidify; and (3) the solid fibers are further processed to obtain their optimum properties.

Fiber production. Depending upon the starting materials, typically one of three procedures is used to produce fibers.

Wet spinning. The major example of this process is the production of rayon by the viscose process: the polymer is dissolved in an applicable reagent to form the fluid (dope). The fluid is then pumped through metal plates (spinnerets) containing many small holes into a liquid bath of appropriate composition. A chemical reaction between the spinning dope and the bath causes the fiber to solidify.

Dry spinning. In this process, the polymer is again dissolved in an appropriate solvent and extruded through a spinneret. However, the liquid bath is replaced by a stream of warm gas (usually air) which evaporates the solvent and allows the polymer to solidify as a filament. Cellulose diacetate and triacetate are produced in this manner.

Melt spinning. Nylon, polyester, and the other thermoplastic fibers are produced by melt spinning. No solvents or reagents are required since the polymer can be melted without appreciable decomposition. Thus, the fluid consists of hot molten polymer which, upon extrusion into a stream of cold air, solidifies into a filament. Depending upon the end use, filaments may be produced in various sizes ranging from finer than a human hair to thick bristles for toothbrushes. They may also be produced with different cross-sectional shapes, such as round, lobed, square, or dogbone (**Fig. 3**).

Filament drawing. After extrusion, filaments are usually stretched (drawn). Drawing causes an increase in order (crystallinity) by extending the molecules of the fiber so that they pack more closely together, and orients the molecules along the longitudinal axis of the fiber. Higher orientation and increased crystallinity raise the strength of the fiber, decrease its stretch, and improve its elasticity.

Texturing. Often, the manufactured fibers are textured to improve their comfort properties. Fabrics made from smooth, straight filament yarns are not as comfortable as those made from yarns spun from the shorter natural fibers. Spun yarns which are more open and have a higher surface area than the filament yarns trap air between the fibers. This trapped air acts as an insulator which improves winter comfort by retaining body heat. The higher surface area facilitates the transport of body moisture within and through a fabric and improves summer comfort by reducing dampness and enhancing evaporative cooling. Because the early manufactured fibers were straight and smooth, yarns made from them provided little insulation and poor moisture transport, so that the early fabrics were cold and clammy in winter and hot and sticky in summer.

Texturing introduces irregularities (crimp) along the length of the filament and leads to bulkier filament yarns which are closer to spun yarns in their performance (**Fig. 4**). There are a number of methods used to introduce crimp into manufactured filament yarns. The most widely used is the false twist method, in which the yarn traveling vertically is looped over a horizontal pin which rotates at very high speed about its vertical axis, or in which the yarn is passed through a set of rotating disks. In either method, the heated yarn is caused to twist as it enters and to untwist as it exits the device.

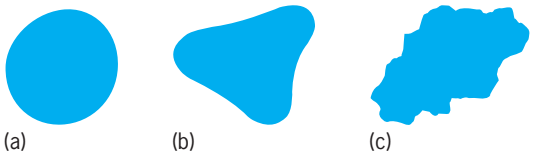

Fig. 3. Typical fiber cross sections. (*a*) Round. (*b*) Lobed. (*c*) Irregular. (*After B. F. Smith and I. Block, Textiles in Perspective, Prentice-Hall, 1982*)

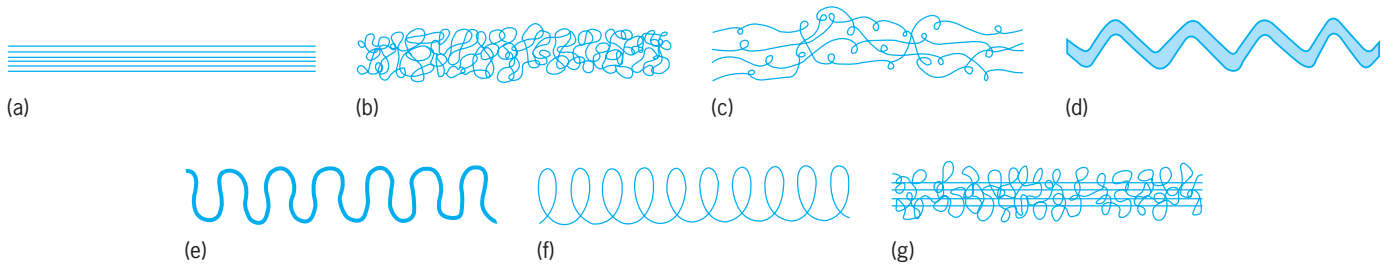

Fig. 4. Comparison of configurations of untextured and textured yarns. (*a*) Untextured filament yarn. (*b*) Entangled yarn. (*c*) Multifilament coil yarn. (*d*) Stuffer box crinkle yarn. (*e*) Knit-de-knit crinkle yarn. (*f*) Monofilament coil yarn. (*g*) Core-bulked yarn. (*After B. F. Smith and I. Block, Textiles in Perspective, Prentice Hall, 1982*)

Beyond the texturing apparatus, the yarn is cooled and the wavy crimp imparted to the individual filaments by the twisting while hot is permanently set. Other methods such as air jet, edge crimp, gear crimp, and stuffer box are also used, depending upon the nature of the fiber and the end use of the yarn (**Fig. 5**). Innovations in the spinning of filament fibers have led to filaments which are

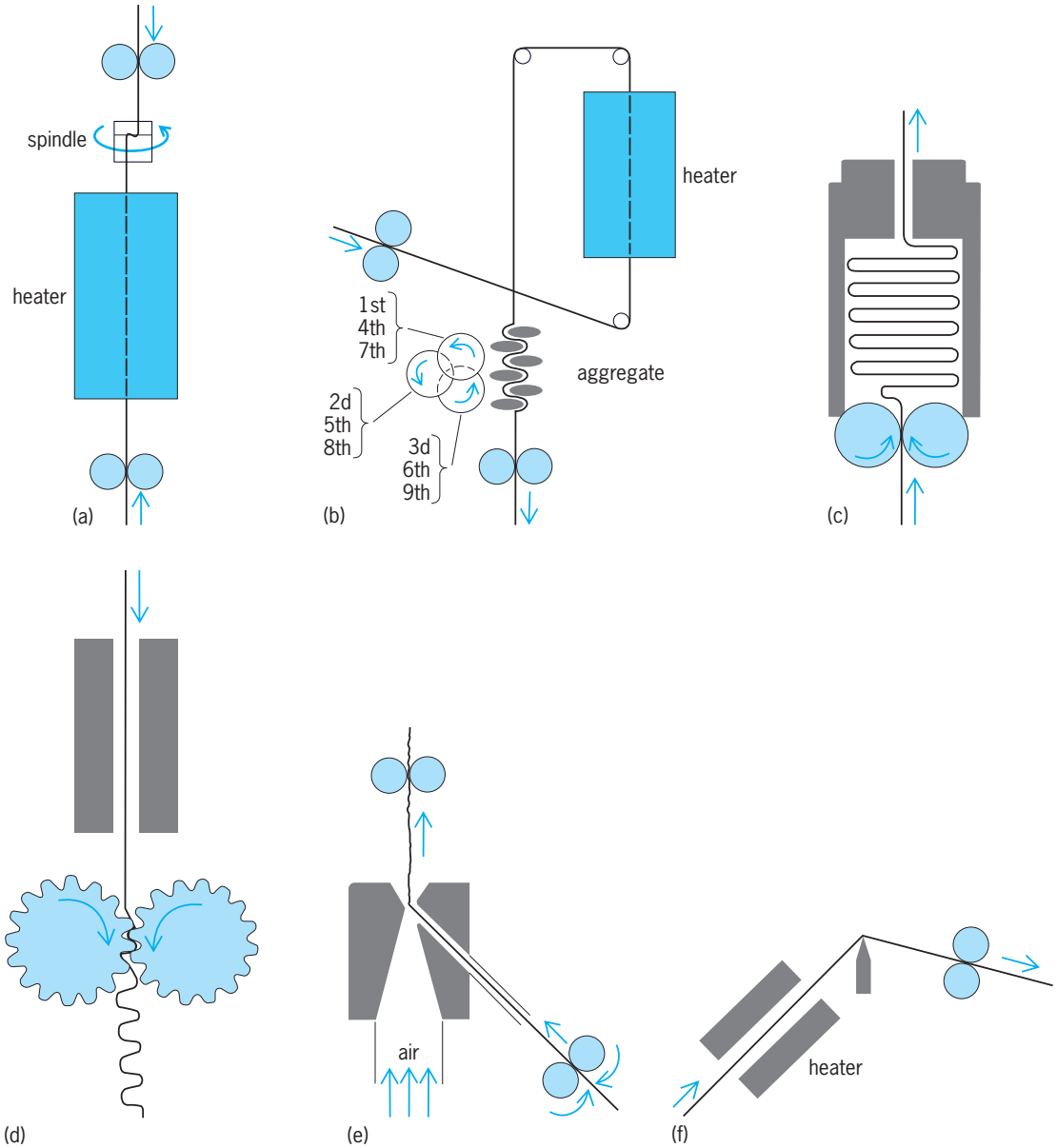

Fig. 5. Methods for crimping manufactured filament yarns. (*a*) Spindle false twisting. (*b*) Friction false twisting. (*c*) Stuffer box method. (*d*) Gear crimping. (*e*) Airjet method. (*f*) Edge crimp method. (*After B. F. Smith and I. Block, Textiles in Perspective, Prentice Hall, 1982*)

self-crimping. It is also possible to cut textured filament fibers to short (staple) length and to twist them into spun yarns, either as 100% manufactured fibers or in blends with natural fibers.

Cellulosic Fibers

These fibers are made from naturally occurring cellulose. They are found as (1) regenerated cellulose (rayon) in which the polymer is converted from the natural staple to manufactured filament while still maintaining its chemical structure; (2) cellulose diacetate (acetate) in which chemical modification converts about 80% of the hydroxyl (S—OH) groups on the cellulose to acetate (S—$OOCCH_3$) groups; and (3) cellulose triacetate in which at least 92% of the hydroxyl groups are converted to acetate. *See* CELLULOSE; WOOD CHEMICALS.

Rayon. Rayon is a manufactured fiber composed of regenerated cellulose in which substituents have replaced not more than 15% of the hydrogens of the hydroxyl groups. To spin these fibers from cotton linters or wood pulp, it is necessary to put the cellulose into fluid form without destroying its chain-like molecular structure. About 1885, Joseph Swan in England and Count Hilaire de Chardonnet in France conducted experiments leading to a process for extruding solutions of nitrocellulose and denitrating the filaments. In 1892, C. F. Cross and E. J. Bevan developed a better method using a solution of cellulose xanthate made from alkali cellulose and carbon disulfide. This is the basis of the viscose process used for the bulk of modern rayon production. Some is made by the regeneration of cuprammonium cellulose.

The viscose process consists of the following steps: steeping cellulose sheets in sodium hydroxide, shredding the sheets, aging the shredded materials to form alkali cellulose, treating it with carbon disulfide to form cellulose xanthate, ripening and filtering the spinning solution, extruding the solution through the small holes of the spinneret into an acid coagulating bath, stretching the filaments, and then washing, drying, and packaging.

Over the years, rayon has steadily improved in quality, versatility, and economic importance. Partly responsible for this expansion was the introduction of staple fiber that could be processed on the same types of equipment used for cotton and wool yarn spinning. For staple fiber, the number of holes in one spinneret is increased from the usual number (20–200) to as many as 20,000. Bundles of fibers from several spinnerets are combined into a rope (tow) which can be cut into short lengths and packaged in bulk. The second significant development was the attainment of high-strength yarns, which led to cords suitable for automobile and truck tires, and to textile yarns with sufficient wet strength for launderable fabrics.

Rayons now offer improved strength coupled with a higher level of comfort. Rayon fibers offer a high luster, a smooth silky hand, and excellent moisture absorption, although strength and abrasion resistance are only fair. The major disadvantages of rayon are its tendency to wrinkle and need for ironing.

Acetate and triacetate. These are manufactured fibers in which the fiber-forming substance is cellulose acetate. The term triacetate may be used as a generic description of the fiber when not less than 92% of the hydroxyl groups are acetylated.

Three hydroxyl groups on each glucose unit of the cellulose molecule can be esterified by acids, such as formic, acetic, propionic, and butyric. The resulting cellulose derivatives can be spun into fibers. Of these, cellulose acetate is most important.

To make acetate, the triacetate is formed first by treating cellulose with acetic anhydride. The triacetate is then converted back to diacetate by partial hydrolysis, precipitated with water, washed acid-free, dried, dissolved in a solvent (usually acetone), and then extruded by a dry-spinning process.

Cellulose triacetate (no hydroxyl groups) is less sensitive to hot water than the diacetate (one hydroxyl group), and garments made of it wrinkle less than those of regular acetate during laundering.

Both acetate and triacetate are less prone to wrinkling and have better wrinkle recovery than rayon. However, they are less durable and not as comfortable as rayon or the natural cellulosic fibers. Where good color, luster, and a silky hand are more important than durability, they have found a useful niche.

Petroleum-Based Fibers

In addition to its use as a fuel, petroleum is a major source of organic molecules which can be converted to synthetic polymers. The petroleum-based fibers constitute the largest class of the manufactured fibers because of the relative ease with which modern technology can prepare the polymers and convert them to fiber. In addition, since the chemical precursors, the polymers, and the fibers can be produced in highly integrated manufacturing plants insensitive to the vagaries of weather and less dependent upon human labor, yarns and cloth made from these fibers are usually less expensive than those produced from natural fibers or from the cellulosics. *See* PETROLEUM.

Nylon. Nylon is a manufactured fiber in which the fiber-forming substance is a long-chain synthetic polyamide having less than 85% of the amide linkages attached directly to two aromatic rings. These fibers are made from polymers based on the condensation polymerization of a dibasic acid with a diamine or of an amino acid with itself. The major one is nylon 6,6, which is made from a six-carbon acid (adipic acid) and a six-carbon diamine. In nylon 6, caprolactam molecules join together by self-condensation to form poly(caproamide). The fiber is produced by melt spinning of the molten polymer.

Nylon fibers have remarkable properties. After spinning, nylon can be drawn to about four times its original length. In addition to producing a finer fiber, the drawing process greatly increases the strength. Another unusual property of nylon fibers compared to earlier manufactured fibers is that nylon can be heat-set. That is, after manufacture of the cloth,

either by weaving or knitting, a heat treatment can be used to relax the residual stresses in the fibers and yarns. Heat setting provides an essentially permanent stability to the fabric so that shrinkage and wrinkling are almost eliminated.

Of major importance to consumers is the high strength, extreme abrasion resistance, excellent resistance to soiling and staining, and ease of care provided by nylon fibers. The major drawback is the slick, synthetic feel of the cloth. However, it is possible to alleviate this problem somewhat by texturing. Alone, or in combination with other fibers such as wool, nylon serves such diverse end uses as intimate apparel, daywear, home furnishings (particularly upholstery, draperies, and carpeting), tire cord, and industrial belting.

Because they are thermoplastic, nylon fibers can be readily spun with cross sections of many shapes. This property was originally exploited by the DuPont Company in its development of trilobal fibers which exhibited an interesting sparkle rivaled only by silk. Others have since developed fibers with Y-shaped, pentalobal and octagonal cross sections, each giving a subtly different combination of esthetics, hand, durability, and ease of care.

An interesting development in nylon technology was Qiana. This fiber, prepared from bis(*p*-aminocyclohexylmethane) and dodecandioic acid is made of molecules having both linear and cyclic components. The ring structures yield a polymer with the strength, durability, and ease of care of nylon, but a hand, luster, and drape which rivals silk. *See* POLYAMIDE RESINS.

Polyester. This is a manufactured fiber composed of at least 85% by weight of a substituted aromatic carboxylic acid, including but not restricted to substituted terephthalate units and parasubstituted hydroxybenzoate units. Although three types of polyester fibers have gained commercial status—poly(ethylene terephthalate) [PET], poly(1,4-cyclohexane dimethyl terephthalate) [PCDT], and poly(ethylene oxybenzoate) [PEB]—the first is by far the largest in sales volume in the United States. Manufactured in a manner similar to that of the nylons, PET is prepared by condensation polymerization of the dimethyl ester of terephthalic acid with ethylene glycol to produce the polymer, and is melt-spun to produce fibers. Like nylon, it is extruded with a variety of cross-sectional shapes and is usually textured prior to use.

Polyester fibers offer high strength and durability but suffer from a tendency to absorb oily stains and poor moisture absorption. Despite these problems, polyester has become the most ubiquitous of the manufactured fibers, appearing as textured filament in knitwear, and in spun blends along with either cotton or wool in both knit and woven apparel. In blends with natural fibers, polyester provides strength, abrasion resistance, and ease of care along with low cost. In home furnishings, polyester is greatly prized as a carpet fiber, but its tendency to absorb oils limits its utility for upholstery.

Acrylic. Acrylic is a manufactured fiber composed of at least 85% by weight of acrylonitrile units. Although the polymer of acrylonitrile had been known for many years, its successful use in fiber form depended upon the discovery of appropriate solvents for dry or wet spinning, and satisfactory methods to dye the fiber. The first commercial acrylic fiber was Orlon. Fibers of this type have attractive properties and a great deal of versatility.

Polyacrylonitrile (PAN) is prepared from the addition polymerization of acrylonitrile. However, the pure polymer is a brittle, glassy substance at room temperature and is not well suited for use as a textile fiber. It was found that by forming copolymers of acrylonitrile and other vinyl monomers the mechanical properties of the fiber would be greatly improved. The fiber is produced by either the wet-spun or the dry-spun process from solution in appropriate solvents. After drawing and texturing, the product is a soft, springy yarn with characteristics similar to wool. It takes dyes well, is extremely resistant to sunlight, is durable, and is relatively easy to maintain. It is widely used by itself and in blends with wool or nylon in sweaters, socks, robes, blankets, carpets, and other products where high bulk and good insulation are important. Because it does not absorb moisture, acrylic suffers from a high buildup of static electricity, particularly in dry weather, which can cause static cling and shock. *See* COPOLYMER; POLYACRYLATE RESIN.

Modacrylic. This is a manufactured fiber composed of less than 85% but at least 35% by weight of acrylonitrile units. The modacrylics are modified acrylic fibers. The addition of large amounts of comonomer with acrylonitrile allows the manufacturer to tailor the fiber to fit specific end-use requirements. These fibers are produced by dissolution of the copolymer in appropriate solvents followed by either wet or dry spinning. The most important uses of modacrylics have been in imitation hair for wigs and imitation furs and fleeces.

Olefin. Olefin is a manufactured fiber composed of at least 85% by weight of ethylene, propylene, or other olefin units. The most useful of the olefin fibers are those produced by the melt spinning of isotactic polypropylene. Because of their rather low melting temperature, these fibers are easily drawn and textured. The product yields fabrics with high bulk and good opacity. In addition, strength, abrasion resistance, and toughness are quite high. The major drawbacks are a slick waxy hand, very poor moisture absorption, and great difficulty in dyeing.

Because of the unappealing hand of cloth made from olefin fibers, they have not been accepted in the general wearing-apparel market. However, thermal undergarments of olefin fiber have been developed. Olefin's most important use is in upholstery and carpeting where its durability, resistance to staining, and ease of cleaning overcome its poor comfort properties and low melting temperature. *See* POLYOLEFIN RESINS.

Spandex. A manufactured fiber identified as spandex must be composed of at least 85% of a segmented polyurethane. In spandex, a long-chain polyester is combined with a short diisocyanate to produce a polymer containing long lengths of a relatively soft material joined by short lengths of a relatively hard material. This is the segmented polyurethane. In the relaxed state, the fibers are rather weak and easily stretched. However, as the fiber is elongated, the hard segments align and form strong, stiff crystals which prevent further extension. Thus, by appropriate combinations of hard and soft segments, these fibers can be engineered to provide any required degree of stretch and strength. The fibers are produced by the wet-spun process from appropriate solvents.

Spandex has found its greatest use as a replacement for rubber, both natural and synthetic. It is particularly useful in intimate apparel and sportswear, where its high elasticity give both figure control and a close fit without restricting the movements of the wearer. *See* RUBBER.

High-performance fibers. Advances in polymer and fiber technology have led to the development of fibers with exceptionally high temperature resistance and extremely high strength. These properties are desirable in applications such as upholstery and floor coverings in aircraft and other mass-transit vehicles, protective clothing for fire fighters and other emergency personnel, body armor for soldiers and police officers, tire cords, and industrial belting.

Aramid. Aramid is a manufactured fiber in which at least 85% of the amide linkages are attached directly to two aromatic rings. The aramids are similar to the nylons. However, the aromatic rings are more stable than linear arrangements of atoms and provide great strength and heat stability. The two most important aramid fibers are Nomex and Kevlar. Nomex is poly(*m*-phenylene isophthalamide), and Kevlar is poly(*p*-phenylene terephthalamide). Both fibers are highly crystalline and very strong—stronger than steel on a weight basis. Abrasion resistance is not very good, so these fibers are most important in specialty uses. Kevlar is also used as a tire cord.

Novoloid. This is a manufactured fiber containing at least 85% by weight of a cross-linked novolac. This fiber is claimed to be more comfortable than the aramids and although not as strong, has been used in clothing and upholstery where fire resistance is very important. Novoloid fiber was chosen for flight suits for astronauts.

Inorganic Fibers

Metallic fibers of silver and gold have been used for millennia to decorate fabrics. Today metallic fibers serve useful as well as decorative purposes. These fibers are formed by drawing metal wires through successively finer dies to achieve the desired diameter. Although gold and silver are the easiest to draw, modern methods have allowed the manufacture of steel, tantalum, and zirconium fibers. Because they are electrical conductors, metal fibers have been blended into fabrics to reduce the tendency to develop static electrical charges.

Glass

The art of heating sand and limestone to form a molten liquid has been known from the time of the ancient Egyptians. It was not until the 1930s, however, that glass fibers fine and flexible enough to be converted into textiles were produced. Glass fibers are prepared by the melt spinning of previously formed glass marbles, and the molten filaments are drawn down to very fine dimensions. It is the fineness of the fibers that gives them their flexibility and allows them to be used in textiles. Unfortunately, the fibers are so stiff that when broken they can penetrate human skin. Thus, they are not well suited to use in apparel or upholstery. Glass is widely used in curtains and drapery because of its total resistance to the degrading effects of sunlight, its low cost, and its flame resistance. It provides a nonrotting, nonsettling insulating material for homes and industrial uses. *See* GLASS.

Heterogeneous Fibers

These are fibers prepared from two or more chemical species which retain their identities after the fiber has been formed. These fibers may be spun with their components side by side, as a sheath of one material surrounding a core of another (sheath-core), or as a matrix of one material surrounding another (**Fig. 6**). The heterogeneous fibers can be tailored to give a variety of performance properties, but the most notable are those which, because of differential response to moisture and temperature, form self-crimping fibers that do not require texturing, and those which incorporate a conductive component to reduce static buildup inhydrophobic fibers.

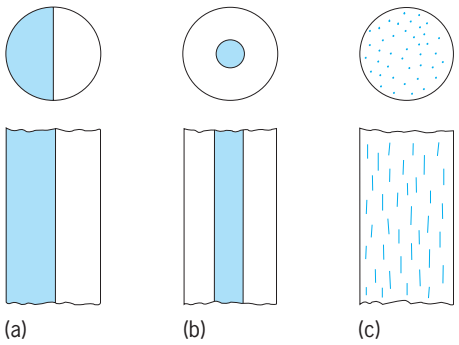

Fig. 6. Heterogeneous fiber types. (*a*) Side-by-side. (*b*) Sheath-core. (*c*) Matrix. (*After B. F. Smith and I. Block, Textiles in Perspective, Prentice Hall, 1982*)

Fiber Properties

These are the physical, mechanical, chemical, biological, and geometrical characteristics of fibers. Some of the more important ones are tensile strength, elongation at break, modulus of elasticity or stiffness,

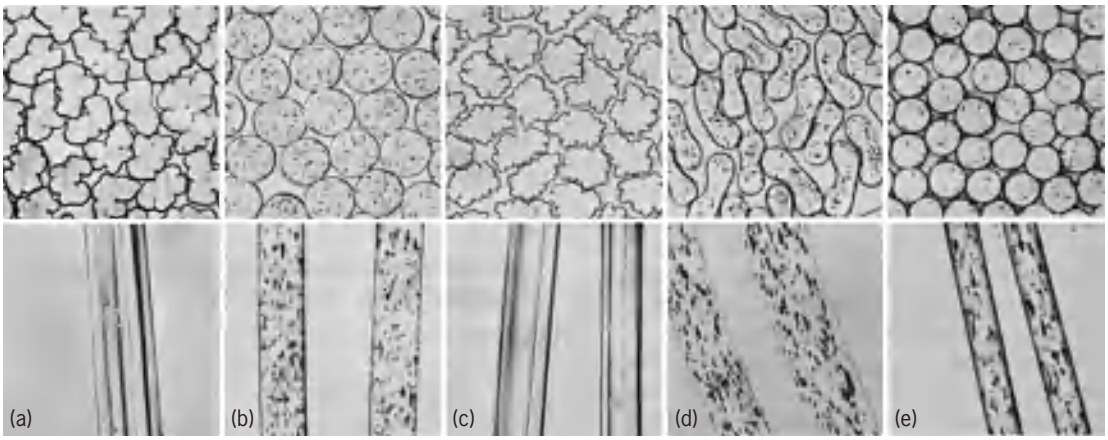

Fig. 7. Cross-sectional and longitudinal views of some typical fibers. (*a*) Acetate fiber. (*b*) Nylon fiber. (*c*) Rayon fiber. (*d*) Orlon fiber. (*e*) Dacron fiber. (*E. I. duPont de Nemours and Company*)

Characteristic attributes of manufactured fibers

Attribute	Properties required	Fiber recommended
Durability and abrasion resistance	High tenacity, high elongation at break	Nylon and polyester fibers
Resistance to soil particles	Smooth, round cross section	Melt-spun fibers
Quick drying	Low moisture content inside fiber	Hydrophobic fibers: polyesters, acrylics, nylon, triacetate
Winter warmth	Ability to entrap air spaces between resilient fibers; thick fabrics	Almost all textured synthetic fibers, the acrylics in particular
Summer comfort	Ability to wick perspiration from wet skin and dry quickly; thin but not-too-smooth fabrics	Acrylics and most rayon-synthetic blends
High resistance to light, solvents, acids, and bases	Compact chemical structure	Fluorocarbons and polyacrylics
Wrinkle resistance and wash-wear	High resilience both wet and dry	Polyesters outstanding; many others good, including acrylics, nylon, and triacetate
Freedom from static and hole melting	High moisture content	Rayon and rayon synthetic blends
Crease and pleat retention	Heat settability	Nylon and many other synthetic fibers

fatigue under repeated stress, resilience or ability to recover from deformation, moisture absorption and wettability, electrostatic properties, friction, color, luster, density, and resistance to light, heat, weathering, abrasion, laundering, mildew, insects, chemicals, and solvents; and finally a number of geometric features, such as diameter, cross-sectional shape, and crimp.

Figure 7 shows cross-sectional and longitudinal views of typical fibers. Such properties play an important part in determining whether or not the fiber can be made into a fabric that will be wrinkle-resistant, pleasing to the touch, comfortable, easy to clean, durable, and attractive in color, luster, drape, and general appearance. With a knowledge of the physical properties of the available fibers, the textile engineer can choose the best fiber or best blend of several fibers to fit the intended use. The final result, however, is also dependent upon the proper choice and control of additional factors such as the yarn and fabric structure, the weave pattern, and the finishing of the cloth.

In spite of the research efforts put into the development of improved fibers, no fiber is perfect for all uses. In many cases, perfect performance represents an incompatible and, therefore, impossible combination, exemplified by the desire for sheer hosiery with heavy-duty wear. Notable among existing deficiencies in some fibers both old and new are poor durability, soiling, static electricity, hole melting, yellowing with age, and poor shape retention. However, for each of these a manufactured fiber can be found which will solve the problem, as indicated in the **table**, which gives a partial listing of the major attributes of manufactured fibers.

All of the properties covered in the table deal with functional performance. An additional category of importance relates to the esthetic properties of fibers and fabrics—certain subtle qualities apparent to the hand, eye, and even to the ear (the whisper of silk or the rustle of a good taffeta). One of the reasons that cellulose acetate was able to compete successfully with its predecessor, viscose rayon, was its superior handle—a soft, luxurious feel—especially desirable in crepes and satins. As the result of much technical effort, manufactured fibers have been able to achieve those esthetic qualities referred to by textile experts in describing the characteristic handle of a cashmere sweater, a wool flannel, or a silk taffeta fabric. *See* NATURAL FIBER; POLYMER; TEXTILE. Ira Block

Bibliography. M. Grayson, *Encyclopedia of Textiles, Fibers, and Non-Woven Fabrics*, 1984; M. Lewin,

Handbook of Fiber Chemistry, 2d ed., 1998; Man-Made Fiber Producers Association, Inc., *Man-Made Fibers Guide*, 1988; H. L. Needles, *Textile Fibers, Dyes, Finishes, and Processes: A Concise Guide*, 1986; B. F. Smith and I. Block, *Textiles in Perspective*, 1996.

Manufacturing engineering

Engineering activities involved in the creation and operation of the technical and economic processes that convert raw materials, energy, and purchased items into components for sale to other manufacturers or into end products for sale to the public. Defined in this way, manufacturing engineering includes product design and manufacturing system design as well as operation of the factory. More specifically, manufacturing engineering involves the analysis and modification of product designs so as to assure manufacturability; the design, selection, specification, and optimization of the required equipment, tooling, processes, and operations; and the determination of other technical matters required to make a given product according to the desired volume, timetable, cost, quality level, and other specifications.

Manufacturing is one of the most complex of human group activities. It comprises hundreds or thousands of simultaneous and serial subactions, some of which occur in fractions of a second while others take hours, months, or even years to have full effect. These actions may be material, technical, informational, social, or economic.

Function. Historically, the function of manufacturing engineering was limited to developing and optimizing the production process. In brief, the manufacturing engineering function bridges the gap between the product design and full production. This can best be understood by considering the total process through which a designer's concept becomes a marketable product: (1) From the results of a needs analysis or market analysis, a product designer conceptualizes a product, and then drawings and one or more prototypes of this product are produced. (2) The finalized prototype and its part drawings are released to the group responsible for the manufacturing engineering function, which starts designing and building an economically justifiable process by which the product will be produced. (3) When the manufacturing process developed by manufacturing engineering has been thoroughly tried and proved workable, it is turned over to the production group, which assumes responsibility for product manufacture.

This method is known as the serial method of product production or the conventional product-production system design process (**Fig. 1**). A method developed after the serial method, known as concurrent design/concurrent engineering, has been practiced in Japan since the late 1950s and since around 1980 in the most progressive companies in the United States. Concurrent design/concurrent engineering means the consideration of issues of product design, manufacture, and use concurrently, but not necessarily simultaneously. As shown in the diagram in **Fig. 2**, a team is formed of all the necessary specialists who meet periodically to review the status of the design, report on the success or difficulties in satisfying their individual functional constraints, and collectively determine their next steps. An alternative arrangement has the team meeting continually. The starting point for a new design includes

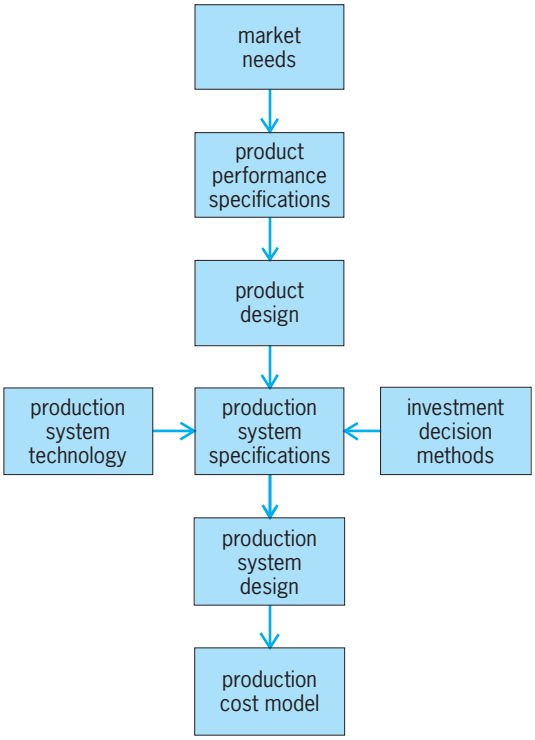

Fig. 1. Flow chart showing the components of the serial method of product production.

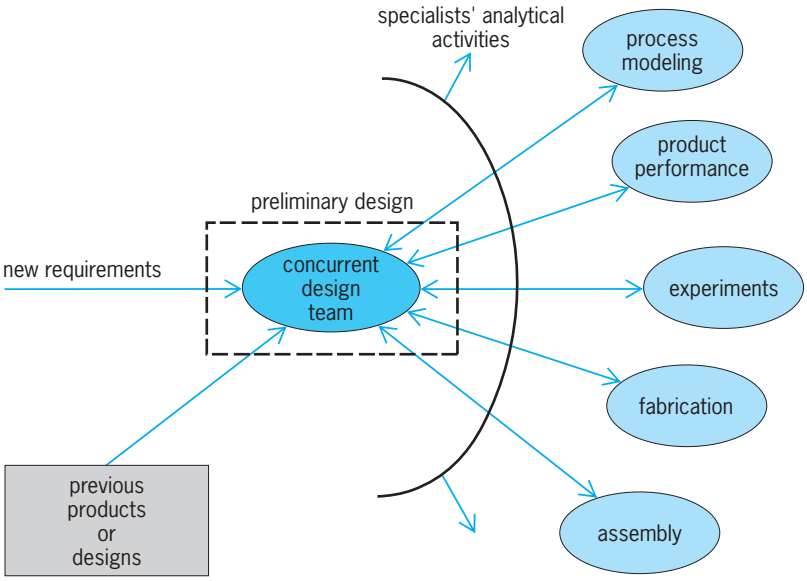

Fig. 2. Diagram showing the various elements of the concurrent design/concurrent engineering (CD/CE) process.

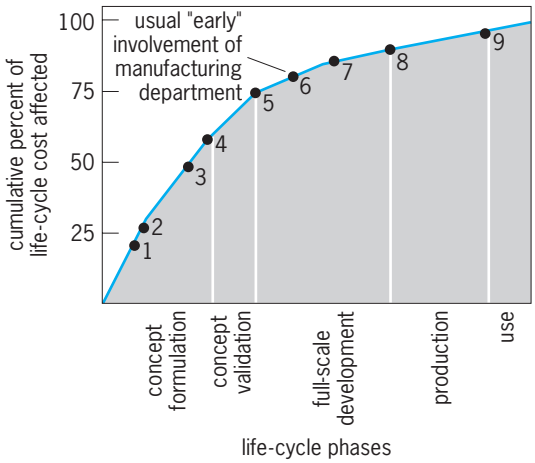

Fig. 3. Chart showing the cumulative percent of life-cycle cost of a product by individual phases of a product life cycle. Numbers represent the following phases: 1, define use patterns; 2, define alternatives; 3, develop alternatives; 4, freeze subsystems; 5, prove feasibility; 6, provide preliminary designs; 7, provide detail designs; 8, provide manufacturing plans; 9, product.

all the history of previous products, how they were designed and manufactured, as well as the institutions of people who did the work and who will do the next design. This concept is important, because, as shown in **Fig. 3**, by the time the concept validation phase is finished, 75% of the total life-cycle cost of a product has been determined. Thus any downstream function can optimize only the remaining 25%, no matter what it does or how much money it spends. *See* PROCESS ENGINEERING; PRODUCTION ENGINEERING.

Models. Manufacturing engineering involves technical as well as economic issues; these can be represented by models. Technical models usually comprise physical parameters, whereas economic models typically comprise time and money. In significant ways, the two types of models interact or may be linked, and in such cases they are known as economic-technical models.

A familiar technical model is that of metal cutting. The metal removal model represents the forces on a cutting tool while cutting a particular material. This model can be used to predict the quantity of heat delivered to the workpiece, the energy required for cutting, the resulting surface finish, and the risk of the tool breaking. A physical model of assembly relates the physical variables (size and shape of workpieces, friction during assembly, stiffness of tools and grippers, and relative error between parts) to the performance variables (likelihood of assembly failure and forces exerted on the parts by tools and each other).

A typical economic model of metal cutting predicts the tool's wear rate, given the same type of physical variables as the technical cutting model, but the quantities of interest are economic: how long the tool lasts, how much a new tool costs, how long it takes to change tools, and so on. The more intensive the cutting, the faster new workpieces are created. However, the work may not be cost-effective if it becomes necessary to stop often because of broken or worn tools.

At their most basic, technical models represent the physical response of machines and materials, perhaps in great detail, while economic models attempt to capture the economic consequences of those responses, typically with less detail. *See* MODEL THEORY.

Work content analysis. Economic models relate an amount of work to the amount of time or money or number of machines needed to do that work. The basic currency of such models is the work content inherent in a manufacturing task. In metal cutting or grinding, the content may be represented by the volume of material that must be removed; in assembly, it is the number of assembly operations needed to make one unit of the product. Since manufacturing may be regarded as a rate process governed by an output quantity in a given time, the work content is usually specified in productive units per unit time (assemblies per hour, for example). Determining the work content of a job is thus a basic step in manufacturing, whether the job is metal cutting, assembly, testing, or transport.

The objective of work content analysis, usually known as process planning, is to describe the elemental operations necessary to completely fabricate each part, to assemble parts into a whole or product, and to properly test the parts and assemblies to verify that the specified quality was achieved. A key aspect of process planning is that it results in an estimate of the time and the number of tools required to complete each fabrication operation. The amount of time that a specific individual piece of a product (piece-part) remains on a machine tool is known as the part cycle time. Part cycle times are used to determine the number of workstations that are necessary in the system in order to achieve a specific mix of parts or products. These times may also be used to determine the maximum allowable service or delivery rate and time available for the material-handling or parts supply system.

Process planning. The formulation of a process plan for a given part has seven aspects: (1) a thorough understanding of processing techniques, their yield and their reliability, precedences, and constraints (both economic as well as technical); (2) the material and tolerances of the part; (3) proper definition of machinability or process data; (4) proper workholding design of the stock or piece part during the fabrication process, a key consideration in generating piece parts of consistent quality; (5) proper tool selection for the task; (6) the capability of the equipment selected; and (7) personnel skills required and available.

Machinability or process data. The cycle times defined in the process plan for each operation are the sum of the cycle times for each tool used during that operation. The process planner determines the amount of time that a cutting tool is actually cutting the metal of a piece-part (in cut) based on the form of the raw material; dimension and surface finish of the final form of a specific feature (such as a hole or

a slot); tools necessary to generate that feature; and cutting speed and its relationship to the feed rate of the material being machined, based on the tool selected and the material to be cut. A principal cause of poor performance in existing systems is the incorrect choice of cutting speed and feed rate for the parts in the system.

Each fabrication process has similar issues and concerns, whether the process is metal forming, composite layup, near-net shape-forming techniques or others.

Work-holding alternatives. Proper conception and design of work-holding devices also known as fixtures are critical to the successful fabrication and assembly of any product. The system designer's objective with respect to fixture definition is to minimize the number of fixtures required throughout the system for fabrication of each component, while assuring proper control and orientation of the component. Adequate support and clamping force must be designed into every fixture. Additionally, the fixture's design should prevent the loading of incorrect parts into the fixture as well as improper orientation of the correct part in the fixture.

Tool selection. The process planner attempts to satisfy the following objectives when selecting tools for processing parts in a fabrication system: (1) accurately generate each feature as quickly as possible; (2) provide sufficient tool longevity to avoid an unnecessary amount of tool replacement (and machine downtime waiting for tools); (3) minimize the number of tools per part; and (4) maximize the commonality of tools among parts.

Cutting time and tool life are inversely related; the faster a tool cuts, the shorter its useful life (between resharpenings). The process planner must balance the speed at which parts are produced with the downtime on the equipment required to replenish tools because of wear. Almost all equipment in a fabrication system will change tools automatically to minimize the direct labor necessary to produce a part. Tool-changing machines have a limited capability for storing tools, often only 40 to 60 tools.

The fewer the tools required by each part, the greater the number of parts a machine can accommodate between tool exchanges. Using the same tools, where appropriate, on as many parts as possible also addresses this problem; as more parts share a greater number of tools, the total number of unique tools needed in the system is reduced. This allows more parts to be produced with less interruption of productive machine time. Additional benefits derived from tool standardization and consolidation include reduced tool-storage space and inventory-tracking requirements, higher-quantity purchases, and less potential for tool unavailability.

Scheduling and routing of parts are also simplified. A part does not have to visit a particular machine just because a unique tool is stored there. Instead, it may visit any available machine that has the common tools it needs.

Significant gains in production efficiency have been made since the mid-1970s. Ensembles of computer-controlled machines, known as flexible manufacturing systems, permit their individual machines to cut metal 65–85% of the time. Manually operated machine tools are much less efficient. *See* ASSEMBLY MACHINES; FLEXIBLE MANUFACTURING SYSTEM; COMPUTER NUMERICAL CONTROL.

Computer-aided process planning. Process planning aids based on computer programs that incorporate a type of spread sheet can be used to reduce significantly the time required to generate individual process plans. For example, systems have been developed that calculate the cycle time for each part as well as the number of tools used per part, the number of unique tools per part set, and the total time for cutting operations per tool type. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Materials-handling equipment selection. In parallel with the definition of process equipment, the manufacturing system designer must determine the most appropriate materials-handling techniques for the transfer of parts from machine to machine of each family of parts. During the manufacture of pieces, the parts are organized by the type of feature desired. The parts are then grouped and manufactured as a family, a method known as group technology. This includes the selection of storage devices appropriate for raw material, work in process, and finished-goods inventory as well as fixtures, gages, and tooling. Materials-handling equipment may be very different for each family, depending on part size and weight, aggregate production volume, part quality considerations during transfer, and ease of loading and unloading candidate machines. Different materials-handling approaches may also be appropriate within individual fabrication systems. *See* MATERIALS-HANDLING EQUIPMENT.

Process quality assurance. A quality assurance philosophy must be developed that emphasizes process control as the means to assure part conformance rather than emphasizing the detection of part nonconformance as a means of detecting an out-of-control process. The success of any fabrication process is based on rigid work-holding devices that are accurately referenced to the machine, accurate tool sizing, and tool position control. The basic way to determine if these three factors are functioning together acceptably is to measure a feature they produce as they produce it or as soon as possible after that feature is machined. The primary objective of this measurement is to determine that the combination is working within acceptable limits (statistical process control); the fact that the part feature is in conformance to print (conformance to tolerance specified on the print/drawing) is a by-product of a process that is in control. *See* QUALITY CONTROL.

Human resource considerations. In a modern manufacturing environment an organization's strategies for highly automated systems and the role for workers in these systems are generally based on one of two distinct philosophical approaches. One approach views workers within the plant as the greatest source of error. This approach uses computer-integrated

manufacturing technology to reduce the workers' influence on the manufacturing process. The second approach uses computer-integrated manufacturing technology to help the workers make the best product possible. It implies that workers use the technology to control variance, detect and correct error, and adapt to a changing marketplace. *See* COMPUTER-INTEGRATED MANUFACTURING.

The best approach utilizes the attributes of employees in the factory to produce products in response to customer demand. This viewpoint enables the employees to exert some control over the system, rather than simply serving it. The employees can then use the system as a tool to achieve production goals.

Prospects. As technology and automation have advanced, it has become necessary for manufacturing engineers to gain a much broader perspective. They must be able to function in an integrated activity involving product design, product manufacture, and product use. They also have to consider how the product will be destroyed as well as the efficient recovery of the materials used in its manufacture.

Manufacturing engineers must also be able to use an increasing array of computerized support tools, ranging from process planning and monitoring to total factory simulation—and in some cases, including models of the total enterprise. *See* SIMULATION.

Consideration is being given to redefine and enlarge the role of manufacturing engineers. There is a tendency to move away from the highly segmented array of specialists that requires industrial engineers to detail tasks, plant engineers to design facilities, engineers to design materials-handling systems, and so forth. J. L. Nevins

Bibliography. B. H. Amstead et al., *Manufacturing Processes*, 8th ed., 1987; G. Boothroyd, P. Dewhurst, and W. Knight, *Product Design for Manufacture and Assembly*, 2d ed., 2001; Committee on the Effective Implementation of Advanced Manufacturing Technology of the Manufacturing Studies Board, *Human Resource Practices for Implementing Advanced Manufacturing Technology*, Commission on Engineering and Technical Systems, 1986; C. S. Draper Laboratory, *Flexible Manufacturing System Handbook*, CSDL Rep. R-1599, 1984; D. T. Koening, *Manufacturing Engineering: Principles for Optimization*, 2d ed., 1994; G. Taguchi et al., *Taguchi's Quality Engineering Handbook*, 2004.

Map design

In contrast to the creative freedom of painting or graphic illustration, map design must observe rigid limitations because the locations of geographic features such as rivers, national boundaries, and coastlines are tied to reality. Map design is defined as the systematic process of arranging and assigning meaning to elements on a map for the purpose of communicating geographic knowledge in a pleasing format. During the twentieth century, standardized symbols and designs evolved for reference maps such as topographic maps. However, nonstandard maps with unique purposes must be designed within a process that encompasses creativity, experimentation, and evaluation. Careful design is crucial to map effectiveness to avoid distorted or inaccurately represented information.

Design process. The first design stage involves determining the type of map to be created for the problem at hand. Decisions must be made about the map's spatial format in terms of size and shape, the basic layout, and the data to be represented. In this step, the experience, cultural background, and educational attainment of the intended audience must be considered. The second stage involves the exploration of preliminary ideas through the manipulation of design parameters such as symbols, color, typography, and line weight. In the third step, alternatives are evaluated and may be accepted or rejected. Under some circumstances, prototype maps may be developed for sample readers as a means of evaluating design scenarios. The last step involves the selection of a final design.

Design considerations include the selection of scale (the relationship between the mapping media format and the area being mapped), symbols to represent geographic features, the system of projection (the method used to translate Earth coordinates to flat media), titles, legends, text, borders, and credits. The process of arranging each map element is referred to as map composition. Success in map composition is achieved when design principles are applied to create a pleasing image with a high degree of information content and readability.

Design principles. Map design principles have emerged from the disciplines of cartography, graphic arts, and psychology. They include balance, the relationship between figure and ground, visual acuity, hierarchy, and contrast.

Balance. Balance refers to arranging major map elements within the map frame in a visually appealing manner. A map is off balance if all elements appear on one side of its frame, leaving a large amount of open space. The visual balance is closely tied to the map's optical center, located slightly above the center of the rectangle formed by the map border.

Figure and ground relationships. As discovered through the work of psychologists, the mind organizes a map's visual area into two separate perceptual fields—figures that the eye focuses on, and the background. The differentiation of land and water through the application of color or shading is a common method for helping to establish these two fields (**Fig. 1**).

Visual acuity. Map elements must be readable by most map users. Visual acuity refers to the threshold minimal size of objects that can be correctly identified. For example, what is the smallest size of lettering that can be interpreted on a map designed for a viewing distance of 18 in. (46 cm).

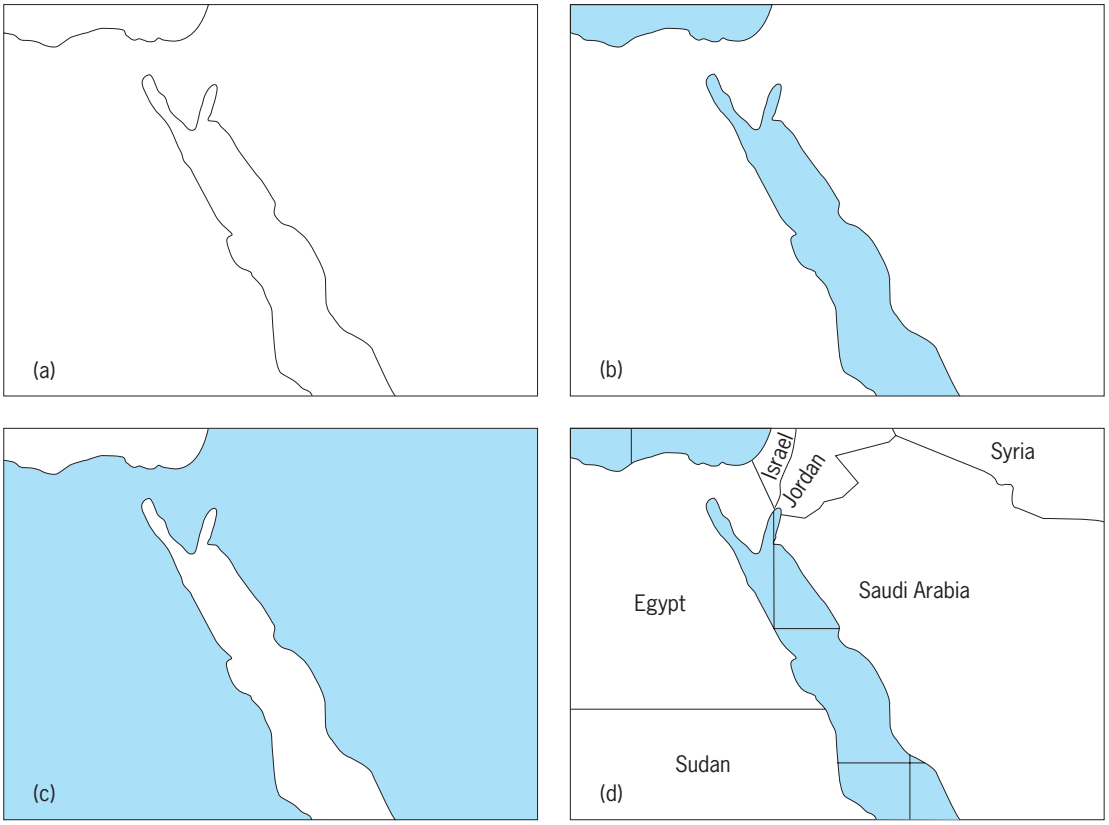

Fig. 1. Figure and ground relationship. (*a*) Map without shading does not assist the user in isolating the subject from the background. (*b*, *c*) Map shaded in two different ways improves definition. (*d*) Addition of labels to show place and of the graticule, clearly differentiates subject and background.

Hierarchy. Maps should exhibit a clear hierarchy of importance for each type of element, the most important elements having the greatest visual prominence. For example, the importance of places can be established using variation in sizes of point symbols or place names.

Contrast. An absence of contrast can result in a map element being lost in its surroundings. Contrast refers to variation in color, line weight, texture, or detail.

Computers. Computers now duplicate all capabilities of manual map design. Design iterations can be explored and evaluated faster and with lower cost, compared to manual drafting methods (**Fig. 2**). The digital environment facilitates the independent storage of map elements that can be combined to form composite images. Automated processes have also impacted map reproduction; it is now possible for the cartographer to prepare a map image for printing without a photographic step.

Research in cartographic design has recently focused on the application of artificial intelligence to the automation of text placement. Other computer innovations such as interactive mapping allow a map user to act as map creator in exploring geographic relationships by selecting and tailoring data sets available through software or the World Wide Web. In addition to its benefits to the design process, mapping software has brought challenges to the mapping

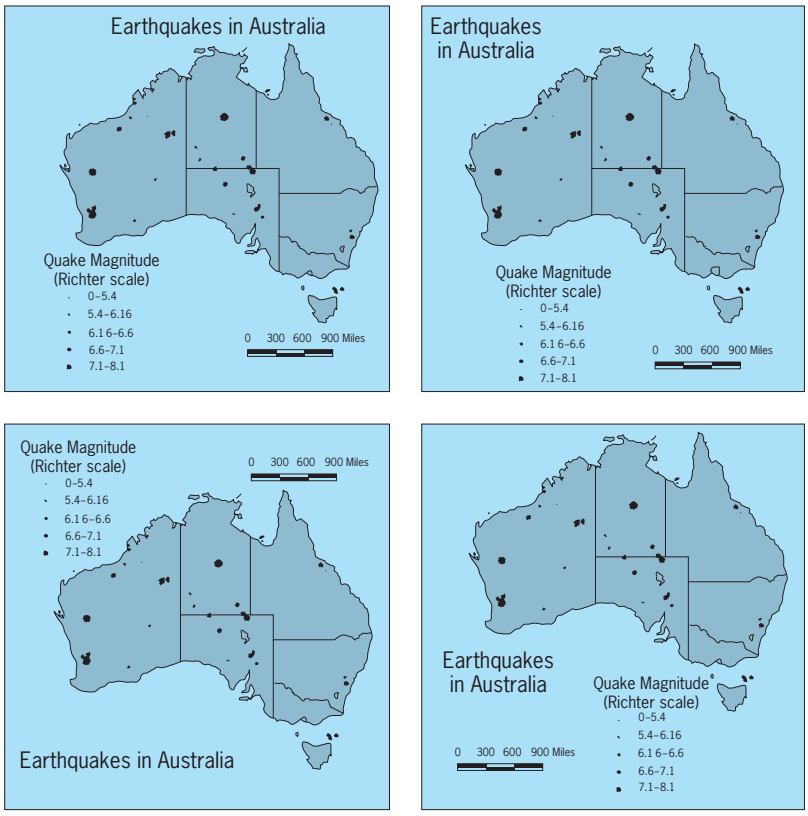

Fig. 2. Design iterations created by a computer.

sciences, including the proliferation of poorly designed maps constructed by persons untrained in cartography. *See* CARTOGRAPHY; COMPUTER GRAPHICS; MAP PROJECTIONS; MAP REPRODUCTION.

Thomas A. Wikle

Bibliography. A. H. Robinson et al., *Elements of Cartography*, 6th ed., Wiley, 1995; C. H. Wood and C. P. Keller, Design: Its place in cartography, *Cartographic Design: Theoretical and Practical Perspectives*, Wiley, 1996.

Map projections

Systematic methods of transforming the spherical representation of parallels, meridians, and geographic features of the Earth's surface to a nonspherical surface, usually a plane. Map projections have been of concern to cartographers, mathematicians, and geographers for centuries because globes and curved-surface reproductions of the Earth are cumbersome, expensive, and difficult to use for making measurements. Although the term "projection" implies that transformation is accomplished by projecting surface features of a sphere to a flat piece of paper using a light source, most projections are devised mathematically and are drawn with computer assistance. The task can be complex because the sphere and plane are not applicable surfaces. As a result, each of the infinite number of possible projections deforms the geometric relationships among the points on a sphere in some way, with directions, distances, areas, and angular relationships on the Earth never being completely recreated on a flat map.

Although the most commonly used map projections are accessible through mapping software, their fundamental attributes are best appreciated by considering the problem of the transformation of a spherical surface to a plane.

Deformation patterns. It is impossible to transfer spherical coordinates to a flat surface without distortion caused by compression, tearing, or shearing of the surface. Conceptually, the transformation may be accomplished in two ways: (1) by geometric transfer to some other surface, such as a tangent or intersecting cylinder, cone, or plane, which can then be developed, that is, cut apart and laid out flat; or (2) by direct mathematical transfer to a plane of the directions and distances among points on the sphere. Patterns of deformation can be evaluated by looking at different projection families. Whether a projection is geometrically or mathematically derived, if its pattern of scale variation is like that which results from geometric transfer, it is classed as cylindrical, conic, or in the case of a plane, azimuthal or zenithal. *See* CARTOGRAPHY; TERRESTRIAL COORDINATE SYSTEM.

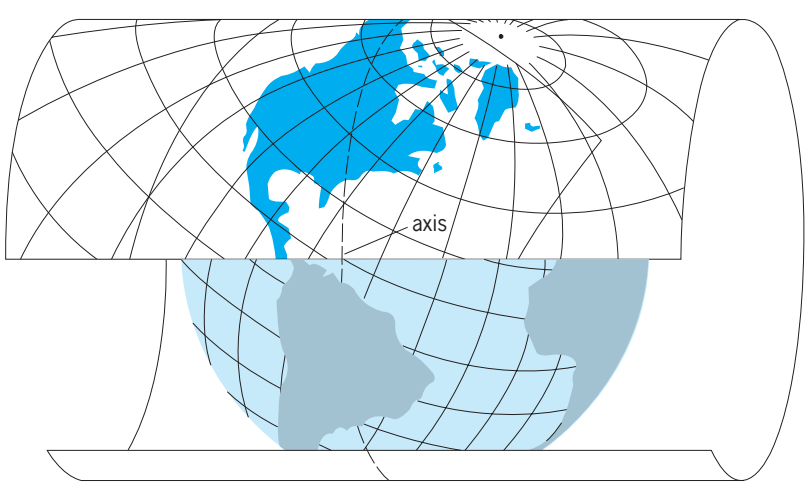

Fig. 1. Oblique Mercator projection on a cylinder. (*After American Oxford Atlas, Oxford, 1951*)

Cylindrical projections. These result from symmetrical transfer of the spherical surface to a tangent or intersecting cylinder. True or correct scale can be obtained along the great circle of tangency or the two homothetic small circles of intersection. If the axis of the cylinder is made parallel to the axis of the Earth, the parallels and meridians appear as perpendicular lines. Points on the Earth equally distant from the tangent great circle (Equator) or small circles of intersection (parallels equally spaced on either side of the Equator) have equal scale departure. The pattern of deformation therefore parallel the parallels, as change in scale occurs in a direction perpendicular to the parallels. A cylinder turned 90° with respect to the Earth's axis creates a transverse projection with a pattern of deformation that is symmetric with respect to a great circle through the Poles. Transverse projections based on the Universal Transverse Mercator grid system are commonly used to represent satellite images, topographic maps, and other digital databases requiring high levels of precision. If the turn of the cylinder is less than 90°, an oblique projection results (**Fig. 1**). All cylindrical projections, whether geometrically or mathematically derived, have similar patterns of deformation. *See* GREAT CIRCLE, TERRESTRIAL.

Conic projections. Transfer to a tangent or intersecting cone is the basis of conic projections. For these projections, true scale can be found along one or two small circles in the same hemisphere. Conic projections are usually arranged with the axis of the cone parallel to the Earth's axis. Consequently, meridians appear as radiating straight lines and parallels as concentric angles. Conical patterns of deformation parallel the parallels; that is, scale departure is uniform along any parallel (**Fig. 2**). Several important conical projections are not true conics in that their derivation either is based upon more than one cone (polyconic) or is based upon one cone with a subsequent rearrangement of scale variation. Because conic projections can be designed to have low levels of distortion in the midlatitudes, they are often preferred for representing countries such as the United States.

Azimuthal projections. These result from the transfer to a tangent or intersecting plane established perpendicular to a right line passing through the center of the Earth. All geometrically developed azimuthal projections are transferred from some point on this line. Points on the Earth equidistant from the point of tangency or the center of the circle of intersection

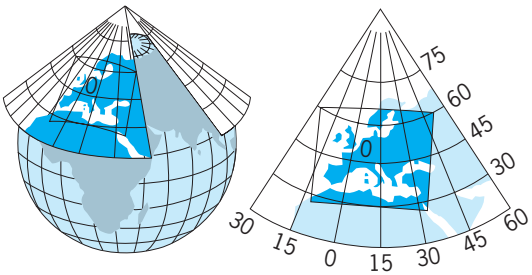

Fig. 2. Conical projection, based on origin 45°N, 10°E. (*After American Oxford Atlas, Oxford, 1951*)

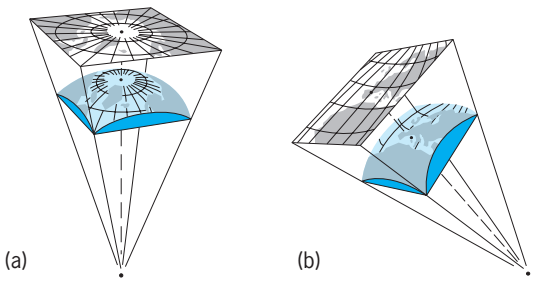

Fig. 3. Azimuthal projections with tangent planes have been raised above the sphere surface. (*a*) Polar azimuthal. (*b*) Oblique azimuthal. (*After American Oxford Atlas, Oxford, 1951*)

have equal scale departure. Hence the pattern of deformation is circular and concentric to the Earth's center. All azimuthal projections, whether geometrically or mathematically derived, have two aspects in common: (1) all great circles that pass through the center of the projection appear as straight lines; and (2) all azimuths from the center are truly displayed (**Fig. 3**).

Nongeometric projections show no consistency in their patterns of deformation; nevertheless, many are useful. One such group, called pseudocylindrical projections, is derived by mathematically arranging the spherical surface within a shape bordered by a series of specified smooth curves, such as ellipses, parabolas, or sine or quartic curves (**Fig. 4**).

Properties. The retention anywhere on the projection of some specific geometric quality of the sphere, such as uniformity of compass rose, or the creation of some useful attribute, such as straight rhumb lines, is called a property. Projections offer possibilities for maintaining single properties or combinations of properties. Ultimately, a map's intended use should determine the properties to be retained in the transformation process. All properties are the result of arranging the magnitudes and directions of scale variation.

Conformality. Conformality (orthomorphism) is the retention of angular relationships at each point, resulting in the preservation of the shapes of small features. It is obtained by arranging the scale so that at each point the scale is uniform in all directions. Since the spheroid and plane are not applicable, scale must change from point to point. Thus, all conformal projections deform the relative sizes of Earth areas (**Fig. 5**). Conformal projections are most widely used for navigational, engineering, and topographic maps, since observable angles may be measured on the map with a protractor.

Equivalence. Equivalence (equal-area) is the retention of the relative sizes of areas. It is obtained by arranging the scale at each point so that the product of the sides of an infinitesimal Earth rectangle projected anywhere on the map is the same. Since the two surfaces are nonapplicable, the scale at a point must be different in different directions. Therefore, the properties of conformality and equivalence are mutually exclusive. Equivalent projections are most widely used for mapping geographic, economic, and similar types of data whenever the areal extent of the phenomena is important. The extent of Earth

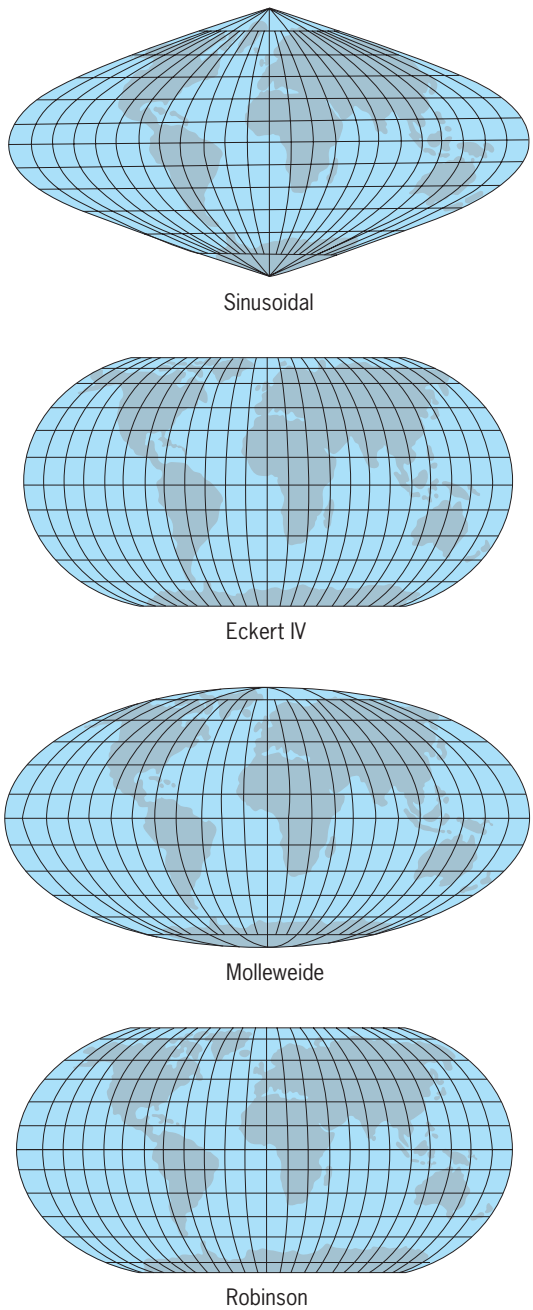

Fig. 4. Four pseudocylindrical projections. (*University of Wisconsin-Madison Cartographic Laboratory*)

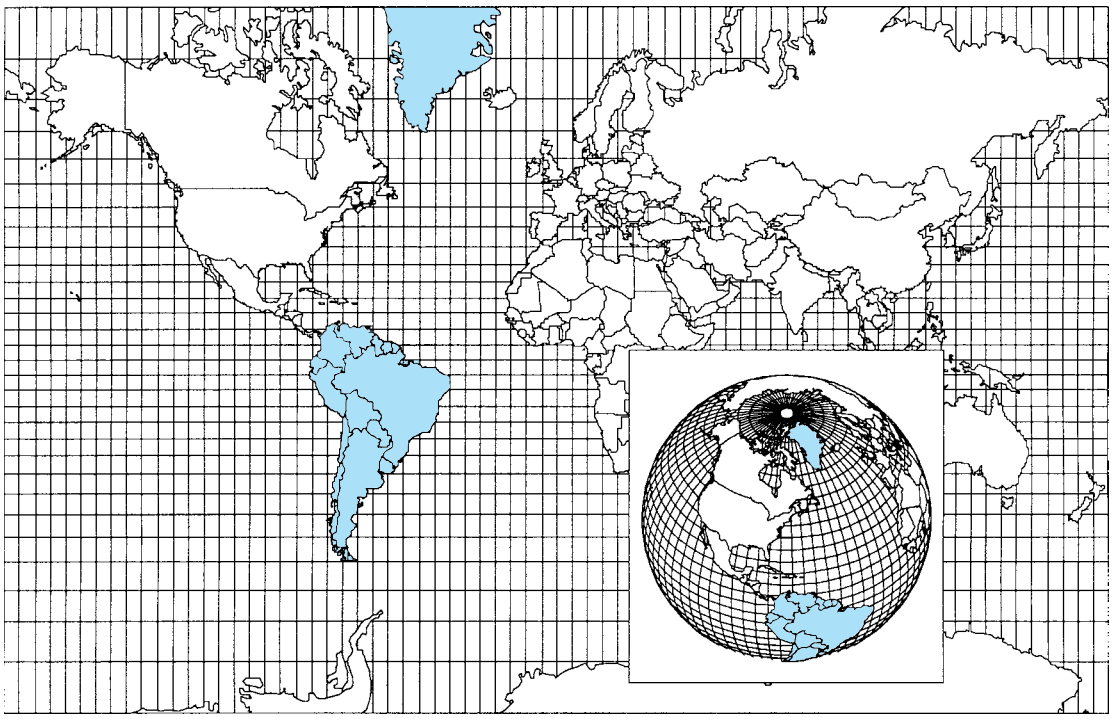

Fig. 5. The property of equivalence is not maintained on the Mercator projection, making Greenland and South America appear similar in size. The inset map shows that South America is actually about 15 times larger than Greenland.

area within administrative subdivisions can best be derived by measurement from equivalent projections.

Azimuthality. Azimuthality is the retention of azimuths (correct directional relations along a great circle bearing) from one point to another. Within limited areas, this, like all properties, can be approached on many projections but can be precisely retained from only one or two points. The centers of all projections theoretically or actually constructed on a tangent plane have this property. Azimuthality is useful in graphic aids to communication and for observations dealing with electronic impulses that travel along great circles.

Significant lines. Significant reference lines on the Earth include great circles, rhumb lines, and small circles, which are utilized in navigation, communication, and area and space analysis. By arranging the scale, these may be made to appear as straight or consistently curved lines on a map. Of special significance are great circles and rhumb lines in air and sea navigation. Because they represent the shortest distance between two locations on the Earth's surface, aircraft follow great circle routes. The practical requirement of guiding movement by compass angle requires that headings be along rhumb lines (loxodromes). Scale on projections can be arranged so that entire families of these kinds of critical lines may be made to appear as straight lines. Similarly, directional concepts, such as a specific cardinal direction, can be made parallel anywhere on the map. Any circle on the Earth can be made to appear as a circle on a map. Many other similar attributes are possible with map projections.

Examples. Although an infinite number of different projections is possible, only a relatively few are widely used. There is no best projection for any specific mapping problem, and in most cases several different projections can be applied to any given project.

Mercator. The Mercator projection is a mathematically derived, cylindrical-type, conformal projection that is conventionally based on the Equator as the tangent great circle. The projection was developed in 1569 by the Flemish cartographer Gerhardus Mercator to represent his world map. Linear scale on the Mercator increases in a direction perpendicular to the Equator as the secant of the latitude on a sphere. All rhumb lines are straight lines on the conventional form, making the Mercator the most widely used projection for nautical navigation. The transverse Mercator (and similar variants for the spheroid, for example, Gauss-Krüger) employs a meridian pair as the tangent great circle. Rhumb lines are not straight lines on the transverse form.

Sinusoidal. The sinusoidal, also called the Sanson-Flamsteed or Sanson-Mercator projection, is an equal-area projection with a straight central meridian and straight, equally spaced parallels. An area of least distortion is found along the Equator and central meridian, making it a popular projection for representing Africa and South America. Since the projection comes to a point at the Poles, shapes are distorted in the high latitudes.

Mollweide. An equal-area projection with straight parallels, the Mollweide resembles the sinusoidal but forms an ellipse without pointed poles. In contrast to the sinusoidal, low distortion is found along the 40th

parallels north and south latitude rather than at the center. When centered on the prime meridian, the Mollweide is a popular projection for representing continental Europe.

Gnomonic. The gnomonic (azimuthal) projection is obtained by geometric transfer from the center of a sphere to a tangent plane. Scale is greatly exaggerated away from its center, making the projection impractical for representing more than 45° of latitude in a single hemisphere. Because all arcs of great circles appear as straight lines, this projection is popular for navigation.

Azimuthal equidistant. The azimuthal equidistant projection is formed by straight meridians radiating from a single point of plane tangency with the surface. Parallels form concentric circles and, when measured from the point of tangency, all points are true in distance and global direction. Although commonly used to represent Arctic and Antarctic regions, the projection can be centered on any location, making it ideal for plotting seismic or radio waves.

Stereographic. The stereographic (azimuthal and conformal) projection is derived by geometric transfer to a tangent plane from a point opposite the point of tangency. All circles on the Earth appear as circles on the map. It is widely used as a framework for topographic maps of small areas and for navigation and grid systems in polar areas.

Lambert conformal conic. The Lambert conformal conic with two true-scale (standard) parallels is widely used for aeronautical and meteorological charts, as the framework for topographic series, and for rectangular grid systems of smaller areas. Scale increases away from the standard parallels, and great circles are nearly straight lines.

Albers equal-area conic. The Albers equal-area conic with two standard parallels is similar in appearance to the Lambert. The projection is commonly used to show middle latitudes for administrative and area research maps, especially for the United States, Europe, and Russia.

Polyconic. Although the polyconic (the ordinary, or American, polyconic) is neither equal-area or conformal; each parallel is true scale. It has been widely used for large-scale topographic maps, each sheet being individually developed and essentially error-free. Sheets fit together north-south but not east-west and cannot be mosaicked in all directions without gaps. The U.S. Geological Survey ceased using the polyconic for topographic maps in favor of either the transverse Mercator or the Lambert conformal conic. The modified polyconic employs straight meridians and distributes scale error over the map sheet.

Goode's homolosine. Interrupted projections, such as the Goode's homolosine, make it possible to maintain a degree of equivalence and conformality on a single world map. The projection is created by inserting breaks over land or ocean areas and is homolographic from 40° north to the North Pole and from 40° south to the South Pole. Elsewhere, from 40° north to 40° south, the projection is sinusoidal. Goode's homolosine is a popular projection for use in atlases and textbooks.

Most other forms of projection are used for small-scale, general or thematic maps. Most widely used for these are equivalent or near-equivalent varieties. The display of imagery from a satellite orbiting about the rotating Earth has required complex mathematical developments in map projections.

Computers. Computers have become extremely valuable tools for the assisting in selecting the most appropriate projection to fit the unique constraints of each mapping problem. Mapping software allows the cartographer to experiment with alternative projections or to make parameter adjustments after a projection has been selected.

Arthur H. Robinson; Thomas A. Wikle

Bibliography. B. D. Dent, *Cartography: Thematic Map Design*, 4th ed., 1996; D. H. Maling, *Coordinate Systems and Map Projections*, 2d ed., 1993; F. Pearson II, *Map Projection Methods*, 1984; A. H. Robinson, *Elements of Cartography*, 6th ed., 1995; J. P. Snyder and P. M. Voxland, *Album of Map Projections*, 1989.

Map reproduction

The processes of producing multiple copies of maps or map drawings. Reproductions are obtained by a number of procedures, including inked-plate printing, photographic processes, office copying equipment, and computers. Choosing a process depends on the use of the map, quality and quantity desired, cost, and availability of equipment. The reproduction process governs the techniques and media for preparing the original artwork. Some processes serve as intermediate stages in map reproduction. Finished reproductions can be on any material, but film and paper are most commonly used and are single color or multicolor.

Inked-plate processes. Of the three major kinds of printing—relief plate (letterpress), intaglio (engraving or gravure), and planographic (lithography or offset)—the planographic process is most commonly used for map reproduction. Before the development of photolithography, the majority of maps were made from engravings. Today, this method is not used, although photogravure maps are made in the printing of newspapers or reproductions of antique maps. Large printing runs of maps may call for use of the letterpress process.

With multicolor maps, registry of colors and images is of great importance both in the making of the artwork and in the printing. Each color is represented by one or more separate artwork overlays which must be matched in printing. Information on each overlay can be depicted by drawing with India ink and adding stick-on symbolization and lettering. The artwork is photographed on film, which is used to make the printing plate. To help reduce the steps leading to the printing plate and have better registry control, another method of preparing the artwork is used, called scribing. Overlays consist of clear, stable plastic coated with opaque material which is engraved (scribed) to establish a clear image. The

scribed overlay can then be used directly as a negative for plate making, bypassing the intermediate photographic process. *See* PRINTING.

Photographic processes. For photographic work, originals can be on transparent or opaque bases and the reproductions may be enlarged, same-size, or reduced copies of the original. Procedures involve a camera, exposure to light-sensitive film or paper as a negative or positive image, and wet developing. Silver salts provide the more common light-sensitive emulsion, although some other processes use different emulsions and light requirements. Reproductions are either single-color or multicolor prints or film transparencies. Often the photographic processes serve as intermediate steps toward the duplication of maps, that is, in the preparation of the artwork and in the production of printing plates. *See* PHOTOGRAPHY.

Diffusion transfer. A high-quality image is produced by first exposing negative paper in a camera or by contact with image light. The negative and receiver material are moistened in a liquid developer, followed by face-to-face contact which allows the image to be transferred from the negative to the receiver sheet. This process often serves as an intermediate step in the production of artwork and printing plates.

Electrostatic. This system, commonly called photocopying (or "xeroxing," from the proprietary name Xerox), uses a chemically coated plate which becomes light-sensitive when charged with electricity. The plate is put in the camera in place of film for exposure, after which the image is coated with a powder or liquid toner. The toner image is transferred to a paper or film and fused to the base by heat. The process can be used as an intermediate step in map production but more commonly serves to produce the printing plate. *See* PHOTOCOPYING PROCESSES.

Contact processes. Generally, these provide reproductions in the same size as the original, involve no optical equipment, and are of lower quality than the above processes.

Spirit duplicating. The map is drawn on a paper master, creating a carbon image on the reverse side. Alcohol is brushed over the carbon image and dissolves a small amount of the image with each revolution. The image is transferred in contact with each passing sheet of paper. This is the only mechanical process producing a multicolor image in one pass through the machine. Different-colored carbons can be combined on one master to prepare the image.

Mimeographing. The map is drawn on a wax-coated sheet of porous paper with styli which open segments of the stencil to allow ink to flow through when the stencil is mounted on a drum on the machine. Centrifugal force moves the ink from inside the drum and through the image area, depositing on the passing sheet of paper. Stencils can also be made from an original drawing by a heat development process and by the electrostatic process.

Diazo. This process produces a good-quality direct print, positive original to positive print, in one exposure and development. A positive transparent paper or film master containing an opaque image is placed in contact with a light-sensitive coated paper or film and exposed to ultraviolet light, neutralizing the coating except where hidden by the master's image. The copy is then passed through ammonia fumes for development, turning the image black.

Blueprinting. The master and exposure requirements are similar to those used with the diazo process. The light-sensitive coating is an iron compound which, after exposure, is washed and fixed in a chemical bath, turning deep blue. The unexposed coating behind the master's image is washed away, and the duplicate image becomes white. Diazo and blueprinting are the only processes to provide duplicate maps of a large size at low cost, limited only by the width of the machine and length of a roll of paper or film.

As an intermediate step in multicolor map construction, diazo or blueprinting serves to make registered duplicate compilation images on overlays to be scribed. Selected information is then scribed on each overlay, representing the color ultimately to be printed. This procedure assures that the various symbols and colors will print in register.

Computer-generated maps. Rapid developments in graphic software programming systems and the accompanying hardware have provided a number of possibilities in map compilation and reproduction of high quality and speed. Geographic information can be stored in data banks, to be called for later to plot images either for study and modification or for producing intermediates for printing plates. The information can be separated into overlays to produce the assigned colors for printing. Film negatives and positives drafted by plotters with ink, with a light beam, or by scribing are the common procedures. Such hardware is more limited to large mapping operations. *See* CARTOGRAPHY; COMPUTER GRAPHICS.

James A. Bier

Bibliography. I. S. Keates, *Cartographic Design and Production*, 2d ed., 1989; A. H. Robinson, *Elements of Cartography*, 6th ed., 1994; D. R. Taylor (ed.), *The Computer in Contemporary Cartography*, vol. 1, 1980.

Maple

A genus, *Acer*, of broad-leaved, deciduous trees including about 115 species in North America, Asia, Europe, and North Africa. This genus is characterized by simple, opposite, usually palmately lobed (rarely pinnate) leaves, generally inconspicuous flowers, and a fruit consisting of two long-winged samaras or keys. The winter buds have several overlapping scales, rarely only two. The most important commercial species is the sugar or rock maple (*A. saccharum*), called hard maple in the lumber market. This tree, attaining a height of 120 ft (37 m), grows in the eastern half of the United States and adjacent Canada. It can be recognized by its gray furrowed bark, sharp-pointed scaly winter buds, and symmetrical oval outline of the crown (**illus.** *a*). *See* SAPINDALES.

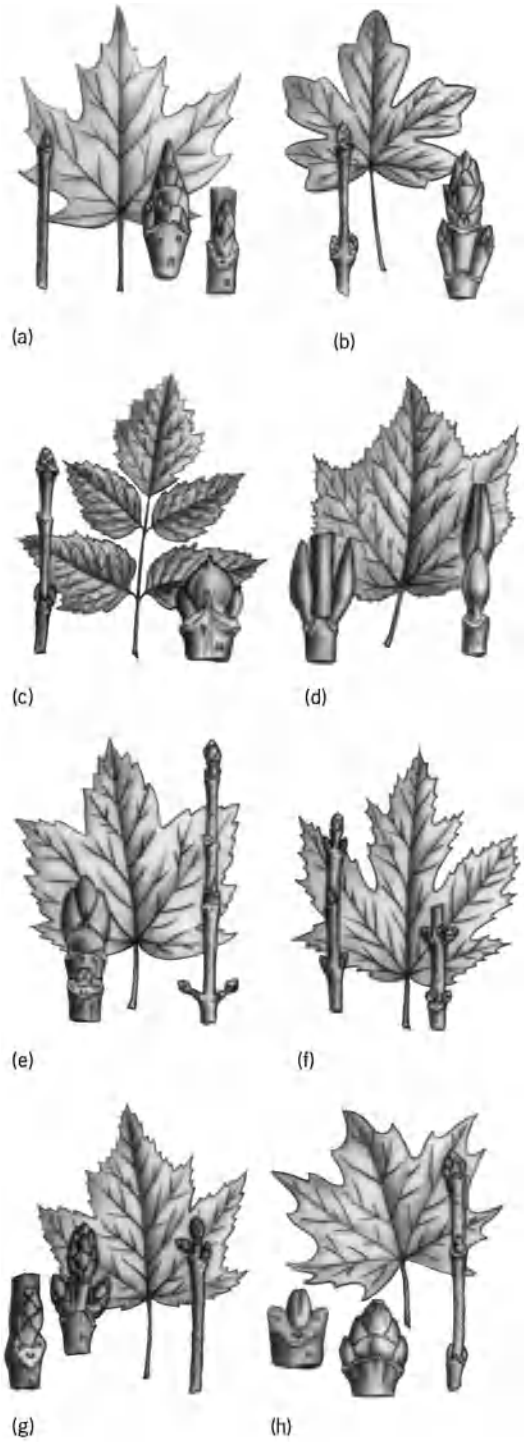

Characteristic maple leaves, twigs, and buds. (*a*) Sugar maple (*Acer saccharum*). (*b*) Hedge maple (*A. campestre*). (*c*) Box elder (*A. negundo*). (*d*) Striped maple (*A. pensylvanicum*). (*e*) Sycamore maple (*A. pseudoplatanus*). (*f*) Silver maple (*A. saccharinum*). (*g*) Red maple (*A. rubrum*). (*h*) Norway maple (*A. platanoides*).

Maple is an important source of hardwood lumber. Hard maple is used for flooring, furniture, boxes, crates, woodenware, spools, bobbins, motor vehicle parts, veneer, railroad ties, and pulpwood. It is the source of maple sugar and syrup and is planted as a shade tree.

The red maple (*A. rubrum*), with much the same botanical range as *A. saccharum*, is second commercially to sugar maple. It can be distinguished by fewer scales on the winter buds, scalier or flakier bark, and the red coloration of all parts of the shoot, leaves, flowers, fruits, and twigs. The leaf lobes are separated by sharp sinuses in contrast to the U-shaped ones of sugar maple (illus. *g*). Because it often grows in moist places, it is commonly known as swamp maple. In the lumber trade it is called soft maple.

The silver maple (*A. saccharinum*), with a range similar to *A. rubrum*, has more deeply cut leaves which are silvery beneath (illus. *f*). Its lumber and that of the red maple are used for furniture, boxes, crates, woodenware, spools, bobbins, railroad ties, and pulpwood. Both silver and red maples are planted as shade trees.

The bigleaf maple (*A. macrophyllum*), which attains a height of 90 ft (27 m) in the Pacific Coast region and north to Canada, can be recognized by its large leaves, which strongly resemble those of the eastern sugar maple, and by its long (about 1 ft or 0.3 m) leafstalks. The lumber is used for veneer, furniture, handles, fixtures, and woodenware. It is also planted as a shade tree.

The box elder (*A. negundo*), with pinnately compound leaves (illus. *c*), is native throughout most of the United States and part of Canada. It is often cultivated as a shade tree.

Two European species are popular street and shade trees: the Norway maple (*A. platanoides*), now becoming naturalized in North America, and the sycamore maple (*A. pseudoplatanus*). The former has copious milky sap in its leaves, twigs, and red bud scales, and close, nonscaly bark; the latter has green winter buds and bark which exfoliates in small polygonal plates (illus. *e* and *h*). *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Marattiales

An order (class Polypodiopsida) of primitive ferns that consists of about seven living genera and 150 species. They exist mostly in tropical regions. These large, coarse, sappy ferns usually have large, highly divided fronds. *Danaea simplicafolia*, however, has an undivided frond consisting of a single blade. Stems of living genera (*Angiopteris*, **illus.** *a*; *Marattia*) are short and bulbous, often a foot or more in diameter. Other genera have smaller, mostly horizontal, and somewhat more elongated stems. The internal vascular system (food and water conduction tissue) of the stems and fronds can be quite complex, apparently as a function of size. Marattialeans are included among the primitive ferns primarily because of the way that the sporangia (spore-bearing organs) arise from a mass of cells rather than from a single cell as in more highly evolved ferns. Sporangia are generally fused together into a synangium. Synangia are borne on the lower surface of the pinnules and are usually elongate. *Christensenia* differs from other genera in having palmately branched fronds (pinnate in other

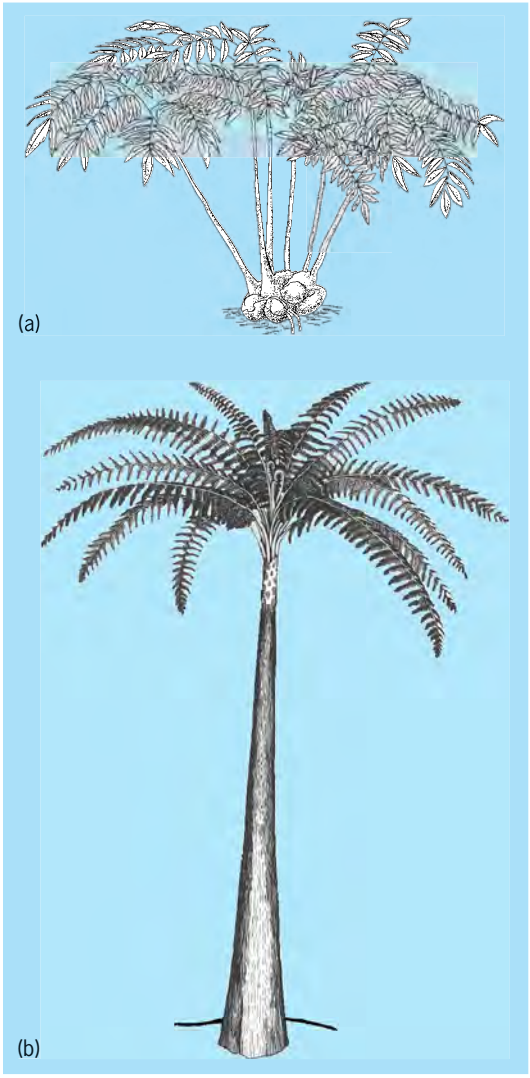

Typical marattialean ferns. (*a*) *Angiopteris*, an extant genus with a short bulbous stem. (*b*) *Psaronius*, a Carboniferous tree fern (*after B. M. Stidd, Morphology and anatomy of the frond of Psaronius, Palaeontographica B, 134:87–123, 1971*).

genera), net venation, and circular rather than linear synangia. Gametophytes (the small plant generation bearing egg- and sperm-producing organs) are large (among ferns) and thick and resemble thallose liverworts.

Psaronius is perhaps the best-known marattialean fern in the fossil record (illus. *b*). It was a prominent member of Carboniferous swamp forests, and much of the fern foliage (known as *Pecopteris*) found on coal mine spoil heaps represents portions of the large fronds borne on *Psaronius* stems. In contrast to most genera, *Psaronius* was more than 33 ft (10 m) tall. The stem was slender, 0.04 in. (1 mm) in diameter at the base, and increased upward, to about 6 in. (15 cm) or more. A plant so slender at the base could not stand erect while bearing the weight of a crown of fronds, each 10–16 ft (3–5 m) long. Consequently, the *Psaronius* trunklike support consisted mostly of a massive cylinder of roots that surrounded the stem. The roots arose from progressively higher levels on the stem as it grew upward; new roots grew downward over older ones and proliferating tissue from root surfaces bound the ensemble into a solid mass, at least in the inner portions of the root mantle. This system of stem encased in a solid layer of roots is highly unusual, if not unique, among plants, both living and fossil. Despite this distinctive feature of *Psaronius*, the anatomy of stem and frond vascular systems is fundamentally similar to that of living marattialean ferns. *See* PALEOBOTANY; PLANT KINGDOM; POLYPODIOPHYTA. Benton M. Stidd

Bibliography. W. N. Stewart and G. W. Rothwell, *Paleobotany and the Evolution of Plants*, 1993; T. N. Taylor and E. L. Taylor, *The Biology and Evolution of Fossil Plants*, 1993.

Marble

A term applied commercially to any limestone or dolomite taking polish. Marble is extensively used for building and ornamental purposes. *See* LIMESTONE; STONE AND STONE PRODUCTS.

In petrography the term marble is applied to metamorphic rocks composed of recrystallized calcite or dolomite. Schistosity, often controlled by the original bedding, is usually weak except in impure micaceous or tremolite-bearing types. Calcite (marble) deforms readily by plastic flow even at low temperatures. Therefore, granulation is rare, and instead of schistosity there develops a flow structure characterized by elongation and bending of the grains concomitant with a strong development of twin lamellae. In calcite the twinning plane is the flat rhombohedron $\{01\bar{1}2\}$; in dolomite twinning is markedly less common and the plane is the steep rhombohedron $\{02\bar{2}1\}$. *See* METAMORPHIC ROCKS; MINERALOGY.

Regional metamorphic marbles are more or less deformed and therefore texturally marked by characteristic deformation patterns—parallel elongation of irregularly bounded lensoid grains of calcite (less typically of dolomite)—and by preferred orientation of mica flakes, if mica is present, producing a tightly woven, close fabric. This makes the rock, particularly the pure calcite (not dolomitic) marble, well suited for building purposes.

Contact metamorphism of limestones and of dolomites produces granoblastic rocks composed mainly of a mosaic of equant grains of calcite. The fabric is loose and the rock not suitable for use as building material.

Pure marbles attaining 99% calcium carbonate ($CaCO_3$) are often formed by simple recrystallization of sedimentary limestone. Dolomite marbles are usually formed by metasomatism. More interesting petrographically are impure marbles in which various reactions between silicates and carbonates have taken place.

If an impure carbonate rock is subjected to slowly increasing temperature at constant pressure, a series of reactions involving progressive elimination of carbon dioxide takes place. Each of these reactions is determined by the temperature, pressure, and bulk composition of the reacting system. In natural rocks,

water and traces of other fugitive compounds are usually present. These facilitate the metamorphism and permit certain reactions to take place at low temperatures. The pressure usually has not great influence.

If at the beginning of the reaction there is a dolomitic limestone at low temperature composed of the phases calcite + dolomite + quartz (+ traces of water), certain mineral sequences will develop at given stages in the metamorphism. If the temperature is raised only a moderate amount, talc will form, followed by tremolite, forsterite, diopside, and brucite, until eventually wollastonite forms according to the reaction shown here. This reaction

$$\underset{\text{Calcite}}{CaCO_3} + \underset{\text{Quartz}}{SiO_2} \rightarrow \underset{\text{Wollastonite}}{CaSiO_3} + \underset{\text{Carbon dioxide}}{CO_2}$$

has been carefully studied, and it can be shown that it takes place at about 840°F (450°C) [assuming no great buildup of CO_2 pressure]. At higher temperatures rare lime-silicates, such as monticellite, akermanite, tilleyite, spurrite, merwinite, and others, develop.

Examinations of natural occurrences show that at the relatively low temperatures of regional metamorphism the diopside stage is rarely passed. At the somewhat higher temperatures of the contact zones about granites, which come from the coolest of magmas, the wollastonite stage may be attained but not passed. It is only at the hotter contacts of syenitic or granodioritic masses that members higher in the series begin to form, and the highest members are associated only with basic, and usually basaltic, rocks. T. F. W. Barth

Marchantiales

An order of the liverwort subclass Marchantiidae. Characteristic features include the differentiation of upper and lower tissues of the gametophyte, ventral scales, and rhizoids of two kinds. The sporophyte is considerably reduced, and the capsule dehisces irregularly. The archegonia, though dorsal, come to be pendant from an elevated receptacle because of differential growth resulting in decurved margins. The stalks of the receptacles are modified branches often with rhizoids in one, two, or rarely four furrows along their length. The order consists of 12 families grouped in two suborders, the more complex Marchantiineae and the simplified Ricciineae. *See* BRYOPHYTA; HEPATICOPSIDA; MARCHANTIIDAE. Howard Crum

Marchantiidae

One of the two subclasses of liverworts (class Hepaticopsida). The gametophytes are ribbonlike or rosette-shaped thalli, usually showing considerable internal tissue differentiation. The rhizoids may be both smooth and internally pegged on the same thalli. Oil bodies, if present, are restricted to scattered cells that lack chloroplasts. The antheridia are usually ovoid, and the archegonia usually consist of six rows of cells. The sporophytes are generally reduced.

The subclass differs from the thallose Metzgeriales of the subclass Jungermanniidae in the internal differentiation of the thallose gametophytes, the smooth and pegged rhizoids, and the variation in the way the capsules dehisce. The subclass is divided into three orders, the Marchantiales, the Monocleales, and the Sphaerocarpales. *See* BRYOPHYTA; HEPATICOPSIDA; JUNGERMANNIIDAE; MARCHANTIALES; METZGERIALES; MONOCLEALES; SPHAEROCARPALES. Howard Crum

Marcus equation

In the late 1950s and early 1960s Rudolph A. Marcus developed a theory for treating the rates of outer-sphere electron-transfer reactions. In outer-sphere reactions an electron is transferred from a donor to an acceptor without any chemical bonds being made or broken. (Electron-transfer reactions in which bonds are made or broken are referred to as inner-sphere reactions.) Marcus derived several useful expressions, one of which has come to be known as the Marcus cross-relation or the Marcus equation. It is widely used for correlating and predicting electron-transfer rates. For contributions to the understanding of electron-transfer reactions, Marcus received the 1992 Nobel prize in chemistry.

Like ordinary chemical reactions, an electron-transfer reaction can be described in terms of the motion of the system on an energy surface. As the reaction proceeds, the system moves from the reactant minimum (initial state) to the product minimum (final state). The nuclear configurations of the reactants and products and the configuration of the surrounding solvent are constantly changing as a consequence of thermal motion. Subject to certain assumptions, these fluctuations can be described in terms of displacements on harmonic free-energy curves that are a function of a single reaction coordinate. Two harmonic free-energy curves are needed to describe the reaction—one refers to the reactants plus surrounding medium, and the other to the products plus surrounding medium. The two free-energy curves have identical force constants, and the reaction coordinate is the difference between the reactant and product free energies at a particular nuclear configuration.

The free energy of the close-contact reactants plus surrounding medium (curve R) and the free energy of the close-contact products plus surrounding medium (curve P) are plotted versus the reaction coordinate (**Fig. 1**). The free-energy curves intersect where the reactants plus surrounding solvent and the products plus surrounding solvent have the same nuclear configurations and energies. This intersection defines the transition state for the reaction: the energy required to reach the intersection is the

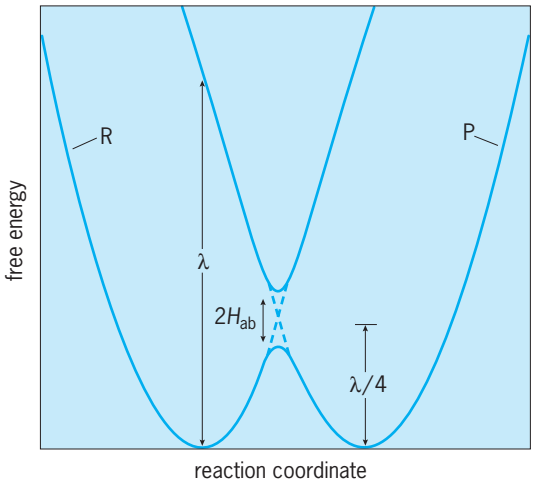

Key:
R = reactants
P = products
λ = reorganization parameter
H_{ab} = electronic coupling matrix element

Fig. 1. Free-energy plot for an electron-transfer reaction with zero standard free-energy change (an electron self-exchange reaction).

free energy of activation for the reaction. The reorganization parameter, λ, is the vertical difference between the free energies of the (noninteracting) reactants and products of a self-exchange reaction at the reactants' equilibrium configuration. Electronic interaction of the reactants gives rise to the splitting of the energy curves (noncrossing) at their intersection. This splitting is equal to $2H_{ab}$, where H_{ab} is the electronic coupling matrix element. The coupling of the reactants is assumed large enough that the electron transfer occurs with unit probability at the transition-state configuration but not large enough to lower the barrier significantly. The splitting is usually neglected in calculating the free energy of activation for the electron transfer.

Self-exchange reactions. The reactants and products of an electron self-exchange reaction are identical. An example is reaction (1), where phen = 1,10-phenanthroline. In terms of the Marcus formalism, the rate constant for a self-exchange reaction is given by Eqs. (2) and (3), where A is the collision

$$\mathrm{Fe(phen)_3^{2+} + Fe(phen)_3^{3+} = Fe(phen)_3^{3+} + Fe(phen)_3^{2+}} \quad (1)$$

$$k = A\exp(-\Delta G^*/RT) \quad (2)$$

$$\Delta G^* = \lambda/4 \quad (3)$$

frequency of the reactants in solution, ΔG^* is the activation free energy for the electron transfer with the two reactants in contact, R is the gas constant, T is the absolute temperature, and λ is the reorganization parameter. The reorganization parameter or intrinsic barrier comprises solvational and intramolecular (vibrational) components. The solvational contribution depends upon the solvent dielectric properties, on the distance separating the donor and acceptor sites and, for a given separation, on the shape of the reactants. The intramolecular contribution depends upon the bond length changes and force constants and is generally treated within a harmonic approximation.

Net electron-transfer reactions. The oxidation of $Ru(NH_3)_6^{2+}$ by $Fe(phen)_3^{3+}$ in reaction (4) is an example of an electron-transfer reaction accompanied by a net chemical change (most electron-transfer reactions fall into this category). Because the reactants and products of a net reaction differ, there is a net free-energy change in the reaction. The free-energy curves for the reactants and products of an exergonic electron-transfer reaction, that is, a reaction accompanied by a decrease in free energy, may be plotted (**Fig. 2**). The free energy of activation for such a reaction is given by Eq. (5), where ΔG° is the standard free-energy change for the reaction with the two reactants in contact. Equation (5) reduces to the expression for the free energy of activation for a self-exchange reaction, Eq. (3), when $\Delta G^\circ = 0$. $|\Delta G^\circ|$

$$\mathrm{Ru(NH_3)_6^{2+} + Fe(phen)_3^{3+} = Ru(NH_3)_6^{3+} + Fe(phen)_3^{2+}} \quad (4)$$

$$\Delta G^* = \frac{(\lambda + \Delta G^\circ)^2}{4\lambda} \quad (5)$$

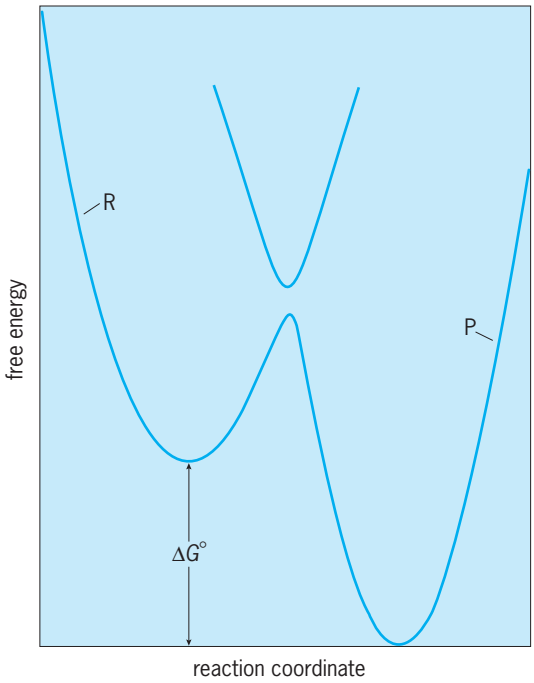

Key:
R = reactants
P = products
ΔG° = standard free-energy change

Fig. 2. Free-energy plot for an electron-transfer reaction when the free-energy change ($|\Delta G^\circ|$) is less than the reorganization parameter (λ).

may be viewed as the driving force for an exergonic electron-transfer reaction.

Three free-energy regimes can be distinguished depending on the relative magnitudes of the reorganization parameter and the driving force. When $|\Delta G^\circ| < \lambda$, the reaction is in the normal regime where ΔG^* decreases, and the rate constant increases, with increasing driving force (as is the case in Fig. 2). The reaction becomes barrierless ($\Delta G^* = 0$) when $|\Delta G^\circ| = \lambda$. If the driving force is increased even further, then $|\Delta G^\circ| > \lambda$ and ΔG^* increases, and the rate constant decreases, with increasing driving force. This is the nonintuitive inverted regime.

Marcus cross-relation. Marcus showed that the rate constant k_{12} for an electron-transfer reaction accompanied by a net chemical change is related to the rate constants k_{11} and k_{22} for the component self exchanges by Eq. (6). Here K_{12} is the equilibrium

$$k_{12} = (k_{11}k_{22}K_{12}f_{12})^{1/2}W_{12} \tag{6}$$

$$\ln f_{12} = \frac{[\ln K_{12} + (w_{12} - w_{21})/RT]^2}{4\left[\ln\left(\frac{k_{11}k_{22}}{Z^2}\right) + \frac{w_{11} + w_{22}}{RT}\right]}$$

$$W_{12} = \exp[-(w_{12} + w_{21} - w_{11} - w_{22})/2RT]$$

constant for the cross-reaction, f_{12} is a factor that becomes significant at large driving forces, W_{12} is a factor that corrects for differences between the electrostatic work terms for the various reactions, w_{12} (w_{21}) is the electrostatic work required to bring together the reactants (products) of the cross-reaction, w_{11} (w_{22}) is the corresponding work to bring together the reactants in the self-exchange reaction, and Z is the collision frequency of two uncharged molecules in solution (10^{11} M^{-1} s^{-1}). Equation (6) is based upon the assumption that the reorganization parameter for a cross-reaction is equal to the average of the reorganization parameters for the component self-exchange reactions, Eq. (7).

$$\lambda_{12} = \frac{\lambda_{11} + \lambda_{22}}{2} \tag{7}$$

For example, if the oxidation of $Ru(NH_3)_6^{2+}$ by $Fe(phen)_3^{3+}$ [reaction (4)] is the cross-reaction (k_{12}, λ_{12}), then the $Ru(NH_3)_6^{2+}$–$Ru(NH_3)_6^{3+}$ reaction (8) and

$$Ru(NH_3)_6^{2+} + Ru(NH_3)_6^{3+} = Ru(NH_3)_6^{3+} + Ru(NH_3)_6^{2+} \tag{8}$$

$Fe(phen)_3^{2+}$–$Fe(phen)_3^{3+}$ reaction (1) are the component self-exchange reactions (k_{11}, λ_{11} and k_{22}, λ_{22}). When the driving force is not too large and the work terms cancel, then $\ln f_{12} \approx 0$ and $W_{12} = 1$. Under these conditions, Eq. (6) reduces to Eq. (9).

$$k_{12} = (k_{11}k_{22}K_{12})^{1/2} \tag{9}$$

Equation (9) is often referred to as the Marcus equation. It has been extensively applied to electron-transfer reactions involving metal complexes. The stable oxidation states of transition-metal complexes generally differ by one electron, and the metal centers and surrounding ligands can be independently varied to encompass a wide range of electron-transfer rates and driving forces. For moderately exergonic electron-transfer reactions between similarly charged complexes, the agreement of the observed rates with those calculated from Eq. (9) is often remarkably good. At high exergonicities or when the net electron transfer involves oppositely charged complexes, it is necessary to use Eq. (6). The rates of highly exergonic reactions calculated with the extended equation are usually somewhat higher than the observed rates in the normal free-energy regime and are slower than the observed rates in the inverted regime. Reasons advanced for the failure of the cross-relation include changes in reaction mechanism, anharmonicity contributions, nuclear tunneling contributions, and noncancellation of nonelectrostatic contributions to the work terms.

Examples of the application of the Marcus equation to organic or organometallic systems include the oxidation of tetraalkyltin, tetraalkyllead, and dialkylmercury by $Fe(phen)_3^{3+}$; the oxidation of various methylarenes, $ArCH_3$, by substituted $Fe(phen)_3^{3+}$ complexes, reaction (10); and the oxidation of

$$ArCH_3 + Fe(phen)_3^{3+} = ArCH_3^+\cdot + Fe(phen)_3^{2+} \tag{10}$$

N-alkylpyridinyl radicals by [Ir(1,5-cyclooctadiene)(μ-pyrazolyl)]$_2^+$, reaction (11), where very good

$$C_6H_5N\text{–}R\cdot + Ir_2^+\cdot = C_6H_5N\text{–}R^+ + Ir_2 \tag{11}$$

agreement with the predictions for highly exergonic electron transfer is obtained. The Marcus equation has also been succesfully applied to reactions of alkylhydrazines, ferrocene derivatives, and *p*-phenylenediamine derivatives.

Because the reorganization barrier is an intrinsic property of a redox couple and independent of its reaction partner, a simple relationship should exist between the homogeneous self-exchange rate constant of a redox couple and its exchange rate constant at an electrode. The relationship between homogeneous and heterogeneous electron-transfer rate constants predicted by the Marcus formalism also holds up well in practice.

Proton, hydride, atom, and group transfers. Although originally derived for electron-transfer reactions, the Marcus formalism can also be applied to reactions in which bond making and bond breaking occur. If the net reaction involves transfer of a covalently attached X from a donor to an acceptor, then the cross-reaction and self-exchanges need to be redefined in terms of X-transfer reactions as shown in Eqs. (12).

$$\begin{aligned} AX + B &= A + XB \\ AX + A &= A + XA \\ BX + B &= B + XB \end{aligned} \tag{12}$$

For example, the oxidation of a methylarene, reaction (10), is followed by proton transfer from the methylarene cation radical to a substituted pyridine,

reaction (13). In this case, $X = H^+$, $A = ArCH_2^{\cdot}$, and

$$ArCH_3^+\cdot + Y\text{-}py = ArCH_2\cdot + Y\text{-}pyH^+ \qquad (13)$$

$B = Y\text{-}py$. The Marcus formalism yields a consistent value for the intrinsic barrier for the proton transfer. Other applications include hydride transfers, reaction (14), where A_iN^+R is a nitrogen heterocycle; gas-phase atom transfers, reaction (15); and nucleophilic displacements (methyl transfers), reaction (16).

$$A_1(H)NR + A_2N^+R = A_1N^+R + A_2(H)NR \qquad (14)$$

$$RH + CH_3\cdot = R\cdot + CH_4 \qquad (15)$$

$$ClCH_3 + F^- = Cl^- + CH_3F \qquad (16)$$

Since the transfer of the heavier particles generally involves strong electronic interactions, the X transfers are not well described by intersecting harmonic free-energy curves. Use of a bond-energy bond-order (BEBO) model for the X transfers results in a free-energy expression similar to Eq. (5) when $|\Delta G^\circ|/\lambda$ is small, but differs from it in not predicting an inverted regime at large $|\Delta G^\circ|/\lambda$. The λ parameter is found to have the same property as Eq. (5), and a cross-relation analogous to Eq. (6) is obtained.

The expression for the free-energy of activation, Eq. (5), is applicable to reactions that can be described in terms of two weakly interacting harmonic free-energy curves. It does not apply when there is an appreciable contribution from bond making or rupture in the rate-determining step. The barrier averaging relation, Eq. (7), is, however, more general: intrinsic barrier additivity obtains in both electron and heavier-particle transfers. The factors determining reaction rates, particularly the additivity of intrinsic barriers, have proved invaluable in understanding and predicting the rates of a variety of electron-, atom-, and group-transfer reactions. Norman Sutin

Bibliography. D. A. Astruc, *Electron Transfer and Radical Processes in Transition-Metal Chemistry*, 1995; L. Eberson, *Electron Transfer Reactions in Organic Chemistry*, 1987; R. A. Marcus, *Annu. Rev. Phys. Chem.*, 15:155–196, 1964; R. A. Marcus, *Pure Appl. Chem.*, 69:13–29, 1997; R. A. Marcus and N. Sutin, *Biochim. Biophys. Acta*, 811:265–322, 1985; S. F. Nelsen et al., *J. Amer. Chem. Soc.*, 119:5900–5907, 1997; N. Sutin, *Prog. Inorg. Chem.*, 30:441–498, 1983.

Margarine

An emulsified fatty food product used as a spread and a baking and cooking fat, consisting of an aqueous phase dispersed in the fat as a continuous phase. Hyppolyte Mège-Mouriés invented margarine in 1869 and won the prize offered by Louis Napoleon for the development of an inexpensive butter substitute. Margarine is now considered a food in its own right and is manufactured in forms unknown to butter, such as plastic, soft, or fluid. However, the interrelationship is still acknowledged since margarine is colored, flavored, fortified with vitamins, and otherwise formulated to have the same or similar taste, appearance, and nutritional value as butter. *See* BUTTER.

Margarine is regulated by law in practically every country in which it is manufactured or sold. In 1977 the U.S. Food and Drug Administration revised and broadened its margarine Standards of Identity to conform to the agreements reached by the Codex Alimentarius Commission, an international body of food and legal authorities charged with the responsibility of unifying foods sold in international trade.

Margarine produced in the United States must contain not less than 80% fat. The fats and oils must be edible but may be from any vegetable or animal carcass source, natural or hydrogenated. The required aqueous phase may be water, milk products, or solutions of dairy or vegetable protein, and must be pasteurized. Milk may be cultured after pasteurization. Vitamin A must be added to yield a finished margarine with not less than 15,000 international units per pound (0.45 kg). Optional ingredients include salt or potassium chloride for low-sodium diets, nutritive carbohydrate sweeteners, fatty emulsifiers, antioxidants, preservatives, edible colors, flavors, vitamin D, acids, and alkalies. *See* PASTEURIZATION.

Margarines are formulated by blending two or more fats and oils in proportions designed to obtain the desired texture, consistency, and melting characteristics in the finished product. The various types of components include unhydrogenated or liquid, lightly, moderately, and highly hydrogenated oils. Meat fats are used in various bakery margarine formulations. *See* FAT AND OIL (FOOD).

Low-calorie spreads designed to resemble margarine contain only 40–60% fat. The higher level of aqueous phase in the product requires the addition of emulsion stabilizers. These may be high levels of milk solids, casein, vegetable proteins, vegetable gums, or various emulsifiers. Low-fat spreads are not covered by U.S. Standards of Identity and cannot be labeled as margarine. *See* FOOD ENGINEERING.

Theodore J. Weiss

Bibliography. *Code of Federal Regulations*, 21, 166.110, 1983; H. Lawson, *Food Oils and Fats: Technology, Utilization, and Nutrition*, 1994; T. J. Weiss, *Food Oils and Their Uses*, 2d ed., 1983.

Marine biological sampling

The collection and observation of living organisms in the sea, including the quantitative determination of their abundance in time and space. The biological survey of the ocean largely depends on specially equipped vessels. Sampling in intertidal regions at low tide is one of the few instances where it is possible to observe and collect marine organisms without special apparatus.

A primary aim of marine biology is to discover how ocean phenomena control the distribution of organisms. Sampling is the means by which this aim is accomplished. Suppositions about the dynamics of

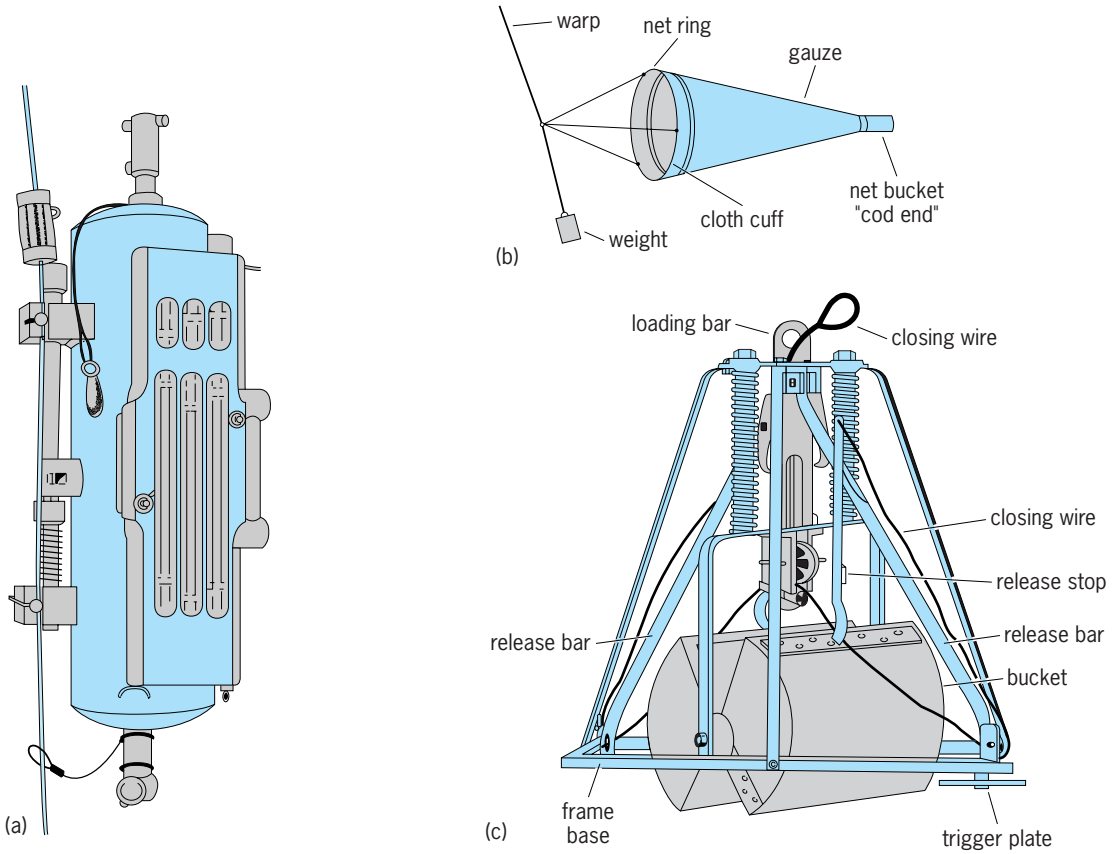

Fig. 1. Marine sampling devices. (*a*) Water sampling bottle with attached thermometer rack for recording water temperature at sampling depth. The brass messenger on the wire has just tripped the spring-loaded end closures (*after S. Niskin, A reversing-thermometer mechanism for attachment to oceanographic devices, Limnol. Oceanogr., 9:591–594, 1964*). (*b*) Conventional plankton net (*after C. Schlieper, ed., Research Methods in Marine Biology, University of Washington Press, pp. 46–63, 1972*). (*c*) Spring-loaded grab sampler (*after W. Smith and A. D. MacIntyre, J. Mar. Biol. Ass. U.K., 33:257–264, 1954*).

ocean life rest entirely on sampling methodology. Traditional techniques use samplers attached to wires lowered over the side of a ship by hydraulic winches (**Fig. 1**). These samplers include bottles designed for collecting seawater samples from particular depths, fine-meshed nets that are towed behind the ship to sieve out plankton and fish, and grabs or dredges that collect animals inhabiting the ocean bottom. These types of gear are relied upon in many circumstances; however, they illustrate some of the problems common to all methods by which the ocean is sampled. First, sampling is never synoptic, which means that it is not possible to sample an area of ocean so that conditions can be considered equivalent at each point. Usually, it is assumed that this is so. Second, there are marine organisms for which there exists no sampling methodology. For example, knowledge of the larger species of squid is confined to the few animals that have been washed ashore, or to the rare video recording. Third, the representativeness of the samples collected is problematic. The open ocean has no easily definable boundaries, and organisms are not uniformly distributed. The actual sampling is regularly done out of view of the observer; thus, sampling effectiveness is often difficult to determine. Furthermore, navigational systems are not error-free, so the position of the sample is never precisely known. All developments in methods for sampling the ocean try to resolve one or more of these difficulties by improving synopticity, devising more efficient sampling gear, or devising methods for observation such that more meaningful samples can be obtained.

Plankton. Single samples for phytoplankton and zooplankton are obtained by using sampling bottles or by towing a plankton net through the water horizontally, vertically, or obliquely (Fig. 1). These samples are preserved and then counted and identified using a microscope. Two developments in sampling planktonic organisms are continuous-recording systems, and multiple-net systems that collect zooplankton from various depths. *See* PHYTOPLANKTON; ZOOPLANKTON.

Continuously recording systems. These methods for phytoplankton take advantage of the fact that chlorophyll *a*, the primary light-sensitive pigment used in photosynthesis, is a fluorescent molecule. This means that when chlorophyll-containing plant cells are illuminated (excited) by a particular wavelength of light, they reemit (fluoresce) light at another wavelength. Chlorophyll is excited by blue light and fluoresces red. Seawater is pumped aboard ship and passed through an instrument called a fluorometer that illuminates the water and records the level of fluorescence due to chlorophyll *a* in the

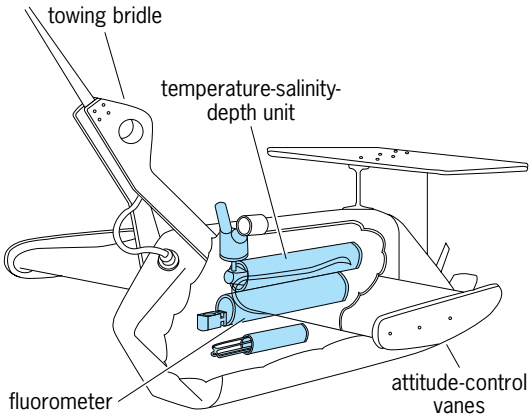

Fig. 2. Towed undulating vehicle system with some of its sensors. (*Chelsea Instruments, Ltd., United Kingdom*)

phytoplankton. This technique thus provides a continuous estimate of the total amount of plant biomass as the water is pumped through the fluorometer. Some fluorometers can be lowered over the side of the ship to measure directly the fluorescence in the water. A strobe light is used to illuminate a small volume near the device, the fluorescence is sensed, and a signal is sent up a conducting cable to shipboard recording devices. *See* CHLOROPHYLL; FLUORESCENCE; LUMINESCENCE ANALYSIS.

Acoustic methods and video cameras can be used for estimating zooplankton abundances continuously. The sound sources used are derived from fish-finder sonars developed for commercial fishing. Sound energy is transmitted at multiple ultrahigh frequencies, usually between 10 kHz and 1 MHz, insonifying a volume of water within 1–2 m (3.3–6.6 ft) of the transducer; and the sound that is reflected, or backscattered, from the organisms is then

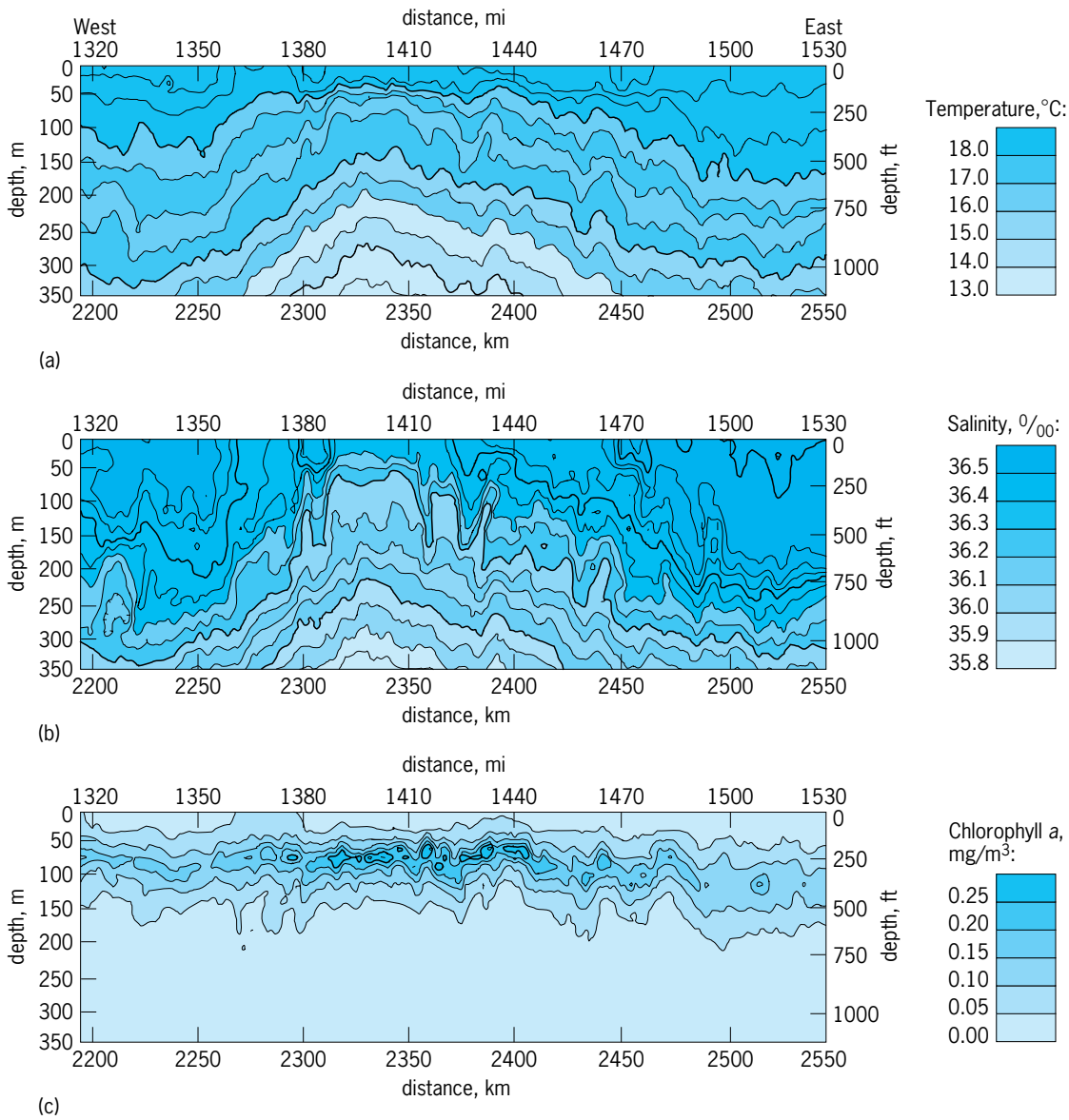

Fig. 3. Data collected by a towed undulating vehicle along a transect in the ocean. (*a*) Temperature; $^\circ F = (^\circ C \times 1.8) + 32$. (*b*) Salinity. (*c*) Chlorophyll *a*. (*Chelsea Instruments, Ltd., United Kingdom*)

recorded electronically. In this range, organisms 0.1–10 mm (0.004–0.4 in.) in size can be detected. Usually, the transducers for each frequency are triggered sequentially, and the acoustic backscatter measured from each. Backscattered sound depends on size. Smaller organisms can be detected only with high frequencies, while larger organisms can be detected at all frequencies. In the video method, a video camera with a synchronized strobe light is used to sample a volume of water at 60 frames per second. The recorded images are digitized and scanned for zooplankton.

The fluorescence method for phytoplankton and the acoustic and video methods for zooplankton can be used while the ship is under way. For phytoplankton, water is normally pumped continuously to shipboard labs, whereas acoustic transducers are mounted to the hull or are towed just behind the ship near the surface. *See* SCATTERING LAYER; SONAR.

The towed-undulating vehicle instrument can incorporate all three of the continuous-recording techniques as an undulating instrument package towed behind the ship (**Fig. 2**). Variations in plankton abundance (as acoustic backscatter or optical images for zooplankton, and as chlorophyll *a* for phytoplankton) along with environmental parameters such as temperature and salinity are recorded simultaneously in two dimensions (**Fig. 3**).

These techniques are indirect and, with the exception of video methods, provide little information on what species are present. They do provide a detailed, nearly synoptic estimate of the small-scale variation in abundances of plankton over wide areas and with depth. Species analysis of phytoplankton requires bottle sampling (or in special cases, nets). Study of particular zooplankton organisms or groups can be accomplished by use of multiple-net systems.

Multiple-net systems. These types of sampling gear are used to characterize the small-scale vertical distribution of zooplankton species. There are several designs in use. One of the most successful is the multiple opening/closing net environmental sensing system (MOCNESS). MOCNESS carries nine nets that are sequentially opened and closed at desired depths by sending electrical command signals down a conducting cable. The system is deployed with one net opened and is lowered to the desired depth. As MOCNESS is hauled in, nets are opened and closed sequentially at desired depths. The opening of one net automatically closes the previous net. For example, MOCNESS can be sent down to 1000 m (3000 ft) with one net open, and the remaining eight nets could be opened/closed at increments of 125 m (400 ft) as the net is hauled in. An electronic instrument package is mounted on the main frame and contains salinity, temperature, pressure (depth), light, and dissolved oxygen sensors for the recording of environmental data, as well as sensors that detect the flow of water through the net and the attitude of the net as it is towed. During a tow, all this information is recorded on a computer on the ship for later analysis. Every tow provides a series of eight zooplankton samples, each from a different stratum, and one sample representing organisms collected over the entire depth that the net fished.

Nekton. For catching the more active fishes and cephalopods and the larger crustaceans, large-sized nets are essential. The mouth opening of the MOCNESS has been enlarged for this purpose, to 215 ft^2 (20 m^2) for example, and represents an advance in nekton sampling because multiple depths can be fished. A similar design is the rectangular midwater trawl, combining multiple nets of 11 ft^2 (1 m^2) and others of 88 ft^2 (8 m^2), the two-sized nets being used to sample plankton and nekton, respectively. Other enhancements to these nets include specially designed cod ends that maintain the pressure and temperature of the sample as it is collected. By using this last modification, deep-living fish and crustaceans can be brought to the surface alive. A net that will catch larger squids and fishes inhabiting the deep has yet to be developed. For these organisms, video recording and direct viewing from deep-sea vehicles are the best options. *See* DEEP-SEA FAUNA.

Benthos. Organisms that live on the bottom are sampled by grabs, dredges, trawls, and box corers. Grabs are basically a pair of jaws, triggered in various ways to close when they touch bottom, capturing a small (about 0.1 m^2 or 3.3 ft^2) sample of sediment. Several designs are available, depending on bottom type and depth of water (Fig. 1). The sample, after being returned to the deck, is washed and sieved for the organisms it contains.

Dredges have a bag of coarse-meshed netting attached to a rectangular wrought-iron frame (0.6–0.9 m or 2-3 ft wide) that forms the mouth of the net. The dredge is towed along the bottom by means of a movable bridle arrangement, while the lower side of the frame digs into the sediment to scoop up organisms living just below the surface.

Bottom trawls are baglike nets tapering to a cod end. The mouth is kept open by otter boards, to which the two towing warps are attached. The otter boards also serve to disturb organisms off the bottom so that they become trapped in the net. Sometimes, counting wheels are fitted to trawls or dredges to give an indication of distance traveled.

The box corer (**Fig. 4**) is a sediment-coring device. However, the same features that make it useful for the study of ocean sediments make it a valuable piece of equipment for benthic biological research as well. The advantages of the box corer over grabs and dredges are the precision with which it can obtain a bottom sample and the fact that it can bring a largely undisturbed sediment sample to the surface. Therefore, a better characterization of the environment of the organisms can be obtained. In addition, there is little bow wave as the corer approaches the bottom, and it can work in all types of sediment. The disadvantage is that it is large and heavy and thus requires handling by the most powerful winches.

The box corer is allowed to free-fall to the ocean bottom. As the weight of the corer drives it into the bottom, the wire slackens. This action releases the spade, which swings to close off the bottom

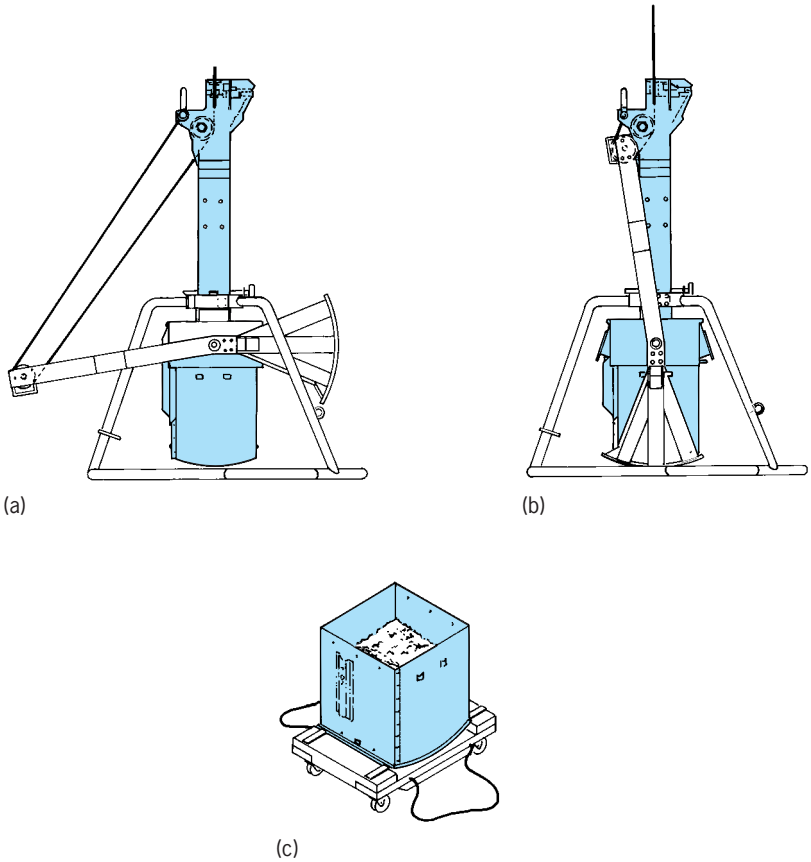

Fig. 4. United States Naval Electronics Laboratory box corer. (*a*) Side view with spade in open position. (*b*) Side view with spade in closed position. (*c*) Disengaged core box for processing sample. (*After R. R. Hessler and P. A. Jumars, Abyssal community analysis from replicate box cores in the central North Pacific, Deep-Sea Res., 21:185–209, 1974*)

of the corer. The box corer is then brought to the surface. *See* MARINE SEDIMENTS.

Bacteria and viruses. The use of fluorescent dyes, or fluorochromes, simplifies sampling and enumeration of marine bacteria. For this technique, a known volume of uncontaminated seawater or sediment (collected by water sampling bottles, grabs, or corers) is incubated for a short time with the fluorochrome, while bacteria cells absorb some of the dye. The cells are then collected on a filter and viewed under a high-powered microscope equipped with a light source that causes the dye to fluoresce. The bacteria appear as bright objects against a dark background and are counted directly. Acridine orange, a fluorochrome that binds to nucleic acids (RNA and DNA), is commonly used for enumerating marine bacteria. Not all bacterial cells seen under the fluorescent microscope are alive or capable of growth in laboratory culture, although the abundances measured correlate with other plankton variables. Viruses must be enumerated with an electron microscope, usually on samples that have been concentrated and filtered for larger forms. *See* MARINE ECOLOGY; MARINE MICROBIOLOGY; STAIN (MICROBIOLOGY).

Other methods and observations. Direct and remote observation methods provide valuable information on the undersea environment and thus on the representativeness of various sampling techniques. Personnel-operated deep-submergence research vessels (DSRVs) are increasingly used to observe ocean life at depth and on the bottom, and to determine appropriate sampling schemes. The deep-submergence research vessels are used with cameras and television recording equipment and are fitted with coring devices, seawater samplers, and sensors of various types. Other cameras are operated unattended at the bottom for months at a time, recording changes occurring there. Scuba diving is playing a larger role, especially in open-ocean areas, and is used to observe marine organisms in their natural habitat as well as to collect the more fragile marine planktonic forms such as foraminifera, radiolaria, and jellyfish. Autonomous underwater vehicles (AUVs) continue to assume greater importance in sampling programs since they can go to greater depths than can divers, and they overcome a limitation in diving in that AUVs can be operated at night and can sample extensively under ice. Optical sensors carried aboard Earth-orbiting satellites can provide images of ocean color over wide areas. Ocean color is related to the turbidity as well as the amount of plant material in the seawater. This establishes a means by which sampling programs from ships can be optimized. *See* DIVING; SEAWATER FERTILITY; UNDERWATER PHOTOGRAPHY; UNDERWATER TELEVISION; UNDERWATER VEHICLE.

John Marra

Bibliography. P. F. Brodie, D. D. Sameoto, and R. W. Sheldon, Population densities of euphausiids off Nova Scotia as indicated by net samples, whale stomach contents, and sonar, *Limnol. Oceanog.*, 23:1264–1267, 1978; C. S. Davis et al., Rapid visualization of plankton abundance and taxonomic composition using the Video Plankton Recorder, *Deep Sea Res. II*, 43:1947–1970, 1996; K. L. Denman and D. L. Mackas, Collection and analysis of underway data and related physical measurements, pp. 85–109 in J. H. Steele (ed.), *Spaital Pattern in Plankton Communities*, Plenum, New York, 1978; L. R. Haury, J. A. McGowan, and P. H. Wiebe, Patterns and processes in the time-space scales of plankton distributions, pp. 277–327 in J. H. Steele (ed.), *Spaital Pattern in Plankton Communities*, Plenum, New York, 1978; J. C. Kelly, Sampling the sea, pp. 361–387 in D. H. Cushing and J. J. Walsh (eds.), *The Ecology of the Seas*, W. B. Saunders, Philadelphia, 1976; J. Sherman et al., The autonomous underwater glider "Spray," *IEEE J. Oceanic Eng.*, 26:437–446, 2001.

Marine boiler

A steam boiler designed to suit the marine environment and generally arranged to supply steam to the main propulsion machinery, ship's service electric generators, feed-pump drivers, and other auxiliary services. *See* MARINE MACHINERY.

Types of units. Marine boilers are usually of the two-drum water-tube type with water-cooled furnaces, superheaters, desuperheaters, and heat

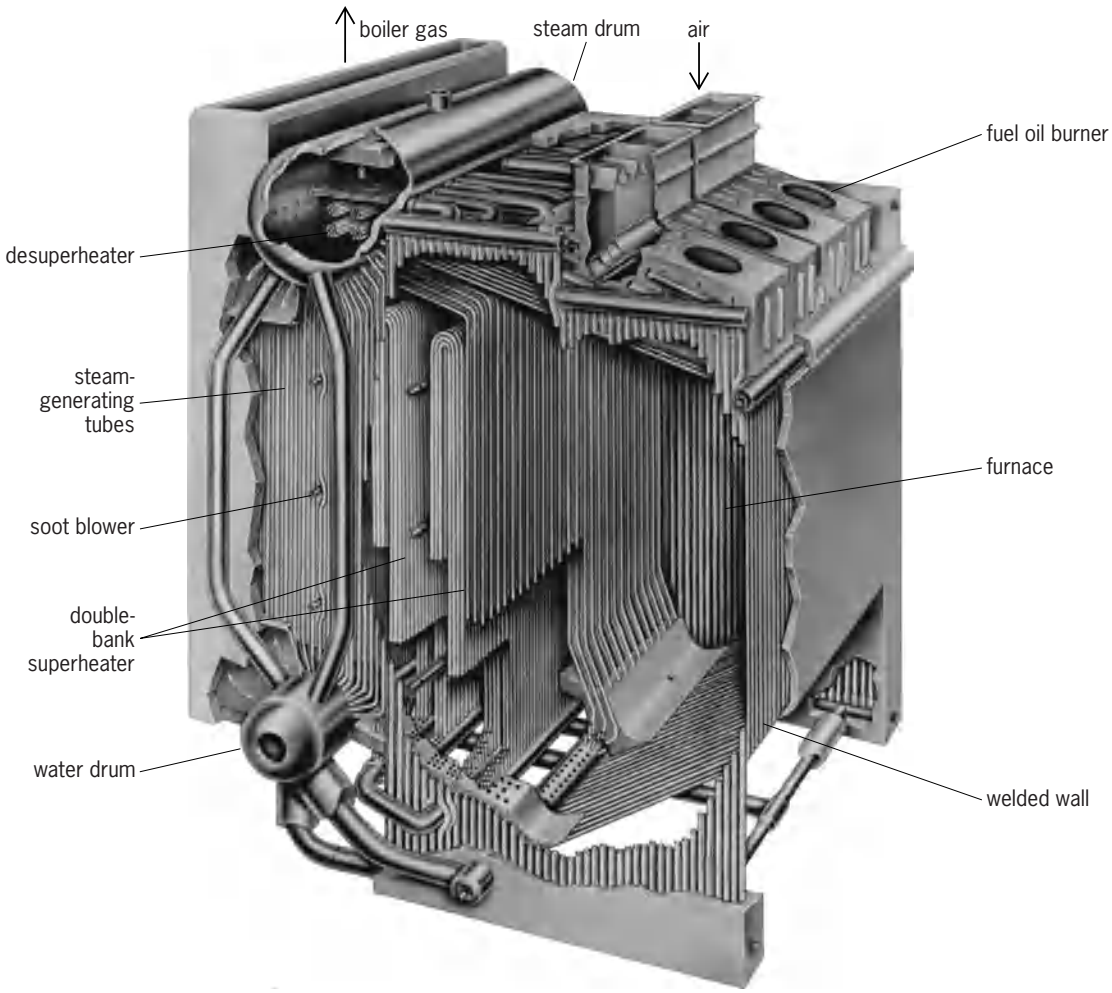

Fig. 1. Typical nonreheat, two-drum, down-fired welded-wall, double-superheater marine boiler. (*Foster Wheeler Boiler Corp.*)

recovery equipment of the economizer or air-heater type. The older fire-tube boilers, commonly referred to as Scotch boilers, and water-tube header types are no longer manufactured for marine service, although some of these boilers are still in use on older ships. *See* FIRE-TUBE BOILER; WATER-TUBE BOILER.

The majority of ships are fitted with two boilers, although some large passenger ships may have three or more. Some cargo ships are fitted with only one boiler, and in some of these cases a smaller auxiliary boiler may be fitted for emergency or in-port steaming use.

Design considerations. The pressure and temperature of the steam produced by marine boilers are dependent upon the steam cycle selected, which in turn relates to economic considerations such as the cost of fuel and the installed cost for the equipment for the propulsion power plant. Since the cost of fuel (energy) is high and rising as it relates to the installed cost of the equipment, there is a need for efficient steam plants, which in turn require high steam pressures and temperatures and efficient boilers. To attain high boiler efficiency it is necessary to minimize the heat loss up the boiler stack. This can be accomplished by utilizing heat recovery equipment such as economizers or air heaters. *See* AIR HEATER; BOILER ECONOMIZER.

Worldwide, most modern ships generate steam at a pressure of 850 lb/in.^{2} gage (5.86 megapascals) and a temperature of 955°F (513°C) at the superheater outlet and employ rotary regenerative air preheaters for the heat recovery equipment. However, in the United States most power plants employ the less efficient economizer as the heat recovery unit.

Figure 1 shows a cutaway view of a typical two-drum marine boiler. It has a welded wall construction in way of the furnace and superheater; that is, the tubes are welded together and joined by ligaments so as to make the furnace and area in way of the superheaters gas- and airtight. *See* SUPERHEATER.

In the upper boiler drum a submerged desuperheater is shown. The purpose of the desuperheater is to lower the steam temperature for a portion of the superheated steam that may be used for auxiliary purposes.

Although they are not popular in the United States, a number of marine power plants employ a steam cycle using boiler steam reheaters. With this arrangement, steam that has passed through a portion of the turbine is directed back into the boiler for

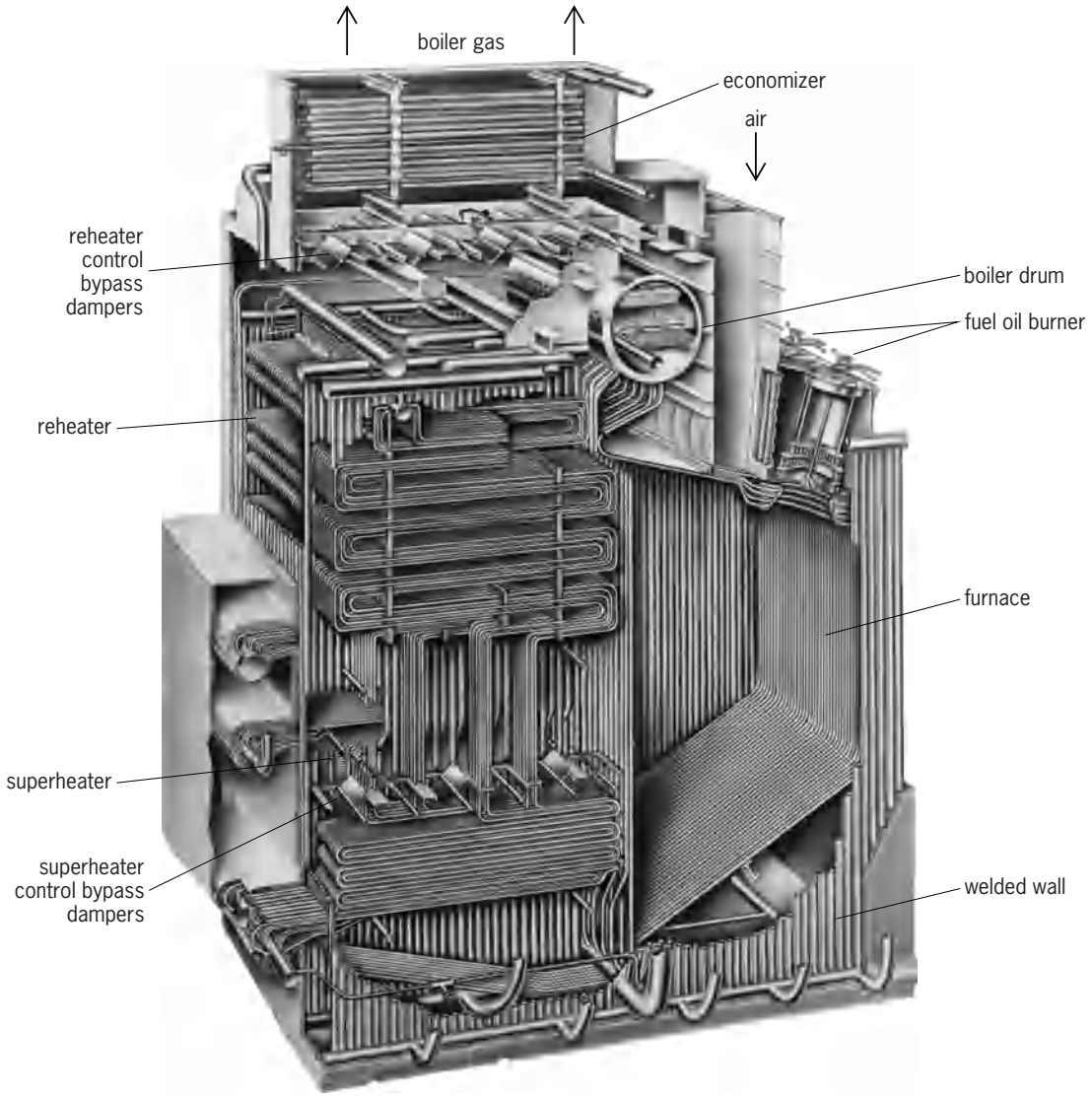

Fig. 2. Typical reheat, down-fired, welded-wall marine boiler containing primary superheater and reheater with economizer and gas damper control. The boiler gas air heater is not shown. (*Foster Wheeler Boiler Corp.*)

reheating in a device similar to the initial boiler superheater, except that the pressure is lower. The advantage of using this arrangement is that higher initial steam pressure may be employed, generally around 1450 lb/in.^{2} gage (10.0 MPa), and this advantage, coupled with the advantage of reheating the steam, can result in propulsion power plants where the fuel consumption may be 5–8% less than the conventional arrangement discussed above. *See* REHEATING.

Figure 2 shows an economizer in the gas path; this would be installed in series with a gas air heater of the type shown in **Fig. 3**. The air heater shown is a rotary regenerative type and consists of a drum filled with a plate-type heat exchange surface that is rotated first through the boiler exhaust gas path, which heats the surface while cooling the gas, and then through the air that is discharged from the boiler-forced draft fan, which heats the air while cooling the surface. The heated air then passes to the boiler, where it supplies the air necessary for combustion. The use of an air heater of this type can permit boiler efficiencies to be in the range of 91%, based on the lower heating value of the fuel.

Although the heat exchange surfaces are coated with acid-resistant enamel, they deteriorate in time and must be replaced, particularly in the colder part of the heater. Thus arrangements must be made for easy replacement of the surfaces.

Where economizers alone are fitted to serve as the boiler heat-recovery equipment, it is common to use steam air heaters. With this arrangement the air for combustion is forced through a heat exchanger containing extended surface tubes where heat is extracted from the steam to warm the combustion air.

Special design requirements. In the design of marine boilers, special requirements must be considered. The units must be compact and economical with regard to space and weight. Also, they must be built to withstand the effect of the pitching and rolling motions of the ship, as well as vibration which is caused by the ship's propeller. Reliability must be inherent in the design because of the requirements of safety at sea.

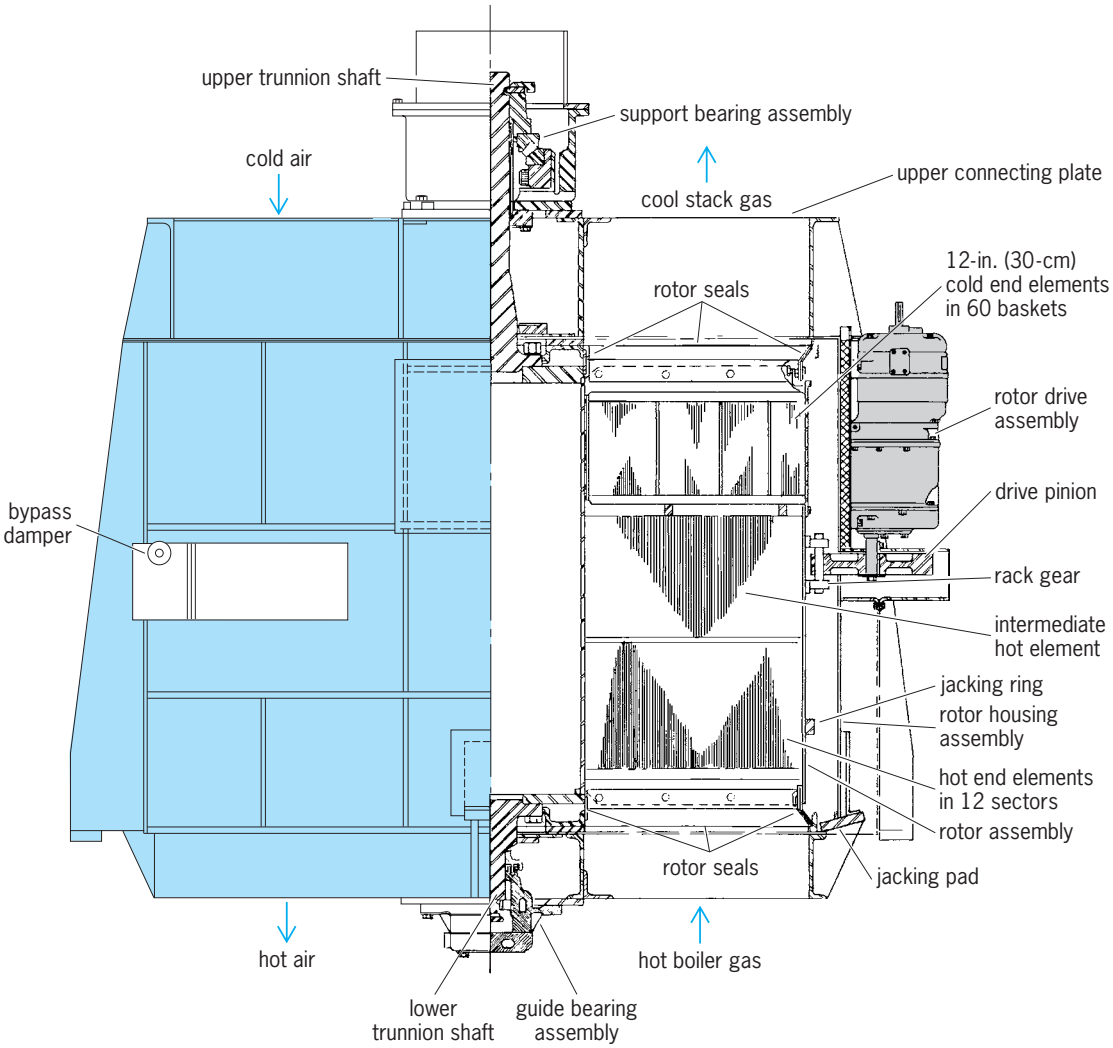

Fig. 3. Vertical rotary regenerative gas air heater. (*Air Preheater Company, Inc.*)

Fuels. Marine boilers are generally arranged for oil firing. Oil is used extensively because of its simplicity of handling and storing and the fact that it can be stored in spaces that often cannot be used for carrying cargo. For economic reasons, coal was used in the United States for boats operating on the Great Lakes, but most of these installations have been converted to use oil as a fuel.

Oil-fired boilers are arranged for forced draft; that is, the air for combustion is blown or forced into the boiler furnace by use of centrifugal, motor-driven fans. Coal firing generally uses a combination of forced and induced draft fans, and these boilers are operated such that the gas pressure in the boiler furnace is slightly negative with respect to the surrounding atmosphere.

The oil normally burned is a residual fuel called No. 6 fuel oil or sometimes referred to as Bunker C. This fuel has the lowest cost of liquid fuel since it is what is left over after the refining process, which has extracted better and cleaner products such as gasoline, diesel oil, and home heating oil.

The fuel from the ship's bunker tanks is generally transferred to a settling tank, which contains about a day's supply. While in the tank, the fuel is heated, and water and heavy solids such as sand can be removed by settling before the fuel is pumped at high pressure through fuel oil heaters and to the burners. The purpose of the fuel oil heaters is to reduce the viscosity of the oil so that it may be atomized for burning. To assist this mechanical atomizing process, most fuel oil burners utilize steam under pressure to further break down the oil particles so as to provide better combustion characteristics. The oil being atomized at the burner is then mixed with the heated combustion air so that the combustion process can take place in the furnace.

Coal is normally burned by using mechanical stokers and moving chain grates, although some consideration is being given to burning powdered coal, either in the solid form or as a powder mixed with oil in the form of a slurry.

Automatic controls. Automatic controls regulate the feedwater flow to the boiler, the superheater outlet steam pressure and temperature, the fuel oil flow, and the quantity of combustion air. Generally the controls are of the solid-state electric type that sends signals to the electric or air power devices for positioning of valves and dampers. *See* BOILER

FEEDWATER REGULATION; STEAM TEMPERATURE CONTROL.

Maintenance and repairs. Reliable and efficient performance of the boiler is required over the economic life of the ship, which is considered to be about 25 years. Normal maintenance and repairs are generally made while the ship is in port, while major repairs are made in a shipyard during the annual inspection periods. *See* DRYDOCKING.

Boiler firesides must be cleaned of soot and ash deposits to obtain high boiler efficiency and proper superheater steam temperature. For this purpose soot blowers using steam as cleaning media are fitted in the superheater and the steam-generating and heat recovery sections of the boiler. These surfaces also must be water-washed periodically.

Scale deposits and corrosion of the water sides of the boiler must be avoided. To accomplish this, chemically treated distilled water must be used for boiler makeup, and leaks of seawater to the boiler feedwater must be avoided. Periodic chemical testing of the boiler water and the feedwater must be conducted by manual or automatic means. *See* BOILER; BOILER FEEDWATER; STEAM-GENERATING UNIT.

Robert P. Giblon

Marine conservation

The management of marine species and ecosystems to prevent their decline and extinction. As in terrestrial conservation, the goal of marine conservation is to preserve and protect biodiversity and ecosystem function through the preservation of species, populations, and habitats. The importance of conserving marine species and ecosystems is growing as a consequence of human activities. Negative impacts on marine biological systems are caused by such actions as overfishing; overutilization, degradation, and loss of coastal and marine habitats; introduction of nonnative species; and intensification of global climate change, which alters oceanic circulation and disrupts existing trophic relationships. Such factors contribute to species decline and extinction in the sea. The growing human population puts direct pressure on marine biological resources through extraction of food and natural resources from the ocean; conversion of coastal habitats for housing, marinas, and other purposes; use of waterways for shipping and recreational purposes; intentional and unintentional translocation of exotic marine species; and pollution of coastal waters and their underlying sediments. In addition, humans put indirect pressure on coastal and marine systems through runoff, erosion, pollution, and eutrophication from terrestrial sources. Additional indirect pressures are imposed on marine systems, and on the entire biosphere, by the production of atmospheric emissions that contribute to global warming and increased penetration of ultraviolet radiation. Marine conservation biologists seek to reduce the negative effects of all these factors by conducting directed research and helping to develop management strategies for particular species, communities, habitats, or ecosystems.

The discipline of marine conservation has lagged behind terrestrial conservation in its development and application. This is due to many factors, including the logistic difficulties of working in the marine environment; practical difficulties of observing, counting, and tracking marine organisms and their offspring; and a relative lack of knowledge concerning many marine communities and habitats, especially those of the open ocean and deep sea.

A variety of approaches and tools are used in marine conservation. These include population assessment; mitigation, recovery, and restoration efforts; establishment of marine protected areas; and monitoring programs. Many of these approaches overlap with those in terrestrial conservation. However, fundamental differences between terrestrial and marine environments in spatial dimension, habitat type, and organismal life history require that basic conservation techniques be modified for application to the marine environment. *See* BIODIVERSITY; ECOSYSTEM; MARINE ECOLOGY; OCEANOGRAPHY.

Population assessment. Effective management requires knowledge of the size and status of populations. Trends in abundance can be detected through stock assessment methods first developed for marine fisheries and subsequently modified for application to other marine organisms. These methods use estimates of population size, reproduction, survivorship, and immigration to determine whether populations are increasing, decreasing, or stable. Small populations and populations with steeply declining numbers are typically thought to be at greater risk of extinction than larger, more stable populations. Population viability analysis is a specialized statistical assessment in which demographic and environmental information is used to determine the probability that a population will persist in a particular environment for a specified period of time. Population viability analysis can be used to guide management decisions, and has been used in efforts to manage marine mammals, turtles, seabirds, and other species. *See* ECOLOGICAL COMMUNITIES; ECOLOGICAL METHODS; POPULATION ECOLOGY.

Mitigation, recovery, and restoration. Depleted, threatened, or endangered populations are often subject to mitigation or recovery efforts. The purpose of these efforts is to reduce the immediate threat of extinction or extirpation. This is typically achieved by direct human intervention to increase the size of a population or to prevent further decline in population size. Recovery can be promoted by reducing juvenile and adult mortality and by restoring critical habitat (that is, areas essential for feeding or reproduction). Methods used to achieve recovery of fish and marine invertebrate populations include reducing fishing quotas, restricting the use of certain types of fishing gear, restricting the seasonal or annual distribution of a fishery, or closing fisheries altogether. For species such as some marine mammals, recovery has been effected by legislation that greatly limits or prohibits their capture or lethal take. The eastern North Pacific stock of gray whale and the Bering-Chucki-Beaufort stock of bowhead whale, both of which suffered severe depletion due to

overexploitation in the nineteenth and twentieth centuries, are examples of successful recovery efforts. In contrast, for reasons that are not fully understood, attempts to increase the number of Hawaiian monk seals have been largely unsuccessful, and the species remains in danger of extinction. *See* ENDANGERED SPECIES; FISHERIES ECOLOGY.

Recovery efforts can be most successful if they are based on multispecies or ecosystem-level management strategies. These strategies take into account positive and negative interactions between species, such as facilitation, competition, and predation. They further take into account interactions between species and their environment. An example of a multispecies interaction comes from the red sea urchin and the abalone, both of which have been the target of commercial and recreational fisheries in California. Juvenile abalone recruit most successfully beneath the spine canopies of adult urchins; thus, efforts to aid the recovery of abalone are most effective when the abundance of urchins is sufficiently high. An example of ecosystem-level interactions comes from the Steller sea lion, which has shown continued population decline in the Gulf of Alaska. Despite ongoing research and recovery efforts, the causes of this decline remain unknown, and attempts to increase the number of sea lions have been mostly unsuccessful. However, some believe that the decline has been caused in part by the combined effects of commercial fisheries that target the forage fish required by the sea lions, and by climate change, which has altered oceanic circulation patterns and thereby altered the distribution and reduced the availability of suitable prey.

Key to the success of assessment and recovery programs is identification of the appropriate biological unit for conservation (for example, population, subspecies, stock, or evolutionarily significant unit). Maintaining genetic diversity is an important goal of conservation biology, because genetic diversity confers evolutionary potential. Thus, conservation efforts often are aimed at populations that are genetically distinct from other populations of the same species. Species become and remain genetically distinct through reproductive isolation (that is, by lack of interbreeding). A good example of reproductive isolation in marine species is found among salmon stocks in the Pacific Northwest of the United States. In this region, salmon of diverse stocks and species intermingle while they are feeding at sea. However, members of each stock return to a particular stream or river system to reproduce, thereby creating a high degree of genetic isolation between groups. Effective conservation strategies must take into account this genetic distinction by managing distinct stocks separately.

Restoration efforts are aimed at returning habitats to an ecologically functional condition, usually consistent with some previous, more pristine condition. In the marine realm, restoration efforts have been limited to certain coastal and estuarine habitats. Seagrasses provide an important example of restoration efforts in coastal areas. Seagrasses are marine flowering plants that grow in large stands, or "meadows," in shallow bays and inlets. Seagrass communities are highly productive, and they perform essential ecosystem functions through provision of nursery areas for juvenile fish and other animals, and as feeding areas for both fish and birds. However, seagrass communities are being lost rapidly through coastal development. To offset these losses, restoration activities have been initiated in California, Texas, Florida, and other places. These efforts include transplanting seagrasses from healthy populations to sites where they can replace plants that were lost; regulating activities that contribute to habitat loss, such as the dredging or filling of tidelands; and modifying the design of docks and marinas to have a less deleterious effect on existing seagrass populations. *See* MARINE FISHERIES.

Marine protected areas. Marine protected areas are set aside for the protection or recovery of species, habitats, or ecosystems. They include marine parks, marine reserves, marine sanctuaries, harvest refugia, and voluntary or legislated no-take areas. Some marine protected areas allow for consumptive use (such as fishing) or extraction of resources (for example, oil drilling), while others are closed to most human activities.

The number of marine protected areas is growing. Two of the largest are Great Barrier Reef Marine Park in Australia, established to protect coral reef ecosystems, and the Galápagos Marine Park in Ecuador, which protects a unique temperate ecosystem. In the Antarctic, protection is provided to living and nonrenewable marine resources by principles codified in several international treaties. Other, smaller marine protected areas have been established to protect or promote the recovery of individual species or groups of species such as seahorses, dugongs, sea turtles, Mediterranean monk seals, Hawaiian monk seals, southern sea otters, rockfish, abalone, and seagrasses. However, many of these marine protected areas are so new that their effectiveness has not yet been demonstrated.

Monitoring programs. Monitoring programs are necessary to determine the outcome of specific conservation actions and to guide future conservation decisions. Monitoring programs vary according to the objectives of specific conservation projects but typically include such activities as long-term surveys of population size and status, and the development of mathematical models to help predict specific outcomes. A good example of monitoring efforts are those concerned with the California gray whale. Repeated surveys of migrating whales allowed observers to estimate the number of animals in the population as well as trends in abundance over a 30-year period. The data showed the California gray whale population to be growing steadily and sufficiently large to allow its removal from the U.S. List of Endangered and Threatened Wildlife in June 1994.

Terrie Klinger

Bibliography. T. S. Agardy, *Marine Protected Areas and Ocean Conservation*, Academic Press, San Diego, 1977; E. A. Norse (ed.), *Global Marine Biological Diversity*, Island Press, Washington, DC, 1993.

Marine containers

Standardized rectangular boxes for the transport of marine cargo. Ocean transportation of cargo containers began in 1956, when the converted tanker *Ideal-X* of Pan-Atlantic Steamship Company left Newark, NJ, for Houston, TX, for the first time. On the vessel's deck were fifty-nine 35-ft (11-m) containers holding cargo that until then had been brought to piers by truck or train, placed on pallets or in cargo slings, and lifted into the ships' holds. This method of handling cargo, called break-bulk, was slow and labor-intensive. Containers, in contrast, can be loaded at a remote site, sealed, moved to the pier, and lifted into the ship without the need for intermediate handling. The principal benefits other than reduced pierside handling cost are less cargo damage and reduced pilferage. *See* MERCHANT SHIP.

Containerization has been the most significant change ever made to cargo transportation by sea. The now ubiquitous container transportation system has made break-bulk shipping almost obsolete on the major trade routes of the world. Even on routes serving minor outports, where cargo volumes have been insufficient to justify the landside infrastructure required for the container system, the phenomenon of containerization is increasingly important.

Development. One aspect of the success of container transportation has centered on the standardization of equipment, which has allowed virtually any container to be carried anywhere in the world—by road, rail, or sea or any combination of these modes—without concern for the ability of the transport facilities to deliver the cargo, safely and untouched, to the ultimate destination, however remote it might be.

The predominant materials of construction are steel and aluminum, although a small number of containers have included reinforced plastic side panels. All are built around a steel frame to provide the requisite strength for lifting and stacking, in vessel holds and on deck, as well as during intermodal portions of the journey.

Through the mid-1960s, container transportation was marked by the use of boxes of a variety of lengths according to the needs of individual vessel operators and local statutory limits on over-the-road transportation. Containers with a length of 17, 20, 24, 30, and 35 ft (5.2, 6, 7.3, 9, and 11 m) were used during this period. These soon gave way to containers of International Standards Organization (ISO) dimensions, wherein lengths of 20 and 40 ft (6 and 12 m) were specified with heights of 8, 8.5, and 9.5 ft (2.4, 2.6, and 2.9 m). Width was standardized at 8 ft (2.4 m).

At a later date, containers 45 ft (14 m) long were introduced, and later still, the length was increased to 48 ft (15 m) and finally 53 ft (16 m). The width of the 48- and 53-ft containers is 8.5 ft (2.6 m). Although these non-ISO boxes originated in the United States, where highway transportation practices and most state laws permit their use, they are now common in many other parts of the world. In addition, attempts are being made to adapt the ISO standard containers to accommodate European standard pallets more efficiently; however, these have not been widely used in international shipping.

An increasing number of containers are specially fitted with racks and other devices for dedicated services, of which the transport of motor vehicles and vehicle components is the most common. When completed vehicles are carried in ISO standard containers, they are usually stowed on movable racks to accommodate the maximum load per box. A low-height 53-ft (16-m) container has been developed to permit motor cars to be driven in and secured without the need for special interior fittings.

Where wider-than-ISO containers [mainly 48-ft (15-m) boxes] are to be carried below deck, it has been necessary to construct dedicated cellular guides to accommodate these boxes. On deck, the problem is less onerous, for container support points on deck and on hatch covers can readily be installed. In intermodal traffic, a number of special rail cars will accept containers of either width by the use of adjustable corner fittings.

Although a majority of containers are used for general cargo, an increasing number of specialized boxes carry refrigerated or frozen meats, fish, fruit, and other foods. Refrigerated boxes are also used to carry critical medical supplies such as blood plasma and insulin. These containers are equipped with self-contained refrigeration units powered from the ship's electrical system. Instrumentation systems are used to monitor and control cargo temperature.

Tank containers are used to transport a vast array of oils, chemicals, and other liquids. Notable among these liquids are certain potable commodities and cryogenic gases, the latter typically carried in insulated, double-wall tanks.

High-cube boxes are widely used for low-density commodities such as apparel. Apparel is carried on hangers that may be moved direct to retail outlets without the need for packaging.

Half-height containers are used for special ladings, such as dense commodities. Open-top and open-flat containers are used for outsized loads, while a few double-height boxes have been used for the movement of outsized aircraft components.

Stowage and handling. Containers are carried below deck on ships having cellular holds fitted with steel guides at the four corners, which provide longitudinal and transverse restraint as the ship rolls and pitches at sea. No other restraint is required. Stacking containers up to nine high is common in the holds of the largest ships.

Deck loading, a unique characteristic of container ships, had been unavailable for most cargoes on break-bulk vessels. In the pioneering days of container transportation, deck stacks generally were limited to two tiers. As lashing methods and equipment were perfected, stack heights eventually rose to six high, consistent with adequate ship stability and unimpeded visibility from the wheelhouse. Many operators use the sixth tier to transport empty containers.

Common to all containers are standardized corner fittings, which not only provide structural strength to permit stacking but also provide anchoring points for lashing. Lockable fittings are employed at each corner to provide rigidity throughout the entire stack.

Lashings are conventionally chain or steel rods, with appropriate fittings at each end to mate with deck lashing points or container corners. The lashing assemblies include toggle or turnbucklelike devices to assist in securing the stack as an integral unit.

Several ship operators have substituted structural frames for a portion of the conventional lashing devices in order to minimize time and labor in making deck-loaded containers ready for sea. These lashing frames are anchored to fixed ship structure and typically are placed between the lowermost and second tier of a deck stack, thus eliminating the need for chain or rod lashings on these tiers.

A number of "hatch-coverless" container ships have been built for the purpose of reducing port time and the expense of handling hatch covers while loading and discharging. On such vessels, cell guides are extended upward above the hatch coamings to secure containers stowed above deck.

The carriage of hazardous cargo in containers has received much attention. Regulatory requirements and common sense dictate how and where containers loaded with such commodities should be stowed. An important consideration in transporting hazardous cargo in this manner is to ensure that cargo manifests are accurate and the guidelines for stowage are strictly followed.

Shoreside handling. In loading or discharging a container ship, individual boxes are typically lifted by shoreside or ship-mounted cranes using spreader frames having locking devices at the corners that engage the container corner fittings. The remote-operated locking devices are under the control of the crane operator. Under certain circumstances, two containers are handled simultaneously; two 20-ft (6-m) boxes in tandem may be lifted using a single 40-ft (12-m) spreader frame, or two empty boxes, one above the other and secured to one another by locking fittings, may be lifted.

Shoreside cranes are of the gantry type, capable of reaching to the extreme far side of the ship being worked. Over the first 30 years of the container era, shoreside cranes evolved to the centerpiece of an efficient handling for vessels of up to Panamax (106 ft or 32.2 m) beam, then the industry-wide standard. From the mid-1980s on, ships with a beam of up to 138 ft (42 m) have been built, making them unable to traverse the Panama Canal. As a result, the maximum outreach of newer shoreside cranes has been increased to handle these larger vessels. Cranes with an outreach of up to 180 ft (55 m) are being installed at major terminals, anticipating dimensions of future vessels. *See* HOISTING MACHINES.

Geared container ships, that is, vessels having shipboard cranes, are generally limited to feeder vessels and smaller line-haul ships that call at ports not having extensive container-handling terminals. Although some multipurpose vessels, which may also be used to carry break-bulk or bulk cargoes, have center-line mounted cranes, the cranes of geared container ships are typically installed on one side of the ship so that the vessel may be moored with cranes away from the pier in order to eliminate interference while containers are loaded or discharged by shoreside cranes.

Handling and movement of containers at shore terminals has received close attention in an effort to reduce terminal dwell time. Two separate systems have evolved: placing each container on its own tractor-drawn chassis for eventual dispatch by road, and handling the containers within the terminal with specially designed straddle carriers, some of which can transport two containers one above the other. At some terminals, on-pier marshaling yards permit the loading of rail cars direct from the vessel for intermodal shipment. In-terminal transfer of containers from ship to ship or barge is also used extensively.

The use of containers as an integral part of military logistics systems has received much attention, having been proven in both simulated and actual operations. Prepositioned military support vessels utilize containers for a major part of their cargo stowage. *See* HARBORS AND PORTS.

Intermodalism. The concept of intermodalism, by which containers may be transported successively by sea, rail, and road, was the premise upon which containerization was conceived and achieved its greatest success. Intermodalism was most emphatically demonstrated in the United States, where the growth of the industry coincided with the establishment of the interstate federal highway network.

The railroads, previously handicapped because of the atmosphere engendered by an archaic regulatory climate, eventually assembled an efficient container-on-flat-car transport system by which transocean containers could be carried between ocean terminals and inland points. Special rail cars were devised to permit up to two 40-ft (12-m) containers to be carried on each piece of rolling stock, and dedicated inland terminals were able to handle transfers between rail and road vehicles.

Although the rail system was reasonably efficient, it was not until the development of the double-stack rail car fleet that per-mile costs were decreased sufficiently to make rail transport a truly competitive alternative. Articulated, drop-center cars permitted containers (including high-cube boxes) to be carried one atop the other, thus doubling the load carried per unit of train length. The western railroads of the United States, enjoying more liberal loading-gauge restrictions, were the earliest beneficiaries of this concept. Many eastern railroads, with their more congested facilities, found it necessary to alter bridges, tunnels, and other lineside impediments to take advantage of the double-stack equipment. The concept is in use with success over much of the United States, bringing the heartland within easy reach of deep-water ports. *See* RAILROAD ENGINEERING.

Prospects. It is predicted that there will be significant increases in the level of both international and domestic container trade. Among the developments

in container transportation that must accompany increased traffic will be further increases in the size of the container ship fleet, individual ship capacity, more efficient terminals and intermodal connections, and changes in container technology.

Maximum ship capacity, now around 7000 TEU (twenty-foot equivalent units), may soon reach 8000 to 9000 TEU, and it is possible that true megacarriers, which carry over 10,000 TEU, may become the standard for major line-haul routes.

Maximum ship capacity is closely tied to the ability of shore terminals and intermodal facilities to handle container throughput efficiently. Hence, improvements in the way that containers are handled throughout the landside transportation system, including ship loading and discharging, will be made.

Critical to the ability of developed countries to handle increased traffic is the provision of adaquate and efficient port facilities. Assuming that container dwell time is constant, throughput at a terminal is a function of terminal area. The question must be raised: is sufficient additional area available at existing container ports to accommodate future growth? If not, is the construction of new ports, with appropriate intermodal connections, a viable alternative?

Heretofore, the physical size of container ships has permitted the operation of port-to-port services, in which vessels call at multiple ports at each end of an ocean voyage. Ship size has increased significantly from the mid-1980s, leading to the emergence of hub- and-spoke operations, in which large line-haul ships call at a limited number of hub ports. From these, smaller feeder ships (which may be as large as yesterday's line-haul vessels) distribute containers to nearby smaller outports. The expansion of the hub-and-spoke concept will continue as line-haul ship capacities increase to megacarrier (about 10,000 TEU) dimensions. The success of the hub port concept will depend upon the ability of the ship-port interface to decrease terminal dwell time to a commercially acceptable amount.

There is also likely to be a substantial increase in development of intermodal systems in parts of the world where rail and highway facilities are being expanded. Similarly, the intermodal system will need to be made more efficient in developed countries as cargo volume increases and existing road and rail facilities become overloaded.

Other changes are likely to be seen in the containers themselves. It is foreseen that collapsible containers will be developed for certain commodity loading to minimize the volume occupied in repositioning empty boxes, and specialized containers for dedicated cargoes will be utilized to a greater degree than that at present.

Rod Vulovic

Marine ecology

An integrative science that studies the basic structural and functional relationships within and among living populations and their physical–chemical environments in marine ecosystems. Although an outgrowth of natural history and the life sciences, marine ecology draws on all the major fields within the biological sciences as well as oceanography, physics, geology, and chemistry. *See* ENVIRONMENT.

Historically, and to some extent today, emphasis within the field has been directed at the gathering and analysis of data and descriptive information about taxonomy, species distributions and abundances, natural history and population biology, and the physical–chemical characteristics of marine environments. However, emphasis has evolved toward understanding the rates and controls on ecological processes that govern both short- and long-term events, including population growth and survival, primary and secondary productivity, and community dynamics and stability.

Marine ecology focuses on specific organisms as well as on particular environments or physical settings. Mangrove forests, salt marshes, seagrasses and seaweeds, and coral reefs form unique ecosystems defined by the dominant biological community. Lagoons, estuaries, bays, and the principal ocean environments of continental shelf, slope, abyssal plain, and mid-oceanic ridges form ecosystems based on physical–geological features. Together, these ecosystems occur in water depths ranging from the intertidal zone to the deepest ocean depths, about 33,000 ft (10,000 m), and encompass temperature extremes of 27°F (−3°C) in the deepest ocean waters to 104°F (40°C) in some tropical lagoons. *See* ATOLL; MID-OCEANIC RIDGE.

Marine environments. Classification of marine environments for ecological purposes is based very generally on two criteria, the dominant community or ecosystem type and the physical–geological setting. Those ecosystems identified by their dominant community type include mangrove forests, coastal salt marshes, submersed seagrasses and seaweeds, and tropical coral reefs. Plankton are a principal component in all marine ecosystems and include organisms ranging in size from viruses to larval fishes that permanently, or at some stage in their life cycle, reside in the water column. Other communities include benthic (bottom–dwelling) communities in shallow and deep water, occupants of upwelling areas, and the specialized community that inhabits deep-sea hydrothermal vents. Marine environments identified by their physical–geological setting include estuaries, coastal marine and nearshore zones, and open-ocean–deep-sea regions (**Fig. 1**). *See* DEEP-SEA FAUNA; ECOLOGICAL COMMUNITIES; HYDROTHERMAL VENT; PHYTOPLANKTON; ZOOPLANKTON.

Estuarine ecosystems. An estuary is a semienclosed area or basin with an open outlet to the sea where fresh water from the land mixes with seawater. The ecological consequences of fresh-water input and mixing create strong gradients in physical–chemical characteristics, biological activity and diversity, and the potential for major adverse impacts associated with human activities. Because of the physical forces of tides, wind, waves, and freshwater

input, estuaries are perhaps the most ecologically complex marine environment. They are also the most productive of all marine ecosystems on an area basis and contain within their physical boundaries many of the principal marine ecosystems defined by community type. *See* ESTUARINE OCEANOGRAPHY.

Salinity, temperature, submarine irradiance, that is, underwater light, and dissolved inorganic nutrients determine to large extent the biological composition of estuarine communities. In estuaries, salinity ranges from nearly fresh-water to oceanic concentrations. The actual salinity structure of estuaries varies both among and within estuaries and is determined by fresh-water inflow, tidal energy, and general estuarine circulation patterns. Estuaries are often classified into three groups, based on salinity structure (**Fig. 2**). The ecological significance of this is that most estuarine populations of plants and animals are physiologically limited by their salinity tolerance range, which restricts their distribution to particular zones. Temperature and submarine light play similarly important ecological roles by limiting or controlling important ecological processes. Temperature extremes generally limit the global distribution of species and seasonally determine the rates of metabolic processes within a given ecosystem. Submarine light plays an equally important role in governing both the rates of photosynthesis and the distribution of photosynthetic organisms. Dissolved inorganic nutrients, particularly nitrogen and phosphorus, not only limit the level of primary production but can influence species composition of estuarine communities. *See* SEAWATER.

Estuaries are further characterized by their principal ecosystem components, which often form the basic ecological unit of study. Salt marshes, mangroves, seagrasses, plankton, and benthic ecosystems are principal components of estuaries and vary in relative importance or even occurrence within specific estuaries. Salt marshes and mangroves are ecosystems that are dominated by vascular plants, occur worldwide, and occupy the intertidal zone. Salt marshes have their greatest distribution and abundance in the temperate latitudes, whereas mangroves are restricted to the tropics. Both play important roles in aquatic productivity, trophic structure, and nutrient cycling and serve as natural erosion-control structures. Seagrasses, which are submersed aquatic vascular plants, occur subtidally throughout the world and grow at depths ranging from just below the water surface to more than 30 ft (10 m) in some tropical areas. These plants, evolved from terrestrial ancestors, play many of the same ecological roles as the intertidal salt marsh and mangrove ecosystems, and are limited principally by the availability of submarine light. Planktonic and benthic ecosystems occur in all aquatic systems and for most estuarine environments have the greatest areal extent. Both are important for aquatic productivity and nutrient cycling and support, at least in part, economically valuable fisheries resources. *See* MANGROVE; SALT MARSH.

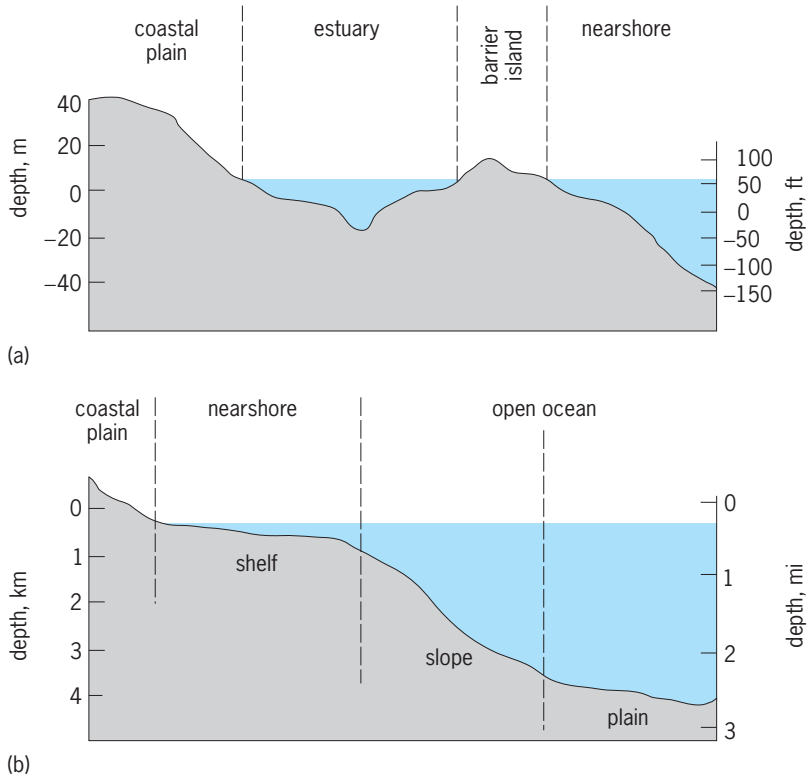

Fig. 1. Cross-sectional profile of (*a*) estuarine and nearshore and (*b*) coastal and open-ocean marine environments. The nearshore region, although not precisely defined, is generally within kilometers of the coastline. 1 m = 3.3 ft.

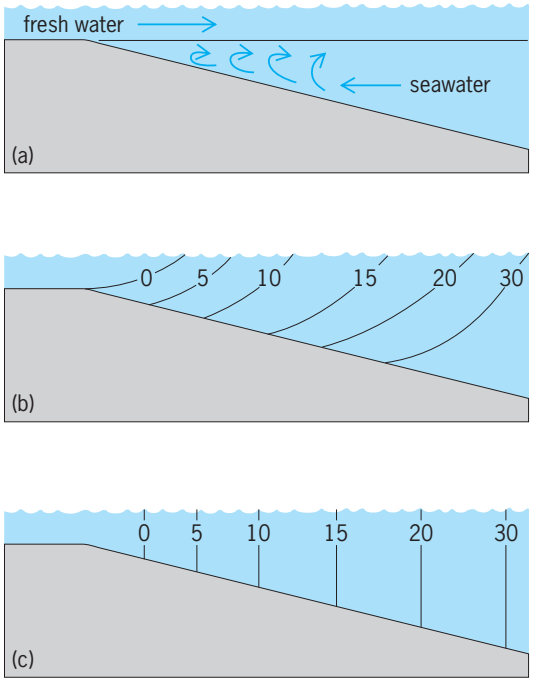

Fig. 2. Salinity structure of the three general types of estuaries found throughout the world. (*a*) Strongly stratified system due to high surface fresh-water inflow. (*b*) Moderated stratified system showing lines of equal salinity in parts per thousand. (*c*) Well-mixed estuary showing vertically homogeneous salinity structure in parts per thousand.

Coastal and nearshore ecosystems. Coastal and nearshore marine ecosystems are generally considered to be marine environments bounded by the coastal land margin (seashore) and the continental shelf 300–600 ft (100–200 m) below sea level. The continental shelf, which occupies the greater area of the two and varies in width from a few to several hundred kilometers, is strongly influenced by physical oceanographic processes that govern general patterns of circulation and the energy associated with waves and currents. *See* CONTINENTAL MARGIN.

Ecologically, the coastal and nearshore zones grade from shallow water depths, influenced by the adjacent landmass and input from coastal rivers and estuaries, to the continental shelf break, where oceanic processes predominate. Biological productivity and species diversity and abundance tend to decrease in an offshore direction as the food web becomes supported only by planktonic production.

Estuarine–nearshore couplings include the advective transport of nutrients and organic matter into nearshore waters, and seasonal migrations of ecologically and economically important species between coastal waters and estuaries. Advective transport contributes organic matter and nutrients, particularly nitrogen, which are thought to augment both primary and secondary production. Net primary production in nearshore waters has often been shown inadequate to support estimated secondary production, and so riverine and estuarine sources provide organic matter and nutrients. In seasonal migrations, many species of shellfish and finfish that spawn offshore migrate or are transported by water into estuaries during the larval or juvenile stage. There, they grow and mature and then migrate offshore as adults to complete the cycle.

Among the unique marine ecosystems associated with coastal and nearshore water bodies are seaweed-dominated communities (for example, kelp "forests"), coral reefs, and upwellings. Seaweed communities are distributed worldwide and occupy coastal areas from the intertidal zone to depths exceeding 30 ft (10 m). Green (Chlorophyceae), brown (Phaeophyceae), and red (Rhodophyceae) macroscopic algae dominate seaweed communities and show strong depth-dependent zonation patterns in response to various physical stresses, light requirements, and grazing pressure. Coral reefs are calcium carbonate structures made by reef-building (hermatypic) corals and crustose coralline algae and are best known for their high species diversity and structural complexity. They are limited to the tropical regions and require high light and temperature environments (greater than 68°F or 20°C). Upwelling regions are coastal areas where nutrient-rich bottom waters are moved toward the surface by persistent winds blowing parallel to the coast. As a result of the Coriolis effect and the Ekman spiral, surface water is transported offshore and replaced by bottom or middepth water, leading to a relatively continuous supply of nutrients to enhance primary production. Upwelling regions have historically been areas of economically important fisheries. *See* REEF; UPWELLING.

Open-ocean – deep-sea environments. Approximately 70% of the Earth's surface is covered by oceans, and more than 80% of the ocean's surface overlies water depths greater than 600 ft (200 m), making open-ocean–deep-sea environments the largest—yet the least ecologically studied and understood— of all marine environments. The major oceans of the world differ in their extent of landmass influence, circulation patterns, and other physical–chemical properties. Other major water bodies included in open-ocean–deep-sea environments are the areas of the oceans that are referred to as seas. A sea is a water body that is smaller than an ocean and has unique physical oceanographic features defined by basin morphology. Because of their circulation patterns and geomorphology, seas are more strongly influenced by the continental landmass and island chain structures than are oceanic environments.

Within the major oceans, as well as seas, various oceanographic environments can be defined. A simple classification would include water column depths receiving sufficient light to support photosynthesis (photic zone); water depths at which light penetration cannot support photosynthesis and which for all ecological purposes are without light (aphotic zone); and the benthos or bottom-dwelling organisms. Occupying the depths between the surface and the average depth of the world's oceans (12,000 ft or 4000 m) are a diverse group of organisms that include microbes, invertebrates, and vertebrates. Classical oceanography defines four depth zones; epipelagic, 0–450 ft (0–150 m), which is variable; mesopelagic, 450–3000 ft (150–1000 m); bathypelagic, 3000–12,000 ft (1000–4000 m); and abyssopelagic, greater than 12,000 ft (4000 m). These depth strata correspond approximately to the depth of sufficient light penetration to support photosynthesis; the zone in which all light is attenuated; the truly aphotic zone; and the deepest oceanic environments. The benthic fauna are classified in a similar manner as continental shelf and slope fauna, ocean basin or abyssal fauna, and fauna unique to the deep ocean basins and trenches.

The open-ocean–deep-sea environments present themselves as both a challenge scientifically and as a world resource that demands cooperative international research and attention. Estuaries and coastal ecosystems, as defined previously, are more localized and lend themselves to regional scientific exploration and understanding. Oceans and seas, on the other hand, are a global resource, crossing all political and national boundaries and requiring international cooperation and participation in ecological studies.

Marine ecological processes. Fundamental to marine ecology is the discovery and understanding of the principles that underlie the organization of marine communities and govern their behavior, such as controls on population growth and stability, quantifying interactions among populations that lead to persistent communities, and coupling of communities

to form viable ecosystems. The basis of this organization is the flow of energy and cycling of materials, beginning with the capture of radiant solar energy through the processes of photosynthesis and ending with the remineralization of organic matter and nutrients.

Photosynthesis and autotrophic production. Photosynthesis in seawater is carried out by various marine organisms that range in size from the microscopic, single-celled marine algae to multicellular vascular plants. Conversion of carbon dioxide into simple organic compounds and the subsequent synthesis of algal or vascular plant biomass forms the ultimate base of most food webs in the marine environment (omitted from this discussion are food webs based on chemosynthetic organisms). Because of that relationship, photosynthesis and primary production remain the focus of much marine ecological research. *See* FOOD WEB.

The rate of photosynthesis, and thus the growth and primary production of marine plants, is dependent on a number of factors, the more important of which are availability and uptake of nutrients, temperature, and intensity and quality of light. Of these three, the last probably is the single most important in governing primary production and the distribution and abundance of marine plants. Consequently, the optical physics of seawater has been an important component of much research on photosynthesis and primary production in marine environments.

The intensity and spectral quality of light in seawater is a function of incoming solar radiation, surface roughness, dissolved substances, suspended particulates, and water depth. Historically, measures of submarine light were based on photometric standards and units that had been developed for the human eye, such as footcandle, lumen, and lux, which have limited and questionable meaning in a biological context. Both physiologically and ecologically, light in the spectral region termed photosynthetically active radiation (PAR), at about 400–700 nanometers, is important and is measured as irradiance in units of energy flux (joules $\cdot$ $m^{-2} \cdot s^{-1}$) or photon flux (microeinsteins $\cdot$ $m^{-2} \cdot s^{-1}$), where one einstein (1 E) equals one mole of quanta or photons (6.02×10^{23}). As solar radiation passes through a water column, it is reduced in energy content and changed in spectral quality because of absorption and scattering. The total effect on intensity, termed attenuation, is probably the most common optical property of water reported and is derived according to the exponential decay function expressed in the equation

$$k_{1,2} = \frac{\ln(I_2) - \ln(I_1)}{z_2 - z_1}$$

where k is the attenuation coefficient between depths 1 and 2, I is the irradiance intensity at depths 1 and 2, and z is the depth. Generally, all depths are measured in meters below the water surface, and the attenuation coefficient is expressed in units of per-meter (m^{-1}). PAR attenuation varies considerably depending on water type, with attenuation

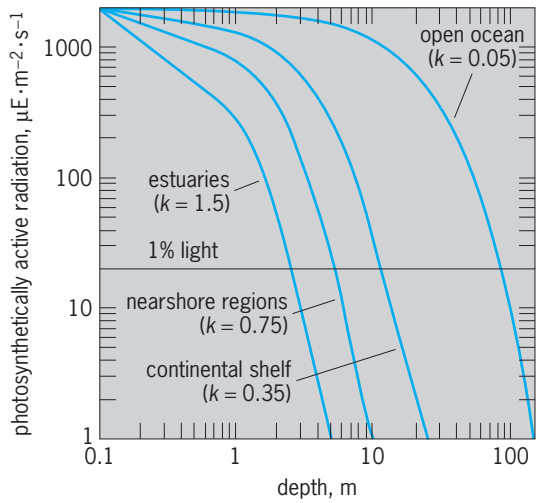

Fig. 3. Photosynthetically active radiation intensity at depth for various values of the attenuation coefficient (*k*). The intersection of the 1% light level (the limiting light intensity for photosynthesis) with the decay curves marks the depth limit for the photic zone. In these examples, the depth limit ranges from approximately 300 ft (100 m) in the open ocean to 10 ft (3 m) in estuaries. 1 m = 3.3 ft.

coefficients of 0.03–0.04 for the clearest open-ocean waters and 3–4 for the most turbid coastal and estuarine waters. PAR attenuation determines water depths at which light intensity can support photosynthesis (photic zone) and ranges from a few meters in estuaries and turbid coastal waters to approximately 450–600 ft (150–200 m) in the clearest open-ocean waters (**Fig. 3**).

The relationship between photosynthesis and light intensity varies among the different primary producers in marine environments but can be defined generally by a hyperbolic relationship. Characteristic of this relationship is (1) an initial positive, linear increase in photosynthesis with increasing light; (2) a nonlinear response above a threshold light intensity where photosynthesis does not increase linearly with increasing light; and (3) an upper limit to the rate of photosynthesis (P_{max}) at which increasing light has no effect and, if light intensity becomes high enough, may actually reduce or inhibit photosynthesis in some plants. **Figure 4** illustrates the photosynthesis–light relationships for two hypothetical marine plants that differ in these fundamental characteristics.

Autotrophic or primary production is the direct result of the biochemical and physiological processes of photosynthesis, but it is not equal to it, and results from the net positive growth of marine plants. Considering the high attenuation of light in water and the relationships between light intensity and photosynthesis, net autotrophic production is confined to relatively shallow water depths. The major primary producers in marine environments are intertidal salt marshes and mangroves, submersed seagrasses and seaweeds, phytoplankton, benthic and attached microalgae, and—for coral reefs—symbiotic algae (zooxanthellae). These principal autotrophic components of marine ecosystems differ in geographic distribution and abundance, rates of organic matter

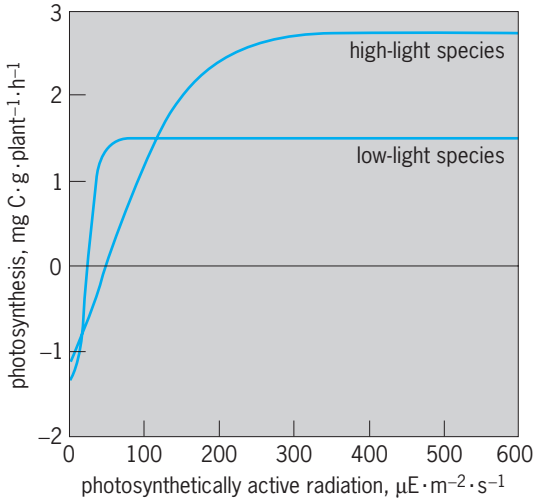

Fig. 4. Photosynthesis–photosynthetically active radiation curves for two marine plants with fundamentally different light relations. The high-light plant, which has a higher P_{max} but requires much more light, is typical of high-light environments such as those of the tropics and shallow water. The low-light plant has a lower P_{max} but requires much less light and photosynthetically saturates more quickly. These plants are typical of more turbid environments or environments where light is limiting for significant portions of the photoperiod.

production, and the food webs they support relative to secondary or heterotrophic production, including fishery yields. Though imprecisely known and highly variable for specific environments, net primary production for the principal marine environments of the world's oceans is summarized in the **table**. On an areal basis, estuaries and nearshore marine ecosystems have the highest annual rates of primary production. From a global perspective, the open oceans are the greatest contributors to total marine primary production because of their overwhelming size.

The two other principal factors that influence photosynthesis and primary production are temperature and nutrient supply. Temperature affects the rate of metabolic reactions, and marine plants show specific optima and tolerance ranges relative to photosynthesis. Nutrients, particularly nitrogen, phosphorus, and silica, are essential for marine plants and influence both the rate of photosynthesis and plant growth. For many phytoplankton-based marine ecosystems, dissolved inorganic nitrogen is considered the principal limiting nutrient for autotrophic production, both in its limiting behavior and in its role in the eutrophication of estuarine and coastal waters. *See* PHOTOSYNTHESIS.

Food webs and heterotrophic production. The concept of a marine food web and heterotrophic, or secondary, production is based on the view that energy, as well as organic matter and nutrients, is transferred between organisms that occupy different levels in a heirarchical trophic or feeding structure. The complexity and specific interactions among populations within this conceptual framework form the basis of much marine ecological research and are a source of controversy among marine ecologists. What has evolved is that marine food webs and the processes leading to secondary production of marine populations can be divided into plankton-based and detritus-based food webs. They approximate phytoplankton-based systems and macrophyte-based systems.

The classical view of the open-ocean food web involved planktonic algae (diatoms) being grazed by planktonic crustaceans (copepods), which were in turn preyed upon by other invertebrates and fish. Based on size, the final links in the food chain were relatively large organisms, and the classical view dictated that both energy flow and productivity in the oceans were due to these larger populations. To a large extent, this view resulted from the sampling gear and techniques available at the time. This simple, linear food-chain model has been replaced by recognition of the role played by smaller organisms, both autotrophic and heterotrophic, in primary and secondary production, and the cycling of carbon and nutrients. For planktonic food webs, current evidence suggests that primary production is partitioned among groups of variously sized organisms, with small organisms, such as cyanobacteria, playing an equal if not dominant role at times in aquatic productivity. The smaller autotrophs, both through excretion of dissolved organic compounds to provide a substrate for bacterial growth and by direct grazing by protozoa (microflagellates and ciliates), create a microbially based food web in aquatic ecosystems that had once been largely ignored. The consequences of the revised view are that the major portion of autotrophic production and secondary utilization in marine food webs may be controlled,

Areal and total estimates for net primary production in the principal marine environments of the world

Region	Area 10^6 km^2*	Net primary production, g · m^{-2} · yr^{-1}	Total net primary production 10^9 t · yr^{-1}	Percent of area	Percent of total net primary production
Open ocean	332	125	41.5	91.9	77.7
Continental shelf	27	300	8.1	7.5	15.2
Estuaries	1.4	1500	2.1	0.4	3.9
Seaweeds and coral	0.6	2500	1.5	0.2	2.6
Upwelling regions	0.4	500	0.2	0.1	0.4
Total	361.4		53.4		

*1 km^2 = 0.4 mi^2.

not by the larger organisms typically described as supporting marine food webs (including those leading to the major coastal and oceanic fisheries), but by microscopic populations, thus giving rise to the paradigm of the so-called microbial loop.

Macrophyte-based food webs, such as those associated with salt marsh, mangrove, and seagrass ecosystems, are not supported by direct grazing of the dominant vascular plant but by the production of detrital matter through plant mortality. The classic example is the detritus-based food webs of coastal salt marsh ecosystems. These ecosystems, which have very high rates of primary production (see table), enter the marine food web as decomposed and fragmented particulate organics. The particulate organics of vascular plant origin support a diverse microbial community that includes bacteria, flagellates, ciliates, and other protozoa. These organisms in turn support higher-level consumers, which include gastropods, mollusks, polychaetes, crustaceans, and fish. The high productivity of estuaries, particularly those in the Mid-Atlantic temperate regions of North America and Europe and the northern coasts of the Gulf of Mexico, results not only from phytoplankton production in the estuaries themselves but also from the contribution of surrounding salt marsh ecosystems.

The highest levels of both primary and secondary production are associated with estuarine and nearshore coastal marine ecosystems. Open-ocean systems are less productive on a per-area basis, and both pelagic (water column) and benthic food webs in deep ocean environments depend on primary production in the overlying water column. For benthic communities, organic matter must reach the bottom by sinking through a deep water column, a process that further reduces its energy content. Thus, in the open ocean, high rates of secondary production, such as fish yields, are associated with areas in which physical–chemical conditions permit and sustain high rates of primary production over long periods of time, as is found in upwelling regions.

Regardless of specific marine environment, microbial processes provide fundamental links in marine food webs that directly or indirectly govern flows of organic matter and nutrients that in turn control ecosystem productivity and stability. *See* BIOLOGICAL PRODUCTIVITY; ECOLOGY; ECOSYSTEM; SEAWATER FERTILITY. Richard Wetzel

Bibliography. J. W. Day, Jr., et al., *Estuarine Ecology*, 1989; M. L. Dring, *The Biology of Marine Plants*, 1982; J. W. Nybakken, *Marine Biology: An Ecological Approach*, 5th ed., 2000; I. Valiela, *Marine Ecological Processes*, 2d ed., 1995.

Marine engine

An engine that propels a waterborne vessel. In all except the smallest boats, the engine is but part of an integrated power plant, which includes auxiliary machinery for propulsion engine support, ship services, and cargo, trade, or mission services. Marine engines in common use are diesel engines, steam turbines, and gas turbines. Gasoline engines are widely used in pleasure craft. Only the propulsion engine will be discussed in this article. *See* BOAT PROPULSION; INTERNAL COMBUSTION ENGINE; MARINE MACHINERY.

Factors in engine selection. The engine type selected for an application should depend on its characteristics, usually with reliability of prime importance. Other factors may include weight, volume, quality and type of fuel required, efficiency, maintenance requirements, durability, simplicity in installation and operation, noise and vibration, initial cost, and ability to sustain the required power levels. In some applications, selection has been made on the basis of expedience or convenience, usually because of the ready availability of an engine type; this explains the continued use of gasoline engines in pleasure craft.

Weight and volume. These are usually the dominant criteria in warship applications and are also very important in fast merchant ships. Efficiency, or fuel economy, is tied to weight and volume of the machinery, since it is the range of operation that determines the amount of fuel that must be carried, by which the ship's payload is correspondingly diminished. In large, relatively slow merchant ships, machinery weight and volume are less important.

Fuels. Mostly, liquid petroleum fuels are used, with distillate fuels (including diesel fuel and gas oil) being the highest quality but most expensive, and residual fuel the cheapest but most difficult to use. Blended fuels of intermediate characteristics are often used, generally tending more toward the residual in cost and difficulty of use. While fuel-oil costs have varied widely over time, heavy blends and residuals have tended to cost about half as much as distillates, although experience has demonstrated an inverse correlation between maintenance cost and fuel quality. Coal and natural gas have application in certain trades, and the use of both is considered likely to increase in the future. Nuclear fuel is no longer considered viable for merchant ships, for economic and social reasons, but remains the dominant choice for the largest surface warships and for long-range submarines. *See* DIESEL FUEL; FUEL OIL; SHIP NUCLEAR PROPULSION.

Power level. Power-level requirements depend on the application. One extreme is represented by most merchant trades, where an engine is required to sustain high power output between ports, without relief until arrival, and yet be capable of fairly rapid response over the whole power range while the ship is maneuvered. Another extreme exists with warships, where sustained performance is usually required at a low power level yielding a modest cruise or endurance speed, as well as at a much higher power for high speeds. The difference in power requirements is so great that two engines of different type may be selected. *See* SHIP POWERING, MANEUVERING, AND SEAKEEPING.

Propellers. Most often, marine engines drive propellers, which must be operated at rotative speeds lower than those of turbines and of all but the

largest diesel engines. A speed-reducing transmission is therefore inherent with most marine engines, with reduction gearing the most common type, and electric drive or, in some small installations, belt or hydraulic drive also used. *See* PROPELLER (MARINE CRAFT).

Diesel engines. Diesel engines of all types and power outputs are in use for propulsion of most merchant ships from the smallest to the largest, most service and utility craft, most naval auxiliary vessels, and most smaller surface warships and shorter-range submarines.

The diesel engines most commonly used fall into either a low-speed category or the medium- and high-speed category. Low-speed engines (**Fig. 1**) are generally intended for the direct drive of propellers without any speed reduction, and therefore are restricted to a range of rotative speeds for which efficient propellers can be designed, generally below 300 revolutions per minute (rpm), and as low as 50 rpm at rated power. Low-speed engines are two-stroke, crosshead engines, with 4–12 cylinders, which are always in-line, turbocharged, and aftercooled. The stroke-to-bore ratio of low-speed engines in current production is about 3:1. The largest engines are rated for power output of over 5000 kW (almost 7500 horsepower) per cylinder at about 100 rpm. These engines are heavy and very large, but they are well suited to operation on low-quality fuels and generally require only modest levels of maintenance. *See* TURBOCHARGER.

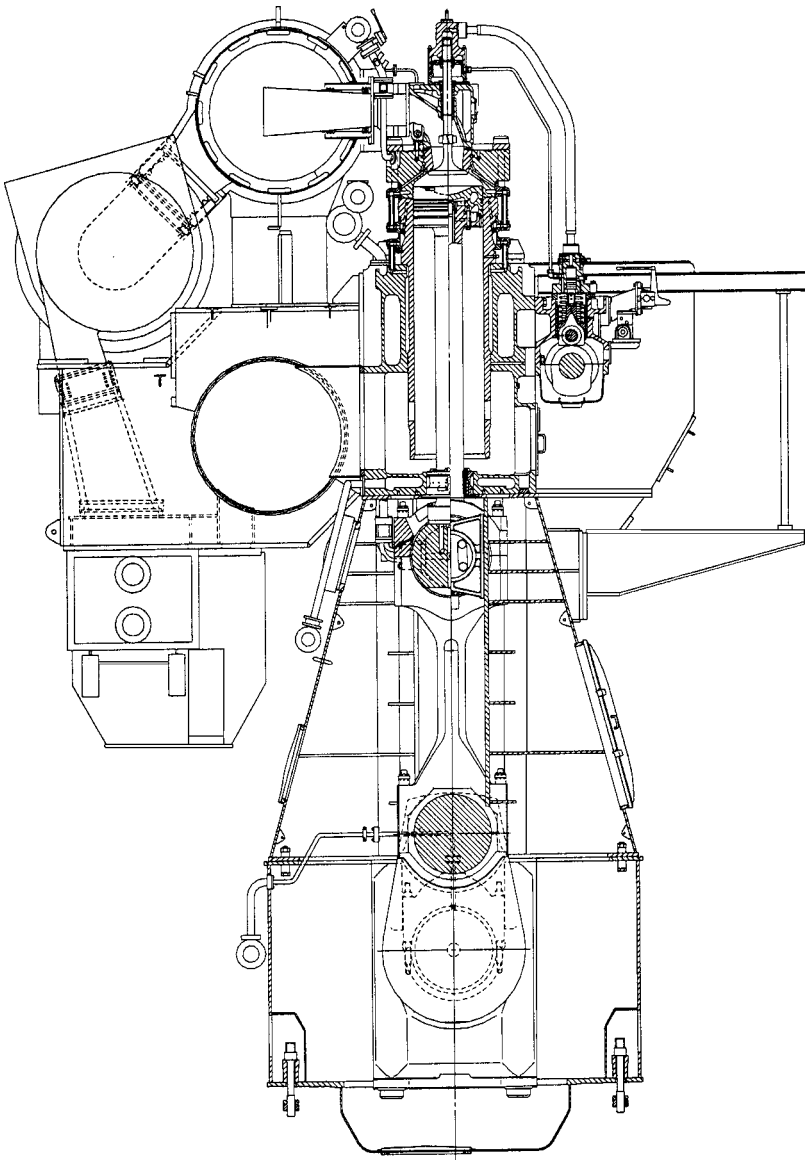

Fig. 1. Cross section of a low-speed diesel engine.

Because of their higher rotative speeds, medium- and high-speed engines (**Fig. 2**) drive propellers through speed-reduction gears, but they are directly connected for driving generators in diesel-electric installations. With few exceptions, these are four-stroke, trunk-piston engines, which have up to 10 cylinders in-line or up to 24 in a V configuration and are mostly turbocharged and aftercooled. Stroke-to-bore ratios range from 1:1 to 1.5:1. Large medium-speed engines are capable of over 1500 kW (2000 hp) per cylinder at about 400 rpm. The upper limit of the medium-speed category, and the start of the high-speed category, is generally placed in the range of 900–1200 rpm, but there are no clear physical features that enable the distinction to be made. Many of these engines have a proven heavy-fuel capability, but most evidence indicates that maintenance costs are higher than those of low-speed engines that are run on fuels of similar poor quality. Some engines, especially those in the higher-speed category, are restricted to distillate fuels.

Outside of the typical categories cited above are four-stroke, trunk-piston engines that are designed to be directly connected to propellers, and are indigenous, in both manufacture and application, to the Far East. These engines are built with six or eight cylinders in-line; they are rated for 70–700 kW (95–950 hp) per cylinder at speeds of 200–500 rpm. Another exception is a series of two-stroke, trunk-piston, medium-speed engines that dominate their field of application in United States waters, despite their requirement for distillate fuels. The highest rated of these engines, with 20 cylinders, has an output of more than 3500 kW (over 4500 hp) at 900 rpm.

All of the low-speed diesel engines and many of the medium- and high-speed engines are direct-reversing engines, that is, capable of being adjusted, after being stopped, so that they can be started and run in the opposite direction. Nonreversing engines use reversing gearing, if reversing is not provided by electric drive or a controllable-pitch propeller.

Most diesel engines have complex vibration characteristics and a tendency toward noisy operation, but much can be done to prevent these characteristics from becoming problematic. All diesel engines consume lubricating oil, generally at a rate of less than 1% of the fuel consumption. The thermal efficiency, which in the largest engines can exceed 45% even without the heat-recovery equipment generally used, is the highest of all current alternatives. The largest, low-speed engines offer this efficiency from

low-quality fuel oils, with modest maintenance requirements and relative simplicity but with penalties of high weight, high volume, and high initial cost. The largest medium-speed engines offer comparable efficiencies but with lower weight, volume, and cost, and with a trade-off between fuel quality and maintenance costs. High-speed engines are lighter and more compact and have good efficiency, but generally require good fuels if intensive maintenance is to be avoided. *See* DIESEL ENGINE.

Steam turbines. While steam-turbine plants cannot achieve the thermal efficiency of diesel engines, steam turbines of moderately high power levels (above about 7500 kW or 10,000 hp) offer efficient energy conversion from steam, which can in turn be produced by combustion of low-quality fuel oil, coal, or natural gas in boilers, or from a nuclear reactor. For high efficiency, high turbine speeds are required, typically 3000–10,000 rpm, with reduction gearing or electric drive used to achieve low propeller rotative speeds. The combination of turbine and reduction gear or electric drive has usually proven robust and durable, so that most oil-fueled steamships currently in service are held over from an earlier era. Others, more recently built, are capitalizing on the availability, in their trade, of a fuel unsuitable for diesel engines.

In modern merchant ship plants, typical steam conditions at the turbine are a pressure of about 6 megapascals (870 lb/in.^{2}), superheated to about 510°C (950°F), with exhaust to a vacuum condenser at about 5 kilopascals (1.5 in. Hg). Some high-powered plants have used reheat turbines, but generally cycle improvements are confined to two to four stages of regenerative feed heating from turbine extractions, and to economizers or rotary regenerative air heaters. *See* MARINE BOILER.

In some applications, especially with electric drive, single-cased turbines are used, but more common is the division of the turbine into a small, high-speed, high-pressure unit and a larger, lower-speed, low-pressure unit, each with 8–10 stages, independently connected to a single propulsion gear set (**Fig. 3**). This arrangement, known as a cross-compound turbine, provides more flexibility in design, allowing the high- and low-pressure stages to be better matched to the properties of the steam, which expands to about 400 times its initial volume from the throttle conditions cited above, to exhaust. For related reasons, impulse blading is favored for the high-pressure stages, with lower-pressure stages having increasing reaction. The first high-pressure stage is the control stage, and has partial admission, with the nozzles in groups, each controlled by a valve, enabling power output to be regulated by the number of nozzles in use. Throttle valve control is reserved for maneuvering. For reversing, separate astern stages are usually incorporated in the casing of the low-pressure turbine (Fig. 3), if reversing is not provided by an electric drive. *See* IMPULSE TURBINE; REACTION TURBINE; STEAM TURBINE.

Gas turbines. Aircraft-derivative gas turbines have become the dominant type of propulsion engine for medium-sized surface warships, including frigates, destroyers, cruisers, and small aircraft carriers. In all cases the turbines are multishaft, simple-cycle engines, with the power turbine geared to the propeller. In some installations, two to four turbines are the sole means of propulsion; in other cases, one or two turbines provide high-speed propulsion, while

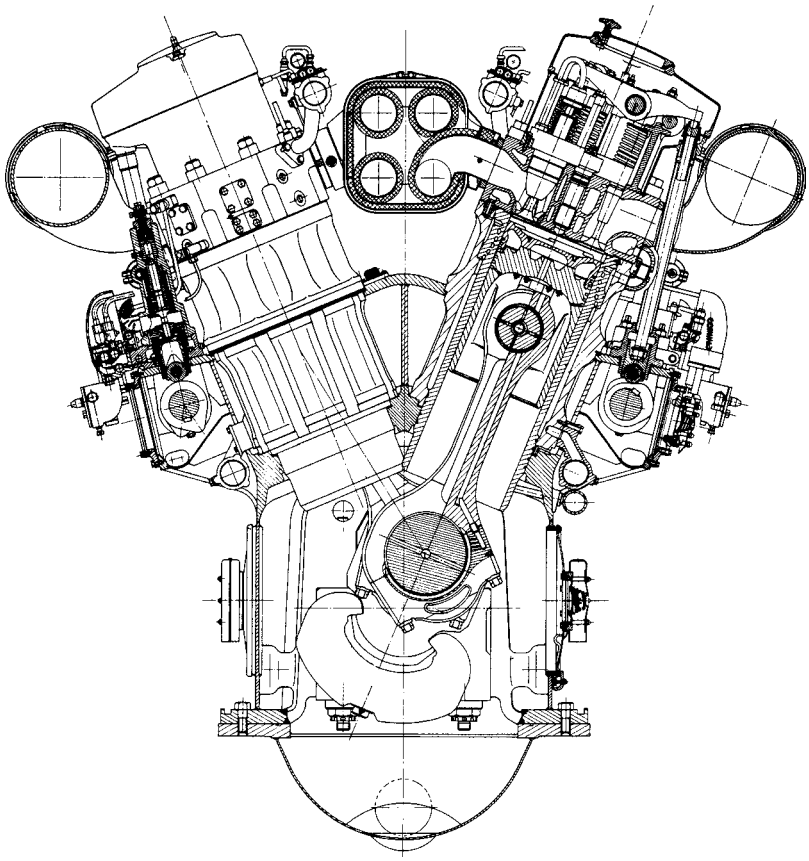

Fig. 2. Cross section of a four-stroke-cycle V-engine.

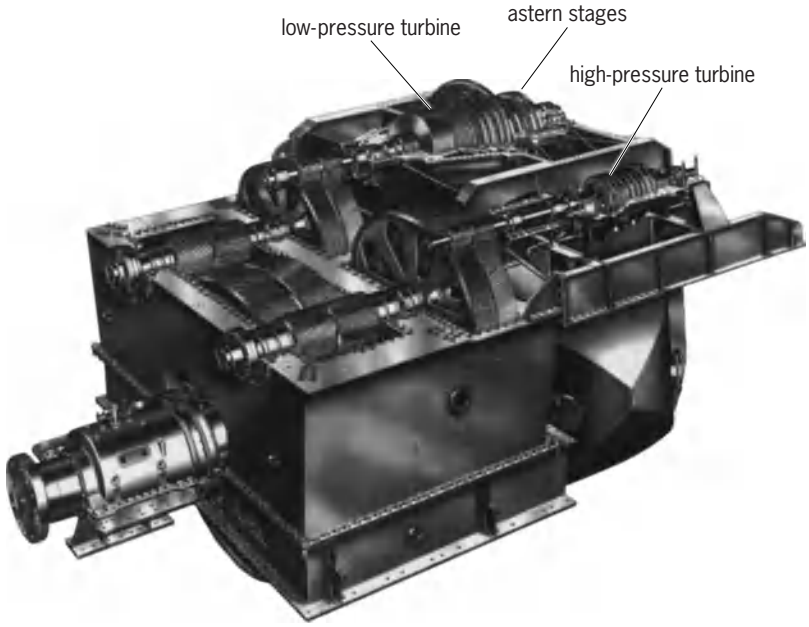

Fig. 3. Typical cross-compound steam turbine, with a double reduction gear, shown with upper halves of the turbine and gear casings removed. (*Westinghouse Electric Corp.*)

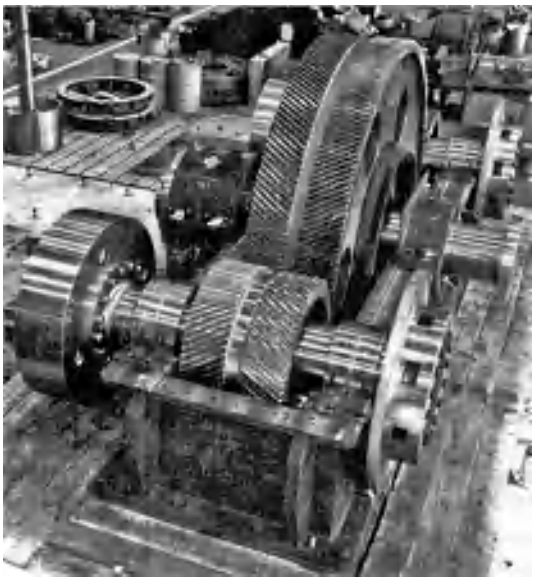

Fig. 4. Gearset for a twin-engine, single-screw ship. (*Philadelphia Gear Corp.*)

diesel engines or smaller gas turbines are used for cruising speeds.

Factors favoring the aircraft-derivative gas turbine in this application are low weight, compact dimensions, high power, rapid start and response, standardization of components, and maintenance by replacement. The last two features are especially important in warships, since an engine can be exchanged in 2 days or less and overhauled by civilian experts ashore, permitting a smaller on-board crew. A major disadvantage of aircraft-derivative gas turbines, their requirement for high-quality distillate fuels, is of diminished importance in the warship application, since most warships, including steamships, use these fuels anyway. Other disadvantages include poorer performance in warmer climates, decreased efficiency at power levels below maximum, large air intake and exhaust ducts, and high exhaust temperatures. Reversing thrust must be provided in the gearbox, by electric drive, or by a controllable-pitch propeller.

A further evolution of the aircraft-derivative gas turbine is an intercooled, regenerative-cycle unit, which offers higher thermal efficiency across the entire upper power output range, while retaining the other advantages cited above. Past applications of aircraft-derivative gas turbines and of simple-cycle and regenerative-cycle heavy-duty gas turbines to ships other than warships have had only limited success, and many of these ships were later converted to diesel propulsion. More recently, gas turbines have been used for very fast merchant ships. *See* GAS TURBINE; TURBINE PROPULSION.

Mechanical reduction gears. Marine propulsion gearing is usually of the parallel-shaft, helical-tooth type. Single-reduction gearing (**Fig. 4**) is usually

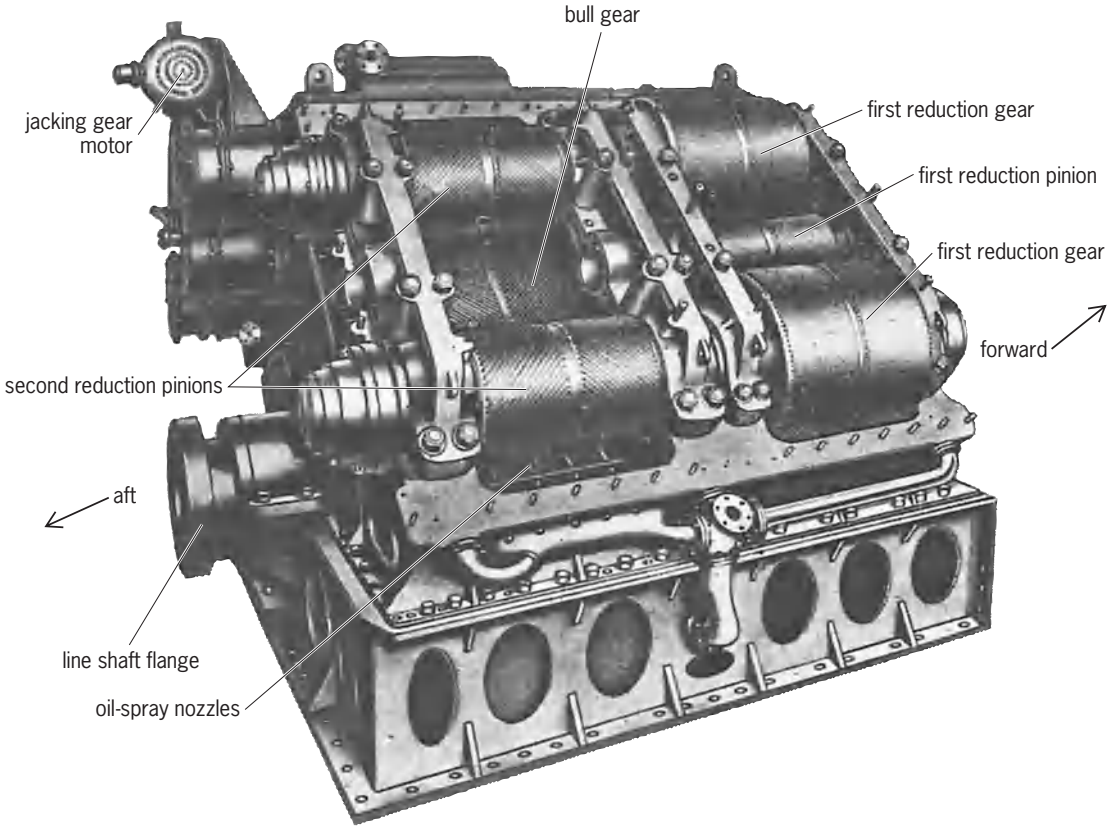

Fig. 5. Locked-train, double-reduction gearing for use with a cross-compound steam turbine or twin gas turbines. (*Naval Machinery, U.S. Naval Academy, 1946*)

adequate for medium- and high-speed diesels, where the rotative speed ratio ranges from about 2:1 to 10:1. In a typical arrangement (Fig. 4), two engines drive a single propeller, with each engine connected to its pinion via one or more flexible couplings to reduce torsional vibration, and a clutch, which is usually pneumatic or hydraulic. Both double-helical teeth (Fig. 4) and single-helical teeth are common.

The high speed of steam and gas turbines requires rotative speed ratios which can reach 100:1, so that double-reduction gearing is necessary (Fig. 3). For higher powers, with both steam and gas turbines, locked-train, double-reduction gearing is used (**Fig. 5**), which divides the input torque into parallel paths, permitting smaller, if more numerous, gears than would otherwise be required. Couplings of limited flexibility are provided between the first and second stages.

Epicyclic gearing has been successfully used in propulsion applications, most often for high-speed stages in double- or triple-reduction steam turbine gearing. Two-speed gears, permitting a selection of rotative speed ratio, have been fitted in low-power applications, with diesel engines, where there is a large range of required ship speeds.

Propulsion gearing is constructed to extremely close tolerances, with teeth of modified involute form, formed by hobbing or grinding processes, and either through-hardened or case-hardened, depending on the severity of the stresses. Forced lubrication of bearings, meshing teeth, and mechanical couplings is provided. Transmission efficiencies are high, with losses rarely exceeding 1–2% per stage. *See* GEAR; GEAR TRAIN.

Electric drive. In this arrangement, the engine is directly coupled to a generator, and the electricity produced drives an electric motor, which is most often of sufficiently low rotative speed to be directly connected to the propeller shaft. Any number of engine-generator sets may be connected to drive one or more propulsion motors. Electric drive has been used with engines of all types, including low-speed diesels. Advantages of electric drive include flexibility of machinery arrangement, elimination of gear noise, high propeller torque at low speed, and inherent reversing capability. In ships with high electric requirements for cargo, mission, or trade services—for example, passenger ships, tankers with electric-motor-driven cargo pumps, or warships with laser weapons—there is an advantage in integrating propulsion and ship service support through a common electric distribution system. However, electric drive is usually heavier, higher in initial cost, and less efficient than direct or geared drive.

Control arrangements. Automation has been widely applied to marine vehicles of all types; propulsion is normally under direct control from the wheelhouse. Alan L. Rowen

Bibliography. R. L. Harrington (ed.), *Marine Engineering*, 1992; S. H. Henshall, *Medium and High Speed Diesel Engines for Marine Use*, 1972; E. C. Hunt et al., (eds.), *Modern Marine Engineer's Manual*, vol. 1, 3d ed., 1999, vol. 2, 2d ed., 1991; C. M. Plumb, *Warship Propulsion System Selection*, 1987; D. A. Taylor, *Introduction to Marine Engineering*, 2d rev. ed., 1996.

Marine engineering

The engineering discipline concerned with the machinery and systems of ships and other marine vehicles and structures. Marine engineers are responsible for the design and selection of equipment and systems, for installation and commissioning, for operation, and for maintenance and repair. They must interface with naval architects, especially during design and construction.

Marine operating engineers form part of the crew of large ships, where they are responsible for the operation and maintenance of the machinery. Aboard merchant ships, operating engineers are ship's officers licensed by government agencies (in the United States, by the Coast Guard).

Marine engineers are likely to have to deal with a wide range of systems, including diesel engines, gas turbines, steam turbines, boilers, heat exchangers, and pumps and compressors; electrical machinery; hydraulic machinery; refrigeration machinery; steam, water, fuel oil, lubricating oil, compressed gas, and electrical systems; equipment for automation and control; equipment for fire fighting and other forms of damage control; and systems for cargo handling. Many marine engineers become involved with structural issues, including inspection and surveying, corrosion protection, and repair.

Marine engineers are generally mechanical engineers or systems engineers who have acquired their marine orientation through professional experience, but programs leading to degrees in marine engineering are offered by colleges and universities in many countries. Most marine operating engineers are graduates of these programs, although others have obtained their training through experience aboard ship. *See* BOAT PROPULSION; MARINE BOILER; MARINE ENGINE; MARINE MACHINERY; MARINE REFRIGERATION; NAVAL ARCHITECTURE; PROPELLER (MARINE CRAFT); SHIP DESIGN; SHIP POWERING, MANEUVERING, AND SEAKEEPING. Alan L. Rowen

Bibliography. C. Drummond, *The Remarkable Life of Victoria Drummond, Marine Engineer*, 1995; R. L. Harrington (ed.), *Marine Engineering*, 1992; E. C. Hunt (ed.), *Modern Marine Engineer's Manual*, 1991; D. A. Taylor, *Introduction to Marine Engineering*, 1990.

Marine fisheries

The harvest of animals and plants from the ocean to provide food and recreation for people, food for animals, and a variety of organic materials for industry. Important products in marine capture fisheries include fish; mollusks such as oysters, clams, and squid; and crustaceans such as crabs and shrimps.

Marine mammals (whales) and reptiles (turtles) have been important historically in marine landings (the weight of the catch landed at the whart); and some plants, mostly seaweeds, are harvested in significant amounts. World landings began to increase substantially during the industrial revolution and rose at a rapid rate after World War II before leveling off in recent years.

Marine Resources

The oceans' fish resources were once thought to be inexhaustible, essentially immune to the effects of fishing. As technological advances were made, bigger, more powerful vessels ranged ever farther from their home ports to exploit valuable fishery resources. The advent of steam and diesel engines, freezing capacity, and fish-finding capabilities added to the fishing power of specialized fleets, which by the 1960s fished for extended periods throughout the world's oceans. By then, it was clear that fishing was a major factor controlling abundances of fish in the sea.

Fishing is the last major food-hunting industry in the world in which wild animals are exploited in large quantities by nations with market economies. Marine fisheries provide >10% annually of animal protein consumed by humans, and additional large amounts of fish are processed to be fed to domestic livestock. As practiced in recent decades, unregulated or poorly managed fisheries can lead to collapse of major fish stocks. *See* FISHERIES ECOLOGY; FOOD MANUFACTURING.

The annual marine fish catch, not including aquaculture and exclusive of plants and marine mammals, increased from 18.5 million metric tons (MMT) in 1950 to more than 93 MMT in 1997. Increases were particularly rapid in the 1960s as the rich ocean upwelling regions off Peru came under exploitation, primarily for anchoveta (*Engraulis ringens*). Global fleets capable of industrial-scale fishing anywhere in the world dominated during this expansion phase (**Fig. 1**). Marine fisheries landings increased 7.2% per year during the 1960s, spurring false optimism that marine fish catches would ensure animal protein for a fast-growing human population. In the period 1969–1973, temporary drops in global landings were recorded, a reflection of combined heavy fishing and unfavorable El Niño climate conditions that devastated the anchoveta population off Peru, which had supported the world's largest single-species fishery, peaking at more than 12 MMT in 1970. Subsequently, marine fish catches rose again, increasing at 1.3% annually in the period 1985–1997 (see **table**). This rate is less than the rate of human population growth in that period. *See* EL NIÑO; SEAWATER FERTILITY; UPWELLING.

Most scientists believe that the global potential for marine fish catches has essentially been reached unless there is a major shift in fishing strategy. For traditional fisheries, approximately 100 MMT appears to be the limit of sustainability, and even that level depends upon effective habitat management and protection. A shift from fishing mostly large, carnivorous species to primarily species lower on the food chain that feed on plants or small animal plankton (which is already occurring to an extent) potentially could increase global landings. Targeting resources now hardly exploited, such as lanternfishes (family Myctophidae) in the deep sea or the shrimplike krill (euphausids) in the Antarctic, could also increase global landings. Whether developing such technologies will allow such shifts in strategy, whether the economics and marketing will be favorable, and whether it is ethically acceptable to develop these

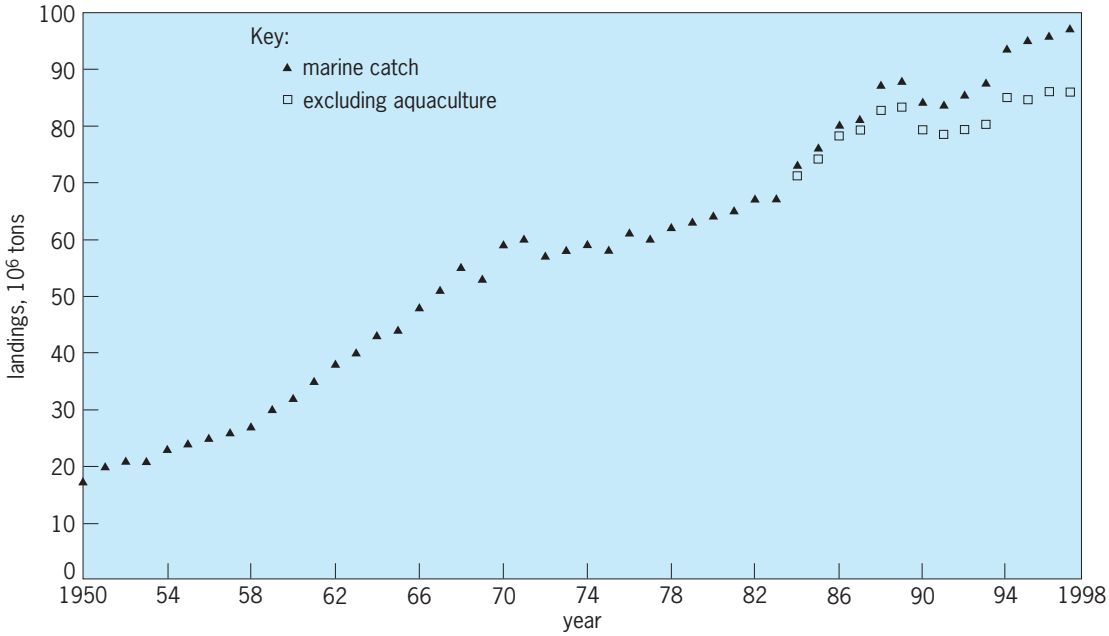

Fig. 1. Global landings from marine capture fisheries and marine aquaculture production, 1950–1997. Marine aquaculture contributed only small amounts until the mid-1980s. Marine plants (8.4 million tons harvested, mostly from aquaculture, in 1997) are not included in the graph. (*Data from United Nations, Food and Agriculture Organization, Rome*)

World fisheries, 1985–1997, in millions of tons*

	1985	1990	1995	1997	Rate of increase, % per year
Inland					
Capture	5.7	6.5	7.2	7.7	2.5
Aquaculture	5.0	8.2	13.8	17.1	10.5
Marine					
Capture	73.2	79.0	84.7	86.0	1.3
Aquaculture	2.7	4.2	7.2	11.2	11.9
Total world	86.6	98.0	112.9	122.0	2.9
Human population (in billions)	4.9	5.3	5.7	6.0	1.7

*Fisheries data from FAO (*http://www.fao.org/docrep/w9900e/9900e00.htm*). Data for 1997 are preliminary and subject to change.

fisheries in sensitive marine ecosystems, are questions that must be answered. To most scientists, it seems unlikely that the sustainable level of marine capture fisheries will be much higher than 100 MMT. *See* ANTARCTIC OCEAN; FOOD WEB.

Biological productivity of ocean regions varies greatly, controlling the potential to produce fisheries products. Coastal regions and estuaries are more productive than the open ocean, which has been described in a comparative sense as a desert. Coastal areas of the sea where nutrient-rich water upwells, especially on the west coasts of continents, support many of the world's greatest fisheries for sardinelike and anchovy species, even though these areas constitute less than 2% of the sea. Despite the infertility and low productivity of the open ocean, some of the world's most valuable fisheries, for example tuna, are found there, often in weak but geographically extensive upwelling regions of the subtropical gyres in the world oceans.

Most of the sea is being fished to some extent today. The deep sea has been relatively untouched because of technological challenges and dilute fish resources. But new fisheries at greater depths are developing that may change this situation. The vast expanses of the southern ocean support major krill (*Euphausia superba*) resources, which have been exploited only lightly to date because they are regulated by international treaties. If these krill were exploited heavily, the annual yield could be millions of tons. However, the role of krill as prey for fish, birds, and mammals in the Antarctic marine ecosystem is critical, and it seems unlikely that krill will support large fisheries in the near future.

Marine aquaculture has become a major factor in marine fisheries production since the 1980s. The marine aquaculture sector has had the greatest relative increase in fisheries production since 1985 (Fig. 1; table) and will continue to increase during the twenty-first century. Ironically, the rapid rise in marine aquaculture production depends in part on landings of marine-capture fish, which are processed into fish meal, a major constituent of aquaculture feeds. Aquaculture is the sector of fisheries where large increases in production remain possible over the next 25 years. *See* AQUACULTURE.

The character of fishing has changed significantly since 1950. Traditional artisanal fisheries still exist in many parts of the world, although developing market economies have transformed these fisheries from subsistence fisheries to market industries. Since 1985, catches by developing nations have exceeded those of developed nations. Industrial fishing, dependent on fleets of large vessels and supported by sophisticated technologies, expanded primarily in developed nations, mostly in the Western Hemisphere, Japan, and the former Soviet Union. It was this sector of marine fisheries, especially its distant-water fleets, that was responsible for the rapid global expansion of fisheries.

Marine recreational fishing has increased in importance. Statistics for recreational fishing are seldom accurate, but these fisheries account for a major fraction of landings for some species. While most recreational fishing is practiced in developed nations, its importance to tourist economies in developing nations has contributed to its increasing popularity. Recreational fishing is a major industry in the United States, where annual landings of recreationally caught marine fish have averaged more than 90,000 tons in recent years, and marine recreational anglers spent more than $7 billion dollars in 1997. Recreational fishers outnumber commercial fishers and have a powerful influence on fisheries management in the United States.

Major components in the catch. The global catch (capture plus aquaculture) consists of hundreds of species partitioned into categories by the Food and Agriculture Organization (FAO). However, on a global basis 40% of the world landings and production of fish products (combined fresh-water, marine, and aquaculture) come from only 20 species. Thus, very few species dominate the marine capture fisheries. In 1997 five species contributed 22.1% to the marine landings: Peru anchoveta 7.7 MMT, Alaska pollock (*Theragra chalcogramma*) 4.4 MMT, Chilean jack mackerel (*Trachurus symmetricus*) 3.6 MMT, Atlantic herring (*Clupea harengus*) 2.5 MMT, and chub mackerel (*Scomber japonicus*) 2.4 MMT. All of these species are small, shoaling fishes that feed on plankton.

Herrings, sardines, and anchovies make up the largest FAO category in global landing and production statistics, with more than 20 MMT landed

annually in recent years. Other major categories include cods, hakes, and pollocks (10 MMT), flatfishes and flounders (1 MMT), salmons (>2 MMT), and tunas, billfishes, and swordfish (nearly 5 MMT). Among the invertebrates, shrimps (>3 MMT), oysters (>3 MMT), squids (>3 MMT), and clams (nearly 3 MMT) are major contributors to these relatively high-value fisheries.

The largest fraction of landings (70–75%) is consumed directly by humans. However, an important component, ranging 25–30% in recent years, is reduced to fishmeal and oil, which is used primarily as feed for domestic livestock. Small, oily fishes with low market value, such as sardines, anchovies, and capelin (*Mallotus villosus*), contribute most to these industrial meal and oil fisheries. A significant fraction of these landings and products are processed as feed for fish and shrimp in the aquaculture industry. Thus, inexpensive fish are converted into products that are used to produce cultured shrimp and salmon which have high market value. Some scientists question the wisdom of this approach, which requires heavy exploitation of small fishes to support a relatively small production of valued aquaculture products.

Global status. Many marine fisheries are overexploited and less productive than they could be if managed better. On a global basis, FAO indicates that approximately 25% of the world's fisheries were overexploited, depleted, or recovering during the mid-1990s. Another 44% were heavily or fully exploited, while approximately 31% were believed to be capable of sustaining heavier fishing. In the United States, the situation was similar in 1999, when 27% of its marine fisheries were known to be overexploited, 37% were heavily or fully exploited, and 15% were capable of sustaining heavier exploitation. The status of 21% of the United States stocks was unknown in 1999.

Global demand. It is estimated that annual demand for fish to be consumed by humans will reach 95 MMT by the year 2010 if per capita consumption (15.7 kg in 1996) is maintained. Approximately 58 MMT of marine capture fish and 7 MMT of marine aquaculture fish are consumed annually in addition to approximately 25 MMT of fresh-water fish (aquaculture plus capture). Increased demand will require an additional 5 MMT to maintain the per capita consumption level. There is concern that wild fish stocks cannot meet that demand, given present fishing strategies and management problems. Marine aquaculture production will continue to increase, although its ultimate potential is uncertain and dependent to an extent upon landings of wild fish. Many authorities believe that an additional 5–10 MMT of marine capture fish could be landed sustainably if fishery management improved and overfished stocks were restored.

Major Fishing Nations

Major marine fishing nations fall into two categories: recent entrants and historical fishing nations. Each of the top five marine fishing nations landed more than 5 MMT in marine capture fisheries in 1997. In that year, the top five nations were China (13.8 MMT), Peru (7.8 MMT), Japan (5.8 MMT), Chile (5.8 MMT), and the United States (5.0 MMT).

Beginning in the 1980s, China has emerged as the number one fishing nation. China also has a major fresh-water and marine aquaculture industry. China's total catch and production of fishery products was 35 MMT in 1997 (28.6% of world total). Although China has a long history of fishing and fresh-water aquaculture, its marine catch increased rapidly in recent years and continues to increase. In contrast, Peru was not a major fishing nation until the 1960s, when its anchoveta stocks came under exploitation. Peruvian fisheries concentrate on small pelagic fishes, including anchoveta, sardines, and jack mackerel, much of which is reduced to meal and oil. Landings and fish production fluctuate significantly under the positive influence of upwelled nutrients and are negatively affected by periodic El Niño oceanographic conditions. As a consequence, annual landings fluctuate markedly in Peru. Chile also is a recent entrant into the major fishing nation category, with landings resembling those of Peru, but with added diversity.

Japan has a long history of fishing and was the number one fishing nation until the 1980s. Traditional fisheries in Japan are characterized by their diversity. Declaration of 200-nautical-mile Exclusive Economic Zones by most maritime nations in the 1970s had a negative impact on Japan's distant-water fisheries. The United States, with extensive coastlines on both the Atlantic and Pacific coasts, also has a long history as a major fishing nation. Although levels of marine landings changed little in the 1990s (4–5 MMT), the United States has maintained its position among the top five or six fishing nations, which it has held for many decades.

Many other nations have economies that depend significantly upon marine fisheries and marine aquaculture or have high per capita consumption of marine fishery products. These nations and their 1997 landings include the Russian Federation (4.4 MMT), Iceland (2.2 MMT), Norway (3.2 MMT), Thailand (3.5 MMT), Indonesia (3.3 MMT) and the Republic of Korea (2.6 MMT). In the past 25 years, there has been a significant shift in the dominance nations producing the most fish products. Historically, nations with developed economies landed the highest percentage of fish, but a shift occurred in the mid-1980s when nations with developing economies became the largest producers annually catching or producing approximately 70% of the global fish.

Managing the Fisheries

Managing marine fisheries is complex and often uncertain. Managers and the agencies or institutions that support them have jurisdiction over living resources that fluctuate naturally in abundance, often are highly mobile, and are only rarely privately owned. Many species (for example, tunas and billfishes) have ranges that extend over thousands of kilometers during their annual migratory cycle. The common property nature of marine fisheries

complicates conventional management approaches that allow open access and require agreement among jurisdictions, and consensus on the best methods to regulate fishing and the allocation of the catch.

Total Allowable Catch versus fishing effort control. The most common management approach is to assign an annual Total Allowable Catch (TAC). This requires setting a quota that, if exceeded, will terminate the fishery for that year. The TAC is set at a level to prevent a catch so large that the stock will be overfished. In an open-access fishery, there may be competition among fishers to secure a large fraction of the TAC, which can lead to "derby" fishing that is difficult to control. In international fisheries, the TAC is commonly apportioned among participating nations based on negotiated allocations and historical landings. Many variations of quota assignments, including dividing the TAC into seasonal allocations or apportioning it among areas, have been used. For example, in Pacific salmon management, individual TACs may be assigned to stocks that are fished near the mouths of spawning rivers. In this way, sufficient adults escape to insure the continued reproductive success of the individual stocks.

A second common management approach is to control the amount of fishing effort. Controlling effort nominally limits the mortality rate attributable to fishing, which is presumed to be directly proportional to fishing effort. Effort regulation, when effective, fixes the fraction of a stock that is caught. Consequently, the number of fish landed may vary among years, with high catches during years of high abundance and lower catches when the stock is less abundant. Many fishery scientists and managers argue that effort management is preferred over TACs because TACs may allow landings that are too high in years when the stock is at low levels. Fishing effort can be controlled by limiting numbers of participants, by limiting amounts of fishing time or amounts of gear, or by legislating inefficiencies into the fisheries. Prohibiting fishing gears or methods that may, because of their efficiency, catch fish too rapidly or effectively may control fishers and allow an even distribution of landings throughout the fishing year.

Frequently, TAC (output) and fishing effort (input) controls are applied together, especially in heavily fished stocks, for example, northwestern Atlantic stocks, such as cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*). Other, less direct methods may be applied as well. For example, size limits may be imposed to protect young fish, allowing them to mature and spawn. In some species (for example, some sturgeons), very large individuals may be protected from fishing to maintain the fecundity and spawning potential of a stock. Other controls which are becoming common include time and area restrictions, such as closure of nursery areas, protection of spawning habitats, and exclusion of fishers from polluted areas. In some cases, managers attempt to maximize the quality of the catch by specifying a fishing season when the resource is at its best condition (for example, oysters). Daily landing quotas, called bag limits, may be assigned to spread the TAC over a season or to ensure that the catch is spread among numerous fishers, rather than being taken by the most powerful fishing vessels. In some managed stocks, only males are allowed to be landed—females carrying eggs must be released to protect spawning potential of the stock. Such measures are particularly popular in some crustacean (crab and lobster) fisheries.

Maximizing sustainable yields. An objective of fisheries management is to achieve yields from a fishery that are both high and sustainable. In marine commercial fisheries, that objective historically was translated into a fishing strategy that allowed a maximum sustainable yield (MSY) from a stock that is in balance with the stock's reproductive and growth capacities under a given set of environmental conditions. If the population growth rate of a species is known and is presumed to be primarily a function of the population's size (expressed as its aggregate weight, or biomass), then the MSY can be predicted in relation to fishing effort (or fishing mortality rate) [**Fig. 2**]. *See* POPULATION ECOLOGY.

For many years, MSY was set as a management target, which in retrospect often allowed catches that were too high, precipitating declines in the stock, leading to an unsustainable, overfished condition.

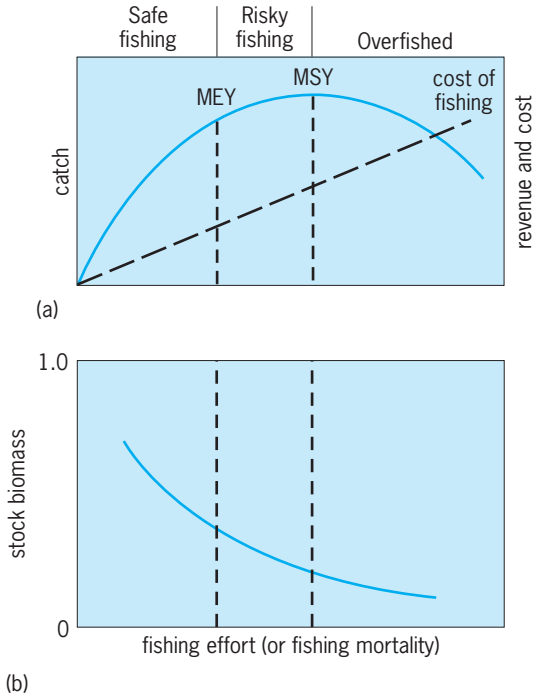

Fig. 2. Concept of maximum sustainable yield (MSY). (*a*) Catch and revenue and cost versus fishing effort. (*b*) Stock biomass versus fishing effort. In a developing fishery, catch increases in relation to fishing effort. At an intermediate level of effort and stock biomass, the MSY can, in theory, be caught. At higher levels of effort, the sustainable level of catch declines and the stock is overfished. Fishing at effort levels near MSY is risky. Safe fishing is practiced at considerably lower levels of effort. The maximum economic yield (MEY) is taken at a level of effort where net profit is maximum. The MEY is taken at effort levels lower than the effort level associated with MSY. In unregulated, open-access fisheries, the level of effort may increase to the point where the "cost of fishing" line intersects the catch (=revenue) line.

Environmental variability and failed reproduction in years of unfavorable conditions make MSY an undesirable and risky strategy. However, estimating MSY remains an important objective, and recent management advice usually sets MSY as a threshold, a level of catch that, if approached, should signal caution rather than being considered a target for the fishery.

Maximizing yield per recruit. A second common management goal has been to maximize or optimize the catch with respect to a given level of recruitment (production of young fish). In this approach, the size (or age) at which fish first enter the fishery is adjusted by regulating mesh sizes of fishing gears, and the fishing mortality rate is controlled by adjusting the amount of fishing effort. Given a fundamental knowledge of fish growth and an estimate of natural mortality rate, maximizing the yield per recruit can be easily achieved in theory, although it may be achieved at the risk of catching too many adults, thus reducing the biomass of the spawning stock to a level so low that future recruitment success is jeopardized (**Fig. 3**). Thus, fishing at a level that takes maximum yield per recruit is risky, and is now considered to be a threshold associated with too high a level of fishing mortality.

The yield-per-recruit model also can be used to estimate spawning biomass per recruit relative to its level if there were no fishing mortality and under specified levels of fishing mortality (Fig. 3). Spawning biomass of a stock falls rapidly as fishing mortality rate increases, indicating that fishing mortalities must be set at relatively low levels to conserve high spawning potential and ensure future recruitment success.

Biological reference points. Managing marine fisheries generally requires the designation of reference points, which serve as targets and thresholds. Biological reference points most often are specified fishing mortality rates, spawning stock biomasses associated with fishing mortality rates, or catch levels relative to specified fishing mortality rates or stock biomass levels. Precautionary management rests on selecting target reference points that allow sustainable fishing at profitable levels, yet do not endanger the stock's capacity to reproduce. In many fisheries, experience has demonstrated that fishing mortality rates must be limited to levels that maintain spawning biomass at levels at least 20–40% of that in an unfished stock. Thus, safe fishing may require limiting yields to levels that are less than the maximum sustainable yield and less than the maximum yield per recruit (Figs. 2 and 3).

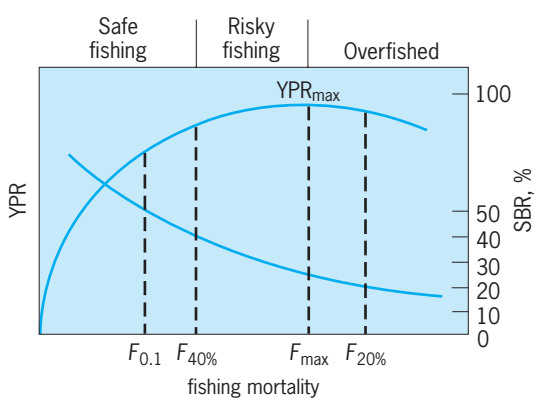

Fig. 3. Concept of maximizing yield per recruit (YPR_{max}). Yield per recruit increases as fishing mortality (or effort) increases until a maximum is reached, beyond which increases in effort result in declining YPR as the stock is overfished. Biological reference points ($F_{\%}$) are fishing mortality rates associated with specified yields per recruit. A stock is overfished at F levels higher than F_{max}. Spawning biomass per recruit (SBR) declines as F increases. Safe fishing may require a relatively low F that ensures higher SBR ($\geq$40% in the example). The $F_{0.1}$ reference point, historically considered a safe target level, sets F at a level where the slope of the YPR curve is 1/10 of that at the origin.

Economics. To be profitable, revenues from catches obviously must exceed costs, although governments often have subsidized fishing, an industry that provides human food. In a simple model, the maximum economic yield, or level of catch that provides the largest net profit, is achieved at an effort level where revenues most exceed costs of fishing. This effort level is considerably lower than the effort required to obtain maximum sustainable yield (Fig. 2). In this sense, it is a safer level at which to fish, in addition to maximizing profits. In open-access fisheries, without strong management control, economists argue that effort will increase in a fishery until it equilibrates at a high level where net profits are zero (Fig. 2)—a clearly undesirable circumstance that also risks the biological collapse of the stock.

Recreational fisheries. Marine recreational fishing is a valuable industry in many developed countries. Allocations of catch to recreational fishers are controversial in fisheries with competing users and different values, althouth principles of conservation and management for recreational and commercial sectors often are similar. In some cases, reference points in fisheries dominated by the recreational sector may be set conservatively to maintain quality (many and big fish) in the fishery. As marine recreational fishing increases in importance globally and contributes significantly to the economies of developing nations, conflicting management philosophy and practice between the recreational and commercial sectors seem certain to increase.

Allocating rights and privileges. Many scientists and managers believe that to secure and stabilize marine fisheries their historical common property nature must be transformed to selectively grant user rights or privileges to participants in the fishery. Many examples of rights-based management are in effect throughout the world, such as licensing and permitting systems, and limiting entry to already overcapitalized fisheries to reduce or place a cap on the numbers of participants. An increasingly common but controversial management approach to limit access is the Individual Fishing Quota, which allocates shares to selected fishers based on their historical presence in a fishery. These shares of the Total Allowable Catch are often treated as a commodity that can be traded or sold. Management by Individual Fishing Quota, works on the assumption that participants,

because they hold access rights and in a sense share ownership of the resource, will not engage in behavior that promotes overfishing or damages the marine environment.

Institutions and agencies. A complex of local, regional, national, and international institutions is charged with management of marine fisheries. Many fish stocks are migratory. Fish stocks do not recognize jurisdictional boundaries and cannot be assessed or managed effectively if some jurisdictions are not included in management plans. The classic case exemplifying the need for interjurisdictional management is that of tunas, swordfish, and billfishes, which have been called "fish without countries" because of their migratory behaviors.

International commissions provide management advice and in some cases hold management authority for highly migratory species, such as tunas, or for less migratory fish stocks in international waters, such as cod, haddock, and herring in the North Sea, Barents Sea, or Baltic Sea ecosystems. Examples of international commissions are the International Commission for the Conservation of Atlantic Tunas (ICCAT) and the Northwest Atlantic Fisheries Organization (NAFO). International commissions vary with respect to their research responsibilities and with respect to their advisory or regulatory roles relative to authority exercised by the signatory nations. *See* ARCTIC OCEAN; BALTIC SEA; NORTH SEA.

Marine fisheries of most nations are usually managed by a national authority. In the United States, that authority rests in the Department of Commerce and its National Marine Fisheries Service. Nations with large, geographically dispersed fisheries may have regional management institutions responsibe for particular fish stocks. In the United States, regional management is vested in eight Regional Fisheries Management Councils, which develop management plans to govern particular fisheries. In addition, other regional institutions may manage marine fisheries in particular zones. For example, in the United States, the Atlantic States Marine Fisheries Commission is responsible for managing migratory stocks, such as striped bass (*Morone saxatilis*) and American shad (*Alosa sapidissima*), that primarily reside within 3 nautical miles of shore, but which may be fished along the coastlines of many states. There are also agencies within individual regions or states (or in local waters) that have authority to regulate fishing. These jurisdictional units sometimes have conflicting goals and approaches to fish stock management. In many cases, conflicts are resolved or avoided by agreements that ensure joint management responsibility among agencies. For example, in the United States, Atlantic coastal stocks of bluefish (*Pomatomus saltatrix*) are jointly managed by the Atlantic States Marine Fisheries Commission, the states, and three Regional Fisheries Management Councils.

Ecosystem Approaches to Management

During the 1990s there was increasing public concern about the status of marine fisheries and exploitation practices. As more fisheries were overexploited, more habitats came under stress, and water quality issues came to the forefront. Most nations and management agencies adopted, at least nominally, precautionary approaches to the exploitation and management of marine fisheries. There emerged a sense that the burden of proof with respect to effects of heavier fishing pressure or implementation of new marine fisheries ought to be borne by fishers and not by management agencies. This shift in philosophy gained a foothold in the 1990s and has slowly become a guiding framework for marine fisheries management. The approach recognizes that fishing is a significant presence in marine ecosystems and that its ecosystem-level effects must be considered.

Multispecies management. New models are emerging to account for multispecies interactions in marine fisheries that previously were neglected in traditional, single-species management. Predatorprey relationships dominate interactions among marine species and are emphasized in these models. Most fish die from being consumed by larger fish. Fishing that significantly alters abundances of predator and prey species can affect productivities of the stocks and their potentials to provide sustainable catches. In cases where several species of fish interact through predator-prey and competitive relationships, the effects of fishing on those interactions and on potential yields of individual species are difficult to predict and not always intuitive. Consequently, managers have been reluctant to adopt complex multispecies models as the basis for management. However, it is probable that such models will become a tool that managers regularly use to forecast potential effects of different fishing strategies on fisheries productivity and marine ecosystem structure.

Ecosystem principles and fisheries management. Beyond multispecies management, the broader consequences of ecosystem effects of fishing need to be addressed. Consideration of fishing practices and their effects must be included in fisheries management plans. Water quality and critical spawning and nursery areas, or areas with critical habitats for prey or predators of a target species, must be protected. Catches of nontargeted organisms, called bycatch, whether they are unwanted species or are small individuals of a targeted species, must be reduced. The bycatch issue extends to the capture of birds, turtles, and marine mammals, adding impetus to solve the problem. The magnitude of the problem is highlighted by records that indicate more than 25 MMT annually of marine bycatch has been caught and discarded in recent years. *See* WATER CONSERVATION.

Accounting for effects of fishing on ecosystems will depend on development of ecosystem models that not only examine a multispecies complex exploited by a fishery but include other components and aspects of ecosystems. In future management scenarios, ecosystem-level effects of fishing practices and alternatives will be explored by running ecosystem models in which fishing is simulated and ecosystem responses are evaluated. Such

integrated management of marine fisheries will depend on developing new ecosystem models and on applying new tools. The concepts of marine protected areas and no-take reserves have become popular, and under some circumstances they promise to conserve, restore, and improve fisheries. Moving toward ecosystem-level management of marine fisheries requires a rigorous hypothesis-based, experimental approach. In this sense, adaptive management, which to date has mostly meant remaining flexible in management responses, will take on a more important role as ecosystem-level models are developed to guide marine fisheries management. *See* MARINE ECOLOGY; MATHEMATICAL ECOLOGY.

Edward D. Houde

Fishing Technology

Fishing technology is the rational application of practical experience together with engineering and related natural sciences (such as physics, chemistry, mathematics, biology, and hydrology) to fishing vessels, gears, and operations.

Until the 1930s the technical development of fishing was largely left to the fishers, the ship builders, and the supplying industry. In industrialized countries the overall technical development has had a significant impact on fishing in terms of increased size, power mechanization, range, complexity, and costs. Consequently the traditional trial-and-error approach had to be replaced by systematic research and development. Industrialized countries with important fisheries have, therefore, established specific government- and industry-sponsored research and development capacities, and fishing technology has thus become a fisheries discipline in its own right.

The introduction of engine power to fishing vessels at the end of the last century considerably increased the range and flexibility of fishing operations. It also enabled the development of auxiliary machinery such as winches and haulers to improve efficiency and safety and to reduce hard labor.

Fig. 4. Example of a long-distance factory trawler (3177 ft or 95 m long, 5000 hp or 3.7 MW) capable of bottom and midwater trawling anywhere in the world. (*Hansetische Hochseefischerei, Germany*)

Trawling. This mechanization—from which all fishing techniques benefited to various degrees—started a spectacular development of trawling which became one of the most productive fishing methods. In trawling, a conical net bag is towed over the ground (bottom trawling) or through the free water at any depth (midwater trawling) and scoops up the fish in its path. The net is usually towed by a vessel called a trawler; sometimes two vessels are used. Trawlers range from outboard-powered canoes, operating on lakes or in coastal waters, to long-distance factory trawlers that are more than 330 ft (100 m) long and have main engines with more than 5000 hp (3.7 MW) (**Fig. 4**). These are among the most expensive and sophisticated commercial vessels of their size. The most common trawl gear is the otter trawl, for which special shearing devices (otter boards) are used to spread the net opening horizontally. Otter boards range from flat rectangular plates to hydrofoil curved designs up to 129 ft^3 (12 m^2) each. Bottom trawl nets have a mouth opening of up to about 2700 ft^2 (250 m^2) and a total net length of about 260 ft (80 m), and midwater trawl nets can be even larger, up to 22,000 ft^2 (2000 m^2) and 590 ft (180 m), respectively. Commercial bottom trawling goes down to more than 3300 ft (1000 m) depth, and midwater trawling to about 2600 ft (800 m). Towing speed is between about 2.5 and 5 knots (4.5 and 9 km/h). As exploitation of ocean resources has grown more intense, trawling in the deep sea (depths >400 m) has been increasingly common.

Aimed trawling. Technical progress in trawlers, trawl gears, and underwater acoustics led during the 1960s to the development of aimed trawling (**Fig. 5**). The gear requires a horizontal echo sounder or sonar to detect and select suitable fish schools ahead of the trawler, the ship's echo sounder to further assess the depth and quantity of the school and the bottom conditions, and a net sounder to monitor the net depth, the opening height, and the fish in, below, and above the net opening. This combination of echo-sounding (acoustic) devices with modern navigational aids such as gyro compass, sensitive speed log, radar, decca navigator, loran, or even satellite navigation significantly improved the chances of hitting a fish school with the trawl gear and of counteracting escape reactions. Aimed trawling is mainly applied in midwater trawling, but the technology is increasingly also used in bottom trawling. *See* UNDERWATER SOUND.

Integrated trawling systems supplement aimed trawling by adding electronic data processing and computer control of certain parts of the aiming process. A virtually automatic interception of trawl and fish and optimum operational efficiency are achieved by computer processing. This processing includes adjustment of course and speed of the trawler and the trawl gear to the position and the movements of selected fish schools, the fixing of sonar to the target, and the adjustment of the trawl winch operation

and the towing speed to correct the net depth. Such complex systems are mainly meant for midwater trawling. Other applications are being developed. *See* SONAR.

Echo sounding. One of the major problems of fishing is to find the fish when they are not visible at the surface. This was largely solved when, starting in the late 1930s, echo sounding was developed not only to measure the water depth but also to show fish. The principle of echo sounding consists in measuring the time that a sound impulse takes to travel to a target and the reflected echo to come back. Half the total time multiplied by the speed of sound (in water, about 0.9 mi/s or 1500 m/s) gives the distance. The echo strength depends, among other things, on the reflection properties of the target; for example, rocks reflect better than sand or mud and large fish (particularly with swim bladder), or fish schools better than small or single fish. Fish-finding echo sounders need more transmission power and echo amplification than purely navigational echo sounders and also a suitable monitor display, Historically, echograms were presented on recording or cathode-ray-tube (CRT) displays for echo discrimination of bottom and various fish sizes and concentrations.

The normal echo sounder, covering the water column below the vessel, is applicable to all fishing methods. Horizontal, or, rather, oblique, echo sounding at various angles between the surface and the bottom forward and around the vessel is also known as sonar (sound navigation and ranging). It mainly serves in midwater trawling and purse seining (see below), for which the advance detection of fish is quite essential. Depending on type and power, the range for fish detection of echo sounders is down to about a 3300-ft (1000-m) depth, and of sonar up to about 9900 ft (3000 m) away from the vessel. The net sounder developed for midwater trawling is an echo sounder with the transducers or sounding units at the net rather than in the ship's bottom (Fig. 3). The connection from net to trawler may be by cable or by an acoustic link (wireless). The more complex multi-net sounder equipment for commercial fisheries includes forward transducers in addition to the down and up transducers, a net filling meter, and a thermometer at the net to check whether the water temperature is according to the preference of the species sought. Scientific multi-net sounder equipment has additional transducers across the trawl opening and further aft in the bag, a speed-through-water meter, and a transponder to study gear performance and fish reactions in detail.

Purse seining. In addition to giving all fishers an eye to see under water and paving the way for aimed trawling, the development of underwater acoustics also opened a new dimension to purse seining. This fishing method uses a net which may be 490 to 4900 ft (150 to 1500 m) long and 100 to 660 ft (30 to 200 m) deep which is set in a circle around a fish school and then closed at the bottom like a purse, catching all fish inside. Unless it is done in connection with fish attraction (such as with light), purse seining is completely dependent on finding suitable

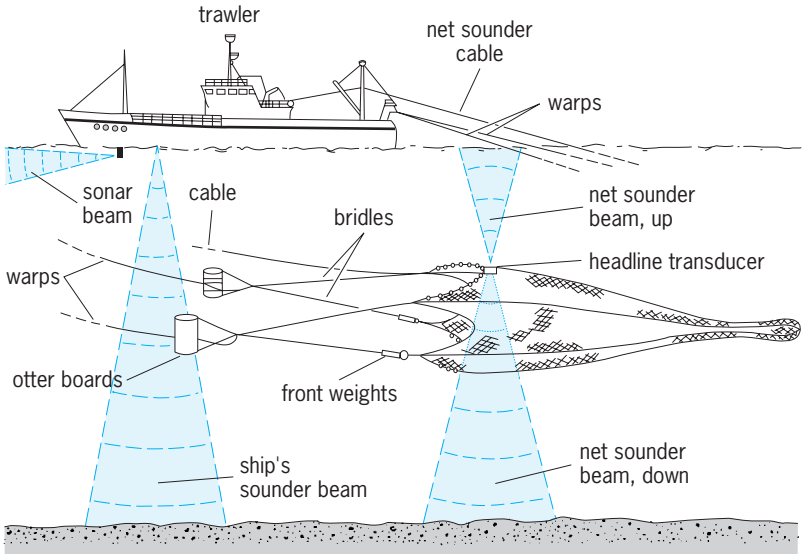

Fig. 5. Typical arrangements of fishing gear and acoustic instrumentation for aimed midwater trawling.

fish schools. Schools of fish may be sited from the vessel or from airplane spotters who locate the schools and direct the vessel to them. For fish which are not visible at the surface, sonar-guided purse seining is the only solution and quickly led to a very significant expansion of this fishery, which is the largest bulk fish producer in the world. As in aimed trawling, the success of sonar-guided purse seining is largely dependent on the skill of the individual skipper. Consequently, computerized systems have been developed to assist the skipper and also improve the complex tasks of detecting and approaching a fish school, assessing its size and depth, and coordinating the relative movements of vessel and fish school during actual fishing. One such system employs a television screen (**Fig. 6**) to show the situation in the horizontal in relative or true motion as desired. The computer can also link the sonar to a selected fish school and calculate the depth of the school.

Longlining. Longlines range (from a few hundred meters to many kilometers in length and carry baited hooks set at intervals. This gear has been adapted since the 1950s to effectively catch migratory species such as tunas and swordfish in the open sea. Development of accurate navigation and the Global Positioning System, and acquisition of sensors to detect temperature and other oceanographic features, have increased the ability of longliners to locate and remain in the environment of the target species. In addition, the use of chemiluminescent sticks light as a fish attractant has increased the effectiveness of the longline technique. *See* REMOTE SENSING; SATELLITE NAVIGATION SYSTEMS.

Vessel trends. In addition to increases in size and horsepower, vessels in modern industrial fishing fleets have become more diversified to allow participation in fisheries that use more than one type of gear and target different kinds of fish. For example, in the Gulf of Alaska and Bering Sea fisheries, vessels may be rigged to alternatively trawl for groundfish, set

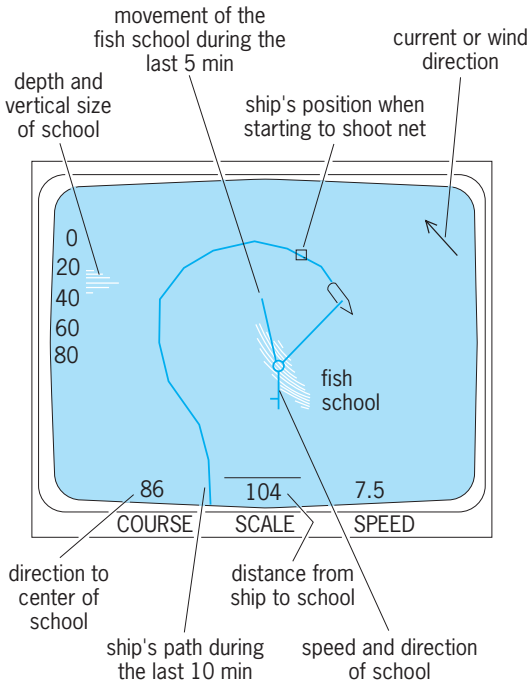

Fig. 6. Automatic tracking and true-motion situation display by a computer-integrated purse-seining system. Course is indicated in degrees, scale in meters, and speed of vessel in knots. Depth scale at left is in meters. 1 m = 3.3 ft. (*SIMRAD AS, Norway*)

traps for crab, or longline for halibut, depending on the season and potential profitablility of a particular fishery. Such versatility often is required to remain profitable in modern, highly competitive fisheries. *See* BERING SEA; OCEANOGRAPHIC VESSELS.

Synthetic fibers. Apart from these special developments in trawling and purse seining, the introduction of synthetic materials for yarns, netting, and ropes had an even more general impact on fishing gears and fishing than the introduction of echo sounding, which started at about the same time. The main advantages of synthetics such as polyamide (for example, nylon), polyethylene, and polypropylene are their strong resistance against rotting. Some also have higher breaking strength, elasticity, and abrasion resistance than natural fibers such as cotton, manila, and sisal used hitherto. Higher breaking strength enables the use of thinner yarns, and synthetic monofilaments are almost invisible in water. This can significantly improve the catching efficiency of, for instance, trawls and other bag nets, gill nets, or angling gear. These advantages compensate for the higher price of synthetics, which have now almost completely replaced the traditional natural-fiber materials. *See* MANUFACTURED FIBER.

Joachim Schärfe

Bibliography. *Ecosystem-Based Fishery Management*, National Oceanographic and Atmospheric Administration, National Marine Fisheries Service, Silver Spring, MD, 1999; *Fisheries of the United States*, National Marine Fisheries Service, Silver Spring, MD, annually; National Research Council, *Improving Fish Stock Assessments*, National Academy Press, Washington, DC, 1998; National Research Council, *Sustaining Marine Fisheries*, National Academy Press, Washington, DC, 1999; *Our Living Oceans 1999*, National Oceanographic and Atmospheric Administration, National Marine Fisheries Service, Silver Spring, MD, 1999; *Yearbook of Fishery Statistics*, Food and Agriculture Organization, Rome, annually.

Marine geology

The study of the portion of the Earth beneath the oceans. Approximately 70% of the Earth's surface is covered with water. Marine geology involves the study of the sea floor; of the sediments, rocks, and structures beneath the sea floor; and of the processes that are responsible for their formation. For the marine geologist, the presence of the oceans masks the principal features of interest. The average depth of the ocean is about 3800 m (12,500 ft), and the greatest depths are in excess of 11,000 m (36,000 ft; the Marianas Trench). The study of the sea floor necessitates employing a complex suite of techniques to measure the characteristic properties of the Earth's surface beneath the oceans. Contrary to popular views, only a minority of marine geological investigations involve the direct observation of the sea floor by scuba diving or in submersibles. Instead, most of the ocean floor has been investigated by surface ships using remote-sensing geophysical techniques, and more recently by satellite observations.

In the mid-1960s to the mid-1970s, the theory of plate tectonics was developed and verified based on geological and geophysical observations of the ocean floor and crust. Key observations included the magnetic characteristics of rocks, the disposition of submarine landforms, the distribution and ages of marine sediments, and the locus and nature of zones of earthquake and volcanic activity. Earthquakes served to define the principal plate boundaries. The axes of mid-oceanic ridges mark sites of plate divergence; trenches mark sites of plate collisions (subduction zones). Diverging plates allow for the birth of new oceanic crust, whereas colliding plates accommodate the destruction of equivalent amounts of crust. Large strike-slip faults exist (such as the San Andreas Fault) where plates slide by one another (**Fig. 1**). *See* EARTHQUAKE; PLATE TECTONICS.

The ocean crust is relatively young, having been formed entirely within the last 200 million years (m.y). The process of renewing or recycling the oceanic crust is the direct consequence of plate tectonics and sea-floor-spreading processes. It is therefore logical, and perhaps essential, that the geologic history of the sea floor be outlined within the framework of plate tectonic tenets. Where plates move apart, molten lava reaches the surface to fill the voids, creating new oceanic crust. Where the plates come together, oceanic crust is thrust back within the interior of the Earth, creating the deep oceanic trenches. These trenches are located primarily around the rim of the Pacific Ocean (Fig. 1). The down-going material can be traced by using the distribution of earthquakes to depths of about 700 km (420 mi). At that level, the character of the subducted lithosphere

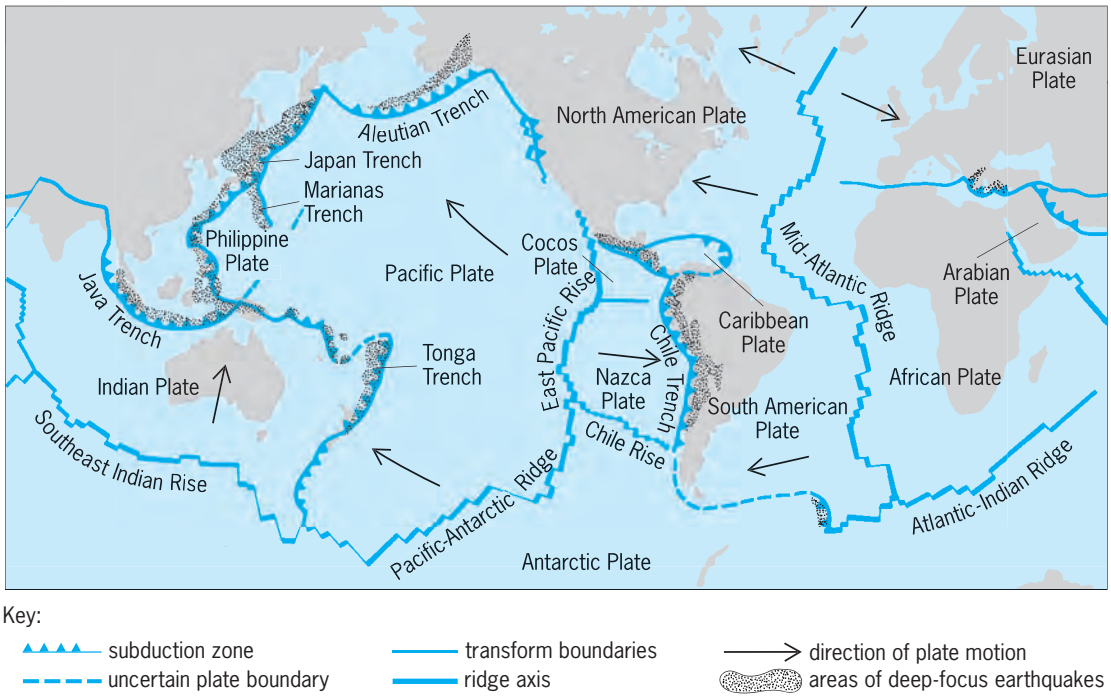

Fig. 1. Map showing the tectonic structure of the Earth's lithosphere, which is composed of about 10 rigid plates, each moving as a distinct unit. (*After J. F. Dewey, Plate tectonics, Sci. Amer., 226(5):56–68, 1972*)

is lost, and this material is presumably remelted and assimilated with the surrounding upper-mantle material. *See* GEODYNAMICS; LITHOSPHERE.

Major Morphologic and Sediment Provinces

The major features of the sea floor are the mid-oceanic ridges, basins, continental margins, large igneous provinces, marginal seas, and anomalous features.

Mid-oceanic ridges. Most of the ocean floor can be classified into three broad physiographic regions, one grading into the other (**Fig. 2**). The approximate centers of the ocean basin are characterized by spectacular, globally encircling mountain ranges, the mid-oceanic ridge (MOR) system, which formed as the direct consequence of the splitting apart of oceanic lithosphere. The small-scale morphologic characteristics of these mountain ranges depend somewhat upon the rate of separation of the plates involved. As is shown in **Fig. 3**, abyssal hill relief, especially within 500 km (300 mi) of ridge crest, is noticeably rougher on the slow-spreading Mid-Atlantic Ridge than on the

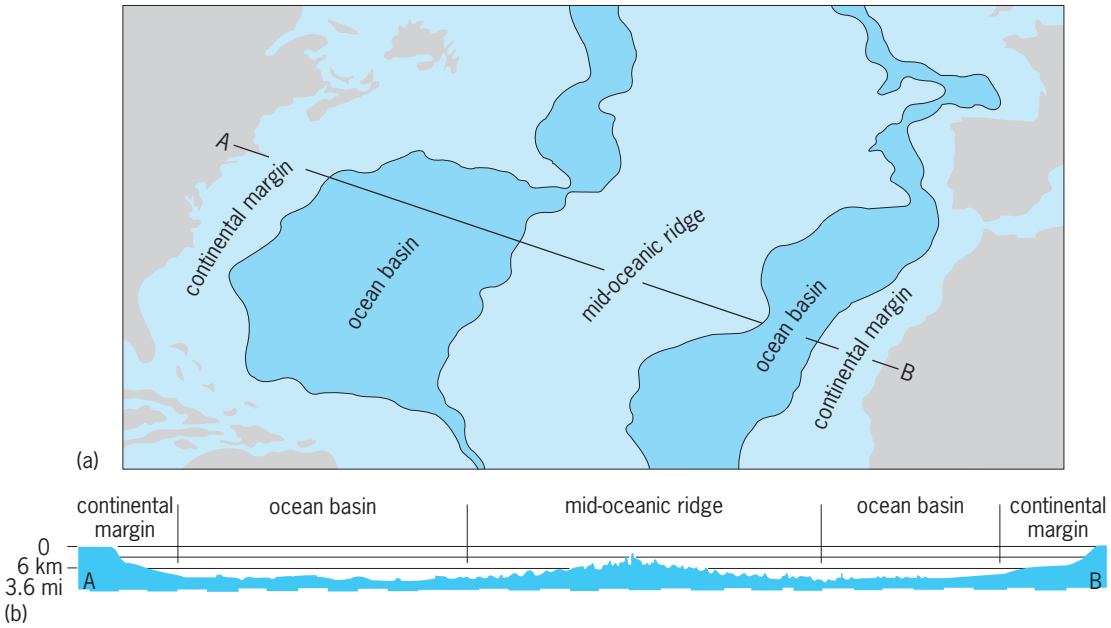

Fig. 2. Geology of the North Atlantic Ocean. (*a*) Physiographic divisions of the ocean floor. (*b*) Principal morphologic features along the profile between North America and Africa. (*Modified from B. C. Heezen et al., GSA Spec. Pap. 65, 1959*)

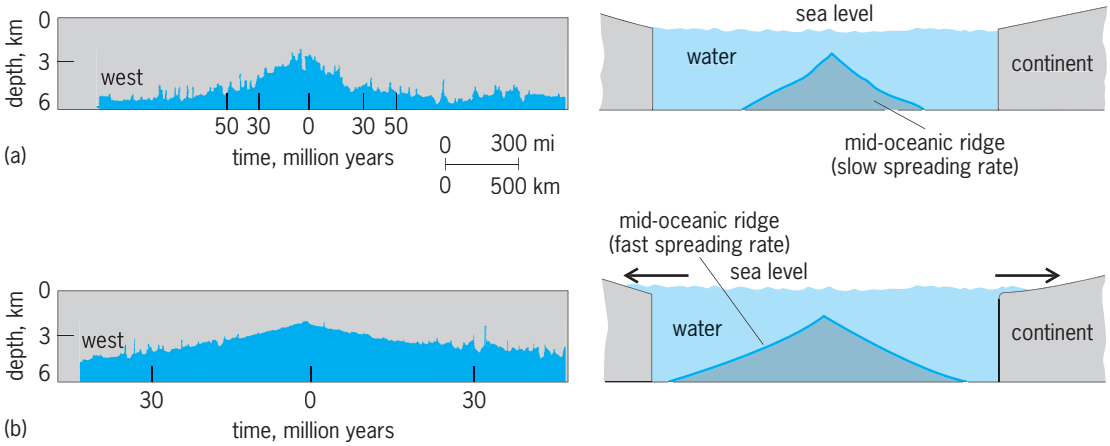

Fig. 3. Topography of the mid-oceanic ridge. (*a*) Mid-Atlantic Ridge (25°N), spreading rate 2.6 cm (1.0 in.) per year. (*b*) East Pacific Rise (55°S), spreading rate 8.8 cm (3.5 in.) per year; sea-level rises causing transgression (arrows). 1 km = 0.6 mi.

fast-spreading East Pacific Rise. The profile of the East Pacific Rise is also broader and shallower than for the Mid-Atlantic Ridge. If the entire mid-oceanic ridge system were spreading rapidly, the expanded volume of the ridge system would displace water from the ocean basins onto the continents, and this may explain a well-known incident of marine transgression during the Cretaceous Period.

The broad cross-sectional shape of this mid-ocean mountain range can be related directly and simply to its age. The depth of the mid-oceanic ridge at any place is a consequence of the steady conduction of heat to the surface and the associated cooling of the oceanic crust and lithosphere. As it cools, contracts, and becomes denser, the oceanic crust plus the oceanic lithosphere sink isostatically (under its own weight) into the more fluid asthenosphere. The depth to the top of the oceanic crust is a predictable function of the age of that crust; departures from such depth predictions represent oceanic depth anomalies (**Fig. 4**). These depth anomalies are presumably formed because of processes other than lithospheric cooling, such as intraplate volcanism. The Hawaiian island chain and the Polynesian island groups are examples of this type of volcanism. *See* ASTHENOSPHERE; ISOSTASY.

Very detailed surveys have been completed for modest sized pieces of the mid-ocean ridge system in the North Pacific, South Pacific, and South Atlantic in an effort to compare and contrast the morphologic features of the crestal zones at fast-, intermediate-, and slow-spreading ridges. The amplitude of the small-scale component of relief is somewhat dependent on the associated rate of sea-floor spreading. This small-scale relief is born at the axis of the ridge. As it moves laterally by spreading, it evolves gradually, becoming more subdued as it is slowly draped by a "snowfall" of pelagic sediments. The resulting sea floor abyssal hills constitute the largest morphologic province in the world.

The mapping studies have revealed that the mid-ocean ridges are characteristically segmented at a variety of along-axis scales ranging from about 10 km (6 mi) up to about 1000 km (600 mi). The depth of the near crestal regions and the degree of segmentation are thought to reflect the relative amount of magmatic melt available for injection and its temporal variability. Mid-ocean ridges' variable subsidence character captures the largest scale component of ridge segmentation. The smaller scales of segmentation are manifest as contrasts in the along-axis depths and in their near-crestal morphology (**Fig. 5**).

The loci of sea-floor volcanic activity create, most notably at or near the axis of the MOR system, an extreme environment of chemistry and temperature, with hydrothermal plumes (vents) having temperatures often in excess of 400°C (750°F) and containing dissolved and particulate minerals such as sulfides, lead, copper, and others but devoid of free oxygen and sunlight. These conditions define an environment previously presumed unsuitable for supporting life. Numerous diving vessel expeditions to these environments have sampled the fluids and mineral particles emanating like "black smoke" from large chimneylike structures. They have also recovered strange life-forms such as enormous clams and giant tubeworms that symbiotically exist with a type of bacteria called chemoautotrophs. Such observations have required geobiologists to reevaluate their presumptions about what other extreme environmental conditions may sustain life-forms. *See* HYDROTHERMAL VENT; MID-OCEANIC RIDGE; OCEANIC ISLANDS; VOLCANOLOGY.

Basins. The deep ocean basins, which lie adjacent to the flanks of the mid-oceanic ridge, represent the older portions of the sea floor that were once the shallower flanks of the ridge (Fig. 2). The bulk of sediments found on the ocean floor can be broadly classified as terrigenous or biogenic. Terrigenous sediments are derived from drainage of adjacent landmasses and are brought to the sea floor through river systems. This sediment load is sometimes transported across the continental shelves, often using as pathways the submarine canyons that dissect the shelves, the continental slope, and the continental rise. Biogenic sediments are found in all parts of the

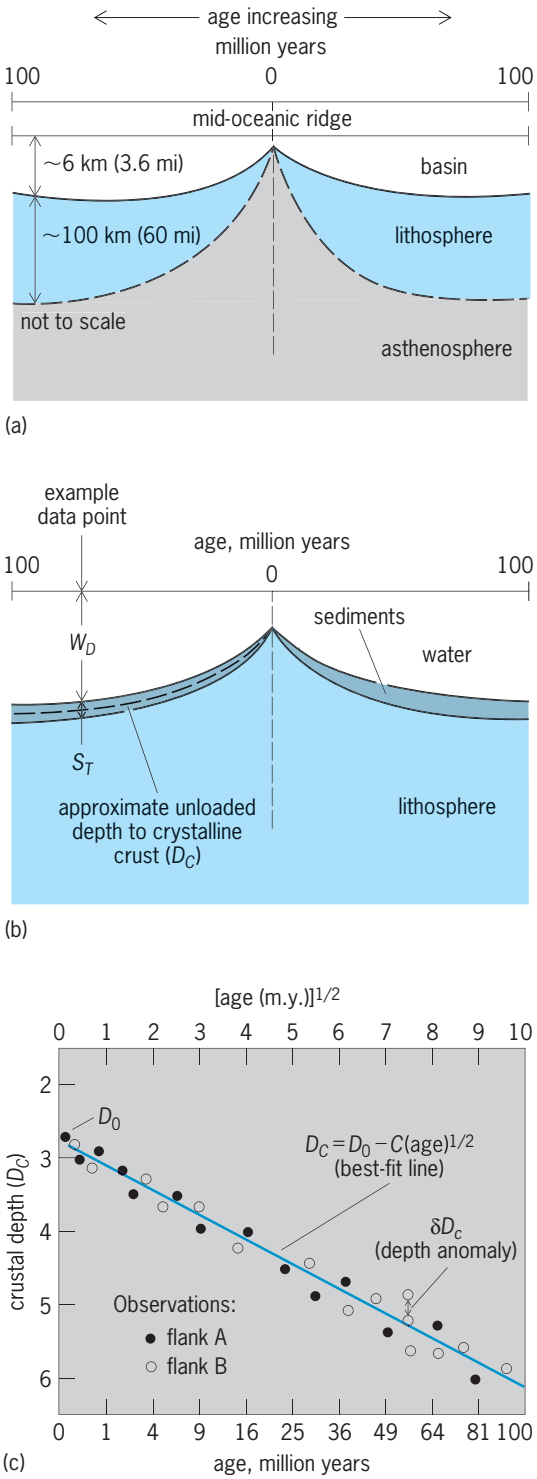

Fig. 4. Ocean depth anomalies. (*a*) Schema of the lithosphere, asthenosphere, and varying ocean depths away from the axis of mid-oceanic ridge spreading system. (*b*) Parameters used to compute adjusted crustal depths and to analyze depth-age relationships. (*c*) Hypothetical crustal depth (D_C) versus (age)$^{1/2}$ plot showing linear relationship predicted by theory. Differences between the actual observations and the statistically best-fitted line represent residual depth anomalies. D_C (predicted) yield positive or negative depth anomalies. W_D = water depth; S_T = sediment thickness; D_0 = crustal depth at zero age crust; δD_C = depth anomalies (difference between the observed and predicted crustal depths). (*After D. E. Hayes, Mapping oceanic depth anomalies: In search for indicators of asthenospheric convection, Lamont-Doherty Geological Observatory Yearbook, 1982–1983*)

ocean, intermixed either with terrigenous sediments or in near "pure form" in those areas inaccessible to terrigenous sedimentation.

Biogenic sediments are composed mostly of the undissolved tests of siliceous and calcareous microorganisms, which settle slowly to the sea floor. This steady so-called pelagic rain typically accumulates at rates of a few centimeters per thousand years. The composition and extent of the input to the biogenic sediment depend upon the composition and abundances of the organisms, which in turn are largely reflective of the water temperature and the available supply of nutrients. The Pacific equatorial zones and certain other regions of deep ocean upwelling are rich in nutrients and correspondingly rich in the microfauna and flora of the surface waters. Such regions are characterized by atypically high pelagic sedimention rates. *See* UPWELLING.

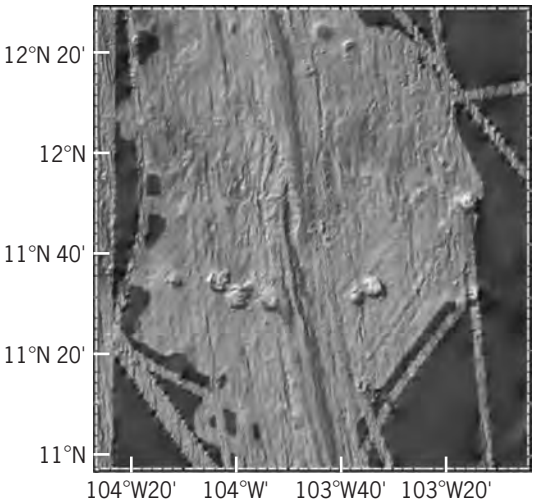

Fig. 5. Digitized contoured map of a ~200 km (120 mi) stretch of the East Pacific Rise (Ridge) near 11°N, 104°W. The data were collected using a multibeam swath mapping system. The depth of the ridge crest varies from south to north and illustrates an intermediate scale segmentation of the ridge. The area slightly north of the map center shows an overlapping spreading center, another example of ridge segmentation. (*Taken from Ridge 2000 MBS data and displayed using GeoMapApp software*)

As the ridge flanks subside to great depths, often in excess of 6 km (4 mi) over a period of 100 m.y., the small-scale topography is gradually modified by pelagic sediments. Once sediments reach the sea floor, their ultimate disposition depends on local environment. Fine sediments can be held in suspension for long periods or can be transported along the bottom, ultimately to be deposited in a more tranquil environment. Spectacular bedforms such as sediment waves sometimes result and provide important information regarding the nature and vigor of present and ancient bottom-current activity. In some cases, the bottom currents can be so strong as to prevent the deposition of fine-grained sediments altogether, essentially scouring the sea floor free of any significant sediment. Over most of the deep-sea floor, bottom currents are sluggish, and most pelagic sediments

are deposited in a relatively tranquil environment as a uniform blanket of sediments on the existing sea-floor relief. If steep sea-floor slopes exist, some sediments may slump off to the sides, effectively smoothing the primary topography upon which they are deposited.

Where the processes of sedimentation have been acting long enough and the deposition locale is accessible to the sediments derived from the continents, the original relief of the ocean crust can become completely buried. Many terrigenous deposits represent relatively impulsive events (for example, major submarine landslides on the continental margin or large turbidity flows), where large volumes of material are transported great distances over very short periods. The principal mechanism of deposition is a leveling process that fills first the topographic lows and, ultimately, can lead to burial of the entire initial sea-floor relief. This result is an extremely flat, topographically featureless abyssal plain. Such abyssal plains characteristically lie between the continental margins and the exposed flanks of the mid-oceanic ridge system. The abyssal plains areas have gradients that are less than 1:1000 and are typically underlain by a variety of discrete layers of transported sediments. *See* BASIN; MARINE SEDIMENTS; TURBIDITY CURRENT.

Continental margins. The continental margins lie at the transition zone between the continents and the ocean basins and mark a major change from deep to shallow water and from thin to thick continental crust. The continental margins consist of two broad categories: rifted margins and convergent margins.

Rifted margins. Good examples of rifted margins are found bounding the Atlantic Ocean (Fig. 2). These margins represent sections of the South American and North American continents that were once contiguous to west Africa and northwest Africa, respectively. These supercontinents were rifted apart 160–200 m.y. ago as the initial stages of sea-floor spreading and the birth of the present Atlantic Ocean sea floor (Fig. 2). *See* CONTINENTS, EVOLUTION OF; SUPERCONTINENT.

Continental margins are proximal to large sources of terrestrial sediments that are the products of continental erosion. The margins are also the regions of very large vertical motions through time. This vertical motion is a consequence of heating and subsequent cooling of the rifted continental lithosphere and subsidence. During initial rifting of the continents, fault-bound rift basins are formed that serve as sites of deposition for large quantities of sediment (**Fig. 6**). These sedimentary basins constitute significant loads on the underlying crust, giving rise to an additional component of margin subsidence. The continental margins are of particular importance also because, as sites of thick sediment accumulations (including organic detritus), they hold considerable potential for the eventual formation and concentration of hydrocarbons. As relatively shallow areas, they are also accessible to offshore exploratory drilling and oil and gas production wells.

Also found in selected continental margin environments are the unusual formations known as gas hydrates. Hydrates are frozen gases, typically composed of methane, higher-order hydrocarbons, and carbon dioxide, which exist in a stable state only within a relatively narrow range of temperature and pressure. The occurrence of such hydrates is now believed to be far greater than past estimates. Hydrates are often associated with a bottom-simulating reflector, which is thought to represent a sharp drop in acoustic impedance created by accumulations of free methane gas just below the base of the gas hydrate zone. *See* HYDRATE.

There is growing evidence that the hydrocarbon gases (mostly methane), entrapped in sub-sea floor gas hydrates, constitute a significant component of global methane and is likely to be the focus of intense ongoing investigations. The potential of gas hydrates to contribute to the global energy inventory of recoverable hydrocarbons remains the subject of debate. *See* OIL AND GAS, OFFSHORE.

Many sedimentary aprons or submarine fans are found seaward of prominent submarine canyons that incise the continental margins. Studies of these sedimentary deposits have revealed a number of unusual surface features that include a complex system of submarine distributary channels, some with levees. The channel systems control and influence sediment distribution by depositional or erosional interchannel flows. Fans also result from major instantaneous sediment inputs caused by large submarine mass slumping and extrachannel turbidity flows. *See* SUBMARINE CANYON.

The present coastline shows no particular geological significance as a boundary, and continental and oceanic crustal structures may or may not lie close to the present shoreline. It has been well established that the sea level has changed many times, fluctuating by as much as 300–400 m (980–1300 ft) during the last 200 m.y. Such changes in sea level are primarily due to the presence or absence of major continental ice sheets, which store vast quantities of water, thereby reducing the total water available to fill the ocean basins. During periods of glaciation, sea level falls. *See* GLACIAL EPOCH.

Low stands of sea level coincide with times of rapid transport of terrestrial sediments across the continental shelves and into the deep ocean basins. During interglacial periods, when little water is stored as continental ice sheets, sea level rises. Accordingly, much of the terrestrial material transported to

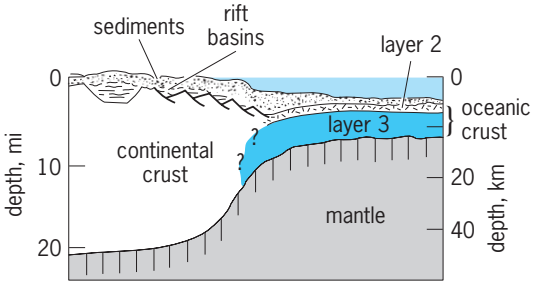

Fig. 6. Schematic cross section of a rifted continental margin. Note dramatic change in total crustal thickness across the margin, fault-bounded sedimentary basins, and large thicknesses of land-derived sediments making up the continental rise.

the sea is trapped on the broad continental shelves, and results in the outbuilding or progradation that characterizes many present-day continental margins. The interplay between sea-level changes and sedimentation yield gives rise to the complicated interfingering of sediment deposits along the continental margins.

Another method of significantly altering sea level is related to large, global differences in the rate at which sea floor is created at the mid-oceanic ridge systems. When sea floor is created extremely rapidly, that is, the plates are moving apart very rapidly, the profile of the mid-oceanic ridge is relatively broad. Because the depths of the mid-oceanic ridges are a direct function of crustal age, fast spreading results in a greater fraction of the ridge appearing at shallow elevations than for a mid-oceanic ridge created by very slow spreading (Fig. 3). Hence, a mid-oceanic ridge system with a broad profile will effectively displace more water from the ocean basin regions than one with a steep profile. The displaced water must move onto the adjacent continents in association with the eustatic sea-level rise. This is one explanation for the occurrence of a major global transgression of the ocean onto the continental areas during the middle Cretaceous period (85–115 m.y.a.). *See* RIFT VALLEY.

Convergent margins. In contrast to the rifted margins, the continental margins that typically surround the Pacific Ocean represent areas where plates are colliding (Fig. 1). As a consequence of these collisions, the oceanic lithosphere is thrust back into the interior of the Earth; the loci of underthrusting are manifest as atypically deep ocean sites known as oceanic trenches (**Fig. 7**). The processes of subducting the oceanic lithosphere give rise to a suite of tectonic and morphologic features characteristically found in association with the oceanic trenches. An upward bulge of the crust is created seaward of the trench that represents the flexing of the rigid oceanic crust as it is bent downward at the trench. The broad zone landward of most trenches is known as the accretionary prism and represents the accumulation of large quantities of sediment that was carried on the oceanic crust to the trench. Because the sediments have relatively little strength, they are not underthrust with the more rigid oceanic crust, but they are scraped off. In effect, they are plastered along the inner wall of the trench system, giving rise to a zone of highly deformed sediments. These sediments derived from the ocean floor are intermixed with sediments transported downslope from the adjacent landmass, thus creating a classic sedimentary melange. *See* CONTINENTAL MARGIN; SEDIMENTOLOGY.

Anomalous features. In addition to the major morphologic and sediment provinces, parts of the sea floor consist of anomalous features that obviously were not formed by fundamental processes of seafloor spreading, plate collisions, or sedimentation. Such features are nonetheless important and represent significant components of the ocean-floor relief. Examples are long, linear chains of seamounts and islands. Many of these chains are thought to reflect

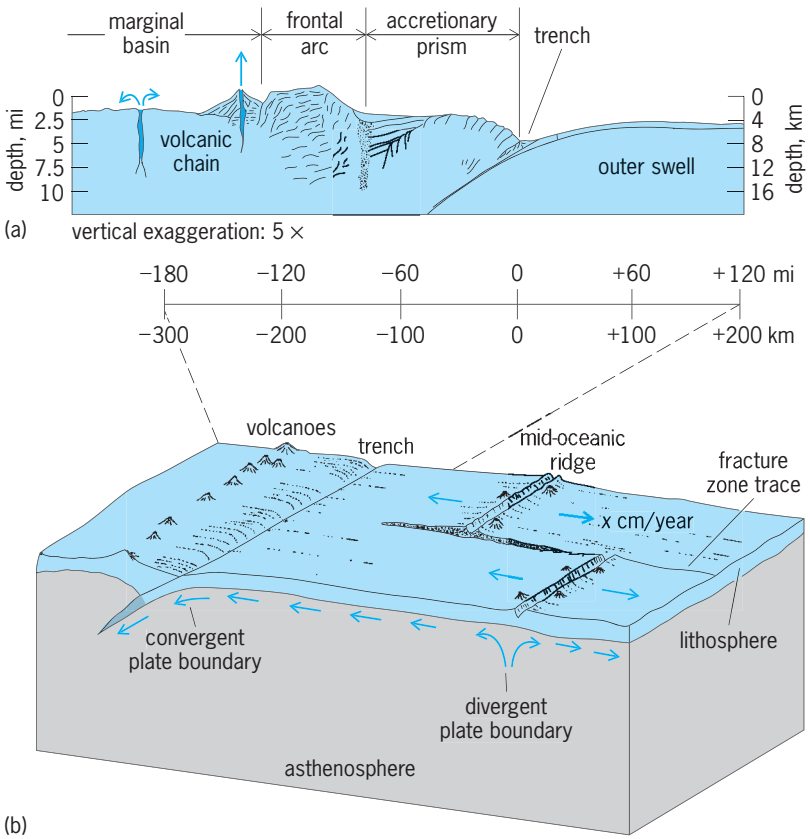

Fig. 7. Oceanic trench and associated tectonic and morphologic features. (*a*) Cross section of typical island arc system showing tectonic units and terminology (*after D. E. Karig and G. E. Sharma III, Subduction and accretion at trenches, Geol. Soc. Amer. Bull., 86:377–389, 1975*). (*b*) Schematic cross section showing the relationship of relative plate motions to plate boundaries and associated sea-floor features. At divergent plate boundaries (mid-oceanic ridges), shallow earthquakes occur. At convergent plate boundaries (trenches), shallow, intermediate, and deep earthquakes occur. Shallow earthquakes occur along active transform zones; fracture zones represent the relict traces of crustal discontinuities originally formed within a transform zone (*after B. Isacks, J. Oliver, and L. R. Sykes, Seismology and the new global tectonics, J. Geophys. Res., 73:5855–5900, 1968*).

the motion of the oceanic plates over hot spots that are fixed within the mantle. Hot spots carry magma through the oceanic crust to the surface, resulting in volcanic trails, which serve to define the relative motion of the plate over the hot spot. *See* MAGMA; SEAMOUNT AND GUYOT.

The presence of large, anomalously shallow regions known as oceanic plateaus may also represent long periods of anomalous regional magmatic activity that may have occurred either near divergent plate boundaries or within the plate. Alternatively, many oceanic plateaus are thought to be small fragments of continental blocks that have been dispersed through the processes of rifting and spreading, and have subsequently subsided below sea level to become part of the submarine terrain.

Large igneous provinces (LIPs) are noteworthy examples of vast outpourings of mafic lava that are found both on continents as large flood basalts and on the ocean floor primarily as large oceanic plateaus. While the existence of these features has been known for some time, their genesis was largely assumed to be the result of perturbations to normal plate tectonic processes. It is now believed that the

emplacement of many LIPs was so rapid (<1–2 my) and the volume of material so large, that creation of new oceanic crust as attributed to sea-floor spreading could not accommodate the rate of magma production represented by the LIPs.

It has been proposed by some investigators that the LIPs were formed by major rising plumes created when rapid overturn of the mantle was initiated by sinking of accumulated "cold material" from the upper mantle at about 650 km (400 mi). These rising plumes then poured out at the Earth's surface, creating LIPs. The last major episode of LIP formation was during the Cretaceous Period. Examples of oceanic LIPs are Kerguelen Plateau, Ontong-Java Plateau, Rio Grande Rise, and Broken Ridge. The reason these and similar features have been the subject of renewed interest is the realization that such huge outpourings of basalt would be accompanied by the extensive release of associated gases (such as CO_2 and SO_2), which in turn could have profoundly impacted the environment and initiated mass extinctions of various life-forms.

Other important features of the ocean floor are the so-called scars represented by fracture zone traces that were formed as part of the mid-oceanic ridge system, where the ridge axis was initially offset. Oceanic crusts on opposite sides of such offsets have different ages and hence they have different crustal depths. A structural-tectonic discontinuity exists across this zone of ridge axis offset known as a transform zone (Fig. 7*b*). Although relative plate motion does not occur outside the transform zone, the contrasting properties represented by the crustal age differences create contrasting topographic and subsurface structural discontinuities, which can sometimes be traced for great distances. Fracture zone traces define the paths of relative motion between the two plates involved. Those mapped by conventional methods of marine survey have provided fundamental information that allows rough reconstructions of the relative positions of the continents and oceans throughout the last 150–200 m.y. The study that deals with the relative motions of the plates is known as plate kinematics. *See* TRANSFORM FAULT.

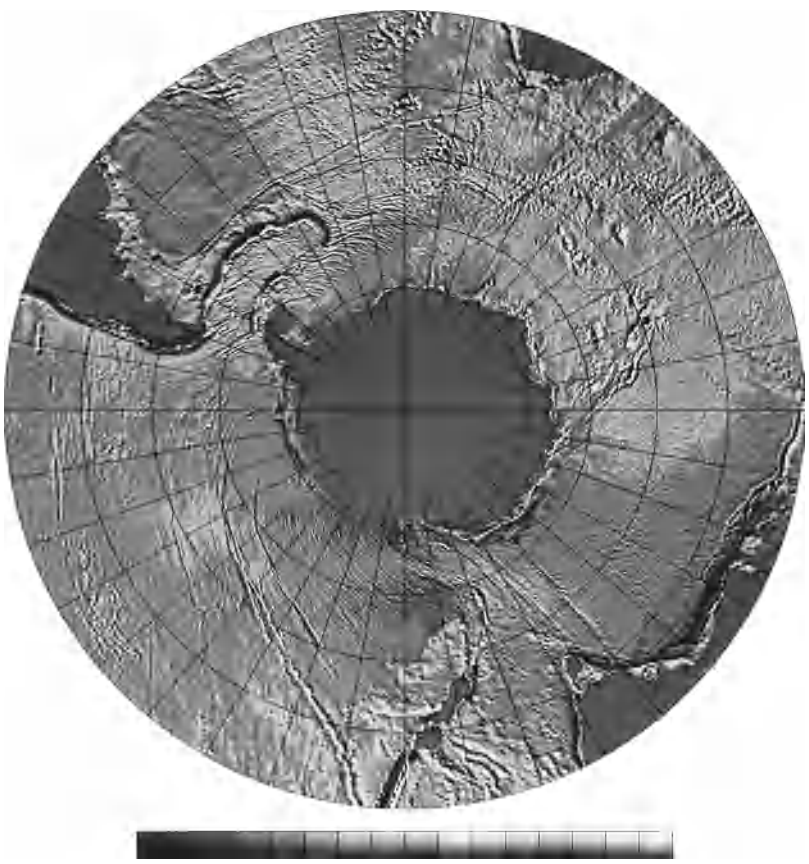

Fig. 8. Free-air gravity map of the Southern Ocean derived from analysis of *SEASAT* and *GEOSAT* altimeter data. Light areas represent positive anomalies; dark areas represent negative anomalies. Note long linear expressions of fracture zone traces. (*From W. Haxby and D. E. Hayes, The Gravity Field of the Circum-Antarctic, Antarctic Research Series, 1990, copyright by the American Geophysical Union*)

In 1978, precise satellite altimeter observations from the *SEASAT* mission, coupled with similar but more precise measurements from a later *GEOSAT* mission in 1987, revealed the detailed variations in sea-surface height over the entire water-covered surface of the Earth. The relief of the sea surface responds to lateral variations of mass within the crust below the ocean. These lateral variations in mass are due to the topography of the sea floor, to the crustal structure associated with such topographic variations, and to other lateral subsurface variations in density. Through sophisticated analysis of the *SEASAT* and *GEOSAT* data, it is possible to create a map of subtle gravity variations (of the order of 1 part in 100,000) over the entire water-covered surface of the Earth.

The resultant gravity maps largely mimic the major topographic and structural features of the sea floor, which lie hundreds to thousands of meters below the ocean surface (**Fig. 8**). Because gravity maps represent not only the major relief features of the sea floor but also the associated subsurface structures, various large-scale topographic features may be manifest in different ways. For example, the fast-spreading mid-oceanic ridge system of the South Pacific shows a far less conspicuous gravity signal associated with it than the more slowly spreading mid-oceanic ridge system found in the South Atlantic (Figs. 1 and 8).

The extraordinary success of this satellite altimeter technique has provided enough detailed information that many fracture zone traces can now be followed, unequivocally, across entire basins. Hence, the earlier relative positions of the plates throughout the last 200 m.y. can be determined with much greater accuracy than was previously possible. Because the satellite coverage is global and consists of fairly closely spaced ground tracks, many features of the ocean floor that were previously unrecognized, because of limited surface-ship data, are now being predicted by utilizing the satellite observations; many have subsequently been verified by conventional surface-ship surveys.

Marginal seas. The sea-floor features described so far are representative of the main ocean basins and reflect their evolution mostly through processes of plate tectonics. Other, more complicated oceanic

regions, typically found in the western Pacific, include a variety of small, marginal seas (back-arc basins) that were formed by the same general processes as the main ocean basins. These regions define a number of small plates whose interaction is also more or less governed by the normal tenets of plate tectonics. One difficulty in studying these small basins is that they are typified by only short-lived phases of evolution. Frequent changes in plate motions interrupt the process, creating tectonic overprints and a new suite of ocean-floor features. Furthermore, conventional methods of analyzing rock magnetism, heat flow, or depths of the sea floor to roughly date the underlying crust do not work well in these small regions. The small dimensions of these seas bring into play relatively large effects of nearby tectonic boundaries and render invalid key assumptions of these analytical techniques. The number of small plates that actually behave as rigid pieces is not well known, but it is probably only 10–20 for the entire world.

Exemplary Techniques

A number of advances have permitted detailed studies and detailed mapping of the sea floor. These include accurate determination of a ship's position and velocity, scanning systems, and comprehensive geophysical measurements.

Determination of ship's position and velocity. The ability to determine a ship's position and velocity is often the factor that controls the scale of features that can be mapped on the sea floor. Since the 1950s, the development of electronic navigational tools has greatly enhanced positioning capabilities. The introduction of navigational satellites (Transit Satellite System) and their availability to the research community in the mid-1960s allowed approximately 10 ship fixes to be obtained each day with an accuracy of about 0.5 km (0.3 mi) or better. This was a vast improvement over the celestial navigation and electronic systems in use at the time.

The Global Positioning System (GPS) now exists which uses a constellation of 24 satellites orbiting the Earth at high elevations. The concept involved is that at any point on the Earth's surface, three or four of these satellites would be in view at all times, thereby allowing position ranging on them. The satellites' positions are known from ground-tracking stations. The Global Positioning System provides continuous position fixing that is accurate to much better than 0.16 km (0.1 mi) [for surface vessels] and is available at all times, under all weather conditions, and over the entire surface of the Earth. This system will allow marine geologists to study effectively features on the sea floor whose horizontal dimensions are of the order of tens of meters or smaller. *See* SATELLITE NAVIGATION SYSTEMS.

Multibeam and side-scan swath systems. Since the mid-1970s very significant improvements have been made in the instruments available for investigating the sea floor and its subsurface structures. The analytical techniques that can be applied to quantitative interpretation and understanding of the key operative processes that shape the Earth have also improved. Since World War II, the conventional method of determining the depth of the ocean has been by the transmission of acoustic signal pulses from the ship to the sea floor, where they are reflected to the sea surface and detected. The elapsed travel time is measured, and either by assuming the velocity of sound in the water or by measuring it, the depths of the ocean can be determined with a relative accuracy of approximately 1 m (3.3 ft) and with an absolute accuracy of a few tens of meters. Historically, the acoustic pulses used have not been very directional, so that the sound waves that reached the sea floor were actually reflected off a fairly large area (the footprint) of the sea floor. Hence, some average depth, taken over the scale of that footprint, was measured. Interpretations of the returning echo were further complicated by the presence of side echos reflected from sea-floor features off to the side of the ship.

Today's advanced systems, generically known as multibeam echo sounders, provide marine scientists with the ability to measure the sea-floor depth with greater precision and over smaller acoustic footprints. The ability to generate an array of directed acoustic pulses, subtending 45 to 90°, across the ship's track (**Fig. 9**) provides a means for the simultaneous measurement of the ocean depths across a swath up to twice the water depth. Because the individual soundings making up this swath are fixed with respect to one another, the details of the sea-floor relief can be determined with remarkable clarity. The data can be analyzed onboard the survey ship in near-real time.

The output of multibeam echo sounders can be displayed as a bathymetric contour map centered on the ship's track (Fig. 9). Contours are generated

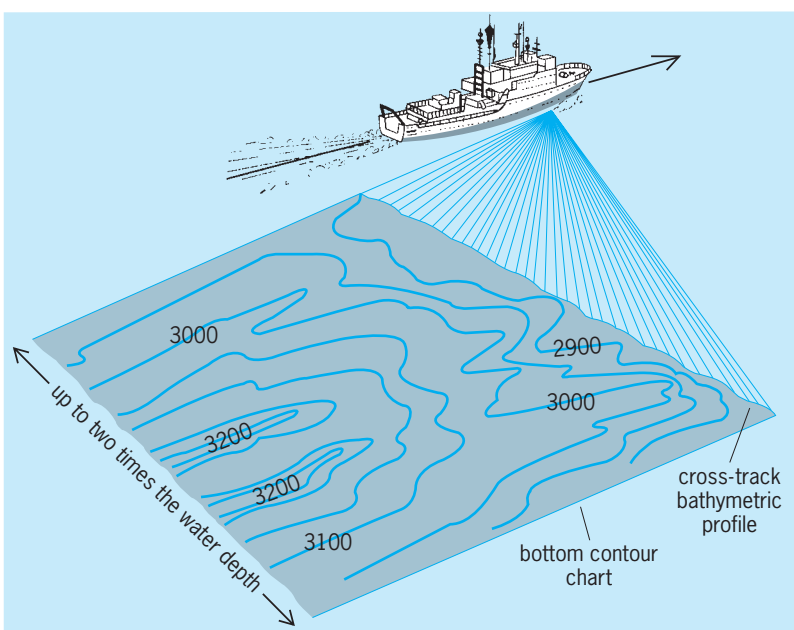

Fig. 9. A multibeam bathymetric swath mapping system. The bottom contour chart is produced on board the vessel in real time. The width of the mapped swath depends on the specific system used but can be up to two times the water depth. Contours are given in meters. 1 m = 3.3 ft. (*Krupp-Atlas Elektronik*)

continuously onboard within a few minutes of the acquisition of the soundings. If a survey is conducted so that tracks are spaced no greater than the swath width, complete sounding coverage of the sea floor is obtained, and small-scale features with bathymetric relief of a few meters can be identified and mapped over large areas in a manner not possible with the older techniques (Fig. 5). *See* ECHO SOUNDER; HYDROPHONE.

Side-scan sonars constitute a similar family of deep- or shallow-towed survey instruments. These systems emit sound at a much higher frequency than conventional single-beam or multibeam echo sounders, and the swaths can extend to several kilometers on either side of the towed instrument. The returned acoustic signals result from backscattering of the sound pulses due to variations in the small-scale roughness and to the varying reflective properties of the sea floor. Side-scan sonar instruments provide a qualitative picture of the sea floor. The near-surface texture and the planform of the sea-floor relief is recorded over a large area, but there is virtually no subsurface information measured. Side-scan instruments provide for maximum areal mapping coverage at minimal expense. The side-scan acoustic image of the sea floor is analogous to an aerial photograph or radar image of a land area. *See* ACOUSTIC SIGNAL PROCESSING; SONAR; UNDERWATER SOUND.

Comprehensive geophysical measurements. Many geophysical properties of the sea floor and its underlying materials are not very diagnostic. For example, a variety of rocks may exhibit the same seismic-wave velocity, the same magnetic properties, or the same density. Gravity variations occur because of lateral differences in the near-surface distribution of mass caused by different rock types, by different structures, and by sea-floor relief. Gravity measurements alone cannot reveal a unique causative geological feature for each observed gravity anomaly, since several different mass distributions could give rise to the same observed variations in gravitational field. But, by adopting reasonable constraints and using a general knowledge of geology and geophysical properties of the Earth, interpretations of a collection of different types of observations can be reduced to a few possibilities with geological plausibility. *See* GEOPHYSICAL EXPLORATION.

Gravity. The ability to measure small variations in the Earth's gravitational field to a few parts in a million in the presence of large disturbing accelerations due to the ship's motion requires very sophisticated instrumentation. Because the measurement of gravity is difficult and the instrumentation expensive, most ships do not carry gravimeters. Consequently, the total gravity track coverage for the entire world oceans is relatively limited.

Satellite observations have allowed the indirect measurement of subtle variations in the Earth's gravitational field over almost the entire world ocean. These variations have been important in refining knowledge about the details of plate kinematics and in identifying the presence of unknown or anomalous topographic or subsurface ocean-floor structures that were previously unknown. The satellites *SEASAT* and *GEOSAT* measure with great accuracy the distance between the satellite and the ocean

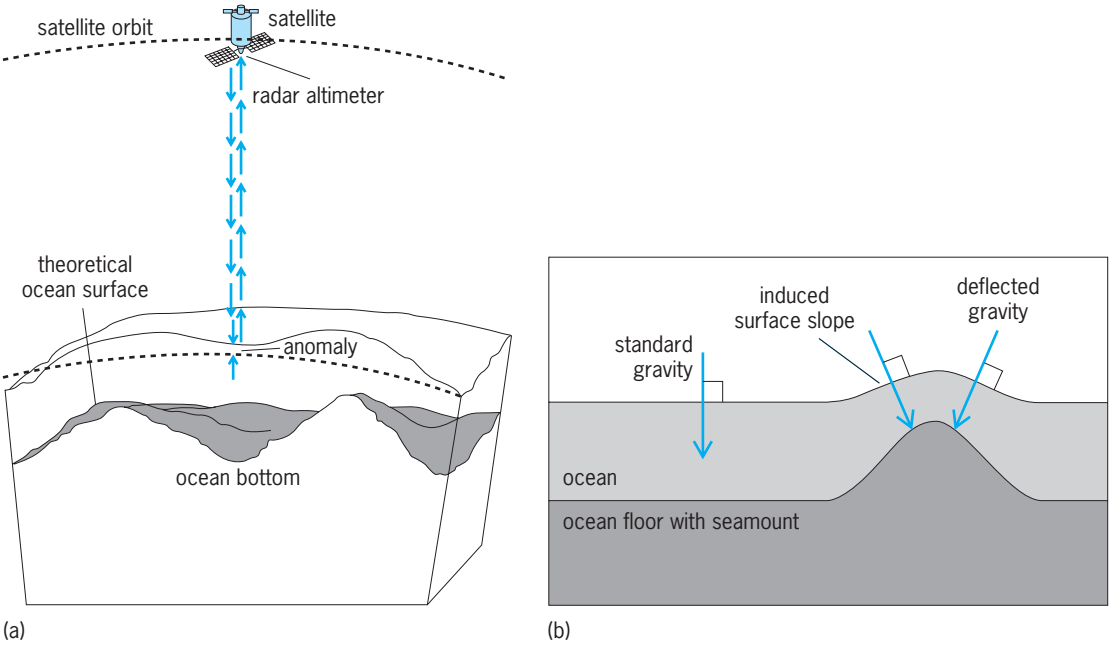

Fig. 10. Satellite-derived bathymetry. (*a*) An Earth-orbiting radar in space cannot see the ocean bottom, but it can measure ocean surface height variations induced by ocean floor topography. A mountain on the ocean floor adds to the pull of Earth's gravity and changes its direction subtly, causing extra water to pile up around the mountain. For example, a mountain on the ocean floor that is 2000 m (6600 ft) tall produces a sea surface bump only 20 cm (8 in.) tall; this is measurable from space. The ultimate resolution of this method is limited by regional ocean depth. (*b*) The tilt in the direction of gravity, called a "deflection of the vertical," is equal to the slope of the sea surface. (*After D. T. Sandwell, S. T. Gille, and W. H. F. Smith, eds., Bathymetry from Space: Oceanography, Geophysics, and Climate, Geoscience Professional Services, Bethesda, MD, June 2002, http://www.igpp.ucsd.edu/bathymetry_workshop*)

surface heights. The detailed shape of the ocean surface is in large part a response to small variations in the Earth's gravitational field. Variations of gravity over distances longer than about 30 km (18 mi) and shorter than about 2000 km (1200 mi) can be determined with a reliability of a few milligals (1 mgal $\simeq 10^{-6}$ of the Earth's average gravity value). This range scale encompasses many of the tectonic and structural and topographic features of interest in the sea floor.

Bathymetry from space. The recent novel applications of satellite altimeter measurements have led to the modeling and prediction of submarine topographic features with wavelengths of 15–2000 km (9–1240 mi). D. T. Sandwell and W. H. F. Smith have applied this analysis on a global basis. By combining this intermediate-scale (interpreted) topography with the deterministic regional topography obtained by ships equipped with echo sounders, a comprehensive model of the sea-floor topography that combines the two approaches can be generated. These model results can be stored digitally for related studies or future revisions.

The Sandwell and Smith methodology involves calculating the free-air gravity anomalies from actual measurements of sea surface undulations. This is accommodated by presuming the sea surface represents a specific gravity equipotential surface known as the geoid. Using the N-S and E-W gradients of the geoid (that is, the surface undulations mapped) the vertical gradient of the geoid can be calculated, thus providing the free-air gravity anomaly. Variations in the free-air gravity anomaly are caused in part by lateral and vertical variations of mass represented by sea-floor topography (for example, seamounts) and the variable subsurface structures associated with them (**Fig. 10**). The degree of "correlation" of free-air anomalies with topography depends greatly on the horizontal dimensions (scale) of the topographic features in question. Generally, the greater the width of the feature, the smaller the correlation between gravity and topography. The transfer functions used to "convert" gravity anomalies to predicted causative topography are dependent on the wavelengths of the features in question. The remote sensing of sea-floor features using satellite altimetry measurements constitutes a significant advance in our knowledge of small- to intermediate-scale submarine topographic relief. This advance is especially significant for the high latitude regions of the Southern Ocean (south of ~30°S) where the bathymetry mapped by conventional ship-borne echo sounders is poorly defined due to the sparseness of survey ship tracks. *See* EARTH, GRAVITY FIELD OF.

Magnetics. Measuring the variation of the magnitude of the Earth's magnetic field is a routine part of most marine survey operations. To make measurements, a simple sensor, usually a nuclear precession magnetometer, is towed several hundred yards behind the vessel. The magnetometer can provide an absolute measure of the total magnetic field strength every few seconds. The reduction and interpretation of the magnetic data can be complicated in part because the Earth's magnetic field undergoes a variety of temporal variations not directly related to the magnetic properties of the Earth's crust.

Magnetic surveys of the world ocean and their subsequent interpretations have provided perhaps the most conclusive evidence for sea-floor spreading and plate tectonics. The Earth's magnetic field is known to reverse its polarity every few hundred thousand years (although not exactly regularly). Newly formed ocean crust at the axis of the mid-oceanic ridge acts like a magnetic tape recorder of the direction and magnitude of the Earth's field at the time the volcanic rocks cool below about 500°C (930°F; the Curie temperature) [**Fig. 11**]. This has led to the creation of small systematic anomalies in the Earth's magnetic field that, when mapped, describe a series of distinctive magnetic lineations (stripes) that are oriented parallel to the axis of the mid-oceanic ridge system. Such magnetic stripes are generally symmetrically disposed about the axis of the mid-oceanic ridge and are offset across transform faults and fracture zones. Because the lineations record the variations

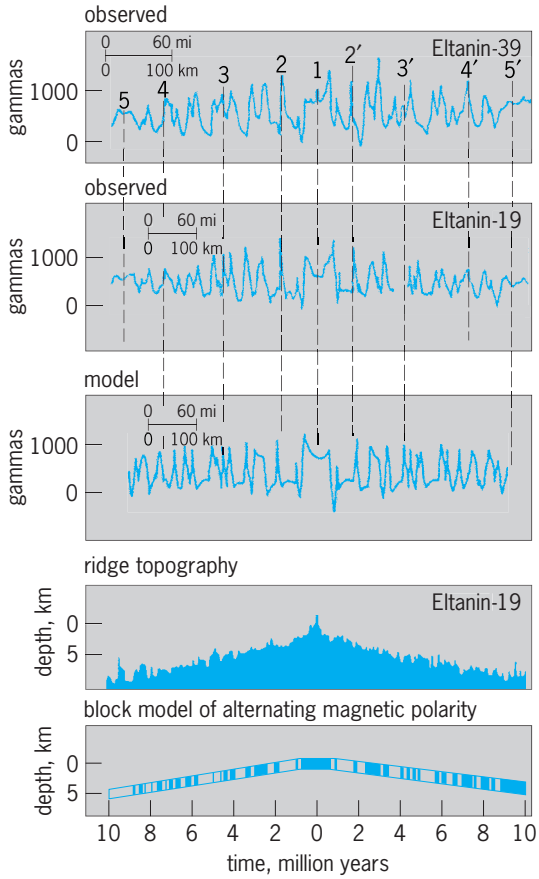

Fig. 11. Observed magnetic anomalies from two *Eltanin* cruises and ridge topography from cruise 19. The theoretical magnetic anomaly (model) computed from the simple block model of alternating magnetic polarity (near the bottom) can be easily correlated with the observed profiles. Profiles Eltanin-39 and Eltanin-19 are from areas more than 10,000 km (6000 mi) apart. Note symmetric distribution of characteristic peaks (2 and 2′, 3 and 3′, and so forth) about the ridge axis. 1 km = 0.6 mi. (*After D. E. Hayes and W. C. Pitman III, Marine geophysics and seafloor spreading in the Pacific-Antarctic area: A review, Antarc. J., 5(3):70–72, 1970*)

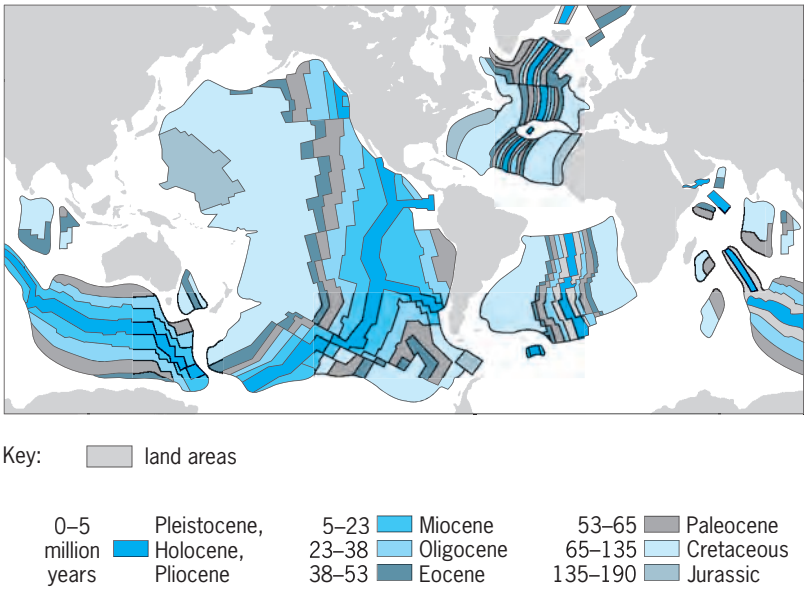

Fig. 12. Map showing age of the ocean crust based on interpretation of magnetic anomaly patterns such as shown in Fig. 10. (*After W. C. Pitman, R. L. Larson, and E. M. Herron, The age of the ocean basins, Geol. Soc. Amer. Map and Chart. Ser. MC-6, 1974*)

in the magnetic polarity of the Earth over time, a sequence of these anomalies can be used to date the formation of the anomalies. Relatively large time slices of the geological polarity history of the Earth can be distinguished from one another. The slightly different pattern of Earth polarity reversals allows the observation of magnetic anomalies to determine when in geologic time the crust underlying those magnetic anomalies was created. This technique has allowed the age of the ocean crust to be mapped on a global basis (**Fig. 12**). *See* CURIE TEMPERATURE; GEOMAGNETISM; PALEOMAGNETISM; ROCK MAGNETISM.

Heat flow. For the deep ocean basins, there are good theoretical and empirical models to account for the observed heat flow from measured geothermal gradients in the ocean sediment and crust. In general, heat flow decreases as a simple function of the age of the crust. Predictable variations in heat flow reflect the heat lost through conductive cooling since the initial time of formation of the underlying crust at the axis of the mid-oceanic ridge system. Young ocean crust is characterized by high heat flow, approximately 500 mW/m^2, whereas old (>100 m.y.) oceanic crust has heat flow that asymptotically approaches a value of about 50 mW/m^2.

In relatively young ocean crust, some heat may be transported by processes other than conduction. Global observations of heat flow suggest that for crust younger than about 40 m.y., observed heat flow often falls well below that predicted by simple conductive cooling models. Hydrothermal circulation in the uppermost parts of the crust may serve to remove significant amounts of heat, thereby resulting in artificially low values of conductive heat flow. As the crust ages and gradually becomes covered by sediments, the effective plumbing system for hydrothermal circulation, the fractured upper crust, eventually becomes sealed, thereby preempting any additional heat loss except by conduction. The amount of heat lost through hydrothermal processes is unknown and can be estimated only by comparing observations with theoretical predictions. Hence, the total integrated annual heat flow through the sea floor is still somewhat uncertain. *See* EARTH, HEAT FLOW IN.

Seismic reflection. The most commonly used technique for investigating the subsurface structure of the sea floor is seismic profiling. In principle, this technique is like depth sounding, but it involves the transmission of low-frequency acoustic energy to the sea floor and subsurface horizons. Low-frequency acoustic energy (about 10–100 Hz) is less rapidly attenuated within the sediments than the high-frequency energy of depth-sounding instruments. Hence, a portion of the outgoing acoustic signal penetrates the sea floor, eventually to be reflected to the surface at major geologic (lithologic) boundaries. Geologic materials with contrasts in density and seismic-wave velocity give rise to the seismic reflections.

Modern seismic techniques involve the recording of the reflected energy at a large number of receiver locations, fixed within a seismic streamer that is towed behind the ship (**Fig. 13**). A typical streamer may consist of 240 or more independent recording channels, equally spaced in the streamer and extending 6 km (3.6 mi) or more behind the vessel. Information from different shots fired along the track and recorded at the various receivers can be gathered

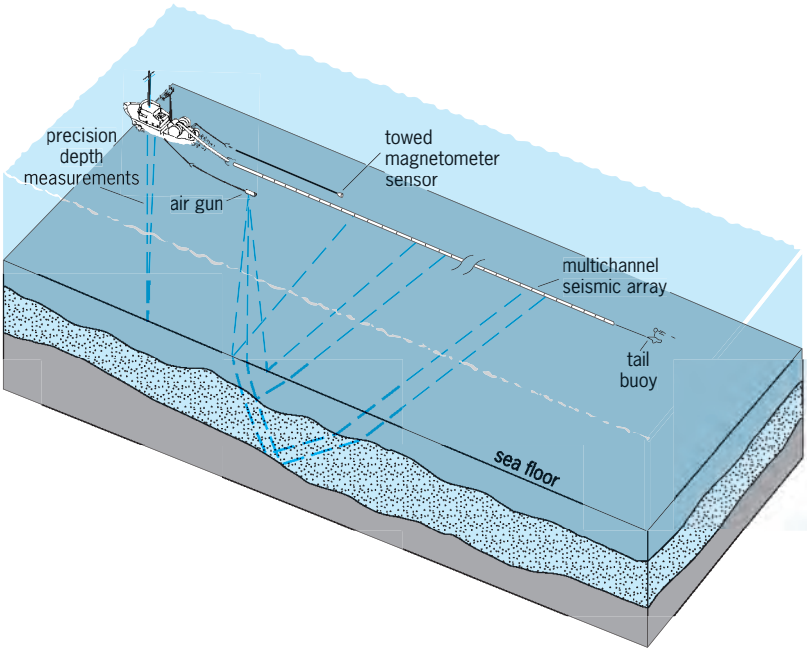

Fig. 13. Diagram of a fully equipped geophysical vessel and the subsurface geology, showing a few selected seismic-ray paths from the source (air gun) to individual sections of the receiving (multichannel seismic) array, whose length is in the range 2.5–6.0 km (1.5–3.6 mi). Precision depth measurements are made by using a high-frequency, hull-mounted single-beam transducer. The gravimeter is mounted in the ship, near its center of rotation. (*After Acquisition of Marine Surveying Technologies, United Nations Publication ST/ESA/17B*)

together by considering signals traveling through different offset distances but reflected from a subsurface common midpoint. This process is used to calculate velocity as a function of depth and to convert the reflection times of key reflecting horizons to depth. Also, by gathering, correcting, and adding the individual receiver signals (stacking), very weak or deep reflecting horizons can be detected.

The common midpoint technique is the most popular method used in offshore oil exploration. It is also used extensively for seismic imaging of the structures of the deep ocean crust and the continental margins. Subsurface structures to depths of several kilometers below the sea floor can be imaged. It is the crustal layer thicknesses, the disposition of the reflecting horizons, and the relative amplitudes of the seismic returns that allow the acoustic properties of the oceanic sediments and crust to be interpreted in terms of geological properties and of marine geological history.

Changes in the chemistry of the seawater, in bottom circulation, and in surface productivity, and impulsive sedimentation events can give rise to abrupt and dramatic changes in the sedimentary environment. Sedimentary deposits bounding such changes are usually characterized by contrasting physical properties. These contrasts in physical properties mark important events in the sedimentary history that give rise to discrete horizons that can be detected with seismic reflection techniques. Seismic techniques not only provide information regarding the total distribution of sediments but also allow correlation of key sedimentological events throughout time. Seismic horizons sampled in a small number of localities, but correlated and traced over very large regions, can be used to extrapolate limited sample results and to interpret the depositional history throughout entire ocean basins. *See* SEISMOLOGY.

Seismic exploration technology continues to produce new and better tools for imaging the sub-sea floor. The Global Positioning System (GPS) has facilitated a higher quality and better resolution of surveys.

The transition from conventional (2D) seismic surveys [Fig. 13] to 3D [**Fig. 14**] has been rapidly incorporated for use in offshore oil exploration. This precise and rapid acquisition methodology has in turn led logically to the time-dependent characterization (that is, 4D seismics) of known oil fields by seismic surveys. The availability of GPS now makes it practical to resurvey areas periodically and to reliably attribute observed differences in the images obtained to the temporal evolution of an oil field or perhaps to structural changes of an area of the sea floor following a tectonic event.

New implications of remote sensing. Ongoing studies of the sea floor and sub-sea floor involve largely the application of sophisticated remote-sensing techniques. Most notable is the use of a variety of high-frequency seismic instrumentation designed to provide higher resolution of acoustic images of the sub-sea floor. Surveys using side-scan tools now have the capacity to map the fine details of the sea floor and its shallow subsurfaces over much greater areal extent than before. Some of these tools provide deeper looks into the structure beneath the sea floor.

Similarly, the improvement of multibeam topographic swath-mapping and associated side-scan sonar has allowed detailed features of the sea floor to be mapped over extensive areas. These map results are helping scientists to discriminate between competing models of volcanism and faulting regarding the formation of the relief of the small-scale topographic features (such as abyssal hills) characteristic of mid-ocean ridges.

Paleobathymetry. Scientists are giving more attention to the importance of sea-floor relief in controlling the pathways of both shallow- and deep-ocean circulation. This circulation largely governs the meridional exchange of heat, which affects the pattern of global climate. With the advent of digital data files for bathymetry, crustal age, and sediment thickness, it is now possible to analyze the third dimension

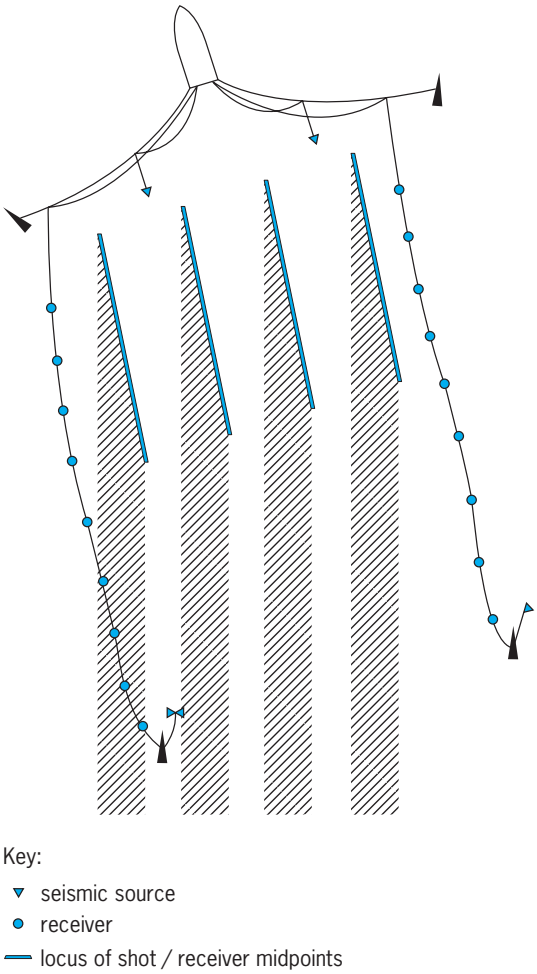

Fig. 14. Seismic profiling using four source arrays and two streamers (*courtesy of Western Geophysical*); hachured areas represent zones of the sea floor "swept over" by lines of shot/receiver midpoints, thereby creating an effective three-dimensional seismic reflection survey. (*E. J. W. Jones, Marine Geophysics, Wiley, 1999*)

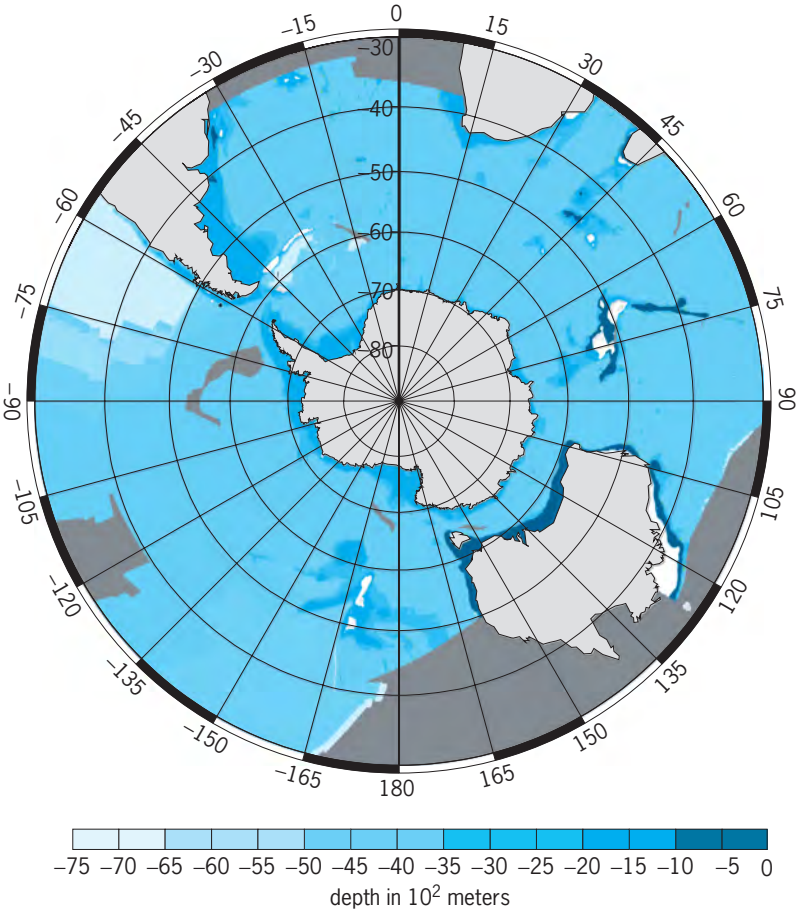

Fig. 15. Paleobathymetry of the circum-Antarctic at 50 m.y. Selected isobaths (in meters) are color-coded. Note the presence of major ocean circulation barriers between Australia and Antarctica, between Cape Horn and the Antarctic Peninsula, and between Antarctica and the Kerguelen Plateau. (*Dennis E. Hayes, Paleobathymetry of the Circum-Antarctic, in prep.*)

of plate reconstructions—paleobathymetry. Such reconstructions reveal the evolution of ridges, basins, and gateways and suggest their influences on variable paleoclimates. For example, although the Antarctic continent was located in a high-latitude position at least for the last 200 m.y., it has only been covered with ice for only about the last 35 m.y. The changing bathymetry of the circum-Antarctic must have played a profound role in modulating the climate in Antarctica and beyond (**Fig. 15**).

Sampling. The ability to sample reflecting horizons, especially deep horizons, is limited. Standard piston cores recover only the 10–20 m (30–60 ft) of the near-surface sediment. Several thousand piston cores have been taken from the world ocean floor. In order to sample horizons lying deep beneath the sea floor, drilling/coring techniques must be used. The initial drilling program began in 1968 as the Deep Sea Drilling Project (DSDP), which transitioned in 1984 into the IPOD (International Phase of Ocean Drilling Program). The most recent phase of cooperative marine geological drilling in the deep sea floor is the Integrated Ocean Drilling Program (IODP). It continues with the goals of solving key problems of Earth science that are best (perhaps only) addressed by direct deep sampling of the sub-sea floor using drilling. The IODP uses a dynamically positioned drilling vessel (the *JOIDES Resolution*) to drill into and sample the sediments and crystalline rock lying as much as 2000 m (6000 ft) below the sea floor.

Over 35 years of international collaborative research is attributable to the ocean drilling program. Since its inception, nearly 800 sites and more than 1700 holes have been drilled. They provide the only direct sampling of the deep sediments and basement rocks that compose the sea floor. These drilling sites have led to breakthrough findings, including evidence of the nature and pattern of paleoclimate change, confirmation of the Earth's paleomagnetic reversal time scale to ~200 mybp, the composition of the upper oceanic crust, the history of deep-sea sedimentation, the nature of past ocean circulation, hydrothermal processes at mid-ocean ridges and within the accretionary sediment wedges associated with subduction, and many others.

The principal difference in the new IPOD phase of the program will be the incorporation of more than one drilling platform. The most dramatic addition will be the Japanese vessel, *CHIKYU* [over 200 m (650 ft) long], that will be capable of riser drilling. Riser drilling includes the ability to "seal off" the drill hole if dangerous conditions are encountered. Riser drilling capability opens up a variety of new target areas previously considered unsafe to drill because of the possibility of unstable sub-sea-floor conditions (such as hydrocarbons or other fluids/gases under high pressure).

The next decade of research will see the implementation of "semipermanent" undersea observatories designed to make *in-situ*, real-time measurements of chemical, biological, and geophysical parameters relevant to sea-floor processes and phenomena. One specific example, the "Neptune Project," will focus on monitoring the activities on and around the Juan de Fuca Plate, offshore NE Pacific Ocean. The proposed network will include ~3000 km (2000 mi) of data and power cables laid on the sea floor which will service some 50 experimental sites through nodes onto the main cable network. Unmanned underwater vehicles will traverse the areas between experiment sites, and can be "repowered" from the cable nodes. They can also be deployed automatically or on command to sites of volcanic or tectonic events, relaying information to shore in real-time via direct links to the internet and facilitating control activities from shore-based labs. *See* UNDERWATER VEHICLES.

The long-term unmanned sea-floor observatories, judiciously deployed in geologically active seafloor regions, will mark one important new approach to marine geological research in the coming decades.

Dennis E. Hayes

Bibliography. C. Emiliani (ed.), *The Sea,* vol. 7: *The Oceanic Lithosphere*, 1981; B. C. Heezen and C. D. Hollister, *The Face of the Deep*, 1971; J. Kennett, *Marine Geology*, 1982; E. Seibold and W. H. Berger, *The Sea Floor: An Introduction to Marine Geology*, 3d ed., 1996.

Marine machinery

All machinery installed on waterborne craft, including engines, transmissions, shafting, propulsors, generators, motors, pumps, compressors, blowers, eductors, centrifuges, boilers and other heat exchangers, winches, cranes, steering gear, and associated piping, tanks, wiring, and controls, used for propulsion, for ship services, and for cargo, trade, or mission services.

Practically all marine machinery elements have nonmarine counterparts; in some cases, the latter were developed from marine applications, while in other cases specific equipment was "marinized." For marine service, machinery may have to meet higher standards of reliability and greater demands for weight and volume reduction and access for maintenance. Marine machinery must be capable of withstanding the marine environment, which tends toward extreme ambient conditions, high humidity, sea-water corrosion, vibration, sea motions, shock, variable demand, and fluctuating support services. Even higher standards may apply for warship machinery. To improve system reliability, essential equipment may be fitted in duplicate or provided with duplicated or alternative support or control systems, while nonessential equipment may be fitted with bypasses, to permit continued operation of a system following a component failure. Isolation valves or circuit breakers are common, enabling immediate repair.

Machinery on modern ships is highly automated, with propulsion usually directly controlled from the wheelhouse, and auxiliary machinery centrally controlled from an air-conditioned, sound-proofed control room, usually in the engine room. In the typical modern merchant ship (but not in passenger ships), the machinery operates automatically, and the controls are unattended at sea, with engineers called out by alarm in the event of malfunctions.

Propulsion Machinery

Propulsion machinery comprises an engine, usually a diesel engine, steam turbine, or gas turbine, with required gearing or other transmission system, and, for steam plants, steam generators. *See* BOAT PROPULSION; MARINE BOILER; MARINE ENGINE; PROPELLER (MARINE CRAFT); SHIP NUCLEAR PROPULSION.

Shafting. Shafting (**Fig. 1**) is of steel, usually solid for fixed-pitch propellers. In modern practice, the propeller is fitted to a matched taper on the propeller shaft by a hydraulic press fit, without a key, since a keyway would be a source of stress concentrations. Shafting sections are flanged and bolted together. For controllable-pitch propellers, the shaft is hollow to accommodate the hydraulic or mechanical linkage, and the propeller is bolted to an outboard flange on the propeller shaft. Because the propeller shaft must then be withdrawn outboard, the inboard connection to the remaining shafting is usually made by a removable sleeve coupling. In modern commercial practice, all bearings, including those in the stern tube, are oil-lubricated, with special seals at both ends of the stern tube; but in many applications, including warships, the stern-tube bearings have nonmetallic linings and are lubricated by water allowed to leak in around the shaft. *See* ANTIFRICTION BEARING; MACHINE KEY.

Shafting is subject to large cyclic variations in load induced by the propeller and, in diesel plants, by the engines. An important consideration in shafting design is to prevent resonant frequencies of the shafting from coinciding with these cyclic variations, within the operating range of shaft speeds.

Thrust bearing. The thrust developed by the propeller is transmitted through the shafting to the hull via a thrust bearing, comprising a collar on the shaft (Fig. 1) and a stationary housing mounted on a substantial foundation integral with the hull structure. For low-speed diesel engines, which are directly coupled to the shaft, the thrust bearing is built into the engine at its after end. In geared installations the bearing may be separate from or integral with the gear case, and forward or aft of the gear. Except in the lowest-powered installations, the bearings are of the pivoted-shoe, oil-lubricated type.

Auxiliary Machinery

In all but the smallest boats, the propulsion machinery is but part of an integrated power plant, which includes auxiliary machinery for propulsion-engine support and other ship services.

Electric generating plant. The electric power required for all continuous or recurring purposes on a ship is about 5% of the propulsion power, when restricted to the supply of motor-driven auxiliaries, lighting, other hotel services including high-voltage alternating current and provision refrigeration, communication, navigation, and control equipment of simple cargo ships. Much higher percentages are required when such services as cargo refrigeration or cargo hold ventilation are provided. On passenger ships the electric load can rival or exceed the propulsion power. Peak power demand may occur intermittently if motor-driven thrusters are fitted for maneuvering.

Ocean-going ships are required to have a minimum of two generators, with either one capable of supporting the essential electrical load, but most ships have at least three in order to improve ship reliability. A ship with a high electrical load may have three or more generators, with two or more operated in parallel to support the load. In addition, most ships are required to have a separate emergency source, usually a diesel generator, located outside the engine room, for emergency lighting, communication, navigation and control services, fire fighting, and often, steering.

Power from the generators is distributed to motors, lighting, and other consumers through switchboards and distribution boards.

On steamships, at least one of the generators, acting as the normal electricity source, is driven by an independent turbine supplied with steam from the same source as the propulsion turbines, and at least one will be a diesel generator. Most diesel-driven

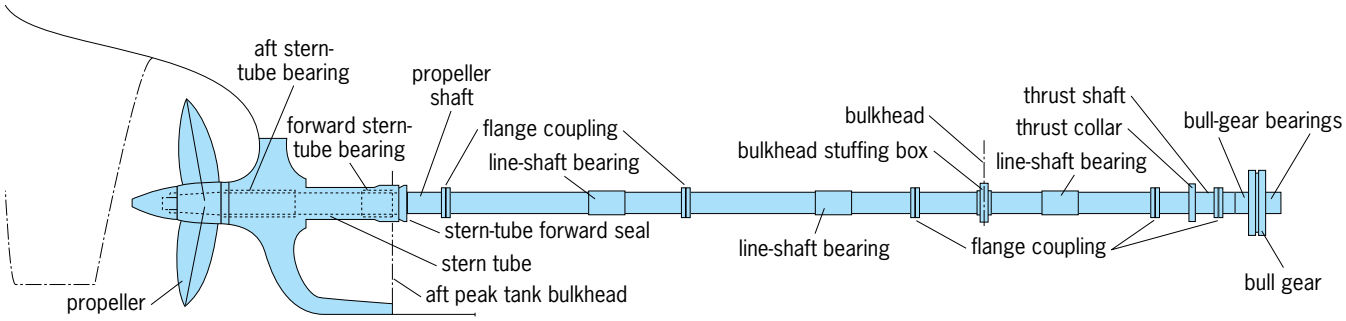

Fig. 1. Propeller shafting for a single-screw merchant ship. (*After R. L. Harrington, ed., Marine Engineering, Society of Naval Architects and Marine Engineers, 1992*)

ships have diesel generators, as do many gas turbine–driven ships, although others have gas turbogenerators. Many ships are fitted with a generator driven from the propulsion gearing or shafting as the normal electricity source. In many high-powered diesel-engine ships, sufficient waste heat can be recovered by an exhaust gas boiler to support a steam turbine, adequate for the normal load. Ships with alternating-current electric drive may take power for ship services from the propulsion generator. *See* DIESEL ENGINE; GAS TURBINE; GENERATOR; STEAM TURBINE.

Fuel systems. Generally, fuel is bunkered into tanks dispersed through the hull of a ship, using centrally controlled piping, and brought daily to settling tanks in or near the engine room by motor-driven transfer pumps, usually located in the engine room and using the same piping. Fuel systems are most complex on merchant ships with diesel engines, burning heavy fuel oil. On these ships, the fuel is heated in the tanks to facilitate pumping, then kept hot in the settlers to encourage the settling, by gravity, of water and solids. After further heating, settled fuel is passed through centrifugal separators to a day tank, from which it is supplied by motor-driven service pumps, through final heaters and fine filters, to the engine. The heating is done by steam in immersed coils in the tanks and in shell-and-tube or plate heat exchangers. In parallel with the heavy fuel system is an abbreviated diesel fuel system supplying the generator engines and package boiler if these are not run on heavy fuel, and supplying the propulsion engine when necessary. The diesel fuel is not heated but is pumped to a settling tank, a centrifugal separator, a day tank, and filters.

For ships using diesel fuel only, fuel heating is usually unnecessary and the system is simplified. Smaller plants are often not fitted with centrifugal separators, eliminating the need for a separate day tank. On steamships burning residual fuel, the fuel must be heated in the tanks, but is not centrifuged, so that the service pumps draw fuel directly from the settling tanks for supply to the burners via heaters and a fine strainer. *See* CENTRIFUGATION; DISPLACEMENT PUMP; HEAT EXCHANGER; PUMPING MACHINERY.

Steam and feedwater systems. On steamships, superheated steam passes to the propulsion turbine and to the turbogenerator, then exhausts to the condenser, a shell-and-tube heat exchanger, which is cooled by seawater inside the tubes. The condenser is maintained under a vacuum, usually 5 kilopascals (1.5 in. Hg), by air ejectors to extend the useful range of energy recovery from the steam. Condensate is drawn from the condenser by motor-driven condensate pumps and discharged through one or more shell-and-tube heaters, each heated regeneratively by steam extracted from the propulsion turbine, to a deaerating feed tank. *See* STEAM CONDENSER; VAPOR CONDENSER.

The deaerating feed tank (**Fig. 2**) uses extraction steam to heat the condensate, now called feedwater, to saturation temperatures, at which air and other noncondensable gases approach insolubility and can be vented from the system. This tank is also a surge

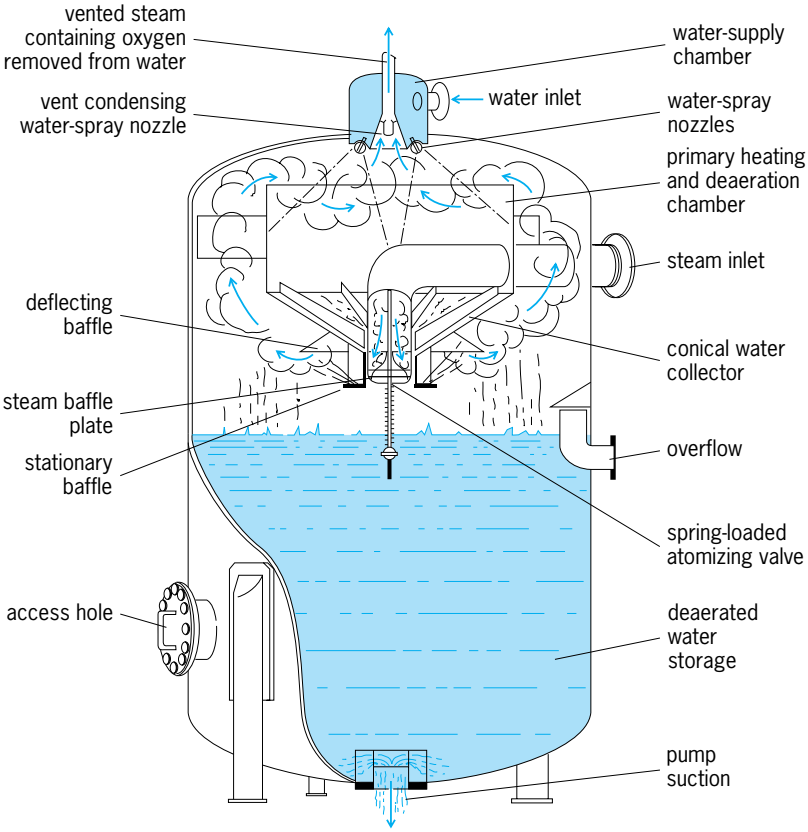

Fig. 2. Deaerating feed heater. (*After R. L. Harrington, ed., Marine Engineering, Society of Naval Architects and Marine Engineers, 1992*)

tank to accommodate differences in total volume of water and steam in the system as the ship is maneuvered. The deaerating feed tank distinguishes the marine steam plant from shoreside steam plants, which are not routinely subject to large load changes, where deaeration and limited surge can be achieved in a large condenser drainwell, not feasible aboard ship because of limited depth below the condenser. The deaerating feed tank is usually located high in the ship to provide a static head at the feedwater pumps, thereby preventing the hot feedwater from flashing to steam. On warships where this location is considered too vulnerable, the tank is lower, and feed booster pumps substitute for the static head. *See* BOILER FEEDWATER.

Feed pumps are usually multistage centrifugal pumps directly driven by compact steam turbines. For boilers fitted with economizers, the feedwater passes directly to the economizer, but when air heaters are fitted the feedwater passes through one or more shell-and-tube, high-pressure heaters, heated regeneratively by steam extracted from the propulsion turbine. *See* AIR HEATER; BOILER ECONOMIZER.

Often the turbogenerator has its own condenser and condensate pump. Sometimes the turbogenerator exhausts to the deaerating feed tank via an auxiliary exhaust line.

Steam for some auxiliary services, such as for fuel heating and the distilling plant, is supplied from turbine extractions as the most efficient source. Steam for the feed-pump turbine is taken from the superheated steam line, in some cases directly, but more frequently via a desuperheater. All exhaust steam and all drains are collected to appropriate points in the cycle to recover the heat and the water.

Steam systems on diesel-powered ships usually comprise a waste-heat boiler in the main engine uptake, in parallel with an oil-fired package boiler. Unless heat recovery is to be maximized, as when a waste-heat turbogenerator is fitted, the steam generated is saturated and is sufficient for fuel heating and other needs. Condensate is collected in a tank for return to the boiler via a motor-driven feed pump.

On tankers, steam may be used for cargo pumps and for heating of the cargo. On steamships the cargo-pump turbines are supplied from the propulsion boilers directly, but to avoid contamination the cargo-heating steam may be supplied indirectly from a steam-to-steam heat exchanger. On diesel-powered tankers, large auxiliary, oil-fired boilers may be used for these services. On ships with no heating requirements for fuel, a low-pressure package boiler may be fitted or an alternative heating plant may be provided.

Water-cooling systems. For steamships, very large quantities of seawater are required for the condensers. The seawater can be pumped through the condenser by large motor-driven centrifugal pumps, or induced to flow by pressure differences across the hull, which may have to be augmented or even created by a protruding lip at the outlet. Other cooling requirements, such as those for lubricating-oil coolers and refrigeration condensers, are also met with seawater, directly or via an intermediate heat exchanger that cools freshwater for circulation to the auxiliaries. *See* CENTRIFUGAL PUMP.

Except for the smallest engines, diesel-engine jackets are fresh-water-cooled in closed circuit, with the fresh water cooled in turn by seawater in a heat exchanger. Engine lubricating-oil coolers and aftercoolers and other auxiliaries may be circulated with seawater directly or via an intermediate freshwater loop. Cooling for gas turbines is largely internal, and cooling-water requirements are limited to lubricating-oil coolers for the engine and its reduction gear, and to the auxiliary machinery.

Seawater and fresh water are circulated by centrifugal pumps, usually motor-driven, but often driven from the engine they serve. Heat exchangers of both shell-and-tube type and plate type are used. The selection of materials for equipment that handles seawater is critical, tending toward bronze pumps with stainless-steel impellers, and nonferrous or plastic-lined piping and heat-exchanger waterboxes. Heat-exchanger tubes are usually copper-nickel, but titanium has been used in naval applications. Plate-type heat exchangers are usually fitted with stainless steel or titanium plates.

Lubricating-oil systems. The most complex systems are those on merchant ships burning heavy fuel in diesel engines. In general, large engines are operated with a dry crankcase, the oil draining to a sump tank below the engine, usually built into the double bottom. The oil is drawn from the sump by circulating pumps, which discharge through a cooler and a fine filter to bearings, pistons (if oil-cooled), camshaft, governor, and valve gear, from which the oil drains to the sump. Each engine has its own lubricating-oil system. Oil is drawn continuously to a centrifugal separator via a heater and then returned to the sump. Low-speed engines are fitted with a separate cylinder lubrication system, storing and supplying a high-viscosity oil of injection directly into the cylinders, where it lubricates the piston rings and cylinder liners before it burns. In trunk-piston engines, some of the circulating oil is used for this purpose.

Reduction gearing is usually fitted with an independent sump, circulating pumps, filter, and cooler. Steam-turbine bearings are supplied from the same system that serves the reduction gear, but gas turbines usually use a different grade of oil and require separate pumps, coolers, and filters. Turbines and connected gearing, which unlike diesels cannot be brought to an immediate stop following an oil failure, are fitted with a gravity tank or other means of maintaining oil flow in an emergency.

Stern tubes, line-shaft bearings, air compressors, auxiliary steam turbines, and some other auxiliaries may use the same grade of oil as the reduction gear on ships with turbines or medium- or high-speed diesel engines, or the same oil as the bearings of low-speed engines. However, diesel generators and refrigeration compressors each require a different grade of

oil. Storage tanks of adequate capacity for several voyages must be provided for each grade of oil, as well as settling tanks for used oils that can be renovated onboard by settling and centrifuging, and used-oil storage tanks and a transfer pump for oils that must be discharged ashore after use.

Lubricating-oil pumps are most often motor-driven rotary types, but centrifugal pumps are occasionally used. Some medium- and high-speed diesel engines are arranged to drive their own oil pumps, as are some reduction gears. Oil coolers may be shell-and-tube or plate type.

Compressed-air systems. Each of these systems comprises one or more motor-driven air compressors and one or more receivers. Most diesel propulsion and auxiliary engines and many gas turbines are started by using compressed air which, because of the volume required, is stored at medium or high pressures, typically 3 megapascals (435 lb/in.^{2}) or higher. Air for miscellaneous services, such as for operating pneumatic hoists and tools, for recharging accumulators, and for supplying pneumatic controls, is usually stored at about 700 kilopascals (100 lb/in.^{2}). The air for pneumatic controls is usually separately compressed and stored, and passed to services via a dryer. Compressors are usually motor-driven reciprocating units, but other types have been used for low-pressure service. *See* COMPRESSOR.

Distilling plants. Fresh water is required for making up losses in steam systems and freshwater cooling circuits, as well as for drinking, washing, and cooking. Most modern ships also use freshwater for toilet flushing. Other ships, such as chemical tankers, require fresh water for cleaning of cargo tanks. In short voyages and in journeys on lakes and rivers, some of these needs might be met by taking water from shore, but most ocean-going ships generate their own freshwater from seawater. Steamships generally use multistage, low-pressure evaporators, heated by steam extracted from the turbine. Diesel-powered ships generally have single-stage evaporators heated by engine cooling water. Gas turbine ships with diesel generators may use the diesel cooling water or steam produced in a waste-heat boiler or an oil-fired package boiler. Alternatives to low-pressure evaporators are vapor-compression plants and reverse-osmosis units. High-pressure evaporators are generally obsolete because of seawater scaling. *See* DISTILLATION.

Bilge and ballast systems. Bilge systems consist of pumps, sometimes assisted by eductors, to remove water and other accumulated liquids from the engine room and other compartments in the ship, other than oil tanks. Where bilge water may be contaminated by oil, it is first pumped to a holding tank and discharged overboard slowly through an oil-and-water separator. Ballast systems comprise the pumps and piping used to fill and empty ballast tanks. (Ballast tanks are filled to improve the trim or stability of a ship which is not carrying cargo.) To reduce pollution, past practices of filling empty fuel and cargo tanks with ballast are no longer allowed. Bilge and ballast pumps are usually motor-driven centrifugal pumps, although steam-reciprocating pumps are sometimes fitted for bilge service, while tanker ballast pumps are often steam-turbine-driven centrifugal units. In smaller plants the bilge pump may be driven via a clutch from the propulsion engine, and may double as a fire pump.

Fire-fighting and inert-gas systems. In addition to portable extinguishers, a ship will be fitted with at least two centrifugal seawater pumps and a fire main to supply hose connections and sprinklers throughout the ship. Fire pumps are usually motor-driven, but in smaller plants a fire pump may be driven via a clutch from the propulsion engine and may double as a bilge pump. Most ships are required to have an emergency fire pump in a separate compartment, driven from the emergency generator or by an independent diesel engine. A separate system is fitted for smothering fires in machinery spaces and dry cargo holds, most often using bottled carbon dioxide and piping for gas distribution. On tankers and other ships with flammable liquid cargoes, a foam system is installed, with tanks of foam concentrate to be injected into seawater from the fire pumps and distributed via nozzles and monitors.

On oil tankers, the space above the cargo in the tanks, or the whole tank when empty, is kept full of inert gas. Most often, the inert gas is boiler exhaust gas, passed through a seawater spray to remove sulfur and then blown into the tanks, but on tankers without boilers of adequate size an oil-burning inert-gas generator is fitted.

Refrigeration, heating, and ventilation. Refrigeration for provisions and air conditioning is generally provided by vapor-compression units, each consisting of a motor-driven compressor and a water-cooled condenser. The halogenated hydrocarbon refrigerants are being replaced by partially chlorinated refrigerants, which have lower ozone-depletion potentials, or by ammonia, previously considered obsolete. In systems suitable for most merchant ships, the refrigerant is expanded directly in evaporator coils in the provision store rooms or in the air-conditioning ducts; but for larger, more distributed systems, as on passenger ships, refrigerated-cargo ships, and fishing boats, an intermediate refrigerant, usually brine, is circulated through piping to the coils. *See* MARINE REFRIGERATION.

Much refrigerated cargo is containerized, with each container equipped with its own motor-driven vapor-compression unit. Ships intended to carry such cargo have increased electrical generating capacity, sockets at the container locations, and increased ventilation capacity for containers carried below deck. *See* MARINE CONTAINERS.

Heating for air conditioning is usually by steam or electric coils in the ducts, supplemented by local convectors. Ventilation systems on all ships are extensive, using motor-driven fans to force airflow through machinery, accommodation, and work spaces, and through cargo areas when appropriate, such as the enclosed vehicle decks of ferries and roll-on/roll-off ships. *See* AIR CONDITIONING; COMFORT HEATING; VENTILATION.

Sewage treatment plants. Overboard discharge of raw sewage is no longer permitted, and craft for which holding tanks are impractical must be fitted with treatment plants. Most installations are of the biological type. In order to reduce the volume of sewage to be treated, clean drains from wash basins and showers are passed directly overboard. Further reductions can be made by using vacuum-flush systems, which are common on ships with large complements or with passenger accommodation. *See* SEWAGE TREATMENT.

Incinerators. Garbage can no longer be dumped overboard, and most ships are therefore fitted with incinerators, most of which burn the garbage in a diesel oil-fired furnace. Most incinerators are equipped to burn used oil and sludge collected at the centrifugal fuel and lubricating-oil separators.

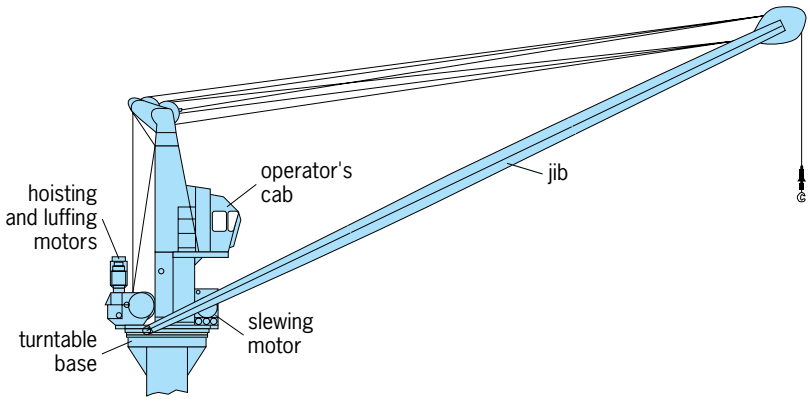

Fig. 3. General-purpose deck crane. (*After D. A. Taylor, Introduction to Marine Engineering, 2d ed., Butterworth, 1990*)

Deck Machinery

This machinery is more appropriately described as hull machinery, as it could encompass all machinery outside the engine room, including much of that cited above as auxiliary machinery. Much of the machinery described below is hydraulic. Such machinery is driven by electrohydraulic power packs, centered about a motor-driven hydraulic pump which supplies fluid under sufficient pressure to operate the rams or rotating hydraulic motors. Direct electric drive is also common, but steam-driven deck machinery is rarely found on modern ships.

Steering machinery. Except for those of the smallest craft, rudders are turned by hydraulic power, usually rams operating directly on the rudder stock. The rams are under command of a telemotor, which transmits the motion of the manually operated steering wheel or lever in the wheelhouse, or signals from the automatic pilot. In general, steering machinery incorporates redundant systems and emergency controls. *See* SHIP POWERING, MANEUVERING, AND SEAKEEPING.

Windlasses, capstans, and mooring winches. A windlass is a power-driven sprocket, permanently meshed to the anchor chain, used to raise the anchor; it is equipped with a brake to allow the anchor to be dropped in an orderly fashion. Capstans are vertical, power-driven drums around which wire or rope lines may be wrapped to assist in pulling and hoisting tasks. Mooring winches are power-driven horizontal spools, equipped with brakes, on which the wire or rope line for securing the craft to a dock is stored, and which are used to pay out the line and draw it taut as required. Some mooring winches are automatic, able to maintain a constant tension in the line. Most mooring winches are fitted with an auxiliary drum, called a warping or gypsy head, which is used like a horizontal capstan to assist in handling loose lines.

Derricks and cranes. These are fitted for cargo handling and utility services on many types of ships, and may range from a boom pivoted at the base of a mast, used in conjunction with winches, to swiveling jib cranes with self-contained hoisting, luffing, and slewing drives, operated from an enclosed cab (**Fig. 3**). Power may be direct-electric or electrohydraulic. *See* DERRICK; HOISTING MACHINES.

Boats, including lifeboats, are usually handled by davits, which are pairs of dedicated cranes that enable a loaded boat to be lowered safely to the surface of the sea. Specialized lifeboats, called free-fall lifeboats, are designed to be launched, fully loaded, from inclined ramps built onto a ship's superstructure.

Access and overhauling equipment. Apart from cargo handling, most ships are designed with carefully considered provisions to facilitate moving stores, garbage, and equipment on and off the ship and about the ship to locations such as workshops. These provisions include stores cranes, hatches or bolted plates, vertical trunks, side ports, monorails with trolleys, ramps for hand trucks or battery-powered forklift trucks, elevators, and strategically located pad eyes. Naval vessels are usually equipped with specialized equipment for underway replenishment of fuel, provisions, ammunition, and spares.

Cargo pumping systems. These systems are installed in tankers and other ships carrying liquid cargo. In a large oil tanker, centrifugal pumps are located at the bottom of a pump room, which is located forward of the engine room bulkhead and extends the full height of the hull. The pumps are usually driven by steam turbines, but in some tankers by electric motors located in the engine room, with the shafts penetrating the bulkhead. The pumps are used to discharge the cargo, via mains with branches from each cargo tank, to flanged connections amidships, at both sides of the main deck, where hoses from the shore can be attached. During discharge, some of the oil is recirculated to tank cleaning machines, which spray it in a pattern to wash residue to the bottom of the tanks. At all times the space above the oil is kept full of inert gas. Stripping eductors, driven by the oil discharged from the pumps, are used to remove the last of the oil from the tanks. The ship is loaded through the same piping, with the oil under pressure from pumps ashore. All valves in the system are remotely controlled from a console in a cargo control room.

Some tankers, generally carriers of refined products or chemicals, have no pump room, but use individual pumps submerged at the bottom of each tank. These pumps may be driven by motors on deck via long shafts, or by hydraulic motors submerged in the tanks with the pumps.

Stabilizers. Many types of ships are fitted with these units, including warships and passenger ships, intended to reduce the rolling motions of the ship by generating a moment to oppose the roll. Fin stabilizers are short, swiveling, streamlined blades which protrude from both sides of the hull, pivoted by hydraulic actuators controlled by a gyroscope. They may retract or fold back against the hull, or they may protrude at all times. Tank stabilizers are pairs of tanks, one on each side of the ship, which are kept half full of water and connected by a large pipe through which water is transferred. They may be passive, relying on gravity to transfer the water as the ship rolls, or active, using a gyroscope-controlled pump or compressed air.

Thrusters. All types of ships are fitted with these units, usually at the bow but often at the stern as well, which function as maneuvering aids by providing a transverse thrust on command. The most common arrangement uses a propeller mounted in a transverse tunnel through the hull, driven by bevel gearing from an electric motor or diesel engine. Other thruster types, which may be rotated, are used for special applications, including dynamically positioned ships. Alan L. Rowen

Bibliography. R. L. Harrington (ed.), *Marine Engineering*, 1992; E. C. Hunt (ed.), *Modern Marine Engineer's Manual*, vol. 2, 2d ed., 1990; A. B. Neild, Jr. (ed.), *Modern Marine Engineer's Manual*, vol. 1, 2d ed., 1965; D. W. Smith, *Marine Auxiliary Machinery,* 1983; D. A. Taylor, *Introduction to Marine Engineering*, 2d ed., 1990.

Marine microbiology

An independent discipline applying the principles and methods of general microbiology to research in marine biology and biogeochemistry. Dealing, by definition, with organisms of microscopic dimensions, marine microbiology focuses primarily on prokaryotic organisms, mainly bacteria.

Every cubic centimeter (0.06) of seawater contains about 1000–100,000 bacteria, and this to 1 million or more in nutrient-rich inshore waters. Because of their small size and easy dispersability, bacteria are virtually ubiquitous in the marine environment. Furthermore, natural populations of marine bacteria comprise a large variety of physiological types, can survive long periods of starvation, and are able to start their metabolic activity as soon as a substrate (food or growth-limiting nutrient) becomes available. As a result, the marine environment, similar to soil, possesses the potential of a large variety of microbial processes that degrade (heterotrophy) but also produce (autotrophy) organic matter. Considering the fact that the marine environment represents about 99% of the biosphere, marine microbial transformations are of tremendous global importance. If the deep sea is defined as waters below 1000 m (0.62 mi), this area still comprises 75% of the global biosphere, with temperatures of 2–4°C (34–39°F), total darkness, and high hydrostatic pressures. *See* BIOSPHERE.

Heterotrophic transformations. Quantitatively, the most important role of microorganisms in the marine environment is heterotrophic decomposition and remineralization of organic matter. It is estimated that about 95% of the photosynthetically produced organic matter is recycled in the upper 300–400 m (1000–1300 ft) of water, while the remaining 5%, largely particulate matter, is further decomposed during sedimentation. Only about 1% of the total organic matter produced in surface waters arrives at the deep-sea floor in particulate form. In other words, the major source of energy and carbon for all marine heterotrophic organisms is distributed over the huge volume of pelagic water mass with an average depth of about 3800 m (2.5 mi). In this highly dilute medium, particulate organic matter is partly replenished from dissolved organic carbon by microbial growth, the so-called microbial loop.

The rate of heterotrophic microbial decomposition of organic matter is primarily affected by three factors: nutrient concentration, temperature, and pressure. As a low-nutrient medium, seawater offers a growth advantage to bacterial populations that are characterized by high enzymatic affinities for their particular growth substrates. Morphological structures, such as increased surface area or stalks and organelles for the attachment of cells to solid surfaces, are other means of coping with the low-nutrient environment. Below depths of 300–400 m (1000–1300 ft) the temperature is uniformly in the range of 2–4°C (34–39°F), furthering growth, and hence creating a predominance of low-temperature-adapted or psychrophilic microorganisms. While most surface-born mesophilic bacteria grow optimally within a range of 20–35°C (68–95°F), marine psychrophilic bacteria have growth optima of 8–16°C (46–61°F), and they are commonly killed at temperatures above 20°C (68°F). *See* SEAWATER.

Hydrostatic pressure increases approximately 1 atm (100 kilopascals) for every 10 m (33 ft) of water depth. Most bacteria, marine and nonmarine, are barotolerant to various degrees and still grow at pressures of 300–600 atm (30–61 megapascals), corresponding to water depths of approximately 3000–6000 m (2–4 mi). Pressure-adapted (barophilic) microorganisms grow beyond these pressures and are characterized by having growth optima at pressures higher than 1 atm (100 kPa). Obligately barophilic bacteria are species that are unable to grow at 1 atm (100 kPa). The various types and degrees of pressure adaptation (**Fig. 1**) do not eliminate a general gradual decrease of microbial metabolic activity with increasing pressure or depth in the ocean. In consideration of microbial adaptations for a deep-sea existence, temperature- and pressure-retaining samplers have been developed (**Fig. 2**).

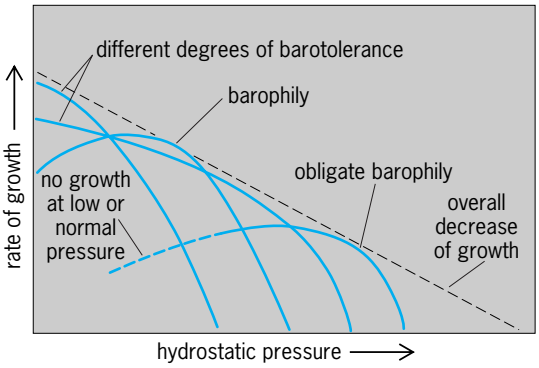

Fig. 1. Pressure-adapted growth in marine bacteria. The broken line represents a general decrease of microbial metabolic activity with increasing hydrostatic pressure (or water depth). (*After H. W. Jannasch, R. E. Marquis, and A. M. Zimmerman, eds., Current Perspectives in High Pressure Biology. Academic Press, 1987*)

Most microorganisms in a given sample of seawater lack their particular growth substrate or are not adapted to the prevailing environmental conditions and are therefore metabolically inactive. As a consequence, the mere number of microorganisms (as determined microscopically or by colony counts on agar plates) provides less information than direct testing of metabolic activity. The latter is done by measuring the rate of incorporation or turnover of a suitable radiolabeled substrate by the natural microbial population under natural conditions. The use of non-growth-stimulating labeled indicators, such as adenine or thymidine as precursors of nucleic acid production, is one of the attempts to assess the unaltered natural activity of whole microbial populations.

Fig. 2. Pressure- and temperature-retaining bacteriological water sampler during recovery from a depth of 6000 m (about 4 mi). Transfer into prepressurized reaction vessels permits the measurement of microbial activity under field conditions. (*Courtesy of S. Molyneaux*)

Of the large variety of organic material decomposed by marine heterotrophic bacteria, oil and related hydrocarbons are of special interest. Other environmentally detrimental pollutants (such as inorganic acids, organic solvents, pesticides, and herbicides) that are directly dumped or reach the ocean as the ultimate sink by land runoff are microbiologically degraded at varying rates. Techniques of molecular genetics are aimed at encoding genes of desirable enzymes into organisms for use as degraders of particular pollutants.

A specifically marine microbiological phenomenon is bacterial bioluminescence, which may function as a respiratory bypass of the electron transport chain. Free-living luminescent bacteria are distinguished from those that live in symbiotic fashion in light organelles of fishes or invertebrates. *See* BIOLUMINESCENCE.

Photoautotrophs and chemoautotrophs. The type of photosynthesis carried out by purple sulfur bacteria is an anaerobic process. It uses hydrogen sulfide (in contrast to water, which is used in green plant photosynthesis) as a source of electrons and thus produces sulfur, not oxygen. Photoautotrophic bacteria are therefore limited to environments where light and hydrogen sulfide occur simultaneously, mostly in lagoons and estuaries. In the presence of sufficient amounts of organic substrates, heterotrophic sulfate-reducing bacteria provide the necessary hydrogen sulfide where oxygen is depleted by decomposition processes. Anoxygenic photosynthesis is also carried out by some blue-green algae, which are now classified as cyanobacteria. A genus of oxygenically photosynthesizing cyanobacteria (*Synechococcus*) was discovered to be common in surface seawater. *See* CYANOBACTERIA; PHOTOSYNTHESIS.

Chemoautotrophic bacteria are able to reduce inorganic carbon to organic carbon (chemosynthesis) by using the chemical energy liberated during the oxidation of inorganic compounds (hydrogen, hydrogen sulfide, elemental sulfur, thiosulfate, ammonia, nitrite, and reduced iron and manganese). Their occurrence, therefore, is not light-limited but depends on the availability of oxygen and the suitable inorganic electron source. Their role as producers of organic carbon is insignificant in comparison with that of photosynthetic producers (exempting the processes found at deep-sea hydrothermal vents). The oxidation of ammonia and nitrite to nitrate (nitrification) furnishes the chemically stable and biologically most available form of inorganic nitrogen for photosynthesis. Genera of nitrifying bacteria include *Nitrosomonas, Nitrosococcus, Nitrosospira, Nitrobacter*, other *Nitrosovibrio, Nitrosolobus*, and *Nitrospina*. *See* NITROGEN CYCLE.

In contrast to soil and fresh water, the effect of acidification by the microbial formation of sulfate and nitrate is minimized in the marine environment by the high buffering capacity of seawater.

By the same token, highly acidophilic sulfur-oxidizing bacteria lose out in competition with facultatively chemoautotrophic bacteria, which produce nonacidifying polythionates as their end product of sulfur oxidation. The generation of methane and acetic acid from hydrogen and carbon dioxide stems from anaerobic bacterial chemosynthesis, and is common in anoxic marine sediments. *See* METHANOGENESIS (BACTERIA).

Marine microbial sulfur cycle. Sulfate is quantitatively the most prominent anion in seawater. Since it can be used by a number of heterotrophic bacteria as an electron acceptor in respiration following the depletion of dissolved oxygen, the resulting sulfate reduction and the further recycling of the reduced sulfur compounds make the marine environment microbiologically distinctly different from freshwater and most soils. Sulfate is far more important in seawater for microbial anaerobic respiration, than nitrate (denitrification, resulting in nitrogen gas production) or carbonate (methanogenesis, resulting in methane production). The assimilatory reduction of sulfate to organic sulfur compounds is quantitatively of minor importance in the marine environment compared to the bacterial respiration of sulfate. The resulting hydrogen sulfide is reoxidized to sulfate in the presence of oxygen with the concomitant production of organic carbon. This chemosynthetic production is considered secondary, since it depends on the primary photosynthetic production of organic carbon and oxygen.

The marine anaerobic, heterotrophic sulfate-reducing bacteria are classified in three genera; *Desulfovibrio, Desulfotomaculum*, and *Clostridium*. Microbial sulfate reduction is particularly important in the Black Sea, the world's largest anoxic basin with a maximum depth of 2000 m (6000 ft), where the oxygen-sulfide interface lies at a depth of about 100 m (300 ft). This microbially produced phenomenon has turned the Black Sea into one of the most studied marine environments. *See* BLACK SEA.

The marine aerobic sulfur-oxidizing bacteria fall into two groups: the thiobacilli (*Thiobacillus* and *Thiomicrospira*) and the morphologically conspicuous filamentous or unicellular organisms (*Beggiatoa, Thiothrix, Thiovulum, Thiospira, Thioploca*, and *Achromatium*). While the former comprise a wide range from obligately to facultatively chemoautotrophic species (requiring none or some organic compounds), few of the latter have been isolated in pure culture, and chemoautotrophy has been demonstrated in only a few. *Thioploca* and *Beggiatoa* have been observed to form massive microbial mats on the sea floor, the former off the coast of Chile in depths of 60–200 m (80–600 ft), the latter in the vicinity of sediment-covered hydrothermal vents at a depth of 2000 m (6000 ft).

Deep-sea hydrothermal vent bacteria. Two types of hydrothermal vents have been investigated: warm vents (8–25°C or 46–77°F) with flow rates of 1–2 cm (0.4–0.8 in.)/s, and hot vents (260–360°C or 500–600°F) with flow rates of 2 m (6.5 ft)/s. In their immediate vicinity, dense communities of benthic invertebrates are found with a biomass that is orders of magnitude higher than that normally found at these depths and dependent on photosynthetic food sources. This phenomenon has been explained by the bacterial primary production of organic carbon through the chemosynthetic oxidation of reduced inorganic compounds. The chemical energy required for this process is analogous to the light energy used in photosynthesis and is provided by the geothermal reduction of inorganic chemical species. The specific compounds contained in the emitted vent waters and suitable for bacterial chemosynthesis are mainly hydrogen sulfide, hydrogen, methane, and reduced iron and manganese.

There appear to be three major sites where microbial growth takes place: within the subsurface vent system, in microbial mats covering surfaces exposed to the vent plume, and in various symbiotic associations. In the first case, bacterial cells and visible cell aggregations are emitted with the vent waters and become available to filter-feeding sessile and planktonic animals; *Thiomicrospira* is the most commonly isolated genus. In the second case, grazing animals can be observed on various types of surfaces feeding on *Beggiatoa* and *Thiothrix*-like filamentous bacteria. Isolated from mat material are manganese-oxidizing, extremely thermophilic methanogenic as well as methylotrophic bacteria (**Fig. 3**). The third case leads to the production of the major portion of animal biomass: prokaryotic chemoautotrophic symbiosis is found in the gill tissue of the vent clam (*Calyptogena magnifica*) and mussel (*Bathymodiolus*), as well as in a special tissue, the trophosome, of

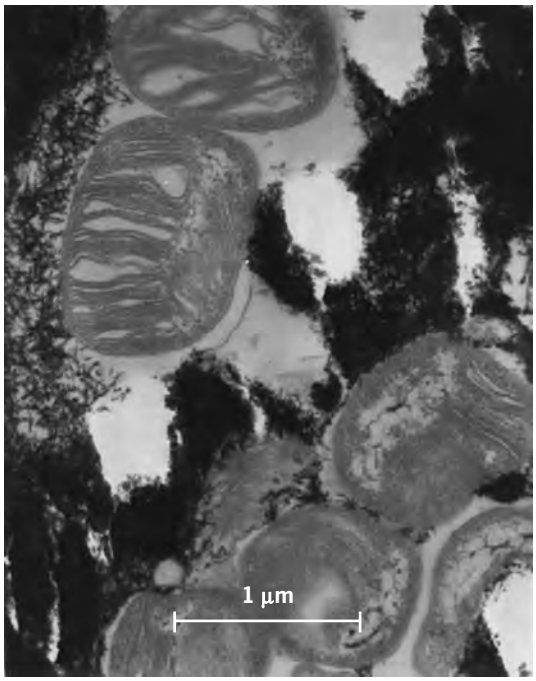

Fig. 3. Transmission electron micrograph of methylotrophic bacteria within iron-manganese deposits of microbial mats from deep-sea hydrothermal vents. (*From P. A. Rona et al., eds., Hydrothermal Processes at Seafloor Spreading Centers, Plenum Press, 1984*)

the mouthless and gutless pogonophoran tubeworm (*Riftia pachyptila*). Similar symbiotic systems have subsequently been found in estuarine environments, where hydrogen sulfide occurs as a product of bacterial sulfate reduction.

The extremely thermophilic microorganisms isolated from hydrothermal vents belong, with the exception of the genus *Thermotoga,* to the Archaebacteria. Of eight archaeal genera, growing within a temperature range of about 75–110°C (165–230°F), three are able to grow beyond the boiling point of water, if the necessary pressure is applied to prevent boiling. These organisms are strictly anaerobic. However, unlike mesophilic bacteria, hyperthermophilic marine isolates tolerate oxygen when cooled below their minimum growth temperature. *See* ARCHAEA; HYDROTHERMAL VENT.

Virology. Observations have shown that seawater may contain 10^5–10^8 virus or phage particles per cubic centimeter (0.62 in.^{3}). While one hypothesis purports that up to 60% of bacterial mortality may by due to viral infection, it is generally assumed that bacterial communities are dominated by cells that are resistant to their co-occurring phages, the latter maintaining themselves by infecting relatively rare sensitive cells. *See* BACTERIOPHAGE.

Biotechnology. Marine microbial biotechnology is an applied science. As an example, thermostable deoxyribonucleic acid–cleaving polymerases from the above-mentioned hyperthermophilic vent bacteria are produced commercially and used extensively in molecular biology, particularly in the polymerase chain reaction. Another example is the use of thermostable enzymes in high-temperature processes, such as the use of amylases from hyperthermophilic archaea in the conversion of cornstarch to glucose and fructose in the beverage industry. *See* BACTERIA; BIOTECHNOLOGY; MICROBIOLOGY.

Holger W. Jannasch

Bibliography. R. M. Atlas and R. Bartha, *Microbial Ecology*, 4th ed., 1997; A. T. D. Brock (ed.), *Thermophiles: General, Molecular and Applied Microbiology*, 1986; R. A. Herbert and G. A. Codd, *Microbes in Extreme Environments*, 1986; G. Rheinheimer, *Aquatic Microbiology*, 4th ed., 1993; H. G. Schlegel and B. Bowien, *Autotrophic Bacteria*, 1989.

Marine mining

The process of recovering mineral wealth from seawater and from deposits on and under the seabed. While the world's demand for mineral commodities is increasing at a rapid rate, most of the developed countries have been thoroughly explored for surface outcroppings of mineral deposits. The mining industry has needed to improve its capabilities for exploring and exploiting low-grade and unconventional sources of ore. Corresponding advances in oceanography have highlighted the importance of the ocean as a source of minerals and indicated that the technology required for their exploitation is in some cases already available.

Manganese nodules. Undersea mining became an important, but unsuccessful, diversification for many major oil and aerospace companies in the late 1960s. Over 80 separate exploration activities were reported in shallow coastal areas worldwide, but few resulted in new operating mines. Attention turned to the deep seabeds, and much of the debate and controversy in the deliberations of the United Nations Conference on the Law of the Sea (UNCLOS) revolved around the exploitation of manganese nodules. Conflicts over the rights of management of these minerals in international waters led to attempted moratoria on deep-seabed mining activities. Anticipating agreement, several countries, including the United States and Germany, prepared interim legislation to permit nationals of the respective countries and reciprocating states to mine the deep seabed for nodules containing manganese, copper, nickel, and cobalt. Several major multinational consortia were active in testing deep-seabed mining systems, including companies from the United States, Canada, United Kingdom, Australia, Germany, France, Belgium, Netherlands, and Japan. *See* MANGANESE NODULES.

Metalliferous sulfides. In 1979, an extraordinary discovery of massive deposits of metalliferous sulfides containing zinc, copper, silver, and gold was made at 21°N on the East Pacific Rise at a depth of 3000 m (10,000 ft). This confirmed projections, based on previous discoveries of metalliferous muds in the Red Sea, that the boundaries between spreading oceanic plates would be host to similar hydrothermal mineral deposits. As oceanic plates move apart, molten basaltic magma is thrust up between them, resulting in extensive fracturing of the adjacent flanks as much as several kilometers from the axis. Cold seawater percolates through the fractures, deep into the oceanic crust, and in its then superheated state dissolves many of the metals contained in the basalt. Normal convection causes the now hot, metal-rich solution to rise back to the seabed along the axis. The temperature and pressure decrease as the solution rises, and the contained minerals may be differentially precipitated as they pass through the cooler rocks. The fluids rise from the seabed at 300–350°C (570–660°F) into cold seawater at 4°C (37°F) to form black or gray smokers (hydrothermal vents), from which the remaining minerals are precipitated to form vent chimneys or to rain down as fine particles of mineral-rich sediments. Intensive examination of parts of the Juan de Fuca/Gorda Ridge spreading centers off the northwestern United States resulted in further discoveries, including gold-rich sulfide deposits off Papua, New Guinea. *See* HYDROTHERMAL VENT; ORE AND MINERAL DEPOSITS; PLATE TECTONICS.

With commercial interest in the manganese nodules waning, because of depressed metal markets and continuing uncertainties over the terms of the UNCLOS, and claims of 370-km (200-nautical-mile) exclusive economic zones (EEZ) proliferating, emphasis changed during the 1980s to assessing the potential of minerals within the EEZs. Prime targets

were the high-cobalt manganese crusts, containing up to 2% cobalt, which were found to be widely distributed throughout the Pacific basin on the flanks of islands and seamounts at depths between 800 and 2400 m (2600 and 7900 ft). Others were the metalliferous sulfide deposits formed at spreading centers and in back-arc basins within the economic zones, as well as the more traditional nearshore deposits of unconsolidated heavy minerals containing gold, tin, and titanium, and extensive bedded deposits of phosphorites projected in the U.S. Atlantic margin (**Fig. 1**). Meantime in certain countries, government interests in the manganese nodules was maintained, mainly in connection with environmental issues. In 1988, France, India, Japan, and the Soviet Union registered as pioneer investors under the terms of the United Nations convention. Since then, China and Korea have been added.

Hydrates. More recently, there has been widespread interest in the research, identification, assessment, exploration, and development of methane hydrate resources. In the United States, methane hydrates had been regarded as an academic curiosity, or as an expensive nuisance causing the clogging of oil and gas pipelines. The frozen, crystalline gas hydrates, also known as clathrates, have been demonstrated to be more widely distributed in the marine environment than was originally anticipated, and that this ubiquitous material represents a potential source of energy, on a global scale, that is significantly greater than all other forms of fossil fuel energy combined. The formation and stability of the crystals are dependent on the combination of high pressure, low temperature, and a source of gas. Fossil hydrates formed during the ice age are found onshore and throughout the continental shelf in high

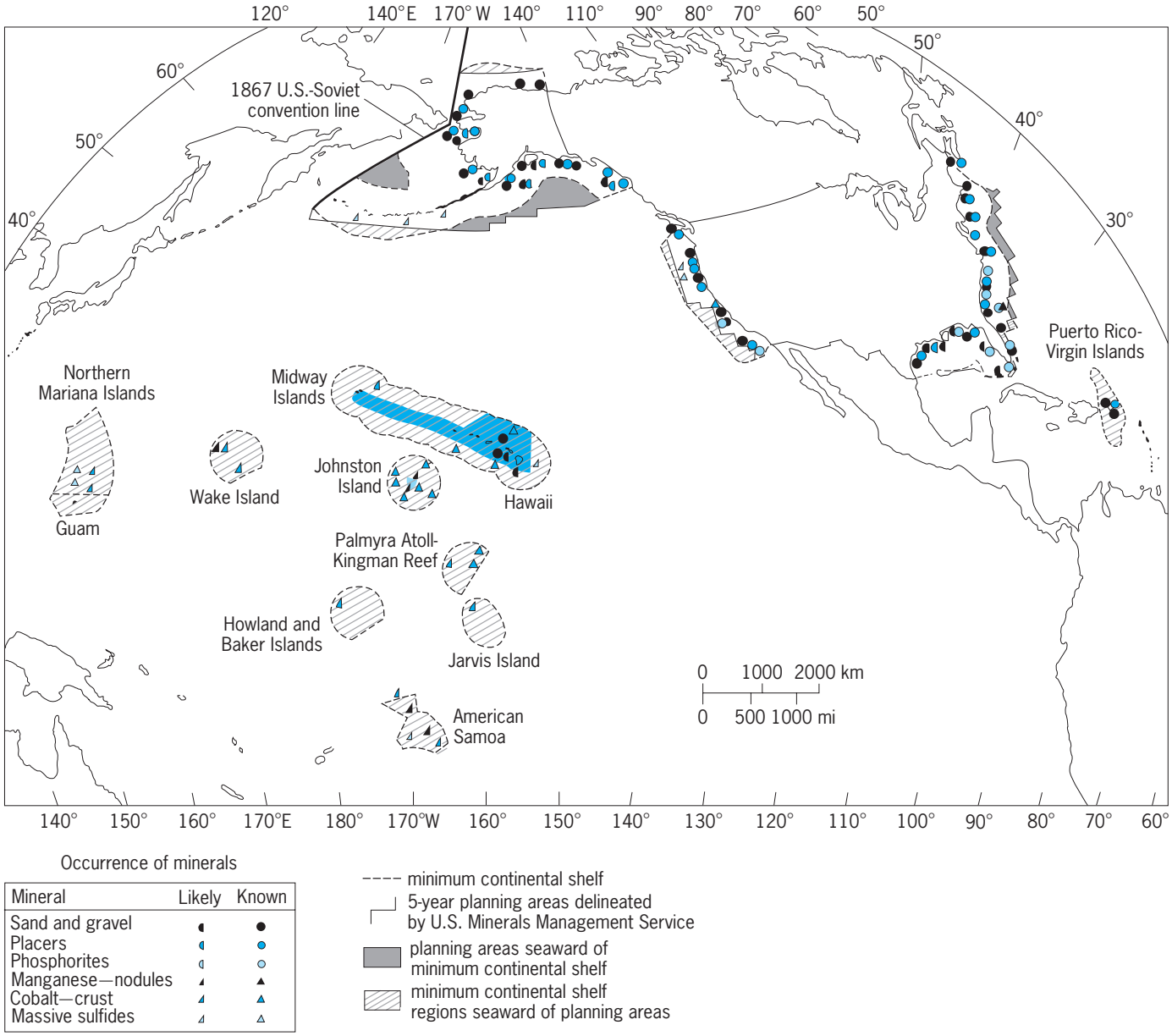

Fig. 1. Offshore mineral areas of the United States.

latitudes. The deep seabeds are host to hydrate deposits that are forming now. Methods of extraction of these deposits, which may be up to hundreds of meters thick over areas of tens of square kilometers, are being studied in India and Japan, where they are considered serious contenders as national sources of energy. One proposed method of recovery would be to drill underneath the deposit and remove the gas below the hydrates. The subsequent reduction in gas pressure would lead to sublimation of the crystalline hydrate cap, resulting in a continuous release of additional gas. The hydrate deposits may also constitute a hazard to oil and gas operations in terms of sea-floor instability, and it is suggested that they may be a significant factor in the presence of atmospheric greenhouse gases as the result of temperature changes in the oceanic environment, causing major gaseous releases from the seabed and influencing global climate. *See* HYDRATE.

Environment and minerals. While mineral resources to the value of trillions of dollars do exist in and under the oceans, their exploitation will not be simple. Many environmental problems must be overcome and many technical advances must be made before the majority of these deposits can be mined in competition with existing land resources. The International Seabed Authority (ISA) has prepared environmental guidelines for the development of manganese nodules and the International Marine Minerals Society (IMMS) has compiled similar guidelines for the development of metalliferous sulfides.

The marine environment may be logically divided into four significant areas: waters, deep seabeds, continental shelf and slope, and seacoast. Of these, the waters are the most significant, both for their mineral content and for their unique properties as a mineral overburden. Not only do they cover the ocean floor with a fluid medium quite different from the solid earth or atmosphere and require entirely different concepts of ground survey and exploration, the constant and often violent movement of the surface waters combined with unusual water depths presents formidable deterrents to the use of conventional seagoing techniques in marine mining operations.

The mineral resources of the marine environment are of three types: dissolved minerals of the ocean waters; unconsolidated mineral deposits of marine beaches, continental shelf, and deep seabed; and consolidated deposits contained within the bedrock underlying the seas. These are described in the classification outline below, which shows also subclasses of seabed and subseabed deposits, characteristics that have a very great influence on the economics of exploration and mining.

I. FLUID
- *Seabed*
  - Freshwater
  - Seawater
    - Magnesium
    - Sodium
    - Uranium
    - Bromine
    - Salts of 26 other elements
- *Subseabed*
  - Hydrothermal fluids

II. UNCONSOLIDATED
- *Seabed*
  - Industrial materials
    - Sand and gravel
    - Shells
    - Aragonite
  - Heavy-mineral placers
    - Magnetite
    - Ilmenite, rutile
    - Chromite, monazite
  - Nodules
    - Manganese
    - Phosphorite
  - Muds and oozes
    - Metalliferous
    - Carbonaceous
    - Siliceous
    - Calcareous
    - Barite
- *Subseabed*
  - Heavy-mineral placers
    - Gold
    - Platinum
    - Cassiterite
    - Gemstones
  - Bedded deposits
    - Phosphorites

III. CONSOLIDATED
- *Seabed*
  - Crusts
    - Phosphorite
    - Cobalt
    - Manganese
  - Mounds and stacks
    - Metallic sulfides
- *Subseabed*
  - Disseminated, stratified, vein, or massive deposits
    - Coal
    - Sulfur
    - Phosphates
    - Carbonates
    - Potash
    - Ironstone
    - Limestone
    - Metallic sulfides
    - Metallic salts
    - Gas hydrates

See MARINE GEOLOGY; SEAWATER.

As with land deposits, the initial stages preceding the production of a marketable commodity include discovery, characterization of the deposit to assess its value and exploitability, and mining, including beneficiation of the material to a salable product.

Exploration. Initial requirements of an exploration program on the continental shelves are a thorough study of the known geology of the shelves and adjacent coastal areas and the extrapolation of known

metallogenic provinces into the offshore areas. The projection of these provinces, characterized by relatively abundant mineralization, generally of one predominant type, has been practiced with some success in the localization of certain mineral commodities, overlain by thick sediments. As a first step, the application of this technique to the continental shelf, overlain by water, is of considerable guidance in localizing more intensive operations. Areas thus delineated are considered to be potentially mineral bearing and are subject to prospecting by geophysical and other methods.

A study of the oceanographic environment may indicate areas favorable to the deposition of authigenic deposits in deep and shallow water. Some deposits may be discovered by chance in the process of other marine activities.

Field exploration prior to or following discovery will involve three major categories of work: ship operation, survey, and sampling.

Ship operation. Conventional seagoing vessels are used for exploration activities with equipment mounted on board to suit the particular type of operation. The use of remotely operated vehicles (ROVs) will eventually augment or replace existing techniques, but they are not yet advanced sufficiently for commercial exploration. *See* UNDERWATER VEHICLE.

One of the most important factors in the location of undersea minerals is accurate navigation. Ore bodies must be relocated after being found and must be accurately delineated and defined. The accuracy of survey required depends upon the phase of operation. Initially, errors of 300 m (1000 ft) or more may be tolerated.

However, once an ore body is believed to exist in a given area, maximum errors of less than 30 m (100 ft) are desirable. These maximum tolerated errors may be further reduced to centimeters in detailed ore-body delineation and extraction.

A variety of types of electronic navigation systems is available, with accuracies from 900 m (3000 ft) down to approximately 1 m (3 ft). Loran, Lorac, and Decca are permanently installed in various locations throughout the world but are becoming obsolete as the differential Global Positioning System (GPS) has completely revolutionized the capabilities for positioning with high accuracy in any part of the world's oceans. *See* SATELLITE NAVIGATION SYSTEMS.

During sampling and mining operations, the vessel must be held steady over a selected spot on the ocean floor. Two procedures that have been fairly well developed for this purpose are multiple anchoring and dynamic positioning.

A three-point anchoring system is of value for a coring vessel working close to the shore. A series of cores may be obtained along the line of operations by winching in the forward anchors and releasing the stern anchor. Good positive control over the vessel can be obtained with this system, and if conditions warrant, a four-point anchoring system may be used. Increased holding power can be obtained by multiple anchoring at each point.

Dynamic positioning is useful in deeper water, where anchoring may not be practical. The ship is kept in position by use of auxiliary outboard propeller drive units or transverse thrusters. These can be placed both fore and aft to provide excellent maneuverability. Sonar transponders are placed on the bottom or held submerged at a depth of minimum disturbance, or the system may be tied to shore stations. The auxiliary power units are then controlled manually or by computer to keep the ranges at a constant value. *See* MARINE NAVIGATION; SONAR; UNDERWATER NAVIGATION.

Survey. The primary aids to exploration for mineral deposits at sea are depth recorders, side scan sonars, subbottom profilers, magnetometers, and bottom and subbottom sampling systems. Their use is dependent upon the characteristics of the ore being sought.

For the initial topographic survey of the seabed, and as an aid to navigation in inshore waters, the depth recorder is indispensable. It is usually carried as standard ship equipment, but precision recorders having a high accuracy are most useful in survey work.

Integrated equipment packages to record bathymetric features using side-scan sonar similar to radar imagery, and precision contouring using multiple, narrow-beam depth recorders have been among the most significant advances in seabed mapping. By using sophisticated image enhancement, a direct fallout from the space program, these systems are able to map in real time at rates up to 25,000 km^2 (9700 mi^2) a day, depending on the precision required. Combined with the GPS, these developments have brought seabed mapping much closer to the capabilities attained in terrestrial mapping from the air, at least where the water is deep enough to allow relatively wide swaths to be recorded.

In the search for marine placer deposits of heavy minerals, the subbottom profiler is probably the most useful of all the exploration aids. It is one of several systems utilizing the reflective characteristics of acoustic or shock waves, and is a development of standard geophysical seismic systems for reflection surveys, used in the oil industry. Profilers for shallow penetration use a variety of energy sources including electric sparks, compressed air, gas explosions, acoustic transducers, and electromechanical (boomer) transducers. The return signals as recorded show a recognizable section of the subbottom. Shallow layers of sediment, configurations in the bedrock, faults, and other features are clearly displayed and require little interpretation. The maximum theoretical penetration is dependent on the time interval between pulses and the wave velocity in the subbottom. A pulse interval of 0.5 s and an average velocity of 2400 m/s (8000 ft/s) will allow a penetration of 600 m (2000 ft), the reflected wave being recorded before the next transmitted pulse. *See* SEISMIC STRATIGRAPHY.

Penetration and resolution are widely variable features on most models of wave velocity profiling systems. In general, high frequencies give high resolution with low penetration, while low frequencies

give low resolution with high penetration. The general range of frequencies for mineral exploration varies from 1 to 10 kHz (typically 3.5 kHz), and the general range of pulse energy is 100–25,000 joules for nonexplosive energy sources. The choice of system will depend very much on the requirements of the survey, but for the location of shallow placer deposits on the continental shelf the smaller low-powered models have been used with considerable success.

With the advent of the flux gate, proton precession, and the rubidium vapor magnetometer, all measuring the Earth's total magnetic field to a high degree of accuracy, this technique has become much more useful in the field of mineral exploration. Anomalies indicative of mineralization such as magnetic bodies, concentrations of magnetic sands, and certain structural features can be detected. Although all three types are adaptable to undersea survey work, the precession magnetometer is more sensitive and more easily handled than the flux gate, and the rubidium vapor type has an extreme degree of sensitivity which enhances its usefulness when used as a gradiometer on the sea surface or submerged. *See* GEOPHYSICAL EXPLORATION.

Once an ore body is indicated by geological, geophysical, or other means, the next step is to sample it in area and in depth. Mineral reserves cannot be confirmed without collecting samples.

Sampling. Mineral deposit sampling involves two stages. First, exploratory or qualitative sampling to locate mineral values and allow preliminary judgment to be made. For marine deposits, this will involve such simple devices as snappers, drop corers, drag dredges, and divers. Accuracy of positioning is not critical at this stage, but of course is dependent on the type of deposit being sampled. Second, the deposit must be characterized in sufficient detail to determine the production technology requirements and to estimate the profitability of its exploitation. This quantitative sampling requires much more sophisticated equipment than does the qualitative type, and for marine work few systems in existence can be considered reliable and accurate. However, in particular cases, systems can be put together using available hardware that give the accuracy required. Specifically, qualitative sampling of any mineral deposit offshore can be carried out with existing equipment. Quantitative sampling of most alluvial deposits of heavy minerals (specific gravity less than 8) can be carried out at shallow depths (less than 107 m or 350 ft overall) using existing equipment, but cannot be carried out with reliability for higher-specific-gravity minerals such as gold (specific gravity 19). Advances have been made in sampling for diamonds.

Any system that will give quantitative samples can be used for qualitative sampling, but in many cases heavy expenses could be avoided by using the simpler equipment.

To obviate the effects of the sea surface environment, the trend is toward the development of fully submerged systems, but the deficiencies in sampling heavy placer minerals are not due to the marine environment. Even on land, the accuracy of placer deposit evaluation is not high and the controlling factors not well understood. There is still a need for more intensive research in this area.

Quantitative sampling of consolidated deposits offshore has been confined to water depths of no more than a few hundred feet mainly because of the expense. Whereas the Ocean Drilling Program has the capability to drill in thousands of meters of water, only one system presently exists that can be placed on the bottom from a conventional vessel, to drill small holes in the rock to depths of about 50 m (165 ft), which is a prerequisite for even the most simple deposit characterization of metalliferous sulfides. Similarly, the characterization of metalliferous crust deposits depends on multiple, accurately positioned samples of the crust material and the underlying bedrock. Until such systems are better developed, the true nature of the many discoveries of deep-seabed consolidated deposits will remain speculative.

Evaluation of surficial nodule deposits on the deep seabed at depths of 5000–6000 m (16,500–19,800 ft) may be carried out by using combinations of optical or acoustic imagery and sampling. Except for box corers that are used for geotechnical sampling, devices lowered from a stationary vessel have given way to the use of free-fall (boomerang) samplers equipped with corers, grabs, or cameras, and discharged at preselected points in groups of 8–12 for later retrieval. Also widely used are towed-television or multiple-shot photographic cameras that give a fully or nearly continuous coverage along selected tracks. Nodules are commonly assayed aboard ship and the data from all systems analyzed with onboard computers. A variety of seabed maps showing bathymetry, seabed topography, nodule distribution, and grade may be produced while the vessel is still on site. Ultimately, methods for continuous in-place assaying while under way should be perfected, which will allow the requisite data maps to be produced in real time.

Exploitation. Despite the intense interest in undersea mining, new activities have been limited mostly to conceptual studies and exploration. The volume of production has shown little change, and publicity has tended to overemphasize some of the smaller, if more newsworthy, operations. All production to date has come from nearshore sources, namely, seawater, beach and nearshore gold and diamond placers, and nearshore-consolidated deposits of coal, sulfur, and potash.

Minerals dissolved in seawater. Commercial separation techniques for the recovery of minerals dissolved in seawater are limited to chemical precipitation and filtration for magnesium and bromine salts and solar evaporation for common salts and freshwater production on a limited scale. Other processes developed in the laboratory on pilot plant scale include electrolysis, electrodialysis, adsorption, ion exchange, chelation, oxidation, chlorination, and solvent extraction. The intensive interest in the extraction of fresh water from the sea has permitted

TABLE 1. Mining regions for dissolved mineral deposits offshore

Mineral	Location
Sodium, NaCl	Worldwide
Magnesium, metal Mg, MgO, $Mg(OH)_2$	United States, United Kingdom, Germany, Russia
Bromine, Br	Worldwide
Freshwater	Middle East, Atlantic region, United States
Heavy water	Canada

additional research on the recovery of minerals, but successful commercial operations will require continued development of the combination of processes involved for each specific mineral.

Three minerals or mineral suites are extracted commercially from seawater: sodium, magnesium, and bromine (**Table 1**). Of these, salt evaporites are the most important. Japan's total production of salt products comes from the sea. Magnesium extracted from seawater accounts for 75% of domestic production in the United States. Freshwater compares with bromine in total production value, and heavy water, or deuterium oxide, has been produced from seawater in Canada for many years. Extensive work has been carried out in Japan on the extraction of uranium.

Unconsolidated deposits. Unconsolidated deposits include all the placer minerals, on the seabed and in the subseabed, as well as the authigenic deposits of manganese nodules found at moderate to great depths. *See* AUTHIGENIC MINERALS.

The mining of unconsolidated deposits became widely publicized with the awareness of the potential of manganese nodules as a source of manganese, copper, nickel, and cobalt, and because of developments in the exploitation of offshore diamonds in South-West Africa in the late 1960s. Unconsolidated deposits for many years presented a major source of tin offshore in South East Asia.

So far, the methods of recovery (which have been used or proposed) have been conventional; that is, dredging using draglines, clamshells, bucket dredges, hydraulic dredges, or airlifts. All these methods (**Fig. 2**) have been used in mining to maximum depths of 60 m (200 ft), and hydraulic dredges for digging to 90 m (300 ft) have been built. Extension to depths much greater than this does not appear to present any insurmountable technical difficulties, and in Japan, reserves of sand and gravel are being mapped for exploitation to depths of 300 m (1000 ft). In the deep ocean three industrial consortia successfully tested dredge systems for manganese nodules in water depths over 4500 m (15000 ft).

Dredging operations have exploited such diverse products as diamonds, gold, heavy mineral sands, iron sands, tin sands, lime sand, and sand and gravel. The most important of all of these commodities is the least exotic; 60% of world production from marine unconsolidated deposits is involved in dredging and mining operations for sand and gravel. Other major contributors to world production are the operations for heavy mineral sands (ilmenite, rutile, and zircon), mostly in Australia, and the tin operations in Thailand and Indonesia, which account for more than 10% of the world's tin.

Economics of these operations are dictated by many conditions. The spectacular range of costs offshore results from the effects of different environmental conditions. In general, offshore operations are more costly than similar operations onshore, but the factoring of secondary environmental costs may change this.

Offshore diamonds represent a $1.5 billion per year industry along the coasts of Namibia and South Africa. Most of the marine diamonds are high-quality gemstones, having traveled thousands of kilometers from their source inland and then been subjected to intensive wave action over several cycles of sea-level change.

Much of the world's heavy mineral sand production is from beach sand operations in Australia, Sri Lanka, and India. Only two oceangoing dredges are used. Most of the others are pontoon-mounted hydraulic dredges, or draglines, with separate washing plants.

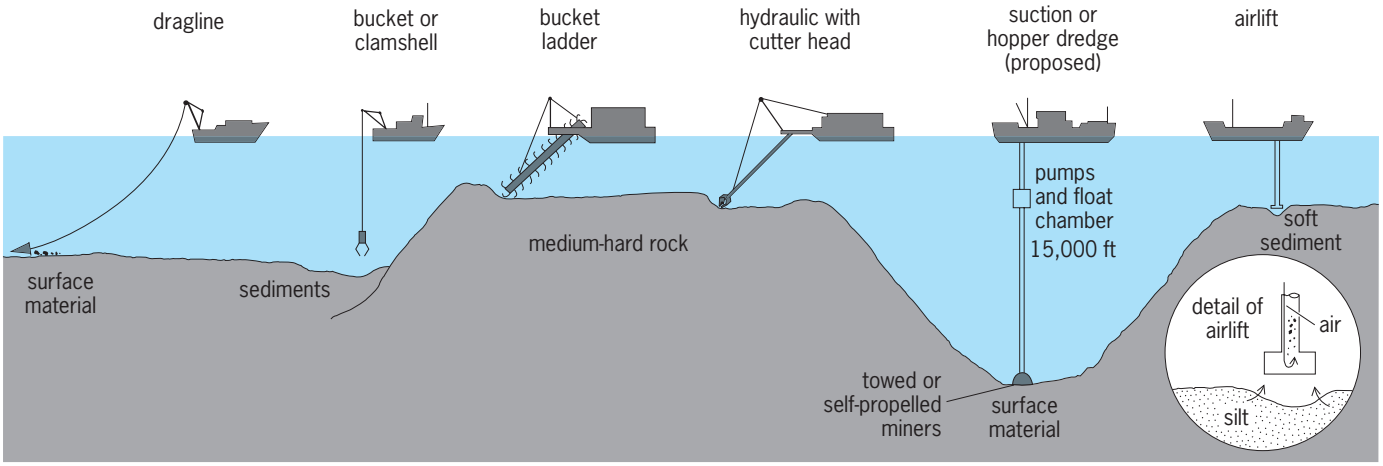

Fig. 2. Methods of dredging used in the exploitation of unconsolidated mineral deposits offshore.

Lime shells are mined as a raw material for portland cement. Two U.S. operations for oyster shells in San Francisco Bay and Louisiana employed barge-mounted hydraulic cutter dredges of 41- and 46-cm (16- and 18-in.) diameter in 9–15 m (30–50 ft) of water. The Iceland Government Cement Works in Akranes uses a 46-m (150-ft) ship to dredge seashells from 40 m (130 ft) of water, with a 61-cm (24-in.) hydraulic drag dredge.

In the United Kingdom, hopper dredges are used for mining undersea reserves of sand and gravel. Drag suction dredges up to 97 cm (38 in.) in diameter are most commonly used with the seagoing hopper hulls. Similar deposits have been mined in the United States, and the same type of dredge is employed for the removal of sand for harbor construction or for beach replenishment. Some sand operations use beach-mounted draglines for removal of material from the surf zone or beyond.

For deep-seabed mining at depths below 5000 m (3 mi), systems have been tested using airlift or suction lift and a mechanical system using a continuous loop of synthetic line from which buckets are suspended every 100 m (330 ft) or so. The gathering devices tested on the bottom with hydraulic systems include towed and self-propelled miners (**Fig. 3**), the latter incorporating a means to crush and slurry the nodules before pumping. Tests to identify the environmental effects of mining have not been conclusive, but tend to indicate that adverse effects are minor and can be mitigated.

Consolidated deposits. The third and last area of offshore mineral resources has an equally long history. The production from in-place mineral deposits under the sea is quite substantial, particularly in coal deposits. Undersea coal accounts for almost 30% of the total coal production in Japan and just less than 10% in Great Britain.

For most of the bedded deposits that extend from shore workings a shaft is sunk on land with access under the sea by tunnel (**Fig. 4*a***). Massive and vein deposits are also worked in this manner. Normal mining methods are used, but precautions must be taken with regard to overhead cover. Near land and in shallow water a shaft is sunk at sea on an artificial island (Fig. 4*b*). The islands are constructed by dredging from the seabed or by transporting fill over causeways. Sinking through the island is accompanied by normal precautions for loose, waterlogged ground, and development and mining are thereafter conventional. The same method is also used in oil drilling. Offshore drilling and in-place mining (Fig. 4*c*) are used only in the mining of sulfur, but this method has considerable possibilities for mining other minerals (such as hydrates, metalliferous sulfides, or deep placers) for which leaching or slurrying are applicable. Petroleum drilling techniques are used throughout, employing stationary platforms constructed on piles driven into the sea floor or floating drill rigs. *See* OIL AND GAS, OFFSHORE; UNDERGROUND MINING.

Minerals produced by offshore mining include dissolved minerals, unconsolidated minerals, and consolidated minerals (**Table 2**). Though this is only a fraction of world mineral production, the percentages will increase as problems of technology, environment, and infrastructure are resolved.

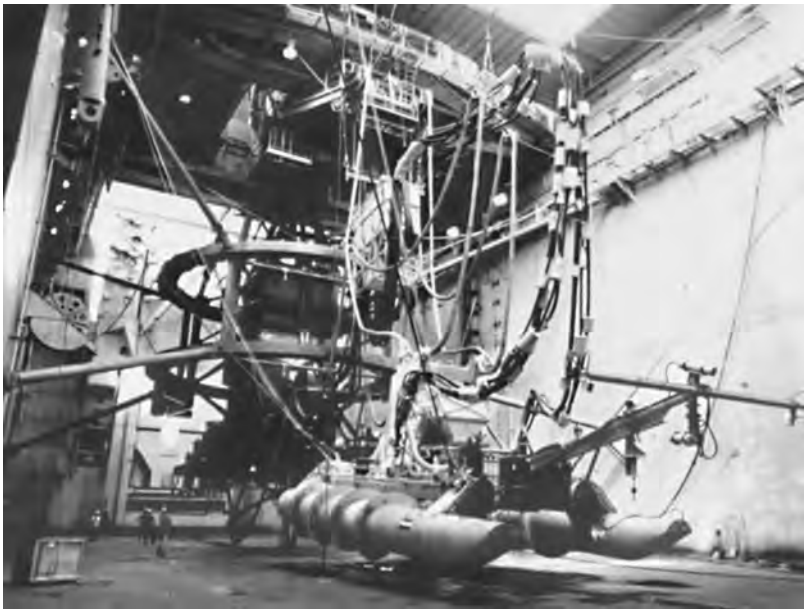

Fig. 3. One-tenth-scale remotely piloted vehicle test miner; note Archimedes' screw drive cyclinders. (*Ocean Mineral Company*)

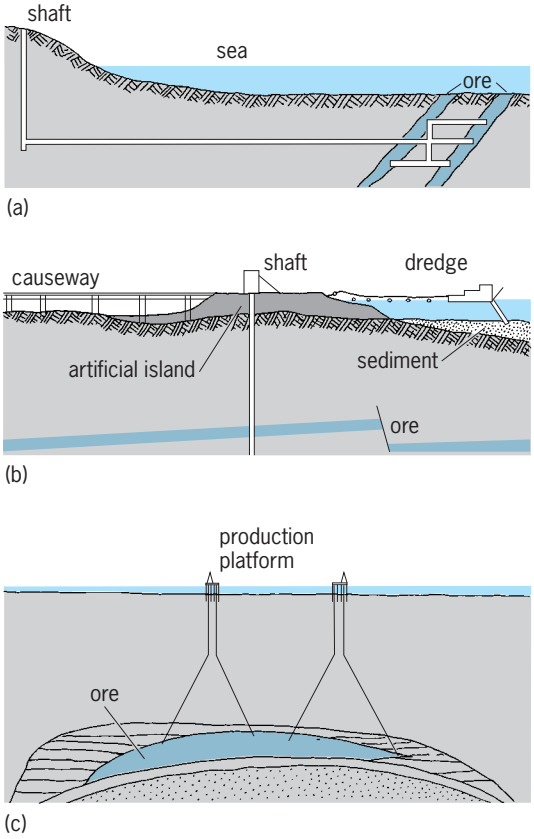

Fig. 4. Methods of mining for exploitation of consolidated mineral deposits offshore. (*a*) Shaft sunk on land, access by tunnel. (*b*) Shaft sunk at sea on artificial island. (*c*) Offshore drilling and in-place mining.

TABLE 2. Summary of production from mineral deposits offshore

Type	Minerals	Percent of world production value
Dissolved minerals	Sodium, magnesium, calcium, bromine	25
Unconsolidated minerals	Diamonds, gold, heavy mineral sands, iron, sands, tin sands, lime shells, sand and gravel	<1
Consolidated minerals	Iron ore, coal, sulfur, hydrates	<1

The future. Despite the technical problems to be overcome, the future of the undersea mining industry is potent. With the advent of national exclusive economic zones claimed by most sovereign coastal states, there is a strong impetus to understand the nature of the resources that are potentially available from these vast new jurisdictional areas. The distinction between shallow and deep-water seabed minerals is no longer political, as in continental-shelf and international seabeds, because many of the exclusive economic zones extend to abyssal depths. The most significant discoveries of marine minerals since the early 1970s have included metalliferous muds in the Red Sea, manganese nodules in the Pacific and Indian oceans, massive metalliferous sulfide deposits in the Atlantic and Pacific oceans, widespread seabed encrustations of high-cobalt metalliferous oxides in the Pacific, and gas hydrates in continental margins and deep ocean basins.

Seawater represents the greatest accessible resource of many elements known and dissolved mineral extraction from seawater by conventional means results in a significant portion of the world's mineral production. As appropriate technology is developed, it is likely that the extraction of minerals directly from the sea will increase. The presence of hydrothermal fluids rich in minerals, associated with sea-floor spreading centers and other volcanic phenomena, presents an additional potential for the production of minerals and geothermal energy.

The mining of unconsolidated deposits nearshore will continue to be significant for the supply of construction materials and heavy minerals and will benefit from improved technology, particularly as deposits are sought in deeper water, further from shore, and in high latitudes. In the deep seabeds, extensive testing of systems for recovering manganese nodules, during the late 1970s, indicated almost conclusively that the nodules are a recoverable resource that await only a suitable economic climate and political infrastructure. Similar conclusions may be drawn from the successful testing of systems to mine the metalliferous muds in the Red Sea.

Consolidated deposits may call for a variety of new mining methods that will depend on the type, grade, and chemistry of the deposit, its distance from land, and the depth of water (**Fig. 5**). The possibility of direct sea-floor access at remote sites through shafts drilled in the sea floor was given consideration under the U.S. Navy's Rocksite program during the 1960s and will be directly applicable to some undersea mining operations. In relatively shallow water, shafts could be sunk by rotary drilling with caissons. In deeper water, the drilling equipment could be placed on the sea floor and the shaft collared on completion. The laying of large-diameter undersea pipelines has been accomplished over distances of 40 km (25 mi) and has been planned for greater distances. Subestuarine road tunnels have been built using prefabricated sections. The sinking of shafts in the sea floor from the extremities of such tunnels should be technically feasible under certain conditions.

The most likely methods for the development of hard-rock submarine ore bodies will be some variation of solution mining technology, involving fracturing of the rock mass in place and removing the valuable materials in solution through a pattern of boreholes. Major problems to be overcome include the control of the fracture pattern and the selective extraction of the desired elements from complex ores. *See* SOLUTION MINING.

In the late 1990s, Australian scientists discovered rich deposits of gold, silver, copper, and zinc sulfide in water at a depth of 1200–1700 m (4000–5700 ft) in the Manus Basin off the coast of Papua New Guinea.

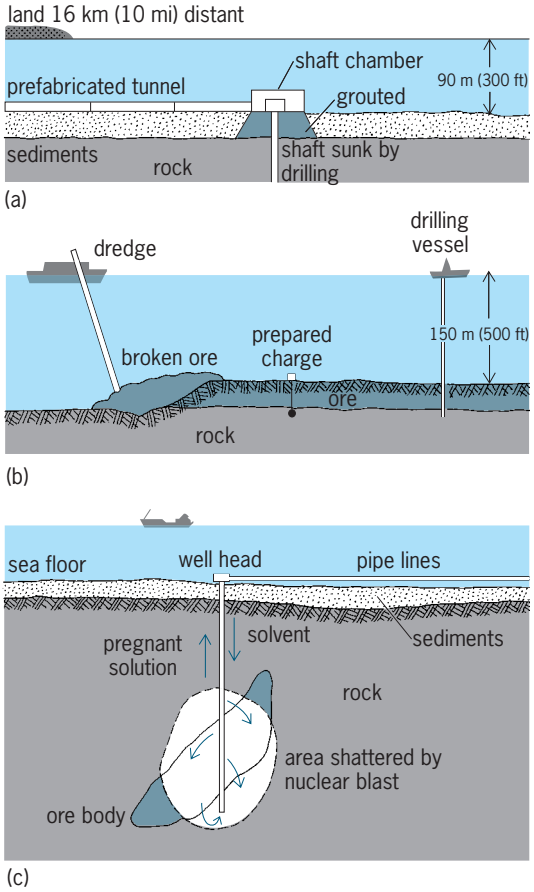

Fig. 5. Possible future methods of mining consolidated mineral deposits under the sea. (*a*) Shaft sinking by rotary drilling from tunnels laid on the seabed. (*b*) Breaking by nuclear blasting and dredging. (*c*) Shattering by nuclear blasting and solution mining.

Metal content of samples from the areas being studied have ranged 10–15% copper, 3–26% zinc, 15–21 g of gold per ton, and 130–200 g of silver per ton. The deposits are surficial masses and chimneys formed around high-temperature hydrothermal vents of mineralized fluids or smokers resulting from tectonic activity associated with the movement of oceanic plates. With drilling for oil and gas now becoming state of the art in similar water depths, the transition to active mining development using off-the-shelf equipment appears likely in the near future. Exploration for similar deposits is being carried out in other areas.

Major technological advances achieved will enhance the possibility of mining in the future, including underwater surveying, advanced underwater vehicle systems, and hydrometallurgical processing. The most important of these may be in the combination of seabed imagery and global positioning, for without adequate maps the characterization and development of the resources of the seabed could not be reasonably pursued.

Outlook. There have been some significant advances in understanding minerals of the seabeds. The ability to manage these minerals as sustainable resources has been enhanced by the passage of legislation. The most important understanding is that, despite the difference in the geological environment of the seabeds compared to land, the potential for mineral occurrence, hectare for hectare, is equivalent in both. On this basis, the global mineral resource base has recently been increased by a factor of 4, and three-quarters of it is underwater. The most important legislation has been the Law of the Sea, in what is probably the most far-reaching redistribution of natural resources in human history. This law has resulted in a peaceful subdivision of jurisdictions, which should significantly affect the world's mineral markets and the balance of economic powers, at least in the Pacific, for as long as humans are dependent on the use of mineral raw materials.

Michael J. Cruickshank

Bibliography. H. R. Cooper, *Practical Dredging*, 1981; D. S. Cronan, *Underwater Minerals*, 1980; M. J. Cruickshank, Law of the Sea and minerals development, *Ocean Yearbook*, vol. 13, pp. 80–106, Dalhousie Law School, 1998; M. J. Cruickshank, Marine mining, in *SME Mining Engineering Handbook*, pp. 1985–2027, 1992; M. J. Cruickshank, Mining, and mineral recovery, *U.S.T. Handbook Directory*, pp. A/15–A/28, 1973; M. J. Cruickshank and C. L. Morgan, *Synthesis and Analysis of Existing Information Regarding Environmental Effects of Marine Mining*, Consulting Report to Continental Shelf Associates for U.S. Department of the Interior, Minerals Management Service, OCS Study MMS 93-0006, 1992; Deep Seabed Polymetallic Nodule Exploration: Development of Environmental Guidelines, *Proceedings of the International Seabed Authority's Workshop*, Sanya, Hainan Island, People's Republic of China, June 1–5, 1998; F. C. F. Earney, *Marine Mineral Resources*, Routledge, 1990; P. Hoagland III and J. Broadus, *Seabed Material Commodity and Resource Summaries*, Woods Hole Oceanog. Inst. Tech. Rep. WHOI-87-43, October 1987; R. H. Joynt, R. Greenshields, and R. Hodgen, Advances in sea and beach diamond mining techniques, *S.A. Min. Eng. J.*, p. 25 ff., 1977; A. Krull, *Ocean Mining and Engineering: New Directions and Opportunities*, 1983; Society of Mining Engineers, *Mineral Resources of the Pacific Rim*, 1982; United Nations (ed.), *Seabed Minerals Series*, 3 vols., 1983; U.S. Congress, Office of Technology Assessment, *Marine Minerals: Exploring Our Ocean Frontier*, 1987.

Marine navigation

The process of directing a watercraft to a destination in a safe and expeditious manner. From a known present position, a course is determined that avoids dangers, and on this course estimates are made of time schedules. The task is to make periodic adjustments using en route measurements of position or heading and speed.

The method used will depend on the type of vessel and on its role or mission. The devices available range from a simple compass to a host of sophisticated electronic systems. In all cases, the navigator must plan and prepare by setting instruments in order and by checking for predictable current and tidal effects and hazards to navigation en route. This preparation includes having the latest, correct charts and reviewing pertinent sections of sailing directions, tabulations of navigational lights, and tide and current tables.

The prudent navigator of a ship planning to cross an ocean will consider the particular requirements of each segment of the voyage. A voyage may be broken into six phases: preparatory, departure, confluence area after departure, high seas, confluence near destination, and terminal.

The methods used to fulfill the requirements of these phases fall under one of the following broad categories of navigation: dead reckoning, piloting, celestial navigation, and electronic navigation. The first three categories have become somewhat standardized; the fourth category has been under constant and innovative development. Electronic navigation is rapidly evolving into integrated systems providing all navigation functions. The publishing of full portfolios of electronic navigational charts and related publications, and the establishment of international standards for systems that display and integrate information, have allowed navigators and maritime administrators to consider reliance on full electronic solutions. The following discussion covers the methods and technologies, and how they may be applied onboard ships.

Dead reckoning. A navigator maintains a reckoning of present and future positions of a ship that is deduced from measurements of direction and speed along the track. In the absence of systems that provide position-fixing capability, this reckoning may be the only way of estimating position on a voyage.

Position fixing is always a past event. Hence, a projection of position requires the use of dead reckoning. It is customary to do this by hand by using pencil, paper, and plotting instruments. However, there are computerized electronic devices that have brought dead reckoning to a high state of sophistication. Among these are marine course computers that will automatically provide such positions besides indicating the command heading, speed made good, distance to destination, estimated time of arrival, and steering signal to gyropilot. *See* DEAD RECKONING.

Inertial navigation is a position-fixing technique that estimates current position by accumulating position changes since the last known position. This self-contained method of navigation eliminates the need for any external references and is passive, that is, nonradiating. The basic principle of inertial navigation is to measure the components of a craft's acceleration along precisely defined axes. If the initial conditions are known, successive integration will produce the velocities and distances along those axes. The key inertial component is the gyroscope with its characteristic "rigidity in space" and its phenomenon of precession. The basic instrument for measuring acceleration is the accelerometer. A digital computer is used for converting accelerations into position. *See* ACCELEROMETER; GYROSCOPE; INERTIAL GUIDANCE SYSTEM.

Another self-contained positioning system is the Doppler sonar navigator. In this system, ultrasonic beams are transmitted to the ocean floor in the forward, aft, starboard, and port directions, and reflected by the ocean floor. Relative motion between the craft and the ocean floor produces Doppler shifts in the frequencies of the reflected signals, the amount of the shift of each beam being directly proportional to the speed of the craft in the direction of that beam. A digital computer converts the Doppler shift of each frequency to speed, and if a reliable reference direction indicator is available, the resultant velocity—speed and direction—of the craft is determined. Distance traveled is determined by integration of the speed with respect to time. Velocity thus determined is with respect to the ocean floor. If the depth of water is greater than about 400 m (220 fathoms or 1320 ft), signals are returned by water reverberation rather than by reflection from the ocean floor, and the velocity is therefore relative to the water. The system can be adapted as an aid in the docking maneuvers of ships. In this use, additional transducers are employed to provide bow, stern, and athwartships velocities. *See* DOPPLER EFFECT; SONAR.

Piloting. This method of navigation is used when the ship is close to land and involves the use of landmarks and seamarks for frequent determination of position. Such marks may be observed by visual, radio, or sound techniques. Visual piloting consists of sighting identifiable landmarks at known positions. The radio system most widely used in this piloting sequence is radar that, besides giving a chartlike pictorial display, can provide bearings and distances for position fixing. The acoustical approach utilizes ultrasonic depth finders. These instruments provide depth measurements that may be compared with those on the chart. Such comparisons can be used to confirm positions or can be used as forewarnings of possible stranding. *See* PILOTING; RADAR.

Celestial navigation. The high-seas phase of a voyage involves the use of either celestial navigation or satellite navigation systems. Sometimes a combination of the two is used. Celestial navigation was formerly considered the primary method of fixing the position of a craft when it was out of sight of land. When using celestial navigation, the navigator follows a routine schedule of a morning twilight observation for a fix, a morning Sun observation of both altitude and azimuth for a morning Sun line of position from which an estimated position can be determined using the dead-reckoning plot, as well as a check on the errors of magnetic compasses and gyrocompasses, an observation of the Sun at meridian transit for a latitude line, and afternoon and evening twilight observations similar to those in the morning. Altitude observations are made with a marine sextant and are timed to the nearest second. For each body observed, the altitude, time, assumed position, and pertinent data from a nautical almanac are converted to a line of position by using prepared tables or a computer. Although two lines of position suitably distributed in azimuth can provide a fix, mariners typically require at least four lines of position unless that is not possible under the prevailing conditions. The celestial bodies observed are the Sun, stars, planets, and the Moon when it is in favorable position relative to other bodies. The principal limitations of celestial navigation are the need for a well-defined natural horizon and for visible celestial bodies. Celestial navigation serves as a backup technique on most vessels and is usually practiced by aficionados or students. The practice of celestial navigation is of value in educating mariners to better understand satellite navigation, since the effects of geometry and measurement quality are similar whether one is navigating by a constellation of satellites or a collection of celestial objects. *See* CELESTIAL NAVIGATION.

Electronic navigation. This division of navigation previously encompassed only radio navigation systems. It now includes navigation involving any electronic device or instrument. A system that utilizes radio signals to provide lines of position or to fix the position of a craft involves the use of transmitters at accurately defined positions on the surface of the Earth or in satellites. Most common today is the use of satellite navigation using the Global Positioning System (GPS). In the near future new signals will be added to GPS, and the European Union is building a compatible system called Galileo. *See* ELECTRONIC NAVIGATION SYSTEMS.

Loran C. Loran C is a hyperbolic system operating at 100 kHz. It combines the pulse-matching techniques of the former Loran A with phase comparison for greater accuracy. Each chain of Loran C stations consists of a master station operating with two or more secondary stations that transmit synchronized

signals. The useful range extends outward to the limit of ground-wave reception, about 1500–2200 km (810–1190 nmi) from the transmitters. Increased accuracy can be obtained by using Loran C in the differential mode. In this mode the reading of a monitor receiver at a known position is compared with the predicted reading, and any difference is then transmitted to vessels in the vicinity and applied as a correction to their readings. Loran C can also be used in a range-range mode, in which position is determined from the ranges to two transmitters, if the user has a stable, accurate time indicator synchronized with the time sources at the transmitters. Loran C has been designated the official system for the United States confluence and coastal regions. In the mid-1990s the advent of satellite navigation led to decisions by the federal government to decommission Loran C by the year 2000. However, the lifetime of the system has now been extended indefinitely. The need for Loran C as a backup system for navigation and timing is being examined by the U.S. Department of Transportation. The Northwest European Loran-C System (NELS) organization is undertaking a similar examination regarding future operations, with further complications arising from the circumstance that several countries host stations. *See* LORAN.

Global Positioning System. The United States' Global Positioning System (GPS) is an advanced navigation satellite system whose space segment consists of 24 satellites in 12-h orbits, 20,200 km (10,907 nmi) above the surface of the Earth. The control segment consists of several monitor stations to continuously track the satellites above the horizon, a master control station (with one or more additional for redundancy) to receive the tracking data and update the orbital data and satellite clock drifts, and upload stations to transmit the data to the satellites daily or more often. The user segment receives the signals transmitted by the satellites and computes pseudorange and pseudorange rate, which it converts to three-dimensional position of the craft. The average velocity of the craft during a measurement interval is the change in carrier phase of the GPS signal averaged over that interval. Time is available to 0.1 microsecond.

With 24 functioning satellites in orbit, five or more are above the horizon virtually all the time. Signals from four satellites are sufficient for computation of three-dimensional positions. Users authorized to have access to the Precise Positioning Service (PPS) can expect absolute accuracy of positions of 5–10 m (16–32 ft). Other users have access to Standard Positioning Service (SPS), providing 20 m (65 ft) accuracy. In the differential mode an accuracy of 3–5 m (10–16 ft) can be achieved. The GPS can be integrated with other systems for redundancy and integrity measurement. *See* SATELLITE NAVIGATION SYSTEMS.

Galileo. The European Union is building a satellite navigation system called Galileo. Galileo is generally similar to GPS. Significant differences in the programs are the level of civilian control of policy in Galileo, a greater number of available services, and the upfront development of applications that take advantage of Galileo.

Automatic Identification System. The Automatic Identification System (AIS) was developed under the guidance of the International Maritime Organization (IMO), primarily to improve maritime safety by assisting the navigation of ships, the protection of the environment, and the operation of vessel traffic services (VTS). AIS serves in a vessel-to-vessel mode for collision avoidance, as a means for coastal states to obtain information about a ship and its cargo, and as a ship-to-shore method for vessel traffic management.

The AIS device includes a Global Positioning System (GPS) receiver, a microprocessor, and a very high frequency–frequency-modulation (VHF-FM) transceiver. The microprocessor takes data from a ship's sensors and packages it with the vessel's identification into a digital signal. The signal is automatically broadcast. When other broadcasts are received, the processor prepares the received data for display. Each device determines its own transmission schedule, and, in the presence of other stations, uses receipt of previous messages to ensure that stations do not transmit simultaneously. This self-organizing logic allows many AIS units to operate without interfering with one another. AIS information includes static, dynamic, and voyage-related elements. Static information is manually entered on installation and seldom changed. Dynamic information is automatically entered from ship sensors, while voyage-related information is manually entered and updated as appropriate. The primary emphasis of AIS is collision avoidance.

AIS is expected to improve vessel-operator situational awareness. In comparison to radar and automatic radar plotting aids ("own ship" information), AIS information is provided directly from the "other" ship's sensors with each vessel's name and call sign. AIS will also provide an efficient method to monitor vessel movements for purposes of environmental safety, coastal security, or traffic management. Though there is as yet no agreement on the definition of an AIS base station, coastal and harbor monitoring will probably be accomplished through a local network of AIS stations forming an "AIS Reception Network." On a regional or national level, the multiple, local networks might be linked through a wide-area network.

Vessel traffic services. Vessel traffic services (VTS) have been established in a number of heavily trafficked ports throughout the world in an attempt to reduce the number of collisions and strandings and safeguard the environment. Generally the service provides marine traffic management of an advisory nature, but in an especially hazardous situation it may be necessary for the VTS to exercise emergency control of vessel movements. The usual system components of such a service include AIS base stations, high-resolution radars, general-purpose computers, and synthetic video displays, all combined to provide real-time forecasts of vessel movements, lane stray alerts, collision and grounding warnings, and a

description of potential congestion areas. A ship-to-shore voice communications link permits information to flow between VTS operations personnel and identified ships within the traffic surveillance area. These voice communications are often recorded on audio tape units. The system tracks stationary and moving traffic to establish position coordinates, and monitors the position of floating navigational aids in the surveillance area to permit detection of misplaced or drifting buoys. *See* VESSEL TRAFFIC SERVICE.

Traffic-separation schemes. Traffic-separation schemes have been established in a number of high-traffic density areas throughout the world, primarily to decrease the risk of collision at sea. International navigation rules govern the operation of vessels within these traffic separation schemes. A typical scheme consists of the establishment of parallel traffic lanes separated by an intervening buffer zone, analogous to a divided freeway on land. In some areas, surveillance centers provide information on the positions and movements of vessels in the area, and other hazards. Although compliance with such schemes is generally advisory only, a dramatic reduction in collisions in the areas where they have been established has contributed to increased acceptance by mariners, and the present trend is toward mandatory compliance, analogous to air-traffic control, but generally leaving the mariner greater freedom of selection of route, speed, and time of encounter of other craft.

Shipboard systems. The human–machine interface for the practice of marine navigation is evolving. Once the navigator was an expert on a complex set of publications, making precise optical measurements, and plotting the results with the accuracy of a draftsman. Now the navigator retains these skills and must also learn to work with complex computer-based systems that integrate several sources of real-time information and compare the results to large quantities of stored hydrography and cartography. The trend is toward more integration with an accompanying assessment of the integrity or quality of separate information sources. This integration is becoming possible through redundancy of the essential information regarding one's own ship and other nearby ships. For example, a crossing situation with another ship can be evaluated with both radar and AIS information. Basic sensors in these systems include GPS receivers, radar, AIS transceivers, a gyrocompass or transmitting magnetic compass, and normally some type of autopilot for steering and perhaps speed control. In ascending order of sophistication, a brief description of the major systems follows:

Electronic chart system. An electronic chart displays on a video screen the same type of hydrographic information that mariners seek in a traditional nautical chart. Both electronic and paper charts offer graphic representations of water depths, shorelines, topographical features, aids to navigation, and hazards.

An electronic chart database and display offers improvement over the paper chart when it is combined with other information. At a minimum, the ship's position (from electronic navigation) and planned track are needed. With such a display, the navigating officer can determine at a glance the ship's position relative to its intended track, the shoreline, the waterway, and any hazards or threats.

Electronic charts integrated with a range of information, and with hardware and software that can process a hydrographic database to support decision making, are classified as electronic chart display and information systems (ECDIS). Thus, electronic charts are a component of an ECDIS. In addition to displaying a real-time picture of the vessel's position in the waterway, an ECDIS manages navigational and piloting information (typically, vessel-route-monitoring, track-keeping, and track-planning information) to support navigational decision making. The term ECDIS is specifically used to indicate an approved device for carriage onboard internationally regulated commercial vessels. The term "electronic chart system" (ECS) is somewhat more generic and is used to indicate a range of systems on various vessels.

Collision avoidance systems. Collision avoidance systems (CAS), also known as automatic radar plotting aids (ARPA), of varying degrees of sophistication have been developed to reduce the work load of the navigator and eliminate human error. Typically, such a system consists of a digital computer that receives inputs from the ship's radar, compass, and log, and determines and displays collision threats, and in some installations provides a recommended avoiding action.

Integrated navigation systems. Integrated navigation systems integrate the functions of the electronic chart system and collision avoidance systems, and normally some level of automated control of vessel course and speed. The advent of the AIS gives a big technical push towards integration in order to correlate the AIS-derived collision avoidance information with the traditional radar tracks. Automation of navigation tasks through an ECDIS combined with the accuracy of plotting intercepts in an integrated collision avoidance system gives the navigator better information and more time to consider and act versus manually plotting and correlating disparate sources. Some systems may provide a visual display of traffic lanes and possible navigation hazards. They may also perform prevoyage route planning and then continuously maintain a dead-reckoning position. Other features available in some are the computation of course to steer to the next waypoint and the use of adaptive automatic steering wherein the vessel's reaction to the outside environment is translated into corrective steering changes designed to optimize the track and decrease the possibility of damage to the ship and its cargo.

Intelligent systems. There are a variety of intelligent systems deployed aboard automated ship's bridges: piloting expert systems, engineering and vibration expert systems, neutral network systems for adaptive and intelligent steering control, and automated intelligent docking systems. Piloting expert systems provide advice for navigation and collision avoidance by

combining the knowledge and reasoning processes of expert ship's captains, ship's officers, and pilots so as to assist navigators during trips. Such systems make use of real-time positioning information, nautical rules of the road, local knowledge, and organizational procedures and regulations to recommend best courses of action for the navigator. *See* EXPERT SYSTEMS; NEURAL NETWORK.

Shipboard piloting expert systems are integrated into conventional and advanced bridge designs so as to provide intelligent ship control systems. Such systems use input from radar, automated radar plotting aid (ARPA), ECDIS, and bridge instrumentation, and employ heuristic or qualitative reasoning, rather than strict application of deterministic rules.

Integrated bridge systems. These systems are designed to allow the wheelhouse to function as the operational center for navigational and supervisory tasks aboard the ship. These bridges in many cases become ship's operations centers, incorporating controls and monitors for all essential vessel functions, including navigation, engine control, and communications. Other ship functions such as cargo loading, monitoring, and damage control may be included. Many routine navigational tasks, such as chart updating, position plotting, and steering, may be automated. The integrated ship's bridge is thus a unified federation of systems supporting vessel navigation, communications, steering, administration, collision avoidance, safety, and monitoring and control of ship's systems. *See* NAVIGATION.

Jay Spalding; Alton B Moody; Martha R. Grabowski; Richard Greenspan

Bibliography. S. F. Appleyard, *Marine Electronic Navigation*, 2d ed., 1988; N. Bowditch, *American Practical Navigator: An Epitome of Navigation*, National Imagery and Mapping Agency, H. O. Publ. 9, 2002; T. J. Cutler, *Dutton's Nautical Navigation*, 15th ed., 2004; R. R. Hobbs, *Marine Navigation*, 4th ed., 1998; F. J. Larkin, *Basic Coastal Navigation*, 2d ed., 1998; A. B Moody, *Navigation Afloat: A Manual For the Seaman*, 1980; National Research Council, *Minding the Helm: Marine Navigation and Piloting*, 1994; H. H. Shufeldt, G. D. Dunlap, and B. A. Bauer, *Piloting and Dead Reckoning*, 4th ed., 1999.

Marine refrigeration

Marine refrigerating equipment is used for shipboard refrigeration of products as well as for air conditioning the quarters of passengers and crew. Shipboard refrigeration is necessary for the preservation of perishables in transit and foodstuffs to be used by passengers and crew. Marine refrigeration is also used for maintaining certain cargo products in liquid form that would otherwise evaporate when stored at ambient conditions. A typical ship stores or cargo refrigeration plant consists of a compressor, motor, condensor, receiver, evaporators, and controls. A typical air-conditioning unit consists of a compressor, motor, condenser, receiver, water chiller, and controls. *See* AIR CONDITIONING; REFRIGERATION.

Refrigerants. The common refrigerants that have been in use for years, known as CFCs (chlorofluorocarbons) and HCFCs (hydrochlorofluorocarbons), are still being phased out because they contain chlorine, which depletes Earth's stratospheric ozone layer. The stratospheric ozone layer protects all life from ultraviolet radiation. In recent years, holes have been detected in the ozone layer in both hemispheres in late winter and early spring. CFC and HCFC refrigerants have been determined to be major contributors to the holes. Common CFC refrigerants include R-11 and R-12. R-22 is a common HCFC refrigerant.

Replacements for these refrigerants known as HFCs, which are non-ozone-depleting, have been developed. These refrigerants contain fluorine in place of chlorine, and do not destroy the ozone layer. The replacement HFC refrigerants normally provide similar to slightly better efficiencies than the CFC refrigerants they replace. The exception is that HFC-134a, which replaces CFC-12 (R-12) in low-temperature applications, decreases plant capacity approximately 35%. This problem is solved by increasing the compressor speed or changing the compressors. For some systems, the cargo itself is used as the refrigerant, such as when carrying ammonia or petroleum products such as propane or butane.

Preservation of perishables. Cargoes such as perishable foodstuffs and chemicals requiring refrigeration are normally carried in insulated holds. The holds are usually cooled by direct-expansion evaporators which may be gravity-type cooling coils or forced-convection cooling units located in the holds. As an alternative to this system, cargoes may be carried in insulated containers stowed on deck.

Liquid cargoes. Liquid cargoes, such as ammonia, and low-pressure petroleum gases, such as propane and butane, are carried in insulated tanks kept cool by allowing the cargo to evaporate. This process involves drawing off the vapors from the tanks, compressing and condensing the fluids, and returning them to the tanks. Hydrocarbon cargoes of very low temperature, such as liquefied natural gas, are stowed in well-insulated tanks and also are kept cold by allowing the cargo to evaporate. The gases are drawn off and may be burned in the ship's power plant as a partial source of energy for propulsion, or used to power an engine-driven refrigerating unit to assist in keeping the product cold while minimizing the amount of product that must be evaporated for cooling.

Refrigerating systems. For refrigerating systems using insulated holds, liquid refrigerant is delivered to cooling coils, where the refrigerant is expanded, and enters the evaporator as a low-pressure liquid. There it evaporates by absorbing heat from the space being cooled. The refrigerant vapor is then drawn from the evaporator by a compressor, which delivers the refrigerant to a condenser as a high-pressure, high-temperature vapor. The condenser, normally a shell-and-tube heat exchanger cooled by seawater flowing through the tubes, removes the heat

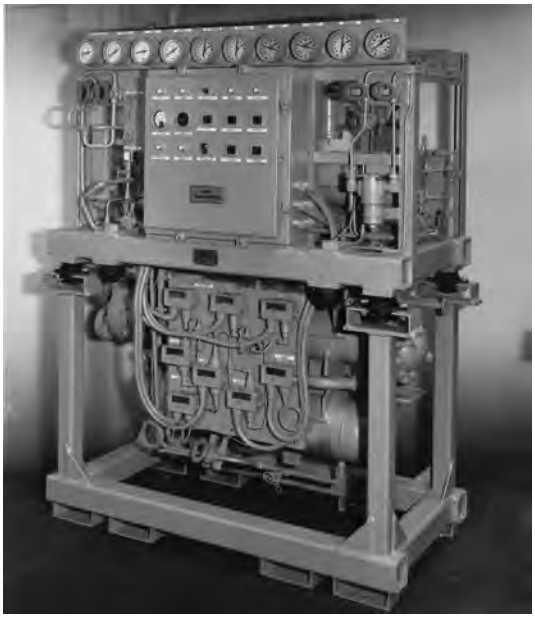

Modern rotary vane HFC-143a ship stores refrigeration unit. (*York International Corp.*)

absorbed in the evaporator and the heat of compression from the refrigerant as it condenses to a high-pressure liquid. The refrigerating cycle is then repeated. For the deck-stowed container system, each container has its own refrigeration system consisting of a compressor, air-cooled condenser, cooling coil, and circulating fan. For liquid cargoes in insulated tanks, the latent heat of vaporization is released as the evaporating liquid changes state to a gas. This is the refrigeration system used in liquid cargo systems.

Ship stores. A modern ship stores refrigeration unit is shown in the **illustration**. Refrigeration systems for ship stores are relatively small compared to cargo refrigeration systems or air-conditioning systems. The refrigerant is circulated through evaporators to maintain the desired temperatures in the various storerooms needed for chilled and frozen products. Two complete refrigeration units are usually employed, with one unit operating and the second on standby.

Air conditioning. For air-conditioning applications, the cooling coils used for refrigeration are replaced with a water chiller. The water chiller is the evaporator where water that is circulated throughout the ship is cooled. The chilled water is circulated to cooling coils situated in various fan rooms for passenger comfort as well as the cooling of equipment. Multiple air-conditioning units are normally employed, and sometimes the units are located in different areas or zones of the ship. For large ships with many passengers, reciprocating compressors are too small; high-tonnage centrifugal-compressor air-conditioning plants are used instead. Reciprocating compressors are normally less than 100 tons, whereas centrifugal compressors can have several hundred tons capacity. Screw compressors can be used for moderate- as well as high-capacity air-conditioning plants.

John H. Merold

Marine sediments

The accumulation of minerals and organic remains on the sea floor. Marine sediments vary widely in composition and physical characteristics as a function of water depth, distance from land, variations in sediment source, and the physical, chemical, and biological characteristics of their environments. The study of marine sediments is an important phase of oceanographic research and, together with the study of sediments and sedimentation processes on land, constitutes the subdivision of geology known as sedimentology. *See* MARINE GEOLOGY; OCEANOGRAPHY; SEDIMENTOLOGY.

Environments of Deposition

Traditionally, marine sediments are subdivided on the basis of their depth of deposition into littoral (0–66 ft or 0–20 m), neritic (66–660 ft or 20–200 m), and bathyal (660–6600 ft or 200–2000 m) deposits. This division overemphasizes depth. More meaningful, although less rigorous, is a distinction between sediments mainly composed of materials derived from land, and sediments composed of biological and mineral material originating in the sea. Moreover, there are significant and general differences between deposits formed along the margins of the continents and large islands, which are influenced strongly by the nearness of land and occur mostly in fairly shallow water, and the pelagic sediments of the deep ocean far from land.

Sediments of continental margins. These include the deposits of the coastal zone, the sediments of the continental shelf, conventionally limited by a maximum depth of 330–660 ft (100–200 m), and those of the continental slope. Because of large differences in sedimentation processes, a useful distinction can be made between the coastal deposits on the one hand (littoral), and the open shelf and slope sediments on the other (neritic and bathyal). Furthermore, significant differences in sediment characteristics and sedimentation patterns exist between areas receiving substantial detrital material from land, and areas where most of the sediment is organic or chemical in origin. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS.

Coastal sediments. These include the deposits of deltas, lagoons, and bays, barrier islands and beaches, and the surf zone. The zone of coastal sediments is limited on the seaward side by the depth to which normal wave action can stir and transport sand, which depends on the exposure of the coast to waves and does not usually exceed 66–99 ft (20–30 m); the width of this zone is normally a few miles. The sediments in the coastal zone are usually land-derived. The material supplied by streams is sorted in the surf zone; the sand fraction is transported along the shore in the surf zone, often over long distances, while the silt and clay fractions are carried offshore into deeper water by currents. Consequently, the beaches and barrier islands are constructed by wave action mainly from material from fairly far away, although local erosion may make a contribution, while the lagoons and

bays behind them receive their sediment from local rivers. The types and patterns of distribution of the sediments are controlled by three factors and their interaction: (1) the rate of continental runoff and sediment supply; (2) the intensity and direction of marine transporting agents, such as waves, tidal currents, and wind; and (3) the rate and direction of sea level changes. The balance between these three determines the types of sediment to be found. *See* DELTA.

On the Texas Gulf Coast, rainfall and continental runoff decrease gradually in a southwesterly direction. The wind regime favors considerable wave action and a southwesterly drift of the nearshore sand from abundant sources in the east. Since sea level has been stable for several thousand years, the conditions have been favorable for the construction of a thick and nearly closed sand barrier that separates a large number of bays from the open Gulf. This barrier is constructed by marine forces from sediments from distant sources and varies little in characteristics along its length. The bays, on the other hand, receive local water and sediment. In the east, the supply of both is fairly abundant, and since the streams are small, the sediment is dominantly fine; the bays have muddy bottoms and brackish waters. Conditions are fairly stable, and a rich, but quantitatively not large fauna is present, including oyster banks. At the southwestern end, continental runoff and sediment supply are negligible.

The only sediment received by the bays comes from washovers from the barriers and is therefore mainly sandy, and the virtually enclosed bays with no runoff are marine to hypersaline. Locally, this yields chemical precipitates such as gypsum and calcium carbonate, and is also conducive to the development of a restricted but very abundant fauna, which produces significant deposits of calcareous material. The sediments of bays and lagoons are often more stratified than those of the open sea as a result of fluctuating conditions. The textural and compositional characteristics depend on local conditions of topography, shore development, and wave and current patterns. They range from coarse gravel and cobbles on rocky beaches fronting the open sea to very fine clayey silt in the interior of quiet lagoons. *See* ESTUARINE OCEANOGRAPHY.

The effects of sea-level changes are imperfectly known, but it is easily comprehended that the development of coastal sediments is to a large extent a function of the duration of this environment in a particular place. Thus if sea level rises or falls rapidly, there is no time for extensive development of beach, barrier, and lagoon deposits, and discontinuous blankets of nearshore sands, with muds behind them, are formed. As the rate of change decreases, open barriers, consisting of widely spaced low sand islands, tend to form, which imperfectly isolate open shallow lagoons in which essentially marine conditions prevail. A prolonged stability is required to produce thick, closed barriers and completely isolate the lagoon environment.

Entirely different nearshore deposits are found on shoals where supply of sediment and freshwater from the land is absent, either because land areas are small (Bahamas), the drainage is directed elsewhere (southern Cuba), or there is no rainfall (Arabian Peninsula). Calcareous muds and coarse calcareous sands then make up the lagoon and beach deposits. If, in addition, the shoal borders directly on the deep ocean without transitional shelf (**Fig. 1**), cool water is driven onto the shoal, where it warms up and precipitates calcium carbonate. In the turbulent water of the shoal, this precipitation either takes place in the form of oolites or it cements organic debris together in small aggregates (grapestone). At the edge of the shoal, the presence of cool, nutrient-rich ocean water is favorable for the growth of coral and algal reefs that are bordered by a zone of skeletal sand derived from broken calcareous organisms. The inner sheltered portions receive the finest calcareous sediments. The types of sediment and their distribution patterns are controlled mainly by the shape of the shoal, in particular the position of its edge, by the prevailing wind, and by the location of sheltered or somewhat deeper quiet areas.

Shelf and slope sediments. The continental shelf is a gently seaward sloping plain of greatly varying width,

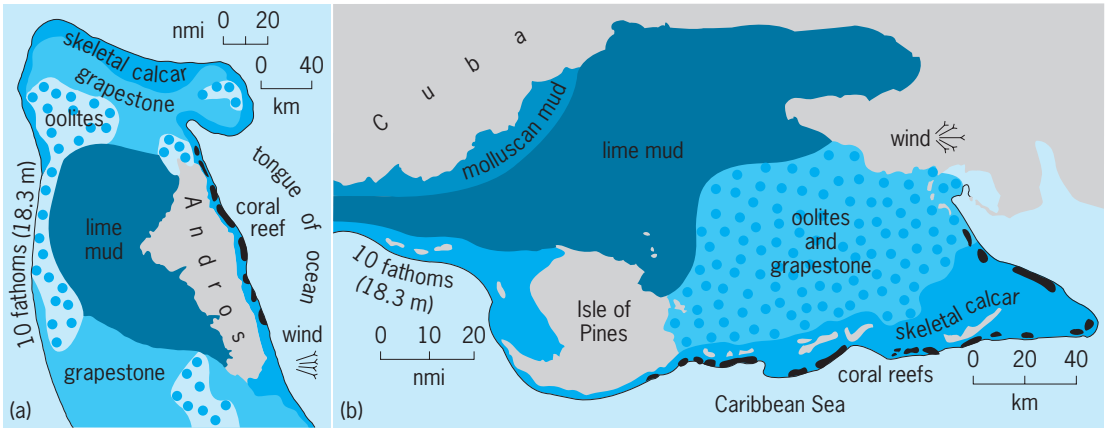

Fig. 1. Shoal water sediments as examples of sedimentation without supply of land-derived sediment. (*a*) The Bahamas. (*b*) The Gulf of Batabano, which is in southwestern Cuba. Both of these shoals border directly on the deep ocean.

ranging from less than a mile (1.6 km) along steep rocky coasts to several hundred miles, for example, in the western Gulf of Mexico. A distinct break in slope at 330–1600-ft (100–500-m) depth marks the transition to the continental slope which descends somewhat more steeply (5–10°) to the deep-sea floor. In areas of active tectonism, for example, off the coast of southern California, the shelf is narrow and separated from the continental slope by a wide zone of deep basins alternating with shallow banks and islands. Submarine canyons cut the edge of the continental shelf in many places and sometimes reach back into the nearshore zone. *See* SUBMARINE CANYON.

During the Pleistocene, the continental shelf was subjected to repeated transgressions and regressions. During each interglacial, sea level was high and the shoreline was located near its present position; during each glacial period, much water was withdrawn from the ocean and the shoreline occurred near the edge of the shelf. The last low sea-level stand occurred approximately 19,000 years ago, and the present shoreline was established as recently as 3000–5000 years ago. On most shelves, equilibrium has not yet been fully established and the sediments reflect to a large extent the recent rise of sea level. Only on narrow shelves with active sedimentation are present environmental conditions alone responsible for the sediment distribution.

Sediments of the continental shelf and slope belong to one or more of the following types: (1) biogenic (derived from organisms and consisting mostly of calcareous material); (2) authigenic (precipitated from seawater or formed by chemical replacement of other particles, for example, glauconite, salt, and phosphorite); (3) residual (locally weathered from underlying rocks); (4) relict (remnants of earlier environments of deposition, for example, deposits formed during the transgression leading to the present high sea-level stand); (5) detrital (products of weathering and erosion of land, supplied by streams and coastal erosion, such as gravels, sand, silt, and clay). *See* AUTHIGENIC MINERALS.

On shelves with abundant land-derived sediment, the coastal zone is composed of deltas, lagoons, bays, and beaches and barriers. Outside the beaches and barriers, a narrow strip of wave-transported sand, usually less than 2 or 3 mi (3 or 5 km) wide, fringes the coast. On the open shelf, the sediment deposited under present conditions is a silty clay, which, near deltas, grades imperceptibly into its bottomset beds. Usually, the silty clay, which results from winnowing near the coast, is carried no more than 20–30 mi (32–48 km) offshore by marine currents, so that the zone of active deposition is restricted. If the shelf is narrow, all of it will fall into this zone, but if it is wide, the outer part will be covered by relict sediments resulting from the recent transgression. These relict sediments were deposited near the migrating shoreline and consist of beach sands and thin lagoonal deposits. They have been extensively churned by burrowing animals and wave action, resulting in a mottled structure, and authigenic glauconite has formed in them. *See* TRACE FOSSILS.

On many shelves, small calcareous reefs (shelf-edge reefs) occur at the outer edge. These reefs apparently depend on the presence of deep water for their growth, although it is not certain that they are growing vigorously at the present time. In the Gulf of Mexico, where they are particularly abundant, they mark the tops of salt domes in the subsurface. Beyond the reefs begins the zone of slope deposition, where in deeper and quiet water silty clays with abundant calcareous remains of open water organisms are being slowly deposited. Thus, there are in principle four parallel zones on each shelf: an inner sandy zone; an intermediate zone of clay deposition; an outer shelf zone of no deposition, where relict sediments occur, terminating in edge reefs; and a slope zone of calcareous clays. This parallel zonation is often strongly modified by special current patterns, which carry fine sediments farther out across the shelf, as in the western Gulf of Mexico; by rapidly advancing deltas that provide a sediment source far out on the shelf, as in the Mississippi delta; or by exposure to unusually vigorous wave action that prevents fine sediments from being deposited, even though a supply is present.

The fine-grained deposits that are being formed tend to be deposited more rapidly near the source than farther away, and as a result contain more biogenous material with increasing distance from the source, so that they become more calcareous (**Fig. 2**). On shelves with little or no land-derived material, the only available sources of sediment are biogenous and authigenic. These sources provide far less material than rivers do, and as a result sedimentation rates are much lower. Even on shelves with abundant supply of land-derived material, the areas of nondeposition are extensive, often 40–50% of the total area. On the calcareous shelves, relict sediment may cover up to 75% of the entire area. Near the shelf

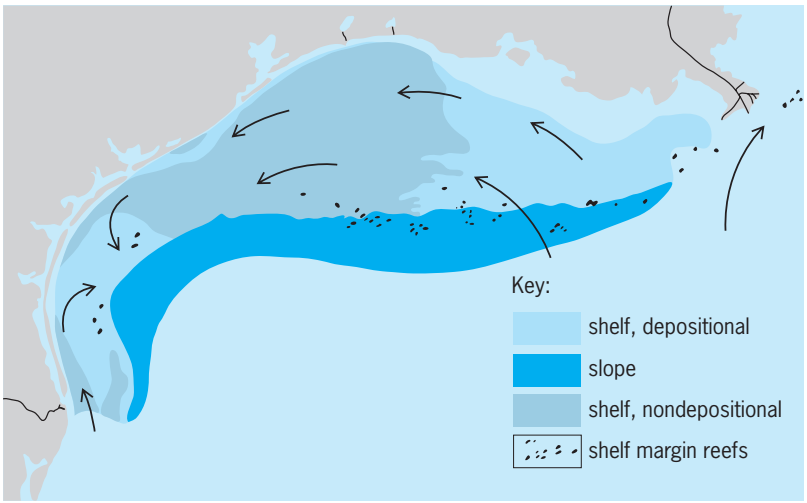

Fig. 2. Sediment distribution of the northwestern Gulf of Mexico as an example of the sediments of a shelf with abundant land-derived sediment. Shelf depositional sediments and slope deposits are silty clays; nondepositional area is covered with relict sediments. Arrows show generalized circulation.

edge, these relict sediments are shallow-water deposits formed during the last low stand of the sea, and consist of small algal reefs and oolites as described for the Bahamas. *See* OOLITE.

Landward of this zone, the deposits, formed when sea level rose rapidly, are thin blankets of calcareous debris, consisting of shell material, bryozoa, coral debris, and so forth (**Fig. 3**). The only active sedimentation zones occur very near the shore, where calcareous and sometimes land-derived sand, silt, and clay are being deposited at present, and at the outermost shelf margin and continental slope, where a blanket of fine-grained calcareous mud very rich in planktonic Foraminiferida is found (foraminiferid calcilutite and calcarenite). Calcareous shelves occur in the Persian Gulf and off the coasts of Australia.

The rise of sea level has, in many regions, severely restricted the supply of sediment to the continental shelf. Along the eastern coast of the United States, the valleys of many rivers have been flooded, and during the present stable sea level, barriers have been built across them which restrict the escape of sediment from the estuaries. As a result, sedimentation on this shelf is slow, and relict transgressive sands occur nearly everywhere at the surface.

Much of the fine-grained sediment transported into the sea by rivers is not permanently deposited on the shelf but kept in suspension by waves. This material is slowly carried across the shelf by currents and by gravity flow down its gentle slope, and is finally deposited either on the continental slope or in the deep sea. If submarine canyons occur in the area, they may intercept these clouds, or suspended material, channel them, and transport them far into the deep ocean as turbidity currents. If the canyons intersect the nearshore zone where sand is transported, they can carry this material also out into deep water over great distances. *See* SUBMARINE CANYON; TURBIDITY CURRENT.

Complex sediment patterns form in areas of considerable relief, for example, the borderland off southern California, where very coarse relict and residual sediments on shallow banks alternate rapidly with silty clays and calcareous deposits in the deep troughs. Such cases, however, are rare along the continental margins. Unimportant, but striking and geologically interesting, are the calcareous sediments associated with coral reefs and atolls. Usually, they occur on islands in mid-ocean, where clear water with abundant nutrients is available, and land-derived sediments are absent; but fringing and barrier reefs with associated calcareous sediments also occur along coasts with low sediment supply. *See* CONTINENTAL MARGIN; REEF. Tjeerd H. Van Andel

Deep-sea sediments. Sediments covering the floor of the deep sea were first systematically described and classified during the late nineteenth century by J. Murray and A. F. Renard (1884, 1891) after their observations during the Challenger Expedition (1872–1876). Their classification included two principal sediment types, terrigenous (sediments deposited near to and derived from continental areas) and

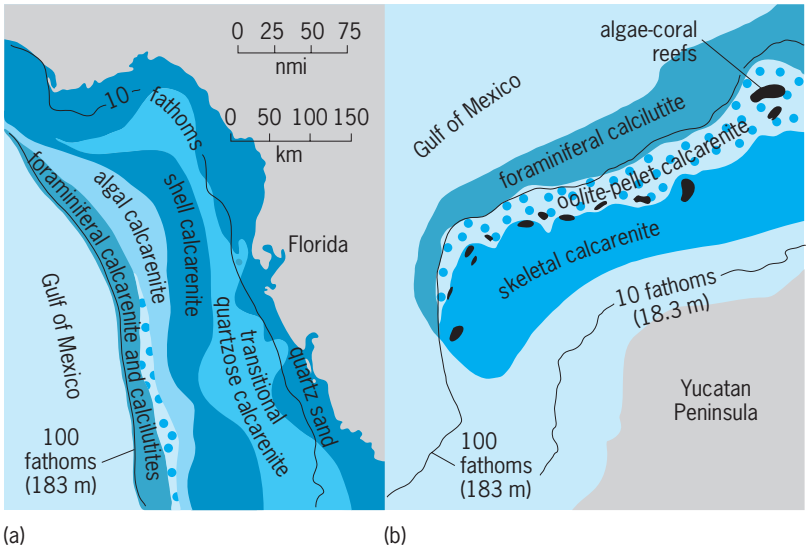

Fig. 3. Sediments on shelves with little land-derived sediment supply. (*a*) Off western coast of Florida. (*b*) Off northwestern Yucatan Peninsula.

pelagic (sediments, principally fine grained, accumulated slowly by settling of suspended material in those parts of the ocean farthest from land). Both categories include biogenic and nonbiogenic material, as well as sediment derived from continents, that is, terrigenous, making the classification, at best, difficult to apply. Some of the terms, such as pelagic red clay, however, have remained in general usage but with meanings somewhat modified from the original definitions. A problem with most classifications has been in distinguishing descriptive categories, for example, red clay, from genetic categories, that is, those that include an interpretation of sediment origins, for example, volcanic mud. In addition, classifications are difficult to apply because so many deep-sea sediments are widely ranging mixtures of two or more end-member sediment types. The following sections will briefly describe the most important end members, their manner of origin, and some of the factors that control their distribution.

Biogenic sediments. Biogenic sediments, those that were formed from the skeletal remains of various kinds of marine organisms, may be distinguished according to the composition of the skeletal material, principally either calcium carbonate or opaline silica. The most abundant contributors of calcium carbonate to the deep-sea sediments are the planktonic foraminiferids, coccolithoforids, and pteropods. Organisms that extract silica from the seawater and whose hard parts eventually are added to the sediment are radiolaria, diatoms, and to a lesser degree, silicoflagellates and sponges. The degree to which deep-sea sediments in any area are composed of one or more of these biogenic types depends on the organic productivity of the various organisms in the surface water, the degree to which the skeletal remains are redissolved by seawater while settling to the bottom, and the rate of sedimentation of other types of sediment material. Where sediments are composed largely of a single type of biogenic

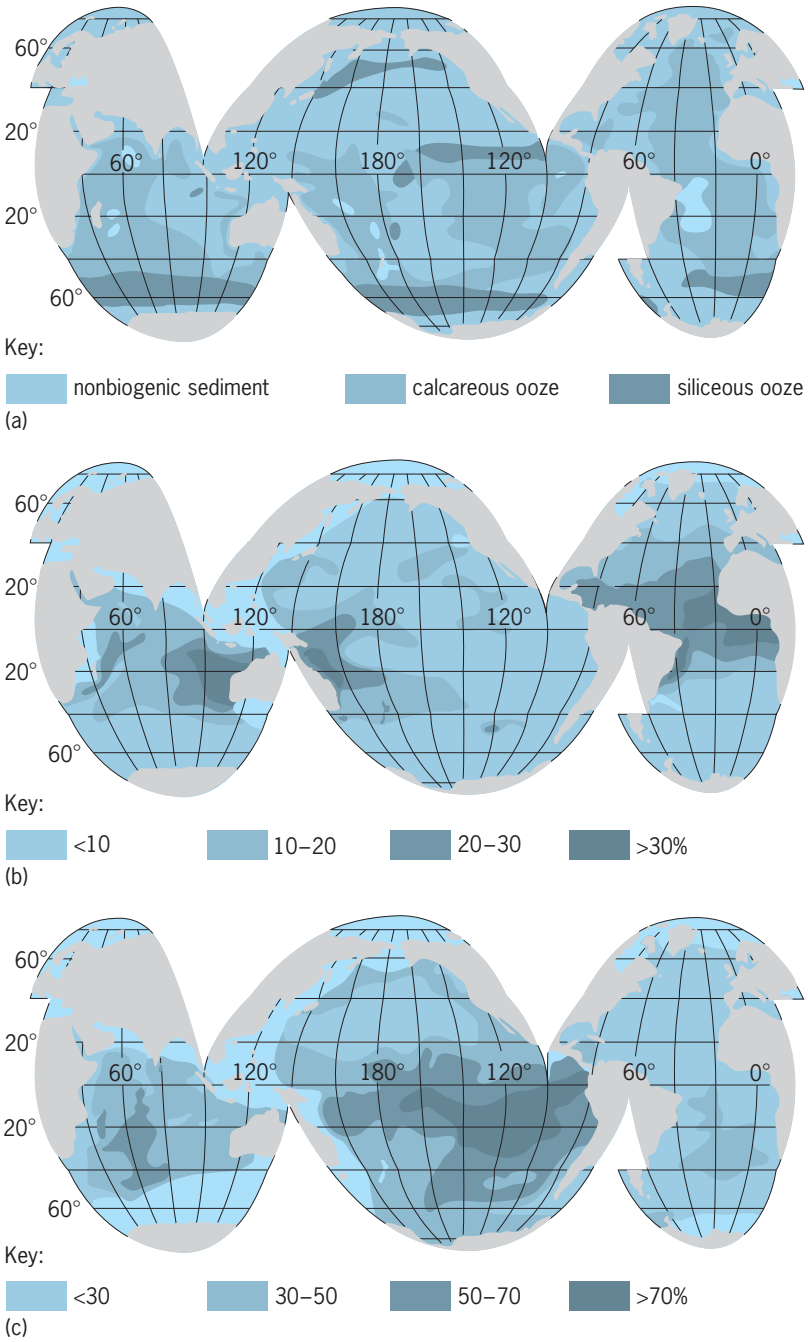

Fig. 4. Distribution of deep-sea sediments. (*a*) Major types. (*b*) Kaolinite concentration. (*c*) Montmorillonite concentrations. The concentrations are in the <2-μm-size fraction on a carbonate-opaline-silica-free basis.

material, it is often referred to as an ooze, after its consistency in place on the ocean floor. Thus, depending on the organism, it may be called a foraminiferal ooze (or globigerina ooze, after one of the most abundant foraminiferid genera, *Globigerina*), coccolith ooze, pteropod ooze, or for the siliceous types, diatomaceous or radiolarian ooze. The distribution map of major sediment types (**Fig. 4*a***) does not distinguish the several types of organism, but indicates that the siliceous oozes are most abundant at high latitudes and in the highly productive areas of upwelling such as the Equatorial Pacific. Calcareous oozes are also abundant beneath zones of upwelling and along topographic highs such as the Mid-Atlantic Ridge. Their relative abundance in shallower water is due to an increased tendency for calcium carbonate (calcite) to redissolve at ocean depths greater than about 13,000 ft (4000 m) because of lower water temperature, greater pressure, and increased dissolved carbon dioxide in the water.

Nonbiogenic sediments. The nonbiogenic sediment constituents are principally silicate materials and, locally, certain oxides. These may be broadly divided into materials which originate on the continents and are transported to the deep sea (detrital constituents) and those which originate in place in the deep sea, either precipitating from solution (authigenic minerals) or forming from the alteration of volcanic or other materials. The coarser constituents of detrital sediments include quartz, feldspars, amphiboles, and a wide spectrum of other common rock-forming minerals. The finer-grained components also include some quartz and feldspars, but belong principally to a group of sheet-silicate minerals known as the clay minerals, the most common of which are illite, montmorillonite, kaolinite, and chlorite. The distributions of several of these clay minerals have yielded information about their origins on the continents and, in several cases, clues to their modes of transport to the oceans. For example, the distribution map of kaolinite concentrations in the less than 2-micrometer fraction (Fig. 4*b*) shows that this mineral occurs in the deep sea principally adjacent to continental areas, mostly at low latitudes, where it is a common constituent of tropical soils. Its distribution indicates that it occurs only in the deep sea as a detrital phase and, in the western Equatorial Atlantic, appears to be carried to the ocean only by rivers. On the eastern side of the Equatorial Atlantic, however, its more widespread distribution corresponds with a region of frequent dust storms and suggests that its transport there is both by rivers and by the Northeast Trade Winds.

One of the common clay minerals, montmorillonite, has both a detrital and volcanogenic origin in different parts of the deep sea. In the Atlantic Ocean, its distribution is controlled by its relative abundance in the soils of adjacent continental areas, whereas in the South Pacific Ocean its great abundance results from its formation, along with authigenic minerals called zeolites, as an alteration product of widespread volcanic glass (Fig. 4*c*). The reason these materials—both the volcanic glass shards and their alteration products—are so abundant in the South Pacific is that they are relatively undiluted by detrital constituents, the South Pacific being so large and rimmed by numerous deep trenches which act as traps for seaward-moving detrital sediments. *See* CLAY MINERALS; ZEOLITE.

An authigenic constituent of deep-sea sediments that has received much attention, because of the potential economic value of trace elements that occur with it (principally copper), is manganese oxide (MnO_4). This material occurs in many areas

as manganese-iron nodules. These are spheroidal pellets of MnO_4 and iron(III) oxide (Fe_2O_3), ranging in size from microscopic to tens of centimeters in diameter and covering thousands of square miles of the deep-sea floor, principally in areas where both biogenic and nonbiogenic sedimentation rates are very slow. *See* MANGANESE NODULES.

Both in the case of biogenic and nonbiogenic sedimentation, the old concept of a slow, gentle "snowfall" of sediment particles to the sea floor where they remain forever buried, has been shown to be untrue in many parts of the ocean. Agencies are at work within the deep sea capable of distributing and redistributing sediment from its point of origin or entry in the ocean. Turbidity currents, high-density clouds of entrained sediment, are responsible for rapidly moving downward vast amounts of sediment from the continental slopes and other topographic highs. In addition, the deep oceanic current systems have been shown to be capable of eroding and transporting sediment along the ocean floor. The evidence for this comes from deep-sea photographs that show current ripple marks and scour marks in many parts of the ocean as well as measurements of relatively turbid zones of water traveling near the bottom (the nepheloid layer) in many parts of the ocean; this research suggests a rather continuous process of sediment erosion and redeposition. This process is discussed in a later section on transport and rate of sedimentation.

Ancient sediments. Because intensive studies of deep-sea sediments are a relatively recent phenomenon in the study of the Earth, and because the surface layer of sediment covers nearly two-thirds of the Earth's surface, investigators have only begun to understand the factors that presently control the distribution of sediments. An entire additional dimension, however, is introduced by extending these studies down through the sediment column, that is, back through geologic time. Samples of sediment have been taken in deep-sea cores back through the Pleistocene, Tertiary, and as old as Jurassic, or more than 130,000,000 years old. The studies on these older deep-sea sediments show that sedimentary conditions were very different from those that now obtain. Several studies in the Pacific Ocean and results from the Atlantic Ocean suggest that sediments derived from volcanic sources were possibly much more important in late Mesozoic and during most Tertiary time, and that the importance of detrital sediment sources is a relatively recent phenomenon, beginning as late as Pliocene or Pleistocene time. *See* JURASSIC; MESOZOIC; PLEISTOCENE; PLIOCENE; TERTIARY. Pierre E. Biscaye

Physical Properties

Physical properties of marine sediments, such as density and the elastic constants, depend on many factors. These include grain shapes, sizes, and compositions; the amount of interstitial fluid and its properties; the nature of grain-to-grain contacts; the degree of compaction and consolidation; and the age. Measurable physical quantities appear to be affected to a much greater extent by the fractional volume of fluid in the sediment (porosity) than by sediment type. This is to be expected since density and elastic constants do not differ much for the principal constituents of sediments (silica, calcium carbonate, clay minerals). Water depth affects physical properties only to a slight extent.

Some measurements of physical properties are made on samples recovered from the ocean bottom by coring devices or in sufficiently shallow water by divers. Such observations are limited to sediments lying within a few tens of feet of the water-sediment interface. Properties of deeper-lying sediments are known from seismic refraction measurements of the velocities of elastic waves, by inference from dispersion of surface waves, or from gravity data (**Table 1**).

TABLE 1. Typical measurements of selected properties of marine sediments

Property measured	Fine sand, 17-station average[a]	Clayey fine silt[a]	Gray clay or silt[b]		Cream calcisiltite[c]		Gray clay[c]	Artificially compacted globigerina ooze pressure,[d] kg/cm^2		
								512	768	1024
Medium grain diameter, mm	0.19	0.02			[0.01][e]	[0.01]				
ρ, g/cm^3	1.93	1.60	1.72	1.58	1.60	1.71	1.46	2.14	2.22	2.26
ϕ, %	46.2	65.6	[56]	[65]	65	57	74	[32]	[28]	[26]
α, km/s	[1.68]	[1.46]			1.59	1.68	1.49	2.68	2.89	3.06
β, km/s								1.20	1.42	1.57
σ (Poisson's ratio)	0.44[f]	0.50[f]						0.37	0.34	0.32
κ, 10^{-11} dyne/cm^2	[0.472]	[0.342]						1.13	1.25	1.38
μ, 10^{-11} dyne/cm^2	[0.06][g]	[0.00][g]						0.31	0.45	0.56
Thermal conductivity, 10^{-4} cal/(cm)(°C)(s)			26.8	23.1						
Approximate water depth, fathoms	15	15	1000	1000	2300	1600	2550			

[a]After E. L. Hamilton et al., 1956. [b]After E. Bullard, 1954. [c]After G. H. Sutton et al., 1957. [d]After A. S. Laughton, 1957. [e]Brackets indicate conversion of units from those used in the original publication. [f]fLower limit. [g]Upper limit.

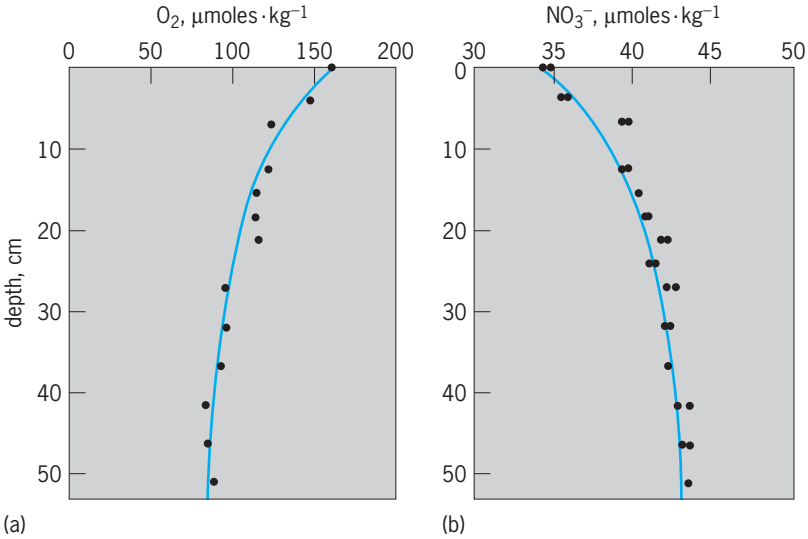

Fig. 5. Dissolved (*a*) oxygen (O_2) and (*b*) nitrate (NO_3^-) in the interstitial water of a pelagic marine sediment. 1 cm = 0.4 in.

Definitions and some useful interrelationships among measurable quantities are given by Eqs. (1)–(5), where ρ_1 and ρ_2 are fluid and average particle

$$\rho = \rho_1\phi + \rho_2(1-\phi) \quad \text{(bulk density)} \tag{1}$$

$$\alpha = \sqrt{(k + {}^4/_3\mu)/\rho} \quad \text{(compressional wave velocity)} \tag{2}$$

$$\beta = \sqrt{\mu/\rho} \quad \text{(shear wave velocity)} \tag{3}$$

$$\left(\frac{\alpha}{\beta}\right)^2 = \frac{2(1-\sigma)}{(1-2\sigma)} \tag{4}$$

$$k = \frac{1}{C} \quad \text{(incompressibility = 1/compressibility)} \tag{5}$$

densities, respectively, ϕ is the porosity ($\phi = 1$ at 100% fluid, $\phi = 0$ at 0% fluid), μ is rigidity, and σ is Poisson's ratio.

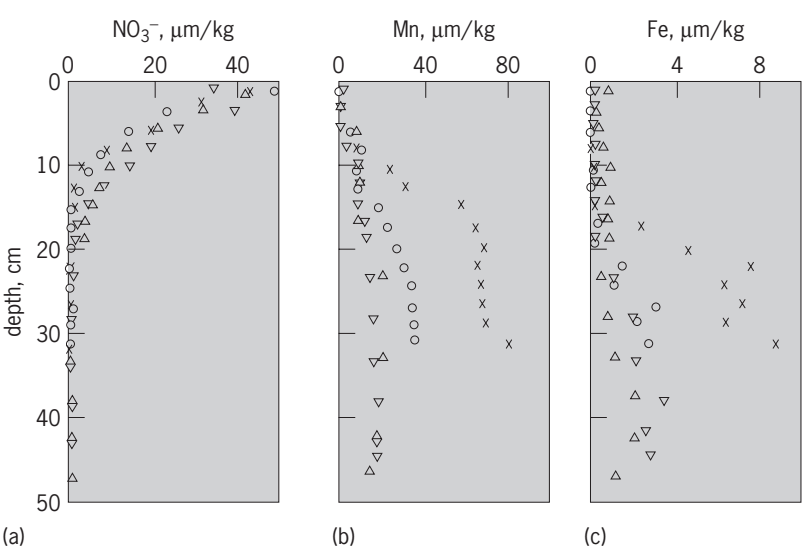

Fig. 6. Dissolved (*a*) nitrate (NO_3^-), (*b*) manganese (Mn), and (*c*) iron (Fe) in the interstitial water from a hemipelagic marine sediment. Symbols represent different cores. 1 cm = 0.4 in.

Several general conclusions may be drawn from observations. Density and porosity are nearly linearly related and for most ocean sediments observations lie between the two lines, as in Eq. (6). Compres-

$$\rho_{cgs} = 1.03 + (1.67 \pm 0.05)(1-\phi) \tag{6}$$

sional wave velocities for $\phi > 0.6$ agree well with the predictions, or equation, of A. B. Wood. This is obtained by inserting into Eq. (2) $\mu = 0$ and $C = 1/k = C_1\phi + C_2(1-\phi)$, where C_1 and C_2 are compressibilities of fluid and particles, respectively. At smaller porosities compressional velocity rises more steeply, the $\phi = 0$ limit being near 6 km/s. Poisson's ratio may be expected to vary from a value near 0.25 at zero porosity to 0.5 at $\mu = 0$. Good agreement of α with the Wood equation for $\phi > 0.6$ indicates that σ reaches its upper limit near $\phi = 0.6$.

Seismic refraction measurements indicate that compressional velocity increases with depth in the sedimentary column, the gradient being from 0.5 to 2.0/s. Thus ρ, μ, k, β should also increase with depth and ϕ and σ decrease. John E. Nafe

Sediment Chemistry

The chemical compositions of interstitial waters are modified from seawater by reactions with organic matter and with the silicate, carbonate, and oxide solid phases in the sediments. While research was formerly concerned primarily with the major ions and silica, emphasis has shifted to the diagenesis of organic matter in marine sediments.

Interstitial waters. Respiration of sedimentary organic matter influences the concentrations of nutrients, alkalinity, and pH. The sequence of reactions for the oxidation of organic matter follows the order of free energy that can be gained during bacterial metabolism. Oxygen is the thermodynamically favored electron acceptor, followed by nitrate, manganese(IV), iron(III), and sulfate. To a first degree, these metabolic processes are segregated into vertically isolated zones. In nature, the boundaries between these zones are obscured by interdiffusion and, in some cases, bioturbation.

Oxidation and nitrification. In the deep sea, little metabolizable organic matter reaches the sediments, the sedimentation rate is slow, and thus the effects of early diagenesis in the top 20 in. (50 cm) of sediments are small. Results on in-place samples from pelagic red clay and carbonate ooze sediments in the Central Equatorial Pacific reveal that aerobic respiration with nitrification is the dominant microbial process. In these sediments, oxygen decreases with depth and the nitrate increases with depth (**Fig. 5**). The stoichiometry of the oxidation reactions must be considered approximate because of uncertainty in the composition of the decomposing organic matter. In all cases from this region, the oxygen profiles level off at some constant value that is greater than 50 μmoles · kg^{-1}.

These profiles can be used to calculate the diffusive flux from the overlying water into the sediments. This steady-state flux can be estimated by using Fick's

first law, as given by Eq. (7), where $D_{O_2,pw}$ is the dif-

$$\text{Flux } O_2 = -D_{O_2,pw}\,\phi\left(\frac{dO_2}{dz}\right)_{z=0} \qquad (7)$$

fusion coefficient for oxygen in the pore water, ϕ is the porosity, and $(dO_2/dz)_{z=0}$ is the gradient at the sediment-water interface. The average diffusive flux of oxygen from seawater into the sediments at these locations was 9.0×10^{-14} moles · cm^{-2} · s^{-1}. The location for aerobic respiration also indicates that alkalinity and pH will decrease during oxygen consumption in red clay sediments. In carbonate ooze sediments, the carbon dioxide (CO_2) produced during respiration dissolves calcium carbonate ($CaCO_3$), and thus the alkalinity will increase. Both of these predicted changes in alkalinity have been observed in pelagic aerobic red clay and carbonate ooze sediments respectively. *See* PH.

Denitrification. Proceeding from aerobic sediments with low organic carbon to sediments with higher sedimentation rates and organic carbon, oxygen is reduced to lower values until it is completely removed. In the absence of dissolved oxygen, bacteria turn to nitrate (NO_3^-) as an electron acceptor and reduce it to nitrogen (N_2) [denitrification]. For example, nitrate decreases to zero by about 8 in. (20 cm) in organic-rich hemipelagic sediments of the Eastern Tropical Pacific (**Fig. 6**). The gradient in NO_3^- results in a flux of NO_3^- from seawater into the sediments, where it is reduced to nitrogen gas. Similar nitrate gradients have been observed in the Equatorial Atlantic. Calculations of the resulting diffusive fluxes of nitrate indicate that hemipelagic sediments may be important sites of marine denitrification.

Manganese and iron reduction. Manganese(IV) reduction and iron(III) reduction are the major microbiologically mediated reactions that follow the reduction of NO_3^-. Thermodynamic calculations suggest that Mn(IV) reduction should precede Fe(III) reductions. These predictions appear to be borne out when the reactions are suitably well spaced over the sediment column. As seen in the hemipelagic sediments of the Eastern Tropical Pacific (Fig. 6), dissolved manganese ion (Mn^{2+}) begins to build up in the interstitial waters after NO_3^- has been almost completely removed. Reduced Fe(II) does not appear until further down in the sediments. The buildup of reduced forms of manganese and iron creates concentration gradients in the interstitial waters that result in a flux toward the sediment–water interface for both metals. Both metals reoxidize when they reach oxidizing conditions. The diagenetic remobilization and upward flux of reduced manganese has been proposed as a source for manganese in ferromanganese nodules, as well as causing an excess of solid manganese in the surface layers of some marine sediments.

The reduction of Mn(IV) and Fe(III) oxides appears to control the interstitial water profiles of other trace metals as well. In particular, copper (Cu) and nickel (Ni) appear to be released when Mn(IV) is reduced, and cadmium (Cd) when Fe(III) is reduced. These metals also diffuse toward the sediment-water interface, where they may accumulate in manganese nodules, readsorb on Mn(IV) and Fe(III) oxides in the sediments, or enter the water column.

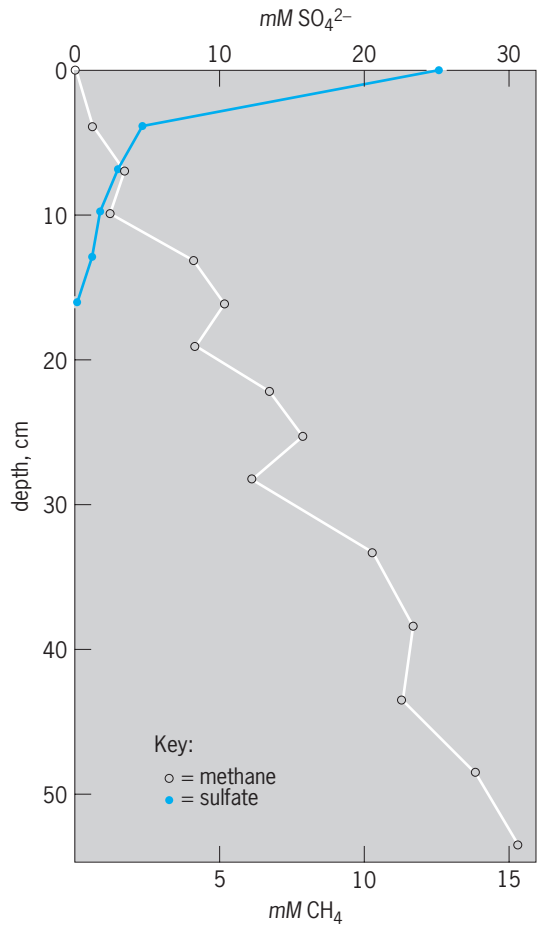

Fig. 7. Dissolved sulfate (SO_4^{2-}) and methane from extremely reducing sediments in Saanich Inlet. 1 cm = 0.4 in.

Sulfate reduction. Following Mn(IV) and Fe(III), sulfate (SO_4^-) is the energetically most favorable electron acceptor. Sulfate reduction is especially intense in rapidly accumulating sediments, rich in organic matter. Sulfate reduction has been intensively studied in the sediments of Long Island Sound and Saanich Inlet (British Columbia, Canada). Sulfate decreases rapidly in these sediments and in some cases is completely consumed by 5.9 in. (15 cm) [**Fig. 7**]. As sulfate is reduced, hydrogen sulfide ion (HS^-) accumulates in the sediments and diffuses toward the sediment–water interface. In some fiords with restricted circulation (such as Saanich Inlet), the HS^- that diffuses out of the sediments can accumulate in the bottom waters.

Following sulfate reduction, methane (CH_4) formation can proceed by two possible pathways: acetic acid fermentation [reaction (8)] and CO_2 reduction [reaction (9)]. The methane produced by these

$$CH_3COOH \longrightarrow CH_4 + CO_2 \qquad (8)$$

$$CO_2 + 8H_2 \longrightarrow CH_4 + 2H_2O \qquad (9)$$

reactions will also diffuse upward and may accumulate in the overlying water (Fig. 7).

Silicate and carbonate reactions. It has been suggested that reactions with silicates and carbonate control the interstitial water concentrations of sodium (Na^+), potassium (K^+), Ca^{2+}, and Mg^{2+} and also influence the alkalinity. Samples have shown that diagenesis leads to the uptake of Mg^{2+} and K^+ and the release of Ca^{2+}, bicarbonate ion (HCO_3^-), and Na^+ by the solid phases. The exact nature of the reactions is uncertain, but reverse weathering reactions may play a role. The calculated fluxes across the sediment–water interface are of the same order of magnitude as river inputs for these ions. In the case of Mg^{2+} and K^+, 60–100% of the river input to the ocean can be balanced by diagenetic uptake in the sediment. Ca^{2+} and HCO_3^- additions to seawater augment the river supply by 25 to 50%. James W. Murray

Carbonate and opal dissolution. Skeletal remains of calcareous planktonic organisms such as foraminifera, coccoliths, and pteropods, and of siliceous planktonic organisms such as radiolarians and diatoms, dissolve partially or entirely in ocean water as they settle through the water column or after they are deposited on the ocean floor. This dissolution, along with that of the remains of benthic shell-secreting organisms, constitutes a major chemical feedback mechanism in the material cycle of calcium, carbon, and silica in the oceans.

Distribution. The concentration of $CaCO_3$ (calcite and aragonite) in recent pelagic sediments (excluding those near the continents and islands) have been measured in various regions of the world oceans as a function of water depth (**Fig. 8**). In shallower depths, down to about 13,000 ft (4000 m), it is nearly constant at about 90 wt % (as $CaCO_3$), and gradually decreases to 5 wt % or less at a water depth of 15,000–16,000 ft (4500–5000 m). The water depth at which the $CaCO_3$ content starts to decrease has been called the sedimentary lysocline, and that at which the $CaCO_3$ content is reduced to 5% or less is called the carbonate compensation depth (CCD). When morphological changes of the tests of certain biological species are used to identify the beginning of $CaCO_3$ loss, the water depth that marks the first appearance of such changes is called the

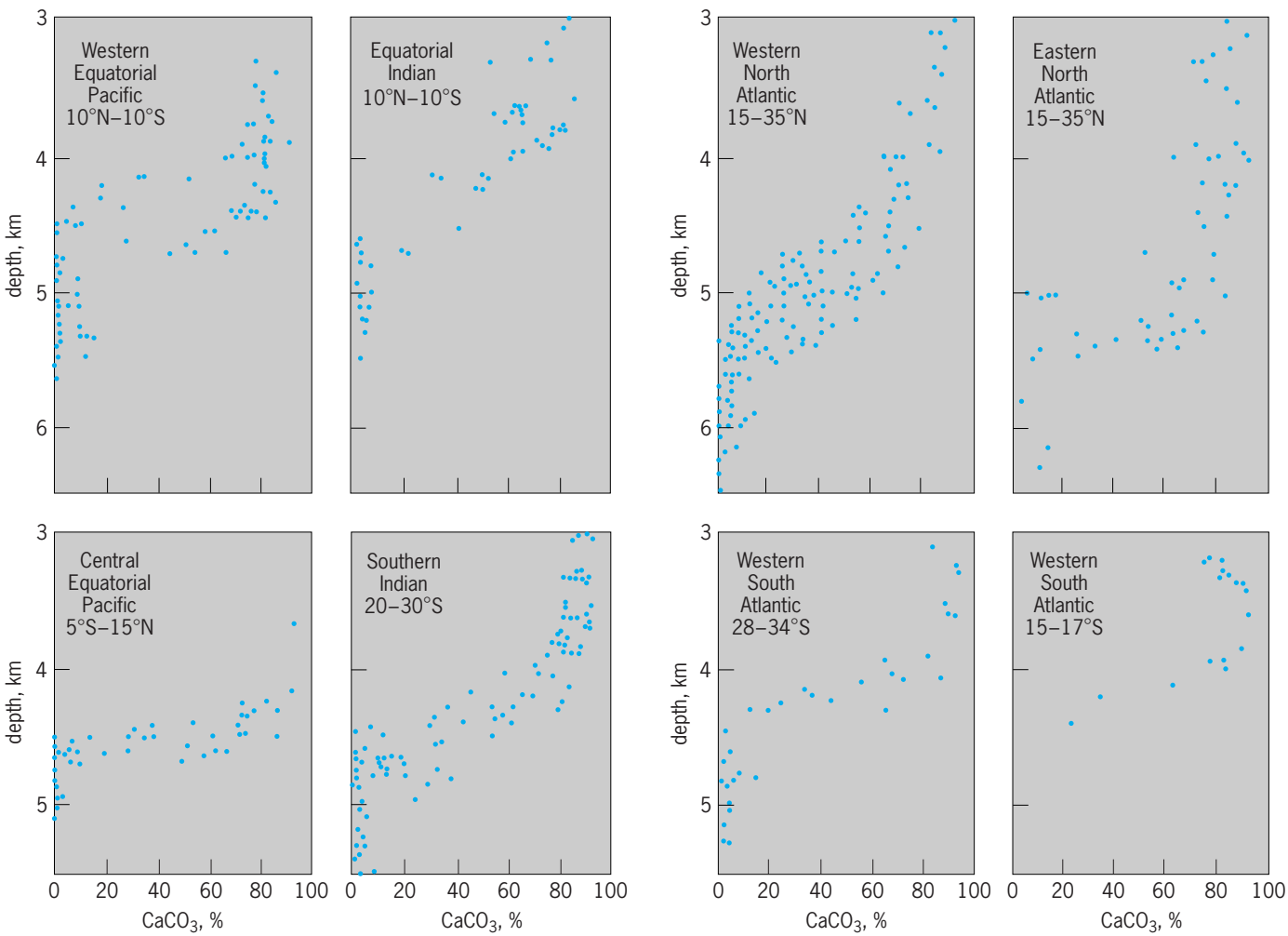

Fig. 8. $CaCO_3$ concentration in recent deep-sea sediments in eight regions of the world oceans as a function of water depths. Both aragonite and calcite are included as $CaCO_3$. 1 km = 0.6 mi. (*After N. Andersen and A. Malahoff, eds., The Fate of Fossil Fuel Carbon Dioxide in the Oceans, Plenum Press, 1977*)

foraminifera or pteropod lysocline according to the organisms used for its identification. Such distinctions are necessary since the lysocline identified on the basis of the $CaCO_3$ content in sediments does not necessarily coincide with that determined on the basis of morphological observations. Such a discrepancy may be attributed to the dissolution of tests that occurs in the guts of some feeders or in the sediments.

In contrast to the clear dependence of the $CaCO_3$ distribution on the water depth, the distribution of opaline silica does not depend upon the water depth. Instead, it appears to correlate well with the biological productivity of silica-secreting organisms in the surface water, which, in turn, depends on the intensity of the vertical circulation and the upwelling of deep waters rich in silica and other nutrient salts. Thus, sediments rich in siliceous tests are found in the belt around the Antarctica continent, in the northernmost Pacific, and in the equatorial zone of the world oceans.

Solubility. The observed decrease in $CaCO_3$ content in pelagic sediments with increasing water depths can be mainly attributed to the undersaturation of the deep waters with respect to calcium carbonate minerals, aragonite (pteropod shells), and calcite (foraminifera and coccolith tests). This undersaturation is caused by increased solubilities of these minerals with increasing pressure and decreasing temperature, and the lower pH. The smaller concentrations of CO_3^{2-} ion in deep waters are therefore caused by the greater concentration of CO_2 that is produced by microbial oxidation of the falling organic debris in the water column.

The concentration of calcium ion in ocean water is nearly constant (10.26 m*M*/kg), and the solubility of aragonite and calcite is expressed in terms of the CO_3^{2-} concentration in seawater. Pressure (or the water depth) affects the solubilities of calcite and aragonite (**Fig. 9**). Waters in the upper layer of the oceans down to about 1600 ft (500 m) in the Pacific and 9800 ft (3000 m) in the Atlantic are supersaturated with respect to aragonite, and below these water depths the water is undersaturated and becomes corrosive to aragonite. Therefore, the aragonitic tests of such organisms as pteropods are dissolved in seawater and are not found in sediments located at water depths greater than several hundred meters in the Pacific and 9800 ft (3000 m) in the Atlantic, although aragonitic organisms appear to be as equally abundant as calcitic ones in plankton tow samples collected in surface waters. Deep ocean waters below 13,000 ft (4000 m) in the Pacific and 15,000 ft (4500 m) in the Atlantic are undersaturated with respect to calcite, the less soluble form of $CaCO_3$. This accounts for the decrease in $CaCO_3$ content in deep-sea sediments. The rate of calcite dissolution increases as a third- to fourth-power function of the degree of saturation, and it increases rapidly with the water depth below the saturation level (see Fig. 9). At the carbonate compensation depth, the rate of deposition of calcite tests becomes equal to (or compensated by) the rate of dissolution, and thus little or no accumulation of $CaCO_3$ takes place below (**Table 2**).

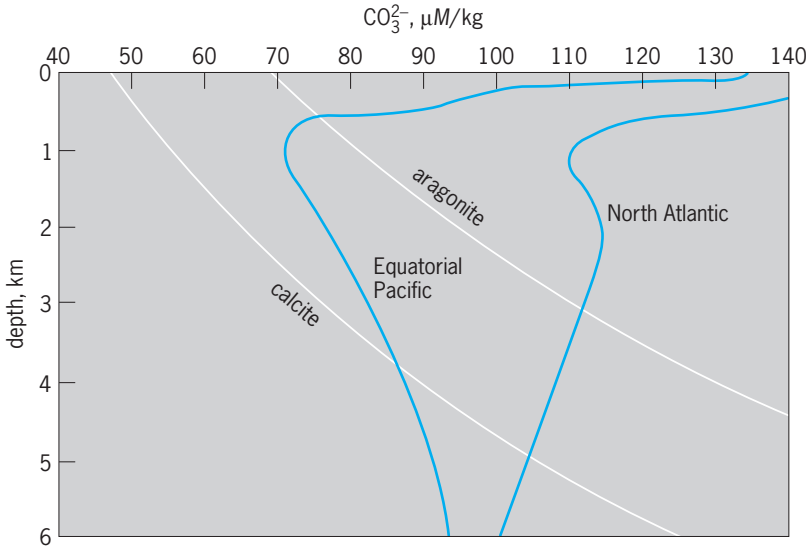

Fig. 9. Solubility of aragonite and calcite [expressed in terms of carbonate (CO_3^{2-}) ion concentration] in seawater at 2°C (36°F) as a function of water depths in the North Atlantic and the eastern Equatorial Pacific. The intersections of the solubility curves with the curves for the concentration of the CO_3^{2-} ion indicate the water depths at which the ocean water is saturated with the respective minerals. 1 km = 0.6 mi.

The solubility of opaline silica, the major constituent of siliceous tests, has not been determined exactly due to its variable stoichiometry, its slow reaction kinetics, and the presence of metastable silica species in seawater. However, it is in the range of 1.7 to 2.3 m*M*/kg at 25°C (77°F) and 1 atm (101 kilopascals) based on measurements made on amorphous silica and silica gel. The solubility increases slightly with increasing pressure. Since the concentration of silica dissolved in ocean water ranges from nearly zero in surface waters to about 0.17 m*M*/kg in the deep water of the North Pacific Ocean, the entire ocean is undersaturated with respect to opaline silica. Thus, the siliceous tests which form in the biologically active surface water dissolve during descent to the ocean floor and while they lie on the sea floor. It has been estimated that only about 4% of the siliceous tests formed in surface waters

TABLE 2. Estimated calcium carbonate balance in the Atlantic and Pacific oceans

Processes	Rates, $CaCO_3$ mg/(cm^2)(yr)	
	Atlantic	Pacific
Production of calcitic tests	1.3	0.9
Production of aragonitic tests	1.3	0.9
Total biogenic calcareous test production (aragonite + calcite)	2.6	1.8
Dissolution of calcareous tests in the oceans (aragonite + calcite)	1.9	1.5
Accumulation of calcareous tests on the ocean floor	0.7	0.3

TABLE 3. Silica budget in the oceans

Processes	Rates (10^{13} g/yr)
Supplies from rivers	43
Addition from alteration of basalts	8
Diffusion out of sediments	57
Burial of silica tests in sediments	104
Loss by inorganic silica adsorption	4
Biological fixation of silica	2500
Dissolution of tests in water column	2155
Dissolution of tests on sea floor	240

escape dissolution and become incorporated permanently into sediments (**Table 3**). Taro Takahashi

Organic geochemistry. Knowledge of the geochemistry of organic matter in sediments is important to the understanding of the global cycle of carbon in both ancient and modern times. Specific topics of importance are the origin of oil and gas, the climatic effects of perturbing the carbon cycle (for example, the CO_2 problem), contemporary sediments as temporary or permanent sinks for pollutant organic compounds (for example, pesticides and petrochemicals), and organic matter in ancient sediments as molecular paleontological markers and even as indicators of the origin of life on Earth.

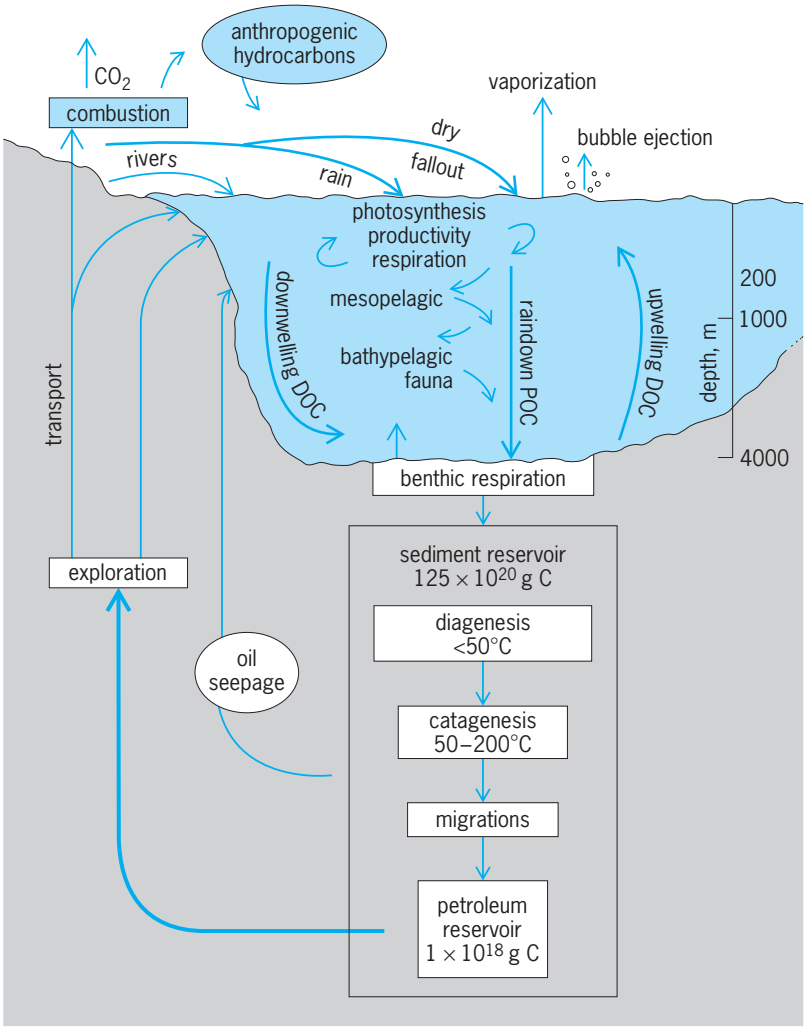

Fig. 10. Global organic carbon cycles. DOC = dissolved organic carbon; POC = particulate organic carbon. °F = (°C × 1.8) + 32; 1 m = 3.3 ft.

Sources of organic matter. In the contemporary marine environment, the carbon cycle involves certain mechanisms (**Fig. 10**). Organic matter in sediments is derived primarily from photosynthesis in the upper layers of the oceans or on land. A portion of this material escapes remineralization processes on land and in the upper layers of the ocean, and is transported by a variety of processes to sediments. Chemosynthesis by bacteria- and mineral-catalyzed reactions are very minor sources of organic matter synthesis.

Organic matter deposited to surface sediments undergoes a complex series of biotic and mineral-catalyzed reactions known as diagenesis (**Fig. 11**). Biologically synthesized polymers, carbohydrates, proteins, and lipids (fats) are broken apart to yield their constituent smaller molecules. Other organic polymers produced by biological systems such as chitinous skeletons of zooplankton and lignin from plants on land are less susceptible to breakdown to their constituent smaller molecules. Bacterial metabolism in sediments contributes other low-molecular-weight organic compounds such as phenols and aldehydes. It is thought that this complex assemblage of low-molecular-weight organic compounds and higher-molecular-weight biopolymers such as chitin and lignin react to produce humic and fulvic materials similar to humic and fulvic materials in soils. The exact mechanisms, rates of reaction, and even the structures of the humic and fulvic acid type of materials are not known. Some evidence is available to support the general scheme (Fig. 11). *See* DIAGENESIS.

Oil and gas formation. Microbial activity decreases and the relative importance of chemical reactions increases as sediments accumulate and the organic matter is buried deeper (Fig. 11). Over periods of thousands to millions of years the humic and fulvic type of materials plus other organic matter are converted to kerogen, itself an ill-defined polymeric structure (Fig. 11). If conditions of heat input, clay catalysis, kerogen composition, and geological structures are appropriate, oil can be formed. Other, slightly different conditions can yield gas. Sometimes the process does not go to completion because of a change in geological conditions, for example, uplifting near the Earth's surface after millions of years of burial. This is exemplified by the oil shales in Colorado and Wyoming. Humans will complete the process of oil formation by heating the shale to drive out oil already trapped and to convert more of the kerogen to oil. *See* KEROGEN; OIL SHALE; PETROLEUM.

Cycle of organic matter. Knowledge of the source of organic matter in sediments is of importance in elucidating the cycle of organic matter in the oceans and sediments. Measurements of stable isotope ratios of carbon to $^{13}C/^{12}C$ have been used extensively for determining the relative importance of land plants and marine plants as sources of organic matter found in

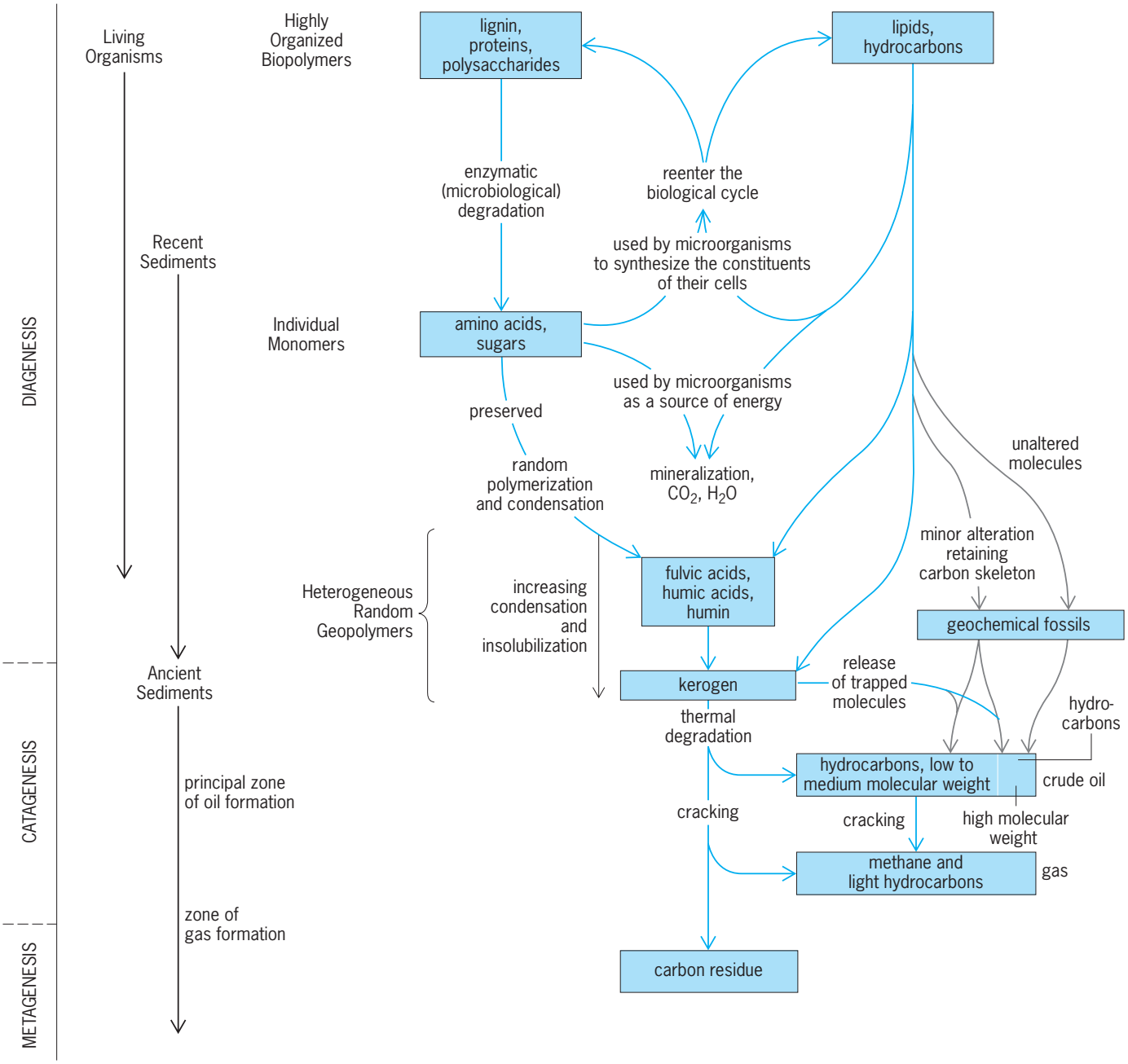

Fig. 11. Processes of change in marine sediments. (*After B. P. Tissot and D. H. Welte, Petroleum Formation and Occurrence, Springer-Verlag, 1978*)

sediments. These ratios are conventionally expressed as shown in Eq. (10). [The standard commonly used

$$\delta^{13}\text{C‰} = \frac{(^{13}\text{C}/^{12}\text{C sample}: {}^{13}\text{C}/^{12}\text{C standard}) \times 1000}{^{13}\text{C}/^{12}\text{C standard}} \quad (10)$$

is the Peedee Belemnite, or PDB—a particular carbonate rock formation.] Terrestrial plant organic matter is usually 5 to 10‰ lighter than marine plant organic matter. This is the result of the differences in the δ^{13}C value of atmospheric carbon dioxide, the source of land plant carbon, and the carbon dioxide–carbonate–bicarbonate source of carbon in marine plants. The differences between land-derived organic matter and marine-derived organic matter become less marked as diagenesis in young sediments and catagenesis in ancient sediments proceeds. Isotope fraction can accompany biochemical reactions in early diagenesis. Biogenic methane is produced with a δ^{13}C of −30 to −50‰ , that is, enriched in the light ^{12}C isotope, while carbon dioxide produced by the same assemblage of microorganisms is enriched in ^{13}C. Different classes of organic matter, for example, lipids, proteins, or carbohydrates, generally have different ^{13}C values. The former class is usually enriched in ^{12}C relative to the latter two classes. Even within a given molecule such as an amino acid, the carbon atom of the carboxyl group has a different ^{13}C value than the other carbon atoms. By carefully

studying the carbon isotope ratios, some of the biological and geological processes which have acted on organic matter in a given sample can be inferred. The radioactive isotope ^{14}C is used to date organic matter in samples ranging from contemporary to 50,000 to 75,000 years old. *See* RADIOCARBON DATING.

Organic structures. The detailed chemical structure of organic matter in sediments is poorly known relative to knowledge of the chemical structure of organic matter in organisms. Various estimates of 10 to 30% are given for that portion of organic matter in seawater or sediments for which information on chemical structures is available since fulvic acid, humic materials, and kerogen are very complex assemblages of organic matter. Nevertheless, sophisticated, sensitive, highly selective analytical chemical tools such as high-resolution gas chromatography–mass spectrometry of the organic compounds in that 10 to 30% of the organic matter has provided a wealth of information about the sources of the organic matter and the processes acting on it prior, during, and after deposition in sediments. Biosynthetic pathways in organisms lead to some very specific structural configurations for the molecules. This structural specificity was a major clue that led to the conclusion that oil in ancient sediments had a biological origin (Fig. 11). Several compounds serve as markers of a particular source of organic matter input, for example, bacterial fatty acids, land-plant *n*-alkanes, land-plant lignin, marine phytoplankton fatty acids and sterols, and amino acids from bacteria. *See* BIOSYNTHESIS.

Geochemical transformations. During diagenesis, several types of geochemical transformations of structures can take place. Some, such as the amino acid racemization, have fairly predictable rates if the thermal history of a sample is known, and thus the ratios of particular isomers give an indication of the age of the sample or at least of the organic matter in the sample. Other reactions convert biological compounds such as sterols and fatty acids to hydrocarbons—aromatic hydrocarbons, steranes (cycloalkanes derived from sterols), and alkanes found in petroleum. *See* AMINO ACID DATING.

Much more research is necessary to unravel the wealth of information about processes active in modern and ancient sedimentary environments that is contained collectively and individually in the many thousands of organic compounds found in ancient and modern sediments. A major advance in analytical chemistry allowed scientists to combine stable isotope measurement with specific compound analyses to determine the isotope ratios of individual compounds in a wide range of biological and geological samples. This approach has provided powerful new information about the carbon cycle and about the processes that yield oil and gas. *See* ORGANIC GEOCHEMISTRY.

John W. Farrington

Transport and Rate of Sedimentation

Rivers, glaciers, wind, and ocean waves and currents carry particles from continents and continental margins to the various environments of deposition in the ocean. The rates of sedimentation can be determined by study of the micropaleontology, radioactivity, or paleomagnetism of the sediments.

Transport of sediments. Most of the sediment eroded from continents enters the oceans at its margins, though some is delivered to the open sea by the wind or by icebergs. Much of this material, and some of the pelagic sediment produced by the life processes of marine plankton, is redistributed by currents within the ocean basins. *See* NEARSHORE PROCESSES.

Turbidity currents. Density flows that derive their energy from high concentrations of suspended sediment are efficient but episodic agents of downslope sediment transport. Their most important role is dispersing sediment from continental margins to abyssal plains and deep-sea fans, but they also may be effective in locally redistributing pelagic sediment in mid-ocean areas of high relief. Turbidity currents are thought to have velocities of several meters per second, especially where they are channelized in submarine canyons and fan valleys, and are competent to transport sediment grains of up to coarse gravel size. Larger blocks of sediment move downslope in debris flows.

Thermohaline currents. Branches of the general oceanic circulation that are driven by pressure differences caused by temperature and salinity distributions transport large amounts of fine sediment in some parts of the deep ocean. They are slower (generally less than 20 in./s or 50 cm/s) but much steadier than turbidity currents, and carry sediment both in suspension, as a so-called bottom nepheloid layer, and as bed load (**Fig. 12**). Because these currents typically flow along rather than across bottom contours, geologists often refer to them as contour currents. They may transport fine suspended clay for hundreds of kilometers along the basin margins even where current velocities are low (<4 in./s or 10 cm/s); bed-load transport, for most of the cohesionless particles available in the deep sea (for example, foraminiferal shells) requires speeds of about 8 in./s (20 cm/s). Where thermohaline currents are locally accelerated, as in constricted passages

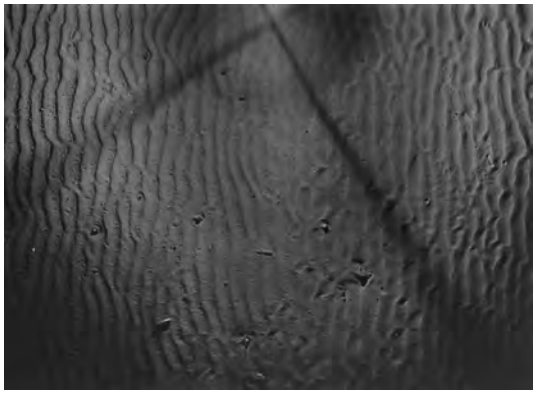

Fig. 12. Rippled sea floor of sand being transported along the bed of a thermohaline current. From a water depth of 8900 ft (2700 m) on the northwest side of the Rockall Plateau, North Atlantic.

between basins, they may attain speeds high enough for effective erosion of their beds.

Tidal currents. These currents are fast enough to entrain and transport silt and sand in some parts of the ocean, even at abyssal depths. Their effects are best known from shallow continental shelves, where alternating ebb and flood currents may build sand ridges and other bedforms. Oscillating sea-bed currents capable of moving sediment are also caused by the passage of surface waves and internal waves. Geological effects of internal waves are poorly known, but may be important at all water depths; normal wind-driven waves have no effect below a depth of about 660 ft (200 m), but tsunamis (seismic waves) can infrequently stir up sediment several kilometers below the ocean surface. *See* TSUNAMI.

Significance. The importance of sediment transport in the ocean is that the thickness of sediment on any part of the floor reflects past and present current regimes, not just crustal age and location of terrigenous and biogenous sediment sources. Some styles of erosion, transport, and deposition produce distinctive types of sediment (for example, turbidites and contourites) and distinctive topographic forms (for example, ripples and mud waves). *See* OCEAN CIRCULATION; TIDALITES; TURBIDITE. P. Lonsdale

Rates of sedimentation. The accumulation rates of sediments in the deep sea are determined either directly or indirectly through the use of natural radioactive nuclides. There are three sources of these nuclides: members of the decay series of the two uranium isotopes ^{235}U and ^{238}U dissolved in seawater; radioactive nuclides, such as carbon-14 (^{14}C), silicon-32 (^{32}Si), beryllium-10 (^{10}Be), and aluminum-26 (^{26}Al), produced in the atmosphere by cosmic rays and delivered to the ocean; and long-lived radionuclides, in particular potassium-40 (^{40}K) and ^{238}U, found in volcanic mineral and glass ejecta from explosive volcanoes (**Table 4**).

In addition to the direct use of radioactive nuclides, the magnetic reversal history of the Earth, as chronicled by the potassium-argon dating of basaltic lava flows on land, is commonly recorded in accumulating deep-sea sediments. So long as the pattern is preserved as a continuous record, sedimentary layers can be dated from the magnetic orientation pattern of the successive layers of sediment. *See* PALEOMAGNETISM; ROCK AGE DETERMINATION.

Sediment type and region. The sediment accumulation rates in the oceans vary depending on sediment type and region. Sediments dominated by clay-size grains are delivered to the ocean surface by streams (passing through estuaries) or winds and accumulate at a rate of about 0.1 in. (0.3 cm)/1000 years. Foraminiferan and coccolith-rich sediments (sometimes called ooze) accumulate at a rate of about 0.8 in. (2 cm)/1000 years. Diatom- and radiolarian-rich sediments also accumulate at these high rates. Detrital sediments supplied from the continental margins to the deep sea by turbidity currents and other forms of bottom transport accumulate in abyssal plains and associated areas at an average rate of about 2 in. (5 cm)/1000 years with a wide range around this value. As a crude estimate, the average rate of sediment accumulation in the deep sea is about 0.4–0.8 in. (1 to 2 cm)/1000 years. Additional sediments accumulate on the continental margins, and the two environments are the repositories of material derived from continental denudation.

Sediment activity. Sediments, once deposited on the sea floor, are hardly inactive. They are subject to biological and physical mixing at the interface, chemical dissolution and mobilization, and perturbation by slumping and scouring by bottom processes such as currents and sediment avalanches. These can affect both the rates of sediment accumulation as compared to the rate of sedimentation and the estimation of rates of accumulation using radionuclides.

Bioturbation. Bioturbation can be identified and its rate determined by the distribution in a sediment core of natural short-lived [lead-210 (^{210}Pb); half-life = 22 years] nuclides sequestered from the water column (**Fig. 13**) or recently introduced radionuclides (plutonium) produced by human activities. The depth of bioturbation is indicated by the effective homogenization of the concentration of the long-lived diagnostic radionuclides used for dating (such as ^{14}C; **Fig. 14**). The rates and depths of bioturbation vary throughout the deep ocean basin, but the average diffusionlike constant is about 8×10^{-8} in.^{2}/s (5×10^{-9} cm^2/s) and the average depth of mixing about 3 in. (8 cm).

Time scales of stability. Sediments can be lost or gained episodically to and from surrounding areas over a range of time scales. The bottoms of fracture zones commonly receive cascades of carbonate-rich sediments from the cliffs above. The stochastics of sediment perturbation on the ocean floor have not yet been studied adequately. Results based on thorium-230 (^{230}Th), ^{14}C, and ^{10}Be indicate time scales of stability ranging from less than 100,000 years (along the Mid-Atlantic Ridge) to 3,000,000 years (on the North Pacific ocean floor). The frequency of postdepositional sediment mobilization by episodic events will influence the geologic record in deep-sea sediments. The understanding

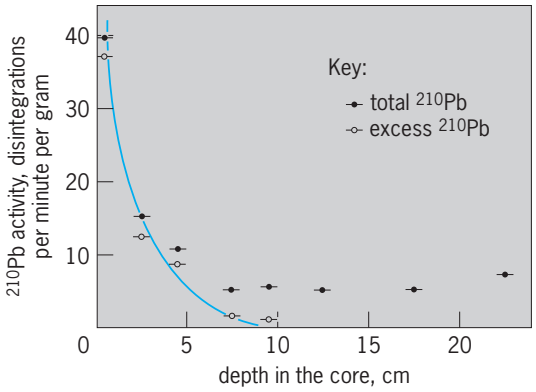

Fig. 13. Lead-210 (^{210}Pb) distribution in a sediment core (527-3) obtained by the research submersible *Alvin* from the *Famous* site (36°48.5′N, 33°15.3′W) on the Mid-Atlantic Ridge. The fit of the excess ^{210}Pb [that is, unsupported by radium-226 (^{226}Ra)] distribution to a diffusion equation yields a constant of 6×10 cm^2/s. 1 cm = 0.4 in. (*After Y. Nozaki et al., Earth Planet Sci. Lett., 34:167, 1977*)

TABLE 4. Radionuclides used in the chronology of deep-sea sediments

Radionuclides	Half-life, years	Useful range, years	Types of sediment
1. Uranium-decay-series nuclides			
Uranium-234 (from ^{238}U)	250,000	0–400,000	All
Thorium-230 (from ^{230}U)	74,500	0–300,000	All
Protactinium-231 (from ^{235}U)	32,000	0–120,000	All
2. Cosmic-ray-produced			
Silicon-32	277 (possibly 100)	0–2,000	Diatom and radiolarian deposits
Carbon-14	5,720	0–40,000	1. Principally calcareous (foram and coccolith deposits) 2. Organic fraction of deposits
Aluminum-26	0.72×10^6	$0–4 \times 10^6$	All
Beryllium-10	1.5×10^6	$0–8 \times 10^6$	All
3. Long-lived from volcanoes on land or tektites			
Potassium-40 (decay to ^{40}Ar)	1.25×10^{16}	>60,000	Feldspars, volcanic glass, tektites
Uranium-238 (fission tracks)	1.01×10^{16} (spontaneous fission)	$>10^6$	Volcanic glass, tektites
4. Magnetic reversal	—	$0–5 \times 10^6$	All with magnetic minerals

of the time scales of stability of sediment piles is also important for engineering uses of the deep-sea floor, such as platform construction and burial of radioactive wastes. Karl K. Turekian

Climatic Record in Sediments

Undisturbed sediments that accumulate at the ocean floor preserve important information regarding past processes in the overlying water column, on adjacent continents, in the atmosphere, as well as on the Earth as a whole. Sediments for climate-change studies are recovered by a variety of coring methods, from piston coring to drilling. The drilling vessel *JOIDES Resolution*, using advanced piston-coring and drilling techniques, has recovered continuous sequences of sediments extending to 100 million years ago (Ma). Typical pelagic sediments accumulate at rates of centimeters every 1000 years, although at some locations pelagic and hemipelagic sediments accumulate much more rapidly, preserving high-resolution records of climate change. Both slowly and rapidly accumulating marine sediments provide important information regarding past patterns of oceanographic and climatic change. Studies of the climate record preserved in marine sediments are collectively referred to as paleoceanography or paleoclimatology. *See* PALEOCEANOGRAPHY; PALEOCLIMATOLOGY.

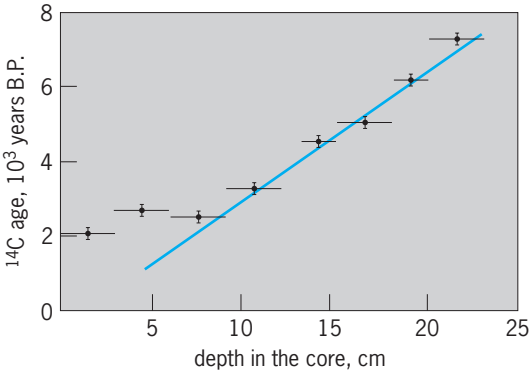

Fig. 14. Distribution of ^{14}C ages with depth in *Famous* core 527-3. The top 3.5 in. (9 cm) yield a constant age of 2400 years before present (B.P.) because of bioturbation. The deeper points define a line yielding a sediment accumulation rate of 2.9 cm/10^3 years. 1 cm = 0.4 in. (*After Y. Nozaki et al., Earth Planet. Sci. Lett., 34:167, 1977*)

Sediment indicators of climatic change. Information from marine sediments regarding past climate change can be inferred from measurements made from marine sediments. These measurements act as indicators, or proxies, of past climate. Proxy indicators for paleoceanographic processes are derived from the lithologic and biotic composition of sediments and from geochemical indices preserved in sedimentary components. Geochemical and isotopic proxies for climate processes are often derived from measurements made by using the microfossil

Problems	Successes
Assume either constant nuclide flux or invariant initial specific activity over time.	^{234}U: Not generally successful in sediment dating by itself but important adjunct to other decay-series nuclides. ^{230}Th: Established glacial climatic chronology of the last 300,000 years from fossiliferous deep-sea sediments in conjunction with uranium-decay-series chronometric studies of fossil hermatypic corals. ^{231}Pa: In conjunction with ^{230}Th, helps distinguish dilution versus decay in sediments.
Time scale comparable to bioturbation time scale. Useful for accumulation rate studies only in fast accumulating (> 0.8 in. or 2 cm/1000 years) sediments.	Established possibility of fast rates of siliceous sediment accumulation around Antarctica and in Gulf of California over hundreds of years.
Assumes constancy of initial $^{14}C/^{12}C$ ratio. Assumes no detrital organic matter. Also needs high organic content in sediment.	Established breaks in sediment accumulation rates of various components over the Pleistocene-Holocene boundary. Fast rates of accumulation in Argentine Basin and abyssal plains determined from organic matter.
Very low activity. Difficult to measure. Assume either constant flux or invariant initial specific activity over time.	Established that ^{26}Al primarily produced in atmosphere. Awaits future uses.
Low activity but successfully measured. Assume either constant nuclide flux or invariant initial specific activity over time.	Showed accumulation rates in North Pacific effectively constant over 3,000,000 years. Confirms magnetic stratigraphy. Also useful for establishing breaks in record due to episodic sediment loss or long-term nondeposition.
Suitable volcanic materials occur only at ocean basin margins close to subduction zones. Record more subject to destruction with increasing age due to subduction.	Useful for calibrating magnetic stratigraphy. Also dating beyond 1,000,000 years.
Depends on preservation of a binary signal without distortion. Signal lost where magnetic minerals absent.	Where applicable, establishes long-term faunal chronologies.

skeletons of foraminifera, a calcareous organism. Although the utility, understanding, and independence of different proxies vary, a number of proxies are available to study different climatic processes. Paleoceanographic and paleoclimatic processes that are studied include past glaciations, surface ocean characteristics, oceanic production and nutrient cycles, deep-ocean circulation, atmospheric circulation, and continental weathering. *See* MICROPALEONTOLOGY.

Ice volume. The ratio of the stable isotopes oxygen-18 (^{18}O) and oxygen-16 (^{16}O) in skeletons of foraminifera records the isotopic composition of the seawater in which the foraminifera lived. The seawater $^{18}O/^{16}O$ ratio is related to local water temperature and to the amount of ^{18}O-depleted water stored as ice on land, a global effect. The $^{18}O/^{16}O$ ratio preserved in foraminifera is dominated by the ice-volume effect and is only secondarily affected by temperature and other factors. Evidence for this is the global similarity of planktonic (surface-dwelling) and benthic (deep-dwelling) foraminiferal $^{18}O/^{16}O$ ratios as well as the similarity of the $^{18}O/^{16}O$ record to sea-level reconstructions. The $^{18}O/^{16}O$ patterns recorded globally in marine sediments are a key indicator of the cyclicity of glaciations of the late Neogene (26–0 Ma; late Oligocene to present). It has also enabled development of a common time scale for marine sediments.

Surface ocean conditions. Estimates of sea surface temperature and surface ocean characteristics such as oceanic upwelling are provided by several proxies. Modern distributions of microfaunal and microfloral assemblages are used to infer past surface ocean conditions. Primary faunal indicators are planktonic foraminifera and radiolaria, and floral indicators are coccolithophorids and diatoms. Foraminifera and coccolithophorids are composed of calcite, while radiolaria and diatoms are composed of opal. Statistical analyses of these microfossil assemblages are used to reconstruct past surface ocean conditions (**Fig. 15**). If the ice-volume component of the $^{18}O/^{16}O$ ratio and vital effects of planktonic foraminifera are accounted for, the planktonic $^{18}O/^{16}O$ ratio yields information about the temperature of the water in which the organisms lived. It is possible to determine oceanic mixed-layer structure from combined assemblage and isotopic compositions of planktonic foraminifera. The isotopic composition of organic matter preserved in marine sediments provides another independent estimate of past temperatures of the surface layer. *See* BACILLARIOPHYCEAE; COCCOLITHOPHORIDA; FORAMINIFERIDA; RADIOLARIA.

Oceanic productivity, CO_2, and nutrients. Multiple sediment proxies yield information regarding the past productivity of surface waters and past oceanic nutrient distributions. A frequently used proxy is the carbon

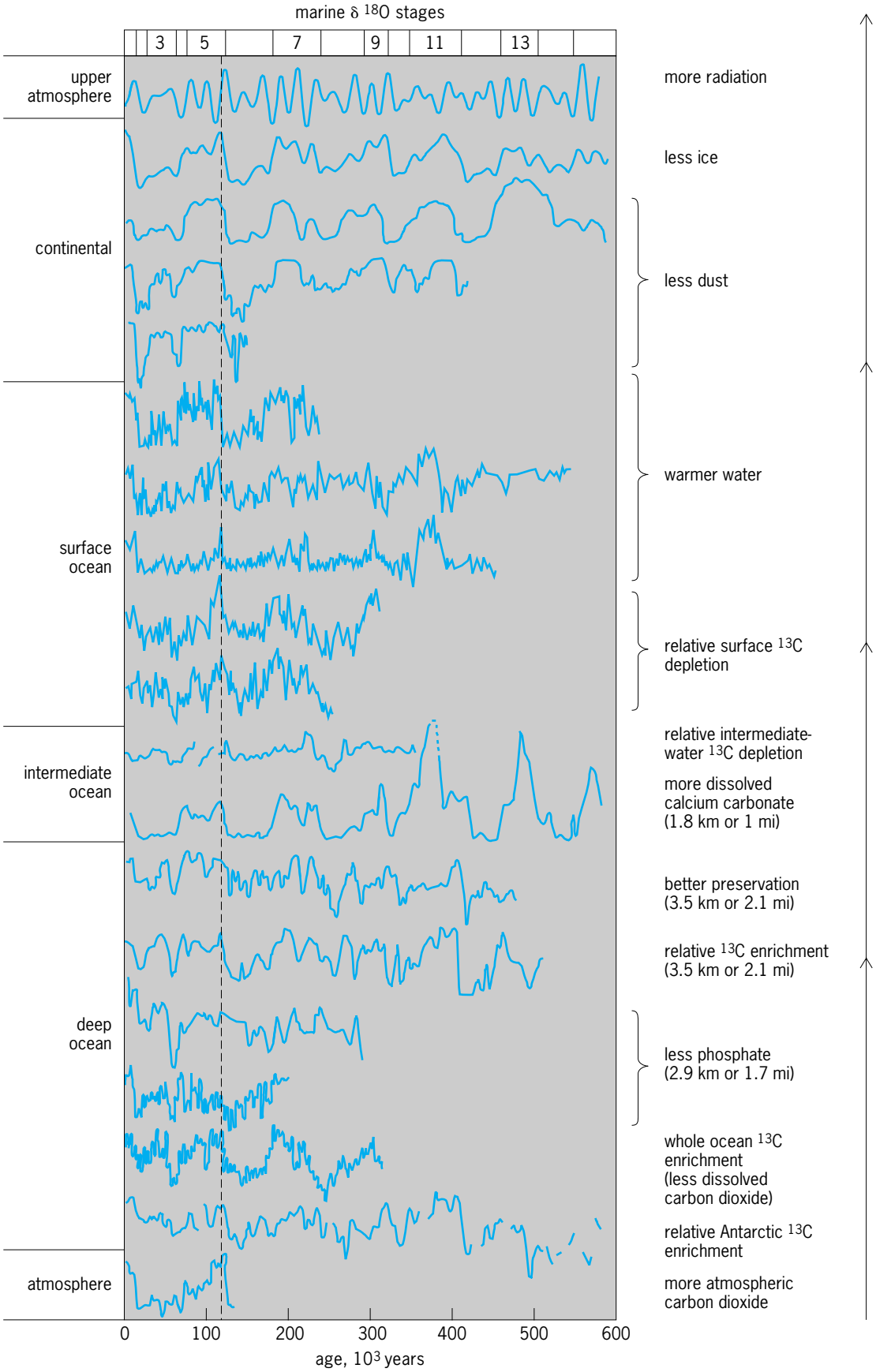

Fig. 15. Summary of major changes in the climate system over the past 800,000 years from a set of 18 proxy indicators. Coherent variations are evident in all components. (*After J. Imbrie et al., On the structure and origin of major glaciation cycles, 1. Linear responses to Milankovitch forcing, Paleoceanography, 7:701–1738, 1993*)

isotopic composition of skeletons of foraminifera. Photosynthetic processes preferentially incorporate ^{12}C and exclude ^{13}C from organic matter; thus changes in oceanic carbon reservoirs are reflected in $^{13}C/^{12}C$ ratios from foraminifera. The $^{13}C/^{12}C$ in benthic and planktonic foraminifera from marine sediments have been used to infer paleoproductivity. Other indicators of past oceanic production are sedimentary organic carbon content, microfossil assemblages, bulk calcite and opal accumulation, and stable nitrogen isotope ratios in organic carbon. Oceanic productivity strongly affects dissolved carbon dioxide (CO_2) levels in the ocean, and hence CO_2 levels in the atmosphere. Past levels of atmospheric CO_2 have been inferred from carbon isotope gradients in planktonic and benthic foraminifera and from specific components of sedimentary organic material. Sedimentary proxies for past nutrient distributions include the above indicators as well as trace-element measurements from fossil foraminifera.

Deep-water circulation. As deep water masses slowly circulate through the world's oceans, their chemistry changes as the waters age. Past patterns and rates of deep-ocean circulation can be inferred from past chemistry changes. Comparison of stable carbon isotope ratios ($^{13}C/^{12}C$) between planktonic and benthic foraminifera, and the geographical variation in these ratios, is a means for reconstructing variations in deep-ocean chemistry and circulation.

Atmospheric circulation. Away from continents and from ocean-bottom thermal vents, the primary source of nonbiogenic materials to the deep sea is wind-borne deposition of terrigenous material or hydrothermal deposition. Terrigenous sediment flux, terrigenous grain size, and the mineralogy of terrigenous sediments provide clues to the strength and location of winds over the ocean, continental aridity, and source regions of terrigenous material.

Weathering fluxes to the ocean. Continental weathering rates are a product of both orogeny (tectonic uplift) and climatic changes. Information regarding past weathering is given by geochemical proxies in marine sediments that indicate the discharge of elements from the continents to the ocean. Proxies for weathering fluxes include ratios of strontium-87/strontium-86 ($^{87}Sr/^{86}Sr$) and germanium/silicon (Ge/Si) in biogenic sediments, accumulation rates of biogenic and terrigenous sediments, and the mineralogical composition of terrigenous sediments. *See* OROGENY.

Climate since the last ice age. On continental margins and in sediment drift deposits where marine sediments accumulate rapidly, detailed studies of climate changes associated with the last ice age are possible. Sediments at these locations fall within the range of radiocarbon dating and can be closely tied to continental and ice-core climate records. Paleoclimate proxies from these sediments indicate that climate has been highly variable, punctuated by rapid cooling events in surface waters, large changes in deep-water circulation, and major iceberg discharges on time scales of centuries to thousands of years. In some marginal basins, where oxygen levels in bottom waters are low, sediments are undisturbed by burrowing organisms and are often laminated. In these specialized settings, the sediments preserve climate records on annual-to-decadal time scales. Climate changes associated with the Little Ice Age and El Niño–Southern Oscillation phenomena have been recorded. *See* EL NIÑO.

Pleistocene climates. The climatic record preserved in marine sediments of the past 2 Ma is dominated by ice ages. Measurements of climate proxies from a wide geographic distribution of sediment cores has documented the timing, amplitude, and spatial sequence of glacial and interglacial cycles (Fig. 15). At the forefront of this effort was the establishment of a consistent time scale for correlations between different sediment cores. This was accomplished by using the global ice-volume signal preserved in oxygen isotope ratios from foraminifera, combined with radiometrically dated magnetic reversals. The ice-volume proxy records indicate that over the past 1 Ma glacial intervals occur cyclically at periods of 100,000 years, 41,000 years, 23,000 years, and 19,000 years. Statistical analyses have demonstrated a strong connection between these major climate cycles and changes in the amount of insolation received by the Earth because of changes in the Earth's orbit. This is the foundation for the astronomical theory of climate change, also known as the Milankovitch hypothesis. Acceptance of this theory led to improvements in the Pleistocene time scale. Variations in every component of the climate system occur sequentially at these astronomically driven periods (Fig. 15). Previous to 1 Ma, the dominance of 100,000-year cycles was diminished, but strong variations in ice volume and other climatic proxies occurred at cycles of 41,000 years, 23,000 years, and 19,000 years. The patterns and amplitudes of Pleistocene (1.88–0 Ma) climatic change can be to first order modeled as a linear response of the climate system to orbitally driven insolation changes. *See* CLIMATE HISTORY; GLACIAL EPOCH; INSOLATION.

Pre-Pleistocene climates. A continuous sequence of marine sediments from the Cretaceous (~136–~66 Ma) through the Holocene (10,000 years ago to present) has been recovered through advanced drilling efforts. The undisturbed nature of some sediment sequences has facilitated extension of the astronomically calibrated time scale. The climatic record preserved in pre-Pleistocene sediments documents a progressive cooling of global climate beginning in the Eocene (57.8–36.6 Ma). It is hypothesized that this cooling is ultimately caused by tectonic boundary condition changes, including uplift of the Tibetan plateau during the past 40 Ma, as weathering associated with these tectonic changes impacts global geochemical cycles. Superimposed on this cooling trend are cycles of warming and cooling both before and after the onset of Northern Hemisphere glaciation at about 2.4 Ma. Despite large boundary condition changes, variations in insolation reaching the Earth have influenced climate variability throughout the Cenozoic (~66–0 Ma). Several periods of abrupt climatic transitions have also occurred during

the Cenozoic. Rapid warming of surface and deep waters occurred in the latest Paleocene, ~55 Ma, whereas global cooling and Antarctic glaciation occurred during the early Oligocene, ~35 Ma. Broad climate studies are focused on understanding the evolution, sensitivity, and variability of climate systems from the beginning of the Cenozoic through today. *See* CENOZOIC; GEOLOGIC THERMOMETRY; GEOLOGIC TIME SCALE. Teresa King Hagelberg

Bibliography. W. S. Broecker and T. H. Peng, *Tracers in the Sea*, 1982; H. Charnock and J. Edmond (eds.), *The Deep Sea Bed: Its Physics, Chemistry, and Biology*, 1991; R. Chester, *Marine Geochemistry*, 2d ed., 2003; R. M. Cline and J. D. Hays (eds.), *Investigation of Late Quaternary Paleoceanography and Paleoclimatology*, Geol. Soc. Amer. Mem. 145, 1976; W. W. Hay (ed.), *Studies in Paleo-Oceanography*, Soc. Econ. Paleontol. Mineralog. Spec. Publ. 20, 1974; B. C. Heezen and C. D. Hollister, *The Face of the Deep*, 1971; J. M. Hunt, *Petroleum Geochemistry and Geology*, 2d ed., 1995; J. Imbrie et al., On the structure and origin of major glaciation cycles, 1. Linear responses to Milankovitch forcing, *Paleoceanography*, 7:701–738, 1992; J. Imbrie and K. P. Imbrie, *Ice Ages: Solving the Mystery*, 1986; J. Kennett, *Marine Geology*, 1982; F. J. Millero and M. L. Sohn, *Chemical Oceanography*, 3d ed., 2005; H. C. Nelson and T. H. Nilsen, *Modern and Ancient Deep Sea Fan Sedimentation*, 1984; Open University Course Team Staff (ed.), *Ocean Chemistry and Deep-Sea Sediments*, vol. 5, 1989; K. E. Peters et al., *The Biomaker Guide*, 2d ed., 2005; M. E. Raymo and W. F. Ruddiman, Tectonic forcing of late Cenozoic climate, *Nature*, 359:117–122, 1992; J. P. Riley (ed.), *Chemical Oceanography*, vol. 9, 1989; B. P. Tissot and D. H. Welte, *Petroleum Formation and Occurrence*, 2d ed., 1984; J. C. Zachos et al., Abrupt climate change and transient climates during the Paleogene: A marine perspective, *J. Geol.*, 101:191–213, 1993.

Maritime meteorology

Those aspects of meteorology that occur over, or are influenced by, ocean areas. Maritime meteorology serves the practical needs of surface and air navigation over the oceans. Phenomena such as heavy weather, high seas, tropical storms, fog, ice accretion, sea ice, and icebergs are especially important because they seriously threaten the safety of ships and personnel. The weather and ocean conditions near the air–ocean interface are also influenced by the atmospheric planetary boundary layer, the ocean mixed layer, and ocean fronts and eddies.

Meteorological phenomena. These include heavy weather, tropical storms, fog, ice accretion, and the atmospheric planetary boundary layer.

Heavy weather. Stormy weather at sea outside the tropics is primarily associated with cyclones that form in certain geographical regions where cold, dry air of continental origin and warm, moist air of tropical origin come together during winter (**Fig. 1**). These extratropical storms are especially dangerous to ships at sea because of their relatively high frequency of occurrence and the fact that they can develop extremely rapidly, reaching full force in less than 24 h. A common storm of this type in high latitudes is known as a polar low. Polar lows are relatively small in size (300–500 km or 180–300 mi in diameter) and quite shallow, extending from the sea surface to a height of only 1 or 2 km (0.6 or 1.2 mi). Sudden, explosive storm development (that is, cyclogenesis) is also very common off the east coast of North America in winter. A mature east coast storm is much larger (1000 km or 600 mi in diameter) and deeper (8–10 km or 5–6 mi deep or more) than a polar low. Mature extratropical storms that bring most of the heavy weather at sea include gale and hurricane force winds, heavy rains, and high seas.

The fact that such powerful and dangerous marine storms can develop so rapidly presents an especially challenging forecasting problem for meteorologists. The explosive cyclogenesis that occurs off the east coast of North America takes place during winter over the warm Gulf Stream waters east of Cape Cod and south of Nova Scotia. The conditions that are most favorable for such explosive cyclogenesis include a strong, horizontal temperature contrast in the lower atmosphere between the cold continent and the warm Gulf Stream; low vertical stability of the air due to the movement of cold, dry continental air out over the warm water; and the release of latent heat of condensation in the ascent of warm, moist air ahead of the developing storm. Since these conditions are usually prevalent throughout the winter, meteorologists consider the movement of an upper-level (midtropospheric) disturbance over the region favorable for cyclone development to be a key ingredient to triggering the growth of such rapidly developing storms. *See* CYCLONE; STORM.

Tropical storms. Cyclones (that is, storms) with very warm, moist air in the center develop in low latitudes over the warmest parts of the tropical oceans (**Fig. 2**). These violent and dangerous storms obtain their energy from the latent heat of condensation that is released in deep, convective rain clouds that form in unstable tropical air. They form when this deep convection becomes organized by the inward-spiraling low-level winds associated with an initially weak tropical disturbance. *See* CLOUD PHYSICS.

The winds in a mature tropical storm in the Northern Hemisphere spiral inward and counterclockwise (clockwise in the Southern Hemisphere) in the lower atmosphere and outward and clockwise (counterclockwise in the Southern Hemisphere) aloft. In intense storms, known as hurricanes in the North Atlantic, Gulf of Mexico, and eastern North Pacific, and typhoons in the western North Pacific, the winds can reach up to 320 km/h (200 mi/h) or more, with a central calm area or eye roughly 25–80 km (15–48 mi) in diameter. These storms generally move from east to west in the trade wind belt, but they occasionally curve poleward into middle latitudes, making forecasting difficult. When a tropical storm

Fig. 1. Cloud patterns in developing extratropical cyclones. (*a*) Photograph (100 mm) taken during a United States space shuttle mission, showing the typical comma-shaped cloud pattern associated with a developing storm off the coast of Nova Scotia (55°N, 50°W). (*b*) Satellite infrared image of a rapidly developing storm in the eastern Atlantic Ocean. The speckled clouds to the northwest of the storm center are seen to originate as thin cloud lines farther upwind to the northwest of the storm (upper left). Such cloud lines are indicative of very cold air that is being drawn into the storm from the northwest by the storm's own cyclonic circulation. (*C. H. Wash, Naval Postgraduate School*)

moves inland, the source of low-level moisture is greatly reduced, and the storm usually weakens or dissipates—however, not without damage and sometimes loss of life due to the strong winds, heavy rains, and flooding storm tides. With modern technology, tropical storms are easily located and tracked by meteorological satellites. However, their movement can be very difficult to predict because of the strong interaction that occurs between the storm and the larger-scale winds in which the storm is embedded. Also, the reason why so few tropical disturbances actually develop into full-fledged hurricanes is still one

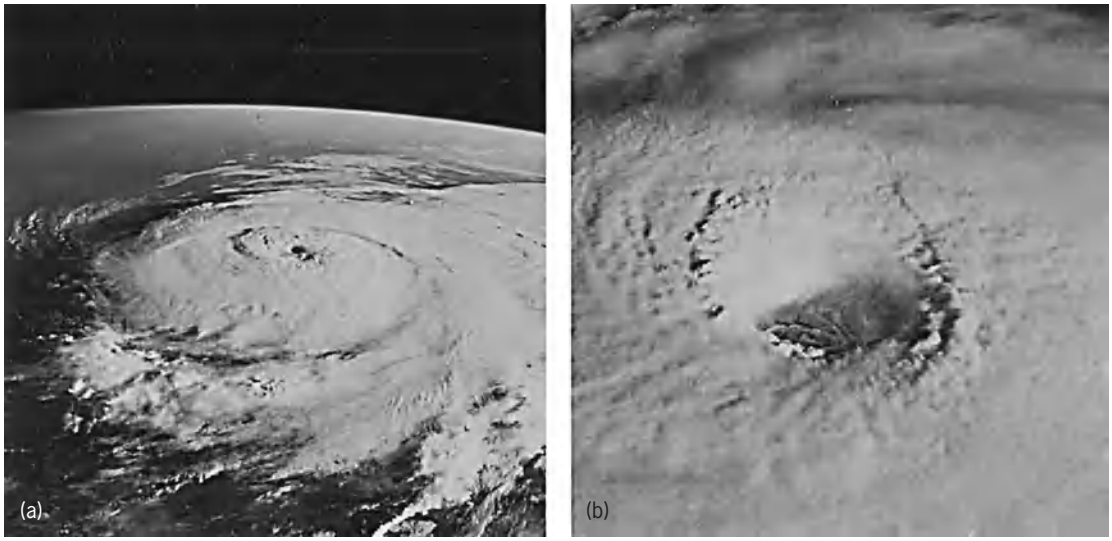

Fig. 2. Photographs (50 mm) of Hurricane Elena in the Gulf of Mexico on September 1, 1988, taken during a United States space shuttle mission. The low viewing angle and sun angle reveal the different cloud structures and heights in this mature hurricane. (*a*) Inward-spiraling cloud bands. (*b*) Close-up of the eye of the storm. (*C. H. Wash, Naval Postgraduate School*)

of the most important unanswered questions in meteorology. *See* HURRICANE; METEOROLOGICAL SATELLITES; SATELLITE METEOROLOGY; STORM DETECTION; TROPICAL METEOROLOGY.

Fog. Fog is important to all seagoing activities because it restricts visibility, making navigation (and aircraft operations from navy aircraft carriers) extremely hazardous. Certain types of marine fog can be very thick and persistent. Steam fog is produced when cold air flows over a warmer water surface. Evaporation of the relatively warm water into the overlying, cooler air, followed by condensation into minute liquid droplets, produces the fog. In high northern latitudes, steam fog is referred to as arctic sea smoke, and it is very common in the marginal ice zone in winter and early spring (**Fig. 3**). Advection fog occurs when relatively warm, moist air flows over cooler water, cooling the air to saturation. Extensive areas of advection fog can develop when warm air flows poleward over colder water at higher latitudes or when such air flows eastward over the cold, coastal ocean currents that usually exist in midlatitudes off the west coast of continents. Advection fog in such coastal regions is especially pronounced in the summer, when the normally cold ocean currents are made even colder by upwelling of cold water near the coast. *See* FOG; UPWELLING.

Ice accretion. Ice accretion, which may occur on ships as well as aircraft, not only decreases the efficiency of operation but also may be extremely hazardous, especially to small vessels. The accumulation of ice reduces freeboard and, more importantly, the stability of the ship by making it top-heavy. Ice accretion affects all exposed equipment on ships, including missile launchers and gun systems on navy ships and radar and radio antennas on all ships. Heavy ice accretion can render such equipment completely inoperable. Ice formation occurs primarily from freezing rain and freezing spray from the sea. The most dangerous accumulations of ice develop with freezing spray, which can occur when the air temperature is several degrees or more below freezing. The colder the air and the sea and the stronger the wind, the more rapid is the accumulation of ice. As much as 45 metric tons (50 tons) or more of ice can accumulate in 24 h depending on the size of the ship, which creates a serious hazard to the ship and crew.

Atmospheric planetary boundary layer. Since the atmosphere is semitransparent to incoming solar radiation, about 50% of the Sun's energy passes directly through the atmosphere to be absorbed at the surface. [Some of the energy in the other 50% is reflected (albedo).] This heat energy is transferred to the atmosphere in the form of sensible and latent heat by conduction and by long-wave radiation, although a large part of the long-wave radiation from the Earth's surface is lost to space and therefore does not affect the atmosphere. The lowest kilometer or so of the atmosphere, in which the air and ocean (or Earth) exchange heat, moisture, and momentum, is known as the atmospheric planetary boundary layer (**Fig. 4**). Except for the lowest millimeter above the sea, the air flow in the atmospheric planetary boundary layer is fully turbulent. This means that the air flow is unsteady, that is, it fluctuates rapidly in space and time; fluid properties mix, that is, they are readily transported from one place to another within the boundary layer; and the forces of inertia dominate those of molecular viscosity. The turbulent nature of the boundary layer greatly increases the vertical transport and the exchange of heat, moisture, and momentum between the atmosphere and the ocean. *See* HEAT BALANCE, TERRESTRIAL ATMOSPHERIC; TURBULENT FLOW.

The atmospheric planetary boundary layer consists of two distinct regions that are of great scientific interest and practical importance to maritime meteorology (Fig. 4). In the lowest 50 m (165 ft) or so, in what is called the surface layer, the vertical

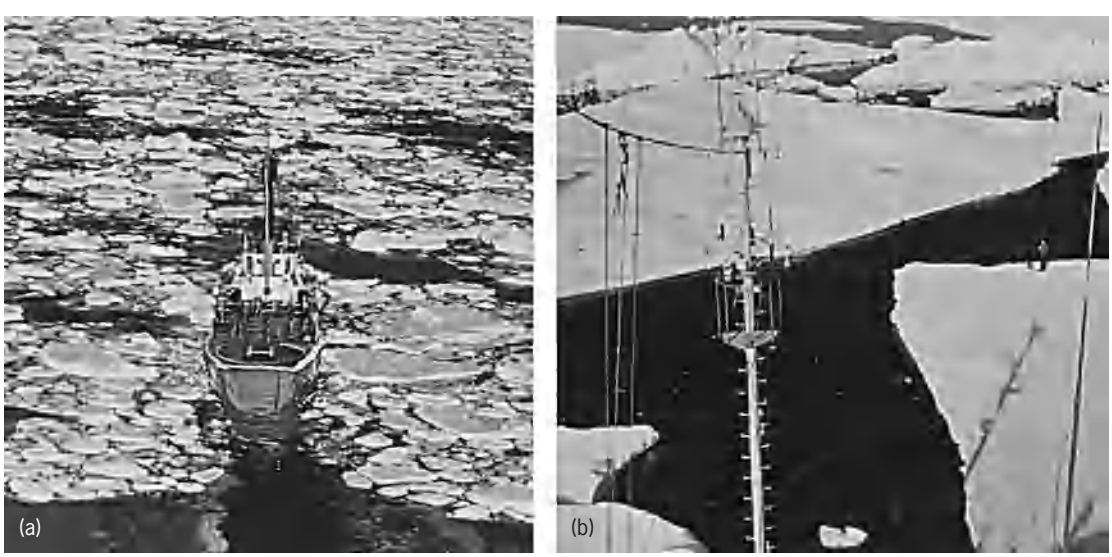

Fig. 3. Sea ice in the marginal ice zone. (*a*) Relatively small cakes of thin ice near the seaward edge of the marginal ice zone. (*b*) Much larger and thicker floes farther back from the edge. (*P. S. Guest, Naval Postgraduate School*)

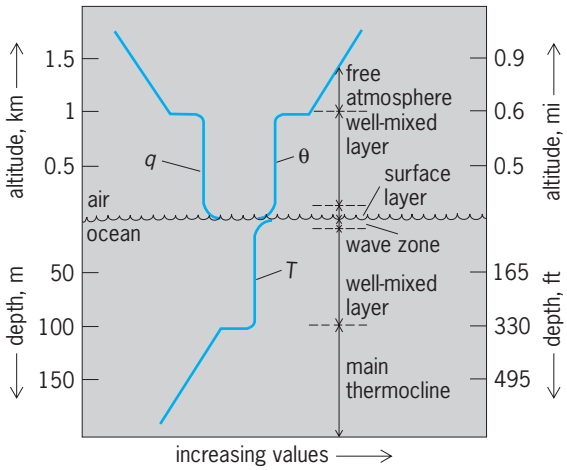

Fig. 4. Schematic drawing of the atmospheric planetary boundary layer and the ocean mixed layer. In the atmosphere, θ is potential temperature (temperature adjusted for the compressibility of the atmosphere), and q is specific humidity (ratio of the mass of water vapor to the mass of air in a unit volume of air). In the ocean, T is temperature. The quantity θ generally increases with height in the atmosphere, while T decreases with depth in the ocean. The wind speed and ocean currents are not shown, but an idealized profile of the winds and currents would tend to follow the θ and T lines, respectively.

turbulent transport of heat, moisture, and momentum is essentially constant in the vertical. As a result, the mean temperature, humidity, and horizontal velocity vary approximately logarithmically with height above the sea surface. Above the surface layer, and extending throughout the rest of the boundary layer, the turbulence is so strong that the mean temperature, moisture, and momentum themselves are well mixed and nearly uniform with height. If the air in the boundary layer is sufficiently moist, clouds (mostly stratocumulus) form in the upper part of the boundary layer, and these clouds sometimes drop to the sea surface as fog. When clouds are present in the boundary layer, both the temperature and the turbulence are affected. Whether the boundary layer is cloudy or not, it usually is capped by a rather thin inversion above which the air is warmer, drier, and much less turbulent. This region above the boundary layer is called the free atmosphere, meaning that the flow there is essentially free of turbulence and largely free from the effects of the underlying surface of the land or ocean. *See* ATMOSPHERE.

Oceanic phenomena. These include high seas, sea ice and icebergs, the mixed layer, and fronts and eddies.

High seas. Winds blowing over the sea generate waves and ocean currents. Waves generated by the wind vary in height, length, and period, depending on the wind speed, the length of water over which the wind is blowing in a more or less uniform direction (fetch), and the duration of time the wind has blown. The winds in a mature extratropical cyclone over the ocean can generate waves 10–15 m (33–49 ft) high. Such high seas can cause great damage to ships, harbors, and beaches, and they produce a large amount of ambient noise (sound) in the ocean that hinders naval antisubmarine warfare operations. Because of the importance of high seas to all seagoing activities, oceanographers have developed numerical wave prediction models to forecast the state of the sea. These models include such processes as wave growth, propagation, and dissipation. Their success depends on an accurate analysis and forecast of the surface wind field. Such wave forecasts are used by private companies and government agencies for various purposes, including merchant marine operations, navy tactics and ship maneuvers, refueling at sea, amphibious landing exercises, oil drilling and oil spill operations, and search and rescue. *See* OCEAN WAVES.

Sea ice and icebergs. Sea ice, formed by the freezing of seawater and floating at the surface of the ocean, is of keen scientific interest and importance to marine meteorology and oceanography. Sea ice plays a key role in air–sea interaction. Since it reflects much of the solar radiation reaching the surface, sea ice greatly reduces the sensible and latent heat exchange between the atmosphere and the ocean, and it severely alters the wind stress acting on the water. At high latitudes, the open and ice-covered regions of the ocean meet in a narrow zone, known as the marginal ice zone (Fig. 3). Environmental parameters of importance in the marginal ice zone include the location, concentration, and thickness of sea ice, and, for naval operations, the small-scale roughness of the top and bottom surfaces of the ice. The ability to predict these parameters depends on a full understanding of how the ocean and the ice exert drag on one another, and how they exchange heat and mass through melting and freezing. Since sea ice rejects most of the salt in the seawater as it freezes, the melting and freezing of sea ice changes the salinity, and hence the density and vertical stability, of the sea immediately below the ice. This, in turn, modifies the dynamic and thermodynamic interaction between the ice and the underlying ocean. In the marginal ice zone, sea ice often drifts, or is episodically forced by the wind, into water that is above freezing, resulting in a rapid melting of the ice. Icebergs, masses of land ice that have broken away from land and are floating on the sea, are of less scientific interest to marine meteorology, but they present a threat to the safety of some ships. They are readily identified and routinely tracked by satellites. *See* ICEBERG; SEA ICE; WIND STRESS.

Mixed layer. Over most of the ocean at low and middle latitudes, solar radiation that penetrates the atmosphere is absorbed in the upper ocean, forming a warm, stable, surface layer (Fig. 4). When the wind blows over the surface of the ocean, or when the surface of the ocean is cooled by heat loss to the atmosphere, turbulent motions are generated which mix the warmed and stably stratified surface waters down into the cooler, denser water below. This one-sided mixing process is called entrainment, and it leads to the formation of a well-mixed upper layer in which the temperature is essentially uniform and at the bottom of which there is a sharp temperature drop (density increase). The situation is analogous to the stable inversion that is formed at the top

of the atmospheric planetary boundary layer. The only difference is that in the atmospheric boundary layer, active entrainment advances upward from below while in the ocean it works downward from above.

When the winds decrease, turbulence in the upper ocean quickly decays, and active entrainment ceases. The absorption of solar radiation along with lower levels of turbulence then can cause a new mixed layer to form at a shallower depth. Over much of the world's ocean, pronounced deepening and shallowing of the mixed layer takes place over an annual cycle, with warm, shallow mixed layers in summer and cold, deep mixed layers in winter. A less pronounced daily cycle of mixed layer deepening at night and shallowing during the day also occurs. Both of these cycles are greatly modified by episodes of strong winds and high surface-heat losses associated with tropical and extratropical cyclones over the ocean. Because turbulence in the upper ocean is so strongly dependent on the local atmospheric forcing, the vertical density structure of the upper ocean preserves a fairly accurate record of the wind and weather of the recent past, modified of course by such factors as horizontal and vertical transports within the upper ocean. Fluctuations in the depth of the mixed layer, and in the magnitude of the temperature drop at the base of the mixed layer, are especially important to many tactical naval operations.

Fronts and eddies. Maritime meteorology is concerned about fronts and eddies in the ocean because of their effect on the overlying atmosphere and because they are associated with significant fluctuations in ocean currents, temperature, salinity, and other ocean variables. These fluctuations have horizontal scales of 10–500 km (6–300 mi), vertical scales of 1–2 km (0.6–1.2 mi), and time scales of several days to several weeks or more. Ocean fronts and eddies are commonly referred to as oceanic weather, because of their dynamic and kinematic similarity to atmospheric weather.

Ocean fronts and eddies come in a wide variety of types. The most significant ones are meanders or cut-off eddies (cold and warm core rings) in major ocean current systems such as the north wall of the Gulf Stream or other strong western boundary current systems. Also, there are eddies in the open ocean driven directly by the wind, eddies generated by the interaction of ocean currents with bottom topography, and fronts in the coastal zones caused by upwelling or by cooling of the shallow water on the continental shelf. The meandering of major western boundary currents and the subsequent pinching-off of eddies and rings occur through mixed baroclinic and barotropic instability in the way that midlatitude cyclones develop and occlude in the atmosphere, but the exact mechanisms responsible for these ocean developments are not completely understood. Variations in ocean depth can influence the direction of propagation of preexisting waves and eddies, and cause eddies to be generated locally when water is forced to flow up or down a topographic slope, and they can significantly modify the dynamic stability of major current systems. Oceanic weather associated with upwelling fronts in coastal regions often takes the form of long filaments or dipole (counterrotating) eddies that extend seaward for several hundred kilometers. The origin and dynamics of many of these features are not completely understood, but they are the subject of active research. *See* FRONT; OCEAN CIRCULATION.

Analysis and forecasting. This involves observations and numerical prediction.

Observations. To support the analysis and forecasting of many meteorological and oceanographic elements over the globe, observations are needed from a depth of roughly 1 km (0.6 mi) in the ocean to a height of 30 km (18 mi) in the atmosphere. In addition, the observations must be plentiful enough in space and time to keep track of the major features of interest, that is, tropical and extratropical weather systems in the atmosphere and fronts and eddies in the ocean. Over populated land areas, there is a fairly dense meteorological network; however, over oceans and uninhabited lands, meteorological observations are scarce and expensive to make, except over the major sea lanes and air routes. Direct observations in the ocean, especially below the sea surface, are insufficient to make a synoptic analysis of the ocean except in very limited regions. Fortunately, remotely sensed data from meteorological and oceanographic satellites are helping to fill in some of these gaps in data. Satellite data can provide useful information on the type and height of clouds, the temperature and humidity structure in the atmosphere, wind velocity at cloud level and at the sea surface, the ocean surface temperature, the height of the sea, and the location of sea ice. Although satellite-borne sensors cannot penetrate below the sea surface, the height of the sea can be used to infer useful information about the density structure of the ocean interior. *See* REMOTE SENSING.

Numerical prediction. The motion of the atmosphere and the ocean is governed by the laws of fluid dynamics and thermodynamics. These laws can be expressed in terms of mathematical equations that can be put on a computer in the form of a numerical model and used to help analyze the present state of the fluid system and to forecast its future state. This is the science of numerical prediction, and it plays a very central role in marine meteorology and physical oceanography.

The first step in numerical prediction is known as data assimilation. This is the procedure by which observations are combined with the most recent numerical prediction valid at the time that the observations are taken. This combination produces an analysis of the present state of the atmosphere and ocean that is better than can be obtained from the observations alone. Data assimilation with a numerical model increases the value of a piece of data, because it spreads the influence of the data in space and time in a dynamically consistent way.

The second step is the numerical forecast itself, in which the model is integrated forward in time

to predict the state of the atmosphere and ocean at a future time. Models of the global atmosphere and world ocean, as well as regional models with higher spatial resolution covering limited geographical areas, are used for this purpose. In meteorology and oceanography the success of numerical prediction depends on collecting sufficient data to keep track of meteorological and oceanographic features of interest (including those in the earliest stages of development), having access to physically complete and accurate numerical models of the atmosphere and ocean, and having computer systems powerful enough to run the models and make timely forecasts. *See* WEATHER FORECASTING AND PREDICTION.

Applications. Two important applications of maritime meteorology are ship routing and tactical naval environmental support.

Ship routing. The headway made by a ship transiting the ocean depends to a great extent on the state of the sea, as well as the force of the winds. The larger the waves and the more directly into the waves, or winds, the ship is headed, the slower the headway. Hence, it is clearly desirable to avoid heavy seas and strong winds, not only to avoid damage and possible injuries but also to save time and fuel in ocean crossing. Empirical data collected from ships' logs have enabled researchers to relate ship speed to sea state and, from such relationships, to construct a least-time track for a ship crossing the ocean. This requires an accurate forecast of the weather along all possible routes the ship might take made several weeks in advance. Since this is generally beyond state-of-the-art forecast skill, preliminary ship tracks are recommended based on 5–7-day predictions and climatology; and then they are modified, if necessary, as the voyage progresses. Ship routing based on weather and sea conditions has been found to improve the efficiency and reduce the cost of shipping; consequently, it has been widely used for both commercial and naval ships of the United States and other countries. *See* SHIP ROUTING.

Tactical naval environmental support. Tactical environmental support refers to describing and predicting those meteorological and oceanographic phenomena that affect the performance of naval weapons systems on time and space scales that are important to tactical naval operations. This aspect of marine meteorology and oceanography is becoming increasingly important to the modern navy because of the increased sophistication and range of naval weapons systems, and the great sensitivity of those systems to variations in the environment. Thus, an accurate knowledge of present and future meteorological and oceanographic conditions can be a significant asset to the operating navy in its conduct of command, control and communication, antisubmarine and electronic warfare, navigation, search and rescue, and many other tactical operations at sea. *See* ANTISUBMARINE WARFARE; NAVIGATION.

The physical environment affects communications and weapon systems primarily by refracting, scattering, and attenuating electromagnetic waves (microwaves, optical, and radio waves) in the atmosphere and acoustic (sound) waves in the ocean. Atmospheric refraction is caused by vertical variations in temperature and humidity, primarily in the lower part of the atmosphere. Such variations can be especially significant in the first few tens of meters above the sea surface and in the sharp transition region between the atmospheric planetary boundary layer and the free atmosphere. As a result, the structure and variability of the atmospheric planetary boundary layer are extremely important to naval meteorology.

An important application of maritime meteorology to naval operations is the prediction of radar (or other electromagnetic or optical) wave propagation in the atmosphere and sound-wave propagation in the ocean (**Fig. 5**). Sound waves can travel great distances through the ocean, at a speed that depends on the temperature, pressure, and salinity

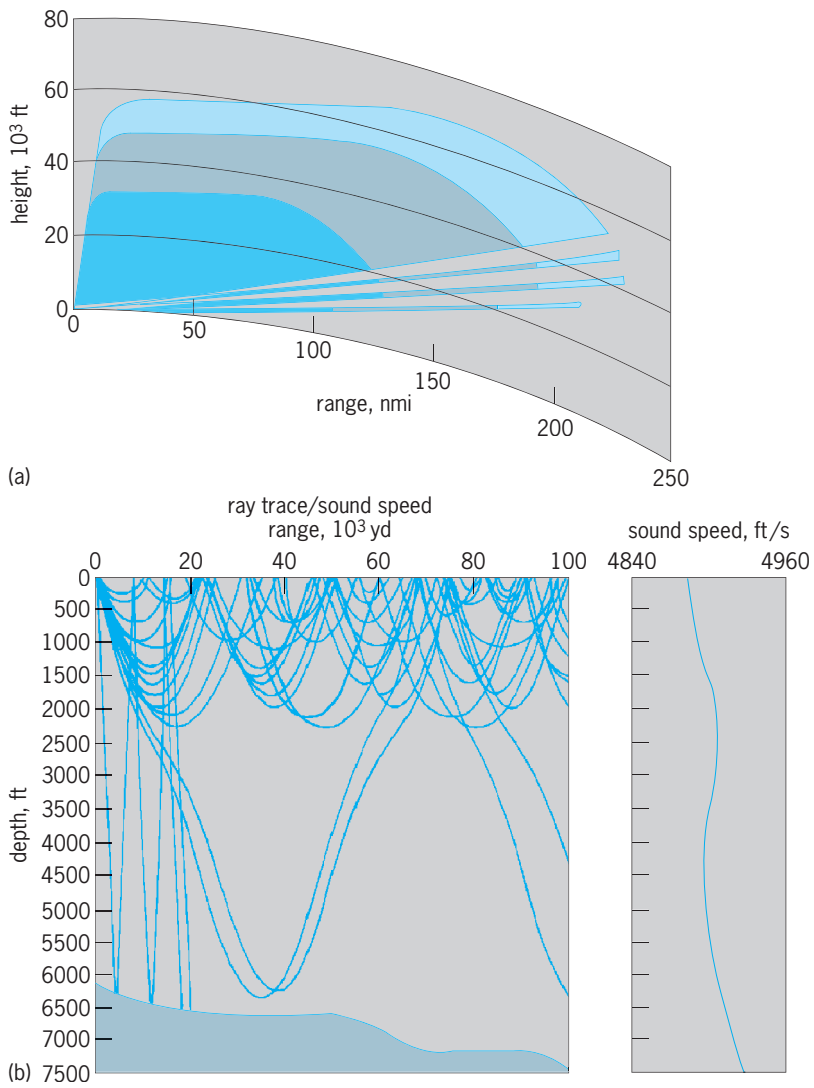

Fig. 5. Prediction of (*a*) radar (or other electromagnetic or optical) wave propagation in the atmosphere and (*b*) sound wave propagation in the ocean. 1 ft = 0.3 m. 1 nautical mile = 1852 m. (*LCDR W. T. "Kim" Curry, U.S. Navy*)

of the water through which the waves are propagating. As a result, variations in these quantities refract sound waves in the ocean in the same way that variations in temperature and humidity refract electromagnetic waves in the atmosphere. Because of refraction, the sound waves emitted from a given location in the ocean will travel along a curved path that depends on the ocean thermal structure. When the waves reach the top or bottom of the ocean, they are reflected into the ocean interior in ways that depend on the properties of the sea surface or sea floor. Because of the variable nature of these environmental factors, acoustic waves leaving a particular sound source will tend to converge into certain areas and diverge from others. The places where sound waves come together because of refraction are called convergence zones, and they represent optimal locations to "listen" for enemy ships. Likewise, places where sound waves tend to diverge from are called shadow zones, and they are good places to "hide" from enemy ships. The location and intensity of convergence zones and shadow zones are strongly affected by ocean fronts and eddies, by fine structure and turbulence in the upper ocean, and by the properties of the sea floor and sea surface. In very high latitudes, the sea ice concentration and the surface wave conditions strongly affect the ambient noise in the ocean, while the roughness of the underside of sea ice is important for acoustic propagation. The application of the science of maritime meteorology and oceanography to naval tactics is a very important aspect of most naval operations at sea. *See* METEOROLOGY; OCEANOGRAPHY; UNDERWATER SOUND.

Robert L. Haney

Bibliography. A. E. Gill, *Atmosphere-Ocean Dynamics*, 1982; G. J. Haltiner and R. T. Williams, *Numerical Prediction and Dynamic Meteorology*, 2d ed., 1988; E. B. Kraus and J. A. Businger, *Atmospheric-Ocean Interaction*, 2d ed., 1994; A. R. Robinson (ed.), *Eddies in Marine Science*, 1983; R. B. Stull, *An Introduction to Boundary Layer Meteorology*, 1988.

Marjoram

The aromatic herb *Majorana hortensis*, a common plant in Mediterranean areas. The spicy camphoraceous odor of marjoram has long been cherished as an addition to a wide variety of foods; in the Middle Ages marjoram was used as an air freshener.

Marjoram is in the mint family (Lamiaceae) and is a close relative of European or Greek oregano, with which it is often confused. There is still controversy concerning the proper taxonomic classification of this plant. Some authors place it in the genus *Origanum*, while others continue to separate it into its own genus *Majorana*. Many ecotypes or varieties of marjoram exist, both in wild populations and under cultivation. *See* OREGANO.

Marjoram is a small perennial (1–2 ft or 30–60 cm tall), and has ovate leaves up to 1 in. (2.5 cm) long. The leaves are slightly hairy, as are the erect somewhat woody stems. Marjoram flowers are white to very light lavender or pink in color but are very small and usually go unnoticed; the entire flower spike or inflorescence is, however, easily noticeable.

Marjoram is native to North Africa and southern Europe; its range stretches into southern Germany. It is cultivated in numerous locations throughout the world, including Europe, Egypt, Israel, Mexico, India, and the United States. Cultivation is best accomplished in a warm environment with a low relative humidity. Marjoram will tolerate a frost, but prolonged cold and dampness will weaken or kill it. Well-drained fertile soil produces the most vigorous plants. In the United States it is grown primarily in California due to the mild winters and low humidity needed for dehydration.

Marjoram is usually directly seeded into the field at a rate of 5–15 lb of seed per acre (2.3–6.8 kg/ha). It will grow enough the first year to be cut twice, and is usually cut three times a year in subsequent years. After 4 years, yields drop off, and it is usually plowed under and then replanted in a new location to avoid disease buildup in the soil. Harvesting is accomplished mechanically or by hand, after which the leaves are dried as quickly as possible at low temperatures (100°F or 37.7°C) to preserve color and flavor.

Because it is a perennial, marjoram is more likely to become infected with soil-borne diseases than annual spices such as basil and savory. Both Phytophora and Verticillium wilt sometimes infect this crop, with resulting dieback and death of individual stems that eventually spreads to whole plants and fields. *See* PLANT PATHOLOGY.

The dried and fresh leaves are used as flavoring for meats (sausage), vegetables, cheeses, poultry stuffing blends, and sauces, especially tomato-based sauces. *See* LAMIALES; SPICE AND FLAVORING.

Seth Kirby

Marl

In the most general use of this term, a sediment that consists of a mixture of calcium carbonate ($CaCO_3$) and any other constituents in varying proportions. Marls usually are fine-grained, so that most marls consist of $CaCO_3$ mixed with silt and clay. The dominant carbonate mineral in most marls is calcite ($CaCO_3$), but other carbonate minerals such as aragonite (another form of $CaCO_3$), dolomite [Ca, Mg $(CO_3)_2$], and siderite ($FeCO_3$) may be present also. *See* ARAGONITE; CALCITE; CARBONATE MINERALS; DOLOMITE; SIDERITE.

Frequently, particularly in North America, the name marl is limited to a lake deposit that is rich in $CaCO_3$, but usually the term is extended to include marine deposits. Deep-sea marls consist of mixtures of clay and the $CaCO_3$ skeletons of microscopic planktonic animals (foraminiferans) and plants (coccoliths and discoasters). Marls deposited in the deeper parts of lakes consist of fine-grained

$CaCO_3$, but marls deposited in shallow water may contain $CaCO_3$ in the form of shells of mollusks and fragments of calcareous algae such as *Chara* (stonewort) mixed with fine-grained $CaCO_3$ that is precipitated on the leaves of rooted aquatic vegetation. The $CaCO_3$ from these various sources is then mixed with sand, silt, or clay brought in by streams from the surrounding drainage basin. *See* MARINE SEDIMENTS.

Although the term marl has been applied to sediments with a greatly variable content of $CaCO_3$, strictly speaking the $CaCO_3$ content should range between 30 and 70%. Because of the high $CaCO_3$ content, most marls are light to medium gray, although they can be almost any color. The high $CaCO_3$ content also tends to make dried marl earthy and crumbly.

The indurated rock equivalent of marl is marlstone. Equivalent terms range from calcareous claystone to argillaceous or impure limestone, depending on the amount of $CaCO_3$ that is present. Marlstone is common in marine sequences of all ages, and is particularly common in ancient lake sequences of Tertiary age in the western United States. For example, the famous "oil shale" of the Green River Formation of Wyoming, Colorado, and Utah is not a shale at all but a marlstone in which the dominant carbonate mineral is dolomite. *See* LIMESTONE; OIL SHALE; SEDIMENTARY ROCKS.

Walter E. Dean

Marmot

A member of the genus *Marmota* in the squirrel family Sciuridae. Fourteen species occur mainly in mountainous areas in the northern hemisphere (see **table**). *See* RODENTIA; SQUIRREL.

General morphology. Marmots are large, heavy-bodied rodents attaining weights of 3–7.5 kg (6.6–16.5 lb) and total lengths of 400–820 mm (16–32.8 in.). Fur color varies considerably among the different species. The upper parts may be gray, brown, or yellowish. The underside, although paler, normally differs little in color from the upper parts, although in *M. flaviventris* the belly is yellow (see **illustration**). Some species have white markings. *Marmota vancouverensis* is melanistic. The texture of the fur varies from thin to long and thick. The head is short and broad, the legs are short, and the densely haired, slightly flattened tail is one-fifth to one-third of the total length. The eyes are small, and the ears are short, broad, rounded, and well haired. The feet have five digits, each terminating in a thick, slightly curved claw. The palms and soles are naked. The dental formula is I 1/1, C 0/0, Pm 2/1, M 3/3 × 2 for a total of 22 teeth.

Habitats and behavior. Marmots are found in open habitats such as pastures, forest edges, alpine meadows, and steppes. Most of their life is spent in burrows, which they excavate in well-drained soil. A mound of fresh soil at the entrance and numerous trails radiating to and from feeding areas is characteristic. Burrows may be as deep as 5–7 m (16–23 ft) and have several entrances. Tunnels measuring over 100 m (330 ft) have been recorded. Winter dens of *M. monax* are normally in wooded or brushy habitats on sloping ground and have only one entrance, whereas summer dens are most often in open areas and have several entrances. Marmots are primarily diurnal and terrestrial, but occasionally climb into shrubs and trees, especially when being pursued. Some are known to be good swimmers. Marmots are herbivorous. Food consists mainly of green vegetation such as grasses and forbs as well as fruits, grains, and legumes. Marmots use visual and auditory senses for the main contacts between members of a group, with some olfactory and tactile cues. A sharp whistle, which functions as an alarm call, is one of several types of vocalizations. Some species sun themselves on warm days near their burrow entrances.

Social behavior varies between the different species. Some, such as *M. monax*, are usually solitary, but sometimes are found in small family groups. Others such as *M. olympus* are highly sociable.

Hibernation. Most marmots inhabit northern mountainous areas where food is unavailable during the winter, thus making hibernation a necessity. They are true hibernators: they do not store food for winter but subsist entirely on fat reserves stored in their bodies. Fat reserves are accumulated during the summer and may account for as much as 20% of the animal's weight when it enters dormancy. Hibernation may last as long as 9 months in some species.

Yellow-bellied marmot, *Marmota flaviventris*. (*Photo by Dr. Lloyd Glenn Ingles, © 1999 California Academy of Sciences*)

Members of the genus *Marmota* and their geographical distribution

Species	Range
M. marmota, alpine marmot	Mountains of Europe
M. camtschatica, black-capped marmot	Mountains of NE Siberia
M. menzbieri, Menzbier's marmot	Mountains of Kyrgyzstan and Kazakhstan
M. bobak, Steppe or Siberian marmot	Poland and Romania to Manchuria and China
M. baibacina, gray marmot	Siberia, Kazakhstan, Kyrgyzstan, Mongolia
M. himalayana, Himalayan marmot	Himalayan region of China, India, Nepal
M. sibirica, Mongolian bobak	Siberia, Mongolia, China
M. broweri, Alaskan marmot	Northern Alaska, Yukon
M. caudata, red marmot	Mountains of Kyrgyzstan, Tajikistan, Afghanistan, India, Pakistan, Tibet
M. monax, woodchuck, groundhog, whistle pig	Alaska to Newfoundland and eastern United States
M. flaviventris, yellow-bellied marmot	SW Canada, western United States
M. caligata, hoary marmot	Alaska to Washington and Montana
M. olympus, Olympic marmot	Olympic Peninsula of western Washington
M. vancouverensis, Vancouver Island marmot	Vancouver Island

In southern parts of its range, the woodchuck may hibernate for only several weeks or not at all. *See* HIBERNATION AND ESTIVATION.

Reproduction and life. Periods of mating, gestation, parturition, and litter size vary as a result of latitude and altitude. Most marmots do not breed until they are 2 to 3 years of age. Breeding normally occurs shortly after emergence from hibernation with a single litter of 4 to 6 young being born annually after a gestation period of approximately 30 days. Female *M. olympus* usually produce a litter every other year. Young marmots are born in a grass-lined nest within the burrow. The young emerge from the burrow at about 4 weeks of age and are weaned at about 6 weeks. The young of some species may disperse as early as 6 weeks of age, whereas others may remain with their parents until their second summer when they are just over 1 year old. Life expectancy in wild *M. marmota* is 4 to 6 years, although one individual in captivity lived for almost 10 years.

Threats. Major natural predators include cougars, coyotes, dogs, foxes, bears, and golden eagles. Young animals are most susceptible to predation since adults are generally able to defend themselves against most natural predators.

Deforestation along with the subsequent expansion of pastureland and the elimination of large predators in the eastern United States has allowed populations of *M. monax* to greatly expand. Some, however, consider it a pest because of its damage to crops and that its burrows pose potential damage to livestock and farm machinery. For example, in agricultural areas, where groundhogs are often considered a pest because of crop depredation and burrowing, they may be shot, trapped, or gassed. Other species have experienced drastic population declines due to loss of subalpine habitat to ski resorts. *Marmota vancouverensis* is classified as Endangered by the International Union for the Conservation of Nature and Natural Resources (IUCN) and the U. S. Department of the Interior. In Central Asia, *M. menzbieri* is classified as vulnerable and *M. caudata* as near threatened by the IUCN.

Donald W. Linzey

Bibliography. G. A. Feldhamer, B. C. Thompson, and J. A. Chapman, *Wild Mammals of North America: Biology, Management, and Conservation*, 2d ed., Johns Hopkins University Press, 2003; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford Limited, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Mars

The fourth planet outward from the Sun. Mars is the planet most like Earth in terms of geology, climate, and suitability for life. It is also the world that has been most thoroughly studied by interplanetary spacecraft.

Orbital characteristics. Mars has a mean heliocentric distance (semimajor axis) of 1.524 astronomical units, equivalent to 141.6×10^6 mi (227.9×10^6 km). Its orbit plane is inclined 1.85° to the ecliptic plane. Its orbital eccentricity of 0.093, one of the largest of the major planets, causes Mars's distance to the Sun to vary from 128×10^6 mi (207×10^6 km) at perihelion to 155×10^6 mi (249×10^6 km) at aphelion. *See* PLANET.

Mars revolves around the Sun with a sidereal period of 1.881 years or 686.93 days at a mean orbital velocity of 15.0 mi/s (24.1 km/s). The planet comes closest to Earth, reaching opposition, every 780 days on average.

The globe. The mean diameter of Mars is 4212 mi (6780 km), or about 53% that of the Earth. The planet has a mass of 7.08×10^{20} tons (6.42×10^{23} kg), about 11% of Earth's. The mean density is thus about 3.91 g/cm^3, less than that of any other terrestrial planet. At the equator the surface gravity is 12.1 ft/s^2 (369 cm/s^2), about 38% of Earth's, and the escape velocity is 3.1 mi/s (5.0 km/s).

Astronomers long ago determined the planet's rotation period very accurately by watching surface markings cross the disk. The sidereal period is 24h

37m 22.7s, corresponding to a mean solar day of 24h 39m 35.2s. Consequently, there are about 669 Martian mean solar days (called sols) in the planet's year. The obliquity of Mars, that is, the inclination of its polar axis to a line perpendicular to the orbital plane, is 25.2°. This value is only slightly greater than that of Earth, so the seasons are similar but longer and also more unequal because of the planet's greater orbital eccentricity. Theoretical work has shown that the obliquity of Mars undergoes large, irregular variations that have ranged from 15° to 35° during the last 10^6 years and from 0° to 60° over geologic time. These chaotic shifts occasionally expose the polar caps to much more direct sunlight and trigger major climatic changes.

Telescopic appearance. Mars appears to the unaided eye as a bright, slightly reddish star. Viewed through a telescope, Mars usually appears as a salmon-hued disk marked by mottled dark regions and white, seasonally variable polar caps (**Fig. 1**). The atmosphere occasionally displays clouds of dust and water ice. In the late 1800s, some telescopic observers mapped a multitude of surface formations that were believed to be canals and oases due to vegetation responding to the Martian seasons. As revealed by spacecraft, however, these features are now known to be complexes of dark rock and light-collored sand whose contrast changes with the introduction of clouds and airborne dust driven by seasonal winds.

The apparent diameter of the disk of Mars varies from 3.5 seconds of arc at conjunction (when on the far side of the Sun) to 25.1 seconds of arc at the most favorable perihelic (near-perihelion) oppositions, when the distance to the Earth is only 35×10^6 mi (56×10^6 km). Perihelic oppositions with Mars occur every 15 to 17 years. During the perihelic opposition on August 27, 2003, Mars was marginally closer to Earth—34,646,418 mi (55,758,006 km)—than at any time in nearly 60,000 years.

As a superior planet (that is, a planet that moves outside Earth's orbit), Mars presents only gibbous phases when it is not in opposition to or conjunction with the Sun; the maximum phase angle (Sun-Mars-Earth angle) is 48°, when Mars is in quadrature. The apparent visual magnitude of Mars in conjunction averages −1.5, but it brightens to −2.7 at perihelic opposition, when Mars is brighter than any other planet except Venus. *See* MAGNITUDE (ASTRONOMY).

The visual albedo of Mars is about 0.15. The planet's reflectivity increases rapidly from a low value of only 5% in the near ultraviolet to more than 30% in the near infrared. This variation accounts for the reddish color of Mars. In the infrared, at wavelengths between 1 and 2 μm, the spectral reflectivity of Mars resembles that of iron oxides, especially goethite and hematite. *See* ALBEDO.

Interior. Measurements of the planet's gravitational field made by orbiting spacecraft have shown that Mars has a dense core and thus is differentiated into a core, mantle, and crust. The crust, composed of silicate rocks, varies irregularly in thickness from about 22 mi (35 km) near the north pole to 45 mi (70 km) near the south pole. The reason for this gradation is not well understood. The surface exhibits no obvious evidence of terrestrial-style plate tectonism.

At present Mars does not have a global magnetic field of appreciable strength. Spacecraft measurements suggest that the equivalent dipole field is at most 30 nanoteslas at the equator, corresponding to a magnetic moment 5000 times weaker than Earth's. However, early in its history Mars must have had a more substantial magnetic field, because its near-surface rocks exhibit strong remanent magnetism. Based on the absence of a remanent magnetic signature in the vicinity of large impact basins, geophysicists believe that the internal dynamo responsible for the planet's intrinsic global magnetic field must have ceased activity about 4×10^9 years ago. *See* ROCK MAGNETISM.

Geology. The Martian surface has been modified extensively by the processes of impact cratering, volcanism, faulting, and fluvial erosion. The terrain in the southern hemisphere is very heavily cratered and thus quite old, having formed some 3.8×10^9 years ago. Much of it stands 1 to 2 mi (1 to 4 km) higher than the planet's mean radius (**Colorplate 1**). Although volcanic sources are not common at southern latitudes, maps of spectral reflectivity made by orbiting spacecraft show that surface rocks in the southern hemisphere are predominantly basaltic (silicon-poor) in mineralogy. The northern hemisphere is dominated by vast, lava-covered plains with relatively few impact craters and an average elevation about 3 mi (5 km) below the mean radius. Mineral compositions are typically andesitic (silicon-rich). This puzzling hemispheric dichotomy has no well-accepted explanation but may be related to the large-scale interior movement early in Martian history.

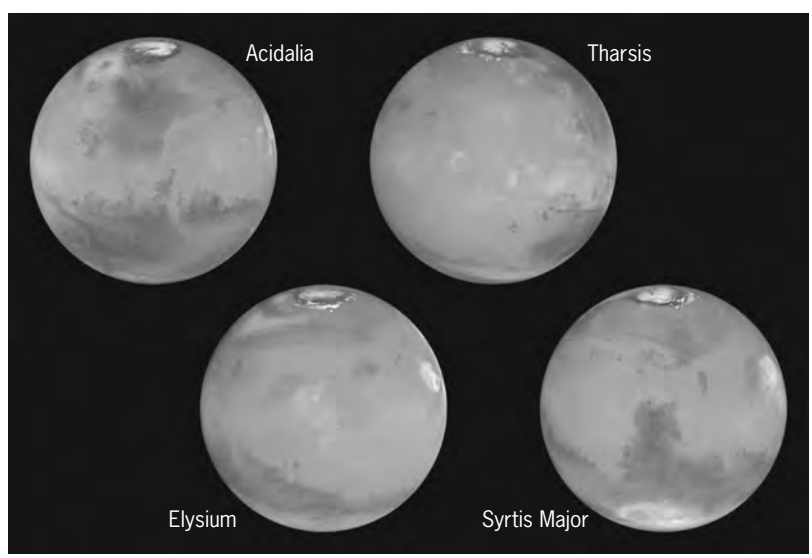

Fig. 1. Images of Mars in 90° longitude steps through one complete rotation. When the *Hubble Space Telescope* obtained these views in 1999, the planet was 54×10^6 mi (87×10^6 km) from Earth. Note the bright and dark albedo markings and bright polar caps. (*NASA/Space Telescope Science Institute*)

Fig. 2. The huge Martian shield volcano Olympus Mons covers an area the size of Arizona. Its peak is 13.2 mi (21.3 km) above Martian "sea level." Somewhat north of the summit is a conspicuous caldera, a complex nest of collapse craters 55 mi (90 km) across. (*NASA*)

The surface also exhibits numerous impact basins. The largest of these, Hellas, is approximately 2500 mi (4000 km) across and 6 mi (9 km) deep; it ranks as the largest known impact structure in the solar system.

The Tharsis ridge, or bulge, is an uplifted portion of the surface that stands several kilometers above the mean elevation of the planet. Most of Mars's tectonic features are associated with Tharsis, which affects approximately one-quarter of the entire surface. The formation of the ridge appears to have begun more than 3×10^9 years ago, long before the volcanoes that now dot its crest and surroundings. The height and mass of Tharsis may be supported below either by an extremely thick, rigid lithosphere or by the force from an upwelling mantle plume.

Volcanoes. The Martian surface exhibits several enormous shield volcanoes. The largest and perhaps the youngest is Olympus Mons, which is nearly 370 mi (600 km) across at its base and stands approximately 16 mi (26 km) above the surrounding terrain (**Fig. 2**). Three other large shield volcanoes—Ascraeus Mons, Pavonis Mons, and Arsia Mons—lie along the nearby Tharsis ridge. The Tharsis volcanoes are approximately 250 mi (400 km) across and 12 mi (20 km) high. By comparison, the largest volcanic constructs on Earth—Mauna Loa and Mauna Kea in Hawaii—together are approximately 125 mi (200 km) across and stand 5.5 mi (9 km) above the ocean floor.

In shape and structure, the Martian shield volcanoes bear a strong resemblance to their Hawaiian counterparts. At the summit of each shield is a complex of calderas, collapsed craterlike features that were once vents for lava. The central smooth-floored caldera of Arsia Mons is the largest, about 80 mi (130 km) in diameter.

About 3000 mi (5000 km) to the west of the Tharsis area is another volcanic field, dominated by the symmetrical shield volcano, Elysium Mons, approximately 140 mi (225 km) across. Many other volcanoes with dimensions of up to 60 mi (100 km) and in various states of preservation are found scattered over the Martian surface, mostly in the northern hemisphere.

A volcanic formation known as a patera (Latin for "saucer"), a shallow, complex crater with scalloped edges, occurs on Mars but has not been observed on any other planet. Alba Patera, northeast of Olympus Mons, is between 900 and 1200 mi (1500 and 2000 km) across but rises only 2 to 3 mi (3 to 5 km) above the surrounding plains. Enormous quantities of highly fluid lava must have emerged from its central caldera, flowing to great distances along channels and tubes. *See* VOLCANO.

Canyons. Perhaps the most spectacular features on the Martian surface are the huge canyons located primarily in the equatorial regions. Valles Marineris, actually a system of canyons, extends for over 3000 mi (5000 km) along the equatorial belt (**Colorplate 2**). In places, the canyon complex is as much as 300 mi (500 km) wide and drops to more than 4 mi (6 km) below the surrounding surface. Dwarfing the Grand Canyon, Valles Marineris is comparable in size to the great East African Rift Valley. In general, the walls of the canyon are precipitous and show evidence of slumping and landslide activity (**Fig. 3**). The principal canyons of Valles Marineris appear to be linked with the formation of the Tharsis uplift. They may

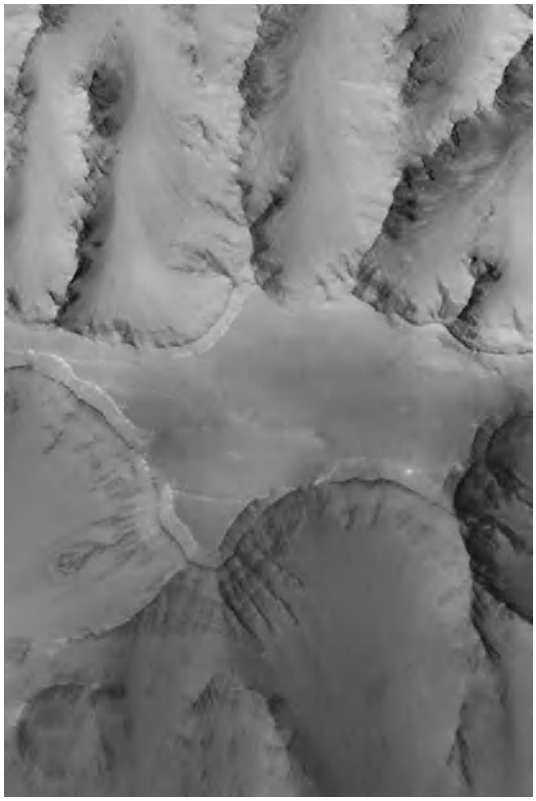

Fig. 3. An image from the *Mars Global Surveyor* resolves features as small as 20 ft (6 m) in a 6-mi-wide (10-km) portion of Coprates Chasma, which is located in the middle of Mars's vast Valles Marineris canyon complex. The high-standing plateau is underlain by multiple rock layers that may be volcanic or sedimentary in nature. (*NASA/Malin Space Science Systems*)

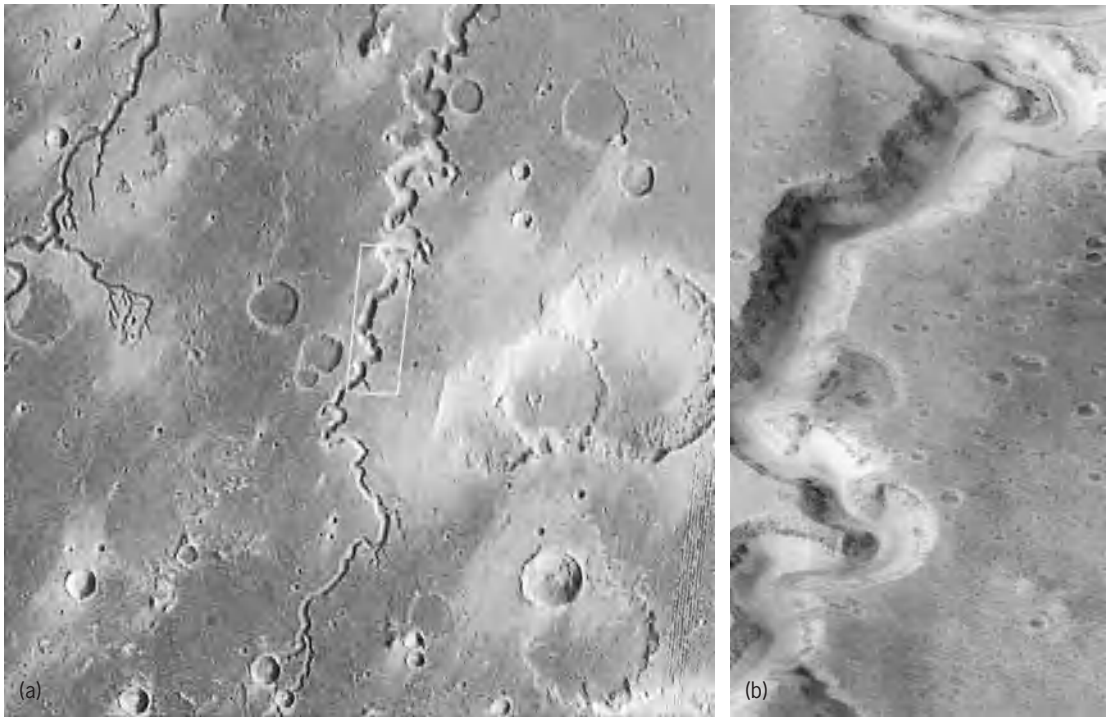

Fig. 4. Nanedi Vallis, a long, water-cut canyon on Mars, which is 1.5 mi (2.5 km) wide. (*a*) Canyon and surrounding terrain. (*b*) Detailed view of area in white rectangle in part a. Rocky outcrops jut from the canyon's walls, and a secondary rivulet is seen in its floor. (*NASA/Malin Space Science Systems*)

have formed by faulting along lines of tension within the crust. In essence, such canyons (called graben) are pairs of parallel faults with a sunken area between them.

Channels. An astonishing class of features on the Martian surface is a widespread network of channels that bear a very strong resemblance to dry river beds (**Fig. 4**). These channels do not coincide with the channels or canals which, in the past, were frequently reported by telescopic observers of Mars.

Ranging in size from broad, sinuous features nearly 40 mi (60 km) wide to small, narrow networks less than 300 ft (100 m) wide, the channels appear to have been created by water erosion. Early analysis assumed the apparently dendritic nature of some channels to be evidence of watershed rain collection and episodic flow. On closer inspection, however, many systems previously considered dendritic now appear to result from more gradual formation through ground-water sapping, not rainfall.

The largest channels must have been formed by enormous torrents of water, presumably released in some catastrophic manner (**Fig. 5**). Yet liquid water at present cannot exist on the surface of Mars, because the atmospheric pressure and surface temperature keep most available water trapped as subsurface ice or locked up in the permanent polar caps.

As a way out of this enigma, it has been suggested that the long-term climate of Mars is variable, and that in the past the planet might have had conditions suitable for free running water over its surface. Alternatively, episodic volcanic activity may have rapidly melted buried ice and disgorged torrents of water. According to some estimates, liquid-water reservoirs could have been maintained under relatively thin ice covers for up to 700×10^6 years after the mean global temperature fell below the freezing point of water. Some geologic evidence indicates that huge lakes and glaciers of ice were present on the Martian surface many millions of years ago.

Images from the *Mars Global Surveyor* spacecraft have revealed the existence of more than 120 small gullies cascading down the slopes of hillsides and crater rims that appear to have been carved by flowing water or mud. These features are typically about 1500 to 3000 ft (0.5 to 1 km) long and usually form on slopes facing away from the equator, which receive less sunlight than their surroundings and are thus colder. Because the gullies appear to be extremely fresh, they imply that liquid water in some form has flowed on the Martian surface within the past 10^6 years and perhaps much more recently. Some scientists have argued that the gullies result from springs or pressurized underground flows, while others believe they were created by trickling water from melting snow packs.

Sedimentary layering. Spacecraft images show that many canyon walls and crater floors exhibit extensive exposures of layered rock, up to 2.5 mi (4 km) thick in some places. Although their stratigraphic age is uncertain, these layers imply that the deposition of sediment was widespread early in Martian history. The frequent presence of layering within craters argues for repeated episodes of deposition in standing water. However, it is also possible that the sediments were transported by the atmosphere, either as massive amounts of windblown dust or as airborne volcanic ash.

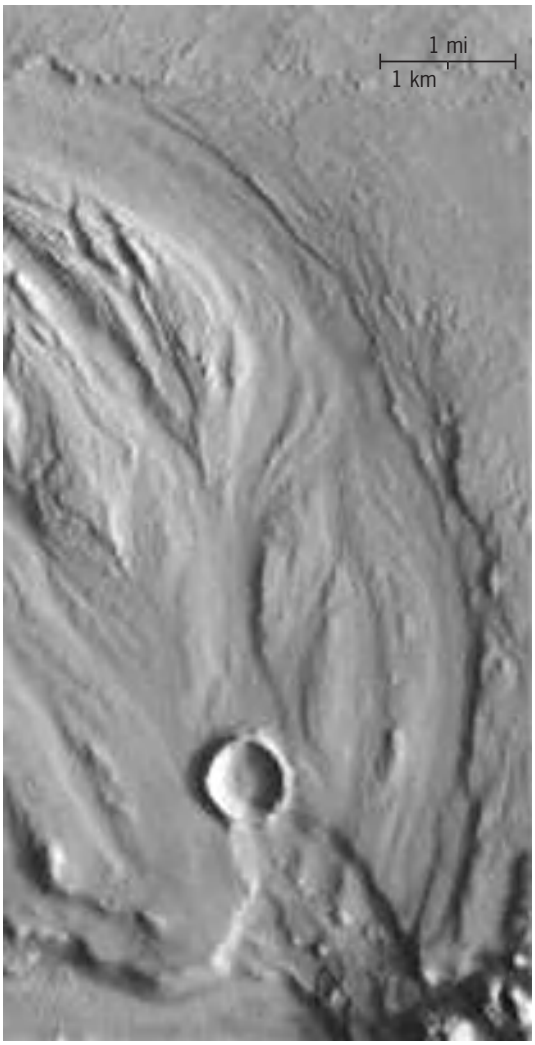

Fig. 5. Section of the Martian flood channel Hrad Vallis, showing streamlined flow features. (*NASA/Malin Space Science Systems*)

Surface characteristics. Five spacecraft have landed safely on the Martian surface and returned significant scientific data about its properties. *See* SPACE PROBE.

Two American *Viking* landers arrived in 1976. *Lander 1* set down in a region known as Chryse Planitia (22.3°N, 48.0°W) on July 20, and *Lander 2* arrived at Utopia Planitia (47.7°N, 225.7°W) on September 3. *Lander 1* sits in a moderately cratered, low-lying volcanic plain near the mouth of a large outflow channel. Numerous rocks of varying size litter the surroundings, and fine-grained material is occasionally seen in large drifts. *Lander 2* is situated in a region of fractured plains about 125 mi (200 km) south of the large crater Mie; it probably sits atop a lobe of the crater's ejecta blanket. This site is very flat and sparsely cratered, though rocks cover a higher proportion of the surface than at the other *Viking* site.

Most of the data concerning the composition of the surface came from x-ray fluorescence spectrometers, which showed that small particles were remarkably similar at both sites, with silicon and iron accounting for about two-thirds of the samples' content. The sulfur content was unexpectedly high, 100 times greater than in the Earth's crust, an indication that much of the surface is coated with very fine grain weathering by-products enriched in salts. In general, the surface material at both sites can be characterized as iron-rich clays. Small particles also clung to magnets on each spacecraft, and these are probably composed of the mineral maghemite. *See* X-RAY FLUORESCENCE ANALYSIS.

On July 4, 1997, *Mars Pathfinder* also landed in Chryse Planitia about 500 mi (800 km) east-southeast of *Viking 1*, near the mouth of the Ares Vallis flood channel at 19.3°N, 33.6°W. The lander's stereo camera recorded a dense accumulation of sizable rocks (**Fig. 6**), which lay on gently sloping ridges and troughs presumably shaped by the flood waters. *Mars Pathfinder* also carried a small, instrumented rover, *Sojourner*, which was directed from Earth via remote control to several rocks and fine-grained drifts in the lander's immediate vicinity.

The most extensive surface exploration has been conducted by two Mars Exploration Rovers (MERs) that landed on Mars in early 2004. *Spirit* arrived on January 3, landing in the large crater Gusev (14.6°S, 184.5°W). *Opportunity* reached the flatland known as Meridiani Planum (2.0°S, 5.5°W) on January 25. Each rover was equipped with panoramic and microscopic cameras, three different spectrometers to characterize surface composition, magnets to collect metallic particles, and a grinding tool to expose fresh rock surfaces. Although the MERs had been designed to operate for 3 months, both were still fully functional in early 2006, more than 2 years after their arrival. In that time *Spirit* traveled more than 3.8 mi (6.1 km) and *Opportunity* logged 4.0 mi (6.5 km).

Both MER landing sites were chosen in the hope of finding evidence for the presence of liquid water at ground level in the geologically recent past. Scientists had chosen Gusev, which is 103 mi (170 km) across, because its floor was once flooded and appeared to be covered with water-borne sediment. Although *Spirit* initially found no evidence of lakebed deposits, it later identified an iron oxide called goethite, which forms only in the presence of water, in elevated terrain called the Columbia Hills about 2 mi (3 km) from its landing site.

Meridiani Planum was chosen for *Opportunity* because its spectral reflectivity strongly implied the presence of coarse-grained hematite, another iron oxide, which forms only in wet environments. After *Opportunity* arrived, scientists discovered that its landing site exhibited countless millimeter-size concretions, nicknamed blueberries, scattered across the ground and embedded in the exposed bedrock (**Fig. 7**). These spherules are the source of the hematite detected spectroscopically from orbit. *Opportunity* also found that the outcrops are rich in sulfate minerals, evidence that the rock had been saturated by salt-laden water; one of these minerals, jarosite, requires highly acidic conditions to form.

Meteorites from Mars. Geologists have come to recognize that a small number of meteorites have reached Earth after being violently ejected from Mars

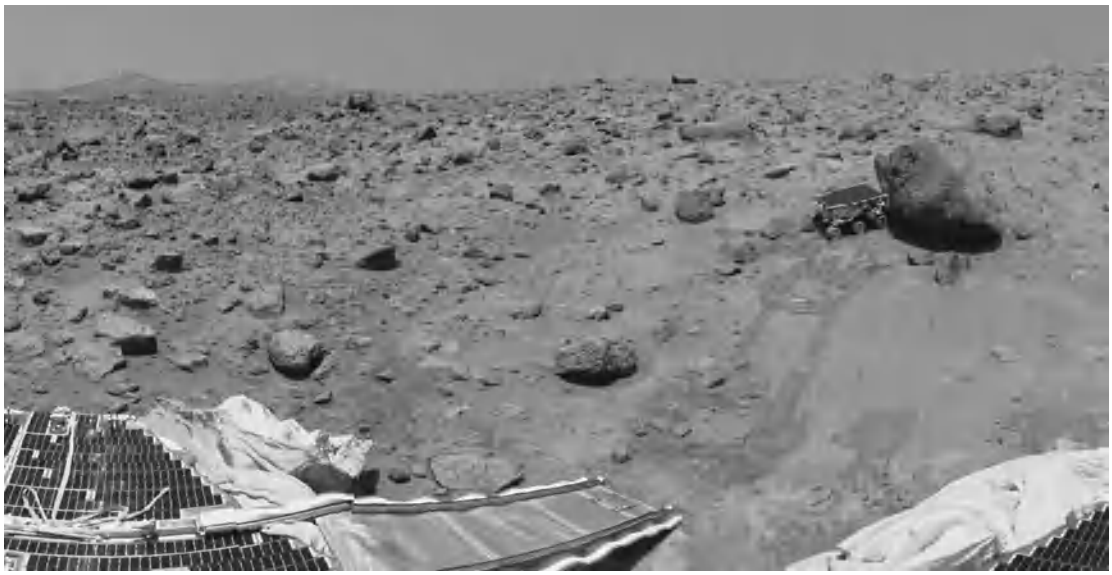

Fig. 6. View of the Ares Vallis flood plain, as recorded by *Mars Pathfinder* in July 1997. The hill on the horizon, near the left edge, is 0.6 mi (1 km) from the lander. At right, the rover *Sojourner* is analyzing the elemental composition of a large boulder nicknamed Yogi. (*NASA*)

during the collision of sizable asteroids or comets. The meteorites are thus samples from Mars itself. The first of these meteorites was recognized in 1982, and 36 distinct falls (some represented by multiple samples) were known by early 2006. They are sometimes collectively termed the SNC meteorites, representing their three subclasses: shergottites, nakhlites, and chassignites. The Martian meteorites contain a distinctive ratio of the isotopes of oxygen found nowhere on Earth, on the Moon, or in other meteorites. Almost all of them crystallized from molten rock 1.3×10^9 years ago, far more recently than the usual 4.6×10^9 years (the age of the solar system) found in all other meteorites. The lone exception to date, designated ALH 84001, crystallized 4.5×10^9 years ago during the planet's infancy.

All but one of the Martian meteorites are basaltic, as would be expected from volcanic flows with relatively young geologic ages. By contrast, ALH 84001 is an igneous rock whose large crystals argue that it cooled slowly within the crust. Some of the Martian meteorites contain tiny amounts of water, hydrated minerals, and carbonates that must have become trapped within the rocks while still on Mars. *See* METEORITE.

Atmosphere. The Martian atmosphere is very thin, with surface pressure averaging about 6.5 millibars (650 Pa), or less than 1% of the pressure at Earth's surface. Atmospheric pressure, temperature, wind velocity, and direction were monitored by both *Viking* landers for more than a full Martian year and by the *Mars Pathfinder* lander for 4 months. Temperatures during these missions ranged from a low of −190°F (−123°C) to a high exceeding 0°F (−18°C), though dust storms tend to moderate these extremes. The globally averaged surface temperature is −81°F (−63°C).

Composition. The atmosphere consists principally of carbon dioxide (95.3%) but contains nitrogen, argon, oxygen, and a trace of water vapor totaling 4.7%. In 2004 a spectrometer aboard *Mars Express*, a spacecraft orbiting the planet, discovered trace amounts of methane in the Martian atmosphere. The mixing ratio of methane varies from 0 to 30 parts per billion by volume, depending on location. Methane cannot exist in equilibrium with the other gases, and scientists are unsure of its source. Speculations include delivery to Mars by a comet, release during a volcanic eruption or by the decomposition of stored methane hydrates, or production by methane-excreting (methanogenic) organisms either at present or in the past.

Clouds. Both carbon dioxide and water form clouds in the Martian atmosphere. Certain water-ice condensations are associated with the largest volcanic

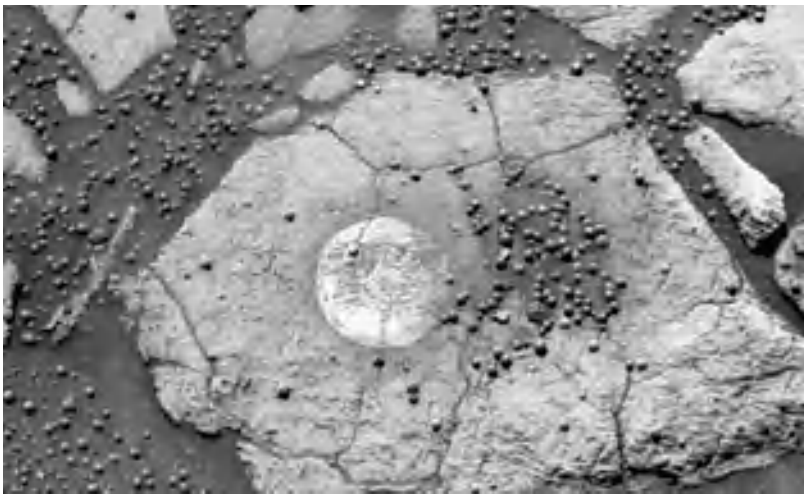

Fig. 7. Image from the roving spacecraft *Opportunity* showing a rock outcrop at Meridiani Planum strewn with small spherical concretions, nicknamed blueberries, that contain the iron-bearing mineral hematite. This hematite type forms only in the presence of water. The light-colored circle, 2 in. (5 cm) across, was created by the rock abrasion tool on *Opportunity*. (*NASA/USGS*)

peaks, while carbon dioxide clouds occur frequently over the polar regions during local winter, when the region is in continuous darkness. Often clouds appear during or just after sunrise, either because vapor has condensed in the atmosphere during the cold night, or because frosts that were deposited on the surface overnight are warmed by the Sun and sublime back into the atmosphere, saturating it with vapor.

Early history. The oldest terrains, in Mars's southern hemisphere, bear systems of small, interconnected channels as well as craters and other landforms that have been heavily eroded and degraded. These features suggest that early in its history the planet had a much thicker atmosphere. However, most or perhaps virtually all of the original atmosphere disappeared beginning about 4×10^9 years ago. The loss of gas was probably due to a combination of impact events, stripping by the solar wind, and chemical reaction with surface rocks.

Dust storms. Localized dust storms appear quite frequently on Mars. Because the Martian atmosphere is extremely thin, wind velocities greater than 90–110 mi/h (40–50 m/s) are needed to set surface dust grains in motion. The most abundant particles in a typical Martian dust storm are estimated to be about 10 μm in diameter, similar to finely powdered talc. The very smallest particles may be carried to heights as great as 30 mi (50 km) above the surface. Localized "dust devils" are apparently widespread and occur frequently.

Some dust storms develop with such intensity that their total extent may be hemispheric or even global. Such was the case of the great dust storm of 1971, which began in September and completely enshrouded the planet by the time *Mariner 9* arrived in November. Major dust storms tend to begin at the end of spring in the southern hemisphere, when Mars is near perihelion. The initial dust-raising disturbance is probably related to temperature instabilities in the Martian atmosphere at a time when solar heating is greatest. Once the dust is in the air, increased absorption of solar energy causes the dust-laden air to rise, producing strong surface winds that rush in to replace the rising air mass. Thus the storm propagates, growing in intensity and extent. When the distribution of dust finally reaches global dimensions, the temperature differences over the planet are diminished, winds correspondingly subside, and dust begins to settle out of the atmosphere. The settling process may take weeks or, in the case of a severe storm, several months. The 1971 storm lasted for approximately 5 months.

Water inventory. If all the water currently present in the atmosphere of Mars were to condense onto the surface, it would form a film only 10 μm thick. This is inconsistent with the amount of water needed to form the planet's huge eroded channels. For example, the volume of water needed to form the flood-related features in and near Chryse Basin alone is roughly 1.0×10^6 mi^3 (4.2×10^6 km^3). Therefore a much larger source of water must exist now or must have existed in the geologically distant past.

However, the surface of Mars exhibits widespread exposures of olivine, an iron-magnesium silicate mineral that readily reacts with water to form other compounds. The frequent occurrence of olivine implies that much of the Martian surface has not been exposed to liquid water for billions of years and argues that Mars has been cold and dry for much of its history.

Polar caps. The seasonal cycle of growth and decay of the bright polar caps has long been taken as evidence of the presence of water on Mars. However, evidence from the *Mariner* spacecraft identified carbon dioxide as the principal constituent of the polar snow, with lesser amounts of water ice also present. Formed during autumn and winter by condensation and deposition of the icy mist covering the polar regions, the polar caps at the end of winter cover a vast area extending to latitude 60° in the southern hemisphere and 70° in the northern hemisphere. The surfaces covered are, respectively, 0.4 and 1.7×10^6 mi^2 (1 and 4.5×10^6 km^2).

During spring and summer the edges of the caps retreat, reaching minimum diameters of about 400 and 300 mi (650 and 450 km), respectively. The north pole's residual cap consists almost entirely of water ice; although likely to be dominated by water ice, the southern cap is covered with frozen carbon dioxide (**Fig. 8**). Measurements by the *Mars Global Survevor* spacecraft showed that both were broad, elevated domes. The volume of water ice within the northern residual cap is likely 290,000–430,000 mi^3 ($1.2–1.7 \times 10^6$ km^3); the estimate for the southern residual cap is 480,000–720,000 mi^3 ($2–3 \times 10^6$ km^3). If all the water present in the polar caps were distributed uniformly over the surface of Mars, it would produce a layer 70–110 ft (22–33 m) deep.

Subterranean ice. Although no spacecraft to date has detected the presence of liquid water directly, there is ample evidence that a great volume of ice lies beneath the surface. Broad, flat plains regions

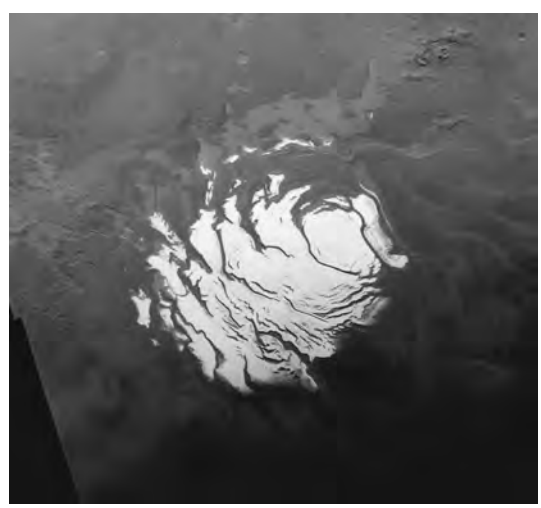

Fig. 8. Mars's south polar cap as it appears near its minimum size, about 250 mi (400 km) across, during local summer. (*NASA/USGS*)

surround each residual cap. This terrain is characterized by stacks of evenly spaced layers from 60 to 150 ft (20 to 50 m) in thickness. Although ice is absent in these layered terrains at ground level, a gamma- ray spectrometer on the 2001 *Mars Odyssey* spacecraft has detected areally extensive subsurface layers enriched in hydrogen. The only plausible source of this hydrogen is water ice, which apparently exists in concentrations of 40 to 73% by volume. The ice-rich layers are located poleward of about latitude 60° in both the northern and southern hemispheres, and the estimated volume of ice exceeds 2400 mi^3 (10,000 km^3), enough to fill Lake Michigan twice. However, *Mars Odyssey's* gamma-ray spectrometer could detect the presence of hydrogen to a depth of only about 3 ft (1 m), so the ice-rich layer may extend to much greater depths. By early 2006 a radar instrument on *Mars Express* had detected the likely presence of massive ice reservoirs buried beneath the plains surrounding both polar caps.

Hypothesized ocean. Even though at present liquid water cannot exist on the Martian surface, the planet's numerous large flood channels suggest that vast amounts of water flowed on the surface early in Martian history. This conclusion, together with water bound in the polar caps and evidence for widespread deposits of subsurface ice, suggests that the total volume of water on Mars would likely cover the planet to a depth of about 0.1–0.3 mi (0.2–0.5 km). The corresponding value for Earth is 1.7 mi (2.7 km). Circumstantial evidence implies that much of the low-lying northern hemisphere was covered with an ocean roughly 4×10^9 years ago. Although unproven, this hypothesis suggests that 10^7 mi^2 (27×10^6 km^2) of the northern plains were inundated. The total volume would have been roughly 3.6×10^6 mi^3 (15×10^6 km^3), about one-third of the water in the Atlantic Ocean.

Possibility of life. The ample evidence that early Mars had an abundance of liquid water, a thicker atmosphere, and a more clement environment has raised speculation that the planet may have once been, or may still be, an abode for primitive life forms.

In 1976, the twin *Viking* landers conducted experiments on the Martian surface to test for the existence of life. Three biological instruments and a gas chromatograph/mass spectrometer sampled the soil within reach of the 10.5-ft (3.2-m) mechanical arm aboard each spacecraft. The spectrometer detected no organic compounds at all (within the limits of its sensitivity). One of the biological experiments gave a positive result—the release of oxygen from a soil sample when humidified—but there is no consensus as to the source of this reaction. Some scientists have suggested inorganic superoxides or other catalysts as the substances involved. *See* GAS CHROMATOGRAPHY; MASS SPECTROMETRY.

Twenty years later, a scientific team claimed that it had found evidence of fossilized microbes within a known Martian meteorite (ALH 84001). They reported that tiny carbonate beads within the meteorite had chemical and mineral characteristics most logically explained by biological activity. However, other research groups dispute this conclusion and have offered various nonbiological explanations for the same features. After years of intensive study, scientists have reached no consensus; the mineral associations within ALH 84001 and other Martian meteorites neither confirm the presence of once-living organisms nor exclude it.

Satellites. Mars has two small satellites, Phobos and Deimos, discovered by Asaph Hall in 1877. Their apparent visual magnitudes at mean opposition are relatively bright, 11.3 and 12.4, respectively, but a large telescope is required to see them because of the bright glare surrounding the planet.

The outer satellite, Deimos, moves at a mean distance from the planet's center of 14,580 mi (23,460 km) in a sidereal period of 30h 18m, only slightly longer than a period of rotation of Mars. The orbital plane of Deimos is inclined to the equatorial plane of Mars about 1.8°. The inner satellite, Phobos, orbits only 3700 mi (6000 km) from the surface of Mars and 5830 mi (9380 km) from the planet's center. The inclination of its orbital plane is 1.9°. Phobos has a sidereal orbital period of 7h 39m, less than one-third of the mean Martian solar day. Thus, as seen from the surface of Mars, Phobos rises in the west and sets in the east twice daily, moving apparently in the direction opposite to all other celestial bodies, including Deimos. Seen overhead near the Martian equator, Phobos would appear less than half the size of the Moon viewed from Earth, and Deimos would be a barely discernible disk.

Neither satellite is massive enough to be gravitationally contracted to a spherical shape (**Fig. 9**). Phobos measures only $11 \times 14 \times 17$ mi ($18 \times 22 \times 27$ km), and Deimos just $6 \times 8 \times 9$ mi ($10 \times 12 \times 15$ km). These tiny satellites have very low surface gravity and an escape velocity of only about 30 ft/s (10 m/s). A person standing on Deimos or Phobos could quite easily throw rocks that would leave the satellite forever.

Both satellites are saturated with impact craters; the largest is Stickney, a 6-mi-diameter (10-km) crater located on Phobos. Had the impacting object that excavated Stickney been much larger, it probably would have destroyed Phobos. In fact, it has been suggested that Deimos and Phobos may be the shattered and, perhaps, partially reconstituted remains

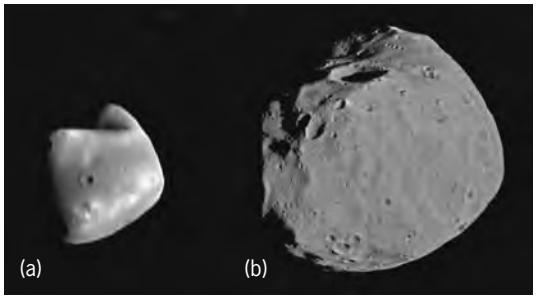

Fig. 9. The two Martian moons, (*a*) Deimos and (*b*) Phobos, have irregular outlines shaped by past collisions. (*NASA*)

of an earlier, single, large satellite of Mars. Both satellites have the same low albedo, approximately 0.05, so that surfaces are as dark as the very darkest asteroids. This low reflectivity suggests that the Martian satellites may have originated in the asteroid belt and been captured by Mars long ago. *See* ASTEROID.

J. Kelly Beatty

Bibliography. A. Albee, The unearthly landscapes of Mars, *Sci. Amer.*, 288(6):44–53, June 2003; J. K. Beatty, C. C. Petersen, and A. Chaikin (eds.), *The New Solar System*, 4th ed., Sky Publishing, Cambridge, MA, and Cambridge University Press, England, 1999; M. H. Carr, *Water on Mars*, Oxford University Press, New York, 1996; W. Sheehan, *The Planet Mars: A History of Observation and Discovery*, University of Arizona Press, Tucson, 1996; S. Squyres, *Roving Mars: Spirit, Opportunity, and the Exploration of the Red Planet*, Hyperion, New York, 2005.

Marsileales

A small order of heterosporous, leptosporangiate ferns (division Polypodiophyta) which grow in water or wet places and are rooted to the substrate. The leaves arise on long stalks from the rhizome and typically have floating blades with four leaflets, suggesting a four-leaved clover. The sporangia are enclosed in modified folded leaves or leaf segments called sporocarps. Megasporangia and microsporangia are borne in the same sporocarp. The order contains a single family and only 3 genera, about 50 species in all, most of them belonging to the widespread genus *Marsilea* (water clover). *See* POLYPODIOPHYTA.

Arthur Cronquist

Marsupialia

An order of animals, long considered the only order of the mammalian infraclass Metatheria.

Morphology. The marsupials are characterized by the presence of a pouch (marsupium) in the female, a skin pocket whose teat-bearing abdominal wall is supported by epipubic bones. The young are born in an embryonic state and crawl unaided to the marsupium, where they attach themselves to the teats and continue their development. In a few species the pouch is vestigial or has disappeared completely. The most complete dental formula for the living genera is exemplified by that of the family of American opossums (Didelphidae): I 5/4 C 1/1 Pm 3/3 M 4/4. The number of teeth may be reduced in other genera. Diprotodonty, or enlargement of the lower median incisors, is characteristic of the Australian herbivorous marsupials (**Fig. 1**). The last (third) premolar is the only replacement tooth in the cheek tooth row. The angle of the mandible is usually inflected. The five-toed plantigrade condition of the hindfeet in most marsupials is curiously modified in some Australian families to the syndactylous condition, in which the second and third digits are contained in a common web of skin (**Fig. 2**). *See* DENTITION.

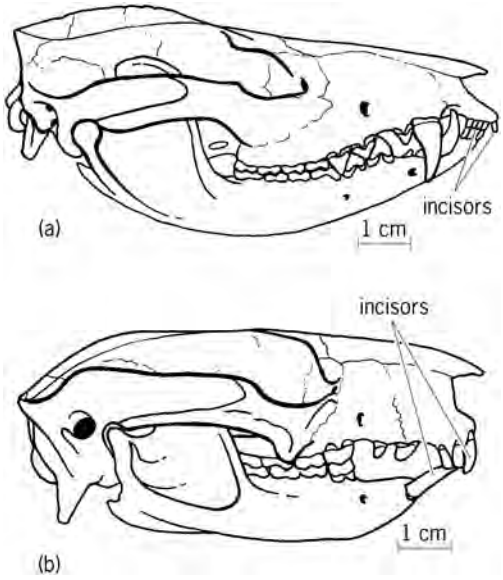

Fig. 1. Skull and jaws of marsupials. (*a*) Didelphidae, *Didelphis* (the carnivorous opossum of North America), Recent, with polyprotodont incisors (I 5/4). (*b*) Phalangeridae, *Trichosurus* (the herbivorous opossum of Australia), Recent, with diprotodont incisors (I 3/1).

A wide variety of terrestrial adaptations are found among the 82 genera of living marsupials. Some spectacular examples of evolution convergent upon placental modes of life are exemplified by the Australian marsupial "moles" (*Notoryctes*), "wolves" (*Thylacinus*), and "flying squirrels" (*Petaurus*). For the most part, the adaptive radiation of the marsupials has paralleled that of the placentals, yielding many ecological analogs, such as kangaroos and wallabies, which are the counterparts of the placental deer and antelopes, although these groups differ widely in structure. *See* EUTHERIA.

Phylogeny. The marsupials and placentals share a common Middle Cretaceous ancestor, and both were undergoing vigorous adaptive radiation in the Late Cretaceous, when the fossil record of the marsupials begins in North America. In the Cenozoic the marsupials diversified most extensively in the southern continents, such as Australia, where they lacked

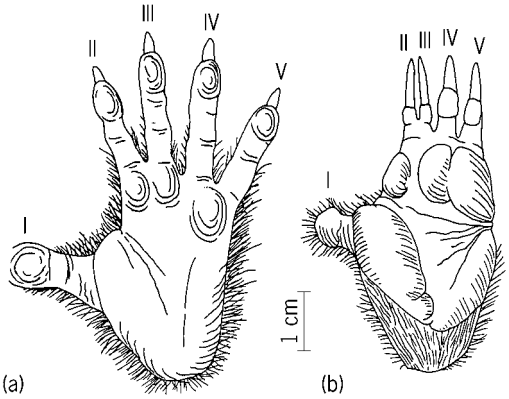

Fig. 2. Views of the soles of the left hindfoot of marsupials. (*a*) Didelphidae, *Didelphis*, representing the five-toed, polydactylous condition. (*b*) Phalangeridae, *Trichosurus*, representing the five-toed, syndactylous condition (digits II and III are bound in one web of skin).

effective placental competition, and South America, where they had only partial competition. Northern Hemisphere Cenozoic records represent only a single family (Didelphidae, the opossums), which is known in North America and Europe from the Paleocene into the early Miocene. The living American opossum (*Didelphis*) is a Pleistocene immigrant from Central and South America. There are no later records of marsupials in Europe and no Cenozoic records at all in Asia and Africa.

The marsupials were well represented during the Cenozoic in South America, where they underwent an adaptive radiation limited by direct competition with contemporary placental herbivores; consequently, the marsupials entered only a restricted number of adaptive zones. The carnivorous families (Borhyaenidae, Didelphidae) were dominant, but insectivorous groups (Caenolestidae, Necrolestidae) were present, and in the early Cenozoic some rodentlike types (Polydolopidae, Groeberiidae) evolved before the appearance of placental rodents in the Oligocene. The extinct carnivorous borhyaenids converged with placental carnivores in hyenalike (*Borhyaena*), wolflike (*Lycopsis*), and saber-toothed (*Thylacosmilus*) adaptations, and even more closely paralleled the carnivorous marsupials of Australia (especially the Thylacinidae). Late in the Cenozoic a curious group of saltatorial caenolestoids (Microtraguliidae) appeared which paralleled the smaller kangaroos of Australia. Only the families Didelphidae and Caenolestidae survived the Pleistocene invasion of South America by the placental carnivores.

It is not known when or how marsupials reached Australia, but evidence from the late Cenozoic faunas indicates that 10 of the 12 recognized families have experienced an evolutionary history that probably extends back into early Cenozoic or even Cretaceous time. Only three families, Thylacoleonidae, Diprotodontidae, and Wynyardiidae, became extinct by the end of Pleistocene time, and the last-mentioned is known only from the early Miocene. The Australian adaptive radiation is often cited as an excellent example of that evolutionary process, since a wide range of adaptive zones were invaded. Families of carnivorous (some Dasyuridae, Thylacinidae, Thylacoleonidae), insectivorous (some Dasyuridae, Peramelidae, Notoryctidae), herbivorous (Phalangeridae, Phascolarctidae, Vombatidae, Wynyardiidae, Diprotodontidae, Macropodidae), and nectar-feeding (Tarsipedidae) marsupials are represented. In addition to almost all the living genera, the Australian Pleistocene fauna included many very large extinct forms, such as giant kangaroos (Sthenurinae) and wombats (*Phascolonus*), diprotodonts such as *Diprotodon*, which was as large as living rhinoceroses, and the so-called marsupial lion *Thylacoleo*.

Taxonomy. Although the marsupials traditionally have been arranged as a single order, Marsupialia, it has long been recognized that their diversity of adaptations and length of geological history parallels that of all 30 orders of placentals. Reviews of living and fossil marsupials and the serology of the living species have given impetus to the formalization of various arrangements of the marsupial families as orders and consequently to the elevation of the term Marsupialia to superordinal rank. The proposed schemes agree in uniting the polydactylous and polyprotodont carnivorous superfamilies (Didelphioidea, Borhyaenoidea, and Dasyuroidea) either as a separate order, Marsupicarnivora Ride (1964), or with the syndactylous and polyprotodont Perameloidea as order Polyprotodonta Owen (1866). The Perameloidea are sometimes separated as order Peramelina Gray (1825). The South American diprotodont and the polydactylous, insectivorous, and rodentlike Caenolestoidea are grouped as order Paucituberculata Ameghino (1894). Most of the diprotodont, syndactylous, herbivorous marsupials of Australia (superfamilies Vombatoidea, Phalangeroidea, and Tarsipedoidea) are placed in the order Diprotodonta Owen (1866). *See* ANTEATER; KANGAROO; KOALA; MAMMALIA; METATHERIA; OPOSSUM.

Richard H. Tedford

Bibliography. M. J. Benton, *Vertebrate Paleontology*, 1991; R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., 1999.

Marten

One of eight groups of carnivorous mammals in the family Mustelidae along with the fisher, weasels, mink, polecats, wolverines, otters, and badgers. Six are found in Europe and Asia and two (marten and fisher) are found in North America (see **table**). The stone marten (*Martes foina*) inhabits coniferous and deciduous woodlands, while the European pine marten (*M. martes*), sable (*M. zibellina*), Japanese marten (*M. melampus*), and American pine marten (*M. americana*) are restricted to northern coniferous forests.

General morphology. The marten (*Martes*) is a medium-sized, dark-colored, semiarboreal mustelid with a moderately elongated shape, sharply pointed face, rounded ears, and a bushy tail (see **illustration**). Some members of this genus possess a white

Geographic distribution of martens

Name	Range
European pine marten *Martes martes*	Central and Northern Europe to western Siberia
Japanese marten *Martes melampus*	Japan, South Korea
Yellow-throated marten *Martes flavigula*	SE Asia, Korea, Borneo, Sumatra, Java
American pine marten *Martes americana*	N. North America to Sierra Nevada and Rockies
Sable *Martes zibellina*	Scandinavia to N. Asia, N. Japanese islands
Stone marten *Martes foina*	Denmark and Spain to Mongolia and Himalayas
Fisher *Martes pennanti*	N. North America to California and West Virginia
Nilgiri marten *Martes gwatkinsi*	Nilgiri Mountains of southern India

American pine marten, *Martes americana*. (*Photo by Gerald and Buff Corsi; © 2001 California Academy of Sciences*)

throat patch. The throat patch in the American pine marten is reddish-orange, while the yellow-throated marten (*M. flavigula*) and the Nilgiri marten (*M. gwatkinsi*) have yellow throat patches. The fisher (*M. pennanti*) lacks a throat patch. Martens are digitigrade (that is, they walk on the digits with the posterior part of the foot raised from the ground) with five toes on each foot. The large paws possess hairy soles and semiretractile claws that are not sheathed. A pair of elongated abdominal scent glands and paired anal scent glands are present in both sexes and are used for marking. Martens range in head-body length 30–75 cm (12–45 in.) and have a 12–45 cm (5–18 in.) tail. They weigh 0.5–6 kg (1–13 lb). Males are considerably larger than females in all species.

Behavior and habitats. Martens are generally solitary. They are among the most agile and graceful members of the weasel family and are able to leap effortlessly from branch to branch. Martens are opportunistic hunters and kill their prey with a bite to the back of the neck.

Martens are crepuscular and are primarily nocturnal, although they may be active any time of the day during all seasons. Dens may be in hollow trees or on or under the ground in rock crevices, hollow logs, tree roots, or under the snow. Squirrel nests and woodpecker holes are important den sites. Red-backed voles (*Clethrionomys*) and red squirrels (*Tamiasciurus*) are its chief prey, although other small mammals such as flying squirrels and shrews, birds, fish, earthworms, insects, and fruit may also be consumed. Martens have few enemies. Coyotes, fisher, foxes, lynx, cougars, eagles, and owls may take an occasional young marten, but humans are the greatest enemy through trapping and habitat destruction. American martens have been known to survive in captivity up to 17 years and in the wild for 19 years.

American pine marten. The American pine marten is about the size of a small house cat. Its pelage, which ranges in color from yellowish brown to dark brown, is soft and lustrous. The tooth formula is I 3/3, C 1/1, Pm 4/4, M 1/2 × 2 for a total of 38 teeth. Formerly found in the northern spruce and balsam fir forests and the mixed deciduous-coniferous forests south to Virginia and California, the species is now extirpated over much of this area due to habitat destruction and overtrapping. It avoids early and mid-successional forests, as well as forest openings.

Reproduction. They are polygamous with one female accepting more than one male. Most species breed in late summer, with litters of 1–5 sparsely furred, blind, and deaf kits being born during early spring. Delayed implantation is known to occur in some northern species (yellow-throated marten, American marten, sable, fisher). In the American marten, for example, gestation is 220–276 days, although the period of active pregnancy is about 27 days. Weaning occurs between 6 and 8 weeks of age. Kits are able to kill prey when 3–4 months old. Sexual maturity is attained at about 15 months.

Fur. The fur of some martens, especially the sable and fisher, is highly valued. It is used in the manufacture of coats, capes, and jackets as well as for trim and for linings for luxury garments. The fur of the American pine marten is sometimes referred to as "American sable." Overharvesting along with habitat destruction has caused some populations to decline. The status of the Nilgiri marten is considered vulnerable. *See* CARNIVORA; FISHER; SABLE. Donald W. Linzey

Bibliography. S. Buskirk et al. (eds.), *Martens, Sables, and Fishers: Biology and Conservation*, Cornell University Press, Ithaca, New York, 1994; T. W. Clark et al., *Martes americana, Mammalian Species*, no. 289, pp. 1–8, American Society of Mammalogists, 1987; G. A. Feldhamer, B. C. Thompson, and J. A. Chapman, *Wild Mammals of North America: Biology, Management, and Conservation*, 2d ed., Johns Hopkins University Press, 2003; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Maser

A device for coherent amplification or generation of electromagnetic waves by use of excitation energy in resonant atomic or molecular systems. "Maser" is an acronym for microwave amplification by stimulated emission of radiation. The device uses an unstable ensemble of atoms or molecules that may be stimulated by an electromagnetic wave to radiate energy at the same frequency and phase as the stimulating wave, thus providing coherent amplification. Amplifiers and oscillators operating on the same principle as the maser exist in many regions of the electromagnetic spectrum. Those operating in the optical region were once called optical masers, but they are now universally called lasers (the "l" stands for "light"). Amplification by maser action is also observed arising naturally from interstellar gases. *See* COHERENCE; LASER.

Maser amplifiers can have exceptionally low internally generated noise, approaching the limiting effective input power of one-half quantum of energy per unit bandwidth. This limit is often described as one quantum per unit bandwidth, but the other half quantum is better attributed to a zero-point input noise field. The limit is consistent with Heisenberg's uncertainty principle, which prohibits exact simultaneous determination of the amplitude and phase of a wave. Their inherently low noise makes maser oscillators that use a narrow atomic or molecular resonance extremely monochromatic, providing a basis for frequency standards. The hydrogen maser, which uses a hyperfine resonance of a gas of hydrogen atoms as the amplification source, is the prime example of this use. Also because of their low noise and consequent high sensitivity, maser amplifiers are particularly useful for reception and detection of very weak signals in radio astronomy, microwave radiometry, and the like. A maser amplifier was used in the experiments that detected the cosmic microwave radiation left over from the big bang that created the universe. *See* COSMIC BACKGROUND RADIATION; ELECTRICAL NOISE; FREQUENCY MEASUREMENT; RADIO ASTRONOMY; UNCERTAINTY PRINCIPLE.

Principle of operation. The quantum theory describes discrete particles such as atoms or molecules as existing in one or more members of a discrete set of energy levels, corresponding to the various possible internal motions of the particle (vibrations, rotations, and so forth). Thermal equilibrium of an ensemble of such particles requires that the number of particles n_1 in a lower energy level 1 be related to the number of particles n_2 in a higher energy level 2 by the Boltzmann distribution, given by Eq. (1), where E_1 and E_2 are the respective energies

$$\frac{n_1}{n_2} = \exp\frac{(E_2 - E_1)}{kT} \tag{1}$$

of the two levels, k is Boltzmann's constant, and T is the absolute (Kelvin) temperature. Thermal equilibrium also requires that the phases of oscillation of the particles, or the relative phases of the quantum-mechanical wave functions of the various states, be random. A collection of such particles in thermal equilibrium absorbs electromagnetic energy at their resonance frequencies $\nu_{21} = (E_2 - E_1)/h$, where h is Planck's constant. Maser amplification requires upset of the thermal equilibrium condition. *See* BOLTZMANN STATISTICS; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Particles may be stimulated by an electromagnetic wave to make transitions from a lower energy level to a higher one, thereby absorbing energy from the wave and decreasing its amplitude, or from a higher energy level to a lower one, thereby giving energy to the wave and increasing its amplitude. These two processes are inverses of each other, and their effects on the stimulating wave add together. The upward and downward transition rates are the same, so that, for example, if the number of particles in the upper and lower energy states is the same, the stimulated emission and absorption processes just cancel. For any substance in thermal equilibrium at a positive (ordinary) temperature, the Boltzmann distribution requires that n_1 be greater than n_2 resulting in net absorption of the wave. If n_2 is greater than n_1, however, there are more particles that emit than those that absorb, so that the particles amplify the wave. In such a case, the ensemble of particles is said to have a negative temperature T, to be consistent with the Boltzmann condition. If there are not too many counterbalancing losses from other sources, this condition allows net amplification. This is the basic description of how a maser amplifies an electromagnetic wave. An energy source is required to create the negative temperature distribution of particles needed for a maser. This source is called the pump.

Gas masers. The first known maser of any kind was realized in 1954. The amplifying medium was a beam of ammonia (NH_3) molecules, and the molecular resonance used was the strongest of the rotation-inversion lines, at a frequency near 23.87 GHz (1.26-cm wavelength). Molecules from a pressurized tank of ammonia issued through an array of small orifices to form a molecular beam in a meter-long vacuum chamber. Spatially varying electric fields in the vacuum chamber created by a cylindrical array of electrodes formed a focusing device, which ejected from the beam the molecules in the lower energy level and directed the molecules in the upper energy level into a metal-walled electromagnetic cavity resonator. When the cavity resonator was tuned to the molecular transition frequency, the number of molecules was sufficiently large to produce net amplification and self-sustained oscillation. This type of maser is particularly useful as a frequency or time standard because of the relative sharpness and invariance of the resonance frequencies of molecules in a dilute gas. *See* CAVITY RESONATOR; MOLECULAR BEAMS.

The condition for oscillation to occur is given by Eq. (2), where m and e are respectively the mass and

$$n_2 - n_1 \geq \frac{\pi m V \nu \Delta\nu}{f e^2 Q} \tag{2}$$

charge of an electron, V and Q are the volume and quality factor of the cavity resonator, while the resonant molecular response is described by its resonant frequency ν, its full width at half maximum $\Delta\nu$, and its oscillator strength f. The quality factor Q is defined as $2\pi\nu$ divided by the fractional rate of loss of field energy stored in the cavity, due to both internal losses and losses through coupling holes. Low loss translates into high Q. The oscillator strength f is the ratio of the integrated response of a molecule in its lower state to that of a harmonically bound electron. Small oscillator strength requires a proportionately greater number of molecules. The maximum power output is the quantum energy $h\nu$ multiplied by the rate at which molecules enter the cavity (for the ammonia maser, this amounted to a few nanowatts). A wave impinging on the cavity at the molecular transition

frequency is amplified in transmission or reflection if Eq. (2) is satisfied, with Q_0 replacing Q, where $Q_0 > Q$ is the quality factor of the cavity representing only its internal losses.

A gas maser that is currently used as a particularly valuable frequency standard is the hydrogen maser mentioned above. First reported in 1960, this maser is similar to the ammonia maser. A beam of hydrogen atoms in a vacuum chamber passes through a state selector, this time using a magnetic field, which directs the upper-state atoms into a small opening in a quartz bulb. The bulb sits in a silver-plated cavity resonator tuned to the hyperfine transition frequency near 1.42 GHz (21.1-cm wavelength). The inside walls of the quartz bulb are specially coated so that the hydrogen atoms can bounce off them many times with little disturbance of the transition frequency or phase. This property leads to an exceptionally narrow resonance line and thus an exceptionally stable frequency of oscillation. *See* ATOMIC CLOCK.

Solid-state masers. Solid-state masers usually involve the electrons of paramagnetic ions in crystalline media immersed in a magnetic field. At least three energy levels are needed for continuous maser action. The energy levels are determined both by the interaction of the electrons with the internal electric fields of the crystal and by the interaction of the magnetic moments of the electrons with the externally applied magnetic field. The resonant frequencies of these materials can be tuned to a desired condition by changing the strength of the applied magnetic field and the orientation of the crystal in the field. An external oscillator, called the pump, excites the transition between levels 1 and 3 [at the frequency $\nu_{31} = (E_3 - E_1)/h$], equalizing their populations. Then, depending on other conditions, the population of the intermediate level 2 may be greater or less than that of levels 1 and 3. If greater, maser amplification can occur at the frequency ν_{21}, or if less, at the frequency ν_{32}. Favorable conditions for this type of maser are obtained only at very low temperature, as in a liquid-helium cryostat. A typical material is synthetic ruby, which contains paramagnetic chromium ions (Cr^{3+}), and has four pertinent energy levels. *See* PARAMAGNETISM.

The important feature of solid-state masers is their sensitivity when used as amplifiers. Such masers have been made with effective input noise temperatures less than about 5 K. They have usually been made using traveling-wave structures rather than resonant cavities. In this case, the gain per unit length (α) produced by the maser medium placed in the structure is given by Eq. (3), where n_2 and n_1 are the numbers

$$\alpha(\nu) = \frac{2f(n_2 - n_1)\pi^2 e^2}{Amv_g} g(\nu) \qquad (3)$$

of ions per unit length, A is the cross-sectional area of the waveguide structure, v_g is the group (energy) velocity of the microwaves in the waveguide, and $g(\nu)$ is the line shape function, normalized so that its integral over frequency is unity. The other symbols are as in Eq. (2). Slow-wave waveguide structures can increase the gain by reducing the group velocity to about one-hundredth of the velocity of light. Typical values of the other parameters can give amplifier gains of a few decibels per centimeter of length, and maximum output powers in the microwatt range. *See* WAVEGUIDE.

Astronomical masers. Powerful, naturally occurring masers have probably existed since the earliest stages of the universe, though that was not realized until a few years after masers were invented and built on Earth. Their existence was first proven by discovery of rather intense 18-cm-wavelength microwave radiation of the free radical hydroxyl (OH) molecule coming from very localized regions of the Milky Way Galaxy.

Microwave radiation of OH reaching the Earth from various directions in outer space had already been studied for a few years before discovery that some of this radiation comes from areas in the Milky Way Galaxy which are very limited in size. This means that in its own locality the radiation is much too intense to be simply due to warm molecules; it had to be brought up to the observed intensity by the process of stimulated emission, or maser-type action. Soon after recognition that OH radiation came from astronomical masers, it was found that the water molecule in interstellar clouds of gas is also producing intense 1.35-cm-wavelength radiation by maser action. This radiation from water and from OH reaches the Earth from many directions in the Milky Way Galaxy, such as from the Orion Nebula, and from many stars that are surrounded by dust particles and gas. Microwave radiation from water molecules has been found to come even from distant galaxies, some of which are broadcasting this radiation at an intensity as large as a few hundred times that of the total radiation at all wavelengths produced by the Sun.

Continued studies of microwave radiation from outer space have revealed many additional masers in various astronomical sources and from a wide variety of molecules. Radiation by astronomical masers at more than 120 different frequencies is now known, produced by more than a dozen different molecules such as OH, water, ammonia, methyl alcohol, and hydrogen cyanide. Masers due to excited hydrogen atoms are also found. Some of these naturally occurring amplifying systems are also at wavelengths shorter than microwaves. In the infrared region, they are generally called lasers rather than masers. There are even lasers in the atmosphere of Mars, due to carbon dioxide that is excited to high energy levels by solar radiation. *See* INTERSTELLAR MATTER.

Although the natural occurrence of masers in astronomical sources had not been considered before they were discovered somewhat accidentally, in hindsight it should have been expected that conditions in astronomical sources would produce them. Radiation from stars, molecular collisions in stellar clouds that are not in thermal equilibrium, and other sources of energy in astronomical objects can easily excite molecules and atoms. It can produce energy

distributions in them that are very different from those of normal materials in thermal equilibrium, and thus provide conditions for maser-type amplification of waves. Some of the mechanisms for this excitation in astronomical situations are understood. An example is the ionization of hydrogen by ultraviolet stellar radiation and subsequent recombination of protons and electrons into excited energy states of hydrogen. In the case of water-vapor masers, it appears that molecular collisions in strong shock waves of rarefied gases are probably the source of their often very powerful maser action. Some masers are recognized to be characteristic of particular types of astronomical locations. For example, masers produced by silicon monoxide (SiO) molecules are found almost always in gas produced by and closely surrounding stars. However, there are also many astronomical masers for which the mechanisms involved are not yet well understood, and undoubtedly there are a number of new types still to be discovered. These masers have existed for many billions of years, and humans have just recently begun to know of them.

Masers in astronomical objects differ from those generally used on Earth in that they involve no resonators or slow-wave structures to contain the radiation and so increase its interaction with the amplifying medium. Instead, the electromagnetic waves in astronomical masers simply travel a very long distance through astronomical clouds of gas, far enough to amplify the waves enormously even on a single pass through the cloud. It is believed that usually these clouds are large enough in all directions that a wave passing through them in any direction can be strongly amplified, and hence astronomical maser radiation emerges from them in all directions. Although some amplifying clouds may be thinner in one direction than in others, and thus amplify unequally in different directions, it is believed that most astronomical masers radiate broadly in a wide range of directions, and it is on this basis that their total power is estimated.

Naturally occurring masers have been important tools for obtaining information about astronomical objects. Since they are very intense localized sources of microwave radiation, their positions around stars or other objects can be determined very accurately with microwave antennas separated by long distances and used as interferometers. This provides information about the location of stars themselves as well as that of the masers often closely surrounding them. The masers' velocity of motion can also be determined by Doppler shifts in their wavelengths. The location and motion of masers surrounding black holes at the centers of galaxies have also provided information on the impressively large mass of these black holes. Astronomical masers often vary in power on time scales of days to years, indicating changing conditions in the regions where they are located. Such masers also give information on likely gas densities, temperature, motions, or other conditions in the rarefied gas of which they are a part. *See* BLACK HOLE; DOPPLER EFFECT.

Charles H. Townes; James P. Gordon

Bibliography. A. W. Clegg and G. E. Nedoluha (eds.), *Astrophysical Masers*, Springer-Verlag, 1993; T. K. Ishii, *Maser and Laser Engineering*, 1983; A. E. Siegman, *An Introduction to Lasers and Masers*, 1971; A. E. Siegman, *Microwave Solid-State Masers*, 1964; A. Yariv, *Quantum Electronics*, 3d ed., John Wiley, 1989.

Masking of sound

Interference with the audibility of a sound caused by the presence of another sound. More specifically, the number of decibels (dB) by which the intensity level of a sound (signal) must be raised above its threshold of audibility, to be heard in the presence of a second sound (masker), is called the masking produced by the masker on the signal. The masker and the signal may be identical or may differ in frequency, complexity, or time.

Masking between pure tones. Masking effects when both the masker and signal are pure tones and are presented simultaneously are shown in **Fig. 1**. The three curves represent cases in which the frequency of the signal is 300, 1000, or 3000 Hz. The tonal signal is brief, and it is presented at a low sensation level in the temporal middle of a longer-duration tonal masker. The level of the tonal masker required for threshold detection of the signal is the dependent variable shown on the ordinate. As the masker is moved through the range of frequencies shown on the abscissa, the amount of masking (or threshold shift) is plotted in the curves. When the tonal signal and masker have the same frequency, a very low level masker is required to mask the signal, indicating significant masking. As the difference in the frequency between the signal and masker increases, the signal is easier to detect, requiring a higher-level masker to mask the signal. Results from these psychophysical tuning curve measures of masking agree very well with data obtained from single auditory neurons in

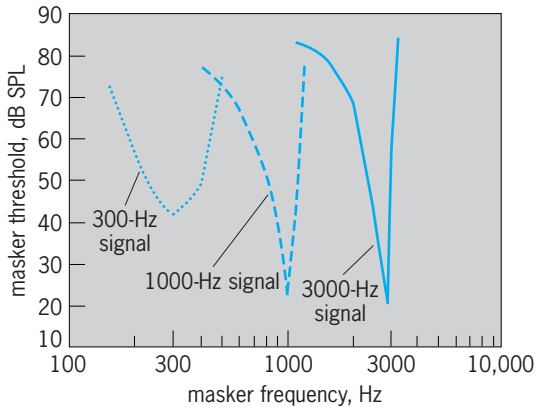

Fig. 1. Three psychophysical tuning curves, showing the tonal masking effects for tonal signals of 300, 1000, or 3000 Hz. (*After F. L. Wightman et al. Factors influencing frequency selectivity in normal and hearing-impaired listeners, in E. F. Evans and J. Wilson, eds., Psychophysics and Physiology of Hearing, Academic Press, London, 1977*)

the auditory periphery, suggesting that tonal masking is mediated by the activity of these peripheral neurons.

Masking between complex sounds. The most widely studied complex masking sound is random noise which has energy at all frequencies and is said to be flat if the level for each 1-Hz bandwidth of the noise is the same. When random-flat noise is used to mask a pure tone, only a narrow frequency band (critical band) of the noise centered at the tonal frequency causes masking. If the bandwidth of the noise is narrower than this critical bandwidth, the tone's intensity can be lowered before the tone is masked. If the bandwidth is wider than this critical bandwidth, further widening of the bandwidth causes no changes in the detectability of the signal. The width of the critical band increases proportionally to its center frequency, that is, to the signal frequency (**Fig. 2**). When noise masks speech, either the detectability of speech or speech intelligibility can be measured. The level for speech intelligibility is about 10–14 dB higher than for speech detectability. Masking by other complex sounds leads to results such as would be expected from the principles shown in Fig. 1 and those based on the existence of critical bands.

Temporal masking. Masking can occur when the signal either precedes or follows the masker in time. In backward masking the signal precedes the masker, while in forward masking the signal follows the masker. **Figure 3** shows the relative amount of masking which exists if the masker is presented for a brief time and the signal is of a shorter duration and appears at various times before, during, and after the masker. As is shown, signal threshold is highest at the temporal edges of the masker (masking overshoot), while more masking is obtained at longer temporal separations between the masker and signal in forward than in backward masking. **Figure 4***a* shows an example of another forward-masking phenomenon. The signal is 1000 Hz and the masker is two tones: a 1000-Hz tone and a second tone whose frequency is shown on the abscissa (Fig. 4*b* shows the temporal relationships). The solid line on the ordinate of Fig. 4*a* represents the signal threshold when only the 1000-Hz tone is the masker (approximately 65 dB). When the frequency of the second tone in the complex two-tone masker is less than 1100 Hz, the second tone yields masking in addition to that obtained for the 1000-Hz masker. When the second tone in the complex two-tone masker is between 1100 and 1600 Hz, the signal threshold is less than that obtained when only the 1000-Hz masker produced the masking. In this frequency region, then, the addition of a second tone to the 1000-Hz masker yields a signal which is easier to detect, and thus its threshold is lower. It is as if the second tone reduces the effectiveness of the 1000-Hz masking stimulus. The physiological basis for this effect, as well as its implications for auditory processing of complex stimuli, is of great interest to the auditory scientist. *See* ACOUSTIC NOISE; HEARING (HUMAN); SOUND. William A. Yost

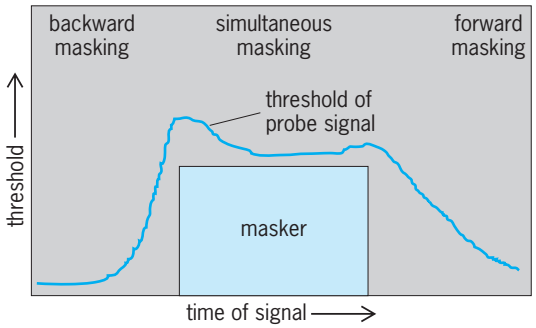

Fig. 3. Relative amount of masking obtained in forward, simultaneous, and backward masking. (*After W. A. Yost, Fundamentals of Hearing: An Introduction, 4th ed., Academic Press, New York, 2000*)

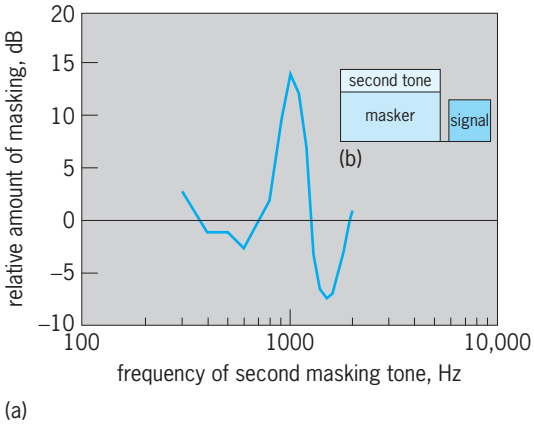

Fig. 4. Masking of a 1000-Hz pure-tone signal in forward masking when the masker is a two-tone complex consisting of a 1000-Hz masker and a second tone. (*a*) Threshold of signal for the two-tone masker relative to that obtained when only the 1000-Hz tone served as the masker, plotted against the frequency of the second masking tone. (*b*) Temporal relationship of signal and maskers. (*After R. V. Shannon, Two-tone unmasking and suppression in a forward-masking situation, J. Acous. Soc. Amer., 59:1460–1471, 1975*)

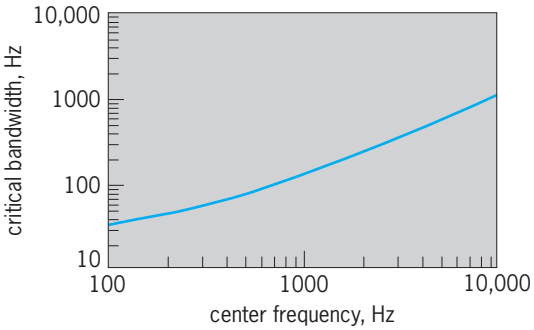

Fig. 2. Width of the critical band in hertz as a function of signal frequency.

Bibliography. W. M. Hartmann, *Signal, Sounds, and Sensation*, Springer-Verlag, New York, 1998; B. C. J. Moore, *Introduction to the Psychology of Hearing*, 4th ed., Academic Press, London, 1997; W. A. Yost, *Fundamentals of Hearing: An Introduction*, 4th ed., Academic Press, New York, 2000.

Masonry

Construction of natural building stone or manufactured units such as brick, concrete block, adobe, glass block, or cast stone that is usually bonded with mortar. Masonry can be used structurally or as cladding or paving. It is strong in compression but requires the incorporation of reinforcing steel to resist tensile and flexural stresses. Masonry veneer cladding can be constructed with adhesive or mechanical bond over a variety of structural frame types and backing walls.

Masonry is noncombustible and can be used as both structural and protective elements in fire-resistive construction. It is durable against wear and abrasion, and most types weather well without protective coatings. The mass and density of masonry also provide efficient thermal and acoustical resistance.

Materials. Brick, concrete block, and stone are the most widely used masonry materials for both interior and exterior applications in bearing and nonbearing construction. Stone masonry can range from small rubble or units of ashlar (a hewn or squared stone) embedded in mortar, to mechanically anchored thin slabs, to ornately carved decorative elements. Hardness dictates both form and cost, and it also affects durability and weathering. Granite, marble, and limestone are the most commonly used commercial building stones. Glass block can be used as security glazing or as elements to produce special daylighting effects. Both clear, colored, and reflective coated glass-block units in smooth and textured surfaces are available. *See* GLASS; GRANITE; LIMESTONE; MARBLE.

Masonry mortar is made from cement, sand, lime, and water. Masonry grout, a more fluid mixture of similar ingredients, is used to fill hollow cores and cavities and to embed reinforcing steel. Anchors and ties are usually of galvanized or stainless steel. Flashing may be of stainless steel, coated copper, heavy rubber sheet, or rubberized asphalt. Weepholes are formed by leaving the mortar out of head joints just above flashing. *See* GROUT; MORTAR.

Wall types. Both brick and concrete block are available in hollow units suitable for walls that are a single unit thick (single wythe). Steel reinforcing bars can be grouted vertically in the hollow cores and horizontally in U-shaped bond beam units. Where walls are required to accommodate electrical conduit or plumbing piping, double-wythe walls with an open or grouted cavity are more appropriate. The backing and facing wythe may be of the same or dissimilar types of masonry.

Veneer applications of masonry are appropriate when the visual appearance of a masonry structure is desired, but a load-bearing wall design is not considered appropriate. Masonry veneers may be used on buildings of wood, steel, or concrete structural frames, as well as cladding on load-bearing masonry buildings. Anchored veneers must be designed with an open cavity for drainage, and require the installation of flashing and weepholes for collection and discharge of moisture.

Important considerations. Masonry walls are permeable. Even when the design is adequate and the workmanship is acceptable, some wind-driven moisture can still penetrate the face shell of single-wythe units or the full thickness of a 4-in. (10-cm) exterior wythe. Cavity walls were initially conceived to provide drainage for this moisture through a system of flashing and weepholes. Cavity walls and anchored veneer walls should have an open separation of at least 2 in. (5 cm) between the exterior masonry wythe and the backing. The open cavity, when it is properly fitted with flashing and weepholes, functions as a drainage system for moisture that penetrates from the exterior or is condensed from water vapor within the wall section. Single-wythe walls and multiwythe solid walls must also be designed with a system of flashing and weepholes to divert moisture to the outside. Moisture protection is maximized when the flashing membrane is continuous, without gaps or voids, and is turned up to form pans at terminations.

Masonry walls are also relatively brittle and contain thousands of linear feet of joints along which cracks can open because of thermal and moisture movements and dissimilar movement between adjacent materials. Concrete products shrink, clay products expand, and metals expand and contract reversibly. Such movement is accommodated through flexible anchorage and the installation of control joints in concrete masonry and expansion joints in clay masonry. Coefficients of thermal expansion and coefficients of moisture expansion can be used to estimate the expected movement of various materials, so that movement joints can be sized and joint sealants selected accordingly. If details do not accommodate the necessary wall movement, excessive moisture can penetrate through the resulting cracks. *See* BRICK; CLAY, COMMERCIAL; CONCRETE; STONE AND STONE PRODUCTS.

Christine Beall

Bibliography. C. Beall, *Masonry Design and Detailing*, 4th ed., 1997; Brick Institute of America, *Technical Notes*; National Concrete Masonry Association, *TEK Notes*; F. A. Robert, R. Schneider, and W. L. Dickey, *Reinforced Masonry Design*, 2d ed., 1987.

Mass

The quantitative measure of the inertia of an object, that is, of its resistance to being accelerated. Thus, the more massive an object is, the harder it is to accelerate it (change its state of motion). Isaac Newton described mass as the "quantity of matter" when formulating his second law of motion. He observed that the rate of change of motion is proportional to the impressed force and in the direction of that force. By "motion" he meant the product of the quantity of matter and the velocity. In modern terms Newton's "motion" is known as momentum. The second law is expressed by Eq. (1), where **F** is the

$$\frac{d\mathbf{p}}{dt} \propto \mathbf{F} \qquad (1)$$

applied force, t the time, and $\mathbf{p}$ the momentum. If the mass of the object is constant, then the rate of change of momentum with time is just the product of the mass with the rate of change of the velocity with time, the acceleration $\mathbf{a}$. For this situation the second law may be written as Eq. (2).

$$\mathbf{a} \propto \frac{\mathbf{F}}{m} \tag{2}$$

See FORCE; LINEAR MOMENTUM; MOMENTUM.

By suitable choice of units the proportionality becomes an equality as shown in Eq. (3). In particu-

$$\mathbf{a} = \frac{\mathbf{F}}{m} \tag{3}$$

lar, the equality holds in the International System of Units (SI), in which acceleration is measured in meters/second, force is in newtons, and the unit of mass is the kilogram (kg). *See* ACCELERATION; UNITS OF MEASUREMENT.

The primary standard of the unit of mass is the international prototype kilogram kept at Le Bureau International des Poids et Mesures in France. The prototype kilogram is a cylinder of a platinum-iridium alloy, a material chosen for its stability. The standard of mass is kept under conditions specified initially by the first Conférence Général des Poids et Mesures in 1889. Secondary standards are maintained in laboratories around the world and are compared with the mass of the prototype by means of balances whose precision reaches 1 part in 10^8 or better. *See* PHYSICAL MEASUREMENT.

Measuring mass. The mass that is described through the second law is known as inertial mass. The inertial mass of any object may be measured by comparing it to the inertial mass of another object. For example, two objects may be made to collide and their accelerations measured. According to Newton's third law of motion, both objects experience forces of equal magnitude. Consequently the ratio of their masses will be related to their accelerations through Eq. (4), where a_1 and a_2 are the magnitudes of ac-

$$\frac{m_1}{m_2} = \frac{a_2}{a_1} \tag{4}$$

celerations. Thus, if the value of one mass is known, the value of the other mass may be determined by experiment.

Masses may be compared or measured by means other than collisions. Systems that oscillate do so with a frequency that depends on their mass. For example, the frequency of a mass oscillating at the end of a spring decreases as the mass is increased. This behavior is utilized to construct inertial balances, devices that measure the inertial mass. *See* HARMONIC MOTION.

Mass and weight. The most common way of measuring mass is by weighing, a process that involves the effect of gravity. With a simple two-pan balance, two masses are compared by comparing their weights. The weight of an object on Earth is the gravitational force of attraction exerted on it by the Earth. The response of the balance to the two masses is really a response to their weights. As a result, in commercial and common use, weight is often expressed in units of kilograms instead of the correct units of newtons. *See* GRAVITATION; WEIGHT.

The gravitational force that leads to weight is not confined to just the interaction between the Earth and objects at its surface. Indeed, the law of universal gravitation, first formulated by Newton in 1684, describes an attractive force that exists between all objects and is proportional to the product of their masses. In effect, mass is the source of the gravitational force of attraction. Consequently, there is another interpretation of mass, one that is based on gravitational force and not on inertia. It is called gravitational mass to distinguish it from inertial mass.

Newton recognized that he had two distinct concepts of mass and devised an experimental test of the equivalence of inertial and gravitational mass. Using crude methods he was unable to detect any difference between the two types of mass. In 1890, Lorand Eötvös devised a sensitive method to test the equivalence of inertial and gravitational mass. He used objects of different materials to compare the effect of the Earth's gravitational force on them with the inertial effects of the Earth's rotation. Eötvös concluded that the two types of mass were the same. More recent experiments with even greater sensitivity continue to support the idea that inertial mass and gravitational mass are identical. The principle of equivalence is the assertion that both masses are exactly equal. Albert Einstein employed the principle of equivalence in his development of the general theory of relativity. *See* RELATIVITY.

Weightlessness and microgravity. Near the Earth's surface, when air resistance is removed, all objects fall with the same acceleration. This constant acceleration is known as the acceleration of gravity, g, and has the standard value 9.80665 m/s^2. Although g is the same for all objects in a given locality, it does vary slightly with altitude and latitude. Newton's second law may be used to relate the gravitational force (the weight) to the acceleration. Thus, the weight w may be expressed by Eq. (5), which gives the relationship

$$w = mg \tag{5}$$

between mass and weight. Weight is a force proportional to the mass of an object and g is the constant of proportionality.

When you stand on a scale, the scale reading gives the magnitude of your weight. Most of the force that comprises your weight is due to the gravitational attraction between you and the Earth. However, because the Earth rotates, your weight is slightly less than it would be if the Earth were not rotating. Consequently, a more precise definition of weight is as follows: The weight of a object in a specified reference frame is the force that, when applied to the object, would give it an acceleration equal to the local acceleration of free fall in that reference frame. On the Earth, the local acceleration of free fall is g. If an object is taken to the Moon, the force of gravity

exerted on it by the Moon is less than the force of gravity on it when it was on Earth. Thus, its weight on the Moon is less than its weight on Earth, even though its mass remains unchanged. *See* REFERENCE FRAME.

An astronaut inside an orbiting spacecraft experiences a condition called weightlessness. In orbit, both the spacecraft and the astronaut are in free fall. They are both being accelerated by gravity at the same rate and there is no noticeable attraction between them. The astronaut floats around freely within the spacecraft. If the astronaut were to stand on a scale attached to the spacecraft, the scale would record essentially zero weight. If an object (including the astronaut) is placed at a location a meter or so from the center of mass of the spacecraft, it will actually drift slightly relative to the center of mass. The drift occurs because the object and the spacecraft experience slightly different accelerations due to the Earth's gravity because of the slight difference in radial distance from the Earth and the corresponding slight difference in orbital path. For the astronaut in a low Earth orbit, 300 km (186 mi) above the surface, this relative motion would seem to be due to a very small force. An object displaced a few meters from the center of mass of the spacecraft would have an acceleration relative to the spacecraft of approximately one millionth the value of g at the Earth's surface. Consequently, this condition of near weightlessness is known as microgravity. *See* FREE FALL; WEIGHTLESSNESS.

In monitoring the health of the crew in sustained orbit, the body mass of the astronauts provides important information. On Earth, their weight would be measured using scales. In the microgravity of orbit, their weight is meaningless. Instead, astronauts measure their mass with an inertial balance that consists of a chair that moves relative to the spacecraft with a simple back-and-forth motion similar to that of a porch glider. The mass is computed from measurements of the frequency of the chair's motion.

Mass and energy. For many years it was thought that mass was a conserved quantity; that is, that it can neither be created nor destroyed. During the eighteenth and nineteenth centuries much of the development of chemistry was based on the conservation of mass. Careful measurements of the mass of constituents of chemical reactions and their products showed that the quantity of matter was unchanged by chemical reactions. However, in the twentieth century, observations and measurements in nuclear and particle physics demonstrated conclusively that mass is not a conserved quantity. For example, radioactive polonium-210 decays into lead-206 by the emission of an energetic alpha particle (a helium nucleus). The combined mass of the lead nucleus and the alpha particle is less than the mass of the original polonium nucleus. The difference in the masses appears as the kinetic energy of the alpha and the lead nucleus. This correspondence between mass and energy was shown by Einstein in Eq. (6). In the case of the alpha decay of polonium, the loss of mass during the decay results an increase in the kinetic energy of the decay products. *See* CONSERVATION OF MASS.

$$E = mc^2 \tag{6}$$

There are many other examples that demonstrate the nonconservation of mass. One of these is the annihilation process that occurs when matter and antimatter combine. An electron can combine with its antiparticle, a positron, and the two annihilate each other. The two massive particles disappear and a pair of massless photons (the quanta of light) emerge. In a coordinate system in which the electron-positron pair is at rest, the two photons are emitted in opposite directions with identical energies as required by momentum conservation. These two photons carry away the energy that was present in the form of the mass of the electron and the positron. As a result, each photon has an energy equivalent to the mass of a single electron. That energy may be found from Eq. (6). Inserting $m = 9.11 \times 10^{-31}$ kg and $c = 2.997 \times 10^8$ m/s gives an energy value of 8.19×10^{-14} joules. In units of electronvolts (eV) the photon energy is 5.11×10^5 eV or 0.511 MeV. *See* ANTIMATTER; ELECTRONVOLT.

The energies released in chemical reactions are typically a fraction of an electronvolt and the resulting mass changes are thus approximately a millionth of the mass of an electron. Because the electron mass is only a small fraction of the atomic mass, the change in mass-energy in chemical reactions is only at the level of parts per billion. The practical result is that on the scale of chemical reactions mass appears to be conserved. At much higher energies, however, the changes in mass-energy become significant. It is the energy release in the fission of heavy elements such as uranium and plutonium that makes nuclear reactors practical sources of energy. *See* NUCLEAR FISSION.

In the present model for understanding fundamental particles and their interactions, that is, the standard model, the masses of particles arise from the energy of the force fields of their constituents. The masses of the nuclear particles, the protons and neutrons, are thought to arise from the interactions of subnuclear particles called quarks. The mass of a proton, for example, comes at least in part from the energy of the force fields of the quarks.

When a particle is accelerated it gains kinetic energy as its speed and momentum increase. As its speed approaches the speed of light, its momentum increases without limit. Further increases in kinetic energy result in increases in momentum but the speed changes more slowly as it approaches the limiting speed, the speed of light. A change of kinetic energy (or work done on the particle) at high speeds does not produce the same change of speed that would result from the identical increment of energy when the particle is moving at a lower speed. In the early twentieth century, the particle was generally regarded as becoming more massive as its speed approached the speed of light. However, the current practice is to take the view that the mass is invariant, unlike the m used in Eq. (6), and is related to

the energy through Eq. (7). Here, E represents the

$$m = \sqrt{\frac{E^2}{c^4} - \frac{p^2}{c^2}} \tag{7}$$

total energy of the particle of mass m moving with momentum p. As the energy, and hence the speed, of the particle increases, the momentum changes according to Eq. (8). In this view, the mass m is the

$$p = \frac{mv}{\sqrt{1 - v^2/c^2}} \tag{8}$$

invariant mass and is a constant. Then at speeds near the speed of light, changes in energy result in negligible changes in speed even though the changes in momentum are large. *See* FLIGHT; MOMENTUM.

Edwin R. Jones

Bibliography. A. Hobson, Teaching $E = mc^2$: Mass without mass, *Phys. Teacher*, 43(2):80–82, 2005; K. R. Symon, *Mechanics*, 3d ed., Addison-Wesley, Boston, 1971; B. N. Taylor (ed.), *The International System of Units (SI)*, NIST Spec. Publ. 330, 2001; E. F. Taylor and J. A. Wheeler, *Spacetime Physics*, 2d ed., Freeman, New York, 1992; F. Wilczek, Whence the force of $F = ma$? II: Rationalizations, *Phys. Today*, 57(12):10–11, December 2004.

Mass defect

The difference between the mass of an atom and the sum of the masses of its individual components in the free (unbound) state. The mass of an atom is always less than the total mass of its constituent particles; this means, according to Albert Einstein's well-known formula, that an energy of $E = mc^2$ has been released in the process of combination, where m is the difference between the total mass of the constituent particles and the mass of the atom, and c is the velocity of light.

The mass defect, when expressed in energy units, is called the binding energy, a term which is perhaps more commonly used. *See* NUCLEAR BINDING ENERGY.

W. W. Watson

Mass–luminosity relation

The relation, observed or predicted by theory, between the quantity of matter a star contains (its mass) and the amount of energy generated in its interior (its luminosity). Because of the great sensitivity of the rate of energy production in a stellar interior to the mass of the star, the mass–luminosity relation provides an important test of theories of stellar interiors. This predicted relation is a corollary of the more general principle that if a star is in equilibrium—it is neither expanding nor contracting—all its other properties depend only on its mass and the distribution of chemical elements within it. For a family of stars with different masses but with the same mixture of chemical elements uniformly distributed throughout the stellar volumes, there should be a unique mass–luminosity relation. The observed relation, obtained from binary stars for which masses and luminosities can be observationally evaluated, is shown in the **illustration**. The majority of stars placed on such a plot are found along this relation, called the main sequence, and it implies that all stars begin their lives with roughly the same chemical composition. It conforms reasonably well to theory. It was this agreement that showed the pioneering studies of stellar structure by A. S. Eddington in the 1920s to be on a sound basis. *See* BINARY STAR; STAR.

Effects of stellar evolution. There are classes of stars not on the main sequence that depart markedly from the mass–luminosity relation for main-sequence stars. The theory of stellar evolution predicts that a star's interior chemical composition will change as a result of the conversion of hydrogen into helium and heavier elements in the process of nuclear energy generation, which is the source of stellar luminosity. Accompanying these changes in composition, there will be changes in luminosity. Examples are the three stars in the illustration lying below the smooth relation. These are white dwarf stars, which are known to be extremely underluminous for their masses, and are thought to be stars in which all sources of nuclear energy have been exhausted. Giants and supergiants are thought to be in an intermediate state of evolution, between main-sequence stars and white dwarfs, with central nuclear energy generation that is markedly different from main-sequence stars. One of the triumphs of twentieth-century astrophysics was the establishment and exploration of stellar evolution theory, which sees stars evolving off the main sequence as they age, and links an individual star's entire evolutionary roadmap to its initial mass and composition. *See* CARBON-NITROGEN-OXYGEN CYCLES; GIANT STAR; HERTZSPRUNG-RUSSELL DIAGRAM; NUCLEOSYNTHESIS; PROTON-PROTON CHAIN; SUPERGIANT STAR; WHITE DWARF STAR.

Stellar mass and luminosity determination. Stellar masses are determined observationally from binary stars. By mapping the orbits of the two components of a binary system, the mass of each star (subject to a correction for the angle at which we are viewing) may be determined from the dynamics of the orbit. Additionally, through comparison of the appearance of that orbit upon the sky with its true physical size, the distance to the binary system may be established. From the apparent brightness and the distance, the actual brightness of the system can be determined, which gives a measure of the luminosity of each component. *See* CELESTIAL MECHANICS; KEPLER'S LAWS; ORBITAL MOTION.

Historically, binary star orbits have been determined though observations of eclipsing binary stars, since we know the angle at which we are observing them. More recently, astronomical interferometers have enabled high-resolution (at scales of milliarcseconds to microarcseconds) mapping of binary

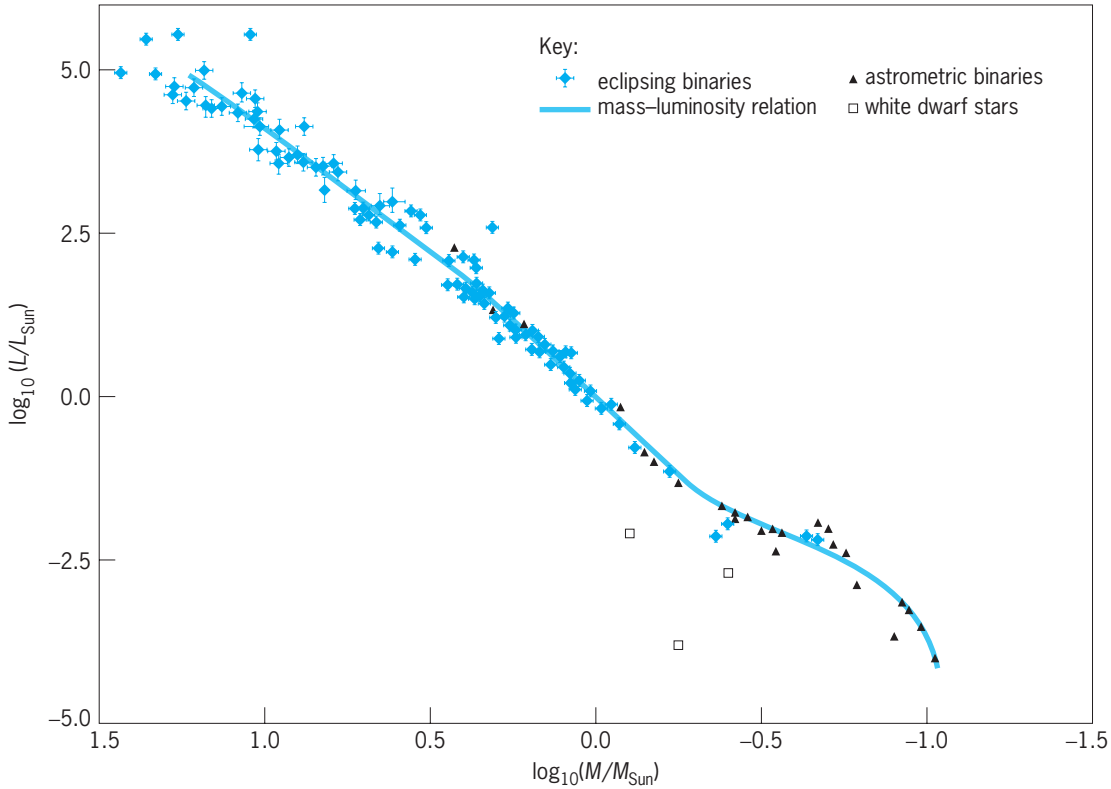

Mass-luminosity relation, where the abscissa is the logarithm of mass *M* and the ordinate is the logarithm of luminosity *L*, both relative to the Sun as unit. This relation is typically plotted with the brightest, most massive stars at the upper left, and the dimmer, least massive stars at the lower right.

star orbits for those binaries that do not happen to eclipse, which has allowed greater discretion in selecting binary stars for observation. Observations of these astrometric binaries (also referred to as visual binaries) has assisted in the observation of orbits of lower-mass binary stars, which are populous in the galaxy, but dim and limited to the few nearest examples for observation. Agreement with the observed mass–luminosity relation is one of the key tests of stellar models that are produced in modern astrophysics. *See* ASTROMETRY; ECLIPSING VARIABLE STARS.

Evolution of close binary systems. If a star is a member of a close binary system, it may be prevented by the presence of its companion from evolving by expansion into a giant star. The evolutionary histories of stars in such systems are predicted to involve mass exchange between the components. The relation between masses and luminosities will depend upon the state of evolution of the individual system. In some cases, gross departures from the mass–luminosity relation for main-sequence stars are predicted and observed. A star that has lost the bulk of its mass to its companion may radiate as much energy as a normal star, even though its mass is much less. X-ray binaries are thought to be extreme examples of advanced evolution of close binary systems. Application of the mass–luminosity relation to the visible components in x-ray binaries leads to the conclusion that in a couple of cases the invisible component may well be a black hole; these systems are the best candidates known for this extreme among stellar configurations. A number of the predictions of the evolution of close binaries appear to be verified observationally. *See* ASTROPHYSICS, HIGH-ENERGY; BLACK HOLE; STELLAR EVOLUTION; X-RAY ASTRONOMY.

Daniel M. Popper; Gerard van Belle

Bibliography. C. W. De Loore, *Structure and Evolution of Single and Binary Stars*, 1992; R. Kippenhahn and A. Weigert, *Stellar Structure and Evolution*, 1991; D. Prialnik, *An Introduction to the Theory of Stellar Structure and Evolution*, 2000; J. Sahade (ed.), *The Realm of Interacting Binary Stars*, 1993.

Mass number

The mass number *A* of an atom is the total number of its nuclear constituents, or nucleons, as the protons and neutrons are collectively called. The mass number is placed before and above the elemental symbol, as in ^{238}U. Because of the approximate equality of the proton and neutron masses, and the relative insignificance of that of the electron, the mass number gives a useful rough figure for the atomic mass; for example, $^{1}H = 1.00783$ atomic mass units (amu), $^{238}U = 238.051$ amu, and so on. The mass number is reduced by four during alpha emission, but is not altered during beta decay or electron capture. *See* ATOMIC NUMBER.

Henry E. Duckworth

Mass spectrometry

An analytical technique for identification of chemical structures, determination of mixtures, and quantitative elemental analysis, based on application of the mass spectrometer. Determination of organic and inorganic molecular structure is based on the fragmentation pattern of the ion formed when the molecule is ionized; further, because such patterns are distinctive, reproducible, and additive, mixtures of known compounds may be analyzed quantitatively. Quantitative elemental analysis of organic compounds requires exact mass values from a high-resolution mass spectrometer; trace analysis of inorganic solids requires a measure of ion intensity as well. *See* MASS SPECTROSCOPE.

Methods of ion production. For analysis of organic compounds the principal methods are electron impact, chemical ionization, field ionization, field desorption, particle bombardment, laser desorption, and electrospray.

Electron impact. When a gaseous sample of a molecular compound is ionized with a beam of energetic (commonly 70-V) electrons, part of the energy is transferred to the ion formed by the collision, as shown in reaction (1). For most molecules, the production of cations is favored over the production of anions by a factor of about 10^4, and the following discussion pertains to cations. The ion corresponding to the simple removal of the electron is commonly called the molecular ion and normally will be the ion of greatest m/e ratio in the spectrum. In the ratio, m is the mass of the ion in atomic mass units (daltons) and e is the charge of the ion measured in terms of the number of electrons removed (or added) during ionization. Occasionally the ion is of vanishing intensity, and sometimes it collides with another molecule to abstract a hydrogen or another group. In these cases, an incorrect assignment of the molecular ion may be made unless further tests are applied. Proper identification gives the molecular weight of the sample.

$$\mathrm{A{-}B{-}C} + e \longrightarrow \mathrm{A{-}B{-}C^+} + e + e \qquad (1)$$

The remaining techniques were devised generally to circumvent the problem of the weak or vanishingly small intensity of a molecular ion.

Chemical ionization. Here the ions to be analyzed are produced by transfer of a heavy particle (H^+, H^-, or heavier) to the sample from ions produced from a reactant gas. Frequently the reactant gas is methane at pressures of 0.2–2.0 torr (27–270 pascals). As above, the initial process in methane (CH_4) upon electron impact is ionization to yield CH_4^+ ions; some of these have enough energy to fragment to $CH_3^+ + H$. These ions in turn react with neutral methane as in reactions (2) and (3). The resulting ions are strong Brønsted acids and react by proton transfer as in reactions (4) and (5) to ionize the molecule of interest. The $C_2H_5^+$ ion also reacts by hydride abstraction, as in reaction (6). Other gases may also be used for ionization; hydrogen (H_2), water (H_2O), ammonia (NH_3), and isobutane (C_4H_{10}) are common. In these the reactive ions are H_3^+, H_3O^+, NH_4^+, and $C_4H_9^+$ [the conjugate acid of $(CH_3)_2C{=}CH_2$], respectively. The proton affinity (**Fig. 1**) is the negative of the enthalpy change for the solvation of a proton by the compound, as in reaction (7). It follows that $\mathrm{A{-}B{-}C{-}H^+}$ protonates molecules with greater proton affinities in the absence of kinetic complications; such reactions are exothermic. Proton affinities have been determined both by high-pressure techniques akin to chemical ionization and by the ion cyclotron resonance technique. Thus H_3^+ transfers the most energy when it protonates a molecule; NH_4^+ transfers the least of the four examples. If the energy transferred (typically 10–50 kcal/mole or 40–200 kilojoules/mole) is great enough, fragmentation can occur, but there is much less in chemical ionization than in electron impact mass spectra.

$$\mathrm{CH_4^+} + \mathrm{CH_4} \longrightarrow \mathrm{CH_5^+} + \mathrm{CH_3} \qquad (2)$$

$$\mathrm{CH_3^+} + \mathrm{CH_4} \longrightarrow \mathrm{C_2H_5^+} + \mathrm{H_2} \qquad (3)$$

$$\mathrm{CH_5^+} + \mathrm{A{-}B{-}C} \longrightarrow \mathrm{H{-}A{-}B{-}C^+} + \mathrm{CH_4} \qquad (4)$$

$$\mathrm{C_2H_5^+} + \mathrm{A{-}B{-}C} \longrightarrow \mathrm{H{-}A{-}B{-}C^+} + \mathrm{C_2H_4} \qquad (5)$$

$$\mathrm{C_2H_5^+} + \mathrm{A{-}B{-}C{-}H} \longrightarrow \mathrm{A{-}B{-}C^+} + \mathrm{C_2H_6} \qquad (6)$$

$$\mathrm{A{-}B{-}C} + \mathrm{H^+} \longrightarrow \mathrm{A{-}B{-}C{-}H^+} \qquad (7)$$

In addition to heavy-particle transfer of other types (transfer of CH_3^+, $C_2H_5^+$, Cl^-), another important method is charge exchange. For a gas with an ionization potential greater than that of the molecule of interest, reaction (8) is exothermic, and the excess energy is transferred as internal energy. Helium ion

$$\mathrm{G^+} + \mathrm{A{-}B{-}C} \longrightarrow \mathrm{A{-}B{-}C^+} + \mathrm{G} \qquad (8)$$

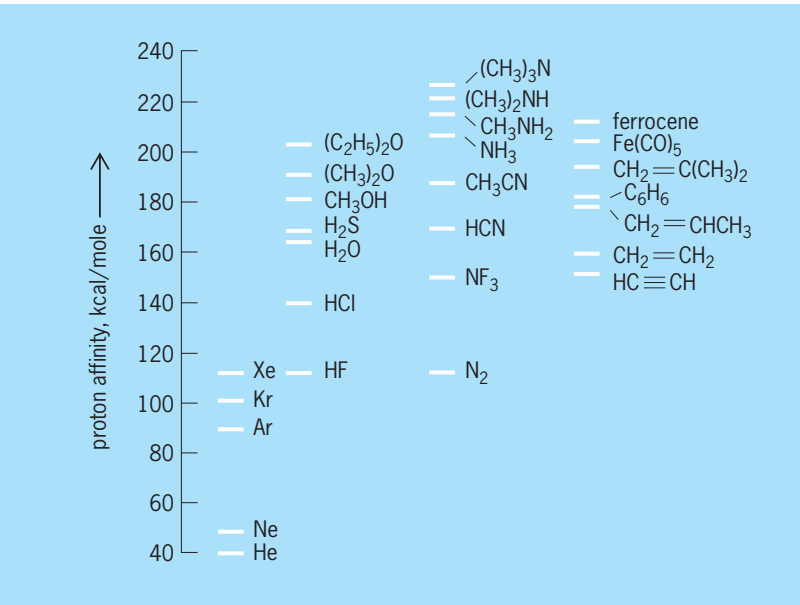

Fig. 1. Representative gas-phase proton affinities of molecules.

(He^+), neon ion (Ne^+), and argon ion (Ar^+) transfer large amounts of internal energy and molecular ions are weak; carbon monoxide (CO) and nitrogen monoxide (NO) are useful gases for this process. Occasionally a combination of chemical ionization and charge exchange is used, with mixtures of Ar and H_2O yielding information about both the molecular weight and important fragments, for example.

The term negative chemical ionization is usually intended to include all ionization processes yielding negative ions under source pressure conditions characteristic of chemical ionization. In some cases, ions may be formed by a true chemical ionization process like proton transfer, as in reaction (9), in

$$RCOOH + OH^- \longrightarrow RCOO^- + H_2O \qquad (9)$$

which OH^- is the reagent ion, produced from H_2O or $N_2O + CH_4$. In others, the reagent gas only mediates the energy of free electrons to low values compatible with capture by neutral molecules to yield negative molecular ions. Molecules with several electronegative atoms have large cross sections for this latter process. Thus the direct analysis of polychlorinated aromatics, and the analysis of small peptides after derivatization by pentafluorobenzoylation, are possible in the femtogram region.

There are applications of ions formed by chemical ionization (atmospheric pressure ionization; plasma chromatography, in which the drift time of the ion through a flowing gas is measured) in which the sample is not depleted before ionization. Because of this preservation of sample molecules, they also have femtogram sensitivities. *See* IONIZATION POTENTIAL; REACTIVE INTERMEDIATES.

Field ionization and field desorption. For less volatile material, the sample is ionized when it is in a very high field gradient (several volts per angstrom) near an electrode surface. The molecular potential well is distorted so that an electron tunnels from the molecule to the anode. The ion thus formed is repelled by the anode (**Fig. 2**). Typically, the lifetime of the ion in the mass spectrometer source is much less (10^{-12} to 10^{-9} s) than in electron impact. Because little energy is transferred as internal energy and the ion is removed rapidly, little fragmentation occurs, and the molecular weight is more easily determined. Field ionization is also used in the time-resolved study of ion fragmentation and rearrangement called field ionization kinetics. This technique permits determination of the fragmentation products at specific times from about 10^{-11} to 10^{-8} s after ionization by energy-focusing of ions at different points in the field ionization source. In this way, simple fragmentations uncomplicated by rearrangements of hydrogen or other atoms in the molecular ion may be observed at the shortest times yet used for sampling ions, and complex rearrangements of ions may be defined from studies over a range of times.

Cations, most commonly lithium ion (Li^+) but often other alkali metals, may be field-desorbed from a salt coating on a wire and pass through a gaseous or adsorbed sample M. MLi^+ ions are produced in some cases where simple field desorption fails.

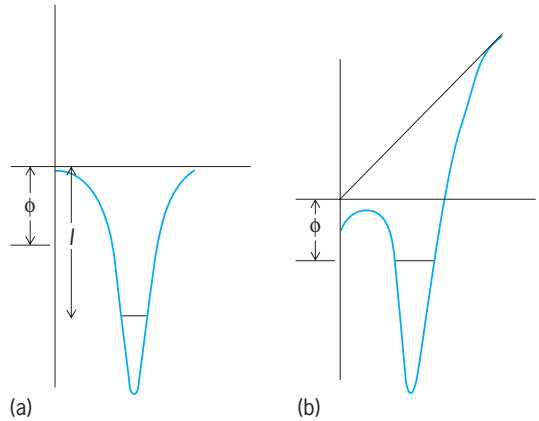

Fig. 2. Potential well of a molecule: (*a*) undistorted, with ionization potential *I* near a metal with work function ϕ in the absence of a field gradient; and (*b*) in a strong field gradient. Near the metal surface the external field raises the most weakly bonded electron to the Fermi level so that tunneling through the small barrier may occur.

Electrohydrodynamic ionization. A high electric-field gradient induces ion emission from a droplet of a liquid solution, that is, the sample and a salt dissolved in a solvent of low volatility. An example is the sample plus sodium iodide (NaI) dissolved in glycerol. The spectra include peaks due to cationized molecules of MNa^+, $MNa(C_3H_8O_3)_n^+$, and $Na(C_3H_8O_3)_n^+$.

Rapid heating methods. A sample heated very rapidly may vaporize before it pyrolyzes. Techniques for heating by raising the temperature of a source probe on which the sample is coated by 200 K (360°F)/s have been developed. Irradiation of an organic sample with laser radiation can move ions of mass up to 1500 daltons into the gas phase for analysis. This technique is the most compatible with analyzers that require particularly low pressures such as ion cyclotron resonance. Time-resolved spectra of surface ejecta are proving to be the most useful kinds of laser desorption spectra available.

Particle bombardment techniques. A solid sample or a sample in a viscous solvent such as glycerol may have ions sputtered from its surface by bombardment with accelerated electrons, ions, or neutrals. Bombardment by electrons is achieved simply by inserting a probe with the sample directly into the electron beam of an electron impact source (in-beam electron ionization). The energetic ions may likewise be the plasma in a chemical ionization source (in-beam chemical ionization), or alternatively, ions in a beam of 2.5-kV Ar^+ directed to a surface coated with the sample (without the coating, this is the inorganic surface analysis technique of secondary ion mass spectrometry). With charge exchange between accelerated Ar^+ (most commonly 6–8 keV) and thermal Ar, a beam of accelerated Ar atoms is produced; desorption of surface layers of glycerol solutions of biological molecules up to 40,000 molecular weight or inorganic salts ($Cs_xI_{(x-1)}^+$ up to 100,000) produces useful spectra by fast atom bombardment. The simplicity of application of this broadly useful technique is

associated with the long duration of spectra because of the constantly regenerated surface of the solution after bombardment, and not so much (as originally thought) with the avoidance of surface charging by using atoms instead of molecules; beams consisting partially or wholly of accelerated Ar^+ ions produce spectra from viscous solutions in techniques more precisely called liquid secondary ion mass spectrometry (SIMS), or liquid-target SIMS. The most powerful method of this variety is the californium-252 (^{252}Cf) plasma desorption source. It must be noted that nomenclature is not unified in this area; some authors also refer to ion-beam chemical ionization as plasma desorption.

The ^{252}Cf source produces fission fragments that penetrate a thin foil on which a film of sample has been coated. The fission fragment deposits energy by electronic excitation in a cylindrical track with a 20-nanometer diameter. Molecular, $(M + 1)^+$, and $(M - 1)^-$ ion ejection from the surface then occurs as the energy is dissipated through the remaining lattice in a shock wave. Molecular ions of mass up to 40,000 daltons have been observed with this technique. *See* SHOCK WAVE.

Laser desorption. The energy of laser photons may be used to remove an analyte from a surface and ionize it for mass-spectrometric analysis. In virtually all cases involving polar compounds of high molecular weight, the analyte is prepared in or on an organic matrix that is coated on the surface to be irradiated. The matrix material assists in the ionization process; the technique is known as matrix-assisted laser desorption ionization (MALDI). Numerous organic materials have been investigated. Molecular weights in excess of 700,000 have been measured with this technique, and the technique is particularly suited for the molecular-weight determination of large polar biological molecules; for example, enzymes and intact antibodies have been analyzed.

Electrospray ionization. A solution sprayed through a nozzle of very small diameter into a vacuum with an electric field having a gradient of several hundred to a thousand volts per centimeter produces gaseous ions from solutes effectively. Electrospray ionization is the only mass-spectrometric technique that produces a large fraction of multiply charged ions from an organic or biological analyte. Since mass spectrometers measure mass-to-charge ratios of ions, not simply their mass, electrospray ionization has the advantage of permitting ions of very high mass to be analyzed without special mass-analysis instrumentation; for example, an ion of mass 120,000 daltons carrying 60 positive charges appears at mass-to-charge 2000, within the range of many mass analyzers. This technique has been used to measure the masses of ions from molecules of masses up to about 200,000 daltons. Since the distribution of charged species reflects to some extent the degree of protonation in solution, a signal that reflects the folding of a protein, there is evidence that solution conformations of proteins can be studied by electrospray ionization mass spectrometry.

Multiphoton ionization of gaseous molecules. If absorption of several photons from a laser produces an excited state of the molecule, molecules accumulate in this state so that absorption of further photons produces especially abundant ions (resonant multiphoton ionization) in contrast to the case where no excited state is reached as an intermediate (nonresonant multiphoton ionization). This tunability for a certain energy to achieve selective ionization of a particular molecule suggests that analysis of mixtures for especially targeted components would be possible. The difficulty is that for even modestly complex molecules the number of states in rotationally and vibrationally broadened bands is large, and the search for selectivity for target molecules has led to use of supersonic nozzle introduction of the sample, where extreme cooling will yield states with insignificant population of higher rotational and vibrational levels, and ideally, electronic levels of nearly line width; thus interference due to excitation of bands of contaminants of the target will be greatly reduced.

Assignment of empirical formula. If the spectrum is a high-resolution spectrum, the deviation of the molecular weight from an integral value is used to determine the elemental composition (for example, $^{12}C_{12}{}^{1}H_{10}$ has molecular weight 154.0782; $^{12}C_9{}^{1}H_{14}{}^{16}O_2$, 154.0994; $^{12}C_{10}{}^{1}H_{16}{}^{16}O$, 154.1358; and $^{12}C_{11}{}^{1}H_{22}$, 154.1721). If it is low-resolution, then the worker takes advantage of the natural abundances of isotopes (for example, ^{12}C, 98.9%; ^{13}C, 1.1%; ^{35}Cl, 75.4%; ^{37}Cl, 24.6%; ^{79}Br, 50.6%; and ^{81}Br, 49.4%) to calculate at least part of the empirical formula by application of the binomial coefficients. For example, if the intensity of the (M + 1) peak in such a spectrum (**Fig. 3**) is 11% that of the M peak, only ions containing one ^{13}C ion contribute to the (M + 1) peak, and the probability that any C chosen randomly is ^{13}C is 1.1%. If 10 C atoms are examined as a group, the probability that one is ^{13}C is 10 times greater, or 11%; hence the molecule contained 10 carbons.

Fragmentation patterns. Since the amount of energy transferred to the molecule is much more than that required simply to ionize it or to break some of the bonds in the remaining ion, some of the molecules fragment after ionization, by competing consecutive decompositions, as indicated in reactions

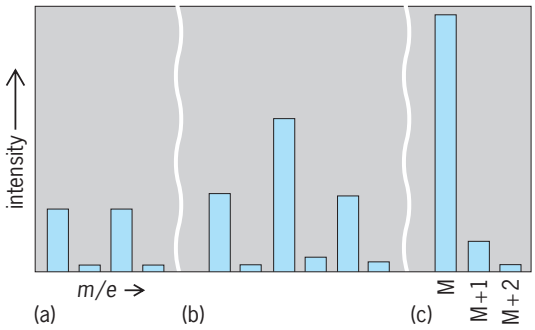

Fig. 3. Use of isotope distribution to determine elemental composition. (*a*) Molecular ion for a compound containing one bromine atom. (*b*) Molecular ion for a compound containing two bromine atoms. (*c*) Molecular ion for a compound containing neither bromine nor chlorine.

(10)–(15). These ions are separated by mass in

$$A{-}B{-}C^{+\cdot} \longrightarrow A{-}B^{+} + C\cdot \qquad (10)$$

$$A{-}B{-}C^{+\cdot} \longrightarrow A{-}B\cdot + C^{+} \qquad (11)$$

$$A{-}B{-}C^{+\cdot} \longrightarrow A\cdot + B{-}C^{+} \qquad (12)$$

$$A{-}B^{+} \longrightarrow A + B^{+} \qquad (13)$$

$$B{-}C^{+} \longrightarrow B^{+} + C \qquad (14)$$

$$A{-}B{-}C^{+\cdot} \longrightarrow A{-}C^{+\cdot} + B \qquad (15)$$

the spectrometer and produce the mass spectrum. Although fragments at almost every mass are usually produced, the most favorable routes for decomposition, which give the most intense peaks in the spectrum, are characteristic of the functional groups in the molecule. Therefore, the intense ions, particularly those at the high-mass end of the spectrum, are used in the assignment of structure to compounds. The common mechanistic rationalizations used for interpretation assume that the charge is localized at the functional group, because the electron lost is most likely to be a π electron or a nonbonding electron. Other less empirical explanations, notably the statistical approach of the quasi-equilibrium theory of mass spectra, have been outlined in the literature. *See* CHEMICAL BONDING.

The reactions of functional groups bearing electronegative atoms are typically loss of the electronegative atom or the group containing it [X = Cl, Br, I, OR, SR, acyl, in reaction (16)] or loss of a substituent group on the α-carbon [R = H or alkyl, X = OR, NR_2, SR, Cl, Br, in reaction (17); R = H, alkyl, or aryl, X = H, OH, OR, NR_2, Cl, Br, I, alkyl, or aryl, in reaction (18)]. These are simple fragmentations

$$R{-}\overset{+\cdot}{X} \longrightarrow R^{+} + X\cdot \qquad (16)$$

$$R{-}{-}CH_2{-}\overset{+\cdot}{X} \longrightarrow R\cdot + CH_2{=}X^{+} \qquad (17)$$

$$R{-}\overset{O^{+\cdot}}{\overset{\|}{C}}{-}X \longrightarrow R{-}C{=}O^{+} + X\cdot \qquad (18)$$

[reactions (10)–(12)] corresponding to losses of a free radical; the remaining ion is an even-electron ion. Except for the molecular ion, which is an odd-electron ion, the principal ions in a spectrum are even-electron ions unless special structural requirements are met. In these less common cases, prominent odd-electron ions, formed by the loss of a molecule, are found. *See* FREE RADICAL.

The most general case of odd-electron ion formation is illustrated in reaction (19), where in general a

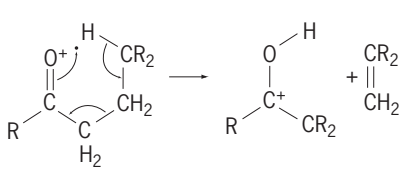

(19)

γ-hydrogen is abstracted by a multiply bonded electronegative atom with rupture of the β bond; this reaction is called the McLafferty rearrangement. Other cases of odd-electron ion formation involve interactions of ortho substituents on aromatic systems and certain decompositions of cyclic molecules, which may or may not be rearrangements [reaction (15)]. The spectrum of the volatile compound acetophenone (**Fig. 4***a*) is characterized only by simple bond-cleavage reactions, but there is an intense peak in the spectrum of valerophenone (Fig. 4*b*) corresponding to a McLafferty rearrangement in which the C_4H_8 group is lost from the molecule.

Even-electron ions, such as many fragment ions found in electron-impact spectra and also (M + 1) ions from chemical ionization, decompose most often to smaller even-electron ions by loss of a small molecule. Illustrations are given by reactions (20) and (21), from the chemical ionization of small pep-

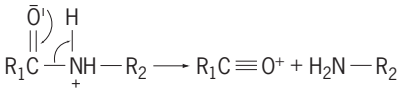

(20)

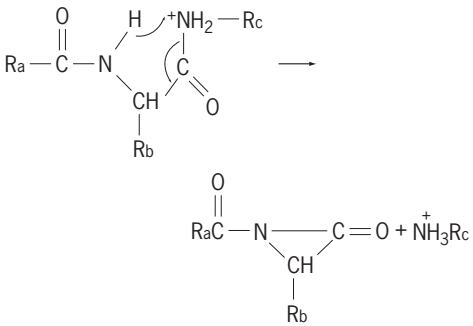

(21)

tides. Complex rearrangements are more competitive with simple cleavages in ions of low internal energy than in ions of higher internal energy; so in chemical ionization spectra, though only a few pathways for loss of small molecules may exist, some of the pathways may be found to be quite remarkable. Occasionally the fragmentations involve rearrangements so poorly understood that the fragment peaks are not of much use in establishing the structure of the molecule.

Spectra of the involatile compound creatine are typical for the various techniques discussed (**Fig. 5**); both the molecular ion (Fig. 5*a*) and the (M + 1) ions (Fig. 5*b*–*d*) lose water by a rearrangement process. It is not universally true that the chance of finding the molecular ion increases as the experiment progresses from electron impact to chemical ionization, field ionization, and field desorption, but these results are typical.

Further useful information about fragments is obtained, first, from the potentials required to ionize the molecule and to form fragments and, second, from observations of shifts of peaks with isotopic labeling techniques. For example, in reaction (19), substitution of D for H at the β position does not increase the mass of the product ion, but when H at the γ position is replaced by D, the mass of

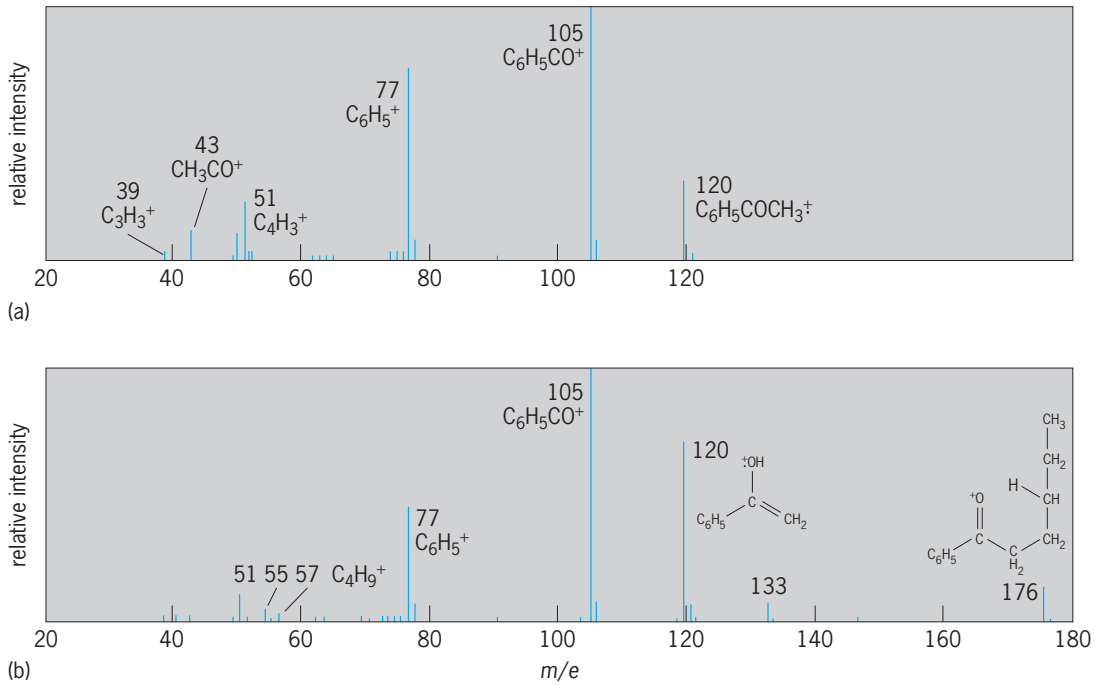

Fig. 4. Mass spectra of (*a*) acetophenone and (*b*) valerophenone, using electron impact.

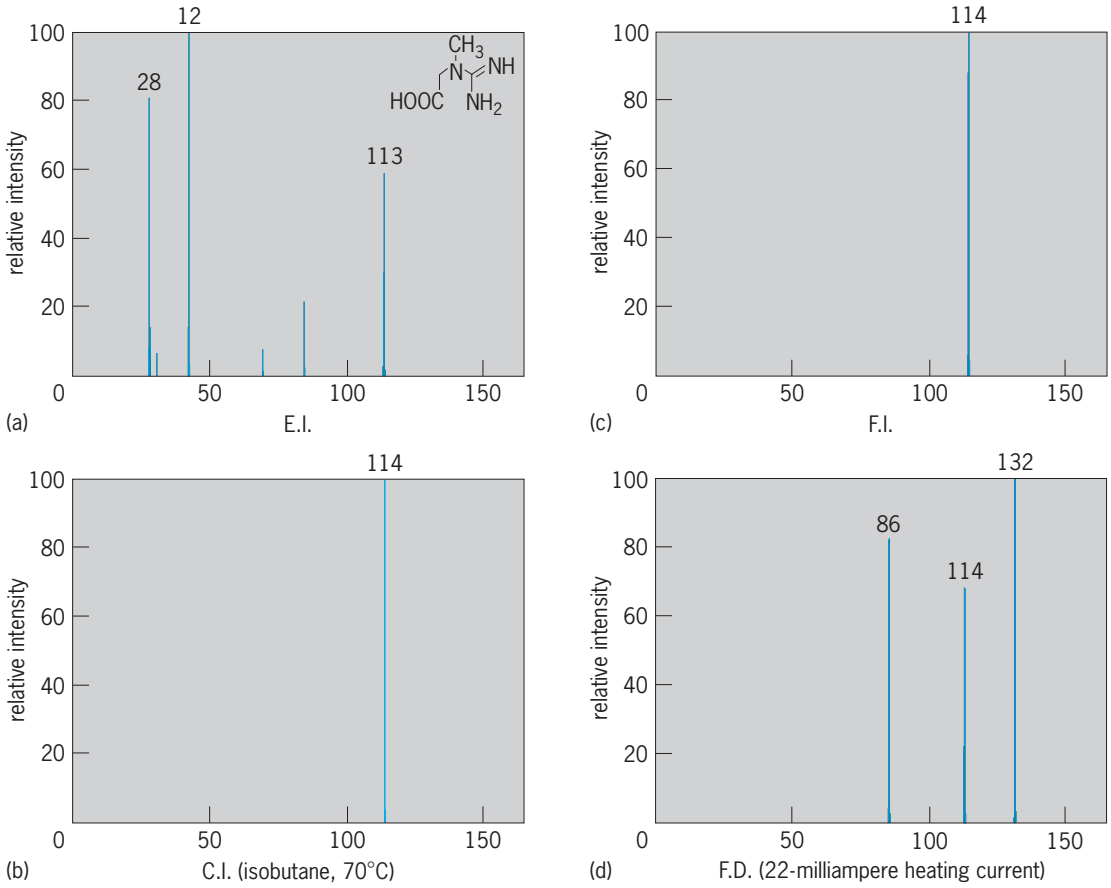

Fig. 5. Mass spectra of creatine (molecular weight 131) obtained by (*a*) electron impact, (*b*) chemical ionization, (*c*) field ionization, and (*d*) field desorption with moderate heating. Both the molecular ion (not observed in *a*) and the M + 1 ion in *b*, *c*, and *d* undergo a loss of water. Field desorption does not always produce the largest ion related to the molecular weight on comparing the four techniques, but this result is typical. (*After H. M. Fales et al., Comparison of mass spectra of some biologically important compounds as obtained by various ionization techniques, Anal. Chem., 47:207–219, 1975*)

the product is shifted to higher mass by one unit. Finally, correlations with trends observed in ordinary chemical reactions yield information. For example, in both aromatic and aliphatic systems, orders of reactivities depending on substitution patterns show correspondence; in some aromatic systems, the correlation is quantitative. If an unknown compound is suspected of belonging to a certain group of compounds, then its structure may often be established by comparison of its spectrum with spectra of compounds of known structure. This technique is particularly valuable in the study of natural products. The shift technique, the interpretation of substitution patterns when the molecular ion and some fragments are shifted by the same amount, for example, 30 units by a methoxy group, is helpful in studying alkaloid spectra. In addition to correlation with typical solution phenomena, relation of fragmentations to pyrolytic reactions and to photochemical processes is studied. The McLafferty reaction is analogous to a well-known photochemical reaction of ketones, the Norrish type II rearrangement, in which an olefin molecule is lost from the alkyl ketone while the enol form of a smaller ketone is produced by the influence of light. Other ionization techniques (by large electric-field gradients, by photons, or by other ions) produce different types of spectra useful in structural analysis. *See* PHOTOCHEMISTRY.

Metastable ions. Correlation of spectra with structure requires the additional use of further decomposition products beyond the first fragmentation steps, but it is difficult to define the origin of products possibly arising by more than one path, for example, B^+ in reactions (13) and (14). A clue to origin of fragments is given by metastable ions: If in reaction (13) some AB^+ ions decompose after they are accelerated but before they are magnetically deflected, a broad peak (**Fig. 6**) appears at an m/e value numerically equal to m_B^2/m_{AB}. The value 56.5 (Fig. 6) equals $77^2/105$; therefore, the metastable peak corresponds to the reaction $105 \rightarrow 77$ and indicates that at least part of the $C_6H_5^+$ ion is formed by the loss of CO from $C_6H_5CO^+$. The order of the decomposition steps is suggested by the collection of metastable ions and bears on the organization of the parts of the molecule into the whole.

Study of metastable ions and other ions that can be made to decompose between acceleration and mass analysis has been carried out through ion kinetic energy spectroscopy (IKES) and mass-analyzed ion kinetic energy spectroscopy (MIKES), also called direct analysis of daughter ions (DADI). In the ion kinetic energy spectroscopy technique, which may be studied with a conventional double-focusing mass spectrometer, those ions which decompose between the accelerating region and the electric sector will form ionic fragments with only a fraction of the kinetic energy imparted to the original ions. The fraction is the ratio of the mass of the fragment to that of the original ion. Scanning the voltage of the electric sector and detecting the ions that leave it provides a scan of ions according to their kinetic energy. Frequently intensities of ions differ substantially for even

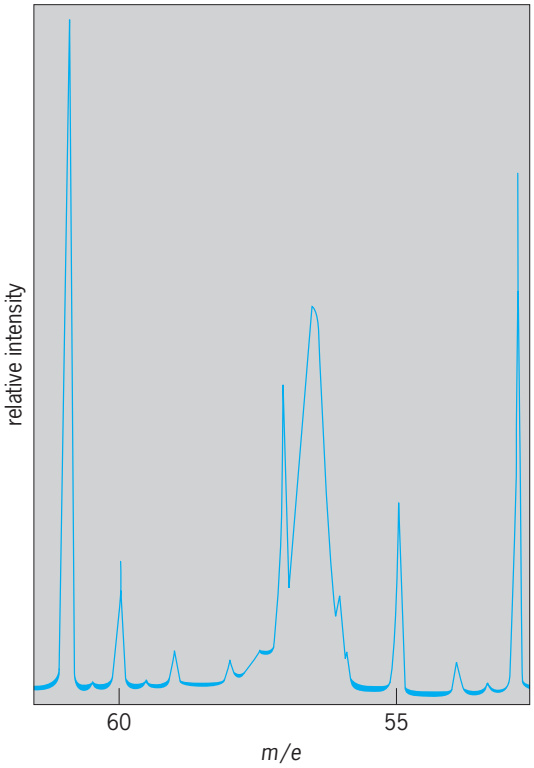

Fig. 6. Scale-expanded portion of acetophenone spectrum showing a broad metastable ion at *m/e* 56.5.

closely related isomers; thus this technique is useful for fingerprinting difficult compounds.

In mass-analyzed ion kinetic energy spectroscopy, the electric sector and the magnetic sector in the double-focusing mass spectrometer are reversed, so that the ions, after acceleration, enter the magnetic sector and are analyzed according to mass before they enter the electric sector. By setting the magnetic field, ions of a certain mass are chosen for study; those that decompose before entering the electric sector have only the fraction of the kinetic energy according to the ratio of fragment and original mass noted above, and by sweeping the electric sector, the amounts of daughter ions from this one particular precursor may be observed. The intensities of ions are useful for structure determination, although they are somewhat affected by the distribution of energies within the fragmenting ion. Much information about energetics of decomposing ions has been gained from studies of the width of peaks due to their products in mass-analyzed ion kinetic energy spectroscopy, for these are always broadened because kinetic energy T is released in the fragmentation. It has been shown that T is a characteristic of molecular structure; for example, the McLafferty product ion formed by the loss of C_2H_4 from the molecular ion of butyrophenone, as in reaction (19), fragments by loss of CH_3, releasing 48 meV (4.6 kilojoules/mole) of T. Since the molecular ion of acetophenone releases 7 meV (0.7 kJ/mole) energy, the McLafferty product does not have the acetophenone structure. Further studies of rearrangements have led to analyses of the partitioning of energy between internal and kinetic

energy in fragmentation as a function of geometry of the activated complex and other parameters. The study of metastable peak shape as a function of ion lifetime, or of ion internal energy, provides information on ion structure.

Unimolecular mass-analyzed ion kinetic energy spectra, that is, those involving metastable ions, are somewhat dependent on ion internal energy. A less energy-dependent method, and therefore a surer guide to structure, is that of collisionally activated decomposition, occasionally also known as collisionally induced dissociation. In this, a collision gas is admitted at a low pressure to the region between the magnetic and electric sector of the reversed instrument. As mass-selected ions travel from the magnetic sector, they collide with the gas molecules, and a fraction of their kinetic energy is transformed into internal energy sufficient to cause rapid decomposition of the ions. The spectra are recorded as in mass-analyzed ion kinetic energy. Charge-exchange processes may also occur on collision, as in reactions (22)–(24). If the electric sector is set at

$$\mathrm{A{-}B{-}C^{2+} + N \longrightarrow A{-}B{-}C^{+} + N^{+}} \qquad (22)$$

$$\mathrm{A{-}B{-}C^{+} + N \longrightarrow A{-}B{-}C^{2+} + N} + e \qquad (23)$$

$$\mathrm{A{-}B{-}C^{+} + N \longrightarrow A{-}B{-}C^{-} + N^{2+}} \qquad (24)$$

twice, one-half, or minus the voltage which transmits normal ions, the products of reactions (22)–(24) can be detected. Thus mass spectra of ions produced by special routes are generated, and the utility of these spectra for interpretive purposes is being explored.

The spectra have been found to depend on the angle by which the products are deflected from their path in the absence of collision, deflections of only a half degree to one degree producing major changes in the spectra in the direction of more energy content in the ions (angle-resolved mass spectrometry).

Collision-induced dissociations in mass-analyzed ion kinetic energy spectrometers involve ion accelerations of 4–8 kV. A similar experiment can be carried out in a triple quadrupole instrument, in which the acceleration is 2–100 eV. In the former, energy transfer is through electronic modes; in the latter through vibrational modes. Because the kinetic energy of the ion strongly influences the collisionally activated decomposition product distribution in the latter experiment, studies of energy-resolved mass spectra of ions by this technique are of much interest, especially in comparison to the angle-resolved mass spectra.

These techniques for producing mass spectra from mass-resolved ions have become of paramount importance to establishing the structures of small ions, particularly in combination with more advanced methods of measuring heats of formation of ions and onset energies for ion formation. Maurice M. Bursey

Analysis techniques. Measurement using a mass spectrometer takes advantage of the mass dependency in the equations of motion of an ion in an electric or magnetic field. Three common mass analyzers are magnetic-sector, quadrupole, and time-of-flight. An ion traversing the magnetic field of a sector instrument successfully reaches the detector when its mass, m, corresponds to $m/z = r^2B^2/2V$, where z is the charge of the ion, r is the radius of the flight tube, B is the magnetic field strength, and V is the acceleration voltage. A mass spectrum can be recorded by varying B or V. A quadrupole is an array of four parallel rods electronically connected such that a radio frequency (RF) is applied to opposing rods, and the waveforms applied to adjacent rods are out of phase. In addition, a direct-current (DC) component is added to the rods such that one opposing pair has +DC and the other −DC. Ions entering the quadrupole follow complex sinusoidal paths. An ion successfully traversing the rods to the detector is mass-selected by the amplitude of the applied radio frequency and the amount of resolution from adjacent masses by the DC/RF ratio. A spectrum is detected by scanning the RF amplitude and the DC voltage in a fixed ratio. The time-of-flight analyzer is a tube with an acceleration field at one end and a detector at the other. Ions are pulsed and accelerated down the tube, starting a clock. Arrival times at the detector are converted to mass by solving the equation $m/z = 2V(t/L)^2$, where V is the acceleration potential, t is the arrival time, and L is the flight tube length.

The mass analyzers are characterized by their mass resolving power (RP), scan or acquisition rate, and mass range. Sector instruments are usually assembled with an additional energy filter and can be operated up to 100,000 RP. Resolving power is defined as $M/\Delta M$, where M is the mass number of the observed mass and ΔM is the mass difference in two signals with a 10% valley between them. Increased resolution in sectors comes at the sacrifice of sensitivity. To increase resolution, the slits are narrowed on the ion beam. It is typical to lose 90% of the signal going from 1000 to 10,000 RP. Magnetic instruments are available with mass ranges up to 10,000 atomic mass units (amu) for singly charged ions at full acceleration potential. To obtain good quality, data sectors are scanned relatively slowly—a practical rate is 5 s per decade (1000–100 amu). Quadrupoles are usually operated at unit resolution. For example, 100 is resolved from 101, or 1000 from 1001. Quadrupoles can be scanned at 2000 amu/s and are available with mass range up to 4000 amu. Fast scanning and high-gas-pressure tolerance make quadrupoles popular for mass spectrometer–chromatographic hybrid systems. Modern time-of-flight analyzers are capable of 10,000 RP, with 50% valley definition of ΔM. Acquisition rates can be 100 spectra per second. The time-of-flight mass range is theoretically unlimited, and singly charged ions over 100,000 amu have been observed. *See* ATOMIC NUMBER; MASS NUMBER. Todd D. Williams

Applications of computers. The principal applications of computers to analysis have been to data acquisition and structure interpretation. High-resolution spectra contain so much data that the

rapid acquisition and presentation of data in a form easily assimilated by the operator has been adapted to the computer. Similarly, in cases where a gas chromatograph effluent passes through the source of the mass spectrometer, the generation of even low-resolution data is so rapid that a dedicated computer is appropriate. A variety of data displays are useful for interpretation. The plot of total ion current versus time produces a reconstructed chromatogram; the plot of ions of a single mass versus time, called a mass fragmentogram, is useful in identifying compound classes among the gas chromatogram peaks if the appropriate mass is chosen, or in identifying compounds directly if some other mass like the molecular weight of a desired component is chosen.

For interpretation, two approaches have been used: library searching and training. Library searches of large collections (over 25,000 spectra) by comparison of the spectrum with known spectra yield degrees of closeness of agreement of the unknown and known spectra. Various algorithms for spectral comparison, using the 10 most intense peaks in the spectrum or the two most intense peaks in each 14-mass-unit segment, for example, have been devised, and the minimum amount of information to be supplied for a good chance of identification has been studied. The other approach involves several pattern recognition approaches in which various features of the spectra are correlated with structural characteristics by methods independent of formal theories of mass spectral interpretation; these include learning-machine and factor-analysis approaches. Hybrid techniques, in which the self-trained computer approach is augmented by selected tests derived from the fragmentation theory noted previously, have also been devised.

Hybrid techniques. The coupling together of different kinds of mass and energy analyzers for specialized purposes produces hybrid mass spectrometers. For example, to achieve unit resolution after collisional activation in a collision cell, an instrument may be designed with a magnetic sector before the collision cell (to choose ions by mass before they enter the collision cell) and a quadrupole mass filter to separate collision products by mass. Another intermediate quadrupole is used between these two analyzers to separate the product–mass-analyzing quadrupole from the high electric field associated with the magnetic sector analysis; ions that have fallen through a high electric field have kinetic energy too great for good separation by a quadrupole. This intermediate quadrupole is operated with only radio-frequency voltages on the rods ("RF-only" mode) and then serves only to collect and move ions of all masses away from the high field, instead of mass-analyzing them. Other hybrid combinations of sectors and assemblies are electric-magnetic-quadrupole-quadrupole (which provides high resolution before the collision cell and unit resolution after) and magnetic-quadrupole–ion cyclotron resonance (which provides high resolution and high mass range after collision). Another instrument related to this discussion contains two electric sectors, so that energy analysis before and after collision, but no mass analysis, can be studied for well-understood systems. Some of these elaborate instruments are commercially available, but they are much more expensive than simpler instruments.

Another use of the term hybrid involves mass spectrometry coupled to a chromatograph or other device for separating compounds; these constitute a major fraction of the so-called hyphenated techniques. The information contained in a rapidly acquired mass spectrum of a chromatographic peak is often enough to identify the structure of the compound eluting from the chromatograph. Among the separation methods that use mass spectrometry for identification are gas chromatography, liquid chromatography, supercritical fluid chromatography, and capillary zone electrophoresis.

Typically, the selective removal of the mobile phase in which the eluting sample is contained is necessary before the mass spectrum is obtained. This may be accomplished by techniques based in part upon rapid pumping. In gas chromatography–mass spectrometry, the gases exiting the chromatograph are also accelerated through a nozzle into the pumping region; the light helium atoms are scattered and pumped away, while the heavier solute molecules tend to stay aligned in a linear path from the nozzle until they enter the orifice of a carefully aligned skimmer, through which they pass to the mass spectrometer source. In liquid chromatography–mass spectrometry a small fraction of the solution exiting the chromatograph may be sprayed into the source, often through a heated nozzle; or it may be deposited on the surface of a moving belt that passes through several vacuum locks into the source, where the sample is evaporated and ionized by heating or particle bombardment. The sample may also be removed from solvent through a selective membrane, or the solution may be mixed with glycerol in a moving stream and subjected to fast atom bombardment in the source. Electrospray methods also serve as a method to interface liquid chromatography and mass spectrometry. The interfaces used for supercritical fluid chromatography–mass spectrometry resemble those for gas chromatography–mass spectrometry, and those for capillary zone electrophoresis–mass spectrometry resemble methods for liquid chromatography–mass spectrometry. *See* ELECTROPHORESIS; GAS CHROMATOGRAPHY; LIQUID CHROMATOGRAPHY; SUPERCRITICAL-FLUID CHROMATOGRAPHY.

Ion cyclotron resonance. In ion cyclotron resonance spectrometry the ions, as charged particles in a magnetic field, complete circular orbits. Their frequency depends upon the strength of the magnetic field and their mass-to-charge ratio [Eq. (25)],

$$\omega = \frac{zB}{m} \tag{25}$$

where ω = frequency, z = charge of the ion, B = magnetic flux density, and m = mass of the ion, and the signal produced by the packet of ions is

detected by tracking of ion motion by image currents in a detector plate. At first, ion cyclotron resonance spectrometry was used for the study of ion–molecule reactions at low pressures because of the long lifetimes of ions in the ion cyclotron resonance cell; it has evolved to serve as the fundamental technique upon which Fourier-transform mass spectrometry is based. In this type of mass spectrometry, ions of all masses are excited by a rapidly swept excitation pulse, so that ions of each mass have coherent motion. The image current consists of a sum of sinusoidal waves of different frequencies; it can be transformed into a spectrum in the frequency domain, and thus by Eq. (25) into a spectrum in the mass domain. Loss of coherence of ion motion results from collisions of ions with neutrals. Although arbitrarily high resolution and mass range are expected, collisions reduce the time available for observation and thus limit resolution. Ions of mass greater than 20,000 daltons have been detected by Fourier-transform mass spectrometry, though other analyzers have also gone above this limit. Resolution greater than 3,000,000 daltons has also been observed in Fourier-transform mass spectrometry, and this is the record for all analyzers. *See* CYCLOTRON RESONANCE EXPERIMENTS; INFRARED SPECTROSCOPY.

Analysis of inorganic solids. The analysis of solid inorganic samples can be made either by vaporization in a Knudsen cell arrangement at very high temperatures or by volatilization of the sample surface so that particles are atomized and ionized with a high-energy spark (for example, 20,000 eV). The wide range of energies given to the particles requires a double-focusing mass spectrometer for analysis. Detection in such instruments is by photographic plate; exposures for different lengths of time are recorded sequentially, and the darkening of the lines on the plate is related empirically to quantitative composition by calibration charts. The method is useful for trace analysis (parts per billion) with accuracy ranging from 10% at higher concentration levels to 50% at trace levels. Methods for improving accuracy, including interruption and sampling of the ion beam, are under development.

Secondary ion mass spectrometry is most commonly used for surface analysis. A primary beam of ions accelerated through a few kilovolts is focused on a surface; ions are among the products sputtered from this surface, and they may be directly analyzed in a quadrupole filter. Sputtered material may also be analyzed by ionization of the neutrals in an inductively coupled plasma and subsequent mass analysis. This method produces ions with a lower energy spread than spark source mass spectrometry and, since detection then becomes less of a problem, has been supplanting spark source methods. *See* LASER SPECTROSCOPY; SECONDARY ION MASS SPECTROMETRY (SIMS); SPECTROSCOPY; SPUTTERING.

Maurice M. Bursey

Bibliography. J. R. Chapman, *Practical Organic Mass Spectrometry: A Guide of Chemical and Biochemical Analysis*, 2d ed., 1993; S. Gaskell (ed.), *Mass Spectrometry in Biomedical Research*, 1986; E. de Hoffmann and V. Stroobant, *Mass Spectrometry: Principles and Applications*, 2d ed., 2001; S.-T. F. Lai, *Gas Chromatography/Mass Spectrometry*, 1988; F. W. McLafferty and F. Turecek, *Interpretation of Mass Spectra*, 4th ed., 1992; J. Throck Watson, *Introduction to Mass Spectrometry*, 3d ed., 1997.

Mass spectroscope

An instrument used for determining the masses of atoms or molecules found in a sample of gas, liquid, or solid. It is analogous to the optical spectroscope, in which a beam of light containing various colors (white light) is sent through a prism to separate it into the spectrum of colors present. In a mass spectroscope, a beam of ions (electrically charged atoms or molecules) is sent through a combination of electric and magnetic fields so arranged that a mass spectrum is produced. If the ions fall on a photographic plate which after development shows the mass spectrum, the instrument is called a mass spectrograph; if the spectrum is allowed to sweep across a slit in front of an electrical detector which records the current, it is called a mass spectrometer.

Operation. A typical mass spectroscope has a continuously pumped vacuum chamber, commonly called the spectrometer tube, into which the gas or vapor to be investigated flows at such a rate that the equilibrium pressure in the chamber is of the order of 10^{-6} mm of mercury (10^{-8} lbf/in.^{2} or 10^{-4} pascal). **Figure 1** is a schematic drawing of a type of mass spectrometer tube widely used for making gas and isotope analyses. The pumping system consists of a mechanical vacuum forepump followed by either an oil or mercury diffusion pump and a cold trap maintained at dry ice or liquid air temperature. A sufficient quantity of the gas to be analyzed is placed in a

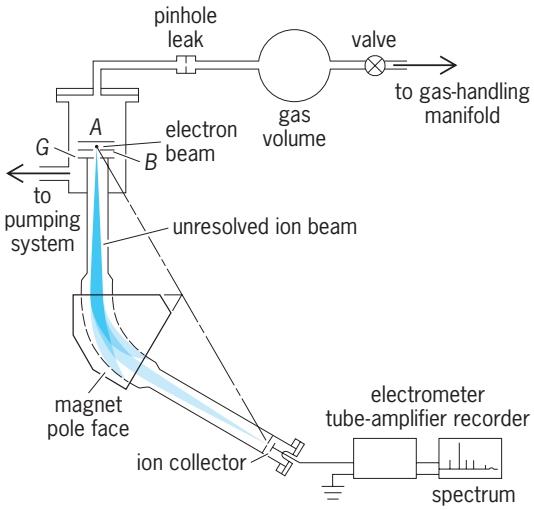

Fig. 1. Schematic diagram of mass spectrometer tube. Ion currents are in the range 10^{-10} to 10^{-15} A and require special electrometer tube amplifiers for their detection. In actual instruments the radius of curvature of ions in a magnetic field is 4–6 in. (10–15 cm).

vacuum chamber, whose pressure is approximately 0.05 mmHg (10^{-3} lbf/in.^{2} or 7 Pa). A pinhole having a diameter of approximately 0.001 in. (0.025 mm) permits this gas to leak continuously into the mass spectrometer tube.

A heated tungsten or rhenium filament, not shown in the diagram, produces an electron beam normal to the plane of the diagram, as shown. The electrons in the beam collide with the molecules of gas present and knock off one or more electrons, thus creating positive ions. *See* ION SOURCES.

In a monatomic gas such as argon, multiply charged positive ions, designated as Ar^+ (singly charged), Ar^{2+} (doubly charged), and so on, are formed. In the case of polyatomic molecules, ionized fragments may also be formed. For example, for methane, CH_4, the ions CH_4^+, CH_3^+, CH_2^+, CH^+, C^+, and H^+ are found. In some cases, negative ions are also created as a result of an electron capture process. Thus for CO, in addition to the ions CO^+, C^+, O^+, the ions C^- and O^- are observed. *See* ELECTRON CAPTURE.

An electric field resulting from the application of a potential difference of several volts between A and B draws the ions through the slit in plate B. Further energy is given the ions by allowing them to fall through an electric potential of several hundreds or thousands of volts applied between plates B and G. In such an instrument setup, plate G is grounded.

The beam of ions travels downward and because of the finite width of the slits in plates B and G, diverges lightly, as shown. The beam passes between the poles of a magnet as indicated. The magnetic field is perpendicular to the plane of the diagram. In the magnetic field the ions experience a force at right angles to their direction of travel given by Eq. (1),

$$f = Bev \tag{1}$$

where f is the force, B the magnetic field intensity, e the charge on the ion, and v its speed, all in SI units. This force results in a circular trajectory, whose radius r is found by equating the force to the product of mass and acceleration according to Newton's laws of motion, as in Eq. (2).

$$Bev = \frac{mv^2}{r} \tag{2}$$

Equation (2) may be put in the more convenient form of Eq. (3) by equating the kinetic energy gained

$$\tfrac{1}{2}mv^2 = eV \tag{3}$$

to the potential energy lost and expressing v in terms of potential difference V through which the ion fell. By combining Eqs. (2) and (3), solving for r, and substituting units which are more convenient for actual calculation, Eqs. (4) are obtained, where r' is now

$$r' = 57(MV/e)^{1/2}B \tag{4a}$$

$$r = \frac{144(mV/e)^{1/2}}{B} \tag{4b}$$

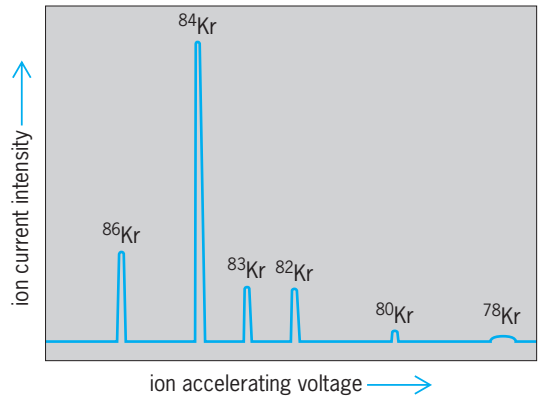

Fig. 2. Mass spectrum of krypton isotopes obtained with an instrument such as that shown in Fig. 1.

in inches, r is in cm, B in gauss (1 gauss = 10^{-4} tesla), V in volts, e is the charge of the ion measured in terms of the number of electrons removed (or added) during ionization, and m is the mass measured in atomic mass units; that is, for hydrogen $m = 1$, for the most abundant isotope of oxygen, ^{16}O, $m = 16$, and so on.

Figure 1 shows the paths of ions having three different masses. Only the intermediate group has the proper mass to reach and pass through the collector slit and to be measured.

If the source of ions, the apex of the wedge-shaped field, and the collector slit all lie on a straight line, the diverging beam of ions is focused as shown. This property of a wedge-shaped magnetic field is analogous to the focusing property of a convex lens in optics.

From Eqs. (4) it is clear that if either B or V is varied continuously, a mass spectrum may be swept. **Figure 2** is a recording of such a spectrum for krypton. The abundances of the several isotopes are proportional to the recording pen deflection.

Routine gas or isotope analyses may be made on 10^{-3} in.^{3} (10^{-2} cm^3) of gas at normal pressure and temperature (NTP), and in special cases as little as 10^{-9} in.^{3} (10^{-8} cm^3) of gas at NTP suffices. Some solids, such as alkali metals or alkali earths, give off ions directly when heated; therefore a filament coated with a salt containing these elements may serve as a source of ions. With this method, analyses have been made on samples as small as 3×10^{-14} oz (10^{-12} g). Electric sparks may also serve as sources of ions in solid analyses. Because of the unsteady nature of a spark, photographic recording of the spectra is generally employed.

In order to obtain higher resolution and a more definite relation between mass and position in a spectrum, an electric field region is employed in series with the magnetic field region. With this arrangement, called double focusing, resolutions of the order of 1 part in 50,000 are obtained, and masses can be determined to 1 part in 10^7, or better. **Figure 3** is a mass spectrum obtained with a double-focusing mass spectrograph.

Applications. Mass spectroscopes are used in both pure and applied science. Atomic masses can be

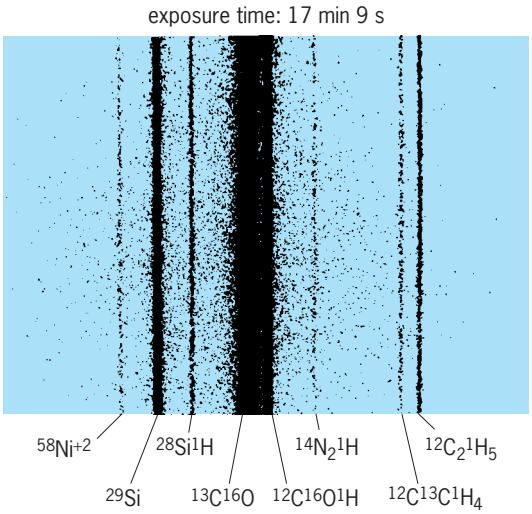

Fig. 3. Mass spectrum observed at mass number 29 when several substances are present simultaneously in a mass spectrograph. Resolution is sufficient to completely separate two close lines such as $^{12}C^{13}C^{1}H_4$ and $^{12}C_2{}^{1}H_5$, which differ in mass by only 1 part in 75,000.

measured very precisely. Because of the equivalence of mass and energy, knowledge of nuclear structure and binding energy of nuclei is thus gained. The relative abundances of the isotopes in naturally occurring or artificially produced elements can be determined. Thus nuclear processes occurring either in nature or in the laboratory may be investigated. Isotopic analyses of elements such as lead, argon, or strontium, which may result from the radioactive decay of other elements, are of particular interest because they make possible the determination of the geological age of the minerals from which the elements are extracted. Other geological effects which cause variations in relative abundances of isotopes may also be investigated.

Empirical and theoretical studies have led to an understanding of the relation between molecular structure and the relative abundances of the fragments observed when a complex molecule, such as a heavy organic compound, is ionized. When a high-resolution instrument is employed, the masses of the molecular or fragment ions can be determined so accurately that identification of the ion can frequently be made from the mass alone. Thus, although $C_2H_4N_4$ and $C_2H_4N_2O$ have molecular weights of approximately 84 when made up of the most abundant isotope in each of the elements concerned, the deviation from integral numbers of the constituents is sufficiently different to cause the exact masses to be 84.04359 and 84.03236 atomic mass units (amu), respectively ($^{12}C = 12.00000$ amu). If the isotopic nature of the elements is also taken into account, additional ions will be observed whose relative abundances depend upon the relative abundances of the isotopes. This fact provides another powerful tool in determining the identification of the ion in question. These methods, supplemented by others, frequently make the determination of the molecular formula of a compound relatively easy. Once this has been accomplished, the actual structure can usually be found by considering the relative abundances of the fragment ions.

Mass spectrometers also make possible isotopic analyses of compounds which have reacted chemically with other compounds containing elements with artificially altered isotopic abundance ratios. Thus the instruments make possible tracer studies of chemical or biochemical reactions.

Because chemical compounds may have mass spectra as unique as fingerprints, mass spectroscopes are widely used in industries such as oil refineries, where analyses of complex hydrocarbon mixtures are required. For further information on applications *see* BETA PARTICLES.

Miscellaneous types. In the instruments thus far discussed, a mass spectrum is obtained by making use of the fact that ions of different mass are deflected through different angles in passing through a magnetic field. In one important modification a superimposed electric and magnetic field is used, resulting in a trochoidal ion path. This has certain advantages, especially in that high resolution is obtained in a very compact apparatus. Various time-of-flight mass spectrometers are also employed. In some of these, ions are accelerated in pulses or by potentials varying sinusoidally with time and then sent through a system of grids where potentials also vary with time. A separation according to mass is effected because ions of different mass require different times to traverse the arrangement. In still others, use is made of the cyclotron principle; that is, the time for an ion to traverse a complete circle in a magnetic field is independent of the energy of the ion and depends only on its mass. In the quadrupole spectrometer, or *Massenfilter*, ions pass along a line of symmetry between four parallel cylindrical rods. An alternating potential superimposed on a steady potential between pairs of rods filters out all ions except those of a predetermined mass.

Alfred O. Nier

Tandem accelerator mass spectrometers. Mass spectrometers of the type described above have limitations in certain circumstances. For example, they may not be able to resolve molecules from equivalent mass number atoms, although the double-focusing mass spectrometers described above generally have sufficient mass resolution to avoid this problem. Such high resolution (1 part in 50,000 or better) is achieved, however, with a sacrifice in efficiency. A more serious problem is the separation of the wanted atom from its isobars. For example, ^{14}C and ^{14}N differ in mass by 1 part in 100,000.

Both these problems are overcome by the use of a tandem electrostatic accelerator as a mass spectrometer (**Fig. 4**). The system starts with negative ions (neutral atoms to which an extra electron is attached) produced in a sputter ion source. These are mass-analyzed to about 1 part in 200 and injected into the low-energy end of a tandem electrostatic accelerator. They are accelerated to the central positive terminal whose voltage can be adjusted up to 12 MV. In the terminal they pass through a differentially pumped

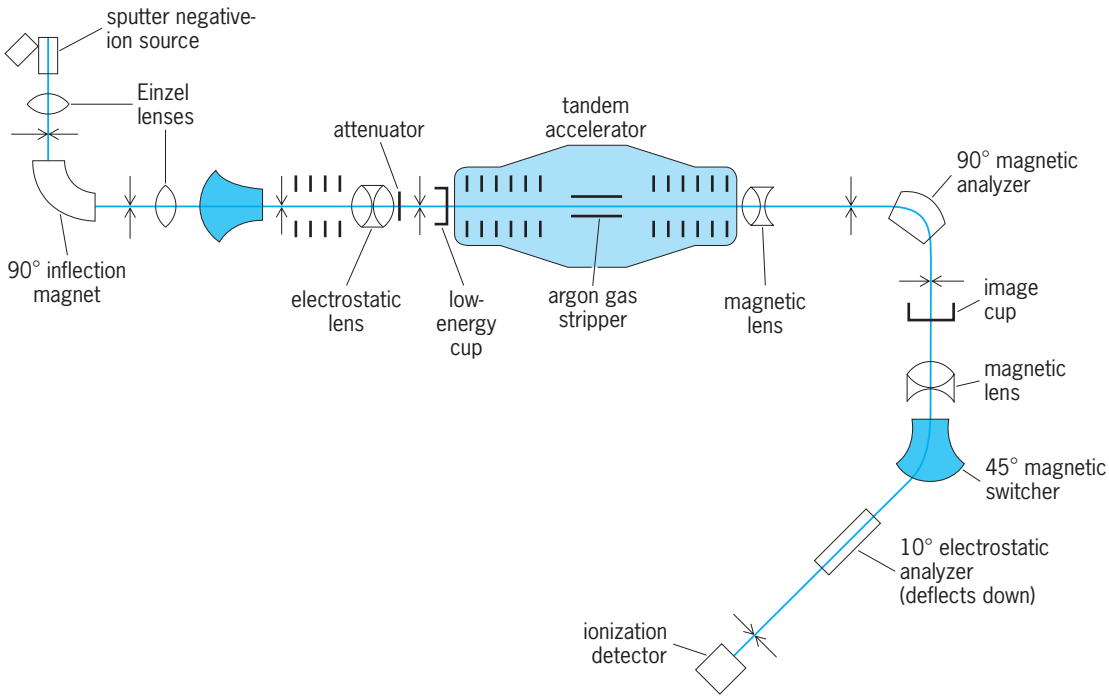

Fig. 4. Typical arrangement of a tandem accelerator mass spectrometry system.

argon gas stripper and emerge with several electrons removed. These positive ions are then accelerated through the second half of the machine and exit at the high-energy end. Here they are subjected to a series of magnetic and electrostatic deflections and finally pass into an ionization detector. Molecules are dissociated in the terminal stripper, and isobars are separated by their different rates of energy loss in the ionization detector due to their different atomic number.

Such systems are used to measure cosmogenic radioisotopes in a wide variety of naturally occurring material as well as stable isotopes in matrices of much more abundant elements. The **table** lists some of the radionuclides that have been measured. The limit of detection is listed in the last column. This limit is the ratio of the number of atoms of the radionuclide to the number of atoms of the stable isotopes. For ^{14}C and ^{16}Cl, the detection limit approaches 1 part in 10^{16}. Such detection limits never have been, and probably never will be, achieved by conventional mass spectrometers.

Radioactive isotopes are normally measured by detecting their decay products. For example, the amount of ^{14}C in a carbonaceous artifact can be used to determine the age of that artifact. To make such a measurement by radioactive decay counting, because of the relatively long half-life of 5730 years, requires carbon samples in the range of 1 to 10 grams. The tandem accelerator technique can measure the ^{14}C contents of a carbon sample by using 10 micrograms to a few milligrams of carbon. There are innumerable examples of carbonaceous artifacts which it would be of great interest to date but whose carbon content or the sample size is much too small to permit radioactive decay counting. The accelerator technique makes their age measurement possible. *See* RADIOCARBON DATING.

Many of the other radioisotopes listed in the table have much larger half-lives than ^{14}C and thus are even more difficult to measure by decay counting. The accelerator mass spectrometry technique, which does not require that the isotope be measured by radioactive decay, can readily be used.

Radionuclides studied in tandem accelerator mass spectrometer

Radio-nuclide	Half-life, years	Stable isotopes	Stable isobars	Chemical form	Charge state	Energy, MeV	Limit of detection
^{10}Be	1.6×10^6	^{9}Be	^{10}B	BeO	3+	33†	7×10^{-5}
^{14}C	5730	$^{13,13}C$	^{14}N*	C	4+	40	0.3×10^{-15}
^{26}Al	7.2×10^5	^{27}Al	^{26}Mg*	Al	5+	48	10×10^{-15}
^{32}Si	108	$^{28,29,30}Si$	^{32}S	SiO_2	5+	55	7×10^{-12}
^{36}Cl	3.0×10^5	$^{35,37}Cl$	^{36}Ar*, ^{36}S	AgCl	7+	80	0.2×10^{-15}
^{129}I	15.9×10^6	^{127}I	^{129}Xe*	AgI	5+	30	0.3×10^{-12}

*Does not form negative ions.
†BeO from source.

Carbon-14 has now been measured by tandem accelerators in a wide variety of important archeological, anthropological, and geological samples. For example, the date of the Turin Shroud was measured by accelerator mass spectrometry. This cloth, believed by many to have been Christ's burial shroud, could not have been dated by the conventional decay-counting technique because a sample the size of a handkerchief would have been required. Three laboratories made the measurements independently, using only a few square centimeters of cloth. They ascertained that the flax from which the linen cloth was made was harvested in A.D. 1325 ± 30.

Chlorine-36 has been measured in meteorites, lunar rocks, ice cores, ground water, and many other naturally occurring materials. Beryllium-10 has been measured in lake and ocean cores, a Moon core, rainwater, ice, and manganese modules. Iodine-129, which is the longest-lived cosmogenic radioisotope, has been measured in meteorites and in lunar rocks and ground water.

The accelerator mass spectrometry technique has also been applied to the measurement of stable isotopes. In particular, platinum and iridium have been measured in samples of geological and anthropological interest.

Accelerator-based ultrasensitive mass spectrometry finds applications in quite different areas of research than conventional mass spectrometry. The two techniques complement one another. *See* MASS SPECTROMETRY.

Harry E. Gove

Bibliography. J. Barker, *Mass Spectrometry*, 2d ed., John Wiley, 1999; J. R. Chapman, *Practical Organic Mass Spectrometry*, 2d ed., Wiley, 1993; E. de Hoffman and V. Stroobart, *Mass Spectrometry Principles and Applications*, 2d ed., Wiley, 2002; H. E. Gove, *From Hiroshima to the Iceman: The Development and Applications of Accelerator Mass Spectrometry*, Institute of Physics, Bristol and Philadelphia, 1998; H. E. Gove, *Relic, Icon or Hoax? Carbon Dating the Turin Shroud*, Institute of Physics, Bristol and Philadelphia, 1996; F. W. McLafferty and F. Turaček, *Interpretation of Mass Spectra*, 4th ed., University Science Books, 1993; C. Tuniz et al., *Accelerator Mass Spectrometry: Ultrasensitive Analysis for Global Science*, CRC Press, Boca Raton, FL, 1998; J. T. Watson, *Introduction to Mass Spectrometry*, 3d ed., Lippincott Williams & Wilkins, 1997.

Mass wasting

A generic term for downslope movement of soil and rock, primarily in response to gravitational body forces. Mass wasting is distinct from other erosive processes in which particles or fragments are carried down by the internal energy of wind, running water, or moving ice and snow. Mass wasting generally creates colluvium, a type of deposit over or at the bottom of many slopes and with characteristics as varied as the mass-wasting processes.

Shear strength and stress. The stability of slope-making materials is lost when their shear strength (or sometimes their tensile strength) is overcome by shear (or tensile) stresses, or when individual particles, fragments, and blocks are induced to topple or tumble. The shear and tensile strength of earth materials depends on their mineralogy and structure. Processes that generally decrease the strength of earth materials include one or more of the following: structural changes, which promote fracturing and faulting, hence, weakening of the mass; weathering, which changes mineralogy and promotes further breakdown; ground water, which softens clays, dissolves some minerals, and, more importantly, develops pressures that oppose the frictional resistance of the mass—a main component of its shear strength; and meteorological changes, which result in melting of ice and snow. Stresses in slopes are increased by steepening, heightening, and external loading due to static and dynamic forces. Processes that increase stresses can be caused by natural or human means. Thus, steepening and heightening may result from erosion or from mining activity; static loading of slopes may be caused by snow, rain, or fill placement; dynamic loading may be the consequence of earthquakes or blasting. Although other classifications exist, these movements can be conveniently classified according to their velocity into two types: creep and landsliding. *See* SOIL MECHANICS.

Creep. Geologically, creep is the imperceptible downslope movement at rates as slow as a fraction of a millimeter per year; its cumulative effects are ubiquitously expressed in slopes as the downhill bending of bedded and foliated rock, bent tree trunks, broken retaining walls, and tilted structures. There are two varieties of geologic creep. Seasonal creep is the slow, episodic movement of the uppermost several centimeters of soil, or fractured and weathered rock. This type of creep occurs when particles or fragments are heaved perpendicular to the slope surface by freezing and allowed to come down roughly in a vertical direction during thawing; the resulting cycle gives the movement a lateral component. Seasonal creep is especially important in regions where the ground is permanently frozen. *See* PERMAFROST.

The second variety, rheologic creep, is sometimes called continuous creep. It is a time-dependent deformation at relatively constant shear stress of masses of rock, soil, ice, and snow. This type of creep affects rock slopes down to depths of a few hundred meters, as well as the surficial layer disturbed by seasonal creep. Continuous creep is most conspicuous in weak rocks and in regions where high horizontal stresses (several tens of bars or several megapascals) are known to exist in rock masses at depths of 300 to 600 ft (100 to 200 m). These conditions are found in shale sections of those parts of the north-central United States, Ontario, and New York adjacent to the Great Lakes. In this region, shales are known to have crept several tens of centimeters into valleys and excavations. In this process, interbeds of sandstone are stretched, fractured, and separated into discrete blocks by the creeping shale above and below

Fig. 1. Landslide of slump type in bank along Mississippi River. (*U.S. Soil Conservation Service*)

the interbeds. Deep-seated creep also occurs along narrow weakness zones in rock masses before such zones become ruptured and the moving masses become landslides.

Landsliding. This includes all perceptible mass movements. Three types are generally recognized on the basis of the type of movement: falls, slides, and flows. Falls involve free-falling material; in slides the moving mass displaces along one or more narrow shear zones (**Fig. 1**); and in flows the distribution of velocities within the moving mass resembles that of a viscous flow. Sometimes flows are not included as a type of landslide but as a separate type of mass movement. All three types can occur in different earth materials and can affect small hill slopes or high mountainsides; often, more than one type occur in a single event. In many instances, a very large rock mass begins to displace as a slide but soon breaks down into progressively smaller blocks and fragments which tend to move very much like sand poured from a container. The most disastrous mass movements in history (debris avalanches) have evolved in this manner (**Fig. 2**).

Fig. 2. Rockslide debris avalanche at Mount Huascaráan, Peru, 1970. The large scar on the side of Mount Huascaráan (background) marks initiation of the rockslide. Santa Valley is seen in the foreground.

Mass wasting is an important consideration in the interaction between humans and the environment. Deforestation accelerates soil creep. Engineering activities such as damming and open-pit mining are known to increase landsliding. On the other hand, enormous natural rock avalanches have buried entire villages and claimed tens of thousands of lives. *See* LANDSLIDE. Alberto S. Nieto

Bibliography. M. J. Crozier, *Landslides*: *Causes, Consequences, and Environment*, 1986; R. L. Schuster and R J. Krizek (eds.), *Landslides*: *Analysis and Control*, NAS-NRC Spec. Rep. 176, 1978; B. Voight (ed.), *Rockslides and Avalanches*, pt. 1, Developments in Geotechnical Engineering, vol. 14a, 1978.

Massif

A block of the Earth's crust commonly consisting of crystalline gneisses and schists, the textural appearance of which is generally markedly different from that of the surrounding rocks. Common usage indicates that a massif has limited areal extent and considerable topographic relief. Structurally, a massif may form the core of an anticline or may be a block bounded by faults or even unconformities. In any case, during the final stages of its development a massif acts as a relatively homogeneous tectonic unit which to some extent controls the structures that surround it. Numerous complex internal structures may be present; many of these are not related to its development as a massif but are the mark of previous deformations. Philip H. Osberg

Mastigophora

A superclass of the Protozoa also known as the Flagellata. The flagellates make up a large and heterogeneous group of Protozoa, varying in size from *Noctiluca*, 1500 micrometers in diameter, to monads, barely 3 micrometers long. The common morphological flagellate type is spherical to cylindrical on an anteroposterior axis. They are probably more nearly representative of the primitive protozoan type than Sarcodina or Ciliophora, but exhibit a more evident relationship to the former. Despite their common morphological plan, based on the flagellum as a means of locomotion, Flagellata are very diverse in shape, colony formation, internal structure, external shell or test, color, physiology, reproduction, and choice of environment. *See* CILIOPHORA; SARCODINA.

The Mastigophora exhibit marked plantlike characteristics, so that texts and references in botany always treat at least some of them. Some workers include all colorless flagellates in the algae; however, flagellates display many distinctly animal features

and sometimes are treated as Protozoa. Thus, flagellates are regarded as a link between the plant and animal kingdoms. *See* PHYTAMASTIGOPHOREA; ZOOMASTIGOPHOREA.

Morphology. Mastigophora possess one common feature, locomotion by means of one or more whiplike protoplasmic extrusions termed flagella. A few have secondarily lost this organelle. In some it is single and directed anteriorly; in a few it is directed posteriorly. Others have one anterior and one trailing flagellum, but there are numerous combinations such as one anterior and two trailing (*Amphimonas*), or up to eight anterior (*Polyblepharides*). Dinoflagellates customarily have two laterally inserted flagella, one posteriorly directed, the other encircling the cell. *Trepomonas* and its relatives have two lateral sets, and the Hypermastigina, symbiotic in termites and wood roaches, have large numbers of very long flagella, mostly lateral. The flagella may be equal or subequal. Usually the flagellar equipment is handed on to one daughter cell at division, a new set being produced in the other. In some species (*Mastigamoeba reptans*) the flagellum may be temporarily withdrawn, at which time the organism moves by ameboid action. Flagella are usually few in number, except in Hypermastigina.

Body form. The flagellate body is typically monaxial and elongate to some degree. Cells are practically spherical in *Monas*, but ovoid, cordiform, pyriform, fusiform, acicular, tubular, or flattened cells are more common. Flattened species typically glide along a surface rather than swim. A particular shape is normally maintained by *Ochromonas* (see **illus.**), whereas *Mastigamoeba* and some others form pseudopodia, and certain species (*Euglena agilis*) undergo frequent distortions termed euglenoid or metabolic changes of shape.

Structurally the flagellate cell is not simple. The cytoplasm is sometimes quite vacuolated (*Collodictyon* or *Trepomonas*) and occasionally colored, although color is generally confined to chromatophores when these are present. One or more contractile vacuoles may be located near the flagellar base.

Colony formation. Most flagellates are single-celled but colony formation is frequent. Cells may be naked (*Oicomonas*), enclosed in a thick cellulose cell wall or pellicle (*Euglena spirogyra*), or in a chitinous, calcareous, or silicious test, the lorica or shell (*Trachelomonas, Distephanus speculum*). The lorica may be smooth, spiny, sculptured, or with attached foreign bodies (*Urceolus sabulosus*). Tests may be clear or colored by impregnated iron. Colony forms are variously flattened (*Platydorina*), spherical (*Volvox*), irregular (*Cladospongia*), linear (*Desmarella*), catenate (*Gonyaulax*), tubular (*Stelexomonas*), or dendroid. Colony walls are usually gelatinous in texture but vary. Single cells are frequently attached by a stalk to some substrate, especially those living in cup-shaped tests. The cell proper rarely shows protuberant parts other than flagella and pseudopodia, but the Protomonadina include many species showing one (*Monosiga, Proterospongia*), or two (*Dicraspedella*) protoplasmic collars ringing the base of the flagellum.

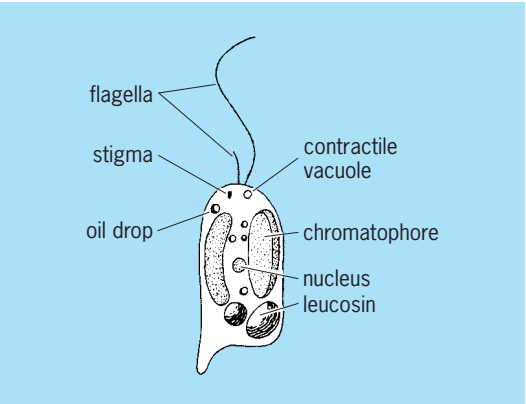

***Ochromonas ludibunda*, showing component parts.**

Flagellum. The flagellum contains a structurally complex longitudinal axial fibril or axoneme. Along its length numerous short lateral hairlike mastigonemes occur which can be seen only in special preparations. The flagellum originates in a kinetic body, the blepharoplast, usually close to the outer surface of the cell. Blepharoplasts for each flagellum are connected by fibrils termed paradesmoses. Another fibril, the rhizoplast, may extend from the blepharoplast to the centriole or to the nucleus. There may be still other structures in the flagellar complex, demonstrable by special techniques. The whole flagellar apparatus produces coordinated movement. Red eye spots (stigmas) may occur near the base of the flagellum. They are considered photoreceptors, and are sometimes accompanied by a lenslike structure; the whole is important in phototactic orientation. *See* CILIA AND FLAGELLA; PHOTORECEPTION.

Nucleus. Nuclei are single except in the distomates (*Hexamitus, Giardia*) and each undergoes mitosis in division. Some contain endosomes of unknown function but which do not contribute directly to chromosome formation. Others contain nucleoli which apparently are identical with, or contribute to, chromatin. Dinoflagellate nuclei are large and vesicular; each chromosome apparently retains its identity during the interphase. In the Euglenoidina the chromatin granules are typically scattered with one or more centrally located endosomes. Nuclei are large in Chloromonadida and are small in Chrysomonadida, Cryptomonadida, Phytomonadida, Zoomastigophorea, Rhizomastigida, and some Protomonadida. In these the chromatin tends to peripheral location with central endosomes.

Nuclear and cell division. In species whose nuclear division is mitotic, details vary, but definite and species-characteristic chromosome numbers occur. In some cases (*Bodopsis godboldi, Dimorpha mutans*), well-defined spindles, within or without the original nuclear membrane, are present. Blepharoplasts function as centrioles. Other mitoses may lack spindles or centrioles (*Oxyrrhis marina, Entosiphon sulcatum*). Mitosis in Hypermastigida is extremely complicated. Most flagellates are haploid; gamete

formation is the exception although sexual processes are known for Phytomonadina. *Pleodorina californica* has visibly different somatic and reproductive cells.

Cell division is generally longitudinal binary fission which sometimes occurs while the organism is encysted, but more often while it is active. Occasionally, multiple fission occurs as in *Chlorogonium* whose trophozoites are sometimes seen to be divided into eight daughter cells. *See* MITOSIS.

Cell inclusions. Flagellates contain various cell inclusions, such as mitochondria and Golgi apparatus. Conspicuous inclusions are chromatophores and reserve food materials. Chromatophores are discoid, cup-shaped, or irregular networks containing one or more of the chlorophylls and often carotene, xanthophyll, or other pigments. Green is characteristic for Phytomonadida and Euglenida, yellow for Chrysomonadida. Cryptomonadida and Dinoflagellida are blue, brown, olive green, pink, or red. Chromatophores may contain pyrenoids as centers of starch formation. Other supposed reserve food inclusions are paramylum (Euglenida); leucosin and volutin (Chrysomonadida); and glycogen (Chloromonadida). Lipids are widely distributed. Occasionally hematochrome granules are sufficiently abundant to color the cell, as when blooms of *Euglena haematoides* redden the surfaces of ponds in hot summer weather. *See* CAROTENOID; CHLOROPHYLL; CHROMATOPHORE; GOLGI APPARATUS; MITOCHONDRIA.

Nematocysts, trichocysts, and trichites occur in flagellates. In *Polykrikos* (Dinoflagellida) nematocysts lie free in the cytoplasm, but their function is unknown. Trichocysts in *Polykrikos*, in *Euglena* species, and in *Chilomonas* may be discharged as long threads by various stimuli, but here too their function is open to speculation.

Nutrition. Nutrition is holophytic when food is manufactured from inorganic substances by those species containing chlorophyll; saprozoic, when dissolved organic materials are absorbed and resynthesized; and holozoic, when solid food is ingested. *Euglena gracilis*, which is holophytic in the light, may become saprozoic in the dark. *Euglena acus* may become saprozoic in the light in a rich organic medium. *Ochromonas ludibunda* apparently is simultaneously holophytic and holozoic. Flagellata perhaps more than any other group vary widely in nutritive processes from those involving intake of specific chemical entities to those of synthesis. Their requirements of trace elements, vitamins, and amino acids vary almost from species to species, as shown by precise experimental investigation. *See* NUTRITION; PHOTOSYNTHESIS.

Respiration. Respiration is complex. Many flagellates adjust to wide oxygen-tension ranges; probably most are facultative (*Peranema trichophorum*), but some (*Gonyostomum semen*) appear to be obligate aerobes, others (*Trepomonas rotans*) obligate anaerobes. Carbon dioxide is excreted through the contractile vacuole as is oxygen. Dense blooms of *Chlamydomonas* or *Euglena* will supersaturate surface waters with oxygen within a few hours after daybreak. Excretion of other metabolites such as antibiotics is demonstrable. Circulation by cyclosis is sometimes observable.

Ecology. Flagellates adjust to wide ranges of pH, osmotic pressure, light, and temperature but optima are found upon investigation. *Euglena mutabilis* grows at pH 0.9–7.2, but densest blooms occur about pH 3.5–4.0. Some flagellate species occur in both fresh and seawater, but cannot be transferred directly from fresh to seawater.

Almost any aqueous ecological niche contains Flagellata. Dense fresh-water blooms of *Euglena* or *Platymonas* are matched by blooms of *Eutreptia* or *Gymnodinium breve* in the ocean. *Trentonia flagellata* is the only green flagellate occurring abundantly in Warm Mineral Springs, an anaerobic sulfur spring of southern Florida. Euglenida and Phytomonadida grow profusely in barnyard pools, Chloromonadida in weakly acid cedar and cypress swamp water, and Phytomastigida and Cryptomonadida in lower-nutrient levels of oceans, streams, and lakes. Dinoflagellida abound in the oceans; Chrysomonadida tend to live in clean water; Zoomastigida and Rhizomastigida favor organically polluted water. There is much overlapping because many are ubiquitous.

Flagellata are near the base of the food pyramid, probably next to bacteria. Their varied synthetic abilities, fast reproductive rate, and huge numbers in both fresh and salt water compensate for their usual small size. The oceanic blooms which kill fish and other animals are principally flagellates, but green flagellates reaerate polluted water. They often cause tastes and odors in potable water, but they readily attack organic matter in natural water as do bacteria. They include many parasites dangerous to humans and animals and are themselves frequently parasitized. *See* FOOD WEB; PROTOZOA. James B. Lackey

Mastitis

Inflammation of the mammary gland. This condition is most frequently caused by infection of the gland with bacteria that are pathogenic for this organ. It has been described in humans, cows, sheep, goats, pigs, horses, and rabbits. Mastitis causes lactating women to experience pain when nursing the child, it damages mammary tissue, and the formation of scar tissue in the breast may cause disfigurement. It also is the most economically important disease of dairy cows.

The mammary gland is composed of a teat and a glandular portion. The number of glands varies with the species. The glandular portion consists of lobules of secretory tissue and a series of collecting ducts. The teat is composed of an external orifice, teat duct, and teat lumen. In humans there are a series of lactose-secreting ducts opening on the nipple surface. The gland has defensive mechanisms to prevent and overcome infection with bacteria. Nonspecific defense mechanisms include teat duct keratin, lactoferrin, lactoperoxidase, and complement. Specific

defense mechanisms are mediated by antibodies and include opsonization, direct lysis of pathogens, and toxin neutralization. Milk contains epithelial cells, macrophages, neutrophils, and lymphocytes. To induce mastitis, a pathogen must first pass through the teat duct to enter the gland, survive the bacteriostatic and bactericidal mechanisms, and multiply. Bacteria possess virulence factors such as capsules and toxins which enable them to withstand these protective mechanisms. *See* LACTATION.

Agents and mode of transmission. Mammary infections in humans are usually caused by *Staphylococcus aureus*. The source of infection is typically the nasopharynx of the infant, although mass outbreaks of mastitis have been documented in maternity wards, suggesting spread between nursing women. Bovine mastitis is classified as either contagious or environmental in nature. Contagious mastitis is caused by *Streptococcus agalactiae* and *Staphylococcus aureus*. The term contagious reflects the role of infected glands as the critical reservoir of infection, with transmission occurring at milking. Environmental mastitis is caused by *Streptococcus uberis, Escherichia coli*, and *Klebsiella* spp., bacteria that are found on other parts of the cow and in bedding, manure, and soil. *Streptococcus aureus* frequently induces mastitis in goats, sheep, and rabbits, with *Pasteurella haemolytica* also being a prime culprit in sheep. In pigs *Escherichia coli* and *Klebsiella* spp. are isolated, while the occasional equine case usually yields streptococci. *See* ESCHERICHIA; KLEBSIELLA; PASTEURELLA; STAPHYLOCOCCUS; STREPTOCOCCUS.

Clinical signs. When bacteria multiply within the gland, there is a release of inflammatory mediators and an influx of neutrophils. The severity of mammary infection is classified according to the clinical signs. In humans, infection occurs during lactation, with clinical episodes most frequent during the first 2 months of lactation. In acute puerperal mastitis the tissue becomes hot, swollen, red, and painful, and a fever may be present. In the absence of treatment, this may progress to a pus-forming mastitis, with the development of breast abscesses.

Clinical mastitis in ruminants causes inflammatory changes as described above, and the milk may contain clots and be discolored. In severe cases the cow may display fever, anorexia, and depression. In subclinical mastitis the milk appears grossly normal, but analysis reveals an elevated cell count and the presence of bacteria. In cows the different etiologic agents vary in the severity of the disease they cause. Coliform bacteria cause a peracute mastitis which can be fatal. Infection with *Staphylococcus aureus* and streptococci often commences with an initial acute clinical episode that subsides into a chronic subclinical state. Occasional clinical flare-ups occur. In rabbits the disease occurs as either a gangrenous or abscess form.

Diagnosis. Clinical mastitis can be diagnosed by examination of the gland and its secretion. Subclinical mastitis cannot be identified by this means but relies on bacteriologic and cytologic analysis. Bacteria cultured from aseptically collected secretion are identified by cultural morphology and biochemical tests. In uninfected glands the predominant cell type in milk is the macrophage. Neutrophils are attracted once bacteria invade the gland, and these neutrophils cause the cell count to increase. The cell count thus serves as an indicator of the health of the gland. Mastitic milk possesses increased levels of sodium and chlorine ions, albumin, and acetylglucosiminidase, and displays altered electrical conductivity.

Treatment. In acute puerperal mastitis of humans suitable antibiotics are administered by the intravenous or intramuscular route, while in the abscess form surgical drainage is provided in addition to antibacterial therapy. In dairy cows the continuous manual removal of the secretion of infected quarters is usually replaced by antibiotic use. Penicillins, cephalosporins, and erythromycin are administered locally into the infected gland after milking for 1–2 days. Additional antibiotics are given systemically, that is, intravenously or intramuscularly, in severe cases of mastitis, and also to improve bacteriologic cure rates. Certain forms of bovine mastitis are unresponsive to therapy, and the animals are isolated from the herd. The milk from cows undergoing treatment is discarded to prevent contamination of milk for human consumption. The pathogenic effects of coliform mastitis in cows are mediated in part by the release of endotoxin from the bacterial cell which induces macrophages to secrete inflammatory cytokines, which in turn cause fever, hypotension, and shock. In this type of mastitis, antibiotics have a secondary role to the administration of fluids and anti-inflammatory agents, and the continuous removal of secretion from the infected gland after injection of oxytocin. *See* ANTIBIOTIC.

Prevention. Proper hygiene during nursing helps prevent breast infections in humans. In cows, good milking practices, milking machine maintenance, and the dipping of teats with disinfectant after milking help reduce infections caused by staphylococci and streptococci. However, these steps have no impact on infection with environmental organisms. For the last 2 months of pregnancy, cows are not milked; this is the so-called dry or nonlactating period during which infections commonly occur. Administering long-acting antibiotic formulations at drying-off helps to prevent new infections and to eliminate existing ones without the need to discard milk. Proper hygiene and sanitary housing are critical in the prevention of environmental mastitis. In the United States, immunization to increase resistance to infection is possible with a vaccine containing killed *Escherichia coli* bacteria.

Neil L. Norcross

Matched-field processing

A set of related signal processing techniques for remote sensing that involve matching measured signals obtained from an array of sensors to synthesized signals obtained from signal propagation

simulations conducted in a model of the actual environment. Here the signals are typically acoustic or electromagnetic waves, but they may be water waves, seismic waves, structural waves, or any other measurable phenomenon that travels in a predictable manner from its source to a distant receiver. Matched-field processing is the extension of elementary sensor-array signal processing techniques for line-of-sight signal propagation in uniform unbounded environments to complicated, but predictable, signal propagation in nonuniform bounded environments. Matched-field processing is superior to these elementary techniques because it can explicitly account for reflection, refraction, and scattering of the signal as it propagates from its source to the sensors. However, this superior performance is possible only when the signal propagation simulations are accurate and the model environment matches the environment in which the signal propagation took place. *See* ACOUSTIC SIGNAL PROCESSING; SIGNAL PROCESSING; SIMULATION; WAVE MOTION.

Remote sensing is common for military, biomedical, surveillance, quality control, and nondestructive evaluation purposes. The primary remote-sensing tasks are typically detection, localization, tracking, and identification of signal sources (or scatterers). These tasks all become more difficult when the environment is complicated and multiple signal-propagation paths are possible between distant signal sources (or scatterers) and the sensors. Such environments occur in applications of radar and sonar where naturally occurring or artificial obstacles and surfaces cause signal echoes and reflections. Matched-field processing is utilized primarily for detection and localization of signal sources, with a greater emphasis on localization. *See* RADAR; REMOTE SENSING; SONAR; UNDERWATER SOUND.

Techniques for remote source localization from sensor-array measurements exist for a variety of remote-sensing scenarios involving different kinds of sources, propagation physics, and environments. For example, signal sources may be relatively small or large compared to the propagation distances, sensor-array size, or average signal wavelength; signal wave fronts may be planar, spherical, or more complicated; and environments may or may not include reflecting surfaces and refraction. Matched-field processing is the preferred source localization technique for small sources, complicated wave fronts, and environments with known reflection and refraction.

Homer Bucker is generally credited with putting matched-field processing into its current form in 1976. He studied remote acoustic source localization in the shallow ocean, and was the first to utilize a realistic environmental model and to introduce a two-dimensional surface as the matched-field processing output. In 1985, Richard Fizell and Stephen Wales reported the first matched-field processing localization results using field measurements of sound propagation under Arctic ice. From these beginnings in sonar, the development of matched-field processing has been boosted by incorporation of ideas from radar and astronomy, so that it is now used or being considered for applications in leak testing, noise-control engineering, seismology, structural monitoring, and even astrophysics.

Basic features. The capabilities of matched-field processing can be illustrated by comparing it to simple plane-wave sensor-array processing in a series of increasingly complicated environments. Consider first a uniform unbounded environment with a linear transducer array (**Fig. 1***a*). If a single plane-wave signal from a distant source propagates past the array at an angle $\theta = \theta_1$, then each transducer will record the same signal waveform at a slightly different time. When the signal propagation speed and the locations of the sensors are known, the recorded signal timing differences between sensors can be used to determine the propagation angle. Here, the plane-wave array-signal-processor output, $A(\theta)$, and the matched-field processing output, $B(\theta)$, will be equivalent and will involve only a single search coordinate, the angle θ (Fig. 1*b*). Both signal processor outputs will peak at the angle θ_1, and the width of this peak and its height relative to A or B values away from θ_1 will be determined by the number and spatial distribution of the array elements, the signal-to-noise ratio and timing accuracy of the signal recordings, and the signal processing algorithms. False peaks and unwanted nonzero values in an array signal processor's output are commonly called side lobes and clutter.

As a second step, consider the same source-array geometry when a reflecting surface causes the original signal to traverse the receiving array in two different directions, θ_1 and θ_2 (**Fig. 2***a*). In this case, the plane-wave array-signal-processor output, A, will indicate two possible source directions, but a matched-field processing scheme based on signal propagation in the actual environment will account for the reflecting surface and will correctly determine the true source direction (Fig. 2*b*). In fact, the matched-field processing scheme does a better job finding the source angle when the reflecting surface is present, compared to when it is absent, because matched-field processing utilizes the signal timing information from both signal paths.

As a final illustration consider an environment with

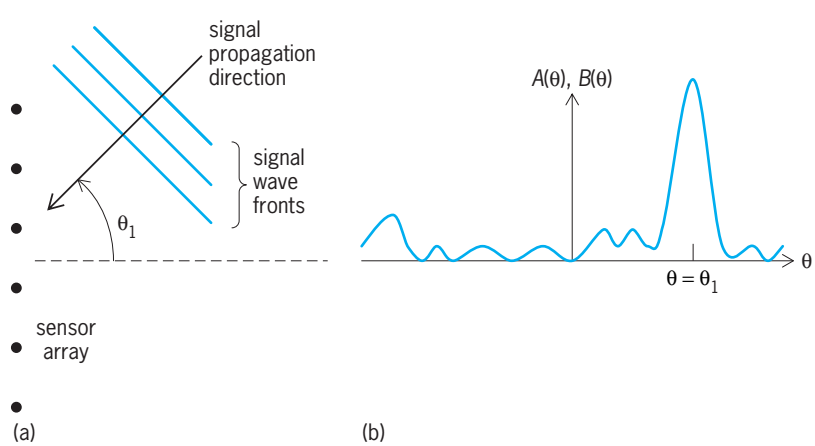

Fig. 1. Plane-wave signal incident on a linear sensor array without any reflections. (*a*) Signal and sensor-array geometry. (*b*) Typical plane-wave array-signal-processor output, $A(\theta)$, and the matched-field processing output, $B(\theta)$, versus signal incident angle, θ.

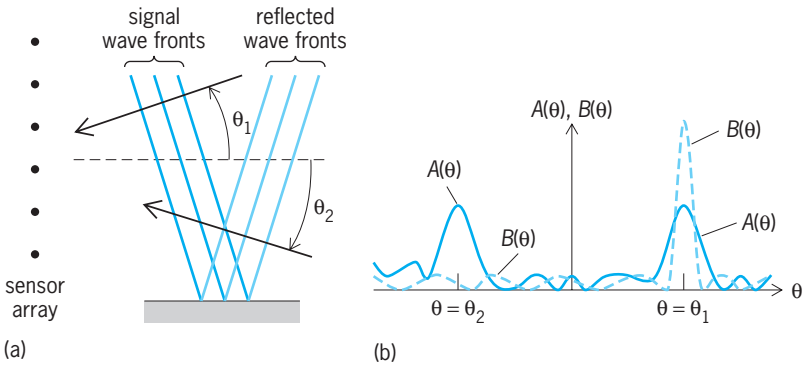

Fig. 2. Plane-wave signal incident on a linear sensor array with one reflection. (*a*) Signal and sensor-array geometry. (*b*) Typical plane-wave array-signal-processor output, *A*(*θ*), and the matched-field processing output, *B*(*θ*), versus signal incident angle, *θ*.

two parallel reflecting surfaces that form a waveguide with four signal paths connecting the source and the sensor array, each path having a different angle (**Fig. 3***a*). This situation is common for ocean sonar operations, but even more source-array paths may be involved. For the case depicted here, the plane-wave array signal processor will indicate four possible source directions (Fig. 3*b*). However, when the environment is known, matched-field processing can exploit the environmental complexity to determine both the horizontal and vertical location of the source. Thus, the matched-field processing output is commonly presented as a surface having two (or more) independent spatial dimensions with color or gray-scale indicating the peaks (Fig. 3*c*). Even though clutter and side lobes still occur with matched-field processing, this example shows how matched-field processing can improve source localization in a multipath environment by eliminating ambiguous source directions in favor of a single source localization peak in range and depth coordinates.

Algorithms. The actual signal processing algorithms used in matched-field processing fall into two broad categories: linear and adaptive (or nonlinear). The primary linear algorithm, often called the Bartlett processor, is implemented by placing a grid of points in the model environment in the region to be searched for the signal source. A synthetic test source is placed at each grid point and signals for each test source location are synthesized for each sensor of the receiving array using a signal propagation simulation. The measured signals are then compared with the synthetic signals. The test source location corresponding to the best match between measured and synthesized signals is the location of the source determined by matched-field processing. The Bartlett processor can also be described in terms of time-forward and time-reverse signal propagation. First, genuine signals travel in the actual environment from the source to the sensors where they are recorded with time increasing, as usual. Next, the direction of time is mathematically reversed for the measured signals and time-reverse propagation is simulated in the model environment. If the propagation simulations are accurate, and the actual and model environments are well matched, then the location in the model environment where the time-reversed signals converge, constructively interfere, and form a peak is the matched-field processing prediction of their source location (Fig. 3*c*). For an oscillating signal in an environment that supports multiple reflections from different directions, linear matched-field processing is capable of localizing a remote source to within a half wavelength. Bartlett matched-field processing is considered linear because the measured signals are directly correlated with the synthesized signals via simple multiplication and sum operations. *See* TIME-REVERSED SIGNAL PROCESSING.

Adaptive matched-field processing differs from linear matched-field processing in the way that the actual and synthetic signals are combined. Although there are many adaptive matched-field processing algorithms, their general intent is to improve source localization resolution and clutter and side-lobe suppression beyond that of linear matched-field processing. Hence, these algorithms are mathematically more complicated, so the descriptions above involving signal correlations and time reversal are no longer directly applicable. However, the implementation procedure for adaptive and linear matched-field processing is essentially the same, but the quality of the results from adaptive matched-field processing can be much better, especially when the signal-to-noise ratio of the original measured signals is high. Under the right conditions, adaptive matched-field processing routines may reduce localization uncertainty by

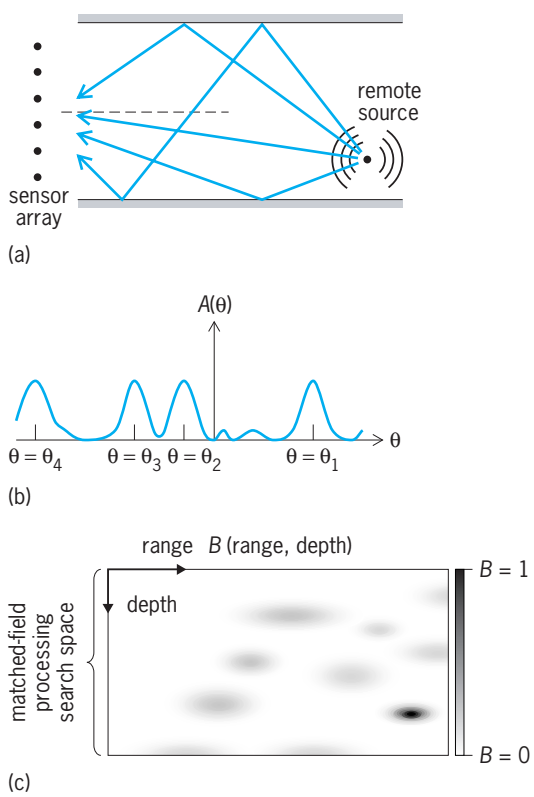

Fig. 3. Signals from a remote source incident on a linear sensor array in a waveguide with two reflecting surfaces and four signal paths connecting the source and the array. (*a*) Signal and sensor-array geometry. (*b*) Typical plane-wave array-signal-processor output, *A*. (*c*) Matched-field output, *B*, in grayscale as a function of range and depth.

a factor of 3 or more, and may suppress side lobes and clutter by more than a factor of 10, when compared to linear matched-field processing. These advantages may be tactically important for correctly discriminating between a single isolated source and two closely spaced sources, or for finding a weak target source in an environment dominated by loud interfering sources.

Limitations. The primary limitations of matched-field processing stem from differences between actual signal propagation in the real environment and simulated signal propagation in the model environment. Here, geometrical differences as small as a fraction of a signal wavelength in the placement of a reflecting surface can lead to erroneous source localization, or degradation (or even loss) of the localization peak in the matched-field processing output. Thus, there is a significant burden on the user to know the environment well and to faithfully simulate signal propagation. Yet, even when all the environmental-specification parameters are not known, matched-field processing is still possible by adding parameter searches to the preexisting search space, an extension of matched-field processing called focalization. However, the main computational burden of matched-field processing commonly results from repeated simulation of signal propagation in the model environment as the test source is moved through the search space. Thus, minimizing the number of signal propagation simulations is typically advantageous. David R. Dowling

Bibliography. A. Baggeroer, W. A. Kuperman, and P. N. Mikhalevsky, An overview of matched field methods in ocean acoustics, *IEEE J. Oceanic Eng.*, 18:401–424, 1993; F. B. Jensen et al., *Computational Ocean Acoustics*, chap. 10, AIP Press, 1994; A. Tolstoy, *Matched Field Processing for Underwater Acoustics*, World Scientific, Singapore, 1993.

Material resource planning

A formal computerized approach to inventory planning, manufacturing scheduling, supplier scheduling, and overall corporate planning. The material requirements planning (MRP) system provides the user with information about timing (when to order) and quantity (how much to order), generates new orders, and reschedules existing orders as necessary to meet the changing requirements of customers and manufacturing. The system is driven by change and constantly recalculates material requirements based on actual forecast orders. It makes adjustments for possible problems prior to their occurrence, as opposed to traditional control systems, which looked at more historical demand and reacted to existing problems. Material resource planning (MRP II) evolved from material requirements planning when it was recognized that most major data needed to manage a manufacturing or distribution firm could be obtained from the material requirements planning information. The ability of a material resource planning system to meet the various needs of manufacturing, materials, and marketing personnel within a changing business environment contributes to its growing implementation by manufacturing companies. *See* MANUFACTURING ENGINEERING.

Dependent and independent demand. The logic of the material requirements planning system is based on the principle of dependent demand, a term describing the direct relationship between demand for one item and demand for a higher-level assembly part or component. For example, the demand for the number of wheel assemblies on a bicycle is directly related to the number of bicycles planned for production; further, the demand for tires is directly dependent on the demand for wheel assemblies. In most manufacturing businesses, the bulk of the raw material and in-process inventories are subject to dependent demand. An important characteristic of dependent demand items that affects timing is known as lumpy demand. Since most manufacturers assemble their products in batches or specific lot sizes to fulfill customer requirements, dependent demand items at lower levels do not exhibit uniform usage but are subject to extreme fluctuation. This creates situations in which no demand exists for weeks, and then a large quantity is required. **Figure 1** illustrates the concept of lumpy demand on three styles of drills. While demand for the end drills is constant, the plant produces them in various lot sizes, thereby creating discontinuous and discrete demands. Applications by corporations of the just-in-time (JIT) philosophy will help smooth lumpiness to a certain extent since JIT argues for running smaller lot sizes. Lumpy demand is handled in material requirements planning through the calculation of future requirements at all levels of the assembled product. Dependent demand quantities are calculated, while independent demand items are forecast. Independent demand is unrelated to a higher-level item that the company manufactures or stocks. Generally, independent demand items are carried in finished goods inventory and subject to uncertain end-customer demand. Spare parts or replacement requirements for a drill press are an example of an independent demand item.

Material resource planning. In essence, material resource planning (also termed closed-loop material

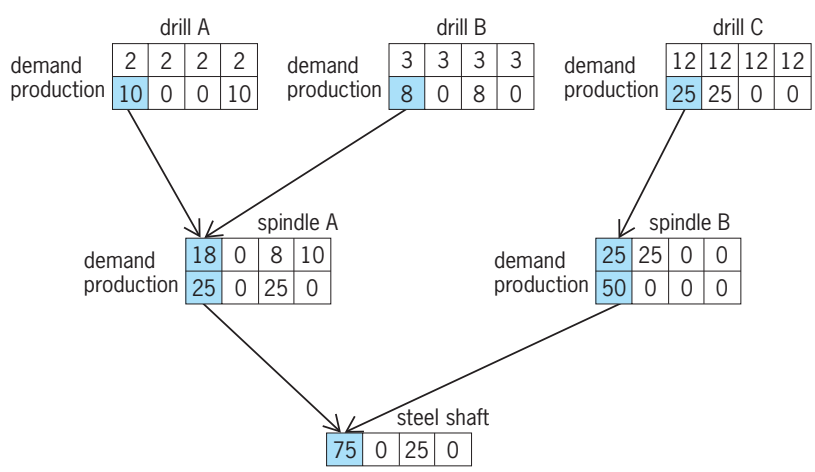

Fig. 1. Lumpy demand in manufacturing for three types of drills. The boxes represent equal time periods.

requirements planning) has the effect of extending the material requirements planning scheduling into all major functions of the firm. Material requirements planning provided firms with the ability to schedule and manage priorities. As the accuracy in material requirements planning increased, firms realized they could use these numbers in their business and financial plans. There was no longer a need for two sets of numbers. Operating and financial data are not kept on one system. Further, top management can use material resource planning as a system since it affects and is affected by sales, inventories, cash flow, production levels, and purchasing activities. Further, it ties key functional groups into the planning cycle.

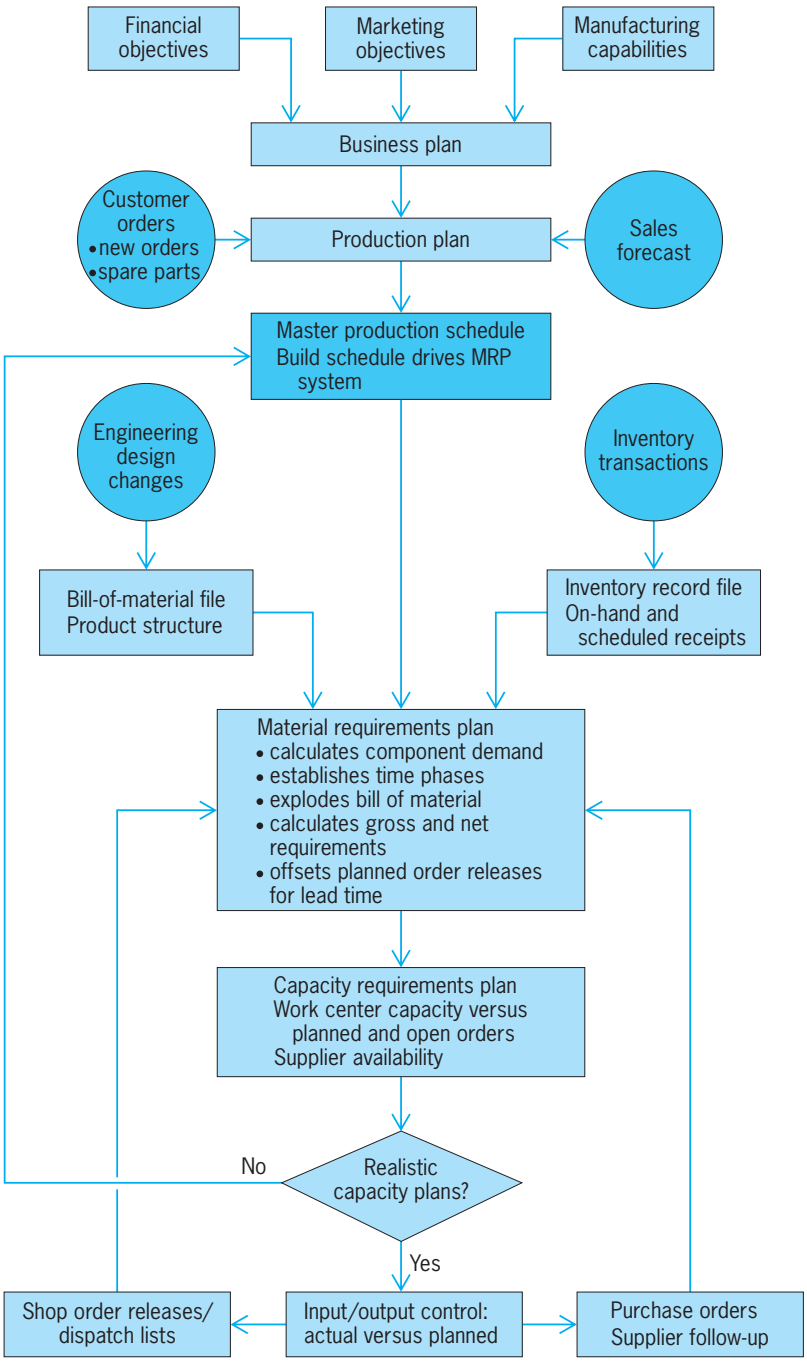

Fig. 2. Flowchart for a model material resource planning system.

For example, the marketing group can use data from master production schedules to plan and promise customer orders. The engineering group's impact on the system has become more visible since changes in design now are reflected by the information system. If drawings are not completed on time, the schedule delay will be visible on the material resource plan. Lastly, this plan permits firms to ask "what if" questions and analyze the effects of various factors on the overall operation.

Model system. A model material resource planning system is shown in **Fig. 2**. The business plan prepared by top management is based on certain expectations of sales from forecasts, firm orders, or a combination of sales forecasts and firm orders. The master production schedule refers to the products that are actually going to be produced in a particular time period. Demand inputs on the typical firm can originate from several sources: customers, warehouses, service parts, safety stock, and interplant orders. Except for make-to-order firms, the marketing forecast is adjusted periodically and reviewed by the plant manager as well as functional managers from the marketing, manufacturing, and materials groups. The marketing forecast is then translated into a production plan. This plan establishes an aggregate production rate which considers planned inventory increases or decreases. In a make-to-order company, production planning adjusts the customer order backlog; in a make-to-stock firm, the inventory levels are adjusted.

Master production schedule. While the production plan is a general plan, the master production schedule is the shop-build schedule which drives the material requirements planning system. The master production schedule is stated in bill-of-material numbers, and is used to generate the material requirements plan, and also generates the capacity requirements plan. This master schedule also lists all the end items (the items as sold by various users) by date (planned shipping period) and their quantity.

The master production schedule gives management an approximation of long-term capacity utilization, labor needs, the impact of new product plans in the shop load, and current priority status.

Bill-of-material file. This contains information about all parts and shows the dependent demand relationship of subassemblies or end items (finished product) down to components and raw material level. For example, a bill of material for a screwdriver would show the plastic handle, forged shaft, and steel from which the shaft was made. If users are to obtain valid system outputs, the bill-of-material file must be accurate and up to date, and must be suitably defined for customers, manufacturing, and purchasing. Users must ensure that the latest engineering changes are incorporated into the bill of material.

Inventory record file. This contains a record of the actual inventory level of each item and part as well as lead times. Inventory accuracy is critical in avoiding erroneous system outputs. Most users strive for 95–99% accuracy and monitor actual inventory through cycle counting, which involves periodically counting inventory.

Master Schedule		week 1	2	3	4	5	6
level 0 skateboards	Planned order release	40	0	50	0	0	60
level 1 axles lead time = 2	Project requirements	80	0	100	0	0	120
	Scheduled receipts						200
	On hand 200	120	120	20	0	0	80
	Planned order release				200		
level 2 steel rod lead time = 3	Projected requirements	0	0	0	200	0	0
	Scheduled receipts				200		
	On hand 10	10	10	10	10	10	10
	Planned order release	200					

offset

offset

Fig. 3. Material requirements plan mechanics and lead-time offsetting for skateboards.

Material requirements plan. The function of this plan is to explode the master schedule against the bill-of-material file to determine the gross requirements necessary to make the product. These gross requirements are then compared to inventory on hand to determine net requirements. The net requirements at the higher level become gross requirements at the next lower level, and so on, down through the bill of material. These requirements are offset for lead time by subtracting the lead time from the date of the net requirements. **Figure 3** shows the material requirements planning format and how lead times are offset. In Fig. 3 the term projected requirements refers to the number of items necessary to make up the next-higher assembly (for example, 40 skateboards require 80 axles). The term scheduled receipts refers to material already on order or scheduled to come in. The inventory of the item is represented by the term on hand. Planned order release (net requirements) means the number of items that manufacturing must make or purchasing must buy to meet higher-level requirements. Level 0 is the highest level and represents the end item. Lead time is the time required to acquire an item after it is ordered or manufactured.

Capacity requirements plan. This part of the system enables manufacturing to determine what capacity will be required by work centers during each time period to meet the master production schedule. The material requirements planning system will indicate which items are required and the timing of the items. The actual work load is then compared to work center capacity. If the available capacity falls short in meeting the master schedule, a decision must be made either to reduce the schedule or to expand capacity by working overtime, subcontracting, or purchasing additional capacity.

Input/output control. This part of the system monitors the levels of work into and out of each work center. Control is achieved by ensuring on-schedule job flow through various work centers. If a job slips behind schedule, its priority should be increased. When input to the factory begins to exceed output, this indicates that the production schedule has been overstated and should be reduced or the capacity expanded.

Dispatch lists and vendor follow-up lists. These help manufacturing and purchasing supervisors to determine accurate priorities. The dispatch list will print out jobs requested to be run and jobs available to be run, and the sequence in which to run them. Dispatch lists, usually generated daily, continually revise priorities on the shop floor. Follow-up lists give the purchasing department a tool to ensure that suppliers will deliver on time by providing information about requirements on a short- and long-term basis.

Advantages. Material requirements and material resource planning systems offer users multiple benefits. From a broad perspective, this type of system provides a tool for total business planning of sales, production, cash flow, and so forth. It also provides a tool for marketing and provides better customer service through higher on-time completion dates, accurate order promise dates, and follow-up information. Other productivity enhancers include more professional purchasing, lower inventories, improved planning by shop supervisors, and increased direct labor productivity. By use of the computer, material resource planning manipulates massive amounts of data to keep schedules up to date and priorities in order and valid.

The technological advances in computing and processing power, the benefits of on-line capabilities, and reduction in computing cost make computerized manufacturing planning and control systems such as material resource planning powerful tools in operating modern manufacturing systems productively. *See* ACTIVITY-BASED COSTING; INDUSTRIAL ENGINEERING; INVENTORY CONTROL; SYSTEMS ENGINEERING.

Larry C. Giunipero

Bibliography. R. B. Chase and N. J. Aquilano, *Production and Operations Management*, 8th ed., 1998; *J. N. Petroff Handbook of MRP II and JIT: Strategies for Total Manufacturing Control*, 1993; R. J. Schonberger and E. M. Knod, *Operations Management: Improving Customer Service*, 5th ed., 1993; O. W. Wight, *MRP II: Unlocking America's Productivity Potential*, 1983.

Materials handling

The movement, protection, storage, and control of materials and products throughout the process of their manufacture, distribution, consumption, and disposal, as defined by the Material Handling Industry of America. Materials handling is not just about equipment for storage and movement, it also includes controls that direct movement, packaging tools, and techniques that protect goods as they flow from point to point. However, no amount of equipment, automation, or information technology can take the place of a well-designed system that facilitates the steady flow of goods through manufacturing and distribution facilities. *See* INDUSTRIAL FACILITIES.

Materials-handling systems are designed to make sure that materials and goods are delivered to the right place, at the right time, in the right condition, with the correct information (labeling and invoices or other paperwork) in a safe and cost-effective manner as they flow through networks of suppliers and customers. Sometimes materials handling is not considered a value-added activity, but it adds value by putting things where they need to be for use at the right time. Items that are damaged or delivered to the wrong location have less value than those delivered when and where they are needed for further processing or use. This is known as the "time-and-place value" of goods and materials.

Types of handling. Materials handling is typically classified as bulk or discrete handling. Bulk handling deals with the controlled movement and storage of materials, such as coal, grains, liquids, or other types of materials, which flow from point to point and are measured in terms of their volume or mass. Mining operations, chemical processing facilities, and oil refineries are bulk-handling operations. Discrete handling deals with individual items or groups of items packaged together to form what is known as a unit load. Unit loads are stored and moved as if they were a single item. They can be items grouped together in a box, a group of boxes on a pallet, or simply one individual item. Unit-load formation is a fundamental concept of discrete handling. Many criteria influence the decision of how to form unit loads, such as size, weight, available equipment, and how often the goods are to be stored or moved. The design of manufacturing facilities, assembly lines, and warehouses are greatly influenced by the tradeoffs necessary in determining the quantity of a unit load. Examples of discrete handling are most often found in manufacturing, assembly, and warehouses. If you have ever picked up your bags from an airport luggage carousel, you have also experienced discrete-parts materials handling. *See* BULK-HANDLING MACHINES.

Equipment. The types of materials handling equipment applied depends on the goals of the facility.

Starting simple. The most common materials-handling equipment is the industrial truck (forklift), which picks up a pallet-load of goods and moves it from place to place, typically between a storage location and a truck at a loading dock. Despite its high labor cost, this is the simplest and most flexible point-to-point movement of goods. The corrugated boxes (cardboard is a different type of material) holding the goods are also part of materials handling. If they are shrink-wrapped to keep them secure, then yet another materials-handling product is being used. It is also possible to use manually powered devices, such as pallet jacks, carts, and many other kinds of equipment, to move the items. Lists of items coming into the facility, sitting in storage, being picked from storage, and being shipped out of the facility can be tracked or recorded on 3×5 index cards if the number of different kinds of items (also known as stock keeping units or SKUs) is small, the storage scheme is simple, and the flow of goods infrequent. However, as needs change and customer demands grow, the typical facility needs to consider other kinds of equipment, techniques, and information technology to meet the challenge. *See* INDUSTRIAL TRUCKS.

Growing complexity. When a facility runs out of space, it might stack pallets higher or consider adding industrial, steel storage racks to use vertical space rather than continuing to spread out horizontally. If customers would order more if they could buy smaller quantities of a greater number of items but shipped more often, then a facility might need to consider shelving, carton flow racks, carousels, and many other kinds of equipment that might better address individual items rather than pallet loads. For moving large, heavy, or bulky items, a facility might consider using cranes, hoists, and various types of attachments. To reduce the amount of packaging consumed to ship goods to regular customers, reusable pallets and containers should be considered. There is a wide range of equipment with features to fit a wide range of handling tasks. Fitting the right equipment and techniques to the task is part of materials handling. *See* ELEVATING MACHINES; HOISTING MACHINES.

New materials handling workhorses. The changing manufacturing and distribution environment requires the high-speed routing of goods through a facility's system of pathways, linking receiving and storage areas with manufacturing, packaging, assembly, and shipping points. The new workhorses of materials handling are "sortation" conveyors—equipment that makes materials flow, then diverts them to another section of conveyor—coupled with automatic identification equipment (for example, barcode scanners). Goods are free to flow to interconnected points because the identification equipment, sensors, controls, and software route and redirect them so they arrive at the right locations at the right time. Packages shipped through any of the major overnight delivery services likely have flowed through miles of conveyors and past many sensors and barcode readers. *See* CONVEYOR; MONORAIL.

Control and information systems. In addition to hardware, a materials-handling system has a control system, as well as information-gathering (sensors) and communication components. In a distribution network, manufacturing, warehousing, and transportation execution systems (often collectively

referred to as supply-chain execution system components) control the movement (when and where) of goods as they flow through the system. These systems not only track the goods in a facility, many of them also track the equipment resources used to move the goods and the labor used to perform tasks. *See* CONTROL SYSTEMS.

Planning and handling techniques. All the handling equipment and information technology must be coupled with good planning and the use of proven materials-handling techniques. For instance, nearly all manufacturing and distribution facilities tend to follow the 80/20 rule (Pareto's law); that is, 80% of the product flow will be the result of 20% of the types of items handled by the facility. As a result, it makes sense to locate the 20% of the fast-moving items closer to the people or machines performing the movements. In a technique called crossdocking, a facility is organized to coordinate the receipt and shipment of items, such that they can be taken directly to a staging area near a truck dock or put directly on a waiting truck, eliminating the need for storage. Another technique, known as task interleaving, alerts workers between tasks that there is a task, nearby or on the way to another, they could perform. Task interleaving increases worker, manufacturing, and distribution productivity. Theses are just a few examples of materials-handling techniques. There are nearly as many techniques as there are types of equipment for performing them. Applying the right equipment, information technologies, and techniques to ensure the timely, safe, and cost-effective delivery of goods and materials is the essence of materials-handling system design. *See* INDUSTRIAL ENGINEERING; PARETO ANALYSIS; PRODUCTION ENGINEERING; PRODUCTION PLANNING; TRANSPORTATION ENGINEERING.

Michael Ogle

Bibliography. R. G. Askin and C. R. Standridge, *Modeling and Analysis of Manufacturing Systems*, Wiley, 1993; S. Heragu, *Facilities Design*, PWS Publishing, 1997; F. E. Meyers and M. P. Stephens, *Manufacturing Facilities Design and Material Handling*, 2d ed., Prentice Hall, NJ, 2000; D. R. Sule, *Manufacturing Facilities: Location, Planning, and Design*, 2d ed., PWS Publishing, 1994; J. A. Tompkins et al., *Facilities Planning*, 3d ed., Wiley, New York, 2003.

Materials science and engineering

The term "materials science and engineering" emerged in the 1960s to represent how the engineering application of materials was increasingly based on scientific principles, rather than the empiricism typically used prior to World War II. Materials science focuses on understanding how the properties (or characteristics) of materials are based on their atomic- and microscopic-scale structure. The term microstructure refers to this microscopic-scale structure. Materials science grew out of the fields of physical chemistry, polymer chemistry, and condensed-matter physics. *See* PHYSICAL CHEMISTRY; SOLID-STATE PHYSICS.

The engineering of materials involves the processing, testing, and selection of materials for various design applications. Materials engineering grew out of the fields of metallurgy, plastics engineering, and ceramic engineering. *See* METALLURGY.

Types of materials. Most engineering designs incorporate a limited "menu" of materials, namely metals, polymers, and ceramics. As material properties are based on the structural arrangement of atoms, these categories follow directly from primary types of atomic bonding, namely metallic, covalent, and ionic.

Metals. Most of the elements in the periodic table exhibit atomic bonding that is metallic in nature. An alloy is a metal composed of more than one element. Copper and its alloys, such as brass (copper plus zinc), are common examples. The metallic bond involves a mobile "gas" of electrons that produces the electrical conductivity, thermal conductivity, and optical reflectivity characteristic of metals. The crystalline (regular and repeating) arrangement of atoms in metals generally provides attractive mechanical behavior in structural applications. Many alloys are reasonably strong and can be formed into practical shapes. *See* ALLOY; METAL; PERIODIC TABLE.

The dominant examples are the iron-based or ferrous alloys, which include carbon and alloy steels and the cast irons. Although ferrous alloys represent the majority of metal applications in engineering designs, nonferrous alloys also are widely used. Major examples include alloys of aluminum, magnesium, titanium, and copper. Other examples include alloys based on zinc, lead (including solders alloyed with tin), refractory metals such as tungsten, and precious metals such as gold, platinum, and silver. *See* FERROALLOY; IRON ALLOYS.

Although most commercial metal alloys are crystalline in nature, rapid processing techniques can produce noncrystalline (amorphous) alloys. The resulting lack of periodic crystalline structure provides these amorphous alloys with some attractive properties, such as high strength combined with high deformability and excellent corrosion resistance.

Polymers. Polymers are composed of a few elements in the periodic table (usually carbon and hydrogen and possibly a few of the nonmetallic elements such as oxygen, nitrogen, fluorine, and silicon) that exhibit covalent atomic bonding. Polymers with carbon backbones (a carbon-carbon bonded chain) are the most common. A very common polymer is polyethylene $(C_2H_4)_n$, where n is the degree of polymerization, a number around 1000 representing the number of ethylene monomers linked by covalent bonding. Other polymers can contain oxygen (for example, acrylics), nitrogen (for example, nylon), fluorine (for example, fluoropolymers), and silicon (silicones). *See* POLYACRYLATE RESIN; POLYAMIDE RESINS; POLYFLUOROOLEFIN RESINS; POLYMER; POLYOLEFIN RESINS; SILICONE RESINS.

Polymers are generally good electrical insulators because directional covalent bonding between adjacent atoms usually does not provide "free" electrons

for conduction. The use of polymeric insulation on electrical wires is a practical example. Polymers also can be optically transparent (for example, clear plastic wrap for food). The alternative name "plastic" comes from the extensive formability of many polymers during fabrication. *See* DIELECTRIC MATERIALS.

Polymers composed of high-molecular-weight molecules, such as $(C_2H_4)_n$ (for example, $n > 100$), often have a "spaghetti-like" structure, with weak secondary nonbonding interactions between molecules. The result is relatively low strength compared to metals. Nonetheless, many polymers have been engineered to have sufficient strength and stiffness which, when combined with their characteristically low density, allows them to be increasingly popular substitutes for traditional structural metals. Contemporary automotive design provides numerous examples of metal substitution by polymers.

Ceramics (and glasses). Ceramics can be defined as nonmetallic and inorganic materials. Ceramics are chemical combinations of at least one metallic and at least one nonmetallic element (especially oxygen, with others possibly including carbon, nitrogen, phosphorous, and sulfur). Aluminum oxide (Al_2O_3) is a common example in which the metallic element aluminum is combined with the nonmetallic element oxygen. In fact, such chemical combinations are a fundamental tendency in nature, where metals chemically combine with nonmetallic elements from the environment. The rusting of iron is an example, both familiar and costly. Engineers also use similar environmental reactions to their advantage. Aluminum is a durable metal for structural applications because it can be anodized, a process in which a thin layer of aluminum oxide is produced on its surface. The aluminum oxide layer, in turn, usually protects the aluminum from further environmental degradation. *See* CERAMICS.

The ionic bond between metallic and nonmetallic atoms is a strong one. In aluminum oxide, the aluminum is a cation, Al^{3+}, and oxygen is an anion, O^{2-}. These oppositely charged ions are strongly attractive in an ionic bond. The thermal stability associated with the ionic bond makes ceramics resistant to high temperatures. The melting point of aluminum is 660°C (1220°F), while the melting point of aluminum oxide is 2020°C (3670°F). This heat resistance leads to the use of ceramics in refractory applications such as furnaces and boilers. *See* REFRACTORY.

Ceramics usually are crystalline in nature. Compared with metals and alloys, the greater structural complexity of ceramic structures leads to generally brittle mechanical behavior, limiting the applications of ceramics in many structural designs. *See* BRITTLENESS.

Although forming amorphous alloys is difficult, forming noncrystalline ceramics, called glasses, is relatively easy, as the crystallization of molten oxides can be a sluggish process. The most common examples are the silicate glasses, based on the mineral silicon dioxide (SiO_2). After melting, many of the silicate materials can be cooled to a rigid state, while retaining the random atomic structure of the melt. Common window and container glasses are made in this way, with a typical composition (by weight) of roughly 75% SiO_2, 15% Na_2O, and 10% CaO. *See* GLASS.

A common theme in processing ceramics and glasses is the presence of porosity (in the form of tiny air bubbles). These small bubbles serve as light-scattering centers, and as little as 3% porosity can turn a transparent material into a completely opaque one. The chalky appearance of many ceramics is the result. Porosity also weakens the material under structural loads.

Composites. Metals, polymers, and ceramics, which are fundamentally based on atomic bonding, cover the basic categories of structural materials. Engineers have developed many advanced materials based on the microscopic-scale combinations of them, resulting in a fourth category called composites. A common example is fiberglass, a composite of glass fibers (a few micrometers in diameter) embedded in a polymer matrix. In this case, the high strength of the small-diameter glass fibers is combined with the flexibility of the polymer matrix. Typical of structural composites, the fiberglass has properties superior to either component by itself. (Compared to fiberglass, the pure glass is more brittle and the pure epoxy is weaker.) More sophisticated composites include poly(*p*-phenylene terephthalamide) [Kevlar®] fibers and graphite fibers in various polymeric matrices, such as epoxy, polyester, polyetheretherketone (PEEK™), and poly(phenylene sulfide) [PPS]. *See* COMPOSITE MATERIAL; METAL MATRIX COMPOSITE; POLYMERIC COMPOSITE.

Not all composites involve fiber reinforcement. Aggregate composites use particles embedded in a matrix. Concrete, with rock and sand in a calcium aluminosilicate (cement) matrix, is a common example. Tungsten carbide particles in a matrix of cobalt make excellent cutting tools. *See* CONCRETE.

Semiconductors. Many structural materials can be classified adequately by the four categories described above. Materials also can be classified based on electrical conductivity rather than atomic bonding. Electrical classification provides a fifth category of materials—semiconductors. While metals are typically good electrical conductors and polymers and ceramics/glasses are typically electrical insulators, composites tend to have electrical properties that are the average of their individual components. Since the invention of the transistor at Bell Labs in 1948, covalently bonded semiconductors, with intermediate levels of electrical conductivity, have played an increasingly critical role in modern technology. The primary example is silicon, a central component of modern, solid-state electronics. Germanium is chemically similar to silicon and is a semiconductor widely used for electronic devices. Chemical compounds of elements close to silicon and germanium in the periodic table often display semiconduction. For example, gallium arsenide (GaAs) is a high-temperature rectifier and laser material. *See* SEMICONDUCTOR.

Advances in materials science and engineering. The field of materials science and engineering is a

dynamic one, with new materials and processing technologies developed by materials scientists constantly being applied by materials engineers to new engineering designs. Recent examples include the application of certain materials with zero electrical resistance (superconductors) as thin-film filters for improved cellular-telephone base stations. Micro-electro-mechanical systems (MEMS) include microscopic-scale gear systems fabricated by technologies developed for integrated circuits. And, finally, nanotechnology involves structures developed on the nanometer (10^{-9} m) scale, rather than the micrometer (10^{-6} m) scale. The resulting nanomaterials can display significantly improved mechanical, electrical, magnetic, thermal, and optical properties, compared to their counterparts on a coarser scale. *See* MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS); NANOTECHNOLOGY; SUPERCONDUCTIVITY.

James F. Shackelford

Bibliography. M. F. Ashby, *Materials Selection in Mechanical Design*, 2d ed., Butterworth-Heinemann, 1999; W. D. Callister, Jr., *Materials Science and Engineering: An Introduction*, 6th ed., Wiley, 2003; J. F. Shackelford, *Introduction to Materials Science for Engineers*, 6th ed., Prentice-Hall, 2005; W. F. Smith, *Foundations of Materials Science and Engineering*, 3d ed., McGraw-Hill, 2005; M. A. White, *Properties of Materials*, Oxford University Press, 1999.

Maternal behavior

The pattern of care given an offspring by its mother. Many species reproduce generation after generation without receiving or providing any parental care. Insects and fish commonly produce vast numbers of offspring that they neither feed nor defend, resulting in the loss of many offspring to predators and to other hazards. Their great numbers, however, ensure that some will survive and reproduce. Also, if the young of a species are self-sufficient at birth or if they mature very rapidly after birth, they can often survive and reproduce with little parental care.

Biological need. Parental behavior is most highly developed in species that produce only a few offspring at once, mature slowly, and have complex behavior patterns that can be learned only through extended practice. Parental behavior can involve as little as hiding the young and never seeing them again or it can involve years of feeding, defending, and teaching. In most species that care for their young after birth, the female does most or all of the work. The biological explanation for this arises from the fact that a male can father vast numbers of offspring, whereas a female can bear only a few. Each offspring that fails to survive or reproduce can represent a significant proportion of the female's lifetime reproductive output. Thus, the females that pass their genes on to subsequent generations in greatest numbers are not those that bear the most young but those that invest more in caring for the few young they produce.

Active care of the young by the male is most common when males of a species limit themselves to one or a few female partners. That is, when males forego the strategy of having many offspring, there is a premium on the success of each one. Examples of paternal investment are seen in many bird species and in monogamous primates such as humans and marmosets. The extended period of infancy and childhood in humans has also favored the evolution of paternal care of offspring. Prolonged immaturity increases the amount of care that each child requires and the number of children that must be cared for at once. Nonetheless, in most human cultures and under most circumstances, mothers contribute more than fathers. *See* REPRODUCTIVE BEHAVIOR.

Eliciting and maintaining maternal behavior. Factors that influence the onset of maternal behavior include physiological and psychological factors related to pregnancy, hormonal influences associated with childbirth, the behavior of the newborn, and cultural factors such as learning and social traditions. Among rodents, exposure to female hormones (estrogens) can cause both females and males to display maternal behavior, but hormone exposure is not necessary. Females who have never given birth and whose estrogen production is blocked will display maternal behavior over a period of time if they are simply exposed to the young. With longer exposure to the young, even males will begin to lick them and retrieve them when they leave the nest. Since the males do not produce female hormones in response to the young, the behavior of the offspring is itself a powerful stimulus for care-giving behavior. *See* ESTROGEN.

Although hormones play necessary roles in recovery from pregnancy and in milk production, there is no evidence that they play a critical role in any other aspect of human maternal behavior. Cultural factors and the behavior of the newborn are the primary cues that prompt human mothers to care for their young. Many factors contribute to maternal competence, that is, the mother's skill and consistency in caring for her infant or child. Hormones may increase her readiness to respond to infant behavior for a short time after birth, but it also seems likely that species' learning abilities have evolved in such a way that patterns of maternal behavior are easily learned. Ultimately, experience and practice are the dominant factors in human maternal behavior.

Human infants have a wide range of behaviors that communicate what their needs are and how urgent they are at a given moment, including smiling, crying, grasping, cuddling against the mother, and even changes in skin color if the infant is cold or extremely upset. Infants do not necessarily perform these behaviors with the intention of securing the mother's care, but the behaviors have the predictable outcome of attracting the mother's attention and care. By the time the child is two years old, language becomes a powerful guide to maternal care. *See* DEVELOPMENTAL PSYCHOLOGY.

Mothers can often tell what a baby needs from the context in which the demand occurs. For example,

the cries of a recently fed baby are unlikely to stem from hunger. There is no evidence that human mothers are biologically endowed with knowledge of appropriate child care. Observation of other mothers, help from parents, and general problem-solving abilities are key factors in human maternal competence. In most cultures, females begin to learn maternal behavior in early childhood. Adults encourage an interest in infants and provide opportunities for girls to practice parental behavior in their play and in caring for younger children. Much of what a mother needs to know is learned by trial and error after the birth of her baby.

Dimensions of maternal behavior. Mothers differ in the patterns of care they provide. Maternal behavior with infants is usually measured in terms of timing and appropriateness, and in terms of cooperation versus interference with the baby's ongoing behavior. After infancy, maternal behavior is typically demonstrated in expressions of warmth rather than hostility, or restrictiveness rather than permissiveness. These variables are modest but consistent predictors of the child's later behavior, just as a lack of supervision is the strongest predictor of delinquency.

Parenting stress and parenting failure. If a mother is very young, living in isolation, under a great deal of stress, or impaired by drugs or mental disorder, she may find caring for her infant to be a difficult learning task, and those same factors are associated with child neglect and physical abuse. Neglect and abuse occur in all social classes, but a limited education, stress, drugs, and mental disorder are more common when families live in isolation and poverty. *See* INSTINCTIVE BEHAVIOR; PERSONALITY THEORY.

Everett Waters; German Posada

Bibliography. L. Hoffman, R. Gandelman, and H. R. Schiffman (eds.), *Parenting*: *Its Causes and Consequences*, 1982; M. Lamb et al. (eds.), *Infant-Mother Attachment*: *The Origins and Developmental Significance of Individual Differences in Strange Situations Behavior*, 1985; P. Mussen and E. M. Hetherington (eds.), *Handbook of Child Psychology*, 1983; R. Trivers, *Social Evolution*, 1985.

Maternal influence

Effects on development that are attributable specifically to the maternal parent or to maternally inherited factors. This influence can take a variety of forms, including maternal-specific inheritance of genetic material (for example, mitochondrial DNA), parent-specific epigenetic modification of genes (genomic imprinting), placement of molecules into the oocyte prior to fertilization, and interactions with the maternal parent during development.

The term "parent-of-origin effects" describes instances where an offspring's phenotype is asymmetrically dependent on the genotypes or phenotypes of its maternal and paternal parents. In most cases where such asymmetry exists, it is the maternal parent that wields the greater influence. This maternal advantage derives in part from the greater size of the oocyte, which means that the mother contributes more cytoplasmic material to the zygote than the father does. This asymmetry is enhanced in species where a part of the offspring's development occurs in physical contact with the mother, as in viviparous animals and many plants. Postnatal care, if present, is often provided preferentially by the mother.

Maternal DNA. One source of maternal influence is the inheritance of DNA specifically from the mother. Organelles that contain their own DNA, such as mitochondria and chloroplasts, are inherited cytoplasmically. Mitochondria are inherited predominantly, if not exclusively, from the mother. Many of the genes encoded by the mitochondrial DNA are involved in oxidative phosphorylation. A number of maternally transmitted diseases are characterized by oxidative phosphorylation defects; they affect in particular the central nervous system and muscle tissues.

The pattern of inheritance for chloroplasts varies, and apparently there may be different plant species in which chloroplasts are inherited exclusively maternally, exclusively paternally, and biparentally. The best-known phenotype associated with chloroplast inheritance occurs in the four-o'clock plant (*Mirabilis jalapa*), where leaf color varies depending on the ability of the chloroplasts to produce chlorophyll. In this species, chloroplasts, and therefore the leaf color trait, are always maternally inherited. *See* DEOXYRIBONUCLEIC ACID (DNA); MITOCHONDRIA; PLANT CELL.

Cytoplasmic factors. Many insect species possess cytoplasmic symbiotic bacteria, such as the well-known members of the *Wolbachia* and *Buchnera* genera. Inheritance of these bacteria bears some similarity to that of the cytoplasmic organelles, which may themselves be descended from bacteria that were once free-living. In many cases, these bacterial endosymbionts are restricted to specific host cells (bacteriocytes). Although cases of horizontal transfer have been documented, the predominant mode of propagation for these bacteria is maternal transmission. In this case, newly formed embryos are colonized by the bacteria present in the mother.

Maternal influence may also be exerted through molecules placed in the egg during oogenesis. Following fertilization, the newly formed genome in a zygote is not immediately transcriptionally active. In some cases, these genes are not transcribed until multiple rounds of cell division have occurred. During this earliest phase of development, the RNA and protein in the embryo predominantly reflect the genes of the mother. Small quantities of protein and mRNA are also carried into the zygote by the sperm, providing a possible source of paternal influence on early embryonic development. However, the much smaller size of the sperm makes the maternal contribution dominant in this context (**Fig. 1**).

Proteins in the egg that are active after fertilization are referred to as maternal-store proteins. In vertebrates, oogenesis is arrested in meiosis I, in some cases for years. These arrested oocytes still contain both of the mother's alleles at each diploid locus.

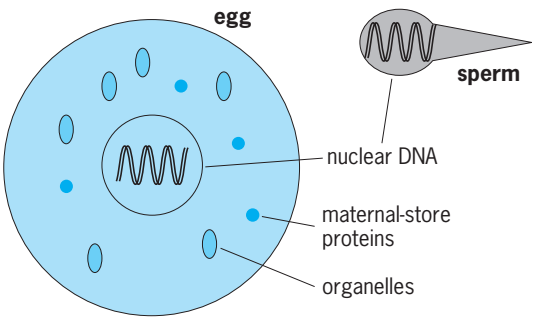

Fig. 1. In cytoplasmic inheritance, the gamete produced by the maternal parent is typically much larger than the gamete of the male parent. As a result, the egg contributes the majority of the cytoplasmic material to the zygote, including organelles that contain their own DNA as well as messenger RNA and proteins.

Maternal-store proteins are synthesized at this stage and inherited by the mature oocyte. In many invertebrates, polyploid "nurse cells" produce mRNA and proteins, which are transferred into oocytes through cytoplasmic bridges.

The reliance on maternal-store factors in early development means that there are certain phenotypes observable in offspring that depend on the genotype of the mother rather than that of the offspring itself. For instance, in preimplantation mouse embryos the maintenance of certain patterns of DNA methylation depends on the enzyme Dnmt1o, an oocyte-specific splice variant of the somatic maintenance methyltransferase Dnmt1. Homozygous knockout (*Dnmt1o*$^{-}$/*Dnmt1o*$^{-}$) offspring of a heterozygous (*Dnmt1o*$^{+}$/*Dnmt1o*$^{-}$) mother are phenotypically normal. Heterozygous offspring of a mother that is homozygous for the *Dnmt1o* knockout suffer methylation defects and embryonic lethality. *See* CELL ORGANIZATION; GENE ACTION; OOGENESIS; SPERMATOGENESIS.

Genomic imprinting. In some species, asymmetric parental influence derives from differential epigenetic modification of genes inherited by the offspring, or "genomic imprinting." An imprinted gene is one where an allele's pattern of expression differs depending on whether it was inherited from the mother or the father. In the simplest instances, one of the two alleles is active and one is transcriptionally silent (**Fig. 2**). In more complex cases, clusters of imprinted genes are coregulated, and various transcripts (including noncoding antisense RNA and splice variants) are produced from the maternally and paternally derived alleles.

The epigenetic modifications involved in imprinting include DNA methylation as well as histone modifications, such as methylation and acetylation. Epigenetic propagation mechanisms maintain these allele-specific modifications across cell divisions, so that each adult somatic cell contains two differentially marked alleles. These modifications are erased and reset in each generation in the germ line.

Genomic imprinting is the result of an evolutionary conflict between the maternally and paternally derived alleles within an individual. The optimal expression strategy for an allele may be different depending upon whether the allele is maternally derived or paternally derived. Alleles at an imprinted locus can be thought of as possessing a unique form of phenotypic plasticity—taking on different strategies depending on their parent of origin.

Imprinted genes have been identified in mammals and angiosperms, where the early part of development occurs in close physical contact with the maternal parent. In these cases, paternally derived genes are selected to place a greater demand for resources on the mother. Maternally derived genes favor more restraint, preserving maternal resources for the mother's other offspring. Many imprinted genes directly influence growth and resource demands during development. At loci where increased expression enhances growth, imprinted genes are paternally expressed and maternally silenced. At growth-suppressing loci, the opposite pattern of expression holds.

The pattern of allele silencing in genomic imprinting results in systematic functional differences between the maternally and paternally derived genomes in mammals and angiosperms. One consequence of this functional nonequivalence is that both the maternally derived and paternally derived chromosomes are required for successful development. Mouse zygotes formed from two maternally derived or two paternally derived pronuclei are not viable, and parthenogenesis (asexual female reproduction) is not seen in mammals. This contrasts with species that lack genomic imprinting, such as birds, amphibians, and insects, where it is possible to form viable offspring with two male or two female parents.

The dependence of normal development on particular epigenetic modifications produces one of the technical difficulties in mammalian cloning. Germ

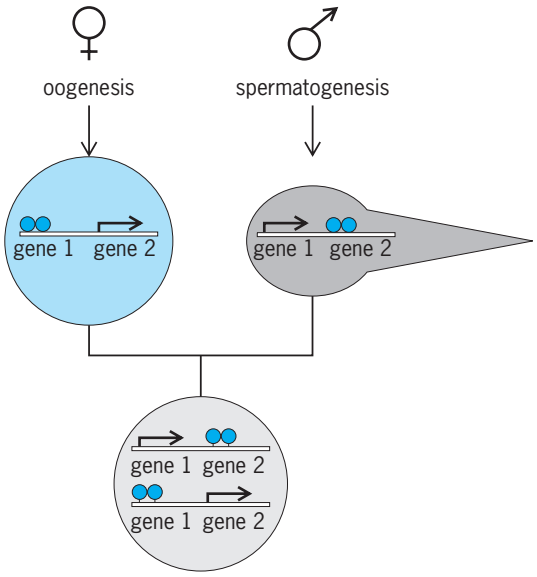

Fig. 2. In genomic imprinting in mammals and angiosperms, some genes are epigenetically modified in the germ line. Typically, only the copy inherited from one of the parents is transcriptionally active. This makes the offspring functionally haploid at these loci.

cells, which have been retained in or returned to a pluripotent state, contain only one of the two required patterns of imprinting. In many cases, the failure of cloning procedure to produce viable tissues or animals is associated with a failure of epigenetic reprogramming at imprinted loci.

Postfertilization influences. A final source of maternal influence results from interactions between the mother and offspring after fertilization. In many plants and in viviparous animals, the early part of development occurs in close physical contact with the mother. In the cases of mammals and angiosperms, specific organs exist (the placenta and endosperm, respectively) that mediate nutrient transfer from mother to offspring. Implantation experiments have demonstrated effects of the intrauterine environment on development in mice.

Further maternal influence arises in postnatal care. For instance, mammalian breast feeding involves the transfer of antibodies as well as nutrients to offspring. An offspring's immunity in the earliest stages of life therefore depends on its mother's genotype at various immune system loci, as well as her history of antigen exposure.

Where additional parental care is provided or where there is significant learning by offspring, the result is additional maternal or paternal influence. A trivial example might be the patrilineal "inheritance" of the last name in many human cultures. A more complex example would be song imitation and learning in songbirds. For example, in the zebra finch (*Taeniopygia guttata*), male birds' songs resemble those of their fathers as a result of cultural transmission. *See* DEVELOPMENTAL BIOLOGY; DEVELOPMENTAL GENETICS.

Bibliography. C. W. Birky, The inheritance of genes in mitochondria and chloroplasts: Laws, mechanisms, and models, *Annu. Rev. Genet.*, 35:125–148, 2001; A. Qvarnstrom and T. D. Price, Maternal effects, paternal effects and sexual selection, *Trends Ecol. Evol.*, 16(2):95–100, 2001; W. M. Rideout, K. Eggan, and R. Jaenisch, Nuclear cloning and epigenetic reprogramming of the genome, *Science*, 293(5532):1093–1098, 2001; W. Tadros and H. D. Lipshitz, Setting the stage for development: mRNA translation and stability during oocyte maturation and egg activation in *Drosophila*, *Dev. Dynamics*, 232(3):593–608, 2005; J. F. Wilkins and D. Haig, What good is genomic imprinting: The function of parent-specific gene expression, *Nat. Rev. Genet.*, 4(5):359–368, 2003.

Mathematical biology

The application of mathematics to biological systems. Just as physics underlies many engineering technologies, biology serves as a foundation for such disciplines as medicine, agriculture, animal husbandry, and environmental sciences, in which mathematical analyses have proved indispensable. Mathematical biology spans all levels of biological organization and biological function, from the configuration of biological macromolecules, as exemplified by the folding of proteins and the recognition properties of active sites in enzymes and receptors, to the entire ecosphere over the course of evolutionary time.

The important role of physics in the development of mathematical biology is natural, considering that physics is a much older science than biology. The influence of physics has been twofold. On the one hand, organisms simply are material systems, albeit of a very special kind, and presumably can be analyzed in the same terms and in the same way as any other material system. Reductionism, the theory that biological processes find their resolution in the particularities of physics, finds its practical embodiment in biophysics. Thus, one of the roots of mathematical biology is what was originally called mathematical biophysics. On the other hand, other early investigations in mathematical biology, such as population dynamics (mathematical ecology), did not attempt such a reductionistic analysis. Instead, they exploited the form of such analyses, such as using differential rate equations, but they expressed their analyses in frankly biological terms rather than physical terms. Such approaches were guided by analogy with mathematical physics rather than by reduction to physics and as a result rest on the form rather than the substance of physics. *See* BIOPHYSICS; PHYSICS.

Both of these approaches are important, especially since organisms possess characteristics (particularly in their development and evolution) that have no obvious counterpart in inorganic systems. As a result, mathematical biology has acquired an independent and unique character. In several important cases, these characteristics have required a reconsideration of physics itself, as in the impact of open systems on classical thermodynamics. *See* OPEN-SYSTEMS THERMODYNAMICS (BIOLOGY).

Surrogacy and models. The idea that something can be learned about a system by studying a different system, or surrogate, is central to all science, experimental as well as theoretical. Indeed, experimental or observational data would be useless if it pertained only to the specific system under observation. The relation between a system and its surrogates is embodied in the concept of a model. *See* MODEL THEORY.

For example, it is commonplace to use a laboratory animal such as a rat as a surrogate for a human. Data drawn from study of the rat can presumably be extrapolated to pertain to the human. It is likewise commonplace to treat a homogenate or extract as a surrogate for the intact material from which the extract was taken.

The basic idea of mathematical biology is that an appropriate formal or mathematical system may similarly be used as a surrogate for a biological system; that is, scientists can learn about an organism by interrogating an appropriate mathematical model the same way that they can learn about a human by interrogating a surrogate experimental animal. The use of mathematical models offers possibilities that, in many important ways, far transcend what can be

done on the basis of observation and experiment alone.

The richness of mathematical biology can become apparent if the mathematical roots of the concept of surrogacy itself are explored. For example, morphological differences between related species can be made to disappear by means of relatively simple coordinate transformations of the space in which the forms are embedded (**Figs. 1** and **2**). Surrogacy explicitly becomes a matter of intertransformability, or similarity, and what is true for morphology also holds true for other functional relationships that are characteristic of organisms, whether they be chemical, physical, or evolutionary. The concept of surrogacy is not confined to biology: the same kinds of ideas underlie the use of scale models in engineering design, the notions of congruence and symmetry in geometry and algebra, and the capability for analog computation in general.

In more modern terminology, these assertions of surrogacy and modeling can be restated: closely related implies similar. This is a nontrivial assertion: "closely related" is a metric relation pertaining to genotypes, whereas "similar" is an equivalence relation based on phenotypes. It is the similarity relation between phenotypes that, of course, provides the basis for surrogacy. Thus the question immediately arises: given a genotype, how far can it be varied or changed or mutated, and still preserve similarity?

Such questions fall mathematically into the province of stability theory, particularly structural stability. The interrelations between metric approximation and similarity is an exceedingly deep and subtle one. Under very general conditions, there exist many genomes that are unstable (bifurcation points) in the sense that however high a degree of metric approximation is chosen, the associated phenotypes may be dissimilar, that is, not intertransformable. That observation by R. Thom provides the basis for his theory of catastrophes and demonstrates the complexity of the surrogacy relationship. The fundamental importance of such ideas for phenomena of development, for evolution (particularly for macroevolution), and for the extrapolation of data from one species to another, or the relation between health and disease, is evident. It is also clear that a coherent mathematical framework is necessary for exploring these ideas. *See* CATASTROPHE THEORY.

Metaphor. A distinct but closely related group of ideas that are characteristic of mathematical biology may be described as metaphoric. One example of a metaphoric approach is the study of brain activities through the application of the properties of neural networks, that is, networks of interconnected boolean (binary-state) switches (**Figs. 3** and **4**). Appropriately configured switching networks are known to exhibit behaviors that are analogous to those that characterize the brain, such as learning, memory, and discrimination. That is, networks of neuronlike units can automatically manifest brainlike behaviors and can be regarded as metaphorical brains. Generally there is no attempt to explicitly model a particular brain as a particular neural network; the idea rather is to develop general properties of such networks, which might lead to a deeper understanding of those particular biological networks called brains.

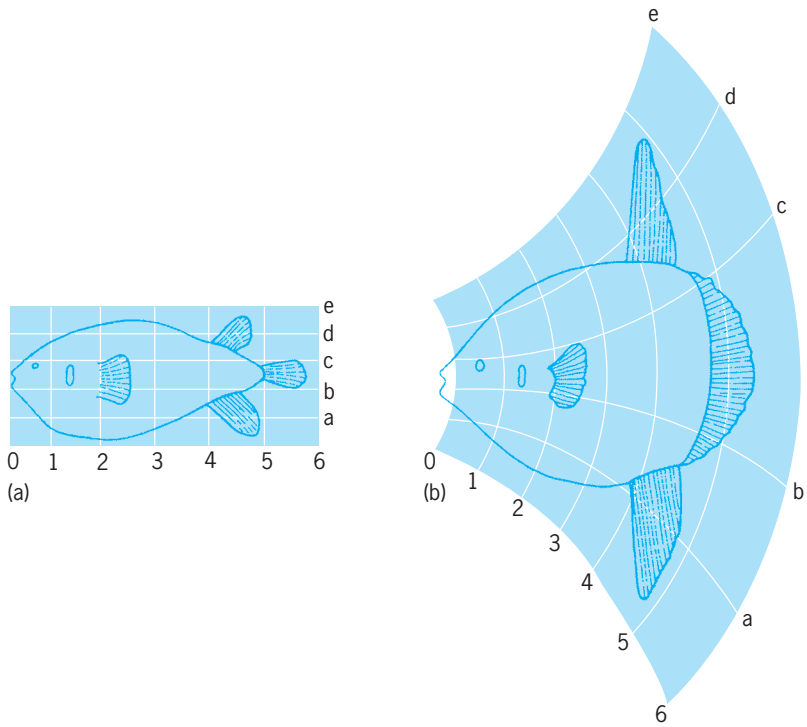

Fig. 1. The fishes (*a*) *Diodon* and (*b*) *Orthagoriscus*, showing the basic relatedness of the two genera by the superimposed grids. (*After D. W. Thompson, On Growth and Form, Cambridge University Press, 1917*)

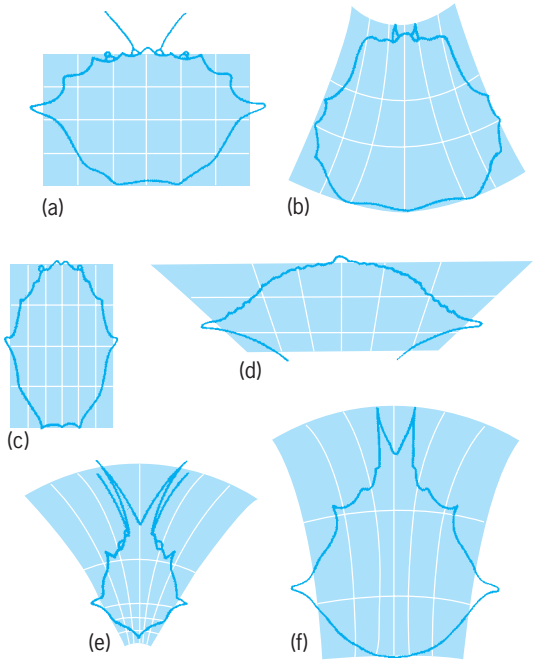

Fig. 2. Crab carapaces from species *b–f* as deformations of *a*. By transforming the grid coordinates, the variation in the shapes of the carapaces is minimized. (*After D. W. Thompson, On Growth and Form, Cambridge University Press, 1917*)

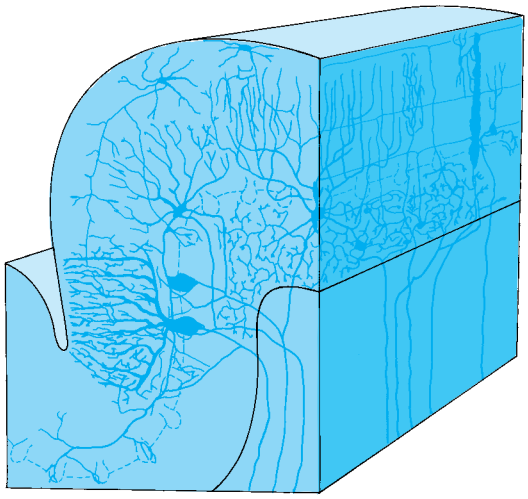

Fig. 3. Section of cerebellum, a "real" network of neurons. (*After K. N. Leibovic, ed., Symposium on Information Processing in the Nervous System, Springer-Verlag, 1969*)

Such metaphorical ideas lead to new ways of relating the behaviors of different biological systems and even of systems that have biological and nonbiological origins. For instance, the boolean neural nets also underlie digital computation, leading to the relationship that is noted between brains and computers. This relationship is explored in the hybrid area of artificial intelligence. The same mathematical formulation of switching networks arises in genetic and developmental phenomena, such as the concept of operon, and in other physiological systems, such as the immune system. *See* ARTIFICIAL INTELLIGENCE; NEURAL NETWORK; OPERON.

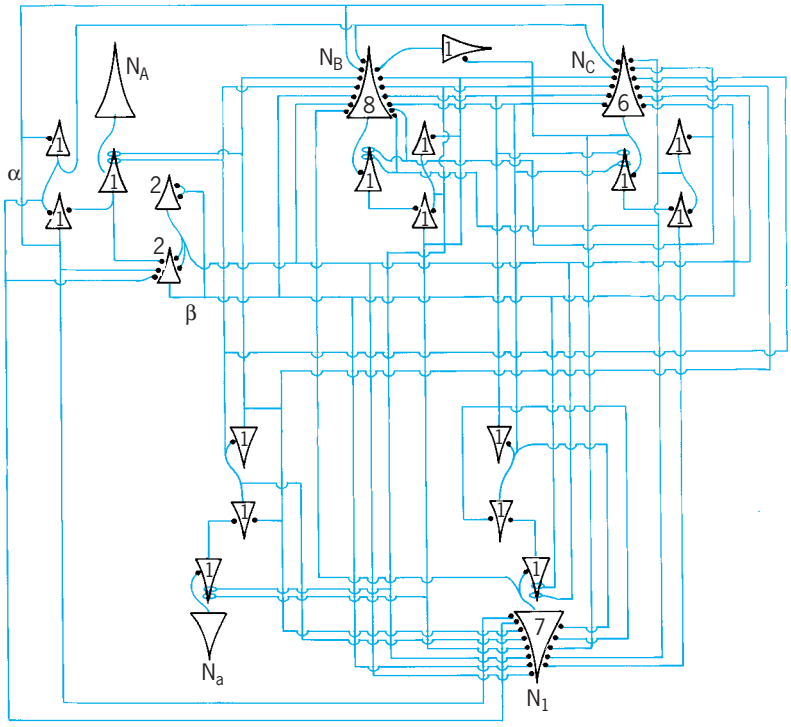

Fig. 4. Neural network that can "reason" syllogistically. Numbers inside neurons (triangles) indicate thresholds, the amount by which excitation must exceed inhibition at one instant in order to fire at the next instant. Triangles without numbers are delays. (*After N. Rashevsky, The neural mechanism of logical thinking, Bull. Math. Biophys., 8:29–40, 1946*)

Another important example of metaphor in biology is morphogenesis, or pattern generation, through the coupling of chemical reactions with physical diffusion. Chemical reactions tend to make systems heterogeneous, diffusion tends to smooth them out, and combining the two can lead to highly complex behaviors. Again, the intention is not to make specific models of specific pattern-generating systems. Rather, since reactions and diffusions typically occur together in biological systems, the idea is to explore the general properties of such systems to illuminate pattern generation in general.

Such ideas turn out to be closely related to those of bifurcation and catastrophe and have a profound impact on physics itself, since they are inherently associated with systems that are thermodynamically open and hence completely outside the realm of classical thermodynamics. The behavior of such open systems can be infinitely more complicated than those that are commonly explored in physics. Open systems may possess large numbers of stable and unstable steady states of various types, as well as more complicated oscillatory steady-state behaviors (limit cycles) and still more general behaviors collectively called chaotic. Changes in initial conditions or in environmental circumstances can result in dramatic switching (bifurcations) between these modes of behavior. *See* CHAOS.

Technological aspects. Biological science underlies a host of basic technologies. Perhaps the biotechnology that has affected everyone most directly is medicine. Medicine can be regarded as a branch of control theory, geared to the maintenance or restoration of a state of health. It is unique in that the systems needed for control are themselves control systems that are far more intricate and complex than any that can be fabricated. In addition to the light it sheds on the processes needed for control, mathematical biology is indispensable for designing the controls themselves and for assessing their costs, benefits, safety, and efficacy.

In general, any theory of control is a matter of design rather than of analysis, and its object is to produce an algorithm, or protocol, that will achieve optimal results. Mathematical biology allows one to relate, and learn about, systems of different characters through the exploitation of their mathematical commonalities. Biology has many optimal designs and optimal controls, which are the products of biological evolution through natural selection. In an exact sense, the design of optimal therapies, for example, is analogous to the generation of optimal organisms, making the concept of surrogacy even more significant. The mathematical theory appropriate for analyzing one discipline of biology, such as evolution, itself becomes transmuted into a theory of control in an entirely different realm.

The same holds true for other biotechnologies. The livelihood of entire nations can rest on

exploitation of biological populations, for example, of fish or forest timber. As with biomedicine, the object is to design strategies for optimally exploiting the resource in question, deriving the maximal economic benefit without irreversibly compromising the resource itself, or inadvertently, the larger system to which the resource belongs and on which it depends. Mathematical biology is called upon to play two roles: it is required to understand the basic biological nature of the resource itself, and to use that understanding in the design of optimal strategies for the harvesting and the maintenance of that resource. *See* MATHEMATICAL ECOLOGY. Robert Rosen

Bibliography. P. L. Antonelli (ed.), *Mathematical Essays on Growth and the Emergence of Form*, 1985; J. A. Metz and O. Diekmann (eds.), *Physiologically Structured Populations*, 1986; L. M. Ricciardi (ed.), *Biomathematics and Related Computational Problems*, 1988.

Mathematical ecology

The application of mathematical theory and technique to ecology. The earliest studies in ecology were by naturalists interested in organisms and their relationships to the environment. Such investigations continue to this day as an active and central part of the subject, and have focused attention on understanding the ecological and evolutionary relationships among species. A large mathematical literature has arisen as an aid to the formalization of concepts and ideas; the underlying goal has been to understand patterns of nature. For the most part, such approaches are retrospective, designed to help in understanding how current ecological relationships developed, and to place that development within appropriate evolutionary context.

The second major branch is applied ecology, and derives from the need to manage the environment and its finite resources. Fisheries, forestry, agriculture, and epidemiology are all highly developed subjects that have arisen from well-defined applied needs. Here the necessity for rigorous mathematical treatments is obvious, but the goals are quite different from those in evolutionary ecology. Management and control are the objectives, and the relevant time horizon lies in the future. The focus is no longer simply to derive understanding and explanation; rather, one seeks methods for prediction and algorithms for control.

That such a dichotomy exists does not mean that pure and applied investigations are disjunct, nor that they are as simply categorized as the above suggests. Yet an appreciation for these opposing tendencies can help to explain the differences among some of the major schools of thought.

Classical theory. The roots of mathematical ecology can be traced to demographic studies as early as the seventeenth century. The well-known arguments of Thomas Malthus and others, that a population growing without bound would soon outstrip the capability of the environment to support it, were based on detailed analyses of births and deaths in human populations. These analyses led to the first important mathematical efforts in ecology, namely those of P. F. Verhulst and others to describe the dependence of population growth rate on population size, and to infer the consequences of such relations. Models of the type developed by Verhulst are called nonlinear, because the graph of growth rate versus population size is not a straight line, but typically bends over due to increased competitive effects (see **illus.**).

The most basic models of growth are iterations, in which population size is recomputed at consecutive times in accordance with a basic underlying growth model. Such a model can be expressed in mathematical form by the difference equation (1), where

$$N(t+h) - N(t) = F[N(t)]h \qquad (1)$$

$N(t)$ is the population size at time t, $N(t+h)$ is the size h units later, and the function F defines the growth rule, as depicted, for example, in the illustration.

Despite the apparent simplicity of such models, their behavior can be exceedingly complex, and they have motivated mathematical investigations even until the present. The time period for which the population growth rate remains fixed determines the characteristics of growth. If the time period h is short, a population governed by such a law typically will come into balance with its environment, equilibrating at its carrying capacity K. However, if the time period is too long, the adjustment of the growth rate is delayed and the population may undergo oscillations, continuing to increase long after

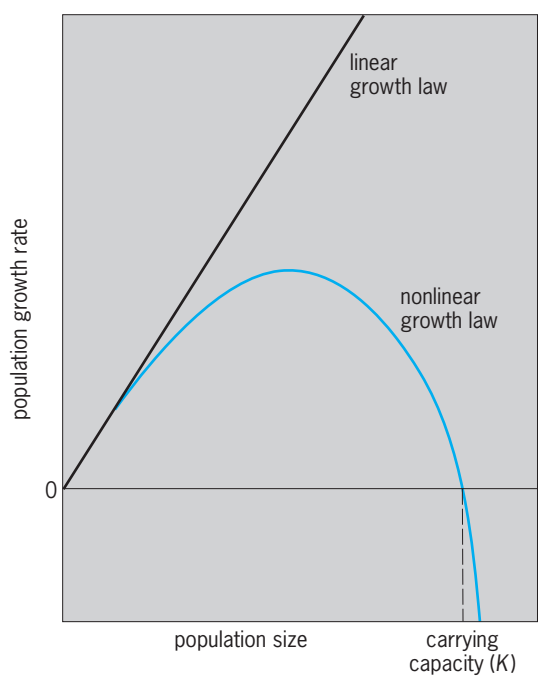

Typical nonlinear curve of population growth rate versus population size, compared with a linear growth law exhibiting the same initial slope (same growth rate at low densities). Note the existence of a carrying capacity, above which population growth rate will be negative.

the resources have been depleted, and then declining until resources recover. Such fluctuations, which may be seen in a variety of mathematical models, mimic the fluctuations observed in many natural populations. When a model based on simple principles gives rise to behavior similar to what is observed in nature, one has an analogy, not a definitive explanation. But such analogies are suggestive, and can be invaluable in helping to develop insights.

When h is small, a powerful approximation to Eq. (1) can be obtained by using differential calculus; in particular, the limiting form is the differential equation (2). If F is continuous, solutions to Eq. (2) never

$$\frac{dN}{dt} = F(N) \tag{2}$$

exhibit oscillations, since they approximate the situation for small h. *See* DIFFERENTIAL EQUATION.

Models of the sort just described are the simplest possible—they ignore all complications due to the fact that a population is composed of a heterogeneous assemblage of individuals, and further that it interacts with other populations and with its abiotic environment. Thus, although these simple models have found application in applied areas such as fishery science, in which one seeks to develop models relating one year's fish population size to the previous year's spawning stock, it is well recognized that more precise answers require consideration of various complications. *See* POPULATION ECOLOGY.

Age-structured models. Perhaps the most obvious potential complicating feature is the spread in ages that may be found within many natural populations. Typically, the production of offspring varies over the lifetime of an individual, as does mortality, the probability that an individual will not survive another year. Such variation is the basis of life tables constructed by life insurance companies to determine premiums; and thus it is not surprising that some of the most important advances in the theory of the dynamics of age-structured populations were made by an actuary, Alfred Lotka, in the early twentieth century. In the approach of Lotka, one describes a population not just by its total size, but also by the number of individuals of each age, summarized mathematically by a vector. This vector is a one-dimensional array [notation (3)], in which n_i denotes the number of

$$n = (n_1, n_2, \ldots, n_A) \tag{3}$$

individuals in age class i, for each i from 1 through the maximum age A.

The description of the dynamics of an age-structured population involves age-specific information concerning both fecundity and mortality. This information may also be expressed in a two-dimensional array, known as the Leslie matrix, which summarizes the contribution, through births or advancement in age, of each age class to each other one. One of the most fundamental results in mathematical ecology is that when the fecundity and mortality schedules for a population remain constant, the population will in general tend to a stable age distribution. In this distribution the relative proportions of the various age classes remain fixed while the population grows exponentially at a rate known as the intrinsic rate of natural increase. The only exception to this rule occurs in the unlikely situation that births are pulsed, taking place only at exact periodic (multiyear) intervals in the life of the organism.

Of course, the discussion above suffers from the same weakness as does discussion of exponential growth: In nature, birth and death parameters will not be constant, but will vary as population size changes. In such situations, the classical results do not apply, and the population may either attain balance or oscillate. As for the earlier case (without age structure), a continuous approximation may be derived, using the calculus of several variables; but unlike the previous case, the solutions to the limiting partial differential equation may sustain oscillations.

Competition and predation models. The most important extensions of these single-species models in the classical literature are to systems of interacting populations, especially the famous differential equations derived independently by Lotka and by the mathematician Vito Volterra. In this description, if X_1 and X_2 denote two interacting populations, the growth relationships are expressed by Eqs. (4), in which r_i denotes the intrinsic rate of natural increase of species i, and K_i denotes its carrying capacity. Further, α_{ij} is a measure of the influence of species j on species i.

$$\begin{aligned} \frac{dX_1}{dt} &= r_1 X_1 \left(1 - \frac{X_1}{K_1} - \alpha_{12}\frac{X_2}{K_1}\right) \\ \frac{dX_2}{dt} &= r_2 X_2 \left(1 - \frac{X_2}{K_2} - \alpha_{21}\frac{X_1}{K_2}\right) \end{aligned} \tag{4}$$

The two most important special cases of Eqs. (4) are competition (in which α_{12} and α_{21} are both positive) and predation (in which α_{12} is positive, but α_{21} is negative to reflect the positive effect of prey density, X_1, on predator growth rates).

The Lotka-Volterra equations have been analyzed in remarkable detail, and continue to fascinate mathematicians. In the case of competition, attention has focused on the conditions under which species can coexist; and when these conditions are violated, Eqs. (4) form the starting point for the theory of competitive exclusion and the closely related theory of the ecological niche. In the case of predation, the most powerful insight came from Volterra's original study, which was designed to show the similarity between the inherently oscillatory character of the solutions to Eqs. (4) and the fluctuations observed in the fisheries of the Adriatic. Later versions have embellished these results, and other models have been developed which demonstrate more stable oscillations, or which extend Eqs. (4) to systems of many interacting populations (this was actually done by the early masters), but no later work has come close to matching the influence of the seminal investigations of Lotka and Volterra.

Modern trends. The classical tradition has been carried into the present, especially regarding the development of an elegant theory of interspecific interactions and evolutionary relationships. Research has extended the theory to an examination of the structure of ecological food webs, and the factors governing their organization, and has taken the subject into diverse mathematical disciplines such as graph theory and dynamical systems theory.

Increasingly, however, mathematical approaches have been recognized as being valuable not only for theoretical investigations but also for the applied problems that society must confront. For example, mathematical descriptions of population dispersal, which are among the oldest models in theoretical ecology, are being used to quantify the movements of agricultural pest species and the rates of advance of exotic species invading new habitats. Also, the theory of island biogeography and more extensive mathematical models of spatially distributed populations are finding application in the design of reserves, in national park planning, and in regional landscape ecology. *See* ISLAND BIOGEOGRAPHY; POPULATION DISPERSAL.

Optimization and control theory are being applied as integral parts of programs for the management of renewable resources, especially fisheries and agricultural production. Such approaches wed biology and economics through mathematical models, and this marriage must become an increasing imperative for society as it faces shortages in energy and depletion of its limited resources. Inherent limitations to predictability are apparent in any rigorous mathematical analysis, and have made essential the development of adaptive management strategies, which couple short-term prediction with continually adjusted management rules. *See* OPTIMIZATION; QUALITY CONTROL.

Epidemiology is an area which has had a firm mathematical basis since the turn of the century. The impact of these models has been limited, although the problems society faces in combating the spread of diseases of humans and other animal and plant species are of overwhelming importance. There has been a dramatic increase in mathematical activity concerning the modeling and control of epidemics, and an increasing recognition of the need to view such problems in their proper ecological context as host-parasite interactions. Very likely this will be one of the most important growth areas in mathematical ecology over the next quarter century. Researchers are using mathematical models to help to understand the factors underlying disease outbreaks, and to develop methods for control such as vaccination strategies. *See* EPIDEMIOLOGY.

Finally, the need for environmental protection in the face of threats from such competing stresses as toxic substances, acid precipitation, and power generation has led to the development of more sophisticated models that address the responses to stress of community and ecosystem characteristics, for example, succession, productivity, and nutrient cycling. Such models, while owing much to their classical origin, typically differ substantially in form, and recognize the importance of explicitly considering environmental factors, physical characteristics of the environment, and nonbiotic system components, and of focusing attention on holistic measures of system response. Examples include large-scale models being developed to examine the likely effects of air pollution on regional patterns of forest productivity, or of power plants upon estuarine communities. *See* ENVIRONMENTAL ENGINEERING.

A major difference between such models and those discussed earlier is that they are in general too large to be fully analyzed, and one must rely heavily on the computer to simulate possible outcomes. However, in such applications, mathematical analyses are as essential as ever, since one must find ways to simplify, to guide simulations, and to derive understanding. Applications will increasingly be needed and will introduce new and vital challenges for mathematical ecology. *See* ECOLOGY; ECOLOGY, APPLIED; SIMULATION.

Simon A. Levin

Bibliography. Y. Cohn (ed.), *Application of Control Theory in Ecology: Lecture Notes in Biomathematics*, 1987; T. G. Hallam and S. A. Levin (eds.), *Mathematical Ecology*, 1986; C. J. Krebs, *Ecological Methodology*, 1988; L. Legendre and P. Legendre (eds.), *Developments in Numerical Ecology*, 1987; A. S. Podolsky, *New Phenology: Elements of Mathematical Forecasting in Ecology*, 1984.

Mathematical geography

The branch of geography that examines human and physical activities on the Earth's surface using models and statistical analysis. The primary areas in which mathematical methods are used include the analysis of spatial patterns, the processes that are responsible for creating and modifying these patterns, and the interactions among spatially separated entities.

What sets geographic methods apart from other quantitative disciplines is geography's focus on place and relative location. Latitude and longitude provide an absolute system of recording spatial data, but geographic databases also typically contain large amounts of relative and relational data about places. Thus, conventional quantitative methods are often insufficient for spatial problems, and geographers have devoted much effort to accounting for spatial interrelations while maintaining consistency with the assumptions of mathematical models and statistical theory. *See* GEOGRAPHY; MODEL THEORY; STATISTICS.

Quantitative methodology has long been part of physical geography due to its ties to geology and other physical sciences, but human geographers were slower in adopting and adapting quantitative methods. In the 1950s a growing dissatisfaction with traditional regional and descriptive approaches to geography, as well as an absence of theory, led geographers to incorporate statistical and mathematical methods and other theoretical models into geographic research. This methodological and philosophical shift is now known as the "Quantitative

Revolution," and quantitative methods remain a cornerstone of geographic research. *See* PHYSICAL GEOGRAPHY.

Spatial pattern. Spatial pattern methodologies attempt to describe the arrangement of phenomena over space. In most cases these phenomena are either point or area features, though computers now allow for advanced three-dimensional modeling as well. Point and area analyses use randomness (or lack of pattern) as a dividing point between two opposite pattern types—dispersed or clustered.

An important innovation in geographic modeling has been the development of spatial autocorrelation techniques. Unlike conventional statistics, in which many tests assume that observations are independent and unrelated, very little spatial data can truly be considered independent. Soil moisture or acidity in one location, for example, is a function of many factors, including the moisture or acidity of nearby points. Because most physical and human phenomena exhibit some form of spatial interrelationships, several statistical methods have been developed to measure this spatial autocorrelation. Once identified, the presence and extent of spatial autocorrelation can be built into the specification of geographical models to more accurately reflect the behavior of spatial phenomena.

Related to the search for spatial pattern is the concept of spatial hierarchy. With respect to cities, Zipf's law or rank-size rule presumes a direct relationship between a city's size and its relative national population rank. The standard form of this rule predicts that the second largest city in a nation should have half the population of the largest city, the third largest city one-third the population of the largest city, and so on. Studies have shown a clear relationship between a nation's conformity to this rule and its development level, with more developed nations more closely adhering to the rank-size rule.

Central place theory is a more comprehensive model of urban structure, focusing on the location and spacing of cities. This theory identifies three primary guiding principles for the arrangement of cities in an urban hierarchy, depending on whether retail, transportation, or administrative efficiency is being optimized. Precise mathematical relationships between the orders of the hierarchy indicate the expected relative market sizes, spacing, and numbers of cities at different levels of the hierarchy.

Analysis of hub-and-spoke networks has become an important field in spatial analysis, due to the proliferation of passenger and parcel transportation via air. Relative location is more important than site-specific characteristics of places, particularly because of the nature of air travel. Hence, a node's interconnections and links to other nodes determine its place in the hierarchy and the overall efficiency of the system.

Spatial processes. Analysis of spatial pattern is motivated by the desire to reveal the processes that create patterns. When central place theory reveals a retail, transportation, or administrative pattern in a landscape, for example, further analysis can uncover the behavioral forces that created such patterns, opening the way to a better understanding of human behavior or of physical properties.

A spatial process that underlies many spatial models is distance decay. Distance decay assumes that interactions diminish with distance, such that entities (cities, individuals, and so on) that are farther apart will interact less than entities in closer proximity. With the electronic and communications advances of the 1990s, though, the impact of distance is shrinking. New avenues of research in geography are exploring the impact of the Internet society on distance decay and on spatial interaction.

Spatial interaction. Perhaps the most frequent application of mathematical methods in geography has been to study the flows of goods, people, and ideas. Early in the Quantitative Revolution, social scientists looked to the physical sciences for analog models that could be adapted to human behavior, giving rise to the term "social physics." The most notable form of social physics has been the use of Newton's gravity model, in which the attraction (or interaction) between two places and the sizes of the places are directly proportional. Additionally, the distance decay effect plays a prominent role as more distant locations are expected to interact less than closer places. Travel time or cost can substitute for simple distance in this model, while some type of population statistic is the most common measure of the size of places.

While a simple gravity model allows for only two places to be modeled simultaneously, spatial interaction models represent an advancement by allowing multiple origins and destinations, multiple attributes of these origins and destinations, and varying effects of distance all to be incorporated into a model of flows or interactions. These models have been heavily used in migration, journey-to-work, and interregional flow studies.

Further integration of mathematics into spatial analysis includes fuzzy logic and chaos theory, better cohesion in spatial statistical theory, behavioral models of individual (as opposed to aggregate) choice, and decision support systems. *See* CHAOS; DECISION SUPPORT SYSTEM. Jonathan C. Comer

Bibliography. B. Boots and A. Getis, *Spatial Point Pattern Analysis*, Sage Publications, Beverly Hills, CA, 1988; A. S. Fotheringham and M. E. O'Kelly, *Spatial Interaction Models: Formulations and Applications*, Kluwer Academic, Dodrecht, Netherlands, 1989; D. A. Griffith and C. G. Amrhein, *Statistical Analysis for Geographers*, Prentice Hall, Englewood Cliffs, NJ, 1991; J. Odland, *Spatial Autocorrelation*, Sage Publications, Newbury Park, CA, 1988; J. Odland, R. G. Golledge, and P. A. Rogerson, Mathematical and statistical analysis in human geography, in G. L. Gaile and C. J. Wilmott (eds.), *Geography in America*, Merrill Publishing, Columbus, OH, 1989.

Mathematical physics

An area of science concerned with the application of mathematical concepts to the physical sciences and the development of mathematical ideas in response

to the needs of physics. Historically, the concept of mathematical physics was synonymous with that of theoretical physics. In present-day terminology, however, a distinction is made between the two. Whereas most of theoretical physics uses a large amount of mathematics as a tool and as a language, Mathematical physics places greater emphasis on mathematical rigor, and devotes attention to the development of areas of mathematics that are, or show promise to be, useful to physics. Henri Poincaré likened physics to a large library in which the books are furnished by the experimenters, and the index, without which the collection would be inaccessible and useless, by the mathematical physicists (in the older sense of the word). If that simile is accepted, there remains the remarkable fact that most of the organization and structure of the library, and even the existence of many of the books, is determined by the nature of the index, since both the language and the content of mathematics continue to have a powerful influence on ideas of the structure of nature. The results obtained by pure mathematicians, with no thought to applications, are almost always found to be both useful and effective in formulating physical theories.

Mathematical physics forms the bridge between physics as the description of nature and its structure on the one hand, and mathematics as a construction of pure logical thought on the other. This bridge between the two disciplines benefits and strengthens both fields enormously. *See* MATHEMATICS; PHYSICS; THEORETICAL PHYSICS.

Historical examples. Isaac Newton invented the differential and integral calculus for the specific purpose of physical application, in the formulation of his equations of motion and for the law of gravitation as applied to extended bodies. In the course of the next two centuries, the calculus grew into the large branch of analysis in mathematics. Within that branch, the development of the science of differential equations was further propelled by the need for solving the equations of motion of objects in mechanics, and specifically the development of partial differential equations by their use in Maxwell's theory of electromagnetism, in continuum mechanics, and in the theory of sound. *See* CALCULUS; CLASSICAL FIELD THEORY; DIFFERENTIAL EQUATION; GRAVITATION; MAXWELL'S EQUATIONS; NEWTON'S LAWS OF MOTION; SOUND.

When the true complication of the newtonian mechanics of the solar system, in which each planet was gravitationally attracted not only by the Sun but also by all the others, became apparent, the urgent problem arose to ascertain mathematically whether this system was stable or whether it might be destined to collapse. The stability of the system was proved by J. L. Lagrange, whose method of proof engendered a new, fruitful branch of mathematics, perturbation theory.

Modern examples. W. Heisenberg's formulation of matrix mechanics included strange noncommuting entities (later called q-numbers by P. A. M. Dirac), which led to his principle of indeterminacy, thus beginning the large-scale introduction of the use of operators on Hilbert space in quantum mechanics. Although D. Hilbert had developed the theory of these infinite-dimensional function spaces, this theory was entirely unfamiliar to physicists at the time. Dirac formulated all of quantum mechanics in the language of that theory, and J. von Neumann put that formulation on a firm mathematical foundation. Today, most of the work of mathematical physicists is concerned with problems in either quantum mechanics or quantum field theory. *See* HILBERT SPACE; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Stability of matter. A problem analogous to that of the stability of the solar system arose in the quantum mechanics of bulk matter: whether a solid (a system of a large number of atoms, consisting of many electrically charged particles, which attract or repel one another, as described by the equations and postulates of quantum mechanics) is stable or will collapse. Since matter is manifestly stable, the problem constituted a mathematically difficult test for quantum mechanics. The stability proof was given in the 1960s. *See* QUANTUM THEORY OF MATTER.

Phase transitions. Numerous other modern problems of mathematical physics also involve many-particle systems. One of these is the question of the existence of a critical temperature at which an abrupt phase transition takes place. Such a phase transition may, as the temperature is lowered or raised, change a liquid into a gas or a solid, cause a ferromagnet to lose its ferromagnetism, change an insulator into a superconductor, or cause any number of other sudden changes of properties. These problems are particularly difficult because such a discontinuous transformation can never occur so long as the number of particles is finite.

As in some other problems of mathematical physics, in the case of ferromagnetism it was the introduction of a very simplified but mathematically nontrivial model that led to a clearer understanding: the Ising model. It consists of a regular, infinite, two-dimensional lattice of elementary magnets that can point only up or down, and that interact with only their nearest neighbors, so as to favor alignment. The calculation of the partition function of statistical mechanics for this model led to a proof that there is, indeed, a Curie temperature, below which there is long-range order and above which that order disappears. However, there is still no proof that even this vastly simplified model has a phase transition when extended to three dimensions. The existence of the most common phase transitions, freezing and boiling of liquids, is an unsolved problem of mathematical physics and so is the problem of why most solid matter is in crystalline form. *See* ISING MODEL; PHASE TRANSITIONS; STATISTICAL MECHANICS.

Scattering theory. Mathematical physicists have also been very active in the development of the quantum theory of scattering, which deals with the scattering of particles under the influence of forces. The mathematical investigation of the analytic properties of wave functions, of the scattering amplitude,

and of the scattering matrix has led to important experimental predictions, such as dispersion relations. The inverse scattering problem is the mathematical question of whether it is possible to infer the interparticle forces from measurements of the scattering, and if so, whether such an inference is unique and how it can be implemented. *See* DISPERSION RELATIONS; INVERSE SCATTERING THEORY; SCATTERING MATRIX.

Quantum field theory. Theories of fundamental particles are all based on quantum field theory, which is fertile ground for problems in mathematical physics. Here the problems usually deal with the question of whether a given theory, formulated in mathematical language, makes mathematical sense. In particular, mathematical physicists investigate whether the postulated equations have solutions, what the properties of these solutions are, and how the solutions can be constructed for actual numerical calculations that can then be compared with experimental data. Sometimes, as in the case of string field theory, such investigations lead to deep new mathematical developments or even to the construction of entirely new areas of mathematics. (In this particular case, the theory of knots received an important stimulus and a new set of ideas.) To find a consistent theory that combines the quantum field theory of elementary particles with the force of gravity is one of the greatest challenges faced by mathematical physics. Entirely novel mathematical concepts have been proposed (such as noncommutative geometry) for this purpose. *See* QUANTUM FIELD THEORY; QUANTUM GRAVITATION; RELATIVITY; SUPERSTRING THEORY.

Quantum electrodynamics. In the instance of quantum electrodynamics (QED), the quantum theory of the interaction of the electromagnetic field and electrons, there is an unsolved puzzle: On the one hand, the theory has produced unprecedented precision in its agreement between calculations that are based on well-defined approximation schemes for solving the equations of QED and experimental results; on the other hand, all attempts at proving that the equations, in fact, have solutions have failed. If a proof should be forthcoming that the equations of QED have no solutions (and there are some indications that this may be the case), it will be the task of mathematical physics to lead to an understanding of what it is that has been calculated and that agrees with nature so remarkably well, if it is not a solution. In other words, the concept of what constitutes a solution of such equations will have to be appropriately altered. *See* QUANTUM ELECTRODYNAMICS.

Computer applications. The use of large-scale computers makes it possible to solve equations that were previously intractable. Even though the computer can rarely provide definitive proofs, it can often provide very fruitful clues to new effects and stimulate ideas. *See* DIGITAL COMPUTER; NUMERICAL ANALYSIS.

Nonlinear differential equations. The study of nonlinear differential equations is a case in point. The equations of classical mechanics are almost always nonlinear and therefore, in practice, unmanageable, with only some very special solutions either numerically calculable by hand or qualitatively comprehensible. In fact, most systems of classical particles evolve into chaos. However, the fact that Maxwell's equations of electromagnetism and those of quantum mechanics are linear led mathematical physicists for almost a century to pay almost exclusive attention to linear phenomena. The introduction of the computer changed this situation, and the study of nonlinear dynamical equations, with the quasi- universal development of chaos, became numerically feasible and consequently a fertile field of research for mathematical physicists. However, turbulence is a highly nonlinear phenomenon of great practical importance that has, so far, resisted the efforts of mathematical physicists. Yet, the borderline between those few solutions of a general nonlinear dynamical system that are said to be regular and those that are chaotic has been shown by the KAM theorem (due to A. N. Kolmogorov, V. I. Arnold, and J. Moser) to be surprisingly fuzzy: when parameters are changed from values for which the system is understood to be quasiperiodic and regular, however it starts, the system does not immediately become chaotic but retains its predictable character for a finite range of parameter values until it finally succumbs to chaos for all initial conditions. *See* CHAOS; NONLINEAR PHYSICS; TURBULENT FLOW.

Solar-system stability. One result of the computer development is that the question of the stability of the solar system has been reexamined without the use of linear methods. It is now feasible to check the stability numerically for several million years. This numerical work has shown that the stability of the orientation of the Earth's axis is assured only because of the presence of the moon.

Solitons. Another example of the influence of the computer upon the development of mathematical physics is the investigation of the Korteweg-De Vries equation, a nonlinear partial differential equation that was devised in the late nineteenth century to describe the motion of shallow water waves in a one-dimensional channel. Computer calculations in the 1960s led to the discovery of solitons, solutions that describe solitary bumps colliding with one another, causing, during the collision, distortions of their shape, but that, after the collision, emerge unscathed, with no change in shape, speed, or amplitude. This quite unexpected effect has subsequently been found for many other nonlinear equations, and its mathematical analysis (rather than numerical study) has given rise to a large field of applied mathematics as well as to an explanation of a number of natural phenomena. *See* SOLITON.

Methods. The methods employed in mathematical physics range over most of mathematics, the areas of analysis and algebra being the most commonly used. Partial differential equations and differential geometry, with heavy use of vector and tensor methods, are of particular importance in the formulation of field theories, and functional analysis as well as operator theory in quantum mechanics. Group theory has become an especially valuable tool in the construction of quantum field theories

and in elementary-particle physics. There has also been an increase in the use of general geometrical approaches and of topology. For solution methods and the calculation of quantities that are amenable to experimental tests, of particular prominence are Fourier analysis, complex analysis, variational methods, the theory of integral equations, and perturbation theory. *See* ABSTRACT ALGEBRA; COMPLEX NUMBERS AND COMPLEX VARIABLES; DIFFERENTIAL GEOMETRY; FOURIER SERIES AND TRANSFORMS; GROUP THEORY; INTEGRAL EQUATION; OPERATOR THEORY; TOPOLOGY; VARIATIONAL METHODS (PHYSICS); VECTOR METHODS (PHYSICS). Roger G. Newton

Bibliography. E. Butkov, *Mathematical Physics*, 1968; R. Geroch, *Mathematical Physics*, 1985; E. H. Lieb and D. C. Mattis, *Mathematical Physics in One Dimension; Exactly Solvable Models of Interacting Particles*, 1966; M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, vol. 1, 2d ed., 1980, vol. 2, 1975, vol. 3, 1979, vol. 4, 1978; R. D. Richtmyer, *Principles of Advanced Mathematical Physics*, 2 vols., 1979, 1981.

Mathematical software

The collection of computer programs that can solve equations or perform mathematical manipulations. The developing of mathematical equations that describe a process is called mathematical modeling. Once these equations are developed, they must be solved, and the solutions to the equations are then analyzed to determine what information they give about the process. Many discoveries have been made by studying how to solve the equations that model a process and by studying the solutions that are obtained.

Before computers, these mathematical equations were usually solved by mathematical manipulation. Frequently, new mathematical techniques had to be discovered in order to solve the equations. In other cases, only the properties of the solutions could be determined. In those cases where solutions could not be obtained, the solutions had to be approximated by using numerical calculations involving only addition, subtraction, multiplication, and division. These methods are called numerical algorithms. These algorithms are often straightforward, but they are usually tedious and require a large number of calculations, usually too many for a human to perform. There are also many cases where there are too many equations to write down. *See* ALGORITHM; NUMERICAL ANALYSIS.

The advent of computers and high-level computer languages has allowed many of the tedious calculations to be performed by a machine. In the cases where there are too many equations, computer programs have been written to manipulate the equations. A numerical algorithm carried out by a computer program can then be applied to these equations to approximate their solutions. Mathematical software is usually divided into two categories: the numerical computation environment and the symbolic computation environment. However, many software packages exist that can perform both numerical and symbolic computation.

Characteristics. Mathematical software that does numerical computations must be accurate, fast, and robust. Accuracy depends on both the algorithm and the machine on which the software is run. Most computers follow the standards of the American National Standards Institute (ANSI) and the Institute of Electrical and Electronics Engineers (IEEE). Most mathematical software uses the most advanced numerical algorithms. Robustness means that the software checks to make sure that the user is inputting reasonable data, and provides information during the performance of the algorithm on the convergence of the calculated numbers to an answer. Mathematical software packages can approximate solutions to a large range of problems in mathematics, including matrix equations, nonlinear equations, ordinary and partial differential equations, integration, and optimization. Mathematical software libraries contain large collections of subroutines that can solve problems in a wide range of mathematics. These subroutines can easily be incorporated into larger programs.

Software packages. For many years, these subroutines were available only in Fortran. Almost all of the Fortran programs are now available in C. IMSL® is a mathematical software package that contains numerous numerical algorithms for solving linear and nonlinear mathematical problems. The IMSL® programs are written in both Fortran and C and are available for downloading. The Numerical Algorithms Group (NAG®) is an organization that puts out a mathematical software package of algorithms in Fortran and C. Two popular texts, *Numerical Recipes in Fortran* and *Numerical Recipes in C*, also contain numerous algorithms. The company The MathWorks has written the software package MATLAB® (Matrix Laboratory). MATLAB® was originally written in Fortran but now is in C. Its numerical software environment contains most of the algorithms in IMSL®, NAG®, and the two *Numerical Recipes* texts. The user of MATLAB® can develop software in MATLAB®. A programmer who has to use a numerical algorithm can save time and effort by using any of these products. *See* PROGRAMMING LANGUAGES.

Symbolic computation. Early computers were used mainly to perform numerical calculations, while the mathematical symbolic manipulations were still done by humans. Now software such as Derive®, Maple®, and Mathematica® is available to perform these mathematical manipulations. Most of the mathematical software packages that perform symbolic manipulations can also perform numerical calculations. Software can be written in the package to perform the numerical calculations, or the calculations can be performed after the symbolic manipulations by putting numbers into the symbolic formulas. Mathematical software that is written to solve a specific problem using a numerical algorithm is usually computationally more efficient than these software environments. However, these software environments can perform almost all the commonly

used numerical and symbolic mathematical manipulations. *See* SYMBOLIC COMPUTING.

Parallel computation. Since most of the problems of interest to society are modeled by complicated mathematical equations, they require a large number of calculations and storage of their results. This has led to industries of people working on building larger and faster computers. Different types of computers have also been developed. Parallel computers have more than one processor that can work on the same problem at the same time. Parallel computing allows a large problem to be distributed over the processors. This allows the problem to be solved in a smaller period of time. Many numerical algorithms have been converted to run on parallel computers. *See* MULTIPROCESSING.

A set of personal computers can be connected into an array that can perform parallel computations. Oak Ridge National Laboratory has developed a software package, Parallel Virtual Machine (PVM), which allows the user to divide a problem among several connected personal computers. The National Science Foundation spearheaded a worldwide operation that united computer scientists and mathematicians from universities and industries to develop a parallel software environment, Message Passing Interface (MPI), which is easy to use and incorporates the best properties of such an environment. Both PVM and MPI can be used with Fortran or C, and both provide users with the power of a supercomputer at a fraction of its cost. Fortran has also been upgraded to a parallel language called High-Performance Fortran (HPF). The National Aeronautics and Space Administration (NASA) has developed a parallel-personal-computer hardware-software environment called Beowulf. The Beowulf system is a configuration of personal computers with a communication environment between them. It can use HPF, PVM, or MPI. *See* COMPUTER PROGRAMMING; CONCURRENT PROCESSING; DIGITAL COMPUTER; DISTRIBUTED SYSTEMS (COMPUTERS); SOFTWARE. James Sochacki

Bibliography. W. J. Kaufmann and L. L. Smarr, *Supercomputing and the Transformation of Science* (*Scientific American Library*, no. 43), W. H. Freeman, 1993; W. H. Press et al., *Numerical Recipes in C: The Art of Scientific Computing*, 2d ed., Cambridge University Press, 1992; W. H. Press et al., *Numerical Recipes in Fortran 90: The Art of Parallel Scientific Computing* (*Fortran Numerical Recipes,* vol. 2), Cambridge University Press, 1996.

Mathematics

Mathematics is frequently encountered in association and interaction with astronomy, physics, and other branches of natural science, and it also has deep-rooted affinities to the humanities. Actually, it is a realm of knowledge entirely unto itself, and one of considerable scope; the word mathematics stems from a root which means learnable knowledge. Mathematical knowledge is commonly deemed to have a high degree of validity, irrespective of culture conditioning and predilection, although it can be argued that, in the past, cultural settings have affected its development noticeably.

Relation to science. As far as the scientist is concerned, mathematics is not a branch of natural science itself. It does not deal with phenomena and objects of the external world and their relations to each other but, strictly speaking, only with objects and relations of its own imagery. Mathematical figures in two- or three-dimensional geometry are largely idealizations of objects occurring in the physical world, but figures in n-dimensional space for general n no longer are such idealization. Integer numbers 1, 2, 3, . . . , and even real numbers in general, can be claimed to be abstractions from quantities occurring in the physical world, but the "imaginary" number $i = \sqrt{-1}$ has received its name from the very fact that it has no basis in the material world, even though the use of complex numbers $a + bi$ is indispensable to science. If a laboratory experiment idealizes a physical system in order to eliminate secondary features not essential to the study at hand, its ultimate objective is an understanding of the unidealized physical system. But even if it were true that the five regular solids as investigated in the thirteenth book of Euclid's *Elements* (cube, tetrahedron, octahedron, icosahedron, and dodecahedron) had been found because of the occurrence of approximating crystals in nature, nevertheless, once found, the idealized geometric figures would still have become primary objects of study in mathematics.

Mathematics is not subordinate to natural science by being a handmaiden of it, and one can practice competently meaningful mathematics without being concerned with science at all. Especially, philosophical attempts to reduce all origin of mathematics to utilitarian motives are wholly unconvincing. However, it is fair to say that mathematics is the language of science in a deep sense. Mathematics is an indispensable medium by which and within which science expresses, formulates, continues, and communicates itself. And just as the language of true literacy not only specifies and expresses thoughts and processes of thinking but also creates them in turn, so does mathematics not only specify, clarify, and make rigorously workable concepts and laws of science, but also at certain crucial instances becomes an indispensable constituent of their creation and emergence as well. In Isaac Newton's formula for the motion of a particle on a straight line, Eq. (1), the mass m and the force F are nonmathematical objects, perhaps. But the instantaneous velocity $v = dx/dt$ and the instantaneous acceleration $a = dv/dt = d^2x/dt^2$ are wholly mathematical, and without a mathematical theory of the infinitesimal calculus, they are not conceivable. I. Newton, the physicist, was driven to creating his version of the calculus because of this. Also, Newton had to have not only the process of differentiation but also

$$m\frac{d^2x}{dt^2} = F \tag{1}$$

the concept of a mathematical function, because only a function can be differentiated. He required not only the path function $x = x(t)$, but, for the second derivative, he had to envisage the velocity $v = dx/dt$ itself, again as a function depending on t, even though by the definition of v this dependence had been reduced to "instantaneousness." This concept of function was given to Newton by the then new theory of analytic geometry of René Descartes, and it is a fact that after Archimedes' work on the motion of a lever, theoretical mechanics stood virtually still for almost 2000 years until the twin mathematical concepts of function and derivative were ready to emerge. *See* CALCULUS.

A striking case of the manner in which mathematics may shape physics has occurred in the theory of relativity. When A. Einstein was pondering the transition from special to general relativity, his attention was drawn to G. Ricci's work on tensor calculus. Einstein found it an excellent medium for his thoughts, adopted it, and made it widely known to mathematicians. In the 1920s the quantum theory could move as rapidly as it did only because certain prefabricated parts from the theory of matrices and differential equations were lying about somewhat idly. *See* DIFFERENTIAL EQUATION; MATRIX THEORY.

Creative formulas. A formula is a string of mathematical symbols subject only to certain general rules of composition. To a working mathematician a string of symbols is a formula if it is something worth remembering. Much mathematics is concentrated in and propelled by certain formulas of unusual import.

The oldest of these is the pythagorean theorem $c^2 = a^2 + b^2$, where a, b are the sides and c is the hypotenuse in a right-angled triangle. If a, b are rational numbers p/q, then c need not be so; for instance, when $a = b = 1$, $c = \sqrt{2}$ and is not a rational number. This compelled the Greeks to study ratios of incommensurable quantities and also quadratic irrationalities. In modern times the formula was extended to $c^2 = a_1{}^2 + \cdots + a_n{}^2$ for the diagonal of a rectangular parallelepiped in n-dimensional euclidean space, and gave Eq. (2) for the distance between two points

$$s^2 = (x^1 - y^1)^2 + \cdots + (x^n - y^n)^2 \qquad (2)$$

$x = (x^1, \ldots, x^n)$, $y = (y^1, \ldots, y^n)$ in cartesian coordinates there. For infinitesimally near points this becomes the line element $ds^2 = (dx^1)^2 + \cdots + (dx^n)^2$, which in the more general affine version, Eq. (3*a*), is the cornerstone of riemannian geometry. [Equation (3*b*) is discussed below.] On the other

$$ds^2 = \sum_{i,j=1}^{n} g_{ij}\, dx^i\, dx^j \qquad (3a)$$

$$ds^2 = \sum_{\alpha,\beta=1}^{n} g_{\alpha\bar{\beta}}\, dz^\alpha\, dz^{\bar{\beta}} \qquad (3b)$$

hand, analytic developments demanded the extension of Eq. (2) to a space of infinitely many dimensions and then even to the case in which coordinates are complex numbers; thus Eq. (4) may be written.

$$s^2 = \left|z^1 - w^1\right|^2 + \left|z^2 - w^2\right|^2 + \cdots \qquad (4)$$

With this definition of distance the space becomes a Hilbert space, the basis of operator theory and quantum mechanics. Finally, again for finite dimension but still complex coordinates, it became significant, in analogy to Eq. (3*a*), to introduce the line element in Eq. (3*b*). It is the basis for the study of riemannian geometry on complex manifolds, and the study of such manifolds is penetrating even relativity and field theory. *See* OPERATOR THEORY; QUANTUM MECHANICS; RIEMANNIAN GEOMETRY.

The simple relation $(1 + x)^2 = 1 + 2x + x^2$ was generalized by Newton to Eq. (5). The coefficients are the binomial coefficients, shown in Eq. (6), and

$$(1 + x)^n = 1 + \binom{n}{1} x + \binom{n}{2} x^2 + \cdots \qquad (5)$$

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{1 \cdot 2 \cdots k} \qquad (6)$$

the formula is a generating relation for them. The expression in Eq. (6) can also be set up if n is any nonintegral number, real or complex, and it was indeed established that Eq. (5) holds for any n for $|x| < 1$. The expansion is usually an infinite series, and it is in fact the Taylor series of $(1 + x)^n$. By studying, for general n, the behavior of the series as $|x| \to 1$, H. Abel laid the foundation for the theories of summability of nonconvergent series. *See* SERIES.

Euclid's theorem that any integer is a unique product of primes (stated in Euclid only implicitly) can be expressed as a formula, Eq. (7). The concept of a

$$a = p_1{}^{n_1} \times p_2{}^{n_2} \times \cdots \times p_k{}^{n_k} \qquad (7)$$

prime factor and the quest for a theorem on unique decomposition into prime factors have been leading principles of arithmetic and algebra on all levels ever since. *See* NUMBER THEORY.

The constant π, widely known from antiquity through the formulas $l = 2\pi r$, $A = \pi r^2$ for a circle and $A = 4\pi r^2$, $V = 4/3\pi r^3$ for a sphere, also enters the sophisticated formula, Eq. (8), derived by G. W.

$$\frac{\pi}{4} = 1 - {}^1/_3 + {}^1/_5 - {}^1/_7 + \cdots \qquad (8)$$

Leibnitz. Both kinds of formulas were much used for the approximate computation of this transcendental number, and a great deal of computational technique (including the use of computing machines) was initiated in the course.

The transcendental number e, the basis of natural logarithms, appeared in the eighteenth century in the formulas given as Eqs. (9), which have been

$$e^{2\pi\sqrt{-1}} = 1 \qquad \cos\varphi + \sqrt{-1}\sin\varphi = e^{\varphi\sqrt{-1}} \qquad (9)$$

of immeasurable consequence to all of mathematics. Even the age-old plane trigonometry was greatly streamlined by them. *See* NUMERICAL ANALYSIS.

Near the end of the eighteenth century, L. Euler gave for convex polyhedrons the formula $e_0 - e_1 + e_2 = 2$, in which e_0 is the number of vertices, e_1 of edges, and e_2 of surfaces. This was a first step in function theory of closed Riemann surfaces and in algebraic geometry, and it was a starting point for H. Poincaré's combinatorial topology. Poincaré's own formula, $B_p = B_{n-p}$ for Betti numbers of compact manifolds, is even today one of the most arresting statements of the theory.

Cauchy's formula, Eq. (10), is a center point from

$$f(z) = \frac{1}{2\pi i} \oint_C \frac{f(\zeta)\, d\zeta}{\zeta - z} \tag{10}$$

which the theory of complex functions radiates. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

Galileo's main achievement is embodied in the formula $s = 1/2gt^2$ for a falling body, and Newton's gravitational law in the formula $F = m_1m_2/r^2$, which also represents Coulomb's law of electricity. Einstein's daring formula $E = mc^2$ from special relativity created the atomic bomb, and W. Heisenberg's interpretation (1927) of the formula $pq - qp = h/2\pi i$ from quantum theory as an indeterminacy relation for physical observables has been an irritant to philosophers ever since. *See* RELATIVITY; UNCERTAINTY PRINCIPLE.

Foundations—mathematical logic. A prime demand on mathematics is that it be deductively rigorous, and a traditional model for intended rigor is Euclid's presentation of mathematical assertions in theorems. For the philosopher G. Spinoza in the seventeenth century, the phrase "more geometrico" is a synonym for "deductively rigorous." A theorem is a proposition which has been proved, excepting certain first theorems called axioms, which are admitted without proof; and to prove a theorem means to obtain it from other theorems by certain procedures of deduction or inference. It had long been commonplace that each branch of mathematics was based on its own axioms, but during the nineteenth century, mathematicians arrived at the insight that even the same branch might have alternate axioms. Specifically, there were envisaged alternate versions of two- and three-dimensional geometry, the axiom varied being the axiom on parallels. It was also recognized that a set of axioms becomes mathematically possible if it is logically consistent, that is, if one cannot deduce from the axioms to theorems one of which, as a proposition, is the negation of the other. *See* EUCLIDEAN GEOMETRY; NONEUCLIDEAN GEOMETRY.

At the same time certain developments led to the realization that not only the axioms but the rules of inference themselves might be, and even ought to be, subjected to variations. Now if axioms and rules of inferences are both viewed as subject to change, it is customary to speak of a mathematical system or also a formal system, and, of course, an irreducible first requirement is that the system be consistent after the manner just stated. Consistency alone is a somewhat negative property. There is a further property, called completeness, which is more positive, and which, if present, is very welcome. A system is complete if for any proposition which can be formulated it can be proved either that it holds or that its negation holds. If a theorem holds in a system and the system is altered, then the same proposition, or what corresponds to it in the new system, may become doubtful if not absolutely false; if it remains valid it may require a new proof because certain axioms or rules of inference are no longer available.

Some of the developments that led to doubt as to whether the traditional rules of inference are inviolate were the following.

1. G. Boole had found in 1854 that the classical Aristotelian connectives "and," "or," "negation of" for propositions follow rules similar to those which the operations addition, multiplication, "the negative of" obey in ordinary algebra (Boole's algebra of propositions); his conclusions took from rules of inference the status of untouchability.

2. G. Cantor, the founder of the theory of sets and operations between them, defined a set (intuitively or naïvely) as the collection of all objects having a certain property which is verbally expressible. Especially, "the set of all sets" is again a set and it has the peculiarity of being a set which contains itself as one of its elements. But this leads to the following contradictory situation (Russell's paradox): Divide the totality of all possible sets into two categories. A set shall belong to category I if it does not contain itself as an element, and to category II if it does contain itself as an element. Now form the set M whose elements are the sets of category I. It can now be reasoned by deductive steps admissible in Cantor's own theory that the set M cannot belong to either of the two categories, although the original division into categories did assign each set to one of them.

3. In 1904 E. Zermelo formulated the following axiom of choice: Given any family of nonvacuous sets $\{S\}$, no matter how (infinitely) large the family may be, it is possible to choose simultaneously an element $x = x_s$ from each given set S and thus to consider the set M consisting of precisely these elements. By the use of this axiom some striking theorems in classical mathematics could be proved which, without the use of the axiom, seemed to be logically out of reach entirely. Mathematicians began to wonder whether a theorem based on the axiom of choice is indeed valid or, at any rate, whether it has the same level of validity as one without it, and as a consequence, theorems employing the axiom of choice were frequently labeled as such. *See* SET THEORY.

Some of the doubts were resolved eventually; the most striking results are the following ones of K. Gödel. (1) Any consistent mathematical system which is sufficient for classical arithmetic must be incomplete. In other words, in a formal system which is expressible within the arithmetical language taught to a student of a secondary school in any country it is possible to formulate assertions which can be neither

proved nor disproved. (2) Any such system remains consistent if one adds to it the axiom of choice, so that working mathematics cannot disprove the axiom of choice. In 1963 P. Cohen showed that the axiom cannot be proved, either. (3) The so-called general continuum hypothesis is also consistent with ordinary mathematics, and, in fact, ordinary mathematics remains consistent if the axiom of choice and the general continuum hypothesis are added simultaneously. Cantor says that two sets A_1, A_2 have the same cardinal number if there exists a one-to-one correspondence between the elements of A_1 and A_2. Any infinite set has the same cardinal number as some proper subset of itself. For instance, the set of all integers $A_1 = \{1,2,3,\dots\}$ has the same cardinal number as the subset of all even integers $A_2 = \{2,4,6,8,\dots\}$ because the association of the number n in A_1 to the number $2n$ in A_2 is one-to-one. However, for a finite set this cannot happen, and the cardinal number of a finite set is the ordinary number of its elements. Next, Cantor says that a set A_2 has a larger cardinal number than A_1 if there is a one-to-one correspondence from A_1 to a subset of A_2 but not one from A_2 to A_1 or a subset of A_1. He then shows that if A is any set and S is the set of its subsets including the empty set, the cardinal number of S is greater than that of A. If A is finite and has n elements, then S has 2^n elements. Now 2^n is indeed greater than n, and for $n \geqq 2$ there are even other integer numbers between n and 2^n. This is true for finite sets. However, for infinite sets the situation seems to change radically, and the generalized continuum hypothesis asserts that there are no other cardinal numbers between those of A and S. *See* LOGIC.

Counterexamples. It follows easily from the definition of a derivative that a differentiable function $y = f(x)$ in $a < x < b$ is also continuous. K. Weierstrass deepened this statement considerably by demonstrating that the converse is not true. He constructed a function $f(x)$ that is continuous in $a < x < b$ but does not have a derivative at any point (his function does have right and left derivatives at all points). A construction of this kind is called a counterexample. Systematic preoccupation with counterexamples, which sometimes require considerable ingenuity, did not arise before the middle of the nineteenth century. Another early counterexample of considerable consequence referred to Fourier series. If $f(x)$ is periodic with period 2π, then its Fourier series is expression (11), from which, by definition, Eqs. (12) may be written. Fourier himself

$$^1/_2 a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx) \qquad (11)$$

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_0^{2\pi} f(t)\ \cos nt\ dt \\ b_n &= \int_0^{2\pi} f(t)\ \sin nt \end{aligned} \qquad (12)$$

stated somewhat vaguely that for "any" function $f(x)$ the partial sums of the Fourier series are convergent to $f(x)$. P. G. L. Dirichlet was the first to give a specific criterion of convergence. In his criterion the function $f(x)$ may have certain simple discontinuities. But in the present discussion these discontinuities are excluded and $f(x)$ must first of all be continuous. Dirichlet proved that the Fourier series converged everywhere to $f(x)$, if $f(x)$ is piecewise monotone, that is, if the interval of periodicity $0 \leqq x < 2\pi$ can be divided into a finite number of smaller intervals such that in each of these the function is either monotonely increasing or monotonely decreasing. Then, P. du Bois Reymond demonstrated that the requirement of monotoneity cannot be dispensed with entirely. He constructed a continuous periodic function $f(x)$ for which the Fourier series fails to converge in at least one point. By a general principle developed by A. Harnack (condensation of singularities) it is then possible to construct a continuous function whose Fourier series fails to converge at countably many points even if they are everywhere dense, for instance, at all rational points $x = p/q$. But to this day it is not known whether there exists a continuous function whose Fourier series fails to converge at all points, or at least at all points with the only exception of a set of Lebesgue measure 0. *See* FOURIER SERIES AND TRANSFORMS.

Counterexamples frequently fall into patterns; that is, the same underlying construction is used over again for different but related purposes. A construction used in many counterexamples is Cantor's ternary set. Divide the closed interval $0 \leqq x \leqq 1$ into three equal parts $(0, {}^1/_3, {}^1/_3, {}^2/_3)$, $({}^2/_3, 1)$ and remove the middle one, but only the open inner part of it ${}^1/_3 < x < {}^2/_3$. This leaves two closed intervals $0 \leqq x \leqq {}^1/_3$, ${}^2/_3 \leqq x \leqq 1$. In each of these remove the open inner third, that is, ${}^1/_9 < x < {}^2/_9$ and ${}^2/_3 + {}^1/_9 < x < {}^2/_3 + {}^2/_9$. This leaves four closed intervals together. From each, remove the open inner third, and so on. After n steps the number of intervals removed is $1 + 2 + 2^2 + \cdots + 2^{n-1}$, and the sum of the length of the removed intervals is Eq. (13). This value is less than

$$\frac{1}{3} + \frac{2}{3^2} + \frac{2^2}{3^3} + \cdots + \frac{2^{n-1}}{3^n} = 1 - \left(\frac{2}{3}\right)^n \qquad (13)$$

1, which must be so, because the total interval has length 1 and the removed intervals do not overlap. However, if one lets n go to ∞ the $({}^2/_3)^n$ tends to 0, and the aggregate length of all infinitely many intervals which have been removed has value 1. Thus, the point set which is left over ought somehow to have a one-dimensional measure which has value 0. This is so indeed in the theory of Lebesgue. Also the leftover set is nowhere dense. However, it is a very big set, by a crude counting of its points, at any rate. More precisely, it turns out that its cardinal number is the same as for the entire interval $0 \leqq x \leqq 1$. This can be deduced from the following fact, which is interesting for many purposes: Any number $0 \leqq x \leqq 1$ can be represented by a ternary expansion instead of by a decimal expansion, that is, by an expansion $0.a_1a_2a_3\dots$ in which each symbol $a_1, a_2, a_3, \dots$, has

one of the values 0, 1, 2. Now the leftover set consists of precisely those points in whose expansion the symbols $a_1, a_2, a_3, \ldots$, assume only the values 0, 2 and not 1.

Constructiveness and approximations. Some mathematicians object to mere existence proofs, and they demand that any proof also be constructive. The interpretations of this demand differ widely. Some proofs closely approach what a practical mathematician welcomes; if, for instance, a theorem asserts the existence of a number or a function, then the proof must also embody a procedure for actual computation of the solution, approximately, at least. Other versions are little more than the negative requirement that certain combinations of inference be avoided. There are also views which combine both; the best known among the last is the intuitionist view. It firmly demands a certain kind of constructiveness, which, however, does not necessarily guarantee the calculation by present-day computing machines. However, the actual stricture by which intuitionism became widely known is that proof by contradiction is not admissible. It is also called proof by double negation, and it is equivalent to the Aristotelian *tertium non datur*. It assumes tentatively that the proposition to be proved is false and from this assumption deduces a contradiction to a previously established theorem.

There is one difference between practical and theoretical demands on constructiveness. Sometimes it is possible to prove the existence of a function nonconstructively and after that, based on the existence statement, to devise a procedure for a manageable construction. To practical mathematicians this is satisfactory; to theoretical mathematicians it is not.

A demand superficially related to constructiveness, and without any overtones of "logic," is a desire to estimate the degree of approximation, whenever a task of approximation is being performed. For instance, a classical result of J. L. Lagrange for the Taylor series with remainder term states that if $f(x)$ has n derivatives in the interval $0 \leqq x \leqq 1$, then $f(x) = P_n(x) + R_n(x)$, where $P_n(x)$ is the polynomial in Eq. (14), and the "error term" $R_n(x)$ is in numerical

$$P_n(x) = \sum_{m=0}^{n} \frac{f^{(m)}(0)}{m!} x^m \tag{14}$$

value bounded by $M_n/n!$, where M_n is the maximum of $|f(n)(x)|$ in $0 \leqq x \leqq 1$. This result leads to the following topic. The polynomial in Eq. (14) can be formed only if $f(x)$ has n derivations. However, there is a theorem of Weierstrass that states that any function $f(x)$ which is continuous in an interval $a \leqq x \leqq b$ (and nothing more) can be approximated uniformly by the polynomials in Eq. (15). Each proof for

$$P_n(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_n \tag{15}$$

this theorem produces its polynomials, and there is an extensive theory for securing those for which the approximation is a best possible one.

The parallel question of approximating not by ordinary polynomials, Eq. (15), but by trigonometric polynomials, Eq. (16), leads to mathematical proce-

$$s_n(x) = {}^1\!/_2 a_0 + \sum_{m=1}^{n} (a_m \cos mx + b_m \sin mx) \tag{16}$$

dures which pervade all of science. If one wishes to make the difference, shown as expression (17),

$$\mathrm{Max}_{0 \leq x < 2\pi} \, |f(x) - s_n(x)| \tag{17}$$

rather small, then the partial sums of the Fourier series, Eq. (11), are not too good, meaning that it is not very advantageous to use the polynomials in Eq. (16) in which the coefficients a_m, b_m are the expressions (12). However, it improves the degree of approximation if one replaces the partial sums by their arithmetic means (Fejer sums), Eq. (18), which are again exponential polynomials, namely Eq. (19).

$$\sigma_n = \frac{s_0 + s_1 + \cdots + s_{n-1}}{n} \tag{18}$$

$$\sigma_n(x) = {}^1\!/_2 a_0 + \sum_{m=1}^{n} \left(1 - \frac{m}{n}\right) \cdot (a_m \cos mx + b_x \sin mx) \tag{19}$$

There are other averages which give even better approximations, and one is led to studying procedures of averaging and smoothing which are of importance to analysis, probability, and statistics. But a major development begins if, following F. W. Bessel and M. A. Parseval (nineteenth century), expression (17) is replaced by expression (20) as a measure for the de-

$$\left[\frac{1}{2\pi} \int_0^{2\pi} |f(x) - s_n(x)|^2 \, dx\right]^{1/2} \tag{20}$$

gree of approximation. In this case the partial sums of the Fourier series are indeed the best approximating sums for any order n, not only for ordinary Fourier series but for so-called orthogonal series in general. Pertinent application to the theory of acoustical and then optical waves led to an interpretation of the Fourier series, Eq. (11), as a spectral resolution of its function $f(x)$, and this interpretation has been gradually extended to a host of related and analogous expansions in analysis, algebra, theory of probability, quantum theory, and other parts of physics. The measuring of approximation, deviation, and dispersion by a square of integral mean—which is expression (20)—has become a strong and pervasive influence on scientific thinking.

Space in mathematics. If geometry is the mathematics of space, then, in a superficial sense, all mathematics began with geometry, because apparently it began with measurements of figures: length, area, volume, and size of angles. It did not concern itself with questions of shape but with clarifying and deciding when figures are equal or substantially equal with regard to form. The first true theory of geometry was the great theory of the Greeks, whose primary concern was study of the basic concept of equality of figures—their congruence and similarity—and the

Greeks were so determined to dissociate their theory from the preceding phase of merely making measurements that Euclid's extensive work, for instance, avoids to a fault any kind of actual measurements. In the case of the Pythagorean theorem $c^2 = a^2 + b^2$, there is no hint that it might be an equality between numbers. It is only an equality between areas based on congruences, two of the squares being cut up into pieces and then put together into the third square after the manner of a jigsaw puzzle. But, for all its lofty purposes, Greek geometry was too rigid and circumscribed to be able really to cope with the mathematical problem of space. It came to a standstill and geometry did not progress further until, with the advent of coordinate systems, introduced by Descartes and his predecessors, a better mathematics of space could be initiated.

If a cartesian coordinate system is etched into two- or three-dimensional euclidean space, then the space becomes a point set, each point being a pair (x^1, x^2) or a triple (x^1, x^2, x^3) of real numbers, and any figure a suitable subset of it. This is a deliberate process of arithmetization of space which unifies space and number at the base. It does not hamper geometry in its task of pursuing problems of shape but instead aids it. In the cartesian plane, two figures are similar if the points of one can be obtained from the points of the other by means of a transformation, Eqs. (21), where Eq. (22) applies and where, for some $\rho > 0$, Eq. (23) can be written. The similarity is a congruence if, and only if, $\rho = 1$ (orthogonal transformation). Now this analytic representation of congruence and similarity suggests a geometric examination of the most general linear transformations, Eq. (21), which are nonsingular, that is, for which the determinant $|\alpha_q^p| \neq 0$. They were virtually unknown to the Greeks, although they highlight the axiom of parallels of Euclid's geometry. A one-to-one transformation of the cartesian plane is such a linear transformation if, and only if, it carries a straight line into a straight line and parallel straight lines into parallel straight lines. Thus, a parallelogram goes into a parallelogram, and, in fact, given any two parallelograms there is a linear transformation which carries the one into the other no matter what the angles and ratios of sides in the two figures are. There is a geometry, the so-called affine geometry, in which any two parallelograms are considered equal. It cannot measure angles, and nonparallel segments cannot be compared for length. This geometry does have conics, however, and it can separate them into ellipses, parabolas, and hyperbolas. *See* CONFORMAL MAPPING.

$$\begin{aligned} y^1 &= a^1 + \alpha_1{}^1 x^1 + \alpha_2{}^1 x^2 \\ y^2 &= a^2 + \alpha_1{}^2 x^1 + \alpha_2{}^2 x^2 \end{aligned} \tag{21}$$

$$\alpha_1{}^1 \cdot \alpha_1{}^2 + \alpha_2{}^1 \cdot \alpha_2{}^2 = 0 \tag{22}$$

$$(\alpha_1{}^1)^2 + (\alpha_2{}^1)^2 = (\alpha_1{}^2)^2 + (\alpha_2{}^2) = \rho^2 \tag{23}$$

The family of all linear transformations constitutes a transitive group, and the subfamily of orthogonal transformations is already a transitive group. F. Klein made the pronouncement, which is generally accepted, that there arises a geometry on a space if on the space there is given a transitive group of transformations; two figures are considered equal whenever one figure can be carried into the other figure by one of the transformations. For the noneuclidean geometries of Bolyai-Lobachevski-Gauss various models have been exhibited which conform to this view, and perhaps the most interesting one is the following. In the plane (x^1, x^2) introduce the complex variable $z = x^1 + ix^2$ and consider the family of transformations, Eq. (24), for all constant complex numbers

$$w = e^{i\theta} \frac{z - a}{1 - \bar{a}z} \tag{24}$$

a for which $|a| < 1$ and all real numbers θ. They give rise to the following noneuclidean geometry. The space is not the entire z plane but only the unit disk: $|z| < 1$. A "point" is an ordinary point in it and a "straight line" of the new geometry is in the euclidean geometry either a diameter of the disk or any circular arc inside the disk whose end points are on the boundary line of the disk and which meets the boundary line of the disk at right angles. Through a point outside a straight line of the geometry go infinitely many straight lines which are "parallel" to it; that is, they do not meet it inside the disk. This makes the geometry a hyperbolic one. If, however, in a geometry any two straight lines always intersect, the geometry is called elliptic. The prototype of an elliptic geometry is a geometry on a surface of a sphere (for example, the surface of the Earth) in which the straight lines are the great circles. As a matter of fact, the hyperbolic geometry just described is in a very precise sense a counterpart to the elliptic geometry of great circles on a sphere.

The arithmetization of space led to a purely mathematical creation of n-dimensional space, euclidean and other, for any integer dimension n, by defining its points generally as n-tuples of real numbers $(x^1, \ldots, x^n)$ with suitable definitions for various geometrical relations between such points. The most eye-catching consequence in science was the four-dimensional space of the theory of relativity, but actually multidimensional geometry had been playing a part in physics before that. If a mechanical system involves M mass points, it was customary in effect to introduce the space of dimension $n = 3M$, whose points are the states of the system, that is, the n-tuples of coordinates $\{x_m{}^1, x_m{}^1, x_m{}^3\}$, $m = 1, \ldots, M$, at any one time point. Also, if there are restraints operative in the system, then the Lagrange-Hamilton theory suitably reduced the dimension of the space by the use of the free parameters of the system instead of the original n coordinates themselves. The use of free parameters spread from mechanical systems to other systems in physics and chemistry, and all so-called equations of state are geared to this. Finally, in quantum theory a state of a system has infinitely many coordinates and the infinitely dimensional space representing it is a Hilbert space. Also, partly under the influence of Hilbert space, mathematicians have become fascinated with

infinitely dimensional spaces in general. They are being studied intensively, and large parts of mathematics are being pressed into these new frames of reference.

The arithmetization of space is also reflected in the ever-widening use of graphs and charts. Any tabulated dependence of a number y on a number is a function $y = f(x)$ and hence representable as a curve on graph paper; a great amount of information is thus illustrated and stored. This is nothing but a practical application of the concept of a manifold in modern topology and differential geometry, and, as is frequently true, the abstract formulation of a mathematical concept is but a circumlocution of what common sense dictates. *See* ALGEBRA; ANALYTIC GEOMETRY; CALCULUS OF VECTORS; COORDINATE SYSTEMS; PROBABILITY; STATISTICS; TOPOLOGY.

Salomon Bochner

Bibliography. N. I. Achiezer, *Theory of Approximation*, 1956, reprint 1992; W. Cheney, *Introduction to Approximation Theory*, 2d ed., 1991; I. M. Copi, *Introduction to Logic*, 9th ed., 1993; H. D. Ebbinghaus, J. Flum, and W. Thomas, *Mathematical Logic*, 1994; V. Klenk, *Understanding Symbolic Logic*, 3d ed., 1994; R. L. Vaught, *Set Theory: An Introduction*, 2d ed., 1994.

Matrix calculus

That branch of mathematics which deals with matrices whose elements are functions of one or more independent variables.

The derivative of a matrix $A(t)$ whose elements $a_{ij}(t)$ are functions of a variable t is defined in Eq. (1),

$$\frac{dA}{dt} = \lim_{\Delta t \to 0} \frac{A(t + \Delta t) - A(t)}{\Delta t} = \left(\frac{da_{ij}}{dt}\right) \tag{1}$$

where da_{ij}/dt represents the matrix whose elements are da_{ij}/dt. Thus dA/dt is formed by replacing the elements of $A(t)$ by their derivatives. *See* MATRIX THEORY.

If the matrices A and B are functions of t, then Eqs. (2) are satisfied by the operation of differentiation.

$$\begin{aligned} \frac{d}{dt}(A + B) &= \frac{dA}{dt} + \frac{dB}{dt} \\ \frac{d}{dt}(AB) &= \frac{dA}{dt}B + A\frac{dB}{dt} \end{aligned} \tag{2}$$

In differentiating a product the order of the factors must be preserved. Thus Eq. (3) is obtained for

$$\frac{d}{dt}A^2 = \frac{dA}{dt}A + A\frac{dA}{dt} \tag{3}$$

$(d/dt)A^2$ and not $2A(dA/dt)$. From $A^{-1}A = I$, Eq. (4) is found.

$$\frac{dA^{-1}}{dt} = -A^{-1}\frac{dA}{dt}A^{-1} \tag{4}$$

The integral of $A(t)$ is defined as the matrix whose elements are integrals of $a_{ij}(t)$.

Every square n by n matrix satisfies a polynomial equation of lowest degree, called its minimum equation, which is shown in Eq. (5). Equation (5) may be

$$A^m + a_1A^{m-1} + \cdots + a_mI = 0 \quad (m \le n) \tag{5}$$

used to express all powers of $A > m - 1$ in terms of $I, A, A^2, \ldots, A^{m-1}$. Therefore any matrix polynomial $f(A)$ of degree $k > m - 1$ may be replaced by a polynomial $F(A)$ of degree $< m$. If

$$f(z) = \sum_{k=0}^{\infty} c_k z^k$$

is a power series whose radius of convergence is r, and if all eigenvalues of A are less than r in absolute value, then (A) is defined as the matrix series shown in Eq. (6).

$$f(A) = c_0I + \sum_{k=0}^{\infty} c_kA^k \qquad |\lambda_A| < r \tag{6}$$

Thus the binomial series yields Eq. (7) for rational n; and since the exponential series for e^z converges for all z, Eq. (8) can be set up for arbitrary

$$(I + A)^n = I + \sum_{k=1}^{\infty} \binom{n}{k} A^k \qquad |\lambda_A| < 1 \tag{7}$$

$$e^A = I + \sum_{k=1}^{\infty} \frac{1}{k!}A^k \tag{8}$$

A. The matrix exponential has the properties $e^0 = I$, $e^Ae^{-A} = I$; also, $e^Ae^B = e^{A+B}$ when $AB = BA$. From the series for e^{tA}, $de^{tA}/dt = Ae^{tA}$.

The minimum equation (5) may also be used to replace a power series for $f(A)$ by a polynomial $F(A)$. If Eq. (5) has distinct roots $\lambda_1, \lambda_2, \ldots, \lambda_m$ (eigenvalues of A), $F(A)$ has the eigenvalues $F(\lambda_i) = f(\lambda_i)$. The polynomial $F(\lambda)$ is completely determined by the m values $f(\lambda_i)$ that it assumes when $\lambda = \lambda_i$. Hence $F(\lambda)$, and also $F(A) = f(A)$, may be determined by Lagrange's interpolation formulas given in Eqs. (9) and (10), where each product has $m - 1$ factors.

$$f(A) = f(\lambda_1)C_1(A) + \cdots + f(\lambda_m)C_m(A) \tag{9}$$

$$C_j(A)\frac{\prod_{i \ne j}(A - \lambda_iI)}{\prod_{i \ne j}(\lambda_j - \lambda_i)} \tag{10}$$

When Eq. (5) has repeated roots, say $\lambda_1 = \lambda_2 = \lambda_3$, the corresponding terms in the modified equation (9) are given by formula (11), and Eq. (10) does

$$f(\lambda_1)C_1(A) + f'(\lambda_1)C_2(A) + f''(\lambda_1)C_3(A) \tag{11}$$

not apply. But since the $C_j(A)$ do not depend on f, they can be found by choosing m functions $f_i(\lambda)$ so that the m linear equations (9) and (10) in C_j have a nonzero determinant, and then solving them for the $C_j(A)$.

Consider, for example, a 4 by 4 matrix A with $\lambda_1 = \lambda_2 = \lambda_3$, $\lambda_4 \ne \lambda_1$. If $m = 4$, the basic equation (9) is

now Eq. (12). By choosing $f(\lambda) = 1, \lambda - \lambda_1, (\lambda - \lambda_1)^2$, and $(\lambda - \lambda_1)^3$, respectively, Eqs. (13) are obtained,

$$f(A) = f(\lambda_1)C_1 + f'(\lambda_1)C_2 + f''(\lambda_1)C_3 + f(\lambda_4)C_4 \quad (12)$$

$$\begin{aligned} I &= C_1 + 0 + 0 + C_4 \\ A - \lambda_1 I &= 0 + C_2 + 0 + (\lambda_4 - \lambda_1)C_4 \\ (A - \lambda_1 I)^2 &= 0 + 0 + 2C_3 + (\lambda_4 - \lambda_1)^2 C_4 \\ (A - \lambda_1 I)^3 &= 0 + 0 + 0 + (\lambda_4 - \lambda_1)^3 C_4 \end{aligned} \quad (13)$$

which give C_4, C_3, C_2, and C_1.

Eigenmatrices. Any n by n matrix A is reduced to its Jordan form J by finding its eigenmatrix E whose columns are the eigenvectors $e_1, \ldots, e_n$ belonging to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. If λ_1 is simple, A has a proper eigenvector e_1 such that $(A - \lambda_1 I)e_1 = 0$. If λ_1 is k-tuple, with one proper eigenvector e_1 and $k - 1$ generalized eigenvectors $e_2, e_3, \ldots, e_k$ which satisfy Eqs. (14), then $E = (e_1|e_2| \ldots |e_n)$ and Eqs. (15) are valid.

$$(A - \lambda_1 I)e_j = e_{j-1} \qquad j = 2, 3, \ldots, k \quad (14)$$

$$E^{-1}AE = J \qquad A = EJE^{-1} \quad (15)$$

If $f(z)$ is an analytic function, then Eq. (16) holds,

$$f(A) = f(EJE^{-1}) = Ef(J)E^{-1} \quad (16)$$

which gives $f(A)$ when the simpler $f(J)$ is known. Thus, when A has distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, J and $f(J)$ have the form shown in Eqs. (17).

$$\begin{aligned} J &= \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \\ f(J) &= \mathrm{diag}[f(\lambda_1), \ldots, f(\lambda_n)] \end{aligned} \quad (17)$$

Equation (16) is an explicit formula for $f(A)$ and agrees with the former interpolation method; it requires, however, the eigenmatrix E and its reciprocal. Both methods are shown in the following example.

The eigenvalues of

$$A = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}$$

are $\lambda_1 = 1, \lambda_2 = 2$ and $e_1 = (1, 1)^\top, e_2 = (1, 2)^\top$. Then Eq. (16) gives Eq. (18).

$$f(A) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} f(1) & 0 \\ 0 & f(2) \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & -1 \\ 2 & -1 \end{pmatrix} f(1) + \begin{pmatrix} -1 & 1 \\ -2 & 2 \end{pmatrix} f(2) \quad (18)$$

The interpolation method uses Eq. (9), which in this case reduces to Eq. (19). Choose $f(\lambda)$ as

$$f(A) = f(1)C_1 + f(2)C_2 \quad (19)$$

$\lambda - 1, \lambda - 2$, respectively; then C_2 and C_1 are given by Eqs. (20) in agreement with Eq. (18).

$$\begin{aligned} C_2 &= A - I = \begin{pmatrix} -1 & 1 \\ -2 & 2 \end{pmatrix} \\ -C_1 &= A - 2I = \begin{pmatrix} -2 & 1 \\ -2 & 1 \end{pmatrix} \end{aligned} \quad (20)$$

As an example with multiple eigenvalues, consider A given by Eq. (21). The modified Eq. (19) is now Eq. (22). Choose $f(\lambda)$ as $1, \lambda - 1, (\lambda - 1)^2$, respectively; then C_1, C_2, and C_3 are given by Eqs. (23).

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \qquad \lambda_1 = \lambda_2 = \lambda_3 = 1 \quad (21)$$

$$f(A) = f(1)C_1 + f'(1)C_2 + f''(1)C_3 \quad (22)$$

$$I = C_1 \qquad A - I = C_2 \qquad (A - I)^2 = 2C_3 \quad (23)$$

Let $f(\lambda) = e^{\lambda t}$; then $f'(\lambda) = te^{\lambda t}$, $f''(\lambda) = t^2 e^{\lambda t}$, and so e^{tA} is given by Eq. (24), which can be expressed as Eq. (25).

$$e^{tA} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} e^t + \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} te^t + \begin{pmatrix} 0 & 0 & 1/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} t^2 e^t \quad (24)$$

$$e^{tA} = \begin{pmatrix} 1 & t & 1/2 t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} e^t \quad (25)$$

Differential equations. Let $x_1(t), \ldots, x(t)$ be n unknown functions and $X(t)$ their column vector. Then if $A(t)$ is an n by n matrix whose elements are continuous functions of t, the system of n linear differential equations (26) may be solved by successive approx-

$$\frac{dX}{dt} = AX \qquad X(0) = X_0 \quad (26)$$

imations. The nth approximation is obtained from the $(n - 1)$th by integrating Eqs. (27) from $t = 0$ to t. Then a solution is obtained in the form of Eq. (28), where Ω is given by Eq. (29) and is called the matricant of A. When A has constant elements, Eq. (29) reduces to Eq. (30).

$$\frac{dX_n}{dt} = AX_{n-1} \qquad n = 1, 2, \ldots \quad (27)$$

$$X = \lim_{n \to \infty} X_n = \Omega X_0 \quad (28)$$

$$\Omega = I + \int_0^t A\,dt + \int_0^t A\,dt \int_0^t A\,dt + \cdots \quad (29)$$

$$\Omega = I + At + \frac{1}{2!}A^2 t^2 + \frac{1}{3!}A^3 t^3 + \cdots = e^{tA} \quad (30)$$

Consider, for example, the system given by Eqs. (31), or in matrix form by Eq. (32). Then

from Eq. (25) one obtains Eq. (33).

$$\begin{aligned} \frac{dx_1}{dt} &= x_1 + x_2 \\ \frac{dx_2}{dt} &= x_2 + x_3 \\ \frac{dx_3}{dt} &= x_3 \end{aligned} \tag{31}$$

$$\frac{dX}{dt} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} X = AX \tag{32}$$

$$X = e^{tA}X_0 = e^t \begin{pmatrix} 1 & t & {}^1\!/_2 t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} X_0 \tag{33}$$

In a linear system with variable coefficients $p_{ij}(t)$, the differential Eq. (26) becomes Eqs. (34). If X_1,

$$\frac{dX}{dt} = P(t)X \qquad P(t) = [p_{ij}(t)] \tag{34}$$

X_2, X_3 are three linearly independent solutions of Eqs. (34), let $Y = (X_1|X_2|X_3)$ denote the 3 by 3 matrix whose columns are X_1, X_2, X_3. Then Eqs. (34) are equivalent to the matrix equation (35).

$$\frac{dY}{dt} = P(t)Y \tag{35}$$

If one differentiates det Y by rows and adds the resulting determinants, one obtains Eq. (36), where

$$\begin{aligned} \frac{d}{dt}(\det Y) &= (p_{11} + p_{22} + p_{33})\det Y \\ &= (\operatorname{tr} P)\det Y \end{aligned} \tag{36}$$

tr P denotes the trace of P and det Y is the determinant of Y. On integration this yields the Jacobi identity given in Eq. (37). Since det Y is not

$$\det Y = c \exp \int_{t_0}^{t} \operatorname{tr} P \, dt \tag{37}$$

identically zero, $c \neq 0$; hence the matrix integral $Y(t)$ is nonsingular; that is, det $Y \neq 0$.

If Y_1 is a particular solution of Eq. (35) and C is an arbitrary 3 by 3 constant matrix, then $Y = Y_1C$ is also a solution. Moreover, all solutions of Eq. (35) are contained in $Y = Y_1C$. For if Y is a solution of Eq. (35), then Eq. (4) gives Eq. (38), hence $Y_1^{-1}Y = C$ and $Y = Y_1C$.

$$\begin{aligned} \frac{d}{dt}(Y_1^{-1}Y) &= \left(-Y_1^{-1}\frac{dY_1}{dt}Y_1^{-1}\right)Y + Y_1^{-1}\frac{dY}{dt} \\ &= -Y_1^{-1}PY + Y_1^{-1}PY = 0 \end{aligned} \tag{38}$$

If the coefficients $p_{ij}(t)$ are periodic with the same period ω, that is, $P(t + \omega) = P(t)$, then the system shown in Eq. (35) can be reduced to one with constant coefficients by a linear transformation $Y = L(t)Z$, where $L(t)$ is a nonsingular matrix of period ω. Let $Y_1(t)$ be the solution of Eq. (35) for which $Y_1(0) = I$; then $Y_1(t + \omega)$ is also a solution, and hence $Y_1(t + \omega) = Y_1(t)C$, where C is a nonsingular matrix of constants. If a matrix B is determined so that $e^{B\omega} = C$, then Eqs. (39) and (40) result, which reduce Eq. (35) to

$$Y = L(t)Z \tag{39}$$

$$L(t) = Y_1(t)e^{-Bt} = L(t + \omega) \tag{40}$$

$dZ/dt = BZ$, whose solution $Z_1(t)e^{Bt}$ corresponds to $Y_1(t)$. Note that $Y_1(t) = Y_1(t)e^{-Bt}Z_1(t)$ requires $Z_1(t) = e^{Bt}$, $dZ_1/dt = BZ_1$; and since all solutions of this equation have the form $Z(t) = CZ_1(t)$, $dZ/dt = BZ$.

Louis Brand

Bibliography. P. Bugl, *Differential Equations: Matrixes and Models*, 1994; R. A. Horn and C. R. Johnson, *Matrix Analysis*, 1985, paper 1990; A. W. Joshi, *Matrices and Tensors in Physics*, 1995; J. R. Magnus and H. Neudecker, *Matrix and Differential Calculus with Applications in Statistics and Econometrics*, 2d ed., 1999.

Matrix isolation

A technique for providing a means of maintaining molecules in an inert medium at low temperature for spectroscopic study. This method is particularly well suited for preserving reactive species in a solid, inert environment. Absorption (infrared, visible, and ultraviolet), electron-spin resonance, and laser-excitation spectroscopes can be used to examine elusive molecular fragments such as free radicals that may be postulated as important controlling intermediates for chemical transformations used in industrial reactions, high-temperature molecules that are in equilibrium with solids at very high temperatures, weak molecular complexes that may be stable at low temperatures, new reactive molecular species, and molecular ions that are produced in plasma discharges or by high-energy radiation. The matrix isolation technique enables spectroscopic data to be obtained for reactive molecular fragments, many of which cannot be studied in the gas phase. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; INFRARED SPECTROSCOPY; SPECTROSCOPY.

Experimental apparatus. The experimental apparatus for matrix isolation experiments is designed with the method of generating the molecular transient and performing the spectroscopy in mind. **Figure 1** shows the cross section of a vacuum vessel used for absorption spectroscopic measurements (such as infrared, visible, and ultraviolet). The optical windows must be transparent to the examining radiation. The rotatable cold window is cooled to 4 to 20 K (−452 to −424°F) by closed-cycle refrigeration or liquid helium. The matrix sample is introduced through the spray-on line at rates of 1–5 millimoles per hour; argon is the most widely used matrix gas, although neon, krypton, xenon, and nitrogen are also used. The reactive species can be generated in a number of ways: mercury-arc photolysis of a trapped precursor molecule through the quartz window, evaporation from a Knudsen cell in the heater, chemical reaction of atoms evaporated from the Knudsen

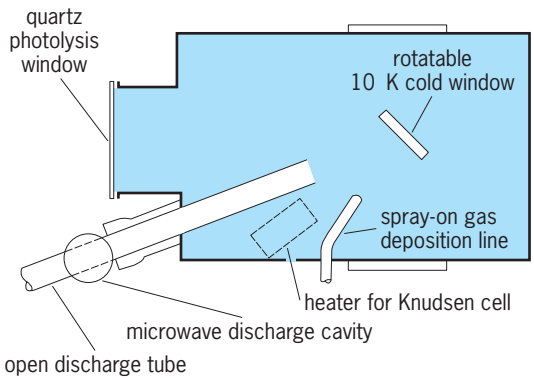

Fig. 1. Vacuum-vessel base cross section for matrix photoionization experiments. 10 K = −263°C or −442°F.

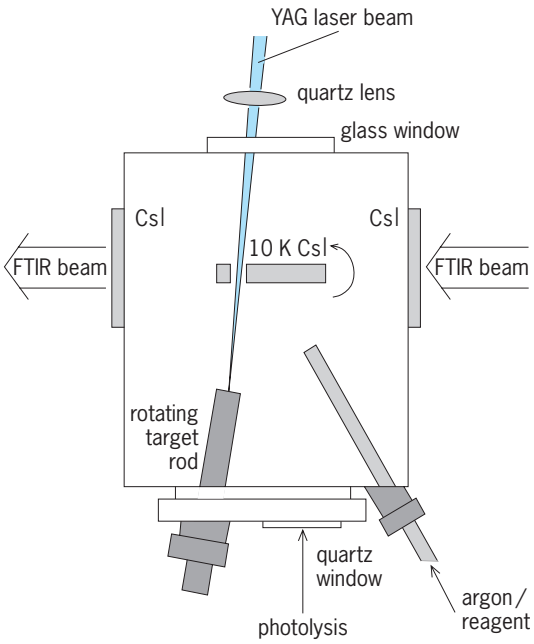

Fig. 2. Schematic plan view of the matrix assembly for pulsed-laser ablation.

cell with molecules deposited through the spray-on line, neodymium-yttrium aluminum garnet (Nd-YAG) laser evaporation of solids, and vacuum-ultraviolet photolysis of molecules deposited from the spray-on line by radiation from discharge-excited atoms flowing through the tube. The cold window is rotated for transmission absorption spectroscopic measurements using ultraviolet-visible or infrared spectrometers.

Figure 2 is a schematic diagram of the laser ablation matrix isolation apparatus. The ablated metal atoms intersect the gas sample during codeposition on the cold cesium iodide (CsI) window, where collisions and reactions occur. The Nd-YAG fundamental at 1064 nanometers in 5–50-millijoule pulses of 10-nanosecond duration was focused (spot size approximately 0.1 nm) onto a rotating metal target. For laser excitation studies, the sample is deposited on a tilted copper wedge which is grazed by the laser beam, and light emitted or scattered at approximately 90° is examined by a spectrograph. In electron-spin resonance studies, the sample is condensed on a sapphire rod that can be lowered into the necessary waveguide and magnet. The chemical versatility of the matrix technique can perhaps be best described by considering a number of important examples. *See* PHOTOLYSIS.

Applications. The first free radical stabilized in sufficient concentration for matrix infrared detection was formyl (HCO·). Hydrogen iodide (HI) was deposited in a carbon monoxide (CO) matrix and photolyzed with a mercury arc; hydrogen atoms produced by dissociation of HI reacted with CO in the cold solid to produce HCO. The infrared spectrum of HCO provided vibrational fundamentals and information about the chemical bonding in this reactive species. Free radicals have been produced in matrices with a variety of techniques such as ultraviolet photolysis [reaction (1)]; vacuum ultraviolet photolysis [reaction (2)]; fluorine atom [reaction (3)]; metal atom [reaction (4)]; and hydrogen atom [reaction (5)].

$$OF_2 + \text{ultraviolet light} \rightarrow OF\cdot + F\cdot \quad (1)$$

$$CHF_3 + \text{vacuum ultraviolet light} \rightarrow CF_3\cdot + H\cdot \quad (2)$$

$$C_2H_2 + F\cdot \rightarrow C_2H_2F\cdot \quad (3)$$

$$CCl_4 + Li \rightarrow CCl_3\cdot + LiCl \quad (4)$$

$$H\cdot + O_2 \rightarrow HO_2\cdot \quad (5)$$

See CHEMICAL BONDING.

The first molecular ionic species characterized in matrices, $Li^+O_2^-$ was formed by the cocondensation reaction of lithium (Li) atoms and oxygen (O_2) molecules at high dilution in argon. The infrared spectrum exhibited a weak $(O \leftrightarrow O)^-$ stretching vibration and two strong $Li^+ \leftrightarrow O_2^-$ stretching vibrations, as shown at the top of **Fig. 3** for the ^7Li and $^{16,18}O_2$ isotopic reaction. The 1-2-1 relative-intensity oxygen isotopic triplets in this experiment showed that the oxygen atomic positions in the molecule are equivalent and indicated an isosceles triangular structure. The ionic model for the bonding

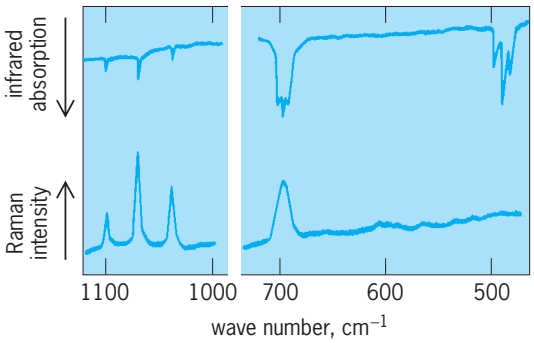

Fig. 3. Infrared and Raman spectra of lithium superoxide, $Li^+O_2^-$, using lithium-7 and 30% $^{18}O_2$, 50% $^{16}O^{18}O$, and 20% $^{16}O_2$ at concentrations of $Ar/O_2 = 100$. The Raman spectrum was recorded by using 200 mW of 488-nm excitation at the sample.

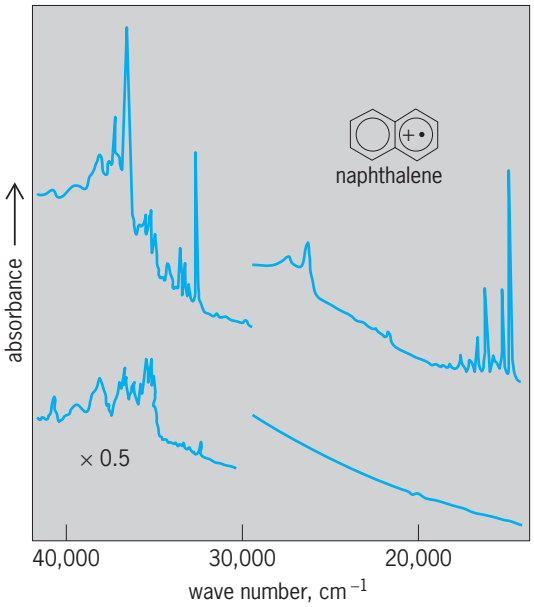

Fig. 4. Visible and ultraviolet absorption spectra of naphthalene samples in solid argon at 20 K. Upper curves show Ar/CCl_4/N = 3000/4/1, and lower curves show the same sample after 15 s of irradiation at 220–1000 nm from a high-pressure mercury arc.

in $Li^+O_2^-$ was confirmed by contrasting intensities between the laser-Raman and infrared-absorption spectra shown in Fig. 3. *See* RAMAN EFFECT.

High-temperature molecules like lithium fluoride (LiF) can be trapped in matrices by evaporating the molecule from the crystalline solid in a Knudsen cell at high temperature or by reacting lithium atoms with fluorine molecules during condensation. The latter method has been used to synthesize the calcium oxide (CaO) molecule from the calcium atom–ozone reaction, which gives the calcium ozonide ion pair $Ca^+O_3^-$ and CaO diatomic products.

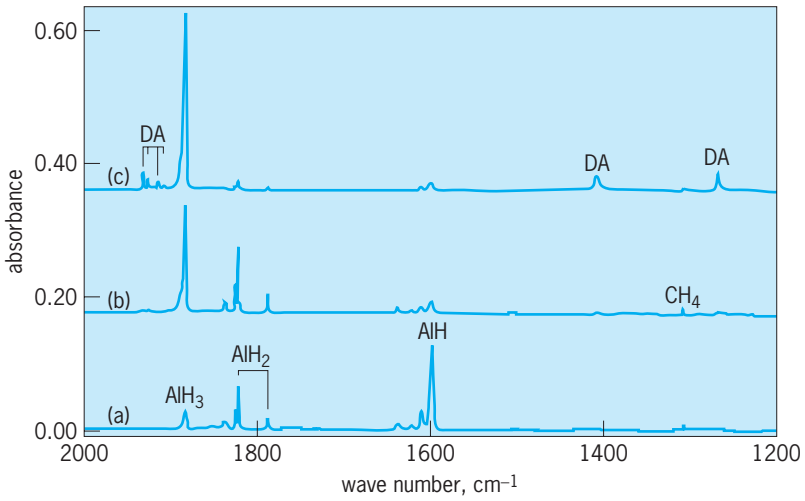

Fig. 5. Infrared spectra in the 2000–1200 cm^{-1} region for the products of the reactions of laser-ablated Al atoms with pure hydrogen co-deposited at 3.5 K. (*a*) Spectrum obtained from initial deposited sample. (*b*) Spectrum after >290 nm irradiation. (*c*) Spectrum after >220 nm irradiation.

Continuous exposure of a condensing sample to argon-resonance radiation during sample condensation has been used to produce molecular cations for spectroscopic study. In the case of fluoroform (CHF_3), photolysis produced the $CF_3\cdot$ radical, which may be photoionized by a second 11.6-eV photon to give CF_3^+. The infrared spectrum of CF_3^+ revealed a very high C-F vibrational fundamental, which indicates substantial pi bonding in this planar carbocation.

The naphthalene ($C_{10}H_8$; mothballs) cation is an interesting case for matrix study since a wealth of data is available from other spectroscopic techniques. Vacuum ultraviolet photoionization of $C_{10}H_8$ during condensation with excess argon produced new absorption band systems at 14,810, 21,697, and 26,240 cm^{-1} for the naphthalene cation. Using a 15-s mercury-arc irradiation with an electron-trapping molecule (CCl_4) present, **Fig. 4** shows that the above three bands plus two others, at 32,510 and 36,417 cm^{-1}, were produced. In the latter case, one ultraviolet photon excited $C_{10}H_8$ and a second photon produced ionization, the removed electron being trapped by CCl_4 (as $CCl_3\cdot$ and Cl^-), and in the former case, one vacuum-ultraviolet photon ionized $C_{10}H_8$ directly to give the $C_{10}H_8^+$ cation trapped in solid argon for observation of its electronic absorption spectrum. Vibronic intervals (vibrational fine structure) in the absorption spectrum provide information on vibrational fundamentals in the cation excited states.

Molecular complexes involving hydrogen fluoride (HF) serve as useful prototypes for the understanding of the important phenomenon of hydrogen bonding. An interesting chemical case is ammonia (NH_3) and HF, which on the macroscopic scale produce the salt ammonium fluoride but on the microscopic scale give the NH_3-HF complex. The infrared spectrum of this complex reveals vibrations for NH_3 and HF perturbed by their association in the complex. Fourier-transformed infrared spectroscopy is particularly advantageous in these studies because of the high vibrational frequencies for HF species and low sample transmission in this region. The ammonia symmetric bending motion is considerably blueshifted, and the hydrogen fluoride stretching fundamental is markedly redshifted. These shifts attest to a strong intermolecular interaction within the complex. *See* HYDROGEN BOND.

Ozone (O_3) is a very important molecule in the upper atmosphere, as it absorbs harmful ultraviolet radiation. In the laboratory, this ultraviolet dissociation of ozone provides oxygen atoms for chemical reactions. Photochemical reactions of elemental phosphorus (P_4), an extremely reactive molecule well suited for matrix isolation studies, and ozone produce a new low oxide of phosphorus, P_4O, which is involved as a reactive intermediate in the striking of a match. The infrared spectrum of P_4O shows a strong terminal PO bond-stretching fundamental and characterizes the P_4O structure as tetrahedral P_4 with an O atom cap. Further photochemical reactions

of phosphorus and oxygen produce and trap the reactive molecule $PO_2\cdot$, which can also be formed by laser evaporation of red phosphorus followed by reaction with a stream of oxygen gas. *See* FREE RADICAL; REACTIVE INTERMEDIATES.

The use of pulsed laser ablation to produce metal atoms for reactions to make new chemical species is illustrated by investigations of metal carbonyls and nitrosyls and metal hydrides. When the cold window (Fig. 1) is cooled below 4 K, neon and pure hydrogen are condensed; and metal atom reactions can be done in an inert solid neon or a reactive solid hydrogen host. This is illustrated by reactions of aluminum (Al) atoms to form in solid hydrogen the dialane (Al_2H_6) molecule,

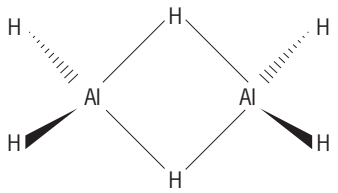

which is isostructural to diborane—a model for bridge bonding. Laser-ablated and excited Al atoms co-deposited with pure hydrogen at 3.5 K react with the hydrogen host to give absorptions for AlH (1599 cm^{-1}), AlH_2 (1822, 1788 cm^{-1}), and AlH_3 (1884 cm^{-1}), which are shown in **Fig. 5***a*. Irradiation of the sample with a mercury-arc lamp filtered by a Pyrex plate (wavelengths >290 nm) promotes the reaction of AlH with H_2 to form AlH_3 as shown in Fig. 5*b*. Further irradiation with the full arc (wavelengths >220 nm; Fig. 5*c*) almost destroys AlH_2, making more AlH_3 and new bands at 1932, 1915, 1408, and 1268 cm^{-1} (labeled DA). These bands shift when the reaction is done in pure D_2 and confirm that a new aluminum hydride species is formed. The important point is that the two higher frequency bands are due to terminal Al-H stretching modes and the 1408 and 1268 cm^{-1} bands are in the region for Al-H-Al bridge-bond-stretching modes, which characterizes a new molecule containing both terminal and bridge aluminum-hydrogen bonds. The new molecule $H_2Al(H)_2AlH_2$ is formed from the combination of two AlH_3 molecules [reactions (6)–(8)] and characterized experimentally from its infrared spectrum.

$$Al + H_2 \rightarrow AlH + H \tag{6}$$

$$AlH + H_2 \rightarrow AlH_3 \tag{7}$$

$$AlH_3 + AlH_3 \rightarrow H_2Al(H)_2AlH_2 \tag{8}$$

The vibrational frequencies assigned to new molecules through matrix isolation infrared spectroscopy can be determined with a high degree of accuracy using density functional theory calculations, a computational quantum-chemistry approach. Such calculations predict that the bridge-bonded dialane molecule has seven infrared active vibrational modes with observable infrared intensity. The above four stretching modes plus three lower-frequency bending modes are observed for dialane in solid hydrogen in excellent agreement with the calculated frequencies. This agreement between calculated and observed frequencies confirms the identification of the novel bridge-bonded dialane molecule in solid hydrogen. *See* BORANE; METAL CARBONYL; METAL HYDRIDES; MOLECULAR SIMULATION; QUANTUM CHEMISTRY.

Lester Andrews

Bibliography. L. Andrews, Fourier transform infrared spectra of HF complexes in solid argon, *J. Phys. Chem.*, 88:2940–2949, 1984; L. Andrews, Matrix infrared spectra and density functional calculations of transition metal hydrides and dihydrogen complexes, *Chem. Soc. Rev.*, 33:123–132, 2004; L. Andrews, W. D. Bare, and G. V. Chertihin, Reactions of laser-ablated V, Cr and Mn atoms with nitrogen atoms and molecules, *J. Phys. Chem.*, 101:8417–8427, 1997; L. Andrews and A. Citra, Infrared spectra and density functional calculations on transition metal nitrosyls, *Chem. Rev.*, 102:885–911, 2002; L. Andrews and M. Moskovits (eds.), *Chemistry and Physics of Matrix Isolated Species*, 1989; L. Andrews and X. Wang, The infrared spectrum of Al_2H_6 in solid hydrogen, *Science*, 299:2049–2052, 2003; V. E. Bondybey, A. M. Smith, and J. Agreiter, New developments in matrix isolation spectroscopy, *Chem. Rev.*, 96:2113–2134, 1996; M. Zhou, L. Andrews, and C. W. Bauschlicher, Jr., Spectroscopic and theoretical investigations of binary unsaturated transition metal carbonyl cations, neutrals, and anions, *Chem. Rev.*, 101:1931–1961, 2001.

Matrix mechanics

A formulation of quantum theory in which the operators are represented by time-dependent matrices. For a discussion of quantum-mechanical operators and other information essential to an understanding of this article *see* NONRELATIVISTIC QUANTUM THEORY

Matrix mechanics is not useful for obtaining quantitative solutions to actual problems; on the other hand, because it is concisely expressed in a form independent of special coordinate systems, matrix mechanics is useful for proving general theorems. For the purposes of the following brief discussion it is sufficient to consider a one-dimensional spinless system, described by a wave function $\psi(x,t)$.

A matrix A is a rectangular array of numbers; the number of rows or columns of the array may be finite or infinite; A_{mn} is the element (number) in the mth row and the nth column; m, n may be discrete or continuous indices.

Symbolic addition $C = A + B$ is defined as meaning that $C_{mn} = A_{mn} + B_{mn}$ for each m and n, with the implication that the rows of A, B have the same indexing m and that the columns also have the same indexing m. Symbolic multiplication $C = AB$ is defined to mean $C_{mn} = \Sigma_j A_{mj} B_{jn}$, summed over all values of j, again with the implication that the columns of A

and the rows of B have the same indexing j; when j is continuous, the sum in the definition of C_{mn} is replaced by an integral. The adjoint $A^\dagger$ of A is obtained by interchanging rows and columns of A and then taking the complex conjugate of each element, that is, $A_{mn}{}^\dagger = (A_{nm})^* \equiv A_{nm}{}^*$; $(AB)^\dagger = B^\dagger A^\dagger$. *See* MATRIX THEORY.

Suppose $\phi(x,t)$ is expanded in terms of any complete set of orthonormal functions u_n, as shown in Eq. (1). The matrix elements (in this u representation) of any quantum-mechanical operator A are defined by Eq. (2). For instance, when $A \equiv p_x$, A is

$$\psi(x, t) = \sum_n \psi_n(t) u_n(x) \tag{1}$$

$$A_{mn} = \int_{-\infty}^{\infty} dx\, u_m{}^*(x) A u_n(x) \tag{2}$$

replaced by $(\hbar/i)\ \partial/\partial x$ in Eq. (2). From Eqs. (1) and the orthonormality of $u_n(x)$, if Eq. (3) holds, then Eq. (4) can be obtained. This result shows that the

$$A\psi \equiv \phi = \sum_n \phi_n u_n \tag{3}$$

$$\phi_m = \sum_n A_{mn} \psi_n \tag{4}$$

equation $\phi = A\psi$, which states that ϕ is the function resulting from the operation A on ψ, can be interpreted equally well as a matrix equation, provided ψ, ϕ are represented by single-column matrices whose elements are respectively ψ_n, ϕ_n. Similarly, the expectation value of the operator A in the state ψ is given by Eq. (5), that is, in matrix notation $\langle A \rangle = \psi^\dagger A \psi$; the adjoint of a column matrix ψ_m is the row $\psi_m{}^*$.

$$\langle A \rangle = \int \psi^* A \psi = \int \psi^* \phi = \sum_{m,n} \psi_m{}^* A_{mn} \psi_n \tag{5}$$

The time-dependent Schrödinger equation implies $\psi(t) = \exp(-iHt/\hbar)\ \psi(0)$. Thus, noting that, because the operator H is hermitian, the matrix H is self-adjoint, $H^\dagger = H$, and the expectation value of A at time t is as shown in Eq. (6). Equation (6) shows that $\langle A \rangle_t$ can be computed in two equivalent ways: (i) the conventional Schrödinger representation, in which ψ is time-dependent and the A_{mn}'s are time-independent; (ii) the Heisenberg representation, in which ψ is time-independent and equal to $\psi(0)$ but A is represented by time-dependent matrix elements given in Eq. (7). Differentiating Eq. (7), one obtains Eq. (8); that is, $dA/dt = (i\hbar)^{-1}\ (A,H)$, where $(A,B) \equiv AB - AB$.

$$\langle A \rangle_t = \psi^\dagger(0) \exp(iHt/\hbar) \times A \exp -(iHt/\hbar)\psi(0) \tag{6}$$

$$A_{mn}(t) = [\exp(iHt/\hbar)A(0)\ \exp -(iHt/\hbar)]_{mn} \tag{7}$$

$$\begin{aligned} \frac{d}{dt}A(t) &= \frac{iH}{\hbar}\ \exp(iHt/\hbar)A(0)\exp -(iHt/\hbar) \\ &\quad - \exp(iHt/\hbar)A(0)\exp -(iHt/\hbar)\frac{iH}{\hbar} \\ &= \frac{1}{i\hbar}\,[A(t)H - HA(t)] \end{aligned} \tag{8}$$

It can be proved that (i) operators can consistently be replaced by equivalent matrices, defined by Eq. (2); and (ii) the Heisenberg time-dependent matrix formulation of quantum theory is completely equivalent to the Schrödinger time-dependent wave function formulation.

Edward Gerjuoy

Matrix theory

The study of matrices and their properties, and of linear transformations on vector spaces, which can be represented by matrices.

Matrices. A matrix is a rectangular array of numbers, with the numbers that appear in the matrix being called entries. For example, A, given by Eq. (1), is a 2×3 matrix (that is, it has 2 rows and 3 columns). The first row of A is given by notation (2), and the second column of A is given by notation (3).

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \tag{1}$$

$$[1 \quad 2 \quad 3] \tag{2}$$

$$\begin{bmatrix} 2 \\ 5 \end{bmatrix} \tag{3}$$

The entry in the ith row and jth column of a matrix is denoted by a_{ij}. For example, in the matrix A, $a_{21} = 4$ and $a_{12} = 2$.

Representation of rotations. Matrices arise naturally as concrete representations of certain geometric transformations in the cartesian plane. An example of such a transformation is one in which any point v in the plane (v is also called a vector) is rotated counterclockwise through 45° around the origin O to a new vector Tv (see **illus.**). The vectors

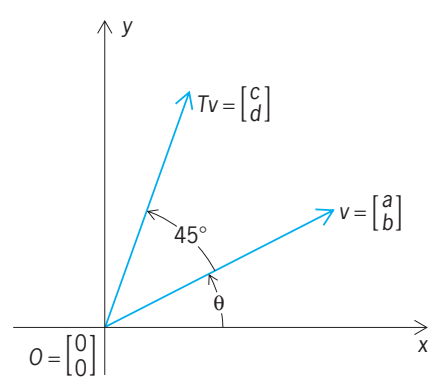

Vector rotation that rotates any vector *v* in the plane counterclockwise through 45° to a new vector *Tv*.

v and Tv can be represented as in Eqs. (4), where

$$v = \begin{bmatrix} a \\ b \end{bmatrix} \qquad Tv = \begin{bmatrix} c \\ d \end{bmatrix} \tag{4}$$

each vector is defined by a pair of numbers: a is the projection of v on the x axis, b is its projection on the y axis, and c and d are the corresponding projections of Tv. If v is a distance r from O and makes an angle θ with the x axis, then the rotated vector Tv is also a distance r from O and makes an angle $\theta + 45^\circ$ with the x axis. Thus, the projections are given by Eqs. (5), and the identities for

$$\begin{aligned} a &= r\cos\theta \\ b &= r\sin\theta \\ c &= r\cos(\theta + 45^\circ) \\ d &= r\sin(\theta + 45^\circ) \end{aligned} \tag{5}$$

the cosine and sine of a sum of angles can be applied to this equation, yielding Eqs. (6). If A is the 2×2 matrix given by Eq. (7), then Eqs. (6) can be conveniently summarized with Eq. (8). Thus, the transformation T is described completely by the matrix A, as in Eq. (8). *See* ANALYTIC GEOMETRY; TRIGONOMETRY.

$$\begin{aligned} c &= r\cos\theta\cos 45^\circ - r\sin\theta\sin 45^\circ \\ &= a\cos 45^\circ - b\sin 45^\circ \\ d &= r\sin\theta\cos 45^\circ + r\cos\theta\sin 45^\circ \\ &= a\sin 45^\circ + b\cos 45^\circ \end{aligned} \tag{6}$$

$$A = \begin{bmatrix} \cos 45^\circ & -\sin 45^\circ \\ \sin 45^\circ & \cos 45^\circ \end{bmatrix} \tag{7}$$

$$\begin{bmatrix} c \\ d \end{bmatrix} = A \begin{bmatrix} a \\ b \end{bmatrix} \tag{8}$$

If S is another counterclockwise rotation through 30°, then it follows similarly that S is described by the matrix B given by Eq. (9).

$$B = \begin{bmatrix} \cos 30^\circ & -\sin 30^\circ \\ \sin 30^\circ & \cos 30^\circ \end{bmatrix} \tag{9}$$

Clearly, the composition of S and T, that is, "T followed by S," denoted by ST, is a counterclockwise rotation through $(45 + 30)^\circ = 75^\circ$. The composition ST has a matrix representation C, which is denoted by Eq. (10). The entries of C can be

$$C = BA \tag{10}$$

conveniently computed directly from the entries of A and B as follows: if A and B are any two 2×2 matrices, given by Eq. (11), then the product of B and A is defined as the matrix in Eq. (12). If the definition in Eq. (12) is applied to the matrices (7) and (9), then Eq. (13) results, and the trigonometric formulas used previously yield Eq. (14). Thus the

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \tag{11}$$

$$BA = \begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{bmatrix} \tag{12}$$

$$BA = \begin{bmatrix} \cos 30^\circ\cos 45^\circ - \sin 30^\circ\sin 45^\circ & \sin 30^\circ\cos 45^\circ + \cos 30^\circ\sin 45^\circ \\ -\cos 30^\circ\sin 45^\circ - \sin 30^\circ\cos 45^\circ & -\sin 30^\circ\sin 45^\circ + \cos 30^\circ\cos 45^\circ \end{bmatrix} \tag{13}$$

$$BA = \begin{bmatrix} \cos 75^\circ & -\sin 75^\circ \\ \sin 75^\circ & \cos 75^\circ \end{bmatrix} \tag{14}$$

definition in Eq. (12) yields the matrix $C = BA$ that represents the composition ST. The matrix C in Eq. (10) is the product of the matrices B and A.

Matrix multiplication. In general, if X is an $m \times n$ matrix and Y is an $n \times q$ matrix, then the product $P = XY$ is an $m \times q$ matrix whose entry in row i and column j is given by Eq. (15) for $i, \ldots, m$ and $j = 1, \ldots, q$.

$$p_{ij} = x_{i1}y_{1j} + x_{i2}y_{2j} + \cdots + x_{in}y_{nj} \tag{15}$$

The number of columns of X must be the same as the number of rows of Y in order that the product matrix P can be defined. Some examples of matrix products are given in Eqs. (16)–(20). The examples in Eqs. (16) and (17) show that, in general, XY and YX need not be the same matrix; that is, matrix multiplication is not always commutative. The examples in Eqs. (18) and (19) show that multiplication of a matrix X by the $n \times n$ matrix with 1 in row i and column i for $i = 1, \ldots, n$, and 0 elsewhere, does not change the matrix X. Finally, the example in Eq. (20) shows that the product of two matrices can have every entry 0 without this being the case for either factor.

$$\begin{aligned} \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 5 & 6 \end{bmatrix} &= \begin{bmatrix} 1\cdot 4 + 2\cdot 5 & 1\cdot 0 + 2\cdot 6 \\ 0\cdot 4 + 3\cdot 5 & 0\cdot 0 + 3\cdot 6 \end{bmatrix} \\ &= \begin{bmatrix} 14 & 12 \\ 15 & 18 \end{bmatrix} \end{aligned} \tag{16}$$

$$\begin{aligned} \begin{bmatrix} 4 & 0 \\ 5 & 6 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} &= \begin{bmatrix} 4\cdot 1 + 0\cdot 0 & 4\cdot 2 + 0\cdot 3 \\ 5\cdot 1 + 6\cdot 0 & 5\cdot 2 + 6\cdot 3 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 8 \\ 5 & 28 \end{bmatrix} \end{aligned} \tag{17}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix} \tag{18}$$

$$\begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{bmatrix} \tag{19}$$

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{20}$$

Vector spaces. Vectors in the plane can be added and multiplied (scaled) by constants. For example, if the vectors u and v and the number c are given by Eq. (21), then $u + v$ is given by Eq. (22) and cu is given by Eq. (23).

$$u = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad v = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \qquad c = -3 \tag{21}$$

$$u + v = \begin{bmatrix} 1+3 \\ 2+4 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix} \tag{22}$$

$$cu = \begin{bmatrix} -3 \cdot 1 \\ -3 \cdot 2 \end{bmatrix} = \begin{bmatrix} -3 \\ -6 \end{bmatrix} \tag{23}$$

Axioms. There are a number of simple axioms that are satisfied by vector addition and scalar multiplication. If u, v, and w are vectors, then [1] Eqs. (24) and (25) are valid; [2] there exists a vector 0 such that Eq. (26) is satisfied; and there exists a unique vector $-u$ such that Eq. (27) is satisfied. If, in addition, c and d are numbers (also called scalars), then Eqs. (28)–(31) are satisfied. A set of vectors V to-

$$u + v = v + u \tag{24}$$

$$u + (v + w) = (u + v) + w \tag{25}$$

$$u + 0 = u \tag{26}$$

$$u + (-u) = 0 \tag{27}$$

$$(cd)\, u = c\,(du) \tag{28}$$

$$1u = u \tag{29}$$

$$c(u + v) = cu + cv \tag{30}$$

$$(c + d)\, u = cu + du \tag{31}$$

gether with a set of numbers F that satisfy these rules is called a vector space over F. The set of numbers F is a field. This means that F has many of the properties associated with rational numbers (fractions), real numbers (decimals), or complex numbers. There are many important mathematical systems that satisfy the preceding axioms for a vector space. Some examples are the two-dimensional space of plane geometry, the three-dimensional space of solid geometry, and polynomials with coefficients from F. *See* FIELD THEORY (MATHEMATICS).

Linear transformations. If V and W are vector spaces over F, then a linear transformation T on V to W is a function that assigns to each vector v in V a unique vector w in W, as in Eq. (32). Moreover, if u and v

$$Tv = w \tag{32}$$

are any vectors in V, and c and d are any scalars in F, then T satisfies Eq. (33). The vector rotation exam-

$$T(cu + dv) = cTu + dTv \tag{33}$$

ples T and S described previously are linear transformations.

A basis of the vector space V is an ordered sequence of vectors $v_1, \ldots, v_n$ in V such that any vector v in V can be written in only one way as in Eq. (34)

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n \tag{34}$$

for appropriate scalars $c_1, \ldots, c_n$. For example, if V is the cartesian plane, then the vectors of notation (35) form a basis, as do the vectors of notation (36).

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{35}$$

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{36}$$

Once bases $v_1, \ldots, v_n$ and $w_1, \ldots, w_m$ for V and W respectively have been chosen, any linear transformation T on V to W can be completely and uniquely described in terms of the entries in an $m \times n$ matrix A by Eq. (37). In other words, each Tv_j, as a vector in

$$Tv_j = a_{1j}w_1 + a_{2j}w_2 + \cdots + a_{mj}w_m \tag{37}$$

$$j = 1, 2, \ldots, n$$

W, can be expressed uniquely as a sum of scalar multiples of the basis $w_1, \ldots, w_m$. The scalars $a_{1j}, \ldots, a_{mj}$ are the entries in column j of A; that is, the entry in row i and column j is a_{ij}. This observation is fundamental because it shows that the study of linear transformations is coextensive with the study of matrices.

Important problems. There are a number of important questions about linear transformations that have been studied extensively, as follows:

1. Given w in W, the problem of determining all vectors v in V (if any) for which Eq. (38) is satisfied.

$$Tv = w \tag{38}$$

2. If $V = W$, the problem of determining all scalars λ and nonzero vectors v for which Eq. (39) is satisfied.

$$Tv = \lambda v \tag{39}$$

3. If $V = W$ and F is any field, the problem of finding a basis of V such that the corresponding matrix representation A for T is as simple as possible.

4. If $V = W$ and F is either the real numbers or the complex numbers, the problem of determining the structure of all T that preserve the lengths of all vectors; that is, for all v, the vectors v and Tv have the same length. This property is possessed by rotations.

The set of all vectors Tv obtained as v varies over V is called the range of T. The range of T is always a vector space. Thus the first problem above is equivalent to determining the range of a linear transformation T. Once bases of V and W have been selected, Eq. (38) is readily seen to be equivalent to solving a system of linear equations of the form of Eqs. (40) for the

determination of the numbers $x_1, \ldots, x_n$.

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{40}$$

Finding the numbers λ and vectors v that satisfy Eq. (39) is called the eigenvalue problem; the number λ in (39) is called an eigenvalue of T, and the corresponding vector v in Eq. (39) is called an eigenvector of T. Equation (39) leads to a system similar to Eqs. (40), but $m = n$ and each b_i must be replaced by λx_i for $i = 1, \ldots, n$. Moreover, both the eigenvalue λ and $x_1, \ldots, x_n$ must be determined. Because of the importance of Eqs. (38) and (39) in many areas of applied mathematics, special-purpose computer programs have been developed to deal with them. *See* EIGENFUNCTION; EIGENVALUE (QUANTUM MECHANICS); LINEAR SYSTEMS OF EQUATIONS.

The problem of finding the basis of a vector space that simplifies the matrix representation of a linear transformation on the vector space to itself was solved in a sequence of papers in the latter part of the nineteenth century and the early part of the twentieth century. *See* LINEAR ALGEBRA.

Results. Because of their applications in the physical and social sciences, certain special classes of linear transformations and matrices have been studied extensively. If A is an $m \times n$ matrix, then the transpose of A is the $n \times m$ matrix whose rows, in succession, are columns of A written as rows. The transpose of A is denoted by A^T. If A has complex number entries, then A^* is the matrix obtained from A^T by replacing each entry by its complex conjugate. An $n \times n$ matrix A is said to be nonsingular (or invertible) if there is a unique $n \times n$ matrix A^{-1} for which $AA^{-1} = A^{-1}A = I_n$; the matrix I_n is the identity matrix: $a_{ii} = 1$ for $i = 1, \ldots, n$, and $a_{ij} = 0$ if $i \neq j$. If $A^* = A$, then A is hermitian; if $A^* = A^{-1}$, then A is unitary; if $A^*A = AA^*$, then A is normal.

An important result is the polar decomposition theorem: if A is an $n \times n$ matrix with complex number entries, then $A = PH$, where P is unitary and H is hermitian with nonnegative real eigenvalues. A triangularization theorem due to I. Schur has many important applications. It states that any $n \times n$ matrix A with complex number entries can be written as $A = P^{-1}BP$, in which P is unitary, the entries b_{ii}, $i = 1, \ldots, n$, are the eigenvalues of A, and $b_{ij} = 0$ if $j > i$. The QR factorization theorem states that any $m \times n$ matrix A, $m \geq n$, can be factored into a product $A = QR$ in which Q is $m \times m$ unitary and R is $m \times n$ upper triangular (that is, $r_{ij} = 0$ if $j > i$). The QR factorization theorem is used to devise practical methods for finding mutually perpendicular vectors of length 1 that comprise a basis of the range of a linear transformation. The singular-value decomposition theorem states that any $m \times n$ matrix A can be factored into a product $A = UDV$, in which U is $m \times m$ unitary, V is $n \times n$ unitary, and D is an $m \times n$ matrix whose only (possibly) nonzero entries are $d_{11} \geq \ldots > d_{pp} \geq 0$, where p is the smaller of m and n. The numbers $d_{11}, \ldots, d_{pp}$ are called the singular values of A. The singular-value decomposition theorem is used extensively in solving least-squares problems.

Matrices whose entries are nonnegative real numbers are important in many applications of matrix theory to problems in the social sciences. The principal result about an $n \times n$ matrix A with positive entries is the Perron-Frobenius theorem: A possesses a unique positive eigenvalue λ and a corresponding eigenvector x with positive entries; that is, $Ax = \lambda x$. Moreover, any other eigenvalue of A is smaller in absolute value than λ.

Matrices with polynomial entries have been studied extensively because of their importance in control theory, systems theory, and other areas of applied mathematics and engineering.

If A is an $n \times n$ complex matrix and x is an $n \times 1$ matrix, then x^*Ax is a scalar (that is, a 1×1 matrix) called the hermitian form associated with A. As x runs over all $n \times 1$ matrices satisfying $x^*x = 1$, the corresponding set of numbers x^*Ax fills up a region in the complex plane called the numerical range or field of values of A, denoted by $W(A)$. The set $W(A)$ is an inclusion region for the eigenvalues of A; that is, every eigenvalue λ of A is contained in $W(A)$. Because of this fact, the set $W(A)$ has been extensively studied ever since the important discovery in 1918 and 1919 by F. Hausdorff and O. Toeplitz that $W(A)$ is always a convex set in the plane. A subset of the plane is convex if the line segment joining any two points in the subset is itself entirely contained in the subset.

Marvin Marcus

Bibliography. W. C. Brown, *Matrices and Vector Spaces*, 1991; I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, 1982; P. R. Halmos, *Finite-Dimensional Vector Spaces*, 2d ed., 1958, reprint 1993; R. A. Horn and C. R. Johnson, *Matrix Analysis*, 1985, paper 1990; D. Lewis, *Matrix Theory*, 1991; M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, 1964, reprint 1992; H. Minc, *Nonnegative Matrices*, 1988; D. J. Winter, *Matrix Algebra*, 1991.

Matter (physics)

A term that traditionally refers to the substance of which all bodies consist. In Aristotelian physics, each type (species) of material body had a distinct "essential form." Early modern scientists, however, asserted that there is one universal type of matter. For Isaac Newton, this matter consisted of "solid, massy, hard, impenetrable, movable," ultimate particles ("atoms"), which were discrete, localized, indivisible bodies. Modern analyses distinguish two types of mass in classical (Newtonian) matter: inertial mass, by which matter retains its state of rest or uniform rectilinear motion in the absence of external forces; and gravitational mass, by which a body exerts forces of attraction on other bodies and by which

it reacts to those forces. Expressed in appropriate units, these two properties are numerically equal—a purely experimental fact, unexplained by theory. Albert Einstein made the equality of inertial and gravitational mass a fundamental principle (principle of equivalence), as one of the two postulates of the theory of general relativity. In the equation $E = mc^2$ (where c is the velocity of light), Einstein recognized the equivalence (interconvertibility) of mass (m) and energy (E), which had been distinguished in classical theories. *See* GRAVITATION; INERTIA; MASS; RELATIVITY; WEIGHT.

Implications of quantum mechanics. Among the departures from classical thinking instituted by quantum mechanics with regard to matter are the following. Matter in classical mechanics is closely identified with mass; in quantum mechanics, mass is only one among many properties (quantum numbers) that a particle can have, for example, electric charge, spin, and parity. The nearest quantum-mechanical analogs of traditional matter are fermions, having half-integral values of spin. Forces are mediated by exchange of bosons, particles having integral spins (such as photons). Fermions correspond to classical matter in exhibiting impenetrability (a consequence of the exclusion principle), but the correspondence is only rough. For example, fermions can also be exchanged in interactions (a photon and an electron can exchange an electron), and they also exhibit wavelike behavior. States of classical matter-particles were given by their positions and momenta, but in quantum mechanics it is impossible to assign simultaneous precise positions and momenta to particles. Finally, two or more particles in quantum theory can become "entangled," so that, in the case of two particles that become entangled and separate by a very great distance, an operation on one particle will be immediately responded to by a corresponding change in the other. This "nonlocality" violates traditional notions of what a piece of matter must be, which led Einstein to refer to quantum mechanics as objectionable in implying a "spooky action at a distance." *See* EXCLUSION PRINCIPLE; NONRELATIVISTIC QUANTUM THEORY; QUANTUM ELECTRODYNAMICS; QUANTUM MECHANICS; QUANTUM STATISTICS.

Matter in the universe. Beginning in the 1920s, astronomers accumulated a vast amount of evidence leading to the conclusion that the density of matter in the universe was much larger than that of visible matter. Important steps in this direction were taken in 1933, when Fritz Zwicky found that the amount of visible matter in a number of galaxy clusters was far too small to hold the galaxies in the cluster, and in 1983, when Vera Rubin and colleagues found that the rotation curves of individual spiral galaxies implied that the galaxies must be surrounded by a sphere of unseen matter, ten times the amount in the visible galaxy itself. It began to appear that the visible matter might constitute less than 1%, or at most 3%, of the amount of matter required, in the simplest theories, to equal the critical density at which the geometry of the universe would be Euclidean ("flat"), and at which the expansion rate of the universe would gradually slow down, asymptotically approaching zero. Some scientists accepted the implication that the universe was "open," with a hyperbolic geometry and continued expansion forever at a rate that did not approach zero.

Others proposed that missing matter, undetectable by current techniques, might equal the critical density, and many proposals were made regarding the nature of the missing matter. The motivation for seeking the missing matter became stronger with the proposal of the theory of inflation, which required a critical density. However, visible matter is baryonic, consisting almost entirely by mass of protons and neutrons. With the total density of baryonic matter so low, the missing nonbaryonic dark matter would have to be some kind of exotic particle. Candidates for this status included axions and the lightest (nondecaying) supersymmetric particles, photinos or perhaps winos. None of these, nor many other candidates, have been observed. *See* DARK MATTER; SUPERSYMMETRY; WEAKLY INTERACTING MASSIVE PARTICLE (WIMP).

It was long realized that such exotic matter was still not enough for closure (critical density), however, and debates arose regarding the nature of the dark energy that was needed to provide the remainder. One popular proposal was that of quintessence, a negative-pressure energy field, perhaps created by particles that had not been suggested in current hypotheses. However, while quintessence theories have not been entirely ruled out, they have been replaced by a different approach. In 1998, results from studies of distant supernovae indicated strongly that the universe is not merely expanding: the rate of expansion is accelerating. It was found that this acceleration has been going on for approximately 5×10^9 years. The new view thus suggested was that some kind of dark antigravitational energy was at work, a product of the vacuum rather than, as with quintessence views, some unknown form of quasi-matter. It thus resembles the cosmological constant that Einstein had proposed as a term in his equations designed to preserve the universe from the possibility that it might expand or contract. He abandoned this term after the discovery of the expansion of the universe. *See* ACCELERATING UNIVERSE; COSMOLOGICAL CONSTANT; DARK ENERGY.

In 2003, one of the great achievements of the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) was to pin down the distribution as follows: 4% of the matter in the universe is baryonic, 23% is dark nonbaryonic matter, and 73% is something else, generally known as dark energy. Though slightly different values have been proposed, all are within an estimated 5% error range. The *WMAP* results were bolstered by ones from the Sloan Digital Sky Survey (SSDS), also announced in 2003. Several future investigations, expanding and improving on these results, are in preparation. *See* COSMOLOGY; SLOAN DIGITAL SKY SURVEY; UNIVERSE; WILKINSON MICROWAVE ANISOTROPY PROBE. Dudley Shapere

Bibliography. N. S. Hetherington (ed.), *Encyclopedia of Cosmology*, Garland Publishing, New York, 1993; M. Jammer, *Concepts of Mass in Classical*

and Modern Physics, 1961, reprint 1993; L. Krauss, *Quintessence: The Mystery of Missing Mass in the Universe*, 2000; E. McMullin, *The Concept of Matter*, 1963; H. C. Ohanian and R. Ruffini, *Gravitation and Spacetime*, 2d ed., 1994; U. Seljak et al., Cosmological parameter analysis including SDSS Lyalpha forest and galaxy bias: Constraints on the primordial spectrum of fluctuations, neutrino mass, and dark energy, *Phys. Rev. D*, 71:103515, 2005; Special issue: The dark side, *Science*, 300:1893–1918, 2003; D. N. Spergel et al., First-year *Wilkinson Microwave Anisotropy Probe* (*WMAP*) observations: Determination of cosmological parameters, *Astrophys. J. Suppl. Ser.*, 148(1):175–194, 2003.

Matthiessen's rule

An empirical rule which states that the total resistivity of a crystalline metallic specimen is the sum of the resistivity due to thermal agitation of the metal ions of the lattice and the resistivity due to the presence of imperfections in the crystal. This rule is a basis for understanding the resistivity behavior of metals and alloys at low temperatures.

The resistivity of a metal results from the scattering of conduction electrons. Lattice vibrations scatter electrons because the vibrations distort the crystal. Imperfections such as impurity atoms, interstitials, dislocations, and grain boundaries scatter conduction electrons because in their immediate vicinity the electrostatic potential differs from that of the perfect crystal. If lattice scattering and imperfection scattering are independent processes and are isotropic, it can be shown that Eq. (1) holds, where

$$\rho(T) = \rho_L(T) + \rho_1 \tag{1}$$

$\rho(T)$ is the resistivity at temperature T, $\rho_L(T)$ is the resistivity due to lattice scattering (the ideal resistivity, which is temperature-dependent), and ρ_1 is the so-called residual resistivity due to imperfections, which is presumably independent of temperature. An equivalent statement is given in Eq. (2).

$$\frac{\partial \rho(T)}{\partial T} = \frac{\partial \rho_L(T)}{\partial T} \tag{2}$$

As exemplified by the **illustration**, Matthiessen's rule is generally obeyed, although some deviations do occur. Such deviations arise for the following reasons: (1) The introduction of impurities generally alters the elastic properties, influences the lattice vibration spectrum, and thereby changes $\rho_L(T)$. (2) Some imperfections, such as dislocations, do not scatter isotropically. (3) Thermal expansion lowers the Fermi energy (since impurity scattering depends on the velocity of the electrons at the Fermi energy, the resistivity due to impurities depends also on the temperature through the thermal expansion). (4) The impurity ions also participate in the thermal vibrations of the crystal lattice. However, since the scattering potential of an impurity differs from that of a solvent ion, lattice scattering from the former also differs from that due to a solvent ion. Frank J. Blatt

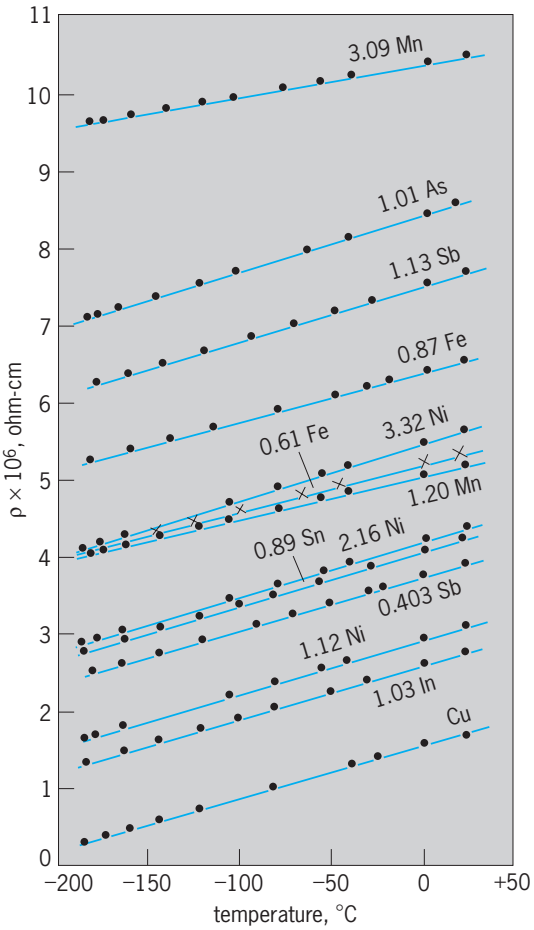

Resistivity of copper and copper alloys. Solute concentrations are given in atomic percent. °F = (°C × 1.8) + 32. (*After J. O. Linde*)

Maxillopoda

A class of Crustacea whose application is gaining moderately widespread use, despite the fact that its validity is not universally accepted by researchers. The constituent subclasses of the Maxillopoda also remain unsettled: before 1956, all major subclasses of the Crustacea were regarded as distinct taxa of equal rank. In addition to the Malacostraca, these included the Cephalocarida, Branchiopoda, Cirripedia, Copepoda, Mystacocarida, Branchiura, and Ostracoda. Use of cladistics, or phylogenetic analysis, has rekindled interest in phylogenetic relationships and hierarchy among the nonmalacostracan taxa.

The class Maxillopoda was proposed for those taxa with six thoracic somites (with some exceptions), a well-developed mandibular palp in adults, well-developed maxillules and maxillae adapted for filtering in filter feeders, and the lack of gnathobases on the appendages of the thorax. The Recent taxa included were the Copepoda, Branchiura, Mystacocarida, and provisionally, the Cirripedia. Subsequently the Cirripedia (subdivided into Cirripedia sensu stricto and Ascothoracida) were unequivocally incorporated. The Ostracoda are now included by some but excluded by others. More recent suggestions have included the Tantulocarida,

and even the probably noncrustacean Pentastomida, within the Maxillopoda. *See* BRANCHIURA; CIRRIPEDIA; COPEPODA; OSTRACODA; PENTASTOMIDA; TANTULOCARIDA.

In addition to the original characters used to diagnose the Maxillopoda, tagmosis (division of the body into five cephalic, six or seven thoracic, and five or four abdominal somites), gonopore position, form and structure of thoracic appendages, evidence from comparative spermatology, and the structure of the nauplius eye have been used to argue for the homogeneity of this class. Opponents have argued that the body segmentation presumed to be common to all subclasses had been misinterpreted and that heterogeneity is apparent during several developmental processes. When morphological characters of members of the various maxillopodan taxa are subjected to cladistic analyses, the results are dependent upon the investigator's interpretations of homologies as opposed to convergences, and primitive versus advanced character states, as well as the number of characters utilized. Whether the Maxillopoda is a monophyletic and meaningful taxon remains to be determined. *See* CRUSTACEA. Patsy A. McLaughlin

Bibliography. G. A. Boxshall and R. Huys, New tantulocarid, *Stygotantulus stocki*, parasitic on harpacticoid copepods, with an analysis of the phylogenetic relationships within the Maxillopoda, *J. Crust. Biol.*, vol. 9, 1989; F. R. Schram, *Crustacea*, 1986; F. R. Schram (ed.), Crustacean phylogeny, *Crust. Iss.*, vol. 1, 1983.

Maxwell's demon

An imaginary being whose action appears to contradict the second law of thermodynamics. The properties of bulk systems are explained in terms of contributions from their constituent molecules. The same, presumably, should be true of certain grand properties of nature, and in particular the second law of thermodynamics, which identifies the natural direction of change with the direction of increasing entropy. The statistical basis of this law, in which entropy is identified with molecular chaos, was established largely through the work of L. Boltzmann. However, there has always been a certain degree of discomfort associated with the acceptance of the law, particularly in relation to the time reversibility of physical laws and the role of molecular fluctuations. In 1867, J. C. Maxwell considered, in this connection, the action of "a finite being who knows the paths and velocities of all the molecules by inspection." This being was later referred to as a demon by Lord Kelvin, and the usage has been generally adopted. *See* ENTROPY; THERMODYNAMIC PRINCIPLES; TIME, ARROW OF.

Trapdoor model. The activity of Maxwell's demon can be modeled by a trapdoor in a partition between two regions full of gas at the same pressure and temperature. The trapdoor needs to be restrained by a light spring to ensure that it is closed unless it is struck by molecules traveling from the left (see **illus.**). Its hinging is such that molecules traveling

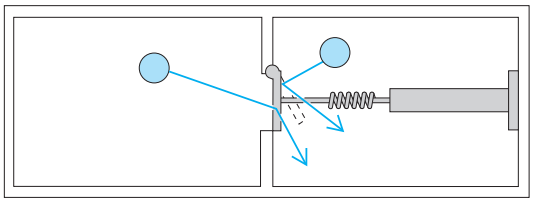

Type of device that emulates mechanically the actions of Maxwell's demon. Molecules traveling to the right can open the trapdoor and enter the right-hand compartment, but those striking it from the right cannot open it, so do not move into the left-hand compartment.

from the right cannot open it. The essential point of Maxwell's vision was that molecules striking the trapdoor from the left would be able to penetrate into the right-hand region but those present on the right would not be able to escape back into the left-hand region. Therefore, the initial equilibrium state of the two regions, that of equal pressures, would be slowly replaced by a state in which the two regions acquired different pressures as molecules accumulated in the right-hand region at the expense of the left-hand region. Only a slightly more elaborate mechanical arrangement is needed to change the apparatus to one in which the temperatures of the two regions move apart. For this to be achieved, the trapdoor needs to admit faster molecules, but not slower molecules, into the right-hand region. In each case, the demonic trapdoor appears to be contriving a change that is contrary to the second law, for an implication of that law is that systems in either mechanical equilibrium (at the same pressure) or thermal equilibrium (at the same temperature) cannot spontaneously diverge from equilibrium.

Resolution of paradox. As frequently occurs in science, the resolution of a paradox or the elimination of an apparent conflict with a firmly based law depends on a detailed analysis of the proposed arrangement. Numerous analyses of this kind have shown that the activities of Maxwell's demon do not in fact result in the overthrow of the second law. An indication of the thrust of these discussions can be obtained by noting that for the trapdoor to act as specified, it must close immediately after it has allowed a molecule to pass, for otherwise the "wrong" molecules will take advantage of its opening. However, the act of opening the door gives the door a certain additional energy, and it is possible for a molecule in the right-hand region to be struck by the door and batted into the left-hand region. Thus, the demonic door undoes what it achieved when it admitted a molecule, and the net effect is the maintenance of mechanical equilibrium. Similar remarks apply to the maintenance of thermal equilibrium.

It is sometimes argued that a true demon, one endowed with intelligence, would elude this argument. However, there are several rebuttals of this claim. One is that a truly intelligent demon would use its intelligence to construct a mechanical device that is equivalent to the trapdoor: hence, the same argument would apply, and the exercise of intelligence would not circumvent the second law. Even if the demon resorts to more elaborate inspection techniques, it fails to circumvent the law. For instance,

if it uses electromagnetic radiation to monitor the approach of candidate molecules for admission to the right-hand region, it needs to unleash and detect at least one photon. However, a photon carries energy, so its emission and detection involve the dissipation of energy and hence a contribution to the entropy. P. W. Atkins

Bibliography. K. G. Denbigh and J. S. Denbigh, *Entropy in Relation to Incomplete Knowledge*, 1985; R. P. Feynman, *Lectures in Physics*, 1966; E. Garber, S. G. Brush, and C. W. F. Everitt (eds.), *Maxwell on Molecules and Gases*, 1985; H. S. Leff and A. F. Rex (eds.), *Maxwell's Demon: Entropy, Information, and Computing*, 1990.

Maxwell's equations

The four equations of electromagnetism completed by J. C. Maxwell in 1864. The equations govern a wide range of phenomena, including much of electronic technology. However, some modern technology requires the quantum-mechanical interpretation of the equations. In their most general form, Eqs. (1)–(4) relate four field quantities, the electric field **E**,

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{1}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \tag{2}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{3}$$

$$\nabla \cdot \mathbf{D} = \rho \tag{4}$$

magnetic induction **B**, magnetic field **H**, and the electric displacement **D**, to the free charge and current densities ρ and **J**. Two auxiliary equations are needed. Microscopic displacements between bound positive and negative charges produce fields on a macroscopic scale when the displacements are the same throughout a macroscopic volume of material. Equation (5) relates **D** to **E** and the polariza-

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \tag{5}$$

tion **P**, which results from the bound charges within a material. The constant ϵ_0 is called the permittivity of free space, and its value is closely approximated by $10^{-11}/(36\pi)$ in SI (meter-kilogram-second-ampere or MKSA) units. The polarization can be a function of **E**, or it can be a spontaneous property of the material. In the latter case, the material is said to be ferroelectric. A material may also contain currents that arise in the atomic and nuclear structure of matter and do not flow over macroscopic distances. The other auxiliary equation (6) relates **B**

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M}) \tag{6}$$

to **H** and the magnetization **M**, which results from the bound currents. The constant μ_0 is called the permeability of free space, defined to be $4\pi(10)^{-7}$ in SI units. The magnetization can be a function of **H** or a spontaneous property in a ferromagnetic material. *See* CURRENT DENSITY; DIFFERENTIAL EQUATION; ELECTRIC CHARGE; ELECTRIC FIELD; FERROELECTRICS; FERROMAGNETISM; MAGNETIZATION; POLARIZATION OF DIELECTRICS.

Lorentz force law. When no polarizable or magnetizable materials are being considered, Maxwell's equations can be solved for the fields **E** and **B** in terms of the sources ρ and **J**. When the fields are known, the motion of charged particles may be found with the help of Eq. (7), which gives the force on a charge

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{7}$$

q moving with velocity **v** as written by H. A. Lorentz. The Lorentz force law depends only on **E** and **B**. For this reason, the other two fields **D** and **H** are not considered fundamental, and in fact Maxwell's equations can be formulated without them. In this more fundamental formulation, there is only one constant in the equations instead of the two ϵ_0 and μ_0. Introducing a second constant, however, allows familiar units to be used for both types of fields. *See* ELECTRICAL UNITS AND STANDARDS.

Divergence and curl. Maxwell's equations can be manipulated easily when the various uses of the gradient operator are known. If $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$ are defined to be unit vectors along the x, y, and z coordinate axes, then the gradient of a scalar function $\phi(x,y,z)$ is compactly written as in Eq. (8). Equation (9) de-

$$\nabla\phi = \hat{\mathbf{i}}\,\frac{\partial\phi}{\partial x} + \hat{\mathbf{j}}\,\frac{\partial\phi}{\partial y} + \hat{\mathbf{k}}\,\frac{\partial\phi}{\partial z} \tag{8}$$

$$\nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \tag{9}$$

fines the dot product of the gradient operator with a vector function of position $\mathbf{A}(x,y,z)$. It is called the divergence of **A**. The cross product of the gradient with **A**, Eq. (10), is called the curl of **A**. The curl of a

$$\nabla \times \mathbf{A} = \hat{\mathbf{i}}\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right) + \hat{\mathbf{j}}\left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right) + \hat{\mathbf{k}}\left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right) \tag{10}$$

curl simplifies to terms involving the gradient, divergence, and laplacian, as shown in Eq. (11). Finally,

$$\nabla \times (\nabla \times \mathbf{A}) = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A} \tag{11}$$

the laplacian, which is defined by Eq. (12), can be

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{12}$$

applied to both scalar and vector functions. A typical use of these rules with Maxwell's equations is to derive a wave equation that can be solved by familiar techniques. *See* CALCULUS OF VECTORS; GRADIENT OF A SCALAR; LAPLACIAN.

Coulomb's law. When there is no time dependence, the four equations uncouple into two for the electric field and two for the magnetic field. Also, some important and instructive special cases of Maxwell's equations can be solved by using symmetry to avoid tedious integrations. The electric field of a charged body with spherical symmetry (**Fig. 1**) can be found by integrating the fourth of Maxwell's

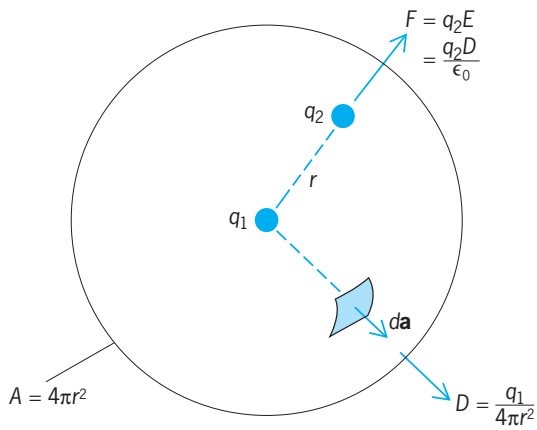

Fig. 1. Electric field of a spherically symmetric charge. Symbols are explained in text.

equations, yielding Eq. (13*a*). The left-hand side of

$$\int d^3x\rho = \int d^3x\Delta.\mathbf{D} \qquad (13a)$$

$$q = \int_{\circ\rightarrow} d\mathbf{a}\cdot\mathbf{D} \qquad (13b)$$

$$q = 4\pi r^2\epsilon_0 E(r) \qquad (13c)$$

this equation gives the total charge q, and the right-hand side is the integral of the divergence of **D** over some volume that includes the charge. By a theorem proved by K. F. Gauss, a volume integral of the divergence of a field can be replaced by an integral of the outward component of the field on the surface that encloses the volume. Thus, Eq. (13*b*) follows, where the closed surface is represented by the circle, and the arrow perpendicular to it represents the direction of the vector element of area $d\mathbf{a}$ perpendicular to the surface. If the closed surface is chosen to be a sphere of radius r, centered on the charge, the outward component of the field is constant on the sphere by symmetry. Therefore the integral reduces to the area of the sphere, $4\pi r^2$, times $\epsilon_0 E(r)$, yielding Eq. (13*c*). The force exerted by one charge on a second charge can then be found from the first term in the Lorentz force law. C. A. de Coulomb first determined experimentally that the force, as shown in Eq. (14), is proportional to both charges, inversely

$$F = \frac{q_1q_2}{4\pi\epsilon_0 r^2} \qquad (14)$$

proportional to the square of the distance between them, and attractive when the charges are of opposite sign. *See* COULOMB'S LAW; ELECTROSTATICS; GAUSS' THEOREM.

Ampére's law. The magnetic field of a cylindrically symmetric, time-independent current distribution can also be found by symmetry. This is done by integrating the second Maxwell equation (**Fig. 2**), yielding Eq. (15*a*). The integral of the outward component of **J** over some surface gives the total current I intersecting that surface. The closed path that bounds the surface is represented by a circle, and the arrow perpendicular to the surface represents the direction of the vector element of area $d\mathbf{a}$. The curl of **H** appears on the right-hand side of Eq. (15*a*). The integral of the outward component of the curl of a vector field over a surface was simplified by G. G. Stokes. Stokes' theorem allows the surface integral of the curl of a vector field to be replaced by an integral of the tangential component of the field around the closed curve bounding the surface, yielding Eq. (15*b*). The right-hand rule is used to determine the direction of $d\mathbf{a}$ (the thumb) relative to the direction of the line integral (the fingers). This directional line integral is represented on the integral sign by a circle with an arrow head embedded in it. *See* STOKES' THEOREM.

$$\int_{\circlearrowleft} d\mathbf{a}\cdot\mathbf{J} = \int_{\circlearrowleft} d\mathbf{a}\cdot\nabla\times\mathbf{H} \qquad (15a)$$

$$I = \oint d\ell\cdot\mathbf{H} \qquad (15b)$$

$$I = \frac{2\pi r}{\mu_0}B(r) \qquad (15c)$$

Since the current distribution is symmetric, rotating it about its axis would not change the field. The magnetic field is therefore tangential in direction and constant in magnitude on any circle centered on the current. Therefore, in the case of such a circular closed path of radius r, the integral reduces to the circumference of the circle, $2\pi r$, times $H(r)$, so that Eq. (15*c*) holds. The second term in the Lorentz force law, Eq. (7), gives the force exerted by the current on a charge q moving with speed v along a parallel line a distance r away, Eq. (16). This force is attractive when the charge is positive and moving in the same direction as the current. When a number of particles are distributed along the line to make up a second current I_2, the charge per unit length is I_2/v. Thus the force per unit length is proportional to both currents and inversely proportional to the distance between them. Equation (17), discovered by A. M. Ampére, is a useful special case of the general rule for the force between two currents. *See* AMPÈRE'S LAW.

$$f = qv\frac{\mu_0 I}{2\pi r} \qquad (16)$$

$$\frac{df}{dl} = \frac{\mu_0 I_1 I_2}{2\pi r} \qquad (17)$$

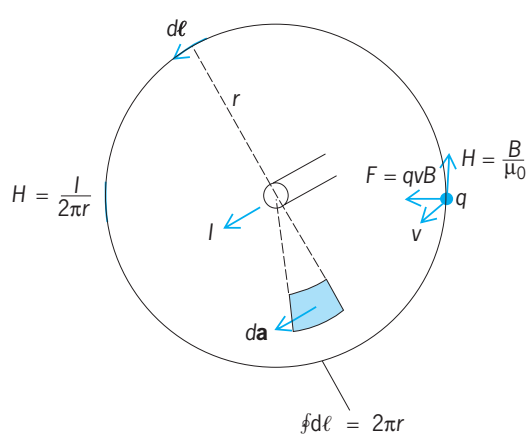

Fig. 2. Magnetic field of a cylindrically symmetric current. Symbols are explained in text.

Faraday's law and Ohm's law. The quantity **B** is also called the magnetic flux density. The reason can be explained by integrating the first Maxwell equation over some surface, yielding Eq. (18*a*). When Stokes'

$$\int_{\mathcal{S}} d\mathbf{a} \cdot \nabla \times \mathbf{E} == \int_{\mathcal{S}} d\mathbf{a} \cdot \frac{\partial \mathbf{B}}{\partial t} \tag{18a}$$

$$\oint dl \cdot \mathbf{E} = -\frac{dF}{dt} \tag{18b}$$

theorem is used, the left-hand side of this equation becomes the line integral of **E** around the path bounding the surface. This line integral known as the electromotive force (emf) is also equal to the work done on a unit charge that moves around the path. The quantity

$$\int_{\mathcal{S}} d\mathbf{a} \cdot \mathbf{B}$$

is called the flux F through the surface. The order of the time derivative and the integration is interchanged to show that the right-hand side of Eq. (18*a*) is the time rate of change of the flux intersecting the path, so that Eq. (18*b*) holds. This relationship between the electromotive force and rate of change of the flux was discovered by M. Faraday and is called electromagnetic induction. *See* ELECTROMAGNETIC INDUCTION; ELECTROMOTIVE FORCE (EMF); FARADAY'S LAW OF INDUCTION.

If the path around the surface in question were traced out by a conducting material, current would flow in response to the electromotive force. Often the current density is linearly related to the electric field, as in Eq. (19), and the coefficient σ is a constant

$$\mathbf{J} = \sigma \mathbf{E} \tag{19}$$

called the conductivity of the material. Integration of this equation around the path yields Eq. (20*a*).

$$\sigma \oint dl \cdot \mathbf{E} = \oint \mathbf{J} \cdot dl \tag{20a}$$

$$V = \frac{IL}{A\sigma} = RI \tag{20b}$$

If it is assumed, for simplicity, that the current density is constant over the cross section of the wire, then the path integral in Eq. (20*a*) relates the current I in the wire to the electromotive force V, as in Eq. (20*b*). The proportionality constant R is called the resistance of the path; it is proportional to the length L and inversely proportional to the cross section A. The unit of resistance is named the ohm, after G. S. Ohm, who formulated the relationship between I and V. *See* ELECTRICAL CONDUCTIVITY OF METALS; ELECTRICAL RESISTANCE; OHM'S LAW.

Continuity equation. By taking the divergence of the second Maxwell equation and the time derivative of the fourth, Eq. (21) is derived. This is called

$$\begin{aligned} 0 = \nabla \cdot (\nabla \times \mathbf{H}) &= \nabla \cdot \mathbf{J} + \nabla \cdot \frac{\partial \mathbf{D}}{\partial \mathbf{t}} \\ &= \nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} \end{aligned} \tag{21}$$

the continuity equation and shows that charge is conserved. This can be shown by integrating over some volume, as in Eq. (22). One term then becomes the

$$0 = \int_{\circ\rightarrow} \mathbf{J} \cdot d\mathbf{a} + \frac{d}{dt} \int d^3x\rho \tag{22}$$

time rate of change of the net charge in that volume. Gauss' theorem can be used to express the other term as the surface integral of the outward component of the current density. The latter is the net current flowing outward through the surface. Thus the continuity equation states that the quantity of charge present in a region can change only through the transport of charge across its boundaries. *See* EQUATION OF CONTINUITY.

Wave equation. Maxwell completed his equations by adding the displacement current, $\partial \mathbf{D}/\partial t$, to the second equation. As was shown above, this was necessary for charge conservation, but it had other far-reaching consequences. Maxwell's equations can be used to calculate the forces between charges and currents, but they also have nontrivial solutions in free space in the absence of sources. This can be seen by taking the curl of the first equation and the time derivative of the second. Indeed, taking the curl of Eq. (1) gives Eq. (23*a*),and applying Eq. (11) to the

$$\nabla \times (\nabla \times \mathbf{E}) = -\nabla \times \frac{\partial \mathbf{B}}{\partial t} \tag{23a}$$

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \tag{23b}$$

$$\left(\nabla^2 - \mu_0 \epsilon_0 \frac{\partial^2}{\partial t^2}\right) \mathbf{E} = 0 \tag{23c}$$

left-hand side of this equation and the time derivative of Eq. (2) to the right-hand side yields Eq. (23*b*). After setting the divergence of the electric field to zero and rearranging terms, Eq. (23*c*), called the wave equation, results. The wave equation contains only one parameter, and is the same in the fundamental units used by Maxwell. This equation has solutions that propagate with speed $1/\sqrt{\mu_0 \epsilon_0}$, which turned out to be numerically equal to the speed of light c, approximately 3×10^8 m/s. Light was identified as an electromagnetic wave, and the theories of electromagnetism and optics were thereby unified. *See* DISPLACEMENT CURRENT; ELECTROMAGNETIC RADIATION; LIGHT; WAVE EQUATION; WAVE MOTION.

Relativity. Maxwell's equations possess wavelike solutions that propagate at the speed of light. However, it was not immediately clear in what coordinate systems the equations should be applied. Two plausible alternatives were that the speed of light should be equal to the constant c (1) relative to the source and (2) relative to some hypothetical medium called ether that extends throughout all space. Experimental refutations of both alternatives were made. A. Einstein based his special theory of relativity on the idea that the speed of light should have the same value when measured by any observer who is not accelerating. Einstein generalized this to the principle of relativity, which states that all natural laws should have

the same form in all nonaccelerating coordinate systems. In Maxwell's equations, c appears as a parameter (which is more obvious in the fundamental formulation mentioned above). It turned out that Maxwell's equations also had the correct form required by the principle of relativity. However, the rules by which the fields transform from one coordinate system to another had to be worked out. *See* RELATIVISTIC ELECTRODYNAMICS; RELATIVITY. T. J. Nelson

Bibliography. R. H. Good, Jr., and T. J. Nelson, *Classical Theory of Electric and Magnetic Fields*, 1971; J. Hendry, *James Clerk Maxwell and the Theory of the Electromagnetic Field*, 1986; J. D. Jackson, *Classical Electrodynamics*, 3d ed., 1998; J. A. Kong, *Electromagnetic Wave Theory*, 3d ed., 2000; P. Lorrain, D. R. Corson, and F. Lorrain, *Electromagnetic Fields and Waves*, 3d ed., 1988; H. P. Neff, Jr., *Basic Electromagnetic Fields*, 2d ed., 1986.

McLeod gage

A type of instrument used to measure vacuum by application of the principle of Boyle's law. *See* BOYLE'S LAW.

A known volume of a gas whose pressure is to be measured is trapped by raising the level of a fluid (mercury or oil) by means of a plunger, by lifting a reservoir, by using pressure, or by tipping the apparatus. As the fluid level is further raised, the gas is compressed into the capillary tube (see **illus.**). Obeying Boyle's law, the compressed gas now exerts enough pressure to support a column of fluid high enough to read. Readings are somewhat independent of the composition of the gas under pressure.

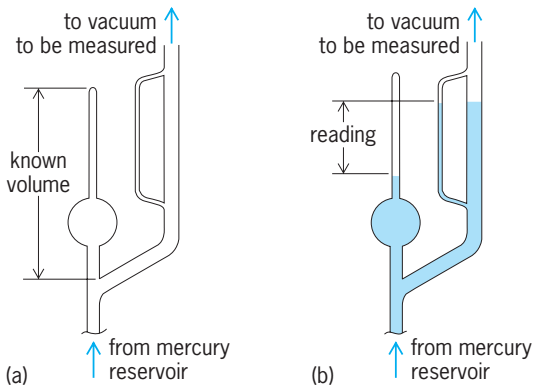

McLeod gage. (*a*) Filling (charging) position. (*b*) Measuring position.

The McLeod gage is simple and inexpensive. It is widely used as a standard of pressure between 10 torr (10^3 pascals) and 10^{-5} torr (10^{-3} pascal) and is often used to calibrate other vacuum instruments. Some disadvantages of this gage are that readings are not continuous, the operator must be skilled in taking visual readings, and mercury or oil vapor may diffuse into the gas being measured. Condensable gases cannot be measured with a McLeod gage, although use of a sorption trap or cold trap may allow some measurements. *See* VACUUM MEASUREMENT. Richard Comeau

Bibliography. T. A. Delchar, *Vacuum Physics and Technology*, 1993; J. F. O'Hanlon, *A User's Guide to Vacuum Technology*, 2d ed., 1989; A. Roth, *Vacuum Technology*, 3d ed., 1990, reprint 1998.

Mean effective pressure

A term commonly used in the evaluation for positive-displacement machinery performance which expresses the average net pressure difference in pounds per square inch (lb/in.^{2} or psi) on the two sides of the piston in engines, pumps, and compressors. It is also known as mean pressure and is abbreviated mep or mp.

In an engine (prime mover) it is the average pressure which urges the piston forward on its stroke. In a pump or compressor it is the average pressure which must be overcome, through the driver, to move the piston against the fluid resistance.

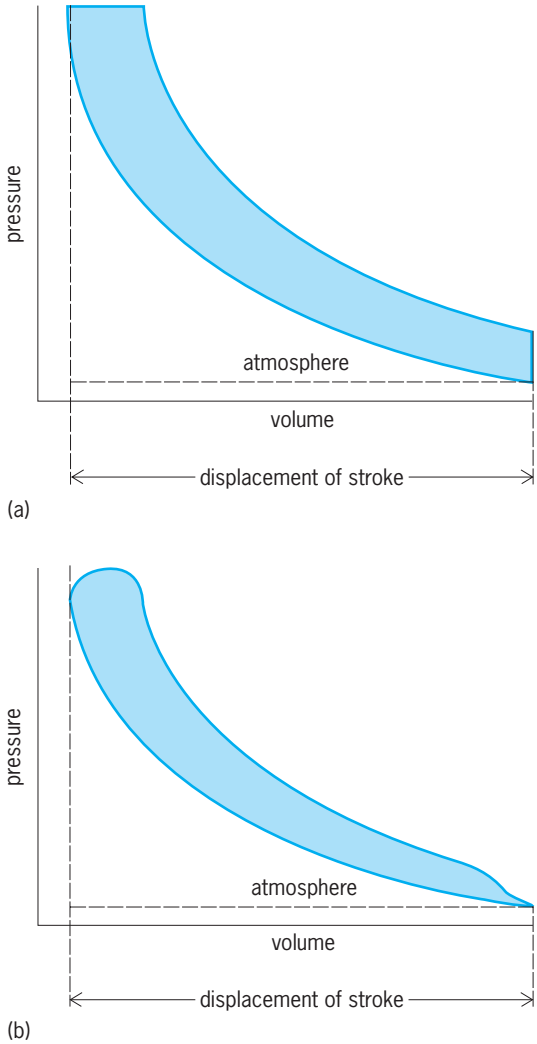

Pressure-volume diagrams (indicator cards) for diesel engine. (*a*) Theoretical (ideal) conditions. (*b*) Actual two-cycle engine card. The colored areas represent the effective work of the cycle.

Some representative values of mean effective pressures at full load

Unit	Value, lb/in.^{2}
Aircraft engine	100–200
Automobile engine	100–140
Diesel engine	75–150
Outboard engine	50–100
Steam engine	50–150
Air compressor	50–100

It can be a theoretic value obtained from the pressure-volume diagram of the thermodynamic or fluid dynamic cycle, such as Otto, diesel, or air-compressor cycle, or it may be the actual value, as measured by an instrument (generally called an engine indicator), which traces the real performance cycle diagram as the machine is running. The **illustration** shows the difference with a two-cycle diesel engine under ideal conditions and under real conditions. The colored area represents the effective work of the cycle. If the area is divided by the length of the diagram (the stroke of the engine), the result is an average height—the mean pressure or mean effective pressure acting throughout the cycle.

The mep can be applied to the cylinder dimensions and to the speed of the engine (or compressor) to give the horsepower. The equation is generally identified as the *plan* equation:

$$\text{Horsepower} = \frac{plan}{33{,}000}$$

$p =$ mep, lb/in.^{2}

$l =$ length of stroke, ft

$a =$ net area of piston, in.^{2}

$n =$ number of cycles completed per minute

Inspection of this equation shows that, for an engine cylinder of given size (l and a) and for a given operating speed (n), the higher the mean effective pressure (p) the greater will be the horsepower output, resulting in the most effective utilization of the engine bulk, weight, and investment. The criterion of mep thus is a vitally convenient device for the evaluation of a reciprocating engine, pump, or compressor design as judged by initial cost, space occupied, and deadweight. Some representative values of mep are given in the **table**. *See* COMPRESSOR; DIESEL CYCLE; THERMODYNAMIC CYCLE; VAPOR CYCLE.

Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers,* 10th ed., 1996.

Mean free path

The average distance traveled between two similar events. The concept of mean free path is met in all fields of science and is classified by the events which take place. The concept is most useful in systems which can be treated statistically, and is most frequently used in the theoretical interpretation of transport phenomena in gases and solids, such as diffusion, viscosity, heat conduction, and electrical conduction. The types of mean free paths which are used most frequently are for elastic collisions of molecules in a gas, of electrons in a crystal, of phonons in a crystal, and of neutrons in a moderator.

An elementary formula for the mean free path for elastic collision of a molecule in a gas can be derived in the following way. In a gas at a pressure P, let n be the average number of molecules per unit volume and a the average radius of a molecule. The distance between centers of the molecules during a collision will be $2a$. Assume that one molecule has a radius $2a$ and travels with a velocity c, and that all other molecules are mass-points at rest. In a time t the moving molecule will "sweep out" a volume $4\pi a^2ct$, and since there are n molecules per unit volume, there will be $4\pi a^2ctn$ collisions. The distance traveled is ct; therefore the average distance between collisions is $1/(4\pi a^2n)$. Other more exact methods of taking averages give values which are closer to those determined experimentally. In hydrogen at normal temperature and pressure the mean free path is 1.7×10^{-5} cm. *See* KINETIC THEORY OF MATTER.

W. Dexter Whitehead

Measles

An acute, highly infectious viral disease with cough, fever, and maculopapular rash. It is of worldwide endemicity. The infective particle is a ribonucleic acid (RNA) virus. Measles, canine distemper, and bovine rinderpest viruses are antigenically related. *See* ANIMAL VIRUS.

The virus enters the body via the respiratory system, multiplies there, and circulates in the blood. Cough, sneezing, conjunctivitis, photophobia, and fever occur, with Koplik's spots in the mouth. Koplik's spots are small red spots containing a bluish-white speck in the center of each, and are located on the mucous membrane of the cheek. Koplik spots are seen early in measles, before the eruption of the skin, and are considered to be a diagnostic indicator of the disease.

A rash appears after 14 days' incubation and persists 5–10 days. Serious complications may occur in 1 out of 15 persons; these are mostly respiratory (bronchitis, pneumonia), but neurological complications are also found. Encephalomyelitis occurs rarely. Permanent disabilities may ensue for a significant number of persons.

Measles is one of the leading causes of death among children in the world, particularly in the developing countries. Complications are frequent, including pneumonia, diarrhea, encephalitis, and blindness, leaving thousands of children disabled each year. Children who survive acute measles have a greatly increased chance of dying from a variety of causes over the next year. The younger the age

at which measles occurs, the greater the effect in delayed mortality.

In unvaccinated populations, immunizing infections occur in early childhood during epidemics which recur after 2–3 years' accumulation of susceptible children. Transmission is by coughing or sneezing. Measles is infectious from the onset of symptoms until a few days after the rash has appeared. Second attacks of measles are very rare. Treatment is symptomatic.

At one time, prevention was limited to use of gamma globulin, which protects for about 4 weeks, and can modify or prevent the disease. *See* IMMUNOGLOBULIN.

Killed virus vaccine was once available, but it should not be used, as certain vaccinees become sensitized and develop local reactions when revaccinated with live attenuated virus, or develop a severe illness upon contracting natural measles. Live attenuated virus vaccine effectively prevents measles; vaccine-induced antibodies persist for years. *See* BIOLOGICALS; HYPERSENSITIVITY.

Measles antibodies cross the placenta and protect the infant during the first 6 months of life. Vaccination with the live virus fails to take during this period. Timely and sustained use of vaccine should lead to the eradication of measles. That goal has been thwarted, however, by repeated outbreaks among unvaccinated inner-city preschool-age children. Lowering the minimum age for routine vaccination from 15 months to 9 months of age could prevent measles outbreaks, but children who receive single-antigen vaccine before their first birthday should be revaccinated with trivalent measles-mumps-rubella (MMR) vaccine at 15 months of age. Vaccination for measles is not recommended in persons with febrile illnesses, with allergies to eggs or other products that are used in production of the vaccine, and with congenital or acquired immune defects. Joseph L. Melnick

Bibliography. F. L. Black, Measles active and passive immunity in a worldwide perspective, *Prog. Med. Virol.*, 36:1–33, 1989; F. T. Cutts et al., Principles of measles control, *Bull. WHO*, 69:1–7, 1991; M. Garenne et al., Child mortality after high-titre measles vaccines: Prospective study in Senegal, *Lancet*, 338:903–906, 1991; A. R. Hinman et al., Elimination of indigenous measles from the United States, *Rev. Infect. Dis.*, 5:538–545, 1983; A. R. Hinman and W. A. Orenstein, Immunization practices in developed countries, *Lancet*, 335:707–710, 1990; T. H. Tulchinsky et al., Measles control in developing and developed countries: The case for a two-dose policy, *Bull. WHO*, 71:93–103, 1993.

Measure

A reference sample used in comparing lengths, areas, volumes, masses, and the like. The measures employed in scientific work are based on the international units of length, mass, and time—the meter, the kilogram, and the second—but decimal multiples and submultiples are commonly employed. Prior to the development of the international metric system, many special-purpose systems of measures had evolved and many still survive, especially in Great Britain and the United States, where metric units have not come into common use. Multiples and submultiples in these systems are usually nondecimal; some are based on successive doubling or halving; others involve subdivision of 60 or 12; still others have multiples and submultiples of unknown origin. A few common systems are listed here; others can be found in engineering handbooks.

Length measures

12 inches = 1 foot (= 0.3048 meter)
3 feet = 1 yard
$5^1/_2$ yards = 1 rod
40 rods = 1 furlong = 220 yards
8 furlongs = 1 mile = 5280 feed
4 inches = 1 hand
9 inches = 1 span

(Surveyor's length measures)

7.92 inches = 1 link
100 links = 1 chain
10 chains = 220 yards = 1 furlong
8 furlongs = 1 mile

(Nautical length measures)

6080.26 feet = 1.15156 statute miles
= 1 nautical mile
3 nautical miles = 1 league
60 nautical miles = 1 degree at Equator
6 feet = 1 fathom
120 fathoms = 1 cable

Measures of area

1 acre = 160 square rods
= 4840 square yards
640 acres = 1 square mile

1 circular inch = area of circle 1 inch in diameter
= 0.7854 square inch
1 circular mil = area of circle 0.001 inch in diam.

Measures of mass (avoirdupois weight)

16 drams = 1 ounce
16 ounce = 7000 grains
= 1 pound (= 453.6 grams)
14 pounds = 1 stone
8 stones = 112 pounds
= 1 hundredweight (cwt)
20 cwt = 2240 pounds = long ton
2000 pounds = 1 short ton

[Troy weight (for gold and silver), carat weight (for precious stones), and apothecaries' weight (for

drugs) are other systems, in all of which the grain is the same as the avoirdupois grain.]

Measures of volume (Liquid measures)

4 gills = 1 pint
2 pints = 1 quart
4 quarts = 1 U.S. gallon
= 231 cubic inches
1 Imperial gallon = 1.20094 U.S. gallons

["Old Liquid Measures," involving 1 barrel = $31\frac{1}{2}$ gallons, and apothecaries' measures, involving 16 fluid ounces = 1 pint, are still in wide use.]

(*Dry measure*)

2 pints = 1 quart
8 quarts = 1 peck
4 pecks = 1 bushel
= 2150.42 cubic inches
105 quarts = 1 barrel
8 Imperial gallons = 1 Imperial bushel

Circular measure

60 seconds = 1 minute
60 minutes = 1 degree
90 degrees = 1 quadrant
360 degrees = 1 revolution
2π radians = 1 revolution

Time

60 seconds = 1 minute
60 minutes = 1 hour
24 hours = 1 day
7 days = 1 week
365 days, 5 hours, 48 min, 45.6 s = 1 year
1 sideral day = 86, 164 s

See METRIC SYSTEM; PHYSICAL MEASUREMENT; TIME; UNITS OF MEASUREMENT; WEIGHT. Dudley Williams

Bibliography. American National Metric Council, *Metric Reference for Consumers*, 1990; J. V. Drazil, *Quantities and Units of Measurement*, 1983.

Measure theory

A branch of mathematical analysis connected with the theory of integration. In order to discuss this subject, a formal definition of the term measure must be given.

Sigma algebras. Let X be an arbitrary set. Let m be a fixed collection of subsets of X satisfying the following conditions:

1. $\phi \in m$. (ϕ is the empty set. The symbol $\in$ indicates that ϕ is an element of m.)

2. If $A \in m$, then $A^c \in m$. (A^c is the complement of A. It consists of those elements of X which do not belong to A.)

3. If $A_1, A_2, \ldots \in m$, then (1) is valid. (This set is

$$\bigcup_{k=1}^{\infty} A_k \in m \tag{1}$$

the union of $A_1, A_2, \ldots$. It consists of those elements of X which belong to at least one of the sets $A_1, A_2, \ldots$.)

In this situation the collection m is called a σ-algebra. Here the term algebra refers to the various set operations (complementation, union, intersection), and the prefix σ to the fact that countably many such operations can be performed with sets in m and still result in sets in m. For example, the three properties mentioned above imply another property:

4. If $A_1, A_2, \ldots \in m$, then (2) is valid. (This set is

$$\bigcap_{k=1}^{\infty} A_k \in m \tag{2}$$

the intersection of $A_1, A_2, \ldots$. It consists of those elements of X which belong to all of the sets $A_1, A_2, \ldots$.) *See* SET THEORY.

Some examples of σ-algebras will be given:

1. If m consists of all subsets of X, then m is a σ-algebra.

2. If m consists only of ϕ and X, then m is a σ-algebra.

These two examples are trivial. The next example is highly nontrivial.

3. Let $X = \mathbf{R}^n$ be n-dimensional euclidean space. Let $\mathcal{O}$ be the collection of all open subsets of $\mathbf{R}^n$. This is not a σ-algebra. However, it is contained in σ-algebras of subsets of $\mathbf{R}^n$ for example, the one mentioned in example. Among all σ-algebras which contain $\mathcal{O}$, there is a smallest one. Namely, it is the intersection of all σ-algebras which contain $\mathcal{O}$. This σ-algebra is said to be generated by $\mathcal{O}$. Let $\mathcal{B}_n$ designate this σ-algebra generated by the open sets of $\mathbf{R}^n$. It is called the Borel σ-algebra. A set in $\mathcal{B}_n$ is called a Borel set. $\mathcal{B}_n$ contains all the open sets in $\mathbf{R}^n$ and also all the closed sets in $\mathbf{R}^n$, but it also contains many more sets.

Measures. Now suppose that X is a set and m is a particular σ-algebra of subsets of X. A measure μ is a function which assigns to each set in m a certain nonnegative real number (or $+\infty$) and which satisfies the following conditions:

1. $\mu(\phi) = 0$.

2. If A and $B \in m$ and are disjoint ($A \cap B = \phi$), then Eq. (3) holds.

$$\mu(A \cup B) = \mu(A) + \mu(B) \tag{3}$$

3. If $A_1, A_2, \ldots \in m$, then Eq. (4) holds.

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} \mu(A_k) \tag{4}$$

Various other properties can be derived from these conditions, such as:

4. If A and $B \in m$ and $A \subset B$ (that is, A is a subset of B), then $\mu(A) \leq \mu(B)$.

5. If $A_1, A_2, \ldots \in \mathcal{M}$ and are mutually disjoint ($A_i \cap A_j = \phi$ if $i \neq j$), then Eq. (5) holds. The last property

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k) \tag{5}$$

is of crucial importance and is called the countable additivity property of the measure μ.

Some examples of measures will be given:

1. Let X be any set and let $\mathcal{M}$ consist of all subsets of X. For any $A \subset X$ define $\mu(A)$ by Eqs. (6). It is not difficult to see that μ is a measure. It is called counting measure.

$$\mu(A) = \begin{cases} \text{number of elements in } A, & \text{if } A \text{ is finite} \\ \infty, & \text{if } A \text{ is infinite} \end{cases} \tag{6}$$

2. Again, let $\mathcal{M}$ be the σ-algebra of all subsets of X. Let $a \in X$ be a fixed element. For any $A \subset X$, define $\mu(A)$ by Eqs. (7). This measure is called the Dirac measure concentrated at a, and is often denoted δ_a.

$$\mu(A) = \begin{cases} 0, & \text{if } a \notin A \\ 1, & \text{if } a \in A \end{cases} \tag{7}$$

The most important example is Lebesgue measure. It is much harder to describe than the previous examples and will be discussed separately.

Lebesgue measure. Let $\mathbf{R}^n$ be n-dimensional euclidean space. A measure λ will now be defined for certain subsets of $\mathbf{R}^n$. If $n = 1$, the measure $\lambda(A)$ will correspond to the length of A; if $n = 2$, $\lambda(A)$ will correspond to the area of A; if $n = 3$, $\lambda(A)$ will correspond to the volume of A. A standard unit of "measurement" must be chosen. To do this, arbitrarily say that the measure of an "interval" is the product of the lengths of its edges. In this discussion an interval is any set of the form given by Eq. (8). The definition is then given by Eq. (9). Thus, $\lambda(I)$ is the length, area, or volume of I if $n = 1$, 2, or 3, respectively.

$$I = \{x \in \mathbf{R}^n \mid a_i \leq x_i \leq b_i,\ i = 1, \ldots, n\} \tag{8}$$

$$\lambda(I) = (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n) \tag{9}$$

The next step is to define the measure of an open set $G \subset \mathbf{R}^n$. This is done by approximating G by smaller finite unions of intervals, as follows. Consider any intervals $I_1, \ldots, I_m$ such that each pair is nonoverlapping and such that each $I_k \subset G$. Then form the number $\lambda(I_1) + \cdots + \lambda(I_m)$. A set of real numbers is obtained in this way. Then $\lambda(G)$ is simply the least upperbound of this set of numbers.

Now let A be any set in $\mathbf{R}^n$. The outer measure of A is defined by approximating A by larger open sets. Specifically, the outer measure of A, $\lambda^*(A)$, is the greatest lower bound of the measures $\lambda(G)$ of open sets G such that $A \subset G$.

It is not the case that λ^* is a measure on the σ-algebra of all subsets of $\mathbf{R}^n$. However, it is a measure on a smaller σ-algebra, which can be defined by a standard procedure that applies to all outer measures, and which will now be described.

Outer measures. Let X be an arbitrary set. An outer measure on X is a function α which assigns to each $A \subset X$ a number $\alpha(A)$ and which satisfies the following conditions:

1. $0 \leq \alpha(A) \leq \infty$
2. $\alpha(\phi) = 0$
3. $\alpha\left(\bigcup_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} i\ \alpha(A_k)$

Every measure defined on the σ-algebra of all subsets of X is an outer measure. There are important outer measures which are not measures, for example, the Lebesgue outer measure λ^* defined above. (It can be verified that λ^* satisfies these conditions.)

Given an outer measure α, C. Carathéodory has defined an important collection of subsets of X associated with α: a set E is said to be measurable (with respect to α) if Eq. (10) is satisfied. The following two facts can then be proved:

$$\alpha(A) = \alpha(A \cap E) + \alpha(A \cap E^c) \quad \text{for all } A \subset X \tag{10}$$

1. The collection of measurable sets is a σ-algebra.
2. When α is restricted to the measurable sets, it is a measure.

Application to Lebesgue measure. As mentioned above, λ^* is an outer measure on $\mathbf{R}^n$. Let $\mathcal{L}_n$ denote the measurable sets which arise from this outer measure by means of the Carathéodory construction. It can be proved that every open set in $\mathbf{R}^n$ is measurable. Since $\mathcal{L}_n$ is thus a σ-algebra which contains all open sets, it also contains $\mathcal{B}_n$ (the smallest σ-algebra which contains all open sets).

This, finally, completes the definition of n-dimensional Lebesgue measure. The set is $\mathbf{R}^n$; the σ-algebra is $\mathcal{L}_n$ and the Lebesgue measure of a (measurable) set $A \in \mathcal{L}_n$ is defined by Eq. (11).

$$\lambda(A) = \lambda^*(A) \tag{11}$$

The importance of Lebesgue measure is due to three factors: (1) $\lambda(A)$ is what an n-dimensional measure should be if A is an open set; (2) a great collection of sets are measurable; and (3) λ is countably additive.

Integration. Measure theory has a great number of important applications. Undoubtedly, the most important is the application to integration.

Let μ be an arbitrary measure on a σ-algebra $\mathcal{M}$ of subsets of X. Let f be a nonnegative simple function on X. That is, there exist disjoint sets $A_1, \ldots, A_m \in \mathcal{M}$ such that Eq. (12) holds and f has a constant value c_k on A_k, where $0 \leq c_k \leq \infty$. Then the integral of f (with respect to μ) is defined by Eq. (13).

$$X = \bigcup_{k=1}^{m} A_k \tag{12}$$

$$\int_X f\, d\mu = \sum_{k=1}^{m} c_k\ \mu(A_k) \tag{13}$$

Now suppose f is a nonnegative measurable function on X. This means that for all $0 \le t < \infty$ the set $S(t)$, consisting of all elements $x \in X$ such that $f(x) \le t$, is a member of m. The integral of f is then defined by Eq. (14), where g is any nonnegative simple function

$$\int_X f\, d\mu = \text{least upper bound of } \int_X g\, d\mu \tag{14}$$

satisfying $g(x) \le f(x)$ for all $x \in X$.

Finally, an arbitrary real-valued function f defined on X is said to be integrable if the two nonnegative functions given by Eqs. (15) are themselves measur-

$$\begin{aligned} f_+(x) &= \max(f(x), 0) \\ f_-(x) &= \max(-f(x), 0) \end{aligned} \tag{15}$$

able and if their integrals are finite. In this case, the integral of f is defined to be the real number given by Eq. (16).

$$\int_X f\, d\mu = \int_X f + d\mu - \int_X f - d\mu \tag{16}$$

In the case of counting measure, the integral is really an infinite sum given by Eq. (17). The function

$$\int_X f\, d\mu = \sum_{x \in X} f(x) \tag{17}$$

f is integrable if and only if it vanishes except on a countable set $\{x_1, x_2, \ldots\}$ and Eq. (18) holds. In

$$\sum_{k=1}^{\infty} |f(x_k)| < \infty \tag{18}$$

the case of the Dirac measure concentrated at a, the integral is given by Eq. (19), and f is integrable if and

$$\int_X f\, d\delta_a = f(a) \tag{19}$$

only if $f(a)$ is a (finite) real number.

Connection with Riemann integration. The integration concept just defined is more valuable in many applications than Riemann integration (the technique discussed in most calculus texts) because more functions are integrable and there are more easily applied limit theorems of the form of Eq. (20).

$$\lim_{k \to \infty} \int_X f_K\, d\mu = \int_X (\lim_{k \to \infty} f_K)\, d\mu \tag{20}$$

In order to illustrate the first point, suppose that $n = 1$ and suppose all functions under consideration vanish outside a fixed interval $[a,b] \subset \mathbf{R}$. Also suppose f is a bounded function. In case f is integrable in the sense of Riemann (as defined in calculus texts), denote its Riemann integral as expression (21). In

$$R \int_a^b f(x)\, dx \tag{21}$$

case f is integrable in the sense of Lebesgue (as defined above), denote its (Lebesgue) integral as expression (22). It can then be proved that if the former

$$L \int_a^b f(x)\, dx \quad \left(= \int_{[a,b]} f\, d\lambda \text{ in the above notation} \right) \tag{22}$$

integral exists, then so does the latter, and also that Eq. (23) holds. But also there are functions which are

$$R \int_a^b f(x)\, dx = L \int_a^b f(x)\, dx \tag{23}$$

Lebesgue-integrable and not Riemann-integrable. For example, if $f(x) = 0$ or 1 according as x is rational or irrational, then f is not Riemann-integrable, although the Lebesgue integral exists and is given by Eq. (24).

$$L \int_a^b f(x)\, dx = b - a \tag{24}$$

Lesbesgue measure also provides an answer to the question of Riemann integrability. The bounded function f on $[a,b]$ is Riemann-integrable if and only if the set of x at which f is discontinuous has zero Lebesgue measure. *See* INTEGRATION.

B. Frank Jones, Jr.

Bibliography. D. Cohn, *Measure Theory*, 1996; J. L. Doob, *Measure Theory and Probability*, 1993; P. R. Halmos, *Measure Theory*, 1950, reprint 1991; F. Morgan, *Geometric Measure Theory*, 3d ed., 2000; H. L. Royden, *Real Analysis*, 3d ed., 1988.

Measured daywork

A tool used primarily in manufacturing facilities as a control device to measure productive output in relation to labor input within a specific time period. The measurement of the work content is accomplished through the use of time standards (which are usually the result of a stopwatch time study), predetermined time standards (methods-time measurement, the work-factor system), or some other form of work-measurement technique designed to measure tasks of labor under normal and average conditions.

Characteristics. A measured daywork plan shares some characteristics of both incentive pay plans and unmeasured daywork plans, thus becoming a unique plan in itself. It is similar to incentive pay plans inasmuch as, in both plans, time standards are used as a device for measuring operator performance and also for various forms of management planning. There the similarity ends, for, with incentive plans, operator earnings are directly related to, and fluctuate accordingly with, productive output. In a measured daywork plan, worker income is based on a fixed hourly rate established by management, and is usually affected only by job classification, shift premiums, and overtime adjustments. Because of the fixed hourly rate in a measured daywork plan, there is a little incentive for a worker to exceed a normal or standard level of performance or productivity. On

the other hand, time standards are more readily acceptable, and become less an item of contention to the employee and bargaining unit (union).

The fixed hourly rate of measured daywork is also characteristic of an unmeasured daywork plan. However, unlike an unmeasured daywork plan, measured daywork does determine worker performance factually. Other advantages of measured daywork compared with unmeasured daywork are that costs can be readily identified for specific jobs or products, accurate estimated costs for new products can be derived if standard data have been compiled, and management planning and control are aided by use of time standards when planning for equipment and worker-power needs and scheduling work through the shop.

Operating principles. The term daywork as used in industry denotes a fixed hourly rate that is not raised or lowered by varying worker performance levels. The hourly rate for a particular job should be a fair one relative to other jobs in the shop, and should also be comparable to rates of pay for similar jobs in the industrial community. In order to ensure an equality of pay rates, a job-evaluation program should exist, and be reviewed and updated as required. Once it is established what a fair rate of pay should be for each job, measurements can be made regarding how long it should take under normal, average conditions to complete the job, task, or operation. Allowances should be made for personal time, unavoidable delays, and, when required, fatigue factors. The result is a time standard as it relates to a specific job, task, or operation. *See* PRODUCTIVITY.

Once measured labor time standards have been established for shop operations, in addition to evaluating operator performance and identifying labor costs, new-product costs can also be determined prior to release to production, worker-power planning and scheduling can be done, equipment capacity requirements can be identified, and planning and make/buy decisions can be facilitated. Most importantly, a system has also been established to aid in identifying those areas, operations, and tasks that offer cost reduction potential.

Program criteria. There are several key criteria that are essential to a sound measured daywork program. First, the labor standards must be realistic and kept current to reflect existing shop conditions. In addition, the labor standards must also be fair and attainable.

Second, the worker should be made aware of what is the standard or production rate/quota. This in itself will often aid in increasing productivity.

Third, since there is no pay incentive for the worker to exceed standard or normal performance, there is the inherent characteristic of employees in a measured daywork system to fall somewhat short of standard or 100% performance level. Consequently it is of the utmost importance to have strong, qualified, and aggressive first-line supervision, or a foreperson who will provide direct supervision to the employee. *See* PERFORMANCE RATING; WAGE INCENTIVES; WORK MEASUREMENT.

Dave H. Scott

Bibliography. C. R. Asfahl, *Industrial Safety and Health Management*, 4th ed., 1998; P. E. Hicks, *Industrial Engineering and Management: A New Perspective*, 2d ed., 1999; R. C. Vaughn, *Introduction to Industrial Engineering*, 1985.

Mechanical advantage

Ratio of the force exerted by a machine (the output) to the force exerted on the machine, usually by an operator (the input). The term is useful in discussing a simple machine, where it becomes a figure of merit. It is not particularly useful, however, when applied to more complicated machines, where other considerations become more important than a simple ratio of forces. *See* EFFICIENCY; SIMPLE MACHINE.

Richard M. Phelan

Mechanical alloying

A materials-processing method that involves the repeated welding, fracturing, and rewelding of a mixture of powder particles, generally in a high-energy ball mill, to produce a controlled, extremely fine microstructure. The mechanical alloying technique allows alloying of elements that are difficult or impossible to combine by conventional melting methods. In general, the process can be viewed as a means of assembling metal constituents with a controlled microstructure. If two metals will form a solid solution, mechanical alloying can be used to achieve this state without the need for a high-temperature excursion. Conversely, if the two metals are insoluble in the liquid or solid state, an extremely fine dispersion of one of the metals in the other can be accomplished. The process of mechanical alloying was originally developed as a means of overcoming the disadvantages associated with using powder metallurgy to alloy elements that are difficult to combine. Using powder metallurgy, homogeneity is dictated by the size of the particle, but contamination and fire hazards become a concern when particle size is very small. *See* POWDER METALLURGY; SOLID SOLUTION.

Process. The process of mechanical alloying consists of repeated flattening, fracture, and rewelding of the powder particles in a high-energy ball charge. Every time two steel balls collide, they trap powder particles between them. This deforms the particles and creates minimal foreign species on the surface so that welding together of adjacent surfaces can occur easily. The alloying is effected by the ball-powder-ball collisions in several stages. In the first stage, intense cold welding predominates, and layered composite particulates of the starting constituents form (**illus.** *a*). This is followed by a hardening of the particles and fracturing and cold welding leading to a finer composite particle size (illus. *b*). Solid-solution formation begins at this stage. Next, in a moderate cold-welding period, the microstructure of the particles gets finer, with a typical spacing between

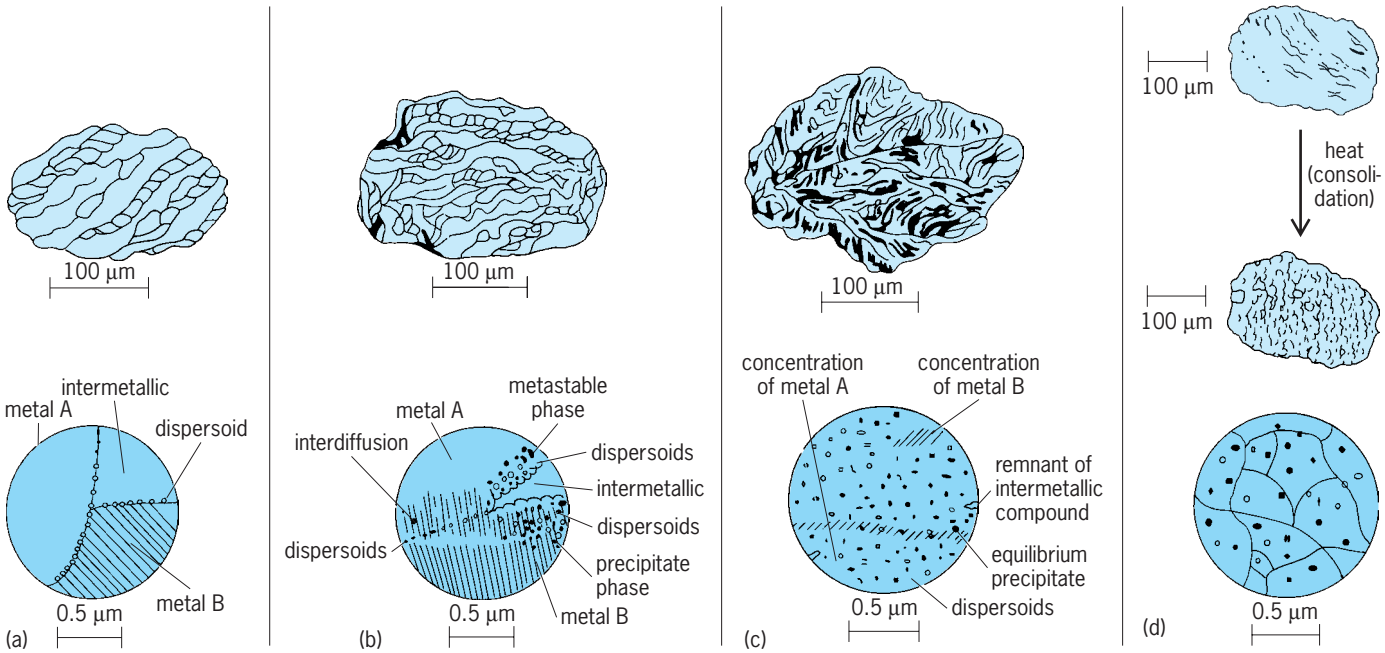

Diagrams showing mechanical alloying; the upper portion in each case represents low magnification, and the lower portion represents a higher magnification. (*a*) First stage of processing, intense cold welding; (*b*) intermediate stage of processing, rapid fracturing; (*c*) final stage of processing, moderate cold welding; (*d*) completion of processing, steady-state. The powder has an extremely deformed metastable structure with dispersoids, where the lamellae are no longer optically resoluble; that is, they cannot be resolved in the optical microscope (in contrast to having to go to a higher magnification using a transmission electron microscope). On consolidation, this leads to a fine grain size with equilibrium distribution of dispersoids.

adjacent regions of 1 micrometer (illus. *c*). The compositions of the individual particles converge to the original blend composition, with the rate of refinement of the internal structure (reduction in scale of the microstructure) of the particles approximately logarithmic with processing time. The final stage is a steady-state period (illus. *d*). By the time processing has been completed, the particles have an extremely deformed metastable structure, which can contain dispersoids. Generally, at this point the microstructural scale is significantly below the micrometer level.

Mechanical alloying can be carried out in small batches on shaker mills (few grams per run), high-energy stirred-ball millers known as attritors (1.1–220 lb or 0.5–100 kg of powder), or on a commercial scale in a ball mill (up to about 4400 lb or 2000 kg). The attritor is capable of generating high energies by utilizing a vertical drum with a series of impellers inside it. A powerful motor rotates the impellers, which in turn agitate the steel balls in the drum. The scaling-up from one process to the next is not direct, and it is considered an art rather than a science. The powders used in mechanical alloying are very hard because of their heavily worked condition; thus they must normally be consolidated to a useful form by hot-processing methods. The powders for mechanical alloying are hot-pressed followed by hot extrusion; in a number of situations this is followed by a hot-rolling operation. *See* CRUSHING AND PULVERIZING; EXTRUSION; METAL FORMING.

Applications. Some oxides are insoluble in molten metals. Mechanical alloying was originally developed to provide a means of dispersing these oxides in the metals. Examples are nickel-based superalloys strengthened with dispersed thorium oxide or yttrium oxide (Y_2O_3), the latter being preferred. These superalloys have excellent strength and corrosion resistance at elevated temperatures, making them attractive candidate materials for use in applications such as jet-engine turbine blades, vanes, and combustors. This class of alloys is known as oxide-dispersion-strengthened nickel-base superalloys. Iron- and aluminum-based dispersion-strengthened alloys have also been developed. Additionally, a number of other potential applications for mechanical alloying material are being explored, including powders for coating applications, alloys of immiscible systems, amorphous alloys, intermetallics, cermets, and organic-ceramic-metallic material systems in general. *See* AMORPHOUS SOLID; CERMET; INTERMETALLIC COMPOUNDS; METAL COATINGS.

Alloys like MA754 (Ni-20Cr-0.3Al-0.5Ti-0.6Y_2O_3) and MA6000 (Ni-15Cr-4W-2Mo-2Ta-4.5Al-2.5Ti-1.1Y_2O_3) are commercially available as bar and plate for use in gas turbine vanes and turbine blades and for other applications in oxidizing or corrosive atmospheres. The vane alloy MA754 is capable of a service temperature of about 2100°F (1100°C) and is used in military engines. The iron-base alloy MA956 (Fe-20Cr-4.5Al-0.5Ti-0.5Y_2O_3), available as plate sheet, bar, and wire, is capable of withstanding temperatures up to 2370°C (1300°C) or even higher in corrosive environments; it has applications in aircraft and industrial gas turbine combustors, in

swirlers and heat exchangers of power generation equipment, and in heat-treatment equipment.

Conventional coatings in advanced gas turbine engines fail by loss of aluminum in the coating or by interdiffusion between the coating and substrate. Coatings possessing a large amount of Y_2O_3, produced by mechanical alloying, could solve such problems. *See* METAL COATINGS.

Liquid or solid immiscible systems are difficult to process by conventional pyrometallurgy, for example, the copper-lead (Cu-Pb) or copper-iron (Cu-Fe) systems. In these cases, mechanical alloying provides a route to obtain a homogeneous distribution in the solid phase. Mechanical alloying has also been used to fabricate the superconducting intermetallic niobium-tin compound Nb_3Sn, which is difficult to produce conventionally because of the large melting-point difference between Nb and Sn. Mechanical alloying can also be used to produce supercorroding magnesium-base alloys. These alloys are designed to corrode at a controlled rate for release of deep-sea equipment at specific depths. The requirement is to have the anode and cathode in proximity by using mechanical alloying. In another interesting application of mechanical alloying, titanium (Ti) and magnesium (Mg) have been combined by using mechanical alloying; usually it is very difficult to produce such an alloy because the boiling point of Mg is lower than the melting point of Ti. This achievement could result in a lower-density titanium alloy for aerospace applications. *See* PYROMETALLURGY.

Amorphous alloys can also be produced by using the mechanical alloying technique. These include Nb_3Sn starting from elemental powders, various rare-earth/cobalt combinations, and titanium-copper (Ti-Cu) and titanium-copper-palladium (Ti-Cu-Pd) systems. *See* ALLOY; HIGH-TEMPERATURE MATERIALS; SOLID-STATE CHEMISTRY; WELDING AND CUTTING OF MATERIALS. F. H. Froes

Bibliography. J. J. De Barbadillo, *Mechanical Alloying for Structural Applications*, 1993; B. L. Bramfitt et al. (eds.), *Oxide Dispersion Strengthened Alloys*, ASTM STP-979, 1988; R. M. German, *Powder Metallurgy Science*, 2d ed., 1994; P. H. Shingu (ed.), *Mechanical Alloying*, 1992; R. Sundaresan and F. H. Froes, Mechanical alloying, *J. Met.*, pp. 22–27, August 1987.

Mechanical classification

A sorting operation in which mixtures of particles of mixed sizes, and often of different specific gravities, are separated into fractions by the action of a stream of fluid. Water is ordinarily used as the sorting fluid, but other liquids or air or other gases may be used.

The main objective of classification is to separate the particles according to size. This function is identical to that of screening, but classification is applicable to smaller particles, especially those that are undersize. For small particles, it is more economical than screening. In classification, the oversize and undersize are called sands and slimes, respectively.

Material also may be mechanically classified by specific gravity, a method that separates substances differing in chemical composition. This is called hydraulic separation. Such classification is based on the fact that, in a fluid, particles of the same specific gravity but of different size or shape settle at different constant speeds. Large, heavy, round particles settle faster than small, light, needlelike ones. If the particles also differ in specific gravity, the speed of settling is further affected. This is the basis for the separation of particles by kind rather than by size alone.

Mechanical classifiers utilize the differences in settling velocities by bringing the feed mixture into contact with a continuous stream of fluid, and allowing the particles to reach their settling velocities with respect to the fluid. Then, by adjusting the rate of flow of the fluid, the small particles, which have the lower settling speeds, are removed from the classifier with the fluid stream before they have had time to settle out of the fluid. The larger particles, which settle through the fluid more rapidly, strike the wall or bottom of the classifier, collect in the bottom of the apparatus, and are rejected as sands.

When the particles in a classifier are so far apart that they do not interfere with one another, the action is called free settling; when they are so close together that they do interfere with one another, the action is called hindered settling. In hindered settling the velocities of settling are less than in free settling, but the capacity of the equipment, measured in tons per square foot per day, is much greater. Also, the separation by specific gravity is improved in hindered settling. This method is used in commercial classification.

Wet classifiers. These are built in two major categories—horizontal flow and vertical flow. The latter are more effective in separations by specific gravity, and the former have greater capacity in separation by size alone.

An example of a horizontal-flow classifier is the Dorr unit, shown in **Fig. 1**. It consists of a rectangular tank with a sloping bottom, a rake mechanism for moving sands uphill along the bottom, an inlet for feed, and outlets for sands and slimes. The feed is admitted in a stream of water behind a baffle and subjected to the action of the stream as it flows toward the slime overflow. Coarse solids (sands) settle

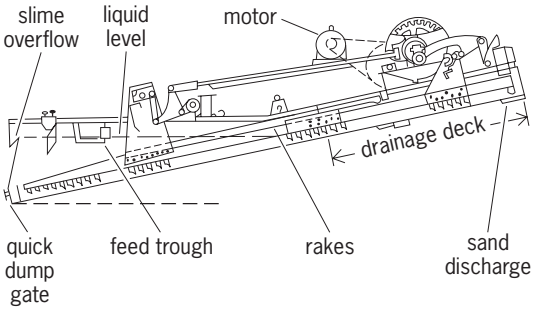

Fig. 1. Dorr classifier. (*After R. R. Perry, ed., Chemical Engineers' Handbook, 4th ed., McGraw-Hill, 1963*)

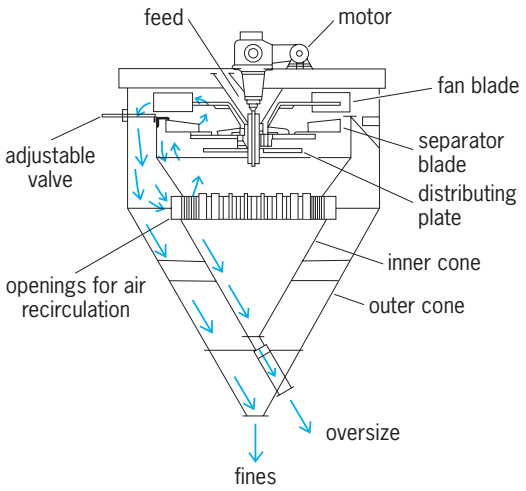

Fig. 2. Double-cone air separator. (*After W. L. McCabe and J. C. Smith, Unit Operations of Chemical Engineering, McGraw-Hill, 1956*)

rapidly enough to reach the bottom of the tank and are raked out of the classifier at the shallow end. The rakes accomplish this by moving slowly along the bottom, dragging the solids with them. After a short motion in this direction, the rakes are lifted clear of the sands by mechanical means, returned, and lowered for another stroke. The action is sufficiently gentle not to disturb the settling in the bulk of the liquid. Fines (slimes) overflow with the liquid at the deep end of the tank.

Vertical-flow classifiers are used when separation by specific gravity is important. These machines have been largely superseded by flotation, which can make closer separations of substances of different chemical composition. *See* FLOTATION.

Air classifiers. These are used to separate particles by size through the action of a stream of air. They make sharper and more accurate separations than wet classifiers. An example is the double-cone air separator in **Fig. 2**. The feed, consisting of a mixture of sizes, is dropped on a horizontal rotating distributor plate, which is driven by a vertical shaft that also carries fan blades. The plate disperses the particles into the air stream, and the fan maintains a circulation of air inside the separator. The air flows downward and then suddenly reverses to flow upward. The larger and heavier particles have sufficient momentum that they do not follow the air, but are thrown against the inner surface of the inner cone, from which they are collected and withdrawn. The smaller particles accompany the air into the annular space between the two cones, where the velocity is low enough to allow the particles to settle and be removed from the bottom of the outside cone. *See* MECHANICAL SEPARATION TECHNIQUES; UNIT OPERATIONS.

Warren L. McCabe

Bibliography. W. L. McCabe, J. C. Smith, and P. Hariott, *Unit Operations in Mechanical Engineering*, 7th ed., 2004; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997.

Mechanical engineering

The branch of engineering that deals with the generation, transmission, and utilization of heat and mechanical power, and with the production of tools and machines. This distinction begins in college where, after 1 year of common engineering studies, a person can be identified as a student of mechanical engineering. After college, mechanical engineers may enter many industries. Common areas of industrial employment for mechanical engineers are private electric power and machinery manufacturing.

To grasp the meaning of mechanical engineering, it is desirable to take a close look at what engineering really is. The Engineers' Council for Professional Development has defined engineering as the profession in which a knowledge of the mathematical and physical sciences gained by study, experience, and practice is applied with judgment to develop ways to utilize economically the materials and forces of nature for the progressive well-being of humankind. Here is a picture of a profession in which study in mathematics and science is blended with experience and judgment for the production of useful things.

Formal training of a mechanical engineer includes mastery of mathematics through the level of differential equations. Training in physical science embraces chemistry, physics, mechanics of materials, fluid mechanics, thermodynamics, statics, and dynamics. Enhancing these subjects are courses in the humanities: literature, economics, philosophy, and history.

Experience, for the mechanical engineer, is not gained solely by the passage of time. For experience to be added, the time spent in engineering must be meaningful. Projects must become more difficult and the consequences of error greater.

Judgment is the hallmark of any competent engineer in whatever specialized branch of engineering is pursued. The exercise of judgment requires the ability to assess and decide between alternate courses of action. In the application of judgment there is constant testing of the engineer's decisions against his or her knowledge of the laws of nature, his or her sense of right and wrong, and economics.

A further essential of an engineer is that efforts be devoted to matters which improve the well-being of humankind. This purpose is kindred to that of all learned professions and is closely allied to the professions of medicine and law.

In relation to other professions the mechanical engineer differs in one large measure. Whereas most doctors and many lawyers are self-employed, almost all engineers are employed by corporations, colleges, or government. This distinction exists for several reasons. Generally the doctor or lawyer is far more concerned with people and their actions than is the engineer. Also, an engineer dealing with machines requires larger sums of money to finance the equipment and facilities used than does the doctor or lawyer.

From the foregoing, a young person thinking of becoming a mechanical engineer should recognize that

there is a long period of preparation necessary. Starting in high school, a full, rigorous academic program should be pursued. Then careful self-examination of results in high school are in order. If high interest and success in mathematics, English, and science are demonstrated, there is room to consider the matter further. Finally, there must be evidence of willingness to work and patience to gain the experience and judgment required in any learned profession. *See* ENGINEERING; MACHINE DESIGN; MACHINERY; TECHNOLOGY.

Robert S. Sherwood

Bibliography. American Society of Mechanical Engineers, *Mechanical Engineering*, monthly; Engineer's Council for Professional Development (ECPD), *So You Want To Be an Engineer*, Pam. EC11; J. R. Henninger (ed. ECDP), *A Professional Guide for Young Engineers*, Pam. EC43; A. P. Johnson (ECPD), *Do I Have Engineering Aptitude?*, Pam. EC14.

Mechanical impedance

For a system executing simple harmonic motion, the mechanical impedance is the ratio of force to particle velocity. If the force is that which drives the system and the velocity is that of the point of application of the force, the ratio is the input or driving-point impedance. If the velocity is that at some other point, the ratio is the transfer impedance corresponding to the two points.

As in the case of electrical impedance, to which it is analogous, mechanical impedance is a complex quantity. The real part, the mechanical resistance, is independent of frequency if the dissipative forces are proportional to velocity; the imaginary part, the mechanical reactance, varies with frequency, becoming zero at the resonant and infinite at the antiresonant frequencies of the system. *See* ACOUSTIC IMPEDANCE; ELECTRICAL IMPEDANCE; FORCED OSCILLATION; HARMONIC MOTION.

Martin Greenspan

Bibliography. S. G. Kelly, *Fundamentals of Mechanical Vibrations*, 2d ed., 2000; W. T. Thomson, *Theory of Vibration with Applications*, 4th ed., 1992.

Mechanical separation techniques

A group of laboratory and production operations whereby the components of a polyphase mixture are separated by mechanical methods into two or more fractions of different mechanical characteristics. The separated fractions may be homogeneous or heterogeneous, particulate or nonparticulate. Mechanical separations are differentiated from two other classes of separations of major importance: those that depend on separation at the molecular level on the basis of thermodynamic or physicochemical properties and by means of such mass-transfer processes as evaporation, dissolution, precipitation, absorption, and intraphase or interphase diffusion; and those that depend on chemical reaction of a component and subsequent separation of the product by mass-transfer or mechanical methods. *See* CHEMICAL SEPARATION TECHNIQUES; UNIT OPERATIONS.

The techniques of mechanical separation are based on differences in phase density, in phase fluidity, and in such mechanical properties of particles as size, shape, and density; and on such particle characteristics as wettability, surface charge, and magnetic susceptibility. Obviously, such techniques are applicable only to the separation of phases in a heterogeneous mixture. They may be applied, however, to all kinds of mixtures containing two or more phases, whether they are liquid-liquid, liquid-gas, liquid-solid, gas-solid, solid-solid, or gas-liquid-solid.

Methods of mechanical separations fall into four general classes: (1) those employing a selective barrier such as a screen or filter cloth; (2) those depending on differences in phase density alone (hydrostatic separators); (3) those depending on fluid and particle mechanics; and (4) those depending on surface or electrical characteristics of particles.

Because of the extensive differences in the nature of mixtures to be separated, in the properties of the components, in the size of the particles, and in the functional character of the tasks performed by mechanical separators, a wide variety of separation devices have been devised and are in use. The more important kinds of equipment are listed in the **table**, grouped according to the phases involved.

Use of selective barriers. In liquid filters, including clarifying and gas (bag or air) filters, solids are removed from the fluid by pumping or drawing the fluid through a filter medium which retains the solids on or within itself. In a centrifugal filter the same process occurs (limited to liquid-solid suspensions), but the driving force is that of a centrifugal field. The solids are removed from the medium, continuously or intermittently, or are discarded with the medium when it is replaced. In screening, screen cloth or perforated sheet or plate is used to pass particles smaller than, and retain particles larger than, the screen aperture (normally performed dry), or to separate relatively large particles from a liquid (wet screening). A special case of separation by a selective barrier is the operation called expression, the squeezing of liquid from a nonfluid solid-liquid system (such as a cider press or cane mill). *See* CENTRIFUGATION; CLARIFICATION; FILTRATION; MEMBRANE SEPARATIONS; SCREENING.

Hydrostatic separators. Batch or continuous settling tanks are employed to retain liquid while an immiscible liquid, a gas, or a solid separates from it. A supernatant liquid or floating solid may be overflowed or skimmed away; a subnatant liquid or settled solid may be pumped away (the latter as a thick sludge). A sedimentation thickener is a continuous hydrostatic settling tank. Sedimentation centrifuges are substantially hydrostatic separators when used as liquid clarifiers or oil-water separators, since density difference is the prime quality that makes separation possible, provided sufficient residence time is allowed, as in settling tanks, to permit complete separation. The same is true of dense-medium separators, in which a pulp of controlled density carries

Types of mechanical separator

Materials separated	Separators
Liquid from liquid	Settling tanks, liquid cyclones, centrifugal decanters, coalescers
Gas from liquid	Still tanks, deaerators, foam breakers
Liquid from gas	Settling chambers, cyclones, electrostatic precipitators, impingement separators
Solid from liquid	Filters, centrifugal filters, clarifiers, thickeners, sedimentation centrifuges, liquid cyclones, wet screens, magnetic separators
Liquid from solid	Presses, centrifugal extractors
Solid from gas	Settling chambers, air filters, bag filters, cyclones, impingement separators, electrostatic and high-tension precipitators
Solid from solid	
By size	Screens, air and wet classifiers, centrifugal classifiers
By other characteristics	Air and wet classifiers, centrifugal classifiers, jigs, tables, spiral concentrators, flotation cells, dense-medium separators, magnetic separators, electrostatic separators

a low-density fraction of solids on its surface while allowing a higher-density fraction of solids to sink and be raked away. *See* SEDIMENTATION (INDUSTRY); THICKENING.

Fluid and particle dynamics. Most of the other methods listed in the table are based on movement of liquid drops or solid particles with respect to a suspending fluid (gas or liquid). Even most solid-solid separations require the presence of a fluid.

Drag forces and gravity, either natural or centrifugal, are the factors involved in establishing the trajectories that make wet and air classifiers, jigs, tables, spiral concentrators, cyclones, and classifying centrifuges effective.

Inertia differences between particles and the suspending fluid are the basis for the operation of impingement separators. When a fluid, ordinarily a gas, is given a sharp change in direction, any particulates therein, because of their inertia, do not conform to the new direction of flow, but continue to move nearly in their original direction. Then they strike the solid surfaces of the equipment, which may consist of a bed of solid shapes, a series of baffles, a bed of liquid-covered fibers, or other collecting surfaces. The particles striking these surfaces coalesce and are removed from the gas stream. *See* DUST AND MIST COLLECTION; MECHANICAL CLASSIFICATION.

Other methods. Froth flotation cells operate on the bases of selective wettability and density difference. Chemicals, called promoters or collectors, cause the particles of one fraction of a suspension to become air-avid and water-repellent. Air introduced to the suspension collects as bubbles on these particles, causing them to float to the surface, while the remaining solids sink and are withdrawn at the bottom.

Electrostatic and electrodynamic (current flow) forces are used in electrostatic precipitators to drive small dust particles or droplets to a collecting electrode, and thus to clean a smoky or foggy gas stream. Magnetic separators, batch or continuous, are used to separate iron-bearing or other magnetic particles from liquids in which they are suspended. *See* ELECTROSTATIC PRECIPITATOR; FLOTATION; MAGNETIC SEPARATION METHODS.

Shelby A. Miller

Bibliography. W. L. McCabe, J. C. Smith, and P. Hariott, *Unit Operations in Mechanical Engineering*, 7th ed., 2004; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; R. W. Rousseau (ed.), *Handbook of Separation Process Technology*, 1987; J. D. Seader and E. J. Henley, *Separation Process Principles*, 2d ed., 2005.

Mechanical vibration

The continuing motion, repetitive and often periodic, of a solid or liquid body within certain spatial limits. Vibration occurs frequently in a variety of natural phenomena such as the tidal motion of the oceans, in rotating and stationary machinery, in structures as varied in nature as buildings and ships, in vehicles, and in combinations of these various elements in larger systems. The sources of vibration and the types of vibratory motion and their propagation are subjects that are complicated and depend a great deal on the particular characteristics of the systems being examined. Further, there is strong coupling between the notions of mechanical vibration and the propagation of vibration and acoustic signals through both the ground and the air so as to create possible sources of discomfort, annoyance, and even physical damage to people and structures adjacent to a source of vibration.

Mass-spring-damper system. Although vibrational phenomena are complex, some basic principles can be recognized in a very simple linear model of a mass-spring-damper system (**Fig. 1**). Such a system contains a mass M, a spring with spring constant k that serves to restore the mass to a neutral position, and a damping element which opposes the motion of the vibratory response with a force proportional to the

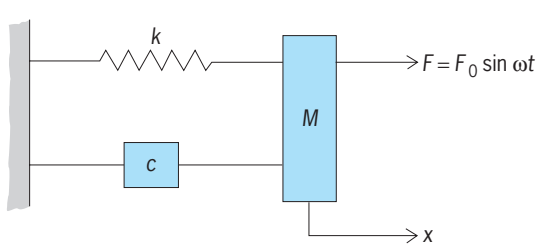

Fig. 1. Vibrating linear system (mass-spring-damper) with one degree of freedom.

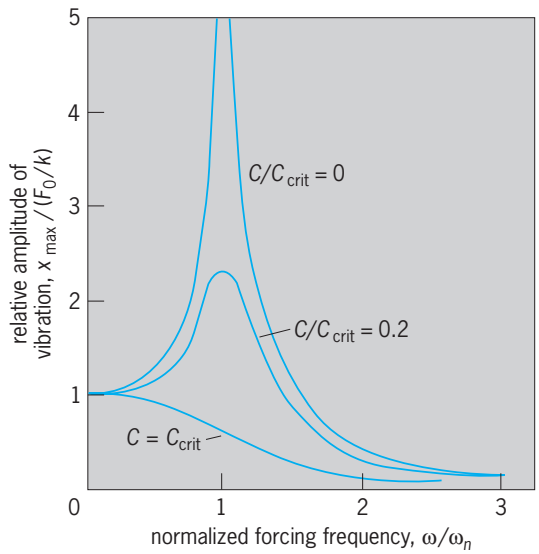

Fig. 2. Variation of amplitude of vibration with forcing frequency and damping. X_{max} = maximum displacement.

velocity of the system, the constant of proportionality being the damping constant c. This damping force is dissipative in nature, and without its presence a response of this mass-spring system would be completely periodic.

Free vibrations. In the absence of an exciting force and of the damping component, the spring-mass system can execute free, periodic vibration at a natural frequency ω_n given by Eq. (1). This is the so-called

$$\omega_n = \sqrt{k/M} \tag{1}$$

angular frequency and must be divided by 2π to get the actual frequency f_n in hertz (with dimensions time^{-1}). Equation (1) demonstrates that the natural frequency, also called the fundamental or resonant frequency, increases with the stiffness of the system and decreases as the mass of the system is increased. Thus, very stiff systems have high natural frequencies, and very massive systems could have very low natural frequencies. In some of the more complex structures referred to above, there are in fact multiple (and even infinite) resonant frequencies, but they all have in common the basic dependence upon the stiffness of the system and the system's mass. *See* HARMONIC MOTION.

Effect of damping. The inclusion of the damping element with the constant of proportionality c results in a change of the natural frequency of the system and of its response characteristics. If the damping is relatively small in magnitude (the scale against which smallness is to be measured will be given below), then the natural frequency of the system does not change substantially, and the free response of a perturbed mass-spring-damper system is periodic within the confines of a slow exponential decay. However, if the damping constant c becomes large, the motion of the system is not vibratory but is in fact one where the mass creeps back, along an exponential curve, to its initial position without any oscillation. The lowest value of damping for which this loss of vibratory response occurs is called critical damping, and it is defined by a damping constant, Eq. (2). Damping is

$$c_{crit} = 2\sqrt{kM} \tag{2}$$

considered to be small when the ratio C/C_{crit} is small compared to unity. *See* DAMPING.

Impedance and resonance. When the mass-spring-damper system is excited by a periodic force, for example, $F = F_0 \sin \omega t$, where the angular frequency ω $(= 2\pi f)$ can be chosen freely, the response of the system is such that it can be displayed in one of two commonly used forms. One of these (**Fig. 2**) is a plot of the relative amplitude of vibration (the dimensionless ratio of the actual motion of the perturbed system divided by the corresponding elastic response of the spring alone to the same force) against the normalized (dimensionless) frequency ratio of the forcing frequency divided by the undamped natural frequency. The curves shown in Fig. 2 indicate the response for no damping ($C/C_{crit} = 0$), small damping, and critical damping. It can be seen that in the undamped case ($C/C_{crit} = 0$), the amplitude of the response is infinite when the forcing frequency is identically equal to the natural frequency. This is the condition of resonance. It can also be seen that when there is damping, even though it may be small, the system responds so as to peak very near the undamped resonant frequency, but the response is not infinite in amplitude. Finally, if the damping is greater than or equal to C_{crit}, the response is simply the decaying curve that exhibits no peak at the resonant frequency.

Another way of presenting these results is to look at the quantity called the impedance Z, which may be defined here as the ratio of the magnitude of the input force to the magnitude of the velocity of the mass-spring system (**Fig. 3**). For the case where there is no damping, the impedance can be shown to be in the form of Eq. (3). It follows from

$$Z = k\left(\frac{1}{\omega} - \frac{M\omega}{k}\right) \tag{3}$$

Eq. (1), which relates the stiffness k and the mass M to the resonant frequency ω_n, and from the impedance formula that when the forcing frequency

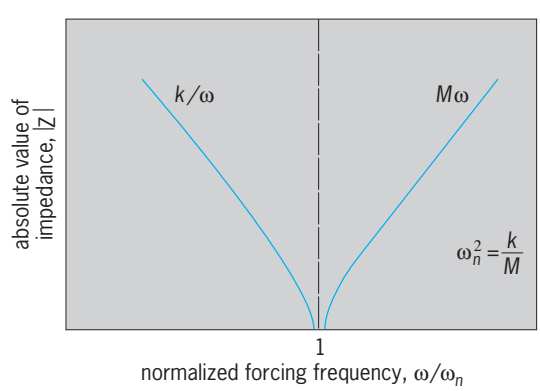

Fig. 3. Impedance of a simple oscillator.

is well below the resonant frequency, the impedance can be approximated by Eq. (4). This is considered

$$Z_k \cong \frac{k}{\omega} \qquad \omega \ll \omega_n \tag{4}$$

the stiffness-controlled region wherein the spring-mass system responds largely as a static spring. For frequencies that are significantly above the resonant frequency, the impedance can be written as Eq. (5).

$$Z_m \cong -M \quad \omega \gg \omega_n \tag{5}$$

In this regime the response of the system is said to be mass-controlled, and at high frequencies, well above the resonant frequency, the dynamics of the system are governed by the mass of the system. It follows from Eq. (3) that the impedance vanishes at resonance in the absence of damping. Thus resonance may also be defined as the condition where a very small input force produces an infinite response in the absence of damping. When damping is incorporated in the calculation of impedance, it then turns out that at resonance the impedance is exactly equal to the damping constant; that is, $Z_{\text{res}} = c$. *See* FORCED OSCILLATION; MECHANICAL IMPEDANCE; RESONANCE (ACOUSTICS AND MECHANICS).

Linear response. One other feature in the simple model is that the response is linear; that is, there is a direct proportionality of the magnitude of the input to the magnitude of the output, and signals of response for different inputs can be superposed to give the total response to a combined signal.

Relation to real systems. This simple model of the response of the single-degree-of-freedom system, as it is commony called, serves as a model or paradigm for almost all mechanical vibration problems. However, real systems are often extremely complex and involve multiple resonances and multiple sources of excitation. Further, the types of vibrational response that can be seen in different solid bodies are very much dependent not only on the nature of the excitation in terms of its time history, but also on the spatial distribution of the exciting force and the geometry of the system that is responding. There are many different ways in which a solid body can vibrate, and these different ways couple into vibratory and acoustic motions of surrounding media such as air or water in very different ways. Some of the sources of vibration, the types of vibrational response exhibited by solid bodies, and their interactions that result in perceptible, audible, and even dangerous vibration and acoustic signals will be discussed. First, however, some further comments about the general complexity of vibration problems are appropriate in order that the context of the above simple model be more clearly and sharply defined.

Analysis of complex systems. The foregoing model of the linear spring-mass-damper system contains within it a number of simplifications that do not reflect conditions of the real world in any obvious way. These simplifications include the periodicity of both the input and, to some extent, the response; the discrete nature of the input, that is, the assumption that it is temporal in nature with no reference to spatial distribution; and the assumption that only a single resonant frequency and a single set of parameters are required to describe the mass, the stiffness, and the damping. The real world is far more complex. Many sources of vibration are not periodic. These include impulsive forces and shock loading, wherein a force is suddenly applied for a very short time to a system; random excitations, wherein the signal fluctuates in time in such a way that its amplitude at any given instant can be expressed only in terms of a probabilistic expectation; and aperiodic motions, wherein the fluctuation in time may be some prescribed nonperiodic function or some other function that is not readily seen to be periodic.

However, in most cases the notions of periodic response to periodic excitation can be used because it is generally possible to describe the variation in time of a nonperiodic forcing function as a sum of periodic components using the tools of Fourier analysis and Fourier series. In such an approach, the nonperiodic signal is expressed as a sum of periodic signals, each of whose amplitudes and individual periods (or resonant frequencies) is tailored to represent the total nonperiodic signal being described. In such a Fourier analysis of a vibrating system, it is customary to look at the response of the complex system as a sum of responses of linear spring-mass-damper systems to each of the given periodic inputs, and then the simple model outlined earlier can be readily applied.

In any real system there are multiple resonances, each reflecting a different balance of stiffness and mass within a complex elastic system. Again, it is possible in general to describe the response of a complicated system in terms of individual spring-mass oscillators, the stiffness and mass of each of which represents some average of the overall stiffness and mass distributions of the complicated system. Thus, the analysis of complex systems can again be reduced to the superposition of systems of the elementary type. Very often this reduction of the complex to the simple not only involves Fourier analysis in time, but also represents a Fourier analysis in space, wherein spatial averages are used to reduce spatial variations of both excitation and response to the discrete forms in which they appear in the simple model. Thus, for a continuous system where the variables depend on space as well as time, a Fourier analysis can be used to generate a system that involves multiple degrees of freedom, each degree of freedom being a single discrete spring-mass-damper oscillator. *See* FOURIER SERIES AND TRANSFORMS.

Sources of vibration. There are many sources of mechanical and structural vibration that the engineer must contend with in both the analysis and the design of engineering systems. The most common form of mechanical vibration problem is motion induced by machinery of varying types, often but not always of the rotating variety; the machinery vibration thus induced can be a problem both for the machinery itself and in terms of vibration and noise propagated to adjacent systems. For example, the vibrations of a

compressor in an air-conditioning system in a building can create fatigue and stress problems for the compressor and adjacent piping itself, and can set into motion the floor and walls of the room wherein the compressor is housed. This in turn can set into vibrational motion other walls and components of the building and thus produce vibration and acoustic signals that are both perceptible and audible throughout the building. This simple example could thus run a gamut of problems from structural failure to human annoyance.

Such machinery-induced vibration can also be propagated into adjacent buildings through the coupling of a building foundation with the ground around it, and it is not uncommon for heavy machinery in a factory to produce perceptible vibration at considerable distances from the factory itself. This is more likely to be the case with extremely heavy pieces of equipment, for example a large-scale shredder used in a solid-waste disposal system.

Another source of vibration wherein groundborne propagation is important is vibration due to construction. Construction vibration is a major source of environmental concern, particularly in urban environments, and especially in the construction of large projects within these environments such as urban transportation systems.

Another source of vibration in urban environments is transportation systems themselves. Included may be vibration from heavy vehicles on conventional pavement as well as vibratory signals from the rail systems common in many metropolitan areas. For such rapid-transit system vibration, the sources include tunnels, at-grade rails, and elevated guideway systems. The propagation characteristics for each of these important vibration sources are different and have been the subject of extensive research.

Other vibrations are induced by natural phenomena, including earthquakes and wind forces. Wave motion is a source of vibration in mechanical and structural systems associated with offshore structures. In the analysis and design of systems that are used in environments where these natural phenomena are important, the basic principles of vibration analysis, such as the elementary model outlined above, remain the same. Thus, although such motion is not in a very strict sense due to ordinary mechanical excitation, it is in terms of practice and design a strongly related phenomenon.

Types of mechanical signals and waves. Due to the varying geometry both of the structural elements involved in a machine or structure and of the loading itself, several different types of vibratory motion may be induced in a solid body. Unlike the acoustic signals that are generated in air and water, wherein all the motion is in the form of waves termed body or compressional waves in which the fundamental mechanism is volumetric contraction or expansion, solid bodies support waves of several different varieties and thus produce several different mechanical vibratory signals. These include torsional vibration, longitudinal vibration, and the lateral vibrations of beams, plates, and membranes. The fundamental mechanisms in each case are different in that the elastic restoring forces that are analogous to the spring constant k in the simple model are very different and are produced by different physical effects in these bodies. Each type of vibration will be discussed briefly, although the major emphasis will be placed on vibrations of and in machinery.

Torsional vibration. An engine-driven system (usually diesel or internal combustion powering a generator, a ship's propeller, or other load) has many degrees of freedom and, hence, many natural frequencies of which only the lowest two or three are of practical importance. The installation may, if excited by alternating components in its torque, execute torsional vibrations between its various parts of a magnitude of a fraction of a degree. This relatively small alternating motion is superposed on, and independent of, the continuous rotation of the engine shaft. Whereas the continuous rotation causes no extreme stress in the shaft, the small (quarter degree) angles of vibration wind up the shaft and relax it, causing alternating stresses that on many occasions have caused failure.

The torque on a crankshaft caused by a single cylinder and piston having one explosion for each one or two revolutions has a highly irregular time history containing many harmonics. (Harmonics are integral multiples of the fundamental frequency.) The torques developed by the various cylinders of a multicylinder engine combine in accordance with their times of occurrences. Thus, the exciting torque contains components of the firing frequency (one cycle per one or two revolutions) and multiples of that firing frequency, which are of practical importance up to the 16th multiple or harmonic. Because these 16 exciting or forced frequencies are proportional to the engine speed and the two or three natural frequencies are independent of it, and each exciting frequency can resonate with each natual one, there are as many as 48 resonances or critical speeds to be considered. Many of these will lie outside the habitual running range of the engine.

The designer may shift the severity of the resonances by using flexible couplings or extra flywheels in the engine, or by changing the arrangement of the cylinders, the V-angle of the engine, and the firing order. These suffice to make an installation satisfactory for any one running speed, but it is usually not possible to avoid all dangerous resonances in a wide range of running speeds. The designer must then resort to dampers to keep the amplitude of motion and the stress down. Various types of dampers exist; the most familiar is the pulley in the front of an automobile engine which drives the fan belt. This pulley is usually a small flywheel coupled to the engine shaft through a rubber insert. The assembly serves simultaneously as a torsional spring and as a dashpot damper. The flywheel is so tuned that it holds the principal torsional critical speed of the engine below dangerous torsional stress levels. Fatigue failures of crankshafts and of other shafting due to torsional vibration were common in the past, but such failures

are avoidable by proper design and are now rare. *See* COUPLING; FLYWHEEL; SHAFTING.

By the principle of action equals reaction, the gas torque acting on the piston-crankshaft assembly is equal and opposite to the gas torque acting on the engine frame. Hence, when the engine frame is rigidly attached to a foundation, that foundation experiences an alternating torque. To protect the foundation and to prevent the vibration from spreading through the structure, the engine is often mounted on metal springs or on rubber bushings. This is now universal practice with automobile and aircraft piston engines. *See* AUTOMOTIVE ENGINE.

Longitudinal vibration. A large ship's shafting excited by the propeller blades may come to resonance. The shafting and propeller must be designed to allow for such resonance. This is one cause of vibration in passenger liners. It occurs at a frequency of the shaft revolutions multiplied by the number of propeller blades. *See* PROPELLER (MARINE CRAFT).

The air or gas column in the suction or discharge lines of internal combustion engines acted upon by the piston motions can come to resonance at certain speeds. This principle has been used to increase the power of the engine by designing those lines so that during the intake more air than usual enters the cylinder, and during exhaust more air than usual goes out, thereby decreasing the back-pressure. *See* INTERNAL COMBUSTION ENGINE.

In internal combustion engines the gas pressure exerted on the piston and thence on a crank throw tends to lengthen the crankshaft. Because this effect is periodic, longitudinal vibrations in internal combustion engine shafts have been observed; however, they are of less importance than the torsional ones and become serious only if a longitudinal and a torsional natural frequency fall close together, causing coupling oscillations.

Lateral vibrations. The most important of all machinery vibrations are the lateral or bending vibrations of shafts and other rotors caused by the centrifugal force of unbalance. The unbalance consists of a very small deviation of the center of gravity of the rotor from the geometric axis connecting the bearing centers, or of a small angular deviation between that bearing center line and a principal axis of inertia of the rotor. Thus the excitation always has the frequency of the shaft rotation, and when that coincides with one of the natural bending frequencies of the rotor, there is resonance or a critical speed where the bending stresses in the shaft can be a hundred times higher than those caused by the centrifugal force of the unbalance directly.

Slow-speed machines are always designed to run well below the lowest or first critical speed, but for high-speed ones this often is not possible. Steam turbines and electric generators of large power stations have for many years been designed to run between the first and second critical speeds; the newest and largest units are between their second and third criticals. The flexibility of the supporting (nonrotating) bearings is an important factor in the calculation of critical speeds or natural frequencies of rotors, because the flexible bearing decreases the natural frequencies by some 10% on the average below those that would be present if rigid bearings were used. For rotors of a high diameter-length ratio such as in steam turbines with large-diameter disks mounted on a comparatively thin shaft, the effect of rotating inertia, sometimes called the gyroscopic effect of the disks, is an important consideration.

Besides the classical unbalance critical speed, which always has a vibration frequency equal to that of the rpm, a number of secondary critical speeds have been observed and explained, in which the frequency of vibration is a multiple of, usually twice, the rpm. The practical importance of these is secondary with respect to the ordinary critical speed.

Beam vibration. Whenever a part of a structure or the entire structure itself has a natural frequency resonating with an alternating excitation, nearby (or sometimes far removed) severe vibrations may result. A typical example is an unbalanced piston machine, such as an air compressor, which frequently causes objectionable vibration in some locations in the building where it is installed. Another example is the vibration of an entire ship in the mode of a free-free (totally unsupported) beam, excited by the propeller blade frequency.

With the advent of jet engines, the effect of high-intensity airborne noise on the very light structures of airplanes and missiles has become important. It is characteristic of jet noise, or indeed of all cases of turbulent flow, that the excitation is distributed continuously over a wide band of frequencies. This type of excitation is random and is playing an increasingly important role in the design of aircraft and missiles.

Membrane vibration. A tightly stretched skin, which has negligible bending stiffness like a drumhead, is a membrane. (The diaphragm in a telephone receiver possesses considerable bending stiffness, and is not stretched; hence technically it is a plate, although sometimes it is also called a membrane.) The theory of vibration of a membrane of circular shape has been known for a century and is one of the more beautiful illustrations of the mathematics of Bessel functions. The lowest frequency of vibration corresponds to a shape where the entire membrane bulges in and out, without nodal lines, the periphery remaining fixed. The higher modes of motion possess nodal lines, which are concentric circles or angularly equidistant diameters. The frequency formula for the circular membrane vibration is given in Eq. (6). Here T is the tension in the membrane,

$$f = \alpha\sqrt{Tg/\gamma r^2} \tag{6}$$

g is the acceleration due to gravity, γ is the surface weight density of the membrane, r is the radius, and a is a numerical factor having the values shown in the **table**.

Plate vibration. Plate vibrations are very important for a variety of reasons. They occur in large, flat, solid surfaces that are excited in planes normal to that surface. The reason that such vibrations are important in mechanical systems is that they occur frequently

Numerical factor for membrane frequency

Number of nodal circles	Number of nodal diameters 0	1	2	3
0	0.383	0.610	0.819	1.02
1	0.880	1.12	1.34	1.55
2	1.38	1.62	1.85	2.08

in practice and in addition they are very good acoustical radiators; that is, they couple extremely well with the air around them and produce significant quantities of vibratory energy that can be both felt and heard as acoustic signals. The basic mechanism by which a plate operates is similar to a beam in that the action consists of the bending of the plate about its middle surface, except that whereas a beam is a one-dimensional element with bending in only one direction, a plate is a surface that has bending in two orthogonal directions as well as a torsional or twisting type of bending at angles between these two orthogonal directions. Thus, more complicated curvature and surface interaction effects are involved.

Self-excited vibrations. In the cases discussed so far, the existence of an alternating exciting force (or torque) has been assumed which would continue to exist by itself even after the vibratory motion was prevented or stopped. These are called forced vibrations. There is another class of motions, called self-excited, in which the exciting force is generated by the vibrating motion itself and hence disappears with that motion.

A familiar mechanical example of self-excited vibration is the piston of an engine. The back-and-forth motion (vibration) is maintained by an alternating gas or steam pressure steered by the valve mechanism. The initial source of energy is without alternating properties (the gasoline supply or the boiler steam) but is made alternating by the valves, which are operated by the vibratory piston itself.

A familiar electrical example is the self-oscillating electronic circuit which has a steady source of energy (for example, a battery), and a transistor that plays the role of the valve.

The oldest practical examples of useful self-excited vibrations are musical instruments. The violin operates on the behavior of the friction between the rosined bow and the string, which has aptly been described as stick-slip friction. Vibrations of this type have appeared repeatedly and still are often met within machinery with insufficient lubrication; the most vexing and difficult case is that of the chattering machine tool cutter which leaves a wavy cut instead of a smooth one. *See* MUSICAL INSTRUMENTS.

The clarinet operates on the passage of air (from the mouth to the instrument) though a narrow leakage opening whose width varies periodically with the vibration of the reed. Serious vibrations of this character occur in steam, gas, or hydraulic turbines, heat exchangers, and other apparatus in which a fluid or gas passes through narrow passages. Other self-excited vibrations are the pulsating flow sometimes observed in fans and blowers, and the shimmying motion of wheels, which has been a serious problem in automobiles and in the landing gear of aircraft. Self-excited vibrations appear in autopilots of aircraft and missiles, and in general are apt to occur in servomechanisms with high gain.

A class of dangerous vibrations is provided by the various phenonena of flutter, whereby an elastic system becomes self-excited in the stream of air, gas, or fluid of sufficient speed. This aviation problem first arose in airplane wings, but as speeds increased, other parts of airplanes and other machinery, such as turbine blades, displayed flutter. Flutter theory has grown into a subject by itself. *See* AEROELASTICITY; FLUTTER (AERONAUTICS).

Serious vibrations have occurred in rocket engines and other types of combustion chambers, whereby the combustion becomes unstable and the burning gas mass enters a state of self-excited vibration. These phenomena have been known for a century and still are not completely understood.

Effect of vibrations. The most serious effect of vibration, especially in the case of machinery, is that sufficiently high alternating stresses can produce fatigue failure in machine and structural parts. Less serious effects include increased wear of parts, general malfunctioning of apparatus, and the propagation of vibration through foundations and buildings to locations where the vibration or its acoustic realization is intolerable either for human comfort or

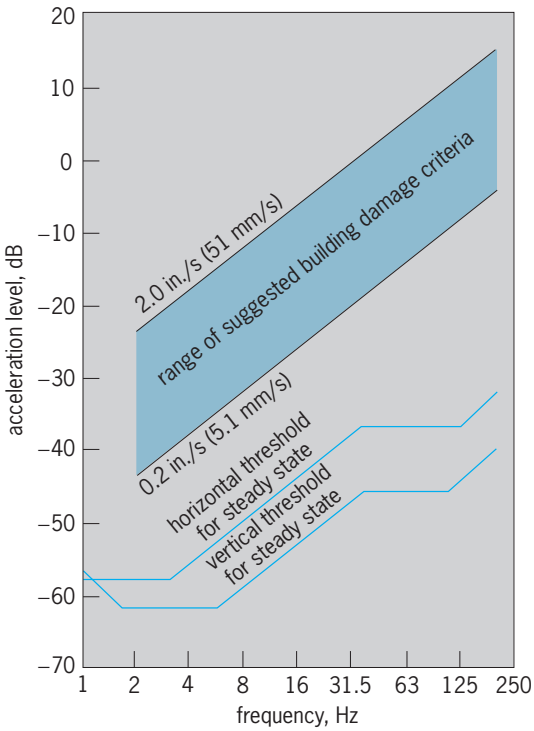

Fig. 4. Comparison of suggested building damage criteria and human threshold of perception curves for whole-body vibration. Acceleration level in dB = 20 log acceleration level/*g*. Because they are negative logarithms, the decibel values on the vertical axis indicate increasing acceleration levels from bottom to top. (*After T. G. Gutowski, L. E. Wittig, and C. L. Dym, Some aspects of the ground vibration problem, Noise Control Eng., 10(3):94–100, 1978*)

for the successful operation of sensitive measuring equipment. *See* STRESS AND STRAIN; WEAR.

Whole-body vibration. **Figure 4** shows a comparison of suggested building damage criteria and human threshold-of-perception curves for whole-body vibration. The threshold curves for human perception are based upon tests wherein people are seated on shakers, so the perception is of motion that affects the entire body. There are a variety of ranges of vibration, and these ranges are strongly dependent upon the frequency (expressed here in hertz, equal to $\omega/2\pi$). The ordinate of this graph is expressed as an acceleration level (in decibels) with reference to g (the acceleration of gravity) as the base acceleration level. The decibel notation indicates a logarithmic dependence such that an acceleration level of −20 dB represents an acceleration level that is 0.1 g in magnitude, an acceleration of −40 dB represents a level that is 0.01 g in level, and so on. The magnitudes of vibration that are felt by humans are extremely small; that is, on the order of −60 dB or 0.001 g in magnitude. While identifying the human threshold of perception from whole-body laboratory experiments is very useful, other factors must be considered in practice. Human annoyance may be related to secondary vibratory effects such as the vibration-induced motion of household items such as plants and mirrors, and so on. Thus, human annoyance could depend on the observer's location and activity as well as other suggested factors (Fig. 4). The values for which damage to a building might be realized are significantly higher than those for which vibration is perceptible by humans. In fact, the difference tends to be on the order of two orders of magnitude (a multiplicative factor of 100 in amplitude).

Generation of acoustic noise. **Figure 5** relates vibration magnitudes both to perceptible vibration level and to audible acoustical levels. These data demonstrate that an acoustic signal can be generated by a mechanical vibration and produce auditory noise that can be very uncomfortable. For example, as discussed above, noise can propagate in a building due to an air-conditioning compressor. The curves displayed in Fig. 5 also show so-called NC-equivalent curves used to describe architectural levels of quietness in a design environment. Thus an NC-20 equivalent curve is the design standard for a first-class concert hall, while an NC-40 criterion would be adequate for a grade-school classroom. The acoustic signals are normally expressed in logarithmic amplitudes of acoustic pressure, but are expressed here in terms of amplitudes of vibration, that is, acceleration levels. An equivalence between the two is based on the recognition that when a large surface vibrates, it simultaneously sets into motion the acoustic medium around it, whether it be air or water. The bending vibrations of plates, for example, couple extremely well with the compressional waves of a surrounding acoustic medium. Thus when a building wall or floor is set into motion by a mechanical source of vibration, it in turn causes a very effective propagation of an acoustical signal, or noise, in the surrounding air. Another significant example of good coupling is the mechanical

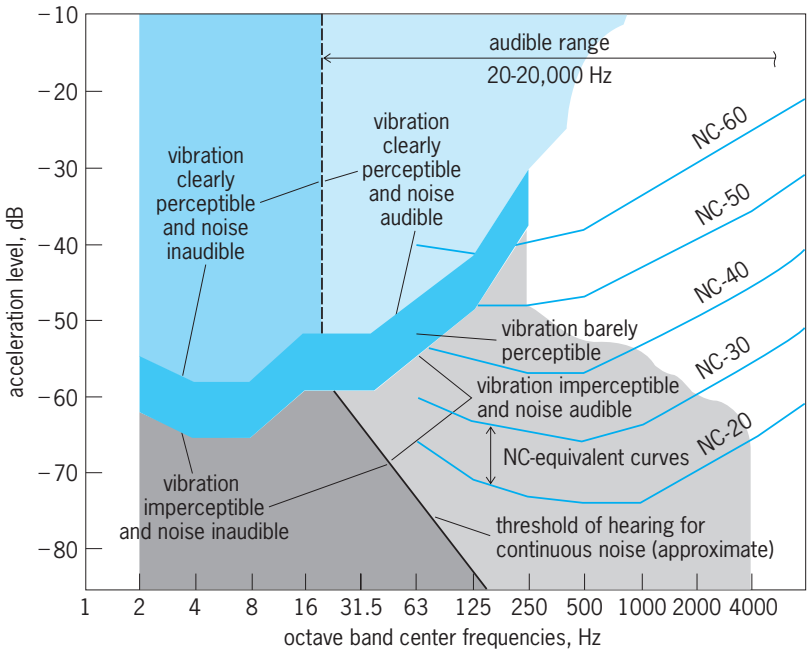

Fig. 5. Levels of perceptibility of vibration and audibility of noise. Acceleration level in dB = 20 log acceleration level/*g*. Because they are negative logarithms, the decibel values on the vertical axis indicate increasing acceleration levels from bottom to top. (*After C. L. Dym and D. Klabin, Architectural implications of structural vibration, Architect. Rec., 157(9):125–127, 1975*)

excitation of a drumhead by a drumstick. The lateral vibrations of the stretched membrane (the drumhead) also couple well with surrounding air to produce the corresponding tympanic sound of the drum which can be heard at great distances from the drum itself.

There are many factors that enter into the coupling of a vibrating structural element and the surrounding acoustical medium. The principal factors are, however, the surface area of the mechanical source that is exposed to the medium and the nature of the mechanical vibration, that is, whether it is perpendicular to the surface (as with bending vibrations of beams and plates or lateral vibrations of membranes) or along the direction of the surface (as in torsional and longitudinal vibration). It is the perpendicular motion of large surfaces that produces the best acoustical coupling and, hence, the best chance of radiation of noise from vibration. *See* ACOUSTIC NOISE; VIBRATION.

Clive L. Dym; J. P. Den Hartog

Bibliography. G. S. Kelly, *Fundamentals of Mechanical Vibrations*, 2d ed., 2000; M. Lalanne and P. Bertnier, *Mechanical Vibrations for Engineers*, 1984; S. S. Rao, *Mechanical Vibrations*, 3d ed., 1995; R. F. Steidel, *An Introduction to Mechanical Vibrations*, 3d ed., 1989.

Mechanics

In its original sense, mechanics refers to the study of the behavior of systems under the action of forces. Statics deals with cases where the forces either produce no motion or the motion is not of interest.

Dynamics deals properly with motions under forces. Mechanics is subdivided according to the types of systems and phenomena involved.

An important distinction is based on the size of the system. Those systems that are large enough can be adequately described by the newtonian laws of classical mechanics. In this category, for example, are celestial mechanics, the study of the motions of planets, stars, and other heavenly bodies, and fluid mechanics, which treats liquids and gases on a macroscopic scale. Fluid mechanics is a part of a larger field called continuum mechanics or (by some physicists) classical field theory, involving any essentially continuous distribution of matter, whether it be rigid, elastic, plastic, or fluid. On the other hand, the behavior of microscopic systems such as molecules, atoms, and nuclei can be interpreted only through the concepts and mathematical methods of quantum mechanics.

From its inception, quantum mechanics had two apparently different mathematical forms—the wave mechanics of E. Schrödinger, which emphasizes the spatial probability distributions in the quantum states, and the matrix mechanics of W. Heisenberg, which emphasizes the transitions between states. These are now known to be equivalent.

Mechanics may also be classified as nonrelativistic or relativistic mechanics, the latter applying to systems with material velocities comparable to the velocity of light. This distinction pertains to both classical and quantum mechanics.

Finally, statistical mechanics uses the methods of statistics for both classical and quantum systems containing very large numbers of similar subsystems to obtain their large-scale properties. *See* CELESTIAL MECHANICS; CLASSICAL FIELD THEORY; CLASSICAL MECHANICS; DYNAMICS; FLUID MECHANICS; QUANTUM MECHANICS; RELATIVISTIC MECHANICS; STATICS; STATISTICAL MECHANICS.

Bernard Goodman

Mechanism

Classically, a mechanical means for the conversion of motion, the transmission of power, or the control of these. Mechanisms are at the core of the workings of many machines and mechanical devices. In modern usage, mechanisms are not always limited to mechanical means. In addition to mechanical elements, they may include pneumatic, hydraulic, electrical, and electronic elements. In this article, the discussion of mechanism is limited to its classical meaning. *See* MACHINE.

Mechanisms are found in internal combustion engines; compressors; locomotives; agricultural, earthmoving, excavating, mining, packaging, textile, and other machinery; machine tools; printing presses; engraving machines; transmissions; ordnance equipment; washing machines; lawn mowers; sewing machines; projectors; pinspotters; toys; pianos; ski bindings; artificial limbs; door locks; nutcrackers; counters; microswitches; speedometers; and innumerable other devices.

Components. Most mechanisms consist of combinations of a relatively small number of basic components. Of these, the most important are cams, gears, links, belts, chains, and logical mechanical elements.

Cams and cam followers. A cam is a specially shaped part designed to impart a prescribed law of motion to a contacting part called the follower. In **Fig. 1** the cam is in the form of a disk, the shape of which governs the motion of the follower. Followers may be translating (Fig. 1*a* and *b*) or swinging (Fig. 1*c*) and may be in sliding (Fig. 1*a*) or rolling (Fig. 1*b* and *c*) contact with the cam. The law of motion is the relation between cam rotation and follower displacement. Contact between cam and follower is often (but not necessarily) maintained by a return spring. *See* CAM MECHANISM.

Gears. Gears are toothed wheels that provide a positive connection between rotating shafts. The most familiar are spur gears, which are used to connect parallel shafts. The teeth of straight spur gears (**Fig. 2***a*) may be imagined mounted on right circular cylinders. When the shaft axes are intersecting, bevel gears can be used. The teeth of the straight bevel gears (Fig. 2*b*) may be imagined mounted on right circular cones. Gear teeth are so shaped that the angular-velocity ratio of the connected shafts is constant. Other forms of gearing include helical (these may be spur gears or crossed helical gears), worm, and hypoid. The latter two, as well as crossed helical gears, are used to connect shafts with nonparallel, nonintersecting axes. *See* GEAR.

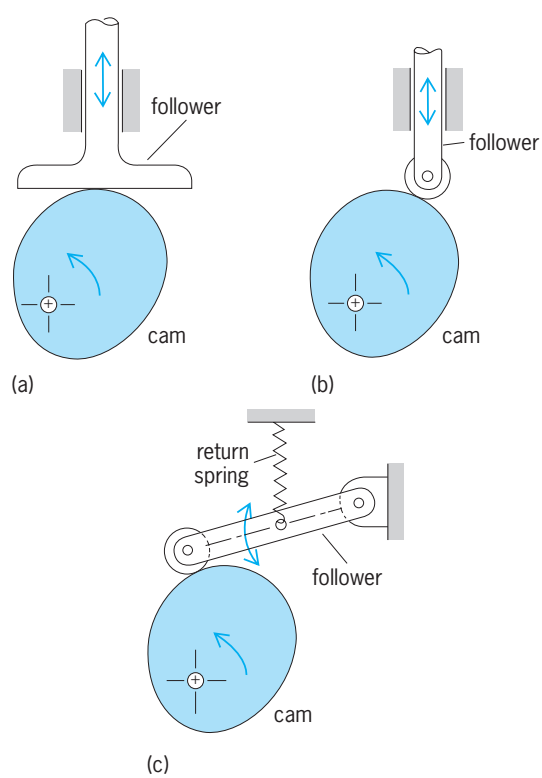

Fig. 1. Cam-follower configurations. (*a*) Disk cam with flat-faced translating follower. (*b*) Disk cam with translating roller follower. (*c*) Disk cam with swinging roller follower.

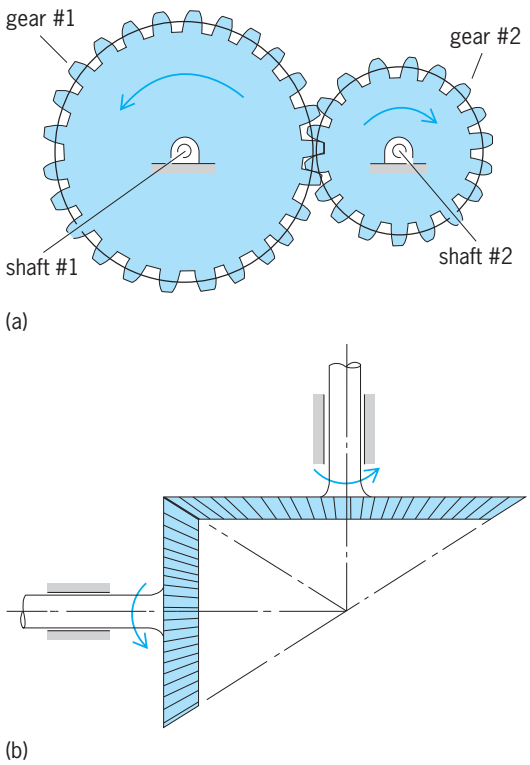

Fig. 2. Gear configurations. (*a*) Spur gears. (*b*) Bevel gears.

Links. For practical purposes, a link may be defined as a rigid body which is connected to other rigid bodies by joints. The most common joints (**Fig. 3**) are the pin joint (symbolized by R), the sliding joint (P), the cylindrical joint (C), the ball joint (S), and the screw joint (H). The tooth connection between gears and the contact between cam and follower are also usually regarded as joints. Typical link configurations are shown in **Fig. 4**. *See* LINKAGE (MECHANISM).

Belts and chains. These flexible connectors include pulleys and flat belts (**Fig. 5**), timing belts and sprockets, and chain belts and sprockets. The first of these is a friction drive, while the latter two are positive drives. All transmit a practically constant angular-velocity ratio between the connected shafts. *See* BELT DRIVE; CHAIN DRIVE.

Logical mechanical elements. These include ratchets, trips, detents, interlocks, and the like. The most interesting of these is the ratchet. **Figure 6** shows a type of ratchet mechanism known as an escapement. The ratchet is the toothed wheel, the rotation of which is controlled by a pawl—a pivoted link which may be in or out of engagement with the ratchet. This arrangement can provide a stop-and-go unidirectional motion, such as that of the hands of a pendulum-driven clock. *See* ESCAPEMENT; RATCHET.

Representative mechanisms. Mechanical components can be combined into mechanisms in an infinite variety of ways. For example, in automobiles there is a slider-crank mechanism (**Fig. 7**), connecting piston to crankshaft; the cam-follower system (**Fig. 8**) for the actuation of the intake or exhaust valves in an overhead valve arrangement; and the differential (**Fig. 9**). The first is a linkage mechanism, the second a cam-follower mechanism, and the third a gear train. The first converts the translation of the piston into crankshaft rotation and transmits power from piston to crankshaft. The second controls the opening and closing of an intake or exhaust valve. Although some power is required to actuate the valve, this is of secondary significance. The differential is a two-degrees-of-freedom gear train, which couples the driveshaft to the tire axles and transmits engine power to the tires. The two degrees of freedom are needed to permit unequal tire speeds when negotiating a turn. In order to understand these and other mechanisms, their degree of freedom, structure, and kinematics must be considered. *See* SLIDER-CRANK MECHANISM.

Degree of freedom. This is conveniently illustrated for mechanisms with rigid links. The discussion is limited to mechanisms which obey the general degree-of-freedom equation (1), where F = degree

$$F = \lambda(l - j - 1) + \Sigma f_i \qquad (1)$$

of freedom of mechanism, l = number of links of mechanism, j = number of joints of mechanism, f_i = degree of freedom of relative motion at ith joint (see Fig. 3), Σ = summation symbol (summation over all joints), and λ = mobility number (the most common cases are $\lambda = 3$ for plane mechanisms and $\lambda = 6$ for spatial mechanisms).

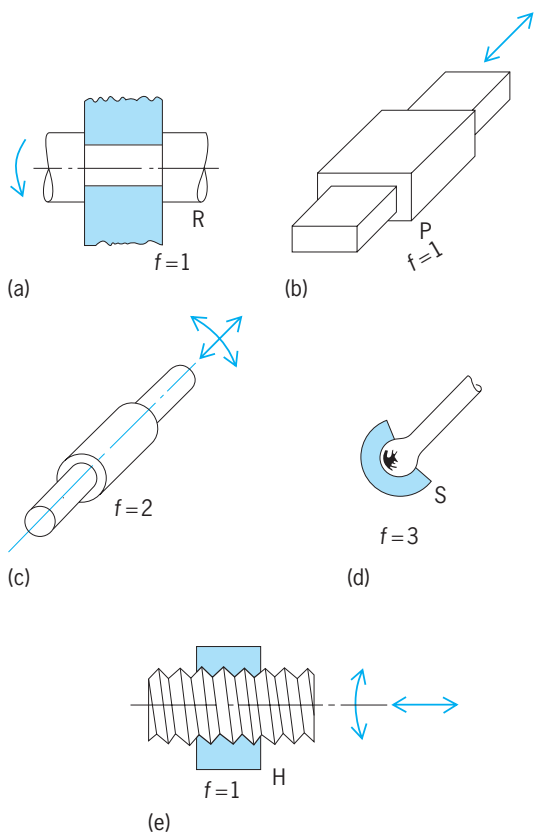

Fig. 3. Common joints, including joint symbols and degree of freedom *f* of relative motion at joint. (*a*) Pin joint. (*b*) Sliding joint. (*c*) Cylindrical joint. (*d*) Ball joint. (*e*) Screw joint.

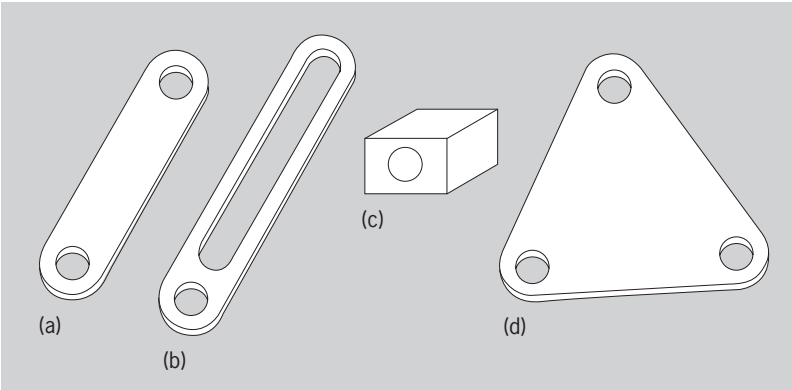

Fig. 4. Typical links. (*a*) Crank. (*b*) Slotted link. (*c*) Slider with provision for pin connection. (*d*) Triangular link.

There are mechanisms, some of them highly significant, which do not obey the general degree-of-freedom equation. The degree of freedom of such mechanisms is generally strongly dependent on their dimensions. Such mechanisms are excluded from the discussion which follows.

For plane mechanisms with pin joints, Eq. (2) ap-

$$F = 3l - 2j - 3 \tag{2}$$

plies, while for gear trains with j_r pin joints and j_g gear meshes Eqs. (3) apply, where L_{ind} denotes the

$$F = j_r - j_g \qquad j_g = L_{\text{ind}} \qquad l = F + L_{\text{ind}} + 1 \tag{3}$$

number of independent loops of the mechanism and is given by Eq. (4) for mechanisms in which the

$$L_{\text{ind}} = j - l + 1 \tag{4}$$

graphs can be drawn without crossing edges. *See* DEGREE OF FREEDOM (MECHANICS).

Kinematic structure. The kinematic structure of a mechanism refers to the identification of the joint connection between its links. Just as chemical compounds can be represented by an abstract formula and electric circuits by schematic diagrams, the kinematic structure of mechanisms can be usefully represented by abstract diagrams.

Mathematical representation. The structure of mechanisms for which each joint connects two links can be represented by a structural diagram, or graph, in which links are denoted by vertices, joints by edges, and in which the edge connection of vertices corresponds to the joint connection of links; edges are labeled according to joint type (and pin axis in the case of gear trains), and the fixed link is identifed as well. Thus the graph of the slider-crank mechanism of Fig. 7*a* is as shown in Fig. 7*b*. In this figure the circle around vertex 1 signifies that link 1 is fixed.

Mechanisms can be considered identical or different, depending on whether their graph representations are identical (isomorphic) or different (nonisomorphic).

Classification and creation of mechanisms. Mechanisms can be classified and enumerated according to their kinematic structure, with the restrictions on their structure for many mechanisms being given by Eqs. (1)–(4).

It can be shown, for example, that when Eqs. (1)–(4) are applicable, there are seven plane single-degree-of-freedom mechanisms with four links and pin joints and up to two sliding joints. The structures of plane pin-jointed single-degree-of-freedom linkages with up to ten members and the structures of single-degree-of-freedom gear trains with up to six members have been determined as well.

The creation of mechanisms according to kinematic structure is also known as type synthesis or structural synthesis. It is based on the separation of kinematic structure from function. Thus, if, for example, mechanisms are sought which function as shaft couplings, the potential kinematic structures of mechanisms having the required degree of freedom, number of parts, and admissible types of joints are considered. The mechanisms corresponding to these structures are then screened by functional considerations. The result is an unbiased collection of potentially useful shaft-coupling configurations, complete within the specification of the search. This procedure can be a useful tool in the conceptual and inventive phases of mechanical design, supplementing atlases and collections of known mechanisms.

Kinematics. The subject is divided into kinematic analysis (analysis of a mechanism of given dimensions) and synthesis (determination of the proportions of a mechanism for given motion requirements). It includes the investigation of finite as well as infinitesimal displacements, velocities, accelerations and higher accelerations, and curvatures and higher curvatures in plane and three-dimensional motions.

Displacements in mechanisms can be found graphically or analytically from the conditions of closure

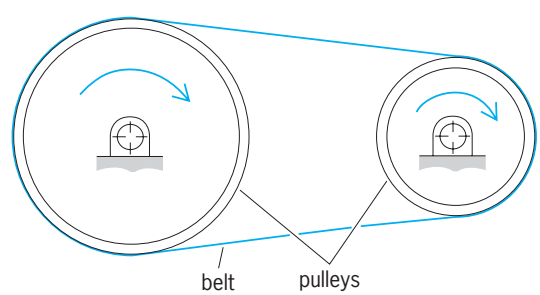

Fig. 5. Belt and pulleys.

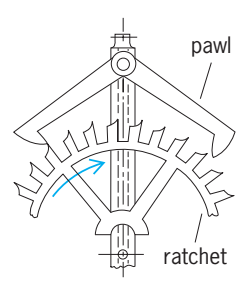

Fig. 6. Escapement.

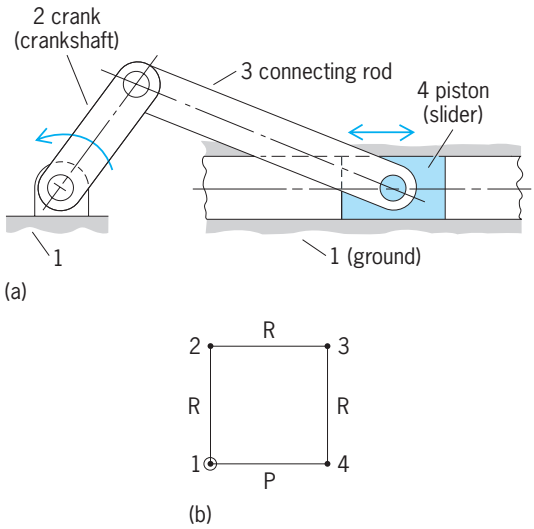

Fig. 7. Slider-crank mechanism. (*a*) Mechanism. (*b*) Graph of mechanism.

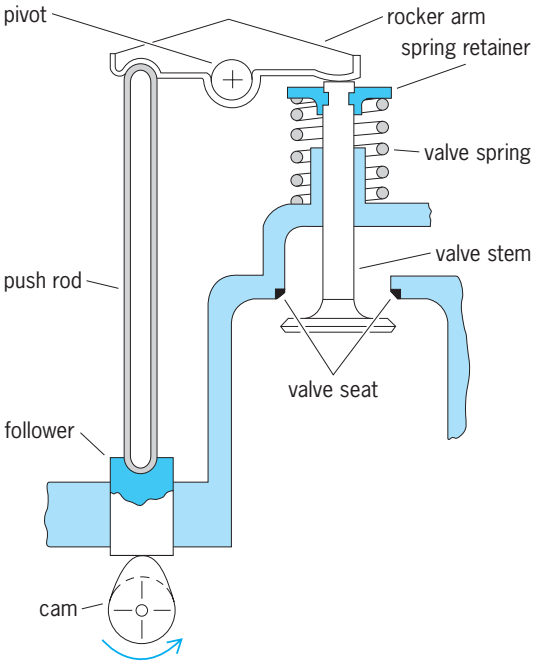

Fig. 8. Valve train of an internal combustion engine. (*Courtesy of Albert Pisano*)

of the independent loops of the mechanism. Velocities, accelerations, and higher-motion derivatives can be found by differentiation. It can be shown that at any instant in a general, continuous plane motion (the term general excludes translations, for example), there is one point for which the velocity vanishes (the instant center), one point for which the acceleration vanishes (the acceleration center), one point for which the acceleration derivative vanishes, and so on. The theory of instantaneous invariants has proved a powerful tool in such investigations.

For the position guidance of a rigid body through a prescribed sequence of finitely separated coplanar positions, one is interested in the loci of points in the body or the frame which may be used to attach cranks or sliders, so that the body can be guided mechanically. For four arbitrary positions these loci include the circlepoint curve (locus of points in the moving body, the four corresponding positions of which lie on a circle), the centerpoint curve (locus of the centers generated by points of the circlepoint curve), and the Ball point (that point in the moving system, the four corresponding positions of which are collinear). Circlepoints and their corresponding centerpoints can be used for locating moving and fixed crankpins, and the Ball point can be used for the attachment of a slider. In this way one can determine the proportions of linkage mechanisms, which can guide a rigid body through a prescribed sequence of positions.

The circlepoint and centerpoint curves are cubic algebraic curves. These curves can be used in the synthesis of linkage mechanisms for a considerable variety of motion requirements. The theory has been extended to five general coplanar positions and to three-dimensional motions. The optimization of link age proportions so as to obtain the closest match of a linkage motion to a desired motion has also been the subject of many investigations, especially with the aid of Chebyshev approximations.

Useful concepts, which occur frequently in kinematic analysis, include kinematic inversion (the process of holding different links of a mechanism fixed), relative motion, toggle positions (positions providing a very high mechanical advantage), lockup positions (positions in which the mechanism cannot be driven from its usual input member), pressure angles and transmission angles (which govern the efficiency of force and power transmission), angular-velocity ratios, and angular-acceleration rations and shock

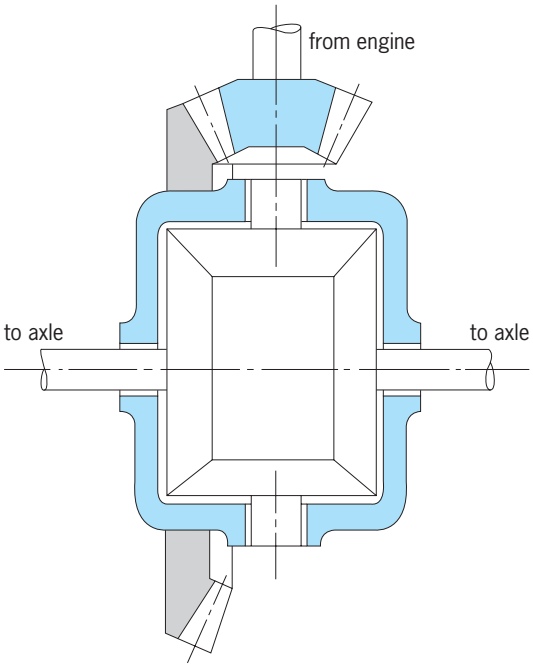

Fig. 9. Automotive differential.

(the derivative of acceleration with respect to time). *See* KINEMATICS.

Design. The design of mechanisms involves many factors. These include their structure, kinematics, dynamics, stress analysis, materials, lubrication, wear, tolerances, production considerations, control and actuation, vibrations, critical speeds, reliability, costs, and environmental considerations.

Compared to other areas of engineering design, the design of mechanisms is complicated by their generally high degree of nonlinearity (with the exception of gear trains), the inability in most instances to limit their operation to small displacements, and the variety of combinations of components and mechanical elements. But, this complexity also provides a challenge to the creativity of the engineer and inventor.

Current trends in the design of mechanisms emphasize economical design analysis by means of computer-aided design techniques, which offer ever-increasing speeds and more sophisticated performance requirements while maintaining reliability and economy of space, weight, and cost. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Ferdinand Freudenstein

Bibliography. A. Erdman and G. N. Sandor, *Mechanisms Design: Analysis and Synthesis*, vol. 1, 2d ed., 1990; F. D. Jones, H. L. Horton, and J. A. Newell (eds.), *Ingenious Mechanisms for Designers and Inventors*, 4 vols., 1968–1971; J. McCarthy, *Introduction to Theoretical Kinematics*, 1990; S. B. Tuttle, *Mechanisms for Engineering Design*, 1967.

Mechanoreceptors

Receptors that provide the organism with information about such mechanical changes in the environment as movement, tension, and pressure. In higher animals receptors are actually the only means by which information of the surroundings is gained and by which reactions to environmental changes are started. *See* SENSATION.

Receptor-effector mechanisms. In primitive forms of life, such as unicellular organisms and sponges, the receptor and effector functions are built into the same cell (**Fig. 1***a* and *b*). The cell is directly excited by and reacts to the environmental stimulus. At higher levels of organization, the receptor separates from the effector and appears specialized in function and structure in the receptor cell of the primitive nerve system of actinians (sea anemones). Stimuli of the environment excite the receptor cell, and this excitation is conveyed to another specialized cell, the effector (Fig. 1*c*). In another step of phylogenetic development, a third element appears, the motor nerve cell, which serves as a link between the receptor and effector. This arrangement appears clearly in coelenterates and is carried through in its essential features up to the vertebrates (Fig. 1*d* and *e*).

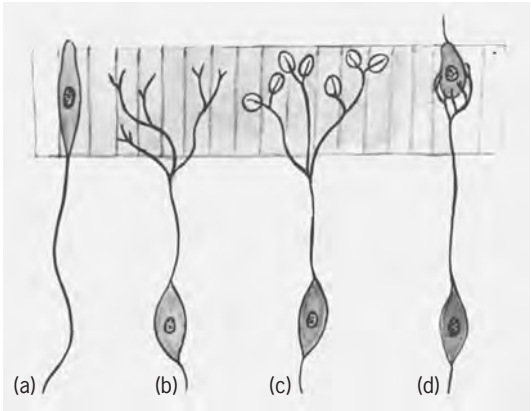

Fig. 2. Differentiation of receptors. (*a*) Primary receptor cell at the periphery. (*b*) Nerve cell buried inside the organism sending out branches which terminate as nude receptor endings. (*c*) Encapsulated receptor endings. (*d*) Secondary receptor cell.

Nerve endings. The primitive receptor cell may remain at the periphery, as in the olfactory receptors of the vertebrates, or it may be buried inside the organism and send protoplasmic branches out to the periphery. A branch of this type, a dendrite or an axon, terminates in its simplest form as a nude ending at the periphery. In more complicated forms, in Pacinian and Meissner corpuscles, the ending is surrounded by accessory tissue of varying structural organization; or the ending appears to terminate on a so-called secondary sense cell, as in certain gustatory, auditory, and visual receptors (**Fig. 2**). But in all receptors known, the nerve ending is invariably present, however complex the accessory structure around it (**Fig. 3**). In the two mechanoreceptors whose fine structure has been examined under the electron microscope (the Pacinian corpuscle and the muscle spindle), the nerve ending is nude, lacking not only a sheath of myelin but one of Schwann cells (**Fig. 4**).

Phylogenetic advances. Survival of an organism depends in large part upon its ability to react to changes in the environment. As organized life climbs up the phylogenetic scale, it develops and subsists within increasing limits of environmental change. In general, the greater the organizational level of the animal, the greater are the environmental changes to which

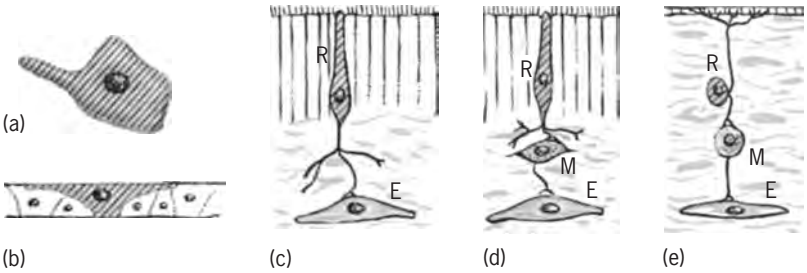

Fig. 1. Evolution of receptor-effector systems. (*a*) Receptor-effector cell of unicellular organism (ameba); (*b*) receptor-effector cell of multicellular organism (sponge); (*c*) specialization shown in separation of receptor and effector cells of actinians (tentacle of sea anemone); (*d*) receptor, motor nerve, and effector cells of coelenterates; (*e*) receptor motor nerve, and effector cells in vertebrates. R = receptor; E = effector; M = motor nerve cell. (*b, c, and d after G. H. Parker, The Elementary Nervous System, Lippincott, 1919*)

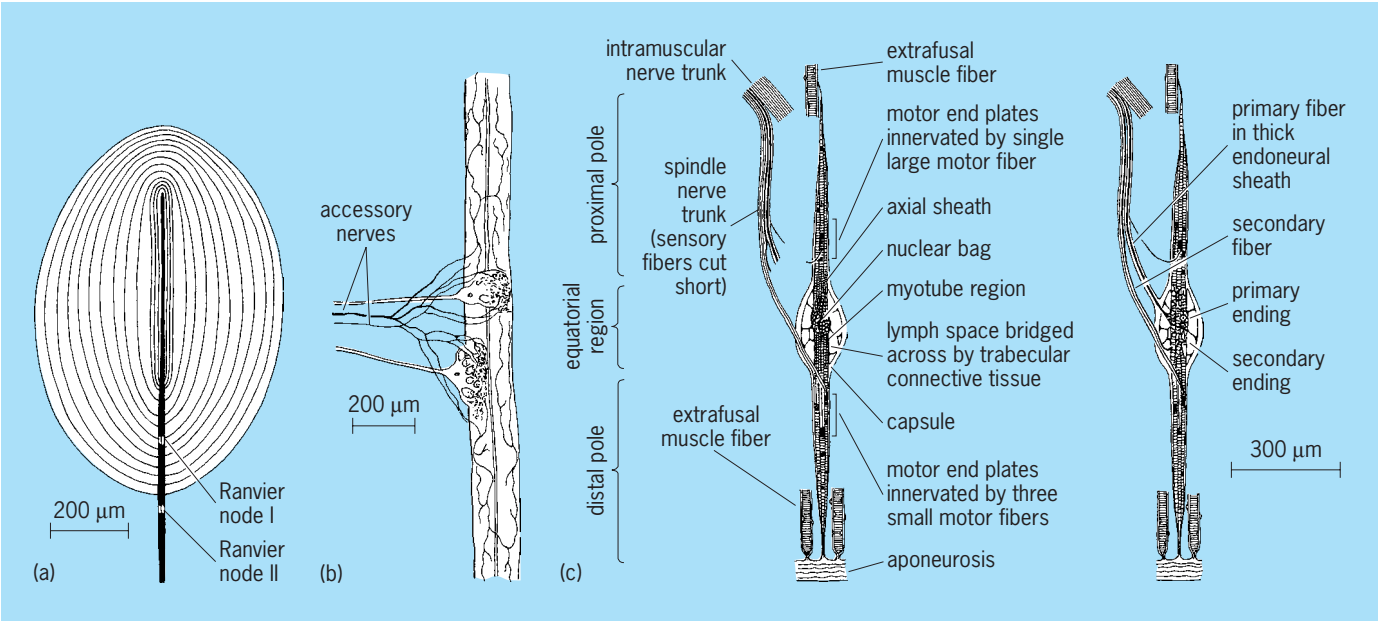

Fig. 3. Schemes of three typical mechanoreceptors. (*a*) Pacinian corpuscle of the cat's mesentery and skin. (*b*) Stretch receptor nerve cells of the abdominal musculature of lobsters (*Homarus vulgaris*) with dendritic endings embedded in muscle strands (*after J. S. Alexandrowicz, Quart. J. Microsc. Sci., 92:163–199, pt. 2, 1951*). (*c*) Muscle spindle of the rabbit (*after D. Barker, Quart. J. Microsc. Sci., 89(2):143–186, 1948*).

it must react and the more varied in form and in degree of functional organization become its receptor organs. Receptors for mechanical changes of the environment (mechanoreceptors) had to develop at an early stage of phylogeny. They were necessary for the avoidance of mechanical objects, for the sensing of water waves transmitted from other animals in the sea, for orientation in space, and for maintenance of a position with respect to gravity. Examples are the lateral line receptors and the later development of the labyrinth receptors. The development of life on land brought development of mechanoreceptors sensitive to sound vibration of the air. With the growth of internal organs with specialized functions and, particularly, with the development of fast regulatory mechanisms for these functions arose the necessity for receptors sensitive to mechanical stimuli originating inside the organs (proprioceptors). Thus, in the vertebrates, mechanoreceptors appear in all organs in which passive or active movements occur, feeding information into the nerve system about movements, tension, or pressure. Mechanoreceptors appear in the digestive tract, lungs, blood vessels, heart, and skeletal musculature. The simple mechanoreceptors of the integument of primitive invertebrates are bare nerve endings in all animals, including the higher vertebrates; they are ubiquitous in the aforementioned organs and in the skin. In addition, around some nerve endings, special enclosures of varying complexity are developed in the skin and other structures of the higher vertebrates (Fig. 3).

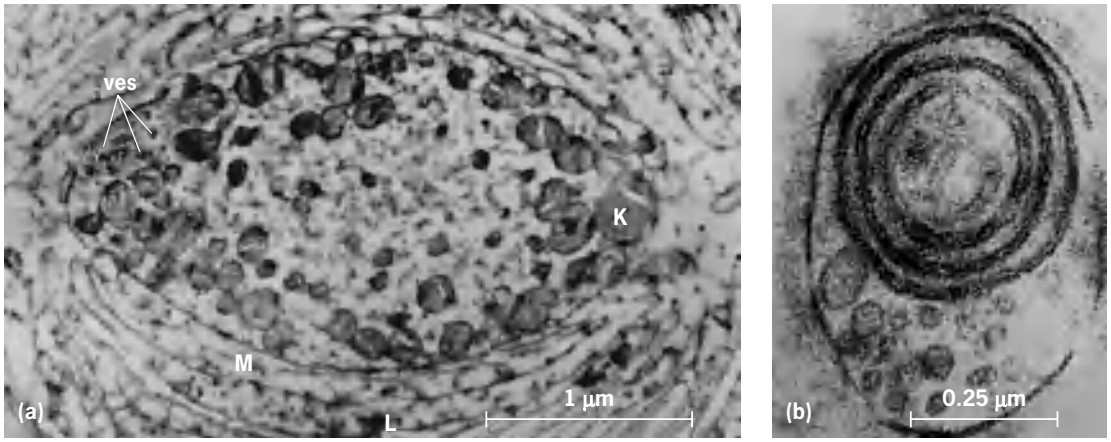

Fig. 4. Electron micrographs showing fine structure of mechanoreceptors. (*a*) Nerve ending of Pacinian corpuscle in the cat, containing vesicular formations (ves), and abundant mitochondria (K) near the ending's membrane (M). L = lamellae of the inner core (*from D. C. Pease and T. A. Quilliam, J. Biophys. Biochem. Cytol., 3(3):331–342, 1957*). (*b*) Nerve ending of the muscle spindle of a frog (*from J. D. Robertson, Biochem. Soc. Symp. no. 16, 3–43, 1959*).

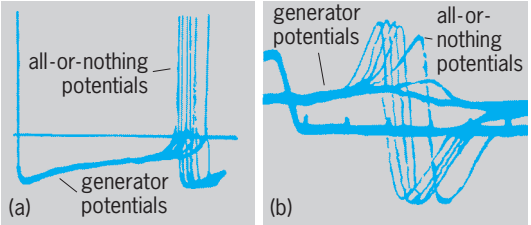

Fig. 5. Generation of nerve impulses. Several responses superposed on successive oscilloscope sweeps. Note firing of impulses from critical level. (*a*) Crustacean stretch receptor (*after C. Eyzaguirre and S. W. Kuffler, J. Gen. Physiol., 39(1):87–119, 1955*). (*b*) Generation in Pacinian corpuscle (cat) (*after J. A. B. Gray and M. Sato, J. Physiol., 122(3):610–636, 1953*).

Mechanoelectric conversion. Mechanoreceptors are excited by mechanical disturbances of their surroundings through deformation of their structure, through pressure or tension, or through a combination of these. In general, little energy is required for mechanical stimuli to cause a detectable excitation in mechanoreceptors. For example, the receptor membrane of Pacinian corpuscles of the cat's mesentery gives detectable generator currents with displacement of the outer layers of its capsule by 10^{-5} cm.

Transducer sequence. From a physical point of view, mechanoreceptors are energy transducers; they convert mechanical into electrical energy, which in turn triggers the nerve impulse. The mechanisms of nerve-impulse production have been studied most extensively in the muscle spindle of frog, in the stretch receptor of Crustacea, and in the Pacinian corpuscle. In the three receptors studied, deformation leads to a sequence of events which may be summarized by the following scheme:

$$\text{Mechanical stimulus} \rightarrow \text{generator current} \rightarrow \text{nerve impulse (actionpotential)}$$

The generator current is the earliest detectable sign of excitation. In Pacinian corpuscles it follows the mechanical stimulus at a fixed and brief latency. The most salient characteristic of the generator current is its graded nature; its amplitude increases continuously, without visible steps, if the stimulus strength is progressively increased. When the generator current reaches a certain critical amplitude, an all-or-nothing potential is discharged in the sense organ which may then propagate as an all-or-nothing nerve impulse along the afferent axon of the receptor (**Fig. 5**).

Bioelectric generator and nerve-impulse initiation sites. Generator and all-or-nothing potentials arise at differently behaving membrane sites. The capsular tissue around the nerve ending, constituting more than 99.9% of the total mass of a Pacinian corpuscle, does not partake actively in the transducer chain outlined previously (**Fig. 6**). Figure 6 illustrates an experiment in which the capsule of a Pacinian corpuscle was removed (Fig. 6*a–c*). The preparation was stimulated with two successive mechanical stimuli, the first producing an all-or-none potential, the second (subthreshold) a generator current. In Fig. 6*c* mechanosensitivity remains unimpaired in spite of the removal of more than 99.99% of the corpuscle. In Fig. 6*d* and e, although the preparation consisted only of a partly isolated ending, it continued to be as mechanosensitive as before. This would probably also occur in other mechanoreceptors whose nerve

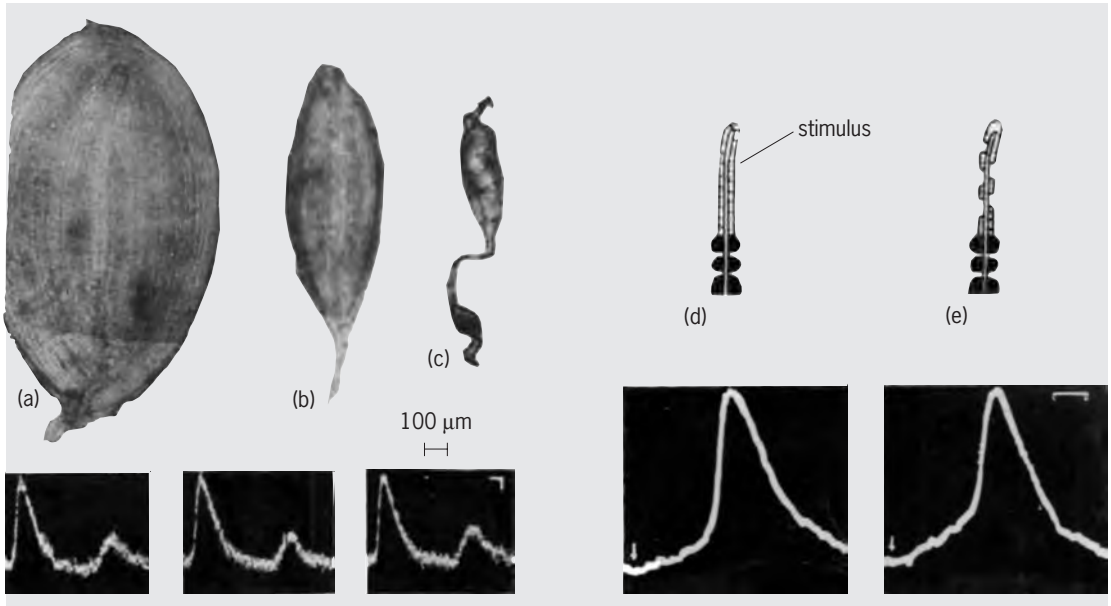

Fig. 6. Localization of transducer element. (*a*) Capsule of intact Pacinian corpuscle. (*b*) Progressive removal of capsule. (*c*) Only ending and surrounding inner core remain; calibration: 1 ms, 25 μV. (*d,e*) Partial destruction of remaining inner core; calibration: 1 ms. Recordings at bottom are generator and action potentials for each stage. (*From W. R. Loewenstein and R. Rathkamp, J. Gen. Physiol., 41(6):1245–1265, 1958*)

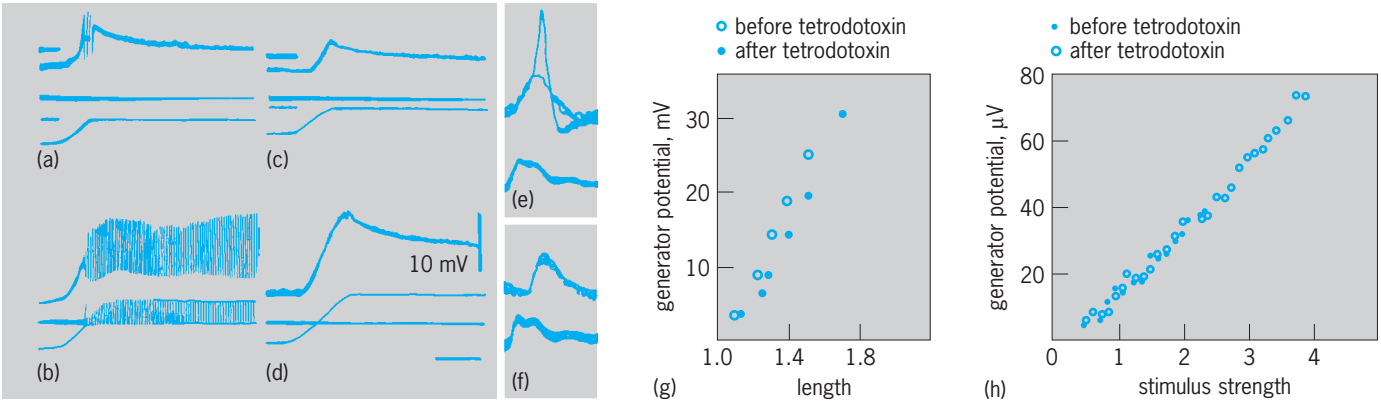

Fig. 7. Effects of tetrodotoxin on (*a–f*) transducer and impulse processes and (*g,h*) generator potentials of receptor organs. (*a–d*) Crayfish stretch receptor neuron. Response to stretch (*a,b*) before and (*c,d*) after application of tetrodotoxin, 5×10^{-6}. Upper beam, intracellular record from cell soma; middle beam, record from axon, about 2 mm away from soma; lower beam, strain gage record of mechanical displacement (stretch) to the receptor organ. Calibration: 2 s; 10 mV. (*e, f*) Pacinian corpuscle (decapsulated). Response to threshold mechanical stimulus (*e*) before and (*f*) after application of tetrodotoxin, 5×10^{-4}. Upper beam is the electrical response recorded externally from the axon near its junction with the nonmyelinated portion of ending; lower beam, photoelectric record of stimulus. Calibration: 1 ms; 15 μV. (*g,h*) Generator potential after tetrodotoxin application (concentration 5×10^{-6}) which blocks action potentials. (*g*) Crustacean stretch receptor. (*h*) Pacinian corpuscle.

endings, like those of Pacinian corpuscles, lie enclosed in a special adventitious tissue. The structure that converts mechanical into electrical energy—that is, the receptor proper—appears to lie within the membrane of the nonmyelinated ending.

The structure in which the all-or-nothing nerve impulse arises is, at least at the molecular scale, different from that giving rise to the generator current. This is indicated by a number of differential properties of the generator and action potentials, particularly by the differential action of tetrodotoxin, a pharmacological agent that blocks rather specifically the Na^+-carried component of the action potential. At low concentrations tetrodotoxin abolishes completely the action potential in the Pacinian corpuscle and the crustacean stretch receptor, leaving the generator current virtually unaltered (**Fig. 7**).

In some receptors the membrane sites of generator current and impulse production seem to be coarsely segregated. In the crustacean stretch receptor the generator current arises at the dendrites, whereas the action potential originates at the initial axon portion of the nerve cell. In other cases, segregation of the sites may be at a much finer scale, and interspersion of the sites on a given membrane may perhaps occur.

Receptor membrane properties. The following picture of nerve impulse generation has emerged from work on Pacinian corpuscles by W. R. Loewenstein and collaborators. There is a resting potential across the receptor membrane of the nonmyelinated nerve ending. Deformation (strain, but not stress) of the membrane leads to a fall in membrane conductance (excitation) and hence to a reduction in membrane potential (depolarization). The excitation is confined to that region of membrane which has been stimulated mechanically; it does not spread to adjacent regions of the same membrane. This is illustrated by the experiment depicted in **Fig. 8**. A small region of receptor membrane is stimulated mechanically, and the resulting generator potential is measured at varying distances from the stimulated region. The region is stimulated with mechanical pulses delivered by a stylus with a tip of about 20-micrometer diameter, while a microelectrode scans the surface of the remaining 800-μm length of receptor membrane. The amplitude of the generator potential is found to decrease exponentially with distance between the stimulated and recording spot; in other words, excitation is localized at and confined to the mechanically stimulated spot. This stands in strik-

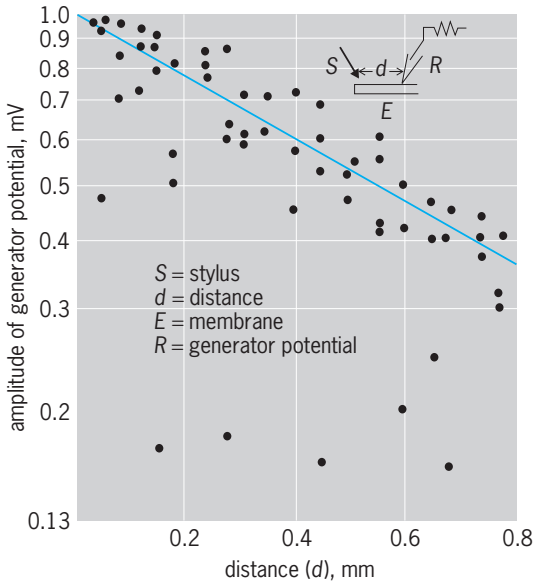

Fig. 8. Local excitation in the receptor membrane. A spot of membrane is stimulated with a stylus with a tip of about 20-μm diameter, and the resulting generator potentials are recorded at varying distances from the stimulated spot. (*After W. R. Loewenstein, Excitation and inactivation in a receptor membrane, Ann. N.Y. Acad. Sci., 94(2):510–534, 1961*)

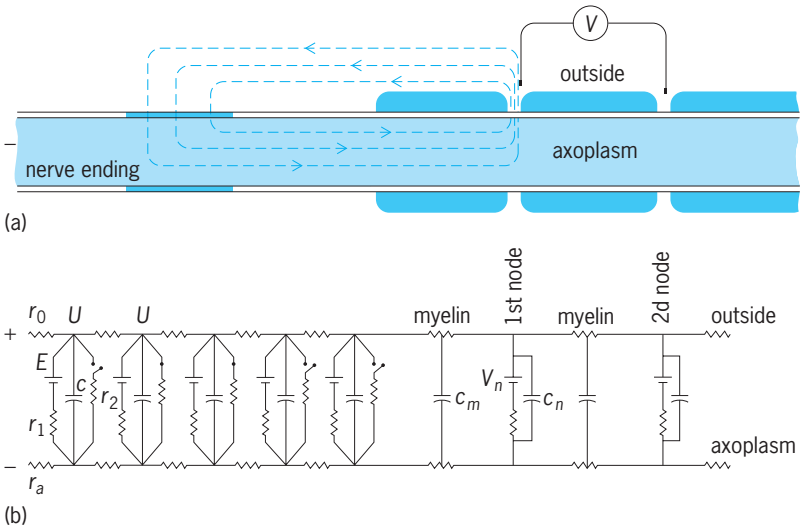

Fig. 9. Receptor analog. The generator potential (*V*) is measured between the first and second nodes of Ranvier. The nerve ending and the associated myelinated axon are represented by the equivalent circuit below, in which *E* represents resting potential, *C* redistribution capacitance, and *r* redistributed resistance. (*After W. R. Loewenstein, Ann. N.Y. Acad. Sci., 94:510–534, 1961*)

ing contrast to the propagated mode of excitation of the conductile portion of the axon. The peculiarity of the receptor membrane is that excitation is produced by the adequate stimulus, the mechanical stimulus in this case, but not by the resulting local flow of current. In this behavior the receptor membrane of mechanoreceptors resembles the membrane of the motor end plate of skeletal musculature and postsynaptic structures of nerve cells which lack regenerative mechanisms and which seem to have within their membranes functional properties of the primitive receptor cell. The result illustrated by Fig. 8 reveals that the conductance change that gives rise to transfer of charges through the membrane, namely the generator current, can be produced in a given spot independently of a neighboring one on the same membrane. *See* SYNAPTIC TRANSMISSION.

Thus, for analytical purposes the receptor membrane may be regarded as being composed of functionally independent units or generator elements (**Fig. 9**). In Fig. 9 membrane capacitance and resistance are assumed to be uniformly distributed over the receptor membrane. The transmembrane potential of a generator element is assumed to drop to a fraction of its resting value when the element is activated mechanically. Each generator element may be excited independently of others and may respond to mechanical stimulation with a (active) drop of membrane potential. This feature of local excitation is represented by switching the shunt resistor (r_2) of the active generator into the circuit. Spatial summation of electric activity along the surface of the receptor membrane is an immediate consequence of this mode of excitation, the total amount of current increasing progressively with the area of membrane excited. An experimental generator potential–membrane area relation is reproduced in **Fig. 10**. The generator potential increases at a diminishing rate with the area of membrane excited. This finding is of considerable physiological interest. The area of membrane excited appears to be the main factor, if not the only one, in determining the intensity of the generator current and, hence, the coding of sensory nerve impulses when the stimulus strength (the deflection amplitude or degree of compression) is varied under physiological conditions (**Fig. 11**).

The generator potential appears to be built up by a statistically fluctuating number of active generator elements. A mechanical stimulus of a given strength excites a statistical population out of a larger finite number of excitable generator elements. The probability for excitation of generator elements increases as a function of stimulus strength. This mode of excitation appears to have the following consequences: (1) The amplitudes of the generator potential in response to stimuli of standard strength fluctuate at random around a mean value; and (2) when two mechanical stimuli are applied in rapid succession, the number of generator elements excitable by the second stimulus is inversely related to the first stimulus.

Temperature. Another factor that influences the generator current of the receptor membrane is temperature (**Fig. 12**). The rate of rise of the generator potential from a constant mean area in response to a constant stimulus increases markedly with temperature; its value is more than doubled for a 118°F (10°C) increment of temperature. In addition, the amplitude of the generator potential also increases at about the same rate, while the duration of the falling phase is practically unchanged. This is another feature of the receptor membrane that contrasts with

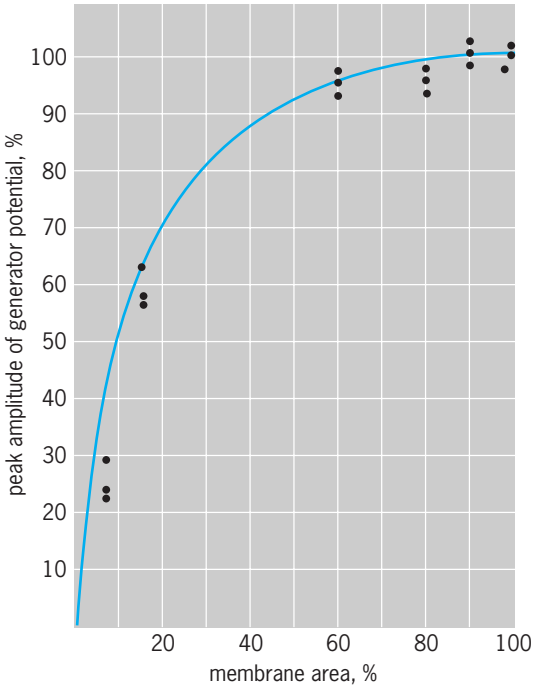

Fig. 10. Amplitude of peak generator potential as a function of the excited membrane area. An increasing area of receptor membrane of a Pacinian corpuscle is excited with styli of increasing diameter. (*After W. W. Loewenstein, Excitation and inactivation in a receptor membrane, Ann. N.Y. Acad. Sci., 94(2):510–534, 1961*)

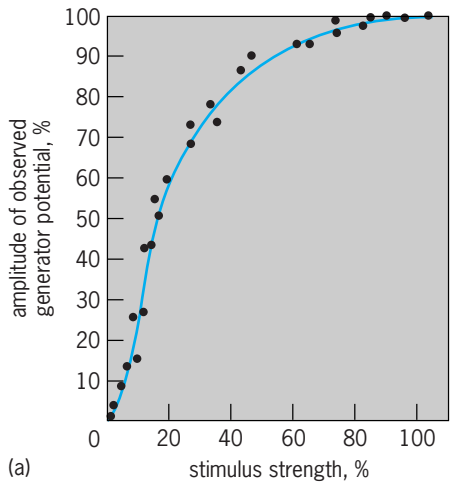

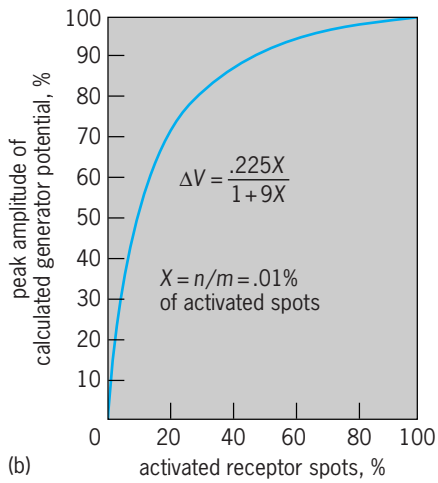

Fig. 11. Input-output relation of receptor membrane. (*a*) The observed generator potential of the receptor membrane of a typical Pacinian corpuscle as a function of stimulus strength. (*b*) The calculated generator potential (ΔV) as a function of the number of active generator elements as computed from the receptor analog of Fig. 9. (*After W. R. Loewenstein, Ann. N.Y. Acad. Sci., 81:367–387, pt. 3, 1959*)

the membrane of the conductile axon. The amplitude of the action potential of the Ranvier nodes adjacent to the receptor membrane, like that of other membranes with regenerative excitability, remains rather constant over a wide range of temperature, while its duration is markedly increased. Temperature appears to increase the conductance change in the excited receptor membrane; a high potential barrier is associated with the conductance change. The energy of activation, as calculated from the temperature dependence of the rate of rise of the generator potential, amounts to about 16,000 cal/mole. The rate-limiting activation energy appears to be associated either with the flow of ions through the receptor membrane or with the alteration in membrane structure that leads to the conductance change. Mechanical energy and thermokinetic energy, however, are not interchangeable in excitation. A change in temperature modifies the conductance change in the mechanically excited membrane but does not initiate it. Initiation of the conductance change, namely excitation, is brought about only by the mechanical stimulus.

Receptor inactivation. After repetitive stimulation at high frequency, the responsiveness of the receptor membrane of Pacinian corpuscles becomes markedly reduced (**Fig. 13**). The effect is fully reversible and has low energy requirements. It reflects a transient inactivation in the receptor membrane. The inactivation increases with frequency, duration of the train of repetitive stimuli, and strength of repetitive stimuli.

Receptor potentiation. With other, more critical parameters of repetitive stimulation, an increase in membrane potential may ensue. Thus, during the period of hyperpolarization (of the order of 10 s) a generator potential in response to a standard test stimulus is augmented (**Fig. 14**).

Production of nerve impulses. When an outward current is passed through the membrane sites of nerve impulse initiation, their resting potential is diminished. As in other excitable tissues with regenerative mechanisms, an all-or-nothing potential is discharged when the resting potential has been lowered to a critical level. This is what happens when the receptor membrane is excited mechanically and a generator current flows between active generator elements and the inactive impulse initiation site. If this current is intense enough, an all-or-nothing potential is discharged which may then propagate as a nerve impulse along the axon toward the nerve centers. This is the way nerve impulses are produced at receptors and how information of the outside world is conveyed to the nerve centers.

Code of sensory messages. Because the conductive elements of the nerve systems of higher organ-

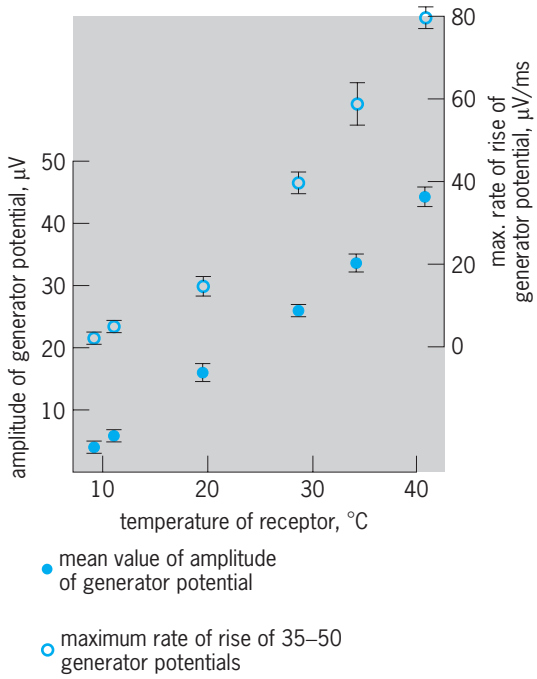

Fig. 12. Effect of temperature on the amplitude and the rate of rise of generator potential in the Pacinian corpuscle. Bars subtend standard error of the mean. °C = (°F − 32)/1.8. (*After N. Ishiko and W. R. Loewenstein, Effects of temperature on the generator and action potentials of a sense organ, J. Gen. Physiol., 45:105–124, 1961*)

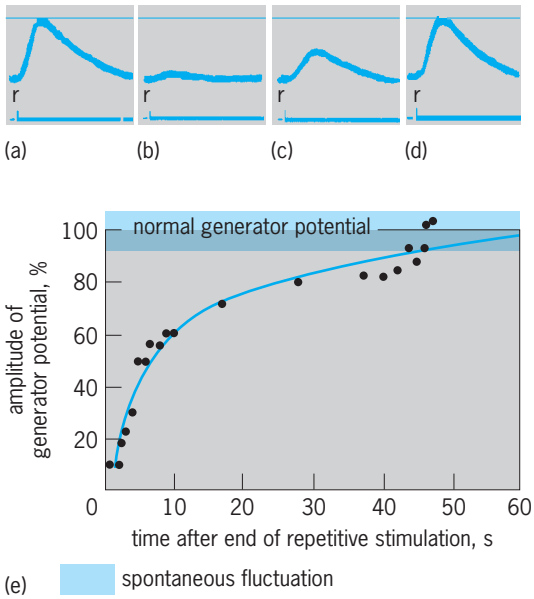

Fig. 13. Aftereffects of activity in receptor membrane: desensitization. (*a*) Generator response to a standard test stimulus of the fully rested receptor membrane (Pacinian corpuscle); (*b*) 0. 2 s, (*c*) 5 s, and (*d*) 50 s after applying a train of repetitive mechanical stimuli at 500/s for 30 s to the receptor membrane. Calibration 20 μV; 0.5 ms. (*e*) Recovery of depressed generator potential as a function of time. (*After W. R. Loewenstein and S. Cohen, J. Gen. Physiol., 43:347, 1959*)

isms, the axons, along which the nerve impulse must propagate have membranes with all-or-nothing type of responsiveness only, there is no possibility for gradation in amplitude of nerve impulses. Nature has endowed organisms with a rather stereotyped information system in which messages are sent in a code of equal all-or-nothing pulses. Thus, the only way of conveying quantitative information in nerve messages in a given information channel is by changing the sequence of its staccato content.

There are two ways by which the central nervous system can analyze the content of a receptor message: by counting the number of parallel information channels engaged in impulse traffic, and by gaging the frequency and sequence pattern of impulses transmitted in each information channel. The uncovering of this form of coding is chiefly due to E. D. Adrian.

Frequency coding. It has been established that the generator current which flows between the active receptor membrane and the impulse initiation site increases in its intensity and in its rate of rise with strength of stimuli. Thus, with suprathreshold stimuli of different strength, the resulting generator currents will cause the nodal membrane potential to drop to the critical level at a rate directly related to stimulus strength. Hence the frequency of discharged impulses is a function of the rate of rise of the generator current and indirectly of the stimulus strength (**Fig. 15**). This is then one way to feed messages of different quantitative content into the nerve system. The greater the stimulus intensity, the higher the impulse frequency of the message.

Spatial code. In animals of higher organization, receptors occur in groups. A pull on a muscle excites not one but many spindle receptors. A weight applied to the skin excites a group of Pacinian corpuscles. If one receptor of a group were to be killed, it would make little or no difference to the total message reaching the central nerve cells. This leads to another manner by which the content of a message may be graded. The greater the deformation of an organ, the greater the probability for excitation of different homonymous receptors. Thus, as a weight applied to the skin is increased, a larger area of skin is deformed, a larger number of receptors are excited, and hence a larger number of parallel axons become involved in impulse conduction. Also, since generally several receptor endings are twigs of one axon, there will be considerable mixing of impulses into the common axon. Consequently, because the number of receptors activated increases with strength, the frequency of total impulses in the common axon may also increase by this mechanism. Further mixing may still occur because of the convergence of several axons into nerve cells along the way to the centers.

Receptor adaptation. One of the striking properties of mechanoreceptors is their ability to adapt to stimuli. Adaptation is readily shown by applying a single mechanical stimulus of a given suprathreshold strength to a receptor organ and maintaining it in this way indefinitely without change. The receptor may respond with an initial burst of impulses at a given frequency; but then, in spite of the fact that

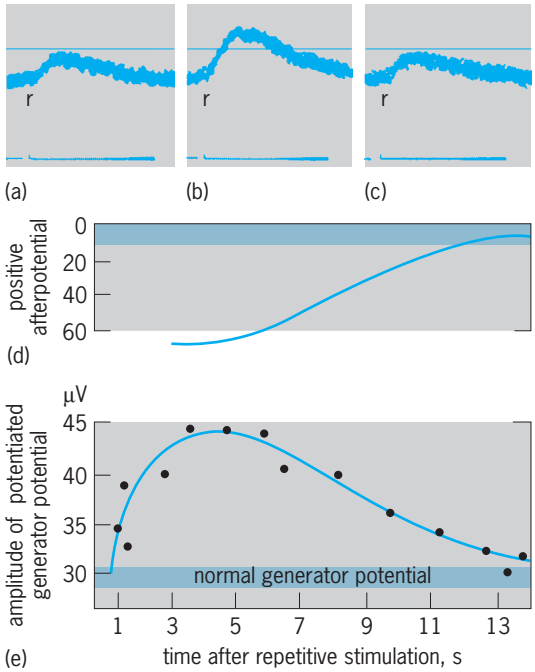

Fig. 14. Aftereffects of activity in receptor membrane: potentiation. Responsiveness of receptor (Pacinian corpuscle) is tested (*a*) with standard satimulus under fully rested conditions, (*b*) 5 s, and (*c*) 18 s after applying a train of repetitive stimuli at 500/s for 5 s. (*d*) Hyperpolarization. (*e*) Time course of resulting potentiation. (*After W. R. Loewenstein and S. Cohen, J. Gen. Physiol. 43:347, 1959*)

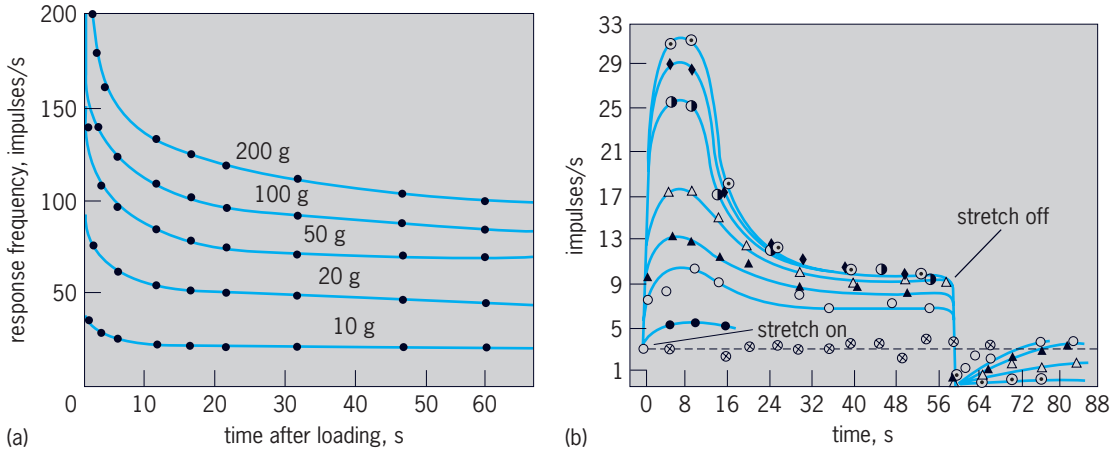

Fig. 15. Impulse frequency as a function of stimulus strength. (*a*) Muscle spindle of the frog at various degrees of tension (*after B. H. C. Matthew, J. Physiol., 78(1):1–54, 1933*). (*b*) Pressure receptor of frog skin at various degrees of stretch (*after W. R. Loewenstein, J. Physiol., 133(3):588–602, 1956*).

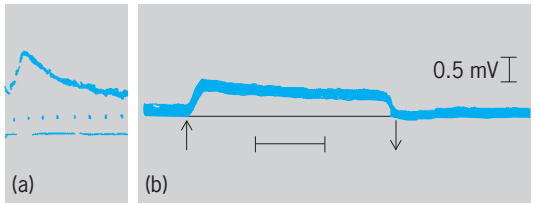

Fig. 16. Adaptation in generator potential. Response to a mechanical stimulus applied and sustained continuously for about 4 s (*a*) in a fast-adapting Pacinian corpuscle (calibration, 1 ms between dots) (*after W. R. Loewenstein, J. Gen. Physiol., 41(4):825–845, 1958*); and (*b*) in a slowly adapting crustacean stretch receptor (calibration, 1 s) (*after C. Eyzaguirre and S. W. Kuffler, J. Gen. Physiol., 39(1):87–119, 1955*).

the stimulus is continuously sustained, the frequency of impulses diminishes and may in some receptors fall to zero. The receptor is said to have adapted to the stimulus. The rate of adaptation varies greatly for different mechanoreceptors. Thus, a Pacinian corpuscle adapts to zero impulses within a few milliseconds, other skin receptors take seconds, and stretch receptors of the skin and of the muscle spindle maintain stationary discharges even after minutes of steady stimulation. It is in the last type that the stimulus strength–impulse frequency relationship is most readily seen (Fig. 15).

Adaptation in generator potential. Adaptation depends, at least in part, upon the ability of the receptor organ to sustain a constant generator potential. In fast-adapting receptors, such as the Pacinian corpuscle, the generator potential decays to zero within 4–7 ms in the face of a continuously sustained suprathreshold stimulus. In slowly adapting or nonadapting receptors, such as the muscle spindle and the slow-stretch receptor of Crustacea, the generator potential is well sustained for some time above the firing threshold of the impulse initiation site (**Fig. 16**). In fast-adapting receptors, impulses can obviously be produced only during the 1–2 ms in which the generator potential is above firing threshold, whereas in slowly adapting or nonadapting receptors a continuously sustained barrage of impulses may result.

The rate-limiting factor in generator current adaptation is a mechanical one. This has been clearly shown for the Pacinian corpuscle, whose capsule operates like a mechanical high-pass filter. The capsule of this receptor organ is made of many concentrically arranged tissue lamellae separated by fluid-filled spaces. The chief elements for transmission of mechanical forces through the capsule are the lamellae, their radial interconnections, and the interlamellar fluid. These are arranged in a transmission chain, at the end of which, at the center of the capsule, lies the nerve ending, the transducer proper. The mechanical equivalent is a system of dashpots and springs (**Fig. 17**). One set of springs, *M*, represents the elastic components of the lamellae, and the other set, *S*,

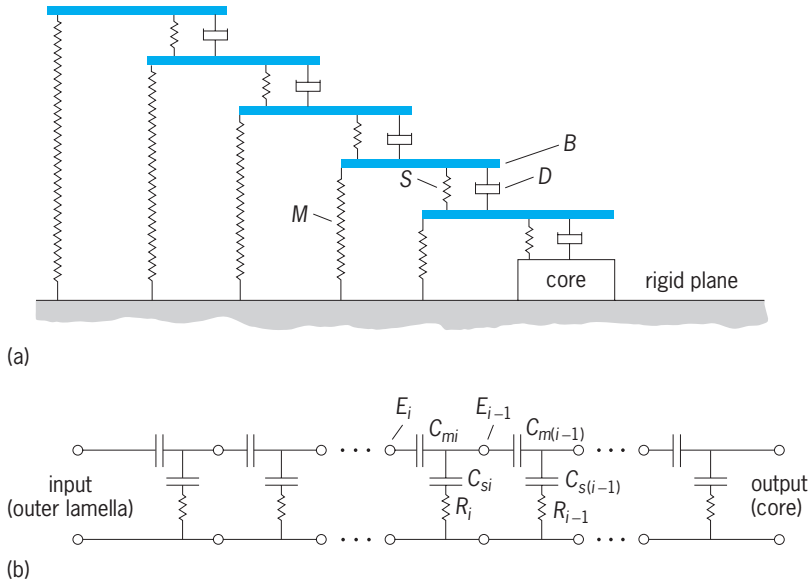

Fig. 17. Analogs of Pacinian corpuscle. (*a*) Mechanical. The rigid components of the lamellae are represented by bars (*B*), their compliance by springs (*M*), their radial spring compliance (interlamellar connections) by springs (*S*), and the fluid resistance by dashpots (*D*). (*b*) Electrical. Lamella compliance is represented by capacitance C_{mi}, radial spring compliance by C_{si}, and fluid resistance by resistance R_i. (*After W. R Loewenstein and S. Kalak, 1966, courtesy of J. Physiol.*)

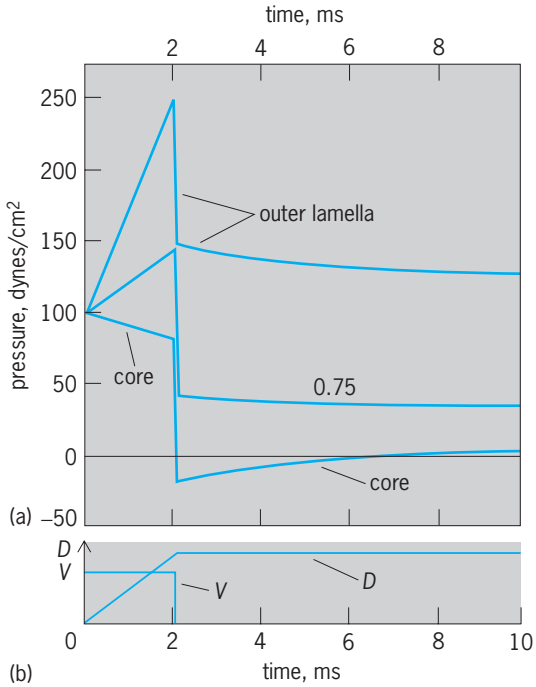

Fig. 18. Dynamic pressures in corpuscle. (*a*) Time course of computed pressures on outermost lamella, at 0.75 of corpuscle radius, and on core for total compression of corpuscle of 20 μm applied uniformly in 2 ms and then maintained. (*b*) Displacement, *D*, and velocity, *V*, at outer lamella. (*After W. R. Loewenstein and S. Kalak, 1966, courtesy of J. Physiol.*)

the much weaker elastic interconnections. During compression the lamellar surfaces operate like a series of dashpots which, in displacing fluid through narrow constraints, generate viscous pressure. The compliance of the lamellar interconnections is much greater than that of the lamellae themselves. As a result, static pressures applied to the surface of the capsule attenuate very rapidly toward the capsule center. When the capsule is loaded statically, only a very small fraction of the external pressure load reaches the center where the transducer nerve ending is located. This is the mechanical basis of adaptation.

During dynamic compression of the capsule the situation is quite different. At this time the fluid velocities set up viscous pressure, which is well propagated through the capsule. This is the only force of import that reaches the transducer ending; elastic forces are filtered out efficiently by the lamellated structure of the capsule. Thus, since viscous force is velocity-dependent, the electrical output of the transducer ending, the generator current, is velocity-dependent. This is generator current adaptation in more general terms.

The physical aspects of adaptation are well illustrated by the situation considered in **Fig. 18**. As long as there are fluid velocities in the system, there are appreciable pressures of viscous origin on the core. With the onset of the static phase of stimulation, fluid velocities and pressures on the transducer nerve ending fall toward zero and remain there for the entire period of static stimulation. This alone accounts for the rate of decay of the generator potential under external static stimulation. Thus, at least one adaptation process resides at the mechanical input side of the transducer chain in the form of a filter for elastic forces.

Another adaptation process is located at the level of nerve impulse initiation, preventing a steady generator current from producing repetitive nerve impulses. It is illustrated by bypassing the normal generator mechanisms and stimulating the site of impulse initiation with an electric current. Under these conditions the number of nerve impulses discharged is more or less independent of the duration of the current, and, typically, only one impulse is generated in the Pacinian corpuscle in response to strong currents. The mechanisms of the second filter action are not yet understood.

Off response. Many fast-adapting mechanoreceptors give generator current to both the on and off phases of the external stimulus. The mechanics of the off response in the Pacinian corpuscle are apparent from the mechanical model of Fig. 11. During the off phase the corpuscle tends to return to its circular cross section under the action of forces generated in the lamellae and lamellar interconnections. Thus energy stored in the elastic elements during compression is set free and viscous pressure is produced anew. During the off phase, compression occurs once again, but this time at right angles to the direction prevailing in the on phase. The transducer ending does not discriminate appreciably the force angle. Thus a second generator current, the off-generator response, results. After removal of the capsule, the receptor ceases to produce off responses.

Efferent control of receptors. Receptor inhibition and facilitation are discussed in the following sections.

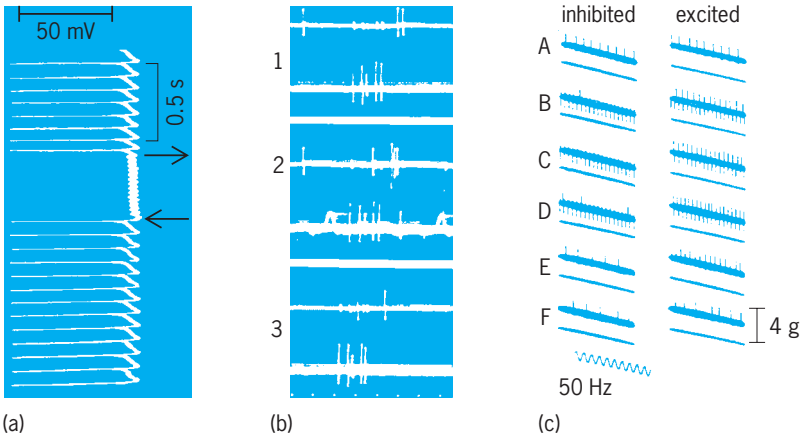

Fig. 19. Efferent control of receptors. (*a*) Intracellular recording from a crustacean stretch receptor submitted to constant stretch (*after C. Eyzaguirre and S. W. Kuffler, J. Gen. Physiol., 39(1):155–184, 1955*). (*b*) Sensitization of touch receptors (frog) by centrifugal sympathetic impulses. Impulse discharge of a group of touch receptors in response to a standard mechanical stimulus: 1, before; 2, during; and 3, after arrival of sympathetic impulses of the skin (*after W. R. Loewenstein, J. Physiol., 132(1):40–60, 1956*). (*c*) Supraspinal control of muscle spindle (cat soleus). Impulse discharge at a standard stretch. A, E, F, before, and B, C, D, after stimulation of inhibitory loci (left) and excitatory loci (right) of the contralateral inferior colliculus (*after E. Eldred et al., J. Physiol., 122(3):498–523, 1953*).

Receptor inhibition. The impulse discharge of certain mechanoreceptors can be modulated by centrifugal impulses. An example of this is the stretch receptor cell of Crustacea. This cell receives accessory nerve fibers which end on the receptor dendrites. S. W. Kuffler and collaborators showed that the effect of impulses conveyed by nerve fibers to the receptor cell is to clamp its membrane potential near resting value by increasing potassium-ion conductance, apparently in a specific way. In this manner, mechanically elicited generator potentials are prevented from reaching the critical level for impulse firing, and sensory discharges in response to stretch are promptly suppressed at the receptor level. In **Fig. 19** the inhibitory axon to the receptor is stimulated (between arrows), thus preventing the generator potential from reaching firing level.

Receptor facilitation. Certain tactile receptors of the frog's skin can be modulated by centrifugal impulses from the sympathetic system. Arrival of sympathetic impulses or application of sympathomimetic substance increases the sensitivity of touch receptors to mechanical stimuli (Fig. 19). In Fig. 19 an excitation of new receptor units, indicating an increase in excitability, is shown; also, there are repetitive discharges of some units during sympathetic stimulation, indicating lowering of the adaptation rate. The sympathetic effect is mediated by an adrenalin-like substance which appears to diffuse from the sympathetic terminal to the touch receptor ending. The anatomical relations between the afferent and efferent systems involved here are not known.

Modulation at accessory structure. A muscle spindle in the rabbit consists schematically of a bundle of muscle fibers (intrafusal muscle fibers) whose central portion (nuclear bag) beds the nude receptor ending (the receptor proper) of the thick myelinated afferent axon (Fig. 3). The intrafusal muscle receives, in addition, a set of motor fibers whose impulses cause it to contract. Stretching out the central nuclear bag excites the receptor ending, causing impulses to travel centripetally along the thick axon. The intrafusal muscle fibers lie parallel to the large muscle fiber of the skeletal muscle in which they are cribbed. Thus, shortening of the skeletal muscle during contraction causes the intrafusal fiber to slacken and thereby the afferent impulse discharge to cease. If, under these conditions, motor impulses are conveyed to the intrafusal fibers, the fibers appear to contract at a relatively fixed length (the ends of the intrafusal fiber are inserted in the endomysium or perimysium of the skeletal muscle). Thus the slack is taken up, the receptor ending is stretched, and impulses are fired again to the centers. The motor innervation of the intrafusal bundle thus appears to control the responsiveness of the sense organ by setting the tension of its accessory structure. *See* BIOPOTENTIALS AND IONIC CURRENTS; NERVOUS SYSTEM (INVERTEBRATE); NERVOUS SYSTEM (VERTEBRATE). Werner R. Loewenstein

Bibliography. C. Barnes and P. W. Kaliuas (eds.), *Sensitization of the Nervous System*, 1988; P. Hnik et al. (eds.), *Mechanoreceptors: Development, Structure, and Function,* 1988; M. Matlin, *Sensation and Perception,* 4th ed., 1996; J. G. Retzloff, *Sensation Research,* 1987; A. Struppler et al. (eds.), *Sensory-Motor Integration: Implication for Neurological Disease*, 1987.

Mecoptera

A small order of insects called the scorpion flies. Characteristic of the adult insect is the peculiar prolongation of the head into a beak, which bears chewing mouthparts. They are small to medium in size. The insects either have two pairs of large, net-veined wings of equal size, often with dark areas, or have short and aborted wings. The legs are long and slender. In some species, the male abdomen has a terminal enlargement which is held recurved over the back so that he resembles a scorpion, thus the common name (see **illus.**). The larvae are eruciform

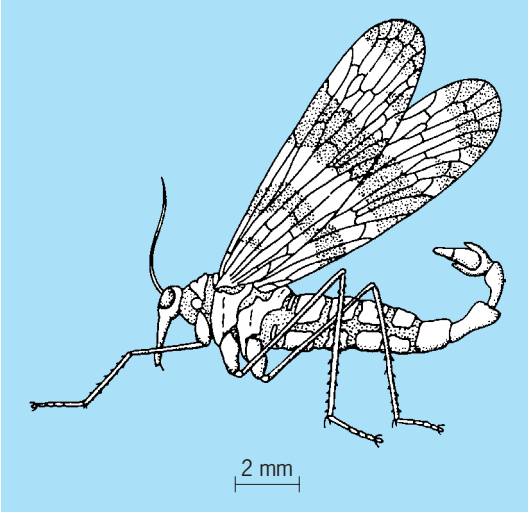

Male scorpion fly (*Panorpa*), order Mecoptera.

(resembling caterpillars), and the pupae exarate; that is, the appendages are free from the body and artificially movable. The Mecoptera are found in moist habitats within densely wooded areas. The adults are omnivorous but feed chiefly on small insects. *See* INSECTA. Bryant E. Rees

Medical bacteriology

Medical bacteriology is the study of bacteria that cause human disease. The field encompasses the detection and identification of bacterial pathogens, determination of the sensitivity and mechanisms of resistance of bacteria to antibiotics, the mechanisms of virulence, and some aspects of immunity to infection. *See* VIRULENCE.

Identification of bacteria. The clinical bacteriology laboratory identifies bacterial pathogens present in specimens such as sputum, pus, urine, feces, blood, and spinal fluid, or from swabs of skin, throat, rectal,

or urogenital surfaces. Identification involves direct staining and microscopic examination of these materials, and isolation of bacteria present in the material by growth in appropriate media. The laboratory must differentiate bacterial pathogens from harmless bacteria that colonize humans.

Many specialized bacteriologic media have been developed for the purpose of selectively permitting the growth of potential pathogens while inhibiting the growth of harmless commensals. Species and virulent strains of bacteria can be identified on the basis of growth properties, metabolic and biochemical tests, and reactivity with specific antibodies.

Recent advances in the field of diagnostic bacteriology have involved automation of biochemical testing; the development of rapid antibody-based detection methods; and the application of molecular biology techniques such as nucleic acid hybridization, the polymerase chain reaction, and deoxyribonucleic acid (DNA) sequencing. Nucleic acid analysis can permit the identification of bacterial pathogens without the necessity of visualization and culture, and has resulted in more rapid identification of slow-growing pathogens. Once a bacterial pathogen has been identified, a major responsibility of the diagnostic bacteriology laboratory is the determination of the sensitivity of the pathogen to antibiotics. This involves observation of the growth of the bacteria in the presence of various concentrations of antibiotics. The process has been made more efficient by the development of automated instrumentation.

An increasingly serious problem in the therapy of infectious diseases is the emergence of antibiotic-resistant strains of bacteria. Bacterial infections caused by such pathogens as *Streptococcus pneumoniae* and *Staphylococcus aureus* that were easily treated in past decades are becoming increasingly difficult to treat due to the emergence of antibiotic resistance. An important area of research is the mechanisms of acquisition of antibiotic resistance and the application of this knowledge to the development of more effective antibiotics. *See* ANTIBIOTIC; ANTIGEN-ANTIBODY REACTION; BACTERIAL PHYSIOLOGY AND METABOLISM; BACTERIAL TAXONOMY; IMMUNOCHEMISTRY.

Infection. Most bacteria that associate with humans are harmless. Bacteria colonize the skin or mucosal epithelium (gastrointestinal tract, respiratory tract, and urogenital tract) and exist as commensals. The intact epithelium serves as the host's primary defense against invasion of deeper tissues. The second line of defense is the presence of white blood cells known as neutrophils and macrophages that ingest and kill any bacteria that manage to penetrate the epithelial surfaces. The process of ingestion of bacteria by white blood cells is termed phagocytosis. A successful bacterial pathogen must defeat these defenses.

Some bacterial pathogens remain at the epithelial surface and produce disease without deeper invasion. This strategy usually involves the production of toxins that damage or alter the function of host tissues. An example is *Vibrio cholerae*, the agent of the potentially fatal dehydrating diarrheal disease known as cholera. This pathogen colonizes the small intestine of the host and produces a toxin that acts on the intestinal epithelium, causing it to abnormally secrete fluid and electrolytes. Another example is *Corynebacterium diphtheriae*. This respiratory pathogen remains on the epithelial surface of the throat and secretes a toxin that damages the respiratory epithelium and enters the circulation and damages the heart. Invasion of the epithelial surface may require prior injury, as is usually the case with skin infections, or may involve the direct bacterial invasion of epithelial cells. A number of bacterial pathogens induce epithelial cells to engulf them. Some remain in the epithelial cell, divide rapidly, and destroy the cell. Examples include *Shigella*, the agent of bacillary dysentery, and *Neisseria gonorrhoeae*, an agent of sexually transmitted disease. Others invade beyond the epithelial surface, enter the circulation, and produce systemic disease.

These pathogens have special properties that help them to defeat phagocytic cells. Some pathogens have surface properties that prevent the phagocytic cell from initiating phagocytosis. An example is *Streptococcus pneumoniae*, an agent of pneumonia and meningitis. This bacterium is encapsulated with a polysaccharide substance that protects it from phagocytic cells. Some pathogens secrete substances known as hemolysins that attack and damage the phagocytic cells. An example is *Streptococcus pyogenes*, an agent of throat and skin infections. The action of hemolysins can be visualized by growth of *S. pyogenes* on medium containing red blood cells, which are destroyed in the vicinity of a bacterial colony.

Some pathogens permit phagocytosis but, once internalized, are not killed. An example is *Mycobacterium tuberculosis*, which survives within the macrophages of the lung. The presence of bacteria in tissue elicits the delivery of phagocytic cells to the site of infection. If the infection is not controlled, the accumulation of phagocytic cells can contribute to the tissue damage caused by bacterial growth. Thus, the body's own defense mechanisms, when defeated by the bacterial pathogen, actually contribute to the disease. Treatment of some types of bacterial meningitis, therefore, involves not only antibiotics to kill the bacteria but also steroidal anti-inflammatory drugs to suppress the destructive action of phagocytic cells. *See* PHAGOCYTOSIS; TOXIN.

Pathogenicity. The study of bacterial pathogenesis involves the fields of molecular genetics, biochemistry, cell biology, and immunology. In cases where the disease is not serious and easily treated, research may involve the deliberate infection of human volunteers. Otherwise, various models of human disease must be utilized. These involve experimental infection of animals and the use of tissue cell culture systems. Modern molecular approaches to the study of bacterial pathogenesis frequently involve the specific mutation or elimination of a bacterial gene thought to encode a virulence property, followed by observation of the mutant bacteria in a

model system of human disease. In this way, relative contributions of specific bacterial traits to different stages of the disease process can be determined. This knowledge permits the design of effective strategies for intervention that will prevent or cure the disease. The studies may also yield significant new basic knowledge of human biology. A recent example is the identification of the mechanism of action of tetanus toxin, a potent bacterial neurotoxin. Identification of the cellular proteins that are targets of tetanus toxin has permitted a much more detailed understanding of the processes of neurotransmitter release by nerve cells. *See* BACTERIAL GENETICS.

Immunity. Infection with many bacterial pathogens renders the host immune to subsequent infections. Frequently, immunity is mediated by specific antibodies. Antibodies are proteins that bind to a specific target, and are the natural result of a first encounter with a specific disease. On a subsequent encounter with the same pathogen, the presence of specific antibodies against bacterial substances will neutralize the virulence properties and the host will be immune to infection. The presence of specific antibodies is frequently useful in the diagnosis of bacterial diseases in which the pathogen is otherwise difficult to detect. An example is the sexually transmitted disease syphilis. The causative pathogen, *Treponema pallidum*, cannot be cultured, and is very difficult to visualize in clinical specimens. The diagnosis of syphilis, therefore, must be confirmed by the demonstration of antibodies specific for *T. pallidum*. *See* ANTIBODY; BIOLOGICALS.

Vaccines are substances that elicit immunity to a specific disease. Most vaccines against bacterial infections elicit the production of protective antibodies in the host. Vaccines provide an initial contact with the pathogen, or components of the pathogen, in such a manner that protective antibodies are elicited in the absence of disease. In cases where disease results from circulation of a bacterial toxin, the vaccine may be an inactivated form of the toxin. Examples are vaccines against tetanus and diphtheria. The bacterial toxins are purified and chemically inactivated but still induce an immune response. Some early vaccines against bacterial diseases consisted of killed bacterial cells. An example still in use is pertussis vaccine, which protects against whooping cough caused by the bacterial pathogen *Bordetella pertussis*. Such vaccines frequently produce undesirable reactions such as fever. Modern molecular research on the pathogenicity of *B. pertussis* has identified specific bacterial virulence components. On the basis of this knowledge, a new type of pertussis vaccine has been designed that consists of some of these components in a purified form. This type of vaccine has been termed subunit vaccine. The vaccine has proven as effective as the killed bacterial cells, and is much less toxic. Another example is the subunit vaccine against *Haemophilus influenzae* type B, which until recently was an agent of meningitis in infants. This pathogen invades the respiratory epithelium, and is resistant to phagocytic cells by virtue of a polysaccharide capsule which prevents interaction with them. A vaccine was developed that consisted of the purified polysaccharide coupled to a protein carrier. This vaccine elicits protective antibodies in infants, has few side effects, and has essentially eliminated *H. influenzae* as a serious pathogen in the United States after only a few years of universal vaccination. *See* VACCINATION.

Immunity to some bacteria that survive intracellularly is not mediated by antibodies but by immune effector cells, known as T cells, that activate infected cells to kill the bacteria that they contain. An active area of research is how bacterial components are presented to the immune system in a way that will induce effective cell-mediated immunity. This research may lead to the development of T-cell vaccines effective against intracellular bacterial pathogens.

For disease entities caused by specific bacteria. *See* ANTHRAX; BOTULISM; BRUCELLOSIS; CHOLERA; DIPHTHERIA; GANGRENE; GLANDERS; GONORRHEA; GRANULOMA INGUINALE; JOHNE'S DISEASE; LEPROSY; LISTERIOSIS; PLAGUE; PSEUDOTUBERCULOSIS; TETANUS; TUBERCULOSIS; TULAREMIA.

For disease entities caused by more than one microorganism *See* FOOD POISONING; INFANT DIARRHEA; MENINGITIS; PNEUMONIA.

For groups of disease-producing bacteria *see* MEDICAL BACTERIOLOGY; HAEMOPHILUS; IMMUNOLOGY; PNEUMOCOCCUS; GANGRENE; STREPTOCOCCUS.

Steve L. Moseley

Bibliography. E. W. Koneman, *Textbook and Atlas of Diagnostic Microbiology*, 5th ed., Lippincott-Raven Publishers, 1997; G. L. Mandell, J. E. Bennett, and R. Dolin (eds.), *Principles and Practice of Infectious Diseases*, 4th ed., Churchill Livingstone, 1994; A. A. Salyers and D. D. Whitt, *Bacterial Pathogenesis: A Molecular Approach*, ASM Press, 1994.

Medical control systems

Physiological and artificial systems that control one or more physiological variables or functions of the human body. Regulation, control processes, and system stability are at the heart of the survival of living organisms, both unicellular and multicellular. In the nineteenth century, C. Bernard concluded that the higher animals, far from being indifferent to their surroundings, must be in close and intimate relation to them. The equilibrium they maintain is the result of compensation established as continually and exactly as if by a very sensitive balance. W. B. Cannon (1929) differentiated the stability properties of biological systems from those of physical systems, and introduced the term homeostasis to describe the steady states in the body that are maintained by complex, coordinated physiological reactions. The condition of homeostasis is achieved either by regulation of supplies (for example, control of blood sugar level) or by regulation of processes (for example, control of body temperature and control of voluntary movements). *See* HOMEOSTASIS.

Medical control systems may be classified into two groups: (1) the physiological control systems in normal or pathological conditions (for example, control of electrolytes, arterial pressure, respiration, body temperature, blood sugar, endocrinal functions, neuromuscular and motor activity, and sensory functions), and (2) the external (artificial) control systems that interface with physiological systems (for example, artificial kidneys or hemodialyzers, blood oxygenators or heart-lung machines used during open-heart surgery, external prosthetics and orthotics, cardiac pacemakers, ventilators, implantable defibrillators, and implantable pumps for drug delivery). For the development and proper functioning of artificial devices, the underlying control mechanisms of the normal and of the disabled physiological systems with which the external devices must interface must be adequately understood. Thus, in its broadest sense, the area of medical control systems encompasses all branches of engineering, mathematical biology, biophysics, physiology, and medicine. *See* BIOMECHANICS; BIOMEDICAL CHEMICAL ENGINEERING; BIOMEDICAL ENGINEERING; CONTROL SYSTEMS; MATHEMATICAL BIOLOGY.

The importance of control systems engineering in medical applications has grown because of the inherent complexity of medical control systems. Although there is no formal definition of complex systems, H. A. Simon's concept of complexity is very appropriate for medical control systems: complex systems are composed of subsystems that in turn have their own subsystems, and so on; and the large number of parts interact in a complicated way so that it is sometimes impossible to infer the properties of the whole from the properties of the parts and their laws of interaction. Indeed, the analytical models developed, using control systems engineering, of the components of a medical system have had limited success in predicting the behavior of the overall system.

Some examples of medical control systems and devices will be discussed. The example of control of voluntary movements emphasizes the complexity of physiological systems. Myoelectric prostheses are replacement devices for lost limbs. External orthoses are used for rehabilitation of patients with acquired disabilities. Implantable devices such as defibrillators and pumps for drug delivery represent advanced technology. Numerous other devices, such as cardiac pacemakers, artificial kidneys, heart-lung machines, and artificial ventilators, have been in routine clinical use for many years.

Control of voluntary movements. The control mechanisms by which the central nervous system produces coordinated or patterned movements has been an area of study from Galen in the second century to current physiologists, psychologists, clinicians, and engineers. In order for the central nervous system to produce any movement, (1) the appropriate muscles must be selected, (2) each participating muscle must be activated or inactivated in proper temporal relationship to the others, and (3) the appropriate amount of excitation must be exerted on each muscle to satisfy the desired goals. These three aspects require spatial, temporal, and quantitative coordination of muscular activities. The achievement of these goals can almost always be accompanied by any of a number of different movements. There is no unique movement trajectory to move from point A to point B in space. In planning and executing a movement trajectory, the central nervous system heavily relies on past experience and sensory feedback (visual, tactile, and proprioceptive inputs) from the environment. In addition, at all times the central nervous system must make appropriate adjustments to maintain the postural stability of the system.

Figure 1 illustrates some selected portions of the motor control system. Given a particular task, the central nervous system chooses a strategy to control the peripheral motor system consisting of the motoneuron pools and the muscles. This system consists of multiple feedback loops that continually provide information for updating the controlling signals. In human experiments, only the peripheral outputs of the motor system are observable. The electromyographic activity arises from the electrical responses in the muscle membranes. The forces produced and the resulting movements are outputs from the mechanical responses of the contracting elements of the muscles and the interacting external loads.

Significant progress has been made in developing mathematical models of subsystems of the motor system. However, an adequate description is far from

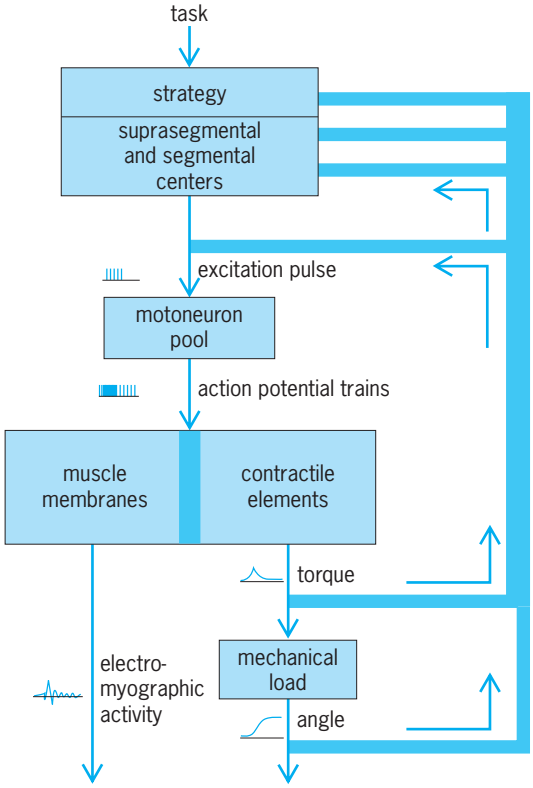

Fig. 1. Diagram of selected portions of the human motor control system. (*After G. L. Gottlieb, D. M. Corcos, and G. C. Agarwal, Strategies for the control of voluntary movements with one mechanical degree of freedom, Behav. Brain Sci., 12:189–250, 1989*)

complete. In order to apply the concepts of control theory to motor coordination, the nature of the input-output variables must be defined. Force, length, velocity, and so forth are the traditional choices of the variables to be controlled in artificial control systems. It is difficult to specify the controlled variables in a human motor control system. The nature of the controlled variable may very well be task-related and the system adapts to its changing environment. *See* MOTOR SYSTEMS.

Myoelectric prostheses. Myoelectric controlled prostheses are most commonly used for below-elbow amputees where the elbow function is retained. In such devices the terminal device is usually designed to look as much like a hand as possible (**Fig. 2**). Myoelectric control uses the electromyographic activity (Fig. 1) of a contracting muscle as a control signal. In a below-elbow amputee, the muscle remnants in the residual limb are used to provide control signals for the powered components.

As the level of amputation progresses to sites above the elbow, the complexity of function and power levels required both increase rapidly, while the capability of an amputee to operate these functions by harnessing body movements decreases rapidly. The human upper extremity has seven degrees of freedom from shoulder to wrist. In such devices, timely processing and adequacy of bioelectric command information from human controller, the amputee, to the machine becomes a critical problem. *See* PROSTHESIS.

Rehabilitation engineering. Rehabilitation engineering is concerned with the application of appropriate science and engineering in rehabilitation of patients who have acquired disabilities by accident, natural causes (including aging), or disease and are physically or mentally impaired. One particular application has been the development of functional electrical stimulation systems for spinal cord injury patients where the controlling signals from the central nervous system are interrupted by lesions in the spinal cord. By using externally applied electrical inputs, some functional capability can be restored in the affected muscles. Functional electrical stimulation applied on its own tends to physically exhaust patients because of the rapid onset of muscle fatigue. Biochemically designed orthoses that provide selective constraints of the hip, knee, and ankle joints combined with functional electrical stimulation have proven to be clinically successful. A hybrid system using both functional electrical stimulation and orthoses is shown in **Fig. 3**. This system involves decomposing the multigoal and multivariable control problem into subsystems organized in hierarchical levels. At the top level, the patient interacts directly with the control system through a command interface comprising a limited number of hand-operated controls. Each locomotion mode is controlled with reference to a finite-state model of the process. This model serves to change the control strategy, as required, at different stages or phases of the locomotion cycle. The control signals, as myoelectric signals, may also be obtained from muscles that are still functional. *See* MULTILEVEL CONTROL THEORY; MULTIVARIABLE CONTROL.

Implantable pacemakers. The implantable cardiac pacemaker is a major contribution to health care. Since the first clinical implant in 1960, the procedures for implanting pacemakers have been greatly simplified. It is no longer necessary to open up the chest to connect stimulating leads to the heart; these leads can be installed through the veins. Over 300,000 new pacemakers are implanted each year throughout the world, and in the United States cardiac pacemaker use has approached 1 per 1000 of population. The performance and reliability requirements of life-support implantable devices, such as pacemakers, often transcend those of traditional high-reliability fields such as the space program. The human body is a very hostile environment for implanted devices, far more so than outer space or the

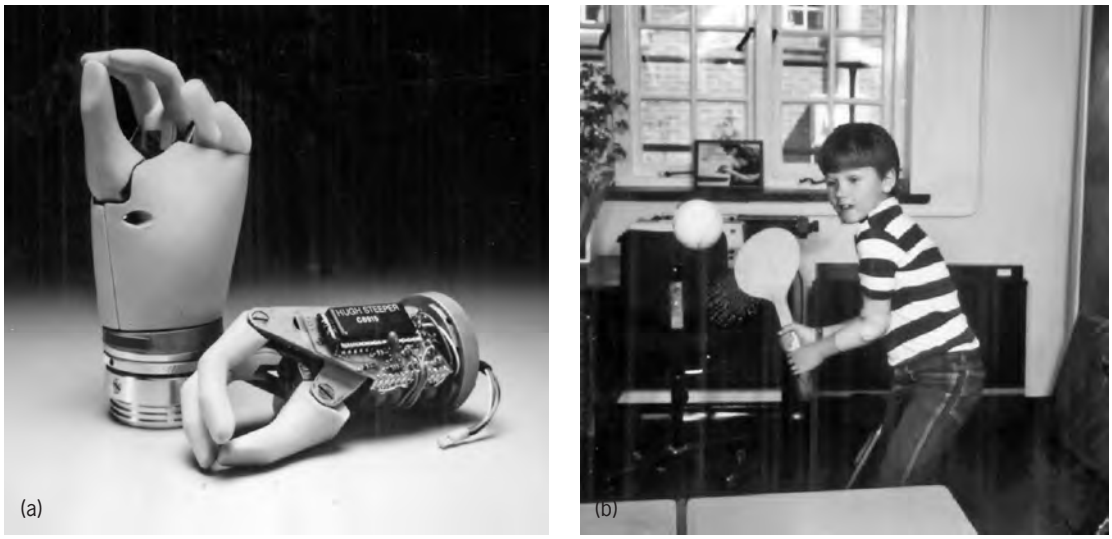

Fig. 2. Electric hand, a myoelectric prosthesis. (*a*) Two views of the hand, 2.5 in. (6 cm) across the knuckles, with the cover removed in one view to show the electric motor drive circuit. (*b*) Child using the electric hand. (*Hugh Steeper Ltd.*)

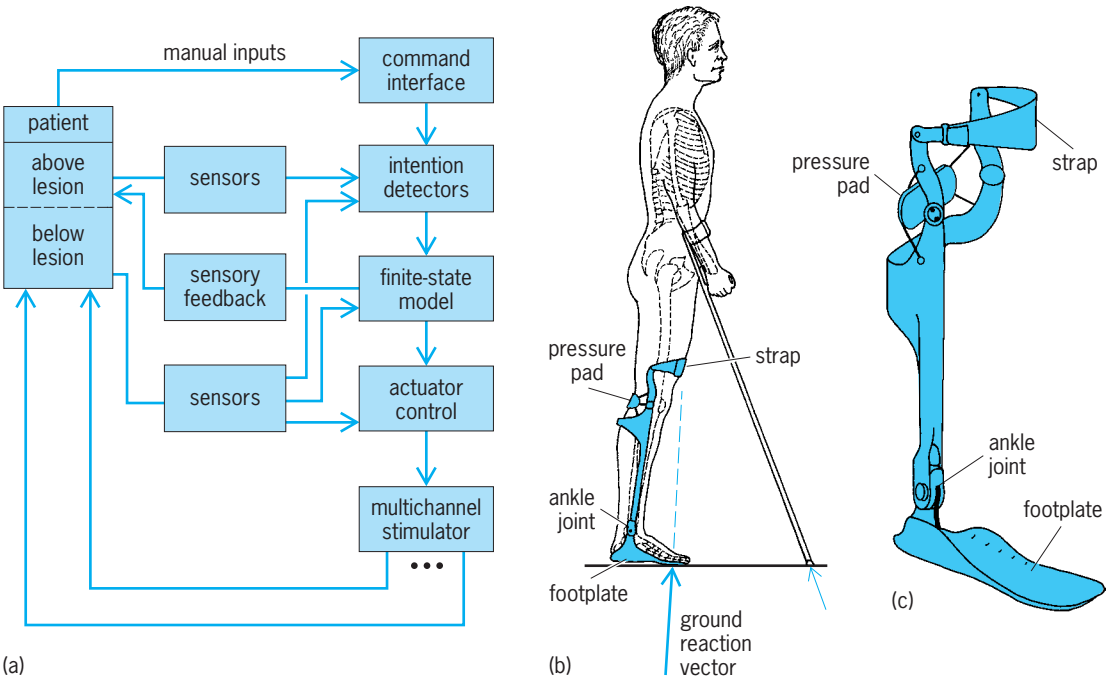

Fig. 3. Hybrid control system using both functional electrical stimulation and orthoses. (*a*) Hierarchical control scheme. (*b*) Braces in use. (*c*) Detail of leg brace. (*After Y. Kim and F. A. Spelman, eds., Images of the Twenty-First Century: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 11, 1989*)

sea bottom. The battery power source has always been the limiting factor in pacemaker performance and reliability. The lithium-iodine battery, invented in 1970, has been the most used battery for pacemakers since 1972. Such batteries have a long life (over 10 years in cardiac pacemakers), are autoclavable, and can be hermetically sealed in a metal, glass, or ceramic material because there is no gas generation. An autoclavable pacemaker with a lithium silver vanadium pentoxide battery is shown in **Fig. 4.** *See* SOLID-STATE BATTERY.

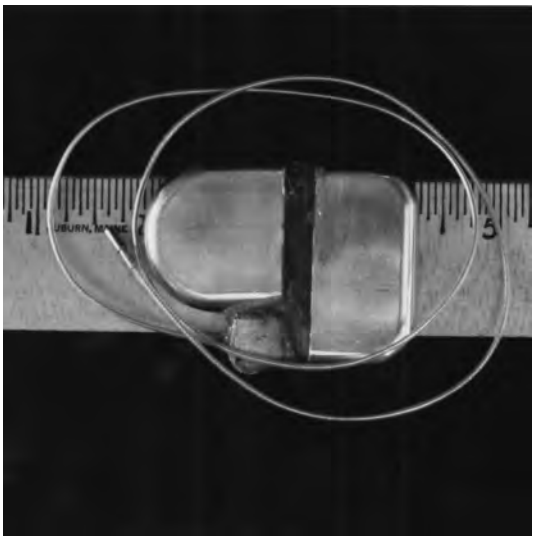

Fig. 4. Autoclavable pacemaker with a lithium silver vanadium pentoxide battery. (*After W. Greatbatch, Pacemaker power sources, IEEE Eng. Med. Biol. Mag., 3(2):15–19, June 1984*)

Developments in pacemaker technology include improving the pacemaker's responsiveness to the body's needs and providing more diagnostic information. A normal healthy heart responds to exercise and stress by increasing cardiac output through increased heart rate and stroke volume. Advanced pacemakers indirectly sense the need for increased heart rate by sensing the P wave, nerve impulses, the Q-T interval, hydrogen-ion concentration (pH), oxygen saturation, respiratory rate, stroke volume, motion, or body temperature. *See* CARDIAC ELECTROPHYSIOLOGY; CARDIOVASCULAR SYSTEM.

Implantable defibrillators. Sudden cardiac death, defined generally as death within 1 h of onset of ventricular tachyarrhythmias (VTs) or ventricular fibrillation (VF), is the cause of death in an estimated 350,000 patients per year in the United States. Fast response time of ambulatory service and immediate external defibrillation by skilled paramedics can save a significant number of so-called at-risk patients. Antiarrhythmic drugs and implantable anti-VT pacemakers are in clinical use. An automatic implantable defibrillator can essentially eliminate cardiac arrhythmic mortality. One such device, which has been approved for clinical use (**Fig. 5**), continually monitors the cardiac rhythm and, upon recognition of a shockable arrhythmia, delivers one or more pulses of 25–30 joules of energy using lithium batteries. An external programmer permits noninvasive interrogation and programming of the implanted device. The first-generation devices were large (volume of 160 mL or 10 in^3 and weight of 290 g or 10.2 oz), were implanted in an abdominal pocket, and required open thoracotomy. More advanced devices

Fig. 5. Automatic implantable cardioverter defibrillator (AICD) system with pulse generator, leads, and external programmer. (*S. Nisam, Cardiac Pacemakers, Inc.*)

do not require thoracotomy and use a transvenous defibrillating lead with sensing electrodes, which are implanted similiar to a standard pacemaker lead. In one study, all-cause mortality in the defibrillator group declined by 46% as compared with conventional pharmacological therapy, and sudden cardiac deaths in the defibrillator group decreased by 85%.

Implantable pumps for drug delivery. In a healthy person, the islet cells of the pancreas continually monitor the blood of all metabolic substrates. Infinitesimal secretions from these cells delicately balance the metabolism over a wide range of food intake rates (from feasting to fasting). Insufficiency of the endocrine pancreas leads to diabetes mellitus. The control of diabetes in most of an estimated 15 million insulin-dependent diabetic patients in the United States is by subcutaneous exogenous injections once or twice daily. Such therapy is open-loop, and feedback control is exerted at regular intervals (mostly weekly or monthly) by blood analysis and dosage adjustments. Negative-feedback closed-loop control systems for blood sugar control have been implemented. The major difficulties have been in finding a suitable and reliable sensor of blood glucose and in miniaturizing the entire device so that it can be implanted. Implantable pumps for variable-rate delivery of drugs are at various stages of development. *See* DIABETES; PANCREAS. Gyan C. Agarwal

Bibliography. N. A. M. Estes III, A. S. Manolis, and P. J. Wang (eds.), *Implantable Cardioverter-Defibrillators*, 1994; U. Laffer, I. Bachmann, and U. Metzger (eds.), *Implantable Drug Delivery Systems*, 1991; G. V. Naccarelli and E. P. Veltri, *Implantable Cardioverter-Defibrillators*, 1993; D. A. Wolfe, D. Kosinski, and B. P. Grubb, Update on implantable cardioverter-defibrillators, *Postgrad. Med.*, 103(1):115–130, January 1998.

Medical imaging

A medical specialty that uses x-rays, gamma rays, high-frequency sound waves, and magnetic fields to produce images of organs and other internal structures of the body. In diagnostic radiology the purpose is to detect and diagnose disease, while in interventional radiology, imaging procedures are combined with other techniques to treat certain diseases and abnormalities. *See* RADIOLOGY.

Contrast media. Contrast media are drugs or compounds that block the x-rays and are used to outline and distinguish certain anatomic structures from surrounding tissues of similar density. The two principal groups of contrast agents are barium compounds and iodine-containing solutions.

Barium compounds are swallowed or are given as an enema to evaluate the different portions of gastrointestinal tract. Iodinated solutions are usually injected into a vein, an artery, or the spinal canal. Thus, with appropriate techniques they are used in angiography to visualize arteries and veins, and in myelography to study the spine, spinal cord, and spinal nerves. Iodinated drugs are also used to study the shape and function of the kidneys and to enhance tissue differences during computed tomography.

Film x-rays. Film x-ray studies, the most common radiologic procedures, are made up of still pictures of the various organs and tissues in the body. In these procedures, x-rays are passed through the body to expose the photographic film that is placed on the opposite side of the body. As they pass through the body, the x-rays are absorbed in different degrees by different types of body tissues. Dense tissue, such as bone, absorbs, or blocks, a greater amount of radiation and causes the film to appear lighter or whiter. Fatty tissue and air-containing structures, such as the lungs, block the least radiation and cause the film to appear darker or blacker. Other tissues appear as varying shades of gray on the film. The changes in film density that result from exposure allow the radiologist to distinguish between normal and abnormal tissue and to diagnose many different disease types.

The chest x-ray is the most frequently performed radiologic study. Other commonly performed procedures are x-ray examinations of the abdomen, spine, extremities, sinuses, skull, and kidneys. *See* X-RAYS.

Mammography is an x-ray examination of the breast. In conjunction with physical examination of the breast, it provides the most effective method for early detection of breast cancer. Specially designed equipment offers a high-quality image of the breast tissue with the lowest possible radiation risk since the dose is very low. Vigorous compression of the breast is required in order to obtain maximum detail of the tissue and to minimize the radiation exposure (**Fig. 1**). Two views of each breast are customary—one from the side and one from the top.

Fluoroscopy. W. C. Roentgen's fluoroscopic unit was developed in 1896 and provided the basis for the modern fluoroscope, which includes an x-ray tube

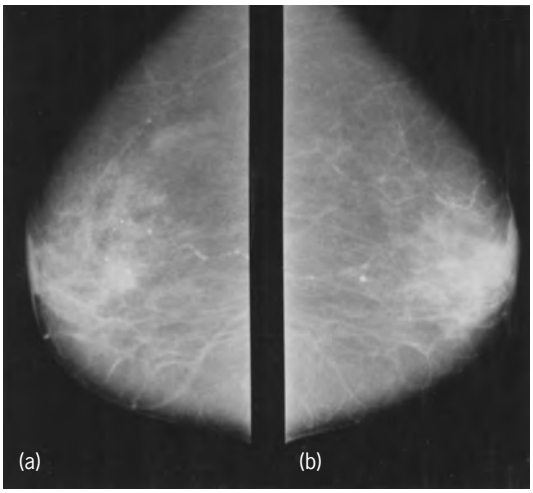

Fig. 1. Mammogram showing lateral views of (*a*) left and (*b*) right breasts. (*American College of Radiology*)

and a fluoroscopic screen. The screen, however, has been incorporated into an image intensifier, which is an electronic instrument that brightens and intensifies the fluoroscopic image many thousands of times. This process has resulted in a dramatic improvement in image quality while markedly reducing radiation exposure. Fluoroscopy is a dynamic x-ray imaging technique that produces a moving image over time. It is essential for evaluating organ movement such as the beating of the heart or movement of the diaphragm.

The gastrointestinal (G.I.) series and the barium enema are the most common fluoroscopic studies. These procedures begin with the administration of a barium mixture either by ingestion or by an enema that fills the stomach or large intestine. The barium mixture, like dense tissues, blocks the x-ray beam. Fluoroscopy then reveals the location of the barium-coated lining of the stomach and intestine and enables the radiologist to observe as they contract and distend. Ulcerations, tumors of the intestine, and many other abnormalities can be seen with this technology. Fluoroscopy is also used in angiography and in interventional radiologic procedures.

Angiography. Angiography is the radiologic study of blood vessels. Because arteries and veins are not normally visible in conventional x-ray studies, an iodinated compound, which is opaque to the x-ray, must be injected into the bloodstream. An arteriogram is an x-ray study of the arteries; a venogram is an x-ray study of the veins. Arteriography is most often used to show the presence and extent that arteries have become clogged and narrowed by arteriosclerosis, which can lead to strokes and heart attacks. It is also performed to locate sites of internal bleeding or tumors or to define many other conditions that may obstruct normal blood flow to body tissues.

A radiologist performs an arteriogram by inserting a long, narrow tube, called a catheter, through a needle puncture in an artery, usually at a point in the groin. Guided by fluoroscopy, the catheter is carefully threaded through the branching arteries of the body to the location to be examined. The contrast solution is then injected through the catheter into the blood vessel. The appearance of the opacified arteries is then recorded on a series of rapidly exposed x-ray film images (two to six pictures per second): an x-ray moving picture, called a cine-radiograph, or videotape recordings. Cine-radiography is frequently used to study the coronary arteries, which supply the heart muscle.

Digital subtraction angiography is a computerized technique for recording x-ray images. Two images are made of the same organ—one before and one after injection of contrast medium into a blood vessel. A computer then subtracts the first image from the second and produces a third image, which shows an unobstructed image of the blood vessels. Digital subtraction angiography helps improve the safety and patient comfort of the angiographic procedure because the amount of contrast solution required is less than that needed for traditional angiography.

Computed tomography. Computed tomography (CT), also called computed axial tomography (CAT), is a scanning technique that combines computer and x-ray technologies. To produce the scans, the patient is positioned on a narrow table (gantry) that slides into the scanner, which is a tubelike structure. An x-ray tube built into the scanner emits a very narrow x-ray beam as it rotates around the body. The radiation that passes through the body is recorded by detectors that lie opposite the x-ray tube. Each detector emits tiny flashes of light in proportion to the amount of radiation it receives. Because x-rays are differentially absorbed by body tissues, the brightness of the light is also a measure of tissue density. A computer measures the flashes of light, magnifies them a thousandfold, and then stores their location and brightness in its memory. From this information, the computer constructs a two-dimensional anatomic image that represents a cross-sectional slice through the body (**Fig. 2**).

Computed tomography can produce images of internal slices of organs, such as the liver and kidney, that are more precise than film x-rays. It is also useful in visualizing the brain and abdomen to detect trauma damage or the size and location of some

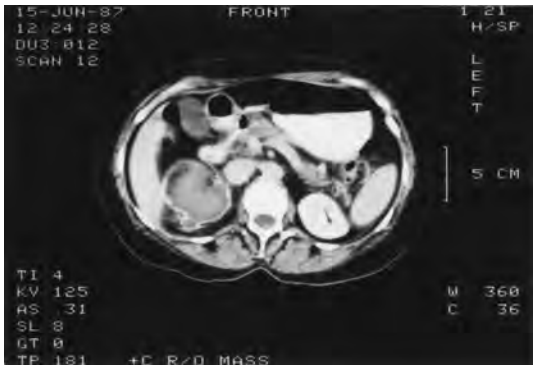

Fig. 2. Computed tomographic image showing a transverse section through the abdomen. (*American College of Radiology*)

tumors. Oral and injected contrast media can further enhance the CT image. Three-dimensional images can be generated by using special computer software (**Fig. 3**). These are especially useful in planning reconstructive orthopedic or plastic surgery. *See* COMPUTER GRAPHICS; COMPUTERIZED TOMOGRAPHY.

Ultrasound. Ultrasound imaging, or sonography, is a diagnostic imaging procedure that uses high frequency sound waves instead of ionizing radiation. During an ultrasound examination, a lightweight transducer is placed on the patient's skin over the part to be imaged. The transducer produces sound waves that penetrate the skin to reach tissues and organs. When the sound waves strike specific tissue surfaces, echoes are produced. The echoes are detected by the transducer and are than electronically converted into an anatomic image that is displayed on a video screen. The image can also be recorded on film or videotape. Some ultrasound studies are done in a real-time format, that is, showing motion as it occurs.

Ultrasound imaging is commonly used in obstetrics to monitor the position and development of the fetus (**Fig. 4**) and also to detect any fetal abnormalities or problems in the pregnancy. Ultrasound is also used to show problems in other internal structures, including the gallbladder, kidney, and heart.

Doppler ultrasound can monitor blood flow through veins and arteries. It is commonly used to study kidney transplants and blood flow to the brain and also to diagnose blocked arteries. *See* DOPPLER EFFECT; MEDICAL ULTRASONIC TOMOGRAPHY; ULTRASONICS.

Magnetic resonance imaging. Magnetic resonance imaging (MRI) is a diagnostic procedure that uses a large, high-strength magnet, radio-frequency signals, and a computer to produce images. During an examination, a patient is placed in an MRI scanner where the body is surrounded by a magnetic field up to 30,000 times stronger than that of the Earth. The nuclei within the atoms that make up the body's tissues are like tiny magnets, each with a north and south pole. The nuclei usually spin in many different directions and at many different angles, but in the presence of the strong external magnetic field generated by the scanner, the north and south poles of the nuclei align themselves with the external magnet. The nuclei continue to spin in many different directions (incoherent spinning). The magnet causes them to vibrate at a specific frequency, so that they act like radios, receiving and transmitting radio-frequency signals. When an external radio-frequency signal from the scanner is transmitted to the area of the body being examined, it changes the alignment of the magnetic poles of the nuclei so that they are no longer in alignment with the external magnet. It also causes the nuclei to spin in the same direction (coherently). When the signal stops, the magnetic poles of the nuclei realign themselves in the direction of the magnet and, at the same time, release their own faint radio signals. The scanner receives the radio signals and, with the aid of a computer, characterizes the different body tissues by the strength and duration of these signals. The computer then reconstructs the information into a two-dimensional anatomic image.

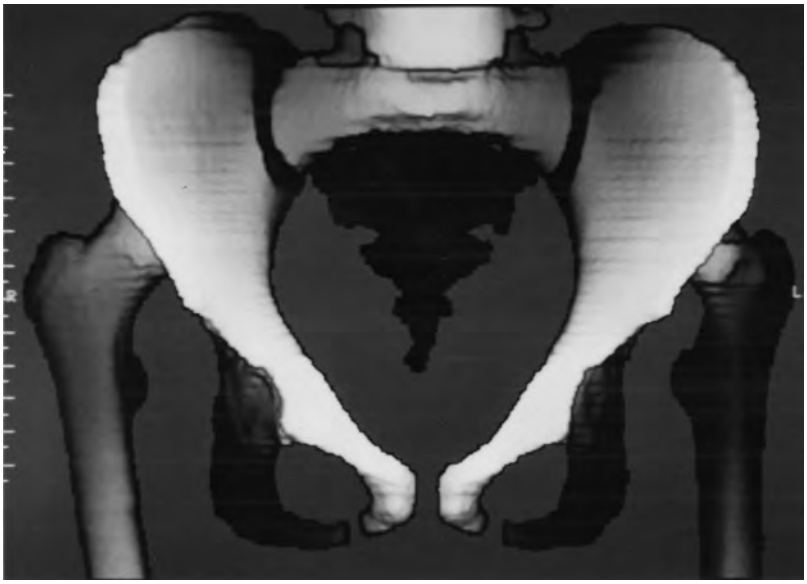

Fig. 3. Three-dimensional computed tomogram of the pelvis. Special computer software has made such imaging available for planning of reconstructive surgery. (*Dimensional Medicine, Inc.*)

The technique of MRI is extremely useful in evaluating diseases of the brain and spine (**Fig. 5**). In some instances it may eliminate the need for arteriography of the brain or myelography of the spinal column. It is also used to evaluate joints, bone and soft tissue abnormalities, as well as abnormalities of the chest, abdomen, and pelvis.

Because of the strong magnetic field involved, some patients are unable to undergo MRI examinations. That is, patients with iron-containing metallic clips in the brain or fragments in the eye might not be scanned as the magnetic field might cause these to move and cause damage. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

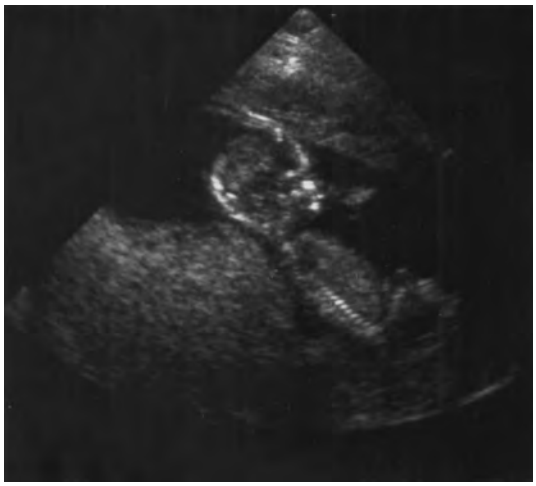

Fig. 4. Ultrasound image of a growing fetus within the uterus. (*American College of Radiology*)

Fig. 5. Magnetic resonance image of a midsagittal section through the skull. (*American College of Radiology*)

Nuclear medicine. Nuclear medicine imaging studies use radioactive compounds called radionuclides of radiopharmaceuticals that emit gamma rays, or some emit beta particles. The chemicals are formulated so that they collect temporarily in the parts of the body to be studied. For most nuclear imaging studies, the radionuclide is injected into the patient; other procedures require that it be inhaled or swallowed. The radiologist chooses the radioactive substance that will produce the best image of the organ or body part under study using the least amount of radiation. Depending on the type of study, the images may be obtained immediately or several hours or days later. The administration of radionuclides for diagnostic procedures is virtually risk-free, but it is not recommended for pregnant women since the fetus appears to be slightly more sensitive to radiation than adults.

The images are taken with a gamma camera suspended above the patient who lies on a table. The camera detects the gamma rays emitted from the radionuclide in the patient's body and uses this information to produce an image that shows the distribution of the radionuclide within the body. The image is recorded on film and is called a scintigram or scan.

Scintigrams of the heart and bone are the two most common nuclear medicine examinations. A cardiac scan is used to evaluate cardiac function and blood flow to the heart muscle. A bone scan (**Fig. 6**) may detect the presence of cancer, infection, or trauma. Other commonly performed radionuclide studies yield scintigrams of the lungs, liver, gallbladder, kidney, thyroid gland, and brain.

Single-photon emission computed tomography. The single-photon emission computed tomography (SPECT) examination uses a computer to obtain two-dimensional images that are thin slices of internal organs such as the heart, brain, and liver. The SPECT images can display organs with much greater detail than conventional scintigrams. Three-dimensional organ displays and cine displays also can be obtained. The cine displays, which resemble moving pictures, are used to evaluate heart muscle contraction.

Positron emission tomography. Positron emission tomography (PET) is a more refined radiologic technique that is used to study the metabolic activity inside an organ. The technique has been shown to be useful in the study of brain-related disorders, such as epilepsy and Alzheimer's disease, and of the vitality of heart tissue.

Interventional procedures. Interventional radiology combines imaging procedures with various injection and catheter techniques to treat tumors, blockages, bleeding vessels, and other abnormalities without extensive surgery. Among the more common interventional procedures is angioplasty, which is used to treat blocked or narrowed arteries. In this procedure a catheter is inserted through the artery to the narrowed segment. There, a balloon at the tip of the catheter is inflated to dilate the narrowed segment. The balloon is then deflated and withdrawn. A combined laser–catheter procedure is also used to treat blocked arteries. The laser vaporizes most of the plaque, and then a balloon is used to dilate any narrowing of the artery.

Interventional radiologists can remove some kidney stones and gallstones by using a specialized

Fig. 6. Bone scan. Typically used to diagnose the spread of cancer or other tumors, it is one of the most common nuclear medicine procedures. (*American College of Radiology*)

catheter as a retrieval device. A catheter can also serve as an aspirator to drain pockets of fluid. *See* RADIOGRAPHY. Michael Lopiano; Raymond Del Fava

Bibliography. D. Avedon, *Electronic Imaging Systems: Design Applications and Management*, 1993; D. Avedon, *Introduction to Medical Imaging*, 1993; R. H. Mohiaddin and D. B. Longmore, *Atlas of Whole Body MRI*, 1992; M. F. Sturman, *Effective Medical Imaging: A Signs and Symptoms Approach*, 1993.

Medical information systems

Standardized methods of collection, evaluation or verification, storage, and retrieval of data about a patient. The three broad areas of any information system—input, data transformation, and output—suffice to describe an existing system, but they are insufficient to design a new one. To these must be added the action that is expected to take place on the basis of the data output from the system, thereby defining the purpose of the system, and the feedback from such action to the system input, which places the system in a specific medical environment. Crucial to any information system, and especially for a medical system, is the accuracy of the input data, which is of more significance than mere precision. A medical information system is a part of necessary time management that is intended to maximize the amount of data and procedure information in order to make the most accurate clinical decisions. As such, an information management system must be interactive to serve clinical decision purposes.

The medical environment is highly data- and information-intensive. Hospitals have always managed data and information and were pioneers in the use of technologic information systems. With scaled-down technology, medical offices are able to incorporate data and information management in the private-practice environment as well.

Input. Data input is provided by the patient and by other information sources involved. Input information is complex because the patient must be described both cross-sectionally, which is current information, and longitudinally, that is back to childhood and to the parents. Historical information obtainable from the patient is subject to increasing error as the individual goes backward in time. Significant error in the medical information system can arise if imperfect recall is not considered and appropriately discounted and historical data are not weighted according to the psychological status of the patient. Unfortunately, this may be done easily only in retrospect. In addition to retrieving facts, a major function of the physician's historical interview is placing all facts in their true light according to the subject's psychological makeup. Data from outside sources are subject to the same errors; hence it is valuable to obtain objective data rather than subjective data from either input source.

Database systems. The medical record database of the Mayo Clinic in Rochester, Minnesota, is one example of systematic cross-sectional and longitudinal recordkeeping efforts. There physicians work with a history of all the patients treated within—and in many cases, outside—the clinic facilities. A standardized coding system is used, and there are two linked segments of the system, medical practice and records: A large referral practice covering every medical and surgical subspecialty, and a primary health care component for the surrounding area residents. All recorded diagnoses made by physicians at the clinic or at affiliated hospitals are indexed. Thus, every formal medical contact of each person (for many from birth to death) is recorded, indexed, and stored in the database. The longitudinal history requires substantiated observation, repeated and confirmed laboratory data, and other facts. The worth of data, however, is dependent on the validity of the information at the time it was acquired.

Mechanical and electronic data collection. Complete and accurate laboratory and procedure data obtained during a cross-sectional evaluation of the patient is the best input for longitudinal patient review and for storage for future retrieval. Thus, the collection of such hard data to confirm, or better, to supplant soft data input is stressed, which means that the data are of the highest quality. Any information system that intends to use engineering technology for processing or storage must stress data acquired directly from a subject by a mechanical or electronic system. Computers and semiautomatic data-gathering systems must serve as dedicated basic components to a complete clinical information system.

Technology provides data on which decisions can be made. The impact of technology on clinical decision making is evident when new techniques are found that can improve the ability to diagnose, cure, or control patients' health problems and concerns. Many technological achievements receive little public notice. For example, because laboratory equipment has become smaller and less expensive to own or operate, private-practice physicians can perform more laboratory tests in their offices. Although the advantage of this to patients is less obvious, in many ways it is even greater than some of the more sophisticated achievements. If a laboratory test result is received quicker, the diagnosis may be quicker so that appropriate therapies can be started, in some cases within the space of one visit to the office.

The goals of technology in the clinical environment are improvement of clinical technique, patient benefit, and cost effectiveness.

In order to achieve these goals, technology must meet the following confidence criteria: (1) Reliability: the degree of assurance that the measurement system will yield reproducible results over a period of time. (2) Reproducibility: that a measurement method applied to an object will yield the same answer over and over again. (3) Precision: that the measurement can be obtained within a reproducible and exacting range. (4) Accuracy: that the numbers produced are precise and relate to some absolute reference point in relationship to a disease. Regardless of the standards used, most assessments of devices by

clinicians focus on whether the device performs as expected. If not, the device has failed.

Characterizing data and information. In most clinical environments, patient-related data and information are stored in a patient's chart. The variety of data required to formulate information to provide medical care includes the following:

- Patient identification data
- Primary medical visit information
  - Patient-described
  - Physician/assistant-described
  - Device-described (laboratory data, electrocardiograph, spirometer, and so on)
  - Action taken (prescription, referral to a specialist, request for additional data such as x-rays, and so on)
- Medical data from other sources
  - Out-of-house laboratory data
  - X-rays
  - Medical information from others (physicians, hospitals, and so on)
  - Observations of others (employers, family members, and so on)
  - Requests for summary of medical information (from nursing care services, insurance companies)
  - Financial data
  - Current bills
  - Ongoing billing
  - Insurance information

The human mind must integrate many of the pieces of these data, at times in a multidimensional sense, but always in a checklist fashion. The goal of a medical information system is to integrate each type of data into a comprehensive, interactive database system. Thus, the medical information system can take care of the checklist and leave people to think, react, and interact creatively. Data gathering, which is performed at the point of decision making by the decision maker (the health care provider), will be subject to less bias when a medical information system provides the checklist for the decision maker.

Much that is considered hard data is subject to error through human reinterpretation; human variability in medical data interpretation is acknowledged to be about 20%. This implies that even the hard output of a laboratory device of known precision and accuracy is softened by the human observer. For example, a computer system reads electrocardiograms with an accuracy of 89% and with the mechanical reliability of the system at 97%. Compounding these two factors results in a performance rating for the total system at 86%. If the system's printout is reinterpreted by an observer with human variability, the performance of the computer–human system decreases further: with 10% human variability, the system could be rated at 77% and with 20% human variability at 68%. This underscores the importance of quality control and the fact that the appropriate role for a human in such circumstances is also to concentrate on that part of the system that cannot be read by the computer.

With increasing use of engineering and other technological advances, the quantity of soft data in use can gradually be diminished. A patient's history of a heart attack, for example, can be accessed, but such a history can be relatively soft data. Hard data would comprise the actual electrocardiogram interpretation by a computer system along with the date of processing. Availability of such minimal information markedly diminishes the need for preserving a large soft-data file of patient-derived information.

In addition to the patient and the data derived directly from the individual, outside sources should also be used to provide information for the medical records, including the family, social service agencies, insurance records, and other medical or clinic records. Actually, surprisingly little data are readily available from outside the immediate family, although in many instances the data may be harder in a factual sense than the patient's data. Since retrieval of such data is often cumbersome and delayed, little benefit is possible to the patient as a result of acquiring it. Regrettably in patient care, a service's value is to a large extent dependent on whether the data are immediately available. Without speed, retrieval can be of little value; hence outside data are little used. Another obstacle that arises in the use of device-generated databases is the lack of comparability standards. For example, a reflectance photometer, which quantitatively measures various serum or plasma constituents as well as blood hemoglobin, is a simple instrument that uses color coding for numerical output in a one-step or two-step process. However, the advantage of its simplicity can be undermined because of the difficulty that is associated with calibration and reliability of resultant values, and also because scales for values of procedures may differ from conventionally used values.

Terminology and nomenclature. Terminology is the language used specifically by a specialized field, and nomenclature is the systematization of that language or vocabulary. In clinical medicine, terminology is well defined and widely used, but the development of a specific nomenclature is essential to useful medical information systems.

Two nomenclature systems that are widely accepted are *Current Procedure Terminology* (CPT) and *International Classification of Diseases* (ICD). The application of the nomenclature, however, requires further standardization, including (1) definition of patient data collection standards; (2) definition of patient identification standards for longitudinal medical records (constrained by computer systems and the perceptions of their programmers); (3) automated conversion of natural language input into machine-oriented language; (4) hierarchy and semantic organization of medical terms; and (5) development of parameters for databases that allow for input and output in a non-machine-oriented mode. If the objective is to develop a bridge between current practices in clinical medicine and interactive databases, the upgrading of

the nomenclature system is important, since terms are the primary carriers of information.

Evaluating databases. Clinical databases can address several evaluation goals, but foremost in medicine is clinical efficacy. In order to assess the worth of a database, three basic aspects must be considered: data collection methods, ongoing management, and output reporting of data. The clinical efficacy of a procedure or technology can be diagnostic, managemental, or prognostic.

The data collected are as good as their collection medium; data obtained with minimal subjectivity and directly from the source are of optimal quality. Database elements must be carefully defined and include descriptions of the patient, test, and outcome. Complete characterization of the patient needs to be made as well, so changes that result from patient management can be identified. The best data are collected by those who have the greatest interest in its completeness, that is, the individuals who are caring for the patient.

Tying data collection to patient care has three advantages: the cost of data collection is reduced by using patient care channels; the physician who is responsible for generating the reports must review the collected data; and the data collection then becomes prospective without extensive chart reviews. When data collection is tied to patient care, the clinicians must be responsible for developing standardized methods to ensure that no time is wasted in collecting useless data and that the reporting methods used do not conflict with those used in reporting patient care. Establishing and maintaining the database is the responsibility of a team headed by the clinicians and including biostatisticians and computer scientists. The clinicians keep the research process focused on the appropriate clinical goals, the biostatisticians analyze the data gathered, and the computer scientists design and employ the appropriate technological systems to achieve the goals of the effort.

Transformation of data. In a medical information system, input information can be processed through one of two parallel, sometimes mutually exclusive routes—administration and medical service organization. Administration encompasses the admitting clerk, administrative officers, pharmacists, and all those who have no direct patient contact. The service route includes the physician, the nurse, and all others who have direct patient contact. The two routes differ in their method and degree of detail. For example, a patient's inability or failure to pay a bill might be considered by the administration in detail in order to set in motion a social service structure. This same information might be of no significance to the nurse or physician, except that the psychological character of the patient is pertinent to the diagnosis or illness under consideration. The passage of such medical information through both routes allows data to reach the output stage. Redundancy, as illustrated by parallel information routes in hospitals, can be a source of economic drain, but a degree of redundancy is necessary for the quality of service required and is a fail-safe feature that is vital to all aspects of quality. Redundancy is particularly important because those in direct contact with the patient are subject to human error or variability. Quality can be maintained by always presenting hard information or by the use of redundancy, but economics may favor redundancy over a constant supply of hard data.

Manufacturers of devices, clinicians (users), engineers, and voluntary and nonvoluntary regulators all have different data requirements even when the object of their evaluation is the same. The principal elements, or uniform terminology, in establishing a database for or evaluation criteria of any problem or device are rationale, principles, population, measurement definitions, protocol, device characteristics, patient characteristics, provider characteristics, format, and outcome. When data are transformed according to these basic elements, each user will be able to answer any questions that arise from the perspective of his or her particular expertise.

Output. The output of a medical information system can be viewed as a control system for actions by clinicians. The output, once it has passed through the transformation media, constitutes the patient's medical records and usually includes the history and physical examination data, progress notes, nursing notes, doctor's orders, and summaries and notations of values from physical laboratory data.

Items from the administrative system are usually not found with the service records. Both represent patient information or input that has been translated through the parallel media of medical personnel and administration, but their separation is indicative of the problems that face planners of good medical information storage and retrieval. Items of output information are interrelated to establish a diagnosis of a course of action. Feedback from that stage to the patient, often in the form of interrelations, can be called patient care.

In addition to its primary function of facilitating patient care, a medical information system can also prove useful for investigation and research into disease and therapy, and the administrative portion of the record can aid in planning future improvements in quality of service. Efficiency should be directed primarily to the medical areas to ensure quality care for the greatest number of patients. *See* DATA COMMUNICATIONS; DATABASE MANAGEMENT SYSTEM; MEDICAL CONTROL SYSTEMS.

Cesar A. Caceres

Bibliography. C. A. Caceres (ed.), *Management and Clinical Engineering*, 1980; C. A. Caceres (ed.), *The Management of Technology in Health and Medical Care,* 1980; C. A. Caceres et al. (eds.), *Medical Devices: Measurement, Quality Assurance, and Standards*, 1983; B. J. Jaeger et al. (eds.), *Multi-Institutional Systems Management*: *Concepts and Cases*, 1987; H. Lamberts and M. Wood, *ICPC International Classification of Primary Care*, 1988; M. Rubin (ed.), *Computerization and Automation in Health Facilities*, 1984; K. Waters and G. Murphy, *Medical Records in Health Information*, 2d ed., 1994.

Medical mycology

The study of fungi (molds and yeasts) that cause human disease. Some pathogenic molds and yeasts normally reside within soil or derive their nutrition from other organic matter until introduced into the body by inhalation or trauma; others are part of the normal body flora or are transmitted from an infected person. Because the immune status of the host plays an important role in susceptibility to fungal infection, highly immunodeficient persons are likely to develop an opportunistic fungal infection. *See* OPPORTUNISTIC INFECTIONS.

Fungal infections are classified according to the site of infection on the body or whether an opportunistic setting is necessary to establish disease. Fungal infections that occur in an opportunistic setting have become more common due to conditions that compromise host defenses, especially cell-mediated immunity. Such conditions include acquired immunodeficiency syndrome (AIDS), cancer, and immunosuppressive therapy to prevent transplant rejection or to control inflammatory syndromes. Additionally, opportunistic fungal infections have become more significant as severely debilitated individuals live longer because of advances in modern medicine, and nosocomial (hospital-acquired) fungal infections are an increasing problem. Early diagnosis with treatment of the fungal infection and control of the predisposing cause are essential.

Antifungal drug therapy is extremely challenging since fungi are eukaryotes, as are their human hosts, leading to a paucity of specific fungal drug targets and also to problems with toxicity or cross-reactivity with host molecules. Most antifungal drugs target the fungal cell membrane or wall. The "gold standard" for therapy of most severe fungal infections is amphotericin B, which binds to ergosterol, a membrane lipid found in most fungi and some other organisms but not in mammals. Unfortunately, minor cross-reactive binding of amphotericin B to cholesterol in mammalian cell membranes can lead to serious toxicity, especially in the kidney where the drug is concentrated. Azole drugs such as ketoconazole interfere with ergosterol biosynthesis by inhibiting the enzyme lanosterol demethylase, but can also cross-react with mammalian P-450 enzymes. Another class of drugs, the echinocandins, inhibit β-glucan synthesis, thereby interfering with a fungal cell wall component not found in mammalian cells. Recent advances in antifungal therapy include the use of liposomal amphotericin B and newer azoles such as fluconazole and itraconazole, which show reduced toxicity or greater specificity. Conversely, drug resistance in pathogenic fungi is an increasing problem, as it is in bacteria.

Opportunistic diseases. Candidiasis is the most common opportunistic fungal infection, and it has also become a major nosocomial infection in hospitalized patients. *Candida albicans* is a dimorphic fungus with a yeast form that is a member of the normal flora of the surface of mucous membranes. In an opportunistic setting, the fungus may proliferate and convert to a hyphal form that invades these tissues, the blood, and other organs. Altered host resistance from different causes usually determines the site of invasion and degree of invasion. Abundant growth on the oral mucosa leads to superficial infection with white patches (thrush). Women, especially those who are diabetic or pregnant, may develop a similar invasion of the surface of the vagina (candida vaginitis). In addition, candidiasis may involve the respiratory or urinary tracts, with colonization followed by serious infection. The disease may extend to the blood or other organs from various infected sites in patients who are suffering from a grave underlying disease or who are immunocompromised.

Aspergilli, which are common saprophytic molds, cause a variety of diseases (aspergillosis). For some individuals, a respiratory tract allergy (allergic rhinitis) develops to airborne spores. For others, asthma may develop. These molds, as well as others, are able to colonize burn wounds and the ear canal and to grow in slowly moving mucus in the bronchi of persons with inhalation allergies (allergic bronchitis) as well as within cavities in the lung caused by tuberculosis (aspergilloma). In severely immune compromised individuals, such as those with acquired immune deficiency syndrome (AIDS) or cancer, eventual extension into surrounding normal tissue or the blood usually follows colonization and can be fatal. Some species of aspergilli produce toxins (aflatoxin) when growing on food (peanut products, stored grain); ingestion of the toxin may result in liver destruction or liver cancer. Toxin-related problems appear more commonly in tropical countries. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); AFLATOXIN; ALLERGY; ASTHMA.

Mucormycosis, also known as zygomycosis, is an opportunistic fungal disease that can be caused by any of a number of related environmental molds. Most commonly, a ketoacidotic diabetic patient develops a mold infection of the sinus that spreads rapidly to the eye and brain (rhinocerebral mucormycosis). Necrosis of the tissue develops, and the patient dies within only a few days unless the predisposing factors are corrected and therapy is begun.

Cryptococcosis is seen in persons with AIDS, in other immunocompromised persons, and occasionally in individuals with no known predisposing conditions. The infection, which has increased greatly in incidence since 1980, is caused by *Cryptococcus neoformans*, an environmental yeast that grows abundantly in dried pigeon droppings. Respiratory disease follows inhalation of the yeast, and subsequently infection may spread to the brain and other organs. Meningitis and brain abscess are the most frequently seen manifestations of disease. *See* MENINGITIS.

Pneumonia caused by *Pneumocystis carinii* is one of the mostcommon and most serious diseases in AIDS patients and some other immunocompromised populations. This organism has presented challenges in laboratory study and in classification. There is currently no continuous in vitro culture system for *P. carinii*, and consequently most studies depend on passage of the organism in infected laboratory

animals, making it impossible to obtain isolated clonal populations of the organism. It was long considered a protozoan parasite due to its morphology and also to other considerations, such as the presence of cholesterol and lack of ergosterol in its membranes, its consequent insensitivity to amphotericin B, and its sensitivity to some antiprotozoal drugs such as pentamidine. However, it does possess some fungal characteristics, and recently nucleic acid–based typing methods have clearly classified *P. carinii* as a fungus. An environmental reservoir or niche has not been identified, but it is thought that the organism is inhaled by humans very commonly, establishing a significant infection only when the host is severely immunocompromised. The most common clinical manifestation is severe pneumonia with significant respiratory compromise that can be fatal, but the fungus can also disseminate elsewhere in the body.

Nonopportunistic systemic diseases. Healthy persons can acquire disease from certain pathogenic fungi following inhalation of their fungal spores. The so-called deep or systemic mycoses are all caused by different species of soil molds; most infections are unrecognized and produce no or few symptoms. However, in some individuals infection may spread to all parts of the body from the lung, and so treatment with amphotericin B or an antifungal azole drug is essential. The deep mycoses are caused by certain white soil molds that change into either budding yeasts or spherules when growing within the body or in culture at 98.6°F (37°C). Each species of mold has a characteristic pathogenic and saprophytic morphology.

Coccidioides immitis, a mold of desert soil, converts into spherules containing endospores when growing within the body and causes coccidioidomycosis or valley fever. Although 60% of infections produce no symptoms, the other 40% develop symptoms that range from a flulike condition to pneumonia. Of the serious pulmonary infections, 1% spreads to other parts of the body. Meningitis and deep skin or bone lesions often develop.

Histoplasma capsulatum grows in moist, rich soil, often concentrated in areas contaminated by bird or bat droppings, and has a worldwide distribution. In the United States, it is associated mainly with the Mississippi and Ohio river valleys in the Midwest and South. In highly endemic areas, nearly the entire population has been infected. Present or past histoplasmosis in a person is indicated by a positive histoplasmin skin test, which uses an extract of the fungus to detect delayed-type hypersensitivity and is similar to tests for tuberculosis. Once infection has occurred by inhalation of mold spores, the fungus grows as a small budding yeast within mononuclear phagocytes of the body. Most infections produce no symptoms, whereas others bring on symptoms that resemble influenza. A few infected persons develop disseminated disease, with lesions in the spleen, liver, and lymph nodes and ulcerated lesions in the mouth. Once acquired, infection may persist for life even if no symptoms occur after primary exposure, and the persistent infection may reactivate and cause significant disease even years later if the human host becomes immunocompromised.

Blastomycosis and paracoccidioidomycosis are two other deep mycoses. Blastomycosis occurs mostly in the eastern United States; paracoccidioidomycosis is endemic to tropical South and Central America. Following pulmonary infection, widespread ulcerous lesions may develop. In blastomycosis, bone lesions and meningitis often develop, whereas in paracoccidioidomycosis, ulcerated lesions of the mouth and gastrointestinal tract are common. Both diseases are caused by species of pathogenic soil molds that have the specific ability to develop into budding yeasts once they are within the host.

Subcutaneous diseases. Other fungal infections develop when certain species of soil molds are inoculated deep into the subcutaneous tissue, such as by a deep thorn prick or other trauma. A specific type of lesion develops with each fungus as it grows within the tissue. Proper wound hygiene will prevent these infections.

Sporotrichosis, caused by *Sporothrix schenckii*, appears as a chain of ulcerated lesions on the arm or leg following inoculation of a soil mold. Within the infected tissue, the mold converts into a yeast. Unlike other fungi, *Sporothrix* is susceptible to treatment with oral potassium iodide.

Chromoblastomycosis and mycetoma also follow inoculation deep into the skin. Each of these clinical diseases can be caused by multiple fungal agents. Chromoblastomycosis appears as persisting ulcerous or wartlike skin lesions and is caused by any of several kinds of black molds. Mycetoma is a chronic tumorlike lesion that drains small grains (colonies) of a mold or actinomycotic agent and can be caused by a number of black or white molds or actinomycetes. Both diseases require early treatment with specific drugs; the responsible agent must be isolated to permit an appropriate choice of drugs.

Actinomycosis and nocardiosis are two bacterial diseases that are caused by agents having characteristics reminiscent of fungi, such as a branching filamentous structure and the production of similar disease patterns. They require specific laboratory identification so that appropriate antibacterial treatment may be given.

Ringworm. Ringworm, also known as dermatophytosis or tinea, is the most common of all fungal infections. Some species of pathogenic molds can grow in the stratum corneum, the dead outermost layer of the skin. Disease results from host hypersensitivity to the metabolic products of the infecting mold as well as from the actual fungal invasion. Tinea corporis, ringworm of the body, appears as a lesion on smooth skin and has a red, circular margin that contains vesicles. The lesion heals with central clearing as the margin advances. On thick stratum corneum, such as the interdigital spaces of the feet, the red, itching lesions, known as athlete's foot or tinea pedis, become more serious if secondary bacterial infection develops. The ringworm fungi may also invade the

hair shaft (tinea capitis) or the nail (onychomycosis). Many pharmaceutical agents are available to treat or arrest such infections, but control of transmission to others is important. *See* FUNGI; YEAST.

Carlyn Halde; Jon P. Woods

Bibliography. N. M. Ampel, Emerging disease issues and fungal pathogens associated with HIV infection, *Emerging Infect. Dis.*, 2:109–116, 1996; E. J. Anaissie (ed.), Focus on fungal infections 5, *Clin. Infect. Dis.*, 22(suppl. 2):S71–S184, 1996; G. S. Deepe, Jr., Prospects for the development of fungal vaccines, *Clin. Microbiol. Rev.*, 10:585–596, 1997; S. K. Fridkin and W. R. Jarvis, Epidemiology of nosocomial fungal infections, *Clin. Microbiol. Rev.*, 9:499–511, 1996; N. H. Georgopapadakou and J. S. Tkacz, The fungal cell wall as a drug target, *Trends Microbiol.*, 3:98–104, 1995; S. L. Newman, Macrophages in host defense against *Histoplasma capsulatum, Trends Microbiol.*, 7:67–71, 1999; D. J. Sheehan, C. A. Hitchcock, and C. M. Sibley, Current and emerging azole antifungal agents, *Clin. Microbiol. Rev.*, 12:40–79, 1999; J. W. Taylor et al., The evolutionary biology and population genetics underlying fungal strain typing, *Clin. Microbiol. Rev.*, 12:126–146, 1999.

Medical parasitology

The study of diseases of humans caused by parasitic agents. It is commonly limited to parasitic worms (helminths) and the protozoa (see **table**). Current usage places the various nonprotozoan microbes in distinct disciplines, such as virology, rickettsiology, and bacteriology. Medical entomology, a branch of medical parasitology, deals with insects that serve as intermediate hosts or vectors of parasites, as well as the noxious effects of the insects themselves. Ticks, mites, and other noninsect arthropods are often included, although they are more accurately combined with insects in the field of medical arthropodology.

Nematodes. The roundworms form an extremely large yet fairly homogeneous assemblage, most of which are free-living (nonparasitic), and are found in virtually every soil and aquatic habitat. Many are also plant parasites, and a large number have become adapted to living parasitically within nearly all vertebrates. Some parasitic nematodes, however, may cause disease in humans (zoonosis), and others cause disease limited to human hosts (anthroponosis). Among the latter, several are enormously abundant and widespread. *See* NEMATA.

Intestinal nematodes. The giant roundworm (*Ascaris lumbricoides*) parasitizes the small intestine, probably affecting over a billion people; and the whipworm (*Trichuris trichiura*) infects the human colon, probably affecting a half billion people throughout the tropics. Similarly, the hookworms of humans, *Necator americanus* in the Americas and the tropical regions of Africa and Asia, and *Ancylostoma duodenale* in temperate Asia, the Mediterranean, and Middle East, suck blood from the small intestine and cause major debilitation, especially among the undernourished. The human pinworm (*Enterobius vermicularis*) infects the large intestine of millions of urban dwellers. Its eggs, viable shortly after being laid on perianal skin, are readily transmitted from hand to mouth and rarely are passed in the stool, and so they are not destroyed in flush-toilet sanitary systems. In contrast, the other intestinal nematodes, which require a period of egg maturation outside the human host before they are infective, are associated with fecal contamination of soil or food crops and are primarily rural in distribution. The small roundworm of humans, *Strongyloides stercoralis*, has three methods of infections: (1) skin penetration by infective third-stage larvae (the method also used by hookworms); (2) penetration by infective larvae after a period of multiplication as adult nonparasitic worms in the soil, whose progeny eventually revert to penetrating larvae; or (3) internal penetration of the gut of immunosuppressed or heavily exposed individuals. This internal infection can be extremely hazardous and even fatal, or it may persist for many years.

Tissue-infecting nematodes. The nonintestinal nematodes are spread by complex life cycles that usually involve bloodsucking insects, such as mosquitos, deerflies, blackflies, or midges. One exception is the guinea worm (*Dracunculus medinensis*), a skin-infecting 2–3-ft (0.6–1-m) worm transmitted by aquatic microcrustaceans (water fleas or copepods) that are ingested in drinking water that has been contaminated by larvae that escape from the skin sores of infected humans. Such bizarre life cycles are typical of many helminths. Other nematodes of humans include (1) the filarial worms, which are transmitted by mosquitoes and may induce enormously enlarged fibrous masses in legs, arms, or genitalia (elephantiasis), and (2) *Onchocerca volvulus*, which is transmitted by blackflies (genus *Simulium*) and forms microscopic embryos (microfilariae) in the eyes causing high incidence of blindness in Africa and parts of central and northern South America.

A more familiar tissue-infecting nematode of temperate regions is *Trichinella spiralis*, the pork or trichina worm, which is the agent of trichinosis. The tiny spiraled larvae encyst in the muscle of pigs, rats, bears, and even Arctic seals and can carry the infection to humans and other carnivorous mammals who eat raw or undercooked infected meat.

Trematodes. Parasites of the class Trematoda vary greatly in size, form, location in the human host, and disease produced, but all go through an initial developmental period in specific kinds of fresh-water snails, where they multiply as highly modified larvae of different types. Ultimately, an infective larval stage (cercaria) escapes in large numbers from the snail and continues the life cycle. Each trematode species follows a highly specific pathway from snail to human host, usually by means of another host or transport mechanism.

Heterophyes heterophyes, the minute intestinal fluke, is transmitted to humans through various fresh-water or estuarine fish eaten raw or undercooked.

The oriental liver fluke (*Clonorchis sinensis*) also infects humans by way of fish, usually the carp, in which the cercariae encyst. Related fish-borne trematodes in the genus *Opisthorchis* also infect the human liver and can survive pickling or even drying of the fish. The human lung fluke (*Paragonimus*) infects the human lung after the encysted cercariae in raw crayfish or freshwater crabs are eaten. Other trematode cercariae swim from their host snails and encyst on aquatic vegetation. The sheep liver fluke [*Fasciola hepatica* (*F. gigantea* in Africa)] may also infect humans, and the giant intestinal fluke (*Fasciolopsis buski*) infects humans (and pigs) in China and Southeast Asia after they ingest infected water chestnuts or other aquatic vegetation. Another important group of trematodes is the blood flukes of the genus *Schistosoma*. *See* SCHISTOSOMIASIS; TREMATODA.

Cestodes. Tapeworms, the other great assemblage of parasitic flatworms, parasitize most vertebrates, with eight or more species found in humans. Their flat ribbonlike body form consists of a chain of hermaphroditic segments. Like the trematodes, their life cycles are complex, although not dependent on a snail host. The enormous beef tapeworm of humans, *Taenia saginata*, is transmitted by infected beef ("measly beef") from cattle that grazed where human feces containing egg-filled tapeworm segments contaminated the soil.

The pork tapeworm (*T. solium*) similarly reaches humans through pork infected after the pig injested human feces that contained the tapeworm segments. Humans can also be infected (usually by drinking contaminated water) from *T. solium* eggs passed in human feces as well as from meat infected with the larvae. The eggs can hatch in the human gut, penetrate the gut wall, and encyst in various tissues, forming cysticerci as in the pig. The resulting disease, cysticercosis, is characterized by scattered and potentially lethal cysticerci, which develop chiefly in the brain, eye, or muscle tissues.

Another tapeworm that infects humans as an intermediate host is the dog tapeworm (*Echinococcus granulosus*) from sheep. In the sheep (or human) the egg hatches in the gut, and larva penetrates the tissues and forms a massive spherical encysted body (hydatid cyst) in the viscera. The hydatid is filled with fluid and forms many thousands of heads (scoleces) of the future adult worms. These can only develop in the dog intestine after the animal eats a cyst in sheep viscera. The dogs's droppings then contain many eggs, which can infect sheep or humans, who may ingest eggs after petting or handling an infected dog. The result is human echinococcosis or hydatid disease.

One common tapeworm of humans seen more frequently in temperate than in tropical regions is *Diphyllobothrium latum*, the broad or fish tapeworm. It reaches monumental size, sometimes 30 ft (10 m) with segments up to 0.8 in. (2 cm) wide; and among some peoples, especially the Finns, a vitamin B_{12} deficiency may develop. If the eggs, which are passed in the feces of an infected person, are discharged into fresh water, they can develop, hatch, be ingested by a water flea (copepod), and form an early larval stage inside the water flea. The small fish that commonly feed on copepods may in turn become infected, with the worm larva progressing into another larval stage, the plerocercoid. The small fish can be eaten by a larger one and in turn a still larger one until the end of the fish food chain—usually a bass, pickerel, or pike—is reached. In each fish host, the worm migrates to the muscles, where it remains unencysted until the fish is eaten by another host. A human or other fish-eating mammal feeding on the fish in raw, undercooked, pickled, or smoked form can acquire the worm, which grows to its full size in the gut in about 3 months. With most tapeworm infections, though, drugs are available to kill and eliminate the worm. If the worm merely breaks, however, leaving the head and anterior segments attached, it can regenerate its original length in 2–3 months. *See* CESTODA.

Protozoa. Protozoa that can infect humans are found in the intestine, various tissues and organs, and the bloodstream. *See* PROTOZOA.

Intestinal protozoa. Of the many protozoa that can reside in the human gut, only the invasive strain of *Entamoeba histolytica* causes serious disease. This parasite, ingested in water contaminated with human feces containing viable cysts of *E. histolytica*, can cause the disease amebiasis, which in its most severe form is known as amebic dysentery. Occasionally, the infection will spread by way of the intestinal veins to the liver, where the amebas can cause hepatic amebiasis, or can spread to other organs as well. In spite of this capacity to invade tissues, most *E. histolytica* remain in the gut lumen as nonpathogenic infections. Another common waterborne intestinal protozoon is the flagellate *Giardia lamblia*, which causes giardiasis, a mild to occasionally serious or long-lasting diarrhea. Campers or skiers may acquire it from animal-fouled water. Human sources frequently initiate outbreaks within institutions, and play or preschool groups are especially prone to a rapidly spreading infection. *See* GIARDIASIS.

Tissue protozoa. Other flagellate parasites infect the human skin, bloodstream, brain, and viscera. The tsetse fly of Africa carries to humans the blood-infecting agents of trypanosomiasis, or African sleeping sickness, *Trypanosoma brucei gambiense* and *T. brucei rhodesiense*. Domesticated animals can suffer from a form of trypanosomiasis as well. The infection can be fatal if the parasites cross the blood–brain barrier, a rare event among infected wild animals, chiefly antelopes and wild pigs, but more common in humans or domestic animals. In Latin America, the flagellate *T. cruzi* is the agent of Chagas' disease, a major cause of debilitation and premature heart disease among those who are poorly housed. The infection is transmitted in the liquid feces of a conenose bug (genus *Triatoma*) and related insects. The infective material is thought to be scratched into the skin or rubbed in the eye, especially by sleeping children. *See* TRYPANOSOMATIDAE.

Clinical aspects of parasitic diseases in humans and biology of disease-producing parasites

Disease (species)	Vector or intermediate host	Infectious stage	Mechanism of infection	Pathology
INTESTINAL PROTOZOA				
Amebiasis (*Entamoeba histolytica*)	None	Four-nucleated cysts (gives eight trophozoites)	Fecal contamination of water or food or oral-anal contact	Most are asymptomatic; 10% show diarrhea, dysentery, ulceration, amebic dysentery, flask-shaped ulcers with lysis and destruction of white blood cells; 1% have liver abscess with anchovy paste pus; other organs can be infected by hematologic destruction; may cause immunosuppression
Balantidiasis (*Balantidium coli*)	None	Cyst	Reservoir (pig) contact with humans; fecal-oral path	Usually in colon lumen; harmless; may ingest host submucosa and cause ulceration in immunosuppressed patients, may cause dome-shaped necrotic ulcer with bowel gangrene and death
Dientamoeba fragilis infection (*Dientamoeba fragilis*)	Possibly within *Enterobius vermicularis* eggs	Trophozoite (no cyst form)	Probably by fecal contamination of food or water; direct contact	Rare pathology; dysentery; flatulence
Trichomoniasis (*Trichomonas vaginalis*)	None	Trophozoite (no cysts)	Sexual transmission	Males are usually asymptomatic; females have vaginal itching and burning, and thick, yellow, blood-tinged discharge
Giardiasis (*Giardia lamblia*)	None	Cyst	Water (organism is not killed by normal chlorination)	Flattened villi in lumen of small intestine; malabsorption, especially of fat; frothy, fatty feces; flatulence; diarrhea, cramping, weight loss
BLOOD AND TISSUE FLAGELLATES				
West African trypanosomiasis [sleeping sickness] (*Trypanosoma brucei* or *T. gambiense*)	*Glossina palpalis* (riverine tsetse fly); both sexes bite; humans only (not zoonosis)	Metacyclic trypomastigote	Tsetse bite lacerates skin; tsetse is "pool feeder"	Exhausts host's antibody production because stumpy form changes antigenic coat every week or so; leptomeningitis, cerebral edema, inflammation with perivascular cuffing; crosses blood–brain barrier to cause sleeping sickness; fatal in 2 years
East African trypanosomiasis [sleeping sickness] (*Trypanosoma brucei* or *T. rhodesiense*)	*Glossina morsitans* (savanna tsetse fly) both sexes bite; antelopes and wild pigs may serve as reservoir hosts (bushbuck is directly implicated); many animals are infected	Metacyclic trypomastigote	Tsetse bite lacerates skin; tsetse is "pool feeder"	Exhausts host's antibody production because stumpy forms change antigenic coat as endless sequence; crosses blood–brain barrier to cause sleeping sickness; fatal in 3–12 months.
Chagas' disease (*Trypanosoma cruzi*)	*Rhodnius* or *Triatoma* species of true bugs (conenose bugs); many reservoir hosts including rodents, armadillo, dogs, etc. (zoonosis)	Trypomastigote in bug feces	Bug feces or vomitus rubbed into bite or eye; blood transfusion; sexual or transplacental transmission	Acute: inflammatory reaction; chronic: tissue destruction of heart, with aneurysm at apex, megaesophagus and megacolon caused by destruction of mesenteric plexuses
Oriental sore, Old World cutaneous leishmaniasis (*Leishmania tropica*)	*Phlebotomus* sand fly (female); no reservoir but humans (anthroponosis)	Promastigote from bite of infected sand fly	Sand fly injection of promastigotes; amastigotes live within vacuole in macrophage, multiply, break out, invade, and are ingested by new macrophages; infected mass of macrophages induce skin lesion with specific characteristics varying with strain or species	Rolled-edge "dry lesion" ulcer with self cure in 1–2 years and immunity only to *L. tropica;* hyperimmune (sensitized) recidivans form in Middle East: incurable granuloma continuously producing satellite lesions and scars
Zoonotic Old World cutaneous leishmaniasis (*Leishmania major*)	*Phlebotomus* sand fly; rodents are usual reservoir hosts (zoonosis)	Promastigote from bite of infected sand fly	Sand fly injection of promastigotes; amastigotes live within vacuole in macrophage, multiply, break out, invade, and are ingested by new macrophages; infected mass of macrophages induce skin lesion with specific characteristics	Rolled-edge "wet lesion" ulcer; usually self cure, then immunity to both *Leishmania major* and *L. tropica* after 6–12 months
Kala azar or visceral leishmaniasis (*Leishmania donovani*)	*Phlebotomus* (Old World) or *Lutzomyia* (New World) sand flies	Promastigote	Injected by infected sand fly; promastigotes change to amastigotes in macrophage, concentrate in spleen, liver, bone marrow (not skin as in cutaneous leishmaniasis)	Splenomegaly, hepatomegaly, leukopenia, wasting, usually death if not treated

Stage that causes pathology	Treatment	Geographic area	Diagnosis
Trophozoite	Diiodohydroxyquin; diloxanide furoate; metronidazole; paromomycin; dehydroemetine	Worldwide, wherever fecal contamination occurs	Cysts or trophozoites in ova and parasite preparation made three times on alternate days; red blood cells in trophozoites suggest pathology
Ciliated trophozoite	Tetracycline	Worldwide, especially in pig-raising areas	Trophozoites or cysts detected in stool culture
Trophozoite	Iodoquinol; tetracycline; paromomycin	Cosmopolitan	Trophozoite in stool
Trophozoite	Metronidazole for both partners	Ubiquitous in sexually active populations	Microscopic examination of urine, discharge, or scraping with normal saline plus methylene blue or liquid culture; Pap test
Trophozoite	Metronidazole; quinacrine	Ubiquitous among campers, skiers, dogs, and wild animals	Fatty diarrhea; trophozoites or cysts in stool; string test on duodenal contents
Trypomastigote stage is found in human blood (extracellular); sleeping sickness sets in after blood–brain barrier is crossed	Hemolytic stage: suramin sodium; late central nervous system involvement: melarsoprol	West African vegetation around rivers	Trypanosomes in blood, cerebrospinal fluid, lymph node aspirate
Trypomastigote stage found in human blood (extracellular) may have direct toxic effect; sleeping sickness sets in after blood–brain barrier is crossed	Hemolytic stage: suramin sodium; late central nervous system involvement: melarsoprol	East Africa	Enlargement of posterior cervical lymph nodes (Winterbottom's sign); blood smear or concentrate
Amastigotes in heart or parasympathetic ganglia cells	Nifurtimox	Latin America (thatched roofs, mud cracks)	Romaña's sign at chagoma (primary lesion); C-shaped trypomastigote in blood; gastric biopsy; xenodiagnosis
Amastigotes (always in macrophage)	Stibogluconate sodium (recidivans form may respond only to heat compresses over several weeks)	Old World: Middle East, India, Africa, former Soviet Union	Needle biopsy at edge of ulcer or culture of scraping
Amastigotes (always in macrophage)	Stibogluconate sodium	Old World: Middle East, India, Africa, former Soviet Union	Needle biopsy at edge of ulcer or culture of scraping
Amastigotes (always in macrophage)	Stibogluconate sodium (post-kala azar dermal leishmaniasis 1–3 years after treatment common in India)	Mediterranean area, Sudan, Kenya, India, China, former Soviet Union, South America	Touch preparation on slide or punch biopsy of spleen or liver or bone-marrow aspirate cultured on blood agar; hyperglobulinemia and no skin reaction during active disease; serodiagnosis by complement fixation, indirect hemagglutination, immunofluorescent assay, or enzyme-linked immunosorbent assay tests

Clinical aspects of parasitic diseases in humans and biology of disease-producing parasites (*cont.*)

Disease (species)	Vector or intermediate host	Infectious stage	Mechanism of infection	Pathology
BLOOD AND TISSUE FLAGELLATES (*cont.*)				
New World cutaneous leishmaniasis (*Leishmania mexicana* complex)	*Lutzomyia* (New World) sand fly; many reservoir hosts, including arboreal mammals (zoonosis)	Promastigote	Sand fly injection of promastigotes; amastigotes live within vacuole in macrophage, multiply, break out, invade, and are ingested by new macrophages; infected mass of macrophages induces skin lesion with specific characteristics; chicle pickers in Yucatan can be infected by bite to ear	Rolled-edge ulcer; immunity to *Leishmania mexicana* only
Mucocutaneous leishmaniasis (*Leishmania braziliensis* complex)	*Lutzomyia* (New World) sand fly; reservoir hosts are monkeys and arboreal rodents (zoonosis)	Promastigote	Sand fly injection of promastigotes; amastigotes live within vacole in macrophage, multiply, break out, invade, and are ingested by new macrophages; infected mass of macrophages induces skin lesion with specific characteristics	Typical lesion, but 3–20 years after primary lesion, may destroy septum in up to 30% or destroy entire face, with death by starvation or asphyxiation; immunity to *Leishmania braziliensis* (if patient survives) or to *L. mexicana* as well (but not to *L. donovani*)
MALARIA AND ALLIED SPOROZOA				
Malignant tertian malaria (*Plasmodium falciparum*)	Female *Anopheles* mosquito (many species); habits vary, an important factor in control	Sporozoites	Enter liver from inoculation into bloodstream by infected mosquito; no relapse	Chills and fever; splenomegaly; severe anemia; disseminated intravascular coagulation; cerebral malaria; blackwater fever; hemoglobinurea; vast numbers of red blood cells are infected and made sticky; stem cells are also destroyed
Benign tertian malaria (*Plasmodium vivax*)	Female *Anopheles* mosquito (many species); habits vary, an important factor in control	Sporozoites	Enter liver from inoculation into bloodstream by infected mosquito; can relapse from hypnozoites in liver	Chills and fever; anemia; splenomegaly
Ovale malaria (*Plasmodium ovale*)	Female *Anopheles* mosquito	Sporozoites	Enter liver from inoculation into bloodstream by infected mosquito; can relapse from hypnozoites in liver	Chills and fever; anemia; splenomegaly
Quartan or malariae malaria (*Plasmodium malariae*)	Female *Anopheles* mosquito	Sporozoites	Enter liver from inoculation into bloodstream by infected mosquito; no relapse, but recrudescence from unknown source, possibly bloodstream	Chills and fever; anemia; splenomegaly; nephrotic syndrome
Toxoplasmosis (*Toxoplasma gondii*)	Cat/mouse is normal cycle; cat is only definitive host; humans and others are intermediate hosts; human is dead-end host	Oocysts from cat feces or cysts in meat; sporozoites relesed in human intestine from cat oocysts; merozoites released from tissue cysts; both penetrate gut, multiply as tachyzoites, encyst in central nervous system or other tissues as tissue cysts with bradyzoites	Ingestion of undercooked meat or fecal-oral exposure from cat litter	Usually strong T-cell-mediated host defense and usually no pathology; fetal central nervous system is invaded if mother is not immune before conceiving, leading to hydrocephalus and cerebral calcification, choroidoretinitis resulting in blindness; possible stillbirth; fatal to immunosuppressed patients; rare retinitis in normal adult, usually with flulike syndrome
Cryptosporiodiosis (*Cryptosporidium* spp.)	One-host cycle, but cattle and possibly others ae reservoir hosts	Sporozoites in oocysts (four per cyst); no sporocyst enclosing sporozoites	Ingestion of cysts, probably by fecal contamination	All stages found in brush border of mucosal epithelia, therefore intracellular; usually mild diarrhea or asymptomatic; in immunosuppressed or immuno-incompetent patients, causes intractable diarrhea
Pneumocystis pneumonia [now reclassified as a fungus] (*Pneumocystis carinii*)	No known vector; reservoir host is unknown	Cysts	Possibly air-borne	Mild or asymptomatic except in immuno-suppressed patients; causes intractable pneumonitis; plasma cell interstitial pneumonitis in premature infants

Stage that causes pathology	Treatment	Geographic area	Diagnosis
Amastigotes (always in macrophage)	Stibogluconate sodium	New World: Mexico, Latin America; disseminated leishmaniasis in Ethiopia and Venezuela induces specific anergy (distinctive syndrome)	Needle biopsy at edge of ulcer or culture of scraping
Amastigotes (always in macrophage)	Stibogluconate sodium; amphotericin B; treatment often unsatisfactory	Amazon basin	Needle biopsy at edge of ulcer or culture of scraping; serodiagnosis is also possible
Merozoite	Prophylaxis: chloroquine; in resistant areas, add daily proguanil or doxycycline; treatment: chloroquine; if resistant, fansidar or quinine (or quinidine); if comatose, intravenous quinine (or quinidine)	All tropics, but especially sub-Saharan Africa	Thick and thin blood smears with ring stage and gametocytes (latter is sausage-shaped); no other stages in peripheral blood (infected cells with trophozoites, adhere to endothelium)
Merozoite	Prophylaxis: chloroquine + primaquine for release (after return); treatment: chloroquine (no resistance observed yet)	Tropics except where *Plasmodium ovale* is found	Schüffner's dots, enlarged host red blood cells with enlarged round gametocytes
Merozoite	Prophylaxis: chloroquine + primaquine (after departure from endemic area); treatment: chloroquine (resistance reported)	Tropical Africa	Same as with *Plasmodium vivax,* but dots appear at earlier stage and red blood cells are often oval and fimbriated (jagged) at one end
Merozoite	Prophylaxis: chloroquine; treatment: chloroquine (no resistance observed yet)	Tropics	Trophozoites are often band-shaped across red blood cell; host cell is not enlarged and no Schüffner's dots are produced; gametocytes are found in nonenlarged red blood cells
Tachyzoites in macrophage or other host cells free in body cavity; bradyzoites fill tissue cysts	Trimethoprimsulfamethoxazole	Ubiquitous; about 40% among humans (mostly asymptomatic)	Serologic tests; xenodiagnosis using mouse inoculation
Merozoite	None satisfactory	Ubiquitous, especially in cattle-raising areas	Stool, seen after special stain or search for minute bodies; mucosal biopsy
Trophozoite	Trimethoprimsulfamethoxazole chemoprophylaxis as nasal-lung aerosol	Universal; perhaps 20–30% of humans	Cysts in sputum sample

Clinical aspects of parasitic diseases in humans and biology of disease-producing parasites *(cont.)*

Disease (species)	Vector or intermediate host	Infectious stage	Mechanism of infection	Pathology
INTESTINAL NEMATODES				
Enterobiasis (*Enterobius vernicularis,* pinworm)	None	Embryonated egg (within 4 h after laid in perianal folds or around anus)	Anal-oral route; eggs in bedding go into air, spread via diaper contact; adults live 2–3 weeks in terminal colon after eggs hatch in duodenum	Symptom-free pruritis; occasional allergic response; possible cause of appendicitis; may cause pelvic inflammatory disease, sleep loss, neurological symptoms
Trichuriasis (*Trichuris trichiura,* whipworm)	None; eggs in soil stick to vegetables or are found in contaminated water	Embryonated egg (after 4–6 weeks in soil); eggs can survive 1–2 years	Ingestion of egg on contaminated vegetables or in water; adults remain in colon (no systemic migration), live up to 9 months	Mild bleeding, diarrhea; with heavy infection, anal prolapse when circular muscles contract spontaneously
Ascariasis (*Ascaris lumbricoides,* giant roundworm)	None; eggs in soil stick to vegetables or are found in contaminated water	Embryonated egg (after 4–6 weeks in soil); eggs can survive 5–10 years in soil	Eggs on contaminated vegetables or in water ingested; hatch in duodenum; larvae penetrate mucosa, carried by circulation to heart and lungs, where they penetrate alveoli, molt, migrate up air ducts to mouth, are swallowed, mature, and mate in small intestine; can live 1 year	Pneumonitis after second infection; hypersensitivity reaction to larvae migrating through lungs; adult may occasionally perforate bowel or obstruct duct openings (bile, pancreatic ducts, inner ear)
Hookworm disease (*Necator americanis* or *Ancylostoma duodenale*)	None	Filariform larva develops in soil	Penetrate skin; carried in bloodstream to lungs and trachea; swallowed; adults in small intestine suck blood from villi; live 1–4 years	Anemia; mucosal scars; possible hypersensitivity with eosinophilia after repeated infections
Stronglyoidiasis (*Strongyloides stercoralis,* small human roundworm)	None; can be free-living in soil	Filariform larvae penetrate skin (from soil); can autoreinfect, especially in immunosuppressed patients	Direct contact of skin with contaminated soil; infection can last a lifetime	Diarrhea, especially with heavy infection; malabsorption; eosinophilia and pneumonitis during larval migration; secondary autoreinfection a major cause of severe pathology and long infection (20–30 years)
Trichinosis (*Trichinella spiralis*)	None	Larvae are ingested with pork, bear, or other animal flesh	Eating undercooked pork; adults mature in small intestine, newborn larvae migrate to muscle fibers, rarely to brain and heart; may be from bear meat or other sources	Secondary infection due to enteric flora; petechial hemorrhages, eosinophilia; periorbital edema; after 7 days, larvae in muscle host cells turn into "nurse cells"
TISSUE-INFECTING NEMATODES				
Filariasis (*Wuchereria bancrofti, Brugia malayi, B. timori*)	Female mosquitoes of *Aedes, Anopheles, Mansonia,* and *Culex* species	Mosquito bites; larvae deposited on skin enter bite wound, migrate to lymph node, mature, mate, produce microfilariae after 1 year	Adults mature in lymphatics; microfilariae stay in lung capillaries during day and peripheral blood at night, where they are available to mosquitoes (nocturnal periodicity)	Elephantiasis in about 10% of cases after years of hypersensitivity; reaction to adults include pain, enlargement, swelling, lymphangina, lymphadenitis, lymphedema, massive fibrosis of genitalia and limbs (elephantiasis)
Onchocerciasis [river blindness] (*Onchocerca volvulus*)	Blackfly (*Simulium damnosum* and other species)	Blackfly takes a blood meal; larvae crawl into bite or hair follicle; mature in 1 year	Adults develop and live in subcutaneous tissues, form dermal nodules; (New World: upper body; Old World: lower body); microfilariae in interstitial fluid are the cause of major pathology in skin and eyes	"Elephant" or "leopard" skin; microfilariae destroy elastic fibers; partial depigmentation; iritis leads to secondary glaucoma, snowflake opacities, corneal and retinal damage, blindness, all caused by reaction to microfilariae in vitreous (after 3–7 years)
Dracunculiasis (*Dracunculus medinensis*)	*Cyclops* species copepods or water fleas	Larva	Drinking water from step-in well, infected by *Cyclops;* larvae are liberated in gut; in 1 year, female travels to deep subcutaneous tissues in leg, lives head-down; induces blister to form around head, which breaks open with cold water; female ejects motile larvae into water	Blisters, ulcers; allergic reaction
CESTODES				
Taeniasis saginata (*Taenia saginata,* beef tapeworm)	Cattle	Cysticerci in beef	Ingestion of undercooked beef	Little pathology; epigastric fullness; nausea; loss of appetite

Stage that causes pathology	Treatment	Geographic area	Diagnosis
Adult: in descending colon, pass through anus to lay eggs in perianal folds	Mebendazole	Ubiquitous in urban crowded conditions; higher incidence in children	Cellophane tape across anus, at night or early morning
Adult	Mebendazole	Worldwide, including tropical areas of southern United States	Eggs in stool; Charcot-Leyden crystals
Larva, adult	Mebendazole	Worldwide	Eggs in stool
Adult worms with cutting plates (*Necator*); adults with daggerlike teeth (*Ancylostoma*)	Mebendazole	*N. americanis:* New World, all tropics; *A duodenale:* Old World, north temperate zones	Unembryonated eggs in stool (usually four-cell stage)
Adult, larva	Thiabendazole	Tropics and subtropics	Rhabditiform larvae in stool
Larval migration and penetration of red muscle fibers; subsequent sensitization and possible toxic effects; strong immunity is induced	Thiabendazole + steroids for severe cases	Wherever pork is consumed	High pork consumption, eosinophilia after 2 weeks; immunologic tests
Adult	Diethylcarbamazine, with care for hypersensitization reaction	Africa, Southeast Asia, New Guinea, South Pacific	Microfilariae in peripheral blood at night (nocturnal periodicity)
Microfilariae in interstitial fluid (not in bloodstream)	Ivermectin	Fast-running streams in Central America and tropical Africa (especially West Africa)	Skin snip shows microfilariae
Adult female; broken worm may cause anaphylaxis	Niridazole; extraction by winding on a stick	Equatorial Africa (rain pools in West Africa); Middle East, India (step-in wells); South America	Head of worm in skin blister; often ulceration if improperly treated
Adult	Niclosamide; praziquantel	Africa, Mexico, Argentina, Europe, United States (beef-eating areas)	Proglottid in stool with more than 12 branches per side of uterus
Adult	Niclosamide; praziquantel	Worldwide in pork-eating populations, especially Latin America	Proglottid in stool with 5–10 branches per side of uterus

Clinical aspects of parasitic diseases in humans and biology of disease-producing parasites *(cont.)*

Disease (species)	Vector or intermediate host	Infectious stage	Mechanism of infection	Pathology
CESTODES (*cont.*)				
Cysticercosis (*Taenia solium*)	Human is accidental abnormal intermediate host (normally is final host)	Eggs develop cysticerci in tissues, either cutaneous or subcutaneous, occasionally in eye	Fecal contamination of water or possible internal reinfection via hatching of eggs in gut	Seizure from cysticerci in brain; eye/muscle space-occupying lesions; calcification after inflammatory reaction
Echinococcosis; unilocular hydatid disease (*Echinococcus granulosus*)	Sheep and dog are intermediate and final host pair; humans can become intermediate host; only dog is final host	Egg	Ingestion of eggs from dog feces on fur or by contaminated water	Eggs hatch, release larvae, which penetrate gut; cysts grow in liver, eye, brain, other organs covered with laminated membrane; form unilocular hydatid cyst
Diphyllobothriasis; broad or fish tapeworm disease (*Diphyllobothrium latum*)	Copepod (water flea) is eaten by fresh-water fish (secondary host); passes through series of fish, from minnow to bass, pike, etc.	Pleurocercoid larva in fish muscle	Ingestion of raw or undercooked fish	Adult lives in small intestine, grows to about 30 ft (10 m), absorbs vitamin B_{12}; rarely may cause pernicious anemia
TREMATODES				
Fasciolopsiasis (*Fasciolopsis buski*)	Fresh-water snail (*Segmentina hippeutis*)	Metacercariae on husks of water chestnut, talpa, and other fresh-water edible plants	Removing water chestnut husk with teeth	Feeds on columnar epithelial cells of intestine; can obstruct ampulla of Vater, common bile duct, small intestine
Fascioliasis (*Fasciola hepatica; F. gigantica*)	Fresh-water snail (*Lymnaea* sp.)	Metacercariae on watercress or other aquatic plants; juveniles penetrate through gut wall and liver into bile ducts	Eating vegetables with metacercariae; possibly by drinking water draining from pasture	Adults live in liver, feed on hepatocytes and blood; cause colic, abscesses; in Middle East, causes halzoun (adult on pharyngeal wall from eating raw liver; worms crawl up from stomach to pharynx)
Clonorchiasis (*Clonorchis sinensis*)	Fresh-water snail (species of operculate snails)	Metacercariae in fish muscle; juveniles migrate via common bile duct to liver	Eating undercooked fish with metacercariae	Adults live in bile ducts, cause inflammation; mild disease, except in heavy infection
Heterophyiasis (*Heterophyes heterophyes*)	Fresh- or brackish-water snail (*Pironella* in Egypt, *Cerethidea* in Japan)	Metacercariae in fish muscle; juveniles migrate to intestine	Eating undercooked infected fish	Diarrhea; rarely eggs get into circulation, cause granulomas to form in heart and brain
Paragonimiasis (*Paragonimus* species, especially *P. westermani,* human lung fluke)	Fresh-water snail (*Semisulcospira brotia*)	Metacercariae on crab or crayfish	Eating undercooked crab or crayfish	In lung, fibrotic lesions, hemoptysis, dyspnea, pleural pain; secondary abscesses, occasionally in brain
Schistosomiasis (*Schistosoma haematobium*)	Fresh-water snail (*Bulinus* species)	Cercariae	Cercariae enter skin from water, migrate to liver; adults live in venous plexus around bladder for 5–8 years	Eggs caught in tissues produce T-cell-dependent granulomas; bladder calcification, hydroureter, and hydronephrosis often associated with bladder cancer
Schistosomiasis (*Schistosoma mansoni*)	Fresh-water snail (*Biomphalaria* species)	Cercariae	Cercariae enter skin from water, migrate to liver; adults live in inferior mesenteric venules for 5–8 years, some over 30 years	Eggs caught in tissues produce T-cell-dependent granulomas, and polyps; liver shows pipe-stem fibrosis, portal hypertension, esophageal varices, cor pulmonale
Schistosomiasis (*Schistosoma japonicum*) and similar species in Southeast Asia)	Fresh-water snail (*Oncomelania* species)	Cercariae	Cercariae enter skin from water, migrate to liver; adults live in inferior and superior mesenteric venules for 5–8 years	Eggs caught in tissues cause thickening of gut and polyps, T-cell-dependent granulomas; eggs (small and laid in clusters) can travel to spinal cord and brain; cor pulmonale; hepatosplenomegaly

Stage that causes pathology	Treatment	Geographic area	Diagnosis
Cysticercus	Surgical removal; anticonvulsants	Pig-raising areas, especially Mexico, South America	Computerized tomography scan, magnetic resonance imaging, x-rays, serology
Hydatid cyst; anaphylaxis and metastatic spread of cysts if ruptured	Surgical removal; mebendazole may shrink cyst	California, Utah, Middle East, Southern Europe sheep-herding areas, especially Basque region	Scans, ultrasound, history, space-occupying lesions, serology
Adult	Niclosamide	Scandinavia, Switzerland, Hungary, former Soviet Union, Great Lakes region of United States	Eggs in stool (proglottids break up in colon)
Adult	Praziquantel	Asia, India	Eggs in stool
Adult	Praziquantel	Sheep-herding areas in Latin America, United States, South America, Africa, Europe, China	Eggs in stool; computerized tomography scan
Adult	Praziquantel	Japan, Korea, Vietnam, China	Eggs in stool (*Opisthorchis felineus* and *O. viverrini* similar to *Clonorchis*)
Adult; heart disease and neurological disturbance (Philippines), mild intestinal disturbance	Praziquantel	Asia, Middle East, Egypt	Eggs in stool
Adult	Praziquantel	Asia, India, Pacific region, South America	Eggs in stool and sputum
Eggs (terminal spine)	Praziquantel; trichlorfon	Africa, Middle East	Eggs in urine
Eggs (lateral spine)	Praziquantel; oxamniquine	Africa, South America, Caribbean region	Eggs in stool
Eggs (small lateral spine)	Praziquantel	China, Philippines, Southeast Asia	Eggs in stool

Another group of parasitic flagellates includes the macrophage-infecting members of the genus *Leishmania*, which are transmitted by blood-sucking midges or sand flies (genus *Phlebotomus* in the Old World, *Lutzomyia* in the New World). Cutaneous leishmaniasis is characterized by masses of infected macrophages in the skin, which induce long-lasting dermal lesions of varying form and severity. A related condition, mucocutaneous leishmaniasis, is found in the Amazon basin and begins as a skin ulceration; ultimately it may attack and destroy the nasal mucosa, cartilage, and soft facial and pharyngeal tissues. Another leishmaniasis disease complex, visceral leishmaniasis or kala azar, occurs in the spleen, liver, bone marrow, and lymph nodes. The broad spectrum of host–parasite interactions is well exemplified by these leishmaniases. The various manifestations of the disease are the result of the particular species of agent and vector, the immunological status of the host, the presence or absence of reservoir hosts, and the pattern of exposure.

Ciliate and sporozoan parasites. Two remaining major groups of protozoa are the ciliates and the sporozoans. The former group is largely free-living, with only a single species, *Balantidium coli*, parasitic in humans (and pigs). This large protozoon, covered with rows of cilia, is found in the large intestine, where it can cause balantidiasis, an ulcerative disease.

The sporozoans on the other hand, are all parasitic and include many parasites of humans. The most important are the agents of malaria. Other disease agents are included in the genera *Isospora, Sarcocystis, Cryptosporidium*, and *Toxoplasma. Pneumocystis*, a major cause of death among persons with acquired immune deficiency syndrome (AIDS), was formerly considered a protozoon of uncertain relationship, but now it is thought to be a member of the Fungi. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); MALARIA; SPOROZOA.

Toxoplasma gondii, the agent of toxoplasmosis, is of special importance because it infects as many as 20% of the world's population and because it can penetrate the placenta and infect the fetus if the mother has not been previously infected and has no antibodies. As with most medically important parasites, the great majority of *Toxoplasma* infections remain undetected and nonpathogenic. The parasite primarily affects individuals lacking immune competence—the very young, the very old, and the immunosuppressed. *See* MEDICAL BACTERIOLOGY; MEDICAL MYCOLOGY; PARASITOLOGY; ZOONOSES.

Donald Heyneman

Bibliography. T. C. Cheng, *General Parasitology*, 2d ed., 1986; R. Goldsmith and D. Heyneman (eds.), *Tropical Medicine and Parasitology*, 1989; A. A. Mahmoud, *Tropical and Geographical Medicine*, 2d ed., 1990; E. K. Markell and M. Voge, *Medical Parasitology*, 8th ed., 1999; G. Piekarski, *Medical Parasitology*, 1989; G. D. Schmidt and L. S. Roberts, *Foundations of Parasitology*, 5th ed., 1995; G. T. Strickland, *Hunter's Tropical Medicine and Emerging Infectious Diseases*, 8th ed., 2000.

Medical ultrasonic tomography

A mapping or imaging technique, used to obtain clinically useful information about the structure and functioning of tissues and organs, in which acoustic pulses are emitted from an acoustoelectric transducer, and echoes are received from acoustic impedance discontinuities along the assumed line-of-sight axial propagation path. A number of different modes of operation have emerged, each having areas of usefulness.

A (amplitude) mode. This technique uses acoustic pulse emissions and echo reception along a single line-of-sight axial propagation path, and thus provides a one-dimensional mapping. The information is most often displayed on a cathode-ray oscilloscope in which the horizontal axis of the display is a linear time base triggered at the time of the transmitted pulse, and the received echoes are manifested as vertical deflections, with vertical displacement a measure of the amplitude or strength of the returning echo. This mode of operation cannot provide identification of structural features. It is, however, a most accurate method of measuring time delays and, therefore, distances between echo-producing structures or distances of structures from transducers, provided the speed of sound propagation in the medium is known. *See* OSCILLOSCOPE.

Since the A-mode display provides only measurement of distance along the line of sight, that is, positional information, it is the task of the user to identify correctly the echo-producing structures and the position of the line of sight in order to ensure that the correct distances are being measured. Because of the difficulties in identifying individual echoes in a complex echo pattern, the technique is limited to objects of relatively simple, and repeatedly obtainable, geometry.

The method is employed in the detection of shifts of midline intracerebral structures, which indicate the presence of space-occupying lesions. In ophthalmology, it is used in determinations of the thicknesses of the cornea and lens and the axial length of the eye. In obstetrics, the growth and development of the fetus can be monitored by measurement of its biparietal diameter in the uterus.

M (motion) mode. The M mode of operation is used to display the movement of time-varying, echo-producing structures by intensity-modulating the trace as it is swept slowly across the oscilloscope screen in a direction at right angles to the fast time-base sweep. This mode of operation is used extensively in diagnosing disorders of the heart. It can be used to determine accurately details of valvular motion and quantitative measurements of heart chamber size and wall thickness. Stiffening of the leaflets of the mitral valve (stenosis), which prevents proper closure, is readily diagnosed by the reduced velocity of the valve leaflets during opening and closure. Another important application is in the identification of pericardial effusion, an accumulation of fluid surrounding the heart. These applications are characterized by identification of the echo-producing

structure by the demonstration of an easily recognizable and unique pattern of movement of the structures of interest. *See* ECHOCARDIOGRAPHY; HEART DISORDERS.

B (brightness) mode. For a two-dimensional picture to be obtained, the line-of-sight propagation path must be scanned and the position and direction of the path monitored and used to form a two-dimensional picture. Typically, the B-mode display is formed by moving the transducer so that the line-of-sight path remains in a single plane. The time-base trace of the cathode-ray oscilloscope screen is moved to correspond, in position and direction, to the ultrasonic line-of-sight propagation path, and echoes are displayed as intensity modulations of the trace. The echogram which results is therefore a two-dimensional representation of distribution of acoustic impedance discontinuities in the interrogated object, with bright lines representing strongly reflecting boundaries and textured gray areas representing the internal structure of the organs. Different kinds of scanning techniques are used. Simple scanning might consist of only a single motion such as sector, arch, or linear translation.

B-mode application is used in obstetrics to confirm intrauteral pregnancy, assess fetal anatomy, determine fetal position and multiple gestation, and monitor needle biopsy. In gynecology, pelvic masses can be identified by using a full urinary bladder as an acoustic window.

B-mode ultrasound is also used for abdominal diagnosis (see **illus.**), for example, in evaluating palpable masses for size, composition, and organ of origin. The kidneys are readily outlined; size and shape can be determined, the presence of structural change, such as tumor or polycystic disease, can be seen, and deviations from the normal echographic picture can lead to a diagnosis of such functional abnormalities as urethral obstruction or intra- or extrahepatic hemorrhage. This eliminates the need for invasive procedures such as angiography.

The relatively simple and regular structural features of the eye allow for easy ultrasonic examination. For example, the extent of the detachment of a retina can be ascertained. B-mode ultrasonography readily differentiates simple cysts from solid masses in the breast and is particularly valuable for diagnosing the dense fibrous breast tissue of younger women, for which examination by x-ray mammography is poorly suited. Examination of the thyroid can reveal enlargement of the thyroid tissue itself or the presence of solid or liquid regions intruding into the organ. Real-time B-scanning, especially useful for cardiac imaging, is made possible by electronic steering. *See* MAMMOGRAPHY.

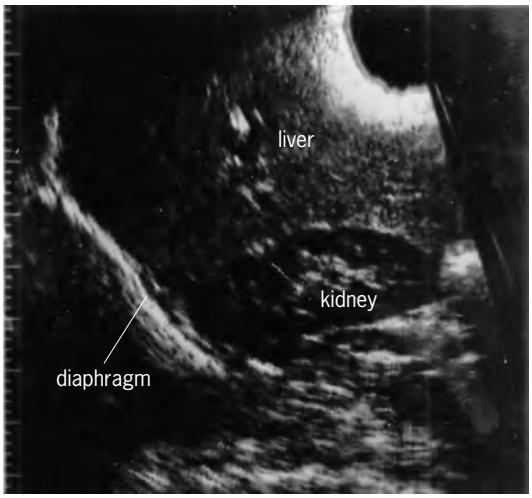

Ultrasonic B-scan of the abdomen.

C (constant-range) mode. Constant-range operation provides a two-dimensional image display at constant time delay, and presumed constant distance, from the ultrasonic transducer. The scanning is arranged so that the point at constant depth along the propagation path (beam axis) traverses a plane. This method provides a convenient display of atrial septal defects by using an intracardiac probe that is mounted at the tip of a catheter. *See* MEDICAL IMAGING; ULTRASONICS.

Floyd Dunn

Bibliography. D. R. Christensen, *Ultrasonic Bioinstrumentation*, 1988; F. J. Fry, *Ultrasound: Its Applications in Medicine and Biology,* 1978; C. R. Hill, *Physical Principles of Medical Ultrasonics*, 1986; P. N. T. Wells, *Biomedical Ultrasonics*, 1977.

Medical waste

Any solid waste that is generated in the diagnosis, treatment, or immunization of human beings or animals, in research pertaining thereto, or in the production or testing of biologicals. Since the development of disposable medical products in the early 1960s, the issue of medical waste has confronted hospitals and regulators. Previously, reusable products included items such as linen, syringes, and bandages; they were sterilized or disinfected prior to reuse, and the principal waste product was limited to human pathological tissue.

Most hazardous substances are described by their relevant properties, such as corrosive, poison, or flammable. Medical waste was originally defined in terms of its infectious properties, and thus it was called infectious waste. However, given the difficulty of identifying pathogenic organisms in waste that might cause disease, it has become standard practice to define medical waste by types or categories. While definitions differ somewhat under different regulations, in the United States the Centers for Disease Control (CDC) cite four categories of infective wastes that should require special handling and treatment: laboratory cultures and stocks, pathology wastes, blood, and items that possess sharp points such as needles and syringes (sharps). These categories, of necessity, require that the generator of these wastes exercise judgment in identifying the material to be included. For example, a small cotton ball or bandage with minimal amounts of dried blood will probably pose no problems or risk, but a dressing saturated with blood causes greater concern.

The waste category that has generated a great deal of interest is sharps. Needles and syringes, in particular, pose risks, since the instruments can penetrate into the body, increasing the potential for disease transmission. Improper disposal of these items in the past has been the catalyst for increased regulation and tighter management control.

Regulations. At the federal level, the Medical Waste Tracking Act (MWTA) established a 2-year demonstration project to be undertaken by the Environment Protection Agency, but this project terminated in 1991. Usually in the United States, requirements for identification, handling, packaging, treatment, storage, and disposal of medical waste are mandated by states and municipalities. This setup allows the various states to address issues on a more specific basis. For example, states with large numbers of rural health facilities may have needs and requirements that are different from those of states having higher and more concentrated population bases with large, centralized health facilities.

Typically, regulations define medical waste by expanding on the list developed by the Centers for Disease Control. Isolation waste from patients with highly communicable diseases is often included in these regulations. For the most part, most standard hospitals do not deal with this type of waste. Many tropical diseases and highly virulent pathogens fall into this category, and such patients are usually cared for at specialized hospitals and treatment centers. Such highly communicable diseases, identified as Biosafety level 4, are caused by those agents listed in Classification 4 in the official classification of etiologic agents of the Centers for Disease Control.

Management. Usually, the regulations mandate that, once identified, medical waste be separated from other waste as close to the source of generation as possible. It is identified by containment in a single, or sometimes double, red bag labeled "Biohazardous Waste" and the international biohazard symbol (see **illus.**). Such waste in stored in an area that can be secured so as to deny access by unauthorized personnel. This area is protected from the elements as well in order to prevent the spread of potentially disease-causing material into the environment.

After separation, the medical waste must be rendered noninfectious. Various treatment options are available, performed either at the source of generation or by a commercial entity off-site. If the waste is to be transported to an off-site location, a transportation permit must be issued and the vehicle registered with the appropriate regulatory agency. Vehicles used to transport medical waste must be fully enclosed and leakproof or leak-resistant in case the waste container ruptures. Bags containing medical waste should be placed in a rigid container for transport, such as a carton, pail, or drum. These containers can be either reusable or disposable; however, before reuse they must be thoroughly disinfected with either chemicals or hot water. The containers and the vehicle must be labeled with the words "Biohazardous Waste" and the international biohazard symbol. Some regulations require that the container as well as the bags also include the identification of the generator of the waste in case tracking the material to a specific source is necessary.

The uniform international biohazard symbol.

Tracking is also done by using a manifest. Manifest requirements vary somewhat, but a manifest document accompanies the waste in transit from generation to final disposal and identifies the custodian of the waste during transport and treatment. When the generator of the waste transfers custody to a transporter, portions of the manifest document that the transfer has been completed; each party retains a copy. When the waste is transferred from transporter to treatment site, additional portions of the manifest are completed to document that transfer. Upon treatment, disposal-site personnel send a completed copy of the manifest back to the generator, documenting that the amount, type, and source of waste generated was treated in accordance with the generator's instructions and the appropriate regulations. Some regulations mandate that if the generator does not receive a copy of the fully completed manifest within 30 days of shipment, the generator is required to notify the enforcement authorities who regulate medical waste in that particular jurisdiction. This procedure is known as exception reporting.

Treatment. Treatment of medical waste constitutes a method for rendering it noninfectious prior to disposal in a landfill or other solid-waste site. Additionally, some states require that medical waste be rendered "nonrecognizable" or destroyed. This requirement may be for esthetic reasons, as in the case of human pathological waste, or for safety reasons, as in the case of sharps. The treatment technologies currently used for medical waste include incineration, sterilization, chemical disinfection, and microwave, as well as others under development.

Incineration. Incineration has long been the traditional method of disposing of medical waste, dating back to when most of the waste was pathological in nature. Typically, these incinerators were designed to burn pathology wastes, which had a very high moisture content and did not create a great deal of heat in the combustion process. Mortuaries or crematories were often used as alternatives to incinerators at hospitals.

Modern incinerators are highly sophisticated, with air-pollution-control devices used to reduce or eliminate the gases created in the combustion process.

These gases result from the high content of plastics in the waste stream, including syringe barrels, suction containers, disposable utensils, and bags containing the waste. Unfortunately, a common perception is that incinerators increase pollution, pose environmental risks, and are not suitable for location at hospitals or in urban areas. These factors have severely reduced public acceptance of incinerators, and as a result other technologies have been developed. *See* AIR POLLUTION.

Sterilization. Steam sterilization has long been used in laboratory settings to sterilize cultures or to prepare samples for analysis or use. Additionally, gas or steam has been used to sterilize surgical instruments and associated items.

In the mid-1970s, the use of steam sterilizers or autoclaves for treating medical waste was begun, largely out of concern over incinerator emissions. Autoclaves, pressure vessels that use a boiler or other source for steam, are used to treat medical waste by hospitals and other generators, as well as by commercial waste-disposal facilities.

Chemical disinfection. Disinfection of medical waste by chemicals is an extension of the long-term hospital practice of applying chemicals to control microorganisms on hard surfaces, such as floors, walls, and tables. Equipment has been developed to use standard chemical agents in a sealed enclosure to treat waste. Often this equipment contains shredders or granulators to render the waste nonrecognizable. Chemical disinfection is growing as a treatment technology for use by generators and commercial entities.

Microwave treatment. Microwave technology combines shredding of the waste material along with exposure to microwaves, which create a moist heat environment for treatment. Waste is sprayed with water upon entering the processor, shredded, and then exposed to the microwaves, which heat the material sufficiently to disinfect it.

Other technologies. Other technologies are being developed that are designed to be environmentally acceptable, including macrowave, plasma-torch, and laser.

Macrowaves, low-frequency electrical waves longer than microwaves, are used to create an internal heat source to disinfect waste. This technology does not require water to create a moist heat environment. The longer wavelengths reduce the requirements for shredding the material, because they penetrate the waste farther than microwaves, thus exposing more of the waste volume.

In plasma-torch technology, gases at elevated temperatures are used to treat the waste. This environment creates a source of heat that can be recovered or generated into electricity.

Laser treatment, a somewhat experimental technology, uses laser-generated heat on the waste. Potential applications include use at the waste-generating site. *See* HAZARDOUS WASTE.

Robert A. Spurgin

Bibliography. E. Kunes, The trashing of America, *Omni Mag.*, pp. 40–44, 92–94, 96, February 1988; R. K. Miller and M. E. Rupnow, *Medical Waste Management*, 1991; U.S. Environmental Protection Agency Staff, *Medical Waste Management and Disposal*, 1990.

Medicine

The term medicine has two general meanings. The first indicates any material, notably in the drug category, which is given to prevent, alleviate, beneficially alter, or stop a disease process. In the second and broader sense, medicine denotes the field of science devoted to healing. Many subdivisions exist and more ramifications appear almost daily. Included in the area of medicine are many clinical specialties, including surgery, pediatrics, psychiatry, obstetrics, and others. Internal medicine is the specialization which deals with internal diseases of a nonsurgical nature. *See* SURGERY.

Related to the clinical specialties, particularly in regard to medical education and research, are the basic medical sciences. These include, among others, anatomy, physiology, pharmacology, biochemistry, and microbiology. Midway between the basic and the clinical sciences lies pathology, the study of the structural and functional alterations caused by diseases or abnormal states. *See* ANATOMY, REGIONAL; BIOCHEMISTRY; MICROBIOLOGY; PATHOLOGY.

An important area in all specialties is preventive medicine and public health. This form of medicine supplies a necessary link with the community, state, or large geographic region in matters of prevention, mass treatment, and statistical appraisals of health matters. It is also concerned with socioeconomic factors related to physical and mental well-being. *See* PUBLIC HEALTH.

The mental health of individuals and nations has received much attention since 1930. Prior to that time mental illness was largely a matter of disgrace, secrecy, and custodial care in primitive, prisonlike hospitals. With the recognition of the importance of the dynamics of psychiatry, the advent of better therapeutic measures, particularly specific drugs which promote mental well-being, and the realization that preventive measures are of vital concern to everyone, the field of mental health at the present time represents one of the largest areas in medicine.

Much attention has been given to aviation and space medicine, both of which consider the problems peculiar to their fields in the light of clinical and experimental information, as well as that supplied by other sciences and disciplines. *See* AEROSPACE MEDICINE.

Socialized medicine is that form which exists under the direct control and financing of the state. The National Health Service of Great Britain is the best-known example, and other systems exist.

Other subdivisions of medicine, with names that are largely self-explanatory, include veterinary, legal, tropical, and military medicine. *See* FORENSIC MEDICINE.

Probably the most familiar aspects of medicine to the nonspecialist are the two clinical aspects,

diagnosis and therapeutics. Diagnosis includes taking a patient's history, a physical examination, and possibly laboratory and special examinations. Therapeutics is the administration of some form of preventive or remedial treatment to the patient.

Although medicine is based parimarily upon scientific information and method, an important feature is the relationship between the physician and the patient. It is at this point that the necessary scientific background of medicine gives way to the art of healing. Edward G. Stuart; N. Karle Mottet

Mediterranean Sea

The Mediterranean Sea lies between Europe, Asia Minor, and Africa. It is completely landlocked except for the Strait of Gibraltar, the Bosporus, and the Suez Canal. The Strait of Gibraltar is 13 mi (21 km) wide at the boundary between the Atlantic and the Mediterranean (Great Europa Point, Gibraltar, to Punta Almina, Ceuta), and the Bosporus is as narrow as 0.5 mi (0.8 km), widening to 2 mi (3 km) at the Black Sea boundary (Rumeli Burnu, European Turkey, to Anadolu Burnu, Asiatic Turkey). The Suez Canal, a human-made channel connecting the Mediterranean with the Red Sea, is 87.6 mi (141 km) long, about 270 ft (90 m) wide, and 35 ft (11 m) deep. In latitude the Mediterranean extends from 30°15′N in the Gulf of Sidra to 45°47′N in the Gulf of Trieste; in longitude, from 5°21′W at Great Europa Point to 36°12′E in the Gulf of Alexandretta.

The Mediterranean is conveniently divided into an eastern basin and a western basin, which are joined by the Strait of Sicily and the Strait of Messina. The Strait of Sicily, between Cap Bon, Tunisia, and Sicily, is 77 mi (124 km) wide, but the Strait of Messina, between Punta Sottile, Sicily, and the Italian mainland, is only 1.7 mi (2.7 km) across.

The major subdivisions of the western basin are the Alboran Sea, which is the portion west of Cabo de Gata, Spain, and is named for the Isla de Alboran; the Balearic or Iberian Sea, between Spain and the Islas Baleares; the Ligurian Sea, the gulf north of Corsica; and the Tyrrhenian Sea, lying between Corsica, Sardinia, Sicily, and the Italian mainland. The eastern basin includes the Adriatic Sea, between Italy down to the heel of the boot (Capo Santa Maria de Leuca) and the Balkan Peninsula north of Corfu; the Ionian Sea, which lies south of the Adriatic as far as the southern tips of Sicily and the Greek mainland; the Aegean Sea, the waters north of Kithera, Andikithera, Crete, Karpathos, and Rhodes, as far as the Dardanelles; and the Sea of Marmara, which extends from the Dardanelles to the Bosporus.

The Mediterranean Sea has numerous islands, particularly in the Ionian Sea (which has the alternate designation of The Archipelago) and the eastern Tyrrhenian and Adriatic. The largest islands are Sicily, 9930 mi^2 (25,710 km^2); Sardinia, 9284 mi^2 (24,040 km^2); Cyprus, 3570 mi^2 (9250 km^2); Corsica, 3370 mi^2 (8720 km^2); and Crete, 3240 mi^2 (8380 km^2).

The total water area of the Mediterranean is 965,000 mi^2 (2,501,000 km^2), and its average depth is 5038 ft (1536 m). The greatest depth in the western basin is 12,200 ft (3719 m), in the Tyrrhenian Sea. The eastern basin is deeper, with a greatest depth of 18,130 ft (5530 m) in the Ionian Sea about 34 mi (55 km) off the Greek mainland. The Adriatic Sea mainly overlies the Continental Shelf, but there is a basin in the southern part with a greatest depth of 4313 ft (1315 m). The sill depth in both the Strait of Gibraltar and the Strait of Sicily is about 900 ft (300 m).

The Atlantic tide disappears in the Strait of Gibraltar. The tides of the Mediterranean are predominantly semidiurnal. The eastern and western basins have standing wave systems, with nodal lines extending roughly from Barcelona to Bougie, and from Kerme Korfezi to Kithera to Tobruch, respectively. The tide range elsewhere is about 0.9 ft (0.3 m). At the ends of the straits of Sicily and Messina, high water occurs in exactly opposite phase, giving rise to strong hydraulic currents, particularly in the narrow Strait of Messina, the legendary seat of Scylla and Charybdis. The Adriatic has a progressive tide, with a range as great as 3 ft (0.9 m) in the northern end, radiating about an amphidromic point near 43°N, 15°E. J. Lyman

Early investigations of ocean currents were stimulated by the continuous inflow of surface water from the Atlantic into the Mediterranean through the Strait of Gibraltar. While it was recognized that evaporation over the Mediterranean greatly exceeds precipitation and river runoff, this inflow is clearly larger than the excess evaporation, and complicated schemes of underground canals were imagined for returning water to the Atlantic. With an ingenious experiment, Count Marsigli suggested in the late seventeenth century that the water should be returned by a deep outflow of Mediterranean water through the Strait of Gibraltar. Marsigli separated two water types of different densities in a tank by a partition. When he removed the partition, the denser water flowed under the lighter water, thereby implying a deep outflow of high-salinity, denser Mediterranean water through the Strait of Gibraltar, in addition to the observed surface inflow of low-salinity, less dense Atlantic water. The deep outflow was finally confirmed in 1870 by lowering a drogue to a depth of 1400 ft (450 m) from a small boat.

The outflow of Mediterranean water can be estimated from consideration of water and salt budgets using evaporation rates and observed salinities of Atlantic and Mediterranean waters in the Strait of Gibraltar. Depending on the assumed evaporation rate, the outflow has been estimated by this method to be between 35 and 63 ft^3/s (1.0×10^6 and 1.8×10^6 m^3/s). During an international experiment in 1960 and 1961, the outflow was measured to be 3.9×10^7 ft^3/s (1.1×10^6 m^3/s) in good agreement with the estimated outflow despite uncertainties in the measured outflow due to spatial and temporal variability.

Because of the excess evaporation over the Mediterranean, surface salinity increases as Atlantic water flows eastward from a value of 36.5‰ in the Strait of Gibraltar to 39.2‰ south of Turkey in the eastern Mediterranean. Surface waters are changed into intermediate and deep waters in wintertime when evaporation and heat lost to the atmosphere make them denser than the water beneath, leading to their downward convection. Levantine intermediate water, which fills the levels between 600 and 3000 ft (200 and 1000 m) depths in most of the Mediterranean, is thought to form near Rhodes in February and March. During the 1970s western Mediterranean deep water was observed to form during March in the same location south of France. Initially, there is a preconditioning phase in which an eddy of slightly colder water forms in the particular location, perhaps trapped by the bottom topography. The surface water inside the eddy becomes more saline and more dense as cold, dry winds, called the mistral, blow down the Rhone valley and over the Mediterranean. Finally, when the surface water becomes as dense as the deep water, convection occurs, and surface water penetrates to great depths and spreads laterally. This formation of deep water is a prototype of deep-water formation in the open ocean which is thought to occur in remote regions of the Arctic and Antarctic.

The outflow of Mediterranean water through the Strait of Gibraltar is made up of a mixture of Levantine intermediate water and western Mediterranean deep water. This outflow is warmer and more saline than Atlantic water of comparable density. These distinctive properties have allowed Mediterranean water influence to be traced westward across the Atlantic to the Gulf Stream and northward into the Norwegian Sea. Harry Bryden

Meiosis

The set of two successive cell divisions that serve to separate homologous chromosome pairs and thus reduce the total number of chromosomes by half. The meiotic process includes two sequential nuclear divisions that must occur prior to the formation of gametes (sperm and eggs). The major purpose of meiosis is the precise reduction in the number of chromosomes by one-half, so a diploid cell can create haploid gametes. To accomplish this reduction, a single cell undergoes two meiotic divisions to produce four daughter cells, each with half the original chromosome complement. The nonmeiotic (or somatic) cells of humans, for instance, have 46 individual chromosomes, or 23 pairs of homologous chromosomes. However, following meiosis, human eggs or sperm have only 23 chromosomes, one member of each pair. Reducing the number of chromosomes in the gametes to 23 allows the fusion of an egg with a sperm (fertilization) to result in an embryo with the requisite 46 chromosomes. Meiosis is therefore a critical component of sexual reproduction. *See* GAMETOGENESIS.

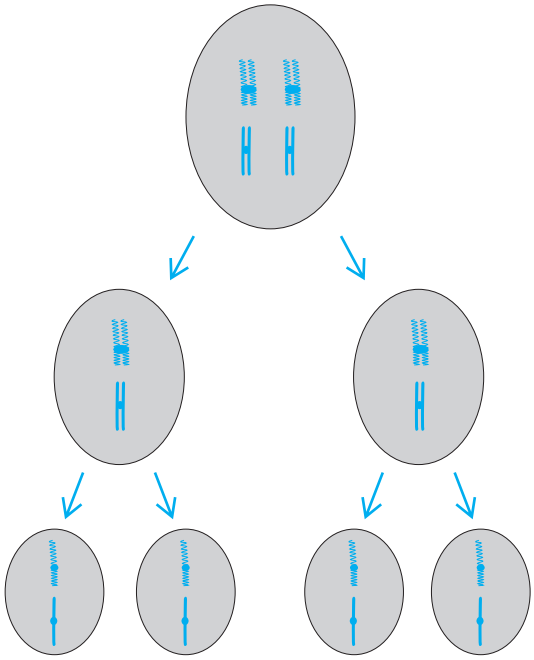

Fig. 1. Simplified view of the two meiotic divisions.

Chromosome behavior. For example, consider an organism that contains only two pairs of chromosomes (**Fig. 1**). The chromosomes in each of these pairs are referred to individually as homologs; one is derived from the father of the organism and the other from the mother. Both homologs carry the same array of genes. As the cell begins meiosis, each chromosome has already duplicated its DNA and carries two identical copies of the DNA molecule. These are visible as two lateral parts, called sister chromatids, which are connected by a centromere.

The basic events of meiosis are actually quite simple. Homologous pairs of chromosomes are first identified and matched. This process, which occurs only in the first of the two meiotic divisions, is called pairing. The matched pairs are then physically interlocked by recombination, which is also known as exchange or crossingover. After recombination, the homologous chromosomes separate from each other, and at the first meiotic division are partitioned into different nuclei. The second meiotic division begins with half of the original number of chromosomes. During this second meiotic division, the sister chromatids of each chromosome separate and migrate to different daughter cells. *See* CHROMOSOME.

The patterns by which genes are inherited are determined by the movement of the chromosomes during the two meiotic divisions. It is a fundamental tenet of Mendelian inheritance that each individual carries two copies of each gene, one derived from its father and one from its mother. Moreover, each of that individual's gametes will carry only one copy of that gene, which is chosen at random. The process by which the two copies of a given gene are distributed into separate gametes is referred to as segregation. Thus, if an individual is heterozygous at the *A* gene for two different alleles, *A* and *a*, his or her gametes

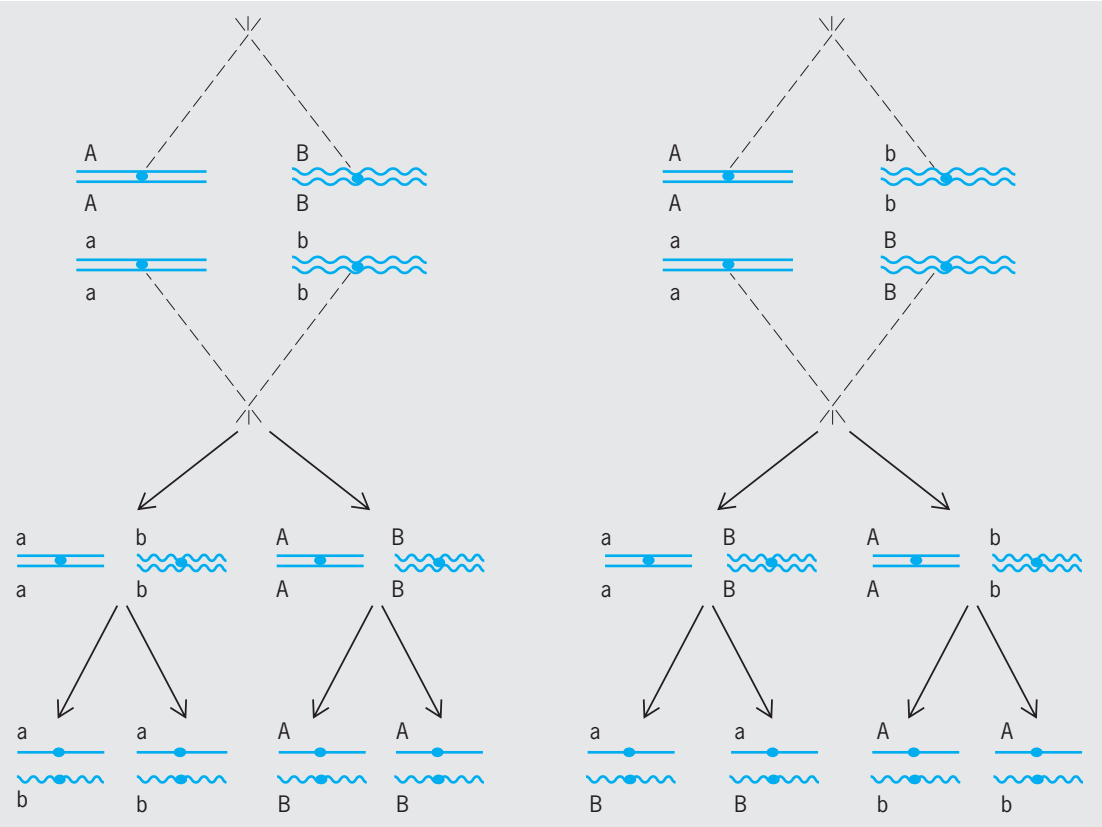

Fig. 2. Behavior of chromosomes at the first meiotic division explains independent assortment.

will be equally likely to carry the *A* allele or the *a* allele, but never both or neither. The fact that homologous chromosomes, and thus homologous genes, segregate to opposite poles at the first meiotic division explains this principle of inheritance. *See* CELL CYCLE.

Mendel's law of independent assortment states that the segregation of two different gene pairs occurs at random with respect to each other. Thus, for an individual of the genotype *AaBb*, the gametes *AB*, *Ab*, *aB*, and *ab* will be formed with equal frequency. This result can be easily understood if the *A* and *B* genes lie on different chromosomes (**Fig. 2**). Because the chromosome pair bearing the *A* gene orients independently of the homolog pair bearing the *B* gene, *AB* ↔ *ab* segregations are as likely as *Ab* ↔ *aB* segregations. Cases where independent assortment does not occur (an effect called linkage) can be understood as resulting from situations where the two gene pairs lie at different positions on the same pair of homologous chromosomes.

Meiotic divisions. The two meiotic divisions may be divided into a number of distinct stages. Meiotic prophase refers to the period after the last cycle of DNA replication, during which time homologous chromosomes pair and recombine. The end of prophase is signaled by the breakdown of the nuclear envelope, and the association of the paired chromosomes with the meiotic spindle. The spindle is made up of microtubules that, with associated motor proteins, mediate chromosome movement. In some cases (such as human sperm formation), the spindle is already formed at the point of nuclear envelope breakdown, and the chromosomes then attach to it. In other systems (such as human female meiosis), the chromosomes themselves organize the spindle.

Metaphase I is the period before the first division during which pairs of interlocked homologous chromosomes, called bivalents, line up on the middle of the meiotic spindle. The chromosomes are primarily (but not exclusively) attached to the spindle by their centromeres such that the centromere of one homolog is attached to spindle fibers emanating from one pole, and the centromere of its partner is attached to spindle fibers from the other pole (**Fig. 3**). The bivalents are physically held together by structures referred to as chiasmata that are the result of meiotic recombination events. In most meiotic systems, meiosis will not continue until all of the homolog pairs are properly oriented at the middle of the spindle, the metaphase plate. It is important to remember that the orientation of each pair of homologs on the spindle occurs in a random fashion, such that the paternally derived homolog of one bivalent may point toward one pole of the spindle, while in the adjacent bivalent the maternally derived homolog is oriented toward the same pole.

Anaphase I refers to the point at which homologous chromosome pairs separate and move to opposite poles. This is accomplished by the release of chiasmata. Although the sister chromatids remain

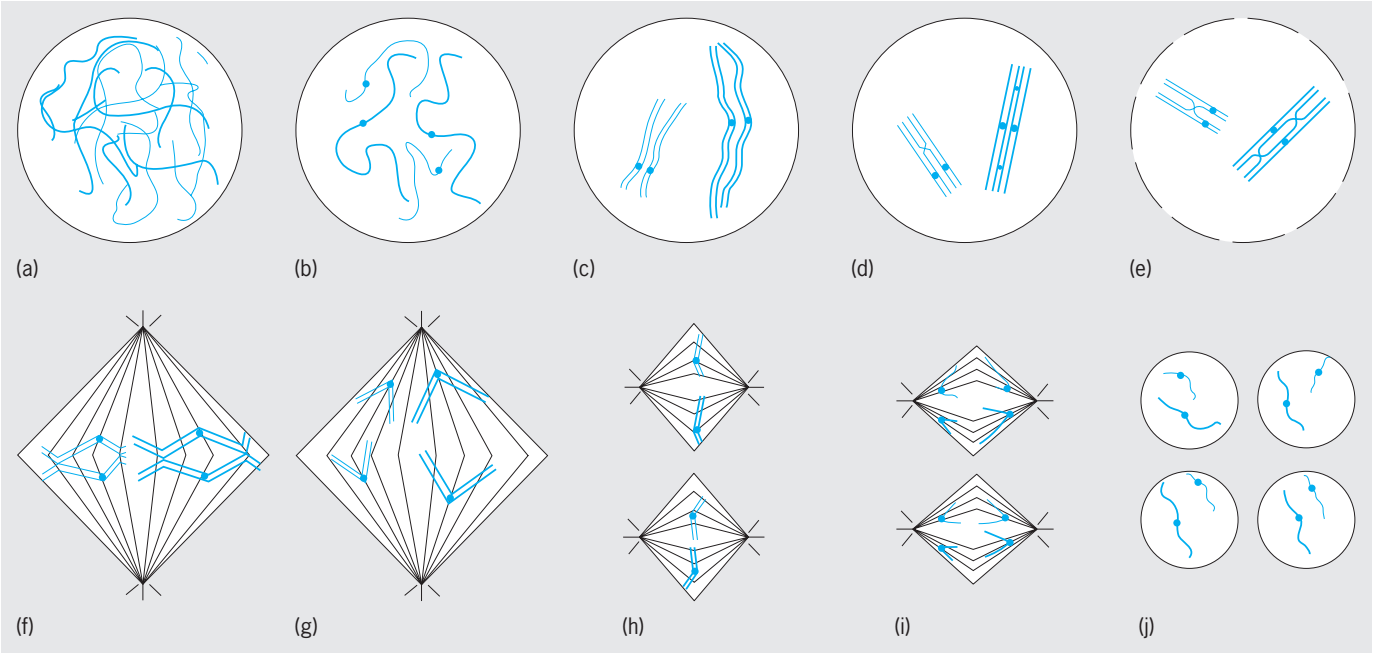

Fig. 3. Stages of meiosis. (*a*) Premeiotic interphase. (*b*) Leptotene (note the appearance of individualized chromosomes). (*c*) Zygotene (signified by the alignment of homologs). (*d*) Pachytene (the chromosomes making up each bivalent are intimately aligned). (*e*) Diplotene/diakinesis. (*f*) Metaphase I. (*g*) Anaphase I. (*h*) Metaphase II. (*i*) Anaphase II. (*j*) Telophase II.

attached around their centromeres, they release each other along the arms of the chromosome, allowing chiasmata to be resolved. Depending on the organism, there may or may not be a true telophase, or a time in which nuclei reform. In most organisms, the first cell division occurs after the completion of anaphase I.

Following the completion of the first meiotic division, the chromosomes align themselves on a new pair of spindles, with their sister chromatids oriented toward opposite poles. The stage at which each chromosome is so aligned is referred to as metaphase II. In some, but not all, organisms, metaphase II is preceded by a brief prophase II. DNA replication does not occur during prophase II; each chromosome still consists of the two sister chromatids. Nor are there opportunities for pairing or recombination at this stage due to the prior separation of homologs at anaphase I.

The start of anaphase II is signaled by the separation of sister centromeres, and the movement of the two sister chromatids to opposite poles. At telophase II, the sisters have reached opposite poles and the nuclei begin to reform. The second cell division usually occurs at this time. Thus, at the end of the second meiotic division, there will be four daughter cells, each with a single copy of each chromosome.

Details of meiotic prophase. Because pairing and recombination occur during the first meiotic prophase, much attention has been focused on this stage of the process. The prophase of the first meiotic division is subdivided into five stages: leptotene, zygotene, pachytene, diplotene, and diakinesis. Homolog recognition, alignment, and synapsis occur during leptotene and zygotene. In the leptotene, an initial phase of chromosome individualization, initial homolog alignments are made. By zygotene, homologous chromosomes have become associated at various points along their length. These associations facilitate a more intimate pairing that results in the homologous chromosomes lying abreast of a tracklike structure called the synaptonemal complex.

The beginning of pachytene is signaled by the completion of a continuous synaptonemal complex running the full length of each bivalent. During diplotene, the attractive forces that mediated homologous pairing disappear, and the homologs begin to repel each other. Luckily, homologs virtually always recombine, and those recombination events can be seen as chiasmata that tether the homologs together. In some organisms, those rare chromosome pairs that have failed to undergo recombination will fly apart prematurely from each other at this stage. The final stage in meiotic prophase is diakinesis, during which the homologs shorten and condense in preparation for nuclear division.

Mechanisms of pairing. Recent models of meiotic pairing suggest that homologous chromosomes initially undergo multiple interactions at many sites along their lengths, and often at their tips. These interactions are the result of several types of homology—finding mechanisms which serve to bring the chromosomes into alignment. However, these associations are apparently weak and transient, and must be stabilized early in meiotic prophase by one of two mechanisms. In the yeast *Saccharomyces cerevisiae*, these pairings can apparently be locked in by the initiation of recombination. The formation of recombination intermediates appears to be fully sufficient to ensure segregation even in genetic

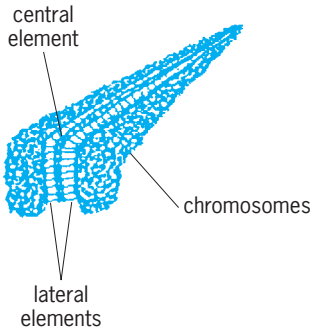

Fig. 4. Synaptonemal complex.

backgrounds in which synapsis and synaptonemal complex formation do not occur. In contrast, recombination is delayed in many higher organisms and its completion appears to require the formation of the synaptonemal complex. In these higher organisms, it is synaptonemal complex formation during zygotene that appears to maintain chromosome pairing during the processes of chromatin compaction and chromosome movement which accompany the progression of meiotic prophase.

The mature synaptonemal complex is a tripartite structure consisting of two lateral elements that flank the chromosomes, and a central element, and running the full length of each bivalent (**Fig. 4**). It is believed that the synaptonemal complex is essential for both synapsis and sister chromatid cohesion, and that proteins involved in maintaining sister chromatid cohesion (known as cohesins) are a major component of the synaptonemal complex itself.

Mechanisms of recombination. Meiotic recombination involves the physical interchange of DNA molecules between the two homologous chromosomes, thus allowing the creation of new combinations of alleles for genes located on that pair of chromosomes. Mechanistically, recombination occurs between homologous chromosomes, and involves the precise breakage and rejoining of two nonsister chromatids. The result is the formation of two recombinant chromatids, each of which carries information from both of the original homologs (**Fig. 5**).

The actual mechanism of recombination involves a complex series of cutting and rejoining of homologous DNA. The initial event appears to be the controlled cleavage of one of the two nonsister chromatids to create a double-strand break, in which both phosphodiester backbones of the double-helical DNA molecule are cleaved. The initial break is then extended to create a gap. Sequences adjacent to the gap seek out identical regions of DNA sequence on the homolog, then use the homolog as a template

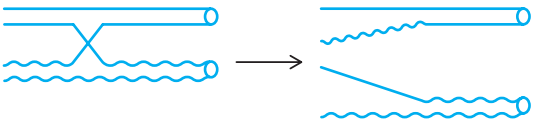

Fig. 5. Recombination involves the physical interchange of genetic material between two nonsister chromatids.

to replace the DNA base pairs removed during gap formation. As a result of this process, the two DNA molecules exchange strands at two closely linked sites and create a structure referred to as a double Holliday intermediate. The resolution of this intermediate can either create a pair of recombinant chromatids or simply result in a short region of DNA from one chromatid being inserted into the other as a result of the gap-filling process. This latter event is called gene conversion.

The number and position of recombination events is very precisely controlled. Exchange occurs only in the gene-rich euchromatin that makes up most of the chromosome arms, never in the heterochromatin that surrounds the centromeres. Moreover, as a result of a process known as interference, the occurrence of one exchange in a given chromosomal region greatly decreases the probability of a second exchange in that region. *See* RECOMBINATION (GENETICS).

Function of genetic recombination. Recent studies have indicated that failures of meiotic chromosome segregation in many organisms, including humans, may arise by a common mechanism. In these organisms, the absence or misplacement of exchange events along the length of a bivalent greatly increases the probability of failed segregation at anaphase I. For certain chromosomes (such as human chromosome 21 and the *Drosophila* X chromosome), exchange events that occur in the distal or most proximal regions of the chromosome frequently result in segregation errors at metaphase I.

Recombination events. The cytological manifestation of an exchange event is a chiasma. Chiasmata can be visualized at diplotene-diakinesis as sites on the bivalent in which two nonsister chromatids appear to cross over from one homolog to the other. Because the sister chromatids display tight cohesion during meiotic prophase and metaphase I, chiasmata hold the paired chromosomes together and thus commit the two centromeres that are physically linked by the exchange to orient to opposite poles of the spindle. Once the two centromeres have oriented toward opposite poles, the chiasmata also function to halt the progression of the homologous centromeres toward opposite poles, thus holding the properly oriented bivalent at the metaphase plate. This represents a stable position in which the bivalent will remain until anaphase I.

Linkage (recombination mapping). Recombination can occur anywhere within the euchromatin but never occurs within the heterochromatin. For two genes on the same chromosome, recombination can be detected by the creation of new combinations of alleles. Consider the case of an *AaBb* double heterozygote, where the *A* and *B* genes lie on the same chromosome arm. In this instance, the *A* and *B* alleles lie on one homolog and the *a* and *b* alleles lie on the other. In the absence of recombination, only *AB* and *ab* gametes can be produced. However, recombination allows the production of *aB* and *Ab* gametes, and the frequency of such gametes will be proportional to the frequency of recombination (**Fig. 6**).

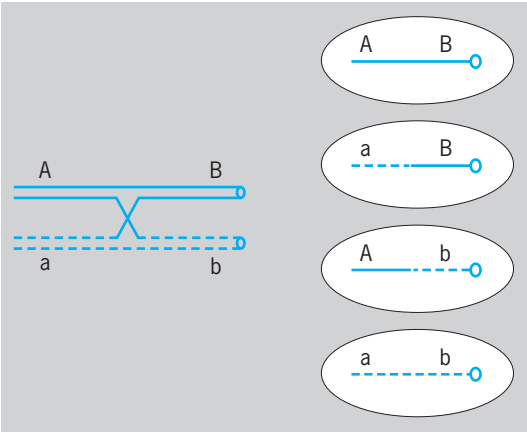

Fig. 6. Recombination leading to linkage.

To a rough approximation, the farther apart two genes are on a given chromosome, the more likely it is that a recombination event will occur between them. Since virtually all homolog pairs undergo at least one recombination event per arm at each meiosis (and the longer chromosome arms may experience three to five exchanges), the probability that two genes far apart on a given arm will have at least one recombination event occur between them is quite high. In contrast, genes lying close to each other will recombine only very rarely. Thus, by determining how often recombination occurs between various pairs of genes, geneticists can build linkage maps and order the position of genes along a given chromosome. *See* LINKAGE (GENETICS).

Sex differences in meiosis. In human male meiosis, all four daughter cells of meiosis will go through a complicated cellular differentiation process called spermiogenesis to become mature functional sperm. In contrast, oogenesis results in only one of the four products of meiosis becoming an egg. The other three products donate their cytoplasm to the chosen oocyte, then die. The oocyte then completes the cellular differentiation process to become a mature egg.

In human females, meiosis begins during fetal development. All of the oocytes (eggs) that a human female will possess in her lifetime are produced during fetal development, but these oocytes are arrested at the end of pachytene. Thus, all of the meiotic recombination a human female will ever do is completed before she is born. These arrested oocytes remain quiescent until the girl enters puberty. At that point, a few oocytes are allowed to begin the maturation process during each menstrual cycle; usually only a single oocyte is ovulated per cycle. The ovulated egg completes anaphase I and then proceeds through meiosis until it arrests again at metaphase II. The completion of the second meiotic division is triggered only by fertilization.

Male meiosis begins at puberty and continues uninterrupted throughout the life of the male. The stem cells that will produce male meiotic cells continue to divide throughout the male's life, constantly producing new populations of spermatocytes. Moreover, once meiosis is initiated in human spermatocytes, it usually proceeds without interruption to produce four daughter cells, all of which differentiate to become mature spermatids.

The molecular mechanisms that ensure meiotic segregation are also different in males and females. In human males, the meiotic spindle is organized by the centrosomes. The chromosomes then attach to the developing spindle after it is formed by the centrosomes. In females, the chromosomes themselves bind to the microtubules, and build the spindle from the inside out without the assistance of the centrosomes. The frequency and distribution of recombination events for each chromosome pair also differ between males and females.

Errors of meiosis. The failure of two chromosomes to segregate properly is called nondisjunction. Nondisjunction occurs either because two homologs failed to pair and/or recombine or because of a failure of the cell to properly move the segregating chromosomes on the meiotic spindle. The result of nondisjunction is the production of gametes that are aneuploid, carrying the wrong number of chromosomes. When such a gamete is involved in a fertilization event, the resulting zygote is also aneuploid. Those cases where the embryo carries an extra copy of a given chromosome are said to be trisomic, while those that carry but one copy are said to be monosomic for that chromosome. Most aneuploid zygotes are not viable and result in early spontaneous abortion. There are no viable monosomies for the human autosomes; however, a few types of trisomic zygotes are capable of survival. These are trisomies for the sex chromosomes (XXX, XXY, XYY), trisomy 21 (Down syndrome), trisomy 18, and trisomy 13. *See* CROSSING-OVER (GENETICS).

Meiosis versus mitosis. The fundamental difference between meiosis and mitosis is that sister chromatids do not separate at the first meiotic division; rather, homologous chromosomes separate from each other with their sister chromatids still attached to each other. Although in some organisms some chromosome pairing does occur in mitotic cells, intimate synapsis along the entire length of chromosomes is usually restricted to meiotic cells. Recombination is frequent in most meiotic cells; however, it occurs only rarely in mitotic cells, usually as part of DNA repair events.

Most critically, DNA synthesis occurs only once within the two meiotic divisions, while there is a complete replication before every mitotic division. This allows mitosis to produce two genetically identical daughter cells, while meiosis produces four daughter cells, each of which have only one-half the number of chromosomes present prior to meiosis. *See* CELL DIVISION; FUNGAL GENETICS; GENE; MITOSIS.

Michelle Y. Walker; R. Scott Hawley

Bibliography. S. L. Page and R. S. Hawley, Chromosome choreography: The meiotic ballet, *Science*, 301(5634):785–789, 2003; M. Petronczki, M. F. Siomos, and K. Nasmyth, Un menage a quatre: The molecular biology of chromosome segregation in meiosis, *Cell*, 112(4):423–440, 2003; D. Zickler

and N. Kleckner, Meiotic chromosomes: integrating structure and function, *Annu. Rev. Genet.*, 33:603–754, 1999.

Meissner effect

The expulsion of magnetic flux from the interior of a superconducting metal when it is cooled in a magnetic field to below the critical temperature, near absolute zero, at which the transition to superconductivity takes place.

Discovery. It was discovered by Walther Meissner and R. Ochsenfeld in 1933, when they measured the magnetic field surrounding two adjacent long cylindrical single crystals of tin and observed that at 3.72 K the Earth's magnetic field was expelled from their interior. This indicated that at the onset of superconductivity they became perfect diamagnets. This discovery showed that the transition to superconductivity is reversible, and that the laws of thermodynamics apply to it. The Meissner effect forms one of the cornerstones in the understanding of superconductivity, and its discovery led F. London and H. London to develop their phenomenological electrodynamics of superconductivity. *See* DIAMAGNETISM; THERMODYNAMIC PRINCIPLES.

The magnetic field is actually not completely expelled, but penetrates a very thin surface layer where currents flow, screening the interior from the magnetic field.

Before the discovery of the effect, it had been pointed out that simple relations exist between the specific heat, a thermal property, and the magnetic properties of superconductors. These relations imply that the change from the normal to the superconducting state of a metal is a thermodynamic transition. The conclusion that a magnetic field would be expelled in the transition to the superconducting state, however, was not drawn.

Shortly before the discovery of the Meissner effect, it was pointed out that because of the inertia of the electrons, screening currents flow in a thin surface layer when a superconductor is suddenly exposed to a magnetic field. Similarly, it was expected that a current passed through a superconducting rod would also be confined to a thin surface layer. This conjecture formed the starting point for Meissner's investigation. By passing the current down one rod and up an identical, adjacent rod, and by measuring the magnetic field surrounding the two rods, it is possible to decide whether the currents flow in a thin surface layer or uniformly over the whole cross section. Since Meissner initially had not canceled the Earth's magnetic field, he was able to observe that the magnetic field was expelled at the onset of superconductivity. Subsequently Meissner also found that indeed the current flows in a very thin surface layer.

Limitations. Meissner and Ochsenfeld prepared two single crystal rods of tin for their experiment. If they had used polycrystalline samples, especially of a noncubic metal, they would not have observed the full diamagnetism. Even when a polycrystalline sample is well annealed and strain-free at room temperature, severe strains are set up in the process of cooling it down, because the thermal contraction is different in different crystal directions. Since the critical temperature shifts slightly with strain, such samples fail to become uniformly superconducting, allowing persistent current to flow in the regions which become superconducting first. These persistent currents then lock in the magnetic flux. *See* MAGNETISM.

In impure samples having a short electronic mean free path, the effect is not observed because the energy associated with a boundary between a superconducting and a normal domain is negative. This allows the formation of a stable array of superconducting flux tubes carrying one quantum of magnetic flux (approximately 2×10^{-7} gauss·cm$^2 = 2 \times 10^{-15}$ weber) each.

The effect is also not observed in a round flat disk with the magnetic field parallel to its axis of rotation, since the magnetic field also would have to be excluded from roughly hemispherical regions above and below the disk. For a disk, however, the difference between the free energies in the superconducting and normal states is too small to balance the change in the magnetic field energy in these two regions. The disk therefore breaks up into normal and superconducting domains or, if extremely thin, exhibits a flux tube structure similar to impure superconductors. *See* SUPERCONDUCTIVITY.

Hans W. Meissner

Meitnerium

The seventeenth of the synthetic transuranium elements. Element 109 falls in column 9 of the periodic table under the elements cobalt, rhodium, and iridium. It is expected to have chemical properties similar to those of iridium. *See* IRIDIUM; PERIODIC TABLE; TRANSURANIUM ELEMENTS.

1																	18
1																	2
H	2											13	14	15	16	17	He
3	4											5	6	7	8	9	10
Li	Be											B	C	N	O	F	Ne
11	12											13	14	15	16	17	18
Na	Mg	3	4	5	6	7	8	9	10	11	12	Al	Si	P	S	Cl	Ar
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
55	56	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86
Cs	Ba	Lu	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
87	88	103	104	105	106	107	108	109	110	111	112	113					
Fr	Ra	Lr	Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg							

lanthanide series	57	58	59	60	61	62	63	64	65	66	67	68	69	70
	La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb
actinide series	89	90	91	92	93	94	95	96	97	98	99	100	101	102
	Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No

Element 109 was discovered in 1982 by a team under P. Armbruster and G. Münzenberg at the

Gesellschaft für Schwerionenforschung (GSI) at Darmstadt, Germany. In a sequence of bombardments of bismuth-209 targets with beams of ions of titanium-50, chromium-54, and iron-58, the compound systems $^{259}105$, $^{263}107$, and $^{267}109$ were produced. The decay analysis of the isotopes produced showed in the case of elements 105 and 107 the production of $^{258}105$ and $^{262}107$ by reaction channels in which one neutron is emitted. These isotopes have odd neutron and proton numbers and possess a special stability against spontaneous fission. It was shown that alpha-particle decay dominated the decay chains. Spontaneous fission occurs through a 30% electron capture branch of $^{256}105$ in $^{258}104$. Three decay chains were observed for the three reactions ending by fission of $^{258}104$, and the decay of the first atom of element 109 was observed. *See* ALPHA PARTICLES; DUBNIUM; NUCLEAR FISSION; NUCLEAR REACTION; RADIOACTIVITY; RUTHERFORDIUM.

The single atom of element 109 was produced at a bombarding energy of 299 MeV in the reaction between iron-58 and bismuth-209. A total dose of 7×10^{17} ions was used to bombard thin layers of bismuth during a 250-h irradiation time. Peter Armbruster

Bibliography. S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; G. Münzenberg, New results on element 109, *Z. Phys.*, A330:435–436, 1998; G. Münzenberg et al., Evidence for element 109 from one correlated decay sequence following the fusion of ^{58}Fe with ^{209}Bi, *Z. Phys.*, A315:145–158, 1984; G. Münzenberg et al., Observation of one correlated alpha decay in the reaction ^{58}Fe on $^{209}Bi \rightarrow {}^{267}109$, *Z. Phys.*, A309:89–90, 1982.

Melanterite

A mineral having composition $FeSO_4 \cdot 7H_2O$. Melanterite occurs mainly in green, fibrous or concretionary masses or in short, monoclinic, prismatic crystals (see **illus.**). Fracture is conchoidal and luster is vitreous. Hardness is 2 on Mohs scale, and specific gravity 1.90. The mineral is similar to chalcanthite, has an astringent taste, and is readily soluble in water.

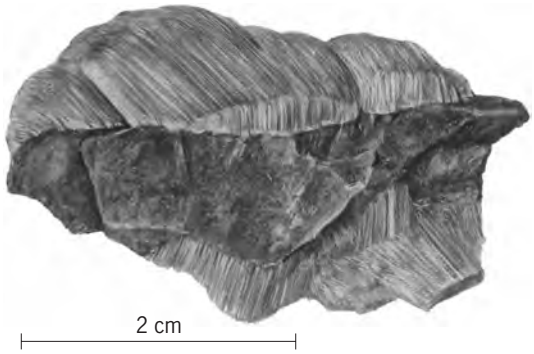

Fibrous crust of melanterite (a common mineral) in a sample from Kremnitz, Hungary. (*Specimen from Department of Geology, Bryn Mawr College*)

Melanterite is a common secondary mineral derived from oxidation and hydration of iron sulfide minerals such as pyrite and marcasite. Its occurrence is widespread. It is not an ore mineral. *See* PYRITE.

Melanterite can be crystallized from aqueous solution at temperatures up to 130°F (56°C). The mineral is unstable in dry air at room temperature, however. It readily dehydrates into a yellowish powder. Edward C. T. Chao

Melilite

A complete solid solution series ranging from gehlenite, $Ca_2Al_2SiO_7$, to akermanite, $Ca_2MgSi_2O_7$, often containing appreciable Na and Fe. Melilites are tetragonal sorosilicates and form short prismatic crystals with poor basal cleavage. The Mohs hardness is 5–6, and the density increases progressively from 2.94 for akermanite to 3.05 for gehlenite. The luster is vitreous to resinous, and the color is white, yellow, greenish, reddish, or brown. Akermanite-rich varieties occur in thermally metamorphosed siliceous limestones and dolomites, but more gehlenite-rich ones result if Al is present, as in Durango, Mexico, and Monzoni, Italy. Melilites are found instead of plagioclase in silica-deficient, feldspathoid-bearing basalts, as at Capo di Bove, Italy; in the Kaiserstuhl, Germany; in the Hawaiian Islands; and at Vesuvius, where SiO_2 was removed from the magma by reaction with

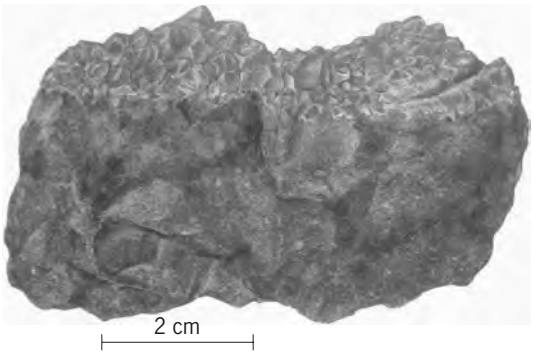

Melilite crystals from Vesuvius, Italy. (*Specimen from Department of Geology, Bryn Mawr College*)

carbonate rocks (see **illus.**). In carbonaceous chondrite meteorites, melilites ranging from pure gehlenite to 80 mole % akermanite occur with hibonite, spinel, perovskite, and pyroxenes in inclusions thought to be ancient condensates from the gas of the solar nebula. *See* SILICATE MINERALS.

Lawrence Grossman

Melting point

The temperature at which a solid changes to a liquid. For pure substances, the melting or fusion process occurs at a single temperature, the temperature rise

with addition of heat being arrested until melting is complete. The direct transition from solid phase to gas phase is not properly called melting, but preferably, sublimation.

Melting points reported in the literature, unless specifically stated otherwise, have been measured under an applied pressure of 1 atm (≃100 kilopascals), usually 1 atm of air. (The solubility of air in the liquid is a complicating factor in precision measurements.) The Clapeyron equation for the pressure dependence of the absolute melting temperature T_m is the equation below, where ΔV_f is the change in

$$\frac{dT_m}{dP} = \frac{T_m \Delta V_f}{\Delta H_f}$$

volume, ΔH_f the heat absorbed during the fusion process, and P the applied pressure. Upon melting, all substances absorb heat, and most substances expand; consequently an increase in pressure normally raises the melting point. A few substances, of which water is the most notable example, contract upon melting; thus, the application of pressure to ice at 0°C (32°F) causes it to melt. Large changes in pressure are required to produce significant shifts in the melting point; a pressure of 10 atm (1 megapascal) lowers the melting point of ice by only 0.075°C (0.135°F).

A sufficient decrease in temperature at ordinary pressures causes all pure substances except helium to freeze to solids; the lowest normal melting point is that of hydrogen, at 14 K (−434°F) and one of the highest is that of rhenium, at 3700 K. Liquid helium can be transformed into a solid only by applying a pressure in excess of 25 atm (2500 kPa).

For solutions of two or more components, the melting process normally occurs over a range of temperatures, and a distinction is made between the melting point, the temperature at which the first trace of liquid appears, and the freezing point, the higher temperature at which the last trace of solid disappears, or equivalently, if one is cooling rather than heating, the temperature at which the first trace of solid appears. Measurement of the freezing point of a solution and the difference between it and the freezing point of the pure solvent provides a convenient method of determining the molecular weight of a dissolved solute, because the freezing point depression of a solution is proportional to the molality of the solute. *See* PHASE EQUILIBRIUM; SOLUTION; SUBLIMATION; TRIPLE POINT.

Robert L. Scott

Membrane distillation

A separation method in which a nonwetting, microporous membrane is used with a liquid feed phase on one side of the membrane and a condensing, permeate phase on the other side. Membrane distillation is also known as transmembrane distillation, membrane evaporation, and thermo-pervaporation. Separation by membrane distillation is based on the relative volatility of various components in the feed solution. The driving force for transport is the partial pressure difference across the membrane. Separation occurs when vapor from components of higher volatility passes through the membrane pores by a convective or diffusive mechanism. *See* CONVECTION (HEAT); DIFFUSION.

Membrane distillation shares some characteristics with another membrane-based separation known as pervaporation, but there also are some vital differences. Both of these membrane separation methods involve direct contact of the membrane with a liquid feed and evaporation of the permeating components. However, while membrane distillation uses porous membranes that have negligible effects on the vapor–liquid equilibria of the components to be separated, pervaporation uses nonporous membranes with the degree of separation influenced by the sorption and diffusion characteristics of permeants in the membrane.

Systems. Membrane distillation systems can be classified broadly into two categories: direct-contact distillation and gas-gap distillation. These terms refer to the permeate or condensing side of the membrane; in both cases the feed is in direct contact with the membrane.

In direct-contact membrane distillation, both sides of the membrane contact a liquid phase; the liquid on the permeate side is used as the condensing medium for the vapors leaving the hot feed solution.

In gas-gap membrane distillation, the condensed permeate is not in direct contact with the membrane. The gas-gap method can be subdivided further, depending on whether the condensing surface is within the module itself or whether condensation of the permeate vapor occurs outside the membrane module. Application of a vacuum or use of a sweep gas are techniques used to carry vapor from the membrane module to an external condensing surface.

Increasing the distance between the hot feed and the condensing surface decreases the heat loss across the membrane while increasing the overall mass-transfer resistance. Choosing the correct gap is important for optimum operation.

For highly porous membranes, the air entrapped in the membrane pores can be a significant component of the total mass-transfer resistance. Thus, deaeration of the feed can be beneficial. It has also been demonstrated that feed pretreatment to remove surfactants can boost transport rates.

Heat losses across the membrane arise through either conduction or convection. Also, the establishment of temperature gradients in the feed (thermal-polarization) can reduce the available driving force for transport. It has been estimated that in membrane distillation 20–70% of the available driving force is actually utilized. Concentration polarization, which is usually significant in other membrane processes, such as ultrafiltration or reverse osmosis, is less important in membrane distillation.

The two important parameters defining the efficiency of a membrane distillation process are the

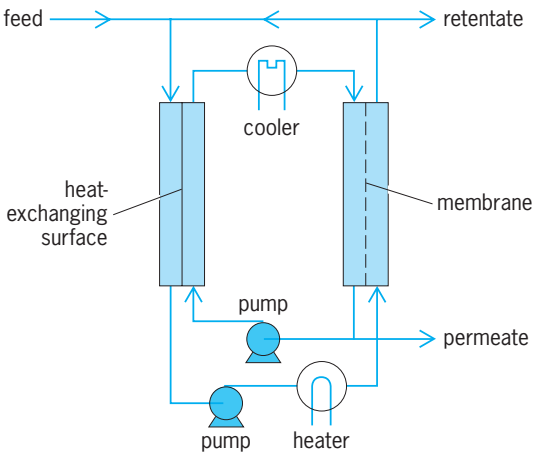

Membrane distillation process with heat recovery. (*After A. C. M. Franken and S. Ripperger, Terminology for Membrane Distillation, European Society of Membrane Science and Technology, 1988*)

energy efficiency [Eq. (1)] and the separation efficiency or concentration factor [Eq. (2)]:

$$\text{Energy efficiency} = \frac{\text{heat used to evaporate permeate}}{\text{total energy input to process}} \quad (1)$$

$$\text{Concentration factor} = \frac{\text{concentration of component in permeate}}{\text{concentration of component in feed}} \quad (2)$$

A schematic diagram of a membrane distillation operation is shown in the **illustration**. The input of energy required to heat and condense vapor is the same for membrane distillation as for traditional evaporation processes. Some degree of heat recovery is possible to improve the overall energy efficiency of the process. Feed can be recycled in order to increase concentration of nonvolatile substances to a desired level.

Membrane materials and modules. Membranes for membrane distillation must be microporous (diameter of pores, 0.2–5 micrometers) and also nonwetting with respect to the feed. Other desirable qualities are chemical and thermal resistance, narrow pore-size distribution, high total porosity, and low thermal conductivity. Water transport fluxes across membranes are typically in the range of 0.2–4 lb/ft^2-h (1–20 kg/m^2-h), depending on temperature differentials of 9–108°F (5–60°C).

Most of the available membranes have been designed for aqueous applications. They are made of hydrophobic polymers such as polypropylene, poly(tetrafluoroethylene), and poly(vinylidene fluoride). Composite membranes consisting of layers of different hydrophobic and hydrophilic materials are also available. Membranes can be produced in either flat-sheet or hollow-fiber/tubular geometries. Microporosity is induced either by mechanical stretching or by a phase-inversion (thermal) process. Higher fractional pore areas give higher transport rates; however, the pore size must be less than the critical pore size at which the liquid would penetrate the membrane. Membrane modules have been developed in various configurations, including spiral-wound, plate-and-frame, and hollow-fiber.

Applications. Membrane distillation has been studied for separations such as sugar solution concentration, dilute blood concentration, and desalination. The economic outlook for membrane distillation compared to either reverse osmosis or multistage flash evaporation has not been determined.

Potential advantages of membrane distillation over traditional evaporation processes include operation at ambient pressures and lower temperatures as well as ease of process scale-up. The additional resistance to mass transfer posed by the membrane can be compensated for to some extent by the high ratio of membrane surface area to volume that can be obtained in commercial membrane modules.

While membrane distillation does not require heating the feed to its boiling point, its value as a commercial process is dependent on the available feed temperature; it works best when a low-grade or waste heat source is available. *See* CHEMICAL SEPARATION TECHNIQUES; MEMBRANE SEPARATIONS; WATER DESALINATION.

S. S Kulkarni; Norman N. Li

Bibliography. S. I. Andersson, N. Kiellander, and B. Rodesjo, Design and field tests of a new membrane distillation desalination process, *Desalination*, 56:345–354, 1985; E. Drioli, V. Calabro, and V. Wu, Microporous membranes in membrane distillation, *Pure Appl. Chem.*, 58(12):1657–1662, 1986; A. G. Fane, R. W. Schofield, and C. J. D. Fell, The efficient use of energy in membrane distillation, *Desalination*, 64:231–243, 1987; A. C. M. Franken and S. Ripperger, *Terminology for Membrane Distillation*, European Society of Membrane Science and Technology, 1988; K. Sakai et al., Effects of temperature and concentration polarization on water vapor permeability for blood in membrane distillation, *Chem. Eng. J.*, 38:B33–39, 1988; S. Timashev, *Physical Chemistry of Membrane Processes*, 1991.

Membrane mimetic chemistry

The study of processes and reactions whose developments have been inspired by the biological membrane. Faithful modeling of the biomembrane is not an objective of membrane mimetic chemistry. Rather, only the essential components of natural systems are recreated from relatively simple, synthesized molecules. Mimetic systems are designed for well-defined purposes. The insight gained into the functioning of real membranes is considered a bonus. (The term membrane mimetic is more restrictive than the term biomimetic. Biomimetic chemistry is directed at the mechanistic elucidation of biochemical reactions and at the development of new compounds modeled on, and expected to mimic, specific biological systems. The focal point of most biomimetic chemistry is enzyme modeling.) *See* CELL MEMBRANES.

Surfactant aggregates. These systems—micelles, monolayers, organized multilayers (Langmuir-Blodgett films), bilayer lipid membranes, and vesicles—have been used, for the most part, in membrane mimetics.

Surfactants, or detergents, contain distinct hydrophobic (apolar) and hydrophilic (polar) regions. Depending on the chemical structure of their hydrophilic polar head groups, surfactants can be neutral, positively charged, or negatively charged. The hydrophobic part can vary in length (typically 8 to 20 carbon atoms), contain multiple bonds, or consist of two or more hydrocarbon chains. Head group–to–tail volume ratios determine, to a first approximation, the type of aggregates formed from a given surfactant. Various types of aggregates formed from surfactants are shown in the **illustration**. *See* DETERGENT.

Aqueous micelles. These spherical aggregates, 4–8 nanometers in diameter, are formed dynamically from surfactants in water above a characteristic concentration, the critical micelle concentration. Micelles break up and reform rapidly by two known processes. The first process occurs on a microsecond time scale and is due to the release of a single surfactant and its subsequent reincorporation in the micelle. The second process occurs on a millisecond time scale and is ascribed to the dissolution of the micelle and to the subsequent reassociation of the monomers. Substrate interaction with the micelle is also dynamic. *See* MICELLE.

Monolayers. These monomolecular layers are formed by spreading naturally occurring lipids or synthetic surfactants, dissolved in volatile solvents, over water in a trough. The polar head groups of the surfactants are in contact with water, the subphase, while their hydrocarbon tails protrude above it. Monolayers are characterized by surface area-surface pressure curves, surface potentials, and surface viscosities. *See* MONOMOLECULAR FILM.

Multilayers. Techniques have been developed for transferring the monolayer onto a solid support and for building up organized multilayer assemblies (Langmuir-Blodgett films) in controlled topological arrangements.

Bilayer lipid membranes. These are formed by brushing an organic solution of a surfactant (or lipid) across a pinhole (2–4 mm or 0.08–0.16 in. in diameter) separating two aqueous phases. Alternatively, bilayer lipid membranes can be formed from monolayers by the Montal-Mueller method. In this method the surfactant, dissolved in an apolar solvent, is spread on the water surface to form a monolayer below the poly(tetrafluoroethylene) partitioning which contains the pinhole (0.1–0.5 mm or 0.004–0.02 in. in diameter). Injection of an appropriate electrolyte solution below the surface raises the water level above the pinhole and brings the monolayer into apposition to form the bilayer lipid membrane.

Vesicles. These are prepared by sonication from naturally occurring phospholipids or from such simple surfactants as dioctadecyldimethylammonium bromide (DODAB) or dihexadecylphosphate (DHP). They are single bilayer spherical aggregates with diameters of 50–100 nm and bilayer thickness of around 5 nm. Unlike micelles, vesicles do not break down on dilution. Nevertheless, they are dynamic

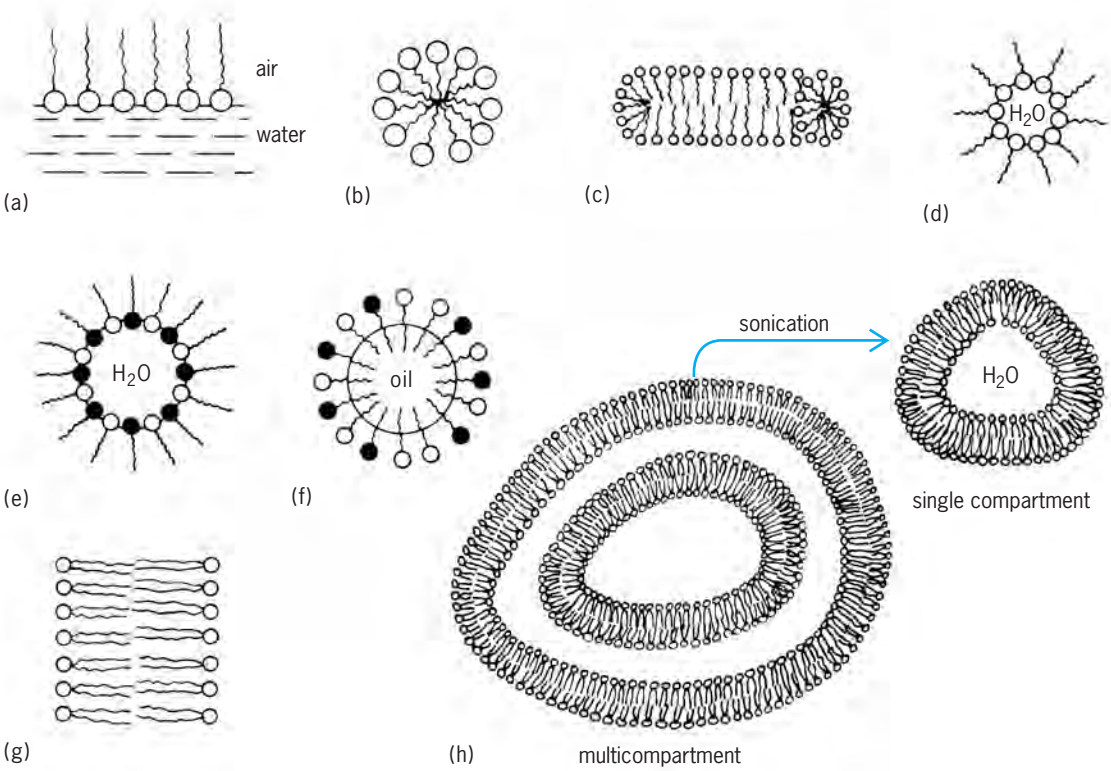

Simplified representation of organized aggregates formed from surfactants. (*a*) Monoloyer. (*b*) Spherical micelli. (*c*) Rodlike micelle. (*d*) Reversed micelli. (*e*) Water-in-oil microemulsion. (*f*) Oil-in-water microemulsion. (*g*) Bilayer lipid membrane. (*h*) Vesicles.

structures. They undergo phase transition, fuse, and are osmotically active. Molecular motions of the individual surfactants in the vesicles involve rotations, kink formation, lateral diffusion on the vesicle plane, and transfer from one interface of the bilayer to the other (flip-flop). Vesicles are capable of organizing a large number of molecules in their compartments. Hydrophobic molecules can be distributed among the hydrocarbon bilayers of vesicles. Polar molecules may move about relatively freely in vesicle-entrapped water pools, particularly if they are electrostatically repelled from the inner surface. Small charged ions can be electrostatically attached to the oppositely charged vesicle surfaces. Species having charges identical with those of the vesicles can be anchored onto the vesicle surface by a long hydrocarbon tail. *See* SONOCHEMISTRY.

The need for increased stabilities, controllable sizes, and permeabilities led to the development of polymerized surfactant vesicles. Vesicle-forming surfactants have been functionalized by vinyl, methacrylate, diacetylene, isocyano, and styrene groups in their hydrocarbon chains or at their head groups. Accordingly, surfactant vesicles could be polymerized in their bilayers or across their head groups. In the latter case, either the outer or the inner vesicle surfaces could be linked separately. All polymerized vesicles show enhanced stabilities compared with their unpolymerized counterparts. They have extensive shelf lives and remain unaffected by the addition of up to 30% alcohol.

Applications. Within a remarkably short time, membrane mimetic chemistry has become a versatile chemical tool. Applications of compartmentalization of reactants in membrane mimetic systems involve altered reaction rates, products, stereochemistries, and isotope distributions. Monolayers and organized multilayers can be employed profitably as molecular electronic devices. Opportunities also exist for using different surfactant aggregates with polymeric membranes for the control and regulation of reverse osmosis and ultrafiltration. The combination of judiciously selected polymeric membranes and functionalized surfactants holds the key for efficient membrane reactors. *See* ULTRAFILTRATION.

Well-constructed surfactant vesicles and polymerized surfactant vesicles can be used as compartments for drugs, offering the possibility of target directing. The exclusive delivery of a given drug in a controllable manner has long been a goal of pharmacists. Much of the activity in this area is based on liposomes (vesicles prepared from phospholipids of biological origin) together with immunological methodologies. Chemists are investigating the use of polymerized vesicles as hosts for drugs. On-command depolymerization (by an enzyme, for example) could trigger the release of drugs at the desired targets. *See* LIPOSOMES; SURFACTANT.

Janos H. Fendler

Bibliography. J. H. Fendler, *Membrane-Mimetic Approach to Advanced Materials*, 2000; J. H. Fendler, *Membrane Mimetic Chemistry*, 1982; J. H. Fendler and P. Tundo, Polymerized surfactant aggregates: Characteristics and utilization, *Acc. Chem. Res.*, 17:3–8, 1984.

Membrane separations

Processes for separating mixtures by using thin barriers (membranes) between two miscible fluids. A suitable driving force across the membrane, for example, concentration or pressure differential, leads to the preferential transport of one or more feed components.

Types of processes. Membrane separation processes are classified under different categories depending on the materials to be separated and the driving force applied: (1) In ultrafiltration, liquids and low-molecular-weight dissolved species pass through porous membranes while colloidal particles and macromolecules are rejected. The driving force is a pressure difference, typically 10–200 lb/in.^{2} (70–1400 kilopascals). (2) In dialysis, low-molecular-weight solutes and ions pass through while colloidal particles and solutes with molecular weights greater than 1000 are rejected under the conditions of a concentration difference across the membrane. (3) In electrodialysis, ions pass through the membrane in preference to all other species, due to a voltage difference. (4) In reverse osmosis, virtually all dissolved and suspended materials are rejected and the permeate is a liquid, typically water. Pressure differences up to 800 lb/in.^{2} (6 megapascals) are used. (5) For gas and liquid separations, unequal rates of transport can be obtained through nonporous membranes by means of a solution and diffusion mechanism. Pressure differences in the range of 15–1500 lb/in.^{2} (0.1–10 MPa) are typical. Pervaporation is a special case of this separation where the feed is in the liquid phase while the permeate, typically drawn under subatmospheric conditions, is in the vapor phase. (6) In facilitated transport, separation is achieved by reversible chemical reaction in the membrane. High selectivity and permeation rate may be obtained because of the reaction scheme. Liquid membranes are used for this type of separation.

Some examples of mixtures separated by membranes are shown in the **table**. Since membranes are permeable to all feed components, though with different permeation rates, a membrane process may require staging to meet specified product concentrations. Various flow and recycle schemes can be used in a multistage process. *See* DIALYSIS; ION-SELECTIVE MEMBRANES AND ELECTRODES; OSMOSIS; ULTRAFILTRATION.

Membrane characteristics. Most membranes are thin polymer films. Since the transport rate through the membrane increases with decreasing membrane thickness, techniques have been developed to make ultrathin, separating films without sacrificing membrane strength. For example, cellulose acetate membranes can be made in an asymmetric form, where a thin dense layer (0.01–0.1 micrometer) is supported by a thick porous substrate (50–100 μm) which provides strength but little resistance to transport. Another technique is to make thin-film-composite membranes where a thin, separating layer of one material is cast on a porous support of another polymer.

Materials other than polymers can also function as membranes. Palladium can separate helium from

Mixtures separated by membrane permeation

Membrane	Temp., °C (°F)	Components	Feed comp., wt %	α B/A
Polyethylene	60 (140)	Methanol	39.1	6.5
		Benzene	60.9	
	25 (77)	Ethylene	50	3.0
		Methane	50	
Polyethylene (conditioned)	30 (86)	*o*-Xylene (A)	30*	1.59 (B/A)
		m-Xylene (B)	65*	1.26 (C/B)
		p-Xylene (C)	5*	2.01 (C/A)
Polypropylene	25 (77)	Ethylene	50	1.95
		Methane	50	84
	60 (140)	*n*-Dodecane	46	
		n-Heptane	54	
Polystyrene	30 (86)	Nitrogen	—	16†
		Helium	—	
Poly(tetrafluoroethylene)	25 (77)	Ethylene	55	1.25
		Methane	45	25*
	30 (86)	Nitrogen	—	
		Helium	—	
Polyethylene terephthalate	25 (77)	Methane	—	170*
		Helium	—	
Nylon	40 (104)	Chloroform	37.5	1.34
		Trichloroethylene	62.5	
Silicone rubber	30 (86)	Nitrogen	—	1.5†
		Helium	—	
Cellulose acetate	—	*n*-Hexane	—	17–130
		Ethanol	—	
	25 (77)	Ethylene	45	1.66
		Methane	55	
Ethylenediamine/poly-epichlorohydrin/iso-phthaloyl chloride	55 (131)	Sodium chloride	3.596.5	167
		Water		
Liquid membrane: surfactant/Isopar M/LIX 64N‡	35 (95)	Water	99.94 0.06	500–2800
		Copper salts		

*Vol %.
†Calculated as ratio permeability coefficients.
‡Surfactant is a polyamine, Isopar M is an isoparaffin solvent, and LIX 64N is a chelating agent.

other gases while silica can be used for hydrogen separation.

Liquid films have been used in two ways as membranes: (1) Immobilized liquid membranes: a liquid is held by a porous (~80% void) polymer support. (2) Liquid surfactant membranes: the membrane is formed by making an emulsion of two immiscible fluids and dispersing the emulsion in a third continuous phase. The immiscible membrane separates the encapsulated phase from the continuous phase.

Advantages and limitations. In general, membrane separation processes have potential advantages for mixtures of thermally unstable components; mixtures of components with similar chemical and physical properties, such as xylene isomers; and separations without phase change, as contrasted to distillation, such as desalination by reverse-osmosis membranes. Membranes may also be used in combination with conventional separation units, for example to break an azeotropic solution before feeding it to a distillation column. Energy savings are possible by using a membrane to concentrate a stream or to make a rough-cut separation before going to an evaporator or distillation unit. *See* WATER DESALINATION.

Usually, membranes are used only under mild temperature conditions. The presence of various chemically active agents or organic solvents in the feed may lead to plasticizing problems and eventual membrane deterioration.

Quantification of separation. The degree of separation obtained in a single membrane module is expressed as a separation factor α defined in terms of concentrations C, as shown in Eq. (1), where sub-

$$\alpha = \frac{(C_A/C_B)_P}{(C_A/C_B)_R} \tag{1}$$

scripts A, B, P, and R refer to a permeate component, reference component, permeate phase, and raffinate phase, respectively. Separation factor α can be a function of temperature, pressure, feed composition, flow rates, and directions of flow on the feed and permeate sides, for example, countercurrent, crossflow, or cocurrent.

For gas permeation, the separation obtainable with a given membrane is often quantified by a selectivity factor (P_A/P_B), where the permeability constant P is defined by the relation given in Eq. (2),

$$F_A = \frac{P_A}{\delta} \Delta p_A \qquad (2)$$

where F is the flux/area, δ is the membrane thickness, and Δp represents the partial pressure difference across the membrane. Using Fick's first law, P can be shown to be the product of an effective diffusion constant and an effective solubility constant. P can be a strong function of pressure (or concentration) for organic gases and vapors. *See* DIFFUSION; TRANSPORT PROCESSES.

Norman N. Li; Sudhir S. Kulkarni

Bibliography. J. R. Critser, Jr., *Membrane Separation Processes*, 1994; H. C. Hayworth and N. N. Li, Extraction of uranium by liquid membranes, *Separat. Sci. Technol.*, 18(6):493, 1983; S. T. Hwang and K. Kammermeyer, *Membranes in Separations*, 1975; N. N. Li and H. Strathmann (eds.), *Separation Technology*, 1988; P. Meares, *Membrane Separation Processes*, 1976.

Memory

The ability to store and access information that has been acquired through experience. Memory is a critical component of practically all aspects of human thinking, including perception, learning, language, and problem solving. Through its elaborate mechanisms for sorting, organizing, and storing information, memory provides people with the building blocks necessary to structure their knowledge of the world and, ultimately, of themselves. *See* PERCEPTION; PROBLEM SOLVING (PSYCHOLOGY).

Stages. The information-processing approach divides memory into three general stages: sensory memory, short-term memory, and long-term memory. Sensory memory refers to the sensations that briefly continue after something has been perceived. For example, if a person sits in a dark room and then is subjected to a flash of light for just a moment, the image will linger after darkness returns. Sensory memories occur for all five senses and are usually very similar to the perceptions from which they arise. Short-term memory includes all of the information that is currently being processed in a person's mind, and is generally thought to have a very limited capacity. Memorizing a telephone number just long enough to dial it is an example of a task that involves short-term memory. Long-term memory is the storehouse of the mind, where all the information that may be used at a later time is kept. Childhood memories of the beach as well as what dinner comprised one night ago are examples of long-term memories.

Sensory memory. Psychologists were long aware of sensory memory without being able to experimentally demonstrate its full capacity. For example, after viewing a briefly flashed slide, subjects reported seeing the entire image gradually fade away. However, by the time they tried to describe what they saw, most of their sensory memory had faded. G. Sperling devised a technique known as the partial report test to demonstrate the full extent of sensory memory. Sperling's experiment had two different conditions for reporting the contents of a briefly flashed slide containing three rows of letters. In the full report condition, subjects saw the slide, and afterward had to name as many of the letters as they could. Subjects in this condition could identify only about a third of the letters. In the partial report condition, Sperling flashed the same slide, but immediately afterward gave subjects a tone indicating which row of letters they were to name. With the partial report technique, subjects could name almost all of the items in any of the three rows. Since subjects did not know which row they would have to name until after the image physically disappeared, they must have had a memory of all of the items in the slide. Subjects in the full report condition probably had the same full memory, but could name only a few letters before their memory faded. Thus, Sperling's experiment suggested that people have a very detailed sensory memory that lasts only a brief time.

A number of interesting facts are known about sensory memory, including the following: (1) sensory memories appear to be associated with mechanisms in the central nervous system rather than at the sensory receptor level, and (2) the amount of attention that a person pays to a stimulus can affect the duration of the sensory memory. Although all of the functions of sensory memory are not understood, one of its most important purposes is to provide people with additional time to determine what should be transferred to the next stage in the memory system, commonly known as short-term memory.

Short-term memory. Information obtained from either sensory memory or long-term memory is processed in short-term memory in order for a person to achieve current goals. In some situations, short-term memory processing simply involves the temporary maintenance of a piece of information, such as remembering a phone number long enough to dial it. Other times, short-term memory can involve elaborate manipulations of information in order to generate new forms. For example, when someone reads 27 + 15, the person manipulates the symbols in short-term memory in order to come up with the solution.

One useful manipulation that can be done in short-term memory is to reorganize items into meaningful chunks. For example, it is a difficult task to keep the letters S K C A U Q K C U D E H T in mind all at once. However, if they are rearranged in short-term memory, in this case reversing them, they can be reduced to a single simple chunk: THE DUCK QUACKS. Short-term memory can accommodate only five to seven chunks at any one time. It therefore becomes very important that each chunk include as much information as possible.

Although, short-term memory holds only a limited number of chunks, the amount of information

contained in each chunk is constrained only by one's practice and ingenuity. In one study, an average college student participated in a 20-month program in which he learned to increase his short-term memory span from 7 to 80 random digits. The chunking trick that he used was associating groups of random numbers with familiar track records, an easy task for him since he was a runner. His accomplishment did not represent a general increase in short-term memory capacity (he could still recall only 7 random letters) but rather an improved strategy for organizing random digits. Thus, in order to increase the amount of information that can be kept in short-term memory at one time, people need to develop specific strategies for organizing that information into meaningful chunks.

Not only does manipulating information in meaningful ways increase the capacity of short-term memory, but it also appears to facilitate the transfer of information from short-term to long-term memory. For example, the running student who managed to chunk 80 random digits in his short-term memory was able to recall 80% of those digits days later. Many other studies have also demonstrated that the transfer of information from short-term to long-term memory is much greater when the information is manipulated rather than simply maintained.

Long-term memory. One can keep massive amounts of information in long-term memory. In general, recall from long-term memory simply involves figuring out the heading under which a memory has been filed.

One of the best ways to ensure that a long-term memory will be found is to memorize it with a very specific file heading. Thus, many tricks for effective retrieval of long-term memories involve associating the memory with another more familiar memory that can serve as an identification tag. This trick of using associations to facilitate remembering is called mnemonics. To make recall easier, a mnemonic called the method of loci may be employed which involves associating to-be-remembered items with series of places (or loci) that are already well fixed in memory. For example, to memorize the presidents of the United States, each president can be associated with a different piece of furniture in the house. By making this association, the furniture can then serve as reminders of the mental file in which the presidents' names can be found. Various other tricks involve associating a familiar memory with an unfamiliar memory. For example, one common way of learning new names is to associate the new person with an old friend or acquaintance with the same name.

Usually people do not make explicit attempts to associate a new memory with an old memory, and yet they are generally reasonably good at accessing long-term memories. This is because long-term memory is organized in a logical manner; that is, long-term memory stores related concepts and incidents in close range of one another. This logical association of memories is indicated by subjects' reaction times for identifying various memories. Generally, people are faster at recalling memories if they have recently recalled a related memory.

Considering that related memories are filed near each other, it is not surprising that one good way to locate a long-term memory is to remember the general situation under which it was stored. Accordingly, techniques that reinstate the context of a memory tend to facilitate remembering. For example, in one study divers learned a list of words underwater and were later tested either underwater or on land. The divers who were tested underwater were able to remember considerably more words. Subjects who learn words while they are in a good mood are more likely to remember them when they are in a good mood. Similarly, words memorized by subjects under the influence of alcohol are better remembered when again under the influence. In each of these cases, reinstating the context allows subjects to better locate the memories in their long-term memory file.

Sometimes information may not have been filed in long-term memory in the first place, or if it has, is inaccessible. In these situations, the long-term memory system often fills in the gaps by using various constructive processes. One common component to memory constructions is a person's expectations. Countless studies have also indicated that memories tend to systematically change in the direction of a prior expectation or inference about what is likely to have occurred.

In addition, a long-term memory often becomes distorted as the result of misleading suggestions. In one study, subjects watched a videotape of an auto accident. Later they were asked either "How fast were the cars going when they hit?" or "How fast were the cars going when they smashed?" On the average, subjects who heard the word "smashed" estimated that the cars were going faster, and were more likely to report having seen nonexistent broken glass. Studies like these indicate that the files in long-term memory are vulnerable to tampering.

Dual memory. The distinction between short- and long-term memory is supported by both physiological and experimental evidence. The experimental support for the distinction comes from a number of sources, including observations of people's recall of words from a memorized word list. When people are given a list of items to remember, the items at the beginning and at the end of the list are remembered well, while items from the middle of the list are remembered less well. The fact that the first few items have a good chance of being remembered is called the primacy effect. The fact that the last few items have a good chance of being remembered is called the recency effect. If a hundred people are given a list to remember, and the number who remember each item is graphed, the classic serial position curve will be seen (**Fig. 1**). The serial position curve can be explained because of the operation of both short-term and long-term memory. Items at the end of the list are remembered because they are in short-term memory at the moment the list begins to be recalled. These words are likely to be in short-term memory because no later items have displaced them. After they are

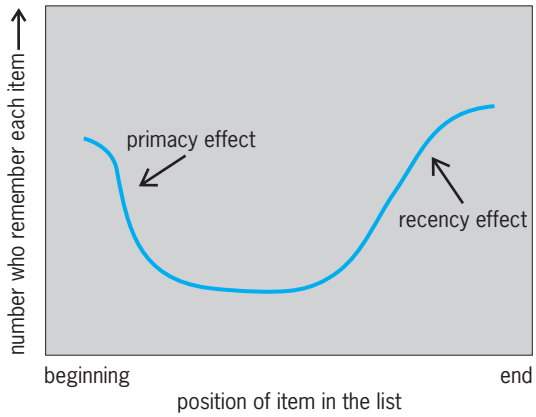

Fig. 1. Serial position curve.

reported, the contents of long-term memory can be searched. The items at the beginning of the list are more likely to be in long-term memory because they get more attention than do other items. Thus, the dual-memory distinction provides a relatively simple explanation for an otherwise puzzling experimental finding.

Although the dual-memory view is widely accepted, there are some theorists who argue that differences between alleged short- and long-term memories can be attributed to differences in how information is processed. One alternative to the dual-memory distinction is known as the levels-of-processing approach. It posits that memories differ only with respect to how "deeply" they have been processed. According to this argument, shallowly processed information will be quickly forgotten, while information that receives a deeper processing will be better remembered. The primary evidence for this view comes from incidental learning situations in which subjects participate in some task involving words. Later, to their surprise, the subjects are asked to recall the words that they previously worked with. Generally, incidental tasks involving "deep" processing (for example, determining the meaning of a word) are associated with better word recall then tasks involving "shallow" processing (for example, determining whether a word is in upper- or lowercase). Both the dual processing and the depth of processing concepts have something to offer. Specifically, depth of processing is an important factor in determining if a memory will last a long time. It is convenient to say that the resulting enduring memory is stored in long-term memory; certain physiological studies support this idea. *See* INFORMATION PROCESSING (PSYCHOLOGY).

Physiology. A number of physiological mechanisms appear to be involved in the formation of memories, and the mechanisms may differ for short-term and long-term memory. There is both direct and indirect evidence suggesting that short-term memory involves the temporary circulation of electrical impulses around complex loops of interconnected neurons. A number of indirect lines of research indicate that short-term memories are eradicated by any event that either suppresses neural activity (for example, a blow to the head or heavy anesthesia) or causes neurons to fire incoherently (for example, electroconvulsive shock). More direct support for the electric circuit model of short-term memory comes from observing electrical brain activity. By implanting electrodes in the brain of experimental animals, researchers have observed that changes in what an animal is watching are associated with different patterns of circulating electrical activity in the brain. These results suggest that different short-term memories may be represented by different electrical patterns. However, the nature of these patterns is not well understood. *See* ELECTROENCEPHALOGRAPHY.

Unlike short-term memories, long-term memories appear to involve some type of permanent structural or chemical change in the composition of the brain. This conclusion is derived both from general observations of the imperviousness of long-term memories and from physiological studies indicating specific changes in brain composition. Even in acute cases of amnesia where massive deficits in long-term memory are reported, often, with time, all long-term memories return. Similarly, although electroconvulsive therapy (shocking the brain in order to produce convulsions) is known to eliminate recent short-term memories, it has practically no effect on memories for events occurring more than an hour prior to shocking. Thus the transfer from a fragile short-term memory to a relatively solid long-term memory occurs within an hour. This process is sometimes called consolidation.

The nature of the "solid" changes associated with long-term memories appears to involve alterations in both the structural (neural connections) and chemical composition of the brain. One study compared the brains of rats that had lived either in enriched environments with lots of toys or in impoverished environments with only an empty cage. The cerebral cortices of the brains of the rats from the enriched environment were thicker, heavier, endowed with more blood vessels, and contained significantly greater amounts of certain brain chemicals (such as the neurotransmitter acetylcholine). Other researchers have observed that brief, high-frequency stimulation of a neuron can produce long-lasting changes in the neuron's communications across synapses.

Researchers believe that different brain structures may be involved in the formation and storage of long-term memories. The hippocampus, thalamus, and amygdala are believed to be critical in the formation of long-term memories (**Fig. 2**). Individuals who have had damage to these structures are able to recall memories prior to the damage, indicating that long-term memory storage is intact; however, they are unable to form new long-term memories, indicating that the long-term memory formation process has been disrupted. It is not known where long-term memories are stored, but they may be localized in the same areas of the brain that participated in the actual learning.

Types of long-term memories. Long-term memory will probably never be completely characterized by

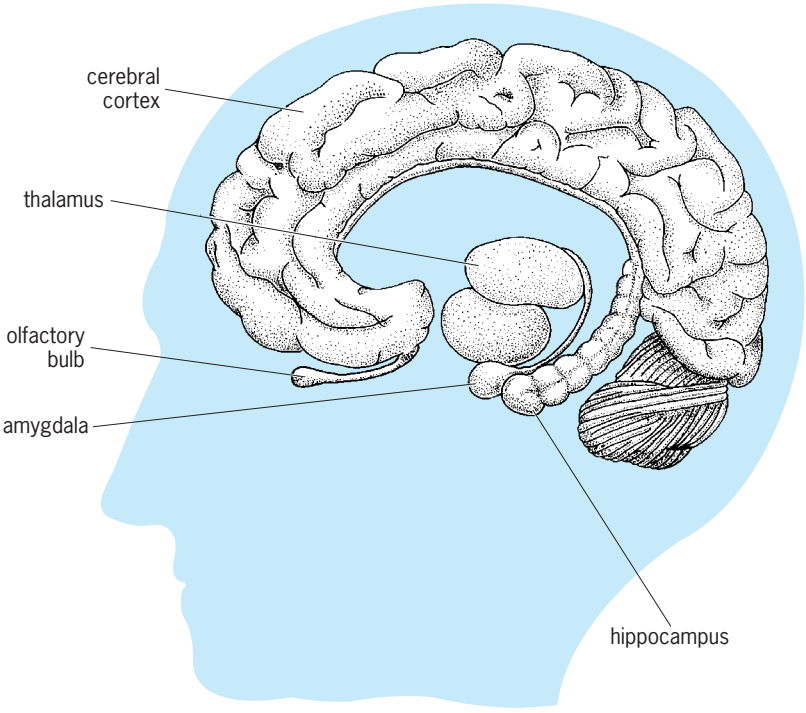

Fig. 2. Selected brain structures associated with memory. (*After C.B. Wortman, E. F. Loftus, and M. E. Marshall, Psychology, 3d ed., Knopf, 1988*)

general principles, because there are many different types of long-term memories.

Visual versus verbal memories. Visual and verbal memories are probably distinct, as suggested by physiological evidence from those who have suffered brain damage. Damage to the thalamus, for example, can result in a loss of the ability to form verbal memories without affecting the formation of visual memories. Visual memories and verbal memories also differ in quality: memory for pictures has been shown to be substantially better than memory for words.

While most people have visual memories, these memories are generally imprecise and are not to be confused with photographic memories—a rare type of memory that is associated with about 5% of young children, and very rarely with adults. People who possess photographic memory, also known as eidetic imagery, report being able to recall remarkable visual details. For example, when remembering a book they claim that they can actually read from the pages in their mind.

Olfactory memory. The unique relationship between memory and smell has long been appreciated, and it is now believed that the relatively direct connection between the olfactory system and the hippocampus, amygdala, and thalamus may account for the unique power of certain odors to elicit long-forgotten memories.

Skill memory. Memories for skills seems to be distinct from memory for general knowledge or events, as has been shown in persons with damage to the amygdala and hippocampus. Even though these individuals cannot generate new long-term memories for events, they still are able to acquire new skills. Skill memory is also different from other forms of long-term memory with respect to the processes of forgetting. Even though memories for most events fade with time, memory for skills, such as riding a bicycle, can last a lifetime.

Forgetting. In addition to investigating the mechanisms affecting memory, researchers have put considerable effort into determining how and why people forget. Forgetting has been associated with three major causes: decay, interference, and motivation.

Decay of memory traces. Decay theories of forgetting assume that memories fade and will gradually be lost if they are not occasionally refreshed. Although this model of forgetting seems intuitively sensible, and is in fact probably the source of some forgetting, there are a number of reasons why it can not fully explain forgetting. First, many memories do not fade even though they have not been recalled for a very long time. A second problem with decay theories of forgetting is that they are very difficult to distinguish from other theories. It is impossible for time to pass without the brain engaging in some type of activity. Consequently, theorists can never be sure whether forgetting is simply due to the passage of time or whether it is due to some activity in that time.

Interference. According to interference theories, forgetting occurs when one memory replaces or becomes confused with another memory. Evidence for interference comes from studies showing that some activities cause greater forgetting than others. Sleeping and being awake, for example, probably involve different levels of interference. If people learn a word list and are then tested 8 h later, they will recall considerably more words if they sleep during those 8 h. Presumably, sleeping reduces the amount of information that people encounter, and consequently reduces the amount of interferences with memory.

Psychologists have identified two different types of interference. Consider the situation of traveling in Europe, where one must learn the same phrases (hello, thank you, how much is the bill) in a number of different languages. If these phrases are learned first in French and later in German, learning them in German will interfere with the ability to recall the earlier learned French words. This situation, in which information that is learned later interferes with information that was learned earlier, is known as retroactive interference. In another type of interference, after having learned a number of phrases in French, it will be more difficult to learn those same words in German. Situations in which old memories inhibit the learning of new memories constitute proactive interference.

Motivated forgetting. The notion that people forget certain things because they want to is a central theme in Freud's psychoanalytic theory. Freud believed that unacceptable or anxiety-provoking thoughts are repressed by the unconscious in order to avoid confronting them directly. Although the processes involved in such motivated forgetting are far from being understood, it is clear that people can forget much about very disturbing events such as violent

crimes or horrendous automobile accidents.

Less dramatic unpleasant experiences are also often forgotten with time. As time passes, people become increasingly more likely to recall the pleasant aspects of an event. In addition, with time people tend to distort various other types of memories in a manner that makes them and their lives appear more worthwhile. For example, people may recall having better-paying jobs, voting more frequently, and donating more to charity than records indicate. Thus, selective forgetting allows memories to be shaped to better fit one's perceptions of the world and of oneself. *See* MOTIVATION.

Permanence. Some widely cited physiological research by W. Penfield was thought to suggest that many seemingly lost memories can be reactivated by direct electrical stimulation. Penfield stimulated particular areas of the brain in patients who were undergoing brain surgery. The patients, who were only under local anesthesia, often reported "reliving" various episodes from earlier times in their life. These observations led Penfield and many other psychologists to conclude that all experiences are kept in memory, and that forgetting simply involves an inability to find a particular memory.

Subsequent consideration of Penfield's findings, however, have shed doubt on this interpretation. Specifically, the experiences that people reported were never verified, so it is not certain that they were really memories at all. The patients might have been experiencing a process like dreaming, involving very lifelike qualities that do not actually correspond to real events. Thus, the question of the permanence of memory is still an open one.

Remembering forgotten memories. While it seems unlikely that everything experienced is recorded in memory, from time to time memories that were throught to be lost forever return. There have been some anecdotal reports that hypnosis can be used as a means of unlocking unavailable memories. However, controlled experimental studies indicate that hypnosis can just as often promote memory fabrications. A more promising technique for recovering difficult-to-remember memories has been developed—cognitive interview. This approach uses four general techniques for jogging memories about specific events such as a witnessed crime, including mentally going back to the scene of the memory; reporting everything, even partial information, regardless of its perceived importance; recounting the events in a variety of orders, such as in reverse; and reporting the event from the perspective of other individuals at the event. The cognitive interview can increase the amount of correct information obtained from witnesses by up to 35% without increasing memory fabrications. *See* LEARNING MECHANISMS.

Jonathan Schooler; Elizabeth F. Loftus

Bibliography. A. Baddeley, *Your Memory: A User's Guide*, 2d ed., 1996; J. Delacour, *Memory System of the Brain*, 1994; R. Fry, *Improve Your Memory*, 4th ed., 2000; D. J. Herrmann, *Memory Improvement Techniques*, 1988; E. F. Loftus, *Memory*, 1980.

Mendelevium

A chemical element, Md, atomic number 101, the twelfth member of the actinide series of elements. Mendelevium does not occur in nature; it was discovered and is prepared by artificial nuclear transmutation of a lighter element. Known isotopes of mendelevium have mass numbers from 248 to 258 and half-lives from a few seconds to about 55 days. They are all produced by charged-particle bombardments of more abundant isotopes. The amounts of mendelevium which are produced and used for studies of chemical and nuclear properties are usually less than about a million atoms; this is of the order of a million times less than a weighable amount. Studies of the chemical properties of mendelevium have been limited to a tracer scale. The behavior of mendelevium in ion-exchange chromatography shows that it exists in aqueous solution primarily in the 3 + oxidation state characteristic of the actinide elements. However, it also has a dipositive (2+) and a monopositive (1+) oxidation state. *See* ACTINIDE ELEMENTS; PERIODIC TABLE; TRANSURANIUM ELEMENTS.

1												13	14	15	16	17	18
1 **H**	2																2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Glenn T. Seaborg

Bibliography. S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; G. T. Seaborg, History of the synthetic actinide elements, *Actinides Rev.*, 1:3–38, 1967; G. T. Seaborg (ed.), *Transuranium Elements: Products of Modern Alchemy*, 1978; G. T. Seaborg and W. D. Loveland, *The Elements Beyond Uranium*, 1990.

Mendelism

Fundamental principles governing the transmission of genetic traits, discovered by an Augustinian monk Gregor Mendel. In 1856, Mendel performed his first set of hybridization experiments with pea plants in the monastery garden at Brno, Moravia. Although the pea plant is normally self-fertilizing, it can be easily crossbred, and grows to maturity in a single season. True breeding strains, each with distinct characteristics, were available from local seed merchants. For his experiments, Mendel chose seven sets of

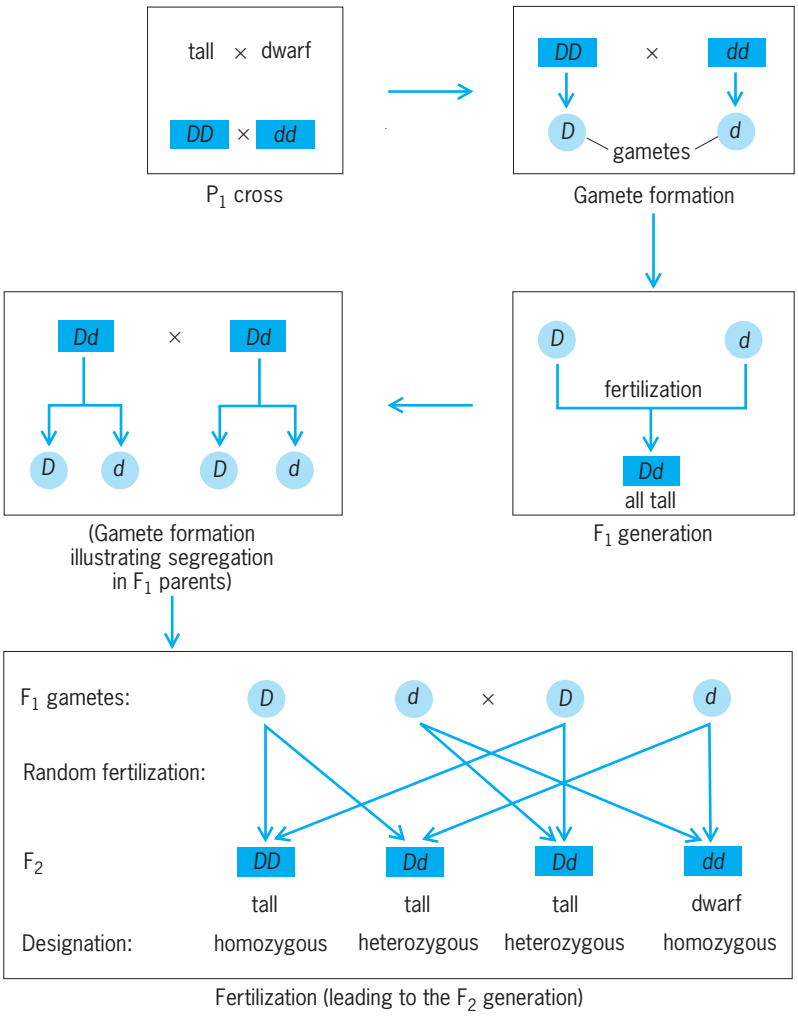

Schematic representation of a monohybrid cross. Pure-bred tall and dwarf strains are crossed, and yield typical 3:1 ratio in the F_2 generation. *D* and *d* represent the tall and dwarf factors (genes), respectively. (*After W. S. Klug and M. R. Cummings, Concepts of Genetics, Charles E. Merrill, 1983*)

contrasting characters or traits. For stem height, the true breeding strains tall (7 ft or 2.1 m) and dwarf (18 in. or 45 cm) were used. He also selected six other sets of traits, involving the shape and color of seeds, pod shape and color and the location of flowers on the plant stem.

The most simple crosses performed by Mendel involved only one pair of traits; each such experiment is known as a monohybrid cross. The plants used as parents in these crosses are known as the P_1 (first parental) generation. When tall and dwarf plants were crossed, the resulting offspring (called the F_1 or first filial generation) were all tall. When members of the F_1 generation were self-crossed, 787 of the resulting 1064 F_2 (second filial generation) plants were tall and 277 were dwarf. The tall trait is expressed in both the F_1 and F_2 generations, while the dwarf trait disappears in the F_1 and reappears in the F_2 generation. The trait expressed in the F_1 generation Mendel called the dominant trait, while the recessive trait is unexpressed in the F_1 but reappears in the F_2. In the F_2, about three-fourths of the offspring are tall and one-fourth are dwarf (a 3:1 ratio). Mendel made similar crosses with plants exhibiting each of the other pairs of traits, and in each case all of the F_1 offspring showed only one of the parental traits and, in the F_2, three-fourths of the plants showed the dominant trait and one-fourth exhibited the recessive trait. In subsequent experiments, Mendel found that the F_2 recessive plants bred true, while among the dominant plants one-third bred true and two-thirds behaved like the F_1 plants. *See* DOMINANCE.

Law of segregation. To explain the results of his monohybrid crosses, Mendel derived several postulates. First, he proposed that each of the traits is controlled by a factor (now called a gene). Since the F_1 tall plants produce both tall and dwarf offspring, they must contain a factor for each, and thus he proposed that each plant contains a pair of factors for each trait. Second, the trait which is expressed in the F_1 generation is controlled by a dominant factor, while the unexpressed trait is controlled by a recessive factor. To prevent the number of factors from being doubled in each generation, Mendel postulated that factors must separate or segregate from each other during gamete formation. Therefore, the F_1 plants can produce two types of gametes, one type containing a factor for tall plants, the other a factor for dwarf plants. At fertilization, the random combination of these gametes can explain the types and ratios of offspring in the F_2 generation (see **illus.**). *See* FERTILIZATION (ANIMAL); GENE.

Independent assortment. Mendel extended his experiments to examine the inheritance of two characters simultaneously. Such a cross, involving two pairs of contrasting traits, is known as a dihybrid cross. For example, Mendel crossed plants with tall stems and round seeds with plants having dwarf stems and wrinkled seeds. The F_1 offspring were all tall and had round seeds. When the F_1 individuals were self-crossed, four types of offspring were produced in the following proportions: 9/16 were tall, round; 3/16 were tall, wrinkled; 3/16 were dwarf, round; and 1/16 were dwarf, wrinkled. On the basis of similar results in other dihybrid crosses, Mendel proposed that during gamete formation, segregating pairs of factors assort independently of one another. As a result of segregation, each gamete receives one member of every pair of factors [this assumes that the factors (genes) are located on different chromosomes]. As a result of independent assortment, all possible combinations of gametes will be found in equal frequency. In other words, during gamete formation, round and wrinkled factors segregate into gametes independently of whether they also contain tall or dwarf factors. *See* GAMETOGENESIS; MEIOSIS.

It might be useful to consider the dihybrid cross as two simultaneous and independent monohybrid crosses. In this case, the predicted F_2 results are 3/4 tall, 1/4 dwarf, and 3/4 round, 1/4 wrinkled. Since the two sets of traits are inherited independently, the number and frequency of phenotypes can be

predicted by combining the two events:

$$3/4 \text{ tall} \begin{cases} 3/4 \text{ round} & (3/4)\,(3/4) \\ & \quad = 9/16 \text{ tall, round} \\ 1/4 \text{ wrinkled} & (3/4)\,(1/4) \\ & \quad = 3/16 \text{ tall, wrinkled} \end{cases}$$

$$1/4 \text{ dwarf} \begin{cases} 3/4 \text{ round} & (1/4)\,(3/4) \\ & \quad = 3/16 \text{ dwarf, round} \\ 1/4 \text{ wrinkled} & (1/4)\,(1/4) \\ & \quad = 1/16 \text{ dwarf, wrinkled} \end{cases}$$

This 9:3:3:1 ratio is known as a dihybrid ratio and is the result of segregation, independent assortment, and random fertilization.

Mendel published his work in 1866, but it went largely unheralded until 1900 when other investigators performed similar hybridization experiments and brought his work to the attention of the scientific world. The rediscovery of his work is regarded as the beginning of genetics as an organized discipline. Since then, genetic crosses with many different organisms have confirmed the fundamental nature and significance of Mendel's work. *See* GENETICS.

Michael R. Cummings

Bibliography. M. R. Cummings, *Human Heredity: Principles and Issues*, 5th ed., 1999; A. R. Hallauer and J. B. Miranda, *Quantitative Genetics and Maize Breeding*, 2d ed., 1988; K. Kim, *Explaining Scientific Consensus: The Case of Mendelian Genetics*, 1994; R. C. Olby, *The Origins of Mendelism*, 2d ed., 1985; G. Poste and S. T. Crooke (eds.), *New Frontiers in the Study of Gene Function*, 1987; T. Sasazuki (ed.), *New Approach to Genetic Diseases*, 1988.

Meninges

In mammals, the three membranes that cover the brain and spinal cord: the dura mater, the arachnoid membrane, and the pia mater.

The outermost, the dura mater, is a tough, fibrous, double-layered structure that is adherent to the skull. The inner layer of the dura mater sends separating sheets between the cerebral hemispheres and between the cerebrum and cerebellum. It also contains large venous sinuses and forms sheaths for nerves leaving the skull. The middle layer, the arachnoid, is a delicate serous layer loosely investing the brain. Below this is the spongy subarachnoid cavity which contains the circulating cerebrospinal fluid. The innermost layer, the pia mater, is a vascular layer which closely follows each convolution of the brain. Together the meninges furnish protection, blood supply, drainage, and cerebrospinal channels for the brain. The meninges around the spinal cord are less complex but basically similar to and continuous with those of the brain.

In submammalian forms, there may be only a single membrane, the meninx primitiva, or two membranes, the dura mater and pia-arachnoid. In forms in which the brain does not fill the cranial cavity, the volume of cerebrospinal fluid and of these membranes is increased. *See* NERVOUS SYSTEM (VERTEBRATE).

Thomas S. Parsons

Meningitis

Inflammation of the meninges. Diagnosis of meningitis is made from examination of a specimen of cerebrospinal fluid that is found in the subarachnoid space. The fluid sample is obtained by lumbar puncture (spinal tap) or, less commonly, by a ventricular tap, which requires passage of a needle into the brain tissue. Certain types of meningitis are associated with distinctive abnormalities in the cerebrospinal fluid. With certain types of meningitis, especially bacterial, the causative organism can usually be recovered from the fluid. *See* MENINGES.

Causes and types. Meningeal inflammation in most cases is caused by invasion of the cerebrospinal fluid by an infectious organism. Noninfectious causes also occur. For example, in immune-mediated disorders antigen-antibody reactions can cause meningeal inflammation. Other noninfectious causes of meningitis are the introduction into the cerebrospinal fluid of foreign substances such as alcohol, detergents, chemotherapeutic agents, or contrast agents used in some radiologic imaging procedures. Meningeal inflammation brought about by such foreign irritants is called chemical meningitis; it sometimes is associated with brain involvement. Inflammation also can occur when cholesterol-containing fluid or lipid-laden material leaks into the cerebrospinal fluid from some intracranial tumors.

The most common and most life-threatening types of meningitis are those caused by infectious organisms, usually bacterial or viral, and less frequently fungal, parasitic, rickettsial, or spirochetal. One type of meningitis is differentiated from another on the basis of symptoms and signs, epidemiologic aspects of the illness, and laboratory examinations, especially of blood and cerebrospinal fluid specimens.

Bacterial meningitis. Bacterial meningitis is among the most feared of human infectious diseases because of its possible seriousness, its rapid progression, its potential for causing severe brain damage, and its frequency of occurrence.

Most cases of bacterial meningitis have an acute onset. Common clinical manifestations are fever, headache, vomiting, stiffness of the neck, confusion, seizures, lethargy, and coma. Symptoms of brain dysfunction are caused by transmission of toxic materials from the infected cerebrospinal fluid into brain tissue and the disruption of arterial perfusion and venous drainage from the brain because of blood vessel inflammation. These factors also provoke cerebral swelling, which increases intracranial pressure and contributes to the symptoms of the illness. Before antibiotics became available, bacterial meningitis was almost invariably fatal.

Most types of acute bacterial meningitis are septic-borne in that they originate when bacteria in the

bloodstream (bacteremia, septicemia) gain entrance into the cerebrospinal fluid. Meningitis arising by this route is called primary bacterial meningitis. Secondary meningitis is that which develops following direct entry of bacteria into the central nervous system, which can occur at the time of neurosurgery, in association with trauma, or through an abnormal communication between the external environment and the cerebrospinal fluid.

Different age groups in childhood are predisposed to meningitis caused by certain specific bacteria. Also, in both children and adults, preexisting medical illnesses tend to be associated with meningitis caused by specific types of organisms. In the neonate, *Escherichia coli* and group B *Streptococcus* are the most common causes of bacterial meningitis. *Haemophilus influenzae* once accounted for most cases of acute bacterial meningitis between 3 months and 3 years of age. However, there is a vaccine that can be given in early infancy and therefore invasive infections caused by *H. influenzae* have become much less common. Other common causes of septic-borne meningitis beyond 2 months of age and including adults include *S. pneumoniae* and *Neisseria meningitidis*. Anaerobic bacteria are infrequent causes of meningitis but can complicate chronic illnesses such as dental sepsis, chronic sinusitis or otitis, and intra-abdominal suppurative diseases. *See* HAEMOPHILUS; MENINGOCOCCUS; PNEUMOCOCCUS; STAPHYLOCOCCUS; STREPTOCOCCUS; TUBERCULOSIS.

Other types. Many viruses can cause meningeal inflammation, a condition referred to as viral aseptic meningitis. The most common viral causes include the enteroviruses, the various herpesviruses, viruses transmitted by arthropods, the human immuno- deficiency virus type I (HIV-1), and formerly, the mumps virus. If the virus attacks mainly the brain rather than the spinal cord, the disorder is termed viral encephalitis. *See* ANIMAL VIRUS; ARBOVIRAL ENCEPHALITIDES; ENTEROVIRUS; HERPES.

Fungal, parasitic, and rickettsial meningitis are less common in the United States than are bacterial and viral. These infections are more likely to be subacute or chronic than those caused by bacteria or viruses; in most cases, the meningeal inflammation is associated with brain involvement. An acute form of aseptic meningitis can occur in the spirochetal diseases, syphilis and Lyme disease. *See* LYME DISEASE; MEDICAL MYCOLOGY; MEDICAL PARASITOLOGY; RICKETTSIOSES; SYPHILIS.

AIDS. An acute inflammatory meningeal syndrome can result either from HIV-1 itself or from various organisms which cause opportunistic infections in individuals with acquired immune deficiency syndrome (AIDS). The HIV-1 organism is believed to enter the brain and cerebrospinal fluid soon after the infection is acquired. The condition initially remains asymptomatic, but in 1–2% of cases it provokes a self-limited acute aseptic meningitis. This is characterized by headaches, malaise, fever, and signs of meningeal irritation, and can persist for several weeks. Among opportunistic infections in AIDS causing meningitis, the most common is of fungal origin, cryptococcosis. Syphilis is not considered to be an opportunistic infection in AIDS cases, but when it occurs, there is a higher incidence of acute syphilitic meningitis than occurs in non-AIDS individuals with syphilis. Symptoms include persistent headache and neck stiffness as well as a common occurrence of cranial nerve dysfunction, especially involving the facial and the acoustic nerves. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS).

Diagnosis and treatment. A diagnosis of bacterial meningitis is made on the basis of cerebrospinal fluid abnormalities, such as the changes in the number and type of cells present and a decrease in the glucose content of the cerebrospinal fluid, as well as recovery of the causative organism from the fluid. Treatment is with high-dose intravenous antibiotics, chosen according to the suspected or proven offending organism. *See* ANTIBIOTIC; MEDICAL BACTERIOLOGY.

William E. Bell

Meningococcus

A major human pathogen belonging to the bacterial genus *Neisseria*, and the cause of meningococcal meningitis and meningococcemia. The official designation is *N. meningitidis*.

Characteristics and serological types. The meningococcus is a gram-negative, aerobic, nonmotile diplococcus, characteristically occurring in smears from clinical material in pairs with flattening of the adjoining sides. It is fastidious in its growth requirements and is very susceptible to adverse physical and chemical conditions. In appearance and biochemical activity, it resembles the gonococcus. The two organisms are most readily distinguished by the ability of the meningococcus to oxidize both glucose and maltose, while the gonococcus oxidizes glucose alone.

Thirteen serogroups of *N. meningitidis* can be distinguished serologically: A, B, C, D, X, Y, Z, 29E, W135, H, I, J, and L. Grouping is based on the antigenic differences between the capsular polysaccharides and is usually accomplished by agglutination reactions. Almost all strains isolated from blood or cerebrospinal fluid of cases can be serogrouped. However, nongroupable (nonencapsulated) strains are frequently isolated from the nasopharynx of healthy carriers. Within each serogroup, serotypes and subtypes can be identified on the basis of antigenic differences between the major outer-membrane proteins. Serotypes are identified by means of monoclonal antibodies. Serotype differences can also be detected among lipopolysaccharide antigens. The eight lipopolysaccharide serotypes are independent of the outer-membrane protein serotypes. Strain differentiation can also be based on genetic relatedness between strains, which is determined by means of multilocus electrophoresis. Differences in the electrophoretic mobility of cytoplasmic enzymes are used to identify allelic

variation in the corresponding genes and enable the recognition of electrophoretic types. *See* SEROLOGY.

Epidemiology. Humans are the only known natural host of the meningococcus. Transmission occurs by droplets directly from person to person. The most frequent site of infection is, asymptomatically, the nasopharynx. This is uncommon in infants and young children but is relatively frequent in adolescents and adults. Fifty percent of cases occur in children 2 months to 4 years, and 5% in children 5 to 14 years of age.

Most sporadic cases are caused by serogroups B and C, with disease to other serogroups being much less common. Almost all large epidemics and pandemics of meningococcal disease are caused by serogroup A. However, the largest known epidemic in a metropolitan area was caused in São Paulo, Brazil, by a serogroup C strain in the early 1970s. Large serogroup A and C outbreaks are associated with poverty and overcrowding. *See* EPIDEMIOLOGY.

Pathogenesis and disease. Susceptibility to bloodstream invasion from the nasopharynx has been shown to correlate with the absence of bactericidal antibody. The development of protective, natural antibody appears to result from acquisition of bacteria other than meningococci which possess antigens that cross-react with the serogroup- and serotype-specific surface antigens of the meningococcus. The most common clinical syndrome caused by the meningococcus is meningitis, which is characterized by fever, headache, nausea, vomiting, and neck stiffness and has a fatality rate of 15% (higher in infants and adults over 60). Disturbance of the state of consciousness quickly occurs, leading to stupor and coma. Many cases also have a typical skin rash. In 10–20% of cases caused by serogroups B and C but less often with serogroup A, a much more fulminant disease occurs in which overwhelming meningococcemia develops. As a result of the rapid release of large quantities of endotoxin, profound shock and hemorrhagic phenomena predominate. Death occurs within 6–24 h in most cases despite available treatment.

Occasionally a pneumonia or a transient febrile illness with self-limited meningococcemia occurs. The bacteremia frequently clears without antibiotic treatment.

Chronic meningococcemia is a rare manifestation, appearing most often in adults. This illness is characterized by recurrent episodes of fever, arthritis, myalgia, and rash. Meningococci are present in the blood only during febrile episodes.

Because of the widespread occurrence of strains of meningococci resistant to sulfonamides, penicillin G is the drug of choice for all forms of meningococcal disease. Penicillin-resistant strains of meningococci have been reported from many parts of the world and have become quite common in some European countries. However, penicillin resistance is still uncommon in North America. A third-generation cephalosporin is the most appropriate alternative for penicillin-resistant strains. Chloramphenicol is an effective alternative in individuals allergic to penicillin.

Prevention. In closed population groups and among household contacts of cases (especially children), there is a significantly increased risk of secondary cases. Antibiotics are recommended for close contacts of cases in order to interrupt transmission by eradicating the carrier state. Because of sulfonamide resistance, sulfonamides cannot be relied upon.

Vaccines against group A, C, Y, and W135 meningococci have been licensed and approved for use in certain circumstances. The vaccines, composed of purified capsular polysaccharides, are very safe and effective in preventing group-specific disease. Routine immunization of children is not recommended because of the limited immunogenicity in infants and the relatively short duration of protection. Routine immunization is used for military recruits in order to prevent meningococcal outbreaks during basic training. The vaccines are also recommended for epidemic control, close contacts of cases, and travelers to areas experiencing epidemics. *See* MEDICAL BACTERIOLOGY.

Ronald Gold

Menopause

The irreversible cessation of regular monthly uterine bleeding in the adult human female, marking the end of her ability to become pregnant. Menopause commonly occurs in the United States between the ages of 47 and 53. It probably occurs because the ovary runs out of eggs and the cyclic rise and fall of brain and ovarian hormones designed to prepare the uterus to receive and nourish pregnancy no longer occur.

Menopause is one event in the climacteric, the period of time during which the reproductive machinery slows down and finally stops. The biochemical hallmark of this period is a reduction in estrogen production by the ovary. Some estrogen continues to be produced by the adrenal gland and the fatty tissues throughout the body, but this amount is very small compared with premenopausal levels. Estrogen has widespread effects on both genital and extragenital systems, and the withdrawal of estrogen accounts for many of the signs and symptoms attributed to the menopause, although these are influenced by both hereditary and social factors.

Postmenopausal changes in the genital organs include atrophy of external structures, including loss of glandular tissue in the breasts as well as shrinkage of fat depots (adipose tissue) throughout the body, which results in wrinkling of overlying skin. The lining of the vagina and urethra becomes thin, easily traumatized, and susceptible to infection. Itching and burning of the vagina, as well as urinary urgency, increased frequency of voiding with small amounts of urine, night voiding, and urine loss upon coughing or laughing may result.

Withdrawal of estrogen may affect the cardiovascular system, with a rise in blood lipids that is

dangerous and with a greater risk of heart disease. Hot flashes and night sweats can occur because of hypothalamic instability of temperature regulation brought on by a lack of estrogen. The skeletal response to withdrawal of estrogen may include loss of both minerals and collagen matrix from bone. The resulting osteoporotic bone is thin, weak, and subject to fractures from minor trauma. *See* OSTEOPOROSIS.

Many psychological problems have been attributed to estrogen deprivation, but well-documented proof of those relationships is lacking. Appearance of the symptoms is probably the result of a loss of self-esteem brought about by the combination of physical changes, loss of reproductive capability, and awareness of aging.

While estrogen can reverse or halt many of the physical changes described, it will not prevent aging or restore reproductive ability. In addition, high-dose, long-term, unopposed estrogen use has been correlated with an increased risk of cancer of the uterus and possibly a slight increase in the risk of breast cancer. Estrogen alone is usually not given to women with a uterus: progesterone is added either continuously at a very low dose or intermittently, usually for 14 days, at a higher dose. Treatment of menopausal symptoms should be undertaken on an individual basis, with careful discussion of the risks and benefits currently known. *See* ESTROGEN; MENSTRUATION. G. Guzinski

Bibliography. H. Judd and W. H. Utian, Current perspectives in the management of the menopausal and postmenopausal patient, *Amer. J. Obstet. Gynecol.*, 156:1279–1356, 1987.

Menstruation

Periodic sloughing of the uterine lining in women of reproductive age. Menstrual bleeding indicates the first day of the menstrual cycle, which lasts an average of 27–30 days, although ranges of 21–60 days have been recorded.

Menarche, the onset of menstruation, occurs between the ages of 9 and 16. The majority of females begin menstruating at ages 12–14. During the first few years, the duration and intensity of menstrual flow and the total cycle length may be quite variable, but regularity is gradually established.

Menstrual cycle. The menstrual cycle consists of cyclic changes in both the ovary and the uterus (**Fig. 1**). These changes are controlled by the interaction of several hormones. These include follicle-stimulating hormone (FSH) and luteinizing hormone (LH), which are secreted by the anterior pituitary, and the steroid hormones estrogen and progesterone, which are secreted by follicles in the ovary. At the beginning of the cycle, the follicle is stimulated by FSH. In response, it grows and secretes estrogen. The amount of estrogen secretion increases rapidly near the middle of the cycle. Estrogen, in turn, stimulates growth of the uterine lining (mucosa), which becomes thicker and fills with blood vessels. In midcycle, the rapid increase in estrogen causes a massive surge of LH release and a smaller release of FSH from the pituitary. This surge causes ovulation, which is the release of the ovum (egg) from the follicle. After ovulation, the follicle undergoes rapid changes and is then called a corpus luteum. The corpus luteum secretes progesterone in response to LH stimulation. Progesterone and estrogen together cause a further thickening of the uterine mucosa, preparing the uterus for pregnancy. If pregnancy does not occur, the corpus luteum degenerates, the uterine mucosa sloughs off, and the cycle begins again (**Fig. 2**).

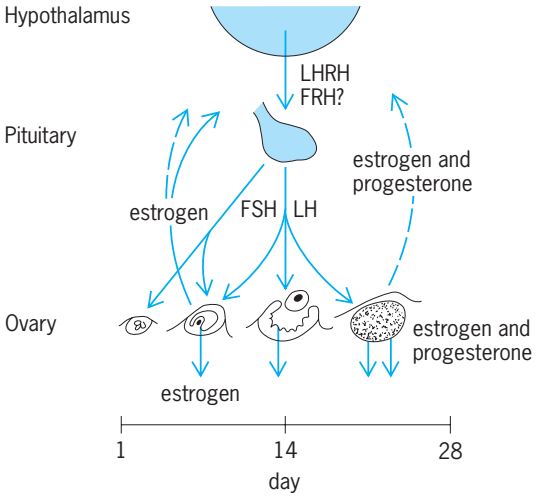

Fig. 1. Hormonal changes during the menstrual cycle. The broken arrows indicate inhibitory effects, and the solid lines, stimulatory effects. LHRH (luteinizing hormone–releasing hormone) and FRH (follice-releasing hormone) are released by the hypothalamus and stimulate the pituitary.

During the menstrual cycle, the uterine mucosa becomes richly supplied with small blood vessels and microscopic glands. When the corpus luteum degenerates, it produces less progesterone. This decline in progesterone at the end of the menstrual cycle causes constriction of the spiral arteries in the uterine lining, restricting blood flow to the mucosa. Without blood flow, the cells in the mucosa die and are sloughed away, causing menstrual bleeding. This menstrual bleeding, or menstrual flow, consists of small amounts of blood and uterine tissue. The microscopic glands in the mucosa produce chemicals called prostaglandins, which promote sloughing and may also cause uterine contractions. Menstrual flow ends when the spiral arteries constrict, and then a new mucosa develops as the cycle begins again. The menstrual flow (menses) lasts 3–7 days, with an average of 5 days, and a loss of approximately 30 ml (roughly 1.5 oz) of blood.

Amenorrhea. There is no menstrual bleeding during pregnancy, as the uterine mucosa is needed for the maintenance of pregnancy. This amenorrhea, or lack of normal ovarian function, sometimes continues during nursing. During nursing (lactation),

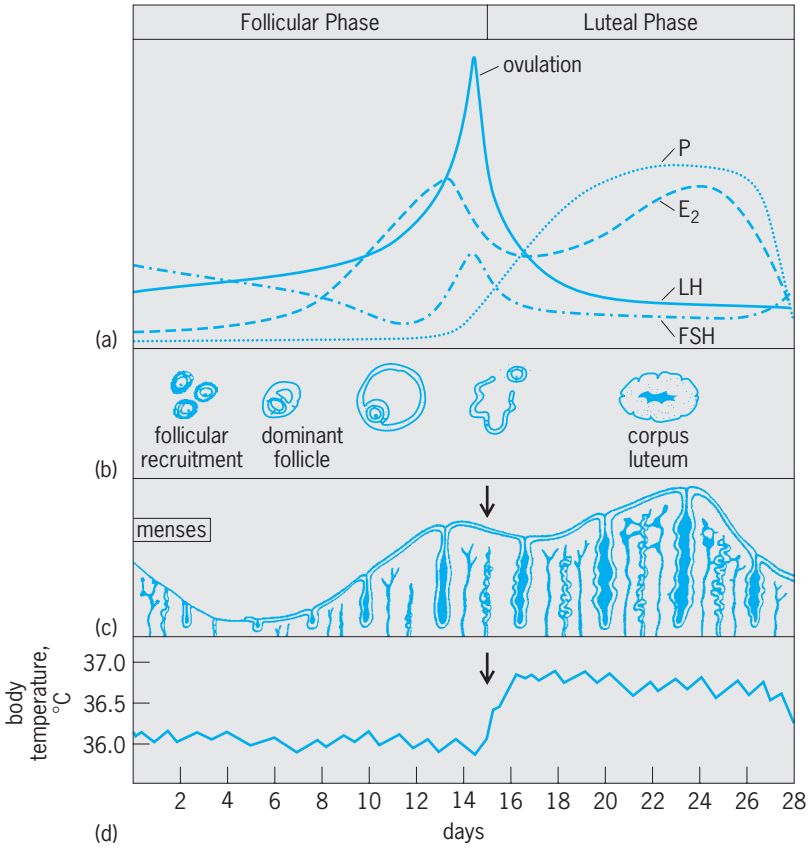

Fig. 2. Hormonal, ovarian, endometrial, and basal body temperature changes and relations throughout the normal menstrual cycle.

there is suppression of gonadotropin-releasing hormone (GnRH), which is needed for the LH surge that causes ovulation. This suppression is probably caused by prolactin, a hormone involved in milk production. Lactational amenorrhea generally lasts for 4–6 months after a woman gives birth. After this, however, ovulation may occur, and conception is possible, even before menses resumes.

Premenstrual tension syndrome. Physical changes before menstruation include a drop in body temperature, retention of fluids which may lead to swelling of the legs and abdomen, tenderness of the breasts, weight gain, and increased skin blemishes. The most common treatment for these symptoms is diuretics, drugs which promote fluid loss through increased urinary volume. Psychological changes may also occur before menstruation, including depression, fatigue, and increased irritability. Up to 40% of women have one or two of these symptoms, whereas approximately 5% have multiple or severe symptoms. The precise cause of premenstrual tension syndrome is not known, but the change in the ratio of estrogen to progesterone caused by falling progesterone levels is the suspected cause.

Menopause. Cessation of menses, or menopause, occurs at an average age of 51, with a range of 42–60 years. Prior to the onset of menopause, secretion of the pituitary hormones LH and particularly FSH increases, and the loss of follicles increases. After menopause, there is a decrease in estrogen production. These alterations of hormone levels cause various physical and sometimes emotional changes. *See* MENOPAUSE.

Janice M. Bahr

Mental retardation

A developmental disability characterized by significantly subaverage general intellectual functioning, with concurrent deficits in adaptive behavior. The causes are many and include both genetic and environmental factors as well as interactions between the two.

Diagnosis. Because the vast majority of persons with mental retardation have mild cognitive deficits (IQ 55 to 70), in most cases the diagnosis is not formally made until children have entered into school settings. School-age children with mental retardation manifest a delayed rate of learning, and may display frustration with the task demands that come with school. In the preschool years, the diagnosis is more likely to be established by evidence of delayed maturation in the areas of sensory-motor, adaptive, cognitive, social, and verbal behaviors. By definition, evidence of mental retardation must exist prior to adulthood, where vocational limitation may be evident, but the need for supervision or support may persist beyond the usual age of social emancipation.

Classification. Mental retardation can be divided into subgroups on the basis of etiology, degree of manifest intellectual retardation, and the need for supports. From the aspect of etiology, mental retardation can be classified by prenatal, perinatal, or postnatal onset. Prenatal causes include genetic disorders (inborn errors of metabolism, excess gene dosage), syndromal disorders (for example, conditions with a likely but as yet unknown genetic cause, or syndromes which result from in utero exposure to teratogens), and developmental disorders of brain formation. Upward of 700 genetic causes have been suggested as associated with the development of mental retardation. Additionally, many environmental influences on the developing fetus, for example, infection, and other unknown errors of development may account for mental retardation.

Perinatal causes include complications at birth, extreme prematurity, infections, and other neonatal disorders. Postnatal causes include trauma, infections, demyelinating and degenerative disorders, consequences of seizure disorders, toxic-metabolic disorders, malnutrition, and environmental deprivation. Even employing this broad classification scheme, often no specific cause can be identified to account for the mental retardation of a particular individual.

Subaverage intellectual functioning. Individuals with mental retardation are typically subclassified in terms of the manifest severity of cognitive disability as reflected by the ratio of mental age to chronological age, or intelligence quotient (IQ). Subaverage intellectual functioning is defined as an IQ score of at least two standard deviations below the mean, or approximately 70 to 75 or below. IQ tests must be individually administered and take into account an individual's educational and sociocultural background. Mild, moderate, severe, and profound degrees of mental retardation refer to two, three, four, or five standard deviations below the normal IQ for the general population.

Limitations in adaptive behavior, as assessed by standardized tests, must also be demonstrable in order to satisfy diagnostic criteria for mental retardation. This criterion is important because certain artistic or other gifts may not be revealed on formal IQ testing, and different levels of learning difficulty may be accentuated by the demands of specific environments. Outside such environments, an individual may navigate a normal course in life.

Comorbidity is frequently the rule in developmental disorders. A specific genetic or other cause of mental retardation may also predispose to other medical or neurologic conditions. Thus, children with trisomy 21 (Down syndrome) are at increased risk for cardiac malformation and thyroid abnormalities. Children with tuberous sclerosis complex (an inherited hardening or fibrosis of brain tissue resulting in mental defect) may have epilepsy and renal involvement. Children with fetal alcohol syndrome often have strabismus (incoordinated action of the muscles of the eyeball). In these circumstances, the comorbid medical conditions may increase the likelihood of emotional or behavioral problems, or contribute to the challenges with which a given child must contend in order to acquire new knowledge and new skills. In addition to medical comorbidity, researchers are increasingly recognizing certain patterns of behavior which associate with specific causes of mental retardation. In Prader-Willi syndrome, children develop an intense preoccupation with food and will eat to morbid obesity. Tantrums, aggression, and self-injury may occur when limits are set surrounding food consumption. In Lesch-Nyhan syndrome, children will compulsively self-injure, typically biting themselves to the point that protective restraint becomes necessary. In Velo-Cardio-Facial syndrome, children may present with extremely labile mood and impulsive behaviors.

Thus, the identification of cause can be important in planning for the medical, educational, and treatment needs of a particular individual.

Prevention and treatment. Over the past few decades, as the understanding of the mechanisms underlying mental retardation has grown, considerable progress has been made in both prevention and treatment. Diet is a method of treatment following early detection of phenylketonuria; warnings regarding alcohol consumption during pregnancy, lead exposure in infancy, and disease immunization and therapy are measures for prevention of retardation. Advances in prenatal, obstetrical, and neonatal care and genetic counseling have had the effect of reducing the incidence or the severity of various conditions. Energetic training and the application of psychosocial techniques (for example, early intervention programs) have resulted in improved social performance and adaptive behavior in many persons with mental retardation.

Bryan H. King

Bibliography. J. C. Harris, *Developmental Neuropsychiatry*, vols. 1 and 2, Oxford University Press, 1995; R. M. Hodapp, Development and disabilities: Mental, motor, and sensory impairments, Cambridge University Press, 1998; B. H. King et al., Mental retardation: A review of the past 10 years, Part I, *J. Amer. Acad. Child Adolescent Psychiat.*, 36:1656–1663, 1997; R. Luckason et al., *Mental Retardation: Definition, Classification, and System of Supports*, American Association on Mental Retardation, 1992; J. L. Matson and J. A. Mulick (eds.), *Manual of Diagnosis and Professional Practice in Mental Retardation*, American Psychological Association, 1996; M. State, B. H. King, and E. Dykens, Mental retardation: A review of the past 10 years, Part II, *J. Amer. Acad. Child Adolescent Psychiat.*, 36:1664–1671, 1997.

Menthol

A cyclic monoterpene alcohol (**1**) with melting point 40°C (104°F), specific optical rotation ($[\alpha]_D$) −50°. Menthol is the major component of the essential oil of peppermint. Since menthol has three chiral centers (asymmetric carbons), structure (**1**) is one of eight stereoisomers. These isomers can be grouped in two sets of four diastereoisomers. In one set the

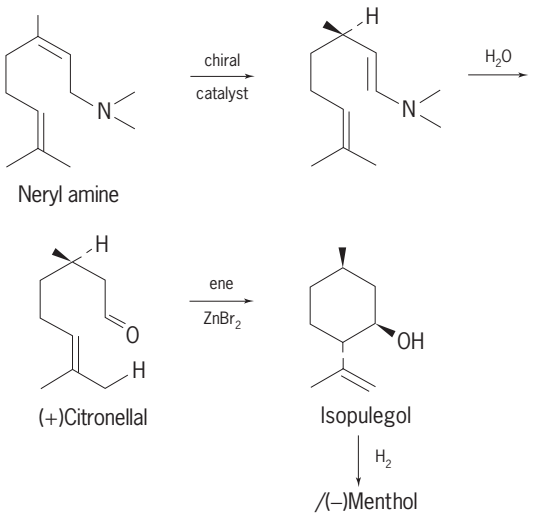

Enantioselective synthesis of /(-) menthol.

methyl group (CH_3) at C-5 is in the *R* configuration, as in structures (**1**)–(**4**). The other set of stereoisomers consists of the enantiomers of (**1**)–(**4**), with opposite configuration at each asymmetric center. *See* POLARIMETRIC ANALYSIS; STEREOCHEMISTRY.

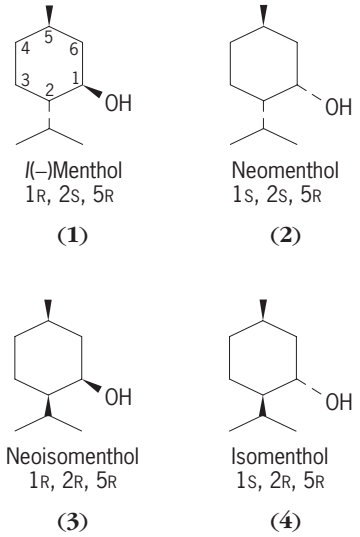

Menthol undergoes esterification, oxidation, and dehydration reactions typical of secondary alcohols. By oxidation to menthone, enolization at C-2, and reduction, an equilibrium mixture of the four diastereoisomers can be obtained. *l*(−)Menthol, in which all three substituents can be in an equatorial conformation, is the most stable of the diastereoisomers (60% at equilibrium). *See* TAUTOMERISM.

l(−)Menthol is produced in large amounts (about 3000 tons per year in 1980) from the oil of Japanese or Brazilian peppermint (*Mentha arvensis*). These oils contain about 80% menthol, much of which can be isolated by direct crystallization from the chilled oil. Neomenthol is a minor constituent.

Syntheses based on nonchiral starting materials such as thymol or citral have been used for the manufacture of racemic menthol [(**1**) and its enantiomer]. *l*(−)Menthol can be prepared by cyclization of (+) citronellal (from citronella oil) in an ene reaction (a pericyclic reaction) followed by hydrogenation. A method has been developed for the enantioselective synthesis of (+)citronellal with a chiral catalyst from an achiral starting material. This process is operating on an industrial scale (see **illus.**). *See* PERICYCLIC REACTION.

Menthol has a cooling, slightly anesthetic effect and is used in pharmaceutical preparations as an antitussive and antipruritic agent, and as a decongestant in nasal sprays. Most of the menthol produced is used in dentrifices, mouthwash, skin lotions, and so forth, and as a flavorant in tobacco, chewing gum, and confectionary products. For most of these purposes, the *l*(−)isomer is preferred. *See* TERPENE. James A. Moore

Mercaptan

One of a group of organosulfur compounds which are also called thiols or thio alcohols and which have the general structure RSH. Aromatic thiols are called thiophenols, and biochemists often refer to thiols as sulfhydryl compounds. The unpleasant odor of volatile thiols causes them to be classed as stenches, but the odors of many solid thiols are not unpleasant.

Mercaptans (1) form salts with bases, (2) are easily oxidized to disulfides and higher oxidation products such as sulfonic acids, (3) react with chlorine (or bromine) to form sulfenyl chlorides (or bromides), and (4) undergo additions to unsaturated compounds, such as olefins, acetylenes, aldehydes, and ketones. The insoluble mercury salts (mercaptides) are used to isolate and identify mercaptans.

The important amino acid cysteine and many pharmaceutical and industrial products contain the mercapto group. Thus, thiosalicylic acid is used in synthesizing the germicide merthiolate; 2-mercaptobenzothiazole is a rubber vulcanization accelerator; 2,3-dimercaptopropanol (British antilewisite) is an antidote for arsenic poisoning; and 6-mercaptopurine is of interest in cancer chemotherapy. The occurrence and removal of mercaptans from petroleum is industrially important. Even the odor of mercaptans finds use; traces of mercaptans added to dangerous gases act as warning agents in case of leaks.

The name thio alcohols suggests that mercaptans are similar to alcohols. Although some properties are analogous, there are decided differences, related to (1) the greater acidities of thiols, (2) the ease with which mercaptans are oxidized, and (3) the ability of mercaptans to enter free-radical reactions. The second difference probably accounts for the absence of thiols, as such, in contrast to alcohols, in nature. *See* ORGANOSULFUR COMPOUND; PETROLEUM PROCESSING AND REFINING; RUBBER. Norman Kharasch

Bibliography. I. L. Finar, *Organic Chemistry: The Fundamental Principles*, 6th ed., 1986; R. T. Morrison and R. N. Boyd, *Organic Chemistry*, 5th ed., 1987; S. T. Preston and R. Pankratz, *A Guide*

to the Analysis of Thioalcohols and Thioethers (Mercaptans and Alkyl Sulfides) by Gas Chromatography, rev. ed., 1981.

Merchant ship

A power-driven ship employed in commercial transport on the oceans and large inland bodies of water such as the Great Lakes. This article treats design considerations of merchant ships and includes detailed information on 10 merchant ship types. The relatively small craft used for inland waterway transportation are not commonly referred to as ships. For a description of such craft *see* INLAND WATERWAYS TRANSPORTATION.

Commodities transported by water are classified as break-bulk, unitized, or bulk (dry or liquid) cargoes. Generally, water cargo transportation is cheaper per ton-mile than land or air transportation. Although the volume of air transportation has increased considerably, approximately 90% of the United States overseas trade revenue is waterborne, while the rest is airborne. The specific cost of transporting break-bulk cargo by air is many times that of ocean transportation. Oceanic passenger transportation is vanishing, except for recreation purposes such as cruises.

Break-bulk cargo includes miscellaneous goods packed in boxes, bales, crates, cases, bags, cartons, barrels, or drums. It may also include lumber, motor vehicles, pipe, steel, and machinery. Vessels engaged pimarily in the transportation of break-bulk cargo are called break-bulk cargo ships (sometimes also called general cargo ships). Partial loadings of refrigerated cargo may be carried in break-bulk cargo ships in specially insulated and refrigerated holds. A ship which primarily carries refrigerated cargo is often called a reefer.

Under the heading of unitized cargo may be grouped cargo carried aboard in pallets, containers, wheeled vehicles (railway cars, trailers, and the like), and barges or lighters. There are pallets of several sizes, from 24 × 32 in. (61 × 81 cm) to 88 × 108 in. (224 × 274 cm), when rectangular, and from 36- to 48-in. (91- to 122-cm) squares. Cargo is stowed on pallets to various heights and secured by steel strapping. Usually, pallets are handled on board by fork-lift trucks. Pallet ships are those used exclusively for the carriage of pallets. There are also several sizes of containers, typically 8 × 8 ft (2.4 × 2.4 m) in cross section and 20 or 40 ft (6 or 12 m) in length.

Containers are built of aluminum, plywood, or steel. Generally, containers are nested in cells underdeck up to nine high in cargo holds, and stacked on deck on top of hatch covers not more than four high. Container ships are called full container ships when all their cargo is carried in containers, and these are almost exclusively handled by shoreside container-handling cranes. Some ships, such as break-bulk cargo ships, carry a partial load of containers. *See* MARINE CONTAINERS.

Cargo-carrying wheeled vehicles are rolled on and off the ship on side or end ramps and stowed on and secured to decks. Ships thus loaded are called roll-on/roll-off ships. In some special cases, wheeled vehicles are lowered or lifted vertically by means of cranes. Although ferryboats may transport cargo-carrying wheeled vehicles, their main purpose is to transport passengers or their vehicles or both.

Barges, or lighters, are handled by special shipboard gantry cranes, elevators, or rarely, heavy lift booms. When the main purpose is the carriage of barges, vessels are called barge-carrying ships. Unitized cargo consists of miscellaneous goods, such as those listed under break-bulk cargo, plus refrigerated cargo carried in insulated and refrigerated containers or trailers. There is a strong tendency toward cargo unitization, especially containerization, because of its inherent advantages, such as faster ship turnaround, interchangeability among different modes of transportation, door-to-door delivery from manufacturer to consumer, and reduction of cargo damage and pilferage.

Examples of bulk dry cargo are ore, coal, sugar, cement, and grain. These items are poured or otherwise loaded into the ship's cargo compartments (holds) without being boxed, bagged, or hand-stowed. In smaller quantities this cargo is also transported in break-bulk cargo ships. Vessels designed specifically for the ore or coal trade are referred to as ore carriers and colliers.

The principal bulk liquid cargoes are petroleum and its by-products. Others include liquefied gases, liquid chemicals, wine, fruit juices, and molten asphalt or sulfur. These liquids are transported in large tank spaces that are integral with, or separated from, the ship's main structure. Ships carrying bulk petroleum products are called tankers. In smaller quantities some bulk liquid cargoes are also transported in deep tanks in other vessels, such as break-bulk cargo ships.

A passenger ship, as defined by International Safety of Life at Sea (SOLAS) rules, carries more than 12 passengers on international voyages. Passenger ships navigating solely on inland waters are excepted from these rules. Passenger vessels that also transport cargo are called passenger-cargo ships.

Design Considerations

With very few exceptions, merchant ships conform to classification society rules. Classification societies are private organizations which issue rules for the construction, equipment, and maintenance of merchant ships. Compliance with their rules assures owners and insurers of the vessel's strength and seaworthiness.

Dimensions. Classification societies define the length between perpendiculars L_{pp} as approximately the distance from the bow to the rudder, but not less than 96% of the load waterline length (the immersed hull length when floating at the maximum permissible draft). The breadth of seagoing cargo ships 425 ft (130 m) long and over is approximately equal to 15% of the length, while smaller ships have relatively larger breadth. The depth of seagoing merchant ships from the keel to the strength deck

(usually the uppermost continuous deck) varies from approximately one-tenth to one-fourteenth the L_{pp}. For Great Lakes vessels, which operate under less severe conditions, the length-to-depth ratio may be as high as 18:1. Also, the length-to-breadth ratio is generally greater than that of oceangoing ships. For the same cargo capacity and speed, a relatively long, slender vessel, such as is found on the Great Lakes, is more costly to construct, but requires slightly less propulsive power and in bad weather need not slow down as much as a shorter vessel to avoid damage. *See* SHIP DESIGN; SHIPBUILDING.

The maximum permissible draft (submerged depth of ship) is limited by international load line regulations. The draft to which cargo ships can be loaded depends principally upon the depth of the ship below the freeboard deck, the length, and other factors, including the length of superstructure (deck erections) such as poop, bridge, and forecastle. The assigned draft corresponds to the load line (waterline at which the ship is allowed to go down in the water), which is distinguished by a mark on the side amidships (load line mark). The assigned freeboard is the distance measured vertically downward amidships from the upper edge of the deck line to the upper edge of the related load line. For dry cargo ships of 500–700-ft (150–210-m) length, the permissible freeboard is approximately 1.6% of the length, while for tankers of the same lengths it is approximately 1.3% of the length. (In other words, tankers are allowed smaller freeboards than are dry cargo ships because of the smaller, easier to keep watertight, deck openings and because of the large number of bulkheads which provide a high degree of subdivision.) Dry cargo ship freeboards may be closer to those of tankers, provided certain standards of subdivision and stability are met. Safety regulations usually require greater freeboard for passenger ships than for cargo ships.

If the draft at the bow is greater than at the stern, a ship is said to trim by the bow (or head). If greater at the stern, the ship trims by the stern. When fully loaded with homogeneous cargo, fuel, and water (no water ballast), ships seldom trim by the bow, but a 2- or 3-ft (0.6- or 0.9-m) trim by the stern is usually acceptable, except in restricted depths of water.

Bulkheads. Ships are divided into watertight compartments by transverse bulkheads (walls) to reduce the extent of seawater flooding in case of damage and to stiffen the hull structure. A normal 250-ft (75-m) general cargo ship has no less than four such bulkheads, and if the ship is 500 ft (150 m) long, eight bulkheads are typical. Passenger vessels usually have more bulkheads, as required by the international safety regulations.

Double bottom. An inner bottom is fitted between the peak (endmost) bulkheads. An inner bottom is horizontal plating, usually 3–5 ft (0.9–1.5 m) above the keel, extending from side to side of the ship. It provides tank space below (double bottom) for fuel oil, fresh water, and seawater ballast, and also minimizes chances of serious flooding in case of minor bottom damage, for example, from grounding.

Superstructures. The partial decks above the main deck, together with the ship's sides, enclose the superstructures. The one at the forward end is called the forecastle and the one aft, poop. Usually, when the machinery is located amidships, the bridge superstructure is placed above it. Above deck spaces that do not extend to the side are called deckhouses. Superstructures may be used partly for cargo stowage and together with the deckhouses provide space for living accommodations, ship's gear, equipment, and

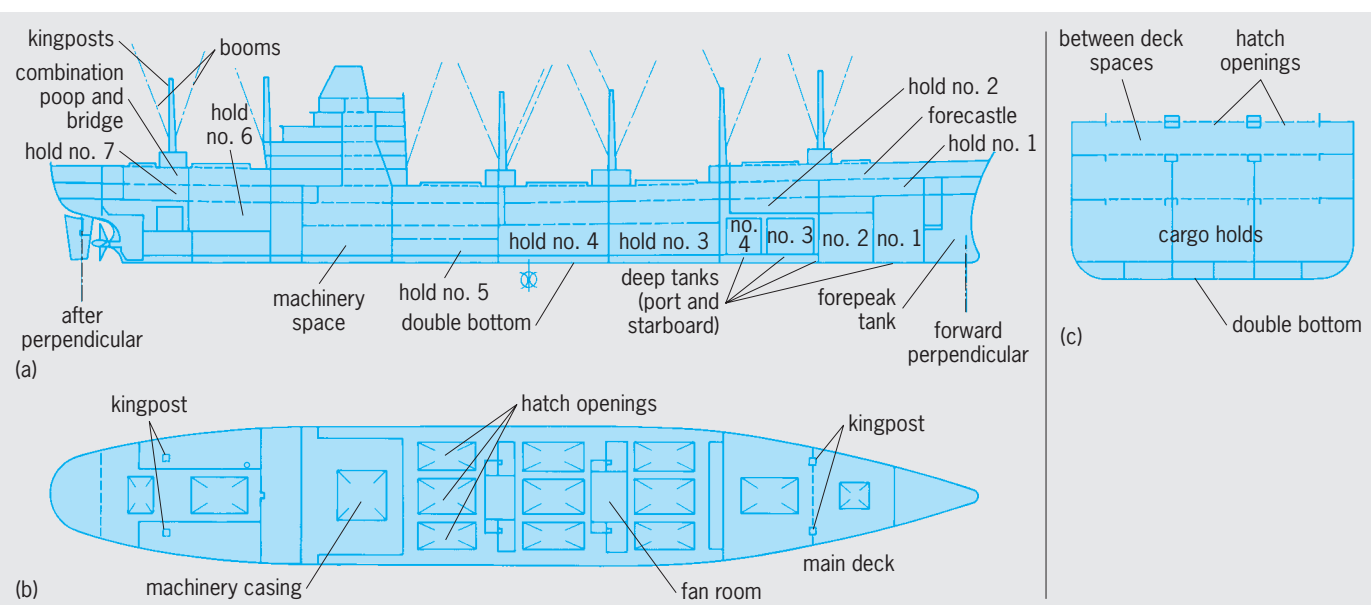

Fig. 1. Break-bulk cargo ship. (*a*) Inboard profile. (*b*) Plan view (deck). (*c*) Typical section. Principal dimensions and other characteristics: length overall = 574 ft = 175 m; length between perpendiculars = 544.5 ft = 166 m; beam = 82 ft = 25 m; depth = 45.5 ft = 14 m; draft = 30.5 ft = 10 m; displacement = 21,235 long tons = 21,576 metric tons; deadweight = 12,932 long tons = 13,140 metric tons; service speed = 23 knots = 43 km/h; maximum shaft horsepower = 24,000 (18 MW); number of crew = 45.

control. The ship in **Fig. 1** has a long forecastle and a long poop combined with the bridge.

Framing. The principal component of the ship's structure is the plating (usually of steel), including the outer watertight envelope. The ship plating is reinforced and stiffened by welded bars (commonly angle, tee, or flat bars). When these bars are welded to the outer envelope, decks, or inner bottom and are repetitive and closely spaced (2–3 ft or 0.6–0.9 m apart), they are called frames and the structural system they compose is called framing. Typical frame dimensions for plating $^1/_2$ to 1 in. (1.3 to 2.5 cm) thick are a depth of 10 in. (25 cm), a flange of 4 in. (10 cm), and a thickness of $^1/_2$ in. (1.3 cm).

Most large land structures such as bridges, building framing, and towers are a complex of struts, rods, beams, and girders, with a negligible amount of plating for secondary functions.

Aircraft structure is similar to ship structure, but its strength and stiffness rest more heavily on the framing system rather than on the plating (shell), which is used more for reasons of aerodynamics than of strength.

Ships have either longitudinal or transverse framing systems or a mixture of both, called combination framing. In a longitudinal framing system the frames are oriented fore and aft and supported transversely by bulkheads and widely spaced deep web frames (6–12 ft or 1.8–3.7 m apart). The transverse system of framing consists essentially of a series of closely spaced ribs (frames) girding the ship, supported longitudinally by the decks, vertical keel, and side and deck girders. In the combination system of framing the bottom plating and often the deck plating are longitudinally framed, while the sides are transversely framed.

Stability. Since in ship design both the breadth and depth affect transverse stability against capsizing, the one must be adjusted in relation to the other. An unusually deep ship must also be unusually broad. Very little stability is risky as the ship may capsize under adverse service conditions. Excessive stability results in objectionable violent rolling in heavy weather.

As fuel and fresh water are consumed en route, the reduction in weight aboard at low levels in the hull reduces stability, possibly risking capsizing if the ship is damaged by collision. In the case of passenger vessels particularly, seawater ballast may be pumped aboard to replace this reduced weight, preferably into low-level compartments not ordinarily used for fuel oil.

Speed factors. Speed in knots is the velocity in nautical miles (6076.1 ft or 1852 m) per hour. Slow seagoing ships (those in which the speed in knots is equal to 50–60% of the square root of the waterline length in feet) have about 30% of ship length of underwater parallel body (no change in shape) and rather bluff fore- and afterbodies. High-speed ships (those in which the speed in knots exceeds the square root of the waterline length in feet) have no parallel underwater body and both ends are quite fine, the forebody waterlines being straight or even hollowed near the bow.

The increased construction cost of weight-saving aluminum alloy in the structure, furniture, and equipment of a high-speed ship is often more than justified by the resultant reduction in horsepower and fuel consumption. The situation pyramids as a reduction in propulsion machinery weight, and reduced fuel weight aboard further reduces the displacement.

An outstanding high-speed liner was the 990-ft (302-m) overall length, quadruple screw, 2000-passenger SS *United States*. During several days of its transatlantic maiden voyage the ship averaged over 36 knots (67 km/h), an unprecedented merchant ship speed.

Propulsion. Cargo ships seldom exceed 40,000 shaft horsepower (shp), equivalent to 30 MW, and one propeller is usual. Higher-powered ships may have two or four propellers. Three are possible but unusual. Ships of less than about 3000 shp (2.2 MW) normally have diesel engines, but for higher powers diesel engines or steam turbines may be used. In the past, steam turbines were more commonly used for ships built in the United States, but the diesel engine is now more often chosen because of fuel economy. Nuclear power, gas turbines, free piston engines, and other power units are also used to propel merchant ships. Sails are almost extinct, although there is considerable interest in a rebirth using more sophisticated, less labor-intensive riggings.

Generally, high-revolution propellers are inefficient. Therefore, if the power plant operates at high revolutions per minute (rpm), it either drives the propeller shaft through reduction gears or powers a direct-drive high-rpm electric generator. The generator in turn drives the propeller shaft through a low-rpm reversible rotation electric motor.

The propulsion machinery of most merchant ships, with the exception of bulk carriers, is located between amidships (midlength of the L_{pp}) and about three-fourths of the L_{pp} from the bow. A general cargo ship normally has fewer holds abaft (sternward from) the machinery space than forward of it. The shaft tunnel and shape of holds near the stern, however, make cargo stowage difficult. There is a trend toward locating the machinery and the deckhouses above it farther aft. Bulk carriers such as tankers, ore carriers, and colliers have the machinery located aft of the cargo spaces. *See* MARINE MACHINERY; PROPELLER (MARINE CRAFT); SHIP POWERING, MANEUVERING, AND SEAKEEPING.

Size-speed relationship. Other factors being equal, a large, fully loaded cargo ship (break-bulk or bulk cargo) will transport cargo at less cost per ton-mile than a small, fully loaded vessel. However, a fully loaded moderate-capacity cargo vessel will transport cargo at less overall cost per ton-mile than a partially loaded, much larger ship of the same type carrying the same quantity of cargo.

Deadweight (dwt) is the weight of cargo, fuel, fresh water, stores, passengers, crew, and baggage aboard a ship. General-cargo transoceanic ships seldom have over 15,000 long tons deadweight capacity, and ships of 8000- to 11,000-ton capacity predominate. In ship design and operation the long ton, 2240 lb (1016.05 kg or 1.016 metric tons), is used. Larger ships have less frequent sailings for a given

annual cargo, and often there is not enough cargo to fully load even an 8000-ton deadweight ship of this type. In ocean service, dry bulk carriers in the neighborhood of 100,000-ton dwt and tankers above that capacity are numerous.

High speed is an important advantage in wartime, and in peacetime it attracts the patronage of both passengers and cargo shippers. High fuel prices, however, have had a dampening effect on the trend toward higher speeds. If two ships of the same cargo capacity are both fully loaded, the cost of shipment per cargo ton-mile is greater for a high-speed vessel. Nevertheless, the faster ship may earn larger dividends if the speed attracts full loads of high-quality cargo. Most cargo ships over 400 ft (122 m) in length have speeds of about 16–18 knots (30–33 km/h). There are many making 20 or more knots (37 km/h). United States ships are a little faster, on the average, than those of other nations.

Cargo handling. The cargo handling time for break-bulk cargo is much longer than for bulk cargo, and this involves costly idle ship time (spending instead of earning) while in port. For vessels of the same basic dimensions and speed, a break-bulk cargo ship can transport more weight or volume of payload cargo than a unitized ship, but the port delay for cargo handling of the break-bulk ship is much greater. Bulk cargo transportation is the most economical per ton-mile, even though the ship is in ballast without cargo on the return passage (except for combination bulk carriers). For relatively long distances, break-bulk cargo ships provide more economical transportation than unitized cargo vessels. This is not the case for short distances. As an extreme example, on a 10-mi (16-km) passage a ferry boat can transport loaded truck vans much more cheaply and quickly than a break-bulk cargo ship, even if more truckloads of cargo were unloaded onto the ferry boat and then reloaded onto trucks at the other end of the passage. For a 10,000-mi (16,500-km) passage, however, it is cheaper to transfer the cargo to a ship and transport only the payload cargo, even though considerable time and money are spent loading and unloading.

Size restrictions. Canal locks and harbor water depths restrict the dimensions of the larger ships. The Panama Canal recommends that vessels not exceed 900-ft (274-m) length overall, 104-ft (32-m) breadth, and 35-ft (10.7 m) draft. The Suez Canal limits the draft to 35 ft (10.7 m), but the length and breadth are unrestricted. The maximum for the St. Lawrence Seaway is 730-ft (223-m) length overall and 75-ft (23-m) breadth. Since the minimum channel depth is 27 ft (8.2 m), about 25.5 ft (7.8 m) is the maximum draft. The minimum clearance under bridges along the Seaway is 120 ft (36.6 m). All these dimensions, except the bridge clearance, can be exceeded slightly in an emergency. The enlarged Soo Lock can accommodate ships 1000 ft (305 m) long and 105 ft (32 m) wide. Although large dry-bulk cargo ships have economic advantages, if the draft exceeds about 38 ft or 11.6 m (few harbors are over 40 ft or 12.2 m in depth), operations may be restricted. Tankers of great draft such as those over 100,000-ton dwt may have to unload partially into a smaller tanker before entering many harbors or may have to unload all cargo at offshore terminals.

Ship Types

The trend brought about by increased transoceanic commerce, higher wages, and keener competition has been toward ship types built for unitized cargo and for certain specific cargoes. The following descriptions of merchant ship types refer to those illustrated in the 10 figures. Ships engaged exclusively in the transportation of passengers are not included because, especially in the United States, air travel has become the predominant mode of transoceanic passenger transportation. Many passenger liners have entered the cruise trade, for ships have many recreational and environmental advantages.

Break-bulk cargo ships. The cargo carrying space in break-bulk cargo ships is divided by transverse bulkheads, spaced 40–70 ft (12–21 m) apart, into a series of cargo compartments of approximately equal volume, generally seven for a ship of about 500-ft (150-m) L_{pp} (Fig. 1). Vertically, these spaces are divided by one or two decks below the uppermost, continuous deck (main or strength deck). The space between the inner bottom and the lowest deck, called the hold, is limited to a height of about 18 ft (5.5 m) to minimize damage to cargo through crushing. Usually the height of each space between decks (termed tween deck space) is 9–10 ft (2.7–3.0 m).

In addition to the previously mentioned double-bottom tanks, most break-bulk cargo ships have deep tanks used for fuel oil, water ballast, or liquid cargoes such as latex, coconut oil, or edible oils.

The cargo is handled through large rectangular deck openings (hatches) over each cargo space. Mechanically operated hatch covers are used to close the openings. The hatch covers in the tween decks are strong enough to support cargo stowed on them. The topside hatch covers are watertight. The tween deck space is generally suitable for break-bulk or palletized cargo. Generally, cargo holds have had one hatch per deck, with a width of 35–50% the ship's breadth and a length of 50–60% the hold length. The trend is toward wider hatches or multiple hatches abreast (Fig. 1) and often longer hatches, to increase cargo handling speed. A multiple hatch arrangement (triple hatch, for instance) is efficiently used for a partial load of containers stowed under deck.

Break-bulk cargo handling between pier and ship is done usually by means of cargo booms installed on board. The booms are raised or lowered by adjustable wire rigging led from the mast or king post to the boom ends. A wire rope leads over sheaves from a winch to the outer end of each boom and terminates in a cargo hook. Cargo can be hoisted using one boom [customarily for very heavy loads of cargo, 10 tons (9 metric tons) or over] or, for faster handling, by a pair of married booms, with one boom end over the hatch and the other over the pier. This cargo handling operation, called burtoning, is customary for loads up to 10 tons. Most break-bulk cargo ships fitted with booms have a pair

of booms at each hatch end to expedite cargo handling. The cargo is often piled together in a large net which is emptied and returned for the next load. Packaged cargo of nearly uniform dimensions may be stacked on pallets which are hoisted aboard individually. The sling load is landed through the hatch opening. The pallets or nets are then unloaded, and each item is individually stowed by the hold gang. Any cargo stowed in the wings of the hold is handled manually unless it is on pallets and handled by a forklift truck. The use of forklift trucks is becoming common practice, and a number of these trucks may be carried on board if they are not available at cargo terminals. The amount of cargo which is handled manually onboard determines largely the ship turnaround and port expenses, and, hence, the profitability of the transportation system.

Most break-bulk cargo ships have provisions for a heavy lift boom of 33–110-ton (30–100-metric ton) capacity for occasional units of heavy cargo.

An increasing number of break-bulk cargo ships are being fitted with revolving deck cargo cranes instead of masts, booms, and winches.

Container ships. Container ships are replacing the conventional break-bulk cargo ship in trade routes where rapid cargo handling is essential. Containers are weatherproof boxes (usually metal) strengthened to withstand stacking and motion at sea. Containers are of standard size, the largest ones weighing up to about 33 tons (30 metric tons) when loaded. The use of standard containers facilitates shipboard stowage, land or waterway transportation, and rental or lease. *See* MARINE CONTAINERS.

A large container ship may be loaded or unloaded completely in about half a day, compared to several days for the same amount of cargo in a break-bulk cargo ship. Generally, the shipper places the cargo in the container, and, except for custom inspection, it is delivered unopened to the consignee. Highway trailers (most commonly), railroad cars, or barges transport containers to and from their land destination and are therefore a part of the same transportation system. For a given payload cargo capacity, container ships are larger and more costly to build than the traditional cargo ship, but both the cargo handling cost and the idle ship time in port are reduced considerably.

Although in some ships containers are moved horizontally for loading and unloading, the predominant arrangement is that illustrated in **Fig. 2**, where containers are stowed in vertical cells and moved vertically in and out of the vessel.

Roll-on/roll-off ships. With a broad interpretation all ships that are designed to handle cargo by rolling it on wheels can be considered under this heading. This would include trailer ships; sea trains (carrying railroad cars or entire trains); auto, truck, and trailer ferries; military vehicle carriers; ships carrying pallets handled by forklift trucks from and to shore; and so on. The following is a description of a ship of this type, which is intended primarily to operate as a trailer ship, although it may handle several types of wheeled vehicles **(Fig. 3)**.

Roll-on/roll-off ships require a high proportion of cubic capacity relative to the amount of cargo and are particularly suited to service with short runs and frequent loading and unloading. They need even shorter port time than container ships, but their building cost is higher.

Because fully loaded roll-on/roll-off ships cannot carry enough cargo to immerse them deeply, their large freeboard allows the fitting of side ports above the waterline for the handling of cargo on wheels by means of ramps. Usually, ships of this type have a transom stern (a square-shaped stern like that of a motorboat) fitted with doors for handling wheeled vehicles on an aft ramp. Roll-on/roll-off ships have several decks, and the cargo is handled on wheels

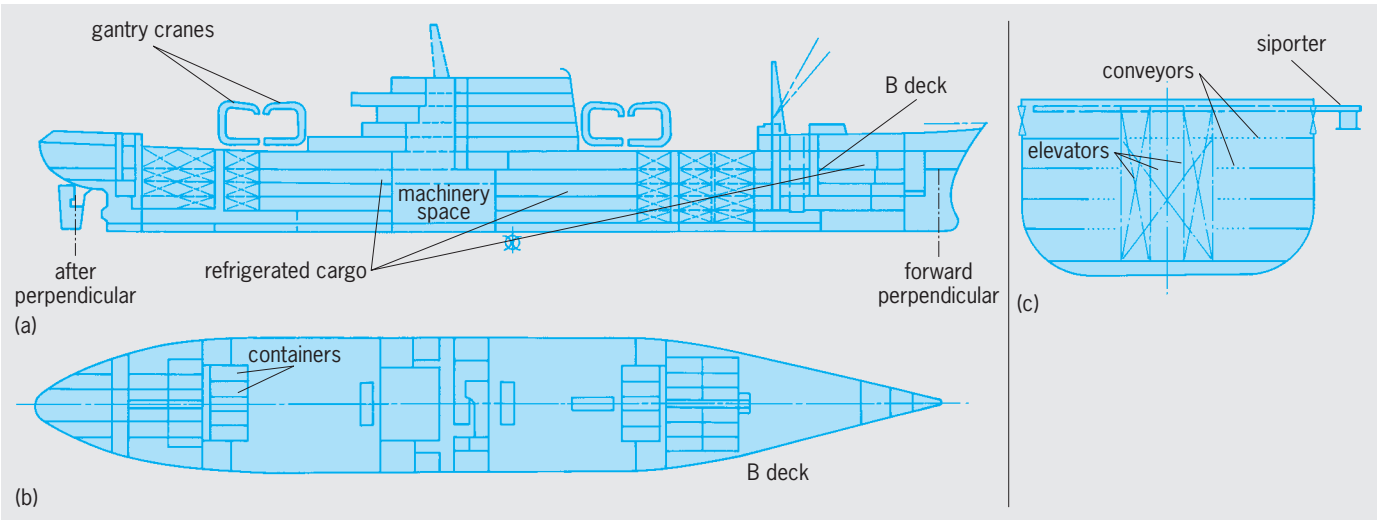

Fig. 2. Container ship. (*a*) Inboard profile. (*b*) Plan view (deck). (*c*) Typical section. Principal dimensions and other characteristics: length overall = 752 ft = 229 m; length between perpendiculars = 705 ft = 215 m; beam = 100 ft = 30 m; depth = 57 ft = 17 m; draft = 31 ft = 9 m; displacement = 33,924 long tons = 34,468 metric tons; deadweight = 19,206 long tons = 19,514 metric tons; service speed = 27 knots = 50 km/h; shaft horsepower (two propellers) = 60,000 (44.7 MW); number of crew = 40.

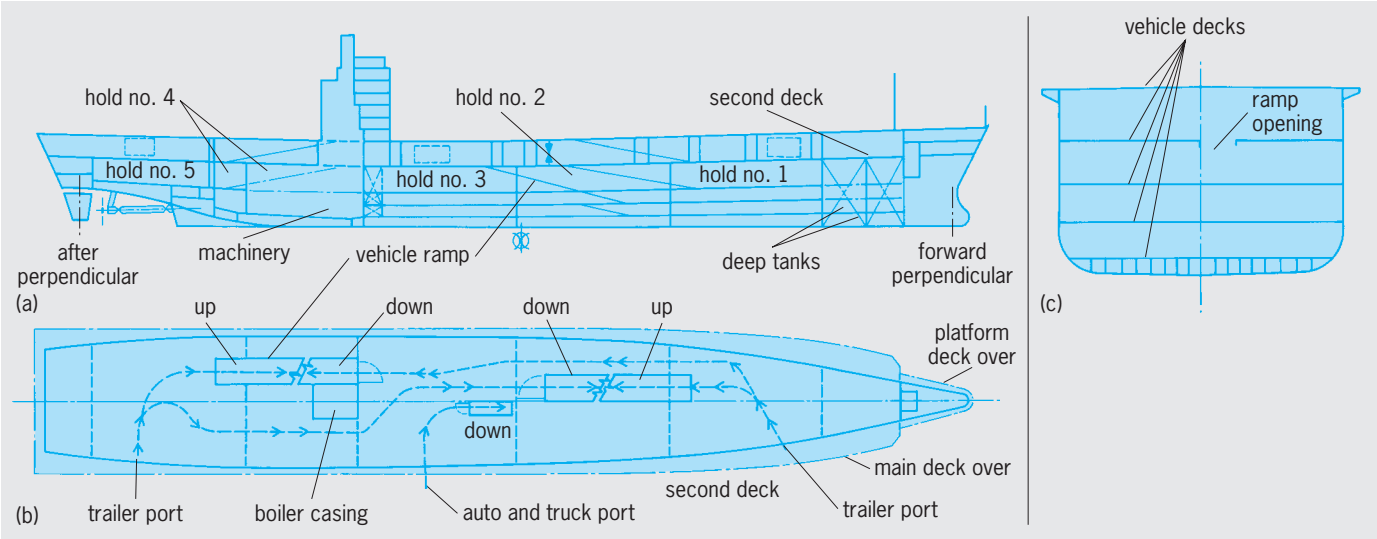

Fig. 3. Roll-on/roll-off ship. (*a*) Inboard profile. (*b*) Plan view (deck). (*c*) Typical section. Principal dimensions and other characteristics: length overall = 700 ft = 213 m; length between perpendiculars 643 ft = 196 m; beam (maximum) = 105 ft = 32 m; depth = 60 ft = 18 m; draft = 27 ft = 8 m; displacement = 24,100 long tons = 24,500 metric tons; deadweight = 13,100 long tons = 13,300 metric tons; service speed = 25 knots = 46 km/h; shaft horsepower = 32,000 (23.9 MW); number of crew = 39.

from the loading deck to other decks by elevators or sloping ramps. Both internal elevators and ramps occupy substantial volume in the ship. The need for clear decks, without interruption by transverse bulkheads, and tween decks for vehicle parking results in a unique structural arrangement.

Barge-carrying ships. This type of ship represents a bold step in the trend toward cargo containerization and port time reductions: Cargo is carried in barges or lighters, each weighing up to 1100 tons (1000 metric tons) when loaded. The lighters are carried below and above deck and are handled by gantry cranes or elevator platforms (**Fig. 4**).

These are among the fastest, largest and costliest ships for the carriage of general cargo. For their size, their payload capacity is less than that of the conventional break-bulk cargo ship. However, they can be loaded and unloaded much faster and with a considerable saving in worker-hours.

Because the lighters can be waterborne and operated as regular barges, these large ships can serve undeveloped ports advantageously.

Using portable fixtures that can be erected quickly, barge-carrying ships can be adapted for the transport of varying amounts of standard containers in addition to or in place of lighters.

Bulk cargo ships. A large proportion of ocean transportation is effected by bulk cargo ships.

Dry bulk cargo includes products such as iron ore, coal, limestone, grain, cement, bauxite, gypsum, and sugar. Most oceangoing dry bulk carriers are loaded and unloaded using shoreside installations. Many dry bulk carriers operating in the Great Lakes have shipboard equipment for the handling of cargo (self-unloaders), and an increasing number of oceangoing ships carrying this type of cargo are being fitted with self-unloading gear.

By far the largest amount of liquid bulk cargo consists of petroleum products, but ocean transportation of other bulk liquid products is increasing in importance: for example, various chemicals, vegetable oils, molasses, latex, liquefied gases, molten sulfur, and even wine and fruit juices. Practically all liquid bulk carriers have pumps for unloading the cargo.

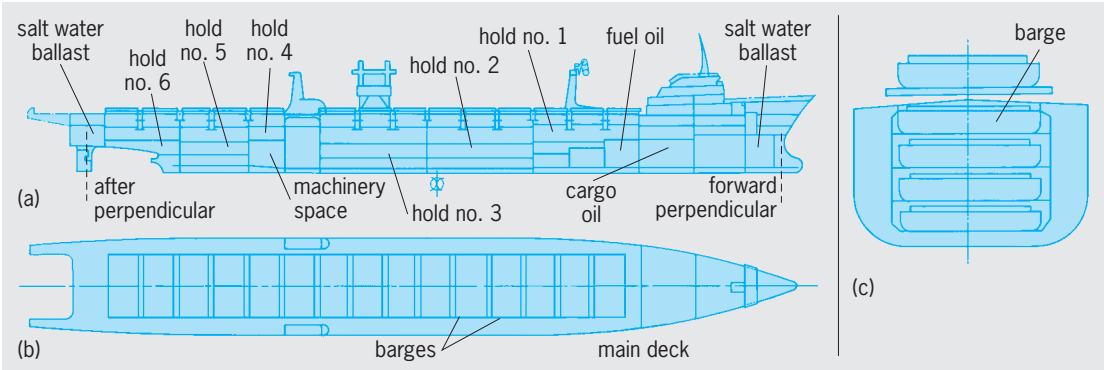

Fig. 4. Barge-carrying ship. (*a*) Inboard profile. (*b*) Plan view (deck). (*c*) Typical section. Principal dimensions and other characteristics: length overall = 820 ft = 250 m; length between perpendiculars = 724 ft = 221 m; beam = 100 ft = 30.5 m; depth = 60 ft = 18 m; draft = 28 ft = 8.5 m; displacement = 32,800 long tons = 33,330 metric tons; deadweight = 18,760 long tons = 19,060 metric tons; service speed = 22.5 knots = 42 km/h; shaft horsepower = 32,000 (24 MW); number of crew = 39.

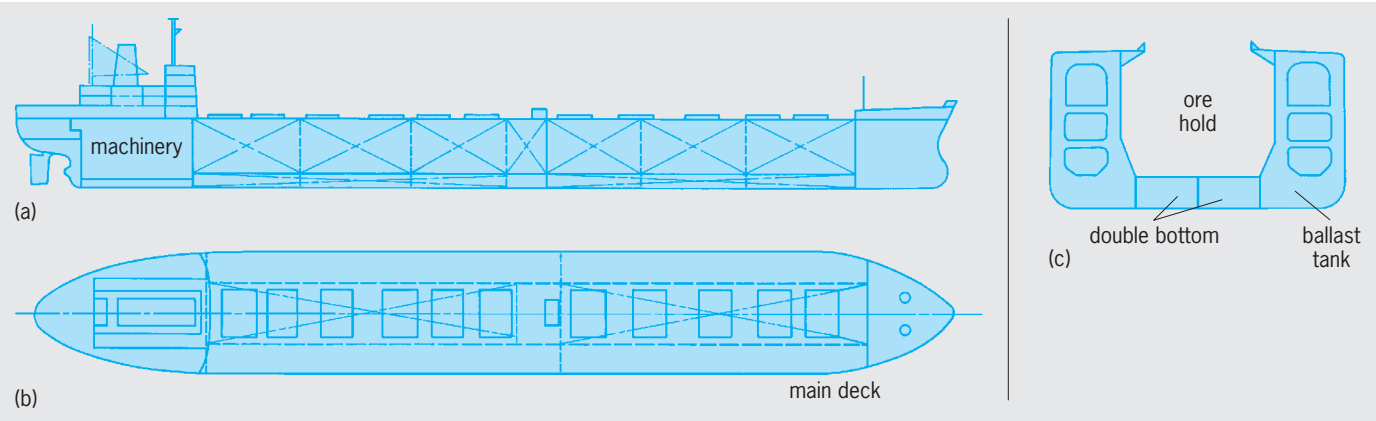

Fig. 5. Oceangoing ore carrier. (*a*) Inboard profile. (*b*) Plan view (deck). (*c*) Typical section. Principal dimensions and other characteristics: length overall = 865 ft = 263 m; length between perpendiculars = 826 ft = 252 m; beam = 124 ft = 39 m; depth = 69 ft = 22 m; draft = 50 ft = 15.5 m; displacement = 124,000 long tons = 126,000 metric tons; deadweight = 104,500 long tons = 106,200 metric tons; service speed = 16 knots = 29 km/h; shaft horsepower = 24,300 (18 MW); number of crew = 40.

Even combination bulk carriers, which may carry simultaneously or alternatively dry and liquid bulk cargo, usually have shipboard pumps for unloading liquids.

Practically all bulk carriers have the machinery compartment, crew accommodations, and conning stations located aft. An exception is the Great Lakes self-unloader with crew accommodations and bridge forward.

The tendency in bulk carriers is toward larger ships, with speeds remaining about constant at a moderate level (16–18 knots or 30–33 km/h for oceangoing ships, lower for Great Lakes vessels).

The oceangoing ore carrier (**Fig. 5**) is characterized by a high double bottom and small volume of cargo hold because of the high density of the ore. Storing the cargo high in the ship decreases stability and prevents excessively quick rolling.

The oceangoing combination bulk carrier (**Fig. 6**) permits low-cost transportation because of its flexibility. It is able to carry many types of bulk cargoes over a variety of sea lanes. This type of ship carries bulk cargoes, such as petroleum products, coal, grain, and ore. The double bottom in bulk carriers is shallow and the volume of cargo holds is large compared to the size of the ship.

The self-unloader (**Fig. 7**) was originally developed in the Great Lakes for the transportation of limestone. A variety of free-flowing dry bulk cargo is carried in this type of ship. Generally, a conveyor belt located low in the ship, so that it can be loaded by gravity flow, carries the cargo to a station at one end of the ship. From there it is transferred automatically to the conveyor of a large boom, which can be swung over the side for unloading ashore.

The tanker (**Fig. 8**) is the characteristic, and by far the most important, liquid bulk carrier both in numbers and tonnage. Tankers carry petroleum products almost exclusively. The very large tankers are used almost entirely for the transport of crude oil. A few tankers are built especially for the transportation of

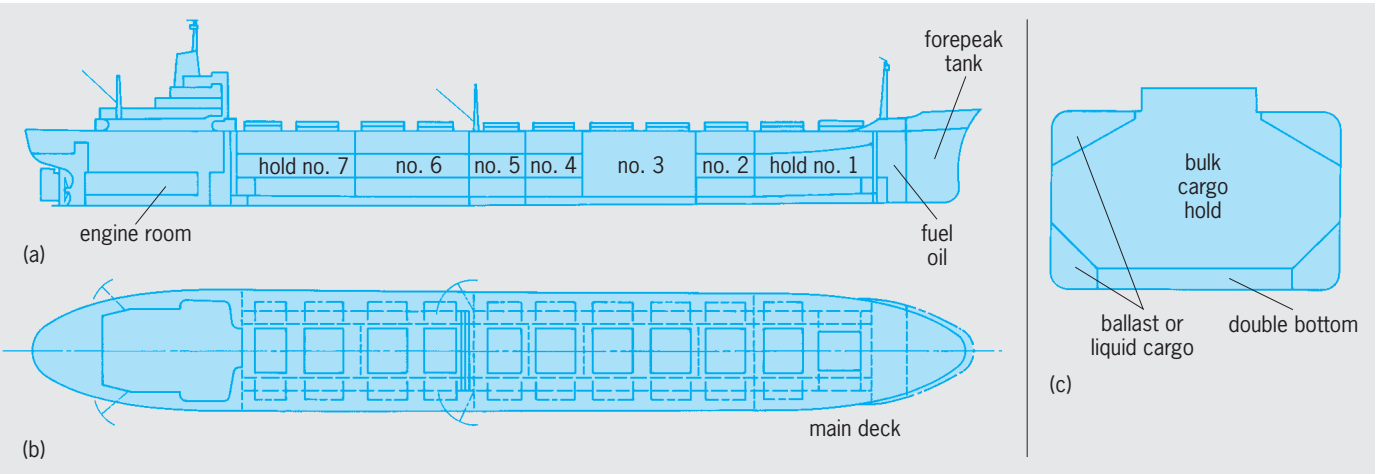

Fig. 6. Oceangoing combination bulk carrier. (*a*) Inboard profile. (*b*) Plan view (deck). (*c*) Typical section. Principal dimensions and other characteristics: length overall = 815 ft = 248 m; length between perpendiculars = 775 ft = 236 m; beam = 104.5 ft = 32 m; depth = 61.5 ft = 19 m; draft = 41 ft = 12.5 m; displacement = 78,938 long tons = 80,205 metric tons; deadweight = 63,410 long tons = 64,428 metric tons; service speed = 16 knots = 29 km/h; shaft horsepower = 18,000 (13.4 MW); number of crew = 41.

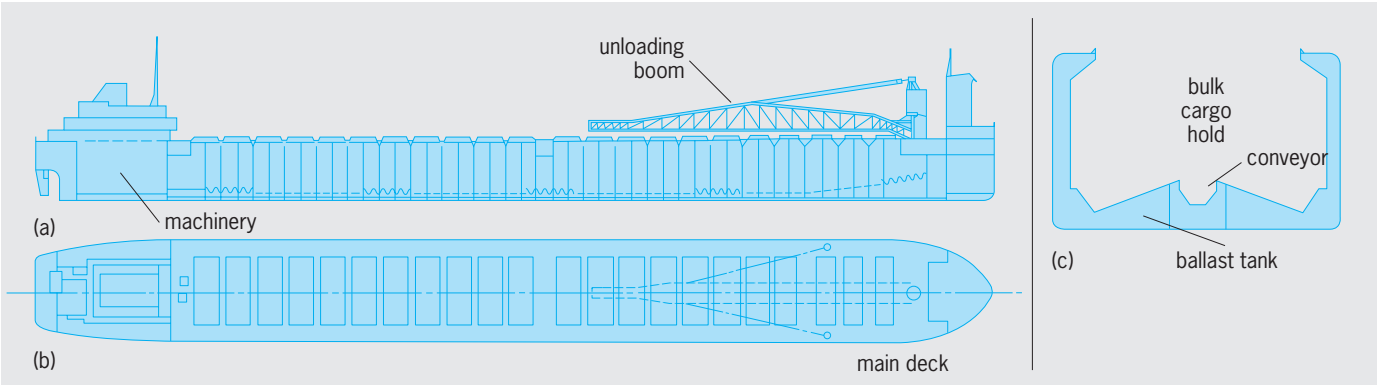

Fig. 7. Great Lakes self-unloader. (*a*) Inboard profile. (*b*) Plan view (deck). (*c*) Typical section. Principal dimensions and other characteristics: length overall = 730 ft = 222.5 m; length between perpendiculars = 724 ft = 221 m; beam = 75 ft = 23 m; depth = 45 ft = 14 m; draft = 28 ft = 8.5 m; displacement = 38,980 long tons = 39,605 metric tons; deadweight = 29,550 long tons = 30,024 metric tons; service speed = 13 knots = 24 km/h; shaft horsepower = 7400 (5.5 MW); number of crew = 32.

chemical products, and others are prepared for alternative loads of grain.

Bulk liquid carriers, with freestanding, rectangular, cylindrical, or spherical cargo tanks separated from the hull, are used for the transportation of molten sulfur and liquefied gases, such as anhydrous ammonia and natural gas (**Fig. 9**). Liquefied natural gas (LNG) is also carried in ships with membrane tanks, i.e., where a thin metallic liner is fitted into a tank composed of ship structure and load-bearing insulation. The transportation of molten sulfur and liquefied gases requires special consideration regarding insulation and high structural soundness of cargo tanks, including the use of high-grade, costly materials for their construction.

Heavy-lift ships. The definition of heavy lift has changed gradually. Whereas 10 short tons (9 metric tons) per unit was once considered heavy-lift cargo, many cargo liners of the 1970s commonly lift 110 tons (100 metric tons) per unit without being classified as heavy-lift ships. Heavy-lift ships considered here are those capable of loading and discharging cargoes having unit weights in the hundreds of tons. The need for these ships has grown due to the expanding market for large industrial equipment.

Heavy-lift cargo includes nuclear power reactors, diesel engines, components for refinery and chemical plants, metalworking machinery, construction equipment, military equipment, offshore oil equipment, locomotives, mining and mineral processing equipment, tugs, and barges.

The large loads carried by heavy-lift ships are handled by high-capacity shipboard jibs and hoisting gear. The factors determining the best derrick for a given design are location of the deck, lifting capacity, outreach, initial cost, and weight. The three main types of lifting gear used are nonrotating derricks, rotating derricks, and gantry cranes.

A commonly used nonrotating heavy-lift derrick has an open double kingpost arrangement which simplifies rigging considerably and allows swinging loads up to 400 metric tons forward and aft of the derrick and to both sides of the ship.

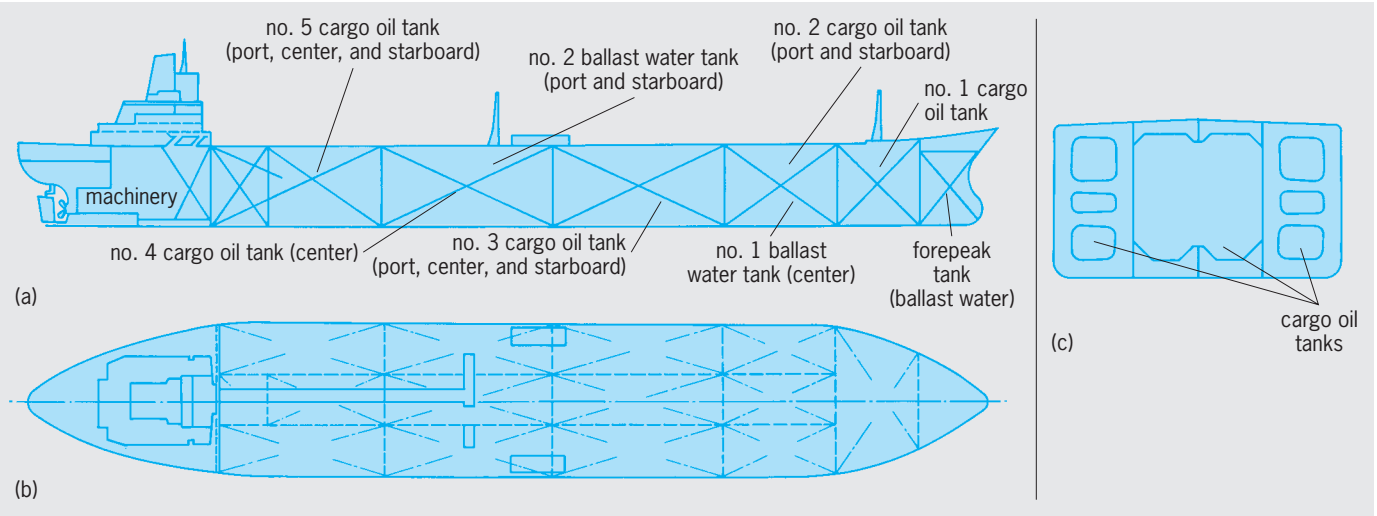

Fig. 8. Tanker. (*a*) Inboard profile. (*b*) Plan view (deck). (*c*) Typical section. Principal dimensions and other characteristics: length overall = 1005 ft = 306.5 m; length between perpendiculars = 951 ft = 290 m; beam = 156 ft = 47.5 m; depth = 79 ft = 24 m; draft = 53 ft = 16 m; displacement = 179,700 long tons = 181,900 metric tons; deadweight = 151,300 long tons = 153,700 metric tons; service speed = 16 knots = 30 km/h; shaft horsepower = 30,000 (22.4 MW); number of crew = 29.

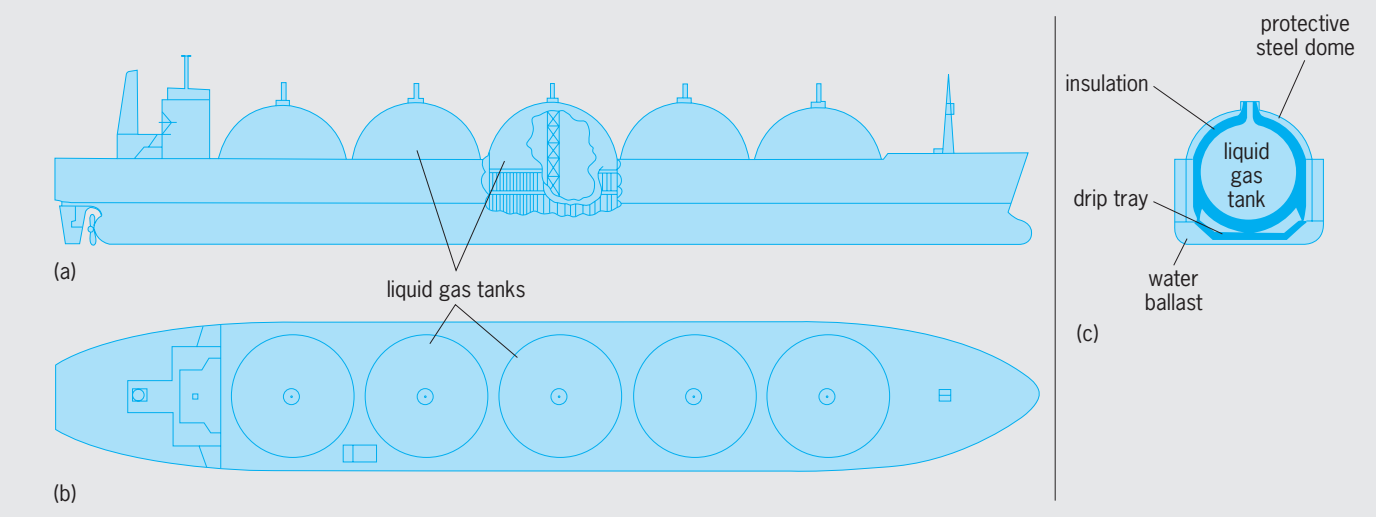

Fig. 9. Spherical-tank liquefied natural gas (LNG) carrier. (*a*) Profile. (*b*) Plan view (main deck). (*c*) Typical section. Principal dimensions and other characteristics: length overall = 924 ft = 282 m; length between perpendiculars = 885 ft = 270 m; beam = 141.5 ft = 43 m; depth = 82 ft = 25 m; draft = 36 ft= 11 m; cargo tank volume = 121,250 m^3; displacement = 91,200 long tons = 92,700 metric tons; deadweight = 61,100 long tons = 62,100 metric tons; maximum speed = 20 knots = 37 km/h; shaft horsepower = 40,000 (29.8 MW); loading or unloading time = 12 h; number of crew = 39.

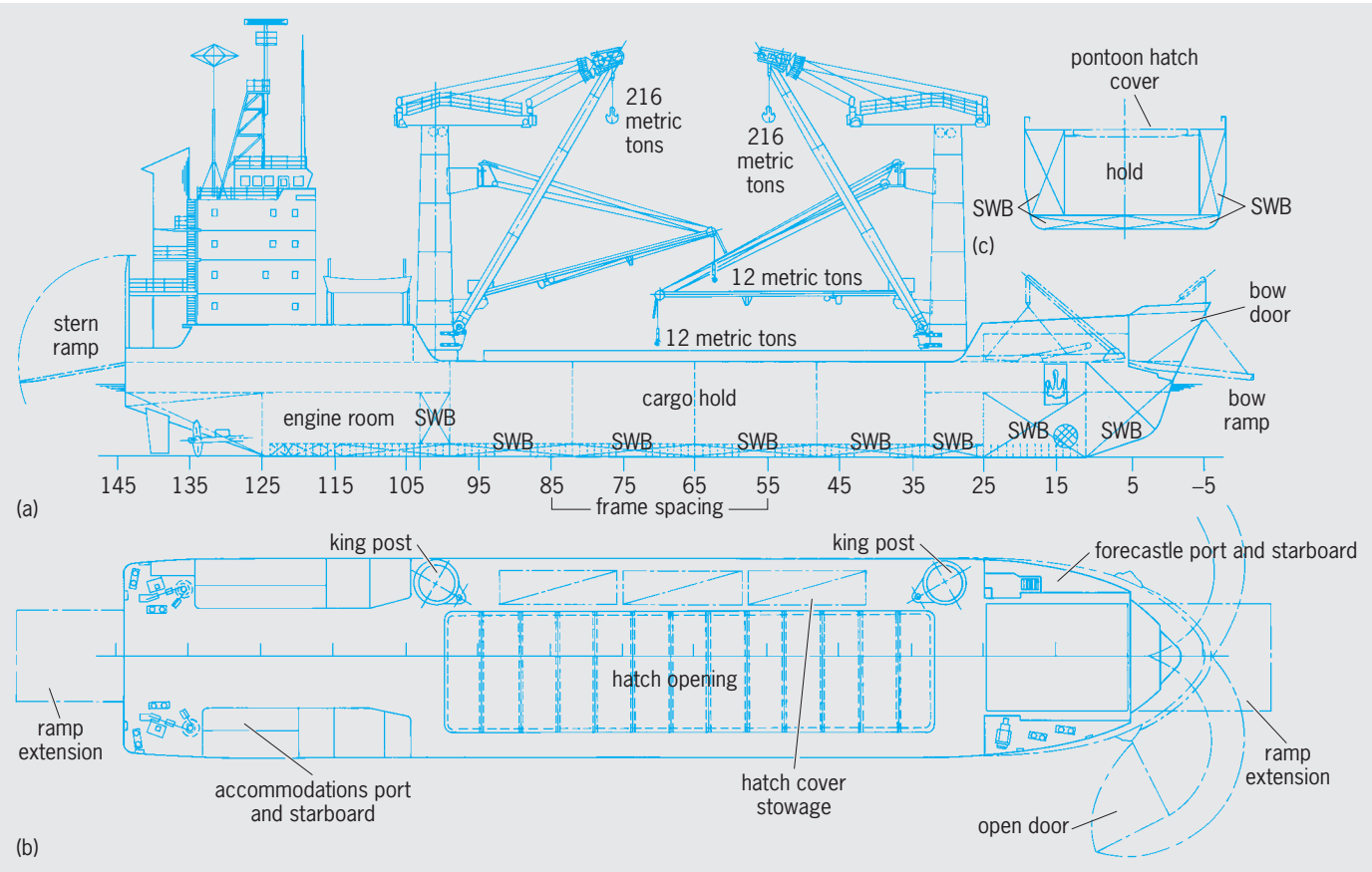

Fig. 10. Heavy-lift ship. (*a*) Profile. (*b*) Plan view (main deck). (*c*) Typical section. SWB = salt water ballast. Principal dimensions and other characteristics: length overall = 300 ft = 91 m; length between perpendiculars = 279 ft = 85 m; beam 55 ft = 17 m; depth = 27 ft = 8 m; draft = 17 ft = 5 m; displacement = 5500 long tons = 5588 metric tons; deadweight = 2947 long tons = 2994 metric tons; service speed = 13.5 knots = 25 km/h; brake horsepower = 5500 (4.1 MW); lifting capacity = 424 long tons = 431 metric tons; number of crew = 21. (*After H. W. Janecke and W. F. Muir, Modern heavy lift ships (state of the art), Trans. Soc. Nav. Architects Mar. Eng., 86:347–374, 1978*)

Rotating derricks have a single tower kingpost with a topping lift and a boom attached at the bottom to a swiveling ring. Their maximum lifting capacity is approximately 770 tons (700 metric tons).

Gantry cranes used on heavy-lift ships carry loads between tracks in traveling the length of the holds. They are generally restricted to a maximum lifting capacity of 550 tons (500 metric tons) and are not utilized in over-the-side lifting, since the outriggers must be retractable.

Handling, transferring, and securing heavy-lift cargo is a specialized operation planned many months in advance. As cargo is lifted, ballast water is transferred to correct for trim and heel computed in advance.

Ships are identified by the method of lifting cargoes or by transferring cargoes on and off. There are two categories of heavy-lift ships: liner-type, general cargo ships utilizing heavy-lift gear, and the heavy-lift ships in the "tramping trade" which provide specialized services such as roll-on/roll-off, float-on/float-off, and slide-on/slide-off. Roll-on/roll-off ships are fitted with doors and ramps on the bow or the stern, including internal ramps between decks. Float-on/float-off ships become submerged at the stern while the cargo—barges, tugs, dredges, or other floating vessels—enters the ship's submerged well. When the well is sealed and pumped dry, the cargo awaits securing. Slide-on/slide-off ships are equipped with two heavy retractable girders extending past the side over which cargo is transferred.

Figure 10 illustrates a heavy-lift ship with rotating derricks and a lifting capacity of 475 tons (424 long tons or 431 metric tons).

Amelio M. D'Arcangelo; Robert B. Zubaly

Bibliography. R. Gardiner and A. Alastair (eds.), *The Shipping Revolution: The Modern Merchant Ship*, 1993; E. V. Lewis (ed.), *Principles of Naval Architecture*, 2d rev., 3 vols., 1988; R. Taggart (ed.), *Ship Design and Construction*, 3d ed., 1980; D. A. Taylor, *Merchant Ship Construction*, 3d ed., 1992; G. Till (ed.), *Merchant Shipping*, 1994; *Transactions of the Society of Naval Architects and Marine Engineers*, annually.

Mercury (element)

A chemical element, Hg, atomic number 80 and atomic weight 200.59. Mercury is a silverwhite liquid at room temperature (melting point −38.89°C or −37.46°F); it boils at 357.25°C (675.05°F) under atmospheric pressure. It is a noble metal that is soluble only in oxidizing solutions. Solid mercury is as soft as lead. The metal and its compounds are very toxic. With some metals (gold, silver, platinum, uranium, copper, lead, sodium, and potassium, for example) mercury forms solutions called amalgams. *See* AMALGAM; PERIODIC TABLE; TRANSITION ELEMENTS.

In its compounds, mercury is found in the 2+, 1+, and lower oxidation states, for example, $HgCl_2$, Hg_2Cl_2, or $Hg_3(AsF_6)_2$. Often the mercury atoms are doubly covalently bonded, for example, Cl—Hg—Cl or Cl—Hg—Hg—Cl. Some mercury(II) salts, for example, $Hg(NO_3)_2$, or $Hg(ClO_4)_2$, are quite soluble in water and dissociate normally. The aqueous solutions of these salts react as strong acids because of hydrolysis. Other mercury(II) salts, for example, $HgCl_2$ or $Hg(CN)_2$, also dissolve in water, but exist in solution as only slightly dissociated molecules. There are compounds in which mercury atoms are bound directly to carbon or nitrogen atoms, for example, H_3C—Hg—CH_3 or H_3C—CO—NH—Hg—NH—CO—CH_3. In complex compounds, for example, $K_2(HgI_4)$, mercury often has three or four bonds.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Metallic mercury is used as a liquid contact material for electrical switches, in vacuum technology as the working fluid of diffusion pumps, for the manufacture of mercury-vapor rectifiers, thermometers, barometers, tachometers, and thermostats, and for the manufacture of mercury-vapor lamps. It finds application for the manufacture of silver amalgams for tooth fillings in dentistry. Of importance in electrochemistry are the standard calomel electrode, used as the reference electrode for the measurement of potentials and for potentiometric titrations, and the Weston standard cell.

Mercury is commonly found as the sulfide, HgS, frequently as the red cinnabar and less often as the black metacinnabar. A less common ore is the mercury(I) chloride. Occasionally the mercury ore contains small drops of metallic mercury. *See* CINNABAR.

The surface tension of liquid mercury is 484 dynes/cm, six times greater than that of water in contact with air. Hence, mercury does not wet surfaces with which it is in contact. In dry air metallic mercury is not oxidized. After long standing in moist air, however, the metal becomes coated with a thin layer of oxide. In air-free hydrochloric acid or in dilute sulfuric acid, the metal does not dissolve. Conversely, it is dissolved by oxidizing acids (nitric acid, concentrated sulfuric acid, and aqua regia).

Klaus Brodersen

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; R. R. Lauwerys and P. Hoet, *Industrial Chemical Exposure: Guidelines for Biological Monitoring*, 3d ed., 2001; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004.

Mercury (planet)

The planet closest to the Sun. It is visible to the unaided eye only shortly after sunset or shortly before sunrise, when it is near its greatest angular distance from the Sun (28°). Mercury is the smallest planet since Pluto is no longer classified as a planet. Its diameter is 3031 mi (4878 km), and its mass is 7.28×10^{23} lb (3.301×10^{23} kg), or 0.055 times the mass of the Earth. Most detailed knowledge of Mercury is derived from data returned by the *Mariner 10* spacecraft, which flew by the planet three times in 1974 and 1975. *Mariner 10* imaged only about 45% of the surface at an average resolution of about 0.6 mi (1 km), and less than 1% at resolutions between about 300 and 1500 ft (100 and 500 m). This coverage and resolution is somewhat comparable to Earth-based telescopic coverage and resolution of the Moon before the advent of space flight. As a consequence, there are still many uncertainties and questions concerning this unusual planet. Mercury represents an end member in solar system origin and evolution because it formed closer to the Sun than any other planet and, therefore, in the hottest part of the solar nebula from which the entire solar system formed. *See* PLUTO; SOLAR SYSTEM.

Motions. Mercury has the most eccentric (0.205) and inclined (7°) orbit of any planet in the solar system. Its average distance from the Sun is 3.60×10^7 mi (5.79×10^7 km), but distance varies from 2.86×10^7 mi (4.60×10^7 km) at perihelion to 4.34×10^7 mi (6.98×10^7 km) at aphelion because of the large eccentricity. The rotation period is 58.646 Earth days, and the orbital period, 87.969. Therefore, Mercury has a 3:2 resonant relationship between its rotational and orbital periods; it makes exactly three rotations around its axis for every two orbits around the Sun. Thus, a solar day (sunrise to sunrise) lasts two mercurian years (176 Earth days).

The motion of Mercury played an important role in the development of the general theory of relativity. Mercury's elliptical orbit is pulled around the Sun so that the perihelion point changes its position in space by about 5600 arc-seconds per century. When all known forces that act on the planet are taken into account, Mercury's perihelion-point motion exceeds that calculated by newtonian gravitational theory by 43 arc-seconds per century. This discrepancy remained unexplained until 1915, when Einstein's general theory of relativity predicted a perihelion motion in exact agreement with observation. *See* RELATIVITY.

Temperature. Although Mercury is closest to the Sun, it is not the hottest planet. The surface of Venus is hotter because of the atmospheric greenhouse effect. However, Mercury experiences the greatest range in surface temperature (1130°F or 628°C) of any planet or satellite in the solar system because of its proximity to the Sun, its long solar day, and its lack of an insulating atmosphere. Its maximum surface temperature is 800°F (427°C) at perihelion on the equator, hot enough to melt zinc. At night, however, the unshielded surface plunges to below −300°F (−184°C). *See* VENUS.

Polar radar features. High-resolution, full-disk radar images of Mercury show very high reflectivities and polarization ratios in the polar regions. The reflectivity and ratio values are similar to outer-planet icy satellites and the residual polar water-ice cap of Mars. Therefore, Mercury's polar radar features are interpreted to be water ice.

Twenty discrete high-reflectivity anomalies have been detected in the polar regions, and most of these correspond to craters viewed by *Mariner 10*. Some of the anomalies occur at latitudes as low as 72°. Because the obliquity of Mercury is near 0°, the planet does not experience seasons, and therefore temperatures in the polar regions should be less than −216°F (−138°C). In permanently shaded areas of the floors and sides of craters, the temperatures could be less than −258°F (−161°C), and water ice should be stable to evaporation on time scales of billions of years. Furthermore, the ice could be covered by a thin layer of regolith and still be detected by radar. The source of the water is not known. It may have been derived from comets or from water-rich asteroid impacts when water vaporized by the impacts is cold-trapped in the polar regions. If this is the source of the water on Mercury, the permanently shaded areas of craters in the polar regions of the Moon should also have water deposits. In 1998 the *Lunar Prospector* spacecraft identified water ice in the permanently shadowed areas of craters in the Moon's polar regions. *See* MOON; RADAR ASTRONOMY.

Magnetic field. *Mariner 10* discovered an intrinsic dipole magnetic field with a dipole moment equal to about 0.004 that of the Earth. Mercury is the only terrestrial planet besides the Earth with a magnetic field. Although weak compared to that of the Earth, the field has sufficient strength to hold off the solar wind, creating a bow shock and accelerating charged particles from the solar wind. As with the Earth, the magnetic axis is inclined about 11° from the rotation axis. Mercury occupies a much larger fraction of the volume of its magnetosphere than do other planets, and the solar wind actually reaches the surface at times of highest solar activity. Because of the small size of Mercury's magnetosphere, magnetic events happen more quickly and repeat more often than in Earth's magnetosphere. The maintenance of planetary magnetic fields apparently requires an electrically conducting fluid outer core. Therefore, Mercury's dipolar magnetic field is evidence that Mercury currently has a fluid outer core of unknown thickness. High-resolution radar measurements of the magnitude of Mercury's librations show that the mantle is detached from the core, indicating the outer core is still in a fluid condition. *See* GEOMAGNETISM; MAGNETISM; MAGNETOSPHERE.

Interior. Mercury's internal structure is unique in the solar system. The planet's mean density is 5.44 g/cm^3 (5.44 times that of water), which is larger than that of any other planet or satellite except Earth (5.52 g/cm^3). Because of Earth's large internal pressures, however, its uncompressed density is only 4.4 g/cm^3 compared to Mercury's uncompressed density of 5.3 g/cm^3. This means that Mercury

contains a much larger fraction of iron than any other planet or satellite in the solar system. The iron core must be about 75% of the planet diameter, or 42% of Mercury's volume. It is surrounded by a silicate mantle and crust only about 370 mi (600 km) thick. Earth's core is only 54% of the planet diameter, or just 16% of the total volume. Because the outer core is at least partly molten at present, a light alloying element in the core must have lowered the melting point and retained a partially molten core over geologic history (4.5×10^9 years); otherwise the core would have solidified long ago. Although oxygen is such an element, it is not sufficiently soluble in iron at Mercury's low internal pressures. Therefore, sulfur is the most reasonable candidate for this alloying element. Mercury probably has between about 0.2 and 7% sulfur in its core, since a sulfur abundance of less than 0.2 would result in an entirely solid core at the present time, whereas an abundance of 7% would result in an entirely fluid core at present. *See* EARTH.

Origin. The origin of Mercury and how it acquired such a large percentage of iron is a major unsolved problem. Although chemical equilibrium condensation models successfully explain the densities of most planets in the solar system, they cannot explain Mercury. For Mercury's position in the solar nebula, these models predict an uncompressed density of only about 4 g/cm^3 rather than the observed 5.3 g/cm^3; they also predict the complete absence of sulfur.

Three different hypotheses have been proposed to account for the large discrepancy between the iron abundance indicated by Mercury's high density and that predicted by equilibrium condensation models. In the selective accretion model, a differential response of iron and silicates to impact fragmentation and aerodynamic sorting leads to iron enrichment owing to higher gas densities and shorter dynamical time scales in the innermost part of the solar nebula. The postaccretion vaporization model proposes that intense bombardment by solar electromagnetic and corpuscular radiation in the earliest phases of the Sun's evolution (the T-Tauri phase) vaporized and drove off much of the planet's silicate mantle. In the giant impact hypothesis, a collision of a planet-sized object with Mercury ejected much of the mantle. Each of these hypotheses has major consequences for the formation of the other terrestrial planets, and each predicts a significantly different chemical composition for the silicate portion of Mercury, which are to be tested by the *MESSENGER* mission, now on its way to Mercury.

Atmosphere. Mercury's atmosphere is very tenuous and is essentially exospheric in that its atoms rarely collide with each other. The atmospheric surface pressure is 10^{12} times less than Earth's. *Mariner 10*'s ultraviolet spectrometer identified hydrogen, helium, oxygen, and argon in the atmosphere, all of which are probably derived largely from the solar wind. Earth-based telescopic observations in 1985 discovered that Mercury is surrounded by a tenuous atmosphere of sodium and potassium that is probably derived from its surface. At the subsolar point the atmospheric density of sodium is 280,000 to 620,000 atoms per cubic inch (17,000 to 38,000 atoms per cubic centimeter), and for potassium it is 8000 atoms per cubic inch (500 atoms per cubic centimeter). The mechanism for removing these elements from the surface and injecting them into the atmosphere is uncertain. Several areas of enhanced sodium and potassium emission have been identified with surface features such as the Caloris Basin and two radar-bright areas on the unimaged part of Mercury. The reason for these emissions is uncertain.

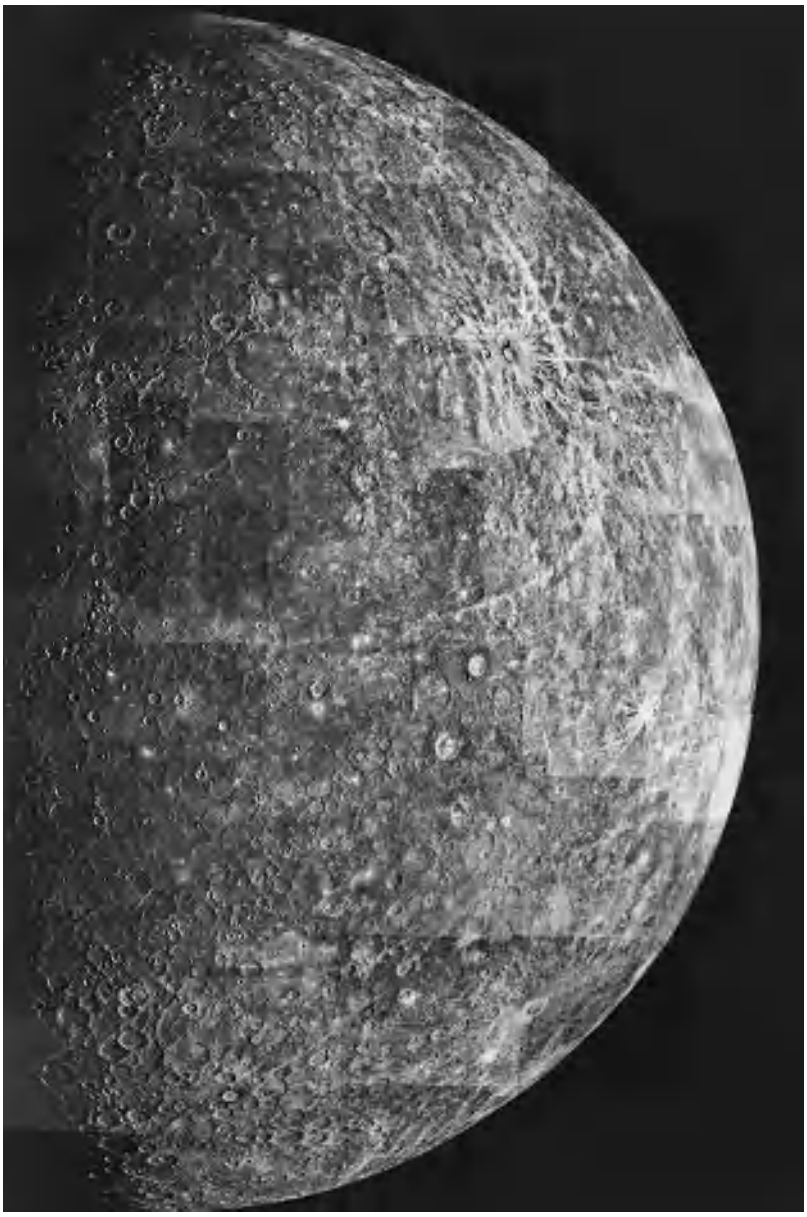

Fig. 1. Photomosaic of Mercury as seen by the outgoing *Mariner 10* spacecraft in March 1974. The terminator, the boundary between the lighted and unlighted halves of the planet, runs vertically along the left side of the photograph. The Caloris Basin is on the terminator, slightly above the middle. It is surrounded and filled by younger smooth plains. (*Jet Propulsion Laboratory*)

Geology. The surface of Mercury superficially resembles that of the Moon (**Fig. 1**). It is heavily cratered, with large expanses of younger smooth plains (similar to the lunar maria) that fill and surround major impact basins. Unlike the Moon surface,

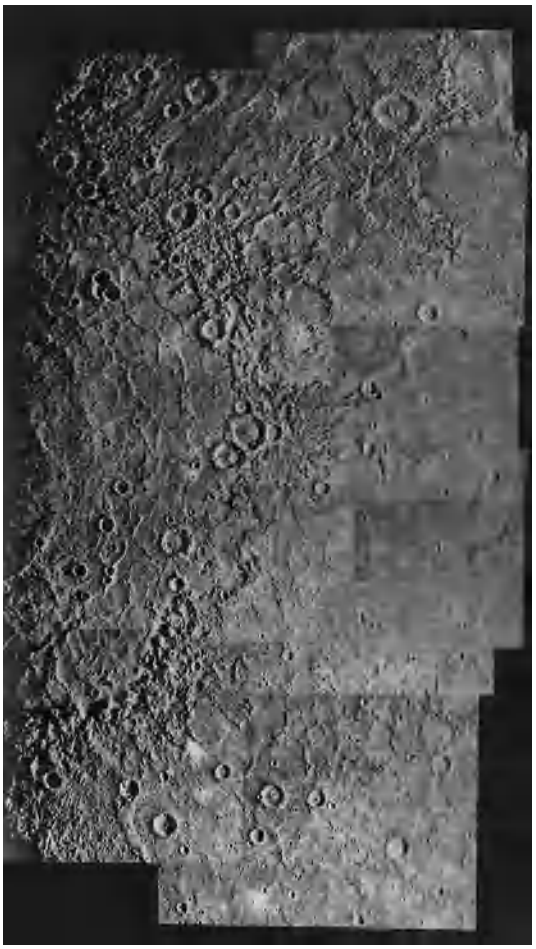

Fig. 2. Photomosaic showing the 800-mi-diameter (1300-km) Caloris Basin, the largest impact structure observed by *Mariner 10*, on the left, midway between top and bottom. (*Jet Propulsion Laboratory*)

Fig. 3. View of the Hilly and Lineated Terrain, taken by *Mariner 10* in March 1974. (*Jet Propulsion Laboratory*)

Mercury's heavily cratered terrain is interspersed with large regions of gently rolling intercrater plains, the major terrain type on the planet. Also unlike the Moon surface, a system of thrust faults, unique in the solar system, transects the surface viewed by *Mariner 10*. The largest structure viewed by *Mariner 10* is the 800-mi-diameter (1300-km) Caloris impact basin (**Fig. 2**). It is filled and surrounded by smooth plains resembling the Moon's maria, which also fill impact basins. Directly opposite (antipodal to) the Caloris Basin, on the other side of Mercury, is a peculiar terrain consisting of hills and valleys, called the Hilly and Lineated Terrain (**Fig. 3**). This terrain is thought to be the result of surface disruption by seismic waves focused at the antipodal region and caused by the Caloris Basin impact. Infrared temperature measurements from *Mariner 10* indicate that the surface is a good thermal insulator and, therefore, must be covered with porous soil or rock powder like the lunar regolith. This is expected on a planet whose surface is shattered and stirred by meteorite impacts.

The heavily cratered terrain records the period of heavy bombardment that occurred throughout the entire inner solar system and ended about 3.8×10^9 years ago on the Moon, and presumably at about the same time on Mercury. The objects responsible for it probably originated among the main-belt asteroids when the inward migration of Jupiter and Saturn resulted in the migration of resonances through the asteroids, ejecting them into the inner solar system.

The oldest plains (intercrater plains) occur in the heavily cratered terrain and are thought to be primarily volcanic deposits erupted at various times during the period of heavy bombardment. The younger smooth plains are also thought to be volcanic deposits that were erupted within or around large impact basins such as Caloris. Based on the size–frequency distribution and density of impact craters superimposed on the smooth plains, they appear to have been formed near the end of heavy bombardment and may be on average about 3.8×10^9 years old. If so, they are, in general, older than the lava deposits that constitute the lunar maria. The unique tectonic framework of Mercury, expressed as a system of thrust faults, is probably the result of global compressive stresses in Mercury's crust caused by a decrease in the planet's radius due to cooling of the core and mantle.

High-resolution radar images of the unimaged side of Mercury show three radar-bright features with different characteristics. Two of the features are fresh impact craters, but the third is different and must await further exploration. The other has no known radar counterpart in the solar system. *See* GEOLOGY; MARS.

Geologic history. The general picture of Mercury's geologic history which has emerged from analyses of *Mariner 10* data is that soon after Mercury formed, it became almost completely melted by heating from the decay of radioactive elements and the inward migration of the large amount of iron to form its enormous core. This led to expansion of the planet

and tensional fracturing of a thin solid lithosphere that provided egress for lavas to reach the surface and form the intercrater plains during the period of heavy bombardment. As the core and mantle began to cool, Mercury's radius decreased by about 0.3 to 1.2 mi (0.5 to 2 km) or more, and the crust was subjected to compressive stresses that resulted in the system of thrust faults. At about this time, the Caloris Basin was formed by a gigantic impact, which caused the Hilly and Lineated Terrain from focused seismic waves in the antipodal regions. Further eruptions of lava within and surrounding the Caloris and other large impact basins formed the smooth plains. Volcanism finally ceased when compressive stresses in the lithosphere became strong enough to close off magma sources. All of these events probably happened very early, perhaps during the first 7–8 $\times$ 10^8 years, in Mercury's history. Since that time, only occasional impacts of comets and asteroids have occurred.

A new mission to Mercury called *MESSENGER* (Mercury Surface, Space Environment, Geochemistry, and Ranging), launched in 2004, is scheduled to orbit the planet in 2011 after two flybys in 2008 and 2009. This mission should answer most of the questions raised by the *Mariner 10* mission and Earth-based observations. *See* PLANET; SPACE PROBE.

Robert G. Strom

Bibliography. J. K. Harmon, Mercury radar studies and lunar comparisons, *Adv. Space Res.*, 19:1487–1496, 1997; R. G. Strom, Mercury: An overview, *Adv. Space Res.*, 19:1471–1485, 1997; R. G. Strom, *Mercury*: *The Elusive Planet*, 1987; Various papers reporting *Mariner 10* results, *Science*, 185:141–180, July 12, 1974, and *J. Geophys. Res.*, 80:2341–2514, June 10, 1975; F. Vilas, C. R. Chapman, and M. S. Matthews (eds.), *Mercury*, 1988.

Mercury-vapor lamp

A vapor or gaseous discharge lamp in which the arc discharge takes place in mercury vapor. This lamp is widely used for roadway and all other forms of illumination, and as a source of ultraviolet radiation for industrial applications.

Construction. The arc discharge takes place in a transparent tube of fused silica, or quartz, and this quartz tube is usually mounted inside a larger bulb of glass (see **illus.**). The outer bulb reduces the ultraviolet radiation of the inner arc tube, encloses and protects the mount structure, and can be coated with phosphors that greatly improve the color of the light emitted.

The arc tube has electrode assemblies mechanically sealed into each end. The main electrodes (one at each end) consist of a molybdenum shank welded to a very thin ribbon of molybdenum and completed by a shank of tungsten wound with tungsten wire and filled with an emission material of metal oxides. At one end of the arc tube, a smaller starting electrode is located close to the main electrode. The starting electrode is connected through a high resistance to the opposite electric polarity of the adjacent main electrode. The arc tube itself is filled with argon gas and a small amount of pure mercury before being sealed.

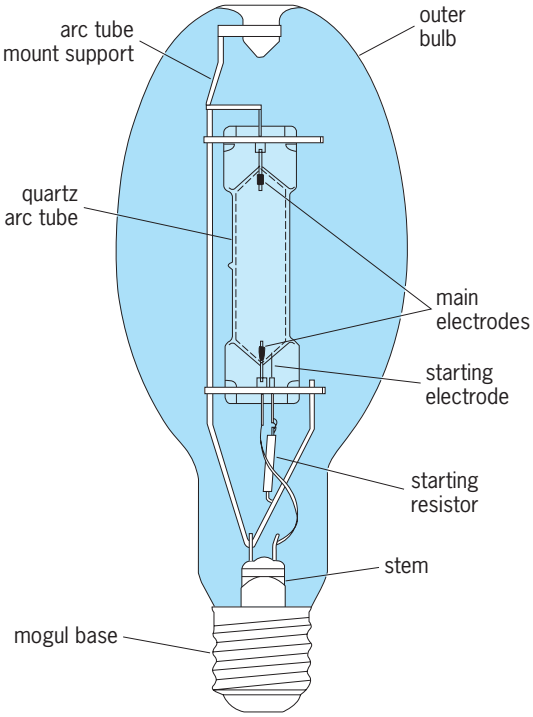

High-pressure mercury-vapor lamp.

The outer bulb supports the mount structure built around the arc tube and is sealed to a glass stem with conducting wires. These wires pass through the outer bulb and are connected to the center post and shell of a metal base.

Starting and operation. Mercury lamps, like other vapor lamps, require sizable voltages to start their arcs, after which the lamp current will continue to increase unless limited by a resistor (which severely reduces efficiency) or a magnetic device (a choke or transformer) called a ballast. The ballast, when energized, supplies proper voltage to the lamp base for starting, and with zero current full starting voltage appears between the main and starting electrode at one end of the lamp. Electrons move across a short gap, ionizing some argon gas in the tube. The ionized gas diffuses until an arc strikes between the main electrodes, where the path now contains less resistance than that between the main and starting electrodes. Once the main arc is formed, the heat from the arc vaporizes the mercury droplets, and they become ionized current carriers. As the lamp current increases to its full current rating, the self-limiting design of the ballast simultaneously reduces the supply voltage to maintain stable operation and extinguishes the glow between starter and main electrode at the reduced voltage. Initial light output requires minutes of warm-up and, similarly, if the ballast is deenergized or power is lost, it is necessary for the lamp to cool and the mercury pressure to be reduced before the starting glow can occur again. *See* ARC DISCHARGE; TRANSFORMER.

Color. Radiation from the mercury arc is confined to four specific wavelengths in the visible portion of

the spectrum and several strong lines in the ultraviolet. Careful design of the arc tube can adjust mercury pressure and cause some shifting of radiation among these wavelengths, thereby changing slightly the color of the illumination.

The visible radiation from clear mercury lamps is all concentrated in the blue, green, and yellow-green regions of the spectrum. This provides light with a distinct blue-green appearance and poor color rendition. Red objects, for example, look brown or black, and human skin looks unattractive. For this reason, most mercury lamps have a phosphor color-correcting coating, and clear bulbs are used only where appearance of colors is secondary to some gain in efficiency.

The phosphors applied to the inner wall of the outer bulb can provide superior color rendition through conversion of the arc tube's ultraviolet radiation to sizable reradiation in the red part of the spectrum or additional yellow, green, and blue-green radiation. One phosphor used for deluxe white lamps has made their use popular in many commercial lighting applications such as department stores and office areas. *See* PHOSPHORESCENCE.

Applications. After the fluorescent lamp, the mercury-vapor lamp is the most popular vapor lamp in use today, with its greatest use in roadway lighting. Many millions of these lamps are produced in the world annually, and despite the fact that other gaseous discharge lamps are more efficient, the long life, reliable performance, and relatively low cost of this lamp and ballast system ensure its use for some years to come.

The sizable radiation of ultraviolet by the bare arc tube suits it well as a sun lamp and as a source for ultraviolet applications in industry, where it is used in reproduction of tracings and to speed chemical reactions such as curing of protective coatings and inks. One wavelength in the near-ultraviolet, not harmful to humans, and transmitted by the bulb glass has the property of causing some dyes and pigments to fluoresce. When the visible light is filtered out and it is used for this fluorescing purpose, it is called a black light.

Consideration should be given to avoiding close proximity of humans to the standard mercury-vapor lamp when the outer bulb is broken since its ultraviolet radiation at close distances can cause erythemal reactions or sunburn, and these standard lamps are now being offered in alternate designs with internal switches or other devices to extinguish the arc if the outer bulb should break. *See* ILLUMINATION; LAMP; METAL HALIDE LAMP; ULTRAVIOLET LAMP; VAPOR LAMP. T. F. Neubecker

Meridian

A line of longitude on Earth or in the sky, perpendicular to the Equator, and extending between the North and South poles.

A meridian of longitude on Earth marks a system of location and timekeeping whose zero was set at Greenwich, England, in the nineteenth century, when it was necessary to regularize time signals in the era of railway transportation. Time is 1 hour different for each 15° of longitude. The International Date Line, where days begin and end, was set 180° away from Greenwich, with jogs for local circumstances. The position of this 180° meridian assumed importance in the discussion of the new millennium, which was widely celebrated on January 1, 2000, in spite of its more accurate definition astronomically as January 1, 2001. Islands and other locations in the Pacific vied for the notoriety of being the first to see the Sun in the new millennium. *See* INTERNATIONAL DATE LINE; LATITUDE AND LONGITUDE; TIME.

Meridians of celestial longitude are the extensions into space of meridians of longitude on Earth. Each celestial meridian is the half of a great circle that goes between the north and south celestial poles and crosses the celestial equator perpendicularly. The unique line known as the celestial meridian is the great circle going from north pole to south pole through a given observer's zenith, the point directly overhead. It thus marks the north-south line, and celestial objects going through it are said to be at their meridian transits or at their culminations. Telescopes known as meridian transits have long been used together with clocks to measure the celestial longitudes of stars, known as their right ascensions. *See* ASTRONOMICAL TRANSIT INSTRUMENT.

Stars sufficiently near the celestial equator transit the meridian once per day, known as upper culmination, and invisibly transit the meridian 180° away in their lower culmination. Circumpolar stars are visible in both upper and lower culminations. *See* ASTRONOMICAL COORDINATE SYSTEMS. Jay M. Pasachoff

Bibliography. J. Mitton, *Cambridge Dictionary of Astronomy*, Cambridge University Press, 2001; J. M. Pasachoff and A. Filippenko, *The Cosmos: Astronomy in the New Millennium*, 3d ed., Brooks/Cole Publishers, Belmont, CA, 2007; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, University Science Books, Mill Valley, CA, 1992.

Meromictic lake

A lake whose water is permanently stratified and therefore does not circulate completely throughout the basin at any time during the year. Normally lakes in the Temperate Zone mix completely during the spring and autumn when water temperatures are approximately the same from top to bottom. In meromictic (*mero*, partial; *mixis*, circulation) lakes, there are no periods of overturn or complete mixing because seasonal changes in the thermal gradient either are small or are overridden by the stability of a chemical gradient, or the deeper waters are physically inaccessible to the mixing energy of the wind. Commonly in meromictic lakes, the vertical stratification in density is stabilized by a chemical gradient.

The upper stratum of water in a meromictic lake is mixed by the wind and is called the mixolimnion.

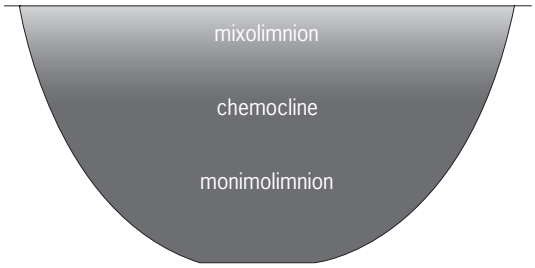

Fig. 1. Cross-sectional diagram of a meromictic lake.

The bottom, denser stratum, which does not mix with the water above, is referred to as the monimolimnion. The transition layer between these strata is called the chemocline (**Fig. 1**).

It is unknown how many meromictic lakes exist. But of the hundreds of thousands of lakes on the Earth, less than 200 have been reported to be meromictic and less than 50 meromictic lakes are reported for North America. Nevertheless, meromictic lakes are widely dispersed across all continents and on islands such as Cuba, the Galápagos and Tasmania. Zige Tangco at 4560 m (15,000 ft) above sea level in the central Tibetan Plateau is the highest known meromictic lake in the world. In North America, meromictic lakes generally are restricted to sheltered basins with proportionally very small surface areas in relation to depth, basins in arid regions, and isolated basins in fiords. Meromixis may occur in lakes with a permanent ice cover. Meromictic lakes frequently contain colored water, which limits penetration of solar radiation. *See* LAKE.

Interest in meromictic lakes is disproportional to their number, not only because they are oddities of nature but because they offer unique opportunities for limnological research. Research activity on meromictic lakes has been focused on the physiology and behavior of organisms, biogeochemistry, studies of lake history, deep-water circulation, and heat flow through the bottom sediments.

Chemical studies. Relatively few detailed chemical studies have been done on meromictic lakes, particularly on a seasonal basis. The most typical chemical characteristic of meromictic lakes is the absence of dissolved oxygen in the monimolimnion. Large quantities of hydrogen sulfide and ammonia may be associated with this anaerobic condition in deep water. J. Kjensmo proposed that the accumulation of ferrous bicarbonate in the deepest layers of some lakes may have initiated meromixis under certain conditions. Separation of salts from surface water as the water freezes, and then the accumulation of these salts in deeper waters, also has led to meromixis in Antarctic lakes. Some human activities have caused meromixis in a few cases. For example, runoff containing salt used for deicing nearby roads can produce permanent stratification in small lakes.

Meromictic lakes are exciting model systems for many important biogeochemical studies. The isolation of the monimolimnetic water makes these studies quite interesting and important. In studies of the biological fractionation of sulfur and carbon isotopes in the monimolimnion of Fayetteville Green Lake, New York, the fractionation factor for sulfur was found to be the highest ever observed.

Sediments from meromictic lakes are among the best for studies of lake history, since there is relatively little disturbance or decomposition of biogenic materials. Such sediments may provide a detailed record of the origin, ecological changes, and physical changes of meromictic lakes and their drainage basins over time. Also, deuterium has been used by several workers in an attempt to unravel the history of meromictic lake water.

Physical studies. Radioactive tracers have been used to show that the monimolimnetic water in a small meromictic lake (Stewart's Dark Lake) in Wisconsin is not stagnant but undergoes significant horizontal movement (**Fig. 2**). While the maximum radial spread was about 16–18 m (52–59 ft) per day, vertical movements were restricted to negligible amounts because of the strong vertical density gradient. A similar pattern of movement was reported in the deep water of a larger meromictic lake (Soap Lake) in Washington. It was determined from these studies that the average horizontal eddy diffusion coefficient is 3.2 cm^2/s (0.50 $in.^2/s$) in the monimolimnion of Soap Lake and 17 cm^2/s (2.6 $in.^2/s$) in Stewart's Dark Lake.

Because vertical mixing is restricted in a meromictic lake, heat from solar radiation may be trapped in the monimolimnion and thereby can produce an anomalous temperature profile. Monimolimnitic water temperatures have been reported as high as 50.5°C (123°F) at a depth of 2 m (7 ft) in shallow Hot Lake, Washington.

Some meromictic lakes are very convenient for studies of geothermal heat flow because deep-water temperatures may be nearly constant year-round. Studies of terrestrial heat flow have been made in the sediments of Stewart's Dark Lake, Wisconsin, and in

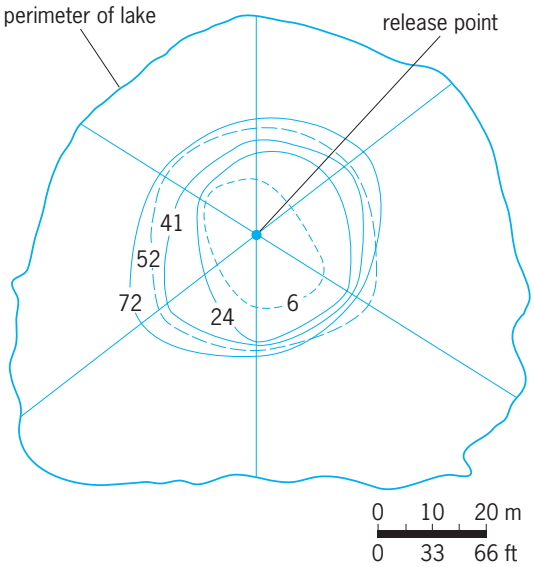

Fig. 2. The outlines show the maximum horizontal displacement of a radiotracer (sodium-24) following its release at 8-m (26-ft) depth in Stewart's Dark Lake, Wisconsin. The numbers indicate the hours elapsed after release.

Fayetteville Green Lake, New York. Steady-state thermal conditions were found in the sediments of Stewart's Dark Lake, and the total heat flow was calculated to be 2.1×10^{-6} cal/cm^2 s (8.8×10^{-2} W/m^2). However, about one-half of this flux was attributed to the temperature contrasts between the rim and the central portion of the lake's basin. It was found that the thermal conductivity of the surface sediments in the center of Stewart's Dark Lake was 1.10×10^{-3} cal · cm/ (cm^2 · s · °C) [0.46 W · m/(m^2 · °C)], a value substantially lower than that for pure water at the same temperature but consistent with measurements on colloidal gels.

Biological studies. Relatively few kinds of organisms can survive in the rigorous, chemically reduced environment of the monimolimnion. However, anaerobic bacteria and larvae of the phantom midge (*Chaoborus* sp.) are common members of this specialized community. By using sonar, vertical migrations of *Chaoborus* sp. larvae have been observed in meromictic lakes. The sound waves are reflected by small gas bladders on the dorsal surface of the larvae. The migration pattern is similar to that shown by the deep-scattering layers in the sea. The larvae come into the surface waters at night when the light intensity is low, and sink into deeper waters during the daylight hours. Biologically significant quantities of chemical nutrients or pollutants can be transported from the deep and relatively inaccessible part of a meromictic lake to the surface and thence to the adjacent terrestrial environment by these organisms. A radiotracer (iodine-131) appeared in flying adult *Chaoborus* sp. along the shoreline of the lake within 20 days after it had been released within the region of the chemocline.

Specialized bacteria, such as the photosynthetic, green sulfur or purple sulfur bacteria (*Chromatium* spp.), utilize the unique characteristics of the upper monimolimnion. Often, these bacterioplanktonic populations become dense enough to impart color to the water.

Spectacular degassing events from the deep waters of two tropical meromictic lakes (Lakes Nyos and Monoun) in Cameroon, West Africa, have occurred on occasion during late summer when vertical stability is reduced due to weather conditions and mixing penetrates deeply in the water column. Large releases of carbon dioxide from Lake Monoun (August 15, 1984) and from Lake Nyos (August 21, 1986) killed over 1700 people living nearby. *See* BIOGEOCHEMISTRY; FRESHWATER ECOSYSTEM; LIMNOLOGY; SONAR. Gene E. Likens

Bibliography. G. J. Brunskill et al., Fayetteville Green Lake, New York, I: Physical and chemical limnology, II: Precipitation and sedimentation of calcite in a meromictic lake with laminated sediments, III: The laminated sediments, IV: Interstitial water chemistry of the sediments, V: Studies of primary production and zooplankton in a meromictic marl lake, *Limnol. Oceanogr.*, 14(6):817–873, 1969; E. S. Deevey, N. Nakai, and M. Stuiver, Fractionation of sulfur and carbon isotopes in a meromictic lake, *Science*, 139:407–408, 1963; L. J. Garcia-Gil, X. Casamitjana, and C. A. Abella, Comparative study of two meromictic basins of Lake Banyoles (Spain) with sulphur phototrophic bacteria, *Hydrobiologia*, 319(3):203–211, 1996; J. A. E. Gibson, The meromictic lakes and stratified marine basins of the Vestfold Hills, East Antarctica, *Antarctic Sci.*, 11(2):175–192, 1999; N. M. Johnson and G. E. Likens, Steady-state thermal gradient in the sediments of a meromictic lake, *J. Geophys. Res.*, 72:3049–3052, 1967; G. W. Kling, Seasonal mixing and catastrophic degassing in tropical lakes, Cameroon, West Africa, *Science*, 237:1022–1024, 1987; K. F. Walker and G. E. Likens, Meromixis and a reconsidered typology of lake circulation patterns, *Verh. Int. Verein. Limnol.*, 19:442–458, 1975.

Merostomata

A class of phylum Arthropoda, subphylum Chelicerata. Merostomes are aquatic chelicerates, characterized by abdominal appendages bearing respiratory organs. Most merostomes are extinct; only the horseshoe crabs, comprising four species and three genera (*Carcinoscorpio* and *Tachypleus* of eastern Asia and *Limulus polyphemus* of eastern North America) survive. The body of a merostome consists of a prosoma, or head, which lacks antennae, has a pair of compound eyes and a pair of simple median eyes, and bears the chelicerae (pincers) and five pairs of uniramous walking legs with gnathobases for mastication. The opisthosoma, or trunk, consists of 12 or fewer segments which may be freely articulating or partly or entirely fused into a solid shield; the opisthosoma bears the respiratory appendages. The telson (tail) is a solid, usually spikelike, structure. *See* CHELICERATA.

A modern classification of the Merostomata follows:

Phylum Arthropoda
- Class Merostomata
  - Subclass Eurypterida
    - Order: Eurypteridida
    - Chasmataspida
  - Subclass Xiphosura
    - Order: Aglaspida
    - Xiphosurida
      - Suborder: Synziphosurina
      - Limulina

The scorpionlike Eurypteridida (Ordovician–Permian) was the most diverse merostome group, inhabiting a variety of nearshore marine and freshwater habitats. Eurypterids were the largest known arthropods, reaching lengths of about 6 ft (2 m) or more in some genera. The prosomal appendages were modified into a variety of shapes for walking, swimming, capturing prey, and perhaps burrowing. *See* EURYPTERIDA.

The Chasmataspids (Ordovician–Devonian) were a heterogeneous group of eurypteridlike merostomes of uncertain affinity.

The Aglaspida (Cambrian–Ordovician) are generally considered to be the most primitive of all known merostomes. Their body form was simple (**Fig. 1***a*),

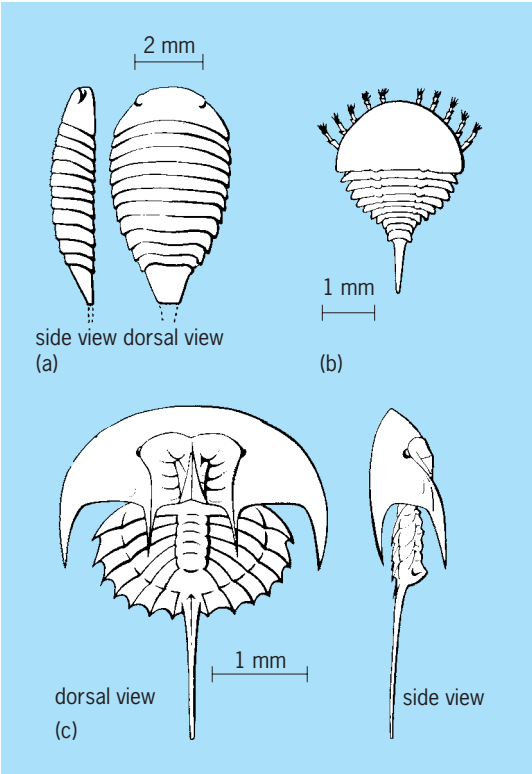

Fig. 1. Some typical extinct species of the class Merostomata. (*a*) *Paleomerus*, order Aglaspida. (*b*) *Weinbergina*, suborder Synziphosurina. (*c*) *Euproops*, order Xiphosurida. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, Geological Society of America, 1955*)

with a small prosoma with compound eyes, an opisthosoma with 11 or 12 freely articulating segments, and a simple telson. The aglaspid cuticle was usually phosphatic, in contrast to the chitinous cuticle of other merostomes. A few species have been described with walking legs (instead of gill-bearing appendages) on the opisthosoma.

The Xiphosurida appeared in the Silurian with certainty, although Cambrian and Ordovician forms may be early representatives of this group. The Xiphosurida are united by a number of characters, including a prosoma and an opisthosoma of nearly equal length and the presence of an ophthalmic ridge, a linear elevation on either side of the prosoma, bearing the lateral compound eyes (Fig. 1*c*). *See* XIPHOSURIDA.

Fig. 2. *Legrandella*, from the Devonian of Bolivia. Note the crescentic ophthalmic ridges and the anterior four freely articulating opisthosomal segments.

The most primitive xiphosuridans, the Synziphosurina (Silurian–Devonian), until recently were poorly known. Their prosomal anatomy was like that of true horseshoe crabs, but they retained the free opisthosomal segments of the primitive merostomes (Fig. 1*b*). Synziphosurans were not blind, as previously supposed. Compare Fig. 1*b*, an old restoration of the synziphosuran *Weinbergina*, with **Fig. 2**, a photograph of *Legrandella*, by far the best-preserved synziphosuran yet found, which possesses ophthalmic ridges, eyes, and a number of other features previously thought to be absent in the Synziphosurina.

The Limulina (Devonian–Recent) are the true horseshoe crabs. Paleozoic representatives lived in both nearshore marine and fresh-water habitats and differed from modern limulines chiefly by their smaller size. During the Mesozoic and Cenozoic, the limulines became larger, whereas the group as a whole became restricted in terms of the number of different species living at any one time; the four species alive today probably represent the typical amount of limuline diversity for at least the past 100,000,000 years. Their survival for so long in relatively few numbers and in a relatively unchanged state qualifies them as living fossils. *See* LIVING FOSSILS.

Niles Eldredge

Bibliography. N. Eldredge, Revision of the suborder Synziphosurina (Chelicerata, Merostomata), with remarks on merostome phylogeny, *Amer. Mus. Novitates,* 2543:41, 1974.

Mesogastropoda

The largest and most diverse order of gastropod mollusks in the subclass Prosobranchia. Mesogastropods are all single-gilled and thus monotocardiac, that is, with only one auricle and asymmetry of other cardiac, renal, and genital structures. Except for the minor superfamily Valvatacea, which retains an aspidobranch or bipectinate "feather" gill like those of the Archaeogastropoda, all mesogastropods have a one-sided comb-shaped or pectinibranch ctenidium or "half-gill" with its axis and blood vessels fused to the mantle wall. This results in a simplified mantle cavity that is asymmetric and hydraulically more efficient than in the Archaeogastropoda, and from which both feces and nitrogenous excretory wastes are discharged to the right side of the snail's head-foot. Such a system can operate in silty waters and, unlike the archaeogastropods which are limited to clean waters over hard substrata, mesogastropods with a wide variety of marine life-styles are found burrowing in soft muds and living exposed to air near high-tide level. There are also estuarine, fresh-water, and terrestrial species in the order. The anatomical asymmetry makes possible a functional separation of genital from renal ducts, which in turn allows the development of internal fertilization, of large

eggs, of ovoviviparity, and of true viviparity (all features advantageous for the evolution of nonmarine stocks).

Certain older classifications designate the mesogastropods as order Taenioglossa, referring to the intermediate type of radula characteristic of this group. The taenioglossan radula is narrower and normally has only seven teeth in each row. Mesogastropods are structurally diverse, and most classifications divide them into 13 superfamilies, 4 of which number several thousand species and encompass some of the world's most abundant, ubiquitous, and cosmopolitan marine snails.

Major groups and species. On the basis of neural anatomy, the large superfamily Cyclophoracea (mainly family Viviparidae) is regarded as comprising the most primitive living mesogastropods, but it is in other ways a highly specialized nonmarine group with its genera well adapted for life in freshwaters and on land (mostly in humid tropical habitats). Similarly, the small superfamily Valvatacea has the primitive gill noted above, but again its members are fresh-water animals, including several ovoviviparous species. One of the largest superfamilies of mollusks is the Rissoacea, encompassing more than 1350 living species and vast numbers of individual snails, mostly tiny and highly motile, in a variety of marine, estuarine, and freshwater habitats. Another huge superfamily is the Cerithacea (at least 1300 living species), dominating the sublittoral marine muds in many parts of the world, mainly with elongate pointed shells like the genus *Turritella*, but also including tropical estuarine and freshwater forms. The smaller superfamily Strombacea includes the pelican's-foot shells such as *Aporrhais*, and *Strombus gigas*, the queen or true-pink conch of the Florida Keys. Another small superfamily, Calyptraeacea, includes relatively sessile microphages such as *Capulus* and *Crepidula*, in which the individual leaflets of the pectinibranch gill have been elongated into filaments for use in filter feeding (paralleling the evolution of the eulamellibranch bivalves).

Among the best-known and most collected tropical sea shells are the 400 species of cowries (superfamily Cypraeacea) with glossy colorful ovate and convolute shells having a long narrow aperture bounded on both sides by inrolled toothed lips. In living cowries, the expanded mantle folds completely conceal the shell, and it is this that results in the outside (as well as the inside) of the shell being "varnished" with a nacreous (pearly) layer. In several parts of the world, cowries have been used as currency. Elsewhere, their resemblance to human female genitalia has made them important in fertility rituals and in the worship of Aphrodite and similar goddesses. Superfamily Heteropoda includes forms like the atlantids, living pelagically in the open ocean with reduced fragile shells and various methods of flotation. Superfamily Ptenoglossa (or Scalacea) comprises both the shell collectors' prized wentletraps and the ianthinids which float on the ocean surface by using a mass of bubbles of hardened mucus.

Possibly the most successful of all seashore animals are the 800 species of the Littorinacea, including several high-littoral air-breathing species of periwinkles, and some tropical terrestrial forms. In the middle levels of the intertidal zone, littorinid snails are alternately aerial and aquatic in habit, and several periwinkle species have both lungs and gills. Other aspects of littorinid physiology, including reproduction and nitrogenous excretion, are involved in the trends for those snail species living at the higher levels of the seashore to be better adapted for aerial or terrestrial conditions.

The superfamily Naticacea comprises the moon snails, about 200 species of predaceous carnivores, moving through sand and mud and preying on bivalves and other snails. Naticids employ a protrusible proboscis and attached boring gland to drill through the shells of their prey by using a combination of mechanical and chemical processes. Larger naticids also use an unusual series of pedal water sinuses to inflate the enormous foot (used not only to move rapidly through sand but also to wrap around the clam prey).

Evolution. In the evolution of the gastropod mollusks, mesogastropods clearly constitute an intermediate grade. It is possible that the order Mesogastropoda is not monophyletic, and that stocks with the asymmetric efficiency of structures which is diagnostic are derived from several archaeogastropod stocks. However, it is clear that the several lines of "higher" gastropods, including not only prosobranchs of the order Neogastropoda but also nonprosobranchs in the subclasses Opisthobranchia and Pulmonata, have each been derived from stocks within the order Mesogastropoda. *See* GASTROPODA; NEOGASTROPODA; PROSOBRANCHIA.

W. D. Russell-Hunter

Bibliography. V. Fretter and A. Graham, *British Prosobranch Molluscs*, 1962; W. D. Russell-Hunter, *A Life of Invertebrates*, 1979; K. M. Wilbur (ed.), *The Mollusca*, vols. 1–12, 1983–1987.

Meso-ionic compound

A member of a class of five-membered ring heterocycles (and their benzo derivatives) which possess a sextet of π electrons in association with the atoms composing the ring but which cannot be represented satisfactorily by any one covalent or polar structure.

Structure. Two main types, depending formally upon the origin of the electrons in the π system, have been identified; they are exemplified by compounds (**1**) and (**2**). In structure (**1**) the nitrogen and

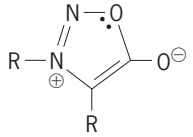

(**1**)

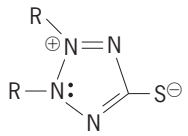

(**2**)

oxygen atoms, 1,3 to each other, are shown as donating two electrons each to the total of eight electrons in the whole π system, whereas in structure (**2**) the two middle nitrogen atoms, 1,2 to each other, are the two-electron donors.

The term "satisfactorily" in the definition refers to the fact that the charge in the ring cannot be associated exclusively with one ring atom. Thus, these compounds are in sharp contrast with other dipolar structures, such as ylides [for example, compound (**3**)], and such compounds are not considered meso-ionic. Meso-ionic compounds are most commonly represented as follows: compound (**1**) as structure (**4**) and compound (**2**) as structure (**5**). The circle represents six π electrons; the positive charge is shared by all the ring atoms.

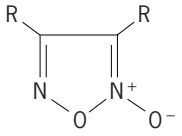

(**3**)

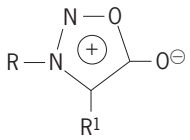

(**4**)

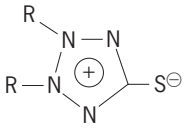

(**5**)

Preparation. Compounds of type (**1**), containing a 1,2,3-oxadiazole nucleus, were the first discovered members of this class and are known as sydnones. They are made by the action of acetic anhydride on the *N*-nitroso derivatives of *N*-substituted α-amino acids [reaction (1)].

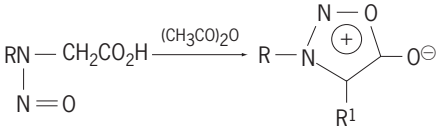

(1)

($R^1 = H$)

Other representatives of this type contain 1,3-oxazole, imidazole, thiazole, 1,3-dithiole, 1,3,4-oxadiazole, 1,3,2-oxathiazole, 1,2,3- and 1,2,4-triazole, and 1,3,4- and 1,2,3-thiadiazole nuclei. There are also derivatives with four hetero atoms.

Compounds of type (**2**) are known as dehydrodithizones and were actually the first meso-ionic compounds discovered. They are made by the oxidation of the acyclic dithizones (**6**) with manganese dioxide [reaction (2)] or isoamyl nitrite. Other nuclei present in compounds related to (**5**) are isoxazole, pyrazole, isothiazole, and 1,2-dithiole.

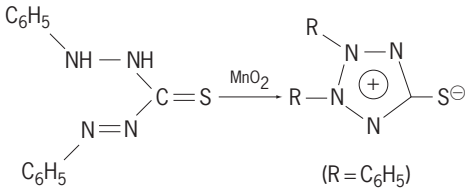

(2)

(**6**) ($R = C_6H_5$) (**5**)

Reactions. The most important reactions of meso-ionic compounds are dipolar cycloaddition reactions in which the meso-ionic compound reacts with alkenes and acetylenes. Sometimes these reactions involve their acyclic tautomers [reaction (3)]. They also are extremely reactive toward nucleophiles, even as mild as water, yielding acyclic products.

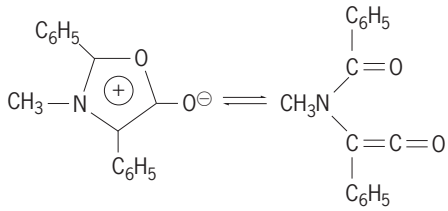

(3)

Meso-ionic compound — Acyclic tautomer

Photochemically, they are highly reactive and extensive rearrangement and degradation has been observed. For example, sydnones can be rearranged to oxadiazolinones, as in reaction (4), where R^2 is C_6H_5 or H.

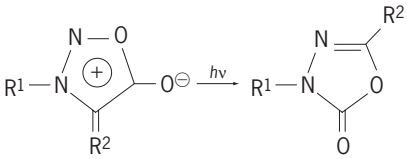

(4)

See PHOTOCHEMISTRY.

Uses. There has been considerable interest in the pharmacological activity of meso-ionic compounds, and derivatives have shown a variety of antibiotic, anthelminthic, antidepressant, and anti-inflammatory properties. *See* HETEROCYCLIC COMPOUNDS.

Jeremiah P. Freeman

Bibliography. J. A. Joule, *Heterocyclic Chemistry*, 4th ed., 2000; A. R. Katrizky (ed.), *Advances in Heterocyclic Chemistry*, vol. 48, 1990; E. C. Taylor and A. R. Katritzky (ed.), *Advances in Heterocyclic Chemistry*, vol. 19, 1975.

Mesometeorology

That portion of meteorology comprising the knowledge of intermediate-scale atmospheric phenomena, that is, in the size range of approximately 1–1200 mi (2–2000 km) and with time periods typically less than 1 day. In addition to these time and space criteria for defining mesoscale, dynamical considerations can be used. For example, a dynamical definition requires that both background rotation and ageostrophic advection be present to a significant degree. These requirements are based upon the relative magnitudes of the two major forces, pressure gradient and Coriolis, that govern movement of the atmosphere. To the extent that these forces are not in balance, air is accelerated. When the forces are of comparable magnitude, accelerations are relatively small and the atmosphere is considered to be in

geostrophic equilibrium. The wind, under these so-called balanced conditions, is called the geostrophic wind. For weather systems larger than mesoscale, the atmosphere is predominantly geostrophic. However, for mesoscale systems, geostrophic balance is not routinely present and accelerations can be quite large. Thus, the wind departs from geostrophic balance, and the departure is termed the ageostrophic wind. The term total wind can then be thought of as the sum of the geostrophic and ageostrophic components, and each component can advect (move, carry, transport, and so on) various properties or constituents of the atmosphere. Therefore, one of the clearly definable dynamic differences between large and mesoscale systems is that ageostrophic advection is negligible for large-scale circulations, while for mesoscale circulations it is not. A convenient way of measuring this condition is through the nondimensional Rossby number (R). This number is the ratio of the net acceleration of the horizontal wind by all the forces to the acceleration produced by the Coriolis force alone. For mesoscale motions, R is on the order of 1, as shown below.

$$R \equiv \frac{\text{net of accelerations by all forces}}{\text{acceleration by Coriolis force}} \sim 1.0$$

For progressively larger weather systems, accelerations tend to become smaller and R decreases. As R approaches the order of 0.1, weather systems are considered synoptic scale instead of mesoscale. *See* DIMENSIONLESS GROUPS.

Observing mesoscale phenomena. Over midlatitude landmasses, the average spacing between atmospheric sounding stations that use balloon-borne instrument packages is about 180–360 mi (300–600 km) and the standard sounding frequency is twice each day. Consequently, only the largest mesoscale phenomena, with wavelengths greater than about 600 mi (1000 km), are routinely observed (resolved) by this network. Information with higher

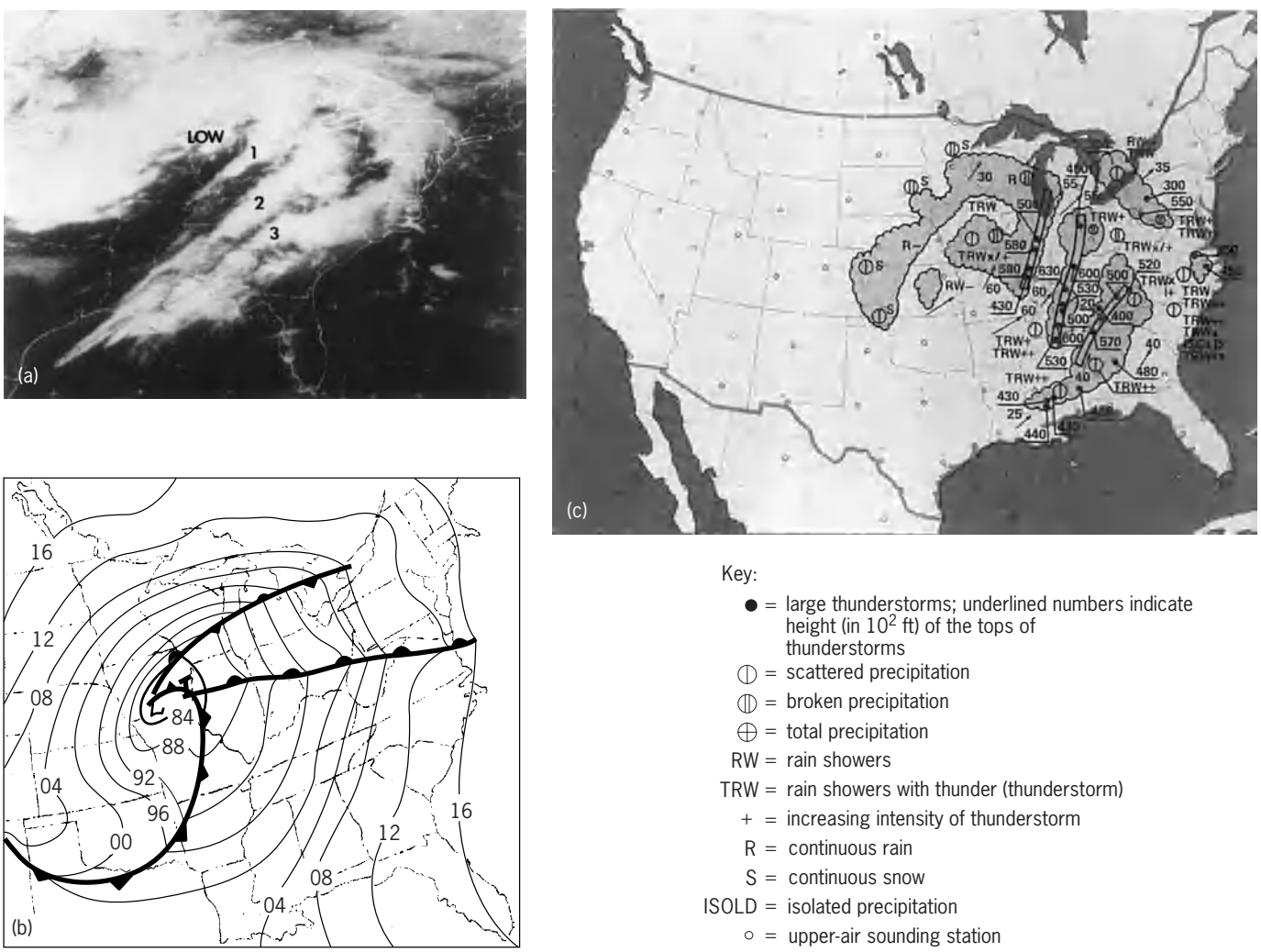

Fig. 1. Mesoscale weather systems. (*a*) Satellite image for 2100 UTC (4:00 PM EST), April 3, 1974, showing the three major squall lines responsible for the greatest tornado outbreak in United States history. (*b*) Graphic of squall lines. (*c*) Easten United States showing detailed weather. (*After L. R. Hoxit and C. F. Chappell, Tornado outbreak of April 3–4, 1974, NOAA Tech. Rep. ERL 338-APCL 37, Department of Commerce*)

time and space resolution is available from aircraft observations, such as those taken by the Aircraft Communications, Addressing and Reporting System (ACARS), and networks of radar stations, profilers, and satellite imagery. The profiler is a ground-based hybrid observing system of vertically pointing radar and microwave radiometry. Unlike the discrete (in time) observations provided by most other observing systems, profilers provide continuous depictions of the vertical profiles of wind, temperature, and humidity. However, only a limited network of profilers (situated mostly over the midwestern United States) is operational. In contrast, satellite and Doppler radars cover virtually all areas of the United States and provide detailed observations of many mesoscale phenomena.

The satellite and radar observations often show that mesoscale weather systems are distinct components of large synoptic-scale cyclones (low-pressure systems) and anticyclones (high-pressure systems). For example, the satellite image in **Fig. 1** shows a large, contiguous area of clouds associated with a synoptic-scale low-pressure system. Within this large cloud mass, three distinct mesoscale weather systems are clearly apparent over the eastern portion of the United States. These systems, squall lines, are composed of lines of thunderstorms acting in a unified manner (individual thunderstorms are discernible in the easternmost squall line) and have characteristic scales of approximately 150 mi (250 km)—much smaller than the resolution of the standard sounding network. *See* CYCLONE; REMOTE SENSING; SQUALL LINE; THUNDERSTORM; WEATHER MAP.

Although satellite imagery and radar data provide extensive areal coverage and clearly reveal the presence of mesoscale systems, they do not provide adequate vertical resolution of certain atmospheric parameters (such as temperature, moisture, and pressure) so that the mesoscale structures and circulations can be examined quantitatively and fully understood. Thus, while mesoscale phenomena can be "observed," they can not be studied and predicted (in the conventional manner) as easily as synoptic-scale systems. *See* METEOROLOGICAL SATELLITES; RADAR METEOROLOGY; WEATHER FORECASTING AND PREDICTION.

Mesoscale weather systems. These are responsible for many of the significant weather events that routinely occur in the midlatitudes and the tropics. These systems naturally fall into two categories: free- and forced-mode systems. As considered here, forced-mode systems most frequently arise from interactions between the atmosphere and the Earth's topography. Free-mode systems develop in situ from the background (large-scale) environment. The **table** lists some commonly recognized free- and forced-mode mesoscale phenomena and their typical wavelength or other appropriate dimension. These phenomena include terrain modulation of precipitation, terrain-forced gravity waves, lake-effect snowstorms, convective complexes, and rainbands.

Terrain modulation of precipitation. Since the absolute humidity of an air parcel tends to increase with temperature, decreasing the temperature of a saturated parcel results in net condensation of water vapor, that is, the formation of clouds. Futhermore, since air parcels expand and cool when they ascend, and compress and warm when they descend, clouds and precipitation are likely to develop or intensify as air ascends mountainous terrain, and to weaken or dissipate as air descends. Thus, for mountainous areas of the world, the distribution of precipitation is strongly modulated by mesoscale terrain features. For example, **Fig. 2***a* shows the local terrain for a mesoscale region in the western United States, where east-west-oriented ridges are separated by river valleys. If forecasting were being done for the single synoptic-scale forecast zone using only synoptic-scale weather features as a guide, the weather would be considered to be essentially the same throughout, irrespective of the mesoscale terrain effects. However, as is apparent from Fig. 2*a*, when the wind is from the north, air alternately ascends the ridges and then descends into the valleys, causing clouds and precipitation over the ridges and clearing in the valleys. Such a situation is shown in Fig. 2*b*, where a large (synoptic-scale) storm is moving east through the Rocky Mountain region. The counterclockwise circulation around the low-pressure center is producing low-level northerly flow across the mesoscale region shown in Fig. 2*a*. There are alternating cloudy and clear areas over the ridges and valleys, respectively. In this particular

Mesoscale phenomena

Phenomena	Wavelength, mi (km)
Force mode	
Gravity waves (amplifying mountain lee-waves)	1.2–1200 (2–2000)
	6–60 (10–100)
Lake-effect snowstorms	6–60 (10–100)
Mountain-valley circulations	Twice distance from mountain crest to valley floor
Sea breeze	6–60 (10–100)
Terrain modulation of precipitation	Scale of terrain feature
Free mode	
Precipitation bands (moist symmetric instability)	6–60 (10–100)
Convective complexes	600–1200 (1000–2000)
Jet streaks	600–1200 (1000–2000)
Polar lows	600–1200 (1000–2000)
Squall lines	12–20 (20–200)
Tropical storms and hurricanes	1200 (2000)

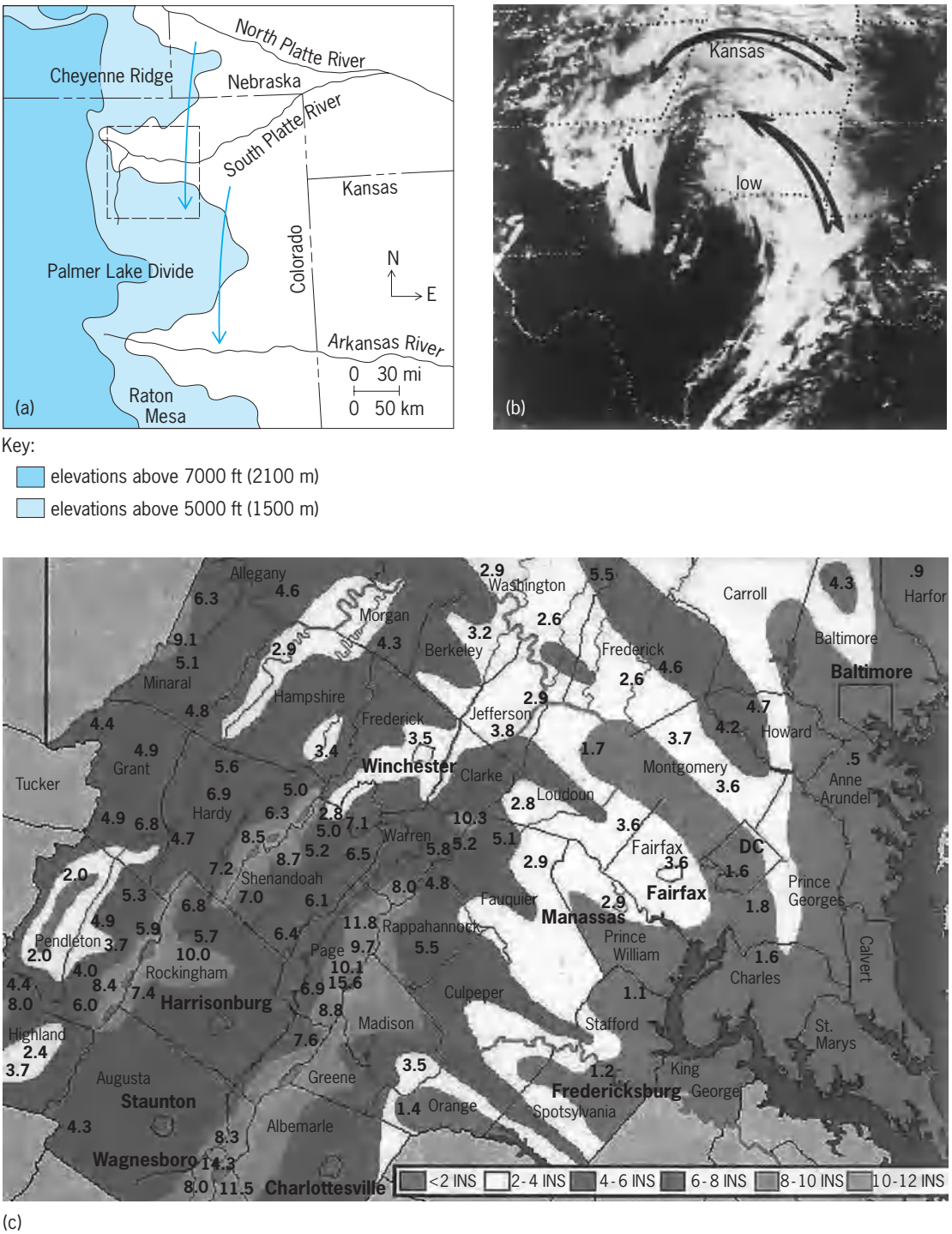

Fig. 2. Terrain modulation of precipitation. (*a*) Mesoscale terrain features (shaded). Arrows indicate direction of mean low-level airflow. Boxed area indicates typical size of a synoptic-scale forecast zone. (*b*) Visible *GOES* satellite imagery for November 8, 1977. Large arrows indicate low-level cyclonic flow around synoptic-scale low-pressure center. (*c*) Rainfall (in inches) from tropical storm Fran on September 6, 1996 (*adapted from Barbara Watson, NOAA*).

instance, well over a foot of snow fell on the Palmer Lake Divide with little or no snow reported in the river valleys on either side. Similarly, the rainfall from tropical storm Fran was modulated by the mesoscale terrain features of the Appalachian Mountains of Virginia. Note in Fig. 2*c* how rainfall totals vary from over a foot (>30 cm) along the ridges to only a few inches in the valleys. Such mesoscale terrain effects are so strong in mountainous areas of the world that average annual precipitation isohyets approximately parallel the terrain-height contours. *See* ATMOSPHERE; CLOUD; CLOUD PHYSICS; PRECIPITATION (METEOROLOGY).

Terrain-forced gravity waves. As an air parcel encounters a mountain barrier and is displaced upward, it undergoes changes in temperature, pressure, and density

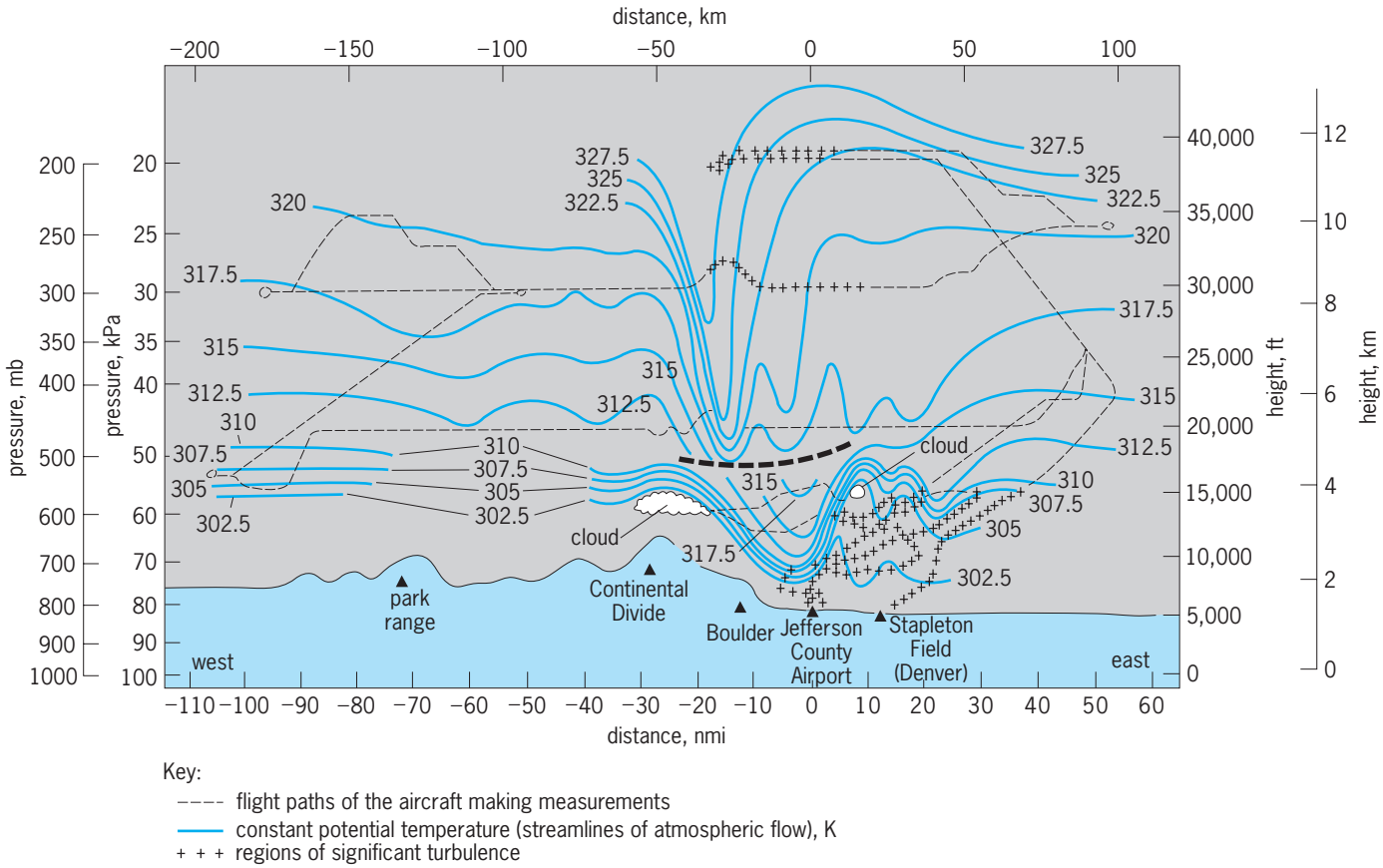

Fig. 3. Cross section of potential temperature (K) on January 11, 1972, on an east-west line through Boulder, Colorado. The heavy broken line separates measurements taken at two different times. °F = 1.8 K − 459.67. (*After D. K. Lilly, A severe downslope windstorm and aircraft turbulence event induced by a mountain wave, J. Atm. Sci., 35:59–77, 1978*)

relative to its surroundings. The parcel is therefore subject to pressure and gravitational forces which act to restore the parcel toward its initial equilibrium position. The undulation produced by alternate displacement followed by restoration, where gravity plays an important role, is commonly referrred to as a gravity wave. Propagation of gravity waves away from a mountain barrier indicates energy being transmitted through the air away from the mountain. The propagation of gravity waves is dependent upon the characteristics of the medium (in this case, air) through which it propagates. If the wave encounters a zone where there is a locally sharp change in the physical characteristics (such as a change in density) of the medium, not all of the energy of the wave may be transmitted into the new medium—some energy may be reflected. In certain situations, the atmosphere is composed of layers of air with quite different characteristics so that a portion of the energy with upward-propagating gravity waves is reflected downward. Depending upon the thickness of these layers and the height of the mountain, reflected energy may return "in phase" with subsequent original waves—this is commonly called resonance. In these instances, the amplitudes (energy) of the original and reflected waves are additive, and the wave grows. A mountain-forced gravity wave (mountain wave) under "ideal" amplification conditions may grow to tremendous heights and produce extreme winds [often in excess of 100 mi (160 km) per hour and, on occasion, as fast as 140 mi (240 km) per hour] in the lee of the mountain barrier. **Figure 3** shows a cross section through such a wave.

In addition to the obvious threat that such amplifying waves and attending severe winds pose to the general public, a second impact area is aviation. Downslope windstorms are capable of destroying light aircraft on the ground, make takeoffs and landings exceedingly hazardous, and generate downdrafts from which small aircraft flying over mountainous terrain may not escape. Pilot reports of vertical currents exceeding 3000 ft (900 m) per minute during these storms are not uncommon. On occasion, vertical currents have been reported to be as strong as about 7900 ft (2400 m) per minute. *See* AERONAUTICAL METEOROLOGY.

Mountain waves and their associated downslope winds occur along the lee slopes of major mountain ranges of the world. The winds are called *foehn* in Europe, *zonda* in Argentina, *puelche* in the Andes, *favogn* in Switzerland, and Canterbury Northwestern in New Zealand. The name varies within the continental United States: chinook along the Rocky Mountains, mono in the Sierras, and Santa Ana along the Coastal Range in southern California. *See* CHINOOK.

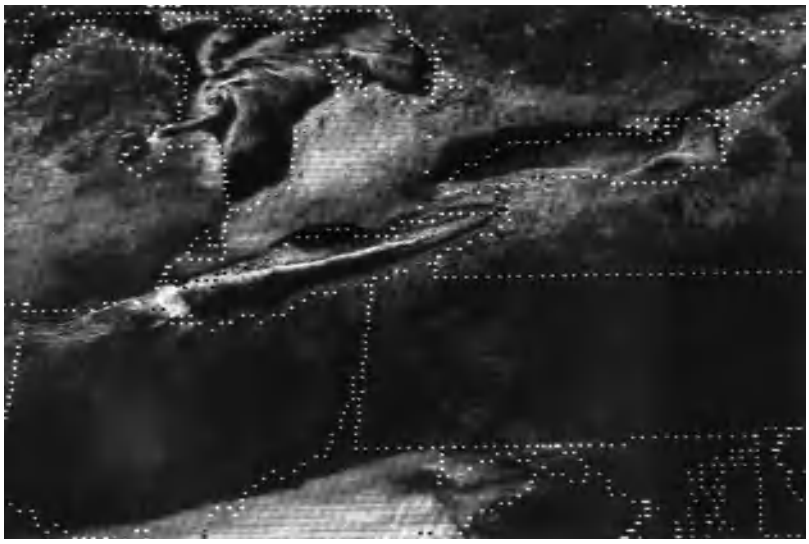

Fig. 4. Visible *GOES* satellite imagery for 1800 UTC (1:00 PM EST), December 20, 1983.

Lake-effect snowstorms. In the fall and early winter, the cooling of large lakes lags behind the cooling of the atmosphere. Consequently, situations develop where very cold air masses pass over relatively warm water. In these instances, large fluxes of heat and moisture from the lake to the air cause the atmosphere to become unstable in an absolute sense; that is, the temperature lapse rate exceeds the dry adiabatic lapse rate and the atmosphere must overturn. This overturning occurs in an organized manner over the lake and the downwind shoreline. Over the lake, the thermally driven overturning often organizes into mesoscale bands of ascending air oriented parallel to the prevailing wind. However, as the air flows from the smooth lake surface to the rough surface of the land, the increase in surface friction causes mass to pile up along the shoreline and forces air upward. Thus, frictionally driven bands tend to form parallel to the shoreline but at an angle to the prevailing wind. These shoreline bands are often enhanced by local terrain features. For a northeasterly flow event, both a lake band and a shoreline band are visible in **Fig. 4**. In most instances, the bands are only a few tens of miles wide but may be hundreds of miles long. *See* CONVECTION (HEAT).

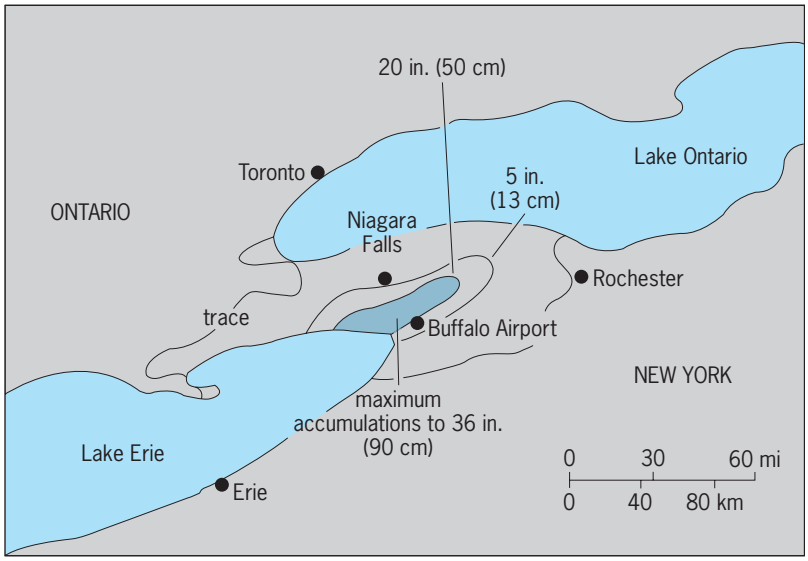

Fig. 5. Snowfall accumulations for a southwesterly flow lake-effect storm in the lee of Lake Erie on November 15, 1974. Note the extremely heavy and localized snowfall centered over the downwind shoreline.

Because cold air masses are usually associated with large regions of high pressure and stable atmospheric conditions, the layer which overturns is shallow (usually less than about 2.5 mi or 4 km deep) relative to overturning in other precipitation-producing weather systems; for example, the vertical circulations in thunderstorms and extratropical cyclones typically extend to around 7 mi (12 km). Nevertheless, lake-effect cloud bands produce intense snowfall rates, sometimes reaching 12 in. (30 cm) per hour and resulting in accumulations of several feet in a day (**Fig. 5**).

Convective complexes. A convective complex is a group of deep convective clouds (usually thunderstorms) acting in a unified manner. Its structure and circulation must be of sufficient size and duration that the Rossby number is on the order of 1.

During the late spring and summer, the large-scale (synoptic) circulations of the atmosphere create convectively unstable environments wherein deep convective clouds can develop. The primary energy source for these clouds is the latent heat released by condensation of water vapor. In most instances, the air supplying the water vapor for cloud growth comes from the Earth's boundary layer (approximately the lowest half mile of the atmosphere); and usually this air must be forcibly lifted (for example, by a warm or cold front) to start the condensation process and subsequent growth of the deep clouds. In some situations, the air feeding the clouds is unusually warm and exceptionally moist so that extremely large amounts of heat are released by condensation and the clouds develop into huge thunderstorms. When several such thunderstorms occur in a group, the tremendous amount of heat and moisture released enhances the upward circulation in the thunderstorm region. This effect forces more boundary-layer air to be lifted to its condensation level and more deep clouds to develop. Thus, an initially small group of thunderstorms can work synergistically with large-scale features to produce an amplifying system culminating in a large complex of deep convective clouds and mesoscale ascent. Moreover, the combination of strong heating and moistening is sometimes sufficient to generate a new mesoscale weather system, a warm-core mesoscale vortex that can persist for several days.

Figures 6 and **7** show the life cycles of two such convective complexes and the resulting development of mesoscale vortices. Figure 6 presents the evolution from the vantage point of a weather satellite, while Fig. 7 shows the evolution as

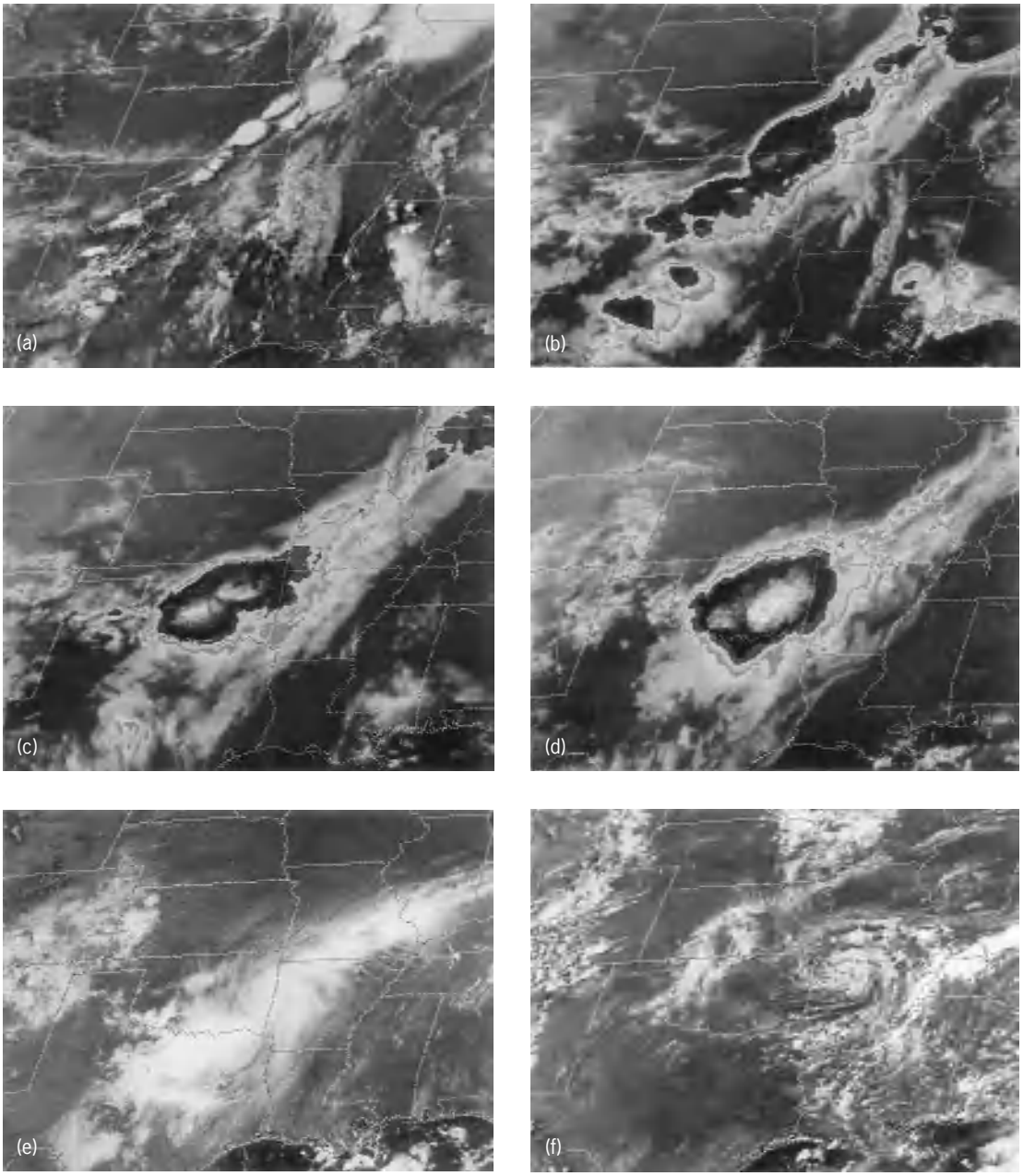

Fig. 6. Life cycle of a mesoscale convective complex. (*a*) Visible image at 2330 UTC (6:30 PM EST), July 6, 1982. (*b*) Enhanced infrared images at 0300 UTC, (*c*) 0600 UTC, and (*d*) 0900 UTC, July 7; interior gray and black contours indicate colder infrared temperatures and higher cloud tops. (*e*) Visible image at 1530 UTC, July 7. (*f*) Visible image at 1830 UTC, July 8.

observed by radar. In Fig. 6*a*, a line of thunderstorms stretches from Missouri southwestward to Texas. Following sundown, the line of thunderstorms consolidates into a concentrated area of intense convective overturning (Fig. 6*b–d*). After daybreak (Fig. 6*e*), the newly created mesoscale vortex becomes evident. The vortex is still evident a day later after it drifts out of Oklahoma and into northwestern Arkansas. In Fig. 7*a*, individual storms over northeastern Nebraska and northwestern Iowa evolve into a nearly solid line of intense convection (Fig. 7*b*,*c*). The entire system develops cyclonic rotation as evident from the comma-shaped cloud in Fig. 7*d*.

Convective complexes, as in Figs. 6 and 7, typically begin in midafternoon and reach their peak development in the middle of the night. By morning, most complexes are weakening and usually dissipate a few hours after sunrise. Although the convective complexes are only mesoscale in size, they occur frequently (on the average, about 40-50 in the United States each year) and produce copious rainfall during the growing season. Convective complexes are observed over all major midlatitude landmasses of the world. Satellite imagery indicates they commonly occur over North and South America, Southeast Asia, Australia, Africa, and, to a lesser degree, over Europe.

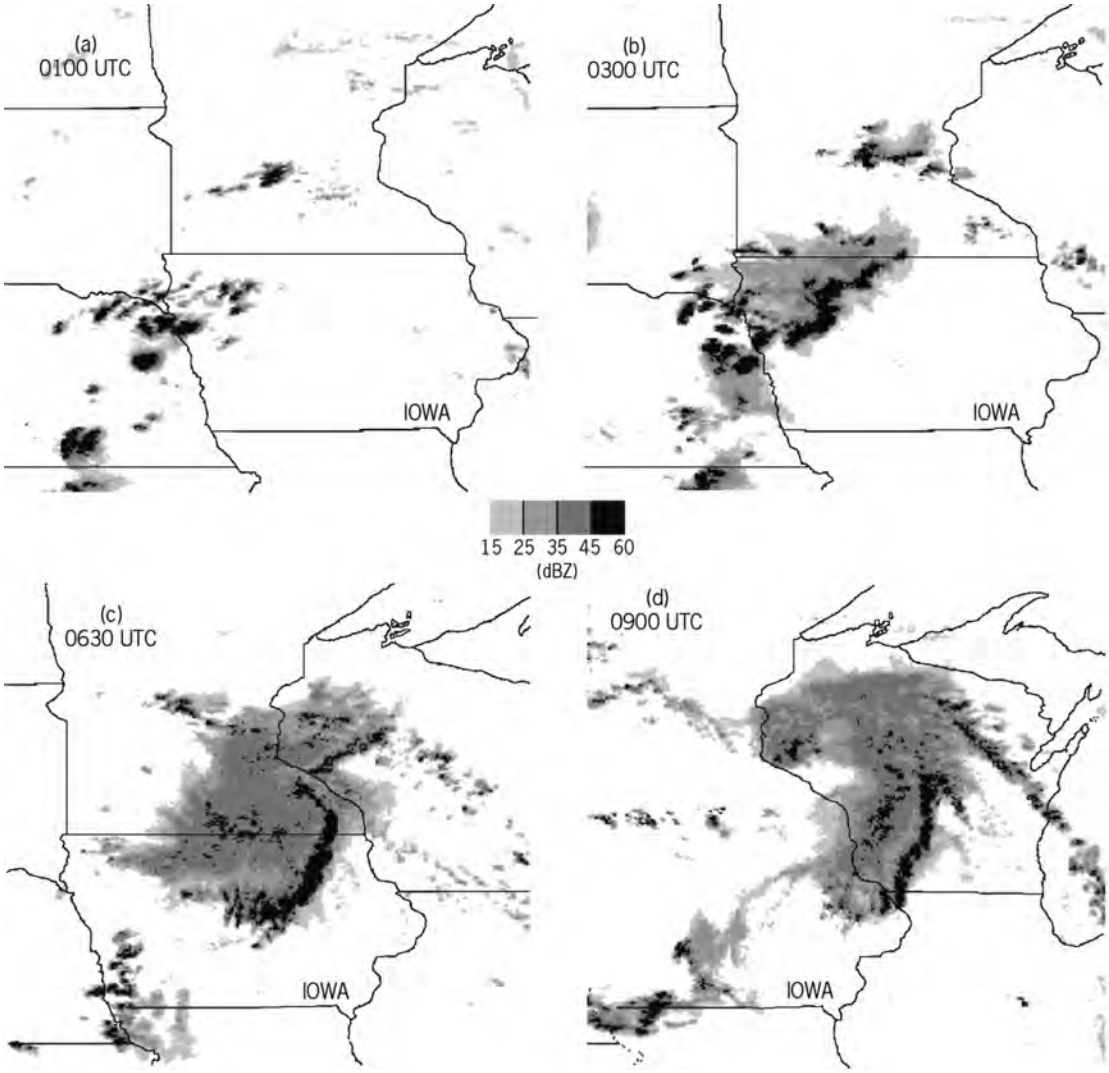

Fig. 7. Radar reflectivity (measured in dBZ) for June 24, 1998, showing the development of a mesoscale convective system and associated mesovortex over the midwestern United States. Values greater than 45 dBZ typically indicate the presence of thunderstorms.

They are also found over warm tropical water, but not as frequently as over land.

Rain/snowbands. These are elongated areas of precipitation (or more intense precipitation) that typically occur in conjunction with extratropical cyclones (**Fig. 8**). Their width usually ranges from about 3 mi (5 km) for bands coinciding with cold fronts to about 30 mi (50 km) for bands occurring in the warm air ahead of the cold front or in the cold air behind. Similar-size bands are also observed to occur ahead of warm fronts in regions of deep warm advection. The lengths of the bands vary considerably, but most are on the order of several hundred miles. Because of the large changes in stability and vertical wind shear from one part of an extratropical cyclone to another, very different mechanisms can produce the various bands. For example, in **Fig. 9**, bands occurred in both the warm and cold sectors and along the cold front. The cold-frontal band most likely can be explained in a conventional manner; that is, it was forced by convergence produced by the cold front advancing into the warm air mass. However, the prefrontal and postfrontal bands result from other processes. The most likely cause of the wide cold-sector bands is moist symmetric instability. This mesoscale instability is present when a parcel of air can be displaced slantwise (between horizontal and vertical) in a region of synoptic-scale vertical wind shear (that is, in a baroclinic atmosphere) and freely continue its slantwise ascent. Since surfaces of constant angular momentum M are also slantwise in a baroclinic atmosphere, a purely horizontal displacement subjects the parcel to different Coriolis and horizontal pressure-gradient forces that tend to restore the parcel to its original position. Similarly, unless the atmospheric lapse rate is dry adiabatic, purely vertical displacements without condensation subject the parcel to vertical pressure-gradient (buoyancy) forces that also tend to restore it to its original position. However, if a parcel with angular momentum M_p is displaced slantwise along the surface $M = M_p$, then

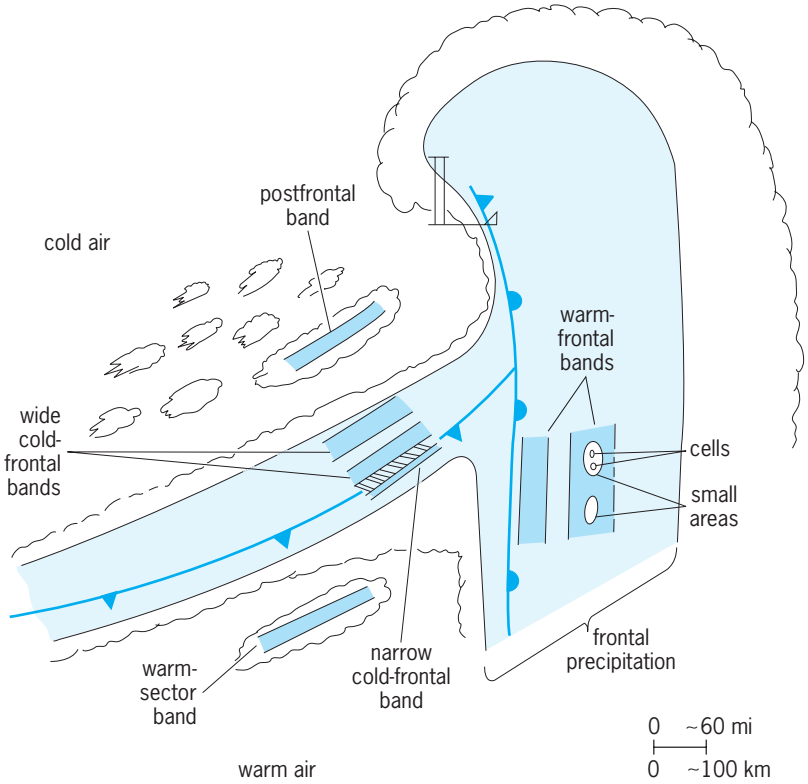

Fig. 8. Types of rainbands observed in extratropical cyclones. (*After R. A. Houze, Structures of atmospheric precipitation systems, Radio Sci., 16:671–689, 1981*)

the horizontal inertial restoring forces are minimal. Furthermore, if the slantwise displacement causes the parcel to become saturated, then the parcel may experience upward-directed buoyancy forces that accelerate it along the ascending slantwise path. In this manner, entire layers of air can ascend slantwise trajectories and produce long bands of precipitation. **Figure 10** shows a radar-observed heavy precipitation band over southeastern New York. The band produced snowfall rates as great as 4 in. (10 cm) per hour at Albany. *See* BAROCLINIC FIELD; CORIOLIS ACCELERATION; GEOSTROPHIC WIND.

Warm-sector rainbands are probably produced by either moist symmetric instability or internal gravity waves enhanced by latent heat release. Internal gravity waves may be generated by momentum imbalances or intense convection along the cold front. As the waves propagate eastward into the warm air mass, the ascending branch of the wave produces condensation and latent heat release. The heating strengthens the ascending circulation and thereby amplifies or helps maintain the propagating gravity wave. Symmetric instability in the warm sector develops in the same manner as described for the wide cold-frontal bands. In the warm sector, however, symmetric instability may originate at lower levels in the atmosphere than in the wide cold-frontal case and therefore has the potential to produce deeper clouds and more intense precipitation rates. *See* ATMOSPHERIC GENERAL CIRCULATION; ATMOSPHERIC WAVES, UPPER SYNOPTIC;

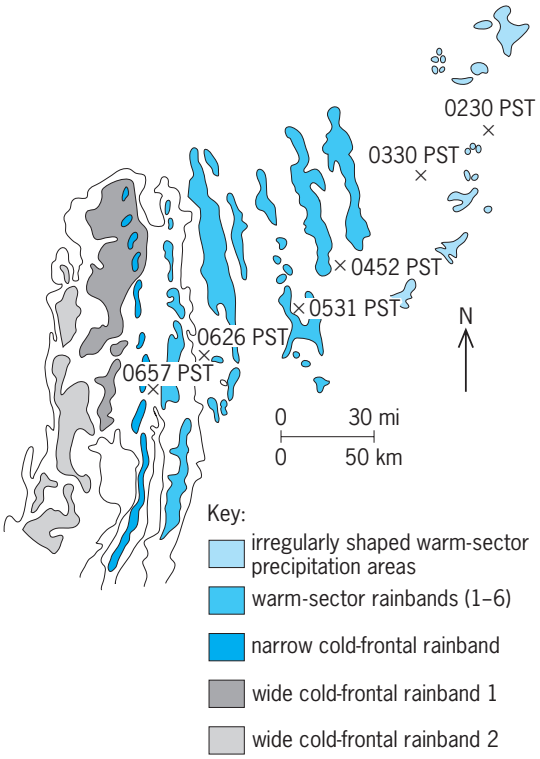

Fig. 9. Time series composite of radar reflectivity PPI displays showing rainbands as they approached the Washington coast. The crosses indicate the positions of the CP-3 radar relative to the rainbands at various times on December 8, 1976. (*After D. B. Parsons and P. V. Hobbs, The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones, XI: Comparisons between observational and theoretical aspects of rainbands, J. Atm. Sci., 40:2377–2397, 1983*)

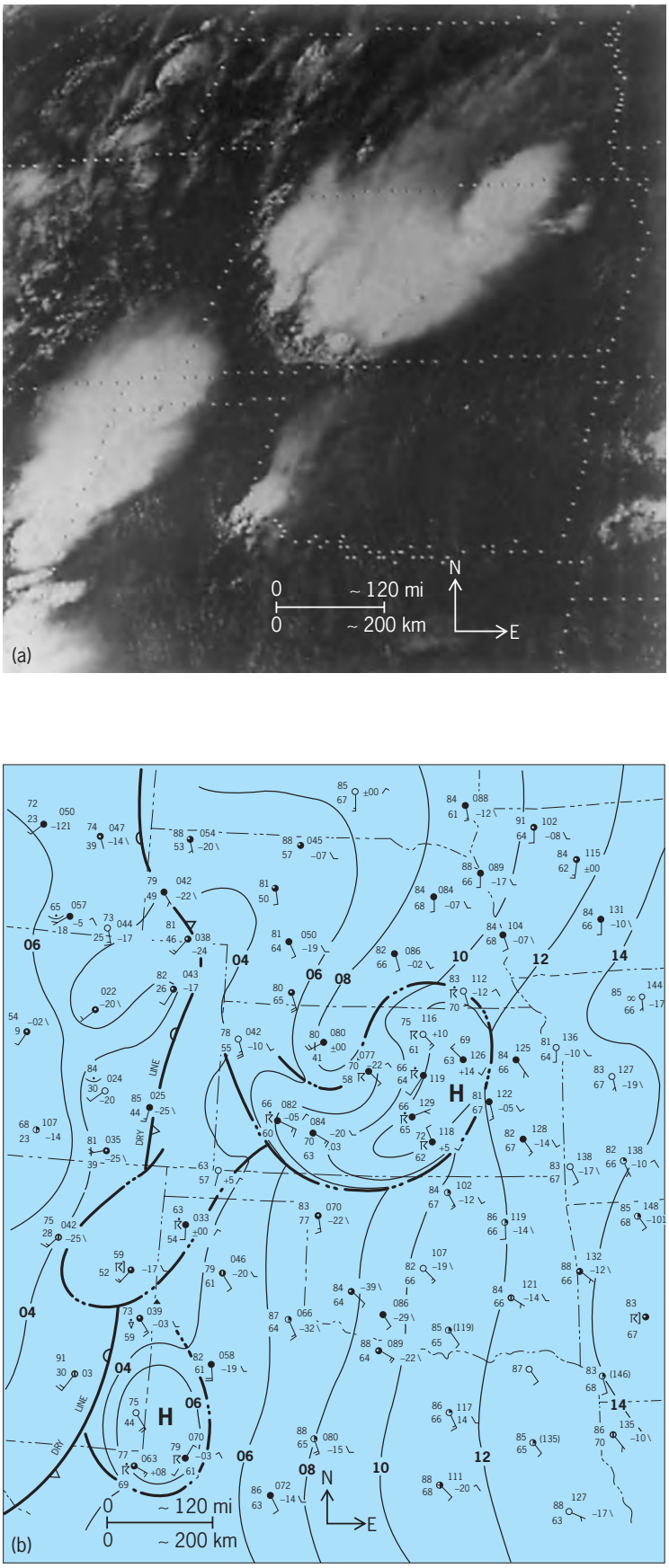

Fig. 10. Mesoanalysis. (*a*) Visible *GOES* satellite imagery, 2132 UTC (4:32 PM EST), May 25, 1978. (*b*) Corresponding surface mesoanalysis, 0000 UTC (7:00 PM EST), May 26, 1978. (*After J. M. Fritsch and R. A. Maddox, Cyclonic thunderstorm outflow, Mon. Weath. Rev., 107:87–89, 1979*)

CONVECTIVE INSTABILITY; DYNAMIC INSTABILITY.

Mesoanalysis. This is the analysis of meteorological data in a manner that reveals the presence and characteristics of mesoscale phenomena. Because the balloon sounding stations are so widely spaced, only the largest of mesoscale systems can be resolved by the balloon sounding network. However, the density of surface observing stations is often satisfactory for identifying mesoscale features or circulations.

Although more detailed mesoanalyses can be performed when surface data are combined with radar and satellite data, it still is not possible to quantify and understand the structure and circulations of mesoscale phenomena without high-density sounding or aircraft data. Consequently, in some research experiments, special high-density sounding networks have been established. These special networks permit higher-resolution three-dimensional analysis wherein the mesoscale structure and circulations of the free troposphere can be determined. Unfortunately, high-density observing systems are much too expensive to maintain continuously over large regions for long time periods. Therefore, mesoanalysis is done only on a sporadic and limited basis. *See* AIR MASS; FRONT; JET STREAM.

Forecasting mesoscale phenomena. Modern weather-forecasting techniques rely heavily upon predictions made by computers. These predictions, commonly called numerical model forecasts, require as input the three-dimensional initial state of the atmosphere. Normally, this initial condition is produced from a combination of the previous forecast (valid at the model initial time) and the most recent observations from radars, satellites, aircraft, and the network of balloon-borne instrument packages. Since the initial conditions for the computer predictions reflect the resolution of the observing network, typically only the largest of mesoscale phenomena are directly included. However, due to the recent advances in computer power, most computer models now have spatial resolution that exceeds the resolution of the observing network and, therefore, some mesoscale phenomena can be generated within the numerical models.

Nevertheless, many of the significant mesoscale phenomena listed in the table are not routinely predicted by the computer models. For these phenomena, meteorologists take an alternate approach to forecasting. It has long been recognized that while mesoscale phenomena have distinct physical characteristics, they are usually strongly modulated or controlled by the synoptic-scale environment. Thus, for many mesoscale phenomena, it has been possible to develop relationships between the large-scale (synoptic) environment and the occurrence of particular types of mesoscale events. By using these relationships, the prediction of the synoptic-scale environment by the numerical models is then used to infer the likelihood of specific mesoscale events, that is, implicit instead of explicit prediction.

This procedure is not as precise as synoptic-scale prediction, and often only a general region and

approximate time period (on the order of 6–12 h) in which the mesoscale event will occur can be determined. Then, by using satellite, radar, and conventional surface observations, the onset of the event is readily detected, and appropriate adjustments to local forecasts are implemented. These adjustments usually come in the form of very short term forecasts, commonly called nowcasts, and typically are valid for only about 3 h. However, depending upon the particular mesoscale phenomena, longer-term (3–12-h) forecasts sometimes are possible. This is particularly true for forced-mode phenomena (for example, lake-effect storms) where the synoptic-scale conditions forcing the mesoscale event are changing very slowly. *See* METEOROLOGY; SATELLITE METEOROLOGY; STORM DETECTION. J. Michael Fritsch

Bibliography. B. Atkinson, *Mesoscale Atmospheric Circulations*, 1981; K. A. Emanuel, On assessing local conditional symmetric instability from atmospheric soundings, *Mon. Weath. Rev.*, 111:2016–2033, 1983; R. L. Lavoie, A mesoscale numerical model of lake-effect storms, *J. Atm. Sci.*, 19:1025–1040, 1972; D. K. Lilly and T. Gal-Chen, *Mesoscale Meteorology: Theories, Observations and Models*, NATO Advanced Science Institute Series, 1983; D. B. Parsons and P. V. Hobbs, The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones, XI: Comparisons between observational and theoretical aspects of rainbands, *J. Atm. Sci.*, 40:2377–2397, 1983; P. S. Ray, *Mesoscale Meteorology and Forecasting*, 1986.

Meson

Strongly interacting particles are either baryons or mesons. Baryons, such as the proton and neutron, have half-integer spins and are fermions, whereas mesons are bosons with integer spin. The lightest meson is the pion, denoted π, which was originally predicted to exist as the carrier of the powerful force that grips the protons and neutrons within atomic nuclei. *See* BARYON; QUANTUM STATISTICS; SPIN (QUANTUM MECHANICS); STRONG NUCLEAR INTERACTIONS.

Like baryons, mesons are very small (roughly 10^{-15} m in diameter) and consist of even smaller elementary particles: quarks, antiquarks, and gluons. Baryons are fundamental particles of matter, made of quarks; antibaryons are their antimatter analogs and are made of antiquarks. Mesons are neither matter nor antimatter: they are made of quarks and a balancing number of antiquarks (in the simplest case a single quark and a single antiquark). Over 200 examples are known of such q–q-bar combinations (q being the generic symbol for a quark; and q-bar, denoted here by $\bar{q}$, for an antiquark). As is the case for baryons, the quarks and antiquarks are bound by the exchange particles of the strong interaction, the gluons. Examples have begun to emerge of mesons that appear to require two quarks and two antiquarks, or even no quarks and antiquarks at all. The former are known as tetraquarks or molecules; and the latter as glueballs, consisting, at least in theory, of just gluons. *See* ANTIMATTER; GLUONS; QUARKS.

Mesons are identified by properties referred to as flavor, spin, and parity, all of which have discrete values. They also have measurable properties whose values form a continuous range of numbers, such as their mass. For more details about these and other properties *see* BARYON.

Instability. Whereas the lightest baryon, the proton, is stable against decay into other particles, all mesons are unstable. Heavy mesons can decay rapidly (in about 10^{-23} s) into lighter mesons so long as the flavor and other properties, such as parity and angular momentum or spin, are preserved. When the lightest meson of a particular flavor is formed, it is nonetheless unstable since its quark and antiquark can mutually annihilate, for example into two photons, an electron, and a positron or neutrino, or even gluons. For example, the lightest meson, the pion, occurs either with no electric charge (π^0) or with negative or positive electric charge (π^- or π^+). When the quark and antiquark within the π^0 annihilate, the products are probably two photons; when the π^- or π^+ decay, they leave a neutrino and an electron or positron, or more often a neutrino and a muon (a heavier analog of the electron and positron). The π^0 survives for only 10^{-16} s—small indeed but over a million times longer than the 10^{-23} s of its heavier cousins. The π^- or π^+ survive for 10^{-8} s. These differences in lifetimes are the result of the forces at work. For heavy mesons, the strong force destroys them; for the π^0, it is the electromagnetic force; and for the electrically charged pions, it is the weak force that is responsible. The large differences in lifetimes reflect the differing strengths of the forces. *See* FUNDAMENTAL INTERACTIONS; LEPTON; NEUTRINO.

Meson flavors. There are six different kinds, or flavors, of quarks, in three families. The lightest family contains the up (u) and down (d) quarks, and charges of $+{}^2/_3$ and $-{}^1/_3$ in units of the electron charge (e), respectively. The corresponding antiquarks, labeled $\bar{u}$ or $\bar{d}$, have the opposite sign of charge to their quark counterparts. Thus, $\bar{u}$ has charge $-{}^2/_3$ and $\bar{d}$ has $+{}^1/_3$. *See* FLAVOR.

A $u\bar{d}$ has charge $+1$, and the lightest example is the π^+. The $d\bar{u}$ has charge -1: the π^-. Both the $u\bar{u}$ and $d\bar{d}$ have charge zero. Quantum theory implies that the particular quantum combination $u\bar{u} - d\bar{d}$ corresponds to the π^0; the alternative combination $u\bar{u} + d\bar{d}$ appears in another meson, the electrically neutral η (eta).

Because the difference between the masses of the up and down quarks and the electromagnetic self energy of quarks or antiquarks in a meson are very small compared to meson masses, properties of mesons are to a very good approximation unchanged by replacing an up quark or antiquark with a down quark or antiquark. This equivalence is referred to as isospin symmetry. In this language, the π^+, π^-, or π^0 are referred to collectively as different charge states of the pion. *See* I-SPIN.

The three varieties of pion correspondingly have masses that differ by only some 4%, the charged pions having a mass of 140 MeV/c^2 (where c is the speed of light) and the π^0 mass being 135 MeV/c^2. The η, however, is some four times more massive at 550 MeV/c^2. In part, this is because the η contains (along with $u\bar{u} + d\bar{d}$) a third flavor of quark, the strange quark, accompanied by its antiquark: thus $s\bar{s}$. The strange quark is some 120 MeV/c^2 heavier than an up or down quark, and so an $s\bar{s}$ has added mass relative to pions, which contain only u or d flavors. Indeed, the additional quark flavor leads to three electrically neutral mesons, π^0, η, and η', which testify to the three possibilities, $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$; the η' mass is 960 MeV/c^2.

It is possible for any of the u, d, or s quarks to be attracted to any of the $\bar{u}$, $\bar{d}$, or $\bar{s}$ antiquarks. Mesons that have an s or $\bar{s}$ partnered by a ($\bar{u}$, $\bar{d}$) or (u, d) will have strangeness. Such "strange" mesons are known as kaons and denoted generically by the symbol K. The kaon, containing a light flavor and a strange, has a mass of around 495 MeV/c^2, which is heavier than a pion but lighter than η or η'. *See* STRANGE PARTICLES.

The remaining three flavors of quarks—the charm c, bottom b, and top t quarks—are heavy on the scale of the mass of the pion, K, or η, and η' mesons. These give rise to further families of mesons known as charmonium and bottomonium (a $c\bar{c}$ or $b\bar{b}$ combination) and also mesons that contain unpaired charm or bottom flavors. The top quark is very massive, causing it to decay so rapidly that it does not have time to form a meson by binding strongly to other quarks or to an antiquark. Thus toponium is unlikely to occur unless there are other forces at work which are yet unknown. *See* CHARM.

Spin. Quarks carry an intrinsic degree of freedom called spin, which is related to the angular momentum carried by a macroscopic object when it rotates. In the very short-distance quantum world, angular momentum is quantized in units of $\hbar = h/(2\pi)$, where h is Planck's constant. The smallest possible angular momentum is $\hbar/2$, and this is the spin of each quark or antiquark making up a meson. Thus quarks and antiquarks are referred to as spin-$^1/_2$ particles. *See* ANGULAR MOMENTUM; SPIN (QUANTUM MECHANICS).

Two spin-$^1/_2$ particles, such as a quark and an antiquark, can have their spins combined to 0 (anti-aligned along an axis called the quantization axis) or to 1 (aligned along the quantization axis), in units of $\hbar$.

The pi mesons, η, η', and K mesons all have a total spin of zero, arising when their quark and antiquark have their spins anti-aligned. If their spins were instead aligned, the total spin would be 1. A family of mesons with spin 1, called the vector mesons, is known. The ρ^+, ρ^0, and ρ^- (rho) are the $u\bar{d}$, $u\bar{u} - d\bar{d}$, $d\bar{u}$ analogs of the spin-0 pi mesons. Whereas the η and η' are complicated mixes of $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$, their vector meson analogs are simpler. The ω (omega) consists of $u\bar{u} + d\bar{d}$, while the φ (phi) is made simply of $s\bar{s}$. The vector strange mesons, denoted K^*, exist, made from $u\bar{s}$, $d\bar{s}$, $s\bar{d}$, and $s\bar{u}$, as for their spin-0 counterparts.

The masses of the spin-1 mesons are greater than their spin-0 counterparts. The ω, made of $u\bar{u} + d\bar{d}$, has mass of 780 MeV/c^2, almost the same as that of the ρ made of $u\bar{u}-d\bar{d}$; the K^* mesons with masses of 890 MeV/c^2 are then midway between the ρ or ω and the φ ($s\bar{s}$), whose mass is 1020 MeV/c^2. This linear pattern is naturally explained by their having respectively only light-flavored u or d quarks (the ρ or ω); one heavier s or $\bar{s}$ quark (the K^*), or both s and $\bar{s}$ quarks (the φ). The greater mass of any individual flavor set with spin 1, such as $u\bar{s}$ for the K^* relative to the spin-0 K, can be thought of as due to a kind of magnetic interaction between the colored, spinning quarks, analogous to the magnetic interaction between charged, spinning electrons. These interactions are repulsive when the quark spins are aligned; this raises the mass of the spin-1 K^*, with both quark spins aligned, relative to that of the spin-0 K, with the quark spins anti-aligned.

In addition to their spin angular momentum, the constituent particles of a meson can have orbital angular momentum as they move around their common center of mass. In the simplest possible (naive quark model) picture, the spin angular momentum of a meson is made up of the orbital angular momentum and intrinsic spin of its constituent quark and antiquark.

Just as the lightest states of the hydrogen atom have the electron and proton in an orbital-angular-momentum-zero, spherically symmetric state, so in this simple picture the lightest states of a meson have zero orbital angular momentum between the quarks, and so their total spin is just that of the quark and antiquark: hence the spin-0 and spin-1 combinations described above.

Mesons with the same quark and antiquark flavor contents as the above exist with spin angular momentum higher than 0 or 1. In most cases this is because the quarks have orbital angular momentum around their common center of mass, but there is also much excitement at the possibility that states known as hybrid mesons exist, where the gluons play a direct role in building up the total spin. Since orbital angular momentum comes in integral multiples of $\hbar$, which is also true of the intrinsic spin of a gluon or quark-antiquark pair, all mesons must have integral spin $J = 0, 1, 2, 3$, and so forth. Mesons with spins up to 6 have been observed.

Parity and charge conjugation. A discrete property of a meson is its parity. The strong interactions between quarks and antiquarks have the property that they do not change when viewed in a mirror. This is essentially the same as spatial inversion symmetry, invariance under the operation of reflecting all of the spatial coordinates through the origin. The quantum wave functions of such states under the operation of inversion of the coordinates either remain the same, called positive parity, or change sign, called negative parity, denoted P. *See* PARITY (QUANTUM MECHANICS).

A similar effect arises when the wave function is matter and antimatter inverted: when quark and antiquark are interchanged in effect. This operation is known as charge conjugation. As was the case for parity, so for charge conjugation: the state can have C be positive or negative. *See* SYMMETRY LAWS (PHYSICS).

For a meson made from a single quark and antiquark, all possible integer spins can occur. However, not all combinations of positive and negative values of C and P can arise. If C and P are both positive, any value of total spin can occur, and if the spin is 1 or more, both C and P can be negative; however, if the total spin is 0, the latter cannot occur. It is traditional to denote the total spin (J), parity P, and charge conjugation C for a meson in the form J^{PC}; so in this language it is said that 0^{--} combinations are forbidden. When P is $-$ and C is $+$, the total spin must be even (giving the values 0^{-+}, 2^{-+}, 4^{-+}, and so forth); when P is $+$ and C is $-$, the total spin must be odd (giving 1^{+-}, 3^{+-}, 5^{+-}, and so forth). Any combinations that are not allowed are known as exotic. For many years, all mesons discovered satisfied these rules. However, evidence has emerged for a meson with spin 1 (odd), negative P, and positive C, namely 1^{-+}, and hence exotic.

Such a meson had been predicted. It turns out that if the quark and antiquark are accompanied by a gluon, it is possible for such otherwise exotic correlations to occur. As mentioned above, mesons where both q and $\bar{q}$, and glue are active are known as hybrids.

Decays. In strong decays, a quark-antiquark pair is created via the strong interaction, with the created quark ending up in one meson and the created antiquark in another meson. This process can happen successively such that an initial meson decays into two, three, or more mesons.

Decays may also proceed via the emission of an energetic particle of light, a gamma-ray photon, since the electromagnetic interaction between quarks, made possible by their electric charges, is the next strongest of the fundamental interactions between them. *See* GAMMA RAYS.

Meson decays may also proceed via the weak interactions mediated by the massive vector-boson particles, W^+, W^-, or Z^0. Such decays can result in the creation from a W boson of a charged lepton and a neutrino of the same family, such as an electron and an electron antineutrino, or their antiparticles, the positron and an electron neutrino. This process is what occurs when a π^- or π^+ decays. *See* INTERMEDIATE VECTOR BOSON; WEAK NUCLEAR INTERACTIONS.

Charmonium and bottomonium. As mentioned above, a charmed quark attracted to a charmed antiquark gives a family of mesons known as charmonium, $c\bar{c}$; and a similar phenomenon occurs for the yet more massive bottom flavor, giving bottomonium, $b\bar{b}$. There are also mesons with manifest charm or bottom flavors, such as the $D(c\bar{u})$ with mass 1865 MeV/c^2 and $B(b\bar{u})$ at 5279 MeV/c^2. Charmonium mesons with masses greater than 3730 MeV/c^2 decay by the strong interaction into pairs of charmed mesons, $D(c\bar{u})$ and $\bar{D}(\bar{c}u)$. However, there are several examples of charmonium mesons that have masses below this limit and so are prevented by energy conservation from decaying this way. Their dominant decays are by shedding energy in the form of photons or pions, eventually ending up as the etacharm (η_c) or J-psi (J/Ψ), the lightest $c\bar{c}$ mesons (the η_c having $c\bar{c}$ spin totaling 0 and the J/Ψ having spin 1, analogous to the pion and rho mesons in the light-flavor examples described earlier).

The J/Ψ is metastable on the time scale of strong interactions. It decays (after about 10^{-20} s) by its c and $\bar{c}$ mutually annihilating, producing photons or gluons; the gluons do not escape free but rapidly create showers of pions or other particles. *See* J/PSI PARTICLE.

There is an interesting possibility that among the decay products of a J/Ψ lurk glueballs. According to theory, the J/Ψ can decay into a photon and two gluons. The energy taken by the photon can vary, and the energy of the two gluons varies in consequence. If this energy equals the mass of a glueball, the two gluons will form a glueball. There are some enigmatic mesons seen in the decay products of the J/Ψ, and it is suspected that these include evidence for glueballs, but definitive proof for their existence is still awaited.

Glueballs, hybrids, and tetraquarks. While over a hundred mesons made of $q\bar{q}$ are known, the attractive forces can also make combinations of two quarks and two antiquarks, $qq\bar{q}\bar{q}$, known as tetraquarks (as noted above). The gluons, which mediate the color force between the quarks, like the photons, which mediate the electromagnetic force between particles, carry a spin angular momentum of $\hbar$ and are referred to as spin-1 or vector particles. As noted above, they can attract one another to form mesons known as glueballs or bind with quarks and antiquarks to make hybrid mesons. While the existence of tetraquarks, glueballs, and hybrids is not yet totally established, several possible examples have been found.

A family of mesons with spin 0 and positive parity (scalar mesons) with masses below 1000 MeV/c^2 have properties that fit well with being examples of tetraquarks. These are made from u, d, and s quark flavors and their antiquarks. Since 2004, examples of further possible tetraquarks have been found that include one or more charm (c) or $\bar{c}$ flavors. Two mesons, with masses 2317 and 2460 MeV/c^2, appear to be the spin-0 and -1 examples of $c\bar{s}u\bar{d}$. A meson with spin 1, positive parity, and mass 3872 MeV/c^2 is probably $c\bar{c}u\bar{u}$; a spin-1 negative parity vector meson at 4260 MeV/c^2 may be a $c\bar{c}s\bar{s}$.

The latter meson may also be a hybrid meson, made from $c\bar{c}$ and excited gluons. Hybrid mesons are predicted to have masses about 1 GeV/c^2 more than their lightest $q\bar{q}$ analogs. The lightest spin-1 $c\bar{c}$ is the J/Ψ at 3095 MeV/c^2, and the

The six flavors of quarks, their charges, and their masses				
Charge = +2e/3	Flavor	up (*u*)	charm (*c*)	top (*t*)
	Mass, MeV/c^2	3	1250	174,000
Charge = −e/3	Flavor	down (*d*)	strange (*s*)	bottom (*b*)
	Mass, MeV/c^2	6	120	4250

4260-MeV/c^2 meson is thus consistent with this rule. Few properties of this newly discovered meson are yet known. To determine which interpretation is correct will require careful study of its decay products. If it is a *ccss* tetraquark, then pairs of D_s "charm strange" mesons, made of $c\bar{s}$ or $s\bar{c}$ with masses around 2000 MeV/c^2, should be dominant. However, for a hybrid $c\bar{c}$-gluon meson, such D_s mesons should not be favored decay products.

There are also examples of possible hybrid mesons containing *u* or *d* flavors, but there is much debate about whether these might instead be tetraquarks.

A detailed spectrum of glueballs is predicted by theory. The lightest of these is expected to be a scalar meson (spin 0, positive parity) with mass around 1500–1700 MeV/c^2. There are scalar mesons around this mass, at 1710, 1500, and 1370 MeV/c^2, which are probably the result of $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$ scalar mesons undergoing quantum-mechanical mixing with the scalar glueball. Extensive studies of their production and decay modes are required to measure the details of this mixing pattern and establish the presence of a glueball.

Meson production. The electrically charged pi mesons and kaons were discovered in cosmic rays; knowledge about other mesons comes from experiments involving collisions of particles. Historically, collisions between protons and nuclei were observed, or dedicated beams of pions or kaons hitting protons and nuclei. Since the 1980s, a large amount of knowledge about mesons has resulted from experiments where electrons collide with and are annihilated by positrons. If the energy of the collision is the same as that of a vector meson, the vector can be produced directly. This is how the J/Ψ—the vector meson of the charmonium family, *cc*, with a mass of 3095 MeV/c^2—was discovered in 1974. The upsilon (Υ), the analogous vector meson of bottomonium, *bb*, can also be made this way. The J/Ψ and Υ decay in about 10^{-20} s and leave a rich pattern of mesons among their decay products. *See* UPSILON PARTICLES.

If the total energy of the electron-positron collisions is around 10 GeV, the result is a pair of *B* mesons, where a *b* flavor is accompanied by up, down, or strange such that the meson has an overall bottom flavor. The *B* mesons decay by the weak interaction. There are several possible combinations of mesons that can result from these decays, and the relative amounts of each are being measured to the highest precision possible. The reason for this interest is that the decays of bottom mesons and their antimatter counterparts, containing a *b* antiquark, are expected to give important clues about the origin of the asymmetry between matter and antimatter.

Experiments with beams of electrons and intense beams of photons in collisions with protons and nuclei are expected to produce vast amounts of data on mesons. These experiments are also expected to give crucial information about the enigmatic hybrid mesons, whose existence is predicted by theory but not yet unambiguously established in experiment. *See* ELEMENTARY PARTICLE.

Frank Close

Bibliography. F. Close, *Particle Physics: A Very Short Introduction*, Oxford, 2004; F. Close, M. Marten, and C. Sutton, *The Particle Odyssey*, Oxford, 2003; S. Eidelman et al. (Particle Data Group), Review of particle physics, *Phys. Lett.*, B592:1–1109, 2004; N. Isgur and G. Karl, Hadron spectroscopy and quarks, *Phys. Today*, 36(11):36–42, November 1983.

Mesosauria

An order of extinct aquatic reptiles, also known as Proganosauria, of latest Carboniferous or earliest Permian time, about 2.8×10^8 years ago. The best-known genus, *Mesosaurus* (see **illus.**), occurs in the Iratí bituminous shales of southern Africa. Geological evidence demonstrates that South America and Africa were part of a single land mass in the late Paleozoic, prior to the formation of the Atlantic Ocean. No other reptiles are known from contemporary beds on either continent. Like many aquatic reptiles, mesosaurs have a very long snout. The teeth are numerous and long, and appear very delicate. They may have served for filter feeding on soft-bodied invertebrates. The girdles and limbs are little modified for aquatic locomotion. The hands and feet are paddle-shaped. The tail is long and laterally compressed. The trunk ribs have the swollen form (pachystosis) characteristic of nothosaurs (aquatic reptiles of the Triassic) and sirenians, both of which show early stages in the adaptation of a formerly terrestrial group toward life in the water. Mesosaurs, an early offshoot of the Carboniferous "stem reptiles" or captorhinomorphs, were the first reptiles to invade marine waters, but apparently soon became extinct, leaving no descendants. *See* REPTILIA. R. L. Carroll

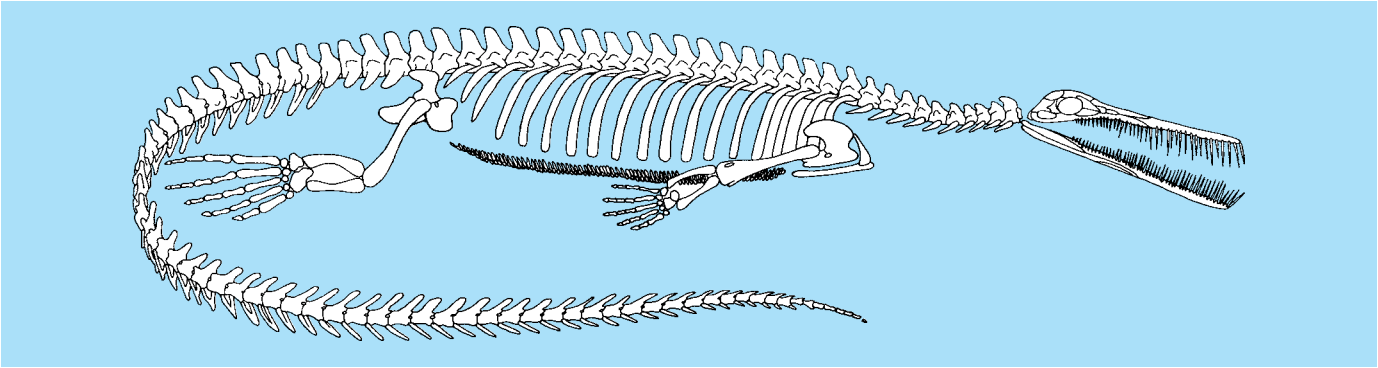

Restoration of the aquatic reptile *Mesosaurus*. Original length was 16 in. (40 cm).

Bibliography. A. S. Romer, *Osteology of the Reptiles,* 1955, reprint 1997.

Mesoscopic physics

A subdiscipline of condensed-matter physics that focuses on the properties of solids in a size range intermediate between bulk matter and individual atoms or molecules. The size scale of interest is determined by the appearance of novel physical phenomena absent in bulk solids and has no rigid definition; however, the systems studied are normally in the range of 100 nanometers (10^{-7} meter, the size of a typical virus) to 1000 nm (the size of a typical bacterium). Other branches of science, such as chemistry and molecular biology, also deal with objects in this size range, but mesoscopic physics has dealt primarily with artificial structures of metal or semiconducting material which have been fabricated by the techniques employed for producing microelectronic circuits. Thus it has a close connection to the fields of nanofabrication and nanotechnology. The boundaries of this field are not sharp; nonetheless, its emergence as a distinct area of investigation was stimulated by the discovery of three categories of new phenomena in such systems: interference effects, quantum size effects, and charging effects. *See* ARTIFICIALLY LAYERED STRUCTURES; NANOSTRUCTURE; QUANTIZED ELECTRONIC STRUCTURE (QUEST); SEMICONDUCTOR HETEROSTRUCTURES.

Interference effects. Although quantum mechanics predicts that any particle should exhibit properties characteristic of waves, electrons in bulk solids may be treated for most purposes as particles. The reason is that the wavelike properties are suppressed by interactions between the electrons and the vibrations of the lattice of ions forming the solid. These interactions become weaker as the temperature of the solid is lowered and thermal vibrations are reduced. At temperatures less than 4 K (−452°F), the vibrations are so rare that electrons can traverse several thousand nanometers (the size of a typical mesoscopic device) without any significant interaction, and the wave properties of the electron affect measurable physical quantities such as the electric conductivity. *See* QUANTUM MECHANICS.

The most important wave property of electrons is that of interference. Unlike a classical particle, an electron wave can split at one point in space into many wavefronts which follow different paths to another point in space. These different wavefronts may add or subtract from each other at that point, leading to constructive or destructive interference. In the former case there will be an increased probability of finding an electron at that point; in the latter case there will be zero probability. The conduction electrons in solids are split into many wavefronts by defects or impurities in the crystal lattice. The density of such defects determines the average electrical conductivity of a bulk solid (at low temperatures) since they impede the flow of the electrons. However, in the mesoscopic regime scattering from defects also induces interference effects which modulate the flow of electrons. This modulation depends not only on the density of impurities but on their precise location relative to one another; hence it differs for each specimen, even if the specimens are made of the same material and fabricated under identical conditions. Thus the concept of conductivity as a property of a given material becomes useless for mesoscopic solids. Instead, the conductance of a given specimen is studied: the ratio of the measured current to the driving voltage. *See* CONDUCTANCE; CONDUCTION (ELECTRICITY); ELECTRICAL RESISTIVITY; INTERFERENCE OF WAVES.

The experimental signature of mesoscopic interference effects is the appearance of reproducible fluctuations in physical quantities. Such time-independent fluctuations have been measured in the magnetization, thermopower coefficient, alternating current, direct current, and nonlinear conductance. For example, the conductance of a given specimen oscillates in an apparently random manner as a function of experimental parameters such as a magnetic field or the electron density (**Fig. 1**). However, the same pattern may be retraced if the experimental parameters are cycled back to their original values; in fact, the patterns observed are reproducible over a period of days. In experiments a magnetic field is most commonly used to demonstrate this effect. A magnetic field induces these fluctuations because

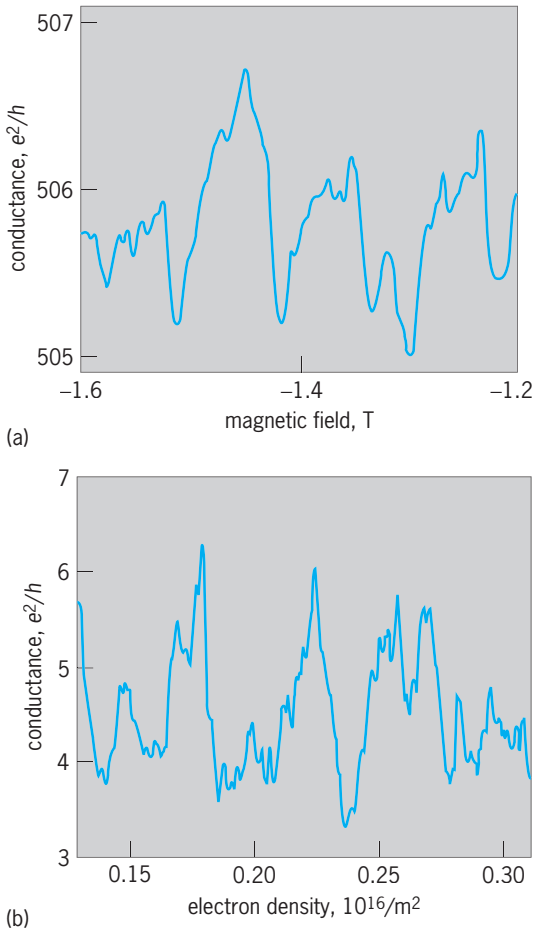

Fig. 1. Conductance fluctuations in mesoscopic conductors. (*a*) Conductance of 2000-nm gold wire as a function of magnetic field measured at temperature 0.04 K (*after R. A. Webb and S. Washburn, Quantum interference fluctuations in disordered materials, Phys. Today, 41(12): 46–55, December 1988*). (*b*) Conductance of 600-nm GaAs/AlGaAs semiconducting heterostructure as a function of electron density (controlled by an electrostatic gate) at temperature 0.1 K (*after M. W. Keller et al. Magnetotransport in a chaotic scattering cavity with a turnable electron density, Surf. Sci., 305:501–506, 1994*). The patterns observed are reproducible over a period of days, and the amplitudes of both fluctuations are of order e^2/h.

according to quantum theory a field alters the phase of the electron waves. The general principle that a magnetic field modulates quantum-mechanical interference is known as the Aharonov-Bohm effect. *See* AHARONOV-BOHM EFFECT.

Quantum size effects. Another prediction of quantum mechanics is that electrons confined to a particular region of space may exist only in a certain set of allowed energy levels. The spacing $\Delta\epsilon$ between these levels increases as the confining region becomes smaller; also, in typical systems the spacing between the levels decreases with increasing overall energy. Conduction electrons in a bulk solid are confined only to the macroscopic dimensions of the solid, and the energy-level spacing becomes so small that the discreteness of the allowed energies is undetectable. In mesoscopic solids, however, the spacing is large enough that the discreteness of the levels affects measurable physical properties. Because conduction electrons in metals have much higher energies than those in semiconductors and the level spacing decreases with energy, it turns out that these quantum size effects are most important in mesoscopic semiconductors. One striking phenomenon which arises from these quantum size effects is the steplike increase of the conductance of electrons flowing through a constriction of several hundred nanometers' width. In classical physics the conductance should increase in proportion to the width of the constriction. In quantum physics the electrons become confined laterally as they enter the constriction and are allowed only certain energies. No further current can be generated as the constriction is widened until a new state appears because of the decrease in the energy-level spacing. The result is that instead of increasing linearly with width the conductance increases in a series of sharp steps. It is found that these steps are of magnitude $2e^2/h$, a result which is expected from the basic quantum theory of the process. *See* ENERGY LEVEL (QUANTUM MECHANICS).

Another mesoscopic system that shows quantum size effects consists of isolated islands of electrons that may be formed at the appropriately patterned interface between two different semiconducting materials. [Gallium arsenide (GaAs) and aluminum gallium arsenide (AlGaAs) are most common.] The electrons typically are confined to a disk-shaped region, and hence these systems have been termed quantum dots. Quantum dots may be regarded as artificial atoms with a new set of energy levels never before realized in nature. The confinement of the electrons in these systems changes their interaction with electromagnetic radiation significantly. The electronic absorption energies are shifted from their values in bulk semiconductors, and the light that is absorbed or emitted by the dot is restricted to certain frequencies because of confinement. Quantum dots can be used to make semiconductor lasers that operate at very low power, a property of great technological interest. *See* LASER.

Charging effects. Isolated mesoscopic solids such as quantum dots or metallic grains on an insulating substrate also show novel effects associated with the discreteness of the charge on the electron. Normally the free-electron charge in a conductor can be regarded as a continuous fluid which is infinitely divisible. This is so both because the amount of charge involved in macroscopic systems is usually very large compared to the charge on a single electron and because quantum mechanics allows even a single electron to be spread out over a very large area. Hence the charge on a capacitor plate is tunable to a small fraction of the electron charge (e) if it is in contact with a reservoir of electron charge such as a battery. However, if the capacitor is mesoscopic, this ceases to be true. *See* CAPACITOR.

The simplest case is that of an isolated mesoscopic grain of metal such as aluminum. The grain has a certain number of atoms and will be in its lowest energy state if it retains the correct number of electrons to neutralize the total positive charge of the atomic nuclei. The energy cost of removing an electron is

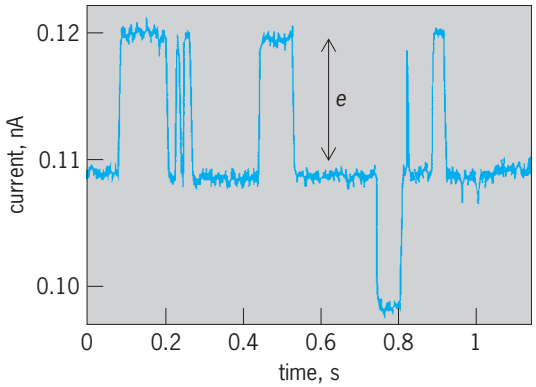

Fig. 2. Time variation of the electric current through a single-electron transistor electrometer, measuring the charge on a mesoscopic metallic grain. Each jump corresponds to the removal or addition of a single electron from the grain. (*After M. H. Devoret, D. Esteve, and C. Urbins, Single-electron transfer in metallic nanostructures, Nature, 360:547–553, 1992*)

inversely proportional to the capacitance of the grain, which decreases with its volume. For a mesoscopic grain this energy exceeds thermal energies at roughly 1 K; below this temperature the number of electrons on the grain will not vary because of thermal fluctuations.

If such a grain is fabricated as part of an electronic circuit, it will block the flow of current until the driving voltage is large enough to compensate for the energy needed to increase the electronic charge on the grain by e. However, if an additional electric field is applied to the grain through an auxiliary capacitor, the singly ionized state of the grain can be made energetically allowed and a current will flow through the circuit. With this principle it has been possible to fabricate circuits for which a current of 10^9 electrons per second (approximately 10^{-9} A) can be switched on and off by the addition of one-half the electron charge to the control capacitor. These devices are known as single-electron transistors (SETs) and are by far the most sensitive electrometers (instruments for measuring electrical charge) presently known. *See* ELECTROMETER; TRANSISTOR.

Building on these concepts, devices employing several single-electron transistors have been demonstrated in which a modulation of the control capacitor with a frequency f leads to an electrical current precisely equal to ef. In such devices, electrons are transferred one by one onto and off the grain. It is even possible to measure jumps in the current corresponding to the transfer of each electron (**Fig. 2**). Such devices are expected to have metrological applications as an absolute standard of electric current. *See* ELECTRICAL UNITS AND STANDARDS.

A. Douglas Stone

Bibliography. B. L. Altshuler, P. A. Lee, and R. A. Webb (eds.), *Mesoscopic Phenomena in Solids*, 1991; M. H. Devoret, D. Esteve, and C. Urbina, Single-electron transfer in metallic nanostructures, *Nature*, 360:547–553, 1992; A. Khurana, Ballistic electron transport through a narrow channel is quantized, *Phys. Today*, 41(11):21–23, 1988; M. A. Reed, Quantum dots, *Sci. Amer.*, 268(1):118–123, January 1993; R. A. Webb and S. Washburn, Quantum interference fluctuations in disordered metals, *Phys. Today*, 41(12):46–55, December 1988.

Mesosphere

A layer within the Earth's atmosphere that extends from about 50 to 85 km (31 to 53 mi) above the surface. The mesosphere is predominantly characterized by its thermal structure. On average, mesospheric temperature decreases with increasing height.

Structure. Temperatures range from as high as 12°C (53°F) at the bottom of the mesosphere to as low as −133°C (−208°F) at its top. The top of the mesosphere, called the mesopause, is the coldest area of the Earth's atmosphere. Temperature increases with increasing altitude above the mesopause in the layer known as the thermosphere, which absorbs the Sun's extreme ultraviolet radiation. In the stratosphere, the atmospheric layer immediately below the mesosphere, the temperature also increases with height. The stratosphere is where ozone, which absorbs ultraviolet radiation from the Sun, is most abundant. The transition zone between the mesosphere and the stratosphere is called the stratopause. Mesospheric temperatures are comparatively cold because very little solar radiation is absorbed in this layer. Meteorologists who predict weather conditions or study the lowest level of the Earth's atmosphere, the troposphere, often refer to the stratosphere, mesosphere, and thermosphere collectively as the upper atmosphere. However, scientists who study these layers distinguish between them; they also refer to the stratosphere and mesosphere as the middle atmosphere. *See* ATMOSPHERE; METEOROLOGY; STRATOSPHERE; THERMOSPHERE; TROPOSPHERE.

In the lower part of the mesosphere, the difference between the temperature at the summer and winter poles is of order 35°C (63°F). Temperatures in the upper mesosphere are colder in summer and warmer in winter, resulting in meridional flow from the summer to the winter hemisphere (**illus.**). Although the temperature gradient in the upper part of the mesosphere remains large, additional complications result in wind speeds that are much slower than they are in the lower part of the mesosphere. Winds in the east-west or zonal direction are greatest at mesospheric middle latitudes. Zonal winds blow toward the west in summer and toward the east in winter. Like their meridional counterparts, zonal winds are comparatively strong near the bottom of the mesosphere and comparatively weak near the top. Thus, on average both temperature and wind speed decrease with increasing height in the mesosphere. *See* ATMOSPHERIC GENERAL CIRCULATION.

Variations. The mesosphere is affected by the atmospheric layers lying above and below it. Above the mesosphere, the thermosphere profoundly responds to solar eruptions and the ensuing storms that occur in space. These geomagnetic storms produce

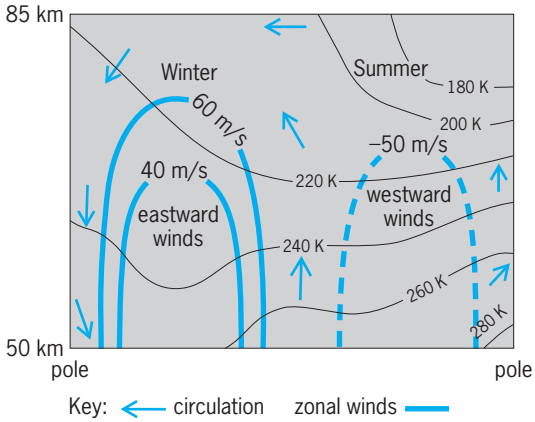

Structure of the mesosphere.

compositional changes in the thermosphere that can be driven down into the upper mesosphere by the associated disturbance winds when the storms are sufficiently strong.

Variations in winds, temperature, pressure, and density that are excited beneath the mesosphere in the troposphere and stratosphere can propagate up. Because density decreases with altitude throughout the Earth's atmosphere, any upward propagating variations will grow in size in order to conserve energy. Upward propagating variations include gravity waves that are excited by winds blowing over mountain ranges or by storm fronts. They also include solar tides which are excited by the periodic absorption of solar infrared and ultraviolet radiation in the troposphere and stratosphere, respectively. Gravity waves, solar tides, and other planetary waves are very large when they enter the mesosphere. They are readily observed by lidars, radars, imagers, and other sensors that measure the mesosphere remotely both from the ground and from satellites. Most waves of lower atmospheric origin dissipate or break in the upper mesosphere. When they do so, their energy and momentum is transferred to the rest of the undisturbed atmosphere. This complicates upper mesospheric circulation. Dissipating waves slow the winds that are driven by the summer-to-winter temperature difference. In addition, the wave dissipation process produces a mesopause temperature anomaly. That is, the summer mesopause is actually colder than the winter mesopause. *See* MIDDLE-ATMOSPHERE DYNAMICS; UPPER-ATMOSPHERE DYNAMICS.

Special features of upper mesosphere. Meteors which enter the Earth's atmosphere vaporize in the upper mesosphere. These meteors contain significant amounts of metallic atoms and molecules which may ionize. Metallic ions combined with ionized water clusters make up a large part of the D-region ionosphere that is embedded in the upper mesosphere. *See* IONOSPHERE.

The upper mesosphere is also where iridescent blue clouds can be seen with the naked eye and photographed in twilight at high summer latitudes when the Sun lights them up in the otherwise darkening sky. These clouds are called noctilucent clouds (NLC). Noctilucent clouds are believed to be tiny ice crystals that grow on bits of meteoric dust. The first reliable reports of noctilucent clouds were made in the 1880s. Since then, the number of sightings of noctilucent clouds has significantly increased, as have their sizes and lifetimes. Changes that are observed in the atmosphere are increasingly linked to human activities. Many scientists think that the depletion of stratospheric ozone, the occurrence of the Antarctic "ozone hole," and global warming in the lower atmosphere may be anthropogenic effects. While it remains to be proven, lower atmospheric global warming could be due to documented increases in atmospheric methane and carbon dioxide that are by-products of agricultural activities and fossil fuel consumption on Earth. These greenhouse gases also affect the mesosphere, but in very different ways. Many experts believe that increases of noctilucent clouds could be related to the documented increases in greenhouse gases. *See* GREENHOUSE EFFECT.

Large-scale atmospheric circulation patterns transport tropospheric air containing methane and carbon dioxide from the lower atmosphere into the middle atmosphere. While carbon dioxide warms the lower atmosphere, it cools the middle and upper atmosphere by releasing heat to space. When it reaches the middle atmosphere, methane undergoes chemical reactions that produce water. If the air is sufficiently cold, the water can freeze and form noctilucent clouds. Temperatures must be below −129°C (−200°F) for noctilucent clouds to form. These conditions are common in the cold summer mesopause region at high latitudes. One way to explain the observed increases in noctilucent clouds over the course of the last century is the plausible but unconfirmed increase in mesospheric water vapor. The most important source of water vapor in the mesosphere, atmospheric methane, has doubled in the last century. Further, any comparable increase in mesospheric carbon dioxide levels only enhances the conditions favorable for noctilucent cloud formation by further cooling the region. These effects combined may eventually lead to larger and more persistent clouds at high latitudes, accompanied by smaller thinner formations at middle latitudes. In June 1999, reliable measurements of noctilucent clouds were made in the continental United States over Colorado and Utah. This unprecedented sighting surprised experts who predicted that noctilucent clouds would not occur at these middle latitude locations until well into the twenty-first century. This recent sighting may be the strongest evidence yet that the mesosphere is where scientists can eventually confirm whether atmospheric changes are linked to human activities. *See* METHANE.

Maura Hagan

Bibliography. D. G. Andrews, J. R. Holton, and C. B. Leovy, *Middle Atmosphere Dynamics*, 1987; G. Brasseur and S. Solomon, *Aeronomy of the Middle Atmosphere*, 1985; J. T. Houghton, *The Physics of Atmospheres*, 1986; G. E. Thomas et al., Relation between increasing methane and the presence of ice clouds at the mesopause, *Nature*, 338:490–492, 1989.

Mesozoa

A division of the animal kingdom sometimes ranked as intermediate between the Protozoa and the Metazoa. These animals are unassignable to any of the better-known phyla, as usually defined. In the absence of proof concerning their relationships, and in view of the disagreement among zoologists relative to their affinities and even with respect to the facts and interpretation of their structure and life cycle, they are treated as a small phylum somewhere between Protozoa and Platyhelminthes. No particular phylogenetic interpretation should be attached to this placement.

The Mesozoa comprise two orders of small, wormlike organisms, the Dicyemida and the Orthonectida. Both are parasitic in marine invertebrates. The body consists of a single layer of ciliated cells enclosing one or more reproductive cells. These body cells are rather constant in number and arrangement for any given species. The internal cells do not correspond to the entoderm of other animals, as they have no digestive function. The life cycles are complex, involving both sexual and asexual generations (metagenesis).

There has been little agreement regarding the affinities of the Mesozoa. Some zoologists believe them to be related most closely to the Protozoa; others believe them derived from some higher group, usually the flatworms. Other groups that have also been suggested as a link with the Mesozoa are the Coelenterata, Echiuroidea, and primitive Aschelminthes. However, the relationship of these groups to the Mesozoa is regarded as less likely by most zoologists. Research indicates that Orthonectida and Dicyemida are more different from each other than formerly realized and probably should not be grouped in one phylum.

The Mesozoa have been united into a class called Planuloidea or Moruloidea because they exhibit the same grade of structural organization as a coelenterate planula larva or an embryonic morula.

A number of other small enigmatic organisms have at times been included in the Mesozoa. Most of these have subsequently been assigned to other phyla. *Salinella*, described by J. Frenzel in 1892, has remained a taxonomic puzzle. It consists of a single layer of cells enclosing a digestive cavity (see **illus.**). There is a mouth at one end and an anal opening at the other. Both internal and external surfaces of all cells are ciliated. Asexual reproduction is by transverse fission. Frenzel observed encystment in pairs and described unicellular young resembling ciliates.

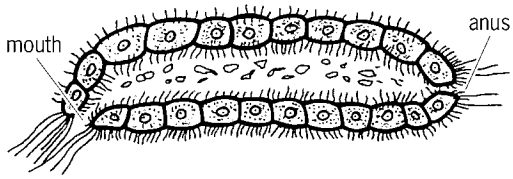

***Salinella*, in longitudinal section.**

Salinella has not been reported again since Frenzel's discovery. *See* ANIMAL KINGDOM; CELL CONSTANCY.

Bayard H. McConnaughey

Bibliography. L. H. Hyman, *The Invertebrates*, vol. 1, 1940; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; H. W. Stunkard, The life history and systematic relations of the Mesozoa, *Quart. Rev. Biol.*, 29(3):230–244, 1954.

Mesozoic

The middle era of the three major divisions of the Phanerozoic Eon (Paleozoic, Mesozoic, and Cenozoic eras) of geologic time, encompassing an interval from 251 to 65 million years ago (Ma) based on various isotopic-age dates. The Mesozoic Era is known also as the Age of the Dinosaurs and the interval of middle life. The Mesozoic Erathem (the largest recognized time-stratigraphic unit) encompasses all sedimentary rocks, body and trace fossils of organisms preserved, metamorphic rocks, and intrusive and extrusive igneous rocks formed during the Mesozoic Era. *See* GEOCHRONOMETRY.

The Mesozoic Era was originally named for one of three principal divisions of the fossil record, or history of life, that was bounded before and after by significant mass extinctions that dramatically changed the biotic composition of the world. In England during the early 1840s, geologist John Phillips introduced the terms Mesozoic Era and Cenozoic Era, in conjunction with geologist Adam Sedwick's term Paleozoic Era, proposed in 1838, to denote the widespread observation that three successive and distinct biotic assemblages were preserved in the fossil record. The Mesozoic Era comprises life intermediate in kind between ancient life-forms (Paleozoic

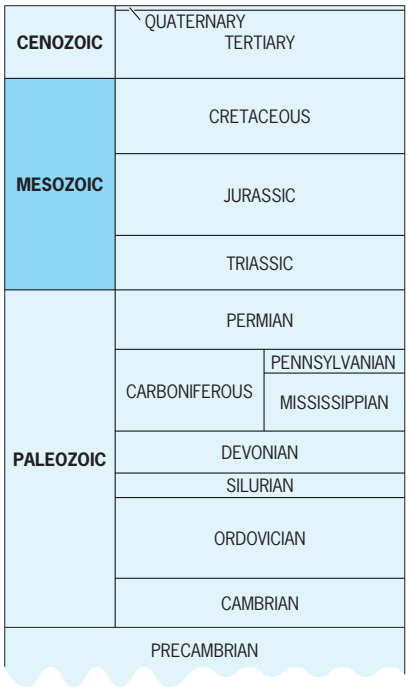

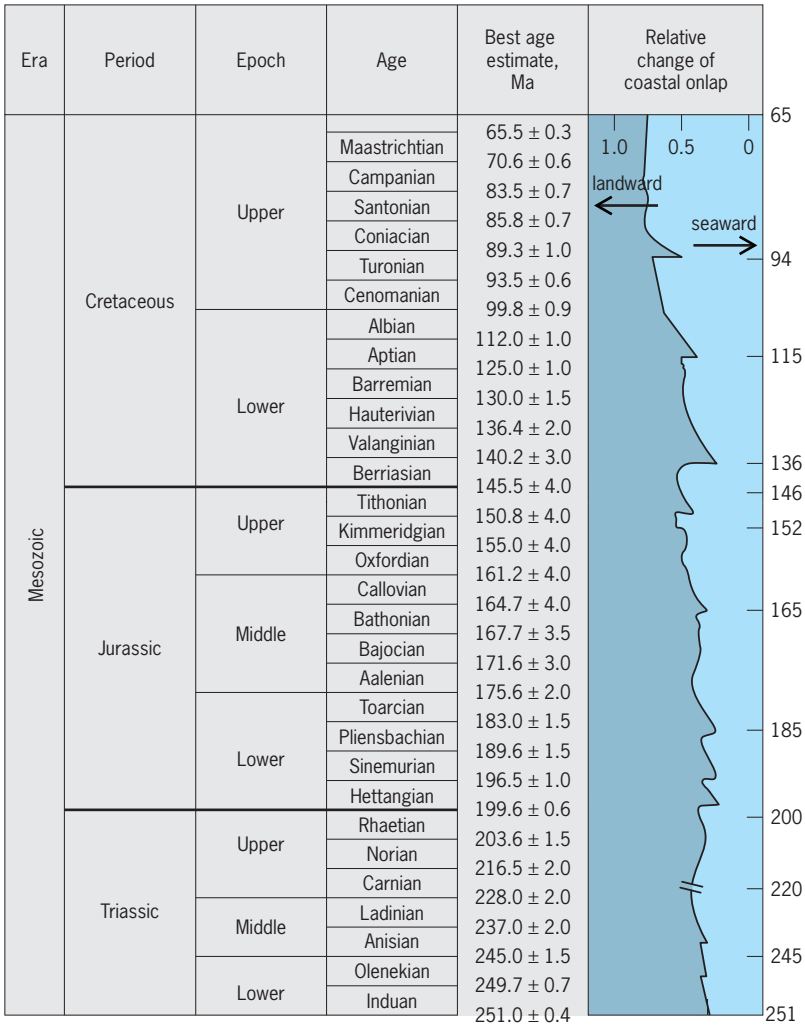

Fig. 1. Subdivisions of the Mesozoic Era, including the best age estimates and the eustatic sea-level curve depicted as relative change in coastal onlap as the shoreline moved landward (sea-level rise) or seaward (sea-level fall).

Era) and recent life-forms (Cenozoic Era). *See* CENOZOIC; PALEOZOIC.

The Mesozoic Era records dramatic changes in the geologic and biologic history of the Earth. At the beginning of the Mesozoic Era, all the continents were amassed into one large supercontinent, Pangaea. Both the marine and continental biotas were impoverished from the mass extinction that marked the boundary between the Permian and Triassic periods, and the end of the Paleozoic Era. This mass extinction was responsible for the loss of over 90% of the species on Earth. During the Mesozoic Era, many significant events were recorded in the geologic and fossil record of the Earth, including the breakup of Pangaea and the evolution of modern ocean basins by continental drift, the rise of the dinosaurs, the ascension of the angiosperms (flowering plants), the diversification of the insects and crustaceans, and the appearance of the mammals and birds. The end of the Mesozoic Era is marked by a major mass extinction at the Cretaceous-Tertiary boundary that records several meteorite impacts, the extinction of the dinosaurs, the rise to dominance of the mammals, and the beginning of the Cenozoic Era and the life-forms dominant today. *See* CONTINENTAL DRIFT; PLATE TECTONICS.

The Mesozoic Era comprises three periods of geologic time: the Triassic Period (251–200 Ma), the Jurassic Period (200–146 Ma), and the Cretaceous Period (146–65 Ma) [**Fig. 1**]. These periods are each subdivided into epochs, formal designations of geologic time described as Early, Middle, and Late (except for the Cretaceous, which has no middle epoch designated yet). The packages of rock themselves are subdivided into series designated Lower, Middle, and Upper (except for Cretaceous). Each epoch is subdivided into ages. Likewise, each series is subdivided into stages, which are time-stratigraphic units whose boundaries are based on unconformities, hiatuses, or erosional surfaces, on correlations to a type section (where rocks are first described), or preferably on changes in the biota that depict true measurable time (for example, evolutionary changes). *See* UNCONFORMITY.

The correlations of time equivalency and age dating in the Mesozoic Era have been accomplished by utilizing biostratigraphic zones based on individual fossil groups or by an acme or composite zonal assemblage based on numerous fossil groups. Marine and continental fossil groups used to describe chronologically Mesozoic rocks include marine foraminifera and nannofossils (shelled protozoa), ammonites (cephalopods), and inoceramids (mollusks); continental plant spores and pollen (palynology); dinosaurs; and mammals. Correlations based on these and other organisms in the Mesozoic Era depend on the faunal and floral succession through origination and extinction of species. *See* CEPHALOPODA; FORAMINIFERIDA; MOLLUSCA; PALYNOLOGY.

The organization of subdivisions based on physical and biological evidence allows geologists and paleontologists to describe both rocks and fossils in specific intervals of time and space. Thus, earth scientists can communicate effectively with one another and characterize more precisely the physical and biotic changes during the Mesozoic Era, as well as the other eras in geologic history. *See* PALEONTOLOGY; STRATIGRAPHY.

Triassic. The Mesozoic Era begins with the Triassic Period, which constitutes nearly one-third the total time of the era and is well exposed especially in Europe and North America, with other important outcrops in India, China, Argentina, and South Africa. The Triassic Period was named originally the Trias in Germany in 1834 by Friedrich August von Alberti for its unique fauna and natural division into three distinct stratigraphic units.

As a result of the unique geography of the single Pangaean landmass, the alteration in oceanic currents produced around one continent, and the monsoonal climatic setting, life changed substantially in both marine and continental ecosystems. The marine ecosystems witnessed the addition of large reptiles

and the modern reef-building corals, the reemergence and diversification of the mollusks, and the emergence of pelagic life in the form of planktonic organisms. Ray-finned and bony fishes and sharks dominated the seas. Placodonts and nothosaurs were aquatic marine reptiles that fed on mollusks and other marine invertebrates. Ichthyosaurs appeared in the oceans for the first time. Freshwater and terrestrial ecosystems were marked by the emergence and diversification of the dinosaurs, flying reptiles, frogs, turtles, terrestrial crocodiles, and birds; the appearance of the mammals, though quite small in size; the emergence of freshwater and terrestrial crayfish; and the emergence of new insects, such as the Isoptera (termites), Diptera (flies), and the Hymenoptera (bees, wasps, and possibly ants), appearing earlier in the Mesozoic than previously thought. Trace fossil evidence for these new insects indicates the advent of social behavior in termites and in primitive bees, prior to the appearance of angiosperms in the Cretaceous. In terrestrial ecosystems ferns and seed ferns were abundant, but gymnosperm floras continued to dominate the landscape. Therapsids rediversified after the Permo-Triassic extinctions, and thecodonts gave rise to the crocodiles and to the first dinosaurs, which were small in stature.

During the Triassic Period the continents were amassed tectonically into one great landmass, the supercontinent Pangaea, that was distributed equally across the paleoequator in both the Northern and Southern hemispheres (**Fig. 2***a*). Since the majority of the enormous Pangaean landmass was inland from the influence of the ocean, and its configuration distributed equally across the Equator, a worldwide monsoonal climate pattern dominated during the Triassic that created alternating wet and dry seasons in many regions. Areas landward of the coasts experienced increased continentality of the climate and produced more pronounced wet and dry seasons. *See* PALEOCLIMATOLOGY; PALEOGEOGRAPHY.

At the end of the Triassic, Pangaea began to break apart and the monsoonal climate pattern began to disintegrate. Evidence for the breakup of Pangaea and the eventual formation of the northern Atlantic Ocean is the presence of rift basins along the east coast of North America and the northwest coast of Africa. A mass extinction defines the boundary between the end of the Triassic and the beginning of the Jurassic. This mass extinction was responsible for the loss of about 60% of the species on Earth. The mass extinctions in the marine and continental realms affected the ocean ecosystem by eliminating the marine conodonts and placodont reptiles, and many species of bivalves, ammonoids, plesiosaurs, and ichthyosaurs disappeared. Most of these groups recovered in the Jurassic. Extinction also claimed the large amphibians and mammallike reptiles from freshwater and terrestrial ecosystems. The cause of these mass extinctions is unknown, though some scientists hypothesize that either a meteorite impact or increasing global aridity caused many genera to go extinct. *See* TRIASSIC.

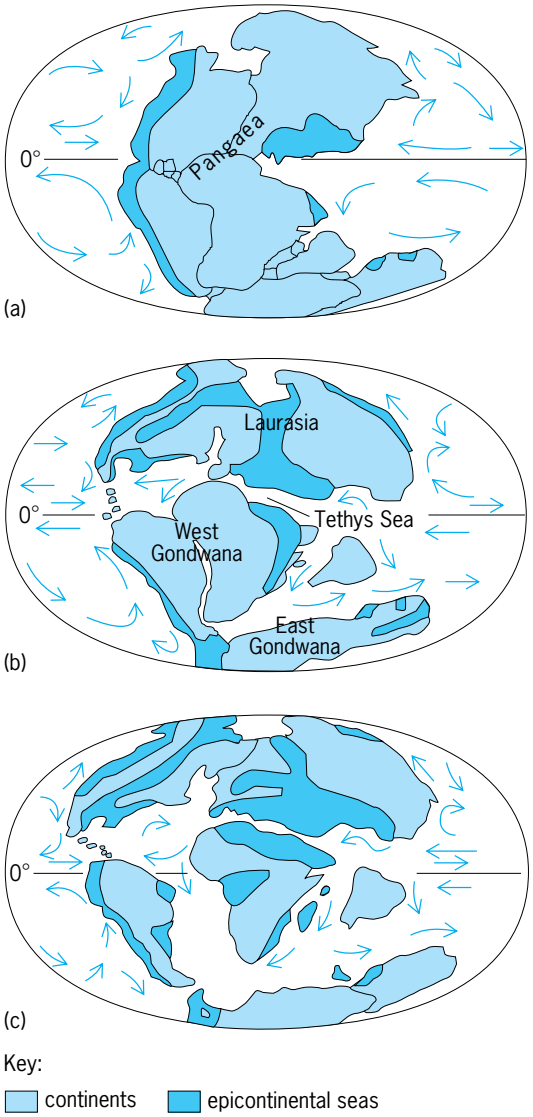

Fig. 2. Schematic reconstruction showing paleogeography of continents, epicontinental seas, and ocean basins (arrows denote ocean currents) on Pangaea in the Mesozoic Era from the (*a*) Triassic (220 Ma), (*b*) Jurassic (155 Ma), and (*c*) Cretaceous (70 Ma) periods.

Jurassic. The middle part of the Mesozoic Era is represented by the Jurassic Period, which constitutes about one-third of the total time of the era. Jurassic rocks are well exposed, especially in North America and Europe, and other important outcrops exist in South America and Asia. In 1839, German geologist Leopold von Buch established the Jurassic as a system for rocks in Switzerland, Germany, and England. The new system was based on descriptions of equivalent rocks made by the German geologist Alexandre von Humboldt (1795) and the English geologist William Smith (1797–1815). They described massive limestones of the Jura Mountains in Switzerland as the Jura-Kalkstein and the Lias-Oolite rock sequences in England and Wales, respectively. *See* LIMESTONE; OOLITE.

During the Jurassic Period the Pangaean landmass continued to separate into two large continen-

tal masses, with one in the Northern Hemisphere and the other in the Southern (Fig. 2*b*). The Northern Hemisphere landmass, Laurasia, was composed of North America and Eurasia, while the Southern Hemisphere landmass, Gondwanda, was composed of South America, Africa, India, Antarctica, and Australia. Continued plate spreading and more rapid sea-level fluctuations late in the Jurassic created the Tethyan Seaway, which extended between Laurasia and Gondwanda, allowing oceans to flow freely between the continents, and caused epicontinental seas to flood large areas of North America and Europe. The opening of the ocean basins and the resulting increased oceanic circulation created a zonal atmospheric climate pattern that ranged from tropical at the Equator to warm temperate near the Poles, with local zones of aridity due to orographic and latitudinal rain shadows. *See* CONTINENTS, EVOLUTION OF; RAIN SHADOW.

Oceanic and continental biotas shifted in composition during tectonic and climatic transformations of the Jurassic Period. Numerous reef communities of modern reef-building corals flourished in shallow tropical oceans along with bivalves, ammonites, belemnoids, sea urchins, and fishes. Planktonic life began to prosper in the warm, shallow seas with the appearance of calcareous nannoplankton. Marine reptiles included plesiosaurs and ichthyosaurs, and the invasion of the oceans by crocodiles. Terrestrial and freshwater ecosystems were dominated by plants such as cycads, cycadeoids, conifers, ginkgos, and to a lesser extent ferns. The Jurassic is also known as the age of the cycads. The dinosaurs' greatest rise to dominance occurred with the radiation of large herbivores such as the sauropods *Apatosaurus, Diplodocus, Camarasaurus,* the plated stegosaurs, and the heavily armored ankylosaurs. The herbivores were pursued by predatory dinosaurs such as *Ceratosaurus* and *Allosaurus*. Many flying reptiles speckled the skies, including pterosaurs and the feathered reptilelike bird, *Archaeopteryx*. Mammals were still small in size but began to increase in diversity. *See* DINOSAUR; PHYTOPLANKTON; REPTILIA; ZOOPLANKTON.

The end of the Jurassic Period was marked by moderate extinctions of biota in both marine and continental ecosystems. In the marine realm, brachiopod diversity declined steadily, other invertebrate faunas varied in diversity, and marine reptiles, such as the ichthyosaur, became nearly extinct. On the continents, the major extinctions eliminated the last of the therapsids and affected the large herbivorous sauropods, stegosaurs, and ankylosaurs, as well as their predators. *See* JURASSIC.

Cretaceous. The last part of the Mesozoic is represented by the Cretaceous Period, which constitutes a little less than one-half the total time of the Mesozoic Era. The Cretaceous is represented well by rocks in North America, South America, Europe, Asia, Africa, and Australia. The Cretaceous Period was named in 1822–1823 by French geologist J. J. d'Omalius d'Halloy for exposures at the White Cliffs of Dover, which are composed of marine chalks and can be traced throughout Europe and North Africa. These widespread chalk units are composed predominantly of microscopic plates of calcareous nannoplankton, and had once been an ancient sea floor. *See* CHALK; MICROPALEONTOLOGY.

During the Cretaceous Period the face of the Earth began to take on an appearance more similar to the present continental and oceanic configuration (Fig. 2*c*). Early in the Cretaceous, both Laurasian and Gondwanan continental masses separated into the continents still recognizable today. The separation of Gondwana in the Early Cretaceous marked the onset to the formation of the South Atlantic Ocean. Increased sea-floor spreading rates, which opened the oceanic gaps between the continents, resulted in the expansion of shallow epicontinental seaways in North America, Africa, and Northern Europe, along with the flooding of most of southern Eurasia. These variations in sea-floor spreading rates caused the sea level to fluctuate constantly throughout most of the Cretaceous. The Pacific and Atlantic ocean basins began taking form, and the Tethyan Seaway, between what is now the western Mediterranean and southeastern Asia, persisted throughout most of the period. New, expanded oceanic realms coupled with the changed continental configurations and increased atmospheric carbon dioxide transformed the early Cretaceous climate into a humid zonal climate, warmer than today. The climate began cooling down, beginning near the end of the Mesozoic.

The biotic composition during the Cretaceous Period contained a mixture of both intermediate and modern forms of life in both marine and continental ecosystems. In the marine realm, modern types of gastropods, bivalves, and modern fishes shared the oceans with marine reptiles such as mosasaurs and plesiosaurs, ammonoids, belemnoids, and other gigantic, coiled oysters and sedentary bivalves. Other marine invertebrates, such as planktonic and benthic foraminifera, flourished together with bryozoans, corals, reef-building rudist bivalves, crabs, lobsters, and other crustaceans. In the continental Cretaceous realm, the greatest change in freshwater and terrestrial ecosystems occurred with the appearance and diversification of angiosperms or flowering plants that became more diverse over the early and middle Mesozoic gymnosperm floras near the end of the Cretaceous. At the same time, freshwater and terrestrial insects continued to diversity and exploit new niches and resources provided by the angiosperms. Many groups of vertebrates, including snakes, modern types of turtles, crocodiles, lizards, and amphibians, diversified from ancient stocks during the Cretaceous to coexist with the dinosaurs that continued to rule the Earth. Mammals continued to evolve and diversify, but remained very small in size in comparison with their modern descendants. *See* MAGNOLIOPHYTA; MAMMALIA.

The dinosaurs diversified for the last time in the Cretaceous and formed ecological communities similar to mammal faunas inhabiting the African plains today. Herbivores such as the great horned

dinosaur *Triceratops* and large duck-billed dinosaurs such as *Hadrosaurus* traveled in herds, roamed the plains, and followed watercourses in seasonally migrating for food. The predators that followed these herds were the largest carnivores of all times, *Albertosaurus* and *Tyrannosaurus*, together with other pack and ambush predators such as the velociraptors and crocodiles, respectively. The few flying reptiles that remained were spectacular, one form attaining a wingspan of nearly 11 m (35 ft), and these creatures shared the skies with modern types of shorebirds and wading birds.

The end of the Cretaceous, and thus the end of the Mesozoic Era, is marked by a mass extinction known as the Cretaceous-Tertiary boundary (Tertiary is the earliest system of the Cenozoic Era, now divided into the Paleogene and Neogene). This mass extinction is widely known for the demise of 60% of the organisms on Earth, including the ammonoids, the rudist corals, marine reptiles, the dinosaurs, and the flying reptiles. There was a large reduction in the diversity of various marine plankton and continental faunas and floras, but they later recovered early in the Cenozoic. There have been many heated debates over the cause of the terminal Mesozoic extinctions, because some of them focus on extraterrestrial causes such as bolide and comet impacts onto the Earth's surface. Intriguing evidence in the form of iridium anomalies comes from marine deposits and continental coal deposits within rocks spanning the Cretaceous and Tertiary boundary worldwide. Iridium is an element that is typically depleted in rocks derived from the Earth's curst but is enriched in extraterrestrial stony meteorites. Some scientists hypothesize that a large 10-km (6-mi) meteorite or comet struck the Earth, exploded on impact, and released a blast greater than all the nuclear weapons on Earth. Such an impact would have thrown enormous volumes of dust and smoke into the atmosphere, blocked a large fraction of sunlight, and thus caused severe hardships to all marine and continental biota by lowering worldwide temperatures and disrupting the food chain. If the impact occurred in the ocean, huge volumes of water would have been vaporized instantly and caused gigantic tsunamis or tidal waves, which could have swept across most of the lowlands of continents. The devastating results of such an impact have been termed nuclear winter, because the effect would be similar to that produced by an all-out nuclear war. Other scientists hypothesize that extensive volcanic outgassing and cooler climates due to plate tectonic movements produced major changes in global climate and environmental disturbances that forced many organisms into extinction. Despite the similarities in global climate change and mass extinction interpreted by both of these hypotheses, other environmentally sensitive and presumably vulnerable groups of organisms were little affected by the mass extinction event. Environmentally sensitive vertebrates, such as crocodiles, lizards, turtles, frogs, and salamanders, were paradoxically spared and made it through the mass extinctions with little loss of species. The birds also persist through the mass extinctions and in fact radiate in the Paleogene and Neogene. *See* CRETACEOUS; EXTINCTION (BIOLOGY); GEOLOGIC TIME SCALE; METEORITE; TERTIARY.

S. T. Hasiotis; R. F. Dubiel

Bibliography. M. V. Caputo, J. A. Peterson, and K. J. Franczyk (eds.), *Mesozoic Systems of the Rocky Mountain Region, USA*, Rocky Mountain Section, Society for Sedimentary Geology, 1994; B. U. Haq and F. W. B. van Eysinga, *Geologic Time Table Chart*, 4th ed., 1994; W. B. Harland et al., *A Geologic Time Scale*, 1989; M. Moullade and A. E. M. Nairn (eds.), *The Phanerozoic Geology of the World II: The Mesozoic*, 1978; S. M. Stanley, *Earth System History*, 2d ed., 1998.

Messier Catalog

An early listing of nebulae and star clusters. Although Charles Messier (1730–1817) was primarily a comet hunter, he is remembered for his catalog of objects. As the first serious and systematic comet hunter (his comet discoveries spanned from 1760 to 1801), Messier learned about the hazards of being misled by the galaxies, nebulae, and star clusters all over the sky that look like comets. In 1758, when Messier found a fuzzy patch near the star beta Tauri that he first thought was a comet but that never moved as a comet would, he began to build a catalog of these objects. "This nebula had such a resemblance to a comet, in its form and brightness," Messier wrote, "that I endeavored to find others, so that astronomers would not confuse these same nebulae with comets just beginning to shine." This first entry is the Crab Nebula, the remnant of the supernova, a near-total destruction of a star, that was observed on July 4, 1054, to be as bright as Venus. (The Crab Nebula was previously discovered by John Bevis.) *See* COMET; CRAB NEBULA; GALAXY, EXTERNAL; NEBULA; PLANETARY NEBULA; STAR CLUSTERS.

Messier published three versions of his catalog. Of the 45 objects of his first catalog that appeared in 1774, no more than 18 were his own discoveries. Ultimately he listed 103 objects.

Evolution. It was not untill 1921 that Messier's catalog underwent some evolution, when Camille Flammarion, a popularizer of astronomy, found some notes about another object in Messier's personal copy of the catalog. That object was the Sombrero Galaxy in Virgo, and was added to the list as M104. In 1947 the astronomer Helen Sawyer Hogg suggested that Messier was aware of additional objects found by his colleague Pierre Méchain. Not long afterward, Owen Gingerich suggested that two galaxies in Ursa Major that were noted in the original copy should be M108 and M109. Kenneth Glyn Jones proposed the addition of the fainter companion to M31, the Andromeda Galaxy, since Messier had observed it in 1773 and drawn it in 1807. Thus, Messier's catalog lists 110 clusters, nebulae, and galaxies (see **table**). Not all those objects belong; M102 seems to be a misplaced position for M101, and M40 is simply two faint stars.

Messier Catalog of nebulae and star clusters*

Messier number†	NGC or (IC)	Right ascension (2000)	Declination (2000)	Apparent visual magnitude	Description
1	1952	5h 34.5m	+22°01′	8.4	Crab Nebula in Taurus; remains of SN 1054
2	7089	21h 33.5m	−0°49′	6.5	Globular cluster in Aquarius
3	5272	13h 42.2m	+28°23′	6.4	Globular cluster in Canes Venatici
4	6121	16h 23.6m	−26°32′	5.9	Globular cluster in Scorpius
5	5904	15h 18.6m	+2°05′	5.8	Globular cluster in Serpens
6	6405	17h 40.1m	−32°13′	4.2	Open cluster in Scorpius
7	6475	17h 53.9m	−34°49′	3.3	Open cluster in Scorpius
8	6523	18h 03.8m	−24°23′	5.8	Lagoon Nebula in Sagittarius
9	6333	17h 19.2m	−18°31′	7.9	Globular cluster in Ophiuchus
10	6254	16h 57.1m	−4°06′	6.6	Globular cluster in Ophiuchus
11	6705	18h 51.1m	−6°16′	5.8	Open cluster in Scutum
12	6218	16h 47.2m	−1°57′	6.6	Globular cluster in Ophiuchus
13	6205	16h 41.7m	+36°28′	5.9	Globular cluster in Hercules
14	6402	17h 37.6m	−3°15′	7.6	Globular cluster in Ophiuchus
15	7078	21h 30.0m	+12°10′	6.4	Globular cluster in Pegasus
16	6611	18h 18.8m	−13°47′	6.0	Open cluster with nebulosity (Eagle Nabula) in Serpens
17	6618	18h 20.8m	−16°11′	7	Swan or Omega Nebula in Sagittarius
18	6613	18h 19.9m	−17°08′	6.9	Open cluster in Sagittarius
19	6273	17h 02.6m	−26°16′	7.2	Globular cluster in Ophiuchus
20	6514	18h 02.6m	−23°02′	8.5	Trifid Nebula in Sagittarius
21	6531	18h 04.6m	−22°30′	5.9	Open cluster in Sagittarius
22	6656	18h 36.4m	−23°54′	5.1	Globular cluster in Sagittarius
23	6494	17h 56.8m	−19°01′	5.5	Open cluster in Sagittarius
24	6603	18h 16.9m	−18°29′	4.5	Open cluster in Sagittarius
25	(4725)	18h 31.6m	−19°15′	4.6	Open cluster in Sagittarius
26	6694	18h 45.2m	−9°24′	8.0	Open cluster in Scutum
27	6853	19h 59.6m	+22°43′	8.1	"Dumbbell" planetary nebula in Vulpecula
28	6626	18h 24.5m	−24°52′	6.9	Globular cluster in Sagittarius
29	6913	20h 23.9m	+38°32′	6.6	Open cluster in Cygnus
30	7099	21h 40.4m	−23°11′	7.5	Globular cluster in Capricornus
31	224	0h 42.7m	+41°16′	3.4	Andromeda Galaxy: spiral galaxy (Sb)
32	221	0h 42.7m	+40°52′	8.2	Elliptical galaxy; companion to M31
33	598	1h 33.9m	+30°39′	5.7	Spiral galaxy (Sc) in Triangulum
34	1039	2h 42.0m	+42°47′	5.2	Open cluster in Perseus
35	2168	6h 08.9m	+24°20′	5.1	Open cluster in Gemini
36	1960	5h 36.1m	+34°08′	6.0	Open cluster in Auriga
37	2099	5h 52.4m	+32°33′	5.6	Open cluster in Auriga
38	1912	5h 28.7m	+35°50′	6.4	Open cluster in Auriga
39	7092	21h 32.2m	+48°26′	4.6	Open cluster in Cygnus
40		12h 22.4m	+58°05′	8	Close double star in Ursa Major
41	2287	6h 47.0m	−20°44′	4.5	Loose open cluster in Canis Major
42	1976	5h 35.4m	−5°27′	4	Orion Nebula
43	1982	5h 35.6m	−5°16′	9	Northeast portion of Orion Nebula
44	2632	8h 40.1m	+19°59′	3.1	Praesepe; open cluster in Cancer
45		3h 47.0m	+24°07′	1.2	The Pleiades; open cluster in Taurus
46	2437	7h 41.8m	−14°49′	6.1	Open cluster in Puppis
47	2422	7h 36.6m	−14°30′	4.4	Loose group of stars in Puppis
48	2548	8h 13.8m	−5°48′	5.8	"Cluster of very small stars" in Hydra
49	4472	12h 29.8m	+8°00′	8.4	Elliptical galaxy in Virgo
50	2323	7h 03.2m	−8°20′	5.9	Loose open cluster in Monoceros
51	5194, 5195	13h 29.9m	+47°12′	8.1	Whirlpool spiral galaxy (Sc) in Canes Venatici
52	7654	23h 24.2m	+61°35′	6.9	Loose open cluster in Cassiopeia
53	5024	13h 12.9m	+18°10′	7.7	Globular cluster in Coma Berenices
54	6715	18h 55.1m	−30°29′	7.7	Globular cluster in Sagittarius
55	6809	19h 40.0m	−30°58′	7.0	Globular cluster in Sagittarius
56	6779	19h 16.6m	+30°11′	8.2	Globular cluster in Lyra
57	6720	18h 53.6m	+33°02′	9.0	Ring Nebula; planetary nebula in Lyra
58	4579	12h 37.7m	+11°49′	9.8	Spiral galaxy (SBb) in Virgo
59	4621	12h 42.0m	+11°39′	9.8	Elliptical galaxy in Virgo
60	4649	12h 43.7m	+11°33′	8.8	Elliptical galaxy in Virgo
61	4303	12h 21.9m	+4°28′	9.7	Spiral galaxy (Sc) in Virgo
62	6266	17h 01.2m	−30°07′	6.6	Globular cluster in Ophiuchus
63	5055	13h 15.8m	+42°02′	8.6	Spiral galaxy (Sb) in Canes Venatici
64	4826	12h 56.7m	+21°41′	8.5	Spiral galaxy (Sb) in Coma Berenices
65	3623	11h 18.9m	+13°05′	9.3	Spiral galaxy (Sa) in Leo

(*cont.*)

Messier Catalog of nebulae and star clusters* (*cont.*)

Messier number‡	NGC or (IC)	Right ascension (2000)	Declination (2000)	Apparent visual magnitude	Description
66	3627	11h 20.2m	+12°59′	9.0	Spiral galaxy (Sb) in Leo; companion to M65
67	2682	8h 50.4m	+11°49′	6.9	Open cluster in Cancer
68	4590	12h 39.5m	−26°45′	8.2	Globular cluster in Hydra
69	6637	18h 31.4m	−32°21′	7.7	Globular cluster in Sagittarius
70	6681	18h 43.2m	−32°18′	8.1	Globular cluster in Sagittarius
71	6838	19h 53.8m	+18°47′	8.3	Globular cluster in Sagitta
72	6981	20h 53.5m	−12°32′	9.4	Globular cluster in Aquarius
73	6994	20h 58.9m	−12°38′	9.1	Group of 4 stars in Aquarius
74	628	1h 36.7m	+15°47′	9.2	Spiral galaxy (Sc) in Pisces
75	6864	20h 06.1m	−21°55′	8.6	Globular cluster in Sagittarius
76	650,651	1h 42.4m	+51°34′	11.5	"Little dumbbell" planetary nebula in Perseus
77	1068	2h 42.7m	−0°01′	8.8	Spiral galaxy (Sb) in Cetus
78	2068	5h 46.7m	0°03′	8	Small emission nebula in Orion
79	1904	5h 24.5m	−24°33′	8.0	Globular cluster in Lepus
80	6093	16h 17.0m	−22°59′	7.2	Globular cluster in Scorpius
81	3031	9h 55.6m	+69°04′	6.8	Spiral galaxy (Sb) in Ursa Major
82	3034	9h 55.8m	+69°41′	8.4	Irregular galaxy in Ursa Major
83	5236	13h 37.0m	−29°52′	7.6	Spiral galaxy (Sc) in Hydra
84	4374	12h 25.1m	+12°53′	9.3	Elliptical galaxy in Virgo
85	4382	12h 25.4m	+18°11′	9.2	Elliptical/lenticular galaxy in Coma Berenices
86	4406	12h 26.2m	+12°57′	9.2	Elliptical galaxy in Virgo
87	4486	12h 30.8m	+12°24′	8.6	Elliptical galaxy in Virgo
88	4501	12h 32.0m	+14°25′	9.5	Spiral galaxy (Sb) in Coma Berenices
89	4552	12h 35.7m	+12°33′	9.8	Elliptical galaxy in Virgo
90	4569	12h 36.8m	+13°10′	9.5	Spiral galaxy (Sb) in Virgo
91	4548	12h 35.4m	+14°30′	10.2	Barred spiral galaxy (SBb) in Coma Berenices
92	6341	17h 17.1m	+43°08′	6.5	Globular cluster in Hercules
93	2447	7h 44.6m	−23°52′	6.2	Open cluster in Puppis
94	4736	12h 50.9m	+41°07′	8.1	Spiral galaxy (Sb) in Canes Venatici
95	3351	10h 44.0m	+11°42′	9.7	Barred spiral galaxy (SBb) in Leo
96	3368	10h 46.8m	+11°49′	9.2	Spiral galaxy (Sa) in Leo
97	3587	11h 14.8m	+55°01′	11.2	Owl Nebula; planetary nebula in Ursa Major
98	4192	12h 13.8m	+14°54′	10.1	Spiral galaxy (Sb) in Coma Berenices
99	4254	12h 18.8m	+14°25′	9.8	Spiral galaxy (Sc) in Coma Berenices
100	4321	12h 22.9m	+15°49′	9.4	Spiral galaxy (Sc) in Coma Berenices
101	5457	14h 03.2m	+54°21′	7.7	Spiral galaxy (Sc) in Ursa Major
102					Probably duplication of M101
103	581	1h 33.2m	+60°42′	7.4	Open cluster in Cassiopeia
104‡	4594	12h 40.0m	−11°37′	8.3	Sombrero Galaxy: Spiral galaxy (Sa) in Virgo
105‡	3379	10h 47.8m	+12°35′	9.3	Elliptical galaxy in Leo
106‡	4258	12h 19.0m	+47°18′	8.3	Spiral galaxy (Sb) in Canes Venatici
107‡	6171	16h 32.5m	−13°03′	8.1	Globular cluster in Ophiuchus
108‡	3556	11h 11.5m	+55°40′	10.0	Spiral galaxy (Sb) in Ursa Major
109‡	3992	11h 57.6m	+53°23′	9.8	Barred spiral galaxy (SBc) in Ursa Major
110‡	205	0h 40.4m	+41°41′	8.5	Elliptical galaxy; companion to M31

*From G. O. Abell, *Drama of the Universe,* Holt, Rinehart and Winston, 1978; J. M. Pasachoff, *A Field Guide to the Stars and Planets,* 4th ed., Houghton Miffiin, 2000; and A. Hirshfeld and R. W. Sinnott, *Sky Catalogue 2000.0*, vol. 2, Sky Publishing, 1985.
†In 2006 K. Graun proposed the addition of NGC869 and NGC884 as "honorary" M111 and M112.
‡Not in Messier's original (1781) list; added later by others.

Other catalogs. As Messier was completing his small catalog, William Herschel was building a far larger one, with thousands of objects seen through his bigger telescopes. That catalog was transformed by John Dreyer into the still more massive, 7000-object *New General Catalogue*, which today is the standard list of distant objects in the sky. *See* ASTRONOMICAL CATALOG.

In more recent years, other scientists have compiled newer catalogs. The *Caldwell Catalog*, which lists 110 other deep-sky objects in both the northern and southern hemisphere sky, was prepared by the astronomer Patrick Moore not from comet hunting but from choosing, in his opinion, the best non-Messier astronomical objects in the sky. A second example is David Levy's catalog of deep-sky objects found during his comet search. Levy built his catalog in the tradition of Messier, from searching for comets for more than 40 years. Some of the objects are truly, as comet hunter Leslie Peltier wrote, comet masqueraders and could easily be mistaken for comets. Others represent interesting objects that Levy has encountered during his search.

Although there are other catalogs, Messier's effort remains the most famous. It contains beautiful objects that are spread all over the sky as seen from

Messier's home, and consequently can be seen from any location in the Northern Hemisphere and from most in the Southern. David H. Levy

Bibliography. K. Graun, *The Next Step: Finding and Viewing Messier's Objects*, Ken Press, 2005; K. G. Jones, *Messier's Nebulae and Star Clusters*, 2d ed., Cambridge University Press, 1991; D. H. Levy, *David Levy's Guide to Observing and Discovering Comets*, Cambridge University Press, 2003; D. H. Levy, *Deep Sky Objects*, Prometheus, 2005; D. Machholz, *The Observing Guide to the Messier Marathon: A Handbook and Atlas*, Cambridge University Press, 2002; J. H. Mallas and E. Kreimer, *The Messier Album*, 1978, reprint, Sky Publishing, 1994; S. J. O'Meara, *Deep-Sky Companions: The Caldwell Objects*, Sky Publishing, 2002; S. J. O'Meara, *Deep-Sky Companions: The Messier Objects*, Sky Publishing, 1998.

Metabolic disorders

Disorders of metabolism principally involve an imbalance in nucleic acids, proteins, lipids, or carbohydrates. They are usually associated with either a deficiency or excess resulting in an imbalance in a particular metabolic pathway. All metabolic disorders have a genetic background, and some of them are expressed as specific genetic diseases. Metabolic processes throughout life are strongly influenced by the genetic component of aging, by diet, and by various types of environmental injury, including infectious agents (bacteria, viruses, fungi, and parasites), physical agents such as radiation and temperature, and chemical agents in the environment.

Other factors affecting metabolism include internal control mechanisms that are superimposed on the genetic background. One of the most important mechanisms is the hormonal control system, which consists of the endocrine, paracrine, and autocrine systems. Endocrine factors released from the hypothalamus stimulate the pituitary to produce somatotropic, corticotropic, gonadotropic, and thyrotropic hormones, which affect secondary endocrine organs such as the thyroid, pancreas, testis, breast, ovary, and adrenal glands. These secondary endocrine organs then secrete another group of systemic hormones (thyroxine, parathyroid hormone, insulin, estrogen, progesterone, testosterone, and epinephrine) that affect target cells. The paracrine system involves hormones or controlling substances secreted by cells affecting adjacent cells. The paracrine system includes neural transmitters affecting transmission between nerve cells; lymphokines secreted by lymphocytes, which affect other blood cells; as well as products of endocrine cells in the intestine, lungs, and pancreas. As an example, glucagon in alpha cells in the pancreatic islets of Langerhans affects the beta cells producing insulin. A third hormonal system is the autocrine system, which involves growth substances secreted by cells that act on themselves, including epithelial growth factor, nerve growth factor, and fibroblast growth factor. *See* ENDOCRINE MECHANISMS; GROWTH FACTOR.

The second control system that has a significant effect on metabolism is the neural control system. It consists of the central nervous system, including the cranial nerves, and the peripheral nervous system, including both sympathetic and parasympathetic nerves arising from the spinal cord. These nervous systems act by stimulating or inhibiting particular activities of organ systems and have a profound effect on the metabolism of tissues and organs. *See* NERVOUS SYSTEM (VERTEBRATE).

The third control system is the immune control system, which relates to both the endocrine and neural systems. The immune response produces antibodies as a defensive reaction against foreign antigens found on membranes of infectious agents, grafts, or tumors. Antigens may be proteins, lipoproteins, glycoproteins, or nucleoproteins; antibodies are glycoproteins. Different cells in the immune system have different functions. The macrophage is responsible for processing the antigen; the B lymphocyte is responsible for production of humoral antibodies; and a T lymphocyte has three functions: (1) a helper function to the B cell in antibody production, (2) a killer cytotoxic cell function in graft and tumor rejection and in some delayed hypersensitivity diseases such as tuberculosis, and (3) a synthetic function, forming lymphokines or paracrine hormones that may either stimulate or inhibit growth and metabolic processes. *See* IMMUNOLOGY.

Genetic background, environmental factors, and the three major control mechanisms, in conjunction with age and sex, bring about profound changes in metabolism, which ultimately result in structural and functional alterations. These may involve individual cells (such as the production of antibody), tissues (such as the inflammatory process associated with infection), or organs (such as the effect of dietary components in obstructing heart vessels and causing organ failure).

Nucleic acids. Abnormalities of nucleic acid metabolism are associated with several diseases, including gout and lupus erythematosus. Gout is a form of arthritis associated with extreme pain, especially around the smaller joints. There are usually accumulations of uric acid in the blood (hyperuricemia) and deposits of uric acid (tophi) in the soft tissue, the cartilage of the ear, the shafts of long bones, the kidneys, and the heart valves. Tophi are collections of sodium urate crystals surrounded by inflammation. Hyperuricemia may result from increased production or decreased excretion of uric acid or a diminished conversion of uric acid to urea. Increased production of uric acid, a degradation product of purines, has been found in some cases. Lupus erythematosus may be considered an abnormality of either protein or nucleic acid, since it involves a disturbance in nucleoprotein metabolism. The etiology is unknown. There is widespread damage to collagen, especially in the heart vessels, kidneys, and skin, and to the pericardial and pleural surfaces. *See* ARTHRITIS; CONNECTIVE TISSUE; DISEASE; GOUT.

All genetic disease implies a defect in nucleic acids, and although some genetic diseases are classified as protein abnormalities, such as amino aciduria or hemoglobinopathies, there is always an inherent defect in the nucleotide. This is either a deficiency, an excess, or a mutation that results in an abnormal protein being formed. Similarly, although lipid storage disease results in an abnormal metabolism of lipids, it is the result of a deficiency of a particular enzyme, which also means a defect in the particular nucleic acid code. In some carbohydrate genetic storage diseases, there is an accumulation of glycogen resulting from deficient metabolism of carbohydrates associated with a deficiency of a particular enzyme, which again results from a nucleotide defect. Certain congenital defects that result in malformations of organ systems are the result of either germ cell or somatic cell deficiencies involving differentiation genes composed of nucleic acid.

Genetic engineering has made possible the replacement of a deficient gene in cells of tissue cultures and in some experimental animals. The appropriate gene is isolated and cloned from a normal healthy cell and then introduced into cells removed from the bone marrow of a deficient animal or human. The gene may be inserted into the cell by direct inoculation with a micropipette or the cell may be incubated in a solution containing the appropriate constituents. *See* GENETIC ENGINEERING.

A genetic disease that has generated the most interest is an immune deficiency called severe combined immunodeficiency syndrome (SCIDS). This results in a deficiency of B lymphocytes, which produce antibody, and T lymphocytes, which are responsible for graft and tumor rejection as a result of their cytotoxic effect. This abnormality is associated with a deficiency of the enzyme adenosine deaminase. Chimpanzees that have this deficiency, when corrected with proper gene insertion into the stem cell, have been able to produce the enzyme for 127 days. The reason for the lack of success in other genetic diseases has been attributed to the complicated control mechanisms operating in the body and the fact that some defective cells, such as brain cells, cannot be isolated. In addition, many of the enzymes have several chains and have complicated control mechanisms. *See* HUMAN GENETICS; IMMUNOLOGICAL DEFICIENCY; NUCLEIC ACID.

Proteins. Proteins are in a continual state of dynamic equilibrium, and each particular protein has a specific rate of turnover. The degradation of proteins may occur within the cell itself in lysosomal autophagic vacuoles, which are responsible for turnover of normal cellular organelles and constituents, or extracellularly by the release of extracellular proteases that can degrade extracellular proteins such as collagen, elastin, and membranes of dead cells. The initial stage in degradation of a protein is called denaturation, which implies unfolding of the tertiary structure and usually results in the loss of function.

It has been noted that 90% of the solid constituents of cells and of the body are proteins. They are associated with structural components of the cell, extracellular structures such as collagen and elastase, and some hormones such as insulin and parathyroid hormone. They also constitute the principal antigens of the body either as simple proteins or in conjunction with lipids (lipoprotein), nucleic acid (nucleoprotein), or carbohydrate (glycoprotein). They constitute 99% of all the enzymes that act as catalysts responsible for controlling metabolism, and it is known that ribonucleic acid (RNA) may also act as an enzyme, and that the muscle protein myosin is both a structural protein and an enzymatic protein. Proteins are responsible for many of the transport mechanisms in the cytoskeleton within the cell, and comprise the transport proteins in vessels, such as the lipoproteins, albumin, and the gamma globulins.

Abnormal protein diseases. The diseases associated with protein abnormalities include those associated with increased production of proteins, decreased production of proteins, production of abnormal proteins, and excretion of unusual amounts of amino acids. Most diseases cause changes in protein metabolism and, to a greater or lesser extent, result in cellular changes such as cloudy swelling and hydropic degeneration. When a cell is injured slightly, the normal reserve components of the cell, plus the ability to resynthesize necessary constituents, usually result in a quick return to normal cell function and an inability to recognize any changes morphologically. Damage by a few cells in an organ as large as the kidney, which contains millions of cells, is of little importance in relation to total kidney function and presumably is a common occurrence.

Protein synthesis. Often called macroglobulinemia, hyperproteinemia results in an increase in beta or gamma globulins, but possibly with less total protein. Hyperglobulinemia diseases include multiple myeloma, kala-azar, Hodgkin's disease, lymphogranuloma inguinale, sarcoidosis, cirrhosis, and amyloid disease. They usually involve stem cell lines of the bone marrow macrophages or B cells. *See* CIRRHOSIS; HODGKIN'S DISEASE.

A decrease in the amount of protein (hypoproteinemia) can result from a lack of amino acids for protein synthesis, a metabolic block, or other interference with normal protein synthesis. Increased excretion of protein, particularly in chronic renal disease with a loss of albumin in the urine (albuminuria), is another common cause of hypoproteinemia. Kwashiorkor is the best example of hypoproteinemia resulting from dietary deficiency. Since albumin is synthesized in the liver, severe liver inflammation (hepatitis) or scarring (cirrhosis) will result in decreased synthesis and hypoalbuminemia with decreased levels of albumin in the blood. In hypogammaglobulinemia and agammaglobulinemia, which may also be classified under hypoproteinemia, the total serum albumin and globulin are not markedly depressed, but the gamma globulins may fall from a normal of 15–20% of total protein to 0.4%. Antibodies are normally part of this fraction, and the individuals are unable to synthesize antibodies against antigens. *See* ANTIBODY; IMMUNOGLOBULIN.

Abnormal proteins. Many diseases are characterized by abnormal proteins, including multiple myeloma, the hemoglobinopathies, and the various amyloid disturbances. Multiple myeloma, a neoplastic growth of plasma cells, particularly in the bone marrow and lymph nodes, is representative of a disease in which an abnormal protein is produced. A particular light chain of the normal gamma globulin, called Bence-Jones protein, is found in the urine of persons with multiple myeloma and differs markedly from any protein found normally in the serum or urine. This protein has a molecular weight of approximately one-fourth that of normal gamma globulin and is not merely a degradation product of the plasma protein but is a new or intermediate protein.

Another large group of abnormal proteins includes the hemoglobinopathies. Hemoglobin functions in the transport of oxygen in the blood. Several different hemoglobins in red blood cells in mammals have been described. Fetal hemoglobin has two alpha chains and two fetal hemoglobin secondary chains. Adult hemoglobin has two alpha and two beta chains. In sickle-cell hemoglobin there is a single base change in the beta chain of one nucleotide which results in a massive alteration in tertiary folding of the hemoglobin molecule and ultimately in a complete distortion of the normal blood cell membrane under low oxygen tension. This sickle cell is unable to carry normal amounts of oxygen and results in hemorrhage, inflammation, and occasionally sickle-cell crisis. Scores of different hemoglobins have been identified, but sickle cell in the United States and thalassemia (Mediterranean anemia) in Europe are two of the most common hemoglobin abnormalities. *See* HEMOGLOBIN; SICKLE CELL DISEASE.

A third type of abnormal protein is found in amyloid disease, of which there are many variations. In multiple myeloma, the light chains produced by the plasma cells are seen as amyloid. In diabetes mellitus, proinsulin in the islet cells of the islets of Langerhans of the pancreas appears as amyloid. In medullary carcinoma of the thyroid, calcitonin, produced by the C cells, appears as an amyloid, with proper staining. Another type of amyloid has been found in the brain and other tissues in Alzheimer's disease. Typically, amyloid has been associated with chronic disease states such as tuberculosis and rheumatoid arthritis, but this type is apparently a transformation of an abnormal globulin protein in the bloodstream. It results in deposition, particularly in blood vessels and in joint spaces. *See* ALZHEIMER'S DISEASE; AMYLOIDOSIS.

Amino acid excretion. A fourth group with abnormal protein metabolism is associated with a change in particular amino acids resulting either from an overflow mechanism, where the concentration of amino acids in the serum surpasses the renal threshold of the glomerular membrane, or from defective absorption of amino acids in renal tubules. *See* KIDNEY DISORDERS.

Of the 20 amino acids important for protein synthesis, many are also important in particular disease states. Tyrosine appears to be one of the most critical amino acids, and its metabolism is related to four key diseases including phenylketonuria, hypothyroidism, albinism, and alkaptonuria. Phenylketonuria, a disease that results in mental deficiency, is caused by an absence of the oxidase enzyme that converts phenylalanine to tyrosine. The accumulation of phenylalanine eventually results in accentuation of alternate pathways of phenylalanine degradation leading to phenylpyruvic acid, phenylacetic acid, and phenyl lactate acid. The latter products are present in such excessive quantities in the serum that they are excreted in the urine. Hypothyroidism is caused by a lack of thyroxine, which markedly influences metabolism. A deficiency of tyrosine results in a lack of thyroglobulin and eventually a lack of thyroxine. Infantile hypothyroidism is known as cretinism, and adult hypothyroidism is also known as myxedema. Tyrosine is also involved in the formation of dopa, which is an intermediate in the formation of melanin. A deficiency of tyrosine may express itself as a lack of melanin pigment of the skin and results in the congenital disease of albinism. Deficiencies in dopa, which is a precursor of epinephrine, may result in other endocrine disturbances. A failure of a complete breakdown of phenylalanine and tyrosine (alkaptonuria) results in the synthesis of an intermediate product, homogentisic acid, which accumulates in tissue and serum and is excreted in the urine. It is likely that the enzyme homogentisase, which converts homogentisic acid to fumaryl acetoacetic acid, is inhibited. Deposits of homogentisic acid in the tissues, especially in cartilage, tendons, ligaments, and sclerae of the eyes, impart a bluish discoloration known as ochronosis. *See* PHENYLKETONURIA.

The liver plays a major role in the deamination of amino acids. Advanced hepatitis and cirrhosis may lead to increased levels of amino acids in the blood and excretion in the urine. Other diseases with amino aciduria that are believed to be the result of defective kidney function include cystinuria (the failure to reabsorb cystine, lysine, arginine, and ornithine), Wilson's disease (a degeneration involving copper metabolism in the liver and brain), Fanconi's syndrome, galactosemia, scurvy, rickets, and lead, cresol, or benzene poisoning. *See* LIVER DISORDERS.

Sickle-cell disease and thalassemia would appear to be prime candidates for correction through genetic engineering. Replacement of the defective gene seems feasible, particularly since in both of these diseases a deficiency of a single base with a single amino acid change is the cause. The sole attempt, in a case of thalassemia, was not successful. One possible reason is that the hemoglobin molecule consists of four chains, and its synthesis, assembly, and release must be perfectly coordinated. A better candidate appears to be a single-chain protein such as the adenosine deaminase associated with severe combined immunodeficiency syndrome. *See* PROTEIN METABOLISM; SICKLE CELL DISEASE.

Lipids. Fatty acids may be synthesized from acetyl coenzyme A, which is an intermediate product and

a central site for amino acid, lipid, and carbohydrate metabolism. Lipids may be degraded by lipases in the serum, blood vessels, or tissues. Lipids are generally transported with proteins as lipoproteins or in a structure called a chylomicron. Lipids may be incorporated into the cell by pinocytosis after transport in the serum.

Fat has both a structural and a metabolic function. In the cell, fats are degraded, oxidized for energy, or used for synthesis of other lipids, carbohydrates, or proteins. Most cell membranes are composed of lipoprotein complexes. As an energy source, lipids at 9 calories per gram, compared with carbohydrates at 4 calories per gram, remain the second important source of adenosine triphosphate (ATP). Cholesterol is an important structural component of membranes. *See* ADENOSINE TRIPHOSPHATE (ATP); ADIPOSE TISSUE; CELL MEMBRANES; CHOLESTEROL.

Although lipid stores remain a secondary energy reserve in starvation, the breakdown of lipids associated with diabetes and starvation results in the production of ketone bodies in the urine with general acidosis (ketosis) in the serum. Of concern to many people is an increase in lipids associated with obesity. Although diet remains the primary control of obesity, some metabolic disturbances associated with endocrine imbalances make the problem of obesity difficult or impossible to control through diet. *See* OBESITY; PITUITARY GLAND DISORDERS; THYROID GLAND DISORDERS.

Hyperlipemia, an excess of lipid in the blood, is often secondary to uncontrollable diabetes, hypothyroidism, biliary cirrhosis, and lipoid nephrosis. Hyperlipemia may result in transfer of lipid into cells, particularly of the liver (as in cirrhosis of the liver), the kidney (as in lipoid nephrosis, now called minimal disease or nil disease), the heart (where it can interfere with contraction of the muscle), and vessels (in arteriosclerosis). Most cells are capable of synthesizing lipid, and some cells apparently produce excessive amounts under abnormal conditions which may be released to the extracellular fluid and thence to the bloodstream. Elevated blood lipids and cholesterol may be relatively independent of diet. Excess lipids present when dietary lipids are severely restricted must be produced by cells, although not necessarily by the cells in which they ultimately appear. If cells do not manufacture and store lipid within themselves, they may obtain the lipid from extracellular sources, presumably the lipoproteins of the serum. Deposits of lipid with fibrosis may occur on the eyelids in older people with high cholesterol levels. Excess proliferation of fat cells is known as a lipoma, which occasionally becomes malignant, producing a liposarcoma. *See* ARTERIOSCLEROSIS.

In a large group of genetic lipid storage diseases, lipid accumulates because of a disturbance in lipid metabolism that is independent of external stimuli. These genetic diseases are inherent nucleic acid defects that result in abnormal enzymatic proteins, which then result in abnormal lipid metabolism. In these diseases, a large accumulation of lipid appears in many cells, but particularly the reticuloendothelial cells of the lymph nodes, liver, spleen, and bone marrow. The lipid seems to be distributed throughout the cell and is not well localized. Abnormal lipid storage occurs in Niemann-Pick, Gaucher's, and Tay-Sachs diseases. Niemann-Pick disease is characterized by the excess accumulation of sphingomyelin, a cell membrane constituent, in the nervous system. Gaucher's disease encompasses a group of inborn errors of glycosphingolipid metabolism associated with an enzyme deficiency, and Tay-Sachs disease also results from an enzyme deficiency. All three diseases are hereditary and are fatal.

In cirrhosis of the liver, scarring is secondary to destruction of liver cells and nutritional deficiencies commonly associated with alcoholism. Lipid appears as small or large droplets in close physical association with plasma lipids. In association with protein deficiencies, normal substrates that are usually involved in protein metabolism may be diverted into lipid synthesis. Lipid accumulation in alcoholic hepatitis and perhaps in cirrhosis may be reversible.

The genetic disease familial hypercholesteremia appears in 1 of every 500 people in the heterozygous population and in 1 of every 1 million in the homozygous population. Heterozygous hypercholesteremia, with cholesterol levels two to four times normal, is characterized by arteriosclerotic lesions of the coronary vessels that cause myocardial infarct and death in the 35–50-year age range. In the homozygous condition, with cholesterol levels eight to ten times normal, death occurs usually before the age of 15. The defect resides in the receptor for low-density lipoprotein (LDL). This receptor defect in the liver and other tissues results in a lack of transport of low-density lipoprotein containing high cholesterol amounts from the bloodstream. Similar defects have been shown with genetic abnormalities in the apoproteins, which bind to specific receptors or transport proteins on the cell membranes and direct the lipoproteins to the metabolic site. Such defects result in a lack of low-density lipoprotein uptake. The receptor may not bind to the low-density lipoprotein, there may not be sufficient number of receptors present, or the internalization process of the low-density lipoprotein once attached to the receptor may be defective.

Familial hypercholesteremia serves as a model for hypercholesteremia associated with nonheritable causes. A reduction in smoking, an increase in exercise, close attention to blood glucose levels in diabetes, a diet low in fat, and careful monitoring of cholesterol levels, are important. When preventive measures fail, any of a large group of drugs can divert the normal antecedents of cholesterol, block the synthesis of cholesterol, or immobilize the cholesterol before entrance into the lipoprotein. The degree to which efforts to reduce cholesterol have played a role remains controversial. *See* CHOLESTEROL; HEART DISORDERS; LIPID METABOLISM.

Carbohydrates. Carbohydrate deficiency is seen in starvation. When insufficient energy is derived from carbohydrate ingestion, the body turns to its fatty acid, breaking down lipid cells and generating

adenosine triphosphate as well as excess ketones. Eventually the body may degrade protein, and amino acids may be diverted into the acetyl coenzyme A cycle with the production of adenosine triphosphate. Excess carbohydrate is stored in the liver and muscle as glycogen, available for use when needed. *See* MALNUTRITION.

Abnormal carbohydrate diseases include the genetic diseases that represent a deficiency in nucleotide and eventually protein enzymatic activity. Of the six common carbohydrate storage diseases, two examples are von Gierke's disease, marked by glycogen storage in the heart, and Pompes' disease, in which the carbohydrate is stored in the liver. Variants of carbohydrate disease involve storage of mucopolysaccharides, as in Hurler's disease, and storage of galactose, as in galactosemia.

The most important disease associated with carbohydrate metabolism is diabetes mellitus, which affects approximately 10% of the United States population. However, only 1% is type I or juvenile diabetes; the other 9% is identified as type II or maturity-onset diabetes. Type I diabetes is associated with a specific human lymphocyte antigen (HLA) profile on the cell. Following injury to the beta cells of the islets of Langerhans, which produce insulin, the altered membranes of the cells act as an antigen to produce autoantibodies. This antibody results in the destruction of other viable beta cells and is a classic example of autoimmune disease. This results in an absolute deficiency of insulin. Type II diabetes results not from an absolute deficiency of insulin but from ineffectiveness of the insulin on cells. There appears to be no genetic HLA susceptibility. It has been postulated, with some evidence, that there is a defect in the insulin receptors on the target cells, either a decrease in number or an inability to bind the appropriate amounts of insulin. Because the autoimmune component is small, dietary control of the disease is possible. If allowed to remain untreated, hyperglycocemia may result in damage to blood vessels, kidneys, and retinas and increased infection. *See* AUTOIMMUNITY; CARBOHYDRATE METABOLISM; DIABETES; METABOLISM.

Donald W. King

Bibliography. G. S. Eisenbarth, Type I diabetes: Clinical implications of autoimmunity, *Hosp. Prac.*, 22:135–152, 1987; F. H. Epstein (ed.), Specificity spillover at the hormone receptor: Exploring its role in human disease, *N. Engl. J. Med.*, 320:640–645, 1989; C. M. Feek and C. R. Edwards, *Endocrine and Metabolic Disease*, 1988; *Harrison's Principles of Internal Medicine,* McGraw-Hill, 16th ed., 2005; V. A. McKusick, The new genetics and clinical medicine: A summing up, *Hosp. Prac.*, pp. 79–84, July 1988; P. N. Plowman, *Endocrinology and Metabolic Diseases*, 1987.

Metabolism

All the physical and chemical processes by which living, organized substance is produced and maintained and the transformations by which energy is made available for use by an organism.

In defining metabolism, it is customary to distinguish between energy metabolism and intermediary metabolism, although the two are, in fact, inseparable. Energy metabolism is primarily concerned with overall heat production in an organism, while intermediary metabolism deals with chemical reactions within cells and tissues. In general, the term metabolism is interpreted to mean intermediary metabolism. *See* ENERGY METABOLISM.

Metabolism thus includes all biochemical processes within cells and tissues which are concerned with their building up, breaking down, and functioning. The synthesis and maintenance of tissue structure generally involves the union of smaller into larger molecules. This part of metabolism, the building of tissues, is termed anabolism. The process of breaking down tissue, of splitting larger protoplasmic molecules into smaller ones, is termed catabolism. Growth or weight gain occurs when anabolism exceeds catabolism. On the other hand, weight loss results if catabolism proceeds more rapidly than anabolism, as in periods of starvation, serious injury, or disease. When the two processes are balanced, tissue mass remains the same.

Energy production and utilization. Various components of the tissues are continually undergoing degradation in catabolism and resynthesis in anabolism. Some chemical reactions involved in these metabolic processes are exergonic; that is, they are accompanied by liberation of energy. Others are endergonic; that is, they require exogenous energy. After energy has been introduced into an endergonic process to move substances from a metastable state or state of intermediate stability, an exergonic reaction may occur. Two or more reactions may be coupled in such a way that an exergonic reaction may be used to drive an endergonic one. Within cells, for example, there exists a mechanism whereby free energy, available usually from oxidative reactions, may be used to spark endergonic processes. In some cases, a gain in free energy occurs. Adenosine triphosphate (ATP), which contains high-energy phosphate bonds (~P), is used in many such instances. In some oxidative reactions, however, it seems that the released energy cannot be used for endergonic purposes. The aerobic dehydrogenases of cytoplasm act as catalysts in the oxidation of a number of substrates, but apparently they cannot be utilized in endergonic reactions. *See* ADENOSINE TRIPHOSPHATE (ATP); BIOLOGICAL OXIDATION.

Metabolism is largely concerned, particularly in the case of synthetic or anabolic phenomena, with the need for providing energy-producing mechanisms for the construction of new tissues and their constituents. The energy is often obtained from simultaneously occurring degradative, or catabolic, reactions. Exergonic transformations are utilized for synthesis of all structures, for maintenance of body temperature, and for providing energy for specific functioning of various tissues, as in muscle contraction, nerve impulse conduction, glandular secretion, absorption, light production, and execretory

processes. To replace the energy thus utilized, a regular provision of new external sources of energy in the form of various foodstuffs must be supplied.

Metabolic pathways. The metabolism of the three major foodstuffs, carbohydrate, fats, and proteins, is intimately interrelated, so any clear-cut division of the three is arbitrary and inaccurate. Thus the metabolism of protoplasm is concerned with all three of these foodstuffs. However, each function of an organism has its own particular metabolic needs. For example, the chemical reactions of major importance in muscle contraction are those which yield energy for transmission to and operation of the contractile elements in muscles. The high-energy phosphate bonds (~P), formed during metabolic degradation and transformation of carbohydrate, fatty acids, and amino acids, transfer their energy to the muscle fibrils to provide the energy for contraction. *See* MUSCLE.

The metabolic pathways of carbohydrates, fats, and proteins cross at many points; thus certain pathways of metabolism are shared in common by fragments of these different classes of foodstuffs. Carbohydrate, protein, and fat may all contribute to the metabolic pool of acetyl coenzyme A (active acetate) which can be oxidized to CO_2 and H_2O (tricarboxylic acid cycle) with release of energy or used for the synthesis of body substances such as hormones. To make the endergonic metabolic processes possible, chemical energy in the form of ~P is produced; CO_2 and H_2O are by-products of these reactions.

Some of the metabolic processes of the protoplasm of both plant and animal cells occur along common pathways; carbohydrate metabolism in plants is similar in many details to carbohydrate metabolism in animals. Thus the study of metabolism in any organism is, in a sense, the study of metabolism in all protoplasm. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; CARBOHYDRATE METABOLISM; CITRIC ACID CYCLE; CYTOCHROME; LIPID METABOLISM; PLANT METABOLISM; PROTEIN METABOLISM.

Mary B. McCann

Metadata

Data that describe or define all objects that make up a data source. That is, metadata is data about the data.

Widely used within the context of database management systems (DBMS), the term metadata refers to the complete description of all database structures and their constraints (conditions that need to be satisfied by the data stored in the database). This metadata is stored in the data dictionary or catalog. The data dictionary is an integral part of any DBMS environment and can be considered a database in its own right (albeit a system database instead of a user database). For this reason, a DBMS is sometimes referred to as a self-describing collection of integrated data since it holds both the data and the data description. A comprehensive data dictionary contains metadata about all database objects, including connections, users, users' privileges, schemas (external, conceptual, and so on), mappings between the different schemes, and all types of security and integrity constraints.

In general, the DBMS software during all its operations actively uses the data dictionary. For example, even in the most basic configurations, whenever a user needs to interact with a DBMS he or she needs to log on to the database server—the computer where the DBMS and the data reside—by specifying a username and a password. During this logon process, the DBMS consults the data dictionary to validate the user. That is, the DBMS verifies that the username and password specified by the user already exist as such in the data dictionary. If the user then proceeds to query, for instance, a table of a relational database, the data dictionary is further consulted to confirm, first, that the table exists and, second, that the user has the privileges (rights) to operate on that table. If the data requested by the user need to be retrieved from physical files, this metadata is also contained in the data dictionary.

To speed up the system responses in some DBMSs, the metadata of the data dictionary is examined in a slightly different manner. In these DBMSs, all objects that have been recently referenced by the database users are temporarily saved, along with their corresponding results if applicable, in designated areas of main memory sometimes called the data dictionary cache and the database buffer cache, respectively. This way, whenever a user makes a query, the DBMS checks first the contents of the data dictionary cache to see if an identical query has already been processed and its results are still in the data buffer cache. If an identical query is present in the data dictionary cache, the DBMS, unless some special data formatting is needed, immediately returns the results to the user. Otherwise, the data are retrieved directly from the physical data files that store the user data. *See* DATABASE MANAGEMENT SYSTEM.

Ramon A. Mata-Toledo

Bibliography. T. M. Connolly and C. E. Begg, *Database Systems: A Practical Approach to Design, Implementation and Management*, 4th ed., Pearson Education, 2004; C. J. Date, *An Introduction to Database System*, 8th ed., Addison-Wesley, 2003.

Metal

A solid or liquid (molten), opaque material with a lustrous surface and good electrical and thermal conductivities. Solid metal is usually crystalline and ductile and can be permanently deformed by shear on crystal planes; permanent deformation is accompanied by an increase in strength (work hardening). Metallic properties are related to the arrangements of positively charged ions bonded through a surrounding field (sea) of free electrons that draw the ions into a close-packed crystalline structure with planes appropriate for slip. Liquids are nearly close packed, noncrystalline, with a thermal energy great enough

to activate random, free movements of atoms. *See* LIQUID.

The rapid development of the metals industry since the late nineteenth century required an in-depth understanding of the fundamentals of metallurgy to improve metal properties and expand their usefulness. Much of the available technical information depends upon empirical correlations, although the rules of thermodynamics, quantum mechanics, and solid-state physics have established a fundamental basis for the science. With quantitative experimental techniques such as x-ray and electron diffraction along with other procedures generated by the rapid advances in semiconductor devices, the synergistic growth in solid-state studies contributed to better comprehension of the structure of metals. *See* ELECTRON DIFFRACTION; SEMICONDUCTOR; THERMODYNAMIC PRINCIPLES; X-RAY DIFFRACTION.

Characteristics of solids. The unique characteristics of the different classes of solids are directly related to the electron pattern and the types of bonds between the atoms of the solid. Atomic bonding is classified in two major categories: strong attraction (covalent, ionic, or metallic) and weaker, secondary bonding (van der Waals or hydrogen bonding). In the equilibrium state, interactions of the outermost orbital electrons and bond restrictions of the aggregate of atoms have the greatest influence on most of the physical and mechanical properties of the solid. Atoms of a solid in a lattice arrangement or in an amorphous distribution vibrate at frequencies about 10^{13} Hz at amplitudes a few percent of the interatomic spacing and at temperatures in the range of one-third of the metal's absolute melting temperature. The higher the temperature, the greater the amplitude of vibration of the atoms. Thermal and elastic pulses are transferred as in a series of coupled vibrating sources exerting forces through to bonds to adjacent atoms. In addition, in metals the free electrons contribute to the transfer of the pulses. *See* CHEMICAL BONDING; ELECTRON; FREE-ELECTRON THEORY OF METALS; HYDROGEN BOND; INTERMOLECULAR FORCES.

In a given pressure range, as the temperature of the element is increased from absolute zero (0 K), stable solid metals (and some nonmetals) tend to change phase into liquid metals at the melting temperature of the specific element or compound. The resulting liquids convert to gases at the vaporization temperature of the specific element. The melting temperature and, to a greater extent, the vaporization temperature are dependent on pressure on the system. The atoms or molecules of gas are too far apart to retain metallic features. Thus, properties of a solid at high and low temperatures are best compared in a temperature range relative to a characteristic temperature of the specific element or compound; in metals some fraction of the melting temperature (in Kelvin) is usually used. In equilibrium conditions, metals in the solid state occur in a limited number of crystal structures; however, depending on the temperature and pressure of the system, some metals have different crystal structures—a phenomenon known as allotropy. Furthermore, at nonequilibrium conditions where the rate for reorganization of atoms into the new equilibrium structure is too slow because of low temperatures—such as the consequence of rapid cooling, heavy deformation on certain metals, or radiation by high-energy particles—the atoms of some metals may be frozen or forced into an intermediate, metastable, higher free-energy state than attained at equilibrium conditions. Examples of these effects occur in the quenching of metals, deformation of austenite into martensite, or high-energy neutron bombardment of metals. An extreme case is the solidifying of the amorphous arrangement of the liquid phase by very rapid cooling of some molten metals. *See* MELTING POINT; METALLIC GLASSES; SOLID-STATE PHYSICS.

Quantum effects. In metals the characteristic crystal structure and physical properties depend mainly on metallic bonding influenced by the tendency of the bonds to incorporate other nonmetallic types of cohesion. Bonding of atoms of an element is determined mainly by the outer electrons of the free atom. For a free atom the electrons with the highest free energy are the valence electrons, and these combine in lowest, nonoccupied, available, quantum energy states of the aggregate of ions.

Because of the small mass of an electron, classical laws of mechanics are not satisfactory to explain the dynamics of the electrons as well as quantum mechanics does. Quantum mechanics represents an electron in a solid as a three-dimensional smear of negative charge, the magnitude of which is equivalent to certain probabilities of being in particular regions of space at some time in the aggregate of ions. The scope and restrictions of quantum-mechanical laws describe satisfactorily the spatial location of the electron charge, and the available energy levels (quantum-energy levels) allowed for electrons. Since stable bonding depends upon the lowest-energy configuration of electrons with respect to the collection of atoms, the modeling by this wave-mechanical procedure along with certain physical limitations (such as the Pauli exclusion principle and the Heisenberg uncertainty principle) require specific available energy levels and, conversely, impose energy ranges forbidden to electrons. Moreover, the locations permitted for the electrons restrict the type of bonds and the possible aggregate geometry (crystalline or liquid at equilibrium or amorphous at nonequilibrium). *See* ELECTRON CONFIGURATION; ENERGY LEVEL (QUANTUM MECHANICS); EXCLUSION PRINCIPLE; QUANTUM MECHANICS; SOLID-STATE CHEMISTRY; UNCERTAINTY PRINCIPLE.

Energy bands. If atoms of the same element are at large enough distances that their electric fields do not influence each other, the discrete electron orbital potential energies in each shell for individual atoms are defined by quantum-mechanical rules for a single atom. At large separation of the atoms, the allowed-electron-potential-energy levels do not intrude on one another, and they are considered free atoms. Consequently, each atom has the same discrete allowed-energy levels as determined by the

laws of quantum mechanics for a single atom. If several atoms are moved near to each other so that the higher energy levels including the valence electron energy tiers interfere with each other, the discrete energy levels are slightly displaced into several other values because only two electrons at any allowable-energy level are possible (Pauli exclusion principle). As before, each level is separated by an energy interval (gap) that cannot contain the electrons for the particular element (forbidden-energy gaps). Since there are many atoms in a metal aggregate—about 10^{23} atoms per cubic centimeter—there are about 10^{23} more energy levels for the aggregate than for the energy levels associated with one free atom. Since in an aggregate the large number of energy levels are relatively close together compared to those in a free atom, the spread of energy levels in the collection of atoms is called energy bands. For some materials, the multitude of energy levels of two or more shells overlap. These conditions are illustrated by energy bands that represent many near discrete-electron-energy levels (**Fig. 1***a*). The potential-energy bands for the equilibrium aggregate of atoms are conveniently shown in an energy-band diagram with the higher energies increasing in the *Y* direction (Fig. 1*b*).

Because the available electrons seek the lowest energy levels, the higher allowed-energy bands do not contain electrons. The average maximum energy of the valence electrons contained in a solid is called the Fermi energy (Fig. 1*b*). At the temperature of absolute zero (no thermal agitation of electrons exist), the Fermi energy is the maximum energy level of the electrons in a nonperturbed solid. In a metal there are always many empty available energy levels minutely separated above the Fermi energy; as a consequence, only small energy gaps between levels act as barriers to the electrons and must be overcome to permit the flow of electrons through the metal. Thus, very small energy increases supplied by an electric field or thermal energy result in a relatively free flow of electrons down a voltage or temperature gradient. The condition is characterized by high electrical conductivity and high heat conductivity by electrons in metals. Furthermore, under conditions of high pressure in which atoms and thus electron energy bands are pressed together (typically in stars or high-pressure experimental devices), nonmetals will exhibit the properties of metals, because the electrons in the Fermi level are able to move freely in energy levels closer to the atom core.

The usual metallic luster—silver, gold, copper, or the many brass metallic colors—are the results of almost total reflection of the long-wavelength portion of the incident light spectrum by the dense plasmalike cloud of conduction electrons throughout the crystal at the appropriate energy band to reflect the proper energy photons. For silver, most of the spectrum, including the visible portion, is reflected except in ultraviolet; whereas with gold the longer wavelengths, including red and yellow, are reflected while the shorter wavelengths of green and below enter the metal. *See* BAND THEORY OF SOLIDS; FERMI SURFACE; PHOTON.

Directional bonds. A covalent bond between two or more atoms is a negative charge directed in space and equivalent in magnitude to one or more electrons. As a consequence, at the lowest energy state the corresponding bond is directional and usually strong, and considerable energy must be supplied to the material to free an electron from this bond or shift an ion from one fixed position to another. For covalent materials, bonding valence electrons require large energies to transport a current composed of many electrons through the aggregate atoms. Thus electrical conductivity and thermal conductivity by electrons are

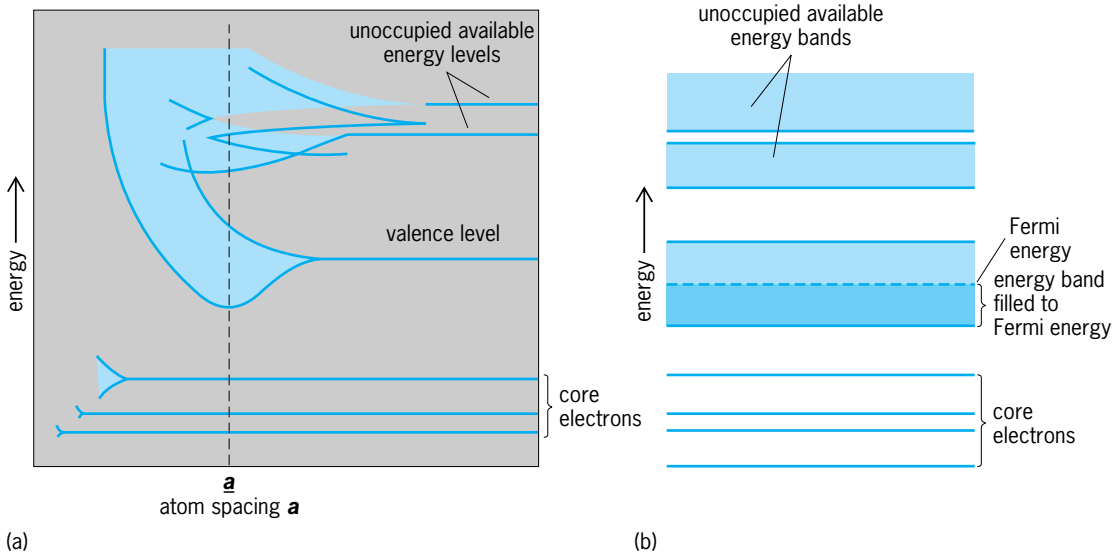

Fig. 1. Available energy levels and bands of energy levels for electrons in a univalent metal. (*a*) Atomic spacing. Allowed and forbidden quantum energies in an aggregate of atoms related to the distance between atoms, *a*; the spaces between horizontal lines signify the energy gaps not allowed for electrons. (*b*) Available energy bands, levels, and gaps for equilibrium spacing, *a*, characteristic of metals with only one electron in the valence level.

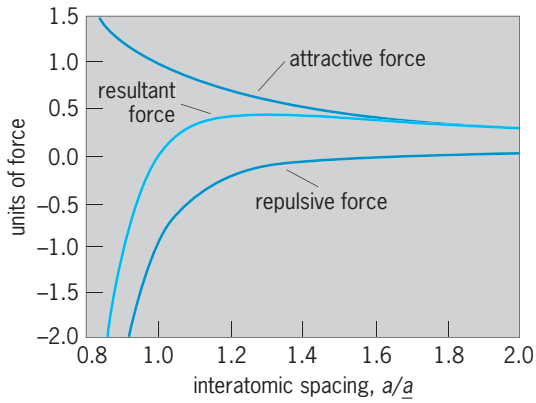

Fig. 2. Ratio of atomic separation, *a*, to equilibrium spacing, *a*, plotted against the sum of the repulsive and the attractive forces. At the ratio of interatomic spacings *a*/*a* = 1.0, equilibrium spacing, the resultant force is zero.

low. This situation is demonstrated in energy-band diagrams by a filled allowable-energy band with a large energy gap over which electrons must transverse to another allowed-energy band to become conduction electrons.

Metallic bonding. In contrast to directional bonds, metallic bonds result from linking ions each having centrosymmetrical attraction to a surrounding atmosphere of the distributed charges of nearly free electrons. These electrons, usually valence electrons, have the same likelihood of being associated with any ion in the aggregate. However, there are quantum-mechanical constraints defining the probability of the charge being at some position with some range of energy for a given time period. Depending upon the spatial distribution of the negative charges and positive ions, characteristic equilibrium distances (lattice spacings) are established between the ions of the metal. This equilibrium condition can be illustrated by plotting the magnitude of the resulting forces, which can be split into two parts: an attractive force that increases as ions are drawn closer in response to the attraction between the negative charge of the electron cloud of one ion to the positive charge of another core; and a shorter-range repulsive force that increases strongly when the ion charges pull the atoms close enough to cause interference between inner electron energy levels. The equilibrium position is at the spacing at which the forces are balanced and the resultant force is zero (**Fig. 2**). Inordinately high pressures on nonmetals, which force atoms to be closer together, distort the existing bonds tending to arrange the electrons into a configuration comparable to metals and convert the behavior of the element to metallic properties. Just as with covalent bond materials, the force necessary to displace the atoms a small distance from the equilibrium positions—elastic displacement—is nearly linear. Thus, for compressive, tensile, or torsion stresses of the aggregates of metal atoms, it is found that elastic properties relating stress to strain for small deformations are nearly linear; and consequently, linear elasticity is convenient for solutions of many problems. Actually, fine measurements of deformation in the elastic region generally indicate nonlinear behavior.

Permanent displacement of atoms from each other as occurs in fracture or plastic deformation depends on the relative maximum stress required to pull apart the bonds to the maximum shear stress to translate one part of the solid from one position to another by transferring the bonds to other positions. The nondirectional bonds in a metal are easily switched from one position to another (ductility), whereas with a covalent solid the bonds are parted and only fracture is possible (brittle). The transition metals are dominated by *d*-level electrons, which among many effects, such as properties of magnetism, impart some covalent features to the bond. Although the maximum cleavage stresses are higher than with the noble metals (gold, silver, and copper), the maximum shear stresses for the transitional metals are low enough to allow the crystal to shear and retain continuity before fracture. *See* BRITTLENESS; PLASTIC DEFORMATION OF METAL.

Crystal structure. To visualize groups of atoms in an aggregate, a qualitative geometric model, the atoms are considered as spheres of the same radius (**Fig. 3**). This representation of metal atoms is reasonable, since bonds originating from metal atom cores with the superimposition of the electron charges in the intervening space are manifested by a nearly

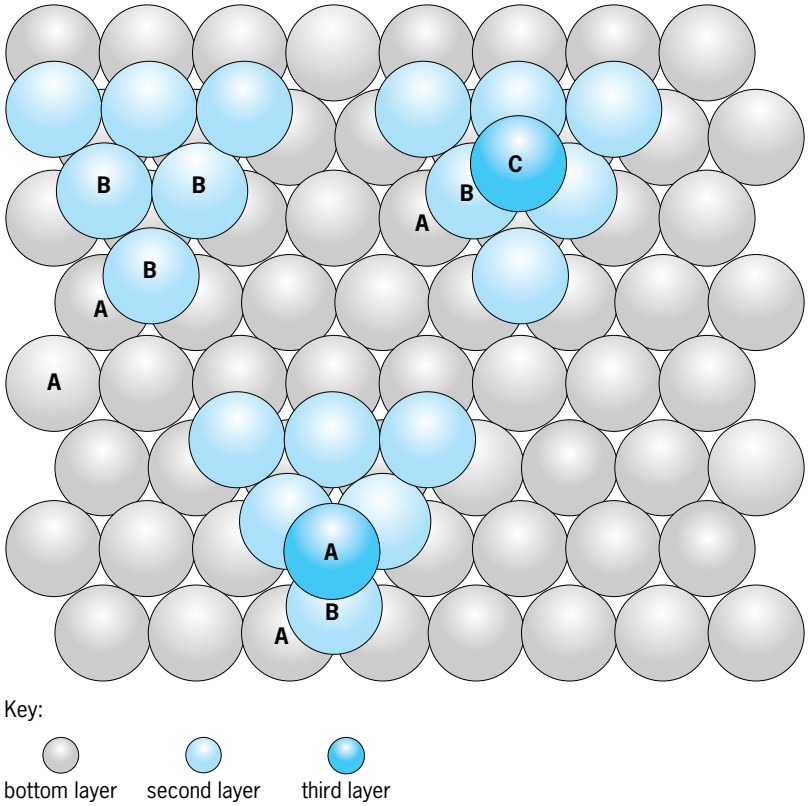

Fig. 3. Sequences of close-packed planes for face-centered-cubic and close-packed-hexagonal. If atoms in the bottom layer are in position A, the second layer is in position B, the third layer can be either in position C or A. If the progression is ABCABC. . . , the structure is face-centered-cubic; if the progression is ABABABAB. . . , the structure is close-packed-hexagonal.

spherically symmetrical attraction between the ions. Instructive models of aggregates of these metal atoms are demonstrated by packing equal-size spheres in the densest arrays per unit volume. The two possible close-packed cases are satisfied by either a face-centered-cubic or an ideal close-packed-hexagonal crystal structure. In both of these arrangements each atom is in contact with 12 other atoms, and in the same ratio of 1.633 between alternate close-packed planes divided by the diameter of the atoms. In close-packed-hexagonal structures, this ratio is called the axial ratio, c/a, where *c* is the distance between close-packed planes and *a* is the diameter of the atom in the close-packed plane. Although the density of packing and the number of nearest neighbors are the same, the face-centered cubic structure contains four close-packed planes and three close-packed directions on each plane, whereas a close-packed hexagonal structure contains only one close-packed plane with three directions on that plane.

Among metals, the face-centered-cubic crystalline structure is found in aluminum, nickel, copper, rhodium, palladium, silver, iridium, platinum, gold, and lead. However, no ideal close-packed-hexagonal metals with the same ratio of 1.633 are observed. The distance between the first nearest-neighbor atoms in both crystal structures is the same, but the distance and position of the second nearest atoms are different in the ideal close-packed-hexagonal form from face-centered-cubic structure. The interaction energy associated with the second nearest neighbors in the ideal close-packed hexagonal must be higher than in face-centered-cubic. This higher energy, because of the position of second nearest neighbors, deters a stable formation of an ideal close-packed-hexagonal crystal structure. However, thin layers of close-packed-hexagonal stacking errors found in face-centered-cubic can extend to the surface of the crystal or must end at a dislocation. These planar surface defects are called stacking faults, and they are more frequent if the energy of the stacking fault is not much larger than the energy of normal stacking of the atoms of the particular element (low stacking-fault energy). Similar close-packed-hexagonal layers occur as coherent twin boundaries parallel to close-packed planes.

If the stacking-fault energy of copper is taken as the median, face-centered-cubic metals with low stacking-fault energies include silver, gold, and solid-solution alloys of copper-zinc (brass) and some stainless steels. High stacking-fault pure metals include nickel and aluminum. Also, stacking faults are found in other crystal types such as close-packed-hexagonal.

Because some metals display characteristics of other bonding types, they occur in other crystal structures primarily because of the interference of electrons at energy levels that are not centrally symmetric. All close-packed-hexagonal metals belong to two categories that can be illustrated by packing elongated or compressed spheres (ellipsoids of revolution). The close-packed equatorial circular cross sections constitute a close-packed plane. An ideal hexagonal-close-packed structure illustrates the hexagonal stacking of spheres, whereas spheres compressed in a direction perpendicular to a close-packed plane in the *c* direction have an axial ratio less than the ideal of 1.633; examples are cobalt, rhenium, titanium, gadolinium, hafnium, and beryllium. Finally, if spheres are extended in the *c* direction (prolate ellipsoids with major axis in *c* direction), the axial ratio is greater than the ideal of 1.633; examples are lanthanum, cadmium, and zinc.

Other metals such as iron, titanium, tungsten, molybdenum, niobium, tantalum, potassium, sodium, vanadium, chromium, and zirconium (alkali and transition metals) are slightly less closely packed, with a body-centered-cubic crystal structure that is 8.1% lower in packing density than face-centered-cubic crystals. Furthermore, the equilibrium crystal structure for many metals is dependent upon temperature and pressure. Consequently, a large number of elements are allotropic; that is, the element will appear as different crystal structures at various ranges of pressure-temperature conditions. Included in this group are calcium, titanium, manganese, iron, cobalt, yttrium, zirconium, tin, hafnium, thallium, polonium, thorium, and uranium. *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Anisotropy of crystalline materials. In spite of the same bond attraction in all directions for equivalent atoms, when atoms are packed in crystalline structures, physical properties such as electrical conductivity, thermal conductivity, elastic modulus, and coefficient of thermal expansion are unique for different crystallographic directions. The characteristic of distinct properties for crystallographic directions is called anisotropy. It is a result of contrasts in linear densities of atoms in the lattice and the quantum-mechanical constraints on the relative ease of movement of electrons and phonons in specific directions. *See* CRYSTAL; PHONON.

The close-packed planes and close-packed directions are of special interest in crystal nucleation and growth from vapor or liquid, and in the solid. Moreover, most static and dynamic properties of crystalline materials are related to these unique combinations of direction and planes. For example, in plastic deformation one portion of a crystal translates with respect to another on a local portion of a slip plane, producing permanent deformation in a substantial part of the crystal. The permanent plastic deformation is a consequence of the movement of dislocations, usually on the close-packed plane in a close-packed direction.

Defects in metal crystals. Many imperfections of various types in the lattice periodicity are found in metal crystals. Some are produced as a result of growth errors during the formation of solid from the molten liquid or from the vapor states. These defects play a major role in the magnitude of electrical, mechanical, and thermal properties of metals. Defects are classed in four categories: point, line, surface, and volume. Any displacement of a normal lattice atom to an intermediate position because of any strain,

including internal strains as would be generated by lattice defects, increases the energy of the configuration of atoms.

Point defects. Point defects are considered to originate from some small, nearly one-dimensional error inside the crystal, although, at a lattice point, defects produce a three-dimensional short-range strain in the crystal. For a missing atom (vacancy) there is an average inward shift on the surrounding atoms. For a different-size atom that fits in a lattice site (substitutional atom), an inward or outward localized shift depending on the effective size of the foreign atom follows. For regions that are not lattice locations, interstitial atoms of the same element or foreign elements produce strains depending on the relative size of the foreign element and its location in the lattice. Lattice atoms adjacent to the defect rearrange to minimize the resultant increase in energy from the strained lattice by the rebalancing of the altered interaction-force array.

Line defects. A line defect is a linear imperfection extending from a surface, internal or external, to a point on the same surface, another surface, or a dislocation. The significant strain field perpendicular to the line defect extends to about 10 atom distances from the center and along the direction of the dislocation. Although the dislocation is a linear defect, it is seldom a straight line; and the features of its strain field are related to the orientations in the lattice. This line imperfection is a singularity in that it is surrounded by a multiply connected crystal. When one atom-to-atom circuit, away from any strained region, is made around the dislocation, the resultant displacement from the starting atom to the finishing atom can be compared to the equivalent circuit in a perfect crystal. The resultant displacement is the closure failure in a multiply connected crystal. This vector, from the starting atom to the last atom, is defined as the Burgers vector of the dislocation.

Surface defects. Surface defects are boundaries, usually between one crystal and another of the same phase or different phases. This type of imperfection also includes the other surface that surrounds the crystal, twin boundaries, or any other two-dimensional surface that is a consequence of a surface discontinuity to the lattice.

Volume defects. Volume defects can be composed of small groups of foreign atoms with a crystal structure coherent with the surrounding crystal or voids assembled by uniting vacancies or gas atoms. Because of the modification of the lattice spacing, strained zones surround all types of imperfections.

Role of defects. Any defect that disrupts the lattice periodicity interferes with the transfer of wavelike information, such as electromagnetic radiation or elastic waves (phonons), hence disturbing the usual constructive diffraction processes for these quantum-mechanical wavelike particles. Thus, the more defects present in the metal solid, the more obstructions to the flow of fundamental particles, and, for example, the lower the magnitudes of electrical and thermal conductivities by electrons and phonons. Also, more defects in the solid present more obstacles to the movement of dislocations. However, the properties that are directly associated with the strength of the bonds and not the electron flow, such as elastic constants and coefficient of thermal expansion, are affected to a minor degree. *See* CRYSTAL DEFECTS.

Atom movements. Because of the negligible influence of nondirectional bonds to a particular location in the lattice, the small, temporary increase in energy produced during the translation of an atom from one low-energy lattice position to another is due to the localized atom strain during the transfer from one stable lattice site to another in spite of the increase in entropy from the loss in regional periodicity. The energy that must be applied to transfer an atom into a vacant position (diffusion) can be supplied by relatively low thermal activations. The rate of atoms jumping from lattice sites to vacant lattice sites is dependent exponentially on the absolute temperature of the solid. Not only are atom positions easily changed in a metal by thermal activation, but the reduction of strain energy or other metastable energy states also drives atom diffusion (one of the mechanisms in the creep process). *See* CREEP (MATERIALS); DIFFUSION; LATTICE VIBRATIONS.

The most effective method of reducing shear-strain energies in a crystal is shear translation of a portion of a crystal, usually on the slip plane in the slip direction, by action of dislocations. After a dislocation advances, the part of the crystal that is moved on the slip plane is reformed into a perfect lattice by merely translating the portions of the crystal on either side of the plane of the dislocation by one lattice position (Burgers vector). This shear process is affected in a minor way by thermal activation.

Polycrystalline metals. Common metallurgical production methods inevitably lead to solids made up of many small crystals separated by higher-energy boundaries containing non-lattice-positioned arrangement of atoms. The crystals in these polycrystalline metals are called grains, and the boundaries are grain boundaries. Since the grain boundaries interrupt periodicity of the lattice, they have a large effect on many of the physical and mechanical properties. They are origins (sources) and terminators (sinks) to all other defects, especially vacancies, solute atoms, and dislocations.

In an ideal polycrystalline arrangement composed of many randomly oriented grains, the anisotropic features of each grain are averaged; and the resulting metal solid would have isotropic properties. Actually, metal manufacturing processes such as casting and deformation forming (forging, rolling, and so forth) tend to develop grains with specific crystallographic directions distributed about certain directions that are distinctive of the forming process and the geometry of the bulk product. This nonrandom arrangement of grains is called preferred orientation; this contributes to anisotropy of properties in the final polycrystalline product.

Because of the change in crystal orientation across the grain boundary and the nonperiodic atom

distribution in the boundary region, the advancement of dislocations is obstructed. Therefore, grain boundaries increase the stress at which plastic deformation takes place. Furthermore, any process (such as plastic deformation and electrical and thermal conductivity) whose flow rate is impeded by lattice imperfections is decreased a greater amount as the average grain size becomes smaller (area of boundary per unit volume is larger). *See* GRAIN BOUNDARIES.

Metal alloys. An alloy has metallic features and is composed of two or more chemical elements. Macroscopically, it appears homogeneous, but microscopically various thermodynamically stable arrangements of the chemical elements are possible, depending on the characteristics of the components. If the sizes of the unlike, non-compound-forming atoms are about the same or the additional element is small enough to fit between the lattice atoms without creating substantial lattice strain, they will intersperse in the host atom lattice and form a solid solution. However, there will be a difference in the average distance between the atoms because of the slight strains, and the properties of the metal will be altered (conductivities will decrease and the movement of dislocations will be impeded). If the attraction of the species of dissimilar atoms is greater than the attraction of the like atoms, a periodic arrangement of a different atomic assortment called an ordered solid solution can result. If the elements greatly differ electrochemically, giving the bonds a partly ionic characteristic, the unlike atoms will form an intermetallic compound. If unlike elements tend to repel one another, a heterogeneous microstructure of discrete phases that is separated by grain boundaries develops. The phases of an alloy system as a function of temperature, composition, and pressure are clearly shown in equilibrium phase diagrams. These may be considered maps, based on thermodynamic principles, the demonstrate the stable state of the alloy for a given temperature, composition, pressure, and number of components. *See* ALLOY; ALLOY STRUCTURES; INTERMETALLIC COMPOUNDS; METALLOGRAPHY; METALLURGY; PHASE EQUILIBRIUM; SOLID SOLUTION.

Irwin G. Greenfield

Bibliography. C. S. Barrett and T. B. Massalski, *Structure of Metals*, 3d rev. ed., 1980; A. H. Cottrell, *Introduction to the Modern Theory of Metals*, 1988; A. H. Cottrell, *Theoretical Structural Metallurgy*, 1957; A. Kelly and N. H. Macmillan, *Strong Solids*, 1986; C. S. Smith, Materials, *Sci. Amer.*, 217(3):69–79, September 1967; W. A. Tiller, *The Science of Crystallization: Microscopic Interfacial Phenomena*, 1991; L. H. Van Vlack, *Elements of Materials Science and Engineering*, 6th ed., 1989; R. Zallen, *The Physics of Amorphous Solids*, 1983.

Metal, mechanical properties of

Commonly measured properties of metals (such as tensile strength, hardness, fracture toughness, creep, and fatigue strength) associated with the way that metals behave when subjected to various states of stress. The properties are discussed independently of theories of elasticity and plasticity, which refer to the distribution of stress and strain throughout a body subjected to external forces.

Stress states. Stress is defined as the internal resistance, per unit area, of a body subjected to external forces. The forces may be distributed over the surface of a body (surface forces) or may be distributed over the volume (body forces); examples of body forces are gravity, magnetic forces, and centrifugal forces. Forces are generally not uniformly distributed over any cross section of the body upon which they act; a complete description of the state of stress at a point requires the magnitudes and directions of the force intensities on each surface of a vanishingly small body surrounding the point. All forces acting on a point may be resolved into components normal and parallel to faces of the body surrounding the point. When force intensity vectors act perpendicular to the surface of the reference body, they are described as normal stresses. When the force intensity vectors are parallel to the surface, they describe a state of shear stress. Normal stresses are positive, when they act to extend a line (tension). Shear stresses always occur in equal pairs of opposite signs.

A complete description of the state of stress requires knowledge of magnitudes and directions of only three normal stresses, known as principal stresses, acting on reference faces at right angles to each other and constituting the bounding faces of a reference parallelepiped. Three such mutually perpendicular planes may always be found in a body acted upon by both normal and shear forces; along these planes there is no shear stress, but on other planes either shear or shear and tensile forces will exist.

The shear stress is a maximum on a plane bisecting the right angle between the principal planes on which act the largest and smallest (algebraic) principal stresses. The largest normal stress in the body is equal to the greatest principal stress. The magnitude and orientation of the maximum shear stress determine the direction and can control the rate of the inelastic shear processes, such as slip or twinning, which occur in metals. Shear stresses also play a role in crack nucleation and propagation, but the magnitude and direction of the maximum normal stress more often control fracture processes in metals capable only of limited plastic deformation.

It is often useful to characterize stress or strain states under boundary conditions of either plane stress (stresses applied only in the plane of a thin sheet) or plane strain (stresses applied to relatively thick bodies under conditions of zero transverse strain). These two extreme conditions illustrate that strains can occur in the absense of stress in that direction, and vice versa.

Tension and torsion. In simple tension, two of the three principal stresses are reduced to zero, so that there is only one principal stress, and the maximum shear stress in numerically half the maximum normal stress. Because of the symmetry in simple tension,

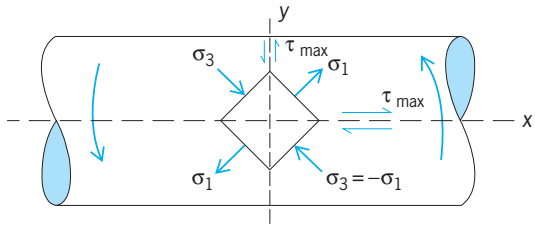

Fig. 1. State of stress in torsion. (*After G. E. Dieter, Mechanical Metallurgy, McGraw-Hill, 1965*)

every plane at 45° to the tensile axis is subjected to the maximum shear stress. For other kinds of loading, the relationship between the maximum shear stress and the principal stresses are obtained using the same method, with the results depending upon the loading condition.

For example, in simple torsion, the maximum principal stress is inclined 45° to the axis of the bar being twisted. The least principal stress (algebraically) is perpendicular to this, at 45° to the bar axis, but equal to and opposite in sign to the first principal stress—that is, it is compressive (**Fig. 1**). Both of these are in a plane perpendicular to the radial direction, the direction of the intermediate principal stress, which in this case has the magnitude zero. Every free external surface of a body is a principal plane on which the principal stress is zero. In torsion, the maximum shear stress occurs on all planes perpendicular to and parallel with the axis of the twisted bar. But because the principal stresses are equal but of opposite sign, the maximum shear stress is numerically equal to the maximum normal stress, instead of to half of it, as in simple tension. This means that in torsion one may expect more ductility (the capacity to deform before fracture) than in tension. Materials that are brittle (exhibiting little capacity for plastic deformation before rupture) in tension may be ductile in torsion. This is because in tension the critical normal stress for fracture may be reached before the critical shear stress for plastic deformation is reached; in torsion, because the maximum shear stress is equal to the maximum normal stress instead of half of it as in tension, the critical shear stress for plastic deformation is reached before the critical maximum normal stress for fracture.

Another important situation arises at the root of a geometrical discontinuity in a loaded body. Such a discontinuity is referred to as a notch. At the root of a notch, immediately beneath but very near the surface, the magnitude of the maximum normal stress may increase to values considerably greater than the applied normal stress; the maximum shear stress, instead of being equal to the maximum normal stress as in torsion, or half of it as in tension, may be a considerably smaller fraction. Also, the critical normal stress for fracture is more easily reached before the critical shear stress for deformation than in torsion and tension, and one may expect notches to have an embrittling effect. This embrittling effect of notches is of very great practical importance and accounts for the majority of service failures in engineering structures. The depth and radius of notches are parameters of overriding importance in determining fracture resistance.

Stress-strain curve. If a metal is strained at a fixed rate, the resistance to deformation increases with straining at a diminishing rate. A plot of the resistance to deformation against the amount of strain is called a stress-strain curve, and the slope of this curve at any point is called the rate of strain hardening. Such a curve is normally obtained by making a tension test and plotting the maximum normal stress against the maximum normal strain. These are identical with the ordinary tensile stress (the tensile force divided by the area of the specimen) and the ordinary tensile strain (the extension per unit of length). The curve produced is known as an engineering stress-strain curve and is widely used to provide design strength information and as an acceptance test for materials. For small strains, the original area and the original gage length may be used without serious error in calculating the stress and strain. Where the strains are large, it is better to use the true average stress, obtained everywhere on the curve by dividing the tensile force by the actual area (which decreases with increasing deformation), and the so-called true strain. The true strain is obtained for every point on the curve by integration of dl/l, between the limits of original length l_o and actual length l, thus taking into account the changing gage length. The true strain is thus the natural logarithm of the ratio of the final length to the original length as expressed in Eq. (1).

$$\int_{l_o}^{l} \frac{dl}{l} = \ln \frac{l}{l_o} \qquad (1)$$

The ratio of the final to the original length may be replaced by the ratio of the original area to the final area, because of the constancy of volume in inelastic deformation. True stress-strain curves are more likely to be utilized in circumstances where fundamental information on deformation mechanisms is desired.

For single-phase metals which have not suffered any prior permanent deformation, the stress-strain curve in simple compression is usually nearly identical with the stress-strain curve in simple tension. This is not true, however, for two-phase materials where the second phase is present as dispersed hard particles, nor is it true for materials within which occur residual stresses resulting from, for example, nonuniform thermal contraction from the processing temperature to the test temperature in the coexisting phases. For other states of stress, additional complications arise. In torsion, for instance, the maximum normal stress reaches a value, at corresponding strains, of less than 60% of the value reached in simple tension. A more generally applicable stress-strain relationship is obtained by plotting the von Mises function of the three principal stresses σ_1, σ_2, and σ_3 (the effective stress $\overline{\sigma}$), as in Eq. (2), against an

$$\overline{\sigma} = [{}^1/_2(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2]^{1/2} \qquad (2)$$

easily derivable function of the three principal strains that may be called the effective strain $\bar{\epsilon}$, given in Eq. (3), where $\bar{\epsilon}$ is the effective strain and ϵ_1, ϵ_2, and

$$\bar{\epsilon} = [{}^2/_3(\epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2)]^{1/2} \tag{3}$$

ϵ_3 are the principal strains. Both the effective stress and the effective strain reduce to the ordinary tensile stress and tensile strain for simple tension testing, so that the tensile stress-strain curve without change in units may be used as the effective stress-strain curve.

Unless some microstructural change occurs in the metal during straining, such as recrystallization or precipitation of a second phase, the effective stress-strain curve usually is describable by the power function $\bar{\sigma} = k\bar{\epsilon}^n$, where k is the effective stress required to produce unit effective strain, and n is an exponent that is always less than unity and commonly has a value between 0.1 and 0.4. The exponent n is called the strain-hardening index because differentiation of $\bar{\sigma} = k\bar{\epsilon}^n$ to find the rate of strain hardening $d\bar{\sigma}/d\bar{\epsilon}$ yields for this expression $n(\bar{\sigma}/\bar{\epsilon})$, so that the slope of the stress-strain curve (the strain-hardening rate) is proportional to n as well as to the ratio of the coordinates of the point on the curve. Higher values of n are associated with annealed pure metals and annealed single-phase alloys; lower values of n are associated with heat-treated alloys such as quenched and tempered steels.

Modern theory makes use of line imperfections (dislocations) to explain the low resistance to initial plastic deformation of annealed metal crystals and polycrystalline aggregates, as well as strain hardening and viscous effects. Imperfection-free crystals deform only elastically and fracture at very high stress levels. This has been demonstrated with metal whiskers of high purity and of extremely small diameter. Larger crystals invariably contain dislocations which move through the crystal lattice at low stress levels, but pile up at barriers such as hard particles or surfaces of contact with other crystals (grain boundaries). The dislocations are surrounded by stress fields which interact with each other, and strain hardening is ascribed to interactions between dislocations moving on different slip systems as well as interactions with other barriers. Viscous effects are associated with the rate at which line imperfections can break loose from their obstacles, usually other line imperfections, with the help of thermal motion. *See* CRYSTAL STRUCTURE; ELASTICITY; PLASTIC DEFORMATION OF METAL; STRESS AND STRAIN.

Tension test. To achieve uniformity of distribution of stress and strain in a tension test requires that the specimen be subjected to no bending moment. This is usually accomplished by providing flexible connections at each end through which the force is applied. The specimen is stretched at a controllable rate, and the force required to deform it is observed with an appropriate dynamometer. The strain is measured by observing the extension between gage marks adequately remote from the ends, or by measuring the diameter and calculating the change in length by using the constancy of volume that characterizes plastic deformation. Diameter measurements are applicable even after necking-down has begun. The elastic properties are seldom determined since these are structure-insensitive. Special refinements are necessary when the elastic stress-strain relationships are in question because the strains are very small and because elastic deformation is not characterized by constancy of volume.

Yield strength. The elastic limit is rarely determined. Metals are seldom if ever ideally elastic, and the value obtained for the elastic limit depends on the sensitivity of strain measurement. The proportional limit, describing the limit of applicability of Hooke's law of linear dependence of stress on strain, is similarly difficult to determine. Modern practice is to determine the stress required to produce a prescribed inelastic strain, which is called the yield strength. The amount of strain used to define the yield strength varies with the application, but is most commonly taken as 0.2% (a unit strain of 0.002 in./in.; 1 in. = 25 mm). Because upon unloading the behavior is almost linearly elastic, it is possible to use the offset method of determining the yield strength from a plotted stress-strain curve; a line with a slope equal to the elastic slope is drawn, displaced from the stress-strain curve at low stress levels by the amount of strain used in the definition of the yield strength, and the stress at the intersection of this line with the stress-strain curve is taken as the yield strength (**Fig. 2**). The elastic slope will not be equal to Young's modulus unless an exceedingly stiff tensile machine is employed or strain gages are attached to the specimen.

The stress-strain curve obtained at room temperature is not appreciably affected by changes in the rate of straining in ordinary tensile tests. At higher temperatures, however, strain rate effects are much more important. A general relation between flow stress and strain rate $\dot{\epsilon}$, at constant temperature and strain, is shown in Eq. (4), where m is the strain-rate

$$\sigma = C\dot{\epsilon}^m \tag{4}$$

sensitivity. For most metals, m increases with both test temperature and strain.

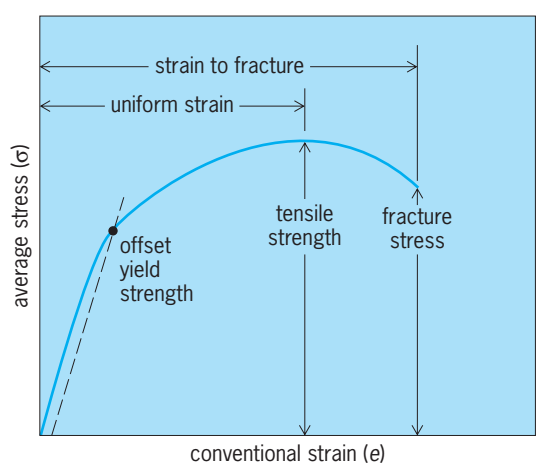

Fig. 2. Engineering stress-strain curve. (*After G. E. Dieter, Mechanical Metallurgy, McGraw-Hill, 1965*)

Test temperature at constant strain rate has a very important influence on yield strength. For most metals, strength increases and ductility decreases with decreasing temperature. For face-centered cubic metals such as copper, nickel, and silver, these effects are small, but for iron, steels, and other body-centered cubic metals and alloys, increases in strength of up to 800% can occur between room temperature and −321°F (77 K). Ductility also may change rapidly in this range. The temperature dependence of yielding in many metals below room temperature may be influenced by impurities or intentional alloying elements.

Changes in strength with temperature will occur also when microstructural changes such as precipitation, strain aging (clustering of solute atoms around dislocations), or recrystallization occur. Extended exposure at elevated temperatures may cause creep, as discussed in a later section of this article.

The stress-strain properties of metals also may be altered drastically by changes in grain (crystal) size. A linear relation is observed for most metals between yield stress σ_y and the reciprocal square root of grain size $d^{-1/2}$, as given by Eq. (5), where σ_o is

$$\sigma_y = \sigma_o + k_y d^{-1/2} \tag{5}$$

a measure of the stress to plastically deform a single crystal, while k_y is a parameter which reflects the ease of propagating slip from one grain to another. For metals in which slip is transmitted with difficulty across grain boundaries (for example, hexagonal close-packed metals such as zinc, cadmium, magnesium, and certain face-centered cubic alloys), k_y is relatively large. However, for pure face-centered cubic or body-centered cubic metals, k_y is a small number. Both σ_o and k_y tend to increase with decreasing temperature. Thus grain boundary strengthening is particularly effective in metals such as zinc, cadmium, and magnesium at low test temperatures.

Tensile strength. Tensile strength, usually called the ultimate tensile strength, is calculated by dividing the maximum load by the original cross-sectional area of the specimen. It is, therefore, not the maximum value of the true tensile stress, which increases continuously to fracture and which is always higher than the nominal tensile stress because the area continuously diminishes. For ductile materials the maximum load, upon which the tensile strength is based, is the load at which necking-down begins. Beyond this point, the true tensile stress continues to increase, but the force on the specimen diminishes. This is because the rate of strain hardening has fallen to a value less than the rate at which the stress is increasing because of the diminution of area. If at any time during the extension of a specimen the force required to extend it diminishes, the strain will become localized. The condition for necking-down is therefore that the force go through a maximum; that is, $\partial F/\partial l$ (the partial derivative of the force with respect to extension) goes through zero. When force and length are expressed as functions of stress and strain, it is easy to show that this condition is satisfied when the true rate of strain hardening falls to a value numerically equal to the stress; that is, $\partial\sigma/\partial\epsilon = \sigma$. For the stress-strain relationship used above, $\overline{\sigma} = k\overline{\epsilon}^n$, this slope is reached when $\epsilon = n$. Necking-down, therefore, will occur when the tensile strain reaches a value equal to n (usually between 10 and 40% extension), and the tensile strength for a ductile material has nothing to do with fracture. It describes only the stress required to produce the tensile instability evidenced by necking-down. For a relatively brittle material, one which fractures before the strain reaches the necking-down value, the tensile strength reflects the fracture strength, but again is not equal to it.

Yield point. A considerable number of alloys, including those of iron, molybdenum, tungsten, cadmium, zinc, and copper, exhibit a sharp transition between elastic and plastic flow. The stress at which this occurs is known as the upper yield point. A sharp drop in load to the lower yield point accompanies yielding, followed, in ideal circumstances, by a flat region of yield elongation; subsequently, normal strain hardening is observed (**Fig. 3**). In the case of polycrystalline mild steels as well as concentrated solid solutions of some single-crystal alloys, the yield elongation marks a period in which strain occurs by the movement of a plastic strain wave along the specimen length. The force on the specimen remains constant during the advance of this strain wave because plastic strain occurs only at the advancing edge of the wave. The wave front is visible on the specimen surface because of surface relief; the markings are known as Lüder's bands. When the strain wave has completely traversed the specimen, the extension of the entire specimen is identical to that achieved at the wave front.

Straining of metals beyond the yield point causes the latter to disappear upon subsequent unloading and restraining. However, annealing at an appropriate low temperature causes the yield point to return. This observation has led to the conclusion that yield points in metal single crystals and polycrystals are associated with the locking of dislocations by interstitial or substitutional impurities. Yielding occurs

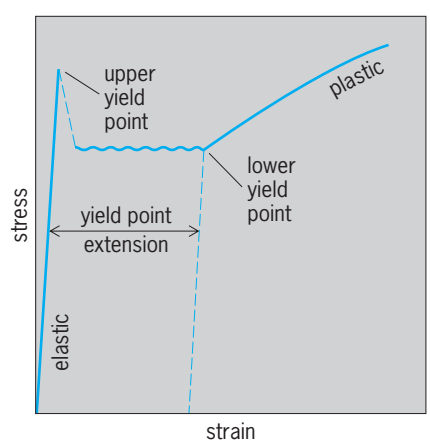

Fig. 3. Yield point of mild steel.

by a process of tearing away dislocations from their impurity "atmospheres," resulting in a sharp drop in load. Annealing of the crystal allows impurities to migrate back to dislocations so that the process may be repeated. It should be noted that other mechanisms for a sharp yield point have been identified as being applicable to metal whiskers and to nonmetallic crystals.

Elongation. The tensile test provides a measure of ductility, by which is meant the capacity to deform by extension. The elongation to the point of necking-down is called the uniform strain or elongation because, until that point on the stress-strain curve, the elongation is uniformly distributed along the gage length (Fig. 2). The strain to fracture or total elongation includes the extension accompanying local necking. Since the necking extension is a fixed amount, independent of gage length, it is obvious that the total elongation will depend upon the gage length, and will be greater for short gage lengths and less for long gage lengths. The American standard ratio of gage length to diameter for a cylindrical test specimen is 2:0.505 [the standard bar has a 2-in. (5-cm) gage length with a diameter of 0.505 in. (1.28 cm) corresponding to an area of 0.2 in.^{2} (1.3 cm^2)]. Specimen configurations have been compiled also for sheet, plate, tube, and wire specimens by the American Society for Testing and Materials. Absolute size does not affect the distribution of deformation. The shape of the cross section has no effect up to ratios of width to thickness of 5:1; rectangular cross sections elongate the same amount as circular cross sections of the same area. The reduction of area at the point of fracture, expressed usually as percentage reduction of area, is independent of gage length, and from it can be calculated the true strain at fracture; this makes the tensile reduction of area a more satisfactory measure of ductility than elongation unless one is interested in the amount that a metal will stretch before it begins to neck down. The reduction of area in the tensile test correlates well with measure of ductility obtained in other ways, such as the strain to fracture on the outside of a bend, provided the strain gradient is not too great, such as in notched-bar testing.

Ductile-to-brittle transition. Many metals and alloys, including iron, zinc, molybdenum, tungsten, chromium, and structural steels, exhibit a transition temperature, below which the metal is brittle and above which it is ductile. The form of the transition for polycrystalline iron and several iron-cobalt alloys is shown in **Fig. 4**. The transition temperature very clearly is sensitive to alloy content, but it will vary even for the same material, depending upon such external test conditions as stress state and strain rate, and microstructural variables such as purity and grain size. The ductility transition frequently is accompanied by a change in the mechanism of fracture (as in iron and steels or zinc), but this need not be so. In general, face-centered-cubic (fcc) materials do not exhibit a ductility transition. Among high-purity body-centered cubic (bcc) metals, the group vanadium, niobium, and tantalum in the periodic table normally are ductile to quite low temperatures; similarly, pure hexagonal close-packed (hcp) titanium and zirconium do not exhibit a ductility transition. However, the presence of interstitial impurities or certain substitutional alloying elements may produce a ductile-to-brittle transition in all of these metals. Moreover, it should be noted that ductile-to-brittle transitions can be induced even in fcc metals by testing in certain aggressive surface environments (for example, alpha brass in mercury) or by adding appropriate small impurities (for example, antimony in copper).

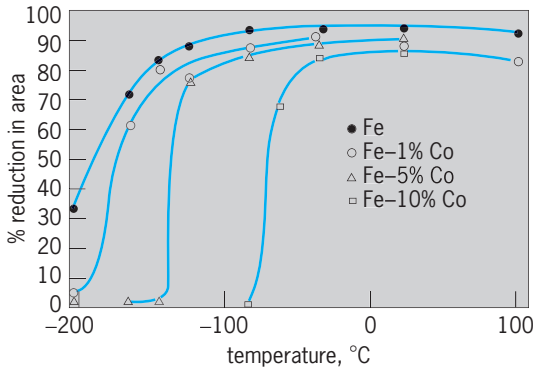

Fig. 4. Ductile-to-brittle transition, iron-cobalt alloys. °F = (°C × 1.8) + 32. (*After N. S. Stoloff, R. G. Davies, and R. C. Ku, Low temperature yielding and fracture in Fe-Co and Fe-V alloys, Trans. Metallurg. Soc. AIME, 233:1500–1508, 1965*)

The transition temperature is different for differing states of stress because the ratio of the maximum shear stress to the maximum normal stress is different. The transition temperature is elevated by the presence of a severe notch and is depressed by testing in torsion. In the case of steels, transition temperatures may vary from near −423°F (20 K) for unnotched low-carbon grades to near −9°F (250 K) in structures which contain defects such as weld cracks or geometrical stress raisers. All large structures contain such defects, and the relationship between defect sizes and fracture stresses or energies embodies the concept of fracture mechanics (see below). There is also a size effect in steels resulting from inhomogeneities in microstructure or coarser grain sizes resulting from slower cooling rates in large structures. Consequently, thick steel plates are characterized by high transition temperatures.

Among the factors which decrease the transition temperature of metals, the following are of general importance to all susceptible metals: (1) increased purity, particularly for bcc metals in the case of interstitial elements (C, O, N, H) and elements which segregate preferentially at grain boundaries (Sb, P, Sn); (2) reduced grain size; (3) reduced strain rate; (4) introduction of very small (diameter of 100 nanometers), hard particles.

In addition, alloying elements in solid solution may be utilized to reduce the transition temperature. Nickel and manganese often are added to steels to improve low-temperature toughness (although manganese is effective only in steels containing relatively

large amounts of carbon). Similarly, lithium lowers the transition temperature of magnesium, and rhenium (a transition metal of the platinum group) improves low-temperature ductility of chromium and tungsten.

Notch tensile test. Notch sensitivity in metals cannot be detected by the ordinary tension test on smooth bar specimens. Either a notched sample may be used in a tension test or a notched-bar impact test may be conducted. The notch in a tensile bar is characterized by the sharpness $d/2r$ and the depth, defined as $1 - d^2/D^2$, where d is the diameter at the root of the notch, D is the original diameter, and r is the notch root radius. Notches produce triaxial stresses under the notch root as tensile forces are applied, thereby decreasing the ratio of shear stress to normal stress and increasing the likelihood of fracture.

Materials are evaluated by a quantity, notch strength, which is the analog of the ultimate tensile strength in an ordinary tensile test. The notch strength is defined as the maximum load divided by the original cross-sectional area at the notch root. For cylindrical bars, a notch ductility also may be defined as in Eq. (6), where $A_{o(n)}$ = original cross-sectional

$$q_n = \frac{A_{o(n)} - A_{f(n)}}{A_{o(n)}} \tag{6}$$

area at notch and $A_{o(f)}$ = cross-sectional area at notch after fracture. This quantity is often very small and is difficult to measure accurately. The ratio of notch tensile strength to ultimate tensile strength is called the notch sensitivity ratio (NSR). If the NSR is less than 1, the material is considered to be notch-sensitive. The notch sensitivity of steels generally increases with increasing tensile strength.

Sharp-notch tension testing is applicable not only to ordinary cylindrical tensile bars but to high-strength sheet materials as well. The method provides a comparative measure of the resistance of sheet materials to unstable fracture resulting from the presence of cracklike defects, and may be utilized in research and development of metals to study effects of composition, processing, and heat treatment. The method also permits comparative evaluation of crack propagation resistance of a number of materials under consideration for a particular application.

Compression test. Very brittle metals, or metals to be utilized in products which are to be formed by compressive loading (rolling, forging), often are tested in compression to obtain yield strength or yield point information. Compression test specimens are generally in the form of solid circular cylinders. The ratio of specimen length to diameter is critical in that high ratios increase the likelihood of buckling during a test, thereby invalidating the test results. Proper specimen alignment is important for the same reason. In addition, care must be taken to lubricate specimen ends to avoid spurious effects from friction between the specimen ends and the testing machine. In the case of a metal which fails in compression by a shattering fracture (for example, cast iron), a quantity known as the compressive strength may be reproducibly obtained by dividing the maximum load carried by the specimen by its cross-sectional area. For materials which do not fail in compression by shattering, the compressive strength is arbitrarily defined as the maximum load at or prior to a specified compressive deformation.

Notched-bar impact testing. Notched-bar impact tests are conducted to estimate the resistance to fracture of structures which may contain defects. The common procedure is to measure the work required to break a standardized specimen, and to express the results in work units, such as foot-pounds or newton-meters. The notched-bar impact test does not provide design information regarding the resistance of a material to crack propagation. Rather, it is a comparative test, useful for preliminary screening of materials or evaluation of processing variables. The notch behavior indicated in a single test applies only to the specimen size, notch geometry, and test conditions involved and is not generally applicable to other specimen sizes and conditions. The test is most useful when conducted over a range of temperatures so that the ductile-to-brittle transition can be determined. In some materials, notably pure polycrystalline bcc metals, there is an abrupt change in energy absorbed with temperature, the curve being similar in shape to that for the change in ductility with test temperature shown in Fig. 4, and an unambiguous transition temperature may be defined.

However, most structural alloys, including steels, exhibit a gradual decrease in impact energy absorption with decreasing temperature, so that arbitrary definitions of transition temperature, based upon a fixed energy level, a fixed contraction at the root of the notch, and average energy, or a change in fracture appearance must be devised. In the case of face-centered cubic metals, there is little change in energy absorption and no change in fracture mechanism with temperature, and no transition can be defined.

Notched-bar tests are usually made in either a simple beam (Charpy) or a cantilever beam (Izod) apparatus, in both of which the specimen is broken by a freely swinging pendulum; the work done is obtained by comparing the position of the pendulum before it is released with the position to which it swings after striking and breaking the specimen (**Fig. 5**). In the Izod test (**Fig. 6***a*), the specimen is held in a vise, with the notch at the level of the top of the vise, and broken as a cantilever beam in bending with the notch on the tension side. In the Charpy test (Fig. 6*b*), the specimen is laid loosely on a support in the path of the pendulum and broken as a beam loaded at three points; the tup (striking edge) strikes the middle of the specimen, with the notch opposite the tup, that is, on the tension side. Both tests give substantially the same result with the same specimen unless the material is very ductile, a situation in which there is little interest.

Specimen geometry and test conditions. The choice of specimen depends to some degree upon the nature of the material; therefore several specimen types

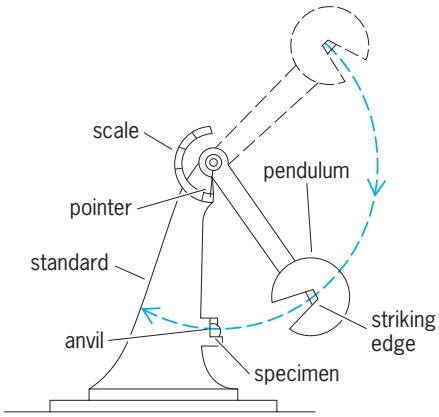

Fig. 5. Diagram of Charpy impact machine.

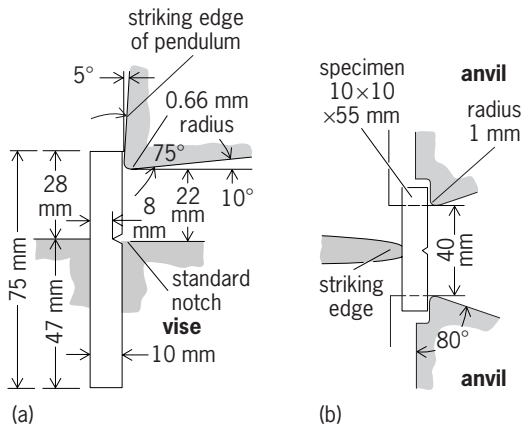

Fig. 6. Mounting of notched-bar test specimens. 1 mm = 0.04 in. (*a*) Izod test. (*b*) Charpy test.

are recognized. Sharp, deep notches are generally required to compare properties of more ductile materials. Typical Charpy bars (Fig. 6*b*) are 0.394 in. (10 mm) square by 2.165 in. (55 mm) long, with either a 45° V notch, or a parallel-sided notch ending in a 0.079-in.-diameter (2-mm) hole (keyhole notch). Keyhole notches usually provide a more abrupt transition at lower temperatures than do V notches, as shown in **Fig. 7**. Notches may also be made with a milling cutter, resulting in a 0.079-in.-wide (2-mm) U-shaped notch. Izod specimens typically are 0.39 in. (10 mm) square by 3.0 in. (75 mm) long, with a 45° V notch 1.102 in. (28 mm) from one end. The transition temperature for a given steel will be different for different tests and specimen geometries.

When testing at temperatures other than room temperature, the Charpy three-point loading arrangement is more convenient than the Izod configuration, in which the specimen is clamped in a vise. It has been established that the temperature of the specimen will not change significantly at the root of the notch for several seconds, so that the specimen may be brought to temperature in a controlled temperature bath and transferred to the testing machine for breaking.

The V-notch transition temperature correlates fairly directly with practical defects in structures, on the basis that brittle fractures in service are seldom observed when a specimen exhibits some minimum level of energy absorbed in the V-notch test. The number of foot-pounds that are needed to ensure freedom from brittle fracture in a structure depends on the material being tested, for the resistance to deformation as well as the ductility enters into the work required to break the specimens, along with the distribution of deformation, which varies from one material to another. For this reason the impact test cannot provide design data.

As in all tests involving brittle behavior, because brittle fracture is concerned with properties in a small region rather than the average throughout a large volume, considerable statistical variation is to be expected in notched-bar testing. Therefore, enough tests must be made to obtain a reliable result by statistical standards, especially when testing in the vicinity of the transition temperature.

Hardness testing. Tension testing provides a complete description of the relationship between stress and strain in plastic deformation, some useful information about ductility, and—in conjunction with the Griffith theory, which relates fracture stress to defect size in an elastically strained body—some information about the resistance to crack propagation under conditions in which the metal is embrittled by the presence of a notch or crack. Often these methods of testing are more time-consuming and more expensive than is necessary, in particular when the only information that is needed is the comparison of the resistance to deformation of a particular sample or lot with a standard material. For such purposes, indentation hardness tests are used. They are relatively inexpensive and fast. They tell nothing about ductility and little about the relationship between stress and strain, for in making the indent the stress and strain are nonuniformly distributed.

In all hardness tests, a standardized load is applied to a standardized indenter, and the dimensions of the indent are measured. This applies to such methods as scratch hardness testing, in which a loaded diamond is dragged across a surface to produce, by plastic deformation, a furrow whose width is measured, and the scleroscope hardness test, in which

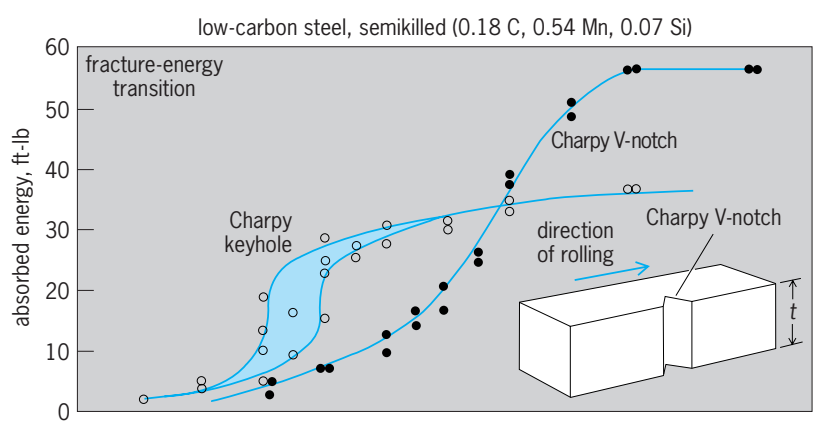

Fig. 7. Fracture energy transition for Charpy keyhole and V notch, showing a more abrupt transition for the keyhole notch. 1 ft-lb = 1.36 J. (*After American Society for Metals, Metals Handbook, vol. 10, 8th ed., 1975*)

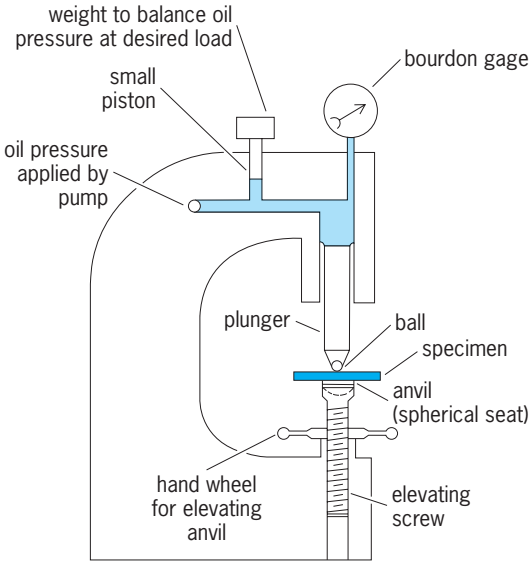

Fig. 8. Features of hydraulic-type Brinell machine.

an indent is produced by dropping a mass with a spherical tup onto a surface. The dimensions of the indent are proportional to the work done in producing it, and the ratio of the height of rebound to the height from which the tup was dropped serves as an indirect measure of the hardness.

Test methods. In the Brinell hardness test, a hardened steel ball is forced into a surface by a force appropriate to the hardness of the material being tested, for a standard time, usually 30 s (**Fig. 8**). With steel and other hard materials, a 6600-lb (3000-kg) force is applied to a 0.394-in.-diameter (10-mm) steel or tungsten carbide ball; with softer alloys a 1100-lb (500-kg) load is often employed. In every instance, the diameter of the indentation crater is kept between one-fourth and one-half the diameter of the ball. The Brinell hardness number (BHN) is defined, in Eq. (7),

$$\mathrm{BHN} = \frac{2P}{\pi D[D - (D^2 - d^2)]^{1/2}} \tag{7}$$

as the force applied divided by the area of contact between the ball and the test piece after the ball has been removed, where P = applied load, kg; D = diameter of ball, mm; and d = diameter of indentation, mm. The units of Brinell hardness are kg/mm^2. The surface on which the impression is made should be relatively smooth and free from scale or dirt. The Brinell number will vary with applied load and ball diameter. In order to obtain a constant hardness number, it is necessary to keep d/D = constant; this can be approximated when P/D^2 is maintained constant.

The Meyer hardness number is defined as the force applied divided by the area of the indent projected onto the plane of the original surface of the test piece. This number, just as useful as the Brinell hardness number, is easier to calculate but has not replaced the standard Brinell hardness number, the calculation of which is everywhere facilitated by easily available tables. The Meyer hardness is less sensitive to load than is the Brinell hardness. A Meyer analysis of the Brinell test is carried out by plotting the logarithm of the applied force against the logarithm of the diameter of the indent. This plot is linear and thereby demonstrates the validity of the relationship $F = ad^n$, where F is the force and a and n are the Meyer constants. Meyer's a is the intercept of the line at unit diameter (1 mm, normally), and Meyer's n is the slope of the logarithmic plot; it reflects the strain-hardening characteristics of the metal, provided that a minimum load level is applied.

The indentation hardness number is independent of the force applied to the indenter when a conical or pyramidal indenter is used. A durable diamond indenter is used in the Vickers penetration hardness testing method. It is a square-base diamond pyramid with a dihedral angle between opposite faces of 138°. This is the contact angle in the Brinell test when the indent is three-eighths of the ball diameter, in the middle of the range of indent sizes used in that test. This choice of angle yields hardness numbers in the Vickers penetration test about the same as those obtained in the Brinell test, with the advantage that the hardness number is independent of the load used down to quite small loads and quite small indents. The diamond pyramid hardness number (DPH) or Vickers hardness number (VHN) is defined as the applied load divided by the surface area of the impression; loads typically range from 2.2 to 260 lb (1 to 120 kg) depending upon the hardness of the material under investigation. The length of diagonals of the impression are determined by microscopic measurement, and the DPH is determined from Eq. (8),

$$\mathrm{DPH} = \frac{1.854P}{L^2} \tag{8}$$

where P is measured in kg and L is the average length of diagonals in mm. In practice, the DPH is read from a chart once the diagonals are known. The Vickers test is widely accepted for research applications because a continuous scale of hardness from 5 DPH for very soft metals to 1000 DPH for very hard metals is achieved. Below loads of about 0.88 oz (25 g), departures are observed. The Knoop hardness test is a modification using a diamond pyramid whose base is a rhombus, the diagonals of which are in the ratio of 7:1. The plastic deformation produced by this indenter is nearly plane strain, which can be accomplished in brittle materials without fracture in the adjoining regions under tensile stress, whereas cracks are produced by the nonplane strain produced by symmetrical indenters. Low loads are applied (under 2.2 lb or 1 kg) and, since surface impressions are about one-half the depth of those for Vickers tests under the same load, the test is particularly suitable for analysis of surface layer properties. For the same reason, it is necessary to prepare surface layers much more carefully in the case of Knoop hardness than for tests involving heavier loads.

Rockwell hardness. The most popular hardness test is the Rockwell test, which is carried out quite rapidly in a convenient machine by sacrificing any attempt to express the hardness as a resistance to deformation

in units of force per area. An indenter chosen from two shapes and various sizes is forced into the test-piece surface, first by a minor load, under which the position of the indentor is established as a reference point, and then by a major load which deepens the indent. Upon removing the major load and leaving the minor load applied, the amount by which the indent has been deepened is established. The hardness number is simply the amount of deepening on a linear scale, with the scale reversed so that soft materials having deeper indents are characterized by smaller numbers. The two shapes of indenters are steel spheres of various diameters, for which the starting point on the linear scale for zero deepening is 130; and a conical diamond with a spherical tip for which the starting point on the scale is 100. The hardness number is the number read from the dial gage indicating the depth of the indent on this reversed scale; it must be accompanied by a letter indicating the kind of indenter and load used. For example, for a load of 220 lb (100 kg) applied to a spherical steel indenter of $^1/_{16}$-in. (0.16-cm) diameter, the letter B is used, so that a hardness number might be Rockwell B80; for the diamond indenter and a load of 330 lb (150 kg), the letter C is used, so that a typical hardness number would be Rockwell C55. Tables are available to convert hardness data for steels from one Rockwell scale to another, and to other hardness scales or to tensile strength. *See* HARDNESS SCALES.

Fatigue. Fatigue is a process involving cumulative damage to a material from repeated stress (or strain) applications (cycles), none of which exceed the ultimate tensile strength. The number of cycles required to produce failure decreases as the stress or strain level per cycle is increased. When cyclic stresses are applied, the results are expressed in the form of an *S-N* curve in which stress of logarithm of stress is plotted against the number of cycles to cause failure. The fatigue strength or fatigue limit is defined as the stress amplitude which will cause failure in a specified number of cycles, usually 10^7 cycles. For a few metals, notably steels and titanium alloys, an endurance limit exists, below which it is not possible to produce fatigue failures no matter how often stresses are applied.

It is often desirable to impose fixed cyclic strains on a material, rather than fixed stresses. Under these circumstances, fatigue data are plotted as logarithm of plastic strain per cycle, $\Delta\epsilon_p$, versus logarithm of cycles to failure, N_f. Most engineering metals exhibit a linear dependence of $\Delta\epsilon_p$ on N_f, known as the Coffin-Manson relationship, shown in Eq. (9), where

$$\frac{\Delta\epsilon_p}{2} = \epsilon_f' N_f^{\alpha} \tag{9}$$

ϵ_f' is a material constant known as the fatigue ductility coefficient, and σ is a constant which for most metals tested at room temperature in air is near -0.5; ϵ_f' typically approximates the ductility of the material in monotonic tension, indicating that fatigue life increases with ductility of the material. Equation (9) is obeyed best at lives shorter than 10^4 cycles, the so-called low-cycle fatigue range, as shown for a compilation of several metals tested at a variety of temperatures in **Fig. 9**.

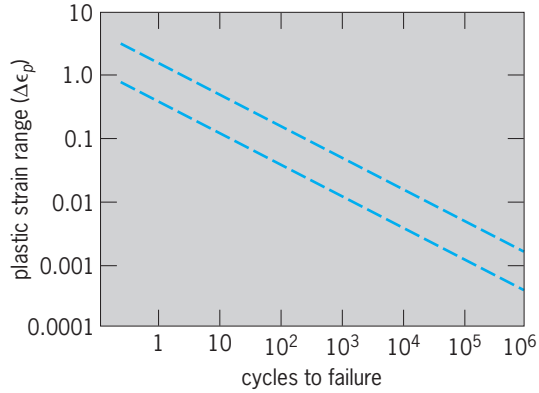

Fig. 9. Low-cycle fatigue data for several alloys. A single band incorporates all data for 11 metals at a variety of test temperatures and test atmospheres. (*After L. F. Coffin, The effect of high vacuum on the low cycle fatigue law, Metallurg. Trans., 3:1777–1788, 1972*)

Fatigue testing in stress control is frequently done with a mean stress of zero, although mean positive stresses are not uncommon. A stress that varies sinusoidally with time is commonly applied, sometimes in direct tension and compression but more usually in bending by utilizing a sample of circular cross section rotated under a constant bending moment. Other waveforms also may be used in tension-compression, most notably triangular or square waves, usually for specialized research applications.

Test frequency is not an important variable at ordinary temperatures, but may assume overriding importance at temperatures sufficiently elevated to permit the introduction of cyclic creep or surface oxidation. Under either of the latter circumstances, low-frequency tests result in shorter lives for the same applied loads. Mean stress superposed upon cyclic stresses can have significant effects, particularly in steel and other alloys sensitive to brittle fracture under tensile loading. In this situation, life decreases as R, the ratio of minimum to maximum cyclic stress, is increased.

Cracks may form very early in a fatigue test, and then grow exceedingly slowly, or, alternatively, crack nucleation may be delayed until very near the total number of cycles to failure. Cracks invariably initiate at surface or near-surface origins such as slip bands, inclusion or precipitate particles, scratches, notches, or other stress raisers, or at material defects such as casting pores. Slip-band cracking is particularly prevalent in pure metals, and under such circumstances the nature of the slip process is a critical factor in determining fatigue life. The greater the multiplicity of active slip systems, the easier it is to initiate a crack; on the other hand, the predominance of one slip system, as in suitably oriented single crystals of zinc and other hcp metals, renders crack nucleation so difficult that it is sometimes impossible to achieve fatigue cracking at all.

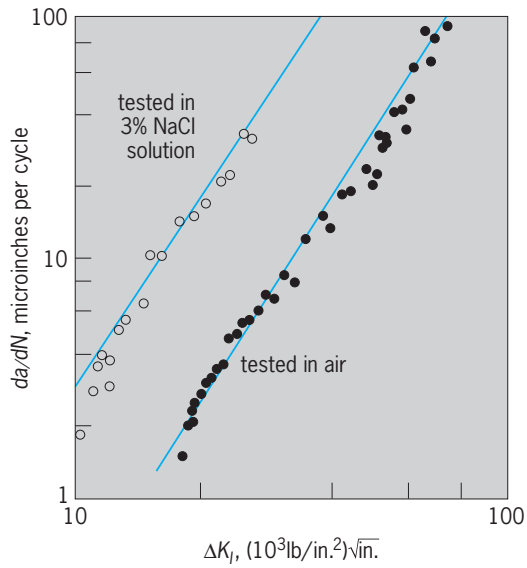

Fig. 10. Crack growth rate of 4340 steel in air and in 3% NaCl. (*After American Society for Metals, Metals Handbook, redrawn from Imhof and Barsom, vol. 10, 8th ed., 1975*)

Crack propagation occurs at a rate da/dN (crack extension per cycle) which depends upon the change in stress intensity K (stress × square root of crack length) per cycle raised to a power m, as expressed in Eq. (10), where C and m are constants

$$\frac{da}{dN} = C(\Delta K)^m \tag{10}$$

that depend upon material and mean stress as well as environment (**Fig. 10**). The value of m has been observed to average near 4 for a wide variety of metals. The fatigue crack continues to grow until the remaining cross-sectional area is insufficient to support the peak stress, fracture then occurring in a single cycle by overload. The appearance of the fracture surface in the fatigue and overload zones differs considerably for ductile materials in that parallel bands known as fatigue striations often occur perpendicular to the direction of fatigue crack propagation. The overload region contains no striations and, for ductile metals, generally comprises microvoids centered about inclusion or precipitate particles. Each striation marks the position of the fatigue crack at the end of a cycle, and the distance between striations serves as a useful measure of crack growth rate. Usually, however, these must be supplemented by microscopic observations of the crack profile with time in order for crack propagation relations such as that of Eq. (10) to be obtained. This is particularly true in materials where striations are not always well developed, such as structural steels and other high-strength alloys with complex microstructures.

Notches may significantly lower fatigue life, particularly in stronger, less ductile alloys. The notch sensitivity factor q in fatigue is given by Eq. (11),

$$q = \frac{K_f - 1}{K_t - 1} \tag{11}$$

where K_t is the theoretical stress concentration factor for ideal elastic behavior, and K_f is the ratio of fatigue strength of an unnotched specimen divided by the fatigue strength of a notched specimen. Both K_f and K_t are greater than unity always. The value of q varies between 0 and 1 and is highest for strong alloys.

Fatigue strengths of engineering alloys tend to be proportional to tensile strength. Therefore any factor which increases monotonic strength, for example, testing at low temperatures or utilizing samples with finer grain size, tends to increase fatigue life. While for many metals the fatigue strength is on the order of one-half the tensile strength, this ratio is reduced for some precipitation-hardened alloys (nickel-base superalloys, aluminum alloys) to as little as one-fourth the tensile strength. Ratios of fatigue strength to tensile strength approaching 0.8 have been noted, however, in some composite materials; for these materials, cracks may be deflected out of their paths by the presence of weak interfaces between phases (artificial composites such as boron reinforced with aluminum) or due to the inability of cracks to penetrate strong fibers (nickel-base or cobalt-base eutectic composites which contain strong nonmetallic reinforcing fibers).

Fatigue life can also be affected by environmental reactions, as shown in Fig. 10. Corrosion acting simultaneously with cyclic deformation can produce appreciable lowering of fatigue strength (corrosion fatigue). The adsorption of liquid metals on solid metals (such as mercury on copper-aluminum alloys) also can lead to striking decreases in fatigue life. Another important factor is residual stresses; tensile stresses act adversely and compressive stresses favorably. Compressive stresses are deliberately applied to metal surfaces to delay crack initiation and thereby prolong fatigue life.

Fracture toughness. Many modifications of conventional test procedures have arisen from an effort to obtain laboratory information which would be useful to designers of engineering structures. Since such structures often contain macroscopic defects, in particular weld flaws, large inclusion particles, or other manufacturing flaws, most effort has been directed toward studying the properties of notched bars. Unfortunately, neither the notched tensile test nor the several impact tests described above provide sufficiently quantitative relations between defect size, stress state, and the likelihood of brittle or otherwise unstable fracture, particularly in cases where crack propagation can occur below the macroscopic yield stress as measured in a tensile test. In the case of materials which exhibit little or no plastic deformation at fracture, the most successful approach has been that of A. A. Griffith, who in 1929 deduced Eq. (12)

$$\sigma_F^2 = \frac{kE\gamma}{\pi c} \tag{12}$$

from both thermodynamic and stress considerations. Here σ_F is the nominal applied stress to cause fracture, k is a numerical constant which depends upon stress state and crack geometry, E is Young's modulus, c is the depth of an edge crack or the half-width

of an internal crack, and γ is the elastic energy necessary to create a new surface. While this equation correctly predicts a lowering of fracture stress as flaw size increases, and in fact provides an accurate measure of elastic work to cause fracture in brittle solids such as glass, several critical problems arise when the equation is applied to metals. When fracture stresses are related to flaw sizes, calculated values of work for fracture, γ_p, are several orders of magnitude higher than those deduced for elastic work by other means. E. Orowan first pointed out that plastic work accompanying fracture must be incorporated into the Griffith equation. This is best done by considering that the total work for fracture, γ_p, is related to γ, the elastic energy, through a proportionality factor which is of magnitude equal to crack tip radius ρ divided by interatomic spacing a. The modified Griffith relation, Eq. (13), is then obtained, where k' is a numerical

$$\sigma_F^2 = \frac{k' E \gamma \rho}{\pi a c} \tag{13}$$

constant. This equation reduces to the Griffith condition, Eq. (12), when ρ approaches a in magnitude, that is, as cracks become very sharp. In practice, ρ is difficult to measure, and a more generally applicable approach has been suggested by G. R. Irwin. A stressed material in which a crack propagates releases strain energy at a rate G. A transition from slow to rapid fracture is denoted by a critical value G_c, which is known as the critical strain energy release rate. Alternatively, G_c may be obtained from a relation between elastic compliance, $1/M$, and crack length c in a series of precracked specimens of differing initial crack lengths as shown in Eq. (14), where

$$G_c = \frac{P^2}{2B} \frac{d(1/M)}{dc} \tag{14}$$

$d(1/M)/dc$ is the change in compliance with increasing length of precrack, B is the thickness, and G_c is a measure of work necessary to cause the crack to propagate. G_c is, therefore, equivalent to γ_p, the distinction being that G_c is an experimentally measured parameter which is not dependent upon assumptions as to its origin. G_c is typically of magnitude 100–1000 in.-lb/in.^{2} for metals. Part of this energy may be dissipated in forming new surface, the rest being converted to plastic work or kinetic energy of the crack.

An alternative and completely complementary approach, based on stress concentrations near crack tips, also has been developed by Irwin, and has been much more widely applied to engineering design situations. Equations for the normal and shear stresses near the crack edge for bodies of various geometries to which forces are applied normal or parallel to the crack edge have been developed. In all cases, the local stresses at the crack edge fall off with distance away from the crack. Fracture of the body occurs when a critical stress distribution, described by the stress intensity factor K, is reached. K depends upon applied stress and crack geometry; for most situations it may be defined as Eq. (15), where σ is a

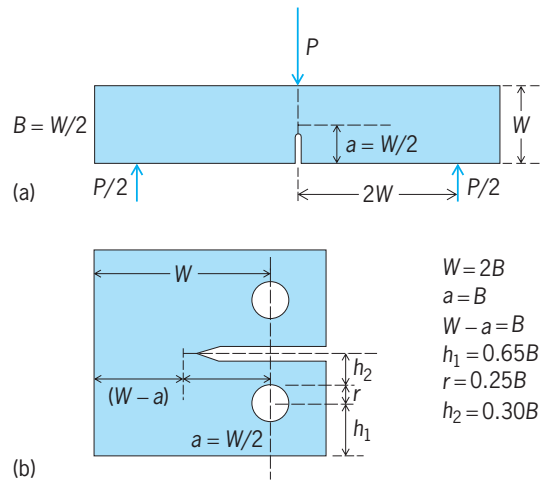

Fig. 11. Standard fracture toughness specimens. (*a*) Edge-notched bend. (*b*) Compact tension. (*After N. H. Polakowski and E. J. Ripley, Strength and Structure of Engineering Materials, Prentice-Hall, 1965*)

$$K = \sigma(\alpha \pi c)^{1/2} \tag{15}$$

parameter depending upon specimen and crack geometry. For a crack in an infinite plate, subjected to loading perpendicular to the crack edge, $\sigma = 1$. K, which is expressed in units of (lb/in.^{2})$\sqrt{\text{in.}}$, is determined experimentally by stressing a precracked specimen in bending or in tension (**Fig. 11**), and noting the critical stress intensity K_c at which unstable (rapid) crack propagation occurs. K and G are related as shown in Eq. (16), where $A = (1 - \nu^2)$ for

$$G = A\frac{K^2}{E} \tag{16}$$

plane strain and $A = 1$ for plane stress; ν is Poisson's ratio.

Critical values of stress intensity K_c and strain energy release rate G_c are related similarly. In addition, since K_c corresponds to a fracture condition, Eq. (16) may be rewritten as Eq. (17), which, when combined with Eq. (15), provides the relation given in Eq. (18),

$$K_c^2 = \alpha \sigma_F^2 \pi c \tag{17}$$

$$\sigma_F^2 = \frac{E G_c}{A \alpha \pi c} \tag{18}$$

which is identical in form to the Griffith condition, Eq. (12).

Unlike the yield stress, which is a reproducible material property for specimens of identical composition and microstructure, K_c and G_c vary with the depth of the precrack relative to the dimensions of the test specimen, particularly width W and thickness B. Specimens with relatively long cracks fracture with appreciable plasticity occurring at the crack tip. The contribution of plastic work is included in G_c or K_c, and is particularly high in thin sections and in low-strength alloys. For high-strength material, and lower-strength alloys in which c/W and c/B are small, K_c and G_c decrease to lower limiting

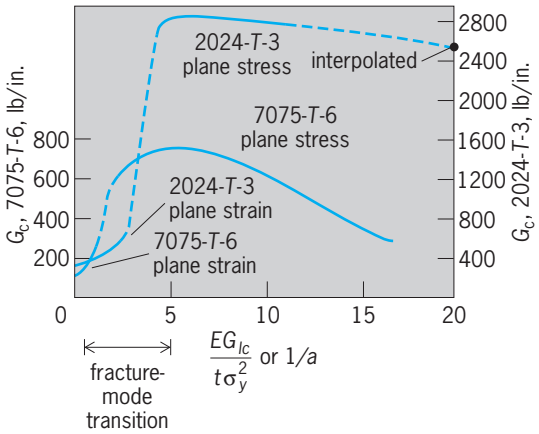

Fig. 12. Critical strain energy release rate as function of thickness. 1 lb/in. = 0.113 N·m. (*After J. N. Goodier and N. J. Hoff, eds., Structural Mechanics, Pergamon Press, 1960*)

values K_{I_c} and G_{I_c}, respectively (**Fig. 12**). These values refer to plane-strain conditions, and the K_{I_c} parameter is known as the plane-strain fracture toughness. K_{I_c} appears to be a material constant insofar as specimen-crack geometry is concerned, and therefore is extremely useful for designers who must estimate structural parameters which are necessary to avoid catastrophic failures. However, K_{I_c} may vary with test temperature, strain rate, and microstructure of the metal. In general, fracture toughness decreases as strength increases, so that K_{I_c} tends to be low for quenched and tempered alloy steels, and much higher for medium-strength carbon steels or aluminum alloys.

Specimen designs which are currently recommended in the United States and Great Britain are shown in Fig. 11*a* and *b*, although additional designs have been utilized in specific situations. Both are single-edge notched specimens which are precracked in fatigue prior to the fracture toughness determination. The bend specimen in Fig. 11*a* is deformed under three-point loading, while the compact tension specimen in Fig. 11*b* is loaded through pins above and below the crack faces. In order to relate applied loads to stress intensity, test specimens must be calibrated. A compliance coefficient (Y function) is published for each specimen-crack-load geometry, which allows load to be converted directly to a K value. The primary variable in specimen preparation is the ratio of crack length c to specimen width W. Calibrations are presented either in tabular or polynomial form as functions of c/W. Plasticity corrections are necessary for situations in which fracture events are not precisely localized at the crack tip, that is, when plastic flow spreads ahead of an advancing crack. The specification of size of the region in which the stress intensity accurately describes the stress field is a critical aspect of fracture toughness testing. In practice, it is usually estimated to be about $^1/_{50}$ the crack length. Therefore the plastic zone at fracture must be limited to $0.02c$ or less. This is accomplished by ensuring that c and $B \geq 2.5\ (K_{I_c}/\sigma_y)^2$, where σ_y is the 0.2% offset yield stress of the material. Calculation of K_{I_c} is based upon the lowest load at which significant measureable extension of the crack occurs. The method has been applied with considerable success to high- and medium-strength steels, as well as to high-strength aluminum- and titanium-base alloys. Modifications of the standard fracture toughness test have also been applied to lower-strength materials, but the test procedures have not yet been standardized.

Creep and stress rupture. Time-dependent deformation under constant load or stress is measured in a creep test. Creep tests are those in which the deformation is recorded with time, while stress rupture tests involve the measurement of time for fracture to occur. The two types of test provide complementary information, and indeed both creep rate and time may be recorded in a single test. Closely related are stress relaxation tests, in which the decay of load with time is noted for a body under a fixed state of strain. Although creep deformation may occur under any state of stress, it is most commonly measured in uniaxial tension. Test durations vary from seconds or minutes to tens of thousand of hours. Appreciable deformation occurs in structural materials only at elevated temperatures, while pure metals may creep at temperatures well below room temperature. In general, the temperature at which creep processes become significant exceeds approximately one-half of the melting temperature on the absolute scale. Under constant load conditions, three stages of inelastic deformation are noted, as shown in **Fig. 13**. These are primary creep (stage I), secondary or steady-state creep (stage II), and tertiary creep (stage III). A considerable portion of primary creep is anelastic and is sometimes referred to as transient creep. The second stage marks a balance between strain hardening of the material and recovery of the microstructure and strength; a constant, minimum creep rate may be noted over most of the life of the specimen, and it is the steady-state creep rate, therefore, which is often determined for structural alloys. Tertiary creep resembles necking in a tensile test in that both are instabilities arising from prefracture events. In this stage, strain hardening cannot compensate for the

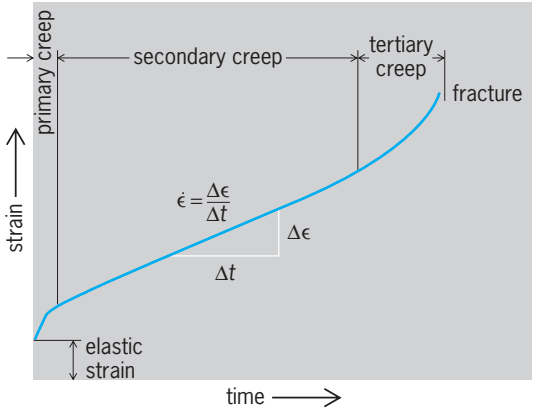

Fig. 13. Creep curve. (*After M. E. Eisenstadt, Introduction to Mechanical Properties of Materials, Macmillan, 1971*)

loss in cross-sectional area during straining as well as the formation of fine cracks. Information usually derived from a series of creep experiments may include, aside from the steady-state creep rate, data on the stress versus logarithm of time for total deformations of up to 5%, as initial stress versus logarithm of time to rupture, and logarithm of stress versus logarithm of creep rate. It should be noted that since the cross-sectional area of a creep specimen decreases with time, the stress increases under constant load conditions. It is possible to run constant stress tests by providing means for load to be reduced as the specimen elongates.

Since creep deformation and rupture time are temperature- and stress-dependent, it is usually necessary to test a material at several stresses and temperatures in order to establish the creep or stress-rupture properties in adequate detail. Interpolation of creep and rupture data is frequently employed to predict behavior at stresses and temperatures which are not included in the test program. However, no generally accepted methods are available for extrapolating creep and rupture data to temperatures, times, and stresses outside of the tested range. Specimen size has little effect on creep and rupture properties. Consequently, test specimen configurations usually resemble those of flat or round tensile bars with reduced gage sections. Temperature measurement must be precise; the difference between indicated and actual test temperature should not exceed ±3.6°F (±2°C) up to 1830°F (1000°C) and ±5.4°F (±3°C) at higher temperatures. Creep strain is measured by an extensometer for tests in which creep rates are to be determined. In the case of stress-rupture tests, dial gages often are employed to provide approximate creep elongations to supplement stress-to-rupture data. *See* METAL; METALLURGY.

N. S. Stoloff

Bibliography. American Society for Metals, *Metals Handbook*, vol. 1: *Properties and Selection, Iron, Steel and High-Performance Alloys*, 10th ed., 1991, and vol. 11: *Failure Analysis and Prevention*, 9th ed., 1988; American Society for Testing and Materials, Standards A370–72a, E92–72, E10–66, E9–70, E8–69, E399–72, E6–66, E23–72; G. E. Dieter, *Mechanical Metallurgy*, 3d ed., 1986; F. A. McClintock and A. S. Argon, *Mechanical Behavior of Materials*, 1966; N. H. Polakowski and E. J. Ripling, *Strength and Structure of Engineering Materials*, 1966.

Metal and mineral processing

The obtaining of useful metals from ores. In the **table** are summarized one or more routes, but not necessarily the only commercial route, used to obtain each of the useful metals by extractive metallurgy. For additional information see the separate articles on each of the metals. *See also* CHEMICAL SEPARATION

Occurrence, extraction, and refining of metals

Metals	Occurrence	Beneficiation operations	Extraction processes: Operation	Extraction processes: Intermediate product	Extraction processes: Product used in refining	Refining (reduction) operations
Aluminum, Al	Bauxite, Al_2O_3, 55–65% Al Alunite, $KAl_3(OH)_6(SO_4)_2$	Dehydration roast Dehydration roast	Caustic (Bayer process) Reduction roast, caustic digestion	Sodium aluminate Alumina hydrate	Pure alumina Pure alumina	Fused salt electrolysis
Antimony, Sb	Stibnite, Sb_2S_3, 50–60% Sb	Liquation		Pure Sb_2S_3	Iron precipitation	
Arsenic, As	Enargite, $3Cu_2S \cdot As_2S_5$ Arsenopyrite, FeAsS		Roasting	AS_2O_3		Charcoal reduction at 1200–1300°F (650–700°C)
Barium, Ba	Barite, $BaSO_4$	Flotation		Pure $BaSO_4$	$BaCl_2$	Fused salt electrolysis
Beryllium, Be	Beryl, $3BeO \cdot Al_2O_3 \cdot 6SiO_2$, 10–12% Be Bertrandite, $Be_4Si_2O_7(OH)_2$	Hand cobbing	Acid sinter plus leaching	BeO	BeF_2	Magnesium at 1652–2372°F (900–1300°C)
Bismuth, Bi	Lead ores	Flotation, gravity	Smelting	Lead bullion	Anode residue from electrolyzed bullion	Chlorine at 932°F (500°C) to remove Pb and Zn Reduction with charcoal and flux
Boron, B	Borax, $Na_2B_4O_7 \cdot 10H_2O$ Kernite, $Na_2B_4O_7 \cdot 4H_2O$ Colemanite, $Ca_2B_6O_{11} \cdot 5H_2O$	Evaporation Brine purification	Crystallization from purified brine, calcining	B_2O_3	Potassium fluoborate	Fused salt electrolysis Reduction with hydrogen Reduction with sodium or magnesium
Cadmium, Cd	Sphalerite, ZnS, 1% Cd	Roasting, sintering	Leaching	Cd-Zn solution	Cd sponge by Zn dust addition	Distillation
		Refuming	Leaching		Purified solution	Electrolysis
Calcium, Ca	Limestone, $CaCO_3$		Calcination	CaO		Aluminum reduction in vacuum
Cerium, Ce	Monazite Bastnaesite	Gravity, flotation	Caustic or acid leaching	$Ce_2(SO_4)_3$	Tetravalent Ce salts Cerium oxides Cerium fluorides	Electrolysis of fused chloride at 1292°F (700°C), or of fused fluoride baths
Cesium, Cs	Lepidolite, 1% Cs	Hand cobbing	Acid leaching	Cesium alum	CsCl Cs hydroxide	Calcium, aluminum, or magnesium reduction
	Pollucite, $Cs_4Al_4Si_9O_{26} \cdot H_2O$	Selective mining of pegmatites	Direct reduction			
Chromium, Cr	Chromite, $FeO \cdot Cr_2O_3$, 35–50% Cr	Gravity and magnetic separation	Electric furnace treatment	Ferrochrome	Chromium alum solution Chromium oxide	Electrolysis Aluminum reduction

Occurrence, extraction, and refining of metals ***(cont.)***

Metals	Occurrence	Beneficiation operations	Extraction processes: Operation	Extraction processes: Intermediate product	Extraction processes: Product used in refining	Refining (reduction) operations
Cobalt, Co	Cobalt or nickel ores, 0.8–10.0% Co Lateritic ores	Flotation, leaching	Acid roast or smelting	$CoSO_4$ solution	$Co(OH)_2$ by milk of lime	Electrolysis of aqueous solution
Columbium, Cb (niobium)	Columbite, $Fe(Cb \cdot TaO_3)_2$	Gravity High-intensity magnetic or electrostatic separation	Caustic fusion, then HCl, then KF Chlorination Solvent extraction Hydrometallurgy	$CbKOF_2$	Columbium oxides	Columbium carbide in vacuum Columbium metal
Copper, Cu	Copper ores, 0.4–6.0% Cu	Flotation, fluid-bed roasting, leaching, pressure leaching, solvent extraction	Smelting, electrolysis, liquid ion exchange, cementation	Cu anodes Ferric chloride leach solution	Cu cathodes Stripped solvent Purified solutions	Electrolysis Other electrowinning methods to produce powdered metal
Dysprosium, Dy	Monazite	Gravity	Sulfuric acid treatment	$Dy_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Erbium, Er	Monazite	Gravity	Sulfuric acid treatment	$Er_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Europium, Eu	Monazite	Gravity	Sulfuric acid treatment	$Eu_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Gadolinium, Gd	Monazite	Gravity	Sulfuric acid treatment	$Gd_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Gallium, Ga	Al and Zn ores	Caustic	CO_2 treatment in acid	Ga_2O_3 and $GaCl_3$		Electrolysis
Germanium, Ge	Sphalerite, ZnS, 0.01–0.015% Ge	Flotation	Roasting	Germanium oxide	Germanium chloride, then hydrolysis to pure oxide	Hydrogen reduction
Gold, Au	Elemental form, 0.001% Au	Gravity, flotation, leaching	Cyanide or smelting	AuCN or Au	Pure gold	Zinc reduction
Hafnium, Hf	Zircon, $ZrO_2 \cdot SiO_2$	Gravity	Acid or alkali leaching	Hf-Zr solutions	Hf solutions by ion exchange to $HfCl_4$; solvent extraction	Magnesium reduction
Holmium, Ho	Monazite	Gravity	Sulfuric acid treatment	$Ho_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Indium, In	Zinc ores	Flotation	Acid roasting	In sulfate	Water solutions	Displacement by zinc metal
Iridium, Ir	Cu-Ni sulfide ores	Flotation, roasting, extraction	Electrolytic refining	Anode slimes	$(NH_4)_2IrCl_6$	Hydrogen reduction
Iron, Fe	Iron ores, 30–60% Fe	Gravity, flotation, magnetic and electrostatic separation, pelletizing	Blast furnace Prereduction or metallizing furnaces	Pig iron Prereduced or metallized pellets		Coke reduction in blast furnace Electric furnace reduction Basic oxygen furnace Direct reduction with CO, H_2, and CH_4
Lanthanum, La	Monazite	Gravity	Sulfuric acid treatment	$La_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Lead, Pb	Galena, PbS, 6–10% Pb	Gravity or flotation	Blast furnace Imperial furnace	Lead bullion	Lead free of Cu, Sn, As, and Sb by treatment with S	Desilvering with Zn, Sb removal
Lithium, Li	Spodumene, $Li_2O \cdot Al_2O_3$ SiO_2 Brines, 0.03–0.24% LiCl	Flotation Evaporation, leaching, flotation	Acid roasting	Li_2SO_4	Li_2,CO_3, then LiCl LiCl	Electrolysis of fused chloride
Lutetium, Lu	Monazite	Gravity	Sulfuric acid treatment	$Lu_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Magnesium, Mg	Seawater, 0.13% Mg Brines Dolomite, $CaMg(CO_3)_2$		Lime slurry Evaporation	$Mg(OH)_2$	MgCl with HCl MgCl	Electrolysis of fused chloride Thermal reduction of dolomite with ferrosilicon
Manganese, Mn	Manganese ores, 45–55% Mn	Flotation, gravity	Roasting in reducing atmosphere	Mn in solution	Pure Mn solutions by filtration	Electrolysis of aqueous solution
Mercury, Hg	Cinnabar, HgS, 1–3% Hg	Sorting, screening	Roasting	Hg		Retorting at 1290°F (700°C)
Molybdenum, Mo	Molybdenite, MoS_2, 1–3% Mo	Flotation	Roasting	MoO_3	Mo by hydrogen reduction	Power metallurgy or are melting
Neodymium, Nd	Monazite	Gravity	Sulfuric acid treatment	$Nd_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Nickel, Ni	Nickel ores, 1–35% Ni Lateritic ores, lower nickel content	Smelting, flotation of Cu-Ni-Fe product Leaching	Roasting, leaching Pressure leach with NH_3	(NiCuFe)S or Ni	Pure NiS by molten Na_2S extraction	Carbon reduction Electrolytic refining Precipitation of Ni as powder from leach solutions
Osmium, Os	Cu-Ni sulfide ores	Flotation, roasting, extraction	Electrolytic refining	Anode slimes	$OsO_2(NH_3)_4Cl_2$	Hydrogen reduction
Palladium, Pd	Cu-Ni sulfide ores	Flotation, roasting, extraction	Electrolytic refining	Anode slimes	$Pd(NH_3)_2Cl_2$	Hydrogen reduction

Occurrence, extraction, and refining of metals *(cont.)*

Metals	Occurrence	Beneficiation operations	Extraction processes: Operation	Extraction processes: Intermediate product	Extraction processes: Product used in refining	Refining (reduction) operations
Platinum, Pt	Cu-Ni sulfide ores, 0.001% Pt	Flotation, roasting, extraction	Electrolytic refining	Anode slimes	$(NH_4)_2PtCl_2$	Hydrogen reduction
Potassium, K	Sylvite, KCl	Flotation, gravity				Sodium reduction
Praseodymium, Pr	Monazite	Gravity	Sulfuric acid treatment	$Pr_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Radium, Ra	Pitchblende	Gravity, flotation	Acid digestion	Ra sulfate	Radium bromide	Repeated fractional crystallization
Rhenium, Re	Molybdenite, MoS_2	Flotation	Roasting	Re_2O_7	$KReO_4$	Hydrogen reduction
Rhodium, Rh	Cu-Ni sulfide ores	Flotation, roasting, extraction	Electrolytic refining	Anode slimes	$(NH_4)_3Rh(NO_2)_6$	Hydrogen reduction
Rubidium, Rb	Lepidolite Pollucite	Hand cobbing	Chlorostannate process followed by pyrolysis, electrolysis, or reduction	Rubidium perchlorate	Chloride Carbonate or hydroxide	High-temperature vacuum reduction with Na or Mg
Ruthenium, Ru	Cu-Ni sulfide ores	Flotation, roasting, extraction	Electrolytic refining	Anode slimes	RuO_2	Hydrogen reduction
Samarium, Sm	Monazite	Gravity	Sulfuric acid treatment	$Sm_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Scandium, Sc	Thortveitite, $(ScY)_2Si_2O_7$ Uranium ores	Leaching with H_2SO_4, then solvent extraction	Roasting with C Stripping solvent with Hf	Scandium carbide Sc_2O_3	Scandium oxide, then ScF_3 ScF_3	Electrolysis of fused salt Direct reduction of fluoride and distillation of Sc metal in vacuum induction furnace
Selenium, Se	Chalcopyrite, $CuFeS_2$	Flotation, roasting	Electrolytic refining	Anode muds	Sodium selenate	SO_2 precipitation from aqueous solution
Silicon, Si	Silica, SiO_2	Flotation, gravity	Smelting			Carbon reduction in electric furnace
Silver, Ag	Nonferrous ores, 0.07% Ag	Gravity or flotation	Blast furnace	Lead bullion	Ag-Zn mixture by Zn treatment	Zn distillation, electrolysis of Ag
Sodium, Na	Halite, NaCl	Flotation, leaching				Electrolysis of fused chloride
Strontium, Sr	Celestite, $SrSO_4$	Hand cobbing, flotation	Digestion with soda ash Calcining with coal; sulfide then leached with lime	$SrCO_3$ $SrCO_3$	$SrCl_2$	Electrolysis of fused chloride
Tantalum, Ta	Tantalite, $Fe(Cb_2TaO_3)_2$ Tin slags	Gravity, electrostatic, magnetic separation	Digestion with HF or caustic, then KF	K_2TaF_7	Ta by sodium reduction	Vacuum sintering at 1290˚F (700˚C)
Tellurium, Te	Chalcopyrite, $CuFeS_2$	Flotation, roasting	Electrolytic refining	Anode muds	Sodium telluride-tellurate	SO_2 from aqueous solution
Terbium, Tb	Monazite	Gravity	Sulfuric acid treatment	$Tb_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Thallium, Tl	Sphalerite, ZnS	Roasting, sintering	Leaching of flue dust	Thallium chloride	TlS by H_2S treatment, then converted to sulfate	Electrolysis of aqueous solution
Thorium, Th	Monazite	Gravity	Sulfuric acid treatment	$Th(SO_4)_2$	Pure salts by fractional crystallization and ion exchange	Reduction of ThF_4 by calcium
Thulium, Tm	Monazite	Gravity	Sulfuric acid treatment	$Tm_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Tin, Sn	Cassiterite, SnO_2, 1.5–5.0% Sn	Gravity, flotation	Reverberatory furnace	Impure tin		Electrolytic refining
Titanium, Ti	Rutile, TiO_2, ilmenite, 1–8% Ti	Gravity, electrostatic, magnetic separation	Chlorination	$TiCl_4$		Magnesium or sodium reduction Electrolysis of $TiCl_4$ in fused salt bath
Tungsten, W	Tungsten ores, 60–70% W	Gravity, flotation	Caustic fusion	Sodium tungstate	WO_3	Hydrogen reduction
Uranium, U	Uranium ores	Leaching with acid or caustic	Ion exchange, solvent extraction		UF_4	Ca Al, Mg, or Na reduction
Vanadium, V	Carnotite, $K_2O \cdot 2UO_3 \cdot V_2O$	Salt roasting	Leaching, solvent extraction, ion exchange	Sodium vanadate	V_2O_5	Thermic reduction followed by electron-beam purification
Ytterbium, Yb	Monazite	Gravity	Sulfuric acid treatment	$Yb_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Yttrium, Y	Monazite	Gravity	Sulfuric acid treatment	$Y_2(SO_4)_3$	Pure salts by fractional crystallization and ion exchange	Electrolysis of fused chloride
Zinc, Zn	Sphalerite, ZnS, 10–30% Zn	Flotation, gravity, pelletizing, roasting	Sulfuric acid leach Fluxing for pyrolytic processes	Purified solution		Electrolysis to produce high-purity cathodes Electrolytic reduction Imperial smelting process
Zirconium, Zr	Zircon, $ZrO_2 \cdot SiO_2$	Electrostatic, magnetic, or gravity concentration	Chlorination, or plasma fusion, then caustic leach		$ZrCl_4$	Magnesium reduction

TECHNIQUES; ELECTROMETALLURGY; FLOTATION; HYDROMETALLURGY; IRON METALLURGY; LEACHING; MECHANICAL SEPARATION TECHNIQUES; METALLURGY; ORE DRESSING; PYROMETALLURGY, NONFERROUS.

Alvin W. Knoerr

Metal-base fuel

A fuel containing a metal of high heat of combustion as a principal constituent. High propellant performance in either a rocket or an air-breathing engine is obtained when the heat of combusion of the fuel is high. Chemically, high heats of combustion are attained by the oxidation of the low-atomic-weight metals in the upper left-hand corner of the periodic table. The generally preferred candidates are lithium, beryllium, boron, carbon, magnesium, and aluminum.

Performance parameters. For the rocket where the total propellant containing both fuel and oxidizer is carried on board, the performance parameter of interest is the heat of combustion per unit mass of combustion product. In air-breathing engines, however, where only the fuel is carried on board and the oxidizer in the form of air is attained free, the figure of merit is the heat of combustion per unit mass of fuel. There is still a third figure of merit for the air-breathing engine in which most of the flight involves a cruise operation in which the thrust just balances the drag. In this case, it is desirable to minimize the cross-sectional area (consequently volume) and maximize the heat of combustion so that a given total energy output can be realized in a minimum-volume vehicle to present low drag. The figure of merit, in this case, is the heat of combustion per unit volume of fuel. *See* PROPULSION.

The **table** shows a listing of the heats of combustion of various candidate metals along with three non-metallized high-performance fuels for comparison. The three different heats of combustion described above are figures of performance merit, depending upon the vehicle mission and type of propulsion system employed. In general terms, for an accelerating rocket, the important figure of merit is Btu/lb pdts (pound of products); for the accelerating air breather, it is Btu/lb fuel; and for the cruising air breather, it is Btu/in.^{3} fuel. *See* PROPELLANT.

The metal with the highest heat release for all applications is beryllium, whereas boron shows essentially equivalent performance for air-breathing applications. For rocket applications where Btu/lb pdts is important, lithium, boron, and aluminum are close to beryllium. As discussed later, however, because of inefficiencies in the combustion heat release as well as thermodynamic limitations, not all of these candidates are equivalent.

Method of use. The metallized additive can be used in either a liquid or solid propellant. When the pure metal is added to liquid fuels, an emulsifying or gelling agent is employed which maintains the particles in uniform suspension. Such gelling agents are added in small concentrations and so do not negate the beneficial effect of the fuel. Particle sizes in such suspensions are generally in the range 0.1–50 micrometers, and stable suspensions have been made which do not separate for long periods of time. When a metallic compound is employed, the compounding group is chosen so that the additive is soluble in the fuel. Thus, if gasoline is the fuel base, the suitable soluble boron compound is ethyldecaborane and, for aluminum, triethylaluminum may be employed. In both cases, the metallic additive contains a hydrocarbon constituent which provides solubility in the hydrocarbon base fuel.

When used in composite solid rocket propellants, the metal powder is usually mixed with the oxidizer and unpolymerized fuel, and the propellant is then processed in the usual way. If a compound of the metal additive is employed, it is convenient to dissolve it in the fuel, and then to process the fuel containing the metallic compound and oxidizer according to normal procedures. Homogeneous solid propellants also employ metallizing constituents, both as the pure metal and as a compound.

Compounds of interest. Early compounding of metallized propellants employed the free metal itself. As a result of extensive research in metalloorganic compounds, however, several classes of metallic compounds have been employed. Two major reasons for the use of such compounds are that the solubility of the metal in the fuel can be realized, resulting in

Heats of combustion of the lightweight metals compared with those of some other fuels

Fuel	Specific gravity	Product	Heat of combustion		
			Btu/lb pdts*	Btu/lb fuel*	Btu/in.^{3} fuel†
Lithium	0.53	Li_2O	8,570	18,500	360
Beryllium	1.85	BeO	10,450	29,100	1,950
Boron	2.35	B_2O_3	7,820	25,200	2,140
Carbon	2.00	CO_2	3,840	14,100	1,020
Magnesium	1.74	MgO	6,430	10,700	670
Aluminum	2.70	Al_2O_3	7,060	13,400	1,310
Gasoline	0.75	CO_2, H_2O	4,600	20,400	550
Hydrazine	1.00	N_2, H_2O	4,180	9,360	340
Hydrogen	0.07	H_2O	6,830	61,500	155

*1 Btu/lb = 2326 joules/kg.
†1 Btu/in.^{3} = 64.4 megajoules/m^3.

a homogeneous propellant, and performance higher than that for the pure metal can be obtained in some cases.

The major classes of metallic compounds of interest as high-performance propellants include the hydrides, amides, and hydrocarbons. Additional classes include mixtures either of two metals or of two chemical groups, such as an amine hydrocarbon. Some examples of the hydrides include LiH, BeH_2, B_2H_6, B_5H_9, $B_{10}H_{14}$, MgH_2, and AlH_3. These compounds form the more exotic class of metallics because they are capable of yielding the highest performances, exceeding those of the pure metals. The presence of hydrogen provides the low-molecular-weight gases needed for high thermodynamic performance. Examples of the propellant amides include $LiNH_2$ and $B(NH_2)_2BNHCH_3$. The largest class is the hydrocarbons, which includes LiC_2H_5, $Be(CH_3)_2$, $B(CH_3)_3$, and $Al(C_2H_5)_3$. The major advantage of this class is the potential solubility in conventional hydrocarbon fuels. *See* BORANE.

Operational limitations. Most metallic fuels are costly, and many, because of particle-size requirements or synthesis in a specific compound, are in limited supply. The combustion gases all produce smoky exhausts which may be objectionable in use. Engine development problems are also increased because of the appearance of smoke and deposits in the engine. These engine deposits can sometimes occur on the injector face, nozzle, and turbine blades of turbojet engines. Careful adjustment of operating conditions is required in such cases to minimize these difficulties with resultant added development costs.

Combustion losses. In rocket engines, where the flame temperatures are high, vaporization and dissociation are particularly pronounced. In the tabular listing of total heats of combustion, the heat value is given as measured in a laboratory calorimeter, the end product being the condensed oxide. At high flame temperatures, however, the oxides vaporize, producing a lower effective heat of combustion. In the case of boron oxide, B_2O_3, for example, the normal boiling point is about 2500 K (4000°F), whereas the flame temperature in most high-performing rocket engines exceeds 3000 K (4900°F). These effects are shown in the **illustration**, which compares

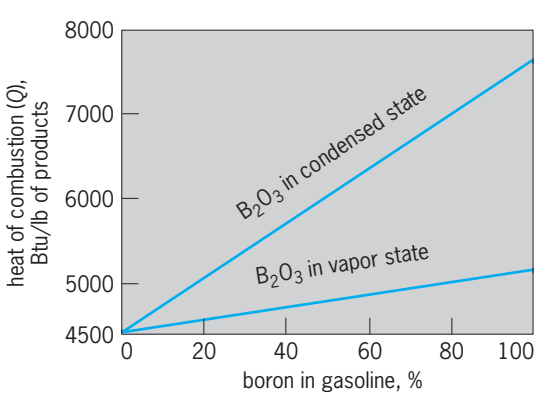

Total heat of combustion of gasoline-boron mixtures. 1 Btu/lb = 2326 joules/kg.

the total heat of combustion Q of gasoline-boron mixtures when the B_3O_3 is solid or liquid to the heat when it is a vapor. At still higher flame temperatures, in addition to vaporization, the oxides undergo endothermic dissociation such as shown in Eqs. (1), (2), and (3).

$$B_2O_3(g) = B_2O_2(g) + {}^1/_2O_2 - 108 \text{ kcal} \quad (1)$$

$$B_2O_2(g) = 2BO(g) - 126 \text{ kcal} \quad (2)$$

$$Al_2O_3(l) = Al_2O(g) + O_2 - 330 \text{ kcal} \quad (3)$$

The net result of these reactions is a further reduction in flame temperature to an extent that sensitively depends on the temperature itself. At temperatures below about 2500 K (4000°F), as in air-breathing engines operating below Mach 3, vaporization and dissociation for the oxides of beryllium, aluminum, boron, and magnesium occur hardly at all, and these limitations do not apply.

In actual engines, a certain fraction of the useful energy which is lost by vaporization and dissociation is recovered by condensation and recombination in nozzles. If thermodynamic equilibrium prevails and if the nozzle area ratio is infinite, all of this energy loss is recovered. In actual cases, however, because of the finite area ratio and finite reaction and condensation rates, only a part of the energy is recovered for propulsion. *See* ROCKET PROPULSION.

Two-phase flow losses. A second type of propulsive efficiency loss occurs as a result of the appearance of a condensed phase. Because the acceleration of the gases in the nozzle results from pressure-volume work, a reduction in gas content reduces the thermodynamic efficiency. The combustion gas behaves as if the molecular weight were increased. A second consideration involves the momentum exchange between condensed phase and gas. If the particles are large and lag appreciably behind the gas in velocity, impulse is lost. This effect can be appreciable when the percentage of condensed phase is high.

The net result of the thermodynamic limitations is to partially reduce in some cases the improvement in performance expected on the basis of heat of combustion alone. As additives, metals are most effective when mixed with the lower-performing propellants because the lower temperatures minimize the vaporization and dissociation losses. For this reason, the greatest use of metal additives has been in solid propellants, which develop 10–20% lower specific impulses than do liquid propellants. The result of the use of such additives has been to close the gap in performance between the solid and liquid propellants. The second attractive application is in air-breathing engines, where flame temperatures are relatively low.

David Altman

Metal carbonyl

A complex of a transition metal combined with carbon monoxide (CO). In a metal carbonyl, the CO groups form sigma bonds to the metal through

lone pairs of electrons on carbon; the metal, in turn, donates electrons to antibonding pi orbitals on CO. In so doing, the metal usually attains the electronic configuration of an inert gas (the 18-electron rule). Thus elemental chromium (Cr^0), with six valence electrons, combines with six molecules of CO, each donating two electrons, to afford chromium hexacarbonyl [$Cr(CO)_6$], in which Cr has 18 valence electrons and is isoelectronic to krypton (Kr). *See* CHEMICAL BONDING; COORDINATION COMPLEXES; VALENCE.

Metal carbonyls usually are electrically neutral and nonpolar, exhibiting many of the physical properties of organic compounds. Typically they are volatile solids or liquids, soluble in common organic solvents but insoluble in water. They are highly toxic.

Anionic and cationic species also are known. Metal carbonyls also may contain one (mononuclear) or more (polynuclear) metal atoms. Polynuclear metal carbonyls may exhibit covalent metal-metal bonds, and in them CO also may function as bridging or capping ligands, forming bonds to two or three metal atoms, respectively. Mononuclear metal carbonyls tend to exhibit geometries expected on the basis of minimization of interligand repulsions (valence shell electron-pair repulsion model); their polynuclear analogs often may be envisioned structurally as arising from the fusion of mononuclear polyhedra. Thus nickel tetracarbonyl [$Ni(CO)_4$], iron pentacarbonyl [$Fe(CO)_5$], and chromium hexacarbonyl [$Cr(CO)_6$] exhibit tetrahedral, trigonal bipyramidal, and octahedral geometries, respectively [structures (**1**)–(**3**)].

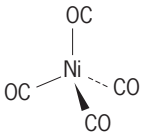

(**1**)

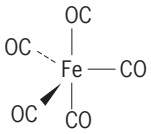

(**2**)

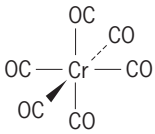

(**3**)

The polynuclear manganese carbonyl [$Mn_2(CO)_{10}$] has a structure obtained by joining two octahedra at a vertex [structure (**4**)], and the structure of diiron en-

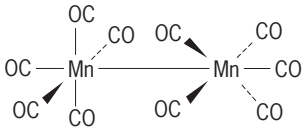

(**4**)

neacarbonyl [$Fe_2(CO)_9$] is viewed as two octahedra sharing three bridging carbonyls at a face [structure (**5**)]. The chemical formulas of representative metal

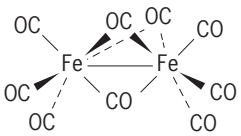

(**5**)

carbonyls are given below; the transition metals in brackets also form the preceding complex.

$V(CO)_6$; $V(CO)_6^-$ [Nb, Ta]; $Cr(CO)_6$ [Mo, W]; $Mn(CO)_6^+$ [Tc, Re]
$Cr_2(CO)_{10}^{2-}$ [Mo, W]; $Mn_2(CO)_{10}$ [Tc, Re]
$Mn(CO)_5^-$ [Tc, Re]; $Fe(CO)_5$ [Ru, Os]
$Fe_2(CO)_9$ [Ru, Os]; $Fe_3(CO)_{12}$ [Ru, Os]
$Os_5(CO)_{16}$; $Os_6(CO)_{18}$; $Os_7(CO)_{21}$; $Os_8(CO)_{23}$
$Co_2(CO)_8$ [Rh]; $Co_4(CO)_{12}$ [Rh]; $Rh_6(CO)_{16}$ [Ir]
$Fe(CO)_4^{2-}$; $Co(CO)_4^-$; $Ni(CO)_4$
$Ni_2(CO)_6^{2-}$; $Ni_3(CO)_8^{2-}$ $Ni_4(CO)_9^{2-}$; $Ni_5(CO)_9^{3-}$

[V = vanadium, Nb = niobium, Ta = tantalum, Cr = chromium, Mn = manganese, Tc = technetium, Re = rhenium, Fe = iron, Ru = ruthenium, Os = osmium, Co = cobalt, Rh = rhodium, Ir = iridium, Ni = nickel.]

The first metal carbonyl, $Ni(CO)_4$, was discovered by L. Mond in 1890; it is synthesized by the direct combination of metallic nickel and carbon monoxide under ambient conditions. Other metal carbonyls can be prepared directly from the metals and CO at elevated temperatures and pressures, or through the reaction of transition-metal salts with CO under reducing conditions, also at elevated temperatures and pressures [reactions (1)–(3)].

$$\underset{\text{Chromium(III) chloride}}{CrCl_3} + Al + 6CO \xrightarrow[\text{benzene}]{\text{aluminum chloride } (AlCl_3)} \underset{\text{Chromium hexacarbonyl}}{CrCl_3} + AlCl_3 \quad (1)$$

$$\underset{\text{Cobalt(III) carbonate}}{2CoCO_3} + 2H_2 + 8CO \xrightarrow[\text{octane}]{} \underset{\text{Discobalt(III) octacarbonyl}}{Co_2(CO)_8} + 2CO_2 + 2H_2O \quad (2)$$

$$\underset{\text{Manganese(II) acetate}}{Mn(CaH_3O_2)_2} + \underset{\text{Triethyl-aluminum}}{(C_2H_5)_3Al} + CO \xrightarrow[\text{di(isopropyl)-ether}]{} \underset{\text{Dimanganese decacarbonyl}}{Mn_2(CO)_{10}} \quad (3)$$

Polynuclear metal carbonyls can also be synthesized from other metal carbonyls by photolysis or thermolysis [reactions (4) and (5)].

$$\underset{\text{Iron pentacarbonyl}}{2Fe(CO)_5} \xrightarrow{h\nu,\ \text{acetic acid}} \underset{\text{Diiron ennacarbonyl}}{Fe_2(CO)_9} + CO \quad (4)$$

$$3Fe_2(CO)_9 \xrightarrow[\text{toluene}]{95^\circ C\ (203^\circ F)} \underset{\text{Triiron dodecacarbonyl}}{2Fe_3(CO)_{12}} + 3CO \quad (5)$$

Metal carbonyls react in solution under relatively mild conditions to form a variety of substitution products, with Lewis bases such as amines, phosphines,

sulfides, and nitric oxide and with organic molecules such as alkenes, alkynes, arenes, and other aromatic compounds. Reactions of metal carbonyls can also afford derivatives containing alkyl, acyl, allyl, halide, hydride, carbene, carbyne, and other substituents. They thus are the most important precursors to the synthesis of organotransition metal complexes. Their thermal decomposition can afford pure metals and CO.

Metal carbonyls and their derivatives also function as reagents in a wide range of organic synthetic applications and as precursors to homogeneous catalysts of considerable industrial importance. For example, in the oxo process, which can employ $Co_2(CO)_8$, several million metric tons of C_{n+1} aldehydes are produced annually from C_n alkenes, CO, and H_2. *See* CARBONYL; HOMOGENEOUS CATALYSIS; ORGANIC CHEMISTRY; ORGANIC SYNTHESIS. Gerard R. Dobson

Bibliography. J. P. Collman et al., *Principles and Applications of Organotransition Metal Chemistry*, 1987; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; R. A. Henderson, *The Mechanisms of Reactions at Transition Metal Sites*, 1994.

Metal casting

A metal-forming process whereby molten metal is poured into a cavity or mold and, when cooled, solidifies and takes on the characteristic shape of the mold. Casting offers several advantages over other methods of metal forming: it is adaptable to intricate shapes, to extremely large pieces, and to mass production; it can provide parts with uniform physical and mechanical properties throughout; and depending on the particular material being cast, the design of the part, and the quantity being produced, it can be more economical.

Categories

The two broad categories of metal-casting processes are ingot casting, which includes continuous casting, and casting to shape. Ingot castings are produced by pouring molten metal into a permanent or reusable mold. Following solidification, the ingots (bars, slabs, or billets) are processed mechanically into many new shapes. Casting to shape involves pouring molten metal into molds in which the cavity provides the final useful shape, followed by heat treatment and machining or welding, depending upon the specific application. Additionally, zero-gravity casting has been developed for processing in space stations.

Ingot casting. Ingot castings make up the majority of metal castings and fall into three categories: static cast ingots, semicontinuous or direct-chill cast ingots, and continuous cast ingots.

Static cast ingots. Static ingot casting simply involves pouring molten metal into a permanent mold (**Fig. 1**). After solidification, the ingot is withdrawn from the mold, and the mold can be reused. This method, while still important for certain grades of steel, is used only for a small amount of steel production. Since 1980, steel production in the United States has changed almost completely from static cast ingots to continuous casting. *See* STEEL MANUFACTURE.

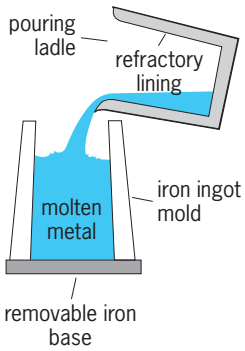

Fig. 1. Static ingot casting.

Semicontinuous cast ingots. A semicontinuous casting process is used in the aluminum industry to produce most of the cast alloys from which rod, sheet, strip, and plate are made. In this process, molten aluminum is transferred to partially fill a water-cooled permanent mold (**Fig. 2***a*) which has a movable base mounted on a long piston. After the solidification has progressed from the mold surface so that a solid "skin" is formed, the piston is moved down, and more metal is added to fill the reservoir (Fig. 2*b*). When the piston has moved its entire length, the process ends. Typically, the aluminum industry uses suitably lubricated metal molds. However, technological advances have allowed major aluminum alloy producers to replace the metal mold (at least in part) with an electromagnetic field so that molten metal touches the metal mold only briefly, thereby making a product with a much smoother finish. This method is also used to produce large brass ingots which are subsequently extruded into various shapes. *See* ALUMINUM; BRASS.

Continuous cast ingots. Continuous casting provides the major source of cast material for the steel and copper industries. In this process, molten metal is delivered to a permanent mold, and is cast in much the same way as in semicontinuous casting. However,

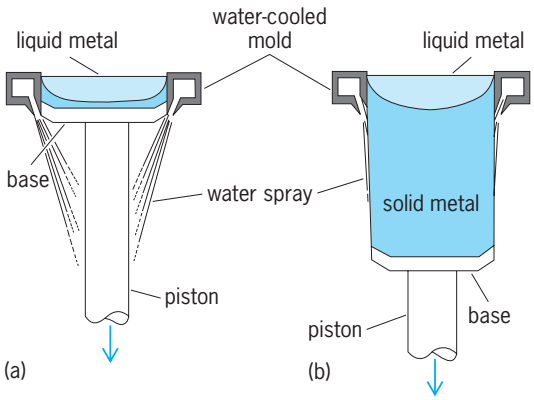

Fig. 2. Semicontinuous (direct-chill) casting. (*a*) Molten aluminum solidifies in a water-cooled mold with a movable base. (*b*) The piston is moved down so more molten metal can be poured into the reservoir.

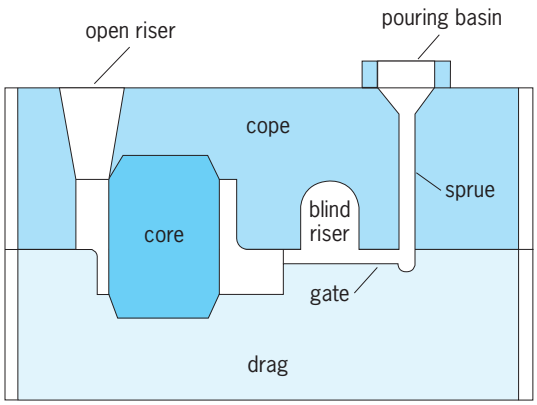

Fig. 3. Section through a sand mold showing the gating system and the risers.

instead of the process stopping after a certain length of time, the solidified ingot is continually sheared or cut into lengths and removed during casting. Thus the solidified billet, plate, or strip is removed as rapidly as it is cast. This method has many economic advantages over conventional casting techniques. As a result, all modern steel mills produce continuous cast products. *See* COPPER; STEEL.

Casting to shape. Casting to shape is generally classified according to the molding process, molding material, or method of feeding the mold. There are five basic shape casting processes: sand, permanent-mold, die, centrifugal, and investment.

Sand casting. This traditional method still produces the largest tonnage of cast-to-shape pieces. It uses a mixture of sand grains, water, clay, and other materials to make high-quality molds into which molten metal is poured. This "green sand" mixture is compacted around a pattern (wood, plaster, or metal), usually by machine, to 20–80% of its bulk density. The basic components of a sand mold, and other molds as well, are shown in **Fig. 3**. The two halves of the mold (the cope and drag) are closed over the cores that form the internal cavities, and the whole assembly is weighted or clamped to prevent floating of the cope when the metal is poured. Examples of sand castings include gray iron diesel engine blocks and disc brake rotors, as well as some very complex aluminum alloy castings for the aerospace industry. *See* CAST IRON.

Other casting processes which use sands as a basic component are the shell, carbon dioxide, ceramic molding, and plaster molding processes (see **table**), as well as a large number of chemically bonded sands which are increasingly important. In addition, both investment casting and evaporative pattern casting (lost foam) use sand in special ways.

Precision casting processes. This casting method produces near net shape products (or as close as possible), with little or no machining required after casting. In general, because of the required tight dimensional tolerances, these processes have molds which are very rigid, such as permanent molds or very hard expendable molds. Precision casting processes include investment casting, die casting, permanent molding processes, lost foam processes, and the Cosworth process. Die casting and permanent molding processes involve using permanent molds together with either gravity casting (permanent molds) or pressure-assisted casting (die casting). The Cosworth process uses disposable molds made with aggregate ceramic materials.

The Cosworth process, or precision sand process, is a method developed in the mid-1970s for producing aluminum castings to meet the highly specialized needs of Formula One racing cars. In this process,

Details of casting processes that utilize sands as a basic component

Process name	Pattern type	Molding aggregate	Type of bond	Relative cost
Green sand	Wood Plaster Metal	Sand Clay Water	Due to clay plasticity and compaction	Low
Shell	Heated metal	2.5–10% thermosetting resin	Polymerization of resin by pattern heat	Medium
CO_2	Wood Metal	Sand 2–6% sodium silicate	Inorganic bond by chemical reaction of CO_2 with silicate	Medium
Investment	Wax Expendable plastic	Sand Slurry Ceramic binder	Setting ceramic binder	High
V-Process*	Wood Metal	Sand enclosed in plastic with vacuum	Vacuum bond	Medium
No-bake	Wood Metal	Sand Organic compound	Organic bond created by reaction with catalyst	Medium
Full-mold*	Polystyrene coated with thin layer of ceramic	Sand	None	Medium

* Proprietary name.

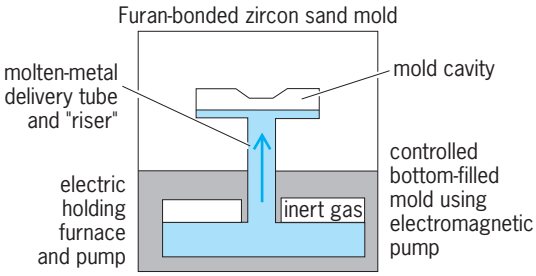

Fig. 4. Cosworth casting process.

molten metal is electromagnetically pumped from the bottom to the top of the mold in a controlled manner (**Fig. 4**). Benefits include improved levels of accuracy and integrity, very little porosity, excellent surface detail, excellent mechanical properties, virtually no cleaning and finishing costs, and minimum machining and high casting yield. The high casting yield results from the minimal gating system and the ability to control molten metal delivery. The delivery tube can act as a riser for the solidifying casting, and the molten metal feeding the shrinkage can, in principle, be returned to the reservoir and pump after the casting has solidified, thereby giving an obvious yield boost. Minimum porosity results from the controlled manner in which the molten metal enters the mold and the inert atmosphere over the molten metal.

Permanent-mold casting. Many high-quality castings are obtained by pouring molten metal into a mold made of cast iron, steel, or bronze. Semipermanent-mold materials, such as aluminum, silicon carbide, and graphite, may also be used. The mold cavity and the gating system are machined to the desired dimensions. The smooth surface from machining gives a good surface finish and dimensional accuracy to the casting. To increase mold life and to make ejection of the casting easier, the surface of the mold cavity is usually coated with carbon soot or a refractory slurry. These coatings also serve as heat barriers and help control the rate of cooling. This casting process is used for cast iron and nonferrous alloys, with advantages over sand casting such as smoother surface finishes, closer tolerances, and higher production rates. Aluminum alloy automotive pistons are an excellent example of permanent-mold castings. *See* BRONZE.

Die casting. Another variation of the permanent molding process is die casting. Molten metal is forced into a die cavity under pressures of 100–100,000 lb/in.^{2} (0.7–700 megapascals). Two basic types of die-casting machines are hot-chamber and cold-chamber. In the hot-chamber machine, a portion of the molten metal is forced into the cavity at pressures up to about 2000 lb/in.^{2} (14 MPa). The process is used for casting low-melting-point alloys such as lead, zinc, and tin alloys. *See* ALLOY; LEAD; TIN; ZINC.

In the cold-chamber process (**Fig. 5**), the molten metal is ladled into an injection cylinder and forced into the cavity under pressures about 10 times higher than those in the hot-chamber process. Higher-melting-point aluminum-, magnesium-, and copper-base alloys are commonly produced by this process. Most aluminum alloy castings are produced by this process. A good example of an aluminum alloy casting produced in large quantities is the marine engine block. Die casting has the advantages of high production rates, high quality and strength, surface finish on the order of 40–100-microinch (1.0–2.5-micrometer) root mean square, and close tolerances with thin sections. *See* MAGNESIUM.

Squeeze casting is a growing technology for producing shrinkage-free castings in permanent-mold dies by applying pressure to the casting as it solidifies, a structural condition that results in vastly improved fatigue properties. This technique successfully produces high-integrity aluminum alloy steering knuckles for automotive applications (**Fig. 6**).

Rheocasting is the casting of a solid and liquid mixture. In this process the alloy to be cast is melted and then allowed to cool until it is about 50% solid and 50% liquid. Vigorous stirring promotes the liquidlike properties of the mixture so that it can be injected in a die-casting operation. A major advantage of this type of casting is reduced die erosion due to the lower casting temperatures.

Centrifugal casting. Inertial forces of rotation distribute molten metal into the mold cavities during centrifugal casting, categorized into true centrifugal casting, semicentrifugal casting, and centrifuging. The first two processes (**Fig. 7**) produce hollow cylindrical shapes and parts with rotational symmetry, respectively. In the third process, the mold cavities are spun at a certain radius from the axis of rotation, and the resulting centrifugal force increases the pressure in the mold cavity.

The rotational speed in centrifugal casting is chosen to give 40–60 *g* acceleration. Dies may be made of forged steel or cast iron. Colloidal graphite is used on the dies to aid casting removal. Karl Rundman

Evaporative pattern casting. This is another casting process that uses sand as a basic component. It is unique in that the casting shape is not defined by molding the sand. Instead, a polystyrene foam pattern or replica is formed into the shape of the component to be cast, and determines the size and shape of the casting. In fact, the casting is a metal copy of

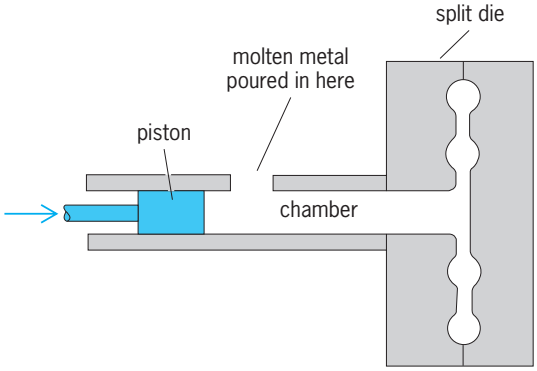

Fig. 5. Cold-chamber die casting machine.

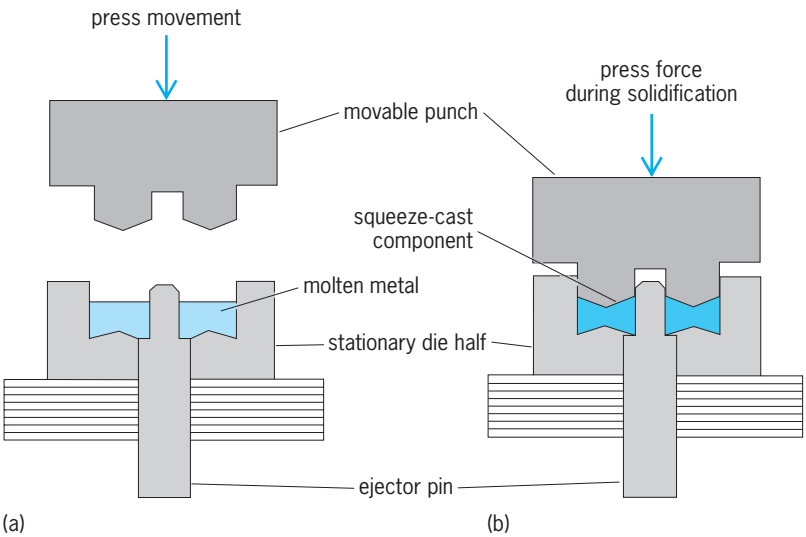

Fig. 6. Production of a squeeze casting. (*a*) Press open. (*b*) Press closed.

the polystyrene foam shape. The sand is used as a backing medium. After the polystyrene replica has been assembled and coated with a refractory wash to prevent burn-in, the coated replica is placed in a flask, and free-flowing sand is compacted around the replica to support the solidifying metal. The sand is mechanically locked in place by vibration, or the replica is immersed in a fluidized bed of sand. In either case, a coated polystyrene foam replica tightly surrounded by sand is now ready for metal pouring. Typically, a pouring cup is placed over the replica. The molding flask is either vented or equipped for vacuum pouring to remove the polystyrene off-gases. These off-gases, created as the polystyrene foam is thermally decomposed by heat from the molten metal, may ignite, depending on the heat of the metal. *See* POLYSTYRENE RESIN.

The evaporative casting process is compatible with bronze, steel, aluminum, and iron castings. In general, the microstructural characteristics of castings made by the evaporative process tend to be the same as those of castings made by conventional sand casting processes. Shrinkage rules for various aluminum alloys and iron types are also the same for evaporative casting and conventional sand molding. Thus, the presence of the polystyrene foam and its decomposition products does not alter casting dimensions and has little effect on microstructure or metal chemistry.

The major advantages of the evaporative casting process are reduced costs and greater flexibility. Savings result from lower investment costs and lower costs to make a casting. The lower investment costs come about because the equipment needed to make polystyrene foam replicas is less expensive than the equipment used to make conventional sand molds and cores. Also, the polystyrene molding equipment operates at higher speeds. Finally, the tooling for polystyrene foam replicas is less expensive and has a longer life than the tooling for conventional sand molds and cores. The cost per casting made by the evaporative process is lower because the process does not require cores and allows for easy reuse of the casting sand (no resins are added to the sand). Also, the casting made by this process is easier to clean than a conventional sand casting because there are no fins and parting lines that need to be ground down.

Evaporative casting offers process advantages in terms of both design flexibility and metals compatibility. The design flexibility results from the greater freedom in planning draft angles, parting lines, locator placement, and transitions from thin to thick sections. Evaporative casting can be used to make one-of-a-kind castings of investment-cast quality as well as high-volume production of automotive castings such as aluminum intake manifolds and iron connecting rods. S. A. Weiner

Investment casting. Investment casting is a process used to make precision castings in a variety of materials. The process has been known for thousands of years as the "lost wax" process. Precision patterns are produced in wax and coated (invested) in a ceramic slurry, which is allowed to dry before subsequent layers are added. Up to seven layers of investment material may be applied before the wax is carefully melted out, leaving a cavity. Next, the invested mold is fired at high temperature to produce a strong, rigid ceramic that is subsequently filled with molten metal. The precision castings are removed from the mold by carefully chipping away the ceramic shell. Perhaps the ultimate development of the investment technique was the controlled growth of investment-cast nickel alloy jet engine turbine blades, where

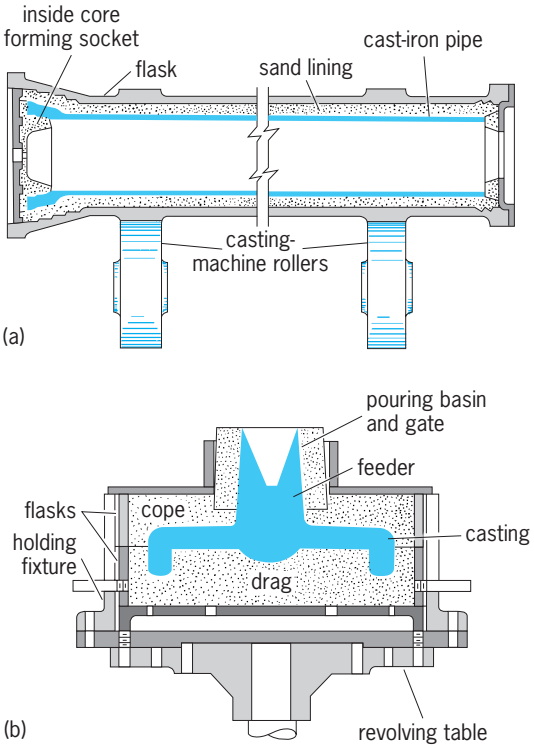

Fig. 7. Centrifugal casting. (*a*) Iron pipe cast in a centrifugal mold. (*b*) Flywheel cast in a semicentrifugal mold.

controlled solidification, starting at the base of the blade and continuing up through a tortuous "pigtail" configuration, allows only one crystal to emerge. The resulting single-crystal blade provides optimum properties (high-temperature creep resistance) in a critical application.

Zero-gravity casting. Zero-gravity casting has come of age with the advent of space stations. In the absence of gravity, the concept of melting and casting is drastically changed. The disadvantages of ground-based processing that appear to be diminished or eliminated by processing under weightless conditions are thermal and solutal convection and sedimentation; contamination of and contact with the container during processing; and hydrostatic pressure effects and deformation at high temperatures. There appears to be a potential for processing liquid metals at zero gravity; however, commercial production is expected to be limited initially to small quantities of high-cost materials whose solidification characteristics are greatly improved by the lack of gravity. Semiconductor materials are thought to be excellent candidates for initial development efforts. *See* SPACE PROCESSING.

Principles and Practice

Successful operation of any metal-casting process requires careful consideration of mold design and metallurgy.

Design considerations. In all types of ingot casting, successful production depends upon several factors. In more technologically complicated semicontinuous or continuous casting processes, the design of cooling systems, choice of lubrication, and rate of slab or billet movement play an important production role.

In the production of castings that are not subsequently machined, the construction of a pattern is an important step. Considerations of casting design, mold design, riser design, and gating practice play a role in the final pattern design from which molds are made. Modern foundries use computer-aided design techniques to simulate metal flow and solidification, incorporating fluid flow and heat transfer principles, before producing a single casting. Computer modeling allows designers and pattern makers to eliminate much trial and error in the casting process, reducing lead times significantly. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Gating system. The gating system (Fig. 3), which includes the pouring basin, downsprue, runners, and gates, must be designed to allow the liquid metal to enter the mold cavity with a minimum of turbulence, slag, and temperature loss. The application of Bernoulli's law, together with the criterion of conservation of mass, allows gas aspiration, turbulence, and slag entrapment to be kept to a minimum. All these factors are capable of being computer-modeled. *See* BERNOULLI'S THEOREM.

Riser design. A riser is a volume attached to the casting (Fig. 3) that provides metal to feed the shrinkage that occurs when liquid cools and solidifies. In fact, there are three contributions to shrinkage that must be accommodated: shrinkage of the liquid as it cools between the pouring temperature and the melting temperature; shrinkage as the liquid transforms to solid; and shrinkage of the solid portion on cooling during the time when some liquid remains. In order to supply liquid metal for the shrinking casting, the riser must also freeze after the cavity has frozen. The solidification time t_s is directly related to the heat transfer rate in the mold and to the volume-to-surface area ratio (V/A) of the casting. Thus, the foundry engineer can ensure the soundness of a casting in situations where the riser and casting have identical mold surroundings by making the V/A ratio of the riser larger than the V/A ratio of the casting. In large bulky castings, the riser may actually be larger than the casting, which increases the cost of production. Reducing this cost can sometimes be accomplished by insulating a smaller riser (changing the mold surroundings), thereby reducing the rate of heat transfer and prolonging the time that the riser is molten. Two types of risers are shown in Fig. 3, an open riser and a blind riser. Again, all these decisions can be made in simulations of the casting process before a drop of molten metal is produced.

Ease of pattern extraction. Sand molding patterns must be designed with a draft of 0.25–2° so that after the mold material has been compacted around the pattern, the pattern can be extracted without damage to the mold. Permanent molding dies must also be constructed so that after solidification the casting can be ejected without damaging either the casting or the die. The wax patterns of investment castings must be capable of being poured out of the cavity during heating so that no pockets of wax remain.

Machining considerations. For castings requiring subsequent machining, the patterns must be designed within certain dimensional tolerances so that enough stock is left for machining.

Metallurgical considerations. While design factors are important for producing sound castings with proper dimensions, factors such as the pouring temperature, alloy content, mode of solidification, gas evolution, and segregation of alloying elements control the final structure of the casting and therefore its mechanical and physical properties. Typically, pouring temperatures are selected within 100–300°F (60–170°C) of an alloy's melting point. Exceedingly high pouring temperatures can result in excessive mold metal reactions, producing numerous casting defects.

Most metals and alloys pass from the liquid state to the solid state in a predictable manner. The rate of solidification is a function of the rate at which heat is withdrawn by the mold and of the size of the casting. In continuous casting or semicontinuous casting, liquid is added at the same rate as solid is formed. Thus, a steady state is established. In the case of static cast ingots or cast-to-shape products, the entire liquid may be poured into the mold before any solidification begins. In either of these cases, however, the manner in which solidification occurs is quite similar for most commercial materials.

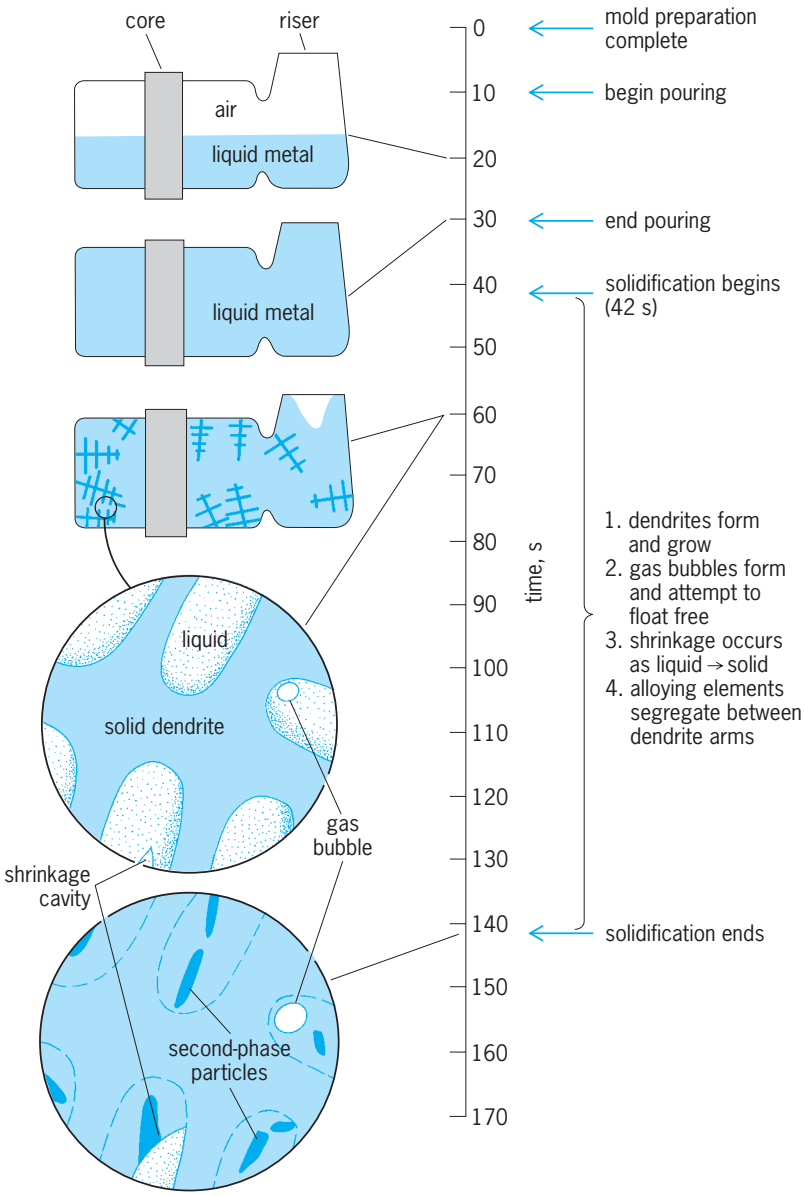

Fig. 8. Solidification and processing history for a cast-to-shape part.

The relationship between the processing history and the solidification is illustrated in **Fig. 8** for a cast-to-shape part. Solidification often begins with the formation of dendrites on the mold walls, which grow with time into the liquid and eventually consume the entire liquid volume. Dendrites are treelike or fernlike solids with primary arms similar to tree trunks and secondary arms resembling tree branches. Dendrites form in almost all commercial metals and are responsible to a large extent for their resulting as-cast properties. In pure metals, dendrites form if supercooling below the melting point occurs. Supercooling refers to the liquid existing below the temperature where it would normally change to solid. Supercooling can be accomplished by cooling the metal rapidly and by providing an extremely clean molten metal bath, thereby hindering nucleation of solid particles.

A particularly effective way of supercooling is through levitation melting, a technique whereby the metal is suspended in a magnetic field so that it touches no foreign object. Thus, no mold surface is available for dendrite growth, and there are no sharp temperature gradients within the metal. The levitation phenomenon is similar to the semicontinuous casting of aluminum alloy ingots, where the molten metal is levitated away from the mold surface. It has been suggested that levitation melting would be a natural method for casting in space. In the absence of gravity, large quantities of molten metals could conceivably be processed and handled by levitation techniques. In addition to the potential for casting in space, levitation melting provides exciting possibilities for producing high-purity materials free from gaseous contaminants and their products.

In metal alloys, solidification usually occurs over a range of temperatures. In alloys, dendrites grow in response to constitutional supercooling of the liquid below the liquidus temperature (the highest temperature at which solidification can occur). Usually it is desired to have the alloying elements distributed uniformly throughout the material after solidification has occurred. Invariably, dendritic growth will cause the alloying elements to be segregated and concentrated either between the dendrite arms (positive segregation) or within the dendrites (negative segregation). The segregation of alloying elements increases with dendrite growth and may ultimately lead to the formation of another solid phase when solidification is complete (Fig. 8). A good example of this type of structure is found in a leaded-copper alloy, containing mostly copper. Due to certain atomic forces between copper and lead atoms, nearly pure copper dendrites will form on solidification, leaving all of the lead as small spheres between the dendrites. This result is desirable in bearings, where the lead acts as a built-in lubricant. Many other commercial cast-to-shape materials have properties that depend on the presence of a second phase in the interdendritic volumes: cast iron structures are basically pure graphite particles (plates or spheres) between iron dendrites or dendrite arms; in aluminum silicon alloys, pure silicon particles (plates or rods) are located between aluminum dendrites. These phases are responsible in large part for the desirable properties of the material. However, in many continuous or semicontinuous cast alloys, particularly aluminum alloys and steels, the alloying elements that solidify between the dendrite arms are undesirable and must be redistributed to obtain the desired properties. This redistribution, or breaking up of the dendritic structure, is accomplished by various combinations of mechanical working and heat treatment.

In addition to creating segregation, the dendritic form of the solid provides a perfect place to trap gas bubbles. In addition, the very tightly spaced dendrite arms restrict liquid flow, often resulting in shrinkage porosity between dendrite arms (Fig. 8). Degassing operations, in which certain gases are bubbled through the melt, may be used to minimize the

gas porosity. As described by Sievert's law, the purging gas bubbles will attract the undesirable gases and remove them from the melt. In the case of shrinkage, the amount of microshrinkage may be reduced by increasing the cooling rate (sometimes called chilling). This action serves to make the dendrites stubbier, thereby providing freer access for metal flow to feed the shrinkage. Ideally, with the use of carefully regulated molding materials, solidification can be controlled so that dendritic growth occurs smoothly from the point farthest away from the riser directly to the riser, so that the last volume to solidify is the riser itself. This directional solidification concept is used in the production of many unique cast metal parts, particularly investment cast turbine blades.

Contraction of the casting during solidification can cause a defect known as hot tearing. Tensile stresses may be set up in the solidifying casting, which causes tears to open up between the dendrites. These types of defects are quite common in both ingot and cast-to-shape products. If the tears are not filled by liquid metal, then cracks will remain in the solidified casting. These defects can usually be avoided by proper mold design (or controlled cooling of continuous cast products) to minimize tensile-stress development. Residual stresses are elastic stresses that may form in solid materials due to uneven heating or cooling. These can cause warping or failure in castings following solidification unless steps are taken to ensure uniform cooling.

Other metal-casting process defects are (1) macrosegregation, caused by significant interdendritic fluid flow; (2) pinhole porosity, a smooth-walled defect elongated in a direction perpendicular to the mold wall and immediately below the casting skin; (3) metal penetration, resulting from liquid metal penetrating sand, and causing a metal-sand crust attached to the casting; (4) blows, cavities resulting from entrapped gases originating from cores; and (5) cold shuts, seams of oxide or slag resulting from low pouring temperatures.

Most or all of the above defects could be minimized or eliminated by processing in the zero-gravity, high-vacuum environment of space. Macrosegregation on Earth is driven primarily by gravity effects in which solute-rich material sinks in response to gravity differences. Thus the nagging inhomogeneities created in this way on Earth will be minimized or eliminated at zero gravity. Porosity and blowholes caused by dissolved gases will be reduced or eliminated by the vacuum melting conditions of space. Mold-metal reaction products can be minimized by containerless casting, use of levitation devices, and the like which eliminate mold-metal contact. Thus, processing at or near zero gravity has the potential for improving the quality of castings. However, the high cost of most such efforts limits their utility to casting small quantities of very specialized alloys.

Application

Almost all metals and alloys used by engineering specialists have at some point been in the molten state and cast. Metallurgists have in general lumped these materials into ferrous and nonferrous categories. Ferrous alloys, cast irons and steels, constitute the largest tonnage of cast metals. Aluminum-, copper-, zinc-, titanium-, cobalt-, and nickel-base alloys are also cast into many forms, but in much smaller quantity than cast iron and steel. Selection of a given material for a certain application will depend upon the physical and chemical properties desired, as well as cost, appearance, and other special requirements.

Aluminum alloy castings have advantages such as low density, resistance to corrosion, high electrical conductivity, ease of machining, and architectural and decorative uses. Magnesium alloy castings have the lowest density of all commercial casting alloys. Both aluminum and magnesium alloy castings are finding increased use in automotive applications. Copper alloy castings, although costly, have advantages such as corrosion resistance, high thermal and electrical conductivity, and wear properties suitable for antifriction-bearing materials. Steel castings have more uniform (isotropic) properties than the same component obtained by mechanical working. Cast irons constitute the largest tonnage of all cast-to-shape metals. Gray iron castings are commonly produced for their low cost, excellent machinability, good damping capacity, excellent thermal conductivity, good wear resistance, and uniformity. Nodular cast irons have significantly higher strength and ductility than gray irons, but are more costly and not as machinable. There is an intermediate grade of cast iron called CG (compacted graphite) iron which shares the best properties of both gray and nodular irons. *See* METAL, MECHANICAL PROPERTIES OF; METAL FORMING.

Karl Rundman

Bibliography. American Foundrymen's Society, *Basic Concepts of the Metalcasting Proccess*, 1987; American Society for Metals, *Casting Metals Hand-book*, vol. 15: *Casting*, 9th ed., 1988; M. C. Flemings, *Solidification Processing*, 1974; R. Johns, *Casting Design*, 1987; H. D. Merchant, D. E. Tyler, and E. H. Chia (eds.), *Continuous Casting of Non-Ferrous Metals and Alloys*, 1989.

Metal cluster compound

A compound in which two or more metal atoms are bonded to one another. Metal cluster compounds bridge the gap between the solid-state chemistry of the metals—or their lower-valent oxides, chalcogenides, and related salts—and the complexes of the metals in which each metal ion is completely surrounded by and bonded to a set of ligands or ions. The latter group comprises the classical coordination chemistry of metal ions.

Interest in metal cluster compounds arises from unique features of their chemistry: (1) Cluster compounds provide models for studying fundamental reactions on surfaces. (2) There is a hope that cluster compounds may provide entry to new classes of catalysts that may be tailored to specific syntheses and may thus be more selective than existing processes.

(3) The nature of the bonding in cluster compounds is an area wherein experiment and theory are continuously challenging each other. (4) The systematic synthesis of mixed metal clusters may provide for the development of new types of supported catalysts (the discrete clusters are deposited on supports such as alumina, silica, or zeolites).

Metal clusters can be classified into three categories: those whose only existence is in the solid state; those that exist both in the solid state and in a solution in which each cluster unit of metal atoms is surrounded by a set of ligands; and the so-called naked metal clusters, which carry a charge but are otherwise surrounded only by solvent molecules when in solution. Often, naked clusters may have a form different from that exhibited in the solid state. For each of these classes the metal atoms may be the same or the clusters may contain more than one type of metallic element: they are known respectively as homometallic or heterometallic clusters.

The term "bonded" is the key in understanding the nature of metallic clusters. The bonding may be localized, where each metal–metal interaction can be counted as an electron pair, or the bonding may be a type in which the electrons are delocalized over the metallic (M_x) subunit so that it is impossible to assign precise electron-pair interactions. The strength of the metal–metal interactions can range from strong metal–metal bonding, as in the metallic state of the elements, to weak interactions in which the metal atoms are coupled only loosely in an electronic sense. The latter is often the case in metal-containing enzymes, where more than one metal atom is required for a specific task. *See* CHEMICAL BONDING; DELOCALIZATION.

Metallic clusters in bioinorganic molecules. In a number of iron-sulfur proteins, such as the ferrodoxins, there are two, three, or four iron (Fe) atoms linked by sulfido sulfur (S) atoms (inorganic sulfurs) and also bonded to the protein by the sulfur atoms of cysteine residues. These are redox active centers, and they are involved in a variety of electron transfer reactions. Other bioinorganic molecules that utilize two or more electronically coupled metal ions are found in hemerythrins (Fe_2 units used for oxygen transport in marine worms), photosystem II [manganese (Mn_4) units used to catalyze the splitting of water (H_2O) to hydrogen (H_2) and oxygen (O_2)], and hemocyanins [copper (Cu_2) subunits involved in the oxygen transport]. *See* ELECTRON-TRANSFER REACTION; OXIDATION-REDUCTION.

While the metal–metal interactions in all known bioinorganic molecules are weak, they are strong enough to result in important cooperative effects. In some cases the detailed nature of both structure and function remains uncertain; an example is nitrogenase, an enzyme used to reduce nitrogen (N_2) to ammonia (NH_3) in a variety of plants. In addition to containing a four-iron/four-sulfur cluster component, nitrogenase contains an iron-molybdenum cluster subunit, the so-called iron–molybdenum cofactor (FeMoco), having two molybdenum (Mo) atoms and six or eight Fe atoms and approximately six inorganic sulfur atoms. The unique way in which this inorganic cluster unit functions within the large biomolecule remains a subject of intense research activity. *See* BIOINORGANIC CHEMISTRY; ENZYME; PROTEIN.

Cluster units in the solid state. A wide variety of lower-valent (subvalent) metal halides, oxides, and chalcogenides contain cluster subunits of metal atoms. These are well exemplified by the so-called Chevrel phases that have the formula $M_xMo_6X_8$, where M = lead (Pb), tin (Sn), Ln (lanthanon; a lanthanide element), potassium (K), calcium (Ca), Cu, lithium (Li), and others (in all, over 40 elements) and X is usually sulfur, selenium (Se), or tellurium (Te) and occasionally bromide (Br^-) or iodide (I^-). Research in this field was driven by the discovery that these cluster-containing compounds become superconducting at low temperatures. These were among the first group of nonmetals and nonmetal alloys to exhibit superconductivity. The basic building block for most of these solid-state clusters is an octahedron of metal atoms inscribed in a cubic arrangement of ligands, either as $M_6(\mu_3\text{-}L)_8$, where the positions L define the corners of a cube and cap the 8 faces of the octahedron, or the $M_6(\mu_3\text{-}L)_{12}$ structure, where each of the 12 edges of the octahedron of metal atoms contains a bridging group L. In the latter case the ligands occupy center positions of the 12 edges of a cube. The three-dimensional lattice structure is built upon these M_6 cluster units through the sharing of common bridging atoms or by the fusing of the M_6 octahedra so that apices, edges, or faces are shared (see **illus.**). Solid-state clusters are also found based on rhomboidal M_4 and triangular M_3 subunits, and these may be viewed as fragments of the M_6 octahedron. In some instances, interstitial atoms such as carbon (C), boron (B), and nitrogen (N) are found inside the M_6 unit, and these atoms provide additional so-called molecular glue by supplying additional electrons to bind the metal atoms together. Similarly, hydride (H^-), carbide (C^{4-} or C_2^{x-}), and nitride (N^{3-}) are commonly found in molecular clusters. *See* CHEVREL PHASES; CRYSTAL STRUCTURE; SUPERCONDUCTIVITY.

Molecular clusters. These compounds may be neutral or charged, but they are characterized by a grouping of metal atoms surrounded by a set of ligands such as carbon monoxide (CO), tertiary phosphines (R_3P; R = functional group), and alkoxides (organic compounds in which the H of an OH group has been replaced by a metal). Some examples include $Au_{55}[P(C_6H_5)_3]_{12}Cl_{12}$; $[Os_{10}(C)(CO)_{24}]^{2-}$; $[Pt_3(CO)_6]_n^{2-}$, where n = 2, 3, and 5; and $W_4(OC_2H_5)_{16}$. Interest in these clusters has centered on understanding and predicting the geometries and bonding and then developing their reaction chemistry. The clusters that possess many metal atoms bridge the gap between metals and molecules, since the ligands bonded to the clusters may resemble the molecules and fragments of molecules bound to metal surfaces. For example, in an Os_3 cluster it has been possible to monitor the interconversion of hydride (H^-) and methyl (CH_3) ligands with dihydrido-methylene

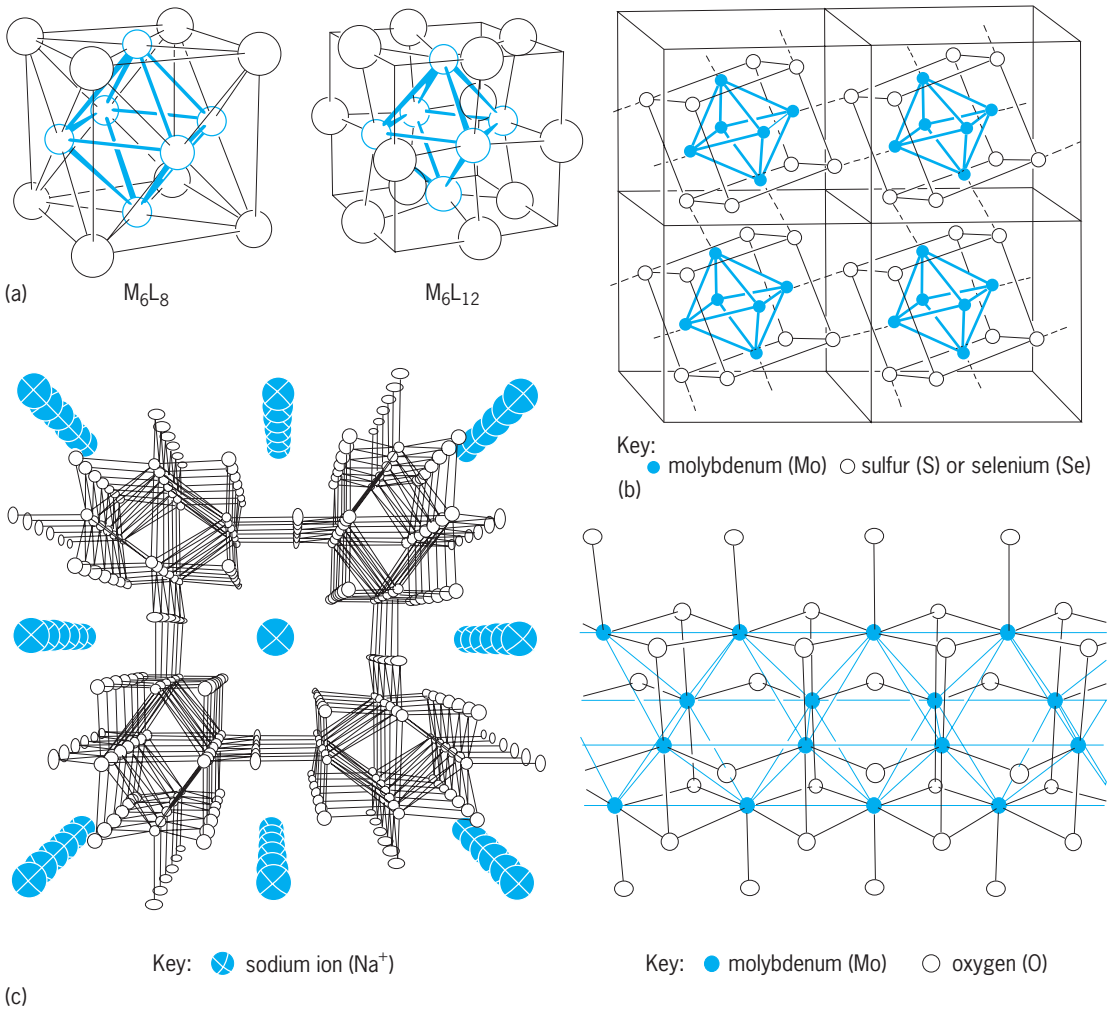

Structural representations of some metal clusters. (*a*) Two common octahedral M_6 cluster units. On the left, 8 ligand atoms lie at the corners of a cube and cap the 8 faces of the octahedron; on the right, 12 ligand atoms occupy the centers of the 12 edges of the cube-octahedron. (*b*) View of the M_6L_8 cluster units in the binary molybdenum compounds Mo_6X_8, where X = sulfur or selenium. (*c*) Two views of the structure of sodium molybdate ($NaMo_4O_6$), which contains Mo_6 octahedra-sharing edges. The view on the left is looking down the crystallographic *c* axis and shows the ordering of the sodium (Na^+) ions in channels between the extended chains of Mo_6 cluster units shown on the right.

[$(H)_2$—CH_2] and trihydrido-methylidyne [$(H)_3$—CH]. These provide models for the stepwise activation of methane (CH_4) by a grouping of metal atoms in which successive C-H bonds are substituted by metal-carbon bonds. It takes more than one metal atom to bring about a specific chemical reaction, which is referred to as the ensemble effect. Collectively, a grouping of metal atoms may act as an electron source or sink and provide a template for substrate activation. For example, even carbon monoxide, which has the highest bond dissociation energy (256 $K \cdot cal \cdot mol^{-1}$), can be cleaved in reactions with certain M_4 clusters. The reactivity and dynamic properties of species bound to molecular clusters provide model studies for many catalytic reactions involving metal surfaces or metal oxides. *See* HETEROGENEOUS CATALYSIS; HOMOGENEOUS CATALYSIS; METAL HYDRIDES.

Naked metal clusters. These are known in heterogeneous phases of main-group metals, transition metals, and mixed metals. Many supported metals (deposited on alumina or silica), for example, rhodium/tin and platinum/tin, are used in the petroleum industry as reforming catalysts. Cationic naked metal clusters such as the bismuth cluster Bi_5^{3+} exist in molten salts, and many others have been made and studied either by low-temperature matrix isolation or in the gas phase. By contrast, anionic naked metal clusters are obtained when certain alloys, for example, sodium/tin, sodium/lead, and sodium/antimony, are extracted into solution with liquid ammonia or ethylenediamine; examples include the lead cluster Pb_5^{2-}, Sn_9^{2-}, and the antimony cluster Sb_7^{3-}. Naked metal cluster anions of metals of the main groups in the periodic table adopt polyhedral structures reminiscent of the boron hydrides. In solution these clusters exhibit strong metal–metal interactions and weak metal–solvent interactions. These species may be useful for the preparation of thin films of metal atoms and as novel reducing agents. *See* CRACKING; MATRIX ISOLATION; PERIODIC TABLE.

Characterization. Generally, the ultimate characterization of a metal cluster rests upon an x-ray structural determination. By this method the detailed

arrangement of metal atoms, the nature of metal–ligand interactions, and the precise stoichiometry can be seen in one picture. However, it is often important to determine that what is seen in the solid state is also present in solution, particularly for molecular clusters that have solubility in various solvents. A variety of complementary techniques can be used for this purpose; nuclear magnetic resonance and infrared and Raman spectroscopy are most important for correlating solid-state and solution properties. In many instances, other techniques may provide specialized information concerning the oxidation states of the metal atoms (x-ray photoelectron spectroscopy, Möossbauer spectroscopy) or the electronic structure within a grouping of M-M molecular orbitals (electron spin resonance and ultraviolet or visible light spectroscopy). Extended x-ray absorption fine structure (EXAFS) has also proved useful in obtaining local information of metal-ligand bondings where single-crystal x-ray studies have not been possible, for example, in probing the nature of nitrogenase; but there are always limitations to the details provided by such less definitive spectroscopic techniques. Neutron diffraction has proved useful for the structural determination of both finely divided but microcrystalline solids and in the absolute location of light atoms, particularly hydrides, in molecular clusters. *See* COORDINATION CHEMISTRY; COORDINATION COMPLEXES; EXTENDED X-RAY ABSORPTION FINE STRUCTURE (EXAFS); INFRARED SPECTROSCOPY; MÖSSBAUER EFFECT; NUCLEAR MAGNETIC RESONANCE (NMR); RAMAN EFFECT; SOLID-STATE CHEMISTRY; SPECTROSCOPY; SURFACE AND INTERFACIAL CHEMISTRY.

Malcolm H. Chisholm

Bibliography. M. H. Chisholm and F. A. Cotton, Bonds between metal atoms: A new mode of transition metal chemistry, *Chem. Eng. News*, 60:40–54, June 28, 1982; F. A. Cotton and R. A. Walton, *Multiple Bonds Between Metal Atoms*, 1984; L. J. De Jongh (ed.), *Physics and Chemistry of Cluster Compounds: Model Systems for Small Metal Particles*, 1994; J. Lewis, Metal clusters revisited, *Chemistry in Britain*, 24:795–800, 1988.

Metal coatings

Thin films of material bonded to metals in order to add specific surface properties, such as corrosion or oxidation resistance, color, attractive appearance, wear resistance, optical properties, electrical resistance, or thermal protection. In all cases proper surface preparation is essential to effective bonding between coating and basis metal, so that coated metals can function as duplex materials. This article discusses various methods of applying either metallic coatings (see **table**) or nonmetallic coatings, such as vitreous enamel and ceramics, and the conversion of surfaces to suitable reaction-product coatings. For other methods for the protection of metal surfaces *see* CLADDING; ELECTROLESS PLATING; ELECTROPLATING OF METALS; LACQUER; PAINT AND COATINGS.

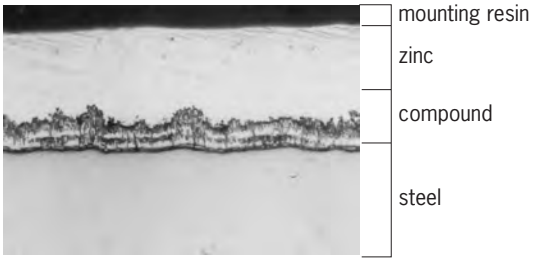

Fig. 1. Cross section of galvanized steel.

Hot-dipped coatings. Low-melting metals provide inexpensive protection to the surfaces of a variety of steel articles. To form hot-dipped coatings, thoroughly cleaned work is immersed in a molten bath of the coating metal. The coating consists of a thin alloy layer together with relatively pure coating metal that adheres to the work as it is withdrawn from the bath (**Fig. 1**). Sheet, strip, and wire are processed on a continuous basis at speeds of several hundred feet per minute. On the other hand, hardware and hollow ware are handled individually or in batches.

Coating weights for zinc coatings are expressed as ounces of coating per square foot (1 oz/ft^2 is equivalent to 0.305 kg/m^2 or 43-micrometer thickness on hardware or wire, and to one-half this thickness on sheet). Hot-dipped tin and lead-tin (Terne) coatings are expressed as pounds per base box. A base box is 31,360 in.^{2} (20.23 m^2); thus 1 lb (0.4536 kg) per base box is equivalent to 1.50 μm of tin.

Galvanized steel, that is, steel hot-dipped in zinc, is used for roofing, structural shapes, hardware, sheet, strip, and wire products. Coatings normally range up to 3 oz/ft^2 (0.9 kg/m^2) for outdoor service. Hot-dipped tinplate is now largely supplanted by electroplated tinplate for tin cans. Terne plate, with coatings up to 20 μm, is used for roofing, chemical cabinets, and gasoline tanks. Hot-dipped aluminum-coated (aluminized) steel with coatings up to 10 μm thick are used for oil refinery equipment and furnace and appliance parts where protection at temperatures up to 1000°F (538°C) is required.

Sprayed coatings. A particular advantage of sprayed coatings is that they can be applied with portable equipment. The technique permits the coating of assembled steel structures to obtain corrosion resistance, the building up of worn machine parts for rejuvenation, and the application of highly refractory coatings with melting points in excess of 3000°F (1650°C).

Nearly any metal or refractory compound can be applied by spraying. Coating material in the form of wire or powder is fed through a specially designed gun, where it is melted and subjected to a high-velocity gas blast that propels the atomized particles against the surface to be coated (**Fig. 2**). The surface is initially blasted with sharp abrasive to provide a roughened surface to which the particles can adhere. Coatings such as zinc and aluminum are applied with a gun which provides heat by burning acetylene, propane, or hydrogen in oxygen. A compressed air blast atomizes and propels the coating metals

Commonly used methods for applying metallic coatings

Coating metal	Hot dip	Electroplate	Spray	Cementation	Vapor deposition	Cladding	Immersion
Zinc	X	X	X	X	—	—	—
Aluminum	X	—	X	X	X	X	—
Tin	X	X	—	—	—	—	X
Nickel	—	X	X	—	X	X	X
Chromium	—	X	—	X	X	—	—
Stainless steel	—	—	X	—	—	X	—
Cadmium	—	X	—	—	X	—	—
Copper	—	X	—	—	—	X	X
Lead	X	X	—	—	—	X	—
Gold	—	X	—	—	X	X	X
Silver	—	X	—	—	X	X	X
Platinum metals	—	X	—	—	X	X	X
Refractory metals	—	X	X	—	—	X	—

onto the surface. Highly refractory coating materials, such as oxides, carbides, and nitrides, can be applied by plasma-arc spraying (**Fig. 3**). In this process temperatures of 20,000°F (12,000°C) or more may be produced by partially ionizing a gas (nitrogen or argon) in an electric arc and passing the gas through a small orifice to produce a jet of hot gas moving at high velocity. Another variation for applying refractory coatings is detonation-flame plating. In this process a mixture of oxygen and gas-suspended fine particles are fired four times per second by a timed spark.

Sprayed zinc or aluminum coatings up to 250 μm are used to protect towers, tanks, and bridges. Such coatings are normally sealed with an organic resin to enhance protection.

Sprayed refractory coatings have been developed for high temperatures experienced in aerospace applications. They are also used for wear resistance, heat resistance, and electrical insulation.

Cementation coatings. These are surface alloys formed by diffusion of the coating metal into the base metal, producing little dimensional change. Parts are heated in contact with powdered coating material that diffuses into the surface to form an alloy coating, whose thickness depends on the time and the temperature of treatment. A zinc alloy coating of 25 μm is formed on steel in 2–3 h at 700°F (375°C). A chromium alloy (chromized) coating of 100 μm is formed in 1 h at 1830°F (1000°C).

Chromized coatings on steel protect aircraft parts and combustion equipment. Sherardized (zinc-iron alloy) coatings are used in threaded parts and castings. Calorized (aluminum-iron alloy) coatings protect chemical equipment and furnace parts. Diffusion coatings are used to provide oxidation resistance to refractory metals, such as molybdenum and tungsten, in aerospace applications where reentry temperatures may exceed 3000°F (1650°C). In addition to the pack process described above, such coatings may be applied in a fluidized bed. In forming disilicide coatings on molybdenum, the bed consists of silicon particles suspended in a stream of heated argon flowing at 0.5 ft/s (0.15 m/s), to which a small amount of iodine is added. The hot gases react with the silicon to form SiI_4, which in turn reacts with the molybdenum to form $MoSi_2$.

Vapor deposition. A thin specular coating is formed on metals, plastics, paper, glass, and even fabrics. Coatings form by condensation of metal vapor originating from molten metal, from high-voltage (500–2000 V) discharge between electrodes (cathode sputtering), or from chemical means such as hydrogen reduction or thermal decomposition (gas plating) of metal halides. Vacuums up to 10^{-6} mmHg (10^{-4} pascal) often are required.

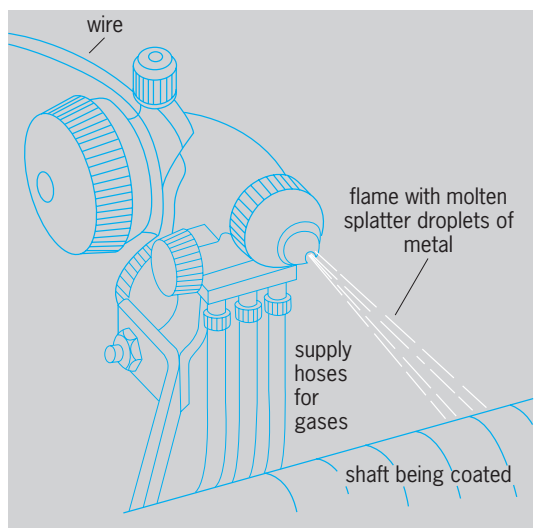

Fig. 2. Schooping gun with the coating material furnished in the form of wire.

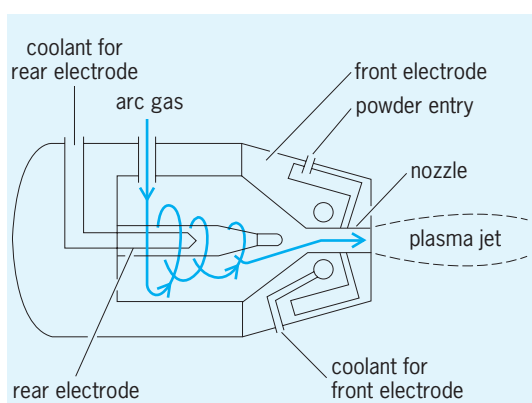

Fig. 3. Cross-sectional diagram of plasma-spray jet.

Aluminum coatings of 0.125 μm are formed on zinc, steel, costume jewelry, plastics, and optical reflectors. Chemical methods can form relatively thick coatings, up to 250 μm. The 200-in. (5-m) mirror for the Mount Palomar telescope was prepared by vapor coating with aluminum in a vacuum chamber 19 ft (6 m) in diameter and 7 ft (2 m) high.

Immersion coatings. Either by direct chemical displacement or for thicker coatings by chemical reduction (electroless coating), metal ions plate out of solution on to the workpiece.

Tin coatings are displaced onto brass and steel notions and on aluminum-alloy pistons as an aid during the breaking-in period. Displacement nickel coatings of 1.25 μm are formed on steel articles. Electroless nickel, involving the reduction of a nickel salt to metallic nickel (actually a nickel-phosphorus alloy), permits the formation of relatively thick uniform coatings up to 250 μm on parts with recessed or hidden surfaces difficult to reach by electroplating.

Vitreous enamel coatings. Glassy but noncrystalline coatings for attractive durable service in chemical, atmospheric, or moderately high-temperature environments are provided by enamel or porcelain coating. In wet enameling a slip is prepared of a water suspension of crushed glass flux, suspending agent, refractory compound, and coloring agents or opacifiers. The slip is applied by dipping or flow coating; it is then fired at a temperature at which it fuses into a continuous vitreous coating. For multiple coats the first or ground coat contains an oxide of cobalt, nickel, or molybdenum to promote adherence.

Dry enameling is used for castings, such as bathtubs. The casting is heated to a high temperature, and then dry enamel powder is sprinkled over the surface, where it fuses.

Firing temperatures for conventional enameling of iron or steel ranges up to 1600°F (870°C). Low-temperature enamels have been developed, permitting the enameling of aluminum and magnesium.

Coatings of 75–500 μm are used for kitchenware, bathroom fixtures, highway signs, and water heaters. Vitreous coatings with crystalline refractory additives can protect stainless steel equipment at temperatures up to 1740°F (950°C).

Ceramic coatings. Essentially crystalline, ceramic coatings are used for high-temperature protection above 2000°F (1100°C). The coatings may be formed by spraying refractory materials such as aluminum oxide or zirconium oxide, or by the cementation processes for coatings of intermetallic compounds such as molybdenum disilicide. Cermets are intimate mixtures of ceramic and metal, such as zirconium boride particles, dispersed throughout an electroplated coating of chromium. *See* CERAMICS; CERMET.

Surface-conversion coatings. An insulating barrier of low solubility is formed on steel, zinc, aluminum, or magnesium without electric current. The article to be coated is either immersed in or sprayed with an aqueous solution, which converts the surface into a phosphate, an oxide, or a chromate. Modern solutions react so rapidly that sheet and strip materials can be treated on continuous lines.

Phosphate coatings, equivalent to 100–400 mg/ft^2 (1–4 g/m^2), are applied to bare or galvanized steel and to zinc-base die castings as preparation for painting. The coating enhances paint adhesion and prevents underfilm corrosion. Phosphate coatings, containing up to 4000 mg/ft^2 or 40 g/m^2 (lubricated), serve as an aid in deep-drawing steel and in other friction-producing processes or applications. Iridescent chromate coatings on zinc-coated steel improve appearance and reduce zinc corrosion. Chromate, phosphate, and oxide coatings on aluminum or magnesium are used to prepare the surface before painting.

Anodic coatings. Coatings of protective oxide may be formed on aluminum or magnesium by making them the anode in an electrolytic cell. Anodized coatings on aluminum up to 75 μm thick are formed in sulfuric acid to form a porous oxide that may be sealed in boiling water or steam to provide a clear, abrasion-resistant, protective coating. If permanent color is required, the coating is impregnated with a dye prior to sealing. Such coatings are used widely on aluminum furniture, automobile trim, and architectural shapes. Thin, nonporous, electrically resistant coatings are formed on aluminum in a boric acid bath in the production of electrolytic capacitors. Anodized coatings on magnesium are thicker (up to 75 μm) and harder than those formed by chemical conversion. Anodic coatings 7.5 μm thick are often used as a paint base. William W. Bradley

Powder coatings. The term powder coating refers to a process whereby organic polymers such as acrylic, polyester, and epoxies are applied to substrates for protection and beautification. It is essentially an industrial painting process which uses a powdered version (25–50-μm particle size) of the resin rather than the solvent solution represented by industrial baking enamels.

In the widest type of commercial usage, these powders are applied to electrically grounded substrates by means of an electrostatic spray gun. In these guns, high-voltage (60–100-kV) low-amperage charges are applied to the powders. The powder particles are attracted to and adhere to the substrate until it can be transported to an oven, where the powder particles melt, coalesce, flow, and form a smooth coating. Typical baking temperatures are 375°F (90°C) for 20 min. *See* ELECTROSTATICS.

Coatings can be applied at thicknesses ranging from 25 to 250 μm, depending upon end-use requirements. These thicknesses can be achieved in a single application. While electrostatic spray is the principal method of commercial application, other techniques exist, including fluidized-bed coating, where a preheated part is immersed into an aerated bed of powder particles, and electrostatic fluidized bed, where a substrate is passed through an "electrified" cloud of the powder.

Powder coating uses no organic solvent, and the oversprayed material can be recaptured and reused. Thus, there are several notable advantages over conventional industrial solvent–based painting:

(1) There is no hydrocarbon baking by-products, and the process is environmentally acceptable. (2) Material conservation and economics are optimized because of the recycling of oversprayed material. (3) Powders are easy to apply and readily lend themselves to automation.

A wide variety of parts are finished with powder coatings. Outdoor lawn and patio furniture coated in this process display good weathering and abuse resistance. Powder-coated electrical transformers are insulated electrically and provided with corrosion protection. Powder coatings have also been developed for finishing major appliances and for automotive coatings. Also, laundry tops and lids have displayed excellent detergent resistance where organic powder coatings have replaced porcelain finishes. *See* CORROSION; SURFACE COATING; SURFACE HARDENING OF STEEL.

R. F. Farrell

Bibliography. A. I. Andrews, *Porcelain Enamel*, 2d ed., 1961; A. I. Borisenko, *High Temperature Protective Coatings*, 1987; R. M. Burns and W. W. Bradley, *Protective Coatings for Metals*, 3d ed., 1975; U. R. Evans, *An Introduction to Metallic Corrosion*, 1981; J. Huminik (ed.), *High Temperature Inorganic Coatings*, 1963; A. T. Kuhn, *Techniques in Electrochemistry, Corrosion, and Metal Finishing: A Handbook*, 1987; *Materials in Design Engineering*, Metals Selector Issue, vol. 80, no. 4, September 1974; C. F. Powell, J. H. Oxley, and J. M. Blocher (eds.), *Vapor Deposition*, 1966; H. H. Uhlig and R. W. Revie, *Corrosion Handbook*, 2d ed., 2000.

Metal forming

Manufacturing processes by which parts or components are fabricated from metal stock. In the specific technical sense, metal forming involves changing the shape of a piece of metal. In general terms, however, it may be classified roughly into five categories: mechanical working, such as forging, extrusion, rolling, drawing, and various sheet-forming processes; casting; powder and fiber metal forming; electroforming; and joining processes. The selection of a process, or combination of processes, requires a knowledge of all possible methods of producing the part if a serviceable part is to be produced at the lowest overall cost.

Influence on service behavior. The service behavior of a part depends on the interrelation of the design, material, processing of the material (such as manufacturing process, heat treatment, and surface treatment), and operating and environmental conditions (such as temperature of operation, loading condition, and corrosive atmosphere). The metal-forming process is a factor in satisfactory service performance in that it affects the microstructure of the metal and its surface finish, it may introduce large residual stresses into the part, and it may affect the final design of the part which, in turn, influences its service behavior.

Process selection. The more important factors to be considered in choosing the optimum process (or combination of processes) include: type of material, metallurgical structure effect inherent in the process, size of the part, shape or complexity of the part, tolerances or finish required, quantity to be produced, cost, and production factors such as availability of equipment, rate of production, and time required to initiate production.

Material. The particular metal or alloy specified for a part is of major importance in the selection of the forming process. Some aluminum or copper alloys may be fabricated by practically any of the manufacturing processes; other alloys may be brittle under cold-working conditions but may be hot-worked. Highly refractory materials, such as tungsten and tungsten carbide, are not suitable for casting and must be fabricated by powder metallurgy methods. This process is also used for making porous metal products, or parts requiring combinations of two materials (not an alloy of the two, however). Alloys that are extremely hard, and therefore unsuitable for machining operations, can be precision cast and then ground if extremely close tolerances are required. Higher-melting-point alloys such as steel can be fabricated by most of the major classes of processes but are not suitable for all of the individual processes within a major class. Where a specific material must be used for the part, the choice of the optimum fabricating process may be definitely limited. The formability of a material may be predicted either from the reduction of area or the percent elongation in a tension test; the higher these values, the better is the formability. Other factors that influence formability are rate of deformation, temperature, type of loading, environment, impurities in the metal, and surface condition of the original stock. *See* ALLOY; ALLOY STRUCTURES; METAL, MECHANICAL PROPERTIES OF; METALLURGY; STEEL.

Metallurgical structure. Each manufacturing process has a different effect on the microstructure of the metal and, consequently, on its mechanical properties. Casting processes generally produce a relatively coarse-grained structure and a random orientation of nonmetallic inclusions. The result is isotropic properties, but lower ductility than that of wrought products. Castings may also be porous. Hot-working processes, such as forging, align the inclusions (fiber structure) and thus impart anisotropic properties, with the strength and ductility generally being higher in a direction parallel to that of the inclusions. This orientation may be an advantage or disadvantage, depending upon the direction of the applied loads. Cold-working processes (such as cold rolling) also produce directional properties in the metal because of the tendency of the grains (or crystals) to align in certain directions. In addition the grains become distorted, and the metal becomes harder and stronger but less ductile. Cold-working operations (or any process causing nonuniform deformations) generally leave residual stresses in the part. Residual stresses left in a part algebraically add to the stresses induced by service loading, and in some instances these residual stresses are of major importance.

Size. Metal-forming processes may be limited as to the size of the part they can produce. Among the processes that are limited to relatively small parts are precision casting, die casting, powder metallurgy, and screw machining. Large parts are best produced by sand casting, forging, or building up component sections by welding or other joining processes. Parts with rotational symmetry can be produced by spinning; other large parts, such as domes, made of sheet or plate stock can be formed by explosive-forming techniques developed in the aerospace industries.

Shape. The complexity of a part often dictates the process or combination of processes used in its manufacture. Generally, castings can be more complex than parts made by most other fabrication processes; however, some casting processes (sand and precision investment casting) are capable of producing parts of greater complexity than others. Parts produced by the powder metallurgy process have definite restrictions as to design because of the inability of metal powder to flow like a liquid. In some instances the complexity of a part may require the fabrication of the structure by welding or brazing several sections together. The change from forming a part as a single piece to fabricating it from several sections usually requires design modifications if optimum serviceability is to be attained. The design should be left flexible until all feasible manufacturing processes have been considered.

Tolerances and finish. Parts requiring close tolerances or smooth finishes can be formed directly by precision investment casting, die casting, or such cold-working processes as swaging, drawing, or stamping. If formed by other processes, they can be finished by machining or grinding. Hot-working processes, such as forging, result in relatively rough, oxidized surfaces and relatively low dimensional accuracies. Welding operations generally result in some distortion or dimensional change. The final overall cost often governs whether the desired tolerances should be attained in the original process or obtained in secondary operations. As an example, in hot-working operations, such as rolling, forging, and extrusion, surface roughness ranges from about 100 to as high as 2000 microinches arithmetic average. In cold working, the range is between 10 and 250 μin.; in some machining operations, it is as low as 1 μin., for example, in grinding and honing.

Quantity. The number of parts to be produced is a major factor in determining the method of manufacture, and whether the part should be produced within the plant or subcontracted. Some processes are suitable only for large-quantity production because of high tooling costs; for example, permanent-mold and die casting, certain forging processes, and deep drawing. Processes such as sand casting, spinning, and welding are readily adaptable, but not necessarily restricted, to small-quantity production.

Cost. If the quantity to be produced is large, the overall finished cost of the product is usually a prime consideration in the selection of a process. In many cases cost is the deciding factor in choosing the fabrication process and perhaps the material as well. In determining the overall cost, the replacement cost of the part based on its estimated service life should be included with more immediate factors such as material and tooling costs, labor, and scrap loss.

Production factors. In some instances the time necessary to initiate production may be of significance in selecting the fabrication process. Those methods involving extensive tooling necessarily require a long lead time before production starts. Availability of equipment may be the deciding factor in choosing between two equally feasible processes, especially if it has been decided that the part is to be produced in a certain plant rather than subcontracting it to another organization. Another production factor which may be important is the required rate of production. Processes such as die casting, powder metallurgy, and deep drawing have high production rates. Conversely, sand casting, spinning, hydraulic press forging, and fusion welding are relatively slow processes.

It is thus evident that there are many factors involved in selecting an optimum for material and processes to manufacture a certain part. A very wide range of variables are involved. Strength levels for metals range from as low as 1000 lb/in.^{2} (7 megapascals) to the order of 500,000 lb/in.^{2} (3500 MPa). Size of parts may range from a thousandth of an inch (0.0254 mm) to a few feet. Rates of deformation can be as high as 40,000 ft/min (12,000 m/min), as in explosive forming, while working temperatures can be as low as cryogenic to the range of 4000°F (2200°C). Capacities of equipment for forming metal components are as high as 50,000 tons (45,000 metric tons), with a 200,000-ton (180,000-metric-ton) hydraulic forging press, which is in the design stage. *See* DRAWING OF METAL; EXTRUSION; FORGING; METAL CASTING; METAL ROLLING; POWDER METALLURGY; SHEET-METAL FORMING.

Serope Kalpakjian

Bibliography. American Society for Metals, *Formability and Workability of Metals,* 1984; American Society of Mechanical Engineers, *ASME Handbook Engineering Design,* 1965; A. L. Hoffmanner (ed.), *Metal Forming: Interrelation Between Theory and Practice,* 1971; W. F. Hosford, *Metal Forming: Mechanics and Metallurgy*, 2d ed., 1993; Society of Manufacturing Engineers, *Tool and Manufacturing Engineer's Handbook*, vol. 2: *Forming*, 1984.

Metal halide lamp

A high-pressure discharge lamp that is enclosed in a quartz envelope containing metal halides, usually iodides, and produces high-efficacy white light. These lamps are widely used for sports stadiums, roadways, commercial interiors, and industrial applications. The singular lamp feature is the compact geometry and high efficacy of nearly white light. *See* LUMINOUS EFFICACY.

Principle of operation. The metal halide lamp requires a high voltage in order to start. Once the arc discharge is ignited, the internal vapor pressure begins to increase until it reaches a preset value,

usually around 5 atm (500 kilopascals). These lamps are always connected to an auxiliary power supply called the ballast which supplies the proper voltage and current for starting and operating the lamps. The light output gradually increases over approximately 2 min as the various ingredients begin to vaporize and emit light. The light from the arc discharge comes from the metal components of the iodide compounds, which are typically a mixture of sodium, thallium, indium, scandium, dysprosium, and occasionally tin iodide. A combination of these metals produces a pleasing white light of very high efficiency, between 80 and 120 lumens per watt or even higher. *See* ARC DISCHARGE.

When the arc is struck, the light output starts at a low value and gradually increases as the pressure of the metallic compounds increases inside the quartz envelope. The wall temperature is usually about 1300°F (700°C) and the arc temperature about 9000°F (5000°C). The high temperature of the arc dissociates the halides into the electronegative portion, namely the iodine, and the metallic portion such as sodium.

Since the arc temperature excites the metal atom more easily than the iodide, the visual radiation emitted by the lamp is characteristic of the metal. Several combinations of metals produce nearly white light. Mercury, which is often an ingredient of high-pressure discharge lamps, is hardly excited in metal halide lamps. That is, the energy levels of mercury (and the iodine) do not produce significant radiation. *See* MERCURY-VAPOR LAMP.

The metals do not stay in the discharge, but diffuse and wander to the wall where they recombine with the iodine and reform the compound. However, there are other molecules that have evaporated and go into the discharge, dissociate, and thus continuously produce a fresh vapor of metal atoms which can radiate their characteristic spectrum. In this way, the relatively inefficient spectra of elements such as mercury, argon, and iodine are suppressed and the relatively efficient spectra of metals such as sodium and scandium are improved, to increase the overall luminous efficiency of the arc. The net result is a lamp whose luminous efficiency is very high.

The discharge tube is placed within an outer bulb in order to protect it from the atmosphere and reduce the hazard associated with an extremely hot quartz envelope when the lamp is operating.

Spectra and applications. Metal halide lamps come in a wide variety of sizes, depending upon their application. For general lighting, the lamps usually have closely spaced lines and are designed for long life and nearly white light. Spectra of two common metal halide lamps are shown in the **illustration**. Lamps of this type are used for lighting roads and residential shopping areas, and for interior lighting in commercial and industrial operations. Probably one of the most extensive early uses of metal halide lamps was in sports lighting, where they were instrumental in the success of color television for sports events.

Special ingredients can be used inside these lamps to make them suitable for plant growth, photocopying, and scientific usages. Certain types of compact metal halide lamps are ideal for use with movie and slide projection. In these cases, the ingredients chosen (usually indium) are used because they produce a small, intense high-brightness arc. The spectra of these lamps are nearly continuous and approach the spectra of incandescent lamps. *See* CINEMATOGRAPHY; INCANDESCENT LAMP; OPTICAL PROJECTION SYSTEMS.

In cases where a more diffuse light source is desired, the outer jacket of the lamp can be coated with a white phosphor in order both to diffuse the light and to change the color of the output and in some cases actually improve the efficiency.

Millions of these lamps are produced annually. The development of metal halide lamps is still an

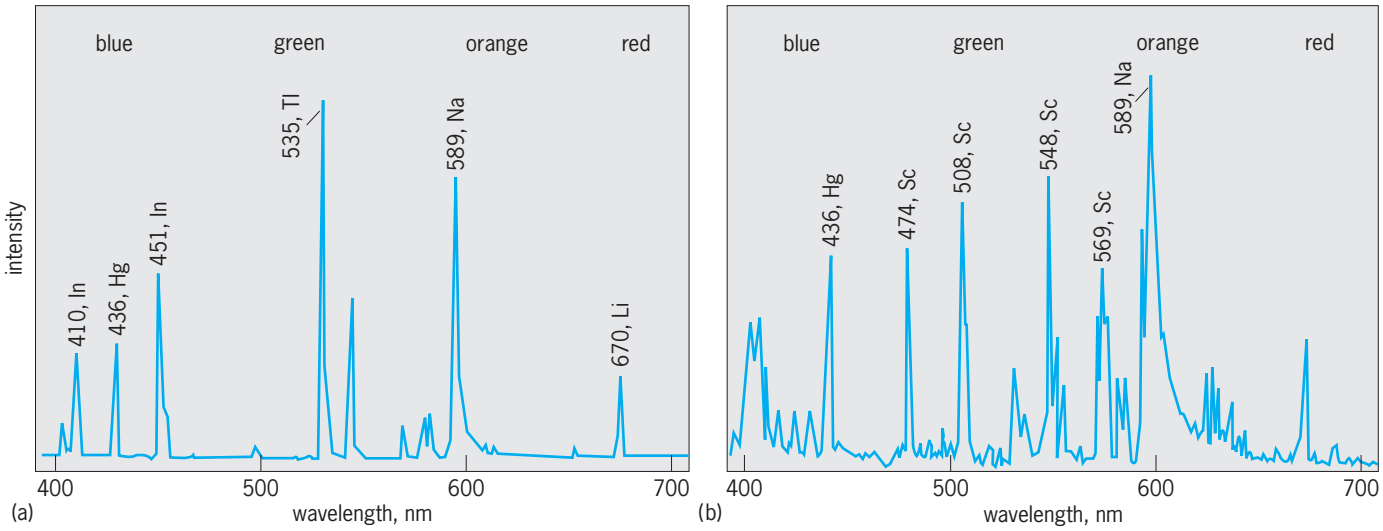

Spectrum of white-colored light from: (*a*) A 400-W metal halide lamp containing sodium iodide (NaI), thallium iodide (TlI), indium iodide (InI_3), mercury, and argon gas. The visible radiation yields an efficacy of 90 lumens per watt and color rendering index of 60. (*b*) A 175-W metal halide lamp containing sodium iodide (NaI), scandium iodide (ScI_3), mercury, and argon gas. The combined visible radiation yields an efficacy of 100 lumens per watt and a color rendering index of 60.

active subject in research laboratories, and improvements in the spectrum and the efficiency have not reached their limits. *See* ARC LAMP; VAPOR LAMP.

Gilbert H. Reiling

Bibliography. Illumination Engineering Society of North America, *IES Lighting Handbook: Reference and Application Volume*, 1993; C. Meyer and H. Nienhuis, *Discharge Lamps*, 1989.

Metal hydrides

A compound in which hydrogen is bonded chemically to a metal or metalloid element. The compounds are classified generally as ionic, transition metal, and covalent hydrides. The periodic table can be arranged according to the type of hydride formed (**Fig. 1**). Covalent hydrides are of two subtypes, binary and complex. Certain hydrides have achieved a position of modest industrial importance, but most are of theoretical interest only. Examples of hydrides are shown in the **table**.

Under extreme conditions such as in electric discharges, many metals form volatile, short-lived transient hydrides of the general formula type MH. Although some of these can be prepared experimentally, most are observed only by their spectra. They are important in studying molecular bonding. The action of atomic hydrogen at low temperatures forms surface films of unstable hydrides with many metals. *See* HYDRIDO COMPLEXES; HYDROGEN.

Ionic hydrides. The most reactive metals (alkali and alkaline-earth metals, with the exception of magnesium and beryllium) combine directly on heating with hydrogen gas at a pressure of 1 atm (10^2 kilopascals). Magnesium reacts at elevated hydrogen pressure. The reactions of hydrogen with sodium and potassium are slow because of the formation of a protective layer of the hydride, but the reaction can be accelerated by surface-active agents or catalysts such as anthracene or by stirring the molten metal under paraffin oil while hydrogen is bubbling in. *See* SURFACTANT.

Ionic (or saline) hydrides, when pure, are white, high-melting solids which have simple formulas. They are believed to consist of metal cations and hydride anions H^-. Ionic hydrides result when the difference in electronegativity between the hydrogen and the metal is large. *See* ELECTRONEGATIVITY.

The crystal structures of hydrides can be determined by the standard means of x-ray diffraction, which reveals the arrangement of metal ions only. To locate the hydrogen (or its isotope, deuterium), neutron diffraction is used. Such studies confirm that

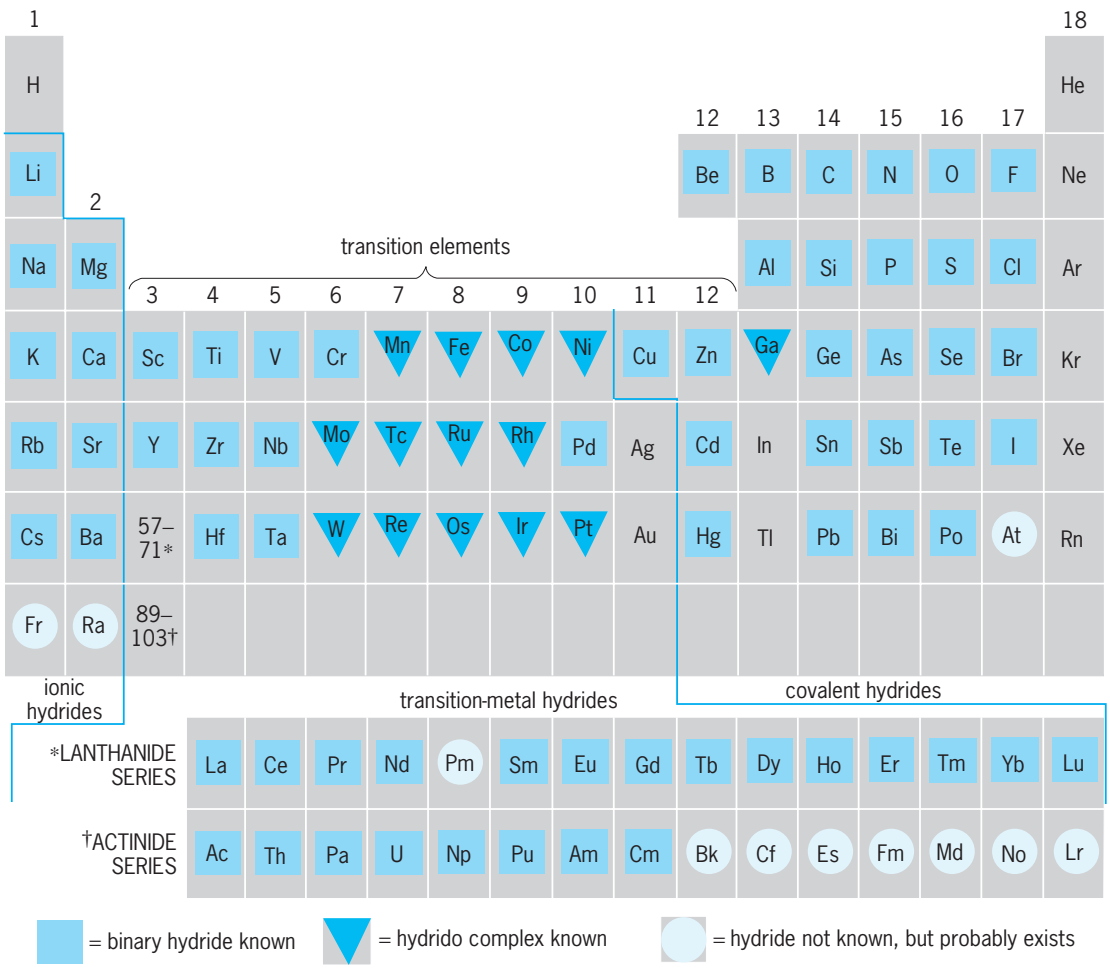

Fig. 1. Periodic table showing hydride types.

Physical properties of representative hydrides

Compound	Formula	Form	Melting point, °C (°F)	Density, g/cm^3	Boiling point, (°F) °C	Temperature at which dissociation pressure is 1 atm or 10^2 kPa, °C (°F)
Lithium hydride	LiH	White crystals	691 (1276)	0.77	—	790 (1454)
Sodium hydride	NaH	White crystals	Decomposes	1.396	—	425 (797)
Calcium hydride	CaH_2	White crystals	>1000 (1830)	1.902	—	960 (1760)
Titanium hydride	TiH_2	Gray powder	Decomposes	3.78	—	630 (1166)
Cerium hydride	CeH_2	Dark gray powder or greenish metallic single crystals	1088 (1990) for $CeH_{1.73}$	5.45	—	750 (1382) for $CeH_{2.2}$
Uranium hydride	UH_2	Black powder	~1050 (1922) at 580 atm (58.8 MPa)	10.95	—	440 (824)
Lanthanum nickel hydride	$LaNi_5H_6$	Dark gray powder	Decomposes	—	—	0 (32) [2.5 atm at 25°C or 0.25 MPa at 77 °F]
Palladium hydride	$PdH_{0.66}$	Metallic	Decomposes	~10.8	—	25 (77) for $PdH_{0.56}$, −78 (−108) for $PdH_{0.83}$
Diborane	B_2H_6	Colorless gas	−165.5 (−265.9)	0.438 at bp	−92.5 (−135)	—
Silane	SiH_4	Colorless gas	−185 (−301)	0.68 at mp	−111.8 (−169.2)	—
Stannane	SnH_4	Colorless gas	−150 (−238)	—	−52 (−62)	—
Arsine	AsH_3	Colorless gas	−113.5 (−172.3)	—	−55 (−67)	—
Stibine	SbH_3	Colorless gas	−88.5 (−127.3)	2.2 at bp	−17 (1.4)	—
Tellurium hydride	TeH_2	Colorless gas	−51 (−59.8)	2.7 at −18 °C (−0.4 °F)	−4 (25)	—
Aluminum hydride	AlH_3	White solid	Decomposes	—	—	—
Copper hydride	CuH	Dark brown solid	Decomposes slowly even at 25 °C (77 °F)	6.39	—	—

sodium hydride has the same type of structure as sodium chloride. *See* CRYSTAL STRUCTURE.

The thermal stability of the alkali metal hydrides decreases, and their chemical reactivity increases, in the order LiH, NaH, KH, RbH, and CsH. The hydride ion in these compounds has a radius of approximately 0.14 nanometer (about the same size as the fluoride ion, F^-). The anion from heavy hydrogen, that is, the deuteride ion D^-, tends to be slightly smaller than H^-, when in an equivalent environment. When molten lithium hydride is electrolyzed, hydrogen is evolved at the anode, since H^- is the migrating species.

A common way to represent a reversible metal-hydrogen system is to plot the composition against hydrogen pressure (called the dissociation pressure) at a fixed temperature. **Figure 2** shows the 500 and 575°C (932 and 1067°F) isotherms for the sodium-hydrogen system. Pure liquid sodium at 575°C (1067°F) exists at point A, and as hydrogen is admitted, it dissolves in the metal up to the limit at point B. This is a solution of NaH (10%) in molten Na (90%). When more hydrogen is pumped in, a solid appears, whose composition is indicated by point C ($NaH_{0.90}$, that is, sodium hydride deficient in hydrogen). During this stage (points B to C), the hydrogen pressure remains constant at 36.6 atm (3.7 MPa). The dissociation pressures at the various plateaus may be computed for any temperature by the equation $\log P = (-5960/T) + 8.59$, where P is the pressure in atmospheres and T the absolute temperature. Further introduction of hydrogen finally yields pure NaH at point D. In the plateau region (B to C), each mole of hydrogen which reacts liberates 27 kcal (113 kilojoules) of heat. More heat is liberated in the cases of lithium hydride (43 kcal or 180 kJ/mole of H_2) and CaH_2 (42 kcal or 176 kJ). *See* PHASE EQUILIBRIUM.

The ionic hydrides are all exceedingly reactive reducing agents. Their reaction with water yields the metal hydroxide and hydrogen, which sometimes ignites at the high temperature resulting from the reaction. Sodium hydride dissolved in molten sodium hydroxide is employed in descaling steel and titanium. Calcium hydride is a source of hydrogen in remote areas, for example, to fill meteorological balloons. CaH_2 is also used as a desiccant in transformer oil and in the reduction of certain oxides to metals, such as tantalum.

Transition metal hydrides. This group of compounds is less well understood than the ionic and

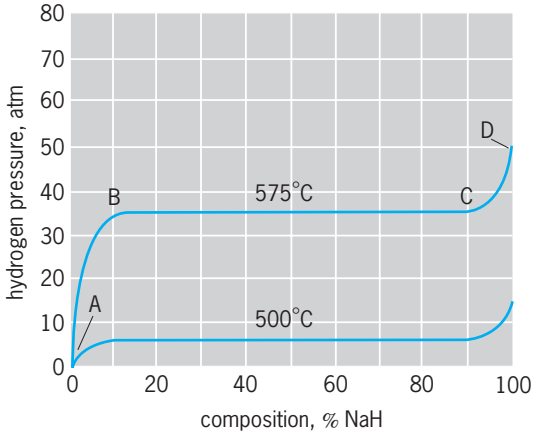

Fig. 2. Sodium-hydrogen system. °F = (°C × 1.8) + 32; 1 atm = 10^5 Pa.

covalent hydrides. When lanthanum is heated in hydrogen, the gas is readily taken up, forming first a material whose composition is approximtely $LaH_{1.86}$ (the "dihydride" phase) and finally LaH_3. The electrical conductivity does not change radically between La and $LaH_{1.86}$, but it falls dramatically from $LaH_{1.86}$ to LaH_3. The hydrogen is generally thought to form hydride ions, with radius about 0.13 nm, but some of the valence electrons remain free or mobile in the dihydride. This accounts for the conductivity of this substance. These mobile electrons are said to be in the conduction band. They are progressively consumed as the dihydride is converted to the trihydride. A single phase exists between $LaH_{1.86}$ and LaH_3, and it is possible to arrest the addition of hydrogen at any intermediate composition, such as $LaH_{2.63}$. Such a substance is sometimes referred to as a solid solution and sometimes as a compound which represents extreme nonstoichiometry. These features (electrical conductivity and nonstoichiometry) characterize the transition metal hydrides. *See* NONSTOICHIOMETRIC COMPOUNDS.

The ionic hydrides, such as NaH, exhibit much more limited nonstoichiometry. When sodium is converted to pure NaH, all electrons in the conduction band are consumed and the product is white and insulating. The hydrides of the rare-earth metals are the easiest to describe in terms of an ionic model. The hydrides of the actinide metals are known through curium, the best known being uranium hydride. All are dark-gray powders which ignite spontaneously in the air. Large single crystals of cerium hydride are fairly stable in air.

Copper hydride, CuH, is unique in that it is the only hydride which can be precipitated from aqueous solution. It forms a dark-brown sediment which decomposes when dried. Copper hydride has a hexagonal crystal structure and is essentially a covalent compound.

Titanium, zirconium, vanadium, and niobium all react with hydrogen, forming dihydrides at the limiting compositions. The materials are conducting at all compositions. Except for palladium, the metals to the right of titanium in the periodic table show decreased reactivity toward hydrogen. Metal of the vanadium family and also chromium absorb hydrogen almost to the monohydride stage. The transition metals in the manganese, iron, cobalt, and nickel groups dissolve only small amounts of hydrogen, except for palladium, which readily absorbs hydrogen to the limit $PdH_{0.66}$. The palladium-hydrogen system has been much studied, because of its importance in catalytic hydrogenation. Thorium hydride, Th_4H_{15}, becomes superconducting at approximately 8 K (−445°F). *See* PALLADIUM; SUPERCONDUCTIVITY.

When hydrogen (protium, ^1H) in a metal hydride is replaced with its isotope deuterium, the product (deuteride) is usually less stable. The dissociation pressures of the deuterides are about twice those of the corresponding hydrides at the same temperature in the great majority of cases. The inverse effect is known, however. For example, at 40°C (104°F) vanadium hydride ($VH_{1.5}$) has a dissociation pressure of 4.8 atm (490 kPa), while $VD_{1.5}$ has a dissociation pressure of 1.8 atm (182 kPa). The uranium compounds of the three isotopes of hydrogen (UH_3, UD_3, and UT_3) are frequently used in the laboratory as sources and reservoirs for these gases and for their purification. Titanium hydride also serves in this capacity. Some of the hydrides serve as a starting point for the synthesis of halides, sulfides, and so forth, of the metal. For example, this is represented by reaction (1). Deuterides and tritides of actinide metals

$$UH_3 + 3HCl \rightarrow UCl_3 + 3H_2 \qquad (1)$$

probably play an important role in thermonuclear (hydrogen) weapons.

Common metals such as steel and copper dissolve small quantities of hydrogen at elevated temperatures. On cooling, the gas comes out of solution and results in severe degradation of the mechanical properties of the metals. This is called hydrogen embrittlement and can be prevented by degassing the metal while it is still molten. Problems from hydrogen embrittlement also arise in high-pressure catalytic units employing hydrogen gas and at electrode surfaces where hydrogen ion is reduced to hydrogen, as in some electroplating processes. The cause of this embrittlement is the introduction of dislocations and stacking faults by the hydrogen atoms. The result is to suppress dislocation glide near cracks, preventing stress relaxation. *See* CRYSTAL DEFECTS.

A number of alloys or intermetallic compounds are known which react with hydrogen and form ternary hydrides. Examples are $LaNi_5H_6$, Mg_2NiH_4, $AlTh_2H_2$, and $CaAg_2H$. $LaNi_5H_6$ permits an attractive means of portable hydrogen storage. This might be important if hydrogen is generated on a massive scale as a pollution-free energy source for fuel cells and internal combustion engines. Storage of hydrogen as compressed gas or as liquid is not convenient. The energy density of $LaNi_5H_6$ is twice that of liquid hydrogen and 12 times that of the compressed gas. The ternary hydride $LaNi_5H_6$ is readily formed reversibly from lanthanum-nickel alloy and hydrogen. Pure rare-earth metals are not necessary in making the nickel alloy; the commercial material (misch metal, mostly Ce, La, and Nd) is satisfactory. After charging with hydrogen, these materials provide practical gas pressure (a few atmospheres) from −30 to +30°C (−22 to 86°F), and are thus attractive for powering hydrogen-burning engines of automobiles.

Owing to the highly reversible character of these metal-hydrogen systems, they have recently been studied as heat pumps and energy storage devices. To operate an air-conditioning unit, hydrogen is pumped from one metal hydride unit (such as $LaNi_5H_6$), which provides cooling, to an alloy unit ($LaNi_5$), where heat is evolved on formation of hydride. Alternatively, solar or waste thermal energy could be converted into mechanical work.

Another important ternary hydride is the Zr-U-H system, which is employed in pulsed reactors of the Triga type. The role of the hydrogen atom is to moderate fast neutrons. Another type of ternary hydride

is illustrated by the thorium carbohydrides, Th_2Ch_2 and Th_3Ch_4. *See* RARE-EARTH ELEMENTS.

Covalent hydrides. Most evidence indicates that in ionic and metallic hydrides an electronic pair is associated primarily with the hydrogen as H^-, while in covalent hydrides the electron pair is shared between the hydrogen atom and an atom of another element. In these compounds hydrogen is considerably smaller (radius 0.03 nm) than in the ionic hydrides. Covalent hydrides usually consist of small molecules, in which case they are gases (SiH_4 and SbH_3), but some form high polymers, in which case they are nonvolatile solids (AlH_3 and ZnH_2). All tend to decompose irreversibly rather easily on heating. Covalent hydrides are generally synthesized indirectly, not from direct combination of the elements. Gaseous hydrides such as SiH_4, PH_3, and AsH_3 can be generated by heating a solid mixture of the corresponding oxide and $LiAlH_4$. *See* CHEMICAL BONDING.

The hydrides of boron deserve special mention, as their bonding posed a difficult problem in chemical theory which has been resolved only recently. In diborane, B_2H_6, four of the hydrogen atoms are bonded to the two boron atoms by ordinary single, covalent bonds. The other two hydrogen atoms are linked to both boron atoms, the four bonds being formed by only two electron pairs. The structure of diborane is illustrated in **Fig. 3**. The hydrogen atoms which bridge between the two boron atoms do not form an ordinary hydrogen bond but another type usually described as an electron-deficient bridging bond. There are nearly two dozen higher boron hydrides, of which B_4H_{10}, B_5H_9, B_6H_{10}, and $B_{10}H_{14}$ are representative. A group of compounds, the carboranes, has been discovered in which some of the boron atoms of the higher hydrides are replaced with carbon atoms. An example is $B_{10}C_2H_{10}$. *See* BORANE; CARBORANE.

Numerous derivatives of the boron hydrides have been prepared, of which sodium tetrahydroborate (or tetrahydridoborate or borohydride), $NaBH_4$, is the most important. It is used as a selective reducing agent in organic chemistry. This hydrido complex is moderately stable in water, especially in alkaline solution. This stability permits its use in recovery of heavy metals from industrial waste streams. *See* HYDROBORATION.

Aluminum hydride, AlH_3, normally exists as an insoluble, nonvolatile polymer. Its derivative, lithium tetrahydroaluminate (or lithium aluminum hydride), $LiAlH_4$, is an important reducing agent in both organic and inorganic chemistry. For example, silane can be prepared

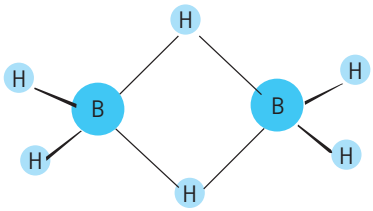

Fig. 3. Structure of diborane, B_2H_6.

as in reaction (2). Lithium tetrahydroaluminate

$$SiCl_4 + LiAlH_4 \xrightarrow{\text{ether}} SiH_4 + LiCl + AlCl_3 \qquad (2)$$

reacts vigorously with water. An ether is generally used as the solvent. Sodium cyanotrihydroborate, $NaBH_3CN$, can be used in acidic medium. A family of aluminum-based reducing agents is now available commercially including sodium diethyl-dihydroaluminate, $NaAlH_2(C_2H_5)_2$; sodium tri-*tert*-butoxohydroaluminate, $NaAlH(O\text{-}t\text{-}C_4H_9)_3$; sodium bis(2-methoxyethoxo)dihydroaluminate, $NaAlH_2(OCO_2CH_2OCH_3)_2$; diethylaluminum hydride, $(C_2H_5)_2AlH$; and di-*iso*-butylaluminum hydride, $(i\text{-}C_4H_9)_2AlH$. All are soluble in aromatic hydrocarbons. They are rather expensive, but their specific reducing powers make them attractive for synthesizing high-value products, such as pharmaceuticals, flavorings, fragrances, dyes, and insecticides.

There are thousands of hydrocarbons, which represent hydrogen compounds of carbon, the first element in group 14. The simplest of these compounds is methane, CH_4. Other 14 hydrogen compounds are silane, SiH_4; germane, GeH_4; stannane, SnH_4; and plumbane, PbH_4. This series is progressively less stable, with stannane decomposing at room temperature. Higher hydrides of silicon are known, Si_2H_6 and so on up to $Si_{10}H_{22}$. These compounds, Si_nH_{2n+2}, are analogous to the corresponding alkanes, C_nH_{2n+2}. Various silane derivates, such as SiH_3HI and $(SiH_3)_3N$, have been synthesized. A parallel situation exists in the case of higher germanium hydrides. Unstable distannane, Sn_2H_6, has also been made. Mixed compounds such as SiH_3GeH_3 can be prepared. The group 15 hydrides show the same stability trends as the preceding group, with stability decreasing in the order NH_3, PH_3, AsH_3, SbH_3, and BiH_3. Group 16 hydrides, including H_2Se, H_2Te, and H_2Po, also grow less stable down the series, the same trends continuing in the case of hydrogen compounds of the halogens. These compounds in the last two series, such as H_2Se and HI, are ordinarily considered as acids. *See* HYDRIDE.

James C. Warf

Bibliography. A. Dedieu, *Transition Metal Hydrides*, 1992; E. L. Muetterties (ed.), *Boron Hydride Chemistry*, 1975.

Metal matrix composite

A material in which a continuous metallic phase (the matrix) is combined with another phase (the reinforcement) that constitutes a few percent to around 50% of the material's total volume. In the strictest sense, metal matrix composite materials are not produced by conventional alloying. This feature differentiates most metal matrix composites from many other multiphase metallic materials, such as pearlitic steels or hypereutectic aluminum-silicon alloys. *See* ALLOY.

The particular benefits exhibited by metal matrix composites, such as lower density, increased specific

strength and stiffness, increased high-temperature performance limits, and improved wear-abrasion resistance, are dependent on the properties of the matrix alloy and of the reinforcing phase. The selection of the matrix is empirically based, using readily available alloys; and the major consideration is the nature of the reinforcing phase.

Matrices and reinforcements. A large variety of metal matrix composite materials exist. The reinforcing phase can be fibrous, platelike, or equiaxed (having equal dimensions in all directions); and its size can also vary widely, from about 0.1 to more than 100 micrometers. Matrices based on most engineering metals have been explored, including aluminum, magnesium, zinc, copper, titanium, nickel, cobalt, iron, and various aluminides. This wide variety of systems has led to an equally wide spectrum of properties for these materials and of processing methods used for their fabrication.

Reinforcements used in metal matrix composites fall in five categories: continuous fibers, short fibers, whiskers, equiaxed particles, and interconnected networks.

Continuous fibers. Several continuous fibers or filaments are used in metal matrix composites. Their elastic moduli vary significantly, depending on the nature of the fiber and its fabrication process. For example, silica-alumina spinels and microcrystalline or amorphous polycarbosilane-derived fibers possess significantly lower elastic moduli than do pure alumina or crystalline β-silicon carbide produced by chemical vapor deposition. Carbon fiber strength and modulus also vary significantly with processing, depending on the level of graphitization of the microstructure. *See* ALUMINUM; ELASTICITY; GRAPHITE; SPINEL.

Short fibers. Short fibers are less expensive, especially when they are mass-produced for other applications such as high-temperature thermal insulation. Their physical properties can be similar to those of continuous fibers; however their reinforcing efficiency in the matrix is also far lower. Short fibers used in engineering practice include chopped carbon fibers and alumina-silica fibers.

Whiskers. Whiskers are single-crystal short fibers, produced to feature highly desirable mechanical properties due to lack of microstructural defects. Whiskers have typically been made of silicon carbide, and they are often priced far higher than short fibers. The high price and toxicity of most whiskers have prevented their application in engineering practice.

Single-crystal whiskers, because of the absence of grain boundary defects, offer much higher tensile strength than other types of discontinuous reinforcements, and thus they are preferred for certain applications of discontinuously reinforced metal matrix composities. The whiskers can be aligned to a preferred orientation by conventional metallurgical processes; higher directional strengths can be achieved in finished components where fabrication is by extrusion, rolling, forging, or superplastic forming. Whiskers tend to produce anisotropic properties due to their alignment during processing, while particulate materials usually produce essentially isotropic properties. *See* CRYSTAL WHISKERS; GRAIN BOUNDARIES; SILICON. *See also* EXTRUSION; FORGING; METAL ROLLING; SUPERPLASTICITY.

Equiaxed particles. Equiaxed particles of several ceramics, including those containing silicon carbide, aluminum oxide, boron carbide, and tungsten carbide, do not provide the possibility for preferential strengthening of the matrix along selected directions; however, their price is low and their combination with the metal is relatively easier. These reinforcements are therefore used in many metal-matrix composite systems, including mass-produced aluminum matrix composites.

Interconnected cellular networks. These can be produced by several methods, such as by chemical vapor deposition of ceramic onto a pyrolizable polymer foam or by conversion of a preceramic polymer foam prior to infiltration with the molten matrix. Alternatively, some processing techniques for in-place metal matrix composites, including directional oxidation of aluminum melts, produce interconnected reinforcing networks.

In-place metal matrix composites comprise those in which the reinforcement is produced from the matrix; although these materials are actually in some cases produced by alloying, their designation as metal matrix composites has resulted from the special morphology of the second phase (for example, some directionally solidified eutectics that have fibrous second phase) or from the ceramic nature of the second phase (for example, titanium-carbide-reinforced steel produced by solidification of iron-titanium-carbon alloys).

Microstructure. The microstructure of a metal matrix composite comprises the structure of matrix and reinforcement, that is, the interface and the distribution of the reinforcement within the matrix.

The distribution of the reinforcement within the composite can be controlled to some extent during processing. Large modifications can be handled individually and distributed with relatively great precision within the composite. Continuous fibers of smaller diameter, in the range of 10 μm or less, are generally combined with the matrix by infiltration. The fibers then tend to concentrate into regions of high-volume fraction, near 50 vol %, within which individual fibers are relatively randomly distributed. Greater control of the reinforcement volume fraction can be exerted with short fibers and whiskers because of bending of fibers between their contact points.

In particle-reinforced metal composites, the reinforcement distribution is affected by solidification of the matrix in liquid-phase processes, because the growing solid phase may push the particles as it grows. In powder metallurgical processes, reinforcement particle clustering may result from unequal sizes of matrix and reinforcement powder particles.

At the interface between matrix and reinforcement, intimate atomic-level bonding is typically achieved between matrix and reinforcement.

Microstructural evolution of the matrix is governed by the same basic laws that govern the microstructural evolution of all metals; however, the reinforcement can exert a strong influence on the course taken by the evolving matrix microstructure. In solidification, for example, the reinforcement has been shown to strongly reduce microsegregation in the matrix, to influence the stability of plane solidification during cooling, and to alter the coarsening of matrix microstructure. In solid-state deformation processes, the reinforcement has been shown to alter the evolution of texture in the matrix, the kinetics of recrystallization, and the grain size after recrystallization. *See* METAL.

Composite properties. Composite properties depend first and foremost on the nature of the composite; however, certain detailed microstructural features of the composite can exert a significant influence on its behavior.

Physical properties of the metal, which can be significantly altered by addition of a reinforcement, are chiefly dependent on the reinforcement distribution. A good example is aluminum-silicon carbide composites, for which the presence of the ceramic increases substantially the elastic modulus of the metal without greatly affecting its density. Elastic moduli for 6061 aluminum matrix composites reinforced with discrete silicon carbide particles or whiskers have been calculated by using the rule of mixtures for the same matrix reinforced with two types of commercial continuous silicon carbide fibers. As a result, several general facts become apparent. First, modulus improvements are significant, even with equiaxed silicon carbide particles, which are far less expensive than fibers or whiskers. However, the level of improvement depends on the shape and alignment of the silicon carbide. Also, it depends on the processing of the reinforcement: for the same reinforcement shape (continuous fibers), microcrystalline polycarbosilane-derived silicon carbide fibers yield much lower improvements than do crystalline β-silicon carbide fibers. These features, of influence of reinforcement shape, orientation, and processing on modulus, are quite general, being also observed for example in metal matrix composites reinforced with aluminum oxide or carbon.

Other properties, such as the strength of metal matrix composites, depend in a much more complex manner on composite microstructure. The strength of a fiber-reinforced composite, for example, is

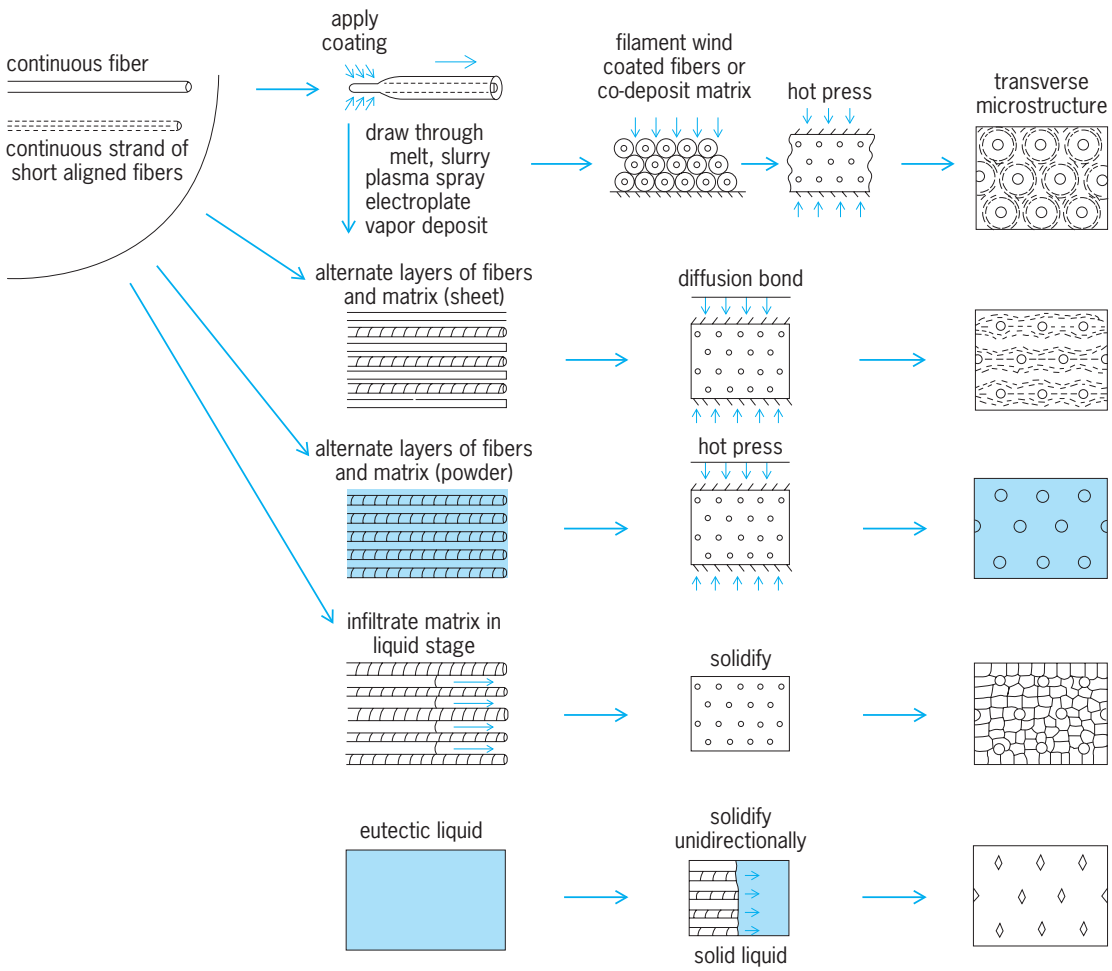

Fig. 1. Examples of methods used to make metal matrix composites. (*Space Sciences Laboratory, General Electric Valley Forge Space Technology Center, King Prussia, Pennsylvania, 1967*)

determined by fracture processes, themselves governed by a combination of microstructural phenomena and features. These include plastic deformation of the matrix, the presence of brittle phases in the matrix, the strength of the interface, the distribution of flaws in the reinforcement, and the distribution of the reinforcement within the composite. Consequently, predicting the strength of the composite from that of its constituent phases is generally difficult. *See* BRITTLENESS; PLASTIC DEFORMATION OF METAL.

Design considerations. In general, metal matrix composites exhibit good creep resistance (deformation under constant stress or load), fatigue properties (response to cyclic loading), and toughness (ability to absorb energy on impact without failure). These properties, in combination with strength and stiffness, are used by the designer in selection of a material for structural applications. Also important to the designer is the fact that metal matrix composites are anisotropic; that is, the mechanical properties depend critically on the orientation of the fibers with respect to the applied force and on the nature of the stress (tensile or compressive). Anisotropy can be an advantage or disadvantage, depending on the state of stress on the part in service. *See* CREEP (MATERIALS); METAL, MECHANICAL PROPERTIES OF; STRESS AND STRAIN.

Production. A variety of techniques are available for the production of continuous or discontinuous metal matrix composites. These may be broadly classified as diffusion processes, deposition processes, and liquid processes (**Fig. 1**).

Diffusion processes. These include the pressing and sintering of a powder metal matrix and bare or coated fibers, and the hot or cold pressing of the reinforcement between thin foils of the matrix metal or alloy.

In diffusion bonding, a stack of matrix foils are pressed around rows of relatively large-diameter fibers. Another solid-state process blends powdered matrix and reinforcement, and the matrix is densified around the reinforcement by conventional sintering or hot-pressing processes. *See* POWDER METALLURGY; SINTERING.

Deposition processes. Primary techniques are electrode-positioning of the matrix around fibers, plasma spraying of the matrix around fibers, or vacuum deposition of the matrix around fibers. In each case, it is usual to complete the process by cold or hot pressing. *See* ELECTROPLATING OF METALS; VACUUM METALLURGY.

Liquid processes. The important techniques in this category are infiltration of the liquid matrix around fibers, the pressing of powdered matrix and fibers at a temperature above the melting point of the matrix, and unidirectional solidification of eutectic alloys. *See* EUTECTICS.

Other processes. Spray casting processes (another class of liquid-state processes) involve spraying the matrix with or onto the reinforcement; solidification is achieved by heat extraction to the supporting substrate. There are additional processes that involve a partly liquid matrix. These include compocasting (a dispersion process using the matrix in semisolid form) and liquid-phase sintering processes.

In-place processes create the reinforcement from within the matrix. This can be accomplished by simple alloying and solidification, as is the case with directionally solidified eutectic alloys. Alternatively, the reinforcement can be formed by reaction of elements contained within the matrix in another phase such as a gas.

Fabrication. Composite processing methods combine the reinforcement with the matrix. This is accomplished while the matrix is either solid or liquid.

Typical liquid-state processes include the dispersion processes, which are casting techniques. A second set of processes involves liquid-metal impregnation; these include squeeze casting, where a preform or a bed of dispersoids is impregnated by molten alloy under hydraulic pressure. A third set comprises spray processes. In one of these, a molten metal stream is fragmented by means of a high-speed cold inert-gas jet passing through a spray gun, and dispersoid powders are simultaneously injected. A stream of molten droplets and dispersoid powders is directed toward a collector substrate where droplets recombine and solidify to form a high-density deposit.

In rapid solidification processing of composites, a jet of liquid alloy–particle slurry impinges under pressure on a water-cooled copper wheel (**Fig. 2**),

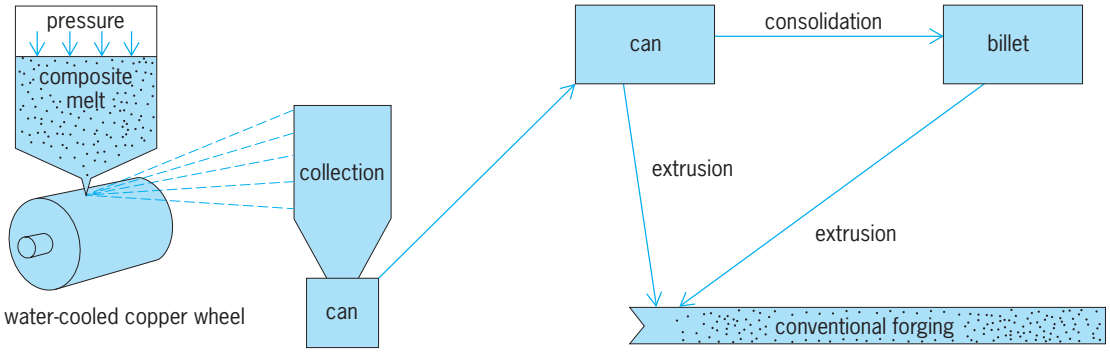

Fig. 2. Steps involved in rapid solidification processing of metal matrix composites.

Some composite components with proven potential

Composite	Components	Advantages
Aluminum–silicon carbide (particle)	Piston	Reduced weight, high strength and wear resistance
	Brake rotor, caliper, liner	High wear resistance, reduced weight
	Propeller shaft	Reduced weight, high specific stiffness
Aluminum–silicon carbide (whiskers)	Connecting rod	Reduced reciprocating mass, high specific strength and stiffness, low coefficient of thermal expansion
Magnesum–silicon carbide (particle)	Sprockets, pulleys, and covers	Reduced weight, high strength and stiffness
Aluminum–aluminum oxide (short fibers)	Piston ring	Wear resistance, high running temperature
	Piston crown (combustion bowl)	Reduced reciprocating mass, high creep and fatigue resistance
Aluminum–aluminum oxide (long fibers)	Connecting rod	Reduced reciprocating mass, improved strength and stiffness
Copper-graphite	Electrical contact strips, electronics packaging, bearings	Low friction and wear, low coefficient of thermal expansion
Aluminum-graphite	Cylinder, liner piston, bearings	Call resistance, reduced friction, wear, and weight
Aluminum–titanium carbide (particle)	Piston, connecting rod	Reduced weight and wear
Aluminum–fiber flax	Piston	Reduced weight and wear
Aluminum–aluminum oxide fibers–carbon fibers	Engine block	Reduced weight, improved strength and wear resistance

and the resulting flake powders are collected. The powder is put into cans and consolidated into billet or extruded to form a dense composite with higher yield strength, ultimate tensile strength, and ultimate ductility. *See* METALLIC GLASSES.

Depending on the process, the desired microstructure, and the desired part, metal matrix composites can be produced to net or near-net shape; or alternatively they can be produced as billet or ingot material for secondary shaping and processing.

Applications. The combined attributes of metal matrix composites, together with the costs of fabrication, vary widely with the nature of the material, the processing methods, and the quality of the product. In engineering, the type of composite used and its application vary significantly, as do the attributes that drive the choice of metal matrix composites in design (see **table**). For example, high specific modulus, low cost, and high weldability of extruded aluminum oxide particle–reinforced aluminum are the properties desirable for bicycle frames. High wear resistance, low weight, low cost, improved high-temperature properties, and the possibility for incorporation in a larger part of unreinforced aluminum are the considerations for design of diesel engine pistons. *See* COMPOSITE MATERIAL; HIGH-TEMPERATURE MATERIALS.

Mel M. Schwartz

Bibliography. American Society for Metals, *ASM Metals Handbook*, vol. 2, 9th ed., 1991; S. M. Lee, *Reference Book for Composites Technology*, 1989; S. Ray, Review synthesis of cast metal matrix particulate composites, *J. Mater. Sci.*, 28:5397–5413, 1993; T. Rienhart, *Engineered Materials Handbook: Composites*, vol. 1, 1988; M. M. Schwartz, *Composite Materials Handbook*, 2d ed., 1992; S. Suresh, A. Mortensen, and A. Needleman (eds.), *Fundamentals of Metal Matrix Composites*, 1993.

Metal rolling

Reducing or changing the cross-sectional area of a workpiece by the compressive forces exerted by rotating rolls. The original material fed into the rolls is usually an ingot from a foundry. The largest product in hot rolling is called a bloom; by successive hot- and cold-rolling operations the bloom is reduced to a billet, slab, plate, sheet, strip, and foil, in decreasing order of thickness and size. The initial breakdown of the ingot by rolling changes the coarse-grained, brittle, and porous structure into a wrought structure with greater ductility and finer grain size.

Process. A schematic presentation of the rolling process, in which the thickness of the metal is reduced as it passes through the rolls, is shown in **Fig. 1***a* . The speed at which the metal moves during rolling changes, as shown in Fig. 1*b*, to keep the volume rate of flow constant throughout the roll gap. Hence, as the thickness decreases, the velocity increases; however, the surface speed of a point on the roll is constant, and there is therefore relative sliding between the roll and the strip. The direction of this relative velocity changes at a point along the contact area, this point being known as the neutral or no-slip point. At the neutral point, roll and strip have the same velocity; to the left of this point (entry side) the strip moves more slowly than the roll; and to the right of this point (exit) it moves faster. Hence, the direction of frictional forces acting on the strip are opposite in these two regions, as shown in Fig. 1*a*. The net frictional force acting on the strip must be in the direction of exit to enable the rolling operation to take place. Although friction is a disadvantage in many metal-working processes, it is a necessity in rolling; without friction the rolls cannot pull the strip through the roll gap. It has been observed that in hot rolling the coefficient of friction may be as

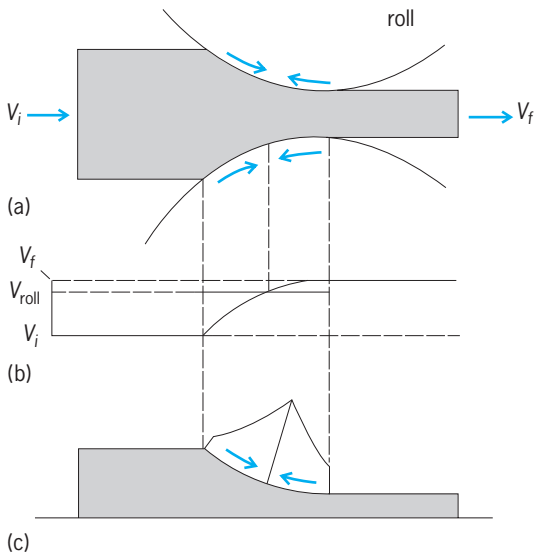

Fig. 1. Rolling process. (*a*) Direction of friction forces in the roll gap. (*b*) Velocity distribution. (*c*) Normal pressure acting on the strip in the roll gap.

much as 0.7, while in cold rolling it generally ranges 0.02–0.3.

The normal pressure distribution on the roll and hence on the strip is of the form shown in Fig. 1*c*. Because of its particular shape this pressure distribution is known as the friction hill. The overall slope of this hill depends on the coefficient of friction and the ratio of roll-strip contact length to the thickness of the strip. The size of the friction hill (area under the curve) represents the magnitude of the roll separating force per unit width that tends to push the rolls apart. This force can be reduced by increasing the workpiece temperature, reducing friction, taking smaller bites, and using smaller roll radii. Another method of reducing this force is by applying tension to the strip; this lowers the apparent compressive flow stress of the material. The tension can be back tension (entry) or front tension (exit) or both. Depending on the magnitude of these tensions, the neutral point shifts from its original position.

The roll-separating force can become so great that it is not possible to reduce further the thickness of the strip during the particular pass. With such a force the rolls deform elastically (roll flattening). In such cases rolling may be accomplished by improving lubrication, rolling a stack of sheets at the same time, or annealing the strip.

The maximum possible draft (the difference between the initial and final thickness of the strip) is a function of the friction coefficient an roll radius; it can be shown analytically that maximum draft is equal to the coefficient of friction squared times the roll radius.

The workpiece may become wider during rolling; this is called spreading. It is found that spreading increases with the thickness-width ratio of the workpiece and with decreasing coefficient of friction and decreasing radius-thickness ratio of the roll.

Practice. A great variety of roll arrangements and equipment are used in rolling. Some basic types are shown in **Fig. 2**. In another method (Steckel rolling) the strip is pulled through idler rolls by front tension only. The proper reduction per pass in rolling depends on the type of material and other factors; for soft, nonferrous metals, reductions are usually high, while for high-strength alloys they are small.

Temperatures in the hot rolling of various metals and alloys are similar to those in forging, namely, for aluminum alloys, 750–850°F (400–450°C); copper alloys, 1500°F (820°C); alloy steels, 1700–2300°F (930–1260°C); titanium alloys, 1400–1800°F (760–980°C); and refractory metal alloys, 1800–3000°F (980–1650°C). Rolling speeds range up to about 5000 ft/min (1500 m/min). Although hot rolling has the advantage of lowering forces and increasing the ductility of the metal, cold rolling gives smoother finish, closer tolerances, and increased strength.

The thickness of a rolled sheet is identified by gage number: The smaller the number the thicker the sheet. Actual thickness depends on the particular standard used; there are different standards, depending on the type of metal. The terminology for

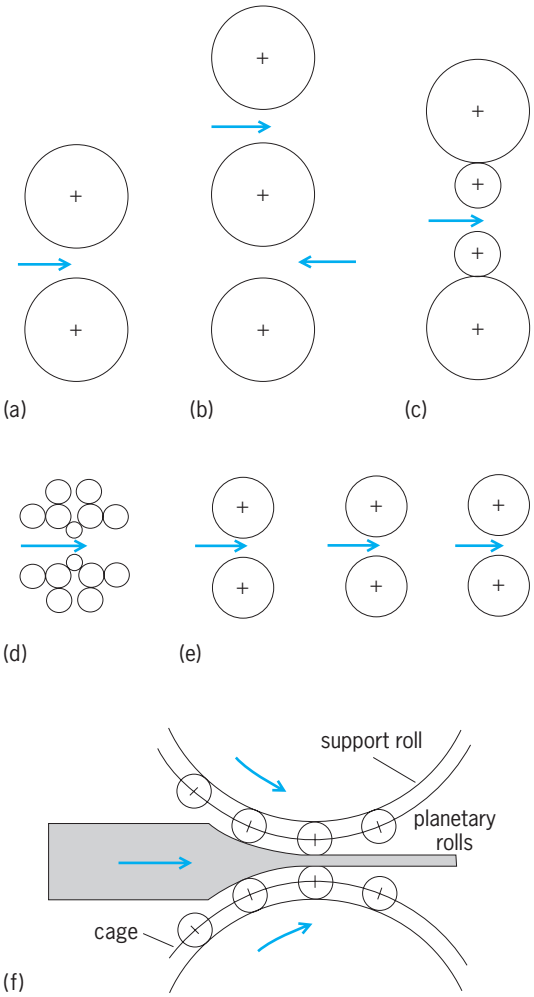

Fig. 2. Basic types of roll arrangement. (*a*) Two-high. (*b*) Three-high. (*c*) Four-high. (*d*) Cluster. (*e*) Tandem rolling with three stands. (*f*) Planetary mill.

hardness for rolled sheets is shown in the **table**. A sheet that is reduced in thickness by one gage number from the annealed condition is called quarter hard, by two gage numbers half hard, and so on, as listed in the table. The dead-soft condition is fully annealed.

Successful rolling practice requires a careful balance of factors such as reduction per pass, control of temperature, lubrication, roll size and finish, and intermediate annealing.

Requirements for roll materials are mainly strength and resistance to wear. Common roll materials are cast iron, cast steel, and forged steel. In cold rolling, cemented carbide is also used for small rolls. To avoid variation in strip thickness due to bending of the rolls, rolls are ground to a particular geometry, called camber, whereby the center of the roll has a diameter of a few thousandths of an inch greater than the diameter of its ends.

Hot rolling is usually carried out without a lubricant, although graphite or grease may be used. Cold rolling is carried out with low-viscosity lubricants, paraffin being a suitable lubricant for nonferrous metals. The type of residual stresses in the strip depends on the reduction and also on roll radius. Small rolls or small reductions produce compressive residual stresses on the surface of the strip and tensile stresses in the central portions, whereas large rolls produce the opposite residual stress pattern. A number of defects (wavy edges, zipper breaks, edge cracks, alligatoring) can result during the rolling process. These are usually eliminated or reduced by changes in operating variables or by special techniques.

Miscellaneous processes. Most seamless tubing is produced by a technique based on the principle that the inside of a round rod is subjected to secondary tensile stresses when compressed radially (**Fig. 3***a*). A simple demonstration of this can be made by rolling an eraser, removed from the end of a pencil, between a flat surface and a ruler that is moved back and forth (Fig. 3*b*). In a short time a hole is produced in the center of the eraser. The roll-piercing of round bars to make seamless tubing (Mannesmann process) is based on this principle (Fig. 3*c*). The compressive radial forces are supplied by two rolls with their axes in parallel planes, but at an angle to each other so as to move the workpiece through the rolls. The rolling action causes the center of the billet to rupture. The purpose of the mandrel is to expand the tube and improve the surface finish of the inside. For this process to be successful, the billet material should be of high quality and free of defects.

Hardness terminology for rolled sheets

Terminology	Increase in gage number	Reduction in thickness, %
Annealed (dead-soft)	0	0
1/8 hard	—	6
1/4 hard	1	11
1/2 hard	2	21
3/4 hard	3	29
Hard	4	37
Extra hard	6	50
Spring hard	8	60
Extra spring hard	10	69

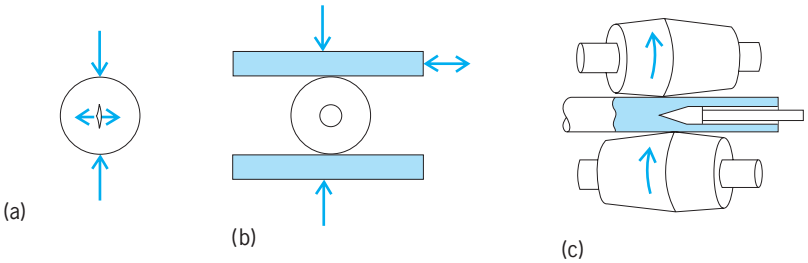

Fig. 3. Development of the Mannesmann process for producing seamless tubing (*a–c* explained in text).

In the ring-rolling process, a thin, large-diameter ring is produced from a thicker and smaller-diameter ring. This is accomplished by placing the ring between two rotating rolls; the reduction in thickness of the ring is compensated by an increase in diameter; there is little or no change in the width of the ring.

Sections such as railroad tracks and I beams are also rolled by passing the stock through a number of specially designed rolls with their axes placed in different directions. Roll-pass design for such sections requires considerable experience to avoid defective products and to obtain desirable properties. Metal powders are also rolled into sheets or strip by special techniques. *See* POWDER METALLURGY.

Serope Kalpakjian

Bibliography. R. Pradhan and G. Lyudkovsky (eds.), *Hot- and Cold-Rolled Sheet Steels*, 1988; W. L. Roberts, *Cold Rolling of Steel*, vol. 2, 1978.

Metallic glasses

Metals and metallic alloys having an amorphous structure on the atomic scale. Although the word "glass" is commonly used to refer to the familiar transparent oxide glasses (such as silicate glasses used to make windowpanes), in a more general sense a glass is any solid obtained from a liquid that does not crystallize upon cooling. As a result, a glass retains an atomic-scale structure in which the atoms are more or less randomly arranged, similar to that of the liquid state. Because most metals crystallize quickly, special care must be taken with processing and alloy design to produce a metallic glass. Metallic glasses produced directly by quenching the liquid are quite similar in terms of structure and properties to amorphous alloys (of the same chemical composition) produced by other means, such as physical vapor deposition, irradiation, mechanical alloying, or solid-state reaction. *See* AMORPHOUS SOLID; CRYSTAL; GLASS; GLASS TRANSITION.

Preparation. When cooled below its equilibrium melting temperature, a liquid becomes unstable and there is a tendency for the stable crystalline phase (or phases) to form. Because of the significant structural change that occurs, crystallization does not happen

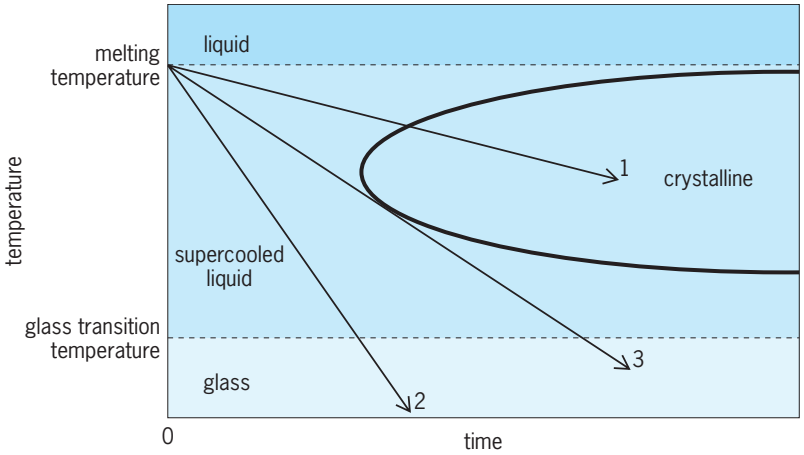

Fig. 1. Time required for crystallization of amorphous alloys below the equilibrium melting temperature.

instantaneously, and it is therefore possible to cool a liquid below its melting temperature without crystallizing it, for a limited time (**Fig. 1**). Such a liquid is said to be supercooled. Crystallization occurs slowly at temperatures just below the melting temperature, where it is difficult to nucleate new crystals, and at low temperature because atomic diffusion is slow. There is an intermediate temperature at which crystallization occurs most rapidly.

If a liquid is cooled slowly but continuously (curve 1 in Fig. 1), crystallization will begin at a temperature below the melting temperature and will be completed at a later time (and lower temperature). A liquid may be cooled so quickly that there is no time for crystallization to occur at all (curve 2). Instead, at a lower temperature called the glass transition temperature, the supercooled liquid transforms to an amorphous solid, or glass. The slowest rate at which the liquid can be cooled and form a glass, while avoiding crystallization, is called the critical cooling rate (curve 3). The critical cooling rate is a measure of how easily a material forms a glass, with low critical cooling rates indicating easy glass formation.

In materials with complicated crystal structures or in which the required atomic rearrangements are difficult, crystallization occurs slowly. Such materials, including many oxides and polymers, form glasses easily and have critical cooling rates on the order of 10^{-2} to 10^{-4} K/s (10^{-2} to 10^{-4} °F/s). Most metals have relatively simple crystal structures and crystallize readily. The critical cooling rate for pure metallic elements is on the order of 10^{13} to 10^{14} K/s (10^{13} to 10^{14} °F/s), rates which can only be achieved in very small volumes of material under specialized laboratory conditions. The critical cooling rate of an alloy can be much lower than that of a pure element, making production of metallic glasses feasible. The first true metallic glass, produced in 1960, was an alloy of gold and silicon and was cooled at a rate of approximately 2×10^6 K/s (4×10^6 °F/s). While still quite high, cooling rates on this order can be easily achieved by a variety of rapid solidification techniques, such as single-roller melt spinning. These techniques enabled the production of commercial quantities of metallic glasses, beginning in the 1970s. A significant limitation of the high cooling rates required was that the early metallic glasses could be produced only in forms in which one dimension of the material was small (typically 50–100 μm), such as ribbons, foils, wires, and powders.

Recently, more complex alloys have been developed in which the critical cooling rate is as low at 1 K/s (2°F/s) or even lower. These cooling rates can be achieved via common metallurgical processing techniques, such as die casting into a metallic mold, enabling the production of metallic glass components with minimum thicknesses in excess of 1 cm (0.4 in.). These easily processed alloys are commonly called bulk metallic glasses. Furthermore, a bulk metallic glass can be reheated to a temperature above its glass transition temperature, without crystallization, causing it to soften dramatically and allowing it to be easily formed into complex shapes in much the same way as is commonly done with amorphous thermoplastic polymers. However, because the crystalline form is still the stable state, all metallic glasses will crystallize if subjected to elevated temperatures for long times. *See* METAL CASTING.

Structure. The differences between the highly regular atomic positions characteristic of a crystal and the disordered structure of a liquid or glass can be explored using techniques based on the scattering of radiation such as x-rays, electrons, or neutrons. In crystalline materials, the regular atomic positions cause constructive or destructive interference among the scattered waves, leading to the formation of sharp, intense diffraction peaks at particular scattering angles that are characteristic of the structure of the material and of the wavelength of the radiation (**Fig. 2***a*). The peaks from amorphous materials by comparison are quite broad and weak (Fig. 2*b*). A truly random structure (such as that of a monatomic gas) would show no peaks at all, so the presence of the peaks is proof that there is short-range order (over a few atomic distances) in the structure of glasses and liquids. Fourier transformation of the scattering data produces a radial distribution function (**Fig. 3**), which describes the atomic environment around an

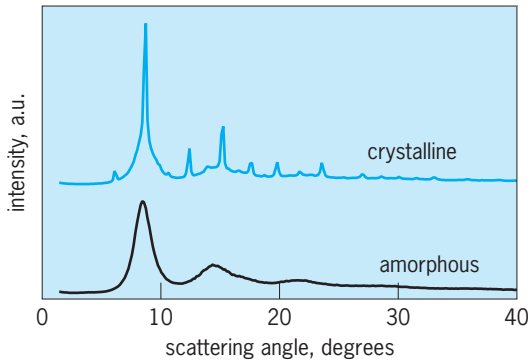

Fig. 2. X-ray diffraction patterns for crystalline and amorphous $Pd_{40}Ni_{40}P_{20}$ alloys.

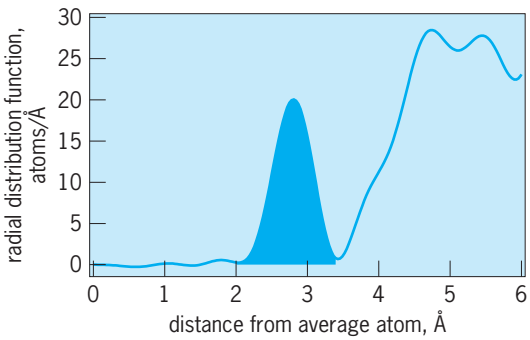

Fig. 3. Radial distribution function for amorphous $Pd_{40}Ni_{40}P_{20}$. The shaded region represents the first coordination shell of atoms around an average atom in the glass.

average atom in the material. From the radial distribution function, one can determine the number of atoms in the nearest-neighbor shell (the coordination number) as well as the distance between the average atom and its nearest neighbors. If the constituent elements are similar in size, then the average coordination number is approximately 12, just as it is in a close-packed crystalline structure. That metallic glasses are nearly as close packed is also reflected in their densities, which are typically just 0.5–2% lower than crystalline phases of similar chemical composition. *See* FOURIER SERIES AND TRANSFORMS; X-RAY DIFFRACTION.

Metallic glasses are also notable for their lack of structure over length scales longer than the atomic. Conventional crystalline alloys have microstructural features such as grains and precipitates with typical sizes of 0.1–100 μm; these features make fabrication of very small components difficult and lead to anisotropic mechanical properties that can be problematic. Because metallic glasses have no structural features on these length scales, they should be well suited for highly miniaturized devices, such as for micro-electro-mechanical systems (MEMS) [**Fig. 4**]. The lack of microstructure also means that metallic glasses have low ability to suppress vibrations (damping capacity). *See* MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS).

Mechanical properties. At low stresses and at temperatures well below the glass transition, metallic glasses display isotropic linear elastic behavior in which the strain ε, Eq. (1), is directly proportional to

$$\varepsilon = \frac{\sigma}{E} \tag{1}$$

the stress σ through Hooke's law, where the proportionality constant E (Young's modulus) describes the resistance to elastic (reversible) deformation, or stiffness. The disordered atomic structure of a metallic glass enables slight rearrangements of the atomic positions in response to an applied load, which would not be possible if the material were crystalline. This has the effect of reducing the stiffness slightly, so that E for metallic glasses is typically 20–30% lower than that for a crystalline alloy of the same composition. Distortion of the atomic bonds requires energy, and the elastic strain energy (per unit volume of material), given by Eq. (2), is called the resilience of the material. The high yield stress of metallic glasses, together with low modulus, means that metallic glasses can sustain large elastic strains of 2–3% (compared with 0.1–0.2% for a typical crystalline metal) and have exceptionally large values of resilience. *See* HOOKE'S LAW; STRESS AND STRAIN; YOUNG'S MODULUS.

$$U_{el} = \frac{1}{2}\sigma\varepsilon = \frac{\sigma^2}{2E} \tag{2}$$

At stresses beyond the elastic limit (or yield stress), metallic glasses (like crystalline alloys) experience permanent, or plastic, deformation. The degree of plastic deformation that a material can sustain before fracture is called ductility. Conventional crystalline alloys that are useful in structural applications combine high strength with good ductility. Both the strength and ductility of crystalline metals are determined largely by the properties of defects in the crystalline structure called dislocations. Although the dislocations reduce the yield strength of a crystalline metal below that predicted solely on the basis of the strength of the atomic bonds, interactions among the dislocations cause the material to strengthen (by strain hardening) as it deforms. This property gives many crystalline metals good ductility and toughness. In contrast, amorphous alloys have no dislocations (due to their disordered structure) and so they can be quite strong, often three to four times as strong as a crystalline alloy of similar composition (see **table**). However, the lack of dislocations means that metallic glasses do not strain-harden, and therefore have almost no ductility in tension. Instead of deforming uniformly, as crystalline metals do, plastic deformation in metallic glasses (at temperatures well below the glass transition temperature) is concentrated into very narrow shear bands. This problem can be partially overcome by reinforcing the metallic glass with particles of a ductile crystalline metal that impede shear band motion, although some strength is sacrificed. *See* PLASTIC DEFORMATION OF METAL.

Besides strength, fracture toughness—the ability of a material that may have flaws (such as small cracks) to support a load without fracturing—is of critical importance for many structural applications. Although plastic deformation (which requires energy and therefore enhances toughness) is concentrated into shear bands, the plastic deformation in each shear band can be extensive, and there can be considerable branching of the shear bands ahead of a crack tip. As a result, the fracture toughness (K_c) of some metallic glasses (20–90 MPa $\cdot$ m$^{1/2}$) is comparable to that of crystalline alloys, although for other amorphous alloys the fracture toughness can be quite low ($\sim$1 MPa $\cdot$ m$^{1/2}$), comparable to engineering ceramics. Furthermore, annealing some metallic glasses tends to make them more brittle, for reasons that are not well understood.

At elevated temperatures, approaching the glass transition temperature, the strength of a metallic glass drops and the deformation becomes homogeneous instead of being concentrated into shear

Mechanical properties of some amorphous alloys, together with properties of common crystalline engineering alloys

	Density, g/cm^3 (lb/in.^{3})	Young's modulus, GPa (10^6 psi)	Poisson's ratio	Yield strength, MPa (10^3 psi)	Tensile strength, MPa (10^3 psi)	Percent elongation	Fracture toughness, MPa · m$^{1/2}$ (ksi · in.$^{1/2}$)
			Amorphous alloys				
$Zr_{41}Ti_{14}Cu_{12.5}Ni_{10}Be_{22.5}$	6.12 (0.221)	101 (14.6)	0.34	1900 (275)	1900 (275)	2*	25 (23)
$Mg_{65}Cu_{25}Tb_{10}$	3.98 (0.144)	48 (7.0)	0.31	700 (100)	700 (100)	1.5*	2 (1.8)
$Pt_{57.5}Cu_{14.7}Ni_{5.3}P_{22.5}$	15.0 (0.541)	95 (14)	0.42	1400 (200)	1470 (213)	1.5*	90 (83)
			Crystalline alloys				
Steel (4340, quenched and tempered)	7.85 (0.283)	207 (30)	0.30	1620 (235)	1760 (255)	12	50 (46)
Titanium (Ti-6Al-4V, solutionized and aged)	4.43 (0.160)	114 (16.5)	0.34	1100 (160)	1170 (170)	10	55 (50)
Aluminum (7075-T6)	2.80 (0.101)	71 (10)	0.33	505 (73)	572 (83)	11	24 (22)

*Elastic deformation; under tensile loading, most metallic glasses fail without measurable plastic deformation.

bands. If heated above the glass transition, a metallic alloy becomes a fluid, the viscosity of which can be controlled over a wide range of temperature. Because the stresses required to cause deformation are low, either of these states can be useful for processing metallic glasses into complex shapes. *See* METAL; MECHANICAL PROPERTIES OF.

Electrical, optical, and magnetic properties. Like crystalline metals, metallic glasses are electrically and thermally conductive, due to the presence of conduction electrons. However, the disordered atomic structure and the high alloy content mean that metallic glasses are not especially good conductors, with an electrical resistivity on the order of $10^{-6}\,\Omega \cdot \mathrm{m}$, as compared to about $10^{-8}\,\Omega \cdot \mathrm{m}$ for copper at room temperature. However, the resistivity of metallic glasses does not vary strongly with temperature (except near absolute zero, where some amorphous alloys become superconducting). The presence of conduction electrons also allows metallic glasses to scatter and absorb incident light, giving metallic glasses the shiny luster typical of metals. Thus, unlike common oxide glasses, metallic glasses are not transparent to visible light. *See* ELECTRICAL CONDUCTIVITY OF METALS.

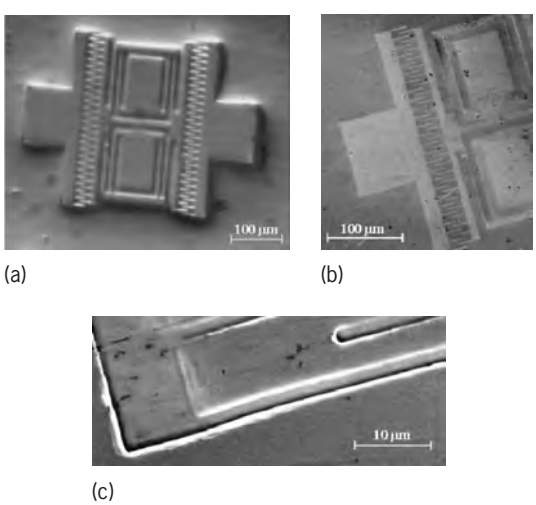

(a) (b) (c)

Fig. 4. Micrographs of (*a*) Replica of silicon micro-electro-mechanical device produced by warm-pressing into a metallic glass. (*b*, *c*) Submicrometer features are cleanly reproduced. [*Courtesy of J. Lewandowski and A. Shamimi Nouri, Case Western Reserve University (J. J. Lewandowski, M. Shazly, and A. Shamimi Nouri, Scripta Materialia, 54:337–341, 2006)*]

Some metallic glasses, particularly alloys based on iron, can be ferromagnetic. Although the presence of alloying elements decreases the saturation magnetization relative to that of a pure ferromagnetic element, for many applications the lack of crystal structure is a distinct advantage. In particular, amorphous alloys based on the transition metals (iron, cobalt, and nickel) have low coercivity because there are no crystalline grain boundaries (which can pin magnetic domain walls) and because there is no magnetocrystalline anisotropy. Also, because the electrical resistivity is high, eddy-current losses associated with high-frequency magnetization and demagnetization are minimized. Many amorphous alloys show large magnetostriction (shape change upon magnetization). *See* MAGNETIC MATERIALS.

Applications. Because the early metallic glasses could only be produced in limited shapes (such as thin ribbons and foils), applications were mostly limited to those making use of the advantageous magnetic properties of iron-based ferromagnetic alloys. For instance, these alloys can be used to make high-efficiency cores for electrical transformers, and the large magnetostriction is useful for making anti-shoplifting tags. The more recent development of bulk glass-forming alloys has opened up a much wider range of applications that make use of the unique mechanical properties of metallic glasses. In particular, the ability to soften the glass by heating it above the glass transition temperature makes it possible to form complex shapes and castings with thin sections. This, together with the high strength and stiffness (compared to polymers), makes metallic glasses attractive for applications such as cases for cellular telephones and other electronic devices. High strength is related to high hardness. This, together with the ability to form fine features, makes metallic glasses well suited for precision knife-edges, such as scalpels. Also, the high resilience and low

damping capacity mean that metallic glasses make good springs. One example is the use of zirconium-based metallic glasses as golf club heads.

Although current bulk metallic glasses tend to be rather expensive in terms of the raw cost of materials (including both the base elements such as zirconium, and the alloying elements such as beryllium), the processing costs can be low because the glasses can be easily processed at temperatures above the glass transition. Future applications are likely to be driven by these economic considerations as much as the unique mechanical properties of metallic glasses. Todd C. Hufnagel

Bibliography. T. H. Courntey, *Mechanical Behavior of Materials*, McGraw-Hill, 2000; S. R. Elliott, *Physics of Amorphous Materials*, Longman, London, 1984; B. Lemley, Glassy metal, *Discover*, pp. 46–51, April 2004; M. Telford, The case for bulk metallic glass, *Mater. Today*, pp. 36–43, March 2004.

Metallocenes

Bis-cyclopentadienyl derivatives of transition metals whose bonding involves overlap of *ns*, $(n - 1)d$, and *np* orbitals of the metal with molecular orbitals of appropriate symmetry of each cyclopentadienyl ring. The resulting complexes often possess two parallel rings (sandwich structure), but in some cases, for example those involving the titanium subgroup of metals, the rings are canted (**Fig. 1**). Numerous monocyclopentadienyl complexes (half-sandwich complexes) also exist, as do systems containing three and even four η^5-bonded cyclopentadienyl rings; these latter systems are not metallocene complexes. Moreover, compounds constituted of parallel ring systems that are not both η^5-bonding cyclopentadienyls (mixed-sandwich complexes) are not metallocene complexes. Metals in the periodic table commonly known to form metallocene complexes are titanium, zirconium, hafnium, vanadium, chromium, molybdenum, tungsten, manganese, iron, ruthenium, osmium, cobalt, rhodium, and nickel. *See* COORDINATION COMPLEXES.

Bonding. Each cyclopentadienyl ligand has three filled bonding molecular orbitals and two low-lying empty antibonding molecular orbitals. The valence levels of iron(0) consists of 4*s*, 3*d*, and 4*p* orbitals containing eight electrons. In the molecular-orbital bonding scheme for ferrocene, three filled molecular orbitals from each of the cyclopentadienyl rings overlap with the 4*s*, the three 4*p*, and two of the 3*d* orbitals of correct symmetry to form six new bonding orbitals. Further stabilization results from back-bonding from the filled *d* orbitals of the iron to the empty antibonding orbitals of the rings. This gives a total of 18 electrons in nine orbitals (18-electron rule), 10 of which come from the two cyclopentadienyl rings (in their conceptual neutral state) and 8 of which come from iron(0). Only six of the orbitals contribute significant bonding. *See* CHEMICAL BONDING.

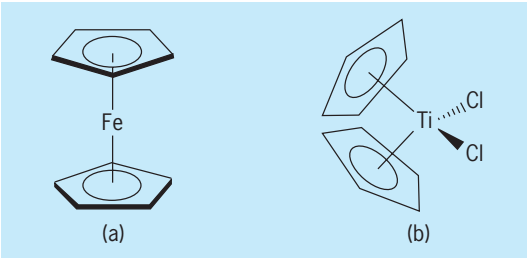

Fig. 1. Metallocene structures. (*a*) Staggered sandwich structure of ferrocene. (*b*) Canted cyclopentadienyl ring structure of titanocene dichloride. The distribution of the ligands about the Ti atom is tetrahedral.

Structure and properties. The first transition-metal series of metallocenes from chromocene through nickelocene forms a continuous series of solid solutions; each metallocene possesses a melting point of 173°C (343°F). In the crystalline state the cyclopentadienyl rings of ferrocene when viewed down the axis joining their centers exhibit the staggered conformation, but in the gas phase the rings are eclipsed. For ruthenocene and osmocene, the conformations in the crystalline state are eclipsed, possibly reflecting the greater ring-metal bond distances. For ferrocene, the gas-phase barrier to rotation has been found to be only 4 ± 1 kJ · mol^{-1}. Ferrocene structurally possesses a C—C bond distance of 0.140 nanometer and a C—Fe bond distance of 0.204 nanometer. Ferrocene can be sublimed under reduced pressure at 100°C (212°F) and is stable to temperatures >400°C (750°F).

Ferrocene is somewhat stable to H_2SO_4; with HNO_3 it forms a cation known as the ferricinium ion. Cobaltocene can be easily oxidized to a very stable cation, $(C_5H_5)_2Co^+$, which can be recovered unchanged from aqua regia. Ferrocene is oxidatively stable, cobaltocene and nickelocene much less so, but they can be handled for brief periods in air. Chromocene, on the other hand, is dangerously pyrophoric when exposed to air. *See* SOLID SOLUTION.

Preparation. The best-known preparation of metallocenes, here exemplified by the formation of ferrocene, is by combination of the appropriate transition-metal salt with the cyclopentadienyl Grignard reagent [reaction (1)]. This Grignard reagent

$$2C_5H_5MgBr + FeCl_2 \longrightarrow (C_5H_5)_2Fe + 2MgBrCl \qquad (1)$$

can be formed by an exchange reaction between cyclopentadiene and C_6H_5MgBr. Useful modifications of this procedure involve treatment of cyclopentadiene with sodium or sodium hydride in tetrahydrofuran to form the sodium salt, followed by addition of a transition-metal halide, carbonyl, or related derivative. The thallium salt, C_5H_5Tl, can also be used. A milder set of conditions is possible; for example, a strong organic base can be used as an acceptor for HCl [reaction (2)].

$$2C_5H_6 + CoCl_2 + 2(C_2H_5)_2NH \longrightarrow (C_5H_5)_2Co + 2(C_2H_5)_2NH_2{}^+Cl^- \qquad (2)$$

See GRIGNARD REACTION.

Reactions. The reactions of metallocenes can be divided into two classes: the first is typified by the

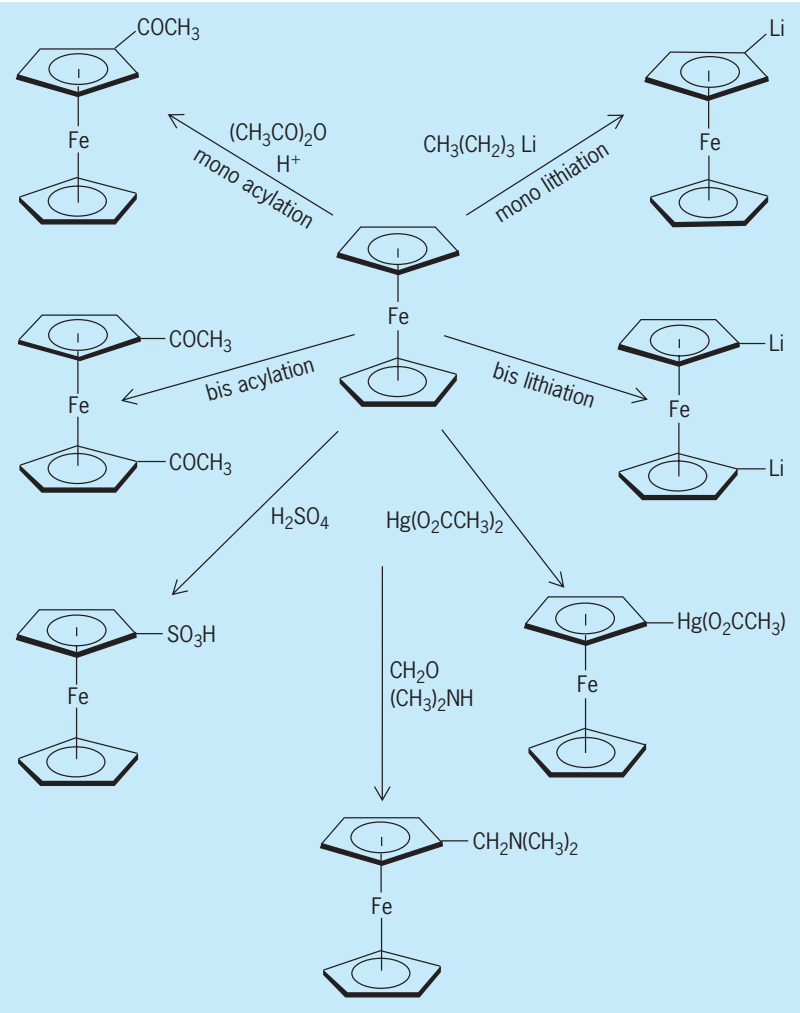

Fig. 2. Aromatic reactions of ferrocene.

iron triad, and comprises essentially the reactions of aromatic molecules; the second consists of the reactions of the other metallocenes where the 18-electron rare-gas configuration is not found. Reactions in these latter systems often lead to a product where the 18-electron rule is obeyed.

Ferrocene is a very electron-rich system, and undergoes electrophilic substitution with great rapidity. For example, acylation proceeds about 10^6 times faster than that of benzene under similar conditions. Ferrocene also undergoes several other typical aromatic substitution reactions besides acylation, including sulfonation, dimethylaminomethylation (Mannich reaction), metalation, and the like (**Fig. 2**). Bis substitution tends to factor a product where each ring is monosubstituted although several cases are known where two substituents are introduced into one ring. Ferrocene is oxidized and deactivated under conditions for nitration and halogenation.

The other metallocenes often react as if to relieve the instability of their electronic configurations; those that are electron-rich lose coordination sites on the ligand (or the entire ligand) in order to reduce their outer-shell electron count to 18, the krypton configuration. Reactions in accord with this formulation are shown in reactions (3) and (4). In each case,

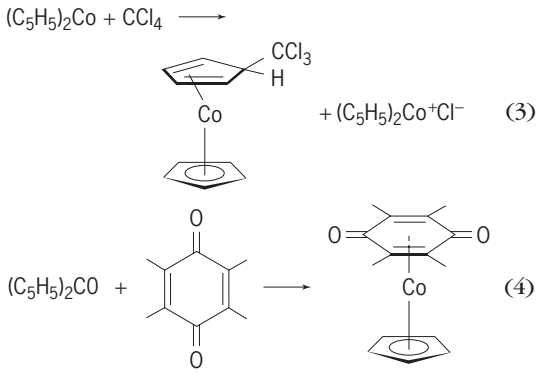

cobalt has gone formally from a Co(II) species to a Co(I) species, and in each of the products cobalt has the 18-electron configuration. Nickelocene reacts in similar fashion to reduce its coordination sphere by a factor of 2 to attain the 18-electron rare-gas configuration [reactions (5) and (6)].

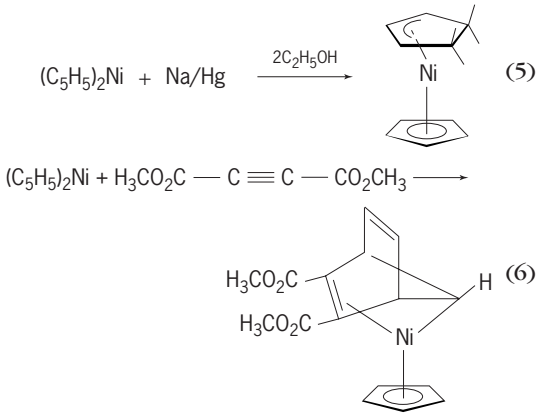

A somewhat similar rationale can be used for the few reactions of manganocene that have been studied. Manganocene, in contrast to the other metallocenes, is only a 17-electron system and so reacts with electron-rich systems such as biphenyl radical anion, resulting in displacement of the cyclopentadienyl ring [reaction (7)].

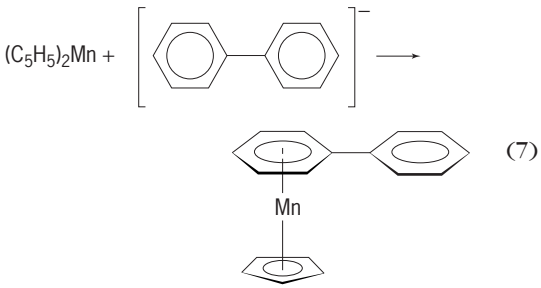

Vanadocene is so electrophilic that it inserts into π bonds in carbenelike fashion [reaction (8)].

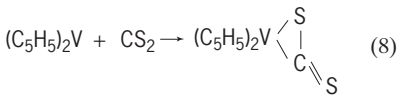

Structures. Metallocenes can possess both axial and chiral atom asymmetry. For example,

α-ketotetramethyleneferrocene (**1**) was the first fer-

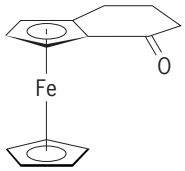

(**1**)

rocene resolved; a number of chiral transition-metal metallocenes in the titanium subgroup such as (**2**)

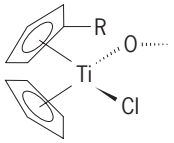

(**2**)

also are known. Other novel structural features are illustrated by the bis-complexation of nickel by pentalene to form (**3**) and by triple-decker metallocenes of type (**4**).

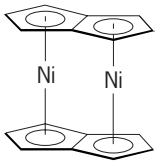

(**3**)

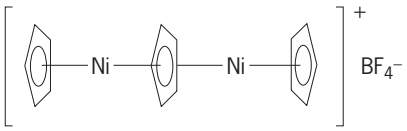

(**4**)

Decasubstituted ferrocenes $[C_5(CH_3)_5]_2Fe$ and $[C_5(C_6H_5)_2]Fe$ and pentasubstituted metallocenes such as $C_5(C_6H_5)_5FeC_5H_5$ have been prepared; the pentasubstituted cyclopentadienyl ligands constitute an interesting extension of the possibilities of metallocene chemistry, offering greater air and thermal stability and other property modification.

Uses. Decamethylzirconocene reacts with nitrogen to form complex (**5**). Acid hydrolysis of this

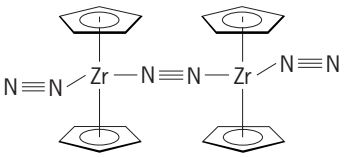

(**5**)

complex provides ammonia and hydrazine. Although not commercial, this species has the potential for acting as a nitrogen-fixing complex that could be used directly in the soil. Chromocene can be reacted with alumina to make a polymerization catalyst for ethylene. Ferrocene and some alkyl-substituted ferrocenes have been used as moderators in high-temperature combustions such as occur in solid rocket fuels. Lastly, a cyclopentadienyl complex, $CH_3C_5H_4Mn(CO)_3$, briefly replaced tetraethyllead as an octane booster and antiknock agent in liquid fuels. *See* ORGANOMETALLIC COMPOUND.

Donald W. Slocum

Bibliography. J. Collman et al., *Principles and Applications of Organotransition Metal Chemistry*, 2d ed., 1987; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; R. H. Crabtree, *The Organometallic Chemistry of the Transition Metals*, 3d ed., 2001; F. G. A. Stone and R. West (eds.), *Advances in Organometallic Chemistry*, vol. 10, 1972; M. Tsutsui et al., *Introduction to Metal π-Complex Chemistry*, 1970.

Metallochaperones

A family of proteins that shuttle metal ions to specific sites within a cell. The target sites for metal delivery include a number of metalloenzymes, or proteins that bind metal ions, such as copper, zinc, or iron, and use these ions as cofactors to carry out essential biochemical reactions. It was long thought that metalloenzymes could acquire their metal ion without the assistance of an accessory metallochaperone. The metal ion was believed to freely diffuse within the cell until it contacted its metalloenzyme target, at which point the metal ion inserted itself into the protein. However, several discoveries in the late 1990s led to a radical change in the thinking of how metalloenzymes acquire their metal ion. The insertion of the metal is now known to require an auxiliary protein, the metallochaperone. Metallochaperones escort the ion to a specific intracellular location and facilitate incorporation of the metal into designated metalloenzymes. *See* BIOINORGANIC CHEMISTRY; CELL (BIOLOGY).

Although the bulk of current knowledge on metallochaperones is restricted to copper, it is reasonable to assume that a distinct class of proteins is responsible for the incorporation of other metal ion cofactors into metalloenzymes. The heavy metals iron, manganese, and zinc are also required as cofactors for metalloenzymes, and like copper, these metals are presumably present at trace quantities inside living cells. It is therefore highly likely that new families of metallochaperones will emerge in the near future.

Copper chaperones. Among the metallochaperones that have been studied in detail are a family of three copper chaperones. These molecules operate in eukaryotic (nucleated) cells to direct copper to distinct intracellular locations: the mitochondria, the secretory pathway, and the cytosol. All three copper chaperones were originally identified through genetic studies of bakers' yeast, *Saccharomyces cerevisiae*. Their discoveries led to rapid identification of the human homolog proteins that are closely related in function and structure. One of the first copper chaperones identified, COX17, is a small protein that specifically directs copper to the mitochondria. Mitochondria, often called the power plants of the cell, are the sites where energy is produced from

respiration. The copper delivered by COX17 is inserted into the metalloenzyme cytochrome oxidase, needed for respiration. A second copper chaperone identified was ATX1, which carries copper to the secretory pathway, a cellular compartment that functions to shuttle proteins toward the cell surface. The metal delivered by ATX1 is incorporated into copper enzymes destined for the cell surface or the extracellular milieu. The most recently identified copper chaperone is CCS, which specifically delivers copper to a single metalloenzyme, superoxide dismutase. This copper-requiring enzyme is located in the soluble cytosolic compartment of the cell and acts to detoxify harmful reactive oxygen species. *See* MITOCHONDRIA.

The COX17 copper chaperone was discovered by D. M. Glerum and A. Tzagoloff in 1996 in their search for yeast genes involved in the assembly of cytochrome oxidase. Each molecule of COX17 can carry two ions of copper by coordination of the metal to cysteine residues (sulfur-containing amino acids) present in the COX17 polypeptide. COX17 has been localized to the cytosolic compartment of the cell as well as to the intermembrane space of the mitochondria, in part explaining its ability to shuttle copper ions across the cytosol to the mitochondria.

ATX1 was originally identified by S. Lin and V. C. Culotta in 1995 as a gene which when expressed to very high levels in yeast would prevent oxidative damage. However, this apparent antioxidant activity was later found not to be the primary function of ATX1. Rather, this protein acts to specifically shuttle copper ions across the cytosol to an intracellular copper transporter located in the secretory pathway. In humans, this intracellular copper transporter or copper pump is known as the Wilson or Menkes protein. A number of copper-requiring metalloenzymes that are exported from the cell rely on the Menkes/Wilson proteins and human ATX1 to acquire their copper. Included in this list are copper enzymes involved in connective tissue formation, brain function, and iron use. The ATX1 copper chaperone is known to bind a single copper ion via two cysteine residues present in the ATX1 amino acid sequence MXCXXC (M = methionine, X = any amino acid, C = cysteine). This copper-binding site is not unique to ATX1, but is also found in the Wilson/Menkes copper transporters that serve as targets for ATX1. Studies by T. V. O'Halloran and colleagues revealed that ATX1 physically interacts with the copper transporter and deposits copper by transferring the metal from the MXCXXC site in the copper chaperone to the same site in the copper pump. *See* AMINO ACIDS.

The CCS copper chaperone discovered by Culotta, J. D. Gitlin, and coworkers in 1997 specifically inserts copper into a single enzyme, the soluble copper- and zinc-requiring superoxide dismutase (SOD1). Interestingly, CCS contains the same MXCXXC copper-binding site found in ATX1, but in addition the CCS protein contains a region that is quite similar to its target of copper delivery, SOD1. CCS is believed to specifically recognize SOD1 by virtue of the SOD1-like region in the CCS protein. This region of CCS has been shown to physically interact with SOD1 and thereby serve as the "molecular glue" that holds SOD1 in place while the MXCXXC copper-binding region of CCS mediates copper transfer.

Although the three copper chaperones vastly differ in their respective targets of copper delivery, they are similar in several important aspects. These molecules are small proteins that bind copper by cysteine residues. They predominantly localize to the cytosol and function to deliver copper from the cell surface to a specific target deep within the cell. Genetic studies in yeast have revealed a family of cell-surface copper transporters (CTRs) that are responsible for the uptake of copper into the cell and delivery of the metal to the copper chaperones. It is presently unclear whether the copper chaperones go directly to the "front door" to pick up copper from the CTRs or whether a "middle molecule" exists that directly supplies the copper chaperones with their metal ion "cargo." An additional property common to all copper chaperones is that these proteins are necessary only when intracellular levels of copper are low. When cells are exposed to high concentrations of copper, the metalloenzymes can acquire their copper in a chaperone-independent manner. Yet intracellular copper is normally present at exquisitely low levels, and activation of copper enzymes is wholly dependent upon copper chaperones. Copper not only is an essential nutrient but also is quite toxic to living cells, and elaborate detoxification mechanisms prevent the free metal ion from accumulating to any substantial degree. The copper-requiring metalloenzymes cannot compete for these vanishingly low levels of available metal, explaining the requirement for the copper metallochaperones (see **illus.**).

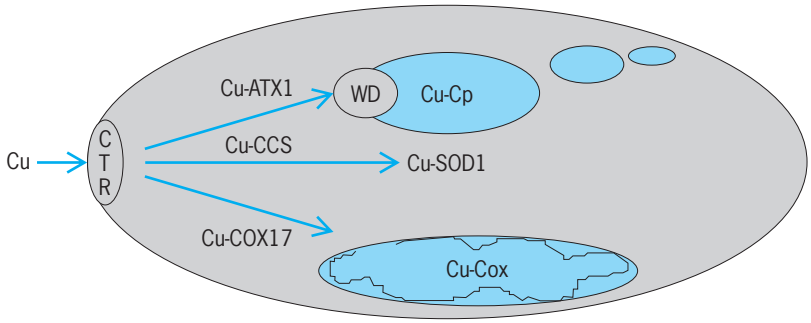

A family of copper metallochaperones direct trafficking of copper ions inside the cell. The schematic shows three distinct pathways of copper movement within a eukaryotic cell. Copper is first transported into the cell by the CTR and is then bound as cargo to one of three copper metallochaperones: (1) ATX1 delivers the metal to an intracellular copper transporter [such as the human Wilson disease (WD) gene product], which in turn pumps the copper into the secretory pathway where the metal is inserted into copper enzymes destined for export from the cell [such as human ceruloplasm (Cp)]. (2) CCS specifically delivers copper to the cytosolic superoxide dismutase enzyme (SOD1). (3) COX17 carries copper to the mitochondria where the metal is incorporated into cytochrome oxidase (Cu-Cox).

Implications for human disease. The copper metallochaperones may have important relevance to specific diseases in humans. For example, a number of neurodegenerative diseases have been associated with deficiencies in mitochondrial cytochrome oxidase, and it is conceivable that mutations in the human COX17 gene may contribute to one or more

of these disorders. With regard to ATX1, the intracellular copper transporters that serve as targets of ATX1 action have already been linked to two important disorders of copper metabolism in humans. Wilson disease results from mutations in the Wilson copper transporter and primarily presents as chronic copper toxicity in the liver. Menkes disease, resulting from mutations in the Menkes copper transporter, is a fatal disorder associated with severe copper deficiency. It is therefore possible that additional inherited disorders of copper metabolism result from abnormalities in human ATX1 itself. Moreover, the CCS copper chaperone may have important implications regarding the fatal motor neuron disease amyotrophic lateral sclerosis (ALS), commonly known as Lou Gehrig's disease. A subset of familial cases of amyotrophic lateral sclerosis have been linked to mutations in the human gene encoding SOD1. This disease does not result from a loss of SOD1 activity but from a new toxic property unique to the mutant enzyme. Although the cause of SOD1-linked amyotrophic lateral sclerosis is not clear, evidence to date suggests that disease results from the toxic copper ion bound to the mutant SOD1 metalloenzyme. Thus, patients suffering from SOD1-linked amyotrophic lateral sclerosis could conceivably benefit from treatments that abrogate the action of the CCS copper chaperone. Valeria Cizewski Culotta

Bibliography. J. R. Conner (ed.), *Metals and Oxidative Damage in Neurological Disorders*, Plenum Publishing, New York, 1997; V. C. Culotta and J. Gitlin, Disorders of copper transport, in C. R. Sriver et al. (eds.), *Molecular and Metabolic Basic of Inherited Disease*, McGraw-Hill, 1999; M. C. Linder, *Biochemistry of Copper*, Biochemistry of the Elements series, vol. 10, Plenum Publishing, New York, 1991.

Metallography

The study of the structure of metals and alloys by various methods, especially light and electron microscopy. Light microscopy of metals is conducted with reflected light on surfaces suitably prepared to reveal structural features. The method is often called optical microscopy or light optical microscopy. A resolution of about 200 nanometers and a linear magnification of at most 2000× can be obtained. Electron microscopy is generally carried out by the scanning electron microscope (SEM) on specimen surfaces or by the transmission electron microscope (TEM) on electron-transparent thin foils prepared from bulk materials. Magnifications can range from 10× to greater than 1,000,000×, sufficient to resolve individual atoms or planes of atoms.

Metallography serves both research and industrial practice. Light microscopy has long been a standard method for observing the morphology of phases resulting from industrial processes that involve phase transformations, such as solidification and heat treatment, and plastic deformation and annealing. Microscopy, both light and electron, is also indispensable for the analysis of the causes of service failures of components and products.

Light microscopy. Microstructural features observed in photomicrographs include the size and shape of the grains (crystals) in single-phase materials (**Fig. 1***a* and **Fig. 2***a*), the structure of alloys containing more than one phase such as steel (Fig. 2*b, c*), the effects of deformation (Fig. 1*b*), microcracking (Fig. 1*c*), and the effects of heat treatment (Fig. 2*b, c*). Other structural features investigated by light microscopy include the morphology and size of precipitates, compositional inhomogeneities (microsegregation), microporosity, corrosion, thickness and structure of surface coatings, and microstructure and defects in welds.

The selection of representative specimens for metallographic examination is important. For example, if a part such as a forged shaft has directional properties, transverse and longitudinal sections should be examined. Sheet, wire, electronic components, and other small specimens are mounted in fixtures or plastic mounts.

A specimen is initially ground on a series of abrasive papers of decreasing grit size, or on laps. It is then polished on one or more cloth-covered wheels with an abrasive such as aluminum oxide or diamond dust. Automated polishing equipment is also used. The specimen surface becomes progressively more scratch-free and mirrorlike; however, polishing, even when properly carried out, leaves a thin layer of distorted metal. *See* ABRASIVE.

Electrolytic polishing, which consists of controlled anodic dissolution, is an alternative to mechanical polishing. Once the operating conditions have been established, electrolytic polishing is a simple operation; and the polished surface may be superior to a mechanically polished surface, particularly for soft metals.

Polished specimens reveal only a few structural features, such as inclusions, microcracks, and microporosity. Etching with an appropriate reagent is generally necessary to reveal microstructural details, and it also removes the distorted layer from mechanically polished specimens. Etching reagents are mainly acids and bases (and occasionally other compounds) dissolved in water or alcohol. Most reagents act by dissolution, a few by selective deposition of reaction products (staining). In a single-phase material, etching is effective because it attacks different parts of the structure selectively. Grain-boundary regions may dissolve preferentially to the body of the grains (Fig. 2*a*); or adjoining grains may develop facets of different orientation, which reflect differing amounts of light (Fig. 1). In multiphase alloys, the phases are attacked selectively (Fig. 2*b, c*).

Etching may also be carried out electrolytically by making the specimen the anode. Etching by heating in a vacuum or oxidizing atmosphere has proved useful. Cathodic etching, in which the specimen is bombarded by ions of an inert gas, is effective with some materials that resist etching by other methods.

In the standard metallographic microscope, a beam of light normal to the surface illuminates the

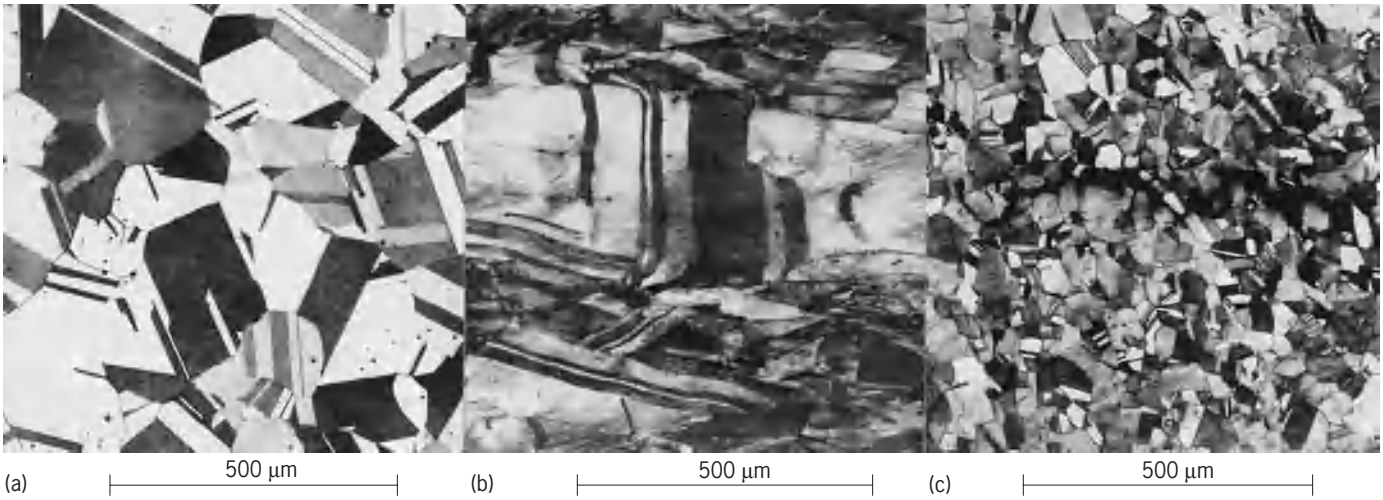

Fig. 1. Photomicrographs of typical microstructures of brass (70% Cu–30% Zn). (*a*) Annealed; grains with twin bands. (*b*) Reduced 40% by cold rolling; distorted grains. (*c*) Stress-corroded, with horizontal crack. *(Courtesy of W. R. Johnson)*

specimen. The objective lens of the microscope serves as a condensing system for the incident light and as the image-forming system for the light reflected by the specimen; this is known as bright-field illumination. In dark-field illumination, a hollow beam of light converges on the specimen; a mirrorlike surface appears dark, whereas a rough surface appears bright. Polarized light reveals differences in constituents that are optically anisotropic; useful (and often beautiful) color variations can be observed with white light. Differential interference contrast is increasingly used; optical interference permits quantitative investigation of surface contours, and color is produced by inserting a sensitive tint plate in the illumination system.

Hot-stage microscopy allows many phase transformations to be followed as they occur. Most specimens must be protected from oxidation by a protective atmosphere or vacuum, and the objective lens must be kept cool. Reflecting objectives are particularly suitable because of their long working distance. Thermal etching reveals the structural features. *See* REFLECTING MICROSCOPE.

Images are normally photographed in black and white or color. Video microscopy is common, and digital image acquisition and image processing methods are also used.

Quantitative microscopy methods can be used for the quantitative characterization of microstructures. The basic problem to be solved is the conversion of measurements on a two-dimensional section into quantities representing the three-dimensional structure of the material. The quantities most frequently measured are grain size, relative amounts of phases in a multiphase material, total area of grain boundaries, particle size of a dispersed phase, and spacing of lamellae and dispersed particles. The use of desktop computers, special software packages, and digital image processing have greatly advanced the speed and accuracy of quantitative microscopy methods. Nevertheless, valid quantitative description of microstructures requires close attention to the selection of representative microsections from the original material or product. *See* GRAIN BOUNDARIES; OPTICAL MICROSCOPE.

Electron microscopy. The electron microscope offers improved depth of field and higher resolution than the light microscope, as well as the possibility of in-place spectroscopy techniques. The scanning electron microscope images the surface of a

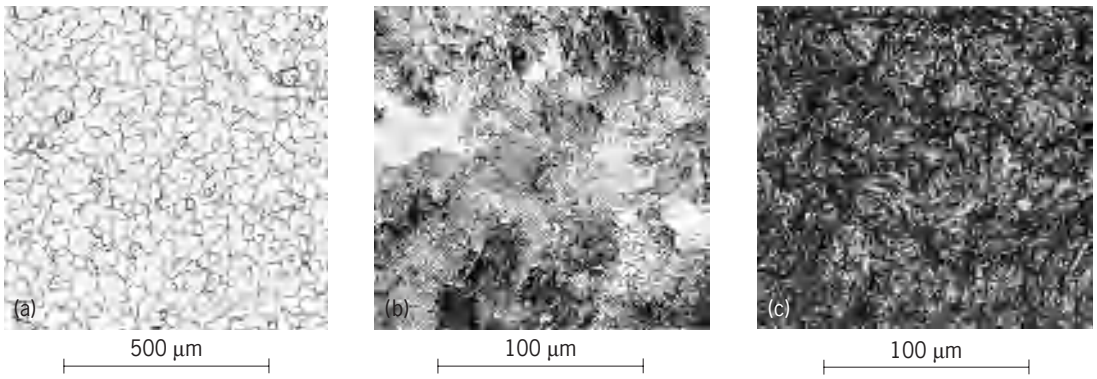

Fig. 2. Photomicrographs of typical structures of iron and steel. (*a*) High-purity iron. (*b*) Steel containing 0.85% C, slowly cooled; pearlite structure, consisting of alternating layers of light ferrite and dark cementite. (*c*) Steel containing 0.80% C, quenched and tempered; dark plates of tempered martensite. (*Courtesy of W. R. Johnson*)

material, while the transmission electron microscope reveals internal microstructure. Images produced by the scanning electron microscope are generally easier to interpret; in addition, the instrument operates at lower voltages, offers lower magnification, and requires less specimen preparation than is necessary for the transmission electron microscope. Consequently it is important to view a specimen with light microscopy and often with the scanning electron microscope before embarking on transmission electron microscopy.

However there are some disadvantages. Electron microscope specimens are viewed under vacuum, the instruments cost significantly more than light microscopes, electron beam damage is always a danger, and representative sampling becomes more difficult as the magnification increases. *See* ELECTRON MICROSCOPE.

Scanning electron microscope. The scanning electron microscope scans a fine beam of electrons across the surface of a specimen; and images are formed by collecting either the electrons that are backscattered from the surface or, more usually, secondary electrons knocked out of the surface atoms. The electron signal is then used to modulate a television screen scanning synchronously with the electron beam. The images contain topographic information, so a major use of the scanning electron microscope is to study rough specimens, such as fracture surfaces and semiconductor devices, which cannot be observed in the light microscope with the whole field of view in focus. Fortunately, secondary electron emission mimics the behavior of reflected light, and the human eye can generally interpret the images produced by a scanning electron microscope in a straightforward manner.

The best scanning electron microscopes offer maximum useful magnifications of about 100,000×. Such a high-resolution instrument, which uses a field emission electron gun, is able to resolve surface detail on a dimensional scale of 1–2 nm. However, most scanning electron microscopes are used in the range of 1000× to 20,000×, resolving detail at the 5–10-nm level with excellent depth of field and improved resolution compared to the light microscope.

Specimen preparation for scanning electron microscopes may consist of metallographic polishing if the topography is unimportant, or if direct comparison has to be made with the light microscope. However, the specimen can usually be observed in its as-received condition with minimal cleaning, if necessary. If electron-beam charging occurs because the specimen is insulating or contaminated, the image resolution degrades catastrophically, so a thin conductive coating is evaporated on the specimen surface (**Fig. 3**). However, field emission gun scanning electron microscopes can be operated at very low voltages (<1 keV) to view insulators such as polymers without conductive coatings, but still resolve fine detail (<10 nm). Most scanning electron microscopes are digitally controlled, so the images reside in a frame store and may be digitally enhanced and processed prior to printing in order to maximize the contrast information. Digital images can easily be analyzed through standard quantitative metallographic software. *See* DIGITAL COMPUTER.

In addition to secondary and backscattered electron images, in the scanning electron microscope it is also possible to image magnetic domains by the way they bend the electrons, to observe crystal orientation by how the electrons channel between crystal planes, and to trace how voltage pulses travel around a semiconductor device and how the electron beam itself changes the semiconducting properties. The principles of scanning imaging, developed for the scanning electron microscope, have also been extended to other metallographic techniques such as scanning ion microscopy, scanning acoustic microscopy, and scanning tunneling microscopy. *See* ACOUSTIC MICROSCOPE; FIELD-EMISSION MICROSCOPY; SCANNING ELECTRON MICROSCOPE; SCANNING TUNNELING MICROSCOPE.

Transmission electron microscope. The transmission electron microscope is primarily designed to present images of the highest possible resolution by observing how high-energy (typically 100–300 keV) electrons are scattered as they traverse a thin (<100-nm) specimen. Thin specimens are made by electropolishing conductive metals, ion-beam-thinning insulating ceramics and two-phase structures such as composites, and ultramicrotoming polymers, none of which is a trivial task. The thinner the specimen, the better the image quality, and it is relatively routine (at magnifications >1,000,000× and a resolution of >0.2–0.3 nm) to reveal the arrangements of atoms at crystal defects such as dislocations and grain boundaries.

More often, the transmission electron microscope is used at lower magnifications to study the distribution of crystal defects and phases, which are

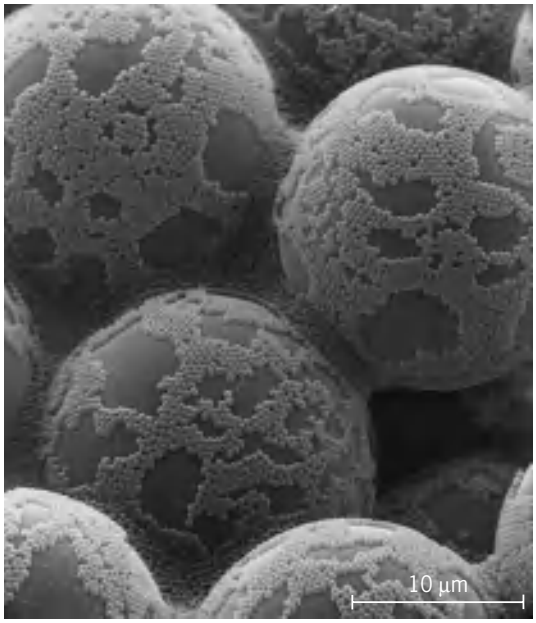

Fig. 3. Scanning electron micrograph of polymer latexes. The acrylic latex spheres of 400-nm diameter are on the surface of 19-μm polystyrene latex. (*Courtesy of O. L. Shaffer*)

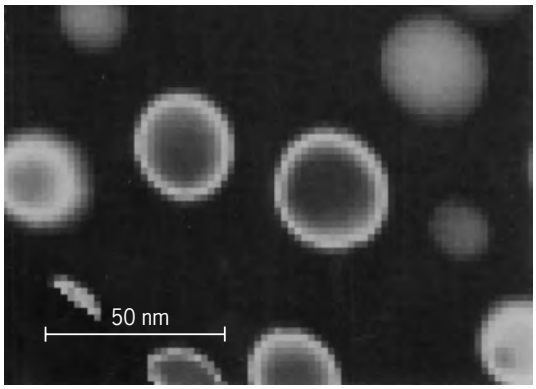

Fig. 4. Electron energy loss spectrum (EELS) image showing spherical precipitates of Al_3Li in an aluminum-lithium alloy. The intensity of each pixel of the digital image represents the energy lost by electrons passing through that region of the thin specimen; the energy loss can then be related to the amount of Li present. (*Courtesy of J. A. Hunt*)

revealed by the different ways they diffract the electron beam (diffraction contrast). Therefore, electron diffraction can be performed simultaneously to determine the crystal structure and orientation of various microstructural features. If the electron beam is converged to a probe, the contrast in the convergent beam diffraction pattern can reveal the lattice parameter, point group, and space group of the thin crystal. The transmission electron microscope image and diffraction pattern can be obtained from the same area of the specimen. *See* CRYSTAL DEFECTS; CRYSTAL STRUCTURE; ELECTRON DIFFRACTION.

Transmission electron microscope images are much more difficult to interpret than those obtained with the scanning electron microscope. The transmission electron microscope image is always a projected image of a three-dimensional sample, and the human eye is not accustomed to viewing in projection. In addition, the diffraction contrast in transmission electron microscope images varies with crystal orientation, which requires careful interpretation. However, in noncrystalline materials, changes in thickness and mass are responsible for contrast, and this is easier to understand. Transmission electron microscopes that operate at high voltage (>1 MeV) are rare; but they offer the possibility of very high resolution, the study of thicker specimens, and the option of carrying out in-place experiments such as straining, heating, or cooling the specimen, in order to watch changes occur dynamically.

Spectrometry. Electron microscopes offer much more information than images and diffraction patterns. The ionizing nature of electron irradiation means that x-ray spectrometry and electron spectrometry, both powerful tools in their own right, can be performed in both scanning electron microscopy and transmission electron microscopy. The various signals detected spectroscopically can also be used to form images of the specimen, which reveal elemental distribution among other information. In particular, the characteristic x-ray signal can be detected and processed to map the elemental distribution quantitatively on a micrometer scale in the scanning electron microscope and a nanometer scale in the transmission electron microscope. Electron spectroscopic signals permit not only elemental images to be formed (**Fig. 4**) but also images that reveal local changes in bonding, dielectric constant, thickness, band gap, and valence state. *See* ELECTRON SPECTROSCOPY.

Limitations. Ultimately the information gathered in both the scanning electron microscope and the transmission electron microscope is limited by electron damage, which is also a consequence of the ionizing nature of the beam. Thus the characteristic of electrons that creates the many useful signals for imaging also places limits on the applicability of the instruments. *See* METALLURGY.

David A. Thomas; David B. Williams

Bibliography. American Society for Metals, *Metals Handbook*, vol. 9; *Metallography and Microstructures*, 9th ed., 1985; J. I. Goldstein et al., *Scanning Electron Microscopy and X-ray Microanalysis*, 2d ed., 1992; L. Reimer, *Transmission Electron Microscopy*, 4th ed., 1997; G. F. Vander Voort, *Metallography, Principles and Practice*, 1999; D. B. Williams et al., *Images of Materials*, 1991.

Metalloid

An element which exhibits the external characteristics of a metal but behaves chemically both as a metal and as a nonmetal. Arsenic and antimony, for example, are hard crystalline solids that are definitely metallic in appearance. They may, however, undergo reactions that are characteristic of both metals and nonmetals. Certain of their oxides dissolve in either acids or bases, and are said to be amphoteric in character because they behave either as a base or an acid. Many elements form compounds that are amphoteric. However, only when this dualistic chemical behavior is very marked and the external appearance metallic is the element commonly called a metalloid. *See* METAL; NONMETAL.

Francis J. Johnston

Metallurgy

The technology and science of metallic materials. Metallurgy as a branch of engineering is concerned with the production of metals and alloys, their adaptation to use, and their performance in service. As a science, metallurgy is concerned with the chemical reactions involved in the processes by which metals are produced and the chemical, physical, and mechanical behavior of metallic materials.

Metallurgy has played an important role in the history of civilization. Metals were first produced more than 6000 years ago. Because only a few metals, principally gold, silver, copper, and meteoric iron, occur in the uncombined state in nature, and then only in small quantities, primitive metallurgists had to discover ways of extracting metals from their ores. Fairly large-scale production of some metals was carried

out with technical competence in early Near Eastern and Mediterranean civilizations and in the Middle Ages in central and northern Europe. Basic metallurgical skills were also developed in other parts of the world.

The winning of metals would have been of little value without the ability to work them. Great craftsmanship in metalworking developed in early times: the objects produced included jewelry, large ornamental and ceremonial objects, tools, and weapons. It may be noted that almost all early materials and techniques that later had important useful applications were discovered and first used in the decorative arts. In the Middle Ages metal working was in the hands of individual or groups of craft workers. The scale and capabilities of metal working developed with growth of industrial organizations. Today's metallurgical plants supply metals and alloys to the manufacturing and construction industries in a variety of forms, such as beams, plates, sheets, bars, wire, and castings. Technologies such as communications, nuclear power, and space exploration continue to demand new techniques of metal production and processing.

The field of metallurgy may be divided into process metallurgy (production metallurgy, extractive metallurgy) and physical metallurgy. According to another system of classification, metallurgy comprises chemical metallurgy, mechanical metallurgy (metal processing and mechanical behavior in service), and physical metallurgy. The more common division into process metallurgy and physical metallurgy, which is adopted in this article, classifies metal processing as a part of process metallurgy and the mechanical behavior of metals as a part of physical metallurgy.

Process metallurgy. Process metallurgy, the science and technology used in the production of metals, employs some of the same unit operations and unit processes as chemical engineering. These operations and processes are carried out with ores, concentrates, scrap metals, fuels, fluxes, slags, solvents, and electrolytes. Different metals require different combinations of operations and processes, but typically the production of a metal involves two major steps. The first is the production of an impure metal from ore minerals, commonly oxides or sulfides, and the second is the refining of the reduced impure metal, for example, by selective oxidation of impurities or by electrolysis. Process metallurgy is continually challenged by the demand for metals which have not been produced previously or are difficult to produce; by the depletion of the deposits of the richer and more easily processed ores of the traditional metals; and by the need for metals of greater purity and higher quality. The mining of leaner ores has greatly enhanced the importance of ore dressing methods for enriching raw materials for metal production. Several nonferrous metals are commonly produced from concentrates. Iron ores are also treated by ore dressing. *See* ELECTROMETALLURGY; HYDROMETALLURGY; ORE DRESSING; PYROMETALLURGY; PYROMETALLURGY, NONFERROUS.

Process metallurgy today mainly involves large-scale operations. A single blast furnace can produce crude iron at the rate of 3000–11,000 tons (2700–9900 metric tons) per day. A basic oxygen furnace for steelmaking can consume 800 tons (720 metric tons) of pure oxygen together with required amounts of crude iron and scrap to produce 12,000 tons (11,000 metric tons) of steel per day. Advanced methods of process analysis and control have been applied to such processing systems. The application of vacuum to extraction and refining processes, the leaching of low-grade ores for the extraction of metals, the use of electrochemical reduction cells, and the refining of reactive metals by processing through the vapor state are other important developments. *See* IRON METALLURGY; STEEL MANUFACTURE.

Because the production of metals employs many different chemical reactions, process metallurgy has been closely associated with inorganic chemistry. Techniques for analyzing ores and metallurgical products originated several centuries ago and represented an early stage of analytical chemistry. Application of physical chemistry to equilibria and kinetics of metallurgical reactions has led to great progress in metallurgical chemistry.

Physical metallurgy. Physical metallurgy investigates the effects of composition and treatment on the structure of metals and the relations of the structure to the properties of metals. Physical metallurgy is also concerned with the engineering applications of scientific principles to the fabrication, mechanical treatment, heat treatment, and service behavior of metals. *See* ALLOY; HEAT TREATMENT (METALLURGY).

The structure of metals consists of their crystal structure, which is investigated by x-ray, electron, and neutron diffraction, their microstructure, which is the subject of metallography, and their macrostructure. Crystal imperfections provide mechanisms for processes occurring in solid metals; for example, the movement of dislocations results in plastic deformation. Crystal imperfections are investigated by x-ray diffraction and metallographic methods, especially electron microscopy. The microstructure is determined by the constituent phases and the geometrical arrangement of the microcrystals (grains) formed by those phases. Macrostructure is important in industrial metals. It involves chemical and physical inhomogeneities on a scale larger than microscopic. Examples are flow lines in steel forgings and blowholes in castings. *See* CRYSTAL STRUCTURE; METALLOGRAPHY; X-RAY DIFFRACTION.

Phase transformations occurring in the solid state underlie many heat-treatment operations. The thermodynamics and kinetics of these transformations are a major concern of physical metallurgy. Physical metallurgy also investigates changes in the structure and properties resulting from mechanical working of metals. *See* PLASTIC DEFORMATION OF METAL.

The composition of metallic objects is often characterized by the presence of impurities, nonuniform distribution (segregation) of solute elements

and nonmetallic inclusions, especially in steel. These phenomena, which originate in the production process, can have important effects on the properties of metals and alloys. They illustrate the close relation between process and physical metallurgy.

The applications for which a metal is intended determine the properties that are of practical interest. For use in machinery and construction, mechanical properties including deformation and fracture behavior are of greatest importance. In transportation equipment the strength-to-weight ratio deserves special consideration. In other applications, electrical or thermal conductivity or magnetic properties may be decisive. Resistance to corrosion is a common requirement and accounts for another close link between metallurgy and chemistry, especially electrochemistry and surface chemistry. *See* CORROSION; METAL, MECHANICAL PROPERTIES OF.

General principles. An old distinction between ferrous and nonferrous metallurgy, although still of some practical significance, is no longer considered fundamental. Certain general principles have become established and apply to all metals. In some respects, however, there are great differences between metals which has led to specialization in research and industrial practice. On the other hand, an underlying structural science is developing to bring all materials—metallic and other inorganic as well as organic materials—within a unified framework.

Metallurgy occupies a position at the juncture of physics, chemistry, mechanical engineering, and chemical engineering. It also borders electrical, civil, aeronautical, and nuclear engineering. Metallurgical knowledge can be relevant to fields that are as removed from engineering as archeology, crime investigation, and orthopedic surgery. In the field of materials science and engineering, metallurgy takes its place as one of the oldest and most highly developed disciplines.

The area of concern of metallurgy widened as problems of materials availability were recognized and focused attention to the need for new production processes and the recycling of secondary metals. Environmental considerations require new technology for pollution abatement. Energy conservation favors more efficient processes in metals production. Conservation of materials also calls for more effective utilization of metals and the substitution of more plentiful for scarce metals.

For more information on metallurgy and some associated techniques see articles on individual metals and their metallurgy.

Michael B. Bever

Bibliography. C. S. Barrett and T. B. Massalski, *Structure of Metals*, 3d rev. ed., 1980; G. F. Carter, *Principles of Physical and Chemical Metallurgy*, 1979; G. E. Dieter, *Mechanical Metallurgy*, 3d ed., 1986; P. Haasen, *Physical Metallurgy*, 2d ed., 1988; M. A. Meyers and K. K. Chawla, *Mechanical Metallurgy: Principles and Applications*, 1984; J. J. Moore, *Chemical Metallurgy*, 2d ed., 1983, reprint 1993; R. E. Reed-Hill and R. Abbaschian, *Physical Metallurgy Principles*, 3d ed., 1992.

Metameres

The successive subdivisions along the length of the body axis in bilaterally symmetrical animals; also called somites or segments. Common examples are the muscles and spinal nerves in the human body and in the body and tail of many mammals, snakes and lizards, salamanders, and fishes. It also occurs in other chordates, and in arthropods and annelid worms. It never involves reproductive organs, and thus differs from strobilization in tapeworms and certain jellyfish. This serial repetition of parts (metamerism or segmentation) arises either from a bilateral series of coelomic pouches which form the segmental muscles, kidneys, and body cavities of lower forms, or from mesoblastic somites which form the skeletal and muscular segments of vertebrates. Repetitive features of the nervous system are acquired secondarily through the influence of mesodermal metameres upon adjoining ectodermal tissues. Several primitive embryonic somites become fused in the heads of adult arthropods and vertebrates. *See* ANIMAL SYMMETRY; COELOM; MUSCULAR SYSTEM; NEURULATION.

Howard L. Hamilton

Metamict state

The state of a special class of amorphous materials that were initially crystalline. W. C. Broegger first used the term *metamikte* in 1893 to describe minerals that were optically isotropic with a "glasslike" fracture but still retained well-formed crystal faces. In 1914 A. Hamburg correctly attributed the transition from the periodic, crystalline state to the aperiodic, metamict state as induced by alpha-decay damage. In minerals, this damage is the result of the decay of naturally occurring radionuclides and their daughter products in the uranium and thorium (^{238}U, ^{235}U, and ^{232}Th) decay series. A wide variety of complex oxides (for example, pyrochlore structure types), silicates (such as zircon, thorite, and yttrialite), and phosphates (for example, xenotime) are reported as occurring in the metamict state. All of these structures can accommodate uranium and thorium. A renewed interest in the metamict state has been stimulated by concern for the long-term stability of crystalline materials (nuclear waste forms) that will serve as hosts for actinides (for example, plutonium, americium, curium, and neptunium). Various crystalline materials (phases) may appear in a single waste form; each phase may or may not suffer radiation damage. For some nuclear waste-form phases, the radiation-induced transformation to the metamict state has been stimulated by doping phases with highly radioactive plutonium-238 or curium-244. *See* ACTINIDE ELEMENTS; ALPHA PARTICLES; RADIOACTIVITY.

Mineral properties. Metamict minerals are generally optically isotropic but may show varying degrees of anisotropy. Metamict phases lack cleavage, and conchoidal fracture is characteristic.

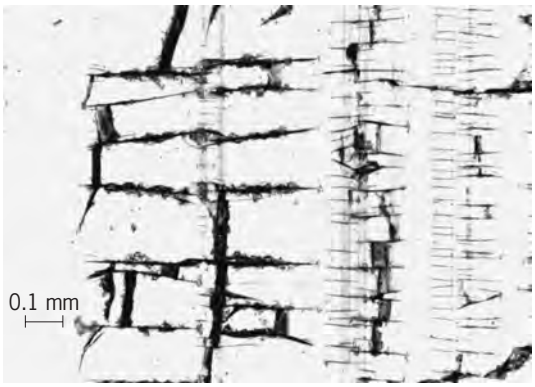

Fig. 1. Optical micrograph of (001) zone in a petrographic thin section of zircon. The variations in the shading are due to differences in the birefringence. (*From W. Lutze and R. C. Ewing, eds., Radioactive Waste Forms for the Future, North-Holland, 1988*)

There is a decrease in the hardness (approximately 30%) and in the elastic modulus with increasing alpha-decay dose. However, the introduction of defects causes a substantial increase in the fracture toughness.

Figure 1 shows a micrograph of (001) zone in a petrographic thin section of zircon that exhibits differences in birefringence. The birefringence decreases with increasing alpha-decay dose until the material becomes optically isotropic. Different zones in this sample have experienced different alpha-decay doses because of variations in the uranium and thorium concentrations. The system of microfractures are perpendicular to the zones, they are restricted to low-dose zones and terminate in zones that have alpha-decay doses $\geq 0.5 \times 10^{16}$ alpha events/mg (approximately equal to 10^{25} alpha events/m^3). Fractures terminate in the high-dose zones in which fracture toughness has increased.

The crystalline structure may be reconstituted by heating. Fully metamict material usually recrystallizes to a polyphase aggregate, but if the material is only partially damaged, it is possible to recrystallize to the original, single crystal. During recrystallization, several new phases may form; the particular phase assemblage is dependent on the conditions of recrystallization (for example, temperature and atmosphere). In many cases, the original premetamict phase may not recrystallize because of compositional changes caused by subsequent alteration. Annealing is exothermic, releasing the stored energy associated with the alpha-decay damage. In rare cases, recrystallization is pyronomic, releasing so much energy that the sample glows incandescently on heating.

Metamict minerals contain uranium and thorium, although contents may be quite low. The presence of uranium and thorium distinguishes metamict minerals from other naturally occurring amorphous materials that have not experienced this radiation-induced transformation. Lanthanide elements are also common (in some cases over 50 wt %) and water of hydration may be high (up to 70 mol %).

Partially crystalline metamict minerals display distinct broadening of diffraction maxima. Shifts in the position of diffraction maxima correspond to the increase in unit cell volume and the decrease in density. Volume expansion can be up to 8% and may cause microfracturing (Fig. 1). Fully damaged material (alpha-decay doses in excess of one displacement per atom) is x-ray- and electron-diffraction amorphous; that is, the fully damaged material is completely aperiodic and thus unable to diffract x-rays or electrons. *See* ELECTRON DIFFRACTION; X-RAY DIFFRACTION.

Alpha-decay damage. The radiation damage caused by the alpha-decay event is the result of two separate but simultaneous processes: (1) An alpha particle with an energy of approximately 4.5 MeV and a range of 10,000 nanometers dissipates most of its energy by ionization; however, at low velocities near the end of its track, it displaces several hundred atoms, creating Frenkel defect pairs. (2) The alpha-recoil atom with an energy of approximately 0.09 MeV and a range of 10 to 20 nm produces several thousand atomic displacements, creating tracks of disordered material. These two damaged areas are separated by thousands of unit cell distances and have different effects on the crystalline structure. Local point defects cause an increase in the distortion; therefore, there is an increase in the strain in the structure. Alpha-recoil tracks create regions of aperiodic material that at high enough alpha-decay doses (usually 10^{24} to 10^{25} alpha-decay events/m^3) overlap and finally lead to the metamict state. The former causes broadening of x-ray diffraction maxima and an increase in unit cell volume (and a decrease in the density); the latter causes a decrease in diffraction peak intensities. Typical changes in density (ρ), unit cell parameter (cell edge; a), and diffraction intensity (I/I_0) are shown in **Fig. 2**. The radiation-induced transition

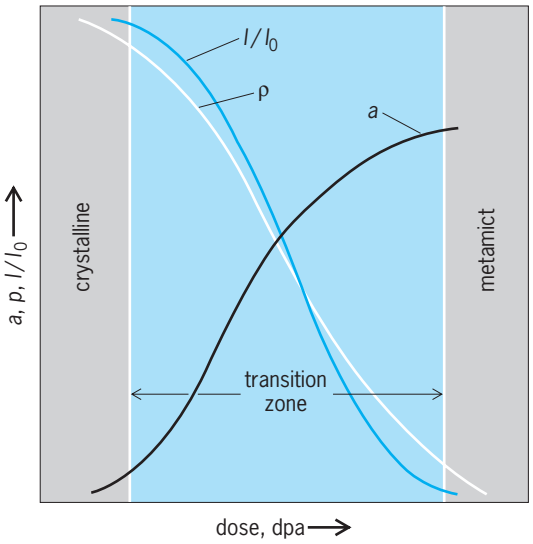

Fig. 2. Schematic representation of the change in the diffraction intensity (I/I_0), cell edge (a), and density (ρ) as a function of increasing alpha-decay dose. (*After G. R. Lumpkin and R. C. Ewing, Alpha-decay damage in minerals of the pyrochlore group, Phys. Chem. Miner., 16:2–20, 1988*)

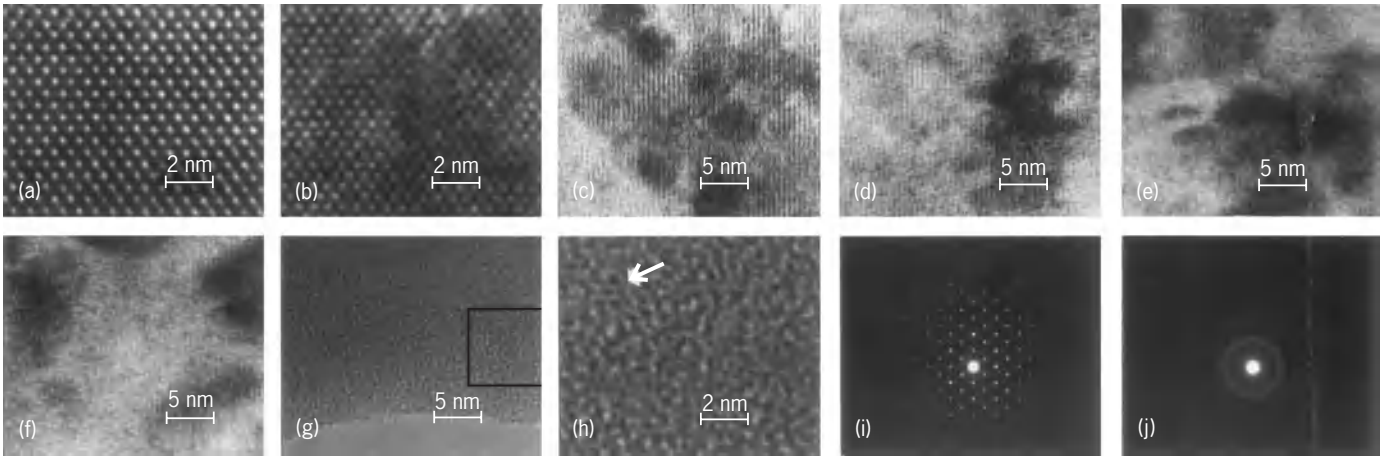

Fig. 3. High-resolution transmission microscopy images showing progressive structural damage to microlite as a function of increasing alpha-event dose (alpha-events/mg, uncorrected for annealing). (*a*) $<10^{15}$, (*b*) 3×10^{16}, (*c*) 3×10^{16}, (*d*) 3×10^{16}, (*e*) 9×10^{16}, (*f*) 10^{17}, (*g*) 4×20^{17}. (*h*) Enlargement of area outlined in *g*; arrow indicates fragment of the original structure. (*i, j*) Selected area diffraction patterns corresponding to images in *a* and *h*. (*From R. C. Ewing et al., Metamict minerals: Natural analogues for radiation damage effects in ceramic nuclear waste forms, Nucl. Instrum. Meth. Phys. Res., B32:487–497, 1988*)

from the crystalline to the metamict state occurs over a narrow range of alpha-decay dose (10^{24}–10^{25} alpha decays/m^3), which corresponds to 0.1 to 1.0 displacements per atom (dpa). *See* CRYSTAL DEFECTS; RADIATION DAMAGE TO MATERIALS.

Damage in-growth. For minerals from the same locality and of the same age but with variable uranium and thorium contents, the uranium/thorium content can be converted to a calculated alpha-decay dose. For such a suite of samples, the effects of increasing alpha-decay dose can be observed directly. These effects can be observed by high-resolution transmission electron microscopy (HRTEM) in the mineral microlite, a complex oxide of tantalum (Ta) and calcium (Ca) [**Fig. 3**]. At 0.9–1.0 wt % uranium oxide (UO_3; 3×10^{16} alpha events/mg), microlite exhibits mottled diffraction contrast and disrupted lattice periodicity within isolated areas of 1 to 3 nm (Fig. 3*b*, *c*). With 1.4 wt % UO_3 (6×10^{16} alpha events/mg), aperiodic areas increase in abundance and fine-scale mosaic textures are observed (Fig. 3*d*). The mosaic pattern consists of crystalline domains (1–5 nm) having lattice misorientations of several degrees. With a uranium content of 2.5 wt % UO_3 (9×10^{16} alpha events/mg), microlite clearly shows regions of coexisting crystalline and aperiodic regions (Fig. 3*e*). The selected area diffraction pattern exhibits prominent diffuse rings. At 2.9 wt % UO_3 (10^{17} alpha events/mg), the mosaic texture is lost, and islands of crystalline microlite are left within an aperiodic matrix (Fig. 3*f*). At 3.6 wt % UO_3 (10^{17} alpha events/mg), no lattice fringes are seen in HRTEM images, and only diffuse rings appear in selected area diffraction patterns (Fig. 3*f*). With 8.9 wt % UO_3 (4×10^{17} alpha events/mg), the microlite exhibits an aperiodic, random network structure (Fig. 3*g*). The images also show that nanometer-sized fragments (see arrow) of the original structure persist beyond the saturation level of the alpha-recoil damage (Fig. 3*h*).

Thus, the transition from the crystalline to the metamict state in microlite can be summarized as follows: (1) At <0.1 wt % UO_3 ($< 10^{15}$ alpha events/mg, < 0.1 dpa), isolated alpha-recoil tracks have little effect on selected area diffraction patterns and are difficult to observe in HRTEM images. (2) At 0.9 wt % UO_3 (3×10^{16} alpha events/mg, 3 dpa), isolated alpha-recoil tracks are observed as 1–3-nm areas of disrupted lattice periodicity. (3) When 2.5 wt % UO_3 (9×10^{16} alpha events/mg, 13 dpa) is reached, alpha-recoil tracks overlap, yielding coexisting areas of crystalline and metamict microlite. (4) The fully metamict state is reached at 2.9 to 3.6 wt % UO_3 (1–1.3×10^{17} alpha events/mg, 15–19 dpa). Beyond this dose, the HRTEM image exhibits a random network structure. The alpha-decay doses calculated for each step of the transition are based simply on the age of the sample and the uranium content. The fact that some areas of crystallinity remain even after reaching calculated doses in excess of one displacement per atom provides strong evidence of annealing of the damage at low temperatures and over long periods of time. *See* ELECTRON MICROSCOPE.

Annealing of radiation damage. One of the most important observations concerning alpha-decay damage in periodic materials is that some naturally occurring phases sustain high alpha-decay doses (in excess of 10 dpa) and still remain essentially crystalline. These materials include uraninite (UO_2), monazite ($CaPO_4$), titanite ($CaTiO_5$), and huttonite ($ThSiO_4$). The apparent stability of these phases to radiation damage is due to annealing of the radiation damage under ambient conditions. Indeed experimental results seem to suggest that the mean alpha-recoil track lifetime τ varies depending on the phase: uraninite and thorite ($\tau = 25{,}000$ years), pyrochlore ($\tau = 100$ million years), zircon ($\tau = 400$ million years), zirconolite ($\tau = 700$ million years). For phases in which the mean track lifetime is short, damage is annealed as quickly as it accumulates, and the phase remains crystalline. For those phases in which the mean track lifetime is great, the tracks accumulate and finally overlap, and the material becomes

metamict. The annealing process is critically important to understanding the crystalline-to-metamict transition and the final microstructure of alpha-decay–damaged materials. *See* AMORPHOUS SOLID; CRYSTAL STRUCTURE; THORITE; ZIRCON.

Rodney C. Ewing

Bibliography. B. C. Chakoumakos et al., *Alpha-Decay-Induced Fracturing in Zircon: The Transition from the Crystalline to the Metamict State*, vol. 236, pp. 1556–1559, 1987; R. C. Ewing et al., Metamict minerals: Natural analogues for radiation damage effects in ceramic nuclear waste forms, *Nucl. Instrum. Meth. Phys. Res.*, B32:487–497, 1988; H. D. Holland and D. Gottfried, The effect of nuclear radiation on the structure of zircon, *Acta Crystallog.*, 8:291–300, 1955; G. R. Lumpkin and R. C. Ewing, Alpha-decay damage in minerals of the pyrochlore group, *Phys. Chem. Miner.*, 16:2–20, 1988; W. J. Weber and G. D. Maupin, Simulation of radiation damage in zircon, *Nucl. Instrum. Meth. Phys. Res.*, B32:512–515, 1988.

Metamorphic rocks

One of the three major groups of rocks that make up the crust of the Earth. The other two groups are igneous rocks and sedimentary rocks. Metamorphic rocks are preexisting rock masses in which new minerals, or textures, or structures are formed at higher temperatures and greater pressures than those normally present at the Earth's surface. *See* IGNEOUS ROCKS; SEDIMENTARY ROCKS.

Two groups of metamorphic rocks may be distinguished; cataclastic rocks, formed by the operation of purely mechanical forces; and recrystallized rocks, or the metamorphic rocks properly so called, formed under the influence of metamorphic pressures and temperatures.

Cataclastic rocks are mechanically sheared and crushed. They represent products of dynamometamorphism, or kinetic metamorphism. Chemical and mineralogical changes generally are negligible. The rocks are characterized by their minute mineral grain size. Each mineral grain is broken up along the edges and is surrounded by a corona of debris or strewn fragments (mortar structure, **Fig. 1***a*). During the early stages of this alteration process the metamorphosed product is known as flaser rock (Fig. 1*b*). Eventually the original mineral grains are entirely gone, as in the mylonites. When seen through the microscope, the comminuted particles consist of a mixture of finely powdered quartz, feldspar, and other minerals with an incipient recrystallization of sericite or chlorite. Pseudotachylite is an extreme end product of this crushing process. *See* METAMORPHISM; MYLONITE.

Structural relations. Metamorphic rocks, properly so called, are recrystallized rocks. The laws of recrystallization are not the same as those of simple crystallization from a liquid, because the crystals can develop freely in a liquid, but during recrystallization the new crystals are encumbered in their growth by the old minerals. Consequently, the structures which develop in metamorphic rocks are distinctive and of great importance, because they reflect the physiochemical environment of recrystallization and thereby the genesis and history of the metamorphic rock.

Crystalloblastic structure. A crystalloblast is a crystal that has grown during the metamorphism of a rock. The majority of minerals are frequently bounded by their own crystal faces (idioblasts). Larger crystals are often packed with small inclusions of other minerals exhibiting the so-called sieve structure (poikilitic or diablastic structure).

Granoblastic refers to a nondirected rock fabric, with minerals forming grains without any preferred shape or dimensional orientation (Fig. 1*c*). Lepidoblastic (Fig. 1*d*), nematoblastic (Fig. 1*e*), and fibroblastic refer to rocks of scaly, rodlike, and fibrous minerals, respectively.

The metamorphic minerals may be arranged in an idioblastic series (crystalloblastic series) in their order of decreasing force of crystallization as follows: (1) sphene, rutile, garnet, tourmaline, staurolite, kyanite; (2) epidote, zoisite; (3) pyroxene, hornblende; (4) ferromagnesite, dolomite, albite; (5) muscovite, biotite, chlorite; (6) calcite; (7) quartz, plagioclase; and (8) orthoclase, microcline. Crystals of any of the listed minerals tend to assume idioblastic outlines at surfaces of contact with simultaneously developed crystals of all minerals of lower position in the series.

Preferred orientation. Certain minerals have a tendency to assume parallel or partially parallel crystallographic orientation. The shape and spatial arrangement of minerals such as mica, hornblende, or augite show a definite relation to the foliation in the schist or gneiss; that is, both foliation and fissility of a metamorphic rock are directly related to the preferred position assumed by the so-called schist-forming minerals, such as mica, hornblende, and chlorite. *See* GNEISS; SCHIST.

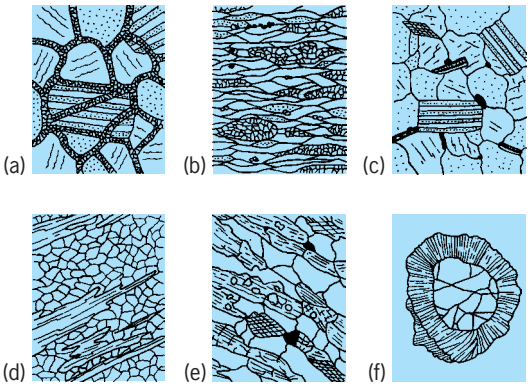

Fig. 1. Fabrics of metamorphic rocks under microscope. (*a*) Mortar fabric; (*b*) flaser or mylonitic fabric; (*c*) granoblastic fabric (*after E. E. Wahlstrom, Petrographic Mineralogy, John Wiley and Sons, 1955*). (*d*) Lepidoblastic fabric; (*e*) nematoblastic fabric; (*f*) porphyroblast with reaction rim (*after T. F. W. Barth, Theoretical Petrology, John Wiley and Sons, 1952*).

Students of structural petrology distinguish between preferred orientation of inequidimensional grains according to their external crystal form, and preferred orientation of equidimensional grains according to their internal or atomic structure. *See* PETROFABRIC ANALYSIS; STRUCTURAL PETROLOGY.

A special microscope technique (universal stage technique) is necessary in most cases to demonstrate in detail the preferred orientation of the mineral grains according to their atomic structure. *See* PETROGRAPHY.

Relic structures. Mineral relics often indicate the temperature and pressure that obtained in the preexisting rock. If a mineral, say quartz, is stable in the earlier rock and is also stable in the later rock, it will be preserved (unless stress action sets in) in its original form as a stable relic. However, when a mineral or a definite association of minerals becomes unstable, it may still escape alteration and appear as an unstable relic. These relics are proterogenic, that is, representative of an earlier, premetamorphic rock, or of an earlier stage of the metamorphism. Hysterogenic products are of later date, and are formed in consequence of changed conditions after the formation of the chief metamorphic minerals.

A common phenomenon, fairly illustrative of the tendency toward equilibria, is the formation of armors or reaction rims around minerals (Fig. 1*f*) which have become unstable in their association but have not been brought beyond their fields of existence in general (the armored relics). Thereby the associations of minerals in actual contact with one another become stable. If, however, the constituent minerals of a rock containing armored relics are named without noting this phenomenon, it may be taken as an unstable association. *See* PORPHYROBLAST.

Structure relics are perhaps of still more importance, directly indicating the nature of the preexisting rock and the mechanism of the metamorphic deformation. The interpretation of relics has been compared to the reading of palimpsests, parchments used for the second time after original writing was nearly erased. Every trace of original structure is important in attempting to reconstruct the history of the rock and in analyzing the causes of its metamorphism.

In sedimentary rocks the most important structure is bedding (stratification or layering) which originally was approximately horizontal. In metamorphic rocks deformed by folding, faulting, or other dislocations, the sum of all deformations can be referred to the original horizontal plane, and the deformations can be analyzed.

Fissility and schistosity. One of the earliest secondary structures to develop in sediments of low metamorphic grade is that of slaty cleavage (also referred to as flow cleavage or fissility), which grades into schistosity which is different from fracture cleavage, or strain-slip cleavage. Slaty cleavage is developed normal to the direction of greatest shortening of the rock mass, and cuts the original bedding at various angles. Tectonic forces acting on a book of sediments

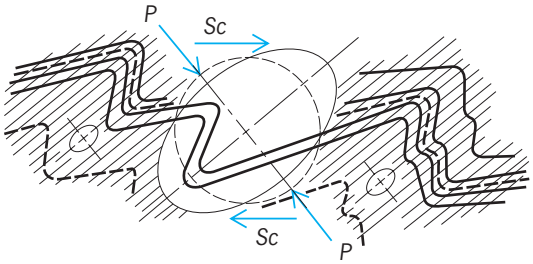

Fig. 2. Diagram showing the general relationship between deformation folding and slaty cleavage caused by pressure *PP* or the couple *ScSc*. Heavy black lines denote the original bedding deformed into folds. Thin lines indicate slaty cleavage which may grow into schistosity (false bedding). (*After G. Wilson, Proceedings of the Geological Association, 1946*)

of heterogeneous layers will throw them into a series of folds, and slaty cleavage develops in response to the stresses imposed on the rock system as a whole, because of the differential resistance of the several layers. Consequently, folding and slaty cleavage have a common parentage, as illustrated in **Fig. 2**.

In the rock series slate-phyllite-schist the slaty cleavage will grade into schistosity. It is a chemical and recrystallization phenomenon, as well as a mechanical one, and the directions of the schistosity become the main avenues of chemical transport. *See* ROCK CLEAVAGE; SLATE.

Contact-metamorphic rocks. Igneous magma at high temperature may penetrate into sedimentary rocks, it may reach the surface, or it may solidify in the form of intrusive bodies (plutons). Heat from such bodies spreads into the surrounding sediments, and because the mineral assemblages of the sediments are adjusted to low temperatures, the heating-up will result in a mineralogical and textural reconstruction known as contact metamorphism. *See* PLUTON.

The width of the thermal aureole of contact metamorphism surrounding igneous bodies varies from almost complete absence in the case of small intrusions (basalt dikes or diabase sills) to several kilometers in the case of large bodies. *See* CONTACT AUREOLE.

The effects produced do not depend only upon the size of the intrusive. Other factors are amount of cover and the closure of the system, composition and texture of the country rock, and the abundance of gaseous and hydrothermal magmatic emanations. The heat conductivity of rocks is so low that gases and vaporous emanations become chiefly responsible for the transportation and transfer of heat into the country rock.

Alteration of stratified rocks. Stratified rocks are altered in the contact zone to what is commonly called hornfels or hornstone. They are hardened, often flinty rocks, usable for road material, and so fine-grained that the mineral components can be discerned only with the microscope. Hornfels used to be regarded as "silicified" sediments. However, T. Kjerulf, in the later half of the nineteenth century, analyzed sedimentary shale and "silicified" shale of the Oslo

region and found that, chemically, they were identical (except for water and carbon dioxide content). Then geologists realized that the "hardening" of the shale took place without appreciable change in the chemical composition. Kjerulf summarized his results by saying that the composition (the shale) was independent of the kind of adjacent igneous rock.

Later H. Rosenbusch arrived at the same conclusion and pronounced that no chemical alterations accompany the formation of hornfelses except for the removal of fugitive constituents. The Kjerulf Rosenbusch rule is useful but needs modification, because chemical changes may ensue from hydrothermal and pneumatolytic action.

The next problem then is to see how the mineral assemblages of the hornfelses depend upon the chemical composition of the original sediments. The chief types of sedimentary rocks are sandstone (sand), shale (clay), and limestone. Among the varieties of hornfelses which may develop from different mixtures of these components, the continuous series from shale to limestone is the most interesting.

Most shales contain some iron- and magnesia-bearing constituents in addition to feldspar and clay minerals. Quartz, SiO_2, is always admixed. Consequently, sufficient SiO_2 is often present in the hornfelses to form highly silicified minerals. Other than SiO_2, the four chief chemical constituents are Al_2O_3, CaO, FeO, and MgO. The last two constituents are grouped together to define a system of three components: alumina, lime, and ferromagnesia.

By applying the mineralogical phase rule, which states that the number of stable minerals in a rock shall not be larger than the number of components, it follows that, except for quartz and some alkali-bearing minerals listed below, no more than three additional minerals should occur in any one (variety) of these hornfelses. Observations have verified this. Thus from alumina, lime, and ferromagnesia, seven minerals will form that are stable under the conditions of contact metamorphism: andalusite, Al_2SiO_5; cordierite, $Mg_2Al_4Si_5O_{18}$; anorthite, $CaAl_2Si_2O_8$; hypersthene, $(Mg,Fe)SiO_3$; diopside, $Ca(Mg,Fe)Si_2O_6$; grossularite, $Ca_3Al_2Si_3O_{12}$; and wollastonite, $CaSiO_3$. Only three (or fewer) of these minerals can occur together. In this way different mineral combinations develop, each combination (plus quartz and an alkali-bearing mineral) representing a natural hornfels. There are 10 such combinations, corresponding to hornfelses of classes 1–10 of V. M. Goldschmidt's terminology. *See* HORNFELS.

Variations from the above scheme are easily explained. Usually enough water and potash are present to produce mica; muscovite may form instead of, or together with, andalusite, and in the hornfelses of classes 4 and 5, biotite is usually present inducing a characteristic chocolate-brown color into the rocks. In hornfelses of class 10 some lime-rich hydrous silicates may develop, for example, vesuvianite (idocrase). The presence of ferric iron may produce andradite, $Ca_3Fe_2Si_3O_{12}$, a yellow to dark-green garnet which will form mixed crystals with grossularite.

Pneumatolysis and metasomatism. Other factors of importance in contact metamorphism are chemical changes that ensue from pneumatolytic and hydrothermal action. These changes are brought about by the magmatic gases and high temperatures that accompany igneous intrusions. The surrounding rocks are deeply penetrated not only by the heat but also by water and other volatile compounds. Because chemical alterations take place in this so-called pneumatolytic or hydrothermal contact zone, the Kjerulf-Rosenbusch rule is not applicable. The width of the affected zone varies from nil to thousands of feet. *See* METASOMATISM; PNEUMATOLYSIS.

The primary magmatic gases are acid and in consequence show high reactivity. If the contact rock is basic, especially limestone, the acid gases will react effectively with it. Limestone acts as a filter, capturing the escaping gases. As a result, a great variety of reaction minerals is formed. The corresponding rocks are known as skarns. If the reaction rocks are limestones composed of lime silicates, the reaction minerals are mainly garnet and pyroxene, often accompanied by phlogopite and fluorite. Sulfides of iron, zinc, lead, or copper may be present, and in some occurrences magnetite is formed. *See* SKARN.

Summary. Contact metamorphism caused by deep-seated magma intrusions is very common, and the products (disregarding the pneumatolytic action) vary regularly in accordance with the chemical compositions of the preexisting contact rock. Another factor of equal importance is the variation in temperature as influenced by the nature of the intruding rock and the distance from the contact. Thus it is possible to distinguish between an inner and an outer contact zone. The zones grade into each other by imperceptible transitions, but the mineral associations in the typical inner contact zone, the only zone considered so far, are markedly different from the associations in the outer contact zone.

These problems involve a consideration of the general relationships between minerals and mineral associations, on the one hand, and the temperature and pressure, on the other. They are discussed further in connection with the facies principle and the general process of regional metamorphism. However, it is important to realize that contact metamorphism, although it appears to be well defined and seems to stand out as an isolated natural phenomenon, is complex and variegated and passes by gradual transitions into other kinds of metamorphism. Geologically, contact metamorphism should be considered in connection with, and as a part of, the general system of rock metamorphism and metasomatism.

Regional metamorphic rocks. Crystalline schists, gneisses, and magmatities are typical products of regional metamorphism and mountain building. If sediments accumulate in a slowly subsiding geosynclinal basin, they are subject to down-warping and deep burial, and thus to gradually increasing temperature and pressure. They become sheared and deformed, and a general recrystallization results. However, subsidence into deeper parts of the crust is not the only reason for increasing temperature. It is

not known what happens at the deeper levels of a live geosyncline, but obviously heat from the interior of the Earth is introduced regionally and locally, partly associated with magmas, partly in the form of "emanations" following certain main avenues, determined by a variety of factors. From this milieu rose the lofty mountain ranges of the world, with their altered beds of thick sediments intercalated with tuffs, lava, and intrusives, all thrown into enormous series of folds and elevated to thousands of meters. Thus were born the crystalline schists with their variants of gneisses and migmatites. *See* EARTH, HEAT FLOW IN; OROGENY.

A. Michel-Lévy (1888) distinguished three main étages in the formation of the crystalline schists; F. Becke and U. Grubenmann (1910) demonstrated that the same original material may produce radically different metamorphic rocks according to the effective temperature and pressure during the metamorphism. Grubenmann distinguished three successive depth zones, epizone, mesozone, and katazone, corresponding to three consecutive steps of progressive metamorphism. In eroded mountain ranges, rocks of the katazone are, generally speaking, encountered in the central parts; toward the marginal parts are found rocks of the mesozone and epizone.

It is of paramount importance to obtain better information about the temperature-pressure conditions of the recrystallization, and thus to show the relation between the chemical and mineralogical composition of all varieties of rocks. A large-scale attempt in this direction was the development of the facies classification of rocks.

Mineral facies. As defined by P. Eskola (1921), a mineral facies "comprises all the rocks that have originated under temperature and pressure conditions so similar that a definite chemical composition has resulted in the same set of minerals, quite regardless of their mode of crystallization, whether from magma or aqueous solution or gas, and whether by direct crystallization from solution . . . or by gradual change of earlier minerals.. . . " To learn which mineral associations were characteristic of high temperature or of low temperature, and to determine which associations combined with high pressure and with low pressure, Eskola studied the mineral associations in the rocks.

It has long been known that in an area of progressive metamorphism each successive stage, or each new zone of metamorphism, is reflected in the appearance of characteristic rock types (G. Barrow, 1893). Rocks within the same zone may be called isofacial, or isograde as proposed by C. E. Tilley (1924) who, furthermore, proposed the term "isograd" for a line of similar degree of metamorphism.

In going from an area of unmetamorphosed sedimentary rocks into an area of progressively more highly metamorphic rocks, new minerals appear in orderly succession. Thus, in a series of argillaceous rocks subjected to progressive metamorphism, the first index mineral to appear is usually chlorite, followed successively by biotite, garnet (almandite), and sillimanite. A line can be drawn on the map indicating where biotite first appears. This line is the biotite isograd. The less metamorphosed argillites on one side of this line lack bioite, whereas the more metamorphosed rocks on the other side contain biotite. An isograd can be drawn for each mineral. Actually the isograds are surfaces, and the lines drawn on the map are the intersections of these surfaces with the surface of the Earth.

Further work along these lines resulted in the conclusion that it was possible to single out a well-defined series of mineral facies. Sedimentary rocks of the lowest metamorphic grade recrystallized to give rocks of the zeolite facies. At slightly higher temperatures the greenschist facies develops—chlorite, albite, and epidote being characteristic minerals. A higher degree of metamorphism produces the epidote-amphibolite facies, and a still higher degree the true amphibolite facies in which hornblende and plagioclase mainly take the place of chlorite and epidote. Representative of the highest regional metamorphic grade is the granulite facies, in which most of the stable minerals are water-free, such as pyroxenes and garnets. Any sedimentary unit will recrystallize according to the rules of the several mineral facies, the complete sequence of events being a progressive change of the sediment by deformation, recrystallization, and alteration in the successive stages: greenschist facies → epidote-amphibolite facies → amphibolite facies → granulite facies. The mineral associations of these rocks are summarized in the next section.

During regional metamorphism a stationary temperature gradient is supposed to be established in the mountain masses. Usually, the outer parts of a geosynclinal region are less affected, and in the ideal case the marginal parts contain unmetamorphosed sediments, clay, sand, and limestone, which gradually change into metamorphic rocks of successively higher facies as they extend into the central and deeper parts. *See* GEOSYNCLINE.

The **table** summarizes the metamorphic series of rocks that develop from the several types of common sediments and usually converge toward a granitic composition regardless of the nature of the original material. Basic igneous rocks (gabbros, basalts) show a composition related to that of marl and yield analogous metamorphic products. Not listed are ultrabasites (peridotites, and others) which by metamorphism become serpentine, chlorite or talc schist, soapstone, hornblende schist, pyroxene, or olivine masses. Original acid igneous rocks (granite, diorite, rhyolite) show a composition related to that of arkose and yield analogous products. Leptite is primarily fine-grained, usually showing tufaceous or blastoporphyric relic structures; or it is derived from argillaceous sediments. Hälleflintas are dense rocks of conchoidal fracture, genetically related to leptites. Kinzigites, characterized by containing aluminum silicates and usually also rich in magnesia, are metasomatic gneisses, but probably argillites also enter into their constitution. Granulite is a gneiss recrystallized in the high-temperature mineral facies group. *See* GRANULITE.

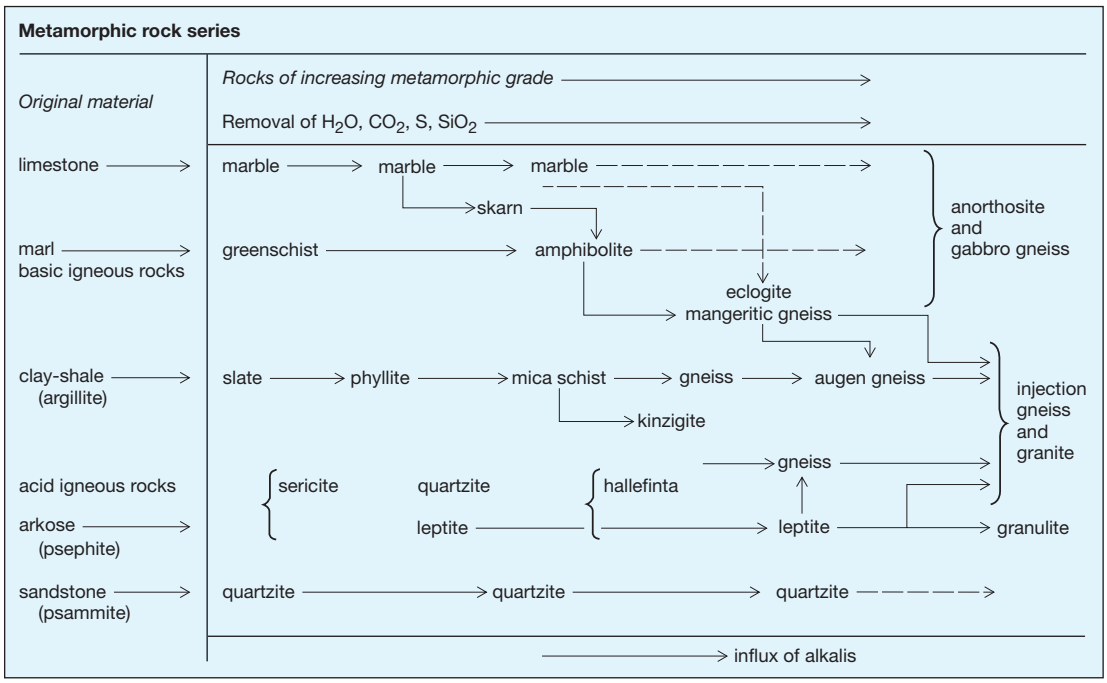

Chemical alterations. The chemical changes in the progressive series are complicated. When a sediment is heated, it obviously loses water and other volatile components, carbon dioxide, halogens, and so forth. These vapors act as carriers of several of the nonvolatile elements, for example, Si, Fe, and Mn. As heat emanates from the central parts of a geosynclinal region, it is heralded by a cloud of these vapors migrating centrifugally through the surrounding sediments. Usually, however, no major chemical alterations take place. Petrographers used to believe that the vapors were rich in alkalis and that large quantities, particularly of sodium, gradually would be deposited and fixed in the sediments. The argument was that the sediments, concomitant with a progressive change by increasing metamorphism, seem to exhibit an increase in their sodium content. However, this is not a real increase, but due to an erroneous sampling of the sedimentary material. Normally the lower stages of metamorphism tend to conserve the original composition.

Mineral associations of the facies. Rocks of the greenschist facies recrystallized at low temperatures and often under high shearing stress. These include chlorite schists, epidote-albite schists, and actinolite schists, all of which are green, hence the name. Other common rocks in this group are serpentinites, talc schists, phyllites, and muscovite (sericite) schists. Plagioclase is not stable in greenschists but breaks up into epidote and albite. Glaucophane schists are rare and probably they represent greenschists formed under conditions of high stress. Rocks of the epidote-amphibolite facies recrystallized in a somewhat higher temperature range. Amphibolites with epidote and either albite or oligoclase are typical. Mica schists (garnet), biotite schists, staurolite schists, and kyanite schists are common. Cordierite-antophyllite (gedrite) schists and chloritoid schists also occur. Sodic plagioclase is stable in these rocks.

Rocks of the amphibolite facies recrystallized at about 1100°F (600°C). Amphibolites of hornblende and plagioclase are typical and often carry quartz or biotite or both, or garnet. Sillimanite-muscovite schists (gneisses) are common. Andalusite and staurolite occur at low pressure; kyanite and staurolite at intermediate and high pressure.

Rocks of the granulite facies have their greatest extension in old Precambrian areas, but are also found in younger deeply eroded mountain chains. Diagnostic association is ortho- and clinopyroxene; hence the alternative name, two-pyroxene facies group. Hypersthene and augite (garnet) plagioclase gneisses, usually with quartz but occasionally with olivine or spinel, are typical. Amphibolites with hypersthene or diopside or both, sillimanite gneisses, and at high pressure kyanite gneisses are found. The temperature range is probably around 1500°F (800°C). The experiments by H. S. Yoder (1952) in the system MgO–Al_2O_3–SiO_2–H_2O indicated that at approximately 1100°F (600°C) and 1200 atm (120 megapascals) it is possible to have different mineral assemblages suggestive of every one of the now accepted metamorphic facies in stable equilibrium. These different mineral assemblages (artificial facies) observed by Yoder are the result of differences in the water content, and are not related to variation in temperature and pressure.

As an example, the mineral clinochlore, corresponding to one of the most common rock-making chlorites, shows an upper limit of stability, either alone or in association with talc, of 1300°F (680°C). This is, indeed, a high temperature for any kind of metamorphism, almost a magmatic temperature. But according to the tenets of the mineral facies, chlorite

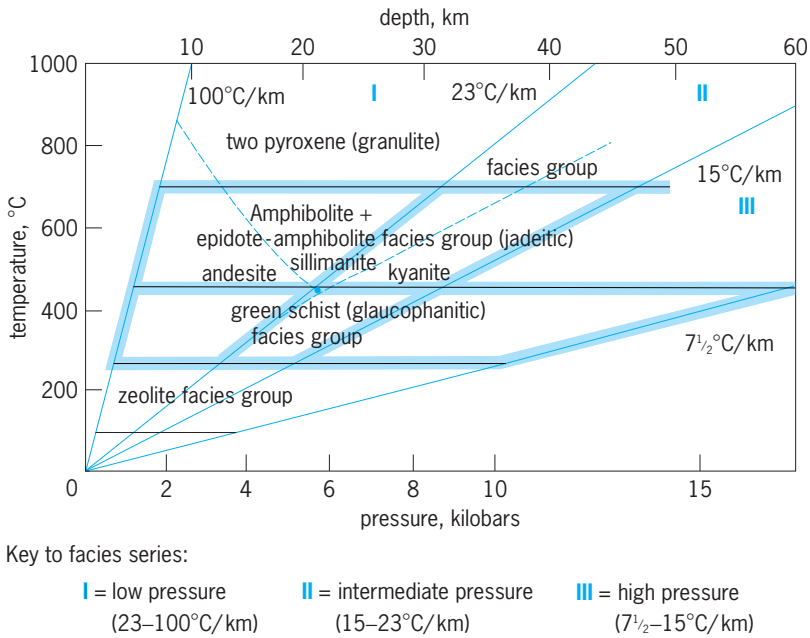

Fig. 3. Mineral facies groups of regional metamorphic rocks showing the temperature and pressure of metamorphism. °F = (°C × 1.8) + 32.1 km = 0.6 mi.

is strictly limited to the greenschist facies, about 390°F (200°C).

Although Yoder has proved that there is no absolute relation between temperature and facies, it appears likely that, to the field geologist and to the laboratory worker as well, the facies will still remain the best system of classification of metamorphic rocks; and in a majority of cases the facies will indicate the temperature-pressure conditions under which the several rocks recrystallized. Generally speaking, there is a regular relation between the chemical activity of water and the facies of the metamorphic rock.

Water content and mineral facies. The role of water in metamorphism is determined by at least four variable, geologically related parameters: rock pressure, temperature, water pressure, and the amount of water present. During a normal progressive regional metamorphism, rock pressure and temperature are interdependent. The amount of water and the pressure of water are related to the encasing sediments and to the degree of metamorphism in such a way that, generally speaking, the low-grade metamorphic facies are characterized by the presence of an excess of water, the medium-grade by some deficiency in water, and the high-grade by virtual absence of water.

In the usual diagrammatic illustration of the mineral facies of rocks, temperature and pressure (depth) are taken as coordinates; in regional-metamorphic rocks a third, dependent coordinate may be added, the activity of water running upward approximately along the geothermal gradient.

Facies series and groups. Metamorphic facies may be divided into facies series depending mainly on pressure, and facies groups depending mainly on temperature.

Three facies series have been proposed: low-pressure, intermediate-pressure, and high-pressure series. The low-pressure series dominates the Hercynian and Svecofennian of Europe, the Paleozoic of Australia, and part of the paired belts in Japan and New Zealand. The intermediate-pressure series (the original Barrowian zones) occurs in the European Caledonides, the Appalachians, the Precambrian Belt series of Idaho, the Himalayas, and parts of Africa. The high-pressure facies series is found in the Alps and the circum-Pacific region—Japan, New Zealand, Celebes, and the United States. Thus each intracontinental orogenic belt is characterized by one facies series which reflects the pressure that prevailed during metamorphism, whereas the circum-Pacific region with its paired metamorphic belts exhibits two facies series.

Four facies groups are recognized from low to high temperature: laumontite and prehnite-pumpellyite; greenschist, including glaucophane schist; amphibolite, including epidote-amphibolite; and the two-pyroxene (granulite) facies group.

This scheme is presented in **Fig. 3**. The "normal" geothermal gradient is in the range 45–70°F per mile (15–23°C per kilometer) depth. The intermediate-pressure facies series is found in areas with this gradient; the high-pressure series (the Alpine series) in areas with lower thermal gradients or with high overpressure (orogenic pressure); and the low-pressure facies series in areas with steep thermal gradients. The phase boundaries of the polymorphic forms of Al_2SiO_5 (andalusite, sillimanite, kyanite) have a central position in this scheme, the triple point being located approximately at 840°F (450°C) and 6 kilobars (60 MPa). The "minimum" melting of granite under water pressure occurs approximately along the boundary between the amphibolite and the two-pyroxene facies group. A separate eclogite facies is not recognized by this scheme. *See* ECLOGITE.

Figure 3 illustrates the distributions and interrelations of the various metamorphic rocks. It also represents a schematic profile through the continental crust down to 40-mi (60-km) depth, that is, down to the Moho discontinuity. Thus the normal continental crust is entirely made up of metamorphic rocks; where thermal, mechanical, and geochemical equilibrium prevails, there are only metamorphic rocks. Border cases of this normal situation occur in the depths where ultrametamorphism brings about differential melting and local formation of magmas. When equilibrium is restored, these magmas congeal and recrystallize to (metamorphic) rocks. At the surface, weathering processes oxidize and disintegrate the rocks superficially and produce sediments as transient products. Thus the cycle is closed; petrology is without a break. All rocks that are found in the continental crust were once metamorphites. *See* AMPHIBOLITE; GREISEN; MARBLE; MIGMATITE; PHYLLITE; QUARTZITE; SCAPOLITE; SERPENTINE; SOAPSTONE.

T. F. W. Barth; Robert C. Newton

Bibliography. M. G. Best, *Igneous and Metamorphic Petrology*, 1982; K. Bucher and M. Frey,

Petrogenesis of Metamorphic Rocks, 6th ed., 1994; R. Mason, *Petrology of the Metamorphic Rocks*, 2d ed., 1990; A. Miyashiro, *Metamorphic Petrology*, 1994; A. Nicolas and J. P. Poirier, *Crystalline Plasticity and Solid-State Flow in Metamorphic Rocks*, 1976; H. Ramberg, *Gravity, Deformation and the Earth's Crust in Theory, Experiments and Geological Application*, 2d ed., 1981.

Metamorphism

Changes in a preexisting rock's mineralogy, texture, and/or composition that occur predominantly in the solid state when the rock is exposed to a physical or chemical environment that is significantly different from that in which it initially formed. Any parameter that can effect such a change is one that could cause metamorphism. These include temperature, pressure, the nature of the fluid phase, and the state of stress. Although the processes of weathering and diagenesis (changes in sediments) are essentially identical to those of metamorphism, and solid metamorphic rock masses can coexist with melts during partial melting, metamorphism is rather arbitrarily confined to conditions between those of diagenesis and large-scale melting. Given that weathering and diagenesis occur only in the thin uppermost veneer of sediments, and that large-scale melting is an exceptional process with respect to normal conditions at depth, metamorphism is considered the dominant process taking place throughout most of the Earth's crust and mantle. Significant uplift and erosion typically is required to expose metamorphic rocks at the surface of the earth. *See* DIAGENESIS; EARTH CRUST; EARTH INTERIOR; METAMORPHIC ROCKS; ROCK; WEATHERING PROCESSES.

Scottish physician and farmer James Hutton (1726–1797) of Edinburgh was the first to propose that some crystalline rocks were originally sedimentary in nature and were subsequently transformed by subterranean heat. British geologist Charles Lyell (1797–1855) adopted the Huttonian theory, and in the first edition of his *Principles of Geology* (1833) proposed the term metamorphic for the altered strata. *See* SEDIMENTARY ROCKS.

Because metamorphic processes occur well beneath the Earth's surface, they cannot be observed directly. It is generally accepted that some metamorphic rocks can be created by simple burial to sufficient depth, but that most metamorphic rocks now exposed at the surface were affected by some additional process such as mountain building (orogeny), alteration by an adjacent igneous body, intense deformation, and so forth. Field observations of uplifted and eroded rocks (exposing the transition from unaltered precursors to thoroughly recrystallized high-temperature rocks) and laboratory simulations at elevated temperature and pressure provide the basis for our understanding of metamorphic processes and conditions. *See* IGNEOUS ROCKS; OROGENY.

The mineralogy of a rock is developed as the chemical constituents combine to form compounds (minerals) that are at equilibrium under the existing conditions. Most natural mineral assemblages are stable over only a limited range of pressure-temperature-composition (*P-T-X*) conditions. For example, the stable form of basalt above 1200°C (2200°F) at the surface of the Earth is liquid lava. It will solidify from lava between about 1200 and 1000°C (2200 and 1800°F), typically to a mixture of the minerals plagioclase, pyroxene, and an Fe-Ti oxide mineral (an igneous basalt). When exposed at the Earth's surface these minerals are no longer the stable assemblage, and gradually are replaced by a mixture of oxides, hydroxides, and clays (weathering products which produce soils). If this mixture is buried and exposed to conditions of progressively higher temperature and pressure, the oxides and clays might become reconstituted to a mineral assemblage dominated by albite, chlorite, and actinolite (low metamorphic grade); or plagioclase and hornblende (medium metamorphic grade). If heated sufficiently, the mineral assemblage may return to the original mineralogy of the basalt (pyroxene and plagioclase) at high metamorphic grade. *See* ALBITE; BASALT; CHLORITE; FELDSPAR; HORNBLENDE; LAVA; MINERAL; PYROXENE.

The texture of a metamorphic rock, likewise, is developed as an approach toward textural equilibrium. The texture of the high-grade metamorphic plagioclase-pyroxene rock discussed above crystallized in the solid state, and should differ significantly from that of the original igneous plagioclase-pyroxene rock that formed from a liquid. Kinetic factors may preclude complete chemical or textural equilibrium, particularly at lower temperatures, resulting in textures exhibiting crystal deformation (with or without partial annealing) or preserving incomplete mineral reactions.

Any transition from one mineral assemblage to another necessarily involves a chemical reaction with some of the initial mineral assemblage as reactants and new minerals as products. These reactions are subject to the principles of thermodynamics, allowing petrologists to use experimentally derived enthalpies, volumes, heat capacities, compressibilities, and chemical mixing parameters to estimate the *P-T-X* stability ranges of a great number of equilibrium mineral assemblages. *See* PETROLOGY.

Types of metamorphism. Different kinds of metamorphism may be defined on the basis of different criteria, including (1) the spatial extent over which metamorphism occurred, (2) the geologic setting of metamorphism, (3) the specific cause of metamorphism, (4) whether a rock equilibrated to a single event or if one can discern superposed overprints, (5) whether or not significant change in composition occurred, and (6) whether all or part of the mineral assemblage developed in response to increasing or decreasing temperature. Classifications of metamorphism therefore vary with the proclivity of the author. The following classification is an attempt at compromise between the various viewpoints.

Regional metamorphism occurs over an area of great extent and affects a large volume of rock. It is thus associated with large-scale processes, such as sea-floor spreading or mountain building (orogeny). Local metamorphism is restricted to far more limited rock volumes and can typically be related to a local cause/source such as a magmatic intrusion, and fault zone. When it is possible to relate metamorphism to a particular cause/source (as is generally the case) the following classification takes precedence.

1. Orogenic metamorphism is a type of regional metamorphism related to the development of mountain belts. Dynamic and thermal effects are combined in varying proportions over a wide range of *P-T* conditions, typically resulting in foliated metamorphic rocks such as slate, schist, and gneiss (listed in order of increasing grade). Multiple deformational and thermal phases may occur, resulting in several metamorphic overprints. Most exposed metamorphic rocks belong to this category. *See* GNEISS; SCHIST; SLATE.

2. Burial metamorphism is a type of regional metamorphism developed in rocks deeply buried under a sedimentary and/or volcanic pile and is typically not associated with orogenic deformation or extensive magmatic intrusion. It commonly involves only low to intermediate metamorphic temperatures and low pressures, resulting in nonfoliated or poorly foliated rocks.

3. Ocean-floor metamorphism is a type of regional metamorphism developed near oceanic spreading centers in response to the high local heat flow and extensive circulation of heated seawater along pervasive fractures.

4. Contact metamorphism is a type of local metamorphism that affects the host rocks adjacent to a magma body. It is typically caused by heat transfer, perhaps accompanied by fluid emanations, from a cooling magmatic intrusion. Metamorphic pressures are generally low, but temperatures vary widely. Textures are predominantly nonfoliated, except in cases of substantial dynamic effects due to the intrusion or when the host rocks are already orogenically metamorphosed and the contact effects are superimposed.

5. Fault-zone metamorphism is a type of local metamorphism associated with fault or shear zones. Dynamic effects predominate, resulting in grain-size reduction and textures governed by the interaction of deformation and recrystallization (typically strain induced). *See* EARTHQUAKE; FAULT AND FAULT STRUCTURES.

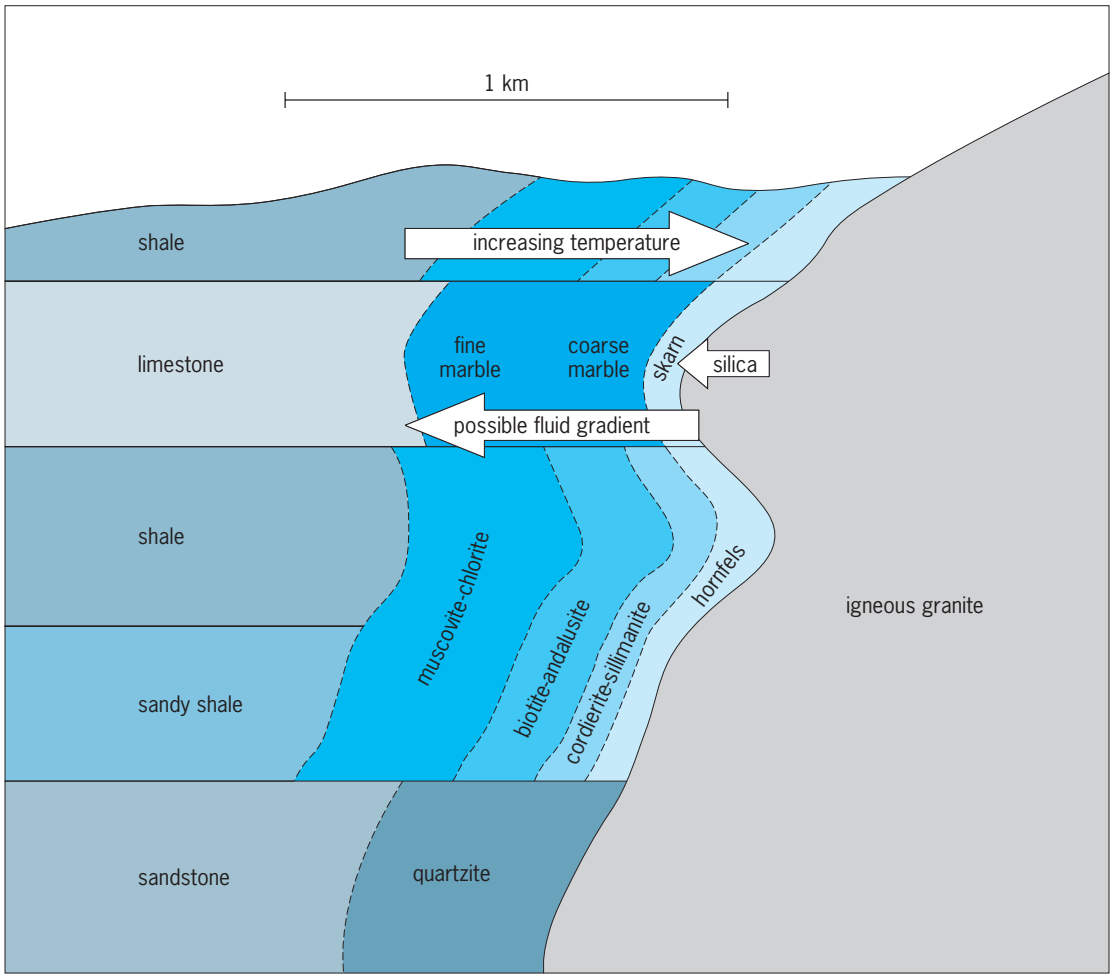

Fig. 1. Diagrammatic cross section through a contact-metamorphic aureole surrounding an intrusive igneous granite.

6. Impact metamorphism is a type of local metamorphism caused by the impact of a meteorite or other projectile. Extreme cases may result in melting and even vaporization.

7. Pyrometamorphism is an extreme type of contact metamorphism characterized by very high temperatures at relatively low pressures, generated by a very hot volcanic or subvolcanic body. It is most typically developed in xenoliths enclosed in such bodies, but may also occur at wall-rock contacts. It is typically accompanied by varying degrees of partial melting. *See* XENOLITH.

8. Hydrothermal metamorphism is a type of metamorphism, typically of local extent, caused by hot water-rich fluids. A common situation for this type of metamorphism is in geothermal fields above cooling magmatic intrusions.

9. Polymetamorphism applies to any situation in which one metamorphic event can be demonstrated to partially overprint a previous metamorphism, while retaining relics of the original event.

10. Metasomatism is any metamorphism accompanied by significant chemical alteration. This typically occurs in situations involving fluid emanations from a crystallizing magmatic intrusion, but may encompass a variety of situations involving contrasting rock compositions and/or fluid mobility. *See* MAGMA; METASOMATISM.

11. Prograde metamorphism refers to the changes in a rock that accompany increasing metamorphic grade (the usual case).

12. Retrograde metamorphism refers to changes that accompany decreasing grade as a body of rock cools and recovers from a metamorphic or igneous event. Equilibrium is better maintained during prograde metamorphism when energy is being added, and retrograde effects are generally minor and incomplete. Most metamorphic rocks thus retain the imprint of the maximum metamorphic temperatures attained.

Metamorphic gradients and zonation. Metamorphism is a response to changes in external conditions. There are gradients in temperature, pressure, and fluid composition in nature, so we can expect some sort of zonation in the mineral assemblages constituting the rocks that equilibrate across an expanse of these gradients. We should thus be able to traverse into an eroded metamorphic area and cross from nonmetamorphosed rocks through zones of progressively higher metamorphic grades. **Figure 1** shows a cross section through a portion of a contact aureole surrounding an intrusive igneous granite. The granite intrudes into the shallow crust at temperatures in excess of 650°C (1200°F), and a gradient in temperature develops as the granite heats the adjacent sediments. Most granites also release water as they cool and crystallize, which may also set up a gradient in fluid composition across the aureole. Because the different sedimentary layers are of contrasting composition, they respond differently to the thermal and fluid gradients. At the outer limits of rocks affected, shale may develop small crystals of muscovite and chlorite. At higher temperatures inward toward the granite, shale may progressively develop larger grains of biotite + andalusite, cordierite + sillimanite. Very near the contact, the original shale may be thoroughly recrystallized to a hornfels containing coarse quartz, cordierite, and sillimanite. Because limestone and sandstone are composed of virtually a single mineral each [limestone is calcite ($CaCO_3$) and sandstone is quartz (SiO_2)], they are less reactive and the principal change developed in these layers is recrystallization of the existing minerals to produce progressively coarser marble and quartzite at higher temperatures. Only very near the granite may some of the released silica-bearing

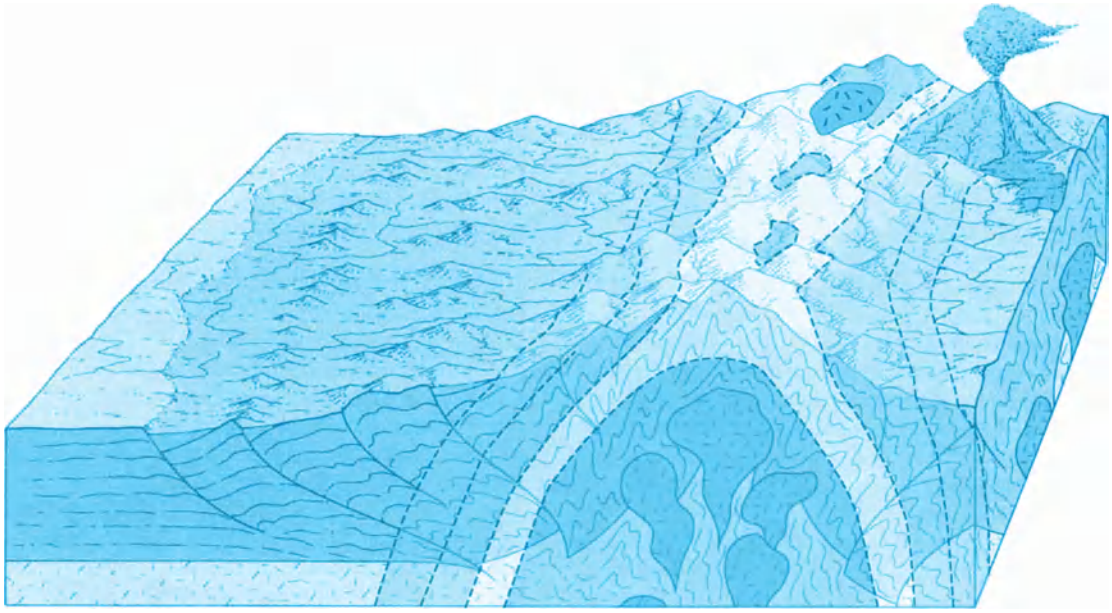

Fig. 2. Cross section through an uplifted and eroded mountain belt, exposing the orogenically metamorphosed rocks.

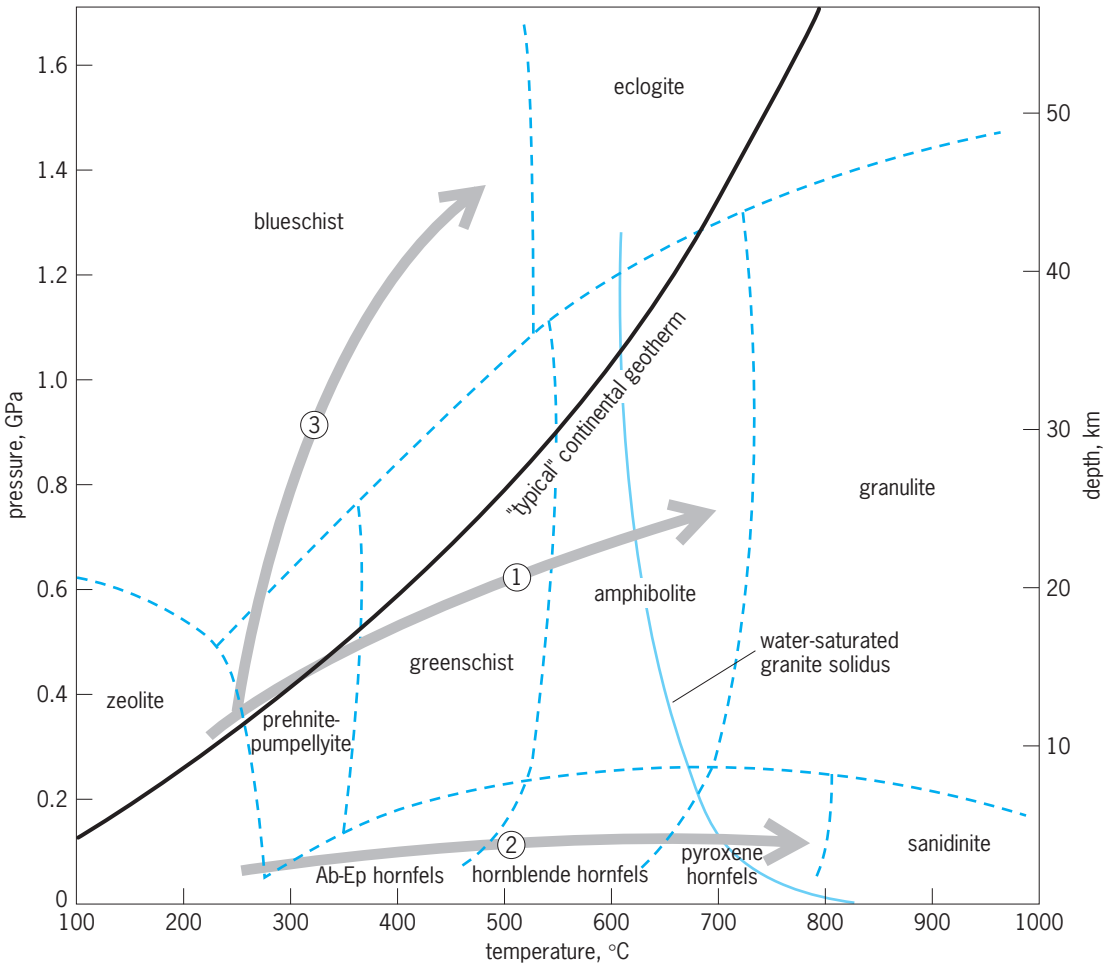

Fig. 3. Pressure-temperature diagram showing the generally accepted limits of the various facies. Boundaries are approximate and gradational. Included are the typical, or average, continental geotherm and the minimum melting curve for water-saturated granite. Broad gray swaths (1–3) represent metamorphic field gradients from three contrasting terranes. 1, average gradient for orogenic metamorphism in the Scottish Highlands; 2, gradient developed in contact aureoles above large granitic plutons; 3, gradiant from a subduction zone complex (Franciscan formation of California).

aqueous fluids interact with the limestone to form calcium-silicate minerals (skarn). **Figure 2** shows a similar cross section through a mountain range that has been uplifted and eroded to expose rocks that have experienced regional metamorphism. Here gradients in both temperature and pressure over much larger distances (typically hundreds of km) cause similar, but larger-scale transitions in the mineral assemblages and textures of the affected rocks. *See* ANDALUSITE; CONTACT AUREOLE; CORDIERITE; GRANITE; LIMESTONE; MUSCOVITE; SANDSTONE; SHALE; SILLIMANITE; SKARN.

The shales in Fig. 1 exhibit four metamorphic zones of progressively higher-grade metamorphism, and the regionally metamorphosed area in Fig. 2 exhibits four zones of much greater areal extent. The transition from one zone to another is typically caused by changing conditions; hence, a reaction between some or all of the lower-grade minerals to produce the higher-grade assemblage. The transition in the field across which this reaction occurs is called an isograd (line of constant metamorphic grade). George Barrow pioneered the mapping of such zones in orogenically metamorphosed shales in the Scottish Highlands. Barrow named the Scottish zones based on the first appearance of a characteristic index mineral developed by a metamorphic reaction marking the transition into that zone. The lower limit of any zone in the field was thus marked by an isograd named for the index mineral that first appeared at that isograd. Barrow's now-classical isograds and zones (in order of increasing grade) are chlorite, biotite, garnet, staurolite, kyanite, and sillimanite, but it is common practice to name zones on locally recognized mineral sequences. *See* GARNET; KYANITE; STAUROLITE.

Metamorphic facies. Pentii Eskola developed the concept of metamorphic facies as an extension of the zonal method of characterizing metamorphism and the grades developed. This approach is based on the typical zones developed in metamorphosed basaltic igneous rocks (which span relatively broad ranges of conditions) and attempts to assign pressure and temperature limits to those zones. **Figure 3** is a pressure-temperature (*P-T*) diagram illustrating the most commonly accepted metamorphic facies and the *P-T* conditions appropriate to each. *See* FACIES (GEOLOGY).

In Fig. 3, *P-T* trajectory 1 is representative of orogenic regional metamorphism and suggests that the typical facies series begins in the low-grade zeolite facies and extends progressively through the prehnite-pumpellyite facies, the greenschist facies, and the amphibolite facies. Large-scale melting is common in the upper amphibolite facies if sufficient water is available, and only dehydrated rocks make it into the granulite facies. Contact metamorphism, developed in aureoles around magmatic plutons, follows a *P-T* trajectory similar to 2 in Fig. 3, with a series of hornfels facies. Regional metamorphism accompanying the subduction of a cool plate is characterized by an unusually high *P/T* ratio (similar to trajectory 3 in Fig. 3), resulting in a series culminating in the blueschist facies. Very deep rocks, generally at mantle pressures, are typically metamorphosed in the eclogite facies (along the normal geotherm). It is common practice to combine the facies and zone approaches. Barrow's chlorite zone, for example, is in the greenschist facies and his staurolite zone is in the amphibolite facies. *See* AMPHIBOLITE; BLUESCHIST; ECLOGITE; GRANULITE; PLUTON; PREHNITE; ZEOLITE.

John D. Winter

Bibliography. G. Barrow, On an intrusion of muscovite biotite gneiss in the S. E. Highlands of Scotland and its accompanying metamorphism, *Quart. J. Geol. Soc.*, 49:350–358, 1893; G. Barrow, On the geology of the lower Deeside and the southern highland border, *Proc. Geol. Ass.*, 23:268–284, 1912; K. Bucher and M. Frey, *Petrogenesis of Metamorphic Rocks*, 7th ed., Springer, 2002; P. Eskola, On the relations between the chemical and mineralogical composition in the metamorphic rocks of the Orijärvi region, *Bull. Commis. Geol. Finlande*, vol. 44, 1915; C. Lyell, *Principles of Geology*, James Murray, London, 1833; J. Winter, *An Introduction to Igneous and Metamorphic Petrology*, Prentice Hall, 2001; B. W. D. Yardley, *An Introduction to Metamorphic Petrology*, Longman, Essex, 1989.

Metamorphosis

A pronounced change in both the internal and external morphology of an animal that takes place in a short amount of time, triggered by some combination of external and internal cues. The extent of morphological change varies considerably among species. Even when morphological changes are relatively slight, metamorphosis typically brings about a pronounced shift in habitat and lifestyle. Consider, for example, the transformation of a sluggish leaf-eating caterpillar into a nectar-drinking flying butterfly; or of a nonfeeding but free-swimming microscopic larva in a bivalved shell into a suspension-feeding sedentary barnacle; or of an aquatic suspension-feeding tailed tadpole larva into a carnivorous semiterrestrial frog or toad. In many species, metamorphosis is truly a morphological, ecological, and physiological revolution. The precise morphological, physiological, and biochemical changes that constitute metamorphosis; the neural, hormonal, and genetic mechanisms through which those changes are controlled; and the ecological consequences of those changes and when they take place continue to be studied in a wide variety of animals. The hormonal and genetic control of metamorphosis has been best examined in a few species of insect, amphibian, and fish (such as flounder), but other aspects of metamorphosis have been investigated for other insect, amphibian, and fish species as well as for crabs, barnacles, gastropods, bivalves, bryozoans, echinoderms, sea squirts—indeed, representatives of essentially every animal phylum.

Amphibians. Amphibians exhibit extensive tissue remodeling during metamorphosis, including resorption of the tail musculature and skeletal system; major reconstruction of the digestive tract; degeneration of larval skin and pronounced alteration in skin chemical composition; growth of the hind and fore limbs; degeneration of the gills and associated support structures; shifts in mode of nitrogen excretion, from ammonia to urea; alteration in visual system biochemistry; replacement of larval hemoglobin with adult hemoglobin; and differential growth of the cerebellum. In a number of species, all of these changes can be brought about or accelerated by exposing advanced tadpole larvae to increasing concentrations of hormones produced by the thyroid gland. The responsiveness of individual tissues depends on the concentration of thyroid hormones in the blood, the presence of receptors that bind the hormones, and possibly the number of receptors in particular tissues. Thyroid activity is controlled by the central nervous system via the hypothalamus. The effects of thyroid hormone can be modified to various degrees in different tissues by the pituitary hormone prolactin; for example, injections of prolactin reduce the rate of thyroid-induced tail and gill regression, which acts to maintain larval form. The role of prolactin in other species is less clear.

In some salamander species, the animals never metamorphose but attain sexual maturity while maintaining the larval form. The target tissues can no longer respond to thyroid hormones, or the thyroid no longer secretes the appropriate hormones, or it secretes them in insufficient amounts, depending on the species. In some species, larvae metamorphose at some temperatures and humidities but maintain the larval form through adulthood under other conditions.

Rates of metamorphosis may also be modulated by melatonin, a secretion of the pineal gland. *See* AMPHIBIA; ANURA; THYROID GLAND.

Insects. Metamorphosis among insects is associated primarily with wing development. Bristletails and other species that do not develop wings and are not descended from winged ancestors exhibit no pronounced metamorphosis, only a gradual increase in size and eventual arrival at reproductive maturity. Metamorphosis is most dramatic among holometabolous species, which pass through a distinctive and largely inactive pupal stage; in such species, all of the transformations separating the larval morphology and physiology from that of the adult

take place in the pupa. Wings, compound eyes, external reproductive parts, and thoracic walking legs develop from discrete infolded pockets of tissue (imaginal discs) that form during larval development.

As with other arthropods, and indeed other animals that secrete a rigid external cuticle, size increases occur only during times of periodic molting—the shedding of one cuticle followed by the secretion and hardening of a new cuticle enclosing a significantly enlarged body. The molting process is controlled by hormones (ecdysone and related ecdysteroids) secreted by prothoracic glands located in the animal's thorax. These inactive precursors are quickly converted to active form elsewhere in the larva. Whether or not the larva changes its morphology and physiology during this process, and the extent to which changes occur, is determined by the production of a second hormone, juvenile hormone by a pair of glands located just behind the brain, the corpora allata. In the presence of juvenile hormone at critical periods ("gates") in the molting process, the larval form is maintained; when this hormone is absent at critical periods, differentiation proceeds under control of prothoracic gland hormones. Whether a particular tissue responds to the presence of juvenile hormone during a particular critical period depends on whether receptors are present or active in those tissues at that time. Prothoracic gland activity is itself stimulated by a hormone secreted by the brain, the prothoracicotropic hormone. Thus, the action of juvenile hormone among insects resembles that of prolactin among amphibians. Only in the absence of juvenile hormone do ecdysteroids suppress the activity of larva-specific genes and prepare pupa-specific genes for their expression. It is not yet completely clear what causes the corpora allata to cease juvenile hormone production. Once the corpora allata are deactivated, however, a juvenile hormone–specific enzyme is secreted that rapidly degrades all juvenile hormone remaining in circulation.

The imaginal discs of holometabolous insects increase in size as the larva grows, but the cells remain undifferentiated until juvenile hormone production stops for a short time during a critical period during pupal life. *See* INSECTA; MOLTING (ARTHROPODA).

Fish. Although coral reef fish exhibit changes in pigmentation and fin structure, number, and location during metamorphosis, the most dramatic metamorphic changes are seen among flounder and other flatfish: in such species, during metamorphosis a symmetrical fish larva becomes an asymmetrical adult, with both eyes displaced to the dorsal surface. The transformation of leptocephalus larvae into juvenile eels is also dramatic; such transformation includes a shift in the position of the urinary and digestive tracts from posterior to anterior. The hormonal control of metamorphosis in fish is not yet well understood, although thyroid gland secretions are clearly involved. *See* EEL; PLEURONECTIFORMES.

Marine invertebrates. The control of metamorphosis among crabs, barnacles, gastropods, bivalves, bryozoans, echinoderms, sea squirts, and other marine invertebrates is poorly understood, partly due to the very small size of the larvae—they rarely exceed 1 mm in length, and most are less than 0.5 mm. There is some evidence that metamorphosis can be stimulated by juvenile hormone or related chemicals (such as methyl farnesoate) in some crustacean species. Most studies have investigated how metamorphosis is initiated by external chemical factors and how those signals are then transduced after their perception.

The larvae of some marine invertebrate species are triggered to metamorphose by specific substances associated with adults of the same species, or with the algae or animals on which they prey. It is not yet clear where the receptors are to which these substances bind, or how the signals result in the genetic switching that drives metamorphosis. In a number of different phyla, including the Porifera (sponges), Mollusca, Annelida, and Echinodermata, metamorphosis can be induced artificially by exposing larvae to certain ions (notably potassium, K^+), neuroactive compounds such as serotonin and gamma aminobutyric acid (GABA), and pharmacological agents such as the nonspecific cyclic nucleotide phosphodiesterase inhibitor isobutyl methylxanthine (IBMX). However, there is much uncertainty in the interpretation of such data as it is not usually clear where in the metamorphic pathway the substances act, or what their precise effects are in the animals under study. Although biologists have long assumed that metamorphosis is brought about through a stimulatory pathway, the possibility that metamorphosis is initiated by the removal of some inhibitory factor cannot be ruled out; indeed, limited evidence suggests such insectlike inhibitory control of metamorphosis in at least one marine gastropod species. *See* ANNELIDA; BIVALVIA; CRAB; DECAPODA (CRUSTACEA); ECHINODERMATA; GASTROPODA; MOLLUSCA.

Timing. Among insects, when metamorphosis occurs is influenced by environmental factors such as temperature, humidity, photoperiod, pheromone production by neighboring individuals, and the nutritional quality of the diet. In a number of species, larvae can undergo a prolonged period of developmental arrest (a diapause) in response to unfavorable environmental conditions, so that metamorphosis can be delayed for many months or even years. The hormonal basis for such effects has been at least partly worked out for a number of insect species.

Among marine invertebrates and in at least some fish species, there is also considerable flexibility in the timing of metamorphosis. At some point in the development of marine invertebrates and apparently also in the development of some coral reef fishes, individual larvae become "competent" to metamorphose. It is not yet clear what makes larvae competent; the development of external receptor cells, or the completion of specific neural pathways, or the activation of hormonal systems or their receptors are likely possibilities. The ability of competent larvae to delay metamorphosis extends 24–48 h in some species to many months in others. Why species differ in this capacity is not well understood.

Deferring metamorphosis in the absence of particular environmental cues should increase the likelihood of metamorphosing into a habitat especially suitable for juvenile survival and growth. In at least a few species, however, juvenile growth rates and competitive ability are reduced if metamorphosis has been delayed too long, so that the potential advantages of delaying metamorphosis may not be fully realized. The capacity of marine invertebrate larvae to delay metamorphosis apparently increases the likelihood of their transport to new, very distant habitats in ship ballast water, and may be an important contributing factor in marine biological invasions.

The metamorphosis of several amphibian species occurs sooner if the temporary ponds occupied by the larvae are drying out, if food supplies become reduced, or if the larvae become too crowded. In consequence, the larvae metamorphose sooner but at a smaller size than normal. Under more benign environmental conditions, the tadpoles metamorphose later and at a larger size. Presumably it is generally advantageous to metamorphose at a larger size; larger frogs are better able to avoid predators, for example. Whether precocious metamorphosis at small size is an adaptive response to pond drying and related conditions (such as increased crowding) is not yet clear, and the hormonal control of such events has not yet been established. Moreover, it is not clear whether the amphibian larvae are delaying their metamorphosis under favorable conditions in the sense that marine invertebrate larvae delay their metamorphosis under conditions unfavorable to subsequent stages; although it is commonly assumed that frog tadpoles become competent to metamorphose at a particular size, there has been no direct confirmation. It seems at least as likely that under some conditions the rate of physiological differentiation is simply accelerated relative to growth. *See* ENDOCRINE MECHANISMS; ENDOCRINE SYSTEM (INVERTEBRATE); INVERTEBRATE EMBRYOLOGY.

Jan A. Pechenik

Bibliography. F.-S. Chia and M. E. Rice, *Settlement and Metamorphosis of Marine Invertebrate Larvae*, Elsevier, New York, 1978; L. I. Gilbert, J. R. Tata, and B. G. Atkinson, *Metamorphosis: Postembryonic Reprogramming of Gene Expression in Amphibian and Insect Cells*, Academic Press, New York, 1996; B. K. Hall and M. H. Wake, *The Origin and Evolution of Larval Forms*, Academic Press, New York, 1999; J. R. Pawlik, Chemical ecology of the settlement of benthic marine invertebrates, *Oceanog. Mar. Biol. Annu. Rev.*, 30:273–335, 1992.

Metasomatism

The process by which the bulk chemical composition of a rock is changed from some previous state by the introduction of components from an external source. In contrast with metamorphism, where rocks are converted to a new set of minerals with little or no change in bulk composition, metasomatism involves the import and export of chemical components through the agency of a chemically active fluid.

Geologic environments and scales. Metasomatism is best developed in environments characterized by extreme physical and chemical gradients and high fluid flux. At the centimeter scale, chemical contrasts along shale-limestone contacts lead to diffusive exchange of components on heating during regional or contact metamorphism. At the kilometer scale in mid-ocean rifts, island arcs, and continental-margin plutonic arcs, metasomatism results from emplacement of magma at depths of a few kilometers and infiltration of hot, saline, aqueous fluids through fractured rocks. At global scales, metasomatism accompanies mass fluxing between the crust and the mantle, such as on emplacement of mantle plumes into the lower crust or subduction of oceanic crust into the mantle. *See* ASTHENOSPHERE; EARTH CRUST; MAGMA; SUBDUCTION ZONES.

Classification and terminology. Classification of types of metasomatism can be based on a combination of factors, such as the types of rocks or fluids involved and the geochemical environment represented by the metasomatic product. Metasomatic types can be identified by the dominant chemical component added to the rock. In some cases, it is more convenient to refer to types of wall-rock alteration with terminology that employs the name of the dominant process (for example, hydrolysis), dominant new mineral formed (for example, sericitic) or component added (for example, hydrogen ion). Thus, the terms hydrogen-ion metasomatism, hydrolysis, hydrolytic alteration, and sericitic alteration all are used to refer to roughly the same metasomatic process.

Quartzo-feldspathic rocks. Clastic rocks and mafic-to-felsic igneous rocks react in similar ways with metasomatic fluids by exchange of alkalies, alkaline earths, and hydrogen. Magnesian alteration (chloritization) occurs in sea-floor basalts and island-arc andesites, where seawater exchanges magnesium for iron and calcium. Chloritization is commonly associated with submarine volcanogenic massive sulfide deposits of copper and zinc. Sodic alteration (albitization) also is common in submarine volcanic rocks and related dioritic plutons, where sodium in seawater is exchanged for potassium and calcium as seawater is heated in the vicinity of magma chambers. Typical minerals produced are albite, chlorite, and actinolite. Sodic alteration accompanies some iron ore (magnetite-apatite) deposits in arc volcanics. Similar albite-rich rocks (albitites) accompany greisen tin-tungsten deposits in peraluminous granites. Calcic alteration of igneous rocks of island-arc settings yields epidote-rich rocks (epidosites) and other calc-silicate rocks (for example, garnet + pyroxene) that resemble skarns. Potassic alteration, the replacement of plagioclase and alkali feldspar by potassium feldspar and of hornblende by biotite, is common in and around apical portions of granitic plutons. In porphyry copper deposits, potassic alteration, accompanied by iron-, copper-, and molybdenum-sulfides, results from cooling of magmatic-hydrothermal brines

released during final crystallization of granitic magmas. *See* ANDESITE; BASALT; GREISEN; PLUTON; PNEUMATOLYSIS.

Hydrolytic alteration dominates lower-temperature metasomatic processes. Hydrolysis involves hydrogen-ion metasomatism—exchange of hydrogen ion for potassium, sodium, and calcium. This is a widespread and varied wall-rock alteration type associated with hydrothermal ore deposits of base metals and precious metals, volcanic fumaroles, and acid hot springs in continental settings. Hydrothermal fluids are either magmatic or meteoric in origin and are acidic and oxidized. White mica (muscovite or sericite), clays (kaolinite, pyrophyllite), and quartz replace feldspars and mafic minerals. Water (as hydrous minerals) and sulfur (as pyrite or sulfate minerals) are commonly added. Subcategories of hydrolytic alteration that represent a series toward increasing acidity of metasomatic fluids and increased leaching of rock constituents are intermediate argillic, sericitic (or phyllic), and advanced argillic. Silicic alteration, in which silica is added to the rock or most other components are removed, is commonly associated with intense hydrolysis. *See* ORE AND MINERAL DEPOSITS; VOLCANO.

Carbonate and ultramafic rocks. Metasomatism of aluminum-poor rocks generally involves addition of silica, metals, and alumina. Iron-silica metasomatism of carbonate rocks at high temperatures (above 300–400°C or 570–750°F) near contacts with intermediate-to-felsic igneous plutons forms skarn, consisting dominantly of iron-rich calc-silicate minerals (for example, garnet, and wollastonite in limestone; olivine and spinel in dolomite). At lower temperatures, iron-silica metasomatism of carbonate rocks forms jasperoid, consisting largely of quartz and hematite. Both of these types of replacement bodies may be accompanied by metal sulfides and sulfate minerals. In some cases, sulfides or iron oxides form separate replacement orebodies near skarn or jasperoid. Ultramafic rocks undergo varying degrees of silication, carbonatization, and hydration, yielding silica-carbonate alteration associated with hot-spring mercury deposits, or talc-carbonate alteration typical of gold-quartz veins in Archean greenstone belts. *See* METAMORPHIC ROCKS; METAMORPHISM; MINERAL; SKARN.

Marco T. Einaudi

Bibliography. P. H. Abelson (ed.), *Researches in Geochemistry*, 1959; H. L. Barnes (ed.), *Geochemistry of Hydrothermal Ore Deposits*, 3d ed., 1997; D. M. Kerrick (ed.), *Contact Metamorphism*, Reviews in Mineralogy, vol. 26, 1991; D. S. Korzhinskii, *Theory of Metasomatic Zoning*, 1970.

Metastable state

In quantum mechanics, a state that is not truly stationary but is almost stationary, as defined below.

Definition. A discrete (not in the continuum) stationary state of a quantum-mechanical system has a wave function that is a product of a function that depends only on the time, and a function that depends only on the other coordinates describing the particles constituting the system, such as their positions and spins. The time-dependent function is a complex number whose modulus remains equal to unity, and which rotates in the complex plane at a rate proportional to E/h, where the purely real quantity E is the energy of the state, and h is Planck's constant, equal to 6.6×10^{-34} joule second. As a result, the probability of finding the system with some specified values of the coordinates is the square of the modulus of the coordinate-dependent function, independent of time, which is why the state is termed stationary. In a metastable (also often called quasistationary) state, the wave function is still well approximated over an extended period of time by the same product of a time-dependent and a coordinate-dependent function, except that now E no longer is a purely real quantity, but is the sum of a purely real number, E_r, which is the resonant energy (often simply termed the energy) of the metastable state and a purely imaginary quantity, $-i\Gamma/2$, where the purely real number $\Gamma > 0$ is the width of the state. This means that the time-dependent function not only rotates in the complex plane at a rate determined by the value of E_r but also decreases exponentially in modulus at a rate determined by the value of Γ. In a metastable state, therefore, the probability of finding the system with some specified values of the coordinates decreases exponentially with increasing time without changes in the state's coordinate probability distribution. In other words, the metastable state decays; the decay lifetime of the metastable state, that is, the time interval required for the probability to decrease to $1/e$ of its original value (where $e = 2.71828$ is the natural base of logarithms), is inversely proportional to Γ. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES; E (MATHEMATICS).

Although the foregoing is quite generally correct, in practice, especially in atomic and nuclear physics applications, the designation metastable state usually is reserved for states whose lifetimes are unusually long. For example, the excited states of atoms usually decay with the emission of a single photon, in a time of the order of 10^{-8} s. However, the necessity for angular momentum and parity conservation forces the second excited state ($2S_{1/2}$) of atomic hydrogen to decay by simultaneous emission of two photons; consequently, the lifetime is increased to an estimated value of 0.15 s. Thus, the $2S_{1/2}$ state of atomic hydrogen is usually termed metastable, but most other hydrogenic states are not. Similarly, emission of a gamma-ray photon by an excited nucleus usually occurs in 10^{-13} s or less; however, the lifetime of one excited state of the ^{113}In nucleus, the state that customarily is termed metastable, is about 100 min. Since radiative transition probabilities for emission of photons generally decrease rapidly with decreasing frequency, a low-lying atomic or nuclear excited state may have a lifetime longer than most excited states of atoms and nuclei and yet not be metastable in the practical sense just described, because photon emission from the state may not be

hindered by any general requirement or selection rule, such as is invoked for the $2S_{1/2}$ state of hydrogen. *See* CONSERVATION LAWS (PHYSICS); SELECTION RULES (PHYSICS).

Physical interpretation. Mathematical analysis supports the following physical interpretation of the existence of metastable states in quantum-mechanical systems: Decay of the metastable state characteristically involves escape of some part of the system to infinity. In radiative decay, for instance, photons are emitted to infinity; the spontaneous decay of uranium-238, with a lifetime of billions of years, is accompanied by emission of an alpha particle. In essence, the system is in a metastable state rather than a discrete truly stationary state (wherein Γ would be zero and the decay lifetime infinite) because there exists some mechanism coupling the putative discrete stationary state to the continuum. In radiative decay, this coupling mechanism is the interaction between the charged particles in the system and the radiation field; in alpha-particle emission the coupling mechanism is the quantum-mechanical ability of the alpha particle to tunnel through a potential barrier at the nuclear surface, a barrier which the particle could not possibly traverse in classical mechanics.

In this alpha-particle emission illustration, the quantum-mechanical metastability can be associated with the fact that the corresponding classical motion of the particle inside the nucleus is stable with respect to small energy perturbations (which keep the particle energy below the barrier peak) but unstable to large energy perturbations (which permit the particle to rise in energy above the barrier peak and thereby to roll down to infinity along the extranuclear tail of the potential). Other illustrative classical physics analogs of quantum-mechanical metastable states include the resonant modes of the electromagnetic field inside a cavity with imperfectly reflecting walls (so that some fraction of the radiation incident on the walls from within the cavity can escape the cavity) and the oscillating modes of a weightless, frictionless string, one of whose end points is attached to a very heavy chain rather than, as is more usual, tied to a fixed immovable point (which permits some of the energy on the string to escape to infinity along the chain). *See* CAVITY RESONATOR; EXCITED STATE; NONRELATIVISTIC QUANTUM THEORY; NUCLEAR ISOMERISM; RADIOACTIVITY; VIBRATION.

Edward Gerjuoy

Bibliography. R. B. Leighton, *Principles of Modern Physics*, 1959; A. Messiah, *Quantum Mechanics*, vol. 1, 1961; A. Temkin (ed.), *Autoionization*, 1966.

Metatheria

An infraclass of therian mammals including a single order, the Marsupialia. The Metatheria are distinguished from the Eutheria (the placental mammals) by numerous characters. The full metatherian dentition is I 5/4 C 1/1 Pm 3/3 M 4/4, for a total of 50 teeth. The braincase is small, the angular process of the mandible is inflected, and a pair of marsupial bones articulates with the pelvis. Almost all living marsupials have a pouch on the belly of the female in which the young are carried after birth. The Metatheria arose from unknown therians in the Cretaceous or earlier at about the same time as eutherian mammals. For a time the two groups evolved side by side, but the marsupials were unable to compete with the more progressive later placental forms and died out except in South America and Australia, where they were isolated by water barriers. *See* DENTITION; EUTHERIA; MAMMALIA; THERIA.

D. Dwight Davis; Frederick S. Szalay

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, 6th ed., 1999.

Metazoa

The monophyletic group of eukaryotic organisms comprising all multicellular animals, but not those in which their cells form colonies, such as choanoflagellates. In colonies, all the cells need to feed because they are not in contact with each other, such that nutrients cannot be transported between them. In multicellular organisms, cells are engaged in a more specialized division of labor because nutrients can be transported between them. Therefore metazoans have specific characters related to cell recognition, adhesion, communication, and skeletal elements responsible for maintenance of the shape of the whole body and its organs. The need for cell-to-cell communication via neurotransmitters evolved into one of the most intriguing systems in all organisms, the nervous system. Other specific characteristics of metazoans are connected to sexual reproduction, such as the origin and structure of haploid gametes, fertilization, and development from the zygote to the adult organism. For most recent authors, Metazoa is a synonym of Animalia.

General morphology. Metazoans are divided into about 35 major lineages that represent unique body plans. Each one of these lineages (named phyla; phylum in singular) can be defined by certain unique characteristics not found in any other phyla. The simplest metazoans are the sponges (phylum Porifera) without a well-defined axis of symmetry, few cell types, with collar cells (choanocytes; flagellate cells lining the cavities of a sponge) similar to those of choanoflagellates, and without true tissues. Sponges are sessile and feed by filtering small organisms from the water by using the collar cells. With a few exceptions, the remaining metazoans have well-defined axes of symmetry; some examples are the radial symmetry of the phylum Cnidaria (sea anemones and jellyfish), the pentaradial symmetry of the phylum Echinodermata (sea stars and sea urchins), or the bilateral symmetry of most other Metazoans. *See* PORIFERA.

Complex metazoans have tissues and organs, including muscle cells that allowed the animals to

become independent from the substrate. This motility probably triggered an arms race of predators and prey escape mechanisms. The same race is probably responsible for the evolution of increased body complexity, size, and development of jaws and skeletons. Prominent examples of animals with exoskeletons are arthropods (including crustaceans, arachnids and insects) and mollusks (snails, clams and their relatives). Endoskeletons are less common among the vertebrates, but are a feature of some mollusks and echinoderms. Different types of jawlike structures evolved in many lineages, and probably played a key role in the evolution of vertebrates. *See* ANIMAL KINGDOM; SKELETAL SYSTEM.

History. Metazoans evolved during the Precambrian (Proterozoic) in the primitive oceans. Most preserved Precambrian animals had simple body plans, probably similar to those of extant sponges and cnidarians. It is believed that most of the extant animal body plans also originated later during the Precambrian times, but that these were inconspicuous animals that did not leave well-preserved fossil evidence. Large metazoans flourished during the Cambrian, and many exceptionally preserved fossil remains are well known from famous deposits such as the Burgess Shale in British Columbia, Canada, or the exquisite Chengjiang localities in Yunnan Province, China. Colonization of terrestrial ecosystems by metazoans is evidenced by fossils in the Silurian Period (ca. 420 million years ago), with likely animal traces in soils even older. Even nowadays most phyla are strictly marine, and only a few have colonized terrestrial environments although still having aquatic relatives or even aquatic life stages. Only one animal phylum, the Onychophora (velvet worms), has no aquatic extant representatives. *See* BURGESS SHALE.

Gonzalo Giribet

Bibliography. R. C. Brusca and G. J. Brusca, *Invertebrates*, 2d ed., Sinauer Associates, Sunderland, 2003; K. M. Halanych, The new view of animal phylogeny, *Annu. Rev. Ecol. Evol. System.*, 35:229–256, 2004; C. Nielsen, *Animal Evolution, Interrelationships of the Living Phyla*, 2d ed., Oxford University Press, 2001.

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

11 METE-NIT

www.MHEST.com

M

Meteor — Myzostomida

Meteor

The luminous streak lasting seconds or fractions of a second and seen at night when a solid, natural body plunges into the Earth's (or another planet's) atmosphere. The entering object is called a meteoroid and, if any of it survives atmospheric passage, the remainder is called a meteorite. Cosmic dust particles (with masses of micrograms) entering the atmosphere and leaving very brief, faint trails are called micrometeors, with the surviving pieces known as micrometeorites. If the apparent brightness of a meteor exceeds that of the planet Venus as seen from Earth, it is called a fireball; and when a bright meteor is seen to explode, it is called a bolide. *See* METEORITE; MICROMETEORITE.

Visual observation. Under normal, clear atmospheric conditions and dark skies (no moonlight or artificial lights), an observer will see an average of five meteors per hour. The spatial distribution of meteoroid orbits relative to the Sun, and the circumstances of their intersections with the moving Earth are responsible for pronounced variations in meteor rates.

As the Earth moves in its orbit, its velocity points toward that part of the sky (sometimes called the apex of meteor velocities) which is visible from local midnight through morning to local noon, and points away from that part of the sky which is visible from local noon through evening to local midnight. On average, an observer sees more meteors during the early morning hours as the Earth sweeps up objects in its path than in the early evening hours when meteor-producing objects must catch up with the Earth.

Physical characteristics. The heights of appearance and disappearance of a meteor depend on meteoroid initial velocity, angle of entry with respect to the vertical, initial meteoroid mass, and meteoroid material strength. The average meteor seen by the unaided eye starts with a meteoroid velocity of 18 mi/s (30 km/s) and leaves a luminous trail from 67 to 50 mi (110 to 80 km) high. The fainter is the meteor (the smaller the meteoroid mass), the shorter is the meteor length. The faster the meteoroid travels just before hitting the atmosphere, the higher in the atmosphere the meteor trail occurs.

In the end, most, if not all, of the meteoroid material is vaporized, leaving a deposit of metallic atoms (predominantly sodium, calcium, silicon, and iron) in the upper atmosphere. This deposition is an important mechanism in the wind-shear formation of certain types of highly ionized, radio-reflecting, upper-atmospheric phenomena called the sporadic E layers. *See* IONOSPHERE.

The meteor trails themselves are rapidly expanding columns of atoms, ions, and electrons dislodged from the meteoroid by collisions with air molecules, and can be excited to temperatures of several thousand degrees Celsius. For a time after trail formation, the free electrons are dense enough to reflect radio waves in the very high frequency range, and therefore can be used to transmit radio messages that are brief (0.1–15 s) but high in information content (on account of their large bandwidth) for up to 1300 mi (2200 km). Since meteor-reflected signals are not as subject to ionospheric and other disturbing influences as are other means of radio communications, there has been interest in using meteors for certain military and commercial purposes. *See* RADIO-WAVE PROPAGATION.

Under the right circumstances, particularly with high-power ultrahigh-frequency (UHF) radars, the ionization right around and moving with the meteoroid itself is seen. This is known as the head echo, and a determination of its velocity is the most accurate way to determine radar meteor speeds. If the radar beam is very narrow, such as that at the Arecibo National Observatory, the radiant and the orbit around the Sun can also be accurately determined. *See* RADAR; RADAR ASTRONOMY.

Some meteors show relatively long-lasting glows along their paths. The glows are called meteor trains.

Novice observers often confuse this glow with the brief meteor wake that is sometimes seen just behind the luminous gases surrounding the meteoroid itself. These trains are subject to winds in the upper atmosphere, and for many years their apparent motions (twisting and drifting) were the only probes of wind velocities in the mesosphere, or middle atmosphere (42–72 mi or 70–120 km high). The motions of meteor trains and trails detected through the Doppler effect on reflected radar signals have been extensively studied, and global mesospheric wind patterns have been derived. *See* DOPPLER EFFECT; MESOSPHERE.

Meteoroid orbits and velocities. The Earth moves around the Sun with an average speed of 18 mi/s (30 km/s). According to the laws of celestial mechanics, if a meteoroid comes from beyond the solar system, its velocity at the Earth's distance from the Sun must be greater than 26 mi/s (42 km/s). If such a meteoroid hits the Earth head-on, indications of preatmospheric speeds in excess of 45 mi/s (72 km/s) would be observed. The fact that the vast majority of observed meteoroids have orbits with Earth-approaching velocities of less than 45 mi/s indicates that most of these are comet and asteroid fragments, and are therefore long-term members of the solar system. Early velocity measurements were quite crude, and the large errors led to the belief that a sizable fraction of meteoroids actually had speeds great enough to escape the solar system or come from the outside. Once these errors were reduced, the number of hypervelocity orbits decreased dramatically, with doubts expressed that any meteoroids could come from beyond the solar system. However, in the 1980s and 1990s, a combination of spacecraft and high-power radar observations indicated that hypervelocity micrometeoids do indeed exist with seeming interstellar dust connections. An even more curious connection is the discovery of similarly sized particles with peculiar isotopic abundances within meteorites of presumed asteroidal origin. These micrometeorites appear to predate the solar system with an apparent origin in early supernovae, novae, or supergiant stars. *See* COSMOCHEMISTRY; INTERSTELLAR MATTER.

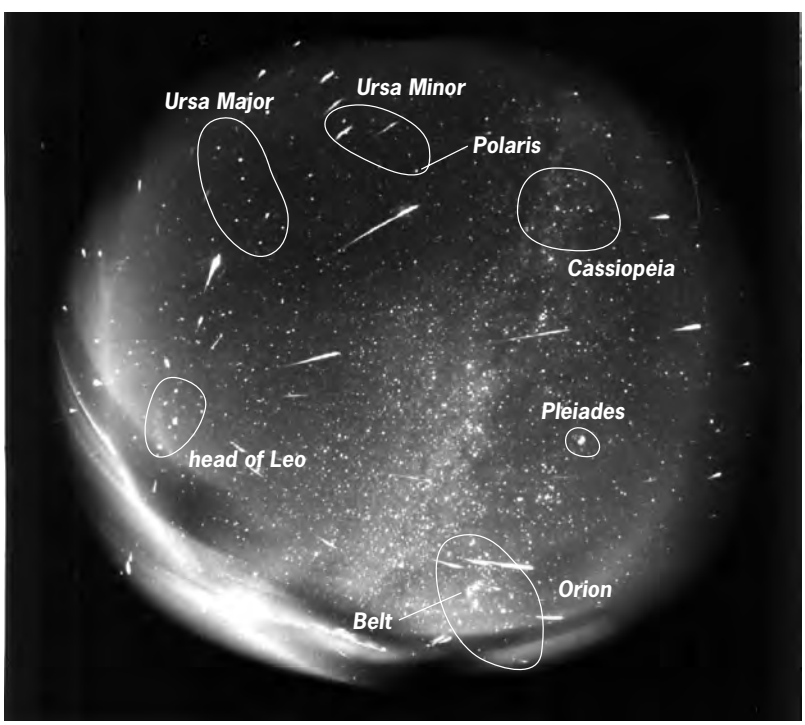

Leonid meteor shower observed at the Modra Astronomical Observatory in Slovakia through a fisheye lens. The exposure on November 16–17, 1998, lasted 4 h. About 150 bolides brighter than magnitude −2 can be seen; the brightest fireball is about magnitude −8. Zenithal hourly rate of the shower estimated from visual observations was about 400. Positions of some well-known constellations and features are indicated. Meteors appear to radiate from the head of Leo. (*Modra Astronomical Observatory*)

The slowest velocity of Earth-meteoroid encounter occurs when the meteoroid has to catch up with the Earth. The gravitational attraction of the Earth keeps such an encounter from producing a zero atmospheric velocity. The velocity that an object will achieve by falling toward the Earth from an infinite distance and in the absence of the Sun's gravitational attraction is 7 mi/s (11 km/s). High-power ultrahigh-frequency radars detect a considerable number of micrometeors with velocities well below the Earth-attraction limit. These are either artificial satellite debris or natural debris captured by the Earth from a large, co-moving interplanetary dust zone that has been discovered in the Earth's orbit. *See* CELESTIAL MECHANICS; ESCAPE VELOCITY; INTERPLANETARY MATTER; ORBITAL MOTION.

Radiants and showers. A combination of the meteoroid's and Earth's velocities of travel around the Sun makes the meteor itself seem to originate from a specific direction in the sky called the radiant. If there are numerous meteoroids in nearly the same orbit (sometimes incorrectly called meteor streams), the Earth sweeps them up at specific times of the year and a so-called meteor shower is observed. Meteor showers are named after the constellation or single star in the sky from which they appear to radiate. While shower meteoroids are really moving nearly parallel through space and result in nearly parallel meteor trails, the effects of perspective make the meteors appear to diverge from the radiant (see **illustration**). If the meteor shower is particularly long-lasting, the radiant will appear to drift slowly in position from night to night as the relative directions of the Earth-meteoroid velocities change. Radiants are not geometric points, but areas in the sky that can be several degrees in diameter. Hence, quoted radiant positions are averages over such areas and often differ somewhat from one compilation to another. Meteors that cannot be shown to be associated with a known shower are termed sporadic meteors.

While all meteor showers show both day-to-day and year-to-year variations, some are more reliable than others, and these are called annual or major showers (**Table 1**). There are over 100 minor showers, some of which are of special interest (**Table 2**). Many minor showers have hourly rates so low that they are often not noticed except by experienced observers or when there is an unusual burst of activity. Some showers have been discovered by radio or radar methods to occur only during the hours of daylight, with no visible counterpart seen at night. Most of the radiants listed refer to geocentric radiants where the effects of the Earth's gravitational

TABLE 1. Major meteor showers

Shower	Duration	Approx. date of maximum	Approx. radiant coordinates, degrees*		Meteoroid orbital speed		Strength†	Suggested parent body	Notes
			Right ascension	Declination	mi/s	km/s			
Quadrantids	Jan. 1–6	Jan. 3	230	+49	27	42	M	2003 EH 1 = extinct (?) nucleus of comet C1490 Y1 (?)	Sharp maximum
Lyrids	Apr. 20–23	Apr. 22	271	+34	30	48	M-W	1861 I	Good in 1982
π Puppids	Apr. 16–25	Apr. 23	110	−45	—	—	W-M	Grigg-Skjellerup	Highly variable
η Aquarids	Apr. 21–May 12	May 4	336	−2	40	64	M	Halley	Second peak May 6
Arietids	May 29–June 19	June 7	44	+23	24	39	S	—	Daytime shower
ζ Perseids	June 1–17	June 7	62	+23	18	29	S	—	Daytime shower
β Taurids	June 24–July 6	June 29	86	+19	20	32	S	Encke	Daytime shower
S. δ Aquarids	July 21–Aug. 25	July 30	333	−16	27	43	M	—	Primary radiant
S. ο Aquarids	July 15–Aug. 25	Aug. 6	333	−15	19	31	W	—	Primary radiant
Perseids	July 23–Aug. 23	Aug. 12	46	+57	37	60	S	1862 III	Best-known shower
Orionids	Oct. 2–Nov. 7	Oct. 21	94	+16	41	66	M	Halley	Trains common
S. Taurids	Sept. 15–Nov. 26	Nov. 3	50	+14	17	27	M	Encke	Known fireball producer
Leonids	Nov. 14–20	Nov. 17	152	+22	45	72	W-S	1866 I	Many peaks seen 1996 to 2003
Puppids-Velids	Nov. 27–Jan.	Dec. 9	135	−48	—	—	M	2102 Tantalus?	Many radiants in region
Geminids	Dec. 4–16	Dec. 13	112	+32	23	36	S	3200 Phaethon	Many bright meteors

* When the stream velocity is unknown or poorly known, the radiant coordinates are those of the apparent radiant rather than the geocentric radiant.
† Estimate of relative meteor hourly rate for visual observers: S = strong (sometimes above 30 per hour at peak); M = moderate (10 to 30 per hour at peak); W = weak (5 to 10 per hour at peak).

attraction have been removed. Radiant positions that have not been corrected for the Earth's attraction (usually because the meteoroid velocities are unknown) are called apparent radiants.

Stationary meteors. Occasionally, a meteor is seen coming directly toward an observer and shows a bright point of light rather than a trail. Such meteors are called, somewhat inaccurately, stationary meteors. The positions of these head-on meteors define the radiant precisely and are quite important to observe, particularly during showers. However, short glints of light from Earth-orbiting spacecraft and space debris are sometimes mistaken for stationary meteors. It has been suggested that a certain

TABLE 2. Minor meteor showers*

Shower	Duration	Approx. date of maximum	Approx. radiant coordinates, degrees†		Meteoroid orbital speed		Suggested parent body	Notes
			Right ascension	Declination	mi/s	km/s		
Coma Berenicids	Dec. 12–Jan. 23	Jan. 17	186	+20	40	65	1913 I	Uncertain radiant position
α Centaurids	Jan. 28–Feb. 23	Feb. 8	209	−59	—	—	—	Colors in bright meteors
δ Leonids	Feb. 5–Mar. 19	Feb. 26	159	+19	14	23	—	Slow, bright meteors
Virginids	Feb. 3–Apr. 15	Mar. 13?	186	+00	21	35	—	Other radiants in region
δ Normids	Feb. 25–Mar. 22	Mar. 14	245	−49	—	—	—	Sharp maximum
δ Pavonids	Mar. 11–Apr. 16	Apr. 6	305	−63	—	—	Grigg-Mellish	Rich in bright meteors
σ Leonids	Mar. 21–May 13	Apr. 17	195	−05	12	20	—	Slow, bright meteors
α Scorpids	Apr. 11–May 12	May 3	240	−22	21	35	—	Other radiants in region
τ Herculids	May 19–June 14	June 3	228	+39	9	15	—	Very slow meteors
Ophiuchids	May 19–July	June 10	270	−23	—	—	—	One of many in region
Corvids	June 25–30	June 26	192	−19	6	11	—	Very low speed
June Draconids	June 5–July 19?	June 28	219	+49	8	14	Pons-Winnecke	Maximum only 1916
Capricornids	July–Aug.	July 8	311	−15	—	—	—	May be multiple
Piscis-Australids	July 15–Aug. 20	July 31	340	−30	—	—	—	Poorly known
α Capricornids	July 15–Aug. 25	Aug. 2	307	−10	14	23	1948 XII (1948n)	Bright meteors
N. α Aquarids	July 14–Aug. 25	Aug. 12	327	−06	26	42	—	Secondary radiant
κ Cygnids	Aug. 9–Oct. 6	Aug. 18	286	+59	15	25	—	Bursts of activity
N. Aquarids	July 15–Sept. 20	Aug. 20	327	−06	19	31	—	Secondary radiant
S. Piscids	Aug. 31–Nov. 2	Sept. 20	6	+00	16	26	—	Primary radiant
Andromedids	Sept. 25–Nov. 12	Oct. 3	20	+34	11	18	—	"Annual" version
October Draconids	Oct. 10	Oct. 10	262	+54	14	23	Giacobini-Zinner	Can be spectacular
N. Piscids	Sept. 25–Oct. 19	Oct. 12	26	+14	18	29	—	Secondary radiant
Leo Minorids	Oct. 22–24	Oct. 24	162	+37	38	62	1739	Probable comet association
μ Pegasids	Oct. 29–Nov. 12	Nov. 12	335	+21	7	11	1819 IV	Probable comet association
Andromedids	Nov. 25	Nov. 25	25	+44	10	17	Biela	Once only, 1885
Phoenicids	Dec. 5	Dec. 5	15	−50	—	—	—	Once only, 1956
Ursids	Dec. 17–24	Dec. 22	217	+76	20	33	—	Good in 1986

* Peak strength for visual observers usually less than 5 per hour. Only showers of special interest are listed here.
† When the stream velocity is unknown or poorly known, the radiant coordinates are those of the apparent radiant rather than the geocentric radiant.

number of visible light pulses from cosmic x-ray and gamma-ray sources are also mistaken for stationary meteors. Such cosmic-ray events are, however, far too brief to be seen with conventional meteor equipment or by the unaided eye. *See* GAMMA-RAY ASTRONOMY; X-RAY ASTRONOMY.

Bright meteors. Extremely bright meteors rivaling even the full moon (often bolides with associated sonic phenomena) are usually not associated with the major showers. Instead, they have meteoroid orbits that are more characteristic of those minor planets and short-period comets that are in highly eccentric orbits in the inner solar system. If the observed luminous-trail end points of these events are less than 12 mi (20 km) high in the atmosphere, there is a good chance that recognizable fragments (meteorites) will fall to Earth and perhaps be recovered. Most fireballs that drop meteorites occur in the afternoon or early evening, when velocities are low. For example, both a meteorite fall on October 9, 1992, which deposited a stone in Peekskill, New York, and one on June 14, 1994, which scattered numerous fragments over Quebec province, east of Montreal, appeared in the early evening. Thus most types of meteorites are believed to be samples of minor planets and possibly certain short-period comets. This connection with minor planets is verified by the fact that the reflecting properties of many minor planets resemble reflections from known meteoritic materials.

It is usually very difficult to detect even the brightest meteors in the atmospheres of other planets, although there have been several instances of spacecraft detection. An exception occurred July 16–22, 1994, when over 20 fragments of Comet Shoemaker-Levy 9 crashed into Jupiter in one of the most spectacular astronomical events ever observed. Even though the impacts happened only on the night side of Jupiter, several of their flashes were seen by reflection from the Jovian satellites. Several spacecraft were able, because of their offset position from the Earth, to view a larger portion of the Jovian night side than was visible from Earth and hence saw the impacts directly. Numerous observatories with large telescopes, including the *Hubble Space Telescope*, watched as hot plumes of material rotated into view at Jupiter's edge. Most impacts left mysterious dark clouds in Jupiter's atmosphere, several of which were visible for months even in small telescopes. Similar impacts on the Earth are thought to be responsible for the extinction of the dinosaurs 65 million years ago, and several other bioextinctions. *See* ASTEROID; COMET; JUPITER.

Origins of shower meteors. A number of meteor showers have been observed to be in orbits that are similar to those traveled by known comets. Thus an association between shower meteors and comets has gradually become a firmly entrenched concept (Tables 1 and 2). There are numerous theoretical scenarios where vaporization of the more volatile cometary ices ejects small solid particles from the surface of the nucleus. A fair proportion of these fragments, particularly the smaller dust-sized ones, escape and take up their own orbits as meteoroids. Cometary nuclei have been known to split into two or more pieces and, when this occurs, it is likely that particles larger than dust size are released as well.

Gravitational attractions of the major planets and the disturbing effect of solar radiation pressure on individual particles tend to spread meteoroids out from the parent object position. Thus "young" showers are those that last only briefly (some as short as an hour or less), while "old" showers may show a few meteors per night but last a month or more. Many showers have a nonuniform structure along their orbits with highest meteoroid densities near the parent body. Lacking orbital synchronism with the Earth's position, these showers do not have an annual appearance at a reliable level. Instead, they show a tendency for strong showings to be separated by intervals roughly equal to the meteoroid orbital periods. The concentration of particles in these orbiting clumps can be relatively high, giving rise to brief deluges called meteor storms, where equivalent rates of thousands of meteors per hour have been noted for times that are at most an hour or so long. These numbers, however, give a false impression of the actual number density of meteoroids in space. With relative velocities on the order of tens of miles per second and a collecting area for each observer of a few hundreds of miles in diameter, the average separation between individual meteoroids is still a few thousand miles. Away from these maxima, however, the meteoroid number densities and hence the observed meteor hourly rates are quite low.

There are some instances where the Earth crosses the meteoroid stream twice per year, giving rise to two separate meteor showers. For example, Comet Halley gives rise to the May Aquarids and the October Orionids, while Comet Encke gives rise to the June Taurids and the November Taurids. *See* HALLEY'S COMET.

While the parent comet idea nicely explains many features of meteor showers, there are problems with this simple picture. First, certain minor planets resemble what might be termed extinct comet nuclei. Some of these have been observed at times with faint atmospheres, a main characteristic of comets. These identifications have been strengthened by the spacecraft observation of the properties of the nucleus of Comet Halley. Second, a perfectly respectable minor planet, 3200 Phaethon, which was discovered by the *Infrared Astronomical Satellite (IRAS)*, has an orbit nearly the same as the prominent Geminid annual shower visible in December. A minor planet, 2003 EH 1 (designated for now as a minor planet, but possibly the extinct nucleus of an ancient comet), has been suggested as the parent of the Quadrantid stream (Table 1). Evidence has also been found for several other minor planet connections with a small number of minor meteor showers. There is convincing evidence that a major source of micrometeoroids with orbits that go out through the asteroid belt is a relatively small number of asteroidal parent objects. There is also accumulating evidence that much of what is termed the sporadic meteor background is really a superposition of millions of ancient meteoroid

streams of various origins—asteroidal, cometary, outer solar system, and possibly interstellar. The meteoroid parent body question is apparently more complicated than previously thought.

The mechanisms that convert cometary or minor-planet fragments into meteoroid streams are not well understood either. The simplistic picture assuming that such fragments are in symmetrical swarms about the parent object are, at best, crude approximations. Thus the prediction of intense meteor storms must be treated with some suspicion. In 1992, when periodic comet Swift-Tuttle, the assumed parent object of the Perseid shower, was again sighted coming in toward the Sun, some astronomers predicted a brief but spectacular Perseid meteor storm for August 1993. In fact, it was nowhere near "storm" level.

Photographic and electronic observations. The strategy of photographic or electronic measurements is to place at least two cameras 10–52 mi (15–85 km) apart over a known baseline, but arranged to examine the same volume of space at a height of about 56 mi (90 km). Each camera has a rotating shutter so that the meteor trail consists of a line of bright dashes (see illustration). Since meteor durations are fairly short, the shutters must spin at speeds of up to 10 times per second to accurately determine the meteor speed. While the early techniques utilized photographic film, ultrasensitive television and other electronic imaging devices have been increasingly used. It has also been customary to place a prism or grating over the front aperture of at least one of the shutter-equipped cameras, so that spectroscopic information can be obtained as well. Television observation of the evolution of meteor trains and wakes has led to a better understanding of upper-atmosphere photochemistry. *See* CAMERA.

Meteor imaging is one of the most difficult areas of astronomical detection, even with ultrafast cameras. Chances for getting an image are better during major showers, because bright meteors are generally more numerous then. Meteor spectroscopy is even more difficult since the light is spread out over areas hundreds of times larger than the meteor trail itself. In spite of these difficulties, several amateur and professional astronomy groups worldwide operate successful bright meteor patrols using photography, but electronic imaging is gaining considerable favor with both professionals and amateurs. *See* ASTRONOMICAL PHOTOGRAPHY; ASTRONOMICAL SPECTROSCOPY.

Radio and radar observations. Radio and radar observations depend on the fact that the initial ion-electron densities in a meteor trail are considerably higher than the average for the ionosphere at an altitude of 56 mi (90 km). For a very high frequency (VHF) or somewhat lower-frequency radar system, the maximum reflected signal occurs when the meteor trail is at right angles to the outgoing wave, with head echoes rarely seen. For a forward-scatter system where the transmitter and receiver are separated, the maximum signal occurs when the meteor trail makes equal angles with the transmitter and receiver lines of sight. As the meteor trail forms, the meteor speed can be inferred from the changing diffraction pattern that results from reflections by different parts of the ionized trail. At ultrahigh frequencies (UHF), radar reflections from the head-echo predominate. From these, high-accuracy radial velocities are determined directly, using the Doppler effect. Thus, original meteoroid speeds can be estimated and, in certain cases, meteoroid radiants and orbits can be determined without reference to optical images. Radio and radar observations generally require a considerable amount of expertise to obtain scientifically valuable results. *See* DOPPLER EFFECT; RADIO ASTRONOMY.

Nontraditional meteor studies and research. Scientific interest in ground-based meteor studies generally declined in North America for three decades from around 1970, when spacecraft methods of directly investigating interplanetary dust became available. In the late 1990s, however, renewed interest in meteor science was generated in the United States with the impending appearance of the Leonid meteor shower produced by comet 1866 I on another of its periodic (33-year intervals) approaches to the Sun. This was due in part to the perceived danger of meteoroids to spacecraft, and in part to rapid advances in applicable research techniques. Along with the usual Leonid ground-based preparations were the increasing professional use of high-power high-frequency (HF), very-high-frequency, and ultrahigh-frequency radars to observe hypervelocity micrometeors, the use of military spy satellites to observe giant bolides, and the increasing use of large jet aircraft as astronomical and geophysical observing platforms.

The 1990s saw claims of interstellar meteors based on ground-based radar investigations, but these particles appeared to have properties greatly different from those seen by the spacecraft instruments. In 2001, it was possible to show that the properties of interstellar micrometeors detected using the Arecibo UHF radar were actually very similar to those of interstellar dust particles observed by the *Ulysses* and *Galileo* space probes, extrapolated to higher masses. *See* SPACE PROBE.

Leonids in 1996–2003. With the appearance of the Leonid comet, it was expected that the strong meteor storm that happened last in 1966 would again make a brief but spectacular appearance. However, perturbations by the outer planets (particularly by Neptune) once again played a significant role in this shower's behavior. Fortunately, this time was unlike the shower's failure to live up to the early predictions in 1899, when virtually nothing happened. The perturbations moved a number of thin meteoroid streams produced by the comet many orbit periods in the past into intersection range of the Earth. This circumstance produced a unique succession of strong peaks covering a span of 7 years. There was a steady increase in the levels of the Leonids starting in 1996, then peaking with an extraordinary number of fireballs in 1999. In those first years (see illustration), a steady rain of fireballs was seen for more than a day worldwide. After 1999, when early predictions had forecast a return to pre-1995 levels, the Leonids kept coming back. They appeared to

decline a bit in 2000, but in 2001 they rose to two spectacular peaks, one rivaling the 1999 peak in Asia, to put on a very spectacular display for both coasts of North America. There was a repeat of two peaks in 2002, but this was not seen in North America. In 2003 two peaks separated by 3 days were predicted, but the observed levels were more modest than those in prior years. The scientific results of this extended display were much greater than anyone had hoped.

Scientific results from the Leonids. While meteor trails had been previously recorded from the space shuttles and other spacecraft, in 1997 the first above-atmosphere, far-ultraviolet spectrum of a bright meteor was recorded during the Leonid shower. Resonance emission lines due to magnesium and ionized magnesium were most prominent in the spectrum.

Spectra obtained from the ground also yielded new information. Both the hydroxyl radical (OH) in the near-ultraviolet and the oxygen molecule in the near-infrared were positively identified for the first time. However, since these are both seen in the airglow of the Earth's upper atmosphere, both may represent the "air" radiation commonly seen in the radiation of high-speed meteors. Such radiation is thought to be produced by constituents in the air surrounding the incoming bodies, rather than from the meteoroids themselves. Searches for the cyanogen radical (CN) have been negative, but now they allow rather strict limits to be placed on its presence and its production in the meteor plasma. In the meteor wake left after the meteoroid passes through the air, infrared spectra show the presence of iron monoxide (FeO). This is likely to have been formed due to the passage of hot meteoric iron through the part of the atmosphere where oxygen is in atomic form (in the mesosphere).

For a number of years, there have been claims of the direct observation of meteoric impacts on the lunar night side, but during several of the 1997–2003 Leonid peaks the number of such claims arose dramatically. For such a claim to be considered seriously, it is not enough to simply record a flash on the Earth-lit side of the Moon. Such flashes can be easily produced by cosmic rays hitting image surfaces of charge-coupled-device (CCD) cameras or photographic film even at ground level. What is needed is confirmation by another image of the same lunar surface, taken simultaneously by another instrument at a different geographic location, showing the flash at the same location on the Moon. During the Leonids, several events satisfying this minimum criterion were reported and so must be considered seriously. What must be sorted out is whether such events were peculiar to the Leonid periods or not. It is possible that detection will not be easily duplicated with other showers with a lesser percentage of fireballs and at a much lesser impact velocity.

The large number of fireballs in the Leonid streams enabled many details of the ablation processes at meteoroid incoming velocities that were higher than average to be recorded with high-speed cameras. Among the more important results is the discovery of a previously unseen shock-wave structure that has little or no "stand-off" distance in front of the meteoroid itself. A second surprising result is that the meteor wake is really two separate structures subject to local air turbulence effects rather than being a single structure with a hollow central zone. The split wakes seen at high resolution appear to be the result of meteoroid fragmentation into two nearly equal parts at heights far above where wake formation takes place.

David D. Meisel

Bibliography. N. Bone, *Meteors*, Sky Publishing, 1993; N. Bone, *Observing Meteors, Comets, Supernovae and Other Transient Phenomena*, Springer, 1998; V. A. Bronshten, *Physics of Meteoric Phenomena*, D. Reidel, Dordrecht, 1983; A. B. C. Lovell, *Meteor Astronomy*, 1954; D. W. R. McKinley, *Meteor Science and Engineering*, 1961; J. M. Pasachoff and W. Tirion, *Peterson Field Guide to the Stars and Planets*, 4th ed., updated, Houghton Miffin, 2004; C. Sumners and C. Allen, *Cosmic Pinball: The Science of Comets, Meteors, and Asteroids*, McGraw-Hill, 2000.

Meteorite

A naturally occurring solid object from interplanetary space that survives impact on a planetary surface. While in space, the object is called a meteoroid, and a meteor if it produces light or other visual effects as it passes through a planetary atmosphere. Various sounds, including hissing and thunderous detonations, have also been reported for large meteors arriving at Earth. Explosive surface impacts by large meteorites are believed to have created the plethora of craters on the solid planets and moons of the solar system. Meteor Crater, Arizona, is Earth's most famous example of an impact crater. *See* METEOR; MICROMETEORITE.

A meteorite seen to strike a surface is known as a fall, whereas a meteorite discovered by chance is known as a find. In both cases, meteorites are named after their geographic places of recovery.

The major classification of meteorites is into chondrites and nonchondrites (or nonchondritic meteorites), and major classes of the nonchondrites are achondrites, stony-irons, and irons, in recognition of their compositions that are dominated by silicate minerals and iron-nickel alloys either alone or as admixtures. Within each of the major categories, detailed classifications are based on distinctive mineralogical and chemical compositions and physical structures (**Table 1**). Group names for unusual types have been chosen based on the established names of first-recognized specimens. For example, shergottites are so named because they belong to the same variety as the meteorite that fell at Shergotty, India, in 1865.

Meteorites represent the most ancient rocks known. Their ages, as determined by radiometric dating, extend to more than 4.5×10^9 years, which is thought to be near the time of solar system formation. As samples of primordial material, stony meteorites known as chondrites are studied for clues about how the solar system formed. In contrast,

TABLE 1. Classification of meteorites[1]

	Nonchondrites[2]			
		Igneous or differentiated		
Chondrites	Primitive	Achondrites	Stony-irons	Irons
Carbonaceous	Acapulcoites[5]	HED[7]	Pallasites[8]	IAB[9]
CI (1)[3]	Lodranites[5]	Howardites	Mesosiderites	IC
CM (1,2)	Winonaites[6]	Eucrites		IIAB
CO (3)		Diogenites		IIC
CR (1,2)		Angrites		IID
CB (3)		Aubrites		IIE
CH (3)		Brachinites		IIF
CV (3,4)		Ureilites		IIIAB
CK (3–6)[4]		Martian		IIICD[9]
Enstatite		Shergottites		IIIE
EH (3–6)		Nakhlites		IIIF
EL (3–6)		Chassignites		IVA[9]
Ordinary		Orthopyroxenites		IVB
H (3–6)		Lunar		Ungrouped
L (3–6)		Feldspathic highlands breccias		
LL (3–6)		Mare basalts		
Other		Mixed breccias		
R (3–6)				
K (3)				

[1]Compiled by Edward R. D. Scott; modified by Michael E. Lipschutz. Not listed are the ungrouped meteorites including 14 carbonaceous chondrites, 85 irons, and a few other igneous meteorites.
[2]Also called nonchondritic meteorites.
[3]Chondrites are also assigned numbers 1–6 called petrologic types, which provide a measure of the degree of low-temperature (0–100°C), aqueous alteration (1,2) or high-temperature, thermal metamorphism (4–6). Type 3 chondrites are the least modified.
[4]There are also numerous C (3–6) samples that are not CK.
[5]These meteorites probably come from the same asteroid.
[6]Closely related to IAB irons with silicate inclusions.
[7]HED meteorites probably come from the asteroid Vesta.
[8]Pallasites are divided into the main group, which are probably related to IIIAB irons, the Eagle Station types, and pyroxene pallasites.
[9]Includes several silicate-rich irons.

achondrites, stony-irons, and irons are samples of melt products formed during processing of solid material in planetary or preplanetary bodies. Chemical analyses of meteorites by geochemists in the 1930s through 1950s supplied first knowledge about abundances of chemical elements (other than hydrogen and helium) in the solar system. Research now includes dissection of meteorites into their many complex components and subsequent analyses of their mineral, microchemical, and isotopic compositions, with the aim of learning when, where, and how the meteorites formed. *See* DATING METHODS; SOLAR SYSTEM.

Meteorite specimens have been recovered from glacial "blue ice" localities in Antarctica that seem to favor surface concentration of meteorites that have fallen on that continent over the past million years. Asteroids are believed to be the sources of most meteorites. In 1982, however, it was conclusively demonstrated that a small achondrite found in Antarctica in 1981 was from the Moon—apparently propelled to Earth at some undetermined time by a large lunar impact event. Since then, many specimens of lunar rocks have been recovered as meteorites from Antarctica and elsewhere, and identified (as of early 2006) with at least 39 meteoroids (many of which fragmented into two or more of the recovered stones on passage through the atmosphere). Even more exciting is the prospect that 34 (as of early 2006) closely related achondrites (24 shergottites, seven nakhlites, two chassignites, and one orthopyroxenite), from various recovery locations around the world, are from Mars; one of them contains trapped gases that are nearly identical to those measured for the Martian atmosphere by the Viking lander in 1976. *See* ASTEROID; MARS; MOON.

Meteoritics, the study of meteors and meteorites, is a premier example of interdisciplinary science, involving chemists, physicists, geologists, and astronomers allied through international research projects. Many advanced laboratory methods, especially for isotope-ratio measurements, have been motivated in large part by meteorite research problems.

James L. Gooding

Chondrites

Chondrites are the most abundant sort of meteorite, over 15,000 being known. They constitute about 82% of all meteorites. Chondrites include three main categories: ordinary, carbonaceous (C), and enstatite (E). (A few Kakangari- and Rumuruti-like chondrites exist; what is known of these 3 K and 19 R chondrites does not modify the overall chondrite picture.) Chondrite categories are based on mineral abundances and compositions; mineral shapes, sizes, and relationships (that is, textures); and bulk chemical compositions. Chondrites are so-named because nearly all contain small (generally 0.5–2 mm) bead-like chondrules. Only CI chondrites are essentially chondrule-free; their matrix texture and chemical

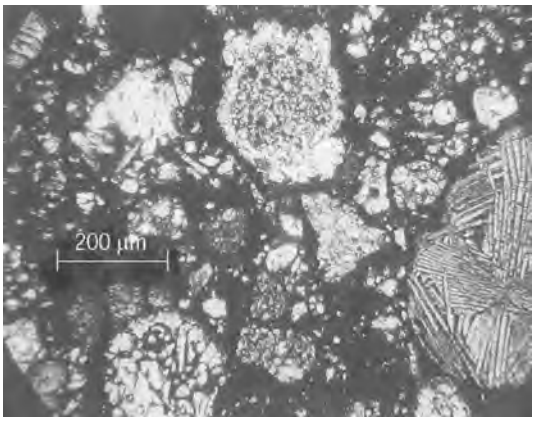

Fig. 1. Typical unequilibrated chondrite: many chondrules are present in a fine-grained mineral matrix.

composition closely relate them to the other chondrule-containing carbonaceous chondrites. The melting of chondrite-like material in parent bodies, followed by segregation and solidification of the resulting iron-nickel metal and silicates, produced the igneous meteorites.

Ordinary chondrites. These constitute 94% of all classified chondrites. Their main constituent minerals are olivine [$(Mg,Fe)_2SiO_4$], low-calcium pyroxene [$(Mg,Fe)SiO_3$], plagioclase [$(Na,Ca)Al(Si,Al)Si_2O_3$], iron-nickel (Fe-Ni) metal, and troilite (FeS). They may contain silicate glass. *See* FELDSPAR; OLIVINE; PYROXENE; PYRRHOTITE.

Chondrules and chondrite matrixes consist of the same minerals: the difference is textural. Chondrule minerals crystallized within molten droplets, and show a variety of shapes consistent with a molten origin involving very rapid (minute-scale) heating and cooling. Matrix minerals are small and granular (**Fig. 1**).

Olivine and low-calcium pyroxene are iron (Fe^{2+}) and magnesium (Mg^{2+}) solid-solution minerals (that is, homogeneous crystalline phases with Fe^{2+} or Mg^{2+} ions located at random lattice points and existing in a range of concentrations). Their compositions in ordinary chondrites vary from fayalite (Fe_2SiO_4) 16 [olivine with 16% Fe^{2+} ions by number of moles, symbolized Fa16] to Fa33, and from ferrosilite ($FeSiO_3$) 15 (low-calcium pyroxene with 15 mole % Fe^{2+}, symbolized Fs15) to Fs25. The variation is discontinuous, and two distinct 1–2% Fa gaps in olivine composition separate ordinary chondrites into three groups: the H group (Fa16 to Fa20; high total and metallic iron), the L group (Fa22 to Fa25; low iron), and the LL group (Fa26 to Fa33; low-low iron). In addition to olivine composition, these groups are characterized by the ratio of iron in metallic minerals (iron-nickel metal and troilite) to total iron (including iron in silicate and oxide minerals). This ratio is 0.6 in the high-iron group, 0.3 in the low-iron group, and 0.1 in the low-low-iron group. High-iron and low-iron chondrites each constitute 40–44% of all chondrites; low-low-iron chondrites are about 9%. *See* SOLID SOLUTION.

Within each ordinary chondrite group, a sequence of textural and mineralogical changes suggesting metamorphism are designated numerically: H3–H6, L3–L6, and LL3–LL6. Each subgroup number connotes similar characteristics. For example, type 3 ordinary chondrites exhibit abundant, very sharply defined chondrules containing fine-grained minerals (olivine, low-calcium pyroxene, with or without metal and troilite) aggregated in silicate glass (Fig. 1). Compositionally, the glass varies from chondrule to chondrule; however, it is plagioclase-like. Within chondrules, and in the surrounding granular matrix, olivine and low-calcium pyroxene exhibit large grain-to-grain compositional differences. Since adjacent grains have different fayalite and ferrosilite contents, these meteorites are called unequilibrated, a characteristic of type 3 ordinary chondrites. The nickel contents of metal grains also vary. Crystalline feldspar is absent; only chondrule glass, compositionally similar to plagioclase, exists.

The characteristics of each subgroup grade into those of the next subgroup. This is particularly apparent in type 3 ordinary chondrites, the subgroups with the most disparate properties. Each subgroup (H3, L3, and LL3) is divided into 10 subtypes (3.0–3.9) using thermoluminescence sensitivity (a measure of plagioclase crystallization from glass), variability of fayalite in olivine and of cobalt in metal, extent of matrix recrystallization, and fayalite and ferrosilite values in matrix relative to whole-rock averages. Thus, subtype 3.0 chondrites, which differ most from those of type 4, have the lowest thermoluminescence, greatest fayalite and cobalt variabilities in olivine and metal, respectively, least-recrystallized matrix, and highest fayalite and ferrosilite values in matrix relative to whole-rock. Ultimately, type 6 chondrites contain no glass but contain crystalline plagioclase. Olivine or low-calcium pyroxene grains are compositionally uniform; thus, type 5 and 6 chondrites are equilibrated. Average grain sizes are larger in such chondrites and chondrules become nearly unrecognizable (**Fig. 2**). Dominant subgroups differ: H5 and LL6 are the most common high-iron and low-low-iron chondrites, respectively, while L6 are the majority of the low-iron group.

Only collisional impacts can disrupt asteroids or planets. A variety of textural alterations, including localized or generalized melting, serve as barometers to estimate peak shock pressures. These can

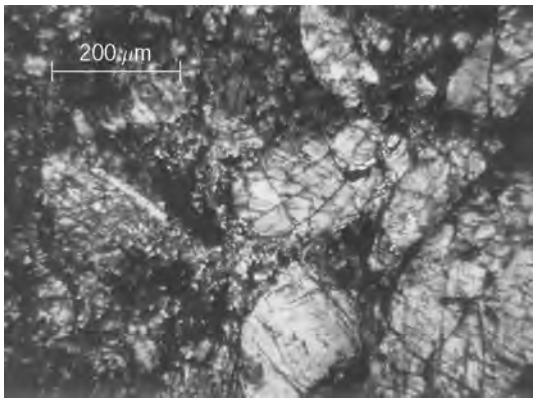

Fig. 2. Typical equilibriated chondrite. No complete chondrules are evident, and the matrix consists of relatively coarse mineral grains with possible poorly defined chondrule fragments.

exceed 500,000 atm (50 GPa), with resultant residual temperatures up to 1500°C (2730°F), persisting for months to years because collisional debris can be large, cooling slowly. About one-third of L4–L6 chondrite falls have been strongly shocked (300,000–500,000 atm); H4–H6, LL4–LL6, and other chondrites exhibit smaller percentages (≤14%).

Carbonaceous chondrites. These constitute less than 5% of chondrite falls and are the least evolved meteorites. The same major chemical elements compose all chondrites, but carbonaceous chondrites can contain substantial carbon, hydrogen, and nitrogen, present at trace levels in ordinary chondrites. In addition to chondrules, they contain many other types of inclusions. Most important, the constituent minerals of chondrules and inclusions in carbonaceous chondrites differ in composition and kind from those composing the surrounding matrix.

In carbonaceous (C) chondrite classification, prototypical resemblance is important, occasioning a second letter. Most carbonaceous chondrites belong to classes CI (mainly types 1 and 2), CV (mainly types 3 and 4), CM, and CO (whose prototypes are meteorites named Ivuna, Vigarano, Mighei, and Ornans, respectively), so these classes are the best studied.

In normal type 1 chondrites, chondrules are extremely rare. Their matrix is 99% clay minerals and similar layer-lattice silicates with water as a structural component; CI chondrites average about 20% water by weight. *See* CLAY MINERALS; SILICATE MINERALS.

The CI matrix contains about 1% small (<100 micrometer) olivine and pyroxene crystals that are refractory minerals formed at high temperatures (750–1500°C or 1380–2730°F). Metal is entirely absent. The clay matrix necessarily formed at low temperatures (less than 120°C or 250°F). Thus, these meteorites are extremely unequilibrated. Epsomite and gypsum veins run through the CI matrix. The matrix includes a complex of organic molecules, most being poorly characterized polymers (designated kerogen), formed nonbiologically. These compounds and graphite account for the high CI carbon content, 3.6%. *See* EPSOMITE; GRAPHITE; GYPSUM; KEROGEN.

CM chondrites consist of about 50% clay and layer-lattice silicates (like those in CI chondrites) containing water, associated with organic compounds. The remaining half are refractory minerals in several forms, including single crystals, crystal fragments, chondrules as sharply defined as those in type 3 chondrites, and irregularly shaped inclusions (Fig. 1 and **Fig. 3**). Olivine and low-calcium pyroxene are the most common refractory minerals, and chondrules constitute only about 2% by volume. Metal is very rare. The textures of some inclusions suggest that they were loose grains aggregated before incorporation into the matrix. A small number of inclusions consist of extremely refractory minerals, composed of titanium, aluminum, and calcium oxides and silicates, formed at temperatures above 1500°C (2730°F). The combination of clay-rich matrix and refractory inclusions means these meteorites are highly unequilibrated (Fig. 3).

C3 chondrites, mainly CO and CV, contain much less water (~1%) and carbon (~0.5%) than do C1 or C2 chondrites, because clays and layer-lattice silicates are rare in them. Instead, 35–40% (by volume) of their matrix consists of submicrometer blades of high-fayalite olivine (~Fa50). The remaining 60–65% consist of larger single crystals, fragments, chondrules, and inclusions of refractory minerals (olivine and low-calcium pyroxene). Highly refractory inclusions are relatively abundant. Thus, these meteorites are also unequilibrated. C4–C6 chondrites are rare, they constitute less than 15% of all carbonaceous chondrites, mainly metamorphosed CK chondrites.

Enstatite chondrites. These make up not more than 2% of all chondrites and consist of 60–80% enstatite (Fs0, that is, $MgSiO_3$) with 10–30% metal and 5–15% troilite. Enstatite chondrites range from those with many distinct chondrules to those having nearly unrecognizable chondrules. Thus, like ordinary chondrites, they fall into subgroups E3–E6. Contents of major elements differ, with EH (mainly types 3 and 4) containing more iron, and so forth, than EL (mainly types 5 and 6). E3 chondrites contain glass but no plagioclase, while E5 and E6 chondrites contain plagioclase but no glass; E4 is transitional. Common minor components are cristobalite, tridymite, and quartz. *See* ENSTATITE.

Minor and trace minerals in the matrix are quite extraordinary. Enstatite chondrites contain oldhamite (CaS), sinoite (Si_2ON_2), and osbornite (TiN) while, in other chondrites, calcium, silicon, and titanium are always combined with oxygen as silicate and oxide minerals. These unusual minerals can exist only under conditions of low oxygen activity, that is, extremely reducing environments. This is consistent with low-calcium pyroxene in them being Fe^{2+}-free (Fs0). The matrix contains abundant microscopic graphite (0.4–0.8% by weight): organic compounds are absent.

Origin. Chondrites contain components from two genetic environments: nebular (those formed from dispersed materials in space) and planetary (those formed within parent bodies).

Carbonaceous chondrites contain the most obvious nebular components. Refractory mineral inclusions in them formed by direct condensation from a hot gas cloud surrounding the primitive Sun. During

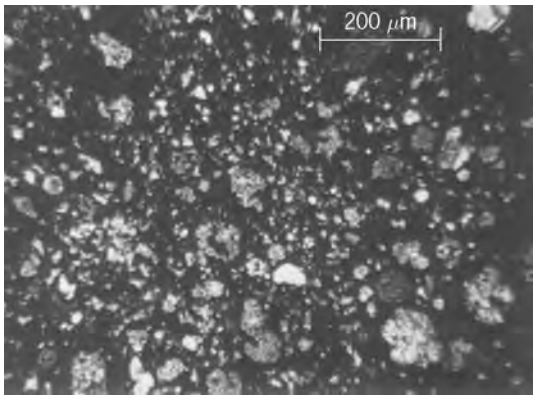

Fig. 3. Type 2 carbonaceous chondrite: numerous small grains of refractory minerals occur in a dense matrix of clay and other layer-lattice silicate minerals.

late-formation stages, as temperatures in the nebular gas fell, lower-temperature minerals formed and accreted as matrix, into asteroidal bodies together with refractory minerals. Thus, of all meteorites, carbonaceous chondrites are generally richest in volatile chemical elements. Most (about 90%) were not subsequently metamorphosed at high temperatures in their parent bodies, which would have erased their primitive characteristics. Carbonaceous chondrites, however, are not entirely pristine. Many experienced some low-temperature aqueous alteration; how much was nebular and how much was planetary is unknown.

Petrologic and chemical alterations in artificially heated chondrites demonstrate that about 10% of the C1–C3 chondrites were thermally metamorphosed at temperatures up to 900°C (1650°F) as open systems (CKs as closed systems) in parent body interiors. Impacts excavated these materials and transported them elsewhere. Spectroscopic studies show that such metamorphosed material comprise surfaces of four asteroid types (C, G, B, and F). A texturally and compositionally unique C1-like chondrite that fell on January 18, 2000, in the Yukon is spectroscopically linked to previously unsampled D-class asteroids. Its orbit could be determined because it was imaged from two locations during atmospheric passage. Like orbits determined previously for an enstatite and five ordinary chondrites, its maximum solar distance was in the asteroid belt.

No consensus exists about the nebular condensation and accretion histories of ordinary and enstatite chondrite parent materials into asteroid-sized bodies but accretion (planetary) processes are broadly understood. The progressive equilibration from type 3 through type 6 apparently reflects planetary metamorphism that obscured chondrules and homogenized mineral grain compositions. Most ordinary chondrites and some enstatite chondrites exhibit brecciation or more severe impact-shock effects caused by parent-body collisions. The one or more collisions involving the one or more L chondrite parents occurred 500 million years ago and must have been exceptionally severe, since radiogenic ^{4}He and ^{40}Ar were lost then as were many volatile trace elements. *See* BRECCIA.

Spectroscopically, only the rare S IV type asteroids seem viable parents for the plentiful ordinary chondrites. By consensus, this paradoxical situation seems due to space weathering, mineral spectral alteration by solar particle impacts on asteroid surfaces.

Each chondritic group contains minerals with characteristic stable oxygen isotopic compositions so that each derives from its own batch of nebular material. It is not known how the very numerous LL, L, H, and E (or very rare R and K) groups are specifically related to each other. The LL-chondrite minerals formed under the highest oxygen activity, and the E-chondrite minerals, the lowest. It is tempting to interpret this as indicating formation at different distances from the primitive Sun: E-chondrites formed nearest the Sun, and the H, L, LL, and carbonaceous chondrites formed progressively farther away. While this is hypothetical, contents of volatile trace elements are lower in mildly shocked (less than 300,000 atm or 30 GPa) H4–H6 chondrites than in analogous L4–L6 samples, suggesting that H4–H6 parent material condensed at higher temperatures, that is, closer to the Sun. Two conclusions are certain: all chondrites could not have come from a single parent, and chondritic groups record different dominant genetic episodes.

Michael E. Lipschutz

Chondrules. These are the most abundant particles in chondrites, with diameters up to about 1 mm (0.04 in.). They are distinguished from other particles (such as unmelted aggregates and calcium-aluminum-rich inclusions or CATs) in the chondrite matrix by textural and compositional criteria, being iron-magnesium silicate spherules with igneous textures. Chondrites themselves, when not modified much by heating in their asteroidal parent bodies, have sedimentary textures. The chondritic asteroids are aggregates of melted droplets, which cooled in the protoplanetary accretion disk (solar nebula). Some major astrophysical (or possibly geological) process clearly operated in the early solar system and caused extensive melting. Chondrules are therefore a key to preplanetary history, but understanding them is complicated by the fact that there are several different kinds, which may not all have formed in the same way.

Mineralogy and textures. The most abundant minerals in chondrules are magnesian to intermediate olivine (fayallite 0 to Fa60), pyroxene (mainly inverted protopyroxene), glass rich in a feldspar component, iron metal, and iron sulfide (FeS). Chondrules are called type I if they have magnesian silicates (Fa $<$ 10) and type II if the have more ferroan silicates (Fa $>$ 10). The less common chondrules with compositions intermediate between the common iron-magnesium-rich kind and refractory inclusions (CAIs) have one or more aluminum-rich mineral such as spinel, plagioclase feldspar, or aluminum-rich pyroxene (fassaite). Many olivine crystals in chondrules are compositionally zoned, with enrichment in iron, calcium, and other elements toward their rims. Some pyroxene solid solutions exsolved on cooling, but the precipitates can be observed only in the electron microscope.

The main textural types (**Fig. 4**) are granular (including porphyritic on a submicroscopic scale), porphyritic (more strictly, microporphyritic), barred (parallel-plate crystals), radial/excentroradial, and cryptocrystalline/glassy, based on abundance, size, and shape of crystals. (Porphyritic texture is characterized by relatively large crystals dispersed in a matrix of finer-grained or glassy material.) A shorthand notation for chondrules combines the texture and one or more dominant minerals; thus there are BO (barred olivine), RP (radial pyroxene), and POP (porphyritic olivine-pyroxene) chondrules. The presence of glass, euhedral and skeletal crystals, and inhomogeneous crystals indicates crystallization of liquids at moderate cooling rates, on a time scale of hours, similar to thin lava flows. *See* IGNEOUS ROCKS.

Chondrules were not all single totally melted droplets. There are rare CAIs and forsteritic

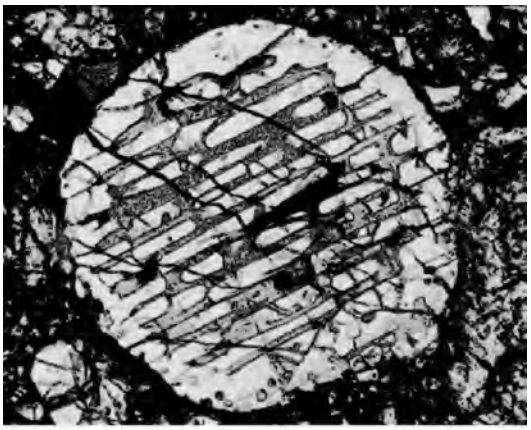

(a)

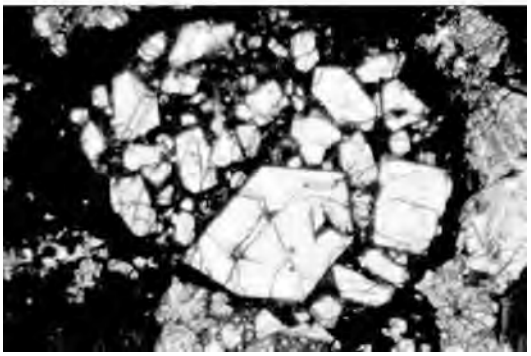

(b)

Fig. 4. Examples of chondrules, showing different textures. (*a*) Barred olivine chondrule. (*b*) Type IIA porphyritic olivine chondrule.

ameboid olivine aggregates inside chondrules. Some chondrules enclose other chondrules. Others are aggregates of patches of somewhat different grain size. Some are layered in a way that suggests accretion. Some have igneous rims. Whether chondrule formation generally involved dustballs that were melted, or a spray of fine droplets and grains that accreted, the occurrence of these complexities on the whole suggests formation of chondrules by incomplete melting of preexisting solids. However, a number of cryptocrystalline chondrules may have formed by condensation of liquid from the gas.

Some chondrules contain grains not crystallized from the chondrule melt. Forsteritic (type I) olivine grains are common in type II chondrules and relatively ferroan (type II) olivine with iron metal inclusions in type I chondrules. The two types of relict olivine grains suggest several generations of type I and type II chondrules, whose debris is incorporated into later chondrules.

Chemical and isotopic data. Type II chondrules, besides being ferroan, have approximately chondritic elemental abundances, with the volatile sodium and potassium present in higher concentrations relative to silicon than in the solar or CI chondrite composition. Type I chondrules are depleted in volatiles and enriched in refractory elements. They have a composition very similar to the primitive Earth mantle and may have been important in its accretion. Both type I and type II chondrules exist in silicon-poor and silicon-rich forms, that is, in olivine-rich and pyroxene-rich varieties, given the suffixes A and B, respectively. Thus, there are type IA and IIA PO chondrules, type IAB POP chondrules, and so forth. Some cryptocrystalline chondrules are depleted in both refractory and volatile elements relative to the major element, silicon. Chondrule compositions do not vary like those of planetary igneous rocks, but are controlled by volatility-related processes, such as evaporation and condensation.

The occurrence of very fine grained type I chondrules rich in sulfur and other volatile elements whose concentrations decrease in coarser-grained similar chondrules suggests evaporative loss with more extended heating. However, the occurrence of olivine-rich cores in some chondrules, surrounded by pyroxene-rich mantles and sometimes silicon-rich rims, suggests condensation of silicon. At chondrule melting temperatures, silicate vapor rather than melt is stable, and chondrules formed in the solar nebula would therefore be expected to at least begin to evaporate. Free evaporation into a vacuum causes isotopic (Rayleigh) mass fractionation, and this process is observed in some refractory inclusions, but not in chondrules. Condensation at the pressures of the solar nebula would normally produce crystals rather than liquids.

The age of the solar system is given by the lead isotopic age of CAIs to be 4567.2 ± 0.6 million years. Similar measurements, and also the ^{26}Al-^{26}Mg relative ages, indicate that chondrule formation continued for about 2 million years after this. *See* COSMOCHEMISTRY.

Chondrules in carbonaceous chondrites are enriched in ^{16}O relative to the Earth but less so than CAIs and ameboid olivine inclusions. Those in ordinary chondrites are depleted. The highest ^{16}O contents in chondrules may be related to the presence of forsteritic olivine relict grains, which have been less modified than the rest of the chondrules by exchange with the nebular gas.

Crystallization experiments. Chondrule simulation has been concerned mostly with duplicating their textures and olivine zoning by melting pellets of pressed mineral powder suspended on platinum wires. Cooling at 10–100°C/h (18–180°F/h) from temperatures a little lower than those which give total melting, about 1500–1600°C (2700–2900°F), is ideal for growing porphyritic textures. The grain size of the synthetic chondrule is very sensitive to the grain size of the starting material, because partial melting leaves a number of nuclei proportional to the initial number of crystals. Chondrule textures have also been produced after total melting of the charge by injecting mineral grains into the melt and inducing crystallization, and by spraying mineral dust into the hot zone of the furnace and collecting the resulting droplets on wire holders.

When chondrule analog are heated at low pressures, about 10^{-5} atm (1 Pa), as in the protoplanetary disk, sulfur is lost almost immediately, and sodium is severely depleted in the first few minutes, but ferrous oxide (FeO) is lost over a period of about 18 h. Iron-poor chondrules would therefore be expected

to be free of sodium, but they are not. At these low pressures, evaporative loss reproduces part of the composition trend of natural chondrules: we get first type IIA compositions, then IAB, and finally IA with up to 18 h heating at about 1550°C (2800°F), all with porphyritic textures. However, the charges show isotopic mass fractionation due to this free evaporation, while natural chondrules do not. When chondrule analog are heated in an atmosphere containing sodium, potassium, or silicon, loss of these elements has been stopped or even reversed. Pyroxene-rich rims on olivine-rich charges were formed by silicon condensation, as in some type I chondrules. Evaporation in the presence of a vapor containing the lithophile (rock-forming) elements greatly reduces their isotopic mass fractionation.

Texture-composition-temperature relationships. Chondrules with skeletal crystals can be reproduced in experiments with melting at near-liquidus temperatures (temperatures of total melting), and porphyritic chondrules with temperatures about 50°C (90°F) below the liquidus. Chondules have a wide range of bulk compositions and therefore a wide range of liquidus temperatures (~1200–1800°C), but are dominantly porphyritic. If chondrules had formed by closed-system melting, this would be surprising: the melting temperature would appear to have been chosen to be always about 50°C (90°F) below the liquidus for each composition. Open-system melting resolves this problem. A porphyritic texture is established during heating and the composition changes during evaporation, giving a liquidus about 200°C (360°F) higher than the peak heating temperature.

Origin. Chondrules formed either by the melting of rock, dustballs, or free dust grains, or by the condensation of gas, and possibly by all of these mechanisms. Melting in the nebula should give mass-fractionated residues and condensation should give crystals. Condensation of liquid and evaporation to yield isotopically normal residues are both possible in the presence of an atmosphere highly enriched (about 100-fold) in rock-forming elements. Such an atmosphere may be achieved by concentrating solids into small domains and then evaporating them. Dust would be totally evaporated before larger aggregates, and both evaporation and condensation or recondensation would be expected. The concentration may possibly be due to accretion of planetesimals, to convection currents in the nebula that concentrate grains in stagnant zones between eddies, or to the delivery of mass to the inner edge of the disk as part of solar accretion. It seems that chondrules must form in less than 1% of the nebula in order to achieve the nonnebular gas composition. This local rather than nebula-wide formation is consistent with the relatively rapid cooling rates of chondrules.

Formation of a solar system involves numerous sources of energy that could have been responsible for chondrule melting. Mechanisms that have been considered include flares from the young Sun, lightning in the nebula, and collisions between hot planetesimals. During solar accretion, much mass is ejected in bipolar outflows, and particles melted by solar radiation might have been thrown back into the disk. The physical conditions during these events are not sufficiently known to see the implications for the properties of chondrules formed in this way. Chondrules might have been formed by heating when shock waves propagated through the disk encountered concentrations of chondrule precursors. This process has been modeled in detail and is the mechanism whose temperature distributions best explain the relative proportions of different kinds of chondrules. How these shocks were generated needs to be established; spiral arms in an unstable massive disk may provide the mechanism. Roger H. Hewins

Achondrites

These are stony meteorites that have few, if any, chondrules and differ chemically from chondrites. They constitute about 8% of all meteorite falls and 1% of all finds. Although achondrites can be divided into several distinct groups based on chemical and isotopic composition, they are generally believed, based on aspects of their textures and composition, to have formed as the result of igneous processes on asteroidal or planetary bodies. Much of the interest in these meteorites derives from the fact that they provide clues into the nature of igneous processes and planetary differentiation early in the history of the solar system on planetary bodies outside the Earth–Moon system and on bodies presumed to be much smaller than the Earth and Moon.

Basaltic achondrites. The eucrites, howardites, and diogenites—often collectively referred to as the basaltic achondrites—are the most abundant achondritic meteorites. They appear to be samples of a series of related igneous rocks and of regolith breccias composed of fragments of these igneous rocks. They (along with the stony-iron meteorites, that is, mesosiderites and certain pallasites, and the iron meteorites) define a coherent group in terms of their oxygen isotope compositions, suggesting they are closely related. With ages near 4.5 billion years, they are products of igneous activity from the earliest history of planetary bodies in the solar system. They have been studied vigorously since the early days of the Apollo program, for they contain the most detailed record of early planetary differentiation, of igneous activity outside the Earth–Moon system, and of igneous activity on asteroid-sized bodies, from which they are presumed to have been derived.

Eucrites are dominantly composed of low-calcium pyroxene (usually pigeonite) and anorthitic plagioclase, with subordinate amounts of chromite, iron-rich metal, high-calcium pyroxene, silica polymorphs, and other minor phases. They are often, but not invariably, brecciated; most of the brecciated samples are monomict, but many of the more recently discovered samples from Antarctica are polymict. Many of the eucrites have pyroxene and plagioclase in ophitic and subophitic textures similar to those observed in terrestrial and lunar basalts. The textures of most of these eucrites and aspects of their chemical compositions, which cover a small

range compared to terrestrial and lunar basalts, suggest that they are samples of crystal-poor magmas. Some of these eucrites have vesicles, suggesting an extrusive or hypabyssal origin. A small number have gabbroic textures and appear to be cumulates from liquids related to the more common eucrites. Many eucrites show features attributed to shock metamorphism and to thermal metamorphism. *See* METAMORPHISM; PIGEONITE.

The diogenites are composed nearly entirely of magnesian orthopyroxene, with subordinate amounts of olivine, silica polymorphs, plagioclase, metal, chromite, and other minor phases. They are typically monomict breccias, but some are unbrecciated, and others contain minor amounts of eucritic material. They range texturally from coarse- to fine-grained and show evidence of thermal metamorphism. They are generally believed to represent cumulates from magmas related to, but more magnesian than, known eucrites.

Based on comparison with lunar regolith breccias, with which they have some similarities, the howardites are generally considered to be samples of the regoliths of the parent planets of the eucrites and diogenites. They are polymict breccias that consist dominantly of angular to subrounded basaltic and pyroxenitic clasts and mineral fragments set in a finer matrix. They also contain chondritic fragments and rare glass beads (presumed to result from impact melting). Most of the fragments found in the howardites resemble eucritic and diogenitic material, and howardite bulk chemical compositions can be approximated as mixtures of these two types of meteorites. However, howardites also contain igneous fragments with textures and chemical compositions unknown among eucrites and diogenites. Many lithic and mineral clasts in howardites experienced a complex range of shock and metamorphic processes prior to incorporation into howardites. Ages of clasts in howardites extend from close to those of the oldest eucrites to 2×10^9 years younger; the young ages are usually interpreted as due to secondary, nonigneous processes.

The eucrites, diogenites, and individual fragments in howardites contain the record of extensive early igneous activity on what are generally assumed to be asteroidal parent bodies. The magmas from which this series of igneous rocks formed are thought to have been produced by partial melting of volatile-depleted, metal-depleted, olivine-rich planetary interiors. Although the relative roles of differing degrees of partial melting and fractional crystallization in the formation of the suite of igneous rocks sampled by these meteorite groups are controversial, it is clear that igneous processes as they are known from study of terrestrial and lunar rocks were active on small bodies very soon after their formation. The heat source for such igneous activity is still under investigation, but it could be the decay of the aluminium isotope ^{26}Al or perhaps heating by electric currents induced in small planets by the passage of an intense solar wind associated with a very active early Sun (T-Tauri phase). The reflectance spectrum of the surface of the asteroid 4 Vesta closely resembles those of eucritic meteorites, and it has been suggested that this could be the source of the basaltic achondrites, although there are dynamical difficulties with such a source. *See* BASALT; IGNEOUS ROCKS; SOLAR WIND.

Shergottites, nakhlites, and chassignites. These are rare meteorites, but there is considerable interest in them because it has been suggested that they may have come from Mars. The 24 shergottites (as of early 2006) are composed primarily of pigeonite and augite pyroxenes plus glass with the composition of an intermediate plagioclase feldspar. The plagioclase glass is known as maskelynite, formed from crystalline plagioclase by shock metamorphism. The shergottites also contain titanomagnetite and rare hydrous amphibole; some contain olivine. The textures are similar to terrestrial diabases, and they have been interpreted as partial cumulates. Determination of their ages is controversial and may be complicated by the effects of shock metamorphism, but the ages are generally between 1.8×10^8 years and 1.3×10^9 years. Aspects of the chemical and isotopic composition of the shergottites suggest that they are derived from a parent body that experienced a complex, multistage history extending over much of the history of the solar system. Although they are basaltic like the eucrites, the shergottites are distinguished by their much higher volatile contents, the presence in them of iron(III) ion (Fe^{3+}, as opposed to metallic iron in the eucrites), and their much younger ages. In these respects the shergottites are more similar to terrestrial basaltic rocks. *See* AMPHIBOLE; DOLERITE; ROCK AGE DETERMINATION.

The seven nakhlites (as of early 2006) are augite-rich rocks whose textures strongly suggest they are cumulates. Augite appears to be the only cumulus phase; other phases include iron-rich olivine, pigeonite, sodic plagioclase feldspar, alkali feldspar, and titanomagnetite. The nakhlites contain hydrous alteration phases that may be preterrestrial in origin. Like the shergottites, the nakhlites have relatively young crystallization ages ($\sim 1.3 \times 10^9$ years) and appear to have formed in oxidizing, volatile-rich, Earth-like environments compared to the basaltic achondrites. Unlike the shergottites, the nakhlites show only minor effects of shock metamorphism.

The Chassigny meteorite (the prototype of the two chassignites, as of early 2006) is a lightly to moderately shocked dunite. Like the shergottites and nakhlites, it is volatile-rich and oxidized and also contains hydrous amphibole. Its age is not well constrained, but a single potassium–argon age is consistent with the young ages of the shergottites and nakhlites. *See* DUNITE.

The young ages of the shergottite, nakhlite, and chassignite meteorites (often referred to as the SNC group), plus the similarity of the bulk compositions of the shergottites to that of the Martian soil as determined by the Viking landers, first led to the suggestion that these meteorites could be derived from Mars. It is difficult to conceive of a heat source for endogenous igneous activity (these meteorites have no

features resembling known impact melts) on an asteroidal parent body at about 1.3×10^9 years ago; and given the limited choice of available larger planets, Mars seemed the most likely choice. The similarity of relative noble gas and nitrogen abundances and isotopic ratios in the Martian atmosphere and shock-produced glass in one shergottite provide strong support for this hypothesis. The very low paleomagnetic intensities of the shergottites are also consistent with a Martian origin. It is still a subject of controversy whether fragments of sufficient size to explain measured cosmic-ray exposure ages could be ejected more or less intact from Mars by impact and subsequently delivered to Earth. It is generally accepted that the question of whether or not these meteorites are from Mars will be resolved only after a sample-return mission to Mars. *See* PALEOMAGNETISM.

Angrites. Until the discovery of two additional examples of this class of meteorite in Antarctica, Angra dos Reis was the only meteorite of this type. Nevertheless, it has been of great interest because of its unusual mineralogy and very primitive isotopic characteristics. Angra dos Reis is composed nearly entirely of an iron-bearing, aluminous clinopyroxene (fassaite), with minor amounts of plagioclase, spinel, calcic olivine, and a variety of other phases. The more recently discovered members of this group are less enriched in fassaite, but their mineralogies are similar. Angra dos Reis is generally interpreted as having formed by accumulation of fassaite from a relatively silica-poor magma, while the Antarctic examples have been interpreted as more closely resembling magmatic compositions. With an age of 4.55×10^9 years and evidence for extinct short-lived radionuclides, Angra dos Reis provides another case of magmatic activity from the very earliest history of the solar system. However, based on the inferred characteristics of the parent liquids of the angrites, this igneous activity was distinct from that recorded by the basaltic achondrites. Due to their rarity, it has been difficult to constrain the details of their petrogenesis, but efforts to do so have generally suggested complex igneous histories. *See* MAGMA.

Ureilites. The dominant silicates of this strange meteorite type are olivine and low-calcium pyroxene (typically pigeonite), usually present as millimeter-sized mosaic aggregates. The silicates are enclosed in a dark matrix that consists mainly of metallic iron, graphite, diamond, lonsdaleite (another high-pressure form of carbon), plus other minor phases. Where the matrix and silicates are in contact, they have reacted at low pressures to form iron-rich metal and more magnesian silicates. Some of the ureilites also exhibit long, oriented cracks or voids. In terms of their oxygen isotopic compositions, the ureilites are distinct from other achondrites, but are similar to certain carbonaceous chondrites.

The silicates often are preferentially oriented. This, along with their relatively coarse grain size, suggests that they are plutonic rocks. However, it is generally difficult to distinguish between accumulations of crystals settled from a magma chamber and residual crystals left behind after the extraction of partial melts from planetary interiors, and both origins have been suggested for the ureilites. In either case, aspects of their trace-element and isotopic compositions suggest a complex igneous history for the ureilites if these features are all igneous in origin. The carbon in the ureilites could be a primary igneous feature, or it could have been introduced subsequent to their igneous history, perhaps during the shock events that led to the mosaicism of the silicates and the presence of the high-pressure phases diamond and lonsdaleite in the matrix. Aspects of the textures of these meteorites tend to favor the former possibility. If so, the coexistence of graphite plus iron-bearing silicates during igneous activity suggests somewhat elevated pressures (that is, more than a few hundred bars). Although no direct connection has been made, many researchers in this field believe that the parent bodies of the ureilites were related to those of carbonaceous chondrites.

Aubrites. These meteorites, also known as enstatite achondrites, are composed almost entirely of polymorphs of enstatite ($MgSiO_3$), with grain sizes sometimes exceeding several centimeters. Most are monomict breccias, but unbrecciated and polymict examples are known. At least some of these meteorites appear to be regolith breccias, based on their textures and the presence of foreign meteorite clasts and solar rare gases. Like the enstatite chondrites, they contain silica polymorphs, metallic iron with dissolved silicon, and an assortment of rare, highly reduced minerals, some of which are known only from these meteorites. The assemblages found in these meteorites indicate that they formed under conditions more reducing than expected in a gas of solar composition.

It is generally accepted that the aubrites are related to the enstatite chondrites, but the relationship is not clear. They are chemically similar to enstatite chondrites, but they are distinctive; in most cases they extend to the more extreme fractionations the same trends observed among the enstatite chondrites. The very coarse grain size of the aubrites suggests that they formed as cumulates, perhaps during differentiation of an enstatite chondrite parent body. Alternatively, it has been suggested that they originated by fractionation processes occurring in the solar nebula and that they represent a continuation of the same processes by which the various nonigneous enstatite chondrites are related. According to this view, the coarse grain size of the aubrites would have resulted from annealing in high-temperature nebular environments. *See* NEBULA.

Edward Stolper

Iron and Stony-Iron Meteorites

Iron and stony-iron meteorites are made largely or partly of metallic iron-nickel and come from deep inside asteroids that were melted around 4560 million years ago. In most of these bodies, the metal formed dense cores that were surrounded by silicate mantles. Iron and stony-iron meteorites are much stronger than other meteorites, so they can survive longer in space. At least 12 of the 20 craters on Earth under 1.2 km (0.75 mi) in diameter were made by iron or stony-iron meteoroids.

Iron meteorites. Iron meteorites are pieces of once molten metallic cores and pools in asteroids that were fragmented by impacts after they cooled slowly. About 700 different iron meteorites have been identified; 40 were seen to fall, and the rest fell during the last million years. The smallest iron meteorites, which weigh only 5–30 g (0.18–1.1 oz), were found in Antarctica and are aerodynamically shaped like certain tektites. The largest single iron meteorite weighs about 60 metric tons (66 tons) and still lies in Namibia. The second largest is from Greenland and weighs 31 metric tons (34 tons). It is on display in the American Museum of Natural History in New York. *See* TEKTITE.

Nine much larger iron masses hit the Earth during the last million years, forming numerous craters 50 to 1200 m (150 to 3900 ft) in diameter. However, the surviving fragments from these meteoroids weigh less than a ton. The largest crater, which is in Arizona, was formed about 50,000 years ago by the impact of a meteoroid measuring about 50 m (150 ft) across. The impact released energy equivalent to about 20–60 megatons of TNT.

Mineralogy. When iron meteorites are sawed, ground, polished, and etched, they typically show a striking geometrical array of oriented crystals known as a Widmanstätten pattern. This pattern results from the crystallographically controlled formation of kamacite plates parallel to the four sets of octahedral planes in the precursor taenite crystals (**Fig. 5**). Kamacite and taenite are the mineral names for the body-centered cubic and face-centered cubic crystal structures of iron alloys. Iron meteorites that show such an oriented array of kamacite plates in taenite are called octahedrites and contain 6–15% nickel. A few iron meteorites known as hexahedrites contain 5–6% nickel and are almost entirely composed of kamacite. Even rarer are the ataxites, which have more than 15% nickel and contain only microscopic kamacite plates. *See* CRYSTAL STRUCTURE.

The large sizes of the precursor taenite crystals and the oriented kamacite plates indicate that iron meteorites were hot when they formed and that they cooled at an extremely slow rate. Thus at one time iron meteorites must have been buried deep inside a large volume of poorly conducting rocky material. The cooling rate in the temperature range 700–400°C (1290–750°F) can be estimated from the thickness of the kamacite plates, the bulk nickel concentration, knowledge of the stability fields of kamacite and taenite, and the rate at which iron and nickel atoms diffuse in these minerals. From these data, computer models have been developed for the diffusion-controlled growth of kamacite plates; these models yield typical cooling rates of 10–100°C (18–180°F) per million years. Therefore, most iron meteorites must have been buried under tens of kilometers of rocky material when they cooled from 700 to 400°C (1290 to 750°F).

Most iron meteorites contain small amounts of a wide variety of other minerals; sulfides, phosphides, carbides, oxides, phosphates, and silicates together make up no more than a few percent by volume of most iron meteorites (**Table 2**).

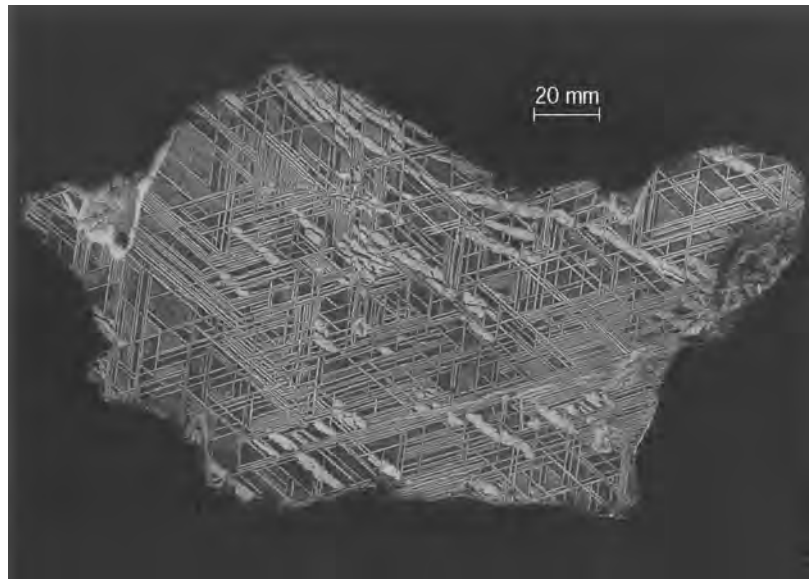

Fig. 5. A piece of the Edmonton (Kentucky) iron meteorite that has been sliced, polished, and etched to show the Widmanstätten pattern. The plates of kamacite 0.3 mm (0. 01 in.) wide are oriented parallel to the faces of an octahedron. (*Smithsonian Institution*)

Although the mineralogy of iron meteorites largely reflects slow cooling at depth, some features result from shock waves of high intensity caused by impacts between asteroids at speeds of several kilometers (1 km = 0.6 mi) per second. Kamacite in many iron meteorites has a distorted structure caused by transient formation of a hexagonal close-packed structure at shock pressures above 13 gigapascals (130 kilobars). Two iron meteorites contain diamonds formed from graphite by shock. Those in the Canyon Diablo iron formed during the impact responsible for the 1200–m (3900-ft) crater in Arizona, whereas diamonds in a closely related octahedrite from Antarctica probably formed during an impact on the parent asteroid. Another iron meteorite contains stishovite, which formed from silica during a collision between asteroids. *See* DIAMOND; SHOCK WAVE; STISHOVITE.

Chemical composition. The chemical compositions of iron meteorites provide important clues to their classification and the manner in which they solidified in molten asteroids. The trace elements gallium and germanium, which are present at concentrations between 0.01 and 1000 parts per million by weight, are

TABLE 2. Common minerals in iron meteorites

Mineral	Composition
Kamacite	Fe-Ni; Ni <7.5%
Taenite	Fe-Ni; Ni >30%
Tetrataenite	FeNi (ordered)
Troilite	FeS
Daubreelite	$FeCr_2S_4$
Schreibersite	$(Fe,Ni)_3P$
Cohenite	Fe_3C
Haxonite	$Fe_{23}C_6$
Graphite	C
Carlsbergite	CrN
Chromite	$FeCr_2O_4$
Whitlockite	$Ca_3(PO_4)_2$
Chlorapatite	$Ca_5(PO_4)_3Cl$

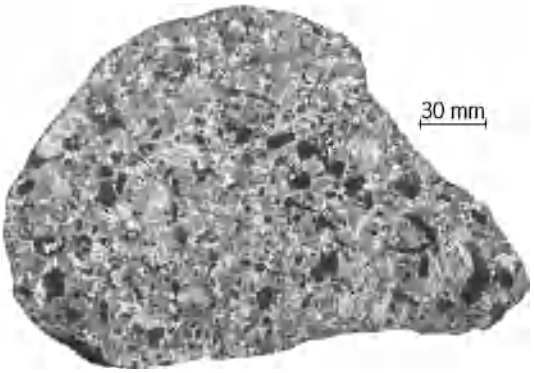

Fig. 6. A slice of the Four Corners (southwestern United States) iron meteorite showing angular black silicate inclusions separating regions with independently oriented Widmanstätten patterns. Like other group IAB members, this meteorite did not form in the core of an asteroid. A large impact probably mixed silicate with molten metallic iron-nickel long before the Widmanstätten pattern formed. (*Smithsonian Institution*)

especially useful in classifying iron meteorites. Most iron meteorites (84%) belong to one of 13 chemical groups, all but two of which have concentrations of gallium and germanium that vary by less than a factor of 2. Each of the 13 groups, which are identified as IAB, IC, IIAB, and so forth, has between 5 and 210 members. This chemical classification of iron meteorites has not made the older, structural classification obsolete, because most iron meteorites (perhaps 70%) can be classified into chemical groups on the basis of structure alone. For example, nearly all hexahedrites belong to group IIA, the medium octahedrites (those having kamacite bandwidths of 0.5 to 1.3 mm or 0.02 to 0.068 in.) largely belong to group IIIAB, and ataxites to group IVB.

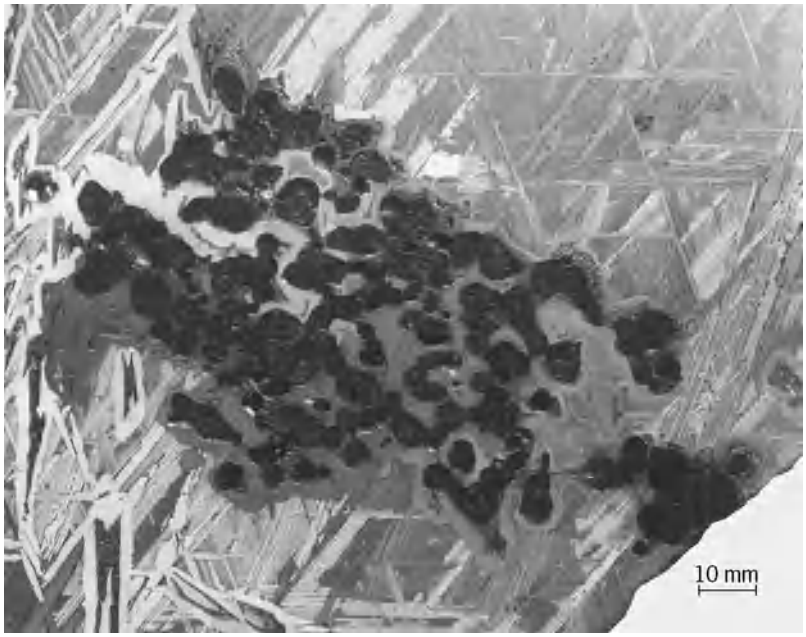

Fig. 7. A part of a slice of the Brenham (Texas) pallasite, showing a cluster of rounded, dark crystals of olivine enclosed in iron-nickel metal that shows a Widmanstätten pattern. Pallasites formed at the core–mantle boundary of asteroids that melted. (*Smithsonian Institution*)

Analyses for other trace elements, such as iridium, gold, and tungsten, and the minor elements phosphorus and nickel can also be used to classify iron meteorites, but they show larger variations within groups. For example, iridium concentrations vary by a factor of 6000 in group IIAB alone. The iridium variations within groups, which are very much larger than the variation in chondritic metal, were produced when molten metal solidified. Thus iridium, tungsten, and other elements that preferentially concentrate in solid metal were enriched in the first metal to solidify, whereas elements that are concentrate in molten metal, such as nickel, phosphorus, and gold, were enriched in the last metal that solidified.

Laboratory measurements of the concentrations of these elements in coexisting solid and liquid metallic iron-nickel allow the chemical trends found within the groups to be accurately modeled. Thus, nearly all of the 13 groups of iron meteorites formed from separate asteroids; the approximately 100 ungrouped iron meteorites probably come from another 40-odd bodies. Cooling rates for samples from a single metallic pool should be uniform, and this is observed for most groups.

Origins. The chemical and mineralogical evidence discussed above shows that iron meteorites formed from molten pools of metal that solidified and then cooled over many millions of years. This evidence is consistent with an origin for iron meteorites in the cores of asteroids that melted and differentiated. When an asteroid is partly melted, iron-nickel and iron sulfide, being denser than the associated silicates, will begin to sink to the center. With sufficient heating, a core of molten sulfur-rich metal will form. Since most iron meteorites contain no silicates and most achondrites have only trivial amounts of metal, it is likely that metallic cores are the source of many iron meteorites.

Two groups of iron meteorites (IAB and IIE) contain several volume percent of silicate inclusions and probably did not form in cores of asteroids. Instead these meteorites are probably derived from many metallic pools distributed within asteroids. Group IAB iron meteorites have angular silicate inclusions (**Fig. 6**), and it is likely that impacts mixed their metal and silicate components while they were hot. In addition, if the planetesimals from which the asteroids formed by accretion had already melted and solidified, it is possible that several metallic cores could have been combined into one asteroid.

At least one group of irons, IVA, formed in an asteroid that probably suffered a catastrophic impact before it had cooled below ~900°C (1600°F). Evidence for the timing and nature of more recent impacts comes from the cosmic-ray exposure ages of the meteorites. Since galactic cosmic rays can penetrate only to depths of about a meter, meteoroids or asteroidal fragments are not significantly exposed to cosmic rays until they are a meter or less in size. Exposure ages of most iron meteorites are $0.1–1 \times 10^9$ years, but two groups show a significant clustering of ages: group IIIAB and IVA iron meteorites

have ages of 6.5×10^8 and 4.5×10^8 years, respectively. Thus about 40% of all irons are fragments of two giant impacts in the asteroid belt that occurred in the last billion years. *See* COSMIC RAYS.

Astronomers believe they have identified many metallic asteroids 1 to 200 km (0.6 to 120 mi) in diameter, but none has been linked yet to a specific group of iron meteorites.

Stony-iron meteorites. There are two major types of stony-iron meteorites: pallasites and mesosiderites. They are among the rarest groups of meteorites, together constituting 1% of meteorite falls. About 50 specimens of each type have been found.

Pallasites consist of 5–80% by volume of olivine [$(Mg,Fe)_2SiO_4$] embedded in metallic iron-nickel, with minor amounts of chromite ($FeCr_2O_4$), troilite (FeS), and pyroxene [$(Mg,Fe)SiO_3$] (**Fig. 7**). Three groups are distinguished: the main group, the Eagle Station types, and the pyroxene-rich pallasites. All three groups formed at the core-mantle boundaries of differentiated asteroids when molten metal was intruded into the surrounding olivine-rich mantle. Most pallasites, which belong to the main group, have metal compositions that match those of the nickel-rich members of the IIIAB group of iron meteorites, implying that they formed in the same asteroid. Molten metal and silicates were largely mixed at the core–mantle boundaries of asteroids possibly as a result of impacts that gouged holes in the crust about 4500 million years ago. Impacts between large asteroids were common in the asteroid belt then as there were vastly more bodies than exist today.

Mesosiderites are enigmatic breccias consisting of fragments of basalt and gabbro mixed with roughly equal amounts of iron-nickel and troilite (FeS). About 50 specimens are known including the Eltanin samples, which came from a 1–5 km (0.6–3 mi) asteroid that was one of the largest to hit the Earth in the last 2 million years. The Eltanin samples were identified in cores of sediment drilled from the southeast Pacific Ocean.

Mesosiderites, like pallasites, formed by mixing of molten metallic iron and nickel with solid silicate. But the impact that created the mesosiderites caused vast quantities of molten metal to spew over fragments of rock from the surface and interior. The silicate-metal mixtures were deeply buried after the impact and cooled at a rate of about 0.4 K per million years, slower than any other group of meteorites.

Olivine-rich meteorites from the mantles of asteroids that melted are much rarer than stony-iron meteorites. Similarly, olivine-rich asteroids are rarer than metal-rich ones. The impacts that excavated the deeply buried iron and stony-iron meteorites and brought fragments to Earth probably crushed most of the olivine-rich mantles to fine dust.

Edward R. D. Scott

Isotopic Anomalies in Meteorites

Primitive meteorites show a large variety of isotopic anomalies, that is, deviations from the average solar system composition (called "normal" composition) that cannot be explained by chemical fractionation and radioactive decay taking place today. They are rather of presolar origin or the result of irradiation from an early active Sun or in interplanetary space. These anomalies provide information about the nucleosynthetic sources of the material that formed the solar system. *See* ISOTOPE.

It is well established that carbon and all the heavier elements are produced in stars by nuclear processes (stellar nucleosynthesis). These elements are ejected into the interstellar medium either from explosions of massive stars (supernovae) or as stellar winds and planetary nebulae from low-to-intermediate mass stars at the end of their evolution. Stars of different masses and ages produce elements with very different isotopic ratios. The solar system was formed from material originating from many stars. However, until the early 1970s it was believed that all this material was thoroughly homogenized into a mix of uniform isotopic composition in a hot solar nebula. Isotopic measurements of samples from the Earth, the Moon, and various meteorites seemed to confirm this belief. Exceptions were well-understood excesses of certain isotopes that were the daughters of long-lived (they still exist today) radioactive isotopes or isotopic fractionation effects as the result of physical and chemical processes. *See* NUCLEOSYNTHESIS.

This belief in isotopic homogeneity was shattered by the discovery of isotopic anomalies in several elements in 1973; and today, thanks to increases in the precision of isotopic analysis and the capability to measure small samples, a plethora of isotopic anomalies are known in meteorites. The largest anomalies are found in small samples, where the effects are not diluted by isotopically normal material.

Isotopic anomalies in meteorites and interplanetary dust particles (IDPs) can be divided into four classes.

(1) *Mass-dependent fractionation caused by physicochemical processes (diffusion, evaporation, condensation, and chemical reactions).* Certain physical and chemical processes can also lead to non-mass-dependent fractionation that might mimic effects of nuclear origin.

(2) *Effects from the decay of radioisotopes.* In addition to effects from the decay of long-lived isotopes, meteorites show also the effects of short-lived, now extinct isotopes.

(3) *Nuclear effects reflecting nucleosynthetic processes in stellar sources.* These effects are found in samples that formed in the solar system but inherited some nucleosynthetic components, as well as in presolar grains, bona fide stardust, that formed in stellar atmospheres and thus represent the isotopic compositions of their parent stars.

(4) *Effects due to irradiation of material by an early active Sun or irradiation of meteorites by galactic and solar cosmic rays.* These effects provide information on the exposure history of meteorites, both on their parent bodies and in interplanetary space during their travel between their asteroidal sources and the Earth.

TABLE 3. Short-lived radioisotopes†

Radioisotope	Half-life, million years	Daughter isotope	Reference isotope	Initial abundance
^{41}Ca	0.10	^{41}K	^{40}Ca	1.5×10^{-8}
^{26}Al	0.74	^{26}Mg	^{27}Al	5×10^{-5}
^{10}Be	1.5	^{10}B	^{9}Be	$\sim 5 \times 10^{-4}$
^{60}Fe	1.5	^{60}Ni	^{56}Fe	$\sim 1.5 \times 10^{-6}$
^{53}Mn	3.7	^{53}Cr	^{55}Mn	$\sim 2 \times 10^{-5}$
^{107}Pd	6.5	^{107}Ag	^{187}Pd	4.5×10^{-5}
^{182}Hf	9	^{182}W	^{180}Hf	2×10^{-4}
^{129}I	16	^{129}Xe	^{127}I	1×10^{-4}
^{244}Pu	81	Fission Xe	^{238}U	$(4–7) \times 10^{-3}$
^{146}Sm	103	^{142}Nd	^{144}Sm	8×10^{-3}

†Short-lived radioisotopes for which evidence has been found in meteorites.

Isotopic fractionation effects. Refractory inclusions (CAIs) from primitive meteorites (mostly carbonaceous and unequilibrated ordinary chondrites) exhibit mass-dependent fractionation effects in the elements oxygen, magnesium, silicon, calcium, and titanium. These effects must be due to evaporation and condensation during the formation of the CAIs and in some cases indicate multistage processing. Fractionation effects are largest in a class of CAIs named FUN inclusions because, in addition to fractionation (F), they show nuclear effects (called UN effects, because they were unknown at the time of discovery) in many elements. It is still not clear, however, why F and UN effects are associated with one another.

Interstellar cloud material. Carbonaceous and unequilibrated ordinary chondrites and interstellar dust particles have large excesses in deuterium (ranging up to 50-fold in certain interstellar dust particles) and ^{15}N. These anomalies appear to be associated with organic matter. Because deuterium is destroyed in stars (it was produced only in the big bang), deuterium excesses cannot be of nucleosynthetic origin. They and the ^{15}N excesses are most likely due to the incorporation of interstellar cloud material into meteorites and interstellar dust particles. This material acquired its isotopic anomalies through fractionation during ion-molecule exchange reactions that took place at very low temperatures (10 to 30 K or −442 to −405°F) in dense molecular clouds. Extremely large deuterium excesses (up to 1 million-fold) relative to the interplanetary medium are observed astronomically in simple molecules from interstellar clouds. *See* DEUTERIUM; INTERSTELLAR MATTER; MOLECULAR CLOUD.

Oxygen. The discovery of large ^{16}O excesses in refractory minerals from the Allende carbonaceous chondrite in 1973 refuted the hot solar nebula model. Large ^{16}O variations dominate the oxygen isotopic effects in meteorites with only minor variations in the $^{17}O/^{18}O$ ratio. CAIs were formed from a reservoir enriched in ^{16}O by about 5% relative to the average isotopic composition of other solar system materials. This enrichment might have a nucleosynthetic origin (supernovae produce huge amounts of ^{16}O); however, among presolar grains (discussed below) almost no oxygen-rich grains with a supernova origin have been detected. Alternatively, the ^{16}O excesses could be the result of non-mass-dependent physicochemical processes or of self-shielding, a photochemical effect in the early solar system that separated ^{16}O from the other two isotopes because of the large difference in the abundance of ^{16}O from those of ^{17}O and ^{18}O. In the case of self-shielding, the Sun is predicted to be enriched in ^{16}O relative to the inner planets and asteroids. Analysis of solar wind samples collected by the *Genesis* mission may settle this issue, despite contamination caused by its crash landing in September 2004. *See* SUPERNOVA.

Short-lived isotopes in the early solar system. Meteorites contain evidence of the presence of short-lived isotopes. These isotopes have half-lives so short that today they are extinct. The initial presence of a short-lived isotope is indicated by excesses in the daughter isotope that are proportional to the parent-to-daughter elemental ratios. There is now solid evidence for a series of short-lived isotopes (**Table 3**). Aluminum-26 has been studied most thoroughly. Its initial presence has been established for many CAIs, most of which have an initial inferred $^{26}Al/^{27}Al$ ratio of 5×10^{-5}. Like other short-lived isotopes (^{53}Mn and ^{129}I are prominent examples), ^{26}Al can in principle serve as a fine-scale chronometer of relative ages of early solar system events. It (and ^{60}Fe) also has been considered as a possible heat source for melting small planetesimals. Both of these functions require its widespread and uniform distribution in the solar nebula. The fact that relative ages between CAIs and other meteoritic samples determined with the ^{26}Al clock agree with age differences obtained with the absolute uranium-lead chronometer indicates that this is indeed the case.

The abundances of the radioisotopes ^{53}Mn, ^{182}Hf, and those with longer half-lives can be explained by continuous equilibrium production by supernovae in the Galaxy. However, the presence of ^{26}Al and ^{60}Fe, and especially that of ^{41}Ca with its much shorter lifetime, in early solar system solids requires production in a stellar event immediately (less than a million years) before solar system formation. The possible candidates are a supernova, whose explosion could have triggered the collapse of the molecular cloud from which the solar system formed, or an asymptotic giant branch (AGB) star, a low-to-intermediate mass (with a mass of less than 8 solar masses) star in the late stages of its evolution. Beryllium-10 is the only short-lived isotope that is not produced by stellar nucleosynthesis but by irradiation of heavier

elements (carbon and oxygen) by energetic nuclear particles, most likely from an early active Sun. *See* RADIOISOTOPE (GEOCHEMISTRY); STELLAR EVOLUTION.

Nuclear anomalies. Many CAIs were found to carry nuclear anomalies. These anomalies are especially large in FUN inclusions. These inclusions show isotopic patterns in the heavy elements (elements heavier than iron) that carry the signatures of distinct nuclear processes. The heavy elements are made either by neutron capture starting from iron, either in the *s* process (slow neutron addition in stellar environments with low neutron density) or in the *r* process (rapid neutron addition in a high-neutron-density environment), or by the *p* process that is responsible for the production of proton-rich isotopes. Whereas the solar (normal) isotopic composition represents a mixture of these different components, FUN inclusions show isotopic signatures of individual components. **Figure 8** shows the isotopic ratios of barium, neodymium, and samarium measured in one FUN inclusion, EK 1-4-1. The patterns show excesses in the isotopes that are produced by the *r* process, and for samarium also an excess in a *p*-process isotope. These patterns are compared with the isotopic patterns measured in presolar silicon carbide (SiC) grains (discussed below). The latter are the inverse of the *r*-process patterns of the FUN inclusions; they indicate production by the *s*-process. These complementary patterns are clear evidence that different stellar sources contributed material to the solar system.

Refractory inclusions have also anomalies in the neutron-rich isotopes of elements in the vicinity of iron in the chart of nuclides (iron-peak elements). These isotopes include ^{48}Ca, ^{50}Ti, ^{54}Cr, ^{58}Fe, ^{64}Ni, and ^{66}Zn. Correlated excesses and depletions are observed. Effects are pronounced for FUN inclusions but, for ^{48}Ca and ^{50}Ti, are largest in inclusions containing the refractory mineral hibonite ($CaO[Al_2O_3]_6$). The elements of the iron peak are all produced in supernovae, but the correlated excesses and depletions of the neutron-rich isotopes indicate contributions from different types of these stellar explosions. Different supernova types are distinguished by astronomical observations.

Improvements in analytical capabilities have led to the discovery of numerous isotopic anomalies in different meteorites and in their chemical separates (components obtained through chemical treatment with acids or other solvents). Although nucleosynthetic components for these anomalies have not been identified in all cases, it is clear that different stellar sources produced the elements and that these components were not completely mixed in the solar system.

Presolar dust grains. Whereas all the solids with isotopic anomalies mentioned so far formed in the solar system and contain only surviving traces of presolar isotopic signatures, primitive meteorites contain also true stellar grains. These grains formed in stellar atmospheres, were ejected into the interstellar medium, survived the formation of the solar system, and were incorporated into meteorites. They

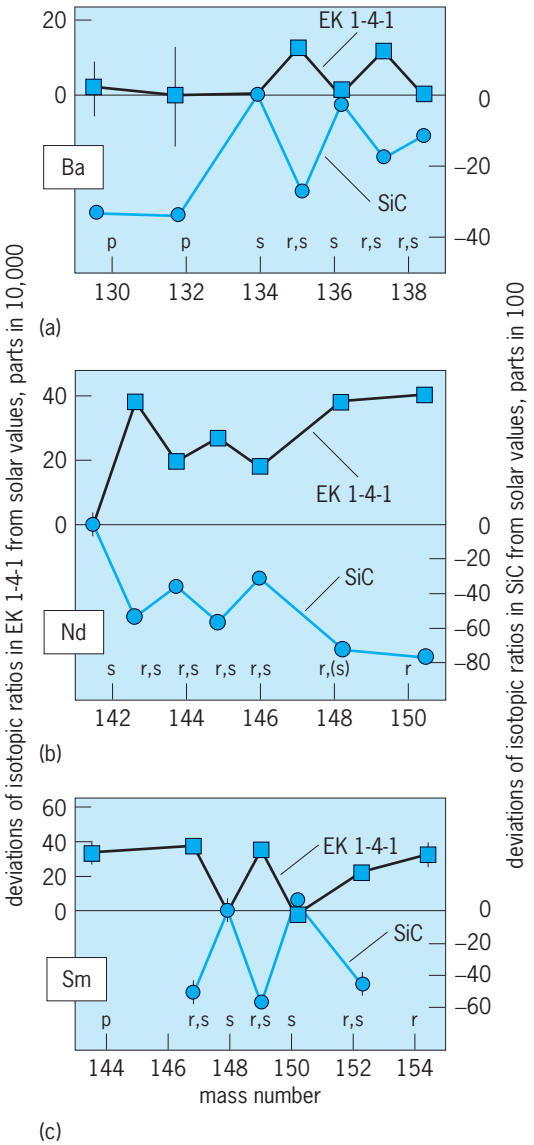

Fig. 8. Isotopic anomalies in (*a*) barium (Ba), (*b*) neodymium (Nd), and (*c*) samarium (Sm) measured in the FUN inclusion EK 1-4-1 and in presolar silicon carbide (SiC). Production of isotopes in the nuclear *s*, *r*, or *p* processes is indicated by letters at the bottoms of the graphs. Whereas EK 1-4-1 exhibits *r*-process patterns, presolar SiC exhibits complementary *s*-process patterns. (*Adapted from F. Begemann, Isotope abundance anomalies and the early solar system: MuSiC vs. FUN, in N. Prantzos, E. Vangioni-Flam, and M. Cassé, eds., Origin and Evolution of the Elements, pp. 517–526, Cambridge University Press, 1993*)

can be located in and extracted from their parent meteorites and studied in detail in the laboratory. Their stellar origin is indicated by their isotopic compositions, which are completely different from those of the solar system and, for some elements, cover extremely wide ranges. For example, in Fig. 8 the anomalies carried by presolar silicon carbide are more than a factor of 100 larger than those in the FUN inclusion EK 1-4-1.

The first presolar grains were discovered because they carry anomalous noble gas components in neon (Ne-E, almost pure ^{22}Ne) and xenon (Xe-S, exhibiting an *s*-process pattern in the form of excesses in ^{128}Xe,

TABLE 4. Presolar grain types

Grain type	Abundance†, parts per million	Size, μm	Stellar sources‡
Silicates in IDPs*	~400	≤1	RG, AGB
Silicates in meteorites	180	≤0.5	RG, AGB
Spinel ($MgAl_2O_4$)	1.5	0.15–2	RG, AGB, SNe
Corundum (Al_2O_3)	0.15	0.15–3	RG, AGB, SNe
Nanodiamonds	1500	0.002	SNe
Mainstream SiC	14	0.3–20	AGB
SiC type A + B	0.5	0.5–5	J stars
SiC type X	0.2	0.3–5	SNe
Graphite	1	1–20	SNe, AGB
Nova grains	0.001	~1	Novae
Silicon nitride (Si_3N_4)	0.002	≤1	SNe
Titanium carbide (TiC)	~0.001	0.01–0.5	SNe, AGB

*Silicates are pyroxene [$(Mg,Fe)_2Si_2O_6$] and olivine [$(Mg, Fe)_2SiO_4$]. IDPs = interplanetary dust particles.
†Abundances vary with meteorite type. Shown here are the maximum values.
‡RG = red giants; AGB = asymptotic giant branch stars; SNe = supernovae; J stars are carbon stars with very low $^{12}C/^{13}C$ ratios.

^{130}Xe, and ^{132}Xe; and Xe-HL, exhibiting excesses in the heavy xenon isotopes ^{132}Xe, ^{134}Xe, and ^{136}Xe, and the light Xe isotopes ^{124}Xe, ^{126}Xe, and ^{128}Xe). These grains turned out to be of carbonaceous nature: diamond, graphite, and silicon carbide. With the exception of diamonds, which are too small, single grains can be analyzed for their isotopic compositions by secondary ion mass spectrometry (SIMS) in the ion microprobe or by resonance ionization mass spectrometry (RIMS), and were found to be anomalous in all their isotopic ratios. This, in addition to the size of the anomalies, identifies them as samples of stellar material. Subsequently, other presolar grain types, not tagged by exotic noble gases, were identified by isotopic analysis of individual grains in the ion microprobe. *See* INERT GASES; RESONANCE IONIZATION SPECTROSCOPY; SECONDARY ION MASS SPECTROMETRY (SIMS).

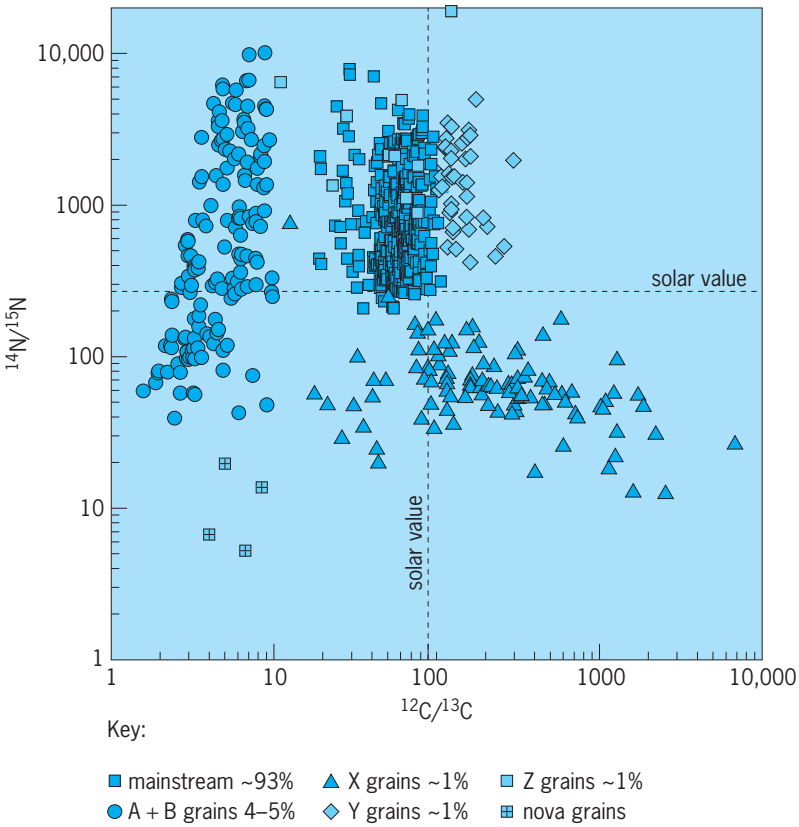

Fig. 9. Nitrogen and carbon isotopic ratios measured in individual presolar SiC grains. The enormous ranges spanned by these ratios are evidence for a stellar, presolar origin of these grains. The nitrogen and carbon isotopic ratios (shown here) and silicon isotopic ratios (shown in Fig. 10) are used to distinguish between different types of grains. The relative abundances of these types are indicated in the key.

Table 4 lists the types of presolar grains for which isotopic ratio measurements have been obtained. The laboratory study of these grains has developed into a new branch of astrophysics because it provides information on stellar evolution and nucleosynthesis, galactic chemical evolution, stellar atmospheres, conditions in interstellar space, the early solar system, and meteorite parent bodies.

Silicon carbide has been studied in greatest detail because it can relatively easily be isolated from meteorites, many grains are several micrometers in size, and it contains many trace elements. The isotopic compositions of the noble gases from helium to xenon, and of the elements carbon, nitrogen, magnesium, silicon, potassium, calcium, titanium, strontium, zirconium, molybdenum, ruthenium, rubidium, barium, neodymium, samarium, and dysprosium, have been measured in presolar silicon carbide. **Figure 9** shows the nitrogen and carbon and **Fig. 10** the silicon isotopic ratios of individual silicon carbide grains. These ratios have been used to classify the grains into different types whose relative abundances are indicated in Fig. 9. Different stellar sources have been identified for these grains. Mainstream grains are from carbon stars, asymptotic giant branch stars that became carbon-rich during their evolution. Y and Z grains are likely also from carbon stars but from stars with low metallicity, that is, depleted in the elements heavier than helium. Grains of type X are from supernovae. Proof of such an origin comes from evidence of initial ^{44}Ti in many X grains. Titanium-44, which has a half life of only 60 years and decays into ^{44}Ca, is produced only in supernovae in an interior layer that consists of almost pure ^{28}Si, in agreement with the ^{28}Si excesses (depletions in ^{29}Si and ^{30}Si) found in X grains (Fig. 10). The low $^{12}C/^{13}C$ ratios defining A + B grains are observed in J-type carbon stars, but it is presently not

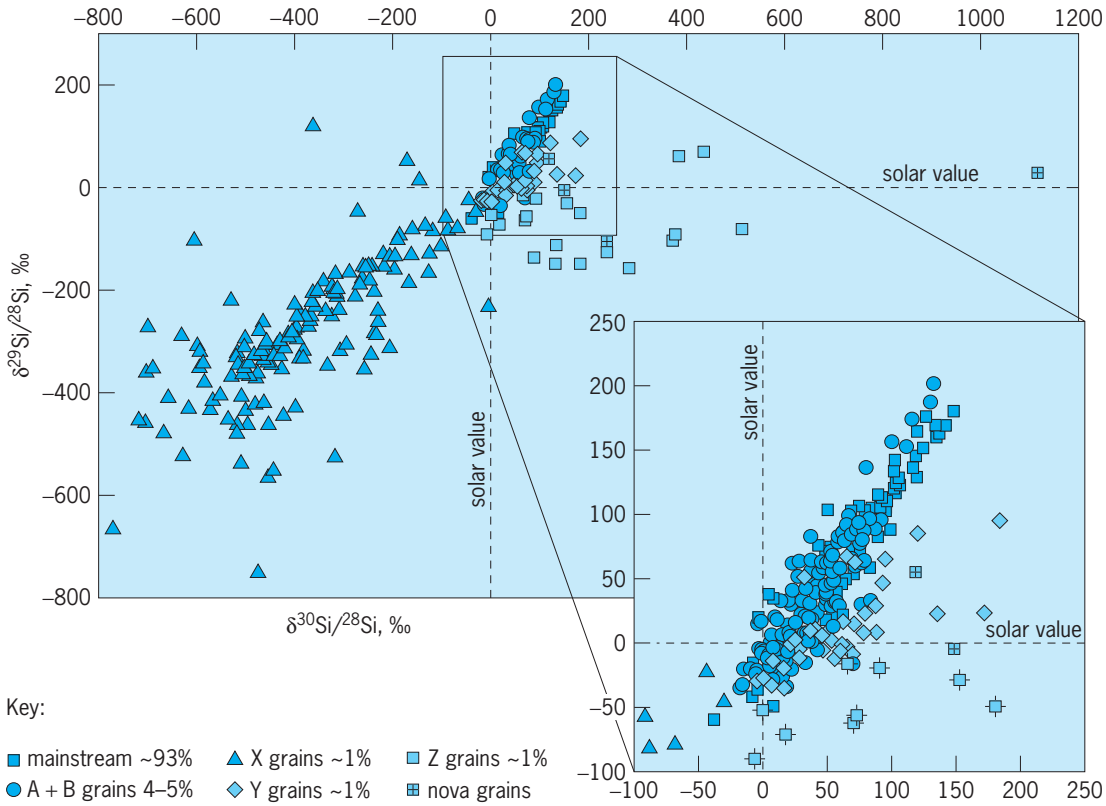

Fig. 10. Silicon isotopic ratios measured in individual presolar SiC grains. The ratios are expressed as delta values, that is, deviations (δ) from the solar values in parts per thousand (‰).

understood how these stars acquired their isotopic compositions. Finally, a handful of silicon carbide grains come from novae, white dwarf stars onto which a companion star dumps material that periodically ignites in a nuclear explosion. *See* CARBON STAR; CATACLYSMIC VARIABLE; NOVA.

Low-density graphite grains also have large ^{44}Ca excesses from ^{44}Ti decay, indicating a supernova origin. Many of these grains have large ^{18}O excesses, which is also a supernova signature originating from a layer where ^{14}N, the product of hydrogen burning (capture of protons) in the carbon-nitrogen-oxygen (CNO) cycle, is turned into ^{18}O by helium burning (capture of helium nuclei). *See* CARBON-NITROGEN-OXYGEN CYCLES.

Whereas all silicon carbide grains found in primitive meteorites are of presolar origin, this is the case for only a small fraction of oxygen-rich grains. The reason is that in the solar system the abundance of oxygen is higher than that of carbon, leading to the formation of oxygen-rich minerals such as silicates, but not to the formation of carbonaceous phases. As a consequence, oxygen-rich presolar grains have to be identified from isotopic measurements in the ion microprobe. The presolar minerals discovered in this way include corundum, spinel, hibonite, and the silicates olivine and pyroxene (Table 4). Most oxygen-rich presolar grains have excesses in ^{17}O and depletions in ^{18}O. This signature has been explained by hydrogen burning in deep, hot layers during the main-sequence phase of the parent star. After exhaustion of hydrogen in the interior of the star, a special mixing process, called the first dredge-up, brought material from these layers to the stellar surface, where grains formed during the subsequent red giant phase when the expanding star lost material in the form of a stellar wind. Some oxygen-rich presolar grains have ^{18}O depletions that are much larger than those that can be explained by the first dredge-up. They indicate additional circulation of envelope material to hot regions close to the thin layer where hydrogen burns during the red giant or asymptotic giant branch phase of evolution. This is one example of isotopic measurements of presolar grains in the laboratory leading to new insights into stellar processes.

Cosmogenic nuclides. During travel from their parent bodies to Earth, meteorites are bombarded by galactic and solar cosmic rays. In some cases, when the meteorites were not buried deeply before their ejection, bombardment also took place on the parent asteroids. This bombardment produces various isotopes, called cosmogenic nuclides, through nuclear reactions. Cosmogenic nuclides include both radionuclides and stable nuclides. Their measurement provides information on exposure ages of meteorites in interplanetary space, terrestrial ages (time between Earth entry and collection), burial history on parent asteroids, and preterrestrial meteorite sizes (most stony meteorites break up upon entry into the terrestrial atmosphere). Whereas the concentrations of radionuclides used to be determined by detecting their radioactive decay, nowadays isotopes such as ^{10}Be, ^{26}Al, ^{36}Cl, and ^{53}Mn

are measured by accelerator mass spectrometry. Stable radionuclides are detected as isotopic anomalies. For example, the three neon isotopes ^{20}Ne, ^{21}Ne, and ^{22}Ne are produced by cosmic rays in approximately equal amounts. However, because in the solar system the abundance ratio of $^{21}Ne/^{20}Ne$ is only 2.4×10^{-3}, any cosmogenic neon shows up as an excess in ^{21}Ne relative to the other two neon isotopes. Similarly, low-abundance isotopes of other noble gases (for example, ^{3}He and ^{126}Xe) are identified as cosmogenic nuclides from isotopic anomalies. *See* COSMOGENIC NUCLIDE.

Ernst Zinner

Meteorite Impact

The process of impact cratering was of fundamental importance for the accumulation of planets in the early solar system, the formation of planetary landscapes, and the Archean geology of the Earth. In addition, meteorite impacts are implicated in the Moon's origin and the extinction of the dinosaurs. However, recognition of the importance of impacts has been achieved only since the 1970s. The dominant role of impact cratering in sculpting the Moon's surface was not widely understood until after the Apollo landings in 1969. The spectacular plumes raised above Jupiter's cloud decks by the impact of fragments of Comet Shoemaker-Levy 9 in July 1994 have led to new investigations of the interaction between meteorites and atmospheres. In July 2005, the National Aeronautics and Space Administration's *Deep Impact* mission created an artificial crater about 200 m (650 ft) in diameter on Comet Tempel 1. *See* COMET.

Cratering mechanics. The precise outcome of a planetary collision depends on the size of the meteorite and conditions on the target planet. Small meteorites striking planets such as the Earth or Venus dissipate most of their energy in the atmosphere and do not strike the surface at high speed. In general, if the mass of the meteorite is small compared to the mass of atmospheric gases displaced during its entry, it will not create an impact crater. On airless bodies such as the Moon, there seems to be no lower limit on impact crater size: Craters as small as a few micrometers in diameter have been discovered on the lunar rocks. On Earth, the atmosphere prevents stony meteorites or comets from making craters smaller than a few kilometers in diameter, and even iron meteorites cannot make high-speed impact craters smaller than a few hundred meters in diameter.

Penetration of large meteorites into the Earth's atmosphere is limited by two effects. The less important is slowing by air drag. More important is crushing and dispersion of incoming objects by the large difference in pressure between the front of the fast-moving meteorite and its rear. Objects ranging in size from about 10 cm (4 in) to 100 m (300 ft) in diameter are broken into many fragments in the atmosphere. When they do finally fall to the surface, these fragments form a characteristic "strewn field" of small craters created by the impact of decelerated fragments that fall at their terminal velocity. Craters in strewn fields on the Earth spread over an elliptical region a few kilometers wide and up to 10 km (6 mi) long, with the largest crater at the downrange end. (On Venus, with its much denser atmosphere, the pattern is the same, but these dimensions are multiplied by a factor of 10.) Most meteorites recovered on Earth originally fell in clusters of this type.

When a large meteorite does penetrate a planet's atmosphere, it initiates a series of swift but orderly processes that eventually create a characteristic landform, an impact crater. Three principal stages are recognized in this process.

Stage 1: contact and compression. The meteorite first plunges into the surface rocks at high speed, compressing the underlying rocks and converting its initial kinetic energy into both heat and kinetic energy of the surface rocks. The high pressures produce a series of characteristic mineralogical changes in the surrounding rocks (**Table 5**) that often permit verification of the impact origin of a suspected crater. The

TABLE 5. Peterographic shock indicators

Material	Indicator	Pressure, GPa
Tonalite (igneous rock)	Shatter cones	2–6
Quartz	Planar elements and fractures	5–35
	Stishovite	15–40
	Coesite	30–50
	Melting	50–65(?)
Plagioclase	Planar elements	13–30
	Maskelynite	30–45
	Melting	45–65(?)
Olivine	Planar elements and fractures	5–45
	Ringwoodite	45
	Recrystallization	45(?)–65(?)
	Melting	<70
Clinopyroxene	Mechanical twinning	5–40(?)
	Majorite	13.5
	Planar elements	30(?)–45
	Melting	45(?)–65(?)
Graphite	Cubic diamond	13
	Hexagonal diamond	70–140

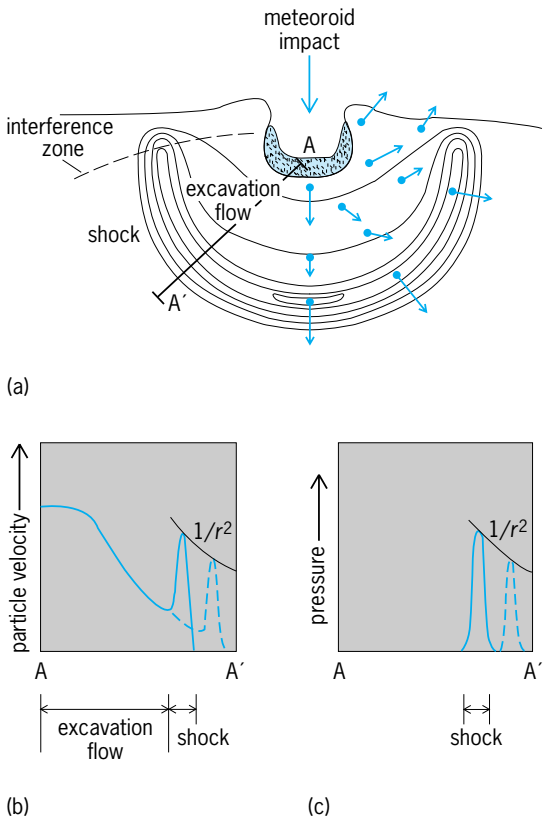

Fig. 11. Cratering by a meteorite. (*a*) Expanding shock wave and excavation flow following a meteorite impact; contours represent pressure at some particular time after impact, with the region of high shock pressure isolated or "detached" on an expanding hemispherical shell. (*b*) Profile of particle velocity and (*c*) profile of pressure along the section AA′. The broken lines show particle velocity and pressure some time later than those shown by the solid lines, and the solid curves connecting the peaks are portions of the "envelopes" of peak particle velocity and peak pressure.

duration of this compression stage is short, however, lasting only as long as it takes the meteorite to travel a distance equal to its own diameter.

Stage 2: excavation. Subsequently, the pent-up pressures in the compressed rocks create an explosion, blasting aside the surrounding rocks as a strong shock wave radiates away from the impact site (**Fig. 11**). The explosive nature of impacts was first recognized in 1924 by A. C. Gifford, who realized that the kinetic energy per unit mass of a meteorite striking at only 3 km/s (1.8 mi/s) is equivalent to the chemical energy of TNT. Because kinetic energy increases as the square of velocity, faster impacts resemble powerful explosions. This explains why most impact craters are circular, even though the most probable angle of impact is 45°. Only very oblique impacts, at angles of less than about 6° to the surface, produce elliptical craters.

The nearly hemispherical shock wave from the impact expands and weakens as time passes, leaving behind outward-moving rock debris that eventually excavate the crater. The time required for a crater of diameter D to open is given roughly by $(D/g)^{1/2}$, where g is the surface acceleration of gravity. Thus, the 1-km diameter (0.6-mi) Arizona Meteor Crater (also known as Barringer Crater) was excavated in about 10 s, whereas the 1000-km diameter (600-mi) Imbrium basin on the Moon took about 13 min to open.

Impact craters are not very efficient at excavating deeply buried material. Although most fresh craters have depth/diameter ratios of about 1/5, the maximum depth of excavation is only about 10% of the crater's diameter, so even the 180-km diameter (110-mi) Chicxulub crater in Yucatan did not excavate rocks from the Earth's mantle. Most debris ejected from an impact crater falls within about one crater diameter of the crater rim, although small quantities of material may be flung thousands of kilometers, or even entirely off the target planet, as demonstrated by the SNC meteorites that were ejected from the surface of Mars by large impacts.

The immediate result of crater excavation is called the transient crater. This is a relatively deep, steep-walled crater that begins to collapse as soon as it forms. The diameter of a transient crater is a function of the velocity of the impact, the angle between the approaching meteorite and the ground, the diameter of the meteorite, and the density of the projectile and of the target surface. Several websites use well-established mathematical relations among these variables to quickly compute transient crater sizes.

Stage 3: modification. Small transient craters are quickly filled by a lens of broken rock that forms from debris that slides down from the rim and pools at the bottom of the crater. Such bowl-shaped craters floored by broken rock are called simple craters.

In larger craters the floor rises as the rim sinks, producing craters with central mounds that are thinly veneered with broken and melted rock. The rims of such craters are scalloped and terraced with great blocks of slumped rock. Termed complex craters, they form at diameters greater than a certain transition size that depends on the gravity and composition of the target planet. The transition from simple to complex craters occurs at about 3 km (1.8 km) on the Earth and about 15 km (9 mi) on the Moon. Still larger craters exhibit circular mountainous rings instead of central peaks.

The very largest impact structures, particularly on the Moon, are surrounded by inward-facing, roughly circular (but often incomplete) mountain rings that probably formed well outside the crater cavity by a process of inward flow and slumping in the fluid asthenosphere beneath the crater. They are termed multi-ring basins, and the ratio between the diameters of the rings is often said to be close to a multiple of $\sqrt{2}$. Multi-ring basins on the Moon range in size from the 410-km-diameter (250-mi) basin Grimaldi to the 2600-km-diameter (1600-mi) South Polar basin. These enormous structures dominate the Moon's surface and form the principal stratigraphic markers on that body. A somewhat different variant of multi-ring basin, in which the center is surrounded by dozens of rings, is observed on the Jovian satellites Callisto and Europa, where the impact occurred in a relatively thin water ice layer overlying a liquid water ocean.

Impact cratering and planetary evolution. It is now believed that the planets formed out of the same cloud of gas and dust from which the Sun formed 4.5 billion years ago. Dust settled toward the midplane of the early solar nebula and accreted into kilometer-scale clumps by some still poorly understood process. The kilometer-size clumps were large enough to perturb each other gravitationally and collide. These collisions were at first gentle, but became more violent as the objects grew and gravitational perturbations became stronger. The last stages of planetary growth are thought to have been characterized by truly titanic collisions between the early protoplanets and objects up to half their diameter. Such collisions must have caused widespread melting and vaporization of the protoplanets and could have blasted material from the surface of the protoplanet, or the colliding object into orbit. One such collision is thought to have created the Moon, when a Mars-size protoplanet (most of which is now merged with the Earth) struck the protoearth and ejected a Moon-sized quantity of material into close orbit about it. Study of extinct radionuclide chronometers dates this event to about 30 million years after the first meteorites formed. The Moon subsequently receded from the Earth as a result of tidal friction, to take up its present orbit about 60 Earth radii away.

The rain of cosmic debris onto the surfaces of the planets diminished after accretion was largely complete, but did not end completely. Many of the most heavily cratered surfaces in the solar system—on the Moon, the southern highlands of Mars, and Mercury—were created during the era of late heavy bombardment, a time between 4.6 and about 3.8 billion years ago, when the cratering rate was hundreds of times larger than at present. Many lunar geologists believe that most of the craters dating from this era formed during a brief but intense cratering event called the terminal cataclysm about 3.9 billion years ago.

In the present era, comets from the outer solar system and asteroids straying from the main asteroid belt occasionally cross the path of the major planets and collide. There are about 170 impact craters currently recognized on the Earth, and new ones are discovered each year. The current cratering rate for the whole Earth is about 2.5 craters greater than 20 km (12 mi) in diameter per million years.

The most dramatic impact event in recent geologic history occurred 65 million years ago, when a 15-km-diameter (9-mi) asteroid or comet struck what is now the northern Yucatan Peninsula and created the 180-km-diameter (110-mi) Chicxulub crater. The impact sprayed vaporized projectile material to a thickness of a few millimeters over the entire Earth, rained melted rock debris over most of North and South America, initiated massive tsunamis and underwater landslides, ignited global wildfires, and may have produced long-term climatic changes by introducing massive amounts of water, carbon dioxide, and sulfates into the upper atmosphere. These sudden and extreme environmental changes caused a tremendous number of animal species to become extinct: Dinosaurs disappeared forever, along with the marine ammonites and many other groups. Marine planktonic communities changed drastically, with the disappearance of many species of foraminifera. *See* EXTINCTION (BIOLOGY); TSUNAMI.

Because at least one major biological extinction was caused by an impact, a vigorous debate has ensued about how large a hazard they pose to humanity. Efforts have begun to detect and chart the orbits and sizes of all asteroids larger than about 1 km (0.6 mi) in diameter that cross the Earth's orbit. To date, about 600 asteroids larger than 1 km in diameter have been charted, and it is estimated that there are another 600 yet to be discovered. Whether or not asteroid or comet impacts are an immediate threat to humanity, it is clear that meteorite impact is far more important than previously supposed for the ancient development of the solar system, the origin of planetary landscapes, and the evolution of life on Earth.

H. J. Melosh

Bibliography. T. J. Bernatowicz and E. Zinner (eds.), *Astrophysical Implications of the Laboratory Study of Presolar Materials*, American Institute of Physics, New York, 1997; A. Bevan and J. R. de Laeter, *Meteorites: A Journey Through Space and Time*, Smithsonian Institution Press, 2002; W. F. Bottke, Jr., et al. (eds.), *Asteroids III*, Arizona University Press, 2002; V. F. Buchwald, *Handbook of Iron Meteorites: Their History, Distribution, Composition and Structure*, 3 vols., University of California Press, 1975; R. N Clayton, R. W. Hinton, and A. M. Davis, Isotopic variations in the rock-forming elements in meteorites, *Phil. Trans. Roy. Soc. London*, A325:483–501, 1988; A. M. Davis (ed.), *Treatise on Geochemistry*, vol. 1: *Meteorites, Comets, and Planets*, Elsevier, 2004; R. T. Dodd, *Thunderstones and Shooting Stars*, 1986; R. A. F. Grieve, Terrestrial impact: The record in the rocks, *Meteoritics*, 26:175–194, 1991; F. Heide and F. Wlotzka, *Meteorites: Messengers from Space*, Springer-Verlag, New York, 1995; R. H. Hewins, R. H. Jones, and E. R. D. Scott (eds.), *Chondrules and the Protoplanetary Disk*, Cambridge University Press, 1996; R. Hutchison, *Meteorites: A Petrologic, Chemical and Isotopic Synthesis*, Cambridge University Press, 2004; R. Hutchison, *The Search for Our Beginning*, 1983; J. F. Kerridge and M. S. Matthews (eds.), *Meteorites and the Early Solar System*, 1988; D. S. Lauretta and H. Y. McSween (eds.), *Meteorites and The Early Solar System II*, University of Arizona Press, 2006; M. E. Lipschutz and L. Schultz, Meteorites, in P. Weissman, L.-A. McFadden, and T. V. Johnson (eds.), *Encyclopedia of the Solar System*, pp. 629–671, Academic Press, 1998; V. Mannings, A. P. Boss, and S. S. Russell (eds.), *Protostars and Planets IV*, University of Arizona Press, Tucson, 2000; H. Y. McSween, Jr., *Meteorites and Their Parent Planets*, 2d ed., 1999; K. Mark, *Meteorite Craters*, 1987; H. J. Melosh, *Impact Cratering: A Geologic Process*, 1989; O. R. Norton, *Rocks from Space*, 2d ed., Mountain Press Publishing, 1998; O. R. Norton, *The Cambridge Encyclopedia of Meteorites*, Cambridge University Press, 2002; A. E. Rubin, *Disturbing the Solar System*, Princeton University Press, 2002; S. R. Taylor, *Solar System Evolution*, 2d ed., Cambridge University Press, 2001;

J. T. Wasson, *Meteorites: Their Record of Early Solar System History*, 1985; D. E. Wilhelms, *The Geologic History of the Moon*, USGS Prof. Pap. 1348, 1987; D. E. Wilhelms, *To a Rocky Moon*, 1994; B. Zanda and M. Rotaru (eds.), *Meteorites: Their Impact on Science and History*, Cambridge University Press, 2001.

Meteorological instrumentation

Devices that measure or estimate properties of the Earth's atmosphere. Meteorological instruments take many forms, from simple mercury thermometers and barometers to complex observing systems that remotely sense winds, thermodynamic properties, and chemical constituents over large volumes of the atmosphere.

The method for monitoring the atmosphere depends strongly on the application. The atmosphere is observed in order to unravel the mysteries of storms, to describe current weather for aviation operations, to provide warnings of tornadoes, hail, and floods, to predict the occurrence of wind-driven oceanic waves, to estimate monsoon precipitation, to anticipate abrupt climate change, as well as to deal with other issues. Types of instruments include thermometers, barometers, hygrometers, anemometers, rain gages, rawinsondes, radiometers, spectrometers, radars, and lidars.

The atmosphere is observed on time scales from millennia (climate change) to milliseconds (aircraft turbulence) and on spatial scales from planetary waves down to the microstructure of ice crystals and the specks of dust (nuclei) on which ice crystals form. A very important middle scale, or mesoscale, of atmospheric circulations contains organized patches of fair or foul weather. At this scale, circulations are particularly adept at suppressing or producing precipitation, large temperature changes, and rapidly varying winds. Mesoscale weather events include temperature fronts, rain and snow storms, thunderstorms, hurricanes, tornadoes, downbursts, mountain-valley circulations, and sea breezes. *See* HURRICANE; MESOMETEOROLOGY; THUNDERSTORM.

Most weather is confined to the troposphere, which is nominally the lowest 8–16 km (5–10 mi) of atmosphere. Just above the Earth's surface and within the lower troposphere is the planetary boundary layer, with unique properties that require special meteorological instrumentation. Most routine atmospheric observations are made near the surface, and most of these are located over land. To fill the data voids, meteorologists use mobile and remote platforms on which to place instrumentation. Examples include satellites, balloons, crewed aircraft, robotic aircraft, ships, buoys, and even oil rigs. *See* TROPOSPHERE.

Variables, processes, and phenomena. The standard meteorological variables are air temperature, air pressure, humidity, winds, and precipitation. Temperature is a measure of molecular kinetic energy. Air pressure is the columnar weight of atmosphere per unit area extending from the altitude of observation to the outer limits of the Earth's atmosphere; it is usually expressed in kilopascals (kPa) or columnar inches of mercury (in. Hg). Humidity is the amount of water vapor in the air, expressed in several ways. Absolute humidity is the mass of water vapor per unit volume of air. Specific humidity is a dimensionless mass known as mixing ratio; for example, X grams (lb) of water vapor per Y grams (lb) of air mixture. The most familiar expression of atmospheric water vapor content is relative humidity, the percentage of water vapor in the air compared to the maximum possible. Instruments measure humidity several ways, including absolute, specific, and relative humidity. *See* AIR PRESSURE; AIR TEMPERATURE; HUMIDITY.

Precipitation is measured as a flux of water mass from the atmosphere to the surface. The water-mass flux is referred to as rainfall rate (mm/h or in./h). The density of ice-phase precipitation is highly variable, depending on the conditions of ice-particle growth in clouds and in the air below. Therefore, when precipitation is in the form of snow or hail, the mass flux is weighed and converted to a rainfall depth equivalent. *See* PRECIPITATION (METEOROLOGY); SNOW SURVEYING.

Wind measurements are critical to nearly all meteorological research and societal applications of meteorology. Measurement of the vertical wind speed, combined with information on humidity and temperature, allows computation of cloud condensate production. Estimating vertical air speed is crucial to quantitative forecasts of precipitation. Horizontal wind is much easier to measure than vertical wind, typically being of order 100 times larger. Precise knowledge of the horizontal wind field allows accurate computation of the vertical wind field. *See* WIND.

Beyond these basic variables, there are many other atmospheric quantities of interest which are measured. Among these are gases that vary in concentration, such as carbon dioxide, methane, ozone, oxides of nitrogen and sulfur, and hydrocarbons.

Particle measurements, including dry aerosols and hydrometeors, are an important category of meteorological instrumentation. Such particles affect the fair-weather electric field and determine the properties of clouds. Nearly all cloud droplets and ice crystals begin their growth on aerosols. Only a small fraction of aerosols are active as cloud condensation nuclei or ice nuclei. *See* AEROSOL.

Hydrometeor measurements are at the heart of meteorology. Types of hydrometeors include cloud waterdroplets, raindrops, individual ice crystals, snowflakes (which are aggregates of crystals), frozen cloud droplets, sleet (which is frozen raindrops), graupel (which are aggregates of frozen cloud droplets), and hail. Measuring the mass of hydrometeors is an extremely important piece of information for reasons related to buoyancy and determining the amount of heat released into the atmosphere by the processes of condensation from vapor to water, freezing from water to ice, and deposition from vapor to ice. *See* CLOUD PHYSICS; HAIL; SNOW.

In weather forecasting and in meteorological

research, the object of individual variable measurements is often the detection, quantification, and understanding of complex phenomena, such as storms, temperature fronts, sea breezes, jet streams, and a host of terrain-induced circulations. To detect and understand the evolution of such phenomena, it is necessary to measure the air motions, kinematics, and to estimate the thermodynamic structure, including temperature deviations, pressure perturbations, and humidity inhomogeneities. When storms produce precipitation, hydrometeor or microphysical measurements complete the characterization of weather systems. The combination of kinematic, thermodynamic, and microphysical information provides the meteorologist with a basis to detect, quantify, understand, and forecast weather events. *See* WEATHER FORECASTING AND PREDICTION.

Sampling considerations. There is no instrument designed to cover the full range of atmospheric temporal or spatial scales. If the purpose is to detect evidence of abrupt global climate warming, it might be sufficient to place mercury thermometers at 50 representative locations throughout the world and measure temperature at a few standard times per day for 20 years. Choosing representative locations is a challenge because of the effects of urbanization and heat island effects. If the purpose is to directly measure the fluxes of sensible heat inside a thunderstorm, independent measurements of temperature might be made from several research aircraft every 0.01 s for a period of 1–2 h. Both objectives require about one million measurements of temperature with similar accuracy and precision. Since the frequency, spacing, duration, and sampling domains differ radically, so do the measurement technologies. *See* URBAN CLIMATOLOGY.

Some research measurements are best conducted under highly controlled conditions in the laboratory, such as detailed microphysical measurements of water and ice nucleation, vapor deposition growth of ice crystals, and rime growth of graupel from the collection of supercooled cloud droplets. Other characteristics such as hydrometeor terminal fallspeed, orientation, and binary interactions among particles are also measured in laboratories. Slowly reacting trace gases and aerosol distributions are spectroscopically analyzed in the laboratory after air samples are collected from research aircraft.

Instruments, observing systems, and platforms. A few broad categories of apparatus are traditional and modern immersion devices and passive and active remote sensors. Traditional immersion devices make contact with the air and are nearly always mechanical. They rely on the expansion properties of materials, on the wind pushing parts, and on mechanical or volumetric measurement of mass. Modern immersion instruments are less mechanical and typically rely on the electrically resistive, capacitive, and radiative properties of sensor materials and of the atmosphere itself (**Fig. 1**). Passive remote-sensing devices usually receive naturally emitted electromagnetic radiation from the atmosphere. Most often, these instruments are referred to as radiometers. Active remote-sensing devices, such as radars, transmit electromagnetic or acoustic radiation at specific frequencies. After a brief delay to allow for electromagnetic wave propagation, these instruments receive a small fraction of the transmitted energy, reflected, refracted, or scattered by the atmosphere. The intensity, frequency, phase, and polarization of the received energy reveal various microphysical, kinematic, and thermodynamic properties of the atmosphere. *See* ACOUSTIC RADIOMETER; METEOROLOGICAL RADAR; METEOROLOGICAL SATELLITES.

Fig. 1. Portable, automated weather station which operates on solar power. Propeller anemometers are at the top of the mast (10 m or 33 ft). The thermometer/psychrometer is located in a beehive radiation shield. Sensor electronics, including aneroid barometer, are in the white box. The geostationary satellite data uplink antenna is above the solar panels. (*National Center for Atmospheric Research*)

Traditional weather station measurements provide a description of conditions near the ground. In addition to the average regional conditions, these measurements also provide local information on mesoscale phenomena such as cold fronts, sea breezes, and disturbed conditions resulting from nearby thunderstorms. Traditional thermodynamic instruments are mechanical or heat-conductive devices relying on the expansion and contraction of metallic and nonmetallic liquids or solid materials as a function of temperature, pressure, and humidity. Among these are the mercury, alcohol, and bimetallic thermometers for measurement of temperature, mercury and metallic bellow (aneroid) barometers for measurement of pressure, human hair hygrometers, and wet/dry-bulb thermometers (called psychrometers) for measurement of relative humidity. Mercury barometers are simply weighing devices that balance the mass of the atmospheric column against the mass of a mercury column. On average, a column of atmosphere weighs the same as 76 cm (29.92 in.) of mercury. Psychrometers measure humidity by means of the wet-bulb depression technique. A moist thermometer is cooled by evaporation when relative humidity is less than 100%. The

temperature difference between wet and dry thermometers is referred to as the wet-bulb depression, a well-known function of relative humidity at standard airflow speeds. A related method of humidity measurement is the chilled mirror technique (dewpointer). A polished surface is cooled to the temperature of water vapor saturation, at which point the cooled surface becomes fogged. Dewpoint saturation uniquely defines humidity at a known temperature and pressure. *See* BAROMETER; DEW POINT; HYGROMETER; PSYCHROMETER; TEMPERATURE MEASUREMENT; THERMOMETER.

Modern in-place thermodynamic measurements usually employ electrical responses to temperature and humidity. A so-called thermocouple naturally generates electric current in a loop between the junctions of two dissimilar metals at two locations of different temperature. The temperature dependence of electrical resistance in metals, such as platinum, or semiconductors is often employed for accurate and fast measurement of temperature. Because water vapor is absorbed by porous materials, electrical resistance or capacitance of substances, such as carbon or thin film polymers, is a strong function of relative humidity. Increasingly, pressure is measured by thin silicon piezo-resistive membranes, which stretch in response to pressure change, thereby changing their electrical resistance properties. Quasi-spectroscopic methods are employed for very fast response applications in measurement of water vapor fluxes, density, and temperature. Energy from calibrated sources of infrared or ultraviolet radiation are absorbed by water vapor along short transmission paths. The energy losses are in direct proportion to the number of water vapor molecules in the path, and these are measured at the receiver. *See* THERMOCOUPLE.

Traditional precipitation measurement devices may be described as precision buckets, which measure the depth or weight of that which falls into them. These gages work best for rainfall, but they are also used in an electrically heated mode for weighing snow. Rulers are routinely used for measurement of snow depth. Time-resolved measurements of rainfall are traditionally made by counting quantum amounts (0.01 in. or 0.25 mm) of rain with a small, mechanically controlled tipping bucket located beneath a large collecting orifice. Modern rain measuring is sometimes performed along short paths via drop-induced scintillations of infrared radiation, which is emitted by a laser. When the raindrop size distribution is needed, optical-shadowing spectrometers are employed, as are momentum-measuring impact distrometers, devices that measure the number density versus the size distribution of raindrops or other hydrometeors (**Fig. 2**). *See* SNOW GAGE.

Traditional wind measurements are performed by anemometers, some of which use wind-driven spinning cups for wind speed determination. Vanes are used in conjunction with cups for indication of wind direction. Alternatively, three-axis propeller anemometers may be employed to provide orthogonal components of the three-dimensional wind vector. Many hybrids of these basic approaches continue to be successfully employed. Fast-response sonic anemometers employ ultrasound transmission, where the apparent propagation speed of sound is measured (**Fig. 3**). The difference between this measured speed and the actual speed for a fluid at rest is the wind speed. Such measurements are made on a time scale 0.01 s and are used to determine the fluxes of momentum, water vapor, sensible heat, and other scalars in the planetary boundary layer. *See* ANEMOMETER; WIND MEASUREMENT.

Balloon-borne vertical profiles or soundings of temperature, humidity, and winds are central to computerized (numerical) weather prediction. Such observations are made simultaneously or synoptically worldwide on a daily basis. The temperature and humidity sensors are lightweight expendable versions of traditional surface station instruments. Balloon drift during ascent provides the wind measurement. The preferred method of tracking these rawinsondes is to use global navigation aid systems such as Omega, Loran-C, and the Global Positioning System.

Fig. 2. Hydrometeor optical spectrometers and temperature sensors on a research aircraft. Optical spectrometers use lasers to create particle shadow images in two dimensions. Particles are imaged as they flow between the blunt arms ahead of the electronics pod beneath the wing. (*University of Wyoming*)

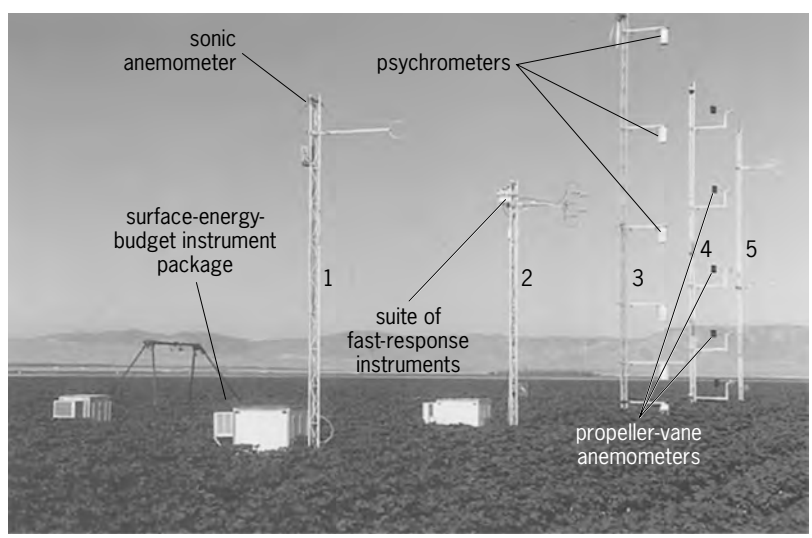

Fig. 3. Atmospheric turbulent energy exchange research facility. Tower 3 makes psychrometric profiles of mean temperature and humidity. Tower 4 measures mean wind profiles with propeller vanes. Towers 1, 2, and 5 support sonic anemometers for high-speed turbulent wind components together with various fast scaler measurements, including temperature, humidity, O_3, and CO_2. The black stand in the background supports visible/infrared radiation measurements. (*After W. F. Dabberdt et al., Atmosphere-surface exchange measurements, Nature, 260:1472–1481, 1993*)

Parachute-borne dropsondes are often released from aircraft in data-sparse regions. *See* LORAN; SATELLITE (ASTRONOMY).

Remote sensing. Remote sensing, principally via electromagnetic radiation, is a mainstay of modern meteorology. Such devices typically operate in the optical, infrared, millimeter-wave, microwave, and high-frequency radio regions of the electromagnetic spectrum.

Passive radiometers typically operate at infrared and microwave frequencies; they are used for estimates of temperature, water vapor, cloud heights, cloud liquid water mass, and trace-gas concentrations. These observations are made from the ground, aircraft, and satellites, usually measuring naturally emitted radiation. Passive instruments are especially effective in Earth orbit because electrical power consumption is low and global coverage is unique. Some radiometers simply estimate columnar amounts of water vapor averaged over the whole troposphere. Multiple-frequency infrared and millimeter-wave radiometers can resolve smoothed profiles of water vapor. More complex radiometers are able to scan and image a horizontal field of view. Advanced radiometers are able to resolve radiation over narrow frequency bands that correspond to molecular absorption lines for various gases, including water vapor, ozone, and methane. Other radiometers can receive radiowaves from human sources that are transmitted by the Global Positioning System of satellites to estimate temperature and water vapor from wave propagation characteristics such as signal phase and angles of refraction over known distance. Microwave radiometers can be tuned to oxygen absorption lines to measure temperature profiles, because the mixing ratio of oxygen in the free troposphere is essentially uniform and known with high precision. Imaging optical radiometers are more commonly known as cameras, and these are routinely mounted on satellites to track the movement of clouds and storm systems. *See* SATELLITE NAVIGATION SYSTEMS.

Radarlike, active remote-sensing devices are among the most powerful tools available to meteorology. Collectively, these instruments are capable of measuring kinematic, microphysical, chemical, and thermodynamic properties of the troposphere at high spatial and temporal resolution. Active meteorological remote sensors are principally deployed on land, ships, and aircraft platforms, as well as aboard satellites. Unlike passive instruments, active remote sensors can precisely resolve the distance at which a measurement is located.

At optical frequencies, lidars measure conditions in relatively clear air. Capabilities include determining the properties of tenuous clouds; determining concentrations of aerosol, ozone, and water vapor; and measuring winds through the Doppler frequency-shift effect. Millimeter-wave radars are used to probe opaque, nonprecipitating clouds. Polarimetric and Doppler techniques reveal hydrometeor type, water mass, and air motions. *See* LIDAR.

The best-known meteorological remote sensor is the microwave weather radar. In addition to measuring rainfall and tracking movement of storms, powerful and sensitive meteorological radars can measure detailed flow fields in and around storms by using

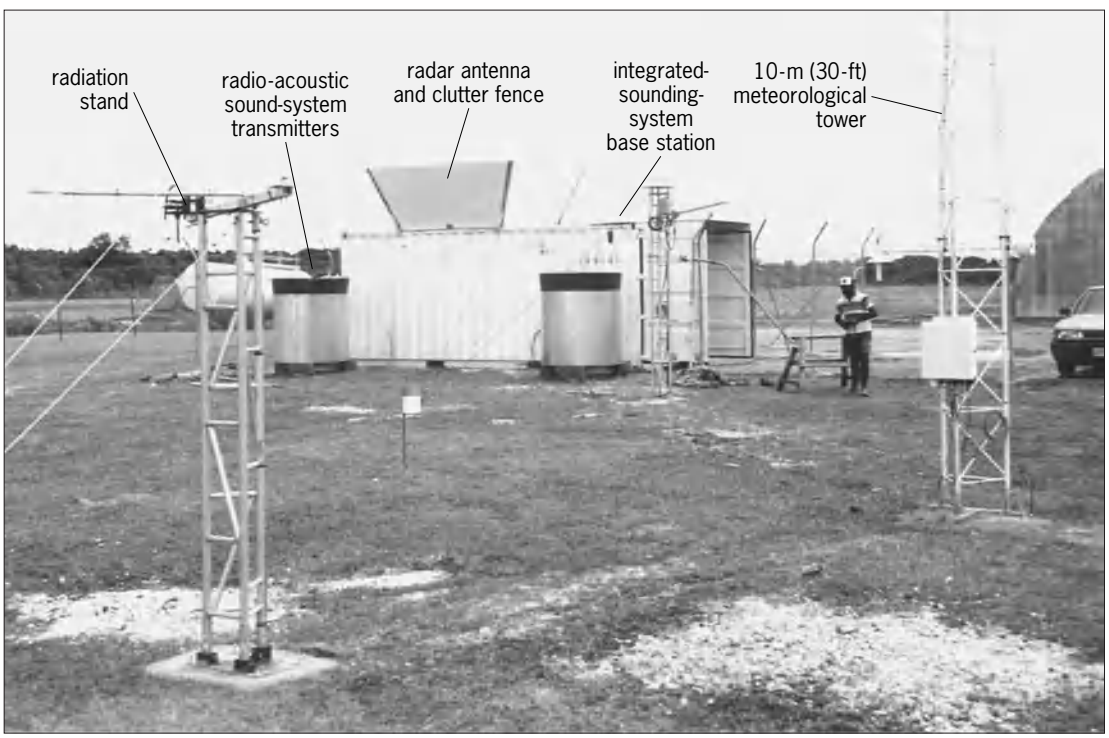

Fig. 4. Integrated sounding/profiling system which contains an ultrahigh-frequency radar wind profiler, acoustic transmission for radio-acoustic air density profiling, a conventional Omega balloon sonde capability (in seatainer), and tower-mounted meteorological instruments. (*After D. W. Parsons et al., The sounding system: Description and preliminary from TOGA COARE, Bull. Amer. Meteorol. Soc., 75:553–567, 1994*)

hydrometeors, insects, and blobs of water vapor as reflective targets. Employing polarimetric methods, these radars can also distinguish between rain, hail, and snow. When Doppler measurements are combined with the atmospheric equations of motion, thermodynamic perturbation fields, such as buoyancy, are revealed inside violent convective storms. At ultrahigh and very high radio frequencies, radars known as wind profilers measure the mean wind as a function of height in the clear and cloudy air (**Fig. 4**). Superior to infrequent weather balloons, radio wind profiling methods permit continuous measurement of winds with regularity and high accuracy. When radio wind profilers are colocated with acoustic transponders, the speed of sound is easily measured through radar tracking of the acoustic wave. This permits the computation of atmospheric density and temperature profiles, on which the speed of sound is strongly dependent. *See* DOPPLER RADAR; METEOROLOGY; RADAR METEOROLOGY; REMOTE SENSING.

Richard E. Carbone

Bibliography. D. Atlas (ed.), *Radar in Meteorology*, 1990; W. F. Dabberdt et al., Atmosphere-surface exchange measurements, *Nature*, 260:1472–1481, 1993; T. P. DeFelice, *An Introduction to Meteorological Instrumentation and Measurement*, 1997; D. H. Lenschow (ed.), *Probing the Atmospheric Boundary Layer*, 1986; D. W. Parsons et al., The integrated sounding system: Description and preliminary observations from TOGA COARE, *Bull. Amer. Meteorol. Soc.*, 75:553–567, 1994; S. Raghavan, *Radar Meteorology*, 2003; R. Ware et al., GPS sounding of the atmosphere from low Earth orbit: Preliminary results, *Bull. Amer. Meteorol. Soc.*, 77:19–40, 1996.

Meteorological optics

The study of optical phenomena occurring in the atmosphere. Many light effects can be seen by looking skyward, and all of them, resulting from the interaction of light with the atmosphere, lie in the province of atmospheric optics or meteorological optics. The subject also includes the effect of light waves too long or too short to be detected by the human eye—light-type radiation in the infrared or ultraviolet regions of the spectrum. Light interacts with the different components of the atmosphere by a variety of physical processes, the most important being scattering, reflection, refraction, diffraction, absorption, and emission. Some other processes involving photochemistry and ionization are not considered in this article. *See* ATMOSPHERE; OPTICS.

Scattering. An observer of light in the sky, while looking in a direction away from the Sun, will see evidence that some process has changed the path of the sunlight to direct it to the observer's eye. For most of the light in the sky this process is that of scattering, by which some of the incident light is sent off in all directions. Scattering by dust in the air makes visible the beam of light coming into a room through a window. Scattering by dust particles or liquid droplets in the atmosphere can add to the brightness of the sky, but the sky would appear light even if the atmosphere were free of all such particles. On a submicroscopic scale, air is not a continuous, uniform fluid but is composed of molecules, which are smaller than the wavelength of visible light by a factor of about a thousand. Such small particles scatter light but do not scatter all wavelengths with equal efficiency. They scatter the short light waves (blue) more strongly than the long waves (red), with the result that the clear sky appears blue.

The Sun's disk, seen high in the sky on a clear day, appears white. As it moves closer to the horizon, its color changes to yellow, and to orange, and perhaps to red. When the low Sun is observed, the light rays that reach the eyes travel a much longer path through the Earth's atmosphere than when the Sun is overhead. Throughout this path, the shorter (blue) wavelengths are selectively scattered in all directions, with the result that the unscattered light that reaches the observer from the Sun is depleted in the blue end of the spectrum and appears orange or red. Thus the blue sky and the red setting Sun are both consequences of the same scattering process.

Larger particles in the atmosphere scatter light more strongly than the gas molecules. If they are of sizes less than the wavelength of visible light, they will selectively scatter the shorter, visible wavelength and contribute to red sunsets. Smoke from large forest fires or fine ejected material from volcanic eruptions can enhance the colors of sunsets thousands of miles away. In fact, following major volcanic explosions, sunsets may be enhanced over the entire Earth as fine ejected material mixes with the atmosphere and circles the Earth with a fallout time of up to a few years.

White clouds are composed of either small transparent waterdroplets or small transparent ice crystals, but cloud particles are small on a scale different from the scatterers of the blue sky. The droplets are typically 50 times larger than the wavelength of visible light, and so would be considered large particles in light scattering. Such large particles scatter light of all visible wavelengths equally well so that the scattered light would have the same spectral distribution as the incident sunlight, which is known as white light. Even if there is some wavelength dependence of the scattered light from one particle, multiple scattering from many particles in a cloud of the appropriate thickness will result in the scattered light appearing white. If the atmosphere of the Earth were made thicker (that is, if air were added), the sky would appear whiter. This effect can be seen even on a clear day: the sky near the horizon is whiter than the sky overhead. There is more air along a sighting path near the horizon than along a vertical path through the atmospheric layer.

The blue color of distant mountains, hazes, or smogs demonstrates the wavelength dependence of light scattering from small particles. Because red or infrared radiation is scattered less by some hazes, red filters or infrared-sensitive film or detectors are sometimes used in aerial photography or satellite image recording for better penetration of haze layers.

See AERIAL PHOTOGRAPH; REMOTE SENSING; SCATTERING OF ELECTROMAGNETIC RADIATION.

Reflection. When light encounters the smooth surface of an ice crystal, some of the light is reflected from the surface and some is transmitted. Sometimes many small ice crystals grow in the atmosphere and slowly fall through the air. Their presence is seen as clouds or as a haze or (ice) fog. The very complicated, many branched ice crystals known as snowflakes usually occur at ground level in temperate regions; simple ice-crystal forms can be found at higher elevation in the sky.

As hexagonal flat-plate crystals fall through the air, they tend to orient with their wide, flat surfaces nearly horizontal. Reflection of sunlight from the nearly horizontal surfaces of such crystals can give rise to a vertical column of light that appears above or below the Sun (or both), when the Sun is near the horizon. These sun pillars are similar in origin to the streak of sunlight (called a glitter path) reflected from the slightly rough surface of a lake or ocean.

Sun pillars can also be produced by hexagonal-column crystals that have the shape of a wooden pencil (before sharpening). Such crystals tend to fall with their long axes horizontal, and light reflection from the long side faces produce sun pillars. Reflection from the vertically oriented faces of falling crystals (the side faces of falling plates or the end faces of falling pencils) can result in the parhelic circle, a band of light parallel to the horizon, passing through the Sun and extending all the way around the sky. Many other effects can result from a combination of reflection and refraction in falling ice crystals or raindrops.

A particular form of a sun pillar, called the subsun, can be seen when looking down from an airplane flying over a layer of ice-crystal clouds (see **illustration**). When the Sun is high in the sky and the crystals are well oriented, a somewhat elongated, bright spot can be seen. It is just a reflected image of the Sun in the nearly horizontal ice-crystal surfaces. *See* REFLECTION OF ELECTROMAGNETIC RADIATION.

Refraction. When a ray of light passes from one transparent medium to another, it is refracted at the surface, that is, its direction changes abruptly. Light from the Sun is thus refracted as it enters an ice crystal and again as it leaves. As a result of these two refractions, light deviates from its original direction. For a hexagonal ice crystal the deviation can bring sunlight to an observer's eye from different directions, producing a variety of spots, arcs, or streaks of light intensity in the sky. Sun dogs and the 22° halo are two of the more commonly observed effects that result from light refracted by falling ice crystals. Light that enters an ice crystal and is internally reflected from one or more crystal faces before emerging can produce a wide variety of halo effects that are observable with the naked eye. *See* HALO; SUN DOG.

In the case of water, where surface tension acts as an elastic skin, squeezing a small falling drop into a spherical shape, light rays that enter the transparent sphere are reflected internally before they emerge to produce a rainbow. *See* RAINBOW.

The subsun, resulting from sunlight reflected off the nearly horizontal faces of falling flat-plate ice crystals. (*From R. Greenler, Rainbows, Halos, and Glories, Cambridge University Press, 1980*)

When a light ray passes through air whose density is not uniform but changes gradually with position, the ray gradually changes its direction, traveling in a smooth curve. The normal variation in atmospheric pressure (and hence of air density) with height causes stars near the horizon to appear elevated above their true position. The curved paths of the light rays enable observation of the Sun, apparently sitting just above the horizon, when it actually is located geometrically just below the horizon. The density of air also changes as a result of the air temperature. Temperature variation of air near the Earth's surface can produce a number of optical distortions that are referred to as mirages. Small-scale variations and temporal fluctuations of air density resulting from air turbulence or temperature variations result in the twinkling of stars or the shimmering of distant scenes. *See* MIRAGE; REFRACTION OF WAVES; TWINKLING STARS.

Diffraction. Diffraction depends on both the wavelength of the light and the size of the particles. If the Moon is viewed through a thin cloud containing waterdroplets or small ice crystals, it seems to be surrounded by colored rings. The rings are a diffraction effect, with the long (red) wavelengths giving rise to a larger set of rings than the shorter (blue) wavelengths. The resulting display is called the corona. A similar display, called the glory, can be seen by looking in exactly the opposite direction from the Sun

or Moon; an observer flying in the sunlight over a cloud layer can see the colored rings of glory around the shadow of the airplane on the clouds below. *See* DIFFRACTION; SOLAR CORONA.

Emission and absorption. Small particles in the air, and the gas molecules that constitute the air, can absorb and emit light. In a heavily polluted atmosphere the color of the sky may be affected by the selective color absorption of smoke particles, but in general there is not much absorption or emission of visible light in the atmosphere. The absorption and emission of infrared radiation, however, are very important processes in establishing the temperature of the Earth's atmosphere and the Earth's surface. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; AIR POLLUTION; GREENHOUSE EFFECT. Robert Greenler

Bibliography. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, 1983; K. L. Coulson, *Polarization and Intensity of Light in the Atmosphere*, 1988; R. Greenler, *Rainbows, Halos, and Glories*, 1989; K. N. Liou, *Radiation and Cloud Physics Processes in the Atmosphere*, 1992; M. Minnaert, *Light and Colour in the Open Air*, 1954; W. Tape, *Atmospheric Halos*, 1994; W. Tape and J Moilanen, *Atmospheric Halos and the Search for Angle X*, 2005; R. A. R. Tricker, *Introduction to Meteorological Optics*, 1970.

Meteorological radar

A remote-sensing device that transmits and receives microwave radiation for the purpose of detecting and measuring weather phenomena. Radar is an acronym for radio detection and ranging. Today, many types of sophisticated radars are used in meteorology, ranging from Doppler radars, which are used to determine air motions (for example, to detect tornadoes), to multiparameter radars, which provide information on the phase (ice or liquid), shape, and size of hydrometeors. Airborne Doppler radars play a vital role in meteorological research. Additionally, weather radar is now orbiting the Earth on the NASA *Tropical Rainfall Measuring Mission* satellite, launched in November 1997. Radars are also used to detect hail, estimate rainfall rates, probe the clear-air atmosphere to monitor wind patterns, and study the electrification processes in thunderstorms that generate lightning discharges.

Doppler radar. Commonly used, pulsed Doppler radar operates in the microwave region, with standard wavelengths of 10, 5, and 3 cm, referred to as S-, C-, and X-band radars, respectively. The electromagnetic radiation is focused into a narrow beam by illuminating a parabolic dish reflector with microwave energy provided by the radar transmitter. S-band radars require the use of large antennas (**Fig. 1**) to generate a narrow beam of microwave energy; transmit high power (peak power of 1 megawatt); and suffer relatively little attenuation as the radar beam passes through regions of heavy rain and hail. X-band radars use much smaller antennas to achieve similar narrow beams, and are highly portable. However, X-band radars suffer from attenuation when used to probe precipitation, which significantly limits their range. Attenuation results when the radar energy is either absorbed by the raindrops or reemitted from the raindrops in directions other than toward the radar. *See* ANTENNA (ELECTROMAGNETISM); DOPPLER RADAR; MICROWAVE.

A pulsed Doppler radar typically emits 1000 electromagnetic pulses per second. These individual pulses are typically 1 microsecond (10^{-6} s) in duration. The pulsing rate defines the pulse repetition frequency (PRF) of the radar, which in turn determines its maximum unambiguous range. For a PRF of 1000 Hz, the maximum range is 150 km (90 mi). Strong echoes beyond this range can still be observed, but are said to be range-folded. The Doppler radar also provides information on the target's velocity, either toward or away from the radar when viewed along the radar beam. The Doppler shift, which is measured as a small difference between the frequency of the transmitted pulse and the frequency of the energy backscattered to the radar, provides a measure of the scatterer's radial motion. The Doppler shift frequency Δf is given by Eq. (1),

$$\Delta f = 2v/\lambda \tag{1}$$

where λ is the wavelength of the radar expressed in meters and v is the radial velocity. Scatterers in the case of meteorological radar include raindrops, ice particles (snowflakes), hailstones, and even insects, providing clear air returns. A Doppler radar

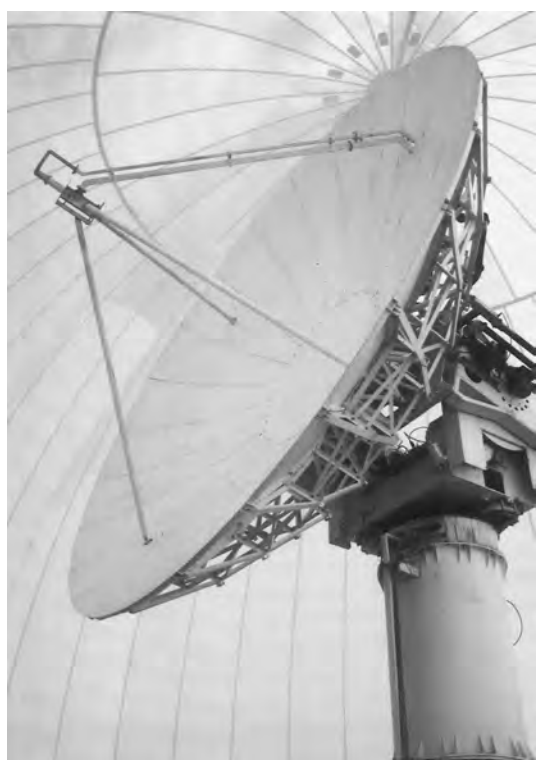

Fig. 1. Parabolic dish antenna of the CSU-CHILL 11-cm multiparameter Doppler radar operating at Colorado State University. The antenna is housed within a large, inflatable radome. (*P. Kennedy, Colorado State University*)

also detects the amplitude of the backscattered signal, which can be used as a measure of storm intensity and as a means of estimating rainfall rates. When the radar wavelength is large compared to the diameter of the meteorological targets, Rayleigh scattering conditions occur, and thus the energy backscattered from a raindrop or ice particle is proportional to the sixth power of the particle diameter. Under Rayleigh scattering conditions, the reflectivity factor Z is defined by Eq. (2), where $N(D)\Delta D$ is the

$$Z = \sum_{0}^{\infty} N(D)D^6 \Delta D \qquad (2)$$

concentration of scatterers (m^{-3}) and D is the particle diameter expressed in millimeters (mm^6). *See* PRECIPITATION (METEOROLOGY); SCATTERING OF ELECTROMAGNETIC RADIATION; STORM; STORM DETECTION.

Airborne Doppler radar. In the early 1980s, researchers began using aircraft equipped with Doppler radar to study storms. Several advantages are provided by airborne Doppler systems compared to surface-based systems, including portability and the immediate integration of other meteorological sensors with the Doppler data, thereby providing comprehensive sampling of storms. However, because of limitations of weight and antenna size, airborne Doppler radars operate at wavelengths of 3 cm or smaller. Since attenuation is significant at these ranges, airborne Doppler radars provide usable data out to about 50 km (30 mi). Currently, airborne Doppler radars are used for meteorological research on aircraft operated by the National Oceanic and Atmospheric Administration (NOAA), NASA, the National Center for Atmospheric Research (NCAR), and the University of Wyoming (the last two facilities are under the sponsorship of the National Science Foundation). NCAR maintains a sophisticated airborne Doppler radar, known as ELDORA (Electra Doppler Radar). ELDORA actually consists of two complete radars (each transmitting five different frequencies) driving a dual-antenna system to direct radar beams both forward and rearward of a line perpendicular to the aircraft's flight path. The antennas are housed within a large, rotatable radome affixed to the tail of the aircraft (**Fig. 2**). These forward- and rearward-looking beams intersect at various distances from the aircraft, and at these points of intersection, both the east-west (u) and north-south (v) components of the horizontal wind can be determined. When the horizontal gradients of u and v are combined with the equation of mass continuity, the vertical wind speed may also be estimated. In this way, a single airborne radar can provide information on the three-dimensional airflow and precipitation structure in storms. The resolution of the ELDORA radar is about several hundred meters, providing unprecedented views of tornadoes and other severe weather.

Fig. 2. Electra aircraft; the antenna radome. (*R. Carbone, National Center for Atmospheric Research*)

Multiparameter radar. During the late 1970s and early 1980s, researchers began to experiment with more advanced Doppler radar systems involving dual-wavelength and polarimetric techniques. Dual-wavelength radar transmits electromagnetic energy at two wavelengths, and it also receives energy at both wavelengths. Typically, S- and X-band wavelengths are used. Dual-wavelength techniques were originally proposed to detect large hail. At S-band, hail is usually a Rayleigh target, whereas at X-band, hail is considered a Mie scatterer. Since radar energy is scattered in various directions by a Mie target, the power returned at the X-band wavelength is reduced relative to that at S-band. The presence of large hail is interpreted on the basis of the ratio of backscattered power at X-band to that at S-band. Dual-wavelength techniques are also used to estimate rainfall rates by comparing the backscattered power at a nonattenuating wavelength (S-band) to that at attenuating wavelengths (X-band). *See* HAIL.

Considerable research is being done with radars that use polarization diversity. Dual-polarization radar is able to transmit and receive both horizontally and vertically polarized radiation (the polarization of a radar beam is defined by the orientation of the electric field vector that comprises an electromagnetic wave). These radars are now used in meteorological research and have largely superseded dual-wavelength radars. A suite of multiparameter variables is being used to infer information on particle phase (ice or water), size, orientation, and shape. One key variable is the differential reflectivity Z_{dr} (decibels, dB), defined by Eq. (3),

$$Z_{dr} = 10 \log_{10}(Z_{hh}/Z_{vv}) \qquad (3)$$

where Z_{hh} is the reflectivity for horizontal transmit to horizontal receive and Z_{vv} is the reflectivity for vertical transmit to vertical receive. For raindrops larger than 1 mm in diameter (D), aerodynamic drag forces act to distort the drops into flattened spheres. Since the reflectivity is proportional to D^6, a flattened raindrop is larger, and therefore backscatters more power, when viewed with horizontally polarized radiation compared to vertical radiation. For raindrops 4 mm (0.16 in.) in diameter, the ratio of the horizontal axis to the vertical axis is about 1.2, which results in a differential reflectivity of approximately 5 dB. Ice particles such as graupel and hail do not deform, and therefore $Z_{dr} = 0$. By examining reflectivity and

differential reflectivity simultaneously, it is possible to distinguish hail from heavy rain, and locate water and ice regions within storms. *See* POLARIZATION OF WAVES.

Another important multiparameter variable is the differential phase Φ_{dp}, and the gradient in range of differential phase, K_{dp}, which is called the specific differential phase. Differential phase is large when flattened targets are present. Differential phase is a measure of the phase shift (measured in degrees) between horizontal and vertical polarization as the waves pass through a region containing flattened particles (such as raindrops). The specific differential phase (degrees per kilometer) is used to provide accurate estimates of intense rainfall rates in storms containing both rain and hail. Since the hail can dominate the reflectivity signature by virtue of its large size, reflectivity alone can overestimate rainfall rates. However, hail does not contribute to differential phase due to its quasi-spherical nature (only oriented particles or flattened raindrops contribute to differential phase shifts). Irregularly shaped hail tumbles as it falls, and appears to the radar as if it were spherical. Multiparameter radars provide a powerful means to study cloud, precipitation, and electrification processes in storms.

Space-borne radar. Radar operating at a wavelength of 2 cm is in low Earth orbit (350 km or 210 m above the surface) and is used for mapping tropical precipitation. Understanding the amount and distribution of tropical rainfall is crucial for better understanding the Earth's climate. This space-borne radar and associated satellite was jointly developed by the United States (NASA) and Japan (National Space Development Agency); it is known as the *Tropical Rainfall Measuring Mission* (*TRMM*) satellite (**Fig. 3**). Space-borne radar presents many challenging problems, including cost, size constraints, reliability issues, and temporal sampling. It is obviously impossible to continuously sample every precipitating cloud in the tropics from radar orbiting the Earth. But the *TRMM* satellite will help scientists develop a statistical distribution of rain rates within a certain area, and calculate the probability of a specific rain rate occurring. Based on this information, it will be possible to generate monthly mean rain amounts within areas of 10^5 km^2. Such information will be vital for the verification of climate models. *See* SATELLITE METEOROLOGY.

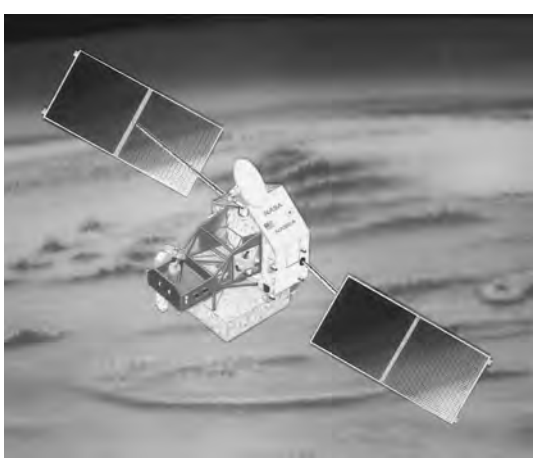

Fig. 3. ***Tropical Rainfall Measuring Mission (TRMM)*** **satellite. (*O. Thiele, National Aeronautics and Space Administration*)**

Clear-air and optical radars. Radars used to probe the clear air, or regions devoid of clouds, are known as profilers. A profiler is essentially a Doppler radar that operates at much longer wavelengths compared to weather radar. Wavelengths of 6 m, 70 cm, and 33 cm are commonly used. In the case of a profiler, the reflected power is not only from hydrometeors but also from gradients in the index of refraction of air, which are caused by turbulent motions in the atmosphere. These turbulent motions in turn cause small fluctuations in air temperature and moisture content, which also change the index of refraction. A profiler can determine the airflow in the cloud-free atmosphere, roughly up to 10 km above the Earth's surface. Optical radars, called lidars, use lasers as the radiation source. At these short wavelengths (0.1–10 μm), the laser beam is scattered by small aerosol particles and air molecules, allowing air motions to be determined, especially in thin, high tropospheric clouds and in the Earth's boundary layer (approximately the lowest 1 km or 0.6 mi of the Earth's atmosphere). *See* AEROSOL; HYDROMETEOROLOGY; LASER; LIDAR.

Operational radars. In the late 1990s, the weather radars used by the National Weather Service to provide warnings of impending severe weather were updated from antiquated WSR-57 and WSR-74 noncoherent radars to NEXRADs (Next Generation Weather Radars). NEXRAD (WSR-88D) radars are state-of-the-art Doppler radars operating at a wavelength of 10 cm. Using NEXRAD's Doppler capability, weather forecasters are able to warn the public sooner of approaching tornadoes and other severe weather. In severe storms, a mesocyclone first develops within the storm. The mesocyclone may be 10 km (6 mi) or more wide, and represents a deep rotating column of air within the storm. Severe and long-lasting tornadoes are often associated with mesocyclones. The mesocyclone is readily detected by a Doppler radar such as NEXRAD. The entire continental United States is covered by the NEXRAD network, consisting of more than 100 radars. NEXRADs along the Gulf Coast, in Florida, and along the eastern seaboard provide warning information on landfalling hurricanes. About 60 of the nation's busiest airports are also equipped with Doppler radars. These radars (operating at a wavelength of 5 cm, known as Terminal Doppler Weather Radars) provide weather-related warnings to air-traffic controllers and pilots. One particularly dangerous weather condition is wind shear, which often occurs as a microburst or intense downdraft. Microbursts can severely affect the flight of landing and departing aircraft, and have been identified as a factor in many aircraft accidents. *See* RADAR; RADAR METEOROLOGY; TORNADO; WEATHER FORECASTING AND PREDICTION.

Steven A. Rutledge

Bibliography. D. Atlas (ed.), *Radar in Meteorology*, 1990; V. N. Bringi and V. Chandrasekar,

Polarimetric Doppler Weather Rada: Principles and Applications, 2005; R. J. Doviak and D. S. Zrnic, *Doppler Radar and Weather Observations*, 2d ed., 1993; P. Meischner (ed.), *Weather Radar: Principles and Advanced Applications*, 2005; R. Meneghini and T. Kozu, *Spaceborne Weather Radar*, 1990; R. E. Rinehart, *Radar For Meteorologists*, 4th ed., 2004; M. I. Skolnik, *Radar Handbook*, 2d ed., 1990.

Meteorological rocket

A small rocket system used for extending observations of the atmosphere above feasible limits for balloon-borne and telemetering instruments. Synoptic exploration of the middle-atmospheric circulation (20–95 km or 12–60 mi altitude) through use of these systems (also known as rocketsondes) matured in the 1960s into a highly productive source of information on atmospheric structure and dynamics. Many thousands of small meteorological rockets have been launched in a coordinated investigation of the wind field and the temperature and ozone structures in the middle atmosphere region at 25–55 km (16–34 mi) altitude.

These data produced dramatic changes in the scientific view of this region of the atmosphere, with a resulting alteration of the structural concepts—which had previously developed without adequate measurements—into a space-age atmospheric model that is primarily characterized by intense dynamics.

The availability of inexpensive, small meteorological rocket systems prompted the idea for simultaneous observations from the many different North American rocket ranges. These simultaneous, or synoptic, observations were scheduled for only a few days each season, but rapidly evolved into three to five launchings a week from many launch ranges. The coordination of these rocket launchings began in 1959 with the formation of the Meteorological Rocket Network (MRN). As other countries (for example, France, Japan, the United Kingdom, and the Soviet Union) developed their particular rocket system and capability, the launch activity became more specific and followed the schedule recommended by recognized scientific organizations such as the Committee on Space Research of the International Council of Scientific Unions.

Meteorological Rocket Network. The Meteorological Rocket Network and other international launch ranges have been very important and extremely effective in obtaining the much-needed temperature and wind data in the middle atmosphere. These data contributed significantly to the knowledge of stratospheric (15–55 km or 10–34 mi) and mesospheric (55–80 km or 34–50 mi) climatology, and they have been used widely in the development of reference atmosphere models. Reference atmospheres have important engineering applications for the design and testing of large launch vehicles such as the space shuttle.

Fig. 1. Transmitter-bead thermistor payload for the Starute.

The development of the small meteorological rocket began in 1959 with a rocketsonde system known as ARCAS. ARCAS permitted a relatively large payload volume to be employed, but its low initial acceleration and slow speed made its trajectory sensitive to low-level winds during launch. The maximum altitude reached by this system was about 60 km (37 mi). A less wind-sensitive rocketsonde system, the Loki-Datasonde (PWN-8B), replaced the ARCAS system during the early 1970s. It was soon replaced with the Super Loki-Datasonde (PWN-11D). The PWN-11D rocketsonde motor burns for 2 s before separation from its inert dart and payload, which are thereby propelled to about 80 km (50 mi) altitude, where the payload is ejected. The payload consists of a small bead thermistor temperature sensor attached to a radio transmitter (**Fig. 1**) that sends the temperature data to a ground receiver, and a Starute parachute. The meteorological measurements are made during payload descent. At launch, the Super Loki-Datasonde (**Fig. 2**) has an overall weight of 31 kg (68 lb), and its length is approximately 4 m (13 ft).

The validity of rocketsonde measurements is highly dependent on the instrument precision. Tests show that the Datasonde measurements have a precision of 1°C (1.8°F) up to 55 km (34 mi) and about 1.5°C (2.7°F) at 60 km (37 mi). Wind speed accuracy is dependent on the quality of the radar tracking system, but it usually is better than 3 m/s (6.6 mi/h).

The total observational system includes a ground station composed of a radar for tracking the metallized Starute and a receiving station for the temperature telemetry data. Again, the equipment of these stations over the Meteorological Rocket Network is highly variable, the best being the precision tracking and telemetry systems found at the missile range.

Because of the increasing cost of rocket motors and payloads, launch equipment, and launch range operations, the Meteorological Rocket Network slowly was dissolved as funds needed to

maintain a viable launch schedule no longer were available. By the latter half of the 1980s, only nine launch ranges remained active. By 1998 only five facilities were still launching small meteorological rockets, but they were no longer using a coordinated schedule. As the number of launch sites was reduced, the coordinated launch schedule was replaced with launch schedules designed to meet unique launch range requirements or to further special research. Meteorological rocket data continue to be needed for verification and calibration check of remotely measuring instruments, that is, satellites and lidar. *See* LIDAR; METEOROLOGICAL INSTRUMENTATION; TELEMETERING.

Results. Meteorological Rocket Network synoptic exploration of the stratospheric circulation revealed an upper atmosphere markedly different from the quiescent, static, drab, and uninteresting place characterized in early models.

A prime example of the synoptic-scale dynamical systems that originate in the stratospheric region is illustrated by the "explosive warmings" first detected in high-altitude balloon flights over Berlin. Explosive warmings are wintertime phenomena, and they appear strongest in the upper stratosphere (30–40 km or 18–24 mi) where large temperature increases of 50°C (90°F) or more occur within a few days. Another phenomenon observed was the 26-month, or quasi-biennial oscillation. This oscillation is noticeable in the east-west wind flow over the Equator. The wind flows from the west for approximately 13 months and then reverses direction and flows from the east for about 13 months. This phenomenon was first observed with balloon soundings; but the rocketsonde observations were, and still are, the only method possible to define its vertical dimension and wind speeds.

Possibly of equal significance at the other end of the turbulence spectrum is the fact that the sensitive sensors of the rocketsonde have revealed a tremendous amount of small-scale variability in the upper atmosphere. These data make clear that assumptions of an inviscous fluid in the atmosphere are highly suspect; the values of the viscous terms in the equations of motion are probably larger than assumed in prior analysis.

Observational data have shown that the small-scale variability is largest in high latitudes during the winter. Temperature variations larger than 10°C (18°F) have been observed to occur within 2–4 h. The source of these extreme perturbations still is vague, but evidence points to gravity-wave activity.

Synoptic-scale circulation systems in the upper atmosphere are demonstrated by the rocketsonde data to be very obviously keyed to the geographic and orographic structures of the Earth's surface. In winter, oceanic regions characteristically have poleward extensions of ridges of high pressure, and continental regions have shifty troughs of low pressure extending equatorward over them. This intimate relationship between the surface and 50 km (31 mi) is most likely the direct result of turbulent energy transport in the vertical direction, and a total understanding of the entire atmospheric system cannot be realized until these factors are incorporated. *See* ATMOSPHERIC GENERAL CIRCULATION; STRATOSPHERE; UPPER-ATMOSPHERE DYNAMICS.

Willis L. Webb; Francis J. Schmidlin

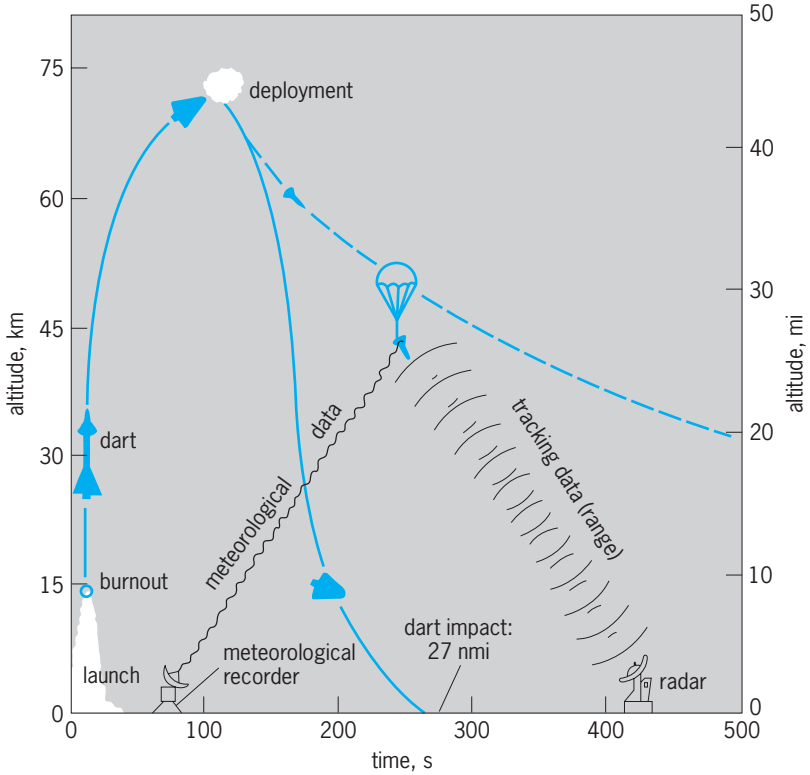

Fig. 2. Diagram of typical rocketsonde launch. Booster impact is within 400 m (1300 ft) under normal wind conditions. (*NASA*)

Bibliography. J. M. Moran and M. D. Moran, *Meteorology: The Atmosphere and the Science of Weather*, 4th ed., 1994; F. J. Schmidlin, Intercomparisons of temperature, density, and wind measurements from *in situ* and satellite techniques. *Adv. Space Res.*, 4(6):101–110, 1984; F. J. Schmidlin, Rocket techniques used to measure the middle atmosphere, in Middle Atmosphere Program, *Handbook for MAP*, vol. 19, 1986; World Data Center A (Asheville, NC), *Meteorological Rocket Network Firings*, vols. 1–97.

Meteorological satellites

Satellites dedicated to the observation of meteorological phenomena and atmospheric or surface properties used for weather forecasting. Operational meteorological satellites provide routine observations of weather conditions as well as an ever-expanding range of environmental properties, such as aerosol, dust and ash clouds from volcanic eruptions, ozone, and land vegetation cover. For this reason, they are known in the United States as operational environmental satellites. *See* METEOROLOGY; SATELLITE (SPACECRAFT); SATELLITE METEOROLOGY.

The United States launched the first meteorological satellite in low Earth orbit, *TIROS 1* (*Television and Infra-Red Observation Satellite*), in April 1960. For the first time, complete pictures of clouds associated with large weather systems were seen by forecasters. A total of 10 *TIROS* were launched, followed by a series of progressively more sophisticated *ESSA* (Environmental Science Services Administration) and *NOAA* (National Oceanic and Atmospheric Administration) satellites, while the National Aeronautics and Space Administration (NASA) was experimenting with new sensors on a succession of *Nimbus* research satellites. In parallel, the Department of Defense initiated the Defense Meteorological Satellite Program (DMSP) used by the Air Force and the Navy for global weather surveillance and forecasting applications. At the same time the Soviet Union implemented its national operational *Meteor* satellite series, which continues to this day. Current plans in the United States are to merge both civilian and military low-altitude meteorological satellite programs into a National Polar-orbiting Operational Environmental Satellite System (NPOESS), beginning around 2008. Altogether, three *NPOESS* Satellites will likely occupy three polar orbits staggered in time to provide global coverage at approximately 4-hour intervals. *See* STORM DETECTION; WEATHER FORECASTING AND PREDICTION.

NASA pioneered Earth observation from geostationary orbit on its first experimental *Application Technology Satellite* launched in 1966, soon to be followed by the operational *SMS* (*Synchronous Meteorological Satellite*) and *GOES* (*Geostationary Operational Environmental Satellite*) series occupying two stations at 75° and 135° west longitude. Similarly, the European Space Agency and Japan National Space Development Agency developed their own equally capable geostationary satellite systems, *Meteosat* operated by EUMETSAT at 0° longitude and *Himawari* (Chrysanthemum) operated by the Japan Meteorological Agency at 140° east longitude, respectively. India and later China have attempted to fill the remaining gap over Asia and the Indian Ocean with their *Insat* and *FY* geostationary satellites.

Optical imaging sensors. Historically, the first recognized application of orbital observation was the visual exploitation of cloud images associated with weather systems. The preferred design of their imaging instruments emphasized spatial resolution and contrast over radiometric accuracy, as opposed to sensors that provide quantitative measurements. The difference in instrument design is beginning to blur in recent instruments, such as the Advanced Very High Resolution Radiometer (AVHRR) on *NOAA* satellites, which is used for a variety of quantitative applications, such as remote sensing of sea surface temperature, monitoring changes in land vegetation, and discriminating between different kinds of clouds. There is a pervasive trend to increase the number of spectral bands in imaging sensors, from 5 channels in the current AVHRR to 36 channels in the experimental Moderate-resolution Imaging Spectroradiometer (MODIS) developed by NASA. These channels sample the full spectrum of backscattered solar radiation in the visible, near-infrared, and long-wave infrared, and a good part of the emitted terrestrial radiation spectrum (thermal infrared). This multiplicity of spectral bands allows the detection of a wide variety of features, from aerosols and smoke in the atmosphere to chlorophyll in the ocean. *See* CLOUD; REMOTE SENSING; TERRESTRIAL RADIATION; WEATHER.

Except for observing polar regions, or providing meteorological support to operations in remote locations worldwide, the ideal platforms for cloud imaging are those in geosynchronous equatorial orbit, also known as geostationary orbit, at the precise altitude (35,900 km) where the orbital period matches the period of rotation of the Earth, so that the satellite appears to hover over a fixed location at the Equator. The international system of four to six geostationary meteorological satellites provides uninterrupted visibility of the global tropics and midlatitudes (up to 60° north and south at the satellite longitude) with the ability to monitor fast-developing weather systems that often are the most dangerous. The sharpness of cloud images (1-km picture elements in the visible), as well as the ability to scan the same scene repeatedly at time intervals as short as 5 minutes, allow for tracking the apparent motion of clouds, deducing wind velocity, and instantaneously assessing the strength of developing storms, a valuable capability in warm climate regions (**Fig. 1**). *See* EARTH ROTATION AND ORBITAL MOTION; TROPICAL METEOROLOGY.

Imaging from geostationary satellites is the principal source of observation that has allowed the National Weather Service to monitor the evolution of strong convective storms and extend the warning time for tornado strikes to about 8 minutes. In general, geostationary satellites provide meteorologists with a view from an ideal vantage point, a perfect synoptic coverage of regional weather, and the ability to continuously track the progression of weather disturbances and extrapolate their arrival time at a particular location.

Imaging microwave radiometers. Also interesting is the detection of diverse atmospheric properties and surface features using multifrequency microwave radiometers with small antenna beams. Water molecule absorption of microwave radiation emitted by the ocean provides an accurate estimation of total precipitable water in the atmospheric column. Microwave radiation emitted by the relatively homogeneous moist atmosphere below is scattered in a recognizable way by waterdrops and ice particles in rain clouds, thus providing an indirect means to estimate precipitation rates. Microwave radiation contrast discriminates ice floes from open ocean water, and wet from dry soil. Microwave radiometry enables diagnostics of sea state and wind strength over the surface of the ocean, or the sea surface temperature. The principal design constraint of imaging microwave radiometers is the diffraction limit of the sensor—large apertures are desirable, but

bulky antennas are a problem because mechanical scanning is needed to preserve radiometric accuracy. In order to achieve reasonably small footprints, microwave sensors are currently deployed in low Earth orbit, such as the Special Instrument/Microwave Imager on *DMSP* satellites and Conical Microwave Imager/Sounder on future *NPOESS* satellites. *See* MICROWAVE; PRECIPITATION (METEOROLOGY); RADIOMETRY.

Sounding sensors. The principal challenge of meteorological observation from space is obtaining quantitative information about atmospheric temperature and pressure, and the concentration of water vapor and other minor constituents, especially the vertical profile of these quantities in the lower 10–20 km.

Radiation emitted by the atmosphere at a given wavelength originates principally from the layer where the optical depth of the overlying atmosphere is equal to unity, meaning that a photon emitted from that region has about one chance out of three to escape. The explanation is that photons emitted much lower in the atmosphere are mostly reabsorbed and few escape to space, while photons emitted much higher in the atmosphere are few because the density of the emitting material decreases exponentially with altitude (**Fig. 2**). The pressure level where the optical depth equals one is determined by the mass of the overlying atmosphere and its specific absorption for the particular wavelength—the stronger the absorption, the higher the level. Thus, the spectrum of terrestrial radiation received by a satellite contains information on thermal emission of the atmosphere at different levels, determined by atmospheric temperature and absorption at different wavelength. *See* ATMOSPHERE; PHOTON.

The retrieval of temperature profile and water vapor information from spectral data is a difficult and not a fully determined mathematical problem. The solutions are highly sensitive to spectral resolution and small errors in radiometric measurements. Nevertheless, considerable progress has been made since the early days of the Nimbus program. The latest Atmospheric Infra-Red Sounder (AIRS) instrument developed by NASA is expected to yield temperature profiles as accurate as balloon measurements, 1°C within each successive 1-km-thick layer of the lower atmosphere. *See* HYDROMETEOROLOGY; INFRARED RADIATION.

Atmospheric sounders operate in the thermal infrared, using the absorption bands of carbon dioxide molecules (3.7–4.9 μm and 13–15 μm), and in the microwave spectrum, using the 54-GHz absorption band of oxygen. Emitted radiation is much weaker and atmospheric sounders correspondingly less sensitive in the microwave region. However, nonprecipitating clouds are largely transparent to such relatively long wavelengths, thus allowing all-weather albeit less accurate observations.

Measurements of temperature and moisture are used mainly to update numerical weather prediction computations that forecast the circulation of the global atmosphere several days in advance. For this quantitative application, a delay of a few hours is immaterial but homogeneous global coverage is essential. Thus, atmospheric sounders are principally deployed on Sun-synchronous polar orbits. The parameters of these circular low Earth orbits are selected from a discrete set of altitudes (800–1000 km) and inclinations (retrograde quasi-polar) that allow

Fig. 1. On May 17, 2000, the first visible image was broadcast by the NOAA *Geostationary Operational Environmental Satellite* (*GOES 11*) launched 2 weeks earlier. A huge storm is visible over the central United States. It dropped 30 cm of snow in Wyoming and 5 cm of hail in Colorado and spawned many tornadoes across Nebraska.

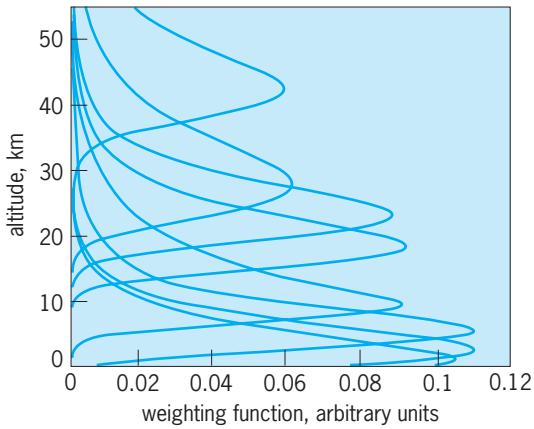

Fig. 2. Outgoing terrestrial radiation observed in space originates from different altitudes in the atmosphere. Depending upon the wavelength, the atmosphere is more or less absorbing. Emitted infrared radiation that is most effectively absorbed emerges only from the uppermost range of altitudes. Shown is a sample of altitude probability distributions or weighting functions for different wavelengths that apply to the advanced Atmospheric Infra-Red Sounder (AIRS) developed by NASA.

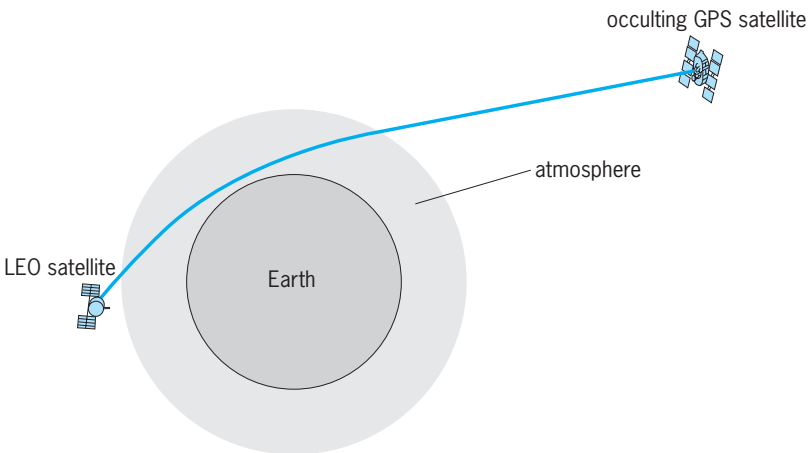

Fig. 3. Air density, pressure, and temperature data are derived from measurements of the increasing microwave refraction index of the atmosphere as altitude above the Earth surface decreases. The index of refraction is deduced from the propagation delay of radio-electric signals between any member of the Global Positioning System (GPS) satellites and a low-altitude orbiting satellite. The effect is significant near the moment of occultation, when the path is close to the limb of the Earth.

the orbital plane to drift by about 1° of longitude per day and match the change in Sun-Earth direction. Thus, a Sun-synchronous satellite crosses the Equator at (nearly) the same local time on every successive orbit.

Active sensors. Satellite sensors are not limited to receiving reflected solar radiation or emission from the atmosphere and clouds. Orbital systems are now powerful enough to probe the atmospheric medium or the surface with beams of electromagnetic radiation generated in space. The first operational sensor of this kind was a coarse radar or scatterometer that measured microwave radiation backscattered by the ocean surface. Backscatter is sensitive to surface roughness and thus provides a measurement of vector wind speed over the ocean (as well as a coarse all-weather mapping of sea ice). NASA developed the NSCAT and Seawinds microwave scatterometers for deployment on experimental Japanese environmental satellites *ADEOS 1* and *2*. A similar operational instrument will be carried by European *METOP* satellites.

Various radar altimeters have been used to map the changing topography of the ocean surface (principally to reconstruct the oceanic circulation from measured altitude gradients). Higher-frequency experimental radar and lidar systems are being tested to profile the distribution and optical properties of aerosol and cloud ice particles and waterdrops. The first demonstration of a space-borne precipitation radar in being conducted with the United States-Japan *Tropical Rain Measuring Mission (TRMM)* launched in 1997. Rain rate can be deduced from the three-dimensional distribution of ice particles and water in rain clouds, as observed by the *TRMM* satellite. *See* LIDAR; METEOROLOGICAL RADAR; RADAR METEOROLOGY.

Yet another promising technology will determine wind velocity in clear air from direct measurements of the frequency shift (Doppler effect) of multiple laser pulses backscattered by aerosol and other diffusive particles. NASA planned to conduct the first flight demonstration of this technique on the space shuttle in year 2001. Global wind measurements will provide an invaluable enhancement of the worldwide meteorological observing network, especially at low latitudes where the wind field cannot be deduced from atmospheric pressure. *See* DOPPLER RADAR.

Bistatic systems. Planetary fly-by missions have provided opportunities to deduce the atmospheric density structure by measuring the propagation delay of a radioelectric signal traversing increasingly longer paths through a planet's atmosphere. Exceedingly precise radioelectric occultation measurements of this type are also possible on Earth by monitoring the propagation delay of navigation signals transmitted by the constellation of Global Positioning System (GPS) satellites and received by low-altitude orbiting satellites whenever the propagation path comes near the Earth limb (apparent outer edge) [**Fig. 3**]. The first GPS-MET flight demonstration, sponsored by the National Science Foundation, accurately retrieved atmospheric density, pressure, and temperature from about 25 km down to about 5 km above the Earth's surface. NASA will continue to test this new observing method with an experimental constellation of GPS-equipped satellites. Future operational satellites such as *NPOESS* and *METOP* will carry advanced versions of these GPS receivers. *See* OCCULTATION; SATELLITE NAVIGATION SYSTEMS.

In the lower atmosphere, the refractive effect of unknown amounts of water vapor invalidates the radioelectric method, unless humidity can be otherwise determined. Alternatively, the method can be used to determine total precipitable water along the propagation path when dry air pressure is known. This meteorological application is being developed for a number of GPS-equipped geodetic stations already in existence to monitor Earth crust deformations.

Pierre Morel

Bibliography. Committee on Earth Observation Satellites, *Towards an Integrated Global Observing Strategy*, European Space Agency, 1997; S. Q. Kidder and T. H. Von der Haar, *Satellite Meteorology: An Introduction*, 1995; W. P. Menzel and J. F. W. Purdom, Introducing GOES-1, The first of a new generation of Geostationary Operational Environmental Satellites, *Bull. Amer. Meteorol. Soc.*, 75:757–781, 1994; National Research Council, *Continuity of NOAA Satellites*, National Academy Press, 1997; P. K. Rao (ed.), *Weather Satellites: Systems, Data and Environmental Applications*, American Meteorological Society, Boston, 1990.

Meteorology

A discipline involving the study of the atmosphere and its phenomena. Meteorology and climatology are rooted in different parent disciplines, the former in physics and the latter in physical geography. They

have, in effect, become interwoven to form a single discipline known as the atmospheric sciences, which is devoted to the understanding and prediction of the evolution of planetary atmospheres and the broad range of phenomena that occur within them. The atmospheric sciences comprise a number of interrelated subdisciplines. *See* CLIMATOLOGY.

Subdisciplines. Atmospheric dynamics (or dynamic meteorology) is concerned with the analysis and interpretation of the three-dimensional, time-varying, macroscale motion field. It is a branch of fluid dynamics, specialized to deal with atmospheric motion systems on scales ranging from the dimensions of clouds up to the scale of the planet itself. The activity within dynamic meteorology that is focused on the description and interpretation of large-scale (greater than 1000 km or 600 mi) tropospheric motion systems such as extratropical cyclones has traditionally been referred to as synoptic meteorology, and that devoted to mesoscale (10–1000 km or 6–600 mi) weather systems such as severe thunderstorm complexes is referred to as mesometeorology. Both synoptic meteorology and mesometeorology are concerned with phenomena of interest in weather forecasting, the former on the day-to-day time scale and the latter on the time scale of minutes to hours. *See* DYNAMIC METEOROLOGY; MESOMETEOROLOGY.

The complementary field of atmospheric physics (or physical meteorology) is concerned with a wide range of processes that are capable of altering the physical properties and the chemical composition of air parcels as they move through the atmosphere. It may be viewed as a branch of physics or chemistry, specializing in processes that are of particular importance within planetary atmospheres. Overlapping subfields within atmospheric physics include cloud physics, which is concerned with the origins, morphology, growth, electrification, and the optical and chemical properties of the droplets within clouds; radiative transfer, which is concerned with the absorption, emission, and scattering of solar and terrestrial radiation by aerosols and radiatively active trace gases within planetary atmospheres; atmospheric chemistry, which deals with a wide range of gas-phase and heterogeneous (that is, involving aerosols or cloud droplets) chemical and photochemical reactions on space scales ranging from individual smokestacks to the global ozone layer; and boundary-layer meteorology or micrometeorology, which is concerned with the vertical transfer of water vapor and other trace constituents, as well as heat and momentum across the interface between the atmosphere and the underlying surfaces and their redistribution within the lowest kilometer of the atmosphere by motions on scales too small to resolve explicitly in global models. Aeronomy is concerned with physical processes in the upper atmosphere (above the 50-km or 30-mi level). *See* AERONOMY; ATMOSPHERIC CHEMISTRY; ATMOSPHERIC ELECTRICITY; ATMOSPHERIC GENERAL CIRCULATION; ATMOSPHERIC WAVES, UPPER SYNOPTIC; CLOUD PHYSICS; METEOROLOGICAL OPTICS; MICROMETEOROLOGY; RADIATIVE TRANSFER; TERRESTRIAL RADIATION.

Although atmospheric dynamics and atmospheric physics in some circumstances can be successfully pursued as separate disciplines, important problems such as the development of numerical weather prediction models and the understanding of the global climate system require a synthesis. Physical processes such as radiative transfer and the condensation of water vapor onto cloud droplets are ultimately responsible for the temperature gradients that drive atmospheric motions, and the motion field, in turn, determines the evolving, three-dimensional setting in which the physical processes take place.

The atmospheric sciences cannot be completely isolated from related disciplines. On time scales longer than a month, the evolution of the state of the atmosphere is influenced by dynamic and thermodynamic interactions with the other elements of the climate system, that is, the oceans, the cryosphere, and the terrestrial biosphere. A notable example is the El Niño–Southern Oscillation phenomenon in the equatorial Pacific Ocean, in which changes in the distribution of surface winds force anomalous ocean currents; the currents can alter the distribution of sea-surface temperature, which in turn can alter the distribution of tropical rainfall, thereby inducing further changes in the surface wind field. On a time scale of decades or longer, the cycling of chemical species such as carbon, nitrogen, and sulfur between these same global reservoirs also influences the evolution of the climate system. Human activities represent an increasingly significant atmospheric source of some of the radiatively active trace gases that play a role in regulating the temperature of the Earth. *See* BIOSPHERE; MARITIME METEOROLOGY; TROPICAL METEOROLOGY.

Vertical profiles of pressure and density. The decrease with height in the pressure (p) and density (ρ) in planetary atmospheres is approximately exponential (**Fig. 1**). This behavior is characteristic of atmospheres composed of ideal gases, in which absolute temperature does not vary strongly with height. If it is assumed that the vertical component of the acceleration of the air is small in comparison to gravity (g) and that the weight of a differential layer of unit horizontal cross section and thickness (dz) in the vertical is equal to the difference between the pressure on its upper and lower surfaces, the hydrostatic equation (1) results. This relationship is valid

$$\frac{dp}{dz} = -\rho g \tag{1}$$

throughout the atmosphere except in the most vigorous convective clouds. Substituting for ρ from the equation of state (2) yields Eq. (3), where R is the

$$p = \rho RT \tag{2}$$

$$d(\ln p) = -\frac{g}{RT}\, dz \tag{3}$$

gas constant appropriate to the chemical composition of the atmosphere and T is the temperature in

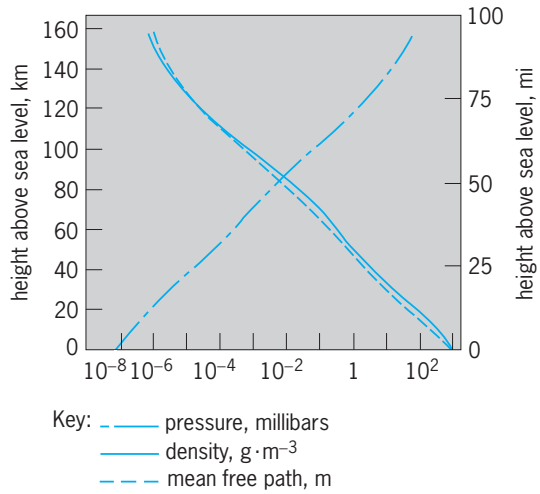

Fig. 1. Vertical profile of pressure , density, and mean free path for typical conditions in the Earth's atmosphere. 1 millibar = 10^2 Pa. 1 m = 3. 28 ft. (*After R. C. Weast, ed., CRC Handbook of Chemistry and Physics, 70th ed., CRC Press, 1989*)

kelvins. If T is assumed to be independent of height, this expression can be integrated from a reference level z_0, at which $p = p_0$, to level z and is expressed in the form shown in Eq. (4), where $H \equiv (RT)/g$

$$p = p_0 \exp -\frac{z}{H} \tag{4}$$

is known as the scale height, that is, the height over which pressure decreases by a factor of e (the base of natural logarithms). By combining Eqs. (2) and (4), it is readily verified that density exhibits the same functional dependence. The gas constant R is obtained by dividing the universal gas constant $R^* = 8313\ \mathrm{J \cdot kmol^{-1} \cdot K^{-1}}$ by the molecular effective weight μ of the mixture of gases of which the atmosphere is composed. For Earth, up to a level of 120 km (72 mi), $\mu = 28.97$ as explained below, $R = 287\ \mathrm{J \cdot kg^{-1} \cdot K^{-1}}$, $g = 9.8\ \mathrm{m \cdot s^{-1}}$, and $T = 240$ K (to within $\pm$ 15%), and so, consistent with the slopes of the pressure and density curves in Fig. 1, $H \cong 7$ km (4 mi). Mean sea-level pressure averaged over the Earth is 1013 millibars (101 kilopascals), and the mean density is on the order of $1.25\ \mathrm{kg \cdot m^{-3}}$. In relating pressure to height in the Earth's atmosphere in the remainder of this section, it is convenient to note that pressure and density drop off by a factor of 10 over a vertical distance of $H \ln(10) \cong 16$ km; and so the 100-mbar (10-kPa) level corresponds to roughly 16 km (10 mi) above sea level, the 10-mbar (1-kPa) level to 32 km (20 mi), the 1.0-mbar (100-Pa) level to 48 km (30 mi), and so forth. Since the pressure at any level is equal to the mass per unit area above that level, divided by g, it follows that 10% of the mass of the Earth's atmosphere lies above the 100-mbar (10-kPa) level, 1% above the 10-mbar (1-kPa) level, and so forth. *See* GAS.

Composition of atmosphere. The atmospheres of the planets are believed to have originated from the outgassing of volatile substances [mainly water (H_2O), carbon dioxide (CO_2), and nitrogen compounds] from their interiors. Most of the water vapor condensed out immediately. On Earth (in contrast to Venus and Mars) nearly all the carbon dioxide has dissolved in the oceans and subsequently been incorporated into carbonate deposits in the crust by shell-forming species of plankton. Most of the nitrogen remains in the atmosphere in the form of molecular nitrogen (dinitrogen; N_2). Photosynthesis by plant life that was buried and fossilized before it had time to decay has generated large quantities of molecular oxygen (dioxygen; O_2).

From the surface up to about 100 km (60 mi), macroscale fluid motions keep the atmosphere well mixed, so that the relative proportions of its major gaseous constituents are nearly constant. Nitrogen, oxygen, and argon are the dominant constituents of dry air: they account for roughly 78%, 21%, and 1% of the molecules, respectively (hence the effective molecular weight of 28.97). Important atmospheric trace constituents include water vapor (up to 4% by volume and highly variable in space and time), ozone (O_3; up to 15 parts per million at the 25-km level), and carbon dioxide (approximately 350 ppm in 1990 and increasing at a rate of about 12 ppm per decade). The latter are triatomic species, which exhibit strong absorption bands associated with rotational–vibrational transitions in the infrared part of the electromagnetic spectrum, the implications of which will be discussed below.

At 100 km (60 mi) above the Earth's surface, the mean free path between collisions reaches 1 m (3.3 ft; Fig. 1) and the characteristic time scale for molecular diffusion becomes comparable to the time scale for mixing by macroscale fluid motions. Above this so-called turbopause, the various atmospheric constituents exhibit gravitational settling under the influence of molecular diffusion, the concentration of each species decreasing exponentially, as shown in Eq. (4), with scale height H inversely proportional to its molecular weight, as if it were the only constituent present. At 500 km (300 mi) the atmosphere is composed primarily of atomic oxygen (O), and above 1000 km (600 mi) helium and hydrogen are the dominant species.

Effects of x-ray and ultraviolet radiation. In **Fig. 2** the marked enhancements of the spectrum of solar radiation relative to the blackbody curve at the very short and very long wavelengths are due to the emission from the solar chromosphere and corona. Ultraviolet radiation with wavelengths (λ) shorter than 0.31 micrometer accounts for less than 2% of the energy emitted by the Sun; radiation with $\lambda < 0.24\ \mu$m, less than 0.1%; and x-rays with $\lambda < 0.1\ \mu$m, only about 3 parts in 10^6—yet the radiation in these wavelength bands has profound influences on the structure and composition of the upper atmosphere.

Several types of photochemical reactions occur within the Earth's atmosphere (**Fig. 3**). X-rays with $\lambda < 0.1\ \mu$m are sufficiently energetic to ionize whatever atomic species happen to be present. Peak ion concentrations (about $10^6\ \mathrm{cm^{-3}}$) are observed about 300 km (180 mi) above the Earth's surface. Virtually all the solar radiation in this wavelength band is

absorbed before it reaches the 60-km (36-mi) level (the base of the ionosphere), by which time ion concentrations have declined by four orders of magnitude relative to those at 300 km (180 mi). *See* IONOSPHERE; PHOTOCHEMISTRY.

Ultraviolet radiation with $\lambda < 0.24\ \mu m$ is sufficiently energetic to photodissociate molecular oxygen. As a consequence of this reaction, most of the oxygen is in the atomic form at levels above 120 km (72 mi). Radiation at these wavelengths is sufficiently depleted by the time it penetrates to the 20–50-km (12–30-mi) level that it produces only trace amounts of atomic oxygen. At these levels the air is sufficiently dense that atomic oxygen quickly combines with molecular oxygen in three-body reaction (5), where

$$O + O_2 + M = O_3 + M \qquad (5)$$

M represents another molecule, to form ozone. The ozone molecules created in this reaction are dissociated by ultraviolet solar radiation with wavelengths $<0.31\ \mu m$, as in reaction (6). The resulting oxygen

$$O_3 + h\nu = O_2 + O \qquad (6)$$

atom quickly recombines with O_2 [reaction (5)].

The net result of reactions (5) and (6) is the absorption of a photon of solar radiation with $0.24 < \lambda < 0.31\ \mu m$, and the heating of the other molecules (M) involved in the three-body collision [reaction (5)]. Through this mechanism, the creation of a single odd oxygen molecule can ultimately result in the absorption of millions of photons of ultraviolet radiation which would otherwise be lethal to life on the surface of the planet. The photochemical reactions associated with ozone chemistry are most active in the 30–60-km (18–36-mi) layer, the so-called ozone layer. Under certain conditions, chlorine radicals created by the photodissociation of anthropogenically produced chlorofluoromethanes at these levels appear to be capable of destroying odd oxygen through catalytic reactions. *See* STRATOSPHERIC OZONE.

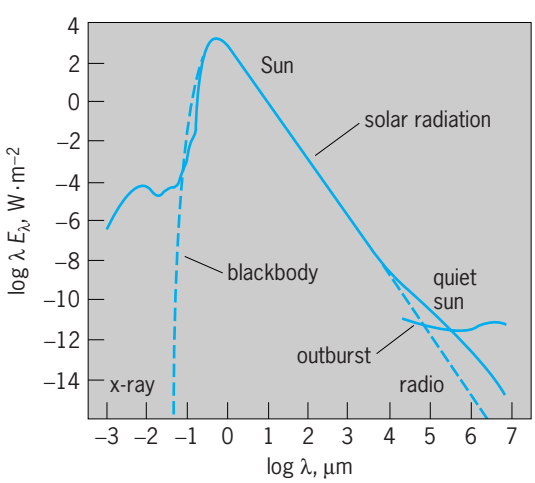

Fig. 2. Spectrum of solar radiation as compared with the spectrum of a blackbody at 5780 K (9934°F). (*After C. W. Allen, Solar radiation, Quart. J. Roy. Meteorol. Soc., 84:311, 1958*)

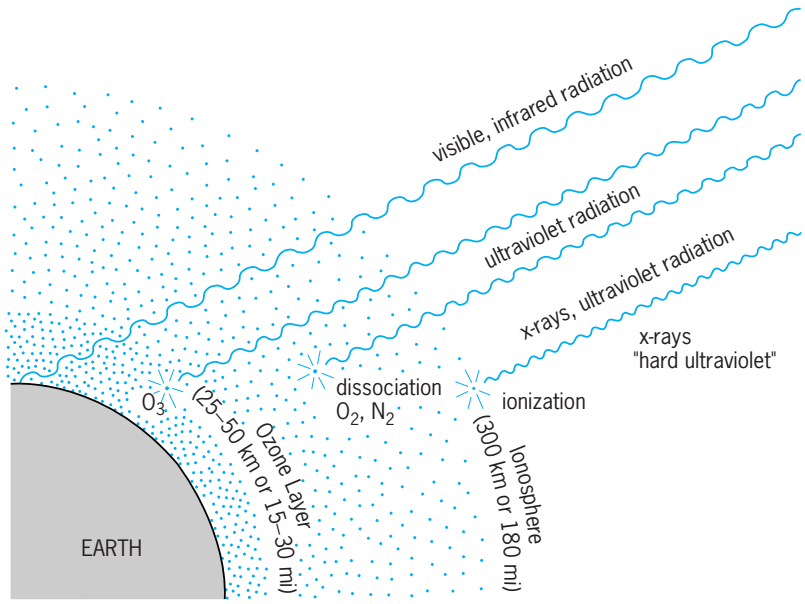

Fig. 3. Schematic representation of the processes responsible for the absorption of solar radiation in the Earth's atmosphere.

Photoionization and photodissociation are strong heat sources above the 100-km (60-mi) level. Even though only a few parts per million of the solar radiation incident on the top of the atmosphere is absorbed at these levels, the energy input is large because the molecules present at these levels account for less than one-millionth of the mass of the atmosphere. Energy absorption per unit mass increases with height in this layer and, in addition, the molecules become less efficient at disposing of energy by emitting infrared radiation as the frequency of molecular collisions decreases. Hence, above about 80 km (50 mi), temperature increases with height, and the downward flux of heat by down-gradient molecular diffusion plays an important role in the energy balance. This outermost layer of the Earth's atmosphere is known as the thermosphere. Analogous layers are observed in the atmospheres of the other planets. Solar output in the x-ray part of the spectrum varies strongly in response to sunspots and solar flares, whose frequency exhibits a remarkable 11-year cycle. During the active part of the cycle, temperatures in the outer thermosphere reach values of 2000 K (3140°F; compared to about 500 K or 440°F in the quiet sun years), and a significant fraction of the hydrogen atoms attain velocities high enough to allow them to escape from the Earth's gravitational field. Over the lifetime of the solar system, appreciable quantities of hydrogen are believed to have escaped from the Earth's atmosphere, and it has been proposed that the hydrogen atoms in the water outgassed from Venus might have been lost in this manner. *See* SOLAR SYSTEM; SUN.

The absorption of ultraviolet radiation by ozone molecules in the 30–60-km (15–36-mi) layer gives rise to a distinct peak in the vertical profile of temperature (**Fig. 4**). It is notable that the atmospheres of Mars and Venus, which lack the oxygen necessary

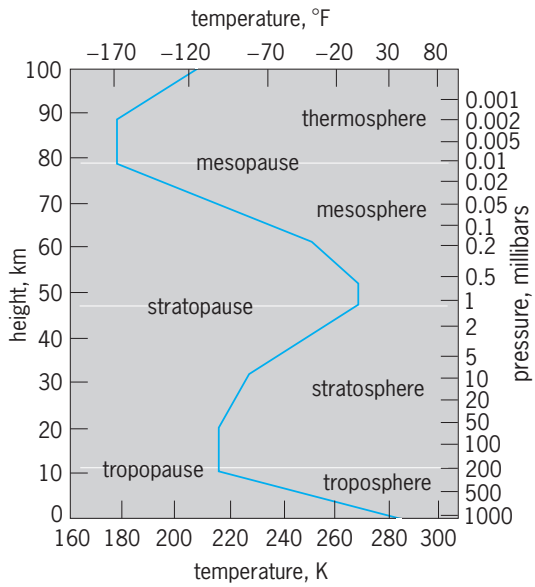

Fig. 4. Typical temperature profile in middle latitudes. 1 millibar = 10^2 Pa. 1 km = 0. 6 mi. (*After J. M. Wallace and P. V. Hobbs, The Atmosphere: An Introductory Survey, Academic Press, 1977*)

to form an ozone layer, do not exhibit such an intermediate temperature maximum between the surface and the thermosphere. The peak in the temperature profile divides the Earth's middle atmosphere into an upper layer (the mesosphere) and a lower layer (the stratosphere). The temperature maximum corresponds to the stratopause, that is, the top of the stratosphere. *See* MESOSPHERE; STRATOSPHERE.

The base of the stratosphere corresponds to a distinct temperature minimum that marks the transition to a lower layer of the atmosphere in which the temperature distribution is maintained by the upward flux of energy from the Earth's surface by macroscale motions. Analogous layers, known as tropospheres, are observed on Venus and Mars. The temperature minimum in the Earth's atmosphere corresponds to the tropopause (that is, the top of the troposphere). In the atmospheres of Venus and Mars a deep isothermal (constant-temperature) layer extends from the tropopause to the base of the thermosphere. *See* TROPOPAUSE; TROPOSPHERE.

Electrical properties. Among the processes that contribute to the production of charged particles in the atmosphere are photoionization of neutral atoms by solar x-rays and ultraviolet rays, and charge separation that takes place when ice particles (or ice and water particles) collide within clouds. The former process gives rise to the ionosphere, which extends upward from the 60-km (36-mi) level to the outermost reaches of the atmosphere, and the latter is responsible for maintaining the fair weather electric field, which exists within the lowest few kilometers above the Earth's surface, and the much stronger fields that sometimes exist locally within and near clouds. Although only a minute fraction of the atoms that make up the Earth's atmosphere are electrically charged, their presence accounts for a wide range of geophysical phenomena.

Ion concentrations increase monotonically with height from the base of the ionosphere to a maximum near 300 km (180 mi). The increase with height tends to be concentrated in a series of layers, labeled (in order of increasing height) D, E, and F. The lower layers are present only during the daytime: in the absence of the Sun's ionizing radiation, the charged particles quickly recombine to form neutral particles. Higher in the ionosphere, recombination is slower because the mean free path between collisions is much longer. Collisions between electrons and neutral particles within the D layer are effective at absorbing AM radio waves propagating upward. When the D layer disappears during the night, radio waves are free to propagate into the upper layers of the ionosphere, where they are reflected to the ground, causing interference and sometimes permitting the reception of distant stations.

The fair weather electric field is strongest in the lowest 100 m (330 ft) of the atmosphere, where it averages 120 V/m (36 V/ft) in the vertical: the atmosphere carries a positive charge relative to the Earth's surface. The Earth and its atmosphere may be viewed as a capacitor whose inner conductor is the Earth and whose outer conductor, which encompasses most of the atmosphere, is referred to as the electrosphere. The conductivity within the electrosphere below 60 km (36 mi) is primarily due to the presence of charged particles generated by cosmic rays colliding with air molecules. Relatively few charged particles are generated within the lowest few kilometers of the atmosphere, and they tend to be immobilized by the presence of large, slow-moving particles. This poorly conducting layer serves as the dielectric of the capacitor. The upward flux of electrons through this "leaky dielectric" would discharge the electrosphere within a matter of minutes were it not continually being recharged by lightning and point discharge currents from the ground within thunderstorms, whose cloud bases carry a strong negative charge relative to the ground because of the charge separation going on within them. *See* SFERICS; STORM ELECTRICITY.

Lapse rates and vertical mixing. The various temperature layers (Fig. 4) have distinctly different dynamical properties, which in turn affect the physical and chemical processes that take place within them. In order to understand why the thermal structure discussed in the previous section is so important, it is necessary to understand the concept of static stability. An idealized air parcel is free to expand and contract in response to the hydrostatic pressure changes [as defined in Eq. (1)] that it encounters as it goes up and down in the atmosphere, and it is thus capable of doing work on its environment as it rises and expands, or having work done on it as it sinks and is compressed; but it is assumed that this air parcel does not exchange heat with its environment. If the parcel does not have any heat source or sink of its own, its temperature will change adiabatically (that is, without the addition or subtraction of heat) as it rises and sinks. The rate of temperature change with height (dT/dz) can be inferred from the first law of thermodynamics, which for an ideal gas

can be written in the form shown in Eq. (7), where

$$dq = c_v\, dT + p\, d\alpha \tag{7}$$

dq is the differential amount of heat added to the parcel, $c_v\, dT$ is the increase in the internal energy of the parcel, $p\, d\alpha$ is the work done by the parcel on the surrounding air as it expands, c_v is the specific heat of air at constant volume (717 J · kg · K^{-1}), and $\alpha = 1/\rho$ is the specific volume of the parcel expressed in m^3 · kg^{-1}. The second term on the right-hand side can be rewritten as $d(p\alpha) - \alpha\, dp$. Substituting from Eq. (2) for $d(p\alpha)$, Eq. (7) can be rewritten as Eq. (8).

$$dq = (c_v + R)\, dT - \alpha\, dp \tag{8}$$

For an isobaric (constant-pressure) process, $dp = 0$, and therefore the specific heat at constant pressure can be defined by expression (9), which is equal to

$$c_p \equiv c_v + R \tag{9}$$

717 + 287 = 1004 J · kg^{-1} · K^{-1}. Substituting into Eq. (8) gives Eq. (10), which is the form of the first

$$dq = c_p\, dT - \alpha\, dp \tag{10}$$

law more frequently used in atmospheric thermodynamics. The first term on the right-hand side of Eq. (10) is referred to as an incremental change in enthalpy or sensible heat. Equation (11) results from

$$-\left(\frac{dT}{dz}\right)_{\text{adiab.}} = \frac{g}{c_p} \tag{11}$$

setting $dq = 0$ and substituting for α from Eq. (1), and this is defined as the dry adiabatic lapse rate, that is, the rate at which the temperature of an air parcel that is not saturated with water vapor drops with increasing height under adiabatic conditions. (Note that the minus sign is implicit in the definition of the term lapse rate.) In the Earth's atmosphere, the dry adiabatic lapse rate has a numerical value of 9.8 K · km^{-1}. *See* ENTHALPY; THERMODYNAMIC PRINCIPLES.

If the environmental lapse rate is the same as the adiabatic lapse rate, an air parcel forcibly displaced upward or downward from its original level would always remain at the same temperature as its environment, and so it would be neutrally buoyant. However, if the environmental lapse rate is less than the adiabatic lapse rate, a rising air parcel will find itself colder (and therefore more dense) than its environment (and vice versa), and so it will encounter a restoring force that will push it back toward its equilibrium level. The larger the difference in lapse rates, the larger the restoring force for a given vertical displacement. The restoring force per unit vertical displacement can be regarded as a measure of the static stability of the atmosphere at that level. Unstable lapse rates (that is, environmental lapse rates larger than the dry adiabatic value) are only very rarely observed in planetary atmospheres, because free convection produces a strong upward transfer of heat whenever the lapse rate reaches the adiabatic value. An isothermal lapse rate ($dT/dz = 0$) represents quite a stable stratification, and a so-called inversion ($dT/dz > 0$) represents an even stronger one.

Thus the stratosphere is much more stably stratified than the troposphere (hence the name), and therefore it is a layer in which vertical mixing of trace substances is strongly suppressed. The characteristic residence time of air parcels in the stratosphere (that is, the time elapsed since they were in the troposphere) ranges from a few months just above the base to many years at the stratopause level. Hence, the stratosphere functions as a long-term reservoir for certain types of pollutants, such as debris from nuclear tests, which are quickly cleansed from tropospheric air by processes discussed below. The high static stability of the stratosphere also limits the height to which plumes of rising air in severe thunderstorms or volcanic eruptions can rise before they become negatively buoyant.

Atmospheric motions. The levels and temperatures of the tropopause, stratopause, and mesopause vary with latitude and season (**Fig. 5**). The stratopause is warmest at the summer pole, where the solar heating is strongest, and coolest at the winter pole, which is in darkness. However, the temperature distribution at the mesopause is the reverse of what would be expected on the basis of arguments based on radiative transfer; the summer pole is much colder than the winter pole. An equally strange distribution (from the point of view of radiative transfer) is observed at the tropopause, which is coldest (−80°C or −112°F or less) on the Equator and warmest at the summer pole and at middle latitudes of the winter hemisphere. At these levels the radiative heating is relatively weak, and dynamical processes are capable of driving temperatures far from their radiative equilibrium values. For example, the extreme coldness of the summer mesopause and the equatorial tropopause is maintained by adiabatic expansion associated with large-scale upward motion.

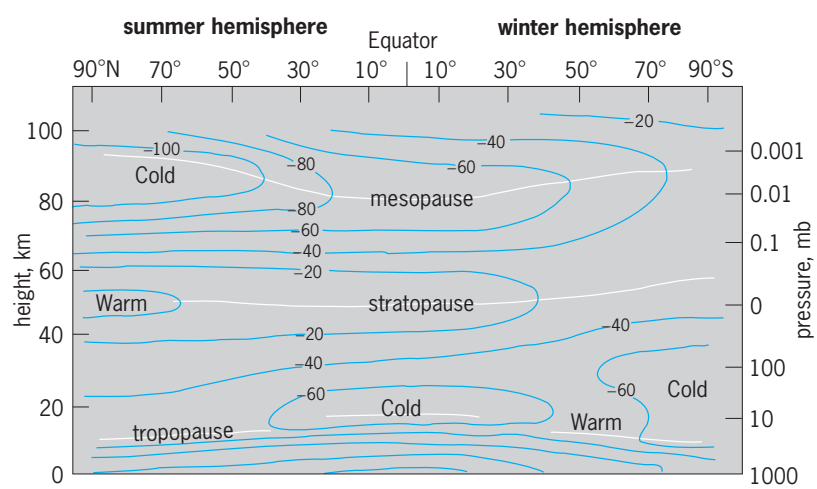

Fig. 5. Meridional cross section of the longitudinally averaged temperature (°C) at the time of the solstices. °F = (°C × 1.8) + 32. 1 millibar = 10^2 Pa. 1 km = 0. 6 mi. (*After J. M. Wallace and P. V. Hobbs, The Atmosphere: An Introductory Survey, Academic Press, 1977*)

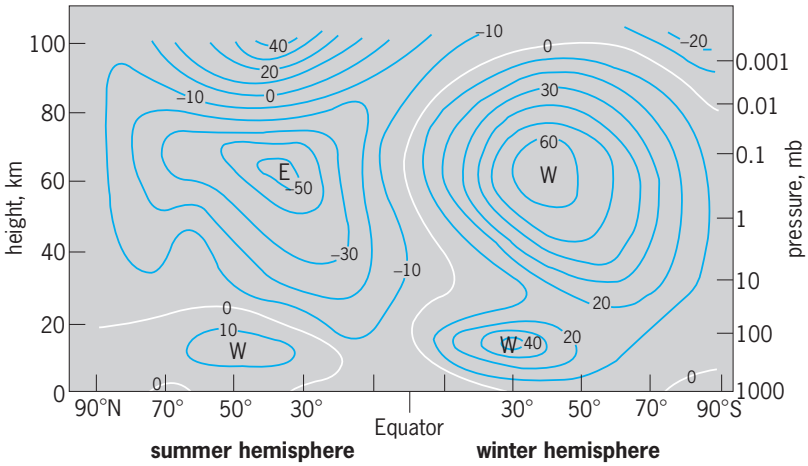

Fig. 6. Meridional cross section of the longitudinally averaged zonal wind ($m \cdot s^{-1}$) at the time of the solstices. Positive zonal winds denote flow from west (W) to east (E). $1\ m \cdot s^{-1} = 2.2$ mi/h. (*After J. M. Wallace and P. V. Hobbs, The Atmosphere: An Introductory Survey, Academic Press, 1977*)

There is a corresponding distribution of the zonal (west-to-east) component of the wind (**Fig. 6**). Wind maxima are observed at the tropopause level (the so-called tropospheric jet streams) and in the mesosphere. The tropospheric jet stream, which is present throughout the year, blows from west to east. It is stronger and somewhat farther equatorward during winter. In the mesosphere the winds blow from the west during winter and from the east during summer. These zonal winds are in thermal wind balance, with the temperature distribution shown in Fig. 6: wherever temperature decreases (increases) with latitude, the zonal wind is becoming more (less) westerly with increasing height. *See* GEOSTROPHIC WIND; JET STREAM.

Winds in the equatorial lower stratosphere (not shown in Fig. 6) exhibit a remarkable 27-month quasiperiodicity, with alternating periods of remarkably persistent easterly and westerly winds, which appear first near the 10-mb or 1-kPa (30-km or 18-mi) level and gradually descend, over the course of a year or so, to the 70-mb or 7-kPa (18-km or 11-mi) level. The peak-to-peak amplitude of the so-called quasibiennial oscillation reaches $45\ m \cdot s^{-1}$ (100 mi/h). The westerly polar vortex in the winter hemisphere is distorted by planetary waves that propagate energy upward from below. In the Northern Hemisphere, these disturbances sometimes become so intense during midwinter that they produce a so-called sudden warming, that is, the disappearance of the cold temperatures normally found over the polar regions at this time of year and the westerly vortex that encircles them.

The tropospheric circulation exhibits a complex array of disturbances on a wide range of space and time scales. Prominent among them are baroclinic waves, whose signature on synoptic charts at the Earth's surface is characterized by migrating extratropical cyclones and anticyclones. Because of the relatively lower static stability at these levels, tropospheric disturbances are characterized by much larger vertical motions than stratospheric disturbances. Lifting is often sufficient to produce widespread condensation of water vapor, which gives rise to clouds and precipitation. *See* DYNAMIC INSTABILITY; UPPER-ATMOSPHERE DYNAMICS; WIND.

Temperature at Earth's surface. With the notable exception of Jupiter, which is emitting substantial amounts of energy released by gravitational compression, the individual planets can be regarded as being in radiative equilibrium with the Sun. They intercept $\pi R^2 S$ units of solar radiation, where R is the radius of the planet and S is the solar irradiance, which decreases with distance from the Sun in accordance with the inverse-square law. Of the intercepted solar radiation, the fraction A, defined as the planetary albedo, is reflected to space; the remaining fraction $(1 - A)$ is absorbed by the planet and its atmosphere. An equal amount of radiation is emitted to space by the planet and its atmosphere. The planetary radiation may be expressed as the amount of radiation emitted by a blackbody at the effective temperature T_E of the planet. T_E can be determined from the balance between incoming and outgoing radiation, Eq. (12), where σ is the Stefan-Boltzmann constant

$$(1 - A)\pi R^2 S = \sigma T_E^{\,4}(4\pi R^2) \qquad (12)$$

$5.67 \times 10^{-8}\ W \cdot m^{-2} \cdot K^{-4}$. For the Earth, the solar irradiance S is $1380\ W \cdot m^{-2}$, $A \approx 0.30$, for which the solution of Eq. (12) yields $T_E = 255$ K. It is readily verified that if nothing else changed, $dT_E/T_E = {}^1/_4 dS/S$ so that, for example, a 1% increase in solar irradiance would raise the effective temperature of the Earth by about 0.64 K (1.2°F). Because of its high planetary albedo (0.78), Venus exhibits an effective temperature of only 227 K (−51°F) despite the fact that it intercepts more than three times as much solar radiation per unit surface area than the Earth does. The effective temperature of Jupiter, as measured from space, is about 125 K (−235°F), compared with a value of 105 K (−271°F) computed from Eq. (12). *See* ALBEDO; HEAT RADIATION; PLANETARY PHYSICS; SOLAR RADIATION.

The peak wavelength λ_m of planetary radiation can be estimated by applying the Wien displacement law to a blackbody at the effective temperature of the planet [Eq. (13)]. For the Earth, it is readily verified

$$\lambda_m = \frac{2897}{T} \qquad (13)$$

that $\lambda_m \approx 15\ \mu$m. Normalized blackbody curves for solar and terrestrial radiation are shown in **Fig. 7** on a logarithmic scale. Since the overlap between the two curves is minimal, solar and terrestrial radiation may be treated separately. *See* BLACKBODY.

The fact that the mean surface temperature of the Earth is about 33 K (59°F) warmer than the effective temperature is due to the absorption and subsequent downward reemission of substantial amounts of terrestrial radiation by water vapor, CO_2, and O_3, and cloud layers, all of which absorb strongly in the infrared (the so-called greenhouse effect). Figure 7*b* shows position of the absorption bands of the atmosphere's major gaseous constituents in relation to

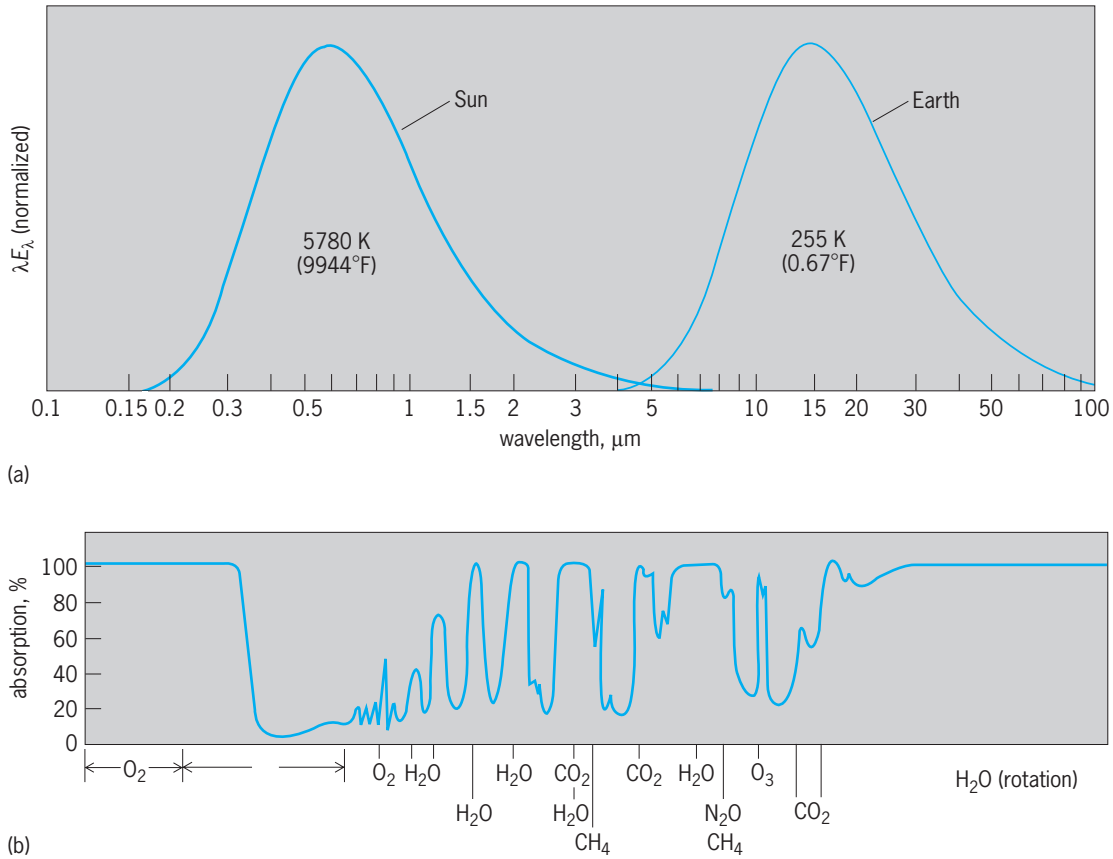

Fig. 7. Radiation curves. (*a*) Normalized blackbody spectra representative of the Sun and the Earth plotted on a logarithmic scale. (*b*) Absorption spectrum for the Earth's atmosphere as a whole. The ordinate denotes the fraction of the radiation incident upon the Earth's atmosphere (either from above or below) that is absorbed during its passage through the atmosphere; the gaseous constituents primarily responsible for the absorption of the radiation at various wavelengths are indicated at the horizontal axis. (*After R. M. Goody, Atmospheric Radiation, Oxford University Press, 1964*)

the spectra of solar and terrestrial radiation. The broad spectral window that enables most of the incident solar radiation (apart from that reflected by clouds) to reach the Earth's surface is readily apparent. The main spectral window through which terrestrial radiation escapes to space is in the vicinity of 10 μm. The increasing concentration of CO_2 in the atmosphere due to the burning of fossil fuels is causing its absorption bands to encroach on the long-wavelength end of this window, leading to predictions of a substantial global greenhouse warming. Methane (CH_4), nitrogen oxides (NO_x), and chlorofluorocarbons, whose concentrations are also increasing in the Earth's atmosphere at least partly as a result of human activity, exhibit absorption bands closer to the middle of this window. Hence, even though their concentrations are orders-of-magnitude smaller than that of CO_2, they are also a source of concern. Based on current projections, increases in these constituents over the next 50 years could cause a global warming comparable to that resulting from a doubling of CO_2. *See* GREENHOUSE EFFECT; HALOGENATED HYDROCARBON.

The mean surface temperature of the Earth is determined by a rather delicate balance between the globally averaged fluxes of solar and terrestrial radiation, sensible heat, and moisture. The 100 units of incoming radiation shown in **Fig. 8** represents the solar irradiance passing through the Earth's orbit (S) times the cross-sectional area of the Earth (πR^2), divided by the total surface area of the Earth ($4\pi R^2$): hence, it is given by $S/4 = 345\ \mathrm{W \cdot m^2}$. The combined effects of the reflection of solar radiation by clouds, air molecules, and the Earth's surface amount to 30 of the 100 units of incoming solar radiation, which accounts for the planetary albedo of 30% alluded to above. Absorption of solar radiation by water-vapor molecules, ozone molecules in the stratosphere, clouds, and aerosols together amounts to 19% of the incoming solar energy. The remaining 51 units are absorbed at the Earth's surface and eventually are returned to the atmosphere by the processes indicated on the right-hand side of the diagram.

The Earth's surface can be regarded as a blackbody at 288 K (15°C or 59°F), for which the Stefan-Boltzmann law predicts in upward irradiance of $390\ \mathrm{W \cdot m^{-2}}$, or 113 of the units in Fig. 8. The irradiance of downward infrared radiation emitted by the atmosphere is 92 units, leaving a net upward irradiance of only 21 units, as shown. If the Earth's surface were in radiative equilibrium (in which case it would have to dispose of the full 51 units of solar

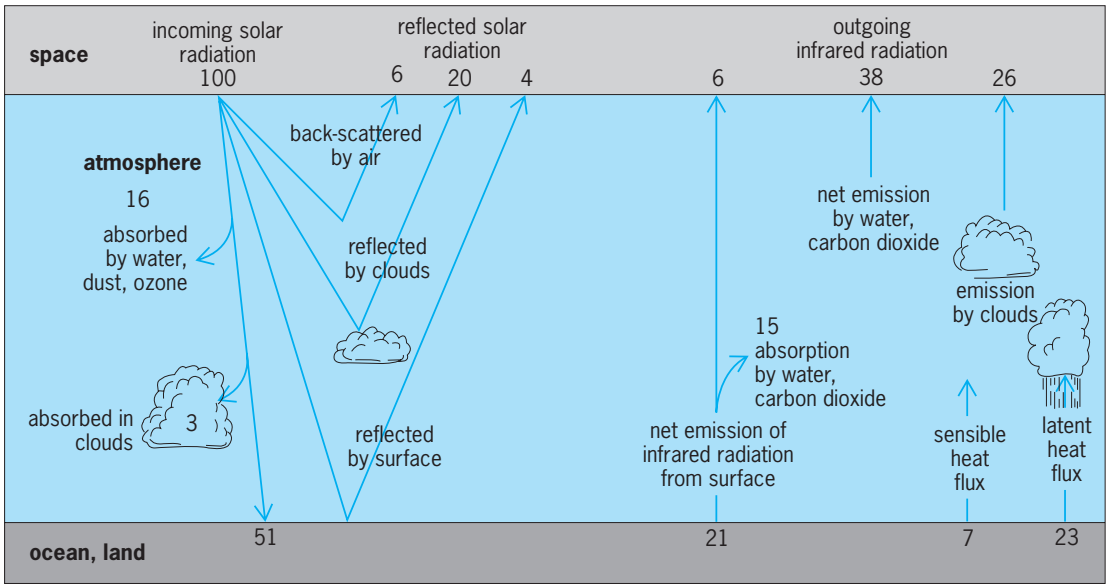

Fig. 8. Annual mean global energy balance of the Earth–atmosphere system. Numbers are given as percentages of the globally averaged solar radiation incident upon the top of the atmosphere. (*After Understanding Climatic Change, U.S. National Academy of Sciences, 1975*)

radiation that it absorbs by a net upward irradiance of infrared radiation), its mean temperature would have to be on the order of 340 K (152°F). The remaining 30 units are transferred to the atmosphere through the fluxes of sensible heat (c_pT) and the latent heat of vaporization of water evaporated at the Earth's surface, which is eventually transferred to the surrounding air molecules when the vapor condenses in clouds. From Fig. 8 it is evident that the effective temperature of the Earth, as viewed from space, is a weighted average of the temperatures of the surface ($^6/_{70}$), the greenhouse gas molecules ($^{38}/_{78}$), and cloud tops ($^{26}/_{70}$). The role of water vapor and clouds will be discussed in more detail in the next section.

The estimates presented in Fig. 8 refer to annual and globally averaged quantities. The latitudinal seasonal variation net incoming solar radiation at the top of the atmosphere is shown in **Fig. 9**. The polar night region is shaded. The small deviations from equatorial symmetry are a consequence of the ellipicity of the Earth's orbit: the Earth is closest to the Sun in January. The equator-to-pole heating gradient is very strong in the winter hemisphere, and it disappears or even reverses for about a month centered on the summer solstice. In the annual average, the polar regions absorb only about 30% as much solar radiation per unit area as the equatorial belt when the high reflectivity of the ice and snow and the persistent cloud cover in those regions are taken into account. The outgoing infrared radiation through the top of the atmosphere (not shown) is proportional to the fourth power of the local effective temperature T_E at which the radiation is emitted. Annual average surface temperature varies from nearly 300 K (80°F) in the tropics to about 250 K (−10°F) in the polar regions, which would be consistent with a 2:1 ratio of outgoing radiation between Equator and poles. Because cloud tops and water vapor, one of the major gaseous sources of terrestrial radiation, extend to higher levels in the tropical atmosphere than in the polar regions and are therefore relatively colder, the actual ratio is closer to 1.3:1.

Averaged over the entire globe and over the whole year, the net downward solar radiation through the top of the atmosphere must very nearly equal the upward infrared radiation. Since the meridional gradient of the former is stronger, it follows that at low latitudes the incoming solar radiation must exceed the local outgoing infrared radiation, while at higher latitudes the opposite situation must prevail. Hence, the atmosphere and oceans must transport energy poleward from a low-latitude source to a high-latitude sink. This north-south heating gradient, which is much larger in the winter hemisphere than in the summer hemisphere, is the main driving force for large-scale atmospheric motions, the so-called general circulation. Land–sea contrasts (differences in heat capacity, thermal conductivity, reflectivity, and availability of moisture between land and sea and between different types of land surfaces) also contribute to the large observed spatial and temporal variability of the local energy balance. The resulting east-west (zonal) heating gradients drive the monsoon circulations in the tropics and subtropics and the stationary planetary waves at higher latitudes. *See* MONSOON METEOROLOGY.

The energy balance arguments used in this section should not be interpreted as indicating that the Earth's climate is static. The climate system has, in fact, emerged from a major ice age less than 20,000 years ago, and it has warmed substantially since the 1880s. However, it can be shown that the energy fluxes required to account for the observed changes in the volume of the polar icecaps and the heat storage in the atmosphere and oceans on these time scales are much smaller than the uncertainties in the

estimates of the current energy balance in Fig. 8. *See* GLACIAL EPOCH.

Impact of hydrologic cycle. Water in its various phases exerts a profound influence not only upon the biosphere but also upon many aspects the behavior of the atmosphere. It is evident in Fig. 8 that evaporation is an important heat sink for the oceans and vegetated land surfaces, and that condensation of water vapor in clouds is the atmosphere's largest single heat source. Hence, the hydrological cycle transfers massive amounts of heat from the Earth's surface into the atmosphere. On a more local scale, condensation plays a major role in generating the buoyancy and the horizontal temperature gradients that drive hurricanes, explosively deepening extratropical cyclones, and severe thunderstorms. It occurs selectively within rising air parcels, which tend to be warmer than the environmental air at the same level, and it causes them to cool more slowly as they expand than they would under adiabatic conditions. Observed lapse rates in the troposphere rarely exceed this so-called moist adiabatic lapse rate.

Condensation of water vapor within the atmosphere involves a complex array of physical processes. Submicrometer aerosol particles known as cloud condensation nuclei play an essential role in the initial formation of droplets of liquid water, which subsequently grow by many orders of magnitude on a time scale of minutes to hours, until the speeds at which they fall become large enough to enable them to reach the ground as rain or snow. The principal mechanisms through which this growth occurs involve the coalescence of smaller droplets into larger ones through collisions and the so-called Bergeron mechanism—the freezing of a small fraction of the droplets and their subsequent growth at the expense of the remaining supercooled (below-freezing) liquid droplets through the diffusion of water-vapor molecules from one to the other.

Liquid water droplets and ice particles tend to be concentrated in macroscale clouds, for which a detailed classification scheme has been devised. Among the most common cloud types are extensive, long-lived cloud layers (cirrostratus, altostratus, stratus, stratocumulus), which cool the Earth through their contribution to the planetary albedo and warm it through their contribution to the greenhouse effect (Fig. 8). On balance, upper-tropospheric (cirrostratus) cloud layers, which emit radiation to space at low effective temperatures, tend to warm the Earth; and low cloud layers (stratus, stratocumulus) tend to cool it. Because of this compensation, the phenomenon of cloud–climate feedbacks represents one of the largest sources of uncertainty in projections of future global climate change due to the buildup of greenhouse gases in the atmosphere.

Cloud droplets provide a hospitable environment for certain types of chemical reactions that would not otherwise take place in the atmosphere. Such reactions figure prominently in the destruction of the stratospheric ozone over the polar regions and in the formation of acid rain downstream from industrial sources of sulfur dioxide (SO_2). Scavenging by cloud droplets that subsequently fall out as precipitation is the primary mechanism for cleansing dust, smoke, and other particles from the troposphere. Charge separation by cloud droplets and ice particles within clouds also has implications for the global atmospheric electric field, and it leads to the buildup of the locally strong potential gradients responsible for lightning discharges. *See* LIGHTNING.

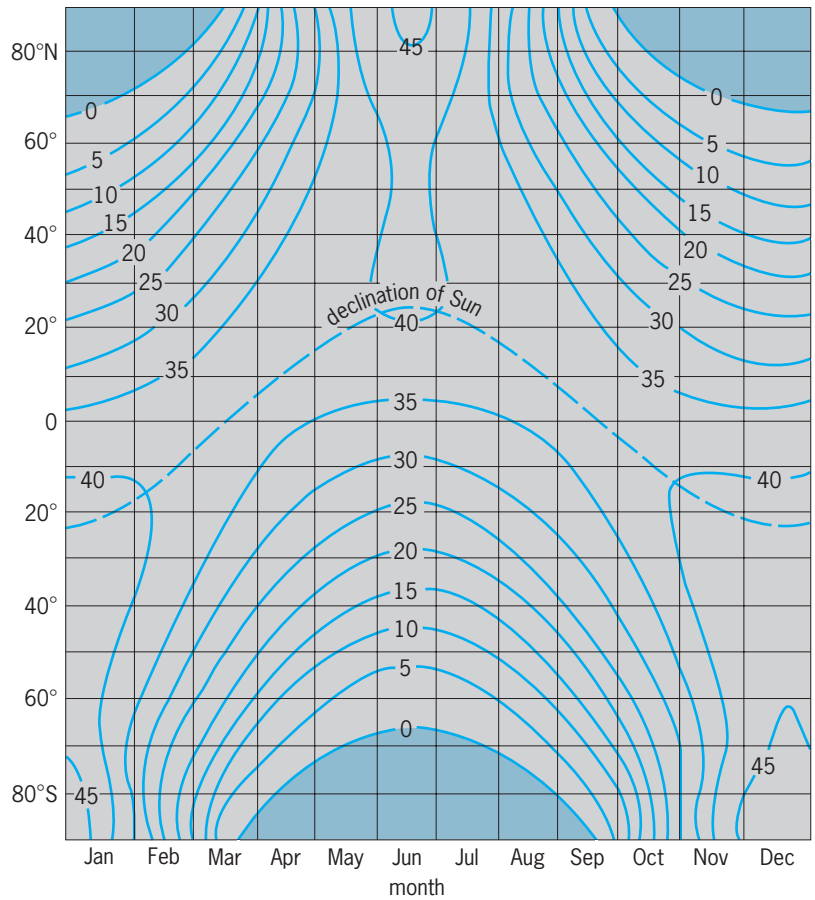

Fig. 9. Solar radiation incident upon a unit horizontal surface at the top of the atmosphere as a function of latitude and calendar date; contours are given in 10^6 J · m^{-2} · day^{-1}. (*After R. J. List, ed., Meteorological Tables, 6th ed., Smithsonian Institution, 1951*)

The 23 units of latent heat flux from the Earth's surface in Fig. 8 are based on an observed evaporation rate of about 0.25 cm (0.1 in.) of liquid water per day averaged over the Earth's surface, which must be equal to the average rate of precipitation. The average precipitable water in the atmosphere is on the order of about 2 cm (0.8 in.) of liquid water, and thus the mean residence time for water-vapor molecules in the atmosphere must be on the order of a week. The mass of liquid water and ice present in the atmosphere in the form of cloud droplets is several-orders-of-magnitude smaller than the amount present in the vapor state. In contrast to the tropospheric air, which is often saturated with water vapor, stratospheric air tends to be remarkably dry because it enters the stratosphere by way of the extremely cold equatorial tropopause (Fig. 5), where most of its water vapor is condensed out.

Stratospheric cloud layers are observed only in the polar night region where temperatures can drop to −80°C (−112°F). *See* CLOUD; HYDROMETEOROLOGY; PRECIPITATION (METEOROLOGY).

Atmospheric prediction. Throughout the atmospheric sciences, prediction is a unifying theme that sets the direction for research and technological development. Prediction on the time scale of minutes to hours is concerned with severe weather events such as tornadoes, hail, and flash floods, which are manifestations of intense mesoscale weather systems, and with urban air-pollution episodes; day-to-day prediction is usually concerned with the more ordinary weather events and changes that attend the passage of synoptic-scale weather systems such as extratropical cyclones; and seasonal prediction is concerned with regional climate anomalies such as drought or recurrent and persistent cold air outbreaks. Prediction on still longer time scales involves issues such as the impact of human activity on the temperature of the Earth, regional climate, the ozone layer, and the chemical makeup of precipitation. *See* CLIMATE MODELING; DROUGHT; HAIL; TORNADO.

The evolution of the atmospheric sciences from a largely descriptive field prior to World War II to a mature, quantitative physical science discipline is apparent in the development of vastly improved predictive capabilities based upon the numerical integration of specialized versions of the Navier-Stokes equations, which include sophisticated parametrizations of physical processes such as radiative transfer, latent heat release, and microscale motions. The so-called numerical weather prediction models have largely replaced the subjective and statistical prediction methods that were widely used as a basis for day-to-day weather forecasting as recently as the 1950s. The state-of-the-art numerical models exhibit significant skill for forecast intervals as long as about a week. *See* NAVIER-STOKES EQUATION.

A distinction is often made between weather prediction, which is largely restricted to the consideration of dynamic and physical processes internal to the atmosphere, and climate prediction, in which interactions between the atmosphere and other elements of the climate system are taken into account. The importance and complexity of these interactions tend to increase with the time scale of the phenomena of interest in the forecast. Weather prediction involves shorter time frames (days to weeks), in which the information contained in the initial conditions is the dominant factor in determining the evolution of the state of the atmosphere; and climate prediction involves longer time frames (seasons and longer), for boundary forcing is the dominant factor in determining the state of the atmosphere.

As in many other systems governed by nonlinear equations, the uncertainties inherent in the definition of the initial state of the atmosphere grow exponentially with time during the forecast until they become as large as typical differences between two arbitrarily chosen states of the atmosphere. Deterministic prediction based solely upon the information contained in the initial conditions is impossible beyond this time frame. The characteristic predictability time of large-scale atmospheric circulation patterns is believed to be on the order of 2 weeks. Prediction on longer time scales such as months or seasons exploits the memory (autocorrelation) and extended predictability of the more slowly varying components of the climate system, which force the atmosphere at its lower boundary. On these longer time scales, prediction is focused on climate anomalies (that is, departures of statistics such as seasonal mean temperature and precipitation from their climatological mean values), which persist or develop in response to the slowly varying boundary forcing. Deterministic prediction of sequences of day-to-day weather changes is not feasible on these extended time scales.

Atmospheric prediction has benefited greatly from major advances in remote sensing. Geostationary and polar orbiting satellites provide continuous surveillance of the global distribution of cloudiness, as viewed with both visible and infrared imagery. These images are used in positioning of features such as cyclones and fronts on synoptic charts. Cloud motion vectors derived from consecutive images provide estimates of winds in regions that have no other data. Passive infrared and microwave sensors aboard satellites also provide information on the distribution of sea-surface temperature, sea state, land-surface vegetation, snow and ice cover, as well as vertical profiles of temperature and moisture in cloud-free regions. Improved ground-based radar imagery and vertical profiling devices provide detailed coverage of convective cells and other significant mesoscale features over land areas. Increasingly sophisticated data assimilation schemes are being developed to incorporate this variety of information into numerical weather prediction models on an operational basis. *See* ATMOSPHERE; CLIMATIC PREDICTION; CYCLONE; FRONT; RADAR METEOROLOGY; SATELLITE METEOROLOGY; WEATHER FORECASTING AND PREDICTION.

John M. Wallace

Bibliography. S. Ackerman and J. A. Knox, *Meteorology: Understanding the Atmosphere*, 2002; C. D. Ahrens, *Meteorology Today*, 2d ed., 2006; W. S. Broecker, *How to Build a Habitable Planet*, 1985; F. K. Lutgens et al., *The Atmosphere: An Introduction to Meteorology*, 9th ed., 2003; V. J. Schaefer and J. Day, *A Field Guide to the Atmosphere*, 1981; R. B. Stull, *Meteorology for Scientists and Engineers*, 2d ed., 1999; J. M. Wallace and P. V. Hobbs, *The Atmosphere: An Introductory Survey*, 1977.

Methane

The simplest compound of carbon and hydrogen. Each molecule of methane contains four atoms of hydrogen bound in a tetrahedral arrangement to one atom of carbon (see **illustration**). At room temperature, methane is a gas less dense than air. The gas liquefies at −164°C (−263°F) and solidifies at −183°C (−297°F). It is not very soluble in water. Methane is combustible, and mixtures of about 5–15% in

Molecular model of methane.

air are explosive. Complete combustion of methane produces carbon dioxide and water. Methane is not toxic when inhaled, but it can produce suffocation by reducing the concentration of oxygen. *See* ALKANE.

Methane is widely distributed in nature. It is the principal component of natural gas, a mixture that by volume contains about 75% methane (CH_4), 15% ethane (C_2H_6), and 5% other hydrocarbons such as propane (C_3H_8) and butane (C_4H_{10}). The combustible gas found in coal mines is chiefly methane. Marsh gas, which is produced under water by anaerobic bacterial decomposition of plant and animal matter, is also methane. Some plants emit methane, and many animals emit methane as a by-product of food digestion. It is estimated that domestic animals account for about 15% of annual methane emissions. Although methane constitutes only about 0.00017% by volume of the atmosphere, its concentration has more than doubled in the last 250 years. Methane is one of the most powerful greenhouse gases in the atmosphere, being more than 60 times as effective as carbon dioxide. *See* COALBED METHANE; GREENHOUSE EFFECT; NATURAL GAS.

Natural gas occurs in reservoirs beneath the surface of the Earth, and it is often found in conjunction with petroleum deposits. Before it is commercially distributed, natural gas usually undergoes some sort of processing in which the heavier hydrocarbons (propane and butane) and nonhydrocarbon gases (such as hydrogen sulfide) are removed. The cleaned gas is then distributed through pipelines. Natural gas is also distributed worldwide in liquid form (liquefied natural gas, or LNG), which is produced by chilling the gas to below its boiling point. Because an undetected natural gas leak could result in an explosion or asphyxiation, local utilities add an odorant to make the gas readily detectable by smell. The odorant is a small amount of a mixture of organic sulfur compounds such as tertiary-butyl mercaptan [$(CH_3)_3CSH$] and dimethyl sulfide (CH_3—S—CH_3). *See* LIQUEFIED NATURAL GAS (LNG).

Methane is synthesized commercially by distilling bituminous coal and by heating a mixture of carbon and hydrogen. It can be produced in the laboratory by heating sodium acetate with sodium hydroxide and by the reaction of aluminum carbide (Al_4C_3) with water.

Uses. The principal use of methane is as a fuel. The combustion of methane is highly exothermic [reaction (1)]. The energy released by the combus-

$$CH_4(g) + 2O_2(g) \rightarrow CO_2(g) + 2H_2O(l) \qquad (1)$$

$$\Delta H = -891 \text{ kJ}$$

tion of methane, in the form of natural gas, is used directly to heat homes and commercial buildings. It is also used in the generation of electric power. In 1996–2005, natural gas accounted for one-fifth of the total energy consumption worldwide and about one-third in the United States.

In the chemical industry, methane is a raw material for the manufacture of methanol (CH_3OH), formaldehyde (CH_2O), nitromethane (CH_3NO_2), chloroform (CH_3Cl), carbon tetrachloride (CCl_4), and some halogenated hydrocarbons (compounds containing carbon and fluorine, and perhaps chlorine and hydrogen). The reactions of methane with chlorine and fluorine are triggered by light, and when exposed to bright visible light, mixtures of methane with chlorine or fluorine can react explosively. Industrially, methane is converted (reformed) catalytically by the Fischer-Tropsch process to a mixture of carbon monoxide and hydrogen gases called syngas [reaction (2)]. Via other catalysts, syngas is converted

$$CH_4(g) + H_2O(g) \rightarrow CO(g) + 3H_2(g) \qquad (2)$$

to liquid diesel fuel. The hydrogen gas produced in the reformation of methane is used in another catalytic process to make ammonia which, in turn, is converted to important fertilizers. *See* FISCHER-TROPSCH PROCESS; SYNTHETIC FUEL.

Hydrates. Methane hydrates are solids consisting of methane and water, and these white solids are less dense than water. Methane hydrates are examples of clathrate hydrates, in which "cages" of water molecules, held together by hydrogen bonds, trap gases (such as N_2, O_2, CO_2, CH_4, H_2S) at high pressure and low temperature. On average, methane hydrates contain about one mole of methane to almost six moles of water. Methane hydrates are found in the oceans at several hundred feet below sea level, where they are trapped in mud and rocks. When brought to the surface, they decompose to water and methane, and the methane burns—hence, the name ice-on-fire for the phenomenon. Methane hydrates are being investigated as a commercial source of methane. *See* CLATHRATE COMPOUNDS; HYDRATE.

Planets. The detection of methane on other planets is of major scientific interest. Whether on Mars or the cold surface of Saturn's largest moon, Titan, the presence of methane and the conjecture of riverlike channels due to liquid methane have intrigued scientists and spurred closer investigations.

Bassam Z. Shakhashiri; Rodney Schreiner

Bibliography. S. Lee, *Methane and Its Derivatives*, 1996; R. T. Morrison and R. N. Boyd, *Organic Chemistry*, 6th ed., 1992; E. D. Sloan, *Clathrate Hydrates of Natural Gases*, 2d ed., 1998.

Methanogenesis (bacteria)

The microbial formation of methane. It is confined to anaerobic habitats where the production of hydrogen, carbon dioxide, formic acid, methanol, methylamines, or acetate—the major substrates used by methanogenic microbes (methanogens)—occurs.

Habitats. In freshwater or marine sediments, in the intestinal tracts of animals, or in habitats engineered by humans such as sewage sludge or biomass digesters, these substrates are the products of anaerobic bacterial metabolism. Methanogens are terminal organisms in the anaerobic microbial food chain—the final product, methane, being poorly soluble, anaerobically inert, and not in equilibrium with the reaction which produces it. In hydrothermal vents, hydrogen formation is a product of geochemistry. In methanogenic habitats, the reducing potential may approximate the hydrogen electrode. The reducing potential in hydrothermal vents is maintained by geochemically produced hydrogen sulfide; and in sediments, especially marine sediments, hydrogen sulfide is a product of bacterial reduction of sulfate. In freshwater sediments and in the intestinal tract of animals, reducing potential is mainly maintained by the hydrogenases of fermentative anaerobes that produce molecular hydrogen. *See* HYDROTHERMAL VENT.

Two highly specialized digestive organs, the rumen and the cecum, have been evolved by herbivores to delay the passage of cellulose fibers so that microbial fermentation may be complete. In these organs, large quantities of methane are produced from hydrogen and carbon dioxide or formic acid by methanogens. From the rumen, which may hold 26 gallons (100 liters) of fermenting biomass, an average cow may belch 26 gallons of methane per day. Methanogens may form endosymbiotic or exosymbiotic relationships with cellulose-ingesting protozoa in the rumen, cecum, hind gut of termites, or sediments.

Isolation and cultivation of methanogens. Methanogens are sensitive to oxygen and cannot be cultivated by ordinary bacteriological techniques; exclusion of oxygen is essential but not sufficient. To plate methanogens on solid agar media in petri dishes, an anaerobic chamber is required for routine bacteriological techniques. Colonies on plates must be examined in an anaerobic chamber. Cell walls of methanogens lack the typical peptidoglycan of bacteria, so methanogens are not sensitive to the common cell-wall antibiotics such as cycloserine and penicillin. Addition of such antibiotics to growth media inhibits the growth of typical bacteria and aids in the selective isolation of methanogens from crude inocula.

Organisms. Methanogens are the only living organisms that produce methane as a way of life. The biochemistry of their metabolism is unique in the biological world and definitively delineates the group. *See* ARCHAEA.

Metabolic reactions are energy-yielding reactions that are carried out in nature. Two reductive biochemical strategies are employed: an eight-electron reduction of carbon dioxide to methane or a two-electron reduction of a methyl group to methane. All methogens form methane by reducing a methyl group.

The major energy-yielding reactions used by methanogens utilize substrates such as hydrogen, formic acid, methanol, acetic acid and methylamine. Dimethyl sulfide, carbon monoxide, and alcohols such as ethanol and propanol are substrates that are used less frequently.

Methanococcus vannielii is a motile coccus, isolated from marine sediment, that grows at 20–40°C. In contrast, *Methanocaldococcus jannaschii* (formerly known as *Methanococcus jannaschii*), isolated from a deep-sea hydrothermal vent, grows at 48–94°C with optimal growth at 85°C.

Of the rod-shaped organisms, *Methanobrevibacter ruminantium*, isolated from a bovine rumen, is a short rod with tapering ends that grows at 37–39°C. The organism has complex nutritional requirements, including B vitamins, coenzyme M, amino acids, acetate, and 2-methylbutyrate. *Methanobacterium formicicum*, isolated from sediment, is a slightly bent rod that grows at 25–37°C. *Methanothermobacter thermoautotrophicus* (formerly known as *Methanobacterium thermoautotrophicum*), isolated from a sewage sludge digester, grows at 50–75°C. This species is easily cultivated in kilogram quantities at 65°C on hydrogen and carbon dioxide in a mineral salts medium. Most of the biochemistry of carbon dioxide reduction to methane was revealed through study of this organism.

Methanothermus fervidus, isolated from a geothermal area in Iceland, is a short rod that grows at 60–97°C. *Methanosphera stadtmaniae*, although closely related to the rod-shaped methanogens, has a spherical morphology and is unique because it cannot reduce carbon dioxide to methane. Instead, hydrogen is oxidized and a methyl group is reduced to methane. The blunt rods of *Methanosaeta concilii* are encased in a sheath. The organism is widely distributed in nature and grows only by an aceticlastic reaction. *Methanosarcina barkeri*, the most metabolically diverse methanogen, also has the ability to use the aceticlastic reaction. Coccoid cells form large clumps. Most isolates are mesophilic; a few will grow at 50°C.

Methanogenium cariaci, a marine coccoid organism that is not closely related to *Methanococcus*, grows optimally at 40°C. *Methanospirillum hungateii*, isolated from a sewage sludge digester, grows at 35–40°C and is the only spiral-shaped methanogen. The most distantly related of all methanogens is the extreme thermophile *Methanopyrus kandleri*, which was isolated from a hydrothermal vent and has an upper limit of growth at 110°C. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; METHANE.

Ralph S. Wolfe

Bibliography. J. G. Ferry (ed.), *Methanogenesis*, 1993.

Methanol

The first member of the homologous series of alkanols, with the formula CH_3OH. This compound was originally obtained by the destructive distillation of wood as a by-product in the preparation of charcoal, hence the older name wood alcohol. Methanol is a highly flammable liquid, boiling point 64.7°C (149°F), and is miscible with water and most organic liquids. It is a highly poisonous substance; sublethal amounts can cause permanent blindness. *See* ALCOHOL.

Methanol is one of the major industrial organic chemicals. It is produced commercially from a mixture of carbon monoxide (CO) and hydrogen (H_2), generally known as synthesis gas, according to reaction (1), where ΔH = change in enthalpy. The pro-

$$CO + 2H_2 \xrightarrow{\text{catalyst}} CH_3OH \qquad (1)$$

$$\Delta H = -90.8 \text{ kJ/mole}$$

cess is carried out at 200–300°C (390–570°F) at moderate pressure (50 atm or 5 MPa) with a copper-based catalyst. *See* CATALYSIS; ENTHALPY.

The synthesis gas used to produce methanol can be produced from coal, natural gas, or petroleum fractions. The major source of synthesis gas is the steam reforming of methane (CH_4) or other hydrocarbons. Reactions (2) and (3) represent the basic

$$CH_4 + H_2O \longrightarrow CO + 3H_2 \qquad (2)$$

$$\text{-}(CH_2)\text{-}_n + nH_2O \longrightarrow nCO + 2nH_2 \qquad (3)$$

steps; the overall process also includes a partial oxidation step to supply heat and adjust the CO/H_2 ratio. *See* REFORMING PROCESSES.

Methanol has the usual chemical properties of a primary alcohol, undergoing traditional reactions to give methyl acetals, amines, ethers, esters, and halides. Addition to isobutene gives methyl *t*-butyl ether (MTBE) [reaction (4)], which is produced on

$$\underset{\text{Isobutene}}{(CH_3)_2C{=}CH_2} + CH_3OH \longrightarrow \underset{\text{MTBE}}{(CH_3)_3COCH_3} \qquad (4)$$

large scale as an octane enhancer to replace lead additives in gasoline.

The major industrial use of methanol is conversion to formaldehyde (CH_2O). This process is carried out by the reaction of methanol with a limited amount of oxygen over a silver catalyst [reaction (5)].

$$CH_3OH + {}^1/_2O_2 \xrightarrow{\text{catalyst}} CH_2O + H_2O \qquad (5)$$

See FORMALDEHYDE.

Since the 1970s, several manufacturing processes have been developed based on transition metal–catalyzed carbonyl insertion reactions. One of the most important of these is the carbonylation of methanol to acetic acid. A rhodium-based catalyst system was the first commercial application, but a less costly process using a nickel catalyst has been reported. Similar carbonylation converts methyl acetate to acetic anhydride [reactions 6].

$$\begin{array}{ccc} CH_3OH & \xrightarrow{CO} & \underset{\text{Acetic acid}}{CH_3CO_2H} \\ \underset{\text{Acetic anhydride}}{(CH_3CO)_2O} & \xrightarrow{CO} & \underset{\text{Methyl acetate}}{CH_3CO_2CH_3} \end{array} \qquad (6)$$

Carbonyl insertion can be coupled with oxidation or reduction, thus leading to an extensive suite of compounds, including such important products as vinyl acetate and ethylene glycol. Methanol or its immediate derivatives formaldehyde and methyl formate are central to these processes, all of which are based on synthesis gas as feedstock. Since the technology exists for obtaining synthesis gas from coal also, this so-called C_1 chemistry provides an important alternative to petroleum- and ethylene-based raw materials for the chemical industry. *See* COAL GASIFICATION; ORGANIC SYNTHESIS.

The potential of methanol as an alternative fuel has been widely discussed. The volatility, ease of transport and distribution, and low combustion emissions combine to make its use as an automotive fuel a realistic possibility. *See* ALCOHOL FUEL. James A. Moore

Bibliography. D. R. Fahey (ed.), *Industrial Chemicals via C_1 Processes*, Amer. Chem. Soc. Symp. Ser. 328, 1987; J. Falbe, *Chemical Feedstocks from Coal*, 1982; C. L. Gray and J. Alson, *Moving America to Methanol*, 1985; R. A. Sheldon, *Chemicals from Synthesis Gas*, 1983; Society of Automotive Engineers, *Fuel Methanol: A Decade of Progress*, 1990.

Methods engineering

A technique used by industrial engineers to improve productivity and quality and to reduce costs in both direct and indirect operations of manufacturing and service organizations. Methods engineering is applicable in any enterprise requiring human effort. It can be defined as the systematic procedure for subjecting all direct and indirect operations to close scrutiny in order to introduce improvements that will make work easier to perform while maintaining or improving quality, and will allow work to be done more smoothly, in less time, with less energy, effort, and fatigue, and with less investment per unit. The ultimate objective of methods engineering is increasing profits, but it is also important in improving worker health and safety.

The terms methods engineering, operation analysis, and work design are frequently used synonymously. In most cases, methods engineering refers to a technique for increasing the production per unit of time or decreasing the cost per unit output—in other words, productivity improvement. However, methods engineering entails analysis work at two different times during the history of a product. Initially, the methods engineer is responsible for designing and developing the various work centers

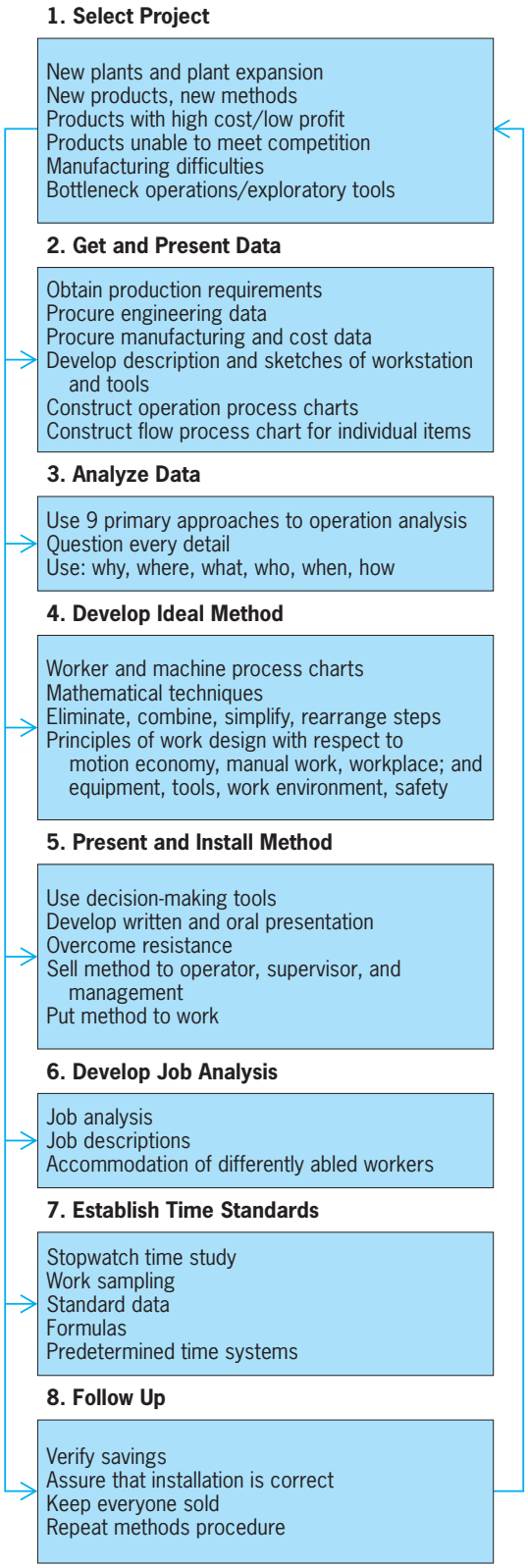

Fig. 1. Principal steps in a methods engineering program.

where the product will be produced. Next, he or she continually restudies the work centers to find a better way to produce the product and/or improve its quality. More recently, this second analysis has been called reengineering or lean manufacturing. *See* LEAN MANUFACTURING; OPERATIONS RESEARCH; PRODUCTIVITY.

Steps. Methods engineers use a systematic procedure (**Fig. 1**) to develop a work center, to produce a product, or to provide a service. This procedure is outlined below. Note that Steps 6 and 7 are not strictly part of a methods study but are necessary in a fully functioning work center.

1. *Select the project*. Typically, projects selected represent either new or existing products that have a high cost of manufacture and a low profit. Also, products that have quality problems and difficulty in meeting competition are logical methods engineering projects.

2. *Get and present the data*. Assemble all the important facts relating to the product or service. These include drawings and specifications, quantity requirements, delivery requirements, and projections about the anticipated life of the product or service. Once all the important information has been acquired, record it in an orderly form for study and analysis. The development of process charts at this point is very helpful.

3. *Analyze the data*. Utilize the primary approaches to operations analysis to decide which alternative will produce the best product or service. These primary approaches include purpose of operation, design of part, tolerances and specifications, materials, manufacturing process, setup and tools, working conditions, materials handling, plant layout, and principles of motion economy.

4. *Develop the ideal method*. Select the best procedure for each operation, inspection, and transportation by considering the various constraints associated with each alternative, including productivity, ergonomic, and health and safety implications.

5. *Present and install the method*. Explain the proposed method in detail to those responsible for its operation and maintenance. Consider all details of the work center to assure that the proposed method will provide the results anticipated.

6. *Develop a job analysis*. Make a job analysis of the installed method to assure that the operator or operators are adequately selected, trained, and rewarded.

7. *Establish time standards*. Establish a fair and equitable standard for the installed method.

8. *Follow up the method*. At regular intervals, audit the installed method to determine if the anticipated productivity and quality are being realized, if costs were correctly projected, and if further improvements can be made.

As a part of developing or maintaining the new method, it is also important to use the principles of work design to ergonomically fit the task and workstation to the human operator. Unfortunately, work design is typically forgotten in the quest for increased productivity. Far too often, overly simplified procedures result in machinelike repetitive jobs for the operators, leading to increased rates of work-related musculoskeletal disorders. Any increases in productivity and reduced costs are more than offset by the

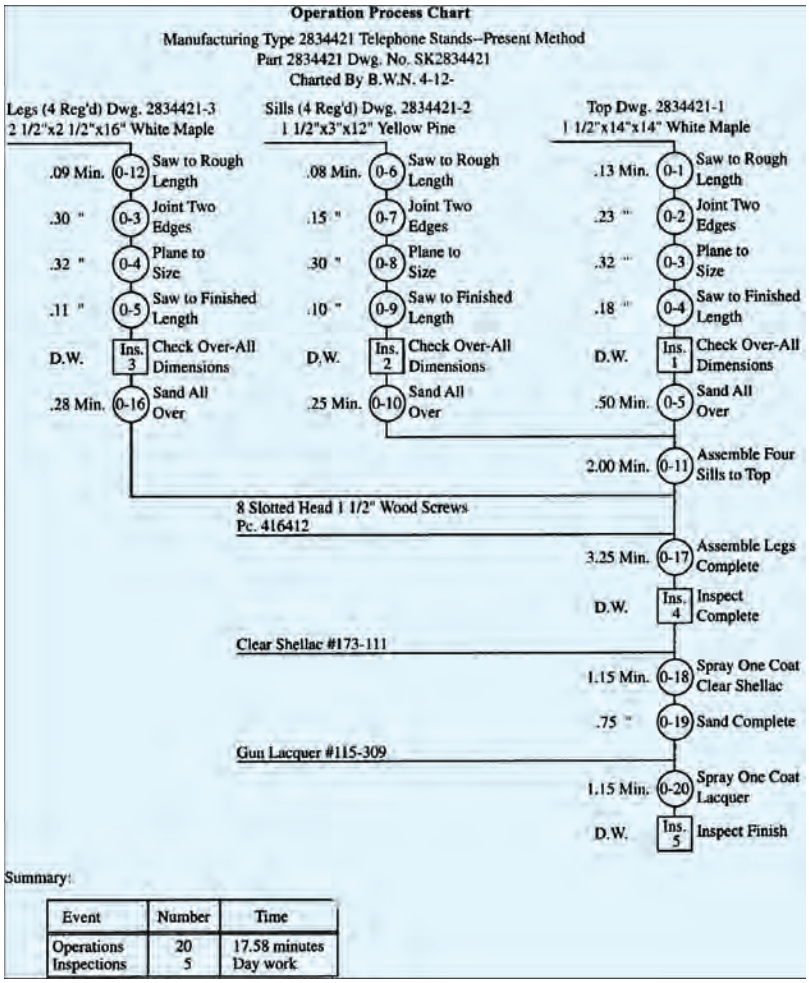

Fig. 2. Operation process chart.

increased medical and workers' compensation costs, especially considering today's ever-escalating health-care trends. Thus, it is necessary for the methods engineer to incorporate the principles of work design into a new method so that the method will be not only more productive but also safe and injury-free for the operator.

Process chart analysis. A process chart is a graphic representation of events occurring during a series of actions or operations. It is most useful in Step 2 (Get and present data) for obtaining an initial understanding of a process, and in Step 3 (Analyze data) and in Step 4 (Develop the ideal method) for analyzing the process and developing a better one. The four process charts most often used are the operation process chart, flow process chart, worker-machine process chart, and operator process chart. The operation process chart and flow process chart are used for analyzing processes involving a number of events or operations. The multiple activity process chart and operator process chart are used to analyze single operations in detail.

An operation process chart shows the chronological sequence of all the operations (and sometimes inspections) in a manufacturing process or service operation (**Fig. 2**). Vertical lines indicate the general flow of the process, while horizontal lines indicate materials or components feeding into the main process. Sometimes, time values for the processes are included. The completed chart helps analysts visualize the present method in hopes of eliminating unnecessary steps. It is particularly useful in connection with plant layout studies.

A flow process chart (**Fig. 3**) represents the sequence of all operations, transportations, inspections, delays, and storages occurring during a process or procedure, and includes the information needed for analysis, such as time required and distance moved. It is used to follow either materials or workers through a process and is an essential tool for materials-handling studies.

A worker-machine process (**Fig. 4**) is used to study, analyze, and improve one workstation at a time. The chart shows the exact time relationship between the working cycle of the person and the operating cycle of the machine. These facts can lead to fuller utilization of both work and machine time, with the goal of having one worker operate more than one machine, known as machine coupling. If both worker and machines are operating on regular cycles, the relationship is known as synchronous servicing, and the

Flow Process Chart Page 1 of 1

Location: Dorben Co.
Activity: Field Inspection of LUX
Date: 4-17-97
Operator: T.Smith Analyst: R. Ruhf
Circle appropriate Method and Type:
Method: (Present) *Proposed*
Type: (Worker) *Material* *Machine*
Remarks:

Summary

Event	Present	Proposed	Savings
Operation	7		
Transport	6		
Delay	2		
Inspection	6		
Storage	0		
Time (min)	32.60		
Distance (ft)	375		
Cost			

Event Description	Symbol	Time (in Minutes)	Distance (in Feet)	Method Recommendation
Leave vehicle, walk to front door, ring bell.	○ ▷ D □ ▽	1.00	75	Call home in advance to reduce waiting delays.
Wait, enter home.	○ ▷ D □ ▽			
Walk to field reservoir.	○ ▷ D □ ▽	.25	25	
Disconnect field reservoir from unit.	○ ▷ D □ ▽	.35		
Inspect for dents, cracks in shroud, cracked glass or missing hardware.	○ ▷ D □ ▽	1.25		This can be done while walking back to vehicle.
Clean unit with approved cleaner and disinfectant.	○ ▷ D □ ▽	2.25		This can be done more effectively at vehicle.
Return to vehicle with empty tank.	○ ▷ D □ ▽	1.00	75	
Unlock vehicle, place empty tank in fixture and connect hardware.	○ ▷ D □ ▽	1.75		
Open valve; begin fill.	○ ▷ D □ ▽	.25		
Wait for tank to fill.	○ ▷ D □ ▽	12.00		Clean unit while being filled.
Check humidifier for proper function.	○ ▷ D □ ▽	.5		Eliminate. No need to do this twice.
Check pressure (indicator).	○ ▷ D □ ▽	.2		
Check reservoir contents (indicator).	○ ▷ D □ ▽	.2		
Return to patient with filled tank.	○ ▷ D □ ▽	1.10	100	
Hook up filled tank.	○ ▷ D □ ▽	1.00		
Check humidifier for proper function.	○ ▷ D □ ▽	.75		
Wait for patient to remove nasal cannula or face mask.	○ ▷ D □ ▽	2.00		
Install new nasal cannula or face mask.	○ ▷ D □ ▽	2.50		
Check flows with patient.	○ ▷ D □ ▽	2.25		
Affix a dated, initialed inspection sticker.	○ ▷ D □ ▽	1.00		Perform this while unit being filled.
Return to vehicle.	○ ▷ D □ ▽	1.00	100	

Fig. 3. Flow process chart, worker type.

optimum number of machines assigned to a worker can be found with the formula

$$N \leq \frac{l+m}{l+w}$$

where N is the number of machines assigned to worker, l is the operator loading/unloading time (both worker and machine interacting), m is the machine running time (independent of worker), and w is the worker time (not directly with the machines).

An operator process chart, sometimes referred to as a two-hand process chart, is a motion-study tool (**Fig. 5**). This chart shows all movements and delays made by the right and left hands, and the relationship between the hands. Each task element performed by the operator can be subdivided into basic motion elements, or therbligs (Gilbreth spelled backward). The therbligs are generally partitioned into effective ones (reach, move, grasp, release, preposition, use, assemble, and disassemble) and ineffective ones (search, select, plan, position, hold, avoidable delay, unavoidable delay, and rest to overcome fatigue). In general, the first eight accomplish necessary work, while the next seven retard the progress of work and should be eliminated when possible. A summary of the effective and ineffective times (corresponding to effective and ineffective therbligs) for left and right hands yields a starting point for methods improvements, the goal being the reduction of the ineffective therbligs. Frequently, therblig analysis can evolve into the setting of standard times using predetermined time systems, most of which have evolved from these same therbligs.

Operations analysis. There are 10 major points of operation analysis that are applied in methods study. Listed in the order in which they usually are considered, they are purpose of operation, design of part, tolerances and specifications, material, process of manufacture, setup and tools, materials handling, plant layout, working conditions, and ergonomics and principles of motion economy.

Where work has not been previously studied in detail, industrial engineers repeatedly find that much of what is being done is unnecessary. Sometimes whole operations are eliminated merely by asking why something is being done and recognizing that

there is no sound answer. The primary technique used to discover such situations is through the analysis of the process charts, considering the methods study approach of purpose of operation.

When the methods analyst considers the design of the part, the objective is to discover manufacturing economics that the design engineer may have overlooked. Considering tolerances and specifications will often suggest better ways of obtaining necessary quality and product reliability. Questioning the materials of which the product is made and the way that the supply materials are used in performing the operations required to make the product will frequently lead to economies. There are usually many different ways to perform a given operation. By questioning the process of manufacture, the methods analyst will consider all alternative processes to determine if the best ones are being used. Software can be developed to evaluate alternative processes based on information such as material being processed, quantity to be produced, geometric configuration of the desired product, and tolerance and finish constraints. Just as there usually are several alternative processes to be considered, there are alternative tooling and setup procedures. The larger the quantity of parts to be produced, the more advanced the tooling should be. As quantity requirements change, existing setup and tooling may become inefficient. Under numerical control, tool setting can be done off-line while the facility is producing. Usually, preassembled fixtures are much better than disassembling and reassembling fixtures during setup. Consequently, the methods analyst should always consider this side of operations analysis. *See* PRODUCTION ENGINEERING.

Materials handling should always be given serious attention. Within a factory or other enterprise, handling of materials adds to the cost of a product without adding to its salable value. Good materials handling provides for the delivery of an adequate inventory of material at the proper time and in the proper condition to the point of use at the least cost. The operation and flow process charts are especially used to give a clear picture of present materials-handling methods and to point the way to new and improved ones. *See* MATERIALS HANDLING.

The physical layout is an important element of an entire production system. The methods analyst will strive to have an effective layout that permits the manufacture of the desired number of products of the desired quality at the least cost. Workstations and facilities, including service facilities, should be arranged to permit the least travel and the most efficient processing of a product with a minimum of handling. Computer programs are used to assist in the development of a sound plant layout. *See* INDUSTRIAL FACILITIES.

Analysis of the working conditions of the immediate environment of each workstation can be another source of savings, productivity, and quality improvement. The methods analyst will study the working conditions of each work center shown on the operation process chart to determine the common possibilities for job improvement. *See* HUMAN-FACTORS ENGINEERING; OCCUPATIONAL HEALTH AND SAFETY.

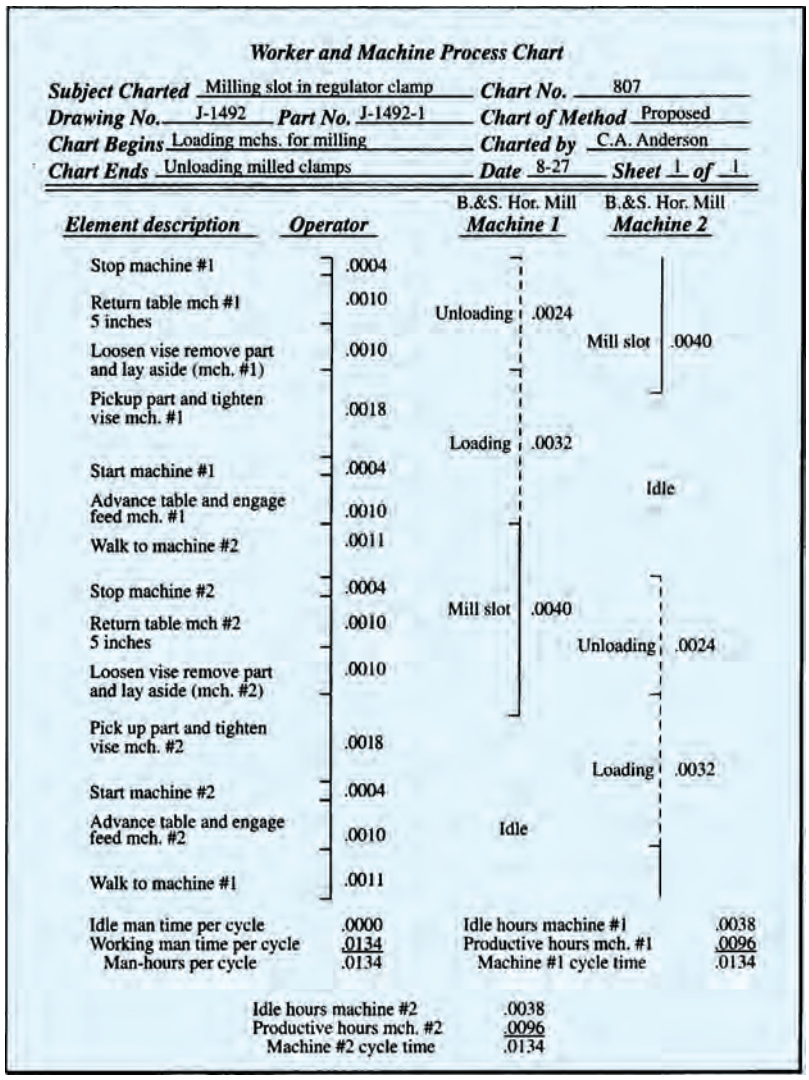

Fig. 4. Worker-machine process chart.

Ergonomics and principles of motion economy involve the study of both manual and eye movements that occur in an operation or work cycle in order to eliminate wasted movements and establish a better sequence of less fatiguing, coordinated movements.

In making a motion study, the methods engineer studies each individual motion in detail and tries to shorten it, combine it with others, or eliminate it altogether. Ineffective motions are identified and replaced, if possible, with effective motions. As a result of this intensive study, easier, more effective, less fatiguing, and quicker ways of doing the work are developed. Motions may be studied individually by direct observation, or they may be observed and timed in groups during the making of a stopwatch time study. When these data are properly arranged on a multiple-activity process chart, possibilities for improvement are often revealed.

In summary, methods engineering can be defined as the systematic close scrutiny of all direct and indirect operations to find improvements, making work easier to perform and allowing work to be done in

Two-hand process chart Page 1 of 1

Operation: Assemble cable clamps	Part: SK-112	Summary	Left hand	Right hand
Operator name and no.: J.B. #1157		Effective time:	2.7	11.6
Analyst: G. Thuering	Date: 6-11-98	Ineffective time:	11.6	2.7
Method (circle choice) (Present)	Proposed	Cycle time =	14.30 sec.	

Sketch:

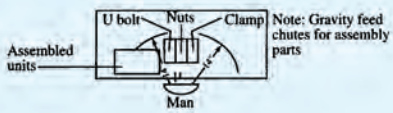

Left hand description	Symbol	Time	Time	Symbol	Right hand description
Get U-bolt (10")	RE G	1.00	1.00	RE G	Get cable clamp (10")
Place U-bolt (10")	M P	1.20	1.20	M P RL	Place cable clamp (10")
Hold U-bolt	H	11.00	1.00	RE G	Get first nut (9")
			1.20	M P	Place first nut (9")
			3.40	U RL	Run down first nut
			1.00	RE G	Get second nut (9")
			1.20	M P	Place second nut (9")
			3.40	U RL	Run down second nut
Dispose of assembly	M RL	1.10	0.90	UD	Wait

Fig. 5. Operator process chart.

less time with less investment per unit. Thus, the main objective of methods engineering is profit improvement.

Andris Freivalds

Bibliography. L. S. Aft, *Work Measurement & Methods Improvement*, Wiley, 2000; S. Konz and S. Johnson, *Work Design: Occupational Ergonomics*, 6th ed., Holcomb Hathaway, 2004; B. W. Niebel and A. Freivalds, *Methods, Standards and Work Design*, 11th ed., McGraw-Hill, 2003; J. A. Tomkins et al., *Facilities Planning*, Wiley, 2003.

Metric system

A system of measurement units used in scientific work and for everyday applications by most countries, with the notable exception of the United States. A prime advantage of the metric system is its international acceptance as a standard of measurement, providing a common measurement language for most of the world's population.

The metric system was developed in France in the late eighteenth century when Louis XVI authorized scientific investigations aimed at the reform of French weights and measures. These investigations led to the development of the first "metric" system. After slow initial acceptance by France, the metric system spread through most of Europe in the nineteenth century and to most of the world in the twentieth century.

The meter-kilogram-second (MKS) and centimeter-gram-second (cgs) systems were major variations of the metric system. Both are now superseded by the SI-metric system adopted in 1960 by the 11th General Conference on Weights and Measures (CGPM). The new SI-metric system, given the official abbreviation SI for Système International d'Unités, is also called the modernized metric system. Most of the original definitions of the SI base units have been superseded as technological advances have permitted more precise determinations. Throughout all these changes, however, the units have remained the same size, only being specified more precisely with each new definition.

SI-metric units have unique symbols, not abbreviations, allowing the units to be recognized in spite of language differences that may affect the spelling of the unit names (such as meter and metre). The symbols represent both the singular and plural of the unit and do not need an added "s" when more than one is meant; nor are the symbols to be followed by periods except at the end of a sentence.

Qualities of the metric system include being decimal and coherent. The decimal nature makes it easy to manipulate these units with the decimal number system used throughout the world. Coherency is a quality of the derivation of SI units from the base units of the system. There is a one-to-one relationship between the equations of physics and the SI units for the quantities expressed in an equation. Metric units and symbols can be manipulated algebraically in the same manner as the physical quantities they represent. There is no need for either multipliers or conversion factors, as is often the case with nonmetric systems of units.

International System of Units. The SI-metric system has seven base units from which all other units are derived (**Table 1**). Two of the base units are named after prominent scientists, A. M. Ampère and Lord Kelvin. These seven units are independent quantities, six of which are now defined in terms of reproducible experimental procedures. Only the kilogram is still defined using an artifact, a platinum-iridium cylinder kept under glass at the International Bureau of Weights and Measures (BIPM) in France, copies

TABLE 1. SI base units

Quantity	Unit	Symbol
Length	meter	m
Mass	kilogram	kg
Time	second	s
Electric current	ampere	A
Thermodynamic temperature	kelvin	K
Amount of substance	mole	mol
Luminous intensity	candela	cd

TABLE 2. SI prefixes

Prefix	Symbol	Multiplication factor	
yotta	Y		10^{24}
zetta	Z		10^{21}
exa	E		10^{18}
peta	P		10^{15}
tera	T		10^{12}
giga	G	1,000,000,000	10^{9}
mega	M	1,000,000	10^{6}
kilo	k	1000	10^{3}
hecto	h	100	10^{2}
deka	da	10	10^{1}
deci	d	0.1	10^{-1}
centi	c	0.01	10^{-2}
milli	m	0.001	10^{-3}
micro	μ	0.000,001	10^{-6}
nano	n	0.000,000,001	10^{-9}
pico	p		10^{-12}
femto	f		10^{-15}
atto	a		10^{-18}
zepto	z		10^{-21}
yocto	y		10^{-24}

of which are available to other countries. For definitions of the SI-metric base units, used as the foundation for all other physical measurements, *see* PHYSICAL MEASUREMENT.

Twenty SI prefixes, from plus to minus the 24th power of 10, cover a range of measurement from the extremely large to the extremely small (**Table 2**). The prefixes are spaced in increments of 10 to the power 3, except for additional powers of +/− 1 and +/− 2. Prefixes for other powers are no longer to be used. The use of the proper SI prefix with a quantity allows values to stay between 1 and 1000. The symbols for prefixes greater than kilo are represented by capital letters, others by lowercase letters, and all symbols are Latin letters except for the Greek mu for micro, 10 to the −6 power. In the case of the kilogram, which for historical reasons already contains a prefix, other prefixes are added to the root word gram or the symbol g.

Derived units in the metric system are formed by combining base and other derived units in a manner similar to the algebraic relationships that link the corresponding quantities. Likewise, the symbols for those derived units are obtained by use of the same relationships. Examples of the many derived units without special names are speed, acceleration, and density. There are, however, 22 SI derived units with special names (**Table 3**). The first two units in the table are dimensionless derived quantities that until 1995 were classified as supplementary units. Seventeen of the derived units are named after prominent scientists. These, together with the two base units ampere and kelvin already mentioned, total 19 SI-metric units that borrow names from scientists. When the metric unit is meant, these names are not capitalized (except for Celsius), but the symbols for these units are always capitalized.

Some other units are in use with SI. These include units that are already used worldwide for time measurement, and for plane angles in degrees, minutes, and seconds (**Table 4**). In addition, there are specially named units for area, volume, and mass that arose from a combination of need and common usage. The liter in particular is different from other metric units in that it has two acceptable symbols, the lowercase l and the uppercase L. The uppercase L, needed to avoid confusion with the number 1, is the preferred symbol in the United States.

Finally, there are yet other units that are to be used only temporarily with SI (**Table 5**). These units could be replaced by metric units, but they are retained because of wide usage. It may be a long time before these units disappear. Complete elimination of these units at this time could be the cause of confusion and possibly dangerous situations, especially for radiation exposure units. *See* PHYSICAL MEASUREMENT.

More recent adoptions. The metric system has become the preferred measurement system for most of the world, starting with Europe and spreading into countries where nonmetric units were previously established. In particular, the Imperial system of weights and measures was used predominantly by British Commonwealth nations until the 1970s. Increases in international trade led to the realization that two competing measurement systems cannot be economically tolerated. One system had to be abandoned in favor of a common and universal measurement language. At first there was some resistance, but eventually the leaders of many of the Commonwealth nations saw the need to change to the metric system.

Notable examples of rather swift and complete transitions to the metric system occurred in Australia, New Zealand, and South Africa. Each country decided on a definite plan and timetable for conversion. In other countries as well, there was a need not only for a decimal measurement system but also for currency systems to match the decimal nature of the base-10 number system used by the entire world. Thus, the change to the metric system in many countries was preceded by adoption of decimal currencies, leaving behind the British system of coinage based on the pound divided into 12 shillings and the shilling divided into 20 pence.

Canada, the United Kingdom, and the United States also embarked on metric transition plans, but with somewhat less success. Many more visible aspects of the metric transition have occurred in Canada and the United Kingdom than in the United States. In the United States, the Metric Conversion Act of 1975 was voluntary in nature, with no conversion deadlines. This is in contrast to the adoption by the United States of the world's first decimal currency when it developed its own monetary system in 1795. Also, the metric system has been legal in the United States since 1866, paving the way for its use alongside conventional units that evolved from, but are significantly different from, those of the Imperial system. In particular, United States liquid measures are not the same as the Imperial namesakes. One of the reasons for adopting metric units was to eliminate major differences that existed between the British and American measurement systems.

The Metric Conversion Act created the U.S. Metric Board (USMB) to coordinate the transition to the

TABLE 3. SI derived units with special names and symbols

Quantity	Name	Symbol	Expression in terms of other SI units
Angle, plane	radian*	rad	m/m = 1
Angle, solid	steradian*	sr	$m^2/m^2 = 1$
Celsius temperature	degree Celsius	°C	K
Electric capacitance	farad	F	C/V
Electric charge, quantity of electricity	coulomb	C	A·s
Electric conductance	siemens	S	A/V
Electric inductance	henry	H	Wb/A
Electric potential difference, electromotive force	volt	V	W/A
Electric resistance	ohm	Ω	V/A
Energy, work, quantity of heat	joule	J	N·m
Force	newton	N	$kg \cdot m/s^2$
Frequency (of a periodic phenomenon)	hertz	Hz	1/s
Illuminance	lux	lx	lm/m^2
Luminous flux	lumen	lm	cd·sr
Magnetic flux	weber	Wb	V·s
Magnetic flux density	tesla	T	Wb/m^2
Power, radiant flux	watt	W	J/s
Pressure, stress	pascal	Pa	N/m^2
Activity (referred to a radionuclide)	becquerel	Bq	1/s
Absorbed dose, specific energy imparted, kerma	gray	Gy	J/kg
Dose equivalent, ambient dose equivalent, directional dose equivalent, personal dose equivalent, organ dose equivalent	sievert	Sv	J/kg
Catalytic activity	katal	kat	mol/s

*The radian and steradian, previously classified as supplementary units, are dimensionless derived units that may be used or omitted in expressing the values of physical quantities.

metric system for Americans. However, after many years of little real progress, the USMB was disbanded in 1982. Metric transition activities of the United States government were handed over to the National Institute of Standards and Technology (NIST) Metric Program office. Private conversion activities are being handled by the U.S. Metric Association, formed in 1916, and dedicated to helping the United States transition to the SI-metric system used by virtually the entire world.

The Omnibus Trade and Competitiveness Act of 1988 strengthened the Metric Conversion Act by making the metric system the preferred system of measurement for the United States government. A Presidential Order issued in 1992 was an additional incentive for federal agencies to convert. However, reasons not to use metric units were still found, such as those contained in a clause stating that metric units need not be used "to the extent that such use is impractical or is likely to cause significant inefficiencies or loss of markets for U.S. firms." As a result, the United States is the least converted of the world's major economies. But the metric system is not dead in the United States, as behind the scenes many consumer products are being designed, built, or packaged in metric.

The units used by the United States have been defined in terms of metric units since 1893. Both the yard and the pound were given equivalent values in terms of metric units by legislation of the U.S. Congress called the Mendenhall Order. In addition, the inch was redefined as exactly 25.4 millimeters in 1959. This change in the definition standardized the nonmetric units in the major inch-using nations of that time: the United States, Canada, and the United Kingdom. Before that change, each country already defined the inch in terms of the metric system as the standard, but differences existed in the sixth

TABLE 4. Units in use with SI

Quantity	Unit	Symbol	Value in SI units
Time	minute	min	1 min = 60 s
	hour	h	1 h = 60 min = 3600 s
	day	d	1 d = 24 h = 86,400 s
	week, month, etc.		
Plane angle	degree*	°	$1° = (\pi/180)$ rad
	minute*	′	$1' = (1/60)° = \pi/10{,}800$ rad
	second*	″	$1'' = (1/60)' = (\pi/648{,}000)$ rad
	revolution, turn	r	$1\ r = 2\pi$ rad
Area	hectare	ha	$1\ ha = 1\ hm^2 = 10^4\ m^2$
Volume	liter†	L, l	$1\ L = 1\ dm^3 = 10^{-3}\ m^3$
Mass	metric ton, tonne‡	t	$1\ t = 10^3\ kg$

*Decimal degrees should be used for division of degrees, except for fields such as astronomy and cartography.
†The symbol L is preferred for use in the United States.
‡Metric ton or tonne is restricted to commercial usage.

TABLE 5. Units in use temporarily with SI

Name	Symbol	Value in SI Units
Nautical mile		1 nautical mile = 1852 m
Knot		1 nautical mile per hour = (1852/3600) m/s
Bar*	bar	1 bar = 100 kPa
Barn	b	1 b = 100 fm^2 = 10^{-28} m^2
Curie	Ci	1 Ci = 3.7×10^{10} Bq
Roentgen	R	1 R = 2.58×10^{-4} C/kg
Rad	rad, rd†	1 rad = 1cGy = 10^{-2} Gy
Rem	rem	1 rem = 1cSv = 10^{-2} Sv

*Use is limited to meteorology.
†When there is risk of confusion with the symbol for radian, rd may be used as the symbol for rad.

decimal place between the conversion factors used by the three countries, enough so that manufacturers using precision measurements were experiencing difficulty in the exchange of inch-based machinery. This problem arose in particular during World War II, but was finally solved by the international agreement to define the inch similarly.

Ever-increasing international cooperation and trade will probably drive future changes to metric units. No countries that have converted to metric have reverted to their former units. Although a cost is involved in the transition to metric, there are also costs in not going metric, such as the costs of dual inventories of products, and the loss of market share for products that are increasingly not accepted in an otherwise metric world. Donald W. Hillger

Bibliography. Bureau International des Poids et Mésures (BIPM), *Le Système International d'Unités (SI): The International System of Units (SI)*, 7th ed., 1998; Institute of Electrical and Electronic Engineers and American Society for Testing and Materials, *Standard for Use of the International System of Units (SI): The Modern Metric System*, IEEE/ASTM SI 10, 2002; National Institute of Standards and Technology, *Guide for the Use of the International System of Units (SI)*, NIST Spec. Publ. 811, 1995; National Institute of Standards and Technology, *The International System of Units (SI)*, NIST Spec. Publ. 330, 2001; U.S. Metric Association, *Guide to the Use of the Metric System (SI Version)*, 15th ed., 2002.

Metzgeriales

An order of liverworts in the subclass Jungermanniidae. Twelve families make up the Metzgeriales, which are also known as the Anacrogynae because of archegonia produced behind the growing apex. The gametophyte plant body is flat, elongated, and usually thallose, with no tissue differentiation or surface pores; less commonly there is a stem with two rows of leaves. The archegonia are generally protected by an involucre formed external to the calyptra and sometimes by additional structures. Capsules dehisce by valves. *See* BRYOPHYTA; HEPATICOPSIDA; JUNGERMANNIIDAE. Howard Crum

Mica

Any one of a group of hydrous aluminum silicate minerals with platy morphology and perfect basal (micaceous) cleavage.

Structure. Sheets can be produced from silicate (SiO_4) tetrahedra by having each tetrahedron share each of its three basal oxygen atoms with a different adjacent SiO_4 tetrahedron. The overall stoichiometry of these sheets will be $(Si_2O_5)^{2-}$. Such sheets can also be considered to be produced by the linking together of double chains of SiO_4 tetrahedra such as occur in the amphibole structure. The structures of trioctahedral phyllosilicates can be thought of as being derived from that of brucite [$Mg_3(OH)_6$], a three-layer structure in which a layer of octahedrally coordinated magnesium ions (Mg^{2+}) is sandwiched between layers of hydroxyl (OH^-) groups. If the sheets described above replace those two OH^- layers in brucite, the mineral talc [$Mg_3(Si_4O_{10})(OH)_2$] is produced; in this structure the Mg^{2+} ions remain octahedrally coordinated, this time by the unshared apical oxygen ions of the sheet-forming tetrahedra and by the remaining hydroxyl ions between the sheets. In the micas, the net negative charge on the sheets is increased by the substitution of one aluminum ion (Al^{3+}) for one of every four silicon ions (Si^{4+}), and this is compensated for by the addition of one alkali metal ion (for each Al^{3+} ion) between the two layers of basal oxygen ions that form the top and bottom of two three-layer sandwiches stacked on top of each other. If potassium ion (K^+) is the alkali, the trioctahedral mica phlogopite [$KMg_3(AlSi_3O_{10})(OH)_2$] results. The structure of dioctahedral micas is related in an analogous way to that of the three-layer structure of gibbsite [$Al_2(OH)_6$], giving such phases as muscovite [$KAl_2(AlSi_3O_{10})(OH)_2$]. In the brittle micas, there is a divalent ion (for example calcium, Ca^{2+}) between the sandwiches. Each mica can exist in several structural types or polytypes. This occurs because the sheets can be rotated relative to one another, and different stacking sequences can result from repeating layers that have been rotated by specific amounts at regular intervals, that is, at every so many layers. Micas crystallize in the monoclinic system, but lepidolite can be trigonal. *See* AMPHIBOLE; CRYSTAL STRUCTURE; LEPIDOLITE; MUSCOVITE; PHLOGOPITE; TALC.

Chemistry. The most common micas are muscovite [$KAl_2(AlSi_3O_{10})(OH)_2$], paragonite [$NaAl_2(AlSi_3O_{10})(OH)_2$], phlogopite [$K(Mg,Fe)_3(AlSi_3O_{10})(OH)_2$], biotite [$K(Fe,Mg)_3(AlSi_3O_{10})(OH)_2$], and lepidolite [$K(Li,Al)_{2.5\text{-}3.0}(Al_{1.0\text{-}0.5}Si_{3.0\text{-}3.5}O_{10})(OH)_2$]. Calcium (Ca), barium (Ba), rubidium (Rb) and cesium (Cs) can substitute for sodium (Na) and potassium (K); manganese (Mn), chromium (Cr), and titanium (Ti) for magnesium (Mg), iron (Fe), and lithium (Li); and fluorine (F) for hydroxyl (OH).

Physical properties. Mica is commonly found as small flakes or lamellar plates without a crystal outline. Muscovite and biotite sometimes occur in thick books, tabular prisms with a hexagonal outline that can be up to several feet across. The prominent

basal cleavage is a consequence of the layered crystal structure. Thin cleavage sheets of micas, particularly muscovite and phlogopite, are flexible, elastic, tough, and translucent to transparent (isinglass). They have low electrical and thermal conductivity and high dielectric strength. Percussion figures may be developed on cleavage plates by striking the surface sharply with a dull-pointed tool. This yields a six-rayed star, the rays of which are parallel to certain crystallographic directions.

Micas have Mohs hardnesses of 2–3 and specific gravities of 2.8–3.2. Upon heating in a closed tube, they evolve water. They have a vitreous-to-pearly luster. Muscovite is colorless to pale shades of brown, green, or gray. Paragonite is colorless to pale yellow. Phlogopite is pale yellow to brown. Biotite is dark green, brown, or black. Lepidolite is most often pale lilac, but it can also be colorless, pale yellow, or pale gray. *See* BIOTITE.

Occurrence. The three major species, muscovite, biotite, and phlogopite, are widely distributed rock-forming minerals, occurring as essential constituents in a variety of igneous, metamorphic, and sedimentary rocks and in many mineral deposits.

Muscovite is found in regionally metamorphosed aluminous rocks that formed under a wide range of physical conditions. In igneous rocks, it occurs in some types of granites, in aplites, and as books in pegmatites. It is a characteristic phase of greisens which are produced when fluorine and other volatiles are introduced from granitic melts into adjacent rocks. Sericite is the name given to fine-grained white mica, usually muscovite. This mineral is a widespread gangue mineral in many hydrothermal ore deposits, either in the deposits themselves or in the adjacent altered wall rocks. Paragonite occurs in schists, gneisses, quartz veins, and fine-grained sediments. Phlogopite is a product of regional metamorphism of impure magnesian limestone. It is also characteristic of mantle-derived kimberlites and inclusions in kimberlites. Of all the micas, biotite occurs in the widest range of geological settings. It is common in the thermally metamorphosed rocks adjacent to granitic intrusions. In regionally metamorphosed rocks, it occurs in schists with chlorite, garnet, staurolite, kyanite, and sillimanite, although not all of these simultaneously. In plutonic igneous rocks, biotite is most common in intermediate and acid rocks, but it even occurs in some norites. Biotite books are found in pegmatites. Lepidolite is the most common lithium-bearing mineral and occurs almost exclusively in pegmatites with beryl, topaz, tourmaline, and other lithium minerals such as spodumene and amblygonite. It is also found occasionally in granites and aplites. *See* APLITE; GRANITE; ORE AND MINERAL DEPOSITS; PEGMATITE.

Uses. Commercial mica is of two main types: sheet, and scrap or flake. Sheet muscovite, mostly from pegmatites, is used as a dielectric in capacitors and vacuum tubes in electronic equipment. Lower-quality muscovite is used as an insulator in home electrical products such as hot plates, toasters, and irons. Scrap and flake mica is ground for use in coatings on roofing materials and waterproof fabrics, and in paint, wallpaper, joint cement, plastics, cosmetics, well drilling products, and a variety of agricultural products. *See* ELECTRIC INSULATOR; SILICATE MINERALS.

Lawrence Grossman; Steven Simon

Bibliography. W. A. Deer, R. A. Howie, and J. Zussman, *An Introduction to the Rock-Forming Minerals*, 2d, ed., 1992; C. Klein, *Manual of Mineralogy*, revised 22nd ed., 2001; J. J. Papike, Chemistry of the rock-forming silicates: Multiple-chain, sheet, and framework structures, *Rev. Geophys.*, 26:407–444, 1988.

Mice and rats

The names associated with a great number of species of mammals in a number of different families of the order Rodentia. The rodents constitute the largest order of mammals, and they presently include 2277 species, or about 42% of the known mammalian species of the world. Most of the smaller rodents are called rats or mice. There is no fundamental taxonomic difference between mice and rats; they differ only in size. (Smaller species are generally called mice, larger ones rats.) Both mice and rats are often in the same family, and there are mice and rats in a number of different rodent families. *See* MAMMALIA; RODENTIA.

Some of the rats are of great economic significance to humans, particularly the Norway rat, *Rattus norvegicus*, and the black or roof rat, *R. rattus*. In addition to harboring many diseases transmissible to humans, such as bubonic plague, endemic typhus, rat-bite fever, and infectious jaundice, infested rats can transmit trichinosis to swine. *See* PLAGUE.

Some species of mice and rats are used for research in biology and medicine. In addition to their use in studying the mechanisms of genetics, they are important in the study of carcinogenesis, effects of drugs, and virology. They are also important experimental animals in studying cell physiology, such as for cell and tissue culture research, and in animal behavior.

There is much current discussion and disagreement concerning the higher classification of rodents. The present classification is as follows. The Rodentia are divided into five suborders, and there are mammals called mice and rats in four of the five. All the animals in one large suborder, the Myomorpha, are rats and mice. The four suborders with mice and rats (Sciuromorpha, Castorimorpha, Myomorpha, Hystricomorpha) are indicated below, with the families including animals classed as mice and rats listed below each of the suborders.

Sciuromorpha

The Sciuromorpha are squirrel-like rodents, including the dormice.

Gliridae. The dormice make up this family. They are one of the oldest living families of rodents,

appearing in Eocene deposits, and they are widespread in the Old World, their 28 species occurring in much of Eurasia and Africa below the Sahara. Dormice are squirrel-like, but unlike most squirrels are generally nocturnal. They live in hollow limbs, rocky crevices, in burrows of other species, and in buildings. They could have been considered with squirrels, but are included here since they are called "mice." They have long, mostly rather bushy tails. Most have large eyes and rounded ears. Most are good climbers and have short, curved claws adapted for climbing. The dental formula is I 1/1 C 0/0 Pm 1/1 M 3/3 = 20. They are generally smaller than squirrels, but larger than mice. The fat or edible dormouse of Europe, *Glis glis*, is about 140–240 mm (5–9 in.) in total length and weighs about 70–180 g (2.5–6.3 oz), which is much smaller than the smallest North American squirrels. This species has been introduced into England. It lives in forests or orchards, usually in cavities in trees, but also it will live in many other protected areas. Some dormice hibernate, but wake occasionally and eat stored food. Food consists of fruit, nuts, insects, eggs, and even small vertebrates, which is similar to that of North American flying squirrels (*Glaucomys*). Most species produce two litters per year, but *Glis* probably produces only one because of the long hibernation, that is, short active period. The meat of *Glis* is edible as the name suggests, and has been a gourmet food in parts of Europe, although the species has become rare in many areas because of destruction of forests.

The African Dormice, subfamily Graphiurinae, with 14 species, are all in the genus *Graphiurus* and all occur in Africa. *Graphiurus murinus* looks much like a deermouse (*Peromyscus*) but with a bushy tail. Some African dormice in southern Africa hibernate. They make twittering sounds and a fairly loud "shriek." The subfamily Leithiinae includes 12 Eurasian and mideastern species: *Chaetocauda*, the Sichuan dormouse of China; *Dryomys*, forest dormice, 3 species in Turkey, Pakistan, and Russia; *Eliomys*, the garden dormice, with 3 species; *Muscardinus*, the hazel dormouse, *Myomimus*, the mouse-tailed dormice, 3 species; and *Selevinia*, the desert dormouse (the latter 4 genera are all of Eurasia). The subfamily Glirinae includes only two species: *Glirulus japonicus*, the Japanese dormouse, and *Glis glis*, the fat or edible dormouse of Eurasia.

Castorimorpha

This suborder comprises the single family Heteromyidae.

Heteromyidae. The Heteromyidae comprise kangaroo rats and mice and pocket mice, with 6 genera and 60 species. The dental formula is I 1/0 C 0/0 Pm 1/1 M 3/3 = 18. The 19 species of kangaroo rats (*Dipodomys*) occur in desert regions of North America. They have elongate hindlegs and hop rather than run, thus resembling a kangaroo. The short front legs are used for holding food, and the long hairy tail acts as a balancing organ or prop when the animal hops or rests. These rodents collect seeds into shallow sand-covered pits for drying by the heat of the sun; the seeds are then carried into the burrows for winter use. These animals do not usually drink water, but obtain fluid from food, an adaptation for water balance in desert environments. Over a period of months with only dry food, there is no change in the water content of their bodies. These species lack sweat glands, the feces have a low water content, and the kidneys are efficient, with urine highly concentrated and excreted in small amounts. Water is conserved and body heat is regulated by these adaptations. When fed high-protein diets with the resulting formation of large amounts of urea, these animals will drink water, even seawater which would poison most animals. However, the kidneys can handle the high salt concentration.

Twenty-six species of pocket mice (*Perognathus*, 9 species, and *Chaetodipus*, 17 species) occur in the United States, with the center of distribution being in the southwest. The ears are small, the tail is moderately long, and a fur-lined cheek pouch opens externally on each side of the mouth. The hindlimbs and hindfeet are moderately enlarged; there are four digits on the forefeet, and five on the hindfeet. Some species are the plains pocket mouse (*P. flavescens*), the Great Basin pocket mouse (*P. parvus*), and the desert pocket mouse (*C. penicillatus*).

There are two species of kangaroo mice (*Microdipodops*), small rodents restricted to the desert areas of the western United States. These are *M. megacephalus*, the dark kangaroo mouse, and *M. pallidus*, the pale kangaroo mouse. Their hindlegs are long and forelegs relatively short, and the mice move like a kangaroo, hopping on the hindlegs and using the tail as a balancing organ. They do not drink regularly, but obtain sufficient moisture from their food.

There are five species of spiny pocket mice (*Liomys*). *Liomys irroratus* enters the United States from its center of distribution in the thorny bushlands of the Rio Grande Valley in Mexico. The ears are small, and the tail is as long as, or longer than, the body. The fur is coarse and consists principally of stiff, flattened spines.

Myomorpha

This suborder includes the families Dipodidae, Platacanthomyidae, Spalacidae, Callomyscidae, Nesomyidae, Cricetidae, and Muridae.

Dipodidae. The Dipodidae now includes three subfamilies: the jumping mice, birchmice, and jerboas.

The jumping mice for years were in their own family, Zapodidae, but now are considered as a subfamily, Zapodinae, of the Dipodidae. The Zapodinae contains three genera and five species. All of the species are entirely North American, except *Eozapus setchwanus*, which is confined to eastern Asia. The meadow jumping mice (*Zapus*, 3 species) occur throughout much of North America. Jumping mice have long tails for balance, grooved upper incisors, and colorful fur. These rodents are adapted for

jumping, with long legs and large feet. *Napaeozapus* can jump up to 10 ft (3 m), *Zapus* usually not more than 2–3 feet (0.6–0.9 m). However, the meadow jumping mice (genus *Zapus*) more often run for cover or move by short hops rather than long leaps. By day, jumping mice hide in clumps of vegetation. Nests may be underground, in a hollow log, or otherwise protected. They hibernate for 6–9 months (the duration varies by species, locality, and elevation) in a nest at the end of a burrow in a bank or other raised area. Gestation is about 17 or 18 days, longer if the female is lactating. Meadow jumping mice eat many seeds, especially of grasses, but species of seeds vary greatly during the season. Subterranean fungi related to *Endogone* form about 35% of the volume of food in the woodland jumping mouse, and about 12% in meadow jumping mice. Among animal foods, the most important are moth larvae (caterpillers, often cutworms) and snout and ground beetles.

Zapus has 18 teeth and the dental formula is I 1/1 C 0/0 Pm 1/0 M 3/3 in. *Napaeozapus insignis*, the woodland jumping mouse, occurs in Canada and eastern United States. It has only 16 teeth; it lacks the upper incisors that appear as a small tooth at the front end of the molariform toothrow in *Zapus*.

The feet of birchmice (subfamily Sicistinae) are only slightly enlarged, and their upper incisors lack the grooves of jumping mice. Their tail and legs are shorter than those of jumping mice, but jumping is their major mode of locomotion. When climbing into bushes, they use their outer toes to cling to vegetation. Birchmice make nests of herbaceous vegetation underground, where they hibernate for about half of the year. Birchmice are active primarily at night, and their main foods are seeds, berries, and insects. Gestation is probably about 18–24 days, with parental care lasting another 4 weeks.

Jerboas (subfamily Dipodinae) are built for jumping. Their hindlegs are at least four times as long as their front, which gives them the ability to escape from predators, although they will move by slow hops. Only the hindlegs are used in moving, allowing the front feet to gather food. Some jerboas can jump 5–10 ft (1.5–3 m), and desert jerboas, *Jaculus*, can jump nearly 3 ft (0.9 m) vertically. Jerboas use their long tails as props when standing upright and as balancing organs when jumping. Jerboa species in sandy areas have tufts of hairs on the undersides of the feet which serve as "snowshoes" to maintain traction and to kick sand backwards when burrowing. Other tufts of hair keep sand out of their ears, and some jerboas, for example, *Jaculus*, have a fold of skin which can be pulled forward over the nose when burrowing. Jerboas feed primarily on seeds, sometimes on succulent vegetation, and also on insects. Jerboas, like pocket mice, manufacture "metabolic water" from food. Some jerboas hibernate during winter, and some enter torpor during hot or dry periods. Northern jerboas mate shortly after emergence from hibernation, but most female jerboas usually produce two litters per year of two to six young. *See* JERBOA.

Platacanthomyidae. The members of this family are the spiny dormice. There are two species, and they are similar to dormice. *Platacanthomys lasiurus* of southern India weighs about 75 g (2.6 oz). Its dorsum is covered with sharp flat spines intermixed with thin underfur. The basal part of the tail is scaly with few hairs and the distal portion forms a brush. It has a pointed snout, medium-sized eyes, and prominent ears. The hindfeet are broad and strong. There are no premolars; the tooth formula is the same as that of the murids, 1/1 0/0 3/3 0/0. Spiny dormice live in rock clefts on rocky hills and in hollows in forested valleys. The long tail is useful as a balancing organ as the animal leaps about in trees. Natives in some areas call it the "pepper rat" because it commonly feeds on peppers, but it also eats fruits, seeds, grains, and roots. The Chinese pygmy dormouse, *Typhlomys cinereus*, occurs in heavy forests in mountainous regions in China and Vietnam. It is mouselike with a long thin tail, prominent ears, and relatively small eyes. The tooth formula is as in *Platacanthomys*. Little is known of its habits, but natives claim that cats will not eat it.

Spalacidae. The members of this family are the mole rats. The family Spalacidae includes six genera: *Eospalax* and *Myospalax*, the mole rats or zocors, with five species in Russia, Mongolia, and China; two genera of bamboo rats, *Cannomys* and *Rhizomys*, with four species in Asia; the blind mole rats (*Spalax*) with 13 species in the Russian region and the mideast; and *Tachoryctes* with 11 species of African mole rats.

Mole rats are short and chunky with short powerful legs. The forefeet have long, heavy claws for digging which are doubled under when the animals walk. They have no external ears and the eyes are tiny and hidden in the fur. Mole rats are fossorial (adapted for digging) and can dig into the ground like moles. They are found in a variety of habitats. Their burrows are very long and include a deep nest chamber, a food storage chamber, and an area for defecation. There are also one to four food tunnels that approach the surface below the food source. Unlike true moles, which feed on animal matter, mole rats feed on plant material such as roots, tubers, and plant stems. They may cause damage in potatoes and sugar beets. Some species of *Spalax* form "breeding mounds" in wet weather in fall and winter. These are elaborate structures that are usually about 160 cm long, 135 cm wide, and 40 cm in height (roughly 5 × 4 ft and 1.3 ft high). The African mole rats have stiff tactile hairs on the face. These animals resemble American pocket gophers. They live in open areas with much rain. They also have burrows and mounds, but burrowing is apparently done with the incisors, with the hindfeet used as braces. The forefeet are used to push the dirt under the animal. Then when ample soil has accumulated, the animal kicks it on back with its forefeet, and finally turns around and pushes the dirt, bulldozer style, using one side of the face and one

forefoot, behaviors all similar to those of true moles. Reproduction may occur throughout the year, but is concentrated during the wet season. *See* BURROWING ANIMALS; MOLE (ZOOLOGY).

Bamboo rats (*Rhizomys* and *Cannomys*, including four species) also resemble American pocket gophers (but lack cheek pouches). They have long naked tails, small eyes and ears, and large orange upper incisors outside the lips. They are generally found in bamboo thickets. They have several burrows and come above ground to feed on their main food—bamboo. They may emerge at night and even climb the bamboo. They will eat other plant foods, and do not drink water if there is adequate moisture in their food. Some species will feed on cultivated tapioca and sugarcane, and in turn are captured by natives for food.

Callomyscidae. This family includes one genus, *Callomyscus*, with eight species, occurring in the Middle East, Syria, and Azerbaijan. They are best called "mouselike hamsters" or just *Callomyscus*. These animals appear most like jerboas or heteromyids, or even *Peromyscus*, with long tufted tails, big ears, and big eyes. There are no cheek pouches. Active mainly at night in summer, they eat seeds, flowers, and leaves, but will readily eat animal food.

Nesomyidae. The nesomyids include 38 species in 12 genera from Africa called the pouched rats (three genera, eight species), swamp mouse (1 species), African climbing mice (3 genera, 14 species), African desert mouse (1 species), African tree mouse (1 species), African fat mice (1 genus, 8 species), African white-tailed rat (1 species), and the pygmy rock mice (4 species). In addition there are 23 species in nine genera in this family all occurring in Madagascar. These include the antsangy (1 species), short-tailed rats (2 species), tufted-tailed rats (10 species), the voalavoanala, votsovotaa, and voalavo (1 species each), the big-footed mice (2 species), and the nasomys (3 species). The African giant rat, *Cricetomys gambianus*, is large, measuring as much as 32 in. (81 cm) from nose to tail tip. It has large cheek pouches and is at times referred to as the Gambian pouched rat. It is a vegetarian and occurs in forest regions, where it nests near the base of trees. It harbors a large, wingless, ectoparasitic cockroach, *Hemimerus talpoides*, which is similar to a louse and is apparently host-specific.

Cricetidae. The major families of mice and rats (Muridae with 730 species and Cricetidae with 130 genera and 681 species), like the higher taxonomic categories of rodents, have also been the subject of much debate and rearrangement. In the early part of the last century, most of the muroid mice and rats were placed in two families, the Cricetidae or "New World rats and mice," and the Muridae, or "Old World mice and rats," although the voles were included in Cricetidae and occurred in both New and Old Worlds. However, in the latter part of the last century almost all of these animals were placed in the Muridae. Today, once again, most of the New World mice are placed in the family Cricetidae, but they are now divided into different subfamilies. The present arrangement of the cricetids follows Carleton and Musser with the major North American genera indicated to help understand the new system. The Cricetidae currently is a very large family of mice, rats, and voles, and includes six subfamilies, as follows.

The Arvicolinae are the voles, lemmings, bog lemmings, and muskrats and were long known as "microtines." They include the typical "meadow mice." Most are burrowers, many with short tails, and have relatively small ears and eyes. Most of these animals feed on green vegetation, a food item that wears teeth down very rapidly. They have evolved two major adaptations for dealing with the tough cellulose in plant food: rootless teeth and flat grinding surfaces on the molariform teeth. Rootless teeth tend to grow throughout the life of the animal and therefore continue to grow from below as they are worn off at the crowns. The flat surfaces help in grinding the food. *See* MUSKRAT.

There are 28 genera and 151 species currently in the Arvicolinae. Arvicolinae are holarctic (that is, they inhabit northern parts of the Old and New Worlds). There are 19 Eurasian genera, including mountain voles (*Aticola*, *Neodon*, *Phaimys*), water voles (*Arvicola*), snow voles (*Chionomys* and *Dineromys*), mole voles (*Ellobius*), Sichuan voles (*Volemys*), and wood lemmings (*Myopus*). The genera *Arborimus*, *Lemmiscus*, *Neofiber*, *Ondatra*, *Phenacomys*, and *Synaptomys* are North American.

Four genera occur in both eastern Eurasia and in North America, *Microtus* (62 species), *Myodes* (was *Clethrionomys*, 12 species), *Lemmus* (5), and *Dicrostonyx* (8). The last two genera are the brown and collared lemmings, which occur in the far north.

Microtus is the largest genus of voles (often called meadow mice), with 62 species. These animals are usually brown and most often live in grassy fields and feed almost entirely on green vegetation. Many species leave characteristic piles of cuttings of grass or other plant stems often including the rachis and seeds of grasses. Some of these animals are very prolific, producing a number of litters per year.

The Cricetinae include the hamsters, with 7 genera and 18 species all of Eurasia. *Cricetus cricetus* is the common hamster often kept as a pet. It has a very short tail, nearly hairless. It lives in burrows and hibernates in winter. It feeds on seeds, roots, green parts of plants, and other wild and cultivated plant items and stores food for use in winter. In addition hamsters will feed on insect larvae and frogs. Hamsters are solitary, with only one animal per burrow. Although they may produce litters throughout the year in captivity, they normally produce two litters of 4–12 young per year in the wild.

The subfamily Lophiomyinae includes one species, the maned rat, *Lophiomys imhausi*, found in Africa often in forests, but not restricted to them. The maned rat is a large species with long hair, and

a long furry tail. The animal is more like a guinea pig than rat. It is variable in color, predominantly gray, but often with black and white spots. The tip of the tail is white. There is a mane of erectile hair from the top of the head to the first quarter of the tail. It is nocturnal and feeds on leaves and tender shoots.

The Neotominae (along with the Arvicolinae) are the big group of North American rats and mice, including 16 genera and 124 species. Mice and rats in this suborder spend more time aboveground than do the voles. Consequently they are more aware of their environment and generally have longer tails, bigger ears, and bigger eyes than the voles and allies.

Seven of the genera occur in North America, the woodrats (*Neotoma*, 22 species), golden mice (*Ochrotomys*, one), grasshopper mice (*Onychomys*, 3), pygmy mice (*Baiomys*, 2), deer mice (*Peromyscus*, 56), Florida deer mice (*Podomys*, 1), and harvest mice (*Reithrodontomys*, 20).

Woodrats, *Neotoma*, are the largest of this group and are often called packrats, as they have a habit of collecting things: food items, pieces of bone, leaves, sticks, etc. They will often carry a bright object to their nests, frequently replacing it with pebbles or other material; because of this practice they are also called trade rats. Most species have long thin tails, except the bushy-tailed woodrat, *Neotoma cinerea*, a western species, common in the Rocky Mountain region, while *N. magister* occurs in the northeast south to the Tennessee River. *Neotoma floridana* ranges along the South Atlantic and Gulf costs, north to the Tennessee River, but interestingly to extreme southern Illinois. Wood rats are found in timbered areas and desert regions. Most are nocturnal.

The golden mouse, *Ochrotomys nuttalli*, is found in the southeastern United States. It is a beautiful golden-cinnamon above, white below. It is highly arboreal, and often climbs trees to 30 ft or more. Its long prehensile tail is used for balance and support. This species often occurs in dense brushy undergrowth such as honeysuckle, greenbriar, grapes, or, when present, in Spanish moss. It makes large bulky nests in the undergrowth and is gregarious.

Three species of grasshopper mice, *Onchrotomys*, occur in the grassland and desert areas of the western United States and Mexico.

Baiomys, the pygmy mice, include two species, *B. musculus* of Mexico and *B. taylori* of the southwestern United States. Pygmy mice are small and short-tailed. They occur in grassy or brushy areas.

There are numerous species, races, varieties, and subspecies of deer mice (*Peromyscus*). They feed on seeds and insects. One of the common species is *P. maniculatus*, the deer mouse, with about 65 subspecies found throughout the United States (see **illustration**). These rodents are small with moderate to long tails and large ears. Another is *P. leucopus*, the white-footed mouse, which occurs in eastern woodlands. The Florida deer mouse, *Podomys floridana*, is very similar to *Peromyscus*, but is much larger. It occurs only in Florida. It seldom makes its own burrow, but usually uses that of the gopher tortoise or pocket gopher.

***Peromyscus maniculatus*; deer mouse. (*Photo by Glenn and Martha Vargas; © 2002 by California Academy of Sciences*)**

Harvest mice (*Reithrodontomys*) are small to very small rodents with a long tail, prominent ears, and long, narrow hindfeet. Five of the 17 living species are found in the United States. The eastern harvest mouse (*R. humulis*) occurs in the deciduous forests of the southeastern states, while *R. montanus* is common in the central grasslands. The salt marsh harvest mouse (*R. raviventris*) is found only in the salt marshes near San Francisco Bay.

Many of the genera are primarily Mexican: *Habromys* (6 species), *Hadomys* (1), *Megadontomys* (3), *Nelsonia* (2), *Neotomodon* (1), and *Osgoodomys* (1). Farther to the south are *Isthmomys* (2 species) and *Scotinomys* (2).

Included in the subfamily Sigmodontinae are 74 genera and 374 species, mostly South American, although a few genera reach into Central America, and two genera, *Sigmodon*, the cotton rats, and *Oryzomys*, the rice rats, occur in the United States. Both of these genera are more common in Central and South America. In much of the northern hemisphere, voles are the dominant small herbivores, whereas in the deep south, cotton rats, *Sigmodon*, fill a similar role. Cotton rats are compact-bodied, coarse-haired rodents, with short rounded ears. *Sigmodon hispidus* is the common cotton rat throughout the southern United States, especially in grassy fields. *Sigmodon* will also feed on eggs and chicks of ground nesting birds such as quail, and will eat crayfish and fiddler crabs. Cotton rats are very prolific. Breeding activity occurs from late winter through late fall, and sometimes throughout the year in the more southern areas. Gestation is about 27 days and up to 15 young may occur in a litter.

The marsh rice rat, *Oryzomys palustris*, lives in the southeastern United States, and it and its relatives occur through east Texas and south through much of Central America. Rice rats usually live in wet marshy areas. They are good swimmers, and feed heavily on rice and other aquatic grasses, but they also will eat insects, and small crabs. They may produce litters of up to six offspring from February to November.

The Galápagos Islands have a relatively impoverished fauna as they are so far from the mainland (about 600 miles west of Ecuador). However, rice rats (*Oryzomys*) probably reached the Galápagos Islands from South America some 3 to 3.5 million

years ago, differentiated into *Nesoryzomys*, and, like "Darwin's finches," spread to the different islands and further differentiated, *N. darwini* on Santa Cruz Island (probably extinct), *N. fernandina* on Fernandina Island, *N. indefessus* (probably extirpated in part of its range), and *N. swarthi* on San Salvador Island. The Galápagos giant rat (*Megaoryzomys*) also occurred on the Galápagos, but it too is extinct.

There are many other Sigmodontinae in the South American region, the "Akodonts" *Abrothrix* (8 species), *Akodon* (40), *Bibimys* (3), *Chelemys* (3), *Deltamys* (1), *Geoxus* (1), *Juscelinomys* (3), *Lenoxus* (1), *Nectomys* (9), *Notiomys* (1), *Pearsonomys* (1), *Phaenomys* (1), *Podoxymys* (1), *Thalpomys* (1), and *Thaptomys* (1). Likewise, the "Pericotes" include the genera, *Andalgalomys* (3), *Auliscomys* (3), *Galenomys* (1), *Graomys* (4), *Loxadontomys* (2), *Paralomys* (1), and *Phyllotis* (13).

The Tylominae, vesper and climbing rats, occur in Mexico and Central America. There are four genera and 10 species.

Muridae. The members of this family are the Old World rats and mice, a huge family (150 genera and 730 species). Except for some introduced species, all occur in the Old World. There is a great biodiversity in this group. They are placed in five subfamilies, including Deomyinae (42 species in four genera), Gerbillinae (105 species in 14 genera), Leimacomyinae (1 species, 1 genus), and Otomyinae (23 species in 3 genera). The rest are in the subfamily Murinae, which includes *Mus* and *Rattus*, which have been introduced into much of the rest of the world. The common house mouse is one of the oldest known species of domestic rodent pests. They begin to breed at 3 months of age, have a gestation of about 3 weeks, and have 4–6 litters of four to eight young per year. Adults have a pointed snout, compact body about 3 in. (8 cm) long, and an equally long tail. The ears are fairly large, as are the legs. These animals have 16 teeth and the dental formula I 1/1 C 0/0 Pm 0/0 M 3/3. House mice are often thought to exist primarily in buildings. However, they are often very common in croplands of the midwestern United States, but require much ground cover. They disappear when the crops are harvested. The other main small mammal species in these croplands is the prairie deermouse, *Peromyscus maniculatus bairdii*. However, this latter species uses the soil as its cover and is the only species that remains when the crops are harvested and even when the ground is plowed. House mice are omnivorous, and will eat grains and other vegetable foods. However, they eat huge quantities of weed seed, foxtail grass, moth larvae such as army worms, and many others. Of the 44 species of *Mus* known, only one species occurs in the United States, and it has become wild in some parts of the country. The familiar white laboratory mouse is an albino form of the house mouse and is used extensively in laboratory research.

Rats (*Rattus*) are usually active at night, feeding on nearly every type of food. These animals have the same dental formula as *Mus*. The teeth are modified for gnawing, with sharp and chisellike incisors that grow continually. The genus *Rattus* contains about 137 species with many varieties and races throughout the world. Fecundity is high with as many as six litters of six to eight young per year; the gestation period is about 21 days; breeding begins at about 3 months. Average lifespan is 2–3 years. Rats have a maximum weight of slightly over 1 lb (0.5 kg). The thin tail is covered with overlapping scales and nearly devoid of hair. The ears and eyes are relatively large, and cheek pouches are lacking. Many species live in burrows, others infest dwellings.

Hystricomorpha

The Hystricomorpha are referred to as the South American rodents, although some of the species occur in Africa.

Octodontidae. The octodonts are South American. Most are not called rats but they are ratlike, and two genera each with two species are called rats, *Octomys* (Viscacha rats), which look like Dipodidae, and *Aconaemys* (rock rats), which look like voles. They make complex burrows and feed on green vegetation.

Abrocomidae. The members of this family are the chinchilla rats. There is one genus and two living species in this family, *Abrocoma bennetti* and *A. cinerea*, both South American. They look similar to chinchillas, but the head and ears are longer, giving them a ratlike appearance. Chinchilla rats have very long intestinal tracts and a huge cecum. These characteristics indicate that they are herbivorous.

Echimyidae. The members of this family are the spiny rats. This group contains 17 genera and 69 species and thus helps fill the mouse and rat niche of South America. They include various sized animals varying from 21 g (0.7 oz), which would easily be called a "mouse," up to 450 g (1 lb). Tail length varies greatly, from less than one-quarter of the body length to more than the body length, and they otherwise exhibit the great morphological variation one would expect from rodents living in a variety of habitats with a variety of niches. An adult female of one species of echimyid weighed 640 g (1.4 lb), which would be a very large rat, and other species in the family can weigh even more. Although the family, collectively, is referred to as the spiny rats, not all are called "rats." For example, members of the genus *Isothrix* are called toros, *Dactylomys* are coro-coros, and *Chaetomys* are thin-spined porcupines. Members of this family are said to be vegetarian, but most are little known and it is likely that some species would be omnivorous or even primarily insectivorous.

Thryonomyidae. The members of this family are the cane rats. The one living genus, *Thyronomys*, contains two species, both in Africa. These are very large for rats, weighing from 4 to 9 kg (9–20 lb). One of the species is semiaquatic (*T. swinderianus*) and lives in marshes, whereas the other lives on dry ground. Both are herbivorous.

Petromuridae. There is but one species, *Petromuris typicus*, the Dessie rat, and it is African. The

fur is soft, but underfur is absent. Hairs grow in clusters of three to five, and these animals have long black whiskers. They live in tiny crevices in rocky areas on hillsides, and their bodies can be greatly compressed.

Bathyergidae. These are the African mole-rats. This family contains 5 genera and 16 species, and all are restricted to Africa. They are molelike in having small ears and eyes and spend all or most of their time in underground burrows. Also, they throw up mounds to get rid of excess dirt when extending their burrows. Their front legs may be enlarged somewhat, but they do not have developed forelimbs as in true moles. Also, true moles feed primarily on earthworms and other invertebrates, whereas mole-rats feed mainly on bulbs and roots and store some in underground chambers for later use. Zech's mole rat (*Cryptomys zechi*) has large protruding incisors, short legs, powerful feet, and a reduced tail, features which make it well adapted for its burrowing existence. Like many desert species, it obtains sufficient water from food.

The strangest and smallest member of this family is *Heterocephalus glaber*, the naked mole-rat of Africa, measuring about 4 in. (10 cm). It lives in hot sandy regions of Africa. It is not totally naked as there are a few light-colored but long hairs on its body and tail and it has well-developed vibrissae. External ears are lacking, and it is blind. It weighs about 30–80 g (1–2.8 oz). This species is colonial and is completely fossorial, digging a complex burrow system with a mound above. This species has the poorest temperature regulation of any known mammal. Another particularly interesting trait is that they will form a chain to dig dirt from deeper areas, with one animal digging, but another animal pushes the soil backward all the way to the surface. The other animals, in turn, will push the soil to the surface, with the "pusher" taking its place at the end of the line. The animals in the chain straddle the pusher as the earth is moved to the surface. John O. Whitaker, Jr.

Bibliography. M. D. Carleton and G. G. Musser, Order Rodentia, pp. 745–752 in D. E. Wilson and D. M. Reeder (eds.), *Mammal Species of the World: A Taxonomic and Geographic Reference*, 3d ed., Johns Hopkins University Press, Baltimore, 2005; R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, Baltimore, 1999; J. O. Whitaker, Jr., and W. J. Hamilton, Jr., *Mammals of the Eastern United States*, Cornell University Press, Ithaca, 1998; D. E. Wilson and S. Ruff, *The Smithsonian Book of North American Mammals*, 1999.

Micelle

A colloidal aggregate of a unique number (50 to 100) of amphipathic molecules, which occurs at a well-defined concentration called the critical micelle concentration. In polar media such as water, the hydrophobic part of the amphiphiles forming the micelle tends to locate away from the polar phase while the polar parts of the molecule (head groups) tend to locate at the polar micelle solvent interface.

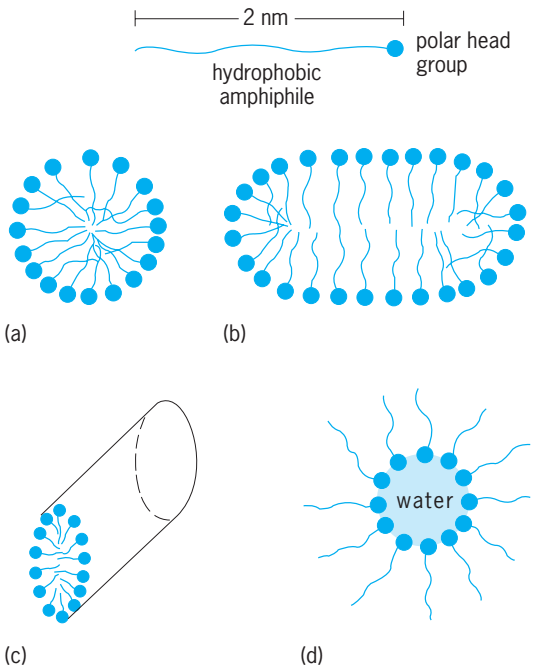

Form of an amphiphile and several forms of micelle: (*a*) spherical, (*b*) disk, (*c*) rod, and (*d*) reversed.

A micelle may take several forms, depending on the conditions and composition of the system, such as distorted spheres, disks, or rods (**illus.** *a*, *b*, and *c*). The dimensions of the particle are derived from those of the amphiphile, for example, a sphere radius of about 2 nanometers or a rod cross-sectional radius of 2 nanometers. Frequently the polar head group is a salt which ionizes in polar media. The micelle may have about 30% of the amphiphiles in the ionized state, giving rise to a highly charged particle surrounded by a cloud of counterions (ions with a charge opposite to that of the micelle ions), and also counterions bound into the micelle surface. Micelles are formed in nonpolar media such as benzene, where the amphiphiles cluster around small water-droplets in the system, forming an assembly known as a reversed micelle (illus. *d*).

When the surfactant contains two long alkyl chains, a vesicle structure rather than a micelle is formed because of geometric restraints. This entity consists of two closed concentric spherical layers of surfactant which enclose an internal volume of water. Such a structure is similar to that of a biological membrane; micelles and vesicles mimic biological systems both in structure and in many kinetic processes. *See* CELL MEMBRANES; MEMBRANE MIMETIC CHEMISTRY.

Micellar systems have the unique property of being able to solubilize both hydrophobic and hydrophilic compounds. They are used extensively in industry for detergency and as solubilizing agents. A strong catalytic action is often associated with these systems and is attributed to the clustering of reactants in the micelle, thereby creating high local

reactant concentration, and also to the strongly charged surface which influences the transition state of a reaction. *See* DETERGENT; SURFACTANT.

The locally high concentration of guest molecules or reactants in a micelle surface leads to enhanced rates of reaction over those observed when the reactants are dispersed in the bulk phase. This catalytic effect can be as large as 1000-fold and can, in some cases, be used selectively. Photolysis of dialkyl ketones in micelles leads to radicals whose local concentrations are high. Only a portion of these radicals can react in micelles, those that have experienced spin relaxation due to their particular isotopic constituency. The unreacted portion with a different isotopic constituency diffuse into the bulk phase and give alternative products. The technique has been used successfully to produce isotopic enrichment. *See* CATALYSIS; FREE RADICAL.

Micelles play an important role in photo-induced reactions, in particular photo-induced electron transfer reactions. Anionic micelles such as sodium dodecyl sulfate strongly promote photoionization of several molecules (for example, phenothiazine, aminopyrene, or tetramethyl benzidine) located at the micelle interface. The ionization threshold is reduced by more than 3.0 eV [>70 kilocalories (300 kilojoules) per mole], an effect not observed in homogeneous solution. The electron of the guest molecule is transferred to the aqueous phase and observed as a hydrated electron, while the cation is stabilized by the micelle. Electron transfer between two guest molecules, giving two radical ions, often occurs most efficiently on micellar systems. The high local concentration and the interface promote electron transfer and lead to efficient reaction. The charged micelle interface stabilizes the ion of opposite charge, while repelling into the bulk water phase the ion of like charge. This latter effect of the micelle surface leads both to efficient charge separation and to long-lived radical ions. The ions can be used to generate useful chemistry via subsequent reaction with other components of the system. In particular, the presence of catalysts such as colloidal platinum can lead to water breakdown, the radical anion donating an electron to the platinum which leads to the formation of hydrogen. These types of micelle-mediated processes are of use in the storage of solar energy. *See* PHOTOCHEMISTRY

. J. K. Thomas

Bibliography. J. H. Fendler, *Membrane Mimetic Chemistry*, 1982; K. Kalyanasundaram, *Photochemistry in Microheterogeneous Systems*, 1987; D. O. Shah (ed.), *Micelles, Microemulsions, and Monolayers*, 1998; J. K. Thomas, *The Chemistry of Excitation at Interfaces*, Amer. Chem. Soc. Monog. Ser. 181, 1984.

Microbial ecology

The study of interrelationships between microorganisms and their living and nonliving environments. Microbial populations are able to tolerate and to grow under varying environmental conditions, including habitats with extreme environmental conditions such as hot springs, salt lakes, and deep-sea thermal vents. Some microorganisms, referred to as extremophiles, grow only under conditions usually considered hostile to life. The archaebacteria (Archaea) have many extremophiles that grow at high temperatures, high salinities, and low pH. Understanding the environmental factors controlling microbial growth and survival offers insight into the distribution of microorganisms in nature, and many studies in microbial ecology are concerned with examining the adaptive features that permit particular microbial species to function in particular habitats.

Within habitats some microorganisms are autochthonous (indigenous), filling the functional niches of the ecosystem, and others are allochthonous (foreign), surviving in the habitat for a period of time but not filling the ecological niches. Because of their diversity and wide distribution, microorganisms are extremely important in ecological processes. The dynamic interactions between microbial populations and their surroundings and the metabolic activities of microorganisms are essential for supporting productivity and maintaining environmental quality of ecosystems. Microorganisms are crucial for the environmental degradation of liquid and solid wastes and various pollutants and for maintaining the ecological balance of ecosystems—essential for preventing environmental problems such as acid mine drainage and eutrophication. *See* ARCHAEA; ECOSYSTEM; EUTROPHICATION.

Population interactions. The various interactions among microbial populations and between microbes, plants, and animals provide stability within the biological community of a given habitat and ensure conservation of the available resources and ecological balance. Interactions between microbial populations can have positive or negative effects, either enhancing the ability of populations to survive or limiting population densities. Sometimes they result in the elimination of a population from a habitat.

Commensal relationships, in which one population benefits and the other is unaffected, occur when one population modifies the habitat to the benefit of a second population. For example, one microbial population may produce a growth factor or may oxidize a substrate and form a metabolic product that a second population can use; one microorganism also may grow on the surface of another organism. Synergism or protocooperation between two populations occurs when both populations benefit from a nonobligatory relationship. Synergism enables microbial populations to reach higher densities in the rhizosphere (soil influenced by plant roots) than in root-free soil, and plants exhibit enhanced growth characteristics as a result of interactions with rhizosphere microbes. Mutualism, an obligatory interrelationship between two populations, allows populations to unite, thereby enabling them to occupy habitats that are unfavorable for the existence of either alone. Mutualistic relationships may lead to the

evolution of new organisms, such as occurs when a fungus unites with either an alga or cyanobacterium to form a lichen. Some animals rely on microorganisms to degrade cellulosic plant residues; ruminants, such as cattle, establish mutualistic relationships with cellulose-degrading microbial populations from which they derived their nutrition. *See* RHIZOSPHERE.

Competition for the same resources results in both populations achieving lower densities than would have occurred in the absence of competition, and prevents populations from occupying the same ecological niche. Amensalism, or antagonism, occurs when one population produces a substance inhibitory to another population, such as the production of an antibiotic by a microbe that inhibits the growth of another microbe. In parasitism, the parasite derives its nutritional requirements from the host, and as a result damages the host. Predation involves the consumption of a prey species by a predatory population for nutrition. Many protozoa nondiscriminantly prey or graze upon bacteria, and protozoa and invertebrates similarly graze on algae. *See* POPULATION ECOLOGY.

Biogeochemical cycling. The transfer of carbon and energy stored in organic compounds between the organisms in the community forms an integrated feeding structure called a food web. Primary producers that form organic matter are at the base of the food web. In food webs based on phytoplankton, algae and cyanobacteria are the primary food sources; in detrital food webs, microbial biomass produced from growth on dead organic matter serves as the primary food source for grazers, the organisms that feed upon primary producers. Grazers are eaten by predators, which in turn may be preyed upon by larger predators. Microbial decomposition of dead plants and animals and partially digested organic matter in the decay portion of a food web is largely responsible for the conversion of organic matter to carbon dioxide. *See* BIOMASS; FOOD WEB.

Biological nitrogen fixation is restricted to specific bacterial and archaebacterial genera. In terrestrial habitats, the microbial fixation of atmospheric nitrogen is carried out by free-living bacteria, such as *Azotobacter*, and by bacteria living in symbiotic association with plants, such as *Rhizobium* or *Bradyrhizobium* living in mutualistic association within nodules on the roots of leguminous plants. In aquatic habitats, cyanobacteria, such as *Anabaena* and *Nostoc*, fix atmospheric nitrogen. The incorporation of the bacterial genes controlling nitrogen fixation into agricultural crops through genetic engineering may help improve yields. *See* NITROGEN FIXATION.

Microorganisms also carry out other processes essential for the biogeochemical cycling of nitrogen. In nitrification, chemolithotrophic microorganisms oxidize ammonium to nitrate. In soil, *Nitrosomonas* oxidizes ammonia to nitrite and *Nitrobacter* oxidizes nitrite to nitrate. The conversion of ammonia to nitrate causes leaching of nitrogen from the soil and results in the loss of soil fertility and nitrate contamination of ground water. Denitrification, the microbial conversion of fixed forms of nitrogen to molecular nitrogen, completes the nitrogen cycle. *See* BIOGEOCHEMISTRY; NITROGEN CYCLE.

Acid mine drainage is a consequence of the metabolism of sulfur- and iron-oxidizing bacteria. When coal mining exposes pyrite ores to atmospheric oxygen, the combination of autoxidation and microbial sulfur and iron oxidation produces large amounts of sulfuric acid. The acid draining from mines kills aquatic life and renders water unsuitable for drinking or for recreation. *See* AUTOXIDATION.

Biodegradation of wastes. The biodegradation (microbial decomposition) of waste is a practical application of microbial metabolism for solving ecological problems. Solid wastes are decomposed by microorganisms in landfills and by composting. In landfills, organic matter is decomposed by anaerobic microorganisms. The products of anaerobic microbial metabolism, including methane and fatty acids, move into the surrounding soil, water, and air, causing the landfill to settle. Eventually, decomposition slows, subsidence ceases, and the land is stabilized. Alternatively, the organic portion of solid waste can be biodegraded aerobically by composting, converting noxious organic waste materials into carbon dioxide, water, and a humuslike product.

Liquid waste (sewage) treatment uses microbes to degrade organic matter, thereby reducing the biochemical oxygen demand (BOD). Otherwise, the BOD would cause oxygen depletion and fish kills when the sewage enters receiving waters. In the trickling filter system, sewage is sprayed over a bed of porous material covered with a slimy film of aerobic bacteria, such as *Zoogloea ramigera*, that degrade the organic matter and thus reduce BOD. Similarly, rotating biological contactors use a film of aerobic bacteria to reduce sewage BOD. In the activated sludge process, which uses a suspension of aerobic bacteria, sewage is introduced into an aeration tank, and sludge from a previous run that contains a massive number of microbes is added as an inoculum. Aerobic microbial metabolism results in the complete degradation of most organic compounds in the sewage. Septic tanks and anaerobic sludge digestors rely on anaerobic microbial metabolism for treating sewage.

Because fecal contamination of potable water from untreated or inadequately treated sewage promotes the rapid dissemination of pathogens, chlorination is used for disinfection following sewage treatment and chloramination is used to disinfect municipal water supplies. The degree of fecal contamination of water is monitored by testing for indicator coliform bacteria. Because the coliform bacterium *Escherichia coli* is present in far greater numbers in human fecal material than are enteropathogens, the coliform test can reliably detect potentially dangerous fecal contamination. *See* ESCHERICHIA; SEWAGE TREATMENT.

Biodegradation of environmental pollutants. Human exploitation of fossil fuels and the production of novel synthetic compounds (xenobiotics) such as plastics and pesticides have introduced many compounds into the environment that are difficult for microorganisms to biodegrade. When petroleum spills occur, for example, microorganisms are faced with the task of degrading tons of hydrocarbons with varying chemical structures. Although it cannot work quickly enough to remove coastal oil spills before beaches are coated and plants and animals are killed, microbial biodegradation will eventually degrade petroleum pollutants, preventing the oceans from being completely covered with oil. Bioremediation, that is, the use of microorganisms to remove pollutants, is used as a biotechnological treatment of oil spills. It also has great potential for the environmental cleanup of other pollutants, including polychlorinated biphenyls (PCBs). *See* BIODEGRADATION.

The ability of a microorganism to degrade an environmental pollutant is highly dependent on the chemical structure of the pollutant. Often a simple change in the substituents of a pesticide may make the difference between biodegradability and recalcitrance (complete resistance to biodegradation). When xenobiotics are recalcitrant and microorganisms fail in their role of "biological incinerators," environmental pollutants accumulate. Synthetic organic compounds can be designed to eliminate obstacles to biodegradation. For example, some alkyl benzyl sulfonates in laundry detergents and plastics have been synthesized so that they are biodegradable. The environmentally safe use of chemical pesticides depends on microbial biodegradation to prevent persistence and biomagnification of these toxic compounds. When microorganisms fail to degrade the pesticide, serious environmental consequences can result far from the site of application, as seen in the case of dichlorodiphenyltrichloroethane (DDT). Biological control, through the regulated establishment of microbial diseases in pest populations, offers a useful alternative to chemical pesticides. *See* DETERGENT; ECOLOGY; ENVIRONMENTAL TOXICOLOGY; BIODEGRADATION; PESTICIDE. Ronald M. Atlas

Bibliography. R. M. Atlas and R. Bartha, *Microbial Ecology: Fundamentals and Applications*, 4th ed., 1997; R. G. Burns and J. H. Slater (eds.), *Experimental Microbial Ecology*, 1982; R. E. Campbell, *Microbial Ecology*, 1983; C. Edwards, *Microbiology of Extreme Environments*, 1990; J. G. Jones (ed.), *Advances in Microbial Ecology*, vol. 13, 1994; J. M. Lynch and J. Hobbie (eds.), *Microorganisms in Action: Concepts and Applications in Microbial Ecology*, 1988.

Microbiology

The multidisciplinary science of microorganisms which began with the study of bacteria. The prefix micro- generally refers to an object sufficiently small that a microscope is required for visualization. In the seventeenth century, Anton van Leeuwenhoek first documented observations of bacteria by using finely ground lenses. Bacteriology, as a precursor science to microbiology, was based on Louis Pasteur's pioneering studies in the nineteenth century, when it was demonstrated that microbes as minute simple living organisms were an integral part of the biosphere involved in fermentation and disease. Microbiology matured into a scientific discipline when students of Pasteur, Robert Koch, and others sustained microbes on various organic substrates and determined that microbes caused chemical changes in the basal nutrients to derive energy for growth. These observations contributed to further understanding of the microbial basis of fermentation. Modern microbiology continued to evolve from bacteriology by encompassing the identification, classification, and study of the structure and function of a wide range of microorganisms. The comprehensive range of organisms is reflected in the major subdivisions of microbiology, which include medical, industrial, agricultural, food, and dairy. *See* BACTERIOLOGY; MICROSCOPE.

Modern microbiology includes protozoa, algae, fungi, viruses, rickettsia, and parasites as well as bacteria. Based on Koch's germ theory of disease, certain microbes are classified as either animal or plant pathogens if, after invasion of the host, a specific disease is manifested. The study of host–parasite relationships gave rise to the concept that the host and parasite coexist in a delicate balance modulated in part by the host's immune system. The latter responds primarily to invasions by agents foreign to the host. Thus, the overlapping disciplines of cell biology, immunology, and modern medicine can trace their origins through microbiology. *See* ALGAE; CELL BIOLOGY; FUNGI; IMMUNOLOGY; MEDICAL BACTERIOLOGY; MEDICAL MYCOLOGY; MEDICAL PARASITOLOGY; PROTOZOA; RICKETTSIOSES; VIRUS.

Microbiology as an interdisciplinary science has directly and indirectly spawned the development of many related fields while retaining a distinct identity. This development can be traced historically from the initial theory of fermentation to microbial metabolism, to enzymology, to biochemistry, to protein chemistry, and finally to molecular biology and biotechnology. Biochemists, geneticists, cytologists, molecular biologists, and immunologists all depend on microbes as basic experimental tools in the study of fundamental processes such as metabolism, photosynthesis, enzymatic catalysis, gene action, population dynamics, and immune responsiveness. *See* BIOCHEMISTRY; BIOTECHNOLOGY; ENZYME; GENETICS; IMMUNOLOGY; METABOLISM; MOLECULAR BIOLOGY; PHOTOSYNTHESIS; POPULATION ECOLOGY.

Microbiology continues as a dynamic science as evidenced by the finding that microbes can no longer be divided into just the two groupings prokaryotes, which have a primitive and dispersed nuclear material, and eukaryotes, which display a distinct nucleus bounded by membrane. A third life form, archaebacteria, is neither prokaryote nor eukaryote in genetic and biochemical properties. Archaebacteria are

usually found in harsh environments such as under extreme temperature, pressure, or salt conditions. Thus, consistent with the dynamic state of microbiology another discipline has been created that will yield important information about the mechanisms whereby cells and their constituents (proteins) function under extreme conditions. *See* ARCHAEA; EUKARYOTAE; PROKARYOTAE.

Microbiological research, both basic and applied, is conducted in the industrial, academic, and government sectors. Industrial microbiological research was stimulated by the discovery of antibiotics of microbial origin. The expertise acquired in microbiology prepared the industrial sector for the advent of biotechnology and recombinant gene cloning, which utilizes bacterial cells for the production of enormous quantities of medically important recombinant proteins. The massive production of recombinant human insulin and bovine growth hormone by genetically engineered bacterial cells provides classical examples. *See* ANTIBIOTIC; GENETIC ENGINEERING; INDUSTRIAL MICROBIOLOGY.

In the academic sector, microbial agents serve as experimental systems to examine the structure and function of various macromolecules such as proteins, polysaccharides, nucleic acids, and lipids. Bacteria and bacterial phage were critical elements leading to the discovery of deoxyribonucleic acid (DNA) as the universal genetic material responsible for the synthesis of proteins. Elucidation of the intricate mechanisms involved in gene expression, control, and regulation are the foundation of modern biotechnology. *See* DEOXYRIBONUCLEIC ACID (DNA).

Government research has focused on the epidemiology, confinement, and eradication of a wide range of endemic diseases of microbial etiology. Federal agencies are primarily responsible for the continuous monitoring of potential epidemics and the eventual production of drugs or vaccines to combat threats posed by such microbial agents as the influenza virus. Federal laboratories also rely on microbiology to better understand food production, geochemistry, and nutrition. International agencies, such as the World Health Organization, are involved in constant efforts to monitor and eradicate microbe-related diseases such as malaria, especially in developing countries. Federal and international agencies must also identify and focus on new diseases which threaten the world's population, as in the case of the viral-induced acquired immune deficiency syndrome (AIDS). *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); EPIDEMIOLOGY; MALARIA.

Many important scientific endeavors are directly traceable to basic microbiological findings and principles. With the discovery of restriction enzymes, which cleave specific sites within chromosomes, specific genes can be identified and isolated, and their structure determined for eventual use in gene cloning experiments. These capabilities led to the founding of the Human Genome Project, which is an attempt to map and determine a base sequence for the DNA in all 23 human chromosomes of a human prototype. Having a complete DNA sequence may allow for the diagnosis and eventual cure of large numbers of gene-linked diseases. *See* CLINICAL MICROBIOLOGY; INDUSTRIAL MICROBIOLOGY; MARINE MICROBIOLOGY; PETROLEUM MICROBIOLOGY; SOIL MICROBIOLOGY; TEXTILE MICROBIOLOGY.

Edward W. Voss, Jr.

Bibliography. T. D. Brock et al., *Biology of Microorganisms*, 9th ed., 1999; P. R. Murray et al., *Medical Microbiology*, 3d ed., 1998.

Microbiota (human)

Microbial flora harbored by normal, healthy persons. The normal fetus is sterile, but during and after birth the infant is exposed to an increasing number of microorganisms. Subsequently, those microorganisms best adapted to survive and colonize particular sites establish themselves and become predominant.

In a healthy human, internal tissues (such as brain, blood, cerebrospinal fluid, and muscle) are normally free of microorganisms. Conversely, surface tissues (such as skin and mucous membranes) are constantly in contact with environmental microorganisms and are readily colonized by certain microbial species. The mixture of microorganisms regularly found at any anatomical site is referred to as the normal microbiota, the indigenous microbial population, the microflora, or the normal flora. For consistency, the term normal microbiota is used here. Because bacteria make up most of the normal microbiota, they are emphasized over the fungi (mainly yeasts) and protozoa. *See* BACTERIA; FUNGI; MICROBIAL ECOLOGY; PROTOZOA; YEAST.

Skin. The adult human is covered with approximately 2 m^2 (21.5 ft^2) of skin. It has been estimated that this surface area supports about 10^{12} bacteria. Commensal microorganisms living on the skin can be either resident (normal) or transient microbiota. Resident organisms normally grow on or in the skin. Their presence becomes fixed in well-defined distribution patterns. Those microorganisms that are temporarily present are transients. Transients usually do not become firmly entrenched and are unable to multiply.

Most skin bacteria are found on the superficial cells, on colonizing dead cells, or are closely associated with oil and sweat glands. Secretions from these glands provide the water, amino acids, urea, electrolytes, and specific fatty acids that serve as nutrients primarily for *Staphylococcus epidermidis* and aerobic corynebacteria. Gram-negative bacteria generally are found in the moister regions. The yeasts *Pityrosporum ovale* and *P. orbiculare* normally occur in the scalp. Some dermatophytic fungi may colonize the skin and produce athlete's foot and ringworm. Other normal microbiota of the skin include coagulase-negative staphylococci, diphtheroids (including *Propionibacterium acnes*), *Staphylococcus aureus*, streptococci (various species), *Bacillus*

spp., *Malassezia furfur, Candida* spp., and *Mycobacterium* spp. (occasionally).

Nose and nasopharynx. The normal microbiota of the nose are found just inside the nostrils. *Staphylococcus aureus* and *S. epidermidis* are the predominant bacteria present and are found in approximately the same numbers as on the skin and face.

The nasopharynx, that part of the pharynx lying above the level of the soft palate, may contain small numbers of potentially pathogenic bacteria such as *Streptococcus pneumoniae, Neisseria meningitidis*, and *Haemophilus influenzae*. Diphtheroids, a large group of nonpathogenic gram-positive bacteria that resemble *Corynebacterium* spp., are commonly found in both the nose and nasopharynx.

Oropharynx. The oropharynx is that division of the pharynx lying between the soft palate and the upper edge of the epiglottis. The most important bacteria found in the oropharynx are the various alpha-hemolytic streptococci (*S. oralis, S. milleri, S. gordoni, S. salivarius*), large numbers of diphtheroids, *Branhamella catarrhalis*, and small gram-negative cocci related to *Neisseria meningitidis*. The palatine and pharyngeal tonsils harbor a similar microbiota, except within the tonsillar crypts, where there is an increase in *Micrococcus* and the anaerobes *Porphyromonas, Prevotella*, and *Fusobacterium*.

Respiratory tract. The upper and lower respiratory tracts (trachea, bronchi, bronchioles, alveoli) do not have a normal microbiota, because microorganisms are removed in at least three ways. First, a continuous stream of mucus is generated by the goblet cells. This mucus entraps microorganisms, and the ciliated epithelial cells continuously move the entrapped microorganisms out of the respiratory tract. Second, alveolar macrophages phagocytize and destroy microorganisms. Finally, a bactericidal effect is exerted by the enzyme lysozyme, which is present in the nasal mucus.

Eye. At birth and throughout life, a small number of bacterial commensals are found on the conjunctiva of the eye. The predominant bacterium is *Staphylococcus epidermidis*, followed by *S. aureus*, *Haemophilus* spp., and *Streptococcus pneumoniae*.

External ear. The normal microbiota of the external ear resemble those of the skin, with coagulase-negative staphylococci and *Corynebacterium* predominating. Mycological studies show the following fungi to be normal microbiota: *Aspergillus, Alternaria, Penicillium, Candidia*, and *Saccharomyces*.

Mouth. The normal microbiota of the mouth or oral cavity contain organisms that resist mechanical removal by adhering to surfaces like the gums and teeth. Those that cannot attach are removed by the mechanical flushing of the oral cavity contents to the stomach, where they are destroyed by hydrochloric acid. The continuous desquamation (shedding) of epithelial cells also removes microorganisms. Those microorganisms able to colonize the mouth find a very confortable environment due to the availability of water and nutrients, the suitability of pH and temperature, and the presence of many other growth factors.

The oral cavity is colonized by microorganisms from the surrounding environment within hours after a human is born. Initially, the microbiota consist mostly of the genera *Streptococcus, Neisseria, Actinomyces, Veillonella*, and *Lactobacillus*. Some yeasts are also present. Most microorganisms that invade the oral cavity initially are aerobes and obligate anaerobes. When the first teeth erupt, the anaerobes (*Porphyromonas, Prevotella*, and *Fusobacterium*) become dominant due to the anaerobic nature of the space between the teeth and gums. As the teeth grow, *Streptococcus parasanguis* and *S. mutans* attach to their enamel surfaces; *S. salivarius* attach to the buccal and gingival epithelial surfaces and colonize the saliva. These streptococci produce a glycocalyx and various other adherence factors that enable them to attach to oral surfaces. The presence of these bacteria contributes to the eventual formation of dental plaque, caries, gingivitis, and peridodontal disease. *See* ANAEROBIC INFECTION.

Stomach. Many microorganisms are washed from the mouth into the stomach. Owing to the very acidic pH values (2 to 3) of the gastric contents, most microorganisms are killed. As a result, the stomach usually contains less than 10 viable bacteria per milliliter of gastric fluid. These are mainly *Streptococcus, Staphylococcus, Lactobacillus, Preptostreptococcus, Helicobacter*, and yeasts such as *Candida* spp. Microorganisms may survive if they pass rapidly through the stomach or if the organisms ingested with food are particularly resistant to gastric pH (mycobacteria). *See* GASTROINTESTINAL TRACT DISORDERS.

Small intestine. The small intestine is divided into three anatomical sections: the duodenum, jejunum, and ileum. The duodenum (the first 25 cm or 10 in. of the small intestine) contains few microorganisms because of the combined influence of the stomach's acidic juices and the inhibitory action of bile and pancreatic secretions. Of the bacteria present, gram-positive cocci and rods make up most of the microbiota. *Enterococcus faecalis*, lactobacilli, diphtheroids, and the yeast *Candidia albicans* are occasionally found in the jejunum. In the distal portion of the small intestine (ileum), the microbiota begin to take on the characteristics of the colon microbiota. It is within the ileum that the pH becomes more alkaline. As a result, anaerobic gram-negative bacteria and members of the family Enterobacteriaceae become established.

Large intestine (colon). The large intestine, or colon, has the largest microbial community in the body. Microscopic counts of feces approach 10^{12} organisms per gram of wet weight. Over 400 different species have been isolated from human feces. The colon can be viewed as a large fermentation vessel, and the microbiota consist primarily of anaerobic, gram-negative, nonsporing bacteria and gram-positive, spore-forming or nonsporing rods. Not only are the vast majority of microorganisms anaerobic,

but many different species are present in large numbers. Several studies have shown the ratio of anaerobic to facultative anaerobic bacteria is approximately 300 to 1. Even the most abundant of the latter, *Escherichia coli*, is only about 0.1% of the total population.

Besides the many bacteria in the colon, the yeast *Candidia albicans* and certain protozoa may occur as harmless commensals. *Trichomonas hominis, Entamoeba hartmanni, Endolimax nana*, and *Iodamoeba butschlii* are common inhabitants.

Genitourinary tract. The upper genitourinary tract (kidneys, ureters, and urinary bladder) is usually free of microorganisms. In both the male and female, a few bacteria (*Staphylococcus epidermidis, Enterococcus faecalis*, and *Corynebacterium* spp.) usually are present in the distal portion of the urethra.

In contrast, the adult female genital tract, because of its large surface area and mucous secretions, has a complex microbiota that constantly changes with the female's menstrual cycle. The major microorganisms are the acid-tolerant lactobacilli, primarily *Lactobacillus acidophilus*, often called Doderlein's bacillus. They ferment glycogen produced by the vaginal epithelium, forming lactic acid. As a result, the pH of the vagina and cervix is maintained between 4.4 and 4.6. John P. Harley

Bibliography. F. Backhed et al., Host-bacterial mutualism in the human intestine, *Science*, 307:1915–1920, 2005; B. S. Drasar et al., *Intestinal Microbiology*, American Society for Microbiology, Washington, DC, 1985; P. B. Ekburg et al., Diversity of the human intestinal microbial flora, *Science*, 308:1635–1638, 2005; P. A. Mackowiak, The normal microbial flora, *N. Engl. J. Med.*, 307:83, 1982; J. Travis, Gut check, *Sci. News*, 163:344–345, 2003.

Microcline

Triclinic potassium feldspar, $KAlSi_3O_8$, that usually contains a few percent sodium feldspar (Ab = $NaAlSi_3O_8$) in solid solution. Its hardness is 6; specific gravity, 2.56; mean refractive index, 1.52; color, white (green varieties are called amazon stone or amazonite; see **illus.**). Good (001) and (010) planar cleavages are inclined at 90.4°. Microcline is found in some relatively high-grade regional metamorphic rocks, but is much more common in pegmatities, granites, and related plutonic igneous rocks. In the last, it often occurs as a microcline perthite, containing exsolved low albite intergrowths. *See* PERTHITE.

Microcline is usually twinned on both the Albite and Pericline laws. This fine-scale, cross-hatched (or tweed) appearance is observed only in thin rock sections examined on a polarizing light microscope. It is the result of ordering of Al and Si from a higher-temperature, monoclinic, sanidinelike arrangement, which may include orthoclase as an intermediate state. Microcline that crystallizes authigenically at relatively low temperatures (<570°F or 300°C), and thus in its own stability field, will lack this twinning entirely. Low microcline is completely ordered,

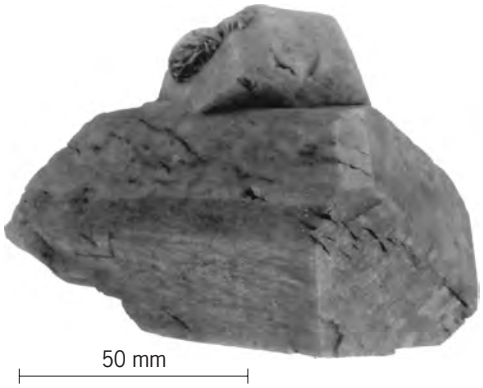

Microcline, variety amazon stone, Pikes Peak, Colorado. (*American Museum of Natural History specimen*)

with Al in one of the four nonequivalent tetrahedral sites and Si in the other three. Intermediate microclines are relatively rare in nature and have complex crystal structures with somewhat disordered Al, Si arrangements. *See* AUTHIGENIC MINERALS; CRYSTAL STRUCTURE; FELDSPAR; IGNEOUS ROCKS; ORTHOCLASE; SOLID SOLUTION. Paul H. Ribbe

Microcomputer

A digital computer whose central processing unit consists of a microprocessor, a single semiconductor integrated circuit chip. Once less powerful than larger computers, microcomputers are now as powerful as the mini- and super-minicomputers of just several years ago. This is due in part to the growing processing power of each successive generation of microprocessor, plus the addition of mainframe computer features to the chip, such as floating-point mathematics, computation hardware, memory management, and multiprocessing support. *See* INTEGRATED CIRCUITS; MICROPROCESSOR; MULTIPROCESSING.

Microcomputers are the driving technology behind the growth of personal computers and workstations. While the hardware distinctions between personal computers and workstations have blurred, the applications they use and the operating system environments are still distinct. Traditionally, personal computers offer lower performance at a lower cost, with applications focusing on business, games, and education. In contrast, workstations use more specialized components that deliver high performance on numerically intensive scientific and engineering computations, such as simulations and three-dimensional (3-D) design.

Early microcomputers relied on ASCII text (characters) for input via a keyboard or paper tape as well as output on a terminal or display. They used floppy disks for storage and had limited computational and memory capacity. Today's microcomputers employ bit-mapped graphical user interfaces on color displays and handle a wide variety of input and output devices such as mice, digitizing tablets,

digital color video cameras, and compact-disk-quality stereo sound. The memory capacities of microcomputers range from tens to hundreds of megabytes, and certain specialized systems may have a gigabyte of memory. Mass storage is handled by hard disk drives with capacities of several gigabytes or more, augmented by optical storage devices such as CD-ROM (compact-disk, read-only memory), magneto-optical storage, and DVD (digital video disk or digital versatile disk). The computational and memory capacity of microcomputers increased from 64 kilobytes in 1974 to over 4 gigabytes in 1998, an increase of four orders of magnitude. Within just several years, microprocessor clock frequencies climbed from 133 MHz to 1 GHz in 2000. The combined memory capacity and the processing speed of today's microcomputers make new applications possible. Such applications include streaming digital video for videoconferencing and entertainment, and the real-time display of three-dimensional graphics for data analysis, games, and virtual reality.

Microcomputers are employed in virtually every sector of society, including commerce, manufacturing, education, government, travel, and the home. General-purpose microcomputers, in the form of personal and professional computers, support a variety of applications, from video games to business reports to laboratory analysis. The microcomputer's network capabilities and storage capacity have fueled the growth of the Internet, a worldwide network of interconnected computers once used only by research laboratories and universities. This network formerly consisted of mainframes and supercomputers, but now microcomputers serve both as the Internet's backbone for storing and transporting information on demand and as the terminals by which information is accessed. The Internet's ease of use and ubiquity have created new ways to gather information and conduct commerce that are still evolving. *See* INTERNET.

The capabilities of today's microprocessors in combination with reduced power consumption have created a new category of microcomputers: handheld devices. Some of these devices are actually general-purpose microcomputers: They have a miniature keyboard and liquid-crystal-display (LCD) screen and use an operating system that runs several general-purpose applications. Many others, however, serve a fixed purpose. Examples of the latter are cellular telephones that provide a display for receiving text-based pager messages and automobile navigation systems that use satellite-positioning signals to plot the vehicle's position. *See* LIQUID CRYSTALS; MOBILE COMMUNICATIONS; RADIO PAGING SYSTEMS; SATELLITE NAVIGATION SYSTEMS.

Hardware. The microcomputer is an electronic device. The microprocessor acts as the microcomputer's central processing unit (CPU), performing all the operations necessary to execute a program. In early microcomputer designs, the microprocessor also managed simple input/output (I/O) devices such as the graphics display and keyboard. Today's microcomputers are complex machines having several distinct subsystems (see **illus.**). The microprocessor still makes up the system's CPU.

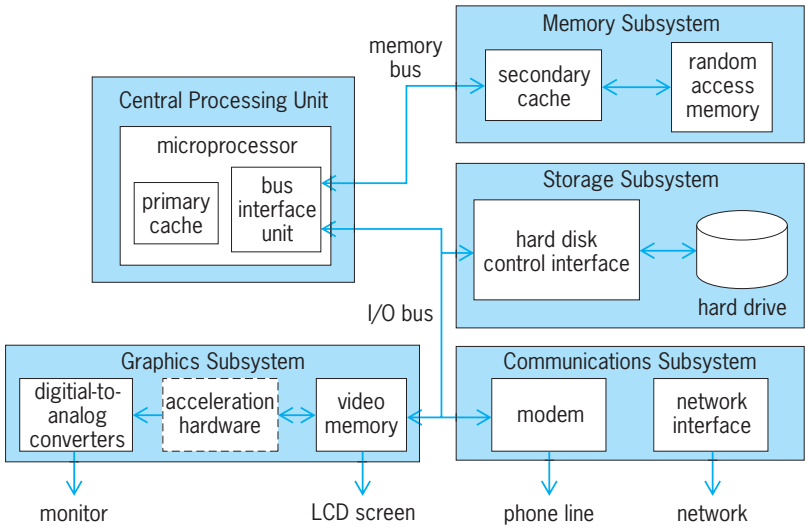

Elements of a microcomputer. The various subsystems are controlled by the central processing unit. Some designs combine the memory bus and bus input/output into a single system bus. The graphics subsystem may contain optional graphics acceleration hardware.

A memory subsystem uses semiconductor random-access memory (RAM) for the temporary storage of data or programs. The memory subsystem may also have a small secondary memory cache that improves the system's performance by storing frequently used data objects or sections of program code in special high-speed RAM.

The graphics subsystem consists of hardware that displays information on a color monitor or LCD screen: a graphics memory buffer stores the images shown on the screen, digital-to-analog convertors (DACs) generate the signals to create an image on an analog monitor, and possibly special hardware accelerates the drawing of two- or three-dimensional graphics. (Since LCD screens are digital devices, the graphics subsystem sends data to the screen directly rather than through the DACs.) *See* DIGITAL-TO-ANALOG CONVERTER.

The storage subsystem uses an internal hard drive or removable media such as floppies, Zip™ or Jazz™ cartridges, or magneto-optical platters for the persistent storage of data.

The communications subsystem consists of a high-speed modem or the electronics necessary to connect the computer to a network. This last subsystem was considered optional in early microcomputer designs; however, the ability to share data among coworkers and to connect to the Internet has become so important that it has become a standard part of microcomputer design.

Architecture. The microprocessor is built of the same components as any digital computer. It has a control unit, an arithmetic and logic unit (ALU), special temporary storage units called registers, and a bus interface unit (BIU). The control unit orchestrates the operation of all the other microprocessor units, having them perform their dedicated tasks simultaneously and at very high speed. The control

unit works by ordering the BIU to fetch instructions from semiconductor memory. It decodes them, typically several at a time, and has the appropriate units perform a sequence of operations that completes the task as directed by the fetched instructions. For example, a set of instructions may have the BIU fetch more data from memory into the microprocessor's registers. Next, the control unit commands the ALU to perform a mathematics calculation (perhaps an add operation) using the values found in two of the registers. Finally, the BIU writes the result of this action back into memory, where it can be used by subsequent instructions. In many microprocessors, a small amount of on-chip memory called the primary cache stores the most recently used instructions or data. If these items are still resident in this cache when the control unit requires them, the BIU can draw the information from the high-speed cache, rather than from the slower off-chip RAM. The performance of microcomputers is often measured in millions of instructions per second (MIPS). Many personal computers operate in the range of 200–400 MIPS.

The BIU thus handles all of the signals necessary for the microprocessor to transfer data between its internal registers and semiconductor memory in the memory subsystem. This electronic interface is known as the memory bus. The BIU also implements the data routes that connect the microprocessor to the other input/output subsystems. Data sent or received from any input/output device in the system typically traverse the memory bus. However, certain special-purpose processor designs use separate buses for memory and input/output, for the sake of design simplicity (to achieve a low cost in a consumer product) or performance (the microprocessor is free to simultaneously decode a compressed video image stored in RAM, while it reads more images from a DVD drive). *See* COMPUTER ARCHITECTURE; DIGITAL COMPUTER.

Word size. An important factor contributing to a microcomputer's processing power is the width of its data paths. These paths usually determine the size of the data elements with which the microcomputer works. The width also determines how fast data can move through the system. The data path width is commonly called the machine's word size and is measured in bits, where each bit represents a single data path. Early microcomputers had 8- or 16-bit data paths, but microcomputers today commonly use 32-bit data paths. Advanced microprocessors using 64-bit data paths are commonly found in workstations or supercomputers (where they make up the elements of a multiprocessor array). Eight- and sixteen-bit processors are still in widespread use as controllers in consumer products such as household appliances. However, 32-bit embedded processors are appearing that meet the processing needs of sophisticated products such as set boxes and smart office copiers. *See* BIT; EMBEDDED SYSTEMS; SUPERCOMPUTER.

Clocks. Another important measure of a microcomputer's computing power is its clock rate. The frequency of the clock signal determines how fast the computer can process information. Commercial microprocessors operate in the range from 25 MHz to 1 GHz and more. Depending upon its architecture, a microprocessor can execute one or more instructions per clock cycle. Consequently, a computation rate of hundreds of millions of instructions per second can be achieved.

Data bus. Previously, communications between the microprocessor and other parts of the microcomputer system occurred over the memory bus. This is because the input/output operations also made use of the memory bus's address and data lines. In this arrangement, the memory bus is commonly called the system bus. The system bus thus provides the routes that collect data from input devices and transfer the data to the semiconductor memory. The CPU then operates on the data, and the system bus transfers the results to an output device. However, in many high-performance microcomputer designs, such as servers, the architecture separates the two functions—accessing memory and device input/output—on two separate buses. This arrangement effectively doubles the data paths through which information can flow through the system, boosting throughput. It also allows the memory subsystem to operate at a higher clock speed than the input/output bus, which further improves performance. The input/output bus is typically based on the peripheral component interconnect (PCI) standard.

In some high-end designs, certain microcomputer subsystems contain their own "local" memory and microprocessors. These intelligent subsystems can handle certain tasks independently of the CPU, using programs stored in the local memory. They use the system bus only when transferring data between the device's local memory and the memory subsystem.

Data storage. Data are stored on a microcomputer in various ways: RAM chips, ferromagnetic hard disk drives, and read-only memory (ROM) such as CD-ROMs. Data bits are stored as electronic "on" or "off" states in the array of transistors that make up the semiconductor chip. RAM is used for the temporary storage of program code and volatile data because of its high access speed (typically in nanoseconds). Programs and data are saved on persistent storage media such as a hard drive. The bits that constitute a program's instructions are encoded as magnetic field changes on the ferromagnetic material. These devices have slower access speeds (measured in milliseconds) but have the advantage of preserving the data when the microcomputer is turned off. *See* COMPACT DISK; COMPUTER STORAGE TECHNOLOGY; MAGNETIC RECORDING; OPTICAL RECORDING; SEMICONDUCTOR MEMORIES.

Input and output. Microcomputers interact with users in a variety of ways. They accept data from input devices such as a keyboard, a mouse, touch screen, microphone, scanner, or video camera. In each case, the data collected by these devices are translated via hardware and software into a form usable by the microcomputer. For example, when

keys are pressed on a keyboard, a controller determines which keys were pressed and converts these actions into data bytes (a byte being a group of bits, usually eight, that encode a character). The microcomputer's operating system converts the bytes into characters, and passes them to a program. The program, perhaps a word processor, then displays the characters on the screen. Talking into a microphone triggers speech recognition software, which parses the tones into specific commands. The software then relays these commands to the operating system, which executes the appropriate actions. A scanner can convert an image into an array of bytes that a drawing program can manipulate. Or an optical character recognition (OCR) program might interpret the patterns of bytes, converting them into text. An example of an advanced input device is a data glove, which senses the position of the fingers on a user's hand. Microcomputer software uses this information to allow a user to control the orientation, position, and rate of objects displayed in a simulation program. When an input device such as the data glove, combined with an output device such as 3-D goggles, enables users to immerse themselves in a simulation program, this type of interaction is called virtual reality. *See* CHARACTER RECOGNITION; COMPUTER PERIPHERAL DEVICES; SPEECH RECOGNITION; VIRTUAL REALITY.

The graphic output produced by the microcomputer is limited only by the resolution of the bit-mapped graphics display hardware. Early display resolutions were either 320 by 240 pixels, or 640 by 480 pixels (where a pixel represents one element on the display). Today, standard screen resolutions range from 800 by 600 pixels to 1024 by 768 pixels. Specialized graphics applications can require displays of even higher resolutions. Most graphics subsystems use an 8- or 24-bit value for each pixel that represents its color. (An 8-bit display has $2^8 = 256$ colors or shades of gray; a 24-bit display has $2^{24} = 16{,}777{,}216$ colors.) Operating systems use these bit-mapped capabilities to present text in any typeface, color, and size. The graphics capability can be used in simulations, in what-you-see-is-what-you-get (WYSIWYG) editors, and in displaying pictures. By presenting a stream of images on a screen rapidly, the graphics subsystem can even present full-motion video. Such multimedia technology allows the microcomputer to combine sound, graphics, and text to enhance the effectiveness of communications. Multimedia-capable microcomputers are part of the reason that users from all walks of life can easily work with information on the Internet. *See* COMPUTER GRAPHICS; ELECTRONIC DISPLAY; HUMAN-COMPUTER INTERACTION; MULTIMEDIA TECHNOLOGY; VOICE RESPONSE.

Networks and electronic services. Attaching computers to each other via computer networks not only multiplies the information available to a microcomputer but also adds the dimension of communications to its capabilities. The incorporation of the network subsystem as standard—rather than optional—hardware has made this task even easier. Network services, such as electronic mail, and dial-up services, such as online systems or Internet service providers, enable users to work remotely from the office and correspond with coworkers or relatives around the globe. Such services have played a major role in the increasing utility and popularity of microcomputers. *See* ELECTRONIC MAIL; LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

Portable computers. The rapid advance of electronic technology has made it possible to build extremely small, light, yet powerful microcomputers. Low-power microprocessors, flat-panel displays, and miniature hard drives have played a critical role. Many laptops weigh from just less than 2 lb (1 kg) up to 7 lb (3 kg). Yet they have RAM, hard-disk capacity, and a display resolution equal to that of stationary desktop computers.

In addition, some portable computers serve as hand-held, fixed-purpose devices. An example is a digital camera that runs an operating system, captures image data, and displays it on a viewfinder's LCD screen in real time. When a picture is taken, the microcomputer applies color corrections to the image, then compresses and stores it on memory cards or a miniature hard drive. All of this is accomplished by an integral microcomputer that is inexpensive and operates on battery power. *See* CAMERA.

Software. Microcomputer software is the logic that makes microcomputers useful. Software consists of programs, which are sets of instructions that direct the microcomputer through a sequence of tasks. A startup program in the microcomputer's ROM initializes all of the devices, loads the operating system software, and starts it. All microcomputers use an operating system that provides basic services such as keyboard input, simple file operations, and the starting or termination of programs. While the operating system used to be one of the major distinctions between personal computers and workstations, today's personal computer operating systems also offer advanced services such as multitasking, networking, and virtual memory. All microcomputers exploit the use of bit-mapped graphics displays to support windowing operating systems. This capability enables the display to act as a series of overlapping windows, each showing the content of a single application program. Windowing operating systems allow an application to be written independently of the actual display size. This lets users allocate screen space as they see fit for multiple applications running simultaneously. *See* OPERATING SYSTEM; SOFTWARE.

Applications. The widespread availability of microcomputers has led to a proliferation of uses. Businesses use microcomputers for payroll, personnel, and other types of management information systems as well as for word processing, electronic mail, presentation formatting, electronic publishing, spreadsheets, project planning, financial projections, scheduling, personal information management, and many other areas. *See* WORD PROCESSING.

The low cost of microcomputers allows them to serve as controllers for scientific instruments. They can collect and process data from the instrument as well as store, analyze, and display the captured

data. Microcomputers are powerful enough to do sophisticated analysis and even maintain databases of previous measurements. The higher-performance microcomputers or workstations are widely used in science and engineering. In science, they can simulate prototype experimental procedures, and predict results that can be compared with the results of actual experiments. Microcomputers have had such a dramatic impact in this area that computation methods based on microcomputers are now considered a "third leg" of science, complementing both theory and experiment. In engineering, microcomputers support design and evaluation of products through computer-aided design (synthesis), evaluation (simulation), and databases (reuse of designs, and the collection of problem reports). In fact, many cars and airplanes are designed and thoroughly tested in simulations before a working prototype is ever built, at tremendous savings in cost and improved safety. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING; DATABASE MANAGEMENT SYSTEM; SIMULATION.

Microcomputers are natural tools for both education and recreation. They are used in schools as a complement to classroom instruction. Software programs can provide specific details in depth, then bring that detail to life interactively. Furthermore, microcomputers can provide personalized responses, customizing lessons by allowing the student to explore where there is a greater interest, and to progress at an individualized pace. The Internet allows students to broaden their educational horizons by providing access to sources of information not available locally, and to converse with other students across the country or around the globe.

Finally, a wide variety of computer games, ranging from flight simulators to war games, fantasy, and problem solving, are available. As the computational capabilities of microcomputers have increased, these games have increased in sophistication and realism. *See* AIRCRAFT TESTING; DIGITAL COMPUTER; VIDEO GAMES. Tom Thompson

Bibliography. P. Freiberger and M. Swaine, *Fire in the Valley: The Making of the Personal Computer*, McGraw-Hill, 1999; H.-P. Messmer, *The Indispensable PC Hardware Book*, 3d ed., Addison-Wesley, 2000; M. Predko, *PC Pb.D.: Inside PC Interfacing*, McGraw-Hill, 1999; M. Rosenthal, *Build Your Own PC*, 2d ed., McGraw-Hill, 2000; S. Veit, *Stan Veit's History of the Personal Computer*, Worldcomm Press, 1993; R. White, *How Computers Work*, 5th ed., MacMillan Computer, 1999.

Microdialysis sampling

An approach for sampling the extracellular space of essentially any tissue or fluid compartment in the body. Continuous sampling can be performed for long periods with minimal perturbation to the experimental animal. Microdialysis provides a route for sampling the extracellular fluid without removing fluid, and administering compounds without adding fluid. The resulting sample is clean and amenable to direct analysis. Microdialysis was initially developed to study neurochemical processes in the brain. The success of this technique in the study of neurotransmitter release has led to the development of microdialysis techniques for general pharmacokinetic and drug distribution studies.

Fundamentals. Microdialysis sampling is performed by implanting a short length of hollow-fiber dialysis membrane at the site of interest. The fiber is slowly perfused with a sampling solution (the perfusate) having an ionic composition and pH that closely matches the extracellular fluid of the tissue being sampled. Low-molecular-weight compounds in the extracellular fluid diffuse into the fiber and are swept to a collection vial for subsequent analysis. The system is analogous to an artificial blood vessel that can deliver compounds and remove the resulting metabolites. A microdialysis system for awake animals consists of a precision microinfusion pump connected to the microdialysis probe through a low-volume liquid swivel. The outlet of the microdialysis probe is then brought back through a second channel of the liquid swivel to a fraction collector. The experimental animal is housed in a containment system so that the tubing does not become entangled as the animal moves. An alternative to the use of a liquid swivel is to place the animal containment system on a turntable that automatically rotates to keep the animal stationary with respect to the microdialysis equipment (**Fig. 1**). *See* DIALYSIS; MEMBRANE SEPARATIONS.

Process. Microdialysis is a diffusion-controlled process. The perfusion rate through the probe is generally in the range of 0.5 to 5.0 ml/min. At this flow rate, there is no net flow of liquid across the dialysis membrane. The driving force for mass transport is the concentration gradient between the extracellular fluid and the fluid in the probe. If the concentration of a compound is higher in the extracellular fluid, some fraction will diffuse into the probe; this is termed a recovery experiment (**Fig. 2***a*). Conversely, if the concentration of a compound is higher in the probe, some fraction will diffuse out of it; this is termed a delivery experiment (Fig. 2*b*). *See* DIFFUSION; TRANSPORT PROCESSES.

At typical perfusion rates, equilibrium is not established across the microdialysis membrane. The concentration of analyte (compound sampled) collected in the dialysate (the perfusate after it has gone through the microdialysis probe) is some fraction of the actual concentration in the extracellular fluid. The relationship between the fraction recovered and the actual concentration is termed the extraction efficiency and is defined as the ratio of the concentration of analyte in the dialysate to the true concentration of analyte in the tissue of interest. The extraction efficiency depends on characteristics of the probe, the sample matrix, and the analyte. Probe characteristics that affect recovery include the type of membrane, the length of membrane, the geometry of the probe, and the perfusion flow rate. The extraction efficiency is influenced by matrix characteristics, such as temperature, tortuosity, and effective

volume. The size, hydrophobicity, charge, and diffusion coefficient of the analyte affect recovery. In addition, metabolic and active transport processes affect extraction efficiency. Usually, the extraction efficiency of a compound is independent of its concentration. Under normal microdialysis sampling conditions, these parameters remain constant; and although equilibrium is not established, a steady state is rapidly achieved.

Probe design. The major design consideration for a microdialysis probe is to provide a short length of hollow-fiber dialysis membrane, through which a solution can be pumped, in a geometry that can be implanted into an animal with minimal tissue damage. Microdialysis membranes are made of cellulose, cellulose acetate, polyacrylonitrile (PAN), and polycarbonate ether. These hydrophilic membranes provide little difference in chemical selectivity. Membranes with nominal molecular weight cutoffs between 5000 and 75,000 daltons (unified atomic mass units, u) are available, although membranes in the range of 10,000 to 20,000 Da are most commonly used. The dialysis efficiency decreases dramatically as the molecular weight of a compound increases. Compounds with molecular weights much above 1000 Da typically have very low efficiencies and are not usually appropriate for microdialysis sampling. A variety of probe designs have been described that provide advantages for use in specific tissues. These designs can be divided into two classes: cannula-style (small-tube) probes with parallel inlet and outlet flow; and linear-style probes in which the inlet, membrane, and outlet are in series (**Fig. 3**).

The most common design is the concentric cannula type that consists of an inner and outer length of stainless steel tubing (Fig. 3*a*). The inner cannula extends beyond the outer cannula and is covered with the dialysis membrane. This design is most appropriate for stereotaxic implantation in the brain. Its rigid nature provides good mechanical stability and allows for precise placement in the brain. The probe can be cemented to the skull to prohibit movement of the probe when the animal moves. Alternatively, a guide cannula can be implanted and glued in place. A concentric cannula microdialysis probe can then be implanted through the guide cannula. This provides the ability to remove the probe and precisely reinsert it or use another probe later.

Microdialysis probes constructed from rigid materials are not appropriate for implantation in peripheral tissue. Unlike probes implanted in the brain, where the skull provides an excellent site to secure the probe and protect it from movement, peripherally implanted probes must be capable of moving as the animal moves without being damaged or causing damage. A modification of the concentric cannula design in which side-by-side pieces of fused-silica are used has been found to be useful for intravenous implantation. In this design, one piece of fused silica extends beyond the other and is covered with the dialysis membrane (Fig. 3*b*). This flexible design is still needlelike and thus can be inserted into blood vessels. However, the fused silica bends as the animal moves so that neither the dialysis membrane nor the blood vessel is punctured.

The most useful probe design for implantation in peripheral tissue is of linear geometry. Several variations of this design are in use, but the general concept is that the hollow-fiber dialysis membrane is connected to small-bore tubing on both ends

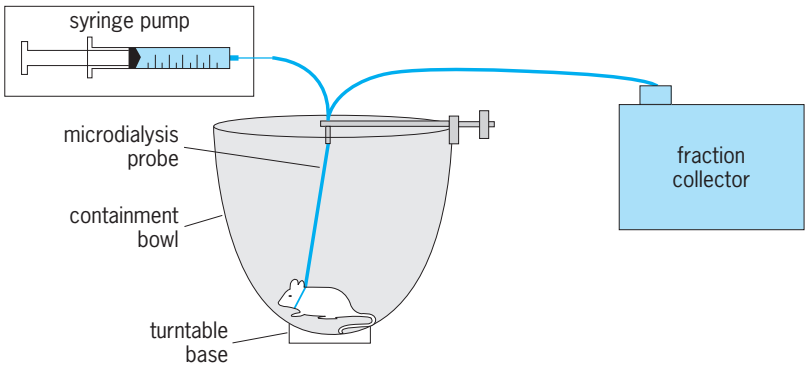

Fig. 1. Microdialysis sampling system for awake animals.

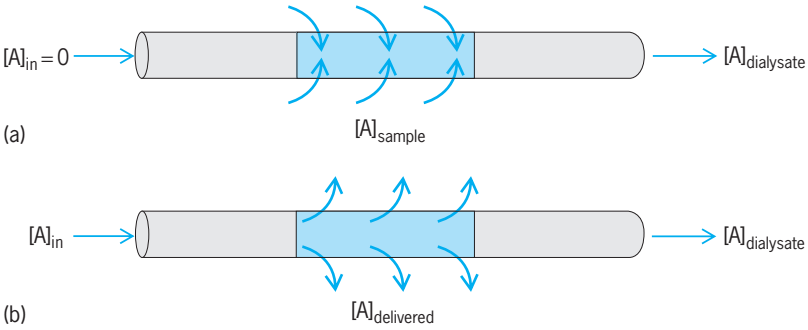

Fig. 2. Mass-transport-driven microdialysis process. (*a*) Recovery. (*b*) Delivery.

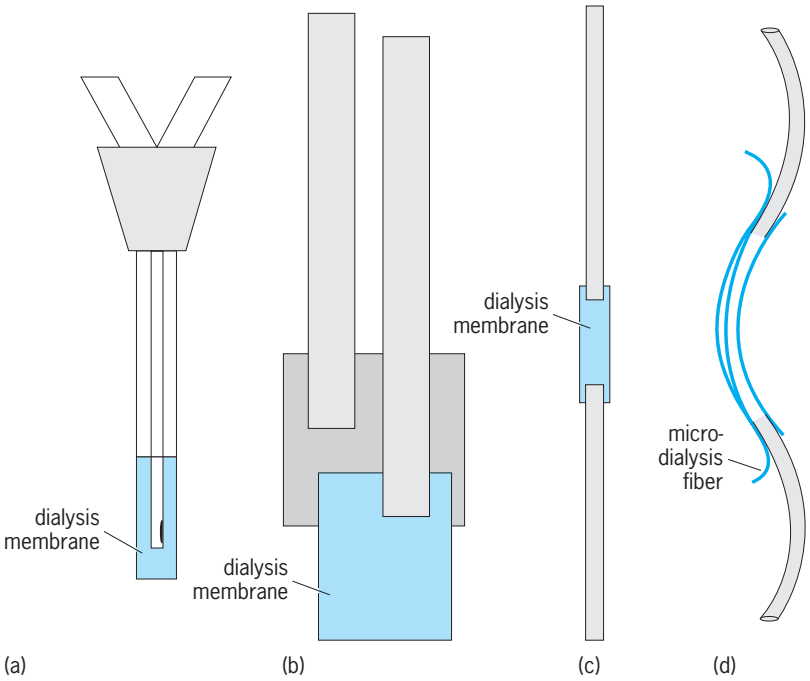

Fig. 3. Microdialysis probes. (*a*) Rigid concentric cannula probe. (*b*) Flexible fused-silica probe. (*c*) Linear probe. (*d*) Flow-through or shunt probe.

(Fig. 3*c*). One end is used as the inlet and the other as the outlet. A variety of tubing types can be used for the inlet and outlet, including Teflon, polyethyletherketone (PEEK), polyethylene (PE), and fused silica. Implantation is accomplished simply by pulling the probe through the tissue like a thread. Often a longer piece of the dialysis fiber than desired for sampling is used because the fiber is more flexible than the fused silica. In this case, most of the fiber is coated with an impermeable silicone resin, leaving a small uncoated window for sampling.

A fourth type of microdialysis probe is termed the flow-through or shunt design. The probe is constructed by inserting the microdialysis fiber inside a length of polyethylene tubing (Fig. 3*d*). The microdialysis fiber is perfused with the sampling solution while the fluid sample flows through the polyethylene tubing. This design is useful for sampling flowing fluids from sites that are too small for implantation of a flexible microdialysis probe. The polyethylene tubing acts as a shunt to bring the biological fluid past the microdialysis membrane. This probe design has been most successfully used to sample bile, but application to blood sampling from arteries or smaller veins may also be possible.

Applications. To date, the greatest use of microdialysis sampling has been in the neurosciences. Microdialysis probes can be implanted in specific brain regions of conscious animals in order to correlate neurochemical activity with behavior. Most studies have focused on determining dopamine or the other monoamine neurotransmitters. Microdialysis sampling has also proven to be a powerful technique for studying excitatory amino acids, such as glutamate, aspartate, and GABA (γ-aminobutyric acid) in the brain. Recently, the use of microdialysis to sample neuropeptides has been explored. An exciting application of microdialysis sampling has been in studying abnormal brain function in humans. Data on the neurochemical processes occurring prior to, during, and after an epileptic seizure have been obtained using microdialysis in humans. *See* AMINO ACIDS; NEUROBIOLOGY.

The use of microdialysis in the pharmaceutical sciences is growing rapidly. Because microdialysis provides continuous sampling without disruption of biological barriers, the technique is particularly well suited for studying the bioavailability of pharmaceutical compounds. Microdialysis probes have been implanted in the skin of experimental animals and humans to determine the transdermal delivery of drugs from ointments. Delivery of anticancer drugs to tumors has been studied using microdialysis. Alternatively, microdialysis may prove a useful technique for delivering toxic drugs, such as anticancer agents, to specific sites without systemic involvement. While continuous sampling from tissues is possible only by microdialysis, sampling from the blood has been the most common approach taken to pharmacokinetic studies using microdialysis. *See* DRUG DELIVERY SYSTEMS; PHARMACOLOGY.

Microdialysis sampling has also been used to study the metabolism of compounds in vivo. Metabolic organs such as the liver and kidneys have been studied by microdialysis sampling. By also sampling the bile by microdialysis, complete metabolic profiles can be obtained from a single experimental animal. This approach dramatically decreases the number of experimental animals needed to assess the metabolism of a new drug.

Craig E. Lunte

Bibliography. W. F. Elmquist and R. J. Sawchuk, Application of microdialysis in pharmacokinetic studies, *Pharm. Res.*, 14:267–288, 1997; D. K. Hansen et al., Pharmacokinetics and metabolism studies using microdialysis sampling, *J. Pharm. Sci.*, 2000; C. E. Lunte, Microdialysis and target organ exposure, in D. E. Johnson (ed.), *Drug Toxicodynamics*, Marcel Dekker, 1999; T. E. Robinson and J. B. Justice, Jr. (eds.), *Microdialysis in the Neurosciences*, Elsevier, 1991.

Micro-electro-mechanical systems (MEMS)

Systems that couple micromechanisms with microelectronics. Such systems are also referred to as microsystems, and the coupling of micromechanisms with microelectronics is also termed micromechatronics. Micromechanics refers to the design and fabrication of micromechanisms that predominantly involve mechanical components with submillimeter dimensions and corresponding tolerances of the order of 1 micrometer or less. The types of systems encompassed by MEMS represent the need for transducers that act between signal and information processing functions, on the one hand, and the mechanical world, on the other. This coupling of a number of engineering areas leads to a highly interdisciplinary field that is commensurately impacting nearly all branches of science and technology in fields such as biology and medicine, telecommunications, automotive engineering, and defense. Ultimately, realization of a "smart" MEMS may be desired for certain applications whereby information processing tasks are integrated with transduction tasks, yielding a device that can autonomously sense and accordingly react to the environment. *See* TRANSDUCER.

Motivating factors behind MEMS include greater independence from packaging shape constraints due to decreased device size, with typical complete packaged devices occupying volumes of less than 0.05 cm^3 (0.003 $in.^3$). In addition, the advantages of repeatable manufacturing processes as well as economic advantages can follow from batch fabrication schemes such as those used in integrated circuit processing, which has formed the basis for MEMS fabrication. Many technical and manufacturing trade-offs, however, come into play in deciding whether an integrated approach is beneficial. In some cases, the device design with the greatest utility is based on a hybrid approach, where mechanical processing and electronic processing are separated until a final packaging step. *See* INTEGRATED CIRCUITS.

Two broad categories of devices follow from the transduction need addressed by MEMS: the input transducer or microsensor, and the output

transducer or microactuator. The specifications for these two categories of devices entail unique fabrication methods that have evolved since about 1970 and have resulted in correspondingly unique process technology and manufacturing methods.

Attributes. An ideal sensor extracts information from the environment without perturbing the environment. Thus, minimal energy exchange is desired. Fast response and high sensitivity with minimal power requirements are also important sensor attributes. All of these qualities are fundamental to microsensors, and the degree to which they may be exploited is a direct consequence of mechanical scaling behavior. The progress in integrated pressure sensors provides an example of how it is possible to capitalize on this behavior. Early micromachined pressure transducers with 25-μm-thick membranes and up to 1 cm^2 area consumed 5×10^{-3} watt of power, using a piezoresistive transducer that was sensitive to forces of barely 10^{-2} newton. Reduction of the deflection membrane thickness to 1 μm resulted in a linear size reduction by a factor of nearly 100 and a corresponding increase in force sensitivity to 10^{-6} N, with slightly reduced piezoresistive power dissipation. The next advance involved changing the transduction mechanism from a piezoresistive element to a micromachined vacuum-encapsulated resonant microbeam, typically 200 μm long, 40 μm wide, and 1 μm thick, that optically sensed force via a shift in resonant frequency. An improvement in force sensitivity to better than 10^{-10} N with an input power requirement of 10^{-13} watt was the result. These improvements may be further extended through sensing concepts based on tunneling and atomic force microscopes, and ultimately through magnetic resonance force microscopes with 60-nanometer-thick resonant beams. These devices have resolved forces of 10^{-18} N (equivalent to masses of tens of attograms), and a thermal noise-limited sensitivity of 10^{-19} N is anticipated. *See* BROWNIAN MOVEMENT; MICROSENSOR; PRESSURE TRANSDUCER; SCANNING TUNNELING MICROSCOPE.

Actuators and corresponding linkage mechanisms are three-dimensional devices that operate via changes in stored energy. The goal is a certain force-versus-displacement character that determines the required volume necessary to be devoted to energy storage, thereby implying a three-dimensional structure. Direct scaling of macroscopic counterparts suggests potential problems for micromechanisms in behavior that is affected by changes in the surface-to-volume ratio. These issues include substantially reduced inertial effects that occur with size reduction, and the relatively greater influence of friction, surface tension, and air damping, which all become more significant relative to mass, particularly in comparison to a 1-horsepower (750-watt) motor or engine, for example. Microactuators and mechanisms based on planar batch integrated circuit fabrication techniques are restricted to essentially two-dimensional elements, a requirement that challenges micro-mechanism design. High-aspect-ratio micromachining (HARM) has been developed to alleviate the two-dimensional constraint. High-aspect-ratio processing also aids in the manufacture of inertial sensors, for example, in defining inertial reference masses for applications such as seismic sensing. *See* SEISMOGRAPHIC INSTRUMENTATION.

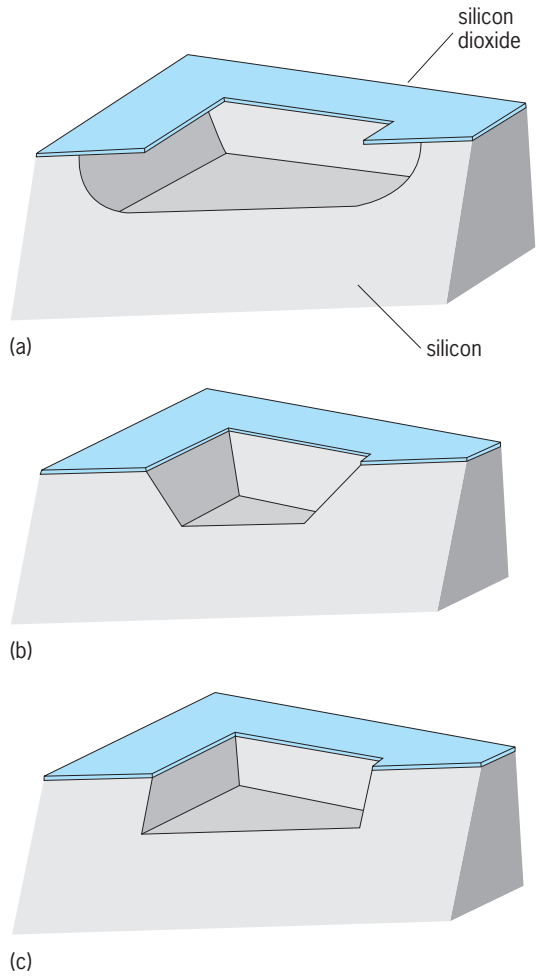

Fig. 1. Silicon etching processes. (*a*) Isotropic etching. (*b*) Anisotropic crystal etching. (*c*) Vertical anisotropic etching.

Microfabrication technology. The development of process tools and materials for MEMS is the pivotal enabler for integration success. A material is chosen and developed for its mechanical attributes and patterned with a process amenable to co-electronic fabrication. Two basic approaches to patterning a material are used. Subtractive techniques pattern via removal of unwanted material, while additive techniques make use of temporary complementary molds within which the resulting structure conforms. In either case, the goal is precision microstructural definition. Both approaches use a mask to transfer a pattern to the desired material. For batch processes, this step typically occurs via photolithography and may itself entail several steps. The basic process is to apply a photoresist, a light-sensitive material, and use a photomask to selectively expose the photoresist in the desired pattern. A solvent chemically develops the photoresist-patterned

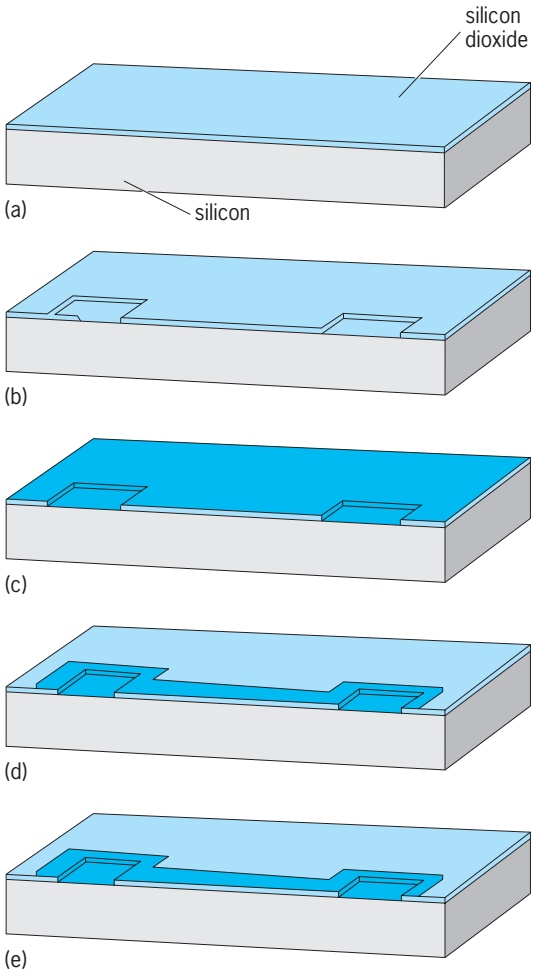

Fig. 2. Basic surface micromachining process. (*a*) Thermal oxidation of silicon. (*b*) Silicon dioxide patterning. (*c*) Polysilicon deposition. (*d*) Polysilicon patterning. (*e*) Release etching of silicon dioxide to render the mechanical structure.

image, which then may be used as a mask for further processing.

Subtractive processing is accomplished via chemical etching. Wet etching occurs in the liquid phase, and dry etching or gas-phase etching may occur in a vapor phase or plasma. A key issue in subtractive patterning is selectivity, which is defined as the ratio of the etch rate of the material that is to be removed to the etch rate of the masking material or any other resident material that the designer wants to be substantially unaffected. *See* PLASMA (PHYSICS).

Bulk micromachining. An omnipresent material in MEMS due to its use as a semiconductor for microelectronics fabrication is single-crystal silicon. In terms of mechanical stability, silicon is also an outstanding material possessing an exceptionally linear mechanical response as well as resistance to the aging effects prevalent in metals and plastics. Single-crystal compound semiconductor materials, such as gallium arsenide, are also of interest due to their use in photonics and microwave circuitry, but are not as widely developed as silicon technology. In either case, these substrate materials are attractive for use in MEMS if precision mechanical structures can be constructed from them in a fashion that does not interfere with their electronic function.

A primary microfabrication technology that has been used for most commercial devices is bulk micromachining, which is the process of removing, or etching, substrate material. The important aspect of precision bulk micromachining is etch directionality. The two limiting cases are isotropic, or directionally insensitive, and anisotropic, or directionally dependent, up to the point of being unidirectional. All cases may be illustrated with single-crystal silicon (**Fig. 1**). In many instances, the photoresist possesses insufficient integrity as an etch mask, and an intermediate layer of material is required to be patterned and used as a mask layer. A submicrometer-thick layer of silicon dioxide or silicon nitride, for example, commonly serves as this mask layer.

Isotropic silicon etching may be carried out through liquid chemistry. For example, a mixture of hydrofluoric acid and nitric acid may be applied, whereby the silicon is oxidized by the nitric acid and the hydrofluoric acid converts the resulting silica to a soluble silicon fluoride compound. A nearly directionally independent etch rate results that may be influenced by agitation (Fig. 1*a*). Alternatively, a plasma with a gas such as sulfur hexafluoride may be used to realize dry silicon isotropic etching, and non-plasma-vapor techniques also exist with xenon difluoride vapor.

Many crystalline materials such as quartz and silicon may be etched preferentially along certain crystal orientations. This attribute may be exploited to chemically machine geometries defined by crystal planes. In the case of silicon, the etch rates in the <100> and <110> crystal directions can be several hundred times greater than in the <111> direction. This type of liquid etching may be accomplished with various alkaline compounds such as potassium hydroxide (KOH) and tetra-methyl ammonium hydroxide (TMAH). The results for a [100] oriented silicon substrate are sidewalls angled at 54.7° with respect to the substrate surface, representing the [111] crystal planes (Fig. 1*b*). A variety of additional orientations are possible with the restriction that crystalline-based anisotropic etching is limited to faceted structures. Nevertheless, anisotropic

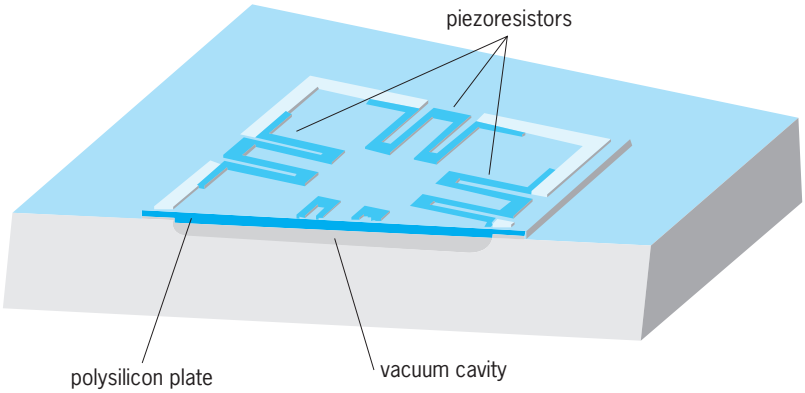

Fig. 3. Surface-micromachined pressure transducer.

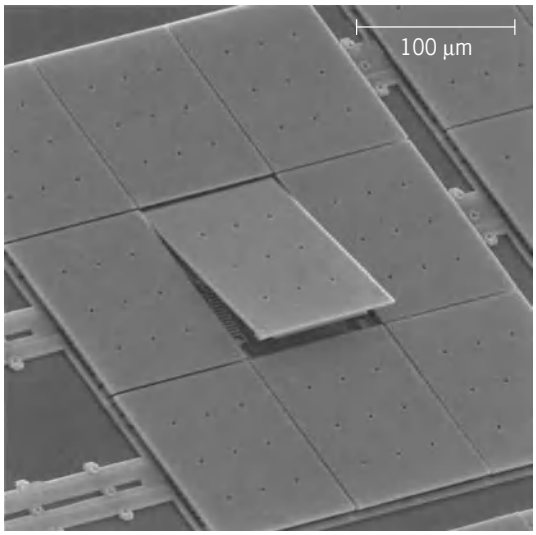

Fig. 4. Surface-micromachined movable mirror array. (*E. Garcia, Sandia National Laboratories*)

etching in conjunction with microelectronic processing has yielded a tremendous number of useful micromachined devices, including disposable ink-jet print heads, pressure transducers, accelerometers, and microelectrode arrays.

Extending the bulk micromachining tool further, anisotropic silicon plasma etching sequences have been devised to allow an arbitrary planar geometry to be nearly perfectly transferred into vertically defined prismatic cuts (Fig. 1*c*). This capability yields great flexibility in defining mechanical elements, with the drawback of a substantially increased tool cost over liquid and vapor-phase chemical etching. An extension of bulk micromachining to yield further structure flexibility is afforded through wafer-to-wafer bonding techniques. Two or more wafers containing etched geometry are aligned and bonded using direct fusion bonding, anodic bonding (with the assistance of a high voltage), or an intermediate adhesive layer. These techniques have resulted in new structures, including precisely defined cavities for pressure transducers and microchannels for microfluidic handling, as well as a batch approach to packaging, a particularly troublesome and expensive part of interfacing mechanical behavior with electronics. *See* ELECTRONIC PACKAGING.

Surface micromachining. An alternative processing approach to bulk microfabrication was driven by the desire to reduce the fraction of the substrate area that had to be devoted to the mechanical components, thereby allowing a larger number of device dies per wafer. The approach, termed surface micromachining (SMM), realizes mechanical structures by depositing and patterning mechanical material layers in conjunction with sacrificial spacer material layers (**Fig. 2**). In contrast to bulk micromachining, where a substrate may be between 250 and 750 μm thick, these deposited films, realized through deposition processes such as sputtering, evaporation, and chemical vapor deposition (CVD), have thicknesses of the order of 1 μm (micrometer). *See* CRYSTAL GROWTH; SEMICONDUCTOR HETEROSTRUCTURES; SPUTTERING.

A substrate material, such as silicon, is used as the basis material for these depositions. The first step in the basic process is to deposit or grow a material to be used as a sacrificial material, such as a silicon dioxide layer (Fig. 2*a*). The sacrificial layer is patterned to open anchor regions for the structural layer (Fig. 2*b*) that is subsequently deposited (Fig. 2*c*). Polysilicon is commonly used for structural layers in surface micromachining, along with other materials that include silicon nitride, silicon carbide, and diamondlike carbon. The structural layer is patterned (Fig. 2*d*), and the sequence of sacrificial and structural depositions and patterning may then be repeated to enable multilayer geometry. The consequence of this repeated process, however, is an increase in topological variation that ultimately poses problems for patterning. Thus, planarization steps may be inserted to maintain a reasonably flat surface on which to continue to perform microlithography. Ultimately, the last step in silicon micromachining is a release etch of the sacrificial material to render the mechanical

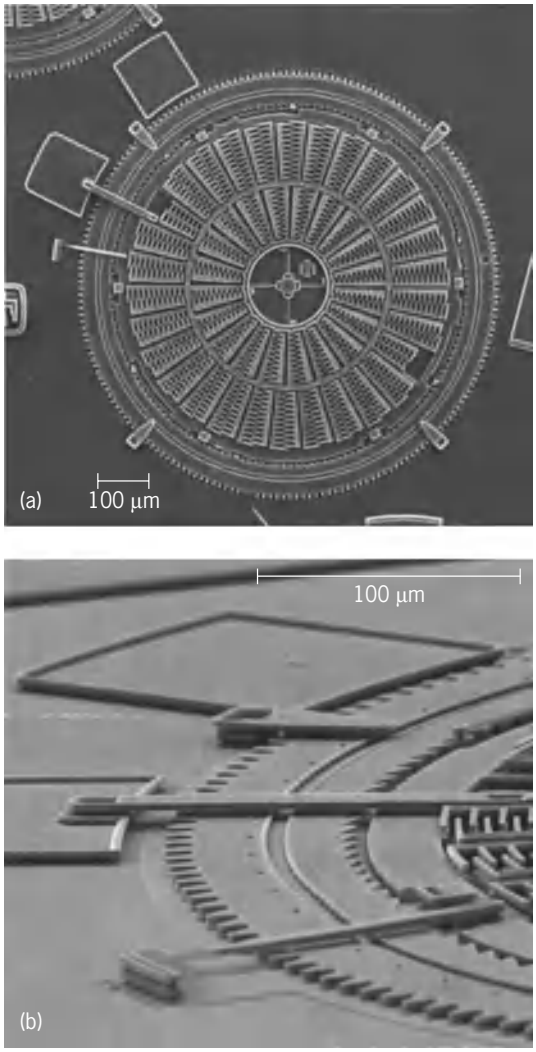

Fig. 5. Torsional ratcheting actuator fabricated by surface micromachining. (*a*) Overview. (*b*) Close-up. (*J. Jakubczala, Sandia National Laboratories*)

structure. In Fig. 2, this structure is a clamped-clamped beam that may be used, for example, in a resonant stain gage (Fig. 2*e*).

Design. The successful use of any microfabrication process is predicated on the ability to control the mechanical properties of the deposited layers and implement these properties in appropriate design models. Material properties of deposited materials not only differ substantially from their bulk counterparts but also vary depending on deposition conditions. Diagnostic structures must therefore be prepared to measure mechanical properties such as yield strength, bulk modulus, and internal strain. Mechanical test microstructures that may be microfabricated with the same process as the actual device of interest have been developed and are used to measure these properties "on chip." The MEMS designer is then faced with the task of identifying appropriate models that identify all physical behavior occurring in the microsystem. This task can be daunting due to the interconnectivity between the many different energy transfer mechanisms, but may be reduced by computer-aided design tools that are being tailored for MEMS design problems. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

The path to consolidating micromachining sequences with microelectronics and components from other fields such as photonics is termed process integration. Careful ordering of the process sequence is required to mediate the effects of high temperatures and different chemistries present in disparate and sometimes incompatible microelectronics and micromechanical processes.

Applications. A highly successful device that is fabricated with both bulk and surface micromachining is the integrated pressure transducer. One surface-micromachined pressure transducer, for example, can measure absolute pressure ranges as high as 70 megapascals (10,000 lb/in.^{2}) and as low as a few pounds per square inch (**Fig. 3**). The process sequence uses surface micromachining techniques to form a polysilicon-plate-covered cavity that is initially open. The open cavity is then reactively vacuum-sealed with a silicon dioxide, silicon nitride, or polysilicon deposition. After the cavity is closed off, the gas remaining in the cavity continues to react, depositing a solid and thereby self-pumping a vacuum in the cavity. The resulting vacuum reference enables measurement of absolute pressures. Plate deflection is monitored with carefully located polysilicon piezoresistors. Application areas include air pressure sensing in automobile engines, environmental monitoring, and blood pressure sensing. Similar processing has resulted in the integration of surface-micromachined polysilicon inertial reference proof masses with microelectronic processing, yielding single-chip force-feedback accelerometers capable of measurement ranges from a few *g* to hundreds of *g*, where *g* is the acceleration of gravity. Bulk-micromachined single-crystal silicon versions exist that are able to measure several hundred thousand *g*, attesting to the robustness of microsensors. *See* ACCELEROMETER.

The use of surface micromachining technology to implement microactuators has resulted in steerable micromirror arrays with as many as 1024 × 768 pixels on a chip. These arrays have revolutionized digital display technology. One device in this category is constructed from four levels of polysilicon and contains 100-μm-square mirrors, separated by 1 μm, that may be tilted up to 10° (**Fig. 4**). Further electrostatic microactuator designs are possible and may be extremely intricate, such as a torsional ratcheting actuator fabricated with five polysilicon levels (**Fig. 5**). These types of devices are suited for a variety of micropositioning applications. Processing based on deep-x-ray lithography has been used to produce precision magnetic microactuators. One such microactuator directly switches a single-mode optical fiber in a 1 × 2 switch configuration (**Fig. 6**). A prime obstacle in manufacturing integrated devices of this sort is that the packaging required to accommodate the interface of the mechanical world with a centralized information network is unique and requires novel design and testing approaches to be economically feasible.

Todd R. Christenson

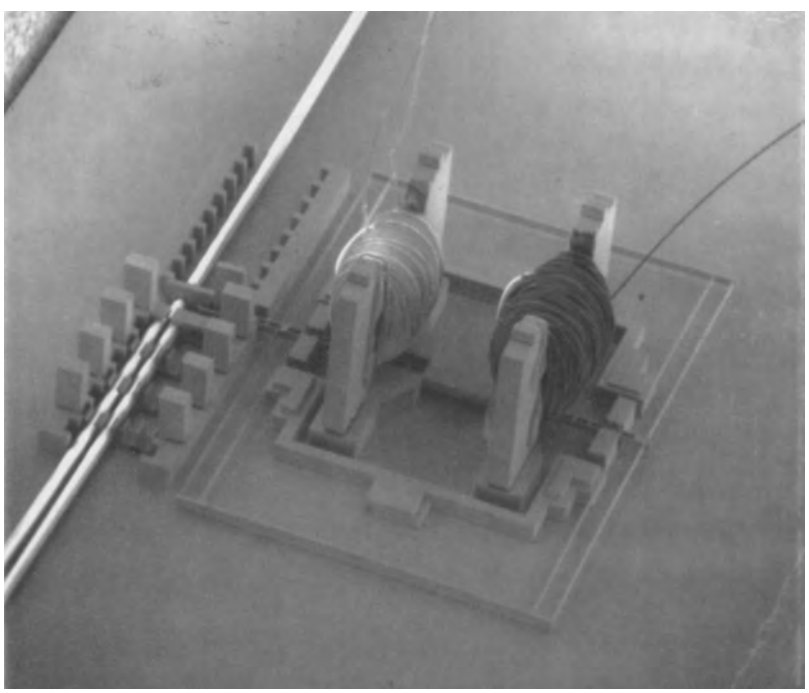

Fig. 6. Magnetic 1 × 2 optical fiber switch fabricated by deep-x-ray lithography. Total device size is approximately 4 mm × 4 mm. (*Henry Guckel, University of Wisconsin*)

Bibliography. G. T. A. Kovacs, *Micromachined Transducers Sourcebook*, McGraw-Hill, Boston, 1998; M. Madou, *Fundamentals of Microfabrication*, CRC Press, Boca Raton, FL, 1997; N. Maluf, *An Introduction to Microelectromechanical Systems Engineering*, Artech House, 1999; P. Rai-Choudhury (ed.), *Handbook of Microlithography, Micromachining, and Microfabrication*, SPIE Optical Engineering Press, Bellingham, WA, 1997; W. Trimmer, *Micromechanics and MEMS: Classic and Seminal Papers to 1990*, IEEE Press, 1997.

Microfluidics

Technology that involves manipulating fluids in structures in which at least one linear dimension is less than a millimeter. Currently, the smallest dimension of common microfluidic elements is 10–500 micrometers. The most mature and commercially successful microfluidic systems are the print heads found in common inkjet printers. More recent efforts have been aimed at developing microfluidic systems for performing chemical processes. The principal target applications of these chemical systems are as portable analyzers of chemicals and biomolecules [proteins and deoxyribonucleic acid (DNA)] for biomedical applications and for defense against chemical and biological weapons. Some chemical syntheses have also been performed in microfluidic devices (**Fig. 1**). Microfluidic technologies have only recently begun to be commercialized. The **table** lists applications of microfluidic systems.

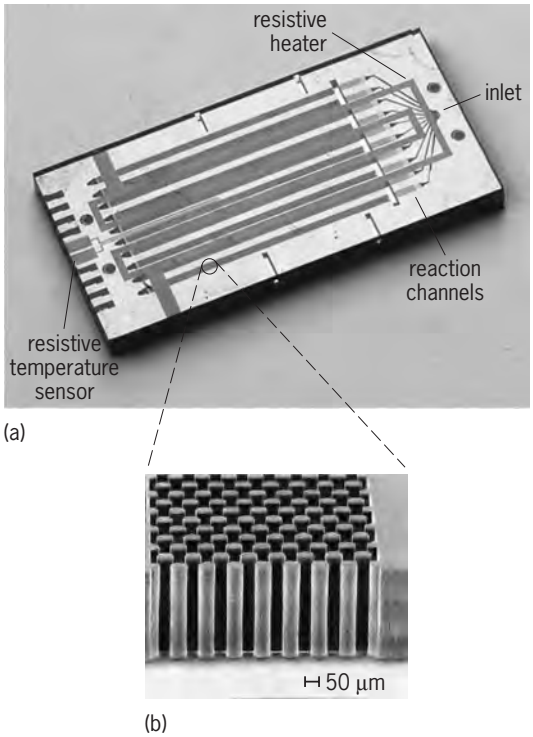

Fig. 1. Microfluidic chip for performing multiphase (liquid/gas) reactions with heterogeneous catalysis. (*a*) Photograph of chip. This system was used for the hydrogenation of alkenes to form alkanes. The channel structures are in silicon. The heater and temperature sensor are thin metal films (*photo by Felice Frankel*). (*b*) Scanning electron micrograph showing some of the micropillars that are fabricated inside each reaction channel. These pillars provide a large surface area on which to present catalytic materials in the reactors (*from M. W. Losey et al., Design and fabrication of microfluidic devices for multiphase mixing and reaction, J. Microelectr. Sys., 11(6):709–717, 2002; © 2002 IEEE*)

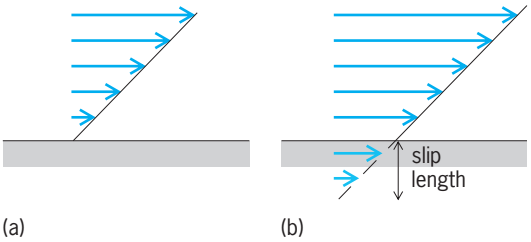

Fig. 2. Flow profiles at a solid boundary (*a*) with no slip and (*b*) with slip.

Flow characteristics. Experimental and theoretical attention has been paid to the possibility that fluid behavior in microchannels deviates from that found in macroscopic flows. Except in unusual cases, such deviations are neither predicted nor observed; for flows of simple liquids, such as water and organic solvents, the same general equations (Navier-Stokes equations) that govern macroscopic flows, with the no-slip condition at solid boundaries (meaning that the fluid is stationary relative to the solid at the boundary; **Fig. 2***a*), are valid on scales above tens of nanometers. One notable exception is the observation of slip of liquids over nonwetting surfaces, that is, surfaces on which the liquid beads up as water does on Teflon® (polytetrafluoroethylene). A proposed mechanism for this slip is that a layer of nanoscopic vapor bubbles covers the surface and acts as a lubricating layer of low viscosity between the liquid and the solid. The largest slip lengths that have been observed are about 1–2 μm (Fig. 2*b*). Thus, slip will lead to deviations from expected behavior only in flows of nonwetting liquids through channels with cross-sectional dimensions of less than 10 μm. In flows of gases (less commonly used than liquids), slip has long been predicted and has been observed in microchannels when the mean free path of molecules in the gas is comparable to the

Applications of microfluidic systems

Application	Advantage of microscopic scale
Inkjet printing	High resolution, high operation speed
Analytical chemistry	Portability, high operation speed, small sample volume (<1 μL)
Protein and DNA analysis for proteomics and genomics	
High-throughput screening of chemicals (pharmaceuticals)	
Portable chemical detection	
Chemical synthesis	Small sample volume, high operation speed, fine control of reaction parameters, rapid transfer of heat and mass
In-situ synthesis of toxic compounds	
Combinatorial chemistry	
Analysis of chemical kinetics	
Fuel cell	Portability
Cooling for microelectronics	Portability, rapid heat transfer
Cell culture/cell sorting	Control of individual cells

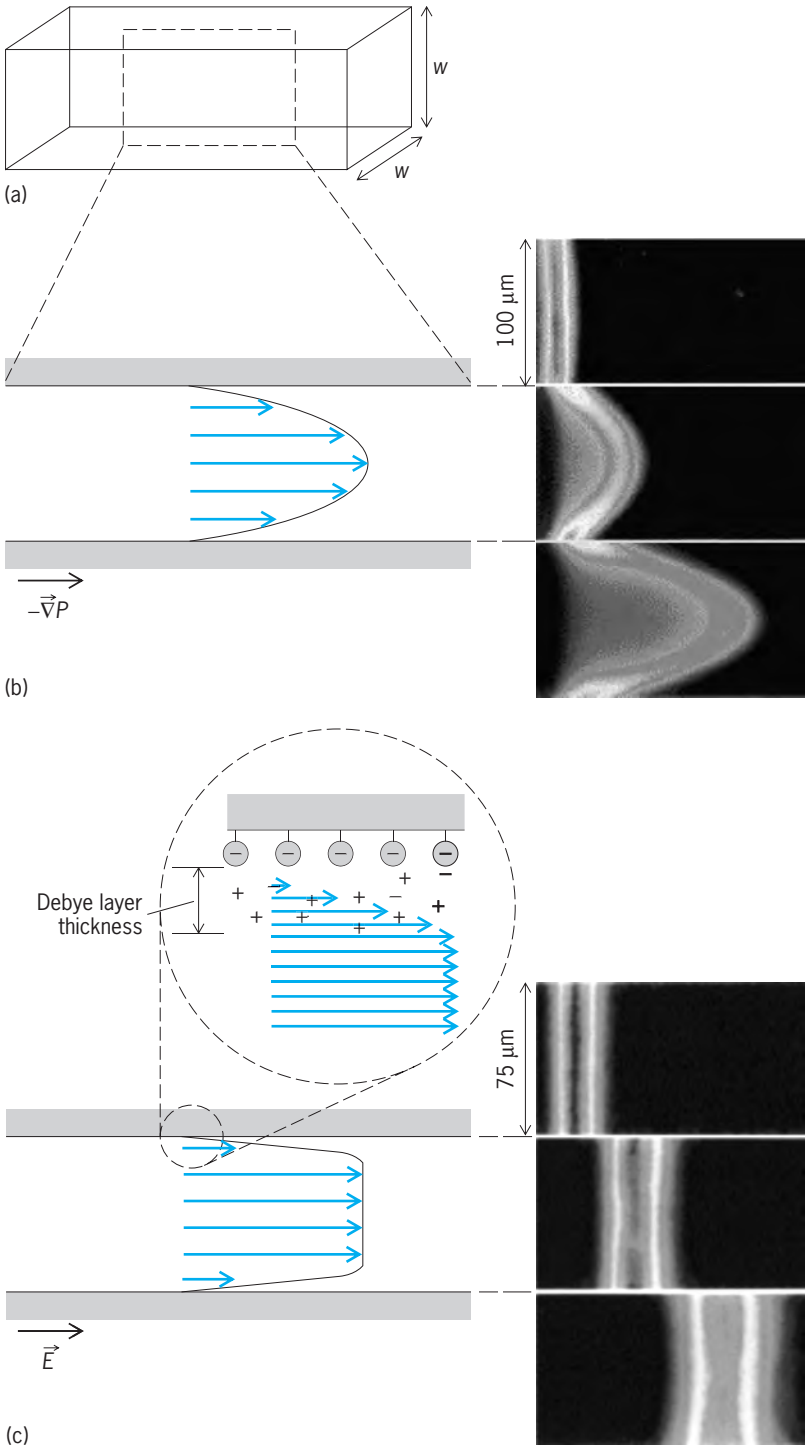

Fig. 3. Flows in microchannels. (*a*) Diagram of a section of a microchannel with a square cross section. (*b*) Flow profile of a pressure-driven flow and (*c*) an electroosmotic flow in a channel. The two series of fluorescent micrographs show the evolution of a band of fluorescent dye in the respective flows. The blurring of the interface between fluorescent and nonfluorescent regions is due to molecular diffusion. (*After P. H. Paul, M. G. Garguilo, and D. J. Rakenstraw, Imaging of pressure and electrokinetically driven flows through open capillaries, Analyt. Chem., 70:2459–2467, 1998*)

dimension of the channel. *See* BOUNDARY-LAYER FLOW; FLUID-FLOW PRINCIPLES; NAVIER-STOKES EQUATION.

While the basic governing equations are the same, there are several general features that distinguish flows in microstructures from flows in common macroscopic systems (such as water faucets and coffee cups). As the dimension of a flow decreases, the importance of forces that act on the volume of the fluid, such as inertia and gravity, diminishes relative to that of forces that act at the surfaces, such as viscous friction and surface tension. The ratio of inertial to viscous forces is expressed as the Reynolds number of the flow, $\mathrm{Re} = \rho v w/\eta$, where ρ (measured in kg/m^3) is the density of the fluid, v (m/s) is the velocity, w (m) is the characteristic dimension of the flow (~volume/surface area of container; **Fig. 3***a*), and η [kg/(m · s)] is the dynamic viscosity of the fluid. In microchannels, $\mathrm{Re} < 100$, so flows are laminar rather than turbulent (flows in channels are typically laminar for $\mathrm{Re} < 2000$). Coflowing streams in a laminar flow intermix only by diffusion because no spontaneous eddies carry mass and momentum between them. This characteristic of laminar flows allows spatial control of solute within the flow, but it hinders rapid mixing. *See* LAMINAR FLOW; REYNOLDS NUMBER; TURBULENT FLOW.

Fabrication. Most methods of fabricating microfluidic systems are adaptations of techniques for fabricating microelectronic structures; photolithography is at the core of these methods. These techniques lead to a flat, chiplike format in which microchannels are constrained in planar layers on a flat substrate (Fig. 1). This format facilitates integration of electronics and optics, but it restricts the geometries that are accessible for the design of fluidic elements. The channel structures are etched into a flat surface of silicon or glass, or molded into a soft or hard plastic sheet; the channels are closed by sealing the structured material to another flat surface. *See* INTEGRATED CIRCUITS; MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS); MICROLITHOGRAPHY.

Flow characterization. Simple characterization of flows in microfluidic devices can be achieved by measuring global variables such as the applied pressure difference (Pa) across a channel and the volumetric flow rate (m^3/s) through the channel. Electronic sensors of pressure and flow speed have been integrated into microchannels to provide local measurements of the flow. Fluorescence microscopy is commonly used to perform characterization of these flows with micrometer-scale resolution. Micro particle image velocimetry (μPIV) maps the streamlines in a flow by following the trajectories of submicrometer fluorescent tracer beads in sequences of micrographs. The evolution of streams in microflows can be followed through a microscope by coinjecting streams of fluorescent solution with streams of clear solution. In steady flows, the three-dimensional evolution of fluorescent streams can be imaged through a confocal microscope (**Fig. 4**). *See* CONFOCAL MICROSCOPY; FLUORESCENCE MICROSCOPE.

Pumping fluids. The two most common methods for driving flows in microchannels are with externally applied pressure gradients and with externally applied electric fields. Capillary, acoustic, and magnetic forces have also been used to move fluids in microchannels.

Pressure gradients are often generated by macroscopic pumps that are connected via tubes to the

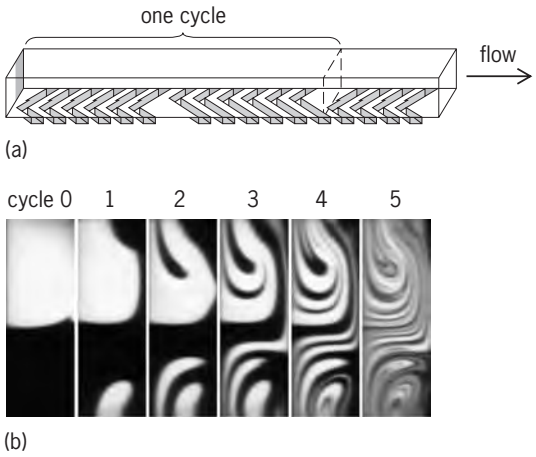

Fig. 4. Passive micromixer. (*a*) Section of a microchannel with grooves in the form of asymmetric herringbones on the bottom wall; the grooves lead to mixing in pressure-driven flows through the channel. One mixing cycle is made up of two regions of asymmetric herringbones; the direction of the asymmetry switches from one region to the next. (*b*) Confocal fluorescence micrographs of the vertical cross sections of flow through such a channel. The evolution of two coflowing streams (one of fluorescent liquid and one of clear liquid) in the mixer is seen: The number of folds approximately doubles with each cycle. The blurring of the interface between fluorescent and nonfluorescent regions is due to molecular diffusion. (*After A. D. Stroock et al., Chaotic mixer for microchannels, Science, 295:647–651, 2002*)

inlets of the microfluidic chip. Integrated microdiaphragm and microperistaltic pumps have also been developed. In pressure-driven flow in channels, the flow speed is given roughly by (1), where ∇P is the

$$v \sim -w^2 \nabla P/\eta \qquad \text{(m/s)} \tag{1}$$

applied pressure gradient (Pa/m). To maintain a given flow speed, the required pressure gradient grows rapidly as the channel shrinks. In pressure-driven flows, the flow speed varies from zero at the wall to its maximum value in the middle of the channel (Fig. 3*b*). This variation leads to dispersion (spreading) of solute along the direction of the flow. This dispersion is unfavorable for transporting narrow bands of solute such as in a chemical separation.

Electric fields are applied by placing electrodes in the inlet and outlet of a microchannel. In electrolyte solutions (such as salty water), flows are generated by the interaction of the applied electric field with ions that accumulate in a thin (less than 1 μm thick) layer of fluid, the Debye layer, adjacent to channel walls; these are called electroosmotic flows (Fig. 3*c*). Fluid in the Debye layer moves due to the electrical body force and entrains the remainder of the fluid. The flow speed in the bulk is given roughly by (2), where μ_{eo} ($m^2/s \times V$) is the electroosmotic

$$v_{eo} \sim \mu_{eo} E \tag{2}$$

mobility that depends on the surface charge density, the concentration of ions in the liquid, and the viscosity of the liquid, and E (V/m) is the magnitude of the electric field. The magnitude of the electroosmotic mobility is given roughly by (3). In uniformly

$$|\mu_{eo}| \sim 1\ \mu\text{m}/(\text{s} \cdot \text{V/m}) \tag{3}$$

charged channels, the flow speed v_{eo} is constant over the cross section except in the Debye layer. This property allows bands of a single type of solute to be transported along the channel with little dispersion. Capillary electrophoresis is a useful tool for chemical analysis that exploits this behavior and is performed in microfluidic devices. The flow speed v_{eo} is also independent of the dimension of the channel down to the thickness of the Debye layer. Electroosmotic pumping is therefore appropriate in channels down to submicrometer dimensions. Disadvantages of electroosmotic pumping include the requirement of high voltages ($V \sim 1$ kV), sensitivity to the chemical characteristics of the walls, and the tendency for species to separate electrophoretically in the applied field. (This tendency is an advantage for analysis but a disadvantage for transporting general chemical mixtures.) *See* ELECTROPHORESIS.

Mixing. In microchannels of common dimensions ($500\ \mu\text{m} > w > 10\ \mu\text{m}$), the motion of solute by diffusion across the laminar flow is often slow relative to the flow speed along the channel. In this situation, mixing is often the slow step in a chemical process, unless transverse flows that stir the fluid are purposely induced. (In turbulent flow, these transverse flows occur spontaneously.) An efficient stirring flow stretches and folds the fluid such that the interface between unmixed regions grows exponentially in time; this type of flow is called chaotic (Fig. 4). Active mixers create transverse motion in the principal flow with local, oscillatory forces generated with bubbles, applied electric or magnetic fields, or flows in cross-channels. These mixers have the potential to be very efficient but require sophisticated controls. Passive mixers use fixed geometrical features in the channel to induce transverse components in the principal flow (Fig. 4). These designs are simple to operate but may not achieve full mixing as quickly as active designs.

Challenges. With the exception of inkjet print heads and, perhaps, on-chip capillary electrophoresis, microfluidic technology is still in its infancy. No consensus has been reached on even the most basic procedures and designs. Optimal designs for pumps, valves, injectors, and so forth must still be invented and characterized, and ideal materials and methods of microfabrication must be developed. Future applications of microfluidic devices must also be conceived of and explored. For example, microfluidic systems could act as implants to deliver drugs and monitor physiological parameters, or as active materials that control environmental parameters in clothing and buildings.

Nanofluidic systems. These systems, in which the characteristic size is less than a micrometer, are already being developed. In this regime, pressure may not be a useful driving force, whereas electroosmosis likely will. Diffusion will often transport solute molecules as rapidly as the flow does; mix-

ing will not be a problem, but careful delivery of solute will. Whereas the design principles for microfluidics were largely borrowed from macroscopic chemical systems (pumps, channels, and so on), the best guides to designing nanofluidic systems may be biological systems such as cells: transport will be controlled on the molecular rather than the hydrodynamic level, with molecular motors carrying chemical building blocks along definite paths between distinct chemical environments. *See* FLUID MECHANICS.

Abraham D. Stroock

Bibliography. C.-M. Ho and Y.-C. Tai, Micro-electro-mechanical-systems and fluid flows, *Annu. Rev. Fluid Mech.*, 30:579–612, 1998; M. W. Losey et al., Design and fabrication of microfluidic devices for multiphase mixing and reaction, *J. Microelectr. Sys.*, 11(6):709–717, 2002; P. H. Paul, M. G. Garguilo, and D. J. Rakestraw, Imaging of pressure and electrokinetically driven flows through open capillaries, *Analyt. Chem.*, 70:2459–2467, 1998; A. D. Stroock et al., Chaotic mixer for microchannels, *Science*, 295:647–651, 2002; D. C. Tretheway and C. D. Meinhart, Apparent fluid slip at hydrophobic channel walls, *Phys. Fluids*, 14:L9–L12, 2002; G. M. Whitesides and A. D. Stroock, Flexible methods for microfluidics, *Phys. Today*, 54(6):42–48, June 2001.

Microlithography

The formation of small three-dimensional relief images on the surface of a substrate for subsequent transfer of this pattern into the substrate itself, as used in such applications as semiconductor fabrication. The fabrication of an integrated circuit (IC) requires a variety of physical and chemical processes performed on a semiconductor (for example, silicon) substrate. In general, the various processes used to make an IC fall into three categories: film deposition, patterning, and semiconductor doping. Films of both conductors (such as polysilicon, aluminum, and more recently copper) and insulators (various forms of silicon dioxide, silicon nitride, and others) are used to connect and isolate transistors and their components. Selective doping of various regions of silicon allows the conductivity of the silicon to be changed with the application of voltage. By creating structures of these various components, millions of transistors can be built and wired together to form the complex circuitry of a modern microelectronic device. Fundamental to all of these processes is lithography. *See* INTEGRATED CIRCUITS; SEMICONDUCTOR.

The word lithography, from Greek, literally means writing on stones. In the case of semiconductor lithography, our stones are silicon wafers and we are writing our patterns in a light-sensitive polymer called a photoresist. To build the complex structures that make up a transistor and the many wires that connect the millions of transistors of a circuit, lithography and etch pattern transfer steps are repeated at least 10 times, but more typically are done 20–30 times in order to make one circuit. Each pattern being printed on the wafer is aligned to the previously formed patterns, and slowly the conductors, insulators, and selectively doped regions are built up to form the final device. *See* POLYMER; TRANSISTOR.

Optical microlithography is a photographic process by which a photoresist is exposed and developed to form three-dimensional images on the substrate. The general sequence of processing steps for a typical optical lithography process is as follows: substrate preparation, photoresist spin coat, prebake, exposure, postexposure bake, development, and postbake. The resist is removed in the final operation of the lithographic process, after the pattern has been transferred into the underlying layer. This sequence is shown in **Fig. 1**, and is generally performed on several tools linked together into a contiguous unit called a lithographic cluster.

Substrate preparation. Substrate preparation is intended to improve the adhesion of the photoresist material to the substrate. This is accomplished by one or more of the following processes: substrate cleaning to remove contamination, dehydration bake to remove water, and addition of an adhesion promoter. One common type of contaminant, adsorbed water, is removed most readily by a high-temperature process called a dehydration bake. A typical dehydration bake does not completely remove water from the surface of silica substrates (including silicon, polysilicon, silicon oxide, and silicon nitride). Surface silicon atoms bond strongly with a monolayer of water forming silanol groups (SiOH). Bake temperatures in excess of 600°C (1112°F) are required to remove this final layer of water. Since this approach is impractical, the preferred method of removing this silanol is by chemical means. Adhesion promoters

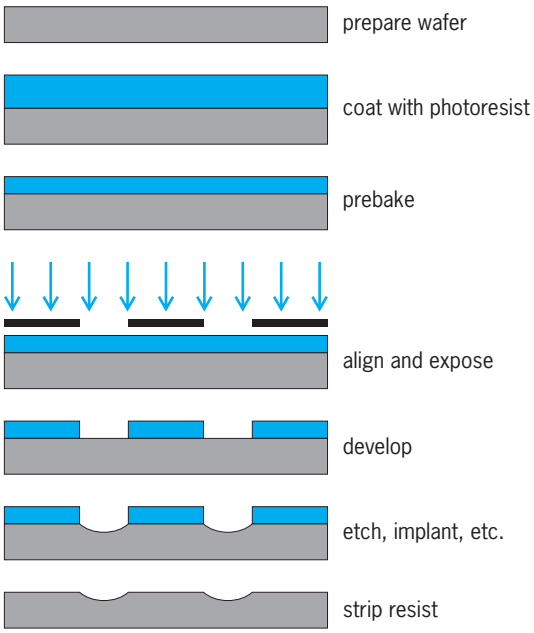

Fig. 1. Example of a typical sequence of lithographic processing steps (with no postexposure bake in this case), illustrated for a positive resist.

are used to react chemically with surface silanol and replace the —OH group with an organic functional group which, unlike the hydroxyl group, offers good adhesion to the photoresist. Silanes are often used for this purpose, the most common being hexamethyldisilizane (HMDS).

Photoresist coating. A thin, uniform coating of photoresist at a specific, well-controlled thickness is accomplished by the seemingly simple process of spin coating. The photoresist, rendered into a liquid by dissolving the solid components in a solvent, is poured onto the wafer and then spun on a turntable at a high speed to produce the desired film. Stringent requirements for thickness control and uniformity and low defect density call for particular attention to be paid to this process, where a large number of parameters can have significant impact on these properties. There is the choice between static dispense (wafer stationary while resist is dispensed) or dynamic dispense (wafer spinning while resist is dispensed), spin speeds and times, and accelerations to each of the spin speeds. Also, the volume of the resist dispensed and properties of the resist (such as viscosity, percent solids, and solvent composition) and the substrate (substrate material and topography) play an important role in the resist thickness uniformity. At the end of this cycle, a solvent-rich film of photoresist covers the wafer, which is ready for the postapply bake. Although theory exists to describe the spin coating process rheologically, in practical terms the variation of photoresist thickness and uniformity with the process parameters must be determined experimentally. A photoresist spin-speed curve is an essential tool for setting the spin speed to obtain the desired resist thickness. The final resist thickness varies as one over the square root of the spin speed and is roughly proportional to the liquid photoresist viscosity.

Postapply bake. After coating, the resulting resist film will contain 20–40% by weight solvent. The postapply bake process, also known as a softbake or a prebake, involves drying the photoresist by removing this excess solvent. There are four major effects of removing the solvent from a photoresist film: (1) film thickness is reduced, (2) postexposure bake and development properties are changed, (3) adhesion is improved, and (4) the film becomes less tacky and thus less susceptible to particulate contamination. Typical prebake processes leave between 3 and 8% residual solvent in the resist film, a sufficiently low amount to keep the film stable during subsequent lithographic processing.

Unfortunately, there are other consequences of baking most photoresists. At temperatures greater than about 70°C (160°F), the photosensitive component of a typical resist mixture, called the photoactive compound, may begin to decompose. Thus, one must search for the optimum prebake conditions which will maximize the benefits of the solvent evaporation and minimize the detriments of resist decomposition. For photoresists known as chemically amplified resists, residual solvent can significantly influence diffusion and reaction properties during the postexposure bake, necessitating careful control over the postapply bake process. Fortunately, these resists do not suffer from significant decomposition of the light-sensitive components during prebake.

Although the use of convection ovens for prebaking the photoresist was once quite common, currently the most popular bake method is the hot plate. The wafer is brought into either intimate vacuum contact with or close proximity to a hot, high-mass metal plate. Due to the high thermal conductivity of silicon, the photoresist is heated to near-hot-plate temperature quickly (in about 5 s for hard contact or in about 20 s for proximity baking). The greatest advantage of this method is an order-of-magnitude decrease in the required bake time (to about 1 min) over convection ovens and the improved uniformity of the bake. In general, proximity baking is preferred to reduce the possibility of particle generation caused by contact with the back side of the wafer.

Alignment and exposure. The basic principle behind the operation of a photoresist is a change in solubility of the resist in developer upon exposure to light (or other type of exposing radiation). For the case of the standard diazonaphthoquinone positive photoresist, the photoactive compound, which is not soluble in the aqueous base developer, is converted to a carboxylic acid upon exposure to ultraviolet (UV) light in the range 350–450 nanometers. The carboxylic acid product is very soluble in the basic developer. Thus, a spatial variation in light energy incident on the photoresist will cause a spatial variation in solubility of the resist in the developer.

Contact and proximity lithography are the simplest methods of exposing a photoresist through a master pattern called a photomask (**Fig. 2**). Contact lithography offers high resolution (down to about the wavelength of the radiation), but practical problems such as mask damage and resulting low yield make this process unusable in most production environments. Proximity printing reduces mask damage by keeping the mask a set distance (for example, 20 micrometers) above the wafer. Unfortunately, the resolution limit is increased to greater than 2–4 μm, making proximity printing insufficient for today's technology. By far the most common method of exposure is projection printing.

Projection lithography derives its name from the fact that an image of the mask is projected onto the wafer. Projection lithography became a viable alternative to contact/proximity printing in the

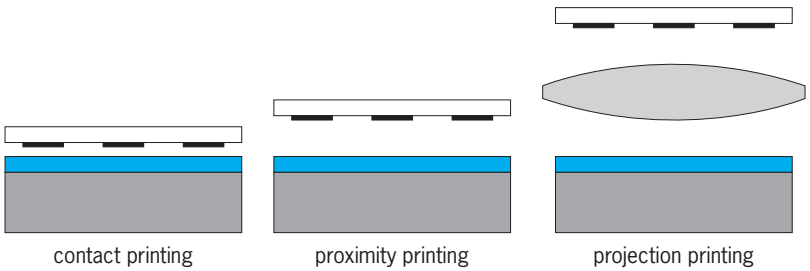

Fig. 2. Lithographic printing in semiconductor manufacturing has evolved from contact printing (in the early 1960s) to projection printing (from the mid 1970s to today).

mid-1970s, when the advent of computer-aided lens design and improved optical materials allowed the production of lens elements of sufficient quality to meet the requirements of the semiconductor industry. These lenses have become so perfect that lens defects, called aberrations, play only a small role in determining the quality of the image. Such an optical system is said to be diffraction-limited, since for the most part diffraction effects and not lens aberrations determine the shape of the image. *See* ABERRATION (OPTICS); DIFFRACTION; LENS (OPTICS); OPTICAL MATERIALS.

There are two major classes of projection lithography tools: scanning and step-and-repeat systems. Scanning projection printing employs reflective optics (that is, mirrors rather than lenses) to project a slit of light from the mask onto the wafer as the mask and wafer are moved simultaneously by the slit. The exposure dose is determined by the intensity of the light, the slit width, and the speed at which the wafer is scanned. These early scanning systems, which use polychromatic light from a mercury arc lamp, are 1:1 (that is, the mask and image sizes are equal). Step-and-repeat cameras (steppers for short) expose the wafer one rectangular section (called the image field) at a time and can be 1:1 or reduction. These systems employ refractive optics (that is, lenses) and are usually quasimonochromatic. Both types of systems (**Fig. 3**) are capable of high-resolution imaging, although reduction imaging is required for the highest resolutions. *See* MIRROR OPTICS.

Scanners replaced proximity printing by the mid-1970s for device geometries below 4–5 μm. By the early 1980s, steppers began to dominate as device designs pushed to 2 μm and below. Steppers continued to dominate lithographic patterning throughout the 1990s as minimum feature sizes reached the 250-nm levels. However, by the early 1990s a hybrid step-and-scan approach was introduced. The step-and-scan approach uses a fraction of a normal stepper field (for example, 25 × 8 mm), then scans this field in one direction to expose the entire 4× reduction mask. The wafer is then stepped to a new location and the scan is repeated. The smaller imaging field simplifies the design and manufacture of the lens, but at the expense of a more complicated reticle and wafer stage. Step-and-scan technology is the technology of choice today for below-250-nm manufacturing.

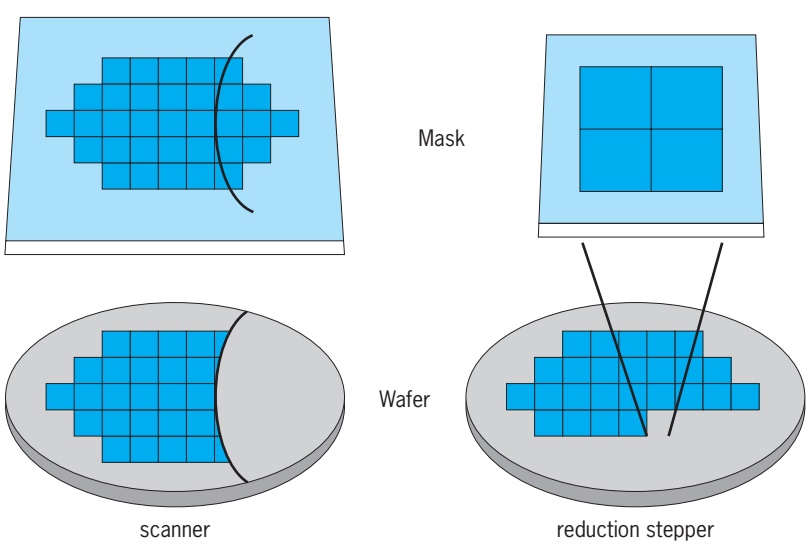

Fig. 3. Scanners and steppers use different techniques for exposing a large wafer with a small image field.

Resolution, the smallest feature that can be printed with adequate control, has two basic limits: the smallest image that can be projected onto the wafer, and the resolving capability of the photoresist to make use of that image. For the projected imaging, resolution is determined by the wavelength of the imaging light (λ) and the numerical aperture (NA) of the projection lens according to the Rayleigh criterion below. Lithography systems have progressed from

$$R \propto \frac{\lambda}{\mathrm{NA}}$$

blue wavelengths (436 nm) to UV (365 nm) to deep-UV (248 nm) to today's mainstream high-resolution wavelength of 193 nm. In the meantime, projection tool numerical apertures have risen from 0.16 for the first scanners to amazingly high 0.85 NA systems today, producing features well under 100 nm in size. *See* OPTICAL IMAGE; RESOLVING POWER (OPTICS).

Before the exposure of the photoresist with an image of the mask can begin, this image must be aligned with the previously defined patterns on the wafer. This alignment and the resulting overlay of the two or more lithographic patterns are critical since tighter overlay control means circuit features can be packed closer together. In the drive toward more functionality per chip, closer packing of devices through better alignment and overlay is nearly as critical as smaller devices through higher resolution.

Another important aspect of photoresist exposure is the standing-wave effect. Monochromatic light, when projected onto a wafer, strikes the photoresist surface over a range of angles, approximating plane waves. This light travels down through the photoresist and, if the substrate is reflective, is reflected back up through the resist. The incoming and reflected light interfere to form a standing-wave pattern of high and low light intensity at different depths in the photoresist. This pattern is replicated in the photoresist, causing ridges in the sidewalls of the resist feature (**Fig. 4**). As pattern dimensions become smaller, these ridges can significantly affect the quality of the feature. The interference that causes standing waves also results in a phenomenon called swing curves, the sinusoidal variation in linewidth with changing resist thickness. These detrimental effects are best cured by coating the substrate with a thin absorbing layer, called a bottom antireflective coating (BARC), that can reduce the reflectivity seen by the photoresist to less than 1%.

Postexposure bake. One method of reducing the standing-wave effect is called the postexposure bake (PEB). Although there is still some debate as to the mechanism, it is believed that the high temperatures used (100–130°C or 212–266°F) cause diffusion of

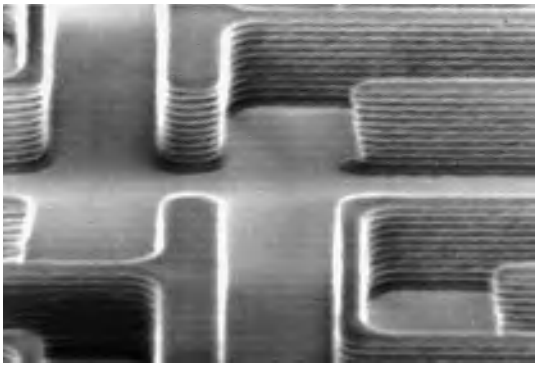

Fig. 4. Photoresist pattern on a silicon substrate showing prominent standing waves.

the exposed photosensitive material, thus smoothing out the standing-wave ridges. It has also been observed that the rate of diffusion is dependent on the prebake conditions since the presence of solvent enhances diffusion during a PEB; that is, a low-temperature prebake results in greater diffusion for a given PEB temperature. For a conventional resist, the main importance of the PEB is diffusion to remove standing waves. For a type of photoresist, called chemically amplified resists, the PEB is an essential part of the chemical reactions that create a solubility differential between exposed and unexposed parts of the resist. For these resists, exposure generates a small amount of a strong acid which does not itself change the solubility of the resist. During the postexposure bake, this photogenerated acid catalyzes a reaction which changes the solubility of the polymer in the resist. Since the photogenerated acid is not consumed in this reaction, it continues to cause more solubility changing events and thus "amplifies" the effects of the exposure. Control of the PEB is extremely critical for chemically amplified resists.

Development. Once exposed, the photoresist must be developed. Most commonly used photoresists use aqueous bases as developers. In particular, tetramethyl ammonium hydroxide (TMAH) is used almost universally at a concentration of 0.26 *N*. Development is undoubtedly one of the most critical steps in the photoresist process. The characteristics of the resist-developer interactions determine to a large extent the shape of the photoresist profile and, more importantly, the control of the sizes of the features being printed.

The method of applying developer to the photoresist is important in controlling the development uniformity and process latitude. Spin development processes, in which developer is poured onto the rotating wafer, use equipment similar to that used for spin coating. The wafer is also rinsed and dried while still spinning. Spray development has been shown to have good results using developers specifically formulated for this method. Using a process identical to spin development, the developer is sprayed on the wafer with a nozzle that produces a fine mist (**Fig. 5**). This technique reduces the amount of developer used and gives more uniform developer coverage. Another in-line development strategy is called puddle development. Again using developers specifically formulated for this process, the developer is poured onto a stationary wafer which is then allowed to sit motionless for the duration of the development time. The wafer is then spin-rinsed and dried. Note that all these processes can be performed in the same piece of equipment with only minor modifications, and combinations of these techniques are frequently used.

Pattern transfer. After the small patterns have been lithographically printed in photoresist, these patterns must be transferred into the substrate. There are three basic pattern transfer approaches: subtractive transfer (etching), additive transfer (selective deposition), and impurity doping (ion implantation). Etching is the most common pattern transfer approach. A uniform layer of the material to be patterned is deposited on the substrate. Lithography is then performed such that the areas to be etched are left unprotected (uncovered) by the photoresist. Etching is performed using wet chemicals (such as acids) or more commonly in a dry plasma environment. The photoresist "resists" the etching and protects the material covered by the resist. When the etching is complete, the resist is stripped, leaving the desired pattern etched into the deposited layer. Additive processes are used whenever workable etching processes are not available, for example, for copper interconnects. Here, the lithographic pattern is used to open areas where the new layer is to be grown (by electroplating, in the case of copper). Stripping of the resist then leaves the new material in a negative version of the patterned photoresist. Doping involves the addition of controlled amounts of contaminants that change the conductive properties of a semiconductor. Ion implantation uses a beam of dopant ions accelerated at the photoresist-patterned substrate. The resist blocks the ions, but the areas that are not covered by the resist are embedded with ions, creating the selectively doped regions that make up the electrical heart of the transistors. *See* ELECTROPLATING OF METALS; ION IMPLANTATION; PLASMA (PHYSICS); TRANSISTOR.

Limits of optical microlithography. To date, optical lithography has been the technology of choice for IC manufacturing. The resolution limit as described

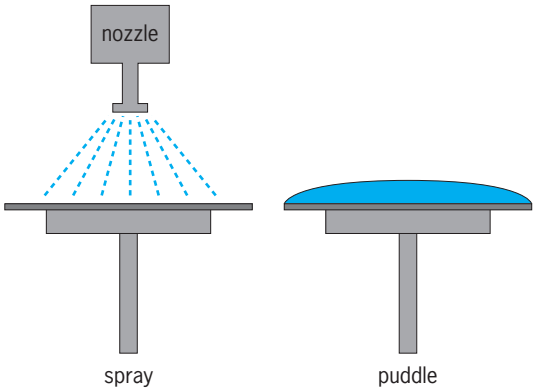

Fig. 5. Different developer application techniques that are commonly used.

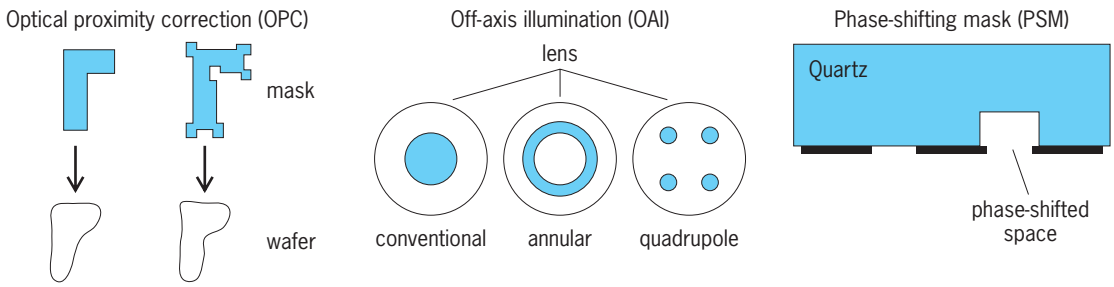

Fig. 6. Examples of several types of resolution enhancement technologies.

earlier in the equation outlines the obvious challenges. Lower wavelengths require expensive, unproven materials (such as superpure fused silica and fluoride salts). Higher numerical apertures result in increasing aberrations, which can only be reduced by more complicated designs and more exacting lens manufacturing processes.

To make the situation more difficult, any improvement in resolution is always accompanied by a decrease in the depth of focus (DOF). According to the Rayleigh criterion, the DOF for small features should decrease as the feature size squared. In reality, empirical results have shown the DOF to decrease as about the feature size to the first power (due to improvements in photoresists and other factors). Today's 100-nm features typically have a DOF of a few tenths of a micrometer. An important requirement of any improvement in the practical resolution of an imaging system is the ability to live within the confines of this reduced DOF. Improvements in wafer and mask flatness, autofocusing and autoleveling systems, and wafer planarization by chemical-mechanical polishing of the wafer are examples of how the industry is coping with reduced DOF.

Attempts to simultaneously improve the resolution and DOF by optical means (sometimes called optical "tricks") include optimizing the mask pattern shape (called optical proximity correction, OPC), optimizing the angles of light illuminating the mask (called off-axis illumination, OAI), adding phase information to the mask in addition to intensity information (called phase shifting masks, PSM), and controlling the polarization of the illumination. Collectively, these optical approaches are known as resolution enhancement technologies (RET) [**Fig. 6**]. OPC predistorts the mask, taking into account the nonlinear nature of the imaging process, to make the final wafer pattern more like the design. OAI improves resolution and depth of focus for small dense features (though often at the expense of larger or more isolated features). PSM adds phase information to the mask, using the property that two beams of light that are 180° out of phase will cancel, creating darkness. Like OAI, PSM can improve resolution and depth of focus, but for a wider range of feature sizes and types. Often several or all of these RET approaches are used together.

The limits of optical lithography may be further pushed with immersion lithography. By replacing the air between the lens and the substrate with a higher index fluid, it is possible to design and build a lens with a numerical aperture greater than 1.0. In fact, numerical apertures up to 1.3 may be possible. These hyper NAs in conjunction with aggressive resolution enhancement technologies should enable 193-nm lithography to extend to resolutions of 45 nm and below.

Chris A. Mack

Bibliography. R. Dammel, *Diazonaphthoquinone-based Resists*, SPIE Tutorial Texts, vol. TT 11, Bellingham, WA, 1993; P. Rai-Choudhury (ed.), *Handbook of Microlithography, Micromachining, and Microfabrication*, vol. 1: *Microlithography*, pp. 597–680, SPIE Press, Bellingham, WA, 1997; J. R. Sheats and B. W. Smith (eds.), *Microlithography Science and Technology*, pp. 109–170, Marcel Dekker, New York, 1998.

Micromanipulation

The technique or practice involving manipulation of objects too small to be easily seen with the unaided eye. When a microscope is used to allow the operator to visually guide the microtools used in the manipulation, the technique is called micrurgy. When an object that is not extremely small needs to be positioned with extreme precision, as when aiming a laser to make a surgical lesion, the technique is called micropositioning.

To the unaided eye, the hand of a skilled person can appear completely steady and thereby seems able to reliably guide a fine-pointed probe onto the thinnest visible line. However, observed under a microscope, a probe held in even the steadiest of trained hands can be seen to drift about constantly within the width of this line, with a frequency around 8–12 Hz, and the rate of movement is magnified in the same proportion as the distance is magnified. Under high magnification, the hand appears capable of only large, fast movements and cannot guide a probe onto a line that is so thin as to be visible only with a microscope.

To manipulate microscopic objects, it is necessary to use a tool called a micromanipulator, which allows relatively coarse hand movements to execute proportionately slower and smaller movements of a probe or microtool. Micromanipulators are essential to much of modern technology. They are used

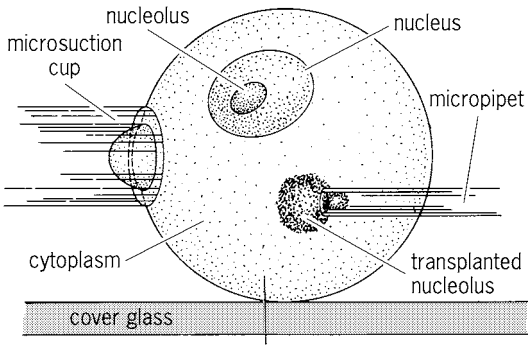

Fig. 1. Diagram of a cell undergoing transplantation of a nucleolus. The microsuction cup is used to hold the cell in place. The micropipet, positioned inside the cell cytoplasm with a micropositioner, holds the nucleolus. A microinjector is used to control both the microsuction cup and the micropipet.

to probe and test integrated circuit chips, to align fiber-optic communication cables, as well as to accomplish techniques in biological research (**Fig. 1**).

The simplest and most common micromanipulators consist of three orthogonal mechanical slides (**Fig. 2**), which can be dovetail or ball- or roller-bearing slides and whose position is controlled by a threaded screw or a rack and pinion. Such devices are relatively inexpensive and adequate for magnifications up to 150 power. However, just as the imperfections of the human hand are revealed by magnification, the imperfections of simple machines are revealed by magnifications greater than 150 power.

Imperfections of simple mechanical micromanipulators. A mechanical object that is pushed along a surface tends to move backward slightly when the force used to advance it is removed. This phenomenon, called backlash, poses a major problem in high-magnification micromanipulation. A probe just inside a cell or one that touches a tiny circuit element may move back out of contact when advance is stopped. To reduce backlash, the relatively coarse micromanipulators usually have one fine drive mounted on top, which is used to make the final advance in the probe direction. These drives have precision-machined and -fitted threaded drives (calibrated in micrometers), as well as a spring to maintain pressure of the slide against the drive. They are adequate to 250-power magnification; for greater precision a more sophisticated micromanipulator is required.

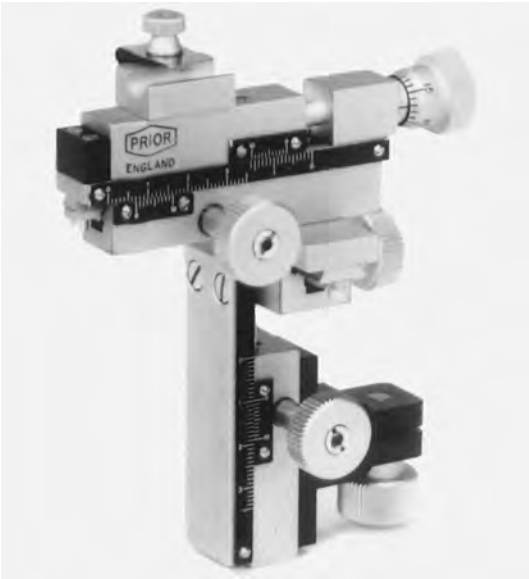

Fig. 2. Simple mechanical micromanipulator, with dovetail slides and rack-and-pinion drives. A fine micrometer drive is mounted on top.

A lag between the beginning of a hand movement and the beginning of a probe movement is called lost motion. Lost motion occurs when a threaded or rack-and-pinion drive is reversed in direction, since there is never a microscopically perfect fit between the nut and the threads, or the rack and pinion gear.

Stiction occurs because the coefficient of static friction is greater than the coefficient of sliding friction. For this reason, greater pressure is required to start a slide moving than to maintain constant motion or even to continue constant acceleration. Stiction results in a tendency to overshoot when a very small initial movement is desired.

A micromanipulator with three orthogonal axes can reach any point in a cubic range of movement, but it must be guided to it by successive approximation on each axis, one knob at a time, which is a slow, painstaking process.

Axial movements of the slide must be accompanied by a certain amount of random lateral deviation since no slide surface is perfectly flat and smooth. Dovetail slides are durable and resistant to damage, but the slide lubrication can attract dirt, which worsens lateral deviation at points on the slide. Ball and roller bearings have relatively less backlash and stiction and tend to force dirt out of their way, but they have less load capacity. Also, the slide track can easily be permanently marred at the point of contact by shock forces such as tipping over on the table. Marring can produce permanent lateral deviation.

Since temperature causes materials to expand or contract, there is inevitably some drift of position of the probe tip with even small temperature change. Drift is negligible in mechanical micromanipulators in a room environment, and when it occurs with other designs it is detrimental only in applications that require long-term holding of a position, as in recording cell intracellular potentials.

Normal movements of the hand that holds the drive knob (the initial reason for employing a micromanipulator) can still create some hand-generated vibration of the probe tip. Such movements can be reduced by adding mass to the micromanipulator, mounting the drive knobs near the base or solid attachment point, or by mechanical isolation of the control mechanism from the probe-carrying mechanism.

Types of micromanipulators. For each of the above imperfections, a micromanipulator has been designed that reduces or virtually eliminates that

limitation. However, a micromanipulator that is the best choice for all applications has yet to be designed.

High-magnification micromanipulators. Hundreds of different designs of micromanipulators have been published or have been commercially produced since the mid-nineteenth century, when such devices began to appear in the literature. Two or three interesting new designs are introduced each year, but only about 12–20 basic concepts are manufactured commercially.

One popular design, first described by T. H. Huxley, replaces the mechanical slides with flexure strips, that is, short ribbons of stiff spring steel that are flexible in only one dimension. These are made to flex by levers that have a 10:1 mechanical reduction and that are operated by three precision micrometers. Movement resolution of 0.1 micrometer is achieved, and position can be read from the micrometer scales. The nearly negligible backlash, stiction, and lost motion of the micrometer are all further reduced in a ratio of 10:1 by the levers. Lateral deviation is zero, as there are no slide surfaces, but the movement range is limited to a tenth of the range of the micrometer. A coarse micromanipulator is therefore usually mounted on top of this micromanipulator for quick initial positioning. Successive approximation is worsened, however, because the probe tip moves around an arc that follows the bend of the flexure strip. The substantial mass of a Huxley-type micromanipulator eliminates hand vibration but occupies a great deal of space.

Several applications are made significantly more efficient by single-lever joystick control of microtool movements, which eliminates the problem of successive approximation. The pneumatic designs of the Cailloux and DeFonbrune manipulators have such controllers mounted separately from the probe actuator, which is linked by pneumatic hoses. Remote controllers eliminate hand-generated vibration of the probe. Vibration of the hand on the joystick produces movement, but it is reduced in the same proportion as the intentional movements, so that they remain negligible as in normal macroscopic movements. The probe actuators are backlash-free in both pneumatic designs. The Cailloux design is unique in that the probe follows hand movement in three dimensions. In other joystick designs, the probe follows the hand in two dimensions while rotation of the stick or some other movement controls the vertical axis. When the load changes, as when attempting to penetrate a cell wall, the air in the cylinders compresses and then abruptly shoves the probe forward. Pneumatic micromanipulators are best for sorting cells or for other applications where the load is relatively constant.

The Leitz micromanipulator employs an eccentrically mounted sphere with a joystick to rotate the sphere in either direction, like a two-dimensional cam. The slides that move the probe have pushers that are held against the sphere surface by springs. The Leitz thus has direct mechanical linkage between the joystick and the probe and gives a very smooth positive movement, but it suffers from backlash. The designs of the Leitz and the DeFonbrune micromanipulators have a significant advantage in that the ratio of movement reduction is adjustable, allowing the same manipulator to have adequate range at low magnification and adequate resolving power at high magnification.

Another micromanipulator that features an adjustable ratio of movement reduction uses four sets of flexure strips in a square called a cloister stage. If one corner is anchored, the opposite corner can move only in a rectilinear, not curvilinear, fashion. The movement is controlled by a joystick. This rather large, complex, and expensive micromanipulator achieves near freedom from mechanical flaws in its movements.

Hydraulic micromanipulators are like the pneumatic instruments but employ an incompressible oil instead of air. Although they are capable of very smooth, forceful movements, they are most subject to drift of all micromanipulator designs, and the magnitude of drift is sufficient to inhibit their use for certain applications.

Electrically powered micromanipulators. There has been increasing use of electrically powered micromanipulators. The three types are direct-current-motorized, stepper-motorized, and piezoelectric.

Direct-current-motor-driven micromanipulators are relatively inexpensive compared to other sophisticated models, and they offer several advantages. They may be operated by remote control with a joystick or single-step push button, which allows for precise control of pulse length. It is easy to adjust the ratio of movement reduction electrically. Lost-motion correction may be built into the controller. They are also relatively light and space-efficient, but at highest magnification the slight vibration of the motors begins to interfere with visualization, making them best suited to medium magnifications.

Micromanipulators with stepper-motor drives provide the ultimate in precise small movements. These are the only micromanipulators capable of fast, direct movements to a specific set of coordinates, and can be directed by a computer over a preplanned route. Movements are precise enough to move around a surface and return to the original set of coordinates with a high degree of accuracy. Stepper-motor-drive micromanipulators can also be joystick- or button-controlled. However, stepper motors broadcast electronic noise when in operation and require large, expensive, high-voltage controllers. Thus, they are not easily adapted to a situation in which high-impedance amplifiers are employed, such as in recording cell- membrane potentials.

Specialized piezoelectric micromanipulators, called cell penetrators, provide movements over a very short range of 1–10 μm. They move very abruptly but are excellent for puncturing small and hard-to-penetrate cell membranes. However, these must be mounted on another form of micromanipulator to bring the membrane within range of the 1-μm punch. *See* ELECTRON MICROSCOPE; MICROSCOPE; OPTICAL MICROSCOPE.

Charles W. Scouten

Micrometeorite

A submillimeter extraterrestrial particle that has survived entry into the atmosphere without melting. Meteoroids are natural interplanetary objects that orbit the Sun, and they range in size from small dust grains to objects that are miles (kilometers) in diameter. Particles below 0.04 in. (1 mm) in diameter are considered micrometeoroids, and the micrometeoroids that enter the atmosphere without melting are called micrometeorites. Meteoroids of all sizes enter the atmosphere with velocities in excess of the Earth's escape velocity of 7.0 mi/s (11.2 km/s), and all but the smallest ones are heated sufficiently by air friction to produce at least partial melting. Micrometeorites survive entry without severe heating because they are small and they totally decelerate from cosmic velocity at high altitudes near 55 mi (90 km). In the thin air at such altitudes the power generated by frictional heating is low enough to be radiated away without a particle reaching its melting point, typically about 2400°F (1300°C) for common meteoritic samples. Larger objects penetrate deeper into the atmosphere before slowing down, and are melted and partially vaporized by friction with the comparatively dense air. Most of the mass of extraterrestrial matter that annually collides with the Earth is in the micrometeoroid size range, a total of about 10^4 tons (10^7 kg), but only a small fraction survives as micrometeorites. Usually only the particles smaller than 0.1 mm survive as true unmelted micrometeorites, although the survival of an individual micrometeorite depends on entry velocity, angle of entry, melting point, and density as well as size. The flux of micrometeorites falling onto the Earth's surface is approximately 1 per square meter per day (0.1 per square foot per day) for particles with diameters of at least 10 micrometers and approximately 1 per square meter per year (0.1 per square foot per year) for particles with diameters of at least 100 μm. *See* METEORITE.

Sources. Micrometeorites are of particular interest because they are samples of comets and asteroids, small primitive bodies that have survived without major change since the earliest history of the solar system. Some of these particles are generated by collisions in the asteroid belt, while others are released from comets when these bodies approach the Sun and ice volatilization releases dust grains and propels them into space. Once released from a parent comet or asteroid, particles survive only for a few thousand to a hundred thousand years, depending on size, before they are either destroyed or collide with a planet. Particles are destroyed either when they collide with other particles or when they spiral into the Sun because of the Poynting-Robertson drag, an effect of sunlight that causes the orbits of small particles to decay. During exposure in space, the small particles accumulate large amounts of helium implanted by the solar wind, and they also are riddled with radiation damage tracks produced by solar cosmic rays, high-energy particles accelerated from solar flares. *See* ASTEROID; COMET; COSMIC RAYS; SOLAR WIND; SUN.

Although it is certain that micrometeorites have both cometary and asteroidal sources, the relative importance of the two sources is not well understood. The dust released from comets can be seen directly as sunlight reflected from the comet tail, and in some cases infrared observations show paths of relatively large particles in the wake of the comet. Most meteors or shooting stars in the night sky are millimeter-sized pieces of comet dust entering the atmosphere at high velocity. Debris from asteroid collisions is observable in the infrared as bands that stretch completely across the sky. Dust from both comets and asteroids can also be seen with the naked eye as the zodiacal light, a glow preceding sunrise and following sunset that is caused by sunlight reflecting off micrometeoroids in interplanetary space. *See* INTERPLANETARY MATTER; METEOR; ZODIACAL LIGHT.

Collection. The collection and laboratory analysis of micrometeorites provide an important source of information on the nature of materials in comets and asteroids. This work complements research on larger conventional meteorites because the dust particles are probably a more representative sampling of early solar-system materials than are the conventional meteorites. Most and possibly all conventional meteorites are believed to be samples of a small number of asteroids from the inner regions of the asteroid belt. These samples are strongly biased because they have to be strong to survive atmospheric entry without fragmenting into dust and they are perturbed to orbits that intersect that of the Earth's only after rare gravitational perturbations, whereas dust particles do not have to be strong to survive atmospheric entry and light-pressure effects allow dust to diffuse through the solar system. Thus, collected micrometeorites include fragile cometary samples and samples of a diverse set of asteroids.

Most micrometeorites are collected in the stratosphere with aircraft such as the U2, which is capable of flying at an altitude of 12 mi (20 km) where terrestrial particles as large as 10 μm are rare. The spatial density of micrometeorites in the stratosphere is exceedingly low, but it is a million times larger than in space because of the low atmospheric fall speeds of micrometeorites relative to their original velocity in space. Micrometeorites are collected from the stratosphere by direct impact onto sticky plates that are extended from aircraft wings into the ambient airstream. With a collection surface area of 5 in.^{2} (30 cm^2), the collection rate is one 10-μm particle per hour of flight time. After a cumulative exposure of many hours, the plates are returned to a clean room where the microscopic particles are picked off with needles and placed onto mounts where they can be studied by electron microscopes (see **illus.**), mass spectrometers, and other instruments. Because of limitations imposed by flux and interference by terrestrial particles, the collection of micrometeorites in the stratosphere is usually limited to the size range from 2 to 100 μm in diameter. Most particles larger than this limit melt to form

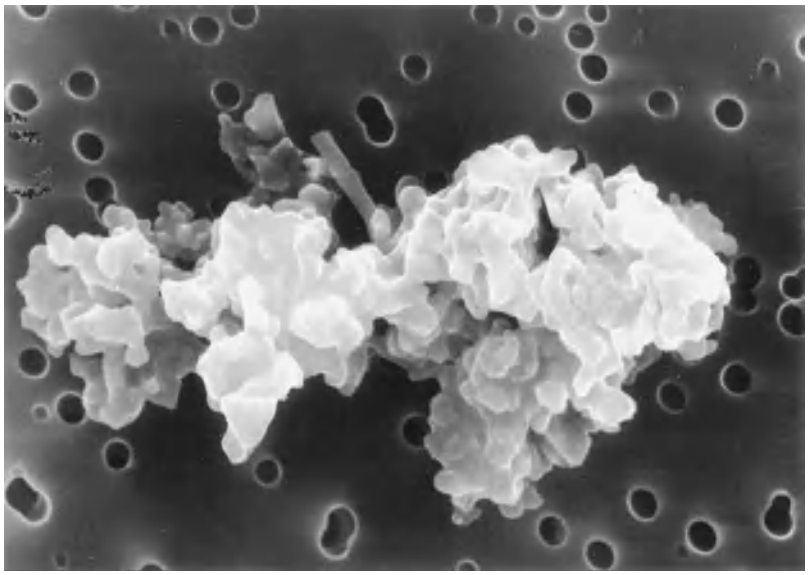

A 10-mm-long micrometeorite collected in the stratosphere with a U2 aircraft. The image was taken with a scanning electron microscope. The holes on the mounting substrate are unrelated to the particle.

cosmic spherules during atmospheric entry and are not true micrometeorites. A small fraction of particles up to 1 mm in diameter do manage to survive as giant micrometeorites. The flux of these larger particles is too low to collect in the air, but they can be collected from a few locations on the Earth's surface where they accumulate without being severely diluted with terrestrial particles. The best location on Earth for collecting particles of this type is ultrapure polar ice that forms in regions that are distant from rock outcrops or other sources of submillimeter particles. *See* COSMIC SPHERULES.

Properties. Typical micrometeorites are small black particulates. Some are composed of a few relatively large mineral grains, but most are aggregates of large numbers of submicrometer mineral grains, plus glass and carbonaceous matter. This latter group is often called chondritic micrometeorites because their elemental composition matches that of chondritic meteorites. The composition matches that of the Sun for condensable elements such as magnesium, iron, silicon, aluminum, sulfur, and sodium. Particles dominated by a small number of mineral grains have elemental compositions similar to that of the largest constituent grain. Most of the chondritic particles have similar elemental compositions, but they vary significantly in mineralogical composition. The two most common mineralogical groups are dominated respectively by hydrous minerals such as serpentine and smectite and anhydrous minerals such as olivine, pyroxene, and iron sulfide. Some of the hydrous particles are similar to carbonaceous chondrite meteorites and are thought to have asteroidal origins. They show evidence of moderate aqueous alteration, which is most likely to have occurred inside a moderately warm asteroidal parent body. Comets are smaller, cold bodies that are close to the Sun only for small fractions of their lifetimes. In comets, the damp and moderately warm internal environments required for aqueous alteration must be very rare. The anhydrous particles are often very porous, show no evidence of aqueous alteration, and are likely to have cometary origins. These particles are unlike any material found in conventional meteorites, and apparently they are samples of meteoroid types that are too fragile to survive atmospheric entry as bodies larger than the size of dust. These particles are being investigated with a broad range of laboratory instruments to determine elemental, chemical, mineralogical, and isotopic compositions and provide clues on the nature and origin of the materials that formed comets and asteroids in the early solar system. Detection of regions inside the particles approximately 1 μm in diameter that are highly enriched in deuterium provide evidence that micrometeorites may also contain records of interstellar materials that predate the origin of the Sun and planets. *See* COSMOCHEMISTRY; ELEMENTS, COSMIC ABUNDANCE OF; MINERAL; SOLAR SYSTEM.

Donald E. Brownlee

Bibliography. D. E. Brownlee, Cosmic dust: Collection and research, *Annu. Rev. Earth Planet. Sci.*, 13:147–173, 1985; J. F. Kerridge and M. S. Matthews (eds.), *Meteorites and the Early Solar System*, 1988; I. D. R. Mackinnon and F. J. M. Rietmeijer, Mineralogy of chondritic interplanetary dust particles, *Rev. Geophys.*, 25:1527–1553, 1987; S. A. Sandford, The collection of extraterrestrial dust particles, *Fundamentals of Cosmic Physics*, 12:1–73, 1987.

Micrometeorology

The study of small-scale meteorological processes associated with the interaction of the atmosphere and the Earth's surface. The lower boundary condition for the atmosphere and the upper boundary condition for the underlying soil or water are determined by interactions occurring in the lowest atmospheric layers. Momentum, heat, water vapor, various gases, and particulate matter are transported vertically by turbulence in the atmospheric boundary layer and thus establish the environment of plants and animals at the surface. These exchanges are important in supplying energy and water vapor to the atmosphere, which ultimately determine large-scale weather and climate patterns. Micrometeorology also includes the study of how air pollutants are diffused and transported within the boundary layer and the deposition of pollutants at the surface.

In many situations, atmospheric motions having time scales between 15 min and 1 h are quite weak. This represents a spectral gap that provides justification for distinguishing micrometeorology from other areas of meteorology. Micrometeorology studies phenomena with time scales shorter than the spectral gap (time scales less than 15 min to 1 h and horizontal length scales less than 2–10 km or 1–6 mi). Some phenomena studied by micrometeorology are dust devils, mirages, dew and frost formation, evaporation, and cloud streets. *See* AIR POLLUTION; ATMOSPHERE; MESOMETEOROLOGY.

Atmospheric stability. The behavior of the lowest part of the atmosphere depends strongly on stability. In an adiabatic process, there is no loss or gain of heat, no mixing of an air parcel with its surroundings, and no condensation or evaporation of water. An important adiabatic process involves expansion (or contraction) of air as the exterior air pressure decreases (or increases), accompanying rising (or descending) motion of an air parcel with temperature decrease (or increase) by 9.8°C for each kilometer (29.4°F for each mile) of ascent (or descent). A decrease of 9.8°C per kilometer is called the adiabatic lapse rate. Many atmospheric motions are approximately adiabatic. *See* ADIABATIC PROCESS.

For any layer in which the air temperature does not decrease as rapidly as the adiabatic lapse rate, an air parcel that is moved adiabatically upward or downward experiences a buoyant force that tends to cause the air to return to its original level. In this case, buoyancy tends to reduce turbulent mixing, and the atmosphere is said to be stable. An inversion is a very stable situation in which temperature actually increases with height. Whenever the temperature decreases more rapidly with height than the adiabatic lapse rate, any small adiabatic displacement of an air parcel results in a buoyant force that causes the air to move still farther away from its starting level. This is known as a superadiabatic lapse rate. This rate is unstable, meaning that the air will rapidly overturn within this layer. A layer in which temperature decreases at a rate equal to the adiabatic lapse rate is said to be neutral. An air parcel that is adiabatically moved upward or downward in a neutral layer will have no buoyant force acting on it. The neutral case is intermediate between the stable and unstable cases. *See* AIR TEMPERATURE; BUOYANCY.

Atmospheric turbulence. The airflow near the Earth's surface is nearly always turbulent. Friction causes the wind to be zero at the Earth's surface regardless of the wind speed higher up. The difference in wind speed near the ground and the air above is one manifestation of wind shear. The large wind shear near the ground is one of the main sources of energy for turbulent motions. Away from the surface, buoyancy increases the turbulent energy during unstable situations and decreases the turbulent energy during stable situations. Turbulent flows are highly chaotic, vary rapidly in time and space, and contain motions of many different sizes. Some appreciation for the complexities of turbulent motion can be gained by watching the rapid changes and the many sizes of motion of smoke flow. Even in carefully controlled laboratory studies, a turbulent flow is never exactly the same during any two repetitions of the same experiment. Because it is not possible to predict the exact details of turbulent flow, statistical methods are used. Wind, temperature, humidity, and other properties of the flow are written as an average part plus a fluctuation from the average. This separation, known as Reynolds decomposition, can be used to produce equations that govern turbulent flows. The rates at which quantities such as momentum and heat are transported by turbulent fluctuations are known as Reynolds fluxes or simply fluxes. *See* AERONAUTICAL METEOROLOGY; REYNOLDS NUMBER; TURBULENT FLOW.

Whenever the surface is warmer than the overlying air, an unstable, superadiabatic layer forms near the surface, and buoyant convection produces an upward heat flux that tends to mix the air vertically and enhances turbulent mixing. For surfaces that are cooler than the air, the surface layer is stable, and buoyant forces reduce the amount of vertical mixing. The degree of stability of the surface layer is quantified by the use of a dimensionless group known as the Richardson number. *See* CONVECTIVE INSTABILITY; DIMENSIONLESS GROUPS; DYNAMIC INSTABILITY; LAMINAR FLOW.

Plant canopy. When plants are growing on the surface, the layer from the ground to the plant top is known as the plant canopy. Within this layer, airflow is largely determined by the geometry of the plants. Wind speed varies with height within the plant canopy for several types of plants. The wind speed within the canopy is less than above the canopy due to the frictional drag of the plant leaves and stems or trunks and branches. The isolated conifer stand has a layer of maximum wind speed at about 0.1 times the height of the canopy. This type of forest has dense needles near the top of the canopy that produce large amounts of drag and slow the flow. The lowest part of the canopy is more open, mostly consisting of trunks that offer less resistance to the flow and allow greater wind speed. The balance between the pressure gradient force and the frictional drag on the trunks determines the flow in this lowest layer. The flow in this layer tends to be perpendicular to the isobars, whereas the flow above the canopy usually crosses the isobars at an angle less than 40°. *See* GEOSTROPHIC WIND; ISOBAR (METEOROLOGY).

The plant canopy absorbs solar radiation. This energy is used for photosynthesis in the leaves and heats the leaves and the surrounding air. Dense canopies shade the lower part of the canopy layer. Leaves of deciduous trees flutter in the wind, allowing sunlight to reach leaves lower in the canopy than would otherwise be possible.

Water vapor exchange between the plants and the air within the canopy is known as evapotranspiration. Rates of evapotranspiration depend on the photosynthetic and respiration rates of the plants and upon the amount of water within them. Much research is being conducted to determine and test techniques that use satellites for remote sensing of the state of surface plants, the evapotranspiration rate, and the surface heat flux. *See* BIOMETEOROLOGY; REMOTE SENSING.

The exchange of momentum, heat, water vapor, and carbon dioxide between the plant canopy and the overlying air is not steady, but occurs in short episodes. Gusts of air from above the canopy called incursions bring high-momentum air into the canopy. Slow-moving air from the canopy is ejected into the faster air above. These incursions and ejections account for 50–60% of the transfer but occur during only about 5% of the time.

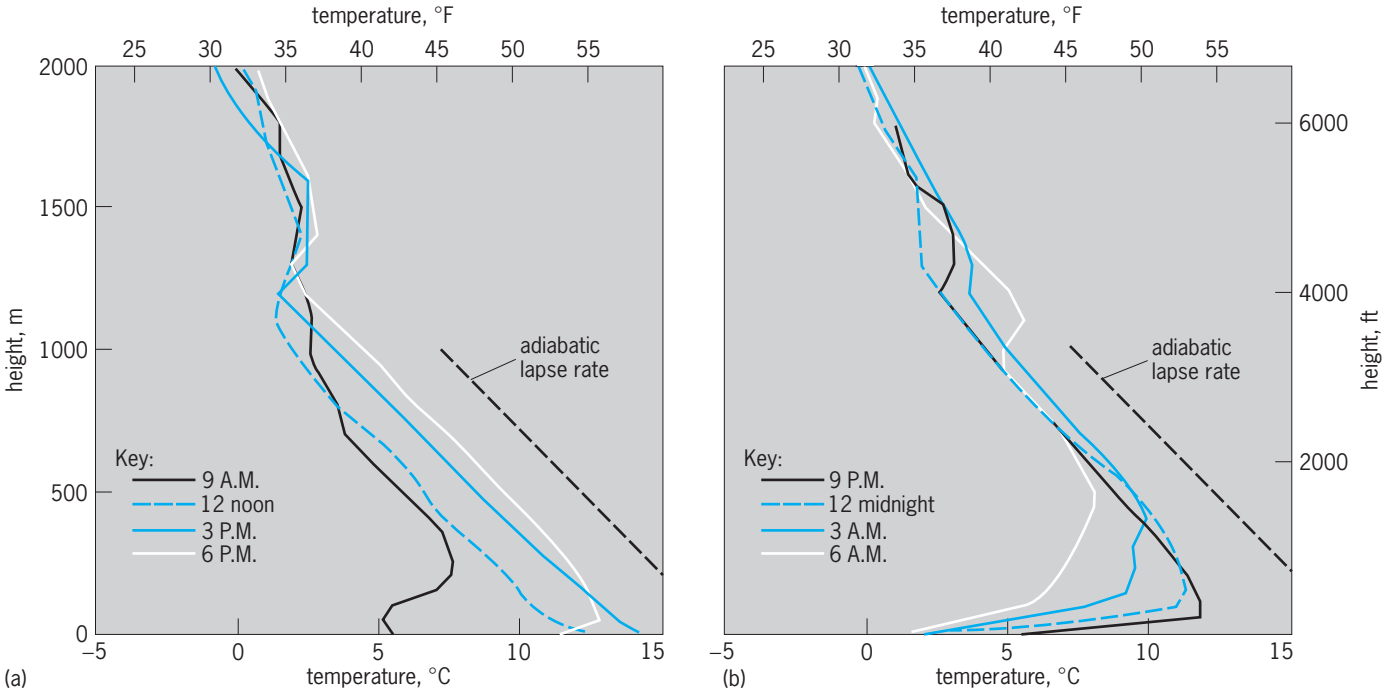

Fig. 1. Soundings of temperature versus height for (*a*) day time and (*b*) nighttime obtained by radiosonde measurements during the Wangara experiment near Hay, New South Wales, Australia, in 1967. Soundings are shown for 9 A.M. through 9 P.M. on day 33 of the experiment and 12 midnight through 6 A.M. on day 34. (*After R. B. Stull, An Introduction to Boundary Layer Meteorology, Kluwer Academic, 1988*)

Surface layer. The surface layer extends from above the canopy to a height of about 10 m (30 ft). Within this layer the transfer rates of momentum, heat, and water vapor are nearly equal to the values they have at the top of the canopy (the surface fluxes of momentum, heat, and vapor). Diabatic profile equations (Businger-Dyer equations) are used frequently to relate the surface fluxes to the changes of average wind speed, temperature, and humidity with height within the surface layer. When average wind speed, temperature, and humidity are measured at one or more heights, these equations can be used to compute surface flux values. *See* GAS DYNAMICS.

Boundary layer. The atmospheric boundary layer extends upward from the surface. It is the entire layer that is directly influenced by the presence of the Earth's surface. It includes the plant canopy and the surface layer and those higher layers in which turbulent mixing takes place. In stable situations the boundary layer may be only a few tens of meters thick. Under convective conditions it is typically 1 km (0.6 mi) or more thick.

The atmospheric boundary layer undergoes distinctive variations during the course of a day (diurnal cycle; **Fig. 1**). As the Sun warms the surface, convection transfers heat upward into the air, increasing the air temperature and mixing the air vertically. Warming and deepening of the boundary layer continues throughout the day as long as the radiation absorbed by the surface from the Sun exceeds the losses by long-wave radiation. In regions of high atmospheric pressure, the deepening of the boundary layer is retarded by subsidence.

A typical sounding of air temperature in the middle of the day may show an unstable superadiabatic layer near the surface (**Fig. 2**). In such a situation, when surface heating is strong, the temperature at the surface can be several degrees Celsius higher than the air above. This strong temperature gradient can cause mirages. In the middle of the mixed layer the rate of temperature change is close to the adiabatic lapse rate, indicating that rapid mixing is taking place. The boundary layer is capped by an inversion layer that limits the extent of vertical mixing and acts like a lid for the convection. During highly convective conditions (Fig. 2), substances are mixed rapidly throughout the boundary layer. These conditions produce

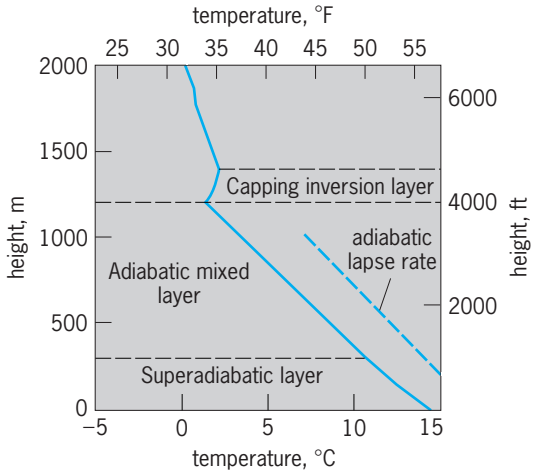

Fig. 2. The 3 P.M. sounding from Fig. 1*a* showing the layers of the daytime convective boundary layer.

rapid dispersion of atmospheric pollutants throughout the boundary layer. Prediction of the height of the inversion is one step in making air-quality forecasts. *See* CONVECTION (HEAT); MIRAGE.

Late in the afternoon, solar heating decreases, the surface heat source shuts off, and convection stops. During the night the surface cools because of infrared radiative heat loss, and a layer of cool air forms near the surface, giving a surface inversion (6 P.M. sounding of Fig. 1). During the night the air near the surface continues to cool (9 P.M. through 6 A.M. soundings of Fig. 1*b*). The surface inversion is stable and tends to suppress turbulence. When the wind speed is low, there is little wind shear and little turbulent mixing. If surface cooling is very strong, the surface layer may become so stable that the layer is no longer turbulent. On clear nights when radiative cooling is strong, the surface can reach low temperatures, and a thin layer of cold air can form near the surface. These conditions favor the formation of dew, frost, and possibly crop-damaging hard freezes. When wind speeds are greater, there is more turbulence in the surface layer, warm air is mixed downward toward the surface, and temperature decline in the surface layer during the night is less extreme. Crops can be protected from frost damage by smudge pots, orchard heaters, and fans. These are used to create turbulent mixing, which warms the air near the surface and thus mitigates crop damage. Spraying water on crops releases latent heat and can prevent frost damage for freezes lasting a few days or less. However, the weight of ice can cause damage if spraying is attempted for longer freezes. *See* AGRICULTURAL METEOROLOGY; DEW; FROST; WIND.

Surface energy budget. Complex interactions take place near the surface of the Earth over land as heat and moisture are exchanged between the soil, the plants, the air within the plant canopy, and the overlying air. During the day, solar radiation and downwelling infrared radiation are absorbed by the soil and by plants. The resulting heat energy is lost by emission of infrared radiation; or is conducted downward into the soil, increasing soil temperature; or provides the latent heat for evaporation of water from the soil into the air and evapotranspiration of water from plants into the air. The remainder of the heat energy causes an increase in the temperature of the plants, and the air within the plant canopy is carried upward by the turbulent heat flux to warm the air above. At night there is no absorption of solar energy, and infrared radiation usually produces a net loss of heat energy from the surface, accompanied by surface cooling, conduction of heat upward from the soil, and turbulent transport of heat downward from the atmosphere. Latent heat associated with evaporation and evapotranspiration often continues to cool the surface during the night; however, dew or frost formation, when present, releases heat and tends to slow the cooling of the surface.

Surface plant cover has several important effects on micrometeorology. The density of the plant canopy, its color, and the size and shape of leaves and stems affect the absorption of solar radiation and the emission of infrared radiation. The plants actively transport moisture upward from the soil. Plants open or close their stomata in an attempt to optimize their own well being. Plants need open stomata to obtain carbon dioxide for photosynthesis. When stomata are open, moisture readily escapes from plants. At night and during times when plants are water-stressed, they close their stomata to reduce water loss. The behavior of plants is also dependent on the stage in the plant life cycle. Modern computer programs to simulate surface micrometeorology contain many equations intended to model the effects of plants. *See* HEAT BALANCE, TERRESTRIAL ATMOSPHERIC; INSOLATION; TERRESTRIAL RADIATION.

Ocean-atmosphere interactions. Over the twothirds of the Earth that is covered by ocean, the boundary layer has a structure somewhat different from that occurring over land. When wind speeds are lower than 3 or 4 $m \cdot s^{-1}$ (7 or 9 mi/h), the water surface is aerodynamically smooth, and viscosity and diffusion are the main mechanisms transferring momentum and heat through the first few millimeters above the surface. The transfer of momentum from the air to the water results in generation of waves on the ocean surface. At greater wind speeds, these waves make the surface aerodynamically rough and increase the frictional drag on the air. The transfer of momentum from the air to the water also can produce ocean currents and oceanic coastal upwelling. *See* DIFFUSION; OCEAN CIRCULATION; UPWELLING; VISCOSITY.

The ocean has a much larger heat capacity than does soil, and therefore does not change temperature as rapidly as land does. As a result, the marine atmospheric boundary layer does not exhibit much diurnal variation, and it is seldom as stable or as unstable as the boundary layer over land. An important exception is that the marine boundary layer can be extremely unstable during outbreaks of cold air, when cold, dry air from over land passes over warm coastal waters. During these episodes, large amounts of heat and moisture enter the air from the ocean. Because of this transfer of heat and moisture, the Atlantic Ocean off eastern North America and the Pacific Ocean off eastern Asia are regions in which many storms are born and grow rapidly. *See* CYCLONE; MARITIME METEOROLOGY; WIND STRESS.

Other applications. Much of the early understanding of micrometeorology was obtained by studying conditions in large, flat, uniform areas that are relatively simple situations. Micrometeorologists have turned their attention to more complex situations that represent conditions over more of the Earth's surface. The micrometeorology of complex terrain, that is, hills and mountains, is important for air pollution in many towns and cities and for visibility in national parks and for locating wind generators. Another interest is the study of micrometeorology in areas of widely varied surface conditions. For instance, several different crops, dry unirrigated lands, lakes, and rivers may be located near one another. In these cases it is important to understand how the micrometeorology associated with each of these surfaces interacts to produce the overall heat and

moisture fluxes of the region so that these areas can be correctly included in weather and climate forecast computer programs. *See* CLIMATOLOGY; MOUNTAIN METEOROLOGY; WEATHER FORECASTING AND PREDICTION.

Research methods. Microscale meteorological features are too small to be observed by the standard national and international weather observing network. Generally, micrometeorological phenomena must be studied during specific experiments by using specially designed instruments. Instruments used to study turbulent fluxes must be able to respond to very rapid fluctuations. Special cup anemometers are made from very light materials, and high-quality bearings are used to minimize drag. Other anemometers use the speed of sound waves or measure the temperature of heated wires to measure wind. Tiny thermometers are used, so that time constants are short. Instruments are usually placed on towers or in aircraft, or are suspended in packages from tethered balloons. Instruments have been developed that can measure turbulence remotely. Wind speed and boundary-layer convection can be measured with Doppler radar, lidar devices using lasers, and sodar (sound detection and ranging) using sound waves. *See* ATMOSPHERIC ACOUSTICS; LIDAR; METEOROLOGICAL INSTRUMENTATION; METEOROLOGICAL RADAR.

Theoretical studies of atmospheric turbulence make extensive use of high-speed digital computers. Because atmospheric turbulent flows have high Reynolds numbers, they have a wide range of sizes and time scales of motion. Even with the fastest modern computers, detailed prediction of only 1 h of boundary-layer turbulence based on the exact equations governing flow would require a time equal to many times the age of the universe to complete. Faced with this obstacle, micrometeorologists have developed approximate techniques based on theoretical and empirical relationships. Higher-order closure techniques use equations that approximate some of the statistics of the flow in terms of simpler quantities and thus obtain a set of equations that can be solved in a reasonable period of time. These models have had good success at reproducing observed behavior of the atmospheric boundary layer.

Large-eddy simulation techniques compute the behavior of the largest features in the flow and use approximations for the smallest features. Even on powerful computers, these techniques require large amounts of time and are thus expensive to use, but they produce very accurate results and are therefore valuable research tools.

Despite the intricate nature of turbulent flow, the rich variety of phenomena involved in micrometeorology, and the detailed computer models available, scientists who make computer weather forecasts are forced to use greatly simplified equations for the boundary layer. These simulate the aspects of the boundary layer that most affect the large-scale weather patterns, including more detailed boundary-layer equations that would be too expensive and time-consuming for routine forecasts. *See* BOUNDARY-LAYER FLOW; CLIMATE MODELING; COMPUTER; METEOROLOGY; SUPERCOMPUTER.

Steven A. Stage

Bibliography. S. P. S. Arya, *Introduction to Micrometeorology*, 1988; D. A. de Vries and N. H. Afgan (eds.), *Heat and Mass Transfer in the Biosphere*, 1975; D. A. Haugen (ed.), *Workshop on Micrometeorology*, 1973; T. R. Oke, *Boundary Layer Climates*, 2d ed., 1988; R. B. Stull, *An Introduction to Boundary Layer Meteorology*, 1988.

Micrometer

A precision instrument used to measure small distances and angles. A common use is on a machinist's caliper, as in the **illustration**. *See* CALIPER.

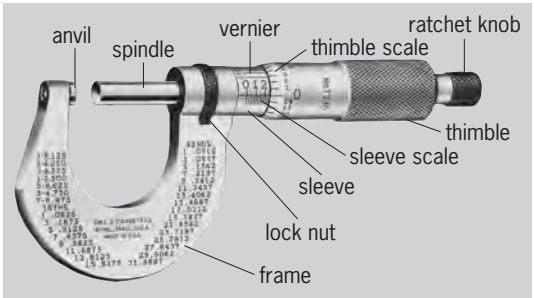

Machinist's outside caliper with micrometer reading 0.250 in. (*L. S. Starrett Co.*)

The spindle of the caliper is an accurately machined screw, which is rotated by the thimble or the ratchet knob until the object to be measured is in contact with both spindle and anvil. The ratchet slips after correct pressure is applied, ensuring consistent, accurate gaging. The number 1 on the sleeve represents 0.1 in.; the smallest divisions are 0.025 in. The thimble makes one complete turn for each 0.025 in. on the sleeve, and the 25 divisions on the thimble allow reading to the nearest 0.001 in. A vernier scale allows accurate reading to 0.0001 in. *See* VERNIER.

Frank H. Rockett

Micro-opto-electro-mechanical systems (MOEMS)

A class of microsystems that combine the functions of optical, mechanical, and electronic components in a single, very small package or assembly. MOEMS devices can vary in size from several micrometers to several millimeters. MOEMS may be thought of as an extension of micro-electro-mechanical systems (MEMS) technology by the provision of some optical functionality. This optical functionality may be in the form of moving optical surfaces such as mirrors or gratings, the integration of guided-wave optics into the device, or the incorporation of optical emitters or detectors into the system. The term may be confused with micro-opto-mechanical systems (MOMS), which more properly

refers to microsystems that do not include electronic functions at the microsystem location. MOEMS is a rapidly growing area of research and commercial development with great potential to impact daily life. The basic concept is the miniaturization of combined optical, mechanical, and electronic functions into an integrated assembly, or monolithically integrated substrate, through the use of micromachining processes derived from those used by the microelectronics industry. These processes, utilizing microlithography and various etch (subtractive) or deposition (additive) steps on a planar substrate, enable the production of extremely precise shapes, structures, and patterns in various materials. *See* INTEGRATED CIRCUITS; INTEGRATED OPTICS; MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS); MICRO-OPTO-MECHANICAL SYSTEMS (MOMS).

Capabilities. The microsystems realized by these techniques can have many unique capabilities. The miniaturization that is realized is useful in itself, allowing the systems to be utilized as sensors or actuators in environments that were not previously accessible, including inside living organisms, in hand-held instruments, or in small spacecraft. The miniaturization also allows for high-speed operation of the system, as the operating speed of mechanical systems is related to their inertial and frictional properties as well as the actuating forces. Optomechanical systems have been historically constrained in this area because of the mass required for stable optical elements and the extremely precise alignment requirements of most opto-mechanical systems, which limits the forces that can be tolerated for rapid motion. In the more integrated forms of MOEMS, the systems are prealigned by the precise fabrication processes, eliminating one of the more expensive aspects of assembling conventional optical systems. The miniaturization along with the scalability of microfabrication processes allows the development of massively parallel opto-mechanical systems, with millions of moving parts, that would not be possible in conventional technologies. MOEMS can incorporate detection and drive electronics in close proximity to provide improvements in signal-to-noise ratio for sensors and simplified interfaces for actuated systems. Ultimately, these electronics may be monolithically integrated in some technologies. Because of the production volumes achievable with micromachining techniques, MOEMS are potentially much less expensive than their conventional counterparts.

Sensors and actuators. Applications of MOEMS devices typically involve sensing of their environment, or the use of actuators to perform useful work. The sensing mechanisms can be any of those applicable from MEMS technology, with the added capability to utilize optical processes to enhance sensitivity or provide optical communication of data. The actuation mechanisms provided by MEMS technology are limited in many applications by the small forces that can be applied. MOEMS applications often involve the control of light energy, which is an ideal role for actuators that cannot deal with large forces, masses, or energies. The actuators can provide linear or rotary motion. MOEMS systems can also include microfluidics, in which liquid solutions can be transported through an optical detection system for analysis or other processes. *See* MICROSENSOR.

Fabrication. The fabrication processes used for MOEMS are the same as for other MEMS technologies. Most involve one of a few basic approaches, sometimes combined with unique materials or combinations of devices that utilize different processes. The fabrication technologies for MOEMS are based on the use of microlithography as a means of defining microscopic structures in large quantities and great precision. Microlithography produces a two-dimensional pattern on the surface of the substrate in a resist material that is coated on the substrate surface. The resist will protect some areas from subsequent process steps or act as a barrier that can be removed later as a sacrificial material. The microlithographic processes are then combined with various micromachining processes to define the mechanical structure. The micromachining processes are generally additive, in which material is deposited in sequence to build up a structure on a substrate, or subtractive, in which deposited films on the substrate or the substrate itself are removed to form the structure.

Surface micromachining in thin-film layers on silicon or other substrates is the most advanced process in terms of miniaturization, integration, and industrial support, as it uses the processes and tools of the silicon integrated circuit industry. The limitations of this process are due to the thin structures that are produced and the weak actuator forces that can be implemented by the thin-film structures. Surface micromachining with compound semiconductors (primarily from group III and group V, that is, columns 13 and 15 of the periodic table), such as gallium arsenide and indium phosphide, allows for light emission in the microsystem, in the form of semiconductor lasers or light-emitting diodes. These materials can also provide photodetectors in wavelength bands where silicon is not sensitive.

Bulk micromachining involves mostly subtractive processes that sculpt the silicon or other substrates to produce simpler, more robust structures that may incorporate thin-film membranes on the silicon surface as well. Variations of this process use deep reactive ion etching to form these structures with more versatility than can be attained by wet etch processes. These processes may be supplemented by wafer bonding to join structures together that have been processed on separate substrates originally.

Another process available is LIGA (a German acronym for *Lithographie, Galvanoformung, und Abformtechnik*), involving the combination of deep x-ray lithography with electroforming of metals and (sometimes) injection molding, to produce mechanical parts that are larger and more robust than the products of semiconductor fabrication processes. The disadvantage of LIGA is its lower level of integration, with more assembly required for functioning systems. Polymer replication processes, including casting, embossing, and injection molding, are

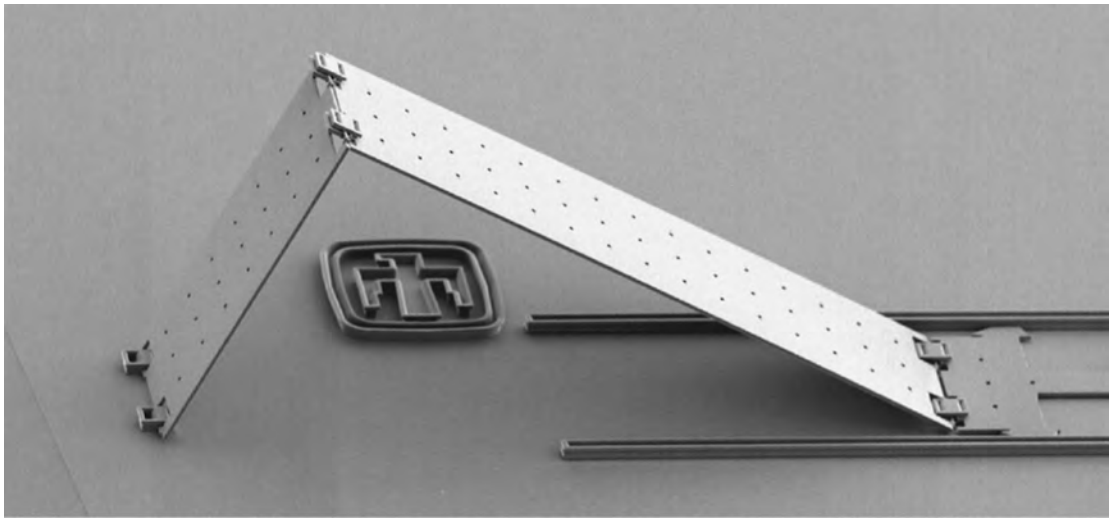

Scanning electron microscope image of a simple MOEMS tilting mirror, actuated by a MEMS rack-and-pinion drive that pushes the hinged mirror up from the right side. The structure is made by suface micromachining in polysilicon. The mirror is approximately 100 μm wide. (*Sandia National Laboratories*)

also applicable to some MOEMS fabrication needs. In all these cases, the use of micromachining, which can be performed in high volume, can produce extremely small and precise devices at low cost per device.

Integration and packaging problems. Some of the major problems that are still unsolved for MOEMS are integration difficulties and packaging challenges.

Although some MOEMS can be manufactured as a monolithically integrated device, in a single material system and process, the combination of different functions in a MOEMS sometimes calls for dissimilar materials and devices that cannot be integrated into a single material and processing line. In those cases, MOEMS are fabricated using hybrid assembly or heterogeneous integration techniques that allow the final system to include parts fabricated by different processes and from different materials to be precisely located on a common substrate and share electrical and optical signals. These assembly processes add expense and complexity in comparison to monolithically integrated systems.

The other major challenge to the development of MOEMS is the packaging of the devices. The mechanical actuators are often very sensitive to contamination and must be protected from the environment. The sealed packages that are required must usually also allow for optical access into and out of the package. In addition, optical systems are often complex three-dimensional structures and MOEMS are usually two-dimensional planar structures. The package is often required to incorporate and hold in position additional optical elements in the system, including optical fibers, lenses, and detectors. This requirement can add great complexity to the packaging. *See* ELECTRONIC PACKAGING.

Applications. The most sophisticated MOEMS device in commercial use is the Digital Micromirror Device (DMD), manufactured by Texas Instruments. This device consists of 2 million individual tilting micromirrors (see **illus.**), each 16 μm × 16 μm in size. Each mirror is attached to a torsion beam suspension that allows it to tilt in one axis. The entire array of mirrors is fabricated by postprocessing of a static random access memory chip that is fabricated in conventional silicon complementary metal-oxide-semiconductor (CMOS) technology. Surface micromachining techniques are used to deposit metal films and sacrificial polymer layers to build the mirrors on top of the memory chip. The cells of the memory chip hold electrical charges, depending on the data stored, that actuate the mirrors above by electrostatic forces. This system is an excellent example of a scalable process and architecture and of monolithic integration. The DMD is used primarily in video projection systems. *See* OPTICAL PROJECTION SYSTEMS.

A new area of application of MOEMS that may have great impact on everyday life is in optical-fiber switching systems. As the need for high-bandwidth data transmission grows with increasing use of the Internet as a means of communication and commerce, fiber-optic communications are increasingly used and will eventually link to homes and offices directly. A necessary component of these optical-fiber networks is a switching system that can independently switch each of the channels carried in the fiber. Conventional opto-mechanical devices are large and slow, and other nonmechanical approaches have not been successful. The MOEMS optical switch, scalable to hundreds or thousands of channels, is under rapid development by many companies and may enable the development of direct access to very high bandwidth communications. *See* OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

Mial E. Warren

Bibliography. M. J. Madou, *Fundamentals of Microfabrication*, CRC Press, Boca Raton, FL, 1997; M. E. Motamedi and R. Goering (eds.), Miniaturized Systems with Micro-Optics and MEMS, *Proc. SPIE*, vol. 3878, 1999; *Proceedings of the 3d International Conference on Micro Opto Electro Mechanical Systems: MOEMS '99*, Mainz, 1999; J. M. Younse,

Mirrors on a chip, *IEEE Spectrum*, pp. 27–31, November 1993.

Micro-opto-mechanical systems (MOMS)

Miniaturized opto-mechanical devices or assemblies that are typically formed using micromachining techniques that borrow heavily from the microelectronics industry. The term may be used to distinguish devices and microsystems that combine optical and mechanical functions without the use of internal electronic devices or signals. Systems that use electronic devices as part of the microsystem may be referred to as MOEMS (micro-opto-electro-mechanical systems). In some cases, these terms may be used synonymously. A related area is MEMS (micro-electro-mechanical systems), in which electronic and mechanical functions are combined in a miniature device or system, but not necessarily implementing optical functions. The progress of MOMS technology has been greatly enabled by the simultaneous development of microelectronics and optical fiber-based telecommunications technology. These two technologies have resulted in a large manufacturing and scientific base from which MOMS has been developed. The commonly used MOMS fabrication processes are the same as are used for variations of the MEMS fabrication processes with some additional materials choices, due to the need to be compatible with optical functions. *See* INTEGRATED CIRCUITS; MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS); MICRO-OPTO-ELECTRO-MECHANICAL SYSTEMS (MOEMS); OPTICAL COMMUNICATIONS.

Advantages. Although similar in concept to MOEMS technologies, MOMS has unique advantages for some applications. The use of only optical energy and signals gives MOMS an inherent immunity to electromagnetic interference (EMI) that is important for applications in electrically noisy or high-voltage environments. The absence of semiconductor electronic devices greatly increases the high-temperature tolerance of the system. MOMS devices can be designed to work immersed in liquids, which is of great importance for chemical sensing and biomedical applications. The fact that the power and signal sources can be remotely provided via an optical fiber, allowing the sensor to be passive, is of great utility and reduces the impact of a MOMS sensor on its local environment. MOMS can be used safely in flammable and explosive environments, making them uniquely valuable in the petrochemical industry.

Sensors and actuators. A micromechanical system is able to function either as a sensor or as an actuator. In sensor design, the goal is to measure a physical parameter with minimum impact on that value from the operation of the sensor and with low sensitivity to other parameters that may be changing. An actuator performs some mechanical work that will either change the local environment or control or enable energy or information to be transported or controlled.

MOMS sensors. MOMS applications have been primarily in sensors, utilizing a number of physical phenomena. MOMS sensors consist of some movable mechanical element that can be detected or measured by optical means. The motion may be coupled to temperature, electrical and magnetic fields, acoustical energy, acceleration, chemical forces on surfaces, and so forth. Many sensors can be made sensitive to various physical or chemical phenomena by addition of special surfaces or coatings to the mechanical element.

The state or position of the mechanical structure can be detected optically by various means. Many sensors use intensity modulation of a light beam. This may be by mechanical interruption of a beam path or by more subtle means, such as defocus or deflection of light to be coupled into an optical fiber. Many sensors use phase modulation as the sensing means. This requires the use of an interferometer to convert phase into a detectable amplitude modulation. The interferometer may be one of several types. Fabry-Perot devices can utilize free-space propagation of beams normal to the mechanical surfaces to detect motion or position of the surfaces. Mach-Zehnder and other designs that require splitting and recombination of light beams are often implemented as guided-wave devices, in which the light is confined in planar thin-film structures. Interferometric detection can be very sensitive, but places additional requirements on the light source and detection electronics. A third mechanism is wavelength, in which a broad spectral source can be spectrally dispersed or filtered. Polarization is also an optical parameter that can be used in some cases. *See* DISPERSION (RADIATION); INTERFEROMETRY; POLARIZATION OF WAVES.

Since many MOMS sensor applications use optical fibers as the sources of power and means of data retrieval, MOMS sensors are often discussed in the context of optical-fiber sensors. MOMS can be distinguished from the more mature field of fiber-optic sensing by their use of miniaturized or integrated mechanical structures, formed by micromachining techniques, as the sensor. A MOMS sensor is extrinsic to the optical fiber, as opposed to intrinsic optical-fiber sensors, in which the fiber is the sensing device. *See* FIBER-OPTIC SENSOR; MICROSENSOR; OPTICAL FIBERS.

MOMS actuators. MOMS actuator technology is much less developed. The actuation mechanisms that are available to allow light to do mechanical work are few. Photothermal mechanisms allow the conversion of light to heat in order to power a thermomechanical actuator. Examples of this process include expansion of a working fluid or deforming a bimetallic strip through differential thermal expansion. Photoexcitation of semiconductor structures can be used to generate electrostatic and surface charge effects or cause photostriction in piezoelectric materials to impart mechanical force. Light can also produce direct radiation pressure, but the forces are small relative to the light intensity. A system can be powered by light through photovoltaic conversion, but then would

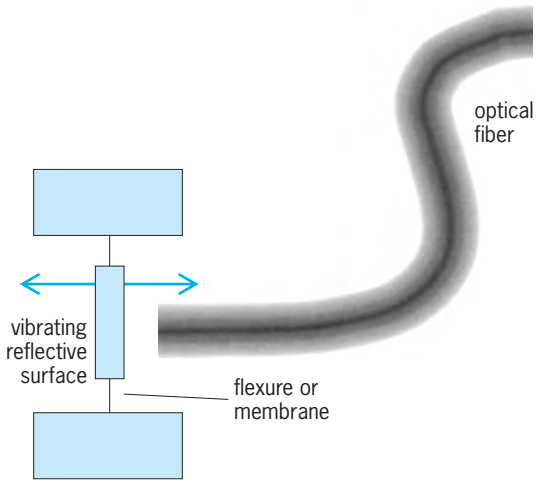

Fig. 1. A simplified MOMS sensor using a reflecting surface on a flexible mount or membrane and the end face of an optical fiber to form an optical interferometer that can sense vibration. The vibration of the flexible membrane allows the reflecting surface to move, changing the resonance wavelength of the interferometer, modulating the intensity of the light that is reflected back into the interferometer. At the other end of the interferometer is a light source and a detector.

not be considered purely opto-mechanical and classified as a MOEMS device. *See* PHOTOVOLTAIC EFFECT; PIEZOELECTRICITY; RADIATION PRESSURE.

Applications. Some examples of MOMS technology include optical pressure transducers—microphones or hydrophones that have a thin mechanical membrane that is one surface in a Fabry-Perot interferometer formed by the reflection from the membrane surface and the reflection from the end of the fiber. (A similar arrangement for sensing vibration is shown in **Fig. 1**.) Other versions have a planar optical waveguide on the surface of a sensitive membrane that is one arm of a two-beam Mach-Zehnder interferometer. Another example is an accelerometer in which a small mass is suspended from flexure attachments to the substrate. Optical fibers are positioned with a small gap in which the moving mass can interrupt the transfer of light from one fiber to another to modulate the light intensity transmitted through the fibers. One of the most well developed MOMS applications is optical sensing of the position of small cantilevers used in scanning tip microscopy processes such as atomic force microscopy. In some versions of these extremely sensitive instruments, a micromachined cantilever or balance mechanism is subject to forces from a sharpened tip on the end of the lever that interacts with the interatomic forces of a surface. The forces bend or tilt the mechanism with displacements of fractions of a nanometer. The minute bending or tilt can be measured optically by angular change of reflected light from the surface of the cantilever beam without perturbing its position (**Fig. 2**). *See* ACCELEROMETER; MICROPHONE; PRESSURE TRANSDUCER; SCANNING TUNNELING MICROSCOPE.

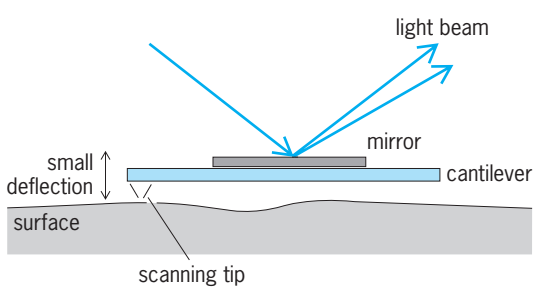

Fig. 2. A simplified representation of how a light beam is used to sense minute deflections of a micromachined cantilever in an atomic force microscope. The sharp tip is scanned very close to a surface. The interatomic forces on the tip deflect the cantilever. The light beam reflected off the top of the cantilever amplifies the motion. This MOMS device can resolve individual atoms on the surface.

Prospects. Micro-opto-mechanical systems are still in their infancy. As optical fiber communications becomes more pervasive, use of MOMS can be expected to increase as a means of optical sensing and in some cases actuation, for situations in which it provides unique value. Some challenges have to be addressed to further improve the performance, cost, and utility of MOMS. Integrating optical and micromechanical functions, which often require different materials and processes, is difficult and limits the sophistication and complexity that can be achieved with MOMS technology. The packaging of MOMS requires input and output coupling of optical signals into a package that often needs to be in intimate contact with its environment for the sensor to function, while still protecting the system. *See* ELECTRONIC PACKAGING.

Mial E. Warren

Bibliography. A. J. Jacobs-Cook, MEMS versus MOMS from a systems point of view, *J. Micromech. Microeng.*, 6:148–156, 1996; M. J. Madou, *Fundamentals of Microfabrication*, CRC Press, Boca Raton, FL, 1997; G. Meyer and N. M. Amer, Novel optical approach to atomic force microscopy, *Appl. Phys. Lett.*, 53:1045–1047, 1988; A. Wang (ed.), Harsh Environment Sensors II, *Proc. SPIE*, vol. 3852, 1999.

Micropaleontology

A branch of paleontology dealing with the fossilized microscopic organic remains (microfossils) of the geologic past, their structure, biology, phylogenetic relations, and distribution in space and time. The study of these microfossils has become an independent scientific field largely because of the following: (1) The size of these fossils requires special methods for collection and examination. (2) Their abundance in geologic formations makes it possible to analyze their spatial distribution and the rates of morphological changes during the course of evolution by means of statistical methods which can be used only under exceptional circumstances in the study of larger fossils. (3) Microfossils have become indispensable tools in certain branches of applied geology, especially in the exploration for oil-bearing strata, because countless numbers of these minute fossils may be obtained from small pieces of subsurface rock recovered from drill holes. (4) The diversity of microfossils, their

wide spatial distribution in varied environments, and their distinctive steps in evolution and the ease of studying them have contributed to make micropaleontology one of the most actively studied branches of the earth sciences. Many organisms are restricted to certain environments best suited to their life activities. The occurrence of certain organisms in sediments, therefore, can in turn provide a clue to a particular environment that existed in the past. Micropaleontologists can obtain important information about depositional environments, such as their ancient environment, water depth, temperature, current systems, water mass distribution, proximity of shore lines and, of course, these organisms are used to date ages of enclosing sediments.

The material subjected to micropaleontological studies forms a spectrum from primitive plants to advanced vertebrates. The only prerequisite for organisms to become the subject of micropaleontological studies is their possession of resistant skeletal components ensuring their preservation in sedimentary strata as fossilized remains even after biological, chemical, or mechanical processes have destroyed the organisms' soft parts.

Most major groups of organisms incorporate, besides organic compounds, hard resistant materials that serve for structural support or protection. The more common substances found among the microfossils are calcium carbonate, silicon dioxide (or silica), calcium phosphate in the form of the mineral apatite (typical of bones and teeth), sporonine (principal constituent of pollen and spore walls), and various complex organic compounds.

Micropaleontological objects representing the plant kingdom (**Fig. 1**) range from primitive unicellular forms such as calcareous nannoplankton, diatoms, charophytes, silicoflagellates, chrysomonadines, dinoflagellates and their cysts, tintinnids, the enigmatic chitinizoans, to portions of advanced plants such as pollen, spores, and seeds. *See* FOSSIL SEEDS AND FRUITS; PALYNOLOGY.

Micropaleontological objects representing the animal kingdom (**Fig. 2**) include almost entire organisms of the unicellular Radiolaria and foraminiferans; carapaces or shells of the minute crustacean Ostracoda; embryonic or immature shells of larger organisms such as bryozoans, corals, or mollusks; individual skeletal elements of advanced organisms, including sponges, corals, sea cucumbers, echinoderms, worms, and fishes; and toothlike skeletal elements of animals of unknown taxonomic affinity known as conodonts, which have been assigned variously to the worms, gastropods, cephalopods, and fishes.

Occurrence and abundance. Microfossils occur in most types of marine and nonmarine sediments and sedimentary rocks, particularly in limy shales, limestones, black shales, siltstones, and fine-grained sandstones. They are usually disseminated through the rock in quantities ranging from a few to several hundred fossils per gram of rock. Certain rock types, however, are composed almost entirely of microfossils. Diatomite is a siliceous earth made up of the bivalved frustules of diatoms; chalk is a soft variety of limestone composed of the calcareous tests of foraminiferans and coccoliths; certain limestones of the Pennsylvanian and Permian systems are essentially composed of fusulinid Foraminiferida; and the Gizeh limestone of the Eocene of Egypt, with which the pyramids were built, is made up almost entirely of the shells of the large foraminiferid *Nummulites*. Many reef limestones in the sedimentary column are built by lime-secreting algae, corals, bryozoans, and stromatoporids. The dense siliceous rock known as chert, in many instances, was formed from accumulations of the shells of diatoms and radiolarians, while plant spores and pollen are believed to be the essential particles composing the rock known as cannel coal. Deep-sea environments harbor special types of sediments entirely made up of shells of microorganisms. These include *Globigerina* ooze, consisting largely of calcareous nannofossils and tests of planktonic foraminifers, and siliceous ooze, of which chief constituents are radiolarians and diatoms. *Globigerina* ooze covers nearly one-half the area of the entire world ocean floor, and siliceous ooze blankets nearly 15%.

Collection and preservation. Surface exposures of sedimentary rock are sampled for micropaleontological analysis by collecting about a pint (a few hundred cubic centimeters) of fresh unweathered rock at the outcrop. Choice of vertical sampling interval depends upon the total thickness of rock exposed, the degree of change in the sedimentary characteristics within the vertical section, and the rate of accumulation of sediments.

Selection of samples containing microfossils larger than 0.02 in. (0.5 mm) may be greatly facilitated by inspecting the sample with a 10× or 15× hand lens. The presence of smaller microfossils cannot be detected in the field without the use of a microscope.

Subsurface samples of fossiliferous rocks are obtained from drill cuttings and from cores obtained in the course of drilling wells for oil, gas, or water. They consist of small chips or rock fragments washed from the borehole during drilling operations, or of cylindrical cores of the rock strata obtained by the use of a special drilling bit designed for such a purpose.

Laboratory treatment to free the microfossils from the enclosing rock matrix varies greatly, depending upon the hardness, texture, and mineral composition of the rock and upon the size and chemical composition of the microfossils. Since most microfossils are quite fragile, the sample is given a minimum of mechanical and chemical treatment to reduce the rock to individual particles and to liberate the microfossils.

Soft shales, siltstones, and chalks are usually soaked and gently boiled in water containing a deflocculant such as sodium hexametaphosphate. This treatment reduces the sample to a thick mud, which is flushed through a series of graduated screen sieves to remove the finest particles. The various-sized fractions are dried and placed in envelopes or vials to await inspection with the microscope. Tougher and more resistant rocks are broken down by first heating

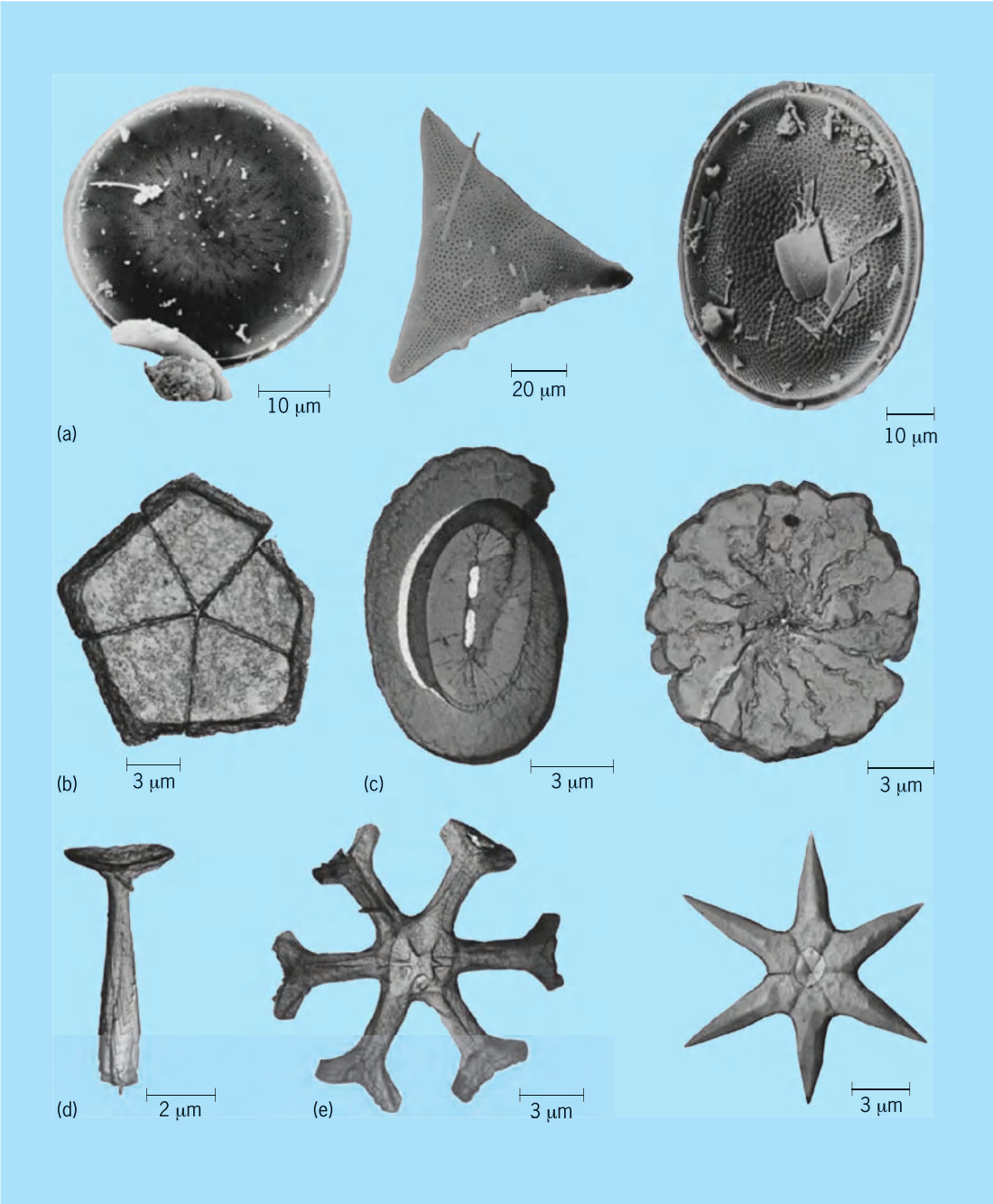

Fig. 1. Representative microfossils of the plant kingdom. (*a*) Diatoms. (*b*–*e*) Examples of calcareous nannoplankton: (*b*) pentalith, (*c*) placoliths, (*d*) rhabdolith, and (*e*) asteroliths.

samples (crushed to a chestnut size) in an oven at a temperature of more than 212°F (100°C) and then soaking them with either a saturated solution of sodium sulfate or kerosine extract. Samples treated with sodium sulfate should be left at room temperature for a few days until the chemical crystallizes, breaking up the rocks. The kerosine-soaked rocks also give a better result when they are left at room temperature for a few days and then boiled in water with a deflocculant.

Limestones and other carbonate rocks are digested with dilute hydrochloric acid or glacial acetic acid to dissolve the carbonate minerals, leaving a residue of insoluble mineral grains and such noncalcareous microfossils as siliceous sponge spicules, conodonts, diatoms, radiolarians, and plant microfossils.

Disaggregation of coal and mudstone to liberate spores, pollen, and other very small microfossils requires special treatment. This includes digestion of the rock sample with hydrofluoric acid and Schulze's solution (a mixture of nitric acid and potassium chlorate), concentration of the fine-grained residues by centrifuging, suspension of smears of the residues in glycerin jelly, selective staining, and mounting of the stained material on glass microslides for microscopic examination. Both the Schulze's solution and

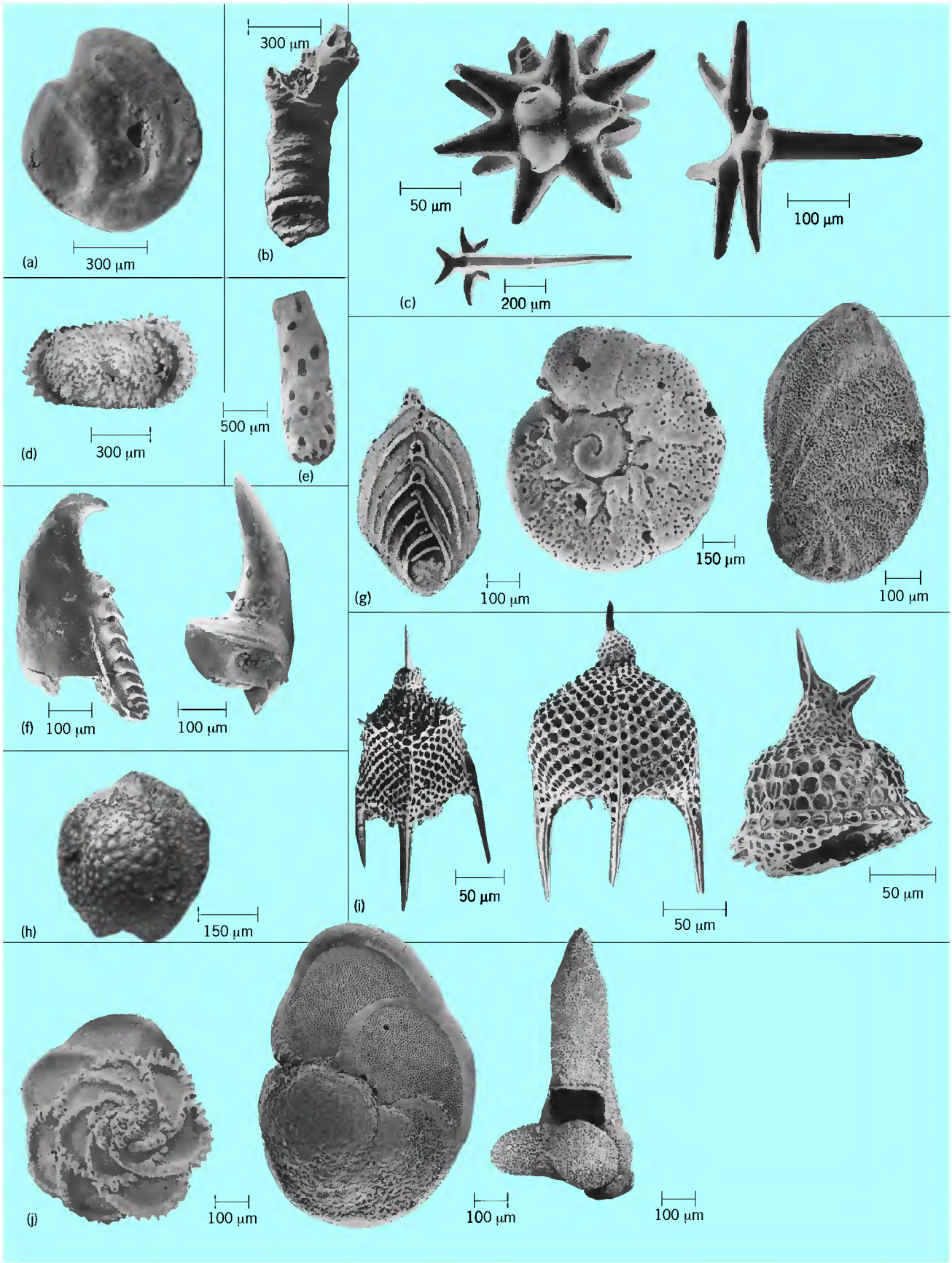

Fig. 2. Microfossils of the animal kingdom. (*a*) Vertebrate remains: fish ear bone. (*b*) Alcyonarian coral. (*c*) Sponge spicules. (*d*) Ostracod. (*e*) Bryozoans. (*f*) Scolecodonts. (*g*) Smaller benthonic foraminiferans. (*h*) Larger benthonic foraminiferans. (*i*) Radiolarians. (*j*) Planktonic foraminiferans.

hydrofluoric acid are highly toxic and corrosive, and extra care must be used when these chemicals are handled.

Microfossils larger than 0.04 in. (0.5 mm) are separated from the dried and size-sorted residues by scattering the grains on a small black tray and inspecting them on the stage of a wide-field stereoscopic binocular microscope at magnifications of 10 or 15 diameters. The individual fossils are selected from the residue by picking them up on the moistened and pointed tip of a small red sable watercolor brush (size 000); transferred to a special micropaleontological slide made of cardboard, which has a holed-out and blacked-in mounting area; and coated with a water-soluble adhesive such as gum tragacanth.

Smaller microfossils are studied in smears or suspensions of fine residues mounted on glass micro slides and protected with cover slips. Depending on the type of study, both the light microscope at magnifications of 100–1500 and the electron microscope at magnifications of 600–25,000 are used to examine microfossils.

Certain large fossils, such as bryozoans, lime-secreting algae, colonial corals, stromatoporids, and large Foraminiferida, are prepared for examination by grinding down an oriented fragment of the large fossil to a thin slice through which light can be transmitted after it is mounted on a glass microslide with a cover slip. Such sections are studied at magnifications of 30–500.

Applied micropaleontology. In the historical development of micropaleontology, researches on various microfossils have progressed at a markedly different pace, reflecting the interests of micropaleontologists working at various periods of time.

Beginning around the 1930s, great impetus was given to the study of Foraminiferida, Ostracoda, and pollen and spores. Micropaleontologists employed by oil companies have found these minute fossils, which could be easily secured from well cuttings and cores, often in great abundance, useful for correlating oil-bearing strata.

Research interest in these areas surged to such an extent during the 1940s that micropaleontologists found it necessary to create ever finer subdivisions of micropaleontology, each of which would be devoted to one particular group of microfossils. For instance, the term palynology, defined as the study of both modern and fossil pollen and other spores and their dispersal, and applications thereof, was introduced in 1944. Probably less well known, a few other terms such as foraminiferology and ostracodology also came into existence to represent the field of research exclusively devoted to a certain group of microorganisms.

The great expansion of oceanographic research during the 1960s gave another advantageous turn in the course of micropaleontological research, inviting earnest investigations of minute plants and animals inhabiting the surface waters of oceans. In contrast to the sessile organisms (termed the benthos) whose habitats are ocean floors, these organisms maintain the state of suspension throughout their life-span and are known collectively as plankton. The free suspensive mode of life of plankton in oceanic waters favors their wide geographic distribution by ocean currents, and fossilized representatives of various planktonic microorganisms have become very important tools in establishing mutual time relationships of marine strata in distant parts of the world.

In contrast to the fast-changing continental environments which were subjected to such forces as advancing glaciers, orogenic upheavals, and inundation by the sea, earth scientists would generally agree that the deep-sea environment is one of the most stable on the Earth. For millions of years, an incessant "snowfall" of tiny particles has been settling over the ocean floor. In deep ocean basins, far removed from the continents, the most significant contributors of these particles are the innumerable planktonic microorganisms, which pass their life in the sunlit surface waters, building up their skeletons or shells from the silica or calcium carbonate present in the water. After the plankton dies, these hard parts, freed from organic matter through bacterial action, sink to the bottom and build up the unique deep-sea sediments known as organic oozes. The oozes, rich in calcium carbonate and largely consisting of the shells of calcareous nannoplankton and foraminiferans, dominate in lower latitudes, while organisms such a diatoms and radiolarians form layers of siliceous oozes in the far north and south, especially around the icebound Antarctic continent.

At any particular location on the deep-ocean floor, if the successively accumulated layers of the organic oozes for millions of years have not been disturbed by slumping or denudation by bottom currents, a core of sediment plugged out of the sea bottom would reveal the succession of floras and faunas that inhabited the oceans of the geologic past. No plant or animal has existed for all geologic time. Each species evolved from some ancestor with a definite time of beginning and, if it is extinct today, also a time of ending. Each extinct species therefore may provide three distinct time levels which can be used for comparing time relationship of marine sediments: (1) the time before it evolved; (2) the time during which it existed; and (3) the time since it became extinct. The evolution of organisms through time thus provides a means by which discrete units of time represented by the material accumulation of sediments can be recognized. Dispersal by ocean currents of these planktonic organisms to distant parts of the world further enables a comparison of the established time units between regions, particularly with the marine sections in Europe on which the presently recognized geologic systems were defined. The long, continuous record of biota in the deep-sea basins was intensively studied in the 1960s because the marine record preserved in the continents is frequently punctuated by breaks in sedimentation.

Planktonic biota in deep-sea basins also play a major role in understanding oceanographic conditions which existed in the geologic past. When marine planktons build their shells from minerals dissolved in the water, they incorporate within their

shells oxygen and carbon isotopes in the proportion which existed at the time of their growth. Proportions of various oxygen and carbon isotopes in oceanic waters are controlled by physicochemical parameters of the water, so that by measuring the isotopic compositions incorporated in microfossils, scientists, in turn, can estimate such parameters as temperature and salinity of past oceans.

When magnetic iron oxide minerals slowly settle out of a fluid, they align their magnetic polarity with any magnetic field that existed during the past. This produces a preferential magnetization in sediments that contain the minerals; this property, which is measurable in the laboratory, is termed remnant magnetism. Study of remnant magnetism in rocks of different ages all over the world indicates that the Earth's North and South poles have abruptly switched positions many times in the geologic past. Since each of these paleomagnetic reversals simultaneously affected the field for the entire Earth, reversal records identified in sediments can be compared with the similar record in continental rocks which can be dated precisely by radiometric age determinations. This permits micropaleontologists to compare rates of microfossil evolution against actual time lines for the first time.

Study of deep-sea sediments represents a unique but active field of research, where micropaleontologists, geochemists, and geophysicists can cooperate in elucidating the interaction of environmental changes with the course of organic evolution. *See* INDEX FOSSIL.

Tsunemasa Saito

Microphone

An electroacoustic device containing a transducer which is actuated by sound waves and delivers electric signals proportional to the sound pressure. Microphones are usually classified with respect to the transducer principle used. Their directional characteristics are also of interest, that is, the voltage output as a function of the direction of incidence for constant sound pressure. *See* DIRECTIVITY; SOUND; SOUND PRESSURE; TRANSDUCER.

In addition to directional characteristics, some other important characteristics of microphones include open-circuit sensitivity, equivalent noise level, dynamic range, and vibration sensitivity.

Open-circuit sensitivity is defined as the ratio of open-circuit output voltage and sound pressure. The pressure sensitivity refers to the actual pressure acting upon the diaphragm of the microphone, while the free-field sensitivity refers to the pressure that existed in the sound field before insertion of the microphone. Pressure sensitivity and free-field sensitivity are equal at low frequencies. Sensitivities are measured in volts/pascal (V/Pa).

Equivalent noise level is equal to the level of a sound pressure which generates an output voltage of the microphone corresponding to its inherent A-weighted noise voltage. It is measured in dB(A).

Dynamic range is defined as the range of sound pressure levels in decibels (dB) extending from the equivalent noise level to the level where the nonlinear distortion reaches 3%.

Vibration sensitivity is defined as the ratio of the output voltage of the microphone as a result of acceleration of its case to the magnitude of the acceleration. Vibration sensitivities are measured in volts/g, where g is the acceleration of the Earth's gravity, or in volts/(m/s^2).

Some aspects of the acoustic performance of a few microphones are given in the **table**. Only microphones of current usage have been considered.

Microphone Types

The various microphone types differ by the transducers used for converting the acoustic into electric signals. Most commonly used are electrostatic, piezoelectric, dynamic, and magnetic transducers. These transducer principles are reversible, that is,

Acoustic characteristics of microphones

Microphone type	Frequency range, Hz	Sensitivity, mV/Pa	Equivalent noise level, dB(A)	Directional characteristics
Condenser				
Measuring microphone, dc-biased, 1-in. (25-mm) diameter	3–8000	50	10	Omnidirectional
Measuring microphone, dc-biased, 0.12-in. (3-mm) diameter	7–140,000	1.0	55	Omnidirectional
Measuring microphone, electret, 0.5-in. (12-mm) diameter	7–18,000	50	14	Omnidirectional
Low-noise measuring microphone, dc-biased, 1-in. (25-mm) diameter	7–10,000	100	−2.5	Omnidirectional
Studio microphone, RF-biased	12–20,000	25	10	Omnidirectional
Lavalier microphone, electret	100–20,000	8	30	Cardioid
Studio shotgun microphone, RF-biased	50–20,000	40	9	Club-shaped
Miniature microphone, electret, 0.16 × 0.12 × 0.08 in. (4 × 3 × 2 mm)	100–5000	10	25	Cardioid
Boundary microphone, electret	20–20,000	20	27	Hemisphere
Silicon (MEMS) microphone	100–10,000	8	35	Omnidirectional
Piezoelectric				
Telephone transmitter, unimorph	300–3400	4		Omnidirectional
Dynamic				
Studio microphone	30–20,000	1.3	30	Omnidirectional
Directional studio	30–20,000	1.8	30	Superdirectional

the microphones can also be used as sound generators. Very recently, micromachining techniques were introduced into transducer production for fabricating some electrostatic and piezoelectric microphones in silicon. These sensors are referred to as MEMS (microelectromechanical systems) or silicon microphones. Another new type is the optical microphone which converts acoustic signals into modulations of light waves; such an acoustic sensor has an optical rather than an electrical output. The classical carbon microphones, while used extensively in telephony in the past, have been almost completely replaced by other types of microphones. *See* LOUDSPEAKER.

Electrostatic (condenser) microphones. These consist of a fixed electrode (the backplate), a movable electrode (the diaphragm), and an air gap between the electrodes. To decrease the acoustic stiffness of the airgap, which is generally about 20 to 30 micrometers (0.8 to 1.2 mils) thick, the backplate is often perforated with holes connecting the air gap to a larger air cavity. The diaphragm is a thin [typically 4 to 6 μm (0.16 to 0.24 mil) nickel, Duralumin, or metalized plastic (for example, Mylar) foil under mechanical tension. The resonance frequency of the microphone is determined by the mass of the diaphragm and by the restoring forces, consisting of the tension of the diaphragm and the stiffnesses of the diaphragm and of the air gap. Below resonance, the system is controlled by the restoring forces and has a frequency-independent sensitivity. To achieve such a frequency-independent sensitivity, the resonance frequency is therefore placed at the upper end of the audio-frequency range. The two electrodes form a capacitor whose capacitance is very small, typically between 2 and 100 picofarads. The electrical impedance of the transducer is therefore relatively large. *See* CAPACITANCE; CAPACITOR; ELECTRICAL IMPEDANCE; RESONANCE (ACOUSTICS AND MECHANICS).

Impinging sound waves produce a motion of the diaphragm and a corresponding variation of the capacitance. This variation can be utilized in various ways to generate electrical output signals. Accordingly, three types of condenser microphones are customary, namely direct-current–biased systems, electret-biased systems, and radio-frequency (RF)–biased systems.

In the dc-biased system (**Fig. 1***a*), a polarizing voltage of about 200 V (or 20 V in portable systems) is applied between the microphone electrodes. Electrical output signals are thus generated by the diaphragm deflections. These are fed into a cathode follower with an input impedance R of typically 10^8 ohms or more. For frequencies larger than $1/RC$, where C is the microphone capacitance, the voltage delivered to the cathode follower is proportional to the sound pressure. To avoid capacitive loading of the transducer by stray capacitances and to eliminate electric pickup of RF signals, the cathode follower is directly built into the microphone case.

In the electret-biased system (Fig. 1*b*), a permanently charged dielectric (electret) is inserted between the two electrodes of the transducer. The electret, usually made of a 12 to 25 μm (0.5 to 1.0 mil) polymer film, such as poly(tetrafluoroethylene), either is cemented onto the backplate or is metalized on one side and used as the diaphragm of the microphone. The amount of charge is chosen so that it corresponds to a bias voltage of about 200 V. As in dc-biased systems, the electric output signal of the transducer is fed into a cathode follower. *See* ELECTRET TRANSDUCER.

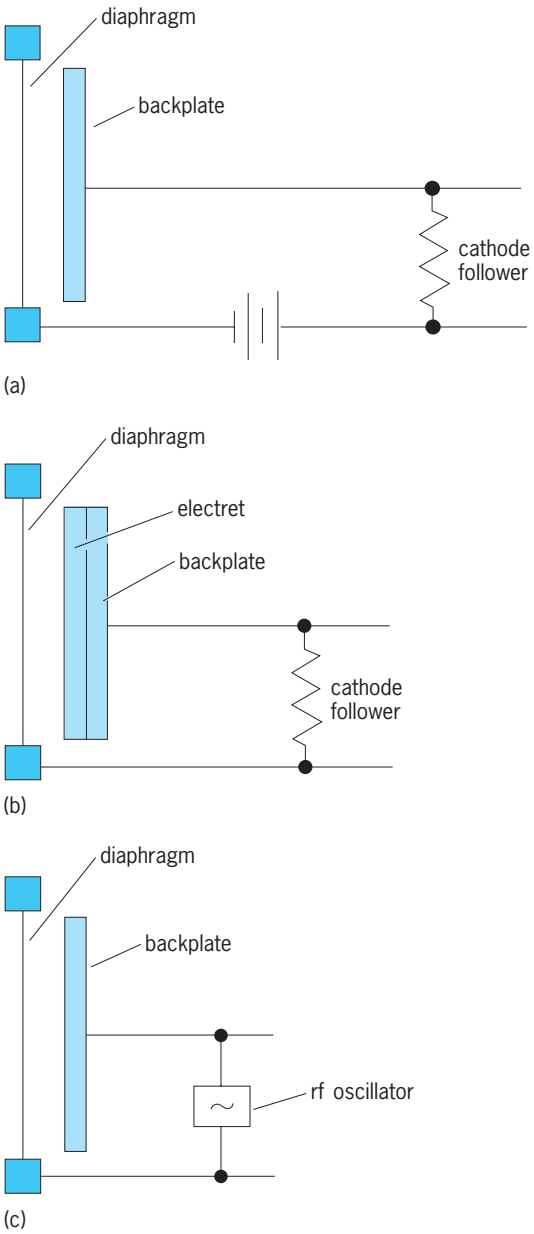

Fig. 1. Electrostatic transducers: (*a*) dc-biased system, (*b*) electret-biased system, (*c*) RF-biased system.

In the RF-biased system (Fig. 1*c*), the capacitance of the transducer determines the frequency of oscillation of an RF circuit, usually about 10 MHz. Motion of the diaphragm produces a frequency-modulated signal. On demodulation, a low-frequency signal corresponding to the sound pressure is obtained. Unlike dc-biased systems, the RF-biased transducer

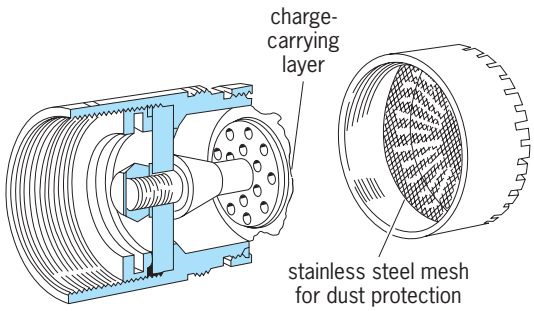

Fig. 2. Measuring condenser microphone. (*Brüel & Kjaer*)

has no low-frequency cutoff. *See* FREQUENCY MODULATION.

Condenser microphones are renowned for their excellent acoustic qualities such as flat frequency response, high sensitivity, large dynamic range, and small vibration sensitivity. Also important is their suitability for miniaturization, with the smallest units having dimensions of only about 0.12 × 0.12 × 0.08 in. (3 × 3 × 2 mm). They can be designed as precision instruments (**Fig. 2**) and as such are widely used in measurement and in high-fidelity sound production. Other applications, particularly of electret microphones, are in telephones, cassette recorders, camcorders, hearing aids, and toys. Electret microphones are the most widely used microphones with worldwide annual production of about 2 billion units (2006). *See* HEARING AID; MAGNETIC RECORDING; SOUND RECORDING; TELEPHONE.

Piezoelectric microphones. These consist of a material having piezoelectric properties. A deformation of the material leads to the generation of a voltage which corresponds to the deformation. Piezoelectric materials can be crystals, polycrystalline ceramics, or semicrystalline polymers. The best-known piezoelectric crystals are quartz and ammonium dihydrogen phosphate (ADP). Representative of polycrystalline ceramics are lead zirconate titanate (PZT) and barium titanate, which are initially electrostrictive; they have to be poled, that is, exposed to a high electric field at elevated temperatures, to become piezoelectric. An example of a semicrystalline polymer is poly(vinylidenefluoride) [PVDF]. It is also made piezoelectric by poling. *See* ELECTRET; ELECTROSTRICTION; PIEZOELECTRICITY.

Piezoelectric microphones for audio and near-ultrasonic frequencies are made from these materials in the form of flexure-mode transducers, such as bimorphs, unimorphs, or curved benders.

The bimorph consists of two piezoelectric disks, usually made of PZT, which are cemented together. Either the disks are poled in the same direction with the outputs connected in parallel (**Fig. 3*a***), or they are poled in opposite directions with the outputs connected in series (Fig. 3*b*). Since bending of the structures shortens one disk and lengthens the other, the voltage signals from the two disks add in both cases. The unimorph (Fig. 3*c*) consists of a single piezoelectric disk cemented to a thin metal plate. It has the advantage of being less fragile than the bimorph. Bimorphs and unimorphs are made in rectangular, square, or circular shapes.

To be operational, these elements have to be mounted in some way. Frequently applied is edge mounting along the entire edge or, for rectangular disks, cantilever mounting at one of the short sides of the rectangle. The transducers are then either directly actuated by the sound waves or actuated by means of a membrane connected by a stylus to some properly chosen point of the system.

The single-element bender (Fig. 3*d*) consists of a curved piezoelectric foil, usually PVDF, which is clamped around part of its edge. Preferred geometries are cylindrical sections or domes. If actuated by a sound wave, a voltage signal due to the piezoelectric effect.

In piezoelectric microphones, the resonance frequency is placed at the upper end of the audio range. At much lower frequencies, the system is stiffness-controlled. Since the output voltage is proportional to the displacement, a constant sensitivity results in this range. Because the resonance is very prominent, it has to be damped to ensure relatively constant sensitivity around this frequency.

Well-designed piezoelectric microphones have acceptable quality. A drawback is the relatively high vibration sensitivity. Directly actuated unimorph microphones are still in occasional use in telephones in some countries. Unimorphs and bimorphs are also employed in the near-ultrasonic range at frequencies up to about 100 kHz.

Dynamic microphones. These consist of a conductor located in the gap of a permanent magnet. Motion of the conductor produces a voltage proportional to its velocity. Depending on the kind of conductor used, moving coil and ribbon microphones are distinguished.

In the moving-coil microphone (**Fig. 4**) the coil, often referred to as voice coil, is connected to a diaphragm actuated by the sound waves. Motion of the coil induces a voltage proportional to its velocity. To obtain a frequency-independent sensitivity, the coil must respond to the sound pressure with frequency-independent velocity. This is accomplished by

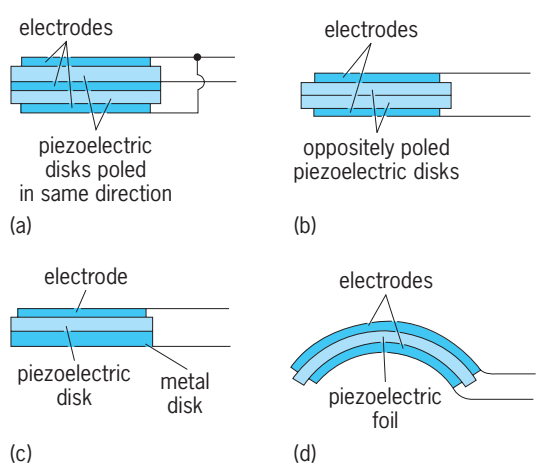

Fig. 3. Piezoelectric transducers: (*a*) parallel bimorph, (*b*) series bimorph, (*c*) unimorph, (*d*) curved bender.

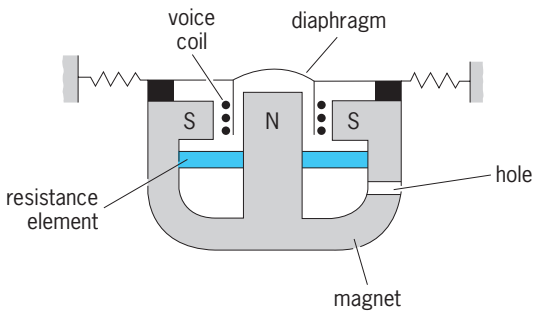

Fig. 4. Dynamic microphone of the moving-coil type.

resistance-controlling the system: the acoustical resistance is made larger in magnitude than the acoustical reactance due to the mass of the diaphragm and coil and due to the compliance of the suspension. A silk cloth or a piece of felt placed behind the voice coil is used for this purpose. The major resonance frequency of the system is usually placed in the range of 500 to 1000 Hz. To equalize the sensitivity in the lower and upper parts of the audio range, other resonances of the microphone are used. In modern moving-coil microphones, the diaphragm is made of a plastic film, for example, 15-μm (0.6-mil) Mylar. The impedance of the voice coil is typically 200 to 1000 ohms. *See* ACOUSTIC IMPEDANCE.

In the ribbon microphone, a metallic ribbon is placed in a magnetic field. As in the moving-coil microphone, the ribbon has to be terminated into an acoustical resistance that is large compared to the reactance of the system if a frequency-independent sensitivity is desired.

Dynamic microphones are relatively complicated systems. If well designed, they are of good quality. Drawbacks are the difficulties encountered in miniaturization and the relatively high vibration sensitivity. Moving-coil microphones are still widely used in high-fidelity, radio, television, and concert applications. In many other areas they have been replaced by electret microphones. Ribbon microphones were very popular in the past but are now rarely employed.

Magnetic microphones. These consist of a diaphragm connected to an armature which, when vibrating, varies the reluctance in a magnetic field. The variation in reluctance leads to a variation in the magnetic flux through a surrounding coil and therefore to an induced voltage. This voltage is proportional to the velocity of the armature. To obtain a frequency-independent sensitivity, the velocity of the armature in response to the sound pressure must be independent of frequency. As in dynamic microphones, this is accomplished by resistance-controlling the system, for example, by placing an acoustic resistance behind the diaphragm.

Magnetic microphones are relatively complicated and have poor frequency response and high vibration sensitivity. While never extensively used, they have now disappeared completely. However, the magnetic principle is still used in telephone receivers and in earphones employed in hearing aids. *See* EARPHONES.

Silicon (MEMS) microphones. Microelectronic processing methods allow fabrication of batch-processed, high-performance sensors. Among these methods are doping, deposition, oxidation, lithography, and, as key technologies, various etching processes. The last are usually subdivided into dry and wet methods. The dry processes, such as plasma, sputter, ion, and laser etching, allow one to fabricate holes with walls perpendicular to the surface of the substrate. The wet methods, subdivided into isotropic and anisotropic processes, usually result in curved walls or walls extending under a certain angle to the surface. An important variant of the etching processes is sacrificial-layer etching where a buried layer of material is etched out through holes in the covering layer. The etching technologies enable fabrication of membranes, holes, pits, and recesses as required in acoustic sensors. The methods in their entirety, if used to make three-dimensional mechanical structures, are referred to as micromachining, and the resulting devices are called microelectromechanical systems, or MEMS.

Acoustic MEMS sensors or silicon microphones utilizing the condenser and the piezoelectric principles have been built since the mid-1980s. In addition, new concepts of transducer design, such as the modulation of the drain current of a field-effect transistor or the modulation of light propagation in an optical waveguide by the sound waves, have been realized in silicon. Silicon microphones based on the condenser principle are now commercially available. *See* FIBER-OPTIC SENSOR; TRANSISTOR.

Silicon condenser microphones can be designed as two-chip or one-chip devices. A typical two-chip transducer (**Fig. 5**) consists of a membrane chip and a back electrode chip where the two chips, which are bonded together, are separated by a thin air layer. Since bonding of two chips complicates the design, modern silicon condenser microphones consist of only a single chip.

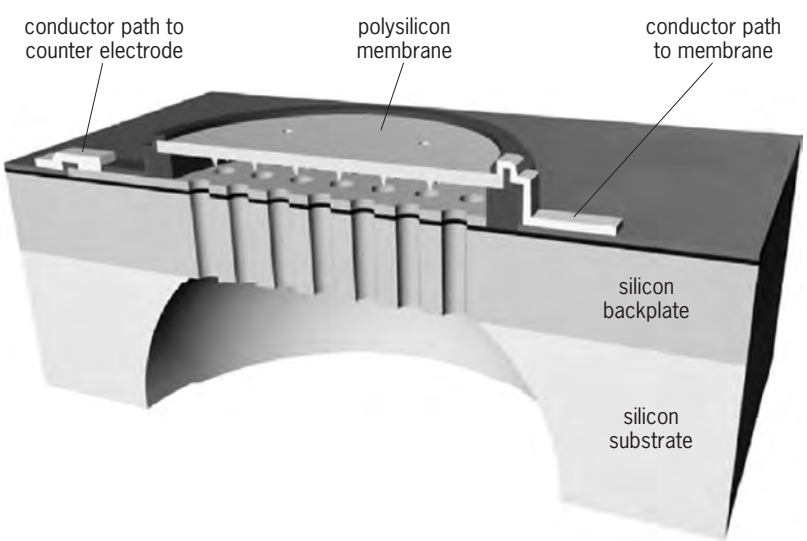

Fig. 5. One-chip silicon condenser microphone consisting of a 1-mm polysilicon membrane, an approximately 2-μm air gap, and a highly perforated silicon back plate.

The cross section of a typical one-chip condenser microphone is shown in Fig. 5. The sensor is fabricated by first growing an epitaxial and doped silicon layer which is highly perforated and which later serves as the backplate. Next, a sacrificial layer of silicon dioxide is deposited whose thickness determines the height of the air gap. On top of this, a polysilicon layer is placed and metalized on its upper side. Thereafter, the wafer backside is wet-etched and the sacrificial oxide is removed through the perforation holes.

The electroacoustic properties of such microphones depend strongly on the geometric, mechanical, and electrical characteristics. For typical microphones with membrane dimensions of 1×1 mm^2 and membrane thicknesses of 0.2 to 0.4 μm, the resonance frequency is in the near-ultrasonic range, and the sensitivity in the audio-frequency range is about 10 mV/Pa for a bias voltage of 3 V (see table).

Compared to conventional condenser microphones, silicon microphones are extremely shock-resistant and have, due to the small mass density of the membrane, a very low vibration sensitivity. The MEMS microphones are also, compared to electret microphones, much less sensitive to exposure to high temperatures. They can be permanently exposed to temperatures of 100°C and are resistant to temperatures up to 260°C for short exposure times. Because they withstand the heat prevailing during standard reflow soldering processes and because of their small height, such microphones can be inserted as surface-mounted devices (SMDs) in circuit boards on automated production lines. Some new silicon SMD microphones possess an integrated AD converter and thus deliver a digital output signal. Silicon microphones are now in commercial use in cellular phones, laptops, PDAs, digital cameras, MP3 recorders, and other devices. Although introduced rather recently, these microphones are already produced in quantities of several 100 million annually.

Piezoelectric silicon microphones are always one-chip designs (**Fig. 6**). They consist of a diaphragm, usually an approximately 1-μm-thick silicon nitride layer, which carries the lower electrode. Onto this, a piezoelectric layer of about equal thickness is deposited. It consists of either a ceramic such as lead zirconate titanate, or a polymer such as polyvinylidenfluoride. Another metal layer forms the upper (second) electrode. Microphones of this kind show somewhat higher noise levels than the condenser types.

Silicon microphones have several advantages as compared to conventional microphones. They can be made considerably smaller with membrane areas of only about 1 mm^2, as opposed to about 5 mm^2 for the smallest conventional transducers. They also have very low vibration sensitivity due to the use of thin diaphragms. They are thus not susceptible to pickup originating from vibrations due to walking or other motions or caused by vibration sources such as motors in cassette recorders or camcorders. Furthermore, they can be produced together with proper signal-processing electronics on the same chip with the same semiconductor methods. Finally, they can be made inexpensively through batch-processing techniques.

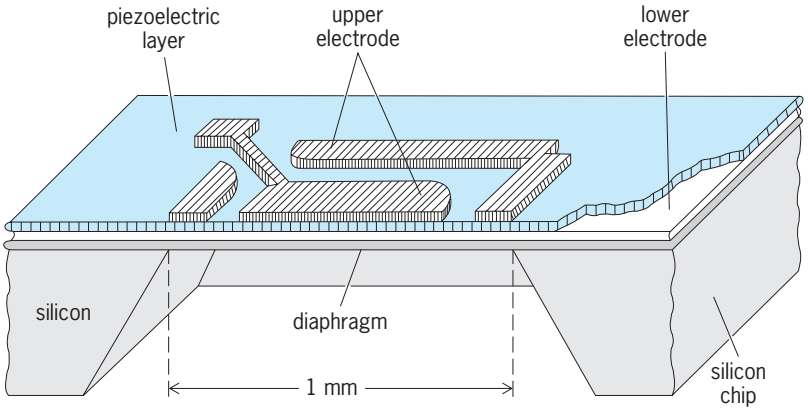

Fig. 6. Piezoelectric microphone consisting of a single silicon chip with a piezoelectric layer integrated into the diaphragm.

Directional Characteristics

The microphones described above are inherently nondirectional (omnidirectional) as long as they are small compared to the wavelength and if the microphone case shields the rear side of the diaphragm from the sound waves. This means that they pick up sound from all directions with equal sensitivity. They are therefore referred to as pressure microphones since pressure is a scalar and not a vector quantity.

In many applications, however, microphones with other-than-omnidirectional characteristics are desired. Such microphones [have a reduced response for all but certain directions and thus are capable of suppressing unwanted sounds. This results in] an improvement of the signal-to-noise ratio [for the desired acoustic signals.] A great variety of directivity patterns can be achieved with microphones based on either the gradient or the dimensional scheme. Combinations of both principles are also useful.

Gradient scheme. The directivity of a first-order gradient microphone depends on the pressure difference between two microphone elements separated by a distance that is small compared to the wavelength. Second-order gradients are formed by subtracting the outputs of two displaced first-order gradients. Analysis shows that the directivity of such systems is always independent of frequency and transducer dimensions as long as the above condition about the separation of the elements holds. Three common directivity patterns achievable with first-order gradients are shown in **Fig. 7**.

A simple gradient microphone with the bidirectional sensitivity pattern shown in Fig. 7*a* is obtained by exposing the diaphragm of a transducer to the sound field from both sides. Such microphones, because of the large sensitivity in the backward direction, are infrequently used.

Of great practical use, however, are the superdirectional and unidirectional microphones based on the gradient principle whose directivity patterns are

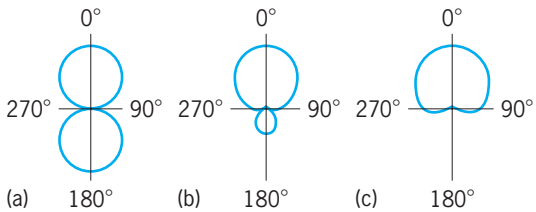

Fig. 7. Polar directivity patterns: (*a*) bidirectional, (*b*) superdirectional (supercardioid), (*c*) unidirectional (cardioid).

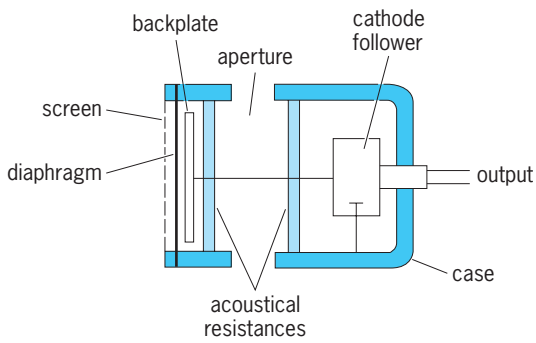

Fig. 8. Sectional view of a unidirectional microphone employing a condenser transducer.

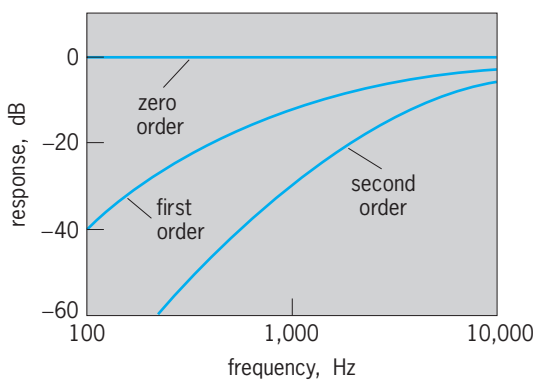

Fig. 9. Frequency-response characteristics of zero-, first-, and second-order gradient microphones, assuming a frequency-independent characteristic of the zero-order system.

shown in Fig. 7*b* and *c*. Such transducers are also made by exposing the diaphragm to the sound field from both sides, but they require an additional delay of the sound waves traveling to the back of the diaphragm. A unidirectional microphone employing a condenser transducer is schematically depicted in **Fig. 8**. An acoustical resistance is located behind the diaphragm to provide resistance control of the vibrating system. The back of the diaphragm is coupled to the acoustic inertance of the ports in the side of the case and the acoustic resistance and acoustic capacitance of the air volume in the case. The phase shift introduced by the acoustic network consisting of the inertance, resistance, and capacitance corresponds to the distance from the ports to the front of the diaphragm. As a result, for sound arriving from the back, the forces on the front and back of the diaphragm are almost equal in phase and amplitude, and the response is low. The phase shift is at a maximum for sound arriving in a forward direction, with resultant maximum sensitivity. Thus, a unidirectional directivity pattern is obtained.

As compared to the sensitivity of a pressure (zero-order gradient) microphone, the sensitivity of a first-order gradient microphone to plane waves from a far sound source decreases with decreasing frequency. This is shown in **Fig. 9** where a flat response for the zero-order gradient is assumed. An even greater sensitivity loss toward low frequencies is found in second-order gradients. However, all gradients show equal sensitivity to the spherical waves of a close sound source. The first- and second-order gradients can thus be used as "close-talking" microphones with frequency-independent sensitivity, having the advantage of suppressing the low-frequency components from far-away noise sources.

Dimensional scheme. The dimensional-directivity scheme depends on interference of signals with phase differences which are exclusively due to the geometry of the system. For the microphone to be directive, the dimensions of the transducer have to be comparable to the wavelength. As opposed to gradient-type microphones, the directivity of such microphones is dependent on the ratio of linear dimensions to wavelength, as shown in **Fig. 10** for a line microphone, that is, a microphone picking up sound along a line. As this ratio increases, the directivity changes from a cardioid to a club-shaped pattern. Other examples of dimensional microphones are transducer arrays and reflectors. *See* INTERFERENCE OF WAVES.

An implementation of the dimensional scheme in the form of a line microphone consists of a long tube which has a large number of sound inlets along its side. An omnidirectional microphone is placed at the end of the tube (**Fig. 11**). For incidence along the axis of the tube in the direction toward the

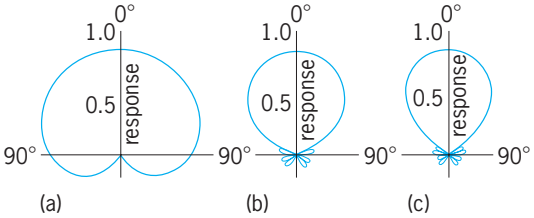

Fig. 10. Directivity patterns for a line microphone at different frequencies. The response maximum is arbitrarily chosen as unity. (*a*) Length of line = $\lambda/2$, where λ = wavelength (low frequencies). (*b*) Length of line = 2λ (intermediate frequencies). (*c*) Length of line = 8λ (high frequencies).

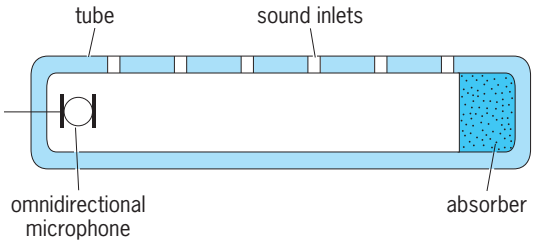

Fig. 11. Schematic section of a line microphone.

omnidirectional microphone, the sound waves entering the tube through any of the holes are all in phase, the interference is constructive, and maximum sensitivity results. For incidence at other angles, phase differences between the sound waves entering the tube through one hole and sound waves entering through another hole appear, and the interference is fully or partially destructive. The directivity patterns are those of Fig. 10. Systems of this kind, referred to as shotgun microphones, are in wide use.

The boundary microphone also uses the dimensional scheme. Such a microphone consists of an omnidirectional transducer set into a disk. At high frequencies, where the wavelength is smaller than the lateral dimension of the disk, sound waves arriving from the front direction experience a 6-dB boost, while sound waves arriving from the rear direction are attenuated. This results in an approximately hemispherical directional (characteristics) of the boundary microphone.

The dimensional scheme is often combined with the gradient principle to improve the directivity of cardioid or supercardioid microphones at higher frequencies. For example, a cardioid capsule can be used in a boundary microphone with the effect of sharpening the unidirectional characteristic.

Sensitivity Calibration

A number of methods may be employed for determining the sensitivity of microphones. The pistonphone and the reciprocity calibration are widely used.

A pistonphone is used to calibrate pressure-type microphones (**Fig. 12**). The small piston is driven by a crank and thus produces a pressure that can be exactly calculated from the geometry of the system. The pistonphone method is useful for calibrating microphones in the low-frequency range. The upper limit is governed by the permissible speed of the mechanical system, which corresponds to approximately 300 Hz. Pressure and free-field sensitivities do not differ in this range. The accuracy of the pistonphone-calibration method is ±0.2 dB.

The reciprocity method allows sensitivity measurements in the entire audio-frequency range and beyond, and can be adapted to yield either the pressure or the free-field sensitivity of microphones. In both cases, an auxiliary reversible transducer S_1 and an auxiliary loudspeaker S_2 are required in addition to the microphone M to be calibrated. A number of experiments have to be performed to determine the sensitivity of M.

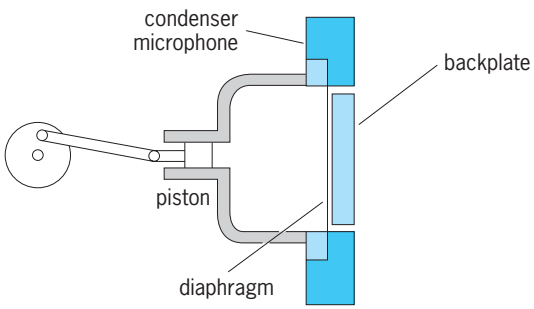

Fig. 12. Sectional view of a pistonphone.

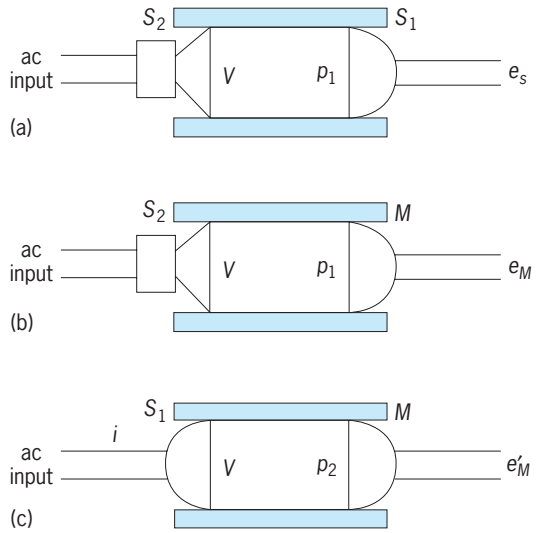

Fig. 13. The three experiments of the reciprocity calibration for obtaining the pressure sensitivity of a microphone. (*a*) Open-circuit voltage e_s of the reversible microphone loudspeaker S_1 when used as a microphone and actuated by a sound pressure p_1. (*b*) Open-circuit voltage e_M of the microphone M to be calibrated, when actuated by a sound pressure p_1. (*c*) Open-circuit voltage e'_M of the microphone M to be calibrated, when actuated by a sound pressure p_2 produced by the reversible microphone loudspeaker S_1 used as a loudspeaker with a current input i in a coupling volume V.

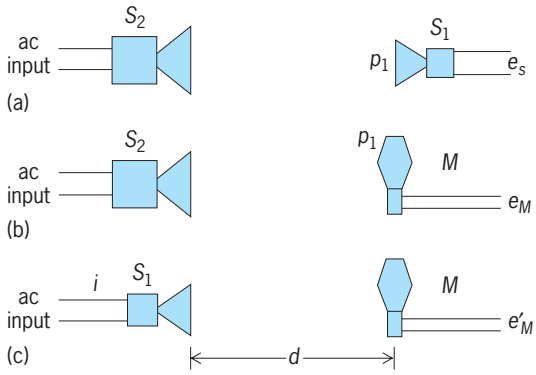

Fig. 14. The three experiments of the reciprocity calibration for obtaining the free-field sensitivity of a microphone. (*a*) Open-circuit voltage e_S of the reversible microphone loudspeaker S_1 when used as a microphone and actuated by a sound pressure p_1. (*b*) Open-circuit voltage e_M of the microphone M to be calibrated, when actuated by a sound pressure p_1. (*c*) Open-circuit voltage e'_M of the microphone M to be calibrated, when actuated by a sound pressure produced by the reversible microphone loudspeaker S_1 used as a loudspeaker with a current input i and a spatial separation d.

The procedure for the pressure-response calibration is schematically shown in **Fig. 13**. In Fig. 13*a*, an alternating current is fed to the loudspeaker S_2. A sound pressure p_1 is produced in the volume V. It generates an open-circuit voltage e_S of S_1. In the next experiment (Fig. 13*b*), the same current is fed to the loudspeaker S_2, resulting in an open-circuit voltage e_M of the microphone M. In the third experiment, a current i is fed to S_1 which is now used as a

loudspeaker. This produces the output voltage e'_M of M. The sensitivity K_M of the microphone M is then given by Eq. (1), where $\omega = 2\pi f$, f is the frequency,

$$K_M = \sqrt{\frac{\omega V e_M e'_M}{\rho_0 c^2 i e_S}} \quad (1)$$

ρ_0 is the density of air, and c is the sound velocity.

The free-field calibration by means of the reciprocity procedure is shown in **Fig. 14**. In this case, the sensitivity K_M of the microphone M is given by Eq. (2), where e_S, e_M, e'_M, and i are obtained from the

$$K_M = \sqrt{\frac{2d\lambda e_M e'_M}{r_A i e_S}} \quad (2)$$

experiments in Fig. 14, d is the distance as shown in the figure, λ is the wavelength of the sound wave, and r_A is the characteristic impedance of air.

The reciprocity calibration is a very accurate method since it depends solely on measurements of electrical quantities and length dimensions. Its accuracy is estimated to be ± 0.02 dB. Gerhard M. Sessler

Bibliography. J. Eargle, *The Microphone Book*, 2d ed., Focal Press–Elsevier, 2004; G. Elko, Microphone arrays, in T. Rossing (ed.), *Springer Handbook of Acoustics*, Springer, 2006; E. L. Hixson and I. J. Busch-Vishniac, Transducer principles, pp. 1375–1388, and I. J. Busch-Vishniac and E. L. Hixson, Types of microphones, pp. 1411–1422, in M. J. Crocker (ed.), *Handbook of Acoustics*, Wiley, 1998; R. Lerch, Sensors for measuring sound, in W. Göpel, J. Hesse, and J. N. Zemel (eds.), *Sensors: A Comprehensive Survey*, vol. 7, Wiley-VCH, 1993; G. M. Sessler, Silicon microphones, *J. Audio Eng. Soc.*, 44:16–22, 1996; G. S. K. Wong and T. F. W. Embleton (eds.), *AIP Handbook of Condenser Microphones*, AIP Press, 1995; G. Wong, Microphones and their calibration, in T. Rossing (ed.), *Springer Handbook of Acoustics*, Springer, 2006.

Microprocessor

A device that integrates the functions of the central processing unit (CPU) of a computer onto one semiconductor chip or integrated circuit (IC). In essence, the microprocessor contains the core elements of a computer system, its computation and control engine. Only a power supply, memory, peripheral interface ICs, and peripherals (typically input/output and storage devices) need be added to build a complete computer system. *See* COMPUTER PERIPHERAL DEVICES.

Internal architecture. A microprocessor consists of multiple internal function units. A basic design has an arithmetic logic unit (ALU), a control unit, a memory interface, an interrupt or exception controller, and an internal cache. More sophisticated microprocessors might also contain extra units that assist in floating-point match calculations, program branching, or vector processing (see **illus.**).

The ALU performs all basic computational operations: arithmetic, logical, and comparisons.

The control unit orchestrates the operation of the other units. It fetches instructions from the on-chip cache, decodes them, and then executes them. Each instruction has the control unit direct the other function units through a sequence of steps that carry out the instruction's intent. The execution path taken by the control unit can depend upon status bits produced by the arithmetic logic unit or the floating-point unit (FPU) after the instruction sequence completes. This capability implements conditional execution control flow, which is a critical element for general-purpose computation. *See* BIT.

The memory interface enables the microprocessor to maintain two-way communication with off-chip semiconductor memory, which stores programs and data. This interface typically supports memory reads and writes in blocks of words (the number of bits that the processor operates on at one time). The block size facilitates burst data transfers to and from the chip's internal cache. *See* SEMICONDUCTOR MEMORIES.

The interrupt or exception controller enables the microprocessor to respond to requests from the external environment or to error conditions by allowing interruptions of the ongoing operation. An interrupt might be an external peripheral requesting service, while an exception typically consists of a floating-point math error or an unrecognized instruction. The interrupt controller can prioritize and selectively handle these interrupts.

The internal cache is an on-chip memory storage area that holds recently used data values or instruction sequences that are likely to be used again in the near future. Since this information is already on-chip, it can be accessed rapidly, thereby accelerating the computation rate. Items not in the cache can take several or more extra operations to access, which significantly degrades the computation rate. Software writers often organize a program's code and data structures so that the most frequently used elements often occupy the cache, thus maintaining a high level of computational throughput. *See* COMPUTER ARCHITECTURE; COMPUTER STORAGE TECHNOLOGY.

Computational power. The computation power of microprocessors can be roughly characterized by their word size and computational rate (operations per second). The earliest microprocessors, introduced in the mid-1970s, were 4- and 8-bit machines operating at a clock rate of 1 MHz. They could process a few hundred thousand operations per second. By the start of the twenty first century, microprocessors had evolved into 32- and 64-bit machines operating at clock speeds of 500 MHz to 1 GHz, and capable of processing rates from 100 million to nearly 2 billion operations per second. *See* DIGITAL COMPUTER; MICROCOMPUTER.

The rapid growth of the microprocessor's computing power, especially in the 1990s, was due to two factors. First, the rapid advance of semiconductor fabrication technology reduced the dimensions of the transistor (the basic building block of a digital

computer) from hundreds of micrometers to fractions of a micrometer. This allows more transistors to be packed onto the chip. The smaller transistors can also switch faster, resulting in faster computations. Today's microprocessors often contain 10 million transistors, with some having more than 20 million transistors. Over half of the transistors in these microprocessors are part of the cache unit, which indicates how critical this unit is in maintaining computational throughput. *See* INTEGRATED CIRCUITS; TRANSISTOR.

Improvements to the microprocessor's architectural design is the second contributor to its improved performance. Such improvements include simplified instruction sets and sophisticated function units that operate in parallel. The latter ability enables the microprocessor to support multiple issue, a technique where several instructions start at the same time. This is done by the control unit issuing instructions to the other autonomous function units. Function units also support pipelining, where an instruction's control sequence is divided into distinct stages (for example, decode, data fetch, execute, data write). Much like an auto assembly line, each stage of the pipeline contains a partially processed instruction. On each tick of the processor clock, one instruction exits the last pipeline stage, other instructions move to subsequent processing stages, and another instruction enters the pipeline. The net result is that the microprocessor is capable of executing one instruction per clock cycle. The combination of multiple instruction issue and instruction pipelining effectively multiplies the amount of work that the microprocessor accomplishes during a clock period. *See* CONCURRENT PROCESSING.

Another trend improving microprocessor performance is the integration of more functions onto the chip. For example, the floating-point unit used to be a separate processor, but it is now part of a microprocessor's complement of function units. This accelerates floating-point math calculations since the floating-point data and instructions now execute on-chip rather than on an external device. Some microprocessors have gone further in this area by integrating either peripheral integrated circuits or vector processing units onto the chip.

RISC versus CISC. The design of instruction sets (the commands that produce basic work when executed by the microprocessor) often influences the design of the microprocessor itself. Instruction sets—and as a consequence, the microprocessor architecture—fall into two camps: reduced instruction set computers (RISC) and complex instruction set computers (CISC). Because of the limits of early computer technology, most computers were by necessity RISC machines. Since most of the software was written in assembly language (that is, a programming language that represented the program's intent in actual machine instructions), there was a drive to build instruction sets of greater sophistication and complexity. These new CISC instruction sets made assembly language programming easier, but they also made it difficult to build high-speed computer hardware. First, CISC instructions were harder to decode (a block move instruction might have many different parameters that the control unit had to recognize). In addition, since CISC instructions involved long and complex operation sequences, they incurred a major cost by requiring more complicated logic to implement. Second, such instructions were also difficult to interrupt or abort if an exception occurred. Finally, such instructions usually carried many data dependencies that made it more difficult to support advanced architectural techniques such as pipelining and multiple issue. By returning to a RISC design, much faster computers can be built. In fact, an enhancement in performance by a factor of 2 to 3 has been attributed to this simple organizational change. To achieve these efficiencies, most of the RISC microprocessor's function units must be kept as busy as possible. This requires optimizing compilers that can translate a program's high-level source code and then reorder the resulting low-level instructions in such a way as to ensure the high throughput. *See* COMPUTER PROGRAMMING; PROGRAMMING LANGUAGES.

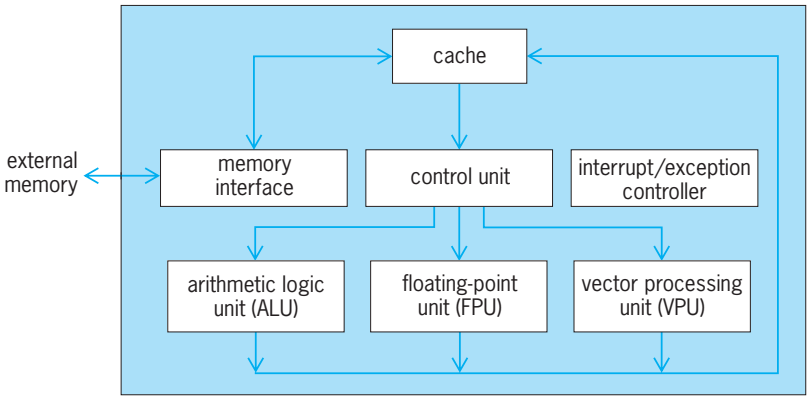

A microprocessor consists of multiple independent function units. The memory interface fetches instructions from, and writes data to, external memory. The control unit issues one or more instructions to other function units. These units process the instructions in parallel to boost performance.

Microprocessor families. There are several widely known microprocessor families. Like all computer families, microprocessor families are organized around software compatibility. Each new member of the family must be capable of executing the programs that run on its progenitors, preserving the substantial investment in software.

The most widely used microprocessor family, the x86, is a CISC architecture. This sacrifices some performance, but offers compatibility with software whose roots can be traced to the early 1980s. (However, many current x86 microprocessors actually decode the CISC instructions into RISC-like "micro operations" so that they can take advantage of pipelining and multiple issue techniques.)

The x86 family consists of Intel's 8080, 8085, 8086, 80286, 80386, 80486, the Pentium, Pentium II, and Pentium III processors. Other members of this family include Advanced Micro Device's K6 and Athalon processors. In this family, the Intel 80386 was the first 32-bit microprocessor. The 80486 integrated the floating-point unit on-chip, and later x86 family

processors used on-chip caches and multiple instruction issue to further boost performance.

The best representative of RISC processors is the PowerPC family, a joint design effort by Motorola and IBM. It consists of the PowerPC 601, 603, 603e, 604, 604e, 750 (also known as G3), and 7400 (also known as G4). The first PowerPC-based desktop computers appeared in 1994. The PowerPC is a 32-bit microprocessor, and features on-chip caches and an integrated floating-point unit. The G4 processor integrates a vector processing unit on-chip, which is valuable for three-dimensional graphics and scientific applications. Although the clock speed of the PowerPC family currently lags behind that of the x86 family, its RISC architecture, combined with instruction pipelining and multiple instruction issue, places its performance on par with that of the fastest x86 processors. *See* COMPUTER GRAPHICS.

Applications. In general, microprocessors can easily and effectively replace custom, large-scale integrated circuits in a variety of applications beyond basic computers. More importantly, the functions and features of a device—such as a network switch—can be readily improved or modified by simply installing new software into the device, rather than replacing custom hardware.

Microprocessors are found in virtually every consumer product that requires electric power, such as microwave ovens, automobiles, video recorders, cellular telephones, digital cameras, and handheld computers. High-performance microprocessors implement the servers that store and distribute Web content, such as streaming audio and video, desktop computers, and the high-speed network switches that constitute the Web's infrastructure. More modest-powered microprocessors are at the heart of notebook computers and electronic games. Low-power microprocessors provide the control and flow logic of hand-held devices, digital cameras, cellular and cordless phones, pagers, and the diagnostic and pollution control of automobile engines. *See* INTERNET; VIDEO GAMES; WIDE-AREA NETWORKS; WORLD WIDE WEB. Tom Thompson

Bibliography. K. Diefendorff, *The Microprocessor Report*, August 2, 1999; T. HalfHill, *BYTE*, April 1995; K. Short, *Microprocessors and Programmed Logic*, 2d ed., 1987; M. Slater, *A Guide to RISC Microprocessors*, 1992; M. Slater, *Microprocessor-Based Design*, 1989; J. Stewart, *Microprocessor Systems*, 1990; D. Tabak, *Advanced Microprocessors*, 2d ed., 1994.

Micropygoida

An order of regular echinoids belonging to the Diadematacea, established for the two recognized species of *Micropyga*. They have an aulodont lantern with grooved teeth. Test plating is imbricate, and ambulacra are composed of trigeminate compound plates in which upper and lower elements are reduced to demiplates. Pore pairs are biserially arranged in ambulacral columns, and there are unique umbrellalike aboral tube feet. *Micropyga* is a deep-water echinoid, found between 480 and 4290 ft (150 and 1340 m) depth in the Indo-West Pacific. There are no fossils that can be placed in this taxon with certainty, but it is possible that the Upper Jurassic *Pedinothuria* may belong here. *See* ECHINODERMATA. Andrew B. Smith

Bibliography. M. Jensen, Morphology and classification of Euechinoidea Bronn, 1860, a cladistic analysis, *Vidensk. Meddr dansk naturh. Foren.*, 143:7–99, 1981.

Microradiography

The process of producing enlarged images of the interior of thin, usually small specimens by penetration of low-energy (0.1–10 keV) x-rays. The magnification can be obtained geometrically during the exposure by subsequent enlargement of the initial image by optical or electronic means, or by a combination of both processes. As with radiography on other size scales, microradiography shows the spatial distribution of mass and elemental composition of the sample. If a pulsed (flash) source or time-gated detector is employed, radiographs also provide stop-motion details of fast-changing objects. Microradiography has numerous applications in biology, material science, the characterization of fabricated microstructures, and assessment of plasma-driven compression of thermonuclear fuel.

Microradiography is largely synonymous with x-ray microscopy, both techniques being concerned with producing enlarged images of opaque objects by using x-rays. However, x-ray microscopy of defects in crystals can be performed by using diffraction, rather than simple absorption, in a technique termed x-ray topography. Microradiography by x-ray absorption is complementary, or related to a variety of other techniques for characterization of microstructures. *See* X-RAY DIFFRACTION; X-RAY MICROSCOPE.

Principles. Radiography on all size scales is based on Beer's law, which relates the x-ray intensity incident on (I_0) and transmitted by (I) an object of density ρ and thickness x, as given by Eq. (1). The

$$\frac{I}{I_0} = \exp\left(-\frac{\mu}{\rho}\rho x\right) \tag{1}$$

mass absorption coefficient μ/ρ (cm^2/g) and the linear absorption coefficient μ (cm^{-1}) depend on the x-ray wavelength and absorbing material.

Constituents of cells such as protein and water have x-ray absorption lengths (the inverse of the linear absorption coefficient) that vary rapidly with energy in the soft x-ray region (**Fig. 1**) and correspond to small penetration ranges (approximately 1–10 micrometers). Radiography with submicrometer spatial resolution requires thin samples in order to avoid blurring and superposition of images, with attendant loss of resolution and contrast, in addition to obtaining the requisite penetration. Usually, microradiographic samples are less than 10 μm thick and 10–1000 μm in lateral dimension.

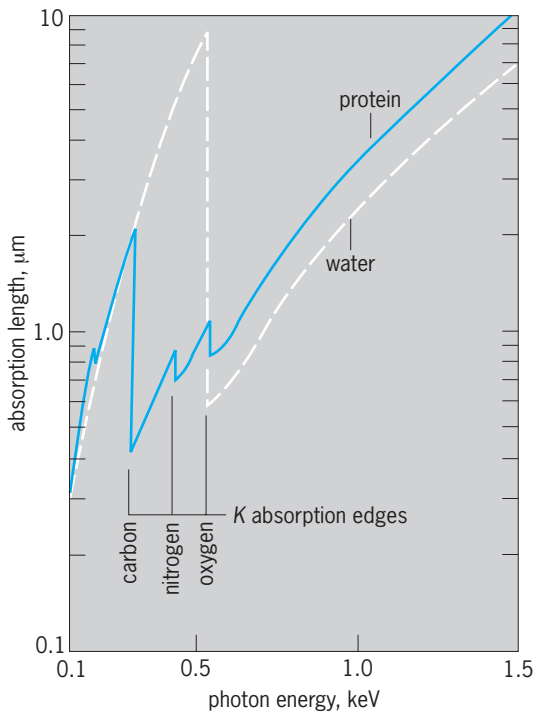

Fig. 1. Soft x-ray absorption lengths (inverse of the linear absorption coefficient) for protein typical of cellular materials and for water. *K*-electron absorption edge jumps are indicated for carbon, nitrogen, and oxygen. The large difference in absorption lengths between the carbon and oxygen edges provides good contrast for live biological specimens. (*After J. C. Solem and G. F. Chapline, X-ray biomicroholography, Opt. Eng., 23: 193–203, 1983*)

The very large steps in absorption length which occur at absorption edges (Fig. 1) are highly significant because they provide chemical specificity and contrast between different materials. These thickness (x) and chemical (μ) sensitivities can be quantified by differentiating Beer's law, as in Eq. (2). A

$$\frac{dI}{I} = -x d\mu - \mu dx \tag{2}$$

relative intensity change of 1% or greater is needed for reliable measurements. Hence, for constant μ, a change in thickness $dx = 10^{-2}/\mu$ is required. In the soft x-ray region, μ is typically of the order of 10^4 cm^{-1}, so a thickness change dx of approximately 10 nanometers is marginally resolvable. For constant thickness x, the required change in linear absorption coefficient is $d\mu = 10^{-2}/x$. Given typical values of $x = 10^{-4}$ cm, a change $d\mu = 100$ cm^{-1} is near the chemical resolution limit. Since μ is given by Eq. (3),

$$\mu = \sum_i W_i \mu_i \tag{3}$$

where W_i is the weight fraction of the ith component of a homogeneous material and μ_i is the corresponding linear absorption coefficient, concentration variations dW_i of 1% weight are barely resolvable in binary samples of $\mu_1 \gg \mu_2$ or vice versa. The technique of detecting small chemical variations by monochromatic microradiography is termed x-ray absorption microanalysis.

Equipment. There are two necessary components of each radiographic exposure, in addition to the object under study: a source of x-rays, and a means of detecting the transmitted x-radiation. Sources are sometimes augmented by x-ray optics to improve the efficiency of radiation collection or to provide monochromatic radiation. Optics are also employed to form images of specimens. Detection systems sometimes include intensifiers in order to shorten exposure times or to permit work with weaker sources.

Sources. There are three classes of sources employed for microradiography: electron-impact devices, storage ring sources of synchrotron radiation, and multimillion-degree plasmas. The first microradiography measurements, and most studies since then, have been done with electron-impact sources. These contain an electron-emitting cathode and an anode in a high vacuum, between which a high voltage is applied. Such devices vary from sealed x-ray tubes to components mounted within special-purpose vacuum systems. Electron-impact sources emit both line and continuum x-radiation. They range in size from about 1 μm to 1 cm. Although usually run continuously, some electron-impact sources can be operated in a pulsed mode. *See* X-RAY TUBE.

High-energy (of the order of gigaelectronvolts) electrons orbiting within an evacuated toroid, called a storage ring, produce intense continuum radiation which is tightly collimated and polarized. Such synchrotron radiation is emitted in nanosecond pulses at MHz rates, and can be 1000 or more times as intense as the x-rays from electron-impact sources. The source cross section is typically 0.2 × 2 mm. A monochromator between the storage ring and sample allows radiation to be tuned over absorption edges of elements in the sample. The use of synchrotron radiation for microradiography began in 1977 and has rapidly expanded. *See* SYNCHROTRON RADIATION.

Very high-temperature plasmas can be heated by high-power lasers or electrical discharges. They are typically 10 μm to 1 mm in size. Multimillion-degree plasmas emit uncollimated line and continuum spectra predominantly in the soft x-ray region, below a few kiloelectronvolts, with pulse lengths of 0.1–100 ns. Stop-motion microradiographs were first made with plasma x-radiation in 1980. Fresnel zone plate x-ray lenses are sometimes used to gather (condense) synchrotron radiation, that is, to image the source, onto a microscopic sample. *See* PLASMA (PHYSICS).

Optics. Devices which monochromatize or focus x-rays incident upon or transmitted by specimens are employed in some microradiographic techniques. The dispersive elements in x-ray monochromators can be reflection or transmission gratings (manufactured) or crystals (natural or grown), depending on the x-ray wavelength. Focusing elements can also be of transmission or reflection character. Fresnel zone plates consist of equal-area alternating open and x-ray opaque annuli. They form soft x-ray images by interference of transmitted radiation. Since Fresnel zone

plates bring different wavelengths to focus at different axial positions, they can also serve as monochromators. Gratings and Fresnel zone plates are often made by using optical or electron beam patterning techniques and materials-processing methods developed for microcircuit production. Multilayer-coated optics, which serve simultaneously to monochromatize and focus x-rays, are produced by highly controlled vapor deposition onto substrates figured by diamond turning, followed by superpolishing. Grazing-incidence focusing optics can also be employed for microradiography. They are also produced by modern techniques for production of precision optics. *See* DIFFRACTION; INTEGRATED CIRCUITS; X-RAY OPTICS.

Detectors. Passive detectors which absorb x-rays and retain the image are most commonly used for microradiography. Photographic film, which provides resolution down to around 0.1 μm, is usually employed. Images recorded on film usually are magnified and observed optically. *See* PHOTOGRAPHIC MATERIALS; PHOTOGRAPHY.

Grainless polymeric media, called photoresists, offer resolution down to 0.01 μm. Absorption of x-radiation alters their solubility in specific solvents. In positive resists, absorption produces polymer chain scissions, reducing the molecular weight and increasing solubility. The reverse is true for negative resists. Development (dissolution) in a solvent matched to the resist results in an image which consists of surface relief in the photoresist. The images are magnified and transferred to film with an electron microscope.

Active electronic detectors of a wide variety can be employed for microradiography. Two-dimensional, "area" detectors employing multiple x-ray collecting anodes, or arrays of small solid-state detectors, or television-style readouts are available. Image converters transduce images from x-rays to electrons or visible light. They sometimes serve as image intensifiers also, reducing required source intensity. The images collected with area detectors can be stored digitally or they can be recorded on film. If a microradiographic image is formed by the two-dimensional raster scanning technique, in which a very fine x-ray beam is moved sequentially over the sample, then electronic detectors which do not offer spatial resolution are employed. *See* CHARGE-COUPLED DEVICES; LIGHT AMPLIFIER.

Techniques. X-ray micrographs may be obtained by four geometrically different techniques. Two of them, the contact and the projection methods, have been in use for decades. However, they have undergone significant development because of the availability of bright synchrotron radiation and plasma radiation sources. The other two techniques, the true imaging and the scanning methods, require the bright sources. Work on these methods goes back only to the mid-1970s, with major developmental efforts underway.

Contact. In this approach, the specimen either touches or is very close to the recording medium (**Fig. 2***a*). Resolution in the one-to-one image is determined primarily by the detector, although penumbral effects can contribute to image blurring for extended sources and relatively thick samples. Photographic film is most commonly employed for recording contact microradiographs, with magnified images being produced from the original record by optical means. However, the film grain size of 0.1 μm or greater sometimes provides an unacceptable resolution. In such cases, for example, imaging of live biological materials, it is necessary to use x-ray-sensitive grainless photoresist media. After coating the developed resist with a thin layer to ensure electrical conductivity, the images are read and enlarged with a scanning electron microscope for a thick resist or substrate, or a transmission electron microscope if the resist is thin and on a very thin substrate. Contact x-ray microradiographs have been made with

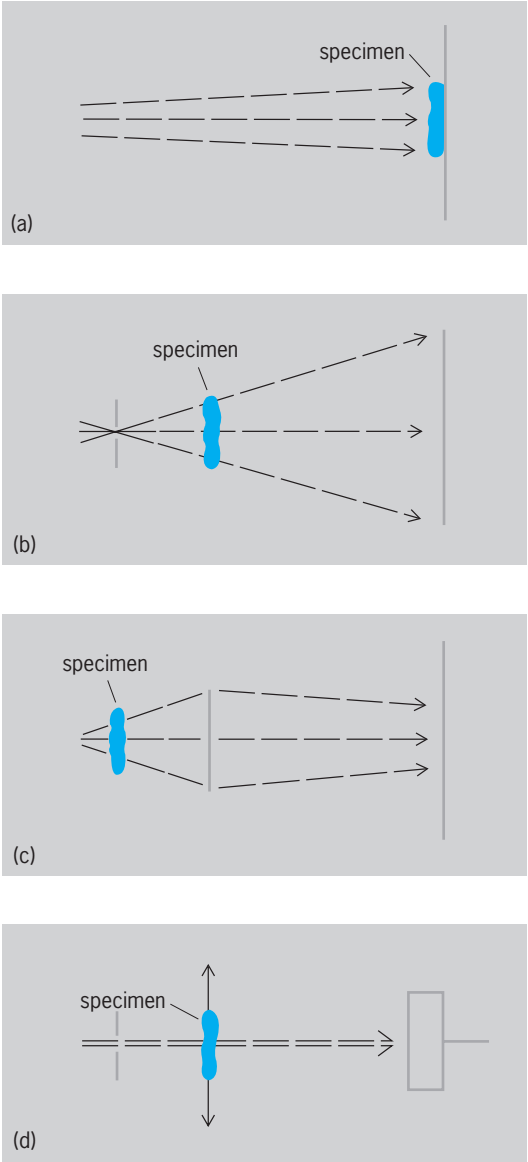

Fig. 2. Schematics of techniques for x-ray microradiography. X-ray paths are indicated by broken lines. Alternative configurations for some techniques are discussed in the text. (*a*) Contact. (*b*) Projection. (*c*) Imaging. (*d*) Scanning.

diverse specimens. Photoresist images of biological specimens exhibit features as fine as 10 nm. Time resolution of about 100 ns has been demonstrated with contact microradiography.

Projection. As with the contact method, the projection approach to microradiography depends on simple geometric shadow casting to form the image. However, in projection, the initial recording is a magnified image of the specimen. This recording is usually on film, although electronic image converters are also employed. The scale of the image depends on the ratio of the source-to-detector and source-to-specimen distances. The source for projection microradiography can be either an intrinsically small x-ray emitter or a larger source limited by a pinhole (Fig. 2*b*). If an extended source and pinhole are employed, the pinhole can be placed between the sample and recording medium. This is tantamount to taking an x-ray pinhole-camera photograph of a backlighted specimen.

The spatial resolution in projection is usually determined by the effective source diameter. Micrometer-sized focused electron-impact sources have long been used to provide striking projection pictures of the interior of complex biological specimens, notably insects. Laser-heated plasma sources of soft x-rays have been used to take static or dynamic microradiographs. Pinholes several micrometers in diameter have been employed to obtain flash microradiographs of laser-driven targets with time resolution of 100 picoseconds, and subpicosecond time resolutions are foreseen. Although the projection method has not been pushed below about 1 μm spatial resolution, it has the potential for resolving 100 nm.

Imaging. Ordinary refractive transmission lenses do not form images in the x-ray region. However, both transmission and reflection x-ray optics are available for x-ray imaging. Fresnel zone plates, as discussed above, with annuli small enough to provide resolution below 100 nm are small and inefficient. Hence, bright synchrotron sources are required for imaging microradiography. Two Fresnel zone plates are employed: the first monochromatizes and condenses x-rays onto the specimen, and the second then forms an image of the sample magnified several hundred times (Fig. 2*c*).

Film has usually been used to record imaged microradiographs with zone plates in a few seconds with synchrotron radiation. Image intensifiers will reduce exposure times to less than 100 ms. Flash microradiographs formed with submicrosecond pulsed plasma x-radiation are also possible. Imaging techniques may eventually yield spatial resolution near 10 nm.

Scanning. The above methods require that the entire sample be exposed to x-rays during image formation in relatively insensitive, high-resolution recording media. This constant exposure produces radiation damage in sensitive specimens, such as biological specimens, if prolonged exposures are made with steady sources. Fast, pulsed sources provide records before the sample is altered by x-ray absorption. A more attractive approach is to use sensitive, photon-counting techniques. However, high-resolution area detectors with the ability to sense individual soft x-ray photons are not available.

Large x-ray detectors, such as gas proportional or scintillation counters, are sensitive and convenient. With them, spatial resolution is achieved by raster-scanning the specimen through a fine x-ray beam while counting transmitted x-ray photons. X-rays strike, and damage, any spot on the sample only when its transmission is being measured. Live cells can be held at lower temperatures to immobilize them during a complete scan. The resolution-defining beam can be simply formed by pinhole aperturing of a larger x-ray beam (Fig. 2*d*). It can also be formed by use of two Fresnel zone plates, one for condensing and another for focusing, as in the imaging method. Reflective x-ray optics, coated with multilayer films which diffract x-rays as do crystals, can also be employed to form submicrometer, monochromatic spots for scanning microscopy.

Spatial resolution of 300 nm has been demonstrated with the scanning technique, and resolutions in the 10–50-nm range are anticipated. Because of the technique's sequential nature, scanning microradiographs require at least 100 s to record, although each picture element (pixel) may require only a few milliseconds to record. Brighter synchrotron radiation sources may reduce the time to record a complete scanning microradiograph to below 1 s. An attractive feature of scanning microradiography concerns the immediate digital nature of the data. Image magnification and analysis is readily performed with modern computer techniques.

Applications. Microradiography was developed in order to observe fine details in the interior of opaque natural (for example, biological) and manufactured (for example, metallurgical) samples. Optical microscopy has long been available to observe the surfaces of opaque specimens, or the interior of clear samples, with spatial resolutions approaching 200 nm. Early interest in microradiography waned for two reasons: weak sources required long exposure times, and electron beam techniques which provide submicrometer details were developed. The availability of bright soft x-ray sources, especially synchrotron radiation, has led to a resurgence of interest in microradiography, especially for cellular samples. Sub-100-nm resolution of live specimens is needed to understand biological structures with sizes between those observable with light microscopes (down to about 200 nm) and the molecular level (about 1 nm) probed by diffraction from ordered arrays of molecules. Microradiography of inorganic natural, manufactured, and dynamic structure remains of interest. *See* ELECTRON MICROSCOPE; OPTICAL MICROSCOPE.

Biology. Microradiographic techniques which would yield images of the ultrastructure of live cells with resolution near 10 nm are under development. Such techniques have immense potential as a research tool and, perhaps, even for clinical medicine. As shown in Fig. 1, radiation at energies in the range

between the carbon K absorption edge (0.284 keV) and the oxygen K edge (0.532 keV) provides excellent contrast between organic material and the water needed to sustain life in the specimen.

Two approaches to microradiography of cells have been undertaken. In one, contact radiographs are made of cells placed on photoresist. A very thin x-ray transmitting window [for example, 100 nm of silicon nitride (Si_3N_4)] retains a layer of water about 1 μm thick around the cells. Specimens for the initial demonstration were blood platelets. The microradiographs were recorded in a resist copolymer of poly(methyl methacrylate) and methacrylic acid. The exposure was made with a 100-ns flash of soft x-rays from a discharge-heated plasma, and a scanning electron micrograph was made of the developed resist (**Fig. 3**). Structures as fine as 10 nm are visible. Details not visible in exposures of dried platelets are observed. In particular, new structures near the base of pseudopods extending from the platelets are seen for the first time.

Scanning microradiography of biological materials is also under development. The use of zone plates to condense and focus radiation to a small spot has been demonstrated in a scanning microscope. **Figure 4** is a digitized (10^4-pixel) image of an algal cell taken with synchrotron radiation in about 1 h with 300-nm pixels. The relatively long exposure was due to the use of monochromatized 0.387-keV radiation. When scanning microradiography with 30-nm resolution is possible, it is expected to yield images exhibiting new structures.

Natural inorganic structures. Interest in microradiography of inorganic materials is not as great as for live biological materials. Inorganic materials can be coated with a conductor and examined by electron microscopes offering resolution to 1 nm. However, microradiography has utility for specimens thicker than 100 nm, especially if chemical information is desired.

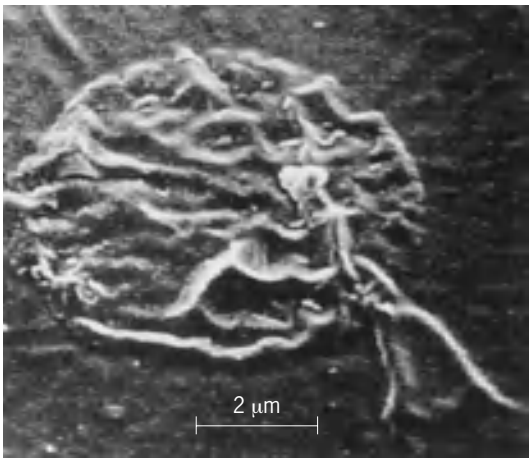

Fig. 3. Scanning electron micrograph of the contact microradiograph of a live human blood platelet recorded in a photoresist with a 100-ns burst of soft x-rays from a plasma heated by an electrical discharge. (*From R. Feder et al., Flash x-ray microscopy, Science, 227: 63–64, 1985*)

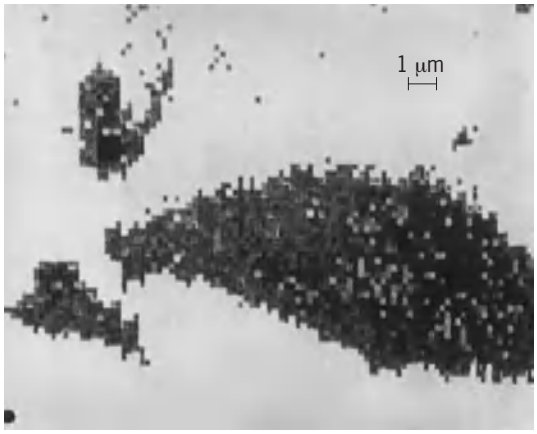

Fig. 4. Digital presentation of a scanning x-ray micrograph of a wet *Phaeodactylum* algal cell taken with synchrotron radiation. (*From H. Rarback et al., Recent results from the Stony Brook scanning microscope, in G. Schmahl and D. Rudolph, eds., X-ray Microscopy, pp. 203–216, Springer-Verlag, 1984*)

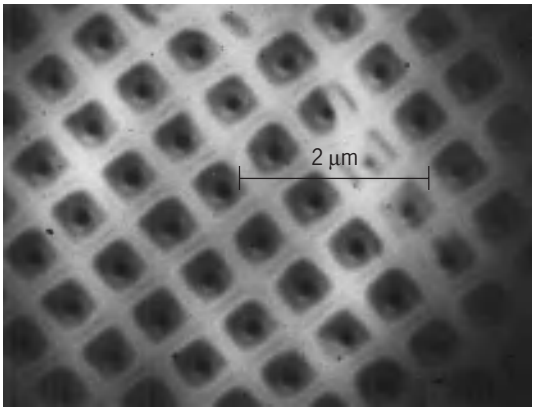

Fig. 5. Soft x-ray image of a diatom recorded on film with synchrotron radiation. (*From G. Schmahl et al., Zone plates for x-ray microscopy, in G. Schmahl and D. Rudolph, eds., X-ray Microscopy, pp. 63–74, Springer-Verlag, 1984*)

Diatoms are calciferous structures with submicrometer detail produced by sea life. They serve as convenient test objects for developmental microradiography techniques. **Figure 5** is an x-ray image of a diatom produced by passing 0.276-keV synchrotron radiation through a pair of Fresnel zone plates. The second plate yielded a 230× magnified image on film in a 3-s exposure. Details as fine as 50 nm can be seen in such images. Microradiographs of wet biological specimens have been taken with the imaging technique used to produce Fig. 5.

Artificial microstructures. A variety of techniques, especially those used in microelectronics production, are available to produce structures with detail on the micrometer scale. As with natural inorganic materials, scanning electron microscopy is usually the tool of choice for examining artificial microstructures. However, again, there is some need for examination of the interior of artificial microstructures which microradiography can fulfill.

Microballoons with diameters of 0.1 to 1 mm, filled with deuterium and tritium, are of interest in

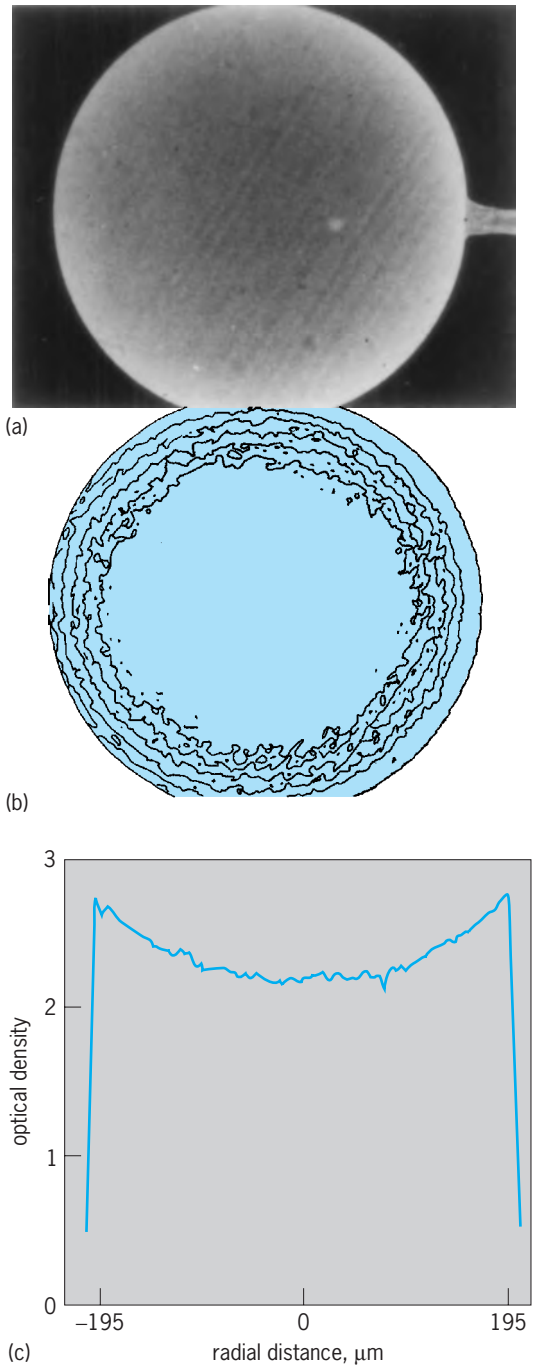

Fig. 6. Examination of a glass microballoon with 390-μm outer diameter. (*a*) Contact x-ray microradiograph of microballoon (mounting stalk on right). (*b*) Two-dimensional contour density representation. (*c*) One-dimensional density scan. (*From H. Kim and M. D. Wittman, X-ray microradiography of inertial fusion targets, J. Vac. Sci. Technol., 3:1262–1265, 1985*)

fusion research. Exterior irradiation of such small spheres (pellets) with very high-power, nanosecond-pulsed lasers blows off the pellet's surface. In response, the pellet and its nuclear fuel are crushed to a small diameter. This raises the fuel density and temperature to conditions required for ignition of thermonuclear burn, as in the hydrogen bomb but on a very much smaller scale. Uniform compression is needed to attain ignition, requiring uniform laser irradiation of pellets with uniform wall thickness. X-ray microradiography is employed, usually with steady electron-impact sources, for preshot examination and selection of microballoons. However, plasma radiation has also been used to inspect microspheres. **Figure 6** shows a 250× magnification of the radiograph of a pellet with a diameter near 400 μm, together with a two-dimensional contour density representation and a one-dimensional density scan. The contact image was recorded on film with 9-keV radiation in a 1-ns pulse from a laser-heated plasma. This rapid exposure, typical of laser-plasma x-ray sources, precludes vibration-induced blurring which may be present when a steady x-ray source is employed. *See* NUCLEAR FUSION.

Dynamic exposures. Very short x-ray pulses are needed to record rapidly changing structures, for example, a laser-driven microballoon. X-rays are employed because optical radiation will not penetrate the plasma which surrounds and drives the pellet inward. Projection microradiography with a pinhole between the microballoon being studied and film is employed. A separate laser beam which can be arbitrarily delayed relative to those striking the pellet provides a stroboscopic x-ray backlighting source. **Figure 7** shows a sequence of flash microradiographs of imploding pellets with initial diameter of 150 μm and final diameters during stagnation (prior to blowup) of about 50 μm. These microradiographs were taken with x-ray pulses of 100 picoseconds, to achieve very high simultaneous spatial resolution (about 10 μm) and time resolution (100 ps).

Related techniques. Microradiography is only one of many techniques useful for structural or chemical characterization of samples.

Electron microscopy, as discussed above, can provide spatial resolution better than 1 nm for the surface of thick samples (measured in the scanning mode) or for thin (<100 nm) samples imaged or scanned in transmission. Scanning and transmission electron microscopy is used to produce enlarged images of contact microradiographs recorded in photoresists (Fig. 3). Emitted x-ray intensity produced when a focused beam of energetic electrons is scanned over a sample can be used to provide maps of elemental distribution. This method, termed electron probe microanalysis, is a form of x-ray microscopy. However, it does not involve x-ray transmission through the sample and, hence, is not microradiography.

X-ray lithography, a method of pattern replication developed for microcircuit and microstructure production, is closely related to contact x-ray microradiography. In lithography, artificial patterns called masks are reproduced rather than natural specimens such as cells. The bright sources developed for fast commercial replication of a lithographic mask, notably dense multimillion-degree plasmas, are now applied to contact microradiography.

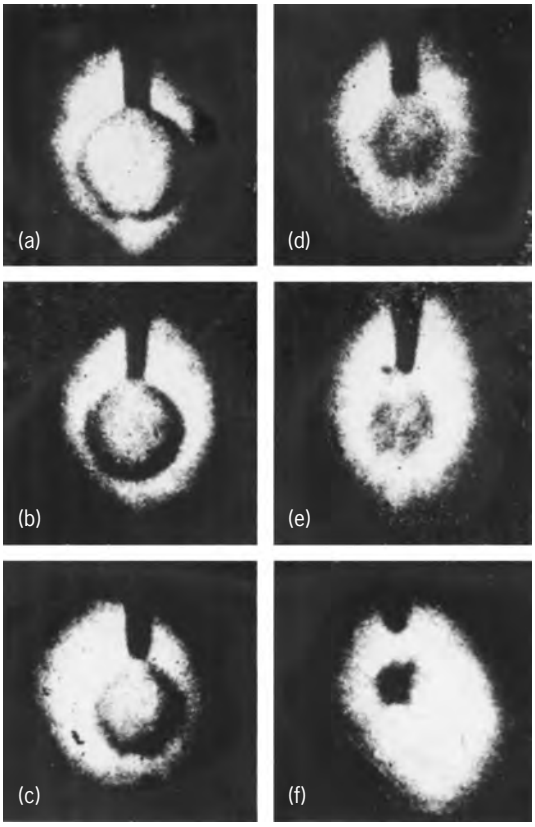

Fig. 7. Projection (pinhole) photographs of laser-irradiated microballoons recorded on film with 100-ps flashes of soft x-rays from oval-shaped laser-heated backlighting plasmas, at time delays between the 100-ps implosion-driving laser pulse and the x-ray strobe pulses of (*a*) 200 ps, (*b*) 350 ps, (*c*) 400 ps, (*d*) 450 ps, (*e*) 500 ps, and (*f*) 1000 ps. (*From C. Yamanaka, Progress of inertial confinement research in Japan, in C. Yamanaka, ed., Advances in Inertial Confinement Systems, pp. 43–56, Institute of Laser Engineering, Osaka, 1980*)

Holography ordinarily involves recording the interference pattern between an optical reference wave and the same wavelength scattered from the surface of an object. Holograms can also be formed by interference of reference and transmitted waves of penetrating x-radiation. Reconstruction of x-ray holograms with longer-wavelength (optical) radiation can provide enlarged, three-dimensional images of the object. X-ray holography requires sources with adequate temporal (longitudinal) and spatial (transverse) coherence to produce the requisite interference. X-ray lasers are under development, but they do not yet have wavelengths short enough to penetrate samples, nor do they exhibit adequate coherence. Bright synchrotron x-radiation sources can be monochromatized (for temporal coherence) and apertured (for spatial coherence) to provide intensity adequate for development of x-ray holographic techniques. X-ray holography promises to be an important technique, distinct from but closely related in its methodology and goals, to conventional microradiography. *See* HOLOGRAPHY.

David J. Nagel

Bibliography. P. C. Cheng and G. Jan, *X-ray Microscopy*, 1987; D. F. Parsons (ed.), Ultrasoft x-ray microscopy: Its application to biological and physical sciences, *Annals of the New York Academy of Sciences*, vol. 342, 1980; G. Schmahl and D. Rudolph (eds.), *X-ray Microscopy*, 1984; K. Shinohara and K. Yada (eds.), *X-ray Microscopy in Biology and Medicine*, 1990.

Microsauria

A diverse order of small, extinct amphibians, known only from the Pennsylvanian and lower Permian of North America and Europe. They range from obligatorily aquatic, perennibranchiate genera with lateral-line canal grooves, to fully terrestrial lizardlike forms. Several families include long-bodied, possibly burrowing, species. Limbs are always retained, and the tail is never specialized as a swimming organ. Microsaurs are recognized by the possession of a broad, strap-shaped occipital condyle, and no more than a single bone in the temporal series. The trunk vertebrae are spool-shaped. Some species have trunk intercentra, and most show caudal hemal arches. The number of presacral vertebrae ranges from 19 to 45. Limb and skull proportions are extremely variable, as is the dentition (see **illus.**). The specific origin of the group remains unknown. They are customarily allied with lysorophids, aistopods, and nectridians in the subclass Lepospondyli, but all the shared, derived characteristics of these groups may be correlated with small body size and do not necessarily

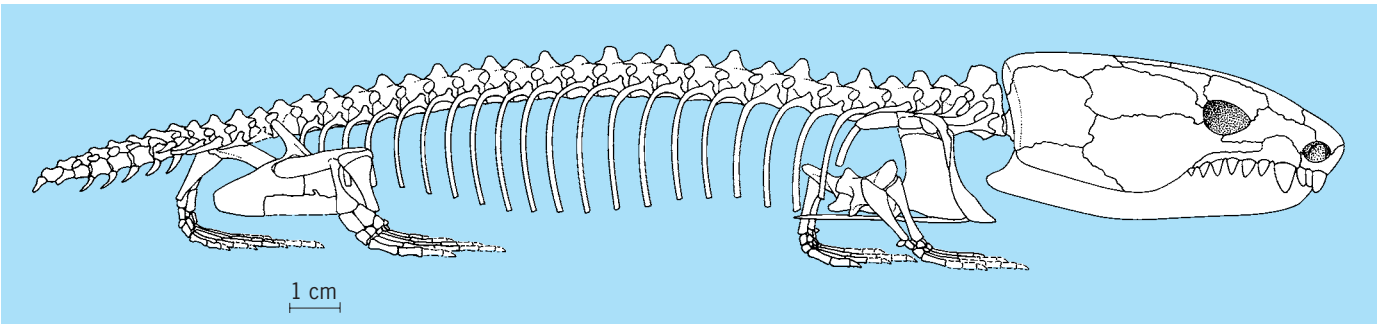

***Pantylus*, a microsaur from the lower Permian of Texas. (*After R. L. Carroll, The postcranial skeleton of the Permian microsaur Pantylus, Can. J. Zool., 46(6):1175–1192, 1968*)**

indicate close relationship. Although intermediate forms are not known, microsaurs appear to be the most probable group of Paleozoic amphibians from which two modern amphibian orders, the apodans and the salamanders, have evolved. *See* AMPHIBIA; LEPOSPONDYLI. Robert L. Carroll

Bibliography. R. L. Carroll and P. Gaskill, The order Microsauria, *Mem. Amer. Phil. Soc.*, 126:1–211, 1978.

Microscope

An instrument used to obtain an enlarged image of a small object. The image may be seen, photographed, or sensed by photocells or other receivers, depending upon the nature of the image and the use to be made of the information of the image. Microscopes are classified as simple or compound according to the kind of radiation used to form the image, and the use for which they are designed.

Simple microscope. A simple microscope, hand lens, or magnifier usually is a round piece of transparent material, ground thinner at the edge than at the center, which can form an enlarged image of a small object. Commonly, simple microscopes are double convex or planoconvex lenses, or systems of lenses acting together to form the image (**Fig. 1**). The lens can be mounted in a simple holder, in a folding case for hand use, or with a support which has a mechanical focusing mechanism, stage, and mirror to make a dissecting microscope. *See* LENS (OPTICS).

Compound microscope. The compound microscope utilizes two lenses or lens systems. One lens system forms an enlarged image of the object and the second magnifies the image formed by the first. The total magnification is then the product of the magnifications of both lens systems. Theoretically, several simple microscopes could be used in line, each to magnify the image of the one before it. Practically, the losses from aberrations, reflections, and other defects limit the compound microscope to two such systems (**Fig. 2**).

The typical compound microscope consists of a stand, a stage to hold the specimen, a movable body-tube containing the two lens systems, and mechanical controls for easy movement of the body and the specimen. The lens system nearest the specimen is called the objective; the one nearest the eye is called the eyepiece or ocular. A mirror is placed under the stage to reflect light into the instrument when the illumination is not built into the stand. For objectives of higher numerical aperture than 0.4, a condenser is provided under the stage to increase the illumination of the specimen. Various optical and mechanical attachments may be added to facilitate the analysis of the information in the doubly enlarged image.

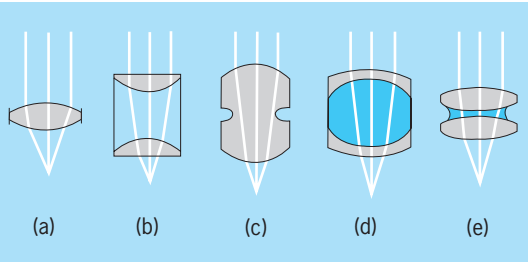

Fig. 1. Some common types of magnifier. (*a*) Double convex. (*b*) Doublet. (*c*) Coddington. (*d*) Hastings triplet. (*e*) Achromat. (*After F. A. Jenkins and H. E. White, Fundamentals of Optics, 3d ed., McGraw-Hill, 1957*)

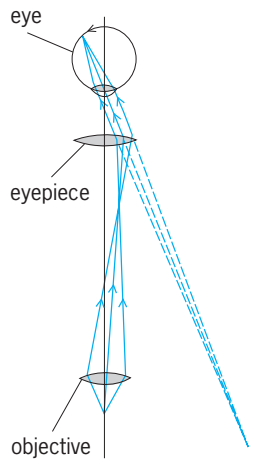

Fig. 2. Diagram of compound microscope. (*After F. A. Jenkins and H. E. White, Fundamentals of Optics, 3d ed., McGraw-Hill, 1957*)

Special compound microscopes have two image-forming systems to give an enlarged image of an image. These instruments utilize electrons, x-rays, sound, or other forms of radiation for image formation, and electromagnetic or electrostatic fields or mirrors to form the enlarged images. Because these images are not visible, photography, television, and special receivers must be used to record and analyze the image. Special microscopes are usually named according to the kind of radiation used, such as electron, x-ray, ion, and ultrasonic.

The compound microscopes which employ a lens system are essentially similar in their working principles. Basically, they are modifications of the ordinary laboratory microscope (bright-field microscope) for specialized or specific purposes. Following are descriptions of some of the different types.

Light or photon microscopes utilize light of wavelengths from 380 to 760 nanometers for image formation. Such microscopes include the laboratory or bright-field microscope, and modifications of it such as the capillary, centrifuge, chemical, comparison, crystallographic, dark-field, dissecting, fluorescence, integrating, interference, inverted microprojection, museum, nuclear track, petrographic, phase, phosphorescence, and profile microscopes.

Reflecting microscopes utilize a mirror rather than a lens system. The infrared microscope uses radiation of wavelengths greater than 700 nm, and the ultraviolet employs light of 180–400 nm. The ultraviolet microscope requires reflecting optics or special quartz and crystal objectives. Color-translating microscopes employ three different wavelengths of light to reveal details produced by ultraviolet or other nonvisible radiation.

In the electron, proton, x-ray, and beta-particle microscopes the image is usually recorded on a fluorescent screen or is photographed.

Mechanical vibrations, generated into an elastic system, provide the basis for the ultrasonic microscope employed for locating foreign bodies or for the analysis of reflecting surfaces. *See* ELECTRON MICROSCOPE; FLUORESCENCE MICROSCOPE; INTERFERENCE MICROSCOPE; OPTICAL MICROSCOPE; PHASE-CONTRAST MICROSCOPE; POLARIZED LIGHT MICROSCOPE; REFLECTING MICROSCOPE; SCANNING TUNNELING MICROSCOPE; X-RAY MICROSCOPE.

Oscar W. Richards

Microsensor

A very small sensor with physical dimensions in the submicrometer to millimeter range. A sensor is a device that converts a nonelectrical physical or chemical quantity, such as pressure, acceleration, temperature, or gas concentration, into an electrical signal. Sensors are an essential element in many measurement, process, and control systems, with countless applications in the automotive, aerospace, biomedical, telecommunications, environmental, agricultural, and other industries. The stimulus to miniaturize sensors lies in the enormous cost benefits that are gained by using semiconductor processing technology, and in the fact that microsensors are generally able to offer a better sensitivity, accuracy, dynamic range, and reliability, as well as lower power consumption, than their larger counterparts.

Technology and materials. Microsensors are typically batch fabricated from silicon wafers by using standard semiconductor process technologies in combination with specially developed processes. This technology, known as micromachining, allows hundreds of complex microsensors (or microstructures) to be produced on a single silicon wafer, resulting in a very low unit cost. This process technology is accurate and repeatable. A wide variety of passive and active materials are used to make microsensors. Single-crystal silicon is commonly used as a passive substrate material, although other substrate materials, such as gallium arsenide, quartz, and silicon carbide, are used for more specialized microsensor applications. These base materials are used in combination with a variety of thin films, including polysilicon, oxides, nitrides, and metals. It is often necessary to deposit an active material which is essential to the operation of the microsensor. Examples include zinc oxide, which is commonly used as the active material in piezoelectric sensors, and conducting polymers, which are commonly used in chemical microsensors. *See* INTEGRATED CIRCUITS; SEMICONDUCTOR.

Engineering a microsensor (sometimes called microengineering) requires the use of a number of micromachining processes, together with an understanding of micromechanics and microelectronics. The term bulk micromachining is often used for a set of processes, such as anisotropic etching, laser ablation, silicon-wafer bonding, and deep reactive ion etching, that enable the three-dimensional sculpting of single-crystal silicon to make a small structure, such as a diaphragm for a pressure sensor or a pressure switch, or a suspended mass structure for an accelerometer and flow sensor. Surface micromachining refers to a set of processes based upon deposition, patterning, and selective etching of thin films to form a free-standing microstructure on the surface of a silicon wafer. **Figure 1** shows a surface-micromachined microflexural resonator that has been integrated onto a thin silicon diaphragm to form a pressure sensor. The resonator is electrostatically driven into a lateral resonant mode of oscillation (around 50 kHz) by using a comb capacitor, and its motion is sensed by using another comb capacitor. The deflection of the diaphragm under an applied pressure stretches the microresonator, thus changing its spring rate and its fundamental resonant frequency. *See* PRESSURE TRANSDUCER.

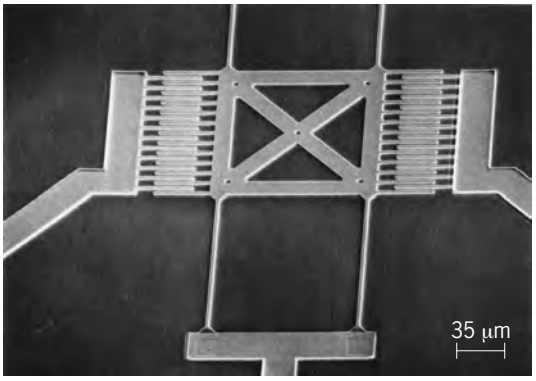

Fig. 1. Scanning electron micrograph of a polysilicon resonant structure made with a surface micromachining technique. ***(From C. J. Welham et al., A laterally driven micromachined resonant pressure sensor, Sensors and Actuators A, 52:86–91, 1996)***

Mechanical microsensors. Mechanical microsensors form perhaps the largest family of microsensors because of their widespread availability. Microsensors have been produced to measure a wide range of mechanical measurands, including force, pressure, displacement, acceleration, rotation, and mass flow. Force sensors generally use a sensing element that converts the applied force into the deformation of the elastic element. **Figure 2** shows a scanning electron micrograph of a simple microcantilever stylus structure that is used in an atomic force microscope for analyzing surfaces at an atomic scale of resolution. The tip of a flexible force-sensing cantilever stylus (with a tip curvature radius of a few nanometers) is scanned over the sample surface. The forces acting between the tip and sample cause minute deflections of the cantilever, which can be detected optically, capacitively, or with integrated piezoresistors to produce an atomistic surface image. *See* SCANNING TUNNELING MICROSCOPE.

Solid-state microsensors. Some microsensors can be made by using conventional bipolar or metal oxide semiconductor (MOS) technology, and a variety of radiation sensors are available that can detect visible and near-infrared radiation by using, for

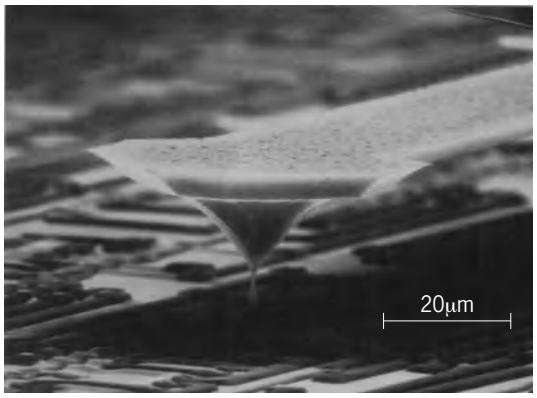

Fig. 2. Micrograph of a silicon microcantilever above an integrated circuit. ***(From J. Brugger et al., Silicon cantilevers and tips for scanning force microscopy, Sensors and Actuators A, 34:193–200, 1992)***

example, silicon *pn* diodes, avalanche photodiodes, and pyroelectric devices. Digital thermometers can be engineered by using a thermally sensitive device, such as a resistor, *pn*-junction diode, or transistor. A variety of magnetic microsensors, based upon Hall-plate devices (which are sensors that measure the magnetic field strength through the Hall effect), magnetoresistors, magnetodiodes, and magnetotransistors, are also available. *See* OPTICAL DETECTORS.

Chemical and biochemical microsensors. Two applications for chemical and biochemical microsensors are environmental monitoring and medicine. Both are relatively undeveloped, but should be of great importance in the future.

Unfortunately, chemical microsensors are the least well developed, for many reasons. For example, it has proved difficult to make a stable sensor that is sensitive to just the one gas of interest. One approach successfully applied to overcome this problem has been to employ an array of nonspecific gas sensors in a microprocessor-based instrument. The microprocessor is able to apply an appropriate pattern-recognition technique to extract the required information from the output of the microsensor array. This approach has opened up an enormous range of applications for gas and odor measurement in many different areas, such as the automotive, medical, and food–beverage industries and environmental monitoring. In addition, there have been advances in the design and synthesis of new materials such as conducting polymers, which coupled with the above approach make high-quality gas microsensors a distinct possibility in the near future. Microsensor array devices are also able to measure multiple measurands and compensate for interfering variables, such as temperature. For example, one chemical microsensor **(Fig. 3)** is employed as the sensing element in an electronic-nose instrument. *See* MICROPROCESSOR.

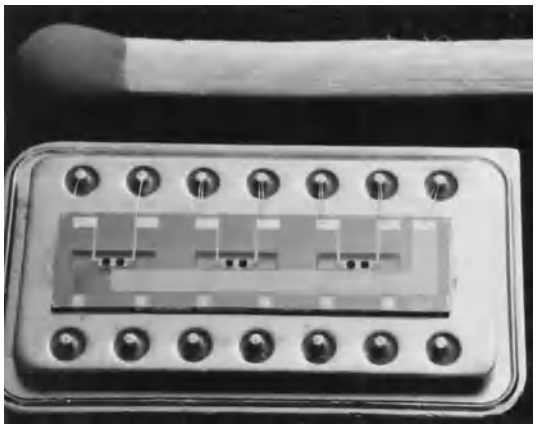

Fig. 3. Micromachined six-element chemical microsensor array in a dual in-line package. ***(From J. W. Gardner et al., Integrated sensor array for detecting organic solvents, Sensors and Actuators B, 26-27:135–139, 1995)***

Applications in the medical industry may involve monitoring blood, urine, and breath, which contain a wealth of information about the patient's state of health. Only a few such devices now exist. Examples include a glucose biochemical microsensor and ion-selective field-effect devices used to measure blood pH. The use of microsensors to gather medical diagnostic information is an attractive proposition, and eventually there may even be implanted microsensors to diagnose health problems, using smell-sensitive array devices. *See* BIOELECTRONICS.

Smart sensors. Since the processes used to fabricate microsensors are similar to the ones used to produce conventional integrated circuits, signal-conditioning electronics such as analog-to-digital converters can be integrated on the same chip, resulting in an integrated microelectromechanical component. Such a smart or intelligent sensor is able either to process information itself or to communicate with an embedded microprocessor. One notable example is a microaccelerometer developed for the automotive industry to control air-bag inflation. This inexpensive monolithic accelerometer comprises a surface-micromachined polysilicon sensor and bipolar–MOS interface circuitry on a single silicon chip. The sensor has been developed for a measurement range of $\pm 5\ g$ (where g is the acceleration of gravity, 32 ft/s^2 or 9.8 m/s^2), operates on a single 5-V supply, and has a power consumption of 40 mW. As the level of integration increases, a single-chip device incorporating microsensors, signal conditioning, and a microprocessor can be expected. *See* ACCELEROMETER.

Andrew C. Pike; Chris J. Welham; Julian W. Gardner

Bibliography. J. W. Gardner, Microsensations, *IEE Rev.*, pp. 185–188, September 1995; J. W. Gardner, *Microsensors: Principles and Applications*, 1994; E. Kress-Rogers (ed.), *Handbook of Biosensors and Electronic Noses*, 1997; S. Middlehoek and S. A. Audet, *Silicon Sensors*, 1989.

Microsporidea

A class of Cnidospora characterized by the production of minute spores with a single intrasporal or one or two intracapsular filaments and a single sporoplasm. The spore membrane of these protozoa is usually a single piece. Microsporidians are mainly

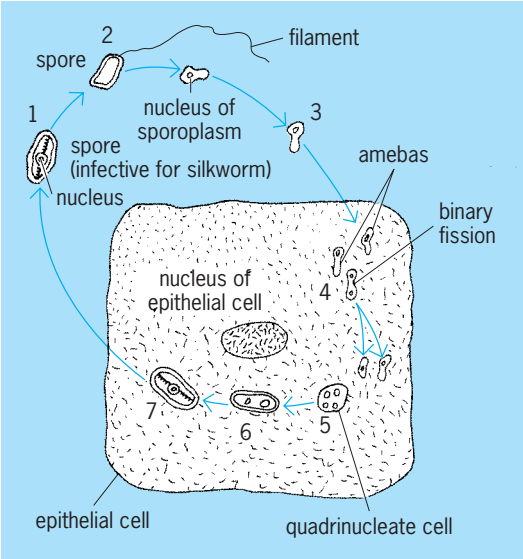

Diagrammatic life cycle of *Nosema bombycis*, cause of pébrine, a fatal disease of silkworms. (1) Spore is typical of the order, with a single filament and sporoplasm. Upon ingestion (2) the filament is extruded and (3) the sporoplasm enters an epithelial cell of the gut where it divides many times to form uninucleate amebas which fill the cell. The cell may dissolve. (4) Amebas attack other cells, perhaps almost all in the worm. (5) Eventually parasites form four-nucleate cells, which (6,7) develop into spore. Silkworms are infected by eating mulberry leaves or other food contaminated by feces or body fragments of infected worms.

intracellular parasites of arthropods and fishes. Microsporida is the only order of this class.

The sporoplasm or amebula, released from the spore in the intestine of the host, passes through the gut wall to reach the site of infection by way of the bloodstream. The amebula enters a cell and becomes a trophozoite, feeding and growing at the expense of the host. Eventually, the trophozoite divides by binary fission, or schizogony. The cells resulting from this asexual division develop into sporonts. A sporont may transform directly into a sporoblast and give rise to a single spore, or it may undergo further nuclear division and produce several sporoblasts, each of which will form a spore. In this order there is no evidence that special cells differentiate in the developing spore to form valves or a capsule as occurs in the Myxosporida and Actinomyxida.

An interesting feature of microsporidiosis is that the parasites may induce an extreme enlargement (hypertrophy) of the cell or nucleus. Thus, certain cells of the stickleback and smelt, infected with *Glugea anomala* and *G. hertwigi*, respectively, often enlarge from a diameter of 8–10 micrometers to one of 5000 micrometers. Such cells are called *Glugea* cysts.

The microsporidian *Nosema bombycis* is a parasite of the silkworm, the larval stage of the insect *Bombyx mori* (see **illus.**). Invasion by the parasite of all the cells of the larva, pupa, and adult causes the lethal disease pébrine which was first studied in detail by L. Pasteur in 1865. *See* CNIDOSPORA; INSECT PATHOLOGY; LEPIDOPTERA; PROTOZOA.

Ross F. Nigrelli

Microtechnique

The art of preparing objects for examination under the microscope and of preserving objects so prepared. Few objects yield useful information if examined without such preparation, which may involve, in addition to preliminary preservation, hardening, rendering transparent, selective coloration of parts, and cutting into thin slices.

The four types of microscope slide commonly made today are wholemounts, smears, squashes, and sections. The last three methods are merely devices to make thinner, or smaller, objects that are unsuitable for the first method. In all four methods, the objects are permanently preserved in a mounting medium between a glass slide, about 1 mm thick, and a glass cover slip 0.2–1 mm thick. The preliminary steps of fixation and preservation, the practice of staining, and the final mounting media are common to all four types.

Fixation and preservation. Biological specimens may be preserved in any environment, physical or chemical, which inhibits enzyme action and thus prevents either autolysis or microbiological growth. Objects thus simply preserved, however, rarely present to microscopists those details of external and internal anatomy which they desire to study. It is therefore customary to "fix" specimens before preservation so as to retain in them a reasonable facsimile of their appearance when alive. A perfect fixative solution would immobilize highly contractile organisms in a fully expanded condition, and at the same time preserve both the relations of the organs and the protoplasmic structure of the cells within them. The immobilization of highly contractile creatures requires instant denaturing of the proteins, while the preservation of protoplasmic structure is best retained by reasonably slow tanning. Actual fixative solutions, which by convention bear the name of the technician who first published them, are compromises (**Table 1**). Violent denaturants, which coagulate proteins, are mixed with less violent tanning agents which do not coagulate proteins. The best-known coagulants are heat, trinitrophenol (picric acid), chromium trioxide (chromic acid), mercuric chloride, and ethanol. Noncoagulating tanning agents in common use are potassium dichromate, osmium tetroxide (osmic acid), acetic acid, and formaldehyde.

No combination of these is perfect. Compositions should be selected not only for their immobilizing, coagulating, or tanning properties but also in regard to the relative importance of nuclear preservation, cytoplasmic preservation, and, in tissues intended for sectioning, the degree of brittleness produced. In general, very acid coagulants result in conventional images of nuclei and chromosomes, while weakly acid tanning agents display cytoplasmic detail and produce the least brittleness. The weakly alkaline fixatives of the electron microscopist produce images which have not proved acceptable to the optical microscopist.

Free fixative is generally removed from specimens

TABLE 1. Common fixatives

Mixture	Example	Penetration	Use
Osmic-chromic-acetic	Flemming's solution	Poor penetration; tissues should be fixed for about 6 h	Excellent for small animals and small pieces of tissue
Chromic-acetic-formaldehyde	Navashin's solution	—	Botanical fixative
Picric-acetic-formaldehyde	Bouin's solution; Allen's solution	Rapid penetration	Excellent immobilizer for contractive forms; produces little hardening
Dichromate-mercuric-acetic	Zenker's solution	Good penetration; results in brilliant staining	Most widely used histological and pathological fixative
Mercuric-acetic	Gilson's solution	Excellent	Excellent general-purpose fixative
	Carnoy and Lebrun's solution	Penetration is rapid	Used where penetration of hard objects is required, as seeds

by washing in water. Trinitrophenol-fixed material must be washed in 70% ethanol to avoid the removal of water-soluble albumen-trinitrophenol complexes. Alcoholic fixatives, such as Carnoy and Lebrun, are washed out in ethanol. After washing, specimens are conventionally preserved in 70% ethanol, although 4% formaldehyde may be equally well used for anything except trinitrophenol-fixed specimens.

Stains and staining. Wholemounts, smears, squashes, and sections are usually stained before being mounted. This process is still justified where it is desired to differentiate chromatically one part of the specimen from another. In many cases, however, this traditional procedure yields no more useful information than the examination of an unstained specimen by phase microscopy.

Dyes used in staining microscopical specimens are acid dyes, which color the nucleus (**Table 2**), and basic dyes, which color other cellular components. Each class contains dyes which attach directly (direct dyes) and those which attach to an intermediary (mordant) which is either applied before, or in the same solution as, the dye.

Orcein, safranin, methylene blue, and crystal violet are typically direct nuclear dyes, while carmine, hematoxylin, and celestin blue B are commonly used with mordants. Acid dyes will also attach to bacteria and cellulose. The various eosins, orange G, ponceau 2R, light green SF, and methyl blue are basic dyes. These give a general background stain when used as direct dyes, but some may be rendered specific to special cytoplasmic substances when used with mordants. Thus methyl blue, applied after a solution of phosphotungstic acid, will remain attached only to collagens.

TABLE 2. Examples of nuclear stains

Stain	Use
Direct nuclear stain	
Lillie's ammonium oxalate crystal violet	Stain bacterial smears
LaCour's acetocarmine	Squashed cells for chromosomes
Gray's celestin blue B	Sections and small objects
Indirect nuclear stain	
Grenacher's alcoholic borax carmine	Wholemounts of small plants and animals
Ehrlich's acid alum hematoxylin	Plant and animal sectioned tissues
Johansen's safranin	Sections of plant and animal tissues

Many basic and acidic dyes form compounds known as neutral dyes which may be used to stain both nucleus and cytoplasm from a single solution. The numerous compounds of various eosins with various thiazins, such as eosin Y–methylene blue, which is used to stain blood films, are in this category.

Basic dyes are commonly applied from simple solutions in water or weak ethanol. Acid dyes, however, usually require compounding into staining solutions, sometimes of considerable complexity and usually of an empirical composition established by tradition. All except orcein, and some oxazines, must be used indirectly; that is, the object is first impregnated with the dye solution and then differentiated, usually with a weak acid, until the dye is removed from the cytoplasm. After this treatment, hematoxylin, which is pink in acid, must be "blued" by treatment with a weak alkali.

Mounting media. A mounting medium should be a preservative. It should have an index of refraction sufficiently high to render objects transparent, and, preferably but not necessarily, it should cause the cover to adhere to the slide. Canada balsam, a natural exudate of *Abies balsamifera*, has all of these properties and is used for most slide preparation. Many proprietary substitutes have been proposed. Glychrogel may be substituted in those cases in which organic solvents would remove cell contents such as fat. Berlese's medium is excellent for small, hard-bodied specimens which are not to be stained. Glycerol may be used for similar purposes, but the edges of the cover slip must be sealed. Fant's medium, applied molten from a heated needle, is satisfactory for sealing glycerol mounts.

Preparation of mounts. Wholemounts, smears, squashes, and sections of tissues are preparations of plant and animal structures for specific purposes. Wholemounts allow the observer to study an entire organism or specific organ structure in some detail. More detailed study of the tissues or cells requires the preparation of sections. Such sectioned material is used in the study of normal and pathological material. Smears are regularly made for bacteriological and blood examinations, while a technique for

chromosome study is the employment of squashes, in which the material to be studied is literally flattened and squashed.

Wholemounts. Wholemounts of objects which can be preserved dry, such as seeds, moss archegonia, microfossils, and small beetles, may be attached with mucilage of tragacanth to a piece of black paper cemented to a slide. A low cylinder, punched from a sheet or cut from a tube, will serve to support the cover slip. Cylinder and cover may be cemented in place with any cellulose ester cement.

Small objects which must be rendered transparent but which need not be stained mount excellently in Berlese's medium. Thus, the majority of small arthropods may be placed, without preliminary fixation or preservation, in a drop of this medium and a cover slip applied.

Most wholemounts, however, are prepared as stained specimens in balsam. This involves the following sequence of operations: fixation, preservation, staining, dehydration, clearing, and mounting.

Fixation, in one of the fluids previously discussed, may have to be preceded by narcotization in the case of contractile invertebrates. Experience alone can determine the most successful narcotic to be used. All but the most delicate specimens, such as Rotatoria, medusae, and fresh-water Bryozoa, should be fixed in hot (50–70°F, or 10–21°C) fluids after narcotization. In many cases, hot fixation of *Hydra* and *Amoeba* without narcotization is quite adequate. The fixed, washed specimens are almost invariably preserved in 70% ethanol.

Wholemounts are best stained in a nuclear stain which differentiates internal organs, either by virtue of their density, which delays the extraction of the dye during differentiation, or because of the greater number of nuclei in small-celled organs. In any case, the specimen is taken from alcohol and placed in, for example, Grenacher's carmine overnight. The specimen is then rinsed in 0.1% hydrochloric in 70% ethanol and transferred to a fresh supply of the same reagent in which the process of differentiation may be watched. These transfers, and those subsequently required in the course of dehydration and clearing, are most readily made with the aid of cloth-bottomed tubes. These tubes are easily prepared by cutting the top from a test tube and tying bolting silk over the inverted rim.

Dehydration is necessary since water would prevent penetration of the balsam. The specimen is accordingly transferred from the acid alcohol to 95% ethanol and thence to pure ethanol. A half-inch-long (1.27-cm) leech will require 24 h in each. A fern prothallus will be perfectly dehydrated after 10 min.

Clearing the dehydrated specimen involves substituting some substance for the alcohol which is miscible with balsam. Wholemounts are usually cleared in essential oils or their synthetic equivalents. Clove oil and thyme oil are widely used, but terpineol renders specimens less brittle.

The specimen is now lifted from the clearing agent and placed in a drop of Canada balsam on a slide. A cover is then lowered vertically in place (**Fig. 1**).

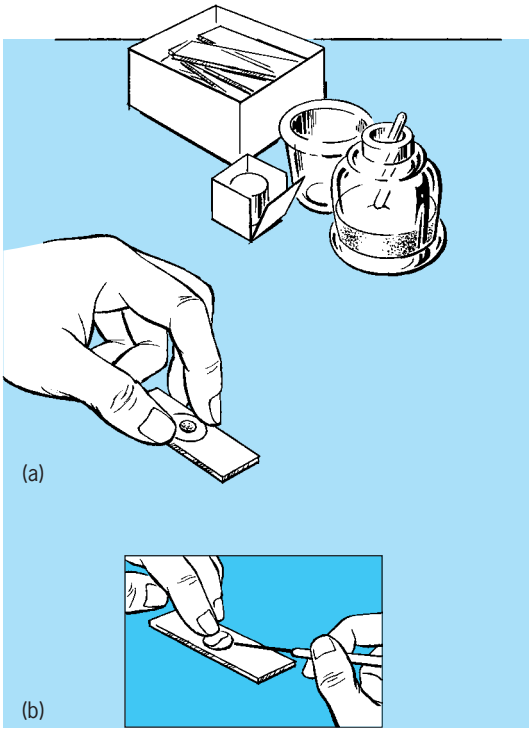

Fig. 1. Materials and method for applying cover slip to balsam wholemount. (*a*) Correct way; notice that the cover slip is held vertically centered over the drop of balsam. (*b*) Wrong way; this method draws the object to one side. (*After P. Gray, Handbook of Basic Microtechnique, 3d ed., McGraw-Hill, 1964*)

Smears and squashes. Smears are the most easily prepared microscope slides. A drop of organic fluid, such as blood, is placed on the end of a clean slide; a second slide is touched to it (**Fig. 2**) and pushed forward to leave a uniform film.

Blood smears of this type are usually stained in a methylene blue–eosinate dissolved in methanol, of which Wright's stain is a typical example. This stain also requires a phosphate buffer with pH of 6.4.

The air-dried smear is flooded with a specific amount of stain and laid on a rack, or across the top of a beaker, for 1 min. Buffer is then added to the stain from a drop bottle in the proportion of two drops of buffer to each drop of stain. After 2 min, the mixture is washed from the slide with distilled water. Blood smears are usually preserved in the dry state.

Squashes are just what their name indicates, and their success is dependent on the condition of the material selected. Anthers of plants, or the testes of insects, may be squashed directly, but root tips or plant ovaries require softening before squashing. Cellulases derived from the snail's stomach are much used for this preliminary.

The salivary glands of *Drosophila* are so widely prepared by this technique that they will serve as an example. Glands are taken from a third instar larva, easily recognized by the sluggish movements with which it crawls on the side of the culture vessel, by pulling them out with the head. This is best performed in a drop of LaCour's acetoorcein. Excess

tissue is rapidly stripped away, a cover slip placed on top, and strong pressure exerted with the forefinger through a sheet of bibulous paper. Inexperienced technicians may find it desirable to dissect out the glands in a drop of saline before transferring them to stain. A successful preparation shows sharply stained chromosomes against a faintly pink background.

These preparations may be rendered permanent. The slide is frozen for about 5 min on a block of dry ice and the cover slip is then gently levered off with a safety razor blade. The squashed material adheres to the slide, which is next placed in a coplin jar containing 70% ethanol before being dehydrated, cleared, and mounted in the same manner as a section.

Sections. All investigations of the structure of cells, or of the relations of cells within organs, require the examination of thin slices known biologically as sections.

Optical microscopy. The useful range of thickness for examination under the optical microscope is from about 25 to 2 micrometers; 10-μm sections are routine in most histological and pathological work.

Thick sections of crisp or waxy structures, particularly plant stems and roots, may be cut by hand. The razor is drawn across the plate with gentle pressure and the section is then washed into a stender dish. Such sections are subsequently stained and mounted, as though they were wholemounts, by the methods described above.

Most biological specimens, however, do not have the necessary consistency for such treatment but must be impregnated with, and embedded in, a substance which sections well. Paraffin is the most widely used. The preparation of paraffin sections requires the following steps: fixation, preservation, dehydration, clearing, impregnation, embedding, cutting, mounting on slides, dewaxing, rehydrating, staining, differentiating, counterstaining, dehydrating, clearing, and mounting.

Fixation, preservation, and dehydration of small pieces of tissue are the same as for the wholemounts described above. Clearing, however, is in this case preliminary to impregnation with molten wax, and xylol is the reagent of choice. Five-millimeter cubes of animal tissue should have at least 2 h in one change of 70% ethanol, two changes of 95% ethanol, two changes of pure ethanol, and two changes of xylol before being placed in molten wax. Plant tissues, in which the protoplasm tends to shrink from the cell wall, should be passed through the following sequence of reagents: (1) 5% ethanol, (2) 10% ethanol, (3) 18% ethanol, (4) 30% ethanol, (5) 40% ethanol plus 10% *tert*-butanol, (6) 50% ethanol plus 20% *tert*-butanol, (7) 50% ethanol plus 35% *tert*-butanol, (8) 40% ethanol plus 55% *tert*-butanol, (9) 25% ethanol, 75% *tert*-butanol, (10) pure *tert*-butanol. It should be understood that, for example, solution (7) of this series is prepared by mixing 50 parts of ethanol with 35 parts of butanol and then "making up" to 100 parts, by addition of water. Five-millimeter cubes of plant tissue will be adequately dehydrated

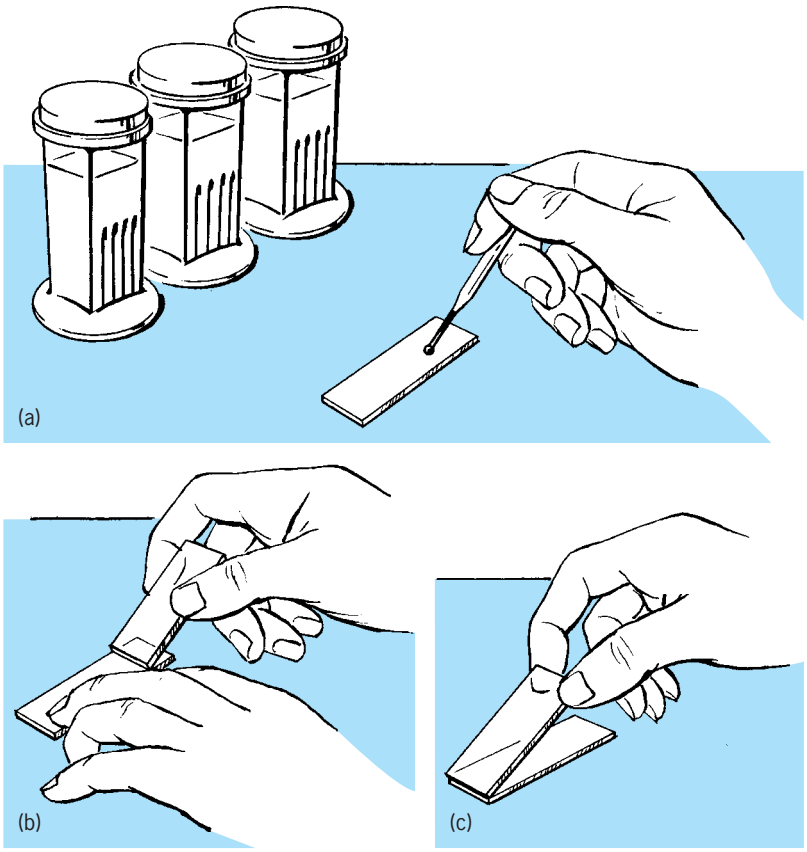

Fig. 2. Making a smear preparation. (*a*) Place the drop about 1 in. (2. 5 cm) from end of slide. (*b*) Apply a second slide just in front of the drop. (*c*) Push slide smoothly forward to spread the smear. (*After P. Gray, Handbook of Basic Microtechnique, 3d ed., McGraw-Hill, 1964*)

and cleared after 1 h in each of the solutions except (6), in which the pieces should remain for 6–12 h.

The tissues, whether cleared in xylol or butanol, must now be transferred to molten paraffin, maintained at just above its melting point in a thermostatically controlled oven or bath. Wax of 129–133°F (54–56°C) melting point, held at 136.4°F (58°C), is

Fig. 3. Rotary microtome produced by the American Optical Co. (*From P. Gray, Handbook of Basic Microtechnique, 3d ed., McGraw-Hill, 1964*)

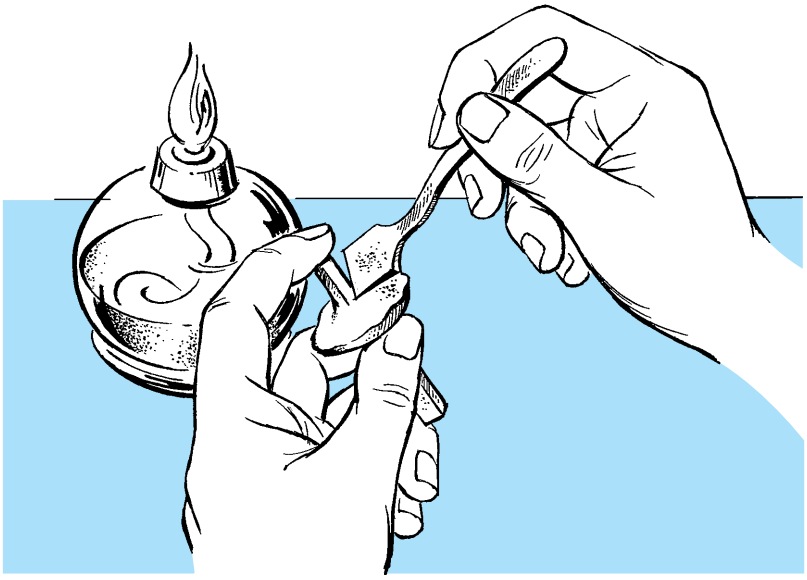

Fig. 4. Method of mounting the wax block on the block holder. (*After P. Gray, Handbook of Basic Microtechnique, 3d ed., McGraw-Hill, 1964*)

conventional. Softer wax, 126–129°F (52–54°C), is difficult to cut but should be used, at 131°F (55°C), for muscular tissues, such as heart and tongue, which tend to harden at 136.4°F (58°C). Harder waxes are rarely used. It is customary to add wax shavings to the tube containing delicate specimens in solvent, to leave this overnight, and then to place the tube in the oven. Pieces of homogeneous tissue like liver may be transferred directly to molten wax.

The impregnated pieces must now be embedded in a wax block. To this end paper boats, or other suitable containers, are filled with molten wax, the specimen is placed in the boat, a heated pipet is used to melt any film which may have formed around the specimen, and the block is then chilled and solidified by partial immersion in water. A satisfactory block is translucent and almost flat on top. A chalky block has either been cooled too slowly or made from solvent-contaminated wax. A block with a deep conical depression in the top has been cooled too fast.

Sections are cut from the block after it has been mounted on a microtome (**Fig. 3**). The revolution of the handle of this device raises and lowers the block vertically against a knife. The block is advanced horizontally at the top of its rise by the action of a micrometer screw advancing a pressure point along a spring-loaded diagonal plate. The extent of the advance, and thus the thickness of the section, is controlled by varying the moment of engagement of the pawl with the ratchet wheel attached to the micrometer screw.

The block of wax containing the object is removed from the container in which it was cast and trimmed until about equal volumes of specimen and wax remain. It is essential that two of the faces be parallel. The trimmed block is fused (**Fig. 4**) to the object carrier of the microtome. The carrier is placed in the holder of the microtome with the parallel faces parallel to the knife edge, adjusted until it almost touches the knife, and the handle turned until the ribbon of sections starts to form. The commonest defects are that the ribbon splits because of a nick in the knife edge or that the ribbons roll into cylinders without

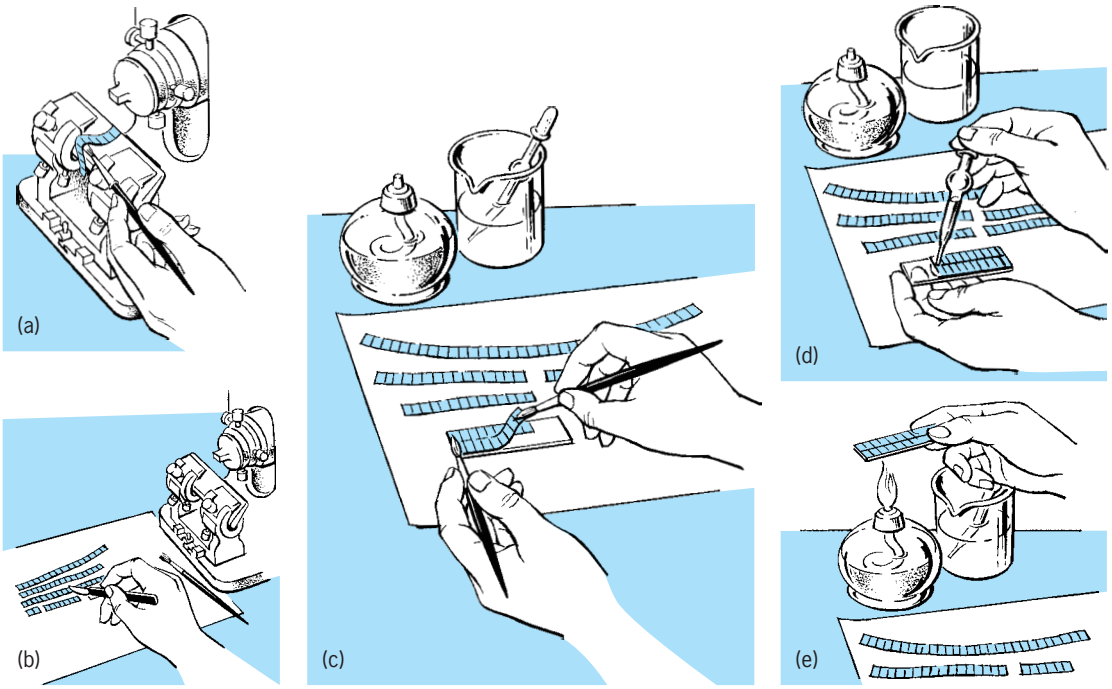

Fig. 5. Basic steps in cutting and mounting paraffin ribbons prior to staining. (*a*) Starting the ribbon. (*b*) Cutting the ribbon in lengths. (*c*) Mounting the dry ribbon. (*d*) Flooding the ribbons. (*e*) Warming the flooded ribbons to flatten them. (*After P. Gray, Handbook of Basic Microtechnique, 3d ed., McGraw-Hill, 1964*)

forming sections. The latter defect is due to a faulty relationship between the hardness of the block and the angle at which the knife strikes it. This relation must be established empirically for each block. The block lifting the ribbon from the knife blade sometimes results from a faulty angle and sometimes from a dirty knife.

When enough sections have been cut, the ribbon is divided (**Fig. 5**) into pieces about 2 in. (5 cm) long in preparation for mounting. A clean slide is lightly smeared with a mixture of equal parts of egg albumen and glycerin. The pieces of ribbon are laid on the prepared slide and flooded with water before being gently warmed until flat. The excess water is then drained off and the slide set on a warming table, at about 95°F (35°C), until dry.

The sections are now ready to be stained. To prepare the conventional celestin blue B–eosin histological slide, the following coplin jars (**Fig. 6**) should be set up: (1) xylol, (2) 50% xylol-ethanol, (3) pure ethanol, (4) 95% ethanol, (5) 70% ethanol, (6) water, (7) Gray's celestin blue B, (8) 0.2% ethyl eosin in 95% ethanol, placed behind the 95% ethanol jar. The slide is warmed until the wax melts, dropped into xylol for 1 min and then passed down the series, with about 30 s in each jar, until it is in water (6). It is then transferred to stain (7) for 1–2 min, rinsed in water, and transferred up the series again until it reaches 95% ethanol (4). The slide is now dipped up and down in the eosin solution (8) until the sections are sufficiently yellow—a point on which opinions vary widely. A quick rinse in 95% ethanol (4) precedes transfer to pure ethanol (3) for about 30 s and thence into xylol. The slide is then withdrawn from xylol, an adequate amount of balsam is placed on the surface, and a cover slip is applied.

Electron microscopy. The ultrathin (20–30 nm) sections used in electron microscopy are prepared by essentially similar methods but require different equipment and materials. Animal tissues are commonly fixed in Palade's fixative, a 1% solution of osmium tetroxide buffered to pH 7.4, but plant tissues are more usually fixed in Luft's fixative, which is a 1% solution of potassium permanganate buffered to the same pH. Both solutions differentially deposit electron-scattering materials in the tissue and thus function also as "stains."

The fixed and washed fragments, usually about 0.25-mm cubes, are dehydrated and "cleared" in acetone or, more rarely, ether before being placed in a plastic monomer which is subsequently catalyzed. Neither the hardness nor crystalline structure of wax permits ultrathin sections, but plastics have the necessary colloidal structure and their hardness can be adjusted at will. Mixtures of methyl methacrylate (very hard) and butyl methacrylate (very soft) were preferred in the past, but the heat generated in the course of polymerization has caused their abandonment in favor of epoxy resins, the hardness of which can be modified by the addition of plasticizers.

The "blocks" are cast in standard pharmacist's gelatin capsules and the end of the block trimmed to a tiny pyramid containing the specimen before mounting in the chuck of the ultramicrotome. This, as the regular microtome, moves the block up and down against a knife edge while advancing the block a given amount between each stroke. The advance is controlled either by a micrometer screw working against a motion-reducing lever or, in most modern instruments, by controlled thermal expansion. The very slight forward motion, coupled with a frequent irregularity of cut, necessitates either withdrawing the block slightly on the upward stroke or giving it a circular motion which avoids the knife on the upward stroke.

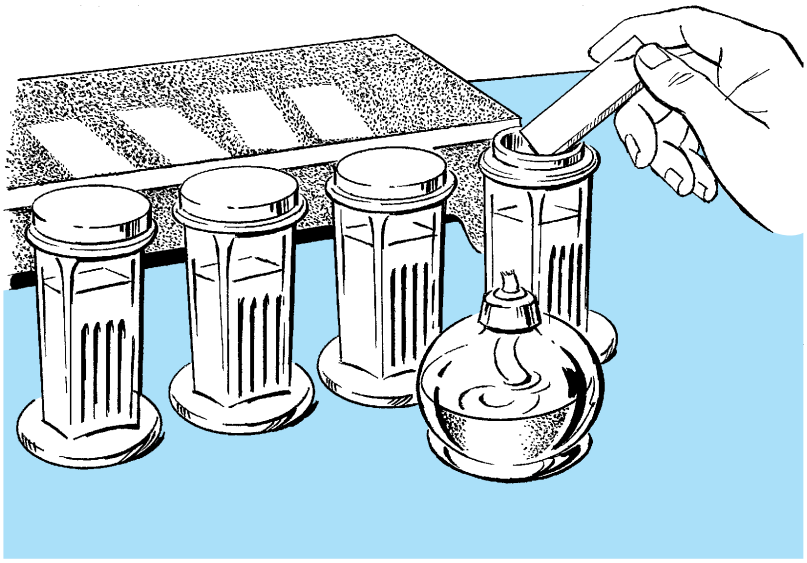

Fig. 6. Coplin jars for starting a slide through the reagent series. (*After P. Gray, Handbook of Basic Microtechnique, 3d ed., McGraw-Hill, 1964*)

The best knife is undoubtedly a freshly broken edge of a 45° wedge of plate glass, but specially ground diamond knives are used for extremely hard materials. The sections, sometimes "stained" in solutions of lead or uranium salts, are collected on wire mesh grids for examination under the electron microscope. *See* ELECTRON MICROSCOPE.

Finishing and storing slides. Mounts made with balsam should be hardened on a warm plate at about 113°F (45°C) for 3 or 4 days. Surplus exuded balsam may then be wiped off with a rag moistened with xylol, Another day of hardening will prepare the slide for a final cleaning in warm soapy water, after which the permanent label may be attached.

Wholemounts must be, and sections are better, stored in flat trays than in vertical grooves. Slides so stored, in a cool dark place, do not alter appreciably in 50 years.

Peter Gray

Bibliography. G. P. Berlyn and J. P. Mikshe, *Botanical Microtechniques and Cytochemistry*, 1976; W. Burrells, *Microscope Techniques*, 1980; P. Gray, *Handbook of Basic Microtechnique*, 3d ed., 1964; P. Gray, *The Microtomists' Formulary and Guide*, 1954, reprint 1975; J. K. Presnell et al., *Humason's Animal Tissue Techniques*, 5th ed., 1997.

Microwave

Electromagnetic energy with wavelengths in free space ranging roughly from 0.3 to 30 cm (**Fig. 1**). Corresponding frequencies range from 1 to 100 GHz. Frequency and wavelength are related by Eq. (1),

$$f\lambda = c \tag{1}$$

where f is the frequency, λ is the free-space wavelength, and c is the velocity of light in vacuum, approximately 3×10^8 m/s. The range shown is an arbitrary one, especially at the higher frequency end, where wavelengths are in the millimeter and submillimeter range and similar techniques are used. *See* ELECTROMAGNETIC RADIATION; FREQUENCY (WAVE MOTION); SUBMILLIMETER-WAVE TECHNOLOGY; WAVE MOTION; WAVELENGTH.

The first portions of the microwave spectrum to be developed were around 30 cm and 10 cm. These bands are designated L (for long) and S (for short), respectively. As other bands were developed, other letter designations were added (see **table**).

Characteristic transmission media for microwaves are hollow-pipe waveguides, where the cross-sectional dimensions are of the order of the wavelength and thus are of convenient size. Coaxial transmission lines are also used, however, especially in the lower-frequency bands, and various stripline techniques are used on microwave integrated circuits. Resonant cavities are commonly used as circuit elements, and radiation or reception of the energy is typically by horns, parabolic reflectors, or arrays. *See* ANTENNA (ELECTROMAGNETISM); CAVITY RESONATOR; COAXIAL CABLE; ELECTROMAGNETIC WAVE TRANSMISSION; TRANSMISSION LINES; WAVEGUIDE.

Generation. For most applications, microwaves are generated in electronic devices that produce oscillations at a frequency in one of the frequency bands (see table). The devices may be single-frequency or tunable, and continuous-wave (cw) or pulsed. Vacuum-tube generators include klystrons, magnetrons, and backward-wave oscillators; solid-state generators include tunnel diodes, Gunn diodes, IMPATT diodes, transistor oscillators, masers, and harmonic generators using varactor diodes. The vacuum-tube generators are used to produce higher powers, which can be as much as thousands of kilowatts. Solid-state generators were formerly limited in power to a few watts, but their power capabilities are continually increasing and now may reach hundreds of watts. *See* GYROTRON; KLYSTRON; MAGNETRON; MICROWAVE SOLID-STATE DEVICES; MICROWAVE TUBE; TRAVELING-WAVE TUBE.

For some applications, microwave noise is useful.

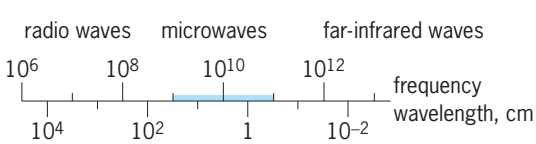

Fig. 1. Portion of the electromagnetic spectrum, with microwave range indicated.

Microwave frequency bands*

Microwave band	Frequency range, GHz
L	1.120–1.700
LS	1.700–2.600
S	2.600–3.950
C (G)	3.950–5.850
XN (J, XC)	5.850–8.200
XB (H, BL)	7.05–10.00
X	8.20–12.40
KU (P)	12.40–18.00
K	18.00–26.5
V (R, KA)	26.5–40
Q (V)	33–50
M (W)	50–75
E (Y)	60–90
F (N)	90–140
G (A)	140–220
R	220–325

*After T. K. Ishii, *Microwave Engineering*, Harcourt Brace Jovanovich, 1989.

This noise, generated in ordinary fluorescent light bulbs mounted in waveguides, is uniform across a wide band of frequencies, and the power emitted is known accurately and is quite stable with respect to time. Generators of this type are used in measuring the noise performance of receivers and, after amplification to high power, as a jamming signal for radar and communication systems. Microwave noise is also generated by current passed through a solid-state diode. *See* ELECTRICAL NOISE; ELECTRONIC WARFARE; JAMMING; MICROWAVE NOISE STANDARDS.

Circuit elements. Physical elements which produce specific effects on microwaves are called circuit elements. Microwave circuit elements bear little resemblance to circuit elements at lower frequencies because the shorter wavelength leads to an entirely different method of transmitting microwaves from place to place within a system (**Fig. 2**).

Waveguide. The most common method of microwave transmission within a system is through hollow circular or rectangular metal tubes of uniform cross section called waveguides. The microwave energy is confined within these tubes and guided along them.

In general, microwave energy is conveyed by an electromagnetic field whose components are the electric and magnetic intensities, denoted by E and H respectively. In a waveguide, only certain specific patterns of E and H, known as modes, can exist. The exact nature of the modes is determined by Maxwell's equations and the boundary conditions; the latter consist of the cross-sectional dimensions and the electromagnetic properties of the waveguide. Two types of mode can exist in an ideal hollow waveguide, the transverse electric (TE) and the transverse magnetic (TM) modes. *See* MAXWELL'S EQUATIONS.

The cross-sectional dimensions of a waveguide determine a frequency and a corresponding free-space wavelength below which no transmission is possible for each mode; this is the cutoff frequency or wavelength. The mode with the lowest cutoff frequency

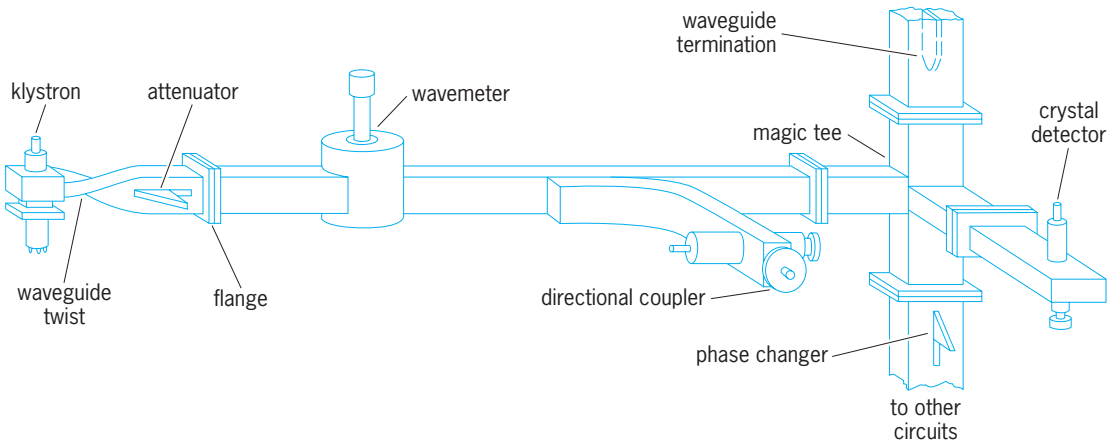

Fig. 2. Sample microwave circuit illustrating some microwave components.

is called the dominant mode; the other modes are called higher-order modes. In most applications, the waveguide dimensions are chosen so that transmission can occur only in the dominant mode, the other modes having cutoff frequencies well above the operating frequency (**Fig. 3**).

Most waveguides are rectangular in cross section. They are used to carry microwaves between other circuit elements and as integral parts of other circuit elements such as attenuators and phase shifters. The points at which microwaves enter or leave a circuit element are called ports; a length of waveguide has two ports.

Stripline. Various forms of stripline are used in interconnecting components on a dielectric or semiconductor substrate when microwave devices are integrated. One important example is the microstrip, in which a metallic strip or ribbon is placed on a thin dielectric, which is in turn backed by a conducting ground plane. Another configuration is the coplanar strip, in which two parallel strips, separated by a gap, are placed on the same surface of a dielectric. Still other variations are possible. These may be regarded as special cases of transmission lines, but since fields penetrate both the dielectric and air, they have dispersive properties; that is, different frequency components propagate with different velocities. Striplines have more losses than hollow-pipe waveguides but are generally used over very short distances. By proper configurations of the strip, typical functional transmission-line elements, such as impedance-matching sections, directional couplers, isolators, and filters, can be made.

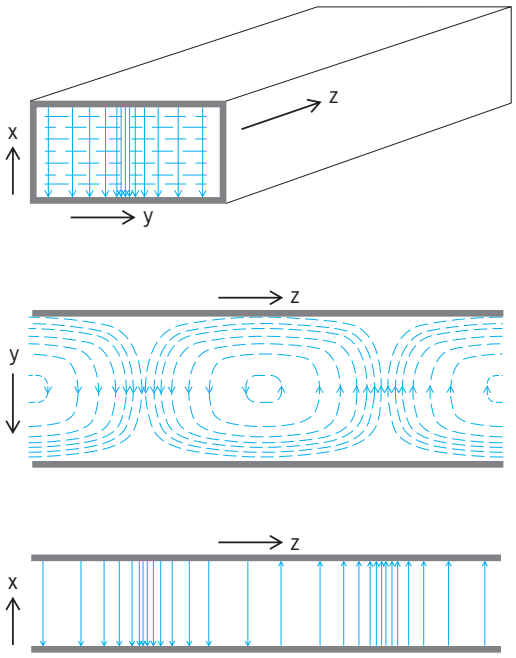

Fig. 3. Instantaneous field pattern for the TE_{01} wave, the dominant TE mode in a rectangular waveguide. The wave propagates in the *z* direction. Solid lines indicate the electronic intensity *E*, and broken lines the magnetic intensity *H*.

Microwave filter. Filters are needed in communication or information-processing systems for blocking of high frequencies (low-pass), blocking of low frequencies (high-pass), elimination of undesired bands (band elimination), or passing of desired bands while attenuating others (band pass). All of these may be made for microwaves by adding periodic perturbations, such as posts, irises, diaphragms, or dimensional variations, to the waveguide or other transmission system. Simple designs may be made by cascading quarter-wave sections of different propagation characteristics. More sophisticated designs use transformations from low-frequency, lumped-element designs. *See* ELECTRIC FILTER; MICROWAVE FILTER.

Attenuator, phase shifter, and termination. A thin sheet of plastic can be used to alter the amplitude or phase of microwaves. If the sheet is coated with powdered carbon with appropriate electrical conductivity and placed in a waveguide with the lossy material parallel to the lines of electrical intensity, it will absorb microwave power. Variable attenuation can be achieved by mechanically inserting more or less of the lossy strip into the path of the wave. *See* ATTENUATION (ELECTRICITY).

A phase shifter changes the phase of a microwave without changing its amplitude. It can be constructed in the same manner as an attenuator without the lossy material.

An attenuator which has a very large loss and is closed at one end is called a termination; it absorbs all the power transmitted into it, reflecting none.

Detector. The most common microwave detector is a silicon diode designed for high frequencies and mounted in a waveguide or a stripline. One terminal of the diode is connected to the waveguide, and the other is connected to a wire post, which is parallel to the direction of the electric field E and brought out of the waveguide through a small hole. The diode rectifies the microwave signal, producing an average current which can be indicated by a direct-current meter connected between the diode terminals. If the microwave signal is modulated in amplitude, the modulation will appear in the output current. *See* AMPLITUDE MODULATION; SEMICONDUCTOR DIODE.

The bolometer is a detector which absorbs microwave power, causing a temperature increase and a corresponding change in resistance. The bolometer does not respond fast enough to detect high-frequency modulation. It is often used as one arm of a resistance bridge circuit in microwave power meters. *See* BOLOMETER; MICROWAVE POWER MEASUREMENT.

Antenna. A transmitting antenna takes microwave power from a waveguide and converts it into a plane wave that propagates through space to a distant receiving antenna. Two important characteristics of antennas are efficiency and directivity, efficiency being the ratio of the power delivered into space to the power available in the waveguide. High directivity is accomplished by large antennas which focus the microwave energy in the same way a searchlight focuses a beam of light. In the parabolic antenna, a small waveguide feed horn (**Fig. 4**) illuminates a large paraboloidal dish which concentrates the energy in a narrow angular beam. In the horn-reflector antenna, a waveguide is expanded into a horn to illuminate a large paraboloidal surface (**Fig. 5**) which reflects the microwave energy in a narrow angular beam. Directivity is described by the beam width W of an antenna which is approximately given in radians by Eq. (2),

$$W = \frac{\lambda}{D} \tag{2}$$

where λ is the wavelength and D is the dimension of the antenna aperture. For circular antennas, D is the diameter. For rectangular antennas, D is either dimension, the beamwidth applying in the plane of the dimension. Directivity is also related to the gain of the antenna, which is the ratio of the power received at a distant receiver to the power that would be received if the transmitted power were radiated from an antenna with a uniform spherical pattern. Most microwave applications require highly directive antennas; gains of several thousand are not unusual, corresponding to beamwidths ranging from a fraction of a degree to several degrees. *See* DIRECTIVITY.

When receiving, an antenna intercepts power from the incident microwave proportional to the area of the antenna aperture. The gain and directivity of an antenna are the same whether it is used to transmit or receive. Efficiency of large antennas ranges from 50% to nearly 100%.

Gyrator, circulator, and isolator. The gyrator is a lossless, nonreciprocal, two-port circuit which has 180° more phase shift in one direction than in the other. This principle is used in the broadband microwave circulator. The three-port circulator has the property that all the power into port 1 exits at port 2, all the power into port 2 exits at port 3, and all the power into port 3 exits at port 1 (**Fig. 6**). The nonreciprocal phase shift is achieved in a magnetic ferrite placed in the waveguide junction and magnetized with an external permanent magnet. An isolator is a circulator with one port terminated, resulting in a circuit which transmits power in one direction and not the other. The input circuit is thus isolated from the output circuit. When the terminated port is internal, the

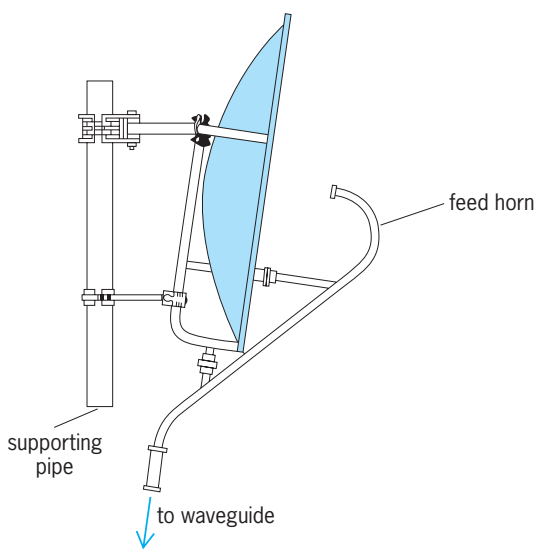

Fig. 4. Parabolic microwave antenna. (*After K. L. Dumas and L. G. Sands, Microwave Systems Planning, Hayden Book Co., 1967*)

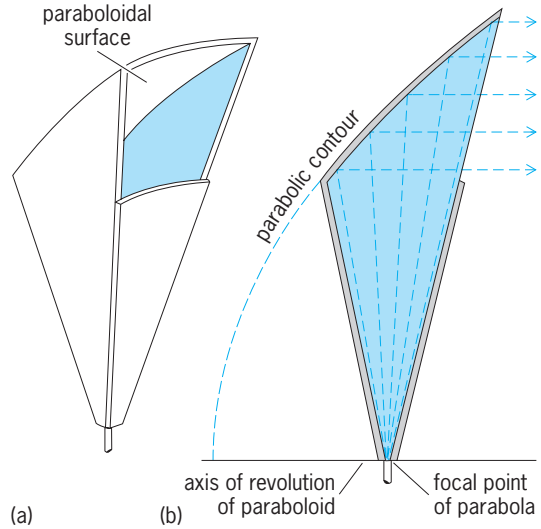

Fig. 5. Horn-reflector antenna. (*a*) Oblique view. (*b*) Cross section. (*After S. A. Schelkunoff and H. T. Friis, Antenna Theory and Practice, John Wiley and Sons, 1952*)

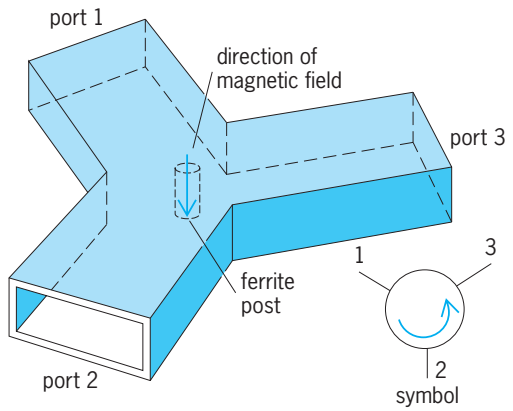

Fig. 6. Microwave circulator. (*After G. J. Wheeler, Introduction to Microwaves, Prentice-Hall, 1963*)

isolator appears to be a two-port element. *See* GYRATOR.

Varactor. This is a solid-state diode whose capacitance changes with applied voltage. Varactors are used as harmonic generators to obtain microwave power efficiently from lower-frequency sources such as quartz crystal–controlled oscillators in the 10–100-MHz range. Varactors are also used in up-converters and down-converters performing the same functions as similar circuits using resistive diodes, but more efficiently and at the expense of narrower bandwidth. *See* OSCILLATOR; VARACTOR.

Amplifier. A microwave amplifier converts a low-power input signal to a higher-power output signal while preserving one or more characteristics. A linear amplifier preserves the amplitude, frequency, and phase of the input signal. When a linear amplifier is overloaded, it becomes saturated: the output amplitude tends to remain constant. The envelope of the input signal becomes distorted, and additional frequency components are introduced in this nonlinear regime.

High power output is achieved with klystron and traveling-wave tube (TWT) amplifiers, both of which can be operated in the linear or saturated modes; the traveling-wave tube amplifier has the larger bandwidth, and the klystron has the higher power. Moderate power can be achieved with transistor amplifiers, and these are continually being extended in their frequency range of usefulness. Silicon bipolar transistors dominate at frequencies up to about 4 GHz, with gallium-arsenide metal-semiconductor field-effect transistors (MESFETs) dominant beyond this. In an injection-locked–oscillator amplifier, a high-amplitude oscillator is locked in phase and frequency to a low-amplitude input signal coupled into the oscillator circuit. The output of the oscillator is unaffected by the input signal; therefore only the phase and frequency of this type of amplifier are preserved. For example, the IMPATT diode can be operated as an injection-locked oscillator amplifier in the saturated mode. These are one-port amplifiers, the output power emerging from the input port, and a circulator is used to convert them into two-port amplifiers.

Very low noise levels are achieved in the maser and the parametric amplifier, both of which require power at a single frequency to pump the active element. In the maser, energy is pumped from one atomic level to another, and in the parametric amplifier a varactor is driven by the pumping signal to provide the nonlinear capacitance required for amplification. *See* AMPLIFIER; MASER; PARAMETRIC AMPLIFIER.

Microwave integrated circuits. Microwave integrated circuits (MICs) are of two types, hybrid MICs and monolithic microwave integrated circuits (MMICs). In the hybrid circuits, some or all of the active and passive devices are added to a dielectric or semiconducting substrate and interconnected by striplines as discussed above. For MMICs, all of the active and passive devices of the functional unit are formed on the substrate, and interconnected by striplines through a variety of microfabrication techniques, including photolithography, epitaxy, ion implantation, etching, diffusion, sputtering, and evaporation. Substrate materials include insulating materials such as alumina or sapphire or semiconducting materials such as silicon or gallium arsenide (GaAs). Conducting materials are gold, silver, copper, or aluminum, and dielectric films are such materials as silicon monoxide (SiO), silicon dioxide (SiO_2), or silicon nitride (Si_3N_4).

MMIC technology promises small size, improved reliability, reproducibility and low cost if produced in volume. MMIC switches using gallium-arsenide field-effect transistor devices have given switching speeds of less than 2 nanoseconds with a bias of only 350 microwatts. An entire X-band microwave receiver including local oscillator, mixer, low-noise amplifier, and intermediate-frequency amplifier has been formed on a single chip of dimension 0.25 × 0.25 in. (6 × 6 mm). *See* INTEGRATED CIRCUITS.

Polarization. An important property of a microwave is its polarization, the direction of its electric field E with respect to a fixed plane. A microwave transmitted parallel to the Earth is said to be vertically polarized if E is perpendicular to the Earth's surface. Two superimposed waves of the same frequency and magnitude, and in both time and phase quadrature, result in a circularly polarized wave in which the electric field direction rotates either clockwise or counterclockwise. *See* POLARIZATION OF WAVES.

Circuit elements can be constructed whose effect is polarization-sensitive. For example, closely spaced, vertical wires in a plane normal to the direction of a wave will completely reflect a vertically polarized wave and pass, without modification, a horizontally polarized wave.

Microwave receiver. The most simple microwave receiver is a silicon diode detector followed by an amplifier; amplitude modulation on the input signal is detected directly and amplified to a suitable level. This receiver is insensitive and noisy and cannot be used with frequency or phase modulation. *See* FREQUENCY MODULATION; MODULATION; PHASE MODULATION.

The first active element in nearly all microwave receivers is a silicon diode operated as a down-converter. In this type of receiver, a strong, continuous-wave local oscillator signal is used to pump the diode over its nonlinear resistance range. In this manner, the local oscillator and the input signal are mixed, shifting the input signal down to an intermediate frequency, which is the difference between the frequencies of the local oscillator signal and the received signal. Intermediate frequencies of a few tens of megahertz are common. Frequency, phase, or amplitude modulation on the received signal appears in the detector output at the intermediate frequency. A bandpass intermediate-frequency amplifier, providing most of the gain of the receiver, follows the detector, after which a demodulator converts the modulation on the intermediate-frequency signal to usable form, for example, an audio or a television signal. *See* RADIO RECEIVER.

If extremely low-noise performance is required, the first active component is a maser or a parametric amplifier. These are not required in most applications.

Microwave transmitter. The main components of a microwave transmitter are a microwave power source, a modulator, and, if necessary, a power amplifier. Among the vacuum-tube power sources, the reflex klystron and the backward-wave oscillator are frequency-modulated, and the magnetron is pulse amplitude-modulated, by a signal on an electrode. In solid-state transmitters, by contrast, the modulator is usually separate from the continuous-wave power source. The modulation can be done directly at microwave frequencies or it can be performed at intermediate frequency and shifted to the microwave frequency in an up-converter, which is very much like a down-converter. A common frequency modulator consists of an intermediate-frequency oscillator tuned by the modulation signal applied to a varactor. *See* PULSE MODULATION.

The power amplifier can be a klystron or traveling-wave tube when higher powers are required. For solid-state transmitters, transistor or injection-locked-oscillator amplifiers are common.

Microwave propagation. In free space, microwaves travel in straight lines as do optical waves. Near the Earth, however, the atmosphere has an index of refraction which normally decreases with distance above the Earth and causes the wave to travel in a circular path which bends slightly toward the Earth. The radius of the circular path is larger than the Earth's radius, so normally a wave traveling above the surface of the Earth and parallel to it will not intercept the Earth. Microwaves are reflected and refracted by objects just as are optical waves. *See* MICROWAVE OPTICS.

Occasionally during the summer, atmospheric conditions cause microwaves transmitted from an antenna to travel to a receiver via two or more paths. These waves interfere at the receiver and may cause large decreases in the received signal amplitude. This phenomenon, called multipath fading, is a serious problem in microwave transmission parallel to the surface of the Earth. Other atmospheric conditions cause the microwave to bend away from the Earth and miss the receiver. This problem is solved by placing the transmitting and receiving antennas on tall towers; the wave initially heads toward the Earth, bends away before intercepting the Earth, and reaches the receiver if it is high enough. If the towers are not sufficiently high, the wave is intercepted by the Earth, resulting in what is known as earth-bulge fading. When microwaves are directed well above the horizon, neither of these two problems occurs. Thus satellite microwave systems do not suffer from either multipath or earth-bulge fading.

At frequencies above about 10 GHz, rain absorbs microwave energy, resulting in large signal losses. Both satellite and point-to-point microwave systems are seriously affected by rain attenuation.

For most frequencies the attenuation of microwaves by the Earth's atmosphere is very small. There are, however, bands of frequencies for which the loss is higher due to molecular absorption. Attenuation due to water vapor occurs in several bands, the first one appearing at about 30 GHz. A larger attenuation, due to oxygen absorption, occurs at 60 GHz. *See* RADIO-WAVE PROPAGATION.

Applications of microwaves. Areas in which microwave radiation is applied include radar, communications, radiometry, medicine, physics, chemistry, and cooking food.

Radar. This is an acronym for radio detection and ranging. In one form a pulse of electromagnetic energy is transmitted in a narrow beam toward a target, and part of the energy is reflected to the receiver. The time difference between the transmitted and received pulses provides the distance to the target, and the direction of the antenna beam provides the target's direction. Radar is used in military applications, commercial aviation, remote sensing of the atmosphere, and astronomy. The high antenna directivity and the excellent propagation characteristics of microwaves in the atmosphere make this the preferred band for radar applications. *See* RADAR.

Electronic countermeasures. Microwaves are also used in electronic countermeasures to radar. The most direct counter is the broadcast of a high-intensity microwave signal, usually with noise modulation, to mask the echoes from the potential target. More sophisticated systems detect the transmitted radar signal, modify it, and send it back in such a way that false information is given concerning the position and direction of the target. *See* ELECTRONIC WARFARE.

Communications. There is at least 100 times as much frequency space available for communications in the microwave band as in the entire spectrum below microwaves. In addition, the high directivity obtainable at microwave frequencies allows reuse of these frequencies many times in the same area, practice not possible at lower frequencies. The high directivity also makes possible communication to satellites and deep-space probes. *See* RADIO SPECTRUM ALLOCATION.

Satellites may act as relay links between ground stations a great distance apart or as direct

broadcast sources to a distribution of ground receivers. Most satellites operate at frequencies between 4 and 12 GHz, although higher frequency bands are available. Light weight, long life, high efficiency of transmitters, and low noise of receivers are of obvious importance for microwave components for satellite use. *See* COMMUNICATIONS SATELLITE.

Communication between the Earth and deep-space probes is accomplished with microwaves because of the great antenna directivity and receiver sensitivity that can be achieved at microwave frequencies. Microwaves are also used in cellular communications, where systems for mobile communications are well established. Higher-frequency systems with cell radii of 0.6 mi (1 km) or less are promising for personal communications. *See* MOBILE COMMUNICATIONS; SPACE COMMUNICATIONS; SPACECRAFT GROUND INSTRUMENTATION.

Radiometry. All objects, including liquids and gases, emit electromagnetic radiation in the form of noise, the amount of the noise being proportional to the absolute temperature of the object. A noise temperature T can be assigned to the object corresponding to the amount of noise radiating from it. The determination of the noise temperatures of selected noise sources, including background radiation from outer space, rain, clouds, stars, and the Earth, is called radiometry. *See* COSMIC BACKGROUND RADIATION; RADIO ASTRONOMY.

A microwave radiometer is a sensitive receiver which measures the noise power received by an antenna; from this measurement, the noise temperature of the source object can be determined. The noise powers measured are so small that receivers with very high gain are required, and the accuracy of the measurements is severely limited by small time variations in the gain. R. H. Dicke solved this problem in 1946 by switching the receiver input between the antenna and a reference noise source at a rate faster than the gain changes in the amplifier. The detected output of the receiver is also switched synchronously at the same rate resulting in two output signals, one due to the noise received by the antenna and the other to the reference noise source. These two signals are subtracted to form an output signal proportional to the difference in the input noise temperatures. The reference noise power can then be adjusted until the output signal is zero, indicating the received noise power is equal to the known reference noise power.

Radiometers are used extensively for remote sensing. Atmospheric temperature, water vapor density over oceans, soil moisture content, and the liquid water content of clouds can be determined from radiometer measurements, some of which are made from satellites. Microwave radiometers are used to study astronomical sources of noise and to observe planets from deep space probes. *See* PASSIVE RADAR; RADIOMETRY; REMOTE SENSING.

Physiological effects. Exposure to microwave power of 100 mW/cm^2 for several minutes can lead to pathophysiologic effects in laboratory animals. The microwaves penetrate beneath the skin and heat the tissue, and tissue destruction can result if the temperature rise is faster than the control mechanisms of the body can handle. It is assumed that human tissue reacts to microwave exposure in the same manner, and safety standards have been proposed. In the United States the maximum recommended exposure is 10 mW/cm^2, although stricter standards are proposed.

Applications of microwaves in medicine include (1) thermography, the measurement of tissue temperature (cancer causes a temperature rise of about 2°F or 1°C, which can be detected with a microwave radiometer); (2) hyperthermia, microwave heating used in the treatment of cancer and in the treatment of hypothermic subjects; and (3) biomedical imaging, the use of microwaves to study the structure of tissue beneath the skin. *See* RADIOLOGY.

Physics and chemistry. Microwave energy is used in large particle accelerators to accelerate charged particles such as electrons and protons to very high energies and cause them to collide. Knowledge of the structure of matter is also obtained from microwave spectroscopy, which is used to study the frequencies and amplitudes of the microwave resonances of molecules. Microwaves are also used in the study of the crystal structure. *See* MICROWAVE SPECTROSCOPY; PARTICLE ACCELERATOR.

Microwave oven. Microwave energy is absorbed in most foods and has been found to be a source of quick, uniform heating or cooking. Microwave ovens based upon this principle are now widely used. The microwave power, supplied by a magnetron, is at a frequency of 2.45 GHz, with rated power from 350 to 750 W. Since oven size is generally much larger than the 12 cm wavelength, standing waves could form, resulting in heating nonuniformities. This problem is avoided by using either a rotating mechanism that shifts or effectively stirs the modes, or a moving platform to shift the food through the standing-wave pattern.

Safety is, of course, the primary concern. Interlocks are provided on all ovens so that the power is off when the door is open, and a backup system should the primary interlock system fail. Leakage levels of microwave energy from the units must meet strict requirements, and all safety features must be operable for at least 100,000 times.

Industrial heating and drying. Microwaves are also used for the industrial heating of foodstuffs and other materials. The advantages are in the rapidity of heating, the accuracy of control, the possibility of selectivity and localization, the clean environment, and the absence of combustion products. For the food industry, cooking, tempering of frozen foods, and drying of pastas and other foodstuffs are examples. Microwaves are aslo used in the pharmaceutical, paper, fabric, rubber, and ceramic industries. The primary frequency ranges in these uses are at 0.9 and 2.45 GHz. *See* FOOD ENGINEERING.

Beyond microwaves. Applications of electromagnetic radiation in the millimeter-wave region, which extends from about 100 to about 300 GHz, are much like those in the microwave region. Beyond

this range, and throughout the long infrared wavelengths, suitable power for many of these applications is not available.

At the shorter infrared wavelengths and in the visible optical range, the light-emitting diode (LED) and the laser provide the power, and solid-state detectors provide the necessary efficiency, for communication and other applications. Optical communication (including the infrared) is feasible through the atmosphere but is seriously affected by fog, haze, rain, and snow. The advent of the optical fiber has greatly accelerated the application of these wavelengths to communications. Fibers are made from glass, are as small as a human hair, and can be made so that they have very low loss, a few tenths of decibel per kilometer. Transmission through fibers is unaffected by atmospheric conditions. The communication capacity of an optical-fiber system can be very large: many billions of pulses per second can be transmitted over a single fiber. Microwaves are used in the modulation of optical-fiber systems. Optical-fiber systems have been built for use inside buildings, between cities, and across oceans. *See* JUNCTION DIODE; LASER; LIGHT-EMITTING DIODE; OPTICAL COMMUNICATIONS; OPTICAL DETECTORS; OPTICAL FIBERS.

Clyde L. Ruthroff; J. R. Whinnery

Bibliography. T. K. Ishii (ed.), *Handbook of Microwave Technology*, 2 vols., 1995; T. K. Ishii, *Microwave Engineering*, 2d ed., 1989; G. M. Kizer, *Microwave Communication*, 1990; Y. Konishi (ed.), *Microwave Integrated Circuits*, 1991; S. Liao, *Microwave Devices and Circuits*, 3d ed., 1996; A. D. Olver, *Microwave and Optical Transmission*, 1992; B. L. Smith and M. H. Carpentier (eds.), *Microwave Engineering Handbook*, vols. 1–3, 1993; J. Thuery, *Microwaves: Industrial, Scientific and Medical Applications*, 1992.

Microwave filter

A two-port component used to provide frequency selectivity in satellite and mobile communications, radar, electronic warfare, metrology, and remote-sensing systems operating at microwave frequencies (1 GHz and above). Four types of filter responses are commonly used in such systems (**Fig. 1**): low-pass, high-pass, band-pass, and band-stop. For example, a low-pass filter allows the passage of signals at frequencies below the cutoff frequency, f_c (the passband), while reflecting signals at higher frequencies (the stopband). While an ideal filter would have zero loss in its passband and infinite attenuation in its stopband, such characteristics can only be approximated for realistic filters. Microwave filters perform the same function as electric filters at lower frequencies, but differ in their implementation because circuit dimensions are on the order of the electrical wavelength at microwave frequencies. Thus, in the microwave regime, distributed circuit elements such as transmission lines must be used in place of the lumped-element inductors and capacitors used at lower frequencies. This can make microwave filter design more difficult, but it also introduces a variety of useful coupling and transmission effects that are not possible at lower frequencies. In addition, the use of distributed elements often introduces spurious stopband regions of low attenuation at frequencies far above the passband.

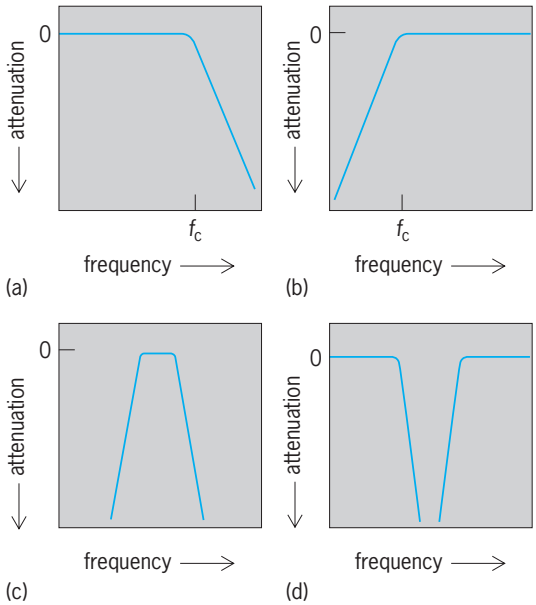

Fig. 1. Common microwave filter attenuation (insertion loss) versus frequency responses. (*a*) Low-pass filter. (*b*) High-pass filter. (*c*) Band-pass filter. (*d*) Band-stop filter. Attenuation or insertion loss increases down the vertical axis.

Design. The majority of modern microwave filters are designed by using the insertion-loss method, whereby the amplitude response of the filter is approximated by using network synthesis techniques that have been extended to accommodate microwave distributed circuit elements. This method allows great flexibility in specifying filter characteristics, including the passband and stopband amplitude responses, the attenuation rate, and the phase response. Commonly used passband responses include maximally flat (binomial), equal-ripple (Chebyshev), and elliptic responses. Other important microwave-filter performance measures are the insertion loss (the passband attenuation, in decibels), the stopband attenuation rate (in decibels per frequency decade), the group delay (in nanoseconds), and the passband impedance match (the input return loss, in decibels). System considerations may include the power capacity or temperature stability of the filter, as well as weight, size, and cost. The order of a filter is a measure of its complexity, and is the number of capacitors and inductors in its low-pass filter prototype; increasing the order of a filter increases the rate of attenuation in the stopband. *See* IMPEDANCE MATCHING.

Microwave filter design by the insertion-loss method follows a general four-step procedure: determination of filter specifications, design of a low-pass prototype filter, scaling and transforming the filter, and implementation (conversion of lumped elements to distributed elements). After a realistic set

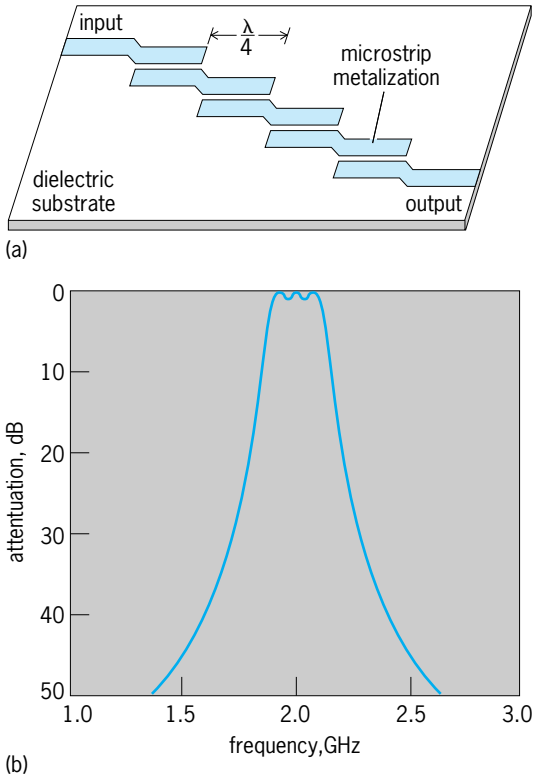

Fig. 2. Third-order (four-section) microwave band-pass filter. (*a*) Layout of parallel coupled stripline filter. (*b*) Calculated frequency response of filter, with center frequency of 2 GHz and 0.5-dB equal-ripple passband.

of filter specifications (such as response type, cutoff frequency, impedance, and order) have been determined, a low-pass prototype filter is designed. This is a normalized filter design using lumped inductors and capacitors, and having a unity cutoff frequency and impedance. A wide variety of design tables for low-pass prototype filters of various orders and responses are available. The low-pass prototype can then easily be transformed to a high-pass, band-pass, or band-stop response, and scaled to the desired cutoff or center frequency. The last step, filter implementation, is what makes microwave filter design a unique and challenging process. There are a wide variety of different techniques for constructing microwave filters, usually specific to the type of transmission-line media being used (such as coaxial line, waveguide, microstrip, or coupled line), but some common principles apply. The Richards transform allows lumped-element capacitors to be replaced with open-circuited transmission-line stubs, and lumped-element inductors to be replaced with short-circuited stubs. The Kuroda identities allow the introduction of transmission-line lengths to provide spacing between filter elements. Use of these techniques is called commensurate line filter synthesis, and works well for low-pass and high-pass microwave filters. An additional filter implementation tool is the impedance or admittance inverter, which can be used to transform series resonators to shunt resonators (and vice versa) and to transform internal impedance levels.

Types. Microwave filters are implemented in many different ways depending on considerations such as cost, size and weight, integration with the rest of the system, and electrical performance issues such as insertion loss, phase characteristics, and stopband attenuation. Waveguide cavity band-pass filters have very low insertion loss, making them preferred for frequency multiplexing in satellite communication systems. Coaxial low-pass filters, made with sections of coaxial line with varying diameters, are compact and inexpensive. Planar filters in microstrip or stripline form (**Fig. 2**) are important for integration with hybrid or monolithic microwave integrated circuits, and can be made by using transmission-line stubs, transmission-line resonators, or coupled transmission lines. Many compact planar microwave filter designs are made in interdigital, comb-line, or hairpin-line form, consisting of multiple sections of parallel coupled transmission lines. While planar filters are usually more cost-effective than waveguide versions, their insertion loss is usually greater. Dielectric resonators can also be used with planar microwave-circuit technology. Computer-aided design procedures are used in the synthesis of more sophisticated amplitude and phase responses, and active microwave devices (field-effect transistors) are used to provide filters with gain or tunable response characteristics. *See* COAXIAL CABLE; COMPUTER-AIDED DESIGN AND MANUFACTURING; ELECTRIC FILTER; MICROWAVE; MICROWAVE SOLID-STATE DEVICES; TRANSMISSION LINES; WAVEGUIDE.

David M. Pozar

Bibliography. I. J. Bahl and P. Bhartia, *Microwave Solid State Circuit Design*, 1988; G. L. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-Matching Networks, and Coupling Structures*, 1980; D. M. Pozar, *Microwave Engineering*, 2d ed., 1997.

Microwave free-field standards

The means for setting up electromagnetic fields of precisely determined intensity at microwave frequencies in unbounded regions of space. Such standards are used to evaluate field probes and antennas for measuring field strength and power density. The standardization of these devices is necessary before they are used for determining the performance of radar and communications systems or for assessing such systems for health and safety risks or electromagnetic compatibility. *See* ELECTROMAGNETIC RADIATION; MICROWAVE.

Near fields and far fields. An electromagnetic field can be launched into an unbounded region by applying an oscillating voltage or current to a radiating structure or antenna. The energy radiating from an antenna passes through several distinct stages before the final pattern emerges. The true near field comprises induction and static fields which decrease in magnitude as the inverse-squared and inverse-cubed of distance R respectively; they thus become negligibly small a few wavelengths from the antenna and

are rarely significant in practice. The radiated field decreases in intensity as the inverse of distance and thus becomes predominant beyond a few wavelengths. *See* ANTENNA (ELECTROMAGNETISM).

Radiating antennas carry alternating currents which can be thought of as an array of monochromatic sources spread over the structure according to the Huygens-Fresnel principle, and these sources will give rise to elemental wavefronts which arrive at a point distant from the antenna with phases which impose an interference-type pattern on the wavefront, those at the extremities of the structure having the greatest phase differences. The complex wavefront is obtained by the superposition of these elementary wavelets. The distance at which a 90° phase shift obtains on the elemental wavefronts is known as the Rayleigh distance, and is equal to $D^2/2\lambda$ where D is the maximum aperture dimension and λ the wavelength of the radiation. The region of the field within this distance is the radiated near field, and because of its complex nature it is commonplace free-field measurement practice to use $2D^2/\lambda$ as a minimum working distance for antenna measurements. However, even at this distance the path phase differences are 22.5° and the power density, being proportional to $1/R^2$, is accurate only to 5%. For 1% uncertainty the required separation is at least $10D^2/\lambda$. *See* HUYGENS' PRINCIPLE; INTERFERENCE OF WAVES; RESOLVING POWER (OPTICS).

Standard antennas. Antennas used in free-field standards are usually either half-wave dipoles or waveguide horns.

Half-wave dipole. The half-wave dipole is a collinear device with a length of approximately one-half of the free-space wavelength of the radiated wave. The radiation pattern is doughnut-shaped and symmetrical about the axis of the dipole. The gain of a simple dipole can be shown to be 2.16 dB.

Waveguide horns. An open-ended waveguide will radiate effectively and, as the aperture dimensions are comparable with the wavelength, its pattern is very broad; an open-ended waveguide has a power gain of about 6 dB. Flaring the aperture decreases the beamwidth and results in a pyramidal horn. Horns can be made with gains ranging from about 8 to about 30 dB, depending on frequency and acceptable size. The gain and pattern characteristics of a pyramidal horn can be calculated to accuracies of about ±0.2 dB; however, the reflections at the throat and aperture discontinuities significantly influence the gain-frequency characteristic and calibration is usually necessary. *See* WAVEGUIDE.

Free-field measuring sites. The ideal environment for making measurements on antennas is a large unobstructed volume which is free of reflecting objects and electromagnetically interfering signals—that is, a free-space condition.

Anechoic chambers. A practical solution is to use an anechoic chamber to set up simulated free-space conditions in a bounded environment. Low reflection of electromagnetic signals from the walls of such a chamber is achieved by the use of an electromagnetic-wave absorbent layer covering all of the reflecting surfaces within the room or chamber, the outer shell of which is a metallic structure to give shielding against interfering signals encroaching on the test region.

Anechoic chambers can be of rectangular or tapered form. For rectangular chambers the width of the chamber is ideally chosen so that the angle of incidence for the reflected ray from the side wall is of the order of 70° or less. The receiving antenna is placed about half the width of the chamber from the back wall so that coupling to the back wall absorbers is minimal. These considerations suggest an optimum length-to-width (and height) ratio of about 3:1. The "tapered" anechoic chamber consists of a pyramidal tapered anechoic section joined to a cubical anechoic section and is an alternative construction suitable for some antenna measurements. The tapered section may be thought of as a large horn antenna terminating in a large waveguide in which a single mode in the form of a plane wave is to be generated.

The size of field strength monitor or antenna which can be measured in an anechoic chamber is determined by the size, and shape, of its quiet zone, which is the volume in which the reflections from the internal surfaces of the chamber are below a specified acceptable signal level compared to the direct radiated value of the field. The measure of the quiet-zone performance, the reflectivity, will vary with position in the zone as well with frequency. For antenna measurements using pseudo-far-field methods the quiet zone needs to be as long as possible with respect to the operating wavelength; this is in order for the separation between receiving and transmitting antennas to be at least equal to several Rayleigh distances at the lowest operating frequency. For normal pattern measurements a reflectivity level of −40 dB in the quiet zone is acceptable; for high-precision gain measurements, with uncertainties of 0.1 dB, −50 dB reflectivity is necessary.

For near-field scanning applications, in which the sampling or detecting probe scans within a few wavelengths of the test antenna, interfering reflections and resonances from the environment which may contaminate the test region can be controlled with a fairly modest application of electromagnetic absorber materials. *See* ANECHOIC CHAMBER.

Open field sites. Clear open sites can be used at low frequencies, or for high-gain antennas where far-field measurements require such large distances that enclosed or anechoic environments are impractical. The horizontal plane can be readily made free of obstruction, as can the upward vertical plane. However, since electromagnetic waves are strongly reflected by a ground plane, this effect must be accommodated, and one of two courses can be adopted. One is to do measurements as far above the ground as possible. The alternative is to make use of the ground plane by working close to it—if necessary, enhancing its reflection with a metal ground plane or grid—and by making allowance for the reflection in the analysis of performance.

Measurements on antennas. Probably the single most important parameter of a standard antenna, when considered for metrological applications, is the boresight or maximum gain, and a number of methods of determining this parameter have been devised.

Three-antenna method. The basis of this method is a measurement of the transmission between two polarization-matched antennas (1 and 2), which is expressible in terms of the antenna gains by the Friis formula (1), where P_t and P_r are the transmit-

$$P_r = P_t G_1 G_2 \left(\frac{\lambda}{4\pi D}\right)^2 \qquad (1)$$

ted and received powers, λ the wavelength, and D the separation. Only the product of the gains can be obtained; however, with three antennas the measured combinations will yield the gain of each antenna uniquely. Small corrections are required due to imperfect matching in the waveguide systems.

Extrapolation method. The extrapolation technique is probably the most accurate method known for determining absolute gain and polarization. The method is one of determining the transmission characteristic between two antennas as the transmission path is increased through about 4 to 10 Rayleigh distances. With good metrology it is possible to get a sufficiently accurate characterization to allow extrapolation to the true far-field range.

Near-field scanning. Antenna metrology has increasingly concentrated on near-field techniques with the objective of improving antenna characterization, particularly at very high (VHF) and ultrahigh (UHF) frequencies where the lack of low-reflectivity far-field ranges has been acute. In this method a probe antenna is used to sample the magnitude and phase, for orthogonal polarizations, of the radiated fields over a well-defined surface, which can be a plane, a cylinder, or a sphere, a few wavelengths from the antenna under test. The objective of the near-field to far-field technique is to mathematically transform the data determined in the near field into the required far-field properties. For a large antenna the number of samples required to define the near field adequately can be large enough to require substantial measurement and computational time. The features of the technique are that complete characteristics can be obtained for any distance, near-field or far-field, for the test antenna and that measurements may be made indoors as close as desired to the antenna in a simple anechoic environment.

Calibration. The essential requirement for the calibration of devices for measuring power flux density or field strength is the creation of a substantially plane wave of known power density which encompasses the effective aperture of the device to be tested. This is effected by launching a known power through an antenna or transverse-electromagnetic (TEM) cell of known characteristics and calculating the field strength or power density from the appropriate equation. The techniques used fall into two groups which are conveniently divided by frequency.

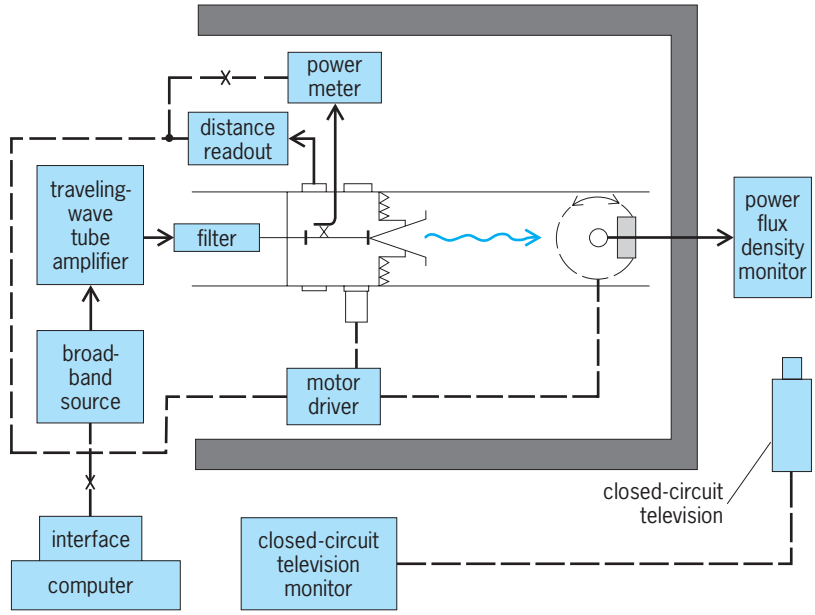

Fig. 1. Microwave free-field calibration system.

Frequencies above 500 MHz. At frequencies above a nominal 500 MHz the conventional technique for calibrating power density or field monitors is to set up a standard field in an anechoic space and position the probe at some prescribed position in front of the waveguide horn.

A practical system for setting up standard fields or power densities involves a small anechoic chamber (**Fig. 1**) about 15 ft (5 m) long, lined on five sides with electromagnetic absorber. One end is left open for access to the generating and measuring systems. The microwave signal is generated by a broadband signal generator operating into a traveling-wave-tube power amplifier. The output of the latter is filtered to remove spurious harmonics and linked to the monitoring coupler and power meter combination and output (horn) antenna by a low-loss coaxial cable. The instrument under test is mounted on a low-loss dielectric pedestal which is positioned about 3 ft (1 m) from the end wall of the chamber and on the axis of the radiating antenna, which places the test probe in the quiet zone. A computer can be used to control the measurement process and apply corrections. The uncertainties associated with this measurement are about 0.3 dB (equivalent to about 7%). *See* SIGNAL GENERATOR; TRAVELING-WAVE TUBE.

Frequencies below 500 MHz. The Crawford transverse-electromagnetic transmission cell is commonly used for setting up radio-frequency fields. This consists of a transmission line with a rectangular outer conductor, 1.5–3 ft (0.5–1 m) in size, and an inner conductor in the form of a flat plate or septum (**Fig. 2**). The input and output ends are tapered so as to form a transition to a conventional coaxial cable. The wave within the cell is essentially plane and can be made to have free-space impedance by proper choice of width and height dimensions. The field distribution within the cell, and thus its uniformity, can be calculated. The electric field within the working area

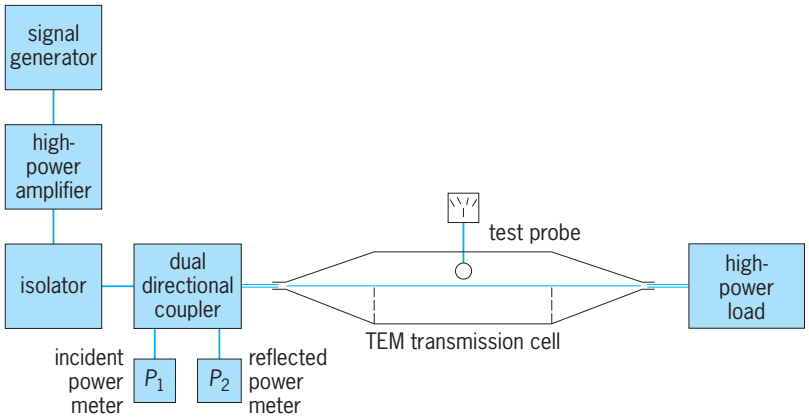

Fig. 2. Guided-wave test facility based on Crawford transverse-electromagnetic (TEM) cell.

is given approximately by dividing the voltage between the plates by the separation of the plates. The voltage is determined by calculation from the power P transmitted through the cell divided by the line impedance Z (equal to 50 ohms). The power density is then expressed by Eq. (2), where Z_0 is the

$$\text{Power density} = \frac{PZ_0}{b^2 Z} \qquad \text{W/m}^2 \qquad (2)$$

impedance of free space (376.7 ohms) and b is the separation of the plates. The uncertainties associated with transverse-electromagnetic cell measurements are of the order of 0.5 to 1 dB. *See* COAXIAL CABLE.

The upper limit to the operating frequency range of transverse-electromagnetic cells is determined by resonances within the cell; also, it is not good practice to more than half fill the gap between the plates. As a consequence, the larger the test object, the larger must be the cell, and thus the lower its maximum frequency. This problem can be reduced to some extent by placing electromagnetic absorber material in the cell, but this expedient renders the cell suitable for comparison purposes only.

Ralph W. Yell

Bibliography. G. E. Evans, *Antenna Measurement Techniques*, 1990; Institute of Electrical and Electronics Engineers, *IEEE Standard Test Procedures for Antennas*, ANSI IEEE Std. 149-1979, 1979, reaffirmed 1990; J. D. Kraus, *Antennas*, 2d ed., 1988; Y. T. Lo and S. W. Lee (eds.), *The Antenna Handbook*, 4 vols., 1993; A. W. Rudge et al., *The Handbook of Antenna Design*, 2 vols., paper 1986.

Microwave impedance measurement

The determination of parameters, associated with microwave propagation in transmission lines or waveguides, which are generalizations of the impedance concept at lower frequencies and are derived from ratios of complex electric or magnetic field amplitudes.

As an outgrowth of popular usage the term impedance measurement, at microwave frequencies, has acquired a more general interpretation than the one associated with it at lower frequencies. Throughout the lower portion of the radio-frequency spectrum, impedance is generally defined as the (complex) ratio between voltage and current. Although the low-frequency concepts of voltage and current can be generalized in such a way as to be useful in the microwave region, the almost universal practice is to use the scattering description. Instead of voltage and current, the basic parameters are now the complex electric- or magnetic-field traveling-wave amplitudes in the (assumed) uniform transmission line or waveguide by which the components of interest are interconnected. The coaxial transmission line and rectangular waveguide (where the center conductor is absent) are common examples. *See* COAXIAL CABLE; TRANSMISSION LINES; WAVEGUIDE.

For a given transmission line or waveguide, there is a preferred band of operating frequencies throughout which the solution of Maxwell's equations is given by a pair of complex numbers a and b which represent the electric or magnetic traveling-wave amplitudes to the left and right (**Fig. 1**). Let a voltage v and current i be defined by Eqs. (1) and (2), where Z_0

$$v = b + a \qquad (1)$$

$$i = \frac{b - a}{Z_0} \qquad (2)$$

is a property of the transmission line, called characteristic impedance. These definitions of voltage and current are consistent with the ones in common use at lower frequencies. Obviously, Eqs. (1) and (2) can be solved for a and b in terms of v and i so that the information content is the same in either description; the mode of expression, however, is different. *See* MAXWELL'S EQUATIONS.

As a counterpart to the lower-frequency definition of impedance (the complex ratio of v to i), a reflection coefficient (which is usually represented by the Greek letter Γ) can now be defined as the complex ratio given by Eq. (3). The measurement

$$\Gamma = \frac{a}{b} \qquad (3)$$

of reflection coefficient is usually included if not primarily implied by the term impedance measurement at microwave frequencies. In its more restricted definition, the impedance Z is given by the ratio of

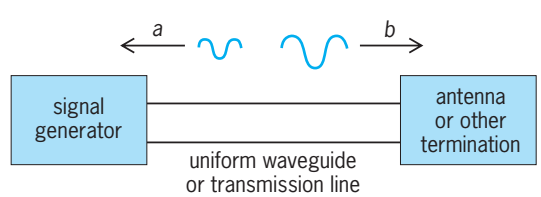

Fig. 1. Microwave system, with complex traveling-wave amplitudes *a* and *b* that are generally used to describe the electric and magnetic fields.

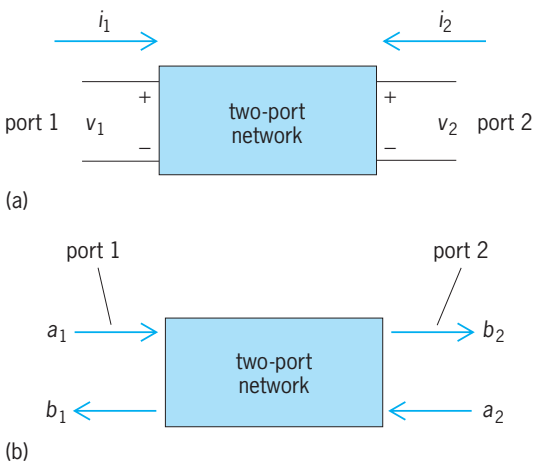

Fig. 2. Two-port network showing (*a*) the voltages and currents used in *y*-parameter definitions, and (*b*) the incident and reflected complex-wave amplitudes used in scattering-parameter definitions.

Eqs. (1) and (2). From here it is easy to show that Eqs. (4) and (5) are valid.

$$\Gamma = \frac{Z - Z_0}{Z + Z_0} \tag{4}$$

$$Z = Z_0 \frac{1 + \Gamma}{1 - \Gamma} \tag{5}$$

Scattering parameters. Linear electrical circuits can be completely characterized by parameters measured at the input and output ports. Several different parameter sets can be used. At low frequencies, h, y, and z parameters are frequently encountered; for example, a two-port network (**Fig. 2***a*) can be characterized by y parameters by using Eqs. (6) and (7),

$$i_1 = y_{11}v_1 + y_{12}v_2 \tag{6}$$

$$i_2 = y_{21}v_1 + y_{22}v_2 \tag{7}$$

where v_1 and v_2 are the voltages across the input and output ports respectively and i_1 and i_2 denote the currents entering the network. *See* NETWORK THEORY.

To measure h, y, or z parameters, it is necessary to excite the two ports in turn with an open-circuit or short-circuit connected to the other port. At microwavelengths (0.3–30 cm) this measurement is difficult to carry out for several reasons: (1) Lead inductance makes it difficult to get a good short circuit. (2) Stray capacitance makes it difficult to create a good open circuit. (3) Microwave amplifiers sometimes oscillate when their ports are open-circuited or short-circuited. (4) Voltages and currents are rarely measured at microwavelengths, but equipment that yields reflection and transmission coefficients is frequently employed.

For all of these reasons, scattering parameters, rather than h, y, and z parameters, are widely used to characterize microwave networks. Let a_1 and a_2 (Fig. 2*b*) represent the complex wave amplitudes entering ports 1 and 2 respectively, and let b_1 and b_2 denote the complex wave amplitudes emerging from ports 1 and 2 respectively. The relationships between these complex wave amplitudes are given by Eqs. (8) and (9).

$$b_1 = s_{11}a_1 + s_{12}a_2 \tag{8}$$

$$b_2 = s_{21}a_1 + s_{22}a_2 \tag{9}$$

Interpretation. When a generator is connected to port 1 and a nonreflecting (that is, perfectly matched) termination is connected to port 2, $a_2 = 0$, and from Eq. (8) it follows that Eq. (10) is satisfied. Thus, s_{11}

$$s_{11} = \frac{b_1}{a_1} \tag{10}$$

is the voltage reflection coefficient looking into port 1 when port 2 is perfectly matched. Under the same conditions, it is seen from Eq. (9) that Eq. (11) is

$$s_{21} = \frac{b_2}{a_1} \tag{11}$$

valid. Thus, s_{21} is the ratio of the complex-wave amplitude emerging from port 2 to the complex-wave amplitude incident on port 1 when port 2 is perfectly matched. After interchanging the generator and matched termination, expressions for s_{22} and s_{12} are found in a similar way.

Matrix notation. The scattering coefficients are often written in matrix notation. For a two-port network, the scattering matrix has the general form shown in Eq. (12).

$$[S] = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} \tag{12}$$

For a device with m ports, the scattering matrix has m rows and m columns. For the nth port, some workers let $|a_n|$ and $|b_n|$ represent the root-mean-square values of the incident and reflected voltages; but the expressions for reciprocity and losslessness with unequal port characteristic impedances become more simple if $|a_n|^2$ and $|b_n|^2$ represent the incident and reflected powers respectively. Using this latter convention, the following rules apply:

1. In a linear reciprocal two-port network, Eq. (13) is valid.

$$s_{12} = s_{21} \tag{13}$$

2. In a linear lossless two-port network, Eqs. (14)–(16) are satisfied, where s_{11}^* denotes the complex

$$s_{11}s_{11}^* + s_{21}s_{21}^* = 1 \tag{14}$$

$$s_{12}s_{12}^* + s_{22}s_{22}^* = 1 \tag{15}$$

$$s_{11}s_{12}^* + s_{21}s_{22}^* = 0 \tag{16}$$

conjugate of s_{11}, and so forth. Equations (14)–(16) can be readily derived by equating the input and output powers, using Eqs. (8) and (9) and remembering that $|a_1|^2 = a_1\, a_1^*$, and so forth. In mathematical language, S is a symmetric, unitary matrix when the two-port network is linear, reciprocal, and lossless. *See* MATRIX THEORY.

These rules remain valid when the two ports have different characteristic impedances. There are corresponding rules for multiport networks. The scattering matrices for many two-, three-, and four-port networks are given in standard works and are widely used when analyzing microwave systems.

The scattering matrix for a perfectly matched lossless line of length x is given by Eq. (17), where

$$[S_x] = \begin{pmatrix} 0 & e^{-j\beta x} \\ e^{-j\beta x} & 0 \end{pmatrix} \tag{17}$$

$\beta = 2\pi/\lambda$ is the phase constant of the line and λ denotes the wavelength in it. The rules stated above are all seen to be satisfied by this matrix. Frank L. Warner

Slotted line. Among the most conceptually simple and oldest methods of measuring microwave impedance is the slotted line. In its basic form, a longitudinal slot is cut in the transmission line of interest, which permits the fields to be sampled by means of a movable probe at different positions along the axis of the line. A mechanical transport mechanism is provided so that the depth of penetration or coupling between the probe and internal fields remains fixed as the probe moves along the slot.

At any instant of time, the phase of a or b depends upon the phase of the generator and the propagation time delay between the generator and the probe position where a or b is measured. As the probe is moved along the line, this time delay will be increased for a and decreased for b, or vice versa, depending upon the position of the generator and the direction of the motion. Thus the phase difference between a and b will depend upon the probe position.

In the existing practice, the designs of most probes are such that their response is proportional only to the total field amplitude which is given by $|a + b|$ or $|v|$ [according to Eq. (1)]. Since motion along the line is accompanied by changes in the phase difference between a and b, the magnitude of the sum changes as required by the laws governing the addition of two complex numbers. In particular, if a and b are in phase, the maximum probe response is obtained, and this is proportional to $|a| + |b|$. At the minimum, the response is proportional to $||a| - |b||$. By definition, the ratio between these two responses is the voltage-standing-wave ratio (VSWR) and is given by Eq. (18). (The reciprocal of this definition is also in use.) The measurement of VSWR also falls within the scope of "impedance measurement" at microwave frequencies. Ordinarily, the directions of wave propagation associated with a and b are chosen such that $|a| \le |b|$. Equations (3) and (18) may now be combined to obtain Eqs. (19) and (20).

$$\text{VSWR} = \frac{|a| + |b|}{||a| - |b||} \tag{18}$$

$$\text{VSWR} = \frac{1 + |\Gamma|}{1 - |\Gamma|} \tag{19}$$

$$|\Gamma| = \frac{\text{VSWR} - 1}{\text{VSWR} + 1} \tag{20}$$

The foregoing procedure determines the magnitude of Γ but not its phase. In many applications, however, the interest in the phase of Γ is minimal. The reason is that the generation of energy is far more difficult at microwave than at the lower frequencies. As a consequence, the interface between different components is usually arranged so as to achieve maximum power transfer. As a general rule, this calls for wave propagation in one direction only. Loosely speaking, the magnitude of Γ usually represents an indication of the extent of departure from the system design objectives.

On the other hand, it is possible to obtain the phase by a rather simple extension of the foregoing technique. While the phases of a and b change with respect to line position, their magnitudes and thus $|\Gamma|$ remain constant. At the position of the minimum probe output, the phase of Γ is 180° since this corresponds to a minimum value of the sum $|a + b|$. To find the phase of Γ at any other position in the line, it is only necessary to note that since the phases of a and b move in opposite directions, the phase of Γ changes at twice the rate of either of them and undergoes a total of 360° between two successive minima. By observing the positions of two minima, the phase of Γ at any other location can then be determined provided that its distance from the position of the minimum is known. *See* PHASE-ANGLE MEASUREMENT.

Reflectometer. In the case of the slotted line, the probe typically consists of a short length of circular rod which projects into the waveguide in a direction perpendicular to its axis. The reflectometer is based on a different type of coupling mechanism (**Fig. 3**). Here the secondary waveguide is coupled to the primary one by means of a pair of small holes spaced a quarter wavelength apart. On an individual basis, each of these holes is a source of waves in the secondary guide which tend to be of equal amplitude in both directions. When taken as a pair of holes, however, the geometry is such that the waves tend to combine in phase in the forward direction and out of phase (and thus cancel) in the reverse direction. (From a given point in the primary guide to the left of the coupling holes, the distance to a point in the secondary guide to the right of these two holes is the

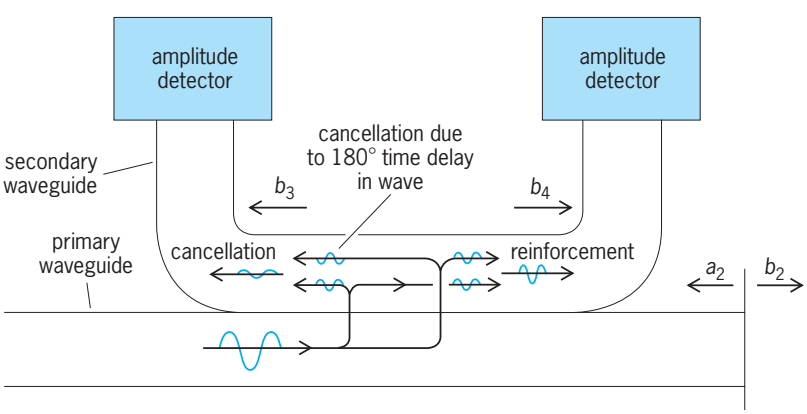

Fig. 3. Directional coupler which forms the basic element of a microwave reflectometer.

same via either of them. But in the reverse direction, that is, to a point in the secondary guide to the left of the two holes, there is a half-wavelength difference in the path length associated with each of the holes. This leads to a reinforcement of the total wave amplitude in one direction and to a cancellation in the other.) *See* INTERFERENCE OF WAVES.

The foregoing waveguide circuit is aptly called a directional coupler, which is the basic element in a reflectometer. If the wave in the primary guide (Fig. 3) is propagating to the left instead of to the right, the coupled wave in the secondary arm will also be primarily to the left. Let each of the secondary arms be terminated by a suitable detector such as a diode or bolometer. These devices respond to the powers which they absorb and which are proportional to $|a|^2$ and $|b|^2$. The square root of the ratio of the detector responses thus gives $|\Gamma|$. *See* DIRECTIONAL COUPLER; MICROWAVE POWER MEASUREMENT.

As before, only the magnitude of Γ is obtained, but this time there is no simple extension of the method which yields the phase. If required, the phase can be obtained by one of the techniques described below.

The foregoing methods have the advantage of requiring only the most simple of microwave detectors, usually either a bolometer or diode, but yield only the amplitude (as contrasted with phase) of the detected signals. On the other hand, the accuracy of these methods is limited by the impossibility of constructing either slotted lines or directional couplers which completely satisfy the simple description outlined above. With regard to the reflectometer, techniques based on the use of tuning transformers have been devised which substantially improve its operation at a single frequency, but these methods become time-consuming in a practical application where measurements at multiple frequencies are required.

Network analyzer. The limitations on accuracy and multiple-frequency operation noted above are largely avoided by the network analyzer, which is a multipurpose instrument with applications in the measurement of attenuation, and possibly power, as well as of impedance.

Let a_2 and b_2 (Fig. 3) represent the wave amplitudes in the primary line, while b_3 and b_4 represent those in the secondary or coupled line. In general, b_3 can be expressed by Eq. (21), where A and B are

$$b_3 = Aa_2 + Bb_2 \tag{21}$$

two complex constants which are functions of the scattering parameters of the directional coupler and its associated detectors. The magnitudes $|A|$ and $|B|$ indicate how tightly the secondary line is coupled to the primary line, while their ratio is a measure of the directive properties of the coupling. In a similar way, b_4 can be expressed by Eq. (22), where C

$$b_4 = Ca_2 + Db_2 \tag{22}$$

and D are additional complex constants. For an ideal reflectometer, one has $B = C = 0$, and taking the ratio of Eqs. (21) and (22) yields Eq. (23), so that if

$$\left|\frac{b_3}{b_4}\right| = \left|\frac{A}{D}\right| \cdot \left|\frac{a_2}{b_2}\right| = \left|\frac{A}{D}\right| \cdot |\Gamma| \tag{23}$$

the detectors measure $|b_3|$ and $|b_4|$ their ratio is proportional to the reflection coefficient. The constant term $|A/D|$ can be easily determined by observing $|b_3/b_4|$ in conjunction with a known value of $|\Gamma|$, usually a short for which $\Gamma = -1$. Thus the system is a reflectometer as described above.

In the network analyzer mode of operation, Eqs. (21) and (22) are combined to yield Eq. (24).

$$\frac{a_2}{b_2} = \Gamma = \frac{\dfrac{b_3}{b_4} - \dfrac{B}{D}}{\dfrac{A}{D} - \dfrac{Cb_3}{Db_4}} \tag{24}$$

Equation (23) gives Γ in terms of the complex ratio b_3/b_4, which means that the detection system must provide the phase difference between b_3 and b_4. In addition, the ratios A/D, B/D, and C/D are also required. These are properties of the coupling network and must be determined at each operating frequency by a suitable calibration procedure. The network analyzer has the advantage that errors due to nonzero values of B and C are in principle eliminated. In addition to the calibration requirement, however, an alternative detection system is required which provides the phase difference between b_3 and b_4. Ordinarily, this calls for heterodyne conversion to a lower frequency which preserves the phase information and where phase detection techniques are well known. *See* HETERODYNE PRINCIPLE.

The typical network analyzer (**Fig. 4**) includes a pair of reflectometers, or test sets, which permits the measurement of a two-port device that has been inserted between the test ports. The analyzer circuit includes a dividing network, which is merely a switch that permits the excitation to be applied to either test set #1 or #2. In the former case, the input reflection is measured by #1, while the transmitted signal is measured by the detectors associated with #2. Given these results, and a similar set with the excitation applied via test set #2, it is possible to obtain the complete set of scattering parameters, and thus the attenuation, of the two-port.

Six-port network analyzer. The six-port network analyzer derives its name from the use of a six-port, rather than four-port, coupling network and the use

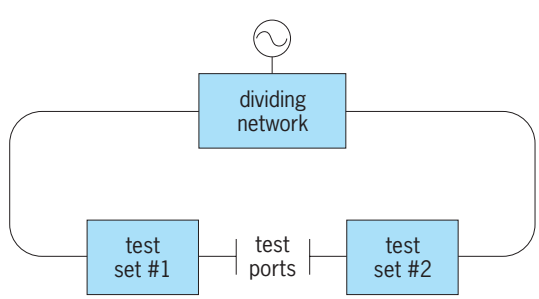

Fig. 4. Basic circuit for a network analyzer.

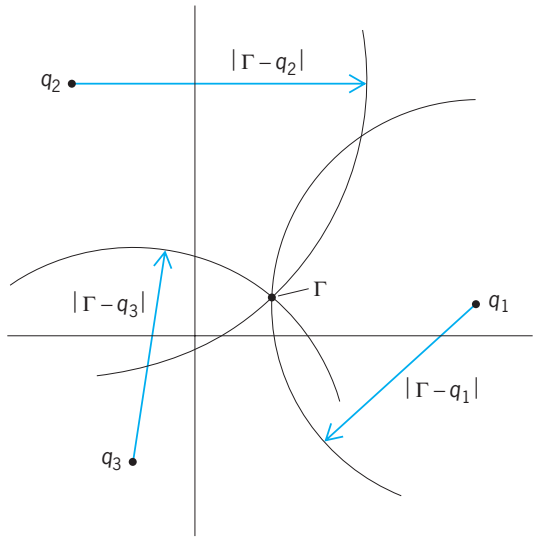

Fig. 5. Geometric construction used with six-port reflectometer to determine the complex reflection coefficient from the intersection of three circles.

of four amplitude detectors rather than a complex-ratio (b_3/b_4) detector. Each detector satisfies Eq. (25),

$$|b_3| = |Aa_2 + Bb_2| \tag{25}$$

similar to Eq. (21), where the absence of phase response is reflected by the use of absolute values.

Equation (25) may be factored to yield Eq. (26), where $q_3 = -(B/A)$.

$$|b_3| = |b_2| \cdot |A| \cdot |\Gamma - q_3| \tag{26}$$

Because of the absolute value sign, it is not possible to solve Eq. (26) for Γ. In the complex plane, $|\Gamma - q_3|$ represents the distance between q_3 and Γ but gives no indication of the direction. The values of Γ which satisfy Eq. (26) lie on a circle whose center is at q_3 and whose radius is determined by $|b_3|$, $|A|$, and $|b_2|$. Here $|b_3|$ is measured by one of the detectors, while $|A|$ and q_3 are parameters of the six-port network.

The term $|b_2|$, however, is an unknown whose value (in common with Γ) depends upon the excitation conditions, and whose determination usually falls within the scope of power rather than impedance measurement. It can be eliminated from consideration in the present context by taking the ratio between two detectors. Here it will prove convenient to assume initially that this second detector is described by Eq. (22) but where $C = 0$. Then, Eq. (27) is satisfied.

$$\left|\frac{b_3}{b_4}\right| = \left|\frac{A}{D}\right| \cdot |\Gamma - q_3| \tag{27}$$

By use of the remaining two detectors, it is possible to form two more equations similar to Eq. (27), but where the six-port design is such that the values of q_3 are different. Each of these equations fixes the distance between Γ and three arbitrary points (circle centers). It is convenient to visualize the operation as providing the value of Γ through the intersection of three circles (**Fig. 5**). Here the circle centers, q_1, q_2, and q_3, are parameters of the six-port, while their radii are given by the detector responses. In practice, where the value of C is usually different from zero, the circle centers as well as their radii depend upon Γ, so that the foregoing description is only an approximate one.

The six-port network analyzer has the advantage of simplicity in the detection system, since the requirement for frequency conversion has been eliminated. But, since the detectors tend to be broadband, it is important for the signal source to be free of harmonics. In addition, if bolometric detectors are employed to achieve the best accuracy, a considerable increase in signal power is required.

The test sets in Fig. 4 may also be implemented as six-port reflectometers, which leads to the dual six-port. In this case, however, the dividing network is such as to provide a simultaneous and nominally equal level of excitation to both test sets, but at several values of phase difference between the two.

Multistate reflectometer. Because most of the associated theory is applicable, the multistate reflectometer may be considered a variant of the six-port network analyzer. A typical implementation (**Fig. 6**) includes a pair of directional couplers, two (rather than four) amplitude detectors, and a moving short in the secondary (or coupled arm) of the coupler on the left. In contrast with the six-port, whose operation requires the ratio of three of the detectors to the fourth, the multistate reflectometer achieves the same result by taking the ratio between the two detectors for three positions of the moving short (whose nominal separation should be a sixth of the wavelength, but are otherwise arbitrary, and must be repeatable). The advantage of the multistate reflectometer is that its operation requires less power, which is an important consideration at the higher microwave frequencies. However, the problems associated with the short adjustment are such as to preclude its operation as a dual six-port.

Calibration methods. As is evident from Eq. (24), the operation of a network anlyzer is characterized by the three complex ratios A/D, B/D, and C/D. Before it can be used to measure impedance, these must first be determined by a suitable calibration. Now Eq. (24) may be written as Eq. (28). Thus, observations of the analyzer response (b_3/b_4) for three

$$\frac{A}{D}\Gamma + \frac{B}{D} - \frac{C}{D}\frac{b_3\Gamma}{b_4} = \frac{b_3}{b_4} \tag{28}$$

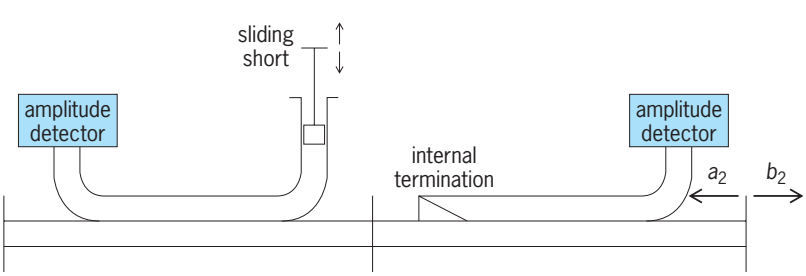

Fig. 6. Basic circuit for a multistate reflectometer.

different terminations for which the values of Γ are known provides a system of three linear equations which may be solved for A/D, B/D, and C/D. In practice, a shorted length of transmission line provides a nominal reflection whose magnitude is unity, while the phase angle may be determined from the line dimensions and the operating frequency. In addition, the response from a nonreflecting ($\Gamma = 0$) termination may be approximated from the response to three or more positions of a weakly reflecting sliding termination. (The latter is a termination of small but unknown reflection which has been designed to slide inside the transmission line or waveguide of interest such that it provides a reflection of constant magnitude but variable phase as its position is changed.) Although this technique may be applied, in turn, to each of the test sets in Fig. 4, the dual configuration permits the use of several alternatives which are based on the thru-reflect-line (TRL) procedure. This procedure calls for observations with the two test ports connected together (thru), then with a strongly reflecting (but otherwise unknown) termination connected, in turn, to each of the test sets (reflect), and finally with a (nominally arbitrary) length of transmission line inserted between the two. Compared with the prior procedure, this technique has the important advantages of eliminating the requirement for terminations of known reflection and sliding terminations, whose implementation is untenable in a line of solid dielectric.

The calibration of the six-port is somewhat more involved. Referring to Eq. (26), the six-port parameters associated with the response, $|b_3|$, are the magnitude $|A|$ and the complex quantity q_3. A similar observation applies to the remaining three detectors. In all there are four real parameters and four complex ones, or (since a complex number may be reduced to two real ones) a total of twelve real ones. As may be inferred from Eq. (27), however, the measurement of the reflection coefficient (as contrasted with the power) requires only the ratio of three of the magnitudes to the fourth. This simplification leaves eleven parameters to be determined in the six-port calibration. From Fig. 5, it is evident that two of the radii (by which Γ is determined) determine the third to the extent of a choice between two possible values. This means that there is a quadratic relationship among the detector responses which may be characterized by five of these eleven parameters and which may be determined by observing the system response to a suitable collection of terminations whose actual values may remain unknown except as required to assure a well-conditioned solution. This procedure is referred to as a six-to-four-port reduction in that the remaining calibration problem is exactly that which has been described earlier. Moreover, the terminations which are utilized in the initial part of this procedure may include those which will be required for the four-port calibration problem.

Swept-frequency reflectometry. In the above discussion, the problems associated with multiple-frequency operation have been largely ignored. As a practical matter, it is frequently important to know how the impedance (or reflection coefficient) varies with frequency. One of the more simple methods of achieving this is with the swept-frequency reflectometer. This is nothing more than the reflectometer described above in conjunction with a swept signal source whose frequency varies smoothly from an initial to a final value and then returns quickly to the starting value for a repeat of the sweep. If the reflectometer output is displayed on an oscilloscope, the reflection coefficient magnitude can be obtained as a function of frequency.

In response to increasing frequency range and accuracy requirements, however, this method has been largely replaced by automated versions of the network analyzers described above. The automatic network analyzer (ANA) typically includes a small digital computer for making the necessary computations and for controlling the signal source. It is thus possible to make a rapid succession of measurements at closely spaced frequencies. Moreover, the phase as well as magnitude of the reflection coefficient is usually obtained.

Time-domain reflectometer. In a typical application, the wave amplitude b (Fig. 1) has its source in a generator (to the left) which is delivering energy, via the transmission line, to an antenna or other termination on the right. Provided that the antenna impedance is equal to Z_0, it is evident from Eqs. (3) and (4) that the amplitude of the wave, a, will equal zero. In a practical application, if a is found to be nonzero, it may be desirable to determine whether the source of this reflected wave is the antenna or possibly some other discontinuity in the transmission line. Although it is possible to answer this question given the results of impedance measurements at a number of different frequencies, a more simple approach, at least conceptually, is provided by the time-domain reflectometer (**Fig. 7**). In one mode of operation the generator output is a series of short pulses. These will travel down the line and be reflected at discontinuities within the line itself (if they exist) or at the antenna (or other termination) if its impedance is different from Z_0. Provided that the pulse width is short in comparison with the transit time from the generator to the discontinuity, or

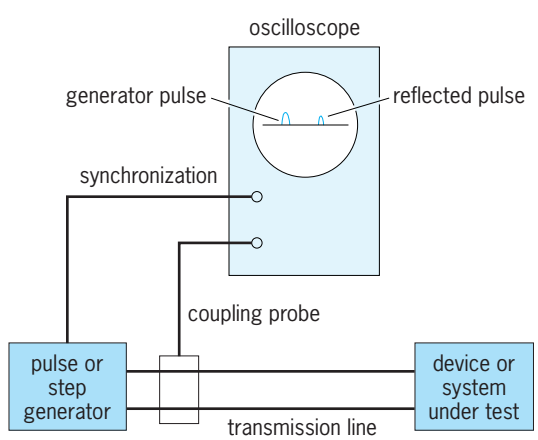

Fig. 7. Basic elements of a time-domain reflectometer.

termination, the oscilloscope responses to the generator pulse and to the reflected pulse will be separated by twice the transit time to the reflection and thus will be displaced from one another along the horizontal time axis. Apart from having confined the transmitted energy to a waveguide rather than free space, the operation is basically the same as radar. Another useful waveform in time-domain reflectometry is the step function. By examining the waveform associated with the reflection, it is possible to learn a great deal about the nature of the discontinuity which produced it. *See* FUNCTION GENERATOR; PULSE GENERATOR; RADAR.

As an alternative to generating these specialized waveforms, it is also possible to synthesize the time-domain response from measurements at a collection of suitably chosen discrete frequencies. The basis for this is found in Fourier analysis. In particular, any arbitrary repetitive waveform (including, for example, a series of short pulses) can be synthesized by the superposition (or addition) of a (usually infinite) number of harmonically related frequencies of the proper amplitude and phase. (As a practical matter, the synthesis is limited to a finite number of frequencies, and the desired waveform is only approximated.) The system response to the synthesized waveform may be obtained by mathematically combining the system responses to this collection of discrete frequencies, usually with the help of a digital computer. *See* FOURIER SERIES AND TRANSFORMS.

Locating reflectometer. Another method of determining the position of waveguide discontinuities is provided by the locating reflectometer, which uses a swept-, rather than stepped-, frequency source. This technique is also based upon Fourier analysis, but the Fourier transformation is obtained by analog methods rather than digital methods, and the need for a (digital) computer is eliminated. On the other hand, the associated waveguide circuit tends to be more complicated.

Comparison reflectometer. As noted above, much of the motivation for the network analyzer is to eliminate errors caused by nonzero values of B and C, which in turn are due to imperfections in the directional coupler. An alternative technique for dealing with this problem is provided by the comparison reflectometer. The solution of Eq. (24) for $|b_3/b_4|$ yields Eq. (29). If $|\Gamma|$ is small, the term $C\Gamma$ may ordinarily

$$\left|\frac{b_3}{b_4}\right| = \frac{|A\Gamma + B|}{|C\Gamma + D|} \tag{29}$$

be neglected in comparison with D. On the other hand, $A\Gamma$ and B may be comparable in magnitude so that the nonzero value of B represents a major source of error. In the comparison reflectometer the directional coupler is separated from the item being measured by a long length of waveguide. This has the effect of causing the phase of Γ to vary rapidly with frequency. A small variation in source frequency will thus cause the terms $A\Gamma$ and B to combine in and out of phase, while ordinarily A and B may be assumed to be constant if the frequency excursion is not too large. The separate contributions of $A\Gamma$ and B to the reflectometer response may then be obtained by techniques similar to those described in conjunction with the slotted line. *See* MICROWAVE MEASUREMENTS; RADIO-FREQUENCY IMPEDANCE MEASUREMENTS.

Glenn F. Engen

Bibliography. A. E. Bailey (ed.), *Microwave Measurements*, 1985; G. H. Bryant, *Principles of Microwave Measurements*, 1988; G. F. Engen, *Microwave Circuit Theory and Foundations of Microwave Metrology*, 1992; K. C. Gupta, *Microwaves*, 1979; I. Kneppo, *Microwave Measurements by Comparison Method*, 1988; T. Laverghetta, *Modern Microwave Measurements and Techniques*, 1988; P. I. Somlo and J. P. Hunter, *Microwave Impedance Measurement*, 1985; F. L. Warner, *Microwave Attenuation Measurement*, 1977.

Microwave landing system (MLS)

An all-weather aircraft landing-guidance system that operates at microwave frequencies and provides deviations from the landing runway centerline using a time-referenced scanning beam (TRSB) technique. The MLS was standardized in 1988 and approved for use in international civil aviation until at least the year 2020. MLS is used to support low-visibility instrument precision approach and landing operations in North America and Europe. In addition to the fixed-base MLS equipment design, a compact mobile microwave landing system (MMLS) equipment design exists. The instrument landing system (ILS) is also standardized internationally and approved for use indefinitely as countries implement their transition to new technologies. Standards for a third landing system, the Global Navigation Satellite System (GNSS), based primarily on Global Positioning System (GPS) technology, exist. Multimode receivers enable an aircraft to conduct an instrument approach using ILS, MLS, or GNSS. *See* INSTRUMENT LANDING SYSTEM (ILS); SATELLITE NAVIGATION SYSTEMS.

Design principles. MLS antennas are designed to generate a narrow fan-shaped main beam with sidelobes (secondary beams) that are typically 25 dB or more below the main beam level. The antenna pattern design used in conjunction with the TRSB technique makes MLS resistive to guidance errors caused by reflection of the radiated signal by objects in the airport environment (such as hangars, buildings, aircraft, or terrain). A reflection of the main lobe will cause negligible guidance errors if it is separated in angle from the line of sight to the aircraft by two main-lobe widths or more. Since a TRSB technique is used, this separation angle equates to different times of arrival at the airborne receiver for the desired signal and undesired reflections. For any landing guidance system sited near the landing runway stop end, the worst-case lateral reflection geometry assumes a large hangar or other reflector offset 500 ft (150 m) from centerline near the threshold of a 15,000-ft (4500-m) runway. Then, the angular offset from runway centerline is about 2°. For this reflector to cause

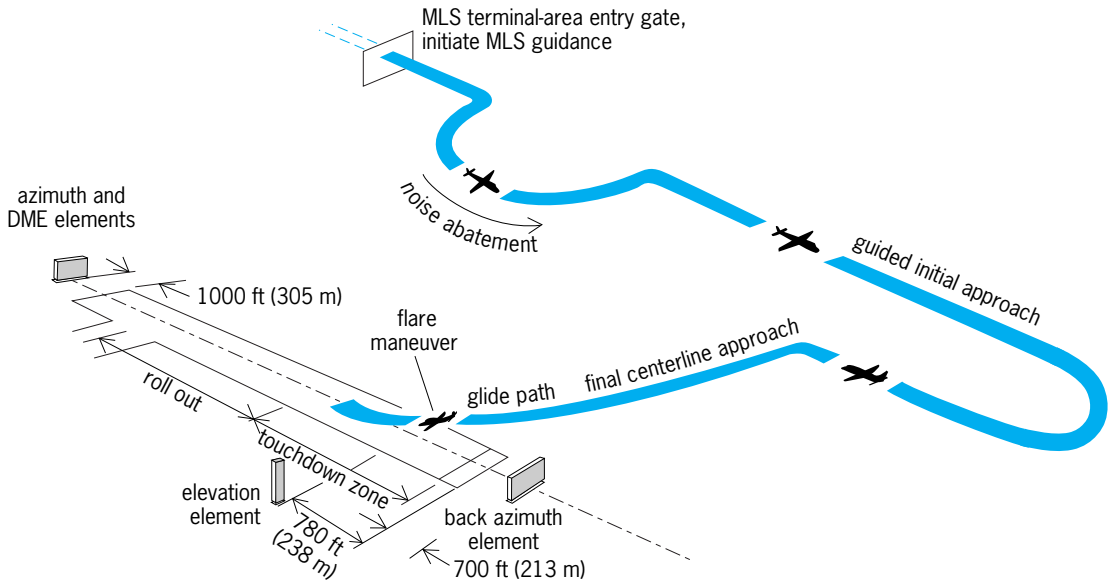

Fig. 1. Landing operation with microwave landing system units located near the runway. The elevation element is sited such that the glide path crosses the threshold above 50 ft (15 m). (*After R. J. Kelly and E. F. C. LaBerge, MLS: A total system approach, IEEE AES Mag., 5(5):27–39 May 1990*)

negligible error, the lateral guidance antenna pattern should have a beamwidth of 1°. In that case, the 2° separation angle approximates two beamwidths. Thus, antennas with beamwidths of 1° keep the lateral guidance errors very small along the centerline approach path. For vertical guidance, the geometry of the reflections from the ground in front of the vertical guidance antenna yields a requirement for antenna pattern beamwidths of 1–2° for a normal approach angle of 3°.

The operating frequencies for MLS lie in a portion of the C band (5030–5091 MHz) designated for use in aeronautical telecommunications. This frequency choice allows a 12-ft (3.6-m) antenna to generate the 1° beamwidth pattern needed to exclude most reflections. The next higher available frequencies (Ku band at 15,000 MHz) would allow the same directivity with an antenna one-third this size (a desirable trade for transportability), but the technological risks increase significantly and the attenuation of signal strength caused by water vapor is severe.

These narrow beams provide accurate vertical guidance down to low heights (8 ft or 2.4 m) above the runway, and thus operations in the lowest-visibility conditions can be supported without limitation from the guidance system. Also, at these microwave frequencies the radio-frequency patterns are well defined at 100–200 ft (30-60 m) from the array, which allows the total radiated signal to be continually validated by field monitors located in protected environments, and provides the high confidence in the signal (high integrity) required for these operations. In order to provide flexible approach paths for a variety of aircraft and rotorcraft, wide lateral and vertical coverage sectors are required. Scanning the narrow microwave beams is an ideal way to generate these coverages, as the numerous reflected signals also generated will be separated, almost always, from the desired signal by a time interval that equates to an angular separation exceeding two beamwidths and thus will not distort the guidance information.

As with ILS, the MLS equipment is sited near the primary runway, with the azimuth transmitter (lateral guidance) and distance-measuring equipment (DME) transponder located near the runway stop end, and the elevation transmitter (vertical guidance) located alongside the runway near landing threshold (**Fig. 1**). This type of siting is referred to as a split-site configuration. With this geometry, the approach course and glide path, generated by the ground equipment, are monitored at the landing runway. Also, the aircraft lateral and vertical displacements due to guidance errors become vanishingly small as the runway is approached and the angular guidance converges to its origin. For special-purpose applications, such as providing guidance to a heliport-only facility, all the MLS equipment may be sited at one common location, known as a collocated site configuration.

Unlike ILS, the 50 times higher frequency of the MLS allows generation of narrow beams by relatively small equipment. Because of this 50:1 scale factor, a 1° beamwidth antenna for MLS requires a 12-ft (3.6-m) antenna, while for ILS a 600-ft (180-m) antenna would be required. *See* ANTENNA (ELECTROMAGNETISM); DISTANCE-MEASURING EQUIPMENT.

Technology. The large coverage volume of MLS is provided by scanning the narrow beams clockwise then counterclockwise for azimuth functions (**Fig. 2**) and up then down for elevation functions. This scanning is electronically controlled at a precise rate of 20,000°/s and fills a lateral sector of 60° (maximum) on each side of the runway centerline and a vertical sector of 30° (maximum). The angular position of the aircraft is decoded by the

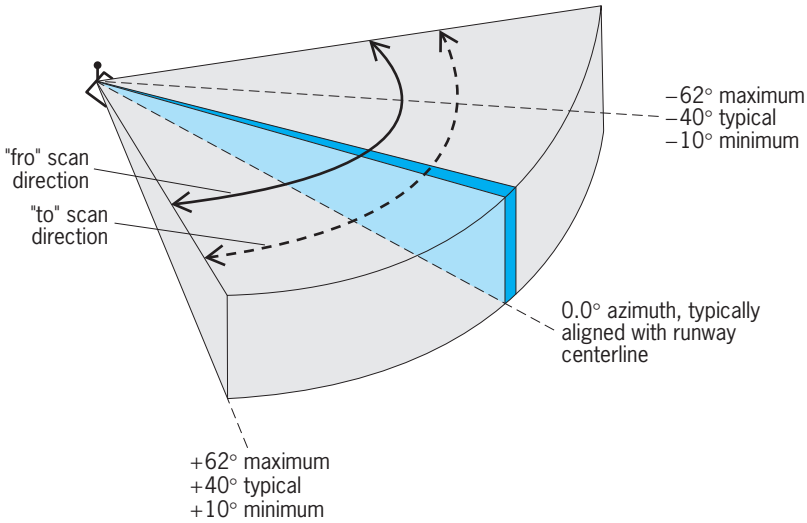

Fig. 2. Azimuth antenna scanning beam and TRSB convention. Maximum, typical, and minimum values of the magnitude of the azimuth at the limits of the coverage volume are indicated. (*After M. F. DiBenedetto, Development of Critical-Area Criteria for Protecting Microwave Landing System Azimuth and Elevation Guidance Signals, Dissertation, Fritz J. and Dolores H. Russ College of Engineering and Technology, Ohio University, Athens, March 1999*)

airborne receiver, which measures the time elapsed between successive passages of azimuth or elevation beams.

The antennas typically used are phased arrays where beam scanning is accomplished by a stored set of commands which, at the appropriate time in the transmission sequence, are directed to variable signal delay devices (phase shifters) associated with each radiating element of the array. These fraction-of-a-wavelength delays are programmed so that the total signal delay increases more rapidly at one end of the array than at the other to cause the radiated phase front to rotate and the pattern to scan in one direction. After a short pause, the program is reversed, and the pattern rotates back to the starting point. The data transmissions use separate fixed-pattern antennas designed to radiate into all parts of the coverage volume simultaneously. Differential phase-shift keying (DPSK) is used to encode the data.

Required navigation performance (RNP). The performance required of an aircraft guidance system may be defined in terms of four parameters: accuracy, integrity, continuity of service, and availability. The criticality of each depends upon the particular flight regime; and for precision approach and landing operations where the aircraft is approaching obstacles (for example, the ground), the emphasis is on guidance signal accuracy and the integrity (believability) of the guidance.

Unlike ILS, the MLS has the same accuracy standard for all categories of operation, which is one result of the reduced environmental effects. Above runway threshold, the lateral error for MLS is limited to ±20 ft (6 m) and the vertical error is limited to ±2 ft (0.6 m). The lateral error limits apply throughout the runway region at 8 ft (2.4 m) above the surface, although increased error is allowed from the threshold out to 20 nautical miles (37 km). The error limits are chosen so that the probability is extremely remote that, with no failures, guidance errors would cause the landing aircraft to stray outside the obstacle-free region near the approach path. The integrity of the guidance is assured by the monitoring system which, for the lowest-visibility operations, is also designed to have an extremely high probability of detecting erroneous guidance and providing an alarm within the required time interval. To meet these very low probabilities, first the ground equipment and the far-field patterns of the radiated MLS signals are continuously monitored, and second the monitor system is verified, end to end, at short intervals. In this manner, the integrity level achieved is controlled by the frequency of the verification.

The MLS receiver provides high integrity by a safety-critical design methodology, including software modules, and by verifying the internal computations at short intervals. For the lowest-visibility applications, the typical aircraft installation may include two or three redundant receivers arranged so that their outputs act as "votes" in an executive monitor.

Mobile microwave landing system (MMLS). The MMLS is a military-grade, all-weather precision instrument approach and landing system. Like MLS, it operates at microwave frequencies and provides deviations from the landing runway centerline using a TRSB technique. The equipment may be sited in either a split-site or a collocated configuration. The MMLS ground equipment is interoperable with both civil and military MLS avionics complying with international standards. The MMLS technology and design principles are very similar to those previously discussed for MLS, with an emphasis on a system design that is compact, mobile, and rapidly deployable. It may be used to support initial deployment of ground forces, for forward-area resupply, for medical evacuation, for rapidly restoring landing service at airports where existing service has been lost due to war or natural disaster, and for other missions not suited for the larger fixed-base civil MLS.

The MMLS equipment is compact and modular in design with a total system weight of less than 600 lb (272 kg). Once on site, the MMLS can be set up and operational in less than 2 hours by a crew of no more than three persons. Achieving the required level of portability mandates the use of antennas that are physically smaller than those used for MLS. This situation results in MMLS antennas with larger antenna pattern beamwidths. Still, antenna beamwidths are typically less than 3.0° for azimuth and less than 2.5° for elevation. Although the coverage and accuracy performance requirements for MMLS are not as stringent as those for MLS, the MMLS performance is well suited for supporting low-visibility approach and landing operations. *See* AIR NAVIGATION; ELECTRONIC NAVIGATION SYSTEMS.

Douglas B. Vickers; Michael F. DiBenedetto

Bibliography. International Civil Aviation Organization, *International Standards, Recommended*

Practices and Procedures for Air Navigation Services: Aeronautical Telecommunications, Annex 10, vol. I, 1985; M. Kayton and W. Fried (eds.), *Avionics Navigation Systems*, 2d ed., Wiley, New York, 1997; R. J. Kelly and E. F. C. LaBerge, MLS: A total system approach, *IEEE AES Mag.*, 5(5):27–39, May 1990; C. Marton (ed.), *Advances in Electronics and Electron Physics*, vol. 57, 1981.

Microwave measurements

A collection of techniques particularly suited for monitoring of devices and systems where physical size of components varies from a significant fraction of an electromagnetic wavelength to many wavelengths. *See* MICROWAVE.

Virtually all microwave devices are coupled together with a transmission line having a uniform cross section. The concept of traveling electromagnetic waves on that transmission line is fundamental to the understanding of microwave measurements. *See* TRANSMISSION LINES.

At any reference plane in a transmission line there are considered to exist two independent traveling electromagnetic waves moving in opposite directions. One is called the forward or incident wave, and the other the reverse or reflected wave. The electromagnetic wave is guided by the transmission line and is composed of electric and magnetic fields with associated electric currents and voltages. Any one of these parameters can be used in considering the traveling waves, but measurements in the early development of microwave technology were made principally on the voltage waves, and this led to the custom of referring only to voltage. One parameter in very common use is the voltage reflection coefficient, Γ, which is related to the incident, V_i, and reflected, V_r, voltage waves by Eq. (1). Often the voltage reflection coefficient is referred to as return loss, which is calculated as $-20 \log_{10} |\Gamma|$ and expressed in decibels (dB). For example, if V_r is 1% of V_i, it is considered that 99.99% of the incident energy is lost in the termination relative to the reflected or returned energy, so the return loss would be 40 dB.

$$\Gamma = \frac{V_r}{V_i} \tag{1}$$

Impedance. The voltage reflection coefficient Γ is related to the impedance terminating the transmission line and to the impedance of the line itself. If a wave is launched to travel in only one direction on a uniform reflectionless transmission line of infinite length, there will be no reflected wave. The input impedance of this infinitely long transmission line is defined as its characteristic impedance Z_0. An arbitrary length of transmission line terminated in an impedance Z_0 will also have an input impedance Z_0. *See* ELECTRICAL IMPEDANCE.

If the transmission line is terminated in the arbitrary complex impedance load Z_L, the complex voltage reflection coefficient Γ_L at the termination is given by Eq. (2).

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0} \tag{2}$$

Even when there is no unique expression for Z_L and Z_0 such as in the case of hollow uniconductor waveguides, the voltage reflection coefficient Γ has a value because it is simply a voltage ratio. In general, the measurement of microwave impedance is the measurement of Γ. Both amplitude and phase of Γ can be measured by direct probing of the voltage standing wave set up along a transmission line by the two opposed traveling waves, but this is a slow technique. Directional couplers have been used for many years to perform much faster swept frequency measurement of the magnitude of Γ, and more recently the use of automatic network analyzers under computer control has made possible rapid, accurate measurements of amplitude and phase of Γ over very broad frequency ranges. *See* DIRECTIONAL COUPLER.

In general, the voltage reflection coefficient Γ measured at the input port of a microwave device derives from a number of reflective discontinuities within the device. The time-domain reflectometer is extremely useful in design work for measuring the positions of the discontinuities and the magnitude of their reflections. A step of voltage with a very short rise time is launched into the input port. Measurement of the time interval between the launching and return of the reflected step gives the position of the discontinuity, and the amplitude of the reflected step gives information on the impedance of the discontinuity. *See* MICROWAVE IMPEDANCE MEASUREMENT.

Power. A required increase in microwave power is expensive whether it be the output from a laboratory signal generator, the power output from a power amplifier on a satellite, or the cooking energy from a microwave oven. To minimize this expense, absolute power must be measured. Most techniques involve conversion of the microwave energy to heat energy which, in turn, causes a temperature rise in a physical body. This temperature rise is measured and is approximately proportional to the power dissipated. The whole device can be calibrated by reference to low-frequency electrical standards and application of appropriate corrections. *See* MICROWAVE POWER MEASUREMENT; RADIOMETRY.

The power sensors are simple and can be made to have a very broad frequency response. A power meter can be connected directly to the output of a generator to measure available power P_A, or a directional coupler may be used to permit measurement of a small fraction of the power actually delivered to the load.

Microwave power measurements provide a good illustration of some of the differences between low-frequency and microwave measurement problems. Consider the simple connection of a load to a generator at low frequency (**Fig. 1**). The equivalent circuit of the generator consists of a source of constant

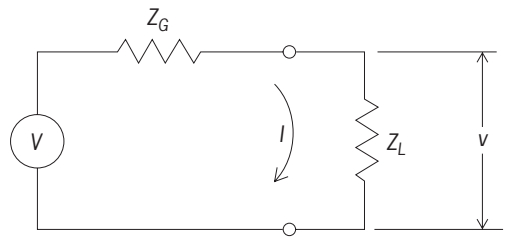

Fig. 1. A load connected to a generator.

voltage V, connected in series with Z_G, the internal impedance of the generator. The load has impedance Z_L.

If the circulating current in Fig. 1 is taken as I and the voltage drop across Z_L as v, Eqs. (3) and (4)

$$I = \frac{V}{Z_G + Z_L} \tag{3}$$

$$v = IZ_L \tag{4}$$

follow from Ohm's law. These are all vector quantities having both a magnitude and a phase angle. *See* ALTERNATING-CURRENT CIRCUIT THEORY.

The power P_L delivered to Z_L is a scalar quantity and can be calculated from Eq. (5), where I^* is the

$$P_L = \mathrm{Re}\{vI^*\} \tag{5}$$

complex conjugate of the current I and the symbol Re{ } stands for the real part of the complex number enclosed inside the braces.

Substituting Eqs. (3) and (4) in Eq. (5) yields Eq. (6) for P_L in terms of the parameters of the circuit.

$$P_L = \mathrm{Re}\left\{\frac{VZ_LV^*}{(Z_G + Z_L)(Z_G + Z_L)^*}\right\} \tag{6}$$

It can be shown that the maximum power which the generator can deliver is available only when the load has impedance Z_G^*, the complex conjugate of the generator impedance. This power is defined as the available power P_A, a fundamental parameter of the generator, and its value is obtained by substituting Z_G^* for Z_L in Eq. (6) and simplifying to obtain Eq. (7). The actual power delivered to Z_L relative to

$$P_A = \frac{VV^*}{4\mathrm{Re}\{Z_G\}} \tag{7}$$

the available power may be determined by taking the ratio of Eq. (6) divided by Eq. (7) to obtain Eq. (8).

$$\frac{P_L}{P_A} = \frac{4\mathrm{Re}\{Z_G\}\mathrm{Re}\{Z_L\}}{|Z_G + Z_L|^2} \tag{8}$$

The same simple circuit can also be considered at microwave frequencies. The same power ratio can be written as Eq. (9), where the voltage reflection

$$\frac{P_L}{P_A} = \frac{\{1 - |\Gamma_L|^2\}\{1 - |\Gamma_G|^2\}}{|1 - \Gamma_L\Gamma_G|^2} \tag{9}$$

coefficient Γ_G of the generator is given by Eq. (10).

$$\Gamma_G = \frac{Z_G - Z_0}{Z_G + Z_0} \tag{10}$$

The expression in Eq. (9) is called a mismatch factor because it would be equal to 1 if the impedances Z_G and Z_L were equal (matched) to Z_0. Since usually the magnitudes $|\Gamma_L|$ and $|\Gamma_G|$ are measured, the denominator of the mismatch factor can commonly be determined only to be within certain limits about 1.

If a uniform transmission line is introduced between the generator and load, the voltage on the line can be considered constant at low frequency, but at microwave frequencies both the phase and the amplitude of the voltage will vary along the line. For example, if the forward-traveling voltage wave has value V_f at a certain reference plane on the transmission line, it will have a value $V_f e^{-\gamma x}$ at a distance x from that reference plane measured in the direction of travel. The parameter γ is the propagation constant and has a real and imaginary part so that the factor $e^{-\gamma x}$ modifies both amplitude and phase of V_f to account for the power loss in the transmission line and the finite time taken for the forward wave to travel the distance x.

Scattering coefficients. While the measurement of absolute power is important, there are many more occasions which require the measurement of relative power which is equivalent to the magnitude of voltage ratio and is related to attenuation. Also, there arises frequently the need to measure the relative phase of two voltages. Measurement systems having this capability are referred to as vector network analyzers, and they are used to measure scattering coefficients of multiport devices. The concept of scattering coefficients is an extension of the voltage reflection coefficient applied to devices having more than one port. The most simple is a two-port device. Its characteristics can be specified completely in terms of a 2×2 scattering matrix (**Fig. 2**). The incident voltage at the reference plane of each port is defined as a, and the reflected voltage is b. Voltages a and b are related by matrix equation (11), where

$$(b_n) = (S_{nm})(a_m) \tag{11}$$

(S_{nm}) is the scattering matrix of the junction. Writing

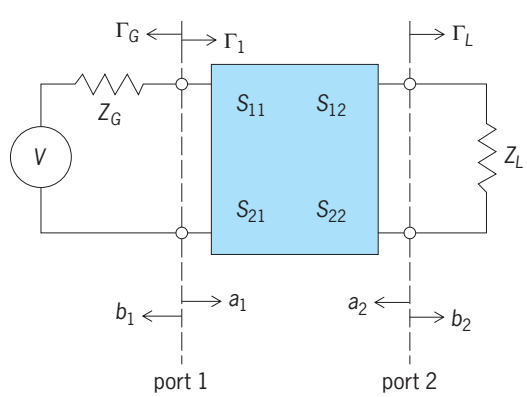

Fig. 2. A two-port inserted between a load and a generator. S_{nm} are the scattering coefficients of the two-port.

Eq. (11) out for a two-port device gives Eqs. (12) and (13). Examination of Eq. (12) shows, for example,

$$b_1 = S_{11}a_1 + S_{12}a_2 \tag{12}$$

$$b_2 = S_{21}a_1 + S_{22}a_2 \tag{13}$$

that S_{11} is the voltage reflection coefficient looking into port 1 if port 2 is terminated with a Z_0 load ($a_2 = 0$). *See* MATRIX THEORY.

Attenuation. To define attenuation and illustrate the use of the scattering equations, it is useful to consider the case in which the two-port is a microwave attenuator. The insertion loss of the attenuator will first be defined, and the attenuation will follow. Consider a source connected to its load as shown in Fig. 1 . The power which will be delivered to Z_L is a function of V, Z_G, and Z_L, and is defined as P_L.

Now if an attenuator is inserted between the source and the load (Fig. 2), the power P_T which will be delivered to Z_L is now a function of V, Z_G, Z_L and the scattering matrix (S_{nm}) of the attenuator. The insertion loss L_I of the attenuator is defined by Eq. (14). Although L_I is a power ratio, independent

$$L_I = 10 \log_{10} \frac{P_L}{P_T} \tag{14}$$

of physical units, it is conveniently expressed as 10 times the logarithm to base 10 and assigned the unit name decibel (dB). *See* DECIBEL.

It can be shown that the insertion loss is given by Eq. (15), where the reflection coefficient Γ_1 is

$$L_I = 10 \log_{10} \left| \frac{(1 - \Gamma_G\Gamma_1)(1 - S_{22}\Gamma_L)}{S_{21}(1 - \Gamma_G\Gamma_L)} \right|^2 \tag{15}$$

the input reflection coefficient of the attenuator-load combination.

The attenuation A of the attenuator may now be defined, and it is the particular value of insertion loss which the attenuator causes when inserted between a generator and load both having zero reflection coefficients. Therefore the expression for A is Eq. (16),

$$A = L_I \Big|_{\Gamma_G, \Gamma_L = 0} = 10 \log_{10} \left| \frac{1}{S_{21}} \right|^2 \tag{16}$$

and the remaining part of the expression in Eq. (15), will be the mismatch factor whose limits may be estimated or whose value may be calculated and applied as a correction if attenuation measurement is being made by using a source and load whose impedances are not matched to the transmission line. *See* ATTENUATION (ELECTRICITY).

Heterodyne. The heterodyne principle is used for scalar attenuation measurements because of its large dynamic range and for vector network analysis because of its phase coherence. The microwave signal at frequency f_s is mixed with a microwave local oscillator at frequency f_{LO} in a nonlinear mixer. The mixer output signal at frequency $f_s - f_{LO}$ is a faithful amplitude and phase reproduction of the original microwave signal but is at a low, fixed frequency so that it can be measured simply with low-frequency techniques. One disadvantage of the heterodyne technique at the highest microwave frequencies is its cost. Consequently, significant effort has been expended in development of multiport network analyzers which use several simple power detectors and a computer analysis approach which allows measurement of both relative voltage amplitude and phase with reduced hardware cost. *See* HETERODYNE PRINCIPLE.

Noise. Microwave noise measurement is important for the communications field and radio astronomy. The noise figure of an amplifier is a measure of the amount of noise added to a signal in the process of amplification. If the amplifier adds a lot of noise, the signal must be correspondingly larger at the input to the amplifier to maintain the minimum signal-to-noise ratio required at the output of the amplifier, for example, to obtain appropriate television picture clarity or insignificant errors in digital communication. If measurement of noise figure shows that one amplifier has a lower noise figure than another, a smaller antenna or a lower transmitter power may be used in a telecommunications link which incorporates the less noisy amplifier.

The measurement of thermal noise at microwave frequencies is essentially the same as low-frequency noise measurement, except that there will be impedance mismatch factors which must be carefully evaluated. The availability of broadband semiconductor noise sources having a stable, high, noise power output has greatly reduced the problems of source impedance mismatch because an impedance-matching attenuator can be inserted between the noise source and the amplifier under test. *See* ELECTRICAL NOISE; ELECTRICAL NOISE GENERATOR; MICROWAVE NOISE STANDARDS.

Sources, detection, and automation. Microwave sources should be stable in both amplitude and frequency for accurate microwave measurements. The amplitude is stabilized by using an amplitude sensor and a negative-feedback loop. Frequency synthesis which permits automatic frequency adjustment with great precision and stability is now fairly common.

Microwave diodes are very often used as detectors. With their extremely low capacitance, their output can follow rapid changes in the amplitude of the microwave input, which is useful for rapid measurements. Although the diode output response varies from square law at low-level inputs to linear at high-level, careful calibration permits use of the diode over a wide dynamic range. A small penalty in accuracy is paid for the simplicity of the detection hardware. *See* DIODE; MICROWAVE SOLID-STATE DEVICES.

The need to apply calculated corrections to obtain the best accuracy in microwave measurement has stimulated the adoption of computers and computer-controlled instruments. An additional benefit of this development is that measurement techniques that are superior in accuracy but too tedious to perform manually can now be considered. An example is the multiport network analyzer using power detectors mentioned above.

Field strength. Measurement of electromagnetic radiated field strength at microwave frequencies requires the guided wave techniques and equipment discussed above plus an antenna of known gain. An antenna to be calibrated may be compared with a standard-gain horn by mounting them interchangeably on an elevated turntable and simply receiving the far-field radiation with a similarly elevated receiving antenna. Measurements of this type on outside antenna ranges have been largely superseded by gain determinations in large anechoic chambers because of the improved possibility of control of both the weather and the electromagnetic environment. *See* MICROWAVE FREE-FIELD STANDARDS.

Uncertainty. As is the norm in all fields of science and engineering, microwave measurements require careful attention to the sources of error that will cause uncertainty. A measurement result is incomplete and, in fact, worthless unless its associated uncertainty is given along with the probability of the measured value falling within the quoted uncertainty range. A cooperative program begun in 1952 involves the national laboratories of many countries in intercomparison of microwave standards to help assure that the microwave measurements made in those countries are meaningful and not in disagreement. *See* PHYSICAL MEASUREMENT. Richard F. Clark

Bibliography. A. E. Bailey (ed.), *Microwave Measurements*, 1985; G. H. Bryant, *Principles of Microwave Measurements*, 1988; G. F. Engen, *Microwave Circuit Theory and Foundations of Microwave Metrology*, 1992; H. Fukui (ed.), *Low Noise Microwave Transistors and Amplifiers*, 1981; K. C. Gupta, *Microwaves*, 1980; T. S. Laverghetta, *Modern Microwave Measurements and Techniques*, 1988; F. L. Warner, *Microwave Attenuation Measurement*, 1977.

Microwave noise standards

Electrical noise generators which produce calculable noise intensities at microwave frequencies, and which are used to calibrate other noise sources by using comparison methods. A factor limiting the performance of microwave communications and radar systems is receiver sensitivity (signal-to-noise ratio). To measure this, a reference noise source is needed. *See* RADAR; SIGNAL-TO-NOISE RATIO.

Blackbody radiator. Noise standards are based upon the blackbody or thermal radiator and generate noise power according to Eq. (1), which is derived

$$S = kTP(f) \qquad \text{W/Hz} \tag{1}$$

from quantum mechanics. Here k is the Boltzmann constant (1.38×10^{-23} joule/kelvin), T the absolute temperature in kelvins, and $P(f)$ the Planck function given by Eq. (2), where h is the Planck constant

$$P(f) = \frac{hf}{kT}[\exp(hf/kT) - 1]^{-1} \tag{2}$$

(6.62×10^{-34} J · s) and f the frequency in Hz. The quantity $TP(f)$ is referred to as the noise temperature and is used as a convenient measure of available noise power from a source. In the microwave region and at room temperature $P(f)$ is nearly equal to 1, and thus $S = kT$ W/Hz. *See* HEAT RADIATION.

The practical realization of a blackbody in the microwave region may best be understood by reference to Kirchhoff's law of radiation, which states that for any body the ratio of emissivity E to absorptivity A is equal to the emissivity E_b of a blackbody; that is, Eq. (3) holds. To realize a blackbody the absorptivity

$$\frac{E}{A} = E_b \tag{3}$$

A is set equal to 1. A microwave absorber with unity absorptivity can be achieved by using a transmission line terminated in its characteristic impedance, or in microwave terminology a matched termination. *See* TRANSMISSION LINES.

Temperature. The range of sources which require calibration and the desire to obtain low uncertainties dictate that microwave thermal noise standards are required with temperatures both above and below the ambient temperature. Sources have been developed with temperatures in the range from 4 to 1300 K (−452 to 1900°F). The low temperatures are normally achieved by immersion of the matched termination in a cryogenic liquid of which liquid nitrogen (77 K or −321°F) is the most common. Standards for measurement of high-temperature sources have the termination in a heated oven. A transition section supports the temperature gradient from the thermal termination to the ambient temperature output which connects to the measurement system (**Fig. 1**). The structures are produced in transmission lines with well-established characteristics such as coaxial lines or waveguides. Complications in the design of microwave thermal noise standards arise principally in the transition and output sections. These are lossy elements with nonzero absorptivity and are thus themselves thermal noise generators. Because the absorptivity A is less than 1 for these elements, the noise power which they contribute is proportional to kAT. If the transmission coefficient of the lossy element is represented by Eq. (4), then for

$$\tau = 1 - A \tag{4}$$

a two-port component at a temperature T_A with an input noise temperature T_{in}, the output noise temperature is given by Eq. (5). The transition and output

$$T_{out} = T_{in}\tau + (1 - \tau)T_A \tag{5}$$

sections thus reduce the noise output from the blackbody termination and add some noise themselves. The combined effect is such that the output noise temperature relative to the termination temperature is lower for an above-ambient source and higher for a below-ambient source. *See* COAXIAL CABLE; WAVEGUIDE.

Radiometers. Measurement of microwave noise involves comparison between an unknown source and a thermal noise source or a secondary standard

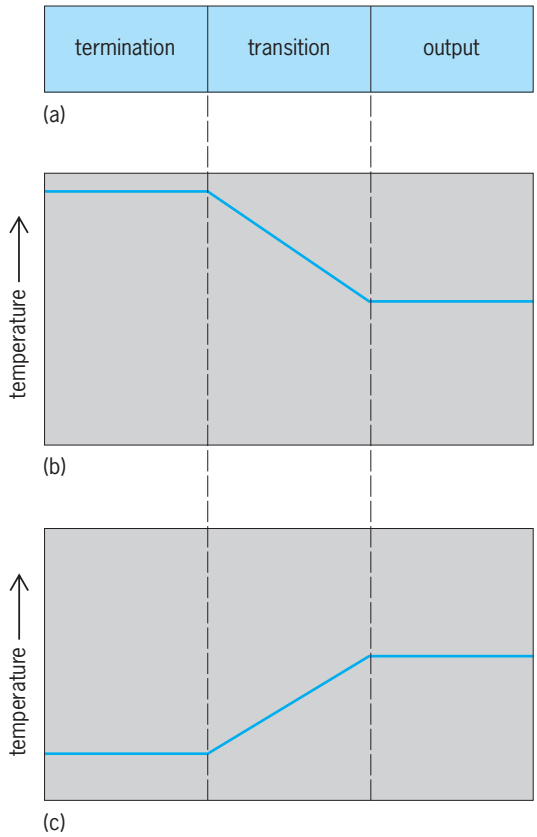

Fig. 1. Microwave noise standard. (*a*) General form. (*b*) Temperature variation in a standard with termination at a temperature above ambient. (*c*) Temperature variation in a cryogenic standard.

which has been compared previously with a thermal source. The measurement systems are classified under the general heading of radiometers, which simply compare the noise powers of the standard and unknown noise sources. The most simple type of radiometer is the total power radiometer (**Fig. 2**).

The standard and unknown noise sources are connected to the input of the amplifier in turn and the output observed on the power meter. The unknown noise source may be calculated by using Eq. (6),

$$T_x = YT_s + T_R(Y - 1) \tag{6}$$

where T_x and T_s are the noise temperatures of the unknown and standard sources, and Y is the ratio of the power meter readings when T_x and T_s are connected,

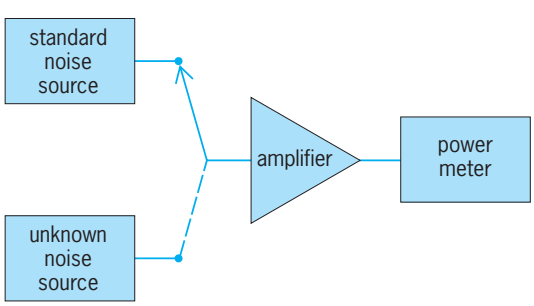

Fig. 2. Total-power radiometer.

respectively. T_R is the input noise temperature of the amplifier system and is determined separately by using two known sources at the system input. The sensitivity of the system is given by Eq. (7), where T_{in}

$$\Delta T = \frac{T_{in}}{(BC)^{1/2}} \tag{7}$$

is the sum of the receiver and source noise temperatures at the input, B is the amplifier bandwidth, and C is the time constant associated with the power meter. ΔT represents the minimum change at the system input which may be detected at the output. The total power radiometer may be used for measurement of above- and below-ambient sources. For the below-ambient (cryogenic) sources the value of T_R must be as low as possible to retain good sensitivity. *See* MICROWAVE; MICROWAVE MEASUREMENTS; RADIOMETRY.

Malcolm W. Sinclair

Bibliography. A. E. Bailey (ed.), *Microwave Measurement*, 1989; G. H. Bryant, *Principles of Microwave Measurement*, 1988.

Microwave optics

The study of those properties of microwaves which are analogous to the properties of light waves in optics. The fact that microwaves and light waves are both electromagnetic waves, the major difference being that of frequency, already suggests that their properties should be alike in many respects. But the reason microwaves behave more like light waves than, for instance, very low frequency waves for electrical power (50 or 60 Hz) is primarily that the microwave wavelengths are usually comparable to or smaller than the ordinary physical dimensions of objects interacting with the waves.

In his classical experiments to verify Maxwell's theory, H. Hertz first demonstrated the optical properties of damped decimeter waves, such as rectilinear propagation, reflection, refraction, and polarization. It is virtually taken for granted that microwaves inherently possess all these properties, and the language of geometrical or physical optics is freely used wherever allowed by the situation.

Rectilinear propagation. As is the case with light, a beam of microwaves propagates along a straight line in a perfectly homogeneous infinite medium. This phenomenon follows directly from a general solution of the wave equation in which the direction of a wave normal does not change in a homogeneous medium. In the use of radar, a microwave beam is justifiably presumed to travel in a straight line before and after reflection by an object. For microwave communication in cases when two distant stations are not along the line of sight, the waves would be blocked by the Earth's surface. The difficulty is remedied by the use of microwave relay links so that straight-line propagation is maintained in each section. *See* WAVE EQUATION.

Reflection and refraction. Consider a plane boundary between two semi-infinite media having different physical properties (**Fig. 1**). If a plane-polarized

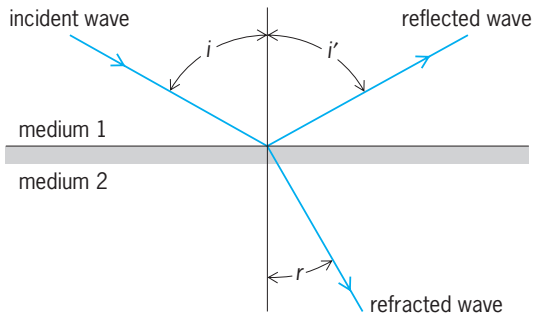

Fig. 1. Reflection and refraction of microwaves at a plane boundary between two insulating media.

microwave is incident from medium 1, the boundary conditions generally require the presence of a reflected wave back to medium 1 and a transmitted (refracted) wave into medium 2. In the case of two insulating media, the familiar relations in optics shown in Eqs. (1) and (2) hold, where the angles i, i', and r

$$i = i' \qquad \text{for reflection} \tag{1}$$

$$\frac{\sin i}{\sin r} = \sqrt{\frac{\epsilon_2}{\epsilon_1}} = N \quad \text{for refraction} \tag{2}$$

are as indicated in Fig. 1, ϵ_1 and ϵ_2 are the dielectric constants of media 1 and 2, respectively, and N is the index of refraction of medium 2 relative to medium 1. The reflected and refracted intensities depend upon whether the incident electric intensity is polarized in the plane of incidence or perpendicular to it. In any case, the well-known Fresnel equations of optics can all be applied to this case. *See* REFLECTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES.

With some modification the laws of reflection and refraction can be applied to the propagation of microwaves inside a dielectric-filled metallic waveguide. The usual case is that of a vertical incidence to the plane boundary between two dielectric media perpendicular to the lengthwise direction of the waveguide. The reflection coefficient can be obtained by measuring the standing-wave ratio in the waveguide. Another interesting application is associated with the microwave analog of total internal reflection in optics. It may be seen from Eq. (2) that if $\epsilon_1 > \epsilon_2$ (that is, if the wave is incident from a denser medium), there is a total internal reflection for the wave when $i > \sin^{-1}\sqrt{\epsilon_2/\epsilon_1}$. This means that a properly designed dielectric rod (without metal walls) can serve as a waveguide by totally reflecting the elementary plane waves. Still another case of interest is that of a microwave lens. By using either a natural dielectric of a certain shape or an artificial dielectric consisting of an array of thin metal plates of a certain design, a microwave lens can be constructed which has the required index of refraction. Such lenses have been used as microwave antennas. *See* ANTENNA (ELECTROMAGNETISM); WAVEGUIDE.

The reflection of a microwave by a conducting plane has all the characteristics of the reflection of a light wave by a metallic mirror. The amplitude of the reflected wave is practically identical to that of the incident wave, with the angle of reflection equal to the angle of incidence. The wave in the conducting medium does not go much beyond a "skin depth" and is of little practical consequence. Examples of reflection of microwaves by conductors are found in parabolic reflectors used as antennas and in targets for radar beams.

Diffraction. In an analogous manner to light, a microwave undergoes diffraction when it encounters an obstacle or an opening which is comparable to or somewhat smaller than its wavelength. Diffraction problems have been much studied but the results are generally too complicated for a simple description. One case of considerable importance, however, may be cited as an illustration. Let two waveguides be coupled through a small hole in a metallic partition as shown in **Fig. 2**. The radius of the hole is assumed to be much smaller than $\lambda/2\pi$, where λ is the wavelength. A wave in one waveguide will leak through the hole by diffraction into the other waveguide. While an exact calculation is difficult, a satisfactory treatment can be worked out by regarding the diffraction effect as being equivalent to the presence of an electric and magnetic dipole placed at the position of the hole. The radiation fields of the dipoles supply the necessary wave coupling between the two waveguides. For additional information on microwave diffraction *See* DIFFRACTION.

Polarization. The polarization of an electromagnetic wave is specified by the direction of the electric and magnetic intensities. For simplicity, consider a plane wave propagating in the z direction. The electric and magnetic intensities are always mutually perpendicular to each other in the xy plane. It is, consequently, only necessary to consider the polarization of the electric intensity alone. If the electric intensity is polarized along one direction, say the x axis, then the wave is said to be plane polarized. If there are components E_x and E_y equal in amplitude but differing in phase by 90°, the wave is circularly polarized

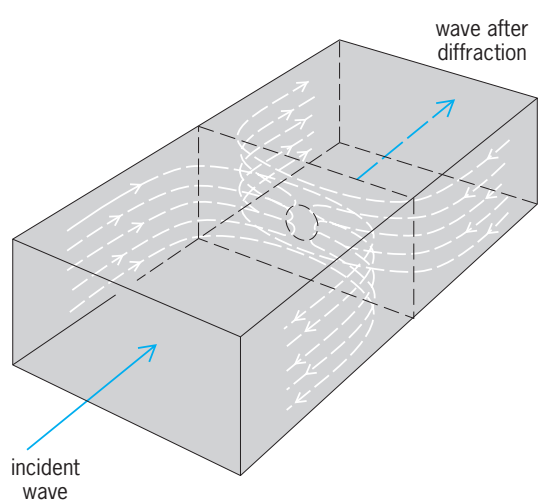

Fig. 2. Diffraction of microwaves through a small aperture between two waveguides. Only magnetic lines of force are shown for the field pattern.

(right-handed for E_y lagging E_x and left-handed for E_y leading E_x). Lastly, if the wave is neither plane nor circularly polarized, it must be elliptically polarized. *See* POLARIZATION OF WAVES; POLARIZED LIGHT.

The preceding essentially optical description holds true for microwaves in free space or wherever there is a transverse electric and magnetic (TEM) wave. Since a hollow waveguide does not support a TEM wave, the description of the polarization of the wave is much more complicated. However, the general notions expressed here are still valid and useful. For instance, the electric intensity of the dominant mode (TE_{01}) in a rectangular waveguide is plane polarized. This situation is also approximately true for the dominant mode (TE_{11}) in a circular waveguide. A circular waveguide is particularly useful in transforming plane-polarized electric intensity into circularly polarized electric intensity by a technique equivalent to the use of a quarter-wave plate in optics. For this purpose, it suffices to use a thin slab of dielectric material, such as mica, to introduce a 90° phase shift for one of the two equal components of the electric intensity. For information on TEM waves *see* MICROWAVE.

Faraday effect for microwaves. In optics, the Faraday effect is the rotation of the plane of polarization of a light beam which propagates in a dense transparent medium placed in a magnetic field along the direction of propagation. In microwaves, a similar phenomenon has been discovered and has led to many interesting applications. *See* FARADAY EFFECT.

Consider, for example, a circular waveguide which contains a slender rod of ferrite (a magnetic material of very low conductivity), as shown in **Fig. 3**. Place a steady magnetic field along the axial direction. If a wave with vertical electric polarization is incident from the left, then in passing through the ferrite zone its plane of polarization is rotated by an angle θ as indicated in the figure. This Faraday effect can be explained by the action of precessing elementary magnets in the magnetized ferrite upon the phase of the propagating wave. The initial plane-polarized electric intensity is equivalent to two oppositely rotating circularly polarized components. Only one of the components is affected (principally in the phase factor) by the precessing magnets because the latter can only precess in a unique direction corresponding to a given direction of the magnetic field. In the

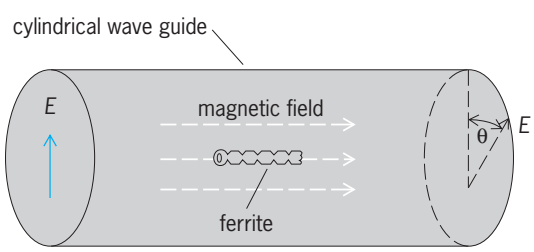

Fig. 3. Rotation of the plane of polarization of a microwave by a ferrite rod in a longitudinal magnetic field. *E* represents electric intensity.

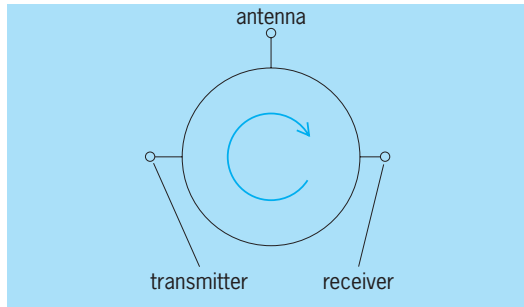

Fig. 4. Circulator formed by one or more gyrator elements for nonreciprocal application.

output, the combination of two circularly polarized waves with a relative phase shift is equal to a plane-polarized wave having its polarization rotated by an angle with respect to the initial plane.

If the output wave is sent back to pass by the ferrite from right to left, the plane of polarization of the backward wave rotates another θ degrees in the same direction as the initial rotation. In other words, the polarization of the new output wave on the left is rotated from the initial plane by 2θ, instead of zero degrees as might be expected by the principle of reciprocity. Thus this system constitutes a nonreciprocal circuit element and is sometimes called a gyrator to signify the gyrating motion of the elementary magnets. *See* GYRATOR; RECIPROCITY PRINCIPLE.

One of the most important applications of a gyrator is found in unidirectional transmission. Suppose a plane-polarized wave is introduced from a rectangular to a circular waveguide containing a gyrator and finally to another rectangular waveguide oriented to suit the polarization of the output wave. If the angle of rotation, which defines the output polarization, is adjusted to 45°, any wave from the output which is fed back to the input will be at right angles with the input intensity and hence will not be accepted by the rectangular waveguide. Another important application of gyrators is known as a circulator, a simple example of which is given in **Fig. 4**. By means of one or more gyrators, a circulator such as the one depicted will allow a wave from the transmitter to go to the antenna but not to the receiver and will let a wave from the antenna go to the receiver but not to the transmitter.

C. K. Jen

Bibliography. S. Cornbleet, *Microwave and Geometric Optics*, 1994; N. Marcuvitz et al., *Waveguide Handbook*, 1986; A. D. Oliver, *Microwave and Optical Transmission*, 1992; D. M. Pozar, *Microwave Engineering*, 2d ed., 1997.

Microwave power measurement

Determination of the rate at which energy is transmitted by microwave propagation. Power is one of the fundamental parameters measured at microwave frequencies. Since the physical dimensions of the microwave system are comparable with the wavelength, the voltage along a conductor is no longer

constant. Further, for energy transmitted inside a hollow, uniconductor waveguide, the voltage difference between two points is not uniquely defined and varies as a function of the path taken to pass from one point to the other. Thus voltage standards, basic to low-frequency electrical measurements, give way to power standards at microwaves. Power meters are connected either temporarily, to measure all the power, in place of the normal termination, or permanently to measure a fixed portion of the power. *See* MICROWAVE; POWER.

Terminating power meters. A terminating power meter absorbs a power P_M at its one input port. The effective efficiency η_e of the terminating power meter is defined by Eq. (1), where P_I is the power in-

$$\eta_e = \frac{P_I}{P_M} \tag{1}$$

dicated by the power meter. A parameter designated the calibration factor K of a terminating power meter is defined by Eq. (2). In this case the power incident

$$K = \frac{P_I}{\text{power incident}} \tag{2}$$

has a rather special meaning. It is the energy per second flowing in the electromagnetic wave incident on the power meter as measured an infinitesimal distance ahead of the power meter input.

In order to relate the two parameters, η_e and K, consider the terminating power meter connected to a short section of transmission line having characteristic impedance Z_0 which has electromagnetic energy fed into its opposite end (**Fig. 1**). If the incident electromagnetic wave is assumed to have voltage V at an infinitesimal distance δ in front of the power meter input, the reflected electromagnetic wave at the same plane will have voltage $\Gamma_M V$, where Γ_M is the voltage reflection coefficient of the power meter. *See* ELECTRICAL IMPEDANCE; MICROWAVE IMPEDANCE MEASUREMENT; REFLECTION AND TRANSMISSION COEFFICIENTS; TRANSMISSION LINES.

Since the energy flow in the traveling waves is proportional to the square of the voltage, assuming an ideal section of transmission line connected to the power meter, it follows that Eq. (3) is valid, where

$$\frac{P_M}{\text{Power incident}} = 1 - |\Gamma_M|^2 \tag{3}$$

$|\Gamma_M|$ is the magnitude of the voltage reflection coefficient. Hence the relation between K and η_e is given by Eq. (4).

$$K = \eta_e(1 - |\Gamma_M|^2) \tag{4}$$

While η_e and K are both useful parameters for a terminating power meter, η_e should be considered the more fundamental since it is independent of the characteristic impedance of the connecting line and of the impedance of the power meter. Furthermore, the two parameters are very simply related, and either one may be determined through knowledge of the other one and an impedance measurement.

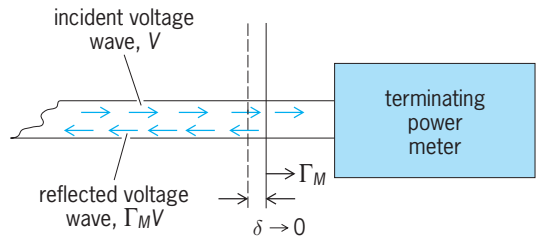

Fig. 1. Microwave power incident on a terminating power meter.

The power P_0 that a microwave source would deliver to a termination with characteristic impedance Z_0, as measured by the terminating power meter, is given by Eq. (5), where Γ_G is the voltage reflection

$$P_0 = \frac{P_I}{K}|1 - \Gamma_G\Gamma_M|^2 \tag{5}$$

coefficient of the source. Often in the past, the impedance portion of the expression in Eq. (5) was not calculated, but its maximum deviation from 1 was included in the error budget of the measurement. Now, as vector network analyzers with their capability to measure both phase and amplitude of Γ have come into wider usage, it has become more common to do the impedance mismatch calculation to reduce the uncertainty of the measurement.

Two-port power monitor. The use of a two-port power monitor was proposed as a technique for assembling a standard power source. The general form of this type of device consists of a stable three-port network converted to a two-port power monitor by permanent connection of a power meter to one of its ports. An excellent power monitor (**Fig. 2**) has a terminating power meter connected to the side-arm, port 3, of a three-port directional coupler. The output from port 3 is proportional to the microwave energy traveling in the main line of the directional coupler in the direction from port 1 to port 2. The fourth port of the directional coupler is terminated internally. A source of radio-frequency energy with adjustable amplitude is connected to port 1. Port 2 becomes the output of a standard source with power available and output impedance given by the power meter reading and the characteristics of the three-port. The variable attenuator may be adjusted manually for a desired power meter reading, or, as suggested by the broken line, an electrical output from the power meter may be used to automatically control the attenuator to

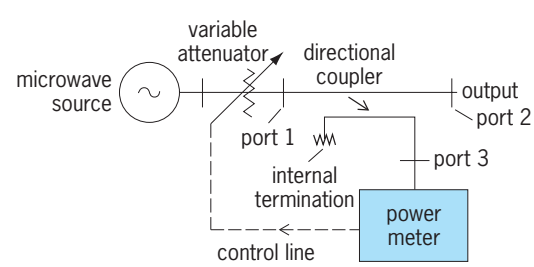

Fig. 2. A standard source of microwave power stabilized by a power meter.

maintain a constant power available from port 2. *See* DIRECTIONAL COUPLER.

A calibration factor may be assigned to the directional coupler with side-arm power monitor. The calibration factor of interest is K_2 and is defined by Eq. (6), where PI_3 is the power indicated by the

$$K_2 = \frac{P_{I_3}}{P_{0_2}} \tag{6}$$

power monitor at port 3 when the power delivered to a load with characteristic impedance Z_0 connected at port 2 is P_{0_2}. Calibration of the coupler-power monitor is thus performed by connection of a power meter having known characteristics to the output port 2 and determination of P_{0_2} for a particular power monitor reading. Use is made of Eq. (5) for this measurement.

The directional coupler-power monitor must now be regarded as a single device and must remain connected together permanently as a calibrated feed-through power monitor. The advantage of this coupler-power monitor is that it can have a very low insertion loss and yet have a very low effective source reflection coefficient. However, its bandwidth is limited by the directional coupler.

For cases in which the low insertion loss is not necessary, it is possible to assemble an extremely broadband feed-through power meter (**Fig. 3**). The power splitter has no low-frequency cutoff, but three-fourths of the power available from the source is lost. In the particular case where R is 50 ohms, the effective source reflection coefficient has magnitude zero in a 50-ohm transmission line system. This is due to the fact that a monitored T junction behaves as a zero impedance voltage source.

Power sensors. A number of techniques have been used to sense microwave power. These include (1) conversion of the microwave power to heat and sensing the thermal effect on a calorimetric body, (2) measurement of the force caused by microwave radiation pressure, (3) measurement of a microwave Hall effect, (4) measurement of electron acceleration caused by a microwave electric field, and (5) rectification of microwave voltage by using diodes. Rectification of microwave voltage is the most simple and most sensitive technique, but conversion of microwave power to heat is by far the most widely used because of its long-term stability. Two versions of detector, based on conversion to heat, are bolometers and thermocouples. *See* DIODE; HALL EFFECT; RADIATION PRESSURE; RECTIFIER.

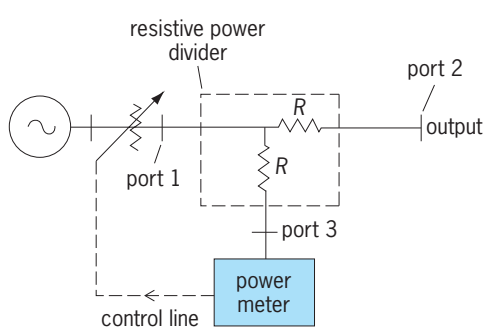

Fig. 3. A standard microwave source having very broad frequency range.

Bolometer mounts. A bolometer is a component whose resistance varies with temperature. Very small bolometers are inside a bolometer mount having a transmission line input. The bolometer is located so that it can simultaneously absorb microwave power and dc power. By using electronic instrumentation, the bolometer resistance is held constant by automatic adjustment of the dc bias power dissipated in the bolometer. When microwave power is fed into the mount, the dc power is automatically reduced to keep the bolometer resistance constant, and the change in dc power required is a measure of the microwave power. The most used type of bolometer is the thermistor with a negative temperature coefficient of resistance. Since this is a power substitution technique, any long-term drift or instability in the thermistor is compensated by an automatic change in dc bias power. *See* BOLOMETER; THERMISTOR.

Thermocouple detectors. A type of power detector in which the microwave power is dissipated directly on the warm junctions of a thermopile has largely replaced the coaxial thermistor mount for general usage. The electronic instrumentation is simply a low-noise dc amplifier-voltmeter combination to measure the thermopile output. Also, manufacturers have achieved a better impedance match over a broad frequency range with this detector than with a thermistor mount. One complication is the inclusion of a stable radio-frequency reference source into the electronic part of the measurement system to permit checking of the long-term stability of the detector. *See* RADIOMETRY; THERMOCOUPLE.

Standards. The power sensors and their associated electronics are the standards for measurement of microwave power throughout industry and in the laboratory. The electronic portion can usually be calibrated by reference to dc standards. The efficiency and calibration factor of the microwave power sensors are compared through a hierarchy of laboratories to fundamental standards of microwave power. These fundamental standards are essentially the calorimetric type mentioned above, but are configured so that their efficiency is either calculable or measurable.

In one type of fundamental standard, referred to as a coaxial calorimeter, microwave power is dissipated in a resistor and the temperature rise of the resistor and its mounting is measured by a thermopile. Direct-current power for the same temperature rise is fed into the calorimeter through the same input port as was used for microwave power. The calorimeter is constructed so that the microwave and direct currents in the resistor are as nearly as possible the same. The difference between the microwave loss and the dc loss in the thermally insulating input to the resistor is measured and used to calculate the reduction in efficiency of the calorimeter at high measurement frequencies.

A second type of fundamental standard involves again a calorimetric measurement, but in this case

a thermistor mount is placed in thermal isolation inside an enclosure. The dc bias power causes the thermistor mount to heat up relative to its immediate surroundings, and a thermopile indicates this relative temperature rise. When microwave power is applied, the dc power automatically decreases to maintain constant power dissipation in the thermistor. However, any microwave power not dissipated in the thermistor itself causes an increase in thermopile output which is interpreted as a measure of the inefficiency of the thermistor mount.

Transfer of the accuracy of the power standard to a power sensor under test is a two-step comparison process. For the first step, the power standard is connected to the output of a source (Figs. 1 and 2). Modern microwave sources are quite stable in amplitude so the controlled variable attenuator would not be necessary here. Instead, simultaneous readings of the power indicated by the standard and by the source power meter permits the calibration of the source as a so-called working-standard power source. The second step involves the connection of the power sensor under test to the working-standard power source. Simultaneous power readings and a calculation of impedance mismatch corrections complete the calibration. *See* ELECTRIC POWER MEASUREMENT; MICROWAVE MEASUREMENTS.

Richard F. Clark

Bibliography. G. H. Bryant, *Principles of Microwave Measurement*, 1988; A. E. Fantom, A. E. Bailey, and A. C. Lynch (eds.), *Radio Frequency and Microwave Power Measurement*, 1990; T. S. Laverghetta, *Handbook of Microwave Testing*, 1981; T. S. Laverghetta, *Modern Microwave Measurements and Techniques*, 1988; T. S. Laverghetta, *Practical Microwaves*, 1996.

Microwave solid-state devices

Semiconductor devices used for the detection, generation, amplification, and control of electromagnetic radiation with wavelengths from 30 cm to 1 mm (frequencies from 1 to 300 GHz). Since 1985, the number and variety of microwave semiconductor devices, used for wireless and satellite communication and optoelectronics, have increased as new techniques, materials, and concepts have been developed and applied. Passive microwave devices, such as *pn* and PIN junctions, Schottky barrier diodes, and varactors, are primarily used for detecting, mixing, modulating, or controlling microwave signals. Step-recovery diodes, transistors, tunnel diodes, and transferred electron devices (TEDs) are active microwave devices that generate power or amplify microwave signals. *See* MICROWAVE; SEMICONDUCTOR DIODE.

Materials

Typical high-frequency semiconductor materials include silicon (Si), germanium (Ge), and compound semiconductors, such as gallium arsenide (GaAs), indium phosphide (InP), silicon germanium (SiGe), silicon carbide (SiC), and gallium nitride (GaN). In general, the compound semiconductors work best for high-frequency applications due to their higher electron mobilities. *See* GALLIUM; GERMANIUM; SEMICONDUCTOR; SILICON.

Silicon microelectronics is a mature technology. Its advantages include good mechanical strength, high thermal conductivity, and a high-quality oxide (SiO_2), allowing for the fabrication of field-effect transistors (FETs). FETs form the basis for complementary metal-oxide semiconductor (CMOS) technology. A major disadvantage of silicon is its low intrinsic resistivity, resulting in very high dielectric losses, which make it almost impossible to transmit microwave signals. To help alleviate the inherent low resistivity of silicon, high-resistivity silicon-on-insulator (SOI) can be used to reduce circuit parasitics (such as parasitic capacitances and resistances). Even with improvements, silicon devices are limited in high-frequency microwave applications, thus requiring inherently faster materials, such as GaAs, InP, or SiGe. The electron mobility and energy bandgap for GaAs and InP are higher than the corresponding values for silicon. A major advantage is that semi-insulating GaAs and InP substrates are easily produced, allowing interconnects to be put over these substrates without introducing large parasitic capacitances. By metallizing the bottom side of the substrate, microstrip lines can be made that integrate the passive microwave circuits with transistors on monolithic microwave integrated circuits (MMICs). The disadvantages of GaAs are low hole mobility, which severely limits its ability to support complementary circuitry, and the inability to form a high-quality oxide on GaAs, a shortcoming that has been partially mitigated by high electron mobility transistors (HEMTs).

Passive Devices

Passive microwave solid-state devices include Schottky-barrier diodes, PIN diodes, and varactor diodes.

PIN and pn diodes. A PIN (*p*-type/intrinsic/*n*-type) diode is a *pn* diode that has an undoped (intrinsic) region between the *p*- and *n*-type regions. The use of an intrinsic region in PIN diodes allows for high-power operation and offers an impedance at microwave frequencies that is controllable by a lower frequency or a direct-current (DC) bias. The PIN diode is one of the most common passive diodes used at microwave frequencies. Under zero and reverse bias the PIN diode has a very high impedance, whereas at moderate forward current its impedance is relatively low. The DC behavior of a PIN diode is basically the same as that of an ordinary *pn* junction. However, at microwave frequencies PIN diodes behave as a frequency-independent conductance that is DC-controlled. The capacitance is mainly determined by the width of the intrinsic region, which is only weakly influenced by the DC bias. The DC control current through a PIN diode can vary the resistance value from over 1 megohm to less than 1 ohm. With reverse or zero bias, the intrinsic region is depleted of carriers and the PIN exhibits very high resistance. When forward bias is applied across the PIN diode, holes from the

p-region and electrons from the *n*-region are injected into the intrinsic region, which increases the conductivity. High off-resistance and low on-resistance make the PIN diode highly attractive for switching applications. The PIN diode switch can be viewed as a resistor in the forward-biased state and a capacitor in the reverse-biased state. PIN diodes are used to switch lengths of transmission line, providing digital increments of phase in individual transmission paths, each capable of carrying kilowatts of peak power. PIN diodes come in a variety of packages for microstrip and stripline packages, and are used as microwave switches, modulators, attenuators, limiters, phase shifters, protectors, and other signal control circuit elements. *See* JUNCTION DIODE.

Schottky barrier diodes. A Schottky barrier diode (SBD) consists of a rectifying metal-semiconductor barrier typically formed by deposition of a metal layer on a semiconductor. The SBD functions in a similar manner to the antiquated point contact diode and the slower-response *pn*-junction diode, and is used for signal mixing and detection. The point contact diode consists of a metal whisker in contact with a semiconductor, forming a rectifying junction. The SBD is more rugged and reliable than the point contact diode. The SBD's main advantage over *pn* diodes is the absence of minority carriers, which limit the response speed in switching applications and the high-frequency performance in mixing and detection applications. SBDs are zero-bias detectors. Frequencies to 40 GHz are available with silicon SBDs, and GaAs SBDs are used for higher-frequency applications. *See* SCHOTTKY EFFECT.

Varactor diodes. The variable-reactance (varactor) diode makes use of the change in capacitance of a *pn* junction or Schottky barrier diode, and is designed to be highly dependent on the applied reverse bias. The capacitance change results from a widening of the depletion layer as the reverse-bias voltage is increased. As variable capacitors, varactor diodes are used in tuned circuits and in voltage-controlled oscillators. For higher-frequency microwave applications, silicon varactors have been replaced with GaAs. Typical applications of varactor diodes are harmonic generation, frequency multiplication, parametric amplification, and electronic tuning. Multipliers are used as local oscillators, low-power transmitters, or transmitter drivers in radar, telemetry, telecommunication, and instrumentation. *See* VARACTOR.

Active Microwave Devices

Transistors are the most widely used active microwave solid-state devices. At very high microwave frequencies, high-frequency effects limit the usefulness of transistors, and two-terminal negative resistance devices, such as transferred-electron devices, avalanche diodes, and tunnel diodes, are sometimes used.

Microwave transistors. Two main categories of transistors are used for microwave applications: bipolar junction transistors (BJTs) and field-effect transistors (FETs). In order to get useful output power at high frequencies, transistors are designed to have a higher periphery-to-area ratio using a simple stripe geometry. The area must be reduced without reducing the periphery, as large area means large interelectrode capacitance. For high-frequency applications the goal is to scale down the size of the device. Narrower widths of the elements within the transistor are the key to superior high-frequency performance. *See* TRANSISTOR.

Microwave bipolar transistors. A BJT consists of three doped regions forming two *pn* junctions. These regions are the emitter, base, and collector in either an *npn* or *pnp* arrangement. Silicon *npn* BJTs have an upper cutoff frequency of about 25 GHz (varies with manufacturing improvements). The cutoff frequency is defined as the frequency at which the current amplification drops to unity as the frequency is raised. The primary limitations to higher frequency are base and emitter resistance, capacitance, and transit time. Modern silicon *npn* BJTs are of two designs: the planar design has its regions implanted or diffused under the surface and patterned using photolithography techniques, and the epitaxial design uses a thin, low-conductivity epitaxial layer for part of the collector region (the remainder is highly conductive, n^+ material).

To operate at microwave frequencies, individual transistor dimensions must be reduced to micrometer or submicrometer size. To maintain current and power capability, various forms of internal paralleling on the chip are used. Three of these geometries are interdigitated fingers that form the emitter and base, the overlaying of emitter and base stripes, and the matrix approach. Silicon BJTs are mainly used in the lower microwave ranges. Their power capability is quite good, but in terms of noise they are inferior to GaAs metal semiconductor field-effect transistors (MESFETs) at frequencies above 1 GHz and are mainly used in power amplifiers and oscillators. They may also be used in small-signal microwave amplifiers when noise performance is not critical. *See* ELECTRICAL NOISE.

Heterojunction bipolar transistors. Heterojunction bipolar transistors (HBTs) have been designed with much higher maximum frequencies than silicon BJTs. HBTs are essentially BJTs that have two or more materials making up the emitter, base, and collector regions (**Fig. 1**). In HBTs, the major goal is to limit the injection of holes into the emitter by using an emitter material with a larger bandgap than the base. The difference in bandgaps manifests itself as a discontinuity in the conduction band or the valence band, or both. For *npn* HBTs, a discontinuity in the valence band is required. In general, to make high-quality heterojunctions, the two materials should have matching lattice constants. For very thin layers, lattice matching is not absolutely necessary as the thin layer can be strained to accommodate the crystal lattice of the other material. Fortunately, the base of a bipolar transistor is designed to be very thin and thus can be made of a strained layer material. Combinations such as AlGaAs/InGaAs and Si/SiGe are possible. *See* BAND THEORY OF SOLIDS; ELECTRICAL CONDUCTIVITY OF METALS; HOLE STATES

IN SOLIDS; SEMICONDUCTOR HETEROSTRUCTURES.

Aluminum gallium arsenide (AlGaAs) and GaAs fulfill the lattice-matching requirement. AlGaAs/GaAs HBTs typically have an AlAs mole fraction of 20–30%. HBTs, in comparison with silicon, have higher electron mobility, resulting in higher cutoff frequencies; a larger bandgap, thus less thermal generation of charge carriers; a semi-insulating substrate for isolation; and the possibility of integration with optoelectronic components. Compared with GaAs-based MESFETs, HBTs have higher transconductance and output current. HBTs can have cutoff frequencies higher than 300 GHz.

The base of the Si/SiGe HBT is made of an alloy of silicon with about 10% germanium. Since germanium has a bandgap of 0.7 electronvolt (eV), as compared with 1.1 eV for silicon, the base will have a lower bandgap than the silicon emitter. A potential problem is that the lattice constant of germanium is larger than that of silicon, which could produce crystal defects at the interfaces. However, if the germanium content is relatively small and the base region is thin enough, the SiGe layer is strained to match the silicon crystal lattice and defects are not a serious problem. The bandgap difference is small and is in the valence band discontinuity. SiGe HBTs are compatible with standard silicon technology. *See* CRYSTAL DEFECTS; CRYSTAL STRUCTURE.

Microwave field-effect transistors. Field-effect transistors (FETs) operate by varying the conductivity of a semiconductor channel through changes in the electric field across the channel. The three basic forms of FETs are the junction FET (JFET), the metal semiconductor FET (MESFET), and the metal-oxide semiconductor FET (MOSFET). All FETs have a channel with a source and drain region at each end and a gate located along the channel, which modulates the channel conduction (**Fig. 2**). Microwave JFETs and MESFETs work by channel depletion. The channel is *n*-type and the gate is *p*-type for JFETs and metal for MESFETs. FET structures are well suited for microwave applications because all contacts are on the surface to keep parasitic capacitances small. The cutoff frequency is mainly determined by the transit time of the electrons under the gate; thus short gate lengths (less than 1 μm) are used.

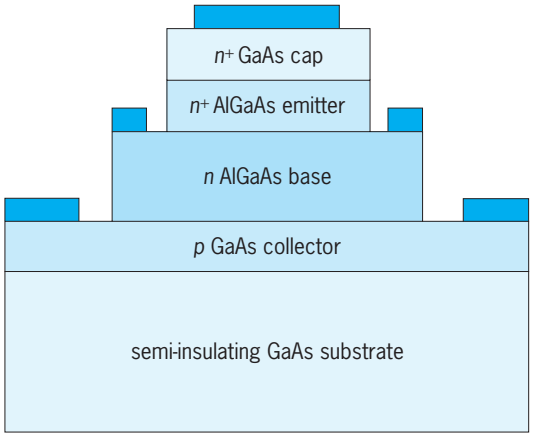

Fig. 1. Materials composition for a heterojunction bipolar transistor (HBT).

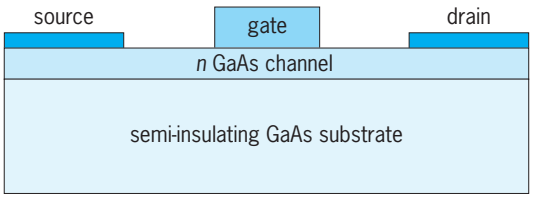

Fig. 2. Gallium arsenide metal semiconductor field-effect transistor (MESFET).

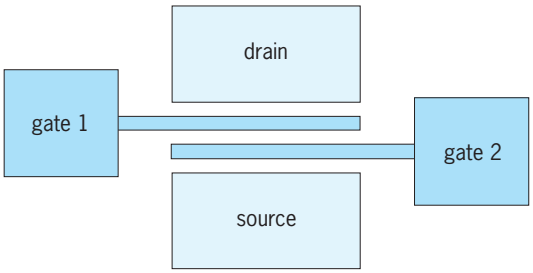

Fig. 3. Top view of a dual-gate MESFET structure.

Power devices consist of a number of MESFETs in parallel with air bridges connecting the sources. GaAs MESFET devices are used in low-noise amplifiers (LNAs), Class C amplifiers, oscillators, and monolithic microwave integrated circuits. The performance of a GaAs FET is determined primarily by the gate width and length. The planar structure of a MESFET makes it straightforward to add a second gate which can be used to control the amplification of the transistor (**Fig. 3**). Dual-gate MESFETs can be used as mixers (with conversion gain) and for control purposes. Applications include heterodyne mixers and amplitude modulation of oscillators. *See* AMPLIFIER; HETERODYNE PRINCIPLE; MIXER; OSCILLATOR.

The MOSFET has a highly insulating silicon dioxide (SiO_2) layer between the semiconductor and the gate; however, silicon MOSFETs are not really considered microwave transistors. Compared with the GaAs MESFET, MOSFETs have lower electron mobility (a decade smaller, resulting in much longer transmit times), larger parasitic resistances, and higher noise levels. Also, since the silicon substrate cannot be made semi-insulating, larger parasitic capacitances result. MOSFETs therefore do not perform very well above 1 GHz. Below this frequency, MOSFETs find application mainly as radio-frequency (RF) power amplifiers where silicon MOS is competitive. Silicon lateral DMOS (diffusion metal-oxide semiconductor) is finding uses in the 10–200-watt power range for cellular base stations operating in the 0.8–1.8 GHz frequency range.

Heterojunction FETs. A disadvantage of the MESFET is that the electron mobility is degraded since electrons are scattered by the ionized impurities in the channel. By using a heterojunction consisting of *n*-type AlGaAs with undoped GaAs, electrons move from the AlGaAs to the GaAs and form a conducting channel

at the interface. The electrons are separated from the donors and have the mobility associated with undoped material. A heterojunction transistor made in this fashion has many different names: high electron mobility transistor (HEMT), two-dimensional electron gas FET (TEGFET), modulation-doped FET (MODFET), selectively doped heterojunction transistor (SDHT), and heterojunction FET (HFET). The HEMT has high power gain at frequencies of 100 GHz or higher with low noise levels. HEMTs are fabricated using ion implantation, molecular beam epitaxy (MBE), or metal organic chemical vapor deposition (MOCVD). GaAs/AlGaAs/InAlGaAs pseudomorphic HEMTs (PHEMTs) have higher mobility than standard MESFETs or HEMTs. In many applications, PHEMTs are replacing MESFETs. Compound semiconductor FETs have low noise levels. The main sources of noise are thermal noise of the parasitic resistances, thermal velocity fluctuations in the channel, fluctuations in the thickness of the channel, and scattering of electrons between valleys in the conduction band at high electric field strengths. *See* CRYSTAL GROWTH; ION IMPLANTATION.

Microwave integrated circuit amplifiers. A monolithic microwave integrated circuit (MMIC) can be made using silicon or GaAs technology with either BJTs or FETs. For high-frequency applications, GaAs FETs are the best choice. A MMIC has both the active and passive devices fabricated directly on the substrate. MMICs are typically used as low-noise amplifiers, as mixers, as modulators, in frequency conversion, in phase detection, and as gain block amplifiers. Silicon MMIC devices operate in the 100-MHz to 3-GHz frequency range. GaAs FET MMICs are typically used in applications above 1 GHz.

Active microwave diodes. Active microwave diodes differ from passive diodes in that they are used as signal sources to generate or amplify microwave frequencies. These include step-recovery, tunnel, Gunn, avalanche, and transit time diodes, such as impact avalanche and transit-time (IMPATT), trapped plasma avalanche triggered transit-time (TRAPATT), barrier injection transit-time (BARITT), and quantum well injection transit time (QWITT) diodes.

Step recovery diodes. A step recovery diode is a special PIN type in which charge storage is used to produce oscillations. When a diode is switched from forward to reverse bias, it remains conducting until the stored charge has been removed by recombination or by the electric field. A step recovery diode is designed to sweep out the carriers by an electric field before any appreciable recombination has taken place. Thus, the transition from the conducting to the nonconducting state is very fast, on the order of picoseconds. Because of the abrupt step, this current is rich in harmonics, so these diodes can be used in frequency multipliers. *See* FREQUENCY MULTIPLIER.

Negative resistance and transferred electron devices. For microwave power generation or amplification, a negative differential resistance (NDR) characteristic at microwave frequencies is necessary. NDR is a phenomenon that occurs when the voltage (V) and current (I) are 180° out of phase. NDR is a dynamic property occurring only under actual circuit conditions; it is not static and cannot be measured with an ohmmeter. Transferred electron devices (TEDs), such as Gunn diodes, and avalanche transit-time devices use NDR for microwave oscillation and amplification. TEDs and avalanche transit-time devices today are among the most important classes of microwave solid-state devices. *See* NEGATIVE-RESISTANCE CIRCUITS.

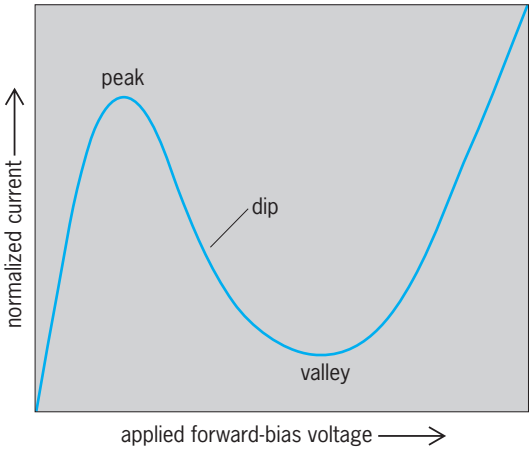

Fig. 4. Typical tunnel diode current-voltage (*I-V*) curve illustrating the negative differential resistance (NDR) phenomenon.

Tunnel diodes. The tunnel diode uses a heavily doped abrupt *pn* junction resulting in an extremely narrow junction that allows electrons to tunnel through the potential barrier at near-zero applied voltage. This results in a dip in the current-voltage (*I-V*) characteristic, which produces NDR (**Fig. 4**). Because this is a majority-carrier effect, the tunnel diode is very fast, permitting response in the millimeter-wave region. Tunnel diodes produce relatively low power. The tunnel diode was the first semiconductor device type found to have NDR. *See* TUNNEL DIODE; TUNNELING IN SOLIDS.

A double-barrier resonant tunneling (DBRT) diode can be grown by MBE with thin layers of AlGaAs and GaAs having sharp interfaces. The layer widths are comparable to the Schrödinger wavelength of the electron, thus permitting resonant behavior. The *I-V* characteristic of a DBRT diode is quite similar to that of the ordinary tunnel diode, showing NDR. Since the structure is symmetric, the *I-V* characteristic is the same for both positive and negative bias. Asymmetry can be achieved using barriers of different widths or heights.

Avalanche diodes. Avalanche diodes are junction devices that produce a negative resistance by appropriately combining impact avalanche breakdown and charge-carrier transit time effects. Avalanche breakdown in semiconductors occurs if the electric field is high enough for the charge carriers to acquire sufficient energy from the field to create electron-hole pairs by impact ionization. For silicon and GaAs,

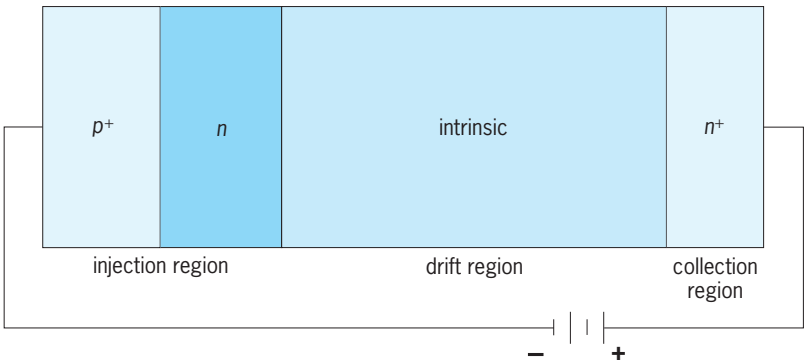

Fig. 5. Impact avalanche and transit-time (IMPATT) p^+nin^+ diode structure and DC bias (Read design).

the semiconductors commonly used to fabricate avalanche diodes, the threshold field for breakdown is on the order of a few hundred thousand volts per centimeter.

The avalanche diode is a *pn*-junction diode reverse-biased into the avalanche region. By setting the DC bias near the avalanche threshold, and superimposing on this an alternating voltage, the diode will swing into avalanche conditions during alternate half-cycles. The hole-electron pairs generated as a result of avalanche action make up the current, with the holes moving into the *p* region, and the electrons into the *n* region. The carriers have a relatively large distance to travel through the depletion region. At high frequencies, where the total time lag for the current is comparable with the period of the voltage, the current pulse will lag the voltage. By making the drift time of the electrons in the depletion region equal to one-half the period of the voltage, the current will be 180° out of phase. This shift in phase of the current with respect to the voltage produces NDR, so that the diode will undergo oscillations when placed in a resonant circuit.

Gunn diodes. A Gunn diode is typically an *n*-type compound semiconductor, such as GaAs or InP, which has a conduction band structure that supports negative differential mobility. Although this device is referred to as a Gunn diode, after its inventor, the device does not contain a *pn* junction and can be viewed as a resistor below the threshold electric field (E_{thres}). For applied voltages that produce electric fields below E_{thres}, the electron velocity increases as the electric field increases according to Ohm's law. For applied voltages that produce electric fields above E_{thres}, conduction band electrons transfer from a region of high mobility to low mobility, hence the general name "transferred electron device." Beyond E_{thres}, the velocity suddenly slows down due to the significant electron transfer to a lower mobility band producing NDR. For GaAs, E_{thres} is about 3 kV/cm. The Gunn effect can be used up to about 80 GHz for GaAs and 160 GHz for InP. Two modes of operation are common: nonresonant bulk (transit-time) and resonant limited space-charge accumulation (LSA). *See* ELECTRIC FIELD.

In the transit-time mode a charge dipole, consisting of an electron accumulation and a depletion layer (Gunn domain), travels through the semiconductor at a frequency dependent on the length of the semiconductor layer and the drift velocity. For typical lengths, the transit time is about 0.1 nanosecond corresponding to a frequency of 10 GHz. The transit-time mode is very inefficient—power efficiencies are about 1–5%. In the LSA mode, the frequency of operation is set by a resonant circuit to be much higher than the transit-time frequency so that domains have insufficient time to form while the field is above threshold. As a result, most of the sample is maintained in the negative conductance state during a large fraction of the voltage cycle. This is a distinct advantage in achieving efficient conversion of power in the upper microwave range. LSA diodes can be operated over a very wide range of frequencies by choosing the proper doping concentration. The efficiency of the LSA mode can be as high as 20%, and LSA diodes can be operated at very high frequencies (100 GHz and above).

IMPATT diodes. Impact avalanche and transit-time diodes (IMPATTs) are NDR devices that operate by a combination of carrier injection and transit time effects. There are several versions of IMPATT diodes, including simple reverse-biased *pn* diodes, complicated reverse-biased multidoped *pn* layered diodes, and reverse-biased PIN diodes including those of the Read design (**Fig. 5**). The IMPATT must be connected to a resonant circuit. At bias turn-on, noise excites the tuned circuit into a natural oscillation frequency. This voltage adds algebraically across the diode's reverse-bias voltage. Near the peak positive half-cycle, the diode experiences impact avalanche breakdown. When the voltage falls below this peak value, avalanche breakdown ceases. A 90° shift occurs between the current pulse and the applied voltage in the avalanche process. A further 90° shift occurs during the transit time, for a total 180° shift which produces NDR. An IMPATT oscillator has higher output power than a Gunn equivalent. However, the Gunn oscillator is relatively noise-free, while the IMPATT is noisy due to avalanche breakdown.

TRAPATT diodes. A trapped plasma avalanche triggered transit-time (TRAPATT) diode is basically a modified IMPATT diode in which the holes and electrons created by impact avalanche ionization multiplication do not completely exit from the transit domain of the diode during the negative half-cycle of the microwave signal. These holes and electrons form a plasma which is trapped in the diode and participates in producing a large microwave current during the positive half-cycle.

BARRITT diodes. A barrier injection transit-time diode (BARRITT) is basically an IMPATT structure that employs a Schottky barrier formed by a metal semiconductor contact instead of a *pn* junction to create similar avalanche electron injection.

QWITT diodes. A major drawback of the IMPATT diode is the fact that the avalanche process, which depends on random impact ionization events, is inherently noisy. A variety of approaches have been investigated to find alternative methods for injecting

carriers into the drift region without relying on the avalanche mechanism. Quantum well injection transit-time diodes (QWITT) employ resonant tunneling through a quantum well to inject electrons into the drift region. The device structure consists of a single GaAs quantum well located between two AlGaAs barriers in series with a drift region of made of undoped GaAs. This structure is then placed between two n^+-GaAs regions to form contacts.

Laurence P. Sadwick

Bibliography. K. Chang, *Microwave Solid-State Circuits and Applications*, Wiley-Interscience, 1994; S. Y. Liao, *Microwave Devices and Circuits*, 3d ed., Prentice Hall, 1996; T. G. Van De Roer, *Microwave Electronic Devices*, Chapman & Hall, 1994; J. S. Yuan, *SiGe, GaAs, and InP Heterojunction Bipolar Transistors*, Wiley-Interscience, 1999.

Microwave spectroscopy

The study of the interaction of matter and electromagnetic radiation in the microwave region of the spectrum. Microwaves are loosely defined as electromagnetic radiation with wavelengths between about 1 mm and 30 cm or frequencies between 1 and 300 GHz. The wavelengths are comparable to the dimensions of experimental apparatus. Experimental techniques make use of ideas from radio-frequency spectroscopy where wavelengths greatly exceed the dimensions of the apparatus, and also techniques from optics where wavelengths are much smaller than the size of the apparatus. *See* MICROWAVE; SPECTROSCOPY.

Apparatus. Microwave circuit elements (waveguides, resonant cavities, directional couplers) cannot be characterized by lumped capacitances (electric field regions) or lumped inductances (magnetic field regions), as in the case of radio-frequency circuits. Because of the short wavelengths of microwaves, both electric and magnetic fields are present in most circuit elements. When many wavelengths are present within a microwave circuit element, the momentum of the electromagnetic wave can become well defined, and momentum conservation or phase-matching conditions can have an important bearing on some spectroscopic techniques. An example is the mixing of microwaves and light waves, where the frequency of the mixed wave must equal the sum of the frequencies of the microwave and the light wave, while the momentum of the mixed wave (the inverse wavelength) must equal the vector sum of the momenta of the microwave and the light wave. *See* CAVITY RESONATOR; DIRECTIONAL COUPLER; WAVEGUIDE.

The transit times of electrons in ordinary electronic tubes are too long for efficient coupling to the rapidly oscillating fields of the microwave region, and special electronic tubes, klystrons, traveling-wave tubes, magnetrons, and so forth have been designed to overcome transit time limitations. Such tubes are often locked in frequency to a high harmonic of a stable quartz crystal oscillator. Solid-state microwave devices are increasingly practical alternatives to electron tube devices for some spectroscopic applications. At the very highest microwave frequencies, alternate methods like time-domain spectroscopy with short-pulse sources and optical analogs like Fourier-transform spectroscopy are more practical than conventional microwave methods. Microwave receivers are extraordinarily sensitive, and 10^{-19} W of microwave power can be detected with a good heterodyne system in a 1-Hz bandwidth. *See* GYROTRON; INFRARED SPECTROSCOPY; KLYSTRON; MAGNETRON; MICROWAVE SOLID-STATE DEVICES; MICROWAVE TUBE; TRAVELING-WAVE TUBE.

Interaction of microwaves with matter. The interaction of microwaves with matter can be detected by observing the attenuation or phase shift of a microwave field as it passes through matter. These are determined by the imaginary or real parts of the microwave susceptibility (the index of refraction). The absorption of microwaves may also trigger a much more easily observed event like the emission of an optical photon in an optical double-resonance experiment or the deflection of a radioactive atom in an atomic beam. *See* MOLECULAR BEAMS.

Microwave energy at a frequency ν is absorbed or emitted according to the Bohr frequency condition, Eq. (1), where h is Planck's constant and E_f

$$h\nu = E_f - E_i \tag{1}$$

and E_i are the energies of the final and initial states of the absorbing system. The initial and final states may be discrete as in the case of rotational states of a molecule, or they may be continuous as in bremsstrahlung in a plasma.

Enhancement of population differences. The characteristic temperature θ of the energy splitting involved in a microwave transition is given by Eq. (2), where

$$\theta = \frac{h\nu}{k} = 0.05\text{–}15\ \mathrm{K} \tag{2}$$

k is Boltzmann's constant. At room temperature T, the relative difference between the populations, N_i and N_f, of the states involved in the transition, $(N_i - N_f)/(N_i + N_f) = \tanh\,[\theta/(2T)] \sim \theta/(2T)$, is a few percent or less. The population difference can be close to 100% at liquid helium temperatures, and microwave spectroscopic experiments are often performed at low temperatures to enhance population differences and to eliminate certain line-broadening mechanisms.

The population differences between the states involved in a microwave transition can be enhanced by artificial means. For example, state selection of atoms or molecules in an atomic beam passing through inhomogeneous magnetic or electric fields can lead to very large population imbalances. Such large population inversions can be prepared so that stimulated emission cross sections for microwaves can exceed the absorption cross section. When the molecules or atoms with inverted populations are placed in an appropriate microwave cavity, microwave oscillations will build up spontaneously in the cavity

through maser (microwave amplification by stimulated emission of radiation) action. Optical pumping can also lead to large artificial population imbalances between the states involved in a microwave transition. *See* MASER; OPTICAL PUMPING.

Applications. The magnetic dipole and electric quadrupole interactions between the nuclei and electrons in atoms and molecules can lead to energy splittings in the microwave region of the spectrum. Thus, microwave spectroscopy has been used extensively for precision determinations of spins and moments of nuclei. The field-independent component of the 9.192-GHz hyperfine transition in the cesium atom is used as a time standard in many laboratories, and the second is defined as the time needed for a free cesium atom at zero magnetic field to make 9,192,631,770 oscillations. *See* ATOMIC CLOCK; HYPERFINE STRUCTURE; NUCLEAR MOMENTS; TIME.

Properties of molecules. The rotational frequencies of molecules often fall within the microwave range, and microwave spectroscopy has contributed a great deal of information about the moments of inertia, the spin coupling mechanisms, and other physical properties of molecules. The rotational frequencies of water and oxygen molecules are responsible for much of the attenuation of microwaves in the atmosphere. Many microwave transitions from simple atoms and molecules are seen from sources in interstellar space, and microwave astronomy has provided much information about the chemistry and molecular composition of various astronomical objects. Inversion frequencies in molecules, for example, the frequency of periodic motion of the nitrogen atom at the apex of the pyramidal NH_3 ammonia molecule through the three hydrogen atoms at the base, often lie in the microwave region. One of the inversion transitions in ammonia was used in the first maser. *See* INTERSTELLAR MATTER; MOLECULAR STRUCTURE AND SPECTRA.

Electron-spin resonance. The magnetic resonance frequencies of electrons in fields of a few thousand gauss (a few tenths of a tesla) lie in the microwave region. Thus, microwave spectroscopy is used in the study of electron-spin resonance or paramagnetic resonance. In the simplest cases, the spins may be well isolated in a dilute gas like oxygen (O_2) or cesium vapor. Then the microwave spectrum provides information about the internal spin couplings of a free atom or molecule. For denser gases or condensed phases, the paramagnetic resonance spectrum provides information about interactions between the spins and their environment. Both static interactions, which determine the resonance frequencies of the microwave transitions, and fluctuating random interactions, which determine resonance line widths, are of interest. Typical static interactions are the electrostatic interactions between the spin and the crystal field. Magnetic dipole interactions and spin exchange interactions between spins at neighboring sites can lead to broadening or narrowing of the resonance lines. Particularly strong interactions between neighboring spins occur in ferromagnetic or antiferromagnetic materials, both of which can be investigated by microwave spectroscopy. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; MAGNETIC RESONANCE.

Cyclotron resonance. The cyclotron resonance frequencies of electrons in solids at magnetic fields of a few thousand gauss (a few tenths of a tesla) lie within the microwave region of the spectrum. The effective mass and the cyclotron frequency of an electron in a solid are greatly modified from those of the free electron by the crystal interactions. Microwave spectroscopy has been used to map out the dependence of the effective mass on the electron momentum. *See* CYCLOTRON RESONANCE EXPERIMENTS.

Plasmas. Gaseous plasmas are strong emitters and absorbers of microwaves. The main coupling mechanisms of charged-particle motion to microwave radiation are bremsstrahlung and cyclotron radiation. Thomson scattering is also important in dense plasmas. Microwaves are completely reflected from plasmas where the electron density n is so high that the plasma resonance frequency f given by Eq. (3) exceeds the microwave frequency. Electron densities and temperatures in plasmas can be determined by microwave spectroscopy. *See* PLASMA (PHYSICS).

$$f = 8980^{1/2}\ \mathrm{Hz} \cdot \mathrm{cm}^{-3/2} \tag{3}$$

Cosmic microwave radiation. Thermal microwaves at a temperature of about 3 K permeate all space and can be detected by sensitive microwave receivers. These microwaves are believed to be the cooled electromagnetic radiation left over from the big bang at the beginning of the universe. *See* COSMIC BACKGROUND RADIATION.

William Happer

Bibliography. G. W. Chantry, *Modern Aspects of Microwave Spectroscopy,* 1980; G. Grüner (ed.), *Millimeter and Submillimeter Wave Spectroscopy of Solids*, Berlin, 1998; C. Kittel, *Introduction to Solid State Physics*, 8th ed., 2005; E. A. Oks, *Plasma Spectroscopy: The Influence of Microwave and Laser Fields*, 1995; C. P. Poole and H. A. Farach, *Handbook of Electron Spin Resonance*, 2 vols., 1994, 1999; C. H. Townes and A. L. Schawlow, *Microwave Spectroscopy*, 1955, reprint 1975.

Microwave tube

A high-vacuum tube designed for operation in the frequency region from approximately 3000 to 300,000 megahertz. Two considerations distinguish a microwave tube from vacuum tubes used at lower frequencies: the dimensions of the tube structure in relation to the wavelength of the signal that it generates or amplifies, and the time during which the electrons interact with the microwave field.

Effect of tube geometry. In a vacuum tube the active region is where the electrons travel through the evacuated space from the cathode, through the grid, to the plate (**Fig. 1**). The circuit in which the tube

operates extends from this active region along the internal tube structure, through the enclosing vacuum-tight envelope, and onto the portions of the circuit external to the tube. In the microwave region wavelengths are in the order of centimeters; resonant circuits are in the forms of transmission lines that extend a quarter of a wavelength from the active region of the microwave tube. With such short circuit dimensions the internal tube structure constitutes an appreciable portion of the circuit.

For these reasons a microwave tube is made to form part of the resonant circuit. Leads from electrodes to external connections are short, and electrodes are parts of surfaces extending through the envelope directly to the external circuit that is often a coaxial transmission line or cavity. Design of the tube and of the circuit in which it is to operate thus become closely related. *See* CAVITY RESONATOR; TRANSMISSION LINES.

Effect of transit time. At frequencies well below the microwave region, the time during which any one electron travels through the active region within a tube is so short compared to the period of the signal as to be negligible. At microwaves the period of the signal is in the range of 0.001–1 nanosecond. If transit time is comparable to the period of the signal, an electron has an inappreciable net change in energy. Even if transit time is reduced to half the signal period, an electron that is in transit when the signal reverses polarity experiences little net change in energy (**Fig. 2**). Only if transit time is less than a quarter of the signal period do significant numbers of electrons exchange appreciable energy with the signal field.

Transit time is reduced in several ways in microwave tubes. Electrodes are closely spaced and made planar in configuration. Spacings of 0.025 mm (0.001 in.) are practical, but to be effective require that electrodes be closely parallel to each other. High interelectrode voltages also decrease transit time by their acceleration of electrons; however, the voltage stresses that glass vacuum seals can withstand place a practical limit on the voltage. *See* VACUUM TUBE.

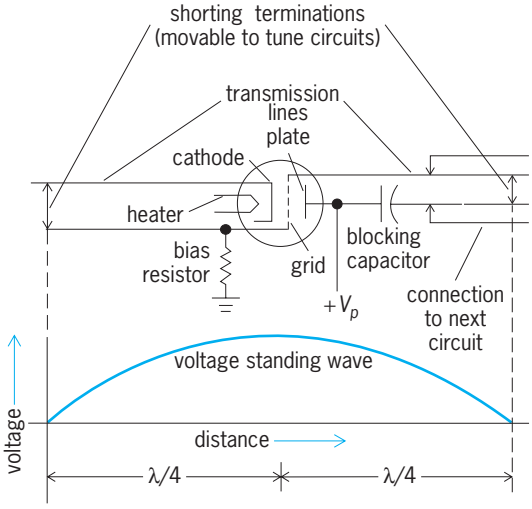

Fig. 1. Schematic of microwave triode in a circuit of two transmission lines. Usually the tube has circular symmetry and fits directly into coaxial lines.

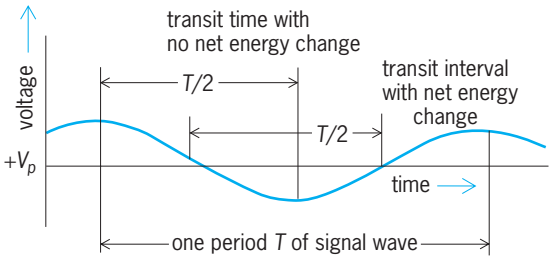

Fig. 2. Diagram of transit time. For efficient tube operation, the transit time during which an electron passes through the signal field needs to be short compared to the period of the signal.

Alternative designs. Tubes designed by the foregoing principles are effective for wavelengths from a few meters to a few centimeters. At longer wavelengths lumped-constant circuits are effective and tubes can be designed for optimum internal characteristics. At shorter wavelengths different principles are necessary. To obtain greater exchange of energy between the electron beam and the electromagnetic field several alternative designs have proved practical.

Instead of collecting the electron beam at a plate formed by the opposite side of the resonant circuit, the beam is allowed to pass into a field-free region before reacting further with an external circuit. The electron could can be deflected by a strong static magnetic field so as to revolve and thereby react several times with the signal field before reaching the plate. *See* KLYSTRON; MAGNETRON.

Instead of producing the field in one or several resonant circuits, the field can be supported by a distributed structure along which it moves at a velocity comparable to the velocity of electrons in the beam. The electron beam is then directed close to this structure so that beam and field interact over an extended interval of time. *See* TRAVELING-WAVE TUBE.

Such structures as these greatly extend the region over which useful gains, low signal-noise ratios, and significant powers can be produced, although the design of any one microwave tube can usually be optimized for only one of these characteristics at a time. Even so, dimensions and tolerances limit the wavelengths for which such tubes can be manufactured. For operation at shorter wavelengths entirely different techniques are necessary, such as those using quantum behavior within molecules or the intermodulation of signals within a nonlinear device. *See* MASER; MICROWAVE SOLID-STATE DEVICES; PARAMETRIC AMPLIFIER.

Frank H. Rockett

Bibliography. R. E. Collin, *Foundations for Microwave Engineering*, 2d ed., 2000; A. S. Gilmour, Jr., *Microwave Tubes*, 1986; C. A. Lee and G. C. Delman, *Microwave Devices, Circuits and Their Interaction*, 1994; S. Y. Liao, *Microwave Devices and Circuits*, 3d ed., 1990; D. M. Pozar, *Microwave Engineering*, 2d ed., 1997.

Mictacea

A proposed order of the Peracarida established for two small crustacean species, *Hirsutia bathyalis* and *Mictocaris halope* (see **illus.**). The two species share many features common to other peracaridans but differ sufficiently to justify their assignment to a distinct order with two monotypic families. Common peracaridan features include a brood pouch formed by basal lamellae of the perepods (oöstegites) in the female; a small movable process (lacinia mobilis) on the mandible; free thoracic somites not fused to a carapace shield; a single maxilliped of typical peracarid form; and partially immobile pereopodal basal segments.

The body is slender and cylindrical, with the cephalon (head) fused to the first thoracic (trunk) somite. A carapace as such is not present, but small lateral carapace folds cover the bases of the postmandibular appendages. Eyes appear to be completely lacking in *Hirsutia*. Eyestalks are present in *Mictocaris*, but again no trace of visual elements has been found. The first thoracic appendage is modified as a maxilliped, whereas pereopods 1–5 in *Mictocaris* or 2–6 in *Hirsutia* are biramous and provided with natatory exopods. Pereopod 7 of *Mictocaris* males is provided with a copulatory structure (males of *Hirsutia* have not been seen). Although both species possess a ventral brood pouch formed from oöstegites, those of *Mictocaris* lack the marginal setae typical of most peracaridans. The oöstegites of *Hirsutia*, while having marginal setae, differ from other peracaridans in that the oöstegites arise posteriorly from pereopod segments rather than medially. Small, one-segmented pleopods are present on the first five abdominal somites in *Mictocaris*, but the identity of similar structures on the abdomen of *Hirsutia* is questionable since no suture separating them from the body wall has been identified. The sexes are distinct in *Mictocaris*, with eggs hatching at the typical peracaridan manaca stage in which the seventh pereopod is not yet developed. *Hirsutia* is presently known from a single, presumably preparatory, female and one juvenile molt. Presumably it is also a sexually dimorphic species.

Mictocaris is a cave-dwelling species, whereas *Hirsutia* has been found only in soft muddy sediment of the deep sea. Nothing is known about the feeding habits of these crustaceans, but with its spined first pereopod and fossorial second, *Hirsutia* is thought to be carnivorous. In contrast, the feeding appendages in *Mictocaris* resemble those of thermosbaenaceans, which scrape food particles from the substrate.

Mictacea appear to be most closely related to the Thermosbaenacea, Spelaogriphacea, and Mysidacea. *See* CRUSTACEA; MYSIDA; PERACARIDA; SPELAEOGRIPHACEA; THERMOSBAENACEA. Patsy A. McLaughlin

Bibliography. T. E. Bowman and T. M. Iliffe, *Mictocaris halope*, a new unusual peracaridan crustacean from marine caves on Bermuda, *J. Crust. Biol.*, 5:58–73, 1985; T. E. Bowman et al., Mictacea, a new order of Crustacea Peracarida, *J. Crust. Biol.*, 5:74–78, 1985; H. L. Sanders, R. R. Hessler, and S. P. Garner, *Hirsutia bathyalis*, a new unusual deep-sea benthic peracaridan crustacean from the tropical Atlantic, *J. Crust. Biol.*, 5:30–57, 1985.

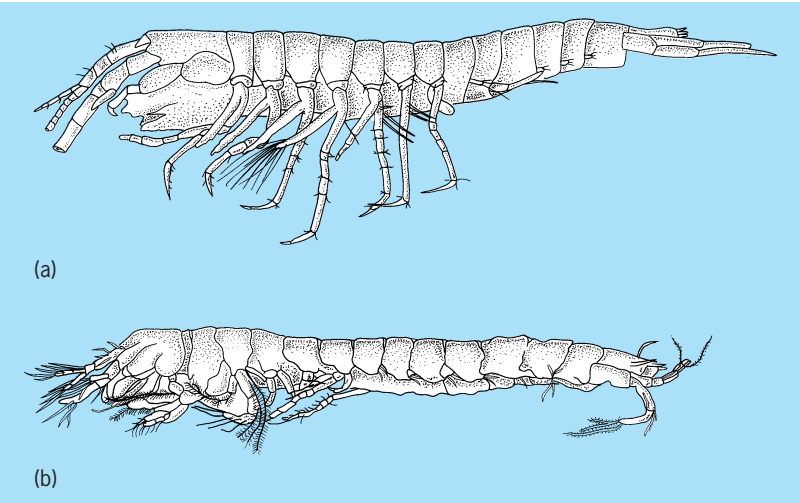

The two species of the order Mictacea. (*a*) *Hirsutia bathyalis* (*after H. L. Sanders et al., Hirsutia bathyalis, a new unusual deep-sea benthis peracaridan crustacean from the tropical Atlantic, J. Crust. Biol., 5:30–57, 1985*). (*b*) *Mictocaris halope* (*after T. E. Bowman and T. M. Iliffe, Mictocaris halope, a new unusual peracaridan crustacean from marine caves on Bermuda, J. Crust. Biol., 5:58–73, 1985*).

Middle-atmosphere dynamics

The motion of that portion of the atmosphere that extends in altitude roughly from 10 to 100 km (6 to 60 mi). Interest in the middle atmosphere has been fueled by concerns over the Earth's environment, and in particular by the discovery of the ozone hole in the Antarctic lower stratosphere, together with predictions of greenhouse warming. Among the most difficult and important tasks in atmospheric science are unequivocal determinations of long-term changes in both amount of ozone and greenhouse-induced global temperature. Both are intimately related to the composition, structure, and dynamical motions in the middle atmosphere.

The Earth's climate is determined by a balance between incoming solar and outgoing Earth thermal radiative energy, both of which must necessarily pass through the middle atmosphere. The lower portion, the stratosphere, contains many greenhouse gases (ozone, water vapor, carbon dioxide, methane, nitrous oxide, chlorofluorocarbons, as well as others); and it is predicted to cool at the same time as the lower atmosphere is warmed by the greenhouse effect. The middle atmosphere is also a focus for effects of emissions from proposed commercial fleets of stratospheric aircraft. In addition to the chemistry involved, dynamical transport modeling and measurements are needed to predict the widespread transport of these important trace gases and emissions over the globe. *See* GREENHOUSE EFFECT; STRATOSPHERE; TERRESTRIAL RADIATION.

Middle-atmosphere structure. The stratosphere comprises the lower part of the middle atmosphere, from about 10 to 50 km (6 to 30 mi) altitude; from about 50 to 80 km (30 to 48 mi) or so lies the mesosphere. The location of the base of the stratosphere (called the tropopause) depends on meteorological conditions, varying on average from about 10 km (6 mi) in altitude at the poles to about 16 km (10 mi) at the Equator.

Actually, the middle atmosphere is not so far removed from everyday experience. The flat-topped anvils of thunderstorms are formed when rapidly ascending air flattens out like a pancake against the bottom of the stratosphere; further vertical penetration is inhibited by negative buoyancy effects due to the rising temperature with height in the stratosphere. A second example is experienced by jet aircraft travelers on middle-latitude or polar routes that often are flown near or just below the tropopause; a traveler looking up into the darker sky from the aircraft window is peering into the stratosphere. A third example involves eruptions of volcanoes. Often after a major volcanic explosion, a beautiful twilight purple glow can be seen a few minutes after sunset, caused by dust ejected into the stratosphere perhaps half a world away, and transported by strong winds in the stratosphere. Except occasionally in winter polar conditions, precipitation does not occur in the stratosphere, and the tiny dust particles many remain there for a year or more.

In the lowest atmospheric layer (called the troposphere) where rain and storms occur, temperature usually decreases with height. The stratosphere is generally a very stable layer in which temperature increases with height. In equatorial regions the tropopause is characterized by, and usually defined by, a narrow altitude range over which temperature stops decreasing with height and begins to increase with height. This definition is not as useful at high latitudes because of different temperature structure. A more general, dynamical definition of the interface between lower and middle atmosphere is in terms of a quantity called potential vorticity, which has large values in the stratosphere but small values in the lower atmosphere. In many cases of interest, the potential vorticity is also an approximately conserved quantity, making it useful for tracing air motions of air parcels.

Ozone. About 90% of the Earth's ozone shield lies in the stratosphere. Ozone is produced mainly at low latitudes (near the Equator) by solar ultraviolet radiation; however, larger concentrations of ozone are found at middle and polar latitudes (**Fig. 1**). The resolution of this apparent anomaly lies in transport from the equatorial source to the polar regions by dynamical processes in the stratosphere. For example, wave-motion-induced transport moves significant amounts of stratospheric ozone to the Antarctic polar region; furthermore, the wave-induced stratospheric conveyance transports the ozone-destroying (mainly chlorine) compounds from low latitudes (where they enter the stratosphere from below) to the polar regions. Without such transport, there would not be a significant Antarctic ozone hole.

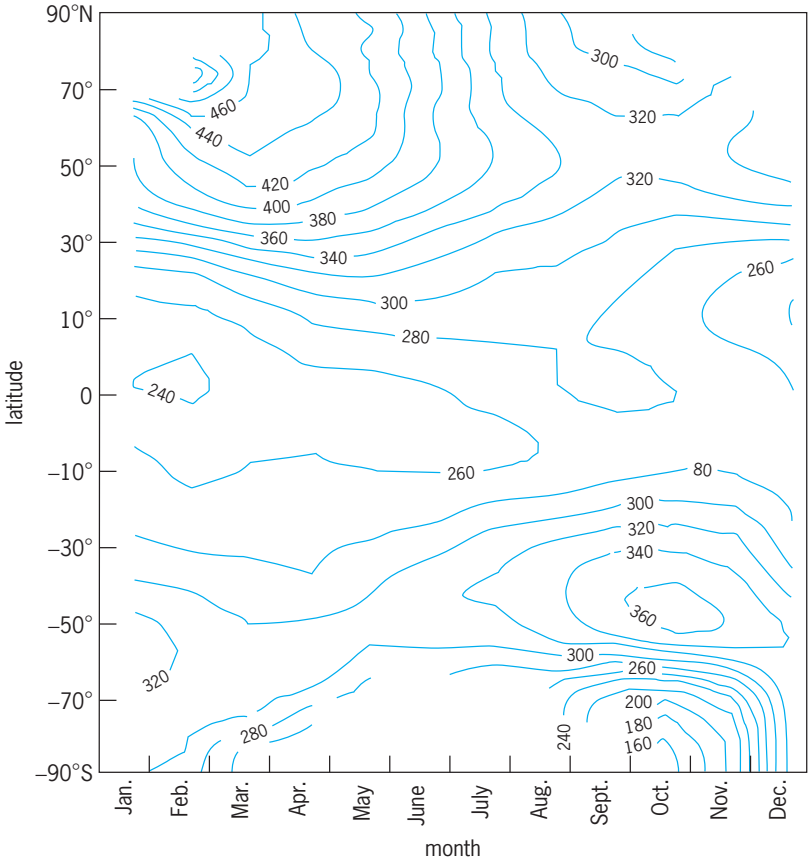

Fig. 1. Time-latitude section showing the seasonal variation of total ozone in Dobson units, based on the Total Ozone Mapping Spectrometer (TOMS) satellite instrument data for 1987. The maxima in middle to high latitudes of each hemisphere's respective spring are due primarily to dynamical transport from the equatorial source region. The very low values of ozone in October near 75–90°S represent the Antarctic ozone hole. (*After World Meterological Organization Global Ozone Research and Monitoring Project, Report of the International Ozone Trends Panel 1988, Rep. 18, vol. 1, 1988*)

Horizontal transport into or out of the polar stratosphere is greatly inhibited, however, when strong stratospheric jet winds develop into a winter vortex core centered over the pole. Shaped somewhat like a vertical cylinder, the winter vortex acts as a barrier to horizontal transport through its walls, dynamically isolating ozone and compounds relevant to ozone depletion. This isolation from outside influences during winter is a major ingredient in the formation of the Antarctic ozone hole, which forms at the end of the Antarctic polar night with the first sunlight in early spring. Although less continuous and less intense, a similar effect often occurs in the north polar winter stratosphere.

Understanding the dynamical and radiative physics underlying middle-atmosphere transport is thus critical for several aspects of predicting global climate change.

Middle-atmosphere waves. Atmospheric gravity waves (not to be confused with the gravity waves of general relativity) result from combined gravitational and pressure gradient forces. Typical characteristics are transverse polarization, vertical wavelengths of

0.1–10 km (0.06–6 mi), horizontal wavelengths of 1–100 km (0.6–60 mi) or more, and periods in the range of 5 min to several hours. These waves may be excited by airflow over orography (mountains) as standing lee waves, by growing clouds, and by large-scale storm complexes in the lower atmosphere; and then they propagate up into the middle atmosphere.

Planetary-scale Rossby waves are large and slowly moving waves affected by the Coriolis effect due to the Earth's rotation. They arise from the variation with latitude of the component of the Earth's rotation vector normal to the planetary surface, and they propagate against the planetary rotation (that is, westward) with periods of several days. They have vertical wavelengths of order 10 km (6 mi) and horizontal wavelengths from hundreds of kilometers up to as large as wave 1, where one wavelength fits around the Earth's circumference (wave 2 refers to two wavelengths fitting around the Earth's circumference, and so forth). Rossby waves are common at middle latitudes in winter, where they can propagate up into the middle atmosphere from excitation regions below. *See* CORIOLIS ACCELERATION.

Near the Equator, hybrid Rossby-gravity waves and also Kelvin waves (a special class of eastward-propagating internal gravity waves having no north-south velocity component) have been observed in the middle atmosphere.

A variety of global-scale normal-mode oscillations (somewhat like the vibrations of a kettledrum head) are also found in the middle atmosphere, prominent examples being wave 1, westward-moving waves with periods of about 5 and 16 days, and a wave 3, a westward moving feature with a period of about 2 days. Another observed oscillation in the middle atmosphere has been found with periods in the range 1–2 months (propagating up from the lower atmosphere).

Waves resulting from fluid dynamical instabilities are also observed: medium-scale (waves 4–7) eastward-moving waves, which are actually the tops of tropospheric storm systems, can dominate the circulation of the summer Southern Hemisphere lower stratosphere. The medium-scale waves have periods of 10–20 days. There are other possible instabilities as well. *See* DYNAMIC INSTABILITY.

Wave driving of the middle atmosphere. Perhaps the most subtle and complex middle-atmosphere phenomena are related to nonlinear effects of large-amplitude waves. An example is the mechanism responsible for the stratospheric-mesospheric jet stream. This striking feature of the middle atmosphere exhibits a nearly zonally symmetric (donut-shaped) vortex with strong winds that maximize in the upper stratosphere and lower mesosphere (**Fig. 2**). The stratospheric-mesospheric jet winds primarily blow from the west in the winter hemisphere and from the east in the summer hemisphere. The primary cause of the jet stream is a combination of the Equator-to-pole difference in solar heating and the Coriolis acceleration due to the Earth's daily rotation. Maximum observed winds (Fig. 2) are usually less than 100 m/s (around 200 mi/h), several times smaller than predicted by radiative transfer models without dynamics.

Radiative equilibrium model. If there were no wave dynamical effects, the structure of the Earth's middle atmosphere would be very different. A radiative equilibrium model using only incoming solar visible light and outgoing infrared radiation plus photochemistry was used to predict the atmospheric temperature field (**Fig. 3**). Wave dynamics were left out of this model on purpose to see what the radiative

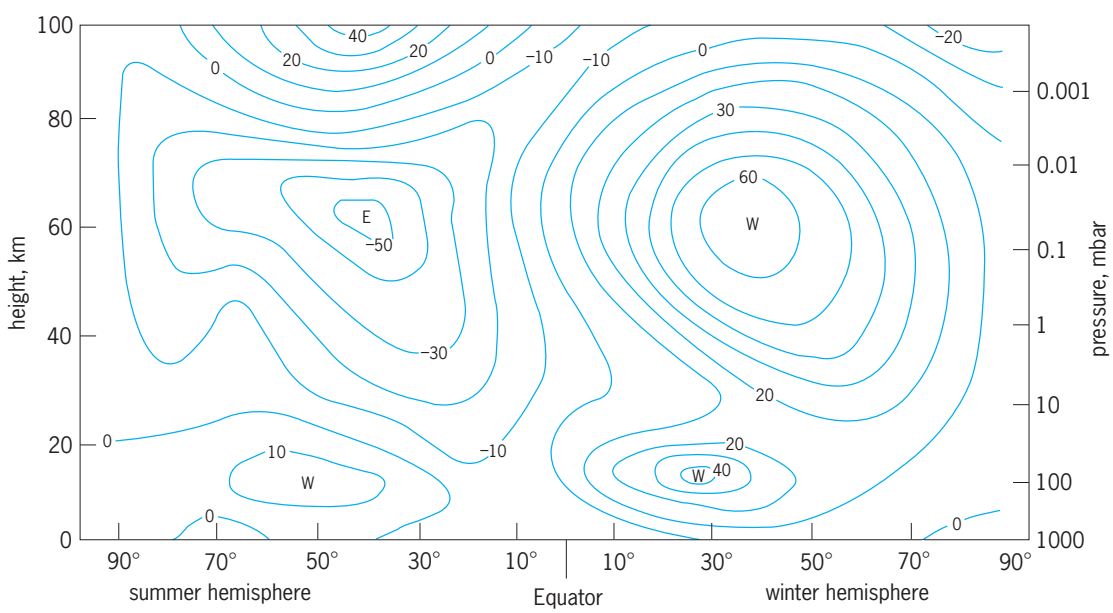

Fig. 2. Latitude-height section of typical observed zonal (east-west) mean wind speeds; contours in m/s (1 m/s ≅ 2 mi/h). W and E denote centers of wind from the west and east. Negative values denote winds from the east. Prominent jets maximize in upper stratosphere and lower mesosphere, decreasing to weak values near 90–100 km (50–60 mi) altitude. (*After D. G. Andrews, J. R. Holton, and C. B. Leovy, Middle Atmosphere Dynamics, Academic Press, 1987*)

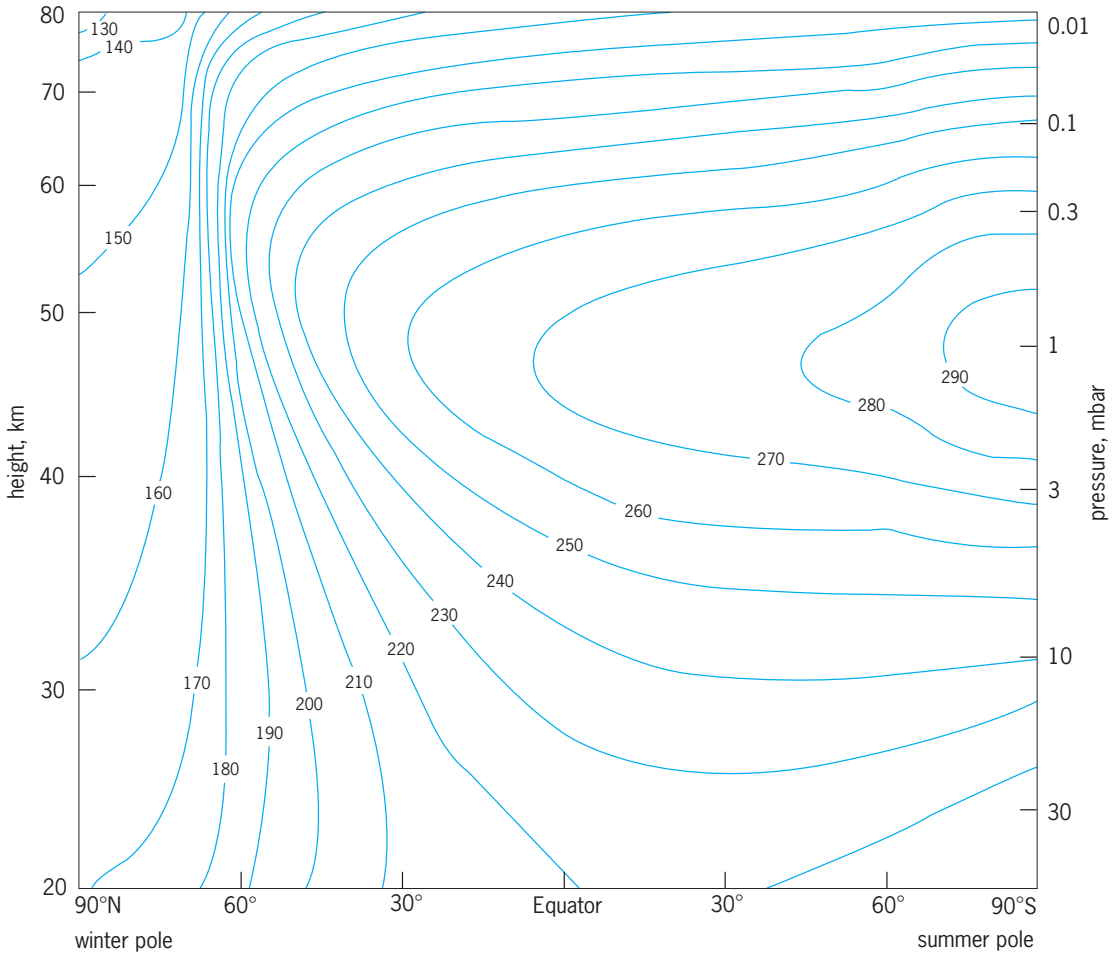

Fig. 3. Zonal mean middle-atmosphere temperatures (K) determined from a model with only radiative and photochemical effects. Vertical axis: height in kilometers (left), atmospheric pressure in millibars (right). °F = (K × 1.8) − 459.67. (*After S. B. Fels, Radiative-dynamical interactions in the middle atmosphere, Adv. Geophys., 28:277–300, 1985*)

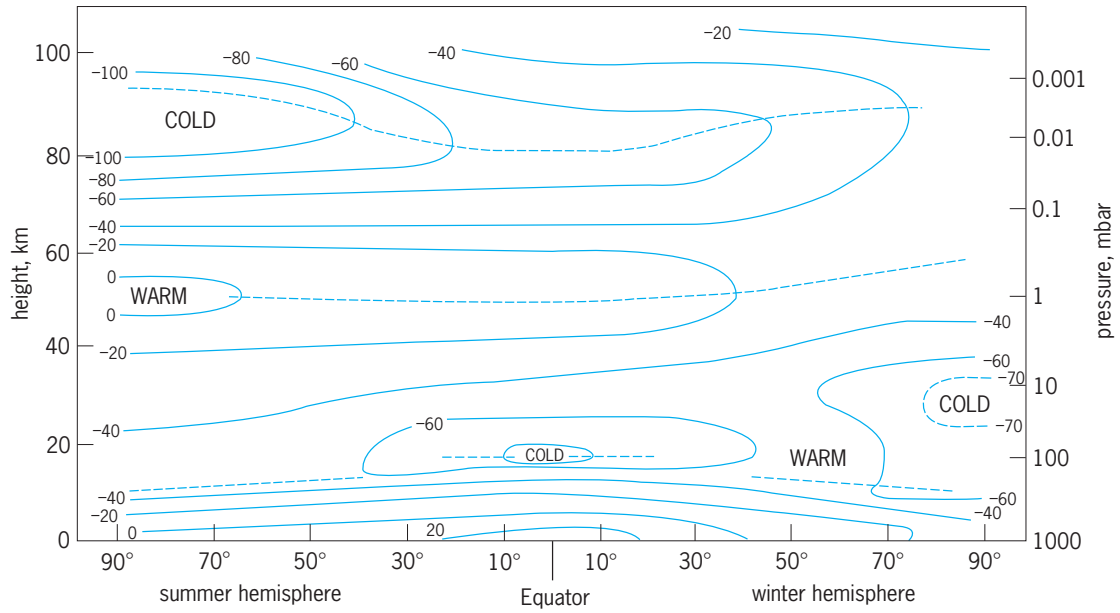

Fig. 4. Latitude versus altitude section of typical observed zonal mean temperatures for solstice conditions. Broken lines indicate, with increasing height, the tropopause, stratopause, and mesopause levels. The temperatures are in °C [°F = (°C × 1.8) + 32°], and winter and summer poles are reversed from those in Fig. 3. (*After D. G. Andrews, J. R. Holton, and C. B. Leovy, Middle Atmosphere Dynamics, Academic Press, 1987*)

equilibrium state would be. The radiative equilibrium state would have very cold temperatures in the sunless polar winter night, and a warm mesosphere (50–80 km or 30–50 mi altitude) in the continuous sunshine at the pole during its summer season. In contrast, observed mean atmospheric temperatures are nearly 100 K (180°F) warmer in the sunless winter mesosphere, and much colder in summer, than predicted by the model (**Fig. 4**).

Middle-atmosphere winds predicted with the radiative equilibrium model are also very different from observed winds (**Fig. 5**). The strong temperature gradients produce pressure gradients that couple with the Earth's rotation (Coriolis torque) to produce strong jet winds. The purely radiative atmosphere would have very high jet-wind speeds of 200–300 m/s (400–600 mi/h) in the upper mesosphere in winter, several times larger than observed jet winds (Fig. 2).

Wave drag. The mechanism that "closes" the jet (preventing extremely large winds) is thought to result from vertical propagation of atmospheric gravity waves, which were not included in the radiative model. Variations in storms and the jet stream in the lower atmosphere excite the atmospheric gravity waves, which then propagate upward into the middle atmosphere. The waves grow in amplitude with increasing height because of decreasing atmospheric density. In the mesosphere the wave amplitudes become so great that they break, somewhat analogous to the breaking surf of ocean waves. When this happens, the waves' momentum per mass (speed) is absorbed by the atmosphere at the breaking altitude. Because they are generated by slow-moving tropospheric storm systems or by airflow over fixed orography, the waves are moving much more slowly than the faster mesospheric jet. The absorption and mixing of the gravity waves causes the mesospheric jet winds to be decelerated. This process, known as wave drag, acts as a dynamical brake. Without it the middle atmosphere jet would reach the very high speeds as predicted by the radiative-only model (Fig. 5).

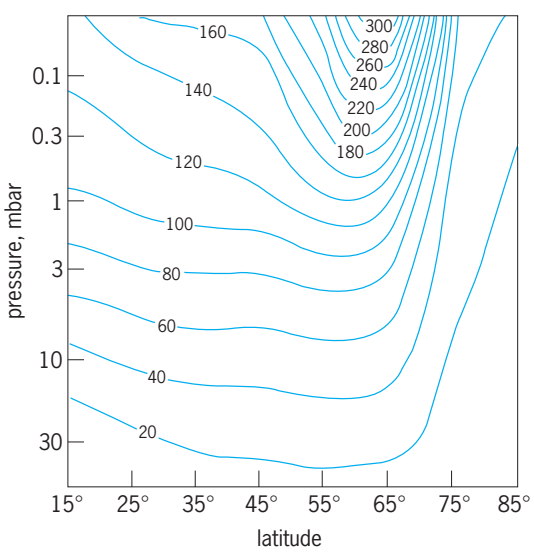

Fig. 5. Mean zonal (east-west) wind calculated from the radiative-equilibrium temperature field in Fig. 4. Contours in m/s (1 m/s ≅ 2 mi/h). (*After D. G. Andrews, J. R. Holton, and C B. Leovy, Middle Atmosphere Dynamics, Academic Press, 1987*)

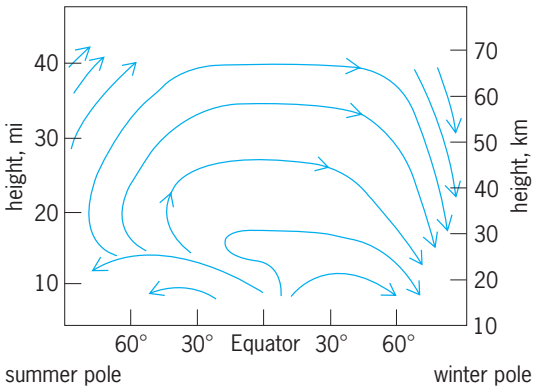

Fig. 6. Schematic flow in the mean meridional circulation of the middle atmosphere. (*After T. Dunkerton, On the mean meridional mass motions of the stratosphere and mesosphere, J. Atm. Sci., 35:2325–2333, 1978*)

In addition to wave drag, wave dynamics can alter the mean wind state of the atmosphere. There are a variety of wave types excited continually in the atmosphere. When these waves propagate vertically, they can encounter critical regions where the wave horizontal-phase speed matches the mean zonal background wind speed. Near such critical regions, waves are absorbed by the atmosphere, leading to dynamic forcing of the mean wind state. Studies have shown that the mean background jet wind can exhibit a seesaw pattern, increasing in strength when waves are absorbed, and decreasing when energy is drawn out of the jet to strengthen waves again. *See* WAVE MOTION.

Wave dynamics also induces a slow but important average meridional circulation of the atmosphere (**Fig. 6**). The basic structure of the meridional circulation has been confirmed, with meridional circulation at high altitudes from summer to winter hemisphere and downward transport into the upper stratosphere at the winter pole. After 1–2 months, air originally high in the summer mesosphere is transported across the Equator and descends deep into the winter polar stratosphere. Although the circulation resembles a simple hot-air-rises cold-air-sinks pattern, in reality the phenomenon requires a rather complex, nonintuitive dynamical wave-driving explanation. *See* JET STREAM; MESOSPHERE.

Sudden stratospheric warming. One of the most dramatic events in the stratosphere occurs about once every other year when the Northern Hemisphere winter polar stratosphere suddenly experiences dramatic warming in a few days. Such a sudden stratospheric warming may lead to a rapid reversal in direction of the stratospheric winter jet, changing it from its usual winter eastward motion (Fig. 2) to westward. Only once has a sudden stratospheric warming been observed in the Southern Hemisphere, the difference between hemispheres

apparently being related to stronger planetary-scale waves (waves 1 and 2) forced by the effects of more extensive mountain ranges and land-ocean temperature contrasts in the Northern Hemisphere.

Quasibiennial oscillation. An unusual, almost periodic, zonally symmetric, quasibiennial oscillation is found in the equatorial stratosphere, manifested as slowly descending winds at low latitudes, switching between westward and eastward directions with an irregular period averaging about 27 months.

The basic mechanism involves a subtle internal oscillation resulting from wave interactions with the mean stratospheric jet flow. Vertically propagating waves, such as gravity, Kelvin, and Rossby-gravity waves (and perhaps others), are radiatively or mechanically damped in the lower stratosphere, leading to the alternating wind regimes of the quasibiennial oscillation. Although the main idea and beauty of the fundamental principle is clear, aspects of the observations, especially why the two quasibiennial oscillation regimes are not symmetric, have yet to be elucidated.

Semiannual oscillation. Another middle-atmosphere phenomenon is the semiannual oscillation, with 6-month period. The equatorial semiannual oscillation, like the quasibiennial oscillation, is thought to involve vertical propagation and absorption of waves. Kelvin and gravity waves are thought to be important, with the latter playing the dominant role. Curiously, the semiannual oscillation winds are out of phase between the upper stratosphere and mesosphere. That is, when semiannual oscillation winds are from the west in the lower layer, they are from the east in the higher layer. More research is needed, but it is thought that this out-of-phase behavior is caused by wave absorption at critical regions in the stratosphere. The forcing waves are absorbed and thus prevented from propagating vertically, or they are allowed to do so, depending on whether the lower altitude winds are moving with or against the waves.

Waves due to instabilities. The dominant wave features in the Southern Hemisphere winter stratosphere are large scale (wave 2) and move eastward with periods of 2–3 weeks. These features are postulated to originate from instabilities that draw their energy from spatial variations in the background wind field. Instabilities have also been predicted to account for curious long-lived perturbations observed circling the pole in about 4 days in the winter middle atmosphere. *See* ATMOSPHERE; DYNAMIC METEOROLOGY; METEOROLOGICAL SATELLITES. J. L. Stanford

Bibliography. D. Andrews, A stratospheric transport system, *Phys. World*, 4(11):41–46, November, 1991; D. G. Andrews, J. R. Holton, and C. B. Leovy, *Middle Atmosphere Dynamics*, 1987; T. Dunkerton, On the mean meridional mass motions of the stratosphere and mesosphere, *J. Atm. Sci.*, 35:2325–2333, 1978; S. B. Fels, Radiative-dynamical interactions in the middle atmosphere, *Adv. Geophys.*, 28A:277–300, 1985; S. Solomon et al., Tracer transport by the diabatic circulation deduced from satellite observations, *J. Atm. Sci.*, 43:1603–1617, 1986.

Midnight sun

The phenomenon occurring when the Sun does not set, but only approaches the horizon at midnight. The effect occurs near the time of the summer solstice, on June 21, for latitudes north of the Arctic Circle. The same effect occurs near the time of the winter solstice, on December 21, for latitudes south of the Antarctic Circle. (Here "summer" and "winter" refer to the Northern Hemisphere; seasons are reversed in the Southern Hemisphere.)

Between the spring equinox and the summer solstice, in the Northern Hemisphere, the locations of sunrise and sunset move northward along the horizon. At the same time, the noontime sun climbs higher in the south each day. The highest altitude is reached at the summer solstice, after which the cycle reverses. On short summer nights, the Sun does not dip far below the northern horizon. A person who travels north along the Earth's curved surface inevitably reaches a point where the Sun becomes visible over the pole above the northern horizon; the midnight sun is observed.

The Earth orbits the Sun on a plane called the ecliptic. The Earth's Equator is inclined with the ecliptic by $23°26'$. As a result, the North and South poles are in turn inclined toward the Sun for 6 months. Close to the summer solstice, on June 21, the Northern Hemisphere reaches its maximum inclination toward the Sun and the Sun illuminates all the polar area down to latitude $+66°34'$. As seen from the polar area, the Sun does not set but only reaches its lowest altitude in the north at midnight. Latitude $+66°34'$ defines the Arctic Circle, which is the southernmost latitude in the Northern Hemisphere where the midnight sun can be observed near the summer solstice. However, atmospheric refraction raises objects on the horizon by about $34'$, and the midnight sun can therefore be seen for a few days from locations 80 km (50 mi) south of the Arctic Circle. Observers at latitudes above the Arctic Circle see the midnight sun higher above the northern horizon, or correspondingly, see the midnight sun before the summer solstice. For example, at latitude $+70°$, uninterrupted daylight starts 5 weeks before the summer solstice. At the North Pole, 6 month-long daylight starts around March 20, the moment of vernal equinox when the Sun travels above the Equator. *See* ECLIPTIC; METEOROLOGICAL OPTICS; REFRACTION OF WAVES.

Near the winter solstice, on December 21, the Earth reaches the opposite point in its orbit and the polar area in the Southern Hemisphere is now most inclined toward the Sun. The South Pole area is illuminated continuously, and the midnight sun phenomenon occurs south of about $-66°34'$. *See* EARTH ROTATION AND ORBITAL MOTION; SEASONS.

Pekka Parviainen

Bibliography. G. O. Abell, D. Morrison, and S. C. Wolff, *Exploration of the Universe*, 6th ed., 1991; J. M. Pasachoff, *Astronomy: From the Earth to the Universe*, 5th ed., 1999.

Mid-Oceanic Ridge

A largely interconnected system of broad submarine rises totaling at least 60,000–80,000 km (37,000–50,000 mi) long, the precise length depending on what is included and how it is measured. Thus the Mid-Oceanic Ridge is the longest mountain range system on the planet. The origin of the Mid-Oceanic Ridge is intimately connected with plate tectonics. Wherever plates move apart sufficiently far and fast for oceanic crust to form in the void between them, a branch of the Mid-Oceanic Ridge will be created (**Fig. 1**). The plate boundary of the Mid-Oceanic Ridge comprises an alternation of spreading centers (or axes or accreting plate boundaries) interrupted or offset by a range of different discontinuities, the most prominent of which are transform faults. As the plates move apart, new oceanic crust is formed along the spreading axes, and the ideal transform fault zones are lines along which plates slip past each other and where oceanic crust is neither created nor destroyed. Bruce Heezen, who pioneered studies of the Mid-Oceanic Ridge, once referred to this feature as a "wound that never heals."

The term Mid-Oceanic Ridge is somewhat a misnomer, having been applied before its tectonic significance (a system of spreading axes connected by transform faults) was appreciated. The ridge generally remains in the middle of an ocean basin only if that basin has formed between two continents rifted apart, and the average spreading rates on each flank have been the same. Some spreading axes are located near the edges of ocean basins and behind island arcs. *See* PLATE TECTONICS; TRANSFORM FAULT.

Separation of plates causes the hot upper mantle to rise along the spreading axes of the Mid-Oceanic Ridge; partial melting of this rising mantle generates magmas of basaltic composition that segregate from the mantle and rise in a narrow zone at the axis of the Mid-Oceanic Ridge to form the oceanic crust. The partially molten mantle "freezes" to the sides and bottoms of the diverging plates to form the mantle lithosphere that, together with the overlying "rind" of oceanic crust, comprises the lithospheric plate. At the axis of the Mid-Oceanic Ridge the underlying column of crust and mantle is hot and thermally expanded; this thermal expansion explains why the Mid-Oceanic Ridge is a ridge. With time, a column of crust plus mantle lithosphere cools and shrinks as it moves away from the ridge axis as part of the plate. The gentle regional slopes of the Mid-Oceanic Ridge (typically from 3 to 50 parts per thousand near the axis, and decreasing smoothly toward the flanks) therefore represent the combined effects of sea-floor spreading (divergent plate motion) and thermal contraction. The height and shape of the average Mid-Oceanic Ridge profile in meters of depth are approximated by the formula $D = 2900 + 350\sqrt{T}$, where T is the age of the crust in millions of years.

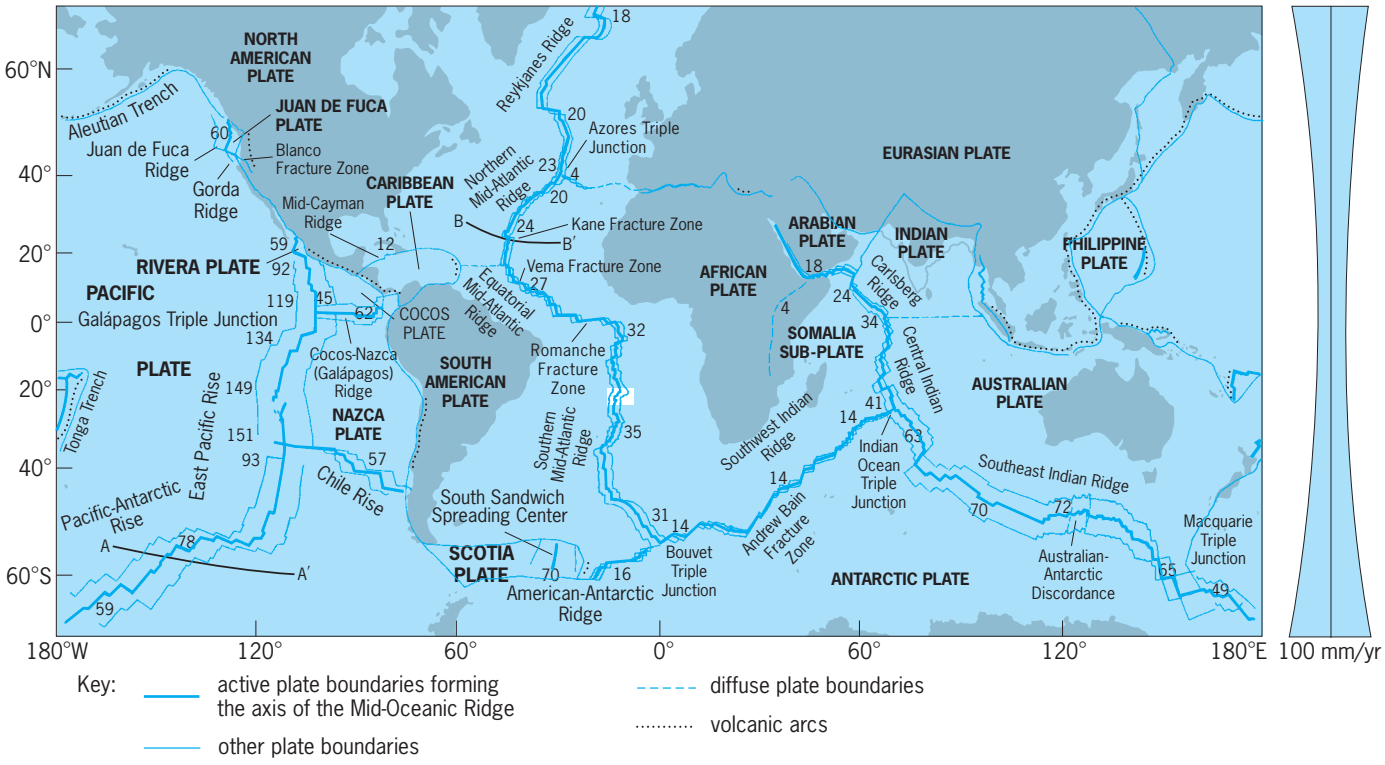

Fig. 1. Mid-Oceanic Ridge system. Paired lines on either side of the axis show the amount of crust generated in the last 10 million years at the current opening rates. Dotted lines show the volcanic arcs, which are lines of volcanoes and volcanic islands formed from magma rising from the subducted plate. The distortion caused by plotting these lines on a Mercator projection is indicated by the hourglass-shaped graph on the right, which gives the amount of crust generated at a 100 mm/yr opening rate (1 mm = 0.04 in.). Lines A-A′ and B-B′ show the locations of the profiles shown in Fig. 2. The white rectangle at 20°S, 12°W defines the location of the contour map of depths over a part of the axis of the ridge in Fig. 3.

For example, the depth of 4-million-year-old crust is $2900 + 350\sqrt{4}$, or 3600 m (12,000 ft). This formula predicts that the Mid-Oceanic Ridge will be bilaterally symmetric as long as spreading is symmetric (that is, crust is accreted at the same rate to both plates). *See* EARTH CRUST; LITHOSPHERE.

Width and related effects. The height and thermal contraction rate of the ridge crest are relatively independent of the rate of sea-floor spreading; thus, the width and regional slopes of the Mid-Oceanic Ridge depend primarily on the rate of plate separation (spreading rate). Where the plates are separating at 2 cm (0.8 in.) per year (for example, along the northern Mid-Atlantic Ridge; Fig. 1), the Mid-Oceanic Ridge has five times the regional slope but only one-fifth the width of a part of the ridge forming where the plates are separating at 10 cm (4 in.) per year (for example, the northern East Pacific Rise; Fig. 1). One consequence of the relation between the width and plate separation rate of the Mid-Oceanic Ridge is that more ocean water is displaced, thereby raising sea level, during times of globally faster plate motion. Observed long-term changes in global (eustatic) sea level probably reflect the combined effect of changing length of the total Mid-Oceanic Ridge and changing average rates of plate motion. About 80 million years ago, in the Late Cretaceous, sea levels were about 300 m (1000 ft) higher than today, primarily because of more rapid plate motion and a wider Mid-Oceanic Ridge.

The modern Mid-Oceanic Ridge is typically 1000–4000 km (600–2500 mi) wide, depending on the rate of plate separation and other factors. Actually, the ridge as a feature of thermal expansion has no sharp outer edge; the plate continues to cool and contract gradually and at ever-decreasing rates. However, the outer edge may be defined functionally as that line or zone beyond which the sea floor, deepening from the axis of the Mid-Oceanic Ridge, ceases to deepen further. Several processes can affect the location of the ridge's outer edge. Where postulated plumes of hot mantle material rise under the plates away from the axis of the Mid-Oceanic Ridge, the crust and mantle lithosphere become reelevated by as much as 1000–2000 m (3300–6600 ft). Whether formed by such a plume or not, an example of such a midplate swell is the Bermuda Rise.

Sedimentation is another process that helps give the Mid-Oceanic Ridge an outer edge. If hemipelagic sedimentation (sediment dropping down on the sea floor from the surface waters) were constant over a given part of the Mid-Oceanic Ridge, sediment thickness would increase linearly with crustal age away from the ridge axis. Although the sediment load depresses the lithospheric plate by a certain amount, the net effect of sedimentation is to make the ocean less deep. The outer edge of the Mid-Oceanic Ridge would be that line or zone where thermal subsidence of the lithospheric plate equals the shoaling effect of sedimentation. With greater distance from the Mid-Oceanic Ridge, the effect of constant sedimentation would exceed that of thermal contraction and the sea floor would rise. In some ocean basins the more dramatic effects of turbidity (suspension) flows have overwhelmed hemipelagic sedimentation by depositing large numbers of turbidites (vertically and horizontally graded sheets of sand, silt, and clay) to form the abyssal plains between the continental margins and the Mid-Oceanic Ridge. As the abyssal plains were built up in this fashion, they simultaneously extended seaward, inundating the lower flanks of the Mid-Oceanic Ridge and displacing its outer edge. Turbidite deposition was greatly accelerated as a result of expanded Plio-Pleistocene glaciation and resultant low sea levels; and so the Mid-Oceanic Ridge, particularly in the North Atlantic, is somewhat narrower than it was prior to this glacial expansion. *See* BASIN; CONTINENTAL MARGIN; ISOSTASY; MARINE SEDIMENTS; TURBIDITE; TURBIDITY CURRENT.

Hot spots. Although the Mid-Oceanic Ridge exhibits little systematic depth variation along much of its length, there are several bulges (swells) of shallower sea floor. The shallow part of the Mid-Atlantic Ridge centered at Iceland is the most prominent example. Such bulges represent 500–3500-m (1600–11,000-ft) shallower-than-normal sea floor along sections of the Mid-Oceanic Ridge that are 500–3000 km (300–1900 mi) long. The features are attributed by some scientists to mantle hot spots, upwelling plumes with warmer-than-average mantle temperature. It may be that off-axis swells like the Bermuda Rise ultimately have the same origin as on-axis ones like the Iceland or Azores swells. For reasons not well understood, the sea-floor bulges are more prominent along parts of the Mid-Oceanic Ridge where the rate of plate separation (spreading rate) is slower; for example, along the northern Mid-Atlantic Ridge and the Southwest Indian Ridge. *See* HOT SPOTS (GEOLOGY).

Local topography. The average (or regional) depth and shape of the Mid-Oceanic Ridge is modulated by more local topographic relief, some of which equals or exceeds the regional ridge in amplitude (just as the local relief represented by Pike's Peak exceeds the regional topographic high of the Rocky Mountains) [**Figs. 2** and **3**]. Whereas the regional elevation of the Mid-Oceanic Ridge is determined by the thermal structure of the upper mantle and in part by average crustal thickness, the local relief is largely the result of tectonic deformation (principally normal faulting) and volcanism. The most prominent local topographic feature is the axis itself, which for opening rates less than 3.4 cm (1.3 in.) per year is a 20–40-km-wide (12–25 mi) rift valley 1000–2800 m (3300–9200 ft) deeper than the flanking rift mountains. This rift valley was first noticed on echo-sounding lines across the Mid-Atlantic Ridge by Marie Tharp (Lamont-Doherty Earth Observatory) in 1953. At opening rates above 9 cm (3.5 in.) per year, there is a topographic high at the axis, and at intermediate rates a rift valley tends to occur only at triple plate junctions and intersections with transform faults. Where a rift valley is present on the axis of the Mid-Oceanic Ridge, the flanks tend to exhibit rough (±500 m or 1600 ft) topography, which is thought to originate by elongate blocks of young crust being downdropped and rotated along normal faults

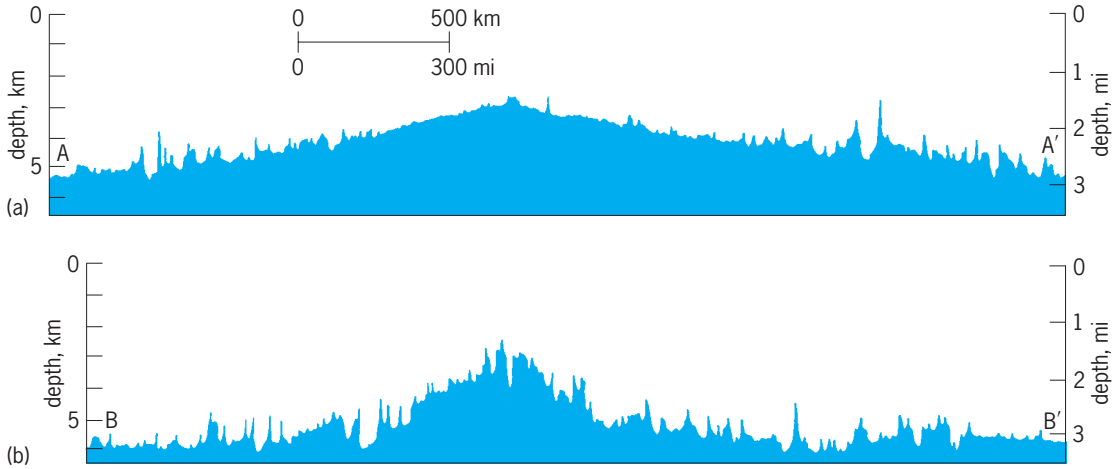

Fig. 2. Topographic profiles comparing (*a*) the broad, smooth, fast-spreading Pacific-Antarctic Ridge with (*b*) the narrow, rough, slow-spreading Mid-Atlantic Ridge. The more prominent ridges and valleys on the flanks are fracture zones (transform fault zones) that were crossed at an oblique angle. Vertical exaggeration 100:1.1 m = 3.3 ft. (*After B. C. Heezen, The deep-sea floor, in S. K. Runcorn, ed., Continental Drift, Academic Press, 1962*)

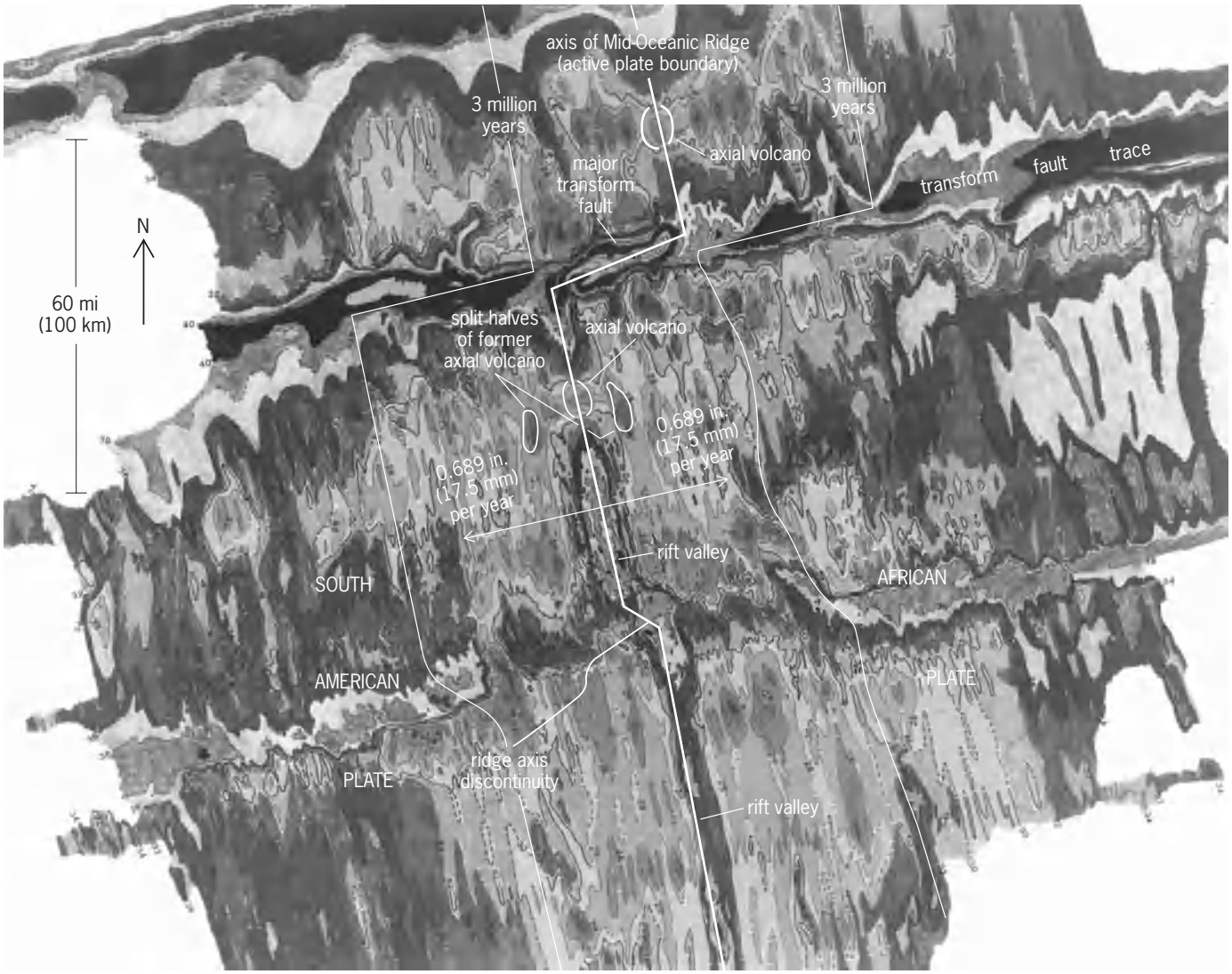

Fig. 3. Topographic (depth) contour over the crest of part of the southern Mid-Atlantic Ridge, illustrating several types of topographic features associated with the ridge. This map was the result of a joint expedition by the Naval Research Laboratory, the Lamont-Doherty Geological Observatory, and the University of Rhode Island. Depth contour intervals are 200 m with small numbers (on the perimeter) in units of 100 m (e.g., 36 = 3600 m).

dipping toward the rift axis. *See* FAULT AND FAULT STRUCTURES; RIFT VALLEY.

Aside from local failure of steep volcanic or tectonic escarpments—generally only on young crust near the axis of the Mid-Oceanic Ridge—and widely scattered off-axis (midplate) volcanism, the oceanic crust and its surface topography are passively carried away, slowly disappearing below the ever-thickening sediment cover. Except for the effects of sedimentation, the local topography on the flanks of the Mid-Oceanic Ridge is similar to that on the crest. Aside from the small area of oceanic crust exposed on Iceland, the topography of the Mid-Oceanic Ridge is essentially protected from significant erosion. Parts of the ridge may be exposed to deep currents up to about 1 knot (50 cm/s), but such currents suffice only to erode and redeposit sea-floor sediments, not the basaltic topography that may be exposed to the currents. *See* VOLCANOLOGY.

When local topographic effects are included, the maximum relief along the Mid-Oceanic Ridge axis is about 10,000 m (33,000 ft). The highest points are central volcanoes constructed on the hot-spot bulges (although generally some tens of kilometers to a hundred kilometers from the active plate boundary). Away from hot-spot bulges, the highest peaks of the Mid-Oceanic Ridge rarely breach the sea surface. More dramatic topographic relief forms near the intersections of transform faults and spreading axes (spreading centers). The rift valley deepens as it approaches its termination at a transform fault, forming a nodal basin. Transform faults are generally marked by prominent valleys, which serve as permanent traces of relative plate motion; these transverse valleys crosscutting the Mid-Oceanic Ridge are continuously generated at the site of the nodal basins. The greater the offset across a transform fault (offset being measured by the age of the old crust next to the nodal basin), the deeper the nodal basin and associated fracture zone valley.

Axial phenomena. The detailed shape of the Mid-Oceanic Ridge is not just a simple alternation of transform faults and spreading centers. Interspersed between transform faults are lesser irregularities ranging from ridge axis discontinuities (significant offset of the axis but without transverse fault development) down to minor departures from linearity. The smaller the irregularity, the more ephemeral the feature. When the offset is small, the tips of the offset spreading center may propagate past each other, forming so-called overlapping spreading centers. This is an unstable configuration and results in deformation of the crust between the overlappers. When one center continues to propagate at the expense of the offset spreading center, the result is a propagating rift. Such rifts tend to propagate away from the regional topographic bulges. Smaller-scale irregularities of the ridge axis are probably related to local volcanic centers developed several tens of kilometers to a hundred kilometers apart along the axis.

The axis of the Mid-Oceanic Ridge—that is, the active plate boundary between two separating plates—is a narrow zone only a few kilometers wide, characterized by frequent earthquakes, intermittent volcanism, and scattered clusters of hydrothermal vents where seawater, percolating downward and heated by proximity to hot rock, is expelled back into the ocean at temperatures as high as 350°C (660°F). Surrounding such vents are deposits of hydrothermal minerals rich in metals, as well as exotic animal communities including, in some vent fields, giant tubeworms and clams. Sulfate-reducing bacteria, rather than photosynthesizing plants, are the base of the food chain at the vents. Life on Earth may have originated in hydrothermal vent environments similar to those presently found along the axis of the Mid-Oceanic Ridge. *See* HYDROTHERMAL VENT; VOLCANO.

Extinction. If a particular section of the Mid-Oceanic Ridge becomes inactive—through cessation of sea-floor spreading—it is referred to as an extinct (or abandoned) spreading axis or center. This may happen on the scale of an entire plate boundary, such as 40 million years ago, when spreading ceased along the Mid-Labrador Sea Ridge between Greenland and North America; as a result, the entire Greenland plate became part of the North American plate. More commonly, only short sections of accreting plate boundary are abandoned as a section of boundary jumps to a new location inside preexisting crust. Propagating rifts create a special category of extinct spreading centers known as failed rifts. As rifts propagate continuously, the extinct spreading center is abandoned continuously and therefore varies in age along its length.

After a section of Mid-Oceanic Ridge becomes extinct, its regional topographic high gradually decays as a result of cooling and thermal contraction. For example, the axis of a ridge that died 36 million years ago is no longer 2900 m (9500 ft) deep, but has subsided to a depth of $2900 + 350\sqrt{36}$, or 5000 m (16,400 ft), and thus is barely recognizable as a ridge. However, the shorter-wavelength topography does not decay, and the site of an extinct ridge will generally still be recognizable by its median rift valley, symmetric magnetic anomaly pattern, and other features.

Deep structure. The properties of the crust and upper mantle below the Mid-Oceanic Ridge have been inferred mostly from seismic waves generated by earthquakes and by artificial sources such as explosions and air guns. By contrast, the deepest oceanic borehole penetration (on the flanks of the Galápagos Ridge) is so far only about 1850 m (6100 ft); on Iceland the deepest penetration is 2820 m (9250 ft), 4300 m (14,000 ft) below the original uneroded land surface. None of these drill holes have penetrated to the Moho (seismic M-discontinuity), the boundary between the crust and upper mantle. However, it is believed that rock assemblages known as ophiolites, exposed in ancient mountain belts, represent remnants of former oceanic crust and uppermost mantle that are tectonically separated from the main down-going slab and thrust into the upper continental crust. At least in terms of physical properties such as seismic-wave velocity, the ophiolites

resemble the modern oceanic crust as determined by geophysical measurements. *See* MOHO (MOHOROVICIC DISCONTINUITY); OPHIOLITE.

Although magma (molten rock) erupts occasionally onto the sea floor at the axis of the Mid-Oceanic Ridge, seismic methods have failed to detect a continuous crustal magma chamber, at least below slow-spreading ridges. However, this hot mantle is partially molten at depths of between about 10 and 50 km (6 and 30 mi) below the axis of the Mid-Oceanic Ridge. Seismic tomography indicates that the anomalously warm, rising mantle below the Mid-Oceanic Ridge probably extends to depths of several hundred kilometers. *See* MAGMA; MARINE GEOLOGY.

Peter R. Vogt

Bibliography. J. Bird (ed.), *Plate Tectonics*, 2d ed., American Geophysical Union, 1980; J. R. Cann (ed.) et al., *Mid-Ocean Ridges: Dynamics of Processes Associated with the Creation of New Oceanic Crust*, Cambridge University Press, 1999; K. C. Condie, *Plate Tectonics*, 4th ed., Butterworth-Heinemann, 1997; A. Cox and R. B. Hart, *Plate Tectonics*, Blackwell Science, 1986; J. Kennett, *Marine Geology*, Prentice Hall, 1982; P. Vogt and B. Tucholke (eds.), *The Western North Atlantic Region*, Geological Society of America, 1986.

Migmatite

Rocks originally defined as of hybrid character due to intimate mixing of older rocks (schist and gneiss) with granitic magma. Now most plutonic rocks of mixed appearance, regardless of how the granitic phase formed, are called migmatites. Commonly they appear as veined gneisses. *See* GNEISS.

Several modes of origin have been proposed. (1) Granitic magma may be intercalated between thin layers of schist (lit-par-lit injection) to form a banded rock called injection gneiss. (2) The granitic magma may form in place by selective melting of the rock components. (3) The granitic layers may develop by metamorphic differentiation (redistribution of minerals in solid rock by recrystallization). (4) The granitic layers may represent selectively replaced or metasomatized portions of the rock.

Veined gneisses include two genetic types: arterites, in which the vein material was injected, and veinites, in which the vein material was secreted from the rock itself. There are many other types of migmatites. Some consist of blocklike masses of various shapes enclosed in granitic rock and resembling fragments of a breccia.

Migmatites are found in zones marginal to intrusive granite and in deep zones of ultrametamorphism. *See* METAMORPHISM; METASOMATISM.

Carleton A. Chapman

Migratory behavior

Regularly occurring, oriented seasonal movements of individuals of many animal species. The term migration is used to refer to a diversity of animal movements, ranging from short-distance dispersal and one-way migration to round-trip migrations occurring on time scales from hours (the vertical movements of aquatic plankton) to years (the return of salmon to their natal streams following several years and thousands of miles of travel in the open sea). This article will concentrate on seasonal movements, usually between breeding and nonbreeding areas, that have clear directionality and are long-distance relative to the sizes and locomotory abilities of the species involved.

Evolution of migration. Organisms that live in highly seasonal environments have two basic options: they may become inactive or dormant for part of the year, or they may move elsewhere. Animals which had evolved means of rapid, long-distance locomotion were preadapted to take advantage of the latter strategy. There may often be significant advantages for an individual that can occupy environments that are available only part of the time. Food and other resources often occur in pulses of great abundance, and at high latitudes summer days are long, giving more time for feeding and caring for the young. In other species such as salmon (*Oncorhynchus*) and eels (*Anguilla*), migrations between fresh and salt water enable the animals to exploit vastly different environments at different stages in their life cycles. In its incipient stages, natural selection should favor the evolution of migratory behavior in a population if those individuals that move seasonally to a new area succeed in producing more offspring than their sedentary counterparts. Beyond this generalization, it has been difficult to reconstruct the details of the evolution of migratory behavior. It is obvious that Pleistocene glaciations had major influences on the patterns of long-distance bird migrations in northern latitudes, but there is no reason to think that migration did not have its origins long before that time. In fact, migratory behavior has likely evolved independently many times, even within the birds; indeed, it is continually appearing and disappearing in present-day populations.

Spectrum of migration. Migratory behavior is amazing in both the breadth of its occurrence in the animal kingdom and the magnitude of the journeys performed by individual species. The feather-light monarch butterfly (*Danaus plexippus*) departs from Canada and the northern United States in early autumn on a 1550-mi (2500-km) flight to Mexican wintering sites used year after year by individuals which have never been there before and which will never return. Some that survive the winter will initiate a return movement in the spring, often crossing the Gulf of Mexico. On the sea floor of the Bahama Bank in autumn, long queues of spiny lobsters (*Panulirus argus*) march southward to deeper water. Green sea turtle (*Chelonia*) hatchlings emerge from their beach nests on Ascension Island in the middle of the South Atlantic Ocean and head out to sea. For 2–3 years they remain at sea, some going as far as the coast of Brazil. Then they return across 1400 mi (2250 km) of seemingly trackless ocean to deposit their own eggs on their natal beaches. For the great whales entire ocean basins may constitute a home

range. Many species regularly move hundreds of kilometers between rich feeding areas and warmer calving sites. Great herds of ungulate mammals migrate over vast distances in many parts of the world. In the Far North, caribou (*Rangifer tarandus*) file along traditional routes across hundreds of kilometers of tundra between summering and wintering grounds. Even in tropical regions such as east Africa, large ungulates, such as the wildebeest (*Connochaetus taurinus*) and zebras (*Equus*), engage in mass movements synchronized with seasonal changes in their food supply.

Animals that have escaped the confines of terrestrial living through the evolution of flight have developed the greatest diversity of migratory strategies. Among mammals, bats are the longest-distance migrants, analogous in many ways to birds. A vast array of birds literally span the globe during their annual migrations. The short-tailed shearwater (*Puffinus tenuirostris*) of the Pacific and great shearwater (*P. griseus*) of the Atlantic breed on tiny islands in the Southern Hemisphere. During the austral winter they perform great clockwise circumnavigations of their respective oceans, returning to the nesting islands with remarkable precision. Arctic terns (*Stena paradisea*) breed northward nearly to the Arctic Circle, and twice each year fly diagonally across the Atlantic Ocean to spend the winter off the southern coast of Africa, some even rounding the Cape of Good Hope into the Indian Ocean. The round-trip distance is about 25,000 mi (40,000 km). Even more remarkable is the autumn migration of the blackpoll warbler (*Dendroica striata*), a 0.42-oz (12-g) songbird that nests in stunted conifers near the tree line in northern Canada. Following the nesting season, the birds move southeastward to the Atlantic coast of maritime Canada and New England. Then, usually following the passage of a cold front, they embark on a nonstop overwater flight to the mainland of South America, a journey that may take up to 4 days.

Control of migratory behavior. Many temperate zone species, including many migrants, are known to respond physiologically to changes in day length with season (photoperiodism). For example, many north temperate organisms are triggered to come into breeding condition by the interaction between the lengthening days in spring and their biological clocks (circadian rhythms). Similar processes, acting through the endocrine system, bring animals into migratory condition. Although these events of the annual cycle seem to be phased by photoperiod, the process is not simple. Many birds, for example, migrate southward in fall and cross the Equator into lengthening days. Yet they do not respond to the increasing photoperiod by coming into spring physiological condition; by some means their systems have become refractory to the potentially stimulating day length. Evidence that has been gathered on several species of vertebrate animals, including migratory birds, suggests that many major events of the annual cycle (hibernation, fattening, molt, migratory restlessness) are controlled by an endogenous timer (circannual rhythm) and will occur spontaneously under a constant photoperiod. *See* BIOLOGICAL CLOCKS; PHOTOPERIODISM.

Mechanics of bird migration. Evidence indicates that many birds imprint on or learn some feature of their birthplace, a prerequisite for them to be able to return to that area following migration. On its first migration, a young bird appears to fly in a given direction for a programmed distance. Upon settling in a wintering area, it will also imprint on that locale and will thereafter show a strong tendency to return to specific sites at both ends of the migratory route. Although some kinds of birds regularly migrate at very high altitudes (over 20,000 ft or 6000 m), and at least the bar-headed goose (*Anser indicus*) regularly crosses the highest Himalayas, most songbirds fly at night at altitudes of less than 3000 ft (900 m) above the ground. During migration, birds usually do not fly every night. The typical pattern for a songbird is to make a flight of 200 mi (300 km) or so one night, and then rest and replenish its fat stores for 2–3 days. Larger water birds often make much longer flights, but may remain on the ground for many days. The initiation of migration by most birds is highly dependent on the ambient weather. Large migratory movements are accompanied by favorable following winds, which can often substantially increase the ground speed of slow-flying songbirds. Many large water birds (ducks, geese, swans, sandpipers, plovers, cranes) migrate in flocks, and sometimes family groups. There is thus opportunity for young, inexperienced individuals to follow or learn from older birds. Most songbirds, however, appear to be randomly dispersed in the night sky, and in some species the sexes and age classes migrate at slightly different times of year.

Orientation mechanisms. To perform regular oriented migrations, animals need some mechanism for determining and maintaining compass bearings. Animals use many environmental cues as sources of directional information, and additional cues may remain to be discovered. Work with birds has shown that species use several compasses; there are backup systems. Much work is devoted to discovering the relationships among the several usable directional cues.

Sun compass. Many species of vertebrates and invertebrates possess a time-compensated Sun compass. With such a system, the animal can determine absolute compass directions at any time of day; that is, its internal biological clock automatically compensates for the changing position of the Sun as the Earth rotates during the day. This can be proved by resetting an animal's internal clock by placing it under an artificial light/dark cycle out of phase with Sun time. When the animal is tested under the Sun, its choice of bearings will be shifted (by about 15° for each hour of clock shift, corresponding to the average rate of change in the Sun's azimuth), because it takes its clock information to be correct and thus misinterprets the azimuth position of the Sun. Interestingly, animals using a Sun compass disregard the elevation of the Sun and use only its azimuth direction. Many arthropods, fish, salamanders, and birds can perceive the plane of polarization of sunlight.

Patterns of polarized skylight are a function of the position of the Sun and can thus be used as a compass. Polarized skylight and perhaps the position of the setting Sun seem to be important in the choice of migratory directions by some night-migrating birds.

Star compass. Only birds that migrate at night have been shown to have a star compass. Unlike the Sun compass, it appears not to be linked to the internal clock. Rather, directions are determined by reference to star patterns, which seem to be learned early in life by observing the axis of stellar rotation of the night sky. Birds in migratory condition will orient in the proper stellar direction even under a stationary planetarium sky.

Magnetic compass. Evidence indicates that several insects, fish, salamanders, birds, and mammals may derive directional information from the weak magnetic field of the Earth. The phenomenon has been best studied in homing pigeons and migratory birds. On cloudy days, the homing orientation of pigeons can be disrupted by attaching magnets to them. The orientation of the nocturnal migratory restlessness, or *Zugunruhe*, of some songbirds can be predictably changed in direction by shifting an Earth-strength magnetic field generated by coils placed around their orientation cages. Birds and salamanders, at least, do not seem to sense the polarity of the magnetic field. Rather, they possess an inclination compass. The horizontal component of the field provides a north-south axis, and in the Northern Hemisphere the axial direction that is characterized by the smaller angle between the resultant magnetic field vector and gravity is taken to be north. *See* MAGNETIC RECEPTION (BIOLOGY).

Wind direction. Wind provides directional information to birds, but because it changes with time, it cannot yield compass information directly. Nocturnal songbird migrants generally fly with a following wind; in part this is because they select a night with favorable winds to initiate migration. Birds sometimes fly in reversed or other inappropriate directions during migration. Usually these flights are downwind, and are especially likely to occur under overcast skies, which prevent the birds from seeing the Sun and stars. Under these conditions, birds often resort to wind direction as their main orientation cue.

Moon. Three beach-dwelling amphipod crustaceans (*Talitrus, Orchestoidea*, and *Talorchestia*) seem to possess a time-compensated lunar compass in addition to a Sun compass.

Homing and navigation. Many kinds of animals show the ability to return to specific sites following a displacement. The phenomenon can usually be explained by familiarity with landmarks near "home" or sensory contact with the goal. For example, salamanders and toads can return to capture sites following displacements of up to a few kilometers by following odor gradients. Salmon are well known for their ability to return to their natal streams after spending several years at sea. Little is known about their orientation at sea, but they recognize the home stream by chemical cues (olfactory) in the water. The young salmon apparently imprint on the odor of the stream in which they were hatched. When they return, they move upstream, choosing the proper tributary at each branch point based on its odor. Young salmon can be imprinted to artificial odors introduced into the hatching stream and later lured into a different stream on the return migration. The origin of the natural odors used in stream recognition remains controversial.

Only in birds can an unequivocal case be made for the existence of true navigation, that is, the ability to return to a goal from an unfamiliar locality in the absence of direct sensory contact with the goal. This process requires both a compass and the analog of a map. Homing pigeons and several seabirds can return to their lofts or nests when experimentally displaced hundreds or thousands of miles. There are two main hypotheses to explain the map component that tells them which direction is toward home. Many experiments with homing pigeons during the 1980s support the olfactory map hypotheses. By perceiving odors at their home loft and associating those odors with the directions of the winds carrying them, the pigeons gradually learn the odor environment over an extensive area. Experiments suggest that such an odor map may be useful in determining the home direction even when the pigeons are displaced hundreds of miles to completely unfamiliar places. The other hypothesis is that pigeons use gradients in one or more parameter of the Earth's magnetic field as a map. The main supporting evidence is that pigeons released at magnetic anomalies are often completely disoriented. These hypotheses are not mutually exclusive. Kenneth P. Able

Bibliography. R. R. Baker, *Bird Navigation*, 1984; P. Berthold (ed.), *Orientation in Birds*, 1991; M. A. Rankin (ed.), *Migration: Mechanisms and Adaptive Significance*, 1985; T. H. Waterman, *Animal Navigation*, 1988.

Military aircraft

Aircraft that are designed for highly specialized military applications. Fixed-wing aircraft, rotary-wing aircraft, free-flight balloons, and blimps have all been used in both crewed and crewless flight modes for military purposes. *See* DRONE.

Designations of military aircraft are standardized by the U.S. Department of Defense (**Table 1**). A letter indicates the purpose of the aircraft, and a following number refers to a particular model. An additional letter before the letter of purpose indicates a modification of function, while a letter following the model number shows a modification of the basic aircraft's equipment. For example, a reconnaissance version of the F-4 fighter with equipment modifications is designated as an RF-4C. Basic types of military aircraft include bombers, fighters, transports, patrol aircraft, trainers, and reconnaissance and observation aircraft.

Modern military aircraft are fitted with complex

TABLE 1. Designations for United States military aircraft

Designation	Type
A	Attack
B	Bomber
C, C*	Cargo and transport
D*	Director for drones
E*	Electronic, early warning
F, F*	Fighter
H	Helicopter
H*	Air rescue modification
K*	Refueling tanker
O	Observation
OV	STOL, observation
P	Patrol
Q	Target and drone
R*	Reconnaissance
SA	Rescue amphibian
SR	Strategic reconnaissance
T	Trainer
U, U*	Utility, observation
V*	Staff administrative transport
W*	Weatherreconnaissance
X*	Experimental
Y*	Service test

*Used as prefix to basic type letter.

equipment for successful accomplishment of their various missions. Many automatic devices assist the crews in obtaining results not possible with purely human abilities. Target-finding, tracking, and weapons-firing equipment to a great degree have replaced manual gunnery, missile launch, and bomb release. Electronic countermeasures equipment automatically counters enemy air defense systems. *See* AIR ARMAMENT; ARMY ARMAMENT; ELECTRONIC WARFARE; NAVAL ARMAMENT.

Operational aircraft are those in active inventory. They have been fully developed and tested in a production configuration (**Table 2**).

An aircraft proceeds through various stages of system engineering during development. The process starts with a definition of system requirements, which leads to concept definition, preliminary design, detailed design, limited-quantity prototyping, test and evaluation, and eventually production design. When an aircraft design is released for serial production, the development phase has been completed. *See* AIRCRAFT DESIGN.

Fig. 1. Rockwell-B-1B strategic bomber. (*North American Aircraft, Rockwell International Corp.*)

Bombers. These are usually characterized by relatively long range, low maneuverability, and large weapon-carrying capability. Bombers are sometimes classified by their range capabilities, such as intercontinental, medium, or short. However, the use of aerial refueling by tanker aircraft gives most bombers a global range. Bombers may be equipped to deliver conventional or nuclear weapons in day, night, or adverse weather.

B-1B. The Reagan administration revived the B-1 bomber program in October 1981, and the first production B-1B (**Fig. 1**) flew in October 1984. All 100 of these aircraft were produced and delivered to the Strategic Air Command (now the Air Combat Command).

The B-1B was designed to operate at speeds up to Mach 0.9 (9/10 the speed of sound) at altitudes near 200 ft (60 m) in the penetration role. One of its defensive attributes is its ability to fly at low altitudes and at high speeds. Another defensive feature is the aircraft's low radar cross section. The B-1B has 1/100

TABLE 2. Table 2. Representative performance characteristics of military aircraft

Designation	Country	Function	First flight	Maximum takeoff weight, lb (kg)	Maximum speed, Mach number (M) or knots* (km/h)	Unrefueled radius, nmi (km)
B-1B	United States	Bomber	Oct. 1984	477,000 (216,558)	M 1.25+	4047 (7500)
TU-160	Russia	Bomber	June 1982	590,000 (267,860)	M 2.0	3940 (7300)
IL-76	Russia	AEW & C†	Mid 1980	375,000 (170,250)	434 (804)	1241 (2300)
E-3A	United States	AEW & C	Oct. 1975	325,000 (147,392)	462 (856)	600 (1112)
F-117A	United States	Stealth fighter	June 1981	35,000 (15,890)	Subsonic	400+ (740+)
F/A-18	United States	Strike fighter	Nov. 1978	50,000 (23,000)	M 1.8+	500+ (927+)
F-16A	United States	Fighter	Dec. 1976	24,000 (10,896)	M 2.0	540 (1000)
F-15E	United States	Fighter/attack	Dec. 1986	81,000 (36,774)	M 2.5	500 (925)
TR-1A	United States	Reconnaissance	Aug. 1981	40,000 (18,160)	375+ (695+)	1303 (2415)
C-5B	United States	Heavy airlift	Sept. 1985	837,000 (379,998)	495 (917)	1490 (2761)
AH-64	United States	Attack helicopter	Sept. 1975	17,400 (7,900)	162 (300)	130 (240)
B-2	United States	Stealth bomber	July 1989	350,000 (158,900)	435 (806)	6000+ (11,119+)
C-17	United States	STOL airlift‡	Mid 1990	570,000 (258,780)	350 (649)	1900 (3521)

*Knot = 1 nautical mile per hour.
†AEW & C = airborne early warning and control.
‡STOL = short takeoff and landing.

Fig. 2. Tupolev TU-160 intercontinental bomber (NATO codename Blackjack). (*TASS from SOVFOTO*)

Fig. 3. Overhead view of Northrop B-2 stealth bomber. (*Photograph by W. G. Hartenstein; from U.S. Air Force, Northrop unveils B-2 next-generation bomber, Aviat. Week Space Technol., 129(22):20–23, November 28, 1988*)

the cross section of a B-52 and 1/7 that of the FB-111. The aircraft has a flight crew of four and is powered by four afterburning turbofan engines. It is capable of carrying short-range attack missiles, nuclear and conventional gravity bombs, and air-launched cruise missiles. *See* TURBOFAN.

Tupolev TU-160. This Russian long-range strategic bomber (NATO codename Blackjack) is one of the world's largest and heaviest bombers. Designed to carry bombs and air-launched cruise missiles, it can cruise subsonically over long distances, perform high-altitude supersonic dash, and attack by utilizing low-altitude, high-subsonic penetration maneuvers. As with the B-1, the TU-160 has a blended wing-body design with a variable swept wing and a single vertical stabilizer (**Fig. 2**). It has an unrefueled combat radius of about 3940 nautical miles (7300 km) and a maximum speed of Mach 2.0.

B-2. One of the most costly and secret United States aircraft development programs has been the B-2 advanced technology bomber or stealth bomber. The B-2 is designed to attack either fixed or mobile targets and fly at high altitudes over lightly defended areas or at low altitudes when warranted by heavy defenses. The flying wing design (**Fig. 3**) provides both long range and high payload capacity in a vehicle of low observability. All exterior features are designed to present the smallest radar, electrooptical, and infrared signatures possible when viewed from any aspect. By carefully integrating features of low observability, the aircraft does not have to rely on the sophisticated offensive and defensive avionic systems carried by the B-1 bomber.

The B-2 carries all weapons internally to minimize drag and radar cross section. The aircraft is powered by four nonafterburning turbofan engines buried in the aircraft body on each side of the bomb bay area. The two-crew-member cockpit is equipped with multifunctional displays and conventional control sticks. The fly-by-wire flight control system is quadruple-redundant. Aircraft range at high altitude is more than 6000 nmi (11,000 km) unrefueled and more than 10,000 nmi (18,500 km) with a single refueling. The bomber is capable of operating subsonically from very low altitudes up to the 50,000-ft (15,000-m) regime. *See* FLIGHT CONTROLS.

Airborne early warning and control. Aircraft that provide airborne early warning and control include a variant of the Russian Ilyushin IL-76 and the United States E-3A.

Ilyushin IL-76. Increasing numbers of the airborne early warning and control variant of the Russian Ilyushin IL-76 aircraft (NATO codename Mainstay) have been produced for early warning against low-altitude penetration and for air battle management (**Fig. 4**). This aircraft has a rotating saucer radome mounted on top of the aft fuselage, an identification-friend-or-foe (IFF) system, comprehensive electronic

Fig. 4. Ilyushin IL-76, airborne early warning and control variant (NATO codename Mainstay). (*From U.S. Department of Defense, Soviet Military Power: An Assessment of the Threat, 1988*)

Fig. 5. Boeing Airborne Warning and Control System (AWACS) E-3A. (*Boeing Aerospace Co.*)

countermeasures, and a flight refueling probe. The combination of Mainstay and advanced Russian counterair fighters, such as the MIG-29, MIG-31, and Sukhoi SU-27, gives the capability to extend strategic air defense far beyond the borders of the country.

E-3A (AWACS). The U.S. Air Force Airborne Warning and Control System (AWACS), designated E-3A, is a versatile surveillance, command, and control center, designated to provide battle management in the conduct of air warfare (**Fig. 5**). AWACS provides full long-range surveillance over all air vehicles, crewed and crewless, at high and low altitudes, during all kinds of weather, and above all kinds of terrain. Its data-storage and processing capability provide an assessment of enemy action as well as the condition and location of available friendly resources. Its primary advantage is that it centralizes the coordination of complex, simultaneous air operations. It can command and control the total air effort: strike, air superiority, support, airlift, reconnaissance, and interdiction. In short, the airborne commander has available all the information needed to detect, assess, and counter any enemy threat.

The Air Force crew complement consists of 13 AWAC system specialists in addition to a flight crew of 4. The AWAC system is installed in a 707-320 aircraft carrying an externally mounted, 30-ft (9-m) rotodome and a powerful, concentrated payload of electronic detection and display equipment. On-board communication gear provides for reception, recording, display, transmission, and relay of a wide variety of signals, both digital and voice, and messages to and from ground and air stations.

The most important feature of AWACS is the amount of surveillance coverage supplied by the radar for command-and-control use. Target detection and other performance characteristics of AWACS are classified. However, it is known that a single AWACS can observe the entire airspace volume over several states. Even in an electronic countermeasures environment, AWACS is able to perform its mission acceptably. The AWACS radar subsystem includes some unique features such as the computerized control and signal-processing functions. Electronic systems are principally digital. The high energy levels required by the radar transmitter affect the design of aircraft power and cooling systems. Although the radar contains approximately 100,000 components and parts, reliability is high. *See* AIRBORNE RADAR; RADAR.

Identification of a target aircraft as a friendly or potential threat vehicle is provided by the identification-friend-or-foe system, which uses an antenna mounted in the rotodome back-to-back with the surveillance radar antenna. This system is a highly directional, interrogate-receive, coded identification system that provides azimuth and range data on coded transponding targets.

External communication with supporting mission elements is provided by high-frequency (HF), very high-frequency (VHF), and ultrahigh-frequency (UHF) communication channels, with information transmitted as clear (unencoded) or secure voice and digital data. A high-powered UHF channel transmits commands to supporting aircraft by data link, and reply information is received on a separate UHF blade antenna. A second UHF high-powered channel is used for communication between AWACS airplanes in the same vicinity, employing voice or digital data. Medium-powered UHF channels also provide voice transmissions for terminals that are beyond the AWACS line of sight. VHF-AM (amplitude modulation) transmits in clear voice modes to friendly aircraft, and VHF-FM (frequency modulation) is used for coordination with ground forces. The AWACS flight crew has its own UHF transceiver and shares access to mission VHF-AM equipment.

Fighters. Unlike bombers and AWACS aircraft, fighters are relatively short-range, highly maneuverable, fast aircraft, designed to destroy enemy aircraft and to attack ground targets. They can carry machine guns, cannons, rockets, guided missiles, and bombs, depending upon the mission. They may be interceptor fighters, designed to shoot down enemy airplanes or missiles during day, night, or adverse weather conditions. Other fighters may be designated for close-in attack of mobile enemy ground forces to provide close support for friendly ground troops. Some fighters, called fighter-bombers, can carry conventional or nuclear weapons several hundreds of kilometers behind enemy lines to strike priority ground targets.

F-117A. This is a unique fighter aircraft in active service in the U.S. Air Force (**Fig. 6**). The F-117A, known as the stealth fighter, is a single-seat, twin-engine aircraft with low radar, infrared, optical, and acoustic

Fig. 6. Lockheed F-117A stealth fighter. (*U.S. Air Force*)

Fig. 7. Two McDonnell Douglas F/A-18 (Hornet) strike fighter aircraft, each carrying two heat-seeking Sidewinder missiles and an external fuel tank. (*McDonnell Aircraft Co., McDonnell Douglas Corp.*)

signatures. It first flew in 1981 and became operational in 1983. The aircraft is designed to evade enemy radar and infrared detection while conducting low-altitude strikes against high-priority targets. The F-117A is about the same size as an aircraft-carrier-based fighter and carries its weapons internally in two armament bays. The aircraft fuselage has a number of flat, vertical surfaces angled outward to deflect and diffuse radar returns. The F-117A also has radar-absorbent composite materials on its external surfaces. Engine exhaust nozzles are located on top of the wing root so that the exhaust flows over the top aft fuselage and is shielded from being viewed by infrared sensors below the aircraft. *See* COMPOSITE MATERIAL; RADAR-ABSORBING MATERIALS.

F/A-18. This aircraft, in service with the U.S. Navy and Marine Corps, is known as the Hornet (**Fig. 7**). It is called a strike fighter or swing aircraft because it can be rapidly reconfigured to fly either as a fighter or attack aircraft. Most of the conventional instrumentation in the cockpit has been replaced by three multifunction displays, an information control panel, and a head-up display in front of the pilot. Every critical switch required for air-to-air and air-to-surface engagements is either in the throttle (left hand) or on the control stick (right hand), so that the pilot is not distracted by having to move the hands to different controls. *See* AIRCRAFT INSTRUMENTATION.

During air-to-air combat the radar can track multiple targets, displaying up to eight target tracks while retaining up to ten in its memory. A raid-assessment mode enables the pilot to discriminate between closely spaced targets. The radar information is displayed on a clutter-free scope in either lookup or lookdown attitude; it also provides range-while-search capability, long-range search and track, and several modes for close-in combat.

A variety of survival enhancement features are designed into the F/A-18. The Hornet is powered by two low-bypass turbofan engines, which provide greater reliability than a single engine. Reticulated foam is used in the fuel tanks to suppress fires and explosions which might result from fragment penetration. The fuel tanks and fuel lines are also self-sealing. The aircraft has two completely separate hydraulic systems. It also has a quadruple, redundant, digital, fly-by-wire flight-control system with direct electrical backup to all control surfaces, and a direct mechanical backup to the stabilizers. *See* STABILIZER (AIRCRAFT).

F-16. The F-16 was developed to replace F-4s in the active force and to modernize the air reserve forces. Advanced techniques incorporated from the start in the single-seat F-16A and two-seat F-16B versions made them two of the most maneuverable fighters ever built. The advances include decreased structural weight through the use of composite materials, decreased drag resulting from a reduced static stability margin, fly-by-wire flight controls with a side-stick controller, a high-visibility cockpit with a 30° reclining seat and single-piece bubble canopy, blended wing-body aerodynamics with forebody strakes, and automatically variable leading-edge wing flaps (**Fig. 8**). The F-16 is powered by a single afterburning turbofan engine. Equipment includes a multimode radar with clutter-free lookdown capability, an advanced radar warning receiver, a head-up display, internal chaff and flare dispensers, and a 500-round 20-mm internal gun. The aircraft also has provisions for electronic countermeasures equipment. The F-16 entered operational service in 1979. *See* STRAKE.

The F-16s has a multimode radar with increased range and advanced electronic countermeasures, and advanced cockpit displays including a wide-angle head-up display with forward-looking infrared video. Shrike antiradiation missiles and

Fig. 8. Lockheed F-16 fighter. (*General Dynamics Corp.*)

Fig. 9. McDonnell Douglas F-15E (Eagle) fighter, fitted with LANTIRN (Low-Altitude Navigation and Targeting Infrared for Night) pods. (*McDonnell Aircraft Co., McDonnell Douglas Corp.*)

multitarget Advanced Medium-Range Air-to-Air Missile (AMRAAM) compatibility have been added to the F-16Cs and Ds. System improvements were also introduced that include installation of an advanced navigation and attack system, digital flight controls, automatic terrain following, advanced identification-friend-or-foe capability, and advanced takeoff weight and maneuvering limits. An advanced radar warning receiver, improved defensive countermeasures, increased-performance engines, and full capability to launch the High-Speed Antiradiation Missile (HARM) and the Shrike missile are part of the current production aircraft.

F-16s are standard equipment with 15 units in the Air Combat Command, U.S. Air Forces Europe, and Pacific Air Forces, and are progressively replacing older aircraft in the Air Force reserves and the Air National Guard. F-16Cs also equip the U.S. Air Force's Thunderbird air demonstration squadron. Eleven hundred more have been delivered to or ordered for the air forces of 17 other countries and the U.S. Navy.

F-15E. The U.S. Air Force's primary all-weather fighter, the F-15, known as the Eagle, has been replacing the F-4 since the mid-1970s. The original single-seat F-15A and two-seat F-15B were followed in 1979 by the F-15C and the F-15D, respectively, with 2000 lb (900 kg) of additional internal fuel and provision for carrying conformal fuel tanks. Basic F-15 equipment includes a lightweight X-band pulse-Doppler radar for long-range detection and tracking of small, high-speed objects down to the tree-top level. *See* DOPPLER RADAR.

The F-15E (**Fig. 9**) is an advanced two-seat, dual-role, totally integrated fighter for all-weather, air-to-air deep interdiction missions. Production F-15Es also have changes in the front cockpit, which include redesigned controls, a head-up display with a wide field of view, and three multipurpose cathode-ray-tube displays. The F-15E is capable of carrying up to 24,500 lb (11,100 kg) of ordnance. The digital, triply redundant flight-control system permits coupled automatic terrain following. Navigational accuracy has been improved with the use of a ring laser gyro. For low-altitude, high-speed penetration and precision attack on tactical targets at night and in adverse weather, the F-15E carries high-resolution APG-70 and LANTIRN (Low-Altitude Navigation and Targeting Infrared for Night) pods, with wide-field, forward-looking infrared (FLIR) imagers. *See* GYROSCOPE; INFRARED IMAGING DEVICES. Robert A. Strohl

Reconnaissance aircraft. The United States reconnaissance program provides capabilities to meet many peacetime and wartime information collection requirements. Reconnaissance resources include strategic, tactical standoff, and penetration aircraft systems that are flexible and responsive. Reconnaissance aircraft carry photographic, infrared, radar, and television sensors. These aircraft may be specially designed or may be modified from a basic fighter or bomber type. Some are equipped with special electronic gear for such purposes as submarine detection; others serve as picket planes for early warning of an enemy approach. Strategic airborne systems have included the U-2R, SR-71, RC-135, and EP-3E aircraft that carried a variety of long-range sensors used on standoff and overflight missions. Tactical aircraft standoff systems include the Air Force TR-1A (**Fig. 10**) and RC-135; the Army RC-12, RU-10, RU-21, EH-1, EH-60, RV-1D, and OV-1D; the Navy EP-3E, EA-3B, and EA-6B; and the Marine Corps EA-6B. The only Air Force tactical reconnaissance aircraft used in a penetrating role is the RF-4C with photo, infrared, and tactical electronic reconnaissance capabilities. The Navy uses the F-14 Tactical Air Reconnaissance Pod System (TARPS), and the Marines employ the RF-4B in a similar role.

The configuration of the TR-1A is basically that of a powered sailplane. Its unusual bicycle landing gear, combined with underwing balancer units, provides stability during takeoff, and both are then jettisoned. Range can be extended by shutting off the engine and gliding. Because of its configuration, the TR-1A requires unusually precise handling during takeoff and landing since there is an extremely small margin between approach speed and stalling speed. After touchdown the aircraft comes to rest on one of its down-turned wingtips, which serve as skids. The single-seat TR-1A is equipped with a variety of electronic sensors to provide continuously available, day or night, high-altitude, all-weather standoff surveillance. The TR-1A (Fig. 10) has the same

Fig. 10. Lockheed TR-1A tactical reconnaissance aircraft. (*Lockheed Aeronautical Systems Co., Lockheed Corp.*)

Fig. 11. Lockheed C-5B strategic transport aircraft. (*Lockheed-Georgia Aircraft Co.*)

basic airframe as the U-2R, but with the significant addition of an advanced synthetic aperture radar system (ASARS), in the form of a side-looking airborne radar (SLAR), and modern electronic countermeasures. The TR-1A is intended primarily for use in Europe, where its SLAR provides the capability to observe approximately 30 nmi (55 km) into hostile territory without the need to overfly an actual or potential battle area.

Transport aircraft. These provide dedicated logistic support to all types of military operations. Transport aircraft carry troops and war supplies. Many are adaptations of airplanes used by commercial airlines. Passenger seats or cargo space and tie-downs are provided as needed. Military transport aircraft, such as the C-5B (**Fig. 11**), are capable of moving large quantities of people and material rapidly to distant points. Cargo may be attached to pallets, which are easily loaded, secured for flight, and quickly unloaded for delivery. Cargo also may be discharged from flying aircraft on parachutes, eliminating the necessity for landing. The aerial tanker is a special-purpose transport aircraft. Fighters, bombers, and helicopters can refuel from tankers while in flight by means of special probe and drogue fittings. Any point on Earth can be brought within range of aerially refueled aircraft. Another special-purpose transport is a gunship. This aircraft is equipped with rapid-fire weapons for saturation attack on ground targets.

C-17. The United States Air Force has in operation a heavy-lift aircraft designated the C-17. The aircraft is the first airlifter designated from the outset to provide both inter- and intratheater airlift. A C-17, with its 172,200-lb (78,100-kg) maximum payload, can carry an M1 Abrams tank or three M2 Bradley fighting vehicles (**Fig. 12**). It is slightly larger in wingspan and length, and nearly 20 ft (6 m) taller, than a C-141. The C-17 is capable of airdropping outsized equipment as well as conducting parachute drops and low-altitude parachute extractions. On the ground the C-17 is able to fuel helicopters and other aircraft directly from the airlifter's wing tanks. The floor can be set up to carry either pallets or tie-down cargo. It can be completely converted in 50 min while in flight. The lower quarters of the fuselage are lined with a bulletproof, composite material to protect troops sitting in the 54 permanently installed seats.

The C-17s are powered by four turbofan engines, each producing 40,700 lb (181 kilonewtons) of thrust. These engines are used to power many 757 civilian aircraft, and the power plants have been in commercial service since September 1987.

The airlifter's short-field landing capability is provided by an externally blown flap system. This technology was pioneered on the YC-15 prototype airlifter in the mid-1970s. The engine exhaust on the C-17 is blown through and down across huge flaps (each as big as the wing of a DC-9 jetliner), creating the effect of a much larger wing surface. This propulsive lift technology allows the C-17 to descend slowly and have a moderate landing speed of 115 knots (59 m/s).

Anotov An-225. The Soviets flight-tested a very heavy airlifter (**Fig. 13**) called the An-225 Mriya (Dream). The An-225 (NATO codename Cossack) has a maximum takeoff gross weight of 1,322,750 lb (600,000 kg), which is much larger than that of any other aircraft currently flying. It is powered by six turbofan engines and has a six-member flight crew. Each main landing gear has seven pairs of wheels in tandem, which gives it superior turning ability on narrow runways.

The aircraft was developed to carry cargo inside the fuselage, which has a large cross section, as well as accommodate oversized payloads externally on

Fig. 12. McDonnell Douglas C-17 transport aircraft. (*Douglas Aircraft Co., McDonnell Douglas Corp.*)

Fig. 13. Antonov An-225 Mriya transport aircraft, carrying the space shuttle orbiter Buran and accompanied by a fighter aircraft. (*From D. North, J. Lenorovitz, and M. Dornheim, Perestroika's changes grip Soviet aerospace industry, Aviat. Week Space Technol., 130(23):34–37, June 5, 1989*)

Fig. 14. McDonnell Douglas AH-64 Apache helicopter. (*Douglas Aircraft, McDonnell Douglas Corp.*)

upper fuselage attach points. The Mriya played a major role in the Soviet space program by carrying loads such as the Buran space shuttle orbiter (Fig. 13) and sections of the giant Energlye rocket launch vehicle. The redesigned, twin-fin tail unit assures optimum directional control when such loads are in place. The An-225 is also designed to ferry heavy military equipment to remote and inhospitable regions.

Helicopters and VTOL aircraft. Helicopters deserve special mention as military aircraft. They are unexcelled for rescue work and for delivery of people and material to otherwise inaccessible areas. Some helicopters are armed and serve as attack aircraft, providing gun and rocket fire against ground targets. Other helicopters deliver assault troops to advanced combat areas and supply them with ammunition and other needs.

AH-64. A unique rotary-wing aircraft is the AH-64 Apache, the U.S. Army's primary attack helicopter (**Fig. 14**). The Apache is a quick-reacting airborne antitank weapon system. Terrain limitations and an unfavorable balance in armor dictate the need for a system that can fly quickly to the position of heaviest enemy penetration and destroy, disrupt, or delay the attack long enough for friendly armor and ground units to reach the scene. The Apache is designed to fly anywhere in the world and survive. It is equipped with a Target Acquisition Designation Sight and Pilot Night Vision Sensor (TADS/PNVS), which permits its two-member crew to navigate and attack in darkness as well as adverse weather conditions. Although the principal mission of the Apache is the destruction of enemy armor with the Hellfire missile, it is also equipped with a 30-mm chain gun and Hydra 70 rockets that are lethal against a wide variety of targets.

V-22. One of the most advanced United States vertical takeoff and landing (VTOL) aircraft in development is the V-22 Osprey tilt-rotor (**Fig. 15**). The aircraft is equipped with wingtip-mounted engines and three-bladed, 38-ft-diameter (11.6-m) rotors/propellers that allow it to take off and land vertically like a helicopter but to cruise at higher speeds as a fixed-wing turboprop aircraft.

The V-22 Osprey makes extensive use of corrosion-resistant advanced composite materials, without which its capacity to perform like a helicopter and a conventional aircraft would not be possible. The aircraft also has triply redundant digital flight management and fly-by-wire flight-control systems.

The Osprey is being developed for all four armed services. The Marine Corps plans to use the aircraft for amphibious troop assault support missions. The Air Force will use its V-22s for long-range placement and pickup of special operations forces. The Navy will use the aircraft for combat search and rescue, special warfare, and fleet logistics support. The Army may order the aircraft at a later date for troop transport. *See* HELICOPTER; VERTICAL TAKEOFF AND LANDING (VTOL).

Research aircraft. These special-purpose aircraft are occasionally designed, assembled, and tested in order to experiment with advanced aerodynamic, structural, avionic, or propulsion concepts that must be validated before they can be applied to other aircraft designs. Research aircraft are usually well instrumented, with performance data telemetered on radio-frequency data links to ground stations located at the test ranges where they are flown. *See* AIRCRAFT TESTING.

X-31. An aircraft capable of maneuvering at extremely high angles of attack without loss of control would be highly advantageous in air combat. Under a joint technology program of the United States and Germany, the concept of enhanced fighter maneuverability has been realized in the form of the X-31 aircraft (**Fig. 16**). The aircraft's ability to perform extremely rapid three-dimensional maneuvers in the poststall regime is expected to give future fighter aircraft a decided tactical advantage in air combat. Head-on or near-head-on missile launches will dominate in the future air-to-air combat arena. For an aircraft to survive in such an environment, it will have to perform extremely agile and tight maneuvers. It must be able to point its nose swiftly in a new direction, especially in the up and down directions (as

Fig. 15. Bell Boeing V-22 Osprey tilt-rotor aircraft. (*Bell Helicopter Textron*)

Fig. 16. Rockwell-Messerschmidt X-31 enhanced maneuverability fighter. (*North American Aircraft, Rockwell International Corp.*)

opposed to side to side), to fire off a shot quickly. In many cases, this will mean moving abruptly into the poststall regime of the flight envelope. This regime is defined to include angles of attack higher than that at which the lift reaches its maximum value (about 35°).

Radius of turn is foreseen as the decisive parameter during close-in, frontal engagements. A smaller radius of turn is likely to lead to preemptive weapon launch as two aircraft are turning against each other. The smallest radius of turn is achieved at maximum lift conditions, which is very close to stall. With current design, the highest-performing contemporary fighter can turn in a radius of about 1600 ft (500 m). Further meaningful reductions in turn radius can probably be achieved only by penetrating the poststall regime at angles of attack much higher than those for maximum lift. Since the X-31 is intended to hold the advantage in a head-on combat engagement, it is being designed for maneuverability in this regime.

A particularly difficult challenge is posed by the design of the X-31's flight-control system for safe and reliable operation at poststall angles of attack. Beyond an aircraft's stall limits, conventional flight controls are insufficient to keep the aircraft from departing into uncontrollable flight, and some means of augmenting control power will be required. Therefore, a vectored thrust system has been integrated into the X-31's flight-control system.

X-29. This X-series research aircraft, built and tested during the 1980s (**Fig. 17**), was designed as a technology demonstrator for the Air Force and NASA to explore the performance potential of a forward-swept-wing design and associated technologies. The aircraft incorporates a composite wing swept forward at an angle of 29.3°. The wing is aeroelastically tailored and incorporates a thin, high lift-to-drag supercritical airfoil. The trailing edge of the wing features double-hinged control surfaces that provide variable camber. The X-29 has close-coupled canards that offer high levels of longitudinal control. The aircraft incorporates a triply redundant, digital, fly-by-wire control system with an analog backup. *See* AEROELASTICITY; SUPERCRITICAL WING.

The first X-29 began flight tests in December 1984. The aircraft completed more flights than any previous X-series aircraft, providing extensive aerodynamic data on the basic forward-swept-wing configuration. It has flown up to 22.5° angle of attack, attained a maximum level speed of Mach 1.47, and reached a maximum altitude of 50,800 ft (15,500 m). The aircraft's dual design points, at which performance is optimized, are Mach 0.9 at 30,000 ft (9144 m) and Mach 1.2 at the same altitude, giving the X-29 a sustained transonic maneuvering capability. It appears that a forward-swept fighter aircraft can be constructed with 10–20% less drag and 5–25% less weight than a conventional aft-swept-wing aircraft.

Unlike the X-31A program, which will examine agility specifically in maneuvers performed at high angles of attack, the X-29 design goal is to provide high levels of maneuvering performance across a wide envelope of operation. Performance associated with the forward-swept wings is expected to permit a number of transonic maneuvering performance improvements, including a high, instantaneous rate of turn and roll control at high angles of attack. The X-29's unique three-surface control configuration (canards, wing control surfaces, and strake flaps) should also provide the aircraft with significant longitudinal control at high angles of attack.

Robert J. Strohl

Fig. 17. Grumman X-29 forward-swept-wing aircraft. (*From S. W. Kandebo, Second X-29 will execute high-angle-of-attack flights, Aviat. Week Space Technol., 129(18):36–42, October 31, 1988*)

Bibliography. *Jane's All the World's Aircraft*, revised periodically; U.S. Department of Defense, *United States Army Weapon Systems*, annually.

Military satellites

Artificial satellites used for a variety of military purposes. Of approximately 4000 satellites successfully launched between 1957 and 1999, about 50% have been either specifically military or usable for military purposes. Major functions of military satellites include communications, positioning and navigation, meteorology, reconnaissance and surveillance, early warning, remote sensing, geodesy, and research.

Communications. Although only certain satellites are used continuously for military purposes, all communication satellites, including commercial, may find use during conflict. All contain the necessary equipment to transmit a signal over great distances to assist in the command, control, administration, and logistic support of military forces. Military communication satellites differ from commercial satellites only in that they contain specialized components, certain capabilities, and multiple redundant systems designed to make them less vulnerable and more effective in a hostile environment.

United States satellites. Several constellations of United States military communication satellites are in operation. They may be divided into three general categories according to the frequencies at which they operate: ultrahigh frequency (UHF, 300 MHz–3 GHz), superhigh frequency (SHF, 3–30 GHz), and extremely high frequency (EHF, 30–300 GHz). *See* RADIO SPECTRUM ALLOCATION.

In the UHF arena are the Fleet Satellite Communication System and the Air Force Satellite Communication System. The Fleet System consists of five active satellites and several spares in geostationary orbit (22,300 mi or 35,800 km altitude), providing worldwide communications (**Fig. 1**). Each spacecraft has 23 channels operating at a variety of UHF frequencies. As these satellites wear out, they are replaced by a new constellation of UHF Follow-on (UFO) satellites with 42 channels providing narrow-band tactical communications and an antijam capability. The Air Force System provides worldwide communications for high-priority users. This system has no satellites of its own. Instead, its transponders ride as passengers on other satellites and provide a communication node independent of other systems on their host satellites.

The Defense Satellite Communication System (DSCS) operates in the SHF arena (**Fig. 2**). These satellites are designed to provide long-haul communications between both fixed and mobile stations for strategic and tactical users. Occupying geostationary positions over the Equator, the DSCS can operate on a great number of SHF channels and has independently operating transponders, a multibeam antenna (allowing the formation of a transmission beam of nearly any size and shape), limited antijam capability, and a 10-year life-span.

Milstar, the most advanced component of the United States' military satellite communication system, operates mostly in the EHF region and consists of extremely large and advanced satellites (**Fig. 3**). It was originally designed only for strategic emergency situations, but the emphasis now is on both strategic and tactical missions. It uses advanced techniques to achieve a high degree of reliability despite advances in electronic warfare. It is highly survivable and jam-resistant, and provides secure worldwide communications. Some advantages of EHF communication for the military user are that it is highly directional, makes better use of available bandwidth, and recovers quickly from the propagation degradation caused

Fig. 1. Fleet Satellite Communication System satellite. (*U.S. Air Force*)

Fig. 2. Defense Satellite Communications System, Phase III (DSCS III) satellite. (*U.S. Air Force*)

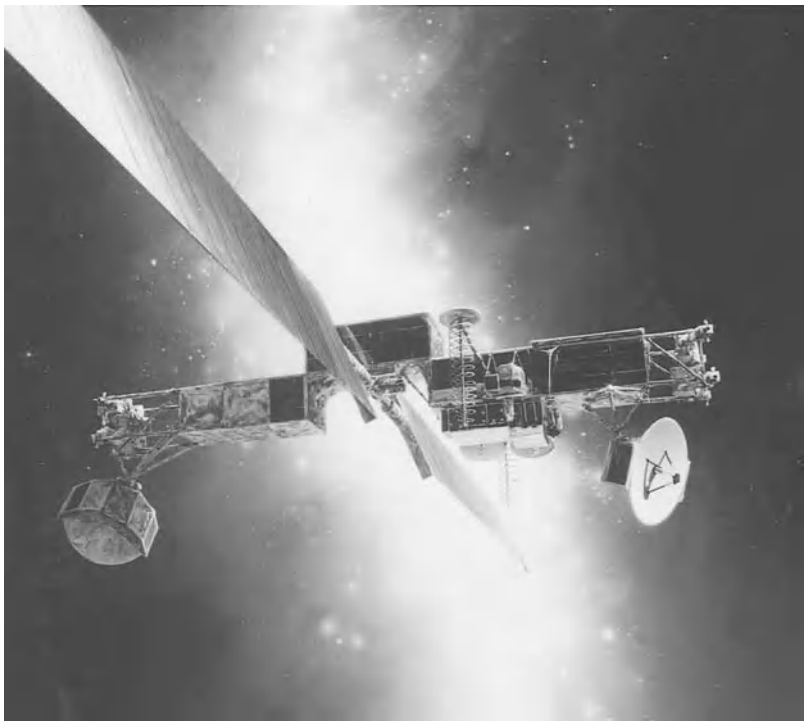

Fig. 3. Milstar satellite. (*TRW*)

by a high-altitude nuclear detonation. Additionally, each satellite can reconfigure its onboard equipment to make best use of multiple transponders and antennas, including spot-beam antennas. Milstar satellites can cross-link with other satellites so that reliance on intermediate ground stations is minimized. They are designed to be backward-compatible with the Fleet and Air Force Systems.

For training purposes, the U.S. Navy operates the Petite Amateur Naval Satellite (PANSAT) to provide operational communication experience for naval officers.

Russian satellites. Russia maintains a sophisticated three-tier network of satellites: small, low-altitude satellites; sophisticated *Molniya* satellites in semisynchronous (12-h) orbits; and satellites in geostationary orbit.

The *Molniya* spacecraft are notable for their innovative orbit, designed to maximize coverage over the high latitudes of the Asian landmass. They fly 12-h, highly elliptical orbits inclined at 63° with respect to the Equator so that any four can be spaced to give virtually continuous coverage over all of Asia, including areas north of 70° latitude that do not normally have satellite coverage. They repeat a ground trace each day over Russia and North America, lingering for 11 h over the Northern Hemisphere and then sweeping low and fast over the Southern Hemisphere.

In geostationary orbit are the *Raduga* spacecraft. At nearly 4500 lb (2000 kg), they provide telephone and telegraph relay and military communications.

Other countries and organizations. Skynet is a British military communication satellite system developed with the assistance of the United States. Several satellites have been launched to complete a constellation in geostationary orbit. They are compatible with the United States' Fleet System and certain ground terminals.

China has a sophisticated satellite communication system of which some channels are devoted to military and government use. France uses a Système de Radio Communications Utilisant un Satellite (SYRACUSE) package on each of its TELECOM satellites to provide secure links for the French Ministry of Defense. Both Spain and Brazil (Brasilsat) have communication satellites in orbit with some channels devoted to government and military purposes. Chile has the *FASAT* satellite, which provides message relay for the Chilean Air Force. The North Atlantic Treaty Organization (NATO) has satellites positioned in geostationary orbit over the Atlantic Ocean that are used for command and control of NATO forces. *See* COMMUNICATIONS SATELLITE.

Positioning and navigation. Military forces must be able to quickly and precisely determine their position on the ground, in the air, or at sea.

GPS. The Navstar Global Positioning System (GPS) [**Fig. 4**] is the most recent addition to positioning and navigation satellite systems and is the most accurate and reliable satellite navigation system available. It includes 25 spacecraft in semisynchronous (12-h) orbits inclined at 55° to the Equator at 11,600 mi (18,700 km) altitude. Of these, 21 are operational and 4 are spares. The inclined orbits provide worldwide coverage, including the Poles. Control stations around the world keep GPS satellites precisely calibrated and their orbits aligned. Each GPS satellite contains an atomic clock and transmits a continuous time signal and other information to receivers on Earth. The receiver must acquire and track these signals, decode the data, and then make pseudorange and relative velocity calculations. The fully operational GPS system allows a user anywhere on Earth to receive the transmissions of at least four satellites at once. Triangulation with these satellites provides a very accurate reading of position and velocity in three dimensions. Advanced versions of GPS spacecraft offer increased redundancy, can be reprogrammed from the ground, and feature cross-link ranging to improve accuracy to less than 20 ft (6 m). GPS helps guide precision weapons such as cruise missiles and also aids search and rescue, air refueling, air combat missions, mapping, geodetic surveys, ground troop movements, and other missions.

Other systems. Russia also maintains global navigation satellite systems. Its Tsikada/Nadezhda low-Earth-orbit system functions similar to the United States' decommissioned Transit system. In addition, Russia operates the GLONASS navigation system. Similar to GPS, the system is less complex, but its satellites have proven less reliable than the United States' version. Receivers are available that will accept navigational data from either GPS or GLONASS.

The international COSPAS/SARSAT system was established in 1979 by a coalition of the United States, France, Canada, and the former Soviet Union. It is a satellite-based search and rescue system. With no satellites of its own, its sensors ride on several

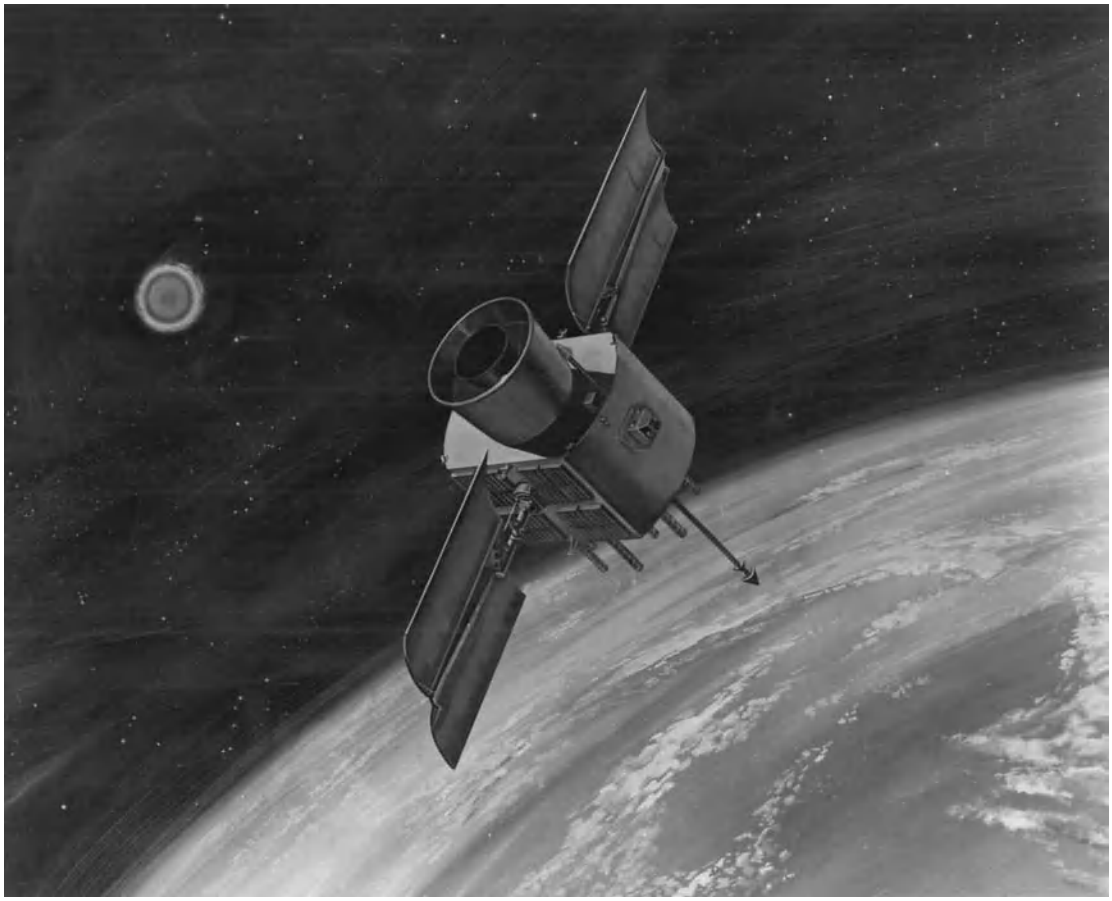

Fig. 4. Global Positioning System (GPS) satellite. (*U.S. Air Force*)

satellites, including U.S. National Oceanographic and Atmospheric Administration (NOAA) weather satellites. *See* SATELLITE NAVIGATION SYSTEMS.

Meteorology. From orbit, it is possible to obtain a wide-field-of-view image of the Earth, its cloud formations, and their movements. This meteorological information is valuable for military planning and operations.

United States meteorological satellites. The *Television Infrared Observation Satellites* (*TIROS*) have, for decades, traveled in Sun-synchronous, low Earth orbits providing images of cloud cover, snow, ice, and the sea surface. The NOAA/TIROS system continues to operate.

The Defense Meteorological Satellite Program (DMSP) [**Fig. 5**] consists of several satellites in low-Earth, Sun-synchronous, polar orbits at an altitude of 517 mi (833 km). These satellites are spaced to provide complete coverage of the Earth at various times of the day and night. The Sun-synchronous orbit allows them to record images of the Earth at the same local time every day. They have some antijam and maneuver capability and view the Earth in infrared and microwave spectra providing information on clouds as well as water vapor, ozone, and temperature. Other sensors examine the near-Earth space environment. The satellites are designed to support military operations through all levels of conflict.

Broader images of an entire hemisphere for military use may be obtained from United States' *Geostationary Operational Environmental Satellite* (*GOES*) spacecraft and other geostationary meteorological satellites.

European and Asian countries. Russia operates the Meteor system of satellites. As many as five satellites are in low Earth orbit similar to *TIROS* with an

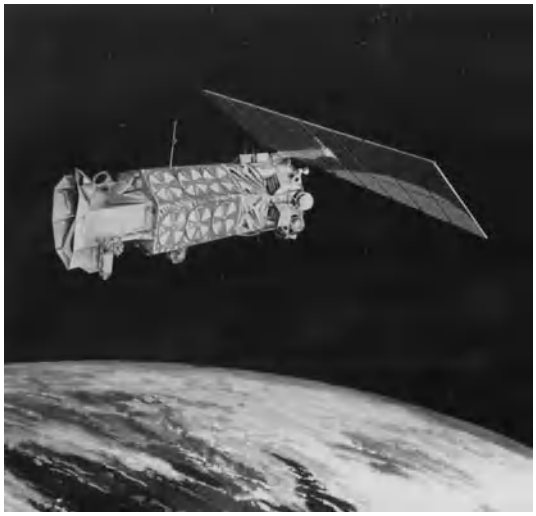

Fig. 5. Defense Meteorological Satellite Program (DMSP) 5D-2 satellite. (*U.S. Air Force*)

orbital inclination of 81.2°. They image in the infrared spectrum.

Several other countries maintain a meteorological satellite capability that provides useful information for military operations, although not specifically designed for military use. Meteosat is a European system whose weather satellites are owned and operated by Eumetsat, a consortium of 17 nations. The satellites are positioned in geostationary orbit providing the best view of the European continent, and provide infrared, visible, and water-vapor imagery for weather forecasting, storm tracking, and data analysis.

China uses *Feng Yun*, a Sun-synchronous satellite that provides weather data for China and east Asia. The signal formats for this satellite are compatible with those of *TIROS*, so receivers designed for *TIROS* can also receive data from *Feng Yun*.

India operates *INSAT*, a multipurpose satellite in geostationary orbit. Among other capabilities, *INSAT* has weather sensors to monitor south Asia.

Japan's *Geostationary Meteorological Satellites* (*GMS*) operate in geostationary orbit over the Pacific Ocean. They have visible and infrared radiometers and space environment monitors. *See* METEOROLOGICAL SATELLITES.

Reconnaissance and surveillance. Military reconnaissance and surveillance satellites offer near-real-time unrestricted access over almost any area on Earth. Operating in many parts of the electromagnetic spectrum, they can be used to observe weapons development and deployment of forces, and to provide warning of attack by ground forces as well as targeting intelligence, technical intelligence on enemy capabilities, electronic intelligence, and bomb damage assessment. Space reconnaissance augments reconnaissance by ground and air systems, resulting in a clearer picture of the battlefield, a more complete intelligence preparation of the battlefield, and more informed strategic and tactical decision making.

Russia frequently launches photoreconnaissance and electronic intelligence satellites. France has the *HELIOS* reconnaissance satellite in low Earth, Sun-synchronous orbit, and China has the *Fanhui Shi Weixing* satellite and others in low Earth orbit for photoreconnaissance and remote sensing purposes.

Early warning. Early warning satellites provide information on missile launch and nuclear detonation that give governments time to make strategic military decisions.

The United States' missile early warning satellite system is the Defense Support Program (DSP) [**Fig. 6**]. Its satellites operate in geostationary orbit and survey the Earth with an infrared sensor that detects heat sources. During Operation Desert Storm in 1991, DSP satellites detected the launch of Iraqi Scud ballistic missiles.

The Nuclear Detonation Detection System (NDDS) sensors on DSP and GPS satellites are a means of detecting, locating, and reporting atmospheric nuclear detonations on a global, near-real-time basis. This system does not have any satellites of its own. Being part of the GPS system allows more than one sensor to view an explosion and, by triangulation, enhances geolocation and quantification. The sensors detect the visible light, x-rays, and electromagnetic pulse given off by a nuclear explosion. *See* ELECTROMAGNETIC PULSE (EMP).

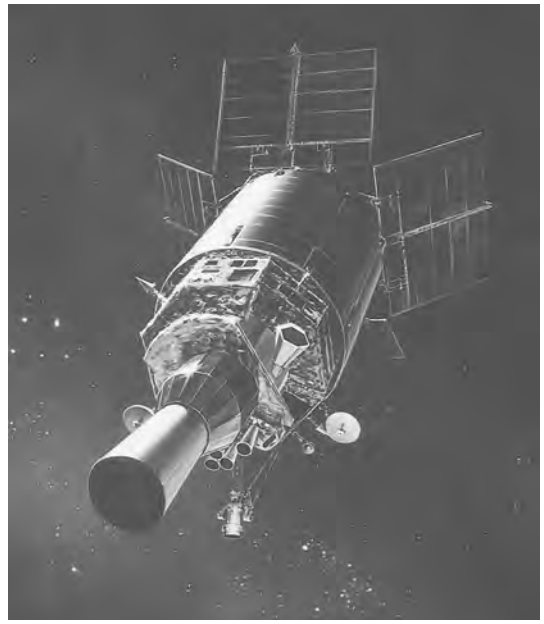

Fig. 6. Defense Support Program (DSP) satellite. (*TRW*)

Russia maintains early warning satellites both in geostationary orbit and in orbital paths similar to that of the 12-h *Molniya* spacecraft.

Remote sensing. Meteorological and reconnaissance-surveillance satellites are not the only Earth-observing satellites used for military purposes. Remote-sensing satellites also afford a unique view of Earth, providing vital information to military forces. The images produced by these satellites are used to conduct route reconnaissance, analyze waterways, assist in exercise and strike planning, and provide up-to-date maps for forces deploying to unfamiliar areas.

Important information is gathered not only in the visible spectrum but in other bands of the electromagnetic spectrum. Multispectral imagery is a process of imaging the Earth at various wavelengths. These data yield information on types of minerals and vegetation as well as trafficability. Those areas that cannot be easily surveyed from the ground or by aircraft make ideal targets for remote-sensing satellites.

The United States maintains *Landsat* satellites in Sun-synchronous orbit. Each satellite carries a multispectral scanner and a thematic mapper. The scanner records images in four different spectral bands; the mapper, in seven. The best resolution is 100 ft (30 m).

Russia currently maintains the *RESURS* series of satellites in low Earth polar orbit for remote sensing photography in both visible and infrared wavelengths. Their best resolution is 7 ft (2 m). They are

designed to be maneuvered on orbit and typically have a 25-day mission life. After the photography period, a film canister separates, reenters the Earth's atmosphere, and is recovered for processing.

France owns and operates *Système Probatoire d'Observation de la Terre* (*SPOT*) in low Earth, Sun-synchronous orbit. Similar to *Landsat*, *SPOT* spacecraft record in fewer spectral bands but with a better spatial resolution of 33 ft (10 m). The Earth images they produce are mostly for use in agriculture, forestry, geology, and resource and environmental management. They orbit at 516 mi (832 km) and at an inclination slightly greater than that of *LANDSAT* at 98.7°.

The Japanese operate the *Marine Observation Satellite* (*MOS*) and *Japan Earth Resources Satellite* (*JERS*). *MOS* is a multispectral imaging spacecraft and, although designed primarily for oceanic studies, its data are useful to topographers. *JERS 1* carries infrared optical sensors and a synthetic aperture radar, allowing it to see through clouds, which simplifies the remote-sensing data-gathering process.

India operates the *Indian Remote Sensing Satellites* (*IRS*) in a low Earth, Sun-synchronous orbit. Current versions of this spacecraft carry a panchromatic camera with 19-ft (5.8-m) resolution, a linear imaging camera for visible and infrared wavelengths at varying resolutions, and a wide-field sensor. *See* REMOTE SENSING.

Geodesy. Geodesy is the study of the Earth's size and shape. Geodetic data are important to the military in that the data affect position determination, navigation, map making, and a variety of other missions. Almost all satellites can be used for geodesy, provided their position in space can be accurately determined by optical or electronic means from the Earth.

GPS satellites are used for geodetic measurements as well as the *GEOS* satellites. This is a joint effort among the U.S. Department of Defense, NASA, the Department of Commerce, and other countries. In addition, the *GEOSAT* follow-on was launched for the U.S. Navy and has radar altimetry as its primary mission.

Russia uses the *Zeya* satellite in low Earth orbit for geodetic purposes. The satellite is a spheroid, slightly smaller than 3.3 ft (1 m) in diameter. It carries 20 laser reflectors for geodetic use and GLONASS and GPS receivers for precise orbit determination.

Research. In addition to the missions discussed above, there have been hundreds of military research and technology spacecraft, as well as thousands of experimental investigations for military purposes on space vehicles launched by all space-faring nations.

Examples of United States research satellites include a secondary use of the *DMSP* weather satellite to map the aurora borealis and australis with their low-light sensors and some military testing of the *Advanced Communication Technology Satellite* (*ACTS*), which incorporates high-speed, high-volume satellite communications. The *Midcourse Space Experiment* (*MSX*) satellite tested the detection of missile signatures. The *Miniature Sensor Technology Initiative* (*MSTI*) satellite tested sensors for missile defense. The *Space Technology Experiment Platform* (*STEP*) tested new technologies for the U.S. Air Force. *Mighty Sat* satellites, launched for the Air Force, test emerging technologies such as advanced solar cells and a composite bus. The 6000-lb (2700-kg) *Advanced Research and Global Observation Satellite* (*ARGOS*) provides the Air Force with a platform to collect data on the Earth's environment and to perform technology demonstrations for military space programs.

Israel uses the *OFEQ* satellite as a test bed for satellite reconnaissance and remote sensing.

Daniel F. Moorer, Jr.

Bibliography. *Interavia Spaceflight Directory*, annually; *Jane's Spaceflight Directory*, annually; TRW, *Space Log*, annually; U.S. Army Space Institute, *Space Reference Text*, 1993.

Milk

The U.S. Food and Drug Administration defines milk as the lacteal secretion, practically free from colostrum, obtained by the complete milking of one or more healthy cows and containing not less than 8.25% milk solids (not fat) and not less than 3.25% milk fat. Among mammals, humans utilize milk as a source of food. The dairy cow supplies the vast majority of milk for human consumption, particularly in the United States; however, milk from goats, water buffalo, and reindeer is also consumed in other countries. Without qualification, the general term milk refers to cow's milk.

Composition. Average composition of milk is 87.2% water, 3.7% fat, 3.5% protein, 4.9% lactose, and 0.7% ash. This average composition varies from cow to cow and breed to breed, as well as during the lactation period and the different seasons of the year, and is dependent upon the feed, nutritional level, age, and health of the animal and mammary gland. *See* DAIRY CATTLE PRODUCTION; LACTOSE; PROTEIN.

Nutritionists state that milk is the most nearly perfect food. Although this is true, there are some limitations, particularly in iron and vitamin C. Whole milk and skim milk are classified as excellent sources of calcium, phosphorus, and riboflavin because 10% of the daily nutritional requirement is supplied by not over 100 kilocalories (420 kilojoules). These two beverages are also classified as good sources of protein and thiamin; and whole milk is a good source of vitamin A. To be classified as good, the source must contribute 10% of a nutrient in not over 200 kcal (840 kJ). Milk is a good source of protein rich in the essential amino acids. *See* FOOD.

Standards for quality. The Food and Drug Administration publishes the *Grade A Milk Ordinance*, which is the basic standard for milk sanitation practice for most states and all interstate milk shippers. This ordinance defines milk and each milk product, establishes the requirements for the production of grade A raw milk, gives instructions for the inspection of dairy farms and processing facilities, and gives

sanitation requirements for the production of raw and pasteurized milk and milk products. The Food and Drug Administration certifies the sanitary quality of all raw milk shipped interstate. To accomplish this, the service inspects farms and collection depots to determine compliance to the code.

Grade A raw milk for pasteurization must be cooled immediately and maintained at 45°F (10°C) or less until processed, cannot exceed 100,000 bacteria/ml per producer, and must have no detectable antibiotic residues. Grade A pasteurized milk and milk products must be cooled immediately and maintained at 45°F (10°C) or less, contain no more than 20,000 bacteria/ml, contain no more than 10 coliforms/ml, and show a negative phosphatase test even at the pull date or expiration date. The laws of some states vary from those established by the code. Agriculture Handbook No. 51 (USDA, Agriculture Marketing Service) shows a compilation of all the federal and state standards for composition of milk products.

Fluid products such as whole milk; low-fat milk; skim milk; flavored milk; coffee, light, whipping, and heavy cream; and half-and-half are all defined by fat and solids contents. Filled (substitution of vegetable fat or milk fat) or imitation (made from nondairy ingredients) products are claiming increased sales. The lower cost of vegetable fat, and therefore lower cost to consumer, and the total lack of standards of identity are two major reasons for the increase. *See* FAT AND OIL (FOOD).

Nearly all dairy processing equipment in use in the United States has milk contact surfaces made from stainless steel, and these surfaces legally must be sanitized before each use. The fact that stainless steel is used in construction also permits the use of strong cleaning solutions, both alkaline and acid, to properly remove residues on these surfaces. Negligible corrosion or leaching of constituents from stainless steel occurs when these detergents and cleaning compounds are used properly. After being cleaned, the entire milk contact surface is sanitized or essentially sterilized with a chemical usually containing either active chlorine or iodine to render the surface free of pathogenic and most other bacteria. *See* DAIRY MACHINERY.

Processing

Most raw milk collected at farms is pumped from calibrated and refrigerated stainless steel tanks into tank trucks for delivery to processing plants. The 10-gallon (38-liter) can that was common some years ago is found only on a few farms.

Collection and intake. The bulk truck drivers are required to check flavor, temperature, and volume of milk in the farm tank and to collect a sample of raw milk for analysis before pumping the milk into the truck. At the receiving area of the processing plant or receiving station, the milk in the farm truck is weighed and pumped into the plant through flexible plastic and stainless steel pipelines.

Separation and clarification. The actual processing of raw milk begins with either separation or clarification. These machines are essentially similar except that in the clarifier the cream and skim milk fractions are not separated. These machines have large high-speed sealed bowls into which whole milk is introduced through ports at the bottom. In the separator milk passes upward through holes in the closely spaced conical discs. At this point the specific gravity is approximately 1.0. Milk fat (sp gr 0.93 at 68°F or 20°C) is forced to the center, but skim milk (sp gr 1.037 at 68°F or 20°C) is forced to the outside. Leukocytes, debris, some bacteria, and sediment carried with the skim milk fraction are deposited in the periphery of the bowl. This single function of a clarifier precludes sediment in homogenized milk. Whole milk or cream and skim milk travel upward to the top of the bowl. A separator contains an upper conical plate without holes to prevent admixing; the clarifier contains no such device. *See* CENTRIFUGATION.

Separators have two discharge pipes, one for cream and one for skim milk. Clarifiers have only one pipe for whole milk. Separators have a device called a cream screw by which the fat content in the cream is regulated. This screw allows more or less cream to pass out through the discharge pipe.

Many processors have units called standardizer-clarifiers which separate only a small fraction of the fat from the raw whole milk. Through manipulation of the cream screw, the amount of fat removed can be regulated. This facilitates the production of milk of standard fat content even though that in the raw product may vary. Recent modifications of separators, clarifiers, and standardizer-clarifiers permit the units to be cleaned without disassembly and the bowls to be self-cleaned during operation. *See* MECHANICAL SEPARATION TECHNIQUES.

Pasteurization. Milk is rendered free of pathogenic bacteria by pasteurization. This is accomplished in a manner so that every particle of milk is heated to a specified temperature and held at that temperature for a specified time. The U.S. Public Health Service stipulates at least 145°F (63°C) for 30 min when milk is pasteurized in a vat or at least 161°F (72°C) for 15 s when milk is pasteurized continuously. Cream and chocolate milk must be heated to either at least 150°F (66°C) for 30 min or 166°F (74°C) for 15 s. Frozen dessert mixes must be heated to at least 150°F (68°C) and held at that temperature for 30 min or 175°F (79°C) for 25 s. The greater heat treatments for the latter products are required because fat and sugar in greater concentrations than that found in whole milk provide heat resistance to bacteria normally found in milk. Most city and state codes follow these requirements.

Pasteurization requirements were originally established to provide for total destruction of *Mycobacterium tuberculosis* with a safety factor added; this bacterium will not survive 145°F (63°C) for 6 min or 155°F (68°C) for 30 s. Certain spore-forming pathogens will survive pasteurization: however, subsequent refrigeration precludes growth. These organisms, such as *Clostridium botulinum*, become hazardous only after production of toxin and growth.

The rickettsia causing Q fever in humans can be transmitted through milk, but is destroyed at 145°F (63°C) for 30 min.

Pasteurization on a batch operation requires a jacketed vat where steam or hot water can circulate and heat the milk. This temperature requires the longer times at lower temperatures to accomplish pasteurization (LTLT pasteurizer). If there is a tendency of the particular product to foam, then a space heater is utilized at the top of the vat to ascertain that every particle of product is adequately heat-treated. Modern methods of processing milk and milk products utilize the high-temperature short-time (HTST) pasteurizer. If milk passing through this pasteurizer is not at a high enough temperature, the flow diversion valve at the end of the holding tube is activated and the milk is diverted back to the surge or float tank to be reprocessed.

In Europe and to a limited amount in the United States, milk and milk products may be ultra-heat-treated (UHT). This process may use equipment similar to that used for HTST or tubular heating equipment. Aseptic equipment is first sterilized by circulating hot water at 295°F (146°C) for 30 to 45 min, then temperatures are adjusted to allow processing of milk, and packaging is done in sterile containers usually made of five laminated layers of paperboard, polyethylene, and aluminum foil. The UHT processing requires a minimum heat treatment of 280°F (138°C) for 2 s. UHT dairy foods have extended shelf life because all of the bacteria that would survive even HTST pasteurization have been destroyed, and thus the food needs no refrigeration. A typical packaging machine for UHT sterilized milk shows the need for absolute sterility.

Homogenization. Fat globules in fluid milk products are broken by homogenization into sizes that are 2 micrometers or less and thus are relatively unaffected by gravitational forces. Most fluid milk is homogenized. The flow of product into the homogenizer head is through suction valves, where it is forced by pistons through discharge valves and finally through valves that cause shearing of the fat globules. These latter two valves can be regulated at different pressures, depending on the product being homogenized. Total pressures as high as 3000 lb/in.^{2} (21 megajoules) are used. This reflects 500–700 lb/in.^{2} (3.4–4.8 mJ) on the second stage and 2300–2500 lb/in.^{2} (16–17 mJ) on the first stage. In order to properly homogenize milk, the product must be heated to at least 140°F (60°C) to liquefy the fat globules and inactivate the enzyme lipase to prevent a rancid flavor. Homogenizers are commonly connected into the HTST pasteurizer where the hot milk comes from the regenerator or heater section.

Several theories explaining homogenization or globule fractionation include shearing, shattering, and cavitation. Shearing is the effect produced as whole milk is forced through a minute orifice at high speed. Shattering is the effect that occurs when whole milk under high velocity strikes a flat surface, such as the impact ring. Cavitation is the effect produced as whole milk changes abruptly from an area of high pressure before the valve to an area of somewhat reduced pressure after the valve. All three are evident in a homogenizer.

Efficiency of homogenization can be ascertained by examining the size of the fat globules under 1000× magnification. The Food and Drug Administration specifies that the fat content of the upper 100 ml of a quart of homogenized milk that has been undisturbed for 48 h cannot differ by more than 10% from that of the remainder.

Vitamin fortification. Most milk is fortified with 400 international units (IU) of vitamin D per quart, and most low-fat and skim milk is fortified also with 2000 IU of vitamin A per quart. These vitamins as concentrates are added either by automatic dispensing with a peristaltic pump into a continuous flow of milk prior to pasteurization or as a single quantity in a batch operation. Vitamin concentrates have a potency range of 3000–200,000 IU/ml.

Nonfat dry milk can be enriched with vitamins A and D by adding the vitamin concentrates to the condensed milk just prior to drying or by blending in a dry beadlet form to the dry milk. *See* VITAMIN; VITAMIN A; VITAMIN D.

Defects

When milk and milk products are mishandled or good management practices with cows are not followed, some defects can be observed in the flavor and appearance of fluid dairy foods.

Flavor defects. There is an apparent delicate balance between the flavor constituents in milk. These are very subtle flavors, with some well below the threshold concentration where organoleptic detection is impossible. Some of the flavoring materials occurring naturally in milk serve only to potentiate other flavors. Humans vary considerably in their ability to detect flavors. Whereas milk normally has a slightly salty character as well as a sweet flavor, the degree to which these background flavors are observed organoleptically varies considerably. *See* TASTE.

Cooked or heated flavor. Prominent in sterilized milks and baby formulations is a heated flavor. To a much lesser degree this flavor can be detected in fluid milk, cream, and ice cream mixes. Conventional pasteurization generally will not cause a noticeable flavor. In fact, a slightly cooked flavor is now desired by most consumers. However, when temperatures and holding times greater than the minimums established for pasteurization are used, this defect is discerned.

Heat causes the whey proteins, β-lactoglobulin and euglobulin, to change from the normally helical or springlike structure to random structures. The energy supplied by the heat breaks the chemical bonds that hold these coils together. The result is that numerous sulfhydril (SH) groups are exposed, which contribute to the flavor.

Prolonged exposure of milk to elevated temperatures will also cause some discoloration or browning. Most of this color change is caused by amino acid–sugar reactions or Maillard-type browning. *See* MAILLARD REACTION.

Feed flavors. Milk is a good absorbent for flavors whether the flavor itself is in the feed of the animal or in the barn air. Changes from one forage crop to another, feeding of silage within 2 h of milking, or allowing cows to feed on cabbage, onion, or other highly flavored crops or some weeds will cause objectionable flavors in raw milk. In the southern sections of the United States, deodorization of milk is mandatory to remove objectionable wild-onion flavors. This process involves flashing or injection of a thin stream of milk at pasteurization temperature into a vacuum chamber which can be included in an HTST pasteurizer between the heating chamber and the flowing diversion valve.

Oxidized and sunlight flavors. Oxidized flavor is probably the most important single flavor defect in milk and milk products. This flavor is also described by terms such as tallowy, metallic, and cardboard. Lipid material is the source, but the specified flavor compounds differ apparently among milk products.

After homogenization or some degree of churning, milk fat itself becomes exposed as the phospholipids no longer totally protect the fat globules; then the fat itself can be oxidized. Of the factors essential to the development of this flavor, atmospheric oxygen and copper are important. An adequate heat treatment is a commonly used process to retard the development of this flavor. The chemistry of this oxidation is a complex phenomenon.

When milk is exposed to light for any period of time, another defect can occur. This is sunlight or activated flavor and is attributed to a reaction with the amino acid, methionine, changing to an aldehyde, methional, in the presence of riboflavin (vitamin B_2). Milk in metal containers is resistant to this change because no light penetrates. Some years ago amber glass bottles were used to prevent this defect in homogenized milk. Glass of this type reduces the amount of energy transmitted but is not totally a preventative.

Rancid flavor. Modern automated methods of handling raw fluid milk promote the onset of this flavor. Each fat globule in milk has a surface coating composed of phospholipids (lecithin and cephalins) and proteins from the serum portion (euglobulin and some casein), among other constituents. Moreover, the enzyme lipase is also present in raw whole milk and presumably is associated with casein.

Raw whole milk is stable to the action of lipase until some physical force disrupts the membrane surrounding the fat globules. Forces such as excessive agitation through prolonged pumping in pipelines, mild heat treatment followed by cooling, or freezing can disrupt this membrane. Lipase acts on fat to cause hydrolysis of the triglycerides into glycerol and free fatty acids. Because milk contains a significant proportion of short-chain fatty acids which have strong aromas, the defect is easily detected by taste. Lipase, as with most enzymes in milk, is inactivated by proper pasteurization.

Physical defects. Fluid whole milk is a rather stable product; however, subjection to adverse conditions can cause some physical changes.

Destabilization. If milk is frozen and then thawed, the fat and protein will aggregate. This can be observed in the thawed fluid product by microscopic flakes of protein and macroscopic fat specks.

Feathering. A defect associated with coffee cream occurs principally because of a salt imbalance in the cream (calcium, magnesium, citrate, or phosphates), too high an acidity, or improper homogenization. Cream is said to feather when a grayish, flaky scum floats to the surface when the cream is added to hot coffee.

Cream plug. Cream and occasionally fluid milk will show some signs of a thick surface. The aggregation of fat that either floats on the surface of fluid milk or forms a thick mass on the surface of cream is caused by churning. Prevention of churning eliminates the defect. Therefore, warming and cooling, freezing, and excessive agitation should be avoided.

Testing for Quality

Laboratories associated with dairy processing operations, or private laboratories that analyze milk and milk products, perform many tests to determine compliance with local, state, and federal regulations; quality control; processing efficiency; and payment to farmers. The test performed in most laboratories for homogenization efficiency is described above in the discussion of homogenization. The subsequent discussion involves other commonly used tests.

Fat determinations. There are several methods used for the determination of fat in milk and its products. The Babcock test developed by S. M. Babcock in 1890 is classical. This test and the European counterpart, the Gerber test, are volumetric measurements of the quantity of fat in a sample. Sulfuric acid is used to digest the 18-g (0.63-oz) sample that was added to the special flask. The heat that develops liquefies the fat. Centrifugal force is used to bring the fat into the calibrated neck of the flask. The Babcock test is used as the basis for payment for raw milk received from farms.

Fat can also be determined by extraction from a sample using specific organic solvents and subsequent evaporation of these solvents to obtain a quantity of fatty residue which is weighed. This procedure is outlined in the Mojonnier and Roese-Gottlieb methods. The solvent extraction methods are considered by far the most accurate ever developed; however, they are time-consuming. One skilled operator can complete 30 samples per 8-h day.

An approved (Association of Official Analytical Chemists) method for fat determination in raw milk involves only the use of the Milko-Tester manufactured by Foss Electric in Denmark. This automated process consists of heating, homogenization to render fat globule size uniform, dilution, and photoelectric sensing of solution opacity. Each analysis requires less than 1 min. Frequent, daily calibration of this instrument is required. Other similar instruments use infrared adsorption to assay fat, protein, and lactose in milk. All automated instruments need calibration using pretested milk to assure accurate analysis.

Total solids or moisture. Whatever is not lost during analysis for moisture in a food product is considered as total solids. There are several different methods approved by the Association of Official Analytical Chemists as legal for the analysis of milk and milk products. Fundamentally, all involve a heat treatment applied to a sample and then vacuum desiccation or subjection of the sample to treatment in a vacuum oven for a stated period. Other less accurate ways involve the use of specialized hydrometers called lactometers which determine the specific gravity of a fluid sample.

Freezing point. The freezing point of milk is a constant at 31°F (−0.55°C) and consequently can be used for the detection of the addition of water. For this determination a cryoscope is employed which contains a bath to lower the temperature of the sample and a precise thermometer or electrical readout system to determine the temperature at which stable freezing occurs. The sample will be slightly supercooled, and when crystallization occurs, the temperature will rise and plateau. This is the moment of measurement.

Pasteurization efficiency. To determine whether milk or cream has been pasteurized properly, a test to determine the amount of remaining phosphatase enzyme is employed. The test involves the mixing and incubation of a small sample of milk, buffered solution, and disodiumphenylphosphate. An indicator is added to show colorimetrically the amount of disodiumphenol liberated. Properly pasteurized milk normally gives a negative test because only 0.1% of the phosphatase remains. However, if this milk were separated and a phosphatase test performed on a sample of the cream, then a positive test would result. This is because this phosphatase enzyme is associated with the fat phase. The phosphatase test is so sensitive that it can determine underpasteurization by 1°F (1.8°C) or adulteration with 0.1% raw milk.

Inhibitory and other foreign substances. On occasion raw milk becomes contaminated with antibiotics used in the control of mastitis and other bovine diseases. Milk containing such foreign substances may not be legally sold.

Detection of antibiotics is mandatory, particularly in cheesemaking and other manufacturing procedures that require bacterial fermentation because approximately 0.2 unit of antibiotic will inactivate most of the fermentating bacteria. Agar which has been inoculated with *Bacillus stereothermophilus* spores is poured into a petri dish and hardened, and filter disks dipped in raw milk are placed on the surface. After incubation, the plates are examined for clear zones around the small disc and a comparison is made with a standard to determine the approximate concentration of antibiotic, if any.

Bacteriological determinations. The microbial content of milk and milk products is an important determination because it relates directly to the sanitary quality and the conditions under which the raw and finished products were handled. A direct microscope technique or an agar plate method is used.

Milk is an excellent growth medium for microorganisms which, if permitted to grow, will produce changes that render the milk unfit for human use. The U.S. Public Health Service Code, which most states and municipalities have adopted, stipulates that pasteurized milk and milk products can contain no more than 20,000 bacteria/ml and no more than 10 coliform bacteria/ml at pull or expiration date. To ascertain compliance with the law and general product quality, processing plants and regulatory agencies analyze samples of milk and its products for bacterial content. A representative sample of the milk is diluted in sterile buffered water and a quantity of the solution (usually 1 or 0.1 ml) is transferred to a petri dish. Bacteriological growth media (plate count agar) is poured into the dish, swirled to distribute the sample, and allowed to gel. The dishes are inverted and incubated at 90°F (32°C) for 48 h. Colonies of bacteria show as white or pigmented growth on the surface or subsurface. Each is counted as one colony times the dilution factor to give total count.

Coliform counts are performed in a similar manner except that a direct sample without dilution is used. Violet-red bile agar is added, swirled, allowed to gel, and overplated with more agar to assure subsurface growth. After the agar has gelled, the plates are inverted to prevent the collection of moisture on the surface of the growth media and the spreading of colonies, and are then incubated 24 h at 90°F (32°C). Coliform bacteria are recognized by their red pigmentation.

One frequent test performed by laboratories associated with processing facilities is to determine the storage stability of the finished product. A container of fluid milk, for example, is stored at 45 or 50°F (7.2 or 10°C) for 1 week. At the end of this period, a standard plate count would determine the quality of the product. If bacterial analyses show the product to be inferior, then an active program is initiated to correct the problem and protect quality.

The significance of bacteria in milk and milk products depends upon the organism and the product. Coliform organisms usually are totally destroyed by pasteurization; therefore, the presence of these organisms in the finished product indicates contamination after pasteurization. With present-day facilities for refrigeration, the psychrophilic bacteria, those that multiply at refrigerator temperatures, present problems. These are normally destroyed by pasteurization and are therefore postpasteurization contaminants. Meticulous cleaning and sanitizing in a processing plant is necessary to keep these microorganisms at a minimum in finished products.

Pasteurization practically precludes the presence of pathogens, and those that survive will not survive refrigerated storage. Milk and its products also contain thermoduric and thermophylic bacteria that do not grow at pasteurization temperatures but can grow at 131°F (55°C), respectively. Thermoduric organisms are transmitted to milk from poorly cleaned equipment, while thermophylic bacteria are usually contaminants from the farm.

Milk Microbiology

Aseptically drawn milk from healthy cows is not sterile. The interior of the udder is open to invasion by bacteria when the opening of the teat comes in contact with the air and bedding. The bacteria present in the udder are distributed internally by their own growth as well as by physical movement. However, only small numbers, averaging about 1000/ml of milk, of a few types are normally found in aseptically drawn milk, although much lower and much higher counts are often reported. During the milking procedure, bacteria are in largest numbers at the beginning and gradually decrease. *See* FOOD MICROBIOLOGY.

Contamination from external sources. There are many sources from which the milk can be contaminated by microorganisms during milking as well as during the subsequent handling of the milk. The most important are given below.

Stable air. This may contain dust in considerable quantities, especially in a dirty stable or when hay is distributed before the milking.

Flies and other insects. A fly may carry as many as 1,000,000 bacteria. Such a fly, falling into a liter of milk, will increase the bacterial count of the milk by 1000/ml, even without bacterial reproduction.

Coat of the animal. Soil, feed, and manure adhere to the cow's coat. During the milking process this material may fall from the coat and a portion of it may get into the milk. Dry manure is a source of heavy contamination. Proper care and preparation of the cow prior to milking precludes most contamination from this external source.

Feed. Hay and silage often contain a great number of spores. Milk may be easily contaminated with these when airborne.

Milk equipment. Equipment, such as pails, cans, coolers, pipelines, bulk tanks, and milking machines, is the most serious source of bacterial contamination. It is very important that utensils be made without seams and sharp corners to facilitate cleaning.

Milking personnel. If the milking personnel are not in good health or have infections on their hands, pathogenic bacteria may be added to the milk. Milk may serve as a carrier of human pathogens from one person to another or animal to human, hence the need for pasteurization.

Kinds of microoganisms. The saprophytic and pathogenic microorganisms in milk are discussed in this section.

Saprophytes. Saprophytic microorganisms live on dead or decaying organic matter. The important ones found in milk and dairy products are as follows:

1. Certain species of bacteria found in milk convert milk sugar into lactic acid and other byproducts. These bacteria are cocci as well as rods, and belong to such genera as *Streptococcus*, *Leuconostoc*, and *Lactobacillus*. Important species are *S. lactis*, causing spontaneous souring of milk and widely used for the making of cheese; *S. thermophilus*, found in fermented milks like yogurt; *Leuconostoc cremoris*, responsible for the butter-like aroma; and *Lactobacillus casei*, present in hard cheese.

2. Certain gram-negative asporogenous (nonsporeforming) rod-shaped bacteria commonly occur in the large intestine of animals. Best known are the species *Escherichia coli* and *Enterobacter aerogenes*. Their presence in pasteurized milk serves as a sensitive index of contamination after pasteurization.

3. The genus *Pseudomonas* is well known for causing spoilage, frequently with pronounced biochemical activity, especially on proteins and fats.

4. The genus *Bacillus* contains aerobic bacilli, like *B. subtilis* and *B. cereus*. Because their spores can survive pasteurizing and sometimes even the sterilizing treatment of milk, such bacteria can be important in causing spoilage of pasteurized and sterilized milk. Some species of the anaerobic genus *Clostridium* attack proteins (*C. sporogenes*) and some produce gas (*C. butyricum*). They are well known for causing defects in cheese.

5. Yeasts sometimes ferment carbohydrates and produce gas, and sometimes are lipolytic (hydrolyze fats). They occur as contamination in sour milk products like yogurt and butter and on cheese rinds.

6. Molds which actively live on carbohydrates, fats, and commonly proteins often spoil milk and dairy products, but some are useful. For example, *Penicillium roqueforti* is used in making blue-veined cheese (Blue, Bleu, Roquefort, Stilton, and Gorgonzola cheese).

Pathogens. The milk of diseased animals may contain living germs or pathogenic microbes, and the consumption of such milk (without heat treatment) by other animals and humans may then cause the disease to be transmitted. Tuberculosis, brucellosis (undulant fever in humans), Q fever (caused by *Coxiella burneti*), foot-and-mouth disease, and causative agents of mastitis may be so propagated. *See* BRUCELLOSIS; FOOT-AND-MOUTH DISEASE; TUBERCULOSIS.

In addition, milk of healthy animals may become contaminated with pathogens of other origin. Milk has been known to transmit in this manner typhoid and paratyphoid fevers, septic sore throat, diphtheria, and scarlet fever.

Some pathogens do not thrive well in milk but remain active a fairly long time. These are easily destroyed by pasteurization.

Cultured Products

Many fermented or cultured products are produced from milk. These fermentations require the use of bacteria that ferment lactose or milk sugar. These bacteria are of two general categories: homofermentative, those that produce only lactic acid from lactose; and heterofermentative, those that produce acetic acid, ethyl alcohol, and carbon dioxide in addition to lactic acid from lactose and a flavor precursor, acetoin, from citric acid. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; FERMENTATION.

Mother or starter cultures. Bacteria such as *Streptococcus lactis*, *S. cremoris*, *S. diacetilactis*, *S. thermophilus*, *Leuconostoc cremoris*, *Lactobacillus*

acidophilis, and *Lactobacillus bulgaricus* are among the common organisms used in cultures. Cultures of these organisms purchased from a commercial supply house are transferred several times to produce an active growth. Selection of the proper culture to use depends upon the talent and judgment of a trained technologist. Larger quantities of milk are inoculated and ripened to develop the appropriate acidity which is directly related to the populations in the culture, and then large vats of milk are inoculated to produce different fermented products. *See* BACTERIOLOGY.

Among the problems associated with fermentations, contamination of a culture with bacteriophage is probably the most devastating. Bacteriophages are viral agents that attack and destroy specific bacteria. To preclude this, manufacturers use special media for control of phage growth. Furthermore, rotation of the strains of the particular organisms used in a culture program is advantageous. Obviously, antibiotics and improper manipulations of cultures can be of serious consequence. *See* BACTERIOPHAGE.

Cultured buttermilk. Skim milk or low-fat milk is pasteurized at 180°F (82°C) for 30 min, cooled to 72°F (22°C), and inoculated with an active starter culture containing *S. lactis* and *Leuconostoc cremoris*. This mixture is incubated at 70°F (21°C) and cooled when acidity is developed to approximately 0.8%. This viscous product is then agitated, packaged, and cooled. The desired flavor is created by volatile acids and diacetyl; the latter is produced by *L. citrovorum*.

So that the modern product will resemble the classical buttermilk, resulting from the churning of cream to butter, manufacturers often spray or otherwise introduce tiny droplets of milk fat into the product during agitation.

Cultured sour cream. Cream containing 18–20% milk fat is processed and inoculated similar to the milk used in the manufacture of buttermilk. One additional procedure involves homogenization of the hot, pasteurized cream prior to cooling to inoculation temperature to control the consistency. Because of its high viscosity sour cream is often ripened in the package to an acidity of 0.6%.

Yogurt. One of the oldest fermented milks known is yogurt. Historically the people of the Middle East relied on yogurt as an important food item. Later, consumption increased rapidly in Europe because of the suspected correlation with longevity.

Yogurt is prepared using whole or low-fat milk with added nonfat milk solids. The milk is heated to approximately 180°F (82°C) for 30 min, homogenized, cooled to 115°F (46°C), inoculated with an active culture, and packaged. Yogurt cultures are mixtures of *S. thermophilus* and *L. bulgaricus* in a 1:1 ratio. Balance of these organisms in the culture is important for production of a quality product. The product after inoculation is incubated until approximately 0.9% acidity has developed and then cooled.

In the United States, greater sales are realized in yogurts that contain added fruit than in the unflavored product. Three types are marketed: fruit mixed throughout, fruit on top, and fruit on bottom. Manufacturing procedures and costs demand an automated process where the product is continuously handled. After the proper acidity is developed during quiescent incubation, the product is agitated, fruit is added, the product is cooled to prevent further acid development, and finally it is pumped to a packaging machine to be filled mostly into small cups.

Yogurt is custardlike in consistency and is generally eaten with a spoon.

Other fermented milks. A few other types of fermented milks produced in other parts of the world are relatively obscure to people in the United States. Bulgarian buttermilk is made by the inoculation of whole milk with a culture of *L. bulgaricus*. It is incubated until approximately 0.9% acidity develops, and then packaged. This product is similar to yogurt except that it is fluid.

Kefir is a fermented milk native to the Caucasus Mountain area of southeastern Europe. Little is produced in the United States. During the fermentation process, lactic acid is produced by *Streptococcus* and *Lactobacillus* organisms, and alcohol (approximately 1%) is produced by lactose-fermenting yeasts. The latter collect in grains that are about the size of wheat and float on the surface.

Kumiss is another milk beverage in which both acid and alcohol are developed. Commonly, milk from mares is used, and the finished product resembles kefir except that no visible grains float on the surface. Kumiss is a mildly acid milk product with relatively little alcohol (0.5–1.5%) and is native to southern Russia.

Concentrated and Dried Milk Products

To reduce costs of transportation and handling, either part or all of the water is removed from milk. Moreover, the partly dehydrated milk can either be sterilized or dried to permit unrefrigerated storage for prolonged periods. Many different milk products are produced for these specific reasons. The composition of some of these is controlled by standards of identity and some by request of the commercial buyer.

In the United States nonfat dry milk is defined by Public Law 244, March 2, 1944, as amended by Public Law 646, July 2, 1956. Such milk is the product resulting from the removal of fat and water. In the dry form, nonfat dry milk can contain no more than 5% water and not over 1.5% fat.

Dry whole milk is the result of removing water from whole milk until the dry product contains no more than 4% water and not less than 26% fat.

Dry buttermilk results from the removal of moisture from fluid buttermilk drained from the churn during the manufacture of butter. It contains no more than 5% water and not less than 4.5% fat.

Evaporated milk is the product that results from removing moisture from whole milk so that the final product contains no less than 7.9% fat or 25.9% total milk solids. Emulsifying salts, to reduce the effect of heat on the protein during sterilization, and vitamin D are optional additives and cannot exceed 0.1% and 400 IU (reconstituted), respectively.

Baby formulations are produced to give a product, when reconstituted, that has the same composition as mother's milk. These formulations contain nonfat milk solids, salts, sugars, and vegetable fat. Baby formulations are marketed in either dry or condensed sterilized liquid form.

Sweetened condensed whole or skim milks are prepared to contain not less than 8.5% fat and 28% total milk solids or 24% total milk solids, respectively. Enough sugar is added to prevent spoilage, usually 42–45%.

Condensed milks and buttermilk are the result of removing some water from the fresh fluid product. The concentration of milk solids is generally requested by the food manufacturer; however, if the concentration of skim milk exceeds 36% and this product is stored, lactose will crystallize.

Condensing or evaporating. Water is removed from the exposed milk product by condensing or evaporating. Initially the fluid product is given a heat treatment to destroy bacteria and enzymes and to impart particular properties to the finished product. For example, ice cream manufacturers want a low-heat product, whereas bakers need a high-heat product to give loaf volume and crust color to bread. Heat treatments vary from 150 to 210°F (66 to 99°C), for a few seconds to as much as 60 min. The preheated product is drawn into the vacuum pan from the hot well or surge tank and is boiled at reduced pressure. To increase the efficiency of this operation, two, three, or four units or effects are connected together. The vapor energy from one pan is used to cause boiling in the next. The vacuum applied to each effect is increased as the final effect is reached.

Sterilization. Evaporated milk is standardized to meet federal specifications, canned, sealed, and sterilized. Temperatures and times of 238–245°F (114–118°C) for 15–17 min are used. To prevent coagulation of the protein during sterilization, the salt balance of the milk must be examined. If the salt balance or concentration of calcium and magnesium is not in proportion to the phosphates and citrates content, then the defect occurs. However, coagulation can be prevented by the addition of phosphate or citrate salts not to exceed 0.1% in the finished product.

Spray drying. Most condensed milks, baby formulations, whole milk, and buttermilk are dried. Condensed fluid product is pumped directly while still hot from the vacuum pan into a surge tank and then to the spray dryer. This piece of equipment can either be in the form of a horizontal stainless steel box or a vertical cylinder that is cone-shaped on the bottom. The fluid product is atomized to facilitate drying. Atomization is accomplished either by centrifugal force where the product is discharged from a rapidly turning horizontal disk, or by forcing the product through a nozzle with a minute orifice. Drying occurs almost instantaneously in the heated atmosphere in the dryer.

The dried product is collected from the drier by several means. One method involves the use of a cyclone, where the dried particles are removed by centrifugal force from a swirling air stream. Other methods used in horizontal dryers involve mechanical sweeps to push the dried powder across the bottom of the dryer to an auger for discharge and a series of cloth filter bags in a separate housing.

The Department of Agriculture has assigned grades to spray-dried nonfat dry milk on the basis of results from flavor, physical, and laboratory examinations. These are U.S. Extra, U.S. Standard, and U.S. Grade Not Assignable. A product in the latter classification is used as animal feed. Further grading is assigned to denote heat treatments that the product receives during preparation. These are U.S. Low Heat, U.S. High Heat, and U.S. Medium Heat, and these reflect the amount of whey protein change or denaturation.

Roller drying. Little milk for human use is roller-dried. The process involves feeding the concentrated milk between two heated and narrowly spaced stainless steel rollers. A thin film of milk adheres and is dried as these rollers turn. A scraper or doctor blade removes the dried product. The chief disadvantage to this process is the excess exposure to heat incurred by the product. The finished product has a characteristic flavor and limited reconstituting characteristics.

Instantizing. To render spray-dried nonfat dry milk more easily reconstituted, the product is exposed to another heat treatment. Unfortunately, the process decreases the solubility of milk proteins. Instantizing creates an agglomerated powder; therefore, the particles are heavier and sink and wet faster than a noninstantized powder. Nonfat dry milk directly from the spray dryer is fed into a hopper and metered into a steam current to cause the product to lump together or agglomerate. This wet (10–15% water) product is redried, and the agglomerates are sieved to the required size. In some commercial operations, sizing rollers are used. Any fine particles are recirculated to the inlet hopper. Most of the nonfat dry milk for fluid consumption is sold in this form. *See* FOOD MANUFACTURING; ICE CREAM. Robert L. Bradley, Jr.

Bibliography. A. W. Farrall, *Engineering for Dairy and Food Products*, 2d ed., 1980; J. L. Henderson, *The Fluid Milk Industry*, 3d ed., 1971; Y. H. Hui (ed.), *Dairy Science and Technology*, 3 vols., 1993; G. H. Richardson (ed.), *Standard Methods for the Examination of Dairy Products*, 16th ed., 1992; N. P. Wong, *Fundamentals of Dairy Chemistry*, 3d ed., 1988; U.S. Department of Agriculture, *Federal and State Standards for the Composition of Milk Products*, Handb. 51, revised 1977.

Milky Way Galaxy

The large disk-shaped aggregation of stars, gas, and dust in which the solar system is located. The term Milky Way is used to refer to the diffuse band of light visible in the night sky emanating from the Milky Way Galaxy. Although the two terms are frequently used interchangeably, Milky Way Galaxy, or simply the Galaxy, refers to the physical object rather than its appearance in the night sky, while Milky Way is used to refer to either.

Fig. 1. Panoramic photograph of the entire Milky Way. The dark patches are nearby clouds containing interstellar dust that block out the background starlight. The photograph conveys the overall impression of the flattened distribution of stars that make up the disk. The direction of the galactic center is in the center but it is hidden by foreground dust. The galactic bulge can be seen as the thickening of the disk in the center. (*European Southern Observatory*)

Appearance. The Milky Way is visible in the night sky to the unaided eye as a broad diffuse band of light stretching from horizon to horizon when viewed from locations away from bright city lights (**Fig. 1**). Nearly all of the visible light is due to individual stars, which in many directions are too numerous to be resolved without a telescope. The patchy appearance of the Milky Way is due to collections of microscopic dust particles which block the light of more distant stars. All of the stars seen by the unaided eye are part of the Milky Way Galaxy and lie relatively close to the Sun. The overall appearance is due to the Sun's location near the midplane of the galactic disk; the diffuse band of light is seen toward directions close to the midplane where there are many more stars along the line of sight than in directions away from the plane.

Structure and contents. The Milky Way Galaxy contains about 2×10^{11} solar masses of visible matter. Roughly 96% is in the form of stars, and about 4% is in the form of interstellar gas. The gas both inside the stars and in the interstellar medium is primarily hydrogen (roughly 87–90% by number of atoms) and helium (about 10%) with a small admixture of all of the heavier atoms (0–3% depending on the location within the Galaxy). The mass of dust is about 1% of the interstellar gas mass and is an insignificant fraction of the total mass of the Galaxy. Its presence, however, limits the view from the Earth in the plane of the Galaxy to a small fraction of the Galaxy's diameter in most directions. *See* INTERSTELLAR MATTER.

The Milky Way Galaxy contains four major structural subdivisions: the nucleus, bulge, disk, and halo. The Sun is located in the disk about halfway between the center and the indistinct outer edge of the disk of stars. The currently accepted value of the distance of the Sun from the galactic center is 8.5 kiloparsecs, although some measurements suggest that the distance may be as small as 7 kpc. (A parsec is equal to 2.06×10^5 times the average distance between the Sun and the Earth, and is approximately equal to the average distance between stars in the solar neighborhood; it is also equal to 3.26 light-years.) **Figure 2** is an infrared image of the Milky Way made from the Sun's position in the disk, showing the bulge and disk of the Galaxy. The infrared image minimizes the obscuring effect of dust. **Figure 3** shows a spiral galaxy that is structurally similar to the Milky Way, giving an approximate idea of how the Milky Way would appear if viewed obliquely. *See* INFRARED ASTRONOMY; PARSEC.

Nucleus. The nucleus of the Milky Way is a region within a few tens of parsecs of the geometric center and is totally obscured at visible wavelengths. The nucleus is the source of very energetic activity detected by means of radio waves and infrared radiation.

At the galactic center, there is a very dense cluster of hot stars observed by means of its infrared radiation. In 1997, astronomers confirmed the existence of a black hole with a mass of about 2.5 million times the mass of the Sun at the position of an unresolved source of radio emission known as Sgr A* in the middle of the central star cluster. The black hole appears to be the dynamical center of the Milky Way, and evidence for its existence appears to be unequivocal. Many spiral galaxies show evidence of black holes at their centers. At a distance of about 1.5 pc from Sgr A* is a ring of gas consisting primarily of molecular hydrogen that surrounds the central star cluster. Within this ring is a three-armed "mini-spiral" of ionized gas that appears to be falling into the central cluster and possibly onto the black hole

Fig. 2. Infrared image of the Milky Way as seen from the location of the Sun, made from the 2MASS infrared sky survey. The image shows both the bulge and the disk of the Milky Way. Absorption by dust is minimized in the infrared, which accounts for the difference between this image and Fig. 1. However, the dust remains visible as the dark band across the disk and other dark patches. The Large and Small Magellanic Clouds are visible in the lower right. The faint vertical wisp coming down from the left part of the bulge is the Sagittarius Dwarf. (*2MASS*)

Fig. 3. NGC 4603, a galaxy with a gross morphology similar to the Milky Way, seen obliquely. The spiral arms and the bulge are visible in the center. The bulge appears to be elongated because the galaxy also contains a bar, an extended feature containing primarily old stars that rotate collectively about the center as if they were a solid body. (*NASA*)

(**Fig. 4**). Molecular hydrogen gas, which is characterized by low temperatures and relatively high densities, is found in great profusion in the inner few hundred parsecs of the Milky Way. Great arcs of gas resulting from the interaction of cosmic rays and magnetic fields have been mapped in the central tens of parsecs of the Milky Way, attesting to the energetic activity taking place there. *See* BLACK HOLE; COSMIC RAYS; RADIO ASTRONOMY.

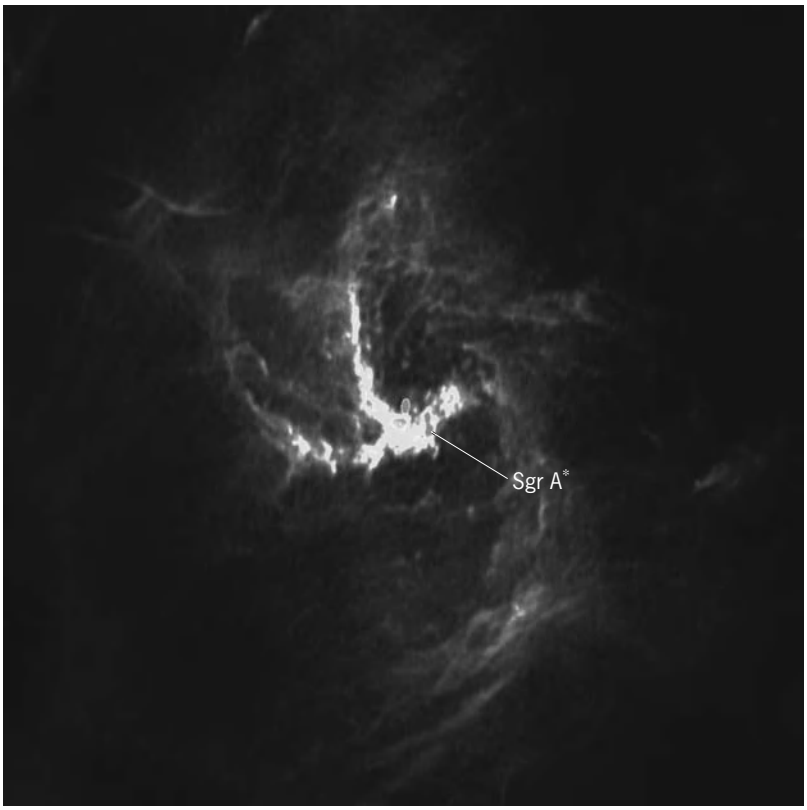

Fig. 4. Radio image of the ionized gas within about 1.5 parsecs of the center of the Galaxy. The image, made with the Very Large Array (VLA), shows what appears to be a three-armed spiral of gas. Other observations indicate that this gas may be falling into the central star cluster (not seen in this image). The position of Sgr A* is indicated by the elongated bright dot seen above the horizontal bar of emission. The dot is thought to be radio emission from hot gas falling into the black hole. (*NRAO, courtesy of N. Killeen and K.-Y. Lo*)

Bulge. The bulge is a spheroidal distribution of stars centered on the nucleus which extends to a distance of about 3 kpc from the center. It contains a relatively old population of stars, very nearly as old as the Milky Way itself. The bulge can be seen by the unaided eye as a thickening of the diffuse band of light that constitutes the Milky Way in the direction of the galactic center, toward the constellation Sagittarius. There is little gas and dust throughout most of the volume of the bulge. Direct imaging with infrared satellites has demonstrated that the bulge is actually an elongated barlike structure with a length about two to three times its width. The Milky Way is thus classified as a barred spiral galaxy, a classification that includes about half of all disk-shaped galaxies. *See* STELLAR POPULATION.

Disk. The disk is a thin distribution of stars and gas orbiting the nucleus of the Galaxy. The disk of stars begins near the end of the bar and can be identified to about 16 kpc from the center of the Galaxy; the disk of gas can be identified to about twice this distance, about 35 kpc from the center. The thicknesses of the disks are characterized by a scale height: the distance from the midplane at which the density of gas and stars falls by a factor of e (2.718...). In the solar vicinity, the scale height is different for each of the components of the disk, varying from about 75 pc for the molecular gas to about 350 pc for the lowest-mass stars. The faint, low-mass stars make up most of the mass of the disk. There is also a thick disk of stars and gas with a scale height of about 1.5 kpc for the stars and about 1 kpc for the atomic gas. The origin of these thick disks is not known. The thin disk of stars contains most of the mass and has a

thickness relative to its diameter similar to that of a commercial compact disk. At this scale, the bulge would have the size of a sausage about 2.5 cm (1 in.) in length superimposed on the nucleus.

The star and gas disks are both warped; the gas disk deviates from a true plane by about 3 kpc (about 10%) in its outermost parts. The disk also becomes thicker with increasing distance from the galactic center. The thickening occurs because there is less gravity to confine the gas and stars to the midplane. The reason for the warping is not yet understood, but warping is a common feature of spiral galaxies.

The disk is the location of the spiral arms that are characteristic of most disk-shaped galaxies, as well as most of the present-day star formation. Attempts to map the location of the spiral arms are hampered by the Sun's location in the disk, but in the parts of the Galaxy beyond the Sun's distance from the center, several long coherent spiral arms have been identified. The inner regions appear to be more chaotic, and no agreed-upon spiral arm structure has been identified. The spiral arms are where giant clouds of molecules are primarily found. These clouds are the most massive objects in the Milky Way, ranging up to about 5×10^6 solar masses; they are the sites of all present-day star formation. The Sun presumably once formed in a giant molecular cloud. *See* MOLECULAR CLOUD.

Halo. The halo is a rarefied spheroidal distribution of stars nearly devoid of the interstellar gas and dust that surrounds the disk. The stars found in the halo are the oldest stars in the Galaxy. The stars are found individually as "field" stars as well as in globular clusters: spherical clusters of up to about a million stars with very low abundances of elements heavier than helium. The extent of the halo is not well determined, but globular clusters with distances of about 40 kpc from the center have been identified. Dynamical evidence suggests that the halo contains nonluminous matter in some unknown form, commonly referred to as dark matter (see below). The dark matter contains most of the mass of the Galaxy, dominating even that in the form of stars. *See* STAR CLUSTERS.

Companion galaxies. The Milky Way has 11 companion galaxies clustered within about 250 kpc of its center. The largest and brightest of these, the Large and Small Magellanic Clouds, are seen as diffuse patches of light roughly 7° and 5° in diameter and separated from the plane of the Milky Way by about 30° and 40°, respectively. They are visible mainly from the Earth's southernmost latitudes. The distance to the Magellanic Clouds is about 55 kpc, not quite twice the diameter of the stars in the disk of the Galaxy. A third small galaxy, known as the Sagittarius Dwarf, was discovered in 1994 and is the nearest galaxy to the Milky Way. It went undiscovered for such a long time because it lies close to the obscuring dust in the midplane. The Sagittarius Dwarf is seen in the general direction of the center of the Galaxy, and its distance from the Sun is about 24 kpc, about three times the distance of the Sun from the center of the Galaxy.

All three galaxies are interacting with the Milky Way. The Magellanic Stream, a narrow band of gas drawn out of the Magellanic Clouds, which stretches over much of the sky, is produced either by the tidal forces of the Milky Way or by the ram pressure of the clouds going through a tenuous halo of galactic gas. The Sagittarius Dwarf shows evidence of being stretched apart by the strong tidal forces of the Milky Way, and is about to collide with the disk, a process that will take millions of years. *See* MAGELLANIC CLOUDS.

Dynamics and kinematics. The stars and gas in the disk of the Milky Way rotate around the galactic center in nearly circular orbits in the plane of the disk. The deviations from circularity are generally small, and give rise to oscillations of stellar motions perpendicular to the disk, epicyclic motions in the plane of the disk, and other, more complex motions. The overall rotation of the disk is differential rather than like a solid body in that the rotation period of the stars and the gas increases with increasing distance from the center. Large-scale deviations from circularity are associated with the spiral arms of the disk.

The barlike bulge produces large noncircular motions in the disk of gas contained within it, providing independent confirmation that the Milky Way contains a "bar" in its inner parts. A bar is a large collections of stars in the inner part of a galaxy in which the stars collectively rotate about the center as a solid body would, even though the individual stellar orbits are quite complex.

The orbits of the stars and clusters in the bulge and halo are not confined to the galactic plane and are, in general, fully three-dimensional. Although the motions of most stars are too small for the orbits to be determined from their changes in position alone, orbits can be inferred from the measured radial velocities of the stars (the velocities measured along the line of sight) and from their positions in the Galaxy. The orbits of stars close to the Sun can be observed directly from their apparent motions in the sky measured over many years. Motions of small numbers of stars as distant as the center of the Galaxy can be measured directly by special techniques.

In the outermost parts of the Milky Way, one expects that the orbital velocities of stars and gas around the galactic center should decrease in a well-determined way, as they do, for example, in the solar system. Measurements show, however, that in the outermost regions of the disk the velocities of the gas and stars do not decrease with increasing distance from the galactic center (**Fig. 5**). These measurements suggest that unseen matter is pulling on the visible matter in the disk and producing the anomalously large velocities. The constancy of the rotation velocity of the Milky Way is generally thought to be the most compelling evidence for large quantities of material in the Milky Way in a form other than stars, gas, and dust. This dark matter is common to most galaxies for which rotation measurements are possible.

Dark matter. The dark matter in the Milky Way constitutes by various estimates $2^1/_2$ to 10 times the total

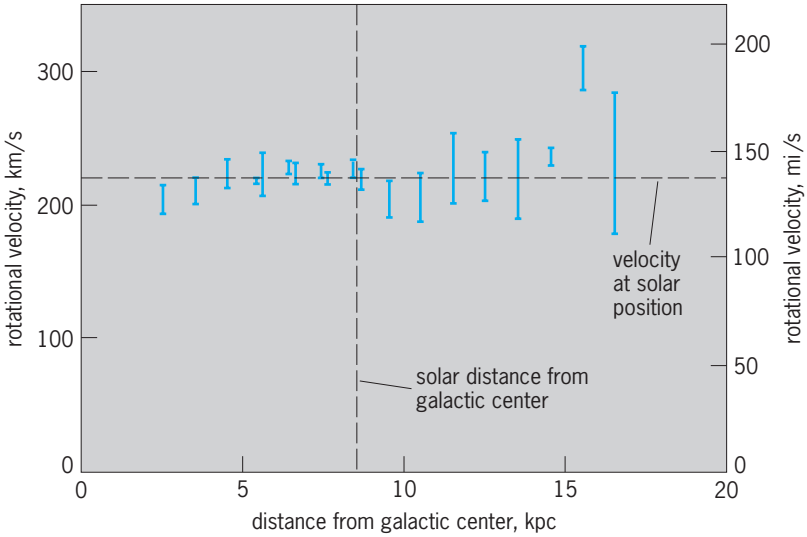

Fig. 5. Plot of the measurement of the rotational velocity of gas and stars about the center of the Milky Way Galaxy as a function of distance from the galactic center, showing that the rotational velocities are essentially constant. Beyond the solar distance, one would expect that the rotational velocity would decrease markedly to a value about two-thirds the value measured at the Sun's distance if only visible matter contributed to the gravity of the Milky Way. (*Courtesy of M. Fich, L. Blitz, and A. A. Stark*)

amount of known matter in the Milky Way, and is consequently the dominant component of mass in the Galaxy. Because the kinematics of most galaxies indicate that they too contain dark matter as their largest mass constituent, the dark matter appears to be the dominant form of matter in the universe. The composition of the dark matter is currently unknown and is one of the major unsolved problems in astronomy.

In the Milky Way, various forms for the dark matter have been ruled out. It cannot be in the form of ordinary stars or in remnants such as white dwarfs, neutron stars, or black holes that are the end products of ordinary stellar evolution. It cannot be in the form of gas in any form, nor can it be in the form of small, solid dust particles. Although a small fraction of the dark matter may reside in the disk, most of it appears to reside in the halo. *See* NEUTRON STAR; WHITE DWARF STAR.

Two possibilities that have not yet been ruled out are that the dark matter consists of small planet-sized bodies that are insufficiently luminous to be detected with present instruments, and that it consists of primordial black holes in the halo. Still other possibilities include weakly interacting massive particles (WIMPs) that are predicted by various elementary particle theories. Observations of the supernova 1987A have apparently ruled out the possibility that neutrinos (massless or very low mass neutral particles) constitute most of the dark matter in the universe. An alternative theory to newtonian gravitation has also been proposed to explain the rotation curve of the Milky Way and other galaxies, but this is generally considered to be a particularly radical explanation. *See* COSMOLOGY; GRAVITATION; NEUTRINO; SUPERNOVA; WEAKLY INTERACTING MASSIVE PARTICLE (WIMP).

Location. The Milky Way is part of a small group of galaxies known as the Local Group. The Local Group also contains two large spiral galaxies, the great nebula in Andromeda (M31) and M33, and 40 small irregular, spheroidal and elliptical galaxies. Some of these small galaxies have been discovered only in the last few years, and it is likely that the census of galaxies in the Local Group is not yet complete. About 95% of the mass in the Local Group is associated with either the Milky Way or M31. The Local Group is itself part of a large supercluster of galaxies known as the Virgo supercluster containing about 1000 known galaxies. The Local Group is an outlying collection of galaxies in the Virgo supercluster and is about 15 Mpc from its center. The Virgo supercluster is one of many such clusters in the universe and occupies no special location within it. *See* ANDROMEDA GALAXY; GALAXY, EXTERNAL; LOCAL GROUP; UNIVERSE.

Formation and evolution. Inferences about the formation and evolution of the Milky Way can be drawn from a large variety of sources, including the theory and observation of the kinematics and dynamics of the gas and stars in the Milky Way, the chemical abundances of the stars, their locations in the Galaxy, and stellar evolution theory. Current evidence suggests that the ages of the oldest stars in the Milky Way are within about 10% of the age of the universe as a whole; thus parts of the Milky Way must have formed early in the history of the universe, about 12 billion years ago. *See* STELLAR EVOLUTION.

There is increasing evidence that the Milky Way formed as a result of the coalescence of small galaxies and protogalaxies, objects with the masses of small dwarf galaxies that are thought to have been among the first objects to form in the universe. The coalescence would have proceeded rapidly at first and then more irregularly as relatively large pieces merged to form the Milky Way. According to this picture, the formation of the Milky Way is not yet complete, and both small galaxies (such as the Magellanic Clouds and the Sagittarius Dwarf) and starless clouds of gas are continuing to rain in on the Milky Way. Direct evidence for this picture comes from the orbit of the Sagittarius Dwarf, which shows it to be colliding with the Milky Way, and from the presence of clouds of hydrogen gas with high velocities and low abundances of heavy elements, which appear to be accreting from intergalactic space.

When the Galaxy was about two-thirds its present age, 4.6 billion years ago, the Sun and the planets formed from the interstellar medium in the disk, the Sun being a rather unremarkable star in an unremarkable location. The part of the disk that has not yet been used up in the process of star formation continues to form stars. *See* SOLAR SYSTEM; SUN.

The Milky Way and M31 are approaching each other at a velocity of 125 km/s (78 mi/s). At that velocity, the two galaxies will collide in about 5 billion years if the transverse velocity of M31 is not too large. The collision will not be violent; the distance between the stars is too vast for more than a few stars to actually collide. The likely outcome is a merged system that resembles an elliptical galaxy. Leo Blitz

Bibliography. J. Binney and M. Merrifield, *Galactic Astronomy*, Princeton University Press, 1998; L. Blitz et al., The centre of the Milky Way, *Nature*, 361:417–424, 1993; L. Blitz and P. Teuben (eds.), *Unsolved Problems of the Milky Way*, Kluwer Academic Publishers, 1996; B. J. Bok and P. F. Bok, *The Milky Way*, 5th ed., Harvard University Press, 1981; A. Eckart and R. Genzel, Stellar proper motions in the central 0.1 pc of the Galaxy, *Mon. Not. Roy. Astron. Soc.*, 284:576–598, 1997; B. K. Gibson, T. S. Axelrod, and M. E. Putman (eds.), *The Galactic Halo*, Astronomical Society of the Pacific, 1999; I. King, G. Gilmore, and P. C. van der Kruit, *The Milky Way as a Galaxy*, University Science Books, 1994.

Milleporina

An order of the class Hydrozoa of the phylum Cnidaria. These are the "stinging corals" of the shallow tropical seas. Their structure is similar to that of hydroids except for the addition of a calcareous exoskeleton (**illus.** *a*). Because of this skeleton, they resemble true corals, which belong, however, to a different class of coelenterates, the Anthozoa. The skeleton is covered by a thin layer of tissue, is penetrated by interconnecting tubes, and is perforated by tiny holes through which the bodies of the "coral animals," or polyps, are extended (illus. *b*).

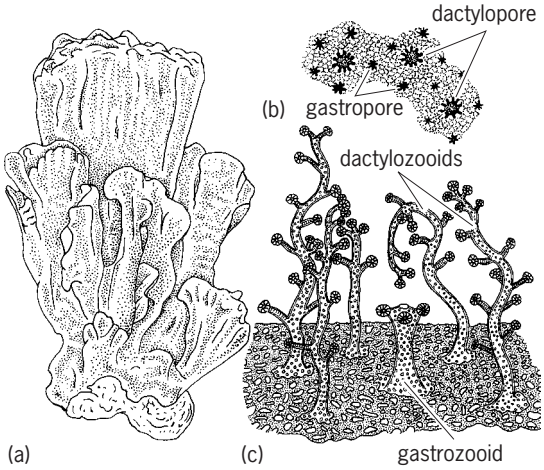

Milleporina. (*a*) Piece of dry *Millepora*, showing typical flabellate shape. (*b*) Same, magnified, showing pores. (*c*) Polyps of *Millepora*. (*After L. H. Hyman, The Invertebrates, vol. 1, McGraw-Hill, 1940*)

The polyps are of two types (illus. *c*): nutritive gastrozooids with tentacles and mouth; and protective polyps, which are long and armed with stinging cells but have no mouth. Their sting is very painful to humans, unlike the relatively harmless sting of most true corals. Millepores produce medusae, or jellyfish, in which sex cells develop.

Some authorities combine the Milleporina with the Stylasterina in a single order, the Hydrocorallina. *See* HYDROZOA. Sears Crowell

Millerite

A mineral having composition NiS and crystallizing in the hexagonal system. Millerite usually occurs in hairlike tufts and radiating groups of slender to capillary crystals (see **illus.**). There is rhombohedral cleavage,

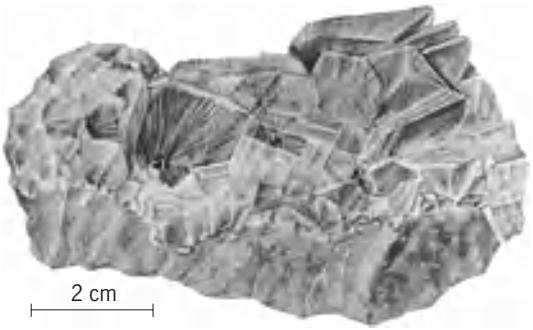

Millerite crystals on calcite from Keokuk, Iowa. (*Specimen from Department of Geology, Bryn Mawr College*)

but it is difficult to observe on the hairlike crystals. The hardness is 3–3.5 (Mohs scale) and the specific gravity is 5.5. The luster is metallic and the color pale brass yellow. Millerite forms at low temperatures, often in cavities and as an alteration of other nickel minerals. It is found in many localities in Europe, notably in Germany and the Czech Republic. In the United States it is found with pyrrhotite at the Gap Mine, Lancaster County, Pennsylvania; with hematite at Antwerp, New York; and in geodes in limestone at Keokuk, Iowa. In Canada large cleavable masses are mined as a nickel ore in Lamotte Township, Quebec. *See* NICKEL; PYRRHOTITE. Cornelius S. Hurlbut, Jr.

Millet

A common name applied to at least five related members of the grass family grown for their edible seeds: foxtail millet (*Setaria italica*), proso millet (*Panicum miliaceum*), pearl or cat-tail millet (*Pennisetum typhoideum*; formerly *P. glaucum*), Japanese barnyard millet (*Echinochloa frumentacea*), raggee or finger millet (*Eleusine coracana*), and koda millet (*Paspalum scrobiculatum*). Millets have been used since prehistoric times as food crops, primarily in regions where the warm growing season is short (60 to 120 days), or in dry regions where rainfall periodicity provides a short period when soil moisture permits growth and ripening of a short-season crop. Under these climatic conditions, one or more millets are grown in such diverse geographic regions as Russia, China, India, Africa, and Latin America.

Use. As a crop for human food, pearl millet is grown widely in the topics and subtropics in regions of limited rainfall where there is a growing season of 90 to 120 days. In Africa and Asia some 1.8×10^7 tons (2×10^7 metric tons) of grain are produced yearly in 40 different countries. This millet is grown

where the limited rainfall or length of growing season is inadequate for grain sorghum or maize. Pearl millet grain yields average 520–620 lb/acre (600–700 kg/hectare), but much higher yields are possible with favorable rainfall and soil fertility. The naked seeds are yellowish to whitish in color and about the size of wheat grain. The dried grain is usually pulverized to make a meal or flour and then cooked in soups, in porridge, or as cakes.

In some Asian countries, finger millet is grown as a short-season food grain crop and is used in similar manner as pearl millet. The other millets also are grown as food grains in dry regions, mostly in subtemperate zones or where there is a very short growing season of 45 to 90 days.

Millet grains are high (55–65%) in starchy components and thus serve as energy foods. The protein content and quality vary greatly among the various types, but they are generally low in lysine amino acid and must be used along with animal products or fish to balance the protein diet. Where widely grown, the millets are major food crops.

In North America millets are used to some extent for forage (foxtail millet in northern regions and pearl millet in warm regions). Proso millet is grown for grain, mostly for livestock feed, in the northern Plains region. The feeding value is about equivalent to corn. There is also an increasing demand for proso and foxtail millet for use as birdseed. *See* GRASS CROPS.

Howard B. Sprague

Diseases. In the United States four general types of disease occur on the millets. These are foliage diseases, head mold, smut, and seedling and root diseases.

Foliage diseases such as Helminthosporium and Cercospora leaf spot diseases of pearl millet usually occur late in the season and, as a rule, damage is slight. Lesions caused by *Helminthosporium* vary in size from brown flecks to large oval or rectangular spots (**Fig. 1***a*), in contrast to the small dark brown, gray, or tan-centered spots caused by *Cercospora* (Fig. 1*b*). Head molds caused by several fungi may seriously affect seed production, particularly in moist weather (**Fig. 2**). Individual seeds or entire heads may be covered with wooly mycelial mats which are black, orange, or white, depending on the specific fungi present. Smut diseases, common in other countries, also affect the seed in the heads of foxtail and proso millet in the United States. These smuts may be controlled by seed treatment. Seedling blights and root rots, caused by a number of soil-inhabiting fungi, may reduce stands under unusually moist conditions. *See* PLANT PATHOLOGY.

Herman A. Rodenhiser

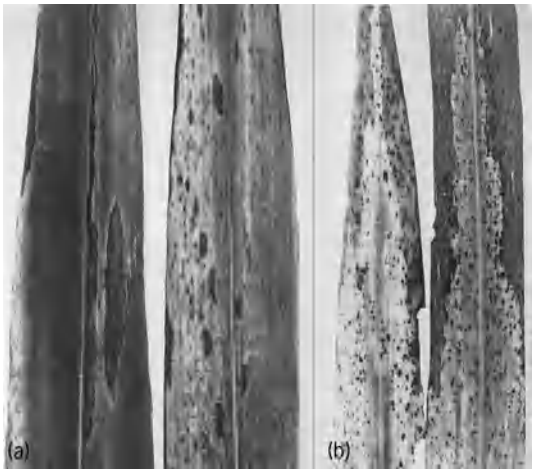

Fig. 1. Leaf diseases of millet. (*a*) Helminthosporium leaf spot of pearl millet (*Helminthosporium stenospilum* or *H. sacchari*). (*b*) Cercospora leaf spot of pearl millet (*Cercospora penniseti*). (*From E. S. Luttrell et al., Diseases of pearl millet in Georgia, Plant Dis. Rep., 38(7):508, 1954*)

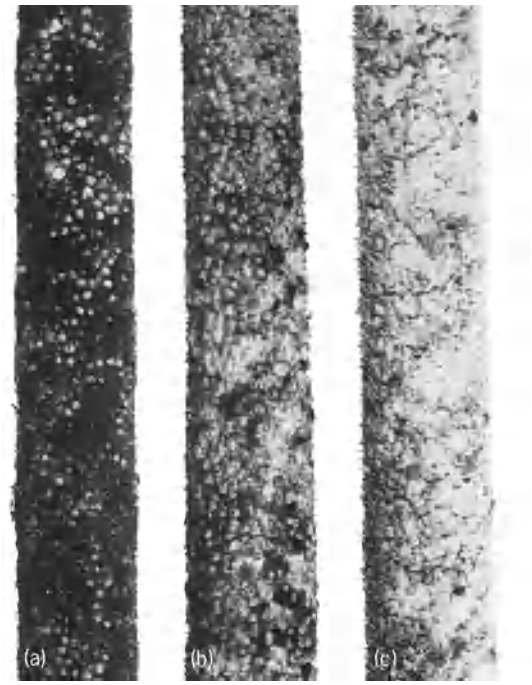

Fig. 2. Head molds of pearl millet. (*a*) Black mold, which is caused by *Helminthosporium* and *Curvularia*; (*b*) orange mold, which is caused by *Fusarium*; and (*c*) white mold, which is caused by *Oidium*.

Mineral

A naturally occurring solid phase with a restricted chemical composition and a definite atomic arrangement. This is really an idealized definition insofar as some minerals (especially those that form at low temperature) have poorly defined atomic structures. A few solid geological materials (such as coal and obsidian) are not minerals because they have no ordered crystal structure and a completely variable composition. Rocks are usually made up of one or more minerals. In a thermodynamic sense, a mineral is a separate phase that is made up of one or more components. For example, the mineral olivine [$(Fe,Mg)_2SiO_4$] can be viewed as being made up of fayalite (Fe_2SiO_4) and forsterite (Mg_2SiO_4) in a solid solution.

Classification and systematics. New minerals are still being discovered. At present, the International Mineralogical Association recognizes over 4000

Examples of important minerals

Silicates	
Orthosilicates	
Olivine $(Fe,Mg)_2SiO_4$	Mafic to ultramafic igneous rocks; a major phase in the upper mantle
Garnet $(Ca,Mg,Fe)_3(Al,Fe)_2(SiO4)_3$	Metamorphic rocks, ultramafic rocks
Kyanite, andalusite, sillimanite $Al_2O(SiO_4)$	Metamorphosed shales: schists and gneiss
Single-chain silicates: pyroxenes	
Orthopyroxenes $(Mg,Fe)_2SiO_3$	Mafic to ultramafic igneous rocks
Clinopyroxenes $(Ca,Fe,Mg)(Mg,Fe)SiO_3$	Mafic to ultramafic igneous rocks
Double-chain silicates: amphiboles	
Hornblende $(Ca,Na)_{2-3}(Mg,Fe,Al)_5Si_6(Si,Al)_2O_{22}(OH)_2$	Metamorphosed mafic rocks, granodiorites
Sheet silicates	
Muscovite $K_2Al_4(Al_2Si_6)O_{20}(OH)_4$	Metamorphic rocks, some granites
Biotite $K_2(Mg,Fe)_6(Al_2Si_6)O_{20}(OH)_4$	Granites, metamorphic rocks
Montmorillonite $(Ca,Mg,Na)_{0.5-0.7}$, $(Mg,Al,Fe)_4(Al_2Si)O_{20}(OH)_4$	Clay mineral: altered volcanic ash, mudstone, soil
Kaolinite $Al_2Si_2O_5(OH)_4$	Clay mineral: altered granites, soil
Framework silicates	
Quartz SiO_2	Major phase in most felsic rocks, sandstone, soil
K-feldspar $KAlSi_3O_8$	Major mineral in the Earth's crust: occurs in granites and metamorphosed sediments
Plagioclase feldspar $(Na,Ca)(Al,Si)AlSi_2O_8$	Most common mineral in the crust, nearly all igneous and metamorphic rocks
Nonsilicates	
Oxides and hydroxides	
Magnetite Fe_3O_4	Minor phase in most igneous and metamorphic rocks
Hematite Fe_2O_3	Soils and sediments, banded-iron formations
Goethite FeOOH	Soils and sediments
Sulfides	
Pyrite FeS_2	Hydrothermal deposits, anoxic sediments
Chalcopyrite $CuFeS_2$	Hydrothermal deposits
Sphalerite ZnS	Hydrothermal deposits
Galena PbS	Hydrothermal deposits
Carbonates	
Calcite $CaCO_3$	Sedimentary rocks, limestone
Aragonite $CaCO_3$	Biogenic shells
Dolomite $CaMg(CO_3)_2$	Sedimentary rocks, hydrothermal deposits
Sulfates	
Gypsum $CaSO_4{\cdot}2H_2O$	Evaporite deposits
Barite $BaSO_4$	Evaporites, hydrothermals deposits
Phosphates	
Apatite $Ca_5(PO_4)_3(F,Cl,OH)$	Minor mineral in igneous, sedimentary, and metamorphic rocks: biogenic (bone and teeth).
Halides	
Halite NaCl	Evaporite deposits
Fluorite CaF_2	Hydrothermal veins
Tungstates	
Scheelite $CaWO_4$	Granite pegmatites, high-temperature hydrothermal deposits
Native elements	
Diamond C	Kimberlites
Gold Au	Hydrothermal and placer deposits

known minerals. Most of these are very rare, and only a few minerals are common enough to be geologically significant. As developed from the scheme of J. D. Dana (1848), minerals are viewed as nominally ionic compounds (salts) and are classified by the nature of the anion (see **table**). The most abundant elements in the Earth's crust are oxygen (O), silicon (Si), and aluminum (Al). These elements combine to form a variety of anionic units based on SiO_4 and AlO_4 tetrahedra. Such aluminosilicate minerals are, by far, the most volumetrically important in the Earth's crust. The silicate minerals are classified by the degree of polymerization of the $(Si,Al)O_4$ tetrahedra. The SiO_4^{4-} tetrahedra can polymerize by sharing oxygen atoms to form $[Si_2O_7^{6-}]$ dimers (sorosilicates), $[(Si_2O_6)^{4-}]_n$ single chains, $[(Si_8O_{22}(OH)_2^{14-}]_n$ double chains, $[(Al,Si)_4O_{10}(OH)_2^{6-}]_n$ sheets, and three-dimensional $[Si_{2-n}Al_nO_4^{n-}]$ frameworks. *See* SILICATE MINERALS.

There are many nonsilicate minerals based on anions such as sulfide, oxide, hydroxide, sulfate, and carbonate, but by volume they are not as significant as silicates in the Earth's crust. Nearly all important ore minerals are nonsilicates. Moreover, the second most abundant mineral in the Earth's lower mantle is the oxide ferropericlase [(Mg,Fe)O].

Thermodynamics and phase transformations. The variety of minerals found in the Earth reflects not just the chemical heterogeneity of crustal rocks but also the variations in pressure and temperature. Minerals formed near the Earth's surface (for example, in sedimentary rocks) will transform to other phases when subjected to the high pressures and temperatures accompanying the collision of crustal terrains by plate tectonics. With increasing pressure, a mineral will tend to transform to a denser phase that is characterized by an atomic structure with higher coordination numbers and increased edge sharing

of coordination polyhedra. With increasing temperature, a mineral will transform to another phase with a structure that has greater entropy. This can result from increased disorder of cations over the crystallographic sites or by adopting a structure with higher symmetry. Consequently, many minerals occur in several different polymorphs that form at different pressures and temperatures. For example, kyanite, andalusite, and sillimanite are all polymorphs of Al_2SiO_5. Phase transformations of minerals can be sluggish. Consequently, the high pressure–high temperature phases that formed deep in the crust or mantle are usually preserved when rocks are brought to the surface during the uplift or erosion of continental crust. The phase transformations that minerals in the Earth's crust undergo are invaluable for the geologist since they provide clues about the temperature and pressure history of a rock. The structural layers of the Earth's deep interior (upper mantle, transition zone, and lower mantle) appear to be defined, at least in part, by pressure-induced phase transformations of the major minerals. The mineral olivine [$(Mg,Fe)_2SiO_4$] transforms to the polymorphs wadsleyite and ringwoodite at high pressure (13.4–23.8 GPa). Because olivine is the dominant mineral in the upper mantle, these phase transformations result in seismic discontinuities at about 400 and 600 km (250 and 370 mi) depth and separate the upper mantle from the transition zone. A further transformation of $(Mg,Fe)_2SiO_4$ into ferropericlase [(Mg,Fe)O] and silicate perovskite [$(Mg,Fe)SiO_3$] occurs at about 660 km (410 mi) depth and defines the onset of the lower mantle. *See* ANDALUSITE; EARTH INTERIOR; KYANITE; OLIVINE; PHASE TRANSITIONS; POLYMORPHISM (CRYSTALLOGRAPHY); SILLIMANITE.

Minerals also react with each other, and at chemical equilibrium the number of minerals that can coexist in a rock is limited by the phase rule. The phase rule states that the number of degrees of freedom (such as temperature and pressure) with which an assemblage of minerals can coexist is given by $f = c - p + 2$, where c is the number of chemical components and p is the number of phases (each mineral being a separate phase). If a mineral assemblage is to be stable over a range of temperature and pressure, $f = 2$, so that $c = p$. Hence, the number of minerals in a rock is limited by the number of independent chemical components. Very rarely, rocks will be found where $p > c$ (or $f < 2$). If the mineral assemblages in such rocks are really at chemical equilibrium, the rock must have formed along the pressure-temperature boundary that separates one phase assemblage from another. For a metamorphic geologist, a map of where such rocks occur can be used to reconstruct pressure-temperature regimes during mountain-building episodes.

Occurrence and formation. Igneous rocks are those that form from a molten state. Partial melting of upper mantle rocks (peridotite) will yield a melt with a basaltic composition. As this melt cools, minerals such as olivine [$(Mg,Fe)_2SiO_4$] and plagioclase feldspar [$(Ca,Na)(AlSi)AlSi_2O_8$] form. If the crystals that form can be separated from the melt, the melt composition will evolve to become more silica rich, and from it minerals such as amphibole, mica, and potassium feldspar may form. The compositional evolution of igneous melts was noted early on in Bowen's reaction series. An interesting system often develops as felsic (granitic) magmas evolve. As the magma solidifies, a fluid-rich phase will remain in which many exotic elements will be concentrated. The high fluid content favors the growth of large crystals. Gem minerals such as tourmaline can be found in pegmatites. *See* AMPHIBOLE; BASALT; FELDSPAR; GEM; IGNEOUS ROCKS; MAGMA; MICA; PEGMATITE; PERIDOTITE; PETROLOGY; TOURMALINE.

When rocks are exposed to the Earth's surface, minerals are broken down by reactions with water and carbon dioxide. These weathering reactions give rise to clay minerals such as montmorillonite and kaolinite. These minerals always occur as colloids and cannot form large crystals. Because of their small particle sizes, clay minerals have very high surface areas. This allows them to play an important role in controlling the concentrations of dissolved ions in seawater and soil solutions. *See* CLAY MINERALS; COLLOID; KAOLINITE; MONTMORILLONITE; SEDIMENTARY ROCKS; WEATHERING PROCESSES.

As crustal plates collide, rocks are slowly subjected to high pressures and temperatures. The original mineral assemblage in a rock will reequilibrate to form new minerals that are stable at the pressures and temperatures of the metamorphic environment. *See* METAMORPHIC ROCKS; METAMORPHISM; PLATE TECTONICS.

Some minerals, especially sulfides, are deposited by hydrothermal solutions. These are usually H_2S-NaCl-rich fluids that are able to dissolve metals such as lead (Pb), zinc (Zn), and copper (Cu) by the formation of metal-chloride complexes. Hydrothermal fluids may have originated by exsolution from magma, or they may result from seawater and ground water that has circulated in a geothermal regime. As these fluids react with minerals such as feldspar ($KAlSi_3O_8$) or calcite ($CaCO_3$) the pH rises and metal sulfides will precipitate from solution either in vein deposits or as disseminated sulfide crystals in the host rock. Subsequent alteration of primary sulfide minerals by oxidized acidic ground water will yield a complex variety of sulfate and oxide minerals. Percolating ground water may also dissolve disseminated sulfides in granitic host rocks and reprecipitate new sulfides at the water table boundary by a process known as supergene enrichment. *See* HYDROTHERMAL ORE DEPOSITS; ORE AND MINERAL DEPOSITS; SULFIDE AND ARSENIDE MINERALS.

Relatively soluble minerals such as halite (NaCl) and gypsum ($CaSO_4 \cdot 2H_2O$) can form in evaporite deposits. Marine evaporates result when seawater is trapped in isolated basins. Smaller evaporite deposits can form on continents when drainage basins feed into extremely arid environments. Evaporation of such waters yields unusual minerals such as nitrates and borates. *See* BORATE MINERALS; GYPSUM; HALITE; NITRATE MINERALS; SALINE EVAPORITES.

Some minerals are formed by biological processes. The most important is aragonite ($CaCO_3$), which is formed by mollusks and planktonic organisms such as foraminfera and coccolithophorids. Aragonite dissolves and reprecipitates as the more stable polymorph calcite ($CaCO_3$) which makes up the vast deposits of limestone, a major rock in the Earth's crust. Iron(III) and manganese(IV) oxides are often formed by Iron- and manganese-oxidizing bacteria. Other bacteria use these oxides, along with sulfate minerals, as electron acceptors during respiration, leading to the formation of various iron(II) oxide and sulfide minerals such as pyrite (FeS_2). The mineral apatite [$Ca_5(PO_4)_3(F,Cl,OH)$] is the inorganic phase in bone and teeth. *See* APATITE; ARAGONITE; CALCITE; LIMESTONE.

Properties. The physical and chemical properties of minerals can be used for their identification. A number of classical properties can be used by amateur collectors and geologists in the field, including color, streak, luster, habit, hardness, and specific gravity (density). The external crystallographic properties of a mineral reflect the symmetry of the internal structural arrangement of atoms. In the laboratory, the most diagnostic means of identifying minerals is by x-ray diffraction and chemical analysis. The optical properties of minerals are very useful for identifying silicate minerals. These properties can be observed when rocks are cut in thin section (35 micrometers) and viewed through transmitted polarized light in a petrographic microscope. *See* PETROGRAPHY; X-RAY DIFFRACTION.

The physical properties of minerals are of fundamental interest insofar as they determine the physical properties of the Earth. The elastic properties of minerals determine the velocities of seismic waves generated by earthquakes. Inversely, seismic velocities are used to infer the elastic properties and mineralogy of the Earth's interior. Most silicates are electrical insulators; however, most sulfides and transition-metal oxides are semiconductors. Their presence in geological bodies can be detected by measuring the conductivity of rocks in the field. This is an important tool for prospecting for sulfide ore bodies. A few minerals are ferrimagnetic. These minerals can record the direction of the Earth's magnetic field that was present when the minerals formed. Such paleomagnetic signatures were an essential line of evidence for continental drift and plate tectonics. *See* PALEOMAGNETISM; ROCK, ELECTRICAL PROPERTIES OF; ROCK MAGNETISM; SEISMOLOGY.

Economic aspects. A few minerals are used as gemstones, and this was the probable first use of minerals (as opposed to mineraloids such as flint) in the Paleolithic. Since the Bronze Age, minerals have provided the raw, natural material from which we extract useful metals. Most economically important metals (such as copper, lead, and zinc) occur as sulfide minerals. Iron ore minerals are oxides such as hematite (Fe_2O_3) and magnetite (Fe_3O_4). The primary ores of aluminum are oxides and oxide hydroxide minerals [for example, diaspore (AlOOH)]. *See* MAGNETITE; OXIDE AND HYDROXIDE MINERALS.

Industrial minerals are those that are economically useful in their natural state. Clay minerals such as kaolinite [$Al_2Si_2O_5(OH)_4$] are used to make china, pottery, bricks, and tile. Gypsum [$CaSO_4 \cdot 2H_2O$] is the main component of plaster and plasterboard. Corundum (Al_2O_3) is a common abrasive. Because of its fibrous habit, chrysotile [$Mg_3Si_2O_5(OH)_4$] is the main mineral of asbestos. The mineral talc [$Mg_3Si_4O_{10}(OH)_2$] is a lubricant owing to its layer structure held together by weak van der Waals bonds. *See* ASBESTOS; MINERALOGY; TALC. David M. Sherman

Bibliography. H. Blatt et al., *Petrology: Igneous, Sedimentary, and Metamorphic*, W. H. Freeman, 3d ed., 2005; C. Klein, *Manual of Mineral Science*, 22d ed., Wiley, 2001; R. V. Gaines et al., *Dana's New Mineralogy: The System of Mineralogy of James Dwight Dana and Edward Salisbury Dana*, Wiley-Interscience, 8th ed., 1997; H. Paquet et al. (eds.), *Soils and Sediments: Mineralogy and Geochemistry*, Springer, 1998.

Mineral resources

Mineral deposits, including ore bodies and potential ore, that are presently recoverable and may be so in the future. Much of what is consumed in modern society has to be obtained from mineral extraction. Countless items, whether coffee cups, camera film, electronic equipment, glasses, motor vehicles, or jewelry, contain minerals that must be extracted from the earth.

Types. Mineral resources generally are categorized into energy and nonfuel resources. Energy resources can be in solid (coal, lignite, uranium) or gaseous and liquid (natural gas, petroleum) form (see **table**). The many uses of these resources have evolved over time. Some resources such as gold have been used from the time of the earliest societies, while other minerals used in products from automobiles to missiles, although also available, have been needed only in recent times. *See* COAL MINING; ENERGY SOURCES.

The demand for minerals remains dependent on technological change and the perceived need for the final product. For example, a decline in the need for missiles would decrease the demand for certain minerals such as aluminum, as would the increased recycling of aluminum cans. Eliminating missiles or replacing aluminum cans with glass containers would have an even more drastic impact on the need to produce aluminum products. For mineral resources, the end-product demand rather than supply determines whether or not they are extracted from the earth.

Sources. Minerals may or may not be easy to find. Some parts of the world have been well explored for these resources, while others have not. Mineral resources can be located in politically stable or unstable countries. Minerals deemed vital for military production and defense purposes have been called strategic minerals. When the minerals are located largely in politically unstable or unfriendly countries, the United States has responded by setting up

Classification of selected metals and minerals and some some of their major uses

Materials	Uses
Fuels	
Bituminous and anthracite coal	Direct fuel, electricity, gas, chemicals
Lignite	Electricity, gas, chemicals
Petroleum	Gasoline, heating, chemicals plastics
Natural gas	Fuel, chemicals
Uranium	Nuclear power, explosives
	Metals
Ferrous metals	
Chromium	Alloys, stainless steels, refractories, chemicals
Cobalt	Alloys, permanent magnets, carbides
Columbium (niobium)	Alloys, stainless steels
Iron	Steels, cast iron
Manganese	Scavenger in steelmaking, batteries
Molybdenum	Alloys
Nickel	Alloys, stainless steels, coinage
Tungsten	Alloys
Vanadium	Alloys
Nonferrous metals	
Copper	Electrical conductors, coinage
Lead	Batteries, gasoline, construction
Tin	Tinplate, solder
Zinc	Galvanizing, die casting, chemicals
Light metals	
Aluminum	Transportation, rockets, building materials
Beryllium	Copper alloys, atomic energy field
Magnesium	Building materials, refractories
Titanium	Pigments, construction, acid-resistant plumbing
Zirconium	Alloys, chemicals, refractories
Precious metals	
Gold	Monetary, jewelry, dental, electronics
Platinum metals	Chemistry, catalysts, automotive
Silver	Photography, electronics, jewelry
Industrial	
Asbestos	Insulation, textiles
Boron	Glass, ceramics, propellants
Clays	Ceramics, filters, absorbents
Corundum	Abrasives
Feldspar	Ceramics, fluxes
Fluorspar	Fluxes, refrigerants, acid
Phosphates	Fertilizers, chemicals
Potassium salts	Fertilizers, chemicals
Salt	Chemicals, foods, glass, metallurgy
Sulfur	Fertilizers, acid, metallurgy, paper, foods, textiles

strategic stockpiles of minerals in case the supply becomes disrupted.

In the United States the use of stockpiles has ranged from storing petroleum in underground salt domes along the Louisiana and Texas coast to maintaining hoards of strategic alloys or even silver in government warehouses. The need for such stockpiles diminishes as the availability of the supply increases or, in the case of strategic minerals, as the threat decreases. For example, the collapse of the Soviet Union (and its diminished threat to the United States) decreases the need for such minerals and their stockpiles. The increase in the number of nations with more democratic governments from South Africa to Chile also provides more stability to the provision of minerals in a worldwide trading system.

Minerals are found by various exploration techniques. Traditionally these techniques have been based on geological theories, on field surveys, and on drilling in test sites. However, remote sensing and analysis of satellite imagery have become more important in prospecting for minerals. *See* GEOPHYSICAL EXPLORATION; PROSPECTING; REMOTE SENSING; SCIENTIFIC AND APPLICATIONS SATELLITES.

Mineral reserves. There have been predictions since before World War II that the United States will not have access to necessary mineral resources during major wars. During the 1960s, another major debate raged on the limits to growth and the depletion of world resources. During the 1970s, prices on a wide range of minerals rose rapidly.

The fear of insufficient resources keeps recurring, because there are not good estimates of how much of any mineral is in reserve throughout the world, or how and when the demand will increase. The uncertainty is over how much there is, and how much can be recovered given current costs and technology.

Estimates of resources include those proven to exist, those that can be recovered, and probable, possible, and potential reserves. The boundaries between these categories can be quite fluid. New discoveries or changes in world politics can dramatically change the availability of resources. New technologies can make feasible the extraction of minerals previously considered too costly to develop.

Development and governmental considerations. The minerals industry has an image and history of disturbing the landscape and creating environmental problems. The image of strip-mining large tracts of land to expose coal seams is not accurate, since most hard-rock mining is done below the surface and disturbs only a small portion of the landscape. Environmental problems associated with such mining have been contamination of ground water and waterways, the disposal of mining wastes, and air pollution from smelting operations. *See* AIR POLLUTION; WATER POLLUTION.

The burden for controlling the negative impacts from mining and mineral processing falls on local, state, and federal governments to regulate the environmental costs resulting from mineral development. In the United States this has been done at the federal level by the passage of environmental and reclamation laws restricting the amount and type of pollution and requiring the restoration of lands to near-original conditions. Certain public lands such as federal wilderness areas have been put off limits to future mineral exploration and development. *See* LAND RECLAMATION.

The danger is that countries with less developed governmental controls will permit environmental degradation in order to foster economic development. The potential for environmental destruction is heightened by the economic restructuring in the minerals industry in the more developed countries such as the United States.

Economic restructuring conservation. The minerals industry in the United States has been in decline and

a period of consolidation because of a decrease in the extraction of resources and in their processing. Much of this activity has moved to countries that are less developed, with lower labor costs, and less stringent or laxly enforced environmental laws and regulation. Instead of importing extracted mineral resources, processing them, and then exporting the products, the United States increasingly imports the processed minerals.

The demand for minerals will increase with world population growth and as other countries become more consumer oriented. The increasing pressures on the world environment will magnify in intensity. This will require more personal efforts and government policies to foster conservation of resources and environmentally sound extraction and processing of minerals. Without such efforts, mineral resource development ultimately will contribute to the decline of quality of life. *See* CONSERVATION OF RESOURCES.

Gundars Rudzitis

Bibliography. S. L. Cutter, H. L. Renwick, and W. H. Renwick, *Exploitation, Conservation, Preservation: A Geographic Perspective on Natural Resource Use*, 3d ed., 1998; H. E. Johansen, O. P. Matthews, and G. Rudzitis (eds.), *Mineral Resource Development: Geopolitics, Economics, and Policy*, 1987; B. O'hUallachain and R. A. Matthews, Economic restructuring in primary industries: Transaction costs and corporate vertical in the Arizona copper industry, 1980–1991, *Ann. Ass. Amer. Geog.*, 84(3):399–417, September 1994.

Mineralogy

The study of the crystalline phases (chemical compounds or pure elements) that make up the Earth and other rocky bodies in the solar system—that is, the terrestrial planets (Mercury, Venus, and Mars), meteorites, asteroids, and planetary satellites. In universities, nearly all mineralogists are found in earth science or geology departments.

History. Before the twentieth century, the science of mineralogy was concerned with identifying new minerals and determining their chemical compositions and physical properties. Investigations of the chemical compositions of minerals played a central role in the early development of chemistry and atomic theory. One of the most striking aspects of minerals is that they often occur as well-formed crystals. This motivated the hypothesis that the external symmetry reflected an internal symmetrical arrangement of atoms. The development of x-ray diffraction in 1915 enabled mineralogists to determine the internal crystal structures of minerals (that is, how the atoms are arranged). Once the chemical compositions and crystal structures of minerals were known, it became possible to systematically determine the thermodynamic properties of minerals and understand their stability as a function of pressure, temperature, and composition. This enabled geologists to understand the stabilities of mineral assemblages and the origins of the different kinds of igneous and metamorphic rocks. *See* CRYSTAL STRUCTURE; GEOLOGY; MINERAL; PETROLOGY; X-RAY DIFFRACTION.

By the 1960s, the crystal structures of most of the rock-forming minerals had been determined by x-ray diffraction. New techniques, however, were used in pursuit of a deeper understanding of the structural chemistry of minerals. Transmission electron microscopy, developed in 1961, has revealed the nanoscale (10–1000 Å) structures of minerals that result from twinning, dislocations, and phase transformations. Starting in the 1960s, a variety of spectroscopic methods (such as Mössbauer, nuclear magnetic resonance, infrared, and Raman) were used to understand how atoms are ordered among the crystallographic sites in minerals and to identify the nature of defects and impurities. *See* ELECTRON MICROSCOPE; SPECTROSCOPY.

With the advent of increasingly powerful computers, mineralogists began investigating the structures and properties of minerals using methods in computational chemistry. In the 1980s and 1990s, these calculations were largely based on classical models for interatomic interactions. In recent years, calculations of mineral stabilities and properties have been done using first-principles calculations based on either the Hartree-Fock approximation or density functional theory. *See* COMPUTATIONAL CHEMISTRY; QUANTUM CHEMISTRY.

Current research areas. Today, mineralogy is concerned with understanding the physics and chemistry of minerals insofar as they control processes in geochemistry and geophysics. Indeed, the boundaries between mineralogy, petrology, and geochemistry have largely disappeared. *See* GEOCHEMISTRY; GEOPHYSICS.

Phase equilibria and thermodynamic properties of minerals. The important minerals that occur deep in the Earth's crust have been discovered; however, their thermodynamic properties are often unknown. Laboratory experiments using high-pressure apparatus (the multianvil press and the diamond-anvil cell) can be used to determine the stable phases at the conditions of the Earth's deep interior. Laboratory experiments show that, in the extreme pressures of the Earth's lower mantle, the dominant mineral phases appear to be $(Mg,Fe)SiO_3$ and $CaSiO_3$ with a perovskite structure, along with $(Mg,Fe)O$ with the rock-salt structure. There are still many other minerals, composed of minor chemical components, which may also be stable in the Earth's mantle. Although these phases may not contribute much to the seismic profile of the Earth, they may be geochemically important as reservoirs of trace components used to understand the chemical evolution of the Earth. The crystal structures of minerals allow a variety of solid solutions and order–disorder transformations. As we understand the thermodynamics and kinetics of such processes, we can develop tools to infer the temperatures and pressures at which a rock formed. Order–disorder transformations can give clues on the rate at which rocks cooled after their formation. These geothermometers, geobarometers, and geospeedometers are invaluable tools for geologists

seeking to unravel the processes of mountain building and metamorphism. *See* GEOLOGIC THERMOMETRY; HIGH-PRESSURE MINERAL SYNTHESIS; METAMORPHISM; OROGENY.

Physical properties of minerals at high pressure and temperature: understanding the composition of the Earth from seismology. A major goal in the earth sciences is to determine the chemical composition of the Earth. An important constraint is provided by seismic data on the density and seismic velocities of the Earth as a function of depth. To derive compositional information from such data, we need to know what mineral phases are stable at high pressures and temperatures and their densities and elastic properties. The multianvil press and the diamond-anvil cells also allow in-situ measurements using x-ray diffraction of minerals confined at high pressure and heated to high temperature. From such experiments, we can determine structural changes and densities of minerals under the conditions of the deep Earth. Elastic properties can be measured using Brillouin scattering; because such experiments are difficult, theoretical calculations based on first principles (quantum mechanics) are also being used. *See* SEISMOLOGY.

Water in nominally anhydrous minerals at high pressure. Some high-pressure minerals such as wadsleyite, β-Mg_2SiO_4, have defect structures that allow some oxygen atoms to be replaced by OH. Many investigators are proposing that a significant amount of water is chemically stored in the Earth's mantle by this mechanism.

Colloid and interface chemistry of minerals: controls on aqueous and environmental geochemistry. As rocks weather, the minerals that formed at high temperature will dissolve and reprecipitate to form a variety of clay minerals along with several oxide and hydroxide minerals. These secondary minerals are important phases in soils and sediments. Because they occur as colloidal-sized (nanocrystalline) particles, they have a very high surface area. The electrostatic charge of the mineral surfaces promotes the sorption of dissolved ions from aqueous solutions. Consequently, clay minerals and authigenic oxide and hydroxide minerals exert major controls on the chemistry of natural waters. In particular, chemical processes at the mineral–water interface control the fate of toxic metals in soil and ground water and the availability of important micronutrients in the oceans. Understanding these processes at a molecular level is an active area of research. The interfacial properties of sulfide minerals are investigated for the development of flotation techniques to separate ore minerals from their host rocks. Reactions at the surfaces of sulfide minerals also are responsible for the formation of acid mine drainage. A variety of analytical techniques are being used to investigate the mineral–water interface. The availability of synchrotron radiation sources has enabled mineralogists to characterize the coordination chemistry at the mineral–water interface using a variety of spectroscopic techniques. Scanning tunneling microscopy and atomic force microscopy are unraveling the atomic structures of mineral surfaces. *See* AUTHIGENIC MINERALS; CLAY MINERALS; OXIDE AND HYDROXIDE MINERALS; SCANNING TUNNELING MICROSCOPE; SULFIDE AND ARSENID MINERALS.

Biomineralization. Current research is investigating the mechanisms by which organisms form inorganic minerals. Biomineralization processes may provide important clues to Earth's history: variations in trace metals and isotopic compositions are found in the shells of marine planktonic organisms, such as foraminifera. It is believed that these proxies may be giving a record of past seawater temperatures.

David M. Sherman

Bibliography. C. Klein, *Manual of Mineral Science*, 22d ed., Wiley, 2001; S. Mann, *Biomineralization*, Oxford University Press, 2002; W. D. Nesse, *Introduction to Mineralogy*, Oxford University Press, 1999.

Minimal principles

In the treatment of physical phenomena, it can sometimes be shown that, of all the processes or conditions which might occur, the ones actually occurring are those for which some characteristic physical quantity assumes a minimum value. These processes or conditions are known as minimal principles. The application of minimal principles provides a powerful method of attacking certain problems that would otherwise prove formidable if approached directly from first principles.

One simple minimal principle asserts that the state of stable equilibrium of any mechanical system is the state for which the potential energy is a minimum. Other general theorems of classical dynamics that are related to minimal principles are Hamilton's principle and the principle of least action. *See* HAMILTON'S PRINCIPLE; LEAST-ACTION PRINCIPLE.

Minimal principles are important in branches of physics other than mechanics. Fermat's principle in optics, for example, states that, of all possible paths of light transmission between two points, the actual path is the path for which the transmission time is a minimum. Minimal principles also find wide application in thermodynamics. *See* CHARGED PARTICLE OPTICS.

Dudley Williams

Bibliography. H. Goldstein, C. E. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; B. A. Kuperschmidt, *The Variational Principles of Dynamics*, 1992; C. Lanczos, *The Variational Principles of Mechanics*, 1970, reprint 1986.

Minimal surfaces

A branch of mathematics belonging to the calculus of variations, differential geometry, and geometric measure theory. A surface, interface, or membrane is called minimal when it has assumed a geometric configuration of least area among those configurations into which it can readily deform. Soap films spanning wire frames or compound soap bubbles

enclosing volumes of trapped air are common examples. *See* CALCULUS OF VARIATIONS; DIFFERENTIAL GEOMETRY; MEASURE THEORY.

Minimal surface equation. If a surface S in xyz space is represented as the graph of a function $z = f(x,y)$ over a region A in the xy plane, the area of S can be computed by formula (1), and the mean curvature of S at a point can be computed by formula (2).

$$\int\int_A \sqrt{1+f_x^2+f_y^2}\,dx\,dy \tag{1}$$

$$\frac{\left(1+f_y^2\right)f_{xx}-2f_xf_yf_{xy}+\left(1+f_x^2\right)f_{yy}}{2\left(1+f_x^2+f_y^2\right)^{3/2}} \tag{2}$$

Geometrically, the mean curvature of a surface at a point is the difference between the maximum upward curvature there and the maximum downward curvature; in particular, a surface of zero mean curvature has such principal curvatures equal and opposite and hence typically appears "saddle-shaped." It turns out that S is a minimal surface; that is, it cannot be perturbed to less area leaving its boundary fixed, provided the mean curvature is zero at each of its points; such a surface could occur, for example, as a soap film spanning a wire frame. The corresponding minimal surface equation is partial differential equation (3). If, alternatively, S were part of a

$$0 = \left(1+f_y^2\right)f_{xx}-2f_xf_yf_{xy}+\left(1+f_x^2\right)f_{yy} \tag{3}$$

soap bubble enclosing trapped air, the mean curvature of S would be proportional to the difference in air pressure between the two sides of S. In the calculus of variations, typically area-minimizing properties of minimal surfaces are emphasized. In differential geometry, minimal surfaces are defined as surfaces of zero mean curvature; surfaces of constant mean curvature are also extensively studied.

Historical development. The minimal surface equation was first written down by J. L. Lagrange in 1762. A list of problems about minimal surfaces posed by J. D. Gergonne in 1816 stimulated mathematical interest in the subject. By the second half of the nineteenth century, many noted mathematicians were studying minimal surfaces; new minimal surfaces were found, and it was discovered that any piece of minimal surface could be parametrized by utilizing a suitable pair of complex functions. In 1873 the physicist J. Plateau published a treatise on the geometry of the interfaces of liquids in equilibrium; in his honor a collection of mathematical problems about the existence of area-minimizing surfaces bear the name Plateau's problem.

The first "solution" to Plateau's problem was given in 1931 by J. Douglas, who showed that every closed curve in space is the boundary of some minimal surface (formally, a harmonic mapping from a disk in the xy plane into xyz space). A decade of renewed mathematical activity followed.

Alternative formulations of Plateau's problem were initiated in 1960, based on geometric measure theory. Several different kinds of mathematical minimal surfaces were invented, including ones resembling actual soap films and compound soap bubbles; these surfaces carry names such as integral current, flat chain, and varifold.

Soap bubble geometry. The possible configurations of soap films and soap bubbles are governed by elementary rules. Any such configuration consists of flat or smoothly curved pieces of surface (each of constant mean curvature) smoothly joined together. Furthermore, these surfaces meet in only two ways; either three surfaces meet along a smooth curve, or six surfaces (together with four curves) meet at a vertex. Finally, when surfaces meet along curves or when curves and surfaces meet at points, they do so at equal angles; in particular, when three surfaces meet along a curve, they do so at angles of 120°, and when four curves meet at a point, they do so at angles of close to 109°. These rules, first formulated by Plateau from his observations, are now known to be a necessary mathematical consequence of area minimization alone.

Minimal surface forms in nature. The two-dimensional surface is dominant in determining shape whenever the energy of a system is changed significantly by a displacement or a change in area of the surface. Such surfaces include the interfaces between crystals in a typical rock or metal, the film of soapy water between the air cells in a soap froth, the membrane separating the cells in living tissue, and the cracks separating basalt columns. Minimization of surface area plays a role in determining the shape of many living organisms. *See* FOAM.

Minimal surfaces in art and design. Soap bubbles have appeared in the works of a number of classical painters, and minimal surfaces have influenced modern artists and sculptors. The architect F. Otto has used soap film surfaces as models for a number of his dramatic roof structures, and the designer P. Pearce has constructed elaborate playground labyrinths inspired by periodic minimal surfaces.

Directions of research. Minimal surfaces are being studied extensively. They provide a tool for understanding the geometry of curved spaces and of curved space-time. The study of higher-dimensional minimal surfaces in higher-dimensional spaces is leading to a wealth of new geometric discoveries and constructions. Methods are also being developed and improved for the computation of minimal surfaces. There is increasing understanding of the relationships between minimal surfaces and their boundary curves; with probability one, for example, a randomly chosen closed curve bounds exactly one surface of least area. Finally, there is a continuing interest in understanding relationships between minimal surfaces and their generalizations and forms of the natural world.

Frederick J. Almgren, Jr.

Bibliography. U. Dierkes, *Minimal Surfaces I: Boundary Value Problems*, 1992; A. T. Fomenko and A. Tuzhilin, *Elements of the Geometry and Topology of Minimal Surfaces in Three-Dimensional Space*, 1991; J. C. Nitsche, *Introduction to Minimal Surfaces*, vol. 1, 1989.

Mining

The taking of minerals from the earth, including production from surface waters and from wells. Usually the oil and gas industries are regarded as separate from the mining industry. The term mining industry commonly includes such functions as exploration, mineral separation, hydrometallurgy, electrolytic reduction, and smelting and refining, even though these are not actually mining operations. *See* HYDROMETALLURGY; METALLURGY; ORE DRESSING.

Methods. Mining is broadly divided into three basic methods: opencast, underground, and fluid mining. Opencast mining is done either from pits or gouged-out slopes or by surface mining, which involves extraction from a series of successive parallel trenches. Dredging is a type of surface mining, with digging done from barges. Hydraulic mining uses jets of water to excavate material.

Underground mining involves extraction from beneath the surface, from depths as great as 10,000 ft (3 km), by any of several methods.

Fluid mining is extraction from natural brines, lakes, oceans, or underground waters; from solutions made by dissolving underground materials and pumping to the surface; from underground oil or gas pools; by melting underground material with hot water and pumping to the surface; or by driving material from well to well by gas drive, water drive, or combustion. Most fluid mining is done by wells. In one experimental type of well mining, insoluble material is washed loose by underground jets and the slurry is pumped to the surface. *See* COAL MINING; OPEN-PIT MINING; PETROLEUM ENGINEERING; PLACER MINING; SOLUTION MINING; SURFACE MINING; UNDERGROUND MINING.

The activities of the mining industry begin with exploration, which, since accidental discoveries or surficially exposed deposits are no longer sufficient, has become a complicated, expensive, and highly technical task. After suitable deposits have been found and their worth proved, development, or preparation for mining, is necessary. For opencast mining, this involves stripping off overburden; and for underground mining, the sinking of shafts, driving of adits and various other underground openings, and providing for drainage and ventilation. For mining by wells, drilling must be done. For all these cases, equipment must be provided for such purposes as blasthole drilling, blasting, loading, transporting, hoisting, power transmission, pumping, ventilation, storage, or casing and connecting wells. Mines may ship their crude products directly to reduction plants, refiners, or consumers, but commonly, concentrating mills are provided to separate useful from useless (gangue) minerals. *See* PROSPECTING.

Depletable resources. A unique feature of mining is the circumstance that mineral deposits undergoing extraction are "wasting assets," meaning that they are not renewable as are other natural resources. This depletability of mineral deposits requires that mining companies must periodically find new deposits and constantly improve their technology in order to stay in business. Depletion means that the supplies of any particular mineral, except those derived from oceanic brine, must be drawn from ever-lower-grade sources.

Evan Just

Underground Excavation

In all types of mining, there are at least three major operations: the ore or waste rock is broken from place, the broken material is transported to its final destination, and the opening caused by the excavation is supported, either naturally or artificially. In underground mining, where it is necessary to go below the surface of the earth to mine the ore body, three other operations must be considered: removing the material from the mine; controlling and pumping underground water; and providing a suitable and healthful working environment for the miners. Generally, all of these operations are involved in underground excavation. Depending on circumstances, some of these operations may be much more important than others. For example, in a dry mine with very strong ore and waste, support and pumping would be of minor importance.

There are two general methods used to mine the ore body below the surface of the earth. If the terrain is amenable, it is possible to excavate horizontal openings called adits into the mountain (**Fig. 1***a*). However, if the mine is located in relatively flat country or the ore lies deeper than the lowest adit level, it becomes necessary to begin by sinking an opening called a shaft from the surface or adit level into the earth (Fig. 1*b*). This shaft may be vertical or inclined even to some small angle (such as 10°) to the horizontal. Low-angle shafts are frequently called declines.

From the shaft at some predetermined spacing, horizontal openings called levels are excavated toward the ore body to prepare the ore for mining. In most cases, the ore is mined so that gravity will cause the broken ore to fall or at least to be directed toward the level below. It can then be transported to the surface through the adits or shafts. Even though a mine may be started through adits, it is usually necessary to sink a shaft inside the mine if the mine is a success. This is called a winze (Fig. 1*a*).

To develop the ore for mining, it is usually necessary to drive openings known as raises from a lower level to a higher level. (Fig. 1*a*, *b*). The raises provide access for personnel and equipment into the stope (the zone where the ore is actually mined) or to pass ore from the mining area to the level below. The other openings except the stope, shafts, levels, and raises are known as development headings.

There are many different types of stoping methods. In practically all methods, the ore must be broken from place and transported to the level below. In some types of stoping, the working area must be made safe for the personnel. In other types, such as caving methods, people must work in the vicinity of the stope in a safe manner so that the ore can be removed from it. Two common methods of stope preparation for ore removal involve work at the bottom of the stope (Fig. 1*c* and *d*). Some ore bodies are nearly horizontal and similar to coal deposits. In

this situation, the ore is mined in a similar way to the room-and-pillar coal mining method.

A brief discussion of underground excavation requires some understanding of shafts, levels, raises, and stopes, because the excavation procedures vary in each of these types of openings. *See* COAL MINING.

Breaking. Ore and waste are broken from place by mechanical means, blasting, or natural forces.

Mechanical means. Mechanical means consist of drilling (augering, percussion, rotary, and abrasion such as with diamonds), dozing or scraping, ripping, chipping, and cutting. Drilling methods are usually employed to place explosives where blasting is necessary. In underground mining, machines designed to auger, scrape, chip, rip, or cut or combinations of these are used to mine softer ores and rocks. These machines are classified as continuous mining machines or long-wall cutting machines. They are used to a great extent in mining coal and some nonmetallic minerals such as trona and salt, or other soft minerals.

Blasting. In most development headings such as shafts, levels, and raises, and in many stopes, the solid material is broken by the force of explosives. It is necessary to place the explosive charge in the rock mass where it can break the solid material efficiently. Hence, holes are drilled in the solid material with some predetermined pattern called drill rounds (**Fig. 2**). The explosives break the material toward a free face.

There is a hole pattern (Fig. 2*a*) that is commonly used in stopes where the explosive force can break toward a free face, which is either horizontal or vertical depending on the type of stoping operation. The holes are then drilled parallel to the free face. This procedure allows the explosive charge to be placed parallel to the face for maximum efficiency.

In a development heading, there is one free face where the holes are drilled; thus, the holes are positioned so they will either break toward this free face or create another free face. The holes in this pattern are known as the cut (Fig. 2*b*–*e*). The cut and hole pattern in a round varies greatly from mine to mine depending on the nature of the ground. As a rule, the maximum depth that a round will break is the shortest dimension of the opening (Fig. 2*c*–*e*).

The burn cut (Fig. 2*b*) is designed to break deeper than the other types of rounds. In the cut, one or more holes are drilled and left unloaded, without an explosive charge. Some holes are left uncharged with explosives. Thus, the unloaded hole is the free face that the loaded hole can break toward. Successive holes then break to this opening. The unloaded hole or holes are often drilled at a larger diameter than the regular loaded holes so that there is a bigger volume in the free face. The depth of burn-cut rounds are generally not limited to the smaller dimensions of the opening.

In all rounds, the order or sequence of detonating the explosive charges, called firing order, is extremely important. The holes nearest the free face must explode first to create and expand the opening (Fig. 2*a* and *e*). *See* EXPLOSIVE.

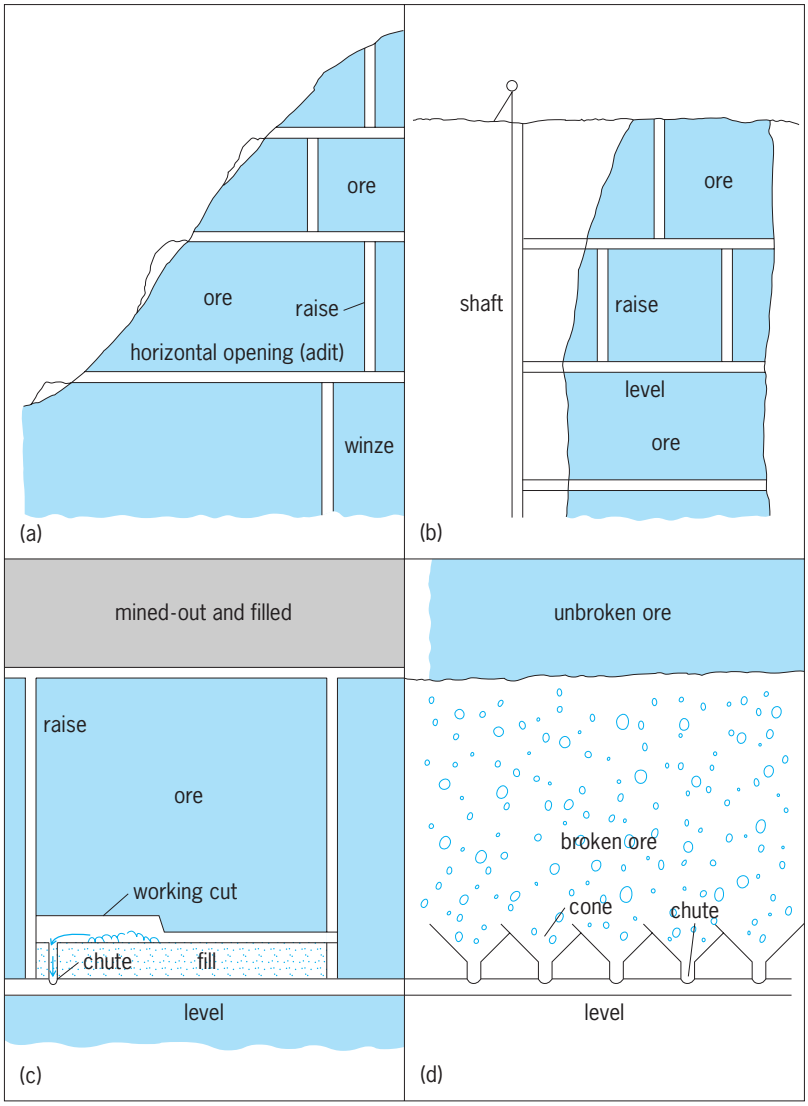

Fig. 1. Long sections of underground mines. (*a*) Ore body developed by adits. (*b*) Ore body developed by a shaft. (*c*) Bottom of stope prepared for horizontal ore movement. (*d*) Bottom of stope prepared by cone-shaped openings for ore movement.

Natural means. Breaking also occurs by natural means if the material surrounding the opening is not adequately supported. Caving methods take advantage of this phenomenon. The general plan is to remove underlying support from a block of ore, which allows the ore to cave into the opening left by the removed supports. The broken ore is directed into a series of cones at the bottom (Fig. 1*d*). As the broken ore is removed, an opening below the unbroken ore develops and the ore caves into this opening. As a rule, broken rock expands to about 140% of its original volume; therefore, as the broken ore develops, it fills in the opening, which stops the caving action. As the broken ore is pulled from the bottom, a new opening in the stope occurs and caving starts again. Where applicable, this system is very cost-effective and produces high tonnage rates of ore mined per day.

Loading and transporting. When the ore or waste is broken, it must be moved or transported to some

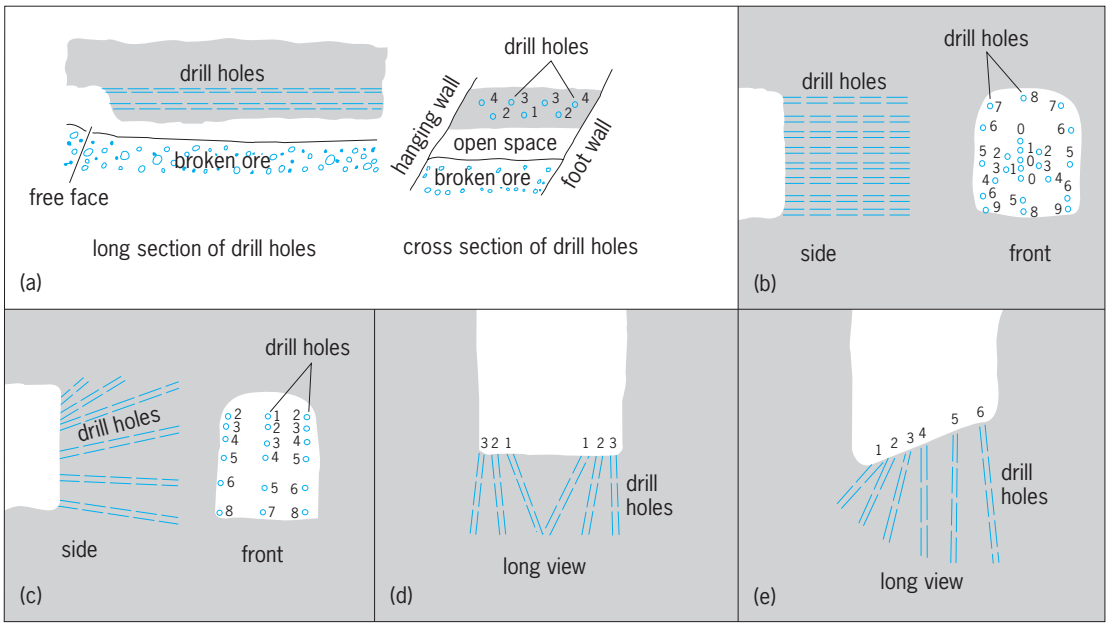

Fig. 2. Long and cross sections of drill rounds. (*a*) Typical stope round. (*b*) Burn-cut round (drift), the 0 holes are left uncharged. (*c*) Drift round with top draw cut. (*d*) V-cut shaft round. (*e*) Sump or bench shaft round. The order of firing follows the number sequence.

other place, either in the mine or to the surface of the mine. Usually this operation consists of loading some type of conveying device such as a truck or railroad track type of car. The mining industry has developed combined loading and transporting equipment called load-haul-dump units (**Fig. 3*a***).

The typical hand shovel, power shovel, and wheel or crawler loader and some underground loaders have a specific type of action (Fig. 3*b*). The bucket or shovel is pushed or crowded into the loose material as it is also lifted through the material, filling the bucket. The bucket is then directed to and dumped into the hauling conveyance. The hauling unit takes the broken material to its next destination point.

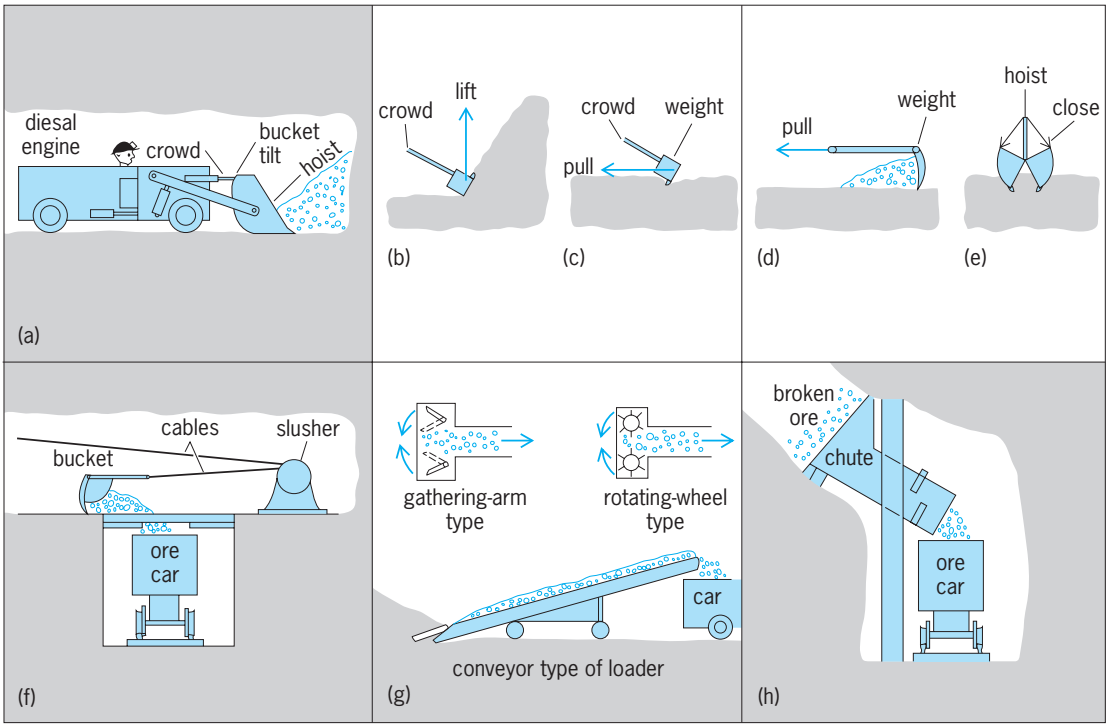

Fig. 3. Loading principles. (*a*) Load-haul-dump unit. (*b*–*e*) Types of loading action: (*b*) filling bucket by shovel action; (*c*) filling backhoe bucket; (*d*) slusher bucket transporting; (*e*) filling clamshell bucket. (*f*) Scraper and gravity loading. (*g*) Gathering and conveyor loading. (*h*) Gravity through chute loading.

In another action (Fig. 3*c* and *d*), the weight of the bucket or the combined weight of the bucket and machine causes the bucket to crowd into the loose material from the top while the pulling action fills the bucket. The loaded bucket can then either be hoisted, as in a pull-shovel or dragline (Fig. 3*e*), or be dragged along a hard surface (Fig. 3*f*) and dumped into a designated conveyance or area. Underground slushers and scrapers use this action and are popular in underground use.

Some machines use a combined crowding, raking, and lifting action (Fig. 3*g*). Machines using this action are high-capacity loaders, more commonly found in coal and nonmetallic soft ore types of mines. Chutes take advantage of gravity to load underground conveyances (Fig. 3*h*).

Loading in larger underground mines is more frequently done by diesel-powered wheel or tractor types of loaders that load into diesel-powered large trucks specially equipped for underground use. However, not all underground mines lend themselves to efficient use of wheel loaders and trucks, and some must rely on the standard types of loaders and conveyances.

Support. In underground excavation, one of the problems frequently encountered is supporting the excavation either during or after the excavation process. This process disturbs the natural forces that occur in the ground mass, and in the readjustment process forces act on the sides of the excavation (**Fig. 4***a*). If the surrounding material is strong enough, no support is required. In other situations, the opening may stay open initially, but force readjustment in the ground mass or surface deterioration may cause the opening to fail at some time after it is excavated.

Rock bolts are probably the most frequently used means of development heading support, either temporary or permanent. Drill holes are placed into the rock mass (Fig. 4*b*), and the holes are filled with rock bolts that are anchored into the rock mass by various means. A head board on the surface of the opening is attached to each bolt and tightened firmly against the ground mass. This action keeps the groundmass firmly in place. Rock bolts are commonly made of steel rods or tubes; occasionally steel cable and fiberglass are used, being anchored in the holes by wedging, expansion shields, concrete, and various types of epoxy glues, and by expansion of split hollow tubes.

Rock bolts are sometimes used as temporary support in stopes before a more permanent support can be installed. However, in some development headings, such as shafts and semipermanent levels, a more lasting support is required. Wood (Fig. 4*c*) and concrete, with or without rock bolts, are frequently used. Timber provides good support for a considerable time and has the advantage of being easily repaired if a failure occurs. In mining situations, ground forces are continually changing; consequently, failures to support systems do occur. Concrete is a more permanent support, but it is very expensive. It is often used for shaft support, but less frequently for level support except in block caving mines where nothing else is satisfactory or where the level is to stay open for a long time.

The type of support required often dictates the mining method used for stoping out the ore. In some cases, the ground will support itself over large spans, and this is called open stoping. In other cases, minimal support is required, and wooden props (stulls) may be enough support until the zone is mined out;

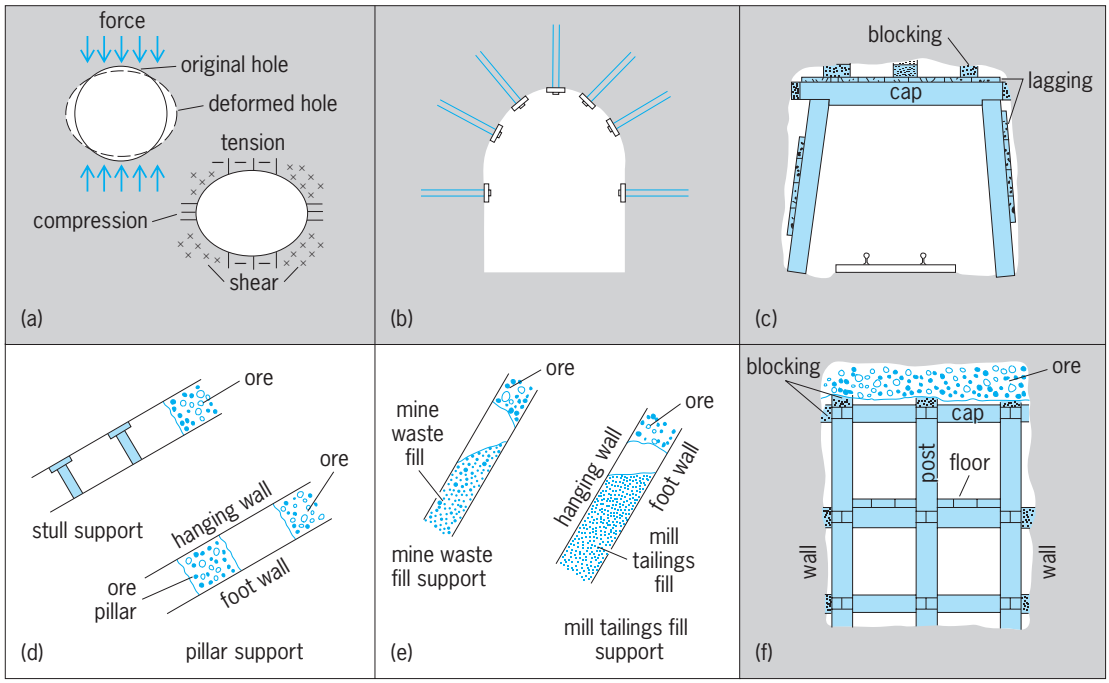

Fig. 4. Sections showing various types of support. (*a*) Deformation caused by forces. (*b*) Bolted cross section of level. (*c*) Timbered cross section of level. (*d*) Support in stopes. (*e*) Mine-waste support. (*f*) Timber support.

pillars usually give more support because they are bigger, but they are composed of ore and not all of the ore can be removed in pillar mining (Fig. 4*d*). Broken waste, ground, or fine mill tailings are often used for stope support (Fig. 4*e*); this is classified as a cut-and-fill system. Timbered systems are used for high-grade ore and in very weak or heavy ground (Fig. 4*f*); the timbered stopes are also filled with mill tailings for additional support. Although, technically speaking, caving systems use no support, the development headings required to extract the caved ore often require very heavy support. Koehler Stout

Machinery

Machinery is used in removing and transporting valuable solid minerals from their place of natural origin to a more accessible location for further processing or transportation. Many of the machines are identical to, or minor adaptations of, those used for excavating in the construction industry. In a wider sense, mining machinery could also include all equipment used in finding (exploring and prospecting), removing (mining: developing and exploiting), and improving (processing: ore dressing, milling, concentrating or beneficiating, and refining) valuable minerals; it could even include metallurgical (smelting) and chemical processing equipment used in extracting or purifying the final product for industry. The term mining machinery is also applied to special equipment for recovery of minerals from beneath the sea. In its usual context, the term does not include apparatus used principally in the petroleum industry. Perhaps those machines most often considered as uniquely mining machinery are drills, mechanical miners, and specially adapted materials-handling equipment for use in mining underground or on the surface (where a large proportion of mines are located). In addition, some unique auxiliary equipment and processing equipment are used in the mining industry.

Design and construction. All components of mining machinery—including primary mechanism, controls, means of powering, and frame—require the following features to a much greater extent than do other machines (with the possible exception of some military, construction, oil well, and marine units).

1. *Ruggedness*. Equipment is handled roughly, frequently receiving severe and sudden shocks from dropping, striking, and blasting vibration; overloading is common, and long life is demanded by the economics of mining.

2. *Weather resistance*. Operations extend over a wide range of climate and altitude.

3. *Abrasion resistance*. Minerals include some of the hardest substances known. Dust and fine particles are always present.

4. *Water and corrosion resistance*. Moisture and water, often acidic, are common in mining operations.

5. *Infrequent and simple maintenance*. Equipment is often widely scattered and in locations with restricted access. Trained mechanics and repair parts are generally limited in availability because of the remoteness of operations.

6. *Easy disassembly and reassembly*. Access to the machinery at the site of operation is frequently limited. Also, the working space near it and mechanical aids to moving or lifting it may be limited or nonexistent.

7. *Safety*. Because mining environments can expose workers to physical risk, most governments have testing bureaus, inspection agencies, and enforcement laws for the approval of mining equipment.

Simplicity of operation, low initial cost, and low operating costs are also desirable features of mining machinery.

Exploration machinery (vehicles, drills, and accessory equipment) is subjected to the same operating conditions as other mining equipment. Mineral-beneficiating equipment must have, above all, abrasion resistance. Smelting equipment has the requirement of heat resistance. Chemical refining process equipment must be highly resistant to corrosion. Reliability of all processing equipment is critical because slurries are commonly handled, and they can cause considerable difficulty in restarting after shutdowns.

Underground requirements. Machinery operated underground must meet special design requirements:

1. *Low-ventilation demand*. Quantity and geometry of passageways for air are rather rigidly fixed, so that high air consumption is a problem and noxious gases cannot be readily dispersed. Heat removal is a problem in deep mines.

2. *Compactness*. Space is at a premium, especially height, particularly in bituminous coal mines.

3. *Easy visibility*. Most operating areas are lighted only by individual cap-mounted or hand-held lamps.

4. *Hand portability*. Units or components must frequently be hand-carried into an operating area.

5. *Absence of spark and flame*. Equipment is often used in or near explosives, timber supports, and combustible gas or dust. In the presence of hydrocarbons, as well as of certain metal ores such as some sulfides, complete absence of open sparks or flames is a major requirement.

Power source. Mining machinery is very commonly powered by compressed air, but electricity is also widely used and is often the basic source. Compressed air has the advantages of simplicity of transmission and safety under wet conditions. It is especially advantageous underground as an aid to ventilation. Machines powered by compressed air can be easily designed to accommodate overloads or jamming, which is desirable on the surface as well as underground. Larger central compressors and extensive pipeline distribution systems are common, especially at underground mines.

Electric power, purchased from public sources or locally generated at a large central station, is common in open-pit and surface mines and dredging operations. Underground coal and saline-mineral mines often use electric-powered production machinery, but in other underground mines electricity is normally used only for pumps and transportation

systems in relatively dry or permanent locations. Direct-current devices have been dominant because of simplicity of speed and power control, but alternating-current apparatus has become common. Mobile equipment is often either battery or cable-reel (having a spring-loaded reel of extension power cable mounted on the machine) type. Processing machinery units are almost exclusively powered by individual electric motors.

Diesel engines are popular for generating small quantities of electric power in remote areas and for transportation units. Underground, abundant ventilation is essential, as well as wet scrubbers, chemical oxidizers, and other accessories to aid the removal of noxious and irritant exhaust gases. Hydraulic (oil) control and driving mechanisms are widely used. Transfer of power by wire rope is common, especially for main vertical transportation.

Drills. Drills make openings, of relatively small cross section and long length, which are used to obtain samples of minerals during exploration, to emplace blasting explosives, and to extract natural or artificial solutions or melts of minerals. Exploration holes are generally vertical or inclined steeply downward, less than 6 in. (15 cm) in diameter and up to 10,000 ft (3 km) long. Blastholes range from 0.75 to 12 in. (2 to 30 cm) in diameter and usually are under 50 ft (15 m) long in any direction, with the larger usually downward. Solution wells normally are vertical and 6 to 12 in. (15 to 30 cm) in diameter. They sometimes reach depths of several thousand feet, and are equipped with several concentric strings of pipe.

Rock drills. Percussion, rotary, or a combination action of a steel rod or pipe, tipped with a harder metal chisel or rolling gearlike bit, chips out holes up to 12 in. (30 cm) or more in diameter by 125 ft (38 m) or more long from the surface, and 1–3 in. (2.5–8 cm) by 5–200 ft (1.5–60 m) from underground. Crawler or wheeled carriers are used, and the smaller drills are often attached to hydraulically maneuvered booms. Air or liquids flush out the chips.

Diamond drills. Rotation of a pipe tipped with a diamond-studded bit is used in exploration to penetrate the hardest rocks. Large units make holes up to 3 in. (8 cm) in diameter and more than 5000 ft (1.5 km) deep; at the other extreme are units so small that they can be pack-carried. A cylindrical core is usually recovered.

Water-jet drills. For exploration and blasting in loose or weakly bonded materials, a water jet washes out a hole as a wall-supporting pipe is inserted.

Jet flame drills. For economical surface blastholes in hard abrasive quartzitic rock, a high-velocity flame is used to spall out a hole.

Mechanical miners. There are many machines designed to excavate the valuable mineral or the access openings by relatively continuous dislodgement of material without resorting to the more common practice of intermittent blasting in drill holes. These units also frequently transport the mineral a short distance, and when designed for weakly bonded minerals, they often become primarily materials-handling equipment.

Continuous miners. For horizontal openings in coal and saline deposits, toothlike lugs on moving chains, or rotating drums or disks rip material from the face of the opening as the assembly crawls ahead.

Longwall mining systems. A longwall mining system consists of a drum shearer (also called a plow), shield supports (movable roof supports) and controls, face conveyors, and crushers. The shearer is used to mill coal from the seam as it is driven along the wall, which may be many hundreds of meters long.

Augers. Coal and soft sediments are mechanically mined by augers up to 5 ft (1.5 m) in diameter and 100 ft (30 m) long, usually used horizontally.

Shaft and raise drills and borers. For vertical and inclined openings up to 8 ft (2.4 m) or even larger in diameter, various rotary coring and fullface boring equipment is used, both in an upward direction (raising) for several hundred feet or downward (sinking) for several thousand feet. These units usually use many rigid teeth or rolling gearlike bits to chip out the mineral.

Tunneling machines. There are rotary boring units for horizontal, or nearly so, openings of any length and up to 35 ft (11 m) in diameter (in soft rock).

Rock saws. To remove large blocks of material, narrow slots or channels are cut by the action of a moving steel band or blade and a slurry of abrasive particles (sometimes diamonds) rather than teeth. Small flame jets are also used.

Hydraulic monitors. Water jets of medium to high pressure (some to 5000 lb/in.^{2} or 34 megapascals) are used to excavate weakly cemented surface material and brittle hydrocarbons on both the surface and underground. In some instances, the jets are pulsed.

Special materials-handling equipment. Loose material (muck) is picked up (mucked or loaded) and transported (hauled or hoisted) by a wide variety of equipment.

LHD (load-haul-dump). LHDs are loaders designed specifically to work in confined underground spaces. They resemble the front-end loaders often found at construction sites, but are not nearly as tall. They are used to transfer muck (broken ore) from the face (or draw point) and deposit it at the ore pass (dump site), often by dumping it through a screen to exclude rocks over a certain size. Because of the confined environment, LHDs are often driven in reverse, and the operator sits sideways to the travel direction to see ahead and behind the vehicle.

Excavator loaders. For confined places underground there are various unique grab-bucket shaft muckers and overcasting shovel tunnel muckers, and gathering-head loading-conveyor units having eccentric arms, lugged chains, and screws or oscillating pans for handling muck in horizontal openings.

Dragline scrapers. Scrapers (slushers) with a flat plowlike blade or partially open bucket pulled by a wire rope are commonly used to move muck up to a couple of hundred feet, especially in underground mines.

Dipper shovels and dragline cranes. In surface mining, single-bucket loads can handle up to 100 yd^3 (75 m^3) of material. Many of the intermediate size (20–40 yd^3 or 15–30 m^3) units move on unique walking shoes.

Bucketline and bucketwheel excavators. These are for surface use and can dig up to 5000 yd^3/h (3800 m^3/h), using a series of buckets on a moving chain or a rotating wheel supported on crawlers, railcars, or floating hulls (dredges).

Suction dredges and pipelines. On the surface, up to 3700 yd^3/h (2800 m^3/h) of moderately loose mineral up to several inches in diameter can be picked up and moved as a slurry (mix of water and solid material) by pumps mounted on floating hulls.

Trucks. Diesel and electric shuttle cars (short-haul trucks) of unusual design, often having very low profile and conveyor bottoms, are used underground. At surface mines, there are diesel and other dump trucks with 330 metric tons capacity.

Railroads. Underground locomotives with electric (storage battery, cable-reel, or trolley), diesel, and sometimes compressed-air power units are used to transport ore cut of the mine. Cars are usually of special design for automatic dumping.

Conveyors. Unique movable, self-propelled, sectional, and extensible conveyors are used in underground mining, especially in longwall and continuous mining operations. In surface mines, conveyors are often used to carry ore over land.

Wire rope hoists. Hoists or winders are used in shafts for vertical or steeply inclined transportation in single lifts of as much as 6000 ft (1.8 km). Of various particular designs, there are two basic types: drum, simply a powered reel of rope; and friction, in which the rope is draped over a powered wheel and a counterweight is attached to one end and a conveyance to the other.

Auxiliary equipment. Drainage pumps handling hundreds of gallons of water per minute at heads of 1000 ft (300 m) or more are used underground. Ventilation fans are capable of moving several hundred thousand cubic feet of air per minute. Crushers can handle pieces of hard rock several feet across in two dimensions. In processing, minerals are sorted by size or density, or both, by a variety of screens, classifiers, and special concentrators, using vibration, fluid flow, centrifugal force, and other principles. Froth flotation and magnetic and electrostatic equipment take advantage of other special properties of minerals. *See* BULK-HANDLING MACHINES; GRINDING MILL.

John P. H. Steele; Lloyd E. Antonides

Mine Automation

Mining automation encompasses all aspects of the mining process from exploration, development, and extraction, to milling and product refinement. Automation of mining processes includes surface and underground operations as well as their supporting technologies, such as autonomous vehicles, communications, remote control, teleoperation, and robotics. Computers and advanced sensing devices are changing the way that materials are mined and processed. In many respects, because mobile mining machines manipulate the earth directly, the new autonomous mining systems being developed represent the most advanced robots produced so far. *See* AUTOMATION.

Underground. Load-haul-dump (LHD) vehicles, which are related to the front end loaders used in construction, are the main tool for transporting ore and waste rock in metal and nonmetal underground mines. Presently, LHDs are operated by miners who are physically located on the machine and drive it much like an automobile, except that the path is a drift and the operator sits sideways to see both ahead and behind the vehicle. In locations where there is significant danger of rock falling from the roof (roof fall), for example in open stopes (ore bodies), the LHD may be operated by remote control using either a tether (cable) or radio signal. The operator stands some distance away but in view of the vehicle and controls the vehicle by manipulating joysticks and switches. While remote operation has successfully protected miners from roof falls, there have been a number of accidents in which operators have been struck by the vehicle due to the tight operating quarters. *See* REMOTE-CONTROL SYSTEM.

New robotic LHDs are being developed and tested in Finland, Sweden, Canada, the United States, and Australia. In Kiruna, Sweden, one mine has several driverless LHDs working in production that perform automated hauling and dumping, and remote control loading. Automated dumping is relatively easy. After arriving at the dump site (typically an ore pass, which is a vertical shaft that funnels the ore to a lower level where it is conveyed out of the mine), the vehicle raises and tips its bucket, dumping the ore into the opening. Various sensors let the vehicle know when it has arrived at the ore pass. Navigating the drift is more challenging because the LHD must traverse the tunnel without hitting the walls, which in many mines are only tens of centimeters away. The most favored sensors for detecting drift walls are scanning lasers in which the time of flight of the light is used to determine the distance from the sensor to the wall. This technique allows the LHD to build up a two-dimensional model of the drift as the vehicle travels through the tunnel. Relatively simple control algorithms are then used to steer the vehicle, and its speed is adjusted as the passage becomes wider or narrower. Automated vehicles can travel through the drift at speeds comparable to those of experienced operators (60 km/h or a little less than 40 mi/h). Automated vehicles are less likely to hit the walls, once the control system is properly tuned, because of the closed-loop control between the laser sensors and the vehicle steering. *See* ALGORITHM; ROBOTICS.

The loading operation is more challenging and so far only teleoperated, because the operator must plan the loading operation and manipulate the bucket and boom while driving into the rock pile, called a muck pile (broken ore or waste rock underground; LHDs are sometimes referred to as muckers). In LHD automation, the operator is stationed in a control shack somewhere other than where the mucker is working (in some cases, on the surface, operating the machine underground). Remote control from a remote location is called teleoperation since the information about the operation must be transmitted, typically as an electronic signal to the operator, who

sends commands back to control the machine. *See* REMOTE MANIPULATORS.

There has been work on automating the loading operation as well. For automated loading, the load in the bucket must be sensed by measuring the pressure in the hydraulic cylinders that lift the boom and bucket. By measuring the forces acting on the bucket and understanding how the load is elevated, a control program (controller) can determine the next action of the boom and bucket. The successful approaches to this problem will likely use artificial intelligence techniques (fuzzy logic). To complete the automation of the loading operation, the robot (LHD) must be able to sense the muck pile and to plan its approach to the pile in order to load it, which requires the ability to sense the muck pile in three dimensions. Researchers are developing stereo vision methods for this sensing. *See* ARTIFICIAL INTELLIGENCE; FUZZY SETS AND SYSTEMS.

Automated loading is dependent on communication with a robot as it works in another location. The ability to send a signal through the earth is limited, so all underground mining operations rely on communications that follow underground passageways (drifts, raises, passes). A leaky feeder is a coaxial cable that allows signals to exit cables all along its length. Such a cable is strung throughout the mine, and two-way communications is possible. However, the technology limits the bandwidth of the signal, and techniques such as spread spectrum communication, like that used for cell phones, are being used to transmit video from the mine. *See* COAXIAL CABLE; MOBILE COMMUNICATIONS.

Once the loading operation is fully automated, there will be true robots working underground. The interaction between the operator and the LHD underground will be at the level of scheduling—directing the LHD to work in certain locations at certain times.

Another underground mining operation that has been automated is drilling. At the Kiruna mine in Sweden, production raise drills (steep-angle drilling machines) are operated via remote control from the surface. The drill rigs are driven into position, and the locations for the drill holes are sited and sent to the machine. Control of the operation is turned over to the machine, and the drilling operation proceeds with the remote operator acting as a supervisor while the drilling of the complete pattern of holes proceeds (these holes will be used to place the blast charges used to fragment the rock). This allows the operator to oversee the operation of several drills at the same time.

Surface mining. Surface mining involves the removal of covering soil (called overburden), the fragmentation of rock or coal by blasting, the transportation of the ore from the mine site to another location for processing (often to a mill), and the processing of the ore to enhance the ore content or to size-reduce the material for shipment. Automation of surface mines has taken several forms, such as automation of the conveyors that are sometimes used to move the ore over long distances, or the automation of the milling operation to reduce the ore to the proper size for the next processing step. The automation that has captured the imagination of many engineers is the development of driverless haul trucks, used to move ore from the active mining site to the next stage in processing. The routes for this transfer are well defined and can be planned ahead of time. The key to making the trucks driverless is the ability to measure a truck's location with great accuracy. This has become possible with the advent of GPS (Global Positioning System) and its refinement, Differential GPS. Depending on the number of satellites within the view of the GPS receiver and their distribution in the sky, GPS can be used to locate a truck within approximately 30 m (100 ft), a solution that is adequate for human navigation, but not good enough to steer a haul truck in a mine. The accuracy of the position measurement can be improved by including the error in the measure as determined by another GPS receiver at a known location. Because the longitude, latitude, and elevation are know at this location, the computed location, based on the signals received from the satellite, includes the measurement error. This error can be passed along to a truck's GPS receiver so that it can correct its computed location. This process reduces the positional error to less than a meter (more accuracy measurements are possible, but will not be discussed). This level of accuracy is adequate for large haul trucks to navigate mine roads without a driver. This technology is currently under development in a mine in Arizona, as well as other locations. *See* SATELLITE NAVIGATION SYSTEMS.

GPS technology is also being applied to dozers. By having a Differential GPS receiver on the dozer and knowing the relationship of the dozer blade to the location of the GPS receiver, and having a map that includes elevation information of the mine area, the dozer operator can be informed about the elevation of the material to be moved relative to the desired final level. In fact, the information from the computer-based map is better than that available from more traditional surveying techniques.

A third activity that has been improved by the development of GPS is the positioning of drill rigs. GPS receivers are placed on the top of the drilling towers so that the location of the hole being drilled (in preparation for loading it with a blast charge) can be more accurately located and the direction of the drilling can be monitored. This results in better control of the blasting operation and more uniform size of the blasted material. Such improvements save energy in the downstream processes and lower production costs.

Another area where automation is having an important impact on mining, both on the surface and underground, is machine health monitoring, sometimes called machine condition monitoring. By having sensors on the machine as it works, a much better understanding of the machine's condition can be developed. Trends or changes in the data can be tracked, and failures anticipated. By knowing when the condition of the machine is changing, catastrophic failures can be prevented and unplanned

downtime can be reduced. This leads to a more efficient and safer mining operation. This technology is making its way into a number of mining systems.

The implementation of automation in mining operations will lead to safer and healthier conditions for miners and to more efficient operations. The nature of the mining enterprise with its constantly changing environment and its lack of structure provides a great challenge to automation developers. Mining continues to challenge engineers to come up with new and cost-effective methods. Automation is one of the principal methods that will be used to continue this tradition of innovation. John P. H. Steele

Mining Power

Mining power refers to the means by which mining equipment is powered from some energy source. The process of extracting minerals from modern underground and surface mines is highly mechanized, requiring mobile mining machines that have variable power demands and adhere to stringent safety requirements. In addition, the surface activities of any mine, which may include shops, changing rooms, offices, ventilation fans, hoisting equipment, mills, preparation plants, and so forth, can have large power requirements. These facilities may receive power from the mine distribution system or at times from a separate substation.

The two predominant power sources for today's mining operations are diesel engines and electricity.

Diesel power. Diesel-powered equipment is used extensively in surface mining for front-end loaders, drills, trucks, locomotives, small shovels, and small draglines, with application essentially the same as in other industries. There is also wide use of diesel-powered equipment in many underground metal and nonmetal mines for front-end loaders, drills, trucks, shuttle cars, and locomotives. The use in underground coal mines is limited and, in these few mines, usually restricted to shuttle cars and locomotives. Because of the confined underground environment, the U.S. Mine Safety and Health Administration enforces strict standards on toxic components in engine exhaust gases, mine-ventilation dilution of exhaust gases, and the handling and storage of diesel fuel. In underground coal and other gassy mines, the diesel equipment used in mining-face areas must also adhere to additional stringent (permissibility) standards so that the machines cannot cause an explosion in the mine atmosphere. *See* DIESEL ENGINE.

Electric power. By far the most popular source of power in underground coal mines is electricity, and one reason is the high horsepower requirements of continuous miners and longwall machinery. Diesel engines cannot supply the mechanical power needed within the available space restrictions on these machines. Electricity is also favored over diesel power for large drills, shovels, and draglines in surface mines because of space restrictions. Large surface excavators can have 12,000 or more connected horsepower (9000 kW), and the largest is over 30,000 hp (22,500 kW).

Three-phase alternating-current (ac) power distribution and motors are used to power the various mining machines that were introduced into underground mines. Alternating current can be distributed at a higher voltage, thereby reducing current, voltage drops, and transmission losses, and providing adequate power delivery to the mining equipment. The use of silicon-diode rectifiers with large current capabilities has made possible the application of direct current (dc) for traction and of ac distribution and utilization on high power loads. For example, while continuous miners normally use ac, part of the ac supply at the power center is rectified to dc, primarily for powering the shuttle cars.

Simplified mine electrical distribution. Most underground-mine power systems have two voltage levels: distribution and utilization (**Fig. 5**).

In remote locations, a mining operation may generate its own power, but it is common practice for a mine to purchase all or most electric power from utility companies whenever it is available. Because utility voltages can range from 24 to 138 kV, a substation transforms the utility voltage down to distribution levels, most often at a voltage greater than 1000 V. In addition to the transformer, substations contain a complex of switches, protection apparatus, and grounding devices, which function as safety equipment. Power at the distribution voltage is distributed from the substation through conductors to a power center; hence, it is termed the distribution system. The power center, actually a portable substation, transforms the voltage to the correct level for equipment use, typically at less than 1000 V. It is at this level that power is normally delivered through conductors to the underground mining equipment. *See* ELECTRIC POWER SUBSTATION.

Originally, ac distribution was made at 2300 or 4160 V. In most mines, these levels were later increased to 7200 V, but a few operations have increased the distribution voltage to 12,470 or 13,200 V for both longwall and continuous mining operations. At first 440 V was the most popular voltage for ac equipment. When the continuous miner proved so successful, its horsepower was progressively increased; this resulted in an increase in the size of its trailing cable, until the weight was almost more than personnel could handle. To compensate, the most usual solution was to raise the rated motor voltages to 550 V. Manufacturers then produced machines with 950-V motors to further overcome the trailing-cable problems.

Underground mine distribution. In an underground mine, there is little freedom for varying system arrangements and the distribution system must almost

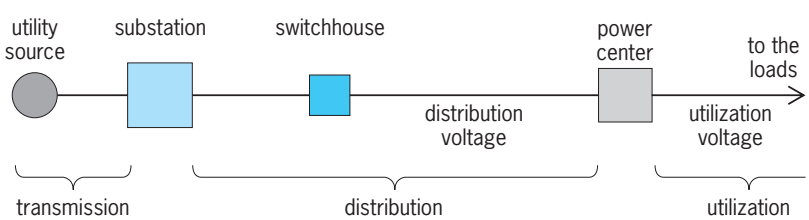

Fig. 5. Simplified electrical distribution system for mining.

always be radial; that is, a single power source and substation supply all equipment, and the system forms a treelike structure spreading out from the substation (**Fig. 6**). Power and mine grounding are fed underground in insulated cables, through a shaft, a borehole, or an entry. From a central disconnect switch, which allows total removal of underground power in an emergency, the power is distributed through cables to power centers that are located as close to the machinery as practical. The centers can power several face machines through couplers and trailing cables.

If belt haulage is used, small power centers are located close to all major conveyor-belt drives. With rail haulage, dedicated dc power centers (rectifiers) are placed along the rail to convert ac from the distribution system to dc. The distribution system can also serve large single motors directly through switchgear. All the power equipment used underground must be rugged, portable, self-contained, and specifically designed for installation and operation in limited spaces. As with diesel power, electrical equipment used in the face areas of gassy mines must be legally approved as permissible.

The primary purpose of any distribution system is to provide a flexible, easily moved or modified power source for the highly mobile mining equipment. The power system must be designed to function as an integral part of the total mine operation. The distribution system in any surface or underground mine that serves portable mining machinery is subject to damage from being run over by mining machinery, and as a result the system must be designed with optimum flexibility and consideration for personnel safety. *See* DIRECT-CURRENT TRANSMISSION; ELECTRIC DISTRIBUTION SYSTEMS; ELECTRIC POWER SYSTEMS.

Surface-mine power requirements. A combination of motor-generator sets driven by synchronous or induction ac motors, the Ward-Leonard speed control system, and dc motors is used on most mining excavators, especially the larger varieties. The standard 13,800-V distribution and machine voltages are for excavators larger than 100 yd^3 (76 m^3), while 23,000-V systems have been designed for machines greater than 200 yd^3 (153 m^3). Production shovels up to 18 yd^3 (14 m^3) commonly utilize 4160 or 7200 V, while in general 4160 V has become standard for machinery with 1500 hp (1125 kW) or less. As a result, more than one voltage level can be required at a mine when excavators of different sizes are employed. *See* ALTERNATING-CURRENT GENERATOR; ALTERNATING-CURRENT MOTOR.

Surface-mine distribution. Simple surface-mine power distribution (Fig. 5) is a very common arrangement in small surface operations where the distribution voltage is usually 4160 V. In the smallest mines, power is purchased at low-voltage utilization (often 480 V ac) and fed to a distribution box to which motors and equipment are connected. At times, simple distribution is also employed in large surface mines where only one machine must be served or where an extensive distribution network cannot be established, as in some contour mining operations.

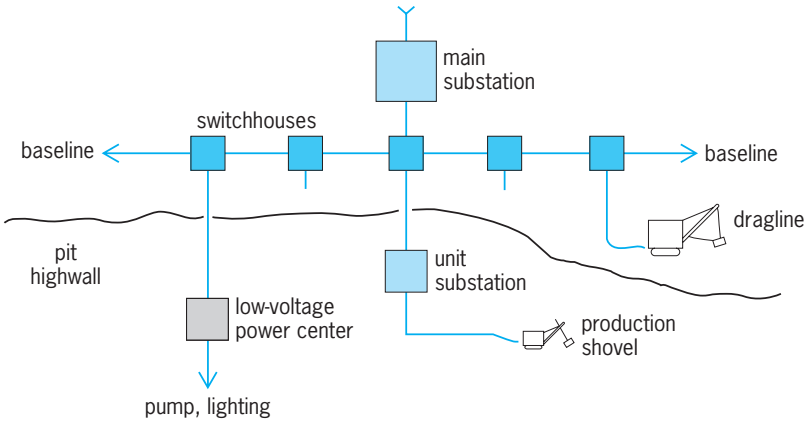

Fig. 6. Radial electrical distribution system for an underground mine.

The majority of surface mines employ radial distribution. In some area-type surface (strip) mines, the distribution system forms a base line which is usually located on the highwall, paralleling the mine pit (**Fig. 7**). Cables connected to the base line deliver power and grounding to skid-mounted switchhouses located on the highwall or in the pit. Specially designed trailing cables then complete the circuit to the machines. As the machines move along the pit, the base-line connections are successively changed to convenient locations. Open-pit power systems are similar with one main exception: the distribution system typically establishes a ring bus or main power line that partially or completely encloses the pit (**Fig. 8**). Radial ties to the bus complete

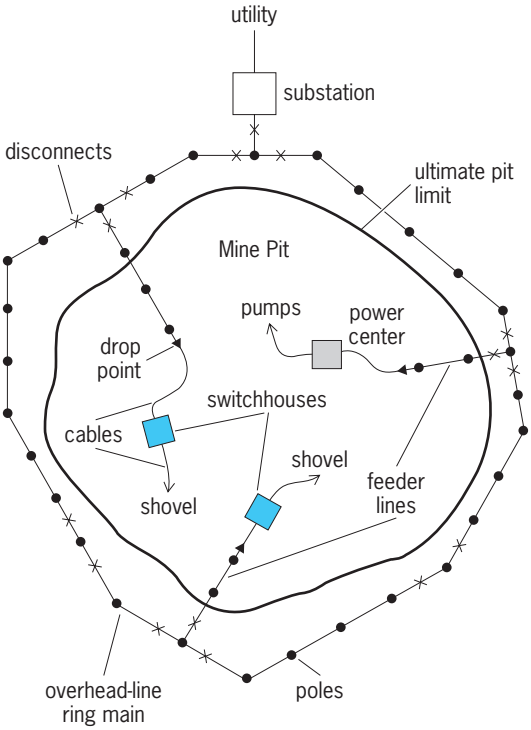

Fig. 7. Electrical distribution system for a surface mine.

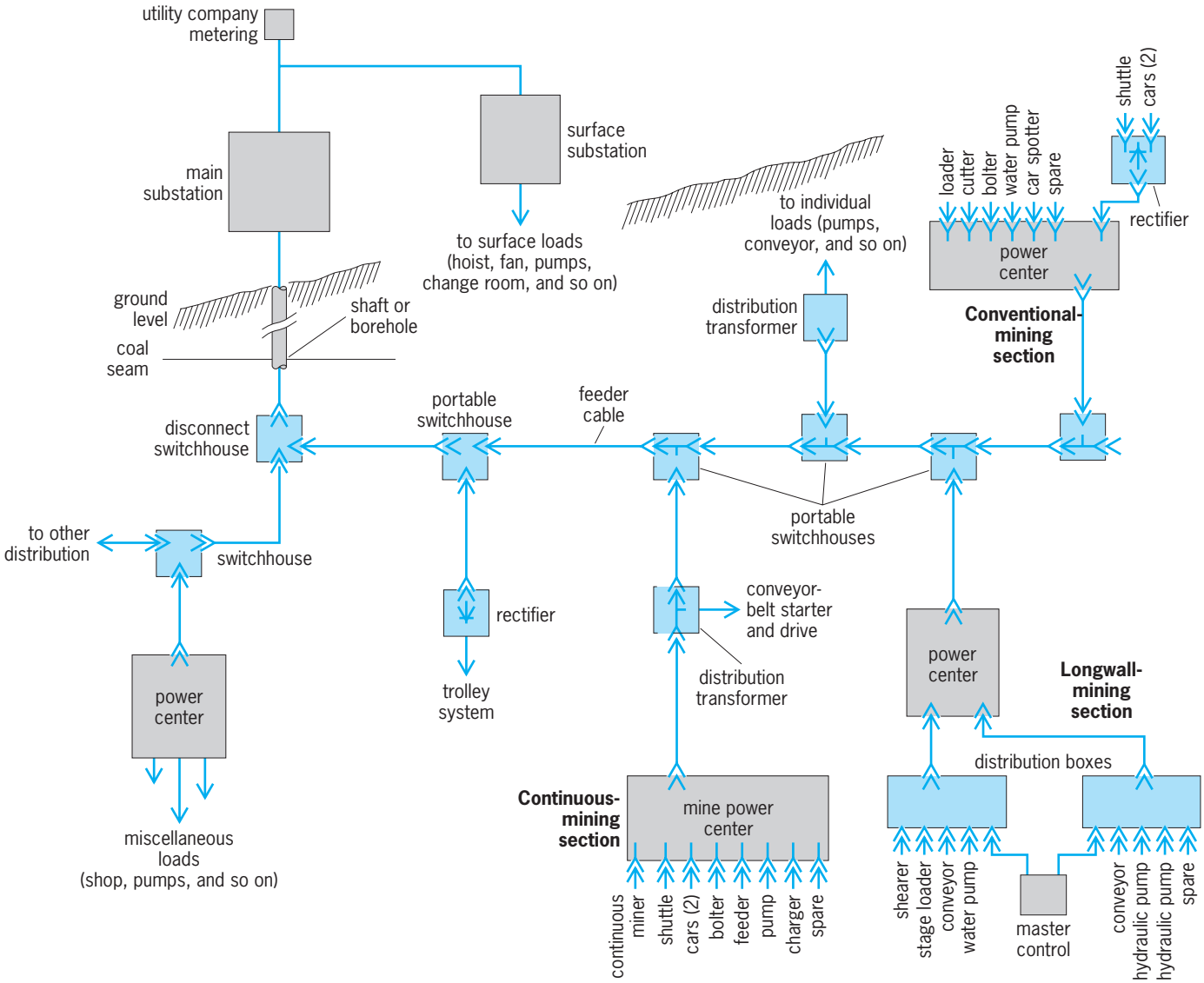

Fig. 8. Open-pit mine power system.

the circuit to switchhouses located in the pit, and portable equipment again uses trailing cables.

Lloyd A. Morley

Mine Ventilation

Mine ventilation consists of the tools, techniques, and methods used to create, maintain, and enhance the underground mine atmosphere for human health and safety. The objective of mine ventilation is to provide adequate quantities of fresh air to workers everywhere in the mine, as well as to render harmless and carry away toxic, noxious, and explosive contaminants, and to maintain heat and humidity at desirable levels. With increasing depth of mining, both ground pressure and virgin rock temperature increase, and are recognized as the limiting factors for operating mines at great depths. Among the areas of concern for worker health and safety underground, none is more critical to human survival than provision of a life-sustaining atmosphere in the hostile subsurface environment. *See* VENTILATION.

The principal atmospheric contaminants in underground mines are particulates (liquids or solids) and nonparticulates (gases, vapors, heat, and radiation). Liquid particulates include mists and fogs, and solid particulates include dust, fumes, smoke, pollen, and bacteria. The most common type of contaminants, however, are gases and dusts, and in deep mines, heat and humidity. Mine air contains several gases and particulates not encountered in normal air. Elevated concentrations of these contaminants are hazardous to worker health and safety. *See* GAS AND ATMOSPHERE ANALYSIS.

Minimum ventilation requirements for underground coal, and metallic and nonmetallic ore mines are prescribed in the Code of Federal Regulations (CFR), Title 30. The velocity and quantity of air required at designated sites (or spaces) are specified, as well as the quality of the air (such as the minimum or maximum concentrations of various gases and dusts). At selected intervals and locations, monitoring of the quality and quantity of a mine's air is required.

Several controls, such as legal, medical, engineering, and administrative, are applied to ensure compliance with standards. Achieving total mine air conditioning requires designing and controlling the quality, quantity, and temperature-humidity of the flowing air. Engineering principles fundamental to the quality and temperature-humidity control are, in order of preference, prevention (or avoidance), removal (or elimination), suppression (or absorption), containment (or isolation), and dilution (or reduction) of the contaminants. Often two or more of these principles are used together to achieve total air conditioning. Mine ventilation, primarily a dilution measure, is the solution to many mine air contaminant control problems. Severe air quality problems can be controlled by application of the other principles as well.

For quantity control, the principles of fluid dynamics are applied to analyze and design airflows in a mine. Airflow is induced from the atmosphere through the mine workings and back to the atmosphere by creating a pressure difference between the mine's openings. The mine ventilation system consists of fans (low pressure pumps creating the pressure differences), shafts, adits, and slopes (openings from the surface to the deposit being mined), airways (interconnections in the deposit between the openings to the surface and the mine working areas), and flow-regulating devices in the airways that control the direction and magnitude of airflow rates in the airways, shafts, and slopes.

Air-quality control. Air contains, by volume, about 78% nitrogen, 21% oxygen, 0.03% carbon dioxide, and less than 1% of argon and other rare gases. Mine air must contain at least 19.5% oxygen and no more than 0.5% carbon dioxide. It may also contain several other gases that are explosive, toxic, or both, and a host of particulates that are also toxic, explosive, or otherwise harmful to health. The limits to the airborne concentrations of these contaminants and durations of exposure are called the threshold limit values (TLVs). The TLVs are recommended by the American Conference of Governmental Industrial Hygienists (ACGIH), based on industrial experience, and animal and human studies. *See* INDUSTRIAL HEALTH AND SAFETY.

The formula to calculate the dilution air quantity required to bring the concentration of a pollutant to below its TLV at any location is Eq. (1). Here $Q =$

$$Q = \frac{Q_g}{(\mathrm{TLV} - B)} \tag{1}$$

the minimum amount of dilution air quantity (m^3/s), $Q_g =$ the rate of generation of the pollutant (m^3/s, mg/s), TLV = threshold limit value of the pollutant (%, mg/m^3), and $B =$ the concentration of the pollutant in the dilution air (%, mg/m^3).

When the amount of air quantity required for dilution is prohibitively high, measures such as prevention, removal, suppression, and containment are used to prevent the contaminant from entering the air stream to which miners are exposed. *See* AIR CONDITIONING.

Mine gases. The gases encountered in mining include methane, carbon monoxide, hydrogen sulfide, oxides of nitrogen, hydrogen, and radon. These gases emanate from several sources, including strata, blasting, internal combustion engines, broken ore, oxidation, incomplete combustion, and fires.

Methane is one of the most dangerous strata gases found in coal and noncoal mines. It does not support life and is explosive in the range 5–15%. The limits for methane are set at 1% for working areas. In some deep coal mines, 5 to 20 metric tons of air are circulated for every metric ton of coal mined to achieve methane control. To decrease this volume of air, and to make mines safer, several deep mines are practicing methane drainage, a process for removing methane through boreholes and pipelines, thereby decreasing the amount of methane entering the ventilating air stream. Radon, hydrogen sulfide, and carbon dioxide are other strata gases, less common than methane.

Carbon monoxide, sulfur dioxide, and oxides of nitrogen are highly toxic gases that arise from blasting and internal combustion engines. The first two can also come from slow oxidation, incomplete combustion, and fires that can occur in the deposit, broken ore, and the mine. Proper selection and use of blasting agents, prevention of oxidation and fires, and proper operation of diesel engines are essential to ensure that gas problems do not evolve into major health and safety hazards.

Mine dusts. The airborne dust in a mine atmosphere is likely to contain the minerals being mined. These particulates are generated during the cutting, drilling, blasting, loading, hauling, and other operations associated with the extraction of the deposit and the subsequent transport of the mined product to the surface.

On the basis of their health effects, mine dusts have been classified as fibrogenic, carcinogenic, toxic, radioactive, explosive, and nuisance. Coal dust is both fibrogenic and explosive. Silica dust is fibrogenic and a suspected carcinogen. Uranium, radium, and thorium dusts are radioactive and toxic. Particulates in diesel engine exhaust contribute to the respirable dust load in the mine air, and are suspected carcinogens.

Coal worker's pneumoconiosis is an occupational disease associated with long-term exposure to high concentrations of coal dust, leading to the formation of fibrous tissues in the lungs. Silicosis is a lung disease associated with inhalation of particulates containing sufficient quantities of silica. For coal mines in the United States, the current permissible exposure limit (PEL) for airborne respirable coal dust is 2 mg/m^3. It is an 8-hour, time-weighted average concentration. This standard is reduced to (10 ÷ percent silica) mg/m^3 when the respirable quartz content exceeds 5% in the airborne dust sample. The ACGIH permissible exposure limit for respirable quartz is 100 $\mu g/m^3$.

Respirable dust, once airborne, is difficult to control. Therefore, all control principles are applied to decrease the generation, reduce the entrainment, dilute the concentration, and control exposure. In

modern coal mining, the use of sharp cutting tools with appropriate power input is combined with the use of adequate water sprays to reduce the generation and entrainment of dust. Further, the cutting machines are equipped with dust extractors, and directional sprays to capture the dust, and if necessary, to steer the dust away from the miners. Operating the machines from a safe distance upwind, using remote control devices, and wearing personal protective equipment, such as respirators and air stream helmets, decrease the concentration of the dust. Despite these advances, respirable dust remains a hazard.

Temperature-humidity control. In very hot and humid working conditions, the body may be unable to cope with the heat stress. Miners' health may be threatened by heat illnesses, such as dehydration, heat rash, heat cramps, heat exhaustion, and heat stroke. Mine air conditioning for heat and humidity control becomes essential when airflow alone is inadequate to control the temperature. *See* AIR COOLING; HUMIDITY CONTROL.

Mine air cooling. The major sources of heat in mines, in order of their contributions, are the heat from the strata, the autocompression of the air as it flows down the intake shafts, and the powered equipment used in the mining process. The rate of change of the virgin rock temperature (VRT) with depth (called the geothermal gradient) is quite variable in different mining districts. The worldwide range is 0.7–6°C for every 100 m depth (**Fig. 9**). The effect due to autocompression is to increase the temperature at approximately 9.66°C for every 1000 m depth. Nearly all the energy input to machinery ends up as heat in the mine air.

The heat removal problem is acute in very deep mines due to the high VRTs, and an increased autocompression effect. As shown in Fig. 9, at an approximate depth of 1980 m, called the critical depth,

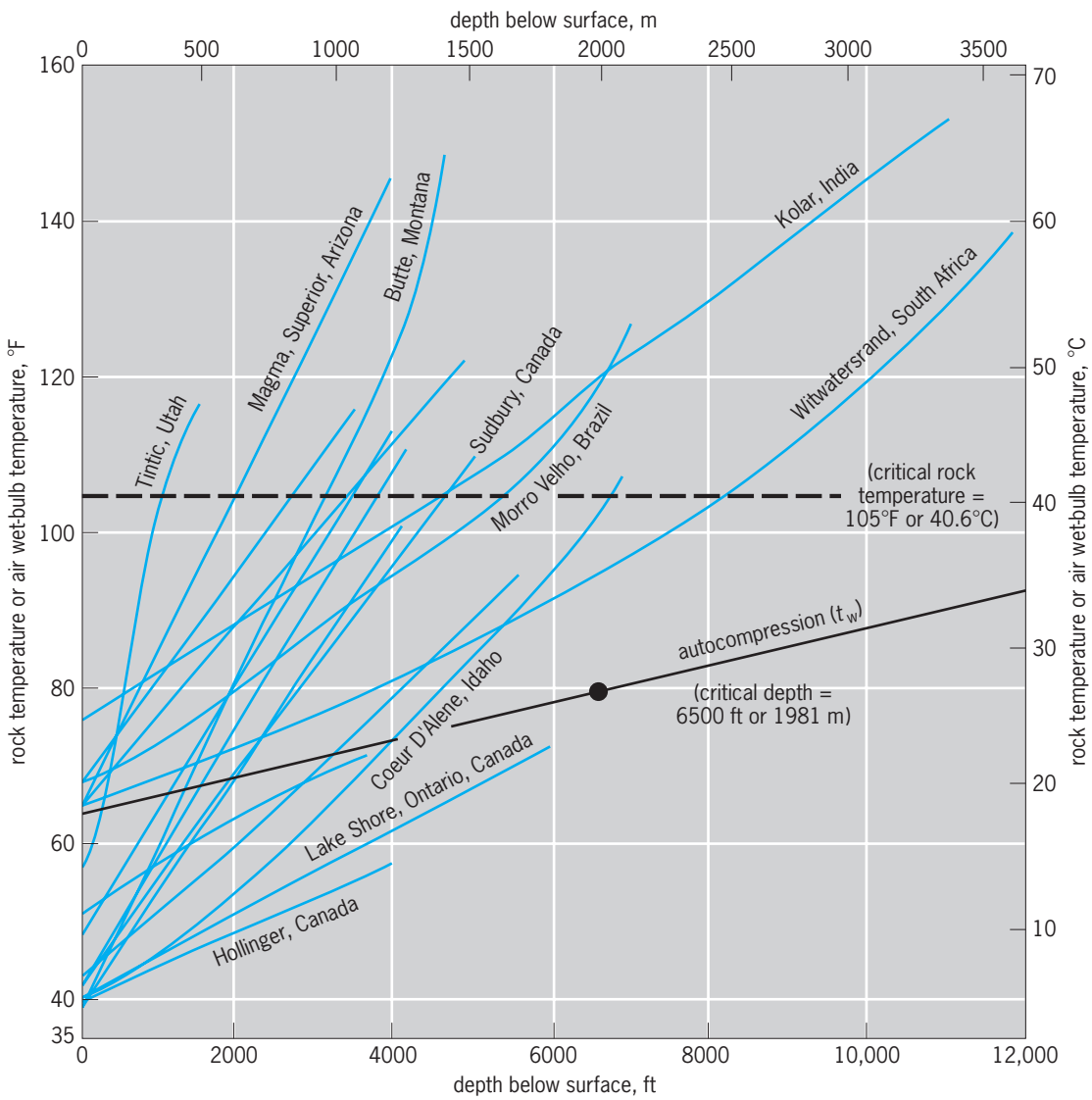

Fig. 9. Average geothermal gradients of various worldwide mining districts and effects of autocompression on air wet-bulb temperature (assumed surface $t_w = 64°F$ or 18°C, and $\delta t_w = 2.4°F/1000$ ft or 4.4°C/1000 m), both as a function of depth below surface. Critical rock temperature and critical depth for t_w are also shown. (*After H. L. Hartman et al., Mine Ventilation and Air Conditioning, Wiley, New York, 1997*)

autocompression alone governs that air conditioning be introduced for worker health and safety. Similarly, a VRT of 40.6°C (called the critical temperature) dictates the installation of air-conditioning systems. The depth associated with this critical temperature, however, varies as a function of the local geothermal gradient.

Mine air heating. Mines in very cold climates or in high altitudes may have very low temperatures. In rare cases, the combination of extremely low temperature and high-velocity air intake can create high wind chill factors, making it necessary to heat the air for the comfort of the miners. Such heating is not a common practice as miners can wear warm clothes. Where required, heating the air to 4–7°C is generally adequate.

Heating the intake air is more often required to protect the equipment and facilities in the intake shafts and slopes from damages due to freezing of water mains, ice buildup on the walls, and cyclical freezing and thawing of concrete linings. Equipment safety and production can be severely affected due to increased operational and maintenance problems in subzero temperatures. Heating is required only when subfreezing temperatures occur, and raising the intake air temperature to 1°C is sufficient.

As opposed to heating in winter, in summer months, mines in the arctic region may require refrigeration techniques to maintain the intake air temperatures to below the melting point of the permafrost.

Quantity control. Quantity control in the mine ventilation system means achieving optimally in the mine airways and working sections the desired quantities of airflow in the desired directions. The term "optimally" as used here is global and not restricted to ventilation only. The ventilation system must complement the mining method, and offer flexibility in meeting emergencies and future needs.

The mine airflow distribution is completely defined by (1) the physical parameters of the airways, including shape, area, length, and surface characteristics; (2) the layout of the mine openings; (3) the pressure sources (for example, fans) in the system, their location and characteristics; and (4) the interconnections between the airways, mine openings, and pressure sources using control devices, such as seals, stoppings, air crossings, regulators, and doors. Most air quality control activities take place at the mine's working faces, where the emanation of contaminants and the need for dilution are generally the greatest. The calculated dilution air must be conducted from the atmosphere to the working sections through the intake openings of the mine. After this air has served its purpose, it must be conducted back to the atmosphere through the return openings of the mine. In the working section, several methods of face ventilation are employed to direct the incoming air to the areas where workers are present and production operations are under way. The face ventilation in a room-and-pillar mine in coal is shown in **Fig. 10**.

The pressure necessary to create and maintain the airflow is provided by mechanical or natural sources.

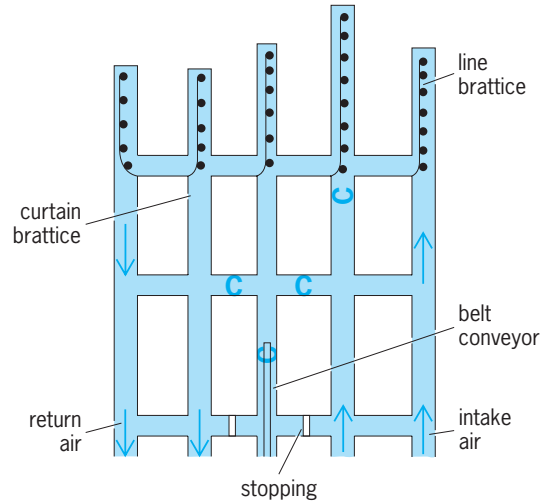

Fig. 10. Horizontal ventilation plan for room-and-pillar mine in coal. (*After H. L. Hartman, Mine Ventilation and Air Conditioning, 2d ed., p. 433, Wiley, 1982*)

The differences in the elevations and air temperatures in the mine openings create a natural draft, called natural ventilation pressure, which varies in both magnitude and direction. Therefore, all underground mines in the United States must be equipped with mechanical ventilators, which in most cases are fans. Compared to the atmospheric pressure, fan pressure is rather small, in most cases of the order of 2–3%. Therefore, it is generally adequate to treat airflow in mines as an incompressible phenomenon.

Some of the air flowing in the higher-pressure intake airways will leak directly into the lower-pressure return airways through the control devices. The quantity of air that is needed for dilution must be adjusted upward to allow for this leakage, and additional air for such needs as the ventilation of idled workings, shops, and battery charging stations must be provided. The volumetric efficiency (defined as the ratio of the volume of air flowing through the working areas to that flowing through the fans) is usually 50–60%. Estimation of the correct fan quantity is critical because the fan pressure and the fan power are proportional to the square and cube of the quantity, respectively. *See* FAN.

Air velocities are critical from two points of view. From the safety viewpoint, inadequate velocities in the airways can cause undesirable accumulations and layering of gases, and very high velocities can raise dust clouds. From the economic viewpoint, the head loss in a mine airway is proportional to the square of the velocity. Definition of the velocity and quantity requirements will automatically determine the number of airways needed to conduct the air once the shape and size of an airway are determined by mining-machine and ground-control conditions.

The frictional resistance of an airway to airflow is given by Eq. (2). Here R = resistance of the airway

$$R = \frac{KLO}{A^3} \tag{2}$$

(N-s^2/m^8), K = friction factor of the airway (N-s^2/m^4),

L = length of the airway (m), O = perimeter of the airway (m), A = area of the airway (m^2).

The frictional pressure loss for a quantity flow is given by Eq. (3). Here H_f = frictional pressure loss (Pa), V = velocity (m/s), Q = quantity (m^3/s).

$$H_f = \frac{KLOV^2}{A} = \frac{KLOQ^2}{A^3} = RQ^2 \qquad (3)$$

In addition to the friction loss, there are shock losses in an airway due to changes in airway cross section and/or airflow direction, as well as to obstructions in the airway. Shock losses can be as high as 30% of the total pressure loss in a mine ventilation system. The parabolic relationship between the quantity and the pressure or head loss for a given mine resistance is called the mine characteristic.

A desirable feature of any quantity distribution plan is to provide each working area with its own supply of fresh air. This is accomplished by splitting the intake air. Airways can be combined in series or in parallel or in combination to course the air to the various mine locations (**Fig. 11**). The combined resistance of n airways in series (R_s) is given by Eq. (4),

$$R_s = R_1 + R_2 + \cdots + R_n \qquad (4)$$

and that of n airways in parallel (R_p) is given by

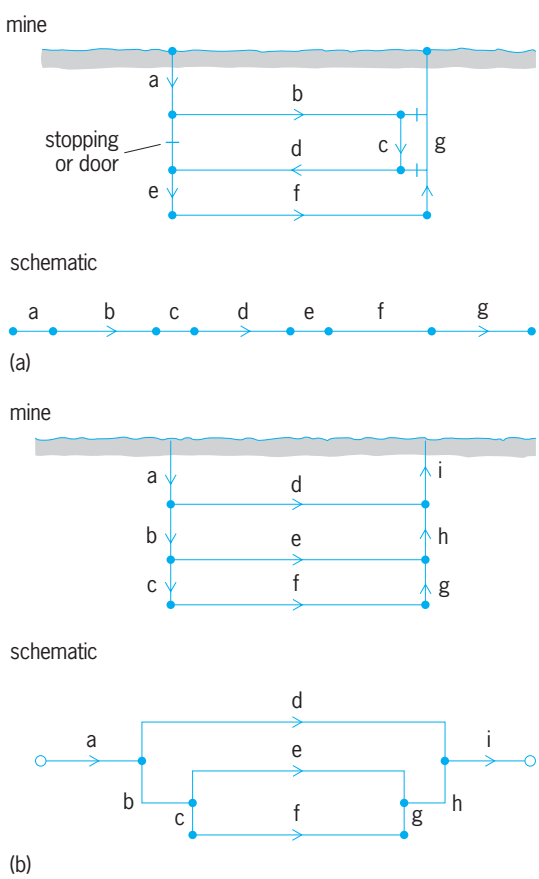

Fig. 11. Ventilation circuits. Letters designate airways. (*a*) Series circuit (*after H. L. Hartman, Mine Ventilation and Air Conditioning, 2d ed., Wiley, p. 174, 1982*). (*b*) Parallel circuit (*After H. L. Hartman, Mine Ventilation and Air Conditioning, p. 128, Ronald, 1961*).

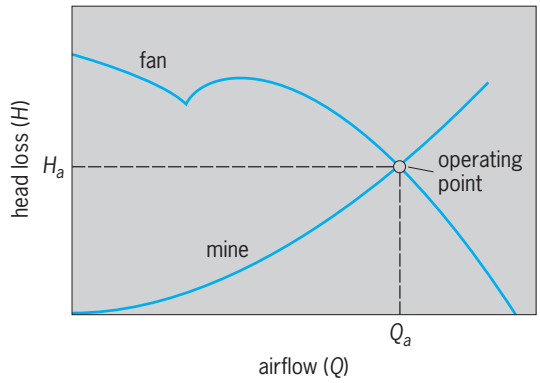

Fig. 12. Fan and mine operating point.

Eq. (5). Parallel flow requires multiple airways be-

$$\frac{1}{\sqrt{R_p}} = \frac{1}{\sqrt{R_1}} + \frac{1}{\sqrt{R_2}} + \cdots + \frac{1}{\sqrt{R_n}} \qquad (5)$$

tween locations but, as shown above, reduces the resistance to flow. It also enhances the opportunity to escape in the event of an emergency in one of the airways. In practice, combinations of series and parallel flows are used to split the intake air.

The pressure losses in individual splits are calculated once the resistance of the airways and the volumes of air circulating in them are known. The head loss in the system is calculated by accumulating the pressure losses in the individual splits using series and parallel flow laws as applicable. A fan of sufficient capacity that will operate at high efficiencies is selected, based on the quantity of the pressure head required for a mine. The pressure generated by a fan for various quantities of flow is called the fan characteristic. The operating point of a fan in a mine occurs at the intersection of the fan and mine characteristics (**Fig. 12**).

Mine ventilation systems. The mine ventilation principles apply equally to coal, metal, and nonmetal mining. In the United States, underground coal mining is confined to relatively flat-lying seams at shallow-to-modest depths. Similar conditions are encountered in some limestone, salt, trona, and lead-zinc ore mines, and their ventilation systems are similar to those in coal mines. However, most metallic ore deposits also have a relatively large vertical dimension, resulting in mining and ventilation practices that are not similar to those used in coal mining.

In coal mines, the primary pollutants are methane gas and coal dust. In metal mines, a wider range of pollutants is often encountered, such as diesel exhaust, radioactivity, and heat, as well as toxic-gas emissions from mining and blasting. In United States coal mines, the two air streams—the intake and the return—must not be allowed to mix, whereas in metallic and nonmetallic mines, under certain conditions, recirculation of some of the return air to the workings is allowed.

Although underground main fans are prohibited in United States coal mines, they are used in coal and metal mines abroad and in metal mines in the

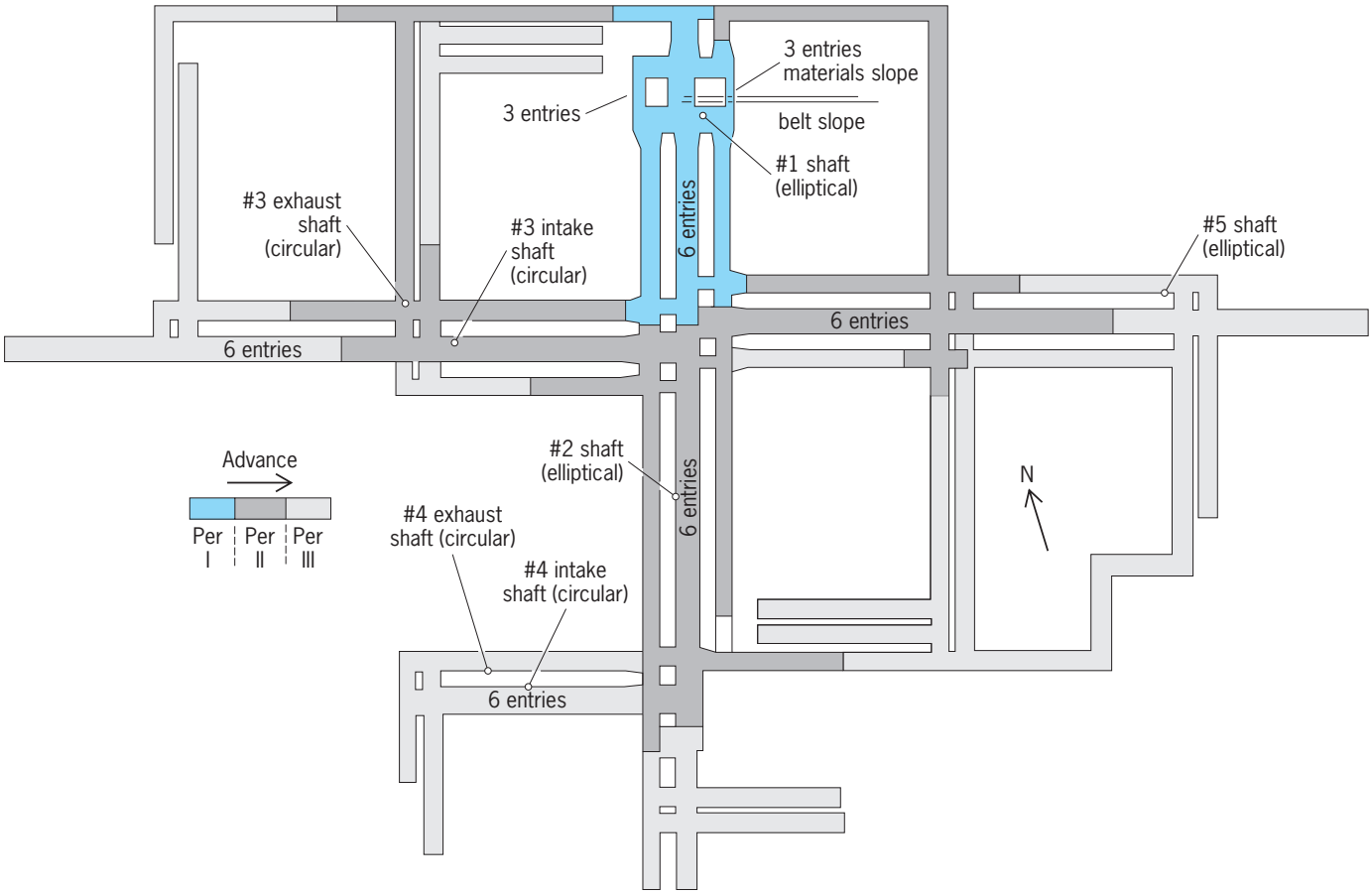

Fig. 13. Ventilation system in a southwestern Pennsylvania coal mine, showing the location of the shafts and slopes, and mine development.

United States. The location of the main fan at the surface is preferable. Still, the complexity of multilevel metal mine ventilation can inhibit the exclusive use of surface fans. Irrespective of coal or noncoal operations, the metric tons of air circulated per metric ton of mined product can vary from 3 to 20 depending on the severity of the air quality and temperature-humidity problems encountered.

Coal mine ventilation system. The ventilation system of a coal mine in the Pittsburgh seam in southwestern Pennsylvania is shown in **Fig. 13**. The seam is about 2 m thick, is fairly level, lies at a depth of 250 m, and is very gassy. The mine was designed for an annual production of 3 million metric tons from 13 continuous miner sections. The evolution of the mine to full production and the associated ventilation system is shown in three periods. In period I, there are five continuous miner sections that are ventilated using two slopes, and one dual compartment shaft that is divided by a partition running down the middle into an intake compartment and a return compartment. In period II, all the continuous miner units are engaged in driving the mains and developing the production panels. At this stage, the mine is ventilated by the two slopes, and four shafts, two of which have dual compartments. In period III, some of the continuous miner units are engaged in pillar extraction. Ventilation is provided by seven shafts and the two slopes. Three of the shafts have dual compartments.

During the mains development, three sections are driven (two outlying and one central), each having in its own belt, intakes, and returns. Entries are 2 m high, 4.5 m wide, and 20 m apart. A solid pillar of coal, about 25 m thick, is left between the sections. At the end of development, the central section, which has six entries, serves as the main intakes, with an isolated belt in one of the entries. The outlying sections, which have five entries each, serve as the main returns.

At peak operation, the mine circulated 1180 m^3/s (air) with a volumetric efficiency of 60% and connected fan power of 4500 kW. The intake mass airflow rate per metric ton of coal produced is over 10 tons. While the mine was initially planned for room-and-pillar mining, the longwall method was later introduced as the main production method. The annual mine production increased to about 6 million metric tons. However, the major aspects of the mine ventilation system did not change significantly.

Metal mine ventilation system. The Homestake gold mine located in the Black Hills of South Dakota is more than 120 years old and is one of the deepest and most extensively developed mine in North America. The ore body consists of six major ledges.

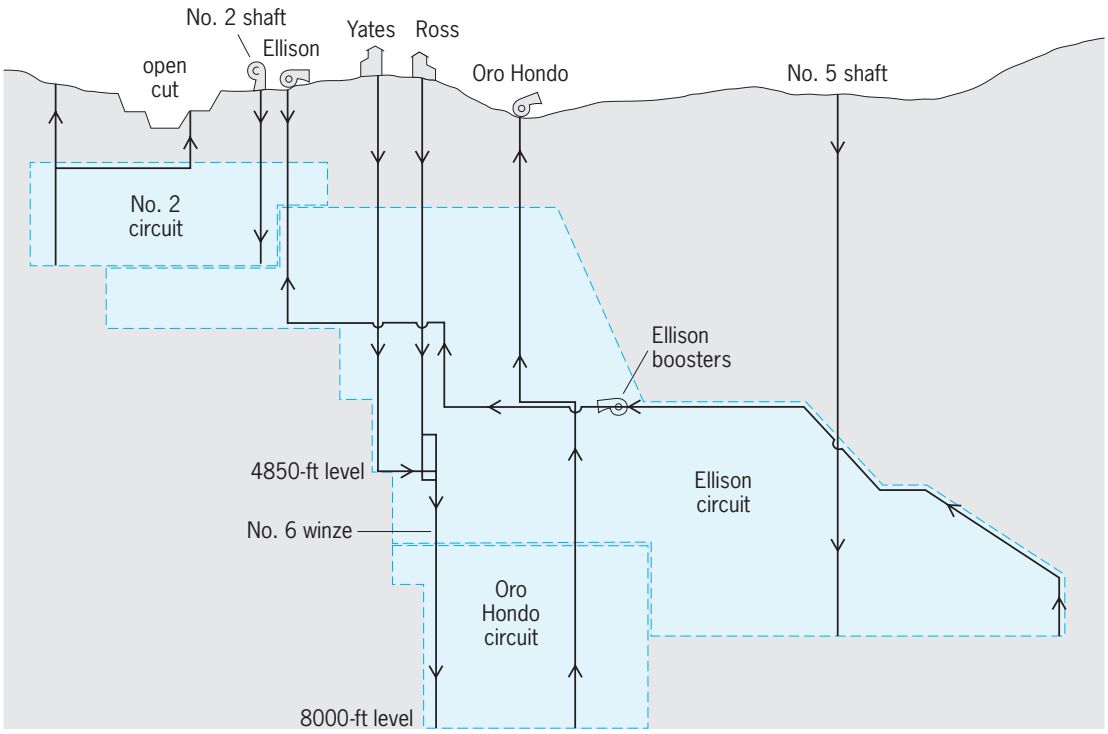

Fig. 14. Homestake Gold Mine, showing the three primary ventilation circuits and major airways.

The mine extends 2500 m below ground, annually producing 10 metric tons of gold from 1.5 million metric tons of ore. The predominant mining methods are mechanized cut-and-fill and vertical crater retreat. The geothermal gradient at Homestake is 2.19°C per 100 m in depth. The mean rock breaking depth is 1750 m, and the virgin rock temperature at this depth is 41°C. The designated reject temperature for the work areas is 29.4°C.

Homestake uses several cooling techniques—chilled water service, portable spot coolers, portable direct-contact spray coolers, refrigeration with water cooling coils, and refrigeration machines in a recirculation circuit. In the deepest mining section, the Virgin rock temperature is between 48 and 56°C. A controlled recirculation circuit, incorporating a refrigeration plant at 6950 ft (about 2000 m below surface).

The mine is ventilated by three relatively independent circuits (**Fig. 14**). For the upper part, No. 2 shaft is the intake. The fan is installed at the 1700 ft level, and exhaust is partly via raises that break into an open-cut mine, and partly through the Ellison shaft. The middepth part of the mine intakes via No. 5, Yates, and Ross shafts with exhaust up the Ellison shaft. Two underground booster fans at the 3500 ft level and two surface fans at the collar of the Ellison provide the motive force. For the deepest part of the mine (the Oro Hondo circuit), intake is via the Yates and Ross shafts, and through No. 6 winze. The exhaust is up the Oro Hondo shaft through an exhaust fan at the shaft collar.

The total intake air quantity is 391 m^3/s, and the total exhaust 445 m^3/s. The intake air mass per metric ton of ore production is 10.92 metric tons. In the recirculation circuit, the fresh intake air volume is 66 m^3/s, whereas the total air volume is 220 m^3/s. The name-plate diesel power used in the mine totals 6200 kW. The intake air volume per diesel kilowatt is 0.063 m^3/s. The total fan power is nearly 5500 kW, and the total air conditioning and refrigeration plant power is also 5500 kW. The mine has an extensive fire detection system, with carbon monoxide sensors at selected locations. The ventilation cost as a fraction of the total underground cost is estimated at 6.2%, and in terms of cost per underground ounce of gold it is $21.20.

Raja V. Ramani

Mine Illumination

Electric lighting of mine working areas may be developed by permanently mounted light sources, machine-mounted lights, and portable light sources such as the miner's battery-powered cap lamp. Different tasks are performed throughout the mine, and different levels of light are needed for the safe performance of the tasks. Typical locations for permanent light sources are repair and maintenance garages, pump rooms, fuel storage areas, crusher stations, and high back stopes. Underground mining machinery may be equipped with headlights and spotlights similar to those on surface vehicles. Portable light sources are used in a wide variety of locations.

The United States has the most stringent lighting standards of any country. Minimum standards for the illumination of the working places in underground mines have been established in the United States and are described in *Code of Federal Regulations*, Title 30 (CFR 30). Illumination system design and implementation must satisfy the requirements of the

miner, operator, machine manufacturer, illumination equipment manufacturer, and the Mine Safety and Health Administration (MSHA) in order to assure a system's acceptance. Each has similar but differing criteria for acceptance. The goal to be obtained is a system that is practical and maintainable and that will enhance the safe and efficient performance of the miner's tasks.

Standards. The standards specify that illumination shall be provided wherever personnel and mobile mining equipment are operating. This illumination is in addition to that provided by personal battery-powered cap lamps and must provide all surfaces in the miner's normal field of vision with a luminous intensity of at least 0.06 footlambert (0.2 candela/m^2). The area to be illuminated varies with the type of mining machinery. For continuous miners and loading machines, the face and ribs, roof, floor, and exposed surfaces of the mining equipment from the face to the outby end (away from the working face) of the machine's bumper is the minimum area to be illuminated. For self-loading haulage equipment such as a scoop, the face and ribs, roof, floor, and exposed surfaces of mining equipment that are between the face and a point 5 ft (1.5 m) outby of the machine is illuminated (**Fig. 15**). Cutting machines and drills have similar requirements. The area to be illuminated for longwall mining equipment consists of the face, the entire length of the self-advancing roof support system, the control station and head and tailpiece of the face conveyor, and the roof and floor for a distance of 5 ft (1.5 m) horizontally from the control station and head and tailpiece (**Fig. 16**). Other pieces of mining machinery have differing illumination area requirements.

All lighting fixtures must be permissible (an MSHA designation), either machine-mounted or stationary. A machine-mounted fixture is powered by the machine. Because of rough service conditions, these fixtures are required to remain effectively grounded when loosened or separated from the machines. Stationary lighting fixtures are attached to the mine roof, ribs, or roof-support equipment, such as chocks or shields on longwalls. Energized stationary lighting fixtures can be contacted by miners; thus, the entire system from power source to light must be designed to avoid hazards.

To increase the visibility for miners, the standards require that paint used on exterior surfaces of mining machines have a minimum reflectance of 30% and that red reflectors or tape be installed on each end of the mining machine. Each person going underground is required to wear an approved cap lamp or equivalent portable light and have a minimum of 6 in.^{2} (39 cm^2) of reflecting material on each side and back of the hard hat.

Lamps. Four basic types of lamps have been tested and are in use in underground mines: common incandescent, fluorescent, mercury-vapor, and high-pressure sodium-vapor. Each lamp has a system of ballasts and cables, and the entire system must gain permissibility before it can be used in an underground mine. The common incandescent system has been thoroughly tested in practice and is widely used; however, other systems are more efficient and can better withstand shocks. The fluorescent system is ruggedly constructed and instantly started without delicate filaments. Fluorescent lamps give off a soft white light with little glare and have three to nine times the life of incandescent lamps. Fluorescent systems accept common machine voltages and provide more lumens per watt than incandescent systems. Mercury-vapor systems are rugged, long-lived, and efficient, but require 3–5 min warmup time and 5–7 min hot restrike delay time. The mercury-vapor lamp produces a blue-white light and provides nearly three times the lumens per watt compared with an incandescent lamp. The high-pressure sodium-vapor lamp is the most efficient of all, providing five times the lumens per watt compared with an incandescent lamp. The sodium-vapor system goes to full

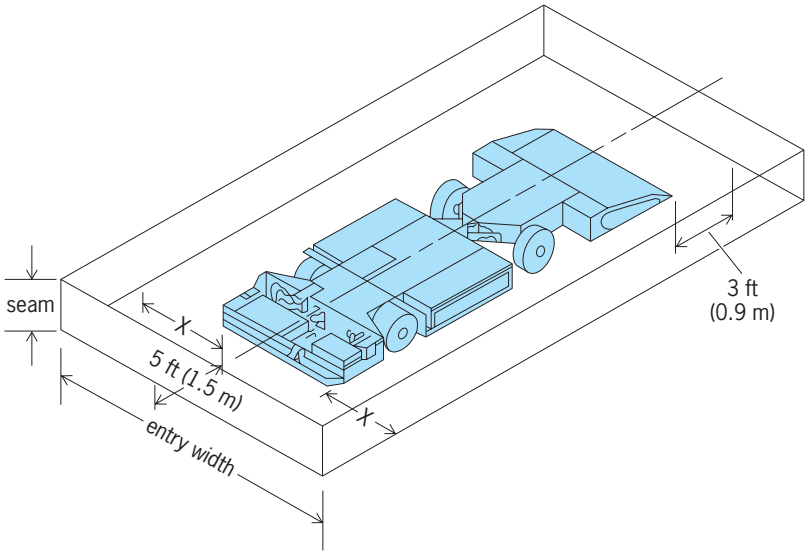

Fig. 15. Area required to be illuminated around a scoop used as a loading machine. Distance *X* is the distance from the side of the machine to rib or sidewall of the mine opening and is variable. (*After Proceedings of the 4th West Virginia University Conference on Coal Mine Electrotechnology, August 2–4, 1978*)

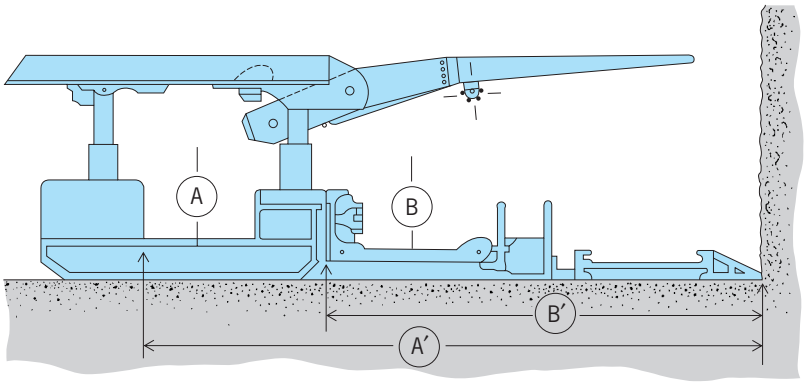

Fig. 16. Longwall illumination system showing area to be illuminated is from the face to the gob side of the travelway, which defines the path by which the miners operating the equipment move from one end of the longwall face to the other. Travelway A is between the hydraulic supports. For the travelway A, the area to be illuminated is A′. Travelway B is under the cantilever roof beam adjacent to the conveyor. For the travelway B, the area to be illuminated is B′. (*After Proceedings of the 4th West Virginia University Conference on Coal Mine Electrotechnology, August 2–4, 1978*)

brightness in 3 min and requires only 30 s for a hot restart. *See* FLUORESCENT LAMP; INCANDESCENT LAMP; MERCURY-VAPOR LAMP; SODIUM-VAPOR LAMP.

Field evaluation of illumination systems indicates that underground lights often dazzle the workers with glare. Louvers and other light diffusers are added to the system to correct this condition. *See* ILLUMINATION.

Malcolm T. Wane

Mine Drainage

Water that is encountered in underground and surface mines has an important effect on the cost and engineering design of many excavations. Its severity varies with local conditions, which relate to geological and hydrological factors, and can be aggravated by topography and climatic conditions. In extreme circumstances, water inflow into mine workings may necessitate the temporary suspension of operations during certain periods of the year. Many mines, even at considerable depth, experience minimum difficulties with water inflow and can be called dry mines.

The existence of water may present hazards in mining and may dictate the methodology of the mining process. Water sources can be planned, such as water delivered to sprays for dust suppression on mining machinery, or can be unplanned, such as water flow from regional rainfall and snowfall, underground aquifers, and abandoned mines.

The dramatic rescue at the Quecreek mine in Somerset County, Pennsylvania, in July 2002 underscores the hazards posed by flooded underground mines. Nine miners accidentally breached a barrier between Quecreek and an abandoned mine thought to contain over 50 million gallons (190,000 m^3) of water. After 77 hours of rescue efforts, including deployment of over a dozen pumps whose combined peak pumping rate exceeded 32,000 gallons/min (120 m^3/min), the trapped miners were brought to safety.

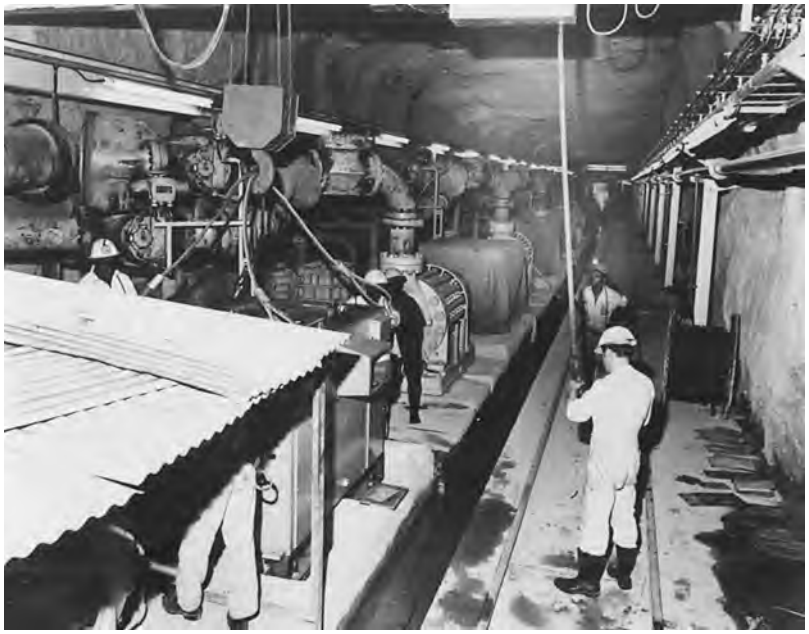

Fig. 17. Pumping installation at a copper mine in Zambia. (*Zambia Consolidated Copper Mines Ltd.*)

Losses from flooding can be devastating to production, life, and property. Therefore, an assessment of the mine drainage problem must be made early in the planning stage. Geologists and hydrologists are needed to identify the sources of water, the probable volume that will be handled, and the corrosive characteristics of the water. Mining engineers take that information and design the mine workings, and install the necessary pumps, discharge lines, and collection points to minimize the impact on the overall mine operations. *See* GROUND-WATER HYDROLOGY.

Pumping principles. Centrifugal pumps are the most commonly used pumps in mine drainage (**Fig. 17**). When a centrifugal pump is primed, or completely filled with water, its rotating impeller creates a partial vacuum. Atmospheric pressure forces more water into the suction side of the pump. Because centrifugal pumps merely create the vacuum that allows atmospheric pressure to push the water into the casing, they must be installed in proximity to the location from which the water will be drawn. Theoretically, if a perfect vacuum could be achieved, the maximum suction lift has a limiting value of about 34 ft (10 m) vertically at sea level, and depends directly on the atmospheric pressure. Since a perfect vacuum cannot be achieved, the height to which atmospheric pressure can push the water in suction is limited to about 22 ft (7 m) at sea level. Pump deficiencies reduce the theoretical height. If the pump is located below the water level, the intake is under positive pressure, and these limitations do not apply.

Centrifugal pumps may be directly connected to an electric or diesel motor and operate at 1750–3500 revolutions per minute. Single-stage pumps can pump about 5000 gallons/min (19 m^3/min) against a 500-ft (150-m) head. Multistage centrifugals, in which pressure is built up by feeding output of the pump successively into several single-stage units in series, can pump to heights over 5000 ft (1500 m). Because the high speeds of rotation invite erosion when gritty water is handled, provisions must be made either to clarify the water or to equip the pump with an erosion-resistant lining and impellers. *See* CENTRIFUGAL PUMP.

Submersible pumps may also be used, and they find their greatest application as well pumps in water supply and dewatering. Special submersible small-diameter motors drive them.

Underground mine drainage. The problems related to water seepage and handling in underground mines are usually more complex than for surface mines. Because operations are conducted at greater depth, the geology and hydrology associated with underground excavations are more varied, and the openings are often located below the regional water table. The water table, in its simplest form, is a nearly horizontal plane when viewed over short distances. However, where a succession of rock layers is present and their permeabilities differ greatly, the geologic cross section may reveal multiple water tables. Thus, it is possible for a deep mine to be completely dry,

operating in impervious rock, while under several water horizons, aquifers, lakes, and even the sea. In general, some water seepage into openings may be expected in most underground mines. Where water flows are large and persistent, an evaluation must be made to strike a balance between the costs of incrementally increased drainage against the benefit of mining in a drier environment. In addition, wet ore often plugs transfer openings and sticks in mine cars, hoist containers, storage bins, trucks, and stockpiles. The general objective is to ensure an adequate drainage scheme for efficient and safe mining, while keeping the operating cost within tolerable limits.

Underground mine drainage strategies consist of two parts: (1) drainage of the mine entrance (portal), which could be a horizontal adit, inclined slope, or vertical shaft, and (2) drainage of the mine itself.

Adits. It is usually cheaper to handle surface precipitation at ground level than to pump the water out of the mine after it has seeped underground. The expected surface inflow can be caught in adequately placed and designed ditches that divert it from the areas where the water enters the mine. A study of surface topography, surface rock lithology, and permeability, and a definition of the drainage basin contributing to the possible inflow are necessary to establish an optimal surface drainage scheme.

Shafts. Shafts, which provide access to the mineral to be extracted, sunk through water-bearing strata and aquifers, present special drainage problems. A number of methods may be applied during the shaft-sinking operation, depending on the thickness of the aquifer, its permeability characteristics and volume of water, the depth of the shaft, and the geologic rock column (**Fig. 18**). The shaft site may be drained through a local depression of the water table by the installation of a system of wells. Pumping is continued until the shaft is completed and made watertight.

An alternative solution, where practicable, is to grout the critical sections of the shaft either before or during shaft sinking. There are many types of grouts, such as chemical grouts or cement slurries. Chemical grouts are low-viscosity resins that are able to penetrate fine cracks and then cure. The grout is pumped through boreholes, specially sunk for this purpose, which form a ring around the shaft. The expectation is that the grout will find its way through cracks and fissures in the rock and effectively seal them off to water. In many cases, grouting is difficult to control, and grouting pressure must be kept within limits to preclude the creation of more fractures than are being sealed. *See* GROUT.

Fig. 18. Dewatering through a mine shaft at a copper mine in Zambia. (*Zambia Consolidated Copper Mines Ltd.*)

Under some circumstances, a decision may be made to freeze the area through which the shaft is to be sunk. Time-consuming and costly, this procedure involves drilling holes outside the perimeter of the proposed shaft and installing special pipes through which a refrigerant is pumped until the strata and rock are frozen to the depth desired. The shaft is then sunk through the frozen ground and sealed to water inflow. Thawing of the ground, which takes considerable time, follows.

Mine workings. Water that enters the mine workings is allowed to drain downgrade through ditches toward a collection and pumping point excavated in the rock called a sump. The sump must have sufficient volume to contain the water inflow over a period during which power is off or the pumping facilities are down for repair. Sumps must have provisions for separating sediments from the water. Multiple-level mines can have several sumps. Engineering considerations dictate that the haulage system in an underground mine be located toward the shaft. This is the main reason why sumps are located in proximity to the shaft, a location with easy access for pump inspection and maintenance.

Surface mining. Water inflow in open pits does not present as many difficulties as in underground mines, although inflow fluctuations tend to be greater. The effect of water on the stability of the pit slopes is an important factor. The locations in which equipment must operate and the adequacy of the haulage roads are greatly affected by the nature of the inflow and the fluctuations in its volume.

Surface mines that operate above the water table and in rock of low permeability contend with surface runoff during rainfall and snowmelt only. The pit layout allows the water to be gathered near the pit bottom, where pumps dispose of it through pipelines with comparatively low pumping heads. This scheme, which is simple and low in cost, places the pumping installation in proximity to the mining operation, making it vulnerable to damage from blasting. The water collection site must be moved as mining progresses.

For open pits that are subjected to large inflows, especially those located below the regional water table, and to constant percolation of water into the pits, consideration must be given to lowering the water table around the mine excavation by means of wells, thereby reducing or eliminating the need to use pumping equipment in the pit. The wells may be located in the pit and/or outside the perimeter of the excavation. If wells are located outside the pit, water can be pumped without interruption during blasting. If the pit is cleared of pumping equipment,

pipes, and power lines, ease of maintenance is assured. In addition, the rate of pumping tends to be more uniform, the water is likely to be clearer (because of filtration through the strata), and the walls of the pit tend to be more stable.

Pit stability and the slope at which mining is done affect the overall cost and, in most cases, are directly related to the water residing in the rock being mined. The presence of water also has a deleterious effect on the haulage roads, resulting in increased operating costs. However, pumping through wells is expensive, and good engineering evaluation and planning must precede the decision to resort to such a dewatering scheme. Christopher J. Bise; Stefan Boshkov

Mine Safety

In the United States, mining operations include underground and surface mines, quarries, sand and gravel pits, and preparation plants. Although accidents in such facilities have been greatly reduced over the years, mining remains a dangerous occupation. During the period 1988–1997, more than 500 lives were lost in the United States coal mining industry and another 500 in metal and nonmetal mines. It may be assumed that 2000 worker-hours represents full-time employment of one miner for 1 year. On this assumption, the average coal mine employee working for one year during the period 1988–1997 had about 1 chance in 10 of experiencing a reportable injury; in the metal and nonmetal industry the corresponding odds were about 1 in 16.

Though they still occur, major mine disasters are no longer common. Most fatal mine accidents involve only one victim. The three most frequent types include falls of ground, powered haulage accidents, and machinery accidents. Less common but significant causes of mine fatalities include electricity; hoisting; falling, rolling or sliding material; and slips or falls of persons. The keys to preventing many fatal accidents are following safe work procedures, regular maintenance, thorough cleanup, and use of personal protective equipment—for example, safety belts and lines where there is danger of falling.

In nonfatal accidents, the parts of the body most frequently injured are the back, fingers, and knee, in that order. The most frequent causes are handling materials (for example, lifting) and slips and falls of persons. The majority of nonfatal accidents can be prevented by such measures as planning the job to fit the human factors, following safe work procedures, personal protective equipment (for example, eye shields), machine guarding, and cleanup.

Safety programs. Safety records of individual mines vary. The safest mines in the United States operate for hundreds of thousands of hours without a serious injury. Several common factors seem to characterize those mine operations that are most successful in preventing accidents. One is commitment to safety by top-level managers, who hold supervisors accountable for safety as well as production. At most mines, management has a formal safety policy and requires that safety rules be followed. Many mines employ a safety director who is likely to be a qualified professional. Many companies have safety departments with technicians and industrial hygienists or other health and safety specialists. The safety function may be linked with loss control or training. Mines with fewer resources may receive assistance with safety and training programs from state or federal programs. Another common factor in safe mines is a cooperative relationship between management and labor. A critical element is thorough, effective, and continued training for all mine employees.

Effective mine safety programs have been found compatible with high productivity. A study published by the National Academy of Sciences (1982) reported that in a survey of underground coal mines the more productive mines also tended to be safer mines. In the 1990s, safety records in United States coal mines continued to improve as coal production reached the highest figures in history.

Under the law, safety in United States mines is primarily the mine operator's responsibility, with the assistance of the miners. Labor also takes a strong interest in safety, often cooperating with management in such areas as training and safety inspections. Safety provisions are included in some union contracts. Many states with extensive mining have state mine agencies whose activities may include training, consultation, rescue, inspections, and investigations. In the federal government, mine safety regulations are administered by the Labor Department's Mine Safety and Health Administration. The National Institute for Occupational Safety and Health conducts research into mine safety and health topics. Others in the mining community who influence safety include equipment manufacturers, educators, safety organizations, industry associations, contractors, and consultants. *See* INDUSTRIAL HEALTH AND SAFETY.

Major safety hazards. These include explosion, fire, toxic or asphyxiating atmospheres, falls of ground, unguarded moving machine parts, operating heavy equipment, and electricity.

Explosion. Methane gas seeps into the air in all underground coal mines and in some metal and nonmetal mines as well, presenting the hazard of a disastrous explosion. Unexpected outbursts may release large volumes of methane. Some underground coal mines liberate over 10^6 ft^3 (28,000 m^3) of methane in 24 h. Methane is explosive at a 5–15% concentration in air; any spark, electric arc, or open flame may set off an explosion. Coal dust suspended in the air increases the explosiveness of a methane-air mixture. It can also be explosive in the absence of methane and may propagate an explosion for hundreds of feet. Winter is the most frequent period for coal mine explosions, due to dry air and air pressure drops preceding storms.

Prevention measures include sound mine design, ventilation, frequent tests with a methanometer, prohibition of smoking and open flames, use of explosion-proof electrical or diesel equipment, cleanup of coal dust accumulations, and blanketing the mine floor and walls with an incombustible rock dust such as powdered limestone.

Fire. Extensive underground fires may be difficult to extinguish, especially in mines where the mineral itself fuels combustion, such as coal operations. In such cases, the mine may have to be sealed to exclude oxygen until the fire can burn itself out. Water or chemical fire extinguishers may be injected through boreholes. On the surface, fires in buildings or on large mining equipment may also endanger life. *See* FIRE EXTINGUISHER.

Causes of fire include faulty electrical equipment, belt conveyor friction, smoking, welding, explosives, and spontaneous combustion. Cleanup of coal dust and refuse is important in preventing fires and discouraging their spread. Flammable materials should be stored in closed containers, and fire-resistant materials should be used for ventilation curtains and conveyor belting. Fire extinguishers are required in buildings, in shops, and on mobile mining equipment. Automatic systems to detect, and in some cases automatically extinguish, fires are used in many underground coal mines, especially along belt lines, and on some large surface equipment. *See* AUTOMATIC SPRINKLER SYSTEM; FIRE DETECTOR.

Toxic or asphyxiating atmospheres. In fires and explosions, many miners have succumbed to smoke inhalation, carbon monoxide, or lack of oxygen. Such atmospheres also may develop in abandoned mines or unventilated areas of active mines. Other sources of this hazard include blasting and the use of internal combustion engines with inadequate ventilation. Carbon monoxide and oxygen-deficient atmospheres are the most frequent of these hazards, but nitric oxide, nitrogen dioxide, sulfur dioxide, and hydrogen sulfide are formed in some mines. Portable and stationary detectors are available to monitor for the presence of these gases.

Adequate ventilation is the key to preventing accumulations of dangerous gases. Miners should not enter unventilated areas and should allow enough time after blasting for gases to disperse. Unexpectedly breaking into a nearby, abandoned mine has caused deaths due to asphyxiation. These accidents—as well as underground mine flooding—can be prevented by accurate mapping and by longhole or other exploratory drilling in advance of mining.

In emergencies, protection from toxic or suffocating atmospheres is provided to underground coal miners by self-contained self-rescuers, breathing devices that provide an hour's oxygen for escape. Underground metal and nonmetal miners are required to carry self-rescuers that filter carbon monoxide from the air. Mine rescue teams, which may have to enter areas with hazardous gases to locate missing miners, must use an approved self-contained breathing apparatus.

Falls of ground. Historically, these have been the single most frequent cause of death in mining. In underground mines, they are falls of the mine roof, face, or ribs (walls). In surface mines, they are falls of the highwall.

Ensuring roof stability can be a difficult task, especially in coal mines where thin, fragile shale strata often overlie the coal seam. Geologic faults, fossil inclusions, and sandstone channels add to roof instability. Steel roof bolts are generally used to pin roof strata to more solid rock above. Bolts, which may be over 10 ft (3 m) long, are inserted by machine into drill holes in the mine roof, and held in place by an expansion shell, a resin grout, or a combination of methods. They may be supplemented with timbers, beams, steel arches, and lagging (planking laid over timbers or arches to prevent small rocks from dropping between them), depending on the roof conditions. Roof control plans should be carefully formulated to match the geological and mining conditions. Loose rock should be taken down before the roof is supported, and the roof's condition should be monitored.

Many roof fall deaths have occurred when miners go under unsupported roofs. Hydraulic roof supports, attached to roof bolting machines, or along the mining face in longwall operations, are used in many underground coal mines. Canopies on mining machines and shuttle cars help protect the operators of mobile equipment from falling rock. Remote-controlled mining machines can allow miners to stay under a supported roof while the machine advances.

In some underground metal and nonmetal mines, the rock requires little or no support; in others the ground may be very loose and extensive ground support is needed. Roof bolts of various types, posts, beams, arches, lagging, and wire mesh are among the support methods employed. In surface mines, rock fall accidents can be prevented by maintaining the highwall at a safe height and angle, stripping off loose material, and minimizing the need for miners to work at the foot.

Unguarded moving machine parts and heavy equipment operation. These are the most frequent causes of death at surface mines and preparation plants. Victims may be struck, crushed, or pulled into machinery. An effective maintenance program is important in guarding against these types of accidents, which may be caused by equipment malfunction or human error. The operation should be planned to minimize these risks; guards, railings, or barriers should protect workers from contact with machinery in motion; good communications should be maintained; and repairs should never be done on equipment in motion.

Electricity. Modern mines have complex electrical systems. Accidents may occur through malfunction or human error. Many electrical accidents occur when personnel knowingly or unknowingly work on energized electrical equipment. Others occur, for example, if high-profile equipment such as a crane contacts an overhead power line. Only qualified persons are allowed to install, maintain, and repair electrical equipment.

Major health hazards. Both mining and mineral processing present health hazards, of which lung diseases are the most important.

Coal mine dust can lead to coal workers' pneumoconiosis (black lung disease), which has killed or disabled thousands of miners. Another lung disease, silicosis, is caused by exposure to crystalline silica

(quartz) dust, which can be present in coal mine strata as well as in metal and nonmetal mines. Cases of silicosis have been identified at silica flour plants, especially in bagging operations, and at coal mines. In underground uranium mines and some nonuranium mines, radon daughters are found. These radioactive decay products of radon gas may cause lung cancer. Lung disease may also be caused by asbestos in the rock or in industrial use, and by large concentrations of other nuisance dusts. *See* ASBESTOS; RADON; RESPIRATORY SYSTEM DISORDERS.

The federal government has set standards limiting exposure to these respiratory hazards. Standards have also been proposed to control diesel particulate, another potential respiratory hazard. The basic method of controlling all these hazards is ventilation. Water sprays on drills, mining machines, and other equipment help to reduce dust. Cleanup of dust accumulations is important, and vacuum-type dust collectors are used in some operations. However, dust control remains a problem in some operations—notably in longwall mining and in the silica flour plants. When miners cannot be adequately protected by engineering controls, respirators are required. Research into better dust control methods is continuing. *See* RESPIRATOR.

A problem in many mines and plants is noise, which can cause hearing loss. Federal regulations limit noise exposure to a maximum of 90 dBA (decibels measured on the A scale of a sound-level meter) over an 8-h shift. Major sources of excessive noise include pneumatic drills, channel burners in dimension stone operations, continuous mining machines, muckers, load-haul-dump machines, bulldozers, draglines, front-end loaders, scrapers, graders, and trucks. At preparation plants, shakeouts used to load railroad cars are the worst sources, generating up to 118 dBA; other noisy equipment includes crushers, sorting screens, chutes, rod-and-ball mills, and kilns. Achieving adequate noise control has not been easy, and a machine-by-machine approach has sometimes been necessary. Compliance also may be achieved by isolating either machinery or operator in a noise-reducing enclosure or by limiting the worker's time in the noisy area. Hearing protection is used when engineering techniques are not adequate. *See* ACOUSTIC NOISE; EAR PROTECTORS.

Another health hazard is heat stress, a danger in some open-pit mines during the summer; prevention measures include acclimatization, drinking liquids, and rest periods. An additional hazard exists in the form of toxic chemicals, such as cyanide and mercury, which are used in some mining and milling processes or contained in the ore.

Training. Inexperienced miners have accounted for a large proportion of fatal accidents. Training and retraining in specific safety and health topics have been required for new and experienced miners by law since 1978. Federal law provides that new underground miners must receive 40 h of training, and new surface miners 24 h. Eight hours' annual refresher training is required for all experienced miners, along with new-task training when changing jobs. Some states have additional training regulations, and some operators' training programs go beyond the legal requirements.

In addition, mine personnel must be qualified or certified to perform critical tasks such as electrical work, hoisting, gas tests, respirable coal mine dust sampling, and work on permissible diesel equipment in coal mines. Most coal mining states have certification requirements for coal mine forepersons and other supervisors, which usually include several years' experience, specialized training, and lengthy examinations.

Specialized training also is required for mine rescue teams. Each underground mine is required to have two such teams on call in case of emergency; federal regulations specify their qualifications, which include 20 h of initial instruction and 40 more hours of training and practice annually. To sharpen the teams' skills, mine rescue contests are held. Local and regional contests are usually sponsored by industry groups, assisted by federal and state agencies; government-sponsored national contests are held for coal and metal-nonmetal teams in alternate years.

Various types of mine safety training are supplied by a number of sources: mining companies, private firms, state agencies, universities, community colleges, vocational-technical schools, and the National Mine Health and Safety Academy.

Legislation. In the United States, legal responsibility for mine safety and health rests primarily with mine operators. The first federal mine safety legislation, in 1910, founded the Bureau of Mines, whose original role was strictly advisory. During World War II, the federal government took over operation of the coal mines during a strike and issued its first code of safety standards. Enforcement powers came in 1952, when Congress provided for closure of underground coal mines employing 15 or more in cases of imminent danger or continued noncompliance with standards. Successive laws expanded these powers; the Federal Mine Safety and Health Act of 1977 provides for enforcement of safety and health standards at virtually all coal, metal and nonmetal mines, quarries, sand and gravel pits, and mineral preparation plants.

Most important mining states also have safety laws, some of which predate federal involvement. Some states, such as Pennsylvania, Kentucky, Arizona, Nevada, and West Virginia, maintain their own inspection program. Others, such as North Carolina, concentrate on training and consultation, leaving most inspections to the federal government.

Federal mine safety and health rules, issued and enforced by MSHA, are published in Title 30 of the *Code of Federal Regulations*. In most cases, separate rules are provided for different types of operations. In issuing new or revised rules, MSHA encourages industry, labor, and others to provide input through written comments, public conferences, and hearings.

The law requires four complete federal inspections of each underground mine per year and two inspections of each surface mine. Additional inspections are to check for correction of violations found on a regular inspection, make spot checks at mines with excessive methane or other special hazards, or focus on specific areas such as electricity or dust

control. The law provides for a representative of the miners to accompany federal inspectors on each regular inspection.

A federal inspector who finds a violation of safety or health standards must issue a citation, which entails a fine that may range from $55 to $55,000, depending upon seriousness and other factors. If violations remain uncorrected or there is imminent danger, the mine or section may be closed. Criminal penalties are possible for knowing and willful violations.

MSHA also investigates fatal accidents, publishes current safety and health information, and supplements enforcement with a wide variety of programs to provide penalty-free help on engineering problems and training and safety programs.

J. Davitt McAteer; Katharine Snyder

Fig. 19. Hopper cars being loaded by a belt conveyor. (*Burlington Northern, Inc.*)

Transportation and Storage

Transportation and storage in any mine, surface or underground, involves the safe, efficient movement of personnel and material at the least possible cost. Total cost includes the cost of lost-time accidents, the expenditure of time and energy, and the effect on the environment. Material includes ore, waste, mining equipment, powder, steel, track, timber, cement, maintenance and plumbing supplies, and other necessities for an operating mine.

Transportation. Loading, hauling, and unloading of ore and waste may involve transfer in stopes, chutes, or ore passes to intermediate levels; haulage to shaft pockets; skip loading or caging of ore cars; and hoisting, haulage, and feeder transfer to storage or active mill operations.

Gravity is employed wherever possible to move the mined rock: from the working face to haulage levels; out of skip pockets, ore bins, or stockpiles; and from one conveyor to another.

Handling of ore and waste at the shaft station and surface is governed by several factors: the underground transport of the ore; the types and sizes of the cars; the inclination of the shaft; the method of hoisting; the tonnage and physical characteristics of the ore; the number of ore classes that must be kept separate during transport; and the surface topography of the site.

The type and size of equipment employed depend upon the different mining conditions and the required results. An equally important factor is the operational cost, as dictated by the character of the material to be moved, the daily tonnage, the methods of loading and dumping, and the distances involved.

In smaller mines, haulage may be accomplished by hand-tramming or animals. Mechanical haulage can be by rail (**Fig. 19**), diesel vehicle (**Fig. 20**), scraper or slusher, or rope, cable, or belt conveyor. Compressed-air machines have an obvious advantage in mines that need ventilation. Haulage by rail is the primary method used in large-scale mines, both surface and underground (**Table 1**).

Conveyors. If there is sufficient dump room or storage capacity near the mine, a system of belt conveyors can handle material at high rates and relatively low costs, but only if proper feed control of a sized material allows a continuous, even flow that matches the system design. Other factors that determine the practicality and size of a conveyor system are the rate at which the material must be handled, the material's density and stickiness, the dusting or degradation on transfer, and the need for the system to handle more than one product.

Slurry transportation. Pipelines have been successfully used to transport many different solid materials: coal, phosphates, limestone, fly ash, copper concentrates, gilsonite, mill tailings, slags, and sand fills. The dry material is first pulped to form a slurry and then pumped to the destination, where it is dewatered. Additional examples of hydraulic movement of material are the loading and unloading of finely divided ore into ore carriers such as ships or sealed cars, the dredging and placing of fill, and the emplacement of backfill in mine stopes.

One notable slurry pipeline completed in 1983 is a 237-mi (382-km) pipeline in northern Mexico, which carries iron ore concentrates from two separate mines to the pelletizing plant. *See* PIPELINE.

Miscellaneous. Rope or cable haulage, particularly aerial trams, are employed to move ore carriers over very rough terrain or along grades that are too steep for locomotives and trucks. The cost of transporting

Fig. 20. A 16-ton (14.5-metric-ton) diesel locomotive with exhaust water scrubber for use in coal mines. (*Brookville Locomotive Division of Pennbro*)

TABLE 1. Summary of haulage systems*

Conditions		Equipment†												
			Scrapers					Truck						
		Bulldozers	Tractor-drawn scraper	Underpowered, rubber-tired scraper	Full-powered single-engine scraper	All-wheel-drive scraper	Rubber-tired tractor with trailer scraper	Rear dump	Semitrailer rear dump	Semitrailer bottom dump	Train	Conveyor	Skip	Pipeline
Material	Rough, blocky	1						1	1		1		1	
	Max. 36 in. (91 cm)	1						1	1		1	4	1	
	Max. 24 in. (61 cm)	1	1	2	2	1	1	1	1	2	1	3	1	
	Fines	1	1	1	1	1	1	1	1	1	1	1	1	4
Length of haul	0–300 ft (0–90 m)	1	1	2	3	3	3	1	3	3		4	4	4
	300–500 ft (90–150 m)	2	1	1	2	2	3	1	2	2		4	4	4
	500–1000 ft (150–300 m)		2	1	1	1	2	1	1	1		4	4	4
	1000–1500 ft (300–450 m)		3	2	1	1	1	1	1	1		4	4	4
	1500–5000 ft (450–1500 m)				1	1	1	1	1	1		3	4	4
	5,000–10,000 ft (1500–3000 m)				3	3	2	1	1	1	2	2		3
	10,000–15,000 ft (3000–4500 m)						3	1	1	1	1	1		3
	Over 15,000 ft (4500 m)							2	2	2	1	1		2
Ground conditions	Good	1	1	1	1	1	1	1	1	1	1	1	1	4
	Wet, soft	1	1			3						3		2
Maximum adverse grade	+ 3%	1	1	1	1	1	1	1	1	1	1	1	4	4
	+ 5%	1	1	3	2	1	2	1	2	2	2	1	4	4
	+ 10%	1	1		3	2	3	1	3	3		1	3	4
	+ 15%	1	1		3	3		1				1	2	4
	+ 20%											1	1	4
	+ 20%											4	1	4
Flexibility under varied conditions	Good	1	1	1	1	1	1	1	1	1				
	Fair	1	1	1	1	1	1	1	1	1	2	3		4
	Poor	1	1	1	1	1	1	1	1	1	1	1	1	4
Daily production rate	Low	1	1	1	1	1	1	1	1	1			3	4
	Medium		3	3	2	1	1	1	1	1	2	2	2	4
	High				2	1	1	1	1	1	1	1	2	4
Total tonnage	Small	1	1	1	1	1	1	1	1	1				
	Medium		3	3	1	1	1	1	1	1	3	2	3	4
	Large				1	1	1	1	1	1	1	1	1	4

*After Society of Mining Engineers of AIME, *SME Mining Engineering Handbook*, vol. 2, pp. 17–23, 1973.
†1. Should be considered. 2. May be considered. 3. May be considered under certain conditions. 4. May be considered, special situation.

material over very rough terrain is essentially confined to the initial cost of the installation.

A scraper is ordinarily employed as a mucking and loading machine, but it can also be used as a conveyor to transport ore and waste hundreds of feet when large quantities of ore must be moved comparatively short distances. Scrapers include bulldozers and scraper-loaders. When dragging is done at higher speeds, scrapers require a smooth base or bottom.

Economics. In any earth-moving operation, more profit may be earned or lost by equipment scheduling than by any other single facet of the project. Ideally, every item of the equipment fleet must be continuously working.

If mechanical haulage can be utilized full-time, the savings in operating costs can be substantial. Idle machines pay no dividends. Open-pit mines in which haulage is by rail should have large ore reserves. The construction of truck and handling facilities for effective operation and maintenance of trains is expensive, however, as is the acquisition of locomotives and cars.

A U.S. Bureau of Mines report has estimated that handling and transportation of ore and waste has accounted for one-fourth of total mining costs.

Proper handling of waste rock can also be economically important. Waste rock is usually disposed of as cheaply and conveniently as possible. This rock, however, could be a submarginal, low-grade ore. Consideration should therefore be given to possible future processing, when improved economics or technical advances allow profitable handling. Many old mine dumps and mill-tailing piles, originally regarded as worthless, have produced valuable products in later years.

Storage. Stockpiles or intermediate storage bins are used to accumulate the product when carrier schedules cannot be diverted to useful development

Fig. 21. A 72-in. (22-m) conveyor belt at the Caridad Mine, Nacozari, Sonora, Mexico.

work or until a means for further transfer of the material becomes available.

It is customary to store large tonnages of crushed ore in piles on the ground because storage in elevated bins is costly. Depending upon the weather and type of ore (for example, dusty or soluble), storage areas may be covered. Each installation is a study in itself and can affect the entire mine operation.

Ore can be reclaimed from a stock-pile by draglines, scrapers or dozers, bucket-wheel excavators combined with stacker and tripper conveyors or mechanical feeders; or from controlled draw points onto conveyors in a tunnel below the stockpile. Again, each installation varies with the type of mine.

A traveling gantry stacker is frequently employed to spread the ore in the manner that is most advantageous for future use. For example, the ore might be carefully bedded for blending, leaching, or expediting reclamation or reloading.

Technological advances. Electronic identification of haulage equipment and amounts of material carried has been under development. In a computerized truck-dispatching system, unit and product data are collected and transmitted without reliance on human operators.

Paved haulage roads are safer and more economical than unpaved roads. Trolley-assisted power systems have come into use in the steep graded areas.

Modern technology has developed multimile conveyors that follow terrain undulations and allow major savings in the civil engineering and earth moving required in construction of mining and milling operations. Regenerative conveyors or declined conveyors can produce, rather than consume, power; one example is the primary ore conveyor belt at La Caridad, Nacozari, Sonora, Mexico (**Fig. 21**), which carries 79,000 tons (72,000 metric tons) per day a distance of 1070 ft (960 m). Under normal conditions, it can generate 1.2 MW of electrical energy per day because of the downhill character of the installation.

Trucks have continued to grow in capacity, efficiency, and operational adaptability.

The planning, packaging, storing, and timing required for the safe, efficient transportation of personnel, equipment, materials, and supplies will always be as essential to a successful mining operation as the movement of ore and waste. David D. Rabb

Field Sampling and Ore Estimation

Sampling is a critical activity in the evaluation, development, and mining of a mineral deposit. The results of sampling are used in defining ore bodies, estimating ore reserves, mine planning, and controlling the quality of ore delivered from the mine to treatment plants. The information provided by sampling is organized into computer databases for geographic information systems (GIS) and ore estimation. *See* GEOGRAPHIC INFORMATION SYSTEMS.

Sampling. A sample is a small portion of material from a mineral deposit, and an accurate and unbiased collection of samples is intended to represent the entire deposit or a specified portion of the deposit.

The quality most commonly determined by sampling and by the laboratory testing or assaying of samples is the content of one or more particular elements or compounds. Amounts of impurities such as phosphorus in iron ore and ratios between certain metals such as chromium and iron in chrome

ore are also determined. Material with a content or grade sufficient to be economically minable is designated as ore. The grade of nonferrous metal ore is generally stated in percent of one or more contained metals such as copper, lead, or zinc. The grade of certain materials such as titanium ore, molybdenum ore, and potash is stated in percent of an index compound: titanium dioxide (TiO_2), molybdenum disulfide (MoS_2), or potassium oxide (K_2O). The grade of precious metal ore and of the precious metal content in other ores is stated in troy ounces per short ton or grams per metric ton of material. The grade in placer deposits is expressed in grams of the recoverable metal (gold, tin, titanium) per cubic meter of material.

Additional qualities determined by sampling a mineral deposit and by laboratory analysis deal with mineralogy and with metallurgical characteristics. An ore may consist of several minerals that contain the desired element but have different recovery characteristics, and an ore may have a composition or texture that affects its amenability to processing. Samples of industrial mineral deposits such as clay, limestone, and sand are tested for desired physical properties as well as for mineralogical components.

The most common methods of taking samples in a mineral deposit involve channel sampling, grab sampling, and bulk sampling in exposed ore and drilling holes into hidden ore.

Channel sampling. In vein and stratiform deposits, the basic procedure is to cut or chisel a linear slot or channel at a uniform width and depth across a fresh and cleaned exposure of the mineralization (**Fig. 22**). The channel is most commonly cut in the face or across the back of a mine working, with the broken material collected in sacks or on a plastic sheet. Where the mineralization consists of distinctive zones, a separate sample is taken from each zone so that the assay results can be composited in various ways to permit selective mining. Where the wall rock on each side of the recognized mineralization may also contain some values and may have to be incorporated as dilution in mining, additional samples are taken. Chip-channel sampling is similar to a more precise channel sampling but is less time-consuming. Chip channels consist of fragments broken in uniform amounts from a broad band across the exposed face or back. Channels are repeated at appropriate intervals along the ore zone.

Grab sampling. This procedure is used more commonly in massive and disseminated deposits where the face, back, or walls of a mine working are entirely in ore mineralization. The sample may consist of fragments chipped at uniform intervals from an exposure (chip samples), material scaled from an exposure (panel samples), a randomly chosen collection of fragments from a freshly blasted mine face, or amounts of material taken in succession from a trench. Samples taken by hand at random from cars and from conveyor belts are also referred to as grab samples.

Bulk sampling. In this procedure, large amounts of material are taken, primarily for the evaluation of dispersed mineralization and for metallurgical testing. In metallurgical pilot plant testing, the entire ore content of an underground mine working or of a surface trench, amounting to hundreds or thousands of tons, may be collected.

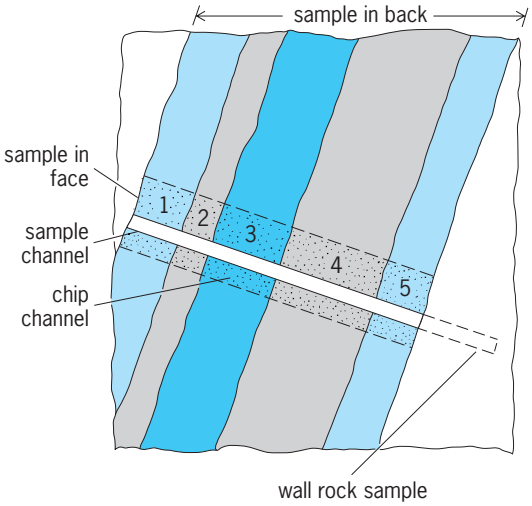

Fig. 22. Sampling channel and chip channel in a zoned vein exposed at a mine face, with a sample from each zone.

Drill-hole sampling. Drill-core and rotary or percussion drill-hole cuttings furnish most of the samples used in delineating ore bodies, evaluating prospects, and extending ore reserves beyond existing mine workings. In addition, cuttings from blastholes are often sampled and studied or quickly assayed for grade control during mining. After geologic examination, sections of drill core are commonly split lengthwise, with half of the core sent to the laboratory for assay and half kept for record. Drill-hole cuttings—chips of material loosened by the drill-hole bit—are examined in sequence at the drilling site and taken in their entirety or in part as samples. *See* DRILLING, GEOTECHNICAL.

Sampling patterns. In defining ore bodies and calculating reserves, the choice of a sampling method and a sampling pattern are dictated by the accessibility, nature, and size of the mineral deposit. The outcropping portion of a deposit and the portion exposed in underground workings will permit channel, grab, and bulk sampling. Hidden deposits and ore projections from exposed deposits permit sampling by drill holes, either from surface sites or from stations in existing underground workings. An exposed vein-type or tabular stratiform deposit with well-defined walls is suited to channel sampling, with preliminary channels taken at wide intervals and subsequent channels taken for detail at intervals of a few meters. The same kind of deposit may also be sampled in successive "fences" of drill holes across the trend; the holes may be drilled from a few sites in fanlike groups designed to cut the ore body at appropriate intervals (**Fig. 23**). Large massive and disseminated deposits are commonly sampled in a systematic grid of drill holes that pass through the ore at intervals of 100–400 ft (30–120 m) and by bulk sampling in underground exploration workings. Placer deposits are bulk-sampled in trenches across the trend, in pits or

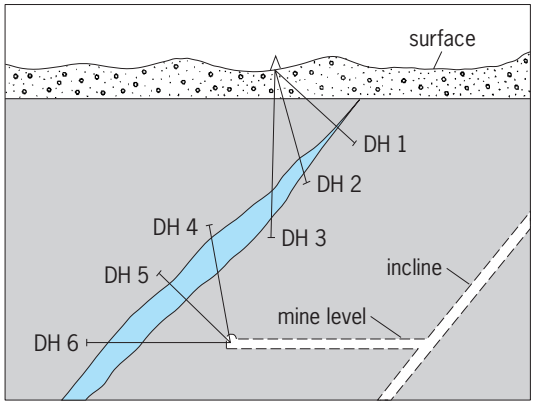

Fig. 23. Cross section of vein sampled by drill holes (DH) from surface and underground sites.

shafts at regular intervals, or by large-diameter noncore drill holes.

Grade control. Sampling for grade control, selectivity, and telemining automation during mining makes use of grab samples, blasthole samples, and samples of broken ore in transit. Broken ore in mine cars or on conveyor belts is commonly sampled by automated and robotic mechanical devices. The grade of uranium and potash mineralization in broken material and in blastholes is commonly estimated on site by the use of gamma-ray spectrometer probes. The grade of silver, lead, zinc, and copper mineralization in broken ore, mine faces, and drill holes can be estimated by the use of x-ray fluorescence scanners or probes. Similar instruments have been used in estimating the in-place grade of gold, nickel, and other metallic ores. *See* X-RAY FLUORESCENCE ANALYSIS.

Sample treatment. Samples must be large enough to be representative of their location in a mineral deposit. In a relatively uniform and fine-grained ore, an appropriate channel sample may weigh 5 lb (2.3 kg); in an irregularly mineralized and coarse-grained ore, a sample on the order of 10–20 lb (4.5–9 kg) may be needed. The sample must subsequently be reduced in amount for laboratory analysis; this is done in stages during which the ore fragments are crushed to successively finer size and thoroughly mixed. In the laboratory, a small portion of finely ground and homogenized sample is taken for instrumental, chemical, or fire assay.

Assaying. An assay (quantitative analysis) of the constituents in a mineral material is made through various combinations of laboratory procedures. The procedures incorporate classic methods of gravimetric and volumetric chemical analysis, while a variety of instrumental methods are used with the requisite accuracy and precision. *See* ANALYTICAL CHEMISTRY.

Fire assaying, the historic and current standard as a gravimetric method, is unique to mining. This method is used for determining the content of gold, silver, and platinum metals in ores. The steps in making a fire assay involve a furnace fusion of the sample with a lead flux, a recovery of the precious metals into a metallic bead by further heating (cupellation), and a weighing of the bead. The precious metals are then identified and quantified by further chemical separation or by instrumental analytical techniques.

Sample and ore calculations. In assigning a composite assay value to the samples taken within a designated channel or interval of drill core, the length of each sample is accounted for in calculating a weighted average (**Table 2**).

In determining the average grade of a deposit or of a zone within a deposit, values at point, channel, and bulk sample locations are projected to reference levels and attitudes and are assigned specific volumes of material. With the use of computer hardware for data collection and graphically oriented mining software, the volumes of material are taken with measurements of density or specific gravity and are calculated into tonnage. In conventional geometric techniques, values at the sample locations may be connected to form a network of triangular prisms, or the value at each sample location may be given a surrounding area and volume of influence extending half-way to the next adjacent sample location. In the latter approach, the deposit is divided into a network of polygonal blocks. In combined geometric-statistical techniques, a search area representing a block of material is moved across the map of sample locations, and the value at the center of each successive search position is influenced by all of the included sample values. In the weighted moving average technique, each sample value is weighted by an inverse power of its distance from the center of the search area so that the nearer samples receive the greater influence.

Statistical techniques for calculating grades within specific volumes of ore make use of grids and polynomial trend surfaces fitted to the overall pattern of sample data, or make use of weighting factors determined from directional trends within the deposit. The latter is a widely used geostatistical approach in which a graph or variogram of values determined for increased sampling intervals in each of several directions is fitted to an idealized curve to identify the range of influence between samples. The estimation of values at individual locations employs the variogram function and kriging, a moving weighted average technique. *See* STATISTICS.

Coal sampling. Coal deposits are sampled to determine their properties as sources of fuel and coke. Characteristics are determined by proximate analysis for volatile matter, fixed carbon, moisture, ash, and heating value, and by ultimate analysis for chemical

TABLE 2. Composite assay values*

Sample	Length, m	Silver, g/metric ton	Length × assay product
1	0.5	220	110
2	0.3	270	81
3	0.7	550	385
4	1.0	350	350
5	0.5	200	100
Total	3.0		1026

*Based on sampling channel shown in Fig. 22. Average grade (1026/3) is 342 g/metric ton.

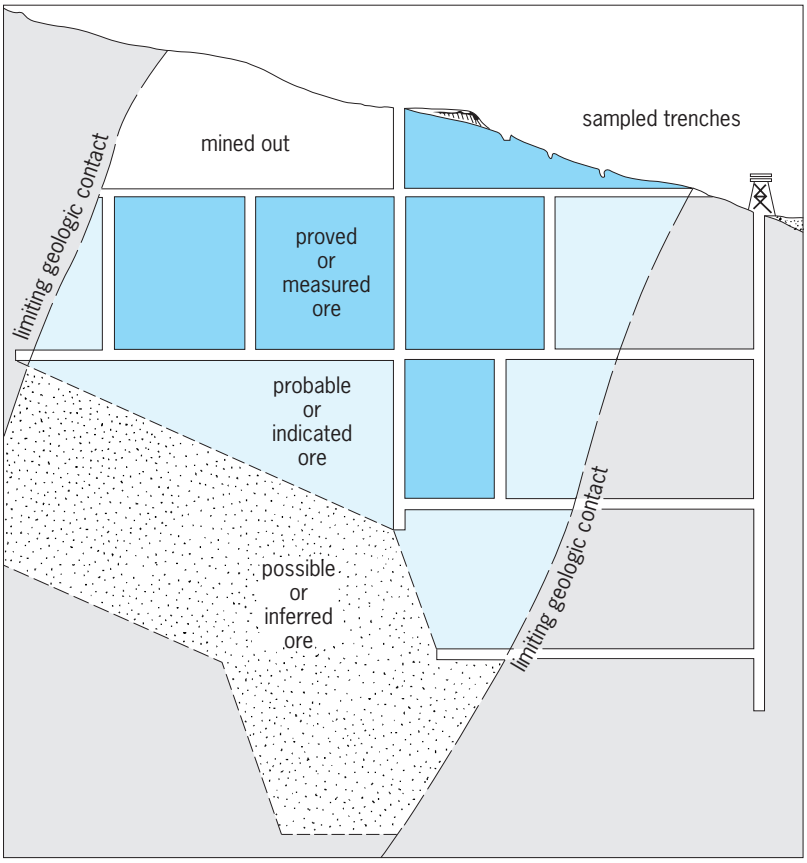

Fig. 24. Longitudinal section of a vein deposit, with ore reserve classifications.

composition. Engineering properties of coal seams are determined from samples by direct physical tests and indirect laboratory tests. Sampling is most commonly done by drill hole and channel methods.

Drill-hole sampling of coal involves core and rotary (chip) drilling, generally from surface locations. Sample intervals are selected according to recognizable coal units within the seam, and the samples of core or cuttings are placed in sealed containers to preserve natural moisture and minimize oxidation. Bulk samples are taken from drill holes by large-diameter core and by the washing of cuttings from induced cavities.

In trenches and underground workings, channel samples across a coal seam are taken as composites of shorter individual samples. Larger bulk samples are taken by blasting of the full seam.

Ore estimation. The quantity of ore in a block of mineralization and in an entire deposit is estimated by measuring areas and thicknesses and by considering the calculated volumes, tonnages, and grades.

A tonnage and grade estimate of the resource or geological reserves in a mineral deposit amounts to an inventory of the in-place mineralization. With graphical software for modeling and mining simulation, an estimate of ore reserves takes the minability of the deposit into account, with appropriate provisions for the method of mining, the percentage of ore that will actually be removed, the dilution of the ore in mining, and the grade and tonnage of material above the minimum economic value or cutoff grade. The rated economic value of ore mineralization will also take into account the level of its metallurgical recovery. For the part of a mineral deposit so thoroughly sampled that its outline, tonnage, and grade are assured, the estimated reserves are called proved or measured. In other parts of the deposit where sampling has been less thorough but where geologic information is sufficient to make reasonably secure projections, reserves are probable or indicated (**Fig. 24**). On the fringes of a deposit where sampling is sparse but where there is geologic evidence of continuity, the reserves are possible or inferred. William C. Peters

Mine Evaluation

Mines represent the culmination of successful investment in discovering and developing economic mineral deposits. To understand mine evaluation, it is necessary to consider the various stages of the mineral supply process, whereby initially unknown mineral deposits are discovered, developed, produced, and reclaimed (**Fig. 25**). In the broad sense, mine evaluation considers assigning value to mineral deposits at all stages of the mineral supply process. Because every mineral deposit is unique, the task of assigning value is complicated. Furthermore, the mineral industry is subject to the offsetting forces of depletion and technology. The finite size of mineral deposits means that mines eventually run out of ore and are closed. Over time, this depletion effect should increase the value of remaining deposits. However, technological advances at all stages of the mineral supply process work to offset the impact of depletion by allowing us to find new deposits and lower the costs of production at existing mines.

Exploration. As shown in Fig. 25, exploration proceeds through several possible stages from early reconnaissance work, where targets are identified in preparation for mine development. At the primary exploration stage, the value of a mineral deposit is highly speculative. This reflects the fact that very few mineral deposits have the requisite geological and technical characteristics to become economic mineral deposits (mines). Several hundred mineral properties may have to be examined for every one that becomes a mine. Overcoming these high odds is both time-consuming and expensive.

At each stage of the exploration process, money is spent to acquire information. This information is then used to make a go/no-go decision. Does the project merit the cost of further exploration? If so, the expectation must be that the probability of success times the expected return is higher than the cost of continuing the exploration program. As projects proceed through the various stages of exploration, the probability of becoming an economic deposit increases, as some of the discovery risk has been resolved. The value assigned to any given property should increase proportionately with the reduction in uncertainty and the increase in information.

Feasibility. As projects move to the delineation phase, the economic factors become more prominent. Deposits that are deemed to have potential

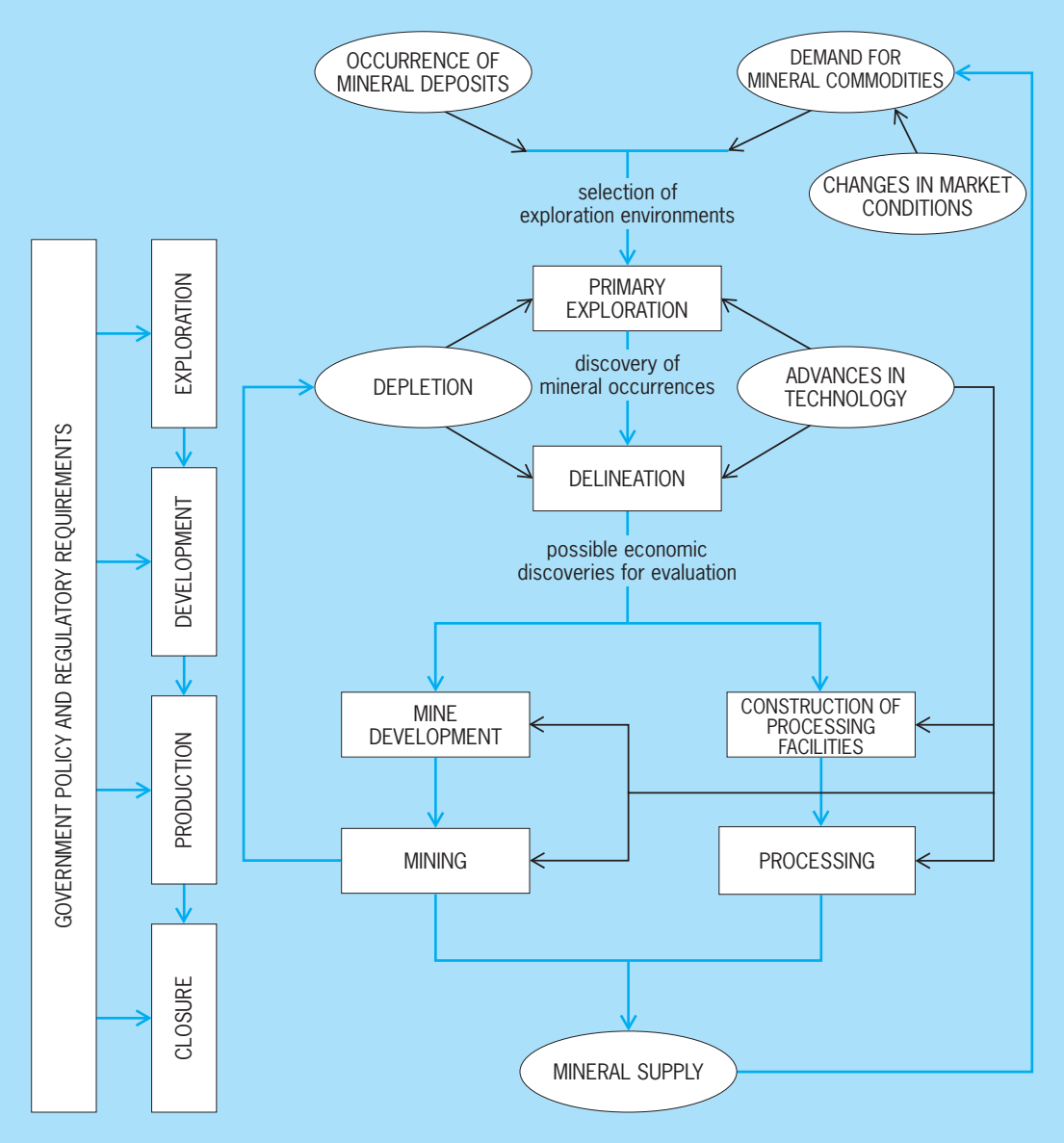

Fig. 25. Mineral supply process.

to become mines are assessed with scoping studies. These early technical and economic assessments are another stage in the go/no-go process. If go is the result, a feasibility study is done to determine if a project should be developed. Feasibility studies usually involve detailed technical evaluations of ore reserves, mining methods, processing recoveries, and waste disposal. Extra information may be collected in the form of core sampling, bulk sampling, or pilot tests. Feasibility studies may take many months or even years to complete and cost many millions of dollars. In the economic portion of a feasibility study, detailed projections are made for the revenues and costs associated with development and production. With respect to revenues, forecasts must be made of metal prices, inflation rates, exchange rates, and annual production of the various payable metals. On the cost side, estimates of capital expenditures are made based on the size of the ore reserve and the production capacity to be installed. These costs can be broadly categorized as mine development, mine equipment, processing facilities, and infrastructure. Production or operating costs are usually subdivided into mining, processing, and overhead/administration, but include detailed breakdowns on individual components such as supplies, labor, power, water, transportation, and refining. Estimates also must be made of the impact of taxes and royalties on project cash flows.

Discounted cash flow measures. The feasibility study is the last step in the go/no-go process. A positive feasibility results in the recommendation that a mine be established on the deposit, with the expectation that the returns from the mine will outweigh the cost of building and financing its construction, as well as closing the mine and reclaiming the site. The feasibility go/no-go decision is traditionally based on a set of discounted cash flow indicators, including

net present value and rate of return on investment. A net present value of greater than zero indicates a positive investment opportunity, meaning that the benefits (positive cash flows) outweigh the costs (negative cash flows), including an allowance for the cost of the capital used in the investment. The expected net present value determined in the feasibility study represents the inherent economic value of the mine. In addition to a positive net present value, most companies impose a profitability requirement in terms of return on investment. For many mining companies, an after-tax rate of return of 15% is considered the minimum level of profitability.

Once a project is in production, the value of the mining operation at any given time is the difference between benefits and costs over the remaining life of the mine. The net present value at any given time, looking forward over the remaining mine life, should represent the value of the mine to the owners. Further investment decisions made after a mine has been put into production are considered in the same manner. The cost of further investment, such as an expansion of mine capacity, is made on its own merits. If the anticipated increase in benefits outweighs the cost of the expansion, it should proceed.

Sensitivity and risk analysis. Uncertainties are a reality of the exploration and mining business. Feasibility study estimates are made on the basis of incomplete and uncertain information with respect to the physical, technical, and financial parameters of mine development. Expected value results need to be considered in the context of this uncertainty. This is done by a sensitivity analysis, whereby one or more of the input values is changed and the impact on the resultant economic criteria is calculated. For example, if the price of copper were 10% less than expected, how much would the net present value of the project be lowered? More complex analysis is possible by considering simultaneous changes to multiple variables. Probabilistic risk analysis and real option studies can be used to measure these impacts.

Michael D. Doggett

Computers in Mining

Computers are used at all stages of a mining operation to help geologists, engineers, designers, and management make informed decisions about the size and location of mineral deposits and how they can be profitably extracted. Computers have helped the mining industry be more systematic, make decisions based on quantifiable information, and implement procedures that are practical, are profitable, and meet stringent environmental safety requirements.

Exploration. Mineral deposits are located by examining surface samples, as well as samples taken from core drilling. Data collected from surface and drill samples are used to define the location of mineralization in three-dimensional space and also to define the elements and minerals composing the rock. Enormous amounts of data are collected, and computers provide a tool for analyzing the data in a timely manner.

The addition of a palmtop or laptop computer to a geologist's field tools is a recent development. Portable computers have high computing capacity and are available in light, easily portable packages. These small computers contain disk storage media and software that the geologist needs to collect data directly in digital form. Data collected in the field is displayed on the computer screen, and analysis can be performed on site before the geologist leaves the area where the data were collected.

Palmtop computers are usually about the size of a large hand-held calculator and include a digital screen. Software on a palmtop computer displays in a form that prompts the geologist for input such as rock type and characteristics, mineralization, outcrop orientation, amount of fracturing, weathering, and other essential descriptive geological information. A geologist may also collect information from core drilling. As core is brought to the surface, geological information is captured. The geologist examines the rocks, and descriptions and measurements are entered directly into a computer. Direct computer entry bypasses errors introduced when a handwritten data sheet is converted to a digital format in a computer. Software for geological data collection is often designed so the geologist can enter codes for rock and other characteristics and minimize typing. Data from a palmtop computer are later transferred to a laptop, to a desktop, or onto a company computer network for installation into a database and production of statistics, maps, and other analysis such as reserve calculations. If a geologist is using a laptop computer, which has more data storage and computing capacity than a palmtop computer, the entire data collection, analysis, and mapping work may be done in the field very soon after or even during data collection.

The data collected are organized in computer databases and used throughout the life of a project. The data stored include geological characteristics, rock types, spatial locations, geochemical analyses, mineral assays and analyses, and any other pertinent information applicable to the type of deposit being evaluated. The digital database is readily available for additions and refinements, ongoing reserve calculations, and use in later analyses for economic evaluation, mine designs, mine plans, operating schedules, and reclamation plans. A mine database is often stored on a central network computer, so that all departments can access it for particular needs, such as short-term mine schedules, location of mine waste, and many other operational requirements.

When enough data are collected for a preliminary definition of a potential ore deposit, computer software helps the geologist interpret the data and calculate mineral quantities in a target area. Mineral reserve calculations are readily updated usin mine-modeling software as more data are collected and added to the database. Continual revisions of the reserve estimates give the geologist clues about where additional information is needed. In the "past," weeks or months of laborious work was required to obtain reserves; and when more information was needed, it had to be collected during the next field season. Computer use has converted geological analysis from a static to a dynamic process. The power

of computer hardware (fast processors, large hard disks) and sophistication of modern graphically oriented software have provided exploration geologists the tools to keep pace with demand for information in a fast-paced industrial environment. *See* DATABASE MANAGEMENT SYSTEM.

The Internet is playing an increasingly important role in data transfer. For example, when a geologic map has been generated, the map and associated data can be sent to a private Web page where others can view it and instantly see the progress of the exploration efforts. Special Internet map server software provides an environment for the map user to look at a map or portions of a map at different scales and to include and overlay other information. For example, a composite map can be produced that shows a satellite image, location of drill holes, interpreted geological information, roads, power lines, political boundaries, and other features.

This nearly instantaneous access to information, and ability to evaluate it, provides a basis for evaluating the progress of an exploration program as the activities are happening rather than months later. Previously, geologists collected field information in the summer and spent the winter evaluating it. An ore body model often is required 6 months or more to complete using hand methods and additional months to adjust for new information. Computer hardware and software tools make the data collection and evaluation process a matter of days and weeks rather than months to complete an analysis of a prospect. *See* REMOTE SENSING.

Planning and operations. When geological reserves are available, the mining engineer begins the planning process to determine what part of a deposit can be mined economically and what type of mining system should be used: surface, underground, or some combination. The mine planner uses a desktop computer with specialized software to build a mathematical and graphical model of an open pit or underground mine design. Formerly, a single mine design often took up to 6 months to produce. An engineer with a computer and specialized mine planning software can produce a multitude of maps, designs, and plans in a few weeks. Each design or plan can be developed with a variety of design criteria, such as minimum grade, different metal prices, physical layouts, and a selection of economic conditions. The resulting surface or underground mine designs can be displayed in color and in three dimensions, and can be shaded to look almost like photographs when enough details are available. The graphics help the planner determine the workability of a design, and can also be used to show management and government regulatory agencies how a mine might look at different stages of development from startup to final reclamation. *See* CONTROL SYSTEMS.

The goal of an open pit mine designer is to find the most economic areas of a deposit, the areas where the value of the ore minerals exceeds the cost of removing waste material to expose valuable ore. Surface mine planning software is used to define ore and waste on each level of the proposed mine, as well as the final outline of the pit after it is mined. Each amount of material in a volume of earth is assigned a net value based on the potential sale price of the ore that can be recovered, the cost of removing the material, and processing or dumping it.

An underground mine design is created using a computer-aided design (CAD) program to graphically lay out the mine openings and extraction areas. The program operator sees a three-dimensional display of the geological model on the monitor screen and can lay out shafts, drifts, entries, and other access to the ore, as well as define the areas of the ore body that will be removed. Ore values and costs of mining are calculated for a design. The design is evaluated, then changed as necessary, until a suitable result is attained that yields the highest return. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

When the mine plans are complete, detailed economic evaluations are performed to determine whether the proposed mine is economical. If so, the next step is to prepare the information for mine permits, a long and complicated process in most countries. The types of software used to evaluate the geology and develop mine plans are also used to prepare documents for permitting. When permits have been obtained, development of a mine operation can proceed.

Surface operations. Computers are used for several aspects of guiding the progress of mining in open pit and strip mines. Surveyors lay out blast-hole patterns based on computer-generated designs. The surveyors use electronic surveying equipment to lay out individual blast holes, identify existing pit outlines, and other surface features and feed these data into a computer for further processing. Computer programs are also used to identify the quantity and type of explosive required to properly fragment the rock. Samples of blast-hole cuttings in an open pit mineral operation are taken during drilling and sent to an automated analysis lab to determine quantities of minerals in the sample. The results are entered into a database and used by the engineer to produce an ore/waste map for the next day's mining. Computers and software provide the ability to produce maps quickly and keep the operation moving while routing mined material to the proper location—ore to the mill, waste to the dump. Maps and ore models developed from the blast-hole information are compared to the model developed from the exploration drilling. The comparison helps the engineer and geologist make adjustments to the modeling method if necessary.

In an open pit, tracking the movement of mobile equipment such as trucks, shovels, and loaders is essential to keep the operation flowing smoothly. Computerized truck dispatching systems are installed in the larger mines to monitor where trucks and other mobile equipment are located at any time and where they travel to be loaded and dump the load. Computers on board equipment monitor activity and movement. Dispatching systems use beacon stations in the pit to sense vehicle movement or use global positioning systems (GPS) to track the location of equipment. The dispatching system provides a printed report

of all activity plus a graphic display showing where the equipment is at any time. Summary reports are available for management at any time if the dispatch data is stored on an in-company network. The dispatch system shows equipment operation hours and is sometimes connected with the maintenance system to help track availability, provide information for preventive maintenance, and record equipment downtime. The system may also collect personnel working hours.

Computer-controlled driverless equipment has been in development for many years, but is not yet widely accepted and used in the mining industry. Several prototype operations have been installed, and the technology is continually being improved. *See* CONTROL SYSTEMS.

Underground operations. Underground mining is a more complicated process than surface mining. Active mining areas are often isolated, and the mine shift bosses and foreman spend most of their time moving from one area to another to make sure extraction is progressing properly. Computer and radio equipment has improved communications underground and helps management keep material flowing smoothly from the active mining faces to the surface.

Sensors installed in an underground mine can detect airflow quantity, presence of gases such as methane or radon, location of equipment, operating condition of equipment, ground movement, and other conditions. The various sensors connected to a computer collect data and display graphs, numerical information, or live video. The information is then available for display on computers in underground offices and surface offices. Management can quickly see at any time the status of the operation and whether equipment is functioning properly, and can provide information to identify possible trouble spots before they become a crisis.

Computerized remote control and robotic equipment is used in some operations such as a low coal seam where it is very difficult or dangerous for a human to work in an area. Automated mining systems and equipment are still being developed and are not yet widely used in mines. *See* ROBOTICS.

Mineral processing. Computer process control is well developed in milling operations. These complex sensor/computer systems provide information to the mill operator about the condition of everything that is happening at any given time at any location in the mill. Process control systems are a dedicated system of computer hardware and specialized software. The process control system contains a central controller, sensors around the mill, one or more operator consoles, and video monitors for different parts of the process. The monitors show numbers such as motor horsepower and other measurable quantities, graphs showing the quantity of material being fed into a crusher, and actual videos when cameras are used to monitor particular parts of the operation. The operator has the ability to change equipment settings from the console without having to physically touch the equipment. *See* PROCESS CONTROL.

Active equipment maintenance in a milling operation is essential for efficient movement of material from the beginning to the end of the process. Properly maintained equipment keeps metal recovery levels high and contributes to a profitable operation. The computers and software used for fixed equipment maintenance, such as found in a mill, is often the same as that used for mobile equipment maintenance. *See* PROCESS CONTROL.

Mining systems. Simulations of all phases of the mining operation can be performed using computer software. Parts or all of a mine operation are simulated using mathematical and probabilistic models. Simulations commonly are run for planning haulage routes, designing a milling operation, evaluating an existing mill or mining operation to identify solutions to problems, and showing how material is actually moved from one place to another. Computer simulation offers the ability to try different mining scenarios at a very low cost. For example, a simulation can help define whether it is economic to use trucks to haul material out of a pit or to install an in-pit crusher and conveyor belt to move material.

Scheduling is an essential function for keeping production at the highest levels. Long-term and short-term production schedules are developed from the overall mine plan. Production targets, consisting of quantity and grade of material, are defined by the engineer, and then the areas that meet those qualities are selected from the mine plan and assigned for extraction to a calendar time. Because short-term scheduling is subject to change (for example, the blast holes may show a different grade than the original exploration drilling), computer programs permit mining engineers to dynamically change the outline for the next production period and calculate new projected grades and volumes. Another approach to automated scheduling is to use a graphical CAD technique where an area is defined and the ore quantity and grade are calculated from the mine's database.

Computerized maintenance and inventory systems are an essential part of modern mines. Maintenance systems produce reports about each piece of equipment (whether stationary or mobile) and when it is due for servicing or major overhaul. A history of each piece of equipment is readily available from the computer maintenance database that shows what parts were installed, any breakdowns that occurred, and when the next regular maintenance is due. The availability of each piece of equipment is reported, and replace or repair decisions can be made. The maintenance system is closely connected with parts and supplies inventory which helps operators identify when to order additional parts or consumables. The system also maintains historical records for parts reliability and equipment failure tracking.

Maintenance and inventory data are stored on networked computers, so it is readily available to the maintenance department personnel and management. Maintenance management systems are often tied to the mine accounting system as well.

Enterprise-wide software solutions are appearing in some of the larger mining operations. This software connects all parts of the mine operation,

including financial accounting, human resources, safety training, maintenance, inventory, facilities management, and material movement. Management can at any time access an enterprise system and see detailed activity reports or summaries of exactly what is happening in the operation at that time. Historical records of past performance help the manager identify changes in production levels and spot potential problem areas.

Mine reclamation. Geographic information systems (GIS) have become important throughout a mine operation, but are especially helpful in preparing maps showing a mine operation in relation to soils, vegetation, wildlife, surface hydrology, and other environmentally sensitive information. GIS combines a wide array of different types of information to produce maps showing how an area will appear during phases of the mining operation and to show final reclamation plans. One use is to show more efficient and effective ways of locating topsoil storage, so it is available in the right place when it is needed for final surface coverage. GIS can also be used for facilities management to keep track of existing infrastructure and future plans.

GIS is used during the exploration phase to combine satellite images with ground data and organize all the spatial data collected from a variety of sources. Surface geochemical data, geological data, geophysical data, assay results, and other spatial data can be posted on a map and analyzed for spatial trends. The information can be shown against a satellite photo background and can be viewed in perspective. *See* GEOGRAPHIC INFORMATION SYSTEMS. Betty Gibbs

Environmental Impact

Historically, mines have caused significant adverse environmental impact. Often, little consideration was given to the effect of the mine on the environment, and rarely was the land reclaimed. As a result, natural weathering processes have caused erosion of unreclaimed land, and reaction of certain rock types with rainwater, forming acid- and metal-laden runoff that may contaminate streams and ground water.

Contemporary mining operations have come a long way environmentally. Today, mining companies recognize that land reclamation and environmental protection are as much a part of mine planning as are the extraction and recovery of the mineral itself. As the practice of environmental assessment and the understanding of the impact of mining wastes on the environment have advanced, the adverse impact on the environment has been reduced.

Laws and regulations. To assure that mining operations act in an environmentally prudent manner, nearly every country has enacted laws and regulations requiring environmental protection. In the United States, laws such as the federal Surface Mining Control and Reclamation Act and various state reclamation acts specifically require mining operations to evaluate environmental effects and develop reclamation plans to mitigate impacts. A few federal laws that directly affect mining but are nonspecific to the industry include the Clean Air Act, the Clean Water Act, and the Resource Conservation and Recovery Act. States also usually have programs similar to these federal laws.

In countries such as Peru, where economic development and public environmental awareness are advancing but are still in the formative process, the laws and regulations are also in the formative stage. These countries typically have an established front-end review and permitting process, but have limited enforcement capability due to limited financial support for the program. In underdeveloped countries where the political environment is unstable and there is very little economic development or public environmental awareness, only very general environmental protection laws may have been enacted.

Developing a mine usually requires a significant investment. Many mining companies must therefore secure mine development financing from international lending institutions. Although they have no environmental regulatory authority, nearly all of the major international lending institutions have specific guidelines for land reclamation and environmental protection. So even in underdeveloped countries where the government has little ability to administer and enforce its environmental protection laws, mine developers are still required to carry out environmental evaluations and reclamation programs in order to secure financing.

Environmental evaluations. Laws and regulations in nearly all countries, as well as most international lending institutions, require a mining company to assess the impacts that its development plans will have on the environment and to propose programs to limit those impacts before construction of the mine begins.

Pre-mine environmental conditions. The first step in evaluating the potential environmental effects of a mining project is to define the preexisting environmental conditions at the location where the mine will be developed. Baseline studies are typically commissioned to characterize the existing physical, biological, and human-interest resources in the area. Physical resource concerns include air quality and climatic and meteorological conditions; surface-water and ground-water quantity, quality, and flow characteristics; the nature, extent, and surrounding geologic media associated with the mineral deposit; and the agronomic capability, land uses, and erosion characteristics of the soils. Biological baselines characterize the presence, abundance, and diversity of plants and animals, including the aquatic species and insects, and the identification of any sensitive or protected plant or animal species. Human-interest resource baselines provide a profile of the population distribution, statistics, education, and income levels of the people near the project site; profile the existing infrastructure and services available; define the local economic and tax structure; define cultural and ethnic beliefs; and identify the presence of historical and archeological resources.

These studies provide a basis against which the impacts associated with mine development can be measured. The development of the environmental baseline data usually requires at least one year, or more, of data collection to address the seasonal

variation of certain resources. For this reason, most baseline studies are initiated several years before mine construction is anticipated.

Predicting environmental impacts. After the baseline conditions have been characterized, project design and development plans need to be analyzed relative to those baseline conditions to predict the effect that project development will have on the environment. This process requires an understanding of how the project designs will be implemented, operated, and maintained, as well as specific knowledge of the environmental resource being evaluated.

One of the most important factors that need to be evaluated in an impact assessment is how mined and processed materials and wastes will behave in the environment. Mining typically exposes to the atmosphere various rock types that have not been exposed for millions of years. Weathering processes can affect certain rock types, resulting in acid generation, metal leaching, and the release of elevated concentrations of pollutants to streams and groundwater systems. It is therefore extremely important to understand the geochemistry of the rock to be mined and the effect that the weathering process will have on the rocks.

Impact assessment is an iterative process in which mine plans are analyzed relative to each specific environmental resource. For example, the effect that mining will have on air quality is analyzed relative to the kinds of emissions that are expected to be generated from each source. Typical emission sources at a mine include exhaust systems from mobile equipment, stacks at processing operations, dust from rock crushing activities, and dust from roads and other surface disturbances. The impact of development is analyzed by understanding the air quality that exists before the mine is developed (the baseline condition), assessing all of the contributing emissions that are expected to be generated by mine development, and predicting what the air quality will be during the various phases of the mine operation. Similar predictions are then made for each of the other environmental resources, such as water, soils, vegetation, and people.

After the effects of project implementation have been predicted for each resource, it is necessary to further evaluate how those changes will fit with other reasonably foreseeable developments in the area. For example, if there are plans to develop other mines and mineral processing facilities over the next several years, the effect of each facility on the environment needs to be evaluated in order to anticipate the combined effect.

Reducing adverse environmental impacts. Most countries and lending institutions have specific limits on how much pollutant can be released from a mining or industrial operation. After the effect of project development has been predicted for each receiving resource, those predictions are compared to the specific requirements for that discharge. If the current plans fail to meet the required levels established for environmental protection, project planners must change their designs to provide a system that will assure compliance.

Mine designers can limit environmental impacts in many ways. Probably the most significant way to reduce the long-term environmental impact of a mining operation is by implementing an effective closure and reclamation plan. Some of the more common ways to control the discharge of pollutants include designing control systems for mining and processing wastes, water and wastewater treatment, and sedimentation control. Disturbances to sensitive plant and animal species, or to cultural or archeological sites, can often be avoided by changing the design location of certain facilities. Where such disturbances are unavoidable, translocation and data recovery programs can often be implemented to reduce the overall impact.

Environmental management systems. Environmental management and monitoring systems are excellent tools for limiting the environmental effects of mining to the extent possible. These systems are designed specifically to monitor the performance of the environmental controls and mitigation measures that were defined in the design stage, and to continuously provide improvements throughout the life of the project.

The environmental management system is a program, to be administered by mine personnel, that defines a systematic process for documenting environmental conditions at the site. It will typically include specific monitoring programs for air quality, meteorology, surface-water and ground-water quality and quantity, the success of revegetation testing, and wildlife presence. It will also clearly establish the lines of authority for implementation and contingency plans for upset conditions. However, because every mine is unique in its environmental effects, the environmental management program must be customized to address what is important for a specific site.

Systematic management and monitoring programs allow mine personnel to monitor the environmental conditions, evaluate their consistency with the predictions made in the impact assessment, and make timely modifications to environmental controls if the conditions vary from the original predictions or do not meet specific environmental requirements. Because these systems are continually being updated and improved, they also allow for new technologies to be easily integrated into the environmental protection program.

Closure and reclamation. Unlike historic mining practices, contemporary mining operations plan for the closure and reclamation of mined lands before the mine is developed. As the mine gets closer to the end of its operational life, the closure and reclamation plans that were defined conceptually in the planning and impact assessment phase must be more specifically defined so they can be implemented.

Closure activities are implemented to assure the long-term chemical stabilization of the site. Examples of closure activities include the rinsing or cleaning of rock piles or equipment that have been exposed to chemical reagents or reactions. The covers that may be required to isolate mine wastes from weathering processes would also be considered clo-

sure components if, without the covers, the wastes would release pollutants to the environment.

Land reclamation restores mine site disturbances to a safe, stable, and productive postmining condition. Reclamation programs are typically initiated after the closure activities are completed. Reclamation activities typically include precluding entry to mined-out workings, grading and shaping the land to a form that is stable over the long term, placing soils over the graded land disturbances, and planting vegetation to allow the final reclaimed surface to blend with the surrounding environment. *See* LAND RECLAMATION.

There are many examples of contemporary mines that have been developed, operated, closed, and reclaimed in an environmentally prudent manner. While it is impossible to mine without affecting the environment, it is possible to limit its environmental effects. With good up-front analysis and impact assessment, effective environmental management systems, and successful site closure and reclamation, the mining industry will continue to balance the need to protect our natural resources with the need to provide the raw materials demanded by our advancing society.

Barbara A. Filas

Restoration of Land

Land disturbance is frequently the most visible impact of mining on the environment. Pits and quarries from surface mining, waste rock dumps, low-grade ore leaching dumps or heaps, and tailings ponds for disposal of processing waste are prominent features in the landscape of a mining area. A variety of techniques are used to reduce or eliminate potential environmental hazards associated with these features.

The goal of reclamation is to return the land surface to as close to its premining condition as possible. When this is not possible, the goal becomes stabilization of the area and preparation of it for an acceptable postmining condition or alternative land use. The basic steps of reclamation usually include movement of materials to achieve a stable land contour, replacement or rebuilding of topsoil, and revegetation. The complexity of reclamation is related to the type of deposit, the mining method used, the type of processing done at the mine site, and the physical and chemical properties of the waste materials involved.

Surface mining. In surface mining, one of the first steps in mine development is removal of the topsoil layer. The topsoil is set aside so that it can be used to cover the disturbed area and be revegetated during reclamation. Shallow flat-lying deposits, such as those mined for phosphate in Florida, are amenable to area mining; that is, the overlying waste material and ore are removed sequentially in long narrow cuts. The waste rock from a cut can be used to fill the previous cut. After several cuts are completed, reclamation can be initiated and then continue throughout the life of the mine. The waste piles in the mined-out areas are recontoured by bulldozers and scrapers. Topsoil is replaced and vegetation is planted. In all types of reclamation, plant species are selected for the soil and climatic conditions on the mine area. When fill materials are not available for the final cut, it may be contoured to form a lake or a wetland. The continuous reclamation process means that only a small portion of the surface area will be disturbed at a time.

Deeper mines. Deeper and more localized deposits, such as those mined for copper, iron, and gold, require the construction of large open pits. As production continues, the pits expand downward and outward, making reclamation impossible until all mining has ceased. Waste rock covering the orebody is removed and placed in permanent waste dumps. The configuration of the waste dumps may be dictated by the topography. In steep mountainous regions, areas for dump construction may be limited; and dumps may be high, making reclamation more difficult. In flatter terrain, such as the desert areas of the southwestern United States, lower-profile dumps are constructed. They are generally more stable, blend in better with the surroundings, and are easier to revegetate. While continuous reclamation may not be possible under these conditions, individual dumps can be completed and reclaimed at various times during the life of a mine. The pit remaining when mining is completed presents a unique problem. Backfilling is usually neither practical nor economically feasible. However, the slope angle of the sides of the ultimate pit can be designed to give maximum stability. The steplike benches that are left on the sides of the pit can catch any loose rock and can be revegetated. Depending on the local hydrology, these pits may also fill with water and form lakes.

Quarries. Quarries and pits mined for construction materials represent yet another type of reclamation challenge. These mines are frequently located in populated areas. The visual impact can be reduced by constructing and landscaping berms of rock and soil around the perimeter of the property. This also helps suppress equipment noise that may be objectionable to neighbors. Deep quarries are treated in the same manner as open-pit metal mines. Shallower sand and gravel pits are often recontoured to make them valuable as areas for residential and commercial development, with the deeper pits left as lakes and ponds. *See* LAND RECLAMATION.

Leach dumps and heaps. In addition to dumps constructed of waste rock, special dumps and heaps are designed for the leaching of low-grade metallic ores. Leaching solutions containing acid or sodium cyanide are typically used for copper or gold, respectively. They are sprayed on the tops of the dumps. As the solutions percolate downward through the pile of rock, the valuable metals are dissolved and carried away in the solutions. The most important environmental concern is contamination of ground water if the solutions are not controlled. Leach dumps and heaps are designed to prevent solution loss. Dumps such as those used for copper leaching can be built on impervious bedrock with a system of channels and sumps constructed for solution collection. Dumps and heaps for leaching gold are constructed on a series of liners that prevent solution loss. The liners may be made of asphalt, synthetic materials,

or impervious clay. Leak-detection instruments are installed between the liners. Monitoring wells surrounding the leaching areas are a further precaution. Leaching operations must provide adequate storage reservoirs to contain solutions during periods of high rainfall or snowmelt.

Tailings ponds. Fine-grained waste-rock tailings from mineral-processing plants are stored in tailings ponds. Well-engineered dams constructed of the coarse waste materials are used to impound finer wastes. The design of tailings dams must provide long-term stability and safety. The tailings are transported to the ponds as slurries. As the solids settle from the slurry, the water is recovered and recycled to the mineral-processing operations. If excess water is generated as drainage from a tailings pond, it must be collected and, if necessary, treated before it is released to a stream. The level of treatment will depend on the chemistry of the water. In some instances, older tailings ponds have contaminated ground water. One method for correcting this problem is to drill wells to intercept the plume of contaminated water and pump the water to a treatment plant. Because of the fine particle size of tailings, dust control is also important. When a tailings pond is full, it must be reclaimed.

Underground mine subsidence. Most underground mining methods do not disturb large areas of the land surface. The underground workings may be backfilled with waste rock or tailings to prevent collapse, or they may be designed to remain open. However, some large-scale mining methods, such as block caving, are designed so that the openings collapse in a controlled manner. When the mine workings collapse, the land surface over the mining area will eventually subside. Depending on the depth and size of the workings, the extent and timing of the subsidence will vary. Although subsidence may take years to stabilize, reclamation of stable areas is possible.

Erosion and sedimentation. Storm-water runoff is the major cause of erosion and the resulting sediments that pollute rivers and streams. Surface water must also be controlled to prevent mine flooding and to reduce the potential for creating contaminated mine drainage. Streams in the vicinity of a mine can be diverted away from the active operations, although provisions may have to be made to accommodate fish. Retention ponds are designed with adequate capacity to hold water during a peak rainfall event and to release the water gradually after the event. Sediment basins are used to permit solids to settle to the bottom before the water is released. These basins are also used in placer mining to prevent the release of sediment into local streams. Revegetation provides another highly effective means of controlling erosion and sediment.

Treating contaminated mine water. The most common form of contaminated mine water is acid mine drainage. It is formed when water, air, sulfide minerals such as iron pyrite, and bacteria react to form sulfuric acid. This acid may in turn result in the release of toxic metals into solution. When acid mine drainage is released into rivers and streams, it can damage water supplies, harm aquatic life, and ultimately render the waterway sterile. Several thousand miles of streams in the United States are contaminated by acid mine drainage from abandoned mines.

Eliminating contamination from acid mine drainage requires water treatment. Depending on the chemical elements involved, treatment processes range in complexity from simple neutralization with lime to more elaborate processes to remove toxic metals such as uranium, cadmium, or lead. Constructed wetlands are being used for treating acid mine drainage. The drainage flows through shallow ponds lined with crushed limestone and organic material such as compost or manure, and then cattails are planted. The metals are removed and become part of the organic substrate. Originally developed for treating coal mine drainage, several large-scale wetlands have been designed to treat metal mine drainage. *See* WETLANDS.

Dust control. Airborne dust can be generated in many parts of a mining operation. The simplest method of dust control is to spray water on mine roads, preventing trucks and other types of mobile equipment from raising clouds of dust. Chemicals and surfactants can also be used to improve control. In crushing plants, water sprays or dust collection systems and filters and used. After mine closure, revegetation reduces airborne dust. L. Michael Kaas

Bibliography. American Geological Institute, *Dictionary of Mining, Mineral, and Related Terms*, 2d ed., 1997; A. E. Annels (ed.), *Mineral Deposit Evaluation: A Practical Approach*, 1991; *Application of Computers and Operations Research in the Mineral Industry* (APCOM), various publishers of proceedings from conferences held since 1967; C. J. Bise, *Mining Engineering Analysis*, 2003; H. L. Hartman (ed.), *SME Mining Engineering Handbook*, 2 vols., 2d ed., 1992; H. L. Hartman et al., *Mine Ventilation and Air Conditioning*, 3d ed., 1997; H. L. Hartman and J. M. Mutmansky, *Introductory Mining Engineering*, 2d ed., 2002; C. C. Heald (ed.), *Cameron Hydraulic Data*, 19th ed., 2002; B. A. Kennedy (ed.), *Surface Mining*, 2d ed., 1990; J. J. Marcus (ed.), *Mining Environmental Handbook*, 1997; M. J. McPherson, *Subsurface Ventilation and Environmental Engineering*, 1993; C. J. Moon et al., *Introduction to Mineral Exploration*, 2d ed., 2006; National Institute for Occupational Safety and Health, *Criteria for a Recommended Standard: Occupational Exposure to Respirable Coal Mine Dust*, DHHS Publ. no. 95-106, 1995; R. V. Ramani, B. K. Mozumdar, and A. B. Samaddar (eds.), *Computers in Mineral Industry*, 1994; E. A. Ripley, R. E. Redman, and A. E. Crowder, *Environmental Effects of Mining*, 1996; I. C. Runge, *Mining Economics and Strategy*, 1998; M. Sengupta, *Mine Environmental Engineering*, vols. I and II, 1989; A. J. Sinclair and G. H. Blackwell, *Applied Mineral Inventory Estimation*, 2002; J. G. Stone and P. G. Dunn, *Ore Reserve Estimates in the Real World*, 3d ed., Soc. Econ. Geol. Spec. Publ. no. 3, 2002; T. F. Torries,

Evaluating Mineral Projects: Applications and Misconceptions, 1998; D. Wyllie and C. W. Mah, *Rock Slope Engineering; Civil and Mining*, 4th ed., 2004.

Mink

A semiaquatic, carnivorous mammal in the family Mustelidae. The American mink (*Mustela vison*) is found in Alaska and Canada and throughout most of the United States. The range extends southward in the western United States to northern New Mexico, northern Nevada, and central California. This species was deliberately introduced into the former Soviet Union, and escaped animals have established populations in France, Spain, Portugal, Germany, Poland, Ireland, Great Britain, and Scandinavia, as well as in Iceland.

The mink has a long, slender, weasellike body and a tail that is bushy at the tip. It is short-legged and almost uniformly dark brown (see **illustration**). The underparts are only slightly paler than the back. Long, glistening guard hairs partially conceal the soft, luxurious underfur. A white patch is usually present on the chin and may extend to the throat and chest. The tail is less than half the length of the body. Males are considerably larger and heavier than females. Females possess eight mammae. Paired anal scent glands are present. Five toes are present on each foot, and the hindfeet may be slightly webbed. The sense of smell is exceptionally well developed, while the senses of hearing and sight are moderately developed. Mink can hear ultrasound within frequencies emitted by potential rodent prey. The dental formula is I 3/3 C 1/1 Pm 3/3 M 1/2 × 2 for a total of 34 teeth. Adult mink are 530–660 mm (20–26 in.) in total length, including a 175–225-mm (7–9-in.) tail. They normally weigh 0.5–1.4 kg (1–3 lb).

The sea mink (*M. macrodon*) formerly occurred along the Atlantic coast from New Brunswick to Massachusetts but is now extinct. It is thought to have been considerably larger than the American mink. No complete specimen is known to exist. Descriptions are based on recorded observations and on numerous bone fragments and teeth found in Indian middens along the New England coast. This species apparently inhabited the rock crevices and ledges along the ocean shore. It became extinct about 1880, possibly as a result of overhunting.

American mink (*Mustela vison*).

American mink are found mainly along the banks of streams, rivers, ponds, and lakes and in swamps and marshes. Forested, log-strewn, or brushy areas are preferred. Males may range several kilometers along a stream. Mink are mainly nocturnal, semiaquatic mammals and seem to be most active around dawn and dusk. They are excellent swimmers and divers. Although they spend a lot of time in the water, mink also travel much on land and on rare occasions climb small trees. Mink are active all year. They are normally solitary, except for family groups of mother and young. The mink may dig its own burrow, or it may appropriate one made by some other animal such as a muskrat or rabbit. Dens may also be located under logs or debris, in rocks, or in muskrat houses.

Mink feed primarily on mice, muskrats, rabbits, birds, frogs, fish, and crayfish. They are able to catch swimming fish. They have also been known to eat bats that were thought to have fallen from the ceiling of a cave. In northern areas, breeding normally occurs during February and March. Ovulation is induced by copulation. The single annual litter is produced in April or May after an average gestation of approximately 51 days (range 40–75 days) because of delayed implantation. Implantation occurs about 30 days before birth. Litters may contain from three to ten young, although the average is four or five per litter. The young are born in a nest of feathers, fur, bones, grass, and plant fibers. Newborn young have their eyes and ears sealed and are covered with fine, silvery-white hair. Within 2 weeks, pale reddish-gray hair replaces the white hair. The teeth begin erupting between 16 and 21 days. The eyes open at about 37 days. Weaning is at about 5 weeks, and the young leave the den when 6–8 weeks old. The young mink remain with the female until late summer or fall, at which time they disperse. Although the male may mate with several females, it usually stays with the last one and assists in caring for the young. Both males and females are capable of breeding during the breeding season following their birth. Mink have few enemies other than humans and their traps. Occasionally, one may be attacked by a fox, bobcat, or great horned owl. Captive ranch mink have lived 10 years, although the life-span in the wild is probably 3 or 4 years. Most of the mink fur used in commerce is produced on farms. Selective breeding has produced colors such as black, white, platinum, and blue, and descriptive trade names such as Aleutian, amber gold, blue iris, pastel, platinum, sapphire, and winter blue. *See* CARNIVORA; DENTITION; MARTEN; WEASEL.

Donald W. Linzey

Bibliography. G. A. Feldhamer, B. C. Thompson, and J. A. Chapman, *Wild Mammals of North America: Biology, Management, and Conservation*, 2d ed., Johns Hopkins University Press, 2003; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Miocene

The second subdivision of the Tertiary Period (Eocene, Miocene, and Pliocene) by Charles Lyell in 1833; the fourth in a more modern sevenfold subdivision (epochs) of the Cenozoic Era; and the first epoch of the Neogene Period (which includes in successive order the Miocene, Pliocene, Pleistocene, and Holocene). The Miocene represents the interval of time from the end of the Oligocene to the beginning of the Pliocene and the rocks (series) formed during this epoch. The Miocene was originally considered as a biostratigraphic (rather than temporal) entity, to unite rocks containing 20–40% extant molluscan species. Lyell based his concept of the Miocene primarily on the shallow marine sediments and associated molluscan faunas in the Superga Hill of northern Italy (near Turin), as well as other Piedmont localities and outcrops in southwestern France (Aquitaine Basin), Touraine, and the Vienna Basin. *See* CENOZOIC; HOLOCENE; PLEISTOCENE; PLIOCENE; OLIGOCENE; TERTIARY.

The Miocene spans the time interval between 23.8 and 5.32 million years ago (Ma) based on integrated astronomical and radioisotopic dating. The Miocene/Pliocene boundary is located in Sicily, just above a major unconformity separating the youngest late Miocene (Messinian) deposits (of the Great Terminal Miocene Salinity Crisis) and the overlying white chalks of the Zanclean. *See* UNCONFORMITY.

Subdivisions. While originally conceived as biostratigraphic units (that is, including rocks containing various broadly contemporaneous marine and terrestrial faunas), the epochs of Lyell (including the Miocene) have come to be considered as chronostratigraphic units (rocks having clearly specified, conceptually temporally isochronous and correlative boundaries) in which fossil faunas and floras are used for the purpose of characterization and correlation. Integrated studies using radiometric dating, paleomagnetic stratigraphy, stable-isotope (oxygen, carbon, strontium) studies together with (micro)paleontologic studies have provided a framework for regional and global correlation of the disjunct stratigraphies established in the marine and terrestrial biosphere (**Figs. 1** and **2**).

An essentially twofold standard subdivision of the lower Miocene is recognized: the Aquitanian and Burdigalian stages, having their type development in the Aquitaine Basin of southwestern France. The middle and upper Miocene have their typical development in the tectonically active basins of Italy. The middle Miocene is represented by the Langhian and Serravallian stages and the upper Miocene by the Tortonian and Messinian stages. Marine sedimentation essentially ceased in the Anglo-Paris-Belgian basins (type areas for the Paleogene stages) in the mid-Oligocene, whereas in the Aquitaine Basin marine sedimentation ceased in the early middle Miocene (Langian Stage). A comparable subdivision has been made for the relatively complete Miocene sequence which is developed in the Paratethys region of central and eastern Europe (Vienna Basin to the Urals).

CENOZOIC	QUATERNARY	Holocene
		Pleistocene
	TERTIARY	Pliocene
		Miocene
		Oligocene
		Eocene
		Paleocene
MESOZOIC	CRETACEOUS	

Tectonics and volcanism. Major orogenic and volcanic events characterize the Miocene (**Fig. 3**). Plate-tectonic motions, originating in the Mesozoic, resulted in the gradual dismemberment of the Tethyan Ocean and the upthrusting of the Alpine-Himalayan orogenic belt in three major phases: the late Eocene (about 40 Ma) and the early (21–17 Ma) and mid-late Miocene (10–7 Ma). Along the eastern margins of the Pacific Ocean, the ocean crust was subducted under the North and South American continents, giving rise to major orogenic movements stretching from the Aleutians to Tierra del Fuego. The Andes range was thrust up during the later part of the Miocene. The Pacific Coast developed as a result of westward drift of North America over, and partial consumption by, the Farallon Plate and collision with the Farallon Ridge. Only two relatively minor plates remain as remnants of the Farallon plate: the Juan de Fuca and Cocos plates between Mexico and Alaska. The plate margin was bounded by transform faults rather than a subduction zone, and northwestern propagation of a major transform fault issuing from the Cocos plate formed the Gulf of California in the late Miocene, and its continued extension northward is familiar to residents of the west coast as the San Andreas Fault System. The latter was responsible for the formation of many of the off- and onshore basins of southern California, some of which contain prolific petroleum resources. Subduction of the Pacific plate at the Middle America Trench during the late Paleogene and Neogene resulted in arc magmatism and eventual uplift of the Central American Isthmus into a series of archipelagos in the late Miocene (about 7 Ma) and eventual fusion into a continuous land bridge in the early Pliocene (about 3 Ma) that resulted in the separation of the Atlantic and Pacific oceans and concomitant disruption in marine faunal communities as well as transcontinental migration of vertebrate animals in the Great American Faunal Interchange. *See* OROGENY; PLATE TECTONICS; SUBDUCTION ZONES; TRANSFORM FAULT.

The Mediterranean, Black, and Caspian seas are remnants of an ancient equatorial ocean called the Tethys that separated Africa and Europe and linked the Atlantic and Pacific oceans for over 100 million years. In the mid-Cenozoic the continued northward movement of the African plate resulted in its collision with Europe in the early Miocene (about 18–19 Ma) along a suture line represented today by the Taurus (Turkey) and Zagros (Iran and Iraq) mountain ranges; the formation of the Alpine, Dinaric, and

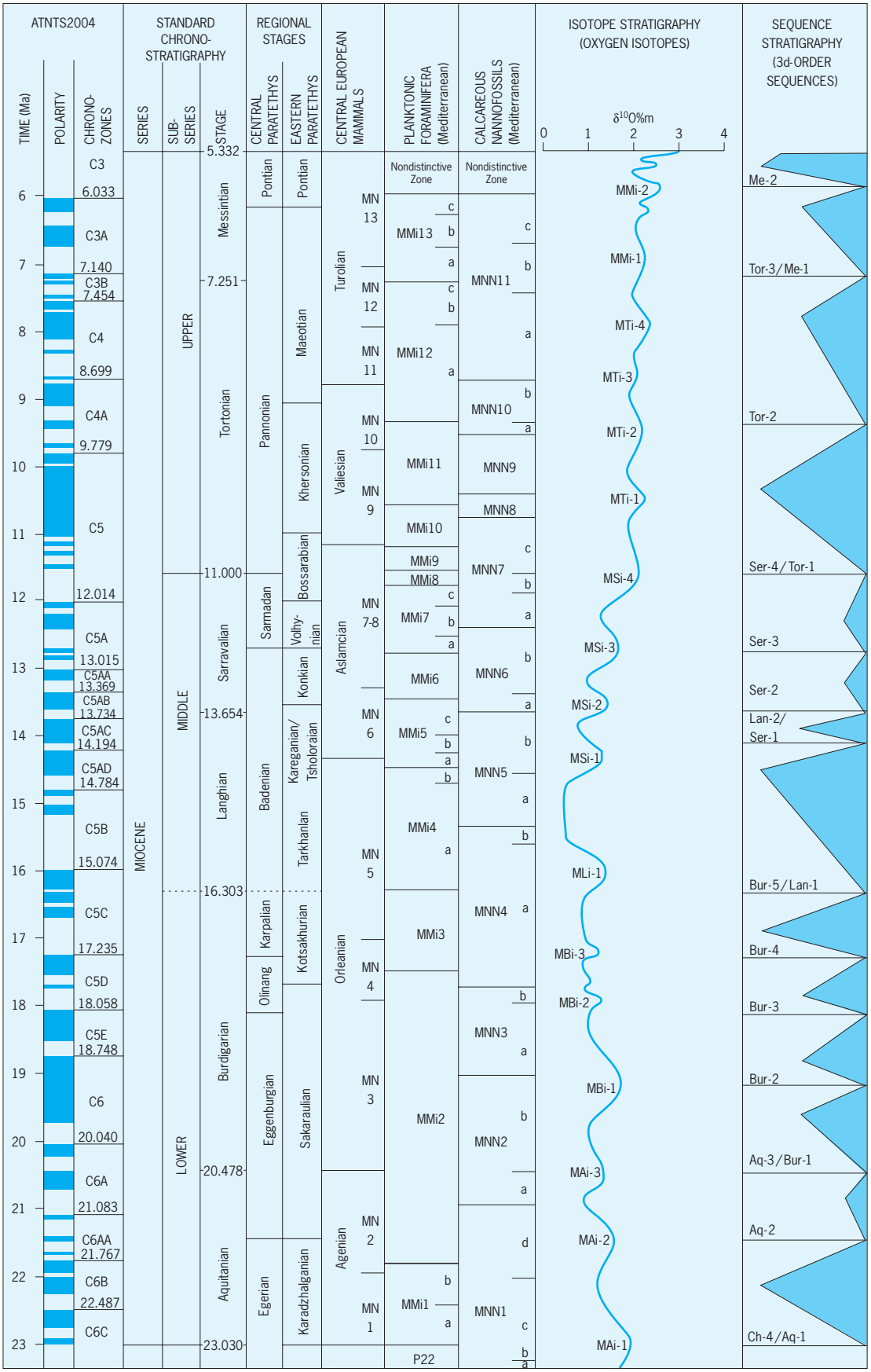

Fig. 1. Time scale of the Miocene.

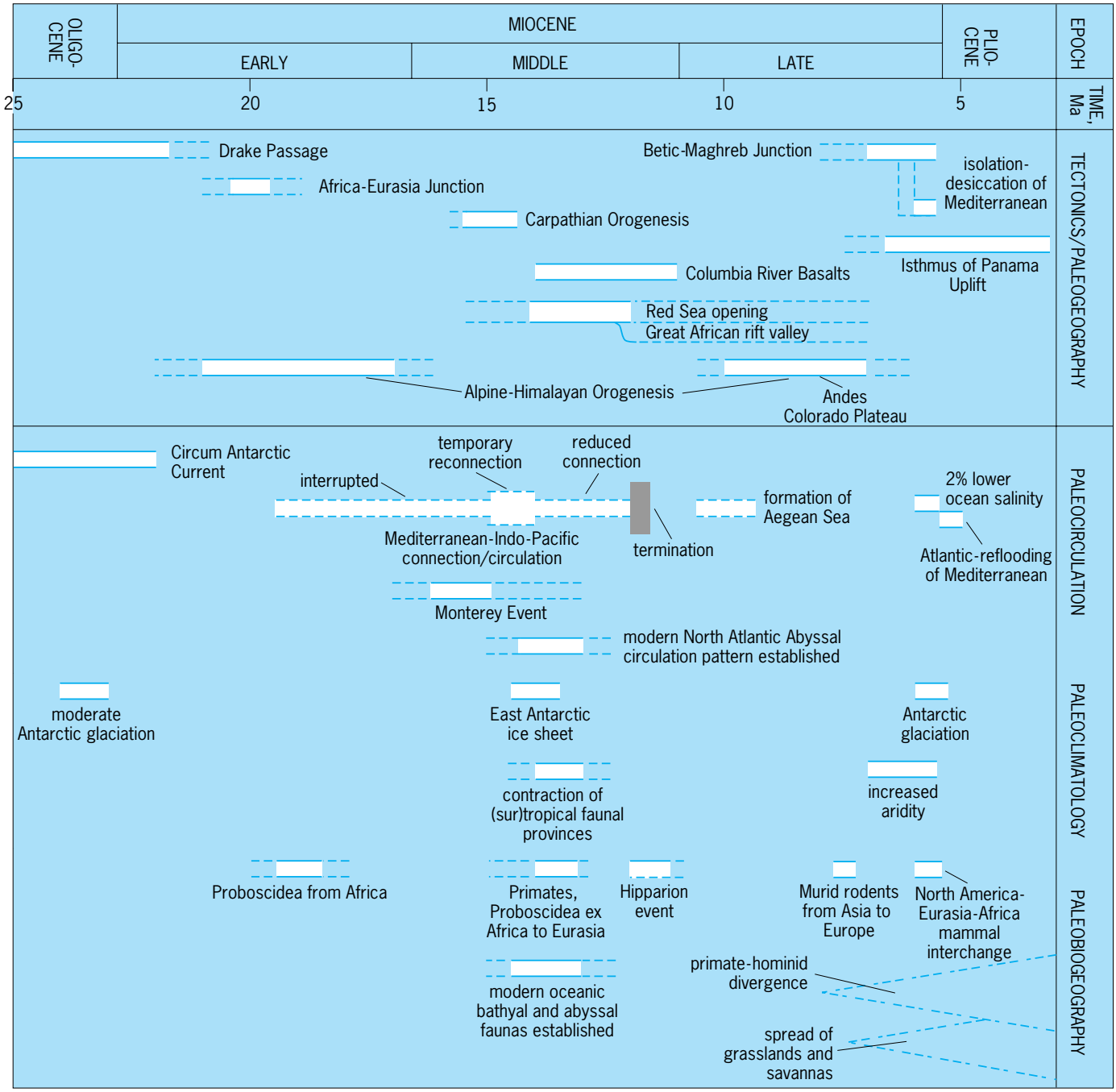

Fig. 2. Major events in the Miocene Epoch.

Hellenic mountain chain; and the eventual separation about 13–14 Ma of the Tethys Ocean into two inland seas: the Miocene (ancestral) Mediterranean (to the south) and the Paratethys (to the north). The Paratethys, which had existed as a shallow marine arm of the Tethys since the late Eocene–early Oligocene, underwent gradual reduction in salinity during the middle Miocene and local development of evaporite (salt) deposits reflecting continued compression, and eventually evolved into a series of both interconnected and separate brackish lakes. Concomitantly, early Miocene clockwise rotation of Africa away from Arabia widened the southern part of the Red Sea (which had opened up during late Oligocene time), and (at least intermittent) communication with the Indian Ocean was maintained throughout the Miocene probably by way of the Afar Depression. The Gulf of Suez and Gulf of Aqaba/Eilat opened during the middle and late Miocene. The southward extension of this major rift zone into the heart of East Africa formed the Great African Rift Valley, an incipient mid-ocean ridge in the making. *See* CONTINENTS, EVOLUTION OF; MID-OCEANIC RIDGE; RIFT VALLEY.

The Colorado Plateau was uplifted from near sea level to nearly its present height of 5000–6000 ft (1500–1800 m) during the late Miocene and Pliocene, and the Colorado River began its downcutting, resulting in Grand Canyon, a feature that exposes 2 billion years of geologic history. The Basin

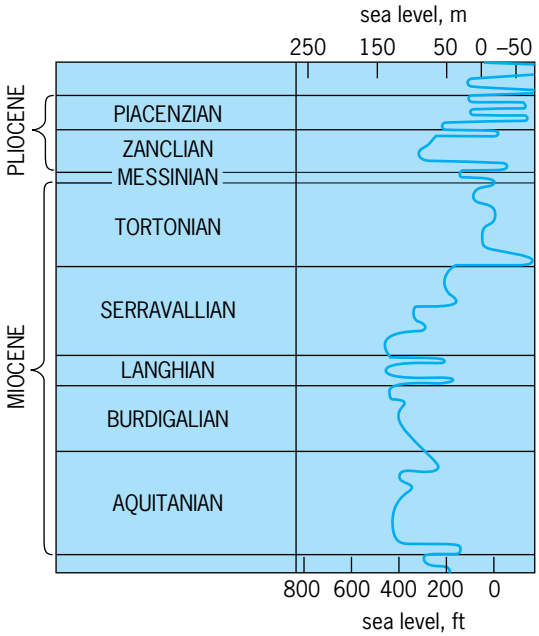

Fig. 3. Eustatic sea-level curve for the Neogene. The zero represents present sea level.

and Range Province of North America was subjected to early Miocene (rhyolite) volcanism, and block faulting and basaltic volcanism during the late Miocene have combined to give the region its characteristic linear topography of alternating mountains separated by broad valleys.

One of the most spectacular volcanic deposits can be found in the Pacific Northwest: the Columbia River basalts. Erupted during a 3–4-million-year span in the middle Miocene, these basalt flows extend over 50,000 mi^2 (200,000 km^2) in eastern and central Washington and have a maximum thickness of over 8250 ft (2500 m). *See* BASALT; VOLCANOLOGY.

Resources. The Miocene is the source for major deposits of phosphates that are mined in North Carolina and Florida as well as extensive deposits of early-middle Miocene diatomaceous earth and phosphates in California which formed as a result of enhanced primary productivity in the eastern Pacific Ocean. Significant oil and gas are produced from Miocene rocks, particularly in the Gulf of Mexico and the west coast basins off California, northern South America, the Indonesian archipelago, and to a lesser extent the Arabian Peninsula and Gulf and northern Iraq and eastern Syria, among other places. *See* NATURAL GAS; PETROLEUM; PHOSPHATE MINERALS.

Ocean currents and climate. Ocean circulation essentially assumed its modern form during the Miocene as enhanced refrigeration in the form of growth of the Antarctic Ice Sheet plunged the Earth inexorably deeper into an icehouse state, although there were some details that were completed during the succeeding Pliocene and Pleistocene epochs. An ice cap has been present on Antarctica, at least intermittently, since at least the early Oligocene (about 34 Ma). The opening of the Drake Passage between South America and Antarctica took place during the latest Oligocene–early Miocene (about 25–23 Ma), allowing the unhindered circulation of ocean currents around the Antarctic continent. The development of the Circum-Antarctic Current thermally isolated high southern latitude waters and the continent of Antarctica from the warmer, low-latitude waters and resulted in the replacement of calcareous oozes (comprising planktonic foraminifera and calcareous nannoplankton) by biosiliceous oozes (diatoms and radiolarians). *See* GLACIAL EPOCH.

During the late early Miocene (about 16 Ma) a significant increase in biologic productivity occurred, and increased extraction and burial of organic carbon in oceanic sediments is interpreted to have lowered atmospheric levels of carbon dioxide and cooled global climates (reverse greenhouse). This has been called the Monterey Event because large amounts of organic matter were deposited around the perimeter of the northern Pacific and particularly in California, off which they are the main source rock for petroleum in the Monterey Formation. An alternative explanation for this event suggests that Early Miocene (about 21–17 Ma) Himalayan uplift caused increased chemical weathering and riverine nutrient flux to the sea, resulting in enhanced productivity and enhanced deposition of organic rich sediments. The latter theory suggests that drawdown of carbon dioxide (and attendant global cooling) is ultimately due to tectonics (and resulting silicate weathering) rather than organic carbon burial and that enhanced delivery of nutrients to the ocean, rather than stronger upwelling, is the driving force behind organic burial.

Global cooling did follow shortly after the Monterey Event. Some 2 million years later, during the Middle Miocene (about 14 Ma) major expansion and permanent emplacement of the East Antarctic Ice Sheet resulted in significant planetary cooling and constriction of tropical and subtropical belts and associated faunas. During the latest Miocene (about 7–5.5 Ma) another major expansion of the ice cap, comparable to, or somewhat less than, present-day volume, resulted in a global sea-level fall of some 230 ft (70 m), extensive glaciation as far away as Patagonia, and development of widespread aridity in midlatitudes of western North America, Australia, and Asia (India, Pakistan). This event coincided with, and is believed to have played an essential role in, the latest Miocene isolation of the Mediterranean Basin from the global ocean and its evaporation and transformation into a desert 2 mi (3 km) below sea level. The subsequent reconnection of the Mediterranean and Atlantic following a rise in sea level at the beginning of the Pliocene and the resulting southward flow of warm, salty surface water from the Mediterranean may have been responsible for the early Pliocene deglaciation of the West Antarctic Ice Sheet.

In the Northern Hemisphere, evidence suggests that ice was present in the polar regions as reflected in the form of glacial dropstones in North Atlantic marine sediments of late Miocene age (about 7 Ma), while tillites suggest the presence of intermittent glaciers at about 10 Ma (early late Miocene) in the Wrangell Mountains of Alaska.

It has been suggested that late middle and late

Miocene (about 10–7 Ma) renewed uplift elevated the Himalayan Plateau to a height where it was able to disrupt the global circulation system, modify the monsoonal airflow over southwestern Asia, and push the Earth climatic system across a delicate threshold which led to the development of permanent Northern Hemisphere glaciation in the mid-Pliocene (about 2.5 Ma). The globally significant expansion of savannah and grasses at this time and the increase in aridity in Africa and development of the arid-adapted flora of the Sahara, which appears to date to about 6–5 Ma (or slightly later), appear to be a consequence of the late Miocene Himalayan uplift. *See* PALEOCLIMATOLOGY; PALEOCEANOGRAPHY.

Sea-level changes. The predominantly regressive tendency of the Oligocene was replaced in the Miocene by an essentially transgressive early Miocene followed by a regressive middle and late Miocene phase (Fig. 3). A significant transgression occurred near the base of the Miocene, and a second led to the highest sea-level stand of the Miocene during the early middle Miocene. Two rapid sea-level falls in the early middle Miocene were followed by a gradual sea-level lowering during the late middle Miocene, which was punctuated in turn by a major, rapid fall near the middle/late Miocene boundary. Major, rapid sea-level falls bracket the terminal Miocene Mediterranean salinity crisis, followed by a rapid sea-level rise at the base of the Pliocene. Indeed, the boundaries of (at least the lower and lower middle) standard Miocene stage as delineated in Europe can be viewed as corresponding to, and reflecting, the major changes in sea level seen on the global sea-level curve.

Tectonic uplift along the Iberian (Spanish) and African continental margins is considered to have been the main cause of the closure of two marine seaways in the Betic and Rifean areas. Global sea-level lowering, associated with an interval of Antarctic glaciation with peak glacials at 5.79 and 5.75 Ma, may have aided in further isolation of the Mediterranean Basin; but it is not considered to have played a significant role in the initial isolation of the Mediterranean Basin from the global ocean inasmuch as these peak glacial stages postdate the initiation of the Messinian Salinity Crisis by ~200 kyr. Peak glacials at 5.79 and 5.75 Ma may have significantly lowered the global sea level, isolated the Sea of Japan, and temporarily turned into a freshwater lake.

Terminal Miocene salinity crisis. About 6 Ma the northwestern part of Africa collided with Europe, closing off the western end of the Mediterranean to the Atlantic. Together with a global lowering of sea level shortly thereafter of approximately 230 ft (70 m), caused by enhanced glaciation on Antarctica, the Mediterraean and Paratethys seas were isolated from the global ocean system and became large, inland lakes of (predominantly) interior drainage with parallel but distinct geologic histories as one of the most dramatic events in geologic history took place over the next half million years. While about 6% of the salt form the world's ocean was extracted, resulting in a lowering of ocean salinity by about 2%, Tethyan faunas emigrated out of the Mediterranean to seek temporary safe haven in the Atlantic as the Mediterraean became hypersaline and deposited some 6600–10,000 ft (2–3 km) of salt, gypsum, and anhydrite over an area of 40,000 mi^2 (10^6 km^2) in a series of discontinuous hypersaline playa lakes and sabkhas (like the present-day Dead Sea or Great Salt Lake) over 10,000 ft (more than 3000 m) below sea level, eventually becoming a gigantic desert like Death Valley. The main difference between the two is that Death Valley is only slightly below sea level. This event has become known to geologists as the Messinian (or Terminal Miocene) Salinity Crisis. Since the Mediterranean salt deposits are over 1.2 mi (2 km) thick, and only 230 ft (70 m) of salt would be produced by a single isolation and evaporative event, intermittent (quasicontinuous) connection with the Atlantic Ocean must have been maintained, probably through either narrow passage (or both) in the Rifian Corridor (northwestern Morocco) or the Betic Straits (southwestern Spain). The Gibraltar Straits did not exist at this time. *See* ANHYDRITE; GYPSUM; HALITE; PLAYA.

The Messinian Stage of the Mediterranean consists of three distinct sedimentary deposits: a lower, normal marine unit exhibiting signs of gradual isolation from the global ocean (7.26–5.96 Ma); a middle unit consisting of two sets of evaporite deposits, the Lower and Upper Evaporite Series, deposited between 5.96 and 5.6 Ma and 5.52 and 5.45 Ma, respectively; and between the cessation of evaporite deposition in the Mediterranean (5.45 Ma) and the early Pliocene deposition of normal marine marls of the Trubi Formation (5.33 Ma)—that is, during the terminal Miocene (Messinian)—a series of nonmarine sediments characterized by a peculiar unique and indigenous megafauna (bivalves, or pelecypods, of the genus *Congeria* and cariids) and microfauna (ostracodes, particularly *Cythereis* and *Loxoconcha djafarovi*) [**Fig. 4**]. These deposits, interpreted as having been formed in brackish-water lakes, have been named Lago Mare—the Italian equivalent of the French term applied for a long time to the middle and late Miocene (Paratetehyan) Lac Mer. While the Mediterranean was undergoing evaporation, the Black Sea was transformed into a shallow brine lake (but in which no salt deposits have ever been found). The typical Lago Mare fauna of the upper part of the Mediterranean Messinian Stage has also been found in the Black Sea. The general interpretation of the transformation from evaporite deposition to the oligo-mesohaline conditions of the Lago Mare deposits was the results of a rearrangement of the drainage pattern of Europe, such that the Mediterranean captured the riverine runoff from the humid regions of central Europe and the Paratetehyan (Black Sea) drainage via the Aegean region. The subsequent infilling of the Mediterranean "bathtub" by the Atlantic occurred at the beginning of the Pliocene Epoch; it was probably abrupt and catastrophic. However, the recent discovery of a fauna of marine euryhaline (95% gobies and trachinoids) and euryhaline and fully marine stenohaline (5% combtooth blennies, gray mullets, herrings, lanternfishes, porgies, pupfishes, silversides,

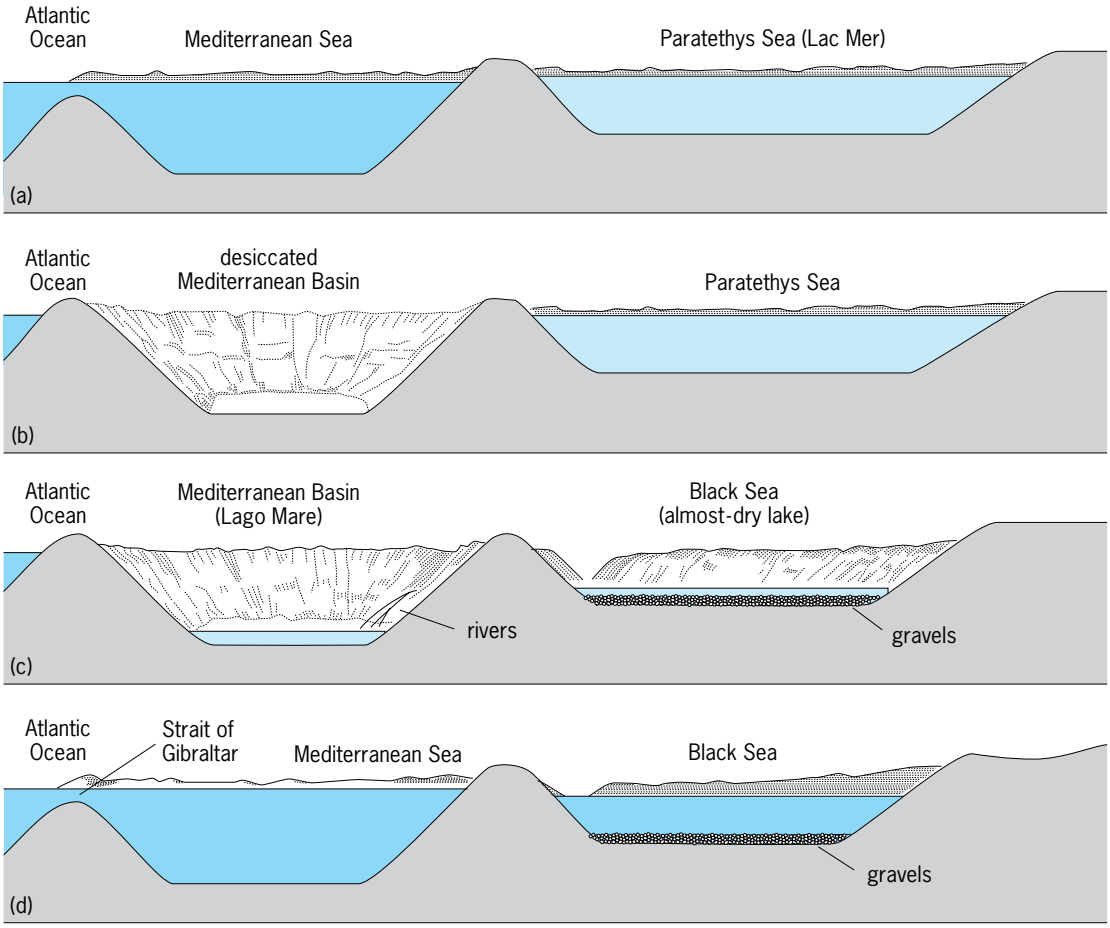

Fig. 4. Changes in Mediterranean Sea and contemporaneous development of the Black Sea. (*a*) Middle Miocene; about 15 Ma. (*b*) Late Miocene; about 6 Ma. (*c*) Late Miocene; about 5.5 Ma. (*d*) Pliocene; about 5 Ma.

toadfishes, viviparous brotulas) fishes belonging to at least 12 teleostean families in the upper part of Lago Mare deposits in Tuscany, Italy, has suggested a modification of long-held beliefs on the nature of the terminal Miocene depositional setting and history of the Lago Mare and of the (supposed) subsequent early Pliocene infilling of the Mediterranean "bathtub," which is considered to have been probably abrupt and catastrophic. The discovery of a diverse marine fish fauna in the Lago Mare deposits indicates that normal marine conditions had been established at least during the upper part of the ~0.10-kyr-long Lago Mare deposition and ~0.10 kyr before the Miocene/Pliocene boundary at 5.33 Ma. In this revised scenario, the Lago Mare brackish ostracode faunas are interpreted as local indicators of marginal basin conditions in a basin that had already undergone the reestablishment of normal marine conditions with the open ocean. The Miocene/Pliocene boundary at 5.33 Ma witnessed the further development of open marine conditions in the Mediterranean Basin with the formation of normal, deep-sea chalk deposits—the Trubi Marls—characterized by normal open ocean marine calcareous protists (planktonic foraminifera and calcareous nanoplankton), and was the results of ongoing tectonic activity inasmuch as the next significant glaciation is recorded at ~4.9 Ma (that is, early Pliocene).

Since the Tethyan Ocean had been cut off from the Indo-Pacific earlier in the Miocene, the coral reefs that thrived in the western Tethys (southern Spain) until the onset of the terminal Miocene evaporative stage never returned to the Mediterranean, and the Mediterranean today is characterized by its low diversity of fauna and low nutrient levels. During the latest Miocene (Messinian) interval of lowered sea level, major rivers around the Mediterranean cut gorges several thousand feet deep into granite in the Nile Valley, and comparable gorges, which have been traced down to the edge of the abyssal plain of the sea floor, have been documented in the Rhone Valley of southern France as far north (180 mi or 300 km from the coast) as Valence, in the Po Valley nearly up the Alps, on the islands of Corsica and Sardinia, and in the eastern (Syria, Israel) and southern (Libya, Algeria) margins of the Mediterranean Sea. *See* BLACK SEA.

Life. During the Miocene, life assumed much of its modern aspect. The spread of grasses and weeds throughout this epoch, but particularly in the late Miocene, and concomitant reduction in and thinning of forests reflected the global Neogene cooling as the Earth entered deeper into an ice house state. In this environment, snakes, frogs, and murids (rats, mice) expanded in diversity and habitat; songbirds reflect the expansion of seed-bearing herbs and, like frogs,

the concomitant diversification of insects, many of which are found entombed in middle Miocene amber from the Dominican Republic. Grazing animals (elephants, rodents, horses, camelids, and rhinos, for example) developed high crowned teeth to resist significant wear caused by silicon fragments in the developing grasses. Some animals assumed gigantic proportions such as *Baluchitherium*, a Eurasian rhino that stood 16 ft (5 m) at the shoulders, and the tallest camel known, a giraffelike form that was over 12 ft (3.5 m) tall.

In at least two adaptive radiations, early proboscideans (elephant family) and mastodonts migrated into Europe from Africa following the docking of Africa and Eurasia about 18 Ma. Over a dozen families of small mammals migrated to Africa, including insectivores and rodents, while larger migrants included ancestors of modern deer, cattle (bovids), antelope, pigs, and large forms such as chalicotheres, rhinos, and early giraffes. Felid-like carnivores (the cat family) originated in the Oligocene Epoch, ~35 Ma. Modern cat species originated in Asia in the late Miocene, ~11 Ma, and replaced the earlier, ancestral creodonts that originated in the Paleogene, with the divergence of the *Panthera* lineage ~10.8 Ma, followed shortly by the bay cat lineage at ~9.4 Ma. An early migration event at ~8.5–5.6 Ma brought an early progenitor of the caracal lineage to Africa, and a second migration at ~8.5–8.0 Ma saw the immigration of a common ancestor to five felid lineages (domestic cat, leopard cat, lynx, ocelot, and puma) to North America via the Bering Land Bridge. A concomitant immigration of Eurasian carnivores (mustelids, procyonids, ursids, and saber-toothed felids) is thought to have crossed from Eurasia to North America. During a migration event about 15 Ma, early apes spread out into the woodland and forests of Eurasia, while hyaenids and shrews infiltrated Africa.

In North America, grazing animals were among the earliest ruminants (camelids, horses), and these were joined shortly afterward by Eurasian migrants by way of the Bering Land Bridge, including over 50 genera of ungulates and large carnivores in the late Miocene. In the late Miocene (about 7 Ma), savannah grasslands underwent a significant areal expansion and major faunal changes occurred which coincided with the interval of the Terminal Miocene Salinity Crisis. Extinctions exceeded replacements, so that the main result was a diminution of faunal diversity by the early Pliocene. *See* MAMMALIA.

Placental ungulates reached South America in the early Paleocene, and by Miocene time they had evolved into a variety of forms and sizes. South America remained isolated from North America for over 100 million years and, with the exception of an apparently brief interval in the Paleocene when minor exchange took place between the two continents, little or no faunal interchange took place between the two until the latest Miocene–early Pliocene. However, parallel evolution took place among many of the terrestrial vertebrates, a good example being the South American Miocene forms *Thoatherium* and *Diadiaphorus*, which bore a striking resemblance to horses of North America. Large carnivorous marsupials were common, including wolflike boryhaenids.

The Neogene evolution of higher primates (anthropoids) was played out in the Southern Hemisphere primarily as late Paleogene cooling associated with modified tropical environments and climates forced the southward displacement of most primates. Old World monkeys evolved in Europe and Asia; New World monkeys evolved in South America. Hominoids [including hyalobatids (gibbons) and panids (African apes, or gorillas, and chimpanzees) and hominids (humans) evolved in Africa, while the orangutan evolved in Asia.

Two groups of apelike hominoids evolved during the Miocene: dryomorphs and ramapithecines (and their late Miocene Asian descendants, also known as sivapithecines). Dryomorphs were small primates that browsed in early Miocene forests and grasslands of East Africa. Ramapithecines and sivapithecines constitute an extinct group of predominantly Eurasian apes that probably evolved from an early Miocene African ancestor.

The relatively free interchange between Eurasia and Africa between 18 and 12 Ma appears to have come to an end in the late Miocene, after which (about 8 Ma) the hominoids of Eurasia and Africa appear to have followed separate and independent lines of evolution: the pongids in Asia, and the panids and hominoids leading eventually to the true hominids in (predominantly East) Africa. This scenario has been linked, in turn, with the development of the East African Rift (and its northward extension into the Red Sea and Gulf of Suez), which would have served as a geographic barrier allowing independent evolution toward forest (panid) and savannah (hominid) adapted forms. With the late Miocene change in climate (7–5 Ma) to cooler, drier conditions and the spread of open savannah and grasslands, monkeys came to dominate the African forest at the expense of dryomorphs. There is a gap in the terrestrial fossil record during this interval of time, and it is only in the early Pliocene (about 4 Ma) that the story of human evolution resumes with the discovery of the earliest true hominids (australopithecines) in East Africa, about 1 million years older than the australopithecine footprints of Laeotili and the skeletons of Lucy and other australopithecines at Hadar in Ethiopia at about 3 million years. Within the past few years, however, there have been reports of fossils of stem hominids in the latest Miocene (~7–5 Ma) in Africa, but these finds await further confirmation. *See* AUSTRALOPITHECINE; FOSSIL HUMANS.

In the marine realm, major radiation of mammals including walruses, seals, sea lions, and whales occurred during the early Miocene. Carnivorous sperm whales with large teeth were among the earliest representatives of the group; and baleen whales, long-nosed dolphins, and sharks, including the large gigantic white shark *Carcharadon*, were among distinctive early Miocene forms. On the sea floor, large bivalve mollusks of the scallop family thrived in the early Miocene, and a distinct horizon of large

pectenids occurs in lower Miocene rocks of Europe and North America and in corresponding levels in the deposits of the Paratethyan Sea in east-central Europe and at least as far to the east as Iran, attesting to an interval of global climatic amelioration. Among the protozoa, planktonic foraminifera experienced a major radiation in the early and middle Miocene following the drastic reduction in diversity during the middle and late Eocene, some 15–20 million years earlier. Mangroves and coral reefs flourished in a circumequatorial belt spanning the Indo-Pacific and Caribbean regions, but the latter were eliminated from the Mediterranean during the terminal Miocene Salinity Crisis, never to return with early Pliocene flushing from the Atlantic. Siliceous oozes expanded and encircled Antarctica (and to a lesser extent the North Pacific) beginning in the early Miocene. This reflected the global cooling of the Neogene and the development of the Circum-Antarctic Current in the early Miocene, which isolated high southern latitudes and the Antarctic continent from the warmer ocean currents to the north for the remainder of the Neogene. *See* FORAMINIFERIDA; MANGROVE; MOLLUSCA; REEF.

A major biogeographic reorganization of tropical, shallow-water corals occurred over a several-million-year span in the early Miocene. About half of these shelf-edge-dwelling reef corals became regionally extinct in the Caribbean, with the majority surviving to the present day in the Indo-Pacific region. Upwelling of cold, nutrient-rich waters, reflected in contemporaneous circum-Caribbean phosphorite deposits, are thought to have been responsible for the local elimination of the corals. An essentially contemporaneous reduction in diversity (and regional extinction) of tropical, photosymbiont-bearing, larger benthic protozoa occurred as well, with a combination of local upwelling (Caribbean) and tectonic closure of ancient seaways (eastern Tethys) having probably played major roles. *See* EXTINCTION (BIOLOGY); GEOLOGIC TIME SCALE; PALEOECOLOGY; PALEONTOLOGY.

W. A. Berggren

Bibliography. G. Carnevale et al., Mare versus Lago-Mare fishes and the Mediterranean environment at the end of the Messinian Salinity Crisis, *J. Geol. Soc.*, 163(1):75–80, 2006; K. J. Hsu, When the Black Sea was drained, *Sci. Amer.*, 258(5):52064, 1978; K. J. Hsu, When the Mediterranean Sea dried up, *Sci. Amer.*, pp. 27–36, December 1972; W. Krijgsman et al., Revised astrochronology for the Ain el Beida section (Atlantic Morocco): No glacio-eustatic control for the onset of the Messinian Salinity Crisis, *Stratigraphy*, 1(1):87–101, 2004; S. R. Stanley, *Earth and Life Through Time*, 1986.

Mira

The first star recognized to have a periodic brightness variation. Mira (officially designated Omicron Ceti, in the constellation Cetus, the Whale) was discovered in 1596 by David Fabricius (or Fabricus), a clergyman and an amateur astronomer from Esens in East Friesland (now in Germany), while he was searching for Mercury. He mistook it for a nova because it later faded from view. However, he saw it reappear 14 years later. Not until 1638 was it recognized to be the first known variable star, when Johann Holwarda, from Franeker, also in Friesland (now in the Netherlands), rediscovered it, and in 1639 determined its period to be 11 months. In 1642, Johannes Hevelius of Danzig, who also observed the star, called it Mira, meaning The Wonderful. Mira is the prototype of an entire class of Mira-type pulsating long-period variables. Although it once resembled the Sun, Mira has evolved into a cool red giant star that is at the end of its life. *See* GIANT STAR; STELLAR EVOLUTION.

Contracting and expanding every 332 days, Mira typically varies in visual brightness from about magnitude 3.4 (brightest) to about 9.3 (faintest). However, an individual maximum can sometimes be as bright as second magnitude, while at other times the maximum may reach barely fifth magnitude. Mira is the brightest long-period variable because it is intrinsically bright and only 420 light-years from the Earth. Angular diameter measurements such as those from the Very Large Telescope (VLT) Interferometer show that the diameter of Mira varies from about 330 to 400 times the diameter of the Sun during its pulsational cycle, and has an average size larger than the orbit of Mars. *See* MAGNITUDE (ASTRONOMY).

Mira is also known as a symbiotic variable: a long-period variable star with a close hot compact companion star. The companion, VZ Ceti (or Mira B), forms a close visual double with Mira (or Mira A; the pair is often known as Mira AB). VZ Ceti was also discovered in 1959 by G. van Biesbroeck to be a variable star in its own right. It is a hotter, bluer, burned-out star, called a white dwarf, surrounded by material captured from Mira's wind. VZ Ceti varies in brightness between magnitude 9.5 and 12.0, and it is only about the size of the Earth. When VZ Ceti is bright, it may affect the apparent brightness of Mira, particularly when Mira is faint. The orbital period is about 498 years. *See* BINARY STAR; SYMBIOTIC STAR; WHITE DWARF STAR.

The *Chandra X-ray Observatory* has observed Mira and its companion (see **illustration**). The separation between Mira and its companion has been measured to be currently about 70 times more than that between the Earth and the Sun (about twice the distance between Pluto and the Sun, and equal to an angular size of 0.6 arcsecond).

The *Chandra* image shows both Mira A and its hot companion, Mira B. Mira A is losing matter from its upper atmosphere via a stellar wind. The image shows a faint bridge of this material between the two stars. Mira A is bright in this x-ray image because it was undergoing a flare; normally, red giants are very faint at x-ray wavelengths. Mira B is likely surrounded by a disk of the accreted material, which is likely the cause of its strange shape in this image.

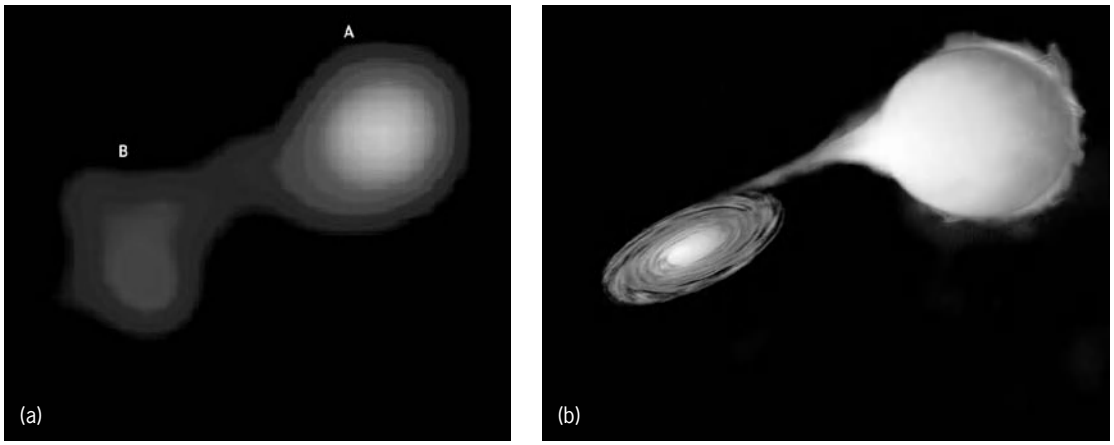

Mira and its companion. (*a*) *Chandra* image, taken in x-ray wavebands, of Mira (labeled A) and its companion (*B*). (*b*) Artist's conception of the Mira AB system. (*M. Karovska, Harvard-Smithsonian Center for Astrophysics; NASA*)

At about 420 light-years from the Earth, Mira is the nearest wind-accreting binary system. *See* STAR; VARIABLE STAR. Arne Henden; Janet Akyüz Mattei

Bibliography. D. Hoffleit, History of the discovery of Mira stars, *J. Amer. Ass. Variable Star Observers*, 25:115–136, 1997; J. E. Isles, Mira's 400th Anniversary, *Sky Telesc.*, 91(2):72–73, February 1996; M. Karovska, High resolution observations of Mira, *J. Amer. Ass. Variable Star Observers*, 25:75–79, 1997; G. van Biesbroeck, Mira as a double star, *Publ. Astron. Soc. Pacific*, 71:462–465, 1959; H. C. Woodruff et al., Interferometric observations of the Mira star o Ceti with the VLTI/VINCI instrument in the near-infrared, *Astron. Astrophys.*, 421:703–714, 2004.

Mirage

A distorted scene of distant objects resulting from the passage of light through a nonuniform medium. In a uniform atmosphere light would travel in straight lines and distant scenes would appear without distortion. Since the density of the air changes with height and with temperature, the atmosphere is not uniform and light rays will follow curved paths. The direction of the curvature is to deflect a ray away from a warmer (less dense) layer of air. **Figure 1** shows some of the features of the most familiar desert or hot-road mirage. The ground is heated by the sun and in turn heats (and rarifies) the layer of air just above the ground. Then rays passing near the ground are bent (refracted) upward. The observer can see the top of the tree in the figure, for example, by looking in two directions. The strongly refracted rays produce an inverted view of the tree. The only other place in nature where such an effect can be seen is in reflection from a water surface, and it is natural to interpret the mirage as evidence of a body of water.

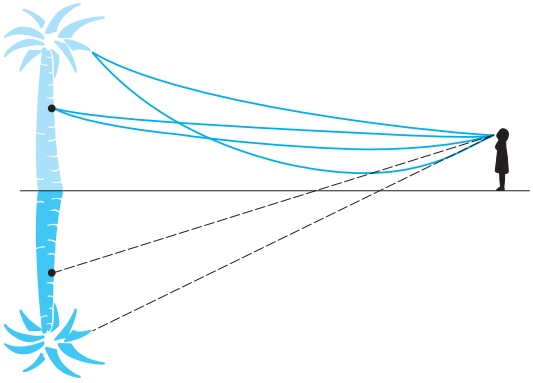

Fig. 1. Rays that give rise to the inverted image in the desert mirage. (*After R. Greenler, Rainbows, Halos, and Glories, Cambridge University Press, 1980*)

Fig. 2. Mirage over the frozen Arctic Ocean, resulting from an inversion layer. (*From R. Greenler, Rainbows, Halos, and Glories, Cambridge University Press, 1980*)

Sometimes a warm layer of air can be trapped between layers of cooler air overhead (a temperature inversion), resulting in light-ray paths that curve downward, so that distorted images of ships or land that normally would lie hidden below the horizon can be seen (**Fig. 2**). *See* METEOROLOGICAL OPTICS; REFRACTION OF WAVES; TEMPERATURE INVERSION. Robert Greenler

Mirror optics

The use of plane or curved reflecting surfaces for the purpose of reverting, directing, or forming images. The most familiar use of reflecting optical surfaces is for the examination of one's own reflected image in a flat or plane mirror. A single reflection in a flat mirror produces a virtual image which is reverted or reversed in appearance. The use of one or more

reflecting surfaces permits light or images to be directed around obstacles, with each successive reflection producing a reversal of the image. A curved mirror, either spherical or conic in form, will produce a real or virtual image in much the same manner as a lens, but generally with reduced aberrations. There will be no chromatic aberrations since the law of reflection is independent of the color or wavelength of the incident light. *See* ABERRATION (OPTICS); OPTICAL IMAGE.

An optical surface which specularly reflects the largest fraction of the incident light is called a reflecting surface. Such surfaces are commonly fabricated by polishing of glass, metal, or plastic substrates, and then coating the surface of the substrate with a thin layer of metal, which may be covered in addition by a single or multiple layers of thin dielectric films. The law of reflection states that the incident and reflected rays will lie in the plane containing the local normal to the reflecting surface and that the angle of the reflected ray from the normal will be equal to the angle of the incident ray from the normal. This law is a special case of the law of refraction in that the angles rather than the sines of the angles of incidence and reflection are equal. Formally, this relation is commonly used in calculations by setting the effective index of refraction prior to incidence on the surface. When this concept is introduced, all of the formulas relating to lenses are applicable to reflective optics. In this article, however, the imaging relations will be described in the most appropriate form for reflecting surfaces. *See* GEOMETRICAL OPTICS.

Plane mirrors. The formation of images in the plane mirrors is easily understood by applying the law of reflection. **Figure 1** illustrates the formation of the image of a point formed by a plane mirror. Each of the reflected rays appears to come from a point image located a distance behind the mirror equal to the distance of the object point in front of the mirror. In Fig. 1, the face of the observer can be considered as a set of points, each of which is imaged by the plane mirror. Since the observer is viewing the facial image from the object side of the mirror, the face will appear to be reversed left for right in

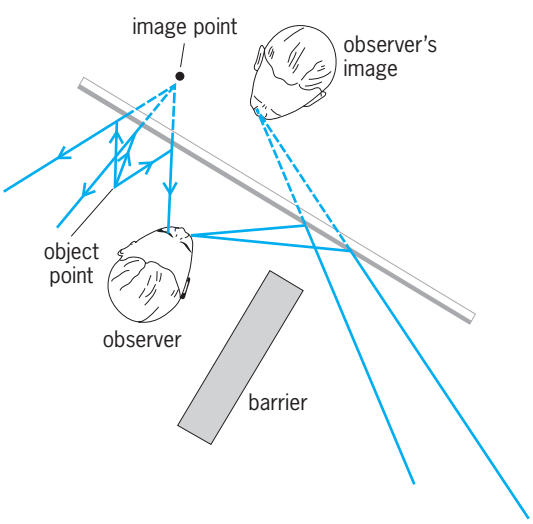

Fig. 1. Formation of images by a plane mirror.

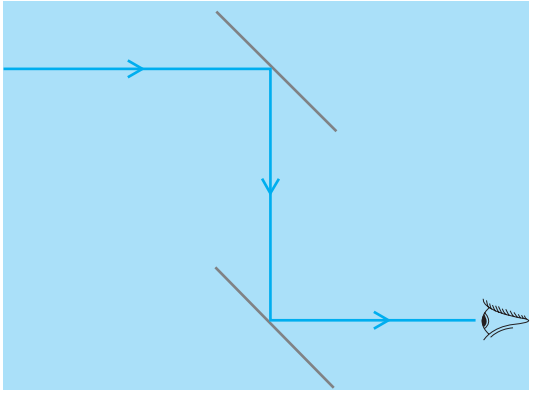

Fig. 2. Simple mirror periscope.

the virtual image formed by the mirror. Such a virtual image cannot, of course, be projected on a screen, but can be viewed by a lens, in this case the eyes of the observer. Figure 1 also indicates the redirection of light by a plane mirror, in that a viewer who cannot observe the object point directly can observe the virtual image of the point formed by the mirror. A simple optical device which is based on this principle is the simple mirror periscope (**Fig. 2**), which uses two mirrors to permit viewing of scenes around an obstacle. In this case two reflections are present and provide an image which is correctly oriented, and not reverted, to the observer. The property of reversion in a complicated mirror system depends upon the location and view direction of the observer, as well as the number of reflections that take place and the orientation of the planes through which the light is directed. *See* PERISCOPE.

Prisms. These are solid-glass optical components that use reflection at the faces to provide redirection of the optical pencils passing through them. The advantage of the use of a prism is that the reflecting surfaces are maintained in accurate location with respect to each other by the integrity of the glass material making up the body of the prism. Difficulties with prisms are that very homogeneous glass is required since the light may make many passes through the prism, and that a prism is optically equivalent to insertion of a long block of glass into the imaging system. The insertion of such a glass block often results in a system which is mechanically shorter in space, but the aberration balance of the imaging system is changed, frequently requiring a redesign of the associated optical components in order to accommodate the increased glass path. It is obvious that the use of solid glass prisms may introduce more weight into the optical design than the equivalent metal mounts required for a similar arrangement of mirrors in air. In certain cases, the angle of incidence on a reflecting surface within the prism may exceed the critical angle of incidence, and no reflective coating may be required on such a surface. **Figure 3** shows some common types of optical prisms with reflecting surfaces. The applications of such prisms range from simple redirection of light to variable angle of rotation of the image passed through the prism and binocular combination of images. A special case of reflection is the use of a beam splitter which permits

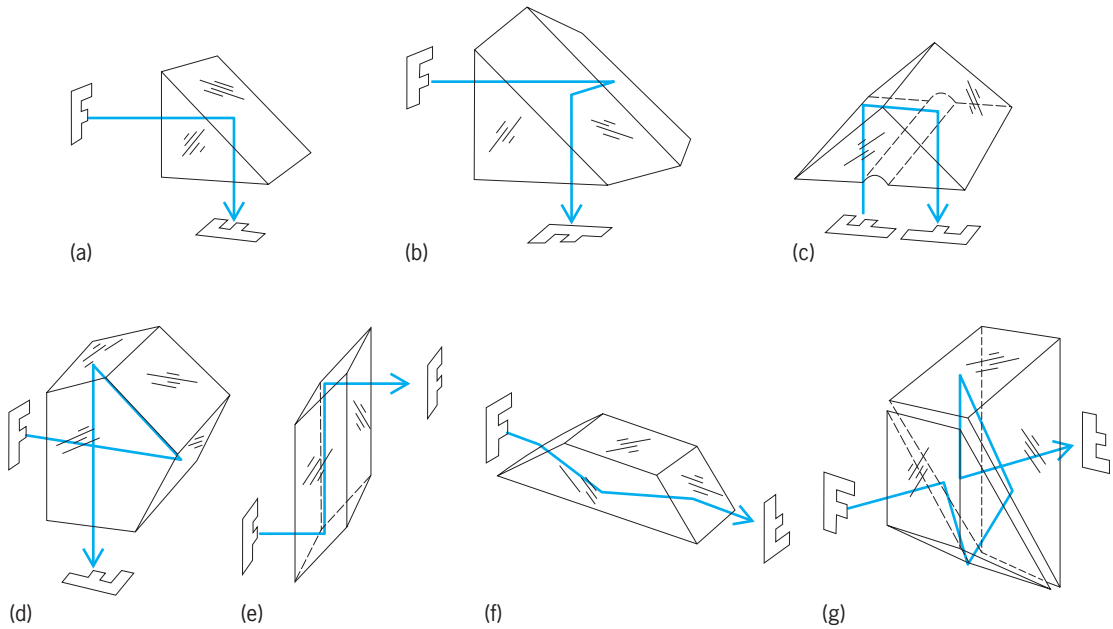

Fig. 3. Prism types. (*a*) Right-angle. (*b*) Amici roof. (*c*) Porro. (*d*) Pentaprism. (*e*) Rhomboid. (*f*) Dove. (*g*) Pechan.

the splitting of light, or the combining of two beams by the use of a surface which is partially reflecting and partially transmitting. *See* BINOCULARS; OPTICAL PRISM.

Spherical mirrors. These are reflecting components that are used in forming images. The optics of such mirrors are almost identical to the properties of lenses, with their ability to form real and virtual images. In the case of spherical mirrors, there is a reversal of the direction of light at the mirror so that real images are formed on the same side of the mirror as the object, while virtual images are viewed from the object side but appear to exist on the opposite side of the mirror. Both concave and convex spherical mirrors are commonly encountered. Only a virtual erect image of a real object will be formed by a convex mirror. Such mirrors are commonly used as wide-angle rearview mirrors in automobiles or on trucks. The image formed appears behind the mirror and is greatly compressed in space, with a demagnification dependent on the curvature of the mirror. A concave spherical mirror can form either real or virtual images. The virtual image will appear to the observer as erect and magnified. A common application is the magnifying shaving mirror frequently found in bathrooms. A real image will be inverted, as is the real image formed by a lens, and will actually appear in space between the observer and the mirror.

Figure 4 shows the formation of real and virtual images by a spherical mirror. The equation which applies to all of the image relations is given below.

$$\frac{1}{S'} + \frac{1}{S} = \frac{2}{R}$$

The distances S and S' are measured from the surface of the spherical mirror; when either is negative, a virtual image is formed behind the mirror. The constant R is the radius of curvature of the mirror. The magnification of the image is the ratio of the image distance S' to the object distance S.

Conic mirrors. These are a special case of the spherical mirror with improved image quality. A spherical mirror will form an image which is not perfect, except for particular conjugate distances. The use of a mirror which has the shape of a rotated conic section, such as a parabola, ellipsoid, or hyperboloid, will form a perfect image for a particular set of object-image conjugate distances and will have reduced aberrations for some range of conjugate relations. Two of the most familiar applications for conic

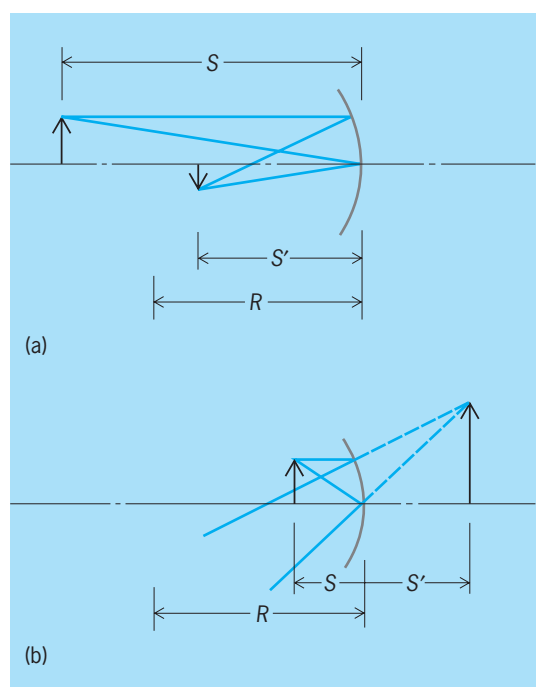

Fig. 4. Formation of images by spherical mirror. (*a*) Real image. (*b*) Virtual image.

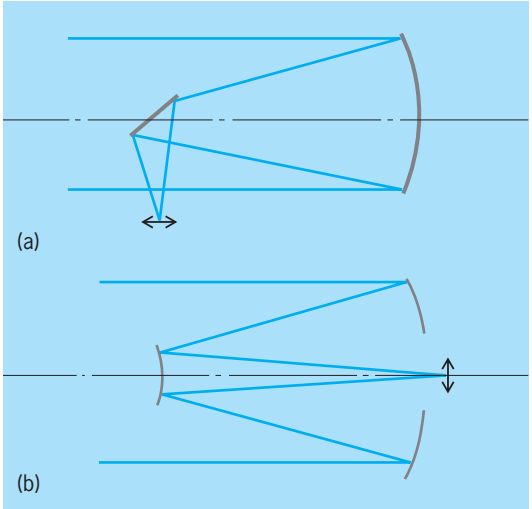

Fig. 5. Applications of conic mirrors. (*a*) Newtonian telescope. (*b*) Cassegrain telescope.

mirrors are shown in **Fig. 5**. Figure 5*a* shows the use of a paraboloid of revolution about the optical axis to form the image of an object at an infinite distance. In this drawing the image to be viewed by the observer at the eyepiece is relayed to the side of the telescope tube by a flat folding mirror in what is called a newtonian form of a telescope. This demonstrates one of the difficulties that is found with the use of reflecting optical components to form real images; namely, that the image must often be relayed out of the incident path on the image-forming mirror, otherwise the observer will block some of the light from the object. Not all reflecting systems carry out this relaying in the same manner. The Cassegrain system uses a curved secondary mirror to achieve magnification of the final image while allowing the image to fall outside the telescope barrel.

Figure 5*b* shows the use of a paraboloid as the primary mirror with the image relayed to the final image location by a hyperbolic secondary mirror. Such a use of two mirrors permits the construction of a long-focal-length telescope within a relatively short space. This latter form, usually referred to as a Cassegrain telescope, serves as the principal type of modern reflecting astronomical telescope. One of the advantages of this type of telescope design is the freedom from chromatic aberration that would be present in a refracting telescope. *See* OPTICAL SURFACES.

Mirror coatings. The reflectivity of a mirror depends on the material used for coating the reflecting surface. The conventional coatings for glass mirror surfaces are silver or aluminum, which are vacuum-deposited or sputtered onto the surface. In some cases, chemical deposition will be used. Most mirrors intended for noncritical uses, such as looking glasses or wall mirrors, will have the reflecting metallic coating placed on the back side of the glass, thus using the glass to protect the coating from oxidation by the atmosphere. Mirrors for most critical or scientific uses require the use of front-surface reflectors, with the reflecting coating on the exposed front surface of the glass. In this case, a hard overcoat of a thin layer of silicon dioxide is frequently deposited over the metal to protect the delicate thin metal surface. The reflectivity of the mirror with respect to wavelength depends on the choice of the metal for the reflector and the material and thickness of material layers in the overcoat. In some cases, a fully dielectric stack will be used as a reflecting coating with special spectral selective properties to form a dichroic beam splitter, as in a color television camera, or as an infrared-transmitting "cold mirror" for a movie projector illumination system. *See* REFLECTION OF ELECTROMAGNETIC RADIATION.

R. R. Shannon

Misophrioida

A copepod order of free-living crustaceans made up of 34 species in 16 genera and 3 families. *Misophria pallida* was the first species described, by Axel Boeck in 1865, from a sample taken close to the sea floor in shallow water. A second genus and species, *Benthomisophria palliata*, was established by Georg Ossian Sars in 1909 for a bathypelagic species from the North Atlantic Ocean. The remaining species have been described since 1964 and reflect an increased interest by taxonomists in this order of copepods. Many of the recently described species have been recovered during the exploration of anchialine caves, which are coastal marine water bodies with subsurface rather than surface hydrologic connections to the sea. Anchialine caves are one category of crevicular habitat in coastal marine geological formations. The number of known species of Misophrioida is expected to increase significantly, as similar habitats, such as the interstices of coral reefs, are explored more systematically. *See* COPEPODA; CRUSTACEA.

Description. Adult females are 0.6–5.7 mm (0.02–0.23 in.) long and males 0.6–3.7 mm (0.02–0.15 in.); the length range includes small species living close to the sea floor and the larger, bathypelagic species. The cuticle of the adult body is distinctly ornamented. The second thoracic somite, bearing the first swimming leg, articulates with the first somite. However, a posterior extension of the cuticle of the first thoracic somite completely covers the second thoracic somite so that these two somites appear unarticulated (see **illustration**). The last two thoracic somites are associated with the posterior part of the body (the podoplean architecture). The internal anatomy of *B. palliata* includes a heart, but the naupliar (larval) eye, a normal feature of adult copepods,

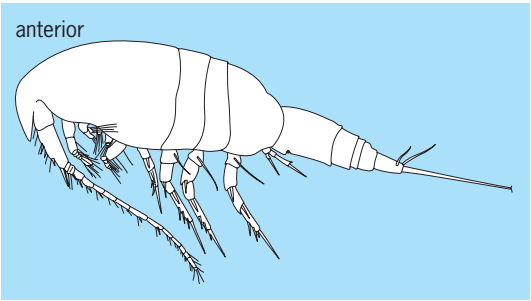

Side view of typical misophrioidan.

is absent. The excretory system is an antennal gland, apparently retained from the naupliar phase of development, rather than a maxillary gland, which is the usual excretory system for postnaupliar stages of copepods. An embryo carried by a female of *M. pallida* hatched to a nauplius (a crustacean larval stage), which molted directly to the first juvenile copepodid. The addition of body somites and limb segments during the copepodid phase of development of *B. palliata* follows a pattern common for many copepods. The vertical distribution of copepodid stages of the bathypelagic *B. palliata* appears uniform between 2500 and 4000 m (8200 and 13,100 ft).

Phylogeny. The origin of the Misophrioida from either shallow marine waters directly over the sea floor or marine waters of anchialine caves continues to be debated. In phylogenetic analyses, the Misophrioida almost invariably are the least derived of the podoplean (the last two thoracic somites with the postenor part of the body) copepod orders, with the gymnoplean calanoids (only the last thoracic somite with the posterior part of the body) as their closest older order. This relationship is based on the presence in misophrioids of a heart, a podoplean body architecture, and a multisegmental configuration of several limb branches. However, the hypothesis has not been examined using the contemporary logic of phylogenetic systematics. *See* CALANOIDA.

Frank D. Ferrari

Bibliography. A. Boeck, Oversigt over de ved Norges Kyster jagttagne Copepoder henhörende til Calanidernes, Cyclopidernes og Harpactidernes Familier, *Forhandl. Vidensk.-Selsk.*, Christiania, vol. 1865; G. A. Boxshall, On the anatomy of the misophrioid copepods, with special reference to *Benthomisophria palliata* Sars, *Phil. Trans. Roy. Soc. London*, ser. B, vol. 297, 1982; G. A. Boxshall and D. Jaume, On the origin of misophrioid copepods from anchialine caves, *Crustaceana*, vol. 72, 1999; G. A. Boxshall and D. Jaume, Discoveries of cave misophrioids (Crustacea: Copepoda) shed new light on the origin of anchialine faunas, *Zool. Anz.*, vol. 239, 2000; G. A. Boxshall and H. S. J. Roe, The life history and ecology of the aberrant bathypelagic genus Benthomisophria Sars, 1909 (Copepoda: Misophrioida), *Bull. Brit. Mus. Nat. Hist.* (*Zool.*), vol. 38, 1980; R. Gurney, Notes on some Copepoda from Plymouth, *J. Mar. Biol. Ass. UK*, vol. 19, 1933; P. Martínez-Arbizu and D. Jaume, New hyperbenthic species of *Misophriopsis* and *Misophriella*, first record of misophrioid copepods (Crustacea) from Antarctic waters, *Helgoland Mar. Res.*, vol. 53, 1999; G. O. Sars, Note préliminaire sur trois formes remarquables de Copépodes, provenant des campagnes de S.A.S. le Prince Albert de Monaco, *Bull. Inst. Océanog.*, Monaco, vol. 147, 1909.

Mississippian

The fifth period of the Paleozoic Era. The Mississippian System (referring to rocks) or Period (referring to time during which these rocks were deposited) is employed in North America as the lower (or older)

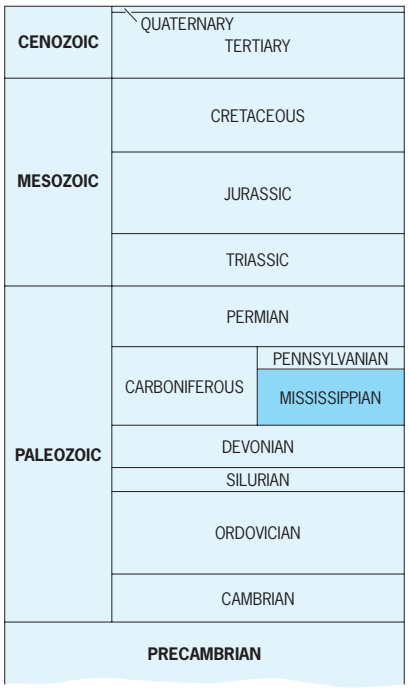

subdivision of the Carboniferous, as used on other continents. The name Mississippian is derived from rock exposures on the banks of the Mississippi River between Illinois and Missouri.

The limits of the Mississippian Period are radiometrically dated. Its start (following the Devonian) is dated as 345–360 million years before the present (Ma). Its end (at the start of the next younger North American period, the Pennsylvanian) is dated as 320–325 Ma. The duration of the Mississippian is generally accepted as 40 million years (m.y.). Biochronologic dating within the Mississippian, based on a combination of conodont (phosphatic teeth of eel- or hagfish-like primitive fish), calcareous foraminiferan, and coral zones permits a relative time resolution to within about 1 m.y. *See* CARBONIFEROUS; PENNSYLVANIAN.

Subdivisions. The Mississippian is divided, in ascending order, into the Lower Mississippian, comprising the Kinderhookian and Osagean, and the Upper Mississippian, comprising the Meramecian and Chesterian. Kinderhookian, Osagean, Meramecian, and Chesterian are used in North America as series (for rocks) and as stages (for time). In Illinois, Valmeyeran is commonly used for the Osagean and Meramecian combined. The type localities for the Kinderhookian, Meramecian, and Chesterian are in the upper Mississippi Valley, whereas the type locality of the Osagean is in western Missouri.

Conodont zones are widely accepted as subdivisions of the Mississippian (**Fig. 1**) and are used for intercontinental correlation. The 10 conodont zones that subdivide the Kinderhookian and Osagean have been proposed as global standard zones, whereas the seven conodont zones that subdivide the Meramecian and Chesterian are used mainly in the United States. Correlated to these zones by detailed biostratigraphic studies are 15 foram (calcareous foraminiferan) zones and 13 coral zones (Fig. 1).

This combined, conodont-foram-coral biochronology provides the best tool for precise correlation within North America. *See* STRATIGRAPHY.

Lithofacies and paleogeography. During much of Mississippian time, the central North American craton (stable part of the continent) was the site of an extensive marine carbonate platform on which mainly limestones and some dolostones and evaporites were deposited. This platform extended either from the present Appalachian Mountains or Mississippi Valley to the present Great Basin, as exemplified by the reconstructed map of this area in mid-Mississippian (middle Osagean) time (**Fig. 2**). The paleoequator passed through Nevada and Wisconsin. Paralleling it to the south was the Transcontinental arch, composed of a peninsula and several islands (extending from Wisconsin Highlands to Zuni-Defiance Island) that separated the carbonate platform into halves. *See* DOLOMITE ROCK; LIMESTONE; SALINE EVAPORITES.

The Mississippian continent was bounded on its present east, south, and west sides by deep troughs and basins in which mainly siliciclastic sediments—siltstones, sandstones, and conglomerates—were deposited as debris produced by erosion of bordering orogenic highlands. In the east, these siliciclastic rocks were deposited as the Pocono clastic wedge and seaward Borden deltaic complex. In the west, the Antler highlands shed clastic sediments into an adjacent trough. More distal parts of the subsiding troughs, which were not overwhelmed by clastics, were the sites of sediment-starved basins, exemplified by the Deseret basin and Caballos-Ouachita trough (Fig. 2), in which mainly shales, cherts, and phosphorites were slowly deposited.

Most Mississippian limestones are bioclastic (composed of parts of organisms), except in the bordering troughs where they are largely calciclastic (turbiditic calcarenites, wackestones, and conglomerates), eroded from preexisting limestones, transported, and then redeposited.

The predominant Mississippian rock type on the carbonate platform is crinoidal limestone or encrinite formed from the skeletal remains of crinoids, blastoids, and other echinoderms. In fact, the Mississippian has been called the Age of Crinoids. The second most common carbonate rock is oolite, formed by aggregates of ooids (sand-size spheres composed of chemically precipitated concentric grains). Another important rock type is very fine grained limestone termed micrite because of its microscopic grains. Evaporitic rocks, mainly anhydrite and some halite, were also deposited in peritidal environments of the Madison shelf (Fig. 2) and Williston basin.

In Kinderhookian and Osagean times, reeflike mudmounds were deposited on the slopes between the carbonate platform and adjacent basins. Fine-grained sediments in these mudmounds were held in place by organisms which, although not preserved, were probably mats of algae or bacteria. Crinoids growing on these mudmounds contributed coarser debris to the sediments. *See* ANHYDRITE; CRINOIDEA; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; ECHINODERMATA; HALITE; MARINE SEDIMENTS; OOLITE.

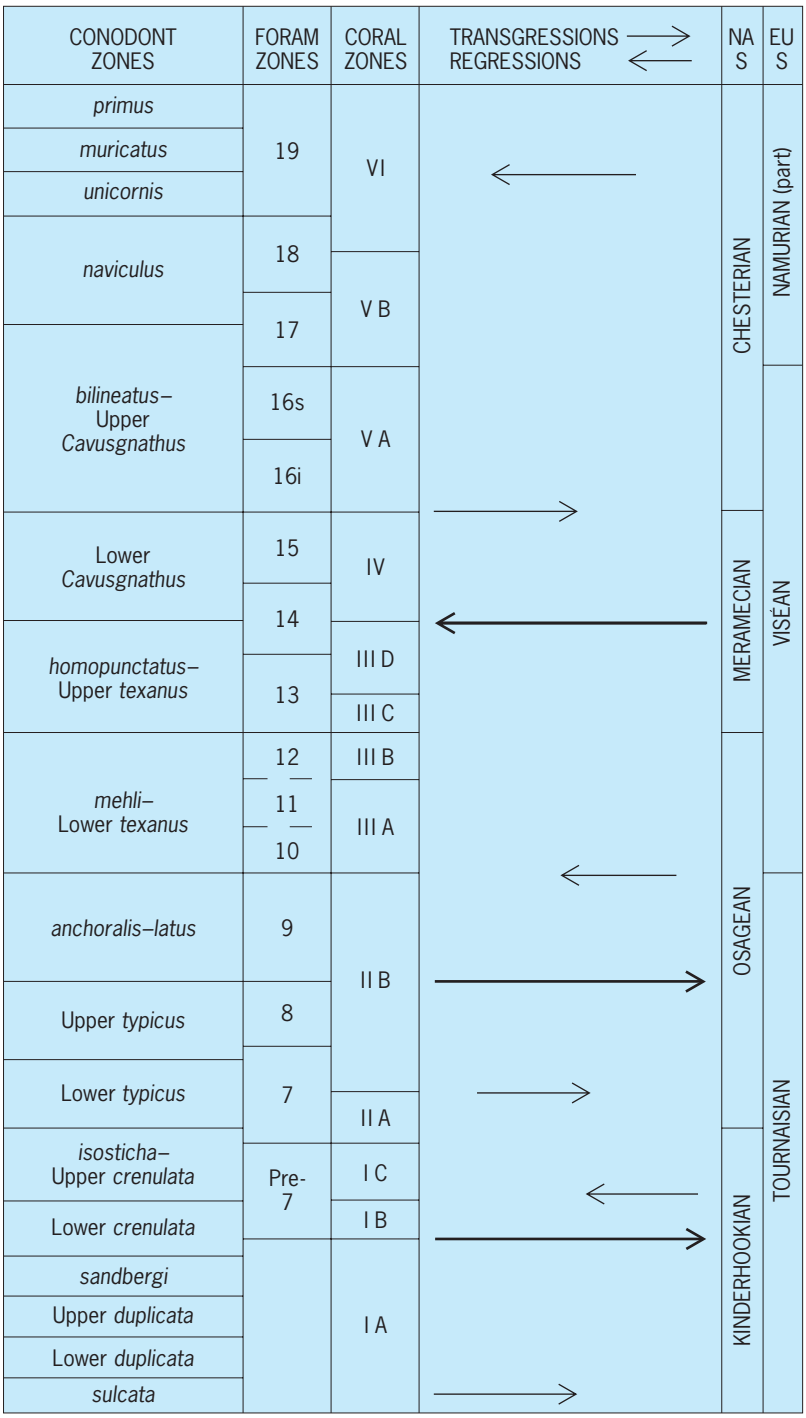

Fig. 1. Conodont-foram-coral biochronology, sea-level changes, and rock series subdivisions of the Mississippian in North America. NA S, North American Series; EU S, European Stages. Length and weight of arrows show relative scale of the major North American sea-level rises (transgressions) and falls (regressions).

Sandstones and siltstones predominate in the Borden deltaic complex (Fig. 2) and in the younger Humbug deltaic complex, which filled the former Deseret basin in Meramecian time. Coarse-to-fine conglomerates composed of boulders, cobbles, and pebbles of deeper marine, older Paleozoic rocks that had been folded and elevated to form the Antler highlands intertongue with shales and argillites in the adjacent flysch trough. *See* SANDSTONE; SEDIMENTARY ROCKS.

In the Deseret basin, fine siliceous rocks termed

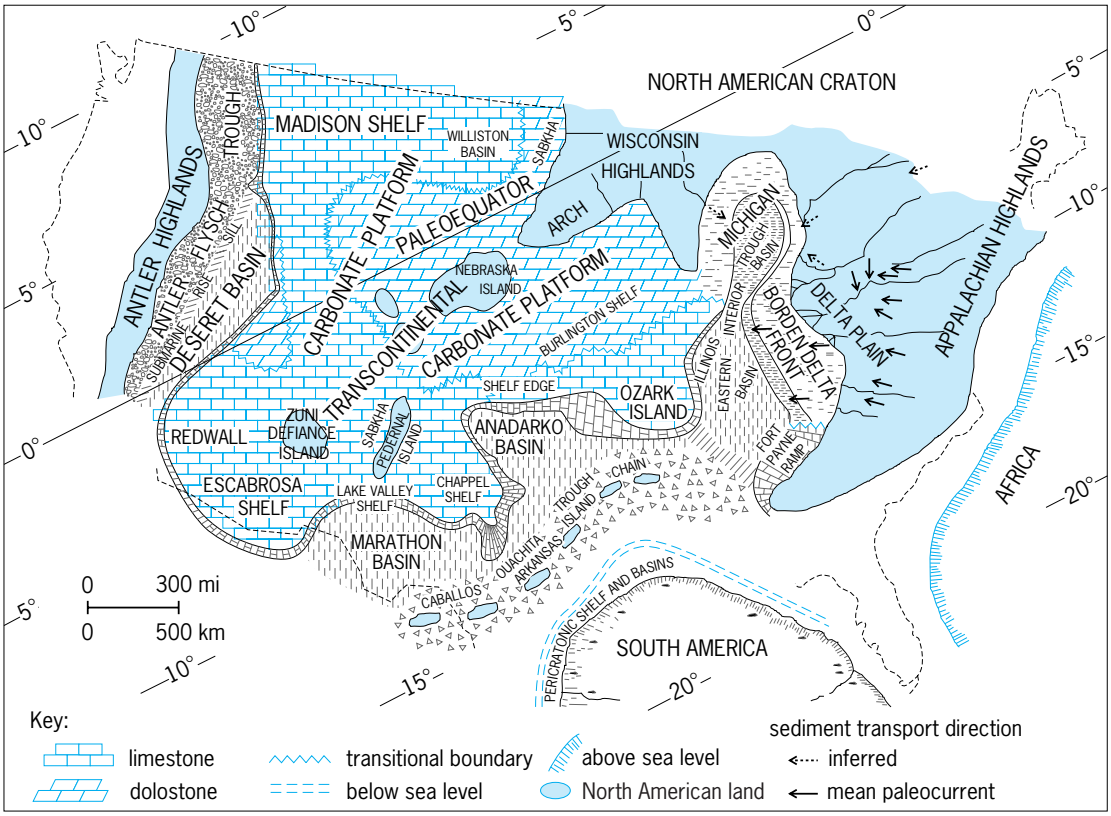

Fig. 2. Mid-Mississippian (middle Osagean) paleogeographic and lithofacies map of conterminous United States showing the widespread carbonate platform, continental spine (Transcontinental arch), bounding highlands and basins, and position of paleoequator. (*After R. C. Gutschick and C. A. Sandberg, Mississippian continental margins of the conterminous United States, Soc. Econ. Paleont. Mineral. Spec. Publ., no. 33:79–96, 1983*)

cherts are black, well bedded, and formed mainly of preserved and dissolved radiolarian tests. The cherts are commonly phosphatic and are interbedded with thicker beds of phosphatic shale and siltstone and thinner beds of oolitic and peloidal phosphorite. Thicker, lighter-colored siliceous rocks in the Caballos-Ouachita trough are termed novaculites.

Terrestrial Mississippian rocks, particularly in the Pocono and Antler clastic wedges, contain thin coal beds. These coal beds are economically insignificant in contrast to those in Pennsylvanian rocks. *See* COAL.

Paleoceanography. The North American craton was covered by shallow, warm, tropical epicontinental seas that had maximum depths of only about 60 m (200 ft) at the shelf edge. A chain of islands occupied the Transcontinental arch between the Madison Sea on the west and the Eastern Interior Sea on the east (Fig. 1). The end of the arch plunged southwestward into the Redwall-Escabrosa Sea. Depths in the Antler trough far exceeded 500 m (1640 ft), whereas those in the Caballos-Ouachita and Eastern Interior troughs exceeded 300 m (1000 ft). *See* PALEOCEANOGRAPHY.

Paleotectonics. During the Mississippian, the North America continent was subjected to three plate-tectonic events—the Acadian, Antler, and proto-Ouachita orogenies—that created orogenic highlands and inboard, deep foreland troughs. These structural features influenced sedimentation on its present-day eastern, southern, and western margins. First, the European and African continental plates collided on the east beginning in Middle Devonian time, producing the Acadian orogeny. Next, the decoupled proto-Pacific oceanic plate collided on the west beginning in latest Devonian time, producing the Antler orogeny. At about the same time, the South American plate began moving toward North America, producing the proto-Ouachita orogeny. This movement created the Caballos-Arkansas island chain (Fig. 1) as a tectonic forebulge during the Mississippian, although actual continental collision did not take place until Pennsylvanian time. *See* OROGENY; PLATE TECTONICS; SEDIMENTOLOGY.

Sea-level changes. Mississippian North America was subjected to a number of sea-level rises, associated with transgressions, and sea-level falls, associated with regressions (Fig. 1). The two major rises, which were probably eustatic (referring to worldwide change of sea level), took place during the late Kinderhookian lower *crenulata* zone and at the start of the middle Osagean *anchoralis–latus* zone. These sea-level rises caused progradation or seaward migration of carbonate platforms as organisms produced buildups that maintained their niches relative to sea level. The rises also caused stratification of the water column in deeper basins, so that bottom conditions became deficient to lacking in oxygen. *See* ANOXIC ZONES; BASIN.

A major sea-level fall occurred at the start of

the middle Meramecian lower *Cavusgnathus* zone (Fig. 1). This fall caused sediments to bypass the carbonate platforms and to be deposited offshore as proximal deltaic deposits and distal prodeltaic deposits. Later, with increased sea-level fall, the carbonate platforms were subaerially exposed and subjected to karstification and cave formation. *See* CAVE; DELTA; KARST TOPOGRAPHY.

The Mississippian, which began with a moderate transgression that continued from latest Devonian time, ended with a moderate regression during the *muricatus* and *primus* zones.

Conodont biofacies and zonation. Conodonts are the phosphatic teeth of primitive fishes that had soft bodies, which rarely are preserved in rocks. These hard parts remained on the sea bottom after the remains of the conodont-bearing animal had decayed, and were generally subjected to postmortem sorting by currents. However, conodonts were only occasionally transported outside their original, broad paleotectonic setting. Different conodont faunas of the same age, which lived in various paleoenvironmental settings, are termed conodont biofacies. Contemporaneous dating of conodont biofacies is enabled by the few conodonts that lived at the fringes of, or were transported just outside, their preferred habitats.

Conodont zonation has become the major biochronologic tool for dating Mississippian rocks. Conodonts lived in most marine environments and many were distributed worldwide. Most genera and species were rapidly evolving and short-ranging, especially in warm tropical waters. Most other organisms evolved slowly, had restricted habitats, were not commonly preserved, or became disarticulated after death and hence difficult to identify.

The Late Devonian and Early Mississippian were the mid-Paleozoic heyday of conodont evolution, but the number and diversity of taxa began to decline gradually beginning with the extreme sea-level fall at the start of the Meramecian lower *Cavusgnathus* zone. The conodont zonation of the Kinderhookian and Osagean is considered a global standard against which other zonations are measured (Fig. 1). However, Meramecian and Chesterian conodont zones are less firmly established and represent longer time intervals. *See* CONODONT.

Life. Crinoids were probably the most abundant biota in Mississippian seas, but are only uncommonly used for correlation because most specific identifications require study of calyxes, which generally disarticulated after death. Crinoid and similar blastoid columnals are major constituents of Mississippian carbonate rocks. *See* CRINOIDEA.

Corals are probably the most widely preserved Mississippian megafossil group, and they are one of the most useful biochronologic tools for constructing biostratigraphic zones for carbonate-platform rocks. The earliest Mississippian is characterized mainly by solitary corals and tubular colonial corals known as *Syringopora*, which had survived the late Frasnian mass extinction. Other colonial corals began a gradual return late in the Kinderhookian. The Late Mississippian was a heyday for reef-building colonial corals and large solitary corals. *See* CORALLINALES; REEF.

Other commonly occurring Mississippian megafossils are the brachiopods, mollusks, bryozoans, and ostracodes. Thick-valved brachiopods, such as spiriferids and productids, are found in abundance in carbonate-platform and upper-slope rocks; thin-valved brachiopods occur in basinal rocks. Among the mollusks, cephalopods, particularly the ammonoids, are most important for biochronologic dating and correlation. Gastropods (snails) and pelecypods (bivalves) occur in most biofacies, but taxa are long-ranging and not generally useful for correlation. Fenestrate (meshlike) and ramose (branched) bryozoans are found in shallower-water rocks. Ostracodes are ubiquitous in all rock types, and deeper-water species locally provide excellent biostratigraphic tools. *See* BRACHIOPODA; BRYOZOA; CEPHALOPODA; MOLLUSCA; OSTRACODA.

Small, phosphatic fish remains (ichthyoliths), including teeth, scales, bones, and dermal denticles, are found in almost all marine Mississippian rocks.

Radiolarians and calcareous and agglutinated foraminferans are common marine Mississippian microfossils. Pelagic radiolarians are useful in dating deep-water rocks. Calcareous forams are important in dating most carbonate-platform and upper-slope rocks. Agglutinated forams (those that build their shells from detritus) provide a biofacies tool. Sponge spicules are common in basinal rocks. Worm jaws (scolecodonts) occur in some fine-grained rocks. *See* FORAMINIFERIDA.

Trace fossils (ichnofossils), including grazing traces, trails, burrows, and fecal pellets, are common in most rocks. Their occurrences are controlled by biofacies, and hence they are useful as depth indicators. Tracks of early amphibians are scarce. *See* TRACE FOSSILS.

Among marine plants, bacterial and algal mats were important in binding sea-bottom sediments, but their remains may not be preserved. However, calcareous algae, calcispheres, algal stromatolites, and trochiliscid charophytes occur commonly in carbonate-platform rocks.

Forests flourished during the Mississippian, and tree trunks, plant stems, roots, and spores occur commonly in terrestrial and peritidal rocks, particularly in coal beds. *Lepidodendron* trunks and *Stigmaria* roots are among the best-known plant remains. *See* GEOLOGIC TIME SCALE; PALEONTOLOGY; PALEOZOIC.

Charles A. Sandberg

Bibliography. E. Atherton, C. Collinson, and J. A. Lineback, Mississippian System: Handbook of Illinois Stratigraphy, *Ill. St. Geol. Surv. Bull.*, 95:123–163, 1975; L. C. Craig and C. W. Connor (coords.), Paleotectonic investigations of the Mississippian System in the United States, *USGS Prof. Pap.*, no. 1010, 1979; R. C. Gutschick and C. A. Sandberg, Mississippian continental margins of the conterminous United States, *Soc. Econ. Paleont. Mineral. Spec. Publ.*, 33:79–96, 1983; F. G. Poole and C. A. Sandberg, Mississippian paleogeography and conodont biostratigraphy of the western United States, *Pacific Section, Soc. Econ. Paleont. Mineral. Spec. Publ.*, 67:107–136, 1991.

Mistletoe

The name given to several species of the mistletoe family (Loranthaceae). The true mistletoe of Europe is *Viscum album*, and among the early nations this was an important ceremonial plant, which probably accounts for the origin of the custom of kissing under the mistletoe. In the United States, the common representative of the group is *Phoradendron flavescens* (see **illus.**). All of the mistletoes are green hemiparasites; that is, they obtain water and minerals from the host plant but manufacture their own food. *See* SANTALALES.

Mistletoe (*Phoradendron flavescens*).

Perry D. Strausbaugh; Earl L. Core

Mitochondria

Specialized compartments (organelles) of all eukaryotic (nucleated) cells that use oxygen (**Fig. 1**). Often called the powerhouses of the cell, mitochondria are responsible for energy generation by the process of oxidative phosphorylation. In this process, electrons produced during the oxidation of simple organic compounds are passed along a chain of four membrane-bound enzymes (the electron transport or respiratory chain; **Fig. 2**), finally reacting with and reducing molecular oxygen to water. The movement of the electrons releases energy that is used to build a gradient of protons across the membrane in which the electron transport chain is situated. Like a stream of water that drives the turbines in a hydroelectric plant, these protons flow back through adenosine triphosphate (ATP) synthase, a membrane-bound enzyme that acts as a molecular turbine. Rotation of part of ATP synthase results in storage of energy in the form of ATP, the universal energy currency of the cell.

Besides their role in energy generation, mitochondria house numerous enzymes that carry out steps essential to metabolism. Defects in mitochondrial assembly or function generally have serious consequences for survival of the cell. In humans, mitochondrial dysfunction is the underlying cause of a wide range of degenerative diseases, with energy-demanding cells such as those of the central nervous and endocrine systems, heart, muscle, and kidney being most severely affected.

Structure. Mitochondria are bounded by two concentric membranes referred to as the outer and the inner. This creates two distinct compartments, the matrix and the intermembrane space.

The outer membrane consists of a bilayer containing about 80% lipid. It is freely permeable to molecules smaller than about 5000 daltons, a feature that is attributable to the presence of many copies of a channel-forming protein called porin.

The inner membrane is also a lipid bilayer. It is extremely rich in protein (about 75%) and is impermeable to even the smallest of ions. This latter property may be due to the presence of a relatively high proportion of cardiolipin (10%) in the lipid bilayer. The inner membrane contains the enzymes of the electron transport chain and the ATP synthase, together with a set of transporter proteins that regulate the movement of metabolites in and out of the matrix space. Mitochondria of cells that depend on a high level of ATP production are usually extensively folded to produce structures called cristae (Fig. 1). These greatly increase the surface area of the inner membrane, allowing many more copies of the enzymes of oxidative phosphorylation. Cristae are generally believed to be baffle-shaped structures in the inner membrane that allow free communication of the intermembrane and intracristal spaces (Fig. 1). In reality, however, as shown by transmission electron microscopic tomography (**Colorplate 1**), cristae may be connected to the peripheral inner membrane by narrow tubular connections that may considerably restrict diffusion of molecules between intermembrane and intracristal spaces.

The intermembrane space contains enzymes

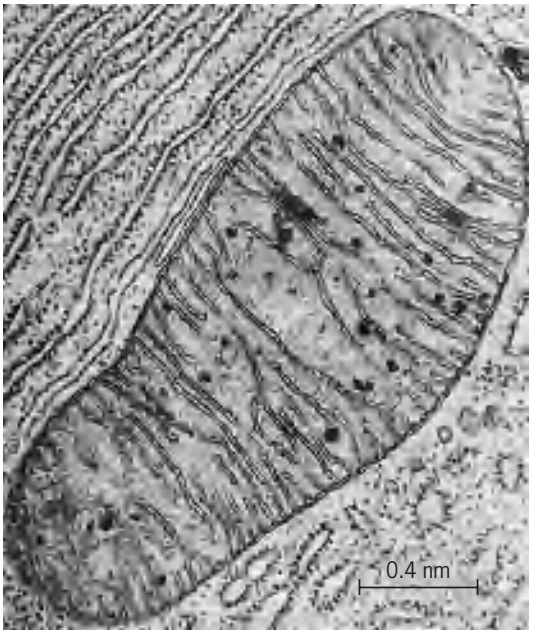

Fig. 1. Electron micrograph of a thin section through the pancreas of a bat, showing a typical mitochondrion in profile. Note how the cristae are formed by extensive folding of the inner membrane. (*Courtesy of K. R. Porter*)

capable of using some of the ATP that is transported out of the matrix to phosphorylate other nucleotides.

The matrix space is packed with a hundred or so water-soluble proteins that form a sort of semisolid gel. They include enzymes of the tricarboxylic acid (Krebs) cycle and enzymes required for the oxidation of pyruvate and fatty acids, comprising steps in the biosynthesis or degradation of amino acids, nucleotides, and steroids.

Dynamics. Textbooks often depict mitochondria as rod- or sausage-shaped structures with diameters of 0.5–1.0 micrometer and lengths of 1–10 μm, depending on cell type (Fig. 1). This convention, based largely on older electron micrographs, fails to bring out the highly dynamic character of the organelle. Studies that have employed confocal laser microscopy and time-lapse photography show mitochondria to be anything but static. Within minutes, they can change shape, grow in length, branch, divide, or fuse, often moving about in the cytoplasm as they do so. In many types of cell, mitochondria accumulate at sites of ATP utilization, for example between adjacent myofibrils in heart muscle cells, or wrapped around the flagellum of a sperm cell. In others, they may cluster about other organelles or structures, thereby facilitating the exchange of metabolic intermediates or ions (**Colorplate 2**). An association between mitochondria and microtubules has also frequently been observed. This may play a role in determining the way in which mitochondria are distributed in the cell and in which they segregate during cell division.

Energy conservation. Most cells derive their energy from the coupled processes of electron transport and oxidative phosphorylation in mitochondria. Energy production is fueled mainly by pyruvate, produced during glycolysis, or by fatty acids. Both compounds are selectively transported into the mitochondrial matrix where, after conversion to two-carbon units in the form of acetyl coenzyme A, they are fed into the citric acid cycle for further oxidative breakdown. Operation of the cycle converts the two-carbon units to CO_2 and H_2O. High-energy electrons extracted during this process are temporarily held by the coenzymes NADH and $FADH_2$ before being passed to molecular oxygen via the enzymes of the respiratory chain. The energy that is released as these electrons pass from one enzyme to the next is used to pump protons across the inner membrane from the matrix space into the intermembrane space (Fig. 2).

The pumping of protons creates an electrochemical gradient across the inner membrane. This gradient represents a form of stored energy, which can be used to perform different kinds of work. Commonly, as mentioned above, backflow of the protons down the gradient is used to drive synthesis of ATP through the action of the enzyme ATP synthase. Other energy-requiring processes can also be driven, including the transport of metabolites or ions against their concentration gradients by coupling their movement to the backflow of protons.

The overall process of oxidative phosphorylation

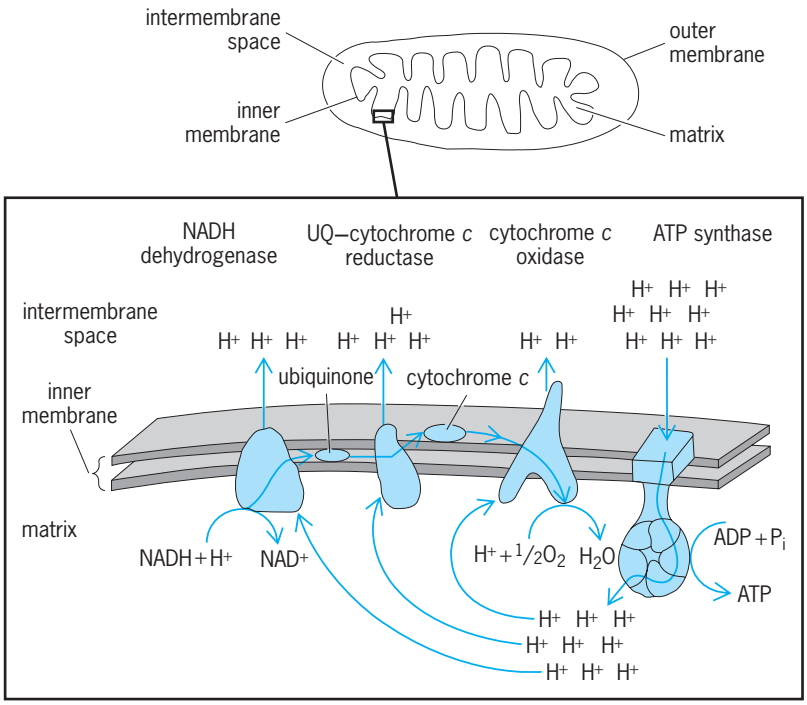

Fig. 2. Mitochondrial inner membrane, showing the pathway of electrons through the electron transport chain and the use of a proton gradient to drive ATP synthesis. Electrons pass from one enzyme to another; each enzyme is one of a series of carriers that are embedded in the membrane. Each carrier in the chain has a progressively higher affinity for electrons, so that these tumble onward, losing energy. Final transfer is to oxygen, which has the highest affinity of all. At two points in the chain, mobile carriers, ubiquinol and cyctochrome *c*, mediate the passage of the electrons. Energy released during electron transfer is used by the enzymes of the chain to pump protons from the matrix into the intermembrane space. (*After B. Alberts et al., Molecular Biology of the Cell, Garland, 1983*)

by mitochondria is remarkably efficient, with about 80% of the energy theoretically available from oxidation of carbohydrates and fats being stored in the form of ATP. Compare this with the 10–20% efficiency of manufactured energy-conversion devices such as the internal combustion engine or electric motor. Lower efficiencies are sometimes found, however, and the mitochondria of brown fat cells are an interesting example. The inner membrane of these mitochondria contains a protein that allows protons to leak down the electrochemical gradient into the matrix. The energy stored by oxidative phosphorylation is thus wasted in the form of heat, but for a reason—such cells are usually found in newborn infants and hibernating animals, where they produce warmth necessary for survival. *See* ADIPOSE TISSUE; CITRIC ACID CYCLE.

Electron transport chain. Understanding the mechanisms of electron transport and oxidative phosphorylation has been increased dramatically by the recent elucidation of the three-dimensional (3D) structures of three of the enzymes involved (bc_1 complex, cytochrome *c* oxidase, and ATP synthase). This was a technical tour de force. First, analysis of the 3D structure required prior crystallization of the enzymes, which was no easy task in this case because of their tendency to aggregate and become insoluble after their release from the membrane. Second, the complexity of the structures required that the analytical tehniques used be pushed to their limits.

Cytochrome bc_1 complex. Cytochrome bc_1 (ubiquinol cytochrome *c* oxidoreductase; complex III) of mammalian mitochondria contains 11 subunits, but only 3 of these carry the redox centers necessary for electron transport (**Colorplate 3**). These are cytochrome *b*, the so-called Rieske iron-sulfur protein, and cytochrome c_1. Both the iron-sulfur protein and cytochrome c_1 are anchored in the inner membrane by their C-terminal tails and possess headpieces that extend into the intermembrane space. The headpiece of the iron-sulfur protein carries an iron-sulfur (Fe_2S_2) cluster, while that of cytochrome c_1 contains a c-type heme group. Cytochrome *b* [encoded by mitochondrial (mtDNA)] contains eight helical regions, which are all contained within the lipid bilayer of the inner membrane. It also contains two heme groups, each of which forms an active site in electron transport. Cytochrome *b* delivers electrons to the iron-sulfur protein, whose headpiece rotates through about 60° to make contact with that of cytochrome c_1. A key component of the bc_1 complex is not a protein at all, but a low-molecular-weight lipid-soluble and highly mobile electron carrier called ubiquinol. In a complicated mechanism that has been designated the Q-cycle, electrons are donated to cytochrome *b* one at a time. Operation of the Q-cycle drives protons across the inner membrane, providing part of the proton-motive force that is later used to synthesize ATP. As occasional byproducts, operation of the Q-cycle also produces highly reactive semiquinol intermediates that can react with molecular oxygen to form various type of reactive oxygen species (ROS). These can cause extensive damage to proteins, lipids, and deoxyribonucleic acid (DNA) and as such are probably responsible for many of the pathological developments in cells that contain dysfunctional respiratory chains.

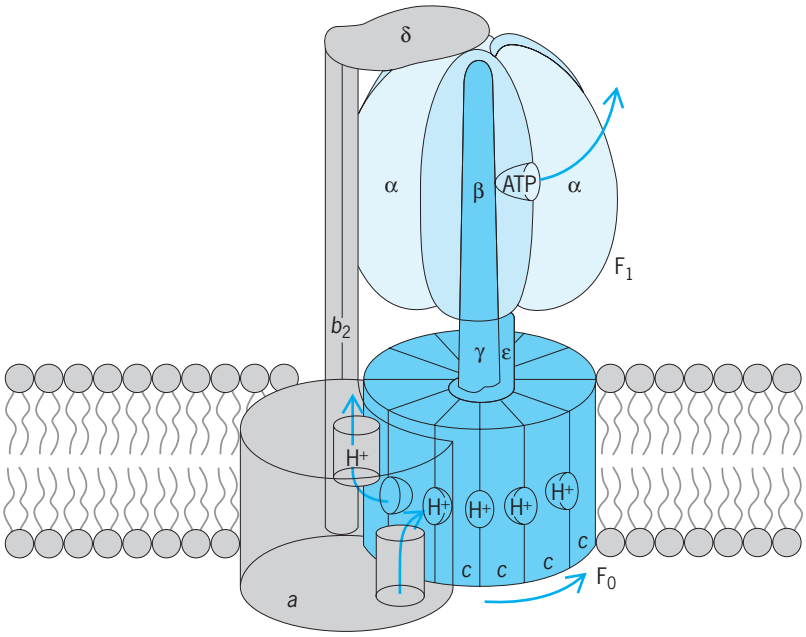

Fig. 3. Three-dimensional structural model of the mitochondrial ATP synthase based in part on x-ray crystallography of bovine F_1 ATPase and of the *Escherichia coli* F_0. In the latter, the a, b, and c subunits correspond to the mitochondrial Atp6, b, and Atp9 subunits. The bovine F_1F_0 ATP synthase probably consists of some 16 different proteins and has a molecular weight of greater than 500 kDa. (*From P. D. Boyer, What makes ATP synthase spin?, Nature, 402: 247-248, 1999*)

Cytochrome c oxidase. The second respiratory enzyme for which detailed 3D structural information is available is cytochrome *c* oxidase (complex IV; **Colorplate 4**). Like the bc_1 complex, cytochrome *c* oxidase acts both as an electron carrier and as a proton pump. In mammalian cells, it consists of 13 subunits, of which only 3 (Cox1, 2, and 3) have clearly defined roles in enzyme activity. These 3 subunits are all encoded by mtDNA. The active site of the enzyme is situated in Cox1, which is buried in the inner membrane, spanning it several times with a series of alpha helices. The substrate for the enzyme is cytochrome *c*, a water-soluble protein with a c-type heme group. The electrons it donates are used by the enzyme to reduce molecular oxygen to two water molecules. Cytochrome *c* is loosely attached to the outer surface of the inner membrane and shuttles electrons between cytochrome c_1 and cytochrome *c* oxidase. Electrons received from reduced cytochrome *c* are picked up by Cox2, which contains a dinuclear, mixed valence copper center (Cu_A). Interaction between the two proteins is possible because Cox2, like cytochrome c_1, is anchored in the inner membrane only by its C-terminal part. The remainder of the protein forms a flexible headpiece that extends into the intermembrane space. The electrons are then passed to the subunit I, the functional core of the enzyme, first to cytochrome *a* (a low-spin heme) and then to the bimetallic cytochrome aa3/Cu_B site. One of the Cu_B ligands is a histidine residue which forms an adduct with a neighboring tyrosine. It has been suggested that free-radical generation by this adduct may play a role in the reduction of molecular oxygen.

Each atom of oxygen is reduced to a molecule of water by acceptance of two electrons and two protons. At the same time, two protons are translocated across the enzyme from the matrix to the intermembrane space. Both protons that are consumed in the reaction, and those that are pumped reach the active center of the enzyme through hydrophilic crevices in the enzyme, possibly by passing along chains of water molecules. Beyond the active site, the mechanism of proton movement is less clear, since this region of the enzyme is strongly hydrophobic.

ATP synthase. Energy stored in the form of a proton gradient across the mitochondrial inner membrane can be converted to ATP by the action of the turbine-like enzyme F_1F_0-ATP synthase (complex V). The enzyme consists of two parts. F_0 is the proton channel. It is embedded in the membrane. F_1 is the catalytic part. It protrudes out of the F_0 complex into the matrix space and consists of five different subunits (α, β, γ, δ, and ϵ) in a ratio of 3:3:1:1:1. Examination of the crystal structure of mitochondrial F_1 (**Fig. 3**) reveals that synthesis of ATP is achieved by rotation of a central stalk consisting of γ and ϵ subunits inside a globular head of α and β subunits. Structural studies on both the mitochondrial and *Escherichia coli* ATP synthases suggest that the head is fixed relative to

the membrane by the stator. This is a complex of a membrane-embedded subunit (a subunit in *E. coli*), together with the β and δ subunits that link to F_0. Gated movement of protons through a complex in F_0 produces structural changes which cause the parts to rotate (rather like the blades in a real turbine). This rotation induces the rotation of the central γ-ϵ stalk, which in turn induces conformational changes in the α-β head, therby driving the conversion of bound adenosine diphosphate (ADP) and inorganic phosphate to ATP and its subsequent release. Elegant studies in which the torque of isolated F_1 has been measured during rotation caused by reversal of the normal enzyme reaction (that is, hydrolysis of ATP) suggest that the ATP synthase operates with almost 100% efficiency.

Mitochondrial genetic systems. Both mitochondria and chloroplasts contain DNA and the machinery necessary to express the information stored there. In both cases, the DNAs are relatively small and simple compared with DNA in the nucleus. While chloroplast DNA tends to be very similar in size in all organisms examined, mtDNA varies widely in complexity, from 16–18 kilobases in metazoa to upward of 2000 kilobases in some higher plants. Much of this variation can be explained by differences in organization: compact in metazoa; expanded in the other organisms, with large intergenic regions and introns in many genes. In the case of higher-plant mtDNAs, some extra sequences appear to have been picked up from chloroplast and nuclear DNAs. MtDNA is generally circular, although some linear exceptions are known among the yeasts, algae, and protozoa.

The information content of most mtDNA is limited. This means that most of the several hundred proteins found in these organelles are encoded by genes located in nuclear DNA. These proteins are synthesized in the cytosol and subsequently transported specifically to the respective organelles.

The contributions of the two genetic systems are usually closely coordinated, so that cells synthesize organelles of more or less constant composition. However, isolated mitochondria and chloroplasts are capable of synthesizing both RNA and protein in vitro and in vivo; both are capable of continuing for a short time with synthesis even after nuclear RNA synthesis or after cytosolic protein synthesis has been blocked.

Both types of activity have provided a means of identifying the products made by each genetic system. Thus the compound cycloheximide inhibits protein synthesis on cytosolic ribosomes but does not affect protein synthesis by either mitochondria or chloroplasts. In contrast, antibiotics such as chloramphenicol, tetracycline, and the macrolides inhibit organellar protein synthesis, but not cytosolic. This inhibition underlies the toxic effects often observed during long-term administration of these antibiotics to patients.

With some minor variations, all mtDNAs contain genes for proteins belonging to enzymes required for electron transport and oxidative phosphorylation. In humans, they include 7 subunits of the NADH dehydrogenase, 3 subunits of cytochrome *c* oxidase, 1 subunit of the cytochrome bc_1 complex (ubiquinol cytochrome *c* reductase), and 2 subunits of the ATP synthase. Genes for distinctive, organellar ribosomal (rRNAs) and transfer RNAs (tRNAs) are also present. Curiously, while some mtDNAs encode a complete set of the latter, others (including plants and trypanosomes) do not, and prevailing evidence suggests that the missing tRNA species have to be specifically imported from the cytoplasm.

The first mitochondrial genome to be completely sequenced was human mtDNA. With a length of 16,569 base pairs, this is one of the most compact of all mtDNAs, and it displays a number of unusual features:

1. Nearly every nucleotide appears to be part of a coding sequence; there is little or no room for regulatory sequences.
2. The genome contains only one major promoter region. Transcripts initiated within one part of this region extend over the full length of both strands (symmetric transcription). Others, initiated a short distance upstream from the first site, display a high tendency to terminate immediately downstream of the genes encoding the two rRNAs. This process of attenuation ensures a higher rate of synthesis of rRNAs relative to other transcripts.
3. Individual rRNAs, tRNAs, and mRNAs are generated by processing, signaled by tRNA sequences which are interspersed between the other genes.
4. DNA sequences of most protein-coding genes lack complete translational stop codons. Full-stop codons are generated by addition of a poly (A) tail.
5. The genetic code differs from the universal genetic code in three important respects: UGA (stop) encodes Trp, AUA (Ile) encodes Met, and AGA/AGG (Arg) encodes stop.
6. Only 22 genes for tRNAs are present. This number is sufficient for mitochondrial protein synthesis, but only because restrictions on codon-anticodon recognition operative in other genetic systems appear to be relaxed.

Inheritance. Many organisms, including humans, show uniparental inheritance of mitochondrial genes because one parent contributes more cytoplasm to the zygote than the other. In humans, it is the egg cell provided by the mother that contributes the cytoplasm. Human mitochondrial genes are thus inherited maternally.

Protein import. Only a small fraction of the total number of mitochondrial proteins is synthesized within the organelle itself. The remainder are specified by nuclear genes and are imported into the growing mitochondrion after their synthesis in the cytoplasm. These proteins are directed to the organelle by specific targeting sequences contained within them. In most cases, these sequences are located at the N-terminus of the newly synthesized protein and are removed after import by a specific protease. A smaller group of imported proteins contains noncleaved targeting sequences that are less well characterized. For both groups, import is initiated by interaction of precursor proteins with receptor complexes located in the outer membrane. Subsequent

steps include transfer to specific translocation complexes of the inner membranes at sites at which inner and outer membranes make transient contact, sorting to one of the intramitochondrial compartments and folding to a native conformation. Both unfolding and refolding of import proteins are carried out by a special class of proteins known as chaperones. Import of all proteins requires ATP and, with the exception of outer membrane proteins, is dependent on an electric potential across the inner membrane.

Diseases. Recent years have seen a growing interest in human diseases that result from mitochondrial dysfunction. A number of these result from mutations in mtDNA. Others are linked to nuclear genes, whose mutation disturbs oxidative phosphorylation or impairs mitochondrial assembly. Mutations in mtDNA are remarkably frequent and lead to a wide range of degenerative, mainly neuromuscular diseases. Most of these diseases are maternally inherited, but some appear to be spontaneous, possibly resulting from error-prone replication of mtDNA. One of the first mitochondrial diseases to be characterized was Leber's hereditary optic neuropathy (LHON), a maternally inherited form of sudden onset of blindness caused by death of the optic nerve. LHON patients possess a missense mutation in one of the mtDNA-encoded subunits of the NADH dehydrogenase (often ND1, ND4, or ND6). Single-base changes in mtDNA-encoded tRNA genes are responsible for the syndromes known as myoclonic epilepsy and ragged-red fibers (MERFF) and mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like (MELAS) episodes. In other myopathies, which include chronic external ophthalmoplegia (CEOP), Kearns-Sayres syndrome (KSS), and Pearson marrow/pancreas syndrome, large-scale deletions of segments of mtDNA are responsible for the disease. The sites of such deletions are often bounded by short repeated sequences in mtDNA, and they may arise as a result of the process of slippage-mispairing during mtDNA replication.

A striking feature of mtDNA-related diseases is the enormous diversity in clinical presentation. This diversity is attributable to two main factors: (1) Heterogeneity in the mtDNA population. Most human tissues contain many thousands of mtDNA molecules per cell. The severity of clinical symptoms depends on the number of mutated molecules present. (2) Dependence of a particular cell type on mitochondrial function (mainly ATP production). Cells with a high requirement for mitochondrially generated ATP (central nervous system, endocrine system, heart, muscle, kidney) are more severely affected than cells with alternative sources of ATP (such as liver).

Apoptosis. Besides their role in metabolism and energy-linked processes, mitochondria have recently been identified as important players in the initiation of apoptosis (programmed cell death). On one hand, the mitochondrial outer membrane houses a number of members of the Bcl-2 family of apoptosis regulatory proteins. On the other hand, release of certain mitochondrial proteins from the intermembrane space is instrumental in activating specialized proteases called caspases. These catalyze a degradative cascade in the cytoplasm that eventually ends in cell death. Cytochrome *c* is just such a caspase-activating protein. Details of how its release is triggered are lacking, but a plausible chain of events includes opening of pores in the outer membrane mitochondrial permeability transition pores, swelling of the inner membrane, and subsequent collapse of the electric potential across the inner membrane. Activation of the mitochondrial permeability transition pores can be induced by excessive uptake of calcium (Ca^{2+}), exposure to reactive oxygen species, or decline in energetic capacity. Since decline in energetic capacity and increase in oxidative stress are common consequences of many mitochondrial diseases, pore opening and consequent triggering of apoptosis is likely to contribute to the pathophysiology of the disease.

Evolution. According to the endosymbiont hypothesis originally formulated by L. Margulis in 1968, mitochondria are proposed to have originated from an aerobic bacterium at a time of increasing oxygen levels in a world inhabited by proto-eukaryotes, the ancestors of present-day eukaryotic cells. These primitive cells survived by engulfing and digesting other organisms. The bacterium is likely to have been an α-proteobacterium similar to present-day *Paracoccus*. Having escaped digestion after engulfment, it is proposed to have developed a stable, symbiotic relationship with its host by providing respiration-derived ATP in exchange for metabolizable substrates and physical protection. As time passed, its genome was drastically reduced by loss of genetic information needed only for independent growth and by transfer of many genes needed for respiration to the host nucleus. The result of this process of reductive evolution is the modern mitochondrion, its DNA encoding a handful of genes that failed to be transferred before divergence of the organellar and nucleo-cytoplasmic genetic systems made further genetic exchange impossible.

The endosymbiont hypothesis has up to now been widely accepted. It received strong support from molecular sequence data published in the period 1975–1995. These showed that both mitochondrial genes and nuclear genes encoding mitochondrial proteins are indeed of α-proteobacterial origin, while many other genes in the eukaryotic nucleus turned out to be of archaebacterial origin.

However, subsequently published sequences, in many cases derived from genomic sequence projects, now reveal that nuclear genomes carry many genes of bacterial origin that have nothing to do with mitochondrial functions. Additionally, primitive amitochondrial eukaryotes, such as the microsporidia, which are considered to be the direct descendants of the ancient proto-eukaryotes, also carry bacterial genes in their nuclear genomes. Plausible alternatives, but still controversial explanations, for both features are offered by what have been termed the hydrogen hypothesis (put forward by W. Martin and M. Müller in early 1998) and the syntrophy and oxygen toxicity hypothesis (both put

forward by S. Andersson and C. Kurland in 1999). Like the endosymbiont hypothesis of Margulis, each of these models proposes that mitochondria indeed have their origins in symbiosis between an α-proteobacterium and a member of the Archaea, and that mitochondrial genome evolution has involved both gene loss and transfer. They each differ, however, in the suggested identity of the driving force behind symbiosis. The endosymbiosis hypothesis proposes that this was respiration-generated ATP; the hydrogen hypothesis proposes the production of hydrogen; the syntrophy hypothesis proposes the utilization of methane; the oxygen toxicity (ox-tox) hypothesis proposes the local removal of oxygen from the intracellular environment of the primitive anaerobic host.

Mammalian mtDNA evolution. Mammalian mtDNAs accumulate mutations at high rates and evolve correspondingly fast (up to 12–15 times faster than single-copy genes in nuclear DNA and up to 100 times faster for rRNA and tRNA genes). This behavior reflects both a high incidence of mutations and a high probability of their fixation. The first is probably related to oxidative damage to mtDNA by oxygen free radicals produced as by-products of electron transfer through the respiratory chain. The second has been attributed to the lack of efficient DNA repair (mitochondria lack nucleotide-excision repair) and to a relatively high tolerance of many mitochondrial gene products to mutational change.

The rapid rate of sequence evolution of mammalian mtDNAs makes these genomes highly sensitive indicators of recent evolutionary relationships. Unlike their nuclear counterparts, mtDNAs do not undergo recombination during sexual transmission and are strictly maternally inherited. Sequence changes in mtDNA therefore provide a clear record of the history of the female lineages through which this DNA has been transmitted. In 1992 the geneticist Allan Wilson compared a short section of these sequences in 182 individuals. His results allowed the construction of an evolutionary tree that indicates that the common ancestor of all surviving human mtDNA types lived in Africa between 140,000 and 290,000 years ago. This ancestor is often misleadingly referred to as Mother Eve, the suggestion being that Eve corresponds to a single woman from whom all living humans have descended. This is incorrect. The root of the tree simply indicates the point in time at which the first mutation occurred in mtDNA. All other ancestral types simply failed to survive. Additionally, the vast majority of human genes are encoded by chromosomes in the nucleus. A large number of individuals, both men and women, are thus likely to have contributed to the genetic makeup of the individual whose mtDNA gave rise to the many variants extant today. *See* BIOLOGICAL OXIDATION; CELL (BIOLOGY); ENZYME.

Les A. Grivell

Bibliography. G. Attardi and A. Chomyn (eds.), *Mitochondrial Biogenesis and Genetics*, vol. 264 of *Methods in Enzymology*, Academic Press, 1996; V. Darley-Usmar and A. H. V. Schapira (eds.), *Mitochondria, DNA, Proteins and Disease*, Portland Press Res. Monog. Ser., no. 5, 1994; I. E. Scheffler, *Mitochondria*, Wiley, 1999; A. Tzagoloff, *Mitochondria*, Plenum, 1982.

Mitosis

The series of visible changes that occur in the nucleus and chromosomes of non-gamete-producing plant and animal cells as they divide. During mitosis, the replicated genes, packaged within the nucleus as chromosomes, are precisely distributed into two genetically identical daughter nuclei (**Fig. 1**). The series of events that prepares the cell for mitosis is known as the cell cycle. When viewed in the context of the cell cycle, the definition of mitosis is often expanded to include cytokinesis, the process by which the cell cytoplasm is partitioned during cell division. Although the continuity of heredity by cell division was first predicted by R. Virchow in 1858, the central role of mitosis was not fully understood until the mid-1900s, when the deoxyribonucleic acid (DNA) within chromosomes was proven to contain the hereditary blueprints for life. *See* CELL CYCLE; DEOXYRIBONUCLEIC ACID (DNA).

Mitotic spindle. Chromosome segregation is mediated in all nonbacterial cells (that is, eukaryotes) by the transient formation of a complex structure known as the mitotic spindle (**Fig. 2**). During mitosis in most higher plants and animals, the nuclear membrane surrounding the replicated chromosomes breaks down, and the spindle is formed in the region previously occupied by the nucleus (open mitosis). In lower organisms, including some protozoa and fungi, the spindle is formed and functions entirely within the nucleus which remains intact throughout the process (closed mitosis).

All spindles are bipolar structures, having two ends or poles. In animal cells, each spindle pole contains an organelle, the centrosome (**Fig. 3**), onto which the spindle focuses. As a result, the spindle in animal cells looks like a football. The polar regions of plant spindles lack centrosomes and, as a result, are much broader. In animals the bipolar nature of the spindle is established by the separation of the centrosomes, which is critical for a successful mitosis; the presence of only one pole produces a monopolar spindle in which chromosome segregation is inhibited. The presence of more than two poles produces multipolar spindles which distribute the chromosomes unequally among three or more nuclei. Centrosomes are duplicated during interphase near the time that the DNA is replicated, but then act as a single functional unit until the onset of mitosis. In plants, and during meiosis in some animals, the two spindle poles are organized by the chromosomes and by molecular motors that order randomly nucleated microtubules into parallel bundles. *See* CENTROSOME; PLANT CELL.

Microtubules. Microtubules are the primary structural components of the mitotic spindle and are required for chromosome motion (Fig. 3). These 25-nanometer-diameter, hollow, tubelike structures are

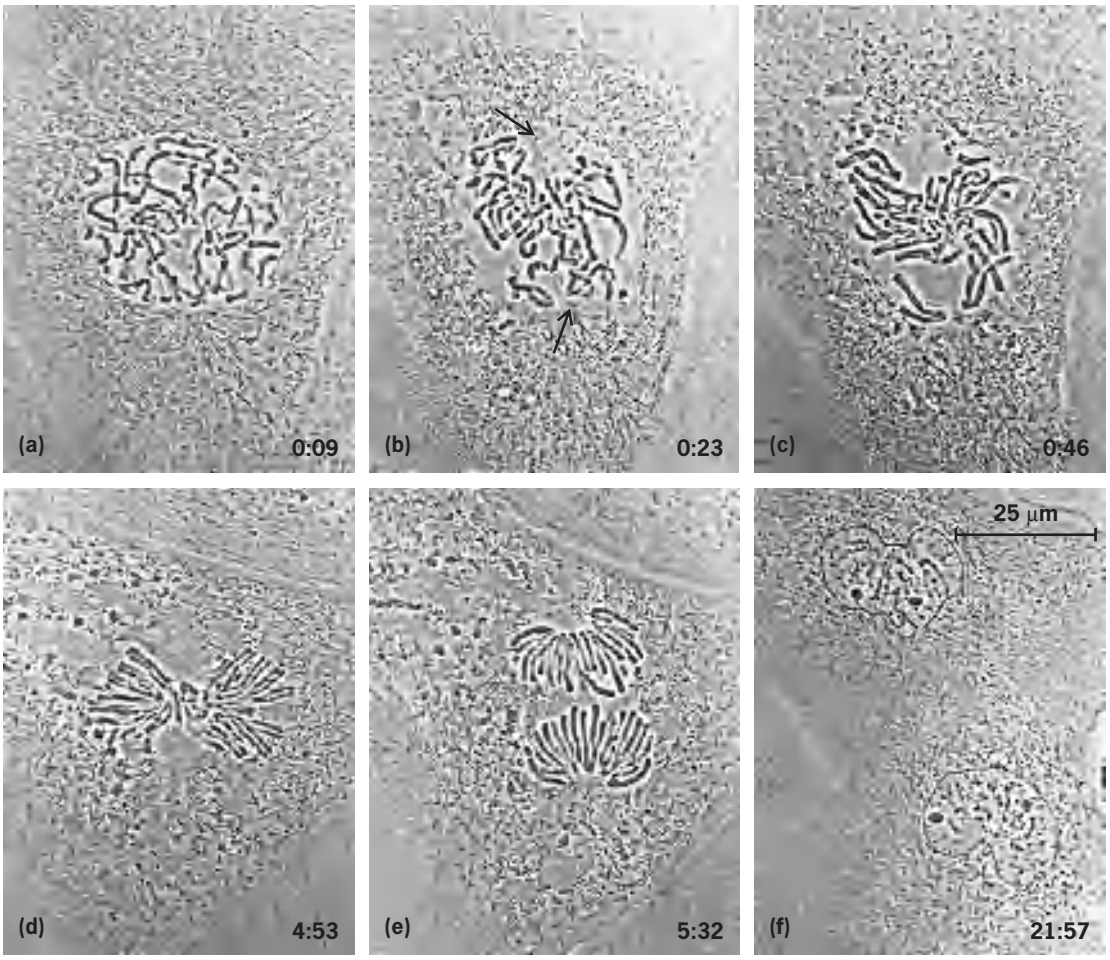

Fig. 1. Selected phase-contrast light micrographs showing changes in chromosome position during mitosis in a living newt lung epithelial cell (elapsed time in hours and minutes). (*a*) Late prophase. (*b*) Prometaphase. (*c*) Mid-prometaphase. (*d*) Metaphase. (*e*) Anaphase. (*f*) Telophase.

formed from the polymerization of protein subunit dimers composed of alpha and beta tubulin. Microtubules are polarized structures that grow much faster at one end than at the other; the fast-growing end is referred to as the positive (+) end, the slow-growing end as the negative (−) end. During interphase, microtubules are distributed throughout the cytoplasm, where they serve to maintain cell shape and also function as polarized roadways for transporting organelles and cell products. As the cell enters mitosis, the cytoplasmic microtubule network is disassembled and replaced by the mitotic spindle. The microtubules in animal cells originate from the centrosome that, like the chromosomes, was inherited during the previous mitosis where it functioned as a spindle pole. The centrosome contains a unique type of microtubule protein (gamma tubulin) that is involved in seeding microtubule assembly. Once growth is initiated, the positive ends of centrosomal microtubules grow away from the centrosome, while the negative ends either remain associated with the centrosome or are shed from it. At any one time, many microtubules within the cell are growing or shrinking as subunits are added or removed from their positive ends. Changes in the parameters that modulate this dynamically unstable behavior of microtubule ends lead to changes in the numbers and average length of microtubules.

The motion associated with microtubules is mediated by several families of molecular motors which bind to and move along the wall of the microtubule. These motors include kinesins that move cargo along microtubules toward their positive ends away from the centrosome, and cytoplasmic dynein (and also some members of the kinesin family) that transports cargo along individual microtubules toward their centrosomal or negative ends. *See* CYTOSKELETON.

Kinetochores. As mitosis begins, each replicated chromosome consists of two identical sister chromatids that are joined along their length. In most cells, chromosomes possess a unique region of highly condensed chromatin (DNA plus protein), known as the centromere, which forms an obvious constriction on the chromosome, referred to as the primary constriction. Spindle microtubules attach to a small specialized structure on the surface of the centromere known as the kinetochore (**Fig. 4**). Fragments of chromosomes lacking a kinetochore do not move poleward; it is always the kinetochore that leads in the poleward motion of the chromosome.

Kinetochores in most plant cells are shaped like a ball, whereas those in animal cells resemble a plate. The centromere region of each replicated chromosome contains two sister kinetochores, one attached to each chromatid, that lie on opposite sides of the primary constriction.

Regulation of mitosis. Cells that are in the cell cycle contain an inactive kinase, known as $p34^{cdc2}$. Once chromosome and centrosome replication is completed at the end of S phase, two cyclin proteins (A and B) are synthesized that rapidly bind to and activate $p34^{cdc2}$. This activated enzyme is called cyclin-dependent kinase 1 (CDK1), formerly known as mitosis or maturation promoting factor (MPF). CDK1 then drives the cell into mitosis by phosphorylating specific proteins involved in the interphase/mitosis transition. In vertebrates, cyclin A is synthesized before cyclin B; and once active CDK1/cyclin A accumulates in the nucleus, it initiates the early events of mitosis including chromosome condensation. The subsequent accumulation of active CDK1/cyclin B then commits the cell to mitosis, and the cell will stay in mitosis as long as CDK1/cyclin B activity remains high. *See* CELL CYCLE.

Stages. Once initiated, mitosis is a continuous process that, depending on the temperature and organism, requires several minutes (*Drosophila*, yeast) to many hours (newts) to complete. Traditionally it has been subdivided into five consecutive stages that are distinguished primarily by chromosome structure, position, and behavior. These stages are prophase, prometaphase, metaphase, anaphase, and telophase.

Prophase. The first visible sign of an impending mitosis occurs within the nucleus as the DNA-containing chromatin begins to condense into chromosomes. This condensation is due to the phosphorylation of chromatin-associated proteins, including histone H1 and H3, by CDK1/cyclin A and the aurora B kinase within the nucleus. Up to a point, this is a reversible process, and cells in early-to-mid prophase can be induced to decondense their chromosomes and return to interphase by a variety of treatments that stress the cell and/or damage the DNA. However, once active CDK1/cyclin B accumulates in the nucleus at late prophase, the cell becomes irreversibly committed to the division process.

By late prophase the individual chromosomes are well defined and the nucleoli begin to dissipate. Near this time the cytoplasmic network of microtubules breaks down, and is replaced in animal cells by two radial astral arrays of microtubules growing from the two spindle poles (centrosomes). When compared to interphase, the microtubules associated with mitotic centrosomes contain more but significantly shorter microtubules. The functional changes that occur in the centrosome as the cell enters mitosis are mediated by the activity of CDK1 and other kinases that modify both centrosome chemistry and the proteins that control microtubule dynamic instability.

As the asters form, they separate and move toward opposite sides of the nucleus. During this separation the microtubule arrays from each aster exhibit considerable overlap, and within this overlap the microtubules from opposing asters are of opposite polarity. In some lower organisms, including yeast and diatoms, the spindle poles are pushed apart by the action of microtubule plus end motors that bind microtubules of opposite polarity within the region of overlap, and then slide them away from one another. However, in higher animal cells, including vertebrates, the asters continue to move apart even when their microtubule arrays no longer overlap. Thus, it is likely that the separation of spindle poles in higher animals occurs via multiple mechanisms, including one that involves pulling.

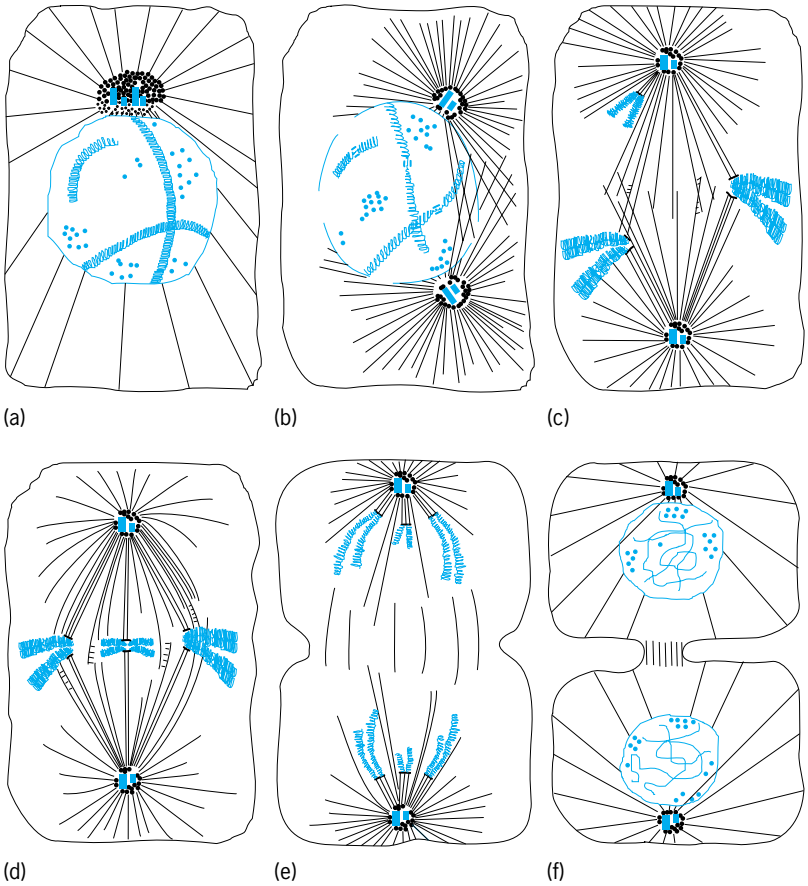

Fig. 2. Changes in microtubule distribution and chromosome position as an animal cell enters into and proceeds through mitosis. (*a*) Mid-prophase: chromosomes are condensing within the nucleus. (*b*) Late prophase/early prometaphase: centrosomes separate; the cytoplasmic microtubule complex is replaced by two radial astral microtubule arrays; microtubules in each aster grow and shorten at their ends away from the centrosome; the nuclear envelope breaks down: kinetochore-containing primary constrictions are sometimes visible on the well-condensed chromosomes. (*c*) Mid-prometaphase: the kinetochores and chromosomes interact with the asters to form the spindle. (*d*) Metaphase: all of the chromosomes are aligned on the spindle equator; sister kinetochores are attached to opposite poles by kinetochore fibers. (*e*) Anaphase: the sister chromatids separate and move toward their respective spindle poles (anaphase A); at the same time the spindle poles move farther apart (anaphase B). (*f*) Telophase: the two groups of sister chromosomes have become two well-separated sister nuclei; the cytoplasm of the cell divides (cytokinesis).

Prometaphase. The sudden activation of CDK1/cyclin B during late prophase results in the rapid phosphorylation of various nuclear and cytoplasmic proteins. The most obvious consequence of this activity is that the nuclear envelope surrounding the chromosomes suddenly swells and dissolves, as many

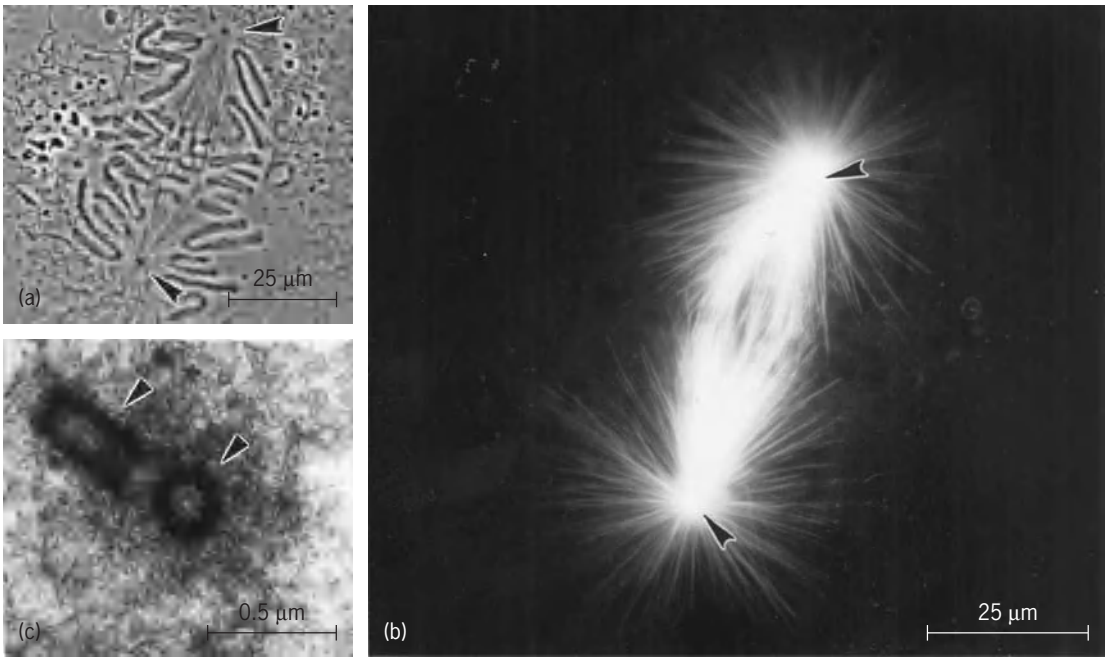

Fig. 3. Phase-contrast light micrograph of a newt lung cell in mid-prometaphase. (*a*) The two centrosomes or spindle poles (arrowheads) and some spindle fibers are visible, as are a number of mono-oriented chromosomes. (*b*) The same cell photographed by fluorescence microscopy. Most of the microtubules are concentrated within the spindle between the poles. (*c*) High-voltage electron micrograph of a thick section through two orthogonally arrayed centrioles (arrows) and surrounding dense material, cut from a mitotic centrosome depleted of its associated microtubule by cold treatment. (*Parts a and b from L. Cassimeris, C. L. Rieder, and E. D. Salmon, Microtubule assembly and kinetochore directional instability in vertebrate monopolar spindles: Implications for the mechanism of chromosome congression, J. Cell Sci., 107:285–297, 1994*)

of its constituent proteins become phosphorylated and soluble. At the same time, CDK1 activity also induces other membrane systems within the cell, including the Golgi apparatus and the endoplasmic reticulum, to fragment into numerous small vesicles that then become randomly distributed.

The breakdown of the nuclear envelope initiates the prometaphase stage of mitosis. During this stage the chromosomes attach to the separating centrosomes as their kinetochores capture dynamically unstable astral microtubules that are randomly probing the cytoplasm. Once an astral microtubule is captured by a kinetochore, it becomes known as a kinetochore microtubule, and its positive end becomes firmly anchored within the kinetochore. A kinetochore begins to move as soon as it encounters its first microtubule. However, as a rule, kinetochores continue to capture as many astral microtubules as their surface area allows and, throughout the capture process, they induce these microtubules to bundle into a kinetochore fiber. In vertebrates, kinetochore fibers rarely form at the same time on sister kinetochores. Instead, the tendency is for a chromosome to attach first to the closest spindle pole as the kinetochore facing that pole acquires microtubules. Because only one of its kinetochores is attached to the spindle, such chromosomes are mono-oriented, and they rapidly move toward the pole to which they have become attached. Over time the opposing sister kinetochore on a mono-oriented chromosome, which lacks an attachment to microtubules, must become associated with the positive ends of microtubules generated from the opposing half-spindle (that is, the far pole). This occurs by chance when it captures one or more microtubule ends growing from the far pole. To increase the chances of this happening in a reasonable time, the unattached kinetochore often becomes laterally associated with the surface of adjacent microtubules residing in the same half-spindle as the mono-oriented chromosome. Once this lateral association is established, microtubule positive-end centromere-associated protein E (CENP-E) motors on the surface of the unattached kinetochore transport it, and its associated chromosome, along the surface of the microtubules toward the spindle equator (that is, away from the closer pole). As the kinetochore approaches the equator, its chances of encountering a positive microtubule end growing from the opposing half-spindle dramatically increase. Once both sister kinetochores are attached to the spindle, so that each kinetochore is attached to different and opposing poles, the chromosome is considered to be bioriented. The failure of a chromosome to acquire this proper biorientation leads to the production of aneuploid cells with unequal numbers of chromosomes, with corresponding detrimental consequences to the organism.

Once a chromosome has achieved biorientation status, it undergoes a complex series of motions, termed congression, that ultimately position its centromere on the spindle equator midway between the poles. In order for congression to occur, the kinetochore fiber microtubules that terminate in the sister kinetochores must elongate and shorten in a coordinated fashion. For example, as a bioriented chromosome moves toward the equator, those kinetochore

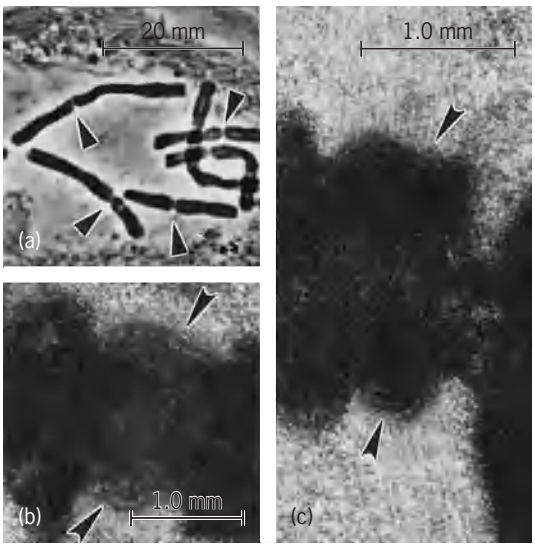

Fig. 4. Kinetochore structure. (*a*) Phase-contrast light micrograph of six unattached chromosomes in a living mitotic newt lung cell. The centromere-containing primary constrictions (arrowheads), and the replicated nature of the chromosomes, are clearly visible. (*b*) High-voltage electron micrograph of a thick section cut through the primary constriction of an unattached rat kangaroo chromosome. Two platelike sister kinetochores (arrowheads) are positioned on opposite sides of the primary constriction. (*c*) Same as b except that each sister kinetochore is attached to opposing spindle poles by a bundle of microtubules. (*Part b from C. L. Rieder and G. G. Borisy, The attachment of kinetochores to the pro-metaphase spindle in PtK$_1$ cells, Chromosoma, 82:693–716, 1981*)

microtubules that attach the chromosome to the proximal pole must elongate, while those that attach it to the distal pole must shorten. In vertebrates, the elongation and shortening of kinetochore fiber microtubules during chromosome motion occur primarily by the addition and deletion of tubulin subunits into microtubules at the kinetochore, but shortening also occurs to a lesser extent from subunit loss at the microtubule minus end associated with the pole.

Mono-oriented chromosomes exhibit constant-velocity, oscillatory motions toward and away from their associated spindle pole. The tendency of the only attached kinetochore on a mono-oriented chromosome to autonomously switch between poleward and away from the pole motion means that the motility of kinetochores is directionally unstable. When the unattached kinetochore of a mono-oriented chromosome attaches to the distal pole, it begins to move toward that pole. In order for this congression motion to continue, its sister kinetochore, attached to the closer pole, must switch into a prolonged movement away from the pole state of motion. The factors that coordinate the motile behavior of sister kinetochores, to allow for congression, remain unknown.

Metaphase. When the last chromosome becomes positioned on the spindle equator, the cell is considered to be in metaphase of mitosis. By this time, many of the astral microtubules have been sequestered into the spindle, which is now a fully mature structure, and cyclin A has been degraded. The metaphase spindle consists of two half-spindles, tethered to one another by the sister kinetochores on the chromosomes and by interactions between their associated overlapping (antipolar) microtubule arrays. Each half-spindle contains the same number of kinetochore fiber microtubule bundles interspersed among more numerous microtubules growing from the pole or free within the spindle.

Anaphase. The sudden and largely synchronous separation (disjunction) of sister chromatids initiates the anaphase stage of mitosis. This disjunction is not dependent on forces generated by the spindle as evidenced by the fact that it also occurs in some cells when spindle formation is inhibited by drugs. In all but embryonic cells, entry into anaphase is controlled by a cell-cycle checkpoint that delays chromatid disjunction until all of the kinetochores have achieved a stable attachment to the spindle. In these cells, unattached or weakly attached kinetochores trigger a signal transduction cascade that delays anaphase by inhibiting the activity of a macromolecular assembly known as the anaphase promoting complex (APC). These large, spindle-associated complexes selectively target several proteins involved in the metaphase/anaphase transition, including securin that maintains the cohesion between the sister chromatids and cyclin B, for proteolytic destruction by ubiquinating them. This "kinetochore attachment" or "spindle assembly" checkpoint ensures that anaphase does not normally start until all of the chromosomes have achieved a bipolar orientation, which is the minimum requirement for equal chromosome segregation.

Once the chromatids disjoin in non-drug-treated cells, they slowly (1–2 μm/min) move toward their respective poles in a process referred to as anaphase A. During this time the once highly organized spindle begins to disassemble. After anaphase A is under way, the poles themselves begin to move farther apart in a process known as anaphase B.

Telophase. The activation of APCs at the metaphase/anaphase transition not only leads to chromatid disjunction but also inactivates the mitotic (CDK1) kinase by destroying its associated cyclin B. As CDK1 activity falls, the various proteins that were phosphorylated during the initial stages of mitosis become dephosphorylated, and pathways are initiated that ultimately return the dividing cell to interphase. During this telophase stage of mitosis, the groups of separated sister chromosomes, which are now positioned near their respective poles, begin to swell and stick to one another. As this occurs, a nuclear envelope is deposited on the surface of the decondensing chromatin, and a Golgi apparatus reforms in association with each of the two centrosomes.

During telophase, new microtubule-based structures, known in animals as stem bodies and in plants as phragmoplasts, also form between the now-separated daughter nuclei. The phragmoplasts are responsible for directing the construction of a new cell wall that ultimately partitions the dividing plant cell into two separate (but connected) entities. In

animals the stem bodies act as a template and catalyst for forming and stabilizing the cytokinetic furrow, which constricts the cell into two independent daughters. At the end of telophase the centrosome inherited by each daughter cell nucleates another interphase complex of cytoplasmic microtubules. *See* CYTOKINESIS.

Chromosome movement. In some cells a kinetochore (and thus its associated chromosome) moves toward a spindle pole because the fiber to which it is attached is pulled toward the pole. This "traction fiber" mechanism accounts for 100% of chromosome poleward motion in some systems, including insect spermatocytes, various oocytes, and perhaps plant cells. It remains to be resolved whether the pulling force for this motion is produced along the entire length of the fiber by microtubule plus end motors anchored in the spindle matrix, or only at its end in the spindle pole. For poleward motion to occur, however, the kinetochore microtubules must shorten as they slide poleward, and this shortening clearly occurs by subunit loss at the microtubule minus ends at the pole. In contrast to a traction fiber mechanism, during poleward chromosome motion in vertebrate somatic cells only 25% of kinetochore fiber shortening can be attributed to subunit loss at the spindle pole. The other 75% occurs by subunit loss at the microtubule plus ends within the kinetochore. Immunological studies show that kinetochores in vertebrates contain cytoplasmic dynein, CENP-E, and other microtubule-based motor molecules, and that during prometaphase they can move rapidly along the surface of a microtubule toward a centrosome (spindle pole). It remains to be determined if the kinetochore-based force for poleward chromosome motion in these cells is produced by microtubule minus-end molecular motors associated with the kinetochore or simply from the disassembly of microtubule minus ends within the kinetochore. Regardless, in humans the kinetochore can be viewed as a complex biological machine that uses its associated kinetochore fiber microtubules to generate the force for chromosome poleward motion while allowing the plus ends of these microtubules to grow and shorten within its confines.

The mechanism which ultimately positions the sister kinetochores of a bioriented chromosome stably on the spindle equator, halfway between the two spindle poles, remains to be solved. For this to occur, the behavior of sister kinetochores must somehow be coordinated, so that when one is moving toward its associated pole, the other is moving away from its pole. At the moment the best guess for how this occurs is that kinetochores somehow "sense" their position and the spindle, and regulate their behavior accordingly. *See* CELL (BIOLOGY); CELL DIVISION; CELL NUCLEUS.

Conly L. Rieder

Bibliography. D. Mazia, Mitosis and the physiology of cell division, pp. 77–412 in *The Cell*, vol. III, ed. by J. Brachet and A. E. Mirsky, Academic Press, New York, 1961; C. L. Rieder and A. Khodjakov, Mitosis through the microscope: Advances in seeing inside live dividing cells, *Science*, 300:91–96, 2003; S. Gadde and R. Heald, Mechanisms and molecules of the mitotic spindle, *Curr. Biol.*, 14:R797–R805, 2004; B. A. A. Weaver and D. W. Cleveland, Decoding the links between mitosis, cancer and chemotherapy: The mitotic checkpoint, adaptation, and cell death, *Cancer Cell*, 8:7–12, 2005.

Mitteniales

An order of the true mosses (subclass Bryidae) found in Australia and New Zealand. The order consists of a single species, *Mittenia plumula*, which is adapted for growth in caves or cavelike places. Branches of the persistent protonema consist of spherical cells which reflect light from a backing of chloroplasts, thus providing a glow. The stems are simple and erect, and the leaves are 2–4-ranked and oblong-lingulate and blunt or apiculate. The midrib ends above the midleaf, and the cells are short and smooth. The sporophytes are terminal, the setae elongate, and the capsules erect and oblong-cylindric. The opercula are conic-subulate. The peristome consists of 16 filiform teeth, inrolled when moist and recurved when dry, and 32 smooth, hyaline, nodulose segments attached to a short membrane. The calyptrae are mitrate and smooth.

The order is remarkably similar to the Schistostegales of North Temperate distribution in protonema and habitat, but the other features of gametophyte and sporophyte are quite different. A distinctive feature is the double peristome with 32 segments derived in part from the outermost cell walls of the endothecium (that is, the inner portion of the embryonic capsule). All other mosses of the subclass Bryidae have peristomes entirely of amphithecial origin. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA; SCHISTOSTEGALES.

Howard Crum

Mixer

A device with two or more signal inputs and one common output. The two primary classes are linear (additive) and nonlinear (multiplicative) mixers. Linear mixers are used to add or blend together two or more signals, nonlinear mixers mainly to shift the spectrum (center frequency) of one signal by the frequency of a second signal.

Linear mixer. Linear mixing is the process of combining signals additively, such as the summing of audio signals in a recording studio. This operation can be accomplished passively by simply using a resistive summing network (**Fig. 1***a*). Although this approach appears very economical, there is a loss in signal strength and an interaction of the signal amplitudes as the gains are adjusted. These problems are demonstrated by the equation below for the output signal, e_o. The numerator contains the proper linear

$$e_o = \frac{R_L\left(\frac{e_1}{R_1} + \frac{e_2}{R_2} + \cdots + \frac{e_N}{R_N}\right)}{1 + R_L\left(\frac{1}{R_1} + \frac{1}{R_2} + \cdots + \frac{1}{R_N}\right)}$$

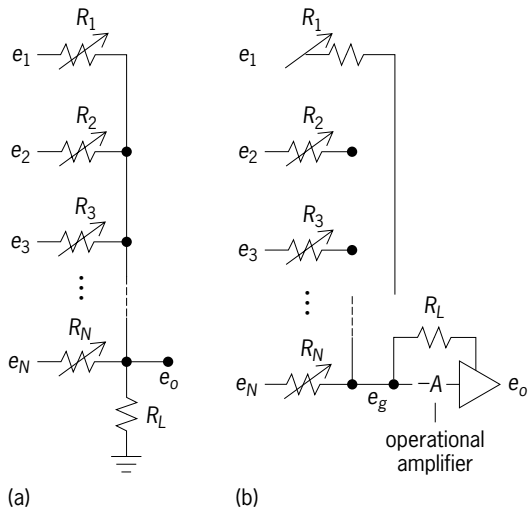

Fig. 1. Linear mixers. (*a*) Passive linear mixer based on a resistive summing network. (*b*) Active linear mixer with an operational amplifier.

mixing terms. The ith signal, e_i, is weighted by R_L/R_i, where R_i is the corresponding input resistance and R_L is the load resistance, and the result of all weighted signals is summed. The denominator contains all of the problems. It provides the overall attenuation of the output, which is further complicated by being a function of the individual R_i settings; one adjustment changes all the gain values.

Inexpensive integrated circuits have solved both problems and have revolutionized this application dramatically. Operational amplifiers of reasonably high quality that will eliminate the adjustment interactions and also provide gain are readily available. The input signals are summed into the virtual ground summing node (signal e_g) at the input of the operational amplifier (Fig. 1*b*). The troublesome denominator in the above equation is replaced by unity, so that interaction and attenuation are eliminated. There is a sign change in the output, but that is a small drawback compared to the advantage of having the virtual ground provided by the operational amplifier. *See* AMPLIFIER; INTEGRATED CIRCUITS; LINEARITY; OPERATIONAL AMPLIFIER.

Nonlinear mixer. Perhaps the most familiar application of nonlinear mixers is in radio and television receivers, but they are widely used in such applications as amplitude modulation (AM) and demodulation, frequency demodulation, phase detection, frequency multiplication, and single-sideband (SSB) generation. The incoming information to a receiver has been transmitted and received at a frequency (say, ω_i) far too high to permit efficient amplification and processing. Therefore the signal is translated or frequency-shifted or heterodyned by a mixer to a lower frequency (ω_o), known as the intermediate frequency (IF), where amplification and processing are performed efficiently by an IF processor, sometimes referred to as the IF strip. *See* AMPLITUDE-MODULATION DETECTOR; AMPLITUDE MODULATOR; FREQUENCY-MODULATION DETECTOR; FREQUENCY MODULATOR; FREQUENCY MULTIPLIER; PHASE-ANGLE MEASUREMENT; RADIO RECEIVER; SINGLE SIDEBAND; TELEVISION RECEIVER.

If the input signal has frequency ω_i, and the adjustable (tuned) local oscillator signal into the mixer has frequency $\omega_i - \omega_o$, then it follows from trigonometry that the mixer output, which is the product of the input and oscillator signals, is the sum of two terms. One of these terms is at a frequency equal to the sum of the input and oscillator frequencies, that is, $2\omega_i - \omega_o$, and since ω_i is much larger than ω_o, this frequency is double that of the input; therefore, this term can be completely filtered out. The frequency of the remaining signal is at the difference of the input and oscillator frequencies, namely ω_o, the IF-channel frequency. *See* HETERODYNE PRINCIPLE.

A second application of a nonlinear mixer is frequency synthesis, where a stable but not easily changed signal at a high frequency, ω_b, is made tunable by mixing it with an easily tunable signal at a low frequency, ω_v, which, perhaps, can be varied in precise increments of any size. As before, the output is the sum of two terms, in this case at frequencies $\omega_b + \omega_v$ and $\omega_b - \omega_v$. The utility of the method is limited by the ability to filter or separate one frequency term from another (the frequency separation of the two terms is only $2\omega_v$), thereby determining the minimum practical value of ω_v for the application.

A mixer is an integral part of an AM-radio integrated circuit which contains virtually all AM-radio functions except filters. A particular type of mixer, the quadrature detector, is included in the frequency-modulation (FM)-radio integrated circuit.

A circuit commonly used to implement the multiplication of two signals to give the product necessary for nonlinear-mixer applications consists of a set of three differential amplifiers (**Fig. 2**). Two of these differential amplifiers (consisting of transistor pairs Q_1-Q_2 and Q_3-Q_4, respectively) are driven out of phase with one another and into saturation by the first input signal. The output points are at the nodes of the cross-coupled collectors (Q_1-Q_3 and Q_2-Q_4). The linear differential drive by the second input signal to the remaining transistor pair (Q_5-Q_6) results in the output being the second signal, square-wave-modulated (or demodulated) by the first signal. The

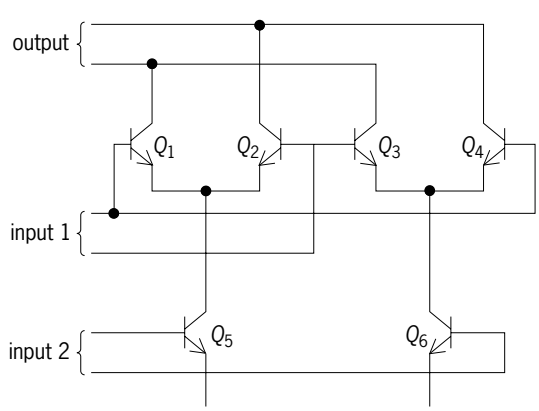

Fig. 2. Balanced modulator (nonlinear mixer). Only signal paths are shown. Biasing must be added.

consequence of this very efficient but nonlinear operation is that the output spectrum contains the desired spectrum, plus images (or aliases or spectral replicas) centered at odd harmonics of the first input signal. These images are removed by subsequent filtering. *See* DIFFERENTIAL AMPLIFIER. Stanley A. White

Bibliography. A. V. Oppenheim and A. S. Willsky, *Signals and Systems*, 2d ed., 1996; M. S. Roden, *Analog and Digital Communication Systems*, 4th ed., 1995; M. Schwartz, *Information Transmission, Modulation, and Noise*, 4th ed., 1990.

Mixing

A common operation to effect distribution, intermingling, and homogeneity of matter. Actually the operation is called agitation, with the term "mixing" being applicable when the goal is blending, that is, homogeneity. Other processes, such as reaction, mass transfer (includes solubility and crystallization), heat transfer, and dispersion, are also promoted by agitation. The type, extent, and intensity of agitation determine both the rates and adequacy of a particular process result. The agitation is accomplished by a variety of equipment.

Most liquid mixing is done by rotating impellers in vertical cylindrical vessels. A typical impeller-type liquid mixer with a variety of features is shown in **Fig. 1**. The internal features, including the vessel itself, are considered as a whole, that is, as the agitated system. The forces applied by the impeller develop overall circulation or bulk flow. Superimposed on this flow pattern, there is molecular diffusion, and if turbulence is present, also turbulent eddies. These provide micromixing. Solids, granular to powder, are mixed in a variety of contrivances.

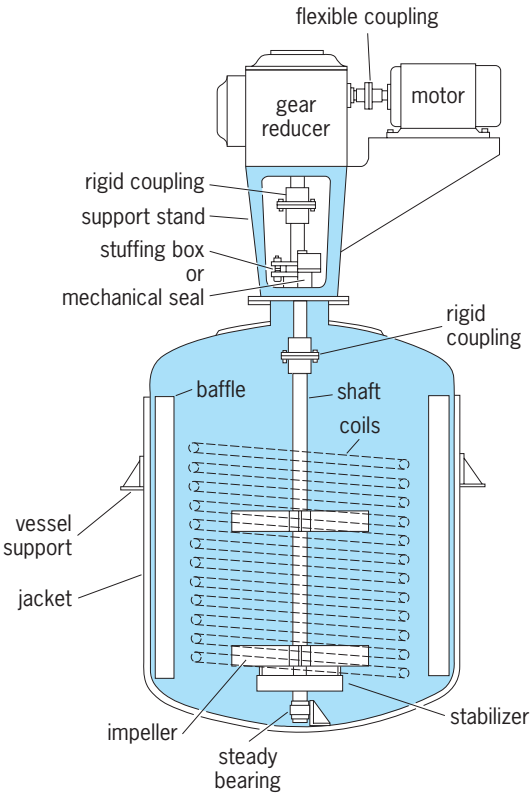

Fig. 1. Typical impeller-type liquid mixer. (*After V. W. Uhl and J. B. Gray, Mixing: Theory and Practice, vol. 2, Academic Press, 1967*)

A classification of operations requiring mechanical agitation and mixing

Physical criteria	General kinds of operations	Chemical and mass transfer criteria
Pumping, circulation	Fluid motion	Heat transfer
Blending	Miscible liquids	Reactions
Emulsions	Immiscible liquids	Extraction, reaction
Dispersions	Liquid-gas systems	Absorption, stripping, reaction
Dispersion, suspension	Liquid-solid systems	Dissolving, crystallization

Fluid mixing. The range of industrial processes for which mechanical agitation is used is conveyed by the **table**. The processes may be batch or continuous. Often several operations are conducted in concert; for example, in the hydrogenation of vegetable oils, the catalyst is suspended, a reaction takes place, and heat is transferred to the jacket wall. The performance criteria for different operations vary, and although several agitation schemes often will accomplish a given process result, certain systems can be shown to be preferable. Here the governing factors generally are the type, size, and speed of the impeller as well as major internals such as baffles.

Fluid motion, both large-scale (bulk circulation) and small-scale (turbulent eddies), is required in turbulent flow. The bulk circulation results when the fluid stream is discharged by the impeller. Turbulence is generated mostly by the velocity discontinuities adjacent to the stream of fluid flowing from the impeller, but also by boundary-and-form separation effects; turbulence spreads throughout the bulk flow and, although attenuated, is carried to all parts of the vessel. It is recognized that some mixing operations should require relatively large bulk or mass flows, whereas others need a high intensity of turbulence (termed shear in this connection). From this it follows that there should be an optimum ratio of flow to shear for a given process result, for example, high flow for blending of miscibles and suspension of solids, and high shear for gas absorption (mass transfer) and dispersion.

Energy must be supplied to produce fluid motion. The power imposed by the mixing impeller is proportional to the flow multiplied by the head developed, which is equated to turbulence (shear). The ratio of the power supplied for flow to that for turbulence can be varied for the same power input and type of impeller; this variation is demonstrated by **Fig. 2**. Impeller shape also has an effect which can be significant, with a rotating disk representing the extreme case for high turbulence or shear.

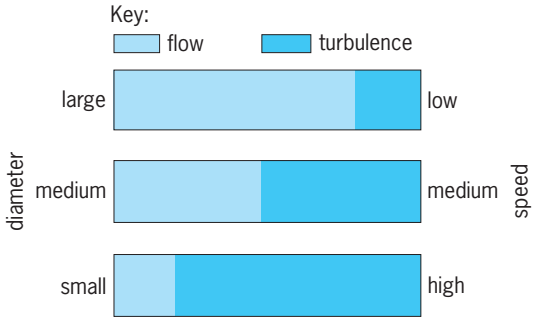

Fig. 2. Effect of impeller size and speed on flow and turbulence (shear) at constant power.

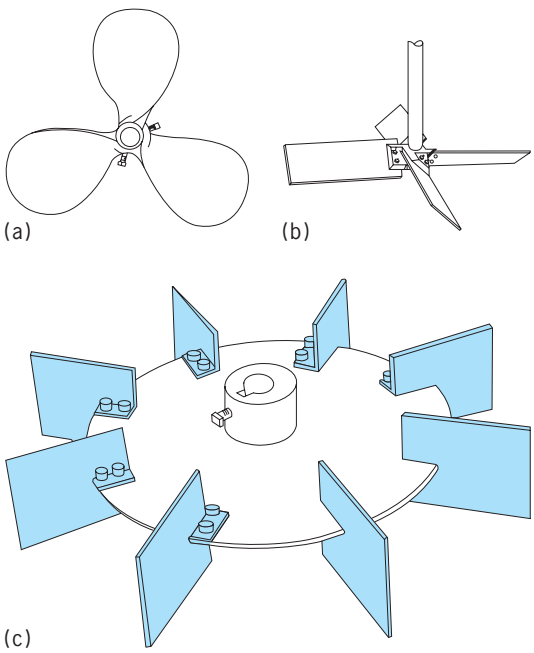

Fig. 3. Common impellers. (*a*) Marine-type mixing propeller. (*b*) Axial-flow (pitched-blade) turbine. (*c*) Flat-blade (Rushton) turbine. (*After R. H. Perry and C. H. Chilton, eds., Chemical Engineers' Handbook, 5th ed., McGraw-Hill, 1973*)

In all bench-scale and pilot-plant work where mixing is important, the type of fluid motion in terms of ratio of flow to shear should be found. This becomes a basis for scale-up in design.

Mixing vessels and impellers. Circulation or bulk flow depends on the shape of the container; the fittings it contains; the type, size, and position of the impeller; and the properties of the fluid. Lateral (radial) and vertical flow currents usually produce the best agitation; these currents must penetrate to all portions of the fluid.

For the laboratory, whenever matters are mixed, cylindrical containers such as beakers should be used. Vertical baffles should be provided for impellers which are centrally located; if a propeller is used in an off-center position, the baffles may be omitted. Round-bottomed flasks are unsatisfactory unless modified by creases blown in the side; this form can be purchased.

The most common and useful types of impellers are the marine-type propeller (**Fig. 3***a*), the axial-flow (pitched-blade) turbine (Fig. 3*b*), the flat-blade (Rushton) turbine (Fig. 3*c*), and the flat paddle (generally two vertical blades relatively large compared to the other types). Any of these impellers centrally positioned produces rotating fluid motion (**Fig. 4**) with a vortex around which liquid swirls. This motion often results in separation or stratification rather than intermingling. Relatively little power can be applied; a dearth of turbulence and of vertical and lateral flow motion results. Inserting projections into the body of the fluid stops the rotary motion, and the vortex disappears. Thus agitation is improved. Such projections at the side of the tank are called baffles (**Fig. 5**). The propeller and axial-flow turbine with tank-wall baffles (Fig. 5*a*) will generate an axial-flow pattern; the paddle and Rushton turbine produce radial flow (Fig. 5*b*). These flow patterns are conducive to good agitation.

Swirl can be obviated in laboratory work by using the propeller off-center. Good vertical and lateral flow motion without swirl and surface vortex can be attained by discharging the propeller downward and positioning as in **Fig. 6**; the exact position is critical but can be found by trial.

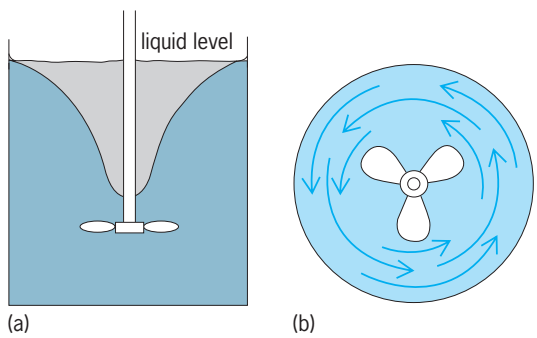

Fig. 4. Swirling flow pattern for impeller of any shape, without baffles. (*a*) Side view. (*b*) Bottom view.

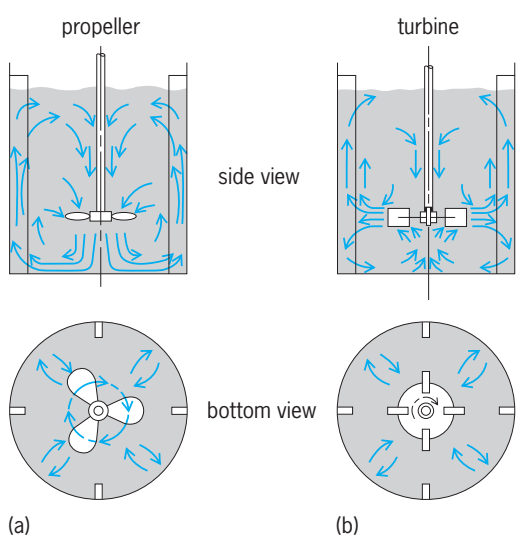

Fig. 5. Flow patterns with baffles. (*a*) Flow pattern for propeller with baffles at tank wall. (*b*) Flow pattern for turbine with baffles at tank wall.

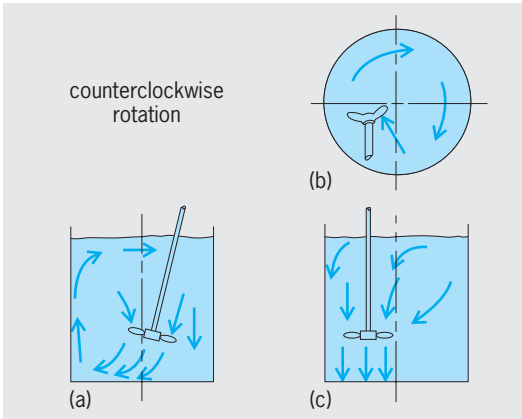

Fig. 6. Flow pattern for top-entering, off-center propeller without baffles: (*a*) front, (*b*) top, and (*c*) side views.

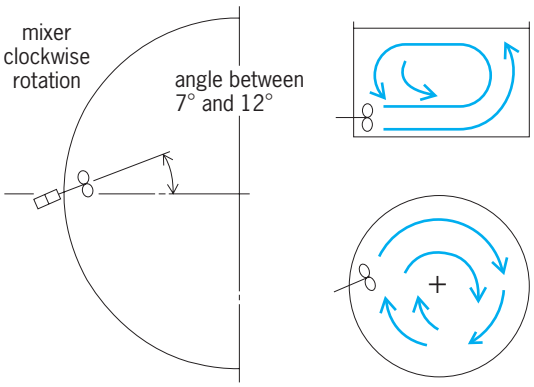

Fig. 7. Side-entering propeller mixer position. With proper propeller position, no vortex will result.

Liquids in large tanks, for example, 3×10^6 gal (11,370 m^3), are often blended by a side-entering propeller. **Figure 7** shows the correct position for a right-hand, clockwise-rotation propeller; proper angular positioning is critical.

The behavior of low-viscosity liquids has been described above and is shown in Fig. 4. High-viscosity liquids, those above 1000 centipoises, will have much less rotary flow; without baffles the motion will approximate that in Fig. 5.

Power and flow. In liquids of low viscosity, power imposed by an impeller in a baffled tank is proportional to the cube of its speed, the fifth power of its diameter, and the density. The fluid regime would be turbulent. For constant power and varying diameter of a given type of impeller, the corresponding speed can be found from the equation $N_r = (1/D_r)^{5/3}$, where N_r is the ratio of speeds and D_r is the ratio of corresponding impeller diameters.

In liquids of high viscosity, the power imposed by the impeller is proportional to the viscosity of the liquid, the square of the speed, and the cube of the diameter. Here the fluid regime is laminar. *See* VISCOSITY.

The discharge rate of flow from impellers is proportional to the speed and to the cube of the diameter. Clearly, for the same power there will be greater flow from the large-diameter, low-speed impeller than from a smaller-diameter, higher-speed one (Fig. 2).

High-viscosity liquids. These liquids and pastes often require different techniques from those used for mixing low-viscosity liquids. Special apparatus is necessary to provide for wiping, stretching, and squeezing, because turbulence cannot be generated in such fluids to provide for the small-scale mass transfer necessary to cause interpenetration of substances. There are few quantitative data available which describe the performance of the various types of equipment. An example of typical equipment is shown in **Fig. 8**.

Continuous processing. For continuous pipeline blending, a baffled mixing cell with one or two impellers on the shaft is available. Centrifugal pumps sometimes provide in-line agitation as an auxiliary service. For continuous two-phase contacting, a column containing compartment separation plates, baffles, and impellers is used.

Solids mixing. Solids of different density and size are mixed in tumblers (a double cone turning end on end) or with agitators (a helical ribbon rotating in a horizontal trough). The duration of mixing is an important additional variable because classification and separation often occur after attainment of the desired distribution if the operation is carried on too long.

Equipment. Portable mixers generally range in size from fractional horsepower to 3 hp (2240 W); they are designed for use in open tanks up to 3000 gal (11.4 m^3) capacity and are clamped to the tank shell. Direct-drive units usually run at 3600, 1750, or 1150 rpm or at variable speeds. Gear-drive units run at about 420 rpm or at variable speeds.

Light-duty, permanently mounted mixers range from fractional horsepower to 3 hp, and are used on open or closed tanks up to 3000 gal (11,400 L) capacity. Heavy-duty, permanently mounted, top-entering mixers are made up to 500 hp (373,000 W). They are normally designed for speeds of 45–200 rpm, made

Fig. 8. Ribbon turbine with double spiral, for high-viscosity liquids.

with independent mounting for the mixer shaft, and connected to the drive by a flexible coupling to protect the speed-reduction gearing. *See* UNIT OPERATIONS. Vincent W. Uhl

Advances, trends, and challenges. The understanding of mixing of fluids and solids has benefited by major advances in computations and chaos theory. These insights help to explain why some designs work and others do not, and also suggest novel designs.

The objective of fluid mixing is to produce the maximum amount of contact or interfacial area between fluids in the minimum amount of time or with the least amount of energy. High interfacial area aids molecular diffusion, ultimately resulting in a homogeneous mixture. The key to effective mixing lies in stretching and folding fluid elements. Stretching and folding is synonymous with chaos and results in an exponential increase of interfacial area with time. Experiments have shown that it is relatively straightforward to produce chaotic flows. Mixing in chaotic flows does not require turbulent motion, and the concepts apply even to very viscous liquids, where turbulence would be difficult to produce. Chaos can be achieved by time modulation of the flow field or by periodic changes in geometry, as in static mixers. Chaos, however, is rarely complete and, in general, chaotic systems contain both chaotic regions, where mixing is effective, as well as regions of poorly mixed fluid (called islands, regular regions, or tubes). The objective of good mixing is the elimination of islands.

Chaotic mixing occurs in impeller-type mixers, in continuous duct-type flows such as static mixers, and even applies to mixing in very small scale systems, such as in microfluidic applications. One example of a duct-type chaotic flow is the so-called partitioned-pipe mixer (**Fig. 9**). This flow consists of a pipe partitioned with a sequence of orthogonally placed rectangular plates. The cross-sectional motion is induced through rotation of the pipe with respect to the assembly of plates. Two streams are injected at the top. In Fig. 9*a*, both streams are in chaotic regions and mix well. In Fig. 9*b*, one of the streams is injected into an island and remains unmixed.

Similar results have been uncovered in stirred tanks where mixing is less efficient than once thought, with only a fraction of the energy input actually participating in mixing. This area has benefited from computational flow codes where rapid advances have taken place. Significant advances have occurred also in areas where the objective is the creation of fine dispersions of drops or particles, such as the blending of immiscible polymers, composites, and nano-composites, and in the mixing of fluids with complex rheological characteristics. Mixing effects in reactive systems are critical in the case of multiple reactions, when reactions are fast, or in viscous media, when mixing is slow compared to the rate of reaction. Improperly designed systems lead to slowing of desired reactions and sometimes even halting before completion, increase of undesired reactions, and reduction of product selectivity.

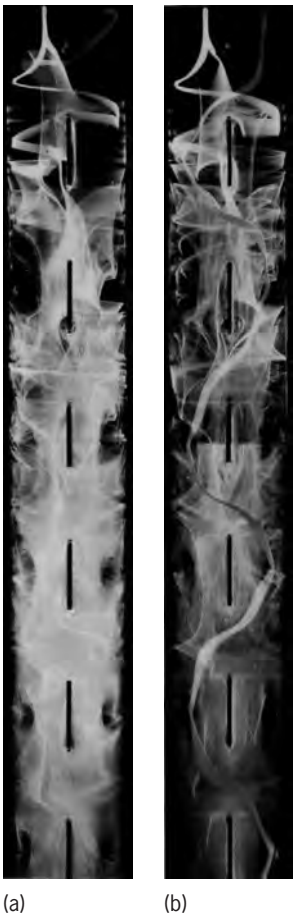

Fig. 9. Coexistence of (*a*) chaotic and (*b*) nonchaotic mixing in a duct-type flow.

Mixing of granular solids is pervasive in industry and natural processes. In contrast to mixing of liquids, mixing of granular materials received disproportionally little attention until recently. Computation and chaos theory concepts have aided in understanding. The simplest case of mixing of dry solids has benefited by simulation via particle dynamics (also called discrete element methods), where each particle is simulated using contact force models that capture friction and inelasticity. Presently it is difficult to simulate mixing with the number of particles that are contained in industrial mixers, such as a double-cone blender and bladed mixers, but laboratory experiments are within reach and simulations with 10^4–10^5 particles are routine. Simulations with 1–2 million particles are now possible.

Mixing of solids is more complicated than mixing of liquids. Mixing happens only when there is flow (there is no molecular diffusion). Also, granular solids can segregate or become unmixed. Unmixing happens if particles differ in size or density (and in practice they offer differ in both). In this case, there may be differential motion of particle types, with the end result of magnifying variations in composition and, in the worst case scenario, nearly complete unmixing (**Fig. 10**). There are numerous processes where a nonuniform composition can have serious consequences. In pharmaceuticals

Fig. 10. Unmixing of particles of two different sizes in a rotating tumbler operating in a time periodic mode.

poor mixing often requires batches of products to be either remixed or disposed. Theory reveals that mixing of solids is highly sensitive to mixer filling. Improper filling may leave entire regions in the mixer unmixed. In general, low fill levels perform better. The shape of the mixer is of critical importance; noncylindrical shapes, such as an equilateral triangle, may mix better than a cylinder, the most common shape. Recent advances have taken place for slurries, where particles are immersed in a liquid, and for powders, where cohesion between particles is important. Julio M. Ottino

Bibliography. N. Harnby, M. F. Edwards, and A. W. Nienow, *Mixing in the Process Industries*, 2d ed., 1997; G. Metcalfe et al., Avalanche mixing of granular materials, *Nature*, 374:39–41, 1995; J. M. Ottino, The mixing of fluids, *Sci. Amer.*, 260:56–67, 1989; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineer's Handbook*, 7th ed., 1997; T. Shinbrot and F. J. Muzzio, Nonequilibrium patterns in granular mixing and segregation, *Phys. Today*, pp. 25–30, March 2000; V. W. Uhl and J. B. Gray, *Mixing*: *Theory and Practice*, vol. 1, 1966, vol. 2, 1967, vol. 3, 1986; J. J. Ulbrecht and J. B. Gray (eds.), *Mixing of Liquids by Mechanical Agitation*, 1985.

Mizar

The second-magnitude star that marks the bend in the handle of the Big Dipper, ζ Ursae Majoris. Mizar was the first telescopic visual binary to be discovered (1650), the first ever to be photographed (1857), and the first star discovered to be a spectroscopic binary (1889). Its distance to the Sun is 24 parsecs (7.5×10^{14} km or 4.7×10^{14} mi). *See* BINARY STAR; CONSTELLATION; URSA MAJOR.

It is one of the finest examples of a multiple star system. Mizar (the Horse) forms a wide visual binary with the 4th-magnitude star Alcor (the Rider), at a separation of 12′. The two stars have nearly the same proper motion, and the pair is visible to the naked eye. Mizar itself is a close visual pair, with bluish-white components of apparent magnitude 2.27 and 3.95, approximately 14″ apart. The spectral types are A2Vp and Am, respectively. The spectrum of the primary star, Mizar A, shows periodic doubling of the absorption lines due to the Doppler effect, caused by the motion of two stars around their common center of mass. The orbital period of this spectroscopic binary is 20.5 days, and the stars have been resolved with an interferometer, the mean angular separation being only 0.012″. By combining the elements of the spectroscopic orbit with those from the astrometric orbit, the masses of the two stars have been derived, and are both about 2.5 times that of the Sun. Mizar B is also a spectroscopic binary with a period of 176 days. *See* DOPPLER EFFECT; SPECTRAL TYPE.

Mizar is a member of the group of stars known as the Ursa Major moving group, which has an estimated age of 3×10^8 years. David W. Latham

Mobile communications

Radio communication in which one or both ends of the communication path are movable. The term "mobile" refers to movement of the radio rather than association with a vehicle (for example, handheld portable radios are included in the definition). The movement need not occur during the communications. The Federal Communications Commission (FCC) licenses and regulates non-federal-government radio activity in the United States, while the National Telecommunications and Information Administration (NTIA) oversees federal-government uses. Other countries have similar agencies. International coordination is afforded through the International Telecommunication Union (ITU) and international treaty. Historically, the term "mobile radio" indicated systems where the user terminal was located in a vehicle. Current usage includes "outdoor" transportable devices that may not be actually moving during a communication session, as would be the case for a person using a laptop at a park bench.

Private land mobile services. Users who lease or purchase radio equipment for personal use fall into this category. Examples are public safety, special emergency, industrial, land transportation, and location radio services. Spectrum over a wide range of radio frequency bands is allocated; for example, low band (30–50 MHz), high band (150–174 MHz), and 900-MHz band (896–901 MHz, paired with 935–940 MHz). Dispatch is the normal mode of operation; that is, all members of the group hear all communications. To accomplish this, high-power, high-site, base station repeaters are generally used so that a single site covers the entire area of interest. Coverage radius varies with frequency band, local terrain, and permissible power levels, but values on the order of 20 mi (32 km) are commonplace. Where areas to be covered are even larger (for example, statewide police systems) or where coverage reliability must be greater than that possible from a single site

(for example, for ambulance communications), multiple sites can simulcast the communications. Current technology allows for data exchanges, vehicle location, and secure digitized voice. *See* RADIO SPECTRUM ALLOCATIONS.

Specialized mobile radio (SMR) is a type of mobile radio service in which individual users with business interests are licensed to operate their mobile, portables, and control stations on channel pairs repeated by specialized mobile radio base stations. Full interconnection to the public switched telephone network (PSTN) is possible. To boost the spectrum efficiency of specialized mobile radios relative to shared repeaters already in use, the FCC required that channels be trunked. Trunking in the context of radio systems means not only sharing equipment but sharing frequencies as well. Trunking channels means that when a user wishes to place a call it can be served by any one of the channel pairs that is available. The peak capacity that is obtained through trunking is no different than that obtained with single-channel operation, but the average grade of service is much improved. For example, with conventional operation, if each channel is busy half the time, the probability of being blocked from service is 50%. However, when five channels are trunked, a new call request will be blocked only if all channels are simultaneously busy. The probability of this is only about 7% according to Erlang-B traffic theory.

Paging service. Although paging is primarily a one-way radio system, two-way operation with such functions as page acknowledgment and short message reply is available. Typical paging operation is as follows. Via the public switched telephone network, the land party calls a general telephone number for the paging system to which the desired party subscribes. When the call is answered, a number unique to the subscriber is entered. Paging control equipment routes the call to appropriate paging transmitters (multiple transmitters are generally required to assure high reliability over large coverage areas). The transmitters simulcast coded radio signals that are recognized only by the desired individual's paging receiver. The user is alerted by a series of beeps or by mechanical vibration.

Some types of paging receivers display digits and letters (alphanumeric displays) that allow the calling party's number and a brief message to be displayed, and a message operator becomes unnecessary. Since display of paging messages involves little information, thousands of users can share a paging channel, thus making the system extremely spectrum efficient. Other types of paging receivers provide for brief voice messages following the alert (tone and voice). *See* RADIO PAGING SYSTEMS.

Public land mobile service. Mobile telephone service (MTS) in the United States initially involved wireline and radio common carriers. The earliest mobile telephone systems operated in push-to-talk mode with calls handled by special operators. The advent of improved mobile telephone service (IMTS) allowed automatic, duplex operation, with customers able to dial to and receive calls from anywhere in the world. Custom features such as call waiting, three-way calling, speed calling, and voice mail became available, as did data information exchange and facsimile transmission when modems were attached to the radios.

Trunked duplex operation (also known as frequency-division duplexing or FDD) was facilitated by assigning the radio channels in pairs, sufficiently separated in frequency such that the transmitted signal did not desensitize the receiver. Thus, conversations in the land-to-mobile direction and the mobile-to-land direction could occur simultaneously. Customers could subscribe for service in one area and travel to other areas and continue service, provided that the companies involved had previously agreed to honor such operation (termed roaming).

Cellular telecommunications service. Until the development of cellular mobile radio, the total number of mobile telephone users was severely limited by the modest number of radio channels allocated. Internationally, many different standards for cellular are in use (**Table 1**). While they differ in frequency, channel bandwidth, and control signaling, all these first-generation systems use analog frequency modulation (FM) supporting a single conversation per channel. *See* FREQUENCY MODULATION.

Cellular technology allows hundreds of thousands of users to be handled in a single metropolitan area. Rather than link into the telephone system from a single high-power, high site that covers the entire metropolitan area (**Fig. 1***a*), users are linked via many low-power, low sites. A single low site, of course, can cover only a limited area, termed a cell, but

TABLE 1. Analog cellular radio systems

Name*	Frequency, MHz: Mobile-to-land	Frequency, MHz: Land-to-mobile	Peak FM deviation, kHz	Channel bandwidth, kHz
AMPS	824–849	869–894	12	30
JTACS	870–885	925–940	5	25
NMT450	453–457.5	463–467.5	4.7	25
NMT900	890–915	935–960	4.7	25
TACS	890–905	935–950	9.5	25
ETACS	872–888	917–933	9.5	25

*AMPS = Advanced Mobile Phone System; JTACS = Japanese Total Access Communication System; NMT = Nordic Mobile Telephone (number following denotes band); TACS = Total Access Communication System; ETACS = Extended TACS.

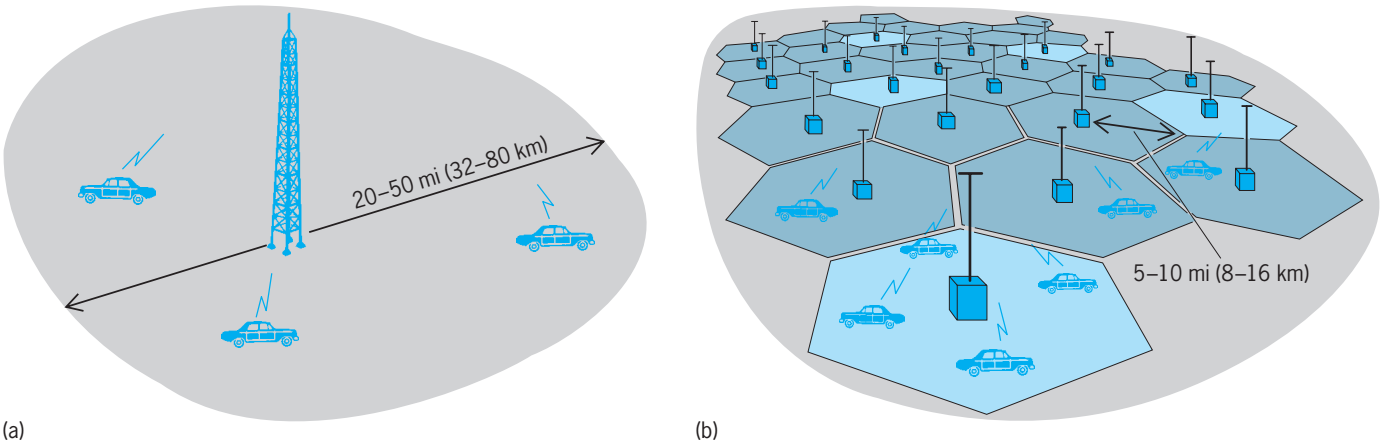

Fig. 1. Comparison of conventional and cellular mobile radio. (*a*) Conventional mobile radio covers a broad geographical region from a single, high site using high-powered base stations. (*b*) Cellular mobile radio divides the service area into cells with low-power, low-antenna base stations, allowing channels to be reused by cells that are sufficiently separated.

many low sites taken together can cover the entire metropolitan area (Fig. 1*b*). Spectrum efficiency stems from reusing the same frequency at all sites that are sufficiently separated. To further limit interference caused by such frequency reuse, each cell may be divided into sectors, and directive antenna patterns may be used.

An attractive feature of cellular radio is the ability to vary the size of cells in accordance with user density; hence, cell size can increase away from city centers. To sustain the reuse pattern with mixed cell sizes, power levels are tailored to produce comparable signal levels at all cell boundaries. Also, as more customers are added, radio channels can be created to serve them by constructing new base stations (hence, new cells) in geographical locations between existing cells. This concept is called cell splitting (**Fig. 2**). Geographical coverage of the system can be expanded as well by constructing new base stations on the periphery of the existing system and assigning frequencies consistent with the original reuse pattern.

Automatic, continuous coverage as users move across cell boundaries is provided by the cell handoff feature of cellular (also termed handover and automatic link transfer). Calls in need of handoff are recognized by monitoring call quality and comparing it to some required threshold. Handoff control procedures for first-generation analog frequency-modulation systems are also in operation.

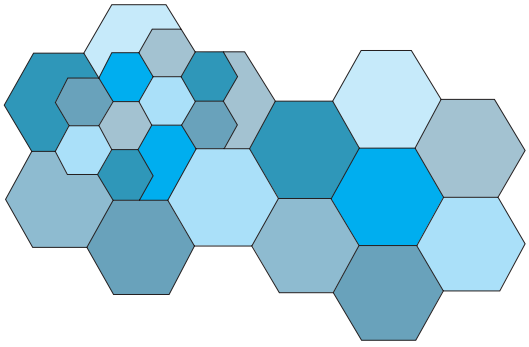

Fig. 2. Cell splitting in a cellular mobile radio system. In this example, nine cells have been created out of parts of four original cells. The additional capacity can be provided without increasing the available radio spectrum, because low-power transmitters in nearby cells can use the same frequencies without unacceptable interference.

Development of mobile radio services. Following more than a decade of trials and standardization, the first commercial cellular service in the United States began in Chicago in 1983, joining Japanese and Nordic systems that became commercial in 1979 and 1985 respectively. Initially, 40 MHz of spectrum was allocated to cellular service, and in 1986 this was extended to 50 MHz, in the 824–849 and 869–894 MHz bands. There are at least two licensed service providers in each of 305 metropolitan statistical areas and 428 rural statistical areas in the United States.

As acceptance of cellular technology has exceeded original expectations, concerns about the ability of the technology to support the large number of users has arisen. Practical limits on the cell-splitting concept occur due to difficulty in controlling interference as cell size shrinks, to capital investment requirements for continued building of new cells, and to local zoning issues. System modifications using split or offset channels have been developed to boost the effective number of simultaneous conversations possible per cell. For example, Narrow-band Advanced Mobile Phone Service (NAMPS) provides three channels in the bandwidth of a single 30-kHz channel of the Advanced Mobile Phone System (AMPS), the original United States analog cellular standard. Likewise, the narrow-band version of the Japanese analog cellular, the Japanese Narrow-band Total Access Communications System (JNTACS), provides two channels in the bandwidth of a single 25-kHz Japanese Total Access Communication System (JTACS) channel. Along with narrowing the channel bandwidth, these modified systems generally have improved handoff by including mobile participation in the process. Data operation has likewise improved with the introduction of cellular digital

TABLE 2. Digital cellular radio systems

Name*	Frequency, MHz: Mobile-to-land	Frequency, MHz: Land-to-mobile	Channel bandwidth	Conversations per channel	Multiplex method
GSM	890–915	935–960	200 kHz	8	TDMA
IS-54	824–849	869–894	30 kHz	3	TDMA
IS-95	824–849	869–894	1.25 MHz	20–60	CDMA
JDC	810–826	940–956	25 kHz	3	TDMA
MIRS	806–824	851–869	25 kHz	6	TDMA

*GSM = Global System for Mobile Communications; IS = Interim Standard; JDC = Japan Digital Cellular; MIRS = Motorola Integrated Radio System.

packet data (CDPD). This allows data traffic to be handled in short pieces or packets during idle periods on the voice servers, a major improvement over circuit-switched data, which require a voice server continuously even though most of the time no real exchange of data is occurring.

2G digital cellular service. Despite the improvements in analog cellular operation, schemes that digitally encode speech and suitably modulate the carrier have been adopted in second-generation (2G) cellular systems (**Table 2**). Such schemes handle data traffic as well as voice traffic. *See* MODULATION.

European digital technology (GSM, or Global System for Mobile Communications) uses 124 carriers of 200 kHz each over 25 MHz of spectrum allocated for each direction of communication (890–915 MHz and 935–960 MHz). Each 200-kHz carrier uses time-division multiple access (TDMA) with eight user time slots per carrier; hence, eight simultaneous conversations can be supported. *See* MULTIPLEXING AND MULTIPLE ACCESS.

Several standards for digital cellular in the United States and Canada exist:

1. Interim Standard 54 (IS-54) specifies a 30-kHz carrier using time-division multiple access with three user time slots per carrier. The existing cellular spectrum allocation is used, with analog FM carriers being gradually displaced by digital channels. This required the initial use of dual-mode radios.

2. Interim Standard 95 (IS-95) spreads all signals simultaneously over the same bandwidth and separates them via code-division multiplex access (CDMA). With the help of precise power to control near-far interference problems (that is, the tendency of nearby users to obliterate base station reception of distant, weak users), the same bandwidth can be reused in each and every cell. This is the means of achieving high spectral efficiency. Speech is encoded at a variable rate according to voice activity. Again, dual-mode radios were initially required.

3. MIRS is the result of a 1990 waiver request to the FCC to operate a cellular system in the frequency bands allocated to specialized mobile radio. Such systems are called enhanced SMRs (ESMRs) and provide six TDMA time slots in 25-kHz channels.

Second-generation digital cellular systems have attempted to improve handoffs by using more meaningful measurements of quality such as bit error rate (BER) and by having the subscriber units assist in the handoff process. IS-95 systems even introduce the concept of soft handoff, wherein the instantaneously best-received signal from a number of cell sites is used.

Cordless telephones. Cordless telephones provide a wireless extension to the public switched telephone network over an extremely limited range. Many units rely on mature analog radio technology that provides only nominal security and almost no interference protection.

In Europe and many Asian countries, there are enhanced cordless telephone services that use digital modulation. The first of these services, known as CT-2, for second-generation cordless telephone, uses frequency-division multiple access (FDMA) and time-division duplexing (TDD); each conversation uses a selected frequency exclusively (FDMA), but only one frequency is used and each direction of communication takes turns using it (TDD). By comparison, analog cordless telephones are also FDMA but use frequency-division duplexing (FDD) in which each direction of communication uses a different radio channel.

Digital European Cordless Telephone (DECT) is another (TDD) system that was developed primarily for in-building business application. In contrast to CT-2, DECT uses TDMA on each radio channel and can support up to 12 user voice channels on a given radio-frequency carrier. DECT uses a dynamic channel allocation (DCA) scheme in which a radio can change carrier frequencies every TDMA frame, if necessary, to avoid interference. In contrast, CT-2 and analog cordless telephones use the same frequency for the duration of a call and thus have no interference avoidance mechanism.

Another TDMA/TDD cordless telephone system, the Personal Handyphone System (PHS), developed in Japan, supports up to four voice circuits in each radio-frequency channel but uses much less bandwidth than DECT; channel spacing for PHS is 300 kHz, while the spacing between DECT carriers is 1.728 MHz.

Personal communication systems (PCS). The great demand for cellular phones and related wireless services has been addressed to some extent by the addition of new spectrum, by the introduction of narrow-band and digital cellular systems, and by cell splitting where practical. Acknowledging that these techniques for increasing capacity would quickly be exhausted, most countries have allocated additional spectrum for mobile and portable communications.

Since these frequency bands have much more spectrum that those previously allocated to cellular service, a greater variety of services are possible.

A spectrum of 120 MHz in the 1850–1910 and 1930–1990 MHz bands was allocated for licensed personal communication system (PCS) operation in the United States, and 20 MHz of spectrum in the 1920–1930 MHz band for unlicensed operation, split evenly between voice (isochronous) and data (asynchronous) applications. The spectrum allocation for licensed PCS operation is divided into six frequency blocks, three of which contain 30 MHz of spectrum, and the other three 10 MHz. It is thus possible in a given region to have as many as six competing service providers, in addition to the two 900-MHz cellular service providers.

The unlicensed PCS band was authorized to spur innovation and to encourage rapid development of service by allowing PCS equipment to be deployed without going through a licensing process. However, equipment designed to operate in this band must comply with a spectrum etiquette developed by the industry. This etiquette is a set of rules and requirements that facilitates simultaneous use of multiple systems in a common portion of the spectrum.

3G systems. In response to the ever-increasing demand for higher capacity and new services, the ITU defined the standards for third-generation systems, also known as International Mobile Telecommunications 2000 (IMT-2000). The ITU specified minimum bit rates of 144 kilobits per second (kbps) in mobile (outdoor) environments and 2 megabits per second (Mbps) in fixed (indoor) environments.

The IMT-2000 initiative gave birth to the Third-Generation Partnership Project (3GPP) for the evolution of GSM networks and the Third-Generation Partnership Project-2 (3GPP2) for the evolution of CDMA networks. The bands 1885–2025 MHz and 2110–2200 MHz are intended for use, on a worldwide basis, by administrations wishing to implement IMT-2000.

The 3GPP is a global initiative under the auspices of the European Telecommunications Standards Institute (ETSI). The 3GPP scope is to generate a global technical specification for a third-generation mobile system based on evolved GSM core networks and their corresponding radio technologies in both FDD and TDD modes. The scope grew to include the evolution of the specifications for the General Packet Radio Service (GPRS) and EDGE (Enhanced Data Rates for GSM Evolution). A permanent project support group called the Mobile Competence Centre (MCC) has been established to ensure the efficient day-to-day running of 3GPP. The MCC is based at the headquarters of the European Telecommunications Standards Institute (ETSI) in Sophia Antipolis, France. A similar partnership project, 3GPP2, was created to define global specifications for a third-generation mobile system based on the evolution of the ANSI/TIA/EIA-41 core network (developed by the American National Standards Institute, the Telecommunications Industry Association, and the Electronic Industries Alliance) and the corresponding radio access technology (Code-Division Multiple Access-2000 or CDMA2000). Members of both 3GPP and 3GPP2 include ARIB (Association of Radio Industries and Businesses, Japan), CCSA (China Communications Standards Association, China), TTA (Telecommunications Technology Association, Korea), and TTC (Telecommunications Technology Committee, Japan).

In 1999, the ITU approved five radio interfaces for IMT-2000 standards as part of the ITU-R M.1457 recommendation. These are:

1. CDMA Direct Spread, also known as UTRA-FDD (Universal Terrestrial Radio Access-Frequency Division Duplexing) or UMTS (Universal Mobile Telecommunications System), including W-CDMA (Wideband CDMA). [The Japanese service provider NTT DoCoMo is using an ARIB-standardized W-CDMA solution.] UMTS was developed by 3GPP. While there are some differences, the terms UMTS and W-CDMA are used to refer to both the ITU and the ARIB W-CDMA standards.
2. IMT-2000 CDMA Multi-carrier, also known as CDMA2000-3xRTT (3 Times Radio Transmission Technology), developed by 3GPP2. IMT-2000 CDMA2000 includes 1xRTT (1 Times Radio Transmission Technology) components, like CDMA2000 1xRTT EV-DO (Evolution Data Optimized).
3. IMT-2000 CDMA TDD, also known as UTRA TDD and TD-SCDMA (Time Division–Synchronous Code-Division Multiple Access). TD-SCDMA is developed in China and supported by the TD-SCDMA Forum.
4. IMT-2000 TDMA Single Carrier, also known as UWC (Universal Wireless Communications)-136 (EDGE), supported by UWCC (Universal Wireless Communications Forum).
5. IMT-2000 DECT, supported by DECT Forum.

There were estimated to be about 2 billion mobile subscribers in March 2005, about one-third the population of the world, with just over 50% subscribers utilizing 2G technologies, a negligible fraction still using first-generation (analog) technology, and just under half utilizing either enhanced digital technologies, known as 2.5G, or full ITU-approved 3G technologies. There were reportedly about 256 million CDMA subscribers in March 2005. The first commercial CDMA2000 system was deployed in October 2000. TD-SCDMA was not yet in commercial deployment as of 2006. DECT is generally considered to be better suited for in-building and enterprise systems.

Two key members of the global family of third-generation systems identified by the ITU are the Universal Mobile Telecommunication System (UMTS) from 3GPP, intended for new spectrum, and the Code-Division Multiple Access–2000 (CDMA2000) from 3GPP2, also known as CDMA2000 1xRTT, intended for deployment by operators within their existing spectrum previously allocated for CDMA2000. Both systems can serve multiple users simultaneously by assigning a distinct code to each user (code-division multiple access). Each code is a unique sequence of pseudorandom sign reversals applied to each user symbol, thereby spreading the user

TABLE 3. Third-generation mobile communication standards

	CDMA2000	W-CDMA (UMTS)	W-CDMA (ARIB)
Frequency band*	Any	1920–1980 MHz uplink, 2110–2170 MHz downlink	1920–1980 MHz uplink, 2110–2170 MHz downlink
Minimum bandwidth required	2 × 1.25 MHz	2 × 5 MHz	2 × 5 MHz
Reuse factor	1	1	1
Chip rate†	1.2288 Mcps	3.84 Mcps	4.096 Mcps
Maximum user data rate	144–230 kbps	384 kbps	2 Mbps
Frame duration	5, 10, 20, 40, or 80 ms	10 ms	10 ms
Spreading factors*	8,..., 128	4,..., 256 uplink, 4,..., 512 downlink	1,..., 256

*Uplink = transmissions from the mobile terminal to the base station; downlink = transmissions from the base station to the mobile terminal.
†Mcps = megachips per second (1 megachip = 2^{20} chips).

symbol across a given number of code chips to obtain a desired spreading factor. Spreading each user symbol with a unique code provides protection against interference from adjacent transmissions, making it possible for all transmissions to use the same frequency with complete reuse. To increase efficiency, multiple user symbols are grouped into frames. The main features of the CDMA2000 and W-CDMA third-generation standards are summarized in **Table 3**. *See* SPREAD SPECTRUM COMMUNICATION.

3G extensions for data. With the increased interest in Internet usage in particular and data applications in general, extensions of the 3G standards have been developed to increase performance by taking advantage of the distinct nature of data as compared to voice traffic. For instance, introducing delay can be very annoying in conversational speech, and thus it is maintained under strict control in voice-optimized systems. In contrast, data packets can be subject to variable delays of up to several hundred milliseconds to exploit better channel conditions with no noticeable degradation in the user experience. The data-optimized evolutions of the third-generation standards are CDMA2000 1xEV-DO (Evolution Data Optimized) and High Speed Down Link Packet Access (HSDPA) for W-CDMA. HSDPA was designed to support simultaneous voice and data traffic in the same 5-MHz channel. HSPDA achieves a peak rate of 14.4 Mbps and, in realistic deployments, it is expected to deliver on average rates in the 400–700 kbps range. 1xEV-DO, first deployed in 2003, devotes a complete 1.25-MHz channel to data users only. This design philosophy is not being strictly followed in that an emerging new service is the delivery of voice in data format, using IP (Internet Protocol) packets (Voice over IP or VoIP). In another departure from previous usage, in CDMA200 1xEV-DO, the multiple CDMA codes are used to increase the bit rate delivered, while communicating with one user at a time (in the downlink, that is, in transmissions from the base station to mobile terminals), instead of using these codes to separate transmissions to multiple users, as is the case in CDMA2000. With this change it is possible to schedule preferentially transmissions to users that experience good channel conditions, in what is known as proportional fair scheduling. This scheduling technique results in throughput improvements, also referred to as multiuser diversity. CDMA2000 1xEV-DO delivers a peak rate of 2.4 Mbps and an average rate in the 300–500 kbps range. Further evolution to 1xEV-DO, known as 1xEV-DO RevA, should achieve faster packet set-up, reduced latency, and improved data rates. *See* VOICE OVER IP.

Smart antenna technologies for 4G systems. With the deployment of third-generation systems well underway, system designers and standard organizations have turned their attention to the evolution of systems "beyond 3G" and to 4G proper. While there is no universal definition of what will constitute a true 4G system, it is possible to extrapolate from the previous generational transitions that a key element will be a significant increase (that is, a tripling or at a least doubling) in system capacity, throughput, and average bit rates. It is expected that such increases will be made possible in large part by advanced antenna technologies.

Antenna diversity. Perhaps one of the best-known multiple-antenna techniques is antenna diversity, where multiple antennas are used to maximize the chances that the radio channel condition to at least one of the antennas will be satisfactory. Clearly, the more antennas, the higher this chance will be. In the technique of selection diversity, the signal from the best antenna is selected. The optimum approach, known a maximal ratio combining, is to perform a sum of all received signals, weighting each in proportion to its signal-to-noise ratio, thus emphasizing the better-quality signals.

This technique, using two antennas, is already in use at the base station receiver for uplink traffic (that is, transmissions from mobile terminals to the base station). It is expected that future base stations may use four-antenna diversity and handsets, and mobile terminals will also employ multiple diversity antennas. While less effective due to overall power limitations, diversity antennas may also be used in transmission.

Beam-forming antennas and future enhancements. Wireless systems can benefit greatly from using antennas that focus radio energy in the geographical direction of interest. In the past, the use of such beam-forming antennas was limited to fixed point-to-point installations. Modern electronics make it possible to coordinate the signals from a set of antenna elements to generate, and even steer, a beam in the

desired direction, by properly phasing the signals associated with each antenna element. Thus, beamforming techniques can be used in cellular systems. The beam can be fixed or electronically steered to track user movements. The beam can be configured in reception as well and in transmission. Most benefit will be observed when the signal directions are constrained to a narrow angle. Thus, beam-forming techniques are, in general, better suited for base station sites high above the surrounding clutter, rather than at the user end, where signals may appear to be arriving from multiple objects reflecting and diffracting the signals in almost all directions.

Spatial-division multiplexing. With sufficient antenna elements, it is possible to form more that a single beam. In this case, multiple messages intended for users in different directions can be associated with correspondingly many beams, generated by the appropriate beam-forming weights for multiple beams in each of the desired directions. By judiciously choosing users that are placed in direction separated by at least the beam width, it is possible to deliver simultaneous messages to all these users without interference. Thus, a system capacity increase factor equal to the number of simultaneous beams may be achieved.

MIMO systems. Link capacity can be greatly increased with a promising technique using multiple antennas at the transmitter and the receiver, known as multiple-input, multiple-output (MIMO). In MIMO systems, the transmitter sends multiple simultaneous messages in the same channel bandwidth without increasing the transmitted power. In a scattering environment, the receiver can process the signals at multiple receive antennas and separate the multiple transmitted messages, thereby creating multiple parallel virtual channels. Hence the link capacity is increased in direct proportion to the number of the multiple virtual channels. In this case, most of the processing is carried out at the receiver. For cost reasons, it is attractive in the downlink to shift the processing burden to the base station transmitter. In this case, the terminal feeds back the propagation channel estimates to the base station. With this information, the base station can compute generalized beam-forming weights to ensure that, at the receiver antenna locations, the multiple messages are received without any significant interference between the multiple transmitted messages. In this case, it is even possible to physically separate the receiving antennas and use this technique for the case in which the (multiple-antenna) base station is sending multiple messages to separate, single-antenna users. Furthermore, with this technique, essentially no changes are needed at the mobile user terminal, making it possible to use unmodified legacy terminals. This technique is sometimes also described as generalized beamforming.

Satellite telephony. There are many scenarios where mobile communications services are needed, yet no network infrastructure, such as base station transceivers, is available. In this case, satellite-based mobile services are extremely valuable. There are three types of satellite-based mobile systems, depending on the orbit in which the satellite base transponder is placed. The first type of system utilizes geostationary Earth orbit (GEO) satellites, where the orbit is chosen such that the satellite appears stationary with respect to Earth. This simplifies the Earth station system design. However, geostationary orbits are at such a large distance from Earth (about seven Earth radii) that large transmission delays are introduced. More recently, lower orbits have been used to reduce delay. These are used in medium Earth orbit (MEO) and low Earth orbit (LEO) types of systems. In these cases, the satellite is in movement with respect to the Earth station location. Hence, the system must provide for user terminal handoffs, as the serving satellite disappears below the horizon and another satellite appears above the horizon. In these systems, many satellites are needed to ensure coverage at all times. *See* COMMUNICATIONS SATELLITE.

Reinaldo A. Valenzuela

Bibliography. J. D. Gibson (ed.), *The Mobile Communications Handbook*, 2d ed., CRC Press, 1999; T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2d ed., Prentice Hall, 2002; R. Steele and L. Hanzo (eds.), *Mobile Radio Communications*, 2d ed., Wiley, 1999.

Mode of vibration

A characteristic manner in which vibration occurs. In a freely vibrating system, oscillation is restricted to certain characteristic patterns of motion at certain characteristic frequencies; these motions are called normal modes of vibration. An ideal string, for example, can vibrate as a whole with a characteristic frequency defined in the equation below, where L

$$f = \frac{1}{2L}\sqrt{\frac{T}{m}}$$

is the length of string between rigid supports, T the tension, and m the mass per unit length of the string. Displacements of different parts of the string are governed by a characteristic shape function. Specifically, the motion of any part of the string is proportional to $\sin(\pi x/L)\sin(2\pi ft)$, where x is the distance of the part of the string from a fixed end and t is the time. This simplest kind of vibration is the first, or fundamental, mode of vibration of the string; its frequency is the fundamental frequency. All parts of the string vibrate with the same frequency, and move outward from equilibrium at the same time.

The string is also capable of vibrating in two segments, one of which goes outward from equilibrium in a positive direction at the same time as the other part goes outward in a negative direction, and conversely. Again, the motion of any part of the string can be described by the product of a space function by a sinusoidal function of time: $\sin(2\pi x/L)\sin(2 \times 2\pi ft)$. All parts of the string move together as a sinusoidal function of the time and at the same frequency; the space function governs the

motion in opposite directions. The frequency of the second mode of vibration is twice that of the first mode. Similarly, modes of higher order have frequencies that are integral multiples of the fundamental frequency.

Because the frequencies are in the ratios 1:2:3 . . . , the modes of vibration of an ideal string are properly called harmonics. Not all vibrating bodies have harmonic modes of vibration, however. The ideal drumhead, for example, vibrates freely with frequencies in the ratios 1:1.59:2.14:2.30 . . . In fact, most real systems vibrating freely have modes of vibration whose frequencies are not exactly in the ratios of integers. *See* HARMONIC (PERIODIC PHENOMENA); VIBRATION.

Robert W. Young

Model theory

The body of knowledge that concerns the fundamental nature, function, development, and use of formal models in science and technology. In its most general sense, a model is a proxy. A model is one entity used to represent some other entity for some well-defined purpose. Examples of models include: (1) An idea (mental model), such as the internalized model of a person's relationships with the environment, used to guide behavior. (2) A picture or drawing (iconic model), such as a map used to record geological data, or a solids model used to design a machine component. (3) A verbal or written description (linguistic model), such as the protocol for a biological experiment or the transcript of a medical operation, used to guide and improve procedures. (4) A physical object (scale model, analog model, or prototype), such as a model airfoil used in the wind-tunnel testing of a new aircraft design, or an electronic circuit used to simulate the neural activity of the brain. (5) A system of equations and logical expressions (mathematical model or computer simulation), such as the mass- and energy-balance equations that predict the end products of a chemical reaction, or a computer program that simulates the flight of a space vehicle. Models are developed and used to help hypothesize, define, explore, understand, simulate, predict, design, or communicate some aspect of the original entity for which the model is a substitute.

Formal models are a mainstay of every scientific and technological discipline. Social and management scientists also make extensive use of models. Indeed, the theory of models and modeling cannot be divorced from broader philosophical issues that concern the origins, nature, methods, and limits of human knowledge (epistemology) and the means of rational inquiry (logic and the scientific method). Philosophical notions of causality are also central to modeling. *See* LOGIC; SCIENTIFIC METHODS.

Advantages of models. Models are so pervasive because a model is usually more accessible to study than the system modeled. Models typically are less costly and less time-consuming to construct and test. Changes in the structure of a model are easier to implement, and changes in the behavior of a model are easier to isolate, understand, and communicate to others. A model can be used to achieve insight when direct experimentation with the actual system is too dangerous, disruptive, or demanding. A model can be used to answer questions about a system that has not yet been observed or built, or even one that cannot be observed or built with present technologies.

Role of model theory. Every discipline develops its own models and its own approach and techniques for studying these models. The specific models developed in different disciplines may differ in subject, form, and intended use. However, basic concepts such as model description, validation, simplification, and simulation are not unique to any particular discipline. Model theory seeks a formal logical and axiomatic understanding of the underlying concepts that are common to all modeling endeavors.

Mathematical models. General and mathematical systems theory have stimulated many of the important developments in model theory. This article outlines the concepts of model theory as these relate to mathematical and computer simulation models of systems. Mathematical models are particularly useful, because of the large body of mathematical theory and technique that exists for the study of logical expressions and the solution of equations. The power and accessibility of digital computers have increased the use and importance of mathematical models and computer simulation in all branches of modern science and technology. A great variety of programming languages and applications software are now available for modeling, computational analysis, and system simulation. Many applications that formerly relied on other types of models now also use mathematical models and computer implementations extensively. *See* DIGITAL COMPUTER; SIMULATION; SYSTEMS ANALYSIS; SYSTEMS ENGINEERING.

Systems as models of systems. A system is a collection of entities that interact. Every system has a boundary. Entities not included within the boundary are known as the environment of the system.

A system is characterized by a set of attributes, that is, a set of quantities which assume values. Static attributes have fixed values and are called the parameters of the system. Dynamic attributes can assume different values at different times and at different points in space and are known as the descriptor variables of the system.

Three kinds of descriptor variables are input variables, output variables, and state variables (**Fig. 1**). Input variables represent the action of the environment on the system. They may be represented by

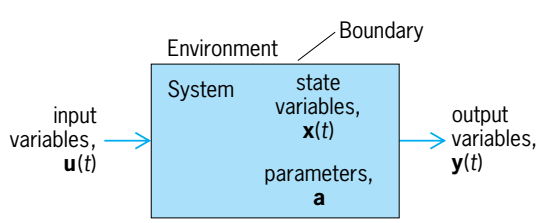

Fig. 1. Conceptual model of a system.

Eq. (1), which indicates that there are m input vari-

$$\mathbf{u}(t) = \{u_1(t),\ u_2(t),\ \ldots,\ u_m(t)\}\epsilon\mathbf{U} \qquad (1)$$

ables, u_1, u_2, ..., u_m; that each depends on time t; that collectively they form an m-component vector $\mathbf{u}$; and that they are members of a range set $\mathbf{U}$. Output variables, Eq. (2), represent the action

$$\mathbf{y}(t) = \{y_1(t),\ y_2(t),\ \ldots,\ y_q(t)\}\epsilon\mathbf{Y} \qquad (2)$$

of the system on the environment. State variables, Eq. (3), represent the internal condition of the sys-

$$\mathbf{x}(t) = \{x_1(t),\ x_2(t),\ \ldots,\ x_n(t)\}\epsilon\mathbf{X} \qquad (3)$$

tem. Changes in the input precipitate changes in the state, which in turn are observed as changes in the output. Changes in the output are assumed to have no direct effect on inputs. *See* CONTROL SYSTEMS; LINEAR SYSTEM ANALYSIS.

While the state of a system is unique, the state variables that describe this internal condition are not. Instead, the state variables are any (minimal) set of variables that has the following two properties:

1. Knowledge of the values of the state variables $\mathbf{x}(t_i)$ at any instant of time t_i is sufficient to determine uniquely the values of any possible set of output variables $\mathbf{y}(t_i)$ at that same instant of time. That is, there exists an output function λ: $\mathbf{x} \rightarrow \mathbf{y}$ defined for all t_i in the time base that maps current states into current outputs.

2. Knowledge of the values of the state variables $\mathbf{x}(t_i)$ at the time t_i, together with knowledge of values of the inputs $\mathbf{u}(t)$ over the time interval $[t_i,t_j]$, $t_i \leq t \leq t_j$, is sufficient to determine uniquely the values of the state variables $\mathbf{x}(t_j)$ at time t_j. That is, there exists a state transition function δ: $(\mathbf{X},\mathbf{U}) \rightarrow \mathbf{X}$ defined for all $[t_i,t_j]$ in the time base that maps current states and inputs into future states. (This is known as the Markov property.)

The pattern of changes exhibited by the state (or output) variables in response to any particular input is known as a state (or output) trajectory of the system. The collection of all possible state (or output) trajectories, corresponding to all possible inputs, is known as the state (or output) behavior of the system. The behavior of the system is generated by the functions δ and λ, which depend on the system parameters and possibly on time and spatial location as well.

The relationships among the input, state, and output variables of a system that give rise to its behavior are known as the structure of the system. A system, therefore, can be completely defined by the algebraic structure $\langle \mathbf{U}, \mathbf{X}, \mathbf{Y}, \delta, \lambda, t\rangle$, where $\mathbf{U}$, $\mathbf{X}$, and $\mathbf{Y}$ are understood to be the range sets of the input, state, and output variables.

If a system S has the structure $\langle \mathbf{U}, \mathbf{X}, \mathbf{Y}, \delta, \lambda, t\rangle$, then a model of S is just some other system S' with structure $\langle \mathbf{U}',\mathbf{X}',\mathbf{Y}',\delta',\lambda',t'\rangle$. The system S' is used as a substitute for S for some specific purpose. Model theory deals with the nature and adequacy of the various possible relationships between these two systems for the purpose at hand.

Modeling elements. Comprising five basic elements, a formal theory of modeling and simulation that builds on the ideas of systems was proposed.

Real system. This is the system modeled. It is simply a source of observable data, in the form of input/output pairs $(\mathbf{u}(t), \mathbf{y}(t))$. Typically, there are no other clues available to determine its structure.

Base model. This is the investigator's image or mental model of the real system. It is a system that is capable (at least hypothetically) of accounting for the complete behavior of the real system. If the real system is highly complex, the base model also is highly complex. The cost, time, and difficulty of realizing the base model explicitly most often are prohibitive, unwarranted, or impossible. As a practical matter, therefore, the structure of the base model is at best partially known to the investigator.

Experiment frame. This is the set of limited circumstances under which the real system is to be observed and understood for the purpose of the modeling exercise. It is a restricted subset of the observed output behaviors, as indicated in Eq. (4).

$$(\mathbf{u}(t),\ \mathbf{y}(t))_e \subseteq (\mathbf{u}(t),\ \mathbf{y}(t)) \qquad (4)$$

Lumped model. This is the concept most often associated with the term model. It is a system which is capable of accounting for the output behavior of the real system, under the experiment frame of interest. It is an explicit simplification and partial realization of the base model. Its structure is completely known to the investigator.

Computer. This is the means by which the behavior of the lumped model is generated. The computer is not necessarily a digital computer. For simple models, it may represent the explicit analytical solution to the model equations, worked out by hand. For more complex models, however, the computer may need to generate individual trajectories step by step, based on instructions provided by the model. This step-by-step process is what is most often associated with the concept of simulation, and is usually conducted by using a digital computer.

Modeling relationships. Model theory addresses the relationships among modeling elements. Three modeling relationships are validation, simplification, and simulation.

Validation. This concerns the relationship between models and the real system. The objective of validation is to ensure that a model matches the system modeled, so that the conclusions drawn about the model are reasonable conclusions about the real system as well. A base model is valid to the extent that it faithfully reproduces the behavior of the real system in all experiment frames. On the other hand, a lumped model is valid to the extent that it faithfully matches the real system under the experimental frame for which it is defined. There can be many different lumped models that are valid, and a lumped model can be valid in one experiment frame and not another.

Model validation is a deep and difficult issue, and there are many different levels and interpretations of validity. For example, a model has face validity simply if it is accepted as reasonable for its intended purpose by people who are knowledgeable about the system under study. A model is replicatively valid if its trajectories match the real system input/output data used in its development. A model is predictively valid if its trajectories also match experimental data not used in its development. A model is structurally valid if it is behaviorally valid and also matches the structure of the base model, that is, if it generates behavior within its experiment frame in substantially the same way that the real system is believed to generate this behavior. There are also many other forms of validity.

Simplification. This concerns the relationships between a base model and its associated lumped models. The objective of simplification is to achieve the most efficient and effective lumped model that is valid within the experiment frame for which it is defined. Simplification can be achieved in many ways. These include dropping relatively insignificant descriptor variables and their associated structures, replacing deterministic variables and structures with random variables and their generating functions, coarsening the range set of descriptive variables, and aggregating descriptor variables and structures into larger blocks. Many of the formal ideas associated with simplification, such as isomorphism and homomorphism, concern the preservation of structural similarities between mathematical systems.

Simulation. This concerns the relationship between models and the computer. The objective of simulation is to ensure that the computer faithfully reproduces the behavior implied by the model. The behavior of a lumped model must be distinguished from the correctness of its computer implementations or solutions, in the same way that the behavior of a real system must be distinguished from the validity of its models. While a valid model may have been developed, it is also necessary to have a correct simulation. Otherwise, the model solution cannot be used to draw conclusions about the real system. Formal ideas associated with simulation include the completeness, consistency, and ambiguity of the computer implementation. The process of matching a simulation to its lumped model sometimes is known as verification. Many of the same techniques used to validate models are also used to verify simulations.

Example. An example that illustrates the concepts of model theory involves the population changes of a region over time. Models of population dynamics are used to help assess future needs for housing, retail and commercial space, and schools and other public services, as well as to forecast revenues that will be available for these needs. The real system in this case might be a county, city, state, or nation, where census data are a source of information about past population trends. The base model comprises all that is known about how and why populations change, including factors that determine natural changes in the population (such as births, deaths, and aging) and influence migration among regions (such as economic and social conditions) [**Fig. 2**].

An experiment frame can be constructed to seek a very general understanding of population changes. The system boundary in this case corresponds to the geographic boundary of the region. A simple lumped model appropriate to this experiment frame might involve the following descriptor variables: total net migration into the region (the input), total regional

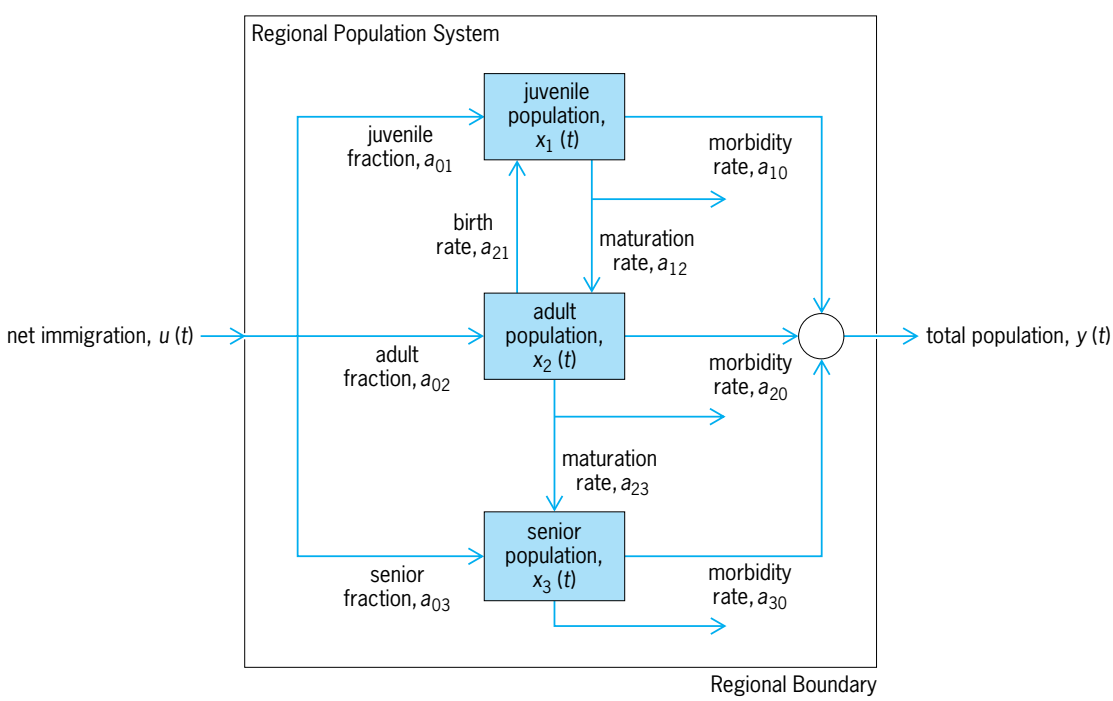

Fig. 2. Model of a regional population system.

population (the output), and the three state variables juvenile population, adult population, and senior population. The interactions among all these variables over time can be expressed in mathematical terms by a set of difference equations that represent a theory about how the state of the system (the values of the state variables) at a given time is determined by the state at some preceding time. The parameters of this model, whose values are determined from census data, reflect survival of the population within each age cohort, maturation from younger cohorts to older cohorts, births, and the age distribution of migrants. The output mapping defines the total population at any time to be the sum of all the population cohorts (juvenile, adult, and senior populations) at that time.

The model equations can be solved analytically, or a digital computer can be used to simulate the evolution of the state and output trajectories. Comparison of the model trajectories against actual census data would then provide a test of the validity of the model. K. Preston White, Jr.

Bibliography. G. Gordon, *System Simulation*, 2d ed., 1978; N. A. Kheir (ed.), *Systems Modeling and Computer Simulation*, 2d ed., 1995; M. Kutz (ed.), *Mechanical Engineers' Handbook*, 2d ed., 1998; A. M. Law and W. D. Kelton, *Simulation Modeling and Analysis*, 3d ed., 1999; M. G. Singh (ed.), *Systems and Control Encyclopedia*, 1988; B. P. Zeigler, *Theory of Modelling and Simulation*, 1976, reprint 1984.

Modeling languages

Graphical or textual languages that are used to build models by describing the qualities with which the models are concerned. A model represents something and includes some but not all qualities of what it represents. For example, an architect's model of a new development area concentrates on the physical appearance, while an aircraft designer's wind-tunnel model concentrates on aerodynamic qualities. When models are used in systems and software engineering, their important qualities are structure, behavior, communication, and requirements.

Need for models. Systems and software engineering are both characterized by a combination of large information content with a need for different people with different backgrounds and interests (system stakeholders) to understand at least some system aspects. During the development of systems, developers can see only a small part of the complete system at a time. They will then need models to assist them in understanding the context of their own work and the requirements on it. Models are useful prior to system development to study requirements and design alternatives, and to ensure that all stakeholders have a common understanding of why a system should be built or modified. They are also useful during and after development of a system to assist developers and users to navigate the system's structure and to understand the system's behavior. Finally,

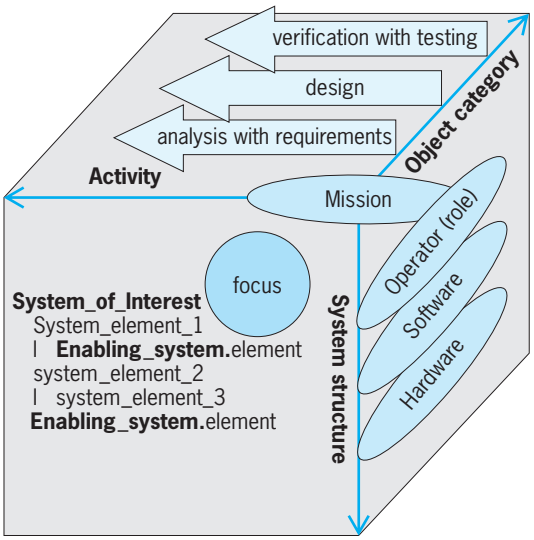

Fig. 1. Systems engineering work space.

models are essential for communication between different participants in projects and as a "corporate project memory." *See* MODEL THEORY.

Systems and software engineering. Systems engineers work with a collection of operator roles (people), software modules, and hardware modules, which are combined to complete one or more missions. Thus, they work in a systems engineering work space with three dimensions (**Fig. 1**): the activity dimension (what they work with), the object category dimension (what kind of object they work with), and the structure dimension (where they work in the system). *See* OBJECT-ORIENTED PROGRAMMING.

It is normally not humanly possible to retain an overall understanding of the system together with a deep insight into the system's details, so there is a need to focus somewhere in this space. A complete system model will then be useful to navigate the system and to ensure that the part that is focused on is consistent with its environment.

Obviously, software is a subset of systems engineering. However, since building the right software requires a correct understanding of the software's environment, it is advisable to include this environment in software engineering models. *See* SOFTWARE ENGINEERING; SYSTEMS ENGINEERING.

Unified Modeling Language diagrams. The Unified Modeling Language (UML™) has emerged as a standard for software engineering. This standard includes a set of diagrams:

Use Case diagram. This diagram shows the interaction between a user and a software system. This is basically a dependency relation where the user depends on the system.

Class diagram. This diagram is basically an enhanced entity-relationship diagram where the entities are classes with names, attributes, and actions. A "software class" can be seen as a "template" from which instances are created to implement the software. The diagram includes relations such as generalization, association, and aggregation. The diagram can

be used for conceptual modeling, for specification, and for documenting a software solution.

Sequence diagram. Message sequence charts have a long tradition for visualizing the interaction between concurrent processes communicating by way of messages. These charts are included in the UML as Sequence diagrams.

Collaboration diagram. This diagram resembles a block diagram with its boxes and arrows. It basically contains the same information as the Sequence diagram.

Package diagram. This diagram shows dependencies between major system components.

State diagram. This diagram visualizes behavior.

Activity diagram. This diagram is a development from flow charts.

Deployment diagram. This diagram shows the distribution of software objects on hardware nodes.

Component diagram. This diagram shows dependencies between components and can also be drawn to show aggregation (components included in other components).

The qualities of a software system that are primarily modeled by the UML are requirements, communication, structure, and behavior. The **table** shows how the diagrams support different qualities.

Extension of UML to systems engineering. At present there is a trend in the systems engineering community toward extended use of the UML for systems engineering. This trend is based on the fact that the UML views software as composed of modules called "classes" or "objects" characterized by their having (1) an offered interface, which contains services offered by the object, (2) an internal behavior, and (3) a required interface, with services required from other objects to complete the internal behavior.

This view is not limited to software, but can be extended to systems engineering when one includes not only software objects but also hardware and operator objects. An example of this extended use is a UML Component diagram (**Fig. 2**) which models a simple system to control the windows in a car.

Figure 2 shows that the car driver requires services from a switch bank and a button to control the car's windows. Further, these hardware control devices depend on software in the car's central computer and in the car's local computers, one in each door. The software in the local computer in the door depends in turn on a set of hardware devices in the door to complete the driver's commands.

Usage of UML diagrams

Diagram	Qualities supported: Requirements	Structure	Communication	Behavior
Use Case	X	—	—	—
Class	—	X	—	—
Sequence	—	—	X	—
Collaboration	—	—	X	—
Package	—	X	—	—
State	—	—	—	X
Activity	—	—	—	X
Deployment	—	X	—	—
Component	—	X	—	—

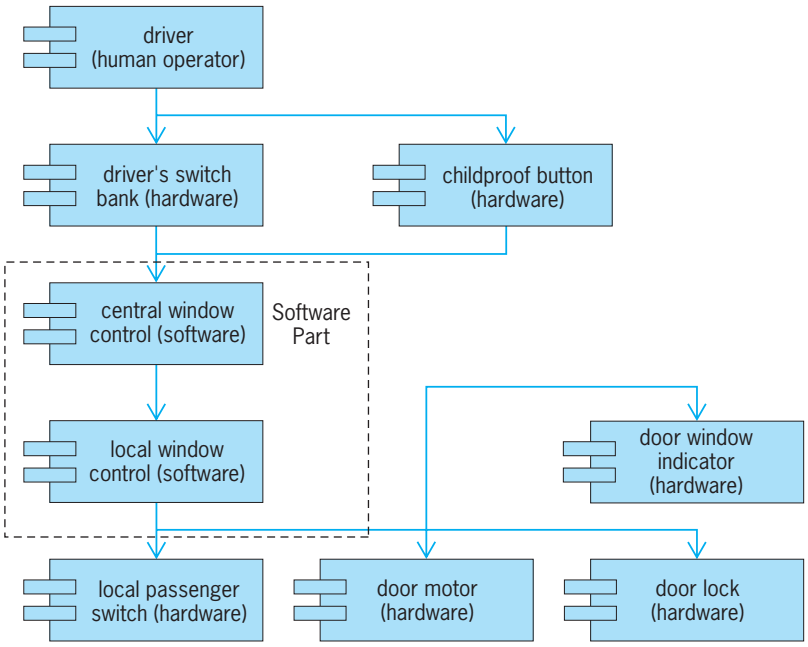

Fig. 2. Component diagram for car window control. The small rectangles signify that this is a component diagram.

Other modeling languages. While the UML is the emerging standard for modeling languages, there are several alternatives, including mathematical formal languages, petri nets, and formalized natural language.

Mathematical formal languages are used in so-called formal methods, with the main advantage being that these languages can be used to create formal specifications, which can then be checked automatically for compliance with a completed software system. The main disadvantage is that the mathematical language used can be difficult to understand for some stakeholders.

Petri nets give a graphical model of how concurrent processes communicate and influence each other through the transmission of "tokens."

A formalized natural language is obtained by simplifying a programming language for modeling use while keeping the "reserved words" and including variables of defined types. An example with part of the behavior for the software module "central window control" in Fig. 2 is:

```
begin
  case direction_up is
    when true =>
      send up_message (window_concerned)
      -- send message to move window
          -- upwards to the local control
      -- software for the concerned window in
          -- the car (driver,
      -- other_front, rear_left or rear_right)
    when false =>
      send down_message (window_concerned)
  end case
end
```

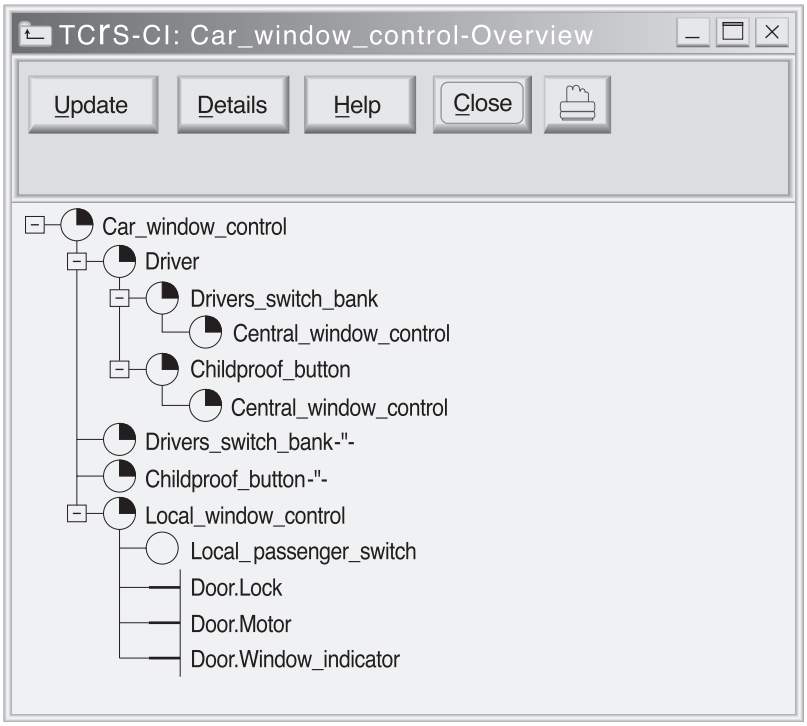

Fig. 3. Example of deep dependency structure.

See NATURAL LANGUAGE PROCESSING; PROGRAMMING LANGUAGES.

Tailoring of standards. The table shows how various diagrams in the UML address different aspects of modeling, and also shows that some aspects are covered by multiple diagrams. The obvious conclusion is that the standard needs tailoring prior to efficient use. Such tailoring should result in a set of modeling languages, some of which may be taken from within the UML, possibly modified, and some of which may be taken from outside the UML. Using a good set of modeling languages can be a blessing for a project, while careless selection, for example through uncritical acceptance of a "proven toolkit," may result in confusion.

An example of tailoring concerns a systems engineering project with the following requirements on the modeling languages used:

1. The system's missions shall be clearly seen.
2. Objects of categories Operator, Software, and Hardware shall be distinguished in the model.
3. The system's "deep dependency structure" shall be clearly seen.
4. System structure and behavior shall be modeled.
5. Stakeholders, without previous systems or software engineering experience, shall be able to understand the models immediately.

It may seem that the UML cannot satisfy much of these requirements, but tailoring can help. The Component diagram in Fig. 2 really shows a dependency structure with objects of different categories. The diagram can be modified by keeping only the objects, ordered according to how they depend on each other, and introducing a new "Mission" object at the top to state the mission "Car window control." The result is the diagram in **Fig. 3**, which is known as the deep dependency structure of the system.

This modified Component diagram has only one component (object) on each line, and each component depends on the components below and one step indented. The "little clocks" indicate completion status for the components, and the quotation marks indicate repeated parts of the diagram, when an object appears a second time. The result is that a way has been found to satisfy the requirements for structural models.

The next step is to construct behavioral models, where the UML diagrams have difficulties with ignorant stakeholders. A possible solution is then to select a text-based modeling language, as described above. Thus, it is possible to find a limited set of modeling languages that can be used for communication between stakeholders with minimum education.

Ingmar Ogren

Bibliography. E. Best, R. Devillers, and M. Koutny, *Petri Net Algebra*, Springer-Verlag, 2001; M. Fowler, *UML Distilled*, 3d ed., Addison-Wesley, 2004; C. Larman, *Applying UML and Patterns*, 3d ed., Prentice Hall, 2005; I. Ogren, On principles for model-based systems engineering, *Sys. Eng.*, 13(1):38–49, 2000; I. Ogren, Possible tailoring of the UML for systems engineering purposes, *Sys. Eng.*, 3(4):212–224, 2000.

Modem

A device that converts the digital signals produced by terminals and computers into the analog signals that telephone circuits are designed to carry. Despite the availability of several all-digital transmission networks, the analog telephone network remains the most readily available facility for voice and data transmission. Since terminals and computers transmit data using digital signaling, whereas telephone circuits are designed to transmit analog signals used to convey human speech, a device is required to convert from one to the other in order to transmit data over telephone circuits. The term modem is a contraction of the two main functions of such a unit, modulation and demodulation. The device is also called a data set. *See* INTEGRATED SERVICES DIGITAL NETWORK (ISDN); MODULATION.

In its most basic form a modem consists of a power supply, transmitter, and receiver. The power supply provides the voltage necessary to operate the modem's circuitry. The transmitter section contains a modulator as well as filtering, wave-shaping, and signal control circuitry that converts digital pulses (often input as a direct-current signal with one level representing a digital one and another level a digital zero) into analog, wave-shaped signals that can be transmitted over a telephone circuit. The receiver section contains a demodulator and associated circuitry that is used to reverse the modulation process by converting the received analog signals back into a series of digital pulses (**Fig. 1**). *See* DEMODULATOR;

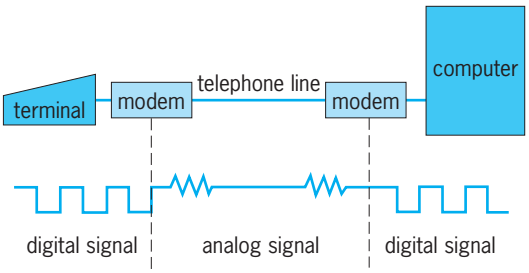

Fig. 1. Signal conversion performed by modems. A modem converts a digital signal to an analog tone (modulation) and reconverts the analog tone into its original digital signal (demodulation).

ELECTRIC FILTER; ELECTRONIC POWER SUPPLY; MODULATOR; WAVE-SHAPING CIRCUITS.

Handshaking. When a modem initiates a transmission sequence, it generates a tone known as a carrier signal. The carrier signal is simply a continuously repeating signal, which by itself conveys no information. However, it serves to inform the device at the opposite end of the telephone circuit that a connection between two modems has occurred. The modem that receives the carrier tone then generates its own carrier (at a different frequency from that of the originating device), which, when recognized by the original sending modem, indicates that a connection in both directions has been established. This process, commonly referred to as handshaking, is for full-duplex modems that are capable of transmitting and receiving data simultaneously. Half-duplex modems that can only send or only receive data at any particular instant switch their carrier signal off and on to reverse the direction of data flow. *See* CARRIER (COMMUNICATIONS).

Modulation process. The modulation process alters the characteristics of the repeating carrier signal and results in the impression of information onto the carrier. Normally, the carrier is a sine wave. The carrier's characteristics that can be altered include its amplitude for amplitude modulation (AM), its frequency for frequency modulation (FM), and its phase angle for phase modulation (PM or ϕM). In addition, a combination of amplitude and phase modulation is commonly employed to transmit high-speed data. *See* AMPLITUDE MODULATION; FREQUENCY MODULATION; PHASE MODULATION.

Amplitude modulation. Amplitude modulation involves varying the magnitude of the carrier signal from a zero level to represent a binary zero to a fixed peak-to-peak voltage level to represent a binary one (**Fig. 2***a*). Amplitude modulation by itself is normally used for very low data rates. However, it is also employed in conjunction with phase modulation to obtain a method of modulating high-speed digital data source.

Frequency modulation. Frequency modulation refers to how frequently a signal repeats itself at a given amplitude. In frequency-shift keying (FSK), the transmitter shifts from one frequency to another as the input digital data changes from a binary one to a binary zero or from a zero to a one. This technique is used primarily by low-speed modems operating at data rates up to 300 bits per second in a full-duplex mode and up to 1200 bits per second in a half-duplex mode (Fig. 2*b*).

Phase modulation. Phase is the position of a waveform of a signal with respect to the origination of the carrier cycle. Thus, phase modulation is the process of varying the carrier signal with respect to the origination of its cycle (Fig. 2*c*). Several forms of phase modulation are used in modems, including single- and multiple-bit phase-shift keying (PSK) and the combination of amplitude and multiple-bit phase-shift keying.

In single-bit phase-shift keying, the transmitter simply shifts the phase of the signal to represent each bit entering the modem. Thus, a binary one might be represented by a 90° phase change, while a zero bit could be represented by a 270° phase change; this technique is known as two-phase modulation.

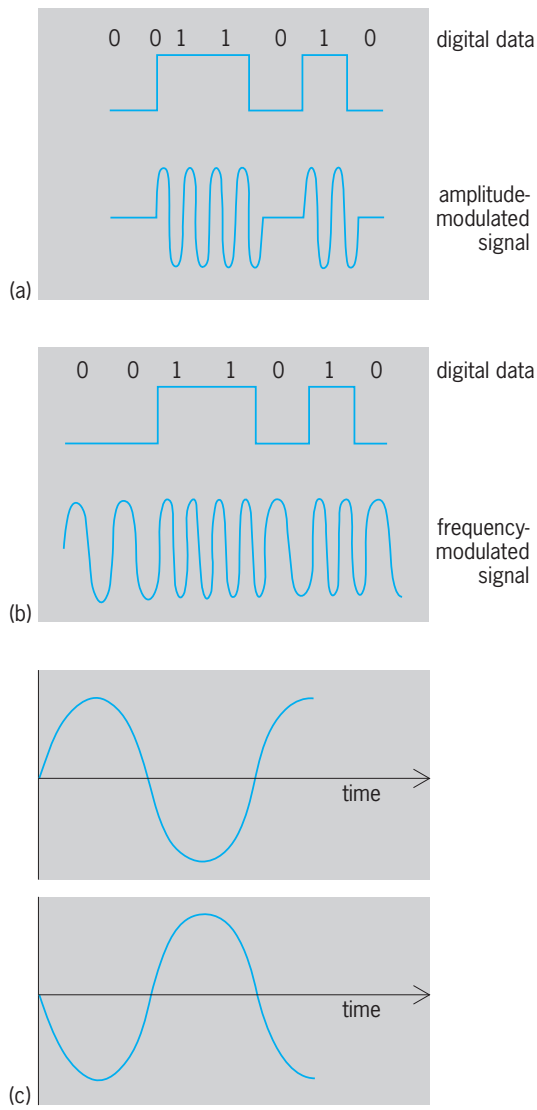

Fig. 2. Use of modulation to encode a digital data stream into analog signals. (*a*) Amplitude modulation. (*b*) Frequency-shift keying, a type of frequency modulation. (*c*) Phase modulation. The bottom wave is 180° out of phase with a normal sine wave illustrated at the top.

In multiple-bit phase-shift keying, two or more bits are grouped together and represented by one phase shift in a signal. Most modems that operate between 600 and 4800 bits per second use multiple-bit phase-shift keying modulation because this technique permits more information to be conveyed in each signal change than single-bit phase-shift keying.

Combined modulation techniques. One of the most frequently used combined modulation methods involves both amplitude and phase modulation. This technique, known as quadrature amplitude modulation (QAM), enables four bits to be represented by one amplitude and phase change. For example, if the modem varies its signal 2400 times a second and each signal represents four bits, then the modem has a data rate of 9600 bits per second.

The first implementation of QAM involved a combination of phase and amplitude modulation, in which 12 values of phase and 3 values of amplitude were employed to produce 16 possible signal states (**Fig. 3**). Most 9600 bit-per-second modems adhere to the Consultative Committee for International Telephone and Telegraph (CCITT) V.29 standard. The V.29 modem uses a 1700-Hz carrier that is varied in both phase and amplitude, resulting in 16 combinations of 8 phase angles and 4 amplitudes.

High-speed modems. The key to the ability to transfer information at high data rates (up to 115.2 kilobits per second) resulted from the use of echo cancellation, trellis coding, and data compression.

Echo cancellation. Echo cancellation permits a modem to operate in a full-duplex transmission mode on a two-line circuit, such as the switched telephone network. Under echo cancellation a modem establishes two high-speed channels in opposite directions. Through the use of digital-signal-processing circuitry, the modem's receiver uses the shape of the modem's transmitter signal to cancel out the effect of its own transmitted signal, enabling the modem to distinguish its sending signal from the signal being received.

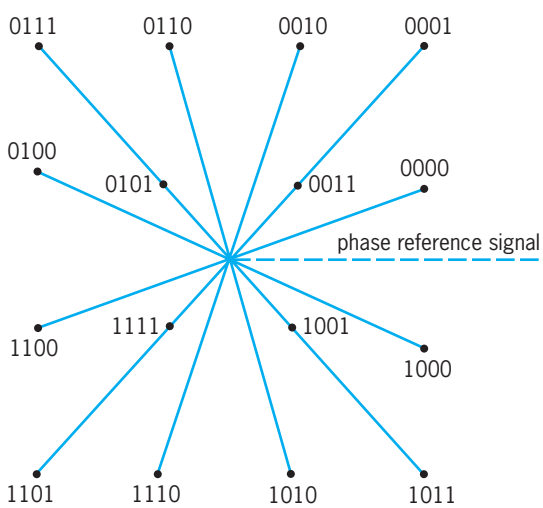

Fig. 3. Early version of quadrature amplitude modulation that produced 16 signal states from a combination of 12 phases (represented by angles of lines with respect to the phase reference signal) and 3 amplitude levels (represented by radial distances of points from the center).

The first modem to use echo cancellation to obtain full-duplex transmission on the switched telephone network at 9600 bits per second was the V.32 device. That modem was followed by the V.32bis (bis meaning the most updated version of the specification), which can operate at up to 14,400 bits per second, and the V.34 modem, which operates at data rates up to 28,800 bits per second. The key to the ability of each of those modems to transfer data at high data rates with extremely low error rates is their use of trellis coding.

Trellis coding. Trellis coding provides a higher tolerance to noise and other transmission impairments. When conventional QAM modems are used, a line impairment displaces the received signal from its appropriate location in the signal constellation. The modem's receiver is designed to select the signal point in the constellation closest to what it received. Hence, an error occurs when line impairments are large enough to cause the received point to be closer to a signal point that is different from the one transmitted.

Under trellis coding the modem's transmitter converts the serial data stream into groups of bits, adding one or more redundant bits to each group based upon the type of trellis coding employed at a particular modem operating rate. For example, a V.32bis modem will convert the serial data stream to be transmitted into 6-bit symbol groups and encode 2 of the 6 bits by using a binary convolutional encoding scheme which adds a code bit to the 2 selected input bits. *See* INFORMATION THEORY.

The redundancy introduced by the encoder means that only certain sequences of signal points are valid. Thus, if an impairment causes a signal point to be shifted, the receiver will compare the observed point to all valid points and select the valid signal point closest to the observed signal. As a result, a modem using trellis coding is only half as susceptible as a conventional QAM modem to an equal amount of noise power, and the resulting error rate is approximately two to three orders of magnitude below that of a QAM modem.

The V.34 modem employs a very complex form of trellis coding, using a 64-state, four-dimensional scheme to make more efficient use of the constellation space, while achieving an error rate below that of conventional 9600-bit-per-second QAM modems.

Data compression. Data-compression technology gained widespread acceptance with the promulgation of the V.42bis standard which defines the use of a modified version of the Lempel-Ziv method. Lempel-Ziv compression is a string-based compression technique in which strings are stored in a dictionary and numeric codes assigned to each string are transmitted in place of the string. Under the V.42bis standard, the algorithm used to develop a dictionary of strings, the number of entries in the dictionary, and its dynamic update are defined.

On the average, a fourfold reduction in data can be expected through the use of V.42bis compression.

This means, for example, that a V.34 modem with V.42bis compression, which transmits data at 28,800 bits per second, can achieve a data throughput of 28,800 × 4 or 115.2 kilobits per second. *See* DATA COMMUNICATIONS; DATA COMPRESSION; ELECTRICAL COMMUNICATIONS. Gilbert Held

Bibliography. G. Held, *The Complete Modem Reference*, 3d ed., 1996; G. Held, *Data Communications Networking Devices*, 4th ed., 1998; G. Held, *Data and Image Compression: Tools and Techniques*, 4th ed., 1996; G. Held and R. Sarch, *Data Communications: A Comprehensive Approach*, 3d ed., 1995.

Modulation

A technique employed in telecommunications transmission systems whereby an electromagnetic signal (the modulating signal) is encoded into one or more of the characteristics of another signal (the carrier signal) to produce a third signal (the modulated signal), whose properties are matched to the characteristics of the medium over which it is to be transmitted. The encoding preserves the original modulating signal in that it can be recovered from the modulated signal at the receiver by the process of demodulation. The main purpose of modulation is to overcome any inherent incompatibilities between the electromagnetic properties of the modulating signal and those of the transmission medium. Of primary importance in this respect is the spectral distribution of power in the modulating signal relative to the passband of the medium. Modulation provides the means for shifting the power of the modulating signal to a part of the frequency spectrum where the medium's transmission characteristics, such as its attenuation, interference, and noise level, are favorable. *See* ELECTROMAGNETIC WAVE TRANSMISSION; RADIO-WAVE PROPAGATION.

Two forms of modulation are generally distinguished, although they have many properties in common: If the modulating signal's amplitude varies continuously with time, it is said to be an analog signal and the modulation is referred to as analog. In the case where the modulating signal may vary its amplitude only between a finite number of values and the change may occur only at discrete moments in time, the modulating signal is said to be a digital signal and the modulation is referred to as digital.

In most applications of modulation the carrier signal is a sine wave, which is completely characterized by its amplitude, its frequency, and its phase relative to some point in time. Modulating the carrier then amounts to varying one or more of these parameters in direct proportion to the amplitude of the modulating signal. In analog modulation systems, varying the amplitude, frequency, or phase of the carrier signal results in amplitude modulation (AM), frequency modulation (FM), or phase modulation (PM), respectively. Since the frequency of a sine wave expressed in radians per second equals the derivative of its phase, frequency modulation and phase modulation are sometimes subsumed under the general term "angle modulation" or "exponential modulation." The **illustration** shows an example of an unmodulated sine-wave carrier signal and the signal resulting from modulating its amplitude, phase, or frequency with the amplitude of an analog modulating signal, which is also taken to be a sine wave.

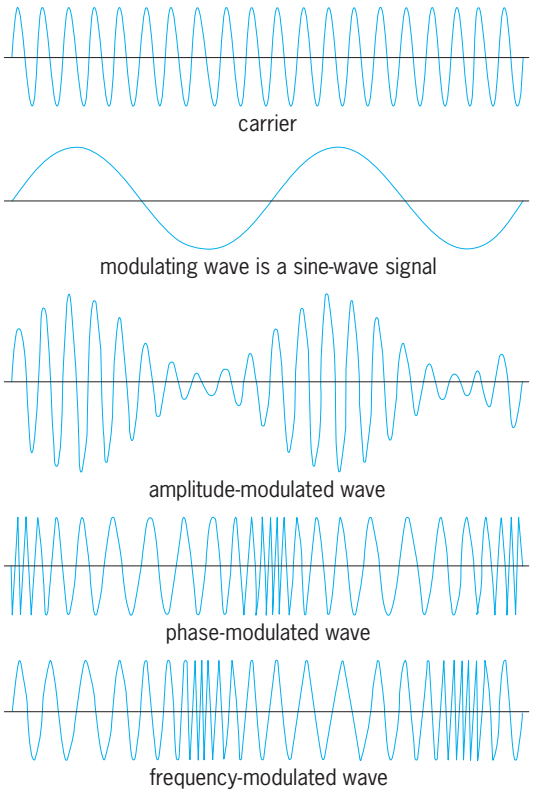

Amplitude, phase, and frequency modulation of a sine-wave carrier by a sine-wave signal. (*After H. S. Black, Modulation Theory, Van Nostrand, 1953*)

Digital modulation. If the modulating signal is digital, the modulation is termed amplitude-shift keying (ASK), frequency-shift keying (FSK), or phase-shift keying (PSK), since in this case the discrete amplitudes of the digital signal can be said to shift the parameter of the carrier signal between a finite number of values. For a modulating signal with only two amplitudes, "binary" is sometimes added before these terms. In a particular form of binary ASK, known as on-off keying (OOK), the carrier signal is turned on or off in accordance with the two amplitudes of the modulating signal.

Digital modulating signals with more than two amplitudes are sometimes encoded into both the amplitude and phase of the carrier signal. For example, if the amplitude of the modulating signal can vary between four different values, each such value can be encoded as a combination of one of two amplitudes and one of two phases of the carrier signal.

Pulse modulation. In certain applications of modulation the carrier signal, rather than being a sine wave, consists of a sequence of electromagnetic pulses of constant amplitude and time duration, which occur at regular points in time. Changing one

or the other of these parameters gives rise to three modulation schemes known as pulse-position modulation (PPM), pulse-duration modulation (PDM), and pulse-amplitude modulation (PAM), in which the time of occurrence of a pulse relative to its nominal occurrence, the time duration of a pulse, or its amplitude are determined by the amplitude of the modulating signal. PAM and PPM are thus in many respects analogous to amplitude modulation and phase modulation. Another form of this type of modulation is pulse-frequency modulation (PFM), in which the rate at which the pulses are transmitted is varied in accordance with the modulating signal. PFM is analogous to frequency modulation of a carrier signal, with its instantaneous frequency being a continuous function of the modulating signal. *See* PULSE MODULATION.

Amplitude modulation. In this form of modulation, the amplitude of the sinusoidal carrier signal is varied in direct proportion to the amplitude of the modulating signal, with the frequency of the carrier signal held constant. If the modulating signal has a bandwidth of B hertz, so that the power in the modulating signal is distributed over the frequency range 0 to B hertz, the spectrum of the resulting modulated signal occupies the range from $f + B$ hertz to $f - B$ hertz, where f is the frequency of the carrier signal. The modulating signal power has therefore been shifted to a part of the spectrum determined by the frequency of the carrier signal, and the bandwidth of the modulated signal is twice the bandwidth of the modulating signal. The part of the spectrum from f to $f + B$ hertz is known as the upper sideband, and the part from f to $f - B$ hertz is known as the lower sideband. Either sideband contains all the information required to recover the original modulating signal. *See* AMPLITUDE MODULATOR; SIDEBAND.

Three special forms of amplitude modulation are frequently employed. In double-sideband, suppressed-carrier modulation (DSB-SC), the power contained in the unmodulated carrier is eliminated prior to transmission of the modulated signal. In single-sideband modulation (SSB), only one of the sidebands is transmitted. In vestigial-sideband modulation (VSB), one complete sideband and a fraction of the other sideband are transmitted. *See* SINGLE SIDEBAND.

Demodulation of an AM signal, which requires the translation of the spectrum of the modulating signal to its original position, can take two different forms. Coherent demodulation, also known as synchronous detection, involves the multiplication of the modulated signal with a precise copy of the carrier signal and subsequent removal of any high-frequency power by filtering. Noncoherent demodulation, also known as envelope detection, involves a nonlinear operation such as the squaring of the modulated signal and subsequent low-pass filtering. *See* AMPLITUDE MODULATION; AMPLITUDE-MODULATION DETECTOR.

Angle modulation. Angle modulation rests on the definition of a generalized sinusoidal signal $s(t) = A \cos \theta(t)$, where $\theta(t)$ is its generalized angle, and the time derivative $\theta'(t) = \omega(t)$ is its instantaneous angular frequency. Two forms of angle modulation can be considered.

In FM, the instantaneous angular frequency $\omega(t)$ is varied in direct proportion to the modulating signal $m(t)$, so that $\omega(t) = \omega_c + km(t)$. Here ω_c is the angular frequency of the unmodulated carrier and k is a constant. If m_p denotes the peak value of $m(t)$, then the maximum deviation of $\Delta\omega$ of ω_c produced by $m(t)$ is given by $\Delta\omega = km_p$.

In PM, the modulating signal $m(t)$ is used to vary the instantaneous phase of the carrier so that $\theta(t) = \omega_c t + km(t)$. Since in this case $\omega(t) = \theta'(t) = \omega_c + km'(t)$, it can be said that modulating the phase of a carrier with $m(t)$ is equivalent to modulating the frequency of the carrier with the time derivative of $m(t)$. The maximum frequency deviation is then $\Delta\omega = km'_p$, where m'_p is the maximum value of $m'(t)$.

For both FM and PM, the modulated signal $s(t)$ has constant amplitude A and constant average power of $A^2/2$, independant of $m(t)$. Its bandwidth W is in theory infinite but for practical purposes is usually calculated from Carson's rule, according to which $W = 2(\Delta f + B)$. Here B is the bandwidth of $m(t)$ or $m'(t)$, and $\Delta f = \Delta\omega/(2\pi)$ is the maximum change in the frequency of the carrier signal produced by $m(t)$ in the case of FM or by $m'(t)$ in the case of PM. Equivalently, by defining $\beta = \Delta f/B$ as the deviation ratio, Carson's Rule can be stated as $W = 2B(\beta + 1)$.

Two forms o FM and PM are normally considered: If β is very much less than 1, the modulation is said to be narrow-band and the bandwidth of $s(t)$ equals $2B$, independant of the peak value of $m(t)$ or $m'(t)$. If β is very much larger than 1, the modulation is referred to as wide-band. The bandwidth of $s(t)$ is then equal to $2\Delta f$. For FM the value of Δf is determined by the peak value of $m'(t)$. Since the latter depends strongly on the spectral distribution of power in $m(t)$, the bandwidth of wide-band PM differs significantly from that of wide-band FM in that it depends on both the amplitude and spectral distribution of $m(t)$.

FM demodulation involves the generation of a signal whose amplitude is linearly proportional to the frequency of the modulated signal. Devices that produce such a response are known as discriminators. Among the most prominent are slope detectors, balanced discriminators, especially the Foster-Seeley discriminator, zero-crossing detectors, and phase-lock loops. *See* FREQUENCY-MODULATION DETECTOR.

Because of the intimate relationship between phase and frequency modulation, any of the FM demodulators can also be employed for the demodulation of a phase-modulated signal, with the additional step of integrating the output of the frequency demodulator. *See* FREQUENCY MODULATION; PHASE MODULATION; PHASE-MODULATION DETECTOR.

Quadrature amplitude modulation (QAM). This is an example of a multilevel digital modulation technique in which the distinct amplitudes of a digital modulating signal are encoded into combinations of phases and amplitudes of the carrier signal. A popular form of this scheme is 16-QAM, in which the carrier signal assumes 12 phases 30° apart, and three amplitudes. Carrier signals with phases of 45°, 135°, 225°, or 315°

assume two of the amplitudes, and the remaining eight phases are assigned the third amplitude. The resulting 16 states of the carrier signal are related to 16 distinct amplitudes of the modulating signal, each of which may represent a group of four binary digits. The bandwidth of any QAM signal is always twice the bandwidth of the modulating signal. Demodulation of a QAM signal involves the measurement at the receiver of both the phase and amplitude of the received signal.

Optical modulation. As in the modulation of electrical carrier signals, the amplitude, frequency, or phase of an optical sinusoidal carrier signal, typically generated by a light-emitting diode (LED) or a semiconductor laser, can be changed in accordance with the amplitude of the modulating signal. The modulating signal can also be used to change the carrier signal's polarization. For binary digital modulating signals the most common form of modulation employed is on-off keying, in which the carrier signal is shifted in power between two values. The ratio of these power levels is known as the extinction ratio. *See* LASER; LIGHT-EMITTING DIODE; POLARIZATION OF WAVES.

In so-called direct modulation schemes the injection current of the laser, which determines the frequency and power of the output signal, is varied directly in accordance with the modulating signal. This method may result in undesirable frequency shifts and other types of signal distortion. Much of this can be avoided through the use of external modulation schemes. Here the amplitude of the laser output is kept at a constant level, but is subsequently varied by inserting an attenuator between the laser and the fiber-optic transmission system whose level of attenuation is controlled by the modulating signal amplitude. External modulators are also used in effecting phase or frequency changes of the carrier signal for phase-modulation and frequency-modulation systems.

Several different methods for the demodulation of optical signals have been implemented. In direct-detection receivers, mostly used in conjunction with on-off keying, a photodetector converts received light energy into an electrical current or voltage, which is then compared to a threshold. Signal values above or below the threshold are decoded as one or the other amplitude of the original modulating signal. Several types of semiconductor photodiodes have been employed as photodetectors. In coherent receivers the received light, together with a monochromatic light signal generated by a local light source, is applied to the photodetector. The latter produces an output signal whose frequency is the difference in the frequency of the two input signals. Comparing this signal to a threshold completes the demodulation process. *See* OPTICAL COMMUNICATIONS; OPTICAL FIBERS; PHOTODIODE.

Performance of modulation systems. The quality of an analog modulation-demodulation system is normally determined by the ratio of signal power to noise power at the input to the demodulator and is measured by the same ratio at the output of the demodulator. For binary digital modulation systems the measure of quality is the bit error rate, the number of incorrect binary digits relative to the total number of binary digits at the output of the demodulator. In general, the performance of a modulation system depends critically on the amount of transmitted power, the bandwidth, the level of attenuation and signal distortion of the channel, the amount of interference present in the received signal, and the method of demodulation. *See* ELECTRICAL COMMUNICATIONS.

Hermann J. Helgert

Bibliography. P. E. Green, Jr., *Fiber Optic Networks*, Prentice Hall, 1993; F. Halsall, *Data Communications, Computer Networks and Open Systems*, 4th ed., Addison-Wesley, 1996; B. P. Lathi, *Modern Digital and Analog Communication Systems*, 3d ed., Oxford University Press, 1998; M. Schwartz, *Information Transmission, Modulation, and Noise*, 4th ed., McGraw-Hill, 1990.

Modulator

A device that combines an electrical information signal with a periodic electrical carrier signal for efficient transmission. The modification of a usually sinusoidal electrical carrier signal (or simply carrier) for the purpose of transmitting information which is carried on a second electrical modulating signal or information signal is called modulation. The device or circuit that performs the modulation is called a modulator. The information may be from any of a variety of sources such as sensor data, speech signals, or image data. The modulator varies the carrier in amplitude, frequency, phase, or some combination in order to imbed the information for efficient transmission. During the process of modulation, information is shifted in frequency from baseband to the carrier frequency for efficient transmission. *See* CARRIER (COMMUNICATIONS).

The primary classes of modulation are amplitude modulation (AM), frequency modulation (FM), and phase modulation (PM). Frequency and phase modulation are highly amplitude-noise resistant because the amplitudes of the received signal are forced to a constant value.

The information signals usually are of analog origin but may be converted to digital format for more exact processing. Transmission is always analog or continuous. Robust and versatile digital signal modulation of any type may be implemented directly with microprocessors or digital components (such as multipliers and adders) by mechanizing the modulation algorithms directly. The precisely controlled modulated digital signal is transformed into an analog signal by a digital-to-analog converter (DAC). Finally, by means of a linear amplifier, its power is increased to a level required for successful radio-frequency (RF) transmission. Because linear RF amplifiers are notoriously inefficient, an alternative lower-cost approach to AM and quadrature amplitude modulation (QAM) is used to first generate the baseband analog signals and then impress them on the RF carriers as described below.

High-quality data resolution is sacrificed in exchange for economy. *See* DIGITAL-TO-ANALOG CONVERTER.

Amplitude modulator. Amplitude modulation is subdivided into the classes double-sideband transmitted carrier (variously called DSTC, AMTC, or "standard" AM); the more complex but power-efficient double-sideband suppressed carrier (DSSC); the even more power- and bandwidth-efficient (and complex) single-sideband suppressed carrier (SSBSC or simply SSB); vestigial sideband (VSB), in which one sideband is only partially removed, leaving only a vestige of itself; and QAM, in which two AM signals are transmitted simultaneously on orthogonal carriers (that is, 90° out of phase with one another), which consequentially stay out of one another's way while sharing a common spectrum. *See* AMPLITUDE MODULATION; SINGLE SIDEBAND.

A sinusoidal carrier signal is given by Eq. (1),

$$c(t) = \cos(2\pi f_c t) \tag{1}$$

where f_c is the carrier frequency in hertz. It is useful to consider the special case where this signal is amplitude-modulated by a sinusoidal information signal given by Eq. (2), where $A \geq 0$ and $B \geq 0$. If $A \geq$

$$s(t) = A + B \cdot \cos(2\pi f_i t) \tag{2}$$

B, the information may be recovered by simple envelope detection. Amplitude modulation is a multiplication process, where the transmitted signal is given by Eq. (3). The arguments of the trigonometric func-

$$\begin{aligned} T(t) = c(t) \cdot s(t) &= A \cdot \cos(2\pi f_c t) \\ &+ B \cdot \cos(2\pi f_i t) \cdot \cos(2\pi f_c t) = A \cdot \cos(2\pi f_c t) \\ &+ {}^1\!/_2 B \cdot \cos[2\pi(f_c + f_i)t] \\ &+ {}^1\!/_2 B \cdot \cos[2\pi(f_c - f_i)t] \end{aligned} \tag{3}$$

tions show that a standard AM or DSTC modulated signal has transmitted components at frequencies f_c, $f_c + f_i$, and $f_c - f_i$ with respective amplitudes of A, ${}^1\!/_2 B$, and ${}^1\!/_2 B$ (**Fig. 1**). One can show that this conclusion is also valid for general bandwidth-limited information signals. *See* DEMODULATOR.

If $A = B$, half of the transmitted power is in the carrier, which contains no information. If we set $A = 0$, the carrier is suppressed, transmitted power is reduced by 50% compared to DSTC, and a DSSC signal results. If the maximum information frequency contained in $s(t)$ is f_{max}, then the bandwidth of the transmitted double-sideband signals is $2f_{max}$. Transmitting power may be halved again, and the bandwidth halved as well, by use of SSB, that is, by sending only one sideband after removing the other by linear band-pass filtering or by actively canceling the unwanted sideband. A little signal manipulation yields Eq. (4). By choosing the plus or minus sign, we can

$$\cos(2\pi f_c t) \cdot \cos(2\pi f_i t) -/+ \sin(2\pi f_c t) \cdot \sin(2\pi f_i t) = \cos[2\pi(f_c +/- f_i)t] \tag{4}$$

elect to generate and to transmit either the upper (+) or lower (−) sideband (**Fig. 2**). *See* ELECTRIC FILTER.

The multiplication function called for in simple and inexpensive analog amplitude modulation is achieved economically by varying, via the modulating signal, the applied power-supply voltage to the output RF amplifier that follows the carrier-frequency oscillator. This can be done by simply transformer-coupling the modulating signal voltage into the amplifier's power-supply output line. *See* AMPLITUDE MODULATOR.

Quadrature amplitude modulator. A simple quadrature amplitude modulator can be implemented by inserting a second signal path from the reference carrier-frequency oscillator output consisting of a 90° phase shifter followed by a second linear amplifier whose power-supply output line is varied by a second transformer-coupled modulating signal voltage. A QAM signal is formed by passively summing the outputs of the two amplifiers delivering amplitude-modulated sine and cosine (quadrature) carriers.

By using analog or digital signals to amplitude-modulate each quadrature carrier, a QAM transmitter can generate an information-bearing carrier varying in both amplitude and phase. For digital signals, a so-called M-ary (that is, each data symbol can be represented by one of M possible states in the complex plane) digital QAM signal occupies the same bandwidth as a PSK signal (discussed below), but the power efficiency is greater. An M-ary digital QAM signal is generated by segmenting the input-signal data into M-bit lengths. These are divided into two sets, A and B, of $M/2$ bits each. A common segmentation approach is to steer the odd-numbered bits of the input-signal data into the A set and the even-numbered bits into the B set. This pair of $M/2$-bit words is modified by complementing the most significant bit of each word and appending a "1" at the end to provide a pair of $M/2 + 1$-bit words that will be treated as two's complement signed numbers, passed through a pair of $M/2 + 1$-bit digital-to-analog converters, and then multiplied (modulated) by $\cos(2\pi f_c t)$ for the A set and by $\sin(2\pi f_c t)$ for the B set. These two sets of analog information are then summed for transmission. *See* NUMBERING SYSTEMS.

This seemingly arcane exercise provides a rectangular M-by-M array of evenly spaced points in the complex plane, centered at the origin, that represent the M^2 possible data states. The case $M = 2$ is, of course, binary. The case $M = 4$ is robust and commonly used. As M increases, both the channel capacity and the sensitivity to noise similarly increase.

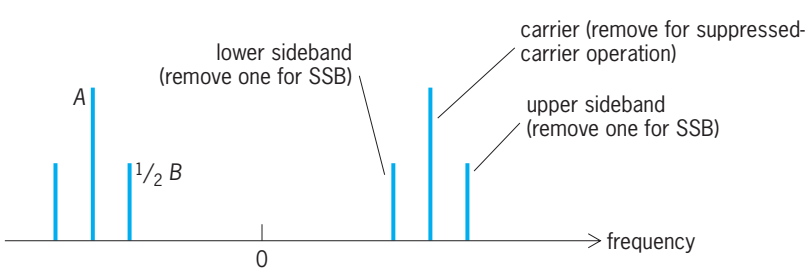

Fig. 1. Amplitude modulation (AM) spectrum.

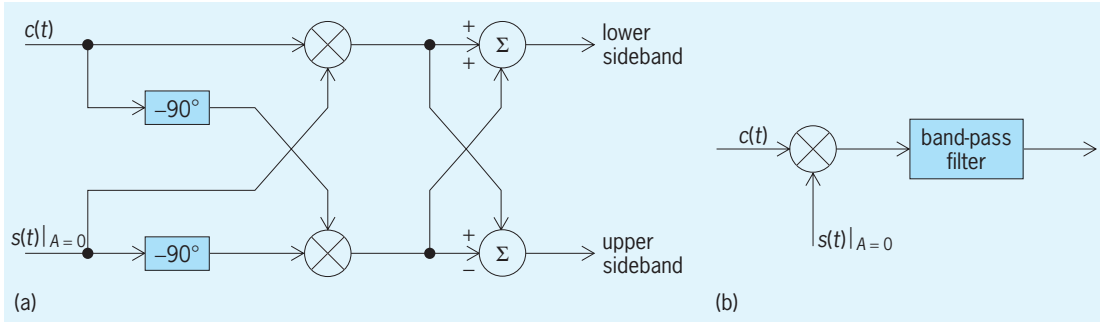

Fig. 2. Single-sideband generation by (*a*) phase shifting and (*b*) band-pass filtering.

Frequency modulator. Amplitude-modulated signals are degraded by additive noise, usually encountered along their propagation paths. This problem is overcome by transmitting information on constant-amplitude carriers and using frequency or phase modulation. The received signals are clipped to a constant amplitude, thereby removing the additive noise.

In frequency modulation, the constant-amplitude signal to be transmitted, given by Eq. (5), is varied in

$$T = \cos\left(2\pi \int_0^1 f_i(\tau)d\tau\right) \tag{5}$$

frequency. The instantaneous frequency is f_i where $f_i = f_c + k_{vf} \cdot s(t)$, f_c is the carrier frequency, $s(t)$ is the information signal, and k_{vf} is the voltage-to-frequency modulation index. If $|s(t)| \leq 1$, k_{vf} is also the peak deviation of the instantaneous frequency from f_c. Equation (5) demonstrates that a time-varying frequency must be integrated with respect to time, not just multiplied by time, in order to obtain phase. The Carson bandwidth (CBW), an approximation to the bandwidth within which virtually all energy lies, is CBW $= 2(k_{vf} + f_{\max})$, where $f_{\max}$ is the highest frequency present in $s(t)$. *See* FREQUENCY MODULATION.

In a frequency modulator, the operating frequency of the carrier-frequency oscillator can be determined by the reactance of a tuned *LC* (inductor-capacitor) circuit. The oscillator frequency is determined by $1/\sqrt{LC}$, so varying either element in the oscillator's tuning circuit varies the frequency. A voltage-controlled capacitor called a varactor can be used in the circuit. The modulating voltage is applied directly to the varactor in order to produce an FM signal. *See* VARACTOR.

Transmission of digital data by frequency modulation is called frequency-shift keying (FSK). The frequency of the transmitted signal is determined by the digitized information signal. *M*-ary FSK employs *M* distinct frequencies. *See* FREQUENCY MODULATOR.

Phase modulator. In phase modulation, the signal to be transmitted is varied in phase according to the information signal, as given by Eq. (6), where k_{vp}

$$T(t) = \cos[2\pi f_c t + k_{vp} \cdot s(t)] \tag{6}$$

is a voltage-to-phase scaling constant. This can be expanded as Eq. (7). If $|s(t)| \leq 1$, k_{vp} is the maximum phase excursion due to the modulating signal. If k_{vp} is small, we have narrow-band phase modulation and the transmitted signal is approximated by Eq. (8), which can be mechanized (**Fig. 3**). Again,

$$T(t) = \cos(2\pi f_c t) \cdot \cos[k_{vp} \cdot s(t)] - \sin(2\pi f_c t) \cdot \sin[k_{vp} \cdot s(t)] \tag{7}$$

$$T(t) = \cos(2\pi f_c t) - k_{vp} \cdot s(t) \cdot \sin(2\pi f_c) \tag{8}$$

the CBW $= 2(k_{pf} + f_{\max})$. *See* PHASE MODULATION.

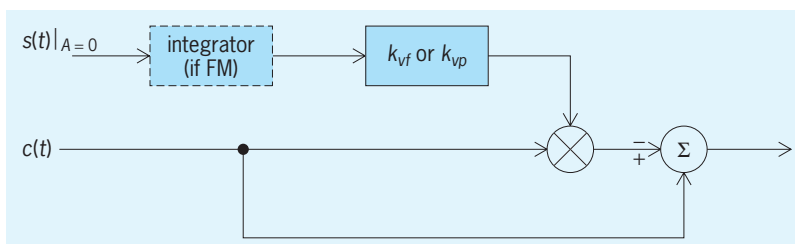

Fig. 3. Narrow-band frequency or phase modulator.

Transmission of digital data by phase-angle modulation is called phase-shift keying (PSK). PSK employs a single frequency. The phase of the transmitted signal is determined by the digitized information signal. Each data symbol in an *M*-ary PSK system may be represented by a point on a unit circle because the amplitude is held constant. The points are separated in angle (phase) by $360°/M$. *See* MODULATION; PHASE MODULATOR.

Stanley A. White

Bibliography. G. A. Breed (ed.), *Wireless Communications Handbook*, Cardiff Publishing, Englewood, CO, 1992; S. Gibilisco, *Handbook of Radio and Wireless Technology*, McGraw-Hill, New York, 1999; J. G. Proakis, *Digital Communications*, 4th ed., McGraw-Hill, New York, 2001; T. S. Rappaport, *Wireless Communications*, 2d ed., Prentice Hall, Upper Saddle River, NJ, 2002; M. S. Roden, *Analog and Digital Communications Systems*, 4th ed., Prentice Hall, Englewood Cliffs, NJ, 1995; M. E. Van Valkenburg and W. M. Middleton (eds.), *Reference Data for Engineers: Radio, Electronics, Computer and Communications*, 9th ed., Butterworth-Heinemann, Woburn, MA, 2002.

Mohair

The long, lustrous hair of the Angora goat, which originated in the area around Ankara (Angora), Turkey. This plateau, about 3000 ft (900 m) high, has a dry climate with extremes in temperature. The Angora goat probably originated before Biblical times, as references to the use of goat's hair are found in the Bible. Angora goats were brought into the United States from Turkey in 1849 and taken westward after the Civil War, especially to Texas and California.

Production and value. About half of the world mohair production is produced in the United States, followed by Turkey, South Africa, and Lesotho (Basutoland). Production in other parts of the world is insignificant. Generally, over half of all mohair produced is exported for processing and manufacturing. Principal importers are the European countries, the United Kingdom, and Japan.

Properties and uses. Mohair is a smooth, strong, durable, and resilient fiber. It enhances softness and luster in fabrics. Mohair absorbs dye evenly and brilliantly, retains color well, and permits unusual decorative effects. It is mainly used as an apparel fiber but may be used in upholstery, draperies, wigs, hairpieces, and rugs. Leather produced from the skin is useful for gloves, purses, and novelties.

Mohair fleeces may be classified by the different types of locks. The ringlet type is represented by long, tight locks or curls (**Fig. 1**). These fleeces have good length but may lack density, resulting in dirt penetration. Flat, wavy locks may be associated with light fleeces (**Fig. 2**). The web type resembles both the flat and ringlet (**Fig. 3**). It may be flat near the skin with a ringlet near the tip.

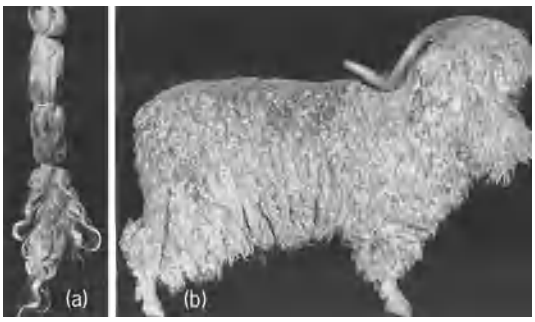

Fig. 1. Ringlet-type Angora: (*a*) lock from (*b*) buck. (*From M. Camp, Texas Sheep and Goat Raisers' Magazine*)

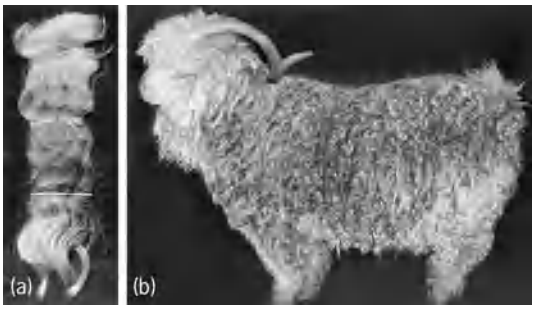

Fig. 2. Flat lock-type Angora: (*a*) lock from (*b*) buck. (*From M. Camp, Texas Sheep and Goat Raisers' Magazine*)

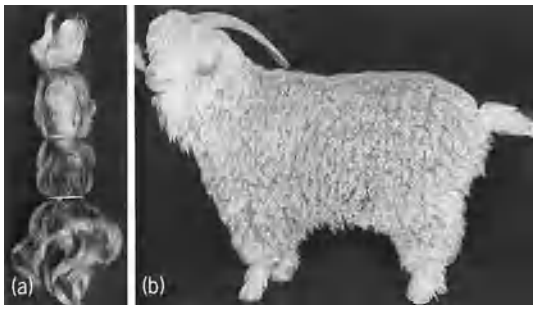

Fig. 3. Intermediate-type Angora: (*a*) lock from (*b*) buck. (*From M. Camp, Texas Sheep and Goat Raisers' Magazine*)

Growth characteristics. Mohair fiber follicles characteristically develop in groups, with connective tissue separating the groups. Three primary or first-appearing follicles are often found in each group. These are associated with an accessory sebaceous gland, a suboriferous gland, and an arrector muscle. Later-developing or secondary follicles generally lack accessory structures, although many have sebaceous glands. The number of secondary fibers per group varies from about 16 to 30. Medullated or hollow fibers are more likely to arise from primary follicles, but generally make up less than 1% of the fibers.

Structure. The mohair fiber is similar to wool in microscopic appearance, with epidermal scales about 20 micrometers long which barely overlap. The mohair fiber has a smooth appearance and is slightly wavy, with less than one crimp per inch. The mohair cortex is made up largely of ortho cells, while wool with its greater numbers of crimps has a cortex of both ortho and para cells. Mohair fibers are circular in cross section and range 10–90 μm in diameter; kid mohair ranges 10–40 μm in diameter. As goats grow older, up to about 8 years, mohair coarsens. Mohair grows at the rate of about 1 in. a month, resulting in a 6-in. (15-cm) staple for a 6-month clip. Large, chalky, brittle fibers known as kemp sometimes occur and have an undesirable effect on spinnability and dyeability. *See* WOOL.

Breeding practices. Angora goats may be improved by selecting for mohair and kid production. Weight of fleece, which is moderately heritable, deserves emphasis in selection. Individual fleece weights, especially for males, should be recorded. Length of staple is important as a quality factor and for its relationship to fleece weight. Fine mohair is preferred. Kemp or colored fibers are undesirable. Fleece density and good covering over the body and belly are preferred. Animals with a tendency to shed should be culled. Size, vigor, and fertility are important in both sexes. Body conformation is important since many goats are sold for meat. Clair E. Terrill

Moho (Mohorovičić discontinuity)

The level in the Earth where the velocity of sonic waves first increases rapidly or discontinuously to a value between 7.6 and 8.6 km/s (4.7 and 5.3 mi/s).

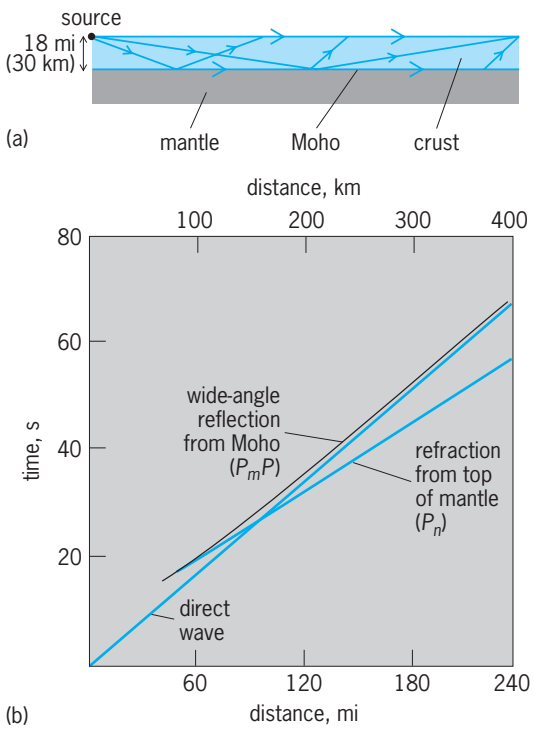

Fig. 1. Characteristics of waves traveling along the Moho. (*a*) Schematic cross section of the continental crust and upper mantle showing typical wave paths from a source of direct waves traveling near the surface of the crust, refracted waves traveling along the Moho (P_n), and wide-angle reflections from the Moho (P_mP). (*b*) Graph showing arrival times for various waves depicted in *a*. (*After M. H. P. Bott, The Interior of the Earth: Its Structure, Constitution and Evolution, 2d ed., Elsevier, 1982*)

A. Mohorovičić discovered this boundary while investigating seismograms of the Zagreb (now capital of Croatia) earthquake of October 8, 1909. He recognized that low-velocity waves traveling directly from the earthquake source were overtaken at large distances by refracted waves traveling through the deeper, high-velocity layer (**Fig. 1**). Modern determinations of the depth and nature of the Moho are commonly made in seismic refraction studies that use artificial seismic sources, such as explosions, rather than earthquakes. This method allows identification of the wave traveling in the high-velocity medium (P_n) and a wide-angle reflection (P_mP) from the boundary (Fig. 1). The Moho is generally assumed to mark the boundary between the crust and mantle, although this need not always be the case. Despite drilling attempts in the early 1960s, the Moho has not been directly sampled. Knowledge of the oceanic and continental Moho is based on interpretation of geophysical data, as well as geological interpretation of ophiolites (believed to be sections of oceanic crust and mantle uplifted onto land). *See* OPHIOLITE.

Formation. The oceanic Moho forms at spreading centers where partially molten material, which has the composition of basalt, both erupts and intrudes to form the oceanic crust, with mafic cumulates and sills forming the lower crust. The higher-density residual ultramafic material cools to form the mantle beneath the crust. Much of the continental crust was formed in the early history of the Earth as low-density, silica-rich materials segregated from the mantle and moved to high levels. Because direct observations of the continental Moho are not available, the mechanics of this segregation process remain enigmatic. However, the continental Moho that formed during early crust formation may not be observable, because the continental Moho is a dynamic feature that can be modified by tectonic and volcanic processes. *See* EARTH CRUST; EARTH INTERIOR.

Continental Moho. Wide-angle and near-vertical reflections from the continental Moho suggest that it is not a first-order discontinuity but a complex zone several kilometers thick composed of interlayered crustal and mantle rock types. A band of strong reflections from the Moho is reported from a few regions, and in others the Moho appears to correspond to a transition from a highly reflective lower crust to a nonreflective upper mantle.

Moho depth, as determined by refraction experiments, varies considerably under the continents as evidenced by the Moho depth (crustal thickness) map of the United States (**Fig. 2**). Worldwide, Moho depths average between 35 and 40 km (22 and 24 mi) and range from 20 to 70 km (12 to 35 mi). These pronounced variations are generally related to tectonic processes. For instance, the Moho is shallow where extension creates thin crust such as in the Basin and Range Province in the western United States. The Moho is deep beneath fold and thrust belts where the crust is thickened by continental collision (for example, the Himalayas and Alps), and under continental magmatic arcs where volcanism and plutonism add new crustal material (for example, the Andes). These observations suggest that the Moho is a dynamic feature that changes its position and characteristics with time in response to tectonism. Certain regions, however, do not fit these generalities. The Moho under the Great Plains, for instance, is deeper than under the adjacent Rocky Mountain region. Significant Moho depth variations in Precambrian shield areas do not correlate with topography but with old tectonic features, indicating that the Moho in these regions represents an unmodified discontinuity established during the Precambrian. *See* GEODYNAMICS; VOLCANOLOGY.

The petrologic nature of the continental Moho is poorly known. Deep continental crust above the Moho is probably composed of a heterogeneous mixture of various metamorphic rocks and igneous rocks with sonic velocities less than 7.5 km/s (4.5 mi/s). Velocities of P_n range from 7.6 to 8.4 km/s (4.7 to 5.0 mi/s), values typical of peridotites, dunites, and eclogites. All these rock types are observed as xenoliths in diatremes, suggesting they are present in the uppermost mantle. In addition to variable upper-mantle composition, significant variation of pressure and temperature at the Moho under different continental regions can cause these lateral variations in velocity. There is evidence that upper-mantle velocities in several regions vary with propagation direction. This anisotropic behavior suggests that, in at least

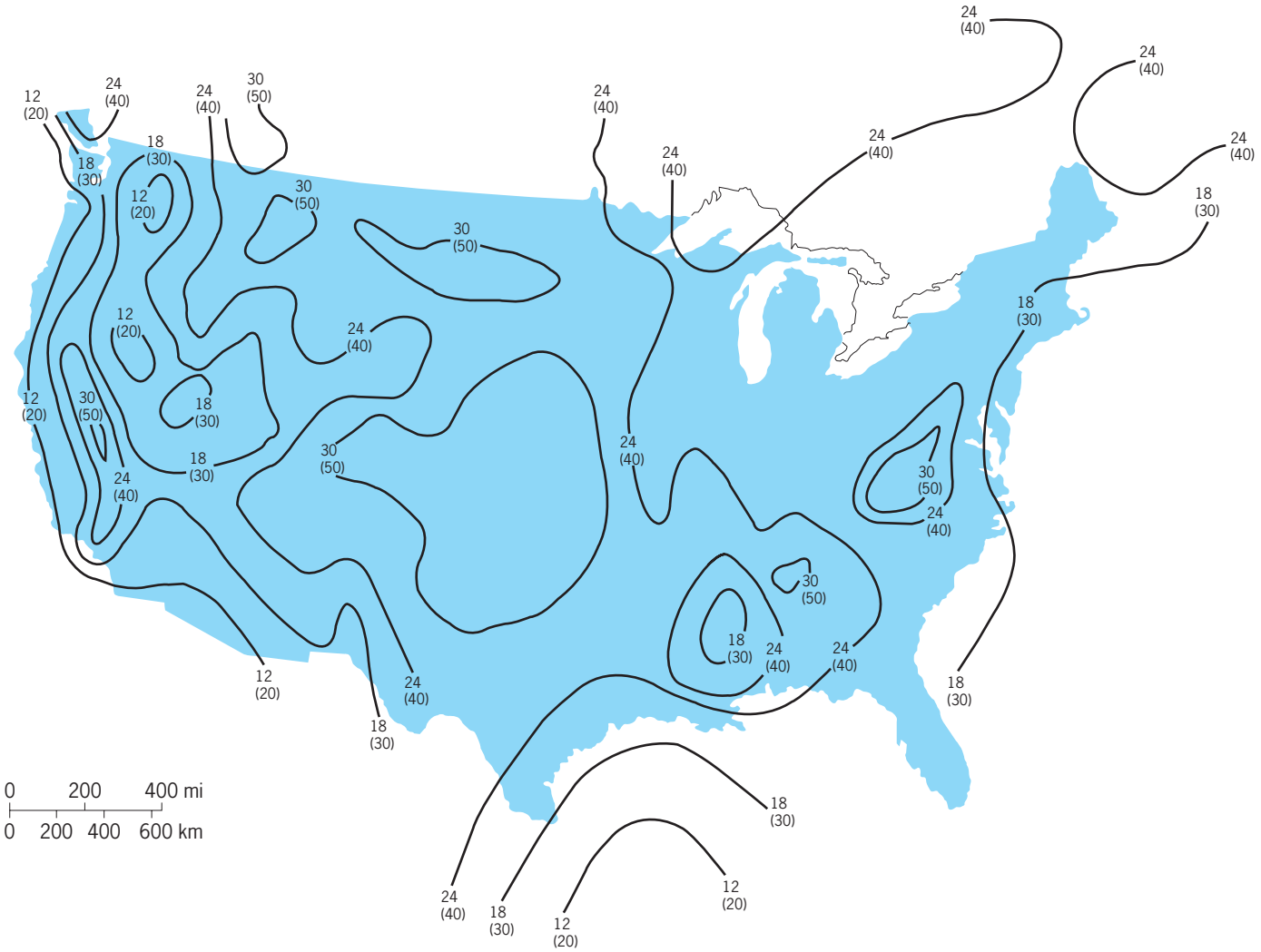

Fig. 2. Contour map of Moho depth in the United States. Contour values are given in miles (kilometers). (*After R. J. Allenby and C. C. Schnetzler, United States Crustal Thickness, Tectonophysics, 93:13–31, 1983*)

some regions, the subcontinental mantle is predominantly composed of peridotites with strong olivine alignment. The continental Moho in some places is probably a compositional boundary separating silica-rich igneous and metamorphic rocks of the crust from denser rock types, such as peridotites, within the mantle.

Results from a Canadian experiment known as LITHOPROBE suggest that four competing hypotheses for the origin of the continental Moho may apply in different regions and may not be mutually exclusive. These are a relict Moho (preserved from an original oceanic Moho); magmatic underplating (from sills); metamorphic front (phase transition); and regional décollment (structural detachment). *See* DUNITE; ECLOGITE; OLIVINE; PERIDOTITE; XENOLITH.

David M. Fountain; Maya Tolstoy

Oceanic Moho. While the definition of Moho in the continents is based solely on seismic refraction observations, the term has been utilized to describe a range of observations pertaining to the oceanic environment. In the strictest sense, Moho refers to the transition boundary where material velocities, as determined from seismic refraction methods, exceed 5 mi/s (8 km/s). Seismic studies conducted throughout the oceans since around 1950 reveal that the arrival of the Moho signal is best modeled as originating from a region (Moho transition zone) of relatively rapid velocity increase from about 4.5 mi/s (7.2 km/s; mafic lower crust) to about 5 mi/s (8 km/s; ultramafic upper mantle). On average, this transition begins at a depth of about 4.0 mi (6.5 km) beneath the top of the oceanic crust (ignoring variable thicknesses of sediments that may also be present), though considerable variability is observed, particularily in the slower-spreading environment. From geological studies of ocean crust and ophiolites (ancient oceanic sections subsequently emplaced on continents), oceanic crust is understood to originate from partial melting of mantle that upwells and decompresses in response to sea-floor spreading. The eruption and intrusion of this melt, and the formation of Moho, occurs in a very narrow zone at the axis of sea-floor spreading. Different mantle conditions can lead to different volumes of melting, accounting for most of the observed variations in crustal thickness (depth to the Moho). Very slow spreading centers, while rare, are expected to form somewhat thinner

crust, largely reflecting the magma-starved nature of these areas. The Moho is a proxy for the transition from crustal materials (for example, basalts, diabase dikes, or gabbros) formed by mafic melts extracted from the mantle to the ultramafic residual mantle (for example, peridotite) that has remained at depth. *See* BASALT; DOLERITE; GABBRO; PETROLOGY; PLATE TECTONICS.

The Moho transition zone can also be imaged by seismic reflection techniques. However, the reflection Moho may originate from the top of a transition zone and thus does not necessarily occur at the same depth as the true (refraction) Moho. Also, seismic reflection imaging is influenced by structural perturbations within the crust. For example, while refraction studies conducted around the relatively fast-spreading East Pacific Rise reveal Moho signal arrivals at a near-constant depth everywhere, reflection studies in the same area are unable to image the Moho in the vicinity of the spreading axis itself because of the presence of a low-velocity zone associated with a small crustal magma chamber perched atop very hot crust. Similarly, seismic reflection studies conducted in the crust formed at the relatively slow-spreading Mid-Atlantic Ridge environment rarely reveal a reflection Moho, even though refraction determinations of depth to Moho, that is, depth to residual mantle, are commonly indistinguishable from those observed in oceanic crust formed elsewhere. The reason is that slower sea-floor spreading is accompanied by a significantly greater degree of extensional faulting within the crust, as revealed by topographic roughness of the young crust, prominent intracrustal reflectors observed in older oceanic crust, and occurrences of hydrated residual mantle (serpentinized peridotite) in slow-spreading environments, particularly near fracture zones. During hydration (serpentinization), the intrinsic velocity of the rock decreases significantly such that the serpentinized peridotite no longer retains a Moho velocity. In such settings, the depth to the reflection Moho can change with time, depending on the depth to which hydration persists. *See* MAGMA.

Results from reflection work provide evidence for frozen magma lenses at and below the Moho supporting earlier refraction, compliance, and ophiolite work suggesting that lenses of gabbroic material at the base of the crust play a significant role in formation of the lower crust. It is thought that the presence or absence of these gabbroic lenses may explain some of the variability in Moho reflectivity within areas of similar topographic relief.

A usage encountered in many texts is petrologic Moho. While quite a departure from the original geophysical derivation of the term Moho, this term has been introduced to distinguish between the depth to the reflection Moho (originating within a Moho transition zone) and the depth to the true residual mantle (beneath any such transition zone); it is often used in descriptions of ophiolites. *See* IGNEOUS ROCKS; MARINE GEOLOGY; MID-OCEANIC RIDGE; SERPENTINE.

Carolyn Z. Mutter; Maya Tolstoy

Bibliography. D. L. Anderson, *Theory of the Earth*, 1989; K. Benn, A. Nicholas, and I. Reuber, Mantle-crust transition zone and origin of wehrlitic magmas: Evidence from the Oman ophiolite, *Tectonophysics*, 151:75–85, 1988; W. E. Bonini and R. R. Bonini, Andrija Mohorovii: Seventy years ago an earthquake shook Zagreb, *EOS Trans. Amer. Geophys. Union*, 60(41):699–701, 1979; M. H. P. Bott, *The Interior of the Earth: Its Structure, Constitution and Evolution*, 2d ed., 1982; W. C. Crawford and S. C. Webb, Variations in the distribution of magma in the lower crust and at the Moho beneath the East Pacific Rise at 9°–10° N., *Earth Planet. Sci. Lett.*, 203:117–130, 2002; D. W. Eaton, Multi-genetic origin of the continental Moho: Insights from LITHOPROBE, *Terra Nova*, 18:34–43, 2006; C. M. Jarchow and G. A. Thompson, The nature of the Mohorovii discontinuity, *Annu. Rev. Earth Planet. Sci.*, 17:475–506, 1989; R. Meissner, T. Weaver, and E. R. Flüh, The Moho in Europe: Implications for crustal development, *Ann. Geophys.*, 5B:357–364, 1987; J. C. Mutter and J. A. Karson, Structural processes at slow spreading ridges, *Science*, 257:627–634, 1992; M. R. Nedimovic et al., Frozen magma lenses below the oceanic crust, *Nature*, 436:1149–1152, 2005; E. E. Vera et al., The structure of 0- to 0.2-m.y.-old oceanic crust at 9°N on the East Pacific Rise from expanded spread profiles, *J. Geophys. Res.*, 95:15529–15556, 1990.

Moiré pattern

When one family of curves is superposed on another family of curves, a new family called the moiré pattern appears. A familiar example of moiré is the pattern one sees on looking through the folds of a nylon curtain. When the curtain is moved slightly, the moiré pattern is observed to move wildly about. This effect raises at once the possibility of using the moiré effect to measure minute displacements. Moreover, the moiré patterns in the curtain are reminiscent of the complex patterns of waves seen at the shore of a lake. This suggests again that moiré might be used to describe in graphical form certain physical phenomena.

To produce moiré patterns, the lines of the overlapping figures must cross at an angle of less than about 45°. The moiré lines are then the locus of points of intersection. **Figure 1** illustrates the case of two identical figures of simple gratings of alternate black and white bars of equal spacing. When the figures are crossed at 90°, a checkerboard pattern with no moiré effect is seen. At crossing angles of less than 45°, however, one sees a moiré pattern of equispaced lines, the moiré fringes. The spacing of the fringes increases with decreasing crossing angle. This provides one with a simple method for measuring extremely small angles (down to 1 second of arc). As the angle of crossing approaches zero, the moiré fringes approach 90° with respect to the original figures.

Even when the spacing of the original figures are far below the resolution of the eye, the moiré fringes will still be readily seen. When two diffraction gratings of the transmitting type are superimposed and

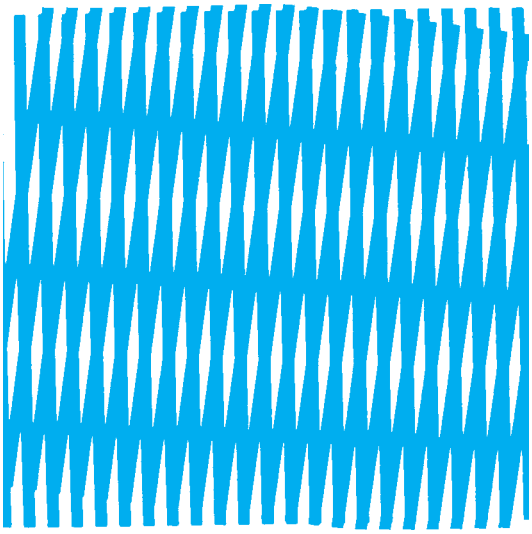

Fig. 1. Two simple gratings crossed at a small angle.

a distant light source viewed through them, moiré fringes can be seen by the unaided eye. This experiment, first performed by Lord Rayleigh in 1874, provides a means of checking the fidelity of a replica of a diffraction grating. If the repeat spacing of one grating differs slightly from that of the other, a beat pattern will be observed when the lines of the two figures are crossed at zero angle. The beat spacing is inversely proportional to the difference in spacings of the two gratings. Thus with nearly identical gratings a slight relative displacement will cause many fringes to pass by. This beat effect is the same as that seen when one is moving past two repetitive structures, such as the railings of a bridge; the separation between the posts of the railing closer to the observer appears to be slightly greater than that of the farther railing. As a consequence, one sees a beat when a post of the nearer railing gets in step (or is in phase) with an open space of the farther railing.

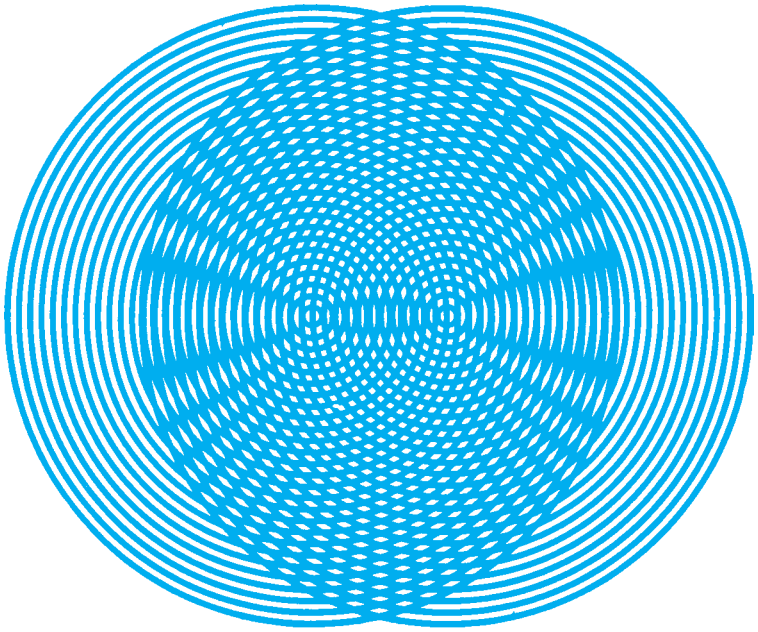

Fig. 2. Superposition of two families of equispaced concentric circles.

The moiré-beat method, using diffraction gratings, has been applied industrially to automotive precision machinery, the fringes being detected by a photocell and the electric pulses so produced fed into a computer. *See* DIFFRACTION GRATING.

Industrial applications. Moiré techniques are used in the stress analysis of metals. One procedure is to produce photographically a copy of a grating on the surface of the material under examination and to view it through the master grating. In the unstressed condition no pattern is observed (that is, the moiré fringes are at infinity). When the sample is subjected to elongation or shear, a moiré pattern appears that can be readily interpreted in terms of the strain distribution. *See* PHOTOELASTICITY.

The degree of flatness of a surface can be determined by the moiré technique. The shadow of a grating on the surface serves as the second grating. If the surface is not perfectly flat, the moiré pattern observed no longer consists of parallel lines, and from the pattern a contour map of the surface is obtained. In this manner the examination of large optical surfaces can be carried out.

Any lens alters the apparent spacings of a grating. If now this altered image is viewed through another grating, the moiré pattern obtained is determined by the focal length of the lens. The aberrations of the lens become apparent, since the fringes are no longer straight and equidistant. When a periodic grating is viewed through a refractive index gradient (for example, sugar molecules diffusing into water), the spacings of the grating are modified. On overlaying this figure with another periodic grating, a moiré pattern is obtained which gives directly the refractive index gradient curve. *See* ABERRATION (OPTICS).

Phenomena in physics. Many concepts in physics may be demonstrated with a moiré kit, which consists of families of curves printed on transparent plastic. A simple grating of black and white bars can be regarded as the representation of a plane periodic wave, whose wavelength is given by the length of the periodic element of the grating. The moiré pattern of Fig. 1 represents the interference (or superposition) of two plane monochromatic waves. This occurs, for example, in x-ray diffraction; the interfringe distance corresponds to the reflecting planes of atoms, and Bragg's law is satisfied. The figure also represents other situations in which two plane waves are crossing, such as can occur with reflection at a seawall, in holography, or in a microwave cavity. The rays (perpendicular to the wavefront) satisfy the condition of reflection, namely, that the angle of incidence equals the angle of reflection. *See* X-RAY DIFFRACTION.

Cylindrical or spherical waves are represented by a figure consisting of equispaced concentric circles. Superposing two such figures gives a moiré pattern (**Fig. 2**) consisting of hyperbolas, and in the central portion a family of ellipses. The hyperbolas give the location of the fringes seen in Thomas Young's famous experiment, in which he illuminated two fine slits with a single source of light. The family of ellipses represents the standing-wave modes in an elliptical cavity. It would represent, for example, the sound patterns in a so-called whispering gallery,

where one speaks at one focus of the room and the voice is heard clearly at the other focus.

Diffraction of light by a straightedge can be represented by the moiré pattern produced by the superposition of a straight-line grating and a circular grating; each point of the straightedge becomes a source of cylindrical waves. The moiré pattern is the interference pattern, or hologram, of a point scatterer (a speck of dust, for example). Indeed, the hologram obtained for more complicated objects may be regarded as the moiré pattern produced by the superposition of many such simple figures. *See* HOLOGRAPHY.

The moiré technique can also be applied to problems of fields and flows. Thus in hydrodynamics the flow of liquid from a source may be represented by a figure of radiating lines; the greater the source strength, the smaller will be the angular openings. As no positive or negative radial direction is indicated, the figure can also be used to represent a sink. Two such figures when superposed give a moiré pattern, which is the resultant flow of liquid for a source near a sink. This is also the field for an electric dipole, a positive charge close to a negative charge. When a simple parallel-line grating is superposed on a radial figure, one obtains the flow pattern for a source of liquid in a moving stream. This result, which is ordinarily difficult to calculate, is readily obtained and in graphic form by the moiré method. *See* HYDRODYNAMICS.

The vortex, or circular motion, of fluids is represented by circles in which the spacing is greater as one goes out from the center. Superposition of two such figures gives the resulting flow for two vortex centers of opposite sense (**Fig. 3**). The moiré pattern also represents the lines of the same voltage (the equipotentials) looking down the ends of two wires which are oppositely charged. It should be noticed that the pattern is turned 90° from that for the radial figures. By superposing this circular figure with other figures, one can obtain the flow of air about an airplane wing, for example. *See* AERODYNAMIC WAVE DRAG; VORTEX.

Mathematical solutions. The moiré technique offers a graphical method for the representation and solution of problems in many branches of mathematics. For example, the problems just described are essentially the solutions of Laplace's differential equation in two dimensions. Consider a simpler case, namely, conic sections. A family of curves represents a contour map of some surface. Thus the simple grating represents the level lines for a plane; the more inclined the plane, the closer are the lines. The figure of equispaced circles is the contour map of a cone. Combining a cone map with that of a plane gives the conic section. Superposing a simple grating with a circular grating of the same spacing gives a family of parabolas as the moiré pattern. If the spacings of the straight-line grating are smaller (that is, the inclination of the plane is higher), one obtains hyperbolas; if the spacings are larger, one gets ellipses.

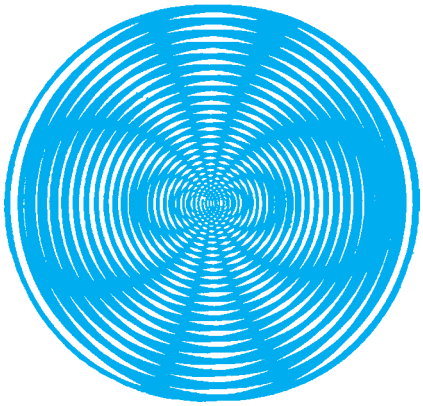

Fig. 3. Moiré pattern produced by superposition of two figures representing vortex or circular motion.

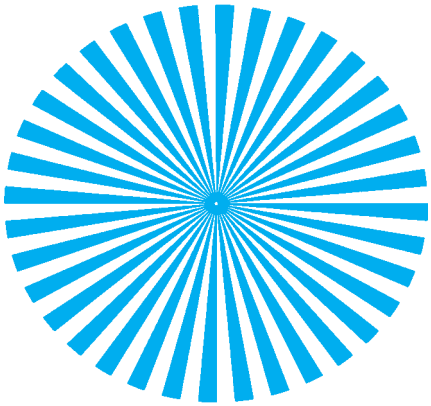

Fig. 4. Radial pattern of 10° openings.

Psychological phenomena. The moiré phenomenon was known to the Chinese in ancient times. Indeed, they produced a luxurious fabric, moiré, or watered silk. It is made by doubling over a glossy, corded silk cloth and pressing the facing surfaces together under steam pressure. When the fabric is unfolded, the ribbed silk with its embossed image superposed creates a moiré pattern. In very recent years moiré has been developed into an art form, and constructions consisting of superposed but separated screens have been exhibited in art museums. The effect on the viewer is one of dynamic motion, because any slight movement of the viewer (even breathing) causes the pattern to move or even change its character.

There are a number of visual psychological phenomena associated with moiré. For example, rapid back and forth movement of a single figure of repetitive elements coarse enough to be resolved by the eye produces a fleeting image. Even when a figure, such as **Fig. 4**, is stationary, the unconscious, rapid, oscillatory movements of the eye give a sparkling character to the figure. The transient moiré pattern produced by the superposition of the immediate image and the transient, displaced afterimage is probably the basis for the liveliness of many examples of so-called optical art (Op art). *See* PERCEPTION; VISION.

Gerald Oster

Bibliography. O. Kafri and I. Glatt, *The Physics of Moiré Metrology*, 1990; K. Patorski, *Handbook of Moiré Fringe Technique*, 1993; D. Post, B. Han, and P. Ifju, *High-Sensitivity Moiré: Experimental Analysis for Mechanics and Materials*, 1997.

Moisture-content measurement

Measurement of the ratio or percentage of water present in a gas, a liquid, or a solid (granular or powdered) material. Nearly all materials contain free water, the relative amount being dependent upon the physical and chemical properties of the material. The primary purpose of determining and maintaining moisture contents within specified limits can usually be traced to economic factors, trade practices, or legal requirements.

Moisture content has a number of synonymous terms, many of which are specific to certain industries, types of product, or material. The water content in solid, granular, or liquid materials is usually referred to as moisture content on either the wet or dry basis; the wet basis is common to most industries. Specifically, moisture content on the wet basis refers to the quantity of water per unit weight or volume of the wet material. A weight basis is preferred. The textile industry uses the dry basis for moisture content of textile fibers. Often referred to as regain moisture content, the dry basis or regain refers to the quantity of water in a material expressed as a percentage of the weight of the bone-dry (thoroughly dried) material. The relationship between the wet and dry moisture-content basis is shown in **Fig. 1**.

The moisture content in air is referred to as humidity, either absolute or relative. Absolute humidity is the number of pounds of water vapor associated with 1 lb of dry air, also called humidity. Relative humidity (RH) is the ratio, usually expressed as a percentage, of the partial pressure of water vapor in the actual atmosphere to the vapor pressure of water at the prevailing temperature. Relative humidity is customarily reported by the U.S. Weather Bureau because it essentially describes the degree of saturation of the air. However, air which is saturated (100% RH) at 50°F (10°C) is quite dry (19% RH) when heated to 100°F (38°C). A changing basis of this type is not convenient for many purposes such as computations used in air-conditioning, combustion, or chemical processing; therefore absolute units, such as dew point or grains of water per pound of dry air, are more acceptable. Dew point is the temperature at which a given mixture of air and water vapor is saturated with water vapor. *See* DEW POINT; HUMIDITY.

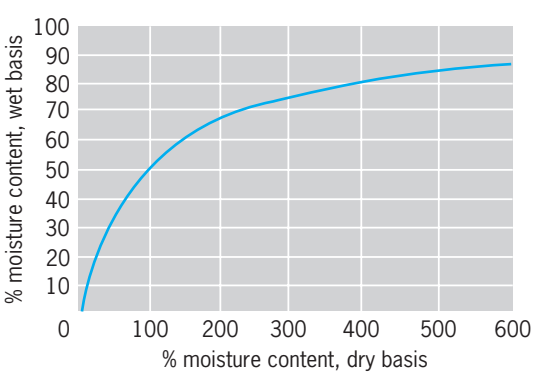

Fig. 1. Dry versus wet moisture-content basis.

Gases

The measurement of water content in gases and mixtures of air and gases is important in industry. A number of commercially manufactured instruments are available for these measurements; their principles of operation include condensation, used in dew- or fog-point indicators; dimensional change, used by hygrometers; thermodynamic equilibrium, used by wet-bulb psychrometers; and absorption methods, which serve as the basic principle for gravimetric and electric conductivity or dielectric types.

The importance of humidity in relation to personal comfort is well known. As a result the air-conditioning industry has grown to considerable proportions by producing equipment to maintain comfortable conditions of temperature and humidity. Considerable industrial air conditioning is also done for process reasons. The textile industry makes wide use of humidity control in rooms for weaving, carding, spinning, and other processes, because the amount of moisture held or absorbed by the textile fibers affects these operations. Paper manufacturers face problems similar to those of the textile industry, as do certain chemical, plastic, and allied processing industries in which control of humidity is important for product quality.

Control of humidity is important in the preservation of materials, especially those which are hygroscopic, and in the storage of food products. In many instances, humidities must be maintained at high levels, as in the storage of apples and vegetables, whereas humidity is maintained at low levels in the storage of dried milks, eggs, and similar products.

Human comfort is affected by high humidities, because the air is so close to its saturation content that it cannot absorb moisture from the surface of the skin and thus cool the individual by evaporation. The higher the temperature of the air, the greater the amount of moisture it can hold. A rule of thumb for human discomfort is that any combination of temperature and relative humidity totaling 130 or higher is uncomfortable. *See* HUMIDITY CONTROL.

Psychrometers. A psychrometer is a device for measuring moisture content of air or gases by means of two thermometers. One thermometer bulb is covered with a wick and maintained wet (the wet bulb); the other bulb is exposed directly to the air or gas (the dry bulb). The evaporation of water from the moistened wick of the wet bulb produces a lowering of its temperature, and by observation of the difference in temperature between the two bulbs the absolute or relative humidity can be determined. For accurate results, the gas or air must have a velocity of 15–20 ft/s (4.5–6 m/s) past the wet bulb. Psychrometric charts or tables are used with the readings obtained from the two thermometers to determine the moisture content of the air or gas. Instruments utilizing this principle are often called wet- and dry-bulb thermometers. *See* PSYCHROMETER; PSYCHROMETRICS.

Hygrometers. Hygrometers measure humidity by the change in dimensions of a hygroscopic material, such as human hair, organic membranes, wood, and plastics. Their most dependable range of operation is from 15 or 20% to 85 or 90% relative humidity at temperatures of 0 to approximately 160°F (−18 to 71°C). Stability of these instruments is better when they are not subjected to extremes of temperature or humidity. There is considerable time lag in the system response to changing humidity conditions when operated at low temperatures. Accuracies of ±3% relative humidity can be expected at normal room temperatures.

Salt conductivity type. Another type of hygrometer utilizes changes in the electrical conductivity of a hygroscopic salt as its operating principle. The conductivity of a solution is dependent upon its concentration (amount of water if the salt content is constant) and its temperature. *See* HYGROMETER.

Gravimetric type. The change in weight of an absorbing material can be used to measure its moisture content, a method known as gravimetric hygrometry. The measurement of moisture content in gases is obtained by passing a known volume of the gas through a suitable desiccant, such as phosphorus pentoxide, silica gel, or similar material, and observing its change in weight. This method is considered a primary standard and is often used in the exact calibration of instruments. It is necessary to make certain that all of the gas or atmosphere has been in contact with the desiccant for a sufficient time to ensure complete absorption of the water vapor. Exact initial and final weighings are also necessary. **Figure 2** shows the basic principle of this instrument.

Hygroscopic materials may also be employed directly to determine changes in water vapor in air. The Aldrich regain indicator makes use of a loose ball of cotton attached on one arm of a sensitive balance; the other arm of the balance serves as a pointer, as in **Fig. 3**. It is important that the sensing material change weight only with change in relative humidity of the surrounding atmosphere.

Dew-point indicators and recorders. When water vapor is cooled, a temperature is reached at which the phase changes to a liquid or solid. This temperature is known as the dew point. The classical method of determining the dew point consists of slowly cooling a polished metal surface until condensation takes place; the temperature at which the first droplet appears is taken as the dew point. The manually operated dew cup is one of the simplest forms of dew-point apparatus. This consists of a polished metal cup containing ether or another volatile liquid into which is placed an accurate mercury-in-glass thermometer. Air is bubbled through the volatile liquid, lowering its temperature and the temperature of the metal cup until dew (condensation) forms on the cup (**Fig. 4**).

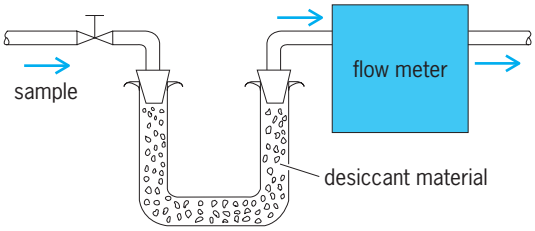

Fig. 2. Gravimetric absorption method.

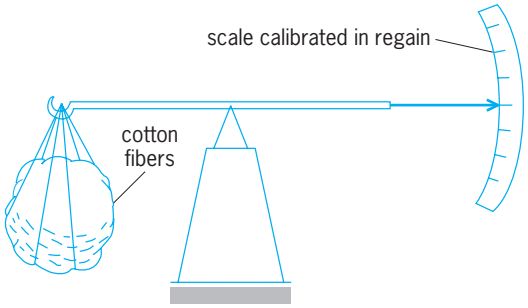

Fig. 3. Aldrich regain indicator.

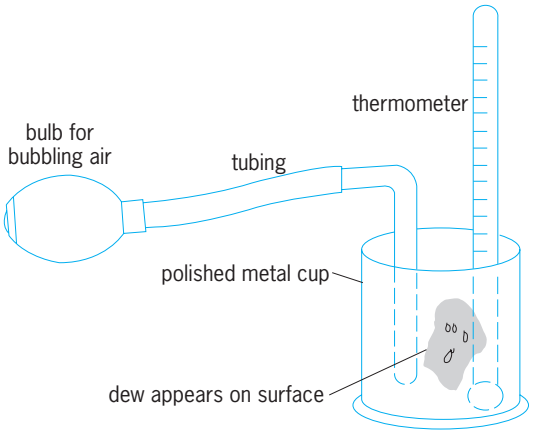

Fig. 4. Manually operated dew-point unit.

These instruments are capable of high accuracy, provided correct techniques are employed. This technique is widely used for measuring water vapor in flue gases, gasoline vapors, furnace gases, compressed gases, and others. More refined instruments utilize refrigeration and closed systems to contain the sample while the condensation is viewed through inspection windows. Others automatically record the condensation point by means of photocells, which alternately control heat or refrigeration to the target area. An adaptation of the dew-point technique makes use of the Wilson cloud chamber principle. The gas sample is compressed and then vented to atmosphere, producing cooling by adiabatic expansion. By repeated trials a pressure ratio can be obtained at which a cloud or fog forms. Dew point is computed from the ratio of the initial to final pressure.

Miscellaneous methods. Several alternate methods exist for determining the amount of water vapor present in gases or air.

The difference in thermal conductivity of dry air and air with water vapor may be determined by measuring the difference in the electrical resistance (temperature) of a hot wire sealed in a small cell. This

method is affected by changes in gas composition. Usually a bridge circuit is used, with a hot-wire cell containing dry air as the reference and a cell with the sample to be tested as the unknown. *See* BRIDGE CIRCUIT.

Spectroscopic methods and index of refraction have been used experimentally, as has the measurement of pressure or volume after the absorption of the water vapor from a sample. *See* SPECTROSCOPY.

Liquids and Solids

The rapid development of instruments for the measurement and control of moisture in liquids and solids was due in a large part to the great need that exists in many processes where the control of a precise moisture is critical. The desirability of a specific moisture content in a product during its preliminary manufacturing process is often required. In general, however, the rigid control of moisture content occurs most frequently in the final product to assure its quality and the fulfillment of legal or trade practices for the individual product.

Instrument types. Instruments suitable for the measurement of moisture content may be classified as periodic and continuous. In general, only those instruments offering continuous measurement are practical for the automatic control of moisture content in a product. The periodic instrument types are generally automated versions of conventional laboratory moisture-analysis procedures. The speed of response for the periodic (intermittent) instruments is typically 2 min or longer, often 15–20 min, making this type of instrument impractical for automatic control. The initial cost of the periodic or intermittent sampling instrument is usually less than for continuous types.

Moisture measuring instruments may also be classified by operating principle. Those instruments employing electrical conductivity (either dc or ac), absorption of electromagnetic energy (radio-frequency regions), electrical capacitance (dielectric constant change), and infrared energy radiations are more readily adapted to continuous measurements inasmuch as the response of these instruments to moisture changes is very fast. Those instruments employing automatic oven drying, chemical titrations (Karl Fischer technique), equilibrium hygrometric methods, distillation methods, and so forth are usually of the intermittent, or periodic, type; the time for analysis is usually 2 min or longer. Most of the more popular methods are outlined in the following discussion.

Electrical conductivity. These methods are based on the relationship between dc resistance and moisture content for such materials as wood, textiles, paper, grain, and similar products. Specific resistance plotted against moisture content results in an approximate straight line up to the moisture saturation point. Beyond the saturation point, where all of the cells and intermediate spaces are saturated with free water, conductivity methods are not reliable. This point varies from approximately 12 to 25% moisture content, depending on the type of product.

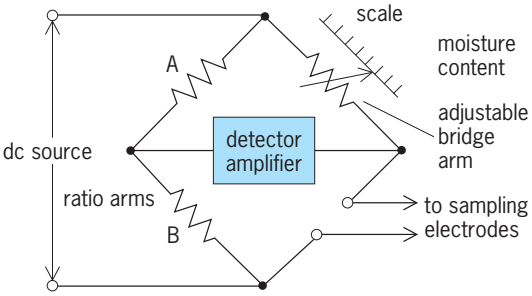

Fig. 5. Basic bridge circuit for conductivity method of moisture-content measurement.

The sample under test is applied to suitable electrodes in the form of needle points for penetrating into woods, plaster, and similar products or flat plates for sheet materials. Granular or fibrous materials may make use of electrodes in the form of a cup or clamp arrangement to confine the material to a fixed volume. The electrodes and sample under test make up one arm of a Wheatstone bridge (**Fig. 5**). The high sensitivity required of the detector dictates the use of electronic amplifiers. The range of resistance values corresponding to normal moisture contents varies from less than 1 to 10,000 megohms or higher, depending upon material, electrode design, and moisture content. Increasing moisture content results in decreasing resistance values. *See* RESISTANCE MEASUREMENT.

Electrical capacitance. These methods are based on the principle of the change in dielectric constant between dry and moist conditions of a material. The dielectric constant of most vegetable organic materials is 2–5 when dry. Water has a dielectric constant of 80; therefore the addition of small amounts of moisture to these materials causes a considerable increase in the dielectric constant. The material being measured forms part of a capacitance bridge circuit, which has radio-frequency power applied from an electronic oscillator (**Fig. 6**). Electronic detectors measure bridge unbalance or frequency change, depending on the method employed.

Electrode design varies with the type of material under test. Parallel-plate types are used for sheet materials, whereas cylindrical electrodes are usually adapted for liquids or powders. Modifications of the parallel-plate capacitor into a rectangular enclosure are used with granular, fibrous, or powdered materials. Range of moisture contents measurable is from

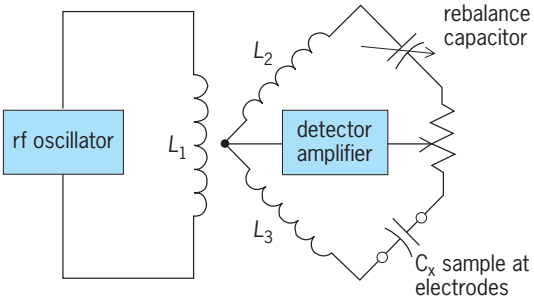

Fig. 6. Basic bridge circuit for capacitance method of moisture-content measurement.

2 or 3% to 15 or 20%, varying with the product. *See* CAPACITANCE; PERMITTIVITY.

Radio-frequency absorption. There are instruments that utilize the attenuation of electromagnetic energy when passed through the material. The radio-frequency types are operated at frequencies below 10 MHz and, for best results, are used with products composed of polar materials. The radio-frequency energy is passed through the polar material and the water molecules absorb some of the energy as molecular motion. *See* ELECTROMAGNETIC RADIATION.

Figure 7 shows the basic block diagram of a typical radio-frequency moisture instrument. These instruments must be calibrated to the particular product under test. By means of suitable electrodes for the sampling technique the instrument can be applied to solid or sheet materials; granular materials are placed into cell-type electrodes. The test results will be affected by any polar material in the sample; however, water of hydration in a product will not be detected. These instruments are mostly available as periodic types but can be adapted for continuous measurements on many products. Moisture content is dependent on the material and ranges from 0.1 to 60%; for certain materials the range will be quite narrow.

Microwave absorption. Instruments employing very high-frequency electromagnetic energy are in limited use for both grab samples and continuous measurement applications. The frequency of the electromagnetic energy is in the region of 1 gigahertz and higher. The 2.45-GHz S band, 8.9–10.68-GHz X band, and 20.3–22.3-GHz K band frequencies have all been used for microwave moisture-analysis instruments. *See* MICROWAVE.

The principle of operation is based upon the fact that the water molecule greatly attenuates the transmitted signal with respect to other molecules in the material in the S and X band frequencies. In the case of the K band microwave frequencies, the water molecule produces molecular resonance. There are no other molecules that respond to this particular resonant frequency, making this frequency most specific to the moisture (free water) in paper products. The wavelength of the K band frequencies are approximately 1.35–1.5 cm long; at these very short wavelengths, the electromagnetic energy may be guided (transmitted) by means of metallic pipes, usually of rectangular shape, called waveguides. Thus the output of a microwave oscillator can be coupled to a horn-shaped opening on the end of the waveguide, propagated through the sample material, and collected by means of an input or receiver horn. The output and input horn are physically separated so that a noncontacting (nondestructive) continuous measurement of the material is possible. The basic block diagram of a microwave moisture instrument is shown in **Fig. 8**. The detection of energy absorption can be made by attenuation (loss) or by phase shift methods, or both measurements may be made simultaneously. Moisture content ranges of less than 1% to 70% or higher can be made. At low moisture contents the S and X bands require a minimum mass of material for practical operation. Accuracies of these instruments are within ±0.5% of indication up to approximately 15% moisture content. *See* MICROWAVE SPECTROSCOPY.

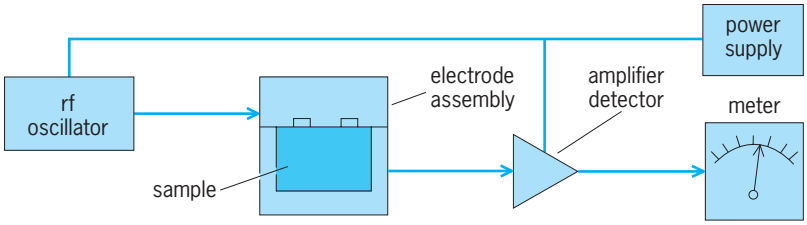

Fig. 7. Radio-frequency-absorption moisture indicator.

Infrared absorption. These instruments operate on the principle of absorption of infrared radiation when passed through the sample material. The water molecule becomes resonant at certain infrared frequencies and thus the amount of energy absorbed by the water absorption band is a measure of the moisture content.

One instrument, used for measuring the moisture content of sheet paper, makes use of infrared energy at two different wavelengths by measuring the differential absorption (**Fig. 9**). Infrared radiation is passed through a rotating disk chopper containing two filters, one filter passing a wavelength of 1.94 micrometers (which covers the water absorption band) and the second filter passing a band of 1.80 micrometers (which is not significantly affected by moisture content). The paper sheet is exposed to the dual-frequency narrow-band radiation and the reflected radiation is collected in an integrating sphere, where it activates a lead sulfide detector. The signals are amplified and the ratio of the two narrow-band signals is read out directly as moisture content. This type of measuring system is quite immune to the effect of paper variables such as basis weight, coatings, density, and composition. The useful range is from 1 to 12% and can be used as either contacting or noncontacting. *See* INFRARED SPECTROSCOPY.

An infrared instrument useful for gas and liquid measurements operates on the principle of selective energy absorption of the sample when compared to a reference material having little or no absorption. The instrument contains two infrared sources (**Fig. 10**). One source is directed through a cell containing the sample material, and the second source is directed through the reference cell, which in gas analysis is

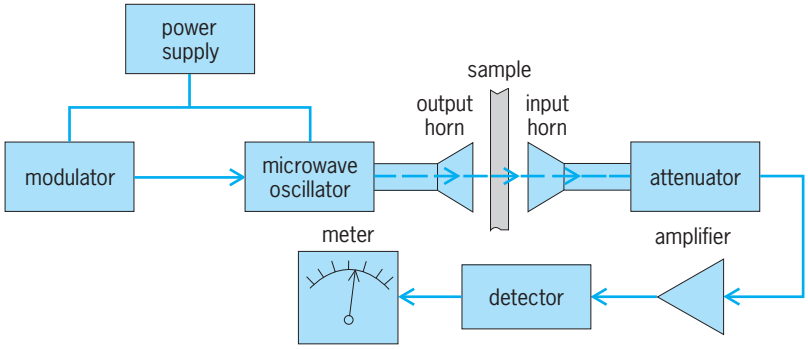

Fig. 8. Microwave-absorption moisture indicator.

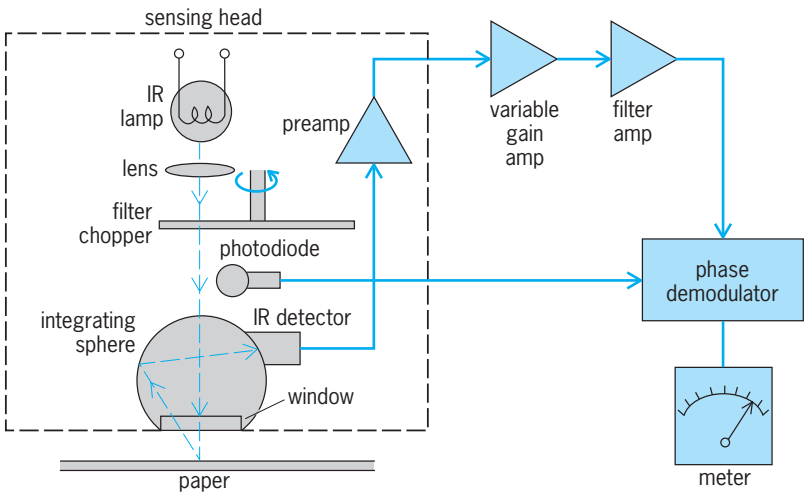

Fig. 9. Infrared-backscatter moisture instrument.

filled with an inert gas, usually air. A detector unit is present which consists of two cavities separated by a metallic diaphragm that acts as a variable capacitor in respect to a fixed plate within the cavity. The operation consists of chopping the infrared energy to the sample and reference cells at 10 Hz. The gas or fluid in the sample cell absorbs the wavelengths which are characteristic of the material and proportional to its concentration. The reference cell does not absorb any appreciable infrared energy. The detector, which has two equal volumes separated by a diaphragm, is filled with gas or vapor of the component of interest. The infrared energy passes through the cells and into the detector, which causes the temperature of gas in the detector to rise; however, if some of the energy has already been absorbed by the sample cell, the sample side of the detector receives less energy than the reference half of the detector. The temperature rise of the gas in the detector causes the pressure to increase and, if the temperature is unequal, the pressure increase is unequal, causing the diaphragm to move in a direction that will tend to equalize the two pressures. This in turn causes the capacitance existing between the diaphragm and its sensing plate to change; this change is used to modulate an oscillator.

The chopper alternately blocks and then passes the energy from each infrared source. Each time the energy is blocked, the pressures equalize in the detector compartments and again the energy is permitted to enter the detector cavities. The modulating frequency is 10 Hz and the magnitude of the capacitance change is directly related to the amount of energy absorbed in the sample cell. The signal from the oscillator is amplified and demodulated for meter indication or applied to a recorder. The range of moisture in gases is from 1000 ppm to 100% water vapor, and in fluids from 6 to 100% water. Response speed is 0.5 s for 90% of full-scale output.

Equilibrium. The equilibrium moisture content of the air at the surface of a material is representative of the moisture within the material. This is particularly true of hygroscopic granular or fibrous materials. Humidity-measuring instruments are used by inserting the measuring element into the material. Moisture content is determined from data of relative humidity plotted against moisture content for the material under test.

Absorption. The water content of liquid organic compounds can be determined by the use of spectrophotometers. The sample is placed in a suitable cell and monochromatic infrared radiation is passed through it. Water vapor absorbs the infrared radiation, but the material absorbs little radiant energy. Thus the radiation varies with the water content. The radiation is focused on a sensitive thermocouple and the resulting voltage amplified. The indication of a sensitive galvanometer can be calibrated in terms of moisture content. This method has been used for accurately detecting 0–10 ppm of water in Freon-12.

Chemical methods. A widely used chemical titration method is the Karl Fischer technique. The method offers extreme sensitivity with good accuracy and covers a wide range of materials. Small samples can be analyzed readily. End points can be determined either visually or electrometrically. The Karl Fischer reagent is added in small increments to a glass flask containing the sample until the color changes from yellow to brown or a change in potential is observed at the end point. Dark-colored solutions require electrical end-point measurement. This method does not differentiate between free and combined water.

Various other methods involve chemical reactions with the free water in liquids or solids by (1) the evolution of free acidic or basic compounds, (2) the evolution of an inert gas, or (3) the formation of an insoluble precipitate. In general, these methods fail to distinguish active hydrogen from water in many materials and are reliable only under laboratory control.

Distillation. A representative sample of the material can be placed in a suitable flask with an excess of liquid which usually has a boiling point higher than water but is immiscible with water. Heat is applied to the flask, and the water and some of the liquid are distilled off. The combined vapors are condensed and collected. The water is separated and

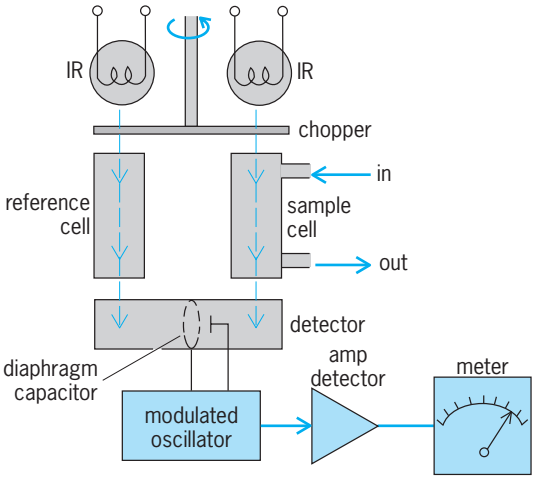

Fig. 10. Infrared gas or liquid analyzer.

measured volumetrically to calculate the moisture content. A variety of laboratory glass apparatus is available for distillation of liquids which are either lighter or heavier than water. The measuring tubes are etched with graduations in cubic centimeters corresponding to grams of water. The moisture content is calculated from the weight of the original sample and the weight of water. This method is slow, and the accuracy depends upon the care exercised and the apparatus used. The method was widely used in the analysis of grains but has largely given way to electrical or oven-drying methods. *See* DISTILLATION.

Oven drying. The oldest and most common analytical method of determining moisture consists of heating the sample to ensure complete drying. Moisture is calculated on the basis of loss in weight between original and dried sample. The method is applicable to many solids and some liquids, and does not require unusual operator skill. Semiautomatic drying and weighing ovens are now available; the moisture content is indicated directly by the weighing scale built into the oven. Problems involve making certain that the material has not lost some of its weight by loss of volatile products other than water, that the material is completely dry at the time of final weighing, and that accurate weighings are made. The oven method is widely used and often serves as the primary standard for calibration of electrical and other indirect methods. Lee E. Cuckler

Bibliography. J. Carmody and J. Lstiburek, *Moisture Control Handbook: Principles and Practices for Residential and Small Commercial Buildings*, 1996; D. M. Considine (ed.), *Process/Industrial Instruments and Control Handbook*, 5th ed., 1999; D. M. Considine and S. D. Ross (eds.), *Handbook of Applied Instrumentation*, 1964, reprint 1984; Instrument Society of America, *Moisture and Humidity*, 1985; A. Pande, *Handbook of Moisture Determination and Control: Principles, Techniques, Applications*, vols. 1–4, 1974–1975.

Mole (chemistry)

A unit (symbolized mol) used to measure the amount of material in a chemical sample. Gram-molecular weight is an older name for the mole and is seldom used. The mole is defined by international agreement as the amount of substance (chemical amount) of a chemical system that contains as many molecules or entities as there are atoms in 12 g of carbon-12 (^{12}C). When the mole is used, the elementary entities need not be molecules, but they must always be specified. They may be atoms, molecules, ions, electrons, or specified groups of such particles.

Quantitative measurements in chemistry are concerned with the amount of material in a given sample. Three obvious ways of measuring the amount are to measure the mass m of the sample (using a balance), to measure the volume V (using a measuring cylinder), or to count the number of molecules in the sample, N. Although it is more difficult to devise an experiment to count molecules, this third way of measuring amount is of special interest to chemists because molecules react in simple rational proportions (for example, one molecule of A may react with one, or two, or three molecules of B, and so forth). For example, in calculating how much air should be provided to completely burn a given amount of octane petroleum fuel to carbon dioxide and water, it is necessary to determine from the molecular formulas and the equation for the reaction how many molecules of oxygen (O_2) react with each molecule of octane, and to count the number of molecules of oxygen in a sample of air and the number of molecules of octane in a sample of the fuel.

However, to count molecules is difficult experimentally and inconvenient in practice because the numbers are so large. For any chemical, a mass of 1 kilogram of the sample contains a large number of molecules, of the order 10^{23}–10^{24}. The mole is defined so that 1 mole of any substance always contains the same number of molecules. This number is large, approximately 6.02×10^{23}, and is known as Avogadro's number. The mole is a more convenient unit in which to measure the amount of a chemical than counting the number of molecules, and it has the same advantages. *See* AVOGADRO'S NUMBER.

The International System of Units (SI) is the internationally accepted system of quantities and units. The mole is the SI unit of the amount of substance (chemical amount), just as the kilogram is the SI unit of mass, and the meter is the SI unit of length. Although amount of substance is the approved name for this quantity, in practice it is not widely used; the name commonly used is number of moles.

The amount of substance (chemical amount) of a sample, n, may be determined in practice by one of three methods.

The value of n (the amount of substance) may be determined from the mass m by dividing by the molar mass M of the sample, as in Eq. (1). If m is

$$n = \frac{m}{M} \tag{1}$$

expressed in g and M in g/mol, then the value of n will be obtained in mol.

For a gas, the value of n may be determined from the volume V, pressure p, and absolute temperature T by using the ideal gas equation (2), where R is the

$$n = \frac{pV}{RT} \tag{2}$$

gas constant ($R = 8.3145$ J K^{-1} mol^{-1}). If pV is expressed in (N m^{-2}) $\times$ (m^3) = J, and RT in J mol^{-1}, then pV/RT gives the value of n in mol. From the ideal gas equation it is determined that at a pressure of 1 atm (= 101,325 pascals) and a temperature of 0°C (= 273.15 K) the volume occupied by 1 mol of an ideal gas, $V = nRT/p$, is 22.414 liters (= 22.414×10^{-3} m^3); and at a pressure of 1 bar (= 10^5 Pa) the volume is 22.711 liters (1 liter = 10^{-3} m^3). *See* AVOGADRO'S LAW; GAS.

For a solution, the amount of solute (or the amount concentration of solution) is frequently determined

by titration: if ν_A molecules of A react with ν_B molecules of B in the titration, then at the end point the amount of A used (n_A) is related to the amount of B (n_B) by Eq. (3), so that if one is known the other

$$n_A = \frac{\nu_A}{\nu_B} n_B \tag{3}$$

may be determined. *See* TITRATION.

The best experimental value for the number of atoms in 12 g of ^{12}C is 6.0221377 (36) $\times 10^{23}$, where the number in parentheses is the uncertainty in the least significant digits. It follows that one mole of any substance contains this number of molecules (or other entities). This leads to the best estimate of the value of the Avogadro constant, N_A, which is $N_A = 6.0221377\ (36) \times 10^{23}\ \text{mol}^{-1}$.

The concentration of a solution may be recorded as (mass of solute)/(volume of solution), in units gram/liter; or as (chemical amount of solute)/(volume of solution) in units mol/liter. Because of the proportionality of chemical amount to number of molecules, the latter is the more useful measure of concentration and is generally used in chemistry and biochemistry. The word concentration without a qualifying adjective is generally taken to mean amount concentration, that is, amount-of-substance concentration. *See* CONCENTRATION SCALES.

Ian. M. Mills

Mole (zoology)

A mammal belonging to the order Soricomorpha (previously Insectivora), family Talpidae (the true moles). There are 39 species in 17 genera of talpids distributed on all continents except Australia. *See* INSECTIVORA; MAMMALIA.

Description. Moles are small, subterranean mammals that are highly specialized for burrowing. The eyes and ears are much reduced, and they have a long pointed snout and short soft fur that can lie either way. The powerful forelimbs are adapted for digging, having very powerful muscles and very large strong claws. Moles have sharp pointed teeth; the dental formula for the star-nosed mole (*Condylura cristata*), hairy-tailed mole (*Parascalops breweri*), and the western moles (*Scapanus*) is I 3/3 C 1/1 Pm 4/4 M 3/3 for a total of 44 teeth. That of the eastern mole (*Scalopus aquaticus*) is I 3/2 C 1/0 P 3/3 M 3/3 = 36, and that of the shrew-mole (*Neurotrichus gibbsii*) is I 3/3 C 1/1 Pm 2/2 M 3/3 = 36. The mole's body is stout and cylindrical with a short neck. *See* DENTITION.

Moles are solitary animals and rarely come aboveground. Earthworms are a staple food of many moles, and they are often stored for later use. The moles paralyze the worms by biting them. There is usually one litter per year, which is consistent with their underground mode of existence; that is, they have few predators. The young approach the size of the adults and begin to fend for themselves when about 4 weeks old. The testes of the males become huge in the mating season.

Molehills are formed of excess dirt pushed up and onto the surface when runways are being dug. Molehills are especially common in wet weather, when the moles are digging new foraging burrows. Foraging burrows are usually temporary and are often visible from above because the soil is pushed up as a ridge. Permanent runs are usually deeper and leave no telltale ridges above the burrows. Moles can turn around in the burrows, and they can run backward in them nearly as fast as they can run forward.

North American moles. The main North American moles—*Condylura* (1 species), *Scalopus* (1), *Parascalops* (1), and *Scapanus* (3)—fall into a related group, the subfamily Scalopinae, separate from the Old World moles, but this group also includes one species from China, *Scapanulus oweni*. The rest of the moles occur mostly in the Old World and are in the subfamily Talpinae, although one genus and species from this group, *Neurotrichus gibbsii*, occurs on the west coast of North America.

The eastern mole (*Scalopus aquaticus*) has the largest range of any of the North American moles: from the Atlantic coast west to southern Minnesota, through Nebraska, and to West Texas. It has a very short, nearly naked tail. The main food of this species, like most moles, is earthworms, but they eat numerous insect larvae, especially "grubworms" (scarabaeid or Junebug larvae), many other invertebrates, and some plant material. They enter ant nests and feed heavily on ant eggs, larvae, and adults. They are active for about 4 or 5 hours, then sleep for a similar time. These moles move rapidly through deeper underground tunnels to a foraging area, then they move rapidly back and forth thus disturbing the soil, which probably causes earthworms and other prey to move and be detected. Mating occurs early in the spring; after a gestation of 30–40 days the single litter of three to seven is born.

The hairy-tailed mole (*Parascalops breweri*) occurs mostly in woodlands from southeast Canada through much of New England and New York. This species has a short, hairy tail. It may come aboveground more than most moles, and may even forage some on the surface, which helps to account for the number caught by cats and other predators. There are many foraging burrows near the surface; permanent burrows are deeper and fewer. Mating is in late March or April, and the young (usually four or five) are produced in late April or May.

Three species of *Scapanus* occur on the west coast of North America—*S. townsendii* and *S. orarius* from southern British Columbia to northwestern California, and *S. latimanus* from southern Oregon to northern Baja California. The tail is short and fairly thick. These moles are often found in forests. Other small mammals, such as mice and shrews, often use their burrows. Nest chambers are lined with dry plant material.

The star-nosed mole of North America (*Condylura cristata*) is unusual in that it is gregarious, is a good swimmer, and prefers wet, mucky areas. It has a long, hairy tail. Encircling the snout is a highly developed sensory organ of 22 projections, or tentacles. These

tentacles were long thought to be sensory, but recent behavioral and experimental evidence suggests that the tentacles are used to detect electrical current. The moles dip their snouts into the water before entering and scan under water with the tentacles. They direct their bites toward areas with strong electrical output, the clitellum ("muscle") and reproductive parts of earthworms. The burrows often end at the water. It is the only North American mole that occurs in mucky situations. Molehills in muckland in the northeastern United States are sure to be those of *Condylura*, although this species sometimes ranges into lawns from associated ponds.

Neurotrichus gibbsii is called the shrew-mole and is closely related to the Japanese shrew-mole (*Urotrichus*). It is much smaller and thinner than other moles. The forefeet are much smaller in *Neurotrichus* than in other talpids, but they do have long curved foreclaws. The tail is about one-third the body length. *Neurotrichus* occurs on the west coast of North America from southern British Columbia to northern California, primarily in deep woodlands. Like other moles, it feeds heavily on earthworms, slugs, snails, centipedes, and insects and their larvae, and has been known to eat salamanders. It is a good swimmer. It climbs low bushes, apparently in search of insects. There may be several litters per year consisting of one to four. Young have been seen in all months except December and January.

Eurasian moles. The desmans—*Desmana moschata* of Russia and the Ukraine and *Galemys pyrenaicus* of the Pyrenees Mountains in France and the northern part of the Iberian peninsula—are, like *Condylura*, long-tailed and highly aquatic. Both have very long, flexible snouts, useful for exploration, and large hindfeet useful for swimming. The Russian desman has a widened flat snout, with grooves above and below. Scent glands at the base of the tail give this animal a musky odor. The hindfeet are webbed, and there are stiff hairs along the edges of the feet. Both of these characters give greater surface area, which is helpful in swimming. This species lives along freshwater streams and ponds and feeds on aquatic foods—fish, crustaceans, insects, and worms. Its den is in a chamber above the water line. A tunnel up to 6 ft (1.8 m) long opens into the water below the surface. The Pyrenean desman is found along swift-flowing streams and is a strong swimmer. Unlike many moles, there are three annual peaks of breeding in this species.

The genus *Talpa* of Eurasia contains nine species. The European mole (*T. europaea*; see **illustration**), found in northern and central Asia as well as Europe, is a rather typical representative of the group and spends most of its life underground, emerging in the spring to gather leaves for its nest and to look for water. The tunnels range from just below the surface to depths of nearly 3 ft (0.9 m). There is usually one litter per year of two to seven. Gestation takes about 28 days.

Other types. Other moles are *Scapanulus* (1 species), *Scaptonyx* (1), and *Uropsilus* (4) of China; *Parascaptor* (1) of Burma; and *Euroscaptor* (6), *Mogera* (5), *Scaptochirus* (1), and *Dymecodon* (1) of Asia.

European common mole (*Talpa europaea*).

In addition, there are two families of molelike mammals that are not talpids. They are the African golden moles, Chrysochloridae, and the marsupial moles, Notoryctidae. These two families occur in Africa and Australia, respectively, parts of the world where talpids do not occur. There they occupy similar niches to the talpids, which means they behave somewhat similarly and have evolved to fit similar conditions.

John O. Whitaker, Jr.

Bibliography. M. L. Gorman and R. D. Stone, The natural history of moles, Comstock Publishing Associates, a division of Cornell University Press, Ithaca, 1990; E. Gould, W. McShea, and T. Grand, Function of the star in the star-nosed mole, *J. Mammal.*, 74:108–116, 1993; G. D. Hartman and T. L. Yates, Moles, Talpidae, pp. 30–55 in G. A. Feldhamer, J. A. Chapman, and B. C. Thompson (eds.), *Wild Mammals of North America: Biology, Management, and Conservation*, 2d ed., Johns Hopkins University Press, Baltimore, 2003; R. Hutterer, Order Soricomorpha, pp. 220–311 in *Mammal Species of the World: A Taxonomic and Geographic Reference*, ed. by D. E. Wilson and D. M. Reeder, 2005; R. M. Nowak, *Walker's Mammals of the World*, 3 vols., Johns Hopkins University Press, Baltimore, 1999; J. O. Whitaker, Jr., and W. J. Hamilton, Jr., *Mammals of the Eastern United States*, Cornell University Press, Ithaca, 1998; D. E. Wilson and S. Ruff, *The Smithsonian Book of North American Mammals*, 1999.

Molecular adhesion

The tendency of dissimilar solids or liquids to cling together as a result of the interatomic forces which they exert upon each other across their common interface. Some factors affecting molecular adhesion are the physical state of the materials, the composition and the topology of the surfaces in contact, the temperature, and the presence of foreign materials such as adsorbed gases on one or both surfaces.

Mutual forces of attraction, electrical in origin, exist between virtually all atoms and molecules. It is these forces which, under proper conditions of temperature and pressure, compel isolated atoms or molecules to condense into solids and liquids. On average, these forces are approximately equally strong in all directions and result in an atom or molecule

achieving its lowest possible energy, that which is thermodynamically favored, when it is surrounded on all sides by other atoms. The presence of a real surface breaks the three-dimensional continuity of the substance, leaving atoms at the surface with an unfulfilled capacity for bonding. It is this capacity for bonding that is responsible for the phenomena of surface tension, adsorption of gases, and adhesion. *See* ADSORPTION; COHESION (PHYSICS); INTERMOLECULAR FORCES; SURFACE TENSION.

In principle, the strength of the adhesive bond between two substances may be quite large—larger, in fact, than the cohesive strength of the weaker material. In practice, however, many factors act to diminish the strength of adhesion. Probably the most significant are incomplete contact of the surfaces due to their microtopography and the presence of adsorbed gases or other surface contaminants. These may diminish the surface's capacity to further bond or simply replace one of the original surfaces with an overlayer which is significantly weaker.

One technological application in which molecular adhesion plays an especially significant role is the deposition by evaporation or sputtering of thin films on substrates of dissimilar composition. A particularly important example is the use of such films to form conducting paths—microscopic wires only a few tens of nanometers thick—in integrated circuits. These films must both adhere well to the underlying substrate and produce contacts with appropriate electrical properties. At present the understanding of adhesion is not sufficiently detailed to be able to predict such properties in advance of experiments. *See* INTEGRATED CIRCUITS; SPUTTERING.

Enhancement has been observed in the degree of adhesion between evaporated layers and the underlying medium as a result of irradiation by heavy ions with energies of a few megaelectronvolts. In these experiments, enhanced adhesion has been observed between metal films and metal, semiconductor, and insulating substrates. Materials as dissimilar as gold and poly(tetrafluoroethylene) have been joined, and it has also been observed that particle irradiation transforms electrically rectifying interfaces into ohmic ones. Further research in ion-beam enhanced adhesion promises not only new technological advances but also improved understanding of molecular adhesion at the most fundamental level. *See* ADHESIVE; INTERFACE OF PHASES; SURFACE TENSION.

Robert A. Weller

Bibliography. K. W. Allen (ed.), *Adhesion*, vols. 1–15, 1977–1991; L. H. Lee, *Fundamentals of Adhesion*, 1991; D. E. Packman (ed.), *Handbook of Adhesion*, 1993; R. L. Patrick (ed.), *Treatise on Adhesion and Adhesives*, vols. 1–7, 1967–1991; M. Prutton, *Introduction to Adhesion*, 1994.

Molecular anthropology

The study of primate phylogeny and human evolution through the genetic information in the deoxyribonucleic acid (DNA) of genomes and in the proteins that genes encode. The first studies in molecular anthropology used immunological and biochemical methods to obtain information from proteins on the degrees of genetic similarity of humans and other primates. These results not only placed chimpanzees and gorillas closest to humans rather than to orangutans but also indicated that the very close kinship between chimpanzees and gorillas was not any closer than the relation of each to humans. Subsequent studies that extracted genetic information directly from DNA extended this original finding. Indeed, the accumulating comparative DNA sequence data provide mounting evidence that the closest genetic kinship is between chimpanzees and humans rather than chimpanzees and gorillas.

DNA data. The results gathered in DNA studies of primate phylogeny challenge the traditional anthropological view that humans are very different from all other animals. DNA results show that, genetically, humans are only slightly remodeled apes. Humans share with their closest relatives, the chimpanzees and bonobos (pygmy chimpanzees), more than 98.3% identity in typical noncoding DNA and probably about 99.5% identity in the active coding sequences of functional nuclear genes. Humans share about 98.0% identity in nuclear genomic DNA with gorillas, 96.5% identity with orangutans, and 95.0% identity with the most distant ape relatives, the gibbons and siamangs. Apes and humans share with the other branch of catarrhines, the Old World monkeys, about 92% identity in nuclear genomic DNA, and with the platyrrhines, the New World monkeys, about 87% identity. Even with nonanthropoid primates, the tarsiers and strepsirhines (lemurs and loriforms, such as bushbabies), the anthropoids (platyrrhines and catarrhines) share a DNA identity in the range of 71–76%. New information on the social lives and high intelligence of chimpanzees and other apes also challenges the traditional view that humans are very different from other animals. Chimpanzees use tools and have material culture, are highly social, and may even have the mental capacity for abstract thought and symbolic communication, for example, for learning rudimentary forms of language. *See* DEOXYRIBONUCLEIC ACID (DNA).

Traditional primate classifications provide a poor guide to the course of primate phylogeny. These classifications use the vague concept of grades of evolutionary advancement to place the smaller-brained primates in the suborder Prosimii (the primitive grade) and the larger-brained primates in the suborder Anthropoidea (the advanced grade). Moreover, on viewing humans as the most advanced primates, traditional primate classifications have humans as the sole living members of family Hominidae, while the great apes of Africa (chimpanzees, bonobos, gorillas) and Asia (orangutans) are members of subfamily Ponginae of family Pongidae, despite overwhelming evidence that the African great apes share a more recent common ancestry with humans rather than with orangutans. This is the traditional anthropocentric view of the place of *Homo sapiens* in the order Primates. In contrast, a strictly objective view based on molecular evidence, but also congruent morphological evidence from both living and fossil primates,

not only places apes with humans within the family Hominidae but also within this family places chimpanzees and bonobos with humans in the genus *Homo*. *See* APES; FOSSIL HUMANS; MAMMALIA.

Primate phylogeny. The course of primate phylogeny, as reconstructed from the molecular and fossil evidence, may be sketched as follows. During the Paleocene geological epoch of 65–57 Ma (Mega annum, or million of years before the present), the ancestors of the modern primates divided first into strepsirhines and haplorhines and then, within the haplorhines, into the tarsiiforms and basal anthropoids. In the early Eocene epoch at about 50 Ma, the strepsirhines divided into lemuriforms and loriforms; and in the middle to later Eocene epoch at about 40 Ma, the basal anthropoids divided into platyrrhines and catarrhines. In the late Oligocene epoch at about 28–25 Ma, while strepsirhines and platyrrhines were dividing into familial clades, the basal catarrhines divided into cercopithecids (Old World monkeys) and hominids (apes and humans).

In the early Miocene epoch at about 23–22 Ma, familial clades divided into subfamilial clades, which by 20–15 Ma divided into tribal clades. Thus, because the last common ancestor of all living apes and humans lived at about 18 Ma, the living hominids are grouped together as subfamily Homininae. This phylogenetic classification not only has each taxon represented by a monophyletic group or clade but also indicates by gradations in taxonomic rank how far back in time is found the last common ancestor of the living members of the clade.

The phylogenetic classification of living hominids is as follows [each age (in Ma units) shown for a taxon is the age of the last common ancestor of all living members of that taxon]:

Family Hominidae
- Subfamily Homininae (18 Ma)
  - Tribe Hylobatini
    - Subtribe Hylobatina (8 Ma)
      - *Symphalangus syndactylus*: siamang
      - *Hylobates lar*: white-handed gibbon
  - Tribe Hominini (14 Ma)
    - Subtribe Pongina
      - *Pongo pygmaeus*: Borneo orangutan
    - Subtribe Hominina (7 Ma)
      - *Gorilla gorilla*: gorilla
      - *Homo* (6 Ma)
        - *H.* (*Pan*) (3 Ma)
          - *H.* (*P.*) *troglodytes*: chimpanzee
          - *H.* (*P.*) *paniscus*: bonobo (pygmy chimpanzee)
        - *H.* (*Homo*) *sapiens*: humankind

After Homininae divided into tribes Hylobatini and Hominini during the late Miocene epoch at about 8 Ma, subtribal hylobatans of the hylobatin clade divided into *Hylobates* (gibbons) and *Symphalangus* (siamangs). The hominin clade in the middle Miocene epoch at about 14 Ma divided into subtribes Pongina, from which *Pongo* (orangutans) evolved, and Hominina. The latter, in the late Miocene epoch at about 7 Ma, divided into *Gorilla* and *Homo*, and at about 6 Ma *Homo* divided into the subgenera *Homo* (*Homo*), out of which humankind evolved, and *Homo* (*Pan*), which in the Pliocene epoch at about 3 Ma separated into *H.* (*Pan*) *troglodytes* (chimpanzees) and *H.* (*Pan*) *paniscus* (bonobos, or pygmy chimpanzees). Other subtribal clades such as those of New World monkeys as well as those of Old World monkeys also divided into genera during the late Miocene epoch at about 10–7 Ma. Then genera divided into subgenera or species groups in the late Miocene to early Pliocene epoch at about 6–4 Ma. Thus, classifying humans, chimpanzees, and bonobos as members of the same genus is equivalent to how other primate clades at the same age are classified.

Ongoing evolution. After the divergence of *Homo* (*Homo*) from *Homo* (*Pan*), humankind's emergence was marked by mutations (such as DNA sequence changes) that spread to fixation in the ancestors of all modern humans. These mutations are the human-specific factors which distinguish the human species genetically from all other species. To obtain a comprehensive catalog of these mutations, genomes of a chimpanzee and a bonobo (closest relatives to humans) and a gorilla (next closest relative) need to be sequenced and compared to the sequence of a human genome. Scientists who call for a Human Genome Evolution Project as a companion to the Human Genome Project envision that the overall goal would be to identify and analyze the functions of human-specific genetic factors involved in the evolution of unique features of human anatomy (for example, bipedal locomotion, greatly enlarged brain) and behavior (such as speech, higher-order cognitive functions).

Ongoing evolution involves mutations that have not spread to fixation, either because they occurred too recently or because natural selection has maintained a polymorphic state. These mutations occur at frequencies that occasionally differ from one human group to another. They account for the genetic diversity found in the human species. Extensive comparative data now exist on the genetic diversity due to mitochondrial haplotypes (each, a set of genetic determinants located on a mitochondrial genome). These mitochondrial genetic variants arise from mutations in the DNA carried by mitochondria, the maternally inherited (from ova) cellular organelles that drive oxidative metabolism. These organelles have their own genomes that encode key proteins and RNA molecules found in mitochondria. Although a mitochondrial genome is quite small (16,000 nucleotide base pairs in length) compared to the nuclear genome (3 billion base pairs in length), each somatic cell contains many mitochondria; thus mitochondrial DNA can be readily prepared and haplotypes of such DNA in different individuals can be determined. The species-wide distribution of mitochondrial haplotypes in present-day human populations compared to the distribution in chimpanzees reveals that humans show less DNA diversity than chimpanzees. In fact, the diversity that humans show is no larger than that shown by a chimpanzee subspecies. While populations in different subspecies of

chimpanzees share only 28% of the total species diversity, human populations in different ethnic groups (European Norwegians and South African hottentots) and even within the same group share 85–90% of the total species diversity. The human genetic diversity due to mutations in nuclear DNA (the DNA carried by the chromosomes of cells) also shows this same pattern in which most of the variation within the human species as a whole is contained within single populations. These findings argue that all living humans are biologically not only members of the same species but also the same subspecies and even the same race, the human race. A further inference from the data on mitochondrial genetic diversity is that the birth of the human race took place in Africa about 200,000 years ago.

Recently, small pieces of mitochondrial DNA were recovered from bones of an undated Neandertal specimen. On comparing mitochondrial haplotypes from present-day humans to the Neandertal haplotype and to chimpanzee haplotypes, and assuming a molecular clock with a reference date of 6 Ma for the age of the last common ancestor of humans and chimpanzees, the estimated date for divergence of the Neandertal haplotype from the lineage to the last common ancestor of the present-day human haplotypes is about 800,000 years ago, as compared to about 200,000 years ago for the age of the last common ancestor of present-day human haplotypes. These age estimates as compared to those for the last common ancestor of haplotypes from different chimpanzee subspecies suggest that Neandertals belonged to the same species as humans do, but a different subspecies. The results so far gathered on nuclear genetic diversity in present-day human populations leave open the possibility that as the founder members of the human subspecies, *Homo sapiens sapiens*, multiplied and spread from one continent to another, the gene pool of these migrants could have been enriched by admixture of some genes from the preexisting archaic human populations. *See* MITOCHONDRIA.

Morris Goodman

Bibliography. M. Goodman, Epilogue: A personal account of the origins of a new paradigm, *Mol. Phylogenet. Evol.*, 5:269–285, 1996; M. Goodman et al., Toward a phylogenetic classification of primates based on DNA evidence complemented by fossil evidence, *Mol. Phylogenet. Evol.*, 9:585–598, 1998; M. Krings et al., Neandertal DNA sequences and the origin of modern humans, *Cell*, 90:19–30, 1997; E. H. McConkey and M. Goodman, A human genome evolution project is needed, *Trends Genet.* 13:350–351, 1997; M. Ruvolo, Genetic diversity in hominoid primates, *Annu. Rev. Anthropol.*, 26:515–540, 1997.

Molecular beams

Utilization of well-directed streams of atoms or molecules in vacuum. This is a cornerstone technique in the investigation of molecular structure and interactions. Molecular beams are usually formed at sufficiently low particle density for the interaction of one beam molecule with another to be negligible. This ensemble of truly isolated molecules is available for the spectroscopic study of molecular energy levels using photon probes from the radiofrequency to optical portions of the electromagnetic spectrum. Some of the best-determined fundamental knowledge of physics comes from spectroscopic molecular-beam experiments. Beyond this, beams can be applied as probes of the multifaceted nature of gases, plasmas, surfaces, and even the structure of solids. An application intermediate in complexity is the study of molecular interactions determining the properties of plasma and electric discharge devices, the nature of the upper atmosphere, and some aspects of the cooler astrophysical regions. *See* SCATTERING EXPERIMENTS (ATOMS AND MOLECULES).

Production and detection. One simple means of forming a beam is to permit gas from an enclosed chamber to escape through a small orifice into a second chamber maintained at high vacuum by means of large pumps. **Figure 1***a* shows such effusion into a collimating chamber from an oven chamber, generally heated to control the vapor pressure of the gas. The molecules coming from the orifice are distributed in angle according to a cosine law, illustrated by the circle downstream of the orifice which represents the envelope of relative beam flux vectors. A useful number of molecules passes forward along the horizontal axis of the apparatus. A well-collimated beam is then formed by requiring that those molecules entering the test chamber where an experiment is to be performed pass not only through the orifice but also through a second small hole separating the collimating and test chambers. A property of beams formed this way is that the velocities of the individual molecules have a large thermodynamic spread in values centered on a mean value of order 10^3 m/s (3.3×10^3 ft/s) determined by the oven temperature.

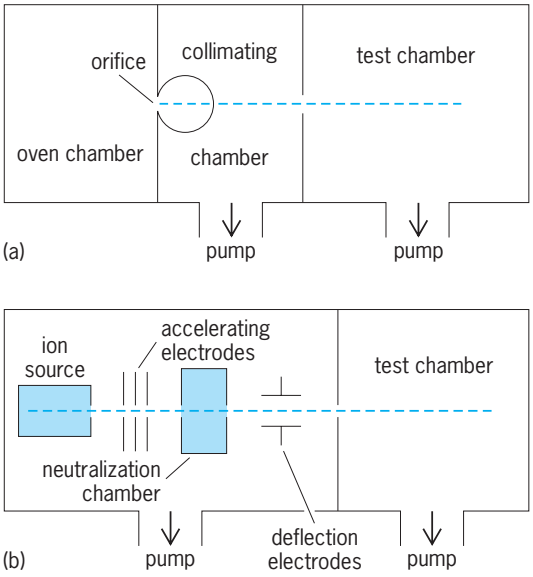

Fig. 1. Schematic diagrams of systems for producing molecular beams. (*a*) Conventional oven-beam system. (*b*) Charge-exchange beam system.

Charge-exchange system. If higher velocities are desired, then a charge-exchange beam system can be used. In this scheme, ions are produced by some ionizing process such as electron impact on atoms within a gas discharge. Since the ions are electrically charged, they can be accelerated to the desired velocity and focused into a beam using electric or magnetic fields. The last step in neutrally charged beam formation is to pass the ions through a neutralizing gas where electrons from the gas molecules are transferred to the beam ions in charge-exchange molecular collisions. If the acceleration voltage is relatively high, then the ion-beam velocity will be 10^5 m/s (3.3×10^5 ft/s) or greater upon entering the neutralizer. In many cases, such energetic charge-exchange collisions produce beam atoms or molecules in internally excited energy levels rather than just in the ground-state level of lowest energy occupied by atoms in low-temperature gases. On the other hand, useful beams of internally excited molecules can be formed from ground-state molecules which have been excited by energetic electron- or ion-beam impact or by photon absorption from a laser or other light beam. *See* EXCITED STATE; GROUND STATE; ION SOURCES.

Secondary electron detection. The faster types of neutral molecular-beam particles are easy to detect by secondary electron ejection from a solid. A collision of a beam molecule with a surface is sufficiently violent for one or more electrons to be ejected. For intense beams, the rate of electron production is so high that the electrons can be collected and measured electronically as an electric current. At lower rates, the effect of each electron can be multiplied by means of an alternating sequence of electron acceleration and surface ejection steps, to produce a burst of 10^6 or more electrons from a single beam molecule. This pulse of current is adequate for electronic pulse counting, leading to a very sensitive overall beam-detection technique useful at beam intensities as low as five molecules per minute.

Other detection techniques. For molecular beams effusing from an oven, a variety of detection techniques have been devised. Alkali atoms have such small ionization potentials that their valence electrons can be transferred to a heated metal surface having a large work function, such as tungsten. The resultant ions can then be detected by current measurement or particle multiplication techniques, depending again upon the beam intensity. A second special technique useful for some beams of reactive or excited molecules is electron ejection from a surface, powered by the internal energy of these special beam molecules. A universal detector for any kind of slow molecule employs initial conversion into ions by electron- or light-beam impact; the ions are then accelerated into a fast beam for easy detection.

Special beams. Occasionally a beam containing atoms or molecules in a specific quantum-mechanical state is needed. Energy-resonant transitions between the ground state and the specific excited state can be utilized, often induced by single-frequency laser radiation. Beams of slow molecules with magnetic (or electric) dipole moments can be selected according to the direction of orientation of their moments; spatial separation into component beams is achieved through orientation-dependent interaction between the dipole moment and an externally applied strong nonuniform magnetic (or electric) field. Some precision spectroscopic molecular-beam experiments use such beams, and form the basis for atomic-clock precision time standards. *See* ATOMIC CLOCK.

Molecular-beam spectroscopy. Much of molecular spectroscopy involves the absorption or emission of light by molecules in a gas sample. The frequency of the light photon is proportional to the separation of molecular energy levels involved in the spectroscopic transition. However, the molecule density in typical gas samples is so high that the energy levels are slightly altered by collisions between molecules, with the transition frequency no longer characteristic of the free molecule. Using low-density molecular beams with their sensitive detection techniques can reduce this collision alteration problem, with the result that atomic properties can be measured to accuracies of parts per million or better. If the simplest atoms or molecules are employed, the basic electromagnetic interactions holding the component electrons and nuclei together can be precisely studied. This is of great importance to fundamental physics, since theoretical understanding of electromagnetic interactions through quantum electrodynamics represents the most successful application of quantum field theory to elementary particle physics problems. *See* QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY.

Properties measured. Among the basic quantities of physics measured by molecular beams are the fine-structure constant $\alpha = e^2/\hbar c = 1/137.0361$ (e is the electron charge in electrostatic units, $\hbar$ is Planck's constant divided by 2π, and c is the speed of light), obtained from microwave spectroscopy studies of the fine and hyperfine energy-level splittings in one- and two-electron atoms; the value 1.5210326×10^{-3} of the magnetic moment of the proton in Bohr magnetons, obtained with the hydrogen maser; the purely quantum-electrodynamic shifts of atomic energy levels called Lamb shifts; the nuclear magnetic dipole moments of several hundred isotopes; the equality in magnitude of the unit electron and nuclear charges to an accuracy of better than parts in 10^{18}; and the absence of intrinsic electric dipole moments for the electron and proton, which is a test of parity and time reversal as symmetry properties obeyed by the electromagnetic interaction. The isolated, unperturbed nature of atoms in a beam is needed for the device used as the length standard, based upon the 605.7-nanometer wavelength of krypton atoms being reproducible to parts in 10^9. The time standard is the cesium-beam atomic clock operating on the ground-state hyperfine splitting microwave transition with a reproducibility over hours of parts in 10^{12}.

Apparatus. The molecular-beam magnetic resonance apparatus that is used in a number of the

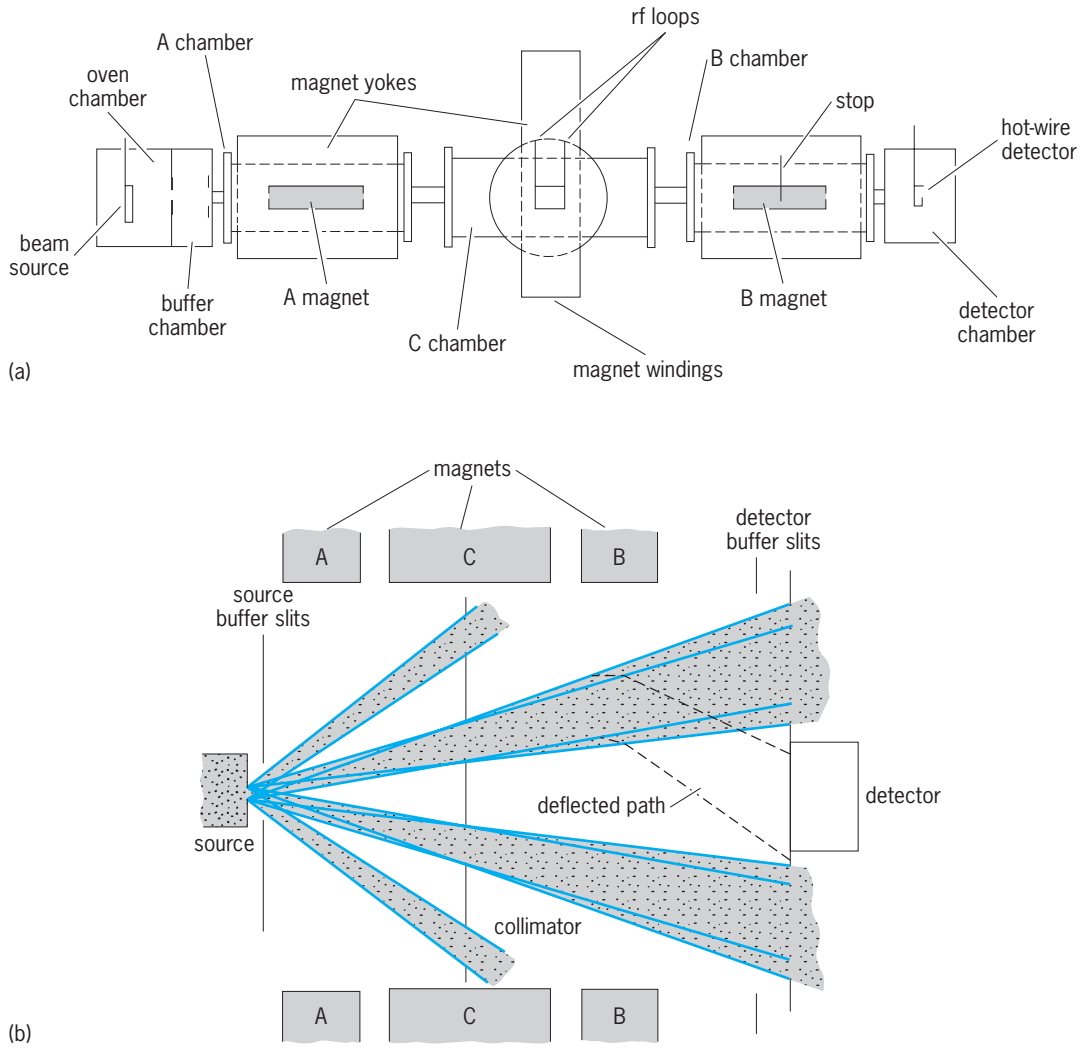

Fig. 2. Molecular-beam magnetic resonance spectroscopy experiment. (*a*) Apparatus used in the experiment. (*b*) Spatially resolved multiple-beam configuration showing typical molecular trajectories.

spectroscopy experiments is shown in **Fig. 2**, along with a diagram showing typical molecular trajectories. On the left end of Fig. 2*a* is an oven beam source, here of alkali atoms; on the right end is a hot-wire surface ionization detector of these atoms. Between the ends are three separate regions where external fields are applied to the beam. The A region is a state selector, and the B region a state analyzer; that is, some property of the beam atoms is well defined in the A region, and then this situation is checked in the B region. In the example of Fig. 2 the selected property is the direction of the atoms' magnetic moment. The center of the apparatus is the transition region marked C, where a resonant radio-frequency transition can be induced between atomic energy levels associated with different directions for the magnetic moment. This alteration of the atoms entering the analyzing region results in a change in atom trajectories, with a corresponding change in the intensity of the beam striking the hot-wire detector. For example, in Fig. 2*b*, molecules undergoing a transition within the C region follow the deflected path and are detected. When the frequency of the radio waves in the C region is not resonant, such a change does not occur. Determining the frequency for maximum effect results in a measure of the atomic energy level splitting, which for the example of Fig. 2 is directly related by theory to the atom's nuclear magnetic moment.

The power of molecular-beam techniques derives, in large measure, from the ability, in principle, of the A and B regions to contain any kind of state-selective device utilizing combinations of static or time-varying electromagnetic fields. Also, the interaction region C can involve resonant interactions with light or other photons or even nonresonant-state destructive interactions such as collisions with atoms in an introduced gas. If the A and B regions do not control beam atom trajectories, but instead some other property such as the state of internal energy, then correspondingly the nature of the beam detector might need to be sensitive to the controlled property.

Use of fast beams. The fast molecular beams produced by charge exchange increasingly have been employed in spectroscopy experiments and offer some advantages in addition to easy detection. Among these is the fractionally well-defined and controllable molecule velocity determined by the ion acceleration voltage, a feature useful in the

study of time-dependent quantum-mechanical interference effects on transition rates. The ability to transport rapidly decaying excited atoms through an apparatus is also enhanced using fast beams. A highly accurate atomic fine-structure measurement employed a fast hydrogen atom beam, as did an experiment on multiphoton microwave transitions between highly excited atomic states.

Scattering experiments. The molecular-beam study of interactions occurring during individual collisions of two species of electrons, ions, atoms, or molecules can be divided into two categories, depending upon whether the targets for the particles in one beam are those in a gas or those in another beam. In addition, the nature of the collisional interaction depends strongly on the velocity of impact of the two particles; if this is very low, all that can happen is that the two particles elastically scatter off each other, with no change in the internal nature of either particle being energetically possible. At higher collision velocities, inelastic processes can occur in which a part of the energy of collision is converted into a change in internal energy of one or both particles. Particle rearrangement can also occur, such as the electron transfer used previously in charge-exchange atomic-beam sources, or proton transfer in chemical reactions between molecules. Collisions involving fast electrons or fast bare nuclei are primarily probes of the structure of the target molecule. The slower collision processes are more concerned with the composite molecular system transiently formed during the collision time. An example is the negative ion state formed when an electron joins a neutral atom or molecule. Such compound states cannot live indefinitely since the total energy of the system is positive; in a rearrangement collision, however, the compound state breaks up into molecules different from the originally colliding ones.

Quantities measured. The simplest quantity to measure in a molecular-beam collision experiment is the probability for a particular elastic or inelastic event to result from a single collision of one beam particle with one target particle. This probability is usually expressed in terms of an apparent size of the target particle for the process, called a total cross section. This quantity is a sum over contributions from all possible distances of closest approach of the two particles, since molecules are so microscopic that individual ones cannot be aimed at other individual ones in any controlled manner. One must work with collections of beam particles and target particles, and divide out their numbers to obtain a cross section. However, distances of approach are often uniquely correlated with definite angles of scattering of beam particles relative to the incident beam direction. Thus, the necessary averaging of the distances of approach contained in a total cross section can be avoided if one instead measured angular distributions of scattered beam particles. Different scattering angles are associated with beam particles probing the strength of the collisional interaction at different particle separations, and so angular distribution studies investigate the detailed nature of molecular interactions in a way similar to angular scattering experiments in nuclear and high-energy physics.

Electron-beam scattering. The apparatus for a comparatively sophisticated electron-beam angular-scattering study is indicated schematically in **Fig. 3**. The target for the electrons is a molecular beam such as previously discussed, and is shown here as a dot in the center of the experiment. The molecular beam is coming out of the paper. No collision-induced changes in the properties of that beam are detected, although this actually can be done in coincidence with observation of electrons scattered from the electron beam.

The electrons are produced by thermionic surface emission from a hot cathode, and have a maxwellian spread in energies of order 0.1 eV. They can be electrostatically accelerated if desired, and then energy-selected with a cylindrical electrostatic analyzer consisting of two 60° electric field plates on arcs of circles about a common origin. The resolution of such analyzers can be as good as 0.01 eV, a value sometimes needed to observe sharp resonance peaks in

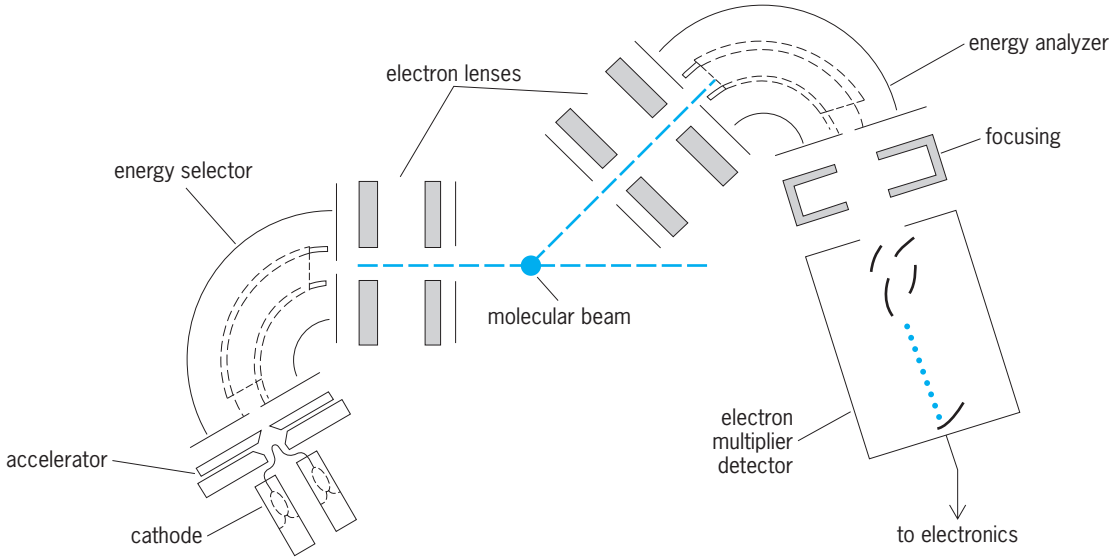

Fig. 3. Apparatus for electron-beam angular scattering study.

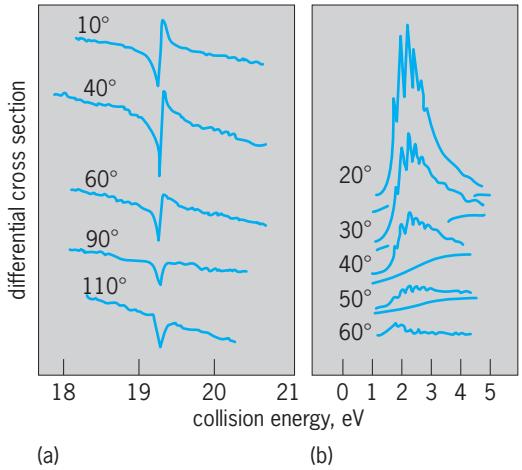

Fig. 4. Resonances observed at various scattering angles for elastic electron scattering by beams of (*a*) He and (*b*) N_2 molecules.

the energy dependence of cross sections that are associated with the excitation of long-lived compound negative-ion states. After focusing, the energy-selected electron beam is scattered by the molecular beam. Electrons scattered at various angles are selected by rotation of the remainder of the apparatus about the molecular-beam axis. Scattered electron focusing is followed by energy analysis in a second cylindrical analyzer and then by acceleration into an electron multiplier for detection by pulse counting. Thus, the experiment is divided into source, state-selector, interaction, state-analyzer, and beam-detector regions just as in the molecular-beam magnetic resonance experiment.

Figure 4 shows the collision energy dependence observed at various scattering angles for elastic electron scattering by beams of He and N_2 molecules. The change with angle in the shape of the He^- resonance near 19.3 eV makes possible a positive identification of the orbital angular momentum of the compound state. The resonances in the N_2 scattering curves are associated with the excitation of different amounts of internuclear vibration in the N_2^- ion. Searches for new states of molecules can be made by using electron scattering by ion beams. Detailed studies of resonances in slow electron scattering are used to investigate with considerable success the nature of energy exchange between the various bonds of molecules as large as benzene.

Double-molecular-beam scattering. **Figure 5** schematically shows a typical double-molecular beam scattering apparatus used in studies of chemical reactions at the molecular level. Elastic scattering can also be studied to obtain data sensitive to the molecule-molecule interaction potential. The beams from the primary and secondary ovens collide in an interaction region. The intensity of the primary beam is monitored by an on-axis detector, and the scattered secondary-beam molecules or chemical reaction products are measured with a second detector. Shown in the primary beam line before the interaction region are two devices commonly used in such experiments: a beam chopper for phase-sensitive beam detection and a velocity selector. The latter is a set of disks rotating together, each with beam-transmission holes placed at different angles about the axis of rotation. Only those beam molecules having a certain velocity pass through all the holes as the disks rotate. The beam chopper breaks the beam into pulses with known timing, and detected scattered particles are electronically checked for the same timing pattern. Thus beam-beam scattering can be detected in the presence of a poor vacuum producing very large amounts of beam scattering in the background gas.

The results of such chemical molecular-beam experiments can tell much about the mechanisms underlying chemical reactions. For instance, the alkali halide product molecule formed in $K + Br_2 \rightarrow KBr + Br$ reactions is produced primarily in small-angle or forward scattering (called a stripping reaction), whereas in $K + CH_3I$ collisions this product is seen largely in backward scattering (a rebound reaction). Large-reaction total cross sections are correlated with stripping reactions, small ones with rebound reactions; this is observed even though the sums of reaction and elastic scattering cross sections are equal in the two cases. One concludes that the total effective sizes of the targets are the same in the two cases, and that the reactive process steals scattering probability away from the elastic process where no particle rearrangement occurs.

A detailed description of the reaction mechanism can be simplified into a sequence of events occurring as the collision progresses. The sequence involves the so-called harpoon model based upon a long-range sudden jumping of the valence electron on the primary alkali atom over to the electronegative halide molecule. The jump occurs at a molecular separation R_c, where the interaction potential energies for the composite systems $K + Br_2$ and $Kr^+ + Br_2^-$ happen to become almost equal, a point where little collision kinetic energy needs to be converted into a change in potential energy. The electron's jump is sufficiently violent for the Br_2^- to be vibrationally excited, and in addition the strong attraction of the two oppositely charged ions polarizes the Br_2^-, two factors leading to the second step in the model: a dissociation of Br_2^- into $Br^- + Br$. The reaction product KBr at last is formed by the attractive interaction of

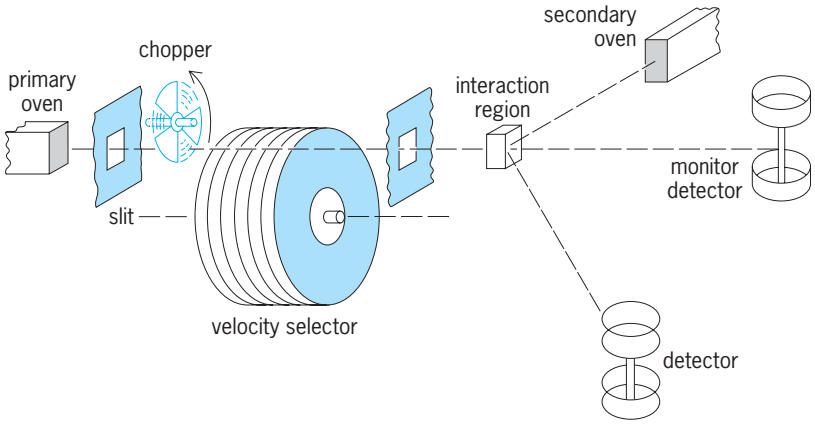

Fig. 5. A double-molecular-beam experiment on chemical reactions.

the K^+ and Br^- ions while the neutral Br atom is still nearby.

The difference between the rebound and stripping cases now can be seen as associated respectively with small and large values of the electron jump distance R_c. When reached, small values of molecular separation correspond to strong molecular interaction and thus to large scattering angles. If R_c itself is small, then reaction occurs only for large scattering angles; at the smaller angles the molecules never become close enough for the electron to jump. The total reaction cross section πR_c^2 will be small in this case since R_c is small. On the other hand, if R_c is large, then the reaction occurs at small angles and the cross section is large as well.

Although the harpoon model is not always applicable, it does characterize the ideas pursued as a result of molecular-beam scattering experiments. Development of apparatus for the scattering of molecules polarized with their electric dipole orientations fixed by external fields has been undertaken, and this should permit further progress in understanding chemical reaction mechanisms, including those of interest to biology.

Laser excitation. The development of tunable, strong laser sources of single-frequency light beams added another dimension to molecular-beam experiments. With laser radiation resonantly tuned to excite a molecule from its normal ground state to one of its infinite number of vibrationally, rotationally, and electronically excited states, the number of possible studies and applications of excited molecular beams becomes enormous. Of basic importance is the fact that excited molecules can be either highly reactive or good carriers of stored potential energy. *See* MOLECULAR STRUCTURE AND SPECTRA; NUCLEAR STRUCTURE.

James E. Bayfield

Bibliography. N. F. Ramsey, *Molecular Beams*, 1956, reprint 1990; G. Scoles et al. (eds.), *Atomic and Molecular Beam Methods*, vol. 1, 1988, vol. 2, 1992.

Molecular biology

The study of structural and functional properties of biological systems, pursued within the context of understanding the roles of the various molecules in living cells and the relationship between them. Historically, molecular biology evolved as a biological discipline with roots derived from biophysics, genetics, and biochemistry. Since its inception, a prime focus of the field has been the molecular basis of genetics, and with the demonstration in the mid-1940s that deoxyribonucleic acid (DNA) is the genetic material, emphasis has been on structure, organization, and regulation of genes. Initially, molecular biologists restricted their studies to bacterial and viral systems, largely because of the systems' genetic and biochemical simplicity. However, a series of conceptual and technological developments occurred rapidly during the late 1970s that permitted molecular biologists to approach a broad spectrum of plant and animal cells with experimental techniques. One of the major factors that rendered higher plant and animal cells amenable to molecular analysis has been the development and applications of genetic engineering. The implementation of recombinant DNA technology allowed the isolation and selective modification of specific genes, thereby reducing both their structural and functional complexity and facilitating the study of gene expression in higher cells.

Systems and Approaches

Although molecular biology initially developed as an independent discipline, the concepts and techniques used by molecular biologists have been rapidly and effectively employed to resolve numerous cellular, biological, and biochemical problems—becoming routine at both the basic and applied levels. The result is that molecular biology has emerged as a high-resolution approach to solving questions previously addressed in more traditional biochemical terms.

Intact cells and cell-free systems. While the ultimate objective of molecular biology is to understand biological function and its regulation at the molecular level, the complexity of intact higher organisms has in most cases precluded direct analysis. It has therefore been necessary to pursue the characterization and functional properties of biologically important molecules in a series of simplified systems. Since the 1960s many such investigations have been carried out using intact cells. *Escherichia coli* has been extensively examined because of its limited number of cellular functions and the corresponding restricted amount of genetic information encoded in the bacterial chromosome. Simple eukaryotic cells, such as protozoa and yeast, offer similar advantages and have also been studied. For these same reasons, bacteriophage and animal viruses have provided molecular biologists with the ability to study the structural and functional properties of molecules in intact cells.

The development of techniques for propagation of higher eukaryotic cells, including human cells, in vessels has permitted assessment of molecular function in living cells under conditions far simpler than those that exist in intact organisms. Although analysis of molecular events and regulatory mechanisms is complicated by the presence of the cell's complete set of genes and by a multitude of complex cellular functions, the ability to work with a single cell type under defined conditions has proved to be extremely advantageous. Cell culture is a widely utilized approach that has become important with advances in the understanding of specific cell growth requirements and of factors necessary for various differentiated functions. *See* TISSUE CULTURE.

Cell-free systems, in which components of cells are isolated and the properties of specific molecules and biochemical reactions are investigated in test tubes, have advanced the understanding of complex cellular events. Systems have been developed that will carry out a number of biologically essential molecular processes such as DNA synthesis, RNA synthesis, and protein synthesis. Most important, these systems allow the specific molecules

participating in the processes to be identified and characterized, frequently by addition, subtraction, or modification of the potentially relevant component. The ability to carry out biologically important reactions in cell-free systems reduces the number of variables and simplifies interpretation of experimental results. Yet, one must accept conclusions from experiments on such systems with the proviso that conditions in a test tube are somewhat artificial.

Gene mapping. The recognition of DNA as the genetic material coupled with the discovery that genes reside in chromosomes resulted in an intensive effort to map genes to specific chromosomes. Initially genes were assigned to chromosomes on the basis of correlations between modifications in cellular function, particularly biochemical defects, and the addition, loss, or modification of specific chromosomes. These early efforts at chromosomal location of genetic traits were largely restricted to lower animals and plants since they have relatively few chromosomes. Chromosome analysis was histochemical, that is, it was done by staining metaphase chromosomes spread on microscope slides and examining the preparations by light microscopy. Despite the limited resolution of this approach, chromosomal assignments of a number of genes were made. In fact, in the absence of knowledge of the biochemical or genetic basis for several human disorders such as Down syndrome, a series of chromosomal aberrations permitted identification of chromosomes encoding genes related to these diseases. Another approach which, early on, permitted assignment of genes to chromosomes was introduction of mutations by radiation or chemicals followed by a correlation between aberrations in metaphase chromosome morphology and biochemical or cellular abnormalities. *See* CHROMOSOME ABERRATION; MUTATION.

A major breakthrough in mapping genes was the development of somatic cell genetics. This is an approach in which, for example, human and hamster cells are fused, resulting in a hybrid cell initially containing the complement of human and hamster chromosomes. As the cells grow and divide in culture, the hamster chromosomes are retained while there is a progressive loss of human chromosomes. By correlating the loss of human biological or biochemical traits with the loss of specific human chromosomes, a number of human genes have been successfully mapped. *See* SOMATIC CELL GENETICS.

The development of methods for isolating genes and for determining the genetic sequences of the DNA in which the genes are encoded, led to rapid advances in gene mapping at several levels of resolution. Localization of specific genes to chromosomes is routinely carried out with cloned genes as probes. For example, isolated human genes are radioactively labeled and used to recognize and interact with an identical copy of the human gene in DNA isolated from human-hamster hybrid cells. Such analysis of a series of human-hamster hybrids, each hybrid containing limited numbers of various human chromosomes, offers an effective method for chromosome assignment. Further information about the segment of a chromosome in which a specific gene resides can be obtained by directly determining the DNA sequences of both the gene itself and the surrounding region.

Chromosome localization of specific genes has numerous applications at both the basic and clinical levels. At the basic level, knowledge of the positions of various genes provides insight into potentially functional relationships. At the clinical level, chromosome aberrations are now routinely used in prenatal diagnosis of an extensive series of human genetic disorders, and several chromosomal modifications have been linked to specific types of cancer. Knowledge of genetic defects at the molecular level has permitted the development of diagnostic procedures that in some instances, such as sickle cell anemia, are based on a single nucleotide change in the DNA. *See* GENETIC MAPPING.

Recombinant DNA. Recombinant DNA technology has provided molecular biology with an extremely powerful tool. Though introduced only in the mid-1970s, genetic engineering has had a major impact on science, medicine, technology, and society in general, providing numerous opportunities for improvement of the quality of human existence. In broad terms, applications of recombinant DNA technology can be divided into four areas—biomedical, basic biological, agricultural, and industrial. Biomedical applications include the elucidation of the cellular and molecular bases of a broad spectrum of diseases, as well as both diagnostic and therapeutic applications in clinical medicine.

In a strictly formal sense, the term recombinant DNA designates the joining or recombination of DNA segments. However, in practice, recombinant DNA has been applied to a series of molecular manipulations whereby segments of DNA are rearranged, added, deleted, or introduced into the genomes of other cells. The processes associated with modifying DNA molecules are extremely precise, permitting addition, subtraction, or alterations of specific genes or of particular regions of genes.

The importance of recombinant DNA can be readily appreciated because of the pivotal role of genetic sequences in defining the structural and functional properties of cells, tissues, and organs. Selective expression of particular genes provides the unique properties characteristic of specialized cells, for example, hemoglobin production by erythropoietic cells and myosin production by muscle cells. A progressive expression of certain genes coupled with selective repression of others is necessary for the successful orchestration of development and differentiation. Prerequisite for maintaining the physiological integrity of cells, tissues, and organs is the capacity to modulate gene expression in response to changes in cellular requirements: for example, activation of genes associated with cell proliferation in conjunction with tissue regeneration, or modifications in the expression of genes encoding metabolic enzymes in association with fluctuations in nutrient levels.

Equally important are the findings that numerous

diseases, including cancer and disorders that involve inborn errors of metabolism, are related to a variety of genetic alterations, including rearrangements or mutations in specific DNA sequences. It therefore follows that by having the capacity to manipulate the organization of genes, the capacity to correct genetic defects is not an unrealistic expectation. And while the latter goal requires technological developments and ethical justifications, the available repertoire of genetic manipulations has enabled scientists to detect several clinically important genetic defects in the fetus (such as sickle-cell anemia) and to produce biologically important molecules in cell-free systems on a scale which allows them to be used for clinical treatment (for example, insulin, growth hormone, and interferon).

Techniques. Historically, the ability to manipulate or "engineer" genetic sequences is based on several developments, as discussed in the following sections.

1. *Methods for breaking and rejoining DNA*. The precise breaking and rejoining of DNA has been made possible by the discovery of restriction endonucleases, enzymes that have the ability to recognize specific DNA sequences and to cleave the double helix precisely at these sites. Also important is the ability to join fragments of DNA together with the enzyme DNA ligase, and the techniques to determine the nucleotide sequence of genes and thereby confirm the identity and location of structural and regulatory sequences.

2. *Carriers for genetic sequences*. Bacterial plasmids, that is, circular double-stranded DNA molecules that replicate extrachromosomally, have been modified so that they can serve as efficient carriers for segments of DNA, complete genes, regions of genes, or sequences contained within several different genes. Bacteriophage (viruses which infect and replicate in bacteria) and animal viruses, such as SV40 (a simian virus), retroviruses (RNA tumor viruses), and bovine papilloma virus, have also been successfully utilized as DNA carriers. In the vernacular of a genetic engineer, these carriers are referred to as cloning vectors. Host cells in which vectors containing cloned genes can replicate range from bacteria to numerous other cells, including normal, transformed, and malignant human cells.

3. *Introduction of recombinant DNA molecules*. Genetic sequences in the form of isolated DNA fragments, or chromosomes, or of DNA molecules cloned in plasmid vectors can be introduced into host cells by a procedure referred to as transfection or DNA-mediated gene transfer—a technique that renders the cell membrane permeable by a brief treatment with calcium phosphate, thereby facilitating DNA uptake. Genes cloned in viruses can also be introduced by infection of host cells.

4. *Selection of cells containing cloned sequences*. Several approaches have been developed for identification of bacterial, plant, or animal cells containing cloned genetic sequences. Bacterial cells containing plasmids with cloned genes can be detected by selective resistance or sensitivity to antibiotics. In addition, the presence of introduced genes in bacterial, plant, or animal cells can be assayed by a procedure known as nucleic acid hybridization, in which a radioactive probe consisting of a region of a cloned gene is permitted to anneal with complementary sequences in the host cells. The recognition of the complementary sequence in the host cell is reflected by formation of a sequence-specific hybrid.

5. *Amplification*. Amplification of genetic sequences cloned in bacterial plasmids is efficiently achieved by treatment of host cells with antibiotics which suppress replication of the bacterial chromosome, yet do not interfere with replication of the plasmid with its cloned gene. Sequences cloned in bacterial or animal viruses are often amplified by virtue of the ability of the virus to replicate preferentially. *See* GENE AMPLIFICATION.

6. *Expression*. Expression of cloned human genes can be mediated by regulatory sequences derived from the natural gene, from exogenous genes, or by host cell sequences. Selection of regulatory sequences affixed to the cloned gene is based on the host cell or the regulatory signal to which expression is desired. For example, production of large quantities of a protein or RNA from a cloned human gene can often be achieved in a bacterial cell if the human gene is under the control of bacterial regulators and regulatory sequences. If a gene is introduced into a human cell to compensate for a genetic defect, however, responsiveness to normal physiological signals is an absolute necessity.

Applications. Two clinically important genes, human insulin and human growth hormone, have been cloned and introduced into bacteria under conditions where biologically active hormones can be produced. It would be difficult to underestimate the clinical relevance of cloned human insulin and growth hormone genes. Availability of biosynthetic human insulin will permit management of those patients with diabetes who cannot be successfully treated with bovine or porcine insulin and will ensure a continuous supply of insulin for everyone with diabetes. The unlimited availability of biosynthetic human growth hormone from cloned genes will effectively eliminate the shortage of this hormone and will make it possible for all children with growth hormone deficiency to be treated. *See* ADENOHYPOPHYSIS HORMONE; INSULIN.

Progress has been made in applications of recombinant DNA technology to the resolution of agricultural problems, especially for the improvement of both crops and livestock. For example, there are possibilities for nitrogen fixation by crops which presently deplete the soil of nutrients, for enhancement of photosynthetic capabilities, and for improvement of the nutritional levels of storage protein in grains. *See* BREEDING (ANIMAL); BREEDING (PLANT); GENETIC ENGINEERING.

Biophysical analysis. Understanding of the structural properties of molecules and the interaction between molecules that constitute biologically important complexes has been facilitated by biophysical analysis. For example, developments in the resolution offered by techniques such as electron

microscopy, x-ray diffraction, and neutron scattering have provided valuable insight into the structure of chromatin, the protein-DNA complex which constitutes the genome of eukaryotic cells. These techniques have also provided clues about modifications in chromatin structure that accompany functional changes. One possible application of biophysical analysis is the diagnosis of human disorders by adaptation of nuclear magnetic resonance for tissue and whole body evaluation of soft tissue tumors, blood flow, and cardiac function. *See* ELECTRON MICROSCOPE; NUCLEAR MAGNETIC RESONANCE (NMR); X-RAY DIFFRACTION.

Flow of Molecular Information

Information for all cellular activities is encoded in DNA; selective elaboration of this information is prerequisite to meeting both structural and biochemical requirements of the cell. In this regard, there are three major areas of investigations by molecular biologists: (1) the composition, structure, and organization of chromatin, the protein-DNA molecular complex in which genetic information is encoded and packaged; (2) the molecular events associated with the expression of genetically encoded information so that specific cellular biochemical requirements can be met; and (3) the molecular signals that trigger the expression of specific genes and the types of communication and feedback operative to monitor and mediate gene control.

Gene composition and organization. In both prokaryotic and eukaryotic cells, genetic information is contained in DNA molecules, but there are differences in the composition and organization of genes which are directly related to variations in expression and control. In all cases, cellular proteins are encoded in DNA sequences utilizing a triplet code of four deoxynucleotide bases (adenine, guanine, thymine, and cytosine), the building blocks of DNA. The DNA sequences also provide for major elements of secondary and possibly higher-order structural features of genes and serve as recognition sites for molecules that participate in gene structure and function. Typically, the central component of the gene is a series of nucleotides encoding a protein, and these protein coding sequences are flanked by stretches of nucleotides which influence expression. *See* GENETIC CODE.

Deoxyribonucleic acid is a long polymer with two polynucleotide chains coiled around a central axis in the configuration of a double helix. The backbone or constant component of each nucleotide chain consists of deoxyribose sugar moieties linked covalently by 3′,5′-phosphodiester bonds. The variable component of DNA which renders the molecule capable of encoding genetic information is its sequence of four nitrogenous bases: two purines (adenine and quanine) and two pyrimidines (cytosine and thymine). The bases are covalently linked to the deoxyribose moiety of the backbone by *N*-glycosidic bonds. The bases of the two nucleotide chains are linked to each other by weak hydrogen bonds, with adenine always paired with thymine and guanine with cytosine. The two polynucleotide chains are antiparallel (the 3′ end of one strand faces the 5′ end of the other) and oriented so that the purines and pyrimidines reside within the helix and the sugar-phosphate backbone is on the outside. The viability of the double-helical model for DNA has been due to its compatibility with two key molecular and biochemical properties of the molecule: the ability to replicate, and the ability to elaborate genetic information by serving as a template for the synthesis of complementary messenger ribonucleic acid (RNA) molecules which then act as template for the synthesis of genetically encoded proteins. *See* DEOXYRIBONUCLEIC ACID (DNA); NUCLEIC ACID.

The representation, organization, and packaging of genes is directly related to their expression; hence, much ongoing effort has been directed toward studying these properties of both prokaryotic and eukaryotic genes. In the bacterium *E. coli* there is a single chromosome containing a circular DNA molecule approximately 0.06 in. (1.5 mm) in length when unfolded, and with about 4×10^6 nucleotides encoding a limited number of genes (approximately 4000), most represented only once. In contrast, human cells contain 23 pairs of chromosomes, each with a linear DNA molecule. Several yards of DNA are found in the nucleus of every human somatic cell, encoding many thousands of genes, some represented only once per set of chromosomes and others represented by up to several hundred copies. *See* CHROMOSOME; GENE.

Protein-DNA complexes. Both prokaryotic and eukaryotic chromosomes exist as protein-DNA complexes. Although the representation of proteins in prokaryotic chromosomes is limited, several bacterial proteins have been extensively characterized and shown to interact with specific regions of genes, thereby regulating expression. Lac repressor is an example of a bacterial protein which interacts with a specific series of nucleotides in the gene encoding β-galactosidase and influences the extent to which β-galactoside messenger RNA (mRNA) can be synthesized.

In contrast, the chromosomes of eukaryotic cells are protein-DNA complexes with a 2:1 protein-to-DNA ratio. A large body of experimental results indicates that the chromosomal proteins of eukaryotic cells are responsible for packaging the huge DNA molecule within the confines of the cell nucleus, which is only several micrometers in diameter. Modifications in the interactions of chromosomal proteins with eukaryotic DNA have also been shown to occur in conjunction with changes in gene activity and have therefore been implicated in the control of gene expression. Accessibility of specific regions of genes to nucleases, enzymes which degrade DNA unprotected by chromosomal proteins, has provided an effective probe for establishing such structural-functional relationships. *See* DEVELOPMENTAL GENETICS.

There are five classes of basic chromosomal proteins known as histones which are enriched in arginine, lysine, and histidine residues. These histone proteins are present in an amount equivalent to DNA and have been shown to interact with one another and with DNA to form the primary unit of

eukaryotic chromatin structure, the nucleosome. The other chromosomal proteins of eukaryotic cells are designated nonhistone proteins and constitute a complex and heterogeneous array of molecules which interact with DNA and appear to participate in structural, enzymatic, and regulatory activities.

DNA rearrangements. It has become evident that genes are not necessarily static. Alterations in the representation and organization of DNA nucleotide sequences can occur in both bacteria and eukaryotic cells. For example, viruses can introduce DNA sequences into chromosomes, and a series of transpositions or movements of DNA sequences from one location to another has been documented in bacterial, plant, and animal cells. Flexibility in gene structure and organization appears to be related to biological function. There are several biological situations in which rearrangements of DNA sequences occur: (1) in conjunction with production of immunoglobulins; (2) with the development and onset of both solid tumors and human leukemias; and (3) with changes in the phenotypic properties of corn.

Extrachromosomal genes. A complete evaluation of a cell's genes and gene functions must take into account the presence of extrachromosomal DNA. Bacterial cells can carry extrachromosomal genes, such as those that confer antibiotic resistance, in the form of plasmids, which are covalently closed circular DNA molecules. Stable sets of extrachromosomal genes in eukaryotic cells are those in the mitochondria, and in chloroplasts, which encode several of the proteins required for the functional integrity of those organelles. Extrachromosomal DNA can also accumulate in eukaryotic cells in conjunction with the development of drug resistance, as has been shown for the amplification of dihydrofolate reductase genes as cells acquire tolerance to the folate antagonist methotrexate. *See* DRUG RESISTANCE; MITOCHONDRIA.

Gene expression. Expression of genetic traits both in bacteria and in eukaryotic cells requires the selective elaboration and processing of genetically encoded information. This occurs by transcription of DNA into RNA molecules that are complementary to the gene; processing of these RNA transcripts into mRNAs; and association of the mRNAs with ribosomes where the nucleotide sequences serve as a template for synthesis of a protein. Three types of RNA are required for the process: mRNA which carries the blueprints for proteins from DNA to ribosomes; ribosomal RNAs (rRNA) which serve as major components of the protein synthesis apparatus; and transfer RNAs (tRNA) which carry amino acids, the primary units of proteins, to the ribosome-associated mRNAs where they are assembled into proteins. *See* RIBONUCLEIC ACID (RNA); RIBOSOMES.

In prokaryotic cells, gene expression is rapid and considerably less complex than in eukaryotic cells. Prokaryotic mRNAs are generally encoded in contiguous DNA nucleotides, and with minimal modifications they function as templates for protein synthesis. In most situations, prokaryotic mRNAs are short-lived and some are polycistronic, that is, several proteins are encoded in a single mRNA molecule regulated by a single set of control sequences. *See* BACTERIAL GENETICS.

Eukaryotic mRNAs in general are not encoded by contiguous DNA nucleotides; rather, a precursor RNA which contains encoded segments of the mRNA (exons) separated by intervening sequences (introns) is transcribed from the DNA. The intervening sequences are excised and the protein coding sequences are joined together, or "spliced." Further aspects of eukaryotic mRNA processing that occur within the nucleus involve enzymatic modifications of the 3′ terminus [addition of a series of adenosine monophosphate (AMP) residues], the 5′ terminus (addition of a methylguanine-containing moiety referred to as a CAP structure), and various bases within the mRNA molecule (for example, methylation). During this processing the mRNA becomes complexed with proteins, forming a ribonucleoprotein particle which is exported to the cytoplasm where it can associate with ribosomes. These complicated eukaryotic mRNA processing steps require from 20 to 45 min; present understanding of the specific molecular mechanisms involved is minimal. The diversity of eukaryotic mRNAs is reflected by the proteins they template, the manner in which they are processed, and their lifetimes which range from several minutes to extremely long periods of time and vary as a function of the biological situation. *See* EXON; INTRON.

Protein synthesis in both prokaryotic and eukaryotic cells is a series of critically orchestrated processes in which the information genetically encoded in the mRNA is utilized for the step-by-step covalent assembly of amino acids into a protein molecule. The process occurs entirely in the cytoplasm of eukaryotic cells after completion of mRNA processing, while in prokaryotic cells protein synthesis can be initiated during completion of mRNA synthesis. Protein synthesis requires availability of ribosomes, tRNAs, and amino acids, along with a series of enzymes, factors, and energy sources. In many instances, after completion of protein synthesis, assembly of protein subunits or enzymatic modifications such as addition of acetate, phosphate, methyl, carbohydrate, or polyadenosine diphosphate-ribose moieties are required to yield a biologically functional protein molecule. *See* AMINO ACIDS; PROTEIN.

Regulation of gene expression. The mechanisms operative in the selective expression of specific genes and those responsible for the inactivity of others are major issues in molecular biology. Attention has been directed primarily toward identification of nucleotide sequences that are involved in regulating gene transcription and toward characterization of molecules that influence the availability of genes for transcription and those that affect the processing of RNA transcripts. In prokaryotic cells, gene expression is largely mediated at the transcriptional level, that is, the extent to which many bacterial proteins are synthesized is directly related to the extent that the gene is transcribed into mRNAs. Yet, posttranscriptional control is also an element in prokaryotic gene regulation, since the extent to which some

bacterial proteins are synthesized is governed by the ability of mRNAs to serve as templates for protein synthesis or by the efficiency of specific biochemical steps which participate in the protein synthetic process.

In eukaryotic cells, regulation of gene expression is both transcriptionally and posttranscriptionally mediated. In those situations where expression is controlled at the transcriptional level, there appear to be similarities in the mechanisms operative in prokaryotic and eukaryotic cells. However, because of the numerous steps associated with the nuclear and cytoplasmic processing of eukaryotic mRNAs there are many potential targets where posttranscriptional control can occur. Data suggest that in normal biological situations processing of mRNA transcripts within the cell nucleus and the ability of cytoplasmic mRNAs to associate with ribosomes and serve as functional templates for protein synthesis can influence the expression of specific genes. Also, the molecular basis of several human diseases probably resides in defects at various posttranscriptional levels of gene expression. For example, in at least one of the β-thalassemias, a human blood disorder in which there is a reduced representation of β-globin, an impairment exists in the splicing of the β-globin mRNA precursor molecule. *See* MOLECULAR PATHOLOGY.

An area of primary relevance to understanding regulation of gene expression, and one where knowledge is extremely limited, is the identification and characterization of cellular and molecular signals which monitor the requirements for specific proteins and which consequently activate the molecular mechanisms required to modulate the transcription, processing, or turnover of the respective mRNAs. Considerable attention has been directed toward the involvement of molecules such as steroid hormones, polypeptide hormones, and cyclic nucleotides as mediators of gene expression. The concept of autogeneous control has also been considered as a mechanism that in some situations can provide the molecular basis for regulating expression of a particular gene by using the gene product to destabilize its mRNA or to block further transcription of the gene. *See* GENETICS.

Gary S. Stein; Janet L. Stein

Bibliography. F. M. Ausubel, *Current Protocols in Molecular Biology*, 2 vols., 1988; A. Darbre, *Introduction to Practical Molecular Biology,*, 1988; E. D. DeRobertis and E. M. DeRobertis, Jr., *The Cell and Molecular Biology*, 8th ed., 1987; A. M. Lesk (ed.), *Computational Molecular Biology: Sources and Methods for Sequence Analysis*, 1989; J. D. Watson et al., *Molecular Biology of the Gene*, vol. 1, 4th ed., 1987.

Molecular chaperone

A specialized cellular protein that binds nonnative forms of other proteins and assists them to reach a functional conformation, in most cases through the expenditure of adenosine triphosphate (ATP). Originally identified by their increased abundance after heat shock, chaperone proteins in general bind to exposed hydrophobic surfaces of nonnative proteins, preventing them from forming intermolecular interactions that lead to irreversible multimolecular aggregation. Thus, the role of chaperone proteins under conditions of stress, such as heat shock, is to protect proteins by binding to incipiently misfolded conformations, preventing aggregation; then, following return of normal conditions, they allow refolding to occur, associated with protein release. *See* ADENOSINE TRIPHOSPHATE (ATP); PROTEIN.

Protein folding. Chaperones also play essential roles in folding under normal conditions (**Fig. 1**). Proteins in general contain in their primary amino acid sequences all the information necessary for proper folding. Under cellular conditions, however, "off-pathway" misfolding becomes a major competing reaction, particularly for multidomain proteins, potentially lodging them in kinetic traps (local energetic minima that are separated by energetic barriers from the global energetic minimum typical of the native state). Chaperones provide kinetic assistance to the folding process, binding such trapped, misfolded states through exposed hydrophobic surfaces and expending energy of ATP to place the protein back on a productive folding path. Thus, chaperones are not true folding catalysts; that is, they do not accelerate on-pathway folding. Rather, as kinetic assistants, they serve to smooth out the energy landscape, thus improving the overall rate and extent of productive folding in vivo.

One of the most extensively investigated roles of chaperones is to assist in the proper folding of newly translated proteins in the cytosol. This involves both cotranslational interactions with translating polypeptide chains and posttranslational interactions. Cotranslational interactions are formed with heat shock protein (Hsp) 70 in the eukaryotic cytosol and with a component that has prolyl isomerase activity, known as trigger factor, in the prokaryotic cytosol. These interactions appear to prevent premature folding of nascent polypeptide chains. In the case of Hsp70, this is likely to be mediated by localized binding of the chaperone to hydrophobic segments of extended nascent chain (**Fig. 2**). Such binding may, therefore, permit the polypeptide to complete translation before folding begins, allowing, for example, the amino terminus and carboxyl terminus to interact if required for proper folding.

Posttranslational chaperone interactions can also involve Hsp70 proteins or partner DnaJ proteins, stabilizing the polypeptide against premature folding or aggregation (**Fig. 3**). Another major class of chaperones, the chaperonin ring complexes, also provide posttranslational assistance for many proteins. In contrast to localized binding to extended segments as in the case of Hsp70, these megadalton oligomeric complexes bind collapsed, globular, nonnative conformations in the central cavity of a ring, affording multivalent hydrophobic interactions between the protein and the subunits of the surrounding chaperonin ring (**Fig. 4**). In the eukaryotic cytosol, for example, newly translated actin and

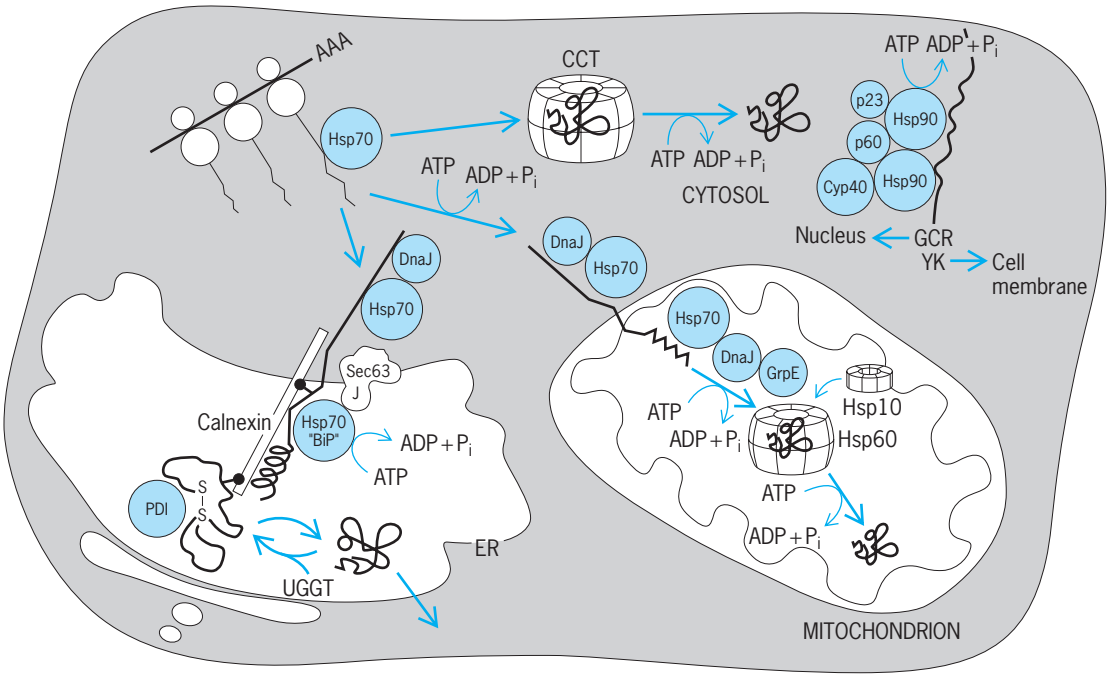

Fig. 1. Chaperone interactions with nonnative proteins in the eukaryotic cell. The interactions are with newly translated proteins in the cytosol and with proteins translocating into the endoplasmic reticulum (ER) and mitochondria, as well as between the Hsp90 complex and proteins involved in signal transduction. The nuclear compartment is not shown. Proteins are generally imported from the cytosol into the nucleus in an already-folded, native state, but chaperones such as Hsp70 proteins and CCT appear to play roles in this compartment under heat stress and in tubulin folding and assembly, respectively. Heavy lines denote nonnative protein; ball-and-stick on ER protein designates monoglucosylated N-linked glycan; PDI, protein disulfide isomerase; UGGT, UDP glucosyltransferase, which participates in the calnexin-calreticulin cycle reglycosylating nonnative N-linked glycan-containing proteins; GCR, glucocorticoid receptor; YK, tyrosine kinase; Cyp40, cyclophilin 40; other members of the Hsp90 complex are designated by molecular size.

tubulin are bound by the chaperonin CCT (containing tailless complex polypeptide 1), which is required for assisting folding into their native form. Similarly, in the bacterial cytoplasm, 5–10% of newly translated proteins interact with the essential GroEL chaperonin following translation. For all of these chaperone interactions, the subsequent binding of ATP to the chaperone allows release of the bound nonnative protein and an attempt at proper folding. In the case of the chaperonins, such release and folding occurs inside the sequestered cavity of a ring, with encapsulation provided for GroEL by association of the lidlike cochaperonin ring, GroES, in the presence of ATP. *See* CYTOPLASM.

Considering that there are often multiple chaperone family members present in the same cellular compartment (for example, both Hsp70 and GroEL reside in the bacterial cytoplasm), multiple chaperones can have significant affinity for the same nonnative protein conformation, setting up a kinetic competition. Thus, a network of chaperone interactions is present in the various cellular compartments, with nonnative states variously partitioning between chaperones that can assist their folding, as well as other components (such as proteases) that can remove irreparably damaged proteins that would otherwise clog the chaperone machinery.

Other functions. In addition to roles in protection against stress and in de novo cytosolic folding, molecular chaperones participate in the following:

1. Maintaining precursor forms of endoplasmic reticulum and mitochondrial proteins in unfolded,

CHAPERONE	TOPOLOGY OF BINDING	ACTION
Hsp100	side top	ATP-dependent disaggregation and unfolding for degradation
Hsp90	Multiprotein complex	Conformational maturation of steroid hormone receptors and signal transducing kinases
Hsp70 (Dnak)		ATP-dependent stabilization of hydrophobic regions in extended polypeptide segments
GroEL (Hsp60)	side top	ATP-dependent facilitation of folding to the native state
Small Hsps (Hsp25, and so on)		Stabilization against aggregation during heat shock
Calnexin-calreticulin	?	Folding of glucosylated proteins in the endoplasmic reticulum in cooperation with glucosyltransferase

Fig. 2. Topology of polypeptide binding and action of chaperone families. Polypeptides are shown as wavy lines, and the thickened segments indicate sites that become directly associated with the chaperone, typically hydrophobic in character. (*Adapted from B. Bukau and A. L. Horwich, Cell, 92:351–366, 1998*)

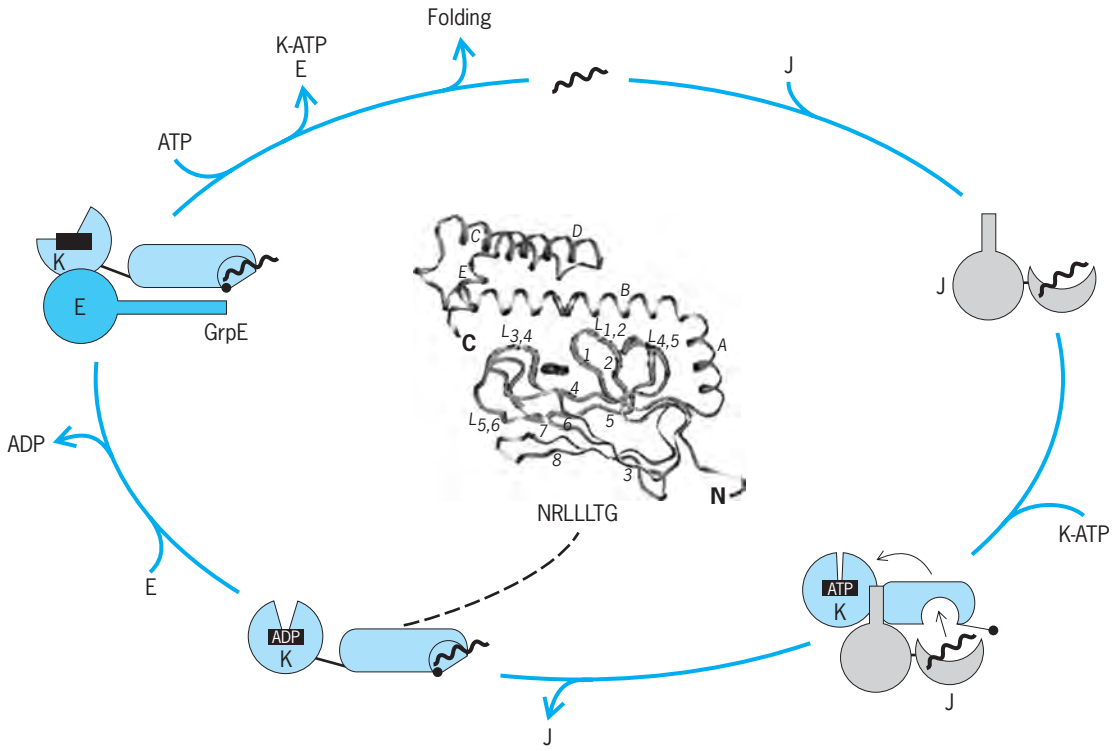

Fig. 3. ATP-directed cycles of nonnative protein binding and folding by DnaK (Hsp70)/DnaJ proteins. DnaJ protein associates with nonnative substrate protein and transfers it to the COOH-terminal peptide binding domain of DnaK. At the same time, association of DnaJ with DnaK stimulates ATP hydrolysis in the NH_2-terminal nucleotide binding domain of DnaK, which allosterically adjusts the COOH-terminal peptide binding domain to "lock in" the transferred substrate protein. (In the structural model in the center of the diagram, hydrophobic peptide, NRLLLTG, is bound to the peptide binding domain of DnaK; the extended peptide is buried in a hydrophobic pocket between loops of a β-sheet; the overlying α-helix may function as a lid.) The protein, stabilized by binding to DnaK, is then released by exchange of nucleotide, with ATP replacing ADP, mediated in the case of bacterial DnaK by an exchange factor known as GrpE, which acts to "pry open" the nucleotide binding site. ATP entry into the nucleotide binding site acts to reduce the affinity of the peptide binding domain, and the bound substrate protein is released to attempt to fold. The homologous eukaryotic Hsp70 proteins carry out similar actions, although the nature of the DnaJ and presence or absence of a GrpE component vary depending on the cellular compartment. (*Model of the reaction cycle is from B. Bukau and A. L. Horwich, Cell, 92:351–366, 1998; the NRLLLTG structure is adapted from X. Zhu et al., Science, 272:1606–1614, 1996*).

extended states in the eukaryotic cytosol, enabling the precursors to be recognized and translocated into these organelles.

2. Enabling translocation of proteins into the endoplasmic reticulum and mitochondria, by action at the trans (inside) aspect of the organellar membrane, mediated by distinct Hsp70 family members inside the organelles.

3. Assisting folding of proteins imported into mitochondria, involving the mitochondrial Hsp70/DnaJ/GrpE system, prolyl isomerases, and the Hsp60/Hsp10 chaperonin system.

4. Assisting folding of some proteins imported into the endoplasmic reticulum, involving another Hsp70 (BiP), thiol oxidoreductases, and the lectins calnexin and calreticulin, in cooperation with a glucosyltransferase (UGGT).

5. Binding near-native forms of signal transduction kinases and nuclear receptors in the cytosol, a role of complexes containing Hsp90 and partner proteins that include prolyl isomerases.

6. Functioning as ATP-driven enzymes to recognize appropriately tagged (for example, ubiquitinated) proteins, unfold them, and then bind and translocate them for degradation into coaxially associated proteolytic cylinders such as the proteasome, mediated by hexameric Hsp100 ring assemblies and the 19S proteasome cap. *See* CELL (BIOLOGY); ENDOPLASMIC RETICULUM; MITOCHONDRIA; PROTEIN DEGRADATION.

Mechanics. The mechanics of ATP-driven action of chaperone proteins are best understood for the Hsp70 and chaperonin ring systems (Figs. 3 and 4). For both systems, it is the energy of ATP binding, not hydrolysis, that drives release of bound polypeptide and initiation of folding. For example, in the case of the GroEL chaperonin, binding of ATP and GroES is associated with large rigid-body movements of the apical and intermediate domains of the bound ring, including a 60° elevation and 90° clockwise turn of the apical domains, all occurring in less than a second. These conformational changes remove the hydrophobic binding surface from the central cavity and release the polypeptide into a cavity whose wall character has become hydrophilic, favoring its folding to the native state. ATP hydrolysis then pushes the chaperonin system forward on its pathway, relaxing the high-affinity interaction between GroEL

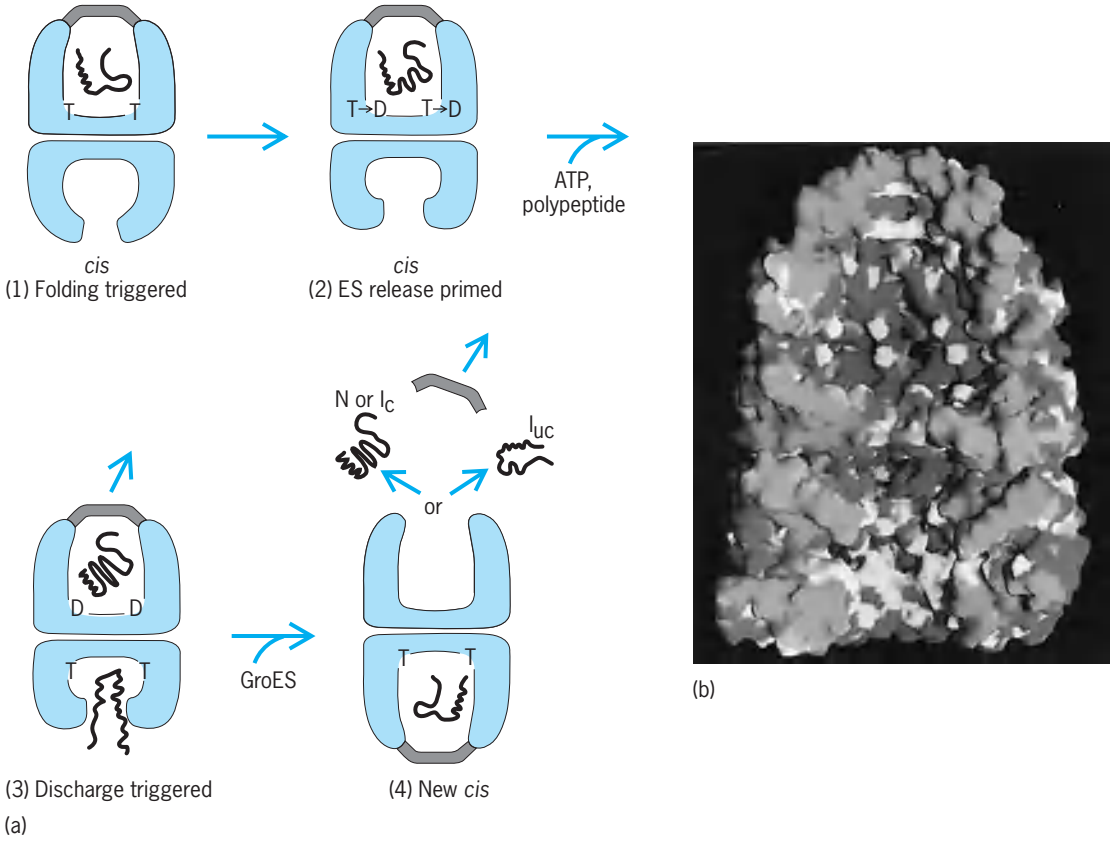

Fig. 4. GroEL/GroES system. (*a*) Nonnative substrate protein becomes bound to the open ring of a GroEL-GroES-ADP asymmetric complex (panels 2 and 3, bottom ring). In the presence of ATP, which also binds to the open ring, the GroES, folded polypeptide, and ADP ligands are rapidly discharged from the opposite ring (panels 3 and 4, top ring). GroES then binds to the polypeptide-ATP-bound ring (panel 4, bottom ring), producing a large conformational change of the ring, which removes the hydrophobic binding sites from contact with polypeptide and discharges it into the encapsulated, now hydrophilic, cavity, where folding commences. Such folding proceeds in this *cis* ATP complex (panel 4, bottom ring, and panel 1, complex inverted), which has no affinity for polypeptide or GroES on its open *trans* ring. ATP hydrolysis occurs after this longest phase of the chaperonin cycle, producing a *cis* ADP complex (panel 2), which has reduced affinity for GroES on its *cis* ring, and can bind polypeptide, ATP, and GroES on its open *trans* ring. While folding continues seamlessly in the *cis* ring of this ADP complex, the lifetime of this complex is short, ended by binding of polypeptide and ATP in the *trans* ring. Thus the chaperonin cycles back and forth in the presence of nonnative polypeptide and ATP, employing one ringful of 7 ATPs per folding cycle, using these ATPs to simultaneously nucleate a new folding-active *cis* complex while dissociating the old one. (*b*) Cutaway view of a crystallographic model of the asymmetric *cis* ADP-GroEL-GroES complex, with the hydrophobic cavity surface of the open *trans* ring, the site of capturing a nonnative protein, shown in white, and the hydrophilic character of the cavity of the folding-active *cis* ring, produced following the large-scale domain movements attendant to binding GroES, shown in dark gray. Dimensions of the complex, 184 Å height, 140 Å diameter. (*Reaction cycle model taken from A. L. Horwich et al., Proc. Nat. Acad. Sci. USA, 96:11033–11040, 1999; GroEL-GroES-ADP model adapted from Z. H. Xu et al., Nature, 388:741–750, 1997*)

and GroES and ultimately enabling the timed release of GroES and of the encapsulated, now-folded polypeptide. Such an orchestrated utilization of energy to carry out molecular movements and associated work, here the assistance of protein folding by alternation of the character of the cavity walls between hydrophobic and hydrophilic, qualifies such chaperones as molecular machines. Further study will delineate the mechanisms of the other chaperones, as well as define the precise effects of these components on protein conformation. *See* CELL (BIOLOGY). Arthur Horwich; Wayne Fenton

Bibliography. R. A. Bradshaw et al. (eds.), *Molecular Biology of Intracellular Protein Sorting and Organelle Assembly*, 1988; A. Horwich (ed.), *Advances in Protein Chemistry*, vol. 59, 2001; G. E. Schultz and R. H. Schirmer, *Principles of Protein Structure*, 1988; W. Voelter et al. (eds.), *Chemistry of Peptides and Proteins*, vol. 3, 1985.

Molecular cloud

A large and relatively dense cloud of cold gas and dust in interstellar space from which new stars are born. Molecular clouds consist primarily of molecular hydrogen (H_2) gas, have temperatures in the range 10–100 K, and contain 10^{31}–10^{36} kg of mass (for comparison the mass of the Sun is 2×10^{30} kg). Molecular clouds are among the most massive gravitationally bound objects in the Milky Way Galaxy.

Molecules. Molecular hydrogen is not directly observable under most conditions in molecular clouds. Therefore, almost all current knowledge about the

properties of molecular clouds has been deduced from observations of trace constituents, mostly simple molecules such as carbon monoxide (CO), which have strong emission lines in the centimeter-, millimeter-, and submillimeter-wavelength portions of the electromagnetic spectrum. The first molecules were detected as optical absorption lines in the spectra of bright stars lying behind molecular clouds. During the 1960s, centimeter-wavelength radio emission lines of hydroxyl (OH), water (H_2O), and ammonia (NH_3) were detected from the vicinity of bright optical nebulae such as the Orion Nebula (Messier 42). In 1970, the most important tracer of molecular gas in space, carbon monoxide, was discovered at a wavelength of 2.6 mm. *See* RADIO ASTRONOMY.

Properties. Molecular clouds are the principal sites of ongoing star formation. Therefore, they tend to be associated with young stars and star-forming regions. The nearest star-forming clouds are found in the constellations Ophiuchus, Taurus, and Perseus, at distances of 125, 140, and 300 parsecs (1 parsec = 3×10^{13} km or 2×10^{13} mi), where the nearest regions of active low- and intermediate-mass star formation are found. These cloud complexes have masses ranging from several thousand to perhaps over 10,000 times the mass of the Sun. Individual cloud cores can have masses as small as the Sun. However, most of the molecular gas in the Milky Way Galaxy is concentrated into giant clouds with masses more than 100,000 times the mass of the Sun. More than 10,000 giant molecular clouds have been found so far. The nearest giant molecular clouds are located at a distance of 460 parsecs toward the constellation Orion (see **colorplate**) where, over the last 10^7 years, they gave birth to tens of thousands of stars, including several dozen relatively rare high-mass stars. Some of the youngest massive stars light up the Great Nebula in Orion. *See* ORION NEBULA; PROTOSTAR.

Most molecular clouds have temperatures of only 10 K. Molecular clouds are orders of magnitude more dense than the general interstellar medium, with gas densities ranging from about 10 molecules per cubic centimeter on large scales to over 10^6 molecules per cubic centimeter in cloud cores. The sizes of individual clouds range from less than 0.1 parsec for small clouds and dense cores to over 100 parsecs for giant molecular clouds. In addition to star-forming molecular clouds concentrated toward the plane of the Milky Way, there are many smaller and lower-density molecular clouds visible most clearly far away from the galactic plane. These so-called high-latitude molecular clouds tend to be nearby, with the closest ones only about 50 parsecs from the Sun. While star-forming molecular clouds are bound by their own gravity, and are therefore long-lived objects with lifetimes ranging from millions to tens of millions of years, high-latitude clouds do not appear to be gravitationally bound. They may rapidly form and dissolve in the turbulent interstellar medium. Thus, these transient objects may survive as molecule bearing clouds for less than 100,000 years.

Structure. Molecular clouds have a very complex internal structure consisting of clumps and filaments of dense gas surrounded by interclump gas of much lower density. Individual clumps usually have supersonic internal motions with a velocity of several kilometers per second. The powerful outflows produced by young stars during the first 100,000 years of their existence may be a major source of these chaotic motions. Magnetic fields which thread molecular clouds may play a role in the longevity of turbulent motions and may support clouds against gravitational collapse. Clumps within a cloud frequently exhibit even larger relative motions produced by the gravitational potential of the entire cloud. Clouds near the galactic center have very large internal motion, frequently with internal velocities of more than 30 km/s (20 mi/s).

Composition. Carbon monoxide is the second most abundant molecule, after hydrogen. There is 1 carbon monoxide molecule for about every 10,000 hydrogen molecules. About 100 different chemical species have been so far identified within molecular clouds, indicating that there is a rich chemistry taking place. Over 1' of the mass of molecular clouds is in the form of interstellar dust grains that absorb starlight, making molecular clouds opaque at visible and ultraviolet wavelengths. Therefore, most nearby clouds can be seen in silhouette against the background of stars. *See* INTERSTELLAR EXTINCTION.

Molecular hydrogen is formed primarily from atomic hydrogen by chemical reactions taking place on the surfaces of dust grains. However, most other chemical species are produced by gas-phase chemical reactions involving the interactions of neutral species with ions. Energetic cosmic-ray particles moving through the molecular gas ionize a small fraction (about 1 part in 10^7) of the molecules, resulting in the formation of the highly reactive H_3^+ ion, which drives the ion-neutral reactions leading to the formation of other observed chemical species. The small abundance of ions (and electrons) in molecular clouds is sufficient to couple cosmic magnetic fields to the gas. The magnetic field, together with the gravitational force, probably regulates the rate at which gas can undergo gravitational collapse to form stars. *See* COSMOCHEMISTRY; MAGNETOHYDRODYNAMICS.

Distribution. Since the discovery of carbon monoxide, surveys of its emission lines have demonstrated that molecular clouds are widespread in the plane of the Milky Way Galaxy. The majority of clouds lie in the broad Molecular Ring encircling the galactic center with an inner radius of about 3 kiloparsecs and an ill-defined outer radius extending to beyond 20 kpc. The number of clouds is greatest near the inner boundary of the Molecular Ring with a gradually diminishing concentration toward the outer Galaxy. While molecular clouds are widely distributed around the ring within 6 kpc of the galactic center, they tend to be highly concentrated in spiral arms in the outer Mikly Way. There are relatively few clouds between the inner edge of the Molecular Ring and the immediate vicinity of the galactic center. However, the inner 0.5-kpc region of the Milky Way contains the greatest concentration of clouds in the entire Galaxy. This relatively small region near the galactic center contains nearly 10' of the molecular

gas in the Milky Way. It is thought that the stars in this part of the Galaxy are concentrated into an elongated cigar-shaped ridge called a galactic bar. Apparently, the rotation of this bar about the galactic center has swept up most of the interstellar gas lying inside the inner edge of the Molecular Ring, concentrating it within 0.5 kpc of the center. Nearly half of all of the mass of gas contained in the Milky Way's interstellar medium which fills the space between the stars is contained in molecular clouds.

Formation, evolution, and destruction. Molecular clouds are believed to survive for several times 10^7 years. The ultraviolet radiation produced by massive stars born from the clouds dissociates molecules, fully ionizes atoms and molecules, and heats the remaining gas. The resulting pressure gradients accelerate and disperse the molecular cloud, leading to its destruction.

The formation of molecular clouds is poorly understood at present. Compression of low-density atomic gas by the passage of a large-scale shock wave or a spiral arm of the Milky Way Galaxy, the tendency of atomic gas to cool by radiation, and the self-gravity of the gas may all play a role in the formation of molecular clouds.

Clouds in galaxies. Most spiral and irregular galaxies that form stars contain molecular clouds. On the other hand, elliptical galaxies contain very little molecular gas. Certain peculiar galaxies that are extremely luminous at infrared wavelengths contain very large amounts of molecular gas and are believed to be forming stars at a rate of many hundreds of solar masses per year. There is evidence that these starburst galaxies are produced by the merging of two gas-rich spiral galaxies. *See* STARBURST GALAXY.

In isolated spiral galaxies such as the Milky Way Galaxy, the total mass of molecular gas in the interstellar medium (at present about 2×10^9 times the mass of the Sun) probably decreases with time as gas is consumed by star formation. In the absence of substantial infall of fresh gas into the galaxy from the intergalactic medium, the mass of gas in the interstellar medium is halved approximately every few billion years. *See* GALAXY, EXTERNAL; INTERSTELLAR MATTER.

John Bally

Bibliography. F. Combes, Distribution of CO in the Milky Way, *Annu. Rev. Astron Astrophys.*, 29:195–237, 1991; N. J. Evans II, Physical conditions in regions of star formation, *Annu. Rev. Astron. Astrophys.*, 37:311–362, 1999; R. A. James and T. J. Millar, Molecular clouds, *Proceedings of a Conference at the Department of Astronomy, University of Manchester, March 26–30, 1990*, Cambridge University Press, 1991; W. B. Latter et al. (eds.), CO: Twenty-Five Years of Millimeter-Wave Spectroscopy, *IAU Symp.*, no. 170, Kluwer Academic, 1997; V. Mannings, A. P. Boss, and S. S. Russell (eds.), *Protostars and Planets IV*, University of Arizona Press, Tucson, 2000; G. Winnewisser and G. C. Pelz, The physics and chemistry of interstellar molecular clouds, *Proceedings of the 2d Cologne-Zermatt Symposium, Zermatt, Switzerland, September 21–24, 1993*, XV, 393, Springer-Verlag, Berlin, 1995.

Molecular electronics

The use of molecules as active electronic devices. As contemporary integrated-circuit technology approaches its limits, molecular electronic devices are being pursued not just as a potential successor to electronic systems but also as a method for introducing new functionality or interfaces.

Quantum transport. Electronic transport in nanoscale systems deviates significantly from that in macroscopic systems due to quantization effects. These quantization phenomena are the quantization of charge due to the discreteness of charge on the electron, the quantization of energy due to quantum-mechanical size effects, and the quantization of conductance due to the quantum-mechanical transmission through a system. *See* QUANTUM MECHANICS.

Charge quantization manifests itself as the impediment or inability of current to flow through a circuit due to the electrostatic repulsion of electrons. When electronic devices are small enough, the addition of even a single electron is proportionally larger than the energy supplied by the bias voltage. This effect, called Coulomb blockade, causes the electrons to shuttle through the device one at a time, analogous to a turnstile.

Energy quantization is a basic tenet of quantum mechanics, the most obvious being the electronic shell structure of atoms. Synthetic structures also can show such quantization, although the energy separation is usually very weak (small) because the effect scales inversely with the length scale. *See* ELECTRON CONFIGURATION; ENERGY LEVEL (QUANTUM MECHANICS).

Conductance quantization is the most subtle effect, only becoming relevant when the length scale of the device is typically less than a few nanometers, which is the length scale over which the electron retains quantum-mechanical coherence. However, the effect is powerful because it predicts that the conductance, G, of a device is related solely to the quantum-mechanical transmission coefficient, T, by the Landauer formula,

$$G = \frac{2e^2}{h} \sum_{n=1}^{N} T_n$$

where e is the electronic charge, h is Planck's constant, and the sum is over the N channels through which electrons can transmit. The transmission coefficient (≤ 1) is the dominant quantity that can be designed or varied over many orders of magnitude.

Depending on details of the system, any combination of these effects can be observed in nanometer-scale devices. Molecular-scale devices are especially susceptible to these effects, which are increasingly important with decreasing length scale.

Molecular transport. The most simple case to consider is a single molecule bridging two electrical contacts. A pioneering advance in the field was the realization that functional end groups can be substituted onto the molecule of choice, allowing the molecule to be placed into a junction by chemical

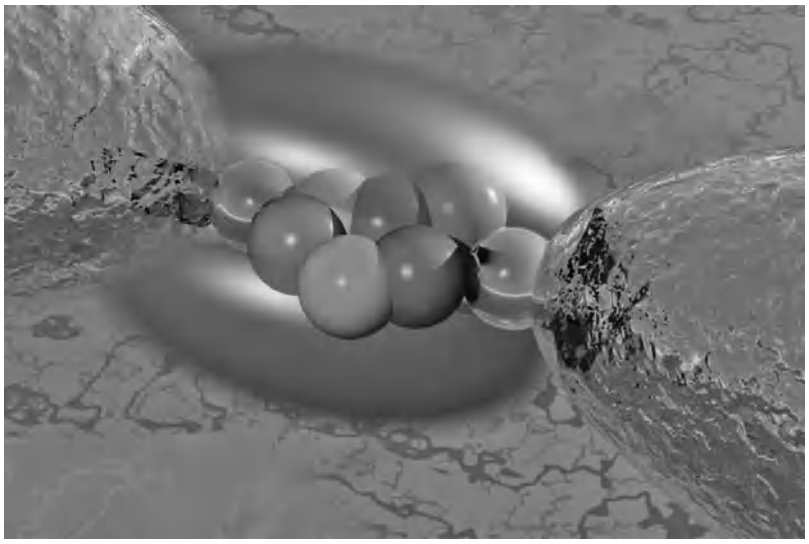

Fig. 1. Schematic of a break-junction apparatus for measuring the electrical properties of a molecular junction. (*Courtesy of Mark A. Reed*)

reaction. The most common example is the bonding of sulfur (thiol) end groups to coinage metals such as gold. An additional advance was the stable creation of nanometer-scale electrode structures bridged by a single molecule, which has been demonstrated by break junctions (**Fig. 1**). Scanning tunneling microscopy serves as a critical diagnostic and spectroscopic tool, although in a different regime as direct physical contact to the counter electrode (tip) is replaced by a vacuum gap. *See* SCANNING TUNNELING MICROSCOPE.

Molecular devices. A combination of experimental results for both single and multiple molecules in parallel (for larger devices, self-assembled molecular layers such as that in **Fig. 2** are used) shows that molecular devices can be tuned through the entire range of quantum transport regimes. In the low-transmission-coefficient regime, the conductance is shown to fit tunneling models well, with expected dependencies on molecular structure and length. If charge centers are placed in the molecule, the Coulomb blockade regime can be achieved; and this can act as the equivalent of an electronic memory. The energy quantization regime in this case reduces to the probing of molecular orbitals, and this potential tentatively has been demonstrated. *See* MOLECULAR ORBITAL THEORY; TUNNELING IN SOLIDS.

Molecular devices have two distinct differences from solid-state electronic devices. First, molecules will nearly always have strong vibronic coupling, which could have a dramatic effect on molecular transport regime or, in some cases, even switch the regime. These vibronic modes have been successfully demonstrated as a diagnostic tool, providing for the first time a "fingerprint" of the molecule in the junction, which had been problematic due to the lack of characterization capability. Second, molecules can be made exquisitely selective to a given reaction. Although this capability has yet to be fully explored and the best device structure has yet to be demonstrated, experiments have been demonstrated in which changing the charge state of the molecule affects the conductance of a nearby conventional electronic device. *See* MOLECULAR STRUCTURE AND SPECTRA.

Applications. Since the mid-1980s, the tremendous growth of molecular electronics is a reflection of the fabrication capability due to surface functionalization as well as novel device approaches that can reach to the ~1 nanometer length scale with repeatability. This progress bodes well not only to the pure intellectual understanding of molecular-scale electronics but also to its eventual system applications. The advantages of molecular device integration into complex systems are an active area of research, as contemporary integrated circuits struggle with power dissipation, materials properties, and interconnect problems. Molecular components offer the possibility of self-assembly processing and different functionality (as well as alternatives to charge-based computation), and represent the ultimate endpoint of any electronic system. Another application issue is precision, mainly of the interface to the active molecules. At the current stage of structural uncertainty, one sees fluctuations in expected device behavior due to this uncertainty. Advances in the field, as well as most other electronics approaches on this length scale, await the development of a truly planar (or 3D) atomic-scale technology with a precision better than 0.1 nm as well as structural characterization tools at this size scale. Mark A. Reed

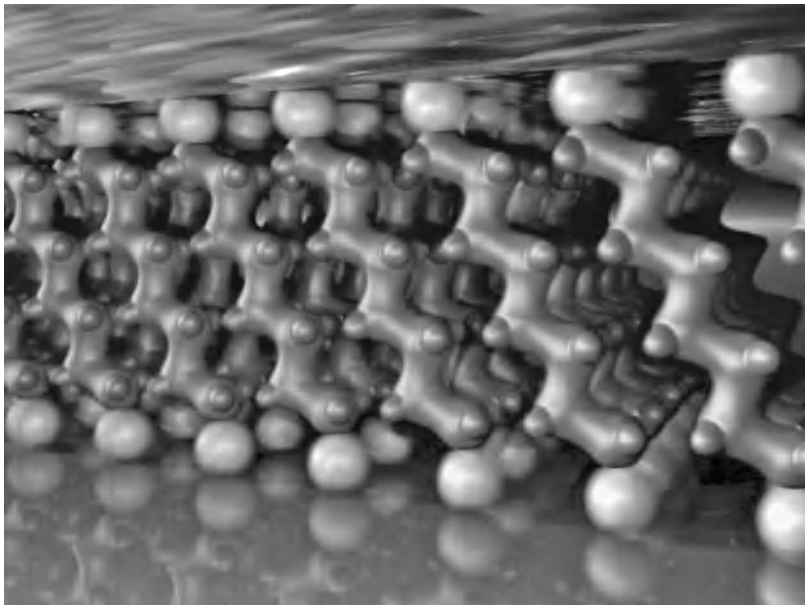

Fig. 2. Illustration of self-assembled monolayer. (*Courtesy of Mark A. Reed*)

Bibliography. G. Cuniberti, G. Fagas, and K. Richter (eds.), *Introducing Molecular Electronics*, Springer, Berlin, 2005; S. T. Pantelides et al. (eds.), *Molecular Electronics*, MRS Symp. Proc., vol. 582, Warrendale, PA, 2001; M. A. Reed and T. Lee (eds.), *Molecular Nanolectronics*, American Scientific Publishers, 2003.

Molecular isomerism

The property of compounds (isomers) which have the same molecular formula but different physical and chemical properties. The difference in properties is caused by a difference in molecular structure (that is, molecular architecture). A typical example is dimethyl ether, CH_3OCH_3, a chemically quite inert gas which condenses at −24°C (−11°F), and ethyl alcohol, CH_3CH_2OH, a liquid of substantial chemical reactivity which boils at 78°C (172°F); both compounds have the molecular formula C_2H_6O.

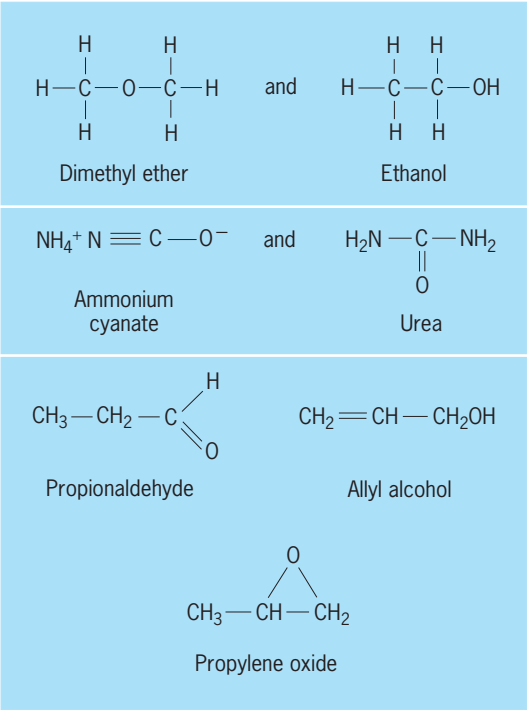

Fig. 1. Functional isomers.

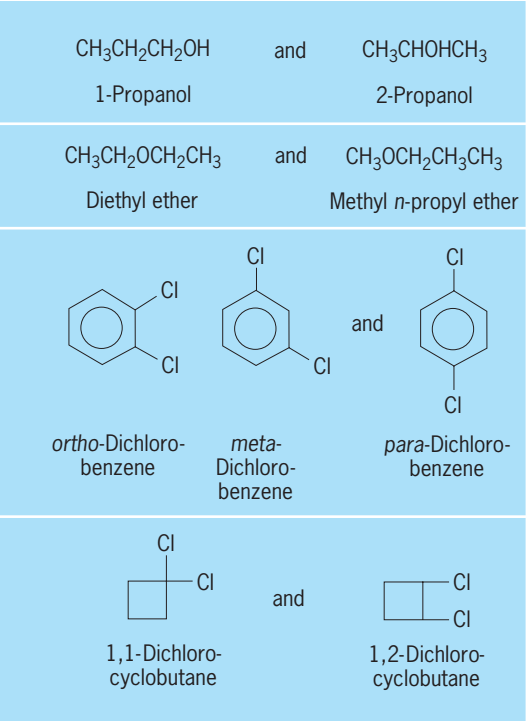

Fig. 2. Positional isomers.

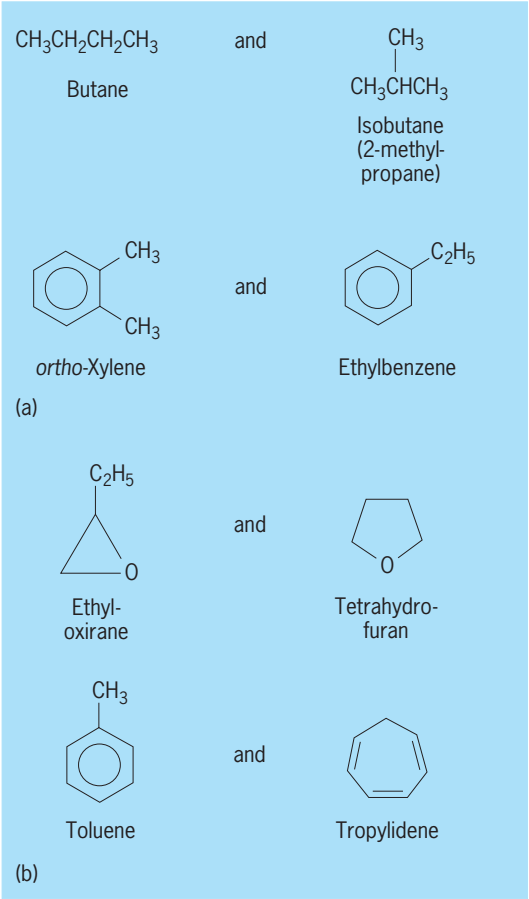

Fig. 3. Skeletal isomers. (*a*) Chain isomers. (*b*) Ring isomers.

Classification. Isomers may be classified as constitutional isomers or stereoisomers. Constitutional isomers differ in constitution or connectedness, relating to the question as to which atoms are linked to which others and how. Dimethyl ether and ethanol (**Fig. 1**) are constitutional isomers. In dimethyl ether each carbon is connected to three hydrogen atoms and the one oxygen atom; the two carbon atoms are thus equivalent. In ethyl alcohol (ethanol) one carbon is linked to three hydrogen atoms and the other carbon; the second carbon is linked to the first carbon, two hydrogens, and the oxygen atom which, in turn, is linked to the sixth hydrogen atom; the two carbon atoms are not equivalent. (Constitutional isomers can often be distinguished, and their constitution recognized, by carbon-13 nuclear magnetic resonance.) Stereoisomers, in contrast, have the same constitution but differ in the three-dimensional array of the atoms in space, called configuration. (In some cases, the difference in three-dimensional arrangement may, however, be due to rotation about single bonds, in which case it is spoken of as a difference in conformation.) *See* NUCLEAR MAGNETIC RESONANCE (NMR).

Constitutional isomers. Constitutional isomers have been subdivided into functional isomers, positional isomers, and chain isomers.

Functional isomers. Functional isomers (Fig. 1) differ in functional group, that is, the group (or groups) most material in determining chemical behavior. The ammonium cyanate–urea pair (CH_4N_2O) shown in Fig. 1 plays an important role in the history of chemistry inasmuch as the conversion of ammonium cyanate to urea by heating, effected by F. Wöhler in 1828, is considered the first example of an organic compound (urea) having been produced in the laboratory from a mineral one (ammonium cyanate). The third example, shown in Fig. 1, that of propionaldehyde (propanal), allyl alcohol (2-propen-1-ol), and propylene oxide (methyloxirane), illustrates the fact that functional isomers do not necessarily come in pairs. The three compounds all correspond to the molecular formula C_3H_6O, but the first one has an aldehyde function, the second combines a double bond with an alcohol function, and the third one has an epoxide function. Indeed, the number of possible isomers corresponding to a given molecular formula is generally remarkably large. Thus, even for the relatively simple composition $C_{10}H_{22}$, a saturated hydrocarbon with 10 carbon atoms, there are 75 isomers, and for a hydrocarbon with twice as many carbon atoms, $C_{20}H_{42}$, the number is well over 300,000.

Positional and chain isomers. Positional isomers (**Fig. 2**) have the same functional group but differ in its position along a chain or in a ring. Closely related are chain isomers which also have the same functional group or groups but differ in the shape of the carbon chain (**Fig. 3***a*); quite similar are ring isomers (Fig. 3*b*) which differ in the size of one or more rings. Ring and chain isomers together are sometimes called skeletal isomers.

Properties. It should be emphasized that these subclassifications of constitutional isomers are made for the convenience of the chemist rather than because of any fundamental importance. All constitutional isomers differ in physical and chemical properties, such as melting and boiling points, density, refractive index, and free energy, as well as in all kinds of spectral properties, such as ultraviolet, infrared and nuclear magnetic resonance spectra, and, to a lesser extent, mass spectra. If crystalline, isomers can generally be assigned their proper structure by x-ray diffraction analysis. The above differences tend to be greatest for functional isomers and more subtle for positional and skeletal isomers. However, the last statement is not universally true; thus, there is a considerable difference between the fairly reactive ethyl oxirane and the fairly inert ring isomer tetrahydrofuran shown in Fig. 3*b*. Indeed, it is often not entirely clear when isomers should be called functional and when they should be called positional or chain isomers. Thus acetone (propanone; **1**), and propionaldehyde (propanal), $CH_3CH_2CH{=}O$, may be considered positionally isometric carbonyl compounds if one chooses not to distinguish between aldehyde and ketone functions. Yet these functions are substantially different (for example, aldehydes are readily oxidized to acids, while ketones are not), so an alternative might be to consider the ketone and the aldehyde as functional isomers. A similar situation occurs with amines: *t*-bulylamine, $(CH_3)_3CNH_2$, and *n*-butylamine, $CH_3CH_2CH_2CH_2NH_2$, are clearly chain isomers, whereas diethylamine (**2**), and methylpropylamine (**3**), should be classified as positional isomers. When one compares *n*-butylamine with diethylamine, however, the situation is less clear-cut. One could argue that these are positional isomers also, but in view of the functional differences (though slight) between primary and secondary amines, they are probably better clasified as functional isomers. Such differences in classification are obviously quite tenuous.

(1)

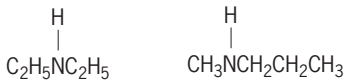

(2) **(3)**

Stereoisomers. Compounds which have not only the same molecular formula but also the same constitution (connectivity of the atoms) but which differ in the disposition of the atoms in space are called stereoisomers. Stereoisomers, in turn, are subdivided into two types: those that are mirror images of each other, called enantiomers, and those which are not mirror images, called diastereomers or diastereoisomers.

Enantiomers. These isomers are unique in that they always come in pairs (**Fig. 4**). Either a molecule is superposable with its mirror image, in which case it does not have an enantiomer, or it is not superposable with its mirror image, in which case it has one

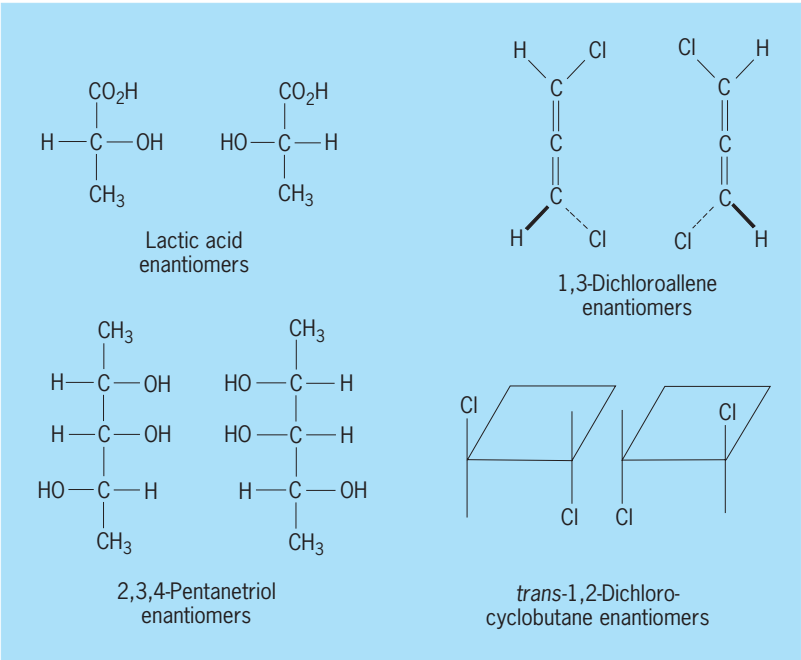

Fig. 4. Enantiomers, mirror-image isomers.

and only one enantiomer (since an object can have only one mirror image). Molecules which are not superposable with their mirror images are called chiral; those which are so superposable are called achiral. Enantiomers are much more alike than are other sets of isomers (constitutional isomers or diastereomers); thus they have the same melting point, boiling point, free energy, spectral properties, x-ray diffraction pattern, and so on. This is because their internal relationships are the same; for example, the distances between corresponding atoms are the same, much as the distances between corresponding fingers are the same in right or left hand. However, enantiomers differ in their behavior toward chiral reagents, much as a right hand and a left hand differ in their relation to a right glove, and they are also different in their behavior toward chiral physical agents, such as circularly polarized light. Thus they differ in circular dichroism and in the direction of rotation which they impart to plane polarized light (optical rotation, optical rotatory dispersion). Such differential physical properties have been called chiroptic properties.

Diastereomers. These isomers have the same constitution but different spatial arrangement and are not mirror images (**Fig. 5**). They resemble constitutional isomers in that there may be more than two isomers in a set and that their physical, energetic, and spectral properties are generally quite distinct. The example of *cis*- and *trans*-1,2-dibromoethene illustrates cis-trans isomerism in olefins. (It is recommended that the term geometrical isomers which was formerly used for this type of isomers be abandoned, just as optical isomers should no longer be used as a synonym for enantiomers.) 1,3-Dichlorocyclobutane and 1,2-dimethylcyclopropane illustrate diastereoisomerism (also of the cis-trans type) in cyclanes. The 1,3-dichlorocyclobutane exists in two achiral diastereomeric forms, whereas the 1,2-dimethylcyclopropane has a pair of (chiral) enantiomers (trans) which are diastereomeric with the (achiral) cis isomer (called meso because it is an achiral diastereomer in a set also containing chiral species). Pentane-2,3,4-triol illustrates a case with one chiral and two achiral (meso) diastereomers. The enantiomeric pair of the chiral 2,3,4-pentanetriols is shown in Fig. 4.

A set of stereoisomers containing n chiral centers will normally contain 2^n members. Since each member of the set has an enantiomer, there will be 2^{n-1} enantiomer pairs which, in relation to each other, are diastereomeric. However, when there is degeneracy, that is, when two or more of the chiral centers are equivalent (as in the 2,3,4-pentanetriols where the chiral centers at C-2 and C-4 are alike), there will be fewer isomers than the formula predicts. Thus there are only four stereoisomeric 2,3,4-pentanetriols instead of the eight (2^3) predicted by the formula (see Figs. 4 and 5). A general method for counting stereoisomers has been developed. *See* STEREOCHEMISTRY.

Limitations. Many isomers are quite stable and if they can be interconverted at all, the barrier between them is quite high (**Fig. 6**). For example, the barrier between *cis*- and *trans*-2-butene, $CH_3CH{=}CHCH_3$,

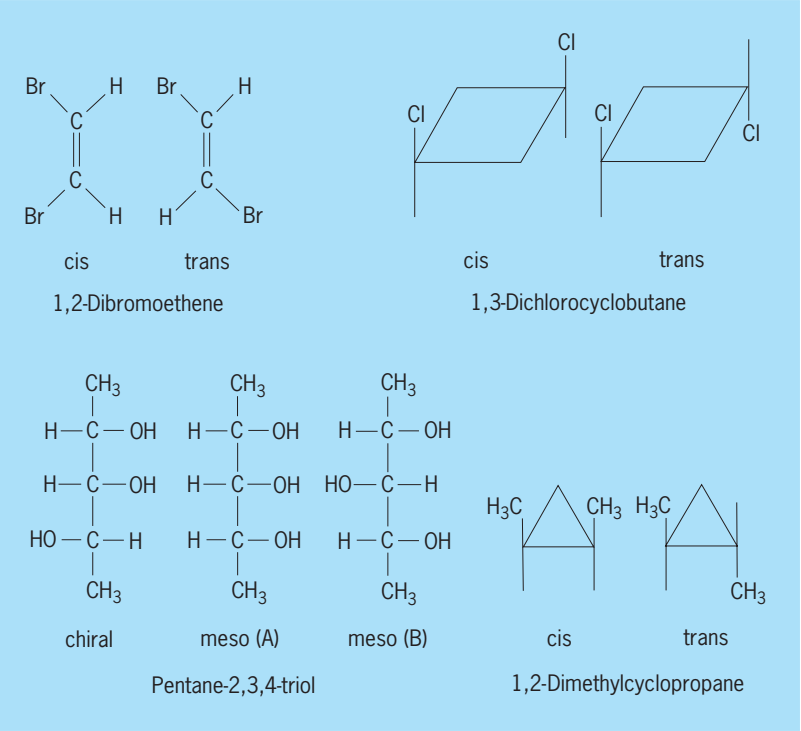

Fig. 5. Diastereomers, isomers having different spatial arrangements of the same groups of atoms.

is of the order of 65 kcal/mol (272 kilojoules/mol), and such isomers (the 1,2-dibromoethenes in Fig. 5 are another example) are spontaneously interconverted only at very high temperatures. In other cases, however, the barriers are intrinsically quite low or are readily lowered by deliberate or adventitious catalysis, frequently by acids or bases. An example relates to the keto and enol forms of ethyl acetoacetate (**Fig. 7***a*). While these isomers can be separated by fractional distillation in clean quartz vessels, they are quickly interconverted in the presence of traces of acids or bases. Such easily interconvertible isomers which differ only in the position of an atom or group (in the case of ethyl acetoacetate, a hydrogen atom which is attached to carbon in the keto form and to oxygen in the enol form) are called tautomers, and the phenomenon is called tautomerism. A closely related type of isomerism in which only bonds shift and atoms remain in place (except for changes in bond distances) is represented by the cyclooctatriene-bicyclooctadiene interconversion (Fig. 7*b*) which takes place readily at room temperature. This kind of rapid isomerization is referred to as valence bond isomerism or valence tautomerism. Two other examples of rapidly interconverting isomers are shown in **Fig. 8**. The first is chlorocyclohexane which exists in an equatorial and an axial conformation rapidly interconverted by reversal (flipping) of the cyclohexane chair. The barrier to this interconversion is about 10 kcal/mol (42 kJ/mol), which is so low that the two conformational isomers (sometimes called conformers) can be isolated only at temperatures as low as −150°C (−238°F). However, the two isomers can be seen separately in nuclear magnetic resonance (NMR) spectra

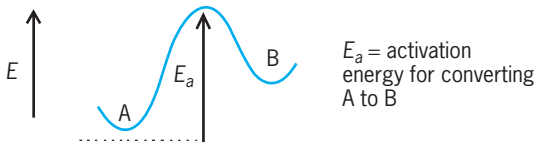

Fig. 6. Energy barriers between isomers A and B.

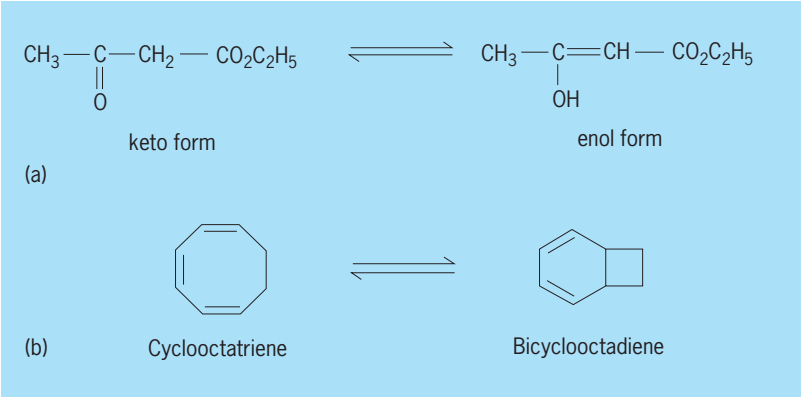

Fig. 7. Easily interconverted isomers. (*a*) Ethyl acetoacetate. (*b*) Cyclooctatriene and bicyclooctadiene.

below about −65°C or −85°F; (the exact temperature depends on the frequency of the instrument and the nucleus observed). The axial and equatorial conformations of chlorocyclohexane are thus diastereomeric. In 1,2-dichloroethane (Fig. 8), three stereoisomeric conformational minima (two enantiomeric gauche conformations and the achiral anticonformation that is diastereomeric to the two others) can be discerned, but the barrier to rotation is so low (about 3 kcal/mol or 13 kJ/mol) that isolation is out of the question, and even many physical techniques (such as electron diffraction, NMR spectroscopy, and dipole moment measurement) yield only weighted average values of the physical properties of the three conformations. However, the gauche isomers and antiisomers can be seen distinctly in the infrared spectrum of the substance and are thus clearly different species.

Ultimately there is the question of what will happen if the barrier to interconversion becomes even lower. A point must come where it is no longer possible to speak of two distinct isomeric molecules but where only a single molecule is deemed to exist. This happens when there is no longer an operational way of demonstrating the existence of two separate energy minima. It has been suggested that there are not two isomeric molecules when the barrier (Fig. 6) is lower than the product of the gas constant R on the absolute temperature T (about 0.6 kcal/mol or 2. 5 kJ/mol at room temperature), since under those circumstances the molecule traverses the barrier in a single molecular vibration, but this limit is clearly somewhat arbitrary. In any case, it is clear that the differentiation between two distinct isomeric molecules and two energy states of a single molecule is not a sharp one in those instances where the energy barrier between isomers is very low. *See* CONFORMATIONAL ANALYSIS; OPTICAL ACTIVITY; TAUTOMERISM.

Ernest L. Eliel

Bibliography. D. Barton and W. D. Ollis (eds.), *Comprehensive Organic Chemistry*, vol. 1, 1979; E. L. Eliel and S. H. Wilen, *Stereochemistry of Organic Compounds*, 1993; Z. Slanina, *Contemporary Theory of Chemical Isomerism*, 1986.

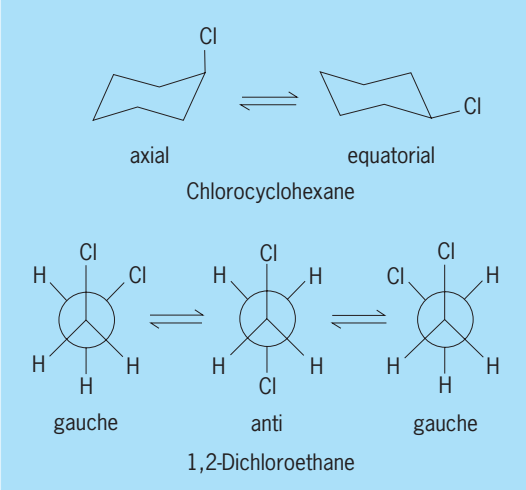

Fig. 8. Conformational isomers.

Molecular machine

A molecular device is an assemblage of a discrete number of molecular components (that is, a supramolecular structure) designed to achieve a specific function. Each molecular component performs a single act, while the entire supramolecular structure performs a more complex function, which results from the cooperation of the various molecular components. Molecular devices operate via electronic or nuclear rearrangements. Like any device, they need energy to operate and signals to communicate with the operator. The extension of the concept of a device, so common on a macromolecular level, to the molecular level is of interest not only for basic research but also for the growth of nanoscience and nanotechnology. *See* NANOTECHNOLOGY.

A molecular machine is a particular type of molecular device in which the component parts can display changes in their relative positions as a result of some external stimulus. Such molecular motions usually result in changes of some chemical or physical property of the supramolecular system, resulting in a "readout" signal that can be used to monitor the operation of the machine (**Fig. 1**). The reversibility of the movement, that is, the possibility to restore the initial situation by means of an opposite stimulus, is an essential feature of a molecular machine. Although there are a number of chemical compounds whose structure or shape can be modified by an external stimulus (for example, photoinduced cis-trans isomerization processes), the term "molecular machines" is used only for systems showing large-amplitude movements of molecular components.

Natural molecular machines. The concept of machines at the molecular level is not new. The human

body can be viewed as a very complex ensemble of molecular-level machines that power motions, repair damage, and orchestrate an inner world of sense, emotion, and thought. Among the most studied natural molecular machines are those based on proteins such as myosin and kynesin, whose motions are driven by adenosine triphosphate (ATP) hydrolysis. One of the most interesting molecular machines of the human body is ATP synthase, a molecular-level rotatory motor. In this machine, a proton flow through a membrane spins a wheellike molecular structure and the attached rodlike species. This changes the structure of catalytic sites, allowing uptake of adenosine diphosphate (ADP) and inorganic phosphate, their reaction to give ATP, and then the release of the synthesized ATP. *See* ADENOSINE TRIPHOSPHATE (ATP).

Artificial molecular machines. An artificial molecular machine performs mechanical movements analogous to those observed in artificial macroscopic machines (for example, tweezers, piston/cylinder, and rotating rings). The problem of the construction of artificial molecular-level machines was first posed by Richard P. Feynman, Nobel Laureate in Physics, in his address "There is Plenty of Room at the Bottom" to the American Physical Society in 1959: "What are the possibilities of small but movable machines?. . . . An internal combustion engine of molecular size is impossible. Other chemical reactions, liberating energy when cold, can be used instead. . . . Lubrication might not be necessary; bearings could run dry; they would not run hot because heat escapes from such a small device very rapidly. . . ." Although clever examples of a few artificial molecular machines (for example, a phototweezer for metal ions) were reported in the 1980s, substantial growth in this research field has occurred only recently, after the development of supramolecular chemistry. Supramolecular species such as pseudorotaxanes, rotaxanes, and catenanes are particularly suitable structures for the design of artificial molecular machines.

Analogously to what happens for macroscopic machines, the energy to make molecular machines work (that is, the stimulus causing the motion of the molecular components of the supramolecular structure) can be supplied as light, electrical energy, or chemical energy. In most cases, the machinelike movement involves two different, well-defined and stable states, and is accompanied by on/off switching of some chemical or physical signal [absorption and emission spectra, nuclear magnetic resonance (NMR), redox potential, or hydronium ion (H_3O^+) concentration]. For this reason, molecular machines can also be regarded as bistable devices for information processing (Fig. 1). *See* ACID AND BASE.

Pseudorotaxanes, rotaxanes, and catenanes. Rotaxanes consist of a linear, dumbbell-shaped molecular component encircled by a macrocyclic component (**Fig. 2***a*). Bulky groups (called stoppers) attached to both ends of the dumbbell-shaped molecule avoid the possibility of dethreading of the macrocyclic component. In pseudorotaxanes (Fig. 2*b*) there are no stoppers at the ends of the thread. Catenanes consist of mechanically interlocked macrocyclic rings (Fig. 2*c*). The rationale and efficient synthetic approaches for the preparation of complex supramolecular systems such as pseudorotaxanes, rotaxanes, and catenanes have been devised only recently. Such strategies usually rely on some kind of interaction (electron donor-acceptor, hydrogen bonding) which allows threading of the wirelike component into the macrocyclic one (pseudorotaxane structure), followed by blocking or cyclization reactions that lead to rotaxane and catenane structures, respectively. *See* CATENANES; MACROCYCLIC COMPOUND.

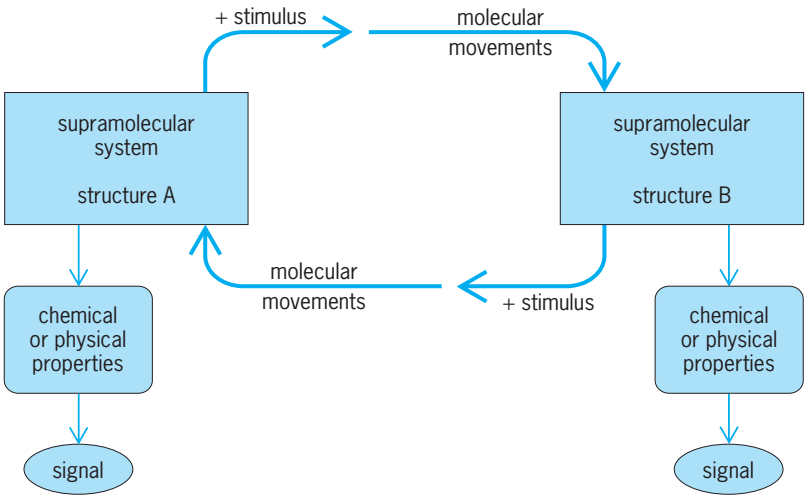

Fig. 1. Working scheme of a molecular machine.

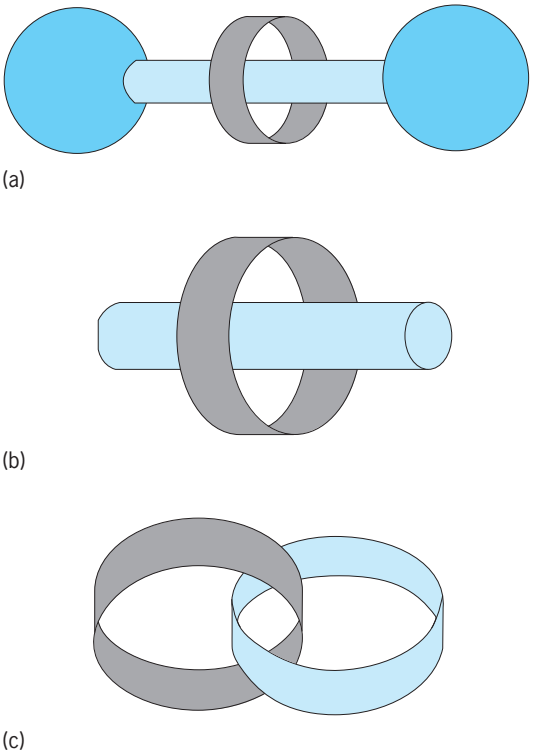

Fig. 2. Schematic representations of (*a*) rotaxanes, (*b*) pseudorotaxanes, and (*c*) catenanes.

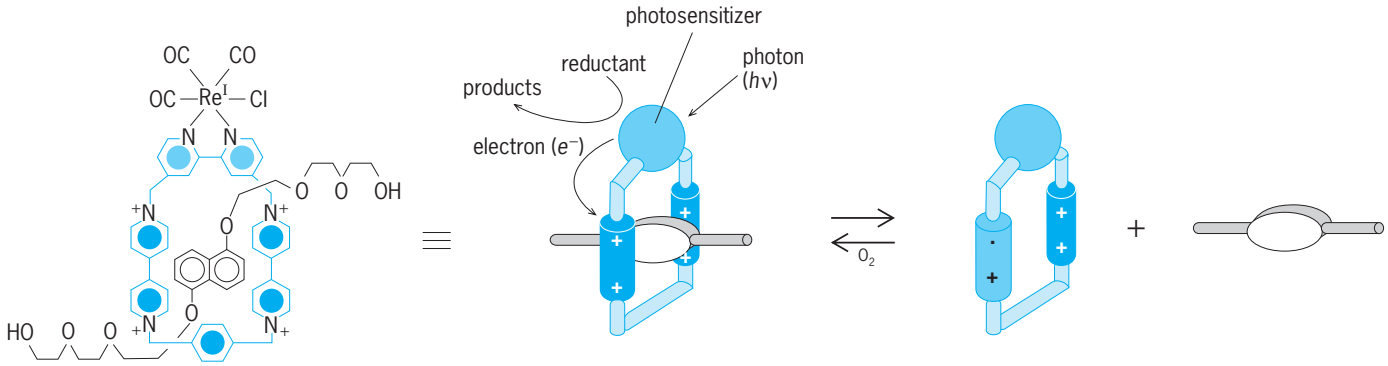

Fig. 3. Light-induced dethreading and oxygen-induced rethreading of a pseudorotaxane.

In the late 1990s several molecular machines based on pseudorotaxanes, rotaxanes, and catenanes were designed and investigated. In all cases the mechanical movements performed by the systems are very simple, so that the term "machine" is perhaps not fully appropriate. However, the extension of the concept of machine to the molecular level is a very stimulating exercise that helps the development of chemistry and underlines new aspects of the bottom-up approach to nanotechnology.

The construction of supramolecular species performing as molecular machines is not an easy task since it requires careful selection of the component units, the design of an appropriate supramolecular structure, and complex synthetic work.

The stimulus causing the motion in molecular machines can be light, electrical energy, or chemical energy. Examples of molecular machines are a pseudorotaxane stimulated by light, a rotaxane stimulated by chemical energy, and a catenane stimulated by electrical energy. The structural changes described below are always accompanied by strong changes in the absorption spectrum, fluorescence, NMR signals, and electrochemical behavior.

Light-fueled "piston/cylinder" molecular machine. In the pseudorotaxane shown in **Fig. 3**, the driving force for self-threading is the interaction between the bipyridinium electron-acceptor units of the cyclic component and the dioxynaphthalene electron-donorunit of the wirelike component. In order to cause dethreading, the electron donor-acceptor interaction must be destroyed. This can be done by reducing the electron acceptor or oxidizing the electron donor. A rhenium (Re) bipyridine complex, which is an excellent electron-transfer photosensitizer, is incorporated into the cyclic component. Light excitation of the photosensitizer is followed by electron transfer from the excited photosensitizer to one of the bipyridinium units of the cycle. The back electron-transfer reaction is prevented by scavenging the oxidized rhenium complex with a suitable sacrificial electron donor. Since a bipyridinium unit is reduced, the donor-acceptor interaction responsible for the stability of the pseudorotaxane structure is weakened so that dethreading occurs. If oxygen is allowed to enter the irradiated solution, the reduced bipyridinium unit is back-oxidized and rethreading takes place. In conclusion,

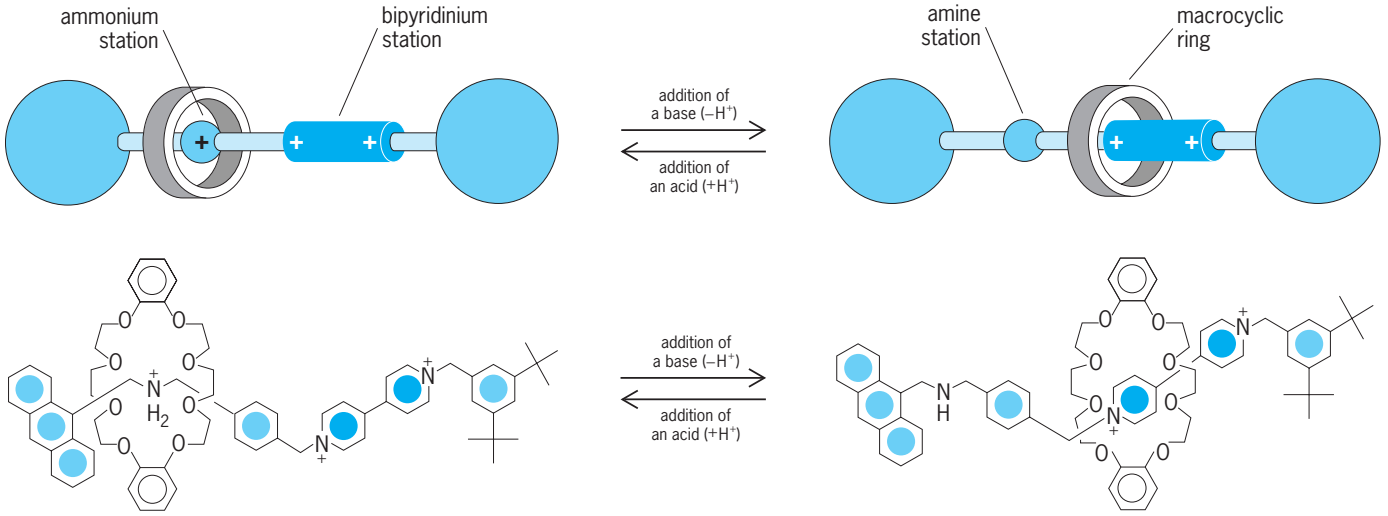

Fig. 4. Acid-base controllable molecular shuttle.

in this system a piston/cylinder-type movement can be caused by a light-fueled motor (namely, the electron-transfer photosensitizer). *See* PHOTOCHEMISTRY.

Chemically driven molecular shuttle. Rotaxanes (**Fig. 4**) have been designed with the purpose of switching the ring component between two different positions (stations) along a wire-type component. The latter component contains two different recognition sites, namely, an ammonium center and an electronacceptor bipyridinium unit. The ring component is a crown ether which incorporates two dioxybenzene units. Such a ring exhibits a strong affinity, based on hydrogen bonding, for the ammonium center and a lower affinity, based on an electron donor-acceptor interaction, for the bipyridinium unit. At the beginning, the rotaxane exists only as the translational isomer where the macrocylic component resides on the ammonium recognition site. Upon addition of a base, the ammonium center is deprotonated and the hydrogen bonding interaction with the ring is destroyed, with consequent displacement of the ring on the bipyridinium station, where stabilization is obtained by donor-acceptor interaction. Addition of acid regenerates the ammonium center, and the cyclic component moves back to its original position. *See* ACID AND BASE; NANOCHEMISTRY.

Electrochemically driven rotation. The catenane shown in **Fig. 5** consists of a symmetric macrocycle containing two electron-acceptor bipyridinium units, and an asymmetric macrocycle containing two different electron-donor units (namely, a tetrathiafulvalene and a dioxynaphthalene unit). Initially, the catenane exists only as the translation isomer, with the better electron donor unit, namely, tetrathiafulvalene, inside the ring containing the two electron-acceptor units. Selective oxidation of the tetrathiafulvalene unit suppresses its electron-donor capacity and therefore destabilizes the original translational isomer with respect to the one in which the best electron donor, which is now the dioxynaphthalene unit, resides inside the electron-acceptor cyclophane. As a consequence, a circumrotation of the asymmetric ring with respect to the symmetric one takes place, with formation of the other translational isomer. Reduction of the oxidized tetrathiafulvalene unit regenerates the original structure.

Information processing. The interest in molecular machines arises not only from their mechanical movements but also from their switching aspects. Computers are based on sets of components constructed by the top-down approach. This approach, however, is now close to its intrinsic limitations. A necessary condition for further miniaturization to increase the power of information processing and computation is the bottom-up construction of molecular-level components capable of performing the functions needed (chemical computer). The molecular machines described above operate according to a binary logic and therefore can be used for switching processes at the molecular level. It has already been shown that suitable designed machinelike systems can be employed to perform complex functions such as multipole switching, plug/socket connection of molecular wires, and XOR logic operation. *See* LOGIC.

V. Balzani

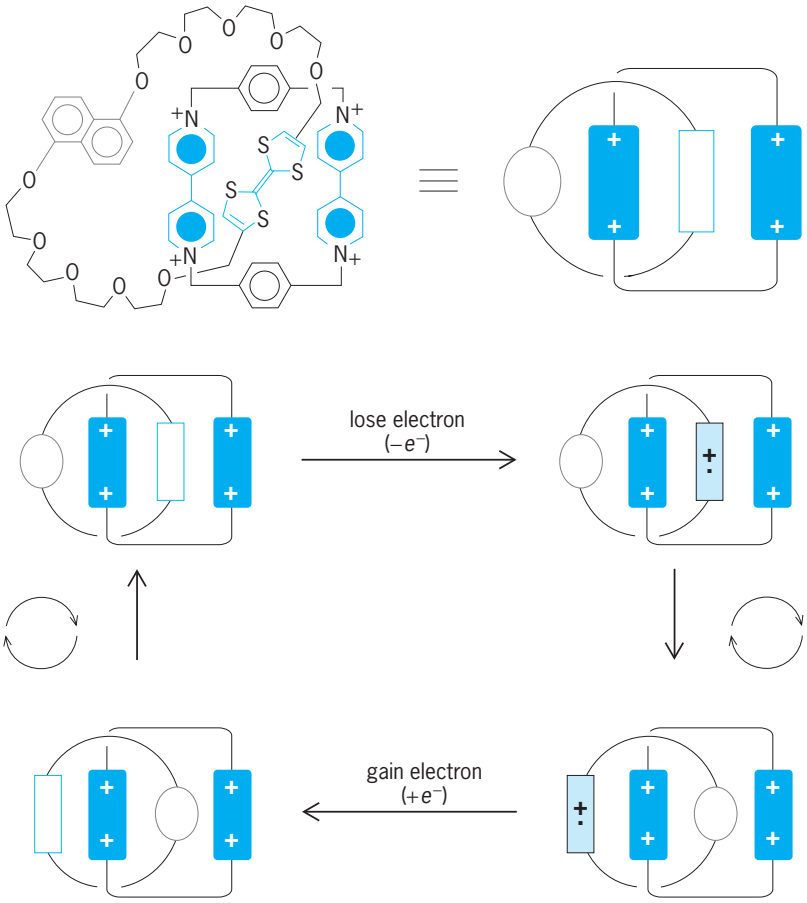

Fig. 5. Electrochemically driven rotation of a ring in a catenane.

Bibliography. V. Balzani, M. Gòmez-Lopez, and J. F. Stoddart, Molecular machines, *Acc. Chem. Res.*, 31:405–414, 1998; A. Credi et al., Logic operations at the molecular level: An XOR gate based on a molecular machine, *J. Amer. Chem. Soc.*, 119:2679–2681, 1997; J.-P. Sauvage, Transition metal-containing rotaxanes and catenanes in motion: Toward molecular machines and motors, *Acc. Chem. Res.*, 31:611–619, 1998.

Molecular mechanics

An empirical computational method that provides structural, energetic, and property information about molecules. One simple way to consider molecules is as a collection of balls (atoms) held together by springs (bonds)—the basis of many molecular model kits (**Fig. 1**). Molecular mechanics offers a way to model the behavior of matter mathematically in this manner. Mechanical spring-based theory begins with a fundamental assumption that matter consists of atoms and that the potential energy of a collection of atoms can be defined for every set of positions. The collection of atoms is treated as a mechanical system moving within this

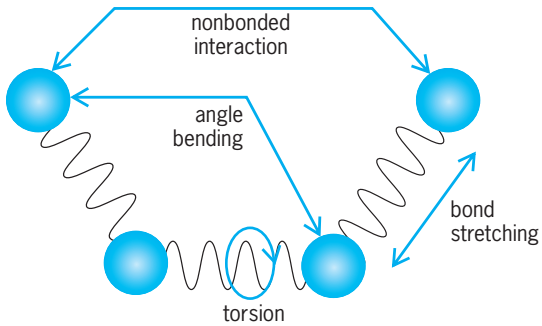

Fig. 1. Ball-and-spring model. Based on the history of physical and chemical analysis of structure and properties, it is logical to think of molecules as mechanical units made up of simple elements like balls (atoms), rods or sticks (bonds), and flexible joints (bond angles and torsion angles).

potential energy, just as a clockwork's motions are determined by its spring potentials. Molecular mechanics methods are a natural outgrowth of concepts of bonding between atoms in molecules and van der Waals forces between nonbonded atoms. In 1930, T. L. Hill proposed the first such potential energy function made up of a simple representation of this type, including van der Waals interactions together with stretching and bending deformation functions. The idea was to minimize this energy function, leading to information about the structure and steric energy in congested molecules. Subsequent studies established more elaborate formulations of the potential-energy functional form, and were used to understand increasingly more detail about molecular systems, such as the equilibrium structure, energetics, thermodynamic properties, and vibrational spectra. Further work in the 1940s verified the theory for more complex reactions. *See* CHEMICAL BONDING; ENERGY; VAN DER WAALS EQUATION.

Unlike quantum-mechanical approaches, electrons are not explicitly included in molecular mechanics calculations. This is possible due to the Born-Oppenheimer approximation, which states that the electronic and nuclear motions can be uncoupled from one another and considered separately. Molecular mechanics assumes that the electrons in a molecule find their optimum distribution, and approaches chemical problems from the standpoint of the nuclear structure. A molecule from this perspective is considered to be a collection of masses that are interacting with each other via nearly ideal harmonic forces—hence, the analogy to the ball-and-spring model (Fig. 1). Potential energy functions are used to describe the interactions between nuclei. With judicious parameterization, the electronic system is implicitly taken into account. The resulting energy function of structure, Energy $= f$ (nuclear positions), is presumed to have a minimum corresponding to the most stable equilibrium geometry. Any deviation of the model from the ideal molecular geometry will correspond to an increase in energy. *See* NUCLEAR STRUCTURE; QUANTUM MECHANICS.

Mathematical formulation. Molecular mechanics methods treat molecules as collections of particles held together by simple forces. The various types of forces are described in terms of simple individual potential functions, each summed over all respective atoms involved in that force. The total of all such contributions then constitutes the overall potential energy, or steric energy, of the molecular system. The resulting potential energy landscape of a molecule is also referred to as an empirical force field, since the derivative of the potential energy determines the forces acting on the atoms when in motion. In fact, the basic principles of molecular mechanics came out of the field of spectroscopy. Spectroscopists frequently use the term force field to mean a similar (but different) set of equations designed to reproduce or predict vibrational spectra. *See* MOLECULAR STRUCTURE AND SPECTRA; SPECTROSCOPY.

In its most simplistic representation, Equation (1)

$$E_T = E_b + E_\varphi + E_t + E_{nb} + E_{bb} + E_{el} + \cdots \qquad (1)$$

expresses the molecular mechanics force field, where E_b is the energy of bond distortion, E_φ that of angular distortion, E_t torsional angle motion, E_{nb} energy associated with atoms that are not bonded together (nonbonded), E_{bb} energy associated with explicit bonds that involve hydrogen, and E_{el} the electrostatic energy, that is, the charge interaction between atoms that are not bonded together. Each of these terms is a mathematical relationship corresponding to a different motion or interaction of the atoms in the molecule, and typically each involves parameters that are obtained from experiment. Other terms may be included in this expression, as indicated by "$+\cdots$," depending on the nature of the force field and the types of molecules and associated properties one would like to be able to use the model for. *See* ATOMIC STRUCTURE AND SPECTRA; INTERMOLECULAR FORCES.

In the simplest force fields, a handful of terms and perhaps a dozen or so parameters are necessary to describe a particular class of molecules, for example, the class of hydrocarbons. The actual mathematical form for the most basic terms in a force field are shown in Eq. 2, along with the various parameters.

$$E_{\text{total}} = \frac{1}{2}\sum_{\text{stretching}} k_b(b-b_0)^2 + \frac{1}{2}\sum_{\text{bending}} k_\theta(\theta-\theta_0)^2 + \sum_{\text{torsion}} k_r(1+[\cos(n\tau-\delta)] + \sum_i\sum_j\left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{q_iq_j}{\varepsilon r_{ij}}\right] + \cdots \qquad (2)$$

Figure 2 illustrates these basic terms in a manner that indicates the motion or interaction implied. The first term in Eq. (2), the stretching energy, is based on Hooke's law, which states that the energy change that results from movement away from some reference state can be well represented by a parabola function—the so-called harmonic model of the energy. In this way, the stretching-potential function estimates the increase in energy associated with either stretching or compressing a bond length, l, away

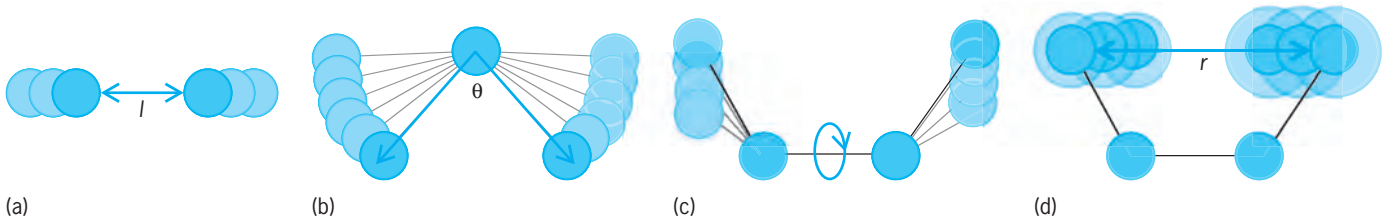

Fig. 2. Ball-and-spring representations of (*a*) bond stretching, (*b*) angle bending, (*c*) torsion, and (*d*) nonbonded interaction.

from resting position, l_0 (Fig. 2*a*). The k_s parameter in this term corresponds to the stiffness of the spring (bond), or the amount of force that would be required to restore the bond to the resting state if distorted. Both l_0 and k_o parameters are assigned from experimental data for each pair of bonded atoms and each possible atom type. The expression compares the current length of a bond (l) to its resting length (l_0), squares that difference, and then multiplies by 1/2 of the restoring force constant, k_o. Then, the summation sign indicates that each bond stretch deformation energy should be summed together for all bonds in the molecule. *See* BOND ANGLE AND DISTANCE; HOOKE'S LAW.

Similarly, the potential function for angle bending is expressed as a Hooke's law potential corresponding to an angle compression and extension away from the reference state with angle θ_0 (Fig. 2*b*). The force constant, k_θ, represents the restoring force for the angle bend. The dihedral angle corresponds to an intramolecular rotation involving four bonded atoms, and is represented by a periodic function. The torsional energy involves several parameters, k_τ, and governs the amplitude of the motion; n is the periodicity of the motion, τ is the degree of the torsional motion, and δ is the phase (Fig. 2*c*). The interaction energy for atoms that are not bonded together (Fig. 2*d*) includes two types of interactions that to some extent provide information regarding the electrons in the molecule, and these are summed over all possible atom-atom pairs in the molecule. The first type describes the van der Waals attractions and repulsions [Eq. (2)] in the molecule, that is, how the nuclei are attracted and repelled by one another. In Fig. 2*d*, a Lennard-Jones potential function is used to best model this interaction. The second type of interaction is called the Coulomb energy and provides an energy associated with the charge–charge interaction on two different atoms. In this expression, the atomic charges, q_i and q_j, are parameters, ε is the macroscopic dielectric constant shielding the interaction due to the environment, and r_{ij} is the distance between the interacting pair of atoms.

The relative order of magnitude of the terms in Eq. (2) is roughly 100:10:1 for bond distortion:angle distortion:nonbonded interactions. The energy associated with deforming a bond is always much larger (60–200 kcal/mol) than that for angle deformation (6–20 kcal/mol), which is again larger than torsional angle deformation (1–5 kcal/mol). The nonbonded energy is expected to contribute overall the least amount of energy, around 2–3 kcal/mol. In general, even a simple force field can provide a reasonable approximation to molecular structures and relative energetic differences provided the parameterization is optimal. In any case, absolute values obtained from such an expression (no matter how many terms) will likely be inferior to actual values measured experimentally.

The description of a molecule in terms of harmonic interactions is only a first approximation. As insights into molecular behavior are gained, and as more and better experimental and theoretical data become available and used, increasingly more sophisticated equations are used to reproduce molecular behavior. As such, there are many variations in the functional forms of force fields, typically falling into general classification types. To better reproduce the available experimental data, further optimization of the parameters may be necessary through modification of the equations that make up the force field. Because the resulting molecular mechanics force field is fit with parameters from experiment and then subsequently is tested against a large class of molecules for reproducibility, the individual parameters become highly interdependent, and the individual energy terms ultimately lose their meaning as individual motion components. Rather, the equation as a whole becomes a representation of the potential energy governing the structure and property of molecules within a particular class—a phenomenon known as transferability of the force field.

Application. Knowledge of the structure and properties of a series of molecules or understanding of very large macromolecular structures, such as enzymes and proteins, are problems particularly well suited for study using molecular mechanics. In fact, the mechanical molecular model formulation was developed out of a need to describe molecular structure and properties at a time when relatively little computer power was available, requiring the resulting method to be as practical as possible. The application to molecular systems with thousands of atoms, even with the advent of computational power, still mandates a more practical method, such as offered by molecular mechanics force-field methods. A variety of mathematical and computational approaches have been used over the years. This continues to be an area of active research.

The resulting equations for the total energy [Eq. (2)] are usually combined with algorithms for locating molecular structures, where the energy is at a minimum value over all possible conformations of

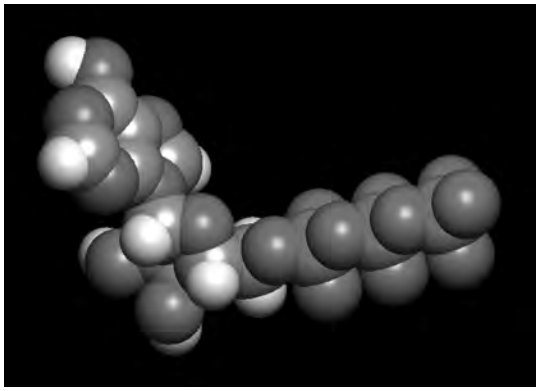

Fig. 3. Small organic molecule.

the molecule. Because this means that many possible conformations of the molecule must be considered, these are iterative procedures. One important consideration is the computational time involved in carrying out a calculation of a very large molecule using the force-field equation. The highest computational cost has to do with the terms in the equation that must sum over all atoms that are nonbonded. In studies involving thousands of atoms, any term involving all atom–atom interactions results in a very large number of evaluations that can ultimately add enormous overhead to the overall computation and, if not programmed efficiently, can bottleneck large computational investigations.

The computer algorithms that carry out the evaluations of the force field and find the minimum energy range from the most simplistic mathematical method (such as steepest descents or conjugate gradients) to sophisticated methods involving higher-order derivatives of the potential energy function (for example, the Newton-Raphson predictor-corrector method) and even grid-based methods. Often the computational procedure invokes a combination of methods to carry out the most practical and accurate method for determining the structure and properties of a molecular system. Many computer programs have been constructed for solving the force-field equations for specialized classes of molecules, and are available to the general research community. *See* COMPUTATIONAL CHEMISTRY.

Promises and perils. Many important and fundamental questions in science can be better understood through investigations using mathematical models. The general range of application of molecular mechanics includes biological or materials-related macromolecules having thousands of atoms, or even a large number of small molecules (**Fig. 3**) within a specialized class, such as organic hydrocarbons, organometallic compounds, or inorganic molecules. Computations involving molecular mechanics are mostly restricted to studies involving no bond-making or bond-breaking processes. This is particularly true if the force-field equation involves functions that well represent molecules only around their resting position, and not when they are stretched to a point of significant change in structure and reactivity.

In many investigations, molecular mechanics is used to study molecules in isolation, that is, studies performed in the gas (vacuum) phase. In biological investigations, it may be necessary to consider the effect that the environment has on the structure and how it might react. The environment is generally treated in one of two ways: an explicit or an implicit representation of the environment. In the explicit model the solvent molecules are actually present and taken into account in the evaluation of the total energy, whereas in the implicit approach the solvent is treated as a continuum governed by a function of the dielectric constant. Likewise, it could be important to consider the salt content of the system or other materials-related effects.

Because molecular mechanics–based methods tend to be computationally very efficient, application is practical for conformational analysis studies to predict the most likely conformations for highly flexible molecules (**Fig. 4**). In structural biology, this is quite important in the protein-folding problem, the goal of which is to predict the three-dimensional structure of a protein from the linear sequence of amino acids. Another wide use of molecular mechanics is in studies involving a small molecule binding into an active site, called docking studies. Estimating correct three-dimensional atomic structures of complexes between proteins and ligands is an important component of the drug-design process in the pharmaceutical industry. Basic aspects of ligand–protein interactions, categorized under the general term molecular recognition, are concerned with the specificity as well as the stability of the ligand to bind into a protein. *See* MOLECULAR RECOGNITION.

Fig. 4. Large flexible biological molecule.

Structural studies performed using the molecular mechanics model takes into account only the potential energy and not the possibility of any kinetic energy, or motion of the structure. In fact, all molecules contain potential energy and kinetic energy, and the atoms of a molecule are constantly undergoing small movement about their resting position. Molecular dynamics is also based on the empirical force field and basic assumptions described above for molecular mechanics, but allows the study of the movement of atoms over time, thereby providing a more realistic understanding of molecular structure and reactivity. In general, the molecular mechanics force-field method is largely applied for basic understanding of energetic and conformational aspects of molecules using strategies such as structure refinement, ligand docking simulations, molecular dynamics simulation, and/or Monte Carlo statistical simulations (similar to molecular dynamics, but based on statistics), all of which rely on the force field that was discussed. An outcome of the relatively long history of all such studies is the ability to construct quantitative structure–activity relationships for general classes of molecules, providing information to support our intuition about the behavior of molecules in terms of their general structure and property features. *See* CHEMICAL DYNAMICS; CHEMOMETRICS; MONTE CARLO METHOD.

Kim K. Baldridge; Celine Amoreira

Bibliography. U. Burkert and N. L. Allinger, *Molecular Mechanics*, ACS Monogr. 177, 1982; D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2002; T. L. Hill, Steric effects, *J. Chem. Phys.*, 14:465, 1946; A. R. Leach, *Molecular Modeling: Principles and Applications*, 1996; A. Hinchliffe, *Molecular Modeling for Beginners*, 2003; K. B. Lipkowitz and D. B. Boyd (eds.), *Reviews in Computational Chemistry*, vol. 2, 1991; T. Schlick, *Molecular Modeling and Simulation: An Interdisciplinary Guide*, 2002.

Molecular orbital theory

A quantum-mechanical model concerned with the description of the discrete energy levels associated with electrons in molecules. One useful way to generate such levels is to assume that the molecular orbital wave function (ψ_j) may be written as a simple weighted sum of the constituent atomic orbitals (χ_i) [Eq. (1)]; this is called the linear combination of

$$\psi_j = \Sigma c_{ij}\chi_i \qquad (1)$$

atomic orbitals approximation. The c_{ij} coefficients may be determined numerically by substitution of Eq. (1) into the Schrödinger equation and application of the variational theorem. The theorem states that an approximate wave function will always be an upper bound to the true energy; thus minimization of the energy of the system given by the wave function of Eq. (1) will provide the best values of c_{ij}. Once the wave function is known, its associated energy may be calculated. The energies of the occupied orbitals in molecules may be probed by using photoelectron spectroscopy, which gives a good check on the accuracy of the theory. There are some simple concepts that contribute to a qualitative understanding of these molecular orbital energy levels and hence an insight into chemical bonding in molecules. They may be illustrated with reference to the hydrogen molecule. *See* CHEMICAL BONDING; ELECTRON SPECTROSCOPY; QUANTUM MECHANICS; SCHRÖDINGER'S WAVE EQUATION.

Hydrogen molecule. First, the basis orbitals (χ_i) used in the expansion of Eq. (1) can usefully be restricted to include the valence orbitals only. For molecular hydrogen (H_2) the 1*s* orbitals on the two hydrogen atoms are then the only two orbitals to be included. Second, since hydrogen atoms are chemically identical, any observable characteristic whose value might be computed with Eq. (1) must be the same for both atoms. This leads to the requirement that $c^2_{1j} = c^2_{2j}$, where the labels 1,2 refer to hydrogen atoms 1 and 2. As a consequence $c_{1j} = \pm c_{2j}$.

When the signs of the two coefficients are the same, the two hydrogen orbitals are mixed in phase; when they are different, the two hydrogen orbitals are mixed out of phase. **Figure 1** shows how the wave function ψ_j and its square (giving the electron probability distribution) may be constructed for those two cases. When the atomic orbitals are mixed in phase, then electron density is built up between the two hydrogen nuclei and the potential energy of the nuclei and electrons is lowered. In fact, a reduction of kinetic energy also occurs. An electron lying in the molecular orbital corresponding to Fig. 1*a* is then of lower energy than an electron associated with an isolated hydrogen 1*s* orbital. It is called a

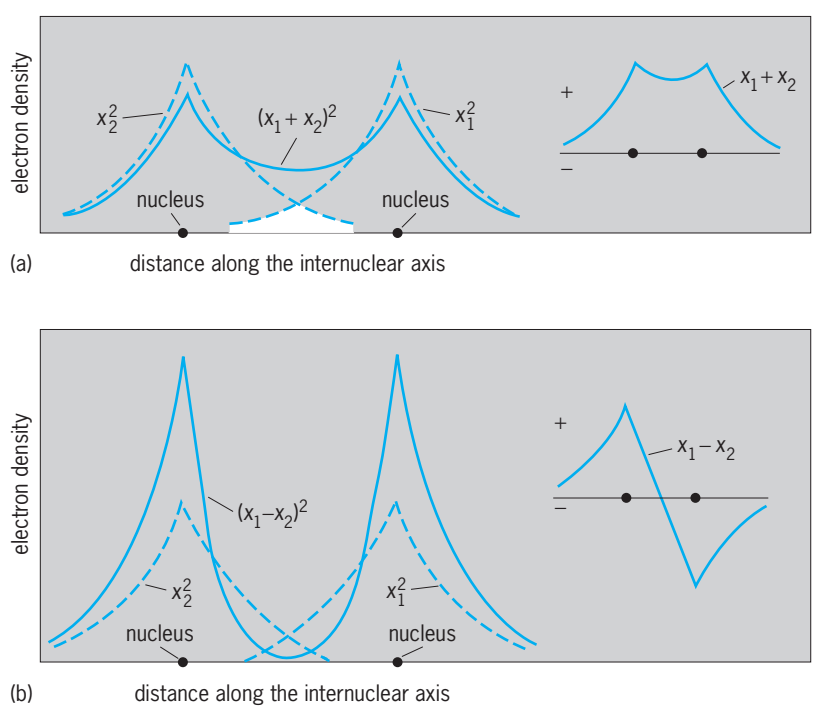

Fig. 1. In-phase and out-of-phase addition of atomic hydrogen wave functions (x_1, x_2) to give (*a*) bonding and (*b*) antibonding molecular orbitals.

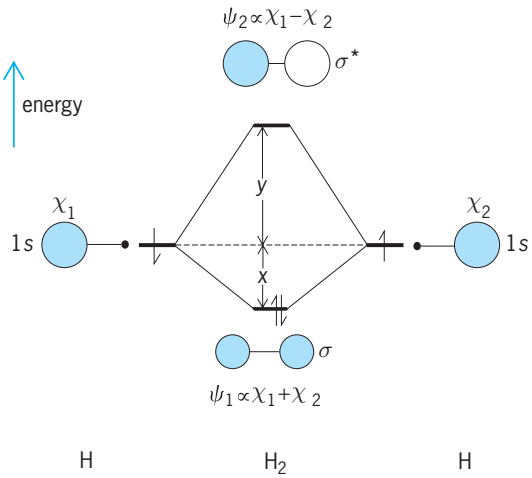

Fig. 2. Molecular orbital diagram of H_2.

bonding orbital. The increase in electron density between the two nuclei is the electronic "glue" holding the nuclei together. When the atomic orbitals are mixed out of phase, the opposite behavior occurs (Fig. 1*b*). Electron density is removed from the region between the two nuclei, resulting in an increase of both potential and kinetic energy of the electrons. An electron lying in such a molecular orbital would experience an energetic destabilization relative to an electron associated with an isolated hydrogen 1*s* orbital.

Such a molecular orbital is called an antibonding orbital. **Figure 2** shows how this information may be collected together on a molecular orbital diagram. The shading convention of the orbitals has been adopted to indicate the in- and out-of-phase mixing of the basis orbitals. Just as the energy levels of atoms are filled in an Aufbau process, so the orbitals of the molecule may be analogously filled up with electrons, each level accommodating two electrons of opposite spin. In H_2 there are two electrons to be accounted for. They lie in the bonding orbital, and the stabilization energy relative to two isolated hydrogen atoms (the bond energy) is $2x$. Antibonding orbitals are invariably destabilized more than their bonding counterparts are stabilized. This is shown in Fig. 2 by making $y > x$. With four electrons to be accommodated in this collection of orbitals (this would correspond to the hypothetical case of the He_2 molecule), one electron pair resides in the bonding orbital and one pair in the antibonding orbital. Since $y > x$, this molecule is less stable relative to two isolated helium atoms, and as a result the molecule does not exist as a stable entity. He_2^+, however, with only three electrons is known.

The size of the interaction energy associated with two atomic orbitals (x in Fig. 2) is controlled by the extent of their spatial overlap. This overlap integral is proportional to the shaded region of Fig. 1*a*, and its magnitude clearly depends upon the internuclear separation. The equilibrium bond length in the hydrogen molecule (and indeed in all molecules) is then a balance between the attractive forces associated with bonding orbital formation and the electrostatic repulsion between the nuclei. Such a molecular parameter is amenable to numerical calculation.

The description of the bonding in the H_2 molecule using this model is one where two electrons occupy a bonding orbital and give rise to a simple two-center–two-electron bond traditionally written as H—H. Since the electron density associated with the bonding orbital is cylindrically and symmetrically located about the H—H axis, this bond is called a σ bond. In Fig. 2 the bonding orbital is labeled with a σ and the corresponding antibonding orbital with a σ^*.

First-row diatomics. Ideas similar to those above are readily extended to diatomic molecules from the first row of the periodic table, such as N_2 and O_2, where the valence orbitals to be considered are the one 2*s* and the three 2*p* orbitals of the atoms. The 2*s* orbitals lie deeper in energy than the triply degenerate 2*p* orbitals. As shown in **Fig. 3**, the atomic 2*s* orbitals form bonding and antibonding orbitals (labeled $s\sigma$ and $s\sigma^*$) just as in the case of elemental hydrogen described above, but the behavior of the 2*p* orbitals is a little different. Here there are three possible types of interaction between the *p* orbitals on one center and those on the other. The end-on overlap of two *p* orbitals gives rise to a σ interaction (**Fig. 4***a*), and the sideways overlap of two *p* orbitals gives rise to a π interaction (Fig. 4*b*). The interaction in Fig. 4*c* can be ignored since the overlap between the two orbitals in this orientation can be seen to be identically zero. The result shown in Fig. 3 is a σ-bonding orbital and a σ-antibonding orbital (labeled

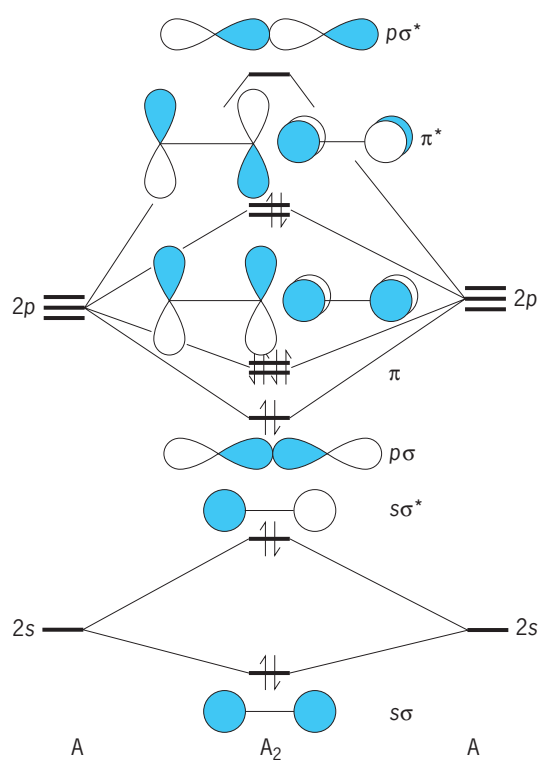

Fig. 3. Molecular orbital diagram for a first-row diatomic molecule (A_2). The electronic configuration corresponds to that of O_2.

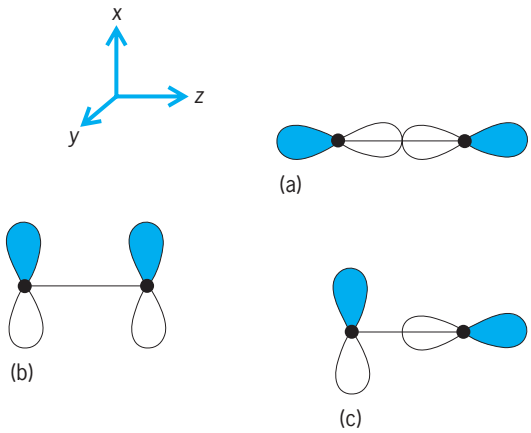

Fig. 4. Possible orientations of the *p* orbitals on adjacent atomic centers. (*a*) End-on overlap. (*b*) Sideways overlap. (*c*) Zero overlap.

$p\sigma$ and $p\sigma^*$), and a pair of π-bonding and a pair of π-antibonding orbitals (labeled π and π^*). The larger interaction energy associated with $p\sigma$ compared to $p\pi$ is due to the larger σ overlap compared to π overlap in Fig. 4.

Filling these orbitals with electrons allows comment on the stability of the resulting diatomics. The molecule Li_2 $(s\sigma)^2$ is known and, like H_2, may be written as Li—Li to emphasize the single, two-center, two-electron bond between the nuclei. The molecule Be_2 which would have the configuration $(s\sigma)^2(s\sigma^*)^2$ is unknown since, just as in He_2, $s\sigma^*$ is destabilized more than $s\sigma$ is stabilized relative to an atomic $2s$ level. If the molecular orbital bond order is written as expression (2), then the bond order in Li_2 is one but the bond order in Be_2 is zero.

$$\begin{matrix}\text{Molecular}\\ \text{bond order}\end{matrix} = \begin{pmatrix}\text{number of bonding}\\ \text{electron pairs}\end{pmatrix} - \begin{pmatrix}\text{number of antibonding}\\ \text{electron pairs}\end{pmatrix} \qquad (2)$$

By filling up the molecular orbital levels derived from the $2p$ orbitals, the bond order associated with the other diatomics may be generated: $B_2(1)$, $C_2(2)$, $N_2(3)$, $O_2(2)$, $F_2(1)$, and $Ne_2(0)$. All of these species are known except Ne_2, which is predicted, like He_2 and Be_2, to have a zero bond order and therefore not to exist as a stable molecule. The molecular orbital bond orders for the three best-known diatomics are consistent with their traditional formulation as N≡N, O=O, and F—F. N_2, for example, would be described as having one σ and two π bonds. Figure 3 shows the electron occupancy for O_2. With the configuration $(s\sigma)^2(s\sigma^*)^2(p\sigma)^2-(\pi)^4(\pi^*)^2$, there are four bonding pairs of electrons and two antibonding pairs giving rise to a net bond order of two. The pair of π^* orbitals is only doubly occupied, whereas there is space for four electrons. Hund's rules (which for the electronic ground state maximize the number of electrons with parallel spins) identify the lowest-energy arrangement as the one where each of the degenerate π^* components is singly occupied, the spins of the two electrons being parallel. Unpaired electrons give rise to paramagnetic behavior, and gaseous oxygen is indeed paramagnetic.

Multicenter bonding. Not all molecules can be described as being held together by simple two-center, two-electron bonds. Multicenter delocalized bonding, for example, is the best way of describing the π bonding which occurs in organic acyclic and cyclic polyenes. **Figure 5** shows the molecular orbitals that result from interaction of the $p\pi$ orbitals that lie perpendicular to the molecular plane in benzene and cyclobutadiene. For benzene there are three bonding orbitals, one lying deeper than the other degenerate pair and three analogous antibonding orbitals. In cyclobutadiene there is a bonding and antibonding orbital in addition to a degenerate pair of nonbonding orbitals. A nonbonding molecular orbital has the

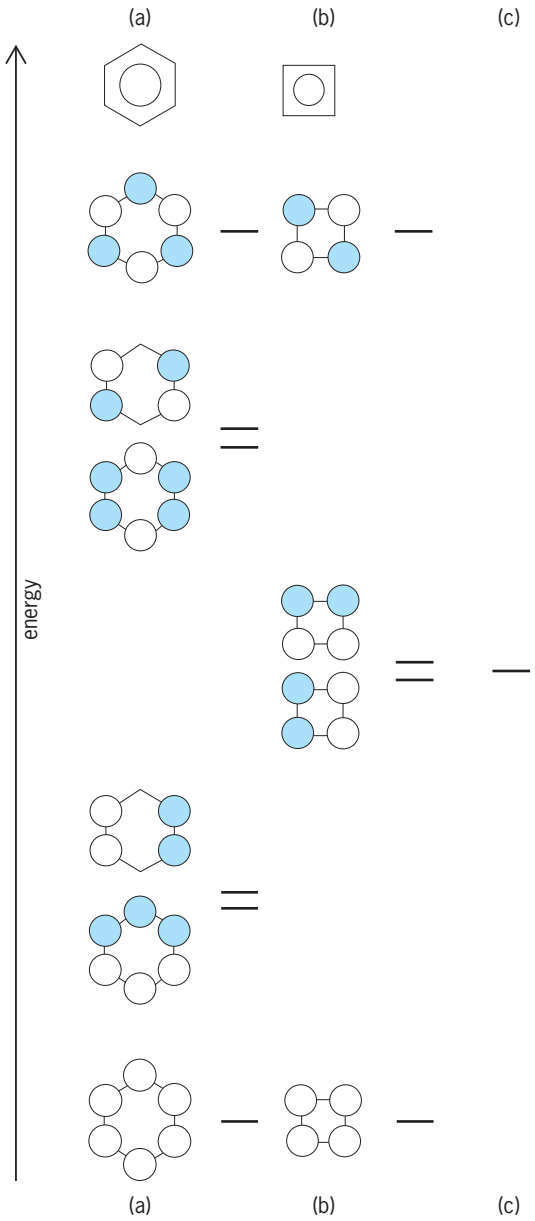

Fig. 5. Molecular π orbitals for (*a*) benzene and (*b*) cyclobutadiene viewed from above so that the relative phases only of the upper half of the $p\pi$ orbitals are shown. (*c*) The energy of an isolated $p\pi$ orbital.

same energy as the isolated atomic orbitals from which it is constructed. In these examples of neutral hydrocarbons the number of molecular orbitals is equal to the number of atomic orbitals used in synthesizing the diagram (Fig. 5). Each carbon atom contributes one electron to this π-orbital manifold. Thus in benzene the lower three orbitals are doubly occupied. As far as the π manifold of benzene is concerned, there are then three bonding pairs but six close contacts between carbon atoms. As a result, the π-bond order is one-half, and the π bonding in this molecule is referred to as being delocalized over the carbon framework. Similar ideas employing delocalized bonding can be applied to three-dimensional molecules such as the boranes, carboranes, and transition-metal cluster compounds. *See* BORANE; CARBORANE; METAL CLUSTER COMPOUND.

There is a simple rule that has its basis in the π-orbital structure of these hydrocarbons. If a molecule possesses $4n + 2\pi$ electrons ($n = 0, 1, 2, \ldots$), then it will be stable as a planar aromatic molecule. If the π-electron count is otherwise, then the planar symmetric molecule will be antiaromatic and unstable. This is called Hückel's rule. Benzene, with six electrons, satisfies the rule ($n = 1$), but cyclobutadiene does not. The latter molecule distorts away from the square geometry, probably to a rhombus, and is kinetically unstable. C_8H_8 is another species which does not satisfy the rule, and this molecule is nonplanar. $C_8H_8^{2-}$ and $C_4H_4^{2-}$ do satisfy Hückel's rule and are found as planar, symmetric units bound to transition metals. The species $S_4^{2+\text{plus;}}$, isoelectronic with $C_4H_4^{2-}$, is also a stable, square molecule. Hückel's rule has its foundation in the requirement that all bonding and nonbonding orbitals of the π network be full of electrons for structural stability. *See* RESONANCE (MOLECULAR STRUCTURE).

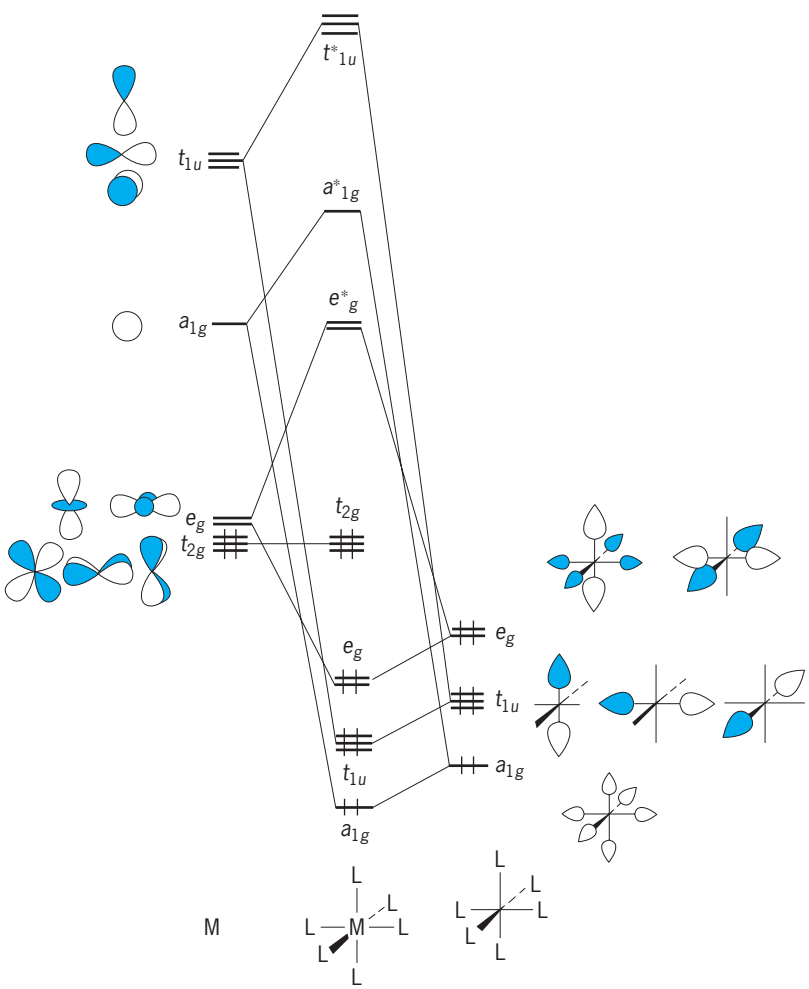

Fig. 6. Molecular orbital diagram for an ML_6 transition-metal complex. The labels a_{1g}, e_g, t_{1u}, and t_{2g} describe the symmetry properties of the orbital levels.

Transition-metal complexes. Similar orbital ideas applied to transition-metal complexes have led to an understanding concerning their structures and reactivity. Molecular orbital models have virtually replaced the older concepts of crystal field and valence bond theory in the area. **Figure 6** shows a molecular orbital diagram for an octahedral ML_6 complex, where L is a ligand that possesses an accessible orbital with which to bond in a σ fashion to the metal. (Examples are hydride, water, and halide.) This orbital diagram is different in several ways from those previously described. The ionization potential for an electron located in a ligand orbital is larger than for a metal-located electron. This is shown by drawing the metal atomic levels higher in energy than the ligand levels. The six ligand orbitals split into three sets when the ligands are coordinated to the metal. (The reasons for this are described in the next section.) The five atomic metal d orbitals split into two sets in ML_6, one doubly degenerate and one triply degenerate. The set labeled t_{2g} is nonbonding between metal and ligands, whereas the set labeled e_g is metal-ligand antibonding. The energy separation between the two is conventionally labeled δ. Electronic transitions between these orbital sets occur in the visible part of the spectrum for many first-row transition-metal complexes (scandium through copper) and are thus the source of their color. The number of electrons that occupy these orbitals is simply given by counting the number of $d + s$ electrons of the corresponding gaseous ion. Thus Cr(II) is a d^4 system while Zn(II) is a d^{10} system. In the latter, the e_g and t_{2g} orbitals are completely filled. In organometallic compounds where δ is large, a rule similar to Hückel's rule is found to apply. If all bonding and nonbonding orbitals are occupied, then kinetic and structural stability is conferred upon the system. This occurs when a total of 18 electrons (or nine electron pairs) are associated with the metal. In Fig. 6 this would occur for six pairs of electrons in ligand-located orbitals plus three pairs of electrons in the nonbonding t_{2g} set. This 18-electron rule has proven to be very useful in understanding organometallic chemistry. *See* ORGANOMETALLIC COMPOUND.

Symmetry considerations. Symmetry is very useful in understanding many aspects of molecular orbital theory. In qualitative terms, the more symmetric a molecule, the larger the number of elements of symmetry it possesses. **Figure 7** shows the symmetry elements—the mirror planes (σ_v), rotation axes (C_x, for a rotation of $360°/x$), and center of symmetry (i)—for the A_2 diatomic molecule shown in Fig. 3.

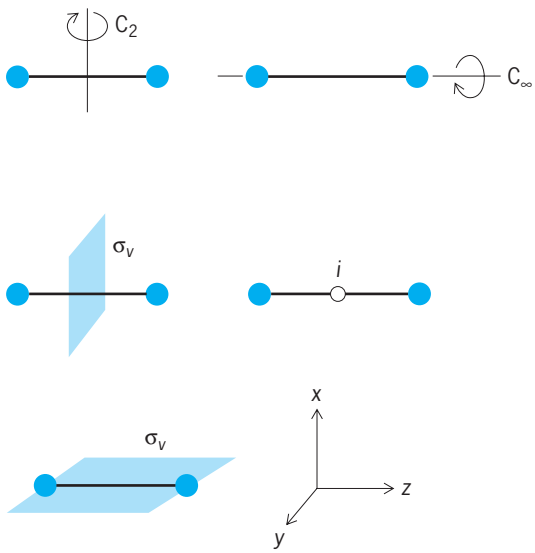

Fig. 7. Symmetry elements for the H_2 or A_2 diatomic molecule.

Although only one mirror plane containing the internuclear axis is shown, there are an infinite number of them, just as there are an infinite number of C_2 axes perpendicular to the internuclear axis.

The mathematical theory that allows manipulation of such observations is called group theory. It can be seen that the directions x and y in Fig. 7 are symmetry-equivalent in such a molecule, but that the z direction is unrelated to either x or y. A more formal way of putting this is to note that, although there are symmetry operations possessed by the molecule which will interconvert x and y, there is no such operation which will interconvert z with either x or y. A direct result of this observation is that the p_x and p_y orbitals will always be found in degenerate orbital situations, but that p_z will be energetically separate. In the linear molecule the symmetry label π is used to refer to such a double degeneracy. The nondegenerate p_z orbital is described by the symmetry label σ. The symmetry description of the three valence p orbitals on an A atom in the A_2 diatomic is then $\sigma + \pi$. As the number of symmetry elements associated with a molecule increases, the size of possible degeneracies also increases. In octahedral and tetrahedral molecules, for example, double and triple degeneracies are possible, and in icosahedral molecules fivefold degeneracies may occur. In the octahedral molecule shown in Fig. 6, the x, y, and z directions are symmetry-equivalent and are described by the symmetry label t_{1u} representing a triple degeneracy. (Greek letters are used for the symmetry labels of linear molecules, for example, σ or π and Roman letters, for example, a_{2g} or t_{2u} for nonlinear molecules.)

Here a useful subscript is used to describe a function's behavior with respect to inversion. The label g applies to functions that are symmetric with respect to such an operation (s and d orbitals on the central atom, for example) and the label u for antisymmetric functions (p and f orbitals on the central atom, for example). The five d orbitals become of $e_g + t_{2g}$ symmetry in the octahedral molecule, just as the three p orbitals become of $\sigma + \pi$ symmetry in the diatomic molecule shown in Fig. 3.

There is one very powerful result which is extremely useful in constructing orbital diagrams. Orbitals of different symmetry have zero overlap and therefore do not interact with each other. For example, in Fig. 4*b* the two p_x orbitals have a nonzero interaction since they are both of π symmetry, but the overlap integral for the combination of Fig. 4*c* is zero since one orbital is of π and the other of σ symmetry. The molecular orbital diagrams shown in Figs. 2, 3, and 6 may then be generated simply by matching orbitals of the same symmetry at the left- and right-hand sides of the picture and constructing bonding and antibonding orbital combinations from each pair. In Fig. 6 the central metal atom orbitals of t_{2g} symmetry find no symmetry match with the orbitals of the ligands and so remain nonbonding. Symmetry arguments such as these are behind many of the basic orbital concepts in molecules. *See* GROUP THEORY; SELECTION RULES (PHYSICS); SYMMETRY LAWS (PHYSICS).

Orbital symmetry. The Woodward-Hoffmann rules, which collect together within a theoretical framework many important organic reactions, are based on the symmetry control of orbital interaction. **Figure 8** shows orbital correlation diagrams linking the orbitals of a 1,4 substituted butadiene and the product of ring closure, the cyclobutene. Two different modes of closure may be envisaged, the conrotatory and disrotatory motion of the groups attached at the 1 and 4 positions. In the former a twofold rotation axis is preserved, and in the latter a mirror plane is preserved during the reaction. In Fig. 8*a* all the levels may then be classified as being either symmetric (S) or antisymmetric (A) with respect to the twofold rotation axis (C_2). Correspondingly in Fig. 8*b* the levels may be classified analogously according to their behavior with respect to the mirror plane operation. Orbitals of the same symmetry will interact with each other and will not cross along the reaction coordinate, but orbitals of different symmetry may cross.

The orbital correlation diagrams of Fig. 8 may thus be generated by connecting pairwise the orbitals at the left- and right-hand sides. Notice how in Fig. 8*b* the highest occupied molecular orbital (HOMO) of the reactant correlates with the lowest unoccupied molecular orbital (LUMO) of the product. For this reaction a high activation energy is expected as a result. Contrast the process in Fig. 8*a* which will experience no such energetic penalty. The thermal reaction of Fig. 8*b* is then orbitally forbidden, whereas that of Fig. 8*a* is orbitally allowed. On photochemical excitation of an electron from HOMO to LUMO of the reactant, the opposite is now true. The disrotatory motion is photochemically allowed, but the conrotatory motion is a photochemically forbidden process. This generalization has had a profound effect on organic chemistry. *See* ORGANIC REACTION MECHANISM; STEREOCHEMISTRY; WOODWARD-HOFFMANN RULE.

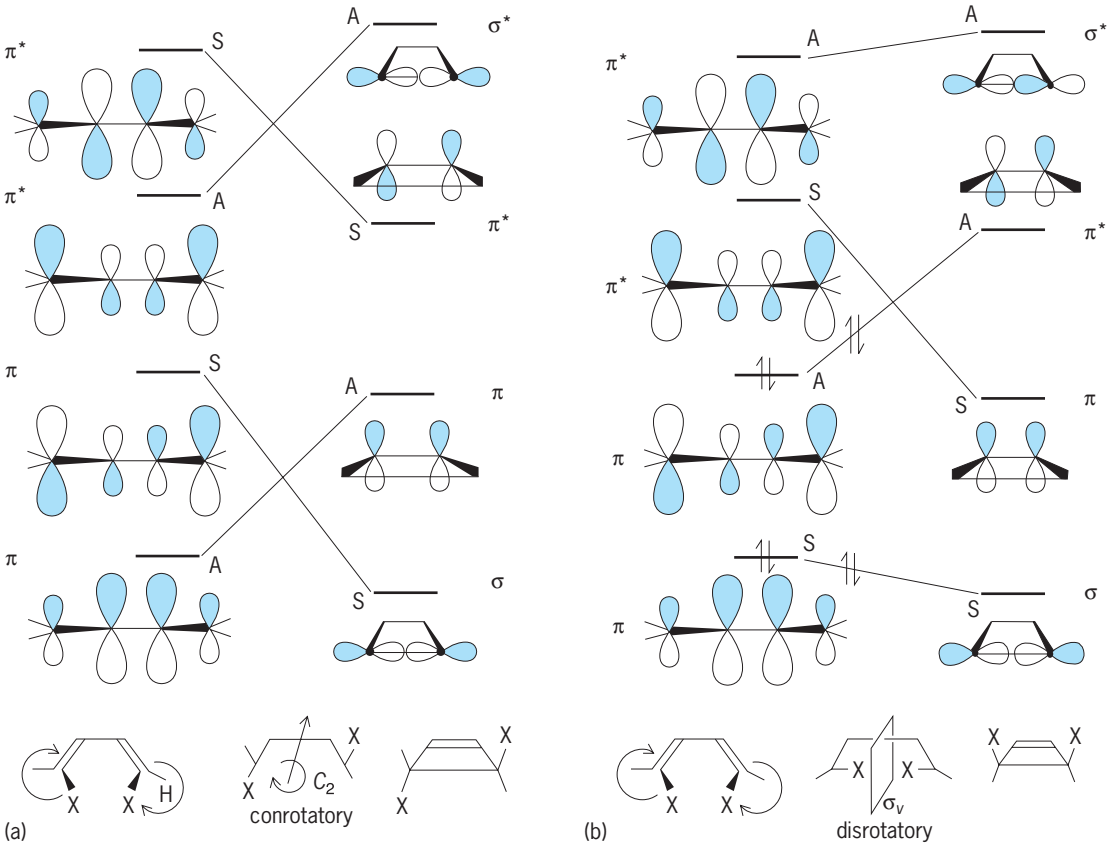

Fig. 8. Orbital correlation diagram for (*a*) the conrotatory and (*b*) the disrotatory ring closing of 1,4-butadienes.

Frontier orbitals. The use of valence orbitals alone is sufficient to build up the molecular orbitals and a qualitative picture of chemical bonding. This occurs because the core orbitals (for example, the 1*s* orbital on the oxygen atom) are very tightly bound and have a negligible overlap with the orbitals of a neighboring atom. When two molecules, or fragments, are brought together, it is often found that their mutual interaction is qualitatively well described by considering, in turn, a rather small number of these valence-orbital–based molecular orbitals. These so-called frontier orbitals are the orbitals of the two molecules which are spatially arranged so that the overlap integral between them is significant. Other interactions are either zero by symmetry, or small because their geometrical arrangement precludes good overlap.

Many aspects of chemical reactivity and structure may be understood by considering the details of their interaction. It often turns out that the energetically most important interactions occur between the HOMO of one molecule and the LUMO of another, giving rise to a stabilizing interaction. **Figure 9** shows how consideration of the frontier orbitals of these molecules allows understanding of the observation that maleic anhydride reacts easily with butadiene but with difficulty with ethylene. The HOMO-LUMO overlap in the former case (Fig. 9*a*) is nonzero and gives rise to a stabilizing interaction. In the latter such overlap is exactly zero since the positive overlap at the left-hand side is canceled by the negative overlap at the right-hand side and the interaction of the reactants is zero (Fig. 9*a*). In the area of organic chemistry, frontier orbital theory allows explanation of many aspects of the Woodward-Hoffmann rules. In the areas of molecular geometry such ideas are very useful in understanding the relative orientation of different parts of a molecule by maximizing such interactions.

Status. In the discussion above, the emphasis has been on the qualitative aspects of molecular orbital theory. Indeed, many of the advances in chemistry that have been led by theoretical considerations have employed arguments such as these, based on symmetry, electronegativity, and overlap. However, the numerical aspects of molecular orbital theory should not be underestimated. Quantitative calculations designed to estimate bond energies and map out the details of reaction pathways are becoming increasingly possible as a result of improvements in the size and speed of computers. The bond dissociation energies of small molecules that contain a few atoms from the second row of the periodic table can be calculated to chemical accuracy (±1 kcal/mole). However, it is not yet possible to predict the outcome of most experiments by calculations of this type. Even as that time approaches, it will still be necessary to have a qualitative understanding of the numbers

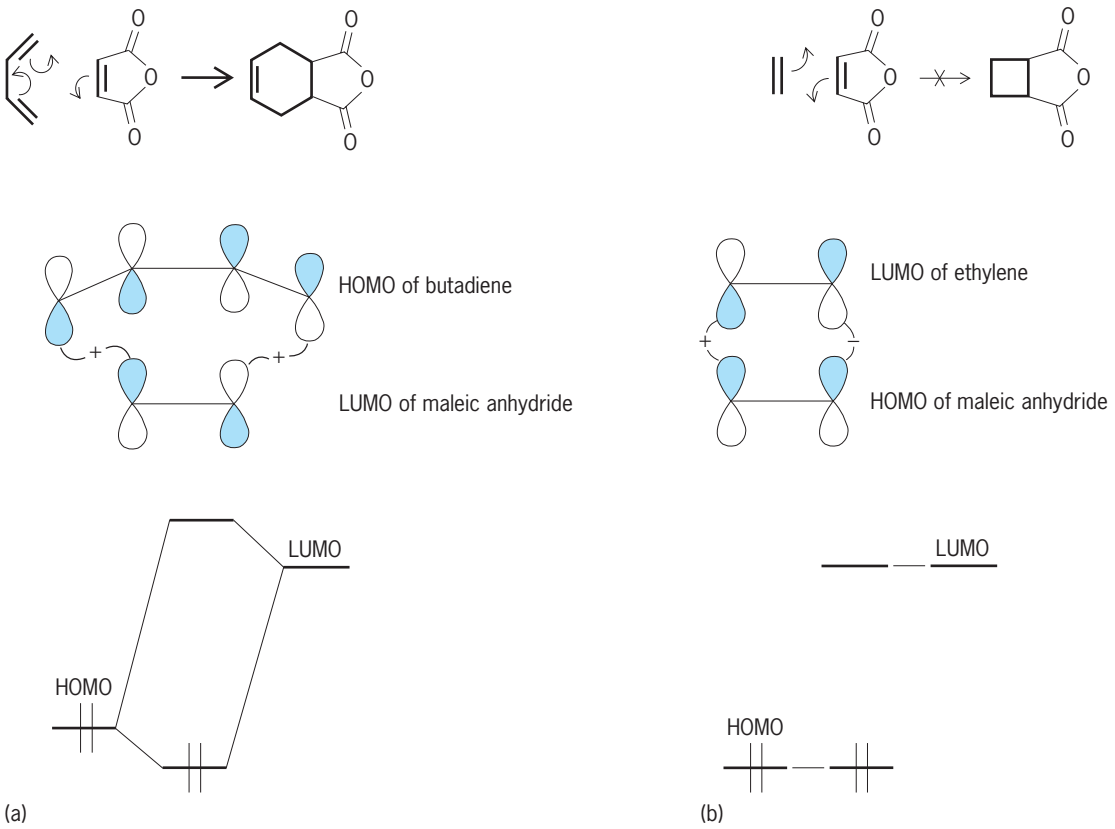

Fig. 9. HOMO-LUMO interactions between maleic anhydride and (*a*) butadiene and (*b*) ethylene.

generated by the computer. *See* BAND THEORY OF SOLIDS; COMPUTATIONAL CHEMISTRY; MOLECULAR STRUCTURE AND SPECTRA; QUANTUM CHEMISTRY.

Jeremy K. Burdett

Bibliography. T. A. Albright, J. K. Burdett, and M. H. Whangbo, *Orbital Interactions in Chemistry*, 1985; O. K. Anderson, V. Kumar, and A. Mookerjee, *Methods of Electronic Structure Calculations*, 1994; F. A. Cotton, *Chemical Applications of Group Theory*, 3d ed., 1990; I. Fleming, *Frontier Orbitals and Organic Chemical Reactions*, 1976; H. Rauk, *Orbital Interaction Theory of Organic Chemistry*, 2d ed., 2000; A. Streitweiser, *Molecular Orbital Theory for Organic Chemists*, 1961; R. B. Woodward and R. Hoffmann, *The Conservation of Orbital Symmetry*, 1970.

Molecular pathology

A discipline that deals with the origins and mechanisms of diseases at their most fundamental level, that of macromolecules such as deoxyribonucleic acid (DNA) and protein, in order to provide precise diagnoses and discover possible avenues for treatment. It is interdisciplinary, including infectious disease, oncology, inherited genetic disease, and legal issues such as parentage determination or forensic identity testing. While a variety of biophysical and biochemical techniques can be applied to study the molecular basis of disease, antibodies and nucleic acid probes are two of the principal approaches. *See* ANTIBODY; NUCLEIC ACID; ONCOLOGY.

Antibodies. Monoclonal antibodies recognize and noncovalently bind specific molecular shapes called epitopes on macromolecules. When these antibodies are either tagged to permit their detection or immobilized on a chromatographic column to purify their specific target molecule, they serve as powerful tools for analyzing pathologic processes. A monoclonal antibody conjugated to an enzyme generating a colored reaction product is the basis of the enzyme-linked immunosorbent assay (ELISA), which is widely applied in many diagnostic tests. *See* MONOCLONAL ANTIBODIES.

Autoimmune diseases result when the body produces antibodies against its own molecular components. Autoantibodies from the sera of affected individuals are often used as highly specific reagents for understanding the nature of these diseases and the role that the affected molecules or organelles normally perform in the cell. For example, autoantibodies directed against the skeletal-muscle acetylcholine receptor result in myasthenia gravis; individuals with scleroderma, a collagen vascular disease associated with excess collagen deposition in the skin and other organs, may have autoantibodies directed against components of the centriole. These autoantibodies can be detected in tissues by using fluorescently tagged antibodies directed against human

immunoglobulins. *See* AUTOIMMUNITY; FLUORESCENCE MICROSCOPE; MYASTHENIA GRAVIS.

In situ hybridization. In an analogous manner to immunohistochemistry, traditional histopathology (that is, the microscopic study of diseased tissues) can be enhanced by using in situ hybridization. With this technique, an infectious agent such as a virus or a specific messenger ribonucleic acid (mRNA) can be localized within a specific cell or tissue. The final appearance of the tissue section resembles an immunohistochemical result, although these distinct molecular probes have different targets; antibodies are generally used to detect proteins, while nucleic acid probes detect DNA or RNA.

The ability to localize specific molecular targets with DNA probes can be applied in cytogenetics by using fluorescent in situ hybridization (FISH). This technique can precisely map genes to a specific region of a chromosome in an appropriately prepared karyotype, or can enumerate chromosomes, or can detect chromosomal deletions, translocations, or gene amplifications in the interphase nucleus of cancer cells.

Nucleic acid probes. Diseases often result from germline or somatic mutations in the individual's DNA, such as are seen in sickle cell disease or cancer, respectively. These abnormalities can be detected by using two basic techniques of molecular genetics: the Southern blot and the polymerase chain reaction (PCR).

Southern blot. For the Southern blot, high-molecular-weight DNA isolated from a specimen (most commonly peripheral-blood white cells) is digested by using an appropriate restriction endonuclease. The resulting fragments are then separated by molecular weight by means of gel electrophoresis. This DNA is transferred in its single-stranded form from the gel to a nylon membrane. This membrane is incubated with a solution containing a specific, labeled probe, also in single-stranded form. Probe-target hybrids formed by annealing of their complementary sequences can be detected by autoradiography or a colorimetric reaction.

The Southern blot technique can detect DNA polymorphisms, mutations (including deletions, expansions, amplifications, or translocations), or the presence of viral, bacterial, or specific sex chromosomes (such as the Y chromosome, indicating male sex). By using a Southern blot, fragile X syndrome, the most common cause of inherited mental retardation, can be diagnosed by the presence of a larger band on the autoradiograph.

Polymerase chain reaction. The polymerase chain reaction has revolutionized molecular genetics and provides a powerful tool for diagnosis. Short oligonucleotide primers flank the specific gene region or RNA sequence to be amplified and are combined with the target specimen and free nucleotides, which are synthesized into new DNA by *Taq* DNA polymerase. An automated thermal cycler repeatedly alters the temperature to denature the target DNA, to allow the primers to reanneal to the target, and then to synthesize the product. This amplified polymerase chain reaction product is typically detectable as a band in a gel. Simultaneous amplification of several genes or gene regions in the same reaction is known as multiplex polymerase chain reaction, which permits more rapid testing.

Cystic fibrosis provides an example of the clinical utility of polymerase chain reaction diagnosis. It is an autosomal recessive disease resulting from a defective gene located on the long arm of chromosome 7. When this gene is mutated, the cell's ability to adjust concentrations of chloride ion is altered, resulting in sweat with elevated chloride levels (the basis for the traditional sweat chloride determination), and in thickened, relatively dehydrated secretions in the body. These secretions obstruct various glands such as the pancreas, ultimately resulting in atrophy and fibrosis, while the lungs become susceptible to repeated infections and subsequent tissue injury. The most common mutation causing cystic fibrosis is a deletion at a position that normally encodes a phenylalanine. After polymerase chain reaction amplification, this deletion is detected as a smaller band when examined in an electrophoretic gel.

Advantages. Understanding diseases at the genetic level has several advantages. The central dogma of molecular biology—that biological information flows from DNA to RNA to protein—permits an impressive understanding of a disease based on the DNA sequence of the involved gene. Even when the gene's protein product is not expressed, definitive diagnoses can be made by using DNA-based techniques. For example, fetal DNA from an amniocentesis permits the prenatal diagnosis of sickle cell disease even though the defective beta chain of hemoglobin is not sufficiently expressed within the fetus to be detected with standard hemoglobin electrophoresis. Diseases that are similar clinically (phenotypically) can, in fact, be due to different mutations (genotypes) within a single gene or due to mutations in different genes, often related in an enzyme complex or as a portion of a group of structural proteins. *See* MUTATION.

Another advantage of molecular diagnosis is the ability to detect phenotypically normal carriers of genetic diseases in order to provide information for appropriate genetic counseling and prenatal diagnosis. *See* HUMAN GENETICS; PATHOLOGY. Mark A. Lovell

Bibliography. N. Lemoine, J. Neoptolemos, and T. Cooke, *Cancer: A Molecular Approach*, 1994; E. Rubin and J. L. Farber, *Pathology*, 3d ed., 1999.

Molecular physics

The study of the physical properties of molecules. Molecules possess a far richer variety of physical and chemical properties than do isolated atoms. This is attributable primarily to the greater complexity of molecular structure, as compared to that of the constituent atoms. Molecules also possess

additional energy modes because they can vibrate; that is, the constituent nuclei oscillate about their equilibrium positions and rotate when unhindered. These modes give rise to additional spectroscopic properties, as compared to those of an atom; molecular spectroscopy in the optical, infrared, and microwave regions is one of the physical chemist's most powerful means of identifying and understanding molecular structure. Molecular spectroscopy has also given rise to the rapidly growing field of molecular astronomy.

Molecular physics is primarily concerned with the study of properties of isolated molecules, as contrasted to the more general study of molecular reactions, which is the domain of physical chemistry. Such properties, in addition to the broad field of spectroscopy, include electron affinities (for the formation of molecular negative ions); polarizabilities (the "distortability" of the molecule along its various symmetry axes by external electric fields); magnetic and electric multipole moments, attributable to the distributions of the electric charge; currents and spins of the molecule; and the (nonreactive) interactions of molecules with other molecules, atoms, and ions.

One of the most important areas of molecular physics is the application of computational techniques to the calculation of molecular wave functions from basic principles and, consequently, of molecular charge distributions and of all the physical properties that are noted above.

Intermolecular forces are responsible for such varied phenomena as condensation, surface tension, gaseous diffusion, and, most importantly, for the formation of crystalline and noncrystalline solids. Thus, the field is basic to achieving an understanding of the physical world. It would be difficult to list all the articles in the present volumes which could be classed under the domain of molecular physics. For a representative sampling *see* COSMOCHEMISTRY; INFRARED SPECTROSCOPY; INTERMOLECULAR FORCES; MICROWAVE SPECTROSCOPY; MOLECULAR BEAMS; MOLECULAR STRUCTURE AND SPECTRA; SPECTROSCOPY.

Benjamin Bederson

Molecular recognition

The ability of biological and chemical systems to distinguish between molecules and regulate behavior accordingly. How molecules fit together is fundamental in disciplines such as biochemistry, medicinal chemistry, materials science, and separation science. A good deal of effort has been expended in trying to evaluate the underlying intermolecular forces. The weak forces that act over short distances (hydrogen bonds, van der Waals interactions, and aryl stacking) provide most of the selectivity observed in biological chemistry and permit molecular recognition. The recognition event initiates behavior such as replication in nucleic acids, immune response in antibodies, signal transduction in receptors, and regulation in enzymes. Most studies of recognition in organic chemistry have been inspired by these biological phenomena. It has been the task of bioorganic chemistry to develop systems capable of such complex behavior with molecules that are comprehensible and manageable in size, that is, with model systems. *See* CHEMORECEPTION; ENZYME; HYDROGEN BOND; INTERMOLECULAR FORCES; NUCLEIC ACID; SYNAPTIC TRANSMISSION.

Macrocycles. The advantage of cyclic structures lies in their ability to restrict conformation or flexibility. A rigid matrix of binding sites, that is, preorganized sites, is usually associated with high selectivity in binding. A flexible matrix tends to accept several binding partners. Although sacrificing selectivity, this has the advantage of transmitting conformational information and is relevant to biological signaling events.

Crown ethers. Early work in the area focused on macrocyclic (crown) ethers. Such structures can bind and transport ions and imitate biological processes involving macrolides. Large ring structures that are lined with oxygen present an inner surface which is complementary to the spherical outer surface of positively charged ions. *See* MACROCYCLIC COMPOUND.

Cyclophane-type structures. Cyclophane-type structures offer considerable rigidity because of the aromatic nuclei. The further incorporation of charged groups on the periphery imparts sufficient water solubility for complexing in aqueous systems. Binding forces between host and guest are largely hydrophobic, with some contribution of aryl-stacking interactions involving the inner surfaces of the cyclophane and the outer π surfaces of the guest species. A typical system is a cyclophane-naphthalene complex (**1**), in which a naphthalene

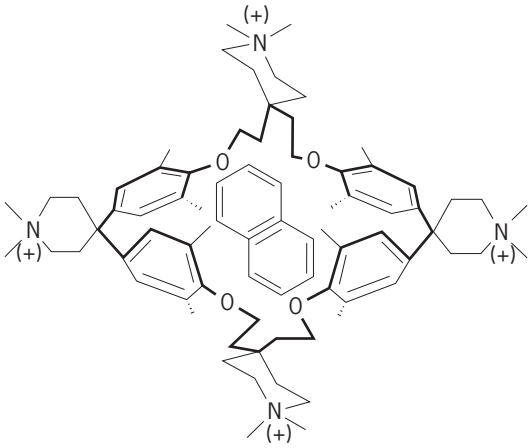

(**1**)

guest is bound by a water-soluble cyclophane derivative. *See* COORDINATION COMPLEXES.

Cyclodextrins. Naturally occurring cycloamylose derivatives (cyclodextrins), such as cyclohexa-

amylose (α-cyclodextrin; **2**), likewise present a

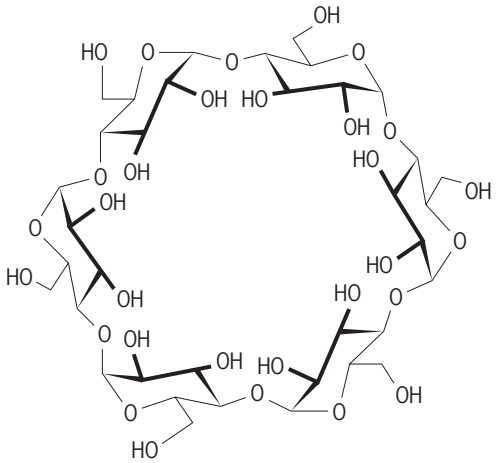

(**2**)

hydrophobic inner surface and a hydroxylic outer surface for complexation of neutral structures in water. Aromatics, adamantanes, and even ferrocene derivatives fit snugly within. Such complexes can place electrophilic sites of the guest near the hydroxyl groups of the host and result in acyl transfer reactions relevant to enzymatic processes. *See* AROMATIC HYDROCARBON.

Hybrid structures. Molecules with large cavities can be assembled in modular fashion by using macrocyclic subunits. In one such supramolecular assembly (**3**), a large aromatic porphyrin surface pro-

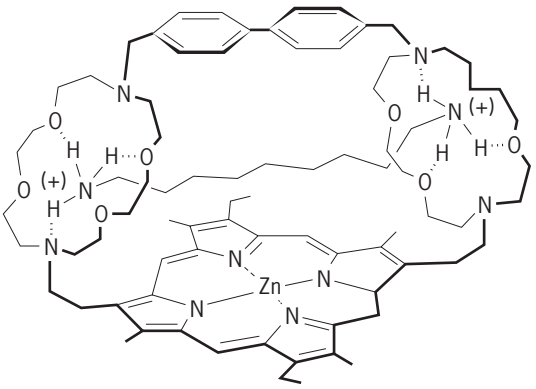

(**3**)

vides the spacer between two crown ether binding sites. A number of rigid spacing elements other than the aryls shown have been used in this way to permit binding to diammonium ions of appropriate size. In addition, macrocycles composed of both carbohydrate and diaryl spacers have been prepared.

Other cyclic structures. The lining of macrocyclic interiors with acidic sites has permitted the complexation of anions. A synthetic receptor capable of finding certain carboxylates has been prepared. These structures, such as (**4**), imitate the vancomycin class

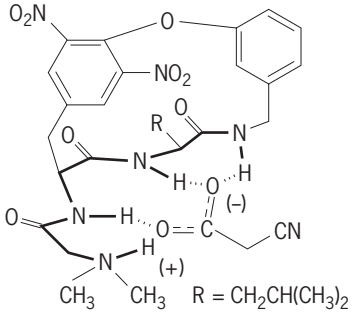

(**4**)

of antibiotics, which bind to peptide acids involved in bacterial cell-wall biosynthesis. A number of modified cyclic structures bind heterocycles such as imidazoles.

Concave structures. Because the encircling of larger, more complex molecules with macrocycles poses structural problems, other molecular shapes have been explored. Cleft molecules offer advantages in this regard. The principle underlying these systems involves the shape of the small organic target molecules: convex in surface and bearing functional groups that diverge from their centers. Accordingly, designing a trap for such targets requires molecules of a concave surface in which functional groups converge. This complementarity is also a feature of the immune system: the "hot spots" of an antigen tend to be convex, whereas the binding sites of the antibody are concave.

Systems featuring a cleft have been developed to bind adenine derivatives and other heterocyclic systems through chelation, as shown in (**5**). The

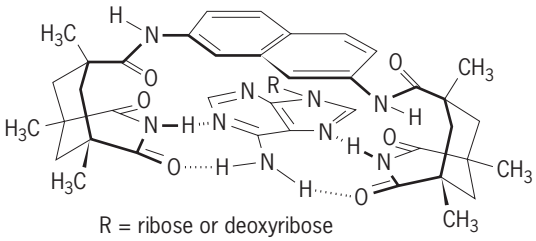

(**5**)

concave surface can provide both Watson-Crick and Hoogsteen base pairing, whereas aromatic stacking interactions help fix the adenine to the synthetic receptor. The high lipophilicity of these structures permits the transport of adenosine across liquid membranes between two aqueous phases. A sickle-shaped system takes advantage of the rigid aromatic structures bearing heterocyclic nitrogen. These provide convergent acidic and basic binding sites closely tailored to uric acid derivatives, which they complex [structure (**6**)] and solubilize in organic media.

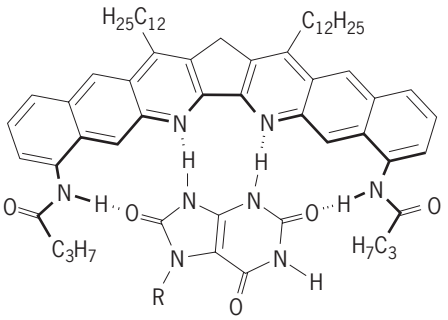

(6)

Related molecules bind urea. *See* CHELATION; UREA; URIC ACID.

Cleft structures also offer hydrophobic and van der Waals forces. In one system, decalin is sandwiched [structure (7)] between the π systems of a rigid syn-

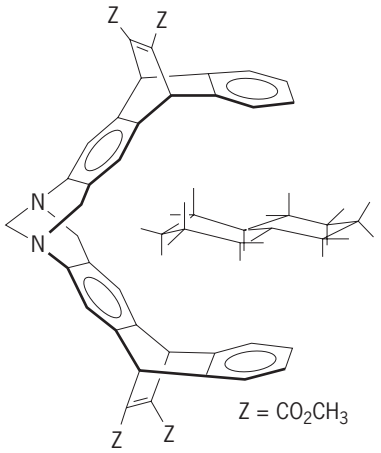

(7)

thetic receptor. At the other extreme, highly polar microenvironments have been engineered into concave surfaces. An example is provided by the metal chelate structure (8). The convergence of the more

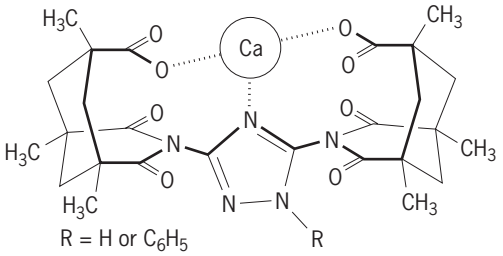

(8)

basic electron pairs of the unusual dicarboxylic acid provides high selectivity (a molecular vise) for divalent ions.

Experimental methods. The complexation of the synthetic receptors with the target molecules is most frequently assessed by observation of solubilization. For example, the extractability of metal ions from aqueous solutions into organic solvents is evidence

Association constants for complexes of the acridine-derived diacid*

Base	K_a (M^{-1})	pK_a (BH^+)
Triethylene diamine*	1.6×10^5	8.2, 4.2
Pyrazine	1.4×10^3	0.65
Quinoxaline	23×10^3	0.56
Phenazine	2.2×10^3	1.2
Pyrimidine	0.7×10^3	1.3
Quinazoline	1.6×10^3	1.9
Imidazole†	$K_1 = 1.0 \times 10^6$	6.9
	$K_2 = 5.5 \times 10^4$	
Benzimidazole†	$K_1 = 1.5 \times 10^4$	5.5
	$K_2 = 7.5 \times 10^3$	
Purine	$\sim 8 \times 10^3$	2.4
Pyridine†	$K_1 = 1.2 \times 10^2$	5.2
	$K_2 < 1$	

* In chloroform-*d* ($CDCl_3$) solution, at 77°F (25°C).
† K_1 and K_2 are association constants for one and two molecules, respectively.

of an unusually favorable relationship between host and guest. Similar observations can be made with host species and amino acids, barbiturates, nucleosides, and uric acid derivatives. It is also possible to perform solid–liquid extractions, because the complexes tend to be more soluble than the individual components. Another way of establishing unusual affinities is through transport studies. Active or passive transport of amino acids, metal ions, and nucleosides can be observed by using liquid membranes that consist of a chloroform layer between two aqueous solutions. *See* AMINO ACIDS; BARBITURATES.

Quantitative determinations are available through titrations. Association constants can be determined between host and guest by spectroscopic techniques. Frequently, nuclear magnetic resonance (NMR) is convenient because it determines response and concentration simultaneously. A special advantage of nuclear magnetic resonance deals with the use of Overhauser effects and two-dimensional spectra. These permit the assignment of contact points between host and guest that map the geometry of the complexes. The **table** shows typical data for titration of molecular clefts featuring convergent acidic sites with appropriate guest species that contain divergent basic sites in chloroform-*d* ($CDCl_3$) solution; the molecular chelation of triethylene diamine (DABCO) within a cleft derived from acridine is shown in (9). Such data reveal features

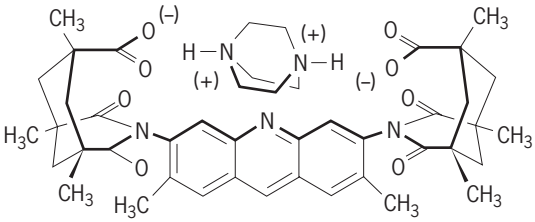

(9)

of size and shape complementarity, acid-base effects, and stereoelectronic details within the complexes. *See* CONFORMATIONAL ANALYSIS; NUCLEAR

MAGNETIC RESONANCE (NMR); SPECTROSCOPY; STEREOCHEMISTRY; TITRATION.

Future applications. Apart from the abstract questions concerning articulation of molecules, some practical applications in the pharmaceutical industry may be envisioned. Many of the target structures are biologically active, and the use of synthetic sequestering agents for metabolic substrates can represent a new approach to biochemical methods and drug delivery. This represents a departure from most current strategies, in which small convex molecules (penicillin, mitomycin, beta blockers) are directed at the folds of a biological macromolecule. In addition, development of sensing devices at the molecular level and the synthesis of new composites are expected. Another issue to be addressed will be how the molecular fit alters the chemical behavior of the interacting partners. Julius Rebek, Jr.

Bibliography. J. L. Atwood (ed.), *Inclusion Phenomena and Molecular Recognition*, 1990; A. D. Buckingham, A. C. Legon, and S. M. Roberts (eds.), *Principles of Molecular Recognition*, 1993; M. Delaage (ed.), *Molecular Recognition Mechanisms*, 1991; E. R. Weber (ed.), *Supramolecular Chemistry, I: Directed Synthesis and Molecular Recognition*, 1993.

Molecular sieve

Any one of the crystalline metal aluminosilicates belonging to a class of minerals known as zeolites. These minerals are found widely scattered in nature in relatively small quantities. Synthetic forms of the naturally occurring minerals, as well as many species having no known natural counterpart, have been prepared by a hydrothermal process. An important characteristic of the zeolites is their ability to undergo dehydration with little or no change in crystal structure. The dehydrated crystals are honeycombed with regularly spaced cavities interlaced by channels of molecular dimensions which offer a very high surface area for the adsorption of foreign molecules.

Structure. The basic formula for all crystalline zeolites can be represented as where M represents a

$$M_{2/n}O{:}\,Al_2O_3{:}\,xSiO_2{:}\,yH_2O$$

metal ion and n its valence. In general, a particular crystalline zeolite will have values for x and y that fall into a definite range. For example, in two synthetic varieties of molecular sieve, designated type A and type X, the values of x are typically about 2.0 and 2.5, respectively. When the crystal is fully dehydrated, the value of y for both types is zero. The crystal structure consists basically of a three-dimensional framework of SiO_4 and AlO_4 tetrahedrons (**Fig. 1**). The tetrahedrons are cross-linked by the sharing of oxygen atoms, so that the ratio of oxygen atoms to the total of silicon and aluminum atoms is equal to 2. The electrovalence of the tetrahedrons containing aluminum is balanced by the inclusion of cations in the crystal. One cation may be exchanged for another by the usual ion-exchange techniques. The size of the cation and its position in the lattice determine the effective diameter of the pore in a given crystal species.

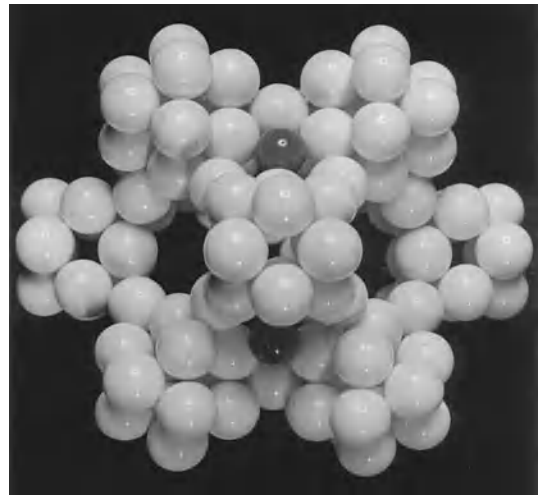

Fig. 1. Molecular sieve type-A crystal model. Dark spheres represent the included cations, and light spheres the SiO_4 or AlO_4 tetrahedrons.

The crystal habit of molecular sieve type X is similar to that of diamond, in which the carbon atoms are replaced by silica-alumina polyhedrons. With alkali metal ions present in the structure, effective pore diameter is between 0.9 and 1.1 nanometers; with the alkaline earth cations present, effective diameter is between 0.8 and 0.9 nm.

Properties. The properties of molecular sieves as absorbents which distinguish them from nonzeolitic adsorbents are (1) the relatively strong coulomb fields generated by the adsorption surface and (2) the uniform pore size; the pore size is controlled, in a given crystal species, by the associated cation. The strong surface forces are reflected in the peculiar shape of the adsorption isotherm, the character of the adsorption isobar, and the relatively high heats of adsorption. The isotherm, found by plotting the capacity for a given adsorbate against pressure or concentration at constant temperature, is of the so-called Langmuir type (**Fig. 2**). The shape of the isotherm is approximately rectangular, rising steeply at low partial pressures or concentrations and leveling off when maximum load is attained. The isobar, found by plotting capacity against temperature at constant pressure, shows that molecular sieves have an unusually high capacity at elevated temperatures. The relatively high exothermic heats characteristic of adsorption on molecular sieves necessitate somewhat higher heat requirements to effect desorption than are necessary with other adsorbents.

The basic characteristics of molecular sieves are utilized commercially in several production and research applications. Their absorption properties make them useful for drying, purification, and separations of gases and liquids. Conversely, molecular sieves can be preloaded with chemical agents, which are thereby isolated from the reactive system

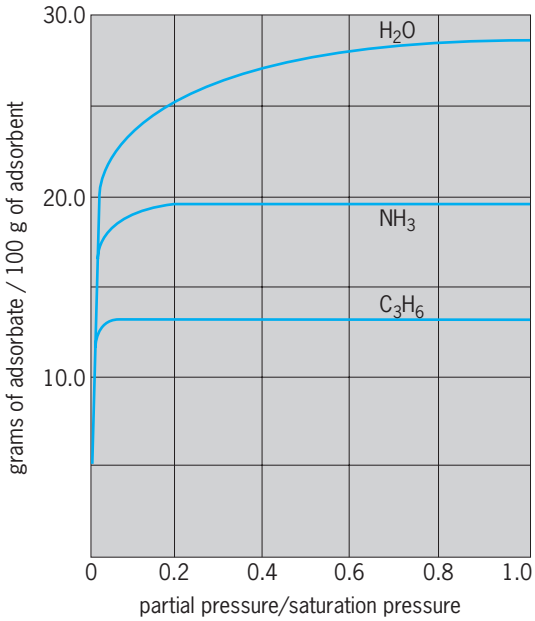

Fig. 2. Isotherms for type-5A molecular sieves at 25°C (77°F).

in which they are dispersed until released from the adsorbent either thermally or by displacement by a more strongly adsorbed compound. The presence in the crystal lattice of an associated exchangeable metal ion provides the basis for their use as a cation exchange medium. Their chemical composition and crystal structure make them novel catalysts and catalyst supports.

As is predictable from the water-adsorption isotherm (Fig. 2), molecular sieves are capable of drying gases and liquids to extremely low residual water concentrations. The isotherm also shows that, even at low initial water concentrations in a gas or liquid, the relative desiccant capacity is high.

By virtue of the uniform pore size of a given molecular-sieve crystal type, molecules having a minimum projected cross section larger than the effective diameter of the zeolite pore are excluded from the internal surface. Molecules having a minimum projected cross section smaller than the effective pore diameter are adsorbed internally. This phenomenon is utilized in separating molecules of fluid mixtures on the basis of their size or shape. For example, the 0.5-nm molecular sieve type has an effective pore size such that straight-chain hydrocarbons are adsorbed and thus effectively separated from branched-chain and cyclic hydrocarbons, which are excluded from the pore system in the selective adsorption process.

In general, when two molecules of similar volatility are sufficiently small to enter the pore system, separation is based on the degree of unsaturation or on the polarity of the molecules. The more unsaturated or more polar molecule is more strongly adsorbed. *See* ADSORPTION; GAS CHROMATOGRAPHY; ION EXCHANGE; ZEOLITE. R. L. Mays

Bibliography. D. W. Breck, *Molecular Sieves: Structure, Chemistry and Use*, 1974; E. M. Flanigen and L. B. Sand (eds.), *Molecular Sieve Zeolites II*, 1971; W. H. Flank and T. E. White (eds.), *Perspectives in Molecular Sieve Science*, 1988; J. R. Katzer (ed.), *Molecular Sieves II*, 1977; W. M. Meier and J. B. Uytterhoeven (eds.), *Molecular Sieves*, 1973.

Molecular simulation

Molecular simulation and computational quantum chemistry are rapidly evolving tools which are having an increasingly significant impact on the chemical, pharmaceutical, materials, and related industries. These tools provide a mechanism for predicting entirely computationally many useful functional properties of systems of interest in these industries. Included are thermodynamic properties (such as equations of state, phase equilibria, and critical constants), thermochemical properties (such as energies of formation and reaction, and reaction pathways), spectroscopic properties (such as dipole moments and vibrational and other spectra), mechanical properties (such as stress-strain relationships and elastic moduli), transport properties (such as viscosity, diffusion, and thermal conductivity), and morphological information (such as location and shape of binding sites on a biomolecule and crystal structure). This list continues to grow as algorithmic and computer hardware advances make it possible to access additional properties.

Definitions. The two main molecular simulation techniques are molecular dynamics and Monte Carlo simulation, both of which are rooted in classical statistical mechanics. Given mathematical models for the internal structure of each molecule (the intramolecular potential which describes the energy of each conformation of the molecule) and the interaction between molecules (the intermolecular potential which describes the energy associated with molecules being in a particular conformation relative to each other), classical statistical mechanics provides a formalism for predicting properties of a macroscopic collection of such molecules based on statistically averaging over the possible microscopic states of the system as it evolves under the rules of classical mechanics. Thus, the building blocks are molecules, the dynamics are described by classical mechanics, and the key concept is statistical averaging. In molecular dynamics, the microscopic states of the system are generated by solving the classical equations of motion as a function of time (typically over a period limited to tens of nanoseconds). Thus, one can observe the relaxation of a system to equilibrium (provided the time for the relaxation falls within the time accessible to molecular dynamics simulation), and so molecular dynamics permits the calculation of transport properties which at the macroscopic scale describe the relaxation of a system in response to inhomogeneities. In Monte Carlo simulation, equilibrium configurations of systems are generated stochastically according to the probabilities rigorously known from classical statistical mechanics. Thus, Monte Carlo simulation

generates equilibrium states directly (which has many advantages, including bypassing configurations which are not characteristic of equilibrium but which may be difficult to escape dynamically) and so can be used to study equilibrium configurations of systems which may be expensive or impossible to access via molecular dynamics. The drawback of Monte Carlo simulation is that it cannot yield the kind of dynamical response information that leads directly to transport properties. Many hybrid methods exist that combine the best features of molecular dynamics and Monte Carlo simulation. Today, intra- and intermolecular forces are determined by a combination of computational quantum chemistry methods and fitting to available experimental data, and the development of these functions (also known as force fields) is a very active subfield of molecular simulation. *See* CHEMICAL DYNAMICS; COMPUTATIONAL CHEMISTRY; MONTE CARLO METHOD; SIMULATION; STOCHASTIC PROCESS.

In computational quantum chemistry, the Schrödinger equation for the electronic degrees of freedom of a collection of atoms (nuclei and electrons) is solved numerically to determine the minimum energy configuration of the atoms. The strength of computational quantum chemistry is that it makes few, if any, assumptions (unlike molecular dynamics and Monte Carlo simulation, in which a model must be assumed for the intra- and intermolecular interactions). The problem is that the computational cost is very high, limiting the methods to relatively small system sizes on current computers. The success of computational quantum chemistry methods resulted in one of the pioneers in the field, John Pople, receiving the 1998 Nobel Prize in Chemistry. One of the developments that has significantly sped up computational quantum chemistry calculations is density functional theory, which solves equations for electron density rather than positions of individual electrons. The primary developer of density functional theory, Walter Kohn, shared the 1998 Nobel prize with Pople. In the course of finding the minimum energy configuration of a collection of molecules, computational quantum chemistry yields information about intra- and intermolecular interactions, and so represents a route to these functions needed by molecular dynamics and Monte Carlo simulation. In fact, by combining molecular dynamics with the density functional theory form of computational quantum chemistry (so that the interactions needed by molecular dynamics are provided by density functional theory calculations) one can perform so-called ab initio molecular dynamics, the best-known version being Car-Parrinello. The Car-Parrinello and similar hybrid methods are presently restricted to fairly small systems simulated over periods of 1–10 picoseconds, yet they offer a glimpse into the future of molecular simulation and computational quantum chemistry when improvements in hardware and algorithms will make ab initio hybrid methods the primary tool for predicting the properties of systems about which little or nothing is known experimentally. *See* COMPUTATIONAL CHEMISTRY; QUANTUM CHEMISTRY.

The use of molecular simulation and computational quantum chemistry has become widespread, well beyond the research specialists, by the development of commercial versions of research codes, often with graphical interfaces that facilitate use of the methods. Companies which market such codes include Molecular Simulations Inc. (molecular simulation and computational quantum chemistry), Gaussian Inc. (computational quantum chemistry), Oxford Molecular (molecular simulation and computational quantum chemistry), and Schrödinger Inc. (molecular simulation/computational quantum chemistry hybrid). This list is not exhaustive. In addition, many noncommercial ventures (located at national laboratories or universities) make available molecular simulation and computational quantum chemistry codes. Examples include the NWChem package from Pacific Northwest National Laboratory (molecular simulation and computational quantum chemistry) and the DL-POLY code (molecular simulation) available from Daresbury Laboratory in the United Kingdom.

Capabilities and impact. Computational quantum chemistry and molecular simulation methods can be used to predict properties that once were only accessible experimentally, resulting in several significant applications in basic and industrial research. These applications include providing estimates of properties for systems for which little or no experimental data are available, which is especially useful in the early stages of chemical process design; yielding insight into the molecular basis for the behavior of particular systems, which is very useful in developing engineering correlations, design rules, or quantitative structure-property relations; and providing guidance for experimental studies by identifying the interesting systems or properties to be measured.

In the fields of chemical engineering and materials science, academic and industrial interest in molecular simulation and computational quantum chemistry has grown exponentially over the past decade. Two developments are making molecular simulation and computational quantum chemistry and practical tools for application to systems of industrial interest. One is the continuing exponential increase in computing power, often referred to as Moore's law, which states that the speed of computers is increasing by a factor of two every 18 months (equivalently, an order of magnitude every 5 years). The practical consequence of Moore's law is that today's desktop work stations and laptops have the computational power of a supercomputer of a decade ago, thus making it possible for an ordinary user to perform calculations that once would have been prohibitively expensive. The second development is the proliferation of efficient algorithms for calculating properties of practical interest. In molecular simulation, for example, the Gibbs ensemble Monte Carlo algorithm, introduced just over a decade ago, made it possible to predict the phase envelope of a fluid system directly from knowledge of the intermolecular forces,

thus opening up the possibility of calculating many properties relevant to the engineering design separations processes (such as distillation). Algorithm advances (which can result in order-of-magnitude or more increases in efficiency), combined with the steady exponential increase in computational power, will continue to dramatically broaden the range of applicability of computational quantum chemistry and molecular simulation methods in the coming years. The expectations for molecular simulation and computational quantum chemistry are very high, and they have been identified as key technologies for the chemical industry to achieve its vision of completely automated product and process design in the year 2020.

Examples. A sampling of companies, and examples of the industrial problems which they have tackled using molecular simulation and computational quantum chemistry techniques, is provided in a National Science Foundation report (1997). Applications are diverse and include molecular design of zeolites for air separation, design and optimization of catalysts for homogeneous and heterogeneous catalysis, evaluation of the feasibility of methane adsorption, molecular modeling of polysaccharide rheology, molecular design of detergent enzymes, and molecular design of photocopying materials. The 2000 report was expected to provide additional comparative information on the ways that European, Japanese, and American companies are applying these tools.

One area which is increasingly benefiting from molecular simulation methods is rheology. Nonequilibrium molecular dynamics methods have become very powerful tools for direct simulation of systems subject to flow fields, such as the shear flow field present in lubrication and the extensional flow fields present in polymer processing. The theory for deriving and validating methods for performing nonequilibrium molecular dynamics simulations has been perfected over the past 15 years. Recently, this technique has been used to predict the viscosity index of several typical lubricant components. The viscosity index is a key performance indicator for lubricants, measuring the degree to which the viscosity diminishes with increasing temperature. Nonequilibrium molecular dynamics methods are also being used to understand how nanoscale confinement, such as in the case of hard-disk drive lubrication, affects the rheological properties of lubricants (**Fig. 1**). These studies also shed light on the lubrication of microelectromechanical systems.

Rational catalyst design is one example of the use of computational quantum chemistry and ab initio molecular dynamics. Density functional methods make it possible to study catalytic processes qualitatively, and in some cases quantitatively, with increasingly sophisticated questions becoming answerable. For example, **Figure 2**, based on ab initio molecular dynamics simulations, shows a snapshot from a simulation study of the effect of water on the dissociation of acetic acid over Pd(III). The protons remain solvated in the water layer (as $H_5O_2^+$ species), while acetate anions are bound to the surface. The solvent lowers the interaction energy significantly. The water solvent interacts with the naked negative charge of the anion, thus decreasing the anion's affinity for the surface.

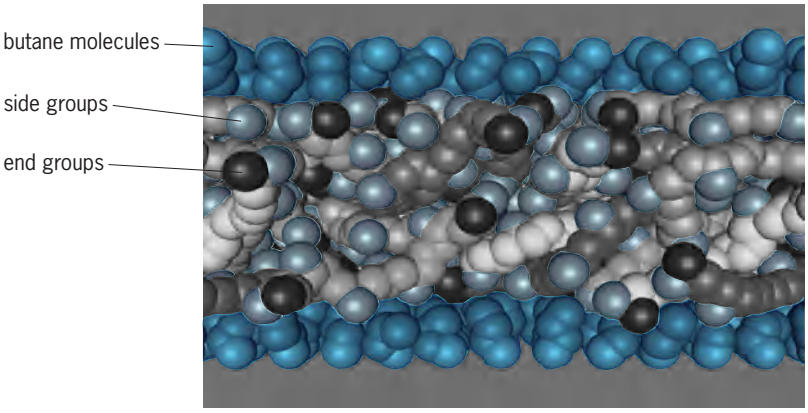

Fig. 1. A lubricant component, squalane (a 30-carbon alkane with a 24-carbon backbone and 6 methyl side groups), is confined to a nanoscale gap created by a surface to which butane molecules have been attached. The visualization depicts the carbons only. Molecules are shaded differently to distinguish them. The goal of such simulations is to understand the changes in lubrication properties induced by confinement to such narrow gaps.

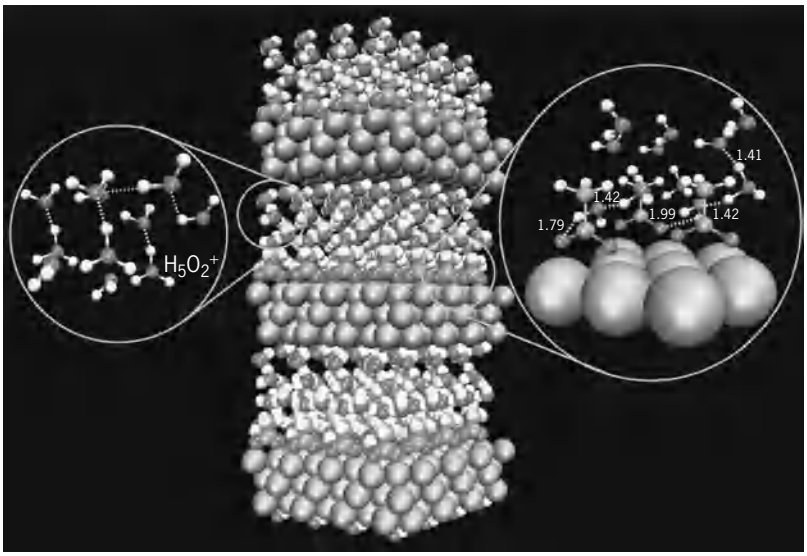

Fig. 2. Snapshot from an ab initio molecular dynamics simulation probing the effect of water on the dissociation of acetic acid over Pd(III). The change in binding energy of the acetate anion on Pd(III) in the liquid phase is measured in the simulation to be −142 kJ/mol, while the same quantity in the vapor phase (that is, in the absence of the aqueous solvent) is −212 kJ/mol. (*Simulation performed by, and image courtesy of, Matt Neurock, University of Virginia*)

The rapidly evolving capabilities of molecular simulation and computational quantum chemistry, and hybrids of these methods, suggest that they are beginning to fulfill their promise as equal and complementary partners with experiments in chemical and related industries.

Peter J. Cummings

Bibliography. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, Oxford, 1987; R. Car and M. Parrinello, Unified approach for molecular dynamics and density-functional theory, *Phys. Rev. Lett.*, 55:2471, 1985;

P. T. Cummings et al., *Report: NSF Workshop on Future Directions for Molecular Modeling and Simulation: Fundamentals and Applications*, National Science Foundation, Arlington, VA, November 3–4, 1997; D. J. Evans and G. P. Morriss, *Statistical Mechanics of Nonequilibrium Liquids*, Academic Press, New York, 1990; A. Z. Panagiotopoulos, Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble, *Mol. Phys.*, 61:813, 1987; A. Z. Panagiotopoulos et al., Phase equilibria by simulation in the Gibbs ensemble: Alternative derivation, generation, and application to mixture and membrane equilibria, *Mol. Phys.*, 63:527, 1988; *Technology Vision 2020: The U.S. Chemical Industry*, a joint report, 1996; S. Sarman, D. J. Evans, and P. T. Cummings, Recent developments in non-equilibrium molecular dynamics, *Phys. Rep.*, 305:1–92, 1998.

Molecular structure and spectra

Until the advent of quantum theory, ideas about the structure of molecules evolved gradually from analysis and interpretation of the facts of chemistry. Chemists developed the concept of molecules as built from atoms in definite proportions, and identified and constructed (synthesized) a great variety of molecules. Later, when the structure of atoms as built from nuclei and electrons began to be understood with the help of quantum theory, a beginning was made in explaining why atoms can combine in definite ways to form molecules; also, infrared spectra began to be used to obtain information about the dimensions and the nuclear motions (vibrations) in molecules. However, a fundamental understanding of chemical binding and molecular structure became possible only by application of the present form of quantum theory, called quantum mechanics. This theory makes it possible to obtain from the spectra of molecules a great deal of information about the nature of molecules in their normal as well as excited states, and about dissociation energies and other characteristics of molecules. For an important aspect of molecular structure which is treated separately *see* CHEMICAL BONDING.

Molecular sizes. The size of a molecule varies approximately in proportion to the numbers and sizes of the atoms in the molecule. Simplest are diatomic molecules. These may be thought of as built of two spherical atoms of radii r and r', flattened where they are joined. The equilibrium value R_e of the distance R between their nuclei is then smaller than the sum of the atomic radii (**Fig. 1**). However, the nuclei of atoms in two different molecules cannot normally approach more closely than a distance $r + r'$; r and r' are called the van der Waals radii of the atoms. The smallest molecule is hydrogen (H_2), with two electrons whose negative charges equal the positive charges of the two nuclei. Here r is about 0.12 nanometer giving $r + r' = 0.24$ nm, but R_e is only 0.074 nm. In HCl $r = 0.12$ nm for H and 0.18 nm for Cl, but R_e is only 0.127 nm.

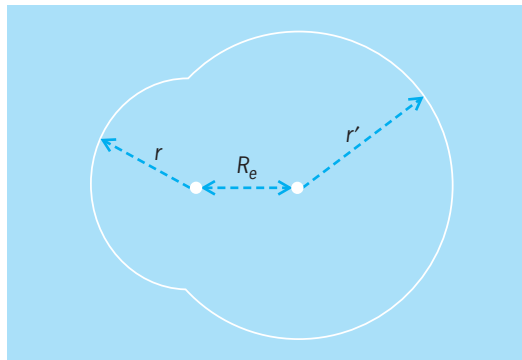

Fig. 1. Diatomic molecule with nuclei at distance R_e apart, built from atoms of radii r and r'.

To describe a polyatomic molecule, one must specify not merely its size but also its shape or configuration. For example, carbon dioxide (CO_2) is a linear symmetrical molecule, the O—C—O angle being 180°. The H—O—H angle in the nonlinear water (H_2O) molecule is 105°. Many molecules which are essential for life contain thousands or even millions of atoms. Proteins are often coiled or twisted and cross-linked in curious ways which are important for their biological functioning.

Dipole moments. Most molecules have an electric dipole moment. In atoms, the electron cloud surrounds the nucleus so symmetrically that its electrical center coincides with the nucleus, giving zero dipole moment; in a molecule, however, these coincidences are disturbed, and a dipole moment usually results.

Thus, when the atoms of HCl come together, there is some shifting of the H-atom electron toward the Cl. A complete shift would give H^+Cl^-, which would constitute an electric dipole of magnitude eR_e, where e is the electronic charge. But in fact the dipole moment is only $0.17eR_e$. This is because the actual electronic shift is only fractional. *See* ELECTRONEGATIVITY.

Although in molecules such as H_2, N_2, and CO_2 partial shifts of electronic charge from the original atoms do take place, these necessarily occur so symmetrically that no dipole moment results. Many larger molecules also have zero dipole moments by virtue of high symmetry. Examples are methane (CH_4), uranium hexafluoride (UF_6), and benzene (C_6H_6).

In general, the dipole moment of a neutral molecule is defined as the vector sum of quantities $+Q\mathbf{S}$ for the nuclei and $-e\mathbf{s}$ for the electrons. Here Q is the charge on any nucleus and $\mathbf{S}$ its vector distance from any fixed point in the molecule; $\mathbf{s}$ is the average vector distance of any electron from the same point. To calculate a dipole moment with these definitions, quantum mechanics must be used.

However, a study of what is known experimentally about molecular dipole moments has led to useful semiempirical generalizations. Bond moments and group moments have been obtained for various types of chemical bonds and of chemical groups or radicals. By adding these vectorially, the actual dipole moment of a large molecule can often be reproduced

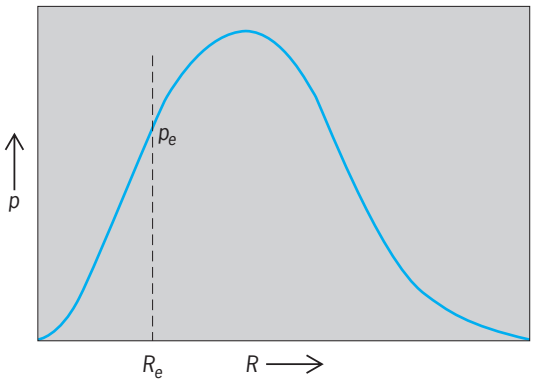

Fig. 2. Electric dipole moment p of typical diatomic molecule as function of internuclear distance *R*; p_e is the dipole moment at the equilibrium distance R_e.

fairly accurately. In CH_4 or CO_2, one can assume a moment for each C—H or C═O bond, even though these cancel out vectorially to give a zero resultant. In the linear molecule OCS, the unequal moments of the C═O and C═S bonds give a nonzero resultant. Because of the zero moment of CH_4, the CH bond moment and the CH_3 group moment must be equal and opposite. In CH_3Cl, the total moment can be thought of as the vector sum of the H_3C group moment and the C—Cl bond moment. *See* PERMITTIVITY.

When molecules vibrate, their dipole moments usually vary. **Figure 2** shows how the dipole moment **p** in a diatomic molecule may vary with *R*; the quantity previously discussed is $\mathbf{p}_e$, the value of **p** at R_e. When $\mathbf{p}_e$ is zero because of symmetry, it remains zero for symmetrical vibrations but, in polyatomic molecules, varies during unsymmetrical vibrations.

Molecules may possess magnetic as well as electric dipole moments. *See* DIPOLE MOMENT; MICROWAVE SPECTROSCOPY.

Molecular polarizability. In the preceding consideration of dipole moments, the discussion has been in terms of atoms and molecules free from external forces. An electric field pulls the electrons of an atom or molecule in one direction and pushes the nuclei in the opposite direction. This action creates a small induced dipole moment, whose magnitude per unit strength of the field is called the polarizability.

Molecular polarizabilities can be expressed as sums of atomic polarizabilities, plus corrections depending on the types of bonds present. Polarizabilities increase rather rapidly in such series as F, Cl, Br, I, and also from HF to HI, or F_2 to I_2.

Molecular polarizabilities can also be expressed approximately as sums of bond polarizabilities. These polarizabilities are anisotropic, being greater along bonds than perpendicular to bonds.

Molecular energy levels. The stationary states of motion of nuclei and electrons in a molecule, or of electrons in an atom, are restricted by quantum mechanics to special forms with definite energies. (Nonstationary states, which vary in the course of time, are constructed by mixing stationary states of different energies.) The state of lowest energy is called the ground state; all others are excited states. In analogy to water levels in a river, the energies of the stationary states are called energy levels. Excited states exist only momentarily, following an electrical or other stimulus. *See* ENERGY LEVEL (QUANTUM MECHANICS); NONRELATIVISTIC QUANTUM THEORY; QUANTUM CHEMISTRY; QUANTUM MECHANICS.

Energy levels are either discrete or continuous. The levels of a self-contained atom or molecule are restricted to special, sharply defined values (discrete levels). When an atom or molecule is ionized, that is, when one of its electrons has enough energy to escape completely, the energy can take on any value exceeding the minimum escape energy. This range of energies is called a continuous level or an ionization continuum. Molecules also have dissociation continua, which are discussed below.

Excitation of an atom consists of a change in the state of motion of its electrons. Electronic excitation of molecules can also occur, but alternatively or additionally, molecules can be excited to discrete states of vibration and rotation.

A diatomic molecule can vibrate similar to the motion of two masses separated by a distance *R* connected by a spring. In a diatomic vibration, *R* varies periodically above and below R_e, the equilibrium internuclear distance. The possible vibration energies E_υ (in joules) are given by Eq. (1),

$$E_\upsilon/h = \omega_e(\upsilon + {}^1/_2) - \omega_e x_e(\upsilon + {}^1/_2)^2 + \cdots \qquad (1)$$

where ω_e is just the small-amplitude vibration frequency (in hertz, with dimensions s^{-1}), and *h* is the Planck constant (6.626×10^{-34} J·s); $\omega_e x_e$ is a small quantity (relative to ω_e) that is nearly always positive. The vibrational quantum number υ can take whole number values 0, 1, 2, and so forth. The $+ \cdots$ in Eq. (1) indicates small correction terms. The zero-point vibration energy ${}^1/_2 h\omega_e - {}^1/_4 h\omega_e x_e$ present in the ground vibration state ($\upsilon = 0$) is a characteristic manifestation of quantum phenomena.

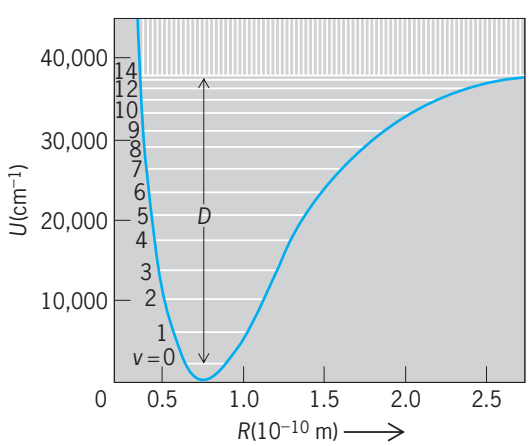

Fig. 3. *U*(*R*) curve of ground electronic state of H_2 with vibrational levels and dissociation continuum. *D* indicates the dissociation energy. Maximum *v* here is 14. (*After G. Herzberg, Molecular Spectra and Molecular Structure, vol. 1, 2d ed., Van Nostrand, 1950*)

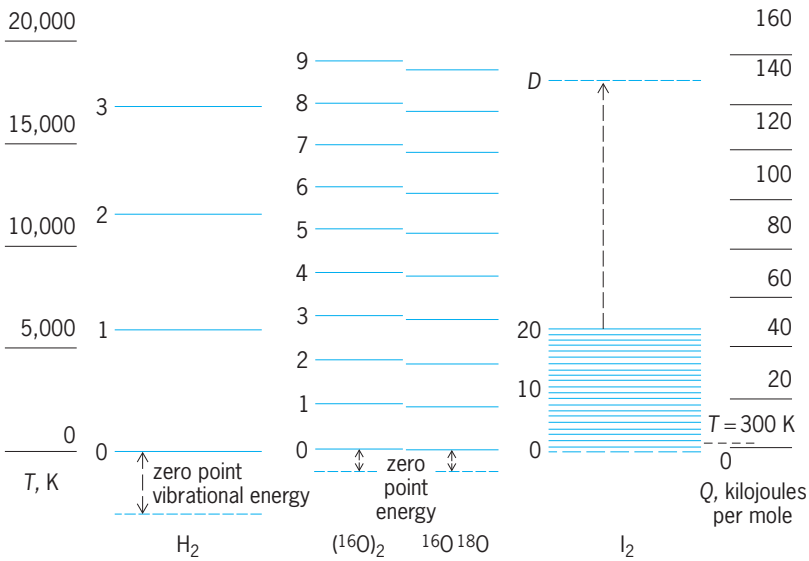

Fig. 4. Lowest vibrational levels of H_2, O_2, and I_2, numbered by vibrational quantum number *v*. Vibration level spacings decrease with increasing *v*. Where spacing reaches zero, the molecule dissociates; dissociation level *D* is indicated for I_2. Energies are given by the scale at right. The scale at left shows the average energy of vibration at various temperatures. °F = (K × 1.8) − 459.67.

Units can be confusing in spectroscopy because energies, vibrational frequencies, and other quantities are often expressed in wavenumbers ($\tilde{\nu}$), traditionally defined as the number of wavelengths per centimeter (that is, $E = h\nu = hc\lambda^{-1} = 100hc\tilde{\nu}$). In physics the symbol ω is generally an angular frequency (units of radians s^{-1}) given by $2\pi\nu$, but in spectroscopy ω_e is a frequency or wavenumber given typically in megahertz (hertz/10^6) or cm^{-1} units. It is also customary to treat $\omega_e x_e$ as a single symbol, rather than as product of ω_e and x_e.

The value of ω_e (in hertz) depends on the masses m_1 and m_2 of the atoms and the force constant k, as shown in Eq. (2), with the reduced mass, μ, defined

$$\omega_e = \frac{1}{2\pi}\sqrt{k/\mu} \qquad (2)$$

as $m_1 m_2/(m_1 + m_2)$.

The vibrational motion of a diatomic molecule is governed by the potential energy curve, which shows how the energy of attraction $U(R)$ of the atoms varies with R (**Fig. 3**). The quantities R_e, k, and the dissociation energy D are the most important properties of a potential curve; k is d^2U/dR^2 taken at R_e. The $U(R)$ curve and vibrational levels for the ground electronic state of H_2 are shown in Fig. 3. Similar curves, but with other R_e, k, and D values, exist for other electronic states and other molecules. Molecules have also repulsive electronic states, whose $U(R)$ curves rise steadily with decreasing R. These are often important for spectroscopy and in atomic collisions. For stable (attractive) $U(R)$ curves, the vibrational levels decrease in spacing as υ increases, until finally, as the spacing approaches zero, a maximum υ is reached just before dissociation; in Fig. 3 this is 14. After a small gap, a dissociation continuum of energy levels then sets in. Here the atoms have enough mutual kinetic energy to fly apart. For repulsive states, there is only a dissociation continuum, with no vibrational levels. **Figure 4** illustrates how strongly vibration level spacings can vary: both k and $1/\mu$, and therefore ω_e, decrease from H_2 to O_2 to I_2. Figure 4 likewise illustrates the effect of mass in isotopic molecules called isotopologues.

The total internal energy of a molecule is a sum of several terms: the largest is the electronic energy, E_{el}, followed by the vibrational energy, E_υ. The total energy can be written as Eq. (3). Both the electronic

$$E = E_{el} + E_\upsilon + (E_r + E_{fs} + E_{hfs} + E_{ext}) \qquad (3)$$

energy E_{el} and vibration energy E_υ can be discrete or continuous. The quantities E_r, E_{fs}, and E_{hfs} denote rotational, fine-structure, and hyperfine-structure energies, respectively. The last two appear as small or minute splittings of the rotation levels. Fine-structure splittings are caused by any unpaired electrons, and hyperfine splittings result from the presence of nuclear spins. The spacings ΔE of adjacent discrete levels of each type are usually in the order given in Eq. (4).

$$\Delta E_{el} \gg \Delta E_\upsilon \gg \Delta E_r \gg \Delta E_{fs} \gg \Delta E_{hfs} \qquad (4)$$

The fine structures of rotational levels differ strongly for different types of electronic states. The simplest diatomic electronic states are called $^1\Sigma$ states, and include $^1\Sigma^+$ and $^1\Sigma^-$ types for heteropolar and $^1\Sigma_g^+$, $^1\Sigma_u^+$, $^1\Sigma_g^-$, and $^1\Sigma_u^-$ for homopolar molecules. Most even-electron diatomic and linear polyatomic molecule ground states are $^1\Sigma^+$ states ($^1\Sigma_g^+$ if homopolar). The rotational levels of $^1\Sigma$ states have no fine structure; hyperfine structure, because of interaction with nuclear spins, is usually on too small a scale to detect by optical or infrared spectroscopy, but is often seen by rotational

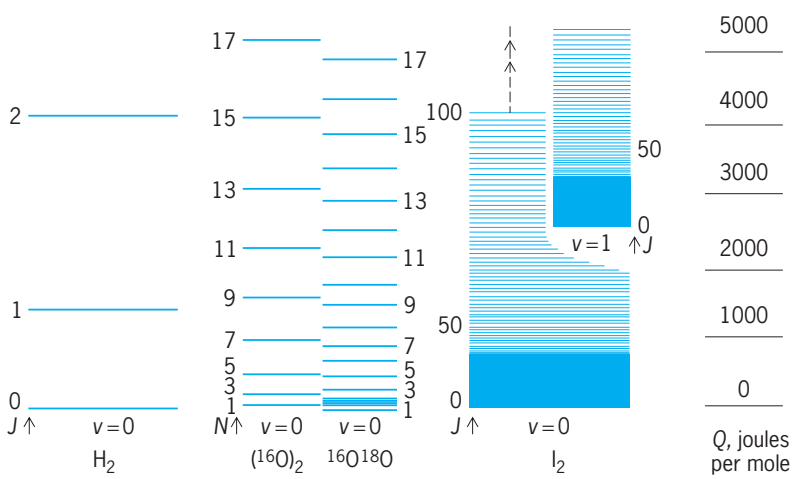

Fig. 5. Lowest rotational levels of H_2, O_2, and I_2. For H_2 and I_2, *J* is the rotational quantum number, according to Eq. (5) in the text. O_2 is in a Hund's case *b* triplet state, and the rotational levels are designated by *N*, where the total angular momentum $J = N + 1$, *N*, and $N - 1$. This narrow spin tripling is indicated for the $N = 1$ level of $^{16}O^{18}O$ only. Energies are given by the scale shown at right.

spectroscopy in the microwave region. The E_{ext} term in Eq. (3) refers to additional fine structure which appears on subjecting molecules to external magnetic fields (Zeeman effect) or electric fields (Stark effect). *See* FINE STRUCTURE (SPECTRAL LINES); HYPERFINE STRUCTURE; STARK EFFECT; ZEEMAN EFFECT.

The rotational levels of any $^1\Sigma$ state are given by Eq. (5). The quantity B_v is related to the moment of inertia I ($I = m_1m_2R^2/(m_1 + m_2) = \mu R^2$ for a diatomic molecule), and to v, by Eq. (6). The small positive quantity D_v in Eq. (5) is the centrifugal distortion constant that accounts for the stretching of the bonds by centrifugal forces as the molecule rotates. The rotational quantum number J can have any whole number value from 0 up, and corresponds to an angular momentum of magnitude $\sqrt{J(J+1)}(h/2\pi)$. The averaging of $1/I$ in Eq. (6) normally results in a slow decrease of B with increasing v (α_e is usually a small positive quantity). The quantity B_e refers to a hypothetical nonvibrating molecule ($R = R_e$). The B, D, and α_e in constants in Eqs. (5) and (6) are given in frequency units (hertz), and the conversion to the traditional wavenumbers ($\tilde{\nu}$ in cm^{-1}) is again accomplished with the equations: $E = h\nu = hc\lambda^{-1} = 100hc\tilde{\nu}$.

$$E_r/h = B_vJ(J+1) - D_v[J(J+1)]^2 + \cdots \quad (5)$$

$$B_v = \frac{h}{8\pi^2}\overline{(1/I)}_v = B_e - \alpha_e(v + 1/2) + \cdots$$
$$= B_0 - \alpha_e v + \cdots \quad (6)$$

Figure 5 illustrates how enormously rotational level spacings can vary because of differences in m and R_e (both are much greater for I_2 than H_2). The effect of mass for isotopic molecules is illustrated for O_2. Comparison with Fig. 4 illustrates the relation $\Delta E_v \gg \Delta E_r$ mentioned earlier.

Polyatomic molecules have much more complicated patterns of vibrational and (usually) of rotational energy levels than diatomic molecules. The number of normal modes (independent motions) of vibration for a molecule with n atoms is $3n - 6$ for nonlinear molecules, and $3n - 5$ for linear molecules. Each normal mode is a cooperative vibration of some or all the atoms moving with the same frequency, characteristic of the mode. Sometimes two or even three modes are so related in form that their frequencies are identical. These are called degenerate vibrations.

Figure 6 depicts the normal modes of H_2O and CO_2. They are labeled by symbols which also denote their frequencies. The arrows indicate the directions of motion of the atoms during one phase of vibration. The CO_2 frequency ν_2 is twofold degenerate: there are two independent modes corresponding to motion in either of two planes at right angles. The other two CO_2 modes, and all three H_2O modes, are nondegenerate.

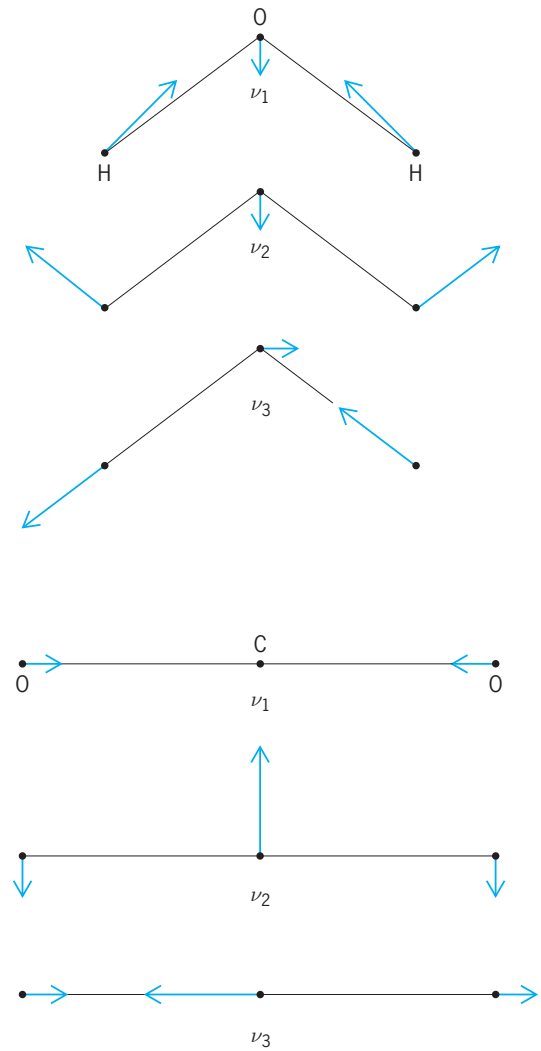

Fig. 6. Normal vibration modes of H_2O and CO_2. Synchronized displacements of atoms occur in proportion to lengths of the arrows. Diagram corresponds to snapshot taken at one phase of vibration.

Molecular spectra. Radiation emitted by a molecule initially in an upper state level E' as it makes a transition to a lower state level E'' obeys the Einstein-Bohr equation (7). Molecular emission

$$h\nu = E' - E'' \quad (7)$$

spectra accompany jumps in energy from higher to lower levels; absorption spectra accompany transitions from lower to higher levels. Both E' and E'' can be either discrete or continuous levels. If both are discrete, they give a spectrum of discrete frequencies; otherwise, they give a continuous spectrum. Discrete spectra are the main type considered here. Discrete frequencies are usually called "lines" because of their appearance when recorded by an optical spectrograph using photographic plates. The wealth and precision of spectroscopic observations have been increased by orders of magnitude by the advent of laser sources and Fourier transform spectrometers. *See* LASER SPECTROSCOPY.

Besides its frequency, the intensity and width of a spectral line are important. Intensities vary over wide ranges. In the extreme case of nearly zero intensity for a spectroscopic transition, the transition is called

forbidden. Only a small minority of all pairs of levels yield allowed transitions. These are governed by selection rules derivable from quantum theory. *See* SELECTION RULES (PHYSICS).

Under disturbing influences, however, some lines that violate these rules are seen weakly. Further, the usual selection rules are electric dipole rules, and additional transitions become very weakly allowed if magnetic dipole, electric quadrupole, and other selection rules are also considered. The following discussion is confined to spectra which obey the electric dipole rules.

Molecular spectra can be classified as fine-structure or low-frequency spectra, rotation spectra, vibration-rotation spectra, and electronic spectra. Low-frequency spectra are discussed elsewhere. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; MAGNETIC RESONANCE; MICROWAVE SPECTROSCOPY; MOLECULAR BEAMS; SPECTROSCOPY.

Pure rotation spectra. Transitions between energy levels differing only in rotational state give rise to pure rotation spectra. For diatomic molecules in $^1\Sigma$ states, Eq. (5), the relation is given by Eq. (8). The transitions

$$\nu = (E_r' - E_r'')/h = B_v[J'(J'+1) - J''(J''+1)] + \cdots \quad (8)$$

obey the selection rule $\Delta J = 1$ (ΔJ means $J' - J''$). Setting $J' = J'' + 1$, Eq. (9) is obtained. Equation (9)

$$\nu = 2B_v(J''+1) + \cdots \quad (9)$$

represents a sequence of lines spaced almost equidistantly ($2B_v, 4B_v, 6B_v, \ldots$), and lying in the far infrared or (for small B or low J'') the microwave region. Their intensities are proportional to $|\mathbf{P}_e|^2$, where $\mathbf{P}_e$ is the electric moment at R_e (Fig. 2); hence homopolar molecules (H_2, N_2, and so on) show no pure rotation spectra. The intensities are proportional also to the lower-state (v'', J'') level population and to ν (for absorption) or ν^4 (for emission).

Pure rotation spectra of linear polyatomic molecules are like those of diatomic molecules. Polyatomic molecules having $\mathbf{P}_e = 0$, whatever their shape (examples are CO_2, CH_4, C_6H_6), have no pure rotation spectra. In other cases, the spectra can be obtained using $h\nu = E_r' - E_r''$ with appropriate E_r expressions and selection rules.

Vibration-rotation bands. Spectra involving only vibrational and rotational state changes lie mainly in the infrared. For a $^1\Sigma$ diatomic state, using Eqs. (1), (5), and (7), Eq. (10) is obtained, with ν_0 defined in Eq. (11). In Eq. (10) B' and B'' mean B_v for v' and v'',

$$\nu = \nu_0(v', v'') + B'J'(J'+1) - B''J''(J''+1) + \cdots \quad (10)$$

$$\nu_0(v', v'') = (\omega_e - \omega_e x_e)(v' - v'') - \omega_e X_e(v'^2 - v''^2) + \cdots \quad (11)$$

respectively. Each band consists of two sets of rotational lines, one on each side of ν_0, which is called the band origin. Each line corresponds to a particular rotational transition conforming to $\Delta J = \pm 1$. The two series (branches) have frequencies defined in Eq. (12) for the R or positive branch ($J' = J'' + 1$), and in Eq. (13) for the P or negative branch ($J' = J'' - 1$). Both can be represented by a single equation, Eq. (14), by letting $m = J'' + 1$ for the R and

$$\nu = \nu_0 + (B' + B'')(J''+1) + (B' - B'')(J''+1)^2 + \cdots \quad (12)$$

$$\nu = \nu_0 - (B' + B'')J'' + (B' - B'')J''^2 + \cdots \quad (13)$$

$$\nu = \nu_0 + (B' + B'')m + (B' - B'')m^2 + \cdots \quad (14)$$

$m = -J''$ for the P branch. Neglecting the term in m^2, Eq. (14) represents a series of equidistant lines with one missing ($m = 0$) at ν_0. **Figure 7** shows how the line positions are related to the upper (v') and lower (v'') sets of rotational levels. (For transitions involving at least one non-$^1\Sigma$ state for diatomic and linear molecules as well as for all other polyatomic molecules, Q branches, corresponding to $\Delta J = 0$, can also exist.)

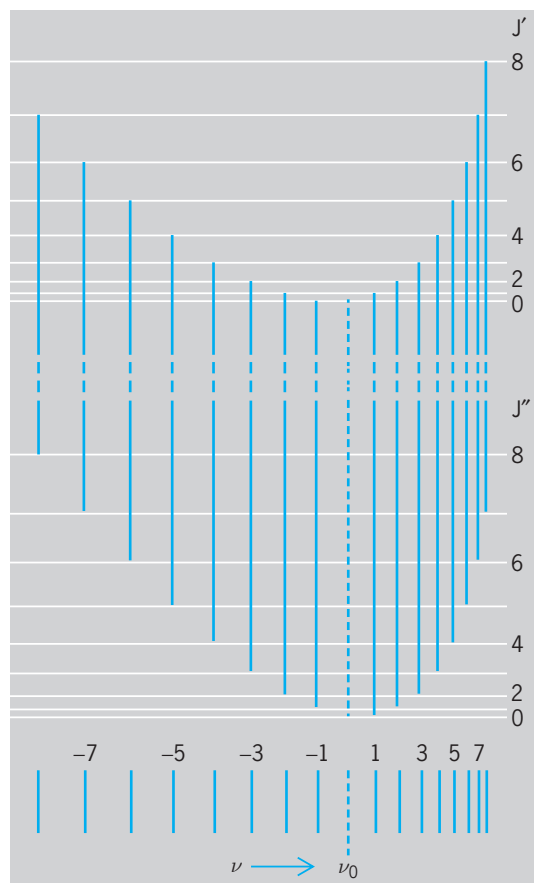

Fig. 7. Relation of band lines (lower part) [see Eqs. (8) and (9)] to rotational levels [see Eq. (6)] for a vibration-rotation band or an electronic band. In the former case, the upper and lower sets of rotational levels belong to two vibrational levels of a $^1\Sigma$ electronic state. In the latter case, they belong to two different $^1\Sigma$ states. Positive *m* values, *R* branch; negative *m* values, *P* branch.

Since $B' - B''$ is a small negative quantity [see Eq. (6), noting that $\upsilon' > \upsilon''$], the m^2 term makes the P line spacing increase and the R line spacing decrease slowly as m increases. This is shown, exaggerated, in Fig. 7. At some large m value, the R branch turns back on itself, but usually the lines have become weak before this value is reached.

The relative intensities of band lines depend primarily on the initial rotational distribution of molecules. More precisely, Eq. (15) holds. Here n in

$$\text{Intensity} = C(\upsilon', \upsilon'')\nu^n(J' + J'' + 1)e^{-E/(kT)} \quad (15)$$

ν^n is 4 for an emission band and 1 for an absorption band, while E is the upper state rotational energy level from which emission occurs or the lower state level for the case of absorption. **Figure 8** shows diagrammatically how the values of B and T affect the appearance of a typical absorption band ($B' = B''$ has been assumed for simplicity in Fig. 8). **Figure 9** shows the appearance of an actual DCl band. The weaker $D^{37}Cl$ lines are at slightly lower frequencies than the $D^{35}Cl$ lines, mainly because ω_e is smaller [see Eqs. (2) and (11)].

The factor $C(\upsilon', \upsilon'')$ is largest by far for fundamental bands (1, 0), and falls rapidly with increasing $\Delta\upsilon$ in the overtone bands or harmonics ($\Delta\upsilon$ is $\upsilon' - \upsilon''$). For fundamental bands, C depends on the slope of the $P(R)$ curve (Fig. 2), being approximately proportional to $(dP/dR)^2$ taken at R_e. For overtone bands, C depends on the detailed shapes of both the $P(R)$ and $U(R)$ curves. Fundamental or overtone bands arising from $\upsilon'' > 0$ are called hot bands.

Vibration absorption bands of liquids and solutions are widely used in chemical analysis. For many purposes, it is sufficient to know empirically the spectrum of each molecule which may be present. Also, groups of atoms which recur in many molecules often have nearly constant frequencies, of use for identification and in determining molecular structure. *See* INFRARED SPECTROSCOPY.

Electronic band spectra. These are the most general type of molecular spectra. The characteristic feature is a change of electronic state. From Eqs. (3) and (7), neglecting fine structure, Eqs. (16) and (17) are

$$h\nu = (E'_{el} - E''_{el}) + (E'_v - E''_v) + (E'_r - E''_r) \quad (16)$$

$$\nu = \nu_{el} + \nu_v + \nu_r = \nu_0 + \nu_r \quad (17)$$

obtained. Diatomic electronic spectra are often observed in emission, while the electronic spectra of polyatomic molecules are usually absorption spectra. Depending mainly on the magnitude of ν_{el}, electronic spectra occur in the infrared, visible, ultraviolet, or vacuum ultraviolet.

For any one electronic transition, the spectrum consists typically of many vibrational bands, each labeled by (υ', υ''). These lie in general at frequencies both above and below ν_{el}, since ν_v can be positive or negative. They constitute a band system. Each band consists of numerous rotational lines arranged in two or more branches and lying on both sides of a central position $\nu_0(\upsilon', \upsilon'')$, called the band origin.

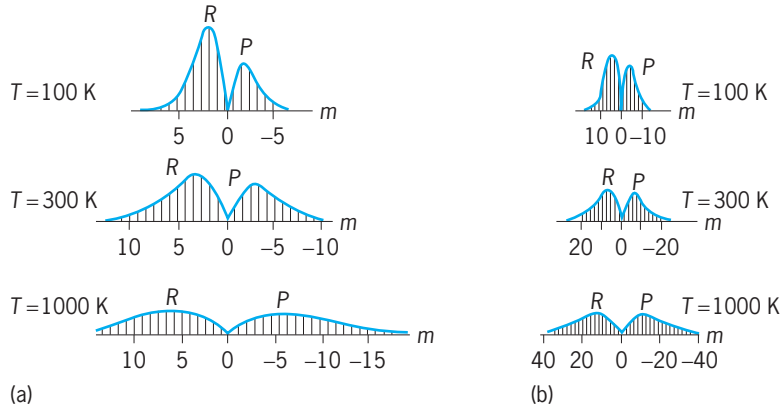

Fig. 8. Intensity distribution at several temperatures for a diatomic absorption band. Line positions are based on Eq. (9) assuming $B' = B''$ for simplicity; frequency increases toward the left (opposite to Fig. 7). (*a*) and (*b*) correspond respectively to B values of HCl ($B = 10.44\ \text{cm}^{-1}$) and of 2 cm^{-1} (approximately the value for CO, for which $B = 1.93\ \text{cm}^{-1}$). 100 K = −280°F; 300 K = 80°F; 1000 K = 1340°F. (*After G. Herzberg, Molecular Spectra and Molecular Structure, vol. 1, 2d ed., Van Nostrand, 1950*)

For diatomic molecules, the band origin ν_0 depends on a single υ' and υ'' and, using Eq. (1) for each electronic state, is given by Eq. (18). Equation (10)

$$\nu_0(\upsilon', \upsilon'') = \nu_{el} + [\omega'_e(\upsilon' + {}^1/_2) - \omega_e x'_e(\upsilon' + {}^1/_2)^2 + \cdots] - [\omega''_e(\upsilon'' + {}^1/_2) - \omega_e x''_e(\upsilon'' + {}^1/_2)^2 + \cdots] \quad (18)$$

also applies to electronic spectra but the band origin, $\nu_0(\upsilon', \upsilon'')$ is given by Eq. (18). The relative intensities of the bands depend on (1) the initial distribution of molecules among vibrational levels, and (2) the relative transition probabilities from any initial to various final vibrational levels.

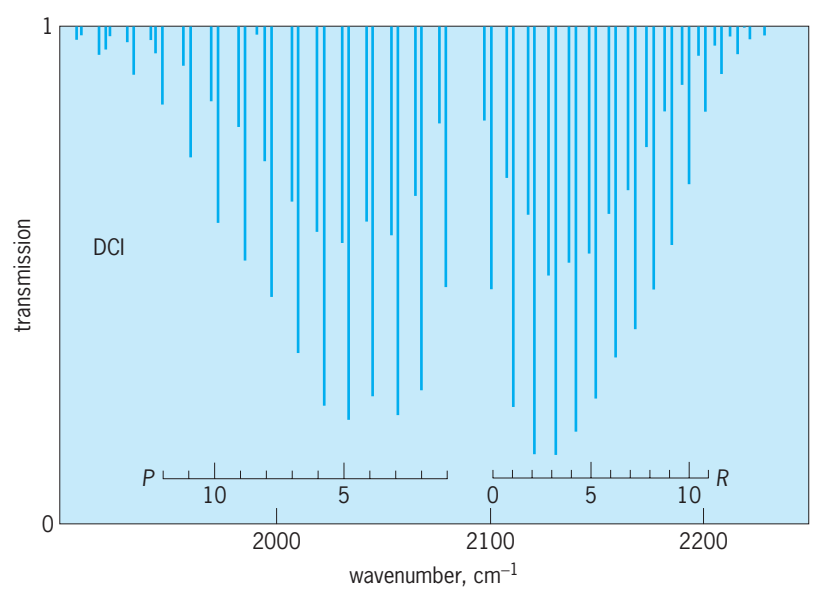

Fig. 9. Fundamental (1,0) vibration-rotation band of DCl in absorption showing the intensity distribution of the lines. The wavenumber scale increases to the right (in contrast to Fig. 8), so the R branch is to the right of the band origin, P branch to the left. The stronger lines are $D^{35}Cl$; the weaker companions, at lower wavenumbers, are $D^{37}Cl$. (*After P. F. Bernath, Spectra of Atoms and Molecules, 2d ed., Oxford University Press, 2005*)

The simplest example is the absorption spectrum of a cool gas of low molecular weight, for which all molecules initially have $\upsilon'' = 0$. The spectrum then consists of one "υ' progression," a single series of bands with various values of υ'; the frequencies are given approximately by $\nu = \nu_0(0, 0) + (\omega'_e - \omega_e x'_e)\upsilon' - \omega_e X'_e \upsilon'^2$, in which $\nu_0(0, 0)$ is the frequency of the (0, 0) band center. For a hot or a heavy gas, additional weaker υ' progressions with $\upsilon'' > 0$ also appear.

In emission spectra, the initial population usually ranges over a number of υ' values, from each of which transitions occur to a number of υ'' values, so that the system contains many bands on both sides of $\nu_0(0, 0)$. In the special case of fluorescence spectra, the molecule is excited to various υ' values by absorbing light; it then emits light belonging to the same (or sometimes another) electronic transition. From each υ', it can descend not only to the original υ'' but also to various other, mainly larger, values. Hence fluorescence bands lie mainly at lower frequencies than the absorption bands used to excite them. *See* FLUORESCENCE.

Relative transition probabilities are governed by the Franck-Condon principle. This takes note of the very great rapidity of electronic motions as compared with those of the far more massive nuclei, and concludes that during the extremely brief time for an electronic transition, the nuclei tend to remain unchanged in their positions and momenta. It is applicable to both polyatomic and diatomic spectra. Consider first a diatomic molecule starting from the $\upsilon'' = 0$ level of a ground state $U(R)$ curve like the lower curves in **Fig. 10**. A vertical line drawn from the bottom point A ($\upsilon'' = 0$ if zero-point vibration is neglected) to point B on any one of the upper curves corresponds to an electronic absorption transition in which the nuclei have not moved. *See* FRANCK-CONDON PRINCIPLE.

In the case of Fig. 10*a*, point B corresponds to $\upsilon' = 0$, and the conclusion is that this is the most probable υ' for $\upsilon'' = 0$. In the case of Fig. 10*b*, point B corresponds to an excited molecule at the inner turning point of a vibration with a υ' of possibly 3 or 4, in a typical case. One then concludes (with J. Franck) that the strongest absorption bands for $\upsilon'' = 0$ have $\upsilon'' = 3$ and 4. To obtain more exact information, a quantum-mechanical calculation (first carried out by E. U. Condon) is necessary.

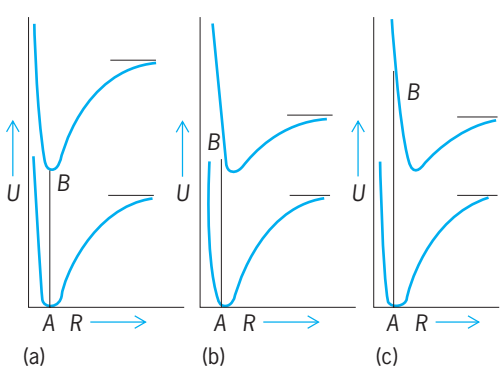

Fig. 10. Diatomic $U(R)$ curves for three cases to explain the vibrational intensity distribution according to the Franck-Condon principle. The asymptote of each curve for large R corresponds to dissociation into atoms, with one or both atoms excited in the case of the upper curves. Starting in each case from the bottom of the lower curve (essentially $v'' = 0$), the most probable transition in absorption is (*a*) to $v' = 0$, (*b*) to $v' = 3$ or 4, and (*c*) to the dissociation continuum, as shown by vertical lines. (*After G. Herzberg, Molecular Spectra and Molecular Structure, vol. 1, 2d ed., Van Nostrand, 1950*)

In the case of Fig. 10*c*, point B corresponds to an energy level in the dissociation continuum above the asymptote of the upper $U(R)$ curve. According to the Franck-Condon principle, the absorption spectrum will have maximum intensity in a continuous range of frequencies, with $h\nu$ about equal to the energy difference AB. The quantum-mechanical calculation shows that the actual spectrum will extend with appreciable intensity over a range of both higher and lower frequencies than this, including, on the lower-frequency side, a number of high-υ'' bands. Actual examples of such spectra (a long υ'' progression followed by a strong continuum) are the far-ultraviolet Schumann-Runge bands of oxygen and the visible bands of iodine. By measuring the frequency at which the continuum begins, one obtains an exact value of the dissociation energy of these molecules. In so doing, any excitation energy of the atomic dissociation products to which the upper $U(R)$ curve leads is subtracted.

The Franck-Condon method is useful in understanding intensity distributions and structure in emission as well as absorption band systems. For diatomic spectra, various patterns of intensity as functions of υ'' and υ'' occur, depending largely on the R_e values of the two $U(R)$ curves and, of course, also on the initial distribution among υ'' levels. Sometimes the upper-state $U(R)$ curve is stable (has a minimum) but the lower state is repulsive. Continuous emission spectra then occur, with the atoms flying apart on reaching the lower state. The H_2 molecule shows such a spectrum, as do rare gas molecules such as He_2 and Kr_2, which are stable only in excited or ionized states.

Molecular electronic states. Before discussing the structures of electronic bands, one must consider the nature of molecular electronic states. Each electronic state has orbital and spin characteristics. The spin quantum number S has a whole-number value if the number of electrons is even, a half-integral value if it is odd. Electronic states with $S = 0$ are called singlet states, all others multiplet states. The orbital characteristics differ sharply for linear (including diatomic) and nonlinear molecules.

For linear molecules only, there is a quantum number Λ such that $\pm\Lambda h/2\pi$ is the component of angular momentum around the line of nuclear centers. Linear-molecule electronic states can be discussed under three headings: (1) singlet states; (2) multiplet states with strong spin coupling (Hund's case *a*); and (3) multiplet states with weak spin coupling (Hund's case *b*). Strictly speaking, actual multiplet states are intermediate between cases *a* and *b*, or between these and certain other cases called *c* and

d. The discussion to follow is largely restricted to singlet electronic states.

Singlet states with $\Lambda = 0$ include $^1\Sigma^+$ and $^1\Sigma^-$ states: states with $\Lambda = 1, 2, \ldots$ are called $^1\Pi$, $^1\Delta$, and so on. In linear molecules with a center of symmetry (H_2, CO_2 and so on), one must further distinguish even and odd (g and u) states: $^1\Sigma_g^+$, $^1\Sigma_u^+$, $^1\Sigma_g^-$, $^1\Sigma_u^-$, $^1\Pi_u$, $^1\Pi_g$, $^1\Delta_g$, $^1\Delta_u$, etc. The rotational levels of singlet states obey formula (19). Here J is restricted to integral values equal to or greater than Λ.

$$E_r/h = B_v[J(J+1) - \Lambda^2] \cdots \tag{19}$$

For $\Lambda > 0$, each rotational level is a narrow doublet (Λ-doubling). Corresponding fine structure [see E_{fs} in Eq. (3)] can usually be detected in electronic bands, but (for ground states only) it can be much more accurately studied in low-frequency spectra. Hyperfine structure [see E_{hfs} in Eq. (3)] is usually on too small a scale to be detected in electronic band lines, but has been found in a few cases. Hyperfine structure is best studied in low-frequency spectra in the microwave region.

Electronic band structure. The simplest electronic bands occur for transitions between singlet electronic states. The possible types of electronic transitions are limited by the selection rule $\Delta\Lambda = 0, \pm 1$. The structure of $^1\Sigma - {}^1\Sigma$ bands is essentially the same as for the $^1\Sigma$-state vibration-rotation bands described earlier. Equations (12) to (15) and Fig. 7, also Fig. 8, for the intensities in absorption are still applicable if Eq. (18) instead of Eq. (11) is used for the band origin ν_0, and it is recognized that B' and B'' now belong to two different electronic states.

The quantity $B' - B''$ in Eq. (14), instead of always being a small negative quantity, may now be either positive or negative, and $(B' - B'')/(B' + B'')$ is often fairly large (although it can also be nearly zero). As a result, it is usual in electronic bands to find so-called heads. A head is a position of maximum or minimum frequency; by using Eq. (14) to obtain $d\nu/dm = 0$, one finds $m_{\text{head}} = (B' + B'')/[2(B' - B'')]$. Then, on inserting m_{head} into Eq. (14), one obtains $\nu_{\text{head}} = \nu_0 - (B' + B'')^2/[4(B' - B'')]$. [Since $(B' + B'')/[2(B' - B'')]$ is not usually a whole number, the actual m_{head} is the nearest whole number m to that calculated.] According to whether $B' - B''$ is negative or positive, the positive (R) or the negative (P) branch forms the head. Figure 7, if continued to somewhat larger m values, illustrates the formation of an R-branch head at a calculated m of 10.5; the actual head is formed by the two coincident lines $m = 10$ and 11.

Although homopolar molecules (H_2, N_2, and so on) have no pure rotation or vibration-rotation spectra, they do have electronic spectra. For homonuclear homopolar molecules, the band lines show alternating intensities. The lines in each branch are alternately stronger and weaker as m increases, this effect being superposed on the otherwise smoothly varying intensity distribution. The alternation ratio depends on the nuclear spin I and has been, in several cases, the means of determining I. When $I = 0$, alternate lines are completely missing. Heteronuclear molecules, even if homopolar (for example, HD or $^{16}O\,^{18}O$) do not show alternating intensities.

Polyatomic electronic spectra. These differ from diatomic electronic spectra because several initial and final vibration quantum numbers are involved, and because the rotational structure (except for linear molecules) is usually much more complicated. However, the detailed structure of the electronic spectra of a number of simple molecules and radicals in the vapor state in emission and in absorption has been studied. Nevertheless, for the most part, the spectra of polyatomic molecules are examined as absorption spectra in solution. The rotational structure is then completely absent, but the vibrational structure can be seen.

The Franck-Condon principle is here a useful guide. One of its corollaries is that only totally symmetric vibrations (vibrations during which the equilibrium symmetry of the molecule is preserved) undergo quantum number changes. This greatly simplifies the vibrational structure, especially of absorption spectra where most molecules are initially mainly in the $\upsilon'' = 0$ state of all vibrations. One finds then mostly υ'' progressions of one or a very few totally symmetric vibrations, and combinations of these.

Rather often, polyatomic band systems do not even show obvious vibrational structure. This can happen for any of several reasons: The upper state may involve dissociation; in CH_3I, for example, the first ultraviolet absorption region yields $CH_3 + I$; there may be so many low-frequency, upper-state vibrations that the spectrum looks continuous; or there may be a combination of these and other reasons. Such continuous or pseudocontinuous band systems are often loosely referred to as bands. For complicated molecules, the spectra of several different electronic transitions often overlap strongly so that it is difficult even to separate one system from another. *See* ATOMIC STRUCTURE AND SPECTRA; ELECTRON SPIN; INTERMOLECULAR FORCES; MOLECULAR WEIGHT; NEUTRON SPECTROMETRY; RAMAN EFFECT; RESONANCE (MOLECULAR STRUCTURE); SCATTERING EXPERIMENTS (ATOMS AND MOLECULES); VALENCE.

Robert S. Mulliken; U. Fano; Peter Bernath

High-resolution spectra. Molecular spectra exhibit varying amounts of detail, depending upon the inherent structure of each band and upon the ability of the recording instrument to distinguish transitions of nearly equal wavelength. This ability, called the resolving power, has increased dramatically since about 1970. Dispersive spectrometers have been superseded by Fourier-transform interferometers and tunable lasers, improving resolution as much as a millionfold. The resulting insight into the details of molecular structure and dynamics has revolutionized the field. *See* INFRARED SPECTROSCOPY; LASER SPECTROSCOPY.

Figure 11 shows a thoroughly studied infrared vibration-rotation absorption band of sulfur hexafluoride (SF_6) at four stages of resolution. This transition is from the vibrational ground state ($\upsilon'' = 0$) to an excited state containing one quantum of the

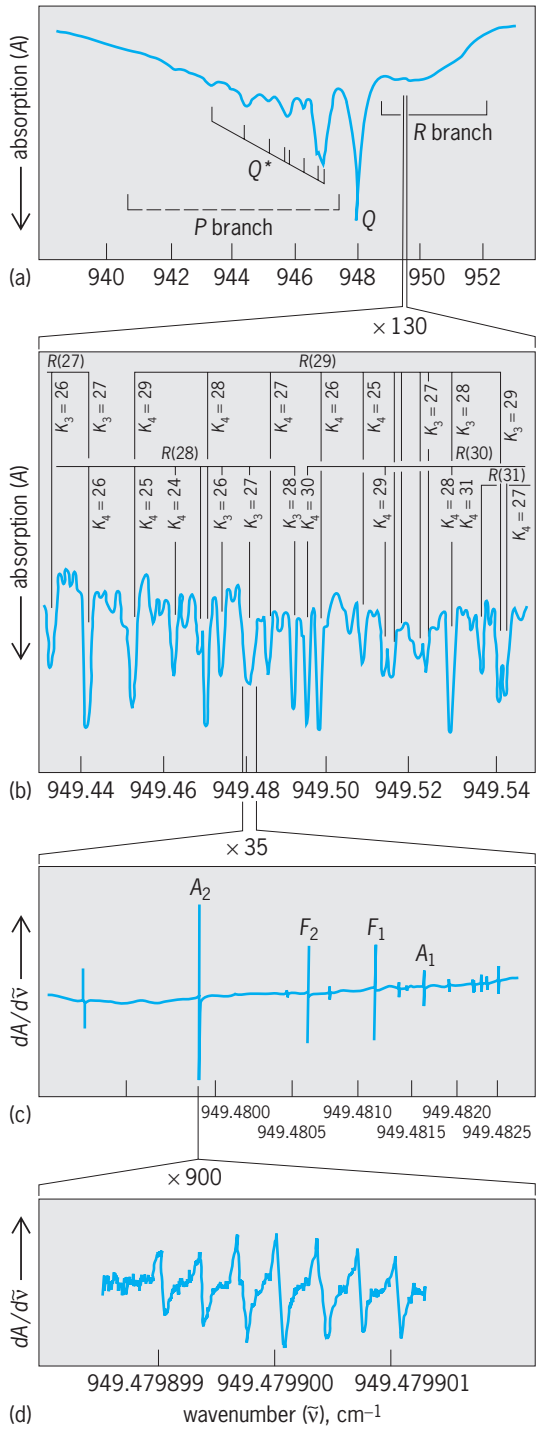

Fig. 11. Assignment of individual transitions in the spectrum of the S-F stretching fundamental ν_3 of sulfur hexafluoride (SF_6) as resolving power increases. (*a*) Fourier transform spectrum; resolution of 0.06 cm^{-1} [2 GHz]. (*b*) Spectrum recorded with tunable semiconductor diode laser; effective resolution limited by the SF_6 Doppler linewidth of 0. 0010 cm^{-1} [30 MHz]. (*Parts a and b after R. S. McDowell et al., Identification of the SF_6 transitions pumped by a CO_2 laser, Opt. Commun., 17:178–183, 1976*) (*c*) Saturation spectrum (derivative trace) recorded with a free-running waveguide carbon dioxide (CO_2) laser at a resolution of about 10^{-6} cm^{-1} [30 kHz]. (*d*) Saturation spectrum from a frequency-controlled CO_2 laser spectrometer with resolution of about 3×10^{-8} cm^{-1} [1 kHz]. (*Parts c and d after J. Bordé and C. J. Bordé, Superfine and hyperfine structures in the ν_3 band of $^{32}SF_6$, Chem. Phys., 71:417–441, 1982*)

infrared-active stretching fundamental v_3 ($v'' = v_3 = 1$). In Fig. 11*a* the unresolved band envelope appears with *P*, *Q*, and *R* branches, corresponding to series of transitions having $\Delta J = -1$, 0, and +1, respectively. The *P* branch is partially obscured by hot-band *Q* branches (labeled *Q**), prominent here because at room temperature 70% of the sulfur hexafluoride molecules are in excited vibrational states ($v'' > 0$) even without any optical excitation.

Higher resolution (Fig. 11*b*) reveals rotational fine structure in a portion of the *R* branch. In *R* (28), for example, the rotational quantum number changes from $J'' = 28$ to $J' = 29$ during the vibrational transition. Centrifugal forces will distort a rotating sulfur hexafluoride molecule from its equilibrium octahedral configuration to a lower symmetry, and so rotations about different molecular axes will have slightly different energies. The states resulting from this dynamic symmetry breaking are identified in Fig. 11*b* by semiclassical quantum numbers K_3 and K_4 that specify the angular momentum about the axes with threefold or fourfold symmetry. These $R(J'')$ manifolds are wider than their mutual separation, and their overlap produces the dense structure observed. Weaker unidentified absorptions arise from hot bands.

Saturation spectroscopy can probe below the Doppler limit to reveal further details. Since the angular momentum can be oriented in either direction along any of the four threefold or three fourfold symmetry axes of sulfur hexafluoride each K_3 or K_4 cluster has eightfold or sixfold degeneracy, respectively, which can be partly removed by quantum tunneling. Figure 11*c* shows this superfine splitting of the $R(28)$ ($K_3 = 27$) cluster into four transitions that are labeled according to the symmetry species of the octahedral rotational group. The $R(28)(K_3 = 27)$ cluster has the structure $A_2 + F_2 + F_1 + A_1$, and since the *F* species are triplets, this accounts for the required eight levels. These four species have statistical weights 10 + 6 + 6 + 2 due to nuclear spin degeneracy, which determines their relative intensities.

Finally Fig. 11*d* shows the $R(28)(K_3 = 27)$ A_2 transition at the highest resolution, revealing magnetic hyperfine structure arising from the interaction of the nuclear spin with the magnetic field of the rotating molecule. The tenfold A_2 spin degeneracy is resolved into a septet and a nearly coincident overlapping triplet.

Many thousands of lines in this one band of sulfur hexafluoride have been assigned to specific transitions, yielding very accurate molecular constants. For example, the moment of inertia implies a S—F bond length of 156.0588 picometers in the vibrational ground state (which differs very slightly from the equilibrium value R_e). The richness and complexity of such spectra pose a rigorous challenge to theoretical analysis, but they reveal subtleties of molecular energy levels and structure that were inaccessible before the development of high-resolution techniques.

Robin S. McDowell

Lasers and synchrotron sources. Modern sources of radiation have changed the field of molecular

spectroscopy enormously. Ultrafast pulsed lasers can offer time resolutions of a few femtoseconds (10^{-15} s) and very high peak powers of the order of terawatts (10^{12} W). Laser linewidths of a few hundred hertz for continuous-wave semiconductor lasers operating at 445 terahertz (674 nm) are readily achieved so that the spectral resolution is limited by Doppler or pressure broadening of the molecular lines of the gas sample. Special laser techniques such as saturation spectroscopy (as described above for SF_6) are required to take advantage of these exceedingly narrow laser linewidths. *See* LASER.

Synchrotron radiation also has many advantages for molecular spectroscopy. This radiation is emitted by a relativistic electron beam (that is, the electrons are traveling at close to the speed of light) that is accelerated in an electron storage ring. The required acceleration can be achieved with dipole magnets (bending magnets), which change the direction of motion of the electron beam, or with special "insertion" devices, which change the direction of motion of the electron beam many times in a short distance. An insertion device (wiggler or undulator) uses a periodic array of magnets (electromagnets or more commonly permanent magnets), with the direction of magnetic field alternating from magnet to magnet (say, field up and field down) to deflect the electrons back and forth through small angles as they travel through the magnetic array. The effect is to produce radiation that is a sum (a coherent sum in the case of an undulator, and an incoherent sum in the case of a wiggler) of the radiation from each deflection, resulting in high brightness (for an undulator) and high flux (for a wiggler). *See* SYNCHROTRON RADIATION.

Synchrotron radiation has long been used at short wavelengths, for example, in the photoionization of molecules and surfaces and in x-ray crystallography. More recently synchrotron radiation has been used also for infrared and far-infrared spectroscopy. Synchrotron radiation is typically used to replace the standard sources (the glowbar in the infrared and the mercury arc lamp in the far infrared) used for absorption spectroscopy with a Fourier transform spectrometer. The far-infrared radiation emitted from the edge of a bending magnet that turns the electron beam is many orders of magnitude brighter than that from a traditional thermal or arc source. In this context, brightness means power (watts) emitted per unit frequency interval, per unit emitting area, and per unit solid angle. The high brightness of the radiation can be used to improve the signal-to-noise ratio or the sensitivity of a typical infrared absorption measurement. Alternatively the high brightness can be used for very fast time resolution or for imaging.

In part the high brightness is due to the very small cross-sectional area (less than a millimeter) of the electron beam emitting the radiation. Synchrotron radiation (like a laser) can therefore be focused to a very small diffraction-limited spot (similar in size to the wavelength of the radiation) without excessive loss of radiation. This small spot can be used for hyperspectral imaging of biological or heterogeneous industrial samples. Hyperspectral imaging means that each pixel of a two-dimensional slice of a sample has an associated spectrum. Hyperspectral imaging creates a three-dimensional data cube. Two of the dimensions are the intensity values at the x and y coordinates of the pixels of the image of the sample. The z direction is then the spectral axis, and each pixel has a spectrum along this third dimension. A typical sample might be a slice of biological tissue and the measurements made with a microscope. The sample is moved in a raster pattern and a spectrum recorded at each position to generate the hyperspectral image. Lasers or synchrotron sources are not required for hyperspectral imaging, but the time required to generate a data cube is generally prohibitively long with conventional light sources.

Peter Bernath

Bibliography. C. N. Banwell and E. M. McCash, *Fundamentals of Molecular Spectroscopy*, 4th ed., 1994; P. F. Bernath, *Spectra of Atoms and Molecules*, 2d ed., 2005; W. Demtröder, *Laser Spectroscopy*, 3d ed., 2003; J. D. Graybeal, *Molecular Spectroscopy*, 1993; W. G. Richards and P. R. Scott, *Structure and Spectra of Molecules*, 1985; W. S. Struve, *Fundamentals of Molecular Spectroscopy*, 1989; S. Svanberg, *Atomic and Molecular Spectroscopy: Basic Aspects and Practical Applications*, 4th ed., 2004; A. P. Thorne, U. Litzén, and S. Johansson, *Spectrophysics: Principles and Applications*, 1999.

Molecular weight

The sum of the atomic weights of all atoms making up a molecule. Actually, what is meant by molecular weight is molecular mass. The use of this expression is historical, however, and will be maintained. The atomic weight is the mass, in atomic mass units, of an atom. It is approximately equal to the total number of nucleons, protons and neutrons, composing the nucleus. Since 1961 the official definition of the atomic mass unit (amu) has been that it is $^1/_{12}$ the mass of the carbon-12 isotope, which is assigned the value 12.000 exactly. *See* ATOMIC MASS; ATOMIC MASS UNIT; RELATIVE ATOMIC MASS; RELATIVE MOLECULAR MASS.

The microscopic atomic weight is connected to the fundamental macroscopic mass unit, the kilogram, through the definition of the mole (mol). The kilogram is defined as the mass of a platinum-iridium bar kept at the International Bureau of Weights and Measures in Paris, France. A mole is an amount of substance containing Avogadro's number, N_A, approximately 6.022×10^{23}, of molecules or atoms. Molecule, in this definition, is understood to be the smallest unit making up the characteristic compound. Originally, the mole was interpreted as that number of particles whose total mass in grams was numerically equivalent to the atomic or molecular weight in atomic mass units, referred to as gram-atomic or gram-molecular weight. This is how the

above value for N_A was calculated. As the ability to make measurements of the absolute masses of single atoms and molecules has improved, however, modern metrology is tending to alter its approach and define Avogadro's number as an exact quantity, thereby changing slightly the definition of the atomic mass unit and removing the need to define atomic weight with respect to a particular isotopic species. The latest and most accurate value for Avogadro's number is $6.0221415(10) \times 10^{23}$ mol^{-1}. *See* AVOGADRO'S NUMBER; MOLE (CHEMISTRY).

Mass spectrometry. The best values for microscopic atomic and molecular masses are derived from measurements utilizing mass spectrometry, an experimental technique employed both for mass detection and for absolute mass determinations based on deflection of charged particles in electric or magnetic fields. Of the two methods, magnetic deflection yields the most accurate results and is the one most often used. A magnetic deflection apparatus which is used for mass determination is called a mass spectrometer. While particular types may vary as to the details of construction, the principle of operation for all is based on the fact that a charged particle moving in a magnetic field will follow a curved path, the curvature of which depends on the mass and charge of the species.

For a spectrometer employing deflection in a constant magnetic field with induction $\vec{B}$, the curvature can be easily calculated. Let $\vec{v}_0$ be the initial velocity with which the particle enters the magnetic field region, m the mass of the particle, and q the charge. The magnetic force does not alter the magnitude of the particle's velocity; the sole action is to cause the particle to execute a circular motion in the field. If $\vec{v}_0$ is perpendicular to $\vec{B}$, so that $\vec{v}_0$ is equal to $\vec{v}_\perp$, the component of the velocity perpendicular to $\vec{B}$, the plane of the orbit will be perpendicular to the region in which $\vec{B}$ is constant. The radius of curvature of the circle r is then given by Eq. (1). Thus, the particle's mass is given by Eq. (2).

$$r = \frac{mv_\perp}{qB} \tag{1}$$

$$m = \frac{qrB}{v_\perp} \tag{2}$$

Mass spectroscopic measurements are made by first ionizing the atom or small molecule by using bombardment with high-energy electrons or absorption of ultraviolet light, then accelerating it across an electric potential difference V to produce a initial velocity given by Eq. (3), and finally allowing it to

$$v = \left(\frac{2qV}{m}\right)^{1/2} \tag{3}$$

be deflected in the magnetic field and strike a detector. With q known, m is extracted by measuring the radius of curvature r of the circle of motion, at least in principle. In practice, r and $\vec{B}$ will be equal to the charge on the electron, as most ionization will produce only the singly ionized species. As the mass of the electron is so much smaller than the mass of the proton, approximately 1/1836, the loss of one electron will not appreciably alter the mass of the atom or molecule.

Mass spectrometry has made possible determinations of absolute masses of virtually all the elements as well as many small molecules. The drawback to this technique is that the species to be measured must be in the vapor phase in order to execute motion in the fields. Its primary advantage is that it, taken together with Avogadro's number, puts the atomic mass scale on an absolute level. *See* MASS SPECTROMETRY; MASS SPECTROSCOPE.

Importance of determination. As the masses of all the atomic species are now known so well, masses of molecules can be determined once the composition of the molecule has been ascertained. Alternatively, if the molecular weight of the molecule is known and enough additional information about composition is available, such as the basic atomic constituents, it is possible to begin to assemble structural information about the molecule. Thus, the determination of the molecular weight is one of the first steps in the analysis of an unknown species. Given the increasing emphasis on the study of biologically important molecules, particular attention has been focused on the determination of molecular weights of larger and larger units. There are a number of methods available, and the particular one chosen will depend on the size and physical state of the molecule. All processes are physical macroscopic measurements and determine the molecular weight directly. Connection to the absolute mass scale is straightforward by using Avogadro's number, although, for extremely large molecules, this connection is often unnecessary or impossible, as the accuracy of the measurements is not that good. The main function of molecular weight determination of large molecules is elucidation of structure.

Weighing of gases. Although molecular weight determination for gases has been largely taken over by the mass spectroscopic analysis, it is still possible to determine molecular weights by direct weighing of gases. This method has some historical value in that much of present-day knowledge about chemical reactions was originally derived on the bases of observations concerning gases. For any real gas, if the pressure is low enough, the relationship between the pressure P, volume V, temperature T, and number of moles n of that gas can be expressed through the ideal gas law in the form of Eq. (4). Here R is the uni-

$$R = \lim_{P \to 0} \frac{PV}{nT} \tag{4}$$

versal gas constant, equal to 8.31441 J/mol K. The number of moles of the gas is related to the molecular weight by Avogadro's number, as discussed above. Thus, by measuring the pressure and weight of a known volume of gas, it is possible to derive the molecular weight M of the gas from Eq. (5). Here

$$M = RT \lim_{P \to 0} \frac{\rho}{P} \tag{5}$$

ρ is the mass density, equal to the total mass of the gas as measured in the experiment divided by the volume V. *See* GAS.

Colligative properties of solutions. Molecular weight determination of materials which are solid or liquid at room temperature is best achieved by taking advantage of one of the colligative properties of solutions, boiling-point elevation, freezing-point lowering, or osmotic pressure, which depend on the number of particles in solution, not on the nature of the particle. The choice of which to use will depend on a number of properties of the substance, the most important of which will be the size. All require that the molecule be small enough to dissolve in the solution but large enough not to participate in the phase change or pass through a semipermeable membrane. Freezing-point lowering is an excellent method for determining molecular weights of smaller organic molecules, and osmometry, as the osmotic pressure determination is called, for determining molecular weights of larger organic molecules, particularly polymeric species. Boiling-point elevation is used less frequently. *See* POLYMER.

The basis of all the methods involving colligative properties of solutions is that the chemical potentials of all phases must be the same. (Chemical potential is the partial change in energy of a system as matter is transferred into or out of it. For two systems in contact at equilibrium, the chemical potentials for each must be equal.) *See* CHEMICAL EQUILIBRIUM; CHEMICAL THERMODYNAMICS.

Phase change methods. For the two methods involving a phase change, analysis requires that it be possible to form a solution of the molecule in some known solvent which is dilute enough so that the solution may be considered ideal but concentrated enough that a measurable raising of the boiling point or lowering of the freezing point will be produced. It must also be possible to make the assumption that there are no solute particles in either the vapor or the solid phase. The balance of chemical potentials then demands that the solvent remain in solution rather than freeze out or vaporize. Under these restricted conditions of a dilute, ideal solution, the extent ΔT of freezing-point lowering or boiling-point elevation is related to the concentration of solute, and alternatively its molecular weight M through Eq. (6). Here R is the univeral gas constant, T_0

$$\Delta T = \frac{RT_0^2 M}{1000\Delta H} c \tag{6}$$

the temperature at which freezing or boiling would occur in the pure solvent, ΔH the enthalpy of fusion or vaporization of the solvent (the amount of energy which must be supplied or removed at constant pressure to cause 1 mole to change phase), and c the molal concentration of the solute. (1 molal equals 1 mole of solute per 1 kg of solvent.) Because of the restrictions on the solutions which must be formed and for which the above expression is valid, combined with the accuracy with which temperature differences can be measured, these two methods are best employed for molecules with molecular weights less than 10,000.

Osmometry. The basis for the application of the osmotic pressure method (osmometry) is that the chemical potentials of a solution and a pure solvent separated from each other by a semipermeable membrane must be the same. The membrane must be coarse enough to allow solvent molecules to pass through, but fine enough to prohibit passage of solute. In an attempt to equalize the chemical potentials on either side of the membrane, solvent will flow from the side containing only solvent to the solution. The resulting change in chemical potential causes an increase in pressure on the side containing the solution, which will continue until the pressure produced prohibits further flow of solvent, that is, the chemical potentials are equal. This pressure Π is a real physical pressure which can be measured and is related to the concentration in the solution through the van't Hoff equation (7). Here R is again

$$\Pi = RTc \tag{7}$$

the universal gas constant, T the temperature, and c the concentration in number of moles per given amount of solvent. Because of restrictions on passage through the membrane, osmometry can be used to determine molecular weights from about 1000 to about 30,000. For larger molecules other methods are necessary. *See* OSMOSIS; SOLUTION.

Sedimentation. Of all techniques available for determination of molecular weights of large molecules, sedimentation, or the settling of large molecules out of solution under the action of an external force, is perhaps the most common. This process may take a number of varied forms, but its successful utilization for weight determination depends on the fact that the solute molecules in a solution are never at rest but are free to move among the solvent molecules. The practical application of sedimentation involves producing a directed motion of these particles. This is done in a high-speed centrifuge. A particle of mass m when rotated in a circle experiences a force which is equal to $m\omega^2 r$, where ω is the angular velocity and r the distance of the particle from the center of rotation. In the absence of any other forces, the particle would tend to move away from the force center and to the edge of the container, where it would be collected as soon as it can move no further. However, as the molecules move away from the center of the rotation, a concentration gradient is set up. This causes a diffusion of the particles in the opposite direction to the gradient, and a tendency for the particles to remain distributed throughout the solvent. In addition, there is also a buoyant force on the particles produced by collisions with solvent molecules which tends to maintain the particles in solution. At equilibrium the external centrifugal force driving the sedimentation must balance the total forces produced by the buoyant effect and the concentration gradient. If V is the volume of the solute molecule, ρ_M its density, and ρ_0 the density of the solvent, this

balancing of forces leads to Eq. (8). Here D is the

$$N_A V \rho_M \left(1 - \rho_M^{-1} \rho_0\right) \omega^2 r = \frac{RT\dot{r}}{D} \tag{8}$$

diffusion coefficient, N_A Avogadro's number, R the universal gas constant, and T the temperature. The molecular weight M is related to the mass density by Eq. (9). At equilibrium the rate of sedimentation

$$M = N_A V \rho_M \tag{9}$$

(rate of deposit of particles across an area in the direction of the rotation axis) will equal the rate of diffusion across that same boundary. Combining this requirement with the force balance yields Eq. (10)

$$M = \frac{2RT}{(1 - \rho_M^{-1} \rho_0)\omega^2} \frac{\ln (c_2/c_1)}{r_2^2 - r_1^2} \tag{10}$$

for the molecular weight. Successful application of this method requires determination of the concentrations c_1 and c_2 at two values of the radius r_1 and r_2. This is carried out by optical refractometry. Light when passing from one medium to another will be refracted so as to follow a different path from the original one. If the second medium happens to be a solution, the angle of refraction will depend on the concentration of the solution. A modern centrifuge has containers for the solution which are constructed of quartz to provide for this measurement while the solution is being centrifuged. *See* REFRACTION OF WAVES.

A variant on the sedimentation equilibrium method described above involves dissolving the molecule in a concentrated salt solution and centrifuging the resulting mixture. The buoyant force produced by the dense salt solution will force the molecular species to be concentrated in that region in which its density is equal to that of the concentrated solution, $\rho_M = \rho_0$. Continuous motion of the molecules, however, prevents the concentration from being a perfect line, but some distribution about the central point r_0 occurs. From the width σ of this distribution about the center and the density gradient $\Delta\rho/\Delta r$ of the salt solution, the molecular weight can be extracted according to Eq. (11).

$$M = \frac{2RT}{\rho_M^{-1}(\Delta\rho/\Delta r)\omega^2 r_0 \sigma^2} \tag{11}$$

See ULTRACENTRIFUGE.

Light scattering. Another measurement from which molecular weights can be obtained is based on the scattering of light from the molecule. A beam of light falling on a molecule will induce in the molecule a dipole moment which in its turn will radiate. The interference between the radiated beam and the incoming beam produces an angular dependence of the scattered radiation which depends on the molecular weight of the molecule. This occurs whether the molecule is free or in solution. While the theory for this effect is complicated and varies according to the size of the molecule, the general result for molecules whose size is considerably less than that of the wavelength λ of the radiation (less than $\lambda/50$) is given by Eq. (12). Here $I(\theta)$ is the

$$\frac{I(\theta)}{I_0} = \text{constant}\,(1 + \cos^2\theta)\,Mc \tag{12}$$

intensity of radiation at angle θ, I_0 the intensity of the incoming beam, M the molecular weight, and c the concentration in grams per cubic centimeter of the molecule. If the molecules are much larger than $\lambda/50$ (about 9 nanometers for visible light), this relationship in this simple form is no longer valid, but the method is still viable with appropriate adjustments to the theory. In fact, it can be used in its extended version even for large aggregates. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

Sample purity. One requirement which all these methods demand is purity of the sample if the resulting weight is to be very accurate. Impurities result in a molecular weight which is not absolute for the species but is the average weight of the species plus the impurities. Depending on whether or not the influence of the impurities is due to their number (for low-molecular-weight impurities) or their weight (high-weight impurities), the average is either a number average M_n or a weight average M_w. The two average molecular weights are defined by Eqs. (13) and (14), where n_i is the number of particles of species i with molecular weight M_i.

$$M_n = \frac{\Sigma_i(n_i M_i)}{\Sigma_i n_i} \tag{13}$$

$$M_w = \frac{\Sigma_i(n_i M_i^2)}{\Sigma_i(n_i M_i)} \tag{14}$$

Measurement of auxiliary parameters. All the methods in use in biophysics require the knowledge of parameters other than the one actually being measured in the experiment. In most cases this will be the volume of the molecule, or ρ_M^{-1}. For these methods to give accurate results, this number must be obtained by other means. One of the best, but often one of the hardest to apply, is x-ray crystallography, the same technique used to determine Avogadro's number. *See* X-RAY CRYSTALLOGRAPHY.

Other determination methods. Other methods of molecular weight determination include end-group labeling, useful for proteins; viscosity, useful for relative molecular weight determination; gel electrophoresis, also useful for relative measurements; and electron microscopy, useful if the molecule is large enough to be well resolved in the microscope and again only for relative measurements. For a truly accurate molecular weight determination, a combination of these techniques may be necessary. *See* ELECTRON MICROSCOPE; ELECTROPHORESIS; VISCOSITY.

C. Denise Caldwell

Bibliography. P. W. Atkins and J. de Paula, *Atkins' Physical Chemistry*, 8th ed., 2006; A. R. Cooper (ed.), *Determination of Molecular Weight*, 1989; D. Freifelder, *Principles of Physical Chemistry with Applications to the Biological Sciences*, 2d ed., 1985; I. N. Levine, *Physical Chemistry*, 5th ed., 2001.

Molecule

The smallest unit of a compound, which consists of atoms bonded together in a unique arrangement. A diatomic molecule consists of two atoms linked together by chemical bonds. Examples include the oxygen and nitrogen of the air (O_2 and N_2, respectively) and the neurotransmitter nitric oxide (NO). A polyatomic molecule consists of more than two linked atoms. Examples include water (H_2O), carbon dioxide (CO_2), ethanol (C_2H_5OH), and the vast molecule of deoxyribonucleic acid (DNA) with thousands of atoms.

The bonds that hold the atoms together are covalent. That is, they consist of pairs of electrons that are shared by the neighboring atoms. This type of bonding is in contrast to ionic bonding, in which there is complete transfer of one or more electrons from one atom to another and large numbers of the resulting ions (charged atoms) clump together, as in sodium chloride (NaCl). Some chemists refer to the smallest unit of an ionic compound (in the case of sodium chloride, a sodium ion, Na^+, and a chloride ion, Cl^-) as constituting a "molecule" of the compound, but this usage is uncommon and the term "formula unit" is preferred. In this article, coverage is confined to covalently bonded molecules, of which there are millions of kinds. *See* CHEMICAL BONDING; IONIC CRYSTALS.

Organic and inorganic molecules. Many molecules are organic. That is, they contain at least one atom and commonly many atoms of carbon. Molecules that do not contain carbon, with a few exceptions, are inorganic. Thus ethanol consists of organic molecules, and water consists of inorganic molecules. The principal exception to this general classification is carbon dioxide, which is treated as an honorary inorganic compound. The term "organic" arose when it was thought erroneously that such compounds could be produced only by living organisms. Advances in chemistry over decades have blurred the distinction between organic and inorganic, and many compounds are now known which, though they contain carbon, are regarded as lying within the domain traditionally regarded as inorganic chemistry (the so-called organometallic compounds), and many organic compounds that are actually produced by organisms contain metal atoms (some enzymes and vitamins, such as vitamin B_{12}, which contains cobalt atoms). Nevertheless, the distinction between organic and inorganic is valuable as a broad-brush classification and is still widely used. This article concentrates on organic molecules. *See* ORGANOMETALLIC COMPOUND; VITAMIN B_{12}.

Multiplicity of bonds. The bonds between atoms may be single, double, or triple (in rare cases, quadruple), in which neighboring atoms share one, two, or three electron pairs. The OH bonds in H_2O are single (represented as H—O—H), the CO bonds in CO_2 are double (represented as O═C═O), and the NN bonds in N_2 are triple (N≡N). Broadly speaking, the greater the multiplicity of bonds, the more tightly bonded the atoms. This tightness of bonding is manifest, for example, in the inertness of molecular nitrogen and its role as a diluent of the dangerously reactive molecular oxygen of the atmosphere. It also accounts for the difficulty of converting atmospheric nitrogen into ammonia and nitrate fertilizers for use in agriculture.

(a) (b) (c)

Fig. 1. Three representations of the ethanol molecule, C_2H_5OH: (*a*) ball-and-stick, (*b*) space-filling, (*c*) surface.

Shapes. Molecules have not only a characteristic atomic composition but also, because the atoms are linked together in a characteristic array, a characteristic shape. Thus, H_2O is an angular molecule in which the two OH bonds make an angle of about 104° to each other, and CO_2 is linear (all three atoms in a straight line). The four bonds that carbon commonly forms (as in methane, CH_4) are typically arranged in a tetrahedral array with bond angle close to 109°. The resulting shape of the ethanol molecule is seen in **Fig. 1**, which shows both the bonds as sticks joining balls representing the atoms (Fig. 1*a*), and a "space-filling" molecule (Fig. 1*b*), which gives a better notion of the overall bulk of the molecule by representing each atom by a sphere that corresponds to its actual size. Computer graphics are enormously useful for displaying molecular shapes, and can take a variety of forms. Figure 1*c* aims to display the general shape of the molecule without distinguishing the individual atoms.

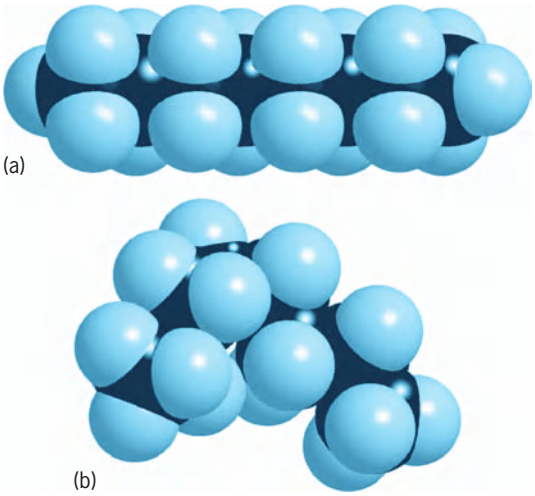

(a) (b)

Fig. 2. Octane molecule (C_8H_{18}) in its (*a*) linear and (*b*) coiled conformations.

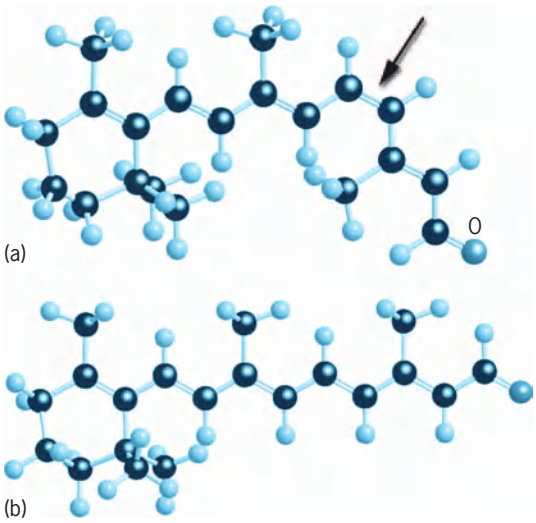

Fig. 3. Ball-and-stick models of (*a*) *cis*-retinal and (*b*) *trans*-retinal. The arrow indicates the double bond that is loosened by the absorption of light.

Flexibility and rigidity. The tetrahedral geometry of the bonds to carbon applies only when all the bonds are single. Moreover, neighboring groups of atoms are free to rotate around a single bond, which acts as a miniature axle. As a result, a polyatomic molecule consisting of many singly bonded atoms should not be thought of as a rigid framework, but as ceaselessly writhing and twisting into different shapes or "conformations." **Figure 2** shows two of the conformations of the hydrocarbon molecule octane (C_8H_{18}), a component of gasoline. *See* CONFORMATIONAL ANALYSIS.

In contrast, the presence of a multiple bond confers structural rigidity because the neighboring groups of atoms joined by a double bond are not free to rotate relative to one another. This structural rigidity has a number of important consequences. For instance, beef fat consists of molecules (tristearin) in which three 18-carbon-atom-long chains are joined together. As all the bonds are single, these long chains are flexible and the entire molecule can roll up into a ball; these ball-like molecules can pack together closely, and tristearin is a solid. On the other hand, similar molecules produced in olives have a double bond in each of their chains. As a result, the chains are less flexible, the molecule cannot roll up into a ball, the molecules cannot pack together as closely, and the compound is an oil rather than a fat. *See* FAT AND OIL.

Role of double bonds between carbon atoms. Double bonds between carbon atoms play important roles in physiology and industry. The primary act of vision, for instance, can be traced to the properties of double bonds. The retinas of human eyes are populated by the molecule *cis*-retinal (**Fig. 3***a*). As can be seen from the line structure, one feature of the molecule is the alternation of single and double bonds between neighboring carbon atoms. This alternation has two consequences: the pairs of electrons forming the bonds are quite loosely attracted to the atomic nuclei and can easily be moved around by incident light; and the molecule is rigid and adopts the shape shown in Fig. 3*a*. However, when light is incident on the molecule, one of the double bonds is weakened, the long side-chain becomes free to rotate, and the molecule snaps into the form known as *trans*-retinal (Fig. 3*b*). This molecule can no longer fit into the protein molecule that normally houses it, and it is expelled. The protein molecule responds by relaxing into a new shape, and as a result a signal is pulsed along the optic nerve into the brain, which interprets it as "vision." Various biochemical reactions then take place to restore the molecule to its original shape, in which it is primed to act again. *See* PHOTORECEPTION.

Polymerization. The industrial importance of double bonds between carbon atoms lies in the reactivity they confer on organic molecules (in contrast to the inertness conferred on nitrogen). One aspect of this reactivity is the ability to combine small double-bonded organic molecules into long chains. That is, certain molecules act as monomers that may be polymerized to form the plastics characteristic of the modern world. The process is illustrated in **Fig. 4**, which shows a variety of polymers based on ethene (ethylene, $CH_2{=}CH_2$). The most primitive of these polymers, but far from being useless, is polyethylene itself, the polymer of ethylene, which consists of long chains of thousands of $—CH_2—CH_2—$ groups, often with branch

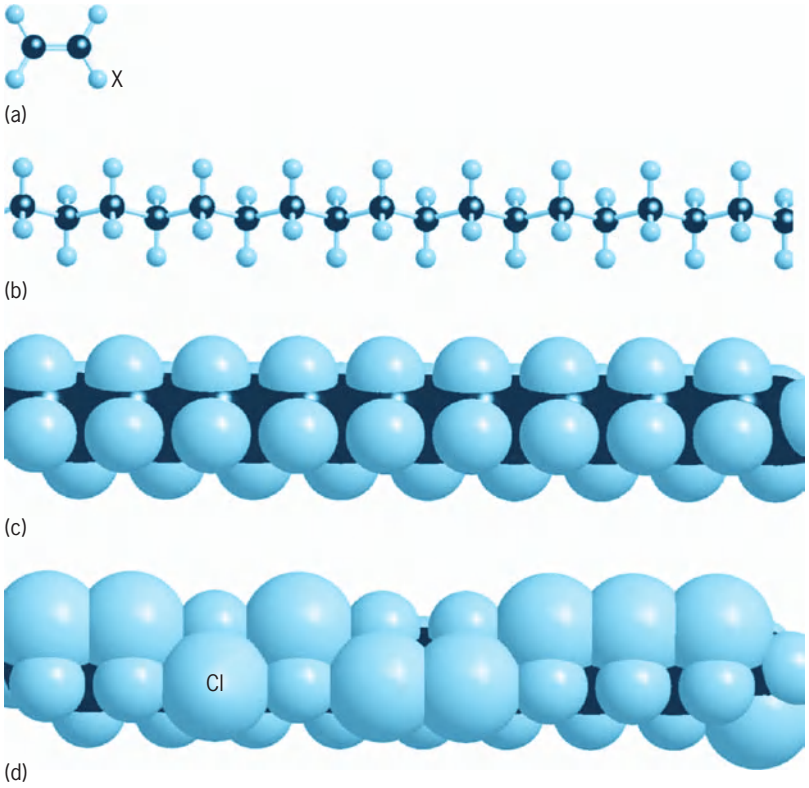

Fig. 4. Polymerization based on ethene (ethylene). (*a*) Ethene molecule; the atom marked X may be replaced by atoms of other elements. (*b*) Ball-and-stick and (*c*) space-filling models of a fragment of a polyethylene chain. (*d*) Space-filling model of a fragment of a polyvinyl chloride (PVC) chain.

Fig. 5. Ball-and-stick model of benzene (C_6H_6) superimposed on a model of its surface.

lines where the polymerization has proceeded in a slightly different way. When the atom marked X in Fig. *4a* is replaced by a chlorine atom to give the monomer $CHCl{=}CH_2$, the polymerization process results in chains of $—CHCl—CH_2—$ groups and the polymer polyvinyl chloride (PVC). When X is a benzene-like ring, the monomer is styrene and the resulting polymer is polystyrene. Chemists are able to modify the monomers in a wide variety of ways and to produce polymers with a correspondingly wide range of properties. Thus, if all the hydrogen atoms of the monomer are replaced by fluorine atoms, to give the monomer $CF_2{=}CF_2$, the resulting polymer is polytetrafluorethylene (PTFE), used in nonstick coatings. *See* ETHYLENE; POLYFLUOROOLEFIN RESINS; POLYMER; POLYMERIZATION; POLYOLEFIN RESINS; POLYSTYRENE RESIN; POLYVINYL RESINS; STYRENE.

Aromatic compounds. The members of one very important class of molecules with double bonds have a greatly reduced reactivity (and thus are like organic counterparts of nitrogen). The parent of these so-called aromatic compounds is benzene, C_6H_6. Like benzene itself, many of these compounds have an aroma, but many do not, and the term applies to any compound that contains a benzene-like group. A benzene molecule consists of a planar hexagonal array of carbon and hydrogen atoms (**Fig. 5**). Although the carbon atoms are joined by alternating single and double bonds, and so might be expected to be highly reactive, this arrangement actually confers stability (for quantum-mechanical reasons) and the molecule is not very reactive. As a result, the ring is a motif that is found in many organic compounds, including polystyrene. *See* AROMATIC HYDROCARBON; BENZENE.

This brief survey only skims the subject of molecules. Whatever we breathe, eat, drink, and wear, whatever medications we take, whatever cosmetics we use, almost whatever we touch and see, is made of molecules. Peter Atkins

Bibliography. P. Atkins, *Atkins' Molecules*, Cambridge University Press, 2003; P. Atkins and L. Jones, *Chemical Principles*, W. H. Freeman, New York, 2007.

Mollusca

A major phylum of the animal kingdom comprising an extreme diversity of external body forms (oysters, clams, chitons, snails, slugs, squid, and octopuses among others), all based on a remarkably uniform basic plan of structure and function. The phylum name is derived from *mollis*, meaning soft, referring to the soft body within a hard calcareous shell, which is usually diagnostic. Soft-bodied mollusks make extensive use of ciliary and mucous mechanisms in feeding, locomotion, and reproduction. Most molluscan species are readily recognizable as such.

The Mollusca constitute a successful phylum; there are probably over 110,000 living species of mollusks, a number second only to that of the phylum Arthropoda, and more than double the number of vertebrate species. More than 99% of living molluscan species belong to two classes: Gastropoda (snails) and Bivalvia. Ecologically, these two classes can make up a dominant fraction of the animal biomass in many natural communities, both marine and fresh-water. Certain bivalve species are the most abundant marine benthic animals, and computations have suggested that one of these may encompass a larger "standing crop" biomass of animal tissue than any other single animal species on the planet.

Classification

The phylum Mollusca is divided into seven distinct extant classes, three of which (Gastropoda, Bivalvia, and Cephalopoda) are of major significance in terms both of species numbers and of ecological bioenergetics, and one extinct class. An outline of their classification follows.

Class Monoplacophora (mainly fossil; but one living genus *Neopilina*)
Class Aplacophora
  Subclass Neomeniomorpha
  Subclass Chaetodermomorpha
Class Polyplacophora
Class Scaphopoda
Class Rostroconchia (fossil only)
Class Gastropoda
  Subclass Prosobranchia
    Order: Archaeogastropoda
      Mesogastropoda
      Neogastropoda
  Subclass Opisthobranchia
    Order: Bullomorpha (or Cephalaspidea)
      Aplysiomorpha (or Anaspidea)
      Thecosomata
      Gymnosomata
      Pleurobranchomorpha (or Notaspidea)
      Acochlidiacea
      Sacoglossa
      Nudibranchia (or Acoela)
  Subclass Pulmonata
    Order: Systellommatophora
      Basommatophora
      Stylommatophora

Class Bivalvia (or Pelecypoda)
- Subclass Protobranchia
- Subclass Lamellibranchia
  - Order: Taxodonta
    - Anisomyaria
    - Heterodonta
    - Schizodonta
    - Adapedonta
    - Anomalodesmata
- Subclass Septibranchia

Class Cephalopoda
- Subclass Nautiloidea
- Subclass Ammonoidea (fossil only)
- Subclass Coleoidea
  - Order: Belemnoidea
    - Sepioidea
    - Teuthoidea
    - Vampyromorpha
    - Octopoda

The subclass Opisthobranchia is sometimes divided into two superorders: Tectibranchia (the first seven orders) and Nudibranchia. The planktonic opisthobranch orders Thecosomata and Gymnosomata are sometimes united in the order Pteropoda. There are several alternative ordinal arrangements for the many well-established superfamilies of lamellibranch bivalves. An attractive one on functional grounds (but unacceptable phyletically) is to divide the subclass into two: Filibranchia and Eulamellibranchia.

The use by paleontologists of subphyla, including Cyrtosoma (for Monoplacophora with Gastropoda and Cephalopoda) and Diasoma (for Rostroconchia with Bivalvia and Scaphopoda) in the classification of the phylum Mollusca has not been generally adopted by students of living forms (neontologists). There is no disagreement over the further division of the phylum into the eight clearly distinguishable classes listed above.

Functional Morphology

The unique basic plan of the Mollusca involves the different modes of growth and of functioning of the three distinct regions of the molluscan body (**Fig. 1**). These are the head-foot with some nerve concentrations, most of the sense organs, and all the locomotory organs; the visceral mass (or hump) containing organs of digestion, reproduction, and excretion; and the mantle (or pallium) hanging from the visceral mass and enfolding it and secreting the shell. In its development and growth, the head-foot shows a bilateral symmetry with an anterioposterior axis of growth. Over and around the visceral mass, however, the mantle-shell shows a biradial symmetry, and always grows by marginal increment around a dorsoventral axis (Fig. 1 and **Fig. 2**). It is of considerable functional importance that a space is left between the mantle-shell and the visceral mass forming a semi-internal cavity; this is the mantle cavity or pallial chamber within which the typical gills of the mollusk, the ctenidia, develop. This mantle cavity is almost diagnostic of the phylum; it is primarily a respiratory chamber housing the ctenidia, but with alimentary, excretory, and genital systems all discharging into it (Fig. 2). The basic functional plan is always recognizable, although it has undergone remarkable modifications of structure and of function in different groups of mollusks.

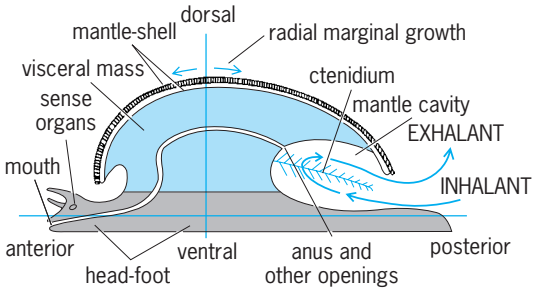

Fig. 1. Generalized model of a stem mollusk (or archetype) in side view. There are three distinct regions in the molluscan body: head-foot, visceral mass, and mantle-shell. Water circulation through the mantle cavity, gills (ctenidia), and pallial complex is from ventral inhalant to dorsal exhalant. (*After W. D. Russell-Hunter, A Life of Invertebrates, Macmillan, 1979*)

Mantle-shell and body form. In looking at any mollusk, it is important to realize that whatever the shape of the shell, it is always underlain by the mantle, a

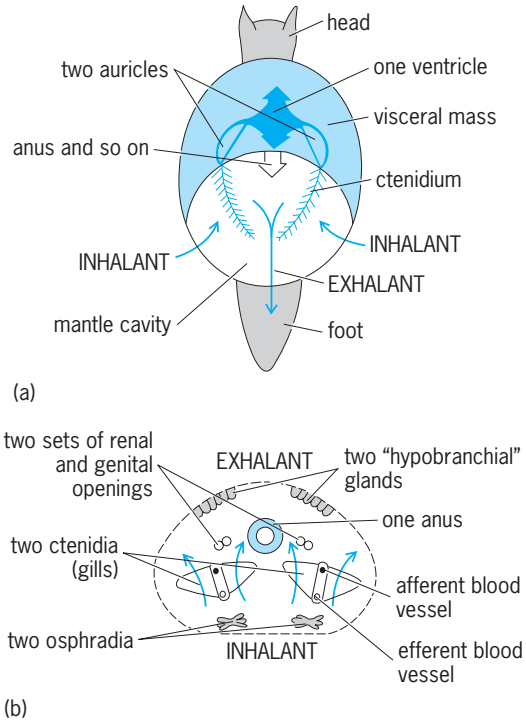

Fig. 2. Mantle cavity and its associated structures in a molluscan archetype. (*a*) Dorsal view of the symmetrical posterior mantle cavity, with a pair of feather gills (aspidobranch condition) and two auricles in the heart (diotocardiac condition). (*b*) Cross section through the mantle cavity of the archetype showing all the elements of the pallial complex in a bilaterally symmetrical array. (*After W. D. Russell-Hunter, A Life of Invertebrates, Macmillan, 1979*)

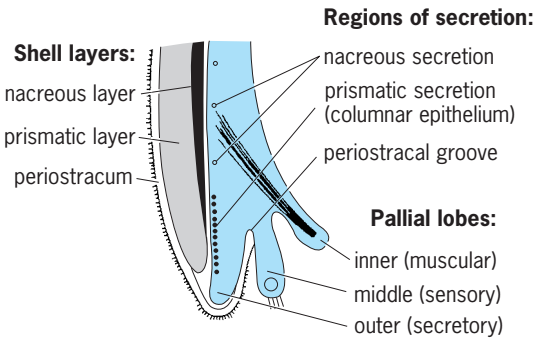

Fig. 3. Arrangement of the mantle and shell layers in the mollusks. The fleshy mantle margin is divided into three lobes, and the outer pallial lobe is largely responsible for secretion of the three layers of the shell. (*After W. D. Russell-Hunter, A Life of Invertebrates, Macmillan, 1979*)

fleshy fold of tissues which has secreted it. The detailed structure of the shell and of the mantle edge (with three functionally distinct lobes; **Fig. 3**) is also consistent throughout the Mollusca. The shell is made up of calcium carbonate crystals enclosed in a meshwork of tanned proteins. It is always in three layers: the outer periostracum, which is highly organic; the prismatic layer, which is the most massive part; and the innermost or nacreous layer, which has the least organic content (Fig. 3).

Each of the eight classes of the Mollusca has a characteristic body form and shell shape. Two classes are enormous (Gastropoda and Bivalvia), one of moderate extent (Cephalopoda), the others being minor by comparison. The Gastropoda constitute a diverse group with the shell usually in one piece. This shell may be coiled as in typical snails—that is, helicoid or turbinate—or it may form a flattened spiral, or a short cone as in the limpets, or it may be secondarily absent as in the slugs. Most gastropods are marine, but many are found in freshwaters and on land; in fact, they are the only successful nonmarine mollusks.

The Bivalvia are a more uniform group, with the shell in the form of two calcareous valves united by an elastic hinge ligament. Mussels, clams, and oysters are familiar bivalves. The group is mainly marine with a few genera in estuaries and in freshwaters. There can be no land bivalves since their basic functional organization is as filter feeders. The third major group, the Cephalopoda, includes the most active and most specialized mollusks. There is a chambered, coiled shell in *Nautilus* and in many fossil forms; this becomes an internal structure in cuttlefish and squids, and is usually entirely absent in octopods.

There are only about 40 genera of chitons, or "coat-of-mail shells" with eight plates, placed in the Polyplacophora. The other three living minor groups (class Rostroconchia is completely extinct) are marine, and there are only a few species in each. The Monoplacophora encompass several fossil families and the one living genus *Neopilina*, all with limpet-like shells. The Scaphopoda are the "elephant's-tusk-shells," and the Aplacophora have a wormlike body whose mantle secretes discrete calcareous spicules but never a shell.

Ctenidia and pallial complex. In the more primitive living snails and bivalves, there are paired feather-like (aspidobranch) ctenidia (**Fig. 4**). Filaments or gill plates alternate on either side of an axis which contains both a dorsal afferent blood vessel (deoxygenated blood) from the body and a ventral efferent blood vessel (oxygenated blood) running into the heart. Water flow (created by the lateral cilia) is always from ventral to dorsal, in the direction opposite to blood flow (Fig. 4). This physiologically efficient counterflow system is functionally homologous throughout the mollusks (with few, adaptively explicable, exceptions), and is based on homologous structural arrangements of ciliated tracts, blood vessels, skeletal elements, and renopericardial connections. A diversity of gill patterns have evolved in the major molluscan groups (**Fig. 5**), paralleling the evolution of the mantle-shell patterns, but all clearly derived from the archetypic ctenidial arrangement (Fig. 4).

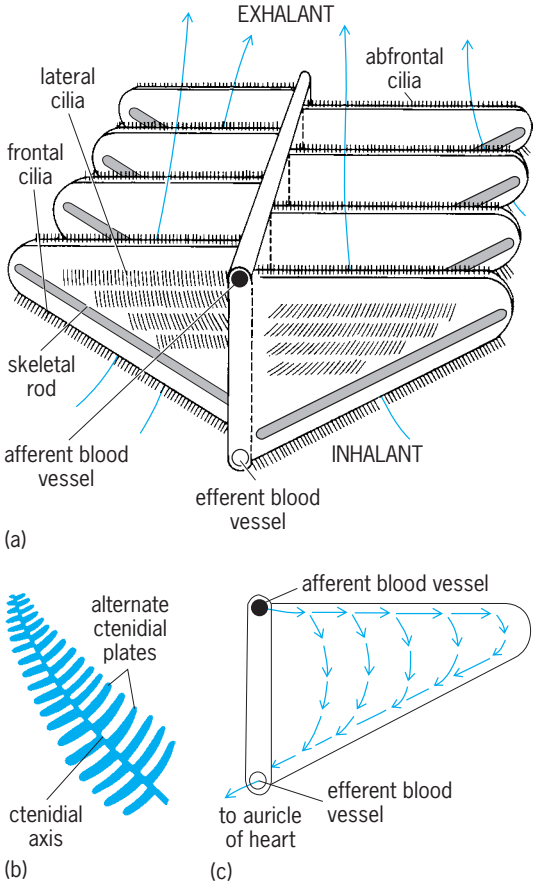

Fig. 4. Aspidobranch ctenidium (feather or plume gill) as found in primitive gastropods and other molluscan groups. (*a*) Stereogram showing the water current from ventral inhalant to dorsal exhalant between adjacent gill plates—a current created by the lateral cilia. (*b*) Entire gill viewed from above. (*c*)A single gill plate showing the counter-current flow of blood within the plate, which results in physiological efficiency of oxygen exchange. (*After W. D. Russell-Hunter, A Life of Invertebrates, Macmillan, 1979*)

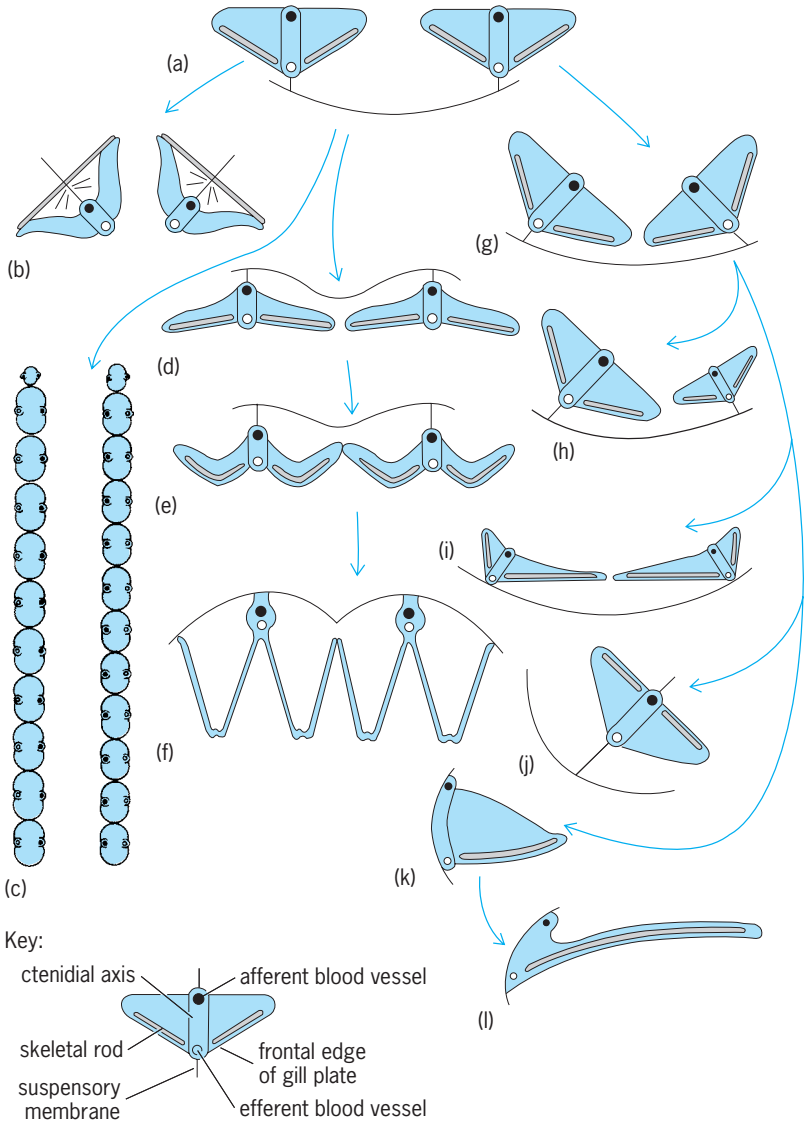

Fig. 5. Evolution of molluscan gills (ctenidia) in cross sections, viewed from the posterior. General (*a*) molluscan and (*g*) gastropod archetypes are illustrated, but the only other hypothetical pattern is (*e*) that shown as an intermediate bivalve condition. All the other gill patterns shown occur in presently living forms: (*b*) in modern cephalopods, (*c*) in chitons, (*d*) in protobranch and (*f*) lamellibranch bivalves, and (*h–l*) in a series of gastropods. (*After W. D. Russell-Hunter, A Life of Invertebrates, Macmillan, 1979*)

The more advanced gastropods show reduction from a pair of aspidobranch ctenidia to a single one, and from that to a one-sided pectinibranch ctenidium (or comb gill), and subsequently to no gill at all in the pulmonate snails. The bivalves show enlargement of gill leaflets to longer filaments and their subsequent folding into the true lamellibranch condition (Fig. 5), used in filter feeding. The gills in the cephalopods, while still structurally homologous, are modified with new skeletal elements to resist the stresses of water pumping by muscles (rather than by cilia). In the chitons, ctenidia are replicated, though not necessarily symmetrically (Fig. 5).

Besides the gills, the other organs of the mantle cavity (termed collectively the pallial complex; Fig. 2) again show morphological and functional consistency throughout the main groups of the Mollusca. The ctenidia form a curtain functionally dividing the mantle cavity into an inhalant part (usually ventral) containing the osphradia (pallial sense organs which sample the incoming water), and an exhalant part (usually dorsal) containing hypobranchial glands and both the anus and the openings of the kidney and genital ducts (Fig. 2). Thus feces, gametes, and nitrogenous wastes are all discharged downstream from the gill curtain in the exhalant current. This arrangement of the pallial complex as a functionally integrated group of structures remains consistent in mollusks whether they have ctenidia asymmetrically reduced (higher Gastropoda), or enlarged into feeding lamellae (higher Bivalvia), or replicated in lateral series (Polyplacophora).

Cardiac and renogenital complex. Throughout the Mollusca, adult kidneys and genital ducts have developed from true mesodermal coelomoducts associated with the pericardial cavity (**Fig. 6**). The cardiac structures of mollusks are also closely linked to the pallial complex. If there is a symmetrical pair of ctenidia, there will be a symmetrical pair of auricles on either side of the muscular ventricle of the heart; if one ctenidium, one auricle; if four ctenidia, four auricles. There are auriculoventricular valves to prevent backflow, and the ventricle drives blood both anteriorly and posteriorly into the major arterial systems, again through nonreturn valves. In return, these serve the hemocoelic spaces with a relatively low-pressure circulation of relatively large volumes of blood. Note that body fluids in mollusks are almost all blood, just as body cavities are almost all hemocoel.

The respiratory pigment is usually hemocyanin in solution, so that neither circulatory efficiency nor blood oxygen-carrying capacity is high (compared to even lower vertebrates like fishes). However, mollusks are mostly sluggish animals with low metabolic (and hence respiratory) rates. The blood in the hemal meshwork of all mollusks has another functional importance, since it is used as a hydraulic skeleton to transmit forces generated by distant skeleton to transmit forces generated by distant muscle contraction. The characteristically extensible soft parts of mollusks, such as tentacles, the foot, and the siphons, can all be rapidly withdrawn by muscular contraction, but are only slowly extended again by blood pressure, by blood being shifted into them from another part of the molluscan body. Mechanically, this is a case of distant sets of antagonistic muscles transmitting forces through a hemocoelic, hydraulic skeleton of fixed volume.

As shown in Fig. 6, the basic arrangement of one pair of gonads and one pair of kidneys is modified in the bivalves and chitons principally by separation of genital from renal functions while retaining paired structures. On the other hand, in gastropod evolution the asymmetry developed in gills and pallial structures (and hence in cardiac organization) has allowed various levels of asymmetry (and hence functional separation) in the renogenital structures. As shown in Fig. 6, in the majority of advanced snails the left coelomoduct becomes the functional renal

organ, while the right one is the gonoduct for the solitary gonad.

Ciliary sorting and gut function. Uniquely molluscan is the use of cilia in "sorting surfaces," which can segregate particles into different size categories and send them to be disposed of in different ways in several parts of the organism. In a simpler type of sorting surface, the epithelium is thrown into a series of ridges and grooves, the cilia in the grooves beating along them and the cilia on the crests of the ridges beating across them. Thus, fine particles impinging on the surface can be carried in the direction of the grooves, while larger particles are carried at right angles. Such sorting surfaces occur both externally on the feeding organs and internally in the gut of many mollusks. For example, on the labial palps of bivalves, they are used to separate the larger sand grains (which are rejected) from the smaller microorganisms which then pass to the mouth. Although such sorting is by size alone, it can result in a separation of valuable food from inedible particles. Some molluscan sorting surfaces can achieve segregation into four or more size categories. Further, many mollusks have the capacity to alter the thresholds of discrimination on these surfaces, since such organs as labial palps can be expanded to a varying extent by blood pressure.

The alimentary canal in mollusks is usually extensively ciliated, with many sorting surfaces. Such a gut is organized to deal slowly but continuously with a steady stream of finely divided plant material passing in from the feeding organs, whether these be filter mechanisms or the characteristic rasping tongue (the radula) of the grazing mollusks. In the majority, digestion is functionally mixed: both extracellular and intracellular breakdown is carried out. In all lamellibranchs, and in both grazing and filter-feeding snails, a peculiar secreted structure termed the crystalline style is used as a cilia-driven gastric stirring rod and enzyme store, allowing a slow, continuous release of amylase and glycogen-breaking enzymes. The digestive diverticula (often miscalled liver lobes) are masses of blind tubules, the cells of which are phagocytic, where final intracellular digestion takes place. In the cephalopods and a few carnivorous gastropods, digestion is simpler, largely extracellular, and cyclic.

Nervous system. The range in levels of complexity of molluscan nervous systems is comparable to that found in the phylum Chordata. The four-strand nervous system with one pair of tiny ganglia found in chitons is not dissimilar to the neural plan in turbellarian flatworms. In contrast, the nervous system and sense organs of a cephalopod like an octopus are equaled and exceeded only by those of some birds and mammals.

In the majority of mollusks the nervous system is in an intermediate condition. This state involves five groups of paired ganglia associated with the buccal structures, the other head organs, the foot, the visceral mass, and the mantle. These five pairs of "local brains" are linked by long connectives and commissures. In gastropods the process of torsion has resulted in twisting of these long trunk nerves (the so-called streptoneurous condition). This decentralized pattern of relative simplicity is probably related to the fact that, in mollusks other than cephalopods, the main effectors controlled by the nervous system are cilia and mucous glands. In fact, apart from the muscles which withdraw it into its shell, the typical mollusk is a slow-working animal with little fast nervous control or quick reflexes. In the brain of modern cephalopods, these paired molluscan ganglia have been fused into a massive structure, with over 300 million neurons and extensive "association" centers providing considerable mnemic and learning capacities. *See* OCTOPUS.

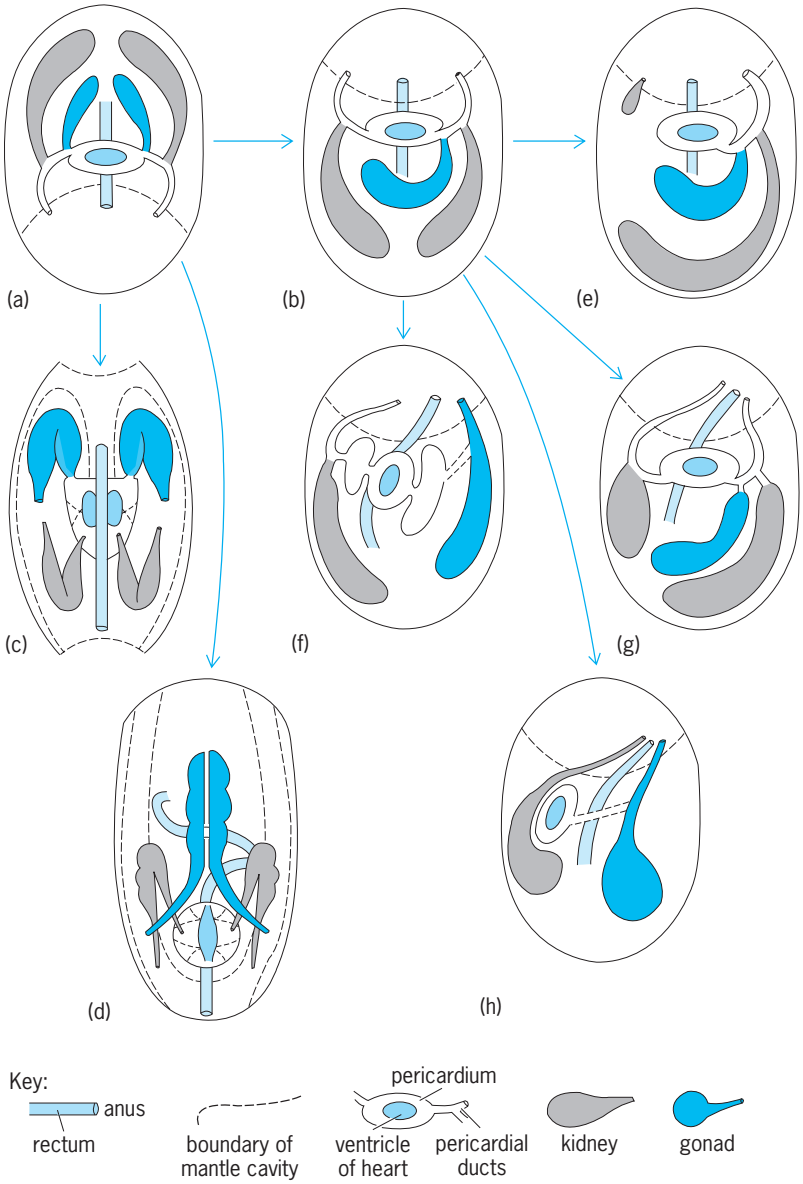

Fig. 6. Evolution of the pericardial cavity and its associated organs (heart, kidneys, gonads, rectum, and various ducts) in a series of representative mollusks. All are dorsal diagrams and, while (*a*) molluscan and (*b*) gastropod archetypes are illustrated first, all the other conditions shown occur as actual states of these organ systems in presently living forms: (*c*) in bivalves, (*d*), chitons, (*e*) fissurellids, (*f*) neritaceans, (*g*) trochids, and (*h*) pectinibranch prosobranchs (and higher gastropods). (*After W. D. Russell-Hunter, A Life of Invertebrates, Macmillan, 1979*)

Reproduction. In all primitive mollusks, the sexes are separate, and external fertilization follows the spawning of eggs and sperm into the sea. The zygote undergoes spiral cleavage and eventually becomes a trochophore larva, with a characteristic ring of locomotory cilia. This later develops into a typical wheel-shaped velum bearing long cilia, and this veliger larva then differentiates a mantle rudiment (secreting a shell) and a distinct headfoot. Molluscan larvae first become recognizable as gastropods, as bivalves, or as chitons when their mantle-shell rudiments first take on the growth patterns characteristic of each class. After this differentiation, the ciliated larva settles to the bottom and metamorphoses into a crawling spat (usually a miniature version of the adult mollusk).

In more advanced mollusks, eggs are larger (and fewer), fertilization may become internal (with complex courtship and copulatory procedures), and larval stages may be sequentially suppressed. A remarkably large number of mollusks (including many higher snails) are hermaphroditic. Although some are truly simultaneous hermaphrodites, many more show various kinds of consecutive sexuality. Most often the male phase occurs first, and these species are said to show protandric hermaphroditism.

Distributional Ecology

Mollusks are largely marine. The extensive use of ciliary and mucous mechanisms in feeding, locomotion, reproduction, and other functions demands a marine environment for the majority of molluscan stocks. Apart from a small number of bivalve genera living in brackish and fresh waters, all nonmarine mollusks are gastropods.

Despite the soft, hydraulically moved bodies and relatively permeable skins typical of all mollusks, some snails are relatively successful as land animals, although they are largely limited to more humid habitats. The primary physiological requirements for life on land concern water control, conversion to air breathing, and temperature regulation. Other nonmarine adaptations involve changes in reproduction (to larger eggs or viviparity) and in nitrogenous excretion (for water conservation). Some groups of littoral snails exhibit the serial acquisition of such terrestrial adaptations, as they occur in the upper seashore. It is a remarkable evolutionary fact that the successful land snails of the class Pulmonata have taken third place (admittedly well behind the arthropods and amniote vertebrates) in the race to exploit nonmarine habitats, both terrestrial and fresh water. Pulmonate land snails occur on suitable mountains at altitudes of nearly 19,700 ft (6000 m). The dominant snails of fresh waters are pulmonates of the higher limnic Basommatophora, representing an air-breathing stock of snails showing progressive readaptation to aquatic life.

In the sea, all classes of mollusks are found, and all habitats have mollusks. Protobranchiate bivalves are found at depths of over 29,500 ft (9000 m). Although ecologically cephalopod mollusks are limited to the sea, there are sound reasons for claiming modern cephalopods as the most highly organized invertebrate animals. The functional efficiencies of jet propulsion and of massive brains in squid, cuttlefish, and octopuses have not been paralleled in their other physiological systems. The processes of excretion and reproduction, and the respiratory function of the blood, are still typically molluscan and could not be readily adapted for nonmarine life. It is clear that the pelagic cephalopods evolved along with the swimming vertebrates, with size and speed ratios being kept in pace by selection pressures of alternate predation and competition. In other ways, pressure to keep up with the competitive vertebrates must have helped evolve the speed, the size, and the brains of modern cephalopods from the molluscan body plan.

In addition to the extreme diversity of external body form exhibited by different mollusks, they show a remarkable diversity in their ecological distribution and life styles. However, the basic molluscan plan of structure and function always remains recognizable. *See* APLACOPHORA; BIVALVIA; CEPHALOPODA; GASTROPODA; LAMELLIBRANCHIA; MONOPLACOPHORA; POLYPLACOPHORA; SCAPHOPODA; SNAIL. W. D. Russell-Hunter

Bibliography. V. Fretter and A. Graham, *British Prosobranch Molluscs*, 1962; J. E. Morton, *Molluscs*, 4th ed., 1967; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; W. D. Russell-Hunter, *A Life of Invertebrates*, 1979; K. M. Wilbur and C. M. Yonge (eds.), *Physiology of Mollusca*, 1964 and 1967; C. M. Yonge and T. E. Thompson, *Living Marine Molluscs*, 1976.

Molybdenum

A chemical element, Mo, atomic number 42, and atomic weight 95.94, in the periodic table in the triad of transition elements that includes chromium (atomic number 24) and tungsten (atomic number 74). Research has revealed it to be one of the most versatile chemical elements, finding applications not only in metallurgy but also in paints, pigments, and dyes; ceramics; electroplating; industrial catalysts; industrial lubricants; and organometallic chemistry. Molybdenum is an essential trace element in soils and in agricultural fertilizers. Molybdenum atoms have been found to perform key functions in enzymes (oxidases and reductases), with particular

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

interest being directed toward its role in nitrogenase, which is employed by bacteria in legumes to convert inert nitrogen (N_2) of the air into biologically useful ammonia (NH_3). *See* NITROGEN FIXATION; PERIODIC TABLE.

Molybdenum is widely distributed in the Earth's crust at a concentration of 1.5 parts per million by weight in the lithosphere and about 10 parts per billion in the sea. It is found in at least 13 minerals, mainly as a sulfide [molybdenite (MoS_2)] or in the form of molybdates [for example, wulfenite ($PbMoO_4$) and magnesium molybdate ($MgMoO_4$)].

Although molybdenum is closer to chromium in atomic weight and atomic number, its chemical behavior is usually very similar to that of tungsten, which has nearly the same atomic radius. (This is due to the so-called lanthanide contraction in which atomic radii decrease for elements 57 to 71 found in the period between molybdenum and tungsten.) *See* CHROMIUM; LANTHANIDE CONTRACTION; TUNGSTEN.

Molybdenum atoms contain six valence electrons ($4d^5 5s^1$), which are employed with great versatility in forming compounds and complexes in which electronic configurations vary from d^0 (no d electrons in oxidation state +6) to d^8 (8 d electrons in oxidation state −2). The +6 state is preferred, but all states from −2 to +6 are known. States usually exhibit a variety of coordination numbers (4 to 9), and include polynuclear complexes and metal-metal bonds in metallic clusters with two to six metal atoms in their metallic cores. Molybdenum forms a very large number of compounds with oxygen. Low-valent molybdenum [for example, $Mo(CO)_6$ and Mo_2, Mo_3, and Mo_6 clusters] has a very rich organometallic chemistry, including clusters that are being studied as models for molybdenum metal surfaces that catalyze organic reactions employed in industrial syntheses and oil refining. The ability of molybdenum atoms to vary oxidation state, coordination number, and coordination geometry and to form metal-metal bonds in clusters accounts in part for the large number of industrial catalysts and biological enzymes in which Mo atoms are found at the active site for catalysis. *See* CHEMICAL BONDING; COORDINATION CHEMISTRY; ELECTRON CONFIGURATION.

Molybdenum is a high-melting silver-gray metal, strong even at high temperatures, hard, and resistant to corrosion (see **table**). It also exhibits high conductivity, a high modulus of elasticity, high thermal conductivity, and a low coefficient of expansion. Its major use is in alloy steels, for example, as tool steels (≤10% molybdenum), stainless steel, and armor plate. Up to 3% molybdenum is added to cast iron to increase strength. Up to 30% molybdenum may be added to iron-, cobalt-, and nickel-based alloys designed for severe heat- and corrosion-resistant applications. It may be used in filaments for light bulbs, and it has many applications in electronic circuitry. *See* ALLOY; IRON ALLOYS; STAINLESS STEEL.

Molybdenum trioxide, molybdates, sulfo-molybdates, and metallic molybdenum are found in thousands of industrial catalysts used in oil refining, ammonia synthesis, and industrial syntheses of organic chemicals. Monomeric molybdenum (IV) in aqueous solution is a powerful catalyst for the reduction of inert oxo-anions such as perchlorate (ClO_4^-) or nitrate (NO_3^-) as well as other oxidized nonmetals such as azide ion (N_3^-) and dinitrogen (N_2). The trinuclear cation $MO_3O_4^{4+}$ is inert, unreactive, and noncatalytic. *See* CATALYSIS; HOMOGENEOUS CATALYSIS.

The molybdenum enzymes comprise two major categories. The first category contains the single, highly important enzyme nitrogenase, which is responsible for biological nitrogen fixation. The second category contains all other known molybdenum enzymes, which are crucial for the metabolism of bacteria, plants, and animals, including humans.

Edward I. Stiefel

Physical properties of molybdenum metal

Property	Value
Density	10.22 g/cm^3 (5.911 oz/in.^{3})
Heat of vaporization	491 kJ/mol
Heat of fusion	28 kJ/mol
Specific heat	0.267 J/g°C
Thermal conductivity	1.246 J/s/cm^2/cm°C (200°C) 0.923 J/s/cm^2/cm°C (2200°C)
Electrical conductivity	34% International Copper Standard
Electrical resistivity	5.2 microhm-cm, 20°C 78.2 microhm-cm, 2525°C
Magnetic susceptibility	0.93×10^{-6} emu, 25°C 1.11×10^{-6} emu, 1825°C
Mean linear expansion coefficient	6.65×10^{-6}/°C, 20–1600°C
Modulus of elasticity	0.324 N/m^2
Lattice parameter	0.314767 nm (body-centered cube)

Bibliography. F. A. Cotton and G. Wilkinson, *Advanced Inorganic Chemistry*, 6th ed., 1999; C. K. Gupta, *Extractive Metallurgy of Molybdenum*, 1993; A. Sigel and H. Sigel (eds.), *Metals Ions in Biological System*, vol. 39: *Molybdenum and Tungsten: Their Roles in Biological Processes*, 2002; T. G. Spiro, *Molybdenum Enzymes*, 1985; E. I. Stiefel, D. Coucouvanis, and W. E. Newton (eds.), *Molybdenum Enzymes, Cofactors, and Model Systems*, ACS Symp. Ser., no. 535, 1993.

Molybdenum alloys

Solid solutions of molybdenum and other metals. Molybdenum is classified as a refractory metal by virtue of its high melting point (2623°C or 4750°F), and many of its applications result from its strength at high temperatures. A number of other physical and mechanical properties make it attractive for use in a wide variety of applications. Molybdenum is used extensively as electrodes in electric-boost furnaces because it erodes very slowly and does not contaminate the glass bath. Its high-temperature strength allows it to support significant structural loads

imposed during operation of the furnaces. *See* MOLYBDENUM; REFRACTORY.

Molybdenum has a body-centered cubic crystal structure and displays the ductile-brittle transition behavior typical of such metals. This behavior is illustrated by bend tests on molybdenum sheet. In this test, a sheet is bent over a mandrel having a radius equal to the sheet thickness. The angle of bend at fracture is a measure of the sheet's ductility. The data from such a test indicate that the microstructure of the material has a strong influence on the ductile-brittle transition temperature. Higher temperatures are required to form recrystallized sheet through a given angle as compared to stress-relieved (recovery-annealed) sheet. Recrystallized microstructures are to be avoided because they are susceptible to brittle intergranular fracture. *See* BRITTLENESS; PLASTIC DEFORMATION OF METAL.

Manufacturing practices. Molybdenum alloys are manufactured and made into usable forms by primary consolidation, forming and machining, and joining.

Primary consolidation. Molybdenum and its alloys are consolidated commercially either by powder metallurgy techniques or by vacuum-arc casting. Both mechanical pressing and cold isostatic pressing are used to consolidate powder metallurgy billets, although most powder metallurgy mill products start as cold-isostatic-pressed billets. Hydrogen atmospheres are typically used to sinter powder metallurgy billets because they aid in the removal of oxygen from the material during sintering by reaction to form water vapor. Vacuum sintering is also employed in production of powder metallurgy billets by some manufacturers. In all cases, oxygen in the finished product must be minimized, because it is quite detrimental to the ductility of molybdenum. The vacuum-arc casting process produces ingots that are low in oxygen and need no further treatment to remove it.

Extrusion and forging can be used to work large powder metallurgy billets; vacuum-arc-cast ingots must first be processed by hot extrusion, because of the as-cast material's propensity for intergranular fracture under tensile stresses. Once large billets and ingots are reduced to convenient sizes, both powder metallurgy and vacuum-arc materials can be processed by conventional techniques such as hot rolling, cold rolling, or swaging. *See* FORGING; METAL CASTING; METAL ROLLING; POWDER METALLURGY; SINTERING.

Forming and machining. Forming or stamping of finished bars or sheets should take into account the ductile-brittle transition-temperature behavior of the material. This will require higher forming temperatures as the sheet gauge or bar diameter increases.

Molybdenum can be machined quite readily, though machinists must be prepared to employ more exacting practices in machining molybdenum and its alloys as compared to conventional materials such as steel, aluminum, and brass. Carbide cutting tools are recommended because of the abrasive nature of the chips formed during machining. Adequate lubrication and appropriate machining practices allow production of intricate and exacting parts. Molybdenum and its alloys are readily machined by electrical-discharge-machining techniques, but the material's high melting point and high thermal conductivity require higher power levels than are encountered in machining conventional materials. *See* MACHINING; METAL, MECHANICAL PROPERTIES OF; METAL FORMING.

Joining. Molybdenum can be joined by both brazing and welding. The fusion and heat-affected zones of welded components are typically very low in ductility because of the high ductile-brittle transition temperature of recrystallized molybdenum. Welding techniques that create narrow heat-affected zones, such as electron-beam or laser welding, are usually more successfully employed than arc welding. Oxygen contamination of welds can result in intergranular oxygen segregation and embrittlement of the already low-ductility fusion zone, and therefore the use of low-oxygen inert-atmosphere glove boxes is necessary for best results when arc welding. (Glove boxes are sealed chambers that are fitted with gloves protruding into the chamber, providing a controlled atmosphere for the operator.) *See* BRAZING; WELDING AND CUTTING OF MATERIALS.

Alloys. Four main classes of commercial molybdenum-base alloys exist. One class relies on the formation of fine metal carbides that strengthen the material by dispersion hardening, and extend the resistance of the microstructure to recrystallization above that of pure molybdenum. A second class relies on solid-solution hardening to strengthen molybdenum. These two classes of materials are typically produced in both vacuum-arc-casting and powder metallurgy grades. The third class uses combinations of carbide formers and solution hardeners to provide improved high-temperature strength. This class of alloys is normally produced by powder metallurgy techniques, but some of the alloys are also amenable to vacuum-arc-casting processing. A final class of alloys, known as dispersion-strengthened alloys, relies on second-phase particles (usually an oxide of a ceramic material) introduced or produced during powder processing to provide resistance to recrystallization and to stabilize the recrystallized grain structure, enhancing high-temperature strength and improving low-temperature ductility. These latter materials by their very nature must be produced by powder metallurgy techniques. *See* HIGH-TEMPERATURE MATERIALS.

Carbide-strengthened alloys. The most common of the carbide-strengthened alloys is known as TZM, containing about 0.5% titanium, 0.08% zirconium, and 0.03% carbon. Other alloys in this class include TZC (1.2% titanium, 0.3% zirconium, 0.1% carbon), MHC (1.2% hafnium, 0.05% carbon), and ZHM (1.2% hafnium, 0.4% zirconium, 0.12% carbon). The high-temperature strength imparted by these alloys is their main reason for existence.

Both TZM and MHC have found application as metalworking tool materials. Their high-temperature strength and high thermal conductivity make them quite resistant to the collapse and thermal

cracking that are common failure mechanisms for tooling materials. A particularly demanding application is the isothermal forging process used to manufacture nickel-base superalloy gas-turbine engine components. In this process the dies and workpiece are both heated to the hot working temperature, and the forging is performed in a vacuum by using large hydraulic presses.

Solid solution alloys. Tungsten and rhenium are the two primary alloy additions in this class of alloys. A variety of compositions may be available as special orders, but the most common compositions are 30% tungsten (Mo-30W), 5% rhenium (Mo-5Re), 41% rhenium (Mo-41Re), and 47.5% rhenium (Mo-50Re). With the exception of the Mo-30W alloy which is available as a vacuum-arc-cast product, these alloys are normally produced by powder metallurgy. The tungsten-containing alloys find application as components in systems handling molten zinc, because of their resistance to this medium. They were developed as a lower-cost, lighter-weight alternative to pure tungsten and have served these applications well over the years. The 5% rhenium alloy is used primarily as thermocouple wire, while the 41% and 47.5% alloys are used in structural aerospace applications. These high-rhenium alloys trace their existence to the discovery of the rhenium effect, which produces a dramatic improvement in molybdenum's ductile-brittle transition temperature when alloyed with rhenium. This effect is also utilized by employing the Mo-50Re alloy as welding rod to improve the ductility of welds in molybdenum. A major stumbling block to the use of rhenium alloys is their high cost, which is related to the expense and price volatility of rhenium. *See* RHENIUM; TUNGSTEN.

Combined alloys. The beneficial effects of solid-solution hardening and dispersion hardening found in the carbide-strengthened alloys have been combined in the HWM-25 alloy (25% tungsten, 1% hafnium, 0.07% carbon). This alloy offers high-temperature strength greater than that of carbide-strengthened molybdenum, but it has not found wide commercial application because of the added cost of tungsten and the expense of processing the material.

Dispersion-strengthened alloys. The dispersion-strengthened alloys rely exclusively on powder metallurgy manufacturing techniques. This allows the production of fine stable dispersions of second phases that stabilize the wrought structure against recrystallization, resulting in a material having improved high-temperature creep strength as compared to pure molybdenum. Once recrystallization occurs, the dispersoids also stabilize the interlocked recrystallized grain structure. This latter effect produces significant improvements in the ductility of the recrystallized material.

The potassium- and silicon-doped alloys such as MH (150 ppm potassium, 300 ppm silicon) and KW (200 ppm potassium, 300 ppm silicon, 100 ppm aluminum) are the oldest of this category; they are analogs to the doped tungsten alloys in common use for tungsten lamp filament. They were first developed to satisfy the requirements of the lighting industry for creep-resistant molybdenum components. They do not possess particularly high strength at low temperatures, but they are quite resistant to recrystallization and they possess excellent creep resistance due to their stable, interlocked recrystallized grain structure. They are used in applications requiring high-temperature creep resistance such as nuclear fuel sintering boats. *See* CREEP (MATERIALS).

Composites. A family of copper-molybdenum-copper (CMC) laminates has been developed to serve the needs of the high-performance electronics industry. Laminating molybdenum with copper raises the coefficient of thermal expansion and improves the heat-transfer characteristics of the composite relative to pure molybdenum. This allows better matching between the composite, which serves as a packaging or mounting material, and the solid-state devices being employed. The coefficient of thermal expansion of the composite can be tailored to a specific application by selection of the copper/molybdenum ratio. Similar control can be exerted on thermal conductivity and elastic modulus, both properties of great interest to circuit designers. *See* ALLOY; CIRCUIT (ELECTRICITY); ELASTICITY; ELECTRICAL CONDUCTIVITY OF METALS; SOLID SOLUTION.

John A. Shields

Bibliography. J. Z. Briggs and R. Q. Barr, Arc-cast molybdenum-base TZM alloy: Properties and applications, *High Temperatures–High Pressures 1971*, vol. 3, pp. 363–409, 1971; Climax Molybdenum Co., *Molybdenum Metal*, 1960; J. B. Conway and P. N. Flagella, *Creep-Rupture Data for the Refractory Metals to High Temperatures*, 1971; J. A. Shields, Jr., Refractory Metals Forum: Molybdenum and Its Alloys, *Adv. Mater. Proc.*, 142(4):28–36, 1992.

Moment of inertia

A measure of the "opposition to rotation" around a fixed axis that a rigid body presents to a torque applied about that axis. The concept is used in analyzing the dynamics of systems in which rotation occurs, for example, in attitude control of spacecraft. It extends naturally into quantum mechanics and the notion of quantization of rotational energy.

Two-dimensional systems. The simplest system to illustrate moment of inertia is the rigid rotor (**illus.** *a*) lying in a horizontal plane. This is a rigid massless rod of length r, pivoted at the origin of a coordinate system, at the end of which is a point mass m. The state of the system is specified by one quantity: the angle coordinate ϕ. The angular velocity ω is given by Eq. (1). It follows from Newton's first law of motion that if no external forces are applied, ω remains constant. However, if a torque of magnitude L is applied to the system, centered on the origin, then from Newton's second law of motion it can be shown that

$$\omega \equiv \frac{d\phi}{dt} \tag{1}$$

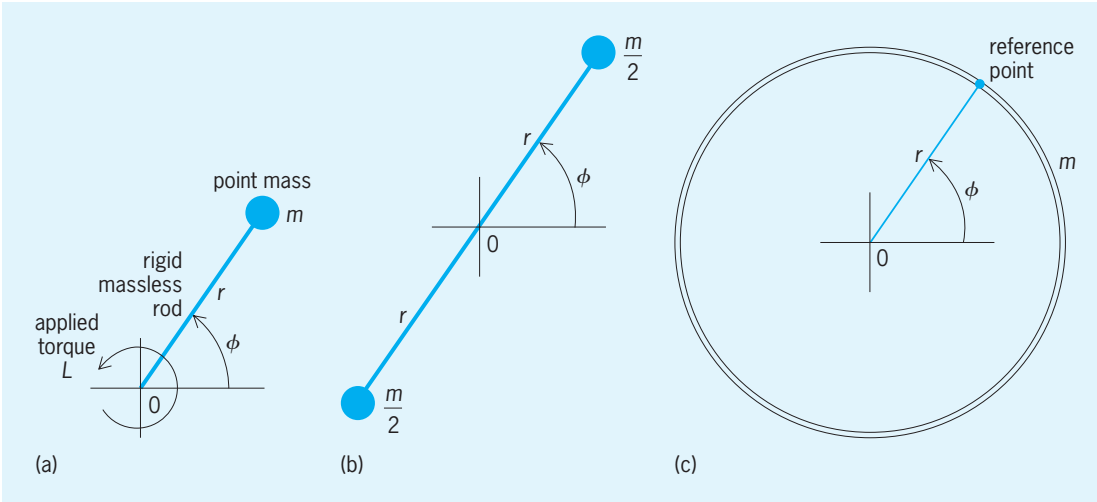

Three rotating systems for which the moments of inertia about the origin are all equal to mr^2. (*a*) Rigid rotor. (*b*) Baton. (*c*) Hoop.

L is related to the angular acceleration by Eq. (2). For

$$L = mr^2 \frac{d\omega}{dt} \tag{2}$$

any given system, the factor preceding the derivative is defined as the moment of inertia, usually symbolized by the letter I. Thus, the angular acceleration is given by Eq. (3). Evidently, for the rigid rotor, the moment of inertia is given by Eq. (4).

$$\frac{d\omega}{dt} = \frac{L}{I} \tag{3}$$

$$I_{\text{rotor}} = mr^2 \tag{4}$$

See ACCELERATION; VELOCITY; TORQUE.

For a rigid system of N mass particles lying in the plane, the total moment of inertia is simply the sum of the individual moments of inertia, that is, the sum given by Eq. (5). Thus, the baton in illus. *b*, having

$$I_{\text{total}} = \sum_{i=1}^{n} m_i r_i^2 = \sum_{i=1}^{n} I_i \tag{5}$$

masses $m/2$ at each end of a rod that is twice as long, has exactly the same moment of inertia as the rigid rotor, as given in Eq. (6).

$$I_{\text{baton}} = \frac{m}{2} r^2 + \frac{m}{2} r^2 = mr^2 = I_{\text{rotor}} \tag{6}$$

Spreading the total mass m evenly around a circle of radius r (illus. *c*) gives a hoop. It can be shown, using calculus, that $I_{\text{hoop}} = mr^2$ as well. However, if the mass is spread uniformly to form a disk of radius r, the moment of inertia is less, as given by Eq. (7).

$$I_{\text{disk}} = \tfrac{1}{2} mr^2 = \tfrac{1}{2} I_{\text{rotor}} \tag{7}$$

The radius of gyration k of a body is defined as the radius at which all its mass could be concentrated to yield the same moment of inertia; that is, k is given by Eq. (8). Thus, for the disk, the radius of gyration is

$$k \equiv \sqrt{\frac{I}{m}} \tag{8}$$

given by Eq. (9). Tables are available listing the values

$$k_{\text{disk}} = \frac{r}{\sqrt{2}} \tag{9}$$

of k for other shapes.

The angular momentum of a rotating system is given by Eq. (10), and its kinetic energy of rotation is given by Eq. (11).

$$J = I\omega \tag{10}$$

$$T_{\text{rotor}} = \tfrac{1}{2} I\omega^2 \tag{11}$$

See ANGULAR MOMENTUM; ENERGY.

Three-dimensional systems. The basic principles discussed above apply equally well to rigid bodies, or to systems of rigidly connected masses in three dimensions. The situation is more complex because not only can one choose any line in space as an axis of rotation, but the body itself may have any orientation with respect to that axis. However, analysis of rotation dynamics is greatly simplified for such systems in two ways. *See* RIGID-BODY DYNAMICS.

The first simplification is by the parallel-axis theorem. Assume that a rigid body of mass M is spun about an axis that passes through its center of mass and for which the moment of inertia is I_{COM}. Now consider a new axis, parallel to the first but shifted to a perpendicular distance R away. The parallel-axis theorem states that the new moment of inertia is given by Eq. (12).

$$I' = I_{\text{COM}} + MR^2 \tag{12}$$

See CENTER OF MASS.

The second simplification arises from the fact that no matter what shape the body may have, only six quantities are needed to allow full analysis. These are

the components of the matrix called the moment of inertia tensor **I**, is given by Eq. (13), where the

$$\mathbf{I} = \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{xy} & I_{yy} & I_{yz} \\ I_{xz} & I_{yz} & I_{zz} \end{pmatrix} \tag{13}$$

terms on the diagonals are the moments of inertia for rotation about the x, y, and z axes, respectively, and the off-diagonal quantities are the products of inertia.

To illustrate the practical meaning of the products of inertia, consider a rigid body rotating on an axle that passes through the center of mass and lies on, say, the x axis. Mechanical bearings at each end of the axle hold it in place. When the body spins, the fact that the products of inertia are not zero means that the bearings will be subjected to rotating sideways forces. In practice, this can lead to undesirable mechanical vibration and wear. For this reason it is not enough to balance automobile wheels statically such that the wheels' centers of mass lie on their axes of rotation: the wheels must also be dynamically balanced to avoid their vibration when the car is moving.

It can be shown that the matrix can always be transformed, through rotation of the assumed coordinate axes, such that the products of inertia are reduced to zero. The resulting axes are known as the principal axes, corresponding to which are the principal moments of inertia, I_a, I_b, and I_c, so the moment of inertia tensor is given by Eq. (14). By convention,

$$\mathbf{I}_{\text{principal}} = \begin{pmatrix} I_a & 0 & 0 \\ 0 & I_b & 0 \\ 0 & 0 & I_c \end{pmatrix} \tag{14}$$

the three principal moments of inertia are ordered in size as $I_a \geq I_b \geq I_c$.

The tennis-racket theorem applies to freely spinning bodies in free fall for which the principal moments of inertia are all different. It states that the body will spin stably about the axes corresponding to the largest principal moment (I_a) or the smallest (I_c); however, an initial spin about the axis corresponding to I_b will be unstable and quickly change to a spin state that is a combination of spins about the other two axes.

A useful visualization of the magnitude of the moment of inertia around various axes is the momental ellipsoid or inertial ellipsoid, given by Eq. (15). (The

$$x^2 I_x + y^2 I_y + z^2 I_z = 1 \tag{15}$$

principal axes are taken here to coincide with the x,y,z-coordinate axes). In the special case of a uniform sphere, any orthogonal set of axes with origin at the sphere's center are principal axes. The three principal moments are equal so that the momental ellipsoid is also a sphere.

Application to molecular dynamics. Among various spectroscopic techniques to deduce the structure of molecules is rotational spectroscopy. Molecules are generally small enough for quantum effects to become evident in that the spectra show discrete rates of rotation, each separated by one quantum of energy at the far-infrared and microwave frequencies. For a diatomic molecule spinning as a rigid rotor confined to a plane, the time-independent part of Schrödinger's equation reduces to Eq. (16), where

$$\frac{d^2\Phi(\phi)}{d\phi^2} = -\frac{2IE}{\hbar^2}\Phi(\phi) \tag{16}$$

Φ gives the variation of the molecule's wavefunction with ϕ, E is rotational energy, $\hbar$ is Planck's constant divided by 2π, and I is the molecule's moment of inertia. The solutions that satisfy the necessary periodicity of Φ with ϕ yield quantized energies given by Eq. (17), where K is restricted to the integer values

$$E_K = \frac{K^2\hbar^2}{2I} \tag{17}$$

$K = 0, \pm 1, \pm 2, \ldots$. *See* MOLECULAR STRUCTURE AND SPECTRA; SCHRÖDINGER'S WAVE EQUATION.

Application to gravitational radiation. Einstein's general theory of relativity predicts the existence of gravitational waves. It is calculated that only very massive physical processes, such as stars collapsing into black holes, can generate gravitational waves powerful enough to be detectable with current technology. In principle, however, a spinning rod of any size will generate gravitational waves. According to theory, the rotational kinetic energy of a freely spinning rod is dissipated, by the gravitational waves it generates, at a rate given by Eq. (18), where G is the

$$\frac{dT_{\text{rotor}}}{dt} = \frac{32G\omega^6}{5c^5} I_{\text{rod}}^2 \tag{18}$$

gravitational constant and c is the speed of light. *See* GRAVITATION; GRAVITATIONAL RADIATION; RELATIVITY.

Andrej Tenne-Sens

Bibliography. J. Foster and J. D. Nightingale, *A Short Course in General Relativity*, 3d ed., Springer-Verlag, 2006; G. R. Fowles and G. L. Cassiday, *Analytical Mechanics*, 7th ed., Thomson Brooks/Cole, 2004; L. N. Hand and J. D. Finch, *Analytical Mechanics*, Cambridge University Press, 1998; G. K. Vemulapalli, *Physical Chemistry*, Prentice Hall, 1993.

Momentum

A foundational quantity in physics that is always conserved in isolated systems, defined in classical mechanics as the product of mass and velocity. An equivalent definition that is better suited to modern physics is that momentum is the rate of change of energy with respect to velocity: this allows extending the conservation laws to include entities of zero mass. Whenever energy is transported from one system to another, it is always associated with some form of momentum. *See* CLASSICAL MECHANICS.

The word momentum (Latin for motion) by itself usually means linear momentum, the vector quantity obtained by multiplying (or scaling) the particle's velocity vector by its mass. The linear

momentum of a system is the sum of the linear momenta of the particles composing it. By Newton's second law of motion, a particle's linear momentum increases with time in direct proportion to the applied force. *See* DYNAMICS; FORCE; LINEAR MOMENTUM; MOTION; NEWTON'S LAWS OF MOTION.

A particle's angular momentum is a vector quantity defined as its moment of (linear) momentum about a specified axis. The angular momentum of a system is the sum of the angular momenta (about the same axis) of the particles composing it. A system's angular momentum increases with time in direct proportion to the total applied moment of force, or torque, about the axis. *See* ANGULAR MOMENTUM; ROTATIONAL MOTION; TORQUE.

The concept of momentum is of fundamental importance because in isolated physical systems (that is, ones not acted on by some outside influence) both linear and angular momentum do not change with time: they are said to be conserved. (This statement also applies to total system energy.) The law of conservation of linear momentum comes from the fact that the laws of physics do not depend on where the system is located (the universe is homogeneous). The law of conservation of angular momentum comes from the fact that the laws of physics do not depend on the orientation of the system (the universe is isotropic). The two conservation laws of momentum along with the law of conservation of energy are the three foundational axioms of classical physics. *See* CONSERVATION LAWS (PHYSICS); CONSERVATION OF ENERGY; CONSERVATION OF MOMENTUM; KINEMATICS; MECHANICS; SYMMETRY LAWS (PHYSICS).

For continuous systems such as semirigid bodies, fluids, and electromagnetic radiation, momentum is often expressed in the form of momentum density, that is, momentum per unit volume.

When velocities approach the speed of light, observed mass increases and Newton's laws of motion must give way to Einstein's more accurate laws of relativity. Nevertheless, the same form of law of conservation of linear momentum remains valid if the relativistic increase of mass is taken into account. Relativity views space and time united into a single entity, called spacetime, and correspondingly unites momentum and energy into momentum-energy, or momenergy. Just as the spacetime interval between events is invariant (the same for all observers), the rest-mass component of a particle's momentum-energy is also. This invariance implies that particles with finite momentum but no mass, such as photons, can travel only at the speed of light. *See* PHOTON; RELATIVITY.

In systems where quantum-mechanical effects are significant, the exact outcome of an experiment cannot be forecast, only the probability of obtaining any one of all the possible outcomes. Nevertheless, each experiment will show that linear and angular momentum are conserved. *See* QUANTUM MECHANICS.

Generalized momentum. The conservation laws of classical mechanics are embodied in the powerful analysis methods of J. L. Lagrange and W. R. Hamilton, applicable to constrained dynamical systems whose degrees of freedom are limited (or at least countable). *See* CONSTRAINT; DEGREE OF FREEDOM (MECHANICS).

The state of a constrained system can be uniquely specified by an appropriate choice of generalized coordinates represented by q_j, one for each degree of freedom. In Lagrangian analysis, the system's total kinetic energy T is written as a function of all the q_j and the generalized velocities $\dot{q}_j \equiv dq_j/dt$. Inserting T into Lagrange's equations (1) yields one differential

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) = \frac{\partial T}{\partial q_j} \tag{1}$$

equation (of second degree in time) for each degree of freedom, $j = 1, 2, 3, \ldots$, and so forth.

The resulting set of simultaneous differential equations are then solved for the desired dynamic variables. The term $\partial T/\partial \dot{q}_j$ is called the generalized momentum; it is conserved in isolated systems. For systems that do not dissipate energy (through friction, say), T is replaced by the Lagrangian function $L \equiv T - V$, where V is the system's total potential energy. *See* LAGRANGE'S EQUATIONS.

As a simple example, consider a bead of mass m constrained to move on a fixed circle of radius R; this system has only one degree of freedom. A natural (but not unique) choice of generalized coordinate is the angle θ, as seen from the circle's center, between the bead and a fixed reference direction. It can be shown that the bead's kinetic energy is $T = \frac{1}{2}m(R\dot{\theta})^2$. Differentiating T with respect to $\dot{\theta}$ gives the generalized momentum: $\partial T/\partial\dot{\theta} = mR^2\dot{\theta}$. For this particular case, the familiar formula for angular momentum is obtained.

It is sometimes convenient to convert the Lagrangian into the Hamiltonian function H, a function of the generalized coordinates and their generalized momenta, represented by p_j. In conservative systems, H equals the total energy $T + V$. Analysis proceeds by plugging H into Hamilton's equations (2), which generate two differential equations

$$\dot{q}_j = \frac{\partial H}{\partial p_j} \qquad \dot{p}_j = -\frac{\partial H}{\partial q_j} \tag{2}$$

(of first degree in time) for each degree of freedom j.

This set of simultaneous differential equations is then solved for the desired results. Because of the symmetry of these equations, p_j and q_j are said to be canonically conjugate, and the p_j are called canonical momenta or conjugate momenta. *See* CANONICAL COORDINATES AND TRANSFORMATIONS; HAMILTON'S EQUATIONS OF MOTION.

In the equations of nonrelativistic quantum mechanics, the generalized coordinates are replaced formally by their corresponding operators. *See* NONRELATIVISTIC QUANTUM THEORY; SCHRÖDINGER'S WAVE EQUATION.

Momentum in unified theories. A major area of research is to reconcile the current contradictions between quantum theory and general relativity. Just as Newton's laws of motion are approximations to those of relativity, it may turn out that the laws of relativity are themselves approximations. For example, the hypothesis of loop quantum gravity, that

spacetime is quantized, carries the implication that the speed of light is not constant. All candidate theories, however, must be compatible with current conservation laws in the limit of everyday mass, time, and length scales. Conversely, any future experiments revealing situations in which momentum (or energy) is not conserved will illuminate the way to a self-consistent unified theory. *See* QUANTUM GRAVITATION. Andrej Tenne-Sens

Bibliography. R. P. Feynman, R. B. Leighton, and M. L. Sands, *The Feynman Lectures on Physics: The Definitive and Extended Edition*, Addison-Wesley, Reading, MA, 2005; H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., Pearson Education, San Francisco, 2002; R. Penrose, *The Road to Reality: A Complete Guide to the Laws of the Universe*, Jonathan Cape, London, 2004; L. Smolin, *Three Roads to Quantum Gravity*, Basic Books, New York, 2001; A. Sommerfeld, *Lectures on Theoretical Physics*, vol. 1: *Mechanics*, Academic Press, New York, 1964; E. F. Taylor and J. A. Wheeler, *Spacetime Physics: Introduction to Special Relativity*, 2d ed., W. H. Freeman, New York, 1992.

Monazite

A rare mineral that incorporates the light rare-earth elements (lanthanum, cerium, praseodymium, neodymium, promethium, samarium, europium, gadolinium) and also yttrium. Monazite has a general formula of (La,Ce,Nd)PO_4, but Pr, Sm, Eu, Gd, and Y substitute for La, Ce, and Nd in solid solution in minor amounts. The dominant rare-earth element in a particular monazite is denoted by the atomic suffix, such as monazite-(Ce) in which cerium exists in amounts greater than other rare-earth atoms. Monazite-(Ce), monazite-(La), and monazite-(Nd) are officially recognized by the International Mineralogical Association. *See* CERIUM; MINERAL; RARE-EARTH ELEMENTS; YTTRIUM.

Monazite is one of the main ore minerals for the rare-earth elements that are used in the manufacture of television and computer screens, fluorescent light bulbs, and highly efficient batteries, among other industrial applications.

Monazite occurs mainly as a minor mineral in granites, and is concentrated in sediments and other rock types following the erosion of the host granites. In the United States, monazite is found in economic amounts in metamorphic rocks and stream deposits in North Carolina. Most of the world's supply of monazite comes from beach sands eroded from granites in Australia, Brazil, India, and Malaysia. *See* GRANITE.

The atomic arrangement of monazite is formed of a packing arrangement of (PO_4) tetrahedra and distorted (REO_9) polyhedra, where RE = the rare-earth elements in the particular monazite mineral. The arrangement is formed of chains of alternating phosphate tetrahedra and RE polyhedra, parallel to the *c* axis. Monazite is similar in structure and chemistry to the tetragonal mineral xenotime, Y(PO_4), that selectively incorporates the heavy rare-earth elements. *See* PHOSPHATE MINERALS.

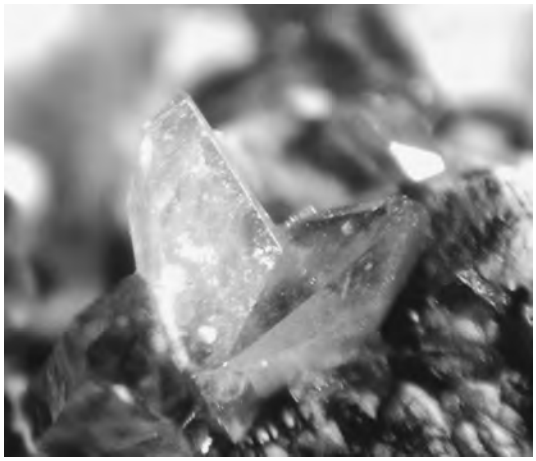

Crystals of monazite.

Monazite crystallizes in the monoclinic crystal system (see **illus.**), crystal class 2/m, space group $P2_1/n$, where $a \approx 6.8$, $b \approx 7.0$, $c \approx 6.5$Å, and $\beta \approx 103.4°$. The mineral is variably green, yellow, brown, or red-brown, and rarely occurs in crystals large enough to discern with the unaided eye. Mohs hardness is 5–5.5, and the specific gravity is 4.6–5.5, varying with substitution of different elements. *See* HARDNESS SCALES.

Monazite is often radioactive because of the substitution of the element thorium for the rare-earth elements. Although monazite typically incorporates relatively minor amounts of thorium, the rarity of this element makes monazite the chief thorium ore. Thorium is used for the generation of atomic energy, although in minor amounts. Its principal use is in the manufacture of mantles for incandescent gas lanterns. *See* RADIOACTIVE MINERALS.

Because of the ability of monazite to incorporate radioactive elements, such as thorium, extensive research has been undertaken to determine if the mineral can be synthesized with large amounts of radioactive elements in it. Much of this work has been pioneered by researchers at the Oak Ridge National Laboratory (Tennessee) in order to determine if synthetic monazite can be used to isolate radioactive waste from nuclear power plants and weapons production facilities. This important work has shown that synthetic monazite can incorporate significant amounts of radioactive waste, and is thus an effective method of encapsulating the waste in a solid form. John M. Hughes; John Rakovan

Bibliography. L. A. Boatner and B. C. Sales, Monazite, in W. Lutze and R.C. Ewing (eds.), *Radioactive Waste Forms for the Future*, pp. 495–564, Elsevier Science, Amsterdam, 1988; L. L. Y. Chang, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, vol. 5B: *Non-silicates, Sulphates, Carbonates, Phosphates and Halides*, Longman Group, England, 1996; C. Klein and C. S. Hurlbut, Jr., *Manual of Mineralogy*, 21st ed., John Wiley, 1993; Y. Ni, J. M. Hughes, and A. N. Mariano, Crystal chemistry of the monazite and xenotime structures, *Amer. Mineral.*, 80:21–26, 1995.

Mongoose

Any of a group of 38 species of carnivorous mammals in the family Herpestidae. Mongooses inhabit a broad band along both sides of the Equator from western Africa to Southeast Asia. Although native to southern Asia, the East Indies, Africa, and Madagascar, one genus (*Herpestes*) was introduced into Spain and Portugal as well as to other parts of Europe and to many islands around the world including the West Indies, Fiji, and the Hawaiian Islands (see **table**).

Description. Mongooses are long slender carnivores with long faces, small rounded ears, short legs, and long tapering bushy tails. Most have long coarse, usually grizzled fur and lack body markings, although a few have stripes (see **illustration**). Color ranges from dark gray through brown to yellowish or reddish. *Galidia elegans* from Madagascar has five to seven dark bands on its tail. Four or five toes are present on each foot, and the claws are nonretractile. Most species have a large anal sac containing at least two glandular openings. Adult mongooses have a head–body length of 24–60 cm (9.5–23 in.), a tail length of 19–44 cm (7.5–17 in.), and a weight ranging from 320 g to 5 kg (11 oz to 11 lb). The dwarf mongoose (*Helogale parvula*) is the smallest species, while the white-tailed mongoose (*Ichneumia albicauda*) is among the largest. The dental formula is I 3/3 C 1/1 Pm 3-4/3-4, M 2/2 × 2 for a total of 36–40 teeth.

Mongooses and their distribution

Name	Range
Malagasy ring-tailed mongoose	
Galidia elegans	E., W., and N. Madagascar
Malagasy broad-striped mongooses	
Galidictis fasciata	E. Madagascar
G. grandidieri	S.W. Madagascar
Malagasy narrow-striped mongoose	
Mungotictis decemlineata	W. and S.W. Madagascar
Malagasy brown-tailed mongoose	
Salanoia concolor	N.E. Madagascar
Mongooses	
Herpestes ichneumon	Spain, Portugal, Morocco, Tunisia, Egypt, Libya, Africa
H. javanicus	Iraq, India, S. China, Malay Peninsula
H. edwardsii	Arabian Peninsula to Afghanistan, India, and Sri Lanka
H. smithii	India, Sri Lanka
H. fuscus	S.W. India, Sri Lanka
H. vitticollis	S.W. India, Sri Lanka
H. urva	Nepal, China, Taiwan, Malaysia
H. semitorquatus	Sumatra, Borneo
H. brachyurus	Malaysia, Sumatra, Borneo
H. nasa	Nigeria, Zaire
Gray and slender mongooses	
Galerella pulverulenta	South Africa
G. nigrita	Angola, Namibia
G. sanguinea	Africa
G. ochracea	Somalia
Banded and Gambian mongooses	
Mungos gambianus	Gambia, Nigeria
M. mungo	Gambia, Ethiopia, South Africa
Cusimanses	
Crossarchus obscurus	Sierra Leone to Ghana
C. platycephalus	Benin, Cameroon, Congo
C. alexandri	Zaire, Uganda, Central African Republic, Zambia
C. ansorgei	Zaire, Angola
Liberian mongoose	
Liberiictis kuhni	Liberia, Ivory Coast
Dwarf mongooses	
Helogale parvula	Ethiopia, Angola, South Africa
H. hirtula	Ethiopia, Somalia, Kenya
African tropical savannah mongoose	
Dologale dybowskii	Central African Republic, Zaire, Sudan, Uganda
Black-legged mongooses	
Bdeogale crassicauda	Yemen, Kenya, Mozambique
B. nigripes	Nigeria, Zaire, Angola
B. jacksoni	Uganda, Kenya
Meller's mongoose	
Rhynchogale melleri	Zaire, Tanzania, South Africa, Angola
White-tailed mongoose	
Ichneumia albicauda	Arabian Peninsula, Senegal, Egypt, Namibia, South Africa
Marsh mongoose (water mongoose)	
Atilax paludinosus	Senegal, Ethiopia, South Africa
Yellow mongoose	
Cynictis penicillata	Angola, Namibia, Botswana, Zimbabwe, South Africa
Gray meerkat (Soule's mongoose)	
Paracynictis selous	Angola, Zambia, Malawi, Namibia, Botswana, Zimbabwe, Mozambique, South Africa
Suricate (slender-tailed meerkat)	
Suricata suricatta	Angola, Namibia, Botswana, South Africa

Banded mongoose, *Mungos mungo*. (*Photo by Gerald and Buff Corsi; © 2004 California Academy of Sciences*)

Mongooses are agile and active. Most species are terrestrial, although some are semiaquatic and arboreal. They seek shelter in hollow logs or trees, in rock crevices, or in holes in the ground. Some, such as *Cynictis* and *Suricata*, live in colonies in underground burrows. They may be either nocturnal or diurnal. Depending upon the species, they inhabit a wide range of habitats ranging from dense forests to open woodland, savanna, semidesert, and desert. The marsh mongoose (*Atilax paludinosus*) is a good swimmer and inhabits marshes, reed-grown streambeds, and tidal estuaries. Mongooses are omnivorous, feeding on insects, crabs, fish, frogs, snakes, birds and their eggs, small mammals, fruits, and other vegetable matter.

Breeding and development. Breeding may be seasonal or continue throughout the year with females having two or three litters annually. Litters of one to four are born after a gestation period ranging from 42 to 105 days. Studies of dwarf mongooses (*Helogale*) in Serengeti National Park revealed groups of four or five adults of each sex living together, with all of the adults cooperating to raise the offspring. Depending upon the species, sexual maturity may be reached as early as 9–10 months or by 2 years of age. A normal life-span in the wild would be about 6–10 years, although in captivity the dwarf mongoose (*Helogale parvula*) has lived 13 years, the marsh mongoose (*Atilax paludinosus*) for 19+ years, and the Malagasy ring-tailed mongoose (*Galidia elegans*) for 24 years 5 months.

Impact. Mongooses, particularly the genus *Herpestes*, have been introduced into many regions of the world. For example, the small Indian Mongoose (*Herpestes javanicus*) has been introduced into Costa Rica, Cuba, Antigua, Dominican Republic, Jamaica, Puerto Rico, West Indies, Fiji, Hawaii, Japan, and other areas. Originally introduced to help control populations of rodents and venomous snakes in sugarcane plantations, they quickly multiplied and spread to become pests since there are no natural predators on many of these islands. They have destroyed not only the rats and snakes but also harmless reptiles, birds, and mammals. They have contributed to the extinction or endangerment of many desirable species of wildlife, as well as preying upon domestic poultry and consuming eggs. They have caused the extinction of seven species of amphibians and reptiles in Puerto Rico, and are one of the major factors in the extinction of a lizard, a snake, two birds, and a rodent in Jamaica. Mongooses are widespread in Hawaii and have had a significant impact on native species. Now the importation or possession of mongooses is forbidden by law in some countries, including the United States and Australia. In many areas, rats now live in trees, where they are safe from mongooses.

Threats. Mongooses may be preyed upon by larger carnivores and raptors. Humans hunt some species as pests. Habitat destruction has caused some species to be listed as vulnerable (*Galidictis fasciata, Mungotictis decemlineata, Salanoia concolor, Bdeogale jacksoni*) or endangered (*Galidictis grandidieri, Liberiictis kuhni, Bdeogale crassicauda omnivora*) by the International Union for the Conservation of Nature and Natural Resources (IUCN). Some species can be easily domesticated. They are fairly intelligent and can be taught simple tricks, so they are often kept as pets to protect the home from vermin. *See* CARNIVORA; SCENT GLAND.

Donald W. Linzey

Bibliography. D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Monhysterida

An order of nematodes in which, generally, the stoma is funnel shaped and lightly cuticularized; however, in some families the stoma is spacious and heavily cuticularized, and is armed with protrusible teeth. The amphids vary from simple spirals to circular forms. Usually the second and third circlets of cephalic sensilla are combined, but in some taxa the third circlet of four distinct setae is separate. The normal pattern of distribution is often disrupted by numerous cervical setae. The near-cylindrical esophagus is sometimes swollen posteriorly. The cuticle may be smooth or may have annuli or ornamentation. When the annuli are distinct, the somatic setae may be long and in four to eight longitudinal rows. The female gonads are outstretched and either single or paired.

There are three monhysterid superfamilies, Linhomoeoidea, Monhysteroidea, and Siphonolaimoidea. The Linhomoeoidea and Siphonolaimoidea are primarily marine forms; feeding habits among the groups are unknown. The Monhysteroidea are free-living nematodes found in all environments, from marine waters to fresh waters and soil; feeding habits are unknown. *See* NEMATA (NEMATODA).

Armand R. Maggenti

Monkey

An adaptive or evolutionary grade among the primates, represented by members of two of the three modern anthropoid superfamilies. The New World, platyrrhine monkeys (Ateloidea) and Old World, catarrhine forms (Cercopithecoidea) probably reached a monkey level of adaptation independently some time after their separation from a common ancestor, perhaps 45 million years ago (Ma; **Fig. 1**). The term monkey is not indicative of taxonomic or phylogenetic relationship: the closest relatives of the cercopithecoids are not the ateloid monkeys but the Old World apes and humans. *See* APES; FOSSIL APES; FOSSIL PRIMATES; PRIMATES.

The Ateloidea comprise two families, while the living Cercopithecoidea are considered today to comprise only one family, with two subfamilies. A modern classification of the Anthropoidea follows:

Hyporder Anthropoidea
- Infraorder Platyrrhini
  - Superfamily Ateloidea (New World or platyrrhine monkeys)
    - Family Atelidae
      - Subfamily Atelinae (howler and spider monkeys)
      - Subfamily Pitheciinae (saki, owl, and titi monkeys)
    - Family Cebidae
      - Subfamily Cebinae (capuchin and squirrel monkeys)
      - Subfamily Callitrichinae (marmosets and tamarins)
    - Family Branisellidae (extinct early ateloids)
- Infraorder Catarrhini (Old World anthropoids)
  - Parvorder Eucatarrhini (modern catarrhines)
    - Superfamily Hominoidea (gibbons, great apes, and humans)
    - Superfamily Cercopithecoidea
      - Family Cercopithecidae (Old World or catarrhine monkeys)
        - Subfamily Cercopithecinae (cheek-pouched monkeys: macaques, baboons, geladas, mangabeys, and guenons)
        - Subfamily Colobinae (leaf eaters: langurs and colobus)
        - Subfamily Victoriapithecinae (extinct early cercopithecids)
  - Parvorder Eocatarrhini (archaic catarrhines)
    - Family Propliopithecidae (early archaic catarrhines)
    - Family Pliopithecidae (later archaic catarrhines)
  - Infraorder Paracatarrhini (extinct early anthropoids)
    - Family Oligopithecidae (extinct archaic anthropoids)
    - Family Parapithecidae (extinct Egyptian monkeys)

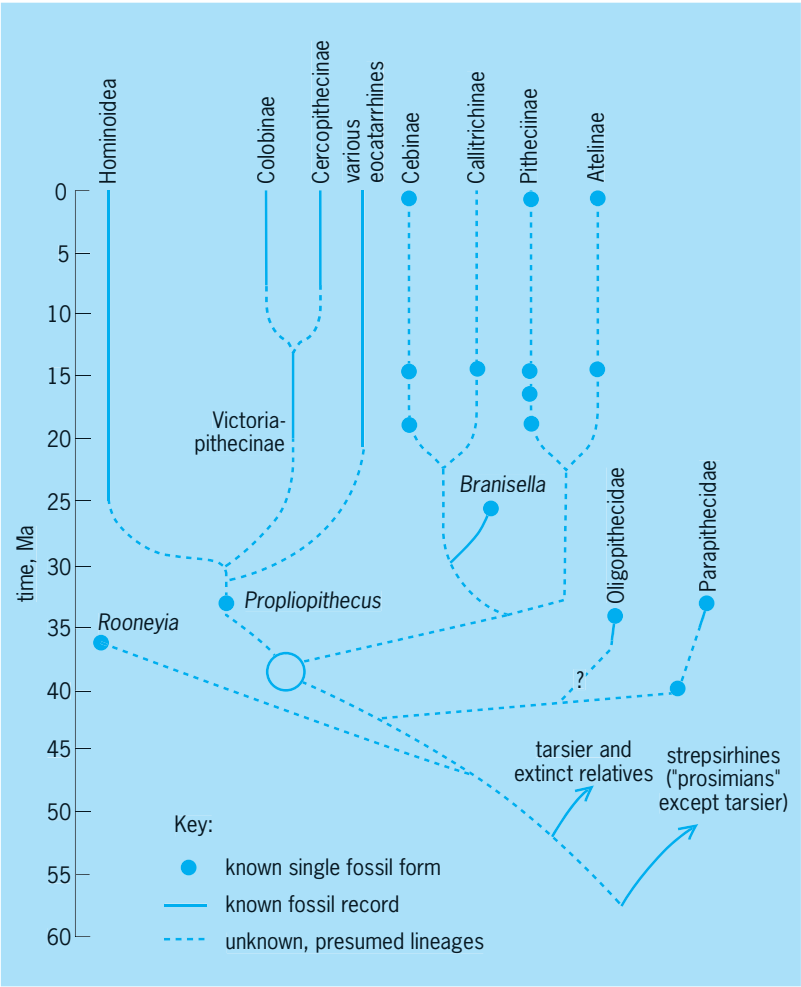

Fig. 1. Evolutionary relationships of monkeys among the primates.

Monkeys are hard to characterize as a group because of their great diversity, and because much of the discussion reflects a comparison with the apes. Both monkeys and apes contrast with the prosimian grade in that they are typically large, diurnal animals that live in social groups. Monkeys differ from apes in their possession of a tail, a smaller brain, quadrupedal pronograde (with long axis of body horizontal) posture, and a usually longer face (**Fig. 2**). They are generally smaller than apes, but large monkeys outweigh gibbons. Like almost all primates, monkeys are pentadactyl (**Fig. 3**), with nails rather than claws on the digits in most cases. They have pectoral mammary glands and well-developed vision. Monkeys are primarily vegetarian and inhabit forested tropical or subtropical regions of Africa, Asia, and South America. The differences between the New and Old World monkeys are summarized in the **table**.

Origins. The origin of the anthropoid or higher primates is still under debate in terms of the source group, area, and timing. The tarsier (*Tarsius*) is probably the living primate sharing the most recent common ancestry with anthropoids, and the tarsier-like extinct Omomyidae is accepted as a likely ancestral stock. The omomyids are known in all northern continents, especially between 55 and

Major contrasts between New and Old World monkeys and special features of each

New World species (Ateloidea)	Old World species (Cercopithecoidea)
Nose platyrrhine (nasal septum wide, nostrils opening to sides)	Nose catarrhine (septum narrow, nostrils opening downward)
Tail long, prehensile only in atelines and *Cebus*	Tail short to long, nonprehensile
3 premolar teeth in each quadrant	2 premolars in each quadrant
24 deciduous, 36 permanent teeth (4 fewer in Callitrichinae), I 2/2 C 1/1 P 3/3 M 3/3/(M 2/2 in Callitrichinae)	20 deciduous, 32 permanent teeth, I 2/2 C 1/1 P 2/2 M 3/3
No ischial callosities	Ischial callosities present
Jaws and teeth lightly built in Cebidae; more robust in Atelidae, with deep lower jaw	Cheek pouches in Cercopithecinae; sacculated stomach in Colobinae
Opening on skull to internal ear circular but flat ("ringlike")	Opening extended into tube
Fingers and toes with curved nails (clawlike in Callitrichinae)	All nails tending to be flattened
Big toe opposable, thumb not fully so and sometimes reduced in Cebidae	Thumb and big toe opposable, thumb reduced in Colobinae

35 Ma. At that time, South America and Africa were much closer, and the ancestors of New World monkeys may have traveled there from Africa on natural vegetation rafts, following paleocurrents across a 600-mi (1000-km) ocean gap. However, this has seemed unlikely to some researchers, because such rafted small primates would probably die of thirst or exposure and because known African fossils were not reasonable ancestors. Instead, it was hypothesized that anthropoid origins are from omomyids in North America or eastern Asia, regions that were connected during the Eocene (50–40 Ma at least) by a Bering land bridge. Early anthropoids might then have entered South America from Central America or the Caribbean by crossing narrow water gaps on rafts or by island hopping. New finds of protoanthropoid primates in northern Africa have reenergized the trans-Atlantic hypothesis, and both views are now strongly supported. Moreover, recent studies have suggested that the omomyids are not a close-knit group but may need to be broken into several distinct units; it is possible that the relatively late (36 Ma) Texas "protoanthropoid" *Rooneyia* might be the closest known relative of the common ancestor of all anthropoids listed above.

In China and Burma, several fossil sites 45–40 Ma old have yielded the remains of primates claimed to be protoanthropoids, but these are better interpreted as early strepsirhines or tarsier relatives. Others are known from North African localities dated about 55–40 Ma. Most of these, again, are either strepsirhines or tarsiiforms. Some authors have suggested that one or more of these newly recovered fossils represent a previously unknown group of early primates (unrelated to the omomyids) which gave rise directly to the anthropoids, while others consider them to be omomyid offshoots. Whichever hypothesis is supported, it is generally agreed that the most complete remains of early anthropoids are known from the Fayum area of northern Egypt.

In this region, three groups of fossil anthropoids have been found. The Parapithecidae (about 37–30 Ma) are monkeylike in adaptation and may be considered a third type of monkey that is not closely related to either living group. They have been suggested as being ancestral to Cercopithecoidea, but are specialized in their own ways while lacking the different specializations of the cercopithecoids (or the ateloids; see table). For example, some parapithecid species had complex teeth with extra cusps, while one had no lower incisor (front) teeth. Probably, they occupied ecological niches later exploited by cercopithecoids, but became extinct without descendants. Previously, it was thought that they represented the earliest known catarrhines, but it now seems more likely that they were part of a radiation of "primitive" anthropoids which preceded the split between catarrhines and platyrrhines. It appears that a number of isolated teeth belonging to early parapithecids have been recovered from the 42–46-Ma

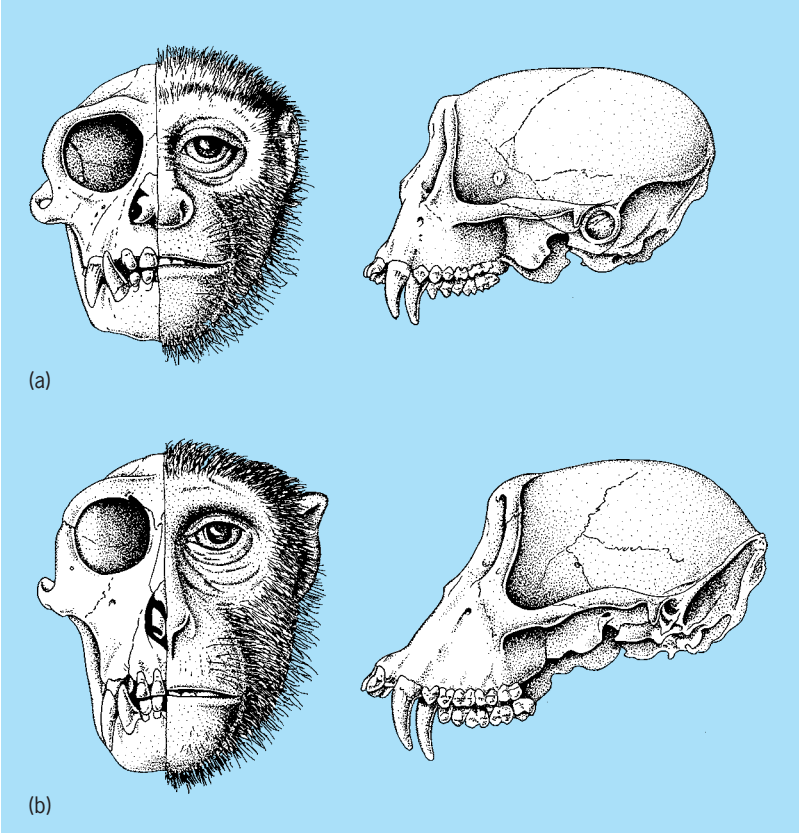

Fig. 2. Side and front views of the heads and skulls of adult male monkeys: (*a*) *Cebus* (New World) and (*b*) *Macaca* (Old World). (*After A. H. Schultz, The Life of Primates, Universe Books, 1969*)

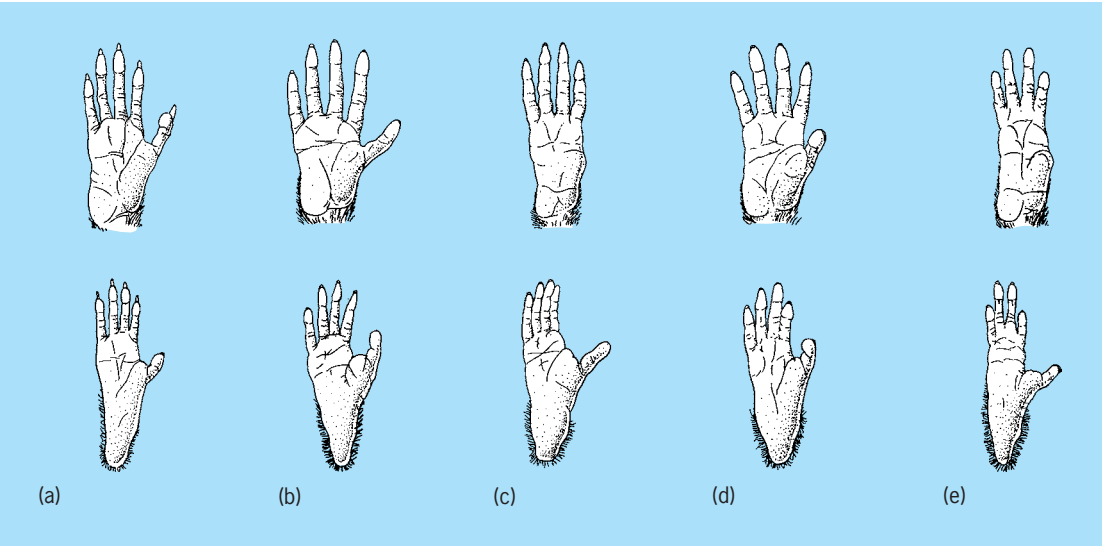

Fig. 3. Right feet (below) and hands (above) of (*a*) *Saguinus*, (*b*) *Cebus*, (*c*) *Ateles*, (*d*) *Macaca*, and (*e*) *Colobus*. (*After A. H. Schultz, The Life of Primates, Universe Books, 1969*)

site of Glib Zegdou, Algeria, which would make them the oldest recognized anthropoids.

Another group of Fayum primates which may belong to the same paracatarrhine radiation is the oligopithecids. Occurrence of these forms ranged around 36–34 Ma, and recently several partial damaged skulls have been recovered. They do not seem to have been monkeylike in adaptation and are really on the border of being anthropoids. Their teeth and lower jaw were conservative, but their skull presents the important anthropoid features of fusion in the middle of the frontal bone (forehead) and complete closure of the back of the orbit (eye socket). It is interesting that the oligopithecids are far more "primitive" than the contemporaneous (and even much older) parapithecids, suggesting that a yet-unknown species of oligopithecid must have lived even earlier than the first parapithecids, perhaps about 45–48 Ma.

The third group of Fayum primates is the Propliopithecidae (32–30 Ma). These species are monkeylike but also in some ways almost apelike; they may well have been close to the common ancestors of cercopithecoid monkeys and hominoids (apes and humans). Although some authors have suggested that the oligopithecids are close relatives of the propliopithecids, this appears unlikely—their similarities are convergent (independent evolutionary developments) rather than homologous (inherited from a recent common ancestor).

Old World species. Today, these animals are found throughout all the warmer regions of the Eastern Hemisphere except Australia and Madagascar. Many of the familiar monkeys are included in this family, such as the rhesus macaque, Barbary "ape," proboscis monkey, baboon, and mandrill. The earliest known representatives of the Cercopithecidae, *Prohylobates* and *Victoriapithecus* (the Victoriapithecinae), are from the early Miocene in north and east Africa, some 20–12 Ma. They probably diverged from their apelike relatives by adapting to a relatively more folivorous (leaf-eating) diet, involving reshaping of the dental chewing surfaces, and also by becoming more terrestrial. Victoriapithecines are often considered to be a common ancestral stock for all later cercopithecids, but new studies suggest they might instead be closely related to the cercopithecines; these alternative hypotheses need to be carefully evaluated.

By the late Miocene, some 10–9 Ma, colobines were present in east Africa (*Microcolobus*) as well as in Europe (*Mesopithecus*), where one line (*Dolichopithecus*) became highly terrestrial before dying out late in the Pliocene. Colobines probably entered Asia in the later Miocene as well, and many fossil forms are known in the African Plio-Pleistocene. Living (and presumably most extinct) members of this subfamily are primarily arboreal leaf eaters and have sharp teeth and specialized stomachs like ruminants to process difficult-to-digest leaf protein. The dispersal of the cercopithecines may have been somewhat later, with the oldest macaque-like forms appearing in north Africa about 7 Ma, whence they spread to Europe and Asia before 5 Ma. In sub-Saharan Africa, possible ancestors of the baboons and mangabeys (the extinct *Parapapio*) are present from the Pliocene, and early relatives of the modern geladas (*Theropithecus*) were widespread in the Pliocene and Pleistocene. The *Cercopithecus* group probably separated from the macaques and baboons in the late Miocene and entered the high forest, but fossils are very rare. All cercopithecines have cheek pouches for temporary food storage, and many are terrestrial.

Macaques. About 12–16 species of the genus *Macaca* are known as macaques, and with the exception of *M. sylvanus*, the Barbary "ape," all are found in southern Asia. The Barbary ape, so called because it lacks an external tail like true apes, occurs in the wild in Algeria and Morocco, and is the only monkey now found in Europe, where it

was introduced to Gibraltar. It is thought to be the remains of a population that originally extended from Britain to the Caucasus, north Africa, and southwestern Asia. *Macaca sylvanus* is a large, robust, and rather terrestrial species with a thick coat that protects it from low temperatures in its natural habitat, where it roams through cedar forests and mountain gorges in large troops. Its diet consists of fruit, leaves, and other vegetation, in addition to small invertebrates such as insects; in fact, this diet is typical of most cercopithecine monkeys. Because these animals raid farm areas and beg for food, they have become a nuisance on Gibraltar. Their gestation period is about 30 weeks, with a single young born almost annually, as in most monkeys. The adult male stands over 2 ft (66 cm) at the shoulder and weighs up to 40 lb (18 kg).

Macaca mulatta is the rhesus monkey, which is found throughout southern continental Asia. Large numbers were previously exported from India each year for medical research, but restrictions have been adopted to protect the species. These monkeys have been used in studies on the effects of space travel and in the development of polio vaccine. The rhesus, or Rh, blood group was so named because the antigen was first found in the red blood cells of rhesus macaques. Further genetic studies have shown that the eastern rhesus of China is in some features more closely related to the macaques of Taiwan and Japan than to the western rhesus of India. This suggests that an offshoot of the eastern rhesus may have given rise to the island populations (now separate species) within the past million years while still remaining part of the same species as the rhesus of India. This is one of the few cases of one living species of primate (or any mammal) being the ancestor of another.

The rhesus is an agile, gregarious species with a short tail, robust limbs of almost equal length, and a stocky build. It is considered sacred in some parts of the Indian subcontinent and has become a scavenger in towns and cities, as well as living in larger troops in its natural forested habitat. Troops of 15–40 are led by an older, experienced male who is dominant to the others in terms of access to food, mates, and living space. These animals do not defend a true territory, but groups generally avoid each other. Macaques have large cheek pouches, which they fill with food. They then wander about eating directly, and when danger approaches, they scatter to the nearest place of safety with a ready supply of food. The adult male is about 2 ft (66 cm) high; he is sexually mature at 4 years and full grown at 5 years, with a maximum lifespan of about 20 years. Mating is promiscuous within a troop, with the dominant male probably siring most offspring. The gestation period is about 23 weeks. On average, males weigh 18 lb (8 kg) and females 11 lb (5 kg).

The natural history of some of the other species of macaques is less well known, although studies are proliferating. Some live at high altitudes, and one (the Japanese *Macaca fuscata*) roams in winter snows. The lion-tailed macaque (*M. silenus*) of southern India is one of the least well known and least typical of the group, being a shy animal that inhabits dense forests. The stump-tailed macaque (*M. arctoides*) is found at high altitudes from Assam to southeastern China. It has long fur, and the naked face is brightly colored, so it is also known as the red-faced monkey. Some researchers have suggested that it is the evolutionary result of hybridization between two other living species.

Fig. 4. Mandrill, a large Old World cercopithecoid monkey.

Mandrills, baboons, and geladas. These are terrestrial animals with long canines in the males, a long muzzle with terminal nostrils, and a great disparity in size between the sexes. Mandrills (*Mandrillus sphinx*) are found on the forest floor of west-central Africa. The males have brightly colored red and blue facial markings for recognition by subordinate troop members (**Fig. 4**). In both sexes, the body is robust, the legs stout, and the tail short. Males weigh up to 77–99 lb (35–45 kg), making them the heaviest of all extant monkeys; females weigh only about one-third as much as large males. Mandrills thus have the highest ratio of sexual dimorphism of all primates. The closely related drill (*M. leucophaeus*) is smaller and less brightly colored than the mandrill but otherwise quite similar. Troop size varies from a dozen (with one adult male) to hundreds.

Typical or "savannah" baboons of the genus *Papio* include at least five or six subspecies of *P. hamadryas*. They were previously thought to represent separate local species, but studies of contact zones have revealed significant interbreeding and hybridization. As full species are defined to be reproductively isolated, the baboon groups are now usually considered only subspecies. *Papio hamadryas* inhabits a "7"-shaped belt of mainly savannah and open woodland from Guinea to Ethiopia (and coastal Arabia) to south Africa. The body and limbs of *Papio* baboons are more slender than the mandrill's (**Fig. 5**), and body size ranges widely: adult males average 29–64 lb (16–29 kg) in the various subspecies, while female averages are 22–34 lb

Fig. 5. Olive baboon (*Papio hamadryas anubis*). (© *California Academy of Sciences; photograph by Dr. Robert Thomas and Margaret Orr*)

(10–16 kg). Troops are often led by a coalition (central hierarchy) of high-ranking males and contain 30–200 members of all ages. In some varieties, especially *P. h. hamadryas*, large troops sleep together but divide during the day into one-male "harems" of 4–14 animals for feeding. Other varieties of social organization may depend on local environmental conditions. *Theropithecus*, commonly called the gelada baboon, is a very distinctive genus; its fossil record goes back to about 4 Ma. Today, it is restricted to the Ethiopian highlands, where it lives in one-male groups on rocky gorges. It eats mostly grasses.

Mangabeys. There are five species in this group, all from central Africa. Two sets of species are recognized, the *torquatus* and *albigena* groups, which were long placed in the genus *Cercocebus*. One feature which linked them to each other is the presence of deep depressions or fossae in the bone below the orbit (eye socket). Recent studies of molecular biology as well as of limb bones and teeth have argued that these two groups are not, in fact, each other's closest relatives, and suggested persuasively that they be placed in separate genera. *Lophocebus albigena* and relatives are almost entirely arboreal, with long tails and long fore- and hindlimbs that make them excellent climbers. They are thought to be evolutionarily linked to the baboons and geladas. *Cercocebus torquatus* and relatives may be more terrestrial, using their strong forelimbs to search through leaf litter, much as do their close relatives, the mandrills. Both groups of mangabeys live in troops with one or several male leaders and are frugivorous to omnivorous. It is not yet clear whether the suborbital fossae found in mangabeys were evolved independently or inherited from a common ancestor whose other cercopithecine decendants lost this feature.

Guenons. These are the dozen or so species of monkey in the genus *Cercopiithecus* and allies. The mona monkey (*C. mona*) is typical of guenons found in the forested areas of west and central Africa. Several species may inhabit a single type of tree, avoiding competition by vertical spacing and slightly different diets, including fruits, some leaves and insects, as well as tree snails. These more arboreal guenons hardly ever come to the ground; they are agile climbers and leap between treetops. Most groups have a harem social organization, with additional males peripheralized. The vervet (*C. aethiops*) is one of the most common of all primates in Africa, where it is widely distributed throughout the grasslands, being equally at home on the ground and in trees. Its yellow-brown fur often has a greenish tinge, and the males engage in a "red, white, and blue" display of the perineal region. Recent studies of DNA (supported by revised work on limb bones) have demonstrated that the partly terrestrial vervets and Lhoest's guenons are in fact close evolutionary relatives of the even more highly terrestrial patas monkey (previously placed in the genus *Erythrocebus*). As a result, these three animals are now placed in a separate genus, *Chlorocebus*, which may in turn be divided into subgenera for each species [for example, *Chlorocebus* (*Erythrocebus*) *patas*]. The tiny talapoin (*Miopithecus*) is the smallest Old World monkey, with most males and females averaging only 3 lb (1.3 kg). The little-known Congo forest form *Allenopithecus* (the swamp monkey) is in some ways a link between the guenon and the baboon-macaque groups; it retains some of the features of their common ancestor also still seen in the latter species while presenting a number of the special characters of the guenons as well.

Langurs. These animals are slender, long-tailed members of the subfamily Colobinae, which are found in forested areas of southern Asia and Indonesia. They may live in large troops or small ones with a single male leader. Two genera of "typical" langurs are now recognized: *Presbytis* for the smaller, mainly island species, and *Semnopithecus* (including *Trachypithecus*) for the larger, mostly mainland forms. Seven species of *Presbytis* occupy peninsular Malaysia, Sumatra, Borneo, eastern Java, and many smaller islands such as the Mentawais. They are excellent leapers and show little size difference between the sexes. *Semnopithecus* species range from Sri Lanka and India into China, through Malaysia and onto Sumatra, Java, Borneo, and some other Indonesian islands. They tend to leap less and are more quadrupedal, with moderate to strong sexual size dimorphism. Eight to ten of these species are placed in the subgenus *Trachypithecus*, while the subgenus *Semnopithecus* includes the Indian langurs, especially *S. entellus*. This is probably the best-known species of the group and the most terrestrial of living colobines. It superficially resembles the spider monkey but of course lacks the prehensile tail. The entellus langurs are common in India up to altitudes of 13,000 ft (4000 m) and are considered sacred by Hindus. It has been observed that males of *P. entellus* and other langurs may kill infants in a troop when ousting a previous leader male. The "adaptive" versus "aberrant" nature of this pattern is debated.

Fig. 6. Black-and-white Colobus monkey.

An unusual Asian colobine is the proboscis monkey (*Nasalis larvatus*) found only in Borneo. This is a powerfully built animal that stands about 2.5 ft (83 cm) tall; weight of a large male averages 44 lb (20 kg). The unusual nose, which may function as a resonating chamber during male calling, is a swollen pendant structure up to 3 in. (7.6 cm) long that, at rest, hangs down over the mouth. These animals live in troops of about 20, feeding on mangroves and palm leaves. They are expert swimmers and have been observed far out to sea. Other related species have upturned noses, such as *N.* (*Simias*) *concolor* of the Mentawai Islands and *Pygathrix* (*Rhinopithecus*) *roxellana*, the snub-nosed monkey, which is an inhabitant of Tibet and southern China and may spend part of its life in regions of perpetual snow. The douc langur (*Pygathrix nemaeus*) of Indochina completes the roster of the group often called the "odd-nosed" colobines.

Colobus. In Africa the Colobinae are represented by two genera, one of which has two marked subgenera. All are forest-dwelling leaf eaters with rudimentary external thumbs (those of the Asian colobines are short but present). The black-and-white *Colobus* (**Fig. 6**) has five species, including *C. polykomos* and *C. guereza*, which have long been hunted for their magnificent pelts. They live in troops of 5–15 with a single male leader and defend a small territory. The genus *Procolobus* includes the so-called red and olive colobus monkeys. *Procolobus* (*Piliocolobus*) *badius* and related species, the red colobus monkeys, does not exhibit territorial behavior but live in multimale groups of 20–50 in a larger home range and have more varied dietary preferences. All these forms range across equatorial Africa, and there is much uncertainty about how many species (or subspecies) to recognize. Less is known of the much smaller olive colobus, *P. verus*, restricted to west Africa.

New World species. The New World monkeys, or ateloids, occupy forested areas from southern Mexico to Argentina. They are divided into two main groups, or families (their major characteristics are given in the table). All are arboreal, including a few with prehensile tails; there is no living form, or any evidence of a fossil form, that has come to the ground habitually. The fossil record of platyrrhines is relatively limited, but documents early diversification of the main lineages, as opposed to the successive replacements seen among catarrhines. The oldest known ateloid, *Branisella*, dates back 26 Ma and may be related to the Cebidae. A variety of monkeys inhabited southern Argentina between 16 and 20 Ma, including pitheciines related to owl monkeys (*Tremacebus*) and titis (*Homunculus*) and two cebines, one apparently close to squirrel monkey ancestry (*Dolichocebus*), and the more "generalized" *Killikaike*, described in 2006 from a partial cranium with a relatively large brain. *Chilecebus* was found in 20-Ma deposits in northern Chile and may be a relative of *Dolichocebus*. In Colombia, 13-Ma deposits at La Venta have yielded forms very close to the living night, squirrel, saki, spider, and howler monkeys and to marmosets. Brazilian cave deposits dated between 100,000 and 10,000 years ago yielded fossils of two large ateline species in the 1990s. Two or three distinctive genera also inhabited the islands of Jamaica, Cuba, and Hispaniola during the past 10,000 years.

Marmosets. The subfamily Callitrichinae consists of a number of small-sized species of four or five genera. Body weight (for males and the similarly sized or even slightly larger females) ranges across many species from 1 to 4.5 oz (110 to 620 g), making callitrichines the smallest of all living anthropoid primates. Most eat insects and fruits, and live in family, or pair-bonded, groups of male, female, and offspring. Births may occur twice a year, and twins are common, so that up to eight young may be born in the 2 years the offspring spend with their parents. Both sexes share equally in the care of the young, the male especially carrying them when traveling. Several forms, particularly the brightly colored golden lion marmoset (*Leontopithecus rosalia*), are severely endangered species. The marmosets were previously thought to be very primitive forms (with claws rather than nails and simple teeth) uniquely separate from all other ateloids. However, research has shown them to be specialized sap feeders, with secondarily evolved clawlike nails and reduced dentition. They share common ancestry with the cebines, which has resulted in the classification of Cebidae and Atelidae as employed here.

Capuchins. The family name Cebidae comes from the capuchin monkey, *Cebus*, which lives in forested areas from Honduras to Argentina in troops of up to 30 individuals (**Fig. 7**). Capuchins are extremely agile leapers and runners in their arboreal habitat. They are omnivorous, feeding on small birds, insects, grubs, leaves, and fruits, and often raid plantations for fruit and grain. Studied in captivity, they have been observed to use simple tools, make drawings, and paint. They have a partly prehensile tail, but it is rather different from that of the atelids.

The squirrel monkeys (*Saimiri*) are close relatives and some of the most widely distributed animals of South America. They are relatively small animals, living in troops of up to 100 individuals, apparently "dominated" by females, with older males peripheralized.

Titis. Three species of *Callicebus* are found in South America in the region of the Amazon, where they live in the finer branches of the high forest trees. These animals (male or female) weigh about 2–3 lb (0.8–1.3 kg), are omnivorous, and live in pair-bonded groups of five or six. They are strongly territorial, and the adults are very dependent on one another, often sitting (or sleeping) side by side on a branch with tails intertwined. A close relative is *Aotus*, the owl or night monkey, the only nocturnal platyrrhine (or anthropoid, for that matter).

Sakis and uakaris. This group of generally little-known pitheciines comprises fruit-eating rainforest dwellers. The sakis have long hair and a bushy tail. There are four species, found in the Guianas and the Amazon basin: *Pithecia monachus*, the hairy saki; *P. pithecia*, the pale-headed saki; *Chiropotes satanas*, the black-nosed saki; and *C. albinasus*, the white-nosed saki. Sakis do not thrive in captivity, and have been known to die simply from the shock of being moved from one cage to another.

The three species of uakari, genus *Cacajao*, are relatively rare and occur in restricted ranges. They are found high in the tops of trees and rarely descend to the ground. The head and face are bright scarlet, and the tail is short and of little use as a balancing organ.

Howler monkeys. All six species of the howler are included in the genus *Alouatta*. They range from South America into Central America and are among the largest of the New World monkeys. Males reach 15–22 lb (7–10 kg). They are primarily leaf-eaters that can hang by the prehensile tail to reach food, much of which they drop after eating partially, and they spend most of their day resting. This is now seen to be an adaptation to processing the large quantity of toxic secondary compounds in their food. Plants which taste bad (that is, are potentially poisonous) may be discarded, while long rest periods allow the howler's stomach to detoxify the hard-to-digest leaf proteins consumed. Their most unusual characteristic is the howling voice, which is perhaps used to space the troops of 20 to 30 apart from each other. It can be heard miles away. The enlarged, cavernous vocal apparatus is a trumpetlike bony box that is accommodated by the enlarged lower jaw and throat.

Fig. 7. Capuchin monkey (*Cebus*).

Fig. 8. Red spider monkey (*Ateles* sp.). (© *California Academy of Sciences; photograph by Dr. Lloyd Glenn Ingles*)

Spider and wooly monkeys. Comprising three genera, these monkeys also have prehensile tails but often lack an external thumb. Four species of spider monkey (*Ateles*; **Fig. 8**) are often found alongside howlers, but they are more active and concentrate on eating fruit, insects, and fewer leaves. They range between Mexico and Bolivia in groups of 15–50, with either one or several adult males. The wooly monkeys (*Lagothrix*) are like heavyset spider monkeys but do not lack a thumb, while the rare wooly spider monkey (*Brachyteles*) is similarly built and lacks a thumb, as do the spiders. Eric Delson

Bibliography. G. Davies and J. F. Oates (eds.), *Colobine Monkeys: Their Ecology, Behaviour and Evolution*, 1994; E. Delson et al., Body mass in Cercopithecidae (Primates, Mammalia): Estimation and scaling in extinct and extant taxa, *Anthropol. Pap. Amer. Mus. Nat. Hist.*, 83:1–159, 2000; E. Delson et al. (eds.), *Encyclopedia of Human Evolution and Prehistory*, 2d ed., 2000; J. G. Fleagle, *Primate Adaptation and Evolution*, 2d ed., 1998; C. Groves, *Primate Taxonomy*, 2001; W. C. Hartwig (ed.), *The Primate Fossil Record*, 2002; N. G. Jablonski (ed.), *The Natural History of the Doucs and Snub-nosed Monkeys*, 1998; N. G. Jablonski (ed.), *Theropithecus: Rise and Fall of a Primate Genus*, 1993; W. G. Kinzey (ed.), *New World Primates: Ecology, Evolution and Behavior*, 1997; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., 1999; N. Rowe, *The Pictorial Guide to the Living Primates*, 1996; R. J. Smith and W. L. Jungers, Body mass in comparative

primatology, *J. Hum. Evol.*, 32:523–559, 1997; F. S. Szalay and E. Delson, *Evolutionary History of the Primates*, 1979; P. F. Whitehead and C. J. Jolly (eds.), *Old World Monkeys*, 2000.

Monocleales

An order of liverworts of the subclass Marchantiidae, consisting of a single genus (*Monoclea*). The gametophyte is among the largest of all liverworts; its thallus consists of homogeneous cells except for scattered oil cells. The rhizoids are smooth and both thick-walled and thin-walled. The sex organs are grouped but not elevated: the antheridia are sunken in cavities and grouped into receptacles, while the archegonia are enclosed in groups by involucres. The very long, massive seta considerably elevates the capsules which dehisce spoonlike by one slit. The lobing of spore mother cells is unique in the subclass Marchantiidae. *See* BRYOPHYTA; HEPATICOPSIDA; MARCHANTIIDAE. Howard Crum

Bibliography. E. O. Campbell, The structure and development of *Monoclea forsteri* Hook, *Trans Roy. Soc. N. Z.*, 82:237–248, 1954; Looking at *Monoclea* again, *J. Hattori Bot. Lab.*, 55:315–319, 1984.

Monoclonal antibodies

Antibody proteins that bind to a specific target molecule (antigen) at one specific site (antigenic site). Monoclonal antibodies are produced by a laboratory culture of cells (a cell line), grown from a single antibody-secreting cell. Development of the technology for making monoclonal antibodies has had an impact in both science and medicine; applications include experiments in basic immunology, medical diagnostics, and therapeutics.

Production. A cell line that produces monoclonal antibodies is created by a combination of immunological and tissue culture techniques (see **illus.**). First, a foreign substance (antigen) is injected into a mouse. The mouse immune system reacts to the antigen by producing antibodies that bind to it. An antibody is a protein molecule which binds to a specific conformational feature (antigenic site) of another molecule via a variable binding region. Immunization with a complex antigen results in the production of a number of different antibodies binding to its different conformational features (antigenic sites). In the spleen each antibody-producing cell makes only one type of antibody, which reacts with a single antigenic site. An immunized mouse will have many antibodies in the serum component of its blood (antiserum) which react with many different features of a complex antigen. Its spleen will contain as many different antibody-producing cells, each making one type of antibody with a single binding specificity for one conformational feature of the antigen. If these spleen cells were removed from the mouse and cultured in the laboratory with nutrients, they would survive only for a few days. In 1975 G. Köhler and C. Milstein developed a technique for immortalizing spleen cells in laboratory culture and simultaneously retaining their antibody secretion ability. This allowed the isolation and culture of individual antibody-secreting cells and the subsequent production of monoclonal antibodies. Köhler and Milstein were awarded the Nobel prize in medicine in 1984 for this contribution.

Köhler and Milstein were working with an immortal cell line, grown from a mouse myeloma tumor which consisted of cancer cells that produced antibody of unknown specificity. They found that when a cell from the myeloma cell line was combined by fusion with a spleen cell from an immunized mouse, a unique type of hybrid cell (hybridoma) was formed. The hybridoma secretes antibody with binding specificity of the antibody made originally by the parent spleen cell. A hybridoma also retains the immortal growth properties of the parent myeloma cell line. Isolation of such hybrid cells is facilitated by a mutational defect in the myeloma cell line which prevents it from growing in a selective medium unless it can use metabolic enzymes provided by a normal cell fused to it. Growth in selective medium to isolate hybrid cells is a classic tissue culture technique, developed by J. Littlefield.

After antibody-secreting hybrid cells are produced, they are diluted in growth medium so that individual hybrid cells arising originally from individual spleen cells are separated from each other and will grow into individual colonies, called clones. Clones are then maintained as separate cultures, each producing only one type of antibody with specificity for a single antigenic site. The antibodies produced by these clones are monoclonal antibodies.

Applications. The ability to produce monoclonal antibodies has revolutionized immunology and the use of immunological techniques in other disciplines. Previously, antibodies could be isolated only from serum of an immunized animal, and analysis of an antibody response reflected the combination of activities and properties of all the different antibodies in that antiserum. But by making monoclonal antibodies from an immunized animal, it has been possible to analyze individual molecular properties of antibodies produced in a single immune response and to understand more about the mechanism by which antibodies are generated.

Production of monoclonal antibodies has also made it possible to take advantage of the exquisite specificity of individual antibodies. In cases where antisera had been used to identify a complex antigen, a monoclonal antibody can be used more confidently and precisely. With a monoclonal antibody, there is never the question of whether more than one antigenic site is being detected. In addition, there is a virtually unlimited supply of monoclonal antibody from cultured hybridoma cells, whereas antiserum from animals never has the same composition of antibodies, even when an immunization protocol is reproduced exactly. Cells are among the most complex antigens which have been used to elicit antibodies,

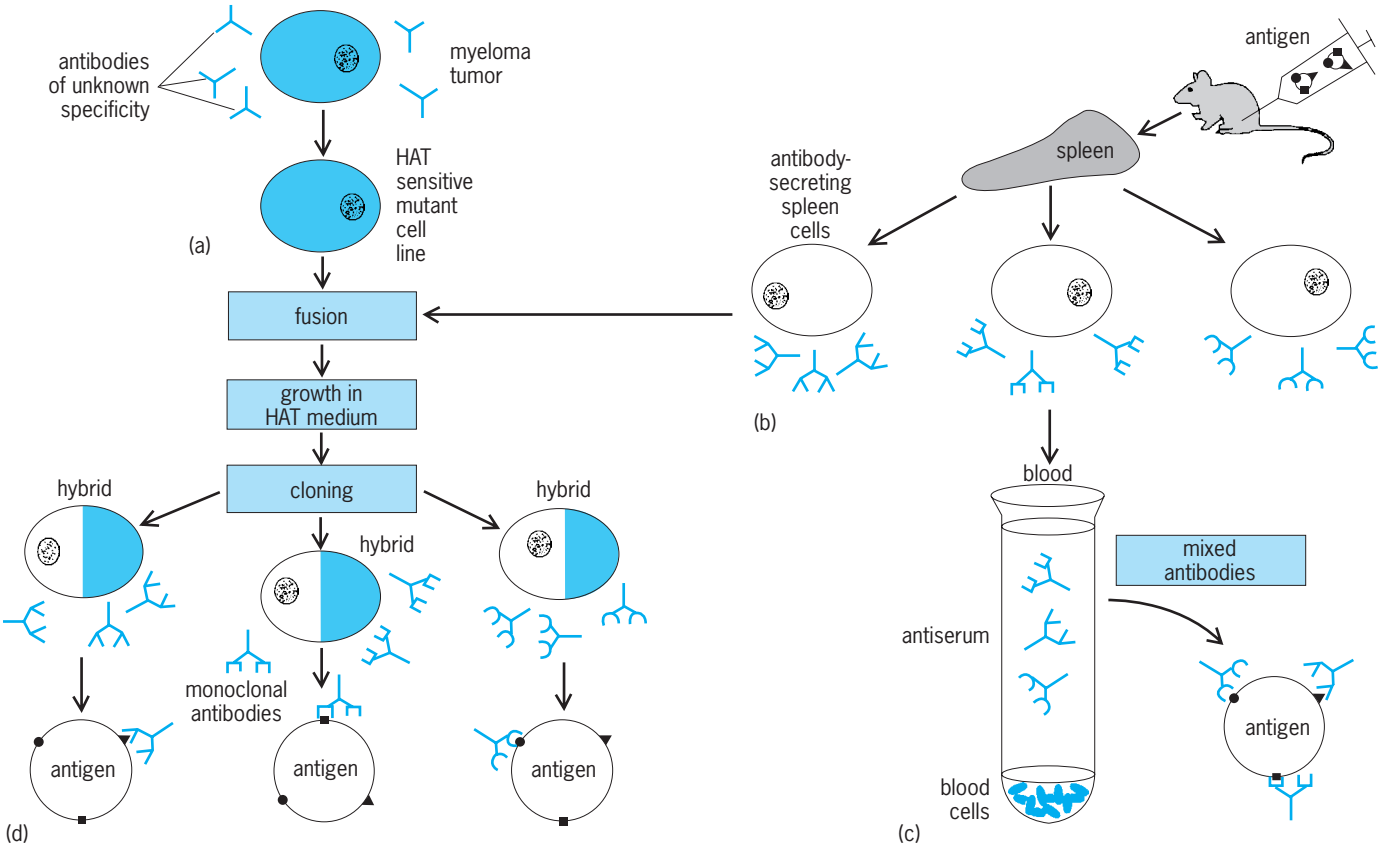

Monoclonal antibody production. (*a*) A myeloma-derived cell line is used to select a mutant that cannot metabolize the nutrients in HAT (hypoxanthine, aminopterin, thymidine) medium. The mutant cell line is fused with (*b*) an antibody-secreting spleen cell. (*c*) Antibodies from the spleen cells are secreted into the serum component of blood, forming an antiserum. (*d*) After fusion, a monoclonal antibody-producing cell line (hybrids) results. (*After C. Milstein, Monoclonal antibodies, Sci. Amer., 243:66–74, 1980*)

and classifying cells by using antisera has been difficult. Monoclonal antibodies reacting with known molecular targets on the surface of cells have helped to distinguish between different types of cells when other criteria have failed. For example, they have been particularly useful in identifying functional subsets of leukocytes.

The potential ability of monoclonal antibodies to recognize structures specific to many different cell types has heightened research interest in the "magic bullet" concept. An antibody with specificity for a particular cell could possibly be used to deliver a toxic substance or nutrient to that cell without affecting surrounding cells. For this reason, monoclonal antibodies may become useful in cancer therapy. The possibility of using monoclonal antibodies for therapy has raised the question of whether the mouse-derived or rat-derived monoclonal antibodies will be tolerated by the human system. Early experiments with mouse monoclonal antibody therapy in humans demonstrated that the recipient does react to the administered antibody. For this reason, laboratories have attempted to produce human-derived monoclonal antibodies, but with only moderate success; no human equivalent of the mouse myeloma parent cell line has been found. There has been some success through using several techniques to immortalize human antibody-secreting cells, but it is difficult to isolate human cells secreting antibody to a selected antigen, given the necessary limitations on immunizations.

While the use of monoclonal antibodies for therapy is still in early phases of development, their application in diagnostics has grown. Many diagnostic tests use antisera to detect cells, infectious organisms, or levels of blood and fluid components (such as hormones and proteins). Appropriate monoclonal antibodies can be used instead of an antiserum in all such tests, with the advantage of unlimited supply and greater reproducibility and specificity. *See* ANTIBODY; ANTIGEN; IMMUNOLOGY; TISSUE CULTURE.

Frances M. Brodsky

Bibliography. G. Köhler and C. Milstein, Continuous cultures of fused cells secreting antibody of predefined specificity, *Nature*, 256:495–497, 1975; J. W. Littlefield, Selection of hybrids from matings of fibroblasts *in vitro* and their presumed recombinants, *Science*, 145:709–710, 1964; C. Milstein, Monoclonal antibodies, *Sci. Amer.*, 243(4):66–74, 1980; J. H. Peters and H. Baumgarten (eds.), *Monoclonal Antibodies: A Practical Guide*, 1992; M. A. Ritter and H. M. Ladgman (eds.), *Monoclonal Antibodies: Production Engineering and Clinical Application*, 1994.

Monocotyledons

This group of flowering plants (angiosperms), with one seed leaf, was previously thought to be one of the two major categories of flowering plants (the other group is dicotyledons). However, deoxyribonucleic acid (DNA) studies have revealed that, although they do constitute a group of closely related families, they are closely related to the magnoliids, with which they share a pollen type with a single aperture. The other dicots, the eudicots, are much more distantly related. In general, monocots can also be recognized by their parallel-veined leaves and three-part flowers. Their roots have disorganized vascular bundles, and if they are treelike (yuccas, aloes, and dracaenas) their wood is unusually structured. Among the important monocots are grasses (including corn, rice, and wheat), lilies, orchids, palms, and sedges. *See* DICOTYLEDONS; EUDICOTYLEDONS; FLOWER; GRASS CROPS; LILIALES; MAGNOLIOPHYTA; ORCHID.

Mark W. Chase; Michael F. Fay

Monogenea

A class of the Platyhelminthes which are ectoparasites of the gills, skin, and orifices of fishes and, less frequently, of the esophageal tracts and bladders of amphibians and turtles. They have conspicuous anterior and posterior holdfasts, the latter usually armed. The terminal genitalia are frequently sclerotized. The group is characterized by sexual reproduction, direct development, and a single host in the life cycle. Egg capsules have terminal filaments.

Taxonomy. The distinctive shapes and sclerotized holdfasts facilitate characterization and arrangement (**Fig. 1**). The most widely used classification employs two subclasses, the Monophisthocotylea, in which the posthaptor is without discrete multiple suckers or clamps, and the Polyopisthocotylea, with suckers or clamps on the posthaptor.

Morphology. Body shapes of the various genera are distinctive, sometimes bizarre, as in *Vallisia*, which is sickle-shaped. Paired external suckers or buccal cavity suckers and adhesive glands occur anteriorly. The posterior holdfast is either solid and armed with central anchors and marginal hooks (Gyrodactyloidea), sucker-shaped with anchors and hooks (Capsaloidea), or solid and bearing suckers or clamps (Polyopisthocotylea). The pharynx is muscular and protrusible, and the esophagus is short, often ramified. The gut is commonly bifid, often much ramified. Multiple testes and numerous vitelline glands are usually present. A genitointestinal canal occurs in many, and the penes are usually armed.

Life history. Monogenea usually have direct development involving simple metamorphosis from the ciliated larval stage to the nonciliated juvenile. Juvenile anchors and hooks may be retained or replaced by adult suckers or clamps. Cross-fertilization or, perhaps less frequently, self-fertilization of hermaphroditic individuals resulting in egg capsules which hatch on the host or in its

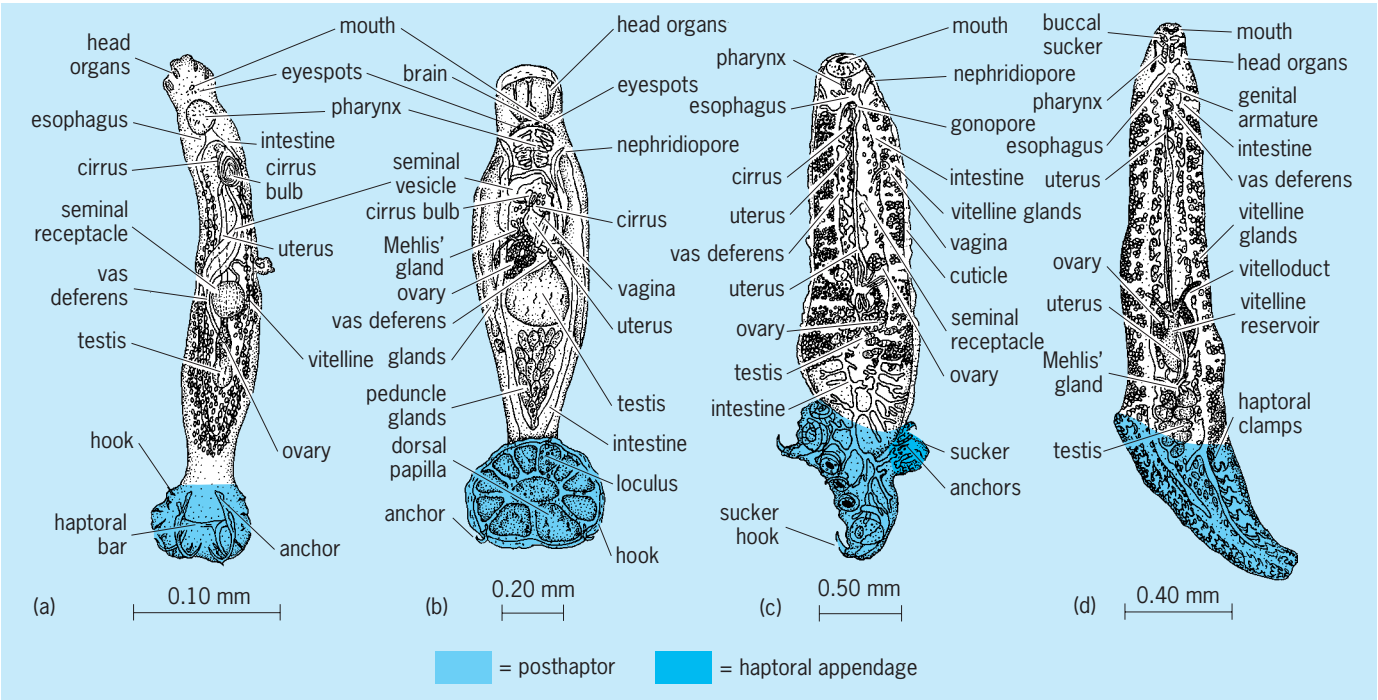

Fig. 1. Representative monogeneids. (*a*) Superfamily Dactylogyroidea, *Pseudohaliotrema carbunculus* from the pinfish (*Lagodon rhomboides*), shown in dorsal view. (*b*) Superfamily Capsaloidea, *Heterocotyle aetobatis* from the spotted eagle ray (*Aetobatis narinari*), ventral view. (*c*) Superfamily Polystomatoidea, *Heteronchocotyle leucas* from the bull shark (*Carcharhinus leucas*), ventral view. (*d*) Superfamily Microcotyloidea, *Heteraxinoides xanthophilis* from the spot (*Leiostomus xanthurus*), ventral view.

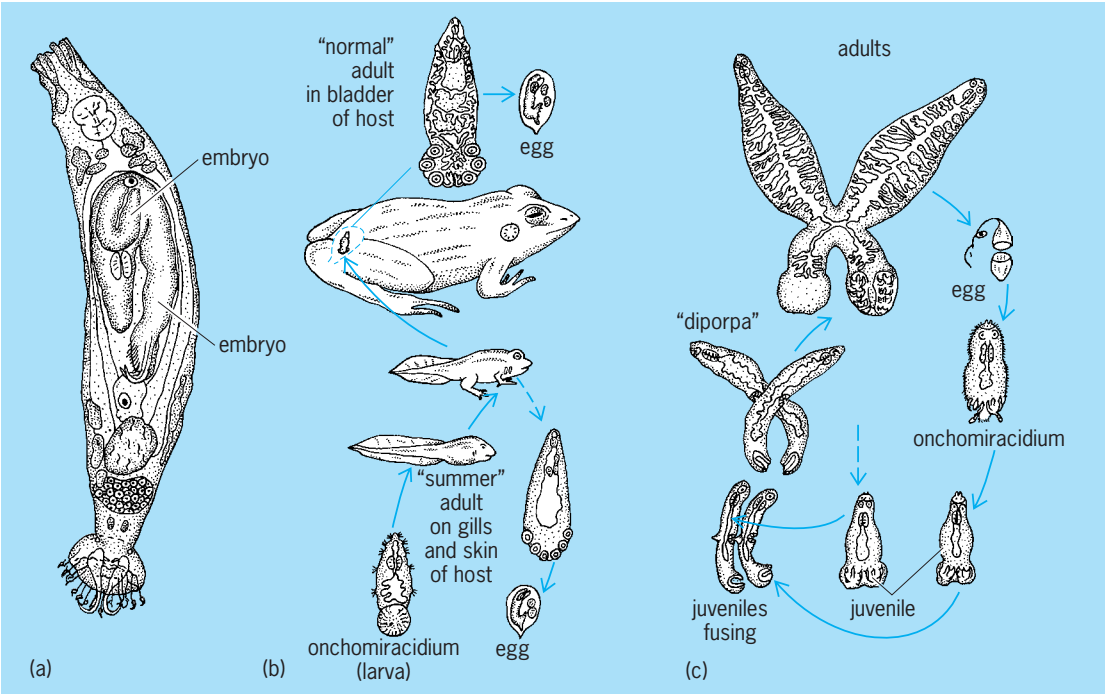

Fig. 2. Three unusual monogeneid life cycles. (*a*) *Gyrodactylus funduli* from the longnose killfish (*Fundulus similis*), dorsal view. In the uterus of the viviparous parent is an embryo within an embryo, thought by some to be the result of polyembryony. (*b*) *Polystoma integerrimum* from the frog bladder. The larva on tadpole gills grows as the tadpole develops. If, when the tadpole metamorphoses, the juvenile fluke does not enter the esophagus of the host, whence it can travel to the bladder, it matures neotenically into a "summer" adult and reproduces the first year. If it does reach the frog's bladder, the fluke matures normally and reproduces in the third year. (*c*) *Diplozoon* sp. The larvae undergo simple metamorphosis into nonciliated juveniles that attach to each other and become fused "diporpa larvae." Fusion of reproductive organs becomes complete as the flukes mature, and cross-fertilization occurs as the occasion demands.

environment is most common; however, interesting exceptions exist (**Fig. 2**).

Physiology. Monogeneids were at one time thought to feed on mucus, but there is good evidence that some are blood feeders. Most live in well-aerated microhabitats and probably require more oxygen than do entoparasites. The male and female organs of some species mature at different times, but the reproductive significance is uncertain.

Ecology. Monogeneids are capable of delicate ontogenetic responses to the microhabitat; for example, the direction of body asymmetry often depends upon which side of the host the worm inhabits. The direct life history seems to favor a higher degree of host specificity than in other trematodes. Often a parasite occurs on a single host species, and closely related fishes bear closely related parasites.

Though usually well accommodated by their hosts, monogeneids may cause intense discomfort and wasting and pave the way for invasion by bacteria, myceliated fungi, and protozoans. Little is known of their effects on healthy, natural populations. Heavy infections cause serious mortality in hatcheries, aquariums, and culture ponds, and sometimes occur in nature. *See* ASPIDOGASTREA; DIGENEA; PLATYHELMINTHES; TREMATODA. William J. Hargis

Bibliography. N. G. Sproston, A synopsis of the monogenetic trematodes, *Zool. Soc.* (*London*), 25(4):185–600, 1946.

Monogononta

A class of the phylum Rotifera which contains the majority of species in this invertebrate class. The organisms of this order are characterized by the presence of a single gonad in both males and females. There is a striking degree of sexual dimorphism, with the males being small and degenerate. The order is

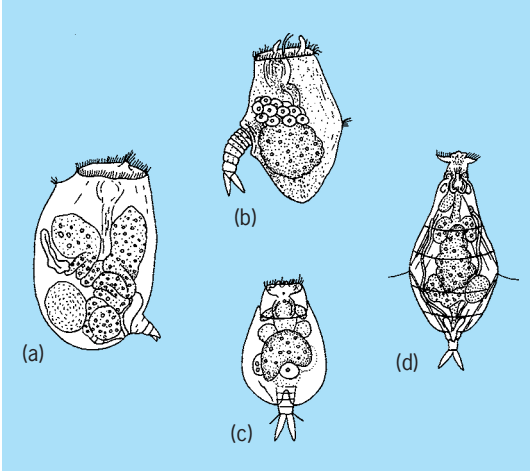

Fig. 1. Ploimates. (*a*) *Asplanchnopus* sp. (*b*) *Ploesoma* sp. (*c*) *Euchlanis dilata*. (*d*) *Notomata* copeus.

made up of three suborders: Ploima, Flosculariacea, and Collothecacea.

In suborder Ploima there is an exceptional diversity of form, varying from soft-bodied wormlike rotifers to species with variously ornamented, loricate shells (**Fig. 1**). Most of the free-swimming benthonic and pelagic rotifers belong to this suborder. Locomotion is by the ciliated corona. No fewer than five kinds of trophi occur: virgate (notommatids), cardate (*Lindia*), forcipate (dicranophorids), malleate (brachionids), and incudate (asplanchnids). Species range in size from 0.002 in. (0.05 mm) in *Colurella* to 0.06 in. (1.60 mm) in *Asplanchna*.

The suborder Flosculariacea contains the spectacular sessile rotifers formerly known as melicertaceans of the family Flosculariidae, as well as a number of equally notable free-swimming forms included in the family Testudinellidae (**Fig. 2**). Among the latter are the globular *Trochosphaera*, hailed as a connecting link between rotifers and annelids, and the appendage-bearing *Pedalia*, once considered as transitional to crustacea. All members of this suborder have a malleoramate mastax. The floscularids often have free-swimming young, and a few kinds are free-swimming as adults. Some of the sessile forms are encased in intricately constructed tubes. Some species occur in spherical or ellipsoidal colonies containing many individuals. Colonies may be over 0.08 in. (2.0 mm) in diameter, but individual floscularians are seldom over 0.04 in. (1.0 mm) in length.

The suborder Collothecacea contains but a single family, the Collothecidae, made up of five genera. Most species of Collothecidae are sessile, and many are encased in gelatinous tubes (**Fig. 3**). The largest genus, *Collotheca*, with an expanded funnel-shaped anterior end, has a wide range of coronal shapes among its 50 or more species. In its simplest form, the corona of *Collotheca* is without lobes, but in some species the corona has 5–7 lobes, which may be long and fingerlike. Three genera lack a corona. All members of this suborder have an uncinate mastax. *C. hoodi*, the largest rotifer known with the exception of Seisonacea, belongs to this group and reaches a length up to 0.1 in. (2.5 mm). *See* ROTIFERA.

Elbert H. Ahlstrom

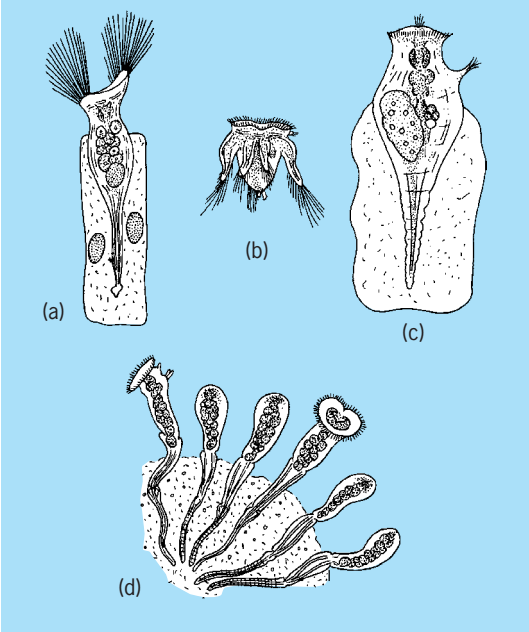

Fig. 2. Flosculariacea. (*a*) *Floscularia mutabilis*. (*b*) *Pedalia mira*. (*c*) *Conochiloides* sp. (*d*) *Lacinularia socialis*.

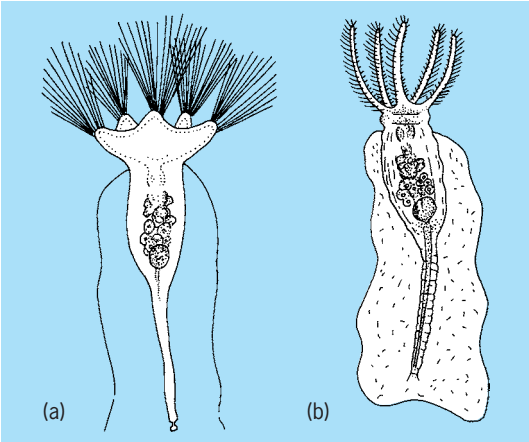

Fig. 3. Members of suborder Collothecacea. (*a*) *Collotheca* sp. (*b*) *Stephanoceros* sp.

Monomolecular film

A film one molecule thick. It is often referred to as a monolayer. Films that form at surfaces or interfaces are of special importance. Such films may reduce friction, wear, or corrosion, or may stabilize emulsions, foams, and solid dispersions. Thin films on water surfaces can reduce evaporation losses, though the spreading of thin films on such surfaces can represent a serious environmental problem. In some instances, the film spreads by itself and is essentially insoluble in the substrate (an insoluble monolayer). In other cases, the film molecules have a low substrate solubility and concentrate at the surface (a Gibbs monolayer).

In many cases, the film molecules are amphipathic, having a dual nature. On water surfaces, polar groups can anchor the film molecules, preventing evaporation, while the nonpolar sections prevent them from going into solution. On solid surfaces, polar interactions, or even chemisorption, may resist film removal and provide an essentially new surface. Based on the design of the film molecules, the newly created surface can be tailored to a number of different tasks. There is a wide range of applications of such films, but two areas stand out. On solid surfaces, the formation of multilayers from monolayers can lead to the creation of new materials, often having properties significantly different from that of typical three-dimensional (3D) solids. On water surfaces, the greatest number of studies relate to biological systems, with an emphasis on biological membranes. Many of the earlier studies were on long-chain fatty acids

Fig. 1. Film-balance apparatus. The barriers of the film balance are made from a combination of Teflon® and Delrin® (acetal resin) to prevent film and liquid leakage and are moved by a microstepping motor (on the right). The trough bottom is roughened to facilitate wetting and has thin Teflon rims to allow for a low volume subphase (55 milliliters with and 20 milliliters without the dipping well). The minature dipper can be used for deposition on small size substrates. The Willhelmy plate consists of a 0.5-mm-diameter metal alloy wire coupled to a miniaturized microbalance. Connecting cables to the A/D board behind are not shown. (*Courtesy of Kibron Inc.*)

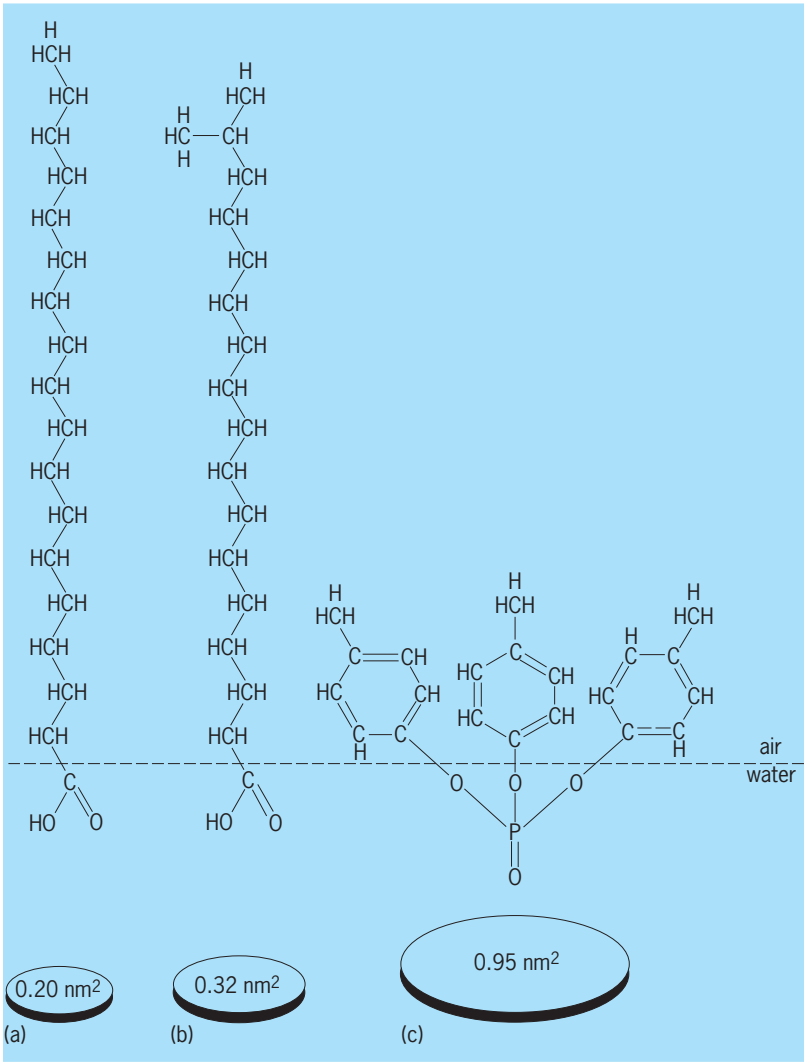

Fig. 2. Molecular orientation and cross-sectional areas of three representative polar organic molecules that are oriented at the water–air interface. (*a*) Stearic acid. (*b*) Isostearic acid. (*c*) Tri-*p*-cresyl phosphate.

and related substances at the air–water interface. Using typical membrane lipids, such as long-chain phospholipids, cholesterol, and membrane proteins, the monolayer constitutes an approximation to one-half of a biological membrane, and has provided considerable information on membrane behavior at the molecular level. *See* CELL MEMBRANES; MEMBRANE MIMETIC CHEMISTRY; POLAR MOLECULE.

Experimental techniques. The film balance provides basic information on molecular geometry and orientation, including the location and strength of polar groups, and the forces of cohesion and adhesion. The basic measurements are those of the surface pressure (or lowering of the surface tension) and the area per film molecule at the surface. Typically, the equipment consists of a thermostated trough having a hydrophobic surface, typically polytetrafluoroethylene (**Fig. 1**). The trough is filled to just above the rim with purified water, and all trace contaminants are carefully removed from the surface. Surface pressure is measured by immersing a thin rectangular plate in the surface prior to spreading the monomolecular film. The plate (a Willhelmy plate) consists of surface-treated platinum or glass or even filter paper, and should be wetted by the aqueous substrate. An ideal system will maintain the depth of the plate constant using an automatic nulling balance. Film molecules are spread from a dilute solution of known concentration using a volatile solvent. The film is confined between hydrophobic surface barriers (one static and one mobile) so that the film can be compressed continuously while the surface pressure and area per molecule isotherm can be recorded using the appropriate software. *See* SURFACE TENSION.

Stearic acid (**Fig. 2***a*) is the classical compound in monolayer studies; it is the simplest structure representative of thousands of important film-forming compounds. The stearic acid molecule consists of a long, straight hydrocarbon chain and a polar group at one extremity. A very similar molecule, isostearic acid (Fig. 2*b*), and a very dissimilar molecule, tri-*p*-cresyl phosphate (Fig. 2*c*), are also used frequently in monolayer studies. The difference between isostearic and stearic acid is very slight—the displacement of the small methyl group (CH_3) at the end of the molecule opposite the polar group. Tri-*p*-cresyl phosphate is greatly different; it has a bulky three-ring hydrocarbon portion attached to a strongly polar phosphate group.

Monolayer studies include graphing pressure-area isotherms (**Fig. 3**). The surface pressure in millinewtons per meter (or dynes per centimeter) is plotted against the average area per molecule. Extrapolation or extension of the steepest part of an isotherm to zero pressure is often used as a measure of molecular area. The point at which the pressure falls or remains constant is called the collapse pressure. Compressibility of the monolayer may be calculated from the slope of the isotherm. Thickness of the monolayer, or the length of the vertically oriented molecules, may be estimated by assuming a density for the monolayer; the volume and area then yield the thickness.

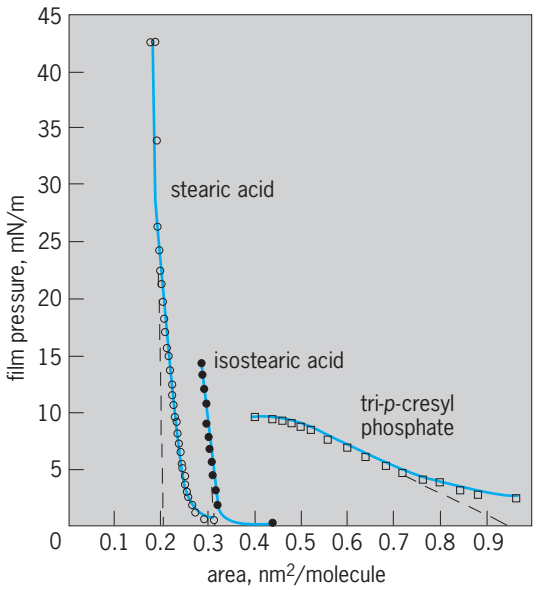

Fig. 3. Pressure-area isotherms for three compounds.

Comparison of the isotherms for stearic and isostearic acids (Fig. 3) demonstrates that the single, small side chain of isostearic acid has increased the cross section from 0.20 nm^2 for the stearic acid molecule to 0.32 nm^2 for isostearic acid, an increase of more than 50%. Collapse pressure falls from 42 mN/m to one-third of this value, 14 mN/m. These are striking differences between molecules that are extremely difficult to distinguish by most chemical methods.

The curve for tri-*p*-cresyl phosphate (Fig. 3) reflects a very different molecular structure. The extrapolated area, 0.95 mn^2/molecule, shows the bulkiness of the three-ring group held close to the surface. The low collapse pressure, 9 mN/m, reflects the weakness of such a film. The gradual slope of the curve, or the high compressibility of the film, indicates poor packing of the molecules.

In addition to the basic surface pressure per area per molecule isotherm, a variety of other properties can be measured simultaneously. Among these are the surface (Volta) potential, the surface viscosity, and the surface elasticity. More recently, absorption/emission spectroscopy, fluorescence microscopy (**Fig. 4**), Brewster angle microscopy (BAM), and grazing incidence diffraction (GIDX) using an intense synchrotron x-ray source have been used. Such techniques are particularly useful where films undergo phase transitions from one state to another. These states range from a two-dimentional (2D) gaseous state, through differing liquid states, to a 2D and a 3D solid state. Thus, both fluorescence microscopy and BAM are capable of distinguishing the different condensed states and providing detailed information on the molecular packing of the film. GIDX, though requiring an intense x-ray source, also provides information on molecular packing and orientation. Following transfer to a solid substrate, several other techniques become available. Most noteworthy are atomic force microscopy (AFM) and, for conducting films, scanning tunneling microscopy. Scanning electron microscopy (SEM) has also proved useful in examining multiphase films deposited on solid surfaces. *See* FLUORESCENCE MICROSCOPE; INTERFACE OF PHASES; SCANNING ELECTRON MICROSCOPE; SPECTROSCOPY; SURFACE AND INTERFACIAL CHEMISTRY; SURFACE PHYSICS.

Transfer of spread monolayers. Once a monolayer has been formed at the water–air interface, it is possible to transfer it quantitatively to another surface, such as that of a smooth solid. Dipping a clean glass slide into and out of the film-balance trough while the monolayer on it is held at constant surface

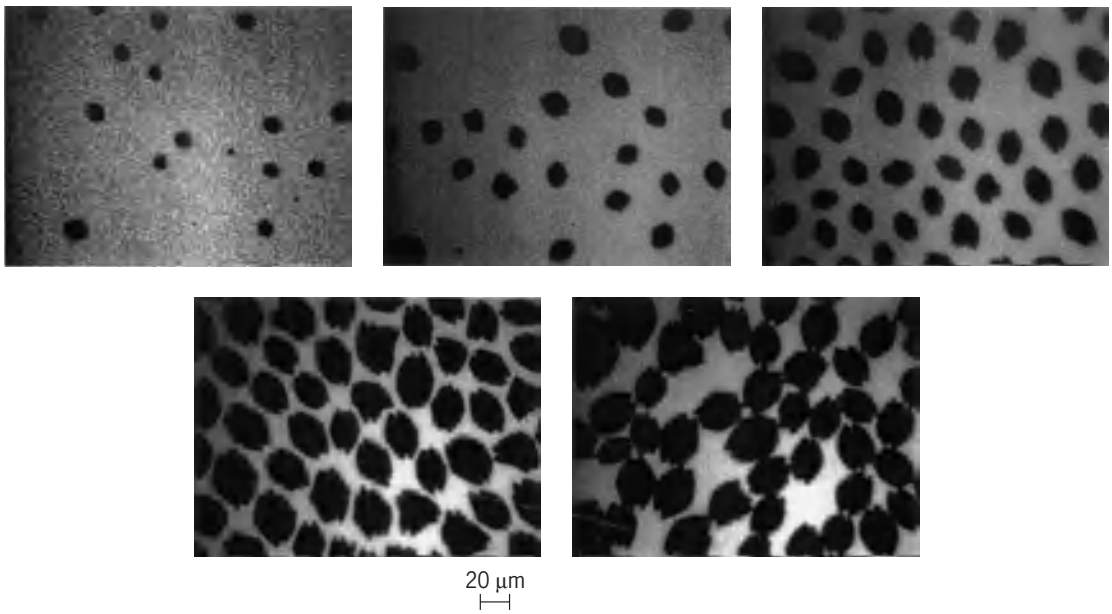

Fig. 4. Fluorescence micrographs of 6-hydroxyoctadecanoic acid (stearic acid) at 22°C during film compression (left to right). The dark regions are those of a liquid condensed phase over the lighter background of the liquid expanded phase. The fluorescence probe was at less than 1% concentration, and it concentrates in the liquid expanded region.

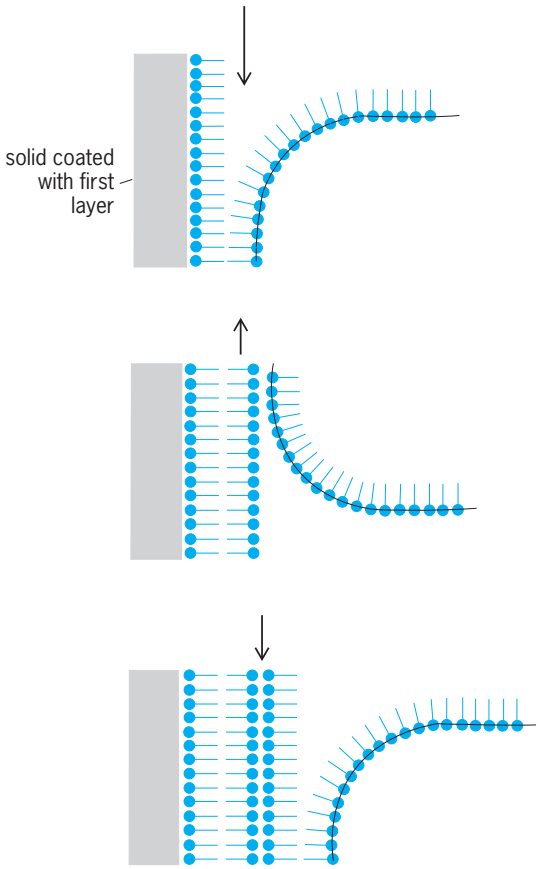

Fig. 5. Successive deposition of monolayers onto a solid plate as it is dipped into and out of the liquid in the Langmuir-Blodgett multilayer deposition method. (*After G. L. Gaines, Jr., Insoluble Monolayers at Liquid-Gas Interfaces, Wiley, 1966*)

pressure (by advancing the confining barrier) leads to such a transfer. The area of monolayer taken up is the same as the surface area of the dipped slide. Properties of the transferred layer, such as its wettability and electron diffraction pattern, indicate that the molecules on the solid surface retain the preferred orientation that they had on the water surface. Monolayers transferred to solids in this way bear close resemblance in structural details both to the precursor liquid-supported monolayer and to films formed on solids by other processes such as adsorption from solutions, but there are usually some differences. The differences, in structure, tightness of bonding, and so on depend on the nature of both the monolayer and the solid support. *See* ADSORPTION.

It is also possible to deposit certain types of monolayers (especially heavy-metal soaps of long-chain fatty acids) sequentially on solid surfaces to form built-up films or multilayers (**Fig. 5**). Since each monolayer is extremely thin but uniform (for barium stearate, for example, almost exactly 2.5 nm), such layer structures are very useful as spacers or thickness gages. Films more than 1000 layers thick have been made. Much interest has developed in the optical and electrical properties of such structures. It is possible to assemble multicomponent structures by using monolayers containing more than one chemical species and by changing from one kind of monolayer to another at different cycles in the dipping process. The location of different molecules at known small distances and in controlled relative orientation has permitted the study of such processes as energy transfer and electron transfers between them. While alterations of the packing may be small for the first few layers of the film, errors tend to increase as the number of layers also increases. It should also be realized that the transfer process depicted in Fig. 5 represents an idealized process, and not all film molecules fit the picture. This is particularly the case where complex molecules are involved. Thus, for polymer films the first layer will not have a perfect polar/nonpolar orientation, and subsequent layers will be less well oriented.

Applications. A large number of monolayer studies involve modification of the film molecules and interpretation of the resultant changes in the film behavior. Increasing the chain length of surfactants will lead to more condensed states. On the other hand, chain branching produces more expanded states, with the effect increasing as the branching entity moves toward the midpoint of the chain. The effects can be even more dramatic when the branching entity is polar, especially at low pressures.

Deuteration similarly leads to an expansion with the effect being roughly proportional to the degree of deuteration. It is clear that monomolecular films can exist in a number of physical states, depending on the nature of the film molecules and the surface pressure and temperature. This is illustrated in **Fig. 6** for L-α-dipalmitoylphosphatidylcholine (dipalmitoyl lecithin, DPPC), a typical biological membrane lipid. Isotherms are shown for the temperatures in the insert. At very high areas per molecule,

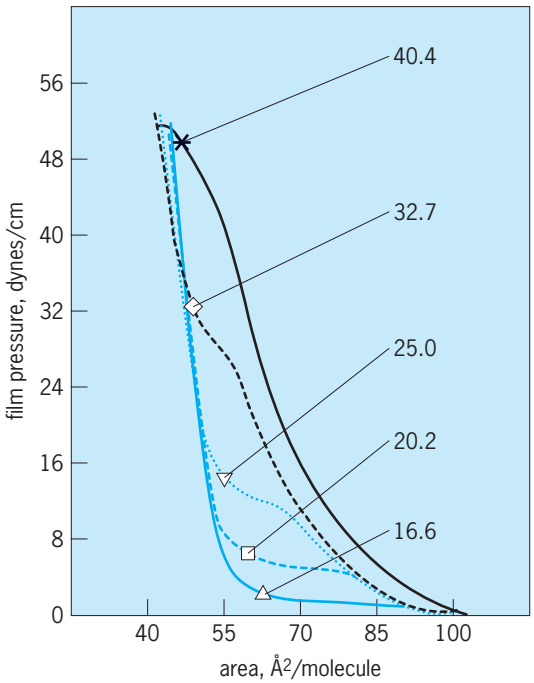

Fig. 6. Pressure-area isotherms for L-α-dipalmitoylphosphatidylcholine.

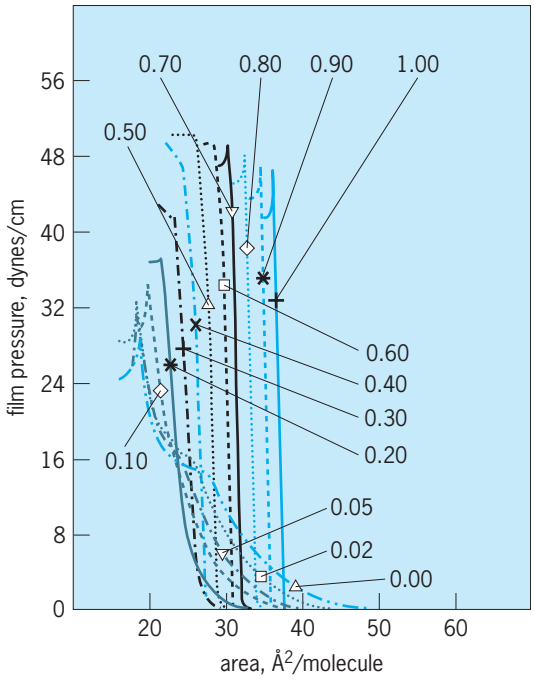

Fig. 7. Pressure-area isotherms for mixed films of myristic acid and cholesterol.

the film molecules are in a 2D gaseous state. While still at low pressures (<1 dyne/cm) and high areas (>100 $Å^2$/molecule), all films are in an equilibrium between a liquid-expanded gaseous state and a gaseous state (these states are not visible in Fig. 6). At some point, around 100 $Å^2$/molecule, all gaseous film has been converted to a liquid expanded (LE) state and the pressure begins to increase rapidly. At increasing pressures with increasing temperature, an inflection in the isotherm is found that denotes a first-order transition to a liquid condensed (LC) state, the inflection becoming less sharp as the temperature approaches 41°C (106°F). Above this temperature, no transition takes place and the film becomes unstable and collapses. The behavior shown is typical for most amphipathic molecules ranging from fatty acids to complex membrane lipids such as lecithins and phosphatidylethanolamines. However, not all surface-active molecules behave in this way, and one example is cholesterol. *See* SURFACTANT.

Mixed films. Typical mixed-film studies are usually limited to two components. **Figure 7** illustrates a series of mixed-film isotherms of myristic acid and cholesterol at 23°C (73°F), the insert numbers indicating the mole % cholesterol. The cholesterol isotherm on the extreme right shows only a fully condensed state at all pressures, with a collapse pressure around 46 dynes/cm. This collapse pressure varies with composition, reaching a maximum around 50 mole % cholesterol. In contrast, myristic (hexadecanoic) acid has a much lower collapse pressure and shows expanded phases at lower pressure similar to those in Fig. 6 for DPPC. With increasing cholesterol concentration, both the LE/LC phase transition and the collapse pressure vary. Variation of a phase transition, including collapse, with composition is indicative of at least partial miscibility of the components of the film. Cholesterol shows that it condenses the expanded myristic acid film but expands the condensed film. The ability of cholesterol to alter the fluidity of a monolayer is indicative of a similar role in biological membranes. At lower pressures, cholesterol is unaffected above about 20 mole %. For DPPC, this occurs at about 43 mole %. This suggests that cholesterol will segregate above these values in the respective mixed films. Mixed film studies can give information on both segregation and interaction of the two components.

Substrate interactions. In many instances, studies of film interactions have been made in which one component is present in the aqueous substrate. Fatty acids have been known to be encapsulated in β-cyclodextrin, and this has been illustrated using the effects of the β-cyclodextrin in the substrate on fatty-acid films. The influence of differing cations on fatty acids and on acidic phospholipids has had numerous studies. Some systems lend themselves to such studies; for example, the ability of lipases in the substrate to hydrolyze biological membrane lipids is an obvious case in point. Many of the hydrolyzed products are soluble and are lost from the monolayer, reducing the surface pressure and thus providing a way of following the reaction. Lipid–protein interactions may similarly be followed. Typical results indicate that in many cases the initial interaction is ionic and that it is often followed by insertion of a hydrophobic chain into the monolayer. Specific types of interactions may be studied by using polypeptides designed for the purpose. Other studies have included lipid–antibiotic interactions and the ability of local anesthetics to penetrate or fluidize monolayers as a model for the effects on membranes and a possible role in anesthesia. In such studies, it is assumed that the lipid monolayer model is sufficiently similar to the behavior of a bilayer and that it approximates a biological membrane.

Pulmonary surfactant. For some time, it has been known that pulmonary (lung) surfactant plays a vital role in reducing the surface tension in the lungs of mature humans. However, this surfactant is missing in the lungs of children born very prematurely and can cause serious lung problems or even death. The need for an artificial pulmonary surfactant has led to many studies of lung surfactant composition, and of mixed films of differing lipid components since these are thought to exist as a monolayer within the lung. It is now known that DPPC is fully capable of reducing the surface tension but has difficulty in respreading. The role of lung-surfactant-specific proteins has been investigated, and it is thought that these proteins play a role in the respreading of the surfactant.

D. Allan Cadenhead; George L. Gaines, Jr.

Self-assembled films. Self-assembled monolayers are monomolecular films formed by the immersion of an appropriate substrate into a solution of an active surfactant (**Fig. 8**). The molecular order in these 2D systems is produced by a spontaneous chemical synthesis at the interface as the system approaches

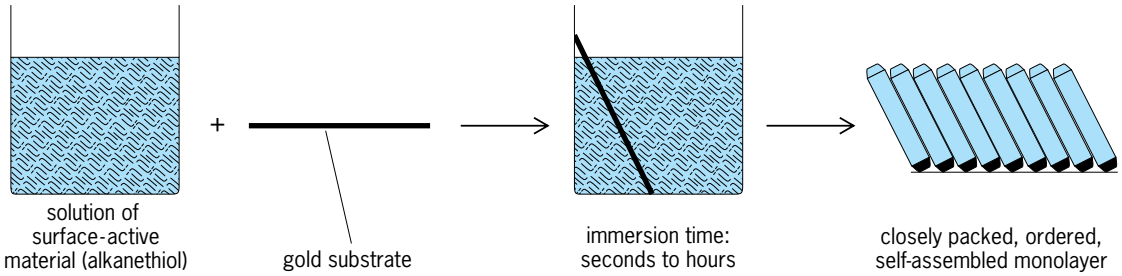

Fig. 8. Steps in forming a self-assembled monolayer.

equilibrium. In contrast to Langmuir-Blodgett monolayers, which require substrate pretreatment for significant bonding and are often metastable, self-assembled monolayers exhibit more uniform and stable behavior. The packing and ordering in self-assembled monolayers—both the conformation of the individual chains within the assembly and their packing and ordering with respect to each other—are determined by the interplay between interchain forces and the interactions with the surface, combined with entropic effects, and specific surface interactions of terminal groups.

There are two cases of monolayer formation. The first is on amorphous surfaces (such as glass), where intermolecular interactions define the packing and ordering; and the second is on crystalline surfaces (for example, the {111} face of gold), where the structure of the overlayer is determined by the surface. *See* CRYSTAL STRUCTURE.

Fatty acids. Long-chain carboxylic acids form monolayers on metal oxide surfaces such as aluminum, silver, and copper oxides via an acid–base reaction; on quartz and glass the mechanism may be ion exchange. They also form monolayers on oxide-free surfaces such as gold and copper in ultrahigh vacuum. The stability of the monolayers increases with increasing basicity of the substrate, in the order quartz $\sim$ gold $<$ glass $\sim$ aluminum $\sim$ copper $<$ silver. In the case of carboxylic acid derivatives with nonpolar functionalities, highly ordered monolayers can be formed. This is the case where vinyl, acetylene, diacetylene, pyrene, benzene, naphthalene, and bipheneyl are incorporated into the chain. Polar functionalities cannot be introduced because of their tendency to adsorb onto the surface, thus competing with the carboxylate group (COOH). *See* ACID AND BASE; ION EXCHANGE.

Very little is known about monolayer formation from the gas phase, and most studies have been carried out by using solution-adsorption. Monolayer formation is a function of both the substrate and the chain length. Thus, while hexanoic acid [$CH_3(CH_2)_5COOH$] forms ordered monolayers on silver oxide (AgO), for aluminum oxide (Al_2O_3) and copper(II) oxide (CuO) at least dodecanoic acid [$CH_3(CH_2)_{11}COOH$] is required. The chemisorption of alkanoic acids on amorphous metal oxide surfaces is not unique. On silver oxide, the two oxygen atoms of the carboxylate bind to the surface nearly symmetrically, and ordered monolayers are formed with a chain tilt angle from the surface normal of 15–25°. On surfaces of copper oxide and aluminum oxide, the carboxylate oxygen atoms bind asymmetrically to the surface, displaying tilt angles close to zero. This suggests that controlling the chemisorption mode determines molecular orientation, and hence physical properties that are orientation-dependent.

A much higher concentration of surfactant is required to produce good monolayers on glass than is required on aluminum oxide. The kinetics is faster for adsorption on glass than on aluminum oxide. It is possible to calculate the Gibbs free energy of adsorption for dodecanoic acid on aluminum oxide and on glass, revealing a stronger binding interaction of the carboxylate to the glass surface. *See* FREE ENERGY.

Alkyltrichlorosilanes. Self-assembled monolayers of alkyltrichlorosilanes (R—$SiCl_3$) require hydroxylated surfaces as substrates [glass, silicon dioxide/silicon (SiO_2/Si), aluminum oxide, and others], since the formation of the self-assembled monolayer is essentially an in-place formation of polysiloxane, which is connected to surface silanol groups (Si—OH) via silicon-oxygen-silicon (Si—O—Si) bonds. The Si···Si spacing [0.44 nm] is determined by the formation of a polysiloxane chain. This spacing leaves little free volume, and the chains are locked perpendicular to the surface. During adsorption, silicon-chlorine (Si—Cl) bonds react with the hydroxyl (OH) groups present on the surface of the substrate, and with trace water, which is adsorbed on the surface, to form a network of (Si—O—Si) bonds.

Self-assembled monolayers from octadecyltrichlorosilane ($C_{18}H_{37}SiC_{l3}$) are temperature-dependent. Complete surface reaction of the —SiC_{l3} groups occurs. The kinetics of the formation of long-chain alkyltrichlorosilane monolayers showed that monolayer formation is complete within $\sim$40 min. Monolayers of alkyltrichlorosilane show remarkable stability.

Construction of multilayers from alkyltrichlorosilane derivatives requires that the monolayer surface be modified to a hydroxylated one. Such surfaces can be prepared by a chemical reaction and the conversion of a nonpolar terminal group to a hydroxyl group. Once a subsequent monolayer is adsorbed on the activated monolayer, multilayer films may be built by repetition of this process. There is a linear relationship between the film thickness and the layer number. However, building thin organic films with interesting physical properties by using

self-assembly is not straightforward. Alkyltrichlorosilane derivatives of large aromatic systems are difficult to purify, and their solutions in organic solvents have to be replaced frequently because of polysiloxane formation. This is especially true for polar systems that are required for building waveguides for nonlinear optical application. Therefore, the unique stability of self-assembled monolayers made from silane derivatives makes them ideal materials for surface modification and functionalization applications, for example, as adhesion promoters, boundary lubricants, and active layers in chemical and biological sensors. *See* HYDROXYL; LUBRICANT; ORGANOSILICON COMPOUND.

Alkanethiols. Organic molecules that contain sulfur or selenium atoms in a neutral or negative oxidation state adsorb spontaneously on metallic surfaces such as gold, silver, platinum, copper, iron, and mercury. However, gold surfaces are the preferred substrate, because they are easily prepared, are simple to handle, and do not have a stable oxide. Examples are alkanethiols, di-*n*-alkyl sulfides, thiophenols, mercaptopyridines, mercaptoanilines, thiophenes, cysteines, alkaneselenoles, xanthates, thiocarbaminates, thiocarbamates, mercaptoimidazoles, and thiourea derivatives. The materials that were first recognized as forming closely packed, ordered self-assembled monolayers on gold substrates were the alkyl disulfides.

Alkanethiolates coordinate very strongly to gold. In many cases, the structure formed by alkanethiolates on the {111} face of gold is commensurate with the underlying gold lattice. *See* ARTIFICIALLY LAYERED STRUCTURES.

Alkanethiolates may have two binding modes at the hollow site of the {111} face of gold. Molecular mechanics energy minimization indicates that these two modes lead to monolayers exhibiting different types of packing arrangements but comparable in their ground-state energies. Therefore, monolayers may consist of two different chemisorption modes ordered in different domains, that is, simultaneously coexisting homogeneous clusters, each characterized by a different conformer in its unit cell. Annealing may result either in a unique chemisorption mode or in a much more complex structure, as the different modes are redistributed. *See* MOLECULAR MECHANICS.

Because of the relatively facile synthesis and purification of alkanethiols and the dense, stable structure of their self-assembled monolayers, these monolayers have potential applications in corrosion prevention, wear protection, and biosensing devices. The ability to tailor both head and tail groups of the constituent molecules makes them ideally suited for a more fundamental understanding of phenomena affected by competing intermolecular, molecule–substrate, and molecule–solvent interactions such as ordering and growth, wetting, adhesion, lubrication, and corrosion. *See* BIOSENSOR.

Mixed self-assembled monolayers of alkanethiolates on gold surfaces are especially suited for the studies of interfacial phenomena because of the fine control of the concentration of surface functional groups. This control is achieved by changing the concentration in the solution from which the mixed monolayers are adsorbed. Self-assembled monolayers are much better systems than polymer surfaces for studies of the effect of molecularly specific interfacial phenomena. The reason is that a polymer exposes a nonhomogeneous, molecularly rough surface because of different degrees of polymerization on different sites, while a self-assembled monolayer exhibits a smooth homogeneous surface. A large number of ω-terminated alkanethiols provide practically endless possibilities for surface engineering. For example, two self-assembled monolayer surfaces can be prepared with the same critical surface tension but with different degrees of hydrogen-bonding ability, for example, a pure chlorine surface and a mixed one of methyl and hydroxyl. Such surfaces, while exhibiting the same contact angle with different liquids, may show differences when contact angles are studied as a function of temperature and relative humidity.

Self-assembled multilayers of diphosphonates. Phosphate salts of tetravalent metal ions are extremely insoluble. These materials have a layered structure in which three of the four phosphate oxygen atoms are bonded to the transition-metal ion, and the fourth phosphorus-hydroxyl (P—OH) group is perpendicular to the sheet at both sides. Thus, replacing the hydroxyl group by an organic moiety should result in a lamellar structure of alkanephosphonates, cross-linked by the transition-metal ions. *See* FILM (CHEMISTRY); MATERIALS SCIENCE AND ENGINEERING.

Abraham Ulman

Bibliography. A. W. Adamson and A. P. Gast, *Physical Chemistry of Surfaces*, 6th ed., 1997; H. Bubert and H. Jenett (eds.), *Surface and Thin Film Analysis: A Compendium of Principles, Instrumentation, and Applications*, 2002; H.-J. Butt, K. Graf, and M. Kappl, *Physics and Chemistry of Interfaces*, 2003; G. L. Gaines, Jr., *Insoluble Monolayers at Liquid-Gas Interfaces*, 1966; R. Miller and D. Mobius (eds.), *Monolayers and Assemblies: Structure, Processes and Function*, 2002; G. A. Somorjai, *Introduction to Surface Chemistry and Catalysis*, 1994.

Mononchida

An order of nematodes having a full complement of cephalic sensilla on the lips in two circlets of 6 and 10. The amphids are small and cuplike, and are located just posterior to the lateral lips; the amphidial aperture is either slitlike or ellipsoidal. The stoma is globular and heavily cuticularized, and is derived primarily from the cheilostome. The stoma bears one or more massive teeth that may be opposed by denticles in either transverse or longitudinal rows. The esophagus is cylindrical conoid, with a heavily cuticularized luminal lining. The excretory system is atrophied. Males have ventromedial supplements and paired spicules. The gubernaculum may possess lateral accessory pieces. Females have one or two

ovaries. Caudal glands and a spinneret are common; however, they may be degenerate or absent.

There are three mononchid superfamilies, Mononchoidea, Bathyodontoidea, and Monochuloidea.

The superfamily Mononchoidea contains some of the most common and easily recognized free-living nonparasitic nematodes that occur in soils and fresh waters throughout the world.

The closely related Bathyodontoidea are inhabitants of soil or fresh water and prey on small microorganisms.

The nonparasitic Monochuloidea comprises both soil and fresh-water species, all of which are predators of microfauna. *See* NEMATA (NEMATODA).

Armand R. Maggenti

Monoplacophora

A class of the phylum Mollusca. Although fossil monoplacophorans had been known since the end of the nineteenth century, it was the discovery of a living species in deep water off Costa Rica in the 1950s that led to universal acceptance of the class. Monoplacophorans are bilaterally symmetrical, univalved mollusks that vanish from the fossil record at the end of the Paleozoic, about 240 million years ago. The living species, such as *Neopilina galatheae*, are rare and inhabit deep water, which may explain their absence from Mesozoic and Cenozoic rocks. Living monoplacophorans have a limpet-shaped shell, a circular foot attached by pairs of retractor muscles, and several gills on each side of the body (**Fig. 1**). An anterior mouth, but no distinct head, is evident, along with a radula for obtaining food, a central gut, and a posterior anus. Many of the organs are repeated as many as eight times down the length of the body, suggesting that the monoplacophorans are derived from some kind of segmented animal. An alternative view, which relates the monoplacophorans and other primitive mollusks to unsegmented flatworms (platyhelminths), has been challenged recently by both anatomical and molecular evidence.

Within the Mollusca, the Monoplacophora is regarded as the most primitive of the shelled, or conchiferan, classes. Viewed as a sister group of the eight-plated Polyplacophora (chitons), the Monoplacophora is seen as the ultimate source of the major molluscan classes, the Gastropoda, Bivalvia, and Cephalopoda. For this reason, much attention has been given to the fossil representatives of the Monoplacophora, and perhaps as a result

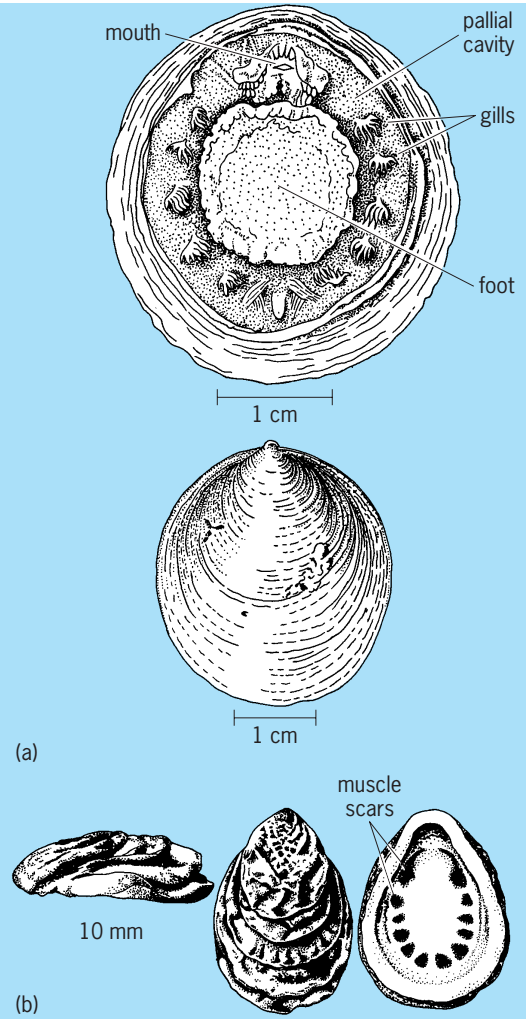

Fig. 1. Living and fossil Monoplacophora. (*a*) *Neopilina galatheae*. (*b*) *Tryblidium reticulatum*. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. I, 1957*)

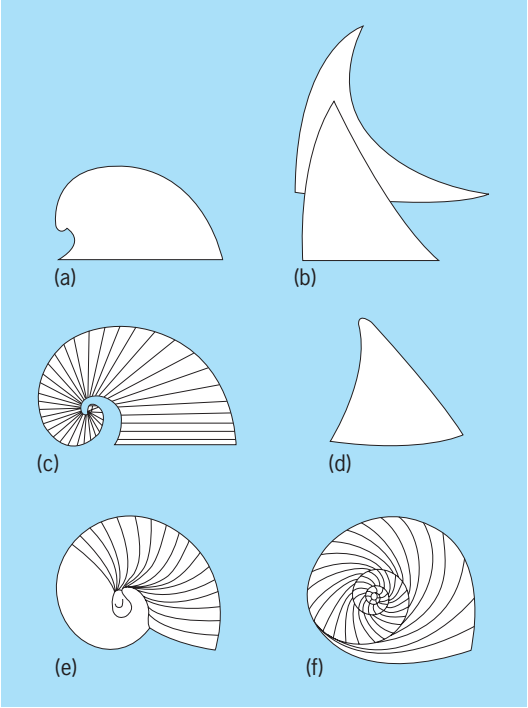

Fig. 2. Side views of various helcionellacean and bellerophont univalved mollusks. (*a*) *Cyrtonellopsis*; (*b*) *Hypseloconus*; (*c*) *Cyclocyrtonella*; (*d*) *Stenotheca*;; (*e*) *Bellerophon*; (*f*) *Tropididodiscus*. (*After B. Runnegar, Class Monoplacophora, in R. S. Boardman, A. H. Cheetham, and A. J. Rowell, eds., Fossil Invertebrates, Blackwell Scientific Publications, 1987*)

there is some controversy about which kinds of fossils should be placed within the class. *See* POLYPLACOPHORA.

All who have studied the group agree that the early Paleozoic cap-shaped shells such as *Tryblidium reticulatum* (Fig. 1) belong in the Monoplacophora. However, other bilaterally symmetrical, univalved mollusks from Paleozoic strata, including the early Paleozoic horn-shaped helcionellaceans and the planispirally coiled bellerophont mollusks, may also be monoplacophorans (**Fig. 2**). In fact, the bellerophonts may have been the most successful group of the Monoplacophora before they became extinct in the Early Triassic.

If bellerophonts were monoplacophorans, they show that gastropodlike features evolved independently in the group. The tight coiling of the shell may have been an adaptation to avoid predation on soft substrates: by being able to withdraw into a coiled shell, univalved mollusks were able to leave the security of rock substrates that acted as a second "valve" for limpet-shaped shells. Thus, many of the characters that are used to relate the bellerophonts to either the Monoplacophora or the Gastropoda are of uncertain phylogenetic significance, and the status of many extinct molluscan univalves remains unresolved.

Some early Paleozoic (Cambrian) monoplacophorans were laterally compressed, a feature that presumably facilitated slicing into soft substrates. Those forms gave rise to the Bivalvia by way of the pseudobivalved Rostroconchia. Other early monoplacophorans, ancestors of the Cephalopoda, had tall, partitioned, conical shells that were preadapted for flotation. A third Cambrian monoplacophoran developed an asymmetric shell and eventually gave rise to the Gastropoda. Thus, all of the major classes of the Mollusca were established within the Monoplacophora at or soon after the beginning of the Paleozoic. *See* GASTROPODA; MOLLUSCA. Bruce Runnegar

Bibliography. R. S. Boardman, A. H. Cheetham, and A. J. Rowell (eds.), *Fossil Invertebrates*, 1987; J. H. McLean, A new monoplacophoran limpet from the continental shelf off southern California, *Contrib. Sci. Nat. Hist. Mus. Los Angeles County*, 307:1–19, 1979; B. Runnegar and P. A. Jell, Australian middle Cambrian molluscs and their bearing on molluscan evolution, *Alcheringa*, 1:109–138, 1976; K. G. Wingstrand, On the anatomy and relationships of Recent Monoplacophora, *Galathea Rep.*, 16:7–94, 1985.

Monopulse radar

Radar capable of estimating target position based on the return from a single pulse. In many radars, precise angular position is estimated by conically scanning a single beam around the initial coarse angle estimate; the orderly amplitude variation of echoes during such scanning provides the refinement. Such measurement is limited, however, by pulse-to-pulse fluctuations in echo strength, a property quite common in radar targets.

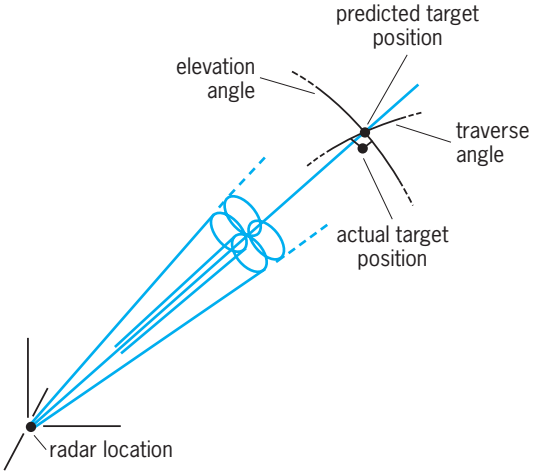

Fig. 1. Use of squinted beams in amplitude-comparison monopulse radar.

Monopulse radars use antennas that provide a local cluster of simultaneous beams (instead of scanning just one beam) to make the same precise angle estimate with each pulse transmitted. Since angle information is contained in each return, fluctuations in echo strength do not significantly degrade the measurement. Monopulse radars with mechanically positioned antennas address only a single target, and average the measurements over many pulses for improved accuracy. Radars using electronic beam steering in stationary phased-array antennas may make such a measurement in a single-pulse "dwell," doing so on dozens of targets, returning to each several times a second if necessary.

Target position may be defined by two angles (azimuth and elevation) and range, as in a radar-centered spherical coordinate system. Local angle deviations are often sensed in elevation and traverse (with subsequent coordinate conversion) depending on the characteristics of the antenna used. An antenna producing a cluster of slightly squinted beams (that is, beams which are not parallel but point in directions that make small angles with each other) [**Fig. 1**] permits angle measurement by comparing the

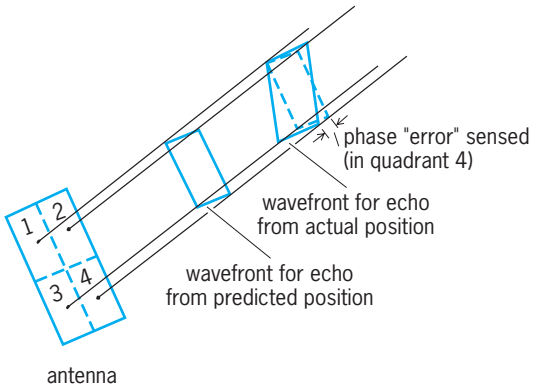

Fig. 2. Use of parallel beams in phase-comparison monopulse radar.

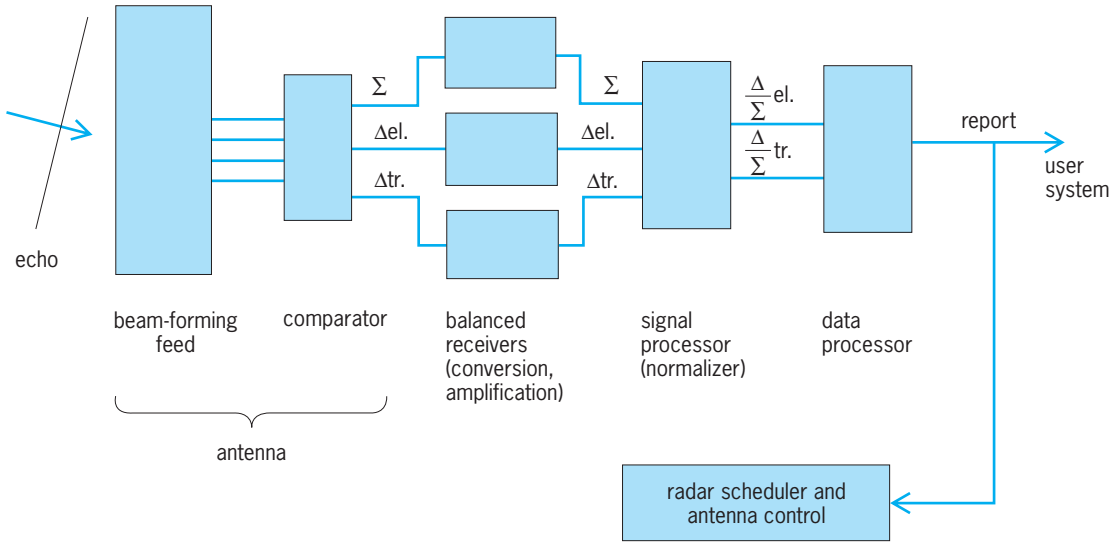

Fig. 3. Generic monopulse radar receiver. el. = elevation; tr. = traverse.

amplitudes of the echoes in the several beams; such is called amplitude-comparison monopulse. In phased-array antennas, it may be more convenient to form beams, one from each quarter of the antenna, all pointing in the same direction (not squinted, but parallel) and to use the relative phase of the echoes as the precise indication of angle. This approach (**Fig. 2**) is called phase-comparison monopulse.

In both approaches, the antenna feed includes a comparator, a device performing the appropriate comparison among the beams. Its outputs are the sum (Σ) of the echoes in the several beams and the differences (Δ) in elevation and traverse. The receiver (**Fig. 3**) processes these outputs to establish the quotient, Δ/Σ, indicative of the difference in each angle between the measurement and the prediction. This normalizing (division by Σ) makes the measurement nearly independent of target amplitude fluctuations.

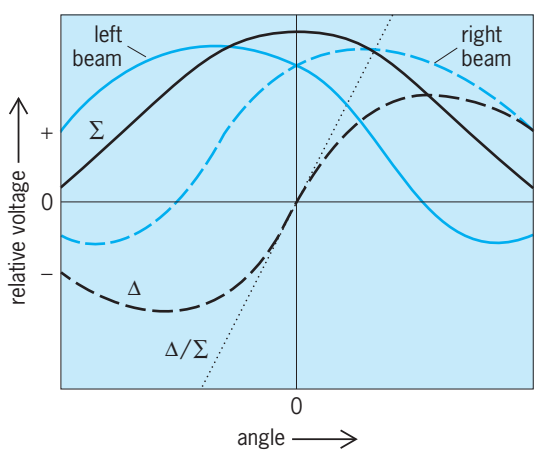

Fig. 4. Antenna patterns for amplitude-comparison monopulse radar in one measurement plane. Patterns of individual beams (left and right), sum (Σ) and difference (Δ) patterns, and the normalized calibration curve (Δ/Σ) are shown.

Antenna patterns formed at the comparator output ports comprise a sum pattern and two difference patterns, having, respectively, even and odd angle symmetry. **Figure 4** illustrates, for amplitude comparison in just one measurement plane, the individual squinted beams, the sum and difference patterns, and the normalized calibration curve, Δ/Σ. A good monopulse radar requires (as does any radar) good antenna patterns—a sum pattern with a narrow main beam, a difference pattern with a steep central slope (a deep null in the power pattern), and low sidelobes in each. The feed system must be well designed and may be somewhat complicated to achieve these qualities jointly.

Angle accuracy may be severely degraded by unresolved multiple targets (targets very close in angle and range) because the signals add together coherently in the separate receiver channels before the normalization. Receivers using only the real part of the complex quotient Δ/Σ can incur large errors in this situation.

A similar single-pulse technique using pairs of range gates to refine range estimates is sometimes called range monopulse. Most radar engineers think of monopulse, however, as the angle-measuring technique described here.

The emerging technique of computer-based digital beamforming in radar antennas will change the physical nature of the antenna considerably, but the basic principle—that of interpolating among multiple simultaneous beams, however formed, to estimate angle—will remain. *See* ANTENNA (ELECTROMAGNETISM); RADAR.

Robert T. Hill

Bibliography. D. K. Barton and H. R. Ward, *Handbook of Radar Measurement*, Artech House, 1984; S. Sherman, *Monopulse Principles and Techniques*, Artech House, 1984; M. I. Skolnik, *Introduction to Radar Systems*, 3d ed., McGraw-Hill, 2001; M. I. Skolnik (ed.), *Radar Handbook*, 2d ed., McGraw-Hill, 1990.

Monorail

A distinctive type of materials-handling machine that provides an overhead, normally horizontal, fixed path of travel in the form of a trackage system and individually propelled hand or powered trolleys which carry their loads suspended freely with an intermittent motion. Because monorails operate over fixed paths rather than over limited areas, they differ from overhead-traveling cranes, and they should not be confused with such overhead conveyors as cableways.

Relatively simple but adequately efficient monorail systems for specialized applications have flat steel bars or galvanized pipes for trackage. However, standard I beams or other similar shapes are used more and more frequently, with the wheels of the trolley bearing directly upon the lower flange of the beam. The latter type are used more for heavy-load service.

Garment manufacturers and cleaning establishments use the simple pipe-rail system for hanging garments on hangers or wheel-equipped trolleys. Switches, crossovers, and other components make setups flexible (see **illus.**). These systems can be arranged to run onto freight elevators, along loading platforms, and directly into carriers.

Of primary importance in a conventional monorail system are the rails. They are connected with butt or lap joints and must be smooth to allow the trolley to move freely. Clamps or brackets suspend the trackage from ceilings or walls. Because monorail systems do not operate with continuous motion, the tracks need not be arranged in self-closing lines. Spurs can be run into paint booths, cooling rooms, and similar work areas. These paths are selected by switches, such as the tongue or the slide (glider) varieties. Specially constructed sections permit 90° changes in trolley travel; turntables permit articles to be turned through 360°. Where monorails pass through fire doorways, liftout sections permit the door to close in the event of fire. Lift and drop sections shift the flow of traffic from one line to another at a different level, thus eliminating the need for inclined and declined tracks.

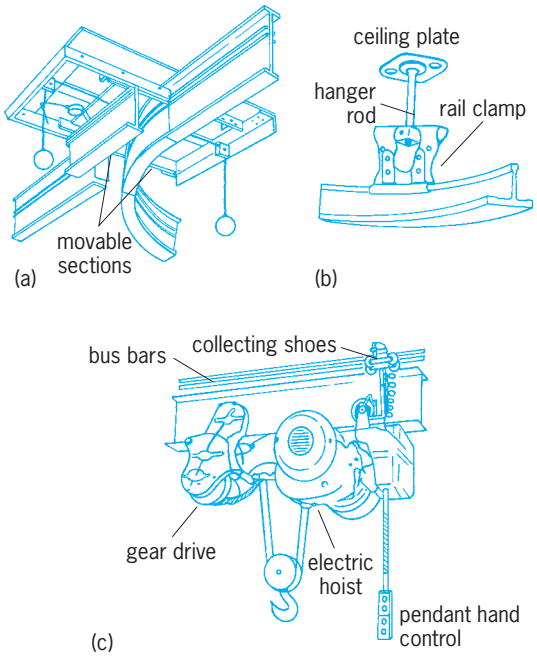

Typical parts of overhead monorails. (*a*) Slide (glider) type. (*b*) I-beam track. (*c*) Powered trolley.

Wheels mounted in trolleys ride on the flanges of the track, in most varieties. Two-wheeled trolleys, connected by bars, combine to make up carriers. The number of wheels depends upon the desired load-carrying capacity. In many instances hand or powered hoists are suspended from the carriers. To provide current for powered hoists, special electrification equipment, such as bus bars and collecting shoes, are used. If powered carriers are required, the drive may be secured in one of several ways. Examples are the motor-driven rubber tire which contacts the under flange of the track, or the trolley wheels which are driven through gear trains by an electric motor. Trolleys of the first type operate without slippage and make possible the introduction of sections with slight inclines and declines. Special carriers are used with monorail systems for handling specific products, such as batches in bakeries and the movement of clothes from one process to another in laundries. Scale sections can be included in a trackage line to weigh such products as batches, textile beams, or rolls of paper without causing a delay in traffic flow. Many below-the-hook devices can be used with monorails. *See* HOISTING MACHINES.

Monorail elements are utilized in the construction of overhead-traveling cranes if distribution over an area rather than along a fixed path is required. Such cranes can not accommodate the heavy loads handled by cranes built with structural girders, but they are adequate for many industrial applications. *See* BULK-HANDLING MACHINES; MATERIALS-HANDLING EQUIPMENT. Arthur M. Perrin

Bibliography. J. M. Apple, *Plant Layout and Materials Handling*, 1977, reprint 1991; F. E. Meyers, *Plant Layout and Material Handling*, 1993.

Monosaccharide

A class of simple sugars containing a chain of 3–10 carbon atoms in the molecule, known as polyhydroxy aldehydes (aldoses) or ketones (ketoses). They are very soluble in water, sparingly soluble in ethanol, and insoluble in ether. The number of monosaccharides known is approximately 70, of which about 20 occur in nature. The remainder are synthetic. The existence of such a large number of compounds is due to the presence of asymmetric carbon atoms in the molecules. Aldohexoses, for example, which include the important sugar glucose, contain no less than four asymmetric atoms, each

of which may be present in either D or L configuration. The number of stereoisomers rapidly increases with each additional asymmetric carbon atom. *See* CARBOHYDRATE; KETONE; OPTICAL ACTIVITY; STEREOCHEMISTRY.

A list of the best-known monosaccharides is given below:

Trioses:	$CH_2OH \cdot CHOH \cdot CHO$, glycerose (glyceric aldehyde)
	$CH_2OH \cdot CO \cdot CH_2OH$, dihydroxy acetone
Tetroses:	$CH_2OH \cdot (CHOH)_2 \cdot CHO$, erythrose
	$CH_2OH \cdot CHOH \cdot CO \cdot CHO$, erythrulose
Pentoses:	$CH_2OH \cdot (CHOH)_3 \cdot CHO$, xylose, arabinose, ribose
	$CH_2OH \cdot (CHOH)_2 \cdot CO \cdot CH_2OH$, xylulose, ribulose
Methyl pentoses (6-deoxyhexoses):	
	$CH_4(CHOH)_4 \cdot CHO$, rhamnose, fucose
Hexoses:	$CH_2OH \cdot (CHOH)_4 \cdot CHO$, glucose, mannose, galactose
	$CH_2OH \cdot (CHOH)_3 \cdot CO \cdot CHOH$, fructose, sorbose
Heptoses:	$CH_2OH \cdot (CHOH)_5 \cdot CHO$, glucoheptose, galamannoheptose
	$CH_2OH \cdot (CHOH)_4 \cdot CO \cdot CH_2OH$, sedoheptulose, mannoheptulose

Aldose monosaccharides having 8, 9, and 10 carbon atoms in their chains have been synthesized.

Reactions. As polyhydroxy aldehydes or ketones, the monosaccharides undergo numerous reactions.

Reduction and oxidation. On reduction, the aldoses take up two atoms of hydrogen and are converted to the corresponding sugar alcohols. A pentitol or pentahydric alcohol is obtained from a pentose, and a hexitol or hexahydric alcohol is obtained from a hexose. On oxidation, the monosaccharides yield carboxylic acids. Mild oxidation converts aldoses first into the corresponding monocarboxylic acids with the same number of carbon atoms; thus, aldopentoses are transformed into pentonic acids, $CH_2OH \cdot (CHOH)_3 \cdot COOH$, and aldohexoses into hexonic acids, $CH_2OH \cdot (CHOH)_4 \cdot COOH$. With stronger oxidizing agents, the process may proceed further, and hexoses, for example, may be oxidized to the corresponding isomeric, saccharic, or tetrahydroxyadipic acids, $COOH \cdot (CHOH)_4 \cdot COOH$. The ketoses, on oxidation, yield acids containing a smaller number of carbon atoms. The reducing properties of the monosaccharides are shown by their behavior with ammoniacal silver nitrate solution, from which metallic silver is precipitated, and particularly with Fehling's solution, from which, on warming, a brick-red precipitate of cuprous oxide is formed. This behavior is characteristic of aldoses as well as ketoses.

Phenylhydrazine. Emil Fischer's introduction in 1884 of phenylhydrazine, $C_6H_5NHNH_2$, as a reagent in sugar chemistry has proved of the greatest value in the separation and identification of the various monosaccharides. When 1 mole of phenylhydrazine

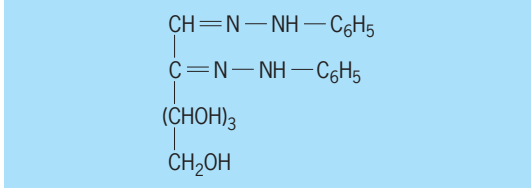

Fig. 1. Structural formula for hexosazone.

reacts with 1 mole of an aldose or ketose sugar, the first product is a hydrazone, as shown in reaction (1).

$$\underset{\text{Glucose}}{CH_2OH{-}(CHOH)_4{-}CHO} + H_2NNHC_6H_5 \longrightarrow \underset{\text{Glucose phenylhydrazone}}{CH_2OH{-}(CHOH)_4{-}CH{=}NHC_6H_5} + H_2O \quad (1)$$

On warming with excess of phenylhydrazine, the hydrazone first formed is oxidized in such a way that the CHOH group adjacent to the original aldehydic or ketonic group is converted into a CO group. The latter then combines with another mole of phenylhydrazine to give a dihydrazone containing group (2). These compounds are termed osazones (**Fig. 1**).

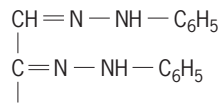

(2)

The osazones are colored compounds which are difficult to purify. For this reason, their melting points and specific rotations cannot be relied upon. However, since the osazones produced by various sugars possess characteristic crystalline forms, they are frequently used for cursory identification of the parent sugar by examining the crystals under the microscope. The osotriazoles (**Fig. 2**) obtained by oxidation of the osazones with copper sulfate are to be preferred, because they are colorless crystalline compounds and have more definite melting points and higher specific rotations than the osazones.

Osazones, like all hydrazones, are hydrolyzed when heated with hydrochloric acid, resulting in regeneration of phenylhydrazine. The original sugar, however, is not recovered, because in the process

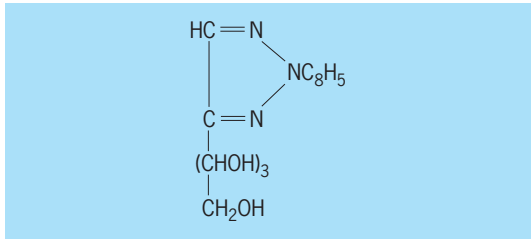

Fig. 2. Structural formula for hexosotriazole.

of regeneration of the sugar, group (3) is converted

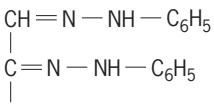

(3)

into the group —CO—CHO. The highly reactive compound so formed is an oxidation product of the original sugar and is termed osone. In the example quoted above, the osazone of glucose yields glucosone, $CH_2OH \cdot (CHOH)_2 \cdot CO \cdot CHO$.

The phenylhydrazine residues from sugar osazones may also be removed with a competing aldehyde, such as benaldehyde or acetaldehyde, resulting in osone formation. On mild reduction of this compound with zinc dust and dilute acetic acid, the aldehydic group alone is attacked and converted into an alcoholic group, the keto group remaining unchanged. In the case of glucosazone, the sugar finally obtained is fructose, $CH_2OH \cdot (CHOH)_3 \cdot CO \cdot CH_2OH$, in place of the glucose used as starting material. These reactions may be used as a general method of transforming an aldose into a ketose, according to scheme (4).

$$\text{Aldose} \xrightarrow{\text{phenylhydrazine}} \text{osazone} \xrightarrow{\text{hydrolysis}} \text{osone} \xrightarrow{\text{reduction}} \text{ketose} \quad (4)$$

Cyanohydrin synthesis. Monosaccharides, such as aldehydes and ketones, react with hydrogen cyanide to form cyanohydrins. By the use of this reaction, which is due to H. Kiliani, the synthesis of a higher from a lower aldose can be effected. The cyanohydrins are first hydrolyzed to hydroxy acids, which are readily converted into lactones. The latter are then reduced to aldoses by means of sodium amalgam. Thus glucose, under these conditions, results in a new seven-carbon sugar, glucoheptose. This is shown in reaction sequence (5).

Similarly, by using the glucoheptose and continuing with the process of cyanohydrin synthesis, an octose can be obtained. The synthesis has been carried as far as glucodecose.

Since cyanohydrin synthesis introduces a new asymmetric carbon atom, two products are obtained from a single monosaccharide. As an example, both L-gluconic and L-mannonic acids are produced from L-arabinose by this process.

Hydroxylamine. Monosaccharides react with hydroxylamine to yield oximes, the aldehydic or ketonic oxygen being replaced by the group =N—OH. Using these compounds. A. Wohl devised a method for effecting degradation of an aldose to one of lower carbon content. Thus, when the aldoxime of D-glucose, $CH_2OH—(CHOH)_4—CH{=}NOH$, is heated with acetic anhydride, it is converted into the acetyl derivative of the nitrile. $CH_2OAc \cdot (CHOAc)_4 \cdot CN$. On treatment with ammoniacal silver nitrate, this compound is deacetylated and the cyanide is released, leaving the corresponding aldopentose, D-arabinose, as the free sugar, as shown in reaction (6).

It is also possible to reduce the number of carbon atoms in an aldose carbon chain by O. Ruff's method, in which the calcium salt of an aldonic acid is oxidized with hydrogen peroxide in the presence of ferric acetate. Thus, from calcium arabinate, erythrose is obtained, as shown in reaction (7).

Epimerization. When two sugars or their derivatives, such as sugar acids, differ only in the configuration of the substituents on the carbon atom adjacent to the reducing group, they are called epimers. The aldonic acids are noteworthy for the ease with which they undergo epimerization or partial inversion of the asymmetry at carbon atom 2 upon heating with a weak base, such as pyridine or quinoline, to produce a mixture of the two epimers.

An aldose is first oxidized to the corresponding monocarboxylic acid, which is then heated with aqueous quinoline or pyridine, to yield a mixture of the epimeric aldohexonic acids. The latter may be separated by fractional crystallization of their lactones. Reduction of the lactone yields the corresponding aldose. This process may be illustrated by the transformation of D-glucose to D-mannose, as shown by reaction (8).

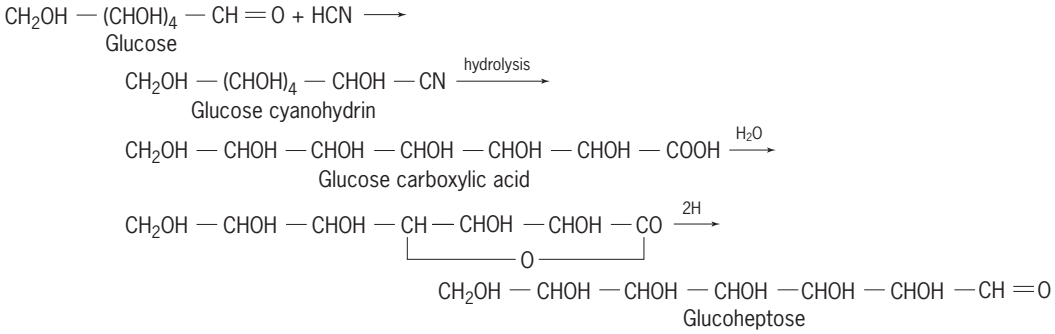

(5)

(6)

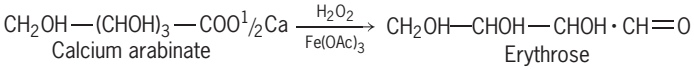

(7)

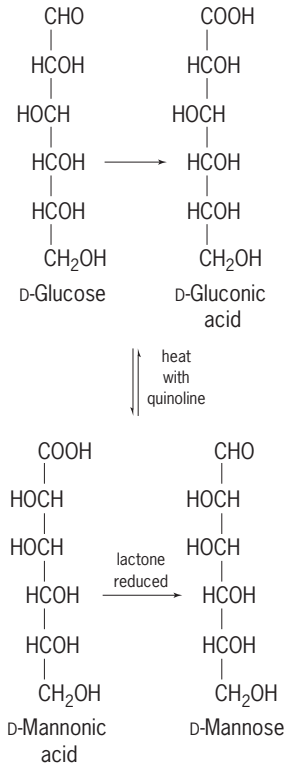

(8)

Enolization. Treatment of an aldose sugar with dilute alkali results in a mixture of an epimeric pair and 2-ketohexose. For example, when either D-glucose, D-mannose, or D-fructose is used, a mixture of these three sugars is obtained. The reaction, known as the Lobry de Bruyn-Ekenstein transformation, which is of general application, is due to the production of enolic forms in the presence of hydroxyl ions, followed by a rearrangement.

The transformation serves to show the close relationship between glucose, fructose, and mannose, as in notation (9). The structural representations of these three sugars are identical below the dotted line of the formula.

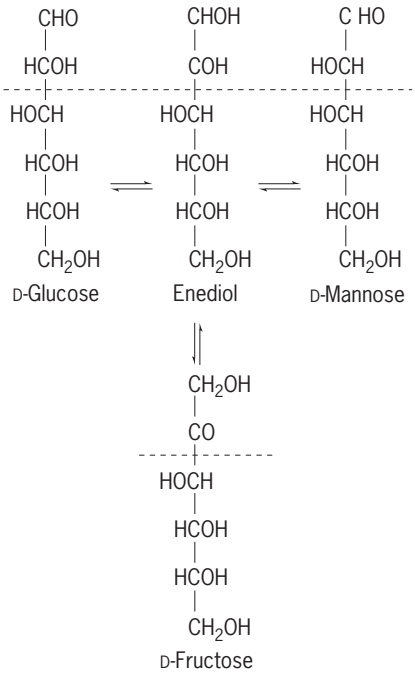

(9)

In a similar manner, the D-galactose series yields a mixture containing D-galactose, its epimer D-talose, and the ketose D-tagatose. When either D-xylose, D-lyxose, or D-xylulose is used, a mixture of the three sugars is obtained.

Derivatives. Derivatives of the monosaccharides are discussed in this section.

Sugar mercaptals. Reducing sugars, except ketoses, react with mercaptans in the presence of concentrated hydrochloric acid to form mercaptals. The reaction with glucose and ethyl mercaptan is given as an illustration in reaction (10).

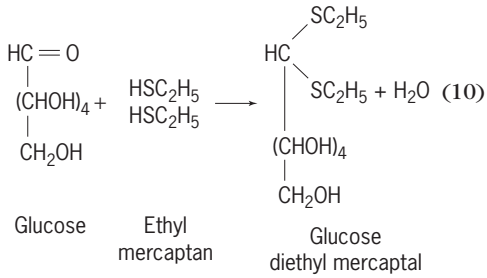

(10)

The mercaptals of the sugars are well-defined crystalline compounds. They are of special interest because they are open-chain compounds and have been found useful for the preparation of other derivatives that have this structure.

Acetone sugars. The reducing sugars yield condensation products with aldehydes and ketones. Those with acetone have played an important role in solving problems of sugar structure. The acetone glucoses such as those shown in **Fig. 3** are obtained by treating glucose in acetone with a condensing reagent such as zinc chloride or sulfuric acid.

Acetylated sugars. Acetylation of all free hydroxyls resulting in the formation of an ester may be accomplished by heating the monosaccharide with acetic anhydride in the presence of a catalyst such as anhydrous sodium acetate or zinc chloride, or by treating the sugar with acetic anhydride in a pyridine solution. When the hydroxyl at carbon 1 of an aldose is acetylated, two isomers (α,β) may be produced (**Fig. 4**). The directive influence of the catalyst used in the reaction is important. Glucose acetylated with

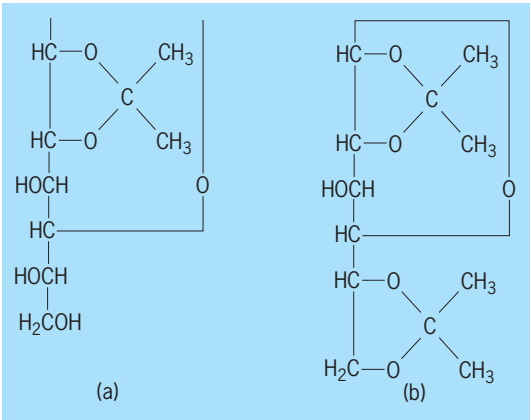

Fig. 3. Structural formulas for two acetone glucoses. Both are furanose derivatives. (*a*) 1,2-Monoacetone-D-glucofuranose. (*b*) 1,2-5,6-Diacetone-D-glucofuranose.

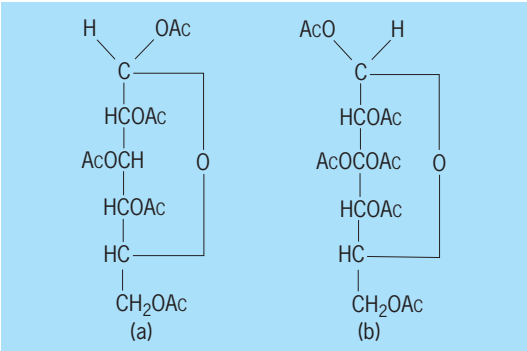

Fig. 4. Structural formulas for two isomers produced from acetylated D-glucose. (*a*) α-D-Glucose pentaacetate. (*b*) β-D-Glucose pentaacetate.

acetic anhydride in the presence of zinc chloride gives the α-pentaacetate, whereas with a sodium acetate catalyst the β form is produced. Furthermore, the β form, on heating with zinc chloride, is converted into the α isomer.

The free sugar may be recovered by deacetylation with dilute sodium hydroxide, or by catalytic deacetylation with barium or sodium methoxide.

Methylated sugars. Treatment of sugars with methyl iodide and silver oxide or with dimethyl sulfate and sodium hydroxide results in the formation of methylated derivatives (ethers) according to reactions (11).

$$\mathrm{ROH} + \mathrm{ICH_3} \rightleftharpoons \mathrm{ROCH_3} + \mathrm{HI(Ag_2O)} \qquad (11a)$$

$$\mathrm{ROH} + \mathrm{(CH_3)_2SO_4} \xrightarrow{\mathrm{NaOH}} \mathrm{ROCH_3} + \mathrm{CH_3{-}HSO_4} \qquad (11b)$$

Reaction (11*a*) is reversible; thus, in order to drive it to completion, the acid end product is removed by the addition of silver oxide.

The methyl ethers are the most common derivatives of this type and can be exemplified by fully methylated D-glucose, in which the hydrogens of all five free hydroxyls, including that of the glucosidic hydroxyl, are substituted by CH_3 groups (**Fig. 5**).

The glucosidic methoxyl group of the acetal is easily hydrolyzed with acid; however, the other methoxyls, which are true ethers, are resistant to acid as well as to alkali hydrolysis. The stability of the methoxyl group in these reagents makes the methylated sugars extremely useful in structural investigations. It is possible by appropriate methods to methylate selectively the hydroxyl groups in a monosaccharide. These partially methylated derivatives are utilized as reference compounds in the investigation of the constitution of oligosaccharides and polysaccharides.

Sugar phosphates. The naturally occurring phosphorylated sugars (including D-fructose-1,6-diphosphate, D-fructose-6-phosphate, D-glucose-6-phosphate, α-D-glucose-1-phosphate, D-glyceraldehyde-3-phosphate, and others) are of great metabolic importance. They function as intermediates in the processes of glycolysis, fermentation, and photosynthesis, and in most oxidative biological processes. Pentose phosphate esters occur as constituents of nucleic acids and a variety of coenzymes. *See* FERMENTATION; PHOTOSYNTHESIS.

Structurally the phosphorylated sugars are esters. The mono- and diesters of sugars are strongly acidic substances usually isolated as barium, calcium, lead, sodium, cyclohexylammonium, or alkaloid salts. D-Glucose-6-phosphate may be prepared by treating the 1,2,3,4-tetra-*O*-acetyl-β-D-glucose with diphenyl chlorophosphonate. The phenyl groups are removed from the resulting tetra-*O*-acetyl-D-glucose-6-diphenylphosphate by catalytic hydrogenolysis employing a platinum catalyst, and the product deacetylated by acid hydrolysis. Other phosphorylated sugars may be similarly prepared from the proper acetylated derivatives.

Aldose-1-phosphates may be prepared by reacting the poly-*O*-acetylglycosyl bromide with trisilver phosphate or silver diphenylphosphate. The resulting phosphate triester is simultaneously deacetylated and hydrolyzed under controlled conditions to give the aldose-1-phosphate. By this procedure, α-D-glucose-1-phosphate, α-D-galactose-1-phosphate, α-D-mannose-1-phosphate, and α-D-xylose-1-phosphate have been prepared. The β anomers may be prepared by reacting the poly-*O*-acetylglycosyl bromide with silver dibenzylphosphate, followed by catalytic hydrogenation to remove the benzyl groups and saponification of the acetyl groups. They also can be synthesized by coupling the poly-*O*-acetylglycosyl bromide with mono silver phosphate as the phosphorylating agent.

Sugar alcohols. These are acyclic linear polyhydric alcohols. They may be considered sugars in which the aldehydic group of the first carbon atom is reduced to a primary alcohol group. They are classified according to the number of the hydroxyl groups in the molecule. Thus, erythritol with four hydroxyls is considered to be a tetritol; arabitol with five hydroxyls is a pentitol.

Sorbital (D-glucitol, sorbite) is one of the most widespread of all the naturally occurring sugar alcohols (**Fig. 6***a*). It is found in higher plants, especially in berries and also in algae (seaweeds). Mannitol (Fig. 6*b*), like sorbitol, is widespread among plants. However, unlike sorbitol, it is frequently found in exudates of plants.

As a group, the sugar alcohols are crystalline substances, having low specific rotations, and ranging in taste from faintly sweet to very sweet. Their

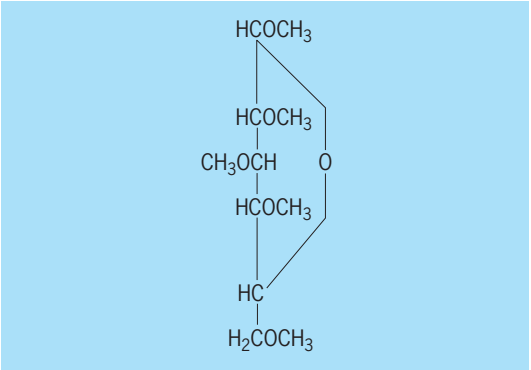

Fig. 5. Structural formula for methyltetra-*O*-methyl-α-D-glucose, a methylated ether.

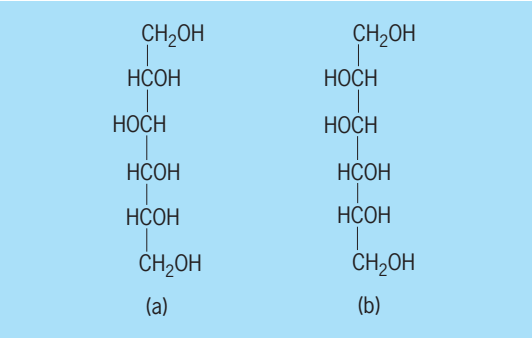

Fig. 6. Structural formulas for two common hexitols. (*a*) Sorbitol. (*b*) D-Mannitol.

distribution in nature is limited exclusively to plants. William Z. Hassid

Bibliography. H. S. Khadem, *Carbohydrate Chemistry: Monosaccharides and Their Oligomers*, 1988; W. W. Pigman, *The Carbohydrates*, 2d ed., vols. 1A, 1B, 2A, and 2B, 1970–1978.

Monosodium glutamate

The single sodium salt of glutamic acid used in foods to accentuate flavors. It is also known as MSG. The molecular structure is:

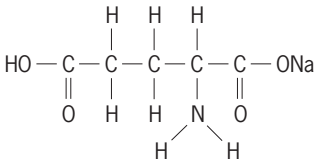

The crystal form available in commerce is the monohydrate, with structure as represented plus one molecule of water of hydration.

Glutamic acid is one of the more common of the amino acids. Its structural formula is the same as the monosodium salt excepting that the —COONa group on the right is replaced by a —COOH group, making the right end similar to the left. *See* AMINO ACIDS.

The formulas for glutamic acid and its salts show an asymmetric carbon atom. This is the fourth carbon atom from the left. It is attached to four entirely different groups. Therefore, the acid itself and each of its salts exist in three forms, the two isomers, L and D, and the racemic of D,L. The form is the so-called natural or active isomer, and its monosodium salt has the power of bringing out or emphasizing flavors, as distinguished from tastes, of certain foods, notably fish, fowl, meat, and vegetables. It is not a flavoring agent but, like salt, aids in developing the savor of foods. As a major constituent of all proteins, it participates in many of the metabolic processes.

Source. Originally produced from seaweed in the Orient, principal modern production is from cereal glutens, such as those of wheat, corn, and soybeans, from solutions evolved in the manufacture of beet sugar, and by microbiological fermentation of carbohydrates. To be commercially feasible as sources, the proteins from cereals must be concentrated and cheap. The two raw materials used for the greater proportion of commercial production are wheat gluten and desugared beet-sugar molasses. The world's largest single producer is a Japanese firm, principally using wheat gluten. In the United States, a number of factories exist; and several work solutions from beet-sugar molasses. The largest United States manufacturer uses a fermentation process and also what is commonly termed concentrated Steffen filtrate (CSF). This results from the multiple effect concentration at the sugar factories of the dilute waste liquor produced in recovering sugar from beet molasses by the Steffen process. Another factory uses liquor from both the Steffen and the barium (Deguide) process of sugar recovery.

Glutamic acid appears in the sugar beet as glutamine. During the sugar process, the glutamine changes to the internal anhydride of glutamic acid, pyrrolidone carboxylic acid. The latter is readily hydrolized by heating with either acid or alkali, and since it is not a protein, the glutamic acid which results is in the desired L form.

Although basically simple, all of the processes used are somewhat complex because of the many organic substances present in the raw materials. However, commercial production attains a high degree of purity of product, over 99.9% monosodium glutamate.

Glutamic acid produced by the microbiological method using a carbohydrate as raw material is manufactured by a United States pharmaceutical manufacturer and a Japanese firm.

Synthesis. A number of methods for synthesizing glutamic acid from several raw materials have been published in the scientific literature. Synthesis, however, invariably results in the racemized or D,L form. Methods for the resolution of this are known, but the few which have been made public are complex and costly. Commercially feasible methods of synthesis and resolution have been developed by two United States chemical manufacturers, one of which is a large producer of monosodium glutamate.

Uses. United States production includes glutamic acids, its hydrochloride, the mono-salts of sodium, potassium, ammonium, and calcium glutamates, all in the L forms. These find uses in medicine. By far the greatest proportion of glutamic acid production is used as raw material in making monosodium glutamate for the food industry. It is used both in the processing of foods and in the institutional and domestic fields. The Food and Drug Administration has approved it as an additive, and since in itself it is not a flavoring agent, its label designation in use is as a constituent and not as an artificial flavor. Originally, the latter was required because the product made then had a meatlike flavor due to impurities.

Monosodium glutamate is recognized as a standard of identity ingredient in several commercial food preparations. While it has been toxicologically cleared for use in food, there have been reports of incidences of allergic reactions. It is also contraindicated for people who must maintain low-sodium

diets; it is not permitted as an additive in baby food for infants under 12 weeks of age.

Its principal use is in the preparation of canned and dried soups, but it also enters into the production of some meat, vegetable, fowl, and fish products.

Monosodium glutamate has been and is of great importance in the Oriental diet. In the Orient, it has even been used as a medium of exchange and is often diluted with lactose or salt in order that it may be sold at a price within reach of the poor. *See* FOOD ENGINEERING. Paul D. V. Manning

Bibliography. H. B. Heath and G. A. Reineccius, *Flavor Chemistry and Technology,* 1986; I. D. Morton and A. J. McLeod (eds.), *Food Flavors*: pt.A: *An Introduction*, 1982; N. D. Pintauro, *Sweeteners and Enhancers*, 1977.

Monotremata

The single order of the mammalian subclass Prototheria. Two living families, the Tachyglossidae and the Ornithorhynchidae, make up this unusual order of quasi-mammals, or mammal-like reptiles.

The known species are highly specialized for their mode of life, and monotreme anatomy is therefore a curious mixture of reptilian and mammalian features along with highly diagnostic monotreme features. The fossil record of the group is virtually unknown, and Pleistocene forms can be classified within the living genera. Although ties with Jurassic multituberculates, docodonts, and other groups have been advocated, monotreme relationships are obscure. It is likely that they represent a divergent form of therapsid reptiles, independent of all other known groups of mammals, without fully attaining a truly mammalian grade of biological organization.

Morphology. The rostrum on the birdlike, toothless (in the adult) skull is covered with a leathery beak; lacrimals and auditory bullae are absent; cervical ribs are present; all vertebrae, except caudals, are without epiphyses; and large epipubic bones are characteristic. The scapula spine is restricted to the anterior ridge and the supraspinous fossa is absent; large precoracoids, coracoids, interclavicles, clavicles, and scapulae form the pectoral girdle. Monotremes are oviparous with telolecithal and meroblastic eggs. The poorly developed uteri are unfused. A cloaca is present, and the ureters open into a papilla in the urogenital sinus. The mammae are teatless; the penis is attached to the ventral wall of the cloaca; the prostate and the seminal vesicle are absent; the testes are abdominal; and there is one pair of bulbourethral glands. There is no corpus callosum, and the nonglandular stomach is lined with stratified epithelium. Endothermy is poorly developed.

Tachyglossidae. The echidnas (spiny anteaters) have relatively large brains with convoluted cerebral hemispheres. The known genera, *Tachyglossus* and *Zaglossus*, are terrestrial, feeding on termites, ants, and other insects. They are capable diggers, both to obtain food and to escape enemies. Like hedgehogs, they can erect their spines and withdraw their limbs when predators threaten. The long, sticky tongue can be extended 6 or 7 in. (15–18 cm) past the snout. Echidnas have well-developed senses of smell and hearing but their vision is poor. Commonly one egg, but occasionally two or even three, is laid directly into the marsupium (pouch) of the mother where it is incubated for up to 10 days. When the young hatch, they are poorly developed and remain in the pouch for about 2 weeks.

Species of *Tachyglossus* live in rocky areas, semideserts, open forests, and scrublands. They are found in Australia, Tasmania, New Guinea, and Salawati Island. Species of *Zaglossus* are found in mountainous, forested areas.

Ornithorhynchidae. The duck-billed platypus has a relatively small brain with smooth cerebral hemispheres. The young have calcified teeth, but in the adult these are replaced by horny plates which form around the teeth in the gums. The snout, as the name indicates, is duck-billed; the tongue is flattened out to work against the palate. The pelage is made of soft hairs, and the well-developed tail is flattened to aid in swimming.

The semiaquatic platypus is a capable swimmer, diver, and digger. When it is diving, its eyes and ears are closed by folds of fur. The bulk of its diet consists of aquatic vegetation, worms, aquatic insect larvae, crustaceans, and mollusks. Two eggs are usually laid by the female into a nest made of damp vegetation. After incubating the eggs for about 10 days the female leaves, and returns only when the eggs are hatched.

The platypus is found in Australia and Tasmania in almost all aquatic habitats, ranging from muddy streams and ponds to clear, rapid mountain streams. *See* MAMMALIA; PROTOTHERIA. Frederick S. Szalay

Bibliography. A. F. DeBlase and R. E. Martin, *A Manual of Mammology*, 2d ed., 1980; R. M. Nowak and J. L. Paradiso (eds.), *Walker's Mammals of the World*, 2 vols., 6th ed., 1999.

Monsoon meteorology

The study of the structure and behavior of the atmosphere in those areas of the world that have monsoon climates. In lay terminology, monsoon connotes the rains of the wet summer season that follows the dry winter. However, for mariners, the term monsoon has come to mean the seasonal wind reversals.

In true monsoon climates, both the wet summer season that follows the dry winter and the seasonal wind reversals should occur. Winds from cooler oceans blow toward heated continents in summer, bringing warm, unsettled, moisture-laden air and the season of rains, the summer monsoon. In winter, winds from the cold heartlands of the continents blow toward the oceans, bringing dry, cool, and sunny weather, the winter monsoon.

Based on these criteria, monsoon climates of the world include almost all of the Eastern Hemisphere tropics and subtropics, which is about 25% of the surface area of the Earth (**Fig. 1**). As shown in Fig. 1, the

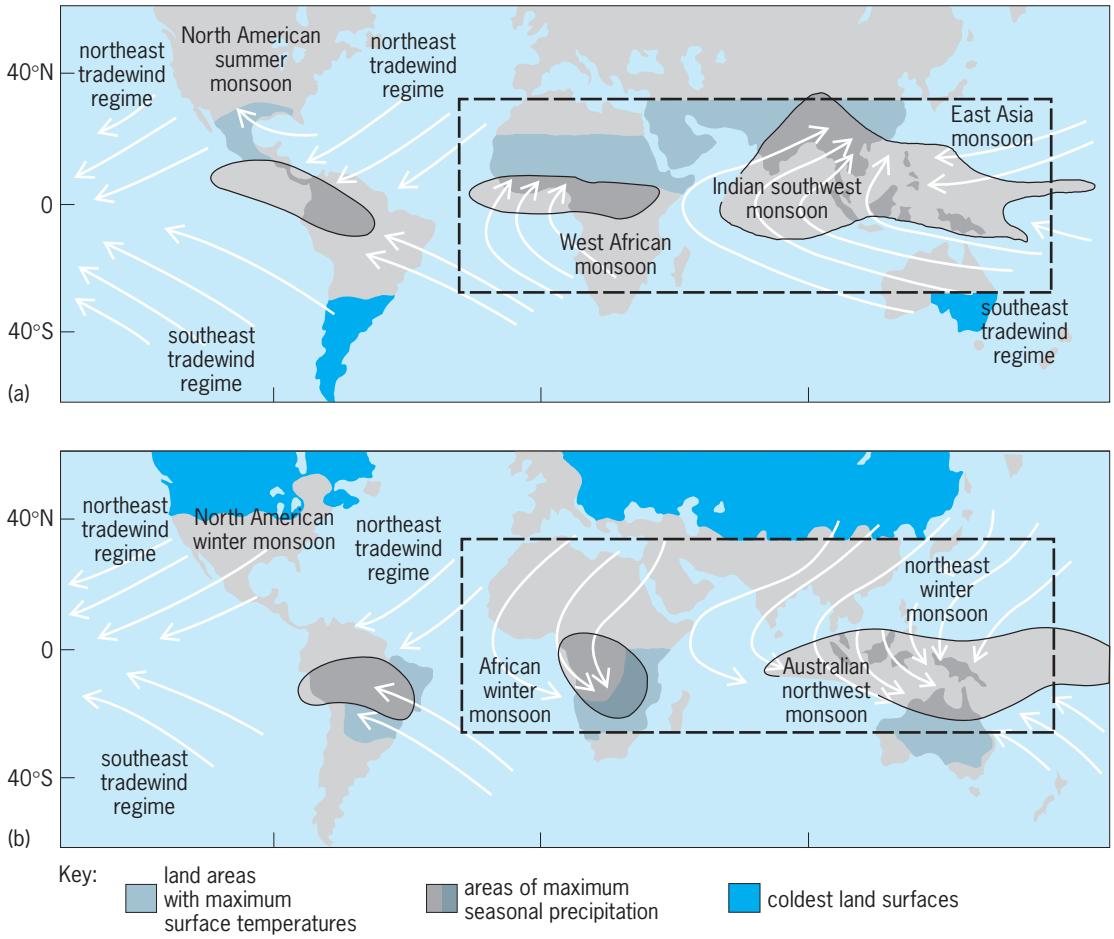

Fig. 1. Domains of the principal monsoon systems during the Northern Hemisphere (*a*) summer and (*b*) winter. Using the criteria of seasonal wind reversal and distinct wet summers and dry winters, the monsoon region is the region outlined by the broken rectangle. (*After P. J. Webster, The elementary monsoon, in J. S. Fein and P. L. Stephens, eds., Monsoons, Wiley, 1987*)

areas of maximum seasonal precipitation straddle or are adjacent to the Equator. Two of the world's areas of maximum precipitation (heavy rainfall) are within the domain of the monsoons: the central and south African region, and the larger south Asia-Australia region. The monsoon surface winds emanate from the cold continents of the winter hemisphere, cross the Equator, and flow toward and over the hot summer-hemisphere landmasses.

India presents the classic example of a monsoon climate region, with an annual cycle that brings southwesterly winds and heavy rains in summer (the Indian southwest monsoon; Fig. 1*a*) and northeasterly winds and dry weather in winter (the northeast winter monsoon; Fig. 1*b*).

Basic driving force. Like all weather systems on Earth, monsoons derive their primary source of energy from the Sun. About 30% of the Sun's energy that enters the top of the atmosphere is transmitted back to space by cloud and surface reflections. Little of the remainder is absorbed directly by the clear atmosphere; it is absorbed at the Earth's surface according to a seasonal cycle. The opposition of seasons in the Northern and Southern hemispheres leads to a slow movement of surface air across the Equator from winter hemisphere to summer hemisphere, forced by horizontal pressure gradients and vertical buoyancy forces resulting from differential seasonal heating. Such a seasonally reversing rhythm is most pronounced in the monsoon regions. *See* ALBEDO; ATMOSPHERE; CLIMATOLOGY; EARTH ROTATION AND ORBITAL MOTION; INSOLATION.

Land and water respond differently to equal amounts of solar energy, for two reasons. First, the specific heat of water is twice that of dry soil; for the same amount of heating, the temperature of dry land increases twice as much as that of water. Second, and more important, the effective heat capacity, that is, the ability of a material to store heat, is many orders of magnitude larger for the oceans than for the land. Being translucent, the ocean transmits radiational energy through its surface layer to some depth. In addition, unlike a solid, which transfers heat solely by molecular diffusion, the oceans transfer heat by turbulent mixing to depths of about 50–100 m (150–300 ft). In contrast, the depth to which dry land transfers heat is about 1 m (3 ft). This difference in distribution of incoming energy in summer results in large horizontal surface-temperature gradients between hot land and cooler oceans within

the monsoon regions. *See* HEAT CAPACITY; SPECIFIC HEAT.

In winter, surfaces on Earth lose more heat through terrestrial radiation to space than they receive from the Sun. Heat stored at depth during the previous summer is transferred to surface layers, where it replaces part of the heat that has been lost. Because the heat stored in the oceans is much greater than that in the land, ocean surfaces cool much less than land surfaces. Thus the oceans act as a heat buffer, like a giant, sluggish flywheel, storing large quantities of heat in summer and releasing it in winter, thereby suppressing extreme seasonal temperature variations such as those that occur over neighboring landmasses. *See* HEAT; HEAT BALANCE, TERRESTRIAL ATMOSPHERIC; TERRESTRIAL RADIATION.

Annual cycle. The Indian monsoon is both the classical and dominant monsoon system of the world. The other monsoon systems shown in Fig. 1 operate under the same physical principles with variations that depend on geography and scale.

Monsoon summer circulation. As summer approaches in a hemisphere, the energy received from the Sun at the Earth's surface becomes greater than its radiant energy loss to space. Land surfaces heat rapidly and intensely. This is particularly true of the empty quarter of Saudi Arabia, one of the great subtropical deserts, and of the Tibetan Plateau, which, at an average elevation of 4 km (2.4 mi), dominates the Asian continent. These heat sources loosely define the western and northern boundaries of the Indian summer monsoon (Fig. 1). By May, northern India, dry since the previous summer, is scorched by air temperatures higher than 100°F (about 40°C). In contrast, land surfaces cool in the opposite (winter) hemisphere, because they radiate more energy than they receive. The oceans experience the same trends as the land in each hemisphere, but to a greatly reduced extent. Also, the maximum and minimum land surface temperatures occur very close to the solstices, while those of the ocean surface waters lag by about 2 months. The net result is intensely heated land surfaces over southern Asia and northern Africa adjacent to a cooler Indian Ocean. *See* SOLSTICE.

As the Asian monsoon regions approach their maximum temperatures, the density of the heated air over land decreases, and a horizontal gradient of surface pressure between ocean and land develops. The pressure gradient and buoyancy forces caused by the heated air produce near-surface convergent motion, which causes moisture-laden air to move across the Equator and Indian Ocean inward and upward over the hot, low-pressure area of southern Asia. Completing the circulation, the rising air diverges aloft out of a large high-pressure area that develops by late May at jet-stream level (about 10 km or 6 mi) over the Tibetan Plateau and south-central Asia. This diverging air then slowly subsides to lower levels, replacing the air being drawn to the continents. *See* JET STREAM.

The actual paths of the winds are curved because of the Coriolis acceleration. Low-level winds spiral counterclockwise into the low-pressure centers over the Indian subcontinent (Fig. 1) and clockwise out of the upper-level high-pressure center over the Tibetan Plateau. These wind patterns are the classical cyclonic and anticyclonic wind circulations around lows and highs. *See* ATMOSPHERIC GENERAL CIRCULATION; CORIOLIS ACCELERATION; CYCLONE; WIND.

Monsoon rains. The rising air currents over the Indian subcontinent enter environments of reduced pressure aloft. The air thus expands and cools, and the moisture carried up from the surface layers in the form of water vapor condenses into clouds that ultimately produce the rains. The condensation process releases the latent heat stored in the water vapor molecules that were evaporated from the surface waters and carried aloft. This heat is a major source of added buoyancy for the monsoon circulation. The Ghats mountain range along the west coast of India and the massive Tibetan Plateau north of India provide orographic (mechanical) forcing, which further enhances condensation and precipitation processes.

The Asian summer monsoon rains last for about 100 days, starting in early to mid-June and ending in September. The date of the start of the rains over India—the onset—is fairly consistent from year to year, but it does vary infrequently by more than 30 days. The average onset date at Kerala, on the southern tip of India (8°N latitude), is June 1, with a standard deviation of about 1 week. The monsoon front advances slowly northwestward through the month of June; the average onset date is June 10 at Bombay (19°N) and June 15 at Delhi (28.5°N). By mid-July, the entire Indian subcontinent is under the influence of the summer monsoon. Water resources are so delicately balanced in India that distress situations occur if the onset is delayed by more than a week. Although onset dates have been observed to fluctuate by more than a month, studies of monsoon onsets over the west coast of India show the interesting and counterintuitive result that rainfall amounts, both for the month of June and for the total summer monsoon season, are not related to the date of onset. Rather, total monsoon seasonal rainfall amounts appear to be affected most by a phenomenon known as the break monsoon.

The phrase "break in the rains" refers to situations of reduced or no rainfall. During break conditions, reductions in rainfall over most of the Indian subcontinent are accompanied by enhanced rainfall over extreme northern India, in the foothills of the Himalayas. These shifts in rainfall patterns are associated with northward shifts of normal surface pressure and circulation patterns, but cause-and-effect relationships remain unknown. A study of 80 years (1888–1969) of monsoons showed that the range of the duration of break conditions was 3–21 days. Within the 80 years analyzed, there were 56 cases of break monsoons in July and 57 in August, with most lasting 3–5 days.

Figure 2 shows the seasonal (June–September) normal average rainfall for India and its standard

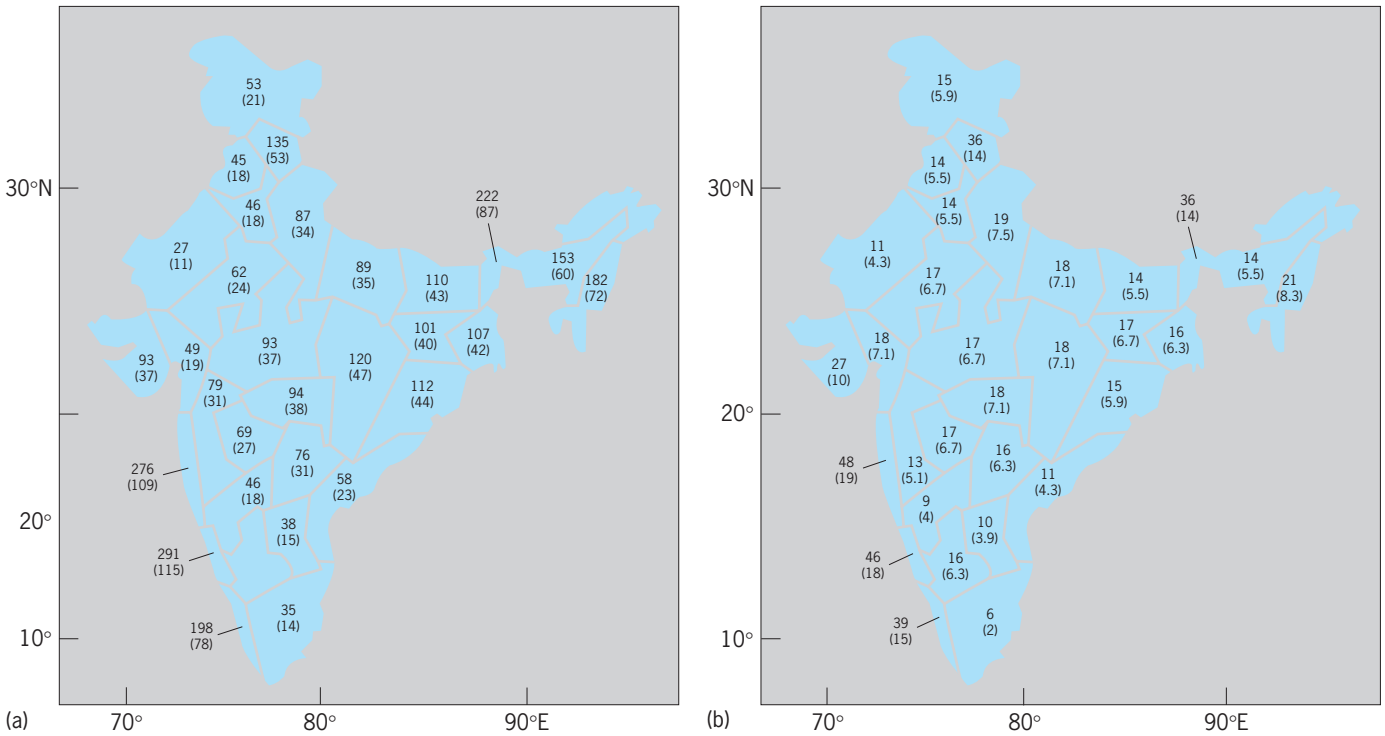

Fig. 2. Rainfall distribution in India, shown as number of centimeters (inches) in each subdivision (the value for one subdivision is not available). (*a*) Seasonal (June–September) mean rainfall calculated from 70 years (1901–1971) of data. (*b*) Standard deviation of seasonal mean rainfall calculated from 81 years (1901–1981) of data. (*After J. Shukla, Interannual variability of monsoons, in J. S. Fein and P. L. Stephens, eds., Monsoons, Wiley, 1987*)

deviation for the outlined subdivisions. Not shown are extreme local values, such as 800 cm (315 in.) at Cherrapunji (25.15°N, 91.44°E, elevation 1313 m or 4308 ft), which is the rainiest site measured on Earth, with more than 1000 cm (390 in.) of rain annually. The largest values of subdivisional mean rainfall occur over the west coast and northern regions of India, where orographic influences are important.

The total summer seasonal rainfall shown in Fig. 2*a* accounts for about 80% of the total annual rainfall for the Indian subcontinent. Of course, natural events are generally much more complex than models created on the basis of human perception. Few places in India are absolutely dry from October to May. Most areas receive up to a few centimeters of rain during these months. However, the state of Tamilnadu, in south India east of the coastal Ghats, and in a rain shadow region during the summer monsoon, receives most of its annual rainfall during October–December. In September, the summer monsoon rains withdraw southward across India. This retreat accompanies the equatorward movement of the regions of lowest surface pressure, the monsoon trough, which during the summer monsoon season is oriented east-west at about 25°N over north-central India. The monsoon trough reaches Tamilnadu in late October, concurrent with the start of its rainy season and, for that region, a wet winter monsoon. *See* RAIN SHADOW.

At this time, the rest of India begins its transition from wet to dry, from southwest to northeast winds, from summer to winter monsoon. By winter solstice, the subcontinent is uniformly cool and dry with northerly winds. These conditions prevail through February, a period of cool, dry, sunny weather over virtually all of India. The transition from March to early June brings hot weather, broken only occasionally by premonsoon thunderstorms. Finally, by early June the southwesterlies invade the southern coast of India and advance northward, once again bringing the rains of the wet season, the summer monsoon, to India. This annual cycle of wet and dry weather, profoundly affects the people of the Eastern Hemisphere. *See* ASIA; METEOROLOGY; PRECIPITATION (METEOROLOGY); TROPICAL METEOROLOGY.

Jay S. Fein

Bibliography. C. P. Chang (ed.), *East Asian Monsoon*, 2004; C. P. Chang and T. N. Krishnamurti (eds.), *Monsoon Meteorology*, 1987; J. S. Fein and P. L. Stephens (eds.), *Monsoons*, 1987; S. Hastenrath, *Climate Dynamics of the Tropics*, 1991; R. N. Keshavamurty, *Physics of Monsoons*, 1992; M. J. Lighthill and R. P Pearce (eds.), *Monsoon Dynamics*, 1981.

Monstrilloida

A copepod order of crustaceans containing 122 species in five genera, *Cymbasoma*, *Guanabaraenia*, *Monstrilla*, *Monstrillopsis*, and *Thaumatohessia*; Monstrillidae is the only family. Previously, the

five species of the family Thaumatopsyllidae had been included as a second family of Monstrilloida, but in 2003 this family was removed to its own order, Thaumatopsylloida, and unpublished observations of the development support this decision. *See* COPEPODA; CRUSTACEA.

Description. Adult females (1.5–2.5 mm or 0.07–0.11 in.) and males (1.0–1.5 mm or 0.04–0.07 in.) of the Monstrilloida are free-swimming and are often collected in plankton samples taken near the surface of shallow marine waters. The body has a podoplean architecture (the last two thoracic somites with the posterior part of the body), although the broader anterior part merges indistinctly with the narrower posterior part (see **illustration**). The second thoracic somite, bearing the first swimming leg, does not articulate with the first thoracic somite. Although antenna 1, the four swimming legs, and legs 5 and 6 are present, the feeding limbs, which include antenna 2, mandible, maxilla 1, maxilla 2, and the maxilliped, fail to form at this stage, and adults are assumed not to feed. Individual embryos often are attached to a long, spinelike attenuation of the female's genital somite. An embryo hatches as a feeble-swimming, simple nauplius (a crustacean larval stage) with well-developed antenna 1, antenna 2, and mandible. This nauplius penetrates the body of a polychaete or a mollusk; a hooklike structure on the mandibular endopod facilitates initial attachment. The next phase of the life cycle is passed within the blood system of the host as a sac-like form with several processes; this endoparasitic phase cannot be identified as a copepod. The monstrilloid leaves its host as an immature copepodid, apparently the fifth copepodid, and molts once to the adult. Multispecific aggregations of adults may occur at dusk, although the advantage to aggregating is not clear. This life cycle has been reported for only a few species and early in the last century; Alphonse Malaquin provides the most complete description. Monstrilloids appear to be found throughout the world's oceans, so it is surprising that their interesting life cycle has not attracted contemporary studies.

Taxonomy. Descriptions of new species continue to remain of interest to taxonomists. For example, 18 of the 122 species, almost 15%, have been described since 1999. This increase has been stimulated, in part, by recent revisions of several genera. Unfortunately, contemporary systematic analyses have not been applied to generic diagnoses, so the number of genera has remained static. The oldest genus, *Monstrilla*, was diagnosed by James Dwight Dana in 1849 and the youngest, *Guanabaraenia*, by Lejeune P. H. de Oliveira in 1945.

Phylogeny. The order Monstrilloida often is proposed as a close relative of the order Siphonostomatoida, despite the fact that the former are missing five limbs which are present on the latter. Closer inspection of this proposal, however, shows that this relationship is based on shared characters that have evolved many times among other copepods, and so do not provide the best evidence of common ancestry. Phylogenetic relationships of Monstrilloida with other copepod orders may remain problematical until an analysis of segment homologies of swimming legs and antenna 1 is undertaken.

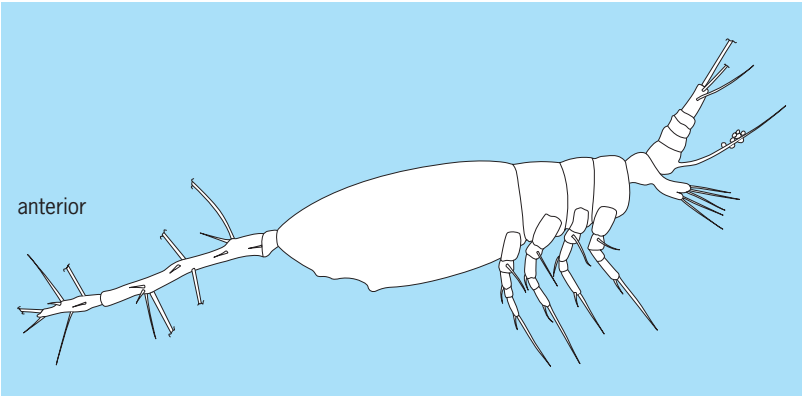

Side view of a monstrilloidan.

Frank D. Ferrari

Bibliography. M. J. Grygier, Annotated chronological bibliography of Monstrilloida (Crustacea: Copepoda), *Galaxea*, 12, 1995; M. J. Grygier, Nomenclature, redescription, and new record from Okinawa of *Cymbasoma morii* Sekiguchi, 1982 (Monstrilloida), *Hydrobiologia*, 292/293, 1994; M. J. Grygier and S. Ohtsuka, SEM observation of the nauplius of *Monstrilla hamatapex*, new species, from Japan and an example of upgraded descriptive standards for monstrilloid copepods, *J. Crust. Biol.*, 15, 1995; A. Malaquin, Le parasitisme évolutif des Monstrillides (Crustacés Copépodes), *Arch. Zool. Exp. Gén.*, Série 3, vol. 9, 1901; E. Suárez-Morales, An aggregation of monstrilloid copepods in a western Caribbean reef area: Ecological and conceptual implications, *Crustaceana*, vol. 74, 2001; E. Suárez-Morales and C. Diaz, A new species of *Monstrilla* (Crustacea: Copepoda: Monstrilloida) from Brazil with notes on *M. brevicornis* Isaac, *Proc. Biol. Soc. Wash.*, vol. 114, 2001.

Monte Carlo method

A technique for estimating the solution, x, of a numerical mathematical problem by means of an artificial sampling experiment. The estimate is usually given as the average value, in a sample, of some statistic whose mathematical expectation is equal to x. In many of the useful applications, the mathematical problem itself arises in a problem of probability in physics or other sciences, operational research, image analysis, general statistics, mathematical economics, or econometrics. The importance of the method arises primarily from the need to solve problems for which other methods are more expensive or impracticable, and from the increased importance of all numerical methods because of the development of the electronic digital computer.

The method as described above is identical with the earlier method known as artificial or model

sampling, or simulation. In fact, the term "simulation" is appropriate when the mathematical problem arises directly from a model of a real-world situation. The main justification for the name Monte Carlo is that since the mid-1940s several tricks have been introduced for improving the efficiency of the method, so that the subject has acquired a special flavor.

One of the earliest examples of artificial sampling was the experimental estimation of π by the French naturalist G. L. L. Buffon in 1773. The method is to throw a needle on a striped tablecloth and see how often it falls touching more than one stripe. If the width of each stripe is equal to the length of the needle, then the proportion of "successes" will be close to $2/\pi$ for a long series of trials.

Classification of Monte Carlo methods. The usual method of applied mathematics is to replace a physical problem P by a mathematical problem M, by assuming an adequate mathematical model; and then to solve the mathematical problem, and thus to solve P (perhaps only approximately). However, sometimes the mathematical problem, when replaced by a numerical problem, is solved with the aid of a calculating machine, so M is replaced by a new physical system, P'. One may think of the matter in this way, especially if the calculating machine is analog, that is, a machine using nondigital arithmetic. The method then becomes $P \to M \to P'$, where the arrow means "is replaced by." If $P = P'$, it can be said that P is solved by the crude method or direct experimental method, in which M is inessential. *See* ANALOG COMPUTER.

A special type of experiment with physical apparatus is a statistical experiment S, which makes use of dice, coins, needles, or Geiger counters. Sometimes random sampling numbers are used, that is, digits generated in such a manner that each selection of a digit has an independent chance of 0.1 (or 0.5) of giving each of the digits 0, 1, 2, . . . , 9 (or 0,1). More often, the digits are pseudorandom; that is, they are produced deterministically but have the appearance of being flat-random, like the digits in the decimal or binary expansion of $\sqrt{2}$. If a distinction is made between physical and other statistical experiments, there are, among others, the following methods of solving problems: $S \to S$, $S \to S'$, $S \to M$, $M \to S$, $S \to M \to S'$.

The method $S \to S$ involves estimating the solution of a real-life statistical problem by direct sampling. It may be regarded as a crude form of the Monte Carlo method. In $S \to S'$, a real-life statistical problem is replaced by a simpler model from which estimation is made by sampling; this method is perhaps best regarded as simulation. $S \to M$ is the ordinary form for the application of mathematical statistics. $M \to S$ is exemplified by Buffon's needle problem and was suggested for serious applications by Enrico Fermi, John von Neumann, Stanislaw Ulam, and Nicholas Metropolis in the 1940s. At that time it might have been reasonably called the true Monte Carlo method, but this name is no longer appropriate because of a change of emphasis. It is the method $S \to M \to S'$ which has been of most interest. This method can be of varying degrees of sophistication, depending on the amount of mathematical ingenuity required in order to transform S into S' via M.

Advantages and disadvantages. The main advantage of Monte Carlo is that other methods can be more costly or impracticable. A familiar example is the estimation of the probability of winning a game of pure chance: Sometimes the only reasonably simple method of estimation is to play the game several times. There are also numerical problems that can be solved by deterministic methods but can be more simply solved approximately by the Monte Carlo method. Sometimes poor approximations are satisfactory because the aim is merely to determine the strategic variables of a problem. This is likely to be a fruitful technique in mathematical economics. To simplify a complicated economic model without losing touch with reality, one could carry out an approximate Monte Carlo solution for the model, find out which of the variables are important, and then perhaps attempt a mathematical solution using only the important variables.

Another situation where a poor approximation is satisfactory occurs when there is available an iterative method of calculation, that is, a method of successive approximation, which converges closely to the right answer in a reasonable time provided that the first trial solution is not too far from the truth. The Monte Carlo method may then perhaps be used for obtaining a first trial solution. Modern Monte Carlo techniques are themselves usually iterative.

Sometimes the expense of a Monte Carlo method does not increase as fast as that of other methods when the dimensionality of a problem is increased. This seems to be true for multiple integration when it cannot be done analytically, and for the solution of Schrödinger's equation for several particles. *See* SCHRÖDINGER'S WAVE EQUATION.

The main disadvantage of some Monte Carlo methods is that for each extra decimal place required, it is necessary to multiply the sample size by 100. Thus, to calculate π to five decimal places by throwing a needle would require about 10^{10} throws, or 1 throw per second for about 300 years.

Applications. The method has been applied to the following problems, among others:

1. Size of cosmic-ray showers.
2. Critical size of nuclear reactors.
3. Other neutron transport problems, concerning, for example, the shielding properties of water or graphite. The probability that a neutron will cause a branching process that penetrates a shield may be as low as 10^{-10}. In a sample of random walks of reasonable size, there will be no "successes" unless some tricks are used.
4. Enumeration of high-polymer molecules or number of self-avoiding walks on a diamond lattice.
5. Percolation of a liquid through a solid.
6. Brownian motion and diffusion.

7. Birth-and-death branching stochastic processes.
8. Autoregressive time series.
9. Theory of queues, and other problems of commercial importance, such as storage, equipment replacement and maintenance, and insurance problems.
10. Laplace's partial differential equation.
11. Schrödinger's partial differential equation.
12. Integral equations.
13. Inversion of matrices.
14. Evaluation of definite integrals.
15. Random rounding-off.
16. Discovery of the t-distribution by Student.

These applications were all well studied before 1960. Since then the range of applications has expanded to include nearly every current numerical topic. *See* LATTICE (MATHEMATICS); QUEUEING THEORY.

Techniques. When a Monte Carlo method is performed on an electronic computer, the random sampling numbers may be replaced conveniently by pseudorandom sampling numbers. To do so has the advantage that every job can be checked precisely, even by other scientists if the "seed" is published. Having generated pseudorandom or random sampling digits, it is trivial to produce random variables that are uniformly distributed in an interval. There are then techniques for producing random variables having other probability distributions. Other techniques deal with the reduction of the variance of estimates. Example 3 above shows the need for such techniques. If x is a random variable with probability density function $p(x)$, the mathematical expectation of a function $f(x)$ may be estimated by replacing f and p by new functions, f^* and p^*, in such a way that $f^*(x)p^*(x) = f(x)p(x)$, where the integral of $p^*(x)$ is unity (so that p^* can be regarded as a probability point function of a new random variable), and where $p^*(x)$ is not small when x is in a region of importance. It is then adequate to estimate the expectation of f^* for the new random variable by averaging values of $f^*(x)$ where x is chosen with probability density $p^*(x)$. This technique, importance sampling, is useful where $f(x)$ is very small. Another technique, called acceptance or rejection sampling, also uses sampling from an adjunct distribution. The necessary adjustment is done by accepting each sampled observation only with an appropriate probability. Some other techniques are familiar in sampling survey work, such as the use of statistical regression and stratified sampling.

Markov-chain Monte Carlo (MCMC). This technique concerns a random vector X which has a known density function $\pi(\cdot)$. The goal is to estimate the expectation of $f(x)$. One way to do this is first to find a Markov chain that has the stationary distribution $\pi(\cdot)$. At first sight it is surprising that this can be done. Then, starting from a somewhat arbitrary initial vector X_0, successive elements (vectors) of the Markov chain can be generated, say $X_1, X_2, X_3, \ldots$. This sequence will usually tend to the stationary distribution $\pi(\cdot)$ from which the expectation of $f(X)$ can be estimated, say by averaging every t-th element in the sequence $X_1, X_2, X_3, \ldots$, where t is large enough to "forget the past." A method for constructing an appropriate Markov chain was provided by W. K. Hastings in 1970, a generalization of a method due to Metropolis in 1953. The applications of interest to Metropolis were in statistical mechanics, but since 1990 the generalization has had numerous applications in statistics, when multiple integrals occur. This often happens in bayesian statistics, including hierarchical Bayes. The Markov-chain Monte Carlo method can, however, take hours to converge, even on a modern computer. *See* BAYESIAN STATISTICS; STOCHASTIC PROCESS.

Gibbs sampling. This is a special but important case of Markov-chain Monte Carlo. It concerns a known density function f of several variables $(x_1, x_2, \ldots, x_n)$, some of which may be parameters. The goal is to estimate the expectations and variances of the marginal distributions of the n variables. If the conditional distribution of each variable is known, given the others, this goal can be achieved by the following iterative technique known as Gibbs sampling. Start from an arbitrary sequence $(x_2, x_3, \ldots, x_n)$ and sample x_1 from the conditional distribution of x_1, given the other $n - 1$ variables. Then do the same for x_2 given $x_1, x_3, \ldots, x_n$, after updating x_1. Then proceed to x_3 given $x_1, x_2, x_4, \ldots, x_n$, and so forth, going around cyclically for many iterations. Then average all the values obtained for x_1 and for x_2 and so on. This gives estimates for the marginal expectations of each of the n components. This description is deliberately oversimplified for the sake of brevity. Gibbs sampling resembles the EM algorithm, in which expectations and maximum likelihood are alternated iteratively.

Random digits. Monte Carlo methods depend on long sequences of random or pseudorandom digits. It is advisable to apply tests of the sequence for apparent flat-randomness. One useful test is the generalized serial test which tests the equiprobability of all m-plets, for any fixed value of m where m is not too large. *See* OPERATIONS RESEARCH; PROBABILITY; PROBABILITY (PHYSICS); STATISTICS.

Irving J. Good

Bibliography. G. S. Fishman, *Monte Carlo: Concepts, Algorithms, and Applications*, 2d ed., Springer-Verlag, 1996; J. E. Gentle, *Random Number Generation and Monte Carlo Methods*, Springer-Verlag, 1998; W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, CRC Press, 1996; J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods*, Chapman and Hall, 1965; D. E. Knuth, *The Art of Computer Programming*, vol. 2: *Seminumerical Algorithms*, 3d ed., Addison-Wesley, 1997; D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge University Press, 2000; M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics*, Oxford University Press, 1999; C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer-Verlag, 1999.

Month

Any of several units of time based on the revolution of the Moon around Earth.

The calendar month is one of the 12 arbitrary periods into which the calendar year is divided. *See* CALENDAR.

The synodic month, the period of the lunar phases, is the average period of revolution of the Moon with respect to the Sun, the same as the average interval between successive full moons. Its duration is 29.531 days. *See* PHASE (ASTRONOMY).

The tropical month is the period required for the mean longitude of the Moon to increase 360°, or 27.322 days.

The sidereal month, 7 s longer than the tropical month, is the average period of revolution of the Moon with respect to a fixed direction in space.

The anomalistic month, 27.555 days in duration, is the average interval between closest approaches of the Moon to Earth. The variation in the Moon's distance from the Earth causes a variation in the apparent size of the Moon and thus in the duration of solar eclipses.

The nodical month, 27.212 days in duration, is the average interval between successive northward passages of the Moon across the ecliptic, points known as nodes. Since eclipses can occur only when the Sun and Moon are near such nodes, this period is also known as a draconic month, after the Chinese mythical dragon that supposedly ate the Sun to cause a solar eclipse.

The fortuitous near-equality of 223 synodic months (6585.32 days), 242 nodical months (6585.36 days), 239 anomalistic months (6585.54 days), and 19 eclipse years gives a repetitive period of similar eclipses, called the saros, of 18 years, $11^1/_3$ days (the date varies +/−1 day, depending on leap years). *See* ECLIPSE; MOON; TIME; YEAR.

Gerald M. Clemence; Jay M. Pasachoff

Bibliography. J. Mitton, *Cambridge Dictionary of Astronomy*, Cambridge University Press, 2001; J. M. Pasachoff, *The Cosmos: Astronomy in the New Millennium*, 3d ed., Brooks/Cole, Belmont, CA, 2007; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, University Science Books, Mill Valley, CA, 1992.

Montmorillonite

A group name for all clay minerals with an expanding structure, except vermiculite, and also a specific mineral name for the high alumina end member of the group. *See* VERMICULITE.

Montmorillonite clays have wide commercial use. The high colloidal, plastic, and binding properties make them especially in demand for bonding molding sands and for oil-well drilling muds. They are also widely used to decolorize oils and as a source of petroleum cracking catalysts. *See* CLAY, COMMERCIAL.

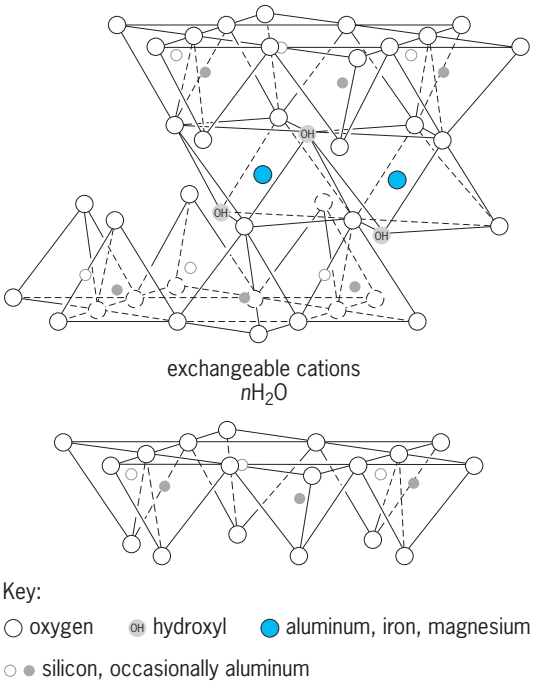

Diagrammatic sketch of structure of montmorillonite. (*After R. E. Grim, Clay Mineralogy, McGraw-Hill, 1953*)

Structure. Because of the extremely small particle size of the montmorillonite minerals, there is still some uncertainty regarding details of their structure. According to the structural concept which is currently accepted, montmorillonite is composed of units made up of two silica tetrahedral sheets with a central alumina octahedral sheet. The atoms in these layers which are common to both sheets become O instead of OH. Montmorillonite is thus referred to as a three-layer clay mineral with tetrahedral-octahedral-tetrahedral layers making up the structural unit (see **illus.**).

These silica-alumina-silica units are continuous in the *a* and *b* crystallographic directions and are stacked one above the other in the *c* direction. In the stacking of these units, the oxygen layers of neighboring units are adjacent. This causes a very weak bond and an excellent cleavage between the units. Water and other polar molecules can enter between the unit layers and cause an expansion of the structure in the *c* direction. Thus montmorillonite does not have a fixed *c*-axis dimension, but can vary considerably depending on the absence or presence of interlayer molecules. The *c*-axis spacing also varies with the nature of the interlayer cation present between the silicate layers. A montmorillonite in an air-dried condition with sodium as the exchange ion frequently has one molecular water layer and a *c*-axis spacing of about 1.25 nm. Under similar conditions there are two molecular water layers with calcium, giving a *c*-axis spacing of about 1.55 nm. The expansion properties of montmorillonite are reversible; however, reexpansion may be difficult after complete structural collapse by removal of all interlayer polar molecules.

Atomic substitution. The theoretical formula for montmorillonite without structural substitutions is $(OH)_4Si_8Al_4O_{20} \cdot nH_2O$ (interlayer). However, montmorillonite always differs from the above theoretical formula because of structural substitution. In the tetrahedral sheet, aluminum and possibly phosphorus substitute for silicon, whereas ions such as magnesium, iron, and lithium substitute for aluminum in octahedral coordination. Total replacement of aluminum by magnesium yields the mineral saponite; replacement of aluminum by iron yields nontronite. If all octahedral positions are filled by ions, the mineral is trioctahedral; if only two-thirds are occupied, the mineral is dioctahedral.

The montmorillonite structure is always unbalanced by the substitutions noted above. The resulting positive net charge deficiency is balanced by exchangeable cations adsorbed between the unit layers and around their edges. The cation-exchange capacity of montmorillonite is normally quite high (100± milliequivalents per 100 g) and is not appreciably affected by particle size. Substitutions within the structure cause about 80% of the total exchange capacity, and broken bonds are responsible for the remainder.

Other properties. Montmorillonite particles are extremely small and may further disperse in water to units approaching single-cell-layer dimensions. Most montmorillonite units are equidimensional flakes. However, nontronite tends to occur in elongate lath-shaped units, and hectorite, the fluorine-bearing magnesium-rich montmorillonite, is found in thin laths.

There is general agreement that the adsorbed interlayer water between the silicate layers has some sort of definite configuration, but the precise nature of this configuration is not agreed upon. The extent and nature of the orientation of the adsorbed water varies with identity of the adsorbed cations.

When montmorillonite is dehydrated, the interlayer water is lost at a relatively low temperature (212–390°F or 100–200°C). The loss of structural (OH) water begins gradually at about 840–930°F (450–500°C), ending at 1110–1290°F (600–750°C). These temperatures vary with the type and amount of structural substitution. The structure of montmorillonite usually persists to temperatures of the order of 1470–1650°F (800–900°C). On further heating montmorillonite, a variety of phases form, such as mullite, cristobalite, and cordierite, depending on the composition and structure prior to fusion at 1830–2700°F (100–1500°C).

Organic ionic compounds enter into cation-exchange reactions with montmorillonite. Polar organic compounds, like glycerol, react by replacing the interlayer water, causing a shift in the *c*-axis spacing of the montmorillonite units. Thus, the identification of montmorillonite by x-ray diffraction is greatly simplified by preliminary treatment with certain organic reagents. The reaction of montmorillonite and organic material is the base of considerable economic use of montmorillonite clays.

Occurrence. Members of the montmorillonite group of clay minerals vary greatly in modes of formation. Alkaline conditions and the presence of magnesium particularly favor the formation of these minerals. Montmorillonites are stable over a wide temperature range and have formed by low-temperature hydrothermal processes, as well as by weathering processes. Several important modes of occurrence are in soils, in bentonites, in mineral veins, in marine shales, and as alteration products of other minerals. Recent sediments have a fairly high montmorillonite content. *See* BENTONITE; CLAY MINERALS; MARINE SEDIMENTS.

Floyd M. Wahl; Ralph E. Grim

Moon

The Earth's natural satellite (**Table 1**). United States and Soviet spacecraft have obtained lunar data and samples, and Americans have orbited, landed, and roved upon the Moon (**Fig. 1**). After a long hiatus, lunar exploration by spacecraft has resumed in the first decade of the twenty-first century and there is even the prospect that humans may revisit and eventually reside upon the Moon. Though the first wave of human exploration has passed, it left a store of information whose meanings are still being deciphered. Many of the Moon's properties are now well understood, but its origin and relations to other planets remain obscure. Theories of its origin include independent condensation and then capture by the Earth; formation in the same cloud of preplanetary matter with the Earth; fission from the Earth; and formation after the impact of a Mars-sized body on the proto-Earth. Because many of the Moon's geologic processes stopped long ago, its surface preserves a record of very ancient events. However, because the Moon's rocks and soils were reworked by geochemical and impact processes, their origins are partly obscured, so that working out the Moon's early history remains a fascinating puzzle.

The apparent motions of the Moon, its waxing and waning, and the visible markings on its face (Fig. 1), are reflected in stories and legends from every early civilization. At the beginning of recorded history on the Earth, it was already known that time could be reckoned by observing the position and phases of the Moon. Attempts to reconcile the repetitive but incommensurate motions of the Moon and Sun led to the construction of calendars in ancient Chinese and Mesopotamian societies and also, a thousand years later, by the Maya. By about 300 B.C., the Babylonian astronomer-priests had accumulated long spans of observational data and so were able to predict eclipses (**Table 2**).

Motions. The Earth and Moon now make one revolution about their barycenter, or common center of mass (a point about 2900 mi or 4670 km from the Earth's center), in $27^d\ 7^h\ 43^m\ 12^s$ (**Table 3**). This sidereal period is slowly lengthening, and the distance (now about 60.27 earth radii) between centers of mass is increasing, because of tidal friction in the oceans of the Earth. The tidal bulges raised by the Moon are dragged eastward

TABLE 1. Characteristics of the Moon

Characteristics	Values and remarks
Diameter (approximate)	2160 mi (3476 km)
Mass	1/81.301 Earth's mass, or 1.62×10^{23} lb (7348×10^{22} kg)
Mean density	0.604 Earth's, or 209 lb/ft^3 (3.34 g/cm^3)
Mean surface gravity	0.165 Earth's, or 5.3 ft/s^2 (162 cm/s^2)
Surface escape velocity	0.213 Earth's, or 1.48 mi/s (2.38 km/s)
Atmosphere	Surface pressure 10^{-12} torr (1.3×10^{-10} Pa); hints of some charged dust particles and occasional venting of volatiles; tenuous sodium and potassium clouds observed
Magnetic field	Dipole field less than $\sim 0.5 \times 10^{-5}$ Earth's; remanent magnetism in rocks shows past field was much stronger
Dielectric properties	Surface material has apparent dielectric constant of 2.8 or less; bulk apparent conductivity is 10^{-5} mho/m or less
Natural radioactivity	Mainly due to solar- and cosmic-ray-induced background (about 1 milliroentgen per hour for quiet Sun)
Seismic activity	Much lower than Earth's; deep moonquakes occur more frequently when the Moon is near perigee; subsurface layer evident
Heat flow	3×10^{-2} W/m^2 (*Apollo 15* site)
Surface composition and properties	Basic silicates, three sites (Table 4); some magnetic material present; soil grain size is 2–60 μm and 50% is less than 10 μm; soil-bearing strength 15 lb/in.^{2} (1 kg/cm^2) at depth of 1–2 in. (a few centimeters)
Rocks	All sizes up to tens of meters present, concentrated in strewn fields; rock samples from Mare Tranquillitatis include fine- and medium-grained igneous and breccia
Surface temperature range	At equator 260°F (400 K) at noon; −315 to −280°F (80–100 K) night minimum; 3 ft (1 m) below surface, −45°F (230 K); at poles −280°F (~100 K)

by the Earth's daily rotation. The displaced water masses exert a gravitational force on the Moon, with a component along its direction of motion, causing the Moon to spiral slowly outward. The Moon, through this same tidal friction, acts to slow the Earth's rotation, lengthening the day. Tidal effects on the Moon itself have caused its rotation to become synchronous with its orbital period, so that it always turns the same face toward the Earth.

Tracing lunar motions backward in time is very difficult, because small errors in the recent data propagate through the lengthy calculations, and because the Earth's own moment of inertia may not have been constant over geologic time. Nevertheless, the attempt is being made by using diverse data sources, such as the old Babylonian eclipse records and the growth rings of fossil shellfish. At its present rate of departure, the Moon would have been quite close to the Earth about 4.6×10^9 years ago, a time which other evidence suggests as the approximate epoch of formation of the Earth.

The Moon's present orbit (**Fig. 2**) is inclined about 5° to the plane of the ecliptic (Table 3). As a result of differential attraction by the Sun on the Earth-Moon system, the Moon's orbital plane rotates slowly relative to the ecliptic (the line of nodes regresses in an average period of 18.60 years) and the Moon's apogee and perigee rotate slowly in the plane of the orbit (the line of apsides advances in a period of 8.850 years). Looking down on the system from the north, the Moon moves counterclockwise. It travels along its orbit at an average speed of nearly 0.6 mi/s (1 km/s) or about 1 lunar diameter per hour; as seen from Earth, its mean motion eastward among the stars is 13°11′ per day.

As a result of the Earth's annual motion around the Sun, the direction of solar illumination changes about 1° per day, so the lunar phases do not repeat in the sidereal period given above but in the synodic period, which averages $29^d\ 12^h\ 44^m$ and varies some 13 h because of the eccentricity of the Moon's orbit. *See* EARTH ROTATION AND ORBITAL MOTION; ORBITAL MOTION.

When the lunar line of nodes (Fig. 2) coincides with the direction to the Sun, and the Moon happens to be near a node, eclipses can occur. Because of the 18.6-year regression of the nodes, groups of eclipses recur with this period. When it passes through the Earth's shadow in a lunar eclipse, the Moon remains dimly visible because of the reddish light scattered through the atmosphere around the limbs of the Earth. When the Moon passes between the Earth and Sun, the solar eclipse may be total or annular. As seen from Earth, the angular diameter of the Moon (31′) is almost the same as that of the Sun, but both apparent diameters vary because of the eccentricities of the orbits of Moon and Earth. Eclipses are annular when the Moon is near apogee and the Earth is near perihelion at the time of eclipse. A partial solar eclipse is seen from places on Earth that are not directly along the track of the Moon's shadow. *See* ECLIPSE.

The Moon's polar axis is inclined slightly to the pole of the lunar orbit (Fig. 2) and rotates with the same 18.6-year period about the ecliptic pole. The rotation of the Moon about its polar axis is nearly uniform, but its orbital motion is not, owing to the finite eccentricity and Kepler's law of equal areas, so that the face of the Moon appears to swing east and west about 8° from its central position every month. This is the apparent libration in longitude. The Moon does rock to and fro in a very small oscillation about its mean rotation rate; this is called the physical libration. There is also a libration in latitude because of the inclination of the Moon's polar axis. The librations make it possible to see about 59% of the Moon's surface from the Earth.

The lunar ephemeris, derived from precise astronomical observations and refined through lengthy

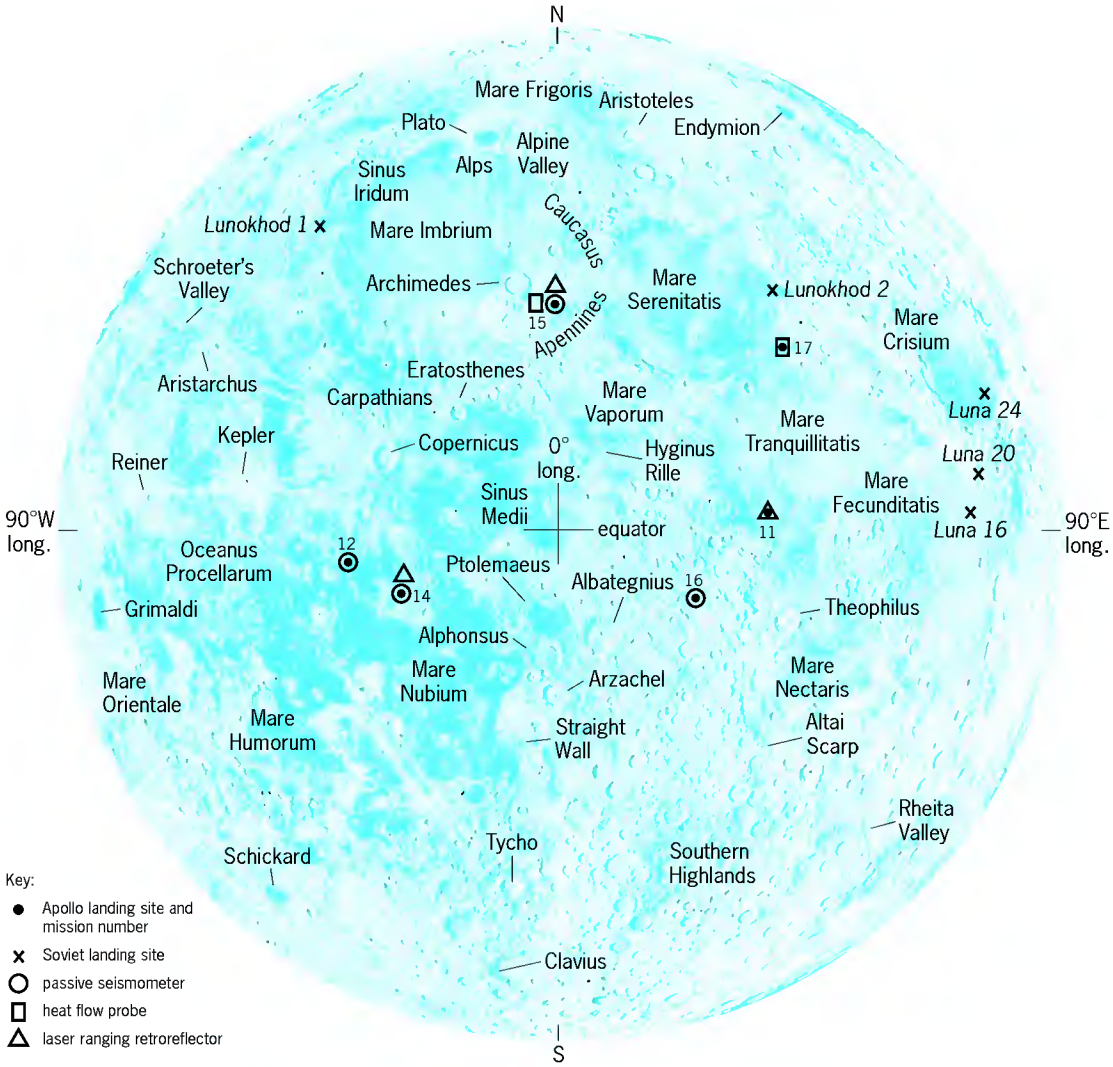

Fig. 1. Map of near side of Moon, showing principal features and some American and Soviet landing sites.

computations of the effects perturbing the movements of the Moon, has now reached a high degree of accuracy in forecasting lunar motions and events such as eclipses. Laser ranging to retroreflectors landed on the Moon, aided by radio ranging to spacecraft, provides measurements of Earth-Moon distances to a precision of the order of meters. *See* PERTURBATION (ASTRONOMY).

Selenodesy. The problem of determining the Moon's true size and shape and its gravitational and inertial properties has been under attack by various methods for centuries (Tables 1 and 2). However, results from space flights have invalidated some of the premises on which the earlier methods were based, and have revealed discrepancies in the older data. The relation between the Moon's shape and its mass distribution is very important to theories of lunar origin and the history of the Earth-Moon system. Radio-tracking data from lunar orbiters indicate that the Moon's gravitational field is ellipsoidal, with the short axis being the polar one (as expected for any rotating body), and with the equatorial section being an ellipse possibly slightly elongated in the Earth-Moon direction. But the Earth-based radar measurements and tracking data from Rangers and Surveyors showed that the Moon's actual surface at the points of landing is about 1.2 mi (2 km) farther from the Earth than expected. Further evidence of an anomalous relationship between mass and shape for the Moon is provided by the mass concentrations in circular maria, discovered through analysis of short-term variations in the 1966 Lunar Orbiter tracking data and then mapped in detail by Apollo tracking. Measurements by *Galileo, Magellan*, and *Clementine* have provided highly accurate knowledge of the shape of the Moon's gravity potential field. By radio altimetry, Apollo confirmed that the Moon's surface on the far side is higher on the average than the near side; that is, the center of mass is offset from the center of figure. The offset is about 1.2 mi (2 km) toward the Earth. These observations suggest that the Moon's crust is thicker on the far side than on the near side. The *Clementine* mission (Table 2) extended measurements to nearly the whole Moon and revealed the depth of a huge basin on the southern far side (**Fig. 3**). *See* SPACE PROBE.

TABLE 2. Growth of human understanding of the Moon

Prehistory	Markings and phases observed, legends created connecting Moon with silver, dark markings with rabbit (shape of maria) or with mud
~ 300 B.C.	Apparent lunar motions recorded and forecast by Babylonians and Chaldeans.
~ 150 B.C.	Phases and eclipses correctly explained, distance to Moon and Sun measured by Hipparchus.
~ A.D. 150	Ancient observations compiled and extended by C. Ptolemy.
~ 700	Ephemeris refined by Arabs.
~ 1600	Empirical laws of planetary motion derived by J. Kepler.
1620	Kepler's dream story, *Somnium,* uses correct description of lunar temperatures.
1609	Lunar craters observed with telescopes by T. Harriot and Galileo.
1650	Moon mapped by J. Hevelius and G. Riccioli; features named by them in system still in use.
1667	Experiments by R. Hooke simulating cratering through impact and volcanism.
1687	Moon's motion ascribed to gravity by I. Newton.
1692	Empirical laws of lunar motion stated by J. D. Cassini.
1700–1800	Lunar librations measured, lunar ephemeris computed using perturbation theory by T. Mayer. Secular changes computed by J. L. Lagrange and P. S. de Laplace. Theory of planetary evolution propounded by I. Kant and Laplace. Many lunar surface features described by J. H. Schroeter and other observers.
1800–1920	Lunar motion theory and observations further refined, leading to understanding of tidal interaction and irregularities in Earth's rotation rate. Photography, photometry, and bolometry applied to description of lunar surface and environment. Lunar atmosphere proved absent. New disciplines of geology and evolution applied to Moon, providing impetus to theories of its origin.
1924	Polarization measured by B. F. Lyot, showing surface to be composed of small particles.
1927–1930	Lunar day, night, and eclipse temperatures measured by E. Pettit and S. B. Nicholson.
1946	First radar return from Moon.
1950–1957	New photographic lunar atlases and geologic reasoning; renewed interest in theories of lunar origin by G. P. Kuiper, H. C. Urey, and E. M. Shoemaker. New methods (for example, isotope dating) applied to meteorites; concepts extended to planetology of Moon. Low subsurface temperatures confirmed by Earth-based microwave radiometry.
1959	Absence of lunar magnetic field (on sunlit side) shown by *Luna 2*.
1960	Eastern far side photographed by *Luna 3*. Slower cooling of Tycho detected during lunar eclipse.
1961	United States commitment to human lunar flight.
1962	Earth-Moon mass ratio measured by *Mariner 2*.
1964	High-resolution pictures sent by *Ranger 7.* Surface temperatures during eclipse measured by Earth-based infrared scan.
1965	Western far side photographed by *Zond 3*.
1966	Surface pictures produced by *Luna 9* and *Surveyor 1*. Radiation dose at surface measured by *Luna 9*. Gamma radioactivity measured by *Luna 10*. High-resolution, broad-area photographs taken by *Lunar Orbiter 1*. Surface strength and density measurements made by *Luna 13*.
1967	Mare soil properties and chemistry measured by *Surveyor 3, 5,* and *6*. Whole front face mapped by *Lunar Orbiter 4*; sites of special scientific interest examined by*Lunar Orbiter 5*. Particle-and-field environment in lunar orbit measured by *Explorer 35*.
1968	Highland soil and rock properties and chemistry measured by *Surveyor 7*. Mass concentrations at circular maria discovered.
1968	Astronauts orbit Moon, return with photographs.
1969	Astronauts land and emplace instruments on Moon, return with lunar samples and photographs.
1969–1972	Lunar seismic and laser retroreflector networks established. Heat flow measured at two sites. Remanent magnetism discovered in lunar rocks. Geologic traverses accomplished. Orbital surveys of natural gamma radioactivity, x-ray fluorescence, gravity, magnetic field, surface elevation, and subsurface electromagnetic properties made at low latitudes. Metric mapping photos obtained. Samples returned by both piloted (United States) and automated (Soviet) missions; sample analyses confirmed early heating and chemical differentiation of Moon, with surface rocks enriched in refractory elements and depleted in volatiles. Age dating of lunar rocks and soils showed that most of the Moon's activity (meteoritic, tectonic, volcanic) occurred more than 3×10^9 years ago.
1975	Giant-impact hypothesis for lunar origin advanced by W. K. Hartmann and D. R. Davis.
1982	Earth-based spectrometry reveals mineral variations over Moon's near side; central peaks of crater Copernicus found to be rich in olivine.
1983	Antarctic meteorite, ALHA 81005, proved to have come from the Moon.
1988–1992	Thin, extended atmosphere of sodium and potassium discovered. Earth-based multispectral observations yield compositional mapping of near side. *Galileo* images Moon during Earth flybys.
1994	*Clementine* maps entire Moon in 11 visible and near-infrared bands. Topography measured by laser altimetry.
1998	Orbital geochemical and gravity mapping by Lunar Prospector spacecraft begins. Increased concentration of near-surface light elements observed at lunar poles, possibly indicating presence of hydrogen in cold-trapped water ice.
2004	Orbital imaging, mineral, chemical, and environmental particles and fields observation by *SMART 1* spacecraft begins.

TABLE 3. Dimensions of Moon's orbit*

Characteristics	Values
Sidereal period (true period of rotation and revolution)	(27.32166140 + 0.000000016*T*) ephemeris days, where *T* is in centuries from 1900
Synodic period (new Moon to new Moon)	(29.5305882 + 0.000000016*T*) ephemeris days
Apogee	252,700 mi or 406,700 km (largest); 251,971 mi or 405,508 km (mean)
Perigee	221,500 mi or 356,400 km (smallest); 225,744 mi or 363,300 km (mean)
Period of rotation of perigee	8.8503 years direct ("direct" meaning that the motion of perigee is in the direction of Moon's motion about the Earth)
Period of regression of nodes	18.5995 years
Eccentricity of orbit	0.054900489 (mean)
Inclination of orbit to ecliptic	5°8′43″ (oscillating ±9′ with period of 173 days)
Inclination of orbit to Earth's Equator	Maximum 28°35′, minimum 18°21′
Inclination of lunar equator	
to ecliptic	1°32′40″
to orbit	6°41′

*In conventional coordinates with origin at the center of the Earth, rather than the Earth-Moon barycenter.

Body properties. The Moon's small size and low mean density (Table 1) result in surface gravity too low to hold a permanent atmosphere, and therefore it was to be expected that lunar surface characteristics would be very different from those of Earth. However, the bulk properties of the Moon are also quite different—the density alone is evidence of that—and the unraveling of the Moon's internal history and constitution is a great challenge to planetologists. *See* EARTH.

The Earth, with its dense metallic fluid core, convective mantle, strong and variable magnetic field with trapped radiation belts, widespread seismic tremors, volcanoes and folded mountain ranges, moving lithospheric plates, and highly differentiated radioactive rocks, is plainly a planet seething with inner activity. Is the Moon also an active, evolving world or is it something very different? The answer lies in a group of related experiments: seismic investigations, heat-flow measurements, surface magnetic and gravity profiles, determination of abundances and ages of the radioactive isotopes in lunar material, and comparison of the latter with those found in the Earth and meteorites. Present theories and experimental data yield the following clues to the problem.

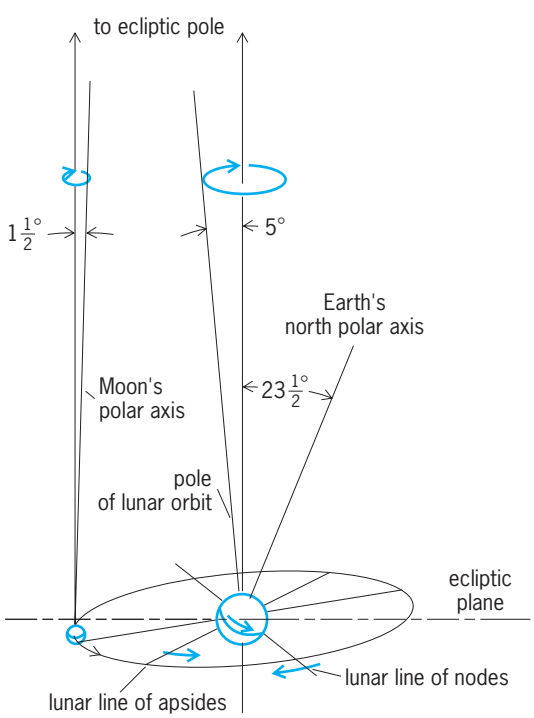

Fig. 2. Sketch of Moon's orbit.

1. The Moon is too small to have compressed its silicates into a metallic phase by gravity; therefore, if it has a dense core at all, the core should be of nickel-iron. But the low mass of the whole Moon does not permit a large core unless the outer layers are of very light material; available data suggest that the Moon's iron core may have a diameter of at most a few hundred kilometers.

2. The Moon has no radiation belts, and behaves as a nonconductor in the presence of the interplanetary field. Moon rocks are magnetized, but the source of the magnetism remains a mystery as there is now little or no general lunar magnetic field.

3. The Moon's natural radioactivity from long-lived isotopes of potassium, thorium, and uranium, expected to provide internal heat sufficient for partial melting, was roughly measured from orbit by *Luna 10*, and the component above the cosmic-ray-induced background radiation was found to be at most that of basic or ultrabasic earthly rock, rather than that of more highly radioactive, differentiated rocks such as granites. *Apollo 11* and *12* rock samples confirmed this result; *Apollo 15*, *16*, and *17* mapped lunar composition and radioactivity from orbit (Fig. 3). X-ray experiments showed higher aluminum-silicon concentration ratios over highland areas and lower values over maria, while magnesium-silicon ratios showed a converse relationship—higher values over maria and lower values over highlands.

4. The Moon is seismically much quieter than the Earth. Moonquakes are small, many of them originate

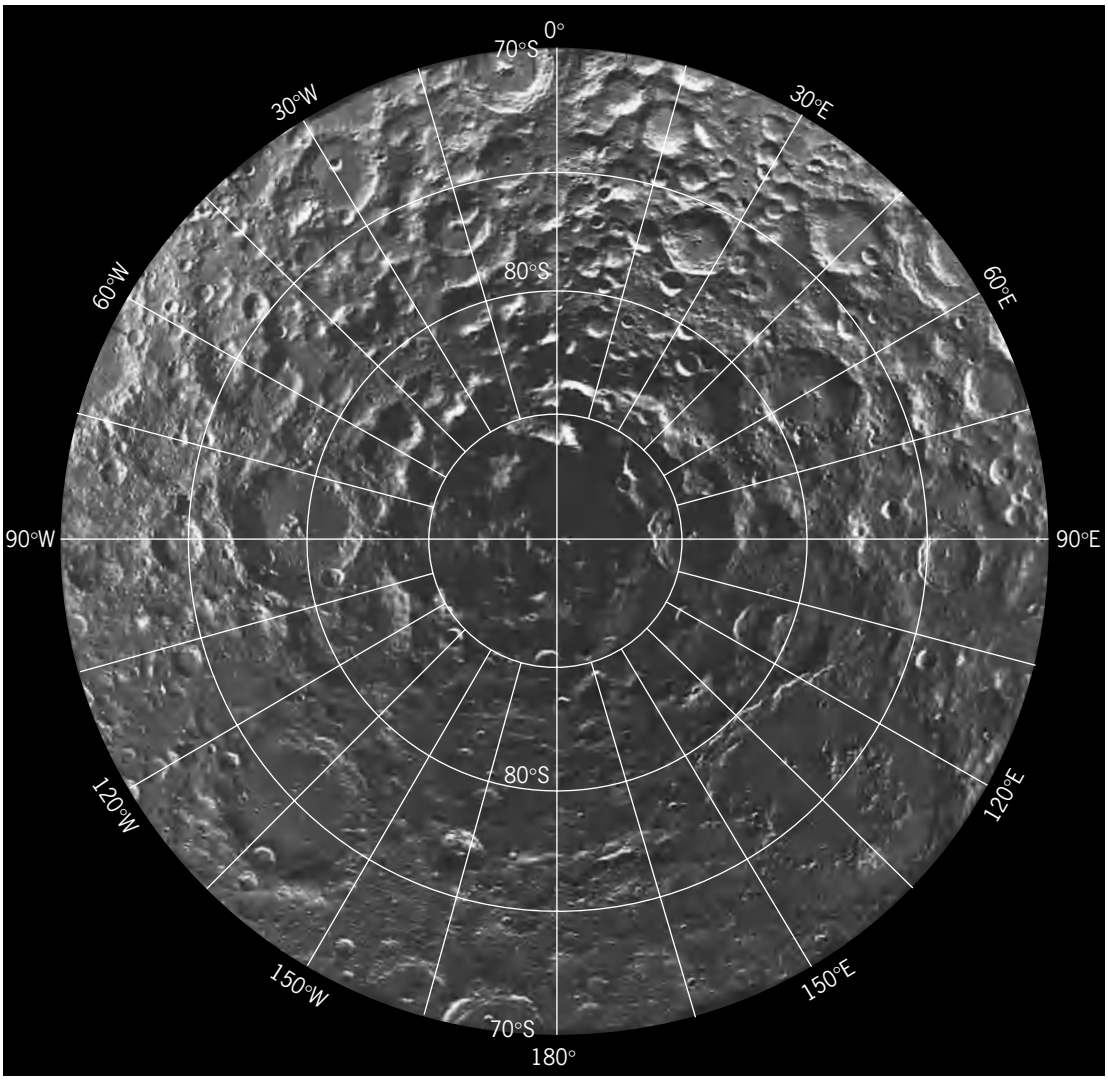

Fig. 3. Mosaic of 1500 *Clementine* images, taken through a red filter, of lunar south polar region. (*Naval Research Laboratory*)

deep in the interior, and activity is correlated with tidal stress: more quakes occur when the Moon is near perigee.

When all of the Apollo observations are taken together, it is evident that the Moon was melted to an unknown depth and chemically differentiated about 4.5×10^9 years ago, leaving the highlands relatively rich in aluminum and an underlying mantle relatively rich in iron and magnesium, with all known lunar materials depleted in volatiles. The subsequent history of impacts and lava flooding includes further episodes of partial melting until about 3.9×10^9 years ago, with the final result being a thick, rigid crust with only minor evidence of recent basaltic extrusions. The temperature profile and physical properties of the Moon's deep interior are, despite the Apollo seismic and heat-flow data, under active debate.

Large-scale surface features. As can be seen from the Earth with the unaided eye, the Moon has two major types of surface: the dark, smooth maria and the lighter, rougher highlands (Fig. 1 and **Fig. 4**). Photography by spacecraft shows that, for some unknown reason, the Moon's far side consists mainly of highlands. Both maria and highlands are covered with craters of all sizes. Craters are more numerous in the highlands than in the maria, except on the steeper slopes, where downhill movement of material apparently tends to obliterate them. Numerous different types of craters can be recognized. Some of them appear very similar to the craters made by explosions on the Earth; they have raised rims, sometimes have central peaks, and are surrounded by fields of hummocky, blocky ejecta. Others are rimless and tend to occur in lines along cracks in the lunar surface. Some of the rimless craters, particularly those with dark halos, may be gas vents bringing dark pyroclastic material to the surface. Most prominent at full moon are the bright ray craters (Fig. 1) whose grayish ejecta appear to have traveled for hundreds of kilometers across the lunar surface. Observers have long recognized that some erosive process has been and is probably still active on the Moon. For example, when craters overlap so that their relative ages are evident, the younger ones are seen to have sharper outlines than the older ones.

Bombardment of the airless Moon by meteoritic matter and solar particles, and extreme temperature cycling, are now considered the most likely erosive agents, but local internal activity is also a possibility. Rocks returned by the Apollo astronauts are covered with tiny glass-lined pits, confirming erosion by small high-speed particles.

The lunar mountains, though very high (26,000 ft or 8000 m or more), are not extremely steep, and lunar explorers see rolling rather than jagged scenery (**Fig. 5**). There are steep slopes (30–40°) on the inside walls and central peaks of recent craters, where the lunar material appears to be resting at its maximum angle of repose, and rocks can be seen to have rolled down to the crater bottoms.

Though widespread networks of cracks are visible, there is no evidence on the Moon of the great mountain-building processes seen on the Earth. There are some low domes suggestive of volcanic activity, but the higher mountains are all part of the gently rolling highlands or the vast circular structures surrounding major basins. One of these is the Mare Orientale (**Fig. 6**), a large concentric structure which is almost invisible from Earth because it lies just past the Moon's western limb; at favorable librations, parts of its basin and mountain ramparts can be seen. A great region of radial sculpture surrounds the Orientale basin, strongly suggesting a catastrophic origin, with huge masses of matter thrown outward from the center. However, the lowest parts of the concentric rings are flooded by dark mare material which displays a gentle appearance. Other basins, namely, Imbrium, Serenitatis, and Crisium, appear more fully flooded (Fig. 1). These maria were created by giant impacts, followed by subsidence of the ejecta and (probably much later) upwelling of lava from inside the Moon. Examination of small variations in Lunar Orbiter motions has revealed that each of the great circular maria is the site of a positive gravity anomaly (excess mass). The old argument about impact versus volcanism as the primary agent in forming the lunar relief, reflected in lunar literature over the past 100 years, appears to be entering a new, more complicated phase with the confirmation of extensive flooding of impact craters by lava on the Moon's near side, while on the far side, where the crust is thicker, the great basins remain mostly empty.

In some of the Moon's mountainous regions bordering on the maria are found sinuous rilles (**Fig. 7**). These winding valleys, some of them known since the eighteenth century, were shown in Lunar Orbiter pictures to have an exquisite fineness of detail. Some of them originate in small circular pits and then wriggle delicately across the Moon's gentle slopes for hundreds of kilometers, detouring around even slight obstacles, before vanishing on the plains. Though their resemblance to meandering rivers is strong, the sinuous rilles have no tributaries or deltas. No explanation for them yet offered has proved entirely convincing.

Other strange large-scale features, observed by telescope and then revealed in more detail by spacecraft cameras, are the ghost craters, circular structures protruding slightly from the maria, and the low, ropy wrinkle ridges that stretch for hundreds of kilometers around some mare borders. Large, light-toned swirly features, with no topography but sometimes with magnetic anomalies, are unexplained (**Fig. 8**).

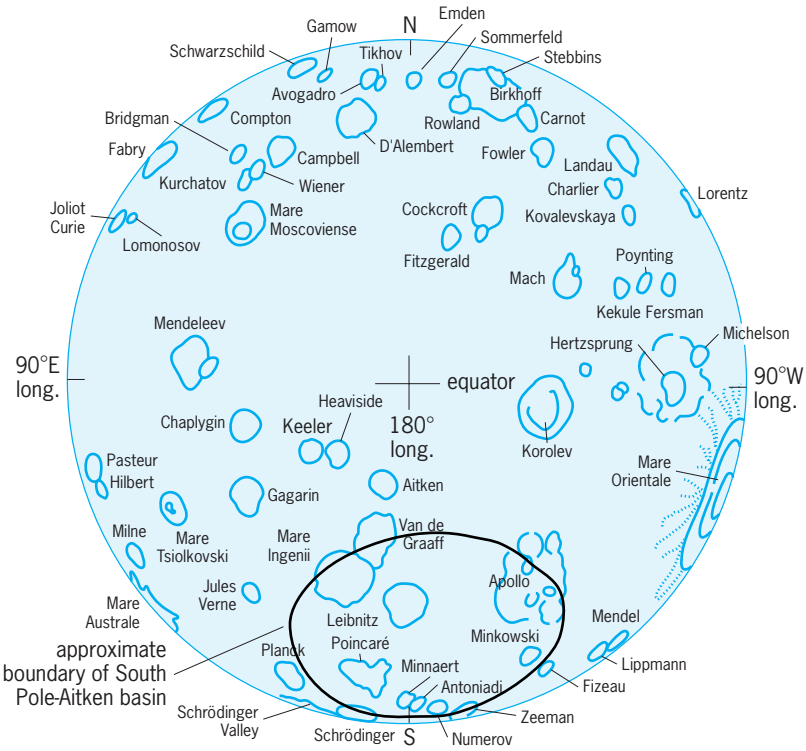

Fig. 4. Map of far side of Moon.

Small-scale surface features. Careful observations, some of them made decades before the beginning of space flight, revealed much about the fine-scale nature of the lunar surface. Since the smallest lunar feature telescopically observable from the Earth is some hundreds of meters in extent, methods other than direct visual observation had to be used. Photometry, polarimetry, and later radiometry and radar

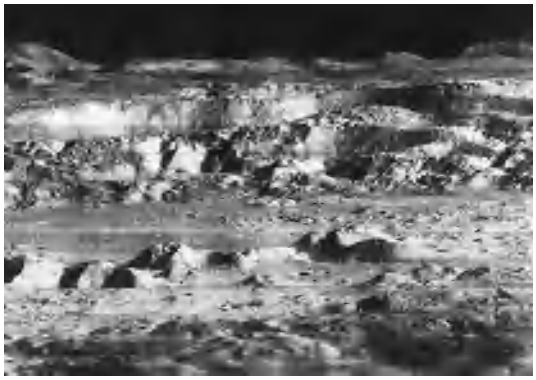

Fig. 5. The crater Copernicus, showing the central peaks, slump terraces, patterned crater walls, and (background) slopes of the Carpathian Mountains. (*Langley Research Center, NASA*)

Fig. 6. Mare Orientale. (*Langley Research Center*, *NASA*)

probing gave the early fine-scale data. Some results of these investigations suggested bizarre characteristics for the Moon. Nevertheless, many of their findings have now been confirmed by spacecraft. The Moon seems to be totally covered, to a depth of at least tens of meters, by a layer of rubble and soil with very peculiar optical and thermal properties. This layer is called the regolith. The observed optical and radio properties are as follows. *See* PHOTOMETRY; POLARIMETRY; RADAR; RADIOMETRY.

1. The Moon reflects only a small portion of the light incident on it (the average albedo of the maria is only 7%). *See* ALBEDO.

2. The full moon is more than 10 times as bright as the half moon.

3. At full moon, the disk is almost equally bright all the way to the edge; that is, there is no "limb darkening" such as is observed for ordinary spheres, whether they be specular or diffuse reflectors.

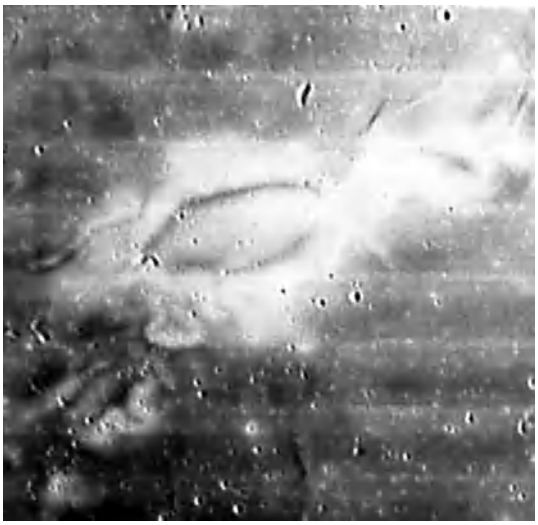

Fig. 8. *Lunar Orbiter* image of Reiner Gamma, the only example of a swirl feature on the Moon's near side. (Only three other such features are known to exist, two on the Moon's far side and one on Mercury.) The feature has an overall dimension of about 43 mi (70 km). It exhibits an almost total lack of elevation and other forms of relief. A relatively strong field of magnetization has been discovered to coincide with the feature. (*NASA*)

Fig. 7. Aristarchus-Harbinger region of the Moon, photographed from the *Apollo 15* spacecraft in lunar orbit, with the craters Aristarchus and Herodotus and Schroeter's Valley, the largest sinuous rille on the Moon. The impact crater Aristarchus, about 25 mi (40 km) in diameter and more than 2.5 mi (4 km) deep, lies at the edge of a mountainous region that shows evidence of volcanic activity. (*NASA*)

4. Color variations are slight; the Moon is a uniform dark gray with a small yellowish cast. Some of the maria are a little redder, some a little bluer, and these differences do correlate with large-scale surface morphology, but the visible color differences are so slight that they are detectable only with special filters. Infrared spectral differences are more pronounced and have provided a method for mapping variations in the Moon's surface composition. Multispectral remote sensing, both from Earth and from the *Clementine* and *SMART 1* spacecraft in polar lunar orbits has greatly advanced the study of compositional variations over the lunar surface.

5. The Moon's polarization properties are those of a surface covered completely by small, opaque grains in the size range of a few micrometers.

6. The material at the lunar surface is an extremely good thermal insulator, better than the most porous terrestrial rocks. The cooling rate as measured by infrared observations during a lunar eclipse is strongly variable; the bright ray craters cool more slowly than their surroundings.

7. The Moon emits thermal radiation in the radio wavelength range; interpretations of this and the

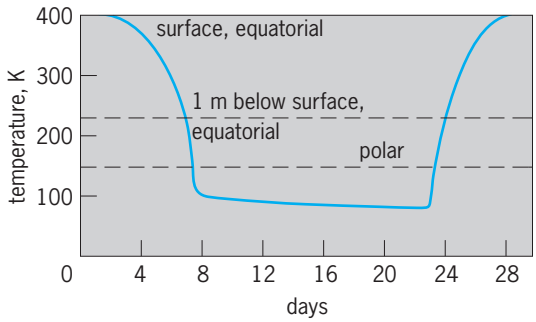

Fig. 9. Lunar surface and subsurface temperatures.

infrared data yield estimates of surface and shallow subsurface temperatures (**Fig. 9**).

8. At wavelengths in the meter range, the Moon appears smooth to radar, with a dielectric constant lower than that of most dry terrestrial rocks. To centimeter waves, the Moon appears rather rough, and at visible light wavelengths it is extremely rough (a conclusion from observations 1–5 above).

These observations all point to a highly porous or underdense structure for at least the top few millimeters of the lunar surface material. The so-called backscatter peak in the photometric function, which describes the sudden brightening near full Moon, is characteristic of surfaces with deep holes or with other roughness elements that are shadowed when the lighting is oblique. However, *Clementine* observations showed that, in addition to shadowing, light reflections within soil particles may contribute to the brightening.

The Ranger, Luna, Surveyor, and Lunar Orbiter missions made it clear that these strange electromagnetic properties are generic characteristics of the dark-gray, fine soil that appears to mantle the entire Moon, softening most surface contours and covering everything except occasional fields of rocks (**Figs. 10** and **11**). This soil, with a slightly cohesive character like that of damp sand and an average chemical composition similar to that of some basic silicates on the Earth, is a product of the radiation, meteoroid, and thermal environment at the lunar surface. A surface texture called patterned ground (Fig. 10) is common on the moderate slopes of the Moon. This widespread phenomenon is unexplained, though there are some similar surfaces developed on the Earth when unconsolidated rock, lava, or glacial ice moves downhill beneath an overburden. At many places on the Moon, there is unmistakable evidence of downward sliding or slumping of material and rolling rocks. There are also a few instances of apparent upwelling, as well as numerous "lakes" where material has collected in depressions.

Magnets on the Surveyors collected magnetic particles from the soil, demonstrating the presence of either meteoritic or native iron minerals at the sites examined. Meteoroid experiments on the Lunar Orbiters showed about the same flux of small particles as is observed at the Earth, so that the lunar soil would be expected to contain a representative sample of meteoritic and possibly also cometary matter. Apollo results confirmed and extended the Surveyor data and also indicated that glassy particles are abundant in and on the soil. Evidence of micrometeoroid bombardment is seen in the many glass-lined microcraters found on lunar rocks. *See* COMET; METEOR.

Chemical, mineral, and isotopic analyses of minerals from the *Apollo 11* site showed that mare rocks there are indeed of the basic igneous class and are very ancient ($3–4 \times 10^9$ years). The *Apollo 12* samples are significantly younger, suggesting that Mare Tranquillitatis and Oceanus Procellarum were formed during a long and complex lunar history. The *Apollo 12* astronauts visited *Surveyor 3* and brought back parts of that spacecraft to permit analysis of the effects of its $2^1/_2$-year exposure on the surface of the Moon. The lunar rock and soil samples returned by the Apollo and Luna missions have yielded much new information on the composition and

Fig. 10. Lunar soil and rocks, and the trenches, 2 in. (5 cm) wide, made by *Surveyor 7*. (*Jet Propulsion Laboratory*, *NASA*)

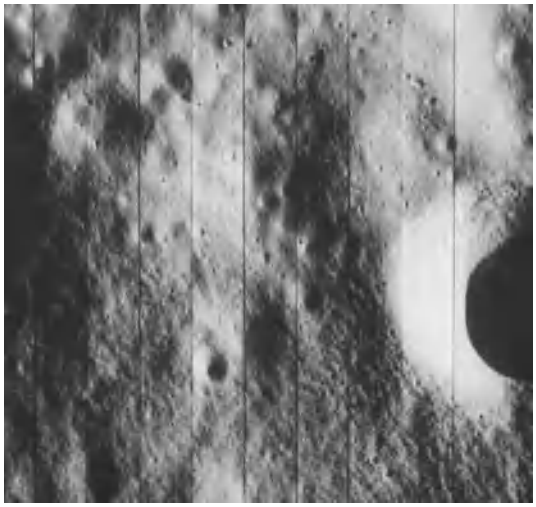

Fig. 11. Lunar patterned ground, a common feature on moderate slopes. (*Langley Research Center*, *NASA*)

TABLE 4. Some selected data from Apollo and Luna missions

Mission	Main sample properties	Other data
Apollo 11	Mare basalts, differentiated from melt at depth 3.7×10^9 years ago. Some crystalline highland fragments in soils. Unexpected abundance of glass. Much evidence of impact shock and microcratering. No water or organic materials.	Study of seismic properties showed low background, much scattering, and low attenuation.
Apollo 12	Basalts 3.2×10^9 years old. One sample 4.0×10^9 years old includes granitic component. Some samples with high potassium, rare-earth elements, and phosphorus (KREEP) may be Copernicus crater ejecta.	Surveyor parts returned showed effects of solar and cosmic bombardment.
Apollo 13	Spacecraft failure–no samples.	Despite emergency, some lunar photos returned.
Luna 16	Basalt 3.4×10^9 years old, relatively high Al content.	
Apollo 14	Shocked highland basalts, probably Imbrium ejecta, 3.95×10^9 years old, higher Al and lower Fe than mare materials.	Deep moonquakes.
Apollo 15	Highland anorthosites including one sample 4.1×10^9 years old, mare basalts similar to *Apollo 11* samples.	Orbital remote sensing began mapping of surface compositions.
Luna 20	Possibly Crisium ejecta, 3.9×10^9 years old.	
Apollo 16	Highland anorthosite breccias $3.9–4 \times 10^9$ years old, also possibly Imbrium ejecta.	Seismic network began recording locations of impacts and deep moonquakes; orbital compositional mapping extended.
Apollo 17	Variety of basalts and anorthosites $3.7–4 \times 10^9$ years old, possibly volcanic glass, few dunite fragments 4.48×10^9 years old, possibly surviving from before the great highland bombardment.	Orbital mapping, and study of sesimic particle-and-field, and subsurface electrical properties yielded comprehensive (but still unexplained) picture of Moon.
Luna 24	Very low titanium basalts from Mare Crisium. Sample includes rock 3.3×10^9 years old.	

history of the Moon (**Table 4**). Among the dominant characteristics of these rocks are enrichment in refractories, depletion in volatiles, much evidence of repeated breaking up and rewelding into breccias, and ages since solidification extending back from the mare flows of $3–4 \times 10^9$ years ago into the period of highland formation more than 4×10^9 years ago, but not as yet including the time of the Moon's original accretion. Soil characteristics, through the analysis of embedded solar-wind atoms, have also yielded information on the ancient history of the Sun.

In 1983 an Antarctic meteorite was found to resemble some of the lunar samples, and it was determined by analysis to have come from the Moon. Many more lunar and Martian meteorites are now known and have been added to the *Apollo* and *Luna* sample collections. *See* METEORITE.

As the Apollo missions progressed, each new landing site was selected with the aim of elucidating more of the Moon's history. A main objective was to sample each of the geologic units mapped by remote observation, either by landing on it or by collecting materials naturally transported from it to the landing site. Although this process did result in collection of both mare and highland materials with a wide range of ages and chemical compositions, it did not result in a complete unraveling of the history of the Moon. Apparently, the great impacts of $3–4 \times 10^9$ years ago erased much of the previous record, resetting radioactive clocks and scrambling minerals of diverse origins into the complicated soils and breccias found today. *See* SPACE FLIGHT.

Atmosphere. Though the Moon may at one time have contained appreciable quantities of the volatile elements and compounds (for example, hydrogen, helium, argon, water, sulfur, and carbon compounds) found in meteorites and on the Earth, its high daytime surface temperature and low gravity would cause rapid escape of the lighter elements. Solar ultraviolet and x-ray irradiation would tend to break down volatile compounds at the surface, and solar charged-particle bombardment would ionize and sweep away even the heavier gas species. Visual observations from the Earth, looking for a twilight glow of the lunar atmosphere just past the terminator on the Moon, and watching radio-star occultations have all been negative, setting an upper limit of 10^{-12} times the Earth's sea-level atmospheric density for any lunar gas envelope. However, subsequent ground-based spectroscopic observations have revealed a tenuous cloud of sodium and potassium atoms extending hundreds of kilometers above the lunar surface. Except for this cloud, either the lunar volatile compounds have vanished into space or they are trapped beneath the surface. The samples returned by Apollo are enriched in refractory elements, depleted in volatiles, and impregnated with hydrogen, helium (including helium-3, a possible nuclear fusion fuel), and other gases from the solar wind. No water appears to have ever been present at any Apollo site, and carbonaceous materials were present, if at all, only in very small amounts.

Occasional luminescent events reported by reliable observers suggest that some volcanic gases are vented from time to time on the Moon, particularly in the regions of the craters Aristarchus and Alphonsus. The *Apollo 17* samples included glassy particles indicative of pyroclastic eruptions. A slight, transient atmosphere does exist on the night side of the Moon as a result of the trapping and release of gas molecules at the very low temperatures prevailing there; also frozen liquids or gases could exist in permanently shadowed crater bottoms near the lunar poles. Data from *Clementine* suggest that favorable trapping conditions may exist near the south pole

(Fig. 3), and neutron spectrometry by the *Lunar Prospector* orbiter confirms concentrations of a light element, most likely hydrogen, within a meter of the surface at both poles. While opinions differ as to the source of this hydrogen, it strongly suggests the presence of cold-trapped water ice. *Lunokhod 2*, a Soviet roving spacecraft, measured a slight glow attributed to a very thin cloud of small dust particles within meters of the surface, which could explain the Surveyor observations of a slight horizon glow after sunset. Also, the ALSEP (Apollo lunar surface experiments packages) experiments landed by Apollo have occasionally detected small gas emanations, including water, from unknown sources; contamination from the landed spacecraft is hard to exclude from these measurements.

Lunar resources. Enough is known about the Moon to show that it is a huge storehouse of metals, oxygen (bound into silicates), and other materials potentially available for future human use in space. Because of the Moon's weak gravity, lunar materials could be placed into orbit at less than one-twentieth of the energy cost for delivering them from Earth. An objective of several lunar missions now planned in the United States, Europe, India, China, and Japan is to augment lunar scientific knowledge and thus add to understanding the possible uses of the natural resources of the Moon.

James D. Burke

Bibliography. D. Baker, *The Moon*, vol. 1: *Physics, Geology and Evolution*, 1992; P. Eckart, *The Lunar Base Handbook*, McGraw-Hill, 2000; W. K. Hartmann, R. J. Phillips, and G. J. Taylor (eds.), *Origin of the Moon*, 1986; G. Heiken, *Lunar Sourcebook: A User's Guide to the Moon*, Cambridge University Press, 1991; M. W. Hutddleston, *Lunar Swirls, Magnetic Anomalies, and Reiner Gamma Formation*, Association of Lunar & Planetary Observers, 1994; Lunar and Planetary Institute, *Proceedings of the Lunar and Planetary Science Conferences*, 1970–2006; W. W. Mendell (ed.), *Lunar Bases and Space Activities of the 21st Century*, Lunar and Planetary Institute, 1985; A. Rukl, *Atlas of the Moon*, 1992; Special section on the *Clementine* mission, *Science*, 266:1835–1862, December 16, 1994; H. H. Schmitt, *Return to the Moon: Exploration, Enterprise, and Energy in the Human Settlement of Space*, Copernicus Books, 2006; P. D. Spudis, *The Once and Future Moon*, Smithsonian Press, 1997; S. R. Taylor, *Planetary Science: A Lunar Perspective*, 1982; D. Wilhelms et al., *Geologic History of the Moon*, U.S. Geol. Sur. Prof. Pap. 1348, 1988.

Moose

An even-toed ungulate (Artiodactyla) which is a member of the deer family, Cervidae. *Alces alces* is the largest member of the family and ranges in the boreal forested areas throughout North America and in northern Eurasia. The moose is known as the elk in Europe and is believed by some authorities to be a race of the American moose (*A. americana*). The adult male is 6 ft (1.8 m) high, weighs over 1200 lb

The moose, a member of the deer family, Cervidae.

(550 kg), and has spatulate antlers which may be over 6 ft (1.8 m) in width (see **illus.**). The legs are long, making the animal well-adapted for its feeding habits of wading for aquatic plants and browsing on trees and bushes. The moose live in small groups during the summer but tend to form larger groups for defense during the winter, since they are susceptible to predation from wolves and even wolverines. During the rutting season in the early fall, the male gathers a number of cows together, and mating takes place. After a gestation period of about 37 weeks, one or two calves are born. The moose is a big-game animal, but hunting restrictions have helped to maintain its numbers; it is abundant in Canada and the northern United States. *See* ARTIODACTYLA.

Charles B. Curtin

Moraxella

A genus of bacteria that are parasites of mucous membranes. Subgenus *Moraxella* is characterized by gram-negative rods that are often very short and plump, frequently resembling a coccus, and usually occurring in pairs. Subgenus *Branhamella* has gram-negative cocci occurring as single cells or in pairs with the adjacent sides flattened. They are usually harmless parasites of humans and other warm-blooded animals and are generally considered not to be highly pathogenic. Most species may be opportunistic pathogens in predisposed or debilitated hosts.

General characteristics. Members of the genus *Moraxella* are generally characterized as being nonspore-forming, nonmotile cells, often demonstrating a tendency to resist Gram decolorization. They are chemoorganotrophic and nutritionally fastidious, strictly aerobic, indophenol oxidase–positive, and usually catalase-positive. They do not produce acid from carbohydrates. Colonies are not pigmented. All species will grow on blood agar media commonly used in clinical laboratories, but some strains of *Moraxella* (*Moraxella*) *lacunata* may require media containing heated blood, such as chocolate agar. Most strains, with a few exceptions of *M.* (*M.*) *lacunata*, will grow on blood agar base media without

blood, but the addition of serum may improve growth of such strains. *Moraxella* (*M.*) *lacunata* and most *M.* (*M.*) *bovis* strains are the only *Moraxella* species that can liquefy the heat-coagulated serum in Loeffler coagulated serum agar slants.

There are presently six species in the subgenus *Moraxella*: *M.* (*M.*) *lacunata* (also known as *Diplobacillus moraxaxenfeld* and *liquefaciens*), *M.* (*M.*) *bovis*, *M.* (*M.*) *nonliquefaciens* (also known as *Bacillus duplex nonliquefaciens*), *M.* (*M.*) *atlantae*, *M.* (*M.*) *phenylpyruvica*, and *M.* (*M.*) *osloensis*. The different species are recognized on the basis of phenotypic properties, including liquefaction of coagulated serum, hemolysis of human blood in blood agar media, nitrate reduction, phenylalanine deaminase activity, urease activity, and growth on mineral salts medium with ammonium ion and acetate as the sole carbon source.

The subgenus *Branhamella* presently contains four species: *M.* (*B.*) *catarrhalis*, *M.* (*B.*) *caviae*, *M.* (*B.*) *ovis*, and *M.* (*B.*) *cuniculi*. The different species are recognized on the basis of hemolysis of human blood in blood agar media, and nitrate and nitrite reduction, among other properties.

Phenotypic relations. Members of *Moraxella* must be distinguished from species of *Neisseria* since both genera have similar phenotypic properties. The latter bacteria are characterized by cocci occurring singly but more often in pairs with adjacent sides flattened. They reduce nitrite, but usually not nitrate, and demonstrate carbonic anhydrase. Many species produce acid from carbohydrates, and some species produce a greenish-yellow carotenoid pigment.

The subgenus *Branhamella* is differentiated from *Neisseria* species by its frequent reduction of nitrate and negative reactions for carbohydrate acidification, pigment production, and carbonic anhydrase synthesis. The subgenus *Moraxella* is distinguished from *Neisseria* species by its rod-shaped appearance. Additionally, species of the subgenus *Moraxella* produce negative reactions for nitrite reduction, carbohydrate acidification, pigment production, and carbonic anhydrase synthesis.

Distinguishing short rods (subgenus *Moraxella*) from cocci (*Branhamella* and *Neisseria*) by microscopic observation sometimes is difficult. A reliable test for differentiation is to culture the organism in the presence of subinhibitory levels of penicillin: rod-shaped organisms form long, stringlike cells, but cocci retain their coccal morphology.

Pathogens. *Moraxella* (*M.*) *lacunata*, the type species of the subgenus *Moraxella*, was described independently by V. Morax in 1896 and T. Axenfeld in 1897, thus the name Morax-Axenfeld bacillus. It was a significant causative agent of human conjunctivitis and keratitis in the past but is only rarely isolated at present. Infectious keratoconjunctivitis in cattle, called pinkeye, is caused by *M.* (*M.*) *bovis*. The natural habitats of both organisms are healthy conjunctiva and sites in the upper respiratory tract of humans and cattle, respectively. *Moraxella* (*M.*) *nonliquefaciens* is most frequently isolated from the nasal cavity of humans, which is probably the natural habitat, as well as other sites of the respiratory tract. It is considered to be a well-established parasite of humans and rarely causes disease, but it has been associated with endophthalmitis and pneumonitis with pulmonary abscess. *Moraxella* (*M.*) *atlantae*, only rarely isolated, has been found in human blood, cerebrospinal fluid, and abscess. Its natural habitat and clinical significance are unknown. *Moraxella* (*M.*) *phenylpyruvica* is isolated from human sources, including urine, blood, cerebrospinal fluid and the genitourethral tract, and from sheep, cattle, goats, and pigs. Its clinical significance is uncertain. *Moraxella* (*M.*) *osloensis* is isolated from the upper respiratory tract, genitourethral tract, blood, and cerebrospinal fluid of humans. Usually a harmless parasite, this organism has been frequently associated with such human infections as osteomyelitis, endocarditis, septicemia, meningitis, stomatitis, and septic arthritis.

Moraxella (*B.*) *catarrhalis*, the type species of *Branhamella*, is the only species of this subgenus recovered from humans. Its natural habitat is the nasal cavity. The organism is considered to be a well-adapted parasite but has been judged the etiologic agent of middle-ear infection, maxillary sinus infection, bronchitis, tracheitis, conjunctivitis, pneumonia, otitis media of infants, respiratory disease in the compromised host, septicemia, meningitis, and endocarditis. The remaining *Branhamella* species (*caviae*, *ovis*, and *cuniculi*) are parasites of guinea pigs, sheep, cattle, and rabbits.

Moraxella species are susceptible to most antimicrobial agents with the exception of the lincomycins. Their usually high susceptibility to the penicillins is a feature that separates them from most other gram-negative rods. *See* CLINICAL MICROBIOLOGY.

Gerald Gilardi

Bibliography. W. A. Clark et al., *Identification of Unusual Pathogenic Gram-Negative Aerobic and Facultatively Anaerobic Bacteria*, 1985; S. D. Henriksen, *Moraxella*, *Acinetobacter*, and the Mimeae, *Bacteriol. Rev.*, 37:522–561, 1973; N. R. Kreig (ed.), *Bergey's Manual of Systematic Bacteriology*, vol. 1, 1984.

Mordant

A chemical substance, usually containing a metallic ion, used to facilitate the fixing of a dye to a fiber. From the earliest days of dyeing, it was observed that for many natural dyes the color imparted to a fabric was enhanced and the fastness (resistance to fading or running) improved if the fabric was first treated with a solution containing a soluble metallic salt of chromium, aluminum, or copper. In the simplest process model, metal ions are incorporated on the surface and within the fibers of the fabric and, upon exposure to dye molecules at the appropriate pH and temperature, form colored metal-dye complexes. The metal ion is called a mordant. The complex formation process is called chelation, and can result in a fast and deeply colored fabric. One

of the oldest mordant dyes is alizarin, derived from the root of the madder plant. Chemically, it is an anthraquinone derivative, and the dye color depends upon the metal with which it is complexed; with barium, a blue color is obtained, and with aluminum a rose-red. *See* CHELATION; CHROMIUM; PH; QUINONE.

In the dyeing process, an organic compound containing a chromogenic structural group is used to impart color to a substrate—fabric, paper, or other object. In order that the compound be useful as a dye, it must be bound to the fabric or surface in such manner that it will remain fixed during washing and cleaning; that is, it must be fast. The actual mechanism of the bonding process may be quite complex, and depends upon the structure of the dye and upon the chemical nature of the surface.

Mordant action is a complex chemical process, and a general mechanism is not applicable. As an example, chrome dyes form colored chelates with Cr(III) ions [Cr(III) represents chromium atoms in an oxidation state of +3]. Chromium is applied to a wool fabric in the form of dichromate ions in which the oxidation state is +6, Cr(VI). Adsorption occurs by the attraction of the negatively charged dichromate ions to the positively charged base groups in the wool. Before chelation with a dye molecule, the Cr(VI) atoms must be reduced to the Cr(III) state. This occurs by the oxidation of the sulfide groups in the fiber by dichromate and the simultaneous reduction of Cr(VI). The resulting metal-dye chelate and color depend upon pH, temperature, and concentration.

When groups on dye molecules can react at sites on the substrate's surface to form covalent bonds, direct dyeing is accomplished. Wool and silk fibers contain molecules comprising chain structures linked together by peptide units. These units contain —C═O (carbonyl) groups, —S—S—(disulfide) linkages, and basic —NH_2 (amino) groups. A dye with a reactive acidic group may form a chemical bond through a direct reaction at the fiber surface. Martius Yellow, for example, is a dye that contains acidic OH (hydroxyl) groups. It is bonded to wool or silk through reaction with the amino groups. *See* PEPTIDE; SILK; WOOL.

In a vat dyeing process, the fabric is immersed in a solution containing dye in a soluble form. Through an oxidation or other chemical transformation, the dye molecules are converted to an insoluble form and dispersed as a precipitate throughout the fabric. A true chemical bonding is not believed to be involved.

The science and practice of mordant dyeing draws heavily upon metal-chelate chemistry. Many synthetic dyes today incorporate metal atoms in the dye structure and are used without mordants. *See* DYE; DYEING.

Francis Johnston

Bibliography. A. Johnson (ed.), *The Theory of Coloration of Textiles*, Society of Dyers and Colourists, West Yorkshire, U.K., 1989; D. Waring and G. Hallas (eds.), *The Chemistry and Application of Dyes*, Plenum Press, New York, 1990.

Mormonilloida

A copepod order of crustaceans containing four species. *Mormonilla minor* and *M. phasma* were described by Wilhelm Giesbrecht in the late 1800s; *M. polaris* and *M. atlantica* were described by Georg Ossian Sars and Richard Norris Wolfenden, respectively, in the early 1900s. The status of a fifth species in a second genus, Thomas Scott's *Corynuropis tenuicaudatus*, remains uncertain. In 1992, males of the order were collected and described for the first time. *See* COPEPODA; CRUSTACEA.

Description. Lengths range from 0.2–1.7 mm (0.01–0.07 in.) for adult females to 0.9–1.4 mm (0.04–0.06 in.) for adult males. The second thoracic somite, bearing the first swimming leg, articulates with the first, and the last two thoracic somites are associated with the posterior part of the body (podoplean architecture) [see **illustration**]. The only internal organ system that has been studied is the musculature; it is simpler than the musculature of copepods from other orders. One or two small embryos are carried in paired embryo sacs attached to the female; nothing else is known of the development. Species of Mormonilloida are deep-water, free-swimming copepods reported from the Arctic, eastern Indian, and eastern North Atlantic oceans, and the Mediterranean, Red, and Arabian seas. Specimens have been collected below 250 m (820 ft) in the Arctic Ocean, between 410 and 700 m (1350 and 2300 ft) in the eastern North Atlantic Ocean, and between 125 and 1850 m (410 and 6080 ft) in the Mediterranean, Red, and Arabian seas and in the eastern Indian Ocean.

Phylogeny. Phylogenetic relationships of Mormonilloida to the other copepod orders remain uncertain because body architecture, as described

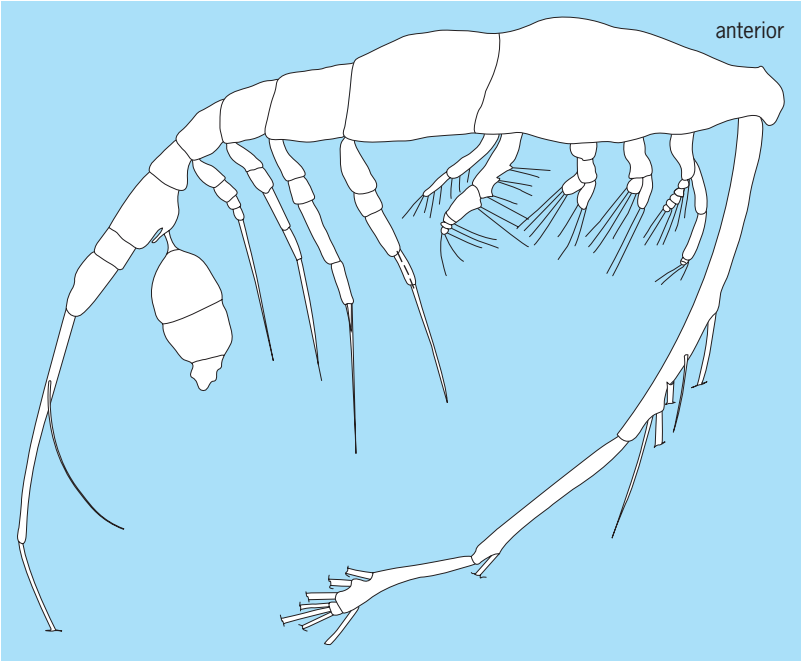

Side view of a mormonilloidan.

by the association of somites along the anterior-posterior axis, places Mormonilloida close to the podoplean copepods, such as cyclopoids and harpacticoids, while the multisegmental configuration of several limbs suggests a close relationship with gymnoplean calanoids (the last thoracic somite with the posterior part of the body). The phylogeny appears unclear in the analyses because neither body architecture nor limb structure is assumed to converge. *See* CALANOIDA; CYCLOPOIDA; HARPACTICOIDA.

Frank D. Ferrari

Bibliography. G. A. Boxshall, The comparative anatomy of two copepods, a predatory calanoid and a particle-feeding mormonilloid, *Phil. Trans. Roy. Soc. London*, Ser. B (Biol. Sci.), vol. 311, 1985; G. A. Boxshall, The planktonic copepods of the northeastern Atlantic Ocean: Harpacticoida, Siphonostomatoida and Mormonilloida), *Bull. Brit. Mus. Nat. Hist. (Zool.)*, vol. 35, 1979; R. A. Buchanan and A. D. Sekerak, Vertical distribution of zooplankton in eastern Lancaster Sound and western Baffin Bay, July-October 1978, *Arctic*, vol. 35, 1982.

Morphine alkaloids

A group of closely related compounds obtained from the dried latex that exudes from the seed capsule of the opium poppy, *Papaver sommiferum*. The major alkaloid is morphine [structure (**1**)], which constitutes about 12% of the weight of the crude opium resin. Some 20 minor alkaloids have been isolated from opium, among them codeine (**1**), thebaine (**2**), and papaverine (**3**). *See* ALKALOID; POPPY.

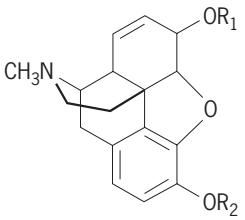

Morphine: $R_1=R_2=H$
Codeine: $R_1=H$, $R_2=CH_3$
Heroin: $R_1=R_2=COCH_3$

(**1**)

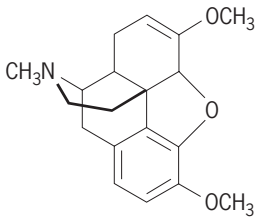

(**2**)

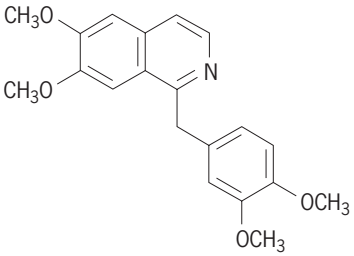

(**3**)

The chemistry of the morphine alkaloids is extremely complex, and the structure of these compounds was not finally established until the early 1950s, when the synthesis of morphine was accomplished. The unique benzomorphane ring system occurs in the plant as a result of oxidative phenol coupling of a benzylisoquinoline, as shown in the **illustration** for the formation of thebaine from reticuline. The discovery of a biosynthetic relationship between the benzylisoquinoline and the morphine ring systems provided a key insight to the structures of these and several other groups of alkaloids. The biosynthetic pathway, beginning with the amino acid tyrosine and proceeding via the steps in the illustration, has been established by the incorporation of isotopic labeling at specific positions in feeding experiments with labeled precursors. The final steps are successive demethylations of thebaine and codeine to the end product morphine. *See* BIOSYNTHESIS.

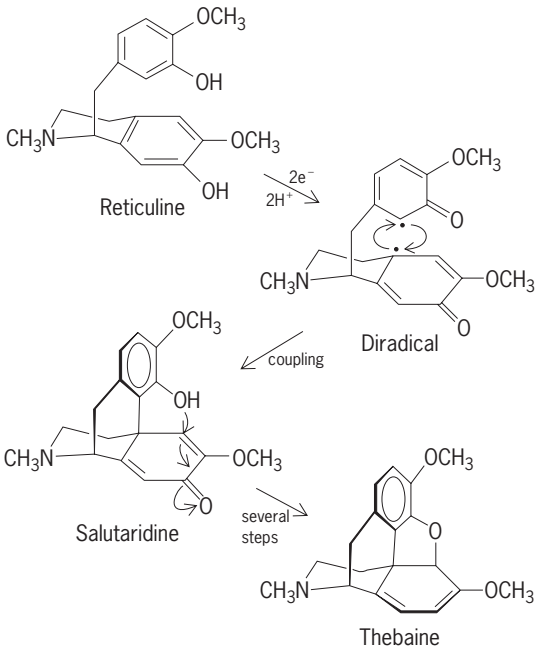

Biosynthesis of thebaine.

Morphine is an analgesic and central nervous system depressant, and is one of the most effective and widely used drugs for the relief of severe or prolonged pain. It is usually administered intramuscularly as the calcium salt. Morphine is also an addicting narcotic; prolonged use leads to dependence and physical addiction.

A number of compounds have been synthesized in an effort to obtain a nonaddicting substitute. Among these are several modifications of the morphine structure, including the simpler systems meperidine (Demerol; **4**) and methadone (**5**). These two compounds have less addiction liability but also have less analgesic potency.

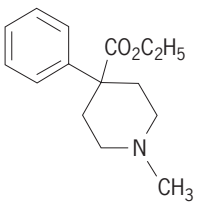

(**4**)

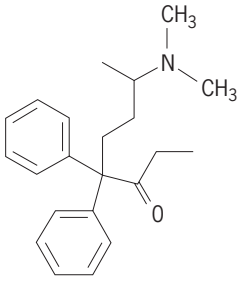

(**5**)

Heroin is prepared readily by acetylation of morphine. This compound is absorbed parenterally much more rapidly than morphine, and the psychological effects and euphoria of opium are more quickly achieved. Heroin is therefore a highly addicting substance, and it causes the most serious problem in illicit opium usage.

Codeine is a somewhat less active analgesic than morphine and is usually used orally in combination with a salicylate such as aspirin. Most of the codeine in medical use is made by methylation of morphine, since the amount present in opium is insufficient. Codeine and derivatives such as dihydro and hydroxy compounds are effective antitussive agents, widely used in cough syrups. *See* ADDICTIVE DISORDERS; ANALGESIC; NARCOTIC; OPIATES. James A. Moore

Bibliography. D. R. Dalton, *The Alkaloids*, 1979; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., 1999; R. N. F. Manske (ed.), *Alkaloids*, vols. 6 and 13, 1960, 1971.

Morphogenesis

The development of form and pattern in animals. Animals have complex shapes and structural patterns that are faithfully reproduced during the embryonic development of each generation. Morphogenesis is a higher outcome of the process of differentiation, defined as the progressive structural and functional diversification and specialization of cells, and recognizable by specific molecular markers and morphological and histological organization. Differentiation is precisely patterned in space to create a three-dimensional functional tissue organization. The process by which this takes place is called pattern formation. Morphogenesis, differentiation, and pattern formation are related: Morphogenesis is accomplished via a complex series of individual and group cell movements called morphogenetic movements. The cell differentiations that lead to these movements are preterminal differentiations that result in coordinated changes in individual cell shape, adhesion, and motility, as well as production of extracellular matrix. The terminal differentiation of cells in the spatial patterns that define tissue relationships within organs and appendages takes place after morphogenetic movements cease, and serves to further refine the shapes generated by these movements. The starting point for these events is a single cell, the fertilized egg, or zygote, that divides repeatedly to form a multicellular embryo capable of carrying out the series of patterned differentiations leading to the morphogenetic movements that shape the embryo. *See* ANIMAL GROWTH; DEVELOPMENTAL BIOLOGY; EMBRYONIC DIFFERENTIATION; MOLECULAR BIOLOGY.

Differentiation. Cell differentiation involves the differential expression of genes in the nuclear DNA that encode proteins specifying the structure and function of each cell type. During cleavage, nuclei divide equivalently so that all cells of the embryo receive the total complement of genes contained in the DNA of the zygote nucleus (**Fig. 1**). However, cells in different regions of the embryos of many animals contain cytoplasm which differs in composition. The cytoplasmic composition of a cell controls which genes will be expressed in each region of the embryo, resulting in a patterned differentiation. Regional cytoplasmic differences are generated during embryonic development by several means. In most animals, regional differentiation of the egg's cytoplasm takes place during its development in the ovary and/or after fertilization. Cleavage segregates these regional differences into separate cells. These differences produce a small number of cell types during cleavage of the zygote. In other animals, mammals being a prime example, the cytoplasm of the egg appears to be homogeneous. Initial differences in cells arise because of the positions they occupy: cells on the outside of the early mammalian embryo become trophoblast cells, whereas cells on the inside become the inner cell mass from which the embryo develops. In all developing multicellular organisms, interactions between these early cell types result in new patterns of gene activity that further increase the diversity of cell types. By means of cascades of such interactions, all of the different cell types of an animal body gradually emerge in the proper spatial patterns. *See* CELL DIFFERENTIATION; CLEAVAGE (EMBRYOLOGY); DEOXYRIBONUCLEIC ACID (DNA); GENE ACTION.

Axial polarity. The structural patterns of animals and their parts exhibit polarity; that is, they display structural differences along anteroposterior (AP), dorsoventral (DV), and right-left axes. In many animals, the axial polarities of the embryo are established during oogenesis and/or the period between fertilization and the first cleavage, by the localization of maternal determinants into specific parts of the egg. For example, in the nematode worm *Caenorhabditis elegans*, the sperm nucleus moves from its locus of entry to the nearest end of the oblong egg. This end of the egg becomes the posterior pole of the AP axis, and the opposite end becomes the anterior pole. Microtubules organized by the sperm centriole act as tracks to position partitioning (PAR) proteins at the anterior and posterior poles of the AP axis. PAR 1 and 3 localize to the posterior pole, and PAR 2 localizes to the anterior pole. The PAR proteins act to localize maternal differentiation factors into different regions of the egg, where

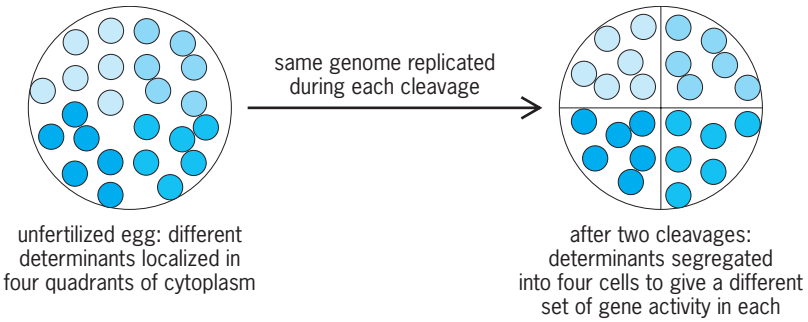

Fig. 1. General concept of how localized maternal determinants of the unfertilized egg give rise during cleavage to differential gene activity in different cells. Left, four types of determinants are each localized to a different quadrant of the egg. During cleavage, the same genome is copied to each cell. Right, two cleavages have partitioned each set of determinants into each of four cells, where they activate different sets of genes.

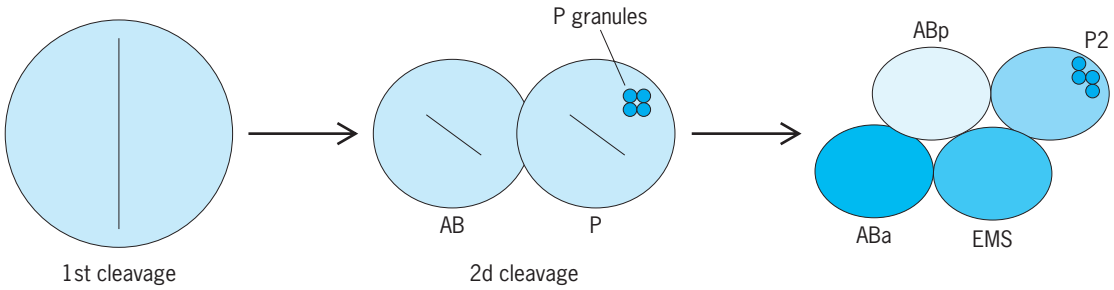

Fig. 2. Cleavage of the nematode worm *Caenorhabditis elegans*. Lines indicate cleavage planes. The first cleavage divides the egg into anterior and posterior blastomeres reflecting the anterior-posterior axial polarity of the fertilized egg. The posterior blastomere contains polar granules that will eventually become germ cells. The second cleavage is at a 45° angle to the first one, creating two dorsal and two ventral cells offset from one another, and establishing the dorsal-ventral axis of the worm.

they will be separated by cleavage. One set of such factors is the P granules, which become localized at the posterior pole. The first cleavage divides the egg into an anterior AB blastomere and a posterior P blastomere, which contains the P granules (**Fig. 2**). Those descendants of the P cells receiving P granules will become germ cells. The remaining descendants of the P cell and those of the AB cell will become the somatic cells of the worm. The DV axis of the nematode is established by the division of the AB and P cells in the plane of the AP axis, but at a 45° angle to that axis, placing two cells with different morphology perpendicular to the AP axis.

The AP and DV axial polarity of the eggs of anamniote and amniote vertebrates is established somewhat differently. The amphibian is the classic example of anamniote development. As the amphibian egg grows in the ovary, one half (the "vegetal" half) becomes laden with yolk, and a dark pigment is deposited in the cortical cytoplasm of the opposite half (the "animal" half) to form an animal-vegetal axis around which the egg is radially symmetrical. The cells derived from the animal half during cleavage will become the ectoderm and mesoderm of the embryo, and the cells derived from the vegetal half will become the endoderm. The animal-vegetal axis approximates the AP axis of the embryo. During oogenesis, the signaling factor Vg1, the receptor for the Wnt signaling pathway (Disheveled) and the transcription factor VegT become localized to the cortical cytoplasm of the vegetal half and will be partitioned to the endodermal cells during cleavage. When the egg is fertilized and shed into the water, it orients with the heavier yolky half down. *See* FERTILIZATION; OOGENESIS; OVUM.

The DV axis of the amphibian embryo is established by a reorganization of the zygote cytoplasm initiated by the events of fertilization (**Fig. 3**). This axis can form in any of the planes that contain the animal-vegetal axis. The plane in which the DV axis actually forms is determined by the meridian at which the sperm enters the egg. This meridian is coincident with the ventral midline of the embryo; the opposite meridian coincides with the dorsal midline.

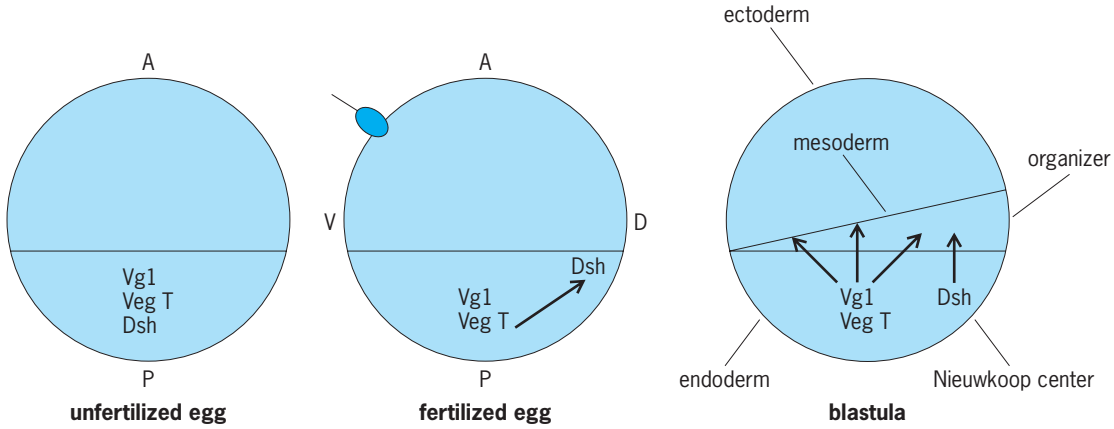

Fig. 3. Summary of axis formation and mesoderm induction during frog development. The unfertilized egg: the maternal determinants vegetal 1 (Vg1), vegetal T (Veg T), and Disheveled (Dsh) are localized to the vegetal cytoplasm of the egg. Fertilization establishes the dorsal-ventral axis. The side of sperm entry becomes the ventral (belly) side of the embryo, and the opposite side becomes the dorsal (back) side. Dorsality is conferred by a 30° upward rotation of the cortical cytoplasm that carries Dsh with it. Blastula stage formed by cleavage. Dsh is contained in the endodermal cells of a signaling center called the Nieuwkoop center and is part of a cascade of gene activation that induces the overlying ectoderm cells as the dorsal mesoderm of the embryo, while Vg1 and Veg T, localized into other endodermal cells, induce other parts of the mesoderm along the dorsal-ventral axis. The three germ layers—ectoderm, mesoderm, and endoderm—are now ready for the morphogenetic movements of gastrulation.

The establishment of the dorsal pole of the DV axis is associated with the upward contraction of the pigmented cortical cytoplasm toward the animal pole (Fig. 3). This contraction is asymmetrical, being greatest dorsally and least ventrally, thus creating a crescent-shaped band of lightly pigmented cytoplasm around the egg, called the gray crescent, the broadest part of which coincides with the dorsal midline. DV polarity is at first labile because it can be wiped out and reestablished in a new plane by tilting the AP axis of the zygote 90° off the vertical so that its ventral side is up. This procedure mimics the cytoplasmic reorganization which occurs after fertilization. A tongue of yolky cytoplasm is formed on the elevated side as the heavy yolk flows down under the influence of gravity, and a new gray crescent appears on this side, its dorsal midline coinciding with the plane in which the AP axis was tilted. The DV axis established by sperm entry becomes determined shortly before the first cleavage, but tilting the egg at this time can still lead to the establishment of a second DV axis. When this is done, a double embryo develops.

Regardless of how it is accomplished or its geographical location, the cortical rotation localizes dorsal maternal determinants to the dorsal pole of the DV axis (Fig. 3). These determinants will be incorporated during cleavage into the dorsalmost blastomeres of the endoderm, a region of the blastula called the Nieuwkoop center. This center plays an important role in inducing formation of the mesoderm, the dorsalmost region of which is called Spemann's organizer because it subsequently determines the AP and DV axial pattern of the embryo. The major determinant that is localized by the cortical rotation into the region that will become the Nieuwkoop center is Disheveled. Disheveled protects the transcription factor β-catenin from being degraded in this region, allowing it to act in combination with another transcription factor, Tcf3, to activate genes critical to the function of the Nieuwkoop center.

In birds, the AP axis is determined during early cleavage stages by gravity. The early embryo of the bird is a disc that lies atop a large mass of yolk. The egg is fertilized, and the cytoplasm on top of the yolk begins to divide as the egg passes down the oviduct. As it moves downward, the egg spins slowly, shifting the yolk so that one side of the radially symmetrical blastoderm is tilted upward. This side becomes the posterior pole of the AP axis. It is the equivalent of the Nieuwkoop center and expresses β-catenin. The equivalent of Spemann's organizer lies just anterior to the Nieuwkoop center. The AP axis in mammals also appears to be determined by gravity. Microgravity environments, such as space, interface with both early avian and early mammalian development. The DV axis of both birds and mammals appears to be determined by a pH and membrane potential difference above and below the blastodisc. The albumin above the blastodisc is basic, and the fluid of the subgerminal cavity below the blastodisc is acidic. The blastodisc pumps water and sodium ions from the albumin into the subgerminal cavity, establishing a potential difference of ~25 mV (positive on the subgerminal cavity side) across the epiblast. The negative basic side of the blastodisc becomes dorsal and the opposite side ventral.

Left-right asymmetry in both invertebrates and vertebrates appears to be the result of a mechanism in which the Notch signal transduction pathway is activated on the left side of the embryo, leading to the expression of a Nodal-related gene in the mesoderm-forming cells on that side. Nodal is a signaling molecule that activates the gene for the transcription factor Pitx2, which specifies the direction of intestinal looping and heart development. In *C. elegans* and the frog, the asymmetric expression of the Nodal-related gene appears to be determined at fertilization by the localization of maternal determinants. In the frog, the determinant may be Vg1, which is activated to a higher degree on the left side of the embryo. In amniotes like birds and mammals, the Notch pathway and *nodal* expression are activated on the left side during gastrulation. The trigger for the asymmetric activation of Notch in the chick appears to be a transient accumulation of extracellular calcium resulting from membrane potentials set up by differential activity of ion pumps.

Determination of pattern. Once axial polarities are established, the pattern of structural differentiation along those axes is elaborated. In *C. elegans*, patterning relies heavily on further segregation of maternal determinants as cleavage proceeds. However, interactions between different cells are also required. For example, if the AB and P cells are isolated from one another, the P cell makes all of the cells it would normally make, but the AB cell makes only some of the cells it would normally make, indicating that interaction with the P cell is required for the AB cell to make a full complement of cell types. Gastrulation begins at the 24-cell stage, when the two cells that will form the endoderm migrate into the center of the embryo, creating a small blastopore. The cells that will form the germ line, the prospective mesenchymal and muscle cells, and the prospective pharyngeal cells migrate in succession through the blastopore into the interior of the embryo, where they take up their appropriate positions relative to one another and their descendants differentiate into the various parts of the worm.

In vertebrates, patterning relies mainly on cell interactions. The first step in patterning is to induce the formation of the mesoderm. At the mid-blastula stage of the frog embryo, when the embryo consists of several thousand cells, an inductive interaction takes place between the endoderm cells and the cells derived from the gray crescent region of the egg, causing the latter to become mesoderm cells. In the endoderm, VegT, Vg1, and β-catenin collaborate to activate the expression of Nodal-related genes. Nodal-related proteins are expressed in a gradient

that is highest dorsally in the Nieuwkoop center and lowest ventrally. β-Catenin and Tcf3 also activate the *siamois* gene in the Nieuwkoop center. Together, the Nodal-related and siamois proteins induce the dorsal mesoderm of Spemann's organizer by activating the *goosecoid* gene. Goosecoid is a transcription factor that activates genes, conferring on the organizer the ability to organize the AP and DV pattern of the mesoderm and initiate the morphogenetic movements of gastrulation. The cells of the organizer form the circular blastopore of the gastrula, through which the prospective endoderm, mesoderm, and endoderm migrate to form three concentric layers, with ectoderm on the outside, endoderm on the inside, and mesoderm in between. The ectoderm develops into the epidermis and nervous system, the endoderm into the organs of the alimentary tract, and the mesoderm into muscles, skeleton, heart, kidneys, and connective tissue. *See* BLASTULATION; GASTRULATION; GERM LAYERS.

Prior to gastrulation, the prospective organ regions of the ectoderm and endoderm are not yet determined, as shown by the fact that they are unable to differentiate in isolation, and that they can develop in ways different from their normal fate when transplanted elsewhere in the embryo. Mesodermal organ regions, however, are highly self-organizing under these conditions. Ectodermal and endodermal organ regions become determined during and after gastrulation by the inductive action of the mesoderm. For example, dorsal mesoderm normally invaginates and stretches out along the dorsal midline where it differentiates as notochord and trunk muscles. The ectoderm overlying the dorsal mesoderm differentiates as the central nervous system. When dorsal mesoderm is transplanted to the flank of an early gastrula, it invaginates and differentiates according to its normal fate. The overlying flank ectoderm, which normally would become skin epidermis, forms a secondary nervous system, indicating that the nervous system is normally determined by dorsal mesodermal factors acting on the overlying ectoderm. Surprisingly, this is not a stimulatory action but an inhibitory one in which the neural default state of the ectoderm is preserved by several proteins inhibitory to the action of bone morphogenetic proteins that induce ectoderm elsewhere to become epidermis. The dorsal mesoderm of the gastrulating embryo also specifies the AP and DV pattern of the neural tube by the region-specific expression of other inducing proteins.

In amniote embryos, such as birds and mammals, the Nieuwkoop center functions in the same way to induce an organizer. In birds, the blastula is a flat disc with an upper layer consisting mostly of prospective ectoderm with prospective mesoderm at its posterior pole, and a lower layer of prospective endoderm. The blastopore in amniote embryos is an elongated structure called the primitive streak, but it functions in the same way as the circular amphibian blastopore to create a three-layered gastrula. In this case, prospective mesoderm is internalized between the prospective ectoderm and endoderm by ingression through the primitive streak. While not as thoroughly studied as the frog embryo, it is clear that many of the genes involved in Nieuwkoop center formation, organizer induction, and axial patterning are the same as in the frog. *See* FATE MAPS (EMBRYOLOGY).

Once induced, organ regions of vertebrate embryos can themselves induce additional organs from undetermined tissue. For example, the retina and iris of the eye develop from a vesicle growing out of the forebrain. This vesicle induces a lens from the overlying head ectoderm, and the lens then induces the cornea from head ectoderm to complete the eye. By means of such cascades of inductive interactions, all the organs of the body are blocked out. *See* EMBRYONIC INDUCTION.

Morphogenetic fields and positional information. Once determined, an organ region constitutes a developmental system, called a morphogenetic field, which specifies the detailed pattern of cell differentiation within the organ. Cells differentiate in patterns dictated by their relative positions within the field. It is generally accepted that graded molecular signals, or cues, are the basis of this positional information. It is proposed that the source of the signals is a set of boundary cells which define the limits of the field. All the cells of a field thus derive their positional information from a common set of boundary cells. The signals can be propagated by extracellular diffusion (paracrine action between different cells or autocrine action of a cell on itself), by intercellular contact via specialized connections between cells (called gap junctions), or by interactions between molecules located on the surfaces of adjacent cells (juxtacrine actions). These signals exert their effect by binding to receptors that transduce the signal to the nucleus by one of several signal transduction pathways. The cell integrates these signals to interpret how it should differentiate, and activates transcription factors that stimulate the appropriate gene activity to carry out the differentiation. Among the most important transcription factors involved in axial patterning are the homeodomain proteins encoded by the homeobox (*Hox*) genes. There are four clusters of 13 *Hox* genes each (A, B, C, and D). Within each cluster, the genes are numbered from 1 to13, and they are expressed in that order from anterior to posterior and from dorsal to ventral.

Cells from different regions of a field can be interchanged within the field boundaries, and the pattern they would have formed is respecified to the pattern dictated by the positional information of their new location. For example, the proximal mesoderm of an embryonic chick leg bud normally forms thigh muscles, but will form toes if grafted to the distal end of the bud. Also, if part of the field boundary is removed, it is restored by some of the remaining cells. For example, if the anterior half of the limb field of a salamander embryo is removed, the posterior half regulates to form a whole limb. Some of the posterior cells take on anterior boundary

properties; the whole positional information gradient is then reestablished over the smaller number of cells, and the half pattern is adjusted to a whole. AP regulation of the limb field does not occur in some other embryos such as birds, which have a fixed determination.

Most fields become inactive after the pattern they specify begins to differentiate; the ability to form normal organs after removal or interchange of cells is then lost. However, in some animals, such as salamanders, the fields of certain organs can be reactivated by loss of a part, even in the adult. The missing part is then redeveloped in a process called regeneration. *See* REGENERATION (BIOLOGY).

Cell matrix and adhesivity. Each cell type of an animal is characterized by a specific set of gene-encoded proteins which determine the structure, shape, mobility, and function of the cell. In addition, different kinds of cells secrete molecules of protein and protein complexed with carbohydrate which constitute equally specific types and patterns of extracellular matrix. The matrix stabilizes tissue and organ structure, guides migrating cells to their proper locations, and is a medium through which cell interactions take place. For example, the supporting strength of bone is due to a compact, mineralized matrix of protein-carbohydrate fibers, whereas the matrix of muscle is a nonmineralized, loose network of fibers which can stretch and contract with the muscle cells.

Cell interactions take place at cell surfaces, the molecular composition of which is distinct from one cell type to another, allowing them to recognize one another. These differences are reflected in varying degrees of adhesivity between different kinds of cells, and between cells and different kinds of extracellular matrix. Differential adhesivity is the property upon which cell migration, clustering, and rearrangement is based, and is thus important in creating the conditions for cell–cell and cell–matrix interactions. This is shown by an experiment in which endoderm, mesoderm, and ectoderm cells of amphibian gastrulae were dissociated, then mixed together in a mass. Each of these cell types sorted out from one another so that their normal concentric spatial relationships were restored. In embryonic stages where organs have begun differentiation, the same sorting behavior is observed among cartilage, heart, retina, epidermis, brain, and liver cells. *See* CELLULAR ADHESION.

Differential growth. Another important mechanism for the development of complex patterns and shapes is the differential growth of organs and their parts. Differential growth is evident as soon as embryonic organ regions begin to develop. During the early part of their development, the growth rates of organs are controlled by intrinsic factors. For example, the vertebrate limb region is at first indistinguishable from the surrounding flank, and the cells of the two regions are dividing with the same frequency. Subsequently, the flank cells divide less frequently than the limb cells, resulting in the elevation and outgrowth of the limb region from the body wall. At later stages of development, including postnatal life, organ growth is largely under the control of hormones secreted by cells of the endocrine glands. When growth ceases at maturity, the relative sizes of the body parts constitute distinctive features of individual and species-specific form.

Cell death. Apoptosis, or cell death by suicide, is an important feature in the refinement of final cell number and shape of developing organs and appendages. For example, a mature *C. elegans* worm has exactly 959 somatic cells, but during its development an additional 115 cells formed that were pruned out by apoptosis. Striking examples of the role of apoptosis in developing vertebrates are the elimination of excess neurons in the spinal cord, and foot development in ducks and chickens. Ducks have webbed toes while chickens do not. This difference results because as the chicken leg bud grows and forms the toes, the cells between the developing digits undergo apoptosis. Various grafting and culturing experiments have indicated that it is a cell's relative position in the limb bud which establishes its fate to die at some later time in development. Apoptosis is either triggered by a signal from other cells or is cell-autonomous, but is preventable if a rescue signal is available. Thus within each cell, there are genes that protect against apoptosis and genes that carry out the death sentence. The protective genes belong to the *Bcl-2* family. The genes that destroy the cell are *Apaf1* (apoptotic protease activating factor 1), which activates the proteases caspase 3 and 9. The caspases digest the cell from within.

David L. Stocum

Bibliography. S. F. Gilbert, *Developmental Biology*, 7th ed., Sinauer, 2003; K. Kalthoff, *Analysis of Biological Development*, McGraw-Hill, 1996; C. D. Stern (ed.), *Gastrulation: From Cells to Embryo*, Cold Spring Harbor Laboratory Press, 2004.

Mortar

A binding agent used in construction of clay brick, concrete masonry, and natural stone masonry walls and, to much less extent, landscape pavements. Modern mortars are improved versions of the lime and sand mixtures historically used in building masonry walls. *See* BRICK; MASONRY.

Mortar types are identified by property or performance specifications of the American Society for Testing and Materials. Selection of the correct type is based on the anticipated exposure—interior or exterior—and structural requirements. Type O, for example, is not designed for exterior load-bearing applications, but may be used in non-load-bearing interior positions.

Masonry mortar is composed of one or more cementitious materials, such as masonry cement or portland cement and lime, clean sand, and sufficient water to produce a plastic, workable mixture. Masonry cements are prepackaged blends of hydraulic

cementitious materials (including portland cement or blended cement), plasticizing additives, and sometimes pigments. They require only the addition of sand and water at the job site. Portland cement-lime mortars are proportioned on the job site from bulk or bagged cement and lime, along with sand, water, and any admixtures. Proportions differ greatly; however, a sample mortar might have one part cementitious material and three parts sand, with water quantity sufficient to create a workable mixture. *See* CEMENT; LIME (INDUSTRY).

Mortars are batched in machine mixers, in larger, mobile plants, or at a central mixing plant. The plastic mixture is then spread over masonry units to form horizontal and vertical joints, typically $^3/_8$ in. (1 cm) wide. Joints are finished (tooled) in a vee, concave, flush, or other desired profile.

Mortars are closely related to concrete but, like grout, generally do not contain coarse aggregate. (Mortar and grout differ, however, and are not interchangeable.) Mortars function with the same calcium silicate-based chemistry as concrete and grouts, bonding with masonry units into a contiguous, weatherproof surface in the process. Masonry cement or portland cement-lime mortars can be formulated not only to meet guidelines for types M, S, N, and O, but also to address job-specific requirements including setting time, rate of hardening, water retentivity, and extended workability. *See* CONCRETE; GROUT.

John Melander

Mosaicism

The condition in which more than one genetically distinct population of cells coexists within one individual. The term mosaic is used for an individual composed of two or more cell lines of different genetic or chromosomal constitutions, where the cell lines originate from one zygote. Individuals derived from more than one zygote or from more than two gametes are called chimeras. However, because the etiology of mosaicism cannot always be established, the term mosaic often is used to refer to any kind of genetic multiplicity within an individual. Mosaicism is a common phenomenon in both the plant and animal kingdoms; it may originate at any state in the course of ontogeny, and in any tissue in which cells proliferate. Cells with a particular genetic constitution form a clone, and this may appear as a mosaic spot embedded in a background containing cells of different genotype. The several types of mosaicism include chromosomal mosaicism, mosiacism due to mitotic recombination or gene mutation, and functional mosaicism.

Chromosomal mosaicism. Abnormal segregation of chromosomes in the course of mitosis results in progeny cells with irregular numbers of chromosomes. Chromosome nondisjunction brings about aneuploidy, a condition in which the cell bears more or fewer chromosomes than normal (**Fig. 1**). Such a condition is often lethal to the cell due to genetic imbalance, especially when more than one large autosome is involved. In humans, trisomic mosiacism for several autosomes has been reported, and trisomy 21, that is, mosaicism for Down syndrome, is the most common. Depending on the ratio and distribution of trisomic and normal diploid tissues, the clinical features can range from syndromes characteristic of complete trisomy to a normal phenotype. Mosaicism can also involve germ cells; such individuals, although they may not manifest symptoms of trisomy, can produce aneuloid progeny. *See* CHROMOSOME ABERRATION; DOWN SYNDROME.

Abnormal chromosome segregation can involve not only one or a few chromosomes but also an entire genome. In these instances, diploid/triploid or diploid/polyploid mosaicism develops. This type of mosiacism can occur spontaneously, or it can be induced by drugs. In *Drosophila* and a number of plant species the mitotic poisons colcemid and colchicine can induce mosaics of this type. Diploid/polyploid mosaicism can involve somatic as well as germ cells. *See* POLYPLOIDY.

The elimination of complete chromosomes or chromosome fragments is a frequent source of mosaicism. As a rule, in some organisms (for instance, the gall midge *Wachtiella* and the nematode *Ascaris*) several chromosomes or fragments are eliminated during embryogenesis from precursors of somatic cells but not from germ cells.

Elimination of sex chromosomes is often the source of sex mosaicism. In *Drosophila* XX/XO female/male sex mosaic gynandromorphs develop from XX zygotes through elimination of one of the X chromosomes (here O represents an absent sex chromosome). Gynandromorphs of other insect species (butterflies, moths, and beetles) are also known. Many hymenopterans, such as the wasp *Habrobracon juglandis*, possess a method of sex determination where unfertilized eggs produce haploid males and fertilized eggs produce diploid females. Gynandromorphs in these species are diploid/haploid mosaics, where the female part is diploid and the male is haploid.

In humans, XX/XO mosaics are females who usually possess symptoms of Turner's syndrome, but some have normal, XX-bearing ovarian cells and are fertile. Besides XX/XO mosaics in humans and in other mammals, XX/XY, XY/XO, XXY/XO, XO/XY/XYY, XYY/XO, XXX/XO, and other mosaics have also been reported. Most of them were noticed as a consequence of the abnormal sexual development of the mosaic individual. Such mosaics may have originated after aggregation of two embryos or as a consequence of "double fertilization," where one sperm nucleus fused with the oocyte and another with a polar body nucleus. In humans most of the XX/XY mosaics are hermaphrodites (persons with both ovarian and testicular tissues). *See* HUMAN GENETICS.

In plants, elimination of chromosomes or chromosome segments that code for the synthesis of chlorophyll or other colored substances results in a colorless cell, and the clonal descendants of this cell form a lightly colored or noncolored patch on the leaves,

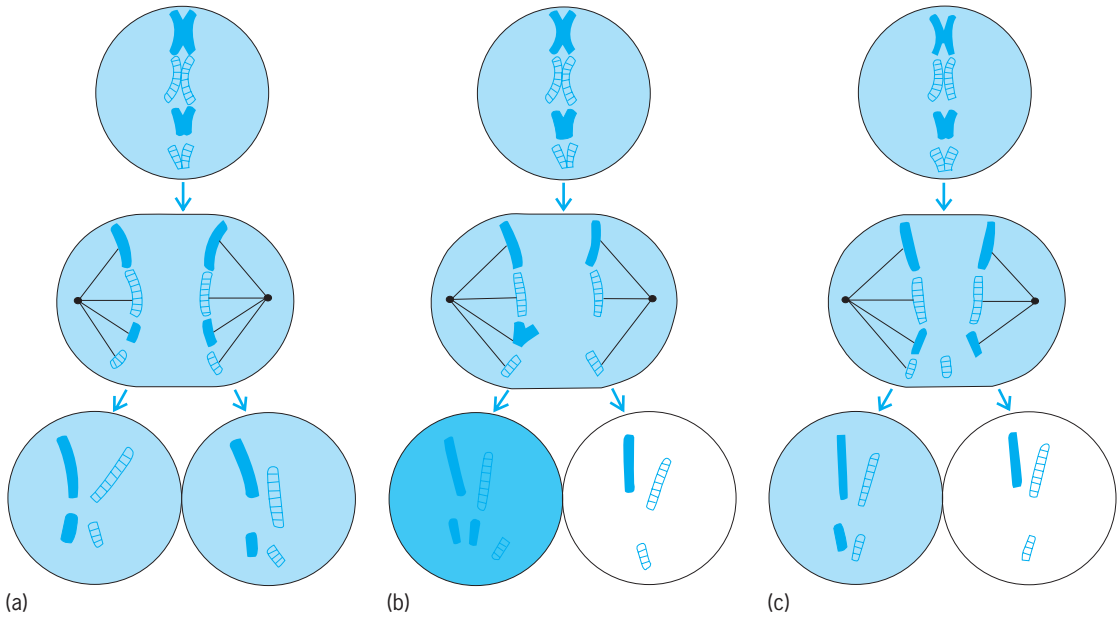

Fig. 1. Chromosomal mosaicism. (*a*) Normal mitosis with two pairs of homologous chromosomes. (*b*) Somatic nondisjunction resulting in aneuploid daughter cells: one bears an additional chromosome (trisomy) and the other lacks one chromosome (monosomy). (*c*) Chromosome elimination leading to monosomy.

flowers, or fruits of the plant. This type of pattern, called variegation, can also occur when normal green and mutant colorless plastids sort out in the course of successive cell divisions. This type of mosaicism is quite distinct in pattern and inheritance from those attributable to any kind of nuclear mutation. It is inherited maternally rather than in mendelian fashion, because plastids are passed to subsequent generations through the female germ line via egg cytoplasm. *See* MATERNAL INFLUENCE.

Loss of chromosomes or chromosome fragments can occur spontaneously (with a low frequency), but it can also be induced by several chemicals, and by x-rays and gamma rays. The effect of radiation is that abnormal chromosomes often form rings or dicentrics which are eliminated during subsequent mitoses.

Mitotic recombination. The exchange of parts of homologous chromosomes often results in cells with a different genetic constitution from that of the mother and the surrounding cells (**Fig. 2**). Progeny of the daughter cells, when properly marked genetically, can be recognized after cell proliferation in both somatic and germ cells. Mitotic recombination (also called somatic crossing-over) is a relatively rare event, but has been known to be the source of spontaneous mosaicism in *Drosophila* and some other insects, in the nematode *Caenorhabditis elegans*, in mice, and in a number of plant species (soybean, tobacco). The frequency of somatic crossing-over can be greatly increased by ultraviolet, x, and gamma rays and also by chemical mutagens and carcinogens (**Fig. 3**). *See* CROSSING-OVER (GENETICS); MUTAGENS AND CARCINOGENS; RECOMBINATION (GENETICS).

Since mitotic recombination can be induced at any time during development, and since it provides a method for the genetic labeling of single cells, it is a useful tool for cell lineage studies. It has been widely employed in *Drosophila* to study gene activity, programmed cell death, cell proliferation, and the development and functioning of different organs.

Gene mutation. In some cells that are heterozygous for mutations, the wild-type alleles may lose their

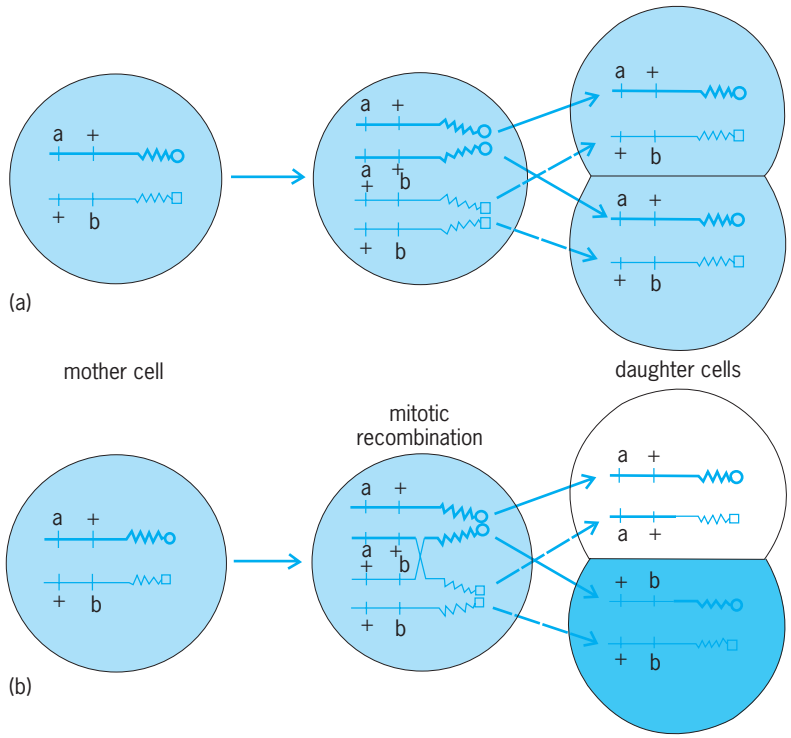

Fig. 2. Mitotic recombination. (*a*) Mitosis produces daughter cells with genotype identical to that of the mother cell. (*b*) After mitotic recombination the daughter cells become homozygous for the marker mutations a and b, and the progeny of these cells form mosaic spots.

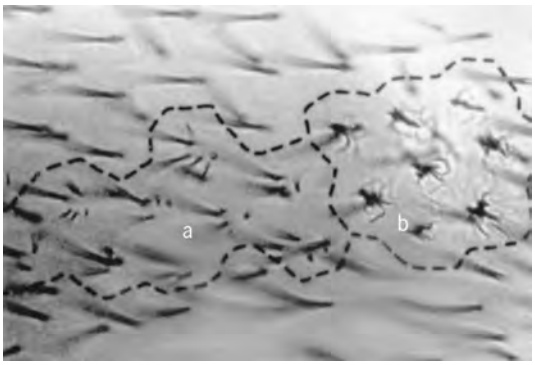

Fig. 3. A twin spot from a wing of *Drosophila melanogaster*, formed after x-ray-induced mitotic recombination. Letters a and b indicate separate clones of cells showing different mutant phenotypes; the background tissue shows the normal phenotype. (*Photo by Andrew W. Wayne*)

activities as a consequence of a new mutation and thus mosaicism can develop. Mutagens and carcinogens are known to induce such mutations and, indeed, these substances can be screened on the basis of this behavior. Activity of the wild-type allele can be suppressed due to nucleotide substitutions, or to deletions or additions of single nucleotides, and also because certain deoxyribonucleic acid sequences, the so-called transposable controlling elements, may become integrated into the gene. Elements of this sort, discovered in maize by B. McClintock in 1950, produce a variegated phenotype. *See* TRANSPOSONS.

Functional mosaicism. Although the genetic constitution is the same (XX) in each of the cells of female eutherian mammals, they show mosaicism for the majority of X-linked genes. This functional mosaicism is due to X-chromosome inactivation. Maternally and paternally derived X chromosomes are turned off randomly during embryogenesis, and consequently recessive X-linked genes can be expressed despite their heterozygous condition. An inactive X chromosome forms a condensed mass of chromatin called a Barr body, after M. L. Barr who first described these bodies in 1949. For example, the calico cat is heterozygous for black and orange alleles of an X-linked coat color gene. Due to X-chromosome inactivation, calico cats exhibit an irregular patchwork of black and orange sectors, with each sector representing a clone of cells derived from one cell. The random inactivation of one of the two X chromosomes in all the somatic cells of females ensures that the somatic cells of males and females generate equivalent levels of X-linked gene products.

Another type of mosaicism develops when genes are juxtaposed to blocks of heterochromatin as a consequence of inversion or translocation of chromosome fragments. These genes may become inactive in some cells but not in others, and thus mosaicism develops. This type of mosaicism is called position effect variegation. *See* GENETICS; SEX-LINKED INHERITANCE.

Janos Szabad

Bibliography. W. Beermann et al. (eds.), *Genetic Mosaics and Cell Differentiation: Results and Problems in Cell Differentiation*, vol. 9, pp. 1–315, 1978; A. Garcia-Bellido, P. A. Lawrence, and G. Morata, Compartments in animal development, *Sci. Amer.*, 24:102–110, 1979; L. B. Russel (ed.), *Genetic Mosaics and Chimeras in Mammals*, vol. 12 of *Basic Life Sciences*, 1978; C. Stern, *Genetic Mosaics and Other Essays*, 1968; R. A. E. Tilney-Bassett, The inheritance and genetic behaviour of plastids, in J. T. O. Kirk and R. A. E. Tilney-Bassett (eds.), *The Plastids*, pp. 251–524, 1978.

Mosquito

Any member of the family Culicidae in the insect order Diptera. Mosquitoes are holometabolous insects and all larval stages are aquatic. Adults are recognized by their long proboscis for piercing and sucking, and characteristic scaled wing venation (**Figs. 1** and **2**).

This is a relatively large group of well-known flies with nearly 3000 species in 34 genera reported in the world. There are 13 genera and 167 recognized species of mosquitoes in North America north of Mexico. Almost 75% of these species belong to three genera: *Aedes* (78 species), *Culex* (29 species), and *Anopheles* (16 species).

Biology. Adult females lay their eggs on or near water: the genera *Culex* and *Culiseta* deposit egg rafts on the surface; *Anopheles* lays eggs singly on the surface; and *Aedes* and *Psorophora* lay batches of eggs on substrates near the water. The latter two genera are called floodwater mosquitoes because their eggs require a period of drying for early embryonation to occur and hatch when rains and flooding cover them with water. Although mosquitoes may be found in nearly all aquatic habitats (ponds, streams, salt lakes, tree holes, artificial containers), the majority use still-water habitats.

Most larvae, or wrigglers, feed on algae and organic debris that they filter from the water with their oral brushes, although certain genera (*Toxorhynchites* and *Psorophora*) may be predaceous and feed on other mosquito larvae. Larvae go through three molts and four instars before pupation. Pupae, or tumblers, are active but nonfeeding stages in which metamorphosis to the adult stage occurs. Both larvae and pupae usually breathe through air tubes at the surface of the water.

Adult male mosquitoes are relatively short-lived, and do not suck blood, but feed primarily on nectar and other plant juices. Females also feed on nectar as their primary energy source, but they require a blood meal for egg production in most species. Some mosquito species are very host-specific, blood-feeding only on humans, birds, mammals, or even reptiles and amphibians, although many species will feed on any available host.

Importance. Mosquitoes are of major importance in both human and veterinary medicine. They can cause severe annoyance and blood loss when they

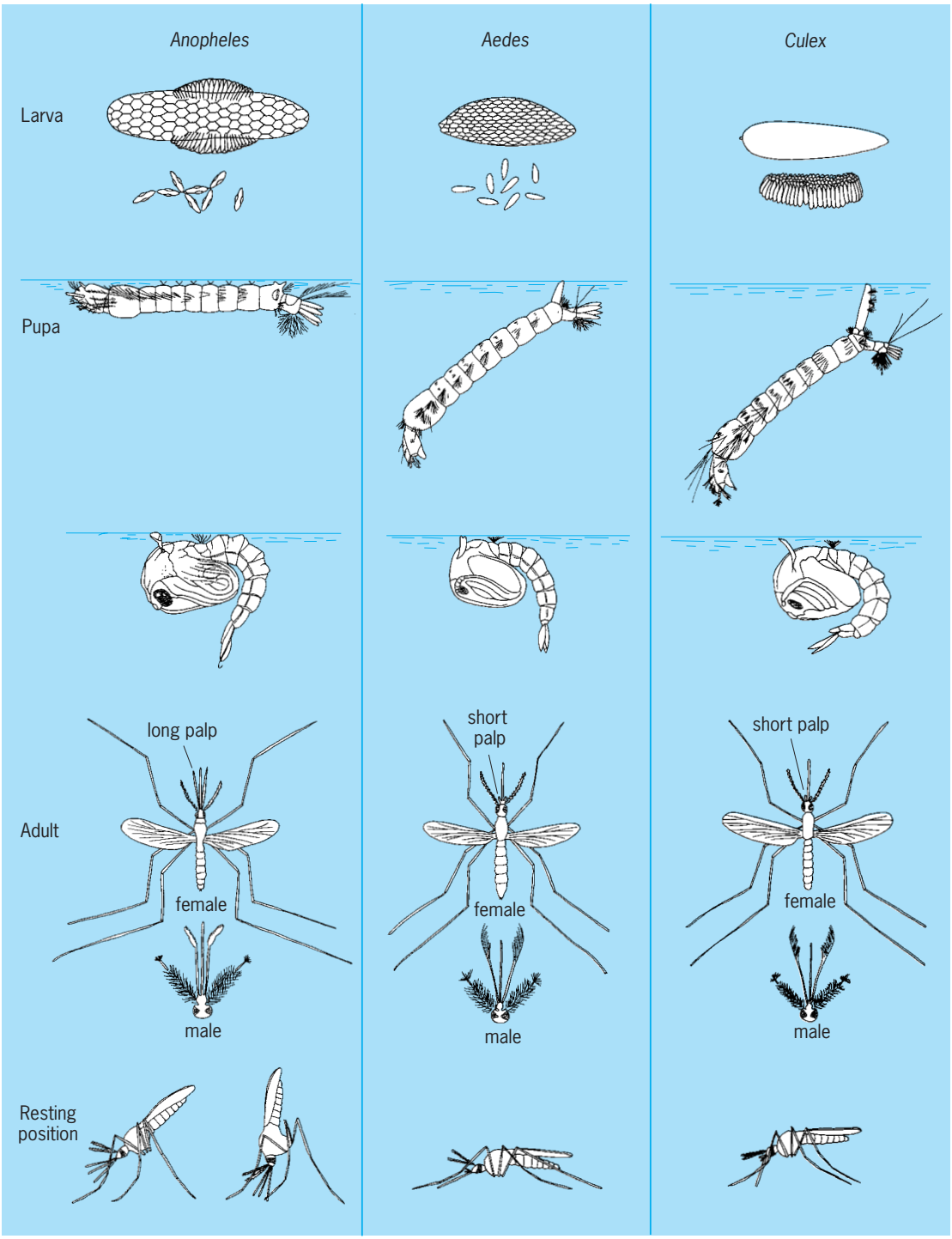

Fig. 1. Morphological and behavioral characteristics of three common genera of mosquitoes. (*After Pictorial Keys to Arthropods, Reptiles, Birds and Mammals of Public Health Significance, PHS Publ. 1955, June 1969*)

occur in dense populations and they act as vectors of three important groups of disease-causing organisms: *Plasmodium*, the protozoan parasite that produces malaria; filarial worms, parasitic nematodes causing elephantiasis in humans and heartworm disease in canines; and arboviruses, which are the causative agents of yellow fever, dengue fever, LaCrosse encephalitis, St. Louis encephalitis, western equine encephalomyelitis, eastern and Venezuelan equine encephalitis, and several other viral diseases. Human malaria is transmitted exclusively by *Anopheles*, filariasis by *Culex*, *Anopheles*, and *Aedes*, and arboviruses primarily by *Culex* and *Aedes* species. *See* ARBOVIRAL ENCEPHALITIDES; HEARTWORMS; MALARIA; YELLOW FEVER.

All of these disease agents enter the mosquito when it feeds on an infected host, and within the vector mosquito the agent undergoes development

Fig. 2. The mosquito, a member of the insect order Diptera.

or multiplication. Following complete development, mosquitoes transmit the agent to susceptible hosts in a subsequent blood feeding; either the parasite is injected with the mosquito saliva (malaria and arboviruses), or the parasite breaks out of the proboscis and crawls into the wound produced during blood feeding (filariasis). *See* ENTOMOLOGY, ECONOMIC; INSECTA; MEDICAL PARASITOLOGY.

Bruce M. Christensen

Bibliography. A. N. Clements, *Biology of Mosquitoes*, I: *Development, Nutrition, and Reproduction*, 1992; R. F. Darsie, Jr., and R. A. Ward, *Identification and Geographical Distribution of the Mosquitoes of North America, North of Mexico*, American Mosquito Control Ass., 1981; K. L. Knight and A. Stone, *A Catalog of the Mosquitoes of the World (Diptera: Culicidae)*, Entomological Society of America, 1977; M. W. Service, *Mosquito Ecology: Field Sampling Methods*, 2d ed., 1993.

Mössbauer effect

Recoil-free gamma-ray resonance absorption. The Mössbauer effect, also called nuclear gamma resonance fluorescence, has become the basis for a type of spectroscopy which has found wide application in nuclear physics, structural and inorganic chemistry, biological sciences, the study of the solid state, and many related areas of science.

Theory of effect. The fundamental physics of this effect involves the transition (decay) of a nucleus from an excited state of energy E_e to a ground state of energy E_g with the emission of a gamma ray of energy E_γ. If the emitting nucleus is free to recoil, so as to conserve momentum, the emitted gamma ray energy is $E_\gamma = (E_e - E_g) - E_r$, where E_r is the recoil energy of the nucleus. The magnitude of E_r is given classically by the relationship $E_r = E_\gamma{}^2/2mc^2$, where m is the mass of the recoiling atom and c is the speed of light. Since E_r is a positive number, the E_γ will always be less than the difference $E_e - E_g$, and if the gamma ray is now absorbed by another nucleus, its energy is insufficient to promote the transition from the nuclear ground state E_g to the excited state E_e.

In 1957 R. L. Mössbauer discovered that if the emitting nucleus is held by strong bonding forces in the lattice of a solid, the whole lattice takes up the recoil energy, and the mass in the recoil energy equation given above becomes the mass of the whole lattice. Since this mass typically corresponds to that of 10^{10} to 10^{20} atoms, the recoil energy is reduced by a factor of 10^{-10} to 10^{-20}, with the important result that $E_r \sim 0$ so that $E_\gamma = E_e - E_g$; that is, the emitted gamma-ray energy is exactly equal to the difference between the nuclear ground-state energy and the excited-state energy. Consequently, absorption of this gamma ray by a nucleus which is also firmly bound to a solid lattice can result in the "pumping" of the absorber nucleus from the ground state to the excited state. The newly excited nucleus remains, on the average, in its upper energy state for a time given by its mean lifetime τ (a quantity dependent on energy, spin, and parity of the nuclear states involved in the deexcitation process) and then falls back to the ground state by reemission of the gamma ray. An important feature of this reemission process is the fact that it is essentially isotropic; that is, it occurs with equal probability in all directions. *See* ENERGY LEVEL (QUANTUM MECHANICS); EXCITED STATE; GAMMA RAYS; GROUND STATE.

Energy modulation. Before this phenomenon of resonance fluorescence can be turned into a spectroscopic technique, it is necessary to provide an appropriate energy modulation of the gamma ray emitted in the initial decay process. An estimate of the energy needed to accomplish this can be calculated from a knowledge of the inherent width or sharpness of the excited-state nuclear level. This is given by the Heisenberg uncertainty principle as $\Gamma = h/2\pi\tau$ (h is Planck's constant and τ is the mean lifetime of the excited state). In the case of ^{57}Fe, a nucleus for which resonance fluorescence is especially easy to observe experimentally, $\Gamma = 4.6 \times 10^{-12}$ keV. In order to modulate the emitted gamma-ray energy, which in this case corresponds to 14.4 keV, one can take advantage of the Doppler phenomenon which states that if a radiation source has a velocity relative to an observer of υ, its energy will be shifted by an amount equal to $E = (\upsilon/c)E_\gamma$. Setting the required Doppler energy equal to the width of the nuclear level and E_γ equal to the nuclear transition energy leads to the equation below. Relative velocities of

$$\upsilon = c\frac{\Gamma}{E_\gamma} = 3 \times 10^{10}\ \text{cm/s} \times \frac{4.6 \times 10^{-12}}{14.4}$$
$$= 0.0096\ \text{cm/s} = 0.0038\ \text{in./s}$$

this order of magnitude can be used to modulate the gamma ray emitted in a typical Mössbauer transition, that is, to "sweep through" the energy width of the

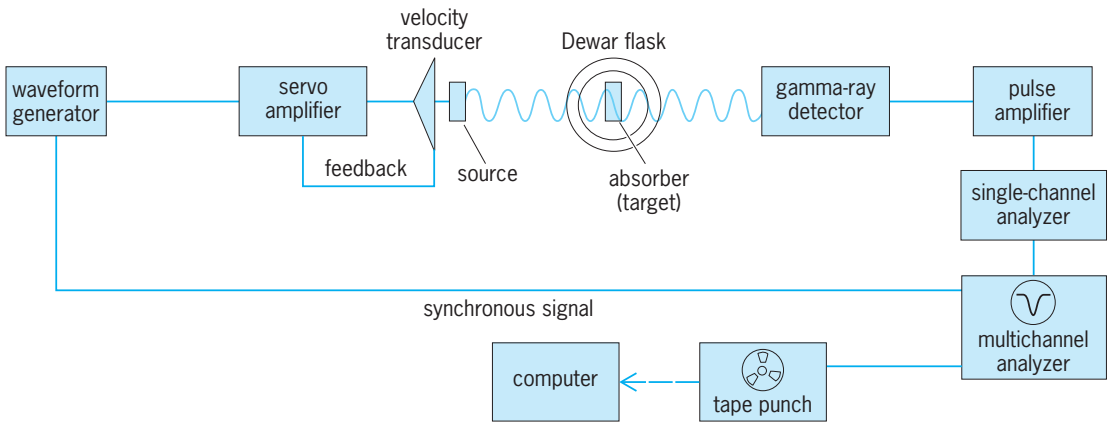

Fig. 1. Experimental arrangement for performing Mössbauer effect spectroscopy. This typical Mössbauer experiment is with ^{57}Fe or ^{119}Sn. (*After R. H. Herber, Mössbauer spectroscopy, Sci. Amer., 255(4):86–95, October, 1971*)

nuclear transition. *See* DOPPLER EFFECT; RADIOACTIVITY; UNCERTAINTY PRINCIPLE.

Experimental realization. The experimental realization of gamma-ray resonance fluorescence can be achieved with the arrangement illustrated schematically in **Fig. 1**. In a typical Mössbauer experiment the radioactive source is mounted on a velocity transducer which imparts a smoothly varying motion (relative to the absorber, which is held stationary), up to a maximum of several centimeters per second, to the source of the gamma rays. These gamma rays are incident on the material to be examined (the absorber). Some of the gamma rays (those for which E_γ is exactly equal to $E_e - E_g$) are absorbed and reemitted in all directions, while the remainder of the gamma rays traverse the absorber and are registered in an appropriate detector which causes one or more pulses to be stored in a multichannel analyzer. The electronics are so arranged that the location (address) in the multichannel analyzer, where the transmitted pulses are stored, is synchronized with the magnitude of the relative motion of source and absorber.

A typical display of a Mössbauer spectrum, which is the result of many repetitive scans through the velocity range of the transducer, is shown in **Fig. 2**. Such a Mössbauer spectrum is characterized by a position δ of the resonance maximum (corresponding to a maximum in the isotropic scattering, and thus a minimum in the intensity of the transmitted radiation), a linewidth Γ, and a resonance effect magnitude ϵ corresponding to the total area A under the resonance curve. *See* GAMMA-RAY DETECTORS.

In the case of the Mössbauer active nuclides ^{57}Fe and ^{119}Sn, among others, two additional features which are of great interest to chemists and physicists may be experimentally elucidated. One of these is the quadrupole coupling which is observed if the Möbauer nuclide is located in an environment where the electric charge distribution does not have cubic (that is, tetrahedral or octahedral) symmetry. Such a spectrum is shown in **Fig. 3**, in which the magnitude of the quadrupole interaction Δ is equal to $e^2qQ/2$, where e is the electron charge, q is the gradient of the electrostatic field at the nucleus, and Q is the nuclear quadrupole moment. Finally, a Mössbauer spectrum can also give information on the magnitude of the magnetic field H_0 acting on the nucleus through the magnetic hyperfine interaction. This is illustrated in **Fig. 4**, where only a single resonance line would be observed in the absence of a magnetic interaction. *See* HYPERFINE STRUCTURE; NUCLEAR MOMENTS.

Moreover, all of these parameters—δ, Δ, Γ, A, and H_0—are temperature-dependent quantities, and their study over a range of temperatures and conditions can shed a great deal of light on the nature of the environment in which the Mössbauer nuclide is located in the sample under investigation. More than one hundred Mössbauer transitions, involving 43 different elements, have been experimentally observed and reported.

Application. Mössbauer effect experiments have been used to elucidate problems in a very wide range of scientific disciplines, and only a few examples can be cited as representative of the information extracted from such studies.

Nuclear physics and chemistry. One of the narrowest resonance lines which has been observed is that from the 6.8-microsecond, 6.2-keV gamma transition in ^{181}Ta, and detailed Mössbauer effect measurements using a source of 140-day ^{181}W have shown that the magnetic moment of the spin-$^9/_2$ excited state

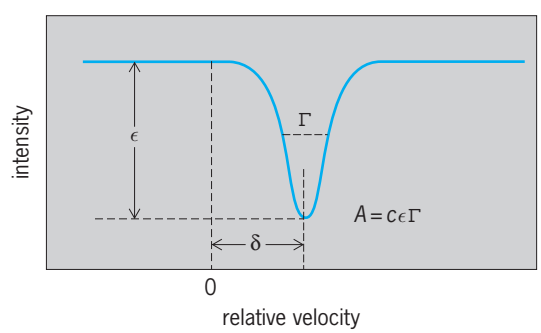

Fig. 2. Mössbauer spectrum of an absorber which gives an upsplit resonance line. The spectrum is characterized by a position δ, a linewidth Γ, and an area A related to the effect magnitude ϵ.

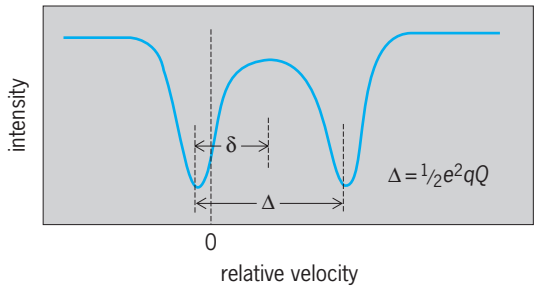

Fig. 3. Mössbauer spectrum of an absorber (containing for example ^{57}Fe or ^{119}Sn) which shows quadrupole splitting Δ.

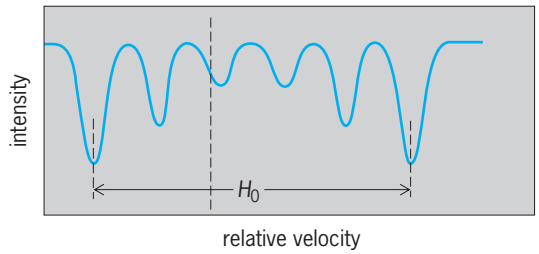

Fig. 4. Mössbauer spectrum of metallic iron showing the splitting of the resonance line by the internal magnetic field (H_0 = 33 T = 330 kG at room temperature).

in ^{181}Ta is $+5.35 \pm 0.09$ nanometers and the nuclear quadrupole moment of this state is $+4.4 \pm 0.05) \times 10^{-24}$ cm^{-2}. Such data are of considerable use to nuclear physicists in refining models which describe the fundamental interaction forces in the nucleus. Similarly, the 93.26-keV resonance in ^{67}Zn has been used to determine the magnetic moment of the 12 state (spin = $^1/_2$, negative parity) in this nuclide and leads to the conclusion that the $^1/_2$ and $^5/_2{}^-$ states can be considered minus-quasiparticle states. Mössbauer spectroscopy has also been a potent technique for the study of the electromagnetic moments of nuclei, in particular the magnetic dipole moment and the (mean-square) charge radius of nuclear states. The nuclides ^{191}Ir and ^{193}Ir are typical of those which have been used for detailed studies of nuclear parameters. Such nuclear information is frequently difficult to obtain by non-Mössbauer-effect methods. *See* NUCLEAR STRUCTURE.

Recoilless gamma-ray resonance experiments have been able to provide detailed information concerning excited-state lifetimes involved in the nuclear decay process. The lifetime values for the nuclides ^{119}Sn, ^{107}Au, and ^{73}Ge, among others, are largely based on Mössbauer effect measurements. *See* NUCLEAR ISOMERISM.

Recoilless gamma fluorescence spectroscopy has also been used to study the chemical consequences of nuclear decay, and the lifetimes of the Mössbauer transition (typically about 10^{-8} s) provide a convenient time scale to distinguish rapid electronic relaxation processes (typically 10^{-12} to 10^{-14} s) from atomic translation processes (typically slower than 10^{-6} s), and thus study the chemical fate of an atom which results from the decay of a radioactive parent nuclide.

Solid-state physics. Mössbauer effect spectroscopy has made significant contributions to the study of problems in solid-state physics, especially of the nature of the magnetic interactions in iron-containing alloys and the dependence of the magnetic field on composition, temperature, pressure, and other parameters which are of importance in metallurgical processes; solid-state-device fabrication; the structural use of metals and alloys; and numerous related problems of great practical importance. *See* FERROMAGNETISM.

Combining Mössbauer effect spectroscopy with vibrational spectroscopic studies has led to a clearer understanding of the nature of inter- and intramolecular forces and the relationship of these forces to the properties of polymeric materials. *See* INTERMOLECULAR FORCES; POLYMER.

It has also been possible, using this technique, to study the effect of high pressure and isotropic compressibility on the chemical properties of materials, especially in the case of experiments with ^{57}Fe, ^{181}Ta, and ^{119}Sn. Such studies have led to the design of high-pressure processes in preparative metallurgy and materials science. *See* HIGH-PRESSURE CHEMISTRY.

A technique which has also found particular application in metallurgy and the study of catalysts is the use of conversion electron spectroscopy to detect the Mössbauer effect. The advantage of this nondestructive technique is that only the surface layers of the material being investigated are probed by the gamma rays. Thus, in conjunction with the standard transmission type of experimental arrangement shown in Fig. 1, it is possible to differentiate between the atoms on the surface of a material and those in the interior. Such studies frequently can lead to a better understanding of the chemical modification of a surface when solid materials are exposed to a reactive environment as in corrosion and in the poisoning of catalysts. *See* SURFACE PHYSICS.

At the extremely low end of the temperature scale, Mössbauer effect spectroscopy has been useful in examining the nature of those materials which become superconductive at sufficiently low temperatures and the relationship between chemical composition and structure on the one hand and the superconductive transition on the other, as in the dichalcogen layer compounds $TaS_2{\cdot}Sn$ and $TaS_2{\cdot}Sn_{1/3}$ (both studied using the nuclide ^{119}Sn). Nb_3Sn, a material widely used in the construction of superconducting magnets, has been subjected to detailed Mössbauer effect investigations. *See* SUPERCONDUCTIVITY.

Structural chemistry. The two Mössbauer nuclides most widely exploited by chemists are ^{57}Fe and ^{119}Sn, although a growing body of data resulting from experiments with ^{129}I, ^{99}Ru, ^{121}Sb, and others has been reported. The position of the resonance maximum δ, also called the isomer or chemical shift, can be related to the systematics of the electron configuration of the atom, and extensive isomer shift correlations for iron- and tin-containing compounds have been tabulated. In particular, the isomer shift of tin compounds is readily related to the oxidation state, since

it has been observed that all stannous (Sn^{2+}) isomer shifts are larger than that observed for metallic tin (β-Sn), while those for stannic compounds (Sn^{4+}) are smaller than this value. This observation allows an assignment of oxidation state to be made on the basis of the isomer shift parameter, as in the two-dimensional layer compound $SnTa_3S_6$ in which the tin atom is clearly identified as a stannous ion, contrary to expectations based on theory. *See* MOLECULAR ISOMERISM; OXIDATION-REDUCTION.

The use of Mössbauer spectroscopy to identify the charge state of a given chemical species has also made major contributions to the understanding of the phenomena of valence fluctuations, valence instabilities, and mixed valencies in solids. Such studies are of particular importance in rare-earth chemistry, and the use of the nuclides ^{149}Sm, ^{152}Sm, ^{153}Eu, and ^{151}Eu has been particularly fruitful. *See* VALENCE.

Similarly, the isomer shifts reported for a number of ruthenium compounds, which have been studied using the 89.36-keV resonance in ^{99}Ru, can be correlated systematically with the number of $4d$ electrons involved in the bonding of the metal atom to its nearest-neighbor ligands. Such Mössbauer effect studies have led to a clearer understanding of the nature of "ruthenium red," a trimeric ammonia ruthenium oxide, and a number of other compounds of this relatively rare transition-metal homolog of iron.

In the field of organometallic chemistry, Mössbauer effect spectroscopy has served to clarify the structure of a number of compounds of iron and tin which are of considerable synthetic and industrial importance, including $Fe_3(CO)_{12}$, $[(\pi C_5H_5)_2Fe(CO)_2]_2$, $[(C_4H_9)_3Sn]_2SO_4$, and the organotin thioglycolates which are used as stabilizers in the plastics industry. *See* ORGANOMETALLIC COMPOUND.

Biological science. Many molecules of biological including hemoproteins, iron-sulfur proteins, and iron storage and transport proteins, offer an ideal system in which Mössbauer effect spectroscopy can be used to elucidate the structure and bonding properties of the metal atom in complex systems. The first measurements on such molecules were reported in 1961, and a very large number of iron-containing systems have been studied since then. Paramagnetic iron compounds can be studied at temperatures below the magnetic ordering point (Néel temperature), and it is thus possible, by means of Mössbauer effect spectroscopy, to determine the sign and magnitude of the magnetic field acting on an iron atom in a complex biological material with molecular weights ranging up to 50,000 or more. *See* PROTEIN.

Success has also been achieved in understanding the magnetic basis by which certain microorganisms can orient themselves in the Earth's magnetic field and thus move in the direction of sediments which contain the nutrients essential to their survival. This magnetic orientation, which is of opposite polarity for organisms living in the Northern and the Southern hemispheres of the Earth, is achieved by the incorporation in the microorganism of particles (magnetosomes) which consist of Fe_3O_4 encased in an organic envelope and which function as a compass in the Earth's magnetic flux lines. The new insights in understanding this phenomenon were achieved in part through the application of Mössbauer spectroscopy involving the isotope ^{57}Fe. *See* MAGNETIC RECEPTION (BIOLOGY).

It is also possible to study antiferromagnetically coupled iron atoms in biological molecules by carrying out the Mössbauer effect measurements in an external magnetic field over a range of temperatures. Typical of such a study is that of oxidized and reduced putidaredoxin, an iron-sulfur protein (molecular weight = 12,500) which acts as a one-electron transfer enzyme. The Mössbauer experiments on this material clearly showed that in the oxidized material the two iron atoms in the molecule occupy chemically equivalent sites. On one electron reduction, one iron atom remains ferric (Fe^{3+}), while the other becomes ferrous (Fe^{2+}), and the two atoms couple antiferro-magnetically to give an electronic ground state of $S = {}^1/_2$. Such detailed knowledge of the chemical behavior of the iron atoms in this molecule can elucidate the action of biological catalysts (enzymes) on the molecular level. *See* ANTIFERROMAGNETISM.

Related fields. Mössbauer effect studies have also played a role in studies in many related fields of science, including archeology, geology, engineering studies, theoretical (relativity) physics, chemical kinetics, and biology. The samples of surface material returned from the Moon by the United States Apollo program have been carefully scrutinized by Mössbauer techniques, as have core samples extracted from deep-drilling experiments on the Earth's outer layer. The geographical distribution of ancient Greek pottery has been traced by making use of characteristic Mössbauer effect data, and the pigments used in painting and decorating have been similarly investigated using this technique.

Rolfe H. Herber

Bibliography. T. E. Cranshaw et al., *Mössbauer Spectroscopy and Its Applications*, 1986; R. S. Drago, *Physical Methods for Chemists*, 2d ed., 1992; D. P. Dickson and F. J. Berry (eds.), *Mössbauer Spectroscopy*, 1987; Yu. M. Kagan and I. S. Lynbutin (eds.), *Applications of the Mössbauer Effect*, 5 vols., 1985; G. J. Long (ed.), *Industrial Applications of the Mössbauer Effect*, 1987; G. J. Long (ed.), *Mössbauer Spectroscopy Applied to Inorganic Chemistry*, 3 vols., 1984, 1987, 1989; G. J. Long and F. Grandjean (eds.), *Mössbauer Spectroscopy Applied to Magnetism and Materials Science*, vol. 1, 1993; S. Mitra, *Applied Mössbauer Spectroscopy: Theory and Practice for Geochemists and Archeologists*, 1993.

Motion

If the position of a material system as measured by a particular observer changes with respect to time, that system is said to be in motion with respect to

the observer. Absolute motion, then, has no significance, and only relative motion may be defined; what one observer measures to be at rest, another observer in a different frame of reference may regard as being in motion. *See* FRAME OF REFERENCE; RELATIVE MOTION.

The time derivatives of the various coordinates used to specify the system may be used to prescribe the motion at any instant of time. How the motion develops in subsequent instants is then determined by the laws of motion. In classical dynamics it is supposed that in principle the motion and configuration of the system may be specified to an arbitrary precision, although in quantum mechanics it is recognized that the measurement of the one disturbs the other.

For a system of f degrees of freedom, the motion may be represented by a point in an f-dimensional velocity space, the coordinates of which are the time rates of change of the coordinates that describe the configuration of the system. For a system under no forces that is described by rectangular cartesian coordinates, these time derivatives are constants. For a single particle, this result is the first of Newton's laws of motion, namely, that a particle remains at rest or in a state of uniform motion unless acted upon by an external force.

The most general theory of motion that has yet been developed is quantum field theory, which combines both quantum mechanics and relativity theory, as well as the experimentally observed fact that elementary particles can be created and annihilated. *See* DEGREE OF FREEDOM (MECHANICS); DYNAMICS; EULER'S EQUATIONS OF MOTION; HAMILTON'S EQUATIONS OF MOTION; HARMONIC MOTION; KINEMATICS; LAGRANGE'S EQUATIONS; NEWTON'S LAWS OF MOTION; OSCILLATION; PERIODIC MOTION; QUANTUM FIELD THEORY; QUANTUM MECHANICS; RECTILINEAR MOTION; RELATIVITY; ROTATIONAL MOTION.

Herbert C. Corben; Bernard Goodman

Motivation

The intentions, desires, goals, and needs that determine human and animal behavior. An inquiry is made into a person's motives in order to explain that person's actions. The answer may have important consequences.

Arousal. Different roles have been assigned to motivational factors in the causation of behavior. Historically, some have defined motivation as a nonspecific energizing of all behavior. Others define it as recruiting and directing behavior, selecting which of many possible actions the organism will perform. Whether there are different states corresponding to different motives or a single state of being motivated is open to question. The likely answer is that both aspects exist. More specific determinants of action may be superimposed on a dimension of activation or arousal that affects a variety of actions nonselectively (**Fig. 1**). The situation determines what the animal does; arousal level affects the vigor, promptness, or persistence with which the animal does it. Many have suggested that behavioral arousal may reflect activity in the reticular formation and related structures in the brainstem, which may alert to prime higher information-processing systems. Finally, there may be an optimal, intermediate level of arousal. One may be too lethargic for effective action if arousal is too low, too frenzied and frantic if it is too high.

Homeostasis and drive. The notion of arousal is not incompatible with the possibility of separate systems for different motives. A specific "hunger drive'' could activate feeding or food-seeking activities, with this effect superimposed upon a more global background of greater or lesser arousal. *See* HUNGER.

Early drive theorists saw motivated behavior as adjunct to physiological mechanisms of homeostasis, that is, the mechanisms by which the body regulates internal variables such as temperature, blood sugar level, and the volume and concentration of body fluids. Faced with dehydration, a land-dwelling animal may slow down its rate of water loss by excreting a more concentrated urine (physiological), but is also becomes disposed to seek and drink water (behavioral). Depleted of energy reserves, an animal may reduce its metabolic rate and so waste fewer calories (physiological). It also becomes disposed to seek and ingest food (behavioral). Thus, motivated behavior forms part of a negative-feedback loop, an arrangement characteristic of regulatory systems.

The physiologist Walter Cannon, the learning theorist Clark Hull, and the clinician Sigmund Freud each proposed variations on homeostasis. As a general theory of motivation, however, the homeostatic model faces difficulties. First, not all "basic biological drives" work this way. For example, there is no evidence of any physiological imbalance that builds up in the absence of sexual activity and corrected by it. In particular, the level of sex hormones in the blood is not reduced by sexual activity and may even be increased by it.

Second, motivated behavior can be influenced by external as well as internal factors. The smell of broiling steak may stimulate hunger, and words and pictures are used to promote sexual arousal. In the rat, provision of specially attractive food can trigger a bout of feeding in the short term, and even produce dietary obesity in the long term; and a male rat's flagging sexual interest may be revived if a new partner is offered. Since these external influences are not coupled with the animal's internal state, they can lead to behavior that does not promote homeostasis and may even threaten it, as in dietary obesity.

To add to the complexity, internal and external factors are not independent and additive; rather they interact with each other. A sodium-deficient rat (internal) becomes hyperresponsive to the taste of salt (external) even if the ingested salt does not relieve the deficiency. The urine of a female guinea pig contains odorous substances (external) that evoke intense interest in intact males but not in castrated males (internal). In such cases, internal influences

affect behavior by setting the animal's responsiveness to certain external signals. The interaction occurs in the opposite direction as well: external signals can affect internal state. Thus, in ring doves the mere sight of a male's courtship display (external) can affect ovarian secretions in the female (internal).

Third, especially in humans, vigorous and persistent goal-directed behavior can occur in the absence of any physiological need. For example, one may strive for success at a task or in one's profession, or to be a good provider, a good student, or a good homemaker. The attempt to explain these motives as "derived from" homeostatic ones has lost its appeal. *See* HOMEOSTASIS.

The brain. These multiple influences must be registered and translated into behavior by the nervous system. It has been suggested that various neuron clusters in the hypothalamus could be command centers, responsible for the excitation or inhibition of the various drive states.

The theory is no longer tenable in its original form. Expression of any action involves systems, not centers, in which multiple levels of the brain cooperate. It remains true, however, that motivational sequences can be initiated from within the hypothalamus. Thus microinjections of norepinephrine into the hypothalamus can trigger hunger in otherwise satiated rats. In rodents (but not in humans), female sexual behavior is promptly and totally abolished by spaying; but it can be restored by implants of tiny amounts of hormones into the hypothalamus. Finally, the more general idea that motivated behavior reflects a balance of excitatory and inhibitory influences is widely accepted.

Cognition. There is another reason for the homeostatic drive concept's loss of popularity. Even relatively simple motives can be influenced by much more than the existing internal and external situation. They respond to potential or expected factors, as registered by cognitive apparatus.

Drives, anticipations, and representations. Even relatively simple motives such as hunger and thirst are responsive to cognitive factors. A rat may learn how much water it is going to need, and adjust its intake to anticipated rather than present requirements. It can learn which of two flavored fluids is going to supply a large caloric load, and drink less of that fluid in anticipation of this. In humans, an image or fantasy of an attractive conspecific can produce the experience, as well the physical signs, of sexual excitement. Such behaviors imply that an animal is capable of representing anticipated or potential states of affairs, as well as perceiving existing ones. The forward-looking or goal-directed nature of behavior may depend on such capabilities. *See* THIRST AND SODIUM APPETITE.

Representations and goal-directed behavior. To a hungry rat, food becomes a goal. The rat will make various responses, including arbitrary learned ones or operants, that lead to contact with food. A rat can be trained to press a lever for food reinforcement, but it could just as easily be trained to run down an alley or dance around the cage or do whatever else is necessary (within its capabilities) to attain its goal. It is this flexibility of goal-directed behavior that justifies the concept of motivation. If an animal will do whatever is necessary to obtain food, it must want food. Internal factors then may act by setting the goal status of environmental commodities: the effect of hunger is to make food a goal.

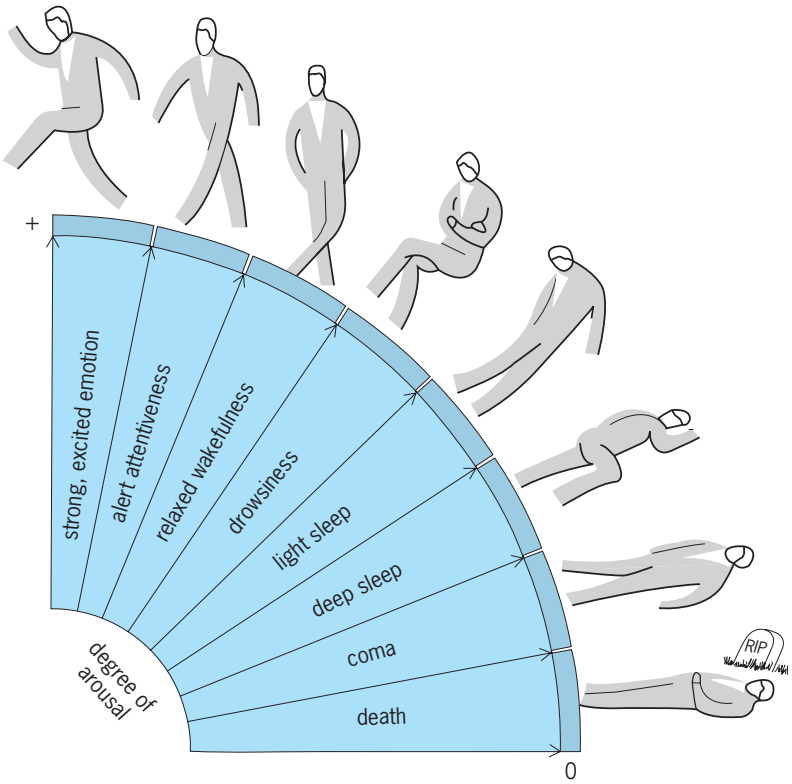

Fig. 1. Diagram of the continuum of arousal.

There is a question as to how behavior can be guided by a state or event (goal attainment) that does not yet exist. Modern approaches to this question lean heavily on cognitive concepts. Mammals, birds, and even some insects can represent to themselves a nonexistent state of affairs. They can represent what a goal object is (search images): a chimpanzee may show behavioral signs of surprise if a different food is substituted for the usual one. They can represent where it is (cognitive maps): a digger wasp remembers the location of its nest relative to arbitrary landmarks, and will fly to the wrong place if the landmarks are moved.

If this idea is generalized, motivated behavior can be thought of as guided by a feedback control system with a set point (**Fig. 2**). A set point, such as a thermostat setting, establishes a goal state which the control system seeks to bring about. Behavior is controlled, not by present external or internal stimuli alone, but by a comparison between the existing state of affairs (registered by the sensory system) and a desired state of affairs, that is, the set point or goal, registered or specified within the brain. The animal then acts to reduce the difference between the existing and the desired state of affairs.

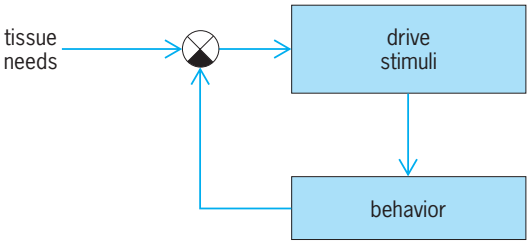

Fig. 2. Motivation as part of a negative-feedback system. Behavior is motivated by drives, which it acts to reduce or remove.

Complex representations in humans. This way of looking at motivation helps bridge the gap between simple motives in animals and complex ones in humans. If to be motivated is to do whatever is necessary to bring about an imagined state of affairs, then human motives can literally be as complex, and be projected as far into the future, as human imaginations permit. Individuals may seek such abstract goals as equitable return for effort, or personal power, or achievement. *See* COGNITION.

Instinctive behavior. Another approach to motivation comes from the work of ethologists, that is, students of behavior in its natural environment. Historically, ethologists tended to emphasize the species-typical, or instinctive, repertoires of their subjects. These were often described as stereotyped action patterns, responsive not to their consequences (like operants) but to their antecedents: sign or releasing stimuli.

However, old distinctions have broken down, and ethology has formed links with cognitive psychology. The broken-wing display of the piping plover provides an example. If a predator approaches a nest with eggs, the parent bird may behave as if injured (hence easy prey) and thus lead the intruder away from the nest. This action pattern is characteristic of the species and unlearned in its gross topography; yet the bird monitors the intruder's behavior and modulates the display accordingly. It may approach more closely and intensify the display if the intruder is not at first diverted from its path. Thus a species-typical action pattern can be used in ways suggestive of purpose and goal direction: the bird modifies it as necessary to promote the goal of diverting the intruder. *See* ETHOLOGY.

Emotion. Motivation and emotion are closely related. Indeed, it has been argued that emotions are the true motivators and that other factors internal, situational, and cognitive take hold of behavior by way of the emotions they evoke. In the simplest case, pleasure and displeasure have been recognized for centuries as having motivational force. If a nondeprived rat eats substantially when offered palatable food, this may simply mean that the food tastes good. In more complex cases, the role of cognitive operations, such as how an individual feels about an event, as well as what is done about it, can depend heavily on how an individual thinks about it.

A further suggestion is that different persons may develop different explanatory styles by which to make causal attributions. A person may habitually explain unfortunate events as resulting from his or her own (internal, stable, uncontrollable) inadequacies, whereas good events may be explained as resulting from luck or chance. Such a pessimistic and self-denigrating explanatory style may play a role in generating the emotional and behavioral symptoms of depression.

Culture. The culture in which an individual is raised has a powerful effect on how the individual behaves. It has been argued that culture teaches its members what to believe are the consequences of a specific action (cognitive), and how the individuals should feel about those consequences or about the actions themselves (emotional/motivational). Douglas G. Mook

Bibliography. R. D'Andrade and C. Strauss (eds.), *Human Motivation and Cultural Models*, 1993; D. G. Mook, *Motivation: The Organization of Action*, 2d ed., 1997; D. Pfaff (ed.), *Physiological Mechanisms of Motivation*, 1982; C. A. Ristau (ed.), *Cognitive Ethology*, 1991; B. Weiner, *Human Motivation*, 1989.

Motor

A machine that converts electrical into mechanical energy. Motors that develop rotational mechanical motion are most common, but linear motors are also used. A rotary motor delivers mechanical power by means of a rotating shaft extending from one or both ends of its enclosure. Such a motor is shown in **Fig. 1**, cut away to show the internal parts. The shaft is attached internally to the rotor. Shaft bearings permit the rotor to turn freely. The rotor is mounted coaxially with the stationary part, or stator, of the motor. The small space between the rotor and stator is called the air gap, even though fluids other than air may fill this gap in certain applications. **Figure 2** illustrates the relationship between the rotor, stator, air gap, and shaft.

Linear motors may have a stationary part that forms an elongated track, along which the movable part carries loads. Alternatively, the stator may be the shorter of the two members, and the movable part may act as a continuous conveyor. Linear motors operate on the same physical principles as rotary alternating-current motors.

In a motor, practically all of the electromechanical energy conversion takes place in the air gap. Although a few electrostatic motors have been built, they require such prohibitively high voltages that all commercial motors employ magnetic fields as the energy link between the electrical input and the mechanical output. The air-gap magnetic field is set up by current-carrying windings located in the rotor or the stator, or by a combination of windings and permanent magnets. The magnetic field exerts forces between the rotor and stator to produce the mechanical shaft torque; at the same time, in accord with Faraday's law, the magnetic field induces voltages in the windings. The voltage induced in the winding connected to the electrical energy source is often called

a countervoltage because it is in opposition to the source voltage. By its magnitude and, in the case of alternating-current (ac) motors, its phase angle, the countervoltage controls the flow of current into the motor's electrical terminals and hence the electrical power input. The physical phenomena underlying motor operation are such that the power input is adjusted automatically to meet the requirements of the mechanical load on the shaft. *See* ELECTROMAGNETIC INDUCTION; MAGNET; MAGNETISM; WINDINGS IN ELECTRIC MACHINERY.

Construction. Both the rotor and stator have a cylindrical core of ferromagnetic material, usually steel. The parts of the core that are subjected to alternating magnetic flux are built up of thin sheet steel laminations that are electrically insulated from each other to impede the flow of eddy currents, which would otherwise greatly reduce motor efficiency. The windings consist of coils of insulated copper or aluminum wire or, in some cases, heavy, rigid insulated conductors. The coils may be placed around pole pieces, called salient poles, projecting into the air gap from one of the cores, or they may be embedded in radial slots cut into the core surface facing the air gap. In a slotted core, the core material remaining between the slots is in the form of teeth, which should not be confused with magnetic poles. A disassembled universal motor, in **Fig. 3**, illustrates several of these features. *See* EDDY CURRENT.

Direct-current (dc) motors usually have salient poles on the stator and slotted rotors. Polyphase ac synchronous motors may have salient poles on the rotor and slotted stators. Rotors and stators are both slotted in induction motors. Permanent magnets may be inserted into salient pole pieces, or they may be cemented to the core surface to form the salient poles.

The windings and permanent magnets produce magnetic poles on the rotor and stator surfaces facing each other across the air gap. If a motor is to develop torque, the number of rotor poles must equal the number of stator poles, and this number must be even because the poles on either member must alternate in polarity (north, south, north, south) circularly around the air gap. If there are four magnetic poles on the rotor, there are also four on the stator, and the motor is said to be a four-pole motor. A high-speed motor has fewer poles than a low-speed motor.

Classification of rotary motors. Motors are classified according to power rating as subfractional horsepower, fractional horsepower, or integral horsepower. They are also classified by application as general purpose, definite purpose, and special purpose. In the United States, definitions of these terms and detailed standards for dimensions and operating characteristics are set by the National Electrical Manufacturers Association and the Small Motors Manufacturers Association. The **table** classifies electric motors in common use according to operating principles; it does not include control motors.

Alternating-current motors. Electric motors operating directly from polyphase (usually three-phase) lines include synchronous motors and induction motors. The speed of synchronous motors is determined precisely by the supply frequency, the speed in revolutions per minute (rpm) being given by $120f/p$, where

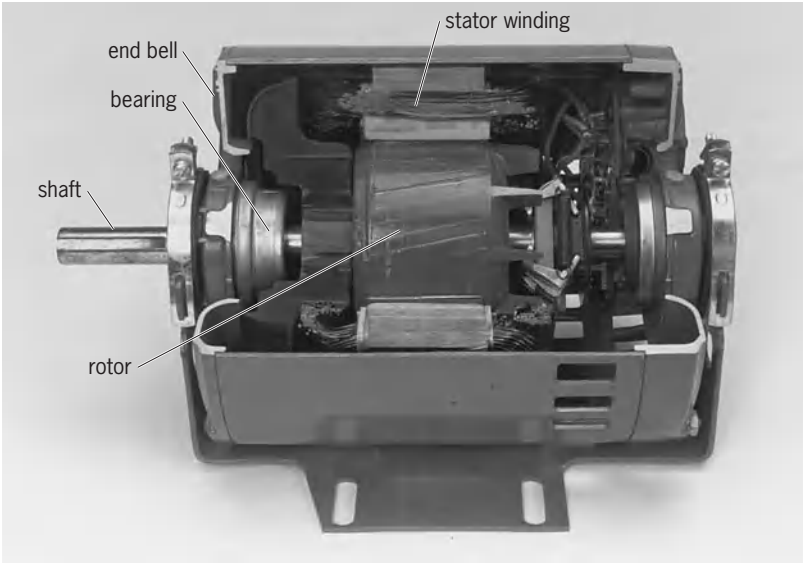

Fig. 1. Cutaway view of a single-phase induction motor. (*Emerson Motor Division*)

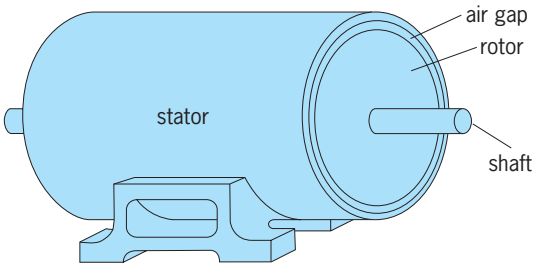

Fig. 2. Elements of an electric motor. (*After G. McPherson and R. D. Laramore, An Introduction to Electrical Machines and Transformers, 2d ed., John Wiley and Sons, 1989*)

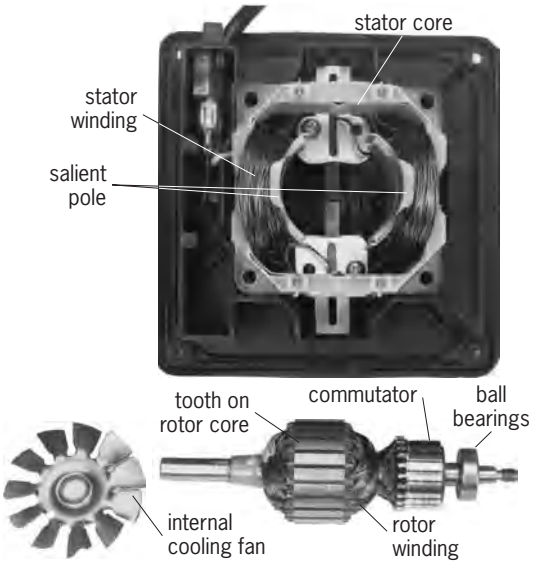

Fig. 3. Disassembled universal motor. (*Emerson Motor Division*)

Varieties of electric motors in common use

Alternating-current motors		
Polyphase	Single phase	Direct-current motors
Synchronous	Synchronous	PMDC (permanent magnet poles)
Induction	Hysteresis	Wound poles
Wound rotor	Reluctance	Shunt
Squirrel cage	Universal (ac or dc) ⟷	Series
Single cage	Induction	Compound
Low resistance (design A)	Shaded pole	
High resistance (design D)	Stepped pole	
Deep bar (design B)	Split phase	
Multicage (design C)	Capacitor start	
	Two-value capacitor	
	Permanent-split (PSC) capacitor	

f is the frequency in hertz and p is the number of poles of the motor. Polyphase induction motors operate at speeds slightly less than those calculated by this formula. *See* SLIP (ELECTRICITY).

Single-phase induction motors have no inherent starting torque and are classified according to the means used to get them started. These classifications include resistance–split-phase (usually simply called split-phase) motors, capacitor-start motors. permanent-split capacitor (PSC) motors, two-value capacitor motors, shaded-pole motors, and stepped-pole motors. Single-phase synchronous motors include reluctance motors and hysteresis motors. *See* ALTERNATING-CURRENT MOTOR; HYSTERESIS MOTOR; INDUCTION MOTOR; RELUCTANCE MOTOR; SYNCHRONOUS MOTOR.

Direct-current motors. Direct-current motors are classified according to the means used to magnetize (excite) their salient field poles. Those employing permanent magnets for this purpose are called PMDC machines. Those having wound poles are called shunt, series, or compound motors, depending on whether their field-pole windings are connected in parallel (shunt = parallel) or in series with the armature, or whether each pole has two windings, one involved in shunt connection and the other in series. A brushless dc motor is a polyphase synchronous motor connected to a dc source through an electronic inverter that is controlled by rotor position sensors in such a way that the system behaves exactly as a shunt dc motor. *See* DIRECT-CURRENT MOTOR.

Universal motors. Universal motors are series dc motors so designed as to make them capable of operating also on single-phase alternating current. They achieve speeds above those possible with synchronous and induction motors, and these high speeds produce high horsepower ratings for very compact motors. *See* UNIVERSAL MOTOR.

Control motors. Motors to be used for precise control of motions are designed to have very low inertia. They include stepper motors, permanent magnet motors, two-phase induction motors, and brushless dc motors. *See* SERVOMECHANISM; STEPPING MOTOR.

Torque development. The windings in motors are largely shielded from the air-gap magnetic field by the steel pole pieces and core teeth, and very little force is exerted by the field on the conductors themselves. Motor torque is best calculated by application of the principle of virtual work to the field energy in the air gap. Under the conditions in the air gap, this principle states that torque is produced if a small rotation θ of the rotor and its magnetic poles results in a change in the energy stored in the air-gap field. If W_{fld} is the energy of the air-gap field, the torque τ is given by Eq. (1). The torque is in newton-meters

$$\tau = \frac{\partial W_{\text{fld}}}{\partial \theta} \tag{1}$$

(N · m) [or pound-feet (lbf · ft)] if W_{fld} is in joules (J) [or foot-pounds (ft · lbf)] and θ in mechanical radians. At least two torque-producing phenomena are evident from Eq. (1). First, torque results from the attraction of rotor and stator poles of unlike polarity, because their alignment would increase the flux per pole and thus increase the field energy. Second, when salient poles are involved, a differential rotation may change the field energy by changing the magnetic reluctance to the flow of flux across the gap. The second phenomenon is the basis for the operation of reluctance and stepping motors and is called reluctance torque. *See* ROTATIONAL MOTION; TORQUE; VIRTUAL WORK PRINCIPLE.

Energy considerations allow formulas to be developed to determine the torque for different motor configurations. Typical is that for a motor with a cylindrical air gap having a diameter D and a length L, and a magnetic flux density that varies sinusoidally with angle around the gap. The formula is written in terms of the specific electric loading (SEL) and the specific magnetic loading (SML). The specific electric loading in this case is the number of conductors of the stator winding per meter of distance measured around the inside surface of the stator, multiplied by the root-mean-square (rms) winding current. The specific magnetic loading is the rms value of the air-gap magnetic flux density. The torque formula (2) applies directly to polyphase induction and

$$\tau = \frac{\pi}{2} D^2 L(\text{SEL})(\text{SML}) k_w \sin \delta_{\text{SR}} \tag{2}$$

cylindrical-rotor synchronous motors, where δ_{SR} is the angular displacement between the poles of magnetomotive force developed by the stator winding and the poles of the air-gap flux field, and k_w is a winding factor. The factor D^2L shows that the torque capability of a motor is proportional to its volume. The torque is independent of the number of poles and hence independent of the rated speed. Thus, for a fixed volume, increasing the number of poles in a motor reduces its horsepower rating because the rated speed is reduced with no change in torque. A similar result is available for dc motors. *See* ELECTRIC ROTATING MACHINERY.

George McPherson, Jr.

Speed Control of AC Motors

The polyphase synchronous motor is a constant-speed (synchronous), variable-torque, doubly excited machine. The stator is supplied with polyphase alternating current of a specific frequency (line frequency), and the rotor field is supplied with direct current. The rotor speed of the synchronous motor is a direct function of the number of stator and rotor field poles and the frequency of the alternating current applied to the stator. Since the number of rotor poles of the polyphase synchronous motor is not easily modified, the change-of-frequency method is the only way to control the synchronous speed of the motor.

The polyphase induction motor is also a doubly excited machine, whose stator is supplied with polyphase alternating current of line frequency and whose rotor is supplied with induced alternating current whose frequency depends on the rotor slip. The speed of the polyphase induction motor is asynchronous, and can be varied by one or more of the following methods: (1) by changing the applied frequency to the stator (the same method as used for the synchronous motor); (2) by controlling the rotor slip by means of rotor resistance control (used for wound-rotor induction motors); (3) by changing the number of poles of both stator and rotor; and (4) by "external voltage" control, obtained by conductively or inductively introducing applied voltages of the proper frequency to the rotor.

A number of electromechanical or purely mechanical methods also provide speed control of alternating-current motors. One is referred to as the Rossman drive, in which an induction motor stator is mounted on trunnion bearings and driven with an auxiliary motor, providing the desired change in slip between stator and rotor. Polyphase induction and synchronous motors having essentially constant speed characteristics at rated voltage are also assembled in packaged drive units employing gears, cylindrical and conical pulleys, and even hydraulic pumps to produce a variable speed output. In addition to reversal of rotation, some of these units employ magnetic slip clutches and solenoids to control the various mechanical and hydraulic arrangements through which a relatively smooth control of speed is achieved.

The principal method of speed control used for fractional-horsepower single-phase induction-type, shaded-pole reluctance, series, and universal motors is the method of primary line-voltage control. It involves a reduction in line voltage applied to the stator winding (of the induction type) or to the armature of series and universal motors. In the former, this produces a reduction of torque and increase in rotor slip. In the latter, it is simply a means of controlling speed by armature voltage control or field flux control or both. The reduction in line voltage is usually accomplished by one of six methods: autotransformer control, series reactance control, tapped main-winding control, saturable reactor (or magnetic amplifier) control, silicon controlled rectifier feedback control, and variable-speed drives. *See* AUTOTRANSFORMER; SATURABLE REACTOR.

Electronic control techniques. The development of power electronic switching devices such as the thyristor, or silicon controlled rectifier (SCR), created unlimited possibilities for control of virtually all types of motors (single-phase, direct-current, and polyphase). Silicon controlled rectifiers in sizes up to 1600 A (rms) with voltage ratings up to 1600 V are available. *See* RECTIFIER; SEMICONDUCTOR RECTIFIER.

SCR control of series motors. Fractional-horse-power universal, alternating-current, and direct-current series motors may be speed-controlled from a single-phase 110- or 220-V supply using the half-wave circuit involving a diode D_1 and silicon controlled rectifier (**Fig. 4***a*). The trigger point of the silicon controlled rectifier is adjusted via the potentiometer R_v, which phase-shifts the gate turn-on voltage of the silicon controlled rectifier whenever its anode A is positive with respect to the cathode K. During the negative half-cycle of input voltage, the silicon controlled rectifier conduction is off and the gating pulse is blocked by the diode D_1. When the positive half-cycle is initiated once again, the capacitor C charges to provide the required gating pulse at the time preset by the time constant R_vC.

Replacing the diode D_1 by a DIAC and the silicon controlled rectifier by a TRIAC converts the circuit from half-wave to full-wave operation and control (Fig. 4*b*). Compared with Fig. 4*a*, this circuit provides almost twice the torque and improved speed regulation. Neither circuit, however, is capable of automatic maintenance of desired speed because of the inherently poor speed regulation of series-type alternating-current, universal, or direct-current motors, with application or variation of motor loading.

Automatic speed regulation is achieved by using the half-wave feedback circuit shown in Fig. 4*c*. This circuit requires, however, that both field leads (f_1, f_2) and both armature leads (A_1, A_2) are separately brought out for connection as shown in Fig. 4*c*. The desired speed is set by the potentiometer R_v, in terms of reference voltage V_d across the diode D_1. The actual motor speed is sensed in terms of the armature voltage V_a. Diode D_1 conducts only when the reference voltage V_d (desired speed) exceeds the

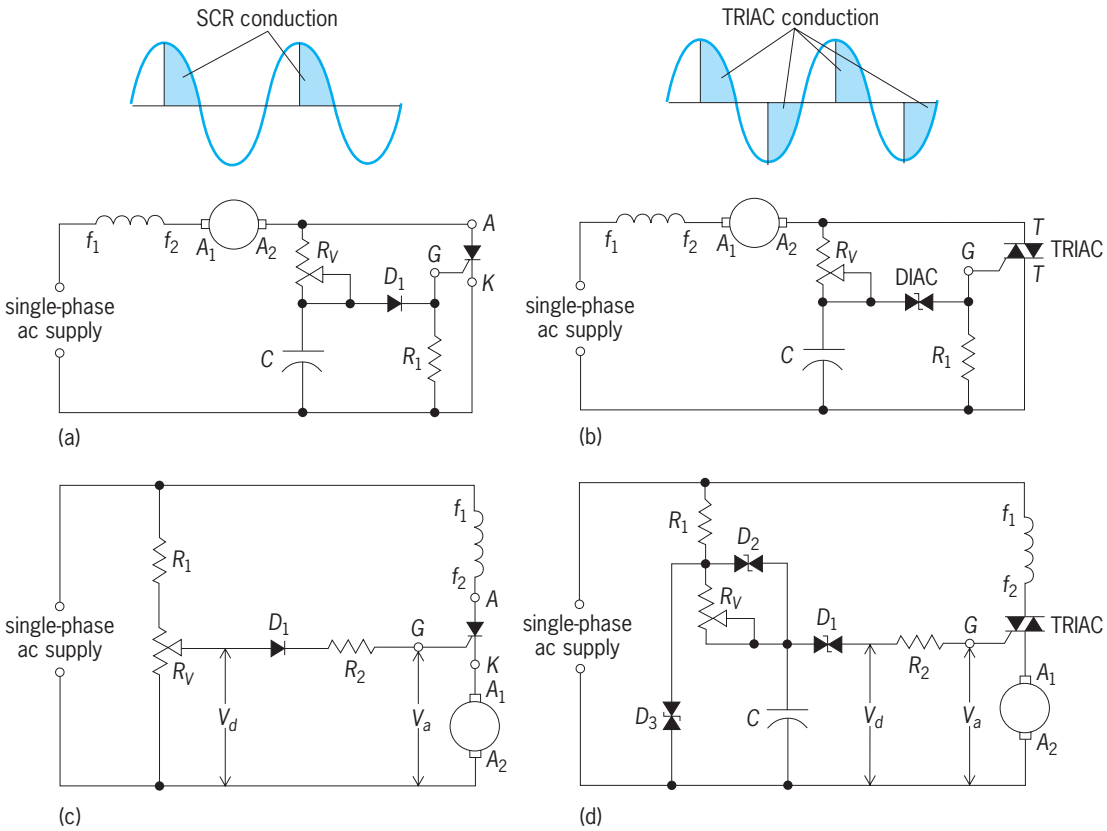

Fig. 4. Single-phase, universal or dc motor control using half- and full-wave (ac) controls, with and without voltage feedback. (*a*) Half-wave circuit. (*b*) Full-wave circuit. (*c*) Half-wave circuit with voltage feedback. (*d*) Full-wave circuit with voltage feedback. (*After I. L. Kosow, Control of Electric Machines, 1973*)

voltage V_a to restore the motor speed to its desired setting. Thus, when the motor load is increased and the speed drops, the silicon controlled rectifier is gated earlier in the cycle because the diode D_1 conducts whenever V_d exceeds V_a.

As with the previous half-wave circuit, the feedback circuit may be converted to full-wave operation by replacing the diode D_1 with a DIAC and the silicon controlled rectifier with a TRIAC (Fig. 4*d*). This circuit also is shown with additional minor modifications to improve the regulation of speed at lower speeds: the addition of the DIACs D_2 and D_3 and the capacitor C. The advent of high-power TRIACs and silicon controlled rectifiers has also extended the use of the circuit shown in Fig. 4*d* to larger, integral-horse-power alternating-current series motors.

When direct-current motors are operated by using the electronic techniques shown in Fig. 4, it is customary to derate motors proportionately because of the heating effect created by alternating-current components in both half- and full-wave waveforms. There is no reversal of direction in either the alternating-current or direct-current series motor under full-wave operation because both the armature and field currents have been reversed.

SCR control of polyphase ac motors. There are fundamentally only three types of polyphase alternating-current motors: synchronous motors (SMs), squirrel-cage induction motors (SCIMs), and wound-rotor induction motors (WRIMs). All three employ identical stator constructions. As a result, larger polyphase motors (up to 10,000 hp or 7500 kW) are controlled by silicon controlled rectifier packages which employ some so-called universal method of speed control. In these methods, both the frequency and stator voltages are varied (in the same proportion) to maintain constant polyphase stator flux densities and avoid operating the motor in the saturation region, thus avoiding overheating.

Two major classes of solid-state adjustable voltage age/frequency drives have emerged, namely the cycloconverter (an ac/ac package) and the rectifier-inverter (an ac-dc-ac package). Both packages convert three-phase fixed-frequency alternating current to three-phase variable voltage/variable frequency.

The half-wave cycloconverter (**Fig. 5***a*) is capable of supplying from 0 to 20 Hz to an alternating-current polyphase motor. It uses six silicon controlled rectifiers per phase but is incapable of either phase-sequence reversal or frequencies above 20 Hz. The full-wave converter (Fig. 5*b*) uses twice the number of silicon-controlled rectifiers (12 per phase) and possesses advantages of wider frequency variation (from +30 Hz down to 0 and up to −30 Hz), potentialities for dynamic braking, and capability of power regeneration.

The symbol for a polyphase squirrel-cage induction motor driven by a cycloconverter is shown in Fig. 5*c*, and a synchronous motor in Fig. 5*d*. The symbols imply frequency/voltage control of identical

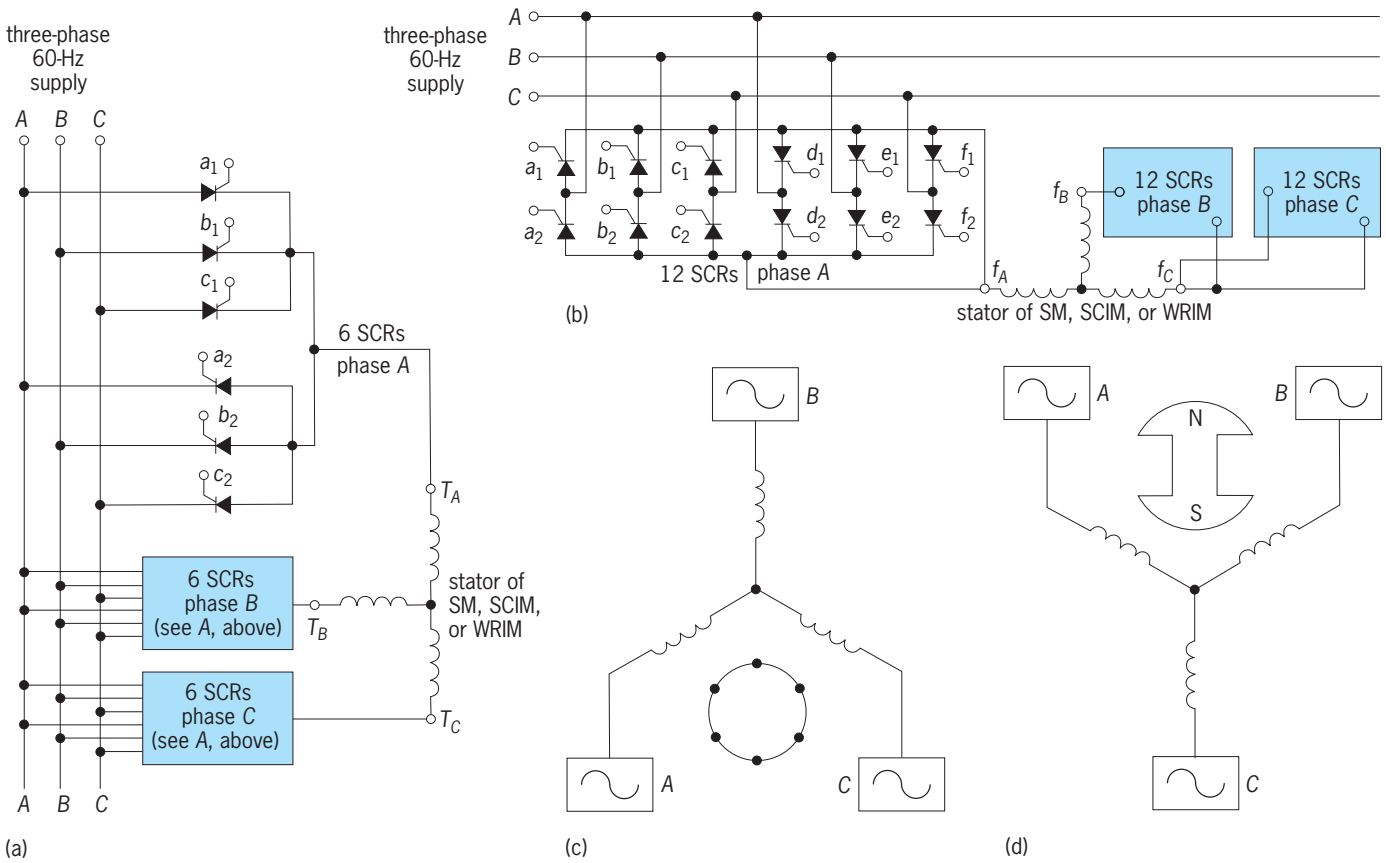

Fig. 5. Half-wave and full-wave cycloconverters with graphical symbols for simplified representation. (*a*) Half-wave cycloconverter (exclusive of triggering of silicon controlled rectifier gates). (*b*) Full-wave cycloconverter (exclusive of silicon controlled rectifier gating). (*c*) Symbol for squirrel-cage induction motor driven by a cycloconverter. (*d*) Symbol for synchronous motor driven by a cycloconverter. (*After I. L. Kosow, Control of Electric Machines, 1973*)

stators, and the nature of the motor is determined solely by its rotor.

The second class of solid-state drive package is the rectifier-inverter, which converts fixed-frequency polyphase alternating current to direct current (variable voltage). The direct current is then inverted by a three-phase inverter (using a minimum of 12 silicon controlled rectifiers commercially) to produce variable-voltage, variable-frequency three-phase alternating current for application to the motor stator. The block diagram of a solid-state rectifier-inverter package is shown in **Fig. 6***a*. The three-phase 60-Hz supply is first rectified to produce variable direct current (Fig. 6*b*) by appropriate phase shifts of silicon controlled rectifier gates. The variable direct current is then applied to the direct-current bus of Fig. 6*a*. Inversion is accomplished by appropriate phase shift of gates of the 12 silicon controlled rectifiers to produce phase and line voltages displaced, respectively, by 120°. This three-phase output voltage is applied to a transformer whose secondary is applied to the stator of the motor (Fig. 6*c*). The rectifier-inverter is capable of power regeneration through the feedback loop shown in Fig. 6*a*.

Brush-shifting motors. By use of a regulating winding, commutator, and a brush-shifting device, the speed of a polyphase induction motor can be controlled similarly to that of a direct-current shunt motor. Such motors are used for knitting and spinning machines, paper mills, and other industrial services that require controlled variable-speed drive. The primary winding on the rotor is supplied from the line through slip rings. The stator windings are the secondary windings, and the third winding, also in the rotor, is an adjusting winding provided with a commutator. Voltages collected from the commutator are fed into the secondary circuit. Three brushes are mounted 120 electrical degrees apart on a movable yoke, and three other brushes are similarly mounted on a separate movable yoke. Each set of brushes can be moved as a group. Thus both the spacing between sets of brushes and the angular position of the brushes are adjustable. Brush spacing determines the magnitude of the voltage applied to the secondary. When brush sets are so adjusted that pairs of brushes are in contact with the same commutator segment, the secondary is short-circuited and no voltage is supplied. Under these conditions the motor behaves as an ordinary induction motor. The speed can be reduced by separating the brushes so that secondary current produces a negative torque. The machine can be operated above synchronism by interchanging the position of the brushes, so the voltage collected is in a direction to produce a positive torque. The motor can be reversed by reversing two of the leads supplying the primary.

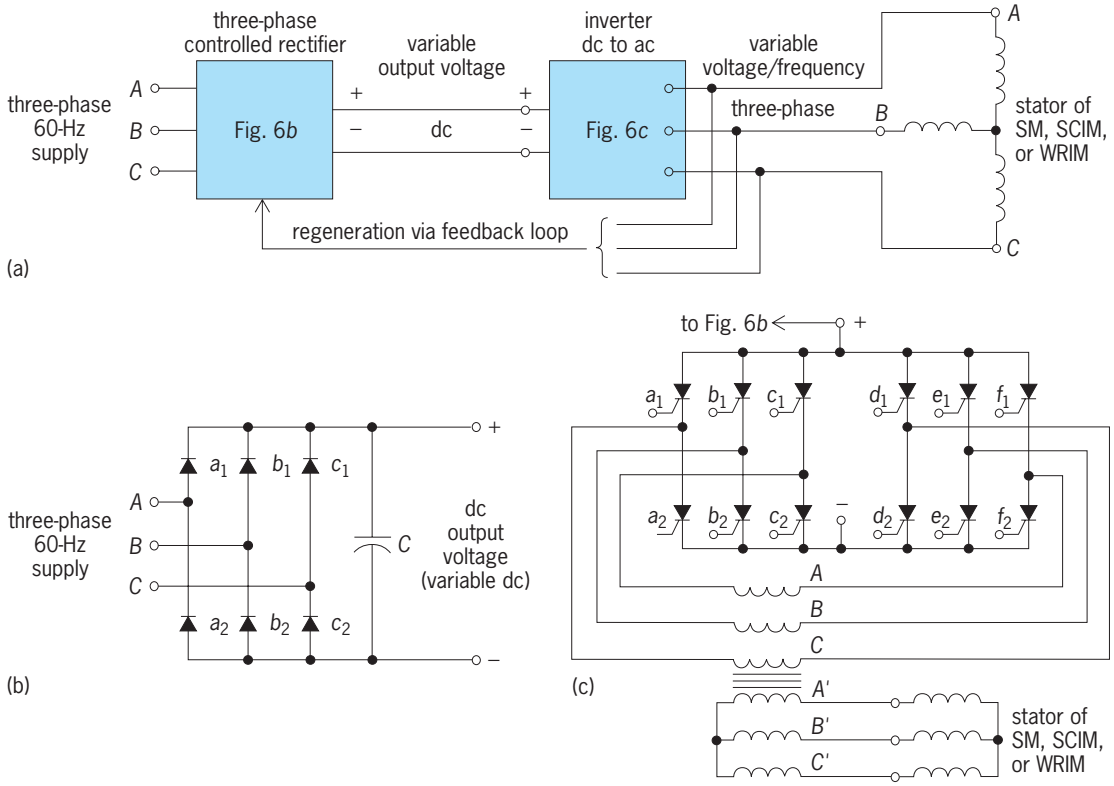

Fig. 6. Rectifier-inverter circuitry for variable voltage/frequency control of three-phase motors. (*a*) Block diagram of rectifier-inverter. (*b*) Three-phase full-wave controlled rectifier. (*c*) Basic inverter dc to three-phase ac for phase-shift voltage control. (*After I. L. Kosow, Control of Electric Machines, 1973*)

Closed-loop control techniques. As mentioned above, slip control of the speed of a wound-rotor induction motor can be achieved by manually controlling an external resistance in series with the rotor winding. This method of speed control has the disadvantages of poor speed regulation at any given resistance setting, reduced motor efficiency due to power lost in the external resistors, and the requirement of a human operator to manually adjust the motor to a desired speed. This method of speed control falls

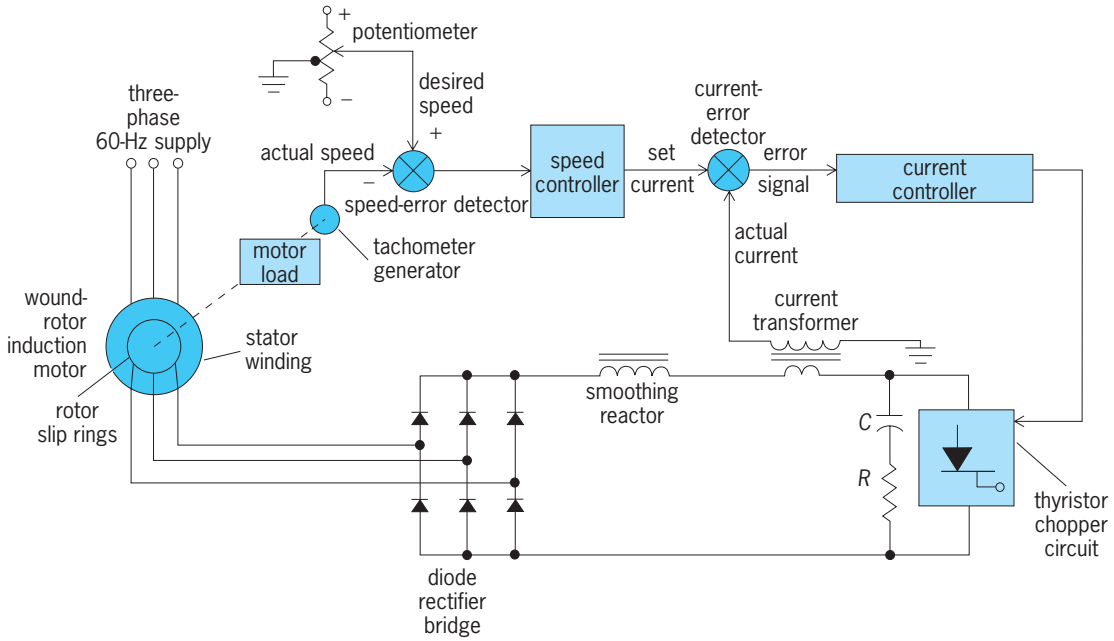

Fig. 7. Speed control of wound-rotor induction motor by static variation of external rotor resistance, using a high-frequency thyristor-controlled chopper circuit.

under the general category of open-loop control.

Thyristor-controlled chopper circuit. A modern method of closed-loop control (**Fig. 7**) uses a high-frequency thyristor-controlled chopper circuit that permits the external rotor resistance to be varied steplessly by using static semiconductor elements. The feedback circuit shown provides a relatively inexpensive variable-speed drive with good dynamic response, a wider range of speed control, and improved speed regulation. *See* CONTROL SYSTEMS.

The desired speed in Fig. 7 is set by means of a potentiometer. The speed-error detector senses the difference between the actual rotor speed and the desired (set) speed, which in turn provides the "set speed" signal to a speed controller. The latter provides a "set current" signal to a second error detector that controls the thyristor chopper and rotor slip power. The thyristor chopper, which parallels the *RC* branch, is switched (on and off) at a frequency of approximately 1 kHz. The ratio of on/off time controls the effective rotor resistance and consequently the motor speed by modification of the torque–speed characteristic.

Cascade drive. The circuit of Fig. 7 uses a single closed-loop method of automatic control of a wound-rotor induction motor. A somewhat more sophisticated method of automatic closed-loop control is the so-called cascade drive system (**Fig. 8**), which uses two closed loops. The inner control loop uses a variation in the thyristor firing angle to control and adjust the rotor current in the cascade circuit. The rotor current actually determines the motor torque and speed, in part. The outer or speed loop senses the rotor current (on the alternating-current side of the thyristor inverter) and uses it as feedback to set the desired motor speed via a speed-error detector and current-error detector, respectively. In effect, the dual closed-loop system permits the speed-error signal to produce the necessary motor torque and rotor current to reduce the error between the desired (set) speed and the actual rotor speed, regardless of changes in motor load. This method of control provides speed regulation of less than 0.1% from no-load to full load at any desired speed setting.

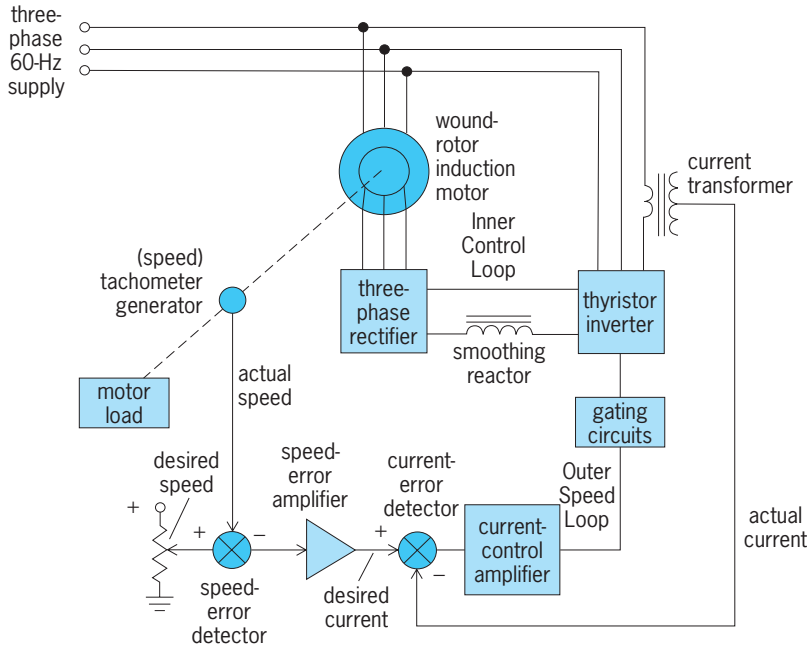

Fig. 8. Closed-loop control of a wound-rotor induction motor by a static cascade drive system.

Current-source inverter. The basic voltage-source inverter (VSI) of Fig. 6*c* is modified in **Fig. 9** to provide a current-source inverter (CSI). Unlike the voltage-source inverter, which uses a wide variety of switching devices and commutation arrangements, the current-source inverter uses the conventional and preferred arrangement shown in Fig. 9. The name current-source inverter stems from the fact that the current to the motor stator is automatically and sequentially commutated from one full-wave thyristor to the next, after each simple gate signal to the oncoming thyristor. The capacitors are selected in conjunction with the single inductor to ensure good commutation. Unlike the voltage-source

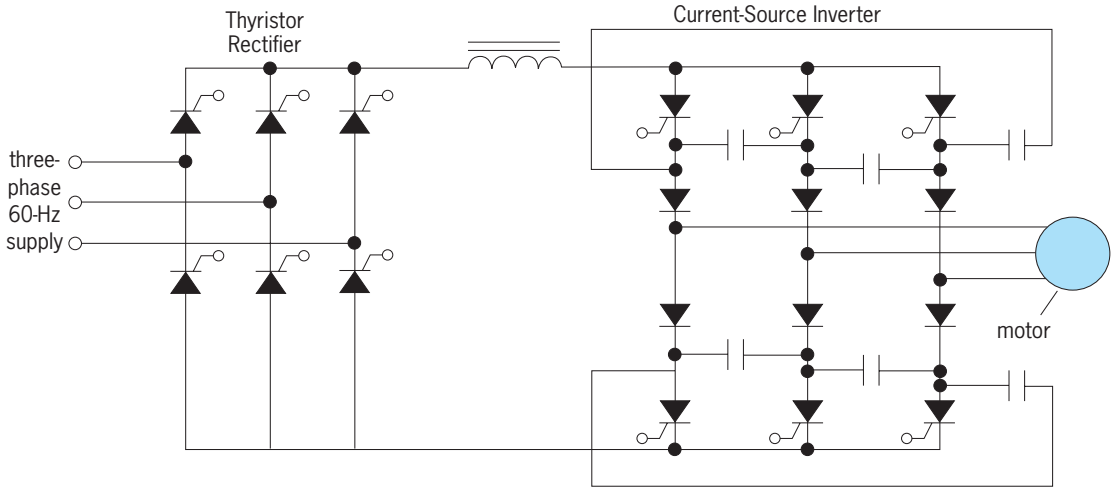

Fig. 9. Full-wave thyristor rectifier and autosequentially commutated current source inverter.

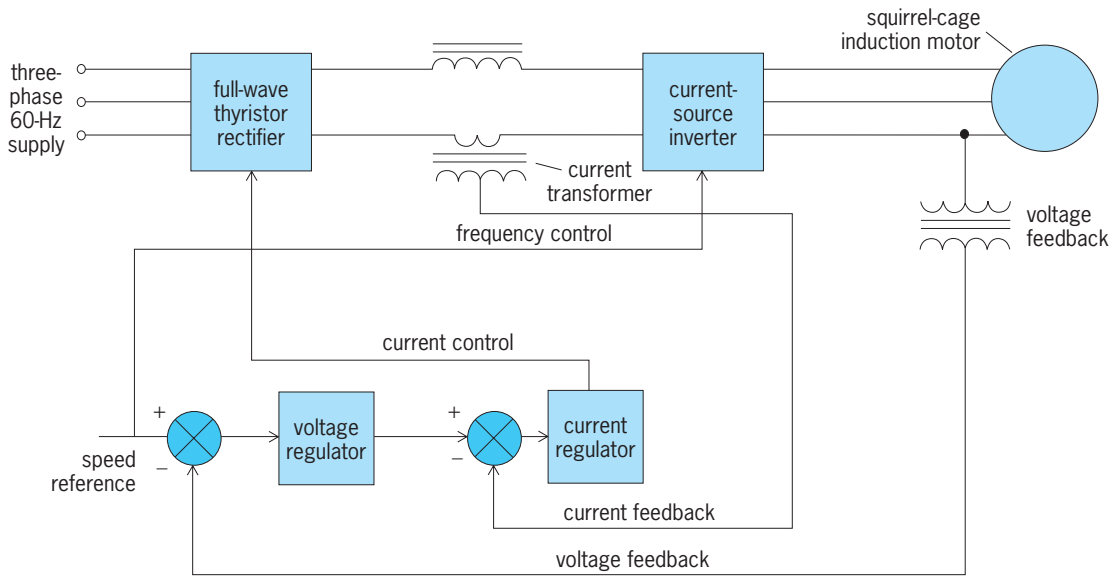

Fig. 10. Closed-loop control of a squirrel-cage induction motor using a current-source inverter with a current-regulating loop and a voltage-regulating loop.

inverter shown in Fig. 6*c*, which provides open-loop control of the stator voltage source, the current-source inverter requires feedback control to permit its use as an adjustable speed drive for a squirrel-cage induction motor or wound-rotor induction motor.

The simplest application of a current-source inverter, providing both current and voltage feedback, is shown in **Fig. 10**. The outer voltage feedback loop senses the stator terminal voltage as an indication of motor speed. The error signal to the voltage regulator serves as a reference to the current-regulator error detector. The current regulator controls the direct-current supply of the thyristor phase-controlled rectifier.

In effect, the circuit of Fig. 10 converts the current-source inverter to an equivalent feedback-controlled voltage source. The advantages of the current-source inverter over conventional pulse-width modulator drives or voltage-source inverter drives are found in three areas: motor harmonics, regeneration, and fault currents. In the current-source inverter, unlike other drives mentioned above, the inverter determines the harmonic currents, which are always proportional to the motor line current at the fundamental frequency. Regeneration is much simpler with the current-source inverter than other drives because only the link voltage reverses in regeneration, rather than the link current, eliminating the need for a dual or other bidirectional input converter. The large link inductor tends to oppose any rapid current changes, making fault protection much simpler than in voltage-source inversion or pulse-width modulation systems, which have large link capacitors.

Irving L. Kosow

Speed Control of DC Motors

Compared to an alternating-current induction motor, the direct-current motor is more expensive and complex. However, the direct-current motor has the advantage of having the capability to develop a wide variety of torque–speed characteristics. Direct-current motors find applications in electric traction and in industrial drives requiring large speed variations in both directions of rotation, or requiring a precise variation of speed over a narrow range. In other words, the direct-current motor is highly controllable.

Speed equation. The direct-current motor has two accessible electric circuits, the field circuit and the armature circuit (**Fig. 11**), which may be used to control the motor speed and other characteristics. The motor steady-state speed n is given by Eq. (3),

$$n = \frac{V_t - I_a(R_a + R_{ax})}{\phi(I_f)} \tag{3}$$

where V_t is the terminal voltage across the armature, I_a is the armature current, R_a is the armature winding resistance, R_{ax} is the external resistance in the armature circuit, and ϕ is some function of the field current I_f. If the magnetic circuit of the motor is not saturated, Eq. (3) may be written as Eq. (4), where k is a constant.

$$n = \frac{V_t - I_a(R_a + R_{ax})}{kI_f} \tag{4}$$

Armature and field control. From Eqs. (3) and (4), it follows that the motor speed may be controlled by varying the terminal voltage V_t, varying the field current I_f by changing the field-circuit external resistance R_{fx}, and varying the armature-circuit external resistance R_{ax}. The method of varying the terminal voltage is known as the armature control, and the method of varying the field-circuit external resistance is called the field control. The method of varying the armature-circuit external resistance is rather inefficient (or lossy) and is suitable for very small motors. From Fig. 11, it is clear that the field current I_f is limited by the field-circuit resistance, $R_f + R_{fx}$. Thus,

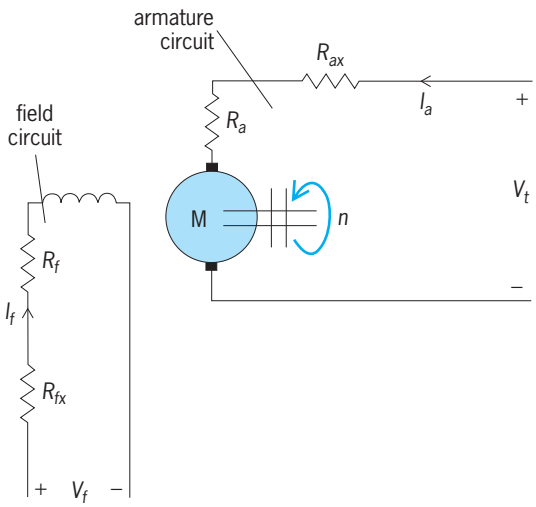

Fig. 11. Field and armature circuits of a dc motor. M = motor; V_t = terminal voltage across the armature; I_a = armature current; R_a = armature winding resistance; R_{ax} = external resistance in armature circuit; V_f = voltage across the field; I_f = field current; R_f = field winding resistance; R_{fx} = external resistance in field circuit; n = speed.

I_f cannot exceed a value determined by R_f given by Eq. (5), where V_f is the voltage across the field.

$$(I_f)_{\max} = \frac{V_f}{R_f} \tag{5}$$

Of course, I_f can be made to decrease below $(I_f)_{\max}$ by increasing the external resistance R_{fx} inserted in the field circuit. Because the speed of the motor depends inversely on the field current, as governed by Eq. (3), for a given voltage V_f, field control can be used only to increase the motor speed beyond the minimum determined by $(I_f)_{\max}$. From Eq. (3), however, it follows that armature control, achieved by varying V_t, may be used to either increase or decrease the motor speed from a given reference. Prior to the advent of power semiconductors, armature control was used in special cases.

Ward-Leonard system. The methods of armature control and field control have been combined to obtain the Ward-Leonard system. In this system a generator driven by an auxiliary motor supplies the motor to be controlled. By varying the generator field-circuit resistance and the motor field-circuit resistance, this system gives a wide range of speed control in both directions of rotation of the motor. The Ward-Leonard system has found applications in mine hoists and rolling mills to drive and reverse large and varying loads. The major drawback of the system is the capital cost of a set of three machines of almost equal ratings, and the space occupied by the motor-generator set. *See* MOTOR-GENERATOR SET.

Solid-state control. Resistance control, either in the field or armature of a direct-current motor, results in poor efficiency. High-power solid-state controllers offer the most practical, reliable, and efficient method of motor control. The most commonly used solid-state devices in motor control are power transistors and thyristors or silicon controlled rectifiers. The principal differences between the two are that the transistor requires a continuous driving signal during conduction, whereas the thyristor requires only a pulse to initiate conduction; and the transistor switches off when the driving signal is removed, but the thyristor turns off when the load current is reduced to zero. Application of a reverse-polarity voltage shortens the recovery time to full reverse-blocking capability. Utilizing power semiconductors, choppers and converters are the most common electronic controllers employed in direct-current drives. *See* TRANSISTOR.

Choppers. In principle, a chopper is an on-off switch connecting the load to and disconnecting it from the battery (or direct-current source), thus producing a chopped voltage across the load. Symbolically, a chopper as a switch is represented in **Fig. 12***a*, and a basic chopper circuit is shown in Fig. 12*b*. In this circuit, when the thyristor does not conduct, the load current flows through the freewheeling diode *D*, which provides a path for the armature current when the silicon controlled rectifier is not conducting. From Fig. 12*a* it is clear that the average voltage across the load, V_0, is given by Eqs. (6), where V is

$$V_0 = \frac{t_{\text{on}}}{t_{\text{on}} + t_{\text{off}}} \qquad V = \frac{t_{\text{on}}}{T} \qquad V = \alpha V \tag{6}$$

the battery voltage, the various times are shown in Fig. 12*a*, T is known as the chopping period, and $\alpha = t_{\text{on}}/T$ is called the duty cycle. Thus the voltage across the load varies with the duty cycle.

There are three ways in which the chopper output voltage can be varied, and these are illustrated in Fig. 12*c*–*e*. The first method, in which the chopping frequency is kept constant and the pulse width (or on-time t_{on}) is varied (Fig. 12*c*), is known as pulse-width modulation. In the second method, called frequency modulation, either t_{on} or t_{off} is fixed, and the chopping period is variable (Fig. 12*d*). The preceding two methods can be combined to obtain pulse-width and frequency modulation (Fig. 12*e*), which is used in current limit control. In a method involving frequency modulation, the frequency must not be decreased to a value that may cause a pulsating effect or a discontinuous armature current; and the frequency should not be increased to such a high value as to result in excessive switching losses. The switching frequency of most choppers for electric drives range from 100 to 1000 pulses per second. The drawback of a high-frequency chopper is that current interruption is accompanied by high-frequency noise. *See* MODULATION.

A simplified form of a chopper circuit is shown in Fig. 12*f*, where the chopper is shown to supply a direct-current series motor (having the field winding in series with the armature). The circuit is a pulse-width modulation circuit, where t_{on} and t_{off} determine the average voltage across the motor, as given by Eq. (6). As mentioned above, the silicon controlled rectifier cannot turn itself off once it begins to conduct. Thus, turning the silicon controlled rectifier off requires a commutating circuit that

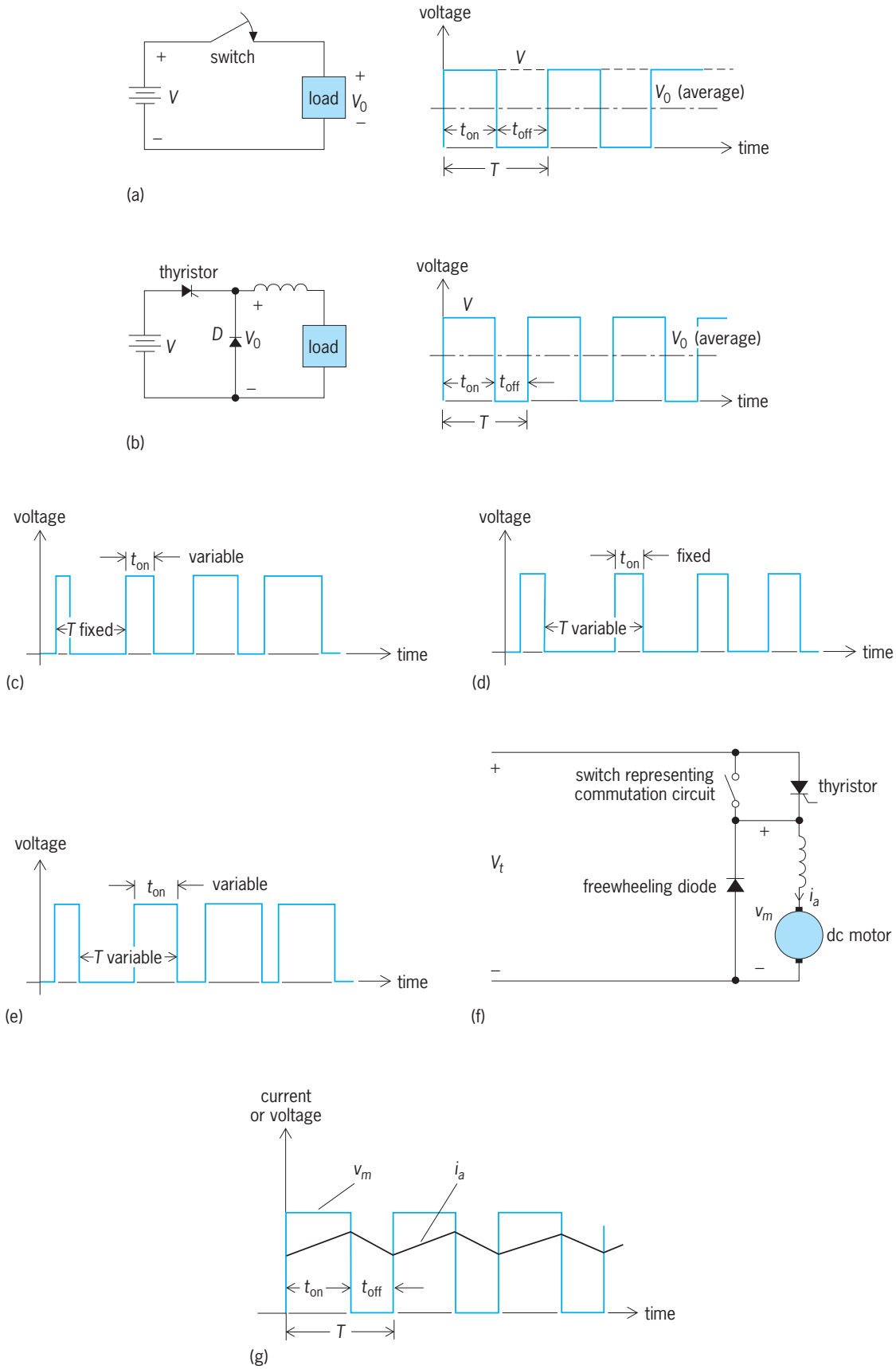

Fig. 12. Chopper control of dc motors. (*a*) Symbolic representation of a chopper. (*b*) Basic chopper circuit. (*c*) Constant-frequency, variable-pulse-width waveform. (*d*) Variable-frequency, constant-pulse-width waveform. (*e*) Variable-frequency, variable-pulse-width waveform. (*f*) A dc motor driven by a chopper. (*g*) Motor current and voltage waveforms. V_t = applied voltage; v_m = motor voltage; i_a = armature current. Lowercase letters denote instantaneous quantities.

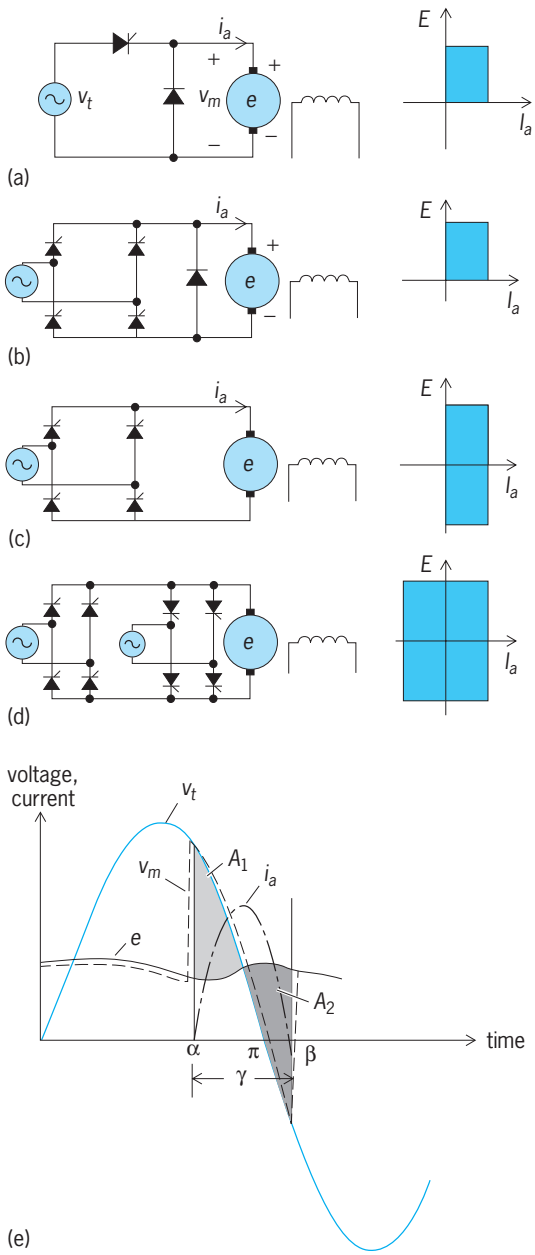

Fig. 13. Single-phase converters and their quadrant operations. (*a*) Half-wave converter. (*b*) Semiconverter. (*c*) Full converter, (*d*) Dual converter. (*e*) Waveforms in a dc motor controlled by a single-phase half-wave converter. E = voltage at dc terminals; I_a = current at dc terminals; v_m = motor voltage; v_t = applied voltage; i_a = armature current; e = motor back electromotive force (emf). Lowercase letters denote instantaneous quantities.

impresses a negative voltage on the silicon controlled rectifier for a very short time (of the order of microseconds). For simplicity, the circuitry for commutation, which is often quite involved, is denoted by a switch in Fig. 12*f*. The motor current and voltage waveforms are shown in Fig. 12*g*. The silicon controlled rectifier is turned on by a gating signal at $t = 0$. The armature current i_a builds up, and the motor starts and picks up speed. After the time t_{on}, the silicon controlled rectifier is turned off, and remains off for a period t_{off}. During this period the armature current continues to drop through the freewheeling diode circuit. Again, at the end of t_{off} the silicon controlled rectifier is turned on, and the on-off cycle continues. The chopper thus acts as a variable-voltage source.

Converters. A converter changes an alternating-current input voltage to a controllable direct-current voltage. Converters have an advantage over choppers in that natural or line commutation is possible in converters; therefore, no complex commutation circuitry is required. Controlled converters use silicon controlled rectifiers and operate either on single-phase or three-phase alternating current. The four types of phase-controlled converters commonly used for direct-current motor control are half-wave converters, semiconverters, full converters, and dual converters. Each of these could be either a single-phase or a three-phase converter. Half-wave converters are used for motors of ratings up to 0.5 hp (0.4 kW). Semiconverters are one-quadrant converters in that the polarities of voltage and current at the direct-current terminals do not reverse. In a full converter, the polarity of the voltage can reverse, but the current is unidirectional. In this sense, a full converter is a two-quadrant converter. Dual converters are four-quadrant converters. The four types of single-phase converters and their respective quadrant operations are shown in **Fig. 13***a–d*.

The simplest converter, the single-phase half-wave converter, will be used to demonstrate the operation of a converter-fed direct-current motor. The waveforms of motor speed, current, and voltage are given in Fig. 13*e*, if the armature resistance is neglected. By controlling the firing angle of the silicon controlled rectifier, the voltage across the motor armature, and hence its speed, is controlled. In Fig. 13*e*, α is the firing angle for the silicon controlled rectifier. However, the armature current i_a does not begin to flow immediately after the silicon controlled rectifier is turned on. The armature current begins to flow only when the line voltage v_t becomes greater than the motor voltage v_m. The current continues to flow for the period γ shown in Fig. 13*e*. This period is also known as the conduction angle, and is determined by the equality of the shaded areas A_1 and A_2. The area A_1 corresponds to the energy stored in the inductance L of the motor circuit while the armature current is building up. This stored energy is returned to the source during the period that the armature current decreases and ultimately becomes zero. The silicon controlled rectifier then blocks until it is turned on again. The motor speed is controlled by the firing angle α. *See* CONVERTER.

Syed A. Nasar

Bibliography. B. J. Chalmers (ed.), *Electric Motor Handbook*, 1988; S. B. Dewan, G. R. Slemon, and A. Straughen, *Power Semiconductor Drives*, 1984; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; A. E. Fitzgerald, C. Kingsley, and S. D. Umans, *Electric Machinery*, 6th ed., 2003; I. L. Kosow, *Electric Machinery and Transformers*, 2d ed., 1991; G. McPherson and R. D. Laramore, *An Introduction to Electrical Machines and Transformers*, 2d ed., 1990; S. A. Nasar, *Electric Machines and Power*

Systems, I: *Electric Machines*, 1995; S. A. Nasar and I. Boldea, *Linear Electric Motors*, 1987; National Electrical Manufacturers Association, *Motors and Generators*, Publ. MG1-2003, 2003; P. C. Sen, *Thyristor DC Drives*, 1981, reprint 1991; H. A. Toliyat and G. B. Kliman (eds.), *Handbook of Electric Motors*, 2d ed., 2004.

Motor-generator set

A motor and one or more generators, with their shafts mechanically coupled, used to convert the voltage or frequency of an available power source to another desired frequency or voltage. The motor of the set is selected so that it operates from the available power supply; the generators are designed to provide the desired output voltage or frequency. Motor-generator sets are also employed to provide special control features for the output voltage.

The principal advantage of a motor-generator set over other conversion systems is the flexibility offered by the use of separate machines for each function. Assemblies of standard machines may often be employed with a minimum of additional engineering required. Since energy conversion is employed twice, electrical to mechanical and back to electrical, the efficiency of this system is lower than that in most other conversion methods. In a two-unit set the efficiency is the product of the efficiencies of the motor and of the generator.

Motor-generator sets are used for a variety of purposes, such as providing a precisely regulated dc current for a welding application, a high-frequency ac power for an induction-heating application, or a continuously and rapidly adjustable dc voltage to the armature of a dc motor employed in a position control system. *See* GENERATOR; MOTOR. Arthur R. Eckels

Motor systems

Those portions of nervous systems that regulate and control the contractile activity of muscle and the secretory activity of glands. Muscles and glands are the two types of organ by which an organism reacts to its environment; together they constitute the machinery of behavior. Cardiac muscle and some smooth muscle and glandular structures can function independently of the nervous system but in a poorly coordinated fashion. Skeletal muscle activity, however, is entirely dependent on neural control. Destruction of the nerves supplying skeletal muscles results in paralysis, or inability to move. The somatic motor system includes those regions of the central nervous system involved in controlling the contraction of skeletal muscles in a manner appropriate to environmental conditions and internal states. This article discusses skeletal muscle innervation and contraction; the motor unit; and neural centers for motor control, particularly the motor cortex. *See* GLAND; MUSCLE; MUSCULAR SYSTEM.

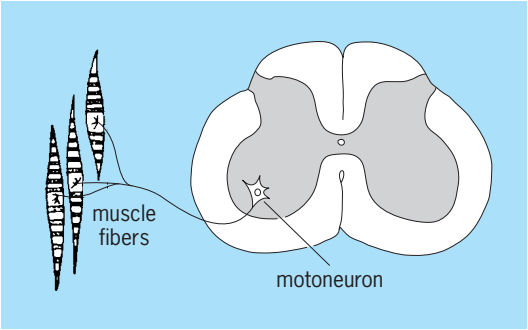

Fig. 1. Diagrammatic representation of a single motor unit. A motoneuron in the ventral horn of the spinal cord innervates multiple muscle fibers (typically more than the three shown).

Skeletal Muscle

A skeletal or striated muscle consists of a bundle of individual contractile elements called muscle fibers. The fibers are held together in the muscle by connective tissue, and their ends are attached via tendons to movable bones; thus, muscle shortening during active contraction results in movement. The nerve supply to skeletal muscles of the limbs and trunk is derived from large nerve cells called motoneurons, whose cell bodies are located in the ventral horn of the spinal cord (**Fig. 1**). Muscles of the face and head are innervated by motoneurons in the brainstem. The axons of the motoneurons traverse the ventral spinal roots (or the appropriate cranial nerve roots) and reach the muscles via peripheral nerve trunks. In the muscle, the axon of every motoneuron divides repeatedly into many terminal branches, each of which innervates a single muscle fiber. The region of innervation, called the neuromuscular junction, or motor end plate, is a secure synaptic contact between the motoneuron terminal and the muscle fiber membrane (**Fig. 2**).

Skeletal muscle contraction. A fixed sequence of events leads from motoneuron activity to contraction of skeletal muscle. An action potential initiated in a motoneuron propagates over the motoneuron axon into its many terminals. At each motor end plate, the action potential causes release of a neuromuscular transmitter agent, which depolarizes the postsynaptic membrane sufficiently to initiate an action potential in the muscle fiber membrane. This muscle fiber action potential propagates along the fiber at a speed of about 5 m/s and triggers the contractile process. *See* BIOPOTENTIALS AND IONIC CURRENTS.

The neuromuscular transmitter agent is acetylcholine, which is synthesized by the nerve terminals and is contained in tiny vesicles seen in electron micrographs. When the terminal is depolarized, some vesicles rupture at the surface, releasing acetylcholine into the synaptic gap. The acetylcholine diffuses across the gap and combines with receptors in the postsynaptic membrane. This combination opens channels in the membrane, and they conduct ions across the membrane, resulting in its depolarization. The presynaptic action potential liberates enough acetylcholine to depolarize the

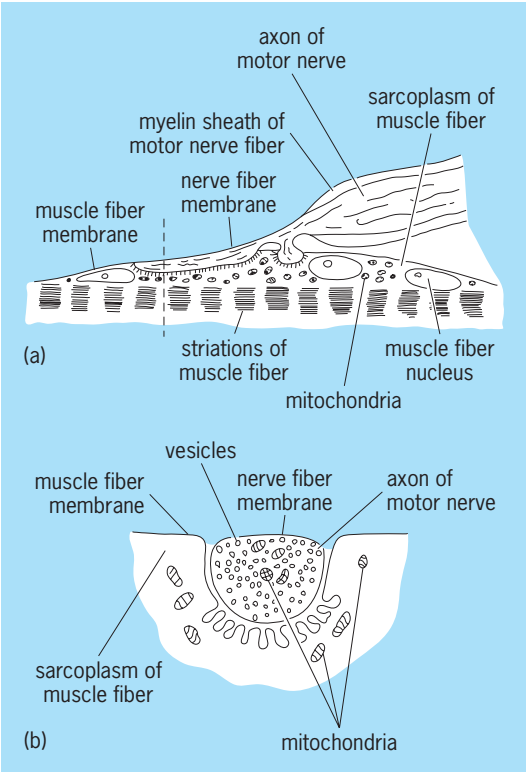

Fig. 2. Diagrams of the neuromuscular junction, or end plate. (*a*) Section cut parallel to long axis of muscle fiber, as seen by light microscopy. (*b*) Section cut perpendicular to long axis of muscle fiber (as through broken line in *a*), as seen by electron microscopy.

muscle membrane past the threshold for initiating an action potential. The receptors in the specialized muscle membrane of the end plate make it about 1000 times more sensitive to acetylcholine than nonjunctional regions of the muscle membrane. *See* ACETYLCHOLINE; SYNAPTIC TRANSMISSION.

The contractile elements in a muscle fiber consist of two proteins, actin and myosin, present in a ratio of about 1:3. The combination, called actomyosin, forms long molecular chains. In the presence of adenosine triphosphate (ATP), actomyosin filaments contract. Each muscle fiber contains up to 10 million such myofilaments of actomyosin. Electron microscopic studies of muscle show that the actin and myosin elements lie parallel to one another and that during contraction the interdigitated elements slide past each other, thus shortening the muscle. *See* ADENOSINE TRIPHOSPHATE (ATP); MUSCLE PROTEINS.

Summation of contraction. The mechanical tension generated by a single contraction of a muscle fiber is much more prolonged than the electrical action potential that initiates it. The muscle fiber action potential lasts only a few milliseconds, but the rise and fall of the resulting twitch tension can take over 100 ms (**Fig. 3**). The muscle action potential, like the action potential of neurons, is an all-or-none phenomenon; its amplitude cannot be modified by changing the strength of the stimulus eliciting it. The contractile mechanism, however, is graded in strength, and the tension produced by a single twitch contraction is less than the maximal possible tension of the fiber. Because the contractile process greatly outlasts the duration of the muscle action potential, the tension elicited by two action potentials briefly separated in time is greater than that produced by a single twitch. The twitch contraction produced by a second action potential sums with that persisting from the first. Several muscle action potentials initiated in rapid succession produce temporal summation of the successive contractile responses. However, the increment of tension of each contraction diminishes as the net tension increases, until a maximum tension, called the tetanic tension, is reached. In some muscles the tetanic tension is nearly four times greater than the twitch tension. The rate of discharge necessary to produce maximal tetanic tension varies in different muscles, from 30 discharges per second for the slow soleus muscle to 350 discharges per second for the rapidly contracting internal rectus muscle of the eye.

Motor units. Since synaptic transmission at the neuromuscular junction is so secure, an action potential in the motoneuron will produce contraction of every muscle fiber that it contacts. For this reason, the motoneuron and all the fibers it innervates form a functional unit, called the motor unit. The number of muscle fibers in a single motor unit may be as small as 6 (for intrinsic eye muscles) or over 700 (for motor units of large limb muscles). For a given muscle the average number of muscle fibers per motor unit can be determined by dividing the total number of fibers in the muscle by the number of motoneurons in its pool. The typical innervation ratio for limb muscles is on the order of 120–150. In general, muscles involved in delicate rapid movements have fewer muscle fibers per motor unit than large muscles concerned with gross movements.

Physiological properties. Even within a given muscle, different motor units have significantly different physiological properties; some motor units are specialized for steady, prolonged contraction, while others are used for brief periods of high tension. The properties of motor units that underlie their functional specialization include the biochemical makeup of their muscle fibers, the speed and magnitude of their twitch tension, their response to repetitive activation, the size of their motoneuron, and physiological constraints on their relative recruitment. Detailed studies on hindlimb muscles of the cat indicate that these properties tend to vary systematically together. The spectrum of motor unit properties may be understood in terms of two contrasting types (Fig. 3). On one end of the functional spectrum are numerous "slow" type-S motor units, so called because their twitch tension takes longer to reach a peak (80–100 ms) and to decay than the more rapid twitch tension of "fast" type-F motor units at the opposite end of the spectrum. The maximum twitch tension of slow units (for example, 1 gram for gastrocnemius muscle) is smaller than the peak tension of fast units (about 20 g). When stimulated repetitively for prolonged periods of time, the tetanic

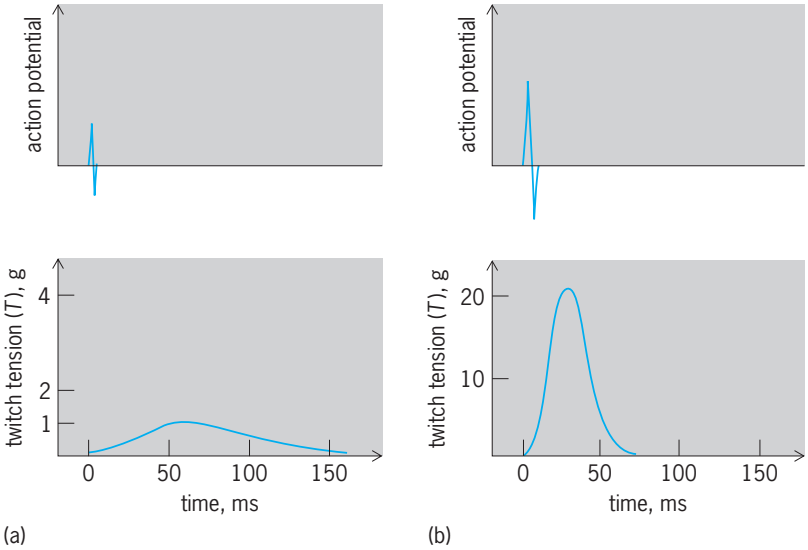

Fig. 3. Electromyographic action potentials and mechanical twitch tensions of two motor units. Type S (slow twitch) on left and type F (fast twitch) on right.

tension of the type-S units is sustained for much longer—on the order of hours—than the tetanic tension of the larger type-F units, which fatigue within minutes. These differences in mechanical performance are related to different metabolic properties of their muscle fibers. As a result, slow motor units are best at providing prolonged periods of tension, whereas fast units are specialized for brief bursts of intense activity.

Motoneuron recruitment. These functional differences are further correlated with differences in the recruitment of their motoneurons. The type-S motor units are generally recruited at lower thresholds in a variety of motor responses than the type-F units. The motoneurons supplying slow motor units tend to have smaller soma size than the large motoneurons of the type-F motor units. In most cases, the motoneurons of a pool are activated in a systematic sequence, in order of increasing size; this recruitment order is the same whether the motoneurons are activated by descending commands from higher centers or by input from peripheral muscle nerves. The fact that properties like twitch tension and recruitment order are related to the motoneuron size has been called the size principle. For example, in isometric contraction of the first dorsal interosseous muscle of the hand, the first motor units to be activated contribute twitch tensions of 0.1 g; as the net muscle force increases, ever larger motor units become active; near the maximum levels of force, the largest units are finally recruited, with a twitch tension of 10 g. Thus, the larger units are recruited at higher levels of net force and also contribute larger twitch tensions. In fact, the size of the motor units' twitch tension is directly proportional to their recruitment level over a wide range of tensions.

Muscle tension. Thus, during voluntary muscle contractions, progressively greater muscle tension is produced by two mechanisms: additional motor units are recruited into activity, in order of small to large; and the firing rates of active motor units increase with net force. In some muscles, the rate of motor unit discharge may increase from 8 to 30 action potentials per second, as the muscle contraction increases from light to maximal effort. At the lower firing rates, the successive twitch tensions of each motor unit produce a series of separate twitches. Since the different motor units of a muscle typically discharge asynchronously, the net tension in the muscle nevertheless varies smoothly. At higher frequencies of motor unit discharge, the twitch tensions of individual motor units begin to summate, producing still further increases in tension. Thus, increased temporal summation of twitches in individual units as well as increased spatial summation of twitches in different units contribute to increased tension. Moreover, the contribution of the large units recruited at high tension is proportionately greater.

Components of Skeletal Motor System

Motoneurons are activated by nerve impulses arriving through many different neural pathways. Some of their neural input originates in peripheral receptor organs located in the muscles themselves, or in receptors in skin or joints. Many muscle receptors discharge in proportion to muscle length or tension; such receptors have relatively potent connections to motoneurons, either direct monosynaptic connections or relays via one or more interneurons. Similarly, stimulation of skin and joints, particularly painful stimulation, can strongly affect motoneurons. Such simple segmental pathways constitute the basis for spinal reflexes.

The other major source of input to motoneurons arises from supraspinal centers. **Figure 4** shows the main nervous system centers involved in controlling the input to motoneurons. Evidence that these regions play a role in control of motor activity comes from three types of observations: damage of cells in these areas by experimental or clinical lesions produces motor deficits; electrial stimulation in these regions evokes motor responses or interferes with

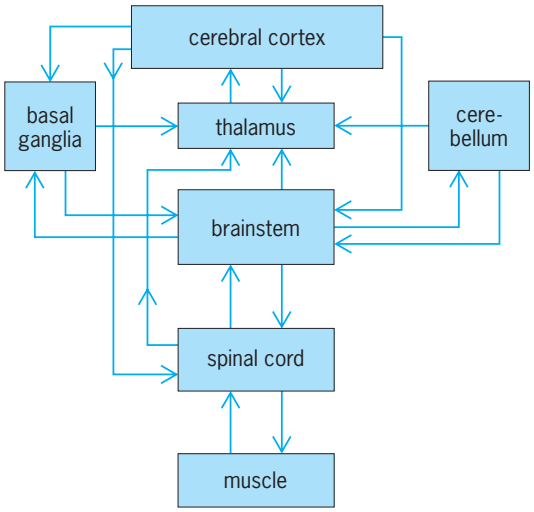

Fig. 4. Schematic diagram of the major components of the vertebrate motor systems. Arrows indicate the main neural connections between regions.

ongoing movements; and activity of cells recorded in these areas in moving animals is clearly related to their motor activity. Figure 4 indicates the major interconnections between cells in these regions; as one might expect, they function together during voluntary movements. The representation of the primate motor system in Fig. 4 is greatly simplified to provide an overview of its major components.

Segmental circuits. At the spinal level, the motoneurons and muscles have a close reciprocal connection. Afferent connections from receptors in the muscles return sensory feedback to the same motoneurons which contract the muscle. Connections to motoneurons of synergist and antagonist muscles are sufficiently potent and appropriately arranged to subserve a variety of reflexes. In animals with all higher centers removed, these segmental circuits may function by themselves to produce simple reflex responses. Under normal conditions, however, the activity of segmental circuits is largely controlled by supraspinal centers. Descending tracts arise from two major supraspinal centers: the cerebral cortex and the brainstem.

Brainstem. The brainstem, which includes medulla and pons, is a major and complex integrating center which combines signals descending from other higher centers, as well as afferent input arising from peripheral receptors. The descending output from brainstem neurons affects motor and sensory cells in the spinal cord. Brainstem centers considerably extend the motor capacity of an animal beyond the stereotyped reflex reactions mediated by the spinal cord. Even without any higher centers, animals with a brainstem can perform integrated activities such as standing, walking, and making appropriate postural adjustments. In contrast to segmental reflexes, these motor responses involve coordination of muscles over the whole body. In decerebrate cats, electrical stimulation of certain brainstem regions evokes a form of locomotion; the rate of walking is influenced by proprioceptive feedback from the limbs. Another major motor function of the brainstem is postural control, exerted via the vestibular nuclei. Sensory receptors in the vestibule of the ear, signaling the orientation of the head in space and changes in head position, have potent influence on the vestibular nuclei, which in turn effectively regulate limb and trunk muscles to respond in a manner that maintains postural stability.

Besides neurons controlling limb muscles, the brainstem also contains a number of important neural centers involved in regulating eye movements. These include motoneurons of the eye muscles and various types of interneurons that mediate the effects of vestibular and visual input on eye movement. For example, one important pathway mediates the vestibular ocular reflex; it transforms vestibular signals, initiated by head movements, to generate compensatory eye movements, with the result that eye position can remain fixed in space. Visual input can generate two different types of eye movements, each mediated by its own neural circuitry: rapid saccadic eye movements, which move the eye as quickly as possible from one visual target to another, and smooth-pursuit eye movements, which follow a slowly moving target. In contrast to limb muscles, which must deal with changing loads, the eye muscles control a fixed and predictable load. This, plus the need for maintaining accurate vision, has led to a highly developed oculomotor system, residing largely in the brainstem.

Cerebellum. Another important coordinating center in the motor system is the cerebellum, an intricately organized network of cells closely interconnected with the brainstem. The cerebellum receives a massive inflow of sensory signals from peripheral receptors in muscles, tendons, joints, and skin, as well as from visual, auditory, and vestibular receptors. Higher centers, particularly the cerebral cortex, also provide extensive input to the cerebellum via pontine brainstem relays. The integration of this massive amount of neural input in the cerebellum somehow serves to smooth out the intended movements and coordinate the activity of muscles. Without the cerebellum, voluntary movements become erratic, and the animal has difficulty accurately terminating and initiating responses. The output of the cerebellum affects primarily brainstem nuclei, but it also provides important signals to the cerebral cortex.

The function of the cerebellum has been likened to a servocontrol mechanism which detects and corrects errors during the course of a movement. Following injury to the cerebellum, movements are ataxic; that is, they suffer from errors in rate, range, force, and direction. Starting, stopping, and changing the direction of movement are especially disturbed; in reaching for an object the hand often misses the target. Voluntary movement is characterized by tremor which increases as the movement progresses. All these disturbances appear to be due to defects, not in generating movement, but in detecting errors and correcting deviations. Thus, if in reaching for an object the hand strays from the correct path, correction is begun too late and proceeds too far, so that the hand overshoots the direct path. A series of such overcorrections results in the oscillating tremor typical of cerebellar injury.

Injury to those portions of the cerebellum which receive afferent input from the vestibular nuclei produces marked disturbances of equilibrium and gait, so that the victim, although showing little uncoordination of movement while lying in bed, is unable to maintain balance and to coordinate limb muscles in walking.

Basal ganglia. At another level of motor system are the basal ganglia. These massive subcortical nuclei receive descending input connections from all parts of the cerebral cortex. Their output projections send recurrent information to cerebral cortex via thalamus, and their other major output is to brainstem cells. Their substantial size suggests that the basal ganglia perform an important motor function, but the precise nature of the function remains unclear. Pathological lesions of the basal ganglia result in two types of motor disturbances, either decreased

motor activity or generation of involuntary movements; such lesions involve no sensory deficits, nor do they impair mental capacities.

Decreased motor activity caused by disruption of basal ganglia, called negative signs, include akinesia (a disinclination to use the affected part of the body), delays and slowing of limb movements, absence of normal postural adjustments necessary for stable standing, and deficits in facial expression. In other cases, basal ganglia disorders produce excessive involuntary motor activity; such positive signs of basal ganglia pathology include muscle rigidity (resistance to passive stretch caused by tonic motor activity), rhythmic muscle tremor, characteristic of parkinsonism, and uncontrollable movements of the hands or feet (alletosis) or the entire limb (ballismus). These motor symptoms can sometimes be alleviated by surgical or chemical treatments affecting the basal ganglia. The fact that both the positive and negative signs of basal ganglia pathology represent a disorder in generation of movements suggests that basal ganglia play a role in the initiation of movements.

Cerebral cortex. At the highest level of the nervous system is the cerebral cortex, which exerts control over the entire motor system. The cerebral cortex performs two kinds of motor function: certain motor areas exert relatively direct control over segmental motoneurons, via a direct corticospinal pathway, the pyramidal tract, and also through extrapyramidal connections via supraspinal motor centers. The second function, performed in various cortical association areas, involves the programming of movements appropriate in the context of sensory information, and the initiation of voluntary movements on the basis of central states. Cortical language areas, for example, contain the circuitry essential to generate the intricate motor patterns of speech. Limb movements to targets in extrapersonal space appear to be programmed in parietal association cortex. Such cortical areas involved in motor programming exert their effects via corticocortical connections to the motor cortex, and by descending connections to subcortical centers, principally basal ganglia and brainstem.

As indicated in Fig. 4, the motor centers are all heavily interconnected, so none really functions in isolation. In fact, some of these connections are so massive that they may form functional loops, acting as subsystems within the motor system. For example, most regions of the cerebral cortex have close reciprocal interconnections with underlying thalamic nuclei, and the corticothalamic system may be considered to form a functional unit. Another example is extensive connection from cerebral cortex to pontine regions of the brainstem, controlling cells that project to the cerebellum, which in turn projects back via the thalamus to cerebral cortex. Such functional loops are at least as important in understanding motor coordination as the individual centers themselves. With this overview in mind, the role of motor cortex in controlling movement will now be considered. The anatomical and physiological information available on the organization of motor cortex also exemplifies the information required to understand other motor centers. *See* BRAIN.

Motor Cortex

The cerebral cortex is an extensive network of nerve cells covering the forebrain. The axons of some cortical cells pass into the underlying white matter and descend through the brainstem to affect the motoneurons of the brainstem or spinal cord. Areas of cortex designated as motor cortical areas have many such corticofugal cells. Electrical stimulation of these motor areas elicits movement, and their excision produces either paralysis or paresis (weakness). Activity of cells recorded in these regions is closely related to limb movements.

Motor areas. Systematic stimulation of the exposed cerebral cortex with electrical pulses reveals several prominent motor areas.

Precentral area. In primates the major one is the precentral motor area, which lies anterior to the central fissure of Rolando. In the monkey it extends over the precentral gyrus (**Fig. 5**); in humans it is largely buried in the anterior wall of the central fissure. Repetitive electrical stimulation within this region can elicit discrete movement, such as flexion of the thumb or a retraction of the lip. The evoked movement is always on the side of the body contralateral (opposite) to the stimulated hemisphere, since the descending pathways cross the midline before reaching the spinal cord. Moreover, the particular muscles activated depend on the stimulated site within the precentral motor area. Beginning at the midline and proceeding laterally along the precentral gyrus, one finds an orderly array of motor points for leg, trunk, arm, and face musculature. The topographical relations may be represented by displaying the body superimposed along the precentral gyrus as shown for the monkey in Fig. 5. In addition to this mediolateral representation of lower to upper

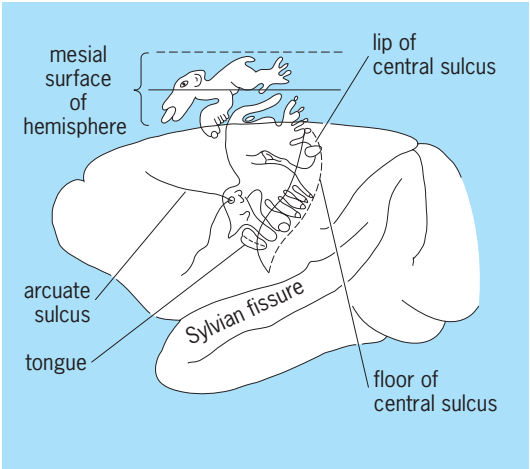

Fig. 5. Brain of monkey, with figurines showing motor representation of body in precentral motor cortex (lower figurine) and supplementary motor area (upper figurine). Parts of the body are drawn superimposed on the cortical regions that regulate their movement. The representation of toes and fingers buried in the precentral bank is shown with the mesial surface of the hemisphere folded out.

body musculature, there is an anterior-posterior organization, with the musculature of the apical portions of the body (fingers, toes, lower lip, and tip of tongue) located posteriorly, mostly in the depths of the central fissure, and the musculature of trunk and back represented more anteriorly, on the free surface of the precentral gyrus. A striking feature of motor cortex organization is the disproportionate area devoted to different muscles. The muscles of fingers, toes, lips, and tongue, which are involved in delicate, precise movements, have relatively large cortical areas devoted to their control. Comparatively smaller cortical areas are devoted to trunk musculature, which explains the distortion of the figurines in Fig. 5.

Precise stimulation of precentral motor cortex sites can sometimes evoke contraction of individual muscles. However, while the threshold response to stimulation may be activation of a single muscle, in most cases such stimuli also evoke subthreshold effects in motoneurons of other muscles. This is due to the fact that the stimuli affect numerous cortical cells projecting to different motoneurons. In the primate, forelimb motoneurons receive direct input from a colony of corticomotoneuronal cells, whose cell bodies are distributed over several square millimeters of cortex. Careful analysis indicates that the cortical colonies converging on motoneurons of different muscles may overlap extensively. In fact, single corticomotoneuronal cells may send divergent terminal projections to motoneurons of several different muscles. This means that the activity of such cells would affect a group of muscles.

Besides having motor output affecting muscles, the cortical cells also receive sensory input from receptors in muscles and skin. The topographic map over precentral cortex in Fig. 5 represents not only the parts of the body that move when the cortex is stimulated, but also the regions of the body from which the cortical cells can be activated. In other words, the sensory map and motor map are superimposed in register: cortical cells receive sensory input from the same regions to which they send motor output. Such input-output relations indicate that normally the cells would be involved in a functional loop interconnecting the cortical site with a peripheral locus.

Premotor and supplementary areas. The precentral motor area is the most elaborate motor cortical area with the most direct output connections to motoneurons. In addition, other motor cortical areas have also been demonstrated. One of these, called the premotor area, is located anterior and adjacent to the primary motor cortex. Cells in the premotor cortex are activated when the monkey prepares to make a movement, suggesting a role in motor programming. Another complete motor representation, the supplementary motor area, lies on the medial surface of the hemisphere (Fig. 5). The threshold for evoking movements by electrical stimulation is higher for supplementary motor area than for precentral motor area; such responses often involve the musculature of both sides of the body.

Postcentral cortex. A fourth significant motor representation is in postcentral cortex. Its topographical organization would be represented by a figurine which is the mirror image of that shown on the precentral cortex in Fig. 5. Like the precentral map, the postcentral map also represents both the sensory input from peripheral receptors as well as motor output to the relevant regions. The precentral and postcentral maps are heavily interconnected by cells making corticocortical connections. Nevertheless, some motor output from postcentral cortex is independent of its connections to precentral cells, since movements can be evoked by stimulating postcentral regions after motor cortex has been removed. The threshold for evoking motor effects from postcentral cortex is higher than from precentral cortex.

Eye and neck movements. Stimulation of the cortex in the region anterior to the arcuate fissure elicits movements of neck and eyes. The typical response consists of conjugate movement of the eyes and turning of the head to the side opposite the stimulated hemisphere. For this reason the area enclosed between the two limbs of the arcuate fissure is often called the frontal eye field. Since the major ascending pathways leading into the cortex cross the midline before reaching the cortex, stimuli applied to the right side of the body generate nerve impulses which reach the left hemisphere. Such impulses activating the motor eye fields on the left would cause movement of the head and eyes toward the right, that is, toward the stimulated side of the body. Such orientational movements may be part of the motor component of attention.

Effects of ablations. Removal of motor cortex within and just anterior to the central fissure causes a partial loss of voluntary movement in the contralateral muscles represented in the ablated cortex. Lesions confined to the leg area cause paralysis of the leg, leaving the arm and face musculature unaffected; thus the topographical organization of motor cortex revealed by stimulation is confirmed by ablation experiments. The paralyzed extremity is flaccid; that is, it hangs limply and displays no resistance to passive flexion or extension. Paralysis of voluntary movement is not permanent, however; the duration and severity of the deficit vary with species, being more severe in the human, ape, and monkey than in lower mammals such as dog and cat. This species difference reflects the increased dominance of the cerebral cortex in phylogenetically advanced forms. Following ablation of the arm motor cortex in monkeys and chimpanzees, movement at the shoulder reappears within a few days. Later, elbow and wrist movements reappear, and finally (8–12 weeks in chimpanzees) crude movements of fingers are restored. However, finger movements never regain the preoperative delicacy; thumb-finger approximation such as occurs in grooming remains permanently defective and awkward. This reflects the critical role of cortex in control of the finger muscles, which have an extensive cortical representation.

Following isolated unilateral ablation of the frontal eye fields in monkeys, the animal's head and eyes

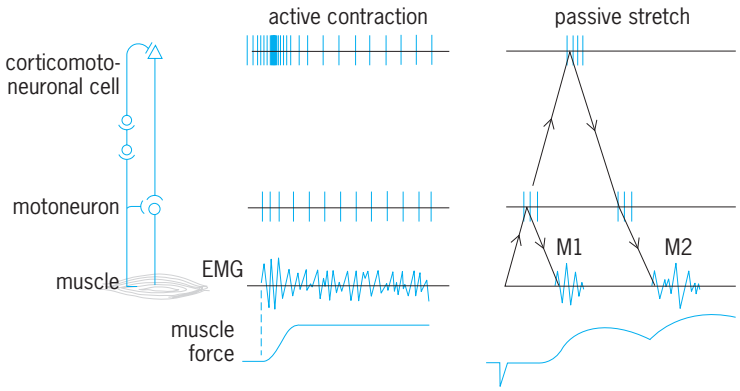

Fig. 6. Response of corticomotoneuronal cell and target muscle during active muscle contraction and in response to passive stretch of muscle. Diagram at left shows neural pathways: input from muscle stretch receptors affects motoneuron and corticomotoneuronal cell; the corticomotoneuronal cell also receives centrally originating input during active movement, and projects to motoneuron. Passive muscle stretch evokes impulses conducted to motoneuron via segmental reflex, producing M1 response, and to corticomotoneuronal cell, contributing to M2 response via transcortical loop.

turn toward the side of the lesion; there is transient paralysis of conjugate deviation of the eyes toward the opposite side. In walking, the animal tends to turn toward the side of the lesion; that is, it tends to follow the eyes and head, so that locomotion is circular. Such circling gait may persist after paralysis of the eye muscles has disappeared.

Activity of cortical neurons. The activity of single neurons recorded in monkeys trained to make specific movements further confirms the role of cerebral cortex cells in initiating and controlling movement. The location of motor cortex neurons which discharge in relation to active movement of particular parts of the body is again represented by the topographic map in Fig. 5. When the monkey initiates a simple hand movement—for example, pressing a key when a light is illuminated—many cells in the precentral hand region become active well before the onset of agonist muscle activity; many neurons continue to fire as long as the related limb muscles are active.

In monkeys making wrist movements requiring different degrees of active force, and independent wrist displacement, experimenters found that motor cortex cell activity was more closely related to the active force of a movement than the resultant limb position. **Figure 6** illustrates the typical response pattern of a corticomotoneuronal cell during an active wrist movement. Such corticomotoneuronal cells facilitate the activity of one or more target muscles. At onset of movement, the cell discharges a burst of activity beginning well before onset of activity in its target muscles. During the static hold period most corticomotoneuronal cells fire tonically at a steady rate; this tonic firing rate was found to increase in proportion to the amount of active force exerted. Thus, the firing pattern of a corticomotoneuronal cell shows a contribution to maintaining steady force during the hold period, as well as an additional burst related to initiating a change in force.

Besides firing during active movement, corticomotoneuronal cells also respond to passive limb movements, typically passive joint movements that stretch their target muscles. Sensory input from stretch receptors in the same muscles which the corticomotoneuronal cell facilitates would make functional sense, since it would help compensate for load changes. A sudden increase in load during an active movement, for example, would stretch the agonist muscles, and would activate corticomotoneuronal cells whose output in turn would facilitate activity of those muscles, tending to overcome the increased load. Such a transcortical load compensation reflex is entirely analogous to the segmental stretch reflex mediated by direct input to motoneurons from stretch receptors.

Figure 6 also illustrates the responses evoked by a sudden stretch of a muscle. The first, short-latency electromyogram (EMG) response (Ml) is mediated by the segmental stretch reflex—the direct connection from stretch receptors to motoneurons. The second EMG response (M2) is mediated by long-loop reflexes, including those through motor cortex. In contrast to the spinal circuit, the gain of the transcortical loop could be changed by the state of higher centers. Indeed, the response of cortical cells to peripheral stimulation is enhanced when the animal is prepared to make a movement in which the cells are involved. Corticomotoneuronal cells with monosynaptic connections to motoneurons are an important component of the pyramidal system, which controls spinal cells directly.

Pyramidal tract. One of the most important corticofugal pathways is the pyramidal tract, or corticospinal tract, which originates from cortical cells and runs without interruption to the spinal cord. The human medullary pyramid contains about 1 million fibers, mostly of small diameter. Almost half of the pyramidal tract originates from neurons in the precentral motor areas. Some of these are the giant pyramidal cells of Betz, large conical cells which histologically distinguish precentral motor cortex. In humans, the motor area of each hemisphere contains about 34,000 Betz cells, enough to account for about 2% of the pyramidal fibers, presumably those of large (11–20-micrometer) diameter. The supplementary motor area also contributes some fibers to the pyramidal tract. The rest of the pyramidal tract originates in the postcentral gyrus, posterior to the central fissure.

The axons of the pyramidal tract leave their cells of origin in cortical layer 5 to enter the subcortical white matter and then pass through the internal capsule, the cerebral peduncles, and the pons (**Fig. 7**). Some fibers terminate in the brainstem on motoneurons of cranial nerve nuclei. Below the pontomedullary junction the tract continues on the ventral surface of the medulla as a recognizable band of fibers called the pyramid. At the caudal border of the medulla most of the fibers cross to the opposite side of the neuraxis to descend in the dorsolateral column of the spinal cord. Some pyramidal fibers do not decussate but continue either in the ventral white funiculi or in the lateral column of the same side. The spinal portion of the corticospinal tract extends throughout the full extent of the spinal cord

but progressively diminishes in size as it passes from cervical to lumbar segments. Terminations are particularly numerous in the cervical and lumbar enlargements, which contain the motoneurons supplying the musculature of upper and lower limbs.

The role of the pyramidal tract in movement in monkeys and chimpanzees has been studied by sectioning the medullary pyramid and observing the resulting changes in motor performance. The most prominent defect following unilateral pyramidotomy is contralateral paresis involving the musculature from the neck down. The affliction is more severe in chimpanzees than in monkeys; in the former, even stereotyped movements of progression are impaired but not abolished. In neither animal is paralysis so grave as to render the affected parts useless, but there is severe impairment of voluntary movement and loss of such fine movements as opposition of thumb and index finger in grooming and individual movements of the fingers in manual exploration. This type of deficit has been observed to persist up to 4 years after operation, and thus may be considered permanent. Associated with the paresis is flaccidity like that seen following precentral cortical ablations; the extremities are limp and offer no resistance to movement at the joints.

Extrapyramidal systems. Isolated destruction of the corticospinal tract does not cause complete loss of voluntary movement. Moreover, electrical stimulation of the motor cortical areas after pyramidotomy still elicits muscular contraction. Therefore, the cortex must give rise to motor pathways other than the pyramidal tract. These other pathways are often referred to collectively as cortically originating

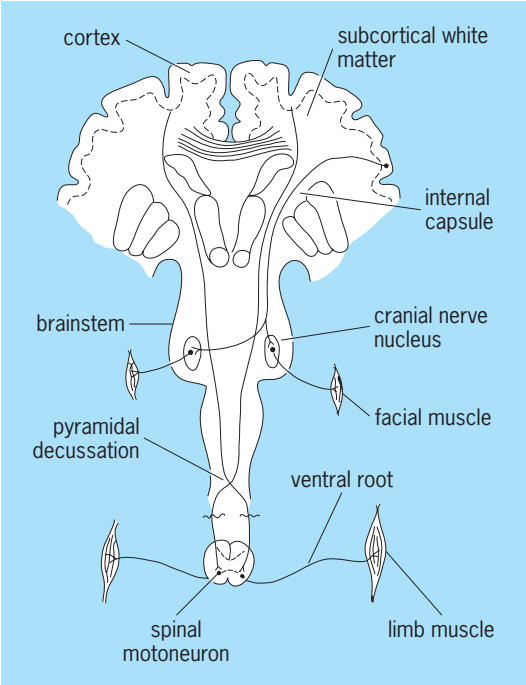

Fig. 7. Diagram of pyramidal system, showing origin and course of corticospinal (pyramidal) tract. In a similar manner, corticobulbar fibers affect cranial motoneurons of facial muscles.

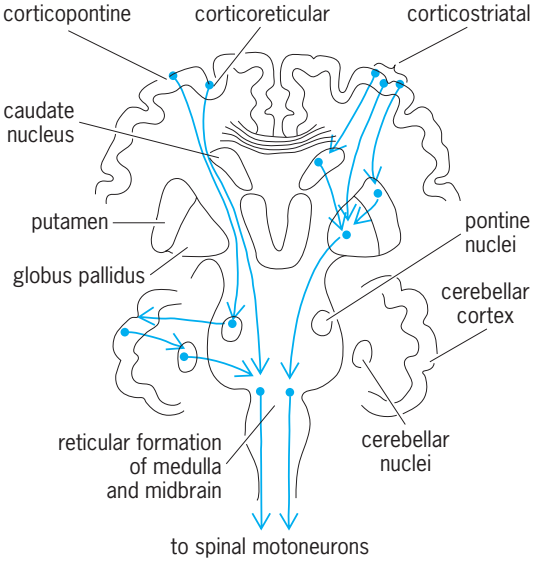

Fig. 8. Diagram of extrapyramidal systems, showing major descending connections from cerebral cortex.

extrapyramidal pathways. Although extrapyramidal projections probably arise from nearly all portions of the cortex, the precentral motor area appears to give rise to particularly large contributions. **Figure 8** illustrates the anatomic relation of some of the better-known extrapyramidal systems.

Corticostriatal and corticopallidal systems. These two systems originate from rostral precentral cortex and project to the caudate nucleus and the putamen, which are portions of the basal ganglia. From the basal ganglia, impulses are relayed to the brainstem, which in turn affects the spinal levels:

Corticoreticular systems. These systems originate from the cortex around the central fissure, especially from the motor area. The axons terminate in the region of the pons and medulla on the diffusely organized neurons constituting the brainstem reticulum. The projection is bilateral and poorly, if at all, organized somatotopically. Impulses are presumably relayed to the spinal cord via the reticulospinal tracts which traverse the ventral and lateral white columns of the cord. *See* RETICULAR FORMATION.

Corticopontine systems. These systems arise from the cortex of each of the four lobes of the brain, although the contribution from the rostral precentral gyrus is most prominent. A significant contribution also comes from the supplementary motor cortex on the mesial surface of the hemispheres. The fibers terminate in the pontine nuclei, which in turn project to the cerebellum. From the cerebellum, projection systems feed impulses into the medullary reticulum, which in turn projects through the reticulospinal tracts to the spinal cord as described above. *See* BRAIN; NERVOUS SYSTEM (VERTEBRATE).

Eberhard E. Fetz

Bibliography. E. R. Kandel, J. H. Schwartz, and T. M. Jessel, *Principles of Natural Science*, 4th ed., 2000; V. B. Mountcastel, *Medical Physiology*, 2 vols., 1980; H. D. Patton et al., *Textbook of Physiology*, 1989; R. Porter and R. Lemon (eds.), *Corticospinal Function and Voluntary Movement*, 1993.

Motorcycle

A two- or three-wheeled self-propelled motor vehicle. Power is usually supplied by an internal combustion engine of one or more cylinders with either air or liquid cooling, although electric-powered motorcycles have been built. Transmissions commonly have four or five forward speeds; motorcycles with automatic transmissions have been commercially available, but have not achieved great popularity. Some three-wheeled motorcycles are equipped with a reverse gear. Power is transmitted from the transmission to the rear wheel by either a roller chain, a belt, or a shaft drive. *See* INTERNAL COMBUSTION ENGINE.

Motorcycles are classified by engine displacement and by type of use. Vehicles powered by engines of 50 cm^3 or lower displacement are classified as mopeds; vehicles powered by engines with displacements of 700 cm^3 and above are classified as heavyweights. Vehicles between these displacements are referred to as lightweight or medium-weight motorcycles depending on where they are in the range. Motorcycles are generally used for recreation or transportation on paved roads. Some light- and medium-weight motorcycles are specifically designed for off-road use.

Recreational motorcycles are further classified as sports or touring vehicles depending on their primary intended use. A typical two-wheeled modern touring motorcycle (see **illus.**) is designed to carry a driver and passenger and is equipped with saddlebags and a tour pack for carrying clothing and supplies, as well as a fairing and windshield for protection from the elements. It is not unusual to find a touring motorcycle also equipped with a stereo radio, a cassette player, a citizens' band radio, and other accessories. As the illustration shows, a two-wheeled motorcycle generally has its engine mounted in a tubular frame and has a rear wheel on a swing arm which can rotate relative to the frame in a vertical plane but which is restrained by one or two coil springs and hydraulic shock absorbers. The front wheel is generally mounted in a fork which contains two telescopically operating spring hydraulic damper systems. *See* SHOCK ABSORBER.

A motorcycle engine is started either by kicking a foot-operated cranking mechanism or by activating an electric starter to rotate the crankshaft. Operating controls are standardized so that the throttle is activated by twisting the grip on the right handlebar, the clutch lever is located on the left handlebar, the gear-shift lever is activated by a foot pedal on the left side of the motorcycle, and the front and rear brakes are operated by hand and foot levers respectively on the right side of the motorcycle. Brakes are of either the disk or drum type, with disk brakes predominating on high-performance motorcycles. A two-wheeled motorcycle is steered by turning the handlebars and leaning in the direction of the turn. The steering of three-wheeled motorcycles or two-wheeled motorcycles equipped with a sidecar requires different techniques. *See* AUTOMOTIVE BRAKE.

J. L. Bleustein

Touring motorcycle. (*Harley-Davidson Motor Co., Inc.*)

Bibliography. J. Bennett, *The Complete Motorcycle Book*, 2d ed., 1999; W. N. Crouse and D. L. Anglin, *Motorcycle Mechanics*, 1982; Society of Automotive Engineers, *Motorcycle Dynamics and Rider Control*, Spec. Publ. 428, February 1978.

Mountain

A feature of the Earth's surface that rises high above its base and has generally steep slopes and a relatively small summit area. Commonly the features designated as mountains have local heights measurable in thousands of feet or many hundreds of meters, lesser features of the same type being called hills, but there are many exceptions. *See* HILL AND MOUNTAIN TERRAIN.

Mountains rarely occur as isolated individuals. Instead they are usually found in roughly circular groups or massifs, such as the Olympic Mountains of northwestern Washington and the Harz Mountains of northern Germany, or in elongated ranges, like the Sierra Nevada of California, the Bighorn Range of Wyoming, or the Sierra de Guadarrama of central Spain. An array of linked ranges and groups, such as the Rocky Mountains, the Alps, or the Himalayas, is a mountain system. North America, South America, and Eurasia possess extensive cordilleran belts, within which the bulk of their higher mountains occur. *See* CORDILLERAN BELT; MASSIF; MOUNTAIN SYSTEMS.

As a rule, mountains represent portions of the Earth's crust that have been raised above their surroundings by upwarping, folding, or buckling, and have been deeply carved by streams or glaciers into their present surface form. Some individual peaks and massifs have been constructed upon the surface by outpourings of lava or eruptions of volcanic ash. *See* VOLCANO.

Edwin H. Hammond

Bibliography. D. R. Butler, S. J. Walsh, and G. P. Malanson (eds.), *Mountain Geomorphology: Integrating Earth Systems*, 2003; J. Gerrard, *Mountain Environments: An Examination of the Physical Geography of Mountains*, 1990; P. N. Owens and O. Slaymaker (eds.), *Mountain Geomorphology*, 2004; M. Price, *Mountains: Geology, Natural History, & Ecosystems*, 2002.

Mountain meteorology

The effects of mountains on the atmosphere, ranging over all scales of motion, including very small (such as turbulence), local (for instance, cloud formations over individual peaks or ridges), and global (such as the monsoons of Asia and North America).

Lee waves. The most readily perceived effects of a mountain, or even of a hill, are related to the blocking of air flow. When there is sufficient wind, the air either goes around the obstacle or over it, causing waves in the flow similar to those in a river washing over a boulder. Since ascending air cools by adiabatic expansion, the saturation point of water vapor may be reached in such waves as they form over an obstacle, and a cloud then forms in the ascending branch of the wave motion. Such a cloud dissipates in the descending branch where adiabatic warming takes place, with a sometimes spectacular display of lenticular cloud formations (**Fig. 1**). *See* CLOUD; CLOUD PHYSICS.

Fig. 1. Lenticular cloud, southwest of Boulder, Colorado, October 24, 1980, 17:50 MST. (*Ron Holle, NOAA/ERL, Boulder*)

The shapes and amplitudes of these lee waves (they form over and to the lee of mountains) depend not only on the thermal stability and on the vertical wind shear in the overlying atmosphere but also on the shape of the underlying terrain. Such waves have been modeled successfully by towing obstacles through a tank in which salt solutions of different concentrations are stacked upon each other, mimicking various degrees of atmospheric stability (**Fig. 2**). They also have been modeled numerically in computers. *See* SIMULATION; WAVE (PHYSICS).

Mountain torque. On a grander scale, mountain ranges, such as the Sierras of North and South America, place an obstacle in the path of the westerly winds (that is, winds from the west), which generally prevail in middle latitudes. Such a blockage tends to generate a high-pressure region upwind from the mountains (this may be viewed as air piling up as it prepares to jump the hurdle), and a low-pressure area downwind. Thus, there is a stronger push against the mountains on the high-pressure western side than on the low-pressure eastern side. The net effect between these stronger and weaker pushes would be to move the mountains toward the east. However, since the mountains are firmly attached to the Earth, an extremely slight acceleration of the Earth's rotation is a result. While minute, this acceleration can be measured by atomic clocks as a variation in the length of day. Since the strength of the westerlies causing this effect varies from day to day, there will be slight variations in the length of day smaller than milliseconds. This variation is known as the mountain torque effect. As this torque speeds up the rotation of the Earth, it will slow down the rotation of the atmosphere, because the total rotational momentum of the earth-atmosphere-ocean system is conserved as the system spins in space. Thus, slowing down the atmospheric flow is precisely what a mountain obstacle does. *See* TORQUE.

Planetary waves. Less subtle than mountain torque effects are the large-scale meanders that develop in the global flow patterns, once they have been perturbed, mainly by the North and South American Andes and by the Plateau of Tibet and its Himalayan mountain ranges. All other mountain ranges on Earth are dwarfed by these gigantic, crustal upheavals. These meanders in the large-scale flow are known as planetary waves. They appear prominently in the pressure patterns of hemispheric or global weather maps. They can also be seen from space as they manifest themselves in large-scale swirls of cloud formations (**Fig. 3**). *See* WEATHER MAP.

Diurnal heating effects. There are important thermal effects attached to mountains in addition to their serving as large obstacles thrown into the path of the atmosphere, a moving fluid. As satellite measurements have shown, surface temperatures of the Earth do not change much with terrain elevation. They depend mostly on the exposure angle to the Sun, the vegetation cover, and the type of soil. Surface temperatures vary dramatically, however, between snow-free and snow-covered ground, mostly because of changes in the reflectivity, or albedo. Thus, the snow-free ground surface temperatures during the day in the Rocky Mountains or in Tibet are not much different from those in the adjacent plains. At night, however, the thinner and drier mountain

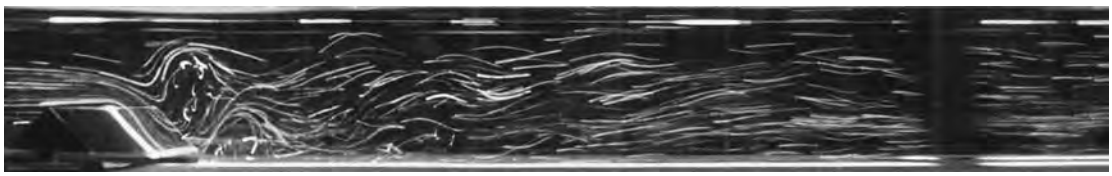

Fig. 2. Lee waves produced by towing an obstacle through a Plexiglas channel filled with stratified salt solutions of various densities, mimicking a stable atmosphere. (*D. L. Boyer and L. Tao, University of Wyoming*)

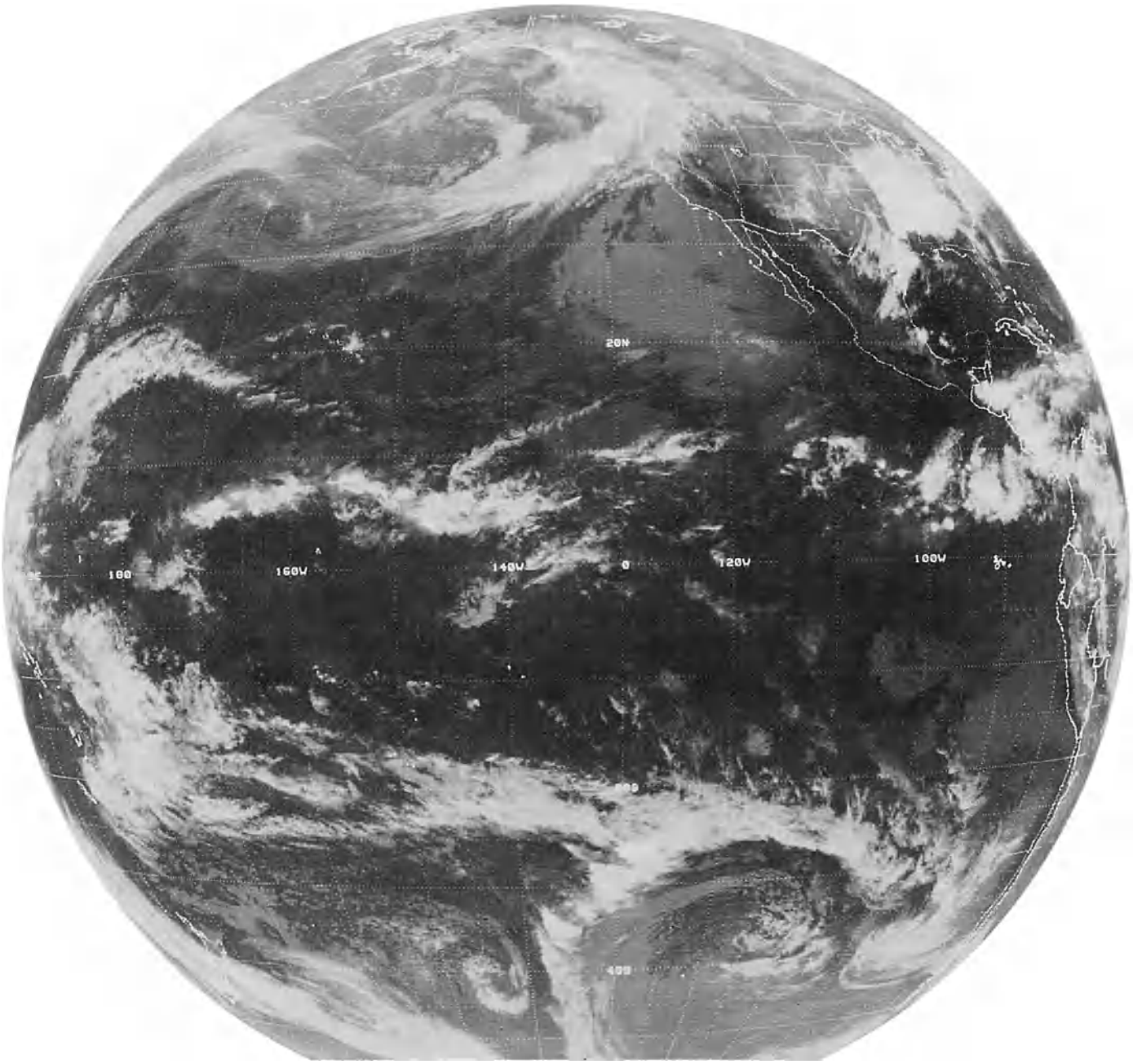

Fig. 3. Infrared image taken by the *GOES West* satellite on May 4, 1979, at 2:45 GMT. The large swirls of clouds in middle latitudes of both hemispheres indicate the position of planetary waves in the jet stream, and of cyclones. (*NASA*)

atmosphere lets more infrared radiation escape to space than does the thicker and moister atmosphere over the plains. Therefore, the ground surface in the mountains cools off more readily than in the plains, causing greater diurnal variations in surface temperatures of elevated terrain. *See* ALBEDO; TERRESTRIAL RADIATION.

The free atmosphere at the same height or pressure levels corresponding to the elevations of adjacent mountain ranges remains cooler during the day, as it is farther away from the surface heat source; and it is warmer during the night, since air does not emit radiation as readily as does the solid surface of the ground. Therefore, during the day there will be a gradient of temperature from the warm mountain surface to the cooler atmosphere at the same altitude over the plains. That gradient reverses at night. As a consequence, during the day a so-called heat low tends to develop over mountain ridges; and it reverses to anticyclonic, high-pressure conditions at night. Since air tends to flow from high to low pressure, especially if the horizontal space dimensions and the time scales involved do not allow the deflecting force of the Earth's rotation (the Coriolis force) to become fully effective, there is flow toward the mountains (valley breezes) during the day and away from the mountains (mountain breezes) at night. *See* CORIOLIS ACCELERATION; CYCLONE.

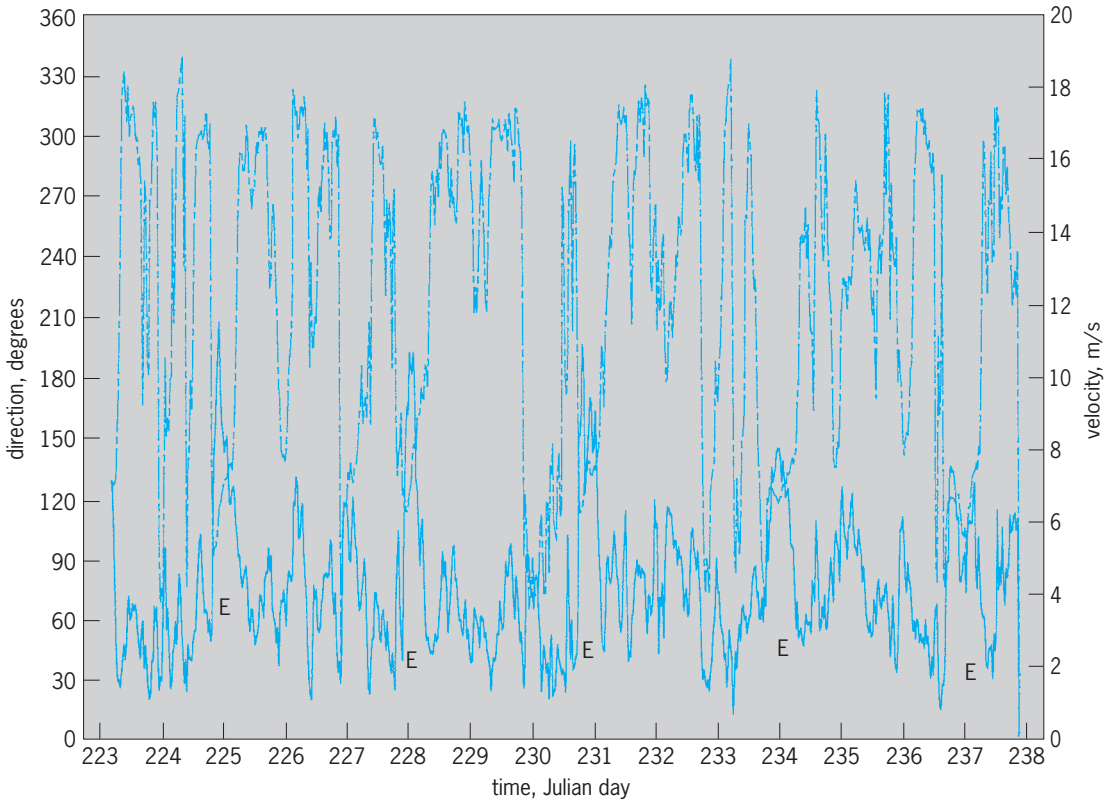

Fig. 4. Wind direction and speed measurements at the peak of Mount Werner, Colorado, during August 1984. During the day, winds tend to be from the northwest; in the evening, they tend to shift to the southeast, at times resulting in rather strong events (E) that might endanger small aircraft flying close to the peak. The layer in which these strong winds are contained typically extends to about 0.6 mi (1 km) above mountain peak elevation. 1 m = 3.3 ft.

Plateau circulation system. A circulation system similar to that generated by diurnal heating effects but on a grander scale develops if large plateaus, such as the Tibetan Plateau or the Rocky Mountains, are involved in the diurnal heating and cooling cycle. The resulting plateau circulation system over North America extends as far as the Mississippi Valley during summer, when this diurnal cycle is best developed. As the air flows toward the mountains, thunderstorms tend to develop over the major ranges and peaks, with a frequency maximum during the early afternoon. With the reversal of the temperature and pressure gradients at night, these thunderstorms tend to collapse, but new ones form at the leading edge of the air mass now flowing away from the mountains. Perhaps it is a result of this circulation that the Great Plains west of the Mississippi show a nocturnal thunderstorm maximum, as late as past midnight. *See* AIR MASS; THUNDERSTORM.

Because the more intense daytime heating of the mountains, as compared to the free atmosphere over the plains, is most prominent in summer, the tendency of air flowing toward the large plateaus and ascending along their periphery also favors that season. Thus the large monsoonal weather systems of Asia and also of North America are caused, at least in part, by the heating effects of the adjacent, large and high plateaus.

Latent heat effects. There is an additional mountain effect associated with the summer thunderstorm and monsoon activities. The ground surface temperature rises rapidly in the morning, especially during summer, producing the temperature and pressure gradients described above, with the resulting wind systems. As air converges toward the mountains and ascends along their slopes, clouds soon form that shade the ground from further insolation and prevent further heating of the soil. However, once cumulus cloud formation has started, especially during the summer thunderstorm season, the latent heat of condensation of water vapor makes the interior of the clouds warmer than their environment, causing them to grow until they develop into full-fledged thunderstorm systems. As heavy rain starts falling from these systems in the late afternoon, its evaporation contributes to the rapid cooling of the subcloud layer, causing a rapid reversal from low-pressure conditions with inflow toward the mountains to high-pressure conditions with flow away from the mountains (**Fig. 4**). *See* AIR PRESSURE; INSOLATION; PRECIPITATION (METEOROLOGY).

Monsoons. It appears that the latent heat cycle—warming by condensation and cooling by evaporation—assumes a dominant part in the formation of the diurnal plateau circulation cycle. A similar concept holds for the monsoons of Asia and North America on a seasonal, rather than diurnal, basis. For instance, the Tibetan Plateau heats in spring; it attracts inflow into a heat-generated low-pressure system near the Earth's surface,

compensated by outflow aloft from an anticyclone that takes up summer residence over the plateau. The steadiness and vigor of this pressure distribution and its attendant monsoonal circulation system, however, is maintained in large part, by the release of latent heat in the torrential rains that fall over India and southwest China whenever the monsoon is at its peak.

It is well known that the major monsoon circulations interact with the global circulation, shaped in part by sea-surface temperature anomalies in the equatorial Pacific. This is why, in certain years, the monsoon rains tend to fail. The various aspects of mountain meteorology, therefore, have to be viewed within the larger picture. There is a continuous interaction between the weather effects on all space and time scales generated by the mountains and the weather patterns that prevail elsewhere on the Earth. Thus, what is occurring over Tibet has an effect on weather over the Pacific and North America, and vice versa; all weather effects are part of a large, complicated system that is only beginning to be elucidated. *See* ATMOSPHERIC GENERAL CIRCULATION; METEOROLOGY; MONSOON METEOROLOGY; WIND.

Elmar R. Reiter

Bibliography. R. G. Barry, *Mountain Weather and Climate*, 2d ed., 1992; E. R. Reiter (ed.), GARP-ALPEX: The Alpine Experiment, *Meteorol. Atm. Phys.*, 36:1–296, 1987; E. R. Reiter et al., ROMPEX-85, The Rocky Mountain Peaks Experiment of 1985, *Bull. Amer. Meteorol. Soc.*, 68:321–328, 1987; E. R. Reiter et al., Tibet revisited: TIPMEX-86, *Bull. Amer. Meteorol. Soc.*, 68:607–615, 1987; E. R. Reiter and M.-C. Tang, Plateau effects on diurnal circulation patterns, *Mon. Weath. Rev.*, 112:638–651, 1984.

Mountain systems

Long, broad, linear to arcuate belts in the Earth's crust where extreme mechanical deformation and thermal activity have been (or are being) concentrated.

Geological properties. Mountain systems in the general sense occur both on continents and in ocean basins, but the geological properties of the systems in continental as opposed to oceanic settings are distinctly different. The mechanical strain in classical, continental mountain systems is expressed in the presence of major folds, faults, and intensive fracturing and cleavage. Thermal effects are in the form of vast volcanic outpourings, intruded bodies of igneous magma, and metamorphism. Uplift and deformation in young mountain systems are conspicuously displayed in the physiographic forms whose bold topographic relief and beauty are so awesome. Where mountain building is presently taking place, the dynamics are partly expressed in warping of the land surface (as in Palmdale Bulge, California) and significant shallow or deep earthquake activity. Locations of ancient mountain systems in continental regions now beveled flat by erosion are clearly disclosed by the presence of highly deformed, intruded, and metamorphosed rocks. Excellent examples of such mountain systems are the Appalachians of eastern North America, the Caledonian system of Norway and the British Isles, the Hercynian of England, France, and Germany, and the Urals of Russia. Additionally, the Precambrian shield provinces of the world contain the roots of ancient mountain systems. *See* VOLCANOLOGY.

Evolution. The generally accepted view among scientists today is that mountain systems evolve through the movements and interference of rigid plates of lithosphere which compose the Earth's crust and upper mantle. The movements and interactions likewise predetermine the history and nature of sediment accumulations which eventually become bound up in the mountain systems. Some plates consist exclusively of oceanic crust (Pacific, Cocos, Nazca, Caribbean, and Philippine), and a few of continental crust (such as Arabia), but the largest are generally some combination of oceanic and continental crust (America, Eurasia, Africa, India, Antarctica, and Somalia). The relation of the present distribution and movements of these plates to active mountain belts and to sedimentation patterns provides the key to understanding the complicated evolution of mountain systems in the past. *See* LITHOSPHERE.

Oceanic systems. Two basic classes of oceanic mountain systems exist. A world-encircling oceanic rift mountain system has been built along the extensional tectonic boundary between plates diverging at rates of 0.8–2.4 in./year (2–6 cm/year) from the mid-oceanic ridges. This rift mountain system, exposed to partial view in Iceland, is about 36,000 mi (60,000 km) long, averages 3200 mi (2000 km) in breadth, and displays a topographic relief of 6600 ft (2000 m). Rocks along the mid-oceanic rift mountain system are mainly faulted basaltic volcanics. The ridge crest of the mountain system, marking the locus of generation of oceanic crust, is offset repeatedly by transform faults. *See* MARINE GEOLOGY.

The second type, island arc mountain systems, occurs in oceanic basins where the crust dives downward at trench sites, thus underthrusting adjacent oceanic crust. The process results in imbricate faulting, folding, gravity sliding, metamorphism, and partial melting of the crust to produce magmas and volcanic piles. Rates of descent are on the order of 2–4.7 in./year (5–12 cm/year). Such island arc mountain building is now taking place in the western Pacific from the Aleutian arc southward through the belt of archipelagos which includes the Japan, Mariana, Philippine, and New Hebrides arcs. The topographic relief of such island arc mountain systems is staggering, in some places exceeding 33,000 ft (10,000 m). All traces of ancient rift and island arc mountain systems which once existed in the ocean basins have been completely destroyed by the dynamics of sea-floor spreading and the coordinated descent (subduction) of oceanic crust into the Earth's mantle. In effect, these mountains, and the rocks of which they were composed, have been recycled. *See* OCEANIC ISLANDS.

Classical systems. The classical, conspicuous mountain systems of the Earth occur at the

continent/ocean interface, for this is the site where plate convergence has led to major sedimentation, subduction of oceanic crust under continents, collision of island arc mountain systems with continents, and head-on collision of continents. Less commonly, the continent–ocean interface may be the site of transform faulting (as in the San Andreas Fault). The South American Cordillera or Andean mountain system has resulted from stresses generated during underthrusting of South America by oceanic crust of the Pacific plate. Melting of the subducted oceanic plate has led to outpourings of volcanic materials which constitute the volcanic edifice of the Andes. The Alpine, Apennine, Carpathian, and Himalayan systems define an east-west-trending Mediterranean system which has resulted from convergence and collision of the African and Indian plates against Eurasia. The ancient Tethys Sea was closed (repeatedly) by the northerly relative movement of Africa, thrusting oceanic crust and sediment onto the European platform. The leading edge of the Indian plate crashed at a rate of 6 in./year (15 cm/year) into Eurasia, producing the spectacular Himalayan system. *See* CORDILLERAN BELT; FAULT AND FAULT STRUCTURES.

Mountain systems then result from mechanical and thermal responses to sometimes complicated and oftentimes superimposed (and perhaps fortuitous) plate movements. The Cordillera of western North America is one of the best examples of a mountain system of long and complex history. Its unusual evolution is partly expressed in the number and diversity of its components (Coast Ranges, Sierra Nevada, Basin and Range, Colorado and Columbia plateaus, Rocky Mountains, and Rio Grande Rift). The geology of the Cordillera mountain system is the product of superimposed rift, compressional, and extensional movements, related mainly to the interaction of the American plate with oceanic plates to the west. *See* OROGENY; PLATE TECTONICS.

George H. Davis

Bibliography. J. B. Bird (ed.), *Plate Tectonics*, rev. ed., 1980; J. Erickson, *Plate Tectonics: Unraveling the Mysteries of the Earth*, 1992, reprint 2001; X. LePichon et al., *Plate Tectonics*, 1976; A. Miyashiro et al., *Orogeny*, 1982; R. B. Singh, *Dynamics of Mountain Geo-Systems*, 1992; S. P. Sychanthavong (ed.), *Crustal Evolution and Orogeny*, 1992; J. T. Wilson, A new class of faults and their bearing on continental drift, *Nature*, 207:343–347, 1965.

Mouth

The oral or buccal cavity and its related structures. The oral cavity forms in the embryo from an inpocketing of the skin, the stomodeum; it is thus lined by ectoderm and is not, properly speaking, part of the digestive tract. Functionally, however, the mouth forms the first portion of both the digestive and respiratory systems. Various special structures are found in, or associated with, the mouths of most vertebrates. *See* DIGESTIVE SYSTEM; RESPIRATORY SYSTEM.

Teeth may be present to help grasp or grind food. In most vertebrates they are relatively simple cones but in some, especially mammals, they are of diverse shapes. Teeth may be present only along the margins of each jaw, as in mammals, or they may be more widely distributed throughout the mouth; they may be set into sockets or fused to the skeleton of the jaws in various ways. *See* DENTITION; TOOTH.

Various glands, salivary glands in a broad sense, are associated with the mouth. These are of infrequent occurrence in fish but are found in most tetrapods. Humans have three pairs of salivary glands: the parotid, submaxillary, and sublingual. In forms such as some snakes salivary glands may produce a poison used to subdue prey.

Other structures also vary greatly. Most tetrapods have a mobile tongue attached to the floor of the mouth, but few fish do. The structure of the roof of the mouth, or palate, is quite different in different groups. Cyclostomes, the lampreys and hagfishes, have especially highly modified oral cavities with "teeth" and a "tongue" quite unlike those of other vertebrates. *See* PALATE; TONGUE.

In mammals, including humans, the margins of the lips mark the junction between the outer skin and the inner mucous lining of the oral cavity. The mucosa of the mouth forms the lining and the gums surrounding the teeth and covers the surface of the tongue. The roof of the mammalian mouth consists of the hard palate and, behind this, the soft palate which merges into the oropharynx. The lateral walls consist of the distensible cheeks, and the floor is formed principally by the tongue and the soft tissues that lie between the two sides of the lower jaw, or mandible.

The posterior limit of the oral cavity of mammals is marked by the fauces, an aperture which leads to the pharynx. On either side of the fauces are two muscular arches covered by mucosa, the glossopalatine and pharyngopalatine arches; between them lie masses of lymphoid tissue, the tonsils. Suspended from the posterior portion of the soft palate is the soft retractable uvula. *See* MOUTH DISORDERS; TONSIL.

Thomas S. Parsons

Mouth disorders

The mouth, or oral cavity, which comprises the lips, tongue, teeth, gums, and related structures, is subject to a large number of disease processes. Periodontitis, an inflammatory disease of the tissues supporting and surrounding the teeth, and dental decay are the most common diseases; together they account for almost all tooth loss. Other diseases of the mouth can be classified as cysts; diseases of the salivary glands; keratotic, inflammatory, ulcerative, and proliferative lesions; oral infections; and an unusual form of rapidly destructive periodontitis observed in some patients with acquired immune deficiency syndrome (AIDS).

Cysts. A cyst is a cavity lined with epithelium and filled or partly filled with fluid. Cysts of the mouth are of odontogenic or nonodontogenic origin. The periapical cyst is the most common odontogenic type

and forms at the apex of the tooth root. Its formation is an attempt by the body to wall off necrotic material emanating from the root of an infected tooth. A follicular cyst arises from the epithelium of the enamel organ of an unerupted tooth and is seen most often around third molars (wisdom teeth). Cysts of nonodontogenic origin develop from epithelium which has been sequestered in bone suture lines during embryonic development, and are named according to their location. The most commonly occurring are nasoalveolar, nasopalatine, globulomaxillary, and incisive canal cysts. Cysts gradually enlarge and may lead to extensive bone destruction. They are treated by surgical excision.

Neoplasms. The oral cavity is made up of many varieties of tissue, each of which can give rise to neoplasms. Thus, several types of malignant and benign neoplasms can be found in and around the mouth. The benign tumors most often seen are fibromas, papillomas, and hemangiomas. Fibromas are tumors of collagenous connective tissue which appear clinically as smooth, elevated areas with broad bases. A papilloma is made up of a thin, connective tissue stalk covered with a thick layer of epithelium. Clinically, this lesion exhibits a verrucous (warty) appearance. Hemangiomas are benign neoplasms of newly formed blood vessels.

Oral squamous-cell carcinoma (epidermoid carcinoma) accounts for over 90% of all malignancies of the mouth and about 5% of all malignancies of the entire body. Its most common site is the lower lip, and it is more common in males than in females. Lesions of the lip are associated with exposure to sunlight, and those in the oral cavity with tobacco use and consumption of alcoholic beverages. Even though leukemia is not a primary tumor of the oral cavity, it can be included among the important mouth disorders because swelling and inflammation of the gingiva may be the first clinical signs of the presence of the disease. *See* ONCOLOGY.

Salivary gland diseases. The major salivary glands are the parotids, submaxillaries, and sublinguals. Mumps is a viral infection of the secretory cells of these glands and occurs most often in one or both parotid glands. The salivary glands can become infected by retrograde passage of bacteria through the secretory duct, causing acute or chronic sialadenitis. Salivary ducts can also become blocked by calculi (stones), resulting in the accumulation of secretions and swelling. This lesion is called a ranula. The most frequent neoplasm is a mixed tumor. Immunosuppression, irradiation, and administration of cytotoxic drugs tend to decrease salivary flow and to result in enhanced caries and periodontal disease and in painful inflammatory lesions of the oral mucosa. *See* ORAL GLANDS.

White and keratotic lesions. Many disorders of the mouth appear clinically as white or hyperkeratotic lesions. Lichen planus is a dermatologic disease which also occurs in the mouth, on the tongue (keratosis linguae), or on the lips (keratosis labialis) as smooth, lacy networks of white lines or, less commonly, as white patches. The patches may become ulcerative. The cause of the disease is not known. Leukoplakia is the most important white lesion of the mouth since it is considered to be premalignant. It appears clinically as granular white patches on the oral mucosa. Only a small percentage of such lesions progress to carcinoma. These lesions exhibit an accumulation of excessive keratin and an alteration in the pattern of maturation of the epithelium. An inflammatory infiltrate develops in the underlying connective tissue. Traumatic injury and burns can also appear clinically as white lesions.

Oral infections. Gingivitis and periodontitis, as well as dental caries, are a result of oral bacterial infections. Mycotic infection by *Candida albicans* is recognized clinically by the presence of red, inflamed mucosa bearing colonies of white fungi with the appearance of milk curds. The disease is called candidiasis or thrush, and it occurs most commonly in children, in adults after extensive antibiotic therapy, and in some patients with acquired immune deficiency syndrome (AIDS).

Infection of the mouth and lips by the virus herpes simplex type I is possibly the most common human viral infection, next to the common cold. Initially, the lesions appear as small blebs or blisters which rupture to form open ulcers. The ulcers are extremely painful and can involve the entire mouth. The disease is self-limiting and usually disappears within 1–2 weeks. *See* HERPES.

Ulcers of the mouth can also be caused by blood dyscrasias, lichen planus, and several other dermatologic diseases.

Health factors. Generalized systemic disease may play a large role in the development of mouth disorders. Vitamin deficiency can result in glossitis (inflammation of the tongue); and diabetes mellitus, stress, and abnormalities in blood leukocytes may predispose the individual to development of gingivitis and periodontitis. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); DIABETES; PERIODONTAL DISEASE; VITAMIN.

Roy C. Page

Bibliography. J. A. Regezi and J. J. Sciubba, *Oral Pathology*, 3d ed., 1999; W. G. Shafer, M. G. Hine, and B. M. Levy, *A Textbook of Oral Pathology*, 4th ed., 1983.

Moving-target indication

A method of presenting pulse-radar echoes in a manner that discriminates in favor of moving targets and suppresses stationary objects. Moving-target indication (MTI) is almost a necessity when moving targets are being sought over a region from which the ground clutter echoes are very strong. The most common presentation of the output of a radar with MTI is a plan position indicator (PPI) display. The moving targets appear as bright echoes, while ground clutter is suppressed.

Principles of operation. MTI is based upon the use of the Doppler effect; that is, the carrier frequency of the echo from a target moving toward or away from the radar shifts by an amount proportional to

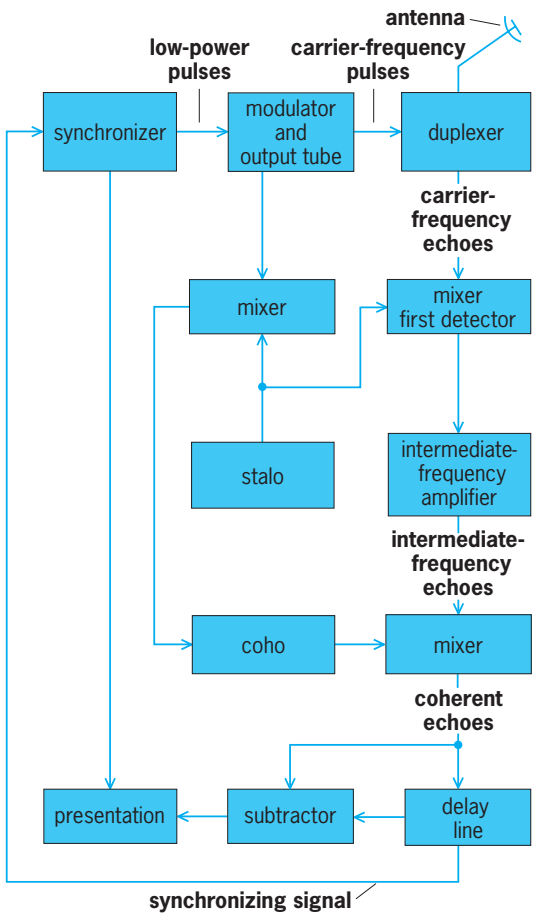

Diagram of radar system with moving-target indication.

the product of radial velocity and transmitted frequency (see **illus.**). A stable oscillator in the receiver is synchronized with the transmitter, providing a continuous reference of the transmitter frequency and phase. Echoes are heterodyned with the reference oscillator. Stationary objects supply echoes having a carrier phase shift which is constant from pulse to pulse, because the time for the signal to travel to a stationary object and back is always the same. Therefore the heterodyner output for stationary echoes does not change in phase from pulse to pulse. However, phase shift in echoes from moving targets changes from pulse to pulse because of the change in signal propagation time. *See* DOPPLER RADAR.

To utilize the lack of pulse-to-pulse phase difference to discriminate against clutter echoes, the heterodyner output is fed to a delay line which stores the signal for a period of time exactly equal to the period between pulses. Then the output of the delay line is subtracted from the freshly produced output of the heterodyner. If the two outputs are identical, as will be the case for clutter, the difference is zero. But this difference is not zero for moving targets because of their pulse-to-pulse phase shift. The difference between consecutive echoes is presented as the MTI output. The delay line usually employed utilizes acoustic propagation within fused quartz.

A variety of internal arrangements are in use for assuring that the reference oscillator is synchronized with the transmitter output. The reference oscillator may be at the carrier frequency but usually operates at an intermediate frequency, because better stability can be obtained in a low-frequency oscillator. This arrangement requires that the local oscillator used for conversion from the carrier frequency to the intermediate frequency be extremely stable, and the term stalo (stable local oscillator) is used to denote circuits that satisfy this requirement. The reference oscillator is called the coho (coherent oscillator) to denote that it remains coherent (fixed phase relation) with the transmitter output.

Measures of merit. The figure of merit of an MTI system, clutter suppression factor, is defined as the ratio of root-mean-square (rms) clutter amplitude before MTI to rms clutter fluctuation after MTI. It is the amount by which the clutter is suppressed. It depends in turn upon several other factors, some of which are system parameters and thus controllable, and some of which are functions of terrain and weather.

Another measure of merit for MTI radar is the signal-to-clutter ratio (SCR). However, it has been found that increasing the SCR simultaneously reduces the signal-to-noise ratio (SNR), a measure of merit of the sensitivity of the radar. Targets in a clutter-free environment are less detectable with MTI than without it because of reduced SNR. A parallel bank of Doppler filters can provide a simultaneous improvement in both SCR and SNR. *See* SIGNAL-TO-NOISE RATIO.

Use of charge-transfer devices. Charge-transfer devices can be utilized to implement MTI with excellent capability for canceling clutter. A charge-transfer device is a capacitorlike semiconductor device of the bucket-brigade variety. As used in MTI, samples of the returned radar signal are stored and processed as discrete quantities of electrical charge. This implementation has become known as discrete-signal moving-target indication (DSMTI). Metal oxide semiconductor field-effect transistors are used as charge-transfer devices to provide the delay that permits subtracting a radar return delayed by one pulse-repetition interval (PRI) from the currently incoming radar return.

Adaptive MTI. Radar systems designers have long appreciated that the best target range and velocity resolution are obtained by using recurrent pulses whose duration is short compared with the PRI. However, this practice gives rise to ambiguities in determining the range and velocity of the target. In fact, it causes the radar to be blind to targets moving at certain velocities when MTI is in use. A way to avoid blind velocities is to implement a random staggering of the radar's pulse-repetition frequency (PRF). Fast changes in the PRF combined with filtering systems and integration methods that adapt automatically to these changes permit ambiguities in determining the range and velocity of the target to be resolved in a single look at the target. Clutter filters switched in synchronism with the PRF also reduce undesirable clutter sidebands even when the PRF is switched at a high rate, and the irregular pulse spacing

alleviates the problem of blind velocities. The clutter filter can be a notch-type filter that has its rejection band centered about the carrier frequency. *See* ELECTRIC FILTER.

This adaptive MTI (AMTI) proved to be useful in achieving acceptable detection of targets in the presence of natural precipitation static and chaff dispersed by the enemy as a defensive measure against radar. *See* ELECTRONIC WARFARE; RADAR.

John M. Carroll

Bibliography. J. L. Eaves and E. K. Reedy, *Principles of Modern Radar*, 1987; M. I. Skolnik, *Introduction to Radar Systems*, 2d ed., 1986; M. I. Skolnik, *Radar Handbook*, 2d ed., 1990.

Mucilage

A naturally occurring, high-molecular-weight (ranging 200,000 and up), organic plant product of unknown detailed structure. The term is loosely used, often interchangably with the term gum. Chemically, mucilage is closely allied to gums and pectins but differs in certain physical properties. Although gums swell in water to form sticky, colloidal dispersions and pectins gelatinize in water, mucilages form slippery, aqueous colloidal dispersions which are optically active and can be hydrolyzed and fermented. Mucilages are not pathological products but are formed in normal plant growth within the plant by mucilage-secreting hairs, sacs, and canals, but they are not found on the surface as exudates as a result of bacterial or fungal action after mechanical injury, as are gums. Mucilages occur in nearly all classes of plants in various parts of the plant, usually in relatively small percentages, and are not infrequently associated with other substances, such as tannins. The most common sources are the root, bark, and seed, but they are also found in the flower, leaf, and cell wall. Any biological functions within the plant are unknown, but they may be considered to aid in water storage, decrease diffusion in aquatic plants, aid in seed dispersal and germination, and act as a membrane thickener and food reserve. Mucilages are commonly identified by physical properties, most recently by infrared spectroscopy.

Most mucilages are considered to be cellulosic polysaccharides containing the same group of sugars as gums and pectins, and commonly occur as salts of acids (acid type). The cations are chiefly of calcium, magnesium, potassium, and sodium. The most common acids are uronic acids, the most important being galacturonic acid, although other acid residues including sulfonic acids are known. Some are known to exist with the acid function absent (neutral type) and consist only of simple sugars connected by a glycosidic linkage. Hydrolysis of mucilages yields pentoses and hexoses, the most common being arabinose, galactose, glucose, mannose, rhamnose, and xylose.

The chief industrial sources of mucilages are Icelandic and Irish moss, linseed, locust bean, slippery elm bark, and quince seed. They are obtained by milling if they occur in the endosperm, but more commonly by extraction with water or dilute sodium carbonate solution when they are found outside the seed coating. They are purified by precipitation with alcohol or salt solutions from the aqueous solution and marketed as powders. Synthetic mucilages can be formed by decomposition reactions, as in the hydrolysis of starch to form dextrins which are mixed with gums to form adhesives. Mucilages find applications in cosmetics (hand lotions and hair sets), medicinals (laxatives and diuretics), pharmaceuticals (emulsifying agents and materials for prevention of precipitation by colloidal suspension), and industry (thickening agents, emulsion stabilizers, binders, sizing of silks, textile printing, and paper manufacture). *See* ADHESIVE; GUM; PECTIN.

Elbert H. Hadley

Bibliography. A. J. Kinloch, *Adhesion and Adhesives: Science and Technology*, 1987; A. A. Lawrence, *Natural Gums for Edible Purposes*, 1977; W. Pigman (ed.), *The Carbohydrates*, 2d ed., vols. 1A–2B, 1970–1978.

Muffler

A device used to attenuate sound while also allowing fluid (usually gas) to flow through it; also known as silencer in British usage. Mufflers are extensively used to reduce the intake and exhaust noise from pumps, fans, compressors, and internal combustion engines. Although active noise control techniques are emerging, most mufflers continue to use passive silencing methods. Passive mufflers are categorized as reactive or dissipative based on their primary method of attenuation. Reactive mufflers reflect sound back toward the noise source, and dissipative mufflers use porous materials to absorb the sound (**Fig. 1**).

Muffler attenuation as a function of frequency is often characterized by the transmission loss (L_{TL}) and insertion loss (L_{IL}), in units of decibels (dB). Transmission loss is defined as $10 \log_{10} (W_{in}/W_{out})$, where W_{in} and W_{out} represent the sound power entering and exiting the muffler. Insertion loss is defined as the difference in sound pressure level at a fixed location caused by inserting the muffler in the system. Transmission loss is convenient for describing the behavior of the muffler itself, while insertion loss gives a better indication of muffler attenuation in a particular system. For some dissipative duct mufflers, acoustic performance is described by the drop in sound pressure level per unit length. *See* DECIBEL; SOUND; SOUND PRESSURE.

In addition to attenuation, several other criteria need to be considered in the design of these elements. The pressure drop due to flow restrictions usually needs to be minimized, since flow losses waste energy and will increase with flow rate. Other design considerations include size, durability, the magnitude of self-noise or flow-generated noise, and the initial, maintenance, and replacement costs.

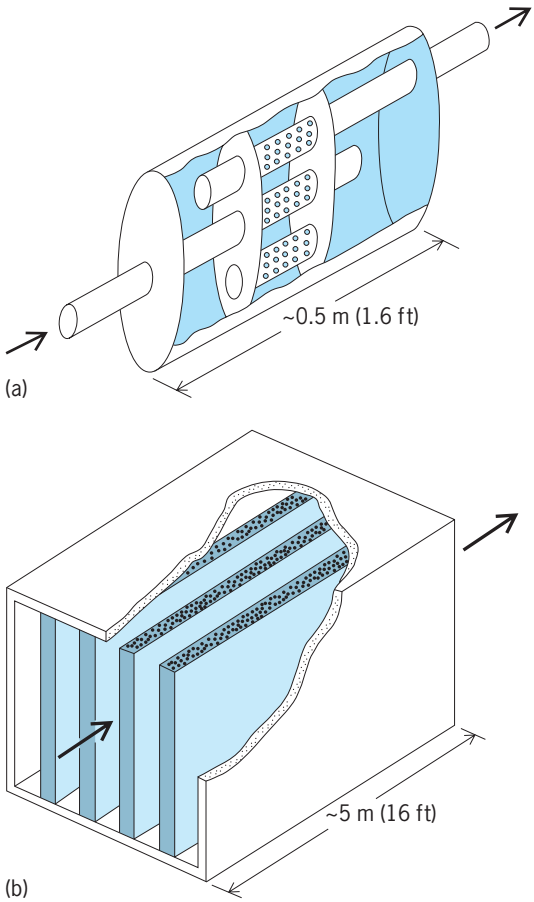

Fig. 1. Mufflers. (*a*) Reactive muffler for automobile exhaust system. (*b*) Dissipative muffler for industrial application.

Often the design criteria conflict: The attenuation generally increases with size and flow restriction, and this increase typically competes with the need to minimize these latter parameters.

Reactive mufflers. Reactive mufflers reflect acoustic waves at locations where a duct expands, contracts, or branches. Often a combination of reactive elements such as expansion chambers, resonators, and flow reversals is used (Fig. 1*a*). Reactive mufflers can be designed to provide better low-frequency attenuation than a dissipative muffler of similar size. Also, reactive mufflers can be used in harsh environments that dissipative or active mufflers might not withstand.

In most cases, reactive mufflers are best suited for low-to-moderate frequencies, where acoustic wavelengths are larger than any cross dimension of the muffler. At these frequencies, mufflers can exhibit resonance and/or broadband attenuation behavior. A resonance suggests a small region of high attenuation centered around a frequency. Resonance behavior is exhibited by the Helmholtz (**Fig. 2***a*) and side-branch (Fig. 2*b*) resonators. Reactive mufflers with broadband behavior, such as a simple expansion chamber, may exhibit lower maximum attenuation than a resonator, but attenuate over a wider frequency range (Fig. 2*c*). The side-branch resonator and the expansion chamber demonstrate another characteristic of many reactive muffler elements where the attenuation is cyclic with respect to frequency. Locations for attenuation maxima and minima are related to ratios between muffler lengths and the acoustic wavelength. *See* RESONANCE (ACOUSTICS AND MECHANICS).

Dissipative mufflers. Dissipative mufflers use absorptive materials that dissipate the acoustic energy into heat. A variety of porous media can be used for absorption, with fibrous materials such as fiberglass being common. The dissipative material can be incorporated in a number of ways. For example, it

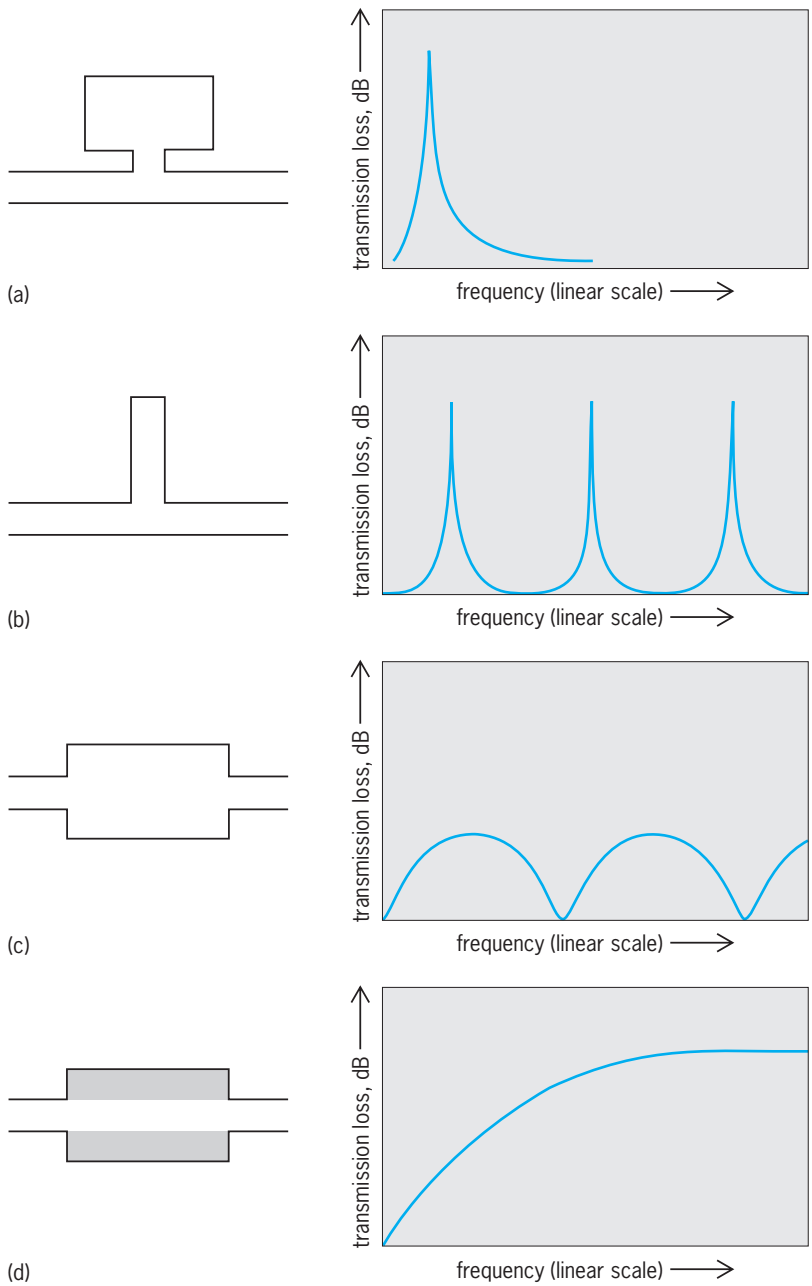

Fig. 2. Simple muffler elements, and their representative transmission losses. (*a*) Helmholtz resonator. (*b*) Side-branch resonator. (*c*) Expansion chamber. (*d*) Expansion chamber with dissipative material.

can fill the outer cavity of an expansion chamber (Fig. 2*d*), or it can be used as linings for the duct walls and baffles that separate a duct into smaller channels (Fig. 1*b*). The linings and baffles can be flat, contoured, constructed from layers of different materials, or mixed and matched for a particular application. Absorptive materials may face challenges due to harsh conditions such as high temperatures and potential clogging from particulate-laden flows. *See* SOUND ABSORPTION.

Dissipative mufflers are best suited for moderate-to-high frequencies, since absorption is less effective at low frequencies. At frequencies where the absorptive materials are effective, the attenuation is broadband, and the passbands exhibited by reactive mufflers are reduced or eliminated (Fig. 2*d*). Compared to reactive mufflers of similar size, dissipative mufflers can have higher attenuation (except at resonances for the reactive muffler) and lower pressure drop. At higher frequencies, where the acoustic wavelength is smaller than the duct width, the attenuation of a dissipative muffler may decrease considerably.

In practice, no muffler is entirely dissipative or reactive, as some amount of both types of behavior is inevitable. Many mufflers combine reactive and dissipative components to capitalize on their respective strengths. Reactive elements are used for low-frequency attenuation, and dissipative elements are used to absorb moderate- and high-frequency noise.

Active mufflers. Active mufflers attenuate unwanted noise by adding sound to counteract it. In a simplified example, a loudspeaker is used to generate sound waves that are an exact mirror image of the incoming noise. The disturbances add algebraically, resulting in a cancellation of the unwanted sound. An active muffler consists of sensors (such as microphones), a controller, and actuators (such as loudspeakers). The controller unit processes the signals from the sensor, and computes an appropriate signal for the actuator. Numerous control systems and strategies exist, and are under continuous development.

Active mufflers are best suited for low frequencies where the sound field is relatively simple. The effectiveness of active mufflers has been demonstrated for a number of situations, but several challenges remain, which are the topic of ongoing research. There is a need for rugged sensors and actuators that can withstand high temperatures and harsh environments. Also, high-intensity disturbances at low frequencies require large-displacement, high-power actuators. *See* ACTIVE SOUND CONTROL.

Design tools. Many predictive tools are available to assist in the design of mufflers. For simple geometries and low frequencies, analytical wave methods are often adequate. For higher frequencies and more complex mufflers, computer models can be used effectively. Currently, however, experimental testing of numerous prototypes is often required to finalize a muffler design. Several factors that can affect muffler behavior are difficult to fully describe and predict. These include mean flow effects, flow-generated noise, intense sound levels, and acoustic propagation through porous interfaces and media. Research is continuing in these and other areas to more fully understand and more accurately predict the muffler behavior. *See* ACOUSTIC NOISE.

Ahmet Selamet

Bibliography. L. L. Beranek and I. L. Vér, *Noise and Vibration Control Engineering*, Wiley, 1992; D. D. Davis, Jr., Acoustical filters and mufflers, in C. M. Harris (ed.), *Handbook of Noise Control*, McGraw-Hill, 1957; M. L. Munjal, *Acoustics of Ducts and Mufflers*, Wiley, 1987.

Mugiliformes

An order in the class Actinopterygii, subdivision Euteleostei, superorder Acanthopterygii. The Mugiliformes comprise 17 genera and about 72 species, all in the single family Mugilidae, the mullets (also known as grey mullets). Past classifications considered mullets as primitive perciforms; however, they are presently deemed subperciforms and are the only order in Mugilomorpha, one of three series making up the Acanthopterygii. *See* ACTINOPTERYGII; PERCIFORMES.

Morphology. The following characters distinguish the family: The head and body sectors are more or less round in cross section and the tail sector is moderately compressed; the dorsal fins are well separated, the first has four weak spines, and the second 8 to 10 soft rays; the pectoral fins are high on the body; the pelvic fins are subabdominal, each with one spine and five rays; the anal fin has two or three spines and 7 to 11 soft rays; an adipose eyelid is usually present; the mouth is small and terminal; teeth are small or absent; gill rakers are long; the stomach is usually muscular (gizzardlike), and the intestine is long and coiled; scales of adults are usually ctenoid (or cycloid in juveniles); and the lateral line is absent or very faint. Adult mullet species range from 15 cm (6 in.) to 1.2 m (4 ft) in maximum length, but most species are about 30 to 70 cm (11.8 to 27.5 in.) long (see **illustration**).

Habitat. Mullets inhabit coastal marine and brackish-water environments in all tropical and temperate seas, as well as freshwater. Some marine species enter freshwater only temporarily and may ascend great distances up rivers, such as the striped mullet (*Mugil cephalus*), a common and economically important species that occurs practically worldwide in warm waters. Adult mountain mullet (*Agonostomus monticola*), a western Atlantic species ranging from North Carolina to Venezuela, including the West Indies, is confined to freshwater streams, but its juveniles occur in offshore coastal waters. Probably the only species that does not spend part of its life in a marine environment is the Abu mullet (*Lisa abu*) of Turkey, Iraq, Pakistan, and southeastern Asia. Most species form schools, some quite large. Their primary food is detritus, plant material, and microorganisms that are filtered from muddy substrates by a pharyngobranchial organ,

Largescale mullet. (*Photo courtesy of John E. Randall*)

macerated by the muscular stomach, and absorbed by the long coiled intestine. Herbert Boschung

Bibliography. I. J. Harrison, Mugilidae, pp. 1071–1085 in K. E. Carpenter (ed.), *The Living Marine Resources of the Western Central Atlantic: FAO Species Identification Guide for Fishery Purposes*, vol. 2, FAO, Rome, 2003 (dated 2002); I. J. Harrison and G. J. Howes, The pharyngobranchial organ of mugilid fishes: Its structure, variability, ontogeny, possible function and taxonomic utility, *Bull. Brit. Mus. Nat. Hist.* (*Zool.*), 57(2):111–132, 1991; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; M. L. J. Stiassny, What are grey mullets?, *Bull. Mar. Sci*, 52(1):197–219, 1993.

Mulberry

A genus, *Morus*, of trees characterized by milky sap and simple, often lobed, alternate leaves. White mulberry (*M. alba*) was introduced into the United States from China during the nineteenth century as a source of food for silkworms. The silkworm project was unsuccessful, but the trees remained. This species has smooth shiny leaves that are usually lobed (**Fig. 1**), and a very irregular habit of branching. The fruit is white and insipid. Red mulberry (*M. rubra*) attains a height of 60 ft (18 m) and grows in the eastern half of the United States and in southern Ontario.

Fig. 1. White mulberry (*Morus alba*).

Its leaves are rough above and soft and pubescent (hairy) beneath, and sometimes have two or three lobes (**Fig. 2**). The fruit is about 1 in. (2.5 cm) long,

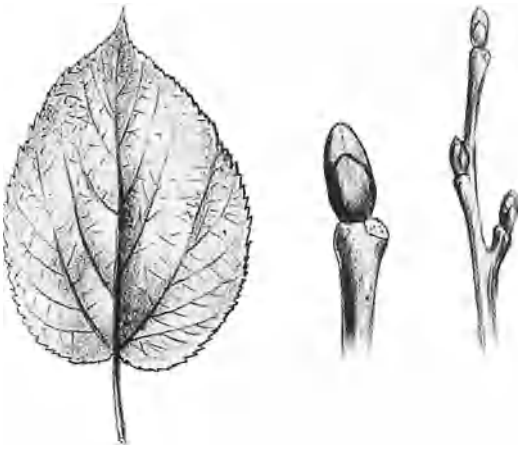

Fig. 2. Red mulberry (*Morus rubra*).

dark purple or black, and sweet, and is used as food by both domestic and wild animals. The wood is used for fence posts, furniture, interior finish, agricultural implements, and barrels. *See* FOREST AND FORESTRY; TREE. Arthur H. Graves; Kenneth P. Davis

Mule

The mule is a hybrid sired by a male ass (*Equus asinus*) out of a female horse (*E. caballus*). The opposite cross, very seldom made, produces the hinny (a hybrid between a stallion and a female ass). The mule and the hinny are usually sterile, but there have been two authenticated cases of mules producing living progeny. Male mules, often called horse mules, are almost always castrated to make them more tractable as work animals.

Jacks and jennets are the males and females, respectively, of the domestic ass family. In the United States, most jacks have been used to sire mules, whereas almost all jennets have been kept solely to produce more jackstock. Small asses, about the size of Shetland ponies, are commonly called donkeys or burros. They are used as pack animals in a few

mountainous areas of the Southwest, and occasionally as children's pets in other regions. *See* HORSE PRODUCTION.

American jacks, sometimes called Mammoth jacks, were developed from a blending of several European strains of jackstock, principally the Catalonian, Maltese, Majorcan, Andalusian, Italian, and Poitou. Large jacks may measure over 15 hands (60 in. or 1.5 m) and weigh 1100 lb (500 kg). Compared to the horse, jacks have shorter hair on the tail and mane, much longer ears, smaller and deeper hoofs, and no callosities (hard or thickened areas) on the inside of the hind legs. The loud, harsh voice of jacks is called a bray. The mule also shows these characteristics, but usually is more full bodied and more smoothly constructed than the jack.

A jack to be used for mule production should be raised with horses after weaning and not permitted to see jennets and mules until he has been broken to breed mares. Otherwise, the jack may be very slow or even unwilling to breed a mare. Heavy draft mares bred to jacks produce big draft mules. Mares of finer quality produce lighter, finer-boned mules of the kind that were formerly used by many tobacco growers and cotton farmers in the South.

Mules are noted for their endurance, sure footedness, and ability to stand hard work in hot weather. They can safely be self-fed in lots or corrals, whereas horses cannot. Usually steady and free from nervous excitability, mules can be handled by inexperienced or careless farm labor. *See* AGRICULTURAL SCIENCE (ANIMAL).

John M. Kays

Multiaccess computer

A computer system in which computational and data resources are made available simultaneously to a number of users. Users access the system through terminal devices, normally on an interactive or conversational basis. A multiaccess computer system may consist of only a single central processor connected directly to a number of terminals (that is, a star configuration), or it may consist of a number of processing systems which are distributed and interconnected with each other as well as with the user terminals.

The primary purpose of multiaccess computer systems is to share resources. The resources being shared may be simply the data-processing capabilities of the central processor, or they may be the programs and the databases they utilize. The earliest examples of the first mode of sharing are the general-purpose, time-sharing, computational services. Examples of the latter mode are airlines reservation systems in which it is essential that all ticket agents have immediate access to current information. Both classes of systems are still popular, with the proportion and importance of distributed application systems continuing to increase in areas such as corporate management and operations. *See* DATABASE MANAGEMENT SYSTEM.

There are a number of economic factors supporting the growth of multiaccess systems. Primary among these are improvements in hardware performance and economics, with the latter being the principal factor. Digital-computer hardware continues to exhibit rapidly decreasing costs while its operational capabilities increase. At the same time most of the other cost factors involved with corporate operations such as labor, travel, and communications are increasing, and the cost of time delays in operations such as inventory management and confirming orders, as well as the cost of investment funding, is becoming significant.

System components. The major hardware components of a multiaccess computer system are terminals or data entry/display devices, communication lines to interconnect the terminals to the central processors, a central processor, and on-line mass storage.

Terminals may be quite simple, providing only the capabilities for entering or displaying data. Also utilized in multiaccess systems are terminals having an appreciable amount of "local intelligence" to support simple operations like editing of the displayed text without requiring the involvement of the central processor. Some terminals provide even more extensive support, such as the local storage of small amounts of text. It is also possible to assemble clusters of terminals to share local logic and storage.

The interconnecting communication lines can be provided by utilizing the common-user telephone system or by obtaining leased, private lines from the telephone company or a specialized carrier. Another important source of data circuits is a value-added network specializing in providing data transmission services. *See* DATA COMMUNICATIONS.

The communications interface for the processor may be provided by an integral hardware component or by a separate device known as a communications controller or front-end (communications) processor. The detailed operations involved in the control of a single communications line are not very great, but servicing a large number of such lines may present an appreciable load to the central processor; thus, separate hardware for communications interfacing is almost essential in a multiaccess computer system.

A central processor suitable for use in a multiaccess system must include the capability to support a large central memory as well as the communications interface mentioned above and the on-line mass storage discussed below.

The desirable characteristics of a mass storage device are: the ability to quickly locate any desired data; the ability to transfer data at a very high speed; the provision of economical storage for a large quantity of data; and a high degree of reliability.

The most common form of mass storage is rotating magnetic disks. Originally, magnetic drums were utilized for this purpose; however, the price, performance, and capacity of disks have resulted in their replacing drums. *See* COMPUTER STORAGE TECHNOLOGY.

System operating requirements. A multiaccess system must include the following functional capabilities: (1) multiline communications capabilities that will support simultaneous conversations with a reasonably large number of remote terminals; (2) concurrent execution of a number of programs

with the ability to quickly switch from executing the program of one user to executing that of another; (3) ability to quickly locate and make available data stored on the mass storage devices while at the same time protecting such data from unauthorized access.

The ability of a system to support a number of simultaneous sessions with remote users is an extension of the capability commonly known as multiprogramming. In order to provide such service, certain hardware and software features should be available in the central processor. Primary among these is the ability to quickly switch from executing one program to another while protecting all programs from interference with one another. These capabilities are normally provided by including a large central memory in the processor, by providing hardware features that support rapid program switching, and by providing high-speed transfers between the central memory and the mass storage unit on which the programs and the data they utilize are stored.

Memory sharing is essential to the efficient operation of a multiaccess system. It permits a number of programs to be simultaneously resident in the central memory so that switching execution between programs involves only changing the contents of the control registers. The programs that are resident in central memory are protected from interfering with one another by a number of techniques. In earlier multiprogramming systems, this was accomplished by assigning contiguous memory space to each program and then checking every access to memory to ensure that a program was accessing only locations in its assigned space. The drawback to this type of memory allocation is that the entire program had to be loaded into memory whenever any portion of it was to be executed. A popular memory management technique is the utilization of paging. The program is broken into a number of fixed-size increments called pages. Similarly, central memory is divided into segments of the same size called page frames. (Typical sizes for pages and page frames are 512 to 4096 bytes.) Under the concept known as demand paging, only those pages that are currently required by the program are loaded into central memory. Page frames may be assigned to a given program in a random checkerboard fashion. Hardware capabilities are provided to automatically manage the assignment of page frames as well as to make them appear to the program as one continuous address space. At the same time, the hardware provides memory access protection.

Software capabilities. The control software component of most interest to an interactive user is the command interpreter. It is only one portion of a larger control program known as the operating system; however, in multiaccess systems it differs greatly from the command interpreter in a system providing batch multiprogramming service on a noninteractive, non-terminal-oriented bases. This routine interacts directly with users, accepting requests for service and translating them into the internal form required by the remainder of the operating system, as well as controlling all interaction with the system.

The operating system must also have the capability for controlling multiprogramming; that is, the concurrent execution of a number of user programs quickly switching from one to another during their execution as well as controlling memory sharing. The capability to page the memory as outlined above can be utilized to provide users with the impression that each has available a memory space much larger than is actually assigned. Such a system is said to provide a virtual memory environment. Similarly, the ability of the operating system to quickly change context from one executing program to another will result in users' receiving the impression that each has an individual processor. This is especially true when considering the large difference between the response and thinking time of a human compared to the computer's processing time. For the interactive users, such a capability results in the impression of having a private virtual machine.

The mass storage device mentioned above as hardware required for a multiaccess system must be supported by an efficient file management system. The latter is responsible for maintaining current information as to the physical location of the data stored on the mass storage device as well as providing a capability for quickly locating those data and controlling their transfer to the central memory for utilization. In addition, the file management system must provide protection of data from unauthorized access. *See* DIGITAL COMPUTER. Philip H. Enslow, Jr.

Bibliography. V. D. Black, *Computer Networks*, 2d ed., 1993; F. Halsall, *Data Communications, Computer Networks, and Open Systems*, 4th ed., 1996; G. Held and R. Sarch, *Data Communications: A Comprehensive Approach*, 6th ed., 1999; W. Stallings, *Data and Computer Communications*, 6th ed., 1999.

Multidimensional scaling

Any of a number of procedures for the fitting of a selected geometric model to a body of multivariate data. In the original form advanced by Warren S. Torgerson, multidimensional scaling dealt with judgments about stimuli, with its major aim being the derivation of psychological dimensions presumed to underlie the judgments, and the determination of scale values for stimuli along the derived dimensions. Classical multidimensional scaling was a significant methodological advance in the characterization of complex stimulus domains. If forced to rely exclusively on unidimensional scaling methods, a researcher could characterize a multidimensional stimulus domain only by first scaling the stimuli on individual dimensions and later assembling the scaled dimensions in a postulated model. Even then, the model would often be incomplete because of insufficient foresight in choice of dimensions, or inability to determine the relations between, and relative importance of, the chosen constituent dimensions.

Advantages. With the advent of multidimensional scaling and related techniques, psychological inquiry benefited in two important ways. First, the data

collection procedures associated with the new techniques did not require prior specification of constituent dimensions before scaling could take place. As a consequence, it was necessary to instruct the subjects in the specific dimensions to be considered in making judgments. Since most multidimensional scaling procedures made use of judgments on the metadimensions of similarity of preference rather than judgments on specific psychological dimensions (such as asking whether something is good or bad), the elicited responses were considered to be less artificially constrained by the investigator, and hence more likely to reflect the true nature of the subject's psychological experience. Second, the multidimensional scaling procedures allowed all relevant stimulus dimensions to be extracted in the same operation and to exhibit their relative strengths and interrelationships within the confines of a single model. This feature thus freed researchers of the requirement of specifying dimensions in advance. These two features, rightly or wrongly, served to promote the idea that multidimensional scaling would allow "recovery" of the cognitive structure underlying the judgments.

Fields other than psychology also found applications for multidimensional scaling. Through the use of metadimensions such as proximity or dominance, analysis could be extended to well-defined surrogates, such as indices of overlap or confusion, measures of net migration, rankings of dominance, or coefficients of interaction. The procedure then allowed psychological scaling methodology to be applied in nonpsychological sciences and permitted results to be interpreted, not in terms of psychological processes, but in terms of processes appropriate for the particular field of application.

Evolution. The evolution of multidimensional scaling from its origins in the classical euclidean model to its present status as a more general model can be chronicled in terms of four somewhat overlapping but distinct generations of methodology.

First generation. In the first generation, that of classical multidimensional scaling, the main models were euclidean in nature, with data having the property of ratio scale required. With scaled dissimilarities considered as distances in euclidean space, or with normed similarities considered as cosines of angles in euclidean space, the classical model used linear algebra to construct basis dimensions through a standard eigenvector decomposition of a scalar product matrix. The resulting axes were orthogonal and were ordered in magnitude according to the square roots of the extracted eigenvalues. Solutions were analytic, as opposed to iterative, and produced the full complement of dimensions simultaneously. Violations of metric axioms in the judged data would sometimes give imaginary dimensions, and nonlinearity of data with respect to the underlying structure often generated extra, spurious dimensions. Data for the classical model were usually aggregated over subjects prior to analysis, since single data sets were not considered adequate for the stringent input data requirements. Little was done to scale subjects individually or to relate one subject's pattern of judgments to that of another. In addition, coordinate dimensions were initially defined solely in terms of the axes of eigenvectors without any consideration for the psychological plausibility of such artificial dimensions.

Second generation. The second distinguishable phase in multidimensional scaling methodology was initiated by the contributions of Roger N. Shepard and has become known as nonmetric multidimensional scaling. Nonmetric multidimensional scaling departed from the classical approach in several respects. First, the procedures incorporated a nonlinear psychological response function directly into the model rather than treating such distortion as simple error or correcting for it prior to the analysis. Although many plausible families of functions, such as exponential or power curves, could have been assumed, the multidimensional scaling packages emerging during this phase made only the very minimal assumption that the psychological function was simple monotonic with respect to the underlying structure. By imposing only ordinal constraints and by allowing input to be transformed as necessary as long as ordinality was preserved, it was found that the fitting of such ordinally transformed data to a space of low dimensionality was accompanied by such a degree of overdetermination that a spatial configuration could be estimated with essentially ratio-level precision.

A further distinctive attribute introduced with nonmetric multidimensional scaling was the algorithmic basis of the solution procedure. Most such procedures were iterative in nature, since the determination of the configuration and estimation of the function by monotone least-squares regression required a succession of approximations programmed on the computer. The main computational problem was the determination of an appropriate loss function for comparing the original to the transformed data, and the determination of an efficient procedure for minimizing the function. Since there were several possibilities for defining the stress, or loss function, as well as a number of heuristics for producing the minimization algorithm, a variety of computer programs emerged. The most widely circulated versions originated from the University of Michigan, the University of North Carolina, and Bell Laboratories. The aims of the various approaches were alike in the sense of attempting to fit a ratio-level geometric model to ordinal-level data. The use of ordinal data relieved the researcher of the requirement to convert original data into ratio-level distance estimates prior to scaling, and was found to be very attractive in psychological research contexts in which data collection procedures could not hope to meet the stringent ratio-level requirements of classical scaling.

A related feature of the nonmetric multidimensional scaling phase was a pluralism resulting from the proliferation of options. Input observations could be judgments of similarities or dissimilarities, conditional or unconditional, complete or partially missing, and so on. The fitted model could be a distance model or a scalar product model; it could be euclidean or noneuclidean, or it could be symmetric or nonsymmetric. And for the fitting procedure,

decision criteria could differ in the initialization, convergence, or termination phases, in the choice of a loss function, or in the specification of dimensionality, to mention only the most relevant. As a result, the hope of agreeing on standard criteria for a unique solution was nearly abandoned. The abundance of options gave rise to a curiously paradoxical state of affairs wherein nonmetric multidimensional scaling seemed at one and the same time to be very precise by producing ratio-level solutions from ordinal data, yet rather imprecise by allowing a whole range of such solutions to be considered about equally acceptable for a given set of data. While the variants of a solution for a given set of data were often highly similar to one another, there was still no easy way to assess the superiority of any one solution over another. Thus, the proliferation of possibilities encouraged a certain permissiveness or tolerance with respect to solutions that seemed to be at variance with one another.

Third generation. In the third generation, that of individual differences models, scaling capability expanded to three-way analyses with the estimation of weights for each subject on individual dimensions. By introducing an additional set of parameters to the estimation process, the analysis was forced to abandon the gradient method of function minimization of the previous stage in favor of using an alternating least-square procedure. This breakthrough, typified by the work of J. Douglas Carroll, offered answers to two long-standing problems in multidimensional scaling. First, inaccuracies and inconsistencies caused by aggregating over subjects were reduced by giving each subject an individual place in the model. Not only that, but by including all subject parameters as an integral part of a single model, a more useful solution could be obtained to give a group space representing a general trend, plus individual spaces with weighted dimensions. A second, and possibly more remarkable, advance was the "dimensional invariance" that offered an answer to the rotation problem. Thus, in three-way weighted euclidean models, the extracted group dimensions were found not to be arbitrary, as they had been in classical and nonmetric multidimensional scaling, but rather, they were fixed by consistent variation underlying the individual subjects' judgments. A further benefit arising from the introduction of three-way analyses was that the technical capability of multidimensional scaling was brought closer in line with current developments in the field of numerical analysis, thereby demonstrating that large parameter sets could be estimated without problems of local minima, and suggesting that other multiple parameter sets could be proposed and successfully programmed by using alternating least-squares procedures. Three-way scaling thus opened the way for a development toward greater computational complexity.

Fourth generation. The fourth and current generation of multidimensional scaling is that of complex model development. Work of this nature actually goes back to the early studies proposing the city-block metric as an alternative to the euclidean metric. As the current phase is gaining momentum, its emphasis is not so much on properties of the input data or on the efficiency of the fitting algorithm, but on the adequacy of the structural model itself for the representation of the kinds of psychological data found in complex stimulus domains. Throughout the development of multidimensional scaling, there has been an acknowledgment that cognitive structures were more than euclidean in nature, and that in certain cases the euclidean spatial model might be totally inappropriate. Early efforts to capture the complexity through noneuclidean spatial modeling were thwarted, however, by lack of computational procedures powerful enough to fit hypothesized structures without being susceptible to degenerate solutions, premature termination of the algorithm at local minima, and other technical problems. In testing two alternative models on the same data, it was often difficult to determine whether one model or the other provided the better fit. But with computation problems and numerical analysis now more under control, the emphasis has turned to development of noneuclidean spatial models and nonspatial geometric models, as well as the exploration of hybrid models that incorporate elements of taxonomic and continuous scaling. To complement previous work on scaling of Minkowski p metrics, research has focused on models for riemannian space, on discrete algebraic models, and on models that are especially appropriate for distinctive feature domains. With increased interest in detailed exploration of cognitive processes, more emphasis has been placed on cluster analysis procedures and on possible means for integrating hierarchical and taxonomic structures within traditional scaling models. Richard Degerman

Bibliography. J. D. Carroll and P. Arabie, Multidimensional scaling. *Annu. Rev. Psychol.*, 31:607–649, 1980; P. E. Green et al., *Multidimensional Scaling: Concepts and Applications*, 1989; W. S. Torgerson, *Theory and Methods of Scaling*, 1958; F. W. Young and R. M. Hamer (eds.), *Multidimensional Scaling: History, Theory and Applications*, 1987.

Multilevel control theory

An approach to the control of large-scale systems based on (1) decomposition of the complex overall control problem into simpler and more easily managed subproblems and (2) coordination of the subproblems so that overall system objectives and constraints are satisfied.

The controllers are organized in a multilevel hierarchical structure according to three basic criteria: functional, plant, and temporal decomposition. Functional and temporal decompositions are often classified as multilayer or vertical structures; plant decomposition is often classified as a multilevel or horizontal structure.

Functional decomposition. In this approach the overall control problem is partitioned into a nested set of generic control functions (**Fig. 1**). The regulatory or direct control function interfaces with the plant to implement the decisions of the optimizing controller inputted in the form of setpoints,

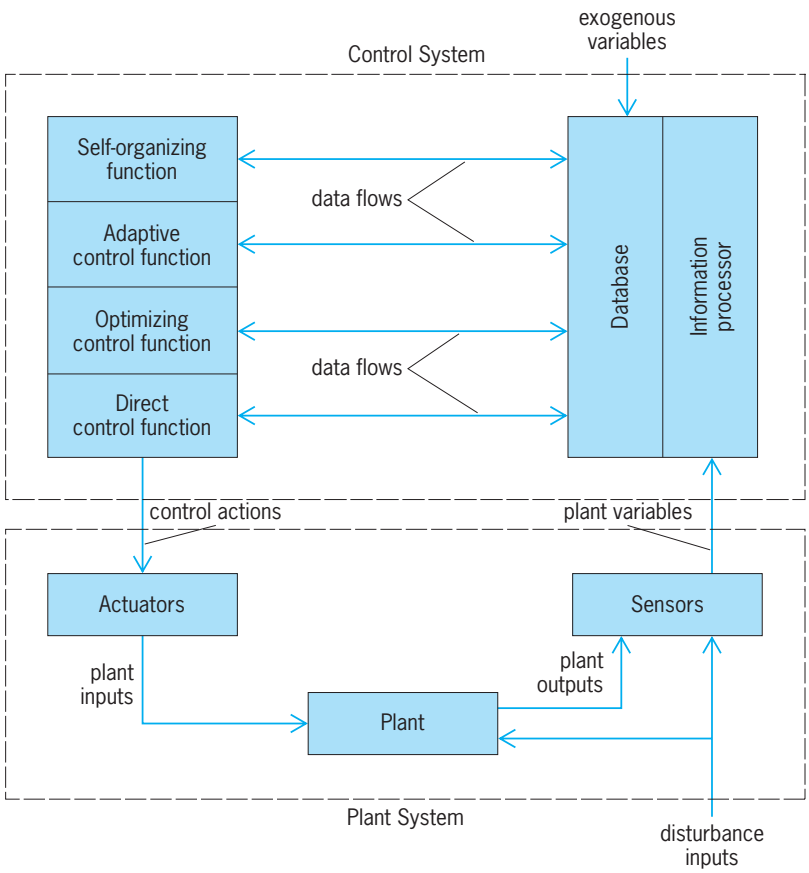

Fig. 1. Functional control hierarchy.

desired trajectories, or targets. The optimizing control function determines the necessary relationships among the variables of the system to achieve an optimal (or perhaps suboptimal but satisfactory) performance based on a given approximate (simplified and aggregated) model of the plant and its environment. The adaptive control function updates parameters of the optimizing control model to achieve a best fit to current plant behavior, and updates parameters of the direct controller to achieve good dynamic response of the closed-loop system. Finally, the self-organizing function is concerned with changing the structure of the control system (for example, nonparametric changes in control, decision-making and information-processing algorithms) in response to changes in system objectives, market conditions, plant characteristics (induced, say, by replacement of obsolete components), major contingency events, and so forth. *See* ADAPTIVE CONTROL; OPTIMAL CONTROL THEORY.

An alternate formulation of the functional hierarchy is characterized by the following distinct layers of control and information processing: direct control, supervisory control, intraarea coordination, production scheduling and operational management, and management information. This formulation differs from the previous one (Fig. 1) in the explicit identification of management information and decision-making functions (level 4), human–machine interfaces (operator, supervisor, and manager consoles), and couplings with other subsystems (resulting from, say, a horizontal decomposition of the overall system).

Plant decomposition. In this multilevel approach the controlled system (plant) is partitioned into subsystems along lines of weak interaction. In a two-level control hierarchy (**Fig. 2**), each subsystem has its own (first-level) controller which acts to satisfy local objectives and constraints. A second-level controller (coordinator) influences the actions of the local controllers to compensate for subsystem interactions so that overall objectives and constraints are satisfied. The approach extends readily to a hierarchy with three or more levels.

The multilevel control problem is typically formulated as a constrained optimization problem where the original complex problem is replaced by a number of optimization subproblems of reduced dimensionality. This simplifies the analysis, and it may also result in reduced computational effort.

There are several methods advanced for coordinating the subproblem solutions. In the goal coordination method (also known as the interaction balance or nonfeasible method), Lagrange multipliers enter into the subsystem cost functions as "shadow" prices. These are adjusted by the second-level controller in an iterative procedure which culminates (if the method is applicable) in the satisfaction of the subsystem coupling relationships. In the interaction prediction method (the feasible method), the interaction variables are specified by the second-level controller according to overall optimality conditions. The subproblems are solved to satisfy local optimality conditions constrained by the specified values of the interaction variables. Again, an iterative procedure is implied.

As far as the plant is concerned, only the final result of the iterative procedure is important. However, since the optimum depends on the exogenous inputs to the system (for example, disturbances), and these are assumed to vary with time, the solution process must be repetitive. Thus, in on-line control applications of the multilevel structure:

1. The first-level controllers compensate for local effects of variations of the exogenous inputs; that is, they maintain local performance close to the optimum while ensuring that local constraints are not violated.

2. The second-level controller modifies the criteria or the constraints, or both, for the first-level controllers in response to changing requirements on the system, so that actions of the local controllers are consistent with the overall objectives of the system. In effect, the second-level controller compensates for the mean effects of variations in the interaction variables.

3. The control signals generated by the first-level controllers represent computed optimizing control conditions to be implemented on the plant. In general, these are implemented via direct controllers with the control signals as setpoints.

Temporal decomposition. In this approach, the control or decision-making problem is partitioned into

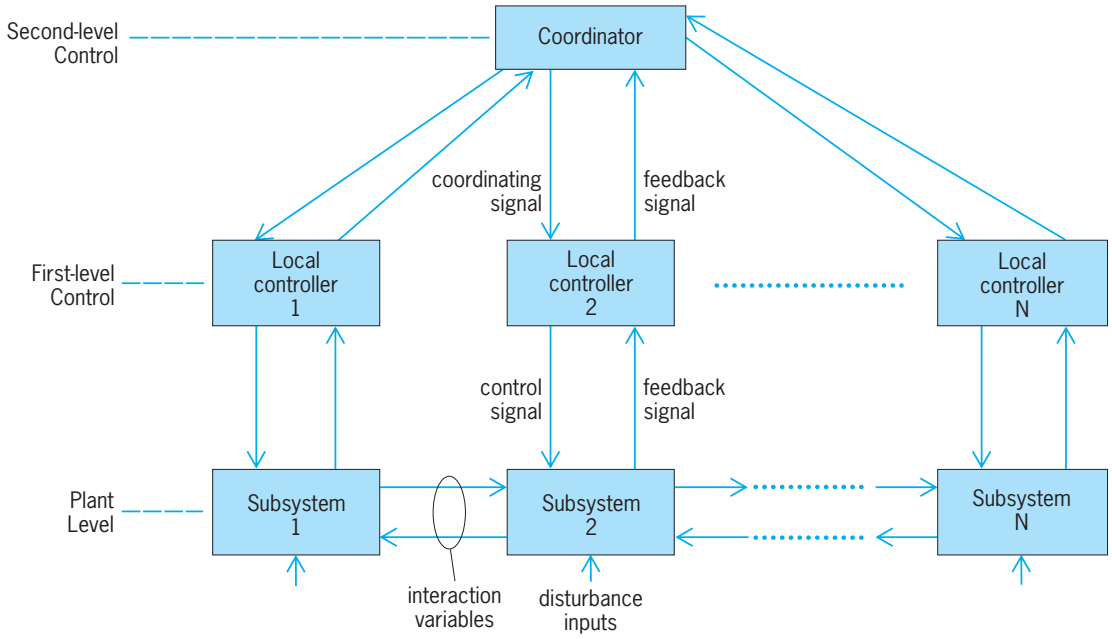

Fig. 2. Two-level control hierarchy.

subproblems based on the different time scales relevant to the associated action functions. These time scales reflect such factors as the response time of the plant, the bandwidth characteristics of the disturbance inputs, and trade-off considerations relating the benefit of control action to its cost. At each layer of the hierarchy, decisions and control actions determined at higher layers (corresponding to lower-frequency events) are treated as constants; disturbances and actions associated with lower layers are treated as noise to be represented by their respective mean values. The temporal hierarchy may embrace a broad spectrum of control and decision-making activities, including process control, production control, scheduling, and planning functions. Time scales may range from seconds to years. Other characteristics of the temporal hierarchy relate to the form of the model, the degree of uncertainty involved in the decision making, the level of aggregation, the information flow requirements, and so forth. In a typical production system hierarchy, three levels of decision making are defined based on time scale: medium-term production planning at the upper level; short-term production planning at the middle level; and scheduling of machines and processes at the lower level. *See* PROCESS CONTROL; PRODUCTION ENGINEERING.

Nontraditional control technologies. Multilevel structures may be used to integrate different knowledge-based and nontraditional control technologies (such as fuzzy control, neural-net control, and expert systems) within a given systems control application. The choice of technology may be based on time-scale requirements, nature of the knowledge base, or characteristics of the control problem. Thus, in one application, fuzzy control is applied at the first level, rule-based control at the second level, and algorithmic-based decision-making at the third. In another application, direct control is carried out by conventional linear controllers under the supervision of a heuristic rule-based controller at the second level. In another structure, a heuristic approach based on fuzzy set theory is applied to the lower three layers of a production planning and scheduling hierarchy organized according to time scale. The fuzzy decision system helps to reduce the effort at the knowledge engineering and design stage, simplify the computer implementation, provide an effective means of handling uncertainties inherent in the problem formulation, and improve coordination between the human operator and the computer control system. *See* EXPERT SYSTEMS; FUZZY SETS AND SYSTEMS; NEURAL NETWORK.

One feature of expert-systems methodology is that it permits the human-machine interface to be shifted to higher layers or levels of the control hierarchy. Therefore, the ability to incorporate heuristics, judgmental, and generally nonquantifiable factors into the decision-making problem means that more of the tasks requiring human intervention may be embedded into the computer control structure.

In the traditional formulation of the hierarchical control/decision-making problem, each subsystem is characterized by a single objective function. However, this simplifying assumption often leads to unsatisfactory or unrealistic results because of multiple, conflicting, or nonquantifiable objectives. In computer-integrated manufacturing systems, for example, the objectives of maximizing product quality, productivity, profit, and production flexibility are often in conflict. The problem is handled by decomposing it into a series of independent multiobjective subproblems arranged in a hierarchy with vertical and horizontal interdependencies. The individual subproblems are solved by standard multiobjective methodologies, in terms of constraints imposed by

the supremal unit (local manager) and feedbacks provided by lower-level units. An interactive approach is applied to handle the vertical interactions, for example, adjusting upper and lower bounds on available resources to achieve local feasible and efficient solutions.

Applications. Hierarchical control has become accepted technology in many industries. Even when the hierarchical structure is not explicitly identified, it is often implicitly embedded in the design of the system, providing the conceptual framework for integrated (or plant-wide) systems control. Motivating considerations are improved productivity, operating efficiency, product quality, or other economic-based objectives. Following is a sampling of industrial applications. The major factor contributing to the economic feasibility of these applications has been the availability of modern, highly sophisticated, distributed microprocessor-based control systems. Hierarchical control has been applied to various nonindustrial systems, such as robotics and freeway systems.

Steel and process industries. Some of the most advanced applications of multilayer and mutlilevel control concepts may be found in the steel industry. Tasks are assigned to a hierarchy of computers according to production unit, time scale, and the nature of the task with respect to computational and informational requirements. Similar applications may be found in chemical, oil, paper, and other large-scale process plants.

Electric power system. The system decomposes naturally with the following levels explicitly identified.

First level: Generating units (steam generator, turbine, electric generator).

Second level: Set of generating units within a power station.

Third level: Set of power stations belonging to a utility company.

Fourth level: Group of utilities interchanging power.

In addition, a temporal hierarchy may interleave the multilevel structure with successive layers (at increasing time scales) associated with economic dispatch, unit commitment, production scheduling, and maintenance scheduling. *See* ELECTRIC POWER SYSTEMS.

Batch processing. Special features of hierarchical control applied to batch process systems (for example, specialty chemicals and pharmaceuticals) include dynamic optimization and superposition of such discrete decision-making functions as scheduling, time allocation, and production planning.

Multimode operations. An occasional requirement of many production systems is operation in different processing modes or stages, with each mode characterized by a distinct set of objectives, constraints, and process relationships. Typical examples are startup and shutdown sequences for large or critical processing units (for example, distillation columns and turbine-generator units), and restorative control of an electric power system to normal operation following a contingency event (such as a generator outage or line fault). In these applications, multilevel control structures provide effective means of controlling the system within each successive mode, as well as coordinating the transitions from one mode to the next.

Robotics. The control problem associated with robotics lends itself naturally to a multilayer decomposition. Typically, first-layer servocontrollers have responsibility for positional and velocity control of the motion components of the robot. The second level coordinates the actions of the servocontrollers so that the robot achieves the specified task or trajectory. The task specifications are determined at the third level so as to optimize overall performance, satisfying constraints (such as avoiding obstacles), and considering interactions with other robots. Finally, at the fourth level, adaptive modifications are incorporated to respond to changing objectives, environmental conditions, and so forth. *See* ROBOTICS.

Manufacturing systems. These systems are characterized by discrete event models and logic-based controls. Computer-integrated manufacturing (CIM) is the conceptual basis for integrating the information flows and decision-making associated with product design, production planning and scheduling, and plant operations. As in the case of integrated control of process-type systems, a hierarchical control architecture forms a natural basis for implementing the integration. In a typical multilevel scheme, individual machines and processes are controlled at level 1; at level 2, related or sequential machines or processes are coordinated to achieve optimum sequencing, maximum throughput, and efficient use of resources; finally, at level 3, the full system is brought together by using resource planning, maintenance planning, and other systems tools.

A natural extension of the hierarchical approach to computer-integrated manufacturing is to automate the entire process from product specification to manufacture by integrating computer-aided design (CAD), computer-aided process planning (CAPP), and computer-aided manufacturing (CAM). This capability reduces workforce requirements and the time required to implement product changes, a valuable asset in job shop production and in achieving rapid response to changing market demand. This approach can be extended to batch processing systems (as in pharmaceutical production). An important step in this direction is the development of a formal approach to the integration problem, that is, the development of generalized formal models, procedures, and protocols for implementing the integration for different classes of manufacturing and production systems. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING; COMPUTER-INTEGRATED MANUFACTURING.

Air-traffic management systems. A multiscaled hierarchical approach has been considered for the next generation of air-traffic management systems (ATM). The proposed structure consists of a horizontal dimension which relates to geographical air space partitions (or control sectors) and a vertical dimension associated with space-time scale decomposition of the decision-making function. The horizontal dimension provides for coordination of the interfaces

between air space partitions with respect to individual users and ground-based control. The vertical hierarchy comprises, at the first level, air-traffic control (ATC), which is concerned with control of individual flights and their interfaces; at the second level, air space traffic flow management, which sets conditions and constraints for air-traffic control; and at the third level, air space management, which oversees the overall air space structure and sets rules assuring successful operation of the lower decentralized levels of control. Increasing uncertainties at each higher level (corresponding to increasing time horizons) are handled by aggregation of the information provided by the lower levels. *See* AIR-TRAFFIC CONTROL; CONTROL SYSTEMS; DISTRIBUTED SYSTEMS (CONTROL SYSTEMS).

Irving Lefkowitz

Bibliography. S. B. Gershwin, Hierarchical flow control: A framework for scheduling and planning discrete events in manufacturing systems, *Proc. IEEE*, 77:195–208, 1989; Y. Y. Haimes et al., *Hierarchical-Multiobjective Analysis of Large-Scale Systems*, 1990; I. Lefkowitz, Integrated control of industrial systems, *Phil. Trans. Roy. Soc. London*, A287:443–465, 1977; S. N. Salthe, *Evolving Hierarchical Systems*, Columbia University Press, New York, 1985; S. G. Tzafestas and J. K. Pal (eds.), *Real Time Microcomputer Control of Industrial Processes*, Kluwer Academic, Dordrecht, 1990; T. J. Williams, *Analysis and Design of Hierarchical Control Systems*, 1985.

Multimedia technology

Computer-based, interactive applications having multiple media elements, including text, graphics, animations, video, and sound. Multimedia technology refers to both the hardware and software used to create and run such systems.

Applications. The earliest multimedia programs, in the 1970s, were arcade and home video games. Their popularity, combined with the introduction of personal computers in the late 1970s and early 1980s, led to more sophisticated and colorful computer-based games. The development of new hardware then made possible CD-ROM (compact disc-read only memory) technology for storage and distribution of programs requiring large amounts of data, such as encyclopedias and other reference works. New technology also permitted the proliferation of graphics, sound files, and the earliest virtual reality programs. The 1990s brought the Internet and the World Wide Web (WWW) into popular use, and this broadened the audience for multimedia applications of all types.

By the beginning of the twenty-first century, the favorite and most widely used multimedia systems still were those designed for entertainment and games. However, multimedia technology is also used to create extensive reference works, such as encyclopedias, dictionaries, and databases. Businesses use multimedia for briefings, advertising, and sales presentations. Teachers and corporate trainers employ multimedia to design lecture aids or tutorial programs.

The mode of delivery for each application depends on the amount of information that must be stored, the privacy desired, and the potential expertise of the users. Applications that require large amounts of data are usually distributed on CD-ROMs, while personal presentations might be made directly from a computer using an attached projector. Advertising and some training materials are often placed on the WWW for easy public access. Museums make use of multimedia kiosks with touch screens and earphones. A listing of the most common delivery modes used by each kind of application is given in the **table**. *See* COMPACT DISK; INTERNET; VIDEO GAMES; WORLD WIDE WEB.

Delivery modes for common multimedia applications

Application	Most common delivery mode
Briefings, presentations, and advertising	Disk, WWW
Reference and databases	CD-ROM, DVD, WWW
Museums, education, and training	Kiosk, Disk, CD-ROM
Entertainment and games	CD-ROM, DVD, WWW

Hardware. Multimedia products may be created and run on the commonly used computer environments, such as Microsoft Windows[R] and Apple Macintosh.[R] Multimedia system users may employ a variety of input devices in addition to the keyboard and mouse, such as joysticks and trackballs. Touch screens provide both input and display capabilities and are often the choice when potentially large numbers of novices may use the system. Other display devices include high-resolution monitors and computer projectors. Generally the abundance of graphics and video in multimedia applications requires the highest resolution and deepest color capacity possible in display devices.

Input devices for the creation of multimedia applications include graphics tablets, which are pressure-sensitive surfaces for drawing with special pens; digital cameras, which take pictures electronically; and scanners, which convert existing pictures and graphics into digital form. Other hardware devices, such as a video card and video digitizing board, are required both to create and to play digital video elements.

The hardware for incorporating sound elements into multimedia systems includes microphones, voice-recognition systems, sound chips within the computer, and speakers, which come in a wide variety of forms with varying capabilities and quality. *See* SPEECH RECOGNITION.

When all the elements of multimedia are united in a single application, the result is a large file requiring a high-capacity storage medium. Large, mass-produced products are usually packaged using CD-ROM or DVD (digital video disc). CD-ROM players and recorders became commonplace in the early 1990s and enabled the delivery of high-capacity products. DVDs, introduced in the middle 1990s, are capable of even larger storage capacity and full-motion video combined with high-quality audio.

Multimedia applications can also be presented using the WWW. The advantage is that users are

not required to have CD-ROM or DVD players on their computers. Moreover, content can be updated frequently. The disadvantage is that the large multimedia applications result in long download time and uneven quality, especially for users with low-bandwidth connections. As the technology surrounding the WWW improves, this kind of delivery should become more common. *See* COMPUTER GRAPHICS; COMPUTER PERIPHERAL DEVICES; COMPUTER STORAGE TECHNOLOGY; ELECTRONIC DISPLAY.

Component development software. Specific software is needed to develop each component of a multimedia program, such as text, graphics and images, animation, digital audio, and digital video.

Text. Text helps the user navigate through the multimedia system; explains pages, screens, or pictures; and communicates the information or content of the program. Designers vary the effect of text by using different fonts, sizes, and colors available in word processing programs. The weight or thickness of each letter greatly affects readability. Text can also be encoded on WWW pages. Many software programs are available to create unusual lettering that is shadowed, warped, curved, or fuzzy. The basic concept in text design is that it should be clear and readable over the background. It should contribute to the intended communication, not detract from it.

Graphics and images. In multimedia presentations the images should be clear and concise without distracting lines and backgrounds. Complex designs should be used only if they are central to the theme of the system. There are two kinds of image files: vector images and bitmaps. Vector images define the picture using mathematical properties of coordinates, length, or size. They result in very small files, but the pictures are not photo-realistic. Bitmap files represent the image as a series of dots or pixels and store information about the color of each pixel. Bitmap images produce high-resolution, clear images, but the files are quite large. Compression is often used to reduce the file size. Some popular formats for compressed files are Graphics Interchange Format (GIF) and Joint Photographic Experts Group (JPEG and JPG). Software is available that allows the designer to draw vector images, create the bitmaps, scan designs or pictures created elsewhere, and edit or crop any digital image. *See* DATA COMPRESSION.

Animation. Images can be moved across the screen or displayed quickly in sequence to create animation. Many editing programs are available to help the user design computer animation files. In fact, most software packages that allow the editing of GIF files also feature the ability to create animated GIF files in a compressed format. Some software allows the user to specify the beginning and ending point for the animation, and then the program fills in the animation required in the middle. Single images can be warped, stretched, or manipulated in any way, and multiple images can be edited, transformed, or morphed, changing from one image to another.

Digital sound. The addition of sound to a multimedia program adds dramatically to its appeal. Computer generated sound has progressed from the early "beep" to realistic playing of symphonic music. Sound files usually come in one of two formats: WAV and MIDI. Digital audio files (WAV) are similar to bitmap images in that the sound is cut into pieces ("sampled"); and attributes, such as frequency and amplitude of the analog sound wave, are stored about individual samples. When more samples are stored, the file becomes bigger and the resulting sound is more realistic when played. Fewer samples result in smaller files but lower-quality sound. Musical Instrument Digital Interface (MIDI) files are used for the recording and playing of musical instruments on the computer. MIDI files contain header sections with information about the arrangement and timing of the music, and one or more tracks are used to combine the sounds of several instruments. MIDI files are smaller than WAV files, but are more difficult to create, are more dependent upon the hardware for quality, and are not appropriate for spoken dialog. WAV files, however, can be used for any sound and have a more consistent quality when played, but are difficult to edit. *See* MUSICAL INSTRUMENTS; VOICE RESPONSE.

Digital video. Digital video combines the technologies of graphics, animation, and sound and is usually the most challenging part of a multimedia system. Video cameras capture the analog signal of color, synchronization, and sound. Video capture cards then convert the video into digital format that can be played back on the computer.

Newer video cameras are available to record directly into digital form. Also called camcorders, these cameras are capable of capturing both high-resolution digital video and still photos. It is possible to record and transfer the video or still images directly to a computer, TV, or VCR. Software can then be employed to copy or edit the video, or to add special effects.

The quality of video is measured by the frame rate (number of frames per second). A higher frame rate yields better-quality video but larger files. Usually the designer attempts to use the smallest acceptable frame rate to keep the file size as small as possible but maintain adequate quality. The file size is also affected by the frame size, or the window size of the video as it is displayed. Full-screen video of television quality and sound results in unmanageably large file sizes. Therefore most digital video makes use of compression to reduce the data stored in the video file. There are two kinds of compression/decompression programs: lossy and lossless. Lossless compression results in quality equivalent to the original during the playback, but the files are larger. Lossy compression algorithms result in a loss of quality between the original video and the video that has been compressed and decompressed for playback. However, lossy algorithms are usually acceptable for digital video sections of multimedia systems. The two main file standards are JPEG and MPEG. JPEG was developed for use with still images but is useful also for video. Moving Picture Experts Group (MPEG) files are widely used for video for WWW and DVD applications.

Graphical user interface. Central to the multimedia system is the graphical user interface. Elements such as buttons, checkboxes, and text boxes allow the system to be interactive. Usually these graphical elements are created and combined with text and/or images for communication purposes. These features also allow users to have more control of their experience with the multimedia application. *See* HUMAN-COMPUTER INTERACTION.

System development software. Once all the individual elements have been prepared, the entire system must be designed and assembled. The most obvious way to create the overall program is to write the entire program using a programming language. The programmer has complete control of the finished product. Many modern programming language environments allow easy creation of a graphical user interface and enable the programmer to include both sound and images with a minimum of effort. However, this kind of programming usually requires considerable skill and experience.

The easiest way to assemble a multimedia system is to use presentation software. Business people with no programming experience can create multimedia presentations featuring text, animation, and sound in a short period of time. However, this kind of system development limits the finished product to the features provided by the particular software.

Authoring software provides an alternative to both these platforms. The developer needs to learn to use the software, which takes much less time than learning computer programming, but still retains control over the system being designed. Some authoring software enables the user to piece together a program by dragging icons into a flowchart. Other software allows the user to specify what happens at each point in the system time line. Whatever kind of software is used, the finished product is the result of a wide variety of segments woven together into a multimedia system.

Trends. The future of multimedia technology is dependent upon the evolution of the hardware. As storage devices get faster and larger, multimedia systems will be able to expand, and increased use of DVD should result in improved quality. Rising network speeds will increase the possibility of delivering multimedia applications over the WWW. Currently, Virtual Reality Modeling Language (VRML) is used for some WWW applications and may drastically expand the multimedia experience. Virtual reality is becoming more realistic and will stretch the multimedia experience to envelop the user. The one certainty in multimedia technology is that it will continue to change, to be faster, better, and more realistic. *See* VIRTUAL REALITY.

Pauline K. Cushman; Robert A. Kolvoord

Bibliography. E. England and A. Finney, *Managing Multimedia: Project Management for Interactive Media*, 2d ed., Addison-Wesley, 1999; D. Hillman, *Multimedia Technology & Applications*, Delmar, 1998; D. D. Peck, *Pocket Guide to Multimedia*, Delmar, 1999; T. Vaughan, *Multimedia: Making It Work*, 3d ed., Macromedia, 1996.

Multimeter

An instrument designed to measure electrical quantities. A typical multimeter can measure alternating- and direct-current potential differences (voltages), current, and resistance, with several full-scale ranges provided for each quantity. Sometimes referred to as a volt-ohm meter (VOM), it is a logical development of the electrical meter, providing a general-purpose instrument. Many kinds of special-purpose multimeters are manufactured to meet the needs of such specialists as telephone engineers and automobile mechanics testing ignition circuits. *See* AMMETER; CURRENT MEASUREMENT; OHMMETER; RESISTANCE MEASUREMENT; VOLTAGE MEASUREMENT; VOLTMETER.

Multimeters originated when all electrical measuring instruments used analog techniques. They were generally based on a moving-coil indicator, in which a pointer moves across a graduated scale. Accuracy was typically limited to about 2%, although models achieving 0.1% were available. Analog multimeters are still preferred for some applications. For most purposes, digital instruments are now used. In these, the measured value is presented as a row of numbers in a window. Inexpensive hand-held models perform at least as well as a good analog design. High-resolution multimeters have short-term errors as low as 0.1 part per million (ppm) and drift less than 5 ppm in one year. Many digital multimeters can be commanded by and send their indications to computers or control equipment.

Analog multimeters. The core of the meter is the indicating element, chosen to be a compromise between sensitivity and robustness. A d'Arsonval movement requiring 1 mA direct current for full-scale deflection can withstand rough treatment, but draws relatively high currents when measuring voltages. Voltage ranges are provided by a selection of series resistors; their values are proportional to the required full-scale voltage. In the above example, this value is 1 kΩ/V; this constant provides a convenient quality rating for multimeters. A highly sensitive but relatively fragile meter might use a 20- or 10-μA full-scale deflection movement. A frequent compromise is a 50-μA movement, giving a 20-kΩ/V rating. Current ranges are achieved by rearranging the connections to the movement so that it appears across the current-measuring terminals, together with a choice of switch-selectable shunts. Alternatively, a series of current range sockets may be provided in conjunction with a so-called universal shunt. *See* SHUNTING.

Resistance measurement requires an internal source of current, usually a dry battery with a terminal voltage V of 1.5–15 V. The battery, movement, and measuring terminals are connected in series, together with an adjustable limiting resistor R_1 (**Fig. 1***a*). The limiting resistor is adjusted with the terminals shorted to bring the meter deflection to full scale, which is marked zero on the ohms ranges. The larger the resistance connected between the measuring terminals, the smaller the meter deflection, which is equal to $V/(R_1 + R_x)$. The resulting scale

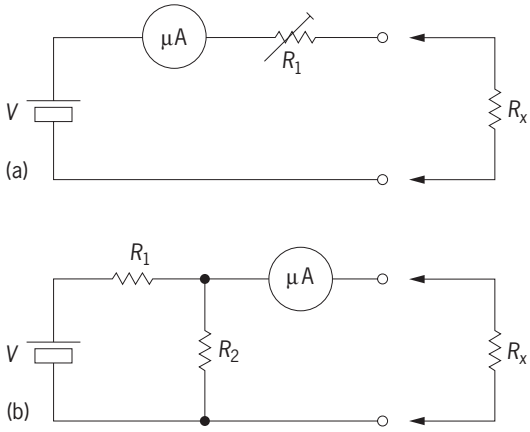

Fig. 1. Circuits for resistance measurements with an analog multimeter over (*a*) normal-resistance range and (*b*) low-resistance ranges.

resembles a logarithmic shape, which conveniently allows a wide range of resistances to be measured on a single range. Other ranges can be obtained with the help of a second battery of a higher voltage or a potential divider (Fig. 1*b*).

Digital test multimeters. Digital multimeters intended for test purposes, often described as hand-held models, are closely related to their analog predecessors and are intended for similar use. They are small, operate from batteries, and provide a digital display that is driven by an analog-to-digital converter. Some instruments also include a quasianalog bar display. In a common version of the dual-slope analog-to-digital converter, the analog input signal is integrated for a fixed time, and the integrator is restored to zero by the application of a reference current. The time taken is measured by counting clock cycles, giving the required digital result. At very little extra cost, a microprocessor can be built into the instrument, allowing a variety of features, such as a bar graph display or automatic ranging. *See* ANALOG-TO-DIGITAL CONVERTER; MICROPROCESSOR.

A power source and active circuits can improve the sensitivity considerably. A typical hand-held unit includes an operational amplifier with an input resistor R_{in} of 10 MΩ to provide all dc ranges (**Fig. 2*a***). Different voltage ranges are provided by switching a feedback resistor R_f, by using field-effect transistors to avoid mechanical contacts and simplify automatic ranging. For ac ranges, a lower value of input resistor may improve the frequency response. *See* OPERATIONAL AMPLIFIER; TRANSISTOR.

Current ranges are provided by measuring the voltage drop across a suitable shunt resistor, using the most sensitive voltage configuration. Restrictions in the power-handling capability of a small shunt limit the current measurement capability of most instruments to 10 A. Resistance ranges are provided by a rearrangement of the input amplifier (Fig. 2*b*). The unknown resistor is used as the feedback element for an operational amplifier which receives a suitably scaled input current from a reference source. The amplifier output is therefore linearly related to the resistor value. A range of reference currents enables a wide variety of resistor values to be measured.

Digital systems multimeters. Systems digital multimeters have totally different requirements. The additional complexity implies higher power consumption, with supplies drawn from the power line. There is little need for economy, except that any excessive heat developed may make the instrument less stable and reliable.

The resolution, linearity, and stability of several designs enable them to make dc measurements within 1 ppm of the standards against which they have been set. This order of accuracy is normally associated with only the best standards laboratories. Although such multimeters contain the same components as hand-held models, the design details of every part are considerably more refined. Special analog-to-digital conversion techniques combine time-division and charge-balance techniques in such a way that errors due to switching transients and dielectric storage are considerably reduced.

Input isolation. The analog-to-digital converter and input circuits are constructed within a screened enclosure (the guard). Their power is provided by carefully isolated transformers. The timing signals that define the measurement result are transmitted to the grounded output circuits through optoisolators or low-capacitance pulse transformers. By this means, the input circuits are well isolated, and the instrument can be used, with certain restrictions, as if it were not operating from the power line. The actual degree of isolation achieved is a measure of the success of the design. It is normally possible for a dc signal applied to both input leads to influence the display as little as a true input signal 140 dB smaller; that is, the common-mode rejection is 140 dB at dc. As it is much more difficult to obtain low capacitance than high resistance, the comparable figure for ac common-mode rejection is much worse, typically 90 dB at 60 Hz, and degrades linearly with frequency.

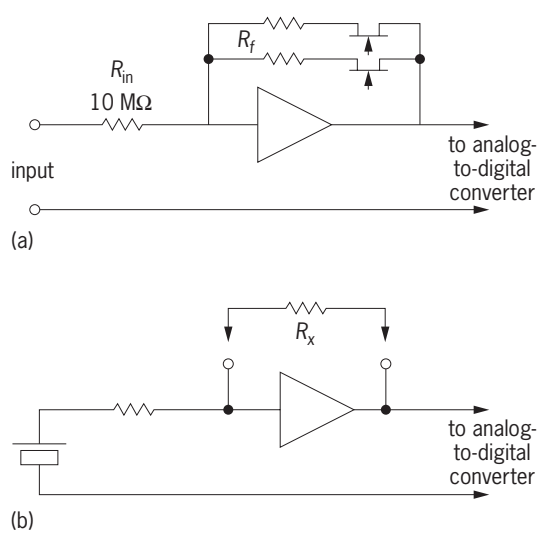

Fig. 2. Digital test multimeter circuits for (*a*) active input ranging and (*b*) active measurement of resistance.

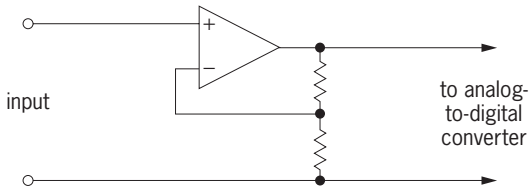

Fig. 3. Circuit for obtaining high input impedance in a digital systems multimeter.

During dc measurements, ac interference is rejected if the integrating time is long or equal to a whole number of periods of the interfering signal. It is also possible to incorporate a filter, resulting in a property defined as normal-mode rejection with a typical value of 70 dB.

Direct-current input circuits. For dc inputs up to about 20 V the signal can be applied to a high-input impedance configuration (**Fig. 3**). Amplifiers can be constructed with noise levels of the order of tens of nanovolts and input resistances higher than 10 GΩ. The offset current, that is, the steady dc current that flows out of the measurement terminals, should be below 1 picoampere if it is not to be the source of significant errors when measurements are made from source resistances above 10 kΩ. So far, few designs have satisfied this requirement. Resistive attenuators provide higher voltage ranges.

Alternating-current input circuits. For ac inputs, the input arrangement resembles that provided for dc in hand-held meters (Fig. 2*a*), except that a lower value of resistor is used in conjunction with a far more elaborate amplifier in order to provide better accuracy of gain and a wide bandwidth. An input resistance of 1 MΩ is typical. In addition, an input capacitance of about 150 picofarads results from stray capacitances and guards designed to reduce frequency errors in the input resistor itself. The ac signal is converted into a dc equivalent [invariably the root-mean-square (rms) value] by a thermal or electronic technique. Thermal techniques depend on the heat generated in some device by the ac signal being matched by that generated by a dc signal automatically adjusted by a feedback arrangement. Electronic methods can involve an analog computation using the accurate logarithmic properties of planar transistors or the multiplying properties of configurations in which the transconductance of a gain stage is controlled by the signal. In the most accurate instruments, errors are reduced by carrying out a subsidiary measurement in which a dc equivalent of the ac signal is passed through the same processing path, and using the resulting output to derive corrections for short-term changes in system gain. Sampling methods analogous to those used in digital sampling wattmeters can also be used. *See* WATTMETER.

Resistance measurements. Resistance measurements are based on principles similar to those of hand-held instruments (Fig. 2*b*), but refined to improve accuracy by many orders of magnitude. Provision is made for the measurement of lower-value resistors (below 100 kΩ) using four-terminal connections. A further refinement is to carry out resistance measurement in two stages, first with and then without the test current in order to eliminate errors caused by thermal electromotive forces. The best instruments achieve accuracy of a few parts per million. *See* THERMOELECTRICITY.

Instrument communication. A vital feature of a systems digital multimeter is its ability to communicate with other instruments. To interconnect instruments from different manufacturers, standard interfaces have been defined, of which two are most commonly found. The IEEE488 bus (also known as the HP-IB and GPIB) carries eight bits of parallel data together with handshaking lines to control data transfer. The RS232 bit serial interface is also widely used, but the control lines and data formats are less well defined, making successful operation a matter of trial and error. USB interfaces are increasingly employed to integrate multimeters under computer control into automated and programmable measurement systems.

R. B. D. Knight

Bibliography. W. E. Boyes (ed.), *Instrumentation Reference Book*, 3d ed., 2002; Institute of Electrical and Electronics Engineers, *IEEE Standard for Higher Performance Protocol for the Standard Digital Interface for Programmable Instrumentation*, IEEE Stand. 488.1–2003, 2003; Institute of Electrical and Electronics Engineers, *IEEE Standard Codes, Formats, Protocols and Common Commands for Use with IEEE Std 488.1–1987, IEEE Standard Digital Interface for Programmable Instrumentation*, IEEE Stand. 488–2, 1992; International Telegraph and Telephone Consultative Committee (RS232), Stand. V24; B. Kibble et al., *A Guide to Measuring Direct and Alternating Current and Voltage Below 1 MHz*, Institute of Measurement and Control, United Kingdom, 2003.

Multiple cropping

Planting two or more species in the same field in the same year. This general term describes a multiplicity of systems that are as complex as agriculture itself. The first conscious planting by early people consisted of mixtures of cereals, grain legumes, roots, and tubers that were located near permanent dwellings. Perhaps this system of planting was patterned after the natural combinations of trees, shrubs, and other plants from which food was gathered in the wild, or resulted from the unintended growth that originated from seed contained in kitchen refuse that germinated and thrived in early fertile garbage dumps.

Preserved through history to maintain biological, economic, and nutritional diversity, these multiple-species systems still are used by the majority of the world's farmers, especially in developing countries. Where farm size is small and the lack of capital has made it difficult to mechanize and expand, farm families that need a low-risk source of food and income often use multiple cropping. These systems maintain a green and growing crop canopy over the soil

through much of the year, the total season depending on rainfall and temperature. Systems with more than one crop frequently make better use of total sunlight, water, and available nutrients than is possible with a single crop. The family has a more diverse supply of food and more than one source of income, with both spread over much of the year.

Multiple-cropping patterns are described by the number of crops per year and the intensity of crop overlap. Double cropping or triple cropping signifies systems with two or three crops planted sequentially with no overlap in growth cycle. Intercropping indicates that two or more crops are planted at the same time, or at least planted so that significant parts of their growth cycles overlap. Relay cropping describes the planting of a second crop after the first crop has flowered; in this system there still may be some competition for water or nutrients. When a crop is harvested and allowed to regrow from the crowns or root systems, the term ratoon cropping is used. Sugarcane, alfalfa, and sudangrass are commonly produced in this way, while the potential exists for such tropical cereals as sorghum and rice. Mixed cropping, strip cropping, associated cropping, and alternative cropping represent variations of these systems.

Efficient resource use. Natural plant communities are highly diverse, and those containing the largest number of species are found in the tropical lowlands. In these ecosystems, depending on rainfall, there is cover over the soil through all or most of the year. In the absence of some cataclysmic event, these systems may be sustained without human intervention over many centuries. Multiple-cropping systems copy some of the characteristics of these natural ecosystems, especially in the use of resources.

Potato, maize, and climbing beans are intercropped by farmers in the Rio Negro area of Antioquia Department in Colombia (**Fig. 1**). Potatoes are planted in January in this highland valley at about 7200 ft (2200 m) elevation, and the crop is hilled up in April and maize is planted alongside the potato rows. The potato crop is harvested in May or June, and farmers hill up the maize while digging the potatoes. Climbing beans are planted in July at the base of the maize plants, and bean plants grow up the maize stalks as the season goes on. Both maize and beans are harvested by hand in December and January, residues are returned to the soil, and the cycle begins again.

Figure 1 shows dry-matter accumulation and overlapping growth curves of three crops, potato, maize, and bean, two of which cover the soil during most of the year. These crops intercept sunlight for photosynthesis, break the impact of the pounding rain to reduce soil erosion, and utilize soil moisture and nutrients for growth. Leaf fall and crop residues increase organic matter and contribute to soil fertility. This multiple-species system is self-sustainable over a long period of time, with minimal use of fertilizers and pesticides. A single crop (monoculture), on the other hand, does not use the natural resources available for growth over as much of the total year, nor would a single crop contribute as much residue or provide as much food and income. The multiple-cropping system clearly can take advantage of the available natural resources to sustain good crop growth.

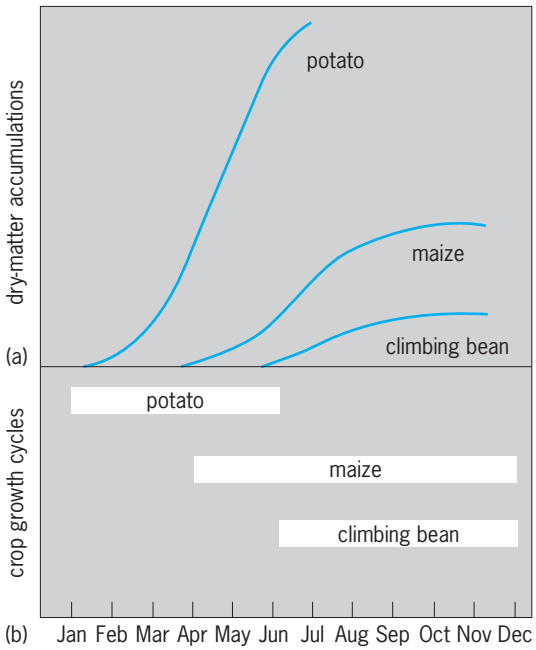

Fig. 1. Graph showing (*a*) dry-matter accumulation curves and (*b*) overlapping crop-growth cycles in a multiple-cropping system of potato, maize, and climbing bean, grown in the same field in Rio Negro, Colombia. Each crop growth cycle is indicated from planting date to harvest date.

Tropical cropping systems. The potato–maize–bean system described is but one example of the diverse group of crops selected for year-round cropping in the tropics. The crops selected and the intensity of the systems depend on such factors as elevation, rainfall, soil, and also food preferences. In highland areas, maize, sorghum, barley, potato, amaranth, and vegetable crops are often components of these systems. Cycles of these crops overlap, and several species may be found in one field at the same time. Because of lower temperatures at these elevations, crop growth is slow. Traditional maize varieties may require 10 to 13 months to mature, while an improved 7- to 8-month variety or hybrid allows the use of more intensive systems such as the potato–maize–bean combination in Fig. 1.

Lowland tropical systems often are more diverse, with more crops planted together and a shorter growth cycle for each crop. An example from the high-rainfall forest zone in Nigeria is shown in **Fig. 2**. Raised mounds are established over a period of years, and upland crops such as cassava, yam, sweet potato, and maize are planted on top of the mounds. Rice is planted in the low areas between mounds where water accumulates and helps control weeds. A similar system in Ecuador in the coastal areas that flood each year has rice planted in the lower areas and part of the way up the sides of

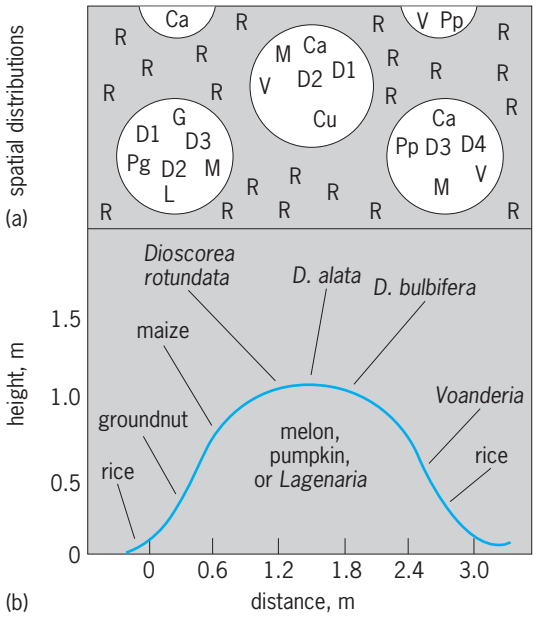

Fig. 2. Diagram showing (*a*) spatial distributions and (*b*) heights of various crop species grown on and between raised mounds in Abakaliki, East Central State, Nigeria. M = maize, R = rice, Ca = cassava, Cu = melon, G = groundnut, L = *Lagenaria*, V = *Vigna*, D1 = *Dioscorea rotundata*, D2 = *D. alata*, D3 = *D. bulbifera*, D4 = *D. cayenensis*, Pp = pumpkin, and Pg = pigeon pea. 1 m = 3.3 ft. (*After C. A. Francis, Variety development for multiple cropping system, CRC Crit. Rev. Plant Sci., 3(2):133–168, 1986*)

mounds. Maize is planted on the higher areas and down the slopes to where rice and maize are both planted together. The area floods each year, but the depth of water is not predictable. In years when there is deep water, the rice on the sides of the mounds does well and maize is killed, while in dry years the maize does well and rice does poorly. This is a risk-avoiding strategy used by small farmers. In Asia the *sorjan* system employs the same principle.

Double and triple cropping of rice in southern China began in the twelfth century when short-cycle varieties were selected. In areas with less rain, short-cycle catch crops of mungbean, soybean, or vegetables can be planted after rice to use residual moisture. Cassava is an 8- to 12-month crop that can be established with late rains and survives through the dry season, providing food and income for families in lowland areas. Other crop combinations are sorghum/maize, sorghum/cowpea or peanut, maize/bean, and annual/perennial mixtures. Agroforestry is another dimension of multiple cropping; in this system useful perennial tree crops are interplanted with shorter-cycle crops. There is a nearly infinite series of mixtures that depend on amount and distribution of rainfall, temperature regime, crop adaptation, and food preference in tropical regions. *See* TERRACING (AGRICULTURE).

Temperate cropping systems. There is less multiple cropping and more monoculture in temperate zones, where mechanization and productivity per unit of labor have shaped agricultural systems. One major exception is pasture systems, such as those in which grasses and legumes are commonly planted together to take advantage of water, nutrients, and light throughout the entire season and to create a more balanced feed source for grazing livestock.

New and less common systems are overseeded or rotated legumes and relay crop patterns of wheat and soybeans (**Fig. 3**). Overseeding legumes into maize provides ground cover through the fall and spring, produces nitrogen for the next cereal crop, and helps prevent erosion of topsoil by breaking the force of raindrops and contributing surface organic matter to absorb rainfall. This practice is gaining in popularity, with hairy vetch and red clover two of the most common legumes in the system. This sequence is illustrated in Fig. 3.

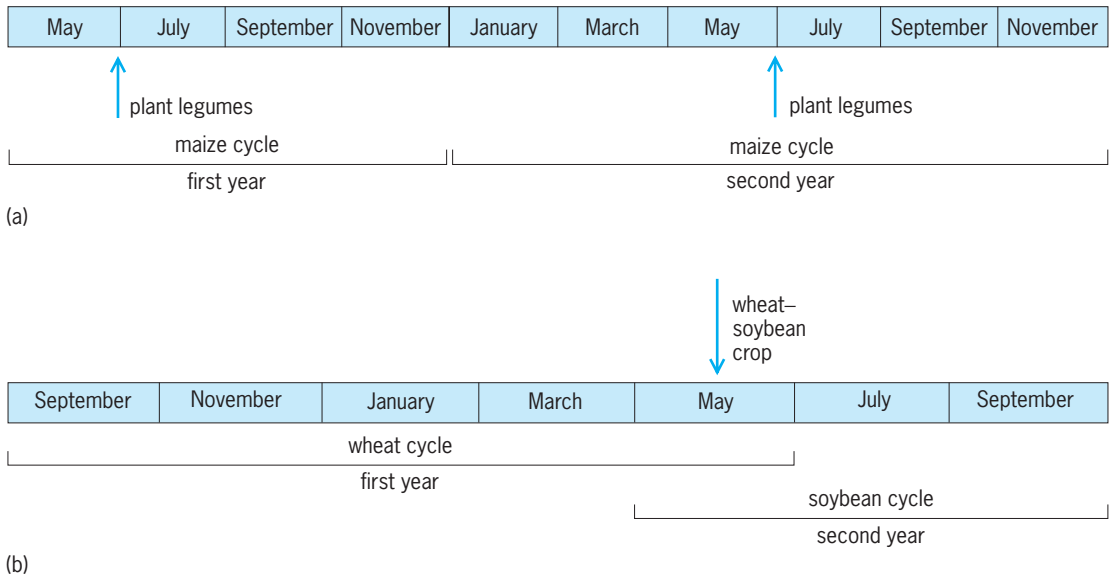

Fig. 3. New multiple-cropping systems for temperate zones in which (*a*) overseeded legumes in maize and (*b*) relay-cropped wheat and soybeans have been planted.

Also shown in Fig. 3 is a relay system where soybeans are seeded into growing wheat in May. The wheat is harvested in July, allowing soybeans to grow to their normal maturity in October. If adequate rainfall occurs, two crops can be produced in one year. Other experimental systems are being tested in temperate zones.

Future of multiple cropping. There will be less multiple cropping in the future in some tropical areas as mechanization leads to consolidation of smaller fields in some favored areas. Complex systems will continue to be popular with farmers who seek maximum food production from small areas and those who engage in subsistence farming. Interest by research and extension specialists has stimulated the search for new varieties and for technologies that can improve the productivity of multiple-species systems.

Greater awareness of the potentials of intensive cropping systems in temperate zones will stimulate their adoption by farmers. Pastures that have an appropriate mixture of species can produce more dry matter and reduce risk over a range of different rainfall patterns. Relay systems provide a potential for greater productivity per unit of soil area and divide the fixed costs of production over two crops. Strip crops planted on the contour can help reduce erosion on sloping lands, stabilizing nutrients and capturing rainfall for the next crops. Multiple-cropping systems are likely to remain an important part of agriculture. *See* AGRICULTURAL SCIENCE (PLANT); AGRICULTURAL SOIL AND CROP PRACTICES; AGRICULTURE; AGRONOMY.

Charles A. Francis

Bibliography. C. A. Francis (ed.), *Multiple Cropping Systems*, 1986; A. A. Gomez and K. A. Gomez, *Multiple Cropping in the Humid Tropics of Asia*, 1983; K. G. Steiner, *Intercropping in Tropical Smallholder Agriculture with Special Reference to West Africa*, 1982.

Multiple proportions, law of

This law states that, when two elements combine together to form more than one compound, the weights of one element that unite with a given weight of the other are in the ratio of small whole numbers. The law can be illustrated by the composition of the five oxides of nitrogen. One gram of nitrogen is combined with 2.85 g of oxygen in nitrogen pentoxide, N_2O_5; with 2.28 g in nitrogen dioxide, NO_2; with 1.71 g in nitrogen trioxide, N_2O_3; with 1.14 g in nitric oxide, NO; and with 0.57 g in nitrous oxide, N_2O. These numbers are in the simple ratio of 5:4:3:2:1. *See* DEFINITE COMPOSITION, LAW OF.

Thomas C. Waddington

Multiple sclerosis

A neuromuscular disorder that characteristically involves the destruction of myelin, the insulating material around nerve fibers. The onset of the disease is unusual in persons under 15 or over 60 years of age, and peak incidence is found in people in their 20s and 30s. Multiple sclerosis affects females more frequently than males by approximately 2:1. Distribution is worldwide, but there is an unusual relationship to latitude, with a much higher incidence at northern latitudes than near the Equator.

Clinical features. The basic mechanism underlying the disease is the destruction of myelin. The effect is analogous to breaking down the insulating material around a telephone wire and exposing the bare wire so that short circuits occur. In multiple sclerosis, only the central nervous system is affected, but both incoming and outgoing processes may be disrupted. Common initial symptoms reflect this underlying disease mechanism. They include blindness in one eye due to disruption of the conduction of the nerve impulse through the optic nerve; weakness of one side of the body due to impairment of the downstream signals from the motor areas of the cerebral cortex through the spinal cord; difficulties with coordination related to problems with cerebellar function; and disturbances in sensation, such as tingling and numbness in an arm or a leg that is related to dysfunction of incoming sensory signals. Multiple sclerosis is a progressive disease, so that over time there is often an accumulation of new symptoms and problems, such as loss of bladder control, difficulty in walking or speaking, and disturbances of mood and cognitive function.

In its classic form, the disease spreads both temporally and anatomically. Temporally, there may be a series of acute attacks, but between attacks a person may recover fully and remain well for some time. Anatomically, areas of disruption of myelin (demyelination) are scattered throughout the nervous system and spinal cord. Thus, symptoms depend upon what part of the nervous system is affected at any given time.

The course of the disease is unpredictable. Some individuals have an acute attack followed by many asymptomatic years. Others have repeated attacks and never quite return to their normal function between attacks. A third group has a chronic progressive form of the disease, usually showing difficulties in walking and use of the hands followed by a subsequent progressive course with no acute exacerbations.

Diagnosis. The diagnosis is based primarily upon the clinical presentation, which makes it very difficult to reach a diagnosis at the time of the first attack. It is the course of the disease that is most helpful in establishing the diagnosis. There are no laboratory tests that define the disease, but abnormalities can aid in the diagnosis. Physiological studies—for example, of the visual, auditory, and sensory systems—can reveal dysfunction of a sensory system that seems clinically normal, indicating that the disease is more widespread than it originally appeared. Areas where myelin has broken down (so-called plaques of demyelination), can be seen by magnetic resonance imaging. However, many people with even a mild form of the disease have the characteristic

lesions of multiple sclerosis on magnetic resonance images.

Lastly, the disturbance of immunological function can be identified by examination of cerebrospinal fluid. There is an increase of immunoglobulins that are synthesized within the nervous system. In addition, the pattern of immunoglobulins shows discrete species of immunoglobulin G (IgG), referred to as oligoclonal bands, which can be detected by electrophoretic studies of cerebrospinal fluid. *See* IMMUNOGLOBULIN.

Etiology. The basic cause of the disease is not known. There clearly are genetic factors in that the incidence of the disease is 20 times higher in first-degree relatives (parents and siblings of an affected individual) than in the general population. If one of a pair of identical twins is affected, the chances of both twins having the disease (the concordance rate) is 28%, whereas in fraternal twins it is 2.5%, which is similar to any group of siblings. Furthermore, certain immunological markers present on lymphocytes are overrepresented in individuals with multiple sclerosis, suggesting a genetic predisposition. However, other factors must be involved, including infection (presumably viral) or possibly an immunological mechanism. The prevailing hypothesis combines these two possible etiological mechanisms to suggest that some type of viral infection occurs early in life to alter the patient's immune system. Thus, the activity and progression of the disease are related to altered immune functions within the central nervous system. *See* HUMAN GENETICS; IMMUNOGENETICS.

Treatment. In keeping with this hypothesis, therapy is aimed at altering the immune status of the affected individual. Therefore, preparations such as prednisone, adrenocorticotropic hormone, cyclophosphamide, or cyclosporine have been employed in attempts to alter the progression of the disease. There is no panacea for multiple sclerosis, however, and therapy is reserved for those who are having acute attacks or who have clear-cut progression of the disease. There is no therapy to prevent further attacks. *See* NERVOUS SYSTEM DISORDERS.

Guy M. McKhann

Bibliography. R. L. Blaylock, *Multiple Sclerosis*, 1980; C. L. Cazzulo et al. (eds.), *Virology and Immunology in Multiple Sclerosis*: *Rationale for Therapy*, 1988; S. U. Kim (ed.), *Myelination and Demyelination*: *Implications for Multiple Sclerosis*, 1989; B. H. Waksman et al., *Research on Multiple Sclerosis*, 3d rev. ed., 1987.

Multiplexing and multiple access

In telecommunications, multiplexing refers to a set of techniques that enable the sharing of the usable electromagnetic spectrum of a telecommunications channel (the channel passband) among multiple users for the transfer of individual information streams. It is assumed that the user information streams join at a common access point to the channel. The term "multiple access" is usually applied to multiplexing schemes by which multiple users who are geographically dispersed gain access to the shared telecommunications facility or channel. Various methods of multiplexing and multiple access are in common use.

Frequency division. In frequency-division multiplexing (FDM) and frequency-division multiple access (FDMA), the passband of a channel is shared among multiple users by assigning distinct and nonoverlapping sections of the electromagnetic spectrum within the passband to individual users. The information stream from a particular user is encoded into a signal whose energy is confined to the part of the passband assigned to that user. All users may transmit their information simultaneously in time and, although their signals thus interfere in time, the composite signal can be separated into its individual user signals at the receiver by frequency-selective filtering. Care must be exercised so that the signals generated by the users do not overlap in the frequency domain. Thus, the energy of each signal must be strictly limited to its section of the passband, and a guard space between adjacent sections may be required. *See* ELECTRIC FILTER; RADIO SPECTRUM ALLOCATION.

Time division. Time-division multiplexing (TDM) and time-division multiple access (TDMA) permit a user access to the full passband of the channel, but only for a limited time, after which the access right is assigned to another user. Normally the access rights are assigned in a cyclical order to the competing users, with each user periodically gaining access for equal amounts of time. In some applications, however, certain users may be given preference over others by allowing more frequent access or access for longer periods of time. Since in a time-division multiplexing system the user information streams are separated in time, the receiver must have accurate information on the start and end times of each user information stream within a cycle. The process of obtaining such information is known as synchronization.

Statistical multiplexing. For applications in which the user information streams occur intermittently in time, the efficiency of utilization of the channel capacity in time-division multiplexing systems may be very low due to the fact that the access time of a particular user in successive cycles is dedicated, irrespective of whether that user requires the channel. Statistical time-division multiplexing overcomes this problem by assigning time on the channel on a demand basis, which typically increases the number of users who may be accommodated on the same channel. On the negative side, the users may experience delay in accessing the channel during periods when the demand exceeds the supply. Furthermore, since a particular user's information may be transported during different time intervals in successive cycles, a certain amount of identifying information must be transferred either by a separate service channel or along with the user's data, thereby increasing the amount of transmission overhead.

Code division. In code-division multiple access (CDMA), all users are assigned the entire passband of the channel and are permitted to transmit their information streams simultaneously. Thus no separation in frequency or time occurs. To maintain the ability to recover the individual signals at the receiver, at the transmitter each signal has impressed on it a characteristic signature. (For example, it may be coded orthogonally to the other signals.) Correlating the composite received signal against the signature of a particular user has the effect of attenuating the signals from all other users to an acceptable level of interference, and thus allows the extraction of a specific signal to a certain degree of accuracy.

Space division. Space-division multiple access (SDMA) refers to the use of the same portion of the electromagnetic spectrum over two or more spatially distinct transmission paths. In most applications of space-division multiple access, the paths are formed by multibeam antennas, in which each beam is directed toward a different geographic area. The performance of such systems is heavily dependent on the degree to which individual beams can be isolated from each other. *See* ANTENNA (ELECTROMAGNETISM).

Wavelength division. Transmission systems that employ the optical portion of the electromagnetic spectrum, such as those using fiber-optic cables as the transmission medium, may share the total available passband of the medium by assigning individual information streams to signals of different wavelengths or "colors." In dense wavelength-division multiplexing (WDM) schemes, a typically large number of such wavelengths is available, and consequently the capacity of such systems to carry multiple information streams is large. *See* OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

Polarization division. Polarization refers to the direction or geometric orientation of the electric field vector of an electromagnetic field. In conventional linear polarization the vector points in either the vertical or horizontal direction in a plane perpendicular to the direction of propagation of the field. In circular polarization the vector rotates in either a clockwise or counterclockwise direction. Polarization-division multiplexing and polarization-division multiple access exploit these differences by assigning electric fields of different polarization to individual channels or users. *See* POLARIZATION OF WAVES; POLARIZED LIGHT.

Performance of multiplexing systems. The performance of the various multiplexing techniques described above is generally stated in terms of the total information-carrying capacity of the channel and the quality of the signals at the receiver. It is heavily dependent on the attenuation characteristics of the channel passband, the distribution of the energy in the signals representing the data streams, the level and type of interference in the channel, and the multiplexing scheme employed. *See* ELECTRICAL COMMUNICATIONS.

Hermann J. Helgert

Bibliography. G. Held, *High Speed Transmission Networking*, Wiley, New York, 1999; S. H. Rowe II, *Telecommunications for Managers*, 4th ed., Prentice Hall, 1999; M. Schwartz, *Information Transmission, Modulation, and Noise*, 4th ed., McGraw-Hill, 1990.

Multiplication

One of the fundamental operations of arithmetic and algebra. The use of the symbol $\times$, commonly employed in arithmetic to denote multiplication, is attributed to the English mathematician William Oughtred (1574–1660). Because of its resemblance to the letter x, it is rarely used in algebra, where multiplication is frequently denoted by a dot (as in $a \cdot b$) or, most often, merely by juxtaposition of letters (for example, ab). Multiplication of numbers (real or complex) is associative, $a(bc) = (ab)c$; commutative, $ab = ba$; and distributive with respect to addition, $a(b + c) = ab + ac$; but the term has been extended to denote binary operations on many other kinds of objects, and these operations need not possess all the properties of ordinary multiplication listed above (for example, multiplication of matrices is not commutative). Much effort was formerly expended in the computation and design of multiplication tables, and the related endeavor of tables of prime numbers. (A whole number is a prime if it is not exactly divisible by any whole number other than itself and 1.) *Crelle's Journal* (1895) gave a 1000×1000 table, and D. H. Lehmer has given a table of all primes from 1 to 10,006,721. Calculators, well adapted for such work, now are used almost exclusively in such computations. *See* ADDITION; ALGEBRA; DIVISION; NUMBERING SYSTEMS; NUMBER THEORY; SUBTRACTION.

Leonard M. Blumenthal

Multipole radiation

Standard patterns of radiation distribution about their source. The term radiation applies primarily to the transport of energy by acoustic, elastic, electromagnetic, or gravitational waves, and extends to the transport of atomic or subatomic particles (as represented by quantum-mechanical wave functions). The time dependence of radiation waves is generally analyzed into sinusoidal components with constant frequency ν and wavelength λ (equal to the wave velocity divided by ν), radiation sources being generally smaller than λ. *See* ELECTROMAGNETIC RADIATION; GRAVITATIONAL RADIATION; QUANTUM MECHANICS; SOUND; WAVE MOTION.

Each multipole pattern reflects the source's geometrical shape (or the shape of a source component). These geometrical features stand out clearly for the static electric potentials generated by fixed charges as shown by the small set of monopole, dipole, and quadrupole charges (see **illus.**), elements of all multipoles being named (in terms of powers of 2)

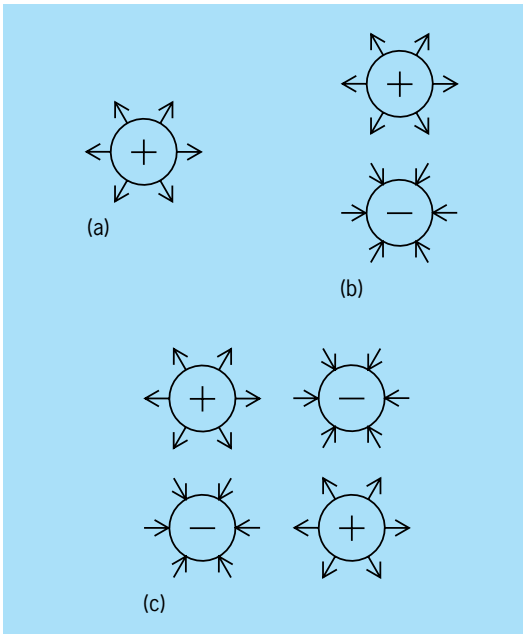

Static electric potentials generated by fixed multipoles. (*a*) Monopole ($l = 0$). (*b*) Dipole ($l = 1$). (*c*) Quadrupole ($l = 2$).

2^l-poles, with l equal to any nonnegative integer. A monopole ($l = 0$) acoustic wave radiates from a perfectly spherical bubble with oscillating radius; higher multipoles would arise from bubble distortions. So-called transverse waves, elastic or electromagnetic (including light), have only $l \geq 1$ components, gravitational waves only $l \geq 2$. The angular distributions, in azimuth (φ) and colatitude (θ), of 2^l-pole waves have amplitudes distributed in directions (θ, φ) in proportion to the spherical harmonic functions $Y_l^m(\theta, \varphi)$. The index m is a positive or negative integer whose absolute value is equal to less than l. *See* COORDINATE SYSTEMS; DIPOLE; SPHERICAL HARMONICS.

Symmetries. The amplitude of each multipole wave is proportional to a function $Y_l^m(\theta, \varphi)$, defined with reference to a symmetry axis, for example, the axis of the simple dipole (illus. *b*). In that case, $l = 1$ and $m = 0$, meaning that the spherical harmonic Y_1^0, which is proportional to $\cos\theta$, does not depend on the azimuth (that is, the longitude) φ. [The function Y_1^0 pertains to the radiation generated by the simple dipole (see illus.), with oscillating length.] For nonzero values of m, Y_l^m depends sinusoidally on $m\varphi$. Changes of the mutual orientation of the coordinate axis and of the radiation source affect only the value of the index m, leaving the source multipolarity index l fixed. The measurable intensity of radiations is proportional to their squared amplitude, which may be proportional to a squared sum over terms proportional to the Y_l^m functions with different m.

Another symmetry concerns the sign reversal of wave amplitudes upon reflection of the source through its center. The dipole's sign reverses upon this reflection, the monopole and quadrupole do not (see illus.). This reflection reverses the sign of electric multipoles only for odd values of l; so-called magnetic multipoles, radiated by current loops with oscillating intensity, reverse their sign only for even values of l. The occurrence of opposite parity classes, represented by $(-1)^l$ or $(-1)^{l+1}$ respectively, extends to all types of multipole radiation. *See* PARITY (QUANTUM MECHANICS); SYMMETRY LAWS (PHYSICS).

Radiation output. The rate of energy radiated, for example, by a radio antenna, depends on the antenna dimension d, wavelength λ, and multipolarity l, the radiation amplitude being proportional to $(d/\lambda)^l$. For light emission by atoms, the ratio d/λ lies below 10^{-3}, making high-multipole radiations very weak. For gamma-ray emission by nuclei, d is on the order of 10^4 lower than for atoms, but λ may be 10^6 times lower, thus favoring the output of higher-multipole radiations. [The amplitude of magnetic multipole emissions is also proportional to $(v/c)^l$, where v stands for the current carrier's speed, and such emissions thus become important only as v approaches the light velocity, c.] For gravitational radiation, from stellar collisions or collapse sources, both d and λ are huge; terrestrial receiver antennas are, of course, far smaller; this mismatch has prevented its detection thus far. *See* ANTENNA (ELECTROMAGNETISM); ATOMIC STRUCTURE AND SPECTRA; GAMMA RAYS; RADIOACTIVITY.

Dynamical elements. The multipolarity index l also represents the number of angular momentum quanta $\hbar$ (Planck's constant divided by 2π) radiated together with each energy quantum $h\nu$ (phonon, photon, graviton, and so forth). Detection and measurement of received energy quanta, together with measurement of their detection rate and mapping of their directional distribution, generally serve to diagnose the mechanics of the radiation source, whether atomic, molecular, nuclear, particle-decay, or astrophysical process. Energy and momentum conservation underlie this analysis; so does the conservation of angular momentum which states that the initial angular momentum of the source ($\vec{Ji}$) equals the vector sum of the final angular momentum of the source ($\vec{Jf}$) and the angular momentum of the radiation ($\vec{l}$). The quantitative implications of this vector relation are studied by the branch of quantum theory called angular momentum algebra. The balancing of parity, that is, of each variable's sign reversal (or persistence) under reflection through the source's center, also contributes to the analysis of experimental data. Further, more complex angular-momentum considerations play a role in the analysis of the behavior of spin-carrying particles. *See* ANGULAR MOMENTUM; CONSERVATION LAWS (PHYSICS); GRAVITON; PHONON; PHOTON; SELECTION RULES (PHYSICS); SPIN (QUANTUM MECHANICS).

Ugo Fano

Bibliography. H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms*, 1957; J. M. Blatt and V. F. Weisskopf, *Theoretical Nuclear Physics*, 2d ed., 1979; R. Courant and D. Hilbert, *Methods of Mathematical Physics*, vol. 1, 1953; J. D. Jackson, *Classical Electrodynamics*, 3d ed., 1998.

Multiprocessing

An organizational technique in which a number of processor units are employed in a single computer system to increase the performance of the system in its application environment above the performance of a single processor. In order to cooperate on a single application or class of applications, the processors share a common resource. Usually this resource is primary memory, and the multiprocessor is called a primary memory multiprocessor. A system in which each processor has a private (local) main memory and shares secondary (global) memory with the others is a secondary memory multiprocessor, sometimes called a multicomputer system because of the looser coupling between processors. The more common multiprocessor systems incorporate only processors of the same type and performance and thus are called homogeneous multiprocessors; however, heterogeneous multiprocessors are also employed. A special case is the attached processor, in which a second processor module is attached to a first processor in a closely coupled fashion so that the first can perform input/output and operating system functions, enabling the attached processor to concentrate on the application workload. *See* COMPUTER STORAGE TECHNOLOGY; OPERATING SYSTEM.

Classification. Multiprocessor systems may be classified into four types: single instruction stream, single data stream (SISD); single instruction stream, multiple data stream (SIMD); multiple instruction stream, single data stream (MISD); and multiple instruction stream, multiple data stream (MIMD). Systems in the MISD category are rarely built. The other three architectures may be distinguished simply by the differences in their respective instruction cycles:

In an SISD architecture there is a single instruction cycle; operands are fetched serially into a single processing unit before execution. Sequential processors fall into this category.

An SIMD architecture also has a single instruction cycle, but multiple sets of operands may be fetched to multiple processing units and may be operated upon simultaneously within a single instruction cycle. Multiple-functional-unit, array, vector, and pipeline processors are in this category. *See* SUPERCOMPUTER.

In an MIMD architecture, several instruction cycles may be active at any given time, each independently fetching instructions and operands into multiple processing units and operating on them in a concurrent fashion. This category includes multiple processor systems in which each processor has its own program control, rather than sharing a single control unit.

MIMD systems can be further classified into throughput-oriented systems, high-availability systems, and response-oriented systems. The goal of throughput-oriented multiprocessing is to obtain high throughput at minimal computing cost (subject to fail-soft equipment redundancy requirements) in a general-purpose computing environment by maximizing the number of independent computing jobs done in parallel. The techniques employed by multiprocessor operating systems to achieve this goal take advantage of an inherent processing versus input/output balance in the workload to produce balanced, uniform loading of system resources with scheduled response.

High-availability multiprocessing systems are generally interactive, often with never-fail real-time online performance requirements. Such application environments are usually centered on a common database and are almost always input/output-limited rather than computer-limited. Tasks are not independent but are often interdependent at the database level. The operating system goal is to maximize the number of cooperating tasks done in parallel. Such systems may also process multiple independent jobs in a background mode. The additional hardware redundancy in a fault-tolerant system over a general-purpose multiprocessor can be considered a trade-off against software complexity and the time required for software check-pointing in a sequential mainframe system. *See* FAULT-TOLERANT SYSTEMS; REAL-TIME SYSTEMS.

The goal of response-oriented multiprocessing (or parallel processing) is to minimize system response time for computational demands. Applications for such systems are naturally computer-intensive, and many such applications can be decomposed into multiple tasks or processes to run concurrently on multiple processors. In the past, successful SIMD and MIMD parallel processors were often special-purpose machines dedicated to a single class of scientific or real-time signal processing applications. The interest in high-performance, low-cost computers able to handle combined numeric, symbolic, and signal processing tasks concurrently, for so-called fifth-generation applications, together with the availability of low-cost very large scale integrated-circuit (VLSI) microprocessors has rekindled interest in response-oriented multiprocessing. *See* CONCURRENT PROCESSING; INTEGRATED CIRCUITS; MICROPROCESSOR.

Throughput-oriented multiprocessing. The performance of a classical shared-memory multiprocessor is limited by the so-called bandwidth of its shared memory (the total data transfer capacity of the memory bus). Access conflicts further reduce effective bandwidth and thus overall system performance. Studies on early multiprocessors showed interesting results on throughput loss as a function of effective memory bandwidth. For example, if a single processor had one unit of throughput, its dual processor had only 10% less throughput than two single processor systems, and a triple processor had 20% less throughput than three individual systems in a multicomputer rather than multiprocessor configuration. This was not a high penalty to pay for fail-soft function in an airline reservation system with 3000 remote agent terminals. Multiprocessors now exhibit similar performance characteristics for up to six processors, that is, up to their effective memory bandwidth performance limit which has been enhanced by a higher degree of memory

interleaving over earlier multiprocessor systems. Operating system software provides the key fail-soft capability in a throughput-oriented multiprocessor. The performance cost of software check-pointing is higher in sequential processors than in a multiprocessor. Since system protective tasks are redundant, the more processors the better, up to the effective bandwidth limits of the system's shared resources.

High-availability multiprocessing. Fault-tolerant multiprocessor systems were a natural development from throughput-oriented multiprocessor systems. While the trade-off that achieves fault tolerance is a hardware one—that is, more hardware units are used in order to achieve greater system availability—the technology employed is primarily a software one. Lower-level hardware redundancy is used in many such systems, but its successful deployment in applications is still a software issue. The basic requisite for a highly available system in most applications is that each major hardware and software component must at least be duplicated. The system requires two or more processors, two paths connecting the processors, and two paths from the processors to the database. The system's disk controllers and communication controllers must be multiported, so that they may be connected to multiple processors. A high-availability database-oriented system requires five essential software ingredients: a network-communication subsystem, a data-communication subsystem, a database manager, a transaction manager, and an operating system. The network communication subsystem supports interprocess communication within a cluster of locally distributed processors. If the highly available system is also a node on a geographically distributed system, the communication subsystem must also support internode communication. *See* DATA COMMUNICATIONS; DATABASE MANAGEMENT SYSTEM.

Response-oriented multiprocessing. The ideal performance characteristic for an N-processor system, on which a given problem could be partitioned into N similar tasks, would be a linear relationship between performance (the rate at which the system can solve a problem, in units of single-processor performance) versus the number of processors. M. Minsky was skeptical of this ideal, conjecturing that for large N the best hope was for $\log_2 N$ performance. In 1967 G. Amdahl suggested Amdahl's (second) law, stating that if a computer has two speeds of operation the slower mode will dominate performance even if the faster mode is infinitely fast. This leads to a $N/\log N$ performance in a multiprocessor performing a single application in multitask mode.

Minsky's conjecture now seems much too pessimistic, and parallel processing performance gains even greater than those predicted by Amdahl's law have become the goal. A sophisticated technique has been developed for the extraction of parallelism from FORTRAN DO-loops that can routinely exceeds Amdahl's law for SIMD machines. While 100% efficiency is probably not attainable, many highly tailored applications based on careful manual extraction of parallelism have achieved efficiency ratings in the 80–90 range. *See* PROGRAMMING LANGUAGES.

Most computer engineers and architects see a high degree of multiprocessing or parallel processing as essential for achieving performance requirements for future computer systems for scientific computation, for fifth-generation or artificial intelligence (AI) applications, and for dedicated and embedded multiprocessors in control or automation systems. It is also widely accepted that a new parallel application and programming technology must be developed to make effective use of multiprocessors having hundreds or thousands of processors. The promise of high-performance systems made up of low-cost high-volume components (for example, microprocessor chips) rather than high-cost-volume components draws computer architects and applications specialists toward the development of parallel processing tecnology. *See* ARTIFICIAL INTELLIGENCE; COMPUTER ARCHITECTURE; DIGITAL COMPUTER; EMBEDDED SYSTEMS.

Peter C. Patton

Bibliography. T. J. Fountain and M. J. Shute (eds.), *Multiprocessor Computer Architectures*, 1990; E. Gelenbe, *Multiprocessor Performance*, 1990; N. Suzuki (ed.), *Shared Memory Multiprocessing*, 1992; D. Tabak, *Multiprocessors*, 1989.

Multituberculata

An extinct order in the class Mammalia, subclass Allotheria, comprising a major group of early mammals, ranging from Late Jurassic to Late Eocene (about 155 to 35 million years ago). The group is best known from North America, Mongolia, and Europe. Most multituberculates were mouselike in size, although the North American Paleocene *Taeniolabis* was larger, probably closer to the woodchuck (*Marmota monax*) in its proportions. Multituberculates appear primarily to have been terrestrial creatures, but some were probably arboreal (living in trees) and others fossorial (adapted for digging). The Eocene decline and extinction of multituberculates may reflect competition from small placental mammals, especially primates and rodents. *See* ARCHAIC UNGULATE; RODENTIA.

Dentition. Although clearly mammalian and often characterized as rodentlike, multituberculates are unlike rodents or other living mammals in several important features. The dentition is distinctive and, with associated jaw fragments, it is all that is known for most species. In the most primitive multituberculates (traditionally classified in the suborder Plagiaulacoidea, Late Jurassic to Early Cretaceous), the upper jaw holds three incisors, a canine, and five premolars, and the lower jaw one incisor and four premolars (the lower canine is never present); the second upper incisor and the lower incisor are somewhat enlarged and procumbent. The anterior upper premolars tend to be peglike, but the last upper premolar is larger and more trenchant; the lower premolars are narrow, serrated blades. Each jaw has two molars; these are low- crowned, with their cusps set

in two parallel anteroposterior rows. *See* DENTITION.

Advanced multituberculates, traditionally classified in the suborders Ptilodontoidea and Taeniolabidoidea (both Late Cretaceous to Eocene), have two upper incisors; the upper canine and the first upper premolar and the anterior lower premolars are absent (a well-developed diastema, a space between two types of teeth, separates incisors from premolars). Molar cusp numbers have increased, and an internal third row of cusps has been added to the upper molars. In ptilodontoids the lower incisor is longer and more slender than in plagiaulacoids, and the last (fourth) lower premolar forms an arcuate, serrated blade that sheared against the bladelike most posterior upper premolar; the third lower premolar is a small peglike tooth. In taeniolabidoids the central incisors are much enlarged and in some species have become gnawing teeth with enamel restricted to an anterior band, as in rodents. The most posterior premolars are often reduced and their shearing function lost, while the molars become comparatively enormous, forming a specialized grinding "mill."

Functional studies using living mammals as analogues suggest that multituberculate incisors were used in grasping, the bladelike premolars in cutting, and the molars in grinding food. Microscopic wear patterns on multituberculate tooth enamel, however, show that jaw motion differed radically from that in other mammals: during chewing, the mandible moved only in the vertical plane, not transversely, and was strongly retracted during the power stroke. These studies also imply that most multituberculates were probably omnivores, although species having gnawing incisors were likely specialized herbivores feeding on tough plant tissues.

Classification. Ongoing research reveals that the traditional classification of multituberculates is inadequate: the relationship between plagiaulacoids and ptilodontoids + taeniolabidoids (cimolodontans) is uncertain, as is that between ptilodontoids and taeniolabidoids. The suborder Taeniolabidoidea itself lacks diagnostic characters and likely includes separate lineages having convergent dental resemblances. Unfortunately, ambiguities of this kind will continue to plague multituberculate systematics until better evidence is available in both the anatomic completeness of collected specimens and the phylogenetic continuity between species.

Multituberculate relationships to other mammals are unclear. Although the multituberculate skull and inner ear differ from those in Theria (marsupials + placentals and their antecedents), they fail to identify a special relationship between multituberculates and alternative groups of other, nontherian mammals. For example, much of the posterolateral wall of the multituberculate braincase is formed by an expanded petrosal bone (another bone, the alisphenoid, occupies this space in therians), but this pattern is widespread among nontherians and its phylogenetic meaning unclear. Aspects of the multituberculate postcranial skeleton are therian-like, implying to some workers that multituberculates and therians are sister groups. Unfortunately, these resemblances are probably nonhomologous: in multituberculates, therian-like postcranial features occur in only a few advanced species, but they also evolved independently in triconodontids, primitive Mesozoic mammals that have no special relationship to therians. Dental evidence is of particular importance to the question of multituberculate relationships: although the dentition of plagiaulacoids is demonstrably more primitive than that of cimolodontans, it preserves no remnant of passage through a stage in which a therian-like reversed triangle pattern of the postcanine teeth was developed. Consequently, the multituberculate dentition is most parsimoniously interpreted as a modified anteroposterior cusp-in-line dentition, from that in more primitive, nontherian mammals. *See* ALLOTHERIA; MAMMALIA; MONOTREMATA; THERIA.

Richard C. Fox; Thomas Martin

Bibliography. F. A. Jenkins, Jr. and D. W. Krause, Adaptations for climbing in North American multituberculates (Mammalia), *Science*, 220:712–715, 1983; Z. Kielan-Jaworowska, R. L. Cifelli, and Z.-X. Luo, *Mammals from the Age of Dinosaurs*, Columbia University Press, 2004; Z. Kielan-Jaworowska and J. H. Hurum, Phylogeny and systematics of multituberculate mammals, *Palaeontology*, 44:389–429, 2001; D. W. Krause, Jaw movement, dental function, and diet in the Paleocene multituberculate *Ptilodus*, *Paleobiology*, 8:265–281, 1982; Z.-X. Luo, Z. Kielan-Jaworowska, and R. L. Cifelli, In quest for a phylogeny of Mesozoic mammals, *Acta Palaeontol. Pol.*, 47:1–78, 2002; F. S. Szalay, M. J. Novacek, and M. C. McKenna (eds.), *Mammal Phylogeny*, vol. 1, Springer-Verlag, 1993.

Multivariable control

The control of systems characterized by multiple inputs, which are usually referred to as the controls; or by multiple outputs, which are often the measured variables and the variables to be controlled (**Fig. 1**); or by both multiple inputs and multiple outputs (MIMO). Automobiles, chemical processing and manufacturing plants, aircraft and aerospace vehicles, biological systems, and the national economy are examples of multivariable systems which require and receive some form of regulation or control, be it mathematically contrived or not. Control systems are based on mathematical models of the behavior to be controlled, and the methodologies rely on sound mathematical theories. This article focuses on

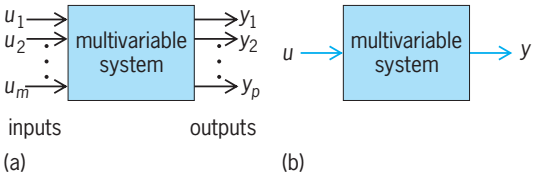

Fig. 1. Two multivariable system representations. (*a*) Input and output variables shown separately. (*b*) Input and output variables shown in vector form.

mathematical techniques which are used to design automatic controllers for multivariable systems.

Most of the mathematical control theory which has been developed for multivariable systems has assumed, as a starting point, some form of known linear, time-invariant, continuous, and causal dynamical mathematical model for the system to be controlled. In that model the manipulated inputs and the outputs to be controlled have been identified. Such a mathematical description of the dynamic behavior of variables of interest may be derived directly from first physical principles and then simplified, for example via linearization of a nonlinear description around an operating point.

The most common description is the state-space or state-variable representation of the form of Eqs. (1),

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{1a}$$

$$y(t) = Cx(t) + Du(t) \tag{1b}$$

where $x(t)$ is the n-dimensional state, $u(t)$ is the m-dimensional input, and $y(t)$ is the p-dimensional output. The corresponding $p \times m$ proper, rational, transfer matrix P of the system is given by Eq. (2),

$$P(s) = C(sI - A)^{-1}B + D \tag{2}$$

where s is the Laplace transform variable. All of the concepts that apply to linear systems, such as the notions of stability, controllability, observability, minimal realizations, and poles and zeros, therefore apply here as well. *See* LINEAR SYSTEM ANALYSIS.

Although the following discussion focuses on linear, time-invariant systems described by Eqs. (1) or (2), some of the multivariable control design methods described below apply to more general systems. For example, the linear quadratic optimal control methodology also applies to time-variant systems, where the parameters—matrices A, B, C, and D in Eq. (1)—depend on time t. Furthermore, all the results discussed below have corresponding results applicable to discrete-time systems described by difference equations instead of differential equations, and transfer matrices expressed in terms of the z-transform variable z instead of the Laplace transform variable s. *See* OPTIMAL CONTROL (LINEAR SYSTEMS); Z TRANSFORM.

Control objectives. The control objectives to be achieved in the multivariable (MIMO) case include those control objectives which are generally sought in the scalar, that is, single input/single output (SISO) case, and also depend on the specific application being considered. In particular, stability is usually a primary concern in the design of any control system, along with various measures of robustness to uncertainties in the plant model.

While controller failure in the scalar case can lead to catastrophic failure of the overall system, such need not be true in the multivariable case due to the interactions between various input/output pairs. More specifically, in certain applications it may be possible to design a multivariable controller "driven" by all p outputs which performs satisfactorily even when feedback information from one or more outputs is lost; or the same control objectives may be attained with fewer than m available control inputs. The integrity of such a multivariable controller would be much better than that of its scalar counterpart. *See* CONTROL SYSTEM STABILITY.

In many control applications, a primary objective is that of tracking, or ensuring that the system output or outputs track a desired input or inputs with little or no steady-state errors. In addition, simultaneous regulation of external disturbances is also often desired; that is, ensuring that any external disturbances affect the plant outputs as little as possible. These performance objectives are to be attained with adequate robustness to uncertainty in the plant and the environment. It is also obvious that simplicity of controller design is a highly desirable design objective in both the scalar and multivariable cases. One final design objective that is unique to the multivariable case is that of minimizing or eliminating the interaction between loops, which is often referred to as decoupling; that is, perfect decoupling would imply a controlled system where each input affects one and only one output or, otherwise stated, a system whose compensated transfer matrix is diagonal and nonsingular.

General considerations. In multivariable control, given a linear, time-invariant, causal system to be controlled described by equations of the type of Eqs. (1) or (2), a linear, time-invariant, causal controller is to be determined that satisfies the control specifications. Here, a causal controller is a controller with proper transfer matrix $C(s)$; $C(s)$ is called proper if $\lim_{s\to\infty} C(s) < \infty$. Note that more complex controllers may be used, such as nonlinear or switching controllers, but they are considered in practice only when absolutely necessary as they may complicate the control design considerably.

It is useful to consider a general, linear, time-invariant controller that has two degrees of freedom. Two transfer matrices, $C_y(s)$ and $C_r(s)$, may be selected, described in Eq. (3), where U, Y, and R are

$$\begin{aligned} U(s) &= C_y(s)Y(s) + C_r(s)R(s) \\ &= [C_y(s), C_r(s)] \begin{bmatrix} Y(s) \\ U(s) \end{bmatrix}^T \end{aligned} \tag{3}$$

the Laplace transformed variables of the input u, output y, and r, an external reference input; in general Y^T denotes the transpose of Y. The controller $C = [C_y(s), C_r(s)]$ is a rational transfer matrix to be determined. Most of the literature on multivariable control explores methodologies to design the feedback controller $C_y(s)$ to satisfy control specifications such as stability and low sensitivity to parameter variations and disturbances. The controller $C_r(s)$ is also important when some specific response to the reference input r is to be attained. Particular control configurations may impose restrictions on $C_y(s)$ and $C_r(s)$ that may not be selected independently any longer; for example, in the unity or error feedback configuration, $C_r(s) = -C_r(s)$.

Under the control law of Eq. (3), the transfer matrix $T(s)$ [such that $Y(s) = T(s)R(s)$] of the controlled system is given by Eq. (4),

$$T(s) = P(s)M(s) \tag{4}$$

where $M = (I - C_yP)^{-1}\, C_r(s)$; here, for simplicity, no disturbances are considered ($Y = PU$). Note that $U = MR$; that is, M is the open-loop equivalent transfer matrix between U and R. It can be shown that the plant P (satisfying $Y = PU$) under the control law of Eq. (3) is internally stable, all poles are stable, and no cancellation of unstable poles occurs if and only if (a) the feedback controller $U = C_yY$ internally stabilizes the system $Y = PU$, and (b) $C_r(s)$ is such that both M and PM are stable, where the rational matrix $M = (I - C_yP)^{-1}C_r$.

In Eq. (4), given a proper and stable T, which determines the desired response of the controlled system, a proper and stable M needs to be determined. Here, M specifies the control action needed, and it is implemented via some combination of feedback and feedforward configuration of proper controllers depending on the problem under consideration. The control input U is almost never implemented via an open-loop controller M. The uncertainties in the plant and environment make it necessary to use a feedback control mechanism. A variety of feedback configurations can be used to implement M, such as linear state feedback with or without state observers or other types of output feedback.

Fundamental limitations. The requirements for stability and causality impose restrictions on attainable control objectives even in the most general case of linear control of Eq. (4); it can be shown that T must contain all the unstable zeros of the plant P, which imposes limitations on the attainable response.

Another fundamental limitation, which is perhaps more widely known, refers to feedback loop properties and is typically expressed in terms of the sensitivity function in the frequency domain via the Bode integral. It states that there is a trade-off in sensitivity; if sensitivity is designed to be low for, say, low frequencies, then unavoidably it will be high at high frequencies, with the situation becoming more restrictive when the system has unstable poles or zeros.

Achievement of objectives. With respect to the achievement of certain or all of the above multivariable design objectives, there are several approaches that may be classified according to the controller type used, whether the approaches are optimal, whether they are time or frequency domain approaches, whether they are driven primarily by industrial applications, or whether they are direct extensions of scalar-case SISO methodologies. Thus, there are state-feedback or output-feedback methodologies that are primarily concerned with pole assignment together with regulation or tracking; methodologies that cast the problems in a cost minimization framework, either in the time domain, such as the linear quadratic regulator problem, or in the frequency domain, such as the H_∞ approach; methodologies such as the model predictive control approach, which has proven itself in industrial applications; and methodologies that are extensions of classical design techniques. These approaches are described below.

Extensions of classical design techniques. There are multivariable controller design techniques that represent extensions of the well-known and thoroughly tested classical, frequency-response techniques such as the root locus and Nyquist stability criteria. Such techniques are usually based on the ability to first decompose the given multivariable system into essentially independent scalar subsystems that can be compensated individually. One of the foremost procedures, the inverse Nyquist array method, involves the concept of a diagonally dominant transfer matrix description of the system to be controlled. More specifically, the $p \times p$ transfer matrix $P(s) = [p_{ij}(s)]$ is said to be diagonally dominant on some contour Γ of the complex plane if for all $i = 1, 2, \ldots, p$, either Eq. (5*a*) or (5*b*) holds, where $|p_{ij}(s)|\Gamma$ represents the

$$|p_{ii}(s)|_\Gamma > \sum_{j \neq i} |p_{ij}(s)|_\Gamma \tag{5a}$$

$$|p_{ii}(s)|_\Gamma > \sum_{j \neq i} |p_{ji}(s)|_\Gamma \tag{5b}$$

magnitude of $p_{ij}(s)$ for any and all s on the contour Γ. Design procedures have been developed for achieving diagonal dominance prior to the employment of the inverse Nyquist array method, which essentially extends the scalar Nyquist stability criteria to the multivariable case.

Another technique that relies on multivariable system decomposition is based on the notion of a dyadic transfer matrix. A $p \times p$ transfer matrix $P(s)$ is said to be dyadic if there exist constant $p \times p$ matrices P_1 and P_2 such that $P(s)$ satisfies Eq. (6). If the dynamical behavior of a system is characterized by a dyadic transfer matrix, it can be decomposed as in **Fig. 2**. Many systems do exhibit dyadic behavior, and in such cases a forward path controller with transfer matrix given by Eq. (7) results in a unity-feedback closed-loop system with the closed-loop transfer matrix given by Eq. (8).

$$P(s) = P_1 \begin{bmatrix} g_1(s) & 0 & \ldots & 0 \\ 0 & g_2(s) & \ldots & 0 \\ 0 & 0 & \ldots & g_p(s) \end{bmatrix} P_2 \tag{6}$$

$$K(s) = P_2^{-1} \begin{bmatrix} k_1(s) & 0 & \ldots & 0 \\ 0 & k_2(s) & \ldots & 0 \\ 0 & 0 & \ldots & k_p(s) \end{bmatrix} P_1^{-1} \tag{7}$$

$$T(s) = P_1 \operatorname{diag} \left\{ \frac{g_i(s)\, k_i(s)}{1 + g_i(s)\, k_i(s)} \right\} P_1^{-1} \tag{8}$$

This implies that the dynamical performance of the closed-loop system is governed by the

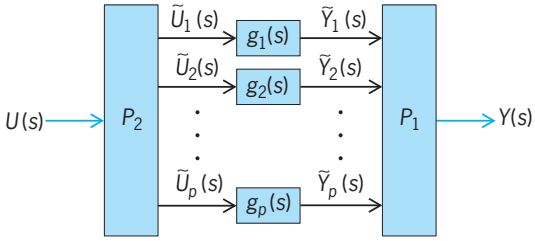

Fig. 2. Decomposition of a dyadic system *P*(s).

performance of the p scalar feedback systems (9), which can be individually designed by using stan-

$$\frac{g_i(s)\,k_i(s)}{1 + g_i(s)\,k_i(s)} \tag{9}$$

dard classical techniques, essentially independent of one another.

Other examples of multivariable extensions of classical design techniques include the characteristic root locus design method and the sequential design method. Virtually all of these classical techniques employ a restrictive form of controller structure in order to facilitate the design.

Control approaches based on state-variable feedback and state estimators. Powerful control design techniques have been developed based on the time-domain, state-variable description of Eq. (1) and the linear state-variable feedback (lsf) control law in Eq. (10), where

$$u(t) = Fx(t) + r(t) \tag{10}$$

$r(t)$ is an m-dimensional external reference input.

Combining Eqs. (1) and (10), it can be seen that the closed-loop dynamics are determined by the eigenvalues of the closed-loop matrix given by Eq. (11).

$$A_F = A + BF \tag{11}$$

The eigenvalues of A_F are exactly the poles of the closed-loop transfer matrix when the system is controllable and observable; the uncontrollable or unobservable eigenvalues cancel in the feedback loop, and in that case the closed-loop poles are only a subset of the eigenvalues of A_F. A celebrated result shows that all the eigenvalues of A_F can be assigned any n desired real or complex-conjugate values if and only if the pair (A, B) is controllable. An important observation is that contrary to the single-input case, given n desired eigenvalues, the feedback matrix F that assigns these eigenvalues to A_F is not unique. This flexibility can be conveniently expressed in terms of the corresponding eigenvectors of A_F where m out of n elements of each eigenvector may be arbitrarily assigned to achieve additional design objectives. These expressions are particularly convenient when the control specifications are expressed in terms of desired eigenvector elements, as is the case in problems such as the flight control of air vehicles.

When the states are not directly measurable (they may be inaccessible or too expensive to measure), they are estimated via an asymptotic state estimator or Luenberger observer, which is driven by the given system's inputs and measurable outputs. The design of the Luenberger observer is the dual to the state-feedback control problem. The separation principle allows the state feedback and the observer to be designed separately, which greatly simplifies the control design. Optimal control and state estimation methodologies, such as the linear quadratic regulator and the Kálmán filtering approaches, which are based on quadratic cost criteria, may be used to determine optimal controllers and observers. *See* ESTIMATION THEORY; STOCHASTIC CONTROL THEORY.

Linear state-variable feedback, with possibly an observer L, is combined with input dynamic compensation to achieve a greater diversity of design goals, such as arbitrary eigenvalue or pole placement, decoupling, model matching, output regulation, and zero-error tracking. An input dynamic compensator may be simply seen as a proper series compensator, $K(s)$, placed in the feedforward loop (**Fig. 3**). Such a control configuration may be seen as a specific implementation instance of the map $M(s)$ of Eq. (4).

Control approaches based on output feedback and frequency domain methods. Linear state-variable feedback–based methods have been used widely in a variety of applications. Alternative methodologies have also been introduced to address stability and performance robustness issues when the control specifications are expressed in the frequency domain.

An important optimal design methodology is the H_∞ optimal control approach, where the cost function is the H_∞ norm of a transfer matrix of interest, typically between some disturbance input and an output. The H_∞ norm provides a measure of the worst possible performance; to illustrate, it can be described as the distance from the origin to the furthest point on the Nyquist plot of the transfer function. In the H_∞ optimal control methodology, a parameterization of all stabilizing controllers is used, and the parameter is selected by minimizing an appropriately selected H_∞ norm, so that the controller not only stabilizes the system but also satisfies additional control objectives expressed in terms of minimizing a norm. The calculations are typically carried out using expressions that involve the constant parameters of some state-variable descriptions. Stability and performance robustness to plant parameter and disturbance uncertainties and the worst case measures in the frequency domain are the important underlying themes of this approach.

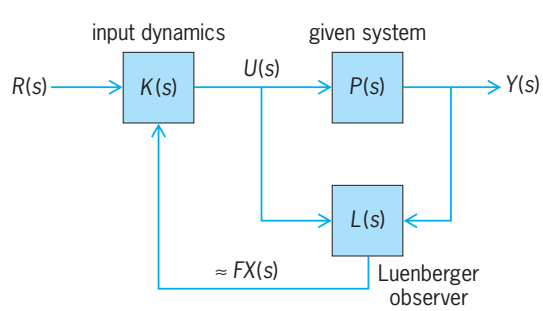

Fig. 3. Linear state variable feedback-input dynamic compensation.

Other frequency domain optimal-control methodologies do exist, for example involving the l_1 norm of appropriate maps, the advantage of the l_1 norm being that solutions may be found efficiently based on linear programming techniques. (The l_1 norm applied to a vector is the sum of the absolute values of all its elements.)

Additional approaches to control design. Optimization theory provides a powerful array of tools that may be used to solve control problems. A relatively recent trend in control design has been the recasting of control problems as linear programming problems or more generally as convex optimization problems, where efficient solution methodologies for large problems exist. Linear matrix inequalities (LMI), which may be solved efficiently by optimization algorithms such as interior point methods, have been used to solve an increasing number of control problems.

An approach that has proven to be very effective in practice, particularly in process control applications, is the model predictive control methodology, which is based on solving constrained optimal control problems on line. In this approach, given the current plant state $x(k)$ in a discrete-time plant description, an open-loop optimal control problem over some future N-step interval is solved subject to current and future constraints on the plant input, state, and output variables. This calculated optimal policy is then applied for only a short interval, one step only, and then the procedure is repeated. (This is a type of receding horizon approach.) The open control law is converted into a closed-loop strategy by using the measured value of $x(k)$ as the current state. This method is robust, deals effectively with constraints, can handle very large problems, and has been applied successfully to nonlinear problems as well. Panos J. Antsaklis; William A. Wolovich

Relative gain array. The relative gain array is an analytical device well known in process control multivariable applications, also called interaction measure and several similar terms. It is based on the comparison of single-loop control to multivariable control. It is expressed as an array (for all possible input/output pairs) of the ratios of a measure of the single-loop behavior between the input and output variable pair, to a related measure of the behavior of the same input/output pair under some idealization of multivariable control. This is appropriate because most practical multivariable control is achieved through multiple single loops. For a square transfer function matrix T the array $M = [\mu_{ij}]$ is defined by $\mu_{ij} = t_{ij}(s)/\hat{p}_{ij}(s) = t_{ij}(s)\hat{t}_{ji}(s)$, where $\hat{t}_{ij}(s)$ are the elements of $T(s) = T^{-1}(s)$, and $\hat{p}_{ij}$, equal by defination to $1/\hat{t}_{ji}$, has the interpretation of the transfer function between the paired input and output variables under perfect control of other variables. A number of alternative forms and extensions have been introduced.

The array has a very strong intuitive appeal, particularly in the steady-state form $M(0) = [\mu_{ij}(0)]$: perfect control is an idealization of real control, and the ratio naturally relates single and multiloop areas. $\mu = 1$ if the loop is isolated from the rest of the system. For

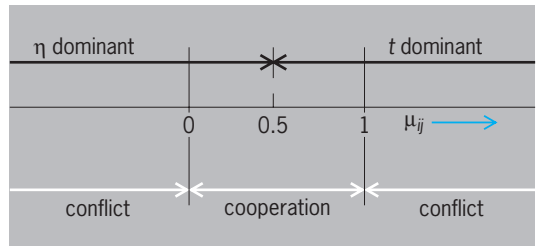

Fig. 4. Classification of ranges of possible values of an element of the relative gain array μ_{ij}.

$\mu < 0$, apparent conflict in behavior should arise. The intuition can be extended by considering $\hat{p}_{ij}$ as being made up of a direct part t_{ij} and an indirect part η_{ij}. The ranges of the values that μ_{ij} can take can be classified in terms of the relative sign and magnitude of μ_{ij} and η_{ij} (**Fig. 4**): For $0 < \mu_{ij} < 1$, $t_{ij}/\eta_{ij} > 0$; t and η cooperate. For $\mu_{ij} < 0$ or $\mu_{ij} > 1$, $t_{ij}/\eta_{ij} < 0$; t and η conflict. For $\mu_{ij} > {}^1/_2$, $|t_{ij}/\eta_{ij}| > 1$; t dominates. For $\mu_{ij} < {}^1/_2$, $|\eta_{ij}/t_{ijg}| > 1$; η dominates. Thus for $\mu < 0$, η dominates and conflicts with t. Intuitively not only is η fighting what would normally be the main path of control, but it overcomes that path. The consequences are indicated as item 2 below.

The intuitive appeal of the array can be formalized in terms of several results, the most important of which are the following:

1. $\mu_{ij}(s)$ is an invariant under scaling and any other single variable compensation or transformation.

2. For normal process and control structures, $\mu_{ij}(0) < 0$ implies that $p_{ij}(s)$ has a right half plane pole or zero, where $p_{ij}(s)$ is the transfer function between the paired input and output variables for any such choice of controls on the rest of the structure. This result can be generalized to relate to consequences in $p_{ij}(s)$ if $\mu_{ij}(s)$ has a right half plane pole or zero.

The obvious use of $\mu(0)$ is to decide on the best pairing of input/output variables for multiloop control using conventional industrial single-loop controllers (for instance, proportional integral or proportional integral derivative controllers). In this case the chosen pairing should generally have the property that the corresponding values of μ_{ij} be positive and preferably be close to 1. The value of μ can usually be related to the expected control behavior and as such can be used to justify more complex controls. Even in transfer function form the measure fails to give information relating to certain issues that are involved in the design of systems, such as detailed dynamic design and disturbance propagation. But experience almost invariably shows that a difficulty in a multivariable control design can be anticipated and diagnosed by use of the array. *See* CONTROL SYSTEMS; PROCESS CONTROL. Edgar H. Bristol

Bibliography. P. J. Antsaklis and A. N. Michel, *Linear Systems*, 2006; E. F. Camacho and C. Bordons, *Model Predictive Control*, 2d ed., 2004; J. M. Maciejowski, *Multivariable Feedback Design*, 1989; S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control: Analysis and Design*, 2d ed., 2005; K. Zhou with J. C. Doyle and K. Glover, *Robust and Optimal Control*, 1996.

Multivibrator

A form of electronic circuit that employs positive feedback to cross-couple two devices so that two distinct states are possible, for example, one device ON and the other device OFF, and in which the states of the two devices can be interchanged either by use of external pulses or by internal capacitance coupling. When the circuit is switched between states, transition times are normally very short compared to the ON and OFF periods. Hence, the output waveforms are essentially rectangular in form.

Multivibrators may be classified as bistable, monostable, or astable. A bistable multivibrator, often referred to as a flip-flop, has two possible stable states, each with one device ON and the other OFF, and the states of the two devices can be interchanged only by the application of external pulses. A monostable multivibrator, sometimes referred to as a one-shot, also has two possible states, only one of which is stable. If it is forced to the opposite state by an externally applied trigger, it will recover to the stable state in a period of time usually controlled by a resistance-capacitance (*RC*) coupling circuit. An astable multivibrator has two possible states, neither of which is stable, and switches between the two states, usually controlled by two *RC* coupling time constants. The astable circuit is one form of relaxation oscillator, which generates recurrent waveforms at a controllable rate.

Symmetrical bistable multivibrator. In bistable multivibrators, either of the two devices in a completely symmetrical circuit may remain conducting, with the other nonconducting, until the application of an external pulse. Such a multivibrator is said to have two stable states.

JFET circuits. The original form of bistable multivibrator made use of vacuum tubes and was known as the Eccles-Jordan circuit, after its inventors. It was also called a flip-flop or binary circuit because of the two alternating output voltage levels. The junction field-effect transistor (JFET) circuit (**Fig. 1**) is a solid-state version of the Eccles-Jordan circuit. Its resistance networks between positive and negative supply voltages are such that, with no current flowing to the drain of the first JFET, the voltage at the gate of the second is slightly negative, zero, or limited to, at most, a slightly positive value. The resultant current in the drain circuit of the second JFET causes a voltage drop across the drain load resistor; this drop in turn lowers the voltage at the gate of the first JFET to a sufficiently negative value to continue to reduce the drain current to zero. This condition of the first device OFF and the second ON will be maintained as long as the circuit remains undisturbed. *See* TRANSISTOR.

If a sharp negative pulse is applied to the gate of the ON transistor, its drain current decreases and its drain voltage rises. A fraction of this rise is applied to the gate of the OFF transistor, causing some drain current to flow. The resultant drop in drain voltage, transferred to the gate of the ON transistor, causes a further rise at its drain. The action is thus one of positive feedback, with nearly instantaneous transfer of conduction from one device to the other. There is one such reversal each time a pulse is applied to the gate of the ON transistor. Normally pulses are applied to both transistors simultaneously so that whichever device is ON will be turned off by the action. The capacitances between the gate of one transistor and the drain of the other play no role other than to improve the high-frequency response of the voltage divider network by compensating for the input capacitances of the transistors and thereby improving the speed of transition.

Bipolar transistor circuits. A bipolar transistor counterpart of the JFET bistable multivibrator with *npn* bipolar transistors is shown in **Fig. 2**. The base of the transistor corresponds to the gate, the emitter to the source, and the collector to the drain. Although

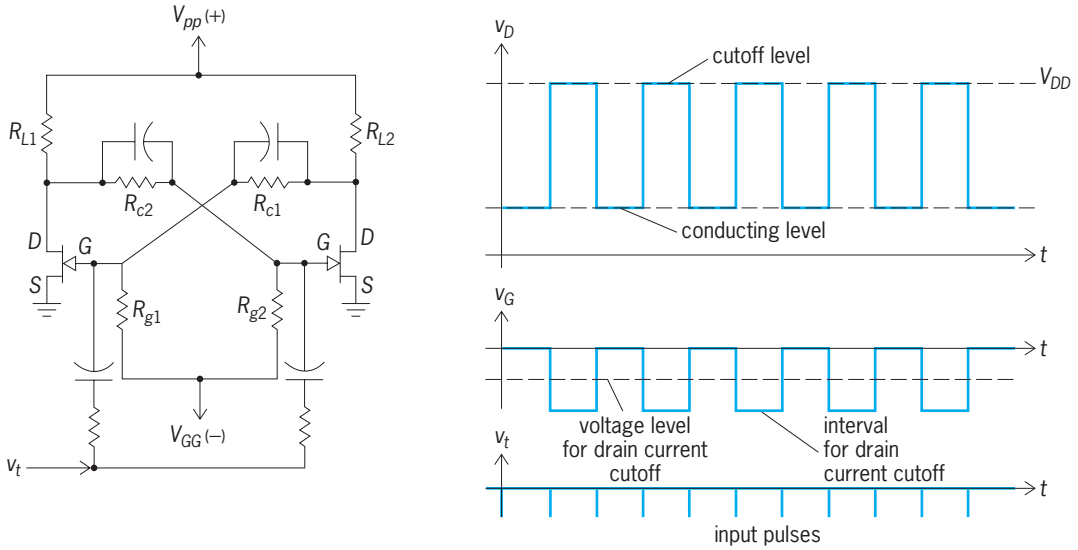

Fig. 1. Bistable multivibrator with triggering, gate, and drain waveforms shown for one transistor.

waveforms are of the same polarity and the action is roughly similar to that of the JFET circuit, there are important differences. The effective resistance of the base-emitter circuit, when it is forward-biased and being used to control collector current, is much lower than the input gate resistance of the JFET when the latter resistance is used to control drain current (a few thousand ohms compared to a few megohms). This fact must be taken into account when the divider networks are designed. If *pnp* transistors are used, all voltage polarities and current directions are reversed.

IGFET circuits. Insulated-gate field-effect transistors (IGFETs) may be used effectively in multivibrators. Use of enhancement-mode insulated-gate field-effect transistors permits direct cross-coupling with considerable simplification (**Fig. 3**). The circuit shown uses *p*-channel FETs with a negative supply voltage. Similarly, *n*-channel FETs may be used with a positive voltage supply.

Unsymmetrical bistable circuits. Bistable action can be obtained in the emitter- or source-coupled circuit with one of the set of cross-coupling elements removed (**Fig. 4**). In this case, regenerative feedback necessary for bistable action is obtained by the one remaining common coupling element, leaving one emitter or gate free for triggering action. Biases can be adjusted such that device 1 is ON, forcing device 2 to be OFF. In this case, a pulse can be applied to the free input in such a direction as to reverse the states. Alternatively, device 1 may initially be OFF with device 2 ON. Then an opposite polarity pulse is required to reverse states. Such an unsymmetrical bistable circuit, historically referred to as the Schmitt trigger circuit, finds widespread use in many applications.

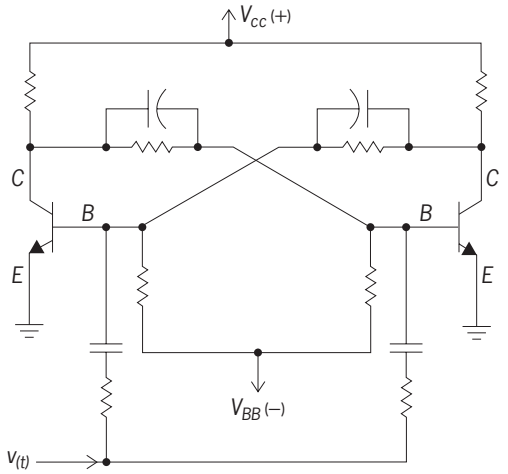

Fig. 2. Bistable multivibrator using *npn* bipolar transistors. Action is similar to vacuum-tube circuit.

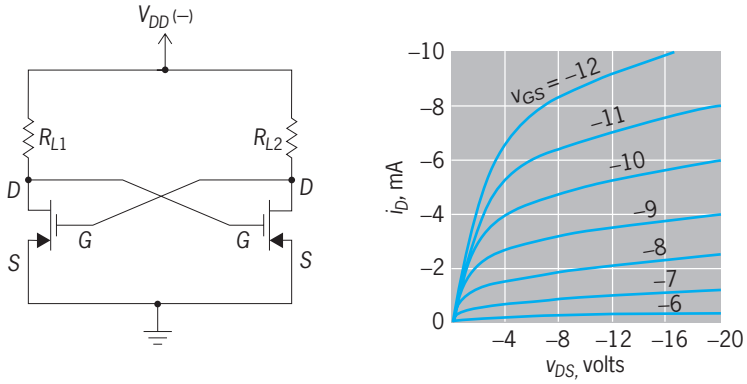

Fig. 3. Direct-coupled, bistable multivibrator which uses *p*-channel enhancement-mode, field-effect transistors, resulting in considerable simplification.

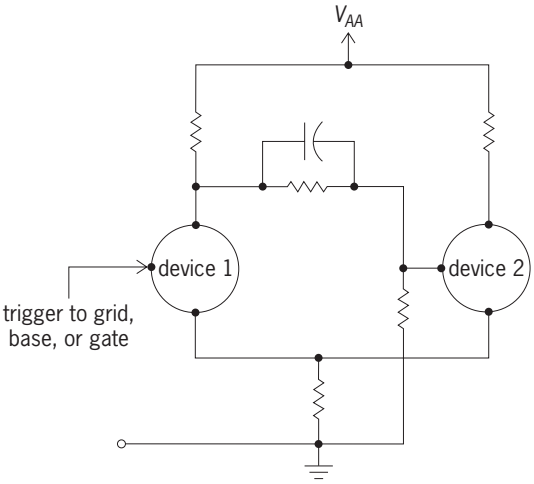

Fig. 4. Unsymmetrical bistable multivibrator.

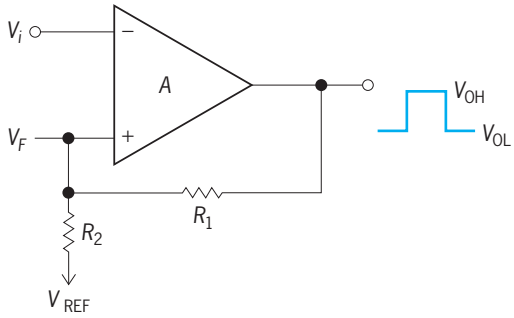

Fig. 5. Operational amplifier comparator used as Schmitt trigger circuit.

A Schmitt trigger circuit often employs a high-gain operational amplifier of the type normally used as a comparator in a positive-feedback or regenerative mode (**Fig. 5**). When $V_i > V_F$, the output is in its low state V_{OL}, with V_F in turn determined by V_{OL}, V_{REF}, R_1, and R_2. When $V_i < V_F$, the output will be in its high state V_{OH}, and V_F will be at a correspondingly higher level. Thus if the input V_i switches between two levels, the output will switch between its low and high states. These switching levels are different depending upon whether V_i is increasing or decreasing. The difference in $V_i = V_F$ for the two levels is known as the hysteresis of the circuit. When the input V_i is between the low and high V_F levels, the output can be in either its high or low state depending upon the previous turn-on history; hence the term bistable circuit. *See* COMPARATOR.

Monostable multivibrator. A monostable or one-shot multivibrator has only one stable state. If one

of the normally active devices is in the conducting state, it remains so until an external pulse is applied to make it nonconducting. The second device is thus made conducting and remains so for a duration dependent upon *RC* time constants within the circuit itself.

A typical monostable multivibrator is shown in **Fig. 6**. The input of field-effect transistor (FET) 1 is capacitance-coupled to the output of FET 2. In the absence of external pulses, FET 1 is conducting, with its gate at zero potential, or limited to a slight positive value by saturation. The resultant drain current limits the drain voltage to a value that makes the gate of FET 2 sufficiently negative to keep FET 2 cut off. If a negative pulse is applied to the gate of FET 1, this FET is cut off and FET 2 conducts. The circuit action is similar to that of the bistable circuits except that the voltage drop at the drain of FET 2 is transferred to the gate of FET 1 through the capacitance C_c. This change, or transition, cannot be maintained indefinitely because the current flowing through C_c and R_{G1} causes a decrease in voltage with time across C_c and rise in voltage at the gate of FET 1, as shown. The initial drop is of the same magnitude as that at the drain of FET 2 at the time the trigger is applied. The ensuing rise is exponential in form and (if R_{G1} is much greater than R_{L2}) is given by Eq. (1). When

$$V_{G1} = (V_{GG} - V_{DD}) \exp\left(\frac{-t}{R_{G1}C_c}\right) + V_{DD} \qquad (1)$$

the rising voltage reaches the level V_{c1}, drain current again flows in FET 1. Positive feedback quickly causes FET 1 to become fully conducting and limited by gate current saturation. The duration of the nonconducting interval for FET 1 time (T_1 in Fig. 6) is found by solving Eq. (1) for time $t = T_1$ and for $V_{G1} = V_{c1}$. The result is Eq. (2). The circuit will remain in this state until another initiating trigger is applied.

$$T_1 = -R_{G1}C_c \ln \frac{V_{c1} - V_{DD}}{V_{GG} - V_{DD}} \qquad (2)$$

Monostable multivibrators are available commercially in integrated chip form, replacing the discrete circuits described here. *See* INTEGRATED CIRCUITS.

Astable multivibrator. The astable multivibrator has capacitance coupling between both of the active devices and therefore has no permanently stable state. Each of the two devices functions in a manner similar to that of the capacitance-coupled half of the monostable multivibrator, as shown in **Fig. 7**. It will therefore generate a periodic rectangular waveform at the output with a period equal to the sum of the OFF periods of the two devices. The duration of each of the two periods is governed by an equation of the form of Eq. (2), with the appropriate values for each of the two parts of the circuit used. The transistor astable multivibrator similarly functions as the combination of two transistor monostable sections coupled together.

Astable multivibrators, although normally free-running, can be synchronized with input pulses re-

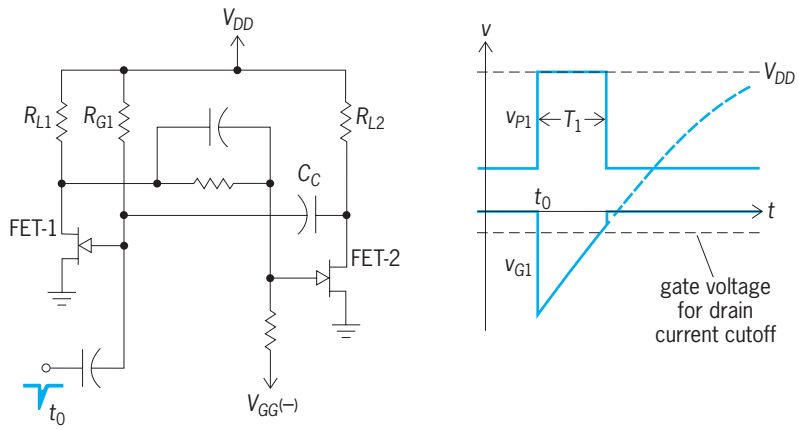

Fig. 6. Basic monostable multivibrator using junction field-effect transistors.

current at a rate slightly faster than the natural recurrence rate of the device itself. This is illustrated in **Fig. 8**, which shows the relation between the internal waveform and the applied synchronizing pulses. If the synchronizing pulses are of sufficient amplitude, they will bring the internal waveform to the conduction level at an earlier than normal time and will thereby determine the recurrence rate.

Triggering of multivibrators. The period of the bistable, the monostable, and the synchronized

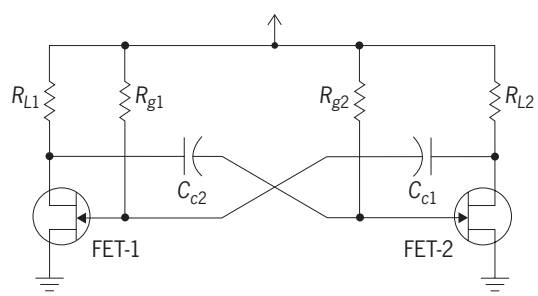

Fig. 7. Astable or free-running multivibrator.

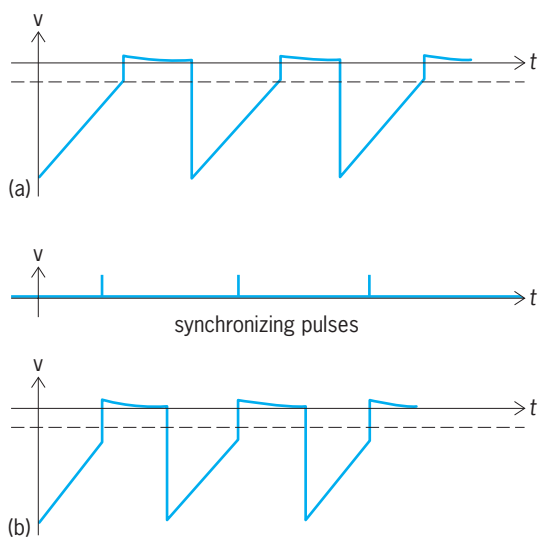

Fig. 8. Comparison of the free-running and synchronized multivibrator waveforms. (*a*) Free-running waveform. (*b*) Synchronized waveform.

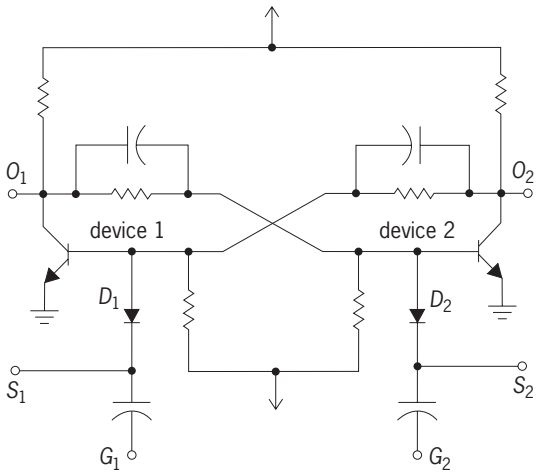

Fig. 9. Multivibrator triggering techniques.

astable multivibrator is controlled by pulses (triggers) from an external source. These triggers may be applied to the circuit in various fashions. The initiating trigger should be sufficiently wide for the circuit to respond (as limited by its high-frequency response) before the pulse is over, but not so wide as to interfere with normal action of the multivibrator once the transition has taken place. The trigger should be coupled to the multivibrator through a small capacitance so that loading by the trigger source is negligible. Usually a faster transition can be achieved if triggers turn off a normally ON device. Triggers are usually applied to the appropriate input, but they may also be applied to the outputs and reach the inputs through the coupling networks. In some cases any coupling between trigger source and multivibrator is objectionable, and an isolating amplifier is used with its drain or collector and that of the multivibrator connected together. Auxiliary diodes are frequently used to provide trigger isolation and to improve triggering stability. *See* TRIGGER CIRCUIT.

Also, particularly where bistable circuits are used in computer logic systems, it is necessary to reset initial states so that the desired device is ON before a subsequent set of triggers is applied. Such a multivibrator provides extra available terminals for triggering purposes, as illustrated in **Fig. 9**.

In modern integrated-circuit digital technology, capacitance coupling of triggering pulses is rarely used. Rather, they are applied through a set of logic circuits to effect the ON-OFF periods.

Logic gate multivibrators. Multivibrators may be formed by using two cross-coupled logic gates, with the unused input terminals used for triggering purposes. The bistable forms of such circuits are usually referred to as flip-flops with the term multivibrator rarely used. For example, the circuit of **Fig. 10**, which is like that of Fig. 2, may be thought of as the cross connection of two, two-input RTL (resistor-transistor-logic) gates. The two extra inputs are available for set-reset triggering functions. The inputs and outputs are at standard logic levels (0) and (1). The output levels after inputs at R and S are removed depend upon the input gate combinations. This is illustrated by the accompanying truth table, where Q_n and $\overline{Q}_n$ are the output states before the input gate signals are applied, and $\overline{Q}_{n+1}$ and $\overline{Q}_{n+1}$ are the outputs after signals at R and S are removed. The outputs are independent of the prior states unless R and S are both zero, in which case they remain unchanged. Also if both inputs are logic (1) the output is indeterminate, since if they both go to (0) at the same time the outputs may return to either bistable

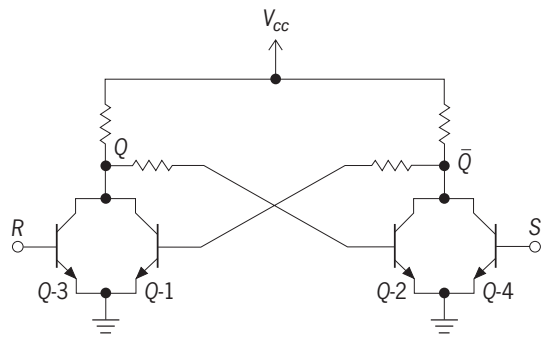

(a)

Q_n	$\overline{Q}_n$	R	S	Q_{n+1}	$\overline{Q}_{n+1}$
1	0	0	0	1	0
0	1	0	0	0	1
1	0	1	0	0	1
0	1	0	1	1	0
0	1	1	0	0	1
1	0	0	1	1	0
1	0	1	1	?	?
0	1	1	1	?	?

(b)

Fig. 10. Cross-coupled RTL gates as *R-S* flip-flop. (*a*) Circuit. (*b*) Truth table.

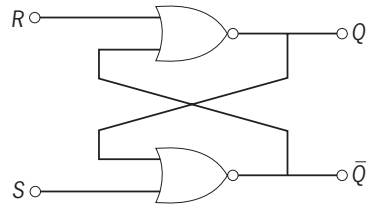

(a)

R	S	Q_{n+1}	$\overline{Q}_{n+1}$
0	0	Q_n	$\overline{Q}_n$
1	0	0	1
0	1	1	0
1	1	not used	

(b)

Fig. 11. NOR gate *R-S* flip-flop. (*a*) Symbolic representation. (*b*) Abridged truth table.

condition. This combination is generally indicated as "not used." The RTL flip-flop is representative of two cross-coupled NOR gates shown symbolically in **Fig. 11** with an abridged truth table leaving out the unnecessary initial state columns. The circuit function illustrated by these two circuits is generally referred to as an *R-S* flip-flop. Such flip-flops identified by the same truth table can be constructed by using cross-coupled NAND gates with inverting stages added at the inputs.

If the *R-S* flip-flop inputs are preceded by AND gates [**Fig. 12**, where R' or S' can be (1) only if both inputs to the AND gate are (1)] with a precisely timed "clock" pulse applied during the existence of the *R* and *S* functions, the output state of the flip-flop is set at the time that the clock pulse is applied. This is called a clocked *R-S* flip-flop and has the same truth table as the unclocked version.

If a clocked *R-S* flip-flop uses a three-input NAND gate at the inputs with a feedback path consisting of

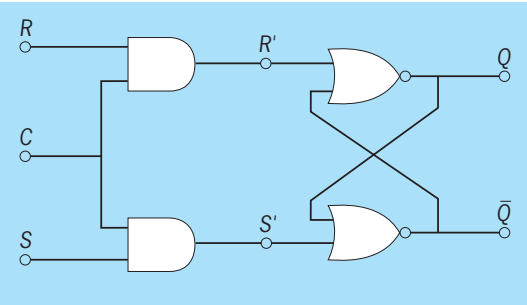

Fig. 12. Clocked *R-S* flip-flop.

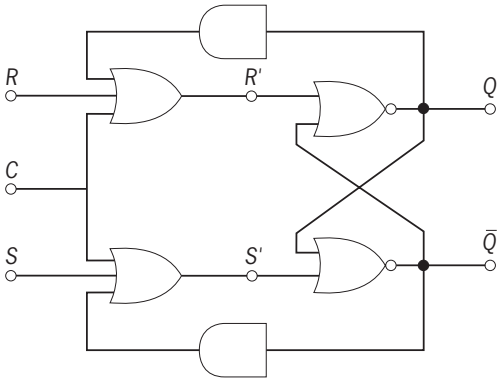

(a)

R	S	Q_{n+1}	$\bar{Q}_{n+1}$
0	0	Q_n	$\bar{Q}_n$
1	0	0	1
0	1	1	0
1	1	$\bar{Q}_n$	Q_n

(b)

Fig. 13. Pulse-triggered *R-S* flip-flop. (*a*) Circuit. (*b*) Truth table.

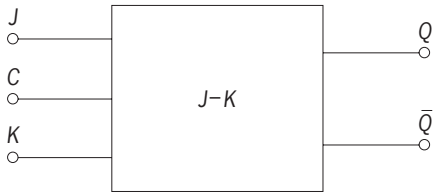

(a)

J	K	Q_{n+1}	$\bar{Q}_{n+1}$
0	0	Q_n	$\bar{Q}_n$
0	1	0	1
1	0	1	0
1	1	$\bar{Q}_n$	Q_n

(b)

Fig. 14. *J-K* flip-flop. (*a*) Symbolic representation. (*b*) Truth table with standard output designation.

a single-gate noninverting delay (**Fig. 13**), the inputs are "steered" in such a way that the not-used state is removed as indicated by the accompanying truth table. This works only if the clock pulse is very narrow so that it is removed before the feedback action is completed. Such a circuit is often referred to as a pulse-triggered *R-S* flip-flop.

Additional circuits which make the operation of the flip-flop independent of pulse width result in a circuit known as a *J-K* flip-flop, shown symbolically in **Fig. 14** with a truth table which identifies *J* and *K* inputs in a standard manner.

Sometimes problems exist with *J-K* flips-flops for slow rising pulses at inputs *J* and *K*. Additional circuits are sometimes used whereby one flip-flop called a master drives another flip-flop called a slave with feedback connections so that relative independence of clock pulse width or rise and fall times is achieved. Such a circuit is called a master-slave *J-K* flip-flop. The truth table is that of the ordinary *J-K* flip-flop.

Bistable, monostable, and astable circuits are all commercially available in integrated-circuit chip form as an alternative to the discrete logic-gate building-block approach. *See* LOGIC CIRCUITS.

Glenn M. Glasford

Bibliography. G. M. Glasford, *Analog Electronic Circuits*, 1986; G. M. Glasford, *Digital Electronic Circuits*, 1988; J. Millman and A. E. Grabel, *Microelectronics*, 2d ed., 1987; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Mumps

An acute contagious viral disease, characterized chiefly by enlargement of the parotid glands (parotitis). The virus, 175–200 nanometers in diameter, is a member of the *Paramyxovirus* genus in the

family *Paramyxoviridae*. It will grow in monkeys, newborn hamsters, embryonated eggs, and tissue cultures. *See* ANIMAL VIRUS.

Besides fever, the chief signs and symptoms are the direct mechanical effect of swelling of glands or organs where the virus localizes. One or both parotids may swell rapidly, producing severe pain when the mouth is opened. In orchitis, the testicle is inflamed but is enclosed by an inelastic membrane and cannot swell; pressure necrosis produces atrophy, and if both testicles are affected, sterility may result. The ovary may enlarge, without sequelae. Meningoencephalitis and thyroiditis are seen, but rarely. Cases without parotitis occur; 30–40% of cases are subclinical.

Diagnosis is by isolation of virus from acute saliva or cerebrospinal fluid in chick embryos or monkey kidney tissue cultures; or by enzyme immunoassay, complement-fixing responses, or neutralizing antibody responses. *See* COMPLEMENT-FIXATION TEST; EMBRYONATED EGG CULTURE; IMMUNOASSAY; TISSUE CULTURE.

Mumps is endemic throughout the world, in all seasons. Epidemics occur under crowded conditions, particularly among 5–15-year-olds, but also among military troops. Humans are the only known reservoir. In cities, a high proportion of the population is immune by the age of 15 years.

Gamma globulin prepared from mumps convalescent serum, given as soon as parotitis is observed, can decrease the incidence of orchitis. An attenuated live virus vaccine, produced by repeated passage through eggs, can induce immunity without parotitis. The vaccine is recommended for children over 1 year of age and for adolescents and adults who have not had mumps parotitis. It is contraindicated in pregnancy and in persons who are allergic to egg protein or neomycin. A single dose of the vaccine given subcutaneously induces detectable antibodies in 95% of vaccinated individuals, and antibody persists for at least 10 years.

Mumps vaccine is available in monovalent form (mumps only) or in combinations with rubella (MR) or with measles and rubella (MMR) live vaccines. Combination live virus vaccines produce antibodies to each of the viruses in about 95% of vaccinated individuals.

In 1967, the year mumps vaccine was licensed, there were about 200,000 mumps cases (and 900 patients with encephalitis) in the United States. In 1985, after 18 years of vaccine use, the number of mumps cases was less than 3000, with fewer than 20 cases of encephalitis. However, mumps cases increased somewhat during the 1986–1988 period, particularly among persons over 14 years old, who in 1987 accounted for 38% of the cases (compared to 8% in the prevaccine period). A number of the outbreaks during this period occurred among college students.

The principal strategy to control mumps in the United States is to achieve and maintain high immunization levels, primarily among infants and young children. Trivalent measles-mumps-rubella (MMR) vaccine is the preferred formulation. All persons thought to be susceptible should be vaccinated, including those without documentation of physician-diagnosed mumps, immunization with live mumps virus vaccine at 12 months of age or older, or laboratory evidence of immunity.

Ensuring immunity for adolescents and young adults is especially important in view of the shift in risk of the disease to these age groups. This trend is probably attributable to the relative underimmunization of individuals born between 1967 and 1977. The shift in risk to older persons is limited to states without comprehensive mumps immunization school laws; this provides further evidence that the relative resurgence of mumps in the United States is not due to vaccine failure but to a failure to vaccinate. *See* VACCINATION. Joseph L. Melnick

Bibliography. R. B. Belshe, *Human Virology*, 2d ed., 1990; A. S. Evans (ed.), *Viral Infections of Humans: Epidemiology and Control*, 4th ed., 1994.

Muonium

An exotic atom, mu or (μ^+e^-), formed when a positively charged muon (μ^+) and an electron are bound by their mutual electrical attraction. It is a light, unstable isotope of hydrogen, with a muon replacing the proton. Muonium has a mass 0.11 times that of a hydrogen atom due to the lighter mass of the muon, and a mean lifetime of 2.2 microseconds, determined by the spontaneous decay of the muon ($\mu^+ \rightarrow e^+\nu_e\overline{\nu}_\mu$).

Muonium is formed when beams of μ^+ produced in particle accelerators are stopped in certain nonmetallic targets. The μ^+ beams are generally spin-polarized; that is, the average spin angular momentum of the muons in the beam points in a definite spatial direction. The muon spin retains this spatial orientation after picking up an electron to form muonium in the reaction $\mu^+ + \mathrm{X} \rightarrow \mathrm{Mu} + \mathrm{X}^+$, where X represents an atom of the target material. Since the positron in μ^+ decay is emitted preferentially along the muon spin direction, the muonium polarization can be monitored very simply by measuring the spatial distribution of the decay positron. This technique of polarization measurement allows extremely small amounts of muonium to be detected and studied. *See* SPIN (QUANTUM MECHANICS).

Muonium was first observed by V. W. Hughes and coworkers in 1960. The characteristic Larmor precession of the muonium polarization in a magnetic field when positive muons were stopped in a target of argon gas indicated the formation of muonium. The observed Larmor frequency of 14 GHz/tesla is a unique signature of the triplet ($F = 1$) bound state of an electron and a muon. *See* LARMOR PRECESSION.

Since muonium is a system consisting only of leptons, it serves as an ideal testing ground for the theory of quantum electrodynamics (QED), which describes the electromagnetic interaction between particles. Indeed, experimental measurement of the hyperfine structure interval of the ground-state

muonium levels have shown that the quantum electrodynamic theory of this system is accurate to the level of one part per million. Such measurements also provide the best available values for the muon mass and magnetic moment. *See* LEPTON; QUANTUM ELECTRODYNAMICS.

Muonium chemistry and muonium spin rotation (MSR) are subfields which concentrate on study of muonium in matter. Chemical reaction rates for a wide variety of muonium reactions have been measured and compared with hydrogen rates. Formation, spin precession, and depolarization of muonium in gases, semiconductors, and insulators have been studied in some detail. Such experiments seek to understand the chemical and physical behavior of a light hydrogen isotope in matter, and to use the extremely sensitive muon polarization measurement technique to probe the structure of materials. *See* POSITRONIUM. Patrick O. Egan

Bibliography. H. J. Ache (ed.), *Positronium and Muonium Chemistry*, 1979; V. Hughes and C. S. Wu (eds.), *Muon Physics*, 3 vols., 1979; A. Schenck, *Muon Spin Rotation Spectroscope*, 1985.

Muscle

The tissue in the body in which cellular contractility has become most apparent. Almost all forms of protoplasm exhibit some degree of contractility, but in muscle fibers specialization has led to the preeminence of this property.

Anatomy

In vertebrates three major types of muscle are recognized: smooth, cardiac, and skeletal.

Smooth muscle. Smooth muscle, also designated visceral and sometimes involuntary, is the simplest type. These muscles consist of elongated fusiform cells which contain a central oval nucleus. The length of such fibers varies greatly, from a few micrometers up to 0.02 in. (0.5 mm). These fibers contract relatively slowly and have the ability to maintain contraction for a long time. Smooth muscle forms the major contractile elements of the viscera, especially those of the respiratory and digestive tracts, and the blood vessels. Smooth muscle fibers in the skin regulate heat loss from the body. Those in the walls of various ducts and tubes in the body act to move the contents to their destinations, as in the biliary system, ureters, and reproductive tubes.

Smooth muscle is usually arranged in sheets or layers, commonly oriented in different directions. The major physiological properties of these muscles are their intrinsic ability to contract spontaneously and their dual regulation by the autonomic nerves of the sympathetic and parasympathetic systems.

Cardiac muscle. Cardiac muscle has many properties in common with smooth muscle; for example, it is innervated by the autonomic system and retains the ability to contract spontaneously. Presumably, cardiac muscle evolved as a specialized type from the general smooth muscle of the circulatory vessels. Its rhythmic contraction begins early in embryonic development and continues until death. Variations in the rate of contraction are induced by autonomic regulation and by many other local and systemic factors. *See* AUTONOMIC NERVOUS SYSTEM.

The cardiac fiber, like smooth muscle, has a central nucleus, but the cell is elongated and not symmetrical. It is a syncytium, a multinuclear cell or a multicellular structure without cell walls. Histologically, cardiac muscle has cross-striations very similar to those of skeletal muscle, and dense transverse bands, the intercalated disks, which occur at short intervals.

The heart contains its own specialized system for initiation and spread of contraction in a wavelike form over the myocardium. This conducting system, which is composed of the sinoauricular and atrioventricular nodes and intervening bundles of special tissue, transmits the primary impulses. It is modified cardiac muscle, sometimes called the neuromuscular system to indicate its dual characteristics. *See* HEART (VERTEBRATE).

Skeletal muscle. Skeletal muscle is also called striated, somatic, and voluntary muscle, depending on whether the description is based on the appearance, the location, or the innervation. The individual cells or fibers are distinct from one another and vary greatly in length from over 6 in. (15 cm) to less than 0.04 in. (1 mm). These fibers do not ordinarily branch, and they are surrounded by a complex membrane, the sarcolemma. Within each fiber are many nuclei; thus it is actually a syncytium formed by the fusion of many precursor cells.

The transverse striations of skeletal muscle form a characteristic pattern of light and dark bands within which are narrower bands. These bands are dependent upon the arrangement of the two sets of sliding filaments and the connections between them.

A number of different types of vertebrate skeletal fibers are known, including twitch and tonus, red and white, fast twitch and slow twitch, large- and small-diametered, and so forth. Details of the correlations between these and other properties of fiber types are far from satisfactorily known.

Morphology. When organized into muscles, skeletal muscle fibers are arranged in an orderly fashion with the axis of the fibers orienting roughly between the two points of muscle attachment. Surrounding the individual muscle fibers is a connective tissue layer, the endomysium. Groups of fibers form bundles, or fasciculi, which are ensheathed with collagenous fibered connective tissue, the perimysium. This perimysium is continuous throughout the muscles and with the epimysium surrounding the entire muscle. These sheaths of connective tissue grade rather abruptly into the dense collagenous structure of tendons at the ends of the muscle.

The size of a muscle—the number and length of fibers—is directly related to the function of that muscle. For a muscle of a fixed volume, the number of fibers is inversely proportional to the length of the fibers. Hence, as the number of fibers increases with the force development of the muscle, their length

decreases with a reduction in the ability of the muscle to shorten. Conversely, as the fibers increase in length, allowing the muscle to shorten over a greater distance, its force decreases with the reduction in the number of fibers. Pinnate muscles have the fibers arranged obliquely to the central tendon of the muscle, permitting a large number of shorter fibers to be packed into a conveniently shaped muscle. These muscles can develop great force but can shorten through a very restricted distance.

Skeletal muscles tend to be distributed about a movable structure (bone) so as to produce antagonistic actions, allowing the force of each muscle to relengthen the opposing muscle. In most parts of the body, movement of the bones is achieved through the synergistic contraction of groups of muscles rather than by the action of individual muscles.

The innervation of skeletal muscles by motor fibers of the peripheral nervous system presents a quite constant pattern. This pattern appears early in embryonic life when the motor nerves grow into the muscle tissue of each body segment. During the ontogeny of limbs and specialized features of the trunk, neck, and head, a great deal of rearrangement occurs, but the early relationships are maintained so long as the muscle persists as a postembryonic structure. *See* MUSCULAR SYSTEM.

Walter Bock

Biophysics

Muscle is a biological system for generating mechanical force and requires the expenditure of free energy. A solution of adenosine triphosphate (ATP) has substantially more free energy than its breakdown products adenosine diphosphate (ADP) and inorganic phosphate (P_i). The loss in free energy when ATP and water are converted into ADP and P_i provides the required expenditure of energy for muscle. Thus ATP and water may be considered the fuel for operating the muscle machine. *See* ADENOSINE TRIPHOSPHATE (ATP).

This energy expenditure may or may not permit the muscle to do useful work. If the force generated is applied to an immovable object, such as a weight too heavy to lift, the result is merely the generation of static tension in the muscle. The muscle is then said to be contracting isometrically—it is prevented by the heavy weight from changing length. If the force generated exceeds the weight, then the muscle, by shortening, will move the weight. Such a muscle is contracting isotonically. An animal normally employs both types of contraction; isometric contraction occurs in applying pressure or in holding the skeleton against gravitational forces, while isotonic contractions result in moving the skeleton.

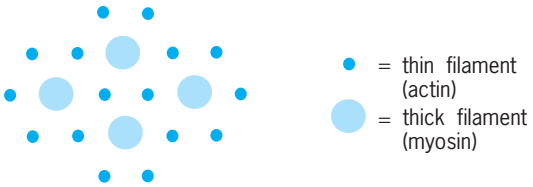

Fig. 2. Cross section of a muscle fibril.

All three types of muscle operate with the same machinery. The following discussion focuses on research on the contractile apparatus of vertebrate skeletal (striated) muscle.

Microstructure. A striated muscle is a cylindrical bundle of fibers; each fiber is, in turn, a cylindrical bundle of fibrils (myofibrils). The myofibrils are segmented into sarcomeres, each a diminutive cylinder about 1 μm in diameter and 2.5 μm long. The sarcomeres of adjacent myofibrils are in register, thus giving the entire fiber a banded (or striated) appearance.

At either end the sarcomeres are bounded by end plates called Z membranes. The central sectors of sarcomeres are highly birefringent and are known as A (for anisotropic) bands, while the sections to either side of the Z members are practically nonbirefringent and are known as I (for isotropic) bands. Both the A and I bands result from the existence of two independent arrays of filaments which interpenetrate extensively in the A band (**Fig. 1**): these filament arrays are hexagonal in cross section (**Fig. 2**). Thin filaments are attached to each Z membrane and extend inward into the sarcomere but are not continuous across it. Thick filaments exist in the central sector. When the fibril is at rest length, the thick and thin filaments interpenetrate, but the degree to which they do so varies with the length of the fibril. The relatively clear central zone separating the two edges of the thin filament arrays is called the H zone (Fig. 1). The Z membrane, which is common to two sarcomeres, is composed of a special weave. From an end-on view it is a cubic lattice, but in longitudinal section it shows thin filaments in a hexagonal array.

Mechanistic model. It is possible to discuss the operation of muscle in mechanistic terms. In this sense a muscle cell has three recognizable machines: the sarcoplasmic reticulum, the calcium sensor, and the contractile apparatus.

Sarcoplasmic reticulum. The sarcolemma (cell membrane) of a single muscle fiber is perforated in a regular manner by tubular invaginations (sarcotubules). In contact with these tubules are elongated flat sacs

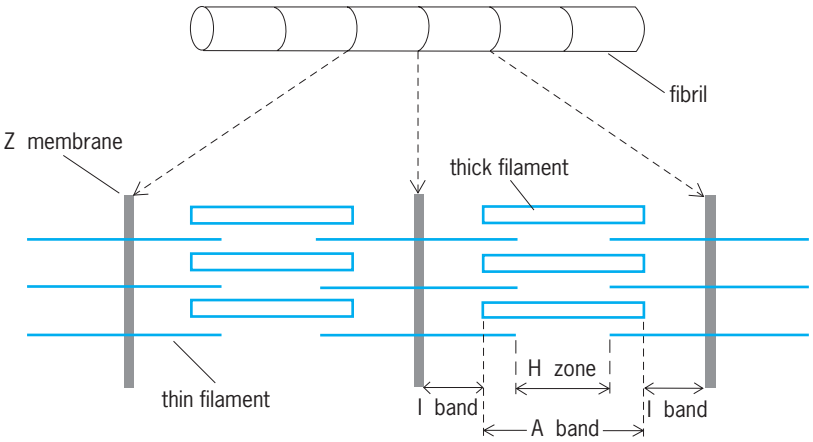

Fig. 1. Schematic representation of two adjacent sarcomeres showing banding patterns as revealed by electron microscopy.

or cisternae which surround subbundles of fibrils like a blanket. Collectively these cisternae are known as the sarcoplasmic reticulum (**Fig. 3**). Both the sarcolemma and the walls of the cisternae are bilayered. There is no direct opening between the tubules and the cisternae; however, at the points of contact there are four distinct layers (Fig. 3*b*). Communication between tubules and cisternae is therefore not by passage of a transmitter substance but perhaps by a field effect. Like bilayered membranes of other cells, the membranous sarcoplasmic reticulum contains essential phospholipid and structural protein, as well as certain enzyme proteins. *See* CELL MEMBRANES.

In response to nerve signals the sarcoplasmic reticulum can explosively release calcium ions into the contractile apparatus. Contractile impulses are initiated at the motor end plates, which are the points of contact of the nerve endings with the muscle cell. The nerve endings release the chemical acetylcholine, and the muscle cell then initiates its own stimulus which is conducted along the sarcolemma to the tubules. The tubules extend inward to surround each myofibril at the Z membrane. From here the electrical signal is somehow transferred to the adjacent anastomosing sheath of cisternae surrounding each myofibril. Sodium and potassium ions are intimately involved in the transmission of the impulses or stimuli. As the stimulus wave travels along the sarcolemma, sodium ions move into the membrane and potassium ions move outward. This shift of cations is not matched by a corresponding shift of anions, and as a result of the change of electrical charge on either side of the membrane an electrical potential change is created which propagates the stimulus wave. As a result of the nerve impulse a transition occurs in the cisternae membranes, causing them to become more permeable and thereby creating a very leaky pump. Because of the ion gradient that exists, calcium ions diffuse out of the cisternae very rapidly. This explains the rapid contractile response to a nerve impulse, since it is the calcium ion (Ca^{2+}) concentration which controls the contractile mechanism at the molecular level. The sarcoplasmic reticulum stores large amounts of Ca^{2+} in the lumen of the cisternae. *See* ACETYLCHOLINE; BIOPOTENTIALS AND IONIC CURRENTS.

After contraction, the electrical effect stops, the impermeability of the cisternae walls is restored, and calcium ion is pumped back into the cisternae, which brings about relaxation. Since calcium must move counter to the electrical gradient, this phase therefore requires work, is associated with ATP hydrolysis, and is another example of active transport.

Calcium sensor. This device is attached to the contractile apparatus and "reads" the ambient Ca^{2+} concentration. When adequate Ca^{2+} is present, it removes blocks and permits the apparatus to work. In the absence of Ca^{2+} it restores these blocks and prevents the apparatus from contracting. This sensor device is described in more detail below.

Contractile apparatus. This machine is constructed from two proteins and runs on energy provided by magnesium ions (Mg^{2+}) and ATP. Muscle fibers shorten or lengthen because their sarcomeres shorten or lengthen, that is, because the filaments within the sarcomeres translate (slide) relative to one another. The self-propelled, active shortening of fibers is always accompanied by the hydrolysis reaction, $ATP + H_2O \rightarrow ADP + P_i$. Fibers lengthen or get stretched only when external forces are applied to them. An emerging generalization seems to be that all biological movement (such as all muscular, ameboid, and ciliary movement) is associated with the translation of filaments and with the hydrolysis of a triphosphate. However, the array arrangement, the proteins constituting the filaments, and the particular phosphate may differ from system to system.

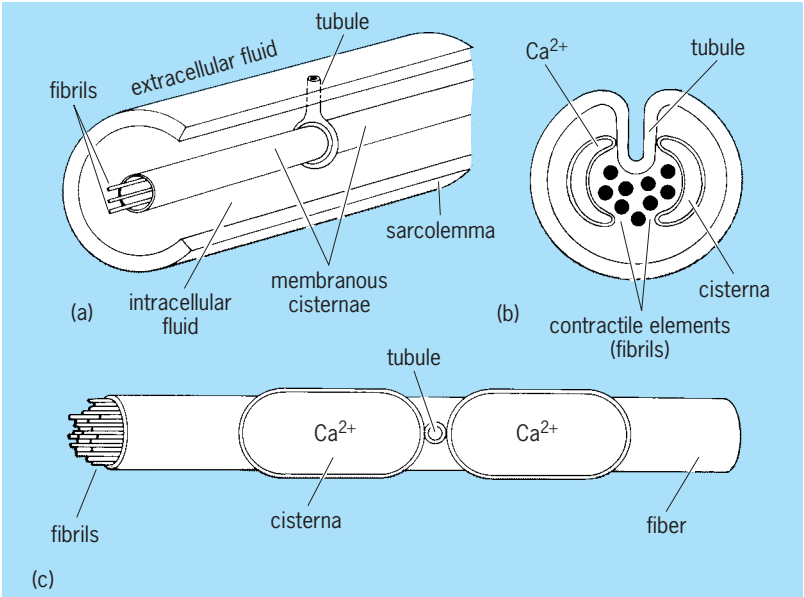

Fig. 3. Schematic representations of the sarcoplasmic and sarcotubular systems. (*a*) Cutaway of a muscle fiber. (*b*) Cross section of a fiber. (*c*) Fiber viewed from above (sarcolemma removed) showing triad formation of tubule and cisternae.

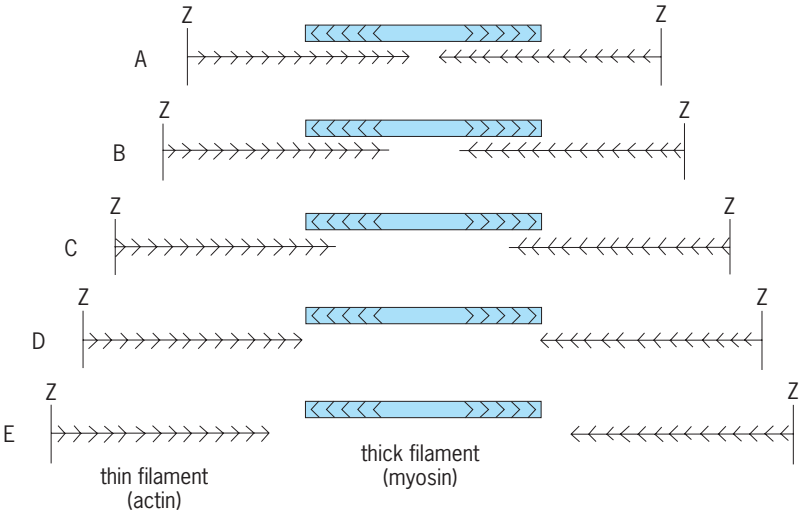

Fig. 4. A sarcomere changing length (Z-Z distance). The thick and thin filaments, though remaining at constant length, translate relative to one another. The chevrons indicate the polarity and location of the molecules constituting the filaments. The number of myosin molecules next to actin differs at different extensions (A, B, C, D, E). Not shown in this schema is the fact that sarcomeres contract at constant volume, so that shortening is accompanied by separation of the filaments in the transverse directions.

The thin and thick filaments are principally composed of the proteins actin and myosin, respectively (**Fig. 4**). Both filaments are polymeric and polar, and each is a serial arrangement of "building blocks" or monomers. One "end" of a monomer is different from the other, and in the serial arrangement of a filament the monomers are assembled "head to tail," so that the filament itself has a direction. The actin filaments can be thought to originate at the Z disks of the sarcomere and to point toward a central midplate. Each thick filament can be thought of as consisting of two coaxial myosin filaments; both originate at the midplane, and each points to one Z disk. The midplate is thus a plane of symmetry for the sarcomere. The directions assigned to the filaments can be taken as the directions of relative motion when a sarcomere shortens.

In both actin and myosin filaments the serial arrangement of monomers is helical. In cross section these actin and myosin helices form a well-ordered lattice. The net result of this geometry is that parallel to the fiber axis some actins of a particular actin filament are closest to some myosins of an adjacent myosin filament (the remainder of the actins are closest to other myosin filaments, and the remainder of the myosins are closest to other actin filaments). The ratio of actins to myosins need not be unity (it is in fact 5) because the pitches of the two types of helices need not be the same. *See* MUSCLE PROTEINS.

Actin monomer. This monomer is a dumbbell-shaped, single-chain globular protein of about 42,000 daltons and a maximum dimension of about 5 nanometers. Actin in monomeric form is known as G-actin, while in the two-stranded helical polymer form it is known as F-actin. Under physiological conditions F-actin is insoluble. In F-actin each monomer is held to the monomer before and after it on the same strand, as well as to two monomers of the other strand. These four interactin bonds are not covalent, but are quite strong. Nevertheless, the F-actin filament is rather flexible, implying that the monomers themselves may be deformable under physiological conditions. Along the filament axis the F-actin double helix repeats every 13 monomers, or every 36 nm.

Myosin monomer. This monomer is a very long (about 150 nm) and very heavy (about 480,000 daltons) multichain molecule resembling a Y with a knob or head at the end of each arm (**Fig. 5**). The main structure of the Y consists of two heavy chains whose C-terminal (carboxyl end) regions are intertwined to form the stem, and whose N-terminal (amino end) regions unravel to form the arms and heads. Attached to each head is a set of two light chains, one of which (the regulatory light chain) contains a high-affinity Ca^{2+}-binding site and a reactive serine that can be phosphorylated. The amino acid composition of the light chains is often different among different muscle cells (for example, fast skeletal muscle, smooth muscle, and embryonic muscle). This circumstance gives rise to many myosin isozymes.

By selective proteolysis it is possible to cut away and isolate just the heads of myosin. Such a head is known as a myosin subfragment 1 or S-1, while the remaining structure is known as a myosin rod. Proteolysis can again be used to cut this rod and generate two more segments; an N-terminal myosin subfragment 2 or S-2, and a C-terminal light meromyosin (LMM). When light meromyosin is cut away from the myosin molecule, the two-headed structure remaining is known as heavy meromyosin (HMM). The places at which the proteases cut are regions of flexibility in the original structure. Specifically, the heads can rotate about the arms, and there is a region of flexibility about 40 nm from the head-rod junction. Light meromyosin and any structure containing light meromyosin are insoluble in intracellular fluid, but all other segments are readily soluble. Myosin molecules are assembled into a thick filament so that there is a central core of strongly

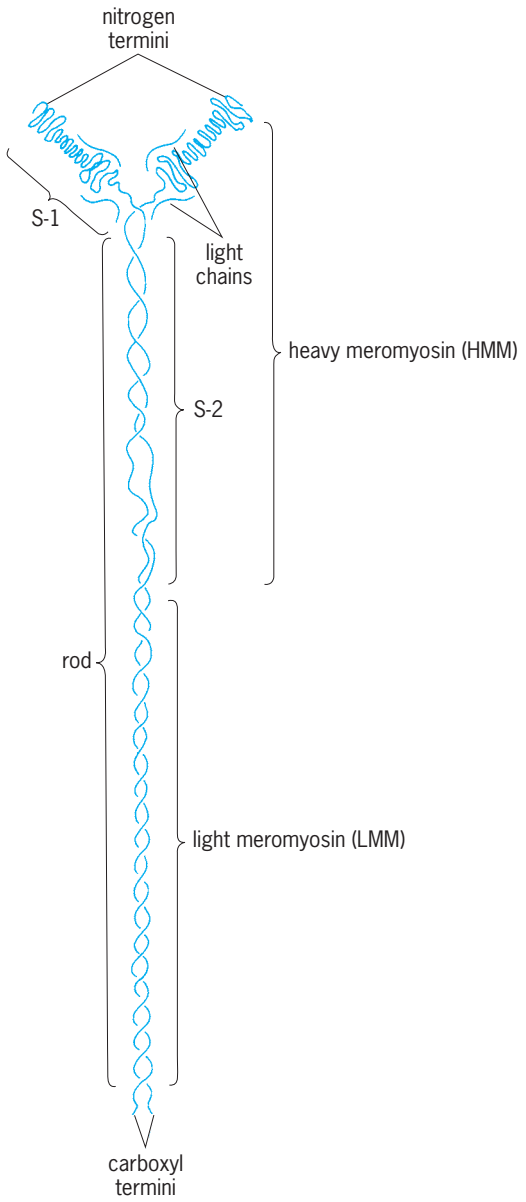

Fig. 5. Diagram of myosin, which is about 15 nm long. The various segments are obtainable after suitable proteolysis; light chains separate out in denaturing solvents. It is currently speculated that S-1 itself contains three domains.

interacting, insoluble light meromyosin segments with radially protruding soluble heavy meromyosin portions. Thus the central core is devoid of radiating heavy meromyosin and the S-1 moieties are situated adjacent to the actin filaments (**Fig. 6**). In the electron microscope the heavy meromyosin moieties appear as projections issuing from the thick filaments which occasionally touch adjacent actin filaments. These projections have been called cross-bridges. A cross-bridge is mainly an S-1 moiety, perhaps with some contribution from the S-2.

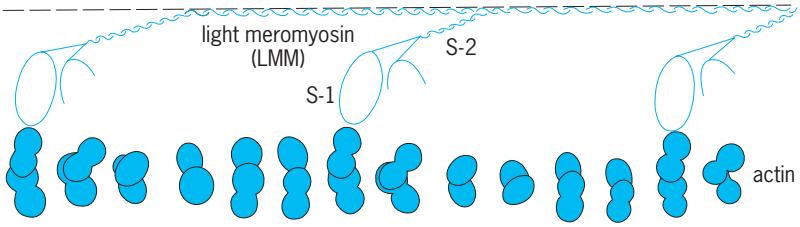

Fig. 6. Diagram of the molecules in a plane of nearest actin-myosin couples. The second S-1 of each myosin head is shown in limited fashion; each would be nearest to a separate actin filament. There is approximately one myosin head to every five actins.

Fueling reaction. It is widely accepted that the hydrolysis of ATP (or its analogs) is the fueling reaction for a great variety of biological machines, including muscle. There are, however, four aspects of this reaction that are especially important in the present context. First, the participants, in this reaction ATP, ADP, and P_i, are highly ionic. Each interacts strongly with both H^+ and Mg^{2+}. This means that the free energy obtainable from the reaction depends on these ionic concentrations. Second, the reaction does not proceed at an appreciable rate in the absence of catalysis, or even in the presence of the catalyst myosin. It proceeds rapidly only when both actin and myosin are present in the catalytic form, actomyosin. These circumstances ensure that fuel is degraded only when work is being performed. Third, enzymatic catalysis always proceeds in steps, so the free energy is dissipated in steps. The algebraic sum of these stepwise decrements must equal the total free energy of the ATP hydrolysis, but the pattern of the steps is characteristic of the particular enzyme. Fourth, the steady-state concentration of MgATP in muscle is around 3 millimolar. The MgATP would be quickly exhausted in heavy exertion were it not that the breakdown products, ADP and P_i, are swiftly synthesized back into ATP by the creatine kinase system (at the expense of creatine phosphate), and more slowly but more extensively synthesized by the glycolytic system (at the expense of stored glycogen) and by mitochondria (at the expense of oxygen and foodstuffs).

Physiology. Under experimental conditions (when the membrane of a muscle fiber has been removed), the chemical composition of the fluid bathing the contractile apparatus can be controlled. It is then possible to correlate unambiguously the composition and the physiological state of a muscle fiber. The results are as follows: (1) When the medium is devoid of ATP, as in exhaustion or death, the fiber is stiff and inextensible, and is said to be in rigor. This physiological state arises because in the absence of ATP the erstwhile high affinity of the S-1 moieties for the actins literally bonds all the cross-bridges to adjacent actins. (2) When the medium contains Mg^{2+} and ATP but no Ca^{2+}, the fiber is flaccid and unresistingly extensible; it is then said to be relaxed. This state comes about because the calcium sensor is preventing any interaction between the cross-bridges and the adjacent actins. (3) When the muscle medium contains Mg^{2+}, ATP, and Ca^{2+}, the fiber is active. If it is prevented from shortening (isometric), it is developing tension and consuming ATP at a low rate, and if it is not restrained (isotonic), it is shortening while at the same time consuming ATP at a rate dependent on the speed of shortening. It is thought that under these circumstances, force interactions tending to produce the filament interdigitation are going on between the cross-bridges and the adjacent actins.

Calcium sensor. The calcium sensor serves as a regulator of muscle action. Along the two grooves of the actin helix in every thin filament, firmly attached to actin, are deployed two long chains of a protein called tropomyosin. Periodically along each tropomyosin strand are clumps of globular proteins called troponins (T, I, and C). There is reason to think that such clumps are attached to both tropomyosin and actin. Troponin C contains several Ca^{2+} binding sites, two of which have a very strong affinity for the Ca^{2+} and strongly prefer it to Mg^{2+}. These sites fit the sensor role.

Experimentation with certain of the Ca^{2+}-binding proteins led to what seemed a well-documented theory of Ca^{2+} regulation in vertebrate skeletal muscle. In relaxation, along its helical course, tropomyosin is bound to actin monomers so that it masks all their S-1 binding sites, thus blocking all contractile activity. As the Ca^{2+} concentration in the bulk interfibrial solution rises (this Ca^{2+} concentration is itself regulated by the sarcoplasmic reticulum), it binds to the troponins which then tilt so as to pull the masks off the actins, thus permitting contraction. There is reason to think that, while elements of this theory remain correct for vertebrate skeletal muscle, other tissues may have other operative components in the Ca^{2+} sensing system. It has been found, for example, that in mollusks, Ca^{2+} sensing is performed solely by the light chains, tropomyosin/troponin being unnecessary. There is mounting evidence that something of the same myosin-linked mechanism operates in vertebrate smooth muscle. Even in vertebrate skeletal muscle the discovery that the light chains must be intact for regulation to occur suggests that these elements work in parallel with tropomyosin/troponin. Evidence has also appeared that not only physical separation of the myosin and actin but also some other means of myosin ATPase inhibition bring about relaxation. Finally, whether and how the Ca^{2+}-binding site of actin has a role in the actin-activation of ATPase remains to be discovered. Perhaps if it does, it too will have a role in regulation. At the present time the field of Ca^{2+} regulation has a surfeit of proposed mechanisms.

Length–tension relationship. The explanation of the active state discussed above arises from a classical physiological experiment known as the length–tension relationship. When the sarcomere is at rest length, every cross-bridge in a plane of nearest actins and myosins has access to actin. This continues to be so for small extensions of the fiber and its sarcomeres. But with increasing stretch a length is reached at which this ceases to be so. For stretches beyond this critical length the number of cross-bridges with adjacent actin decreases in proportion to stretch. At high extensions a second critical length is reached beyond which no cross-bridges have adjacent actins. In the experiment, the fiber, in the relaxed state, is drawn out to the desired length and clamped, then ambient condition (3) above is imposed to activate it (in the case of a membranated fiber, electrical excitation accomplishes the same condition). Two quantities are measured at each fixed length: the static tension developed against the length clamp and the ATPase activity. These two quantities, along with the number of cross-bridges opposite to actins, are then plotted against extension (**Fig. 7**). The graphs are surprisingly similar: flat at first, then falling linearly, finally flat at zero. These results strongly suggest that the ATPase-driven force generators are the cross-bridge–actin couples.

The previous experiment shows that the total force generated by the fiber is simply the sum of the unitary forces generated by a fixed number of cross-bridge–actin couples, but because of clamping, no shortening was permitted. When shortening occurs, the number of couples that can interact can still be kept approximately constant by limiting the experiment to small (around 10%) contractions, but even so a new phenomenon enters, since time changes the relative position of a particular cross-bridge and a particular actin. In this investigation the fiber is allowed

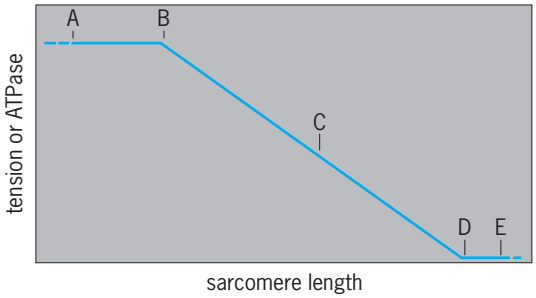

Fig. 7. Sketch of how either tension- or actin-activated ATPase depends on sarcomere length. Letters correspond to sarcomere length in Fig. 4. The correspondence is the basis for thinking that the force-generating unit is a myosin S-1 and its associated actin.

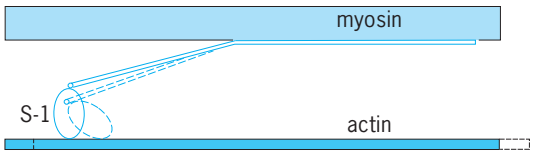

Fig. 8. Diagram suggesting that shortening would result if actin and an S-1 moiety could bind at two different angles.

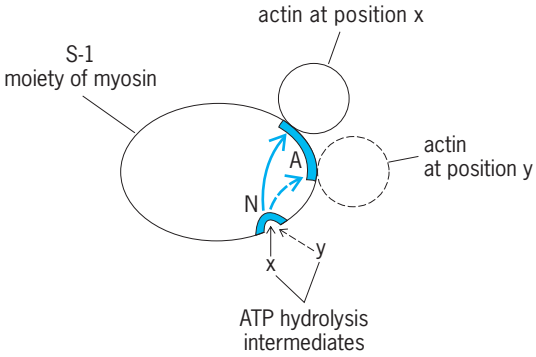

Fig. 9. The hypothesized mechanism of the transformation of chemical energy into mechanical energy in muscle. N is the enzymatic site that catalyzes the hydrolysis of ATP; A is the site at which actin binds; x and y are successive intermediates in ATP hydrolysis. When x is occupying N, the actin binds at position x on the A site; when y is occupying N, the A site binds actin at position y. During catalysis, when x is chemically converted into y, actin is caused to move from one binding position to another. This movement is forced by the chemical conversion, so the product of the distance moved and the force constitutes mechanical work.

to shorten against a constant external force (f) that it can overcome, such as a light weight (force = mass × gravitational acceleration), and the velocity of shortening (υ) is measured. The empirical result is the characteristic equation of isotonic contraction: $(f + a)\upsilon = b(f_0 - f)$, where f = force (load); υ = velocity; f_0 = maximum load; and a,b = constants. The arguments are now more complicated, but this behavior too can be rationalized by assumptions as to how the force interaction between a couple depends on time and distance.

Energy transformation. The structural-functional interpretation of the length–tension diagram (above) suggests that a myosin S-1 piece (attached to the thick filament) reacting with an actin monomer (attached to the thin filament) is the unitary engine causing the interfilament slide (**Fig. 8**). The S-1 piece binds actin through its A site and catalyzes ATPase at its N site (**Fig. 9**). The intermediates in the breakdown of ATP appear sequentially on the N site: first x, then y, then z, and so on. When one ATP is totally degraded, then a fresh ATP initiates the next sequence (x, y, z, and so on). Enzymatic activity is cyclic. Because of the way that S-1 is built, it seems that to each intermediate there corresponds a specific way of holding actin at the A site, namely, in position x, in position y, and so on. Thus, corresponding to endlessly repeated ATPase cycles there are endlessly repeated cycles in the manner of holding actin, that is, mechanical cycles. Each mechanical cycle constitutes a force impulse. When the enzymatic and mechanical activity of all the S-1 pieces acting in concert is summed, a steady force is maintained at the cost of steady ATP hydrolysis. Ultimately, therefore, the energy transduction in muscle is carried out by the molecular structures of the S-1 pieces of myosin. Patricia Rainford

Bibliography. C. Emerson et al., *Molecular Biology of Muscle Development*, 1986; A. Engle and C. Franzini-Armstrong, *Myology*, 2 vols., 2d ed., 1994; G. H. Pollack, *Muscles and Molecules: Uncovering*

the Principles of Biological Motion, 1990; D. J. Schneck, *The Mechanics of Muscle*, 1992; R. J. Stone and J. A. Stone, *Atlas of the Skeletal Muscles*, 3d ed., 1999.

Muscle proteins

Specialized proteins in muscle cells that are the building blocks of the structures constituting the moving and regulatory machinery of muscle. The moving machinery comprises myofilaments that are discernible by electron microscopy. These myofilaments are of two kinds, myosin and actin, and their regular arrangement within the cell gives the striated pattern to skeletal muscle fibers (**Fig. 1**). It is recognized that the sliding of the two sets of filaments relative to each other is the molecular basis of muscle contraction. To understand the ultimate mechanism that causes the movement of these filaments relative to each other, it is necessary to consider the features of the individual molecules that make up these filaments. Practically all nonmuscle cells, although lacking the filaments of muscle, contain proteins similar to those found in muscle; these proteins are likely to be involved in cell motility and in determining properties of cell membranes. *See* CELL MEMBRANES.

The banding pattern of muscle fibers produced by the regular arrangement of myosin and actin is further enhanced by the Z membranes, or disks, which appear as dark bands when the myofibril is viewed longitudinally and which separate adjacent sarcomeres. There is now considerable evidence for elastic connections occurring between the Z disks and interacting with the myosin filaments. The main constituent of these filaments is titin, a protein with a mass of about 3×10^6 daltons; it is also known as connectin. These connections represent individual titin molecules spanning the distance between Z lines and M lines. The length of these molecules varies with the length of the sarcomere. The cDNA encoding large stretches of titin reveal 11 domain superrepeats of titin spanning the 43-nanometer repeat of the thick filament helix. Each superrepeat contains two types of motif; one resembles subunits of fibronectin, a protein involved in establishing links between cells and the extracellular matrix, the other motif being related to the C-2 immunoglobulins. Another giant protein, nebulin (mass 8×10^5 daltons), is also involved in maintaining continuity between Z bands and M filaments forming presumably elastic connections with actin filaments. Nebulin also contains repeating domains, each consisting of 35 amino acid residues and each constituting an actin-binding domain. It has been suggested that nebulin plays a role in determining the length of actin filaments. Both proteins may play a role in maintaining the orderly arrangements of thick and thin filaments.

The main component of the Z disk is the protein α-actinin. Other proteins, including desmin, are also involved. Smooth muscle also contains thick and thin filaments that slide past each other as the muscle contracts. In these muscles, however, the filaments are not in register, and so sarcomeres and Z disks do not exist and striation patterns cannot be seen. Thin filaments are attached to irregularly disposed structures, the so-called dense bodies.

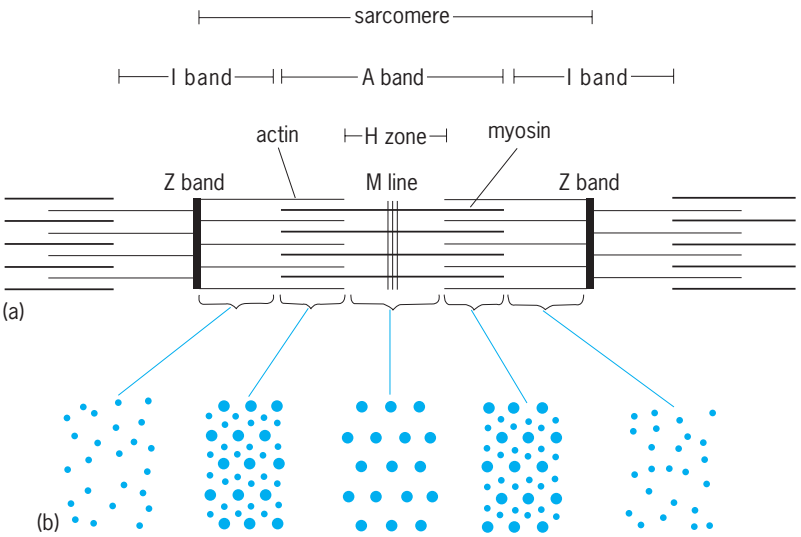

Fig. 1. Origin of striations in muscle. (*a*) Banding pattern of a sarcomere, showing the arrangement of myosin and actin molecules. (*b*) Cross-sectional appearance of sarcomere at different points along its length. (*After R. Craig, The structure of the contractile filaments, in A. G. Engle and B. Q. Banker, eds., Myology, McGraw-Hill, 1986*)

Both chief constituents of the muscle filaments can exist as single molecules (monomeric form) or in self-assembled filamentous structures (polymeric form). When isolated under the usual conditions, both myosin and actin are obtained as monomers.

Myosin. Molecules of myosin, amounting to about 60% of the total muscle protein, are arranged in filaments occupying the central zone of each segment (sarcomere) of the fibril, the A band. Myosin is an elongated molecule whose overall length is about 150 nm (**Fig. 2**). In contrast to many other fibrous molecules, myosin is not uniform throughout its length. It is made up of two intertwined heavy

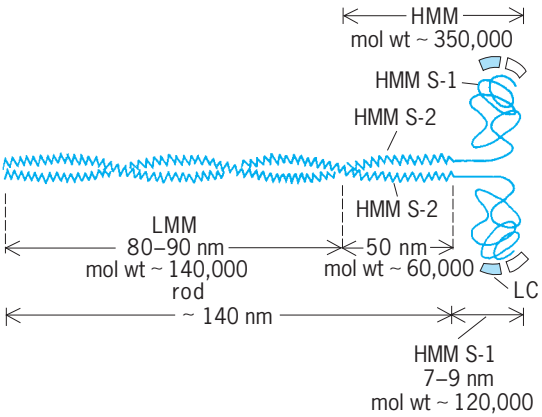

Fig. 2. Schematic representation of the structure of the myosin molecule. Hinge regions postulated in the mechanism of contraction are at the junction of HMM S-1 and HMM S-2 and of HMM S-2 and LMM (light meromyosin). It should be noted that HMM S-1 has one chief polypeptide chain, while the other fragments have two. The scheme suggests the presence of two different subunits in each HMM S-1. (*After J. N. Walton, ed., Disorders of Voluntary Muscle, 3d ed., Longman, 1974*)

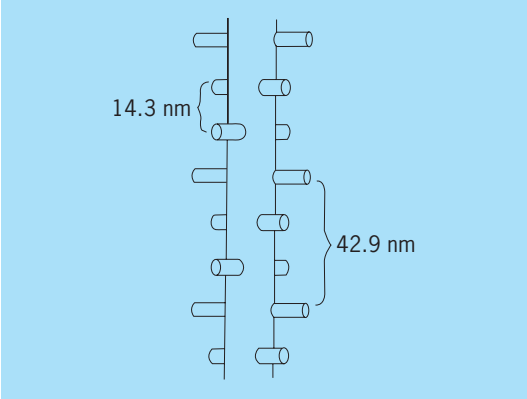

Fig. 3. Schematic diagram of the arrangement of cross bridges in a vertebrate thick filament, showing a three-stranded helix. The peg-shaped bridges are purely schematic and should be interpreted to show only the location and not the shape of the cross bridges. (*After A. T. Bull et al., eds., Companion to Biochemistry: Selected Topics for Further Study, Longman, 1974*)

peptide chains (molecular weight of about 2×10^5) whose ends form two separate globular structures. The intertwined portion has a high α-helical content and forms a rigid rod. Attached to each head is a set of two light chains, one of which can be phosphorylated. A light chain of the latter type is involved in the regulation of smooth muscle and molluscan muscle, hence the name regulatory light chain. The amino acid sequence of both light and heavy chains is different among different muscle cells, such as those of fast skeletal muscle, smooth muscle, cardiac muscle, and embryonic muscle. That gives rise to many myosin forms.

The diversity of myosin is due to the existence of a family of genes coding for the heavy and light chains. These genes show slight variation between species. Within a given species, the expression of these genes shows developmental control as well as tissue specificity in muscles of different function, and it is also susceptible to change depending on the activity of the muscle.

Myosin functions as an ATP-cleaving enzyme, adenosine triphosphatase (ATPase). Each myosin head has an ATP binding site as well as an actin-binding site. The ATPase activity of myosin itself is low in media that contain magnesium ion, but if actin is added to myosin at low ionic strength in the presence of magnesium ion, considerable activation of ATPase takes place. This actin-activated ATPase reaction, which furnishes energy of contraction and the interaction of myosin with actin, is an essential element of the contraction mechanism. *See* ADENOSINE TRIPHOSPHATE (ATP).

Light meromyosin shows a tendency to form regular aggregates, which is thought to be involved in the process leading to the assembly of myosin molecules into a thick filament. Thus there is a central core of strongly interacting light meromyosin segments with radially protruding soluble heavy meromyosin portions (**Fig. 3**). Under the electron microscope, the heavy meromyosin moieties appear as projections issuing from the thick filaments that occasionally touch adjacent actin filaments. These projections have been called cross bridges; a cross bridge is an S-1 moiety attached by an S-2 to the core of the thick filament. In vertebrate muscle there are three cross bridges, separated by 120° angles on planes axially separated by 14.3 nm. The cross-bridge triplets are rotated by 40° on successive planes. The central portions of the thick filaments are devoid of cross bridges.

Actin. In the monomeric form, actin is a bilobular-shaped, nearly globular unit with a diameter of about 6 nm, a height of about 4 nm, and a molecular mass of 4.2×10^4 daltons. Actin, as isolated in the monomeric form, contains a tightly bound atom of magnesium and one bound ATP molecule in a cleft between the two lobes. The monomeric units are arranged in a helical fashion in the thin filament. Since the actin units are not bound by covalent bonds, the helical arrangement can be regarded either as one involving a single helix or one describable as the intertwining of two helices. The change from the globular to the fibrous form can be brought about in the test tube by adding salts and is accompanied by the hydrolysis of the bound ATP to bound adenosine diphosphate (ADP). The transformation of ATP to ADP taking place during polymerization is not involved in muscle contraction. This reaction presumably occurs when actin filaments are laid down in the course of development, growth, or regeneration. *See* ADENOSINE DIPHOSPHATE (ADP).

Activation by calcium. The interaction of actin and myosin in striated muscle is regulated by the cellular calcium level through a protein complex attached at regular intervals to the thin filament. Two protein complexes associated with the actin filaments are tropomyosin and troponin. Tropomyosin is an α-helical protein. A stretch of actin contains 13 monomers in contact with a pair of tropomyosin molecules arranged in a diametrically opposite fashion. Each tropomyosin molecule carries a troponin complex of three proteins, troponin-C, -I, and -T (**Fig. 4**). Troponin-T is the subunit most likely to be involved in association with tropomyosin. Troponin-I is known for its ability to inhibit the interaction of actin and myosin. Troponin-C acts as a calcium sensor to detect the myoplasmic calcium concentration, which in resting muscle is kept low by the calcium pump of the sarcoplasmic reticulum, an intracellular membrane system associated with the T tubule, which in turn is an invagination of the cell

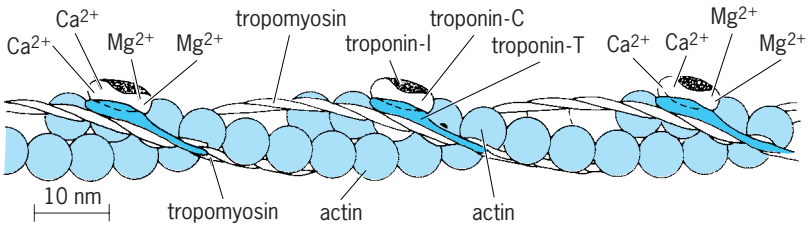

Fig. 4. Schematic representation of regulated thin filament. (*After R. L. Moss, J. D. Allen, and M. L. Greaser, Effects of partial extraction of troponin complex upon the tension-pCa relation in rabbit skeletal muscle: Further evidence that tension development involves cooperative effects within the thin filament, J. Gen. Physiol., 87:761–774, 1986*)

membrane. Other components include a ryanodine receptor, triadine, and a voltage sensor protein participating in the process of calcium ion release. When the calcium level is high enough, the contractile machinery is turned on. As the stimulus ceases, the calcium level drops and the machinery is again turned off. Changes in the protein structure of troponin-C that occur upon binding of calcium ions lead to changes in the other components of the thin filament and finally to activation of the actin–myosin cross-bridge interaction.

The mechanism by which the calcium ion–induced changes in the thin filament lead to activation of muscle contraction is still not fully understood. Earlier views regarded the attachment of myosin heads to actin as the regulated step. In contrast, later biochemical studies of myosin ATPase activated by regulated actin in the test tube suggest that activation involves a shift among states of myosin attached to actin, from one of weak binding to one of strong binding. Time-resolved x-ray diffraction on live muscle shows that the first structural changes in the thin filaments may involve a movement of tropomyosin, but the link between the movement and the changes in actin–myosin interaction require further clarification.

In striated muscles of invertebrates, the regulatory machinery is different. In mollusks, for example, sensing of calcium ion is performed solely by myosin, with the light chains playing a large role and tropomyosin troponin being unnecessary. A myosin-linked mechanism also operates in vertebrate smooth muscle, where phosphorylation triggered by the increase in calcium ion is involved. In smooth muscle, the thin filaments do not contain troponin. Regulation involves the phosphorylation of myosin, the reaction being catalyzed by an enzyme (myosin light-chain kinase), that is activated by the combination of calcium ions released on stimulation with calmodulin, a protein present in all cells. *See* CALCIUM METABOLISM; MUSCLE.

The structure of dystrophin and the associated complex of proteins containing sugarlike molecules (glycoproteins) attached to the cell membrane has been clarified. The absence of dystrophin and associated proteins leads to Duchenne muscular dystrophy. *See* MUSCULAR DYSTROPHY. John Gergely

Bibliography. A. G. Engel and B. Q. Banker, *Myology*, 2d ed., 1994; Z. Grabarek, T. Tao, and J. Gergely, Molecular mechanism of troponin-C function, *J. Muscle Res. Cell Motil.*, 13:383–393, 1992; K. Holmes and W. Kabsch, Muscle proteins: Actin, *Curr. Opin. Struc. Biol.*, 1:120, 1991; K. Matsumura and K. P. Campbell, Dystrophin-glycoprotein complex: Its role in the molecular pathogenesis of muscular dystrophies, *Muscle and Nerve*, 17:2, 1994; I. Rayment et al., Structure of the actin-myosin complex and its implications for muscle contraction, *Science*, 261:58, 1993; I. Rayment et al., The three-dimensional structure of myosin subfragment 1: A molecular motor, *Science*, 261:50, 1993; J. Trinick, Understanding the functions of titin and nebulin, *FEBS Lett.* (Federation of European Biochemical Societies), 307:44–48, 1992; Lord Walton (ed.), *Disorders of Voluntary Muscle*, 6th ed., 1994.

Muscovite

A mineral of the mica group with an ideal composition of $KAl_2(AlSi_3)O_{10}(OH)_2$. Sometimes it is referred to as a white mica or potash mica.

Structure. Muscovite is a dioctahedral mica, where two of the three possible octahedral cation sites are occupied by aluminum atoms (Al; referred to as M2 sites) and the other site (M1) is vacant. The M2 octahedra are formed with six oxygen atoms (O) or hydroxyl groups surrounding the aluminum atoms; these octahedra are linked laterally by the sharing of edges to form an octahedral sheet (see **illus.**). Tetrahedra are located on either side of the octahedral sheet. Each tetrahedron comprises silicon (Si) and aluminum (T site) bonded to four oxygen atoms, with generally 25% random substitution of silicon by aluminum at this site. A two-dimensional network of sixfold rings (ditrigonal) is formed by corner sharing of tetrahedra to produce a tetrahedral sheet. The single octahedral sheet sandwiched between the two tetrahedral sheets forms a 2:1 layer. Potassium (K)

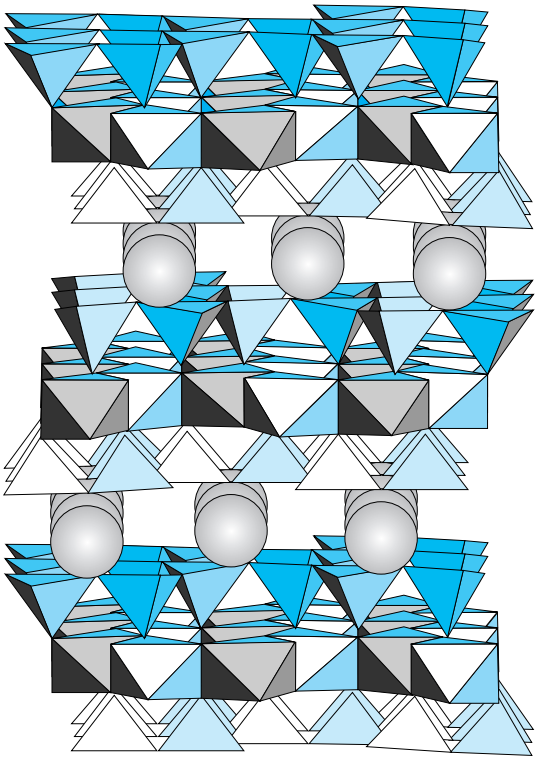

A polyhedral model illustrating the muscovite-*2M₁* atomic structure. The top and bottom layers are identical, but the middle layer differs; note that the octahedral faces are oriented differently. The structural unit includes the center layer and either the top or bottom layer. The corners of the tetrahedra represent oxygen atoms, whereas octahedral corners represent oxygen or hydroxyl groups. The octahedra contain aluminum ions, whereas the tetrahedra contain aluminum or silicon. The polyhedral interiors, the aluminum and (aluminum/silicon) ions, are not shown. Large spheres are interlayer potassium ions, which serve to connect the layers. Hydrogen atoms are not shown.

ions (A site) are located in the interlayer region, where this large cation locks together ditrigonal rings from adjacent tetrahedral sheets.

The stable form of muscovite is a two-layer ($2M_1$) structure, with overall monoclinic symmetry in space group $C2/c$. Regular monoclinic one-layer ($1M$) and disordered forms ($1M_d$) are less abundant, and a trigonal three-layer ($3T$) and a two-layer form ($2M_2$) different from $2M_1$ are rare. An apparent stability sequence exists with $1M_d$ transforming to $1M$ and then to $2M_1$ as either temperature or annealing time is increased. K-deficient forms are generally $1M$ in structure. *See* CRYSTAL STRUCTURE.

Chemistry. The A sites in muscovite may contain K, sodium (Na), barium (Ba), cesium (Cs), and rubidium (Rb); the M2 sites may have minor amounts of magnesium (Mg), ferrous iron [Fe(II)], ferric iron [Fe(III)], lithium (Li), manganese (Mn), titanium (Ti), and chromium (Cr), in addition to aluminum; the T sites contain only silicon and aluminum, and minor amounts of fluorine (F) can replace the hydroxyl group. In the older literature, the term fuchsite was used for the green variety of muscovite, which contains chromium. Solid solutions between other micas, either dioctahedral or trioctahedral, are quite limited. Related minerals include illite, which is similar to muscovite but has more silicon, magnesium, and water (H_2O) and less potassium; and phengite, which contains significant amounts of divalent cations in the M sites and more silicon than muscovite. *See* ILLITE; LEPIDOLITE; SOLID SOLUTION.

Properties. The layerlike qualities of the structure produce a perfect basal (001) cleavage. Specific gravity is 2.76–2.88, hardness on the Mohs scale is 2–2.5, and luster is vitreous to pearly. Thin sheets are flexible and may be colorless, with books (thick crystals) translucent, yellow, brown, reddish, or green. Cation exchange capacity is low because the interlayer cation fully compensates the residual charge on the 2:1 layer. Upon heating in air to temperatures above dehydroxylation [the actual temperatures of dehydroxylation are kinetically controlled, about 800°C (1470°F) for larger crystals, and 500°C (930°F) for fine-grained material], muscovite transforms topotactically to a dehydroxylate structure where the M sites become five-coordinated. *See* COORDINATION CHEMISTRY; HARDNESS SCALES.

Occurrence. Muscovite occurs commonly in all the major rock types, in igneous rocks (granites, pegmatites, and hydrothermal alteration products), in metamorphic rocks (slates, phyllites, schists and gneisses), and in sedimentary rocks (sandstones and other clastic rocks). *See* ROCK.

Uses. As larger flakes, muscovite is used as an electrical insulator, both for its dielectric properties and for its resistance to heat. Ground muscovite is used for fireproofing, as an additive to paint to provide a sheen and for durability, as a filler, and for many other applications. *See* MICA; SILICATE MINERALS.

Stephen Guggenheim

Bibliography. S. W. Bailey (ed.), Micas, *Reviews in Mineralogy*, vol. 13, 1984.

Muscular dystrophy

A group of muscle diseases that are hereditary and characterized by progressive muscle weakness and wasting.

Etiology. The muscular dystrophies are primary diseases of the muscle cells characterized by progressive degeneration and replacement by fibrous tissue, resulting in progressive muscle weakness. In some types of muscular dystrophies, the disease appears to be restricted to the skeletal muscles alone (facioscapulohumeral muscular dystrophy, limb-girdle muscular dystrophy), and in others skeletal-muscle involvement is a part of a more generalized process, with abnormalities in other organ systems as well (Duchenne's muscular dystrophy, myotonic dystrophy). These features as well as the differing patterns of inheritance (see **table**) indicate that the various muscular dystrophies are different diseases with different genetic and biochemical abnormalities underlying them.

The gene for the Duchenne and the Becker muscular dystrophy has been identified and localized to band Xp21 on the short arm of the X chromosome. This gene produces the muscle protein dystrophin, absent in Duchenne's dystrophy and qualitatively altered in Becker's dystrophy. Although the gene for myotonic dystrophy has not been identified, it has been found to be closely linked to the secretor locus and has been mapped to the proximal long arm of chromosome 19. *See* HUMAN GENETICS; MUSCLE PROTEINS.

Pathology. The muscle shows varying degrees of necrosis of individual muscle fibers with phagocytosis, as well as regeneration of some fibers. In addition, in many fibers there is an internal migration of the muscle cell nuclei. The individual muscle fibers become rounded and show considerable variability in their diameter in a random fashion. As the disease progresses, the necrotic muscle fibers are replaced by fibrous connective tissue and fat. The changes in the muscle are much more marked in the more rapidly progressive forms of muscular dystrophy (for example, Duchenne's dystrophy), and may be minimal in the slowly progressive ones (for example, myotonic dystrophy).

As a reflection of muscle degeneration, enzymes originating in skeletal muscle can often be detected in the serum in increased amounts.

Diagnosis. The diagnosis of the various types of muscular dystrophies still remains largely descriptive, and is based on the sex of the patient, the age of onset, the muscle groups initially affected, other muscular and nonmuscular abnormalities, and the pattern of inheritance. Since no specific laboratory test is available for most muscular dystrophies, other than the X-linked types, the number of different types varies depending on the authority consulted. The most common forms are the X-linked recessive Duchenne's muscular dystrophy and the dominantly inherited myotonic dystrophy. Other, less common forms include the milder X-linked Becker's muscular dystrophy, which is allelically related to Duchenne's;

Characterization of the muscular dystrophies

Type	Age	Involved muscles	Other features	Inheritance
Pseudohypertrophic				
Duchenne	Less than 4 years	Pelvic girdle	Pseudohypertrophy, mental retardation, abnormal cardiogram	X-linked
Becker	8–10 years	Pelvic girdle	Pseudohypertrophy, abnormal cardiogram	X-linked
Facioscapulohumeral	Adolescence	Face, shoulder, upper arm	—	Dominant
Myotonic	Young adult	Face, neck, hands	Myotonia, cataract, gonadal atrophy, balding	Dominant
Limb girdle	Adolescence or older	Shoulder and pelvic girdle	—	Recessive
Ocular	Any age	Extraocular muscles	Very variable	Dominant or recessive

dominantly inherited facioscapulohumeral muscular dystrophy; and the recessively inherited limb-girdle muscular dystrophy. The ocular myopathies appear to include several different disease entities, all of which have as a common feature the progressive paralysis of extraocular muscles.

Duchenne's muscular dystrophy. This is the most rapidly progressive form of muscular dystrophy. It affects boys before the age of 4, and is characterized initially by progressive weakness of the hip muscles with difficulty in rising from the floor or chair and in climbing stairs. This is accompanied by enlargement of the calf muscles, which are infiltrated by fat and fibrous tissue (pseudohypertrophy). Weakness of the muscles of the upper arms and shoulder muscles follows. By 10 to 12 years of age, most of the affected children are restricted to a wheelchair, and the majority die in the late teens or early twenties. In the majority of cases, their mothers, who are carriers of this gene, usually have few or no symptoms, but often have elevated serum creatine kinase levels. Although their daughters do not develop this disease, each has a 50% chance of being a carrier of this gene, with a 50% possibility of transmitting this disease to each of her male offspring. In the remaining cases, the affected boy appears to represent a new mutation for this gene and the mother is not the carrier.

Complementary deoxyribonucleic acid (cDNA) probes for the Duchenne muscular dystrophy gene and antibodies to dystrophin now allow identification of many affected male children or fetuses and obligatory female carriers of this gene. The allelically related Becker's muscular dystrophy is also characterized by calf pseudohypertrophy but is much more slowly progressive. The genetic abnormality resides in the same site of the X chromosome as with the Duchenne's type, but the muscle protein dystrophin is present in skeletal muscles, although in altered form.

Myotonic dystrophy. Myotonic dystrophy is very slowly progressive and affects the muscles of the face, neck, and hands. It usually begins in early adulthood. In addition to progressive weakness and wasting of the affected muscles, these individuals also exhibit myotonia, that is, a delayed relaxation of a muscle after forceful contraction. Myotonia is most readily demonstrated in the hands, where there is a slow relaxation of the fingers after forceful grasping. However, the abnormality in myotonic dystrophy is not restricted to the skeletal muscles. Mental retardation, frontal balding, cataracts, and gonadal degeneration are common. Incidence of diabetes mellitus and cardiac disease is higher in patients with myotonic dystrophy than in the general population. The close linkage to genetic markers on chromosome 19 allows the identification of presymptomatic or asymptomatic members of some families.

Treatment. Treatment remains largely symptomatic. However, the identification of the Duchenne muscular dystrophy gene and the resultant absence of dystrophin raise the possibilities of treatment by genetic manipulation or replacement of dystrophin. In the meantime, these findings provide more accurate prenatal diagnosis and genetic counseling. *See* MUSCULAR SYSTEM DISORDERS.

S. M. Sumi; Thomas D. Bird

Bibliography. G. L. Barnett, *Muscular Dystrophy*, 1995; A. E. Emery, *Duchenne Muscular Dystrophy*, 2d ed., 1993; A. G. Engle and C. Franzini-Armstrong, *Myology*, 2 vols., 2d ed., 1994; R. N. Rosenberg and A. E. Harding (eds.), *Molecular Biology of Neurological Disease*, 1988; J. Walton, G. Karpati, and D. Hilton-Jones (eds.), *Disorders of Voluntary Muscle*, 1994.

Muscular system

The muscular system consists of muscle cells, the contractile elements with the specialized property of exerting tension during contraction, and associated connective tissues. The three morphologic types of muscles are voluntary muscle, involuntary muscle, and cardiac muscle. The voluntary, striated, or skeletal muscles are involved with general posture and movements of the head, body, and limbs. The involuntary, nonstriated, or smooth muscles are the muscles of the walls of hollow organs of the digestive,

circulatory, respiratory, and reproductive systems, and other visceral structures. Cardiac muscle is the intrinsic muscle tissue of the heart. *See* MUSCLE.

In this article the comparative embryology of the voluntary and involuntary muscles of the vertebrates will be outlined, followed by the comparative anatomy of the muscular system.

Comparative Anatomy

Phylogenetically speaking, muscles are very plastic organs. They have undergone considerable change during the evolution of vertebrates, correlated in large part with changes in the organisms' environments and in their methods of support and locomotion. Establishment of homologies among muscles is not easy. Adult relationships can be misleading because muscles have subdivided during their evolution, and parts have migrated far from their original positions. Nerve supply is a more reliable criterion, because nerves have tended to follow the muscles through their evolutionary gymnastics, but often homologies cannot be established without recourse to embryonic development. Determination of the development of the thousands of individual muscles among the vertebrate classes is a monumental task. Comparison of muscles among vertebrates is greatly facilitated if the muscular system is subdivided into groups whose homology can be more easily established in the various classes.

Muscle groups are particularly distinct in elasmobranchs and other primitive fishes, and they are generally defined on the basis of their embryonic origin in these animals. Two major groups of skeletal muscles are recognized, somatic (parietal) muscles, which develop from the myotomes, and branchiomeric muscles, which develop in the pharyngeal wall from lateral plate mesoderm. The somatic musculature is subdivided into axial muscles, which develop directly from the myotomes and lie along the longitudinal axis of the body, and appendicular muscles, which develop within the limb bud from mesoderm derived phylogenetically as buds from the myotomes.

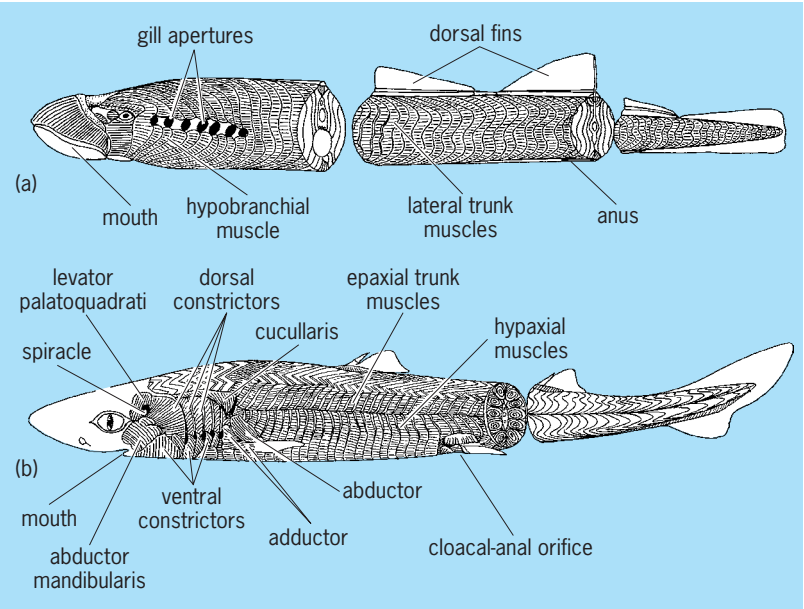

Fig. 1. Superficial muscles. (*a*) Cyclostome (*Petromyzon*). (*b*) Elasmobranch (*Squalus*). (*After H. W. Rand, The Chordates, Blakiston, 1950*)

The vertebrate muscular system is the largest of the organ systems, making up 35–40% of the body weight in humans. The movement of vertebrates is accomplished exclusively by muscular action, and muscles play the major role in transporting materials within the body. Muscles also help to tie the bones of the skeleton together and supplement the skeleton in supporting the body against gravity. *See* SKELETAL SYSTEM.

Axial musculature. Most of the axial musculature is located along the back and flanks of the body, and this part is referred to as trunk musculature. But anteriorly the axial musculature is modified and assigned to other subgroups. Certain of the occipital and neck myotomes form the hypobranchial muscles, and the most anterior myotomes form the extrinsic ocular muscles.

The trunk musculature of cyclostomes consists of a long series of segmental myomeres, each consisting of many longitudinal fibers attaching onto the myosepta (**Fig. 1**). Each is folded in such a way as to appear approximately zigzag-shaped on the surface. The arrangement in jawed fishes is essentially the same, but the folding of the myomeres is more complex, and each is divided by a horizontal connective-tissue septum into dorsal (epaxial) and ventral (hypaxial) portions. A spinal nerve passes to each myomere, the dorsal ramus going to the epaxial portion and the ventral ramus to the hypaxial portion. This pattern of innervation persists in all higher vertebrates.

Epaxial musculature. The epaxial musculature remains powerful in most cases. In amphibians it consists of a group of medial and deep fibers that interlace the vertebrae, and a larger group of superficial fibers (dorsalis trunci). Segmentation is retained and undulations of the trunk and tail still play a role in the locomotion of many amphibians (**Fig. 2**).

In typical reptiles, the epaxial musculature is more complex. A medial and deep group of small, largely segmental muscles bind the vertebrae together and constitute the transversospinalis system; more laterally the musculature is arranged in two more extensive longitudinal groups, the longissimus dorsi, which lies dorsal to the transverse processes, and the iliocostalis, which is attached to the ribs.

These three main divisions persist in mammals, but posteriorly there is a union of the iliocostalis, longissimus, and sometimes the more superficial part of the transversospinalis system to form a powerful erector spinae (sacropinalis) complex that helps to support the vertebral column. In mammals the body is held off the ground by the limbs; thus the backbone is sometimes compared to a girder supported anteriorly and posteriorly by pillars. Much of the epaxial musculature functions as tie members resisting tension stresses along this girder. Anteriorly there is a cleavage of the epaxial divisions into a host

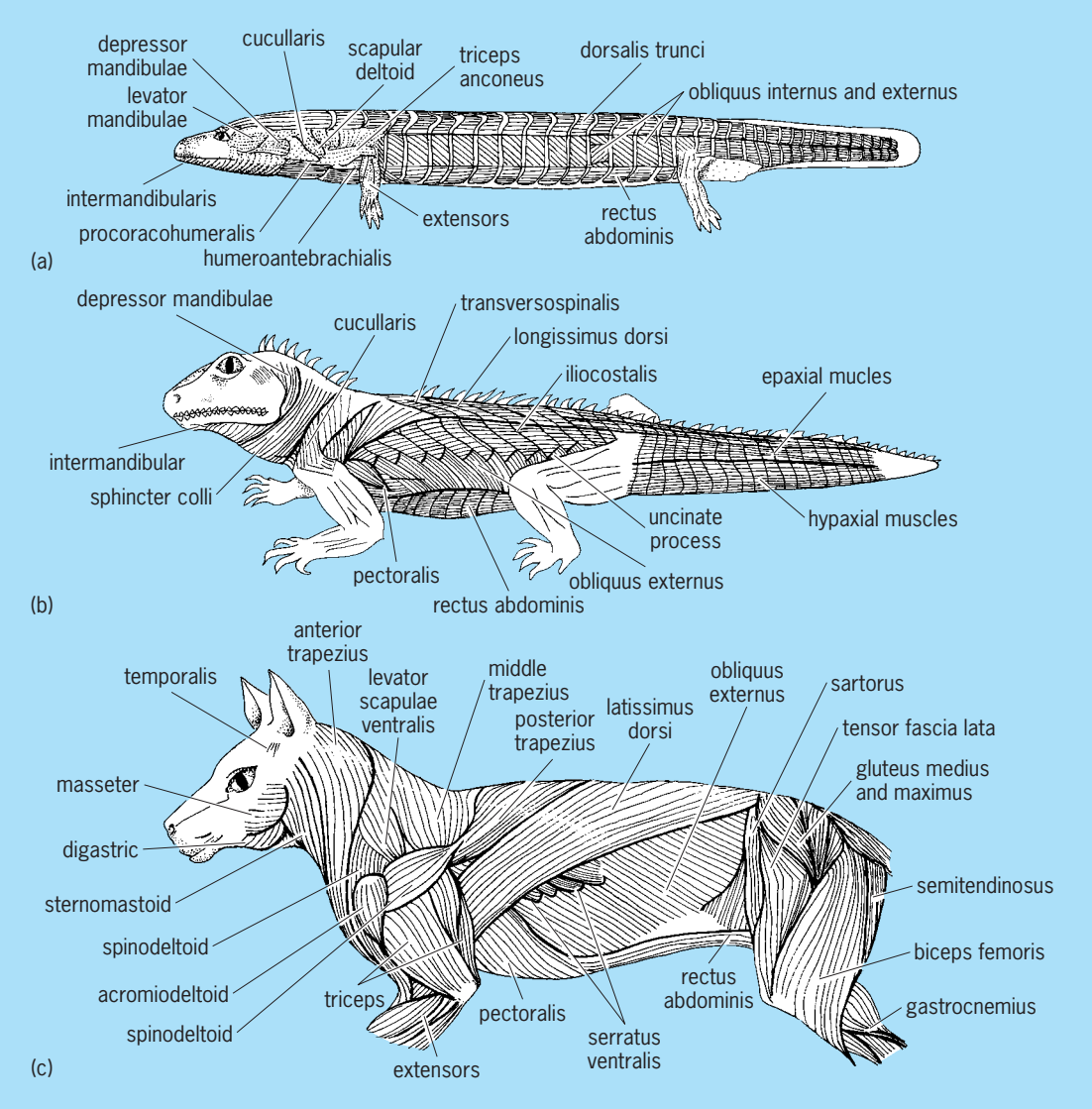

Fig. 2. Superficial muscles of three vertebrates, showing segmentation. (*a*) Amphibian (*Necturus*). (*b*) Reptile (*Sphenodon*). (*c*) Mammal (*Felis*). (*After H. W. Rand, The Chordates, Blakiston, 1950*)

of muscles associated with the complex head and neck movements.

In birds, the epaxial musculature in the trunk is greatly reduced, correlated with a fusion of many of the trunk vertebrae.

Hypaxial musculature. The hypaxial musculature of tetrapods can be subdivided into three groups: (1) a subvertebral (hyposkeletal) group located ventral to the transverse processes and lateral to the centra of the vertebrae, (2) the flank muscles forming the lateral part of the body wall, and (3) the ventral abdominal muscles located on each side of the midventral line.

The subvertebral musculature assists the epaxial muscles in the support and movement of the vertebral column. In mammals, it consists of longitudinal bundles—the longus colli in the neck and the anterior thorax, the quadratus lumborum, and psoas minor more posteriorly.

Most of the flank musculature takes the form of broad, thin sheets of muscle that form much of the body wall and support the viscera. The ancestral, segmental nature of this musculature is retained throughout the trunk in salamanders, but is lost in higher tetrapods except in those parts of the trunk where ribs are well developed (Fig. 2). Three layers can be distinguished in the abdominal region of most tetrapods: a superficial external oblique, whose fibers extend caudally and ventrally; an internal oblique with fibers at right angles to the preceding; and a deep transversus abdominis. This pattern is much the same in the costal region, external intercostals, internal intercostals, and a reduced transversus thoracis being present in mammals. In reptiles the pattern is more complex; the external layer is represented by supracostals, external intercostals, and sometimes a subcutaneous muscle.

Respiratory movements of reptiles and birds are accomplished by the costal and abdominal muscles described above, but in mammals, which have a higher metabolic rate, additional respiratory muscles have evolved from the hypaxial muscles: the

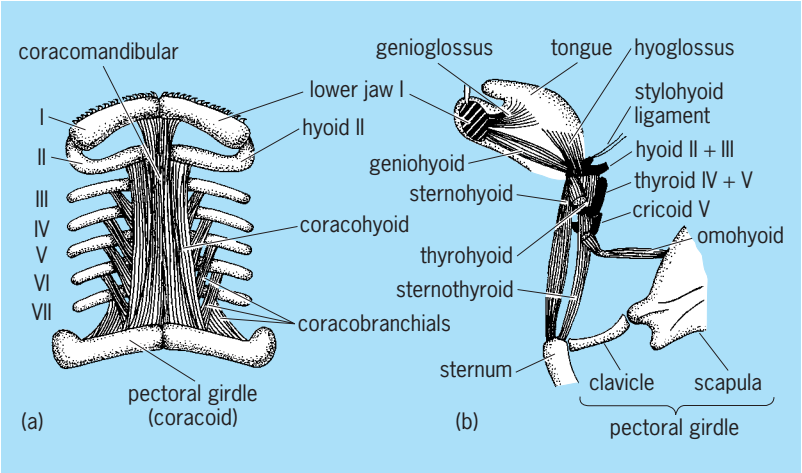

Fig. 3. Diagram of the major hypobranchial muscles. (*a*) Shark, in ventral view. (*b*) Mammal, in lateral view. (*After H. W. Rand, The Chordates, Blakiston, 1950*)

diaphragm (a derivative of cervical myotomes), serratus dorsalis, scalenes, and transversus costarum. *See* RESPIRATORY SYSTEM.

Other parts of the hypaxial flank musculature have gained an attachment to the pectoral girdle where they help to transfer body weight from the vertebral column to the girdle and appendage and help to regulate the movement of the girdle. Only a few muscles of this type, the thoracoscapularis and levator scapulae, for example, are present in primitive tetrapods such as salamanders, and the body is not held far off the ground. In mammals, however, this group includes such large and powerful muscles as the serratus ventralis, rhomboideus, and levator scapulae ventralis. In the pelvic region of tetrapods, weight is transferred to the appendage directly across the sacroiliac joint and not by muscles.

The midventral hypaxial musculature in all tetrapods consists of the rectus abdominis, a longitudinal muscle on each side of the midline that extends from the pelvic region to the anterior part of the trunk (Fig. 2). It has evolved from the oblique flank muscles and in some salamanders remains closely associated with them. Transverse tendinous inscriptions are often present and are believed to represent persistent myosepta.

Hypobranchial musculature. The hypobranchial musculature extends from the pectoral girdle forward along the ventral surface of the neck and pharynx to the hyoid arch, chin, and into the tongue. It is regarded as a continuation of part of the hypaxial trunk musculature, because it develops ontogenetically in most vertebrates from the ventral portion of several occipital and neck myotomes that grow around the back of the gill region and then into the neck and tongue. The innervation of the hypobranchial muscles in amniotes by the cervical nerves and the hypoglossal nerve, which itself has evolved from certain of the spinal and occipital nerves of fishes and amphibians, further indicates the myotomic origin of this group.

The hypobranchial musculature of cyclostomes (Fig. 1) retains its primitive myomeric character, but the myomeres fuse to form longitudinal muscles in higher vertebrates. Traces of myosepta are evident in fishes and some amphibians, but these disappear in amniotes.

In all gnathostomes, the hypobranchial musculature can be divided at the level of the hyoid arch into prehyoid (infrahyoid) and posthyoid (suprahyoid) groups (**Fig. 3**). The prehyoid group of elasmobranchs consists of a single pair of muscles, the coracomandibulars, extending caudally from the jaw symphysis to attach to the posthyoid group slightly anterior to the pectoral girdle. The posthyoid group consists of a superficial mass, which can be subdivided into a coracoarcural and coracohyoid, extending between the pectoral girdle and hyoid arch, and a deeper mass, the coracobranchials, extending from the pectoral girdle to the ventral surface of the branchial arches. The coracobranchials act to expand the pharynx and gill pouches; the others help to support the floor of the pharynx and help to move the hyoid arch and open the mouth.

During the evolution of terrestrial vertebrates, loss of most of the coracobranchials occurs, along with the reduction and loss of many parts of the branchial arches, but the rest of the hypobranchial musculature remains and becomes modified in correlation with the evolution of a muscular tongue and the more complex problem of deglutition in a terrestrial environment. The prehyoid group of amphibians consists primarily of a geniohyoid extending from the chin to the hyoid, but a few muscle fibers have separated from it and enter the tongue. These represent the beginning of a complex group of muscles that manipulates the tongue of amniotes: genioglossus, hyoglossus, styloglossus, and probably the intrinsic lingualis. The posthyoid group of primitive terrestrial vertebrates consists primarily of a rectus cervicis extending between the ventral part of the pectoral girdle and hyoid arch, but several slips separate from it to go to other parts of the girdle or to the remnants of the branchial arches. In higher vertebrates, the rectus cervicis has split into several muscles acting upon the larynx and hyoid: sternohyoid, sternothyroid, thyrohyoid, and omohyoid.

Extrinsic ocular muscles. The extrinsic ocular muscles develop from the prootic somites (head cavities). All vertebrates have six in common (**Fig. 4**). The median (internal), superior, and inferior recti, together with the inferior oblique, develop from the first somite and are innervated by the oculomotor nerve. The superior oblique and lateral (external) rectus develop from the second and third somites, respectively, and are supplied, respectively, by the trochlear and abducens nerves. All of these muscles insert on the eyeball and move it. In addition, most terrestrial vertebrates, with the exception of birds, have a retractor oculi, a cone-shaped muscle lying deep to the recti, that pulls the eyeball deeper into its socket. The retractor oculi has evolved from the lateral rectus and continues to be innervated by the abducens. In many reptiles and in birds, one or more small muscles have also separated from the lateral rectus, or from the retractor oculi, and act upon the nictitating membrane. A levator palpebrae superioris, present in mammals,

completes the ocular group. This muscle elevates the upper eyelid in opposition to the action of certain facial muscles. It has evolved from the superior rectus and is innervated by the oculomotor nerve. *See* EYE (VERTEBRATE).

Appendicular musculature. Limb muscles are often classified as intrinsic if they lie entirely within the confines of the appendage and girdle, and extrinsic if they extend from the girdle or appendage to other parts of the body. This scheme has certain merits, but is misleading from the phylogenetic point of view, because the extrinsic muscles are not all appendicular in the sense in which this group has been defined.

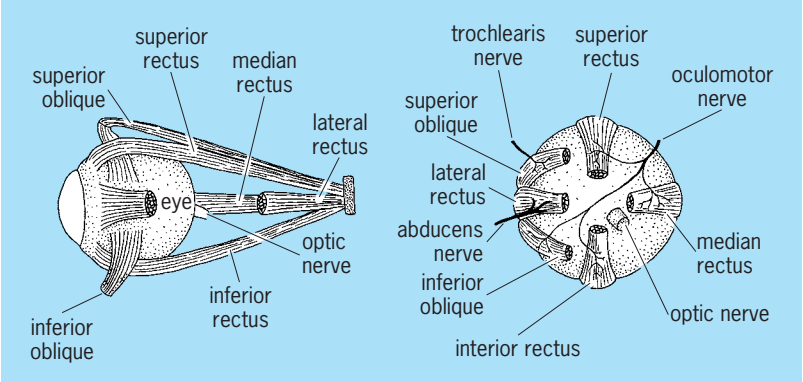

Fig. 4. Human extrinsic ocular muscles, which develop from prootic somites. (*After H. W. Rand, The Chordates, Blakiston, 1950*)

Homologies of appendicular muscles*

	Necturus	Cat
Pectoral muscles		
Dorsal	Latissimus dorsi	Citamepis trimci (part) Latissimus dorsi Teres major
	Subcoracoscapularis	Subscapularis
	Scapular deltoid	Spinodeltoid Acromiodeltoid
	Procoracohumeralis	Clavodeltoid Teres minor
	Triceps	Triceps Epitrochlearis
	Forearm extensors	Forearm extensors
Ventral	Pectoralis	Cutaneous trunci (part) Pectoralis complex
	Supracoracoideus	Supraspinatus Infraspinatus
	Coracoradialis	Biceps brachii
	Humeroantibrachialis (or brachialis)	Brachialis inferior
	Coracobrachialis	Coracobrachialis
	Forearm flexors	Forearm flexors
Pelvic muscles		
Dorsal	Illiotibialis	Sartorius Iliacus
	Puboischiofemoralis internus	Psoas major Pectineus Vasti(?)
	Ilioextensorius	Rectus femoris
	Iliofibularis	Gluteus maximus
	Iliofemoralis	Tensor fasciae latae(?) Gluteus medius Gluteus minimus
	Shank extensors	Shank extensors
Ventral	Puboischiofemoralis externus (adductor femoris, in *Necturus not* clearly separated from the preceding)	Obturator externus Quadratus femoris Adductor brevis of longus
	Pubotibialis	Adductor magnus
	Ischiofermoralis	Obturator internus Gemelli
	Caudofemoralis	Crurococcygeus (absent in some mammals) Pyriformis
	Puboischiotibialis	Gracillis
	Ischioflexorius	Semimembranosus Semitendinosus Biceps femoris(?)
	Shank flexors	Shank flexors

*From W. F. Walker, Jr., *Vertebrate Dissection*, Saunders, 1970.

Some are appendicular muscles that have developed directly from mesoderm in the limb bud, but others are trunk, hypobranchial, and branchiomeric muscles that have secondarily become associated with the girdle. These muscles are considered with the group to which they phylogenetically belong.

Fishes. The paired fins of fishes are primarily horizontal stabilizing keels, but they are also used in deceleration and steering, and to help control the depth at which the fish swims. Fin movements are not complex or powerful and the appendicular muscles in the strictest sense or morphologically simple. A single dorsal muscle (abductor) and a comparable ventral muscle (adductor) extend from the girdle into the fin (Fig. 1). Certain fibers of these muscles are so arranged that they can protract and retract the fin. The appendicular muscles are supplied by the ventral rami of spinal nerves.

Terrestrial vertebrates. In terrestrial vertebrates, the limbs become the main organs for support and locomotion, and the appendicular muscles become correspondingly powerful and complex. The muscles are too numerous to describe individually, but they can be sorted into dorsal and ventral groups, because tetrapod muscles originate embryonically in piscine fashion from a dorsal and a ventral premuscular mass within the limb bud (see **table**). In general, the ventral muscles, which also spread onto the anterior surface of the girdle and appendage, act to protract and adduct the limb and to flex its distal segments; the dorsal muscles, which also extend onto the posterior surface of the girdle and appendage, have the opposite effects (retraction, abduction, and extension). The limb muscles also serve as flexible ties or braces that can fix the bones at a joint and support the body.

Amphibians and reptiles. When at rest, the belly of most amphibians and reptiles is on the ground, and the proximal segment of each limb extends laterally and slightly dorsally from its articulation with the girdle. Locomotion involves the partial adduction of the humerus and femur to raise the body off the ground, as well as their protraction and retraction. Ventral adductor muscles, such as the pectoralis and supracoracoideus in the pectoral region and the puboischiofemoralis externus in the pelvic region, are

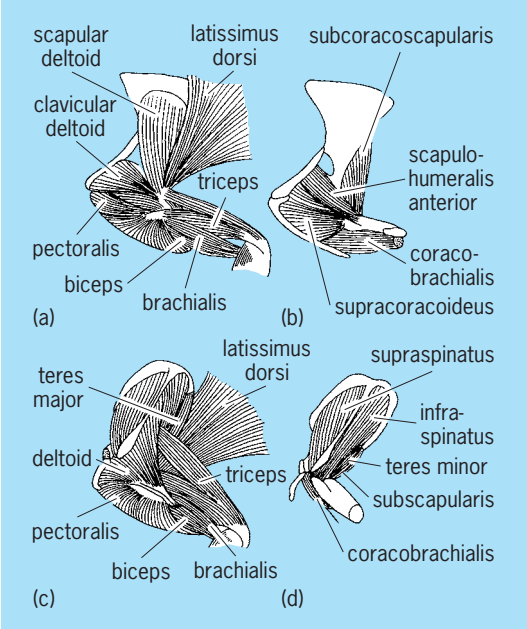

Fig. 5. Lateral views of the shoulder and upper arm muscle structure of two animals. (*a*, *b*) Lizard. (*c*, *d*) Opossum. Superficial muscles are shown in *a* and *c*; deep muscles are shown in *b* and *d*. (*After A. S. Romer, The Vertebrate Body, 3d ed., Saunders, 1962*)

relatively large and powerful. During the evolution of mammals, the limbs have rotated under the body and extend ventrally from the girdle. The body is held off the ground by bony columns braced by muscles. Adductor muscles are less powerful and certain ones of them have migrated to other positions and have assumed other functions. The supracoracoideus, for example, has extended dorsally onto the scapula to form the mammalian supraspinatus and infraspinatus (**Fig. 5**). These muscles now act as braces for the limb in its new position and play a role in its protraction and retraction.

Birds. Flight in birds has entailed a considerable modification of the musculature of the pectoral region. As one example, the ventral adductor muscles are exceedingly large and powerful, and the area from which they arise is increased by the enlargement of the sternum and the evolution of a large sternal keel. Not only does a ventral muscle, the pectoralis, play a major role in the downstroke of the humerus, but a ventral muscle, the supracoracoideus, is active in the upstroke as well. The tendon of insertion of the supracoracoideus has shifted so that it passes through a canal between the clavicle, coracoid, and scapula to attach to the upper surface of the humerus. Its action is analogous to pulling down on a rope that passes over a pulley and down onto a weight.

Branchiomeric musculature. The branchiomeric (branchial) muscles of fishes form a conspicuous part of the muscular system and are rather complex. In jawed fishes they can be subdivided according to the visceral arch with which they are associated. Mandibular muscles act upon the first, or mandibular arch, and are supplied by the trigeminal nerve. The group includes such muscles as the levator palatoquadrati, which in the dogfish helps to support the palatoquadrate cartilage; the adductor mandibulae, the powerful muscles closing the jaws; and the intermandibularis, which together with certain hypobranchial and hyoid muscles, opens the jaws (Fig. 1).

Hyoid muscles act on the second or hyoid arch and are supplied by the facial nerve. The hyoid arch is modified in sharks and many other fishes to help support the palatoquadrate, and its musculature is correspondingly modified. The gill slits of bony fishes are covered by an operculum, which has developed from the gill septum of the hyoid arch, and part of the hyoid musculature controls its movement.

The remaining visceral arches (branchial arches) and their muscles are associated with the gills. The muscles of the third arch are supplied by the glossopharyngeal nerve; those of the fourth through seventh arches by the vagus. Muscles of a typical branchial arch include constrictors, interbranchials, adductors, and interarcuals, all of which act to compress the gill pouches and force water out through the gill slits. The gill pouches are opened primarily by the action of the coracobranchials—a part of the hypobranchial musculature. In addition, each branchial arch has a levator, but the levators have united to form a single muscle, the cucullaris, and their insertion has shifted from all but the last branchial arch onto the pectoral girdle (Fig. 1).

During the evolution of terrestrial vertebrates, gills are lost and the visceral arches become reduced and greatly modified. Most of the mandibular muscles remain associated with the jaws and form the various muscles of mastication. In a mammal these are temporalis, masseter, pterygoids, anterior belly of the digastric, and the mylohyoid (Fig. 2). All but the last two close the jaws. The tensori palati, in the soft palate, and the tensor tympani, which attaches to the malleus (a derivative of the mandibular arch), also belong to this group. Only a few hyoid muscles remain associated with the hyoid arch or its derivatives: stylohyoid, posterior belly of the digastric, and stapedius. Most of the hyoid musculature has spread out beneath the skin of the face and neck to form the platysma and the numerous facial muscles (**Fig. 6**). Most of the musculature of the branchial arches is lost, but parts of it form the intrinsic muscles of the larynx and certain pharyngeal muscles. The cucullaris, in contrast, enlarges and subdivides to form the trapezius and sternocleidomastoid, muscles that act on the pectoral girdle and head (Fig. 2). The mammalian motor nerve to these muscles, the spinal accessory, is homologous to part of the vagus of fishes.

Integumentary musculature. In a number of terrestrial vertebrates, particularly amniotes, certain of the more superficial skeletal muscles of the body have spread out beneath the skin and inserted into it. These may be described as integumentary muscles, but it should be emphasized that they are not a natural phylogenetic group but are derived from several different groups.

Integumentary muscles are particularly well developed in mammals and include the facial muscles and

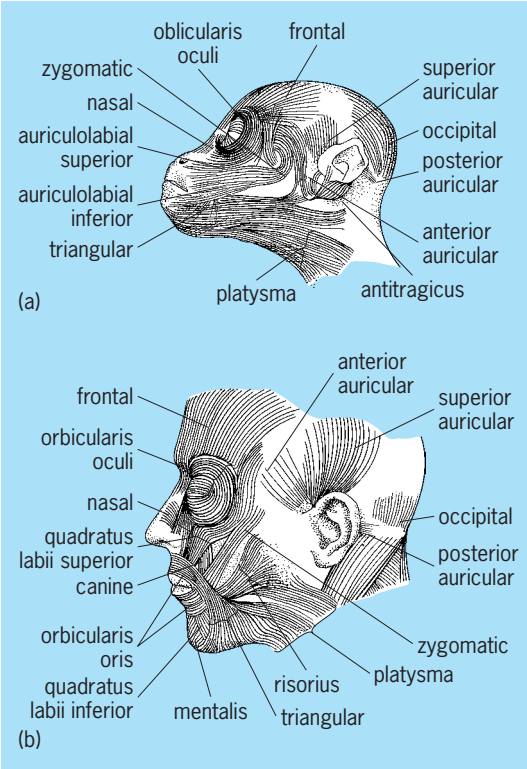

Fig. 6. Facial muscles. (*a*) Monkey. (*b*) Human. (*After H. W. Rand, The Chordates, Blakiston, 1950*)

platysma, derived from the hyoid musculature, and often a large cutaneous trunci. The last is derived from the pectoralis and latissimus dorsi and fans out beneath the skin of the trunk. The twitching of the skin of an ungulate is caused by this muscle.

Birds and reptiles have a sphincter colli (Fig. 2), a superficial neck muscle derived from the hyoid musculature and hence homologous to the platysma, but they lack facial muscles. Other integumentary muscles, derived from appendicular and trunk muscles, attach to the feathers, especially the large flight feathers on the wings and tails. In snakes, costocutaneous muscles extend from the ribs to the large ventral scales and play an important role in locomotion.

Histology

Three histological types of muscle are recognized: smooth, striated (skeletal), and cardiac. Smooth and cardiac muscle fibers generally occur in layers in the walls of organs, but striated muscle fibers are usually grouped into distinct entities, the skeletal muscles of gross anatomy. In some vertebrates, such as the rabbit, red and white skeletal muscles can be distinguished, but in most, an individual muscle contains a variable mixture of red and white fibers. Red fibers contain more sarcoplasm and myoglobin than white fibers. Myoglobin has a greater affinity for oxygen than hemoglobin, hence oxygen can be taken from the blood and stored by red muscle cells. The contraction of red fibers is more sustained, less subject to fatigue, and often slower than the contraction of white fibers. Red muscles, or muscles containing a preponderance of red fibers, tend to be found in situations where the muscles are particularly active in either moving the body or maintaining posture. Some examples are the diaphragm and other respiratory muscles, and the gluteus maximus. White muscles, or muscles with a preponderance of white fibers, are associated with a more intermittent and often a faster and more powerful movement. The biceps and digital muscles are examples.

Each skeletal muscle is individualized by a connective tissue sheath, the epimysium, which is a part of the deep fascia, and by its distinctive attachments onto skeletal elements. The epimysium is continuous with the connective tissue that invests the bundles of fibers within a muscle and with the connective tissue investing the individual fibers. Attachments to the skeleton are made by a continuation of this connective tissue into the periosteum surrounding the bone, and sometimes by connective tissue fibers that penetrate the bone. The connective tissue connection between muscle and bone may take the form of a cord-shaped tendon or a broad sheet known as an aponeurosis; or it may be relatively inconspicuous, in which case the muscle is said to have a fleshy attachment. The attachment of the muscle that tends to be stationary as the muscle contracts is its origin; the opposite attachment, which pulls on a structure that can be moved, is its insertion, but it must be recognized that the force exerted by a muscle during contraction is the same at each end.

Skeletal muscles vary greatly with respect to the number of fibers they contain and in the length and arrangement of their fibers. Important determinants of muscle architecture are the extent and nature of the movement that is to be brought about, the force that must be exerted, and the space available for the muscle in a particular area. The extent to which a muscle shortens is a function of the length of its fibers. Experimentally, muscle fibers can shorten to about one-half of their resting length, but they attach to bones so close to the joints that they seldom shorten this much. The strength of a muscle, on the other hand, is approximately proportional to the number of fibers it contains.

Many muscles, such as the sartorius on the thigh of mammals, have long fibers arranged parallel to each other (**Fig. 7**). An advantage of a strap muscle of this type is that it provides maximum excursion. Quite a different arrangement is seen in a muscle such as

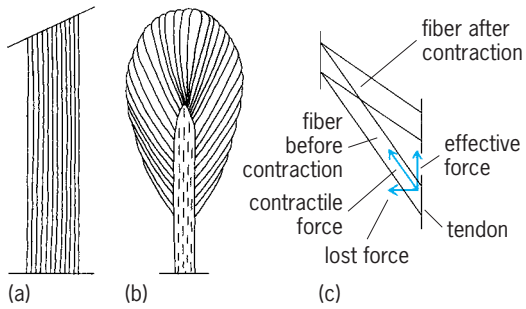

Fig. 7. Diagram showing two types of muscle architecture. (*a*) Parallel muscle. (*b*) Pinnate muscle. (*c*) Pinnate fiber before and after contraction.

the deltoid on the shoulder, which has a central tendon onto which many short, diagonal fibers attach at rather acute angles (Fig. 7). Pinnate muscles have many more fibers than a strap muscle of comparable mass, hence they can exert a greater force, but they shorten over a shorter distance. The full force of the contraction is not realized, however, because the muscle fibers pull on the tendon at an angle. The contractile force is resolved into an effective force along the axis of the tendon and a lost component at right angles to the tendon. This slight disadvantage is outweighed by major advantages. In this way much of the force of many muscle fibers can act through a common tendon upon a restricted area. This is possible because the strength of a tendon is 30–120 times that of a muscle of equal cross section. Another advantage is that the change in angle of the muscle fibers that occurs during contraction (Fig. 7) keeps bulging to a minimum. Some pinnate muscles, such as the tensor tympani in the middle ear of mammals, can act within a confined space.

The examples selected of muscles with fibers arranged completely parallel to one another or acutely pinnate are extremes of a continuum of muscle architecture. Some degree of pinnation is very common. The architecture of any particular skeletal muscle is dependent upon the requirements of packing within a given area, and the amount of force and extent of excursion needed.

Muscles are usually arranged so that one muscle or group of muscles will pull a structure in a certain direction, and an opposing muscle or group will pull the structure in the opposite direction. Several sets of terms describe these antagonistic actions. Flexion is the movement of a distal part of an appendage toward a more proximal part; this occurs at the elbow and knee. It also describes the bending of the head or trunk toward the ventral surface. Extension is the opposite movement. Flexion and extension are sometimes also applied to forward and backward movements of the appendage at the shoulder and hip, but because they have been used in conflicting senses by different authors, the terms protraction for a forward movement and retraction for a backward movement are more appropriate. Abduction is the movement of a part away from some point of reference, and adduction is movement toward it. For the appendages, the reference point is the midventral line of the body. Various types of rotary movement occur. For example, rotation of the bones of the forearm so that the palm of the hand faces up is supination; the opposite movement is pronation.

Muscle Mechanics

Many of the bones serve as lever arms, and the contractions of muscles are forces acting on these arms. The relationship between most muscles and bones is such that the lever systems are classified as third order (**Fig. 8**). The joint, of course, is the fulcrum and it is at one end of the lever. The length of the force arm is the perpendicular distance from the fulcrum to the line of action of the muscle; the length of the work arm is the perpendicular distance from

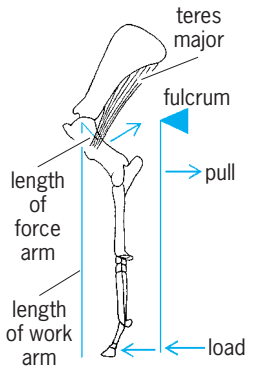

Fig. 8. Typical vertebrate lever system.

the fulcrum to the point of application of the power generated in the lever. Compactness of the body and physiological properties of the muscle necessitates that a muscle attach close to the fulcrum; therefore, the force arm is considerably shorter than the work arm. Most muscles are at a mechanical disadvantage, for they must generate forces greater than the work to be done, but an advantage of this is that a small muscular excursion can induce a much greater movement at the end of the lever.

Slight shifts in the attachments of a muscle that bring it toward or away from the fulcrum, and changes in the length of the work arm, can alter the relationship between force and amount of speed of movement.

In general, the force of a muscle is inversely related to the amount and speed of movement that it can cause. Certain patterns of the skeleton and muscles are adapted for extensive, fast movement at the expense of force, whereas others are adapted for force at the expense of speed. In the limb of a horse, which is adapted for long strides and speed, the muscles that move the limb insert close to the fulcrum and the appendage is long. This provides a short force arm but a very long work arm to the lever system (Fig. 8). In the front leg of a mole, which is adapted for powerful digging, the distance from the fulcrum to the insertion of the muscles is relatively greater and the length of the appendage is less, with the result that the length of the force arm is increased relative to the length of the work arm. *See* BIOMECHANICS.

Warren F. Walker

Locomotion. Different groups of muscles have different mechanical properties. Investigation of these properties in relation to their performance reveals the functional adaptation of the muscular system.

Running mammals. With increased running speed, the fraction of the stride for which the foot is on the ground decreases, so that the feet must exert larger forces while they are on the ground to make the average force over a complete stride match the body weight. Therefore, with increasing speed a greater muscle mass would need to be active to produce the necessary force.

The properties of extensor digitorum longus muscle and soleus muscles, which are components of the calf of the leg, are representative of fast and slow

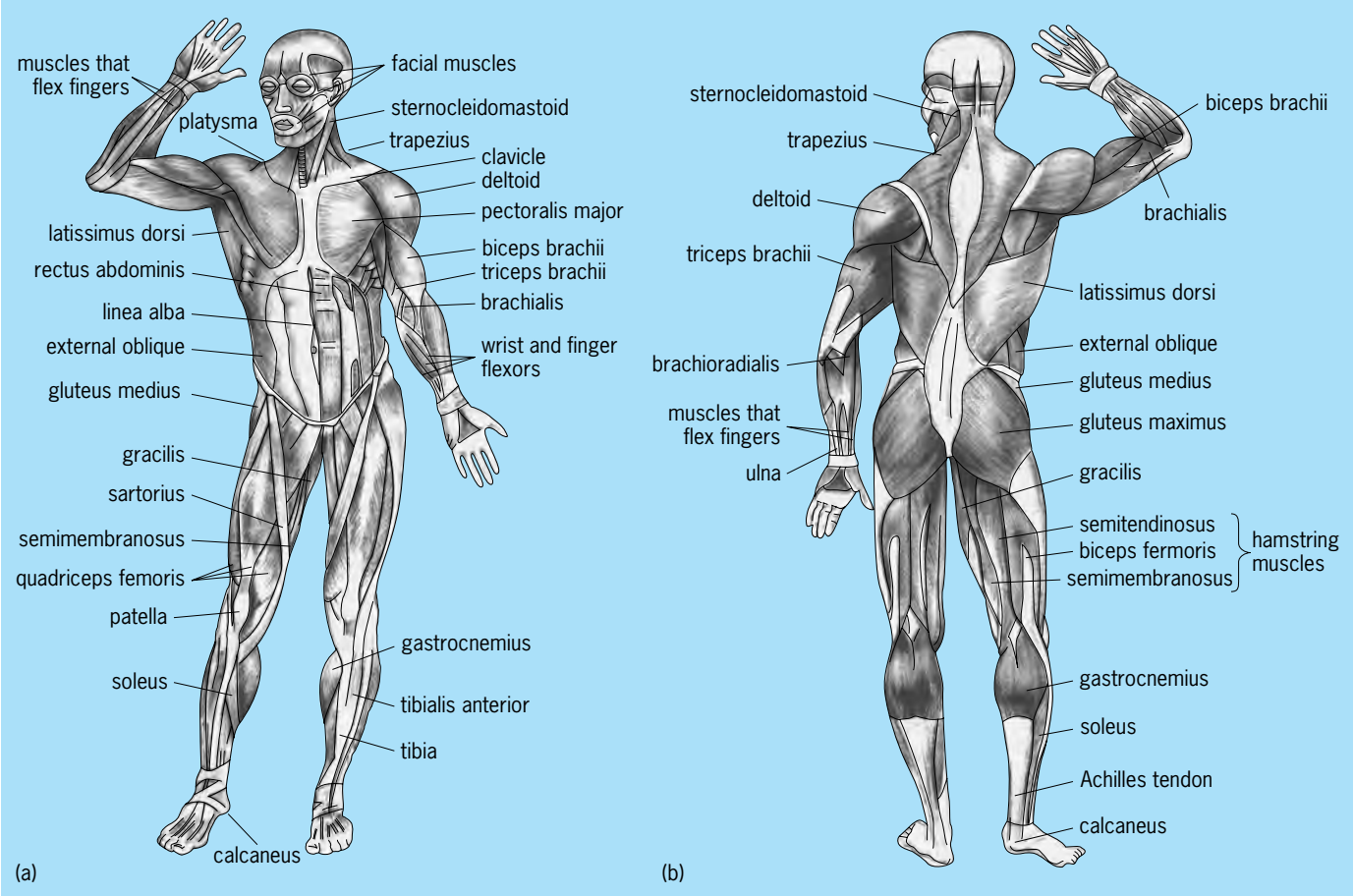

Fig. 9. Superficial muscles of the human body. (*a*) Anterior view. (*b*) Posterior view. (*After C. A. Villee et al., Biology, 2d ed., Saunders, 1989*)

muscle (**Fig. 9**). Extensor digitorum longus muscle consists wholly of fast glycolytic and fast-oxidative-glycolytic fibers in roughly equal proportions. In contrast, the soleus consists of only slow-oxidative and fast-oxidative-glycolytic fibers. These different fiber populations impart different properties to these two muscles. Extensor digitorum longus muscle has a faster contraction time, lower fatigue resistance, and higher power output than soleus. During normal function, these locomotory muscles undergo changes in length, doing work as they shorten and absorbing energy during reextension.

During walking and trotting, the soleus muscle produces 85–100% of its maximal power output. The frequency of maximum power output of extensor digitorum longus muscle coincides with the higher stride frequencies employed during fast galloping. This match between locomotion mechanics and muscle properties has been observed in other mammals and in reptiles—examples of the evolutionary optimization of design.

Fish swimming. As tailbeat frequency and swimming speed increase, there is a sequential recruitment of myotomal muscle (that is, skeletal muscle produced from a somatic cell) from superficial to deep muscle fibers. At slow, sustainable swimming speeds, only the slow muscle fibers close to the lateral line are active. The cost of locomotion is low at these speeds and, consequently, this fiber type typically makes up just a few percent of the total myotomal muscle mass. slow muscle fibers generate maximum power at low tailbeat frequencies and can sustain activity for long periods of time. The power required for swimming increases rapidly with increasing speed and requires the recruitment of fast muscle fibers, which form the bulk of the myotomal muscle. Fast muscle fiber has a higher intrinsic power output that generates maximum power at high tailbeat frequencies, but it fatigues rapidly; therefore, fish can swim at these high speeds of only a brief period of time.

Most fish swim primarily by lateral oscillations of the body. The nature of these oscillations changes among species, and is related to body form (**Fig. 10**). For example, long, slender, round-bodied fish (such as the eel) have a swimming pattern described as anguilliform, that is waves of muscle contraction pass alternately down both sides of the body, causing a series of large-amplitude undulations. As the undulations pass down the body, they push against the water, generating the thrust which moves the fish forward. The power generated by the muscle is converted to hydrodynamic thrust all along the body. This thrust is generated on both sides of the body at any instant, because there is more than one undulatory wave passing down the body at a given time. Fish with a tailblade swim in a different way. In

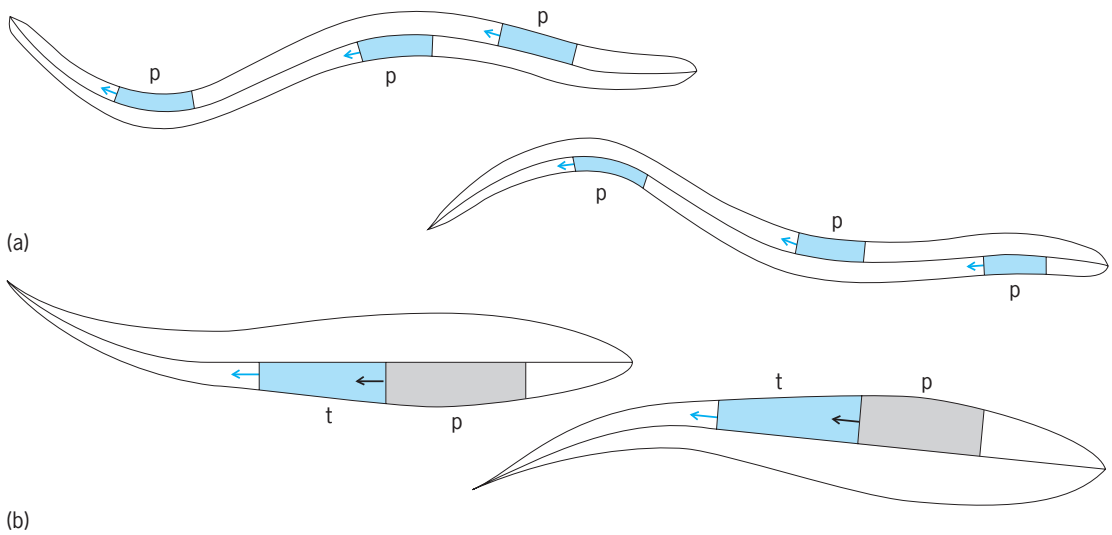

Fig. 10. Function of the myotomal muscle. (*a*) In anguilliform swimmers, waves of muscle contraction pass alternately down the sides of the body, throwing it into a series of large-amplitude undulations pushing against the water, generating power (p) thrust. (*b*) In carangiform swimming, one undulatory wave passes down the body at a given instant; thrust is generated alternately on the two sides and all the thrust is generated at the tailblade. The myotomal muscle generates power (p), which is transmitted down the fish toward the tail by muscle which acts temporarily as a tendon (t).

carangiform swimming, as seen in the mackerel and other fish with similar body forms, all of the thrust is generated at the tailblade.

To swim, the anguilliform and carangiform fish must use their myotomal muscle in different ways. In the eel, myotomal muscle functions in the same way all along the body. As the wave of contraction passes down the body, the muscle generates power, which is passed directly to the water to generate hydrodynamic thrust. In contrast, the function of the active muscle in carangiform swimmers varies along the body. As the wave of contraction starts near the head, myotomal muscle generated power. However, this is not transmitted directly to the water, but down the fish toward the tail by muscle which acts temporarily as a tendon. The muscle transmitting this power is active, but is being stretched because of the complex interaction between the motions of the fish and the water. Under these conditions, muscle tissue is very stiff and is therefore a good power transmission element. As the active region passes down the body, muscle which previously transmitted power may generate power, which is transmitted to the tail by muscle located still further down the body. When wave of power is generated, it passes down the body, and it is transmitted by more posterior, stiffened muscle to the tail, which generates the thrust.

Given the enormous diversity of fish body forms and swimming modes, there are many variations on these basic patterns of muscle use. The scup, for example, is a fish with a short, deep body with high dorsal and ventral fins. When it swims, there is only 0.65 of an undulatory wave on the body at a given instant. Since it has a large tail blade, it might be expected to swim like a mackerel, generating all of its thrust at the tailblade. However, it appears to have some of the characteristics of the eel, and may pass power directly to the water all along its flattened body and through its high fins. A species of sculpin, a bottom-dwelling predator, shows that muscle recruitment patterns vary with the mode of swimming. Fast starts and slow turns, for example, require uses of the myotomal muscle.

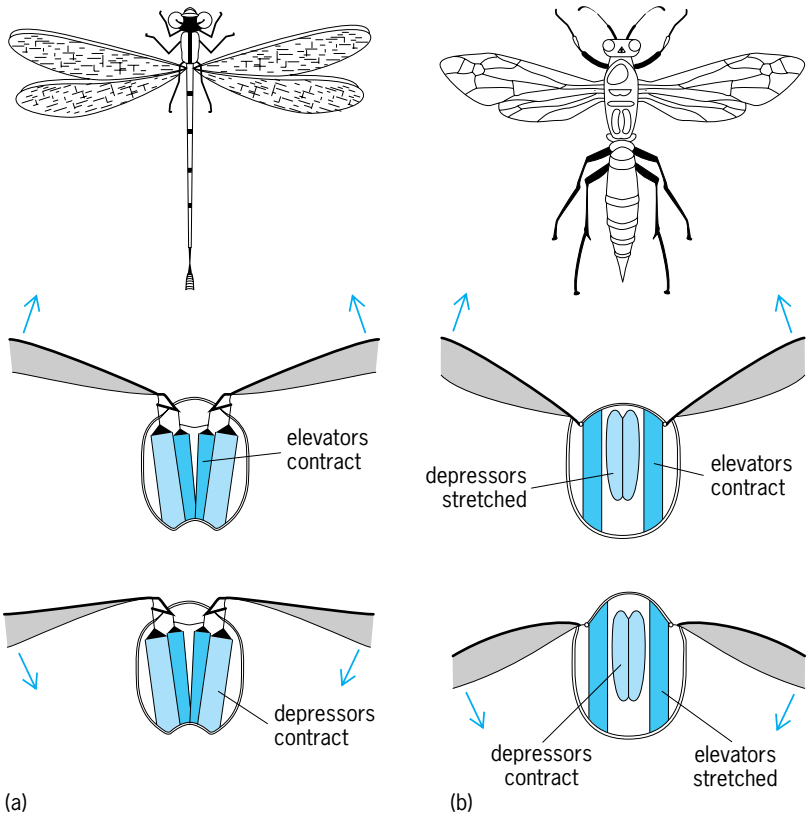

Fig. 11. Two types of insect musculature. (*a*) Direct flight muscle. (*b*) Indirect flight muscle. (*After R. Eckert, D. Randall and G. Augustine, Animal Physiology Mechanisms and Adaptations, 3d ed., W. H. Freeman, 1988*).

Insect flight muscle. Insect flight muscle is divided on structural and physiological grounds into

synchronous and asynchronous types, and on the basis of its mechanical operation within the insect into direct and indirect flight muscles. There is a large degree of structural diversity between types of synchronous muscle, but the functional distinction between it and asynchronous muscle is far more fundamental. The fibers of all insect flight muscles are rich in mitochondria and rely on aerobic metabolism despite the high energy demands of flight.

Direct flight muscle. This type of muscle is present in what are usually regarded as primitive insects, for example, the dragonflies (**Fig. 11***a*). One end of each direct flight muscle is attached to the base of the wing, and the other end to the inside of the thorax. Contraction of the flight muscles, therefore, drives the wings directly. The wingbeat frequencies of insects with direct flight muscles are typically less than 100 Hz. All insects with direct flight muscles have synchronous muscle.

Indirect flight muscle. Indirect muscles attach to the inside of the thorax rather than the wing. Muscle contraction deforms the thorax and, through a complex hinge, moves the wings up and down. The indirect elevator muscles run between the roof and floor, and contraction pulls the roof downward and raises the wings. Contraction of the depressors, which run from front to back, buckles the thorax, raises the roof, and lowers the wings (Fig. 11*b*). The mass of wings, the aerodynamic forces acting upon them, and the elasticity of the thorax act as a resonant system that enables some insects to operate with wingbeat frequencies up to 1000 Hz. In these insects, both synchronous and asynchronous indirect flight muscles are found, but only those with asynchronous muscle can achieve the high wingbeat frequencies.

Synchronous muscle. Synchronous muscle is characterized by an equal and simultaneous neural input and muscular contraction, each mechanical contraction being caused by a burst of neural activity. The rhythm of the wingbeat is thus neurogenic in origin. Insects with this type of muscle rarely have wingbeat frequencies that exceed 100 Hz. This is a limitation imposed by the neurogenic nature, and the time required to activate and relax the muscle in each wingbeat. Some highly modified synchronous flight muscles are found, for example, in the singing muscles of cicadas.

Asynchronous muscle. Very high wingbeat frequencies are found in some species of insects (for example, the wingbeat frequency of some midges is of the order of 1000 Hz). These high frequencies are found in insects that possess asynchronous flight muscles, that is, muscles where the number of nerve impulses to the muscle is much lower than the high wingbeat frequency. The wingbeat frequency of blowfly is approximately 120 Hz, but the frequency of neural input is only 3 Hz. The nervous input to these muscles facilitates rather than controls the frequency of contraction.

The rhythmic contraction of asynchronous muscle that drives the wings results from a unique resonant coupling between the flight muscles and the elastic thorax. When the active muscle is stretched by the recoil of the elastic thorax, it develops force after a short delay. However, this delay is long enough to allow a complete recoil of the thorax, before delayed force development deforms it. This properly of a stretch-induced delayed force development is due to a unique myosinfilament structure. The depressors and elevators work in opposition to oscillate the thorax and beat the wings. The time course of delayed force development enables the muscles and thorax to resonate at the same natural frequency. Because the system is operating at its resonant frequency, a minimum amount of energy is required to maintain the oscillations once the wingbeating has started. Because the muscle does not need to be repeatedly switched on and off, and because the thorax is relatively stiff, the system can operate at high frequencies.

Body size. If evolution has resulted in optimization of muscle function, size-dependent changes in these processes should be reflected in the properties of the muscles which drive them. The main function of muscle is to generate power, normally cyclically and repetitively, for example, contracting and relaxating of the diaphragm during breathing, or of the leg extensor muscles in walking. The body demands the greatest effort from the muscles at particular frequencies, for example, the stride frequency of a sprint. Therefore, under normal circumstances the muscles should perform best over those frequencies. *See* MUSCLE; RESPIRATORY SYSTEM.

Iain S. Young; John D. Altringham

Embryology

The muscles are derived from mesoderm, the middle germ layer. The exceptions are the sphincter and dilator muscles of the iris and the myoepithelial cells of the sweat and mammary glands, which are derived from ectoderm. The embryonic mesoderm that differentiates into muscle tissues includes the dorsal mesoderm, head mesenchyme, intermediate mesoderm, and lateral mesoderm (**Fig. 12**). The dorsal mesoderm that condenses into bilateral columns adjacent to the neural tube forms the segmentally arranged myotomes. Most intrinsic voluntary muscles of the neck and trunk are differentiated from these myotomes. Some voluntary head musculature (muscles of eye and tongue) and the limb musculature are derived from myotomes in the lower vertebrates; the limb muscles in higher vertebrates are mainly derived from lateral plate mesoderm. In tetrapods the myoblast component of all truncal and appendicular muscles originate from somites (somitic origin). The voluntary muscles of the branchial (visceral) arches of the head and neck are derived directly from head mesenchyme.

Cardiac muscle is derived from the splanchnic mesoderm. The involuntary musculature is differentiated from the intermediate mesoderm (mesomere, urogenital mesoderm, or nephrotomic mesoderm) and the lateral mesoderm (splanchnic mesoderm and hypomere). The intermediate mesoderm differentiates into much of the urogenital system, and the lateral mesoderm differentiates into the vascular,

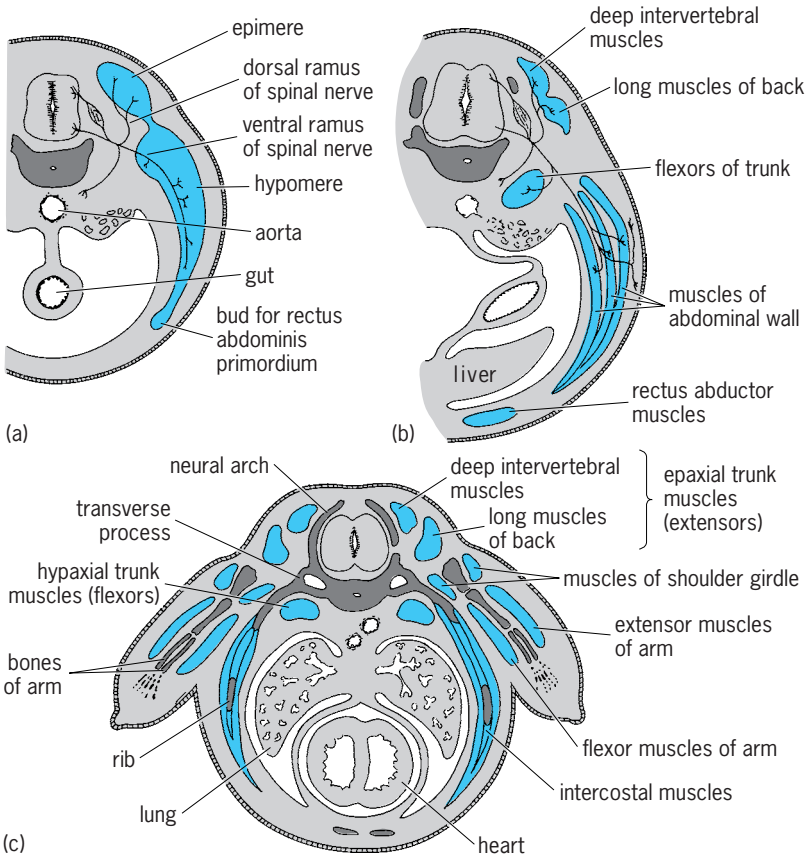

Fig. 12. Differentiation of myomeres and the establishment of the primordia of various muscle groups. (*a*) Early separation of the myomere into an epimere, supplied by the dorsal ramus of the spinal nerve of the level, and a hypomere, supplied by the ventral ramus of the nerve. (*b*) Later stage in the differentiation of muscle primordia at abdominal levels. (*c*) Primordial muscle masses at the level of the arm buds. (*After C. E. Corliss, Patten's Human Embryology, McGraw-Hill, 1976*)

digestive, and respiratory systems and related structures. *See* EMBRYONIC DIFFERENTIATION; EMBRYONIC INDUCTION.

Differentiation of striated muscles. Striated muscles differentiate from the myotomes of the somites and from mesenchyme of nonmyotomic origin. Muscle cells, fascial cells, tendon cells, and aponeurotic cells are derived from these structures. The premuscle cells (myoblasts) migrate as the organism develops. When the myotome is adjacent to the neural tube, it receives its initial innervation which is retained during subsequent migration of the developing muscles. The innervation of the primordial muscle masses and the retention of this innervation during development are significant as a means of determining the homology of muscles of different species. The nerves are probably not concerned with organizing these muscle masses, because muscles will develop without any innervation in certain monsters. *See* NERVOUS SYSTEM (VERTEBRATE).

The segmental pattern of the myotomes may be retained in the adult, as in the intercostal muscles of the thorax of mammals and the trunk muscles of fishes. The pattern may be modified as in the flat muscles of the mammalian abdominal wall. The segmental derivation of the muscle masses may be masked by migration as in the eye and tongue muscles, by fusion of muscles from several segments as in the rectus abdominis muscle (the product of the fusion of muscle masses from successive myotomes), by the splitting of muscles into layers as in the flat abdominal muscles, and by fusion and realignment of muscles as in the back musculature of higher vertebrates where muscle fascicles extend through as many as six segments. During metamorphosis in frogs, the segmental patterns of many muscles in the tadpole are altered by the migration, fusion, and splitting of muscle masses to form many muscles in the adult frog.

Axial musculature. In all vertebrates, the muscles of the neck, trunk, and tail are derived from myotomes, although claims have been made that some myoblasts differentiate in dermatomes. The myotomes develop and migrate laterally and ventrally as the epimere to form the dorsal or epaxial muscle mass (dorsal and lateral back muscles), and as the hypomere to form the ventral or hypaxial muscle mass (lateral and ventral muscles) [Fig. 12]. The subsequent fate of this migration is said to differ in different animals. In the aquatic vertebrates, whose primary means of locomotion is swimming, the musculature is oriented to produce undulatory motion. Each myotome differentiates into a muscle mass which forms a band from the back to the belly. The muscle cells within each band are directed cephalocaudally. The epaxial musculature (muscle mass innervated by the dorsal ramus of a spinal nerve and located dorsal to the vertebral column) is approximately equal in size to the hypaxial musculature (muscle mass innervated by the ventral ramus of a spinal nerve and located ventral to the vertebral column). In most terrestrial animals the bilateral appendages assume primacy in locomotion. In these animals, the myotomes differentiate into the body-wall musculature pattern in which the epaxial musculature is restricted to the back and the hypaxial musculature is present on the lateral and ventral aspect of the body wall. In animals adapted for aerial locomotion, the primitive segmental pattern is modified. The epaxial musculature is greatly reduced in the trunk but is well developed in the nuchal region. The hypaxial musculature is mainly concentrated in the pectoral musculature.

Fate of myotomes. The fate of the myotomes during their differentiation into the axial muscles differs in the various vertebrates (**Fig. 13**). In the fishes the embryonic pattern is retained because the epaxial and hypaxial muscles are relatively equal in size. In the amphibians, two adult patterns are developed. In aquatic amphibians, such as *Necturus*, the embryonic segmental pattern is retained. In terrestrial amphibians, such as the frog, the myotomic derivatives are modified during metamorphosis by migration, fusion, and splitting. As a result, the embryonic pattern is modified as a functional adaptation to life on land. In the other terrestrial animals, reptiles and mammals, the embryonic myotomic segmentation is altered further than in the terrestrial amphibians. The epaxial muscles form the erector spinae back muscles. Some hypaxial muscles retain vestiges of the segmental patterns (intercostal muscles of

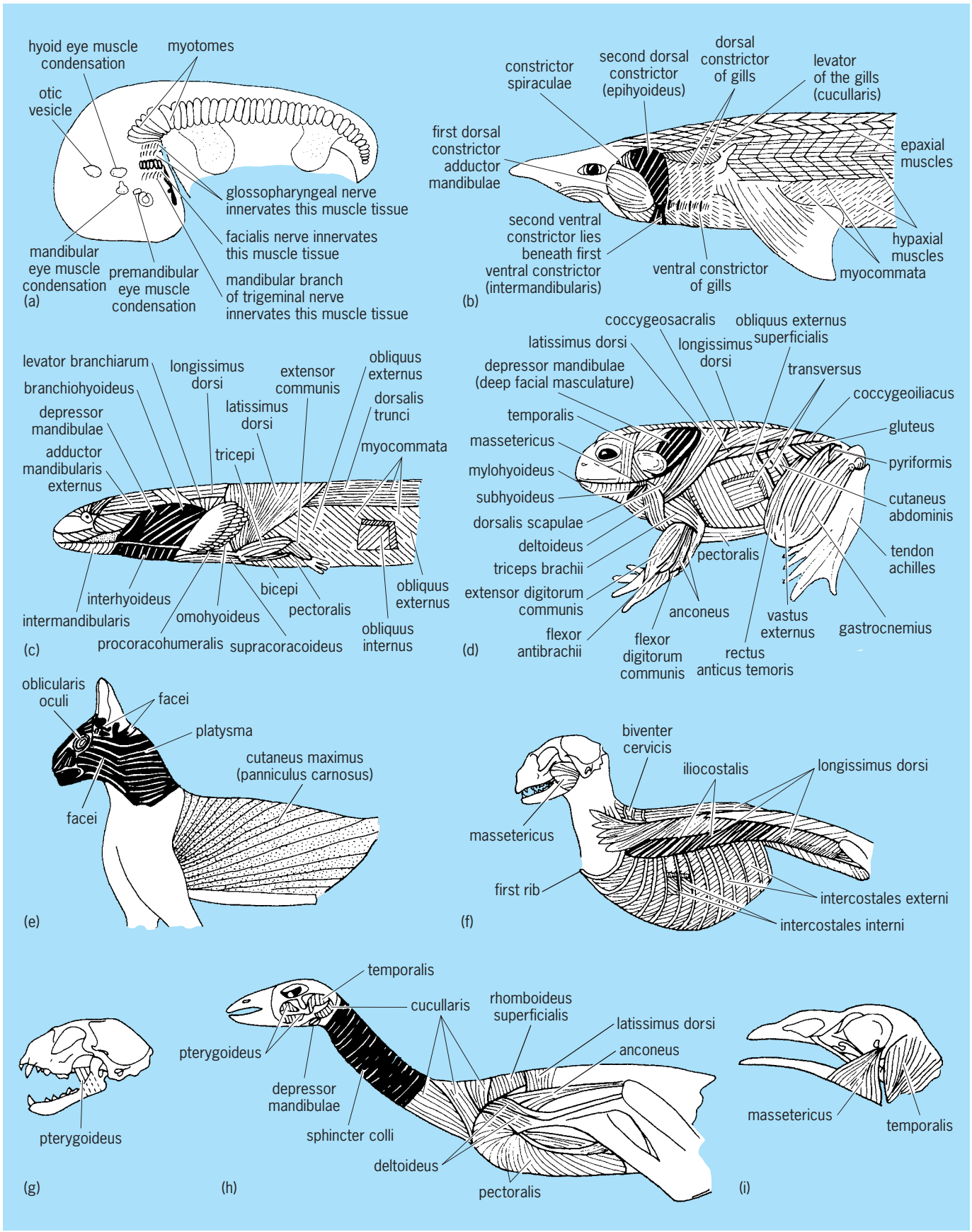

Fig. 13. Development of muscles from the visceral arches and from the myotomes in various vertebrates. (*a*) Basic areas of the embryo from which voluntary muscles develop. (*b*) Shark. (*c*) *Necturus*. (*d*) Frog. (*e–g*) Cat. (*h, i*) Goose. (*After O. E. Nelsen*, *Comparative Embryology of the Vertebrates*, *Blakiston*, *1953*)

abdomen), whereas other muscles are the products of migration, fusion, and splitting during development (flat abdominal muscles). The muscle patterns are altered more drastically in birds.

The tail bud mesoderm differentiates into myotomes in tailed animals that exhibit lateral movements. In these forms—fish, tailed amphibians, crocodiles, and whales—the myotomes differentiate into well-developed epaxial muscles and hypaxial muscles that retain their embryonic metamerism. The musculature of animals with prehensile tails and tails adapted for grasping and wagging movements is partially the result of the migration of myotomes from the hindlimb area. Of significance in mammals is the derivation of the diaphragm from portions of the neck myotomes and their subsequent migration through the neck and thorax to the thoracoabdominal boundary.

Axial muscles. Each axial muscle in the vertebrates is derived from one or more specific myotomes which are numbered according to the vertebral level (cervical, thoracic, lumbar, sacral, caudal, or coccygeal regions) at which the myotome first differentiated. Because the innervation of each myotome occurs during early development, the number in a region of each spinal nerve is identical with that associated with each myotome. In the different classes of animals, the regional number of a myotome and the spinal nerve varies because the number of vertebrae in any region is not similar in the vertebrate classes. Because of this variability, only the segmental numbers of the myotomes in relation to the adult muscles in humans, as an example, will be presented below.

Each myotome expands by growing ventrally to form two myomeres: the epimere and hypomere. The epimere, innervated by the dorsal ramus of a spinal nerve, subdivides into a dorsal segment, which gives rise to the deep intervertebral muscles of the neck and back, and a more ventral segment, which gives rise to the long muscles of the neck and back (Fig. 12). The hypomere, innervated by the ventral ramus of a spinal nerve, differentiates into the flexor muscles of the trunk (hypaxial muscles); intercostal muscles the thorax; muscles of the abdominal wall; and rectus thoracic and abdominal muscles of the ventral body wall (Fig. 12).

Epaxial muscles. The epaxial muscles of humans, the intrinsic extensor muscles of the back, are derived from the dorsal portions of 29 myotomes—the first cervical segment through the fourth sacral segments inclusive. These include 8 cervical, 12 thoracic, 5 lumbar, and 4 sacral segments. The hypaxial cervical musculature is derived from the ventral portions of the 8 cervical myotomes. The myotomes differentiate into a prevertebral portion (immediately in front of the vertebral column), a lateral sheet, and a ventral or rectus column. As a result, the prevertebral portion forms the prevertebral muscles; the lateral sheet forms the scalene muscles; and the rectus column forms the geniohyoid and infrahyoid muscles. The diaphragm is derived from the hypaxial division of the third through fifth cervical myotomes from whence it migrates.

Hypaxial muscles. In humans, the muscles of the ventral and lateral thoracolumbar wall are derived from the hypaxial divisions of the first thoracic through the first lumbar myotomes inclusive. These myotomes differentiate into a main lateral sheet and a rectus column (ventral edge). The intercostal muscles, the oblique abdominal muscles, and the transversus abdominis muscle develop from the main lateral sheet, whereas the rectus abdominis muscle develops from the rectus column. The quadratus lumborum muscle of the posterior abdominal region is derived from the hypaxial divisions of the first through fifth lumbar myotomes. Although direct evidence is difficult to observe, it is probable that the muscles of the pelvic diaphragm (the muscular floor of pelvis, including the coccygeus and levator ani muscles) are derived from the hypaxial divisions of the last four sacral and all coccygeal myotomes.

In the chick, the skeletal muscles of the dorsal and dorsolateral trunk are derived from myotomic mesoderm, whereas the hypaxial muscles differentiate from the somatic mesoderm of the lateral plate.

Paired appendages. Theoretically the paired appendages were derived originally either from gill arches (gill-arch theory) or from fin folds (lateral-fold theory). On the basis of the embryonic development of the fin musculature, the lateral-fold theory is favored because, in fish, the fin musculature is differentiated from the myotomes of somites. Muscle buds of potential muscle cells differentiate at the lower edges of the myotomes and, as slips of tissue, invade the mesenchyme of each limb bud as the dorsal and the ventral premuscle masses of myoblasts. The dorsal premuscle mass differentiates into the dorsal, elevator, and extensor muscles, and the ventral premuscle mass into the ventral, depressor, and adductor muscles of the fins. These muscles are innervated by spinal nerves. Apparently the lateral plate mesoderm is not capable of forming the fin musculature.

In the tetrapods, including the amphibians, reptiles, birds, and mammals, the muscles of the limbs assume a relatively large bulk and a complexity of organization to cope with locomotion on land. The embryonic derivation of most of the myoblasts, which form this musculature in the tetrapods, differs from that in the fishes: much of the limb musculature of tetrapods is said to be derived in place from lateral plate mesoderm and only some from the myotomic mesoderm of somites and pharyngeal mesoderm. This mesenchyme is initially found surrounding the developing skeletal elements (future bones). The standard version is that this mesenchyme does not originate from the myotomic regions of the somites.

On the basis of their mesodermal embryonic precursors and their innervation, the appendicular musculature may be classified into three groups: (1) Mammalian muscles, derived from the mesoderm of the posterodorsal pharyngeal region, include the trapezius and the sternomastoid muscles, which differentiate from the mesoderm of the last branchial arch. These muscles have attachments which extend from the neurocranium and cervical vertebrae to the

proximal bones of the skeleton of the forelimb. They are innervated by the spinal accessory cranial nerve. (2) Muscles derived from myotomic mesoderm of somites include, among others, the rhomboids, pectorals, and serratus muscles of the forelimb and the quadratus lumborum and psoas muscles of the hindlimb. They are the true metameric derivatives of the embryonic myotomes. These muscles have attachments which extend from the vertebrae and ribs to the appendages; they are innervated by the anterior rami of the segmental spinal nerves proximal to where these rami form the brachial plexus of the forelimb and the lumbosacral plexus of the hindlimb. (3) Muscles derived from the nonsegmental core of the limb bud include all the intrinsic limb muscles, muscles with attachments located wholly within the limb (**Fig. 14**). They differentiate within each limb bud and are innervated by the nerves of the brachial plexus and of the lumbosacral plexus. In general, the muscles differentiating from the dorsal premuscle mass of the limb bud become the extensor muscles of the limb, those from the ventral premuscle mass become the flexor muscles, those from the dorsal premuscle mass near the trunk become the abductor muscles, and those from the ventral premuscle mass near the trunk become the adductor muscles.

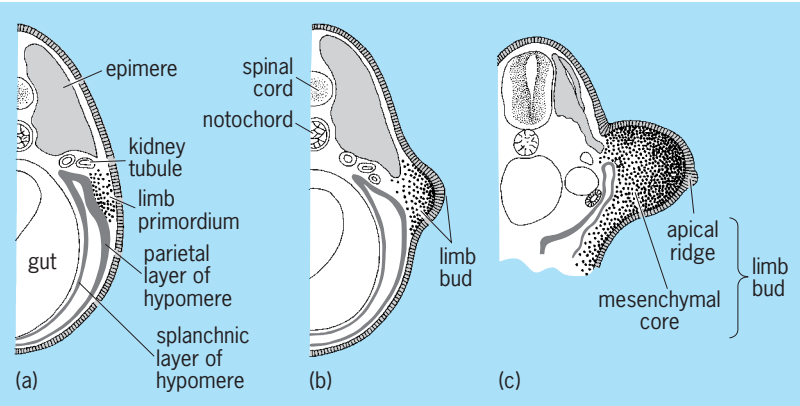

Fig. 14. Diagrams of the origin of a limb. (*a*, *b*) The development of a limb bud of an amphibian. (c) Limb bud of a chick embryo. (*After T. Torrey*, *Morphogenesis of the Vertebrates*, *2d ed.*, *1967*)

Head and visceral arch musculature. The muscles of the head and visceral arches are derived from myotomic mesoderm and from nonmyotomic mesoderm (**Fig. 15**). The extraocular muscles of the eye and the tongue muscles are differentiated from myotomes or from a mesoderm that phylogenetically was originally derived from myotomes. The visceral arch musculature in all vertebrates is derived from mesenchyme of nonmyotomic origin. Some facial muscles are said to originate from neural crest cells.

The extraocular muscles of sharks are derived from the preotic somites of the premandibular myotomes (cranial nerve III), the mandibular myotome (cranial nerve IV), and the hyoid myotome (cranial nerve VI). The extrinsic muscles of the eye include six which are found in all vertebrates: the superior rectus, internal (anterior) rectus, inferior rectus, and inferior oblique muscles, which are innervated by cranial nerve III; the superior oblique muscle, which is innervated by cranial nerve IV, and the external (posterior or lateral) rectus muscle, which is innervated by cranial nerve VI. In addition, the retractor oculi of many mammals and the quadratus muscle and pyramidalis muscle of birds are in this category. The tongue musculature in the sharks develops from six postotic myotomes that migrate ventrally to the hypobranchial region. In higher vertebrates three postotic (occipital) myotomes appear to provide the mesodermal source of this musculature which is innervated in all vertebrates by cranial nerve XII (hypoglossal). Because direct myotomic origin of the extraocular muscles and the tongue muscles is difficult to demonstrate in the higher vertebrates, the literature disagrees.

The mesoderm of the branchial (gill) arches is derived from head mesoderm which develops in place and not from any myotome. The first branchial (mandibular) arch mesoderm differentiates into the muscles of mastication that are innervated by cranial nerve V (trigeminal). The muscles derived from this mesoderm in the fishes are the mandibular adductor muscle and the first ventral constrictor muscle; in the amphibians the temporal, masseter, pterygoid, and mylohyoid muscles; in birds the pterygotemporal, temporal, and digastric muscles; and in mammals the muscles of mastication (temporal, masseter, and pterygoid muscles), mylohyoid muscle, anterior belly of the digastric muscle, tensor palatini muscle, and tensor tympani muscle. The second branchial (hyoid) arch mesoderm differentiates into those muscles innervated by cranial nerve VII (facial). The muscles derived from this mesoderm in fishes are the hyoid gill arch muscles; in amphibians, the subhyoid and mandibular depressor muscles; in birds the sphincters of the neck and the mandibular depressor muscles; and in mammals the muscles of facial expression and other muscles such as the stylohyoid muscle, stapedius muscle, and the posterior belly of the digastric muscle. The mesodermal derivatives of this arch in mammals migrate to the scalp (occipitofrontalis muscle), ear region (auricular muscle), neck (platysma), and the face (orbicularis oculi, orbicularis oris, and others), collectively called the muscles of facial expression. The third visceral (first branchial arch) arch mesoderm differentiates into those muscles innervated by cranial nerve IX (glossopharyngeal). The muscles derived from this mesoderm in fishes are the gill constrictor muscles of this arch and in the higher vertebrates (mammals) the stylopharyngeus muscle and the upper constrictors of the pharynx. The mesoderm of the last three visceral arches (second, third, and fourth branchial arches) differentiates into the muscles innervated by cranial nerve X (vagus). These are the gill constrictor muscles in the fishes and the lower pharyngeal constrictor and laryngeal muscles in the higher vertebrates.

The sternocleidomastoid muscles and trapezius muscles of mammals are innervated by the spinal

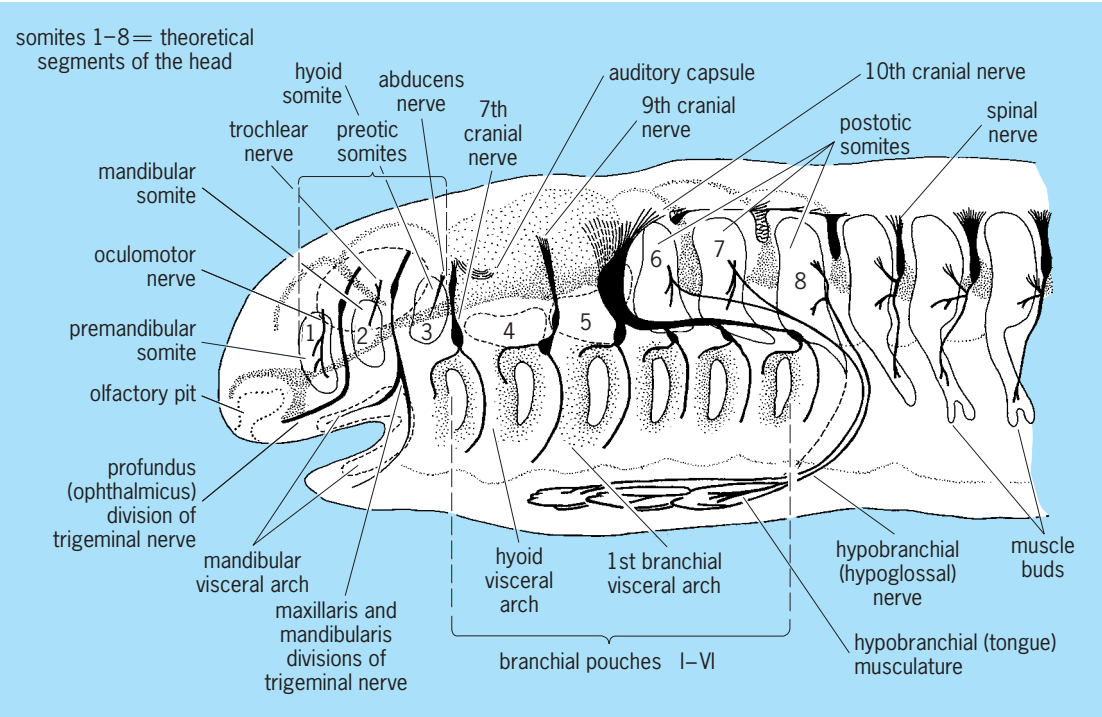

Fig. 15. Basic plan of vertebrate head. (*After O. E. Nelsen, Comparative Embryology of the Vertebrates, Blakiston, 1953*)

cord division of cranial nerve XI (spinal accessory). These muscles are derived either from the mesoderm of the last visceral arch or from postotic myotomes.

Voluntary skin muscles. The skin muscles, the voluntary muscles that move the skin, are divided into two groups, the muscles of facial expression innervated by cranial nerve VII, and the panniculus carnosus of the body-wall skin innervated by the anterior (ventral) thoracic nerves. The panniculus carnosus is derived from mesodermal cells that normally form the pectoral muscles. This muscle, found in such animals as the porcupine, dog, cat, horse, and guinea pig, may have both its origin and insertion in the skin of some species or its origin on the greater tuberosity of the humerus and its insertion in the fascia of the skin of the back and thigh of other species.

Involuntary muscles. Involuntary muscles arise independently of the segmental myotomes and visceral arches. The visceral mesoderm differentiates into the mesenchyme that forms the smooth muscles of the digestive system, respiratory system, and many blood vessels. The heart also differentiates from this mesoderm. The mesenchyme of the somatopleuric mesoderm of the body wall, head, and limb buds differentiates into the smooth muscles of the blood vessels of these regions. Mesenchyme, whatever its origin, is a potential source of smooth muscle. The smooth muscles of many organs develop in place from mesenchymal cells. The smooth muscles of the structures of the urinary system and the genital systems are derived from mesenchymal cells of the intermediate mesoderm (nephrotomic mesoderm). The smooth muscles of ectodermal origin are the dilator muscle of the iris and the myoepithelial cells of the ducts of sweat glands.

Histogenesis of Muscle

Embryonic muscle cells (myoblasts) are derived from mesenchymal cells of mesodermal origin.

Voluntary muscle. The myoblasts of voluntary muscles of either myotomic or nonmyotomic origin are mononucleated spindle-shaped cells with clear cytoplasm. Embryonic stem cells from the mesoderm of the myotomes divide to form bipolar cells. These mononucleated myoblasts cease dividing and develop specialized cell surfaces that render them capable of fusing to form multinucleated elongated myotubes. Myosin and messenger RNA required for its synthesis are present in the myoblasts. After fusion the synthesis of myosin rapidly increases. This is followed by the appearance of contractile filaments, receptor sites for the neurotransmitter (acetylcholine), T-tubular system, and sarcoplasmic reticulum.

The first myofibrils are sparse and coarse. Cross striations are formed almost as soon as the myofibrils are visible. The elongated myotubes have centrally located nuclei and peripherally located myofibrils in the early stages of differentiation. Later, when the myofibrils fill the cells, the nuclei become peripherally located adjacent to the sarcolemma. In humans, at the time of birth, the muscle cells resemble those of the adult except that the nuclei are rounder, the myofibrils are more slender, and the cross striations are less prominent. The multinuclear striated muscle cells are the result of nuclear mitosis which is unaccompanied by cytoplasmic divisions. The postnatal growth of muscle cells is by hypertrophy.

Involuntary muscle. Smooth muscles arise from mesenchymal cells. During embryonic life the cells migrate and concentrate in the vicinity of the epithelial linings of the hollow organs. These myoblasts elongate and orient themselves as in the adult organ. Myofibrils (contractile elements) are visible in the early stages and become more numerous during later development. The iridic muscles (sphincter and dilator pupillae) and myoepithelial cells of sweat and mammary glands are derived from ectodermal cells.

Cardiac muscle. Cardiac muscle develops from the splanchnic mesoderm of the heart tube. The myoblasts adhere to each other, but unlike skeletal muscle, the plasma membranes do not disintegrate—rather the sites of adhesion give rise to the intercalated discs (gap junctions). As in skeletal muscles, myofibrils and the other structural muscle elements differentiate early. Charles R. Noback

Bibliography. N. D. Agnish, Possible role of somites in developing mouse limbs, *in vitro*, *Anat. Rec.*, 184:340–341, 1976; R. McN. Alexander, *Exploring Biomechanics: Animals in Motion*, 1992; B. I. Balinsky, *An Introduction to Embryology*, 5th ed., 1981; A. Chevalier, Role of the somitic mesoderm in the development of the thorax in bird embryos, II: Origin of thoracic and appendicular musculature, *J. Embryol. Exp. Morphol.*, 49:73–88, 1979; K. L. Moore, *The Developing Human*, 6th ed., 1998; R. O'Rahilly and F. Muller, *Developing Stages in Human Embryos*, 1987; A. S. Romer, *The Vertebrate Body*, 6th ed., 1986; T. W. Torrey and A. Feduccia, *Morphogenesis of the Vertebrates*, 4th ed., 1979; W. F. Walker, Jr., *Vertebrate Dissection*, 9th ed., 2000.

Muscular system disorders

Disorders affecting skeletal (voluntary) muscle. Every skeletal muscle consists of a large number of muscle fibers, each of which is a multinucleated cell containing the usual types of cellular inclusions, but also containing the myofibrils, which are the actual contractile elements. The myofibrils have a banded structure which gives the muscle fiber its characteristic striated appearance on histologic examination. Muscle fibers have been separated into two main groups based on their histochemical reaction: type 1 fibers, which are rich in mitochondria and oxidative activity, and type 2 fibers. *See* MUSCLE.

The normal functioning of the skeletal muscle is dependent not only on the integrity of the muscle fibers themselves, but also on that of the motor cortex, the pyramidal tract, and the extrapyramidal system (including the cerebellum). It also depends on innervation by the motoneurons of the brainstem and the spinal cord. In addition, the proper functioning of the other organ systems, such as the endocrine system, and variations in the concentration of various electrolytes may also affect muscle function. *See* NERVOUS SYSTEM (VERTEBRATE).

Damage to the motor cortex or the pyramidal tract produces the type of weakness seen in humans after a stroke or spinal cord injury. Although the paralyzed limb may initially be flaccid (hypotonic), spasticity (hypertonia) eventually develops. Despite the weakness, muscle atrophy is usually not striking. When the extrapyramidal system or the cerebellum is the site of damage, instead of weakness there are uncontrolled movements, difficulty with coordination, or both. In either of these situations there is no characteristic change in the muscle, either grossly or microscopically. At most, atrophy of type 2 fibers is seen.

Motoneuron damage. With damage to the spinal motoneuron or its axon, there is flaccid weakness of the muscle with proportionate wasting. Direct involvement of the spinal motoneuron was typically seen in poliomyelitis, but is now seen more commonly in the progressive spinomuscular atrophies of infancy and childhood, and in amyotrophic lateral sclerosis in adults. Spontaneous twitching of groups of muscle fibers (fasciculation) innervated by the same motoneuron (motor unit) is frequently seen in these disorders.

The axons of these motoneurons may also be damaged as they exit through the spinal nerve root foramina (nerve root compression, radiculopathy) or in the peripheral nerve (peripheral neuropathy). When muscle weakness is due to a radiculopathy, it is restricted to those muscles innervated by that nerve root. In peripheral neuropathy, muscle involvement is usually more diffuse and symmetrical or it occurs throughout the distribution of a particular nerve.

In all the following conditions, in which the motoneurons or their axons are damaged, the affected muscles undergo atrophy. Microscopically, marked variation in the size of individual muscle fibers is seen, in which the atrophic fibers of similar sizes occur in groups (neurogenic atrophy). Myosin adenosinetriphosphatase stain shows that both type 1 and type 2 fibers are involved in the atrophic process.

Poliomyelitis. This is an acute viral infection in which there is selective damage of the motoneurons, resulting in a flaccid paralysis of the muscles innervated by the affected nerve cells. Because the motoneuron damage is patchy and asymmetrical, the distribution of weakness is also quite patchy. Since poliomyelitis is an acute illness, the weakness and paralysis are nonprogressive. *See* POLIOMYELITIS.

Spinomuscular atrophies. These are progressive diseases of unknown cause and may occur in infancy (Werdnig-Hoffmann disease) or later in childhood (Kugelberg-Wielander disease). Both diseases result from degeneration of the motoneurons and appear to be inherited in an autosomal recessive pattern. In the infantile form, the baby is often floppy from birth with generalized weakness, a poor cry, and difficulty in sucking and breathing. Many children succumb in early childhood. In the later childhood forms, the rate of progression is slower and the outlook better. In these children weakness is more marked in the proximal muscles of the limbs.

Amyotrophic lateral sclerosis (ALS). This is a progressive disease of unknown cause, in which both the brainstem and spinal motoneurons, as well as the corticospinal tracts, undergo degeneration. This is a

relatively rapid progressive disease, with death usually occurring within 3 years of diagnosis due to the swallowing difficulty and respiratory failure. Although some cases of ALS appear to be inherited, most cases occur sporadically.

Neuromuscular junction. The most common example of disease at the neuromuscular junction is myasthenia gravis. Other, less common diseases are the myasthenic Eaton-Lambert syndrome and botulism. *See* MYASTHENIA GRAVIS.

In the Eaton-Lambert syndrome limb muscles are most frequently involved, and weakness is the main complaint. However, the muscle strength increases with repeated contractions, and, on repetitive nerve stimulation, there is a progressive increase in the amplitude of the evoked potential on electromyography. Thus, although this is called the myasthenic syndrome, the response to exercise and to repetitive nerve stimulation is opposite to that seen in myasthenia gravis. The neuromuscular defect appears to be due to a decrease in the number of acetylcholine quanta released by a nerve impulse. About 80% of patients with the Eaton-Lambert syndrome have been shown to have a small cell bronchogenic carcinoma. No consistent histologic abnormality has been described in the muscle, but electron microscopy has shown a marked increase in the number and complexity of the secondary synaptic clefts, with a resultant elongation of the postsynaptic membrane. This change also contrasts with the markedly simplified pattern seen in myasthenia gravis.

Botulism is another condition affecting the motor end plate and results from the ingestion of botulinum toxin, most commonly found in improperly canned food. The botulinum toxins block neuromuscular transmission in cholinergic nerve endings, probably by inhibiting the release of acetylcholine. The toxin produces rapidly spreading paralysis of voluntary muscles, including those of swallowing and respiration, leading to death due to respiratory failure. If the patient does not die, full recovery can be expected after a prolonged convalescence. No definite microscopic abnormality has been described in the muscle. *See* ACETYLCHOLINE; BOTULISM; SYNAPTIC TRANSMISSION.

Myopathy. Abnormalities of the muscle itself (myopathy) obviously result in muscle weakness, and muscle diseases fall into two large groups: those with a genetic basis and those which are nongenetic.

Genetic diseases. Congenital myopathies and the muscular dystrophies constitute the genetic muscle diseases. Congenital myopathies are characterized by a generalized weakness which is present at birth. The weakness is usually not progressive and often improves with time. In general, the different types of congenital myopathies cannot be differentiated on clinical grounds, but are distinguished on the basis of morphologic changes seen in muscle biopsy. Muscle fiber necrosis, a characteristic of the muscular dystrophies, is not seen in the congenital myopathies. Instead there are structural abnormalities, such as collections of rodlike particles (nemaline myopathy), central portions of muscle fibers devoid of mitochrondria (central core disease), or centrally placed nuclei in the majority of fibers (centronuclear myopathy). As more infants with muscle weakness are studied by muscle biopsy, other structural abnormalities will probably be identified. *See* MUSCULAR DYSTROPHY.

Nongenetic or acquired diseases. These are all characterized by rapidly progressive weakness of the proximal muscles of the limbs, and so resemble the limb-girdle form of muscular dystrophy in the distribution of weakness. They are often included as inflammatory myopathies. Individuals with these disorders have difficulty arising from a recumbent or sitting position, in climbing stairs, and in lifting heavy objects onto a shelf. They also often have tender and painful muscles, may be febrile, and may have other manifestations of a systemic illness.

In many individuals, the muscle disease appears to be primary (polymyositis), with or without associated skin involvement (dermatomyositis). The cause of this group of muscle diseases is unknown, but they are thought to be due to an altered immune state which may result in direct damage to the muscle, or secondary damage to the capillaries. Such a possibility is also supported by the frequent association of acquired muscle disease with collagen vascular diseases, particularly rheumatoid arthritis, scleroderma, and Sjogren's syndrome.

Histologically, there is considerable variation in the size of the individual muscle fibers. There are large numbers of necrotic or degenerating muscle fibers being invaded by phagocytic cells, as well as small basophilic fibers with large prominent nuclei. The latter are thought to be regenerating cells. These features are qualitatively similar to those seen in the muscular dystrophies, but these changes are quantitatively much more extensive in polymyositis, in which the disease process progresses much more rapidly. In approximately 75% of these patients, there is perivascular infiltration by mononuclear inflammatory cells, mostly lymphocytes and plasma cells.

Acquired muscle diseases may also be secondary to carcinoma, and the pathologic changes in the muscle may be quite similar to what has been described in polymyositis and dermatomyositis Such an association with carcinoma has been noted in up to 20% of adults with polymyositis, and may be even higher in those with dermatomyositis.

Metabolic diseases. A number of metabolic diseases have been associated with muscle symptoms. These include thyroid diseases and certain endocrine diseases, particularly Cushing's syndrome, which are either spontaneous or secondary to therapeutically administered adrenocorticosteroid hormones. No specific histologic changes have been described in the muscle in these disorders.

Muscular symptoms are also seen with some of the glycogen storage diseases, in which specific enzymes necessary for the catabolism of carbohydrates are genetically absent. Muscle symptoms are prominent in Pompe's disease and McArdle's disease; both conditions are inherited in an autosomal recessive pattern. In Pompe's disease acid maltase is the deficient

enzyme. Glycogen accumulates in many organs of the body, including skeletal muscle, the heart, and the central nervous system. The excessive accumulation in the skeletal muscle is seen as a marked vacuolation of the muscle fibers. The child is often weak and floppy from birth and may die from cardiac failure, although milder forms of the disease may occasionally be seen. In McArdle's disease the enzyme myophosphorylase is missing, and the patient is unable to utilize glycogen as a source of energy. Glycogen storage is restricted to skeletal muscule and is seen as subsarcolemmal vacuoles in the muscle fiber. Symptoms are early muscular fatiguing and painful cramps after excessive activity. *See* METABOLIC DISORDERS.

Muscle weakness and paralysis may also be associated with alterations in the level of serum potassium. These are present with episodic or periodic muscle paralysis, and the individual is usually normal between these attacks. These diseases are all hereditary, and during the attack the serum potassium may be low (hypokalemic), elevated (hyperkalemic), or normal (normokalemic). Familial hypokalemic periodic paralysis is the most common form of potassium disorder. It is inherited as an autosomal dominant disease. The muscle fibers contain vacuoles which result from dilatations of the T-tubules and the sarcoplasmic reticulum. The mechanism for the depression of serum potassium during these attacks is unknown, but episodes of weakness and paralysis usually follow excessive exertion followed by rest, or after a heavy carbohydrate meal.

Myotonia. A delayed relaxation of the muscle after forceful contraction (myotonic) is another symptom of muscle disease. This phenomenon is a characteristic feature of myotonic dystrophy and of congenital myotonia. Congenital myotonia may occur as an autosomal dominant (Thomsen's disease) or as a recessive disorder. Myotonia is usually present from early life and is often associated with muscle hypertrophy, but muscle weakness is not a feature of this disease. The muscle shows no definite histologic changes.

Infections. Specific infections of the muscle are not very common and may be caused by bacterial, viral, or parasitic agents. The best known is trichinosis, an infestation by the pork tapeworm, *Trichinella spiralis*. The main muscle symptoms are pain and tenderness, but the systemic symptoms of fever, periorbital edema, and petechial hemorrhages may be more prominent. There is necrosis of muscle fibers and severe inflammation reaction with polymorphonuclear and eosinophilic leukocytes. The parasite can often be seen in the necrotic muscle fiber.

Tumors. Primary tumors of skeletal muscle are quite rare and are of two types: the benign rhabdomyoma and an extremely malignant rhabdomyosarcoma.

Injury. Traumatic injuries to the muscle may be due to laceration, blows, or prolonged pressure and result in focal areas of necrosis (rhabdomyolysis). If the injury is extensive, the muscle respiratory pigment, myoglobin, may be released into the circulation and excreted in the urine (myoglobinuria), with subsequent renal failure. Rhabdomyolysis with myoglobinuria may also follow rigorous running, jumping, or long marches without previous training. Following the exercise, painful swelling of the anterior muscles of the lower legs occurs, accompanied by myoglobinuria. Histologically, there is necrosis of muscle and later healing with fibrous scar formation. It is thought that swelling of these muscles, which are contained in a rather firm fibrous and bony compartment, interferes with circulation and results in subsequent necrosis. *See* MUSCULAR SYSTEM. S. Mark Sumi

Bibliography. M. Adachi, *Neuromuscular Diseases*, 1989; J. R. Anderson, *Atlas of Skeletal Muscle Pathology*, 1985; R. Buschbacher, *Musculoskeletal Disorders: A Practical Guide for Diagnosis and Rehabilitation*, 1993; J. A. Walton, G. Karpati, and D. Hilton-Jones (eds.), *Disorders of Voluntary Muscle*, 1994.

Mushroom

A macroscopic fungus with a fruiting body (also known as a sporocarp). It has been estimated that over 1,500,000 species of fungi exist on the Earth; yet only about 69,000 species have been described to date. Approximately 14% (10,000) described species of fungi are considered mushrooms. Mushrooms grow aboveground or underground. They have a fleshy or nonfleshy texture. Many are edible, and only a small percentage are poisonous.

Reproduction. Mushrooms reproduce via microscopic spheres (spores) that are roughly comparable to the seeds of higher plants. Spores are produced in large numbers on specialized structures in or on the fruiting body. A 3-in.-diameter (7.5-cm) mushroom, for example, may produce as many as 40 million spores an hour. Large numbers of spores are necessary to ensure survival of the species because most spores do not land on a suitable medium that allows germination and growth. Spores that land on a suitable medium absorb moisture, germinate, and produce hyphae that grow and absorb nutrients from the substratum. If suitable mating types are present and the mycelium develops sufficiently to allow fruiting, the life cycle will continue. In nature, completion of the life cycle is dependent on many factors, including temperature, moisture and nutritional status of the substratum, and gas exchange capacity of the medium.

Commercial production of edible mushrooms. Fewer than 20 species of edible mushrooms are cultivated commercially. The most common cultivated mushroom is *Agaricus bisporus* [**Fig. 1**], followed by the oyster mushroom (*Pleurotus* spp.) [**Fig. 2**]. Other major species of commerce include *Lentinula edodes* (shiitake) [**Fig. 3**], *Auricularia* spp. (wood-ear), *Volvariella volvacea* (straw mushroom), *Flammulina velutipes* (winter mushroom), *Tremella fuciformis* (silver-ear), *Hericium erinaceus* (bear's-head), and *Pholiota nameko* (nameko). China is the leading mushroom-producing country; Japan leads

Fig. 1. Common cultivated mushroom (*Agaricus*). (*Courtesy of D. Martinez-Carrera*)

Fig. 2. Oyster mushroom (*Pleurotus*). (*Courtesy of D. Martinez-Carrera*)

the world in number of edible species cultivated commercially.

Cultivation technology. Mushrooms may be cultivated on a wide variety of substrates. The most commonly used materials include straw (wheat, oat, and rice), wood chips and bark, cotton waste (seed hulls and screenings), hay, banana leaves, corn stalks, coffee pulp, and other agricultural waste products. The biodegradation of lignocellulosic materials from agriculture or forestry confers ecological importance on mushroom cultivation, as millions of tons of these by-products are recycled every year as substrates for mushroom growing. Successful cultivation of some mushroom species, for example, *Agaricus bisporus*, requires a composting process to produce a selective medium. Other species usually are cultivated on a noncomposted substratum that has been treated to remove unwanted competitive microorganisms.

Mushrooms are grown from mycelium (the thread-like filaments or hyphae that become interwoven) propagated on a base of steam-sterilized cereal grain (usually rye or millet). This grain and mycelium mixture is called spawn, which is used to seed mushroom substrata. Most spawn is made with mycelium from a stored culture, rather than mycelium whose parent was a spore, because each spore is likely to yield a new strain and its performance would be unpredictable. Spawn making is a rather complex task and is not practicable for the common mushroom grower. Thus, spawn usually is purchased from a number of commercial spawn makers. However, conventional breeding methods, now complemented by powerful molecular techniques, will produce new strains for spawn production capable of giving higher yields and better mushroom quality.

Agaricus bisporus is produced on compost made of straw-bedded horse manure or on synthetic composts made of straw, hay, corn cobs, or other fibrous materials, plus potash and gypsum. These raw materials are combined in piles that are periodically turned, watered, and formed. The mixing and watering serves to equilibrate the environment and promotes a uniform breakdown of the raw materials. The breakdown of materials is accomplished by the growth and reproduction of microbes on them. This breakdown is necessary to achieve a selective medium (one on which only the mushroom will grow) and to provide food for the mushroom in the most efficient form. Composting has developed from a relatively primitive process to a computer-controlled process. Bulk composting in tunnels with computer control of moisture, carbon dioxide, and temperature has allowed the development of a predictable and high-yielding compost. Research work is now focused on the production of environmentally acceptable compost, involving a considerable reduction of odor, ammonia, and water pollution generated by this process.

Other mushroom species may be grown on steam-sterilized or steam-pasteurized substrate. Several wood-inhabiting species of mushrooms that occur in the wild on dead trees or branches may be grown commercially on steam-treated wood chips. Shiitake, oyster mushroom, winter mushroom, wood-ear, silver-ear, bear's-head, and nameko are grown commercially on heat-treated wood chips. In general, sawdust supplemented with nutrient (rice and

Fig. 3. Shiitake mushroom (*Lentinula*). (*Courtesy of D. Martinez-Carrera*)

wheat bran, millet, sugar, and dried yeast) is mixed, moistened, and filled into steam-sterilizable containers such as polypropylene bottles or bags. The substratum is sterilized, cooled, and inoculated with spawn. After a period of growth, the colonized substratum is subjected to conditions conducive to the formation of mushroom fruiting bodies.

Nutritional and medicinal value. Mushrooms contain digestible crude protein, all essential amino acids, vitamins (especially provitamin D-2), and minerals; they are high in potassium and low in sodium, saturated fats, and calories. Although they cannot totally replace meat and other high-protein food in the diet, they can be considered an important dietary supplement and a health food.

Fungi have been used for their medicinal properties for over 2000 years. Li Shih-chen of the Ming Dynasty listed more than 20 species of medicinal fungi in his *Compendium of Materia Medica*, some of which are cultivated today. At present, there may be around 200 species of edible mushrooms having medicinal value; however, only a small proportion of these species is available to the consumer. Significant functional properties found in mushrooms are reduction of blood pressure and blood lipid concentrations, regulation of the immune system, and inhibition of inflammation, microbial action, and tumors.

Although there remains an element of folklore in the use of mushrooms in health and medicine, several important drugs have been isolated from mushroom fruiting bodies and mycelium. The best-known drugs obtained from mushrooms are lentinan from *L. edodes*, grifolin from *Grifola frondosa*, and krestin from *Coriolus versicolor*. These compounds are protein-bound polysaccharides or long chains of glucose, found in the cell walls, and function as antitumor immunomodulatory drugs. Administered orally or intravenously, they stimulate the body's immune system and enhance host-mediated resistance against some viral, bacterial, and fungal pathogens and various types of carcinomas. As immunostimulators, these drugs can enhance the activity of macrophages and T lymphocytes and raise interleukin-2 and antibody levels. Especially important is their ability to improve or stabilize the side effects of chemotherapy and radiation treatments, such as increasing appetite, restoring energy, and reducing nausea and pain. Krestin is commercially produced and sold in Japan for clinical use in the treatments of stomach, esophagus, colon, rectum, lung, and mammary gland cancers, usually in conjunction with chemotherapy or radiation. There are other species of mushrooms that have therapeutic effects similar to those elicited by lentinan, grifolin, and krestin, or unique effects. Another polysaccharide fraction, known as LEM, has been isolated from *L. edodes* mycelium. LEM inhibits human immunodeficiency virus (HIV) infection in vitro, and may be used against acquired immune deficiency syndrome (AIDS). The content and potency of bioactive ingredients are affected by the strain cultivated, substrates, growing conditions, and the way mushrooms are prepared and consumed. It is likely that additional chemically and biologically distinct compounds will be isolated which have medicinal value. *See* FUNGI; MEDICAL MYCOLOGY. D. Martínez-Carrera

Bibliography. C. J. Alexopoulos and C. W. Mims, *Introductory Mycology*, 1979; W. M. Breene, Nutritional and medicinal value of specialty mushrooms, *J. Food Protect.*, 53:883–894, 1990; L. A. Casselton and N. S. Olesnicky, Molecular genetics of mating recognition in basidiomycete fungi, *Microbiol. Mol. Biol. Rev.*, 62:55–70, 1998; S. T. Chang and P. G. Miles, *Edible Mushrooms and Their Cultivation*, 1989; P. Stamets, *Growing Gourmet and Medicinal Mushrooms*, 1993.

Musical acoustics

The branch of acoustics that deals with the generation of sound by musical instruments, the transmission of sound to the listener, and the perception of musical sound. It has fascinated scientists since the time of Pythagoras. It captured the interest of noted scientists such as Hermann von Helmholtz, Lord Rayleigh, Chandrasekhara Raman, John Tyndall, Félix Savart, and Frederick Saunders. It has become an interdisciplinary field joining physics with music. Indeed, important contributions have been made by psychologists, engineers, physiologists, architects, mathematicians, and materials scientists as well as by musicians and physicists.

Generation of musical sound. A main research activity in musical acoustics is the study of the way in which musical instruments vibrate and produce sound. The most common way of classifying musical instruments is according to the nature of the primary vibrator, into string instruments, wind instruments, and percussion instruments. The equivalent classification by musicologists Erich von Hornbostel and Curt Sachs is into chordophones (string instruments), aeorphones (wind instruments), idiophones (xylophones, marimbas, chimes, and so forth), and membranophones (drums; the last two categories are generally considered percussion instruments). To these may be added the electronic synthesizer, the digital computer, and the human voice.

Complex vibrations. The vibrations of a plucked string, a struck membrane, or a blown pipe can be described in terms of normal modes of vibration. Determining the normal modes of a complex vibrator is often termed modal analysis. Much of the progress in understanding how musical instruments generate sound is due to new methods of modal analysis, such as holographic interferometry and experimental modal testing. Holographic interferometry, for example, provides the investigator with a detailed contour map of the vibrational shapes at different frequencies from which the normal modes of vibration can be determined (**Fig. 1**). *See* CAVITY RESONATOR; INTERFEROMETRY; MODE OF VIBRATION; VIBRATION.

Nature of the feedback. In the case of most percussion, plucked string, and struck string instruments, the player delivers energy to the primary vibrator (string,

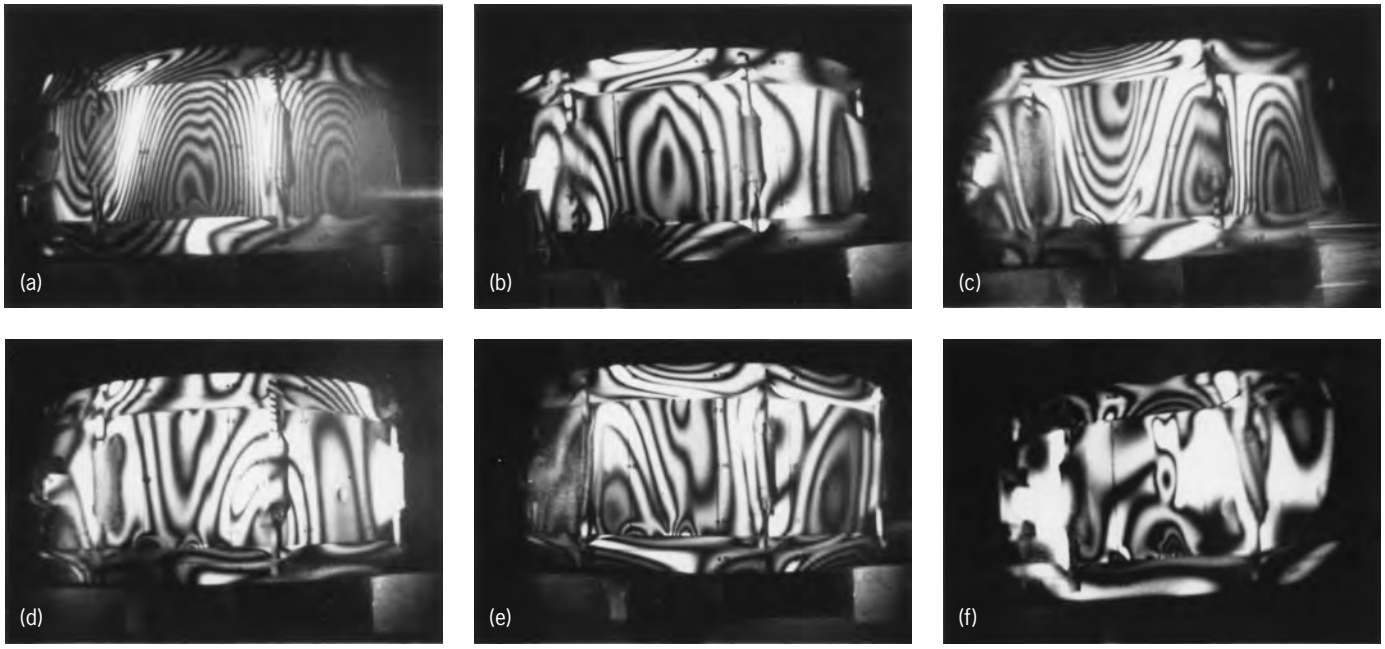

Fig. 1. Holographic interferograms of a snare drum. (*a*) 874 Hz. (*b*) 1064 Hz. (*c*) 1176 Hz. (*d*) 1271 Hz. (*e*) 1408 Hz. (*f*) 1670 Hz.

membrane, bar, or plate) and thereafter has little control over the way it vibrates. In the case of wind and bowed string instruments, however, the continuing flow of energy is controlled by feedback from the vibrating system. In brass and reed woodwinds, pressure feedback opens or closes the input valve. In flutes or flue organ pipes, however, the input valve is flow-controlled. In bowed string instruments, pulses on the string control the stick-slip action of the bow on the string.

Acoustics of singing. In singing, as in speaking, voice sounds are produced by vocal fold vibrations that interrupt the air stream from the lungs, producing a train of air pulses. This signal is modified when passing through the vocal tract, consist of the pharynx and the mouth. The vocal tract, acting as a tube resonator, filters the sound from the vocal folds, adding several strong resonance peaks or formants to the spectrum of the singer's sound. The formants of the singing voice are quite similar to those of the speaking voice, but there are some notable differences, such as the concentration of energy between 2000 and 3500 Hz, referred to as the singer's formant, which gives the trained singer's voice a "shine" and allows the opera singer to be heard in the presence of a loud orchestral accompaniment. *See* MUSICAL INSTRUMENTS; SPEECH.

Transmission of musical sound. Musical acousticians, like architectural acousticians, have always been very interested in the acoustics of concert halls. Since the average person now hears more recorded music than live performance, musical acousticians have paid increasing attention to the recording and reproduction of musical sound. It is impossible to divorce high-fidelity sound reproduction from room acoustics, since the acoustics of both the recording and the listening rooms have a strong influence on the musical sound that reaches the ear of the listener.

One of the problems in designing concert halls is the differences in the individual tastes of both performers and listeners. Some performers want the hall to be reverberant, others put a higher premium on clarity. Even getting the many parties concerned with planning a concert hall to speak the same language can be difficult. Leo Beranek (1996) selected and defined 18 terms that should be understood by both the musical and acoustical planners: intimacy (presence), reverberation (liveness), apparent source width (spaciousness), listener envelopment, clarity, warmth, loudness, acoustic glare, brilliance, balance, blend, ensemble, immediacy of response (attack), texture, freedom from echo, dynamic range, extraneous effects on tone quality, and uniformity of sound. Relating subjective qualities to physically measurable quantities, which the architect can attempt to control, is another important task. *See* ARCHITECTURAL ACOUSTICS.

Rapid advances in the technology of sound recording have had a profound effect on the distribution of music, both classical and popular. Although large libraries of music recorded using analog techniques still exist, nearly all recording now employs digital techniques. The most familiar format for digital recording of music has been the compact disc (CD) digital audio system, introduced in 1982, which now may be replaced by the DVD (digital versatile disk) with greater storage capacity. Digital audio tape recorders (DATs) appeared about 1987. *See* COMPACT DISK; OPTICAL RECORDING; SOUND RECORDING.

Perception of musical sound. Four attributes are frequently used to describe musical sound: loudness, pitch, timbre, and duration. Each of the subjective qualities depends on one or more physical parameters that can be measured. Loudness, for example, depends mainly on sound pressure but also on the

Dependence of subjective qualities of sound on physical parameters

Physical parameter	Subjective quality*			
	Loudness	Pitch	Timbre	Duration
Pressure	+++	+	+	+
Frequency	+	+++	++	+
Spectrum	+	+	+++	+
Duration	+	+	+	+++
Envelope	+	+	++	+

*+ = weakly dependent; ++ = moderately dependent; +++ = strongly dependent.

spectrum of the partials and the physical duration. Pitch depends mainly on frequency, but also shows lesser dependence on sound pressure and envelope. Timbre includes all the attributes by which sounds with the same pitch and loudness are distinguished. The **table** summarizes much of the knowledge about this dependence. Relating the subjective qualities of sound to the physical parameters is a central problem in psychoacoustics, and musical acousticians are concerned with this same problem as it applies to musical sound. *See* PSYCHOACOUSTICS.

Sound pressure level is measured with a sound level meter and is generally expressed on a logarithmic scale of decibels (dB) using an appropriate reference level and weighting network. From measurements of the sound pressure level at different frequencies, it is possible to calculate a subjective loudness, expressed in sones, which describes the sensation of loudness heard by an average listener. Musicians prefer to use dynamic markings ranging from *ppp* (very soft) to *fff* (very loud). *See* DECIBEL; LOUDNESS; SOUND; SOUND PRESSURE.

Pitch is defined as that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high. Although various attempts have been made to establish a psychophysical pitch scale (using units called mels), the scale has never become very popular. Neither has the psychophysical scale based on critical bands of hearing, on which a critical bandwidth is designated one bark (which is nearly the same as 100 mels). Rather, pitch is generally related to a musical scale where the octave, rather than the critical bandwidth, is the "natural" pitch interval. *See* PITCH.

Timbre is defined as that attribute of auditory sensation in terms of which a listener can judge two sounds similarly presented and having the same loudness and pitch as dissimilar. Timbre depends primarily on the spectrum of the sound, but it also depends upon the waveform, the sound pressure, the frequency location of the spectrum, and the temporal characteristics of the sound. It has been found impossible to construct a single subjective scale of timbre (such as the sone scale of loudness or the mel scale of pitch); multidimensional scales have been constructed. The term "tone color" is often used to refer to that part of timbre that is attributable to the steady-state part of the tone, but the time envelope (and especially the attack) has been found to be very important in determining timbre as well.

Combination tones and harmony. Another subject relating to the perception of music is combination tones. When two tones that are close together in frequency are sounded at the same time, beats generally are heard, at a rate that is equal to their frequency difference. When the frequency difference Δf exceeds 15 Hz or so, the beat sensation disappears, and a roughness appears. As Δf increases still further, a point is reached at which the "fused" tone at the average frequency gives way to two tones, still with roughness. The respective resonance regions on the basilar membrane are now separated sufficiently to give two distinct pitches, but the excitations overlap to give a sense of roughness. When the separation Δf exceeds the width of the critical band, the roughness disappears, and the two tones begin to blend.

As the tone separation Δf continues to increase, a difference tone (sometimes called a Tartini tone after the Italian violinist who discovered it around 1714). This difference tone has a pitch corresponding to the frequency difference $f_2 - f_1$. Another combination tone that can usually be heard is the cubic difference tone, having a frequency $2f_1 - f_2$. Other combination tones may be audible, but generally the difference tone and the cubic difference tone are the most prominent. They are shown on a musical staff in **Fig. 2**.

Pythagoras of ancient Greece is considered to have discovered that the tones produced by a string vibrating in two parts with simple ratios such as 2:1, 3:2, or 4:3 sound harmonious. These ratios define the so-called perfect intervals of music, which are considered to have the greatest consonance. Other consonant intervals in music are the major sixth ($f_2/f_1 = 5/3$), the major third ($f_2/f_1 = 5/4$), the major sixth ($f_2/f_1 = 8/5$), and the minor third ($f_2/f_1 = 6/5$).

Why are some intervals more consonant than others? Helmholtz concluded that dissonance (the opposite of consonance) is greatest when partials of the two tones produce 30 to 40 beats per second (which are not heard as beats but produce roughness). The more the partials of one tone coincide in frequency with the partials of the other, the less chance of roughness. This explains why simple frequency ratios define the most consonant intervals.

More recent research has concluded that consonance is related to the critical band. If the frequency difference between two pure tones is greater than a critical band, they sound consonant; if it is less than a critical band, they sound dissonant. The

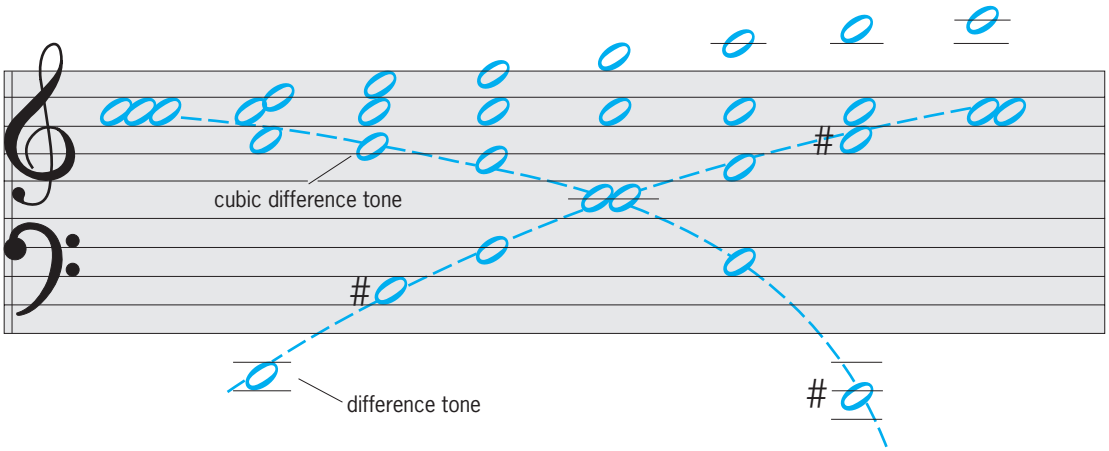

Fig. 2. Combination tones due to two tones of varying separation shown on a musical staff. As the difference tone rises, the cubic difference tone falls, so that they cross at $f_2 = 3/2\, f_1$ (a perfect fifth).

maximum dissonance occurs when Δf is approximately 1/4 of a critical band, which agrees reasonably well with Helmholtz's criterion for tones around 500 Hz.

Musical scales and temperament. Most musical composition is based on musical scales, the most common ones being those with five notes (pentatonic), 12 notes (chromatic), or seven notes (major and minor diatonic). Western music divides the octave into 12 steps called semitones. All the semitones in an octave constitute a chromatic scale or 12-tone scale. However, most music makes use of a scale of seven selected notes, designated as either a major scale or a minor scale.

There are many ways to construct musical scales. The three most important scales are the Pythagorean scale, the just scale, and the scale of equal temperament. The Pythagorean scale creates the greatest number of perfect fourths and fifths. An octave is a fourth plus a fifth ($3/2 \times 4/3 = 2$); therefore, going up a fourth leads to the same letter as going down a fifth, and vice versa. All notes on the scale (sharps and flats included) can be reached by going up or down 12 successive fifths or 12 successive fourths. If 3/2 is multiplied by itself 12 times, the result is 129.7, which means that going up 12 perfect fifths takes one up seven octaves plus one-fourth of a semitone extra. The Pythagorean scale deals with this by using semitones of two different sizes. This leads to rather poor tuning of the thirds.

Because Pythagorean thirds sound out of tune, numerous alterations to the Pythagorean scale have been developed. Nearly all of them flat the third note so that the major third and minor third are close to the corresponding just intervals. Similar compromises form the bases of various meantone temperaments.

The scale of just intonation (or just scale) is based on the major triad, a group of three notes that sound particularly harmonious (for example, C : E : G). The notes of the major triad are spaced in two intervals: a major third (C : E) and a minor third (E : G). When these intervals are made as consonant as possible, the notes in the major triad are found to have frequencies in the ratios 4:5:6. The scale of just intonation assigns this ratio to the three major triads (tonic, subdominant, and dominant).

The ultimate compromise is equal temperament, which makes all semitones the same. The scale of equal temperament consists of five equal whole tones and two semitones; the whole tones are twice the size of the semitones. Twelve equal semitones (each having a ratio $2^{1/12} = 1.05946$) make up an octave. Rather than deal with ratios, it is customary to compare tones by using cents. One cent is 1/100 of a semitone in equal temperament. Thus an octave is 1200 cents, a tempered fifth is 700 cents, and so forth. *See* SCALE (MUSIC).

Thomas D. Rossing

Bibliography. L. Beranek, *Concert and Opera Halls and How They Sound*, 1996; M. J. Crocker (ed.), *Encyclopedia of Acoustics*, vol. 4, 1997; F. A. Everest, *The Master Handbook of Acoustics*, 4th ed., 2000; N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, 2d ed., 1998; W. H. Hartmann, *Signals, Sound, and Sensation*, 1997; K. C. Pohlmann, *Principles of Digital Audio*, 4th ed., 2000; T. D. Rossing, *The Science of Sound*, 2d ed., 1990.

Musical instruments

Musical instruments evolved from simple found objects: hollow logs, perhaps with animal skins stretched over the ends, bamboo tubes or large seashells, and stretched bowstrings. They are now classified as percussion instruments, wind instruments, and string instruments, but other divisions are possible.

Conventional Instruments

A useful distinction is between impulsively excited instruments (bells, drums, plucked strings, and so forth), in which the sound gradually decays after an initial input of energy, and sustained-tone instruments (flutes, trumpets, violins, and so forth), in which energy is supplied continuously. Because extended objects can vibrate in many characteristic ways, termed normal modes, the sounds they

produce are combinations of many different frequencies. The lowest of these frequencies is called the fundamental, and the others are called upper partials. In an impulsively excited instrument, the mode frequencies are determined by the shape of the vibrating object, and there is no necessary relation between them. In a sustained-tone instrument, the frequencies of the upper partials are all exact integer multiples of the fundamental frequency, and they are called harmonics. *See* HARMONIC (PERIODIC PHENOMENA); MODE OF VIBRATION; MUSICAL ACOUSTICS; VIBRATION.

Sustained-tone instruments generate sound by means of a mechanism, such as a moving bow or a pressure-controlled valve, which essentially inserts a negative acoustic resistance into the system. Because of feedback coupling and the fact that the generator mechanism is inherently nonlinear, an exactly repetitive waveform is produced in which all overtones have frequencies that are exact integer multiples of that of the fundamental. Such overtones are called harmonics, the fundamental being the first harmonic. This harmonic structure is, however, not exactly maintained during transients, particularly the attack transient, when the inharmonic normal modes of the resonator may be momentarily excited.

A system diagram for a sustained-tone musical instrument is shown in **Fig. 1**. For an instrument such as the violin all parts of the system are present, while for wind instruments there is no separate sound radiator. For impulsive instruments, such as the guitar or the drum, there is no nonlinear generator and the energy source operates only momentarily. In all cases the sound output power is only about 1% of the input power, which is dissipated in frictional, viscous, and thermal losses.

Percussion instruments. Instruments such as drums, bells, gongs, and xylophones are called percussion instruments because the sound is initiated by a blow from some sort of hammer.

Drums. A vibrating membrane has many closely spaced modes, and the sound of a drum does not give any strong impression of musical pitch. If the membrane is stretched over an air volume, however, the resonances of this volume may emphasize certain frequencies, as in timpani or tom-toms, to produce a pitched sound. The player can vary the pitch by varying membrane tension, and can also modify the sound by striking the drumhead near the center or near the edge to preferentially excite different vibrational modes. *See* PITCH; RESONANCE (ACOUSTICS AND MECHANICS).

Tuned percussion instruments. If a heavy metal object is struck with a metal hammer, it stores a great deal of mechanical energy and can vibrate and radiate sound for many seconds. By careful shaping of the metal, usually bronze, by casting and subsequent machining, the vibrational frequencies can be brought into almost integer ratios, so that object becomes a bell with a pleasant sound and well-defined musical pitch. In a typical church bell, five or more modes are carefully tuned, one of which is in a minor-third frequency ratio 6:5 to the fundamental, to give a characteristic sound. Because cast bells are large and heavy, orchestras use sets of metal tubes to produce an approximation to a bell sound.

A rather flat metal shell struck with a soft hammer is generally called a gong. Only the first one or two modes are excited, and the sound has a soft and well-pitched note. If the shell is made from a rather thin sheet and is slightly conical in form, it is usually called a cymbal. Because the metal is thin, a blow with even a soft hammer can excite it into a quite large vibration, and this causes interactions between vibration modes, giving a characteristic shimmering sound.

Xylophones and marimbas consist of a graded set of slats of wood mounted on a frame and struck with a wooden mallet. The playing range is typically three to four octaves, though larger instruments have been built. Each slat is undercut with a circular arch to tune the first overtone, and in the marimba there is a tube resonator underneath each bar. Similar instruments are also made with metal bars.

Wind instruments. All wind instruments rely upon the vibration of a column of air enclosed by rigid walls. Sound waves propagate along this column and are reflected at its ends, the relative sign of the

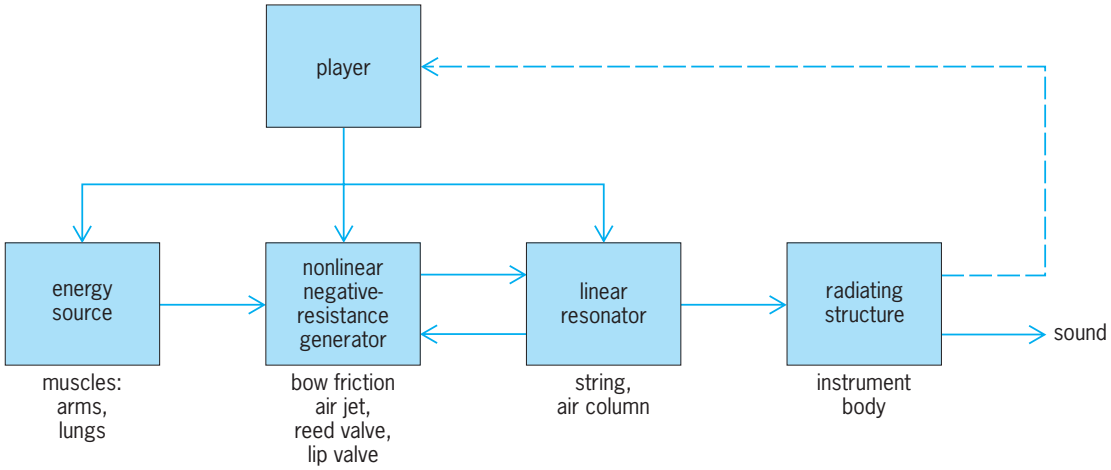

Fig. 1. System diagram for a typical musical instrument.

reflected pressure wave depending upon whether the end is open or closed. For a cylindrical air column open at both ends, the possible vibrations are those with an integral number of half wavelengths of sound along the tube, giving a frequency sequence 2, 4, 6, . . . times $c/4L$, where c is the velocity of sound in air (343 m/s or 1125 ft/s) and L is the length of the tube. The same result applies to an almost complete conical tube, irrespective of whether the narrow end is open or closed. For a cylindrical tube closed at one end, possible vibrations must have an odd number of quarter wavelengths along the tube, giving a sequence 1, 3, 5, . . . for a tube of the same length. These mode frequencies are, however, not quite exact, because of frequency-dependent end corrections. For brass instruments, the air column is nearly cylindrical for one-half to two-thirds of its length, beginning at the mouthpiece end, but flares to a bell at the open end. This flare raises the frequencies of all the modes, compared with a cylindrical tube of the same length, giving a sequence such as 1.4, 4, 6, 8, ... The first mode is not used in playing. *See* CAVITY RESONATOR.

In woodwind instruments the length of the air column, and thus the note to be played, can be changed by opening finger holes along the instrument body. Since there are 12 notes in an octave, and a human has only 10 fingers, this necessitates either compromise fingerings or else a complex system of keys as in modern instruments. In brass instruments all notes must emerge through the flared horn to maintain uniform sound quality, and pitch is changed as discussed below.

Air jet–driven instruments. In instruments of the flute family, a narrow jet of air is blown from the player's mouth, or from a narrow slit in the case of recorders and organ pipes, to cross an aperture near one end of a more or less cylindrical pipe and strike against a sharp edge. As the air jet blows into and out of the pipe mouth, it excites the pipe modes. The deflection of the jet is in turn governed by the airflow into and out of the pipe caused by the sound waves. For the system to work, there must be a match between the frequency of the mode that is excited and the travel time of waves on the jet from the player's lip to the sharp edge. The player can therefore select the mode to be played by varying lip position and blowing pressure. Once the oscillation is established, nonlinear effect generate all harmonics of the fundamental, and these are amplified because of their close match to the resonances of the tube. For a modern flute, the playing range is just over three octaves. Loudness is controlled primarily by varying the area of the lip aperture. *See* JET FLOW.

Reed-driven instruments. Many woodwind instruments use a reed valve to inject an oscillating flow of air into the instrument tube. In clarinets, which have a cylindrical tube, and in saxophones, which have a broad conical bore, the reed is a single piece of thinned cane clamped against an aperture in a carefully shaped mouthpiece. In the oboe and bassoon, which have a narrow conical bore, the reed consists of two similar pieces of thinned cane bound together. The vibration of the reed is controlled by the sound pressure in the instrument, and it effectively closes one end of the tube. Its small vibration allows puffs of air to enter from the player's mouth to maintain the sound oscillation, which is locked to one of the modes of vibration of the air column. Nonlinear effects generate all harmonics of the frequency being played, and these are amplified by the well-aligned resonances of the air column—all harmonics for a conical bore and only odd harmonics for a cylindrical one. For all these instruments, the playing range is about three octaves. Loudness is increased by reducing lip pressure on the reed, allowing it to vibrate to larger amplitude, together with a modest increase in blowing pressure.

Lip-driven instruments. In brass instruments the player's lips provide a vibrating valve under the control of the resonances of the air column. Because of the geometry of the lip-valve motion, it is necessary for the player to adjust lip tension so that the natural frequency of vibration of the lips themselves matches that of the air column for the note to be played. Because of the necessity for this match, it is possible for the player to select any of the natural modes of the air column at will, thus allowing the playing of bugle calls without the use of any valves or slides. In "natural" horns and trumpets with no valves, a whole octave of notes can be played using the 8th to 16th modes of the horn. Again, nonlinear effects generate exact harmonics of the pitch being played, and the player controls loudness primarily by controlling blowing pressure, which can be very high. Notes between the natural modes of the horn are played either by using a slide to add extra cylindrical length, as in the trombone, or by using a set of finger-operated valves to insert extra tube lengths, as in the French horn or trumpet. A set of three valves that lower the pitch by one, two, or three semitones, respectively, is adequate.

Pipe organs. A pipe organ contains as many as 10,000 pipes. Most of these are flutelike flue pipes, but there are also pipes with single metal reeds. Since each pipe need produce only one note, a great range of shapes and tone colors is possible. In addition, depressing the key of one note can sound many pipes at suboctaves, octaves, and higher multiples of the note frequency, as controlled by drawstops. An organ has as many as five keyboards and one pedalboard, and the action either may be completely mechanical or may involve pneumatic, electric, or electropneumatic actuators. Loudness and tone quality are both varied by means of stop selection, but some pipe ranks are enclosed in chambers with louvred fronts controlled by a swell pedal, allowing additional control of loudness.

String instruments. A vibrating string clamped at its two ends has modes in very nearly exact harmonic relationship, the deviation being caused by string stiffness. In addition, the length of a vibrating string can easily be changed by pressing a finger against it. Strings are therefore a good basis for musical instruments. The problem is that, being very thin, a string radiates very little sound. Its vibrations must

therefore be coupled to a much larger structure, the soundboard or body of the instrument, to achieve adequate sound level. The musical quality of the instrument depends largely on the design and resonances of the radiating structure which, generally being of wood, varies considerably from one instrument to another that is outwardly similar. In impulsive instruments the string can be excited either by plucking, as in the guitar or harp, or by hitting with a soft hammer, as in the piano. In sustained tone instruments such as the violin, the string is excited by drawing a horsehair bow across it at a constant velocity.

Plucked string instruments. Since the modes of a thin string are very nearly harmonic, the overtones of plucked strings are also nearly harmonic. The relative amplitudes of different modes can, however, be varied greatly by changing the plucking point. A plucking position near the end of the string increases the relative amplitudes of high-frequency modes and thus brightens the tone. Loudness is inceased by increasing the force of the plucking action. The relative amplitudes of the overtones, and thus the tone quality, are also affected by the resonance properties of the instrument body, and great attention is paid to this by the maker. In guitars the lowest body resonance involves motion of air through the guitar rose as well as motion of the upper plate and back, while for higher modes only the plate vibrations are important.

Bowed string instruments. Strings can be excited by the motion of a bow because the static frictional force between bow and string is greater than the sliding friction. The string successively sticks to the moving bow, or else slips back under the much smaller sliding friction. The vibration frequency of the string is determined only by its length and tension, and the sound quality varies only slightly with bowing position. Loudness is increased by increasing bow speed and contact force on the string. Violins, violas, and cellos have carved arched top and back plates, the shapes of which are carefully adjusted by the maker to achieve the desired body resonance frequencies which are important for tone quality. The shape of the traditional f-holes controls the lowest resonance, which also involves motion of the enclosed air.

Pianos and harpsichords. The notes on a piano have one, two, or three steel strings, the lower strings being overwrapped with brass or copper wire. Because of string stiffness, the octaves are stretched to a little greater than a 2:1 frequency ratio. String vibration is communicated through a bridge to a flat, ribbed wooden soundboard. Felted hammers are made to hit the strings through a complex mechanical action which also raises dampers from individual strings. Two pedals control damper motion, and a third allows only a single string of each note to be struck. Loudness is increased by increasing the force of the finger stroke on the keyboard.

The harpsichord is an earlier and mechanically simpler instrument in which single unwrapped metal strings are plucked by quill (now nylon) plectra. Most harpsichords have two sets of strings, usually with one set pitched an octave above the other. There is also usually a lute stop which applies felt dampers to all the strings to give a soft effect. Large harpsichords have two keyboards and extra sets of strings, often including a set at suboctave pitch. Both tone quality and loudness are varied by varying the string sets used. No variation in plucking force is possible.

Human voice. While not really a musical instrument, the human singing voice is often involved in music. The larynx contains two vocal folds which can open and close like the human lips. When nearly closed, they can be made to vibrate by air pressure in the lungs—the frequency of vibration, and thus the musical pitch of the sound, being controlled by muscular tension. Vocal fold vibration releases successive pulses of air into the base of the vocal tract, giving a sound with high harmonic content. The spectral envelope of the sound is modified by resonances of the vocal tract, giving bands of emphasis called formants. The frequencies of the first three formants, typically around 500, 1500, and 2500 Hz, can be changed by movement of the tongue, jaw, and lips to change the resonance frequencies of the vocal tract, and thus produce characteristic vowel sounds. Consonants are produced as bursts of noise with shaped frequency envelopes. *See* SPEECH. Neville H. Fletcher

Amplification. The sound or vibration of any conventional musical instrument can be picked up, amplified, recorded, or reproduced. Many conventional musical instruments are amplified during live performance. Thus they can be considered electronic in the limited sense that all electronic musical instruments have amplifiers, and deliver their tones to listeners over loudspeakers or earphones. For instruments having mechanical vibrators such as tuned strings, reeds, or bars, vibration pickups can be placed in contact with (or in proximity to) the vibrators (**Fig. 2**). For resonant-air-column instruments such as brass, microphones can be placed in the sound field. The amplified sound can then reinforce or even dominate the direct sound from the instrument. *See* AMPLIFIER; AUDIO AMPLIFIER; MICROPHONE; SOUND RECORDING; SOUND-REPRODUCING SYSTEMS; TRANSDUCER.

Electronic amplification circuitry allows the performer to control the tone of the instrument in unconventional ways, such as changing the tone timbre or even the frequency. However, before such control can be effective, the direct tone from the instrument must be reduced to allow the modified tone to dominate. For this purpose, amplified instruments are usually redesigned to reduce direct sound

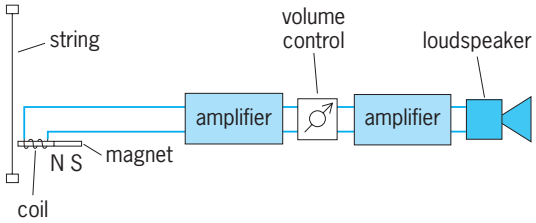

Fig. 2. Electrical scheme of an amplified electropiano. (*After H. F. Olson, Music, Physics and Engineering, Dover, 1967*)

output. Examples are solid-body electric guitars, electropianos with struck strings (or reeds) but without soundboards, amplified enclosed-reed organs, and brass instruments with mutes containing microphones. Such redesign also permits the instrument to be played more softly than normal, either for greater dynamic expression or for practice without disturbing other people, such as in a classroom where music students practice or receive instruction individually or in groups.

Changing the timbre of the conventional tone usually involves alteration of the relative amount of the harmonics in the tone spectrum. One method is to pick up the vibration or sound at different points along the string, for example, either by moving the pickup or by using several pickups mounted at different positions and selecting or mixing the outputs. Another method is to use a single pickup (or set) with tone controls or resonant filter circuits that can be varied or switched manually or by foot pedals during performance. Dynamic use of such controls can provide new timbre effects not possible in the original conventional instrument.

Another kind of dynamic timbre-changing effect is the Leslie effect, used often on the output of electronic keyboard instruments, such as Hammond organs. A Leslie effect is created by rotating one or more directional speakers inside a cabinet such that a mixture of Doppler-shifted reflections is generated. Many electronic "effects" can be considered imitations of the Leslie effect to varying degrees, such as a tremolo circuit, which imparts a simple periodic amplitude modulation on the sound, and the chorus effect, which is typically a sum of electronically delayed copies of the sound, slightly mistuned.

In addition to changing the harmonic mix, new harmonics are often created by using controlled amounts of distortion. Nonlinear distortion is often used by electric guitar players, and sometimes by keyboard players as well. When nonlinear distortion is applied to a harmonic, single-pitched sound, that is, a single note as opposed to a chord, the distortion is harmonic, and generally brightens the tone. If two or more notes are present, at different pitches, then intermodulation distortion occurs, giving rise to many new frequencies, and typically a more distorted sound.

Often in electronic amplification there is a feedback path from the loudspeaker to the vibrating source, such as acoustic coupling from an amplifier loudspeaker to the strings of an electric guitar. Such feedback can cause objectionable "squealing" noises, but it can also be used to provide indefinite sustain for certain harmonics. In effect, the amplifier becomes part of the instrument. *See* FEEDBACK CIRCUIT.

Electrical, Electronic, and Software Instruments

Electrical musical instruments produce electrical tone signals for amplification (and loudspeaker listening) without using tuned mechanical vibrators or air columns. Early electrical instruments predated electronics, generating their tone signals by electromagnetic or electrostatic rotating machinery. Later, most of the instruments became electronic, producing their tone waves from vacuum-tube circuits, then transistor circuits, and now integrated electronic circuits. In the 1980s, electronic instruments transitioned from analog to digital, offering much more precise tuning (and alternative tunings) and repeatability of configuration. The initial purpose of these instruments was to substitute for conventional musical instruments. However, over time, successful electronic instruments evolved their own repertoire, so that now electronic synthesizers typically emulate many synthesizers of the past as well as traditional musical instruments. *See* INTEGRATED CIRCUITS.

The purposes of these electronic instruments have been broadened from simple substitution for conventional instruments to (1) producing new musical sounds, (2) improving instrumentation for music education, (3) increasing the ease of music playing by amateurs, (4) extending the performance capability of trained musicians, and (5) facilitating music composition. Moreover, the means developed for each one of these purposes have also been adapted to other purposes. For example, inadequate electronic tonal substitutions have become new electronic sounds; amplified instruments have been found useful for group instruction; computer-generated tones and compositions have stimulated similar live, real-time tone generation and performances; and easy-play electronic aids for amateurs, such as automatic arpeggios, have extended professional performance beyond previous human manipulative capability.

Sound generation methods. Many different means of sound generation have been used during the evolution of electronic musical instruments, some now obsolete and others invented more recently.

In concept at least, the most simple electronic method would be to substitute directly for each conventional mechanical (string, reed) or acoustical (resonant air column) sound source an electronic circuit source or software algorithm that generates the same complex tone wave. However, in different parts of the musical scale the waveform typically differs for the same instrument. Moreover, keyboard instruments require that a number of tones be produced at the same time. Consequently, the substitution method is less feasible economically than other methods described below. *See* WAVEFORM.

Additive synthesis. Another method is to synthesize each desired complex tone by adding together an appropriate number of simple (sine) waves at harmonically related frequencies (for example, 100, 200, 300 Hz, and so forth), with the relative amplitudes of the harmonic waves adjustable to produce the desired resultant waveform (**Fig. 3**). Additive synthesis is the reverse of Fourier analysis, which enables any repetitive waveform, even a very complex one, to be analyzed into a series of simple sine-wave harmonics. *See* FOURIER SERIES AND TRANSFORMS; HARMONIC ANALYZER.

The additive synthesis method has the advantage that all of the tone-wave circuits generate the same mathematically simple waveform, and that the same

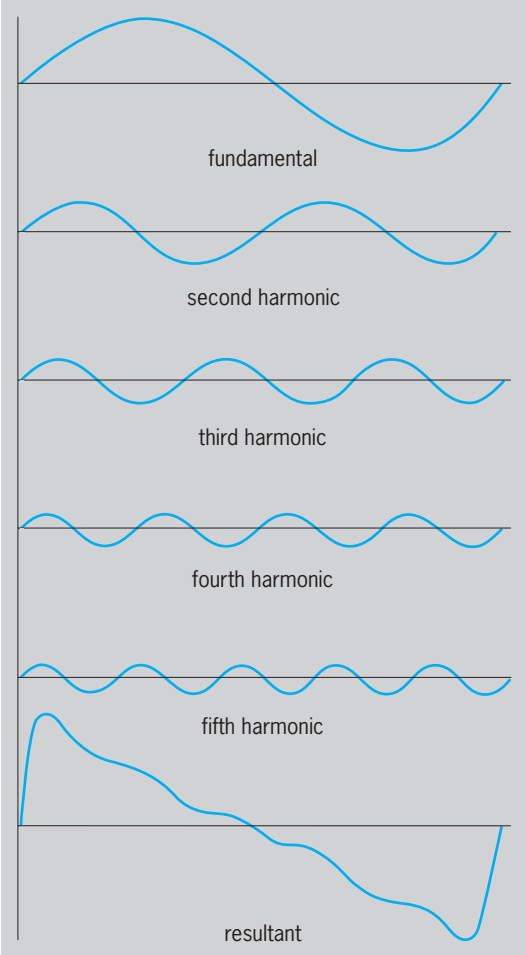

Fig. 3. Synthesis of resultant waveform from five harmonics. (*After H. F. Olson, Music, Physics and Engineering, Dover, 1967*)

harmonic generator can contribute to the synthesis of more than one tone in the musical scale. However, musical tone signals often require many harmonics (for example, string and reed tones), and thus the controls used to adjust the relative amplitudes of the harmonics to the desired values can be very complex and costly. This has typically led designers using this method to make compromises (for example, in the number of available harmonics per tone, and in borrowing inharmonic sine waves).

Additive synthesis has been used extensively in the field of computer music since at least the 1960s, and it has been the basis of a number of limited-production hardware devices. The Hammond organ can be viewed as an analog synthesizer based on additive synthesis, because its tone wheels produce quasi-sinusoidal tones that are summed with weightings provided by sliders. In current music synthesizers, however, there are typically more efficient ways than additive synthesis to achieve desired sounds, and mass-produced electronic musical instruments are generally not based on additive synthesis. However, the concept of additive synthesis remains important, and there are closely related descendents, such as sinusoidal modeling and group additive synthesis, that are both efficient and general in their sound-generating capabilities.

Subtractive synthesis. Subtractive synthesis is similar in principle to that of speech (and singing) sound formation within the human voice system. At each desired musical frequency an electronic circuit produces a standard basic tone waveform (such as a sawtooth wave; **Fig. 4**), which already contains a very large number of harmonics at known relative amplitudes. Then a variety of electric or electronic filters is provided in the instrument wherein switch selection, individually or collectively, converts the basic tone signals into the desired musical tone waveforms. Subtractive synthesis is somewhat the opposite of additive synthesis, since it practically subtracts undesired harmonics instead of adding desired harmonics. *See* ELECTRIC FILTER; FUNCTION GENERATOR; WAVE-SHAPING CIRCUITS.

The subtractive synthesis method can produce a wide variety of tonal timbres using only a small number of signal generators, and it has less switching complexity than harmonic synthesis. However, it does require filters, and the filters are generally more effective on groups of harmonics than on individual harmonics. In order to improve further the variety of available timbres in formant-type instruments, the filter input signals may be derived from generators having more than one basic waveform (for example, square, sawtooth, or pulse waves) used either individually or in additive (or subtractive) combinations.

Subtractive synthesis was used extensively in the first analog synthesizers, and it has been the basis for most voice synthesizers for over half a century. It is also used in the CELP voice coding standard and in musical synthesizers. Several digital synthesizers are devoted to imitating the original analog synthesizers, using virtual analog technology.

FM synthesis. The first all-digital synthesizer was based on frequency-modulation (FM) synthesis.

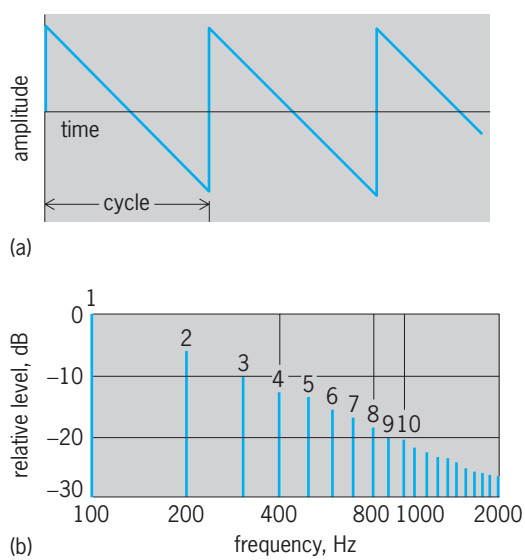

Fig. 4. Sawtooth wave. (*a*) Amplitude as function of time. (*b*) Harmonic spectrum analysis. Numbers above bars indicate harmonics. (*After H. F. Olson, Music, Physics and Engineering, Dover, 1967*)

Simple FM is carried out using two digital oscillators, with the output of one adding to the frequency (or phase) control of the other. In additive synthesis, two oscillators can provide only two harmonics. In FM synthesis, two oscillators can provide any number of harmonics. It is even possible for the output of a digital oscillator to modulate its own phase, giving so-called feedback FM. In the first all-digital synthesizer, several oscillators, called operators, could be combined in various ways, such as each oscillator modulating the next, or several FM pairs in parallel. A major advantage of FM is the synthesis of a wide variety of complex sounds from an extremely simple digital algorithm, thus making an all-digital synthesizer commercially feasible. Disadvantages of FM are its extreme sensitivity to its control parameters, and the fact that many parameter settings sound artificial. FM synthesis remains a valuable complexity-reduction technique in the context of additive synthesis. *See* FREQUENCY MODULATION.

Sampling synthesis. As memory costs came down in the late 1980s, sampling synthesis became more practical. Sampling synthesis can refer to any synthesis method based on playing back digitally recorded sounds. Usually, a sound is recorded at two dynamic levels and at selected pitches, such as every 3–6 semitones or even every octave. The velocity of the keystroke used to play a note is typically used to select a mixture of the two recorded dynamic levels, with a very slow stroke giving the softer recording and a very fast stroke giving the louder one. The duration is altered on playback, when appropriate, by looping playback somewhere in the middle of the recording. The nearest available pitch is altered to the desired pitch on playback either by adjusting the digital-to-analog conversion clock rate, or more commonly by resampling the stored waveform to change its pitch given a fixed sampling rate. In one type of sample-based synthesizer, a custom integrated circuit is used to provide many channels (such as 32 or 64) of simultaneous, high-quality, sampling-rate conversion.

Computer music composers manipulated digitally recorded sound for decades, but they used mainframe computers not available to most musicians. At first, sampling synthesis was available only in very expensive digital music systems, but later less expensive sampling synthesizers appeared, and finally highly affordable sampling synthesizers brought the technique into the mainstream. While all the early sampling synthesizers supported digital audio recording for creating sounds, most advanced synthesizers use samples stored in read-only memory (ROM), and there is no recording ability. Thus, while the synthesis technique is the same, it is better termed sample playback synthesis. It is also widely known as wavetable synthesis.

Today, ROM-based wavetable synthesizers dominate the field, from single-chip devices used in notebook computers to the most expensive synthesizers. With enough digital recordings of an instrument, it is possible to achieve any sound with sampling synthesis. However, approaching full expressive fidelity for most instruments, particularly bowed strings and solo woodwinds, requires prohibitive amounts of memory, system development, and control complexity.

Physical modeling synthesis. A more recent approach to the synthesis of acoustic musical instruments, also made possible by digital technology, is known as physical modeling synthesis. In this approach, the algorithms are based directly on the mathematical physics of the instrument. One variant of physical modeling synthesis, which is especially effective for synthesizing string and wind instruments, is waveguide synthesis. Waveguide synthesis explicitly simulates traveling waves on a string or inside a bore or horn using digital delay lines. The theory is analogous to that of sampled (discrete-time) electric transmission lines. Filters are used in conjuction with digital waveguides to simulate losses due to bridge motion, horn radiation, air absorption, and the like.

The first commercial synthesizers based on waveguide synthesis appeared in 1994. Several companies now make products using this approach, including some implemented purely in personal-computer software. One model of electronic organ uses a combination of sampling synthesis, FM synthesis, and a variant of physical modeling called virtual lead (VL) technology, based on waveguide synthesis.

Simulation of inharmonic partials, noise, and modulation. Musical timbre involves more than the number and relative amplitude of harmonics. The parts (called partials) of a complex percussive tone may be slightly inharmonic (for example, with frequencies of 100, 200.3, 300.7 Hz, and so forth) as in a piano tone, or grossly inharmonic (for example, with frequencies of 100, 165, 210 Hz, and so forth) as in a bell tone. Successful electronic simulation of these tones may require additional frequency generators.

When the number of significant freely ringing partials is small, such as for the marimba, good results are obtained using a variant of subtractive synthesis called modal synthesis. A filter in subtractive synthesis normally provides a formant, that is, a gain boost or cut which affects several successive partials. In modal synthesis, filters are used to generate the partials themselves. The filters are typically organized into a parallel bank of resonators. Each resonator is tuned to an important, audible mode of the main vibrating element (such as the wooden bar in the case of the marimba), and it is excited by a short sound such as a noise burst (corresponding to the strike of the mallet). Modal synthesis is closely related to physical modeling synthesis since the filters are set according to physical parameters of vibration. However, it is usually better considered a parametric (filter-based) model of the time-varying spectrum. In some variants, the mode-simulating filters are excited according to an analysis of the physical excitation geometry; and in this case, modal synthesis behaves as an instance of physical modeling synthesis.

Both percussion tones and continuous tones typically contain small amounts of noise and tone-wave modulation resulting from the tone generation process (for example, blowing, bowing, or striking).

Fully authentic electronic tonal simulation requires that the electronic circuits (or algorithms) generate similar components (for example, broadly or sharply tuned noise) or produce similar modulations, either random or periodic, such as vibrato. Physical modeling synthesis typically obtains the best results along these lines, since the actual physical mechanisms causing noise and modulation are simulated explicitly. Noise due to air turbulence, however, is normally simulated using a digital pseudorandom number generator, filtered to provide the desired spectral content for the noise. *See* ACOUSTIC NOISE; MODULATION.

Envelope control. In conventional percussion tones the amplitude also changes during the course of the tone, rising suddenly at tonal onset and falling during tonal decay. Equivalent electronic tone waves thus require amplitude change, called envelope control. The different partials of a complex percussive tone would ideally have different envelope control. In modal and physical modeling synthesis, the proper amplitude envelopes are provided automatically by the filters used, since they resonate and decay in the same way that the physical instrument does.

The starting of a tone often includes a sound that lasts only for an instant. It may consist of one or more discrete frequencies, or it may be a puff of noise or a click. In pipe organs it is called chiff. In percussion instruments it is the strike. Authentic electronic simulation also requires these components to be under separate envelope control, or else a proper physical model is used as a basis.

Musical Instrument Digital Interface (MIDI). Since the 1980s, the Musical Instrument Digital Interface has emerged as the standard means for automatic control of electronic musical instruments. For example, the keyboard of one instrument can be used to control any number of additional synthesizers connected via MIDI cables. Electronic instruments typically have two or three MIDI connectors: one for MIDI input, one for MIDI output, and sometimes a MIDI "thru" connector which simply copies the MIDI input. A computer can also "play" one or more synthesizers via MIDI, and keyboard performances can be conveniently recorded to the computer. There are dozens of software products available for capturing, processing, and generating MIDI control information. The MIDI protocol has even become standard for interconnecting programs from different manufacturers running on the same computer. Many composers use a sequencer program for managing several MIDI tracks in a manner analogous to a multitrack recording device. Such programs support recording, playback, display, and editing of MIDI data; they provide multiple views such as common music notation (notes on a staff), a "piano roll" view, or simply lists of raw MIDI events, all of which can be conveniently edited.

Audio tracks and MIDI tracks are often manipulated together in a unified manner using a single sequencer program running on a personal computer. Digital audio data are far more demanding than MIDI data, since the former typically require 44,000 16-bit samples per second, while three 8-bit bytes of MIDI data (a single MIDI message) can specify an entire note lasting several seconds on a synthesizer. However, standard personal computers have become powerful enough to handle several channels of digital audio data with ease. Moreover, personal computers are powerful enough to render MIDI data in real time as well, using software synthesis algorithms.

A digital recording from a sequencer program to a compact disk ideally involves mixing the digital audio tracks down to a stereo digital audio stream, and mixing in stereo digital audio streams from each MIDI synthesizer being controlled by the MIDI tracks. Creating a purely digital recording in this manner can be quite expensive if it requires a rack of MIDI synthesizers with digital outputs and a digital mixer. (One reason why digital mixers are expensive is that they need to be able to handle multiple sampling rates, and sampling rates from independent sources which are not synchronized.) If pure software synthesis is used for the MIDI tracks, a digital CD-ROM (compact-disk read-only memory) image file can be created on the computer's hard disk (not necessarily in real time) without any special hardware (aside from what is needed to record digital audio). Thus, a single personal computer can function as virtual all-digital recording studio environment, complete with virtual musical instruments and digital audio effects. *See* COMPACT DISK; COMPUTER STORAGE TECHNOLOGY; MICROCOMPUTER.

Daniel W. Martin; Julius O. Smith III

Bibliography. A. H. Benade, *Fundamentals of Musical Acoustics*, 2d ed., Oxford University Press, New York, 1976, reprint, Dover Publications, 1990; M. J. Crocker (ed.), *Encyclopedia of Acoustics*, vol. 4, sec. 130–139, pp. 1617–1695, John Wiley, New York, 1997; J. M. Eargle, *Music, Sound, and Technology*, 3d ed., Van Nostrand Reinhold, New York, 1996; N. H. Fletcher and T. D. Rossing, *The Physics of Musical Instruments*, 2d ed., Springer-Verlag, New York, 1998; P. Manning, *Electronic and Computer Music*, 2d ed., Oxford University Press, 1995; C. Roads, *Computer Music Tutorial*, MIT Press, 1996.

Musk-ox

An even-toed ungulate, *Ovibos moschatus*, which is a member of the family Bovidae in the mammalian order Artiodactyla. This single species is the northernmost representative of the family, ranging through the tundra areas and snowfields of Canada and Alaska, as well as Greenland.

The musk-ox derives its name from the musky odor it emits. It is a stoutly built animal, stands about 4 ft (1.2 m) high at the shoulder, and may weigh up to 700 lb (320 kg; see **illus.**). It has a coat of long dense hair that is resistant to the extreme cold of the windswept treeless tundra. The large, widely splayed feet are of great advantage to the animal when traveling over snow. These ruminants are herbivores; their food is principally willows during the short summer growing season, and in the winter, mosses and

Musk-ox (*Ovibos moschatus*). (*Brent Huffman/Ultimate Ungulate Images*)

lichens which they uncover from the snow by using their horns and hooves. They do not hibernate and are usually found in herds of 20–100 animals huddled together for warmth. They have 32 teeth and the dental formula is I 0/3 C 0/1 Pm 3/3 M 3/3. *See* DENTITION.

As protection against its natural enemy, the wolf, the musk-ox will form a circle with the young inside. It is hunted for its flesh, fur, and skins by both the Indians and Eskimos. Since a cow produces a single calf every 2 years, the numbers have been so reduced that they are now protected by the Canadian government. Projects to domesticate musk-oxen for wool production have been underway for a number of years at the University of Alaska and elsewhere. Wool produced from the soft undercoat is far superior to the best obtained from cashmere goats. *See* ARTIODACTYLA. Charles B. Curtin

Muskeg

A term derived from Chippewan Indian for "grassy bog." In North America ecologists apply diverse usage, but most include peat bogs or tussock meadows, with variable woody vegetation such as spruce (*Picea* mostly *mariana*) or tamara (*Larix laricina*). The **table** covers a wide variety of organic terrain that is constituted of peat and covering vegetation, which contributes to the underlying peat. Plant remains accumulate when trapped in the water or in media such as sphagnum moss which inhibit decay. The peat might accrue indefinitely or might approach a steady state of raised bogs, blanket bogs, or forest if input became balanced by erosion or by loss as methane and carbon dioxide. *See* PEAT.

Classification. The communities of species making up the covering vegetation of the muskeg suggest kinds or orders of biological organization not treated here. The organization of the table is based on the physical structure observable by nonbotanists, defined and symbolized by class letters for purposes of practical classification primarily for engineers. *See* TAIGA; TUNDRA.

Application of this system led to the identification of 17 major categories of peat structure, distinguished on the basis of relative amounts of granular and fibrous content. The texture of the peat is partly a function of its degree of woodiness, but it is also partly due to arrangement of the fibers. Texture, considered in relation to the entire fossilized mass, indicates the kind and order of characteristic homogeneity of the peat.

Vegetation cover formulas. The class symbols indicating structure in the cover may be combined into formulas. Thus cover formula AFI denotes association of trees more than 15 ft (4.6 m) high with a combination of grasslike plants and moss. The latter two elements of cover differ from the former and from each other as to stature and in being nonwoody. Also, the moss is of a velvety smooth texture in contrast, say, to a lichenaceous cover which is leathery in texture. The latter is symbolized in the reference system as class H and sometimes replaces I in the formula. Class H seldom associates with class F, and thus the formula AFH does not occur in nature. Class H is typically associated with E. Cover formulas AEH and AHE commonly characterize plant community organization in the Hudson Bay Lowlands. Class H usually signifies existence of ice in the underlying peat, indicating permafrost at these latitudes.

Qualitative expression. The position of a class symbol in a formula is important for a composite value of cover class and class predominance for a chosen area of muskeg. After the boundaries for the area are determined, the class offering most cover in the entire area is symbolized in the extreme left of the formula. Vegetation of short stature, if it covers more than that of higher, will be represented at the left in the formula only if there is no overlap of classes. Thus predominance of cover, but not necessarily stature, is progressively less from left to right in the formula.

Properties designating nine pure coverage classes

Coverage type (class)	Woodiness versus nonwoodiness	Stature (approx. height)	Texture (where required)	Growth habit
A	Woody	15 ft (4.6 m) or over	—	Tree form
B	Woody	5–15 ft (1.5–4.6 m)	—	Young or dwarfed tree or bush
C	Nonwoody	2–5 ft (0.6–1.5 m)	—	Tall and grasslike
D	Woody	2–5 ft (0.6–1.5 m)	—	Tall shrub or very dwarfed tree
E	Woody	Up to 2 ft (0.6 m)	—	Low shrub
F	Nonwoody	Up to 2 ft (0.6 m)	—	Mats and clumps or patches, sometimes touching
G	Nonwoody	Up to 2 ft (0.6 m)	—	Singly or loose association
H	Nonwoody	Up to 4 in. (10 cm)	Leathery to crisp	Mostly continuous mats
I	Nonwoody	Up to 4 in. (10 cm)	Soft or velvety	Often continuous mats, sometimes in hummocks

Quantitative expression. Quantitative determination of class coverage is difficult to achieve accurately. Usually qualitative designation based on estimates made by inspection is sufficient for broad terrain analysis. In either case, if field representation of classes is less than 25%, the classes are not regarded as significant and are not symbolized in the formula. However, strips of small area (thin peat over sand or rocky ridges) could have importance for transport, or be impediments (for example, floating peat mats over open water).

Peat. Species of a given culminating community do not contribute equally to peat in the fossilizing process. Nevertheless, structural comparison between the peat and the culminating cover suggests formulas that indirectly symbolize peat structure and texture.

Aerial interpretation. It is possible to interpret subsurface structural constitution from aerial photographs through examination of patterns that the plant communities make in the living layer of vegetation over the peat. When trees occur in the pattern, they impart a stippled appearance and the pattern is called a Stipploid Airform Pattern. The peat beneath Stipploid is coarse and woody-fibrous and contains large chunks or logs (erratics) which on occasion may be as big as fence posts. *See* AERIAL PHOTOGRAPH.

From altitudes of 30,000 ft (9 km) it is possible to identify five different airform patterns which recur over wide expanses of muskeg. By the interpretive method, not only the peat categories but also the presence of microtopographic and water dispersement features can be determined.

Distribution. Organic terrain occurs on every continent. The largest expanses of it are in Russia (especially western Siberia) and Canada (500,000 mi^2 or 1,300,000 km^2). Typical bog or spruce-larch muskeg develops in cool temperature and subarctic to arctic lands. In subtropics and the tropics there is considerable organic terrain, as in Paraguay, Uruguay, and Guyana in South America. Thus, despite differences in climate or floristics, the phenomenon of peat formation persists if local aeration or biochemical conditions hinder decay. Despite climatic and biotic differences, gross peat structure categories seem comparable the world over. Engineering prediction for operations on muskeg is thus possible, and design for operations can be standardized. *See* BOG.

Norman W. Radforth

Bibliography. H. Godwin, *Fenland: Its Ancient Past and Uncertain Future*, 1978; H. Godwin, *The History of the British Flora*, 2d ed., 1975; I. C. MacFarlane, (ed.), *The Muskeg Engineering Handbook*, 1969; P. D. Moore and D. J. Bellamy, *Peatlands*, 1973.

Muskmelon

The edible fruit of *Cucumis melo*, belonging to the gourd family, Cucurbitaceae, as do other vine crops such as cucumber, watermelons, pumpkin, and squash. The muskmelon appears to be indigenous to Africa. There are secondary centers of origin in India, Persia, southern Russia, and China. The muskmelon was a latecomer to the list of domesticated crops. It then exploded into numerous cultivars which were rapidly dispersed throughout Europe and, at an early date, into the Americas. American cultivars encompass the netted, salmon-fleshed cantaloupes; the smooth-skinned, green-fleshed Honey Dew; the green-skinned, bright-orange-fleshed Persian; the delicate-flavored, light-salmon-fleshed Crenshaw; and the wrinkle-skinned, white-fleshed Golden Beauty casaba. Other forms with very different plant and fruit characteristics are used in the Orient for pickling and in India for cooking. All these melons differ only in varietal characters and all intercross freely. *See* CANTALOUPE; HONEY DEW MELON; PERSIAN MELON; VIOLALES.

The plants are annual, trailing vines, with three to five runners that may attain a length of 10 to 12 ft (3 to 3.6 m). The branching vines have coarse, somewhat heartshaped leaves, with almost entire slight-angled, rounded, wavy margins. The runners produce short fruiting branches, which bear the perfect flowers and later the fruits. Most American varieties are andromonoecious, bearing male and perfect flowers (combined male and female). The pollen is heavy and slightly sticky, and therefore insects are required for fertilization. The domestic honeybee is the only known effective pollinator of muskmelon flowers.

Muskmelons maturing on the vine without becoming overripe are superior in quality to those harvested immature. The sugar content, flavor, and texture of the fresh fruit improves very rapidly as the fruit approaches maturity. When mature, the melon is sweet, averages 6 to 8% sugar, and has a slight to distinctly musky odor and flavor, depending upon

cultivar and environment. The flesh is rich in potassium, in vitamin C and, when deep orange, also in vitamin A.

Cultivation and culture. Muskmelon plants during all stages of development are easily killed by frost. They require fairly warm weather and are favored by bright sunshine, low humidity, and absence of rain, which tends to prevent certain diseases that often defoliate the plants in humid areas.

Muskmelons can be grown on several types of soil but not on muck. Peat, heavy clay, or adobe soil are not recommended. The soil should be fairly fertile and free from nematodes, wilt disease, and toxic amounts of alkali. Muskmelons are sensitive to acid soils. They thrive best on neutral or slightly alkaline soil.

Light-textured soils that warm up quickly in the spring favor early maturity. Maturity is also hastened by using glassine paper hot caps to increase soil and air temperature around the plants. In some locations the frost-free period may be too short to grow muskmelons from seed planted directly in the field. Muskmelons can be grown at such locations if the plants are started in greenhouses or hotbeds and transplanted to the field when the danger of frost has passed. Muskmelons are generally grown in rows 5–7 ft (1.5–2.1 m) apart, with a single plant 1 ft (0.3 m) apart in the row, or in hills of two or three plants 2 ft (0.6 m) apart. Under irrigation, the rows are on wide ridges with furrows for irrigation water between them. In the arid West, 2–3 acre-feet (2470–3700 m^3) of water per acre is commonly required to grow a crop of muskmelons.

Harvesting. The time for planting to harvest is 85–125 days, depending upon variety and weather. Cantaloupes are harvested at the "full-slip" stage of maturity. At this stage a thin abscission crack encircles the stem where it is attached to the fruit, and the melon separates easily from the stem. Harvest maturity of Persian melons is determined primarily by skin color changes, because in this variety the abscission layer does not develop or is delayed until the fruit is overripe. The ground spot (area resting on soil) of Persian melons develops a pinkish color when the fruits are mature. Honey Dew melons have achieved harvest maturity when the skin color is white and no waxy skin coating is evident. The surface may feel prickly or hairy. With the casaba varieties, Golden Beauty, Crenshaw, and Santa Claus, maturity can be determined by applying firm pressure with the thumb to the blossom end. A slight yielding or softness indicates maturity.

California, Arizona, and Texas account for approximately 80% of the acreage and 85% of the production of muskmelons in the United States. Frank W. Zink

Muskrat

A large aquatic rodent, *Ondatra zibethicus*, in the family Cricetidae and in the vole subfamily (Arvicolinae). It is known as the common muskrat (see **illustration**). It is dark reddish-brown, with a long, essentially naked, scaly, laterally compressed tail. The pelage is dense and shiny, and consists of a thick coat of underfur over which lies a covering of long, glossy guard hairs. The ears and eyes are relatively small for an animal of this size. The hindfeet bear webbed toes and are much larger than the front feet. The anterior faces of the upper incisors are yellowish-orange. The body length ranges about 14–16 in. (400–640 mm), with the tail about 7–12 in. (180–300 mm). The average weights of muskrats from Indiana are 1215 grams (2.67 lb) in adult males and 1247 grams (2.74 lb) in adult females, which is near average. The dental formula is I 1/1 C 0/0 Pm 0/0 M 3/3 = 16. The muskrat received its name because of its inguinal glands, which produce a musky odor. *See* DENTITION; MAMMALIA; RODENTIA; SCENT GLAND.

Habitat. The muskrat is found only in North America, but is common over much of the United States and Canada where suitable habitat is present. Marshes and other wet areas with an abundance of emergent vegetation, especially cattails, are the preferred habitat of muskrats, but the species also occurs along streams and ditches and about lakes and ponds. The largest populations are usually found in larger marshes. Draining of marshes and other wetlands has undoubtedly caused great decreases in muskrats. Muskrats will even take up residence in small ponds that may later dry up.

Houses. The muskrat is primarily nocturnal but is often active by day, especially in spring and fall. It spends most of the daytime in burrows or houses that it constructs in or near water. The houses are the most conspicuous indicators of the muskrat's presence on a marsh. Most muskrat houses are in water 1–2 ft (0.3–0.6 m) deep and are constructed of emergent and submerged vegetation, cattail being a favorite when it is available. The animals usually clear the emergent vegetation (including roots) from the immediate vicinity of the house, and much of this material goes into house construction. Cattails, burreed, bulrushes, rice cutgrass, watershield, wool grass, smartweed, hornwort, and other plants are used in building; considerable mud is sometimes included. Houses are roughly circular, average 4–5 ft (1.2–1.5 m) in diameter at the water level, and extend 2–3 ft (0.6–0.9 m) above the water surface. They have walls about 1–1.5 ft (0.3–0.5 m) thick. Inside

Common muskrat, *Ondatra zibethicus*. (*Photo by Dr. Lloyd Glenn Ingles; © California Academy of Sciences*)

the house is an irregularly shaped chamber above the water level. An underwater entrance leads to one side of this chamber; the entrance also serves as a plunge hole into which the muskrat dives when danger threatens. The chambers are about a foot in diameter, 6–8 in. (15–20 cm) high, and include nests of dry, shredded cattail leaves and other soft vegetation. Larger houses sometimes have two chambers. House sites are often on a stump, in or on brush, about a tree in the water, on or about some other object in the water, or along a fence. Where no such foundation is present, the muskrats cut vegetation and make a raft, which sinks to the bottom and serves as an anchor for the house. There are two major peaks of house construction—in spring and in late summer or early fall. Houses do not persist for more than a season unless they are well anchored to some durable object and constantly repaired. Rain, wind, and waves quickly destroy them, and the material from which they are made rots rapidly. In the spring, some new houses are built and others are repaired for rearing of the young. Depending on season, from one to ten animals may occupy a house, but there are usually four to five. This number can construct a house in one or two nights. In the spring, during breeding season, only one pair is found per house.

In deep water, bank burrows are used, and they are usually constructed in vertical banks. Burrows usually extend 15–20 ft (4.6–6 m).

Activity. Muskrats tend to enter and leave the water repeatedly at certain places, forming obvious "slides" in the soil along the water. In muddy areas, the footprints and the drag mark made by the tail between the foot tracks can be seen. Tagging studies have shown that many muskrats spend their entire lives within a radius of 200 yards (183 m) of their birthplace. In late summer and fall, however, individuals wander, and some have been observed more than a mile from water. During these overland travels into unfamiliar territory, muskrats are more vulnerable to predation and accidents, and many are killed by automobiles.

Muskrats remain active under the ice in winter, taking advantage of air trapped between the ice and the water. They enter and leave the water where the current prevents freezing, where the ice is broken, or where spaces occur in ice enclosing a tree or other object.

Food. Muskrats are primarily vegetarian but will eat animal matter, even carrion, under certain conditions. Probably, most emergent plants are eaten, and other common foods include cattails, burreed, bulrush, water lilies, pondweeds, smartweeds, duck potato, water plantain, rice cutgrass, hornwort, sedges, grasses, swamp loosestrife, buttonbush, and woolgrass.

Animal foods include winter-killed fishes, frogs, crayfishes, mussels, and other dead animals. Animal material is usually taken in greatest abundance in late fall, winter, and early spring, when succulent plants are not readily available.

Breeding and development. Pair formation is apparently initiated mostly by the female, which swims back and forth before the male and utters squeaking notes. Courtship begins in February, and copulation has been observed as early as the second week of March. Gestation takes about 28–30 days. Most young are probably born in May and June, but there are records of gravid females in late fall and even in early winter. Litter size ranges from 3 to 10, averaging about 7.

Young muskrats at birth are blind, nearly naked, and helpless. They have round tails, which become laterally flattened at about the time of weaning and reach their full shapes at about 3 months of age. The young are reared in dry, warm nests of soft vegetation in the chambers of muskrat houses or burrows. The average weight at birth is about 22 grams (0.05 lb). The eyes open at about 2 weeks, when the young are covered with dark, lead-gray fur; this pelage is replaced by adult-colored hair at about 3 months of age. The young swim at about 3 weeks but have difficulty diving. At 4 weeks, when they average about 200 grams (0.45 lb), they are weaned and on their own. If the female produces a second litter, the young are driven from the house or burrow and must find new living quarters. Muskrats usually have two or three litters per season. The young from the last litter may stay with the parents until the next breeding season. The young grow rapidly and, in good habitat, equal the size of the adults when 4–5 months old. Weights of about 1.37 kg (3 lb) can be attained in 8 months. Reproductive success depends upon habitat conditions and population size. Overpopulation leads to smaller litters and fewer litters per female, because of stress and strife brought on by more frequent encounters between individuals. Some animals have been known to be alive in their fourth year, but only 10–15 percent reach an age of one year.

Threats. Minks can be important predators on muskrats. They sometimes enter muskrat houses by making a hole through the side into the central cavity. Raccoons, dogs, and red foxes prey on them, and dogs are especially destructive in late summer when water levels recede so that the houses are more accessible. Heavy rains, winds, and waves may destroy houses and drown the young. Highway kills may be numerous, especially where heavily traveled roads border good muskrat habitat. Raccoons obtain mostly young muskrats, which they take from their nests after tearing holes in the sides of muskrat houses. Drainage, however, remains the major threat to the species. Severe drought may dry up certain habitats, and the animals may be forced to abandon an area completely. Shallow-water areas that freeze to the bottom cause problems, since muskrats generally do not store food for winter. They may find it difficult or impossible to dig for roots and other foods when ice interferes, and may starve if ice persists for long periods.

Neofiber alleni. The common muskrat does not occur in Florida. However, in the grassy marshes or "prairies" (level, almost treeless bogs or marshes) of Florida and in the Okefenokee Swamp area of southeast Georgia, there is a much smaller

rodent, the round-tailed or Florida muskrat, *Neofiber alleni*. It is suggestive of a small muskrat in general appearance, and it has been characterized as a link between the meadow vole and the muskrat. The ectoparasites (parasites that live on the host's exterior) of these two animals are very different, suggesting that *Ondatra* and *Neofiber* have long been separated. A very young round-tailed muskrat in the grass looks much like a meadow vole. The round-tailed muskrat has a dense, soft, waterproof coat with very thick underfur and dark guard hairs. The small ears are nearly hidden in the fur. The scaly, round tail is almost devoid of hair, and the hindfeet are slightly webbed. The upper parts are generally a dark, rich, uniform brown. Juveniles are lead-gray. The hindfoot is much larger than the front food. The skull is similar to that of *Ondatra* but smaller. The tooth formula is the same as that of the common muskrat.

Neofiber is much less aquatic than *Ondatra*, but it is an excellent swimmer and takes to the water readily, swimming and diving with ease. Sometimes, these animals live in burrows rather than houses. *Neofiber* is active primarily at night.

Neofiber breeds throughout the year. Gestation takes 26–29 days. Four to six litters of one to four are produced per year.

Dogs, cats, water moccasins, harriers (marsh hawks), and barn owls are important predators on this species.

Myocastor coypus. The nutria (family Myocastoridae) is a South American aquatic rodent but is included here because it is very similar and occupies a similar niche to the North American muskrat, *Ondatra zibethicus*. In addition, it has been introduced into the eastern United States where it competes with the muskrat for habitat. The tail is long, but is round rather than being laterally compressed as in the muskrat. As with the muskrat, eyes and ears are small, and the hindfeet are webbed and much larger than the forefeet. There are four mammae, which are far up on the sides of the chest. The nutria is strictly vegetarian. John O. Whitaker

Bibliography. D. Birkenholz, A study of the life history and ecology of the round-tailed muskrat (*Neofiber alleni* True) in north central Florida, *Ecol. Monogr.*, 33:187–213, 1963; D. L. Bounds, M. H. Sherfy, and T. A. Mollett, Nutria *Myocastor coypus*, pp. 1119–1147 in *Wild Mammals of North America: Biology, Management, and Conservation*, ed. by G. A. Feldhamer, B. C. Thompson, and J. A. Chapman, Johns Hopkins University Press, Baltimore, 2003; J. Erb and H. R. Perry, Jr., Muskrats *Ondatra zibethicus* and *Neofiber alleni*, pp. 311–348 in *Wild Mammals of North America: Biology, Management, and Conservation*, ed. by G. A. Feldhamer, B. C. Thompson, and J. A. Chapman, Johns Hopkins University Press, Baltimore, 2003; R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, Baltimore, 1999; J. O. Whitaker, Jr. and W. J. Hamilton, Jr., *Mammals of the Eastern United States*, Cornell University Press, Ithaca, NY, 1998; D. E. Wilson and S. Ruff, *The Smithsonian Book of North American Mammals*, 1999.

Mustard

Any one of a number of annual crucifer species of Asiatic origin belonging to the plant order Capparales. Mustards eaten as greens are *Brassica juncea*, *B. juncea* var. *crispifolia*, and *B. hirta*. Table mustard and oils are obtained from *B. nigra*. Cultural practices are similar to those used for spinach. Southern Giant Curled and Ostrich Plume are popular varieties for greens. Long days and high temperatures favor undesirable seed-stalk development. Harvesting is usually $1^1/_2$–2 months after planting. Important production centers for mustard greens are in the South, where the crop is popular. Montana and the West Coast states are important sources of mustard seed. *See* CAPPARALES; SPINACH. H. John Carew

Mutagens and carcinogens

A mutagen is a substance or agent that induces heritable change in cells or organisms. A carcinogen is a substance that induces unregulated growth processes in cells or tissues of multicellular animals, leading to the disease called cancer. Although mutagen and carcinogen are not synonymous terms, the ability of a substance to induce mutations and its ability to induce cancer are strongly correlated. Mutagenesis refers to processes that result in genetic change, and carcinogenesis (the processes of tumor development) may result from mutagenic events.

Mutagenesis and carcinogenesis. Ionizing radiation was the first agent that was recognized as having the capacity to induce mutations and also to induce cancer. Induction of mutations by ionizing radiation has been recognized to be a product of ionization paths within cells that, when they occur in deoxyribonucleic acid (DNA), cause the breaking of chemical bonds and the alteration of the nucleotide bases. Alterations in properties of cells that occur as a result of exposure to radiation of chemicals are called phenotypic changes; if these changes are the result of DNA alterations and are transmitted to progeny cells, they are mutations. When such mutations affect genes that regulate cellular growth, they may result in cells called transformants, which can give rise to tumors.

Radiation can induce alterations in any segment of DNA, but some of the most important effects occur in genes that are required for survival of cells (injury to one of these genes may kill the cells) or for other important processes such as regulation of cell functions (injury to one of these genes may allow cells to grow uncontrollably). Cancers are disturbances in the growth of cells so that they are no longer under the control of the body. As the altered cells grow, they produce additional cells like themselves, leading to the development of a tumor.

Unlike radiation, many chemicals must undergo a process of enzymatic interaction (metabolism) in cells or tissues. As the cellular enzymes interact with chemicals, they often generate closely related chemical structures that possess mutagenic potential. The

chemical basis for induced mutations was found to be the presence of electrophilic groups in the mutagenic chemicals. *See* MUTATION; RADIATION BIOLOGY.

Mutagens. A mutation is any change in a cell or in an organism that is transmitted to subsequent generations. Mutations can occur spontaneously or be induced by chemical or physical agents. The cause of mutations is usually some form of damage to DNA or chromosomes that results in some change that can be seen or measured. However, damage can occur in a segment of DNA that is a noncoding region (that is, no gene product is produced), and thus will not result in a mutation. Mutations may or may not be harmful, depending upon which function is affected. They may occur in either somatic or germ cells. Mutations that occur in germ cells may be transmitted to subsequent generations, whereas mutations in somatic cells are generally of consequence only to the affected individual.

The impact of induced germ cell mutations on humans has been difficult to determine. Germ cell mutations are the basis of inherited human disorders, but there is uncertainty about what types or frequency of mutations are induced by exposure to mutagens. Most mutagens cause specific chemical changes in the informational content of DNA. Among the specific changes are alterations of purine or pyrimidine bases in DNA by rearrangement, deletion, or insertion events. The consequences of such changes can be the loss or altered activities of gene products.

Not all heritable changes result from damage to DNA. For example, in growth and differentiation of normal cells, major changes in gene expression occur and are transmitted to progeny cells through changes in the signals that control genes that are transcribed into ribonucleic acid (RNA). The mechanisms by which differentiation is controlled are not known. It is possible that chemicals and radiation alter these processes as well. When such an effect is seen in newborns, it is called teratogenic and results in birth defects that are not transmitted to the next generation. However, if the change is transmissible to progeny, it is a mutation, even though it might have arisen from an effect on the way in which the gene is expressed. Thus, chemicals can have somatic effects involving genes regulating cell growth that could lead to the development of cancer, without damaging DNA.

Carcinogens. Cancer arises because of the loss of growth control by dearrangement of regulatory signals. Included in the phenotypic consequences of mutations are alterations in gene regulation brought about by changes either in the regulatory region or in proteins involved with coordinated cellular functions. Altered proteins may exhibit novel interactions with target substrates and thereby lose the ability to provide a regulatory function for the cell or impose altered functions on associated molecules. Through such a complex series of molecular interactions, changes occur in the growth properties of normal cells leading to cancer cells that are not responsive to normal regulatory controls and can eventually give rise to a visible neoplasm or tumor. While mutagens can give rise to neoplasms by a process similar to that described above, not all mutagens induce cancer and not all mutational events result in tumors. Many DNA changes may occur in genes that do not relate to cancer or do not affect gene function. Also, normal cells possess the ability to repair segments of DNA damaged by ionization, chemical adduction, or other events, thus erasing the damage from the DNA sequence. Alternatively, some chemicals that have the potential to form electrophiles may not be taken up by cells, may not be metabolized to an electrophilic intermediate, or may be rapidly detoxified and excreted from the organism.

However, there are carcinogens that do not give rise to electrophiles and are believed to induce tumors by mechanisms that do not involve direct DNA interactions and mutation. The mechanisms by which nonmutagenic carcinogens act are not well understood.

The identification of certain specific types of genes, termed oncogenes, that appear to be causally involved in the neoplastic process has helped to focus mechanistic studies on carcinogenesis. Oncogenes can be classified into a few functionally different groups, and specific mutations in some of the genes have been identified and are believed to be critical in tumorigenesis. Tumor suppressor genes or antioncogenes provide a normal regulatory function; by mutation or other events, the loss of the function of these genes may release cells from normal growth-control processes, allowing them to begin the neoplastic process. *See* ONCOGENES; ONCOLOGY.

Identifying mutagens and carcinogens. There are a number of methods and systems for identifying chemical mutagens. Because mutagenesis involves several steps and since there may be variation between organisms used for detecting mutagens, an agent may be mutagenic only in some systems. Mutations can be detected at a variety of genetic loci (genes) in very diverse organisms, including bacteria (*Salmonella typhimurium*, *Escherichia coli*), insects (*Drosophila melanogaster*), cultured mammalian cells, rodents, and humans. An important advantage in using cultured cells is that they can be used under conditions designed to specifically select for induced mutants. Spontaneous and induced mutations occur very infrequently, the estimated rate being less than 1 in 10,000 per gene per cell generation. This low mutation rate is probably the result of a combination of factors that include the relative inaccessibility of DNA to damaging agents and the ability of cellular processes to repair damage to DNA. By applying methods that can select for mutations occurring in specific genes, it is possible to use fewer cells to detect mutations, and to be able to generate enough mutants to study quantitative aspects of the mutation process.

Factors that contribute to the difficulty in recognizing substances that may be carcinogenic to humans include the prevalence of cancer, the diversity of types of cancer, the generally late-life onset

of most cancers, and the multifactorial nature of the disease process. These factors make the identification of causal factors in human cancer extraordinarily difficult and reduce the certainty of epidemiological methods that attempt to identify risk factors and causal associations. Adding to this uncertainty is the diversity of types and forms of cancers, which suggests different causes for what may be a group of diseases. Furthermore, like the problem of heritable mutations, it is difficult to discriminate between cancers that may be of spontaneous origin and those that may be induced by exogenous substances. With the exception of sunlight (UV), asbestos, and tobacco, the major causes of most human cancers are not known.

Approximately 50 substances have been identified as causes of cancer in humans, but they probably account for only a small portion of the disease incidence. *See* CANCER (MEDICINE); HUMAN GENETICS.

Raymond W. Tennant

Bibliography. J. Ashby and R. W. Tennant, Definitive relationship among chemical structure, carcinogenicity and mutagenicity for 301 chemicals tested by the U.S. NTP, *Mutat. Res.*, 257:229–306, 1991; J. M. Bishop, Molecular themes in oncogenesis, *Cell*, 64:235–248, 1991; L. M. Franks and N. M. Teich (eds.), *Introduction to the Cellular and Molecular Biology of Cancer*, 3d ed., 1997; R. W. Tennant and E. Zeiger, Genetic toxicology: Current status of methods of carcinogen identification, *Environ. Health Perspect.*, 100:307–315, 1993; H. Vainio et al., *Mechanisms of Carcinogenesis in Risk Identification*, 1992.

Mutation

An abrupt, heritable change in genes or chromosomes manifested by changes in the phenotype (the appearance) of an organism. It is theoretically preferable to define mutations as changes in deoxyribonucleic acid (DNA) sequences, but the classical definition remains the operational definition in most circumstances. Genetic segregation and recombination, however, are not mutational processes unless aberrant. *See* DEOXYRIBONUCLEIC ACID (DNA).

The word mutation has two common meanings, one being the process and the other the product (the altered gene or chromosome carries a mutation). The process is also called mutagenesis. An organism bearing a mutation is called a mutant. An agent that induces a mutation is called a mutagen.

The study of mutation has long occupied a central position in genetics. Mutations are the ultimate sources of variability upon which evolution acts, despite being random changes that are far more likely to harm than to improve a complicated and highly evolved organism. Laboratory reconstructions have shown that rapidly mutating microbial populations overtake slowly mutating populations when the two are mixed and placed in a new environment to which neither is fully adapted. Mutation has consistently been the most telling probe into the nature of the gene, and understanding of most aspects of biology has benefited from studies of the properties of mutant organisms. Mutation is also an important component of disease, either causing it directly (for example, through birth defects) or predisposing humans to a vast array of disorders that together constitute a substantial fraction of illnesses. Finally, deliberate selection of mutant plants and animals for economic or esthetic purposes has long been practiced, and has grown into an important aspect of genetics. With advances in molecular genetics, it is now possible to construct specific mutations at will, rather than merely selecting among an array of random mutations for the infrequent useful ones. *See* BREEDING (ANIMAL); BREEDING (PLANT).

Anatomy. Although there are many sites where mutations can occur in even a single chromosome, and many mechanisms to generate them, the products of mutation can be simply cataloged in genetic terms.

Genome or ploidy mutations. The sum of a cell's genes is its genome, and ploidy refers in a general sense to the number of copies of each chromosome in a (nondividing) cell. A cell that has accidentally doubled its number of chromosomes from the normal diploid state (two copies of each chromosome) is called tetraploid; one that has lost a single member of a normal pair is called monosomic; and one that has gained a single member of a normal pair is called trisomic. In higher animals such changes are usually lethal or severely debilitating, the best-known human example being trisomy for chromosome 21 leading to Down syndrome (which occurs at a frequency of roughly 0.001 per birth). *See* CONGENITAL ANOMALIES; POLYPLOIDY.

Chromosome mutations. Chromosome mutations, which alter sections composed of many DNA base pairs, consist of partial losses (deletions or deficiencies), rearrangements, and additions. Like genome mutations, most deletions that remove many genes are highly deleterious. Rearrangements may be less deleterious if they shuffle genes about but do not interrupt them or relocate them to sites where they cannot function well. They involve either inversions (simple reversals of an internal segment of a chromosome) or translocations (transfer of a segment of a chromosome to a new location). Translocations can occur within or between chromosomes, the latter often reciprocally. Even when not directly deleterious, rearrangements lead to anomalies of genetic recombination; a common secondary consequence in humans is sterility. Addition mutations are of two types, duplications and insertions. Duplications usually consist of tandem repeats of a segment of a chromosome, and may range from innocuous to lethal depending upon their location and extent. Insertions occur through the movement of special DNA sequences (transposons) that range from hundreds to thousands of DNA base pairs in length. *See* CHROMOSOME; CHROMOSOME ABERRATION.

Gene mutations. Gene mutations affect only a single gene and consist of intragenic chromosomal mutations, additions or deletions of one or a few base

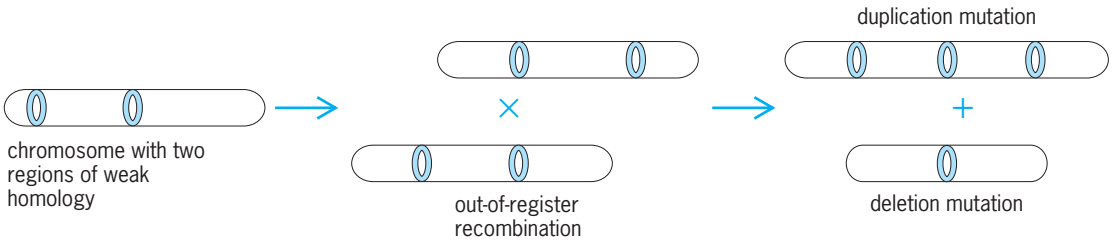

Fig. 1. Formation of duplication and deletion mutations by aberrant recombination.

pairs, base-pair substitutions (point mutations), and complex mutations comprising simultaneously arising clusters of any of the above. The severity of a gene mutation depends on its individual nature and on the importance of the affected gene, and can range from innocuous to lethal.

Base-pair substitutions are divided into two groups called transitions and transversions. In transitions, such as A · T → G · C, the purine·pyrimidine orientation is maintained; in transversions, such as A · T → C · G, this orientation is reversed. (Here the DNA bases are denoted by A = adenine, T = thymine, G = guanine, and C = cytosine, where A and G are purines and T and C are pyrimidines; the dots indicate hydrogen bonding between bases.)

Because the genetic code employs consecutive sets of three DNA base pairs to specify consecutive amino acids in proteins, the addition or deletion of multiples of three base pairs leads to the addition or deletion of one or more amino acids. However, the addition or deletion of one or two base pairs (or any nonmultiple of three) shifts the reading frame, so that everything from that point onward is read out of its normal frame, with drastic consequences for that particular gene. These occurrences are called frameshift mutations. *See* GENETIC CODE.

Forward and reverse mutation. Mutations from a normally functioning reference gene, chromosome, or organism to a mutant condition are called forward mutations. Their reversal by a new mutation that restores the original DNA sequence is called back mutation, or reversion. In addition, new mutations at a site distinct from a forward mutation can sometimes restore the nonmutant phenotype; these are called suppressor mutations. For instance, a base-pair addition in a protein-encoding sequence can sometimes be suppressed by the nearby deletion of a different base pair: the reading frame is restored, and the organism may no longer appear mutant if the associated amino acid changes are innocuous and the protein therefore functions normally.

Soma and germ line. In sexually reproducing multicellular animals, a mutation that arises in a somatic (body) cell cannot be passed to future generations, whereas a germ-line mutation can. Even though somatic mutations cannot harm future generations, they can be important to the individual that carries the mutant cell. They can be deadly (as by leading to cancer) or beneficial (as by generating new antibody molecules).

Scoring mutations. The rarity and sporadic nature of mutations render their study in wild organisms (that is, those in nature) very difficult. Instead, mutagenesis is studied in specialized laboratory organisms, among which microorganisms are favored because the ease of growing large populations (such as 10^9) overrides the low frequencies of mutants (such as 1 in 10^7 organisms). *See* BACTERIAL GENETICS; BACTERIAL GROWTH.

In the bacterial virus T4, the circular plaques formed by viral killing of a lawn of the bacterial host have a characteristic size and fuzzy edge. Visual screening readily detects *r* mutants that make larger, sharp-edged plaques. Among these, the *rII* mutants have been widely studied because their rare revertants (and also recombinants) can be selected by growing on special bacterial hosts resistant to the parental *rII* viruses. *See* BACTERIOPHAGE.

Among bacteria and yeasts, many biochemical traits have been used to score mutations. A frequent approach is first to obtain a mutant that requires for its growth a special nutrient, such as a vitamin or an amino acid. By growing the mutant population in a medium lacking the required growth factor, rare revertants to nutritional independence can be selected. For instance, the widely used Ames test for environmental mutagens uses mutants of the bacterium *Salmonella typhimurium* that require the amino acid histidine for growth, and the method tests for mutagens that can induce reversion.

In the fruit fly *Drosophila melanogaster*, special chromosomes are used to score induced lethal mutations on the X chromosome. Because these mutations are recessive (their effect being masked in the presence of a nonmutated chromosome) and because the X chromosome is a sex chromosome, this system screens for sex-linked recessive lethals. It is the most efficient mutation-scoring system in a higher eukaryote. *See* SEX-LINKED INHERITANCE.

Because of the difficulties of raising and examining mice by the tens of thousands under well-controlled conditions, mutation experiments are infrequently conducted with these or any other mammals. However, their chromosomal mutations can sometimes be scored by cytological (microscopal) analyses, and their gene mutations can be scored in a few systems, such as the specific-locus system in which mutations can be detected in any of seven specific genes determining, for the most part, coat-color traits. (A locus is the position of a gene on a chromosome.) The

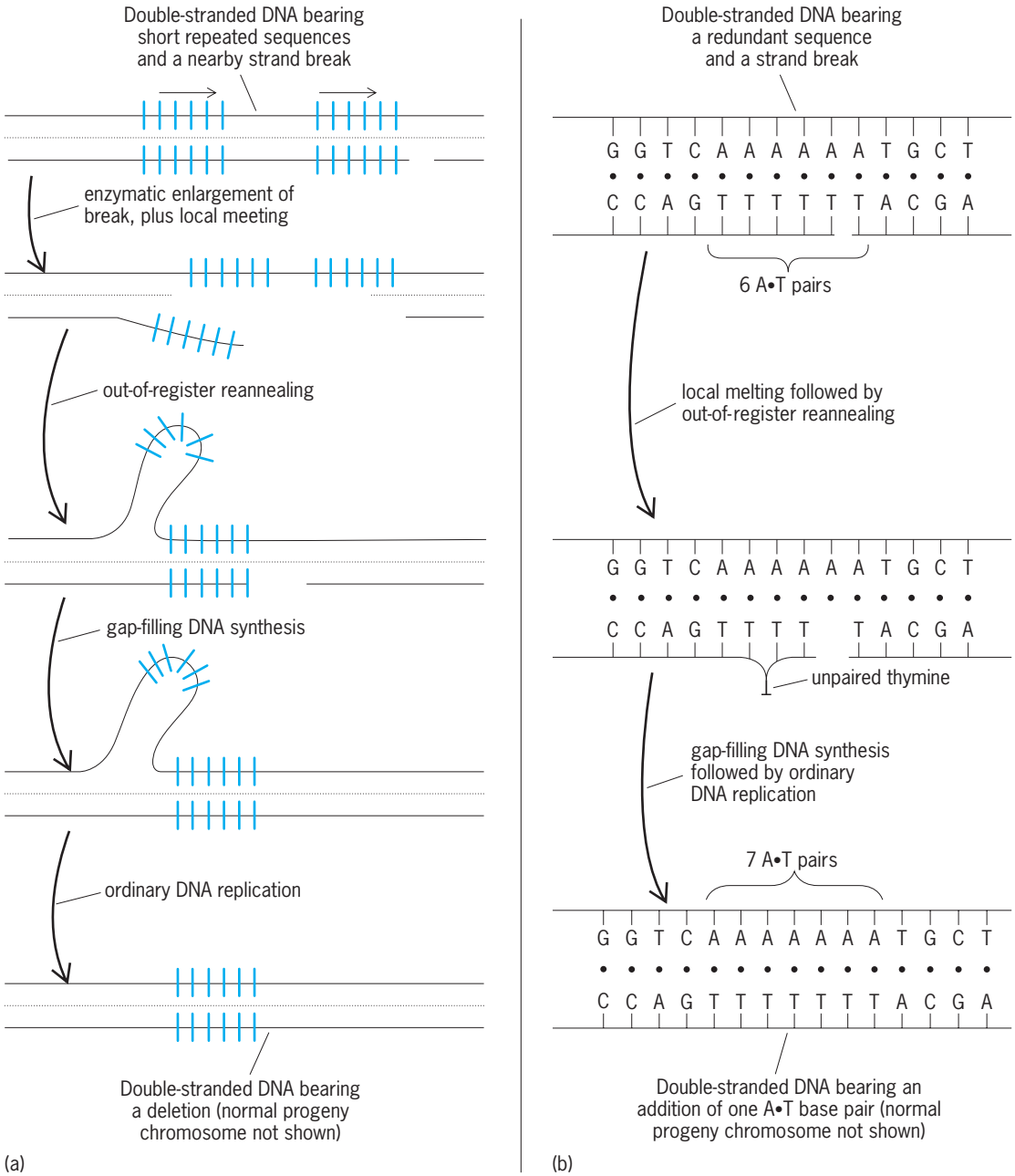

Fig. 2. Schemes for mutagenesis by misalignments within a single chromosome: (*a*) between distant repeated sequences; and (*b*) within a redundant sequence. Continuous lines represent DNA sugar-phosphate backbones, and dots indicate base pairing via hydrogen bonds.

systematic study of mutation in humans is difficult for obvious reasons, and has amounted to little more than recording incidences and examining patterns of inheritance.

Mechanisms. Although important mechanisms of mutagenesis undoubtedly remain to be discovered, many, and perhaps most, of the predominant mechanisms are now known, at least in outline.

Genome mutations. Mutations that alter the number of chromosomes in a cell usually result from the faulty distribution of chromosomes during mitotic or meiotic cell divisions. The fault probably often lies in the systems of spindle fibers that segregate daughter chromosomes into daughter cells; chemicals (such as colchicine) that interfere with such fibers induce aneuploidy at high frequencies. *See* MEIOSIS; MITOSIS.

Chromosome mutations. The larger of these, and probably many of the smaller as well, can be formed by chromosome breakage followed by incorrect patterns of rejoining. Many agents, including ionizing radiations and numerous chemicals, can induce chromosome breaks. As might be expected from their topology, the frequency of chromosome mutations often corresponds to the square of the dose of mutagen, that is, as "two-hit" events (two breaks plus incorrect rejoining). The mechanism of efficient rejoining of broken chromosomes is not understood, but may involve base pairing between repeated DNA sequences in the chromosome. The frequency of

repeated sequences is high in the chromosomes of higher organisms, and chromosomal mutations, particularly deletions, are also frequent relative to point mutations in these organisms. In addition to events triggered by breaks, however, deletions and duplications are triggered by anomalies of genetic recombination between similar but nonhomologous DNA sequences. *See* RADIATION BIOLOGY; RECOMBINATION (GENETICS).

Insertion mutations. These usually arise, not following random chromosome breaks, but through the intrinsic mobility of highly specialized DNA sequences. Called transposons, they have been found to be a major factor in spontaneous mutagenesis, because their transposition into a gene is very likely to inactivate that gene. They seem to play at most a minor role in induced mutagenesis. Transposons come in several types and sizes. Many carry repeated DNA sequences at their ends. Their mobility is often engendered by one or more of their own genes. For example, a DNA copy may be produced and then inserted elsewhere by a specialized recombination mechanism. Alternatively, the transposon may be transcribed into ribonucleic acid (RNA), copied back into DNA by a reverse transcriptase, and then inserted. *See* TRANSPOSONS.

Misalignment mutagenesis. This is a set of mutagenic mechanisms that proceeds through correct DNA base pairing in an incorrect (misaligned) context, generating deletions, duplications, and point mutations.

Consider two DNA sequences, identical or nearly so but separated by several to many base pairs. If the repeated sequences are sufficiently long (perhaps dozens to hundreds of base pairs) to mediate genetic recombination, then "unequal" recombination can occur, generating a duplication or a deletion or both (**Fig. 1**). Even if the repeated sequences are too short for ordinary recombination, they may still mediate anomalies of DNA metabolism that lead to duplications and deletions (**Fig. 2**): a break occurs in one of the two DNA chains, the chains separate (melt) locally, and then they reform a double helix out of register (misanneal). Subsequent DNA synthesis then closes the gap, fixing the mutation in the chromosome. In the extreme example, the DNA sequence repeat lacks intervening bases, and the additions or duplications are of only one or a few bases (generating, if occurring within a protein-coding sequence, frameshift mutations).

If one of the components of a repeated DNA sequence is inverted with respect to the other (both end to end and top strand to bottom strand in order to preserve the chemical polarity of DNA), the result is a DNA palindrome (a sequence that is the same when read in either direction). Just as direct repeats can mediate misalignments, so can palindromes, but with sometimes quite different consequences. For instance, during the transitory stages when DNA becomes single-stranded, as during replication, excision repair, and recombination, palindromic DNA sequences can fold back upon each other to form "hairpin" structures (**Fig. 3**). Such an anomaly can

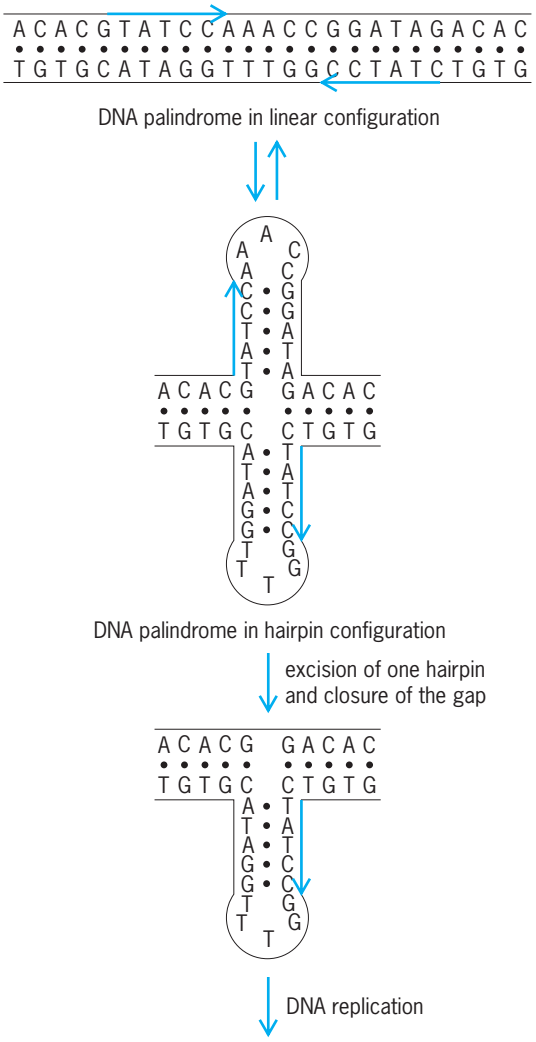

Fig. 3. Palindrome-mediated deletion mutagenesis. Color indicates palindromic sequences.

lead to a deletion, either because synthesis of the complementary strand passes by the hairpin or because the hairpin is recognized as an abnormal DNA structure and excised. In practice, while the ends of deletions often fall in either repeated or palindromic sequences, they also often fall in regions which contain both elements at once, thus providing enhanced misalignment stability.

Palindromes possess an additional property, not shared with direct repeats, that causes them to mediate the formation of point mutations. Consider a palindrome which is imperfect, the inverted repeats not being perfect complements. (Here the intervening bases between the palindromic elements are irrelevant.) If it assumes a hairpin structure, then its stem will encompass mispaired or nonpaired bases (**Fig. 4**). The repair systems that maintain the structural integrity of DNA by excising damaged, abnormal, or mispaired bases can then act upon this imperfect stem, "correcting" one of its strands. The result can be base-pair substitutions, base additions or deletions, and complex mutations. (It should be noted

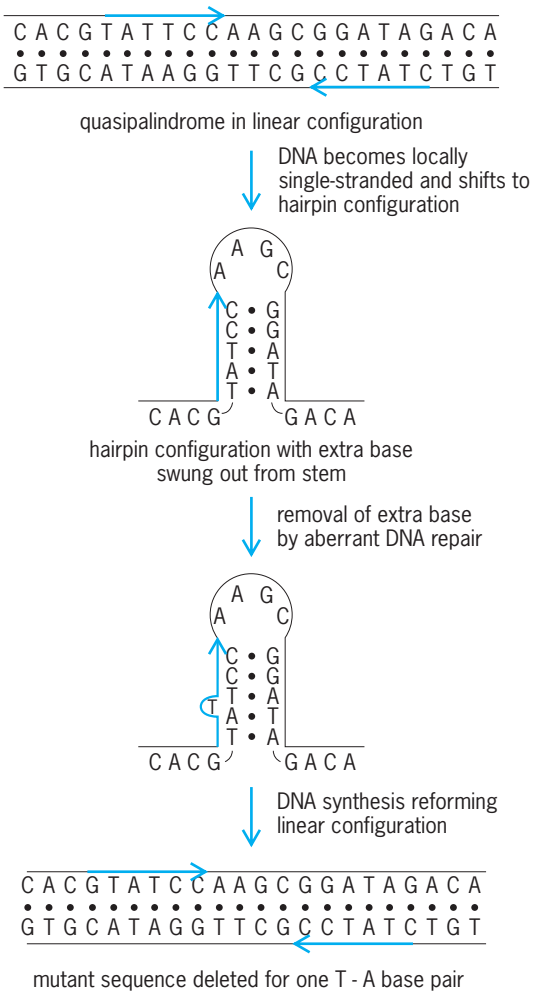

Fig. 4. Palindrome-mediated frameshift mutagenesis. Color indicates imperfectly palindromic sequences.

that the mechanisms discussed here have been set in the hairpin context, but may actually occur by topologically related misalignments at DNA replication forks.)

Mispairing mutagenesis. The normal pairing between the DNA bases (A · T amd G · C) is shown in **Fig. 5***a* and *b*. DNA bases can also pair incorrectly, usually because of transient changes in base structure. An early mispairing proposal invoking base tautomers created by proton migration is shown in Fig. 5*c*. Structural studies now suggest mispairing via "wobble" configurations (Fig. 5*d*) or ionized bases (Fig. 5*e*). Note, however, that the next round of pairing by these bases is likely to be normal, generating one mutant and one nonmutant progeny DNA double helix. *See* PURINE; PYRIMIDINE.

In order to generate a transversion, either two purines or two pyrimidines must mispair. In their usual configurations, however, neither such pair approximates the normal dimensions of a DNA base pair. In practice, most transversions arise via purine·purine mispairs; one example is shown in **Fig. 6**. In a normal base pair, the purines are in the *anti* configuration, which means that their hexagonal ring points toward the complementary pyrimidine. Occasionally, however, a purine (adenine in this example) rotates 180° around its glycosidic bond (the bond leading from the base to the sugar-phosphate backbone) into the *syn* configuration, thus presenting a different part of itself for potential pairing via hydrogen bonds.

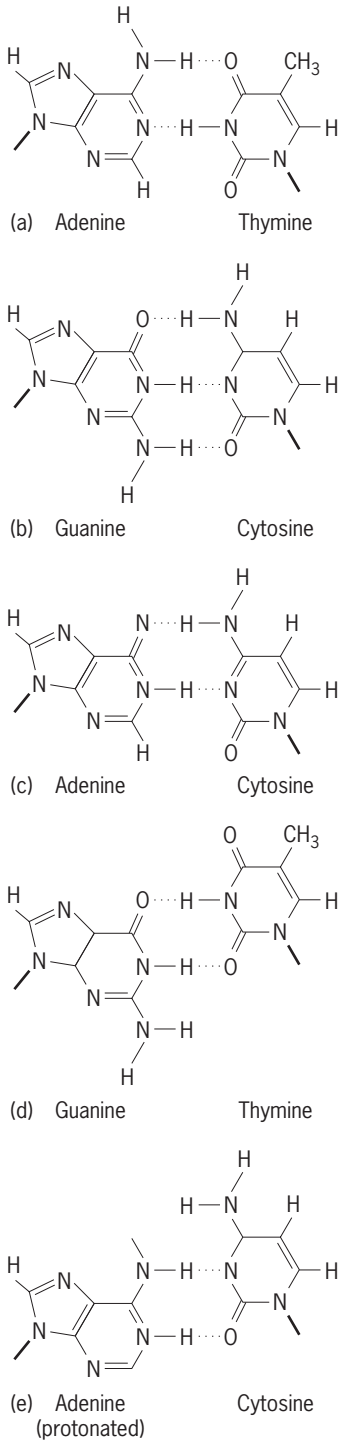

Fig. 5. Normal pairings and mispairings between the DNA bases. (*a*) The normal adenine·thymine base pair; hydrogen bonds are indicated by dotted lines and glycosidic bonds to the DNA backbone by heavy lines. (*b*) The normal guanine·cytosine base pair. (*c*) An adenine·cytosine mispair caused by proton migration in an adenine (arrow), leading to a transition mutation. (*d*) A guanine·thymine mispair involving wobble pairing. (*e*) An adenine·cytosine mispair involving both a protonated base and wobble pairing.

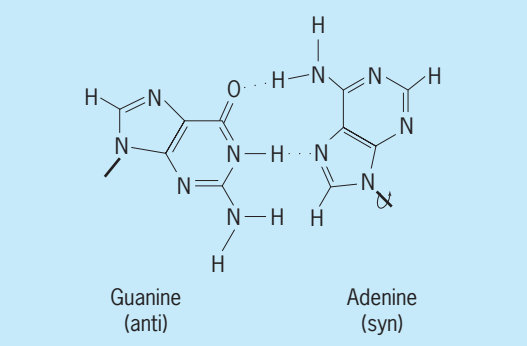

Fig. 6. A guanine·adenine mispair involving rotation of the adenine around its glycosidic bond. The mispair leads to a transversion mutation.

Chemical damage that alters the structure of a base can sometimes greatly enhance mispairing. Adding a methyl or ethyl group to the O^6 position of guanine, for instance, renders the guanine able to pair incorrectly with thymine (**Fig. 7***a*). The deamination of cytosine, which is promoted by heat and by low pH and which replaces the amino group with a keto group, generates uracil, which is identical to thymine in its base-pairing propensities (Fig. 7*b*). Thus, either of these treatments can induce $G \cdot C \rightarrow A \cdot T$ transitions, one via damaged guanine and the other via damaged cytosine.

Many kinds of base damage render bases unfit to pair at all. The dimerization of adjacent pyrimidines by ultraviolet irradiation, the addition of bulky groups to purines by mutagenic carcinogens such as the benzpyrines and aflatoxins, or the loss of a base (usually a purine) cause termination of DNA synthesis rather than mispairing. However, as noted below, mispairing can occur later at such a site.

Cause and prevention. Cells direct a number of repair operations against anomalies of DNA structure that would otherwise result in death or mutation. While some of these repair mechanisms circumvent death, others actually cause mutations.

Prevention. Most DNA damage is detected by excision repair systems before it can interfere with replication. These systems enzymatically remove the damaged bases and their associated sugar-phosphate backbones; the resulting gap is then filled in by DNA repair synthesis, using the complementary strand as a template. However, occasional damaged or transitorily rearranged bases enter the DNA replication fork. There they encounter DNA polymerase, the enzyme responsible for DNA synthesis. Most DNA polymerases are able to discriminate against base mispairs, although by ways that are still poorly understood. Even when a mispaired base has been covalently incorporated, however, it is often removed. This kind of repair can occur at two times. First, a 3′-exonuclease may specifically remove just-incorporated but mispaired nucleotides (nucleotide = base + sugar-phosphate). This enzymatic activity is called proofreading. Second, the few errors that escape proofreading may still be caught by a subsequent mismatch correction system. This system is remarkable in distinguishing between parental and progeny bases in newly synthesized DNA, so that the wrong rather than the right base can be excised.

Mutagenic translesion synthesis. When base damage renders pairing impossible and DNA synthesis terminates, it often reinitiates farther on, leaving behind a gap. Such a gap opposite a useless base is very apt to be lethal, certainly to the cell that inherits the damaged chromosome. Often, however, the cell is able to call up postreplication repair systems to save itself (SOS systems). Some such systems use recombination and are relatively error-free. Others, however, appear to use error-prone translesion synthesis, which is highly mutagenic because the requisite genetic information is unavailable at the damaged site. Such bypass synthesis seems to be responsible for most mutations induced by chemicals and radiations, although perhaps not contributing substantially to spontaneous mutations.

Past and future. Enzymatic systems for preventing (and sometimes causing) mutations have long evolutionary histories, presumably dating back to the genesis of cells. Therefore, mutation rates have evolved to levels that balance the few mutations that are useful, the much greater number that are deleterious, and the investment required to further reduce error rates. In DNA-based organisms ranging from small bacteriophages to the fungus *Neurospora crassa*, this balance produces about 0.3% mutations per genome per replication; organisms with larger genomes have smaller spontaneous mutation rates per base pair. In RNA viruses, the balance produces about one mutation per genome per replication. Spontaneous mutation rates are much less widely sampled in higher organisms, but in the fruitfly there

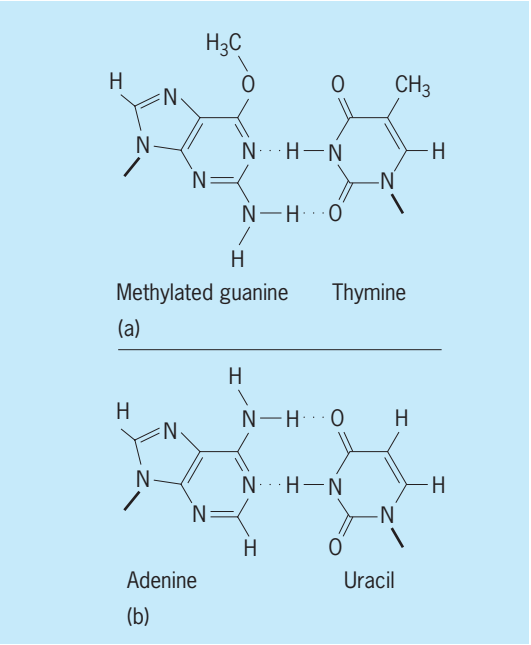

Fig. 7. Mispairing of DNA bases following chemical damage. (*a*) A guanine·thymine mispair promoted by methylation of a guanine (*b*) An adenine·uracil mispair in which the uracil was formed by the deamination of cytosine.

is roughly one new deleterious mutation per diploid genome per sexual generation. The large majority of these mutations, however, are only mildly deleterious, and would be unrecognizable in individual humans; they might, for instance, slightly shorten life-spans, slightly increase the frequency of infections by lowering body resistance, or affect health in other equally diffuse ways.

With the growing recognition (since about 1950) that many chemicals are mutagenic, and that many humans (between about 5 and perhaps 25%, depending on the criteria used) suffer from mutations of ancient or recent origin, much attention has been directed to discovering the mutagenic components of the environment. Many artificial chemicals are mutagenic, although these are still only a small percent of the total; and as a part of their ongoing war against predators, most plants elaborate chemicals that are mutagenic for at least some organisms. The impact of these exposures is largely unassessed, but intensive ongoing research is likely to provide some concrete answers. *See* MUTAGENS AND CARCINOGENS.

While the search for ways to prevent unwanted mutation continues vigorously, other searches seek methods to introduce specific, predetermined, desirable mutations into organisms of economic importance. By using the methods of molecular genetics, it is now usually possible to introduce specific mutations into specific genes in order, for instance, to increase the production of some marketable product. However, this step requires considerable prior knowledge about the organism's genetic and biochemical makeup. A more distant possibility is intervention in the human germ line in order, for instance, to cure specific genetic diseases. *See* GENETIC ENGINEERING; GENETICS; HUMAN GENETICS.

John W. Drake

Bibliography. J. W. Drake, B. W. Glickman, and L. S. Ripley, Updating the theory of mutation, *Amer. Sci.*, 71:621–630, 1983; M. E. Lambert et al. (eds.), *Eukaryotic Transposable Elements as Mutagenic Agents*, 1988.

Mutualism

An interaction between two species that benefits both. Individuals that interact with mutualists experience higher sucess than those that do not. Hence, behaving mutualistically is advantageous to the individual, and it does not require any concern for the well-being of the partner. At one time, mutualisms were thought to be rare curiosities primarily of interest to natural historians. However, it is now believed that every species is involved in one or more mutualisms. Mutualisms are thought to lie at the root of phenomena as diverse as the origin of the eukaryotic cell, the diversification of flowering plants, and the pattern of elevated species diversity in tropical forests.

Mutualisms generally involve an exchange of substances or services that organisms would find difficult or impossible to obtain for themselves. In many cases, mutualists make limited but essential nutrients available to their partners. For instance, *Rhizobium* bacteria found in nodules on the roots of many legume (bean) species fix atmospheric nitrogen into a form (NH_3) that can be taken up by plants. The plant provides the bacteria with carbon in the form of dicarboxylic acids. The carbon is utilized by the bacteria as energy for nitrogen fixation. Consequently, leguminous plants often thrive in nitrogen-poor environments where other plants cannot persist. Another well-known example is lichens, in which fungi take up carbon fixed during photosynthesis of their algae associates. *See* NITROGEN FIXATION.

A second benefit offered within some mutualisms is transportation. Prominent among these mutualisms is biotic pollination, in which certain animals visit flowers to obtain resources (usually food, in the form of nectar) and return a benefit by transporting pollen between the flowers they visit. Biotic seed dispersal, in which animals disperse the seeds of the plants whose fruits they consume, is another example of transportation-based mutualism. A final benefit is protection from one's enemies. For example, ants attack the predators and parasites of certain aphids in exchange for access to the aphids' carbohydrate-rich excretions (honeydew).

Another consideration about mutualisms is whether they are symbiotic. Two species found in intimate physical association for most or all of their lifetimes are considered to be in symbiosis. Not all symbioses are mutualistic; symbioses may benefit both, one, or neither of the partners. Lichens and plant-*Rhizobium* associations are symbiotic mutualisms, whereas plant-pollinator, plant–seed disperser, and ant-aphid associations are nonsymbiotic mutualisms. In some cases, the symbionts become so closely integrated that their individuality becomes difficult and ultimately impossible to distinguish. The eukaryotic cell is now believed to have originated as a symbiosis between a primitive cell and a bacterium ancestral to modern-day mitochondria.

Mutualisms can also be characterized as obligate or facultative (depending on whether or not the partners can survive without each other), and as specialized or generalized (depending on how many species can confer the benefit in question). An obligate, highly specialized mutualism exists between fig trees and the fig wasp pollinators. There is a nearly one-to-one match between fig and fig wasp species, and neither can reproduce in the absence of its partner. In this case, there has been a long history of coevolution; that is, reciprocal evolutionary pressures have led to adaptations of both partners that facilitate the interaction. Most mutualisms are considerably looser than this, however, and the degree to which they are the result of coevolution is more questionable.

Clearly, mutualisms are extremely diverse interactions. However, there are two features common to most mutualisms. First, mutualisms are highly variable in time and space. For example, in the interaction in which ants protect aphids from their

enemies, the benefits to aphids are huge in times and places where their enemies are abundant, but essentially nonexistent where enemies are rare. Furthermore, under some conditions ants will consume aphids rather than protecting them, leading the costs of the interaction to exceed its benefits. Second, mutualisms are susceptible to cheating. Cheaters can be individuals of the mutualist species that profit from their partners' actions without offering anything in return, or else other species that invade the mutualism for their own gain. The best-studied cheaters are nectar-robbing bees, which obtain nectar quickly and efficiently by chewing holes in flowers and thus neither collect nor deposit pollen. Evolutionary biologists are increasingly interested in how mutualistic behaviors, and hence mutualisms, can persist, given the apparent benefits of being a cheater.

An understanding of mutualism has considerable practical significance. Certain mutualisms play central roles in humans' ability to feed the growing population. Cattle and sheep can consume grass because of their mutualisms with gut-inhabiting bacteria that degrade cellulose; soybeans, which make up half the world's grain legume production, are typically inoculated with *Rhizobium* to increase yields. It has been estimated that half the food consumed is the product of biotic pollination. Other mutualisms are critical in structuring natural communities and hence are of great interest from a conservation perspective. For example, fruits produced as a result of the obligate pollination mutualism between figs and fig wasps are relied upon by nearly every fruit-feeding bird and mammal in some tropical forests during times of food scarcity. These animals in turn disperse the seeds of other plants and thus help sustain populations of predators and parasites. Disruption of the pollination mutualism could therefore constitute a serious threat to local biodiversity. Unfortunately, mutualisms are apparently rather sensitive to human-induced environmental changes such as habitat fragmentation, atmospheric pollution, and the introduction of nonnative species. Rather than focusing solely on species in isolation, conservationists increasingly attempt to preserve the critical mutualisms in which endangered organisms are involved. *See* ECOLOGY; PLANT PATHOLOGY. Judith L. Bronstein

Bibliography. D. H. Boucher (ed.), *The Biology of Mutualism*, Oxford University Press, 1985; J. L. Bronstein, Our current understanding of mutualism, *Quart. Rev. Biol.*, 69:31–51,1994; A. E. Douglas, *Symbiotic Interactions*, Oxford University Press,1994; J. N. Thompson, *The Coevolutionary Process*, University of Chicago Press,1994.

Myasthenia gravis

A disease resulting from an abnormality in neuromuscular transmission, characterized by a fluctuating degree of muscle weakness. The weakness is usually aggravated by activity, and there is partial or complete restoration of strength after a period of rest or the administration of anticholinesterase medications. *See* ACETYLCHOLINE.

Etiology. It has been shown that the basic defect in myasthenia gravis is a reduction in the number of acetylcholine receptor sites in the postsynaptic membrane of the neuromuscular junction. It has also been shown that in myasthenia gravis, the individuals often have immunoglobulins in the serum that partially block acetylcholine receptors, and that myasthenic changes can be produced in mice by the injection of serum immunoglobulin from myasthenic patients. The antiacetylcholine receptor antibody is thought to originate in the thymus gland, and the antigenic stimulus is believed to be the acetylcholine receptors in the myoid cells in the thymus gland. *See* AUTOIMMUNITY; IMMUNOGLOBULIN.

Pathology. Abnormalities have been demonstrated in the thymus gland and skeletal muscle in myasthenia gravis. There is an increased incidence of thymoma in myasthenia gravis, and in those without a thymoma, hyperplasia of the germinal centers is a common finding in the thymus gland.

Muscle from myasthenic persons shows variable histologic findings. In some it is completely normal. In others, collections of lymphocytelike cells are present in otherwise normal muscle. In still others patterns suggesting neurogenic atrophy or a myopathy are found. However, by electron microscopy, a simplification of the secondary synaptic clefts and a widening of the primary synaptic cleft of the motor end plate has been demonstrated.

Clinical features. Although the disease affects young women more commonly, usually in the third decade, it can occur in either sex at any age. In approximately 10% of those affected, myasthenia gravis is associated with a thymoma, and the prognosis in such individuals tends to be worse. In the majority of persons, weakness affects muscles of the head, neck, and limbs (generalized myasthenia), but in some the weakness is restricted to the muscles of the eyes (ocular myasthenia), in which case the disease is usually benign. Fifteen percent of babies born of myasthenic mothers have symptoms requiring treatment at birth, but these are transient and disappear in a few weeks (neonatal myasthenia).

The muscles most often affected in generalized myasthenia are those of the eyes, with ptosis and double vision (diplopia), of mastication, swallowing, and speech, and of respiration. Limb muscles are also frequently affected.

Initially, the affected muscles have normal or nearly normal strength. However, with repeated use, progressive weakness develops. Thus, ptosis and diplopia become worse as the day progresses, speech becomes slurred and nasal with continued conversation, or the individual may be unable to finish eating because of increasing difficulty with chewing and swallowing. These symptoms are often only partially reversed with rest. In some individuals myasthenia gravis is a life-threatening illness because of respiratory weakness.

Diagnosis. The diagnosis of myasthenia gravis is based on the demonstration of increasing muscle

weakness with exertion, and reversal of the weakness after the administration of anticholinesterase drugs.

Treatment. The standard treatment for myasthenia gravis has been the use of longer-acting anticholinesterase agents; thymectomy and immunosuppressive drugs are reserved for those individuals with generalized myasthenia that does not respond sufficiently to these agents. With the demonstration that myasthenia gravis is due to thymus-originated antireceptor antibody, thymectomy has been resorted to earlier in the course of the illness, often combined with corticosteroids or other immunosuppressive therapy. In those individuals with significant persistent symptoms, plasmapheresis has also become another mode of treatment. When a thymoma is detected, thymectomy is mandatory.

Persons being treated with anticholinesterase drugs may develop periods of severe muscle weakness with respiratory paralysis. These paralytic periods may be due to inadequate dosage (myasthenic crisis) of, or overmedication (cholinergic crisis) with, these agents, but most commonly they appear to result from development of resistance to the drugs. *See* SYNAPTIC TRANSMISSION. S. Mark Sumi

Bibliography. M. H. De Baets and H. J. Oosterhuis (eds.), *Myasthenia Gravis*, 1993.

Mycobacterial diseases

Diseases caused by mycobacteria, a diffuse group of acid-fast, rod-shaped bacteria in the genus *Mycobacterium*. Some mycobacteria are saprophytes, while others can cause disease in humans. The two most important species are *M. tuberculosis* (the cause of tuberculosis) and *M. leprae* (the cause of leprosy); other species have been called by several names, particularly the atypical mycobacteria or the nontuberculous mycobacteria. This article deals mainly with nontuberculous mycobacteria. *See* LEPROSY; TUBERCULOSIS.

These bacteria are classified according to their pigment formation, rate of growth, and colony morphology. The organisms, and the diseases they cause, are summarized in the **table**. The most commonly involved disease site is the lungs. Nontuberculous mycobacteria are transmitted from natural sources in the environment, rather than from person to person, and thus are not a public health hazard. Nontuberculous mycobacteria have been cultured from various environmental sources. For example, *M. avium intracellulare* has been found in soil and dust; *M. kansasii* has been isolated from water samples; *M. marinum* has been detected in fresh and salt water; and *M. chelonei* has been found in soil and dust. The environmental sources of many other species are unknown.

The diagnosis of disease caused by nontuberculous mycobacteria can be difficult, since colonization or contamination of specimens may be present rather than true infection.

Pulmonary disease. Pulmonary disease resembling tuberculosis is a most important manifestation of disease caused by nontuberculous mycobacteria. The symptoms and chest x-ray findings are similar to those seen in tuberculosis. *Mycobacterium kansasii* and *M. avium intracellulare* are the most common pathogens, followed by *M. fortuitum* and *M. chelonei*. The disease usually occurs in middle-aged men and women with some type of chronic coexisting lung disease such as chronic obstructive pulmonary disease (emphysema), chronic bronchitis, pneumoconiosis (for example, silicosis), or previous tuberculosis. Although the pathogenic mechanisms are obscure, it is presumed that there is an involvement with inhalation of droplet nuclei containing organisms, followed by mycobacterial proliferation in areas of destroyed lung.

Pulmonary infections due to *M. kansasii* can be treated successfully with chemotherapy. The treatment of pulmonary infections due to *M. avium intracellulare* complex is difficult. These organisms are highly resistant to chemotherapy even with multiple drug regimens. Surgical resection of localized disease is considered in selected individuals.

Bone and joint disease. Chronic infection involving joints and bones, bursae, synovia, and tendon sheaths can be caused by various species. Infection usually occurs by direct inoculation of the organism following trauma, surgery, or a joint injection. *Mycobacterium fortuitum*, *M. chelonei*, *M. kansasii*, and *M. avium intracellulare* are the usual pathogens. Treatment of bone and joint infection consists of surgical drainage and removal of tissue. Drug therapy may be useful for the relatively sensitive *M. kansasii*.

Nontuberculous (atypical) mycobacteria

Group	Classification	Characteristics	Representative species	Type of disease
I	Photochromogens	Yellow pigment production when exposed to light	*Mycobacterium kansasii* *M. marinum*	Pulmonary or extrapulmonary Swimming pool granuloma
II	Scotochromogens	Pigment production in dark	*M. scrofulaceum* *M. gordonae*	Cervical adenitis Saprophytic tap-water bacilli
III	Nonphotochromogens *(Battey bacilli)*	No pigment production	*M. avium intracellulare* complex	Pulmonary or extrapulmonary
IV	Rapid growers	Growth in 3–5 days	*M. fortuitum* *M. chelonei*	Saprophytic Wound abscesses

Skin and soft tissue infection. Localized abscesses due to *M. fortuitum* or *M. chelonei* can occur after trauma, after surgical incision, or at injection sites. The usual treatment is surgical incision.

The most common soft tissue infection is caused by *M. marinum*, which may be introduced, following an abrasion or trauma, from handling fish or fish tanks, or around a swimming pool. Treatment is surgical.

Mycobacterium ulcerans causes a destructive skin infection in tropical areas of the world. It is treated by wide excision and skin grafting.

Lymphadenitis. Lymph-node enlargement and inflammation due to *M. avium intracellulare* or *M. scrofulaceum* may occur in children between the ages of 1 and 5 years. Usually, only the submandibular nodes in the neck are involved, but sometimes axillary or inguinal nodes are also affected. The child is usually asymptomatic, but node rupture with drainage and sinus formation can occur rapidly. The condition heals with fibrosis and calcification. Exact pathogenesis is unclear, but the mouth or throat might be the portal of entry. Therapy consists of excising the involved group of nodes.

Disseminated disease. Disseminated disease due to nontuberculous mycobacteria is usually associated with *M. avium intracellulare* and *M. kansasii*. Dissemination occurs most often in individuals with defects in immunity, such as those having leukemia, immunosuppressive drug therapy, chronic renal failure, or malignancy.

Disseminated *M. avium intracellulare* is one of the opportunistic infections seen in the acquired immune deficiency syndrome (AIDS). In individuals with AIDS, the organism has been cultured from lung, brain, cerebrospinal fluid, liver, spleen, intestinal mucosa, and bone marrow. No treatment has yet been effective in this setting. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS).

George M. Lordi; Lee B. Reichman

Bibliography. J. McFadden (ed.), *Molecular Biology of the Mycobacteria*, 1990; J. L. Stanford, *The Biology of the Mycobacterium, III. Clinical Aspects of Mycobacterial Disease*, 1984.

Mycology

The study of organisms of the fungal lineage, including mushrooms, boletes, bracket or shelf fungi, powdery mildew, bread molds, yeasts, puffballs, morels, stinkhorns, truffles, smuts, and rusts. Fungi, in the traditional sense, are an ecological grouping of organisms found in every ecological niche. Mycologists estimate that there are 1.5 million species of fungi, with only 70,000 species now described. Fungi typically have a filamentous-branched somatic structure surrounded by thick cell walls; these structures are known as hyphae. *See* FUNGI; MUSHROOM.

Taxonomy. The names adopted for major taxa (groups) of fungi and fungi-like organisms are as follows:

Eumycota (true fungi)
- Chytridiomycota (Chytridiomycetes)
- Zygomycota (Zygomycetes, Trichomycetes)
- Ascomycota (Archiascomycetes, Hemiascomycetes, Euascomycetes)
- Basidiomycota (Urediniomycetes, Ustilaginomycetes, Hymenomycetes)
- Deuteromycota (Agonomycetes, Balstomycetes, Coelomycetes, Hyphomycetes)

Pseudomycota (fungi-like organisms)
- Oomycota (Oomycetes)
- Hyphochytriomycota (Hyphochytriomycetes)
- Plasmodiophoromycota (Plasmodiophoromycetes)

Myxomycota (slime molds)
- Myxomycota (Mycomycetes)

See CLASSIFICATION, BIOLOGICAL; FUNGAL ECOLOGY.

Description. Fungi are heterotrophic, meaning they cannot make their own food as plants do. They are osmotrophic or nutrition-absorptive, obtaining their food by releasing enzymes into their environment to break down complex organic compounds. Fungi play a significant role in nature as decomposers of wood and wood products, and in forest ecosystems fungi release nutrients back to soil. Fungi also decompose fabrics, leather goods, petroleum products, and foodstuffs. They reproduce through spores produced sexually and asexually. *See* FOREST ECOSYSTEM; WOOD DEGREDATION.

Fungi are often viewed as detrimental to plants, animals, and humans. They cause nearly 80% of all plant disease and, despite the extensive use of fungicides and other control measures, the financial loss due to fungal activity remains enormous. In humans, fungal spores cause allergies in certain individuals; therefore, mold spore levels are carefully monitored and reported. With an increasing immunocompromised population due to human immunodeficiency virus (HIV), malignancies, chemotherapies, and organ transplants, the occurrence of new opportunistic fungal infections has risen. However, along with the increase in human mycoses comes realistic hope for improved diagnosis and therapy. *See* ERGOT AND ERGOTISM; FUNGISTAT AND FUNGICIDE; MEDICAL MYCOLOGY; PLANT PATHOLOGY.

Food. Because of their rapid growth rate and high protein content, fungi are an ideal source of protein for animals and humans. Many mushrooms are edible, and some are cultivated for the canned and fresh market. There are many foods, especially fermented products, that are produced by fungi. The food industry utilizes fungi for production of natural flavor and color compounds.

Uses. Secondary metabolites produced by fungi have been shown to have medicinal uses. Penicillin, an antibiotic extracted from fungi, revolutionized medicine, and many more antibiotics produced by fungi have been discovered. Many herbal medicines are derived from fungi. In addition to their antibiotic secondary metabolites, fungi have a broad potential for producing medically useful

compounds such as metabolic regulators, antitumor drugs, and immunomodulators that enhance host resistance against cancer, immunodeficiency disease, or generalized immunosuppression. *See* ANTIBIOTIC.

Fungi do have some positive roles in agriculture. As an alternative to chemical pesticides, fungi are being tested as biocontrol agents to reduce the crop damage from insects, nematodes, weeds, and plant pathogenic microorganisms. Biofertilizers are specially formulated inoculants of mycorrhizal fungi used to improve crop growth in nutrient-deficient soils. *See* FERTILIZING.

There are many uses for fungi in industry. Biological delignification with fungi can remove the lignin component from wood without destroying the cellulose and hemicellulose important in paper production. Fungal polysaccharide-degrading enzymes can also be used instead of chemicals for removal of noncellulosic materials in the retting of flax fibers, producing high-quality yarns. A wide range of fungi are useful in converting coal to a liquid or to gaseous products that would normally require high temperatures and pressures which are expensive and difficult to achieve. Fungi are also a useful alternative to the conventional scavenging techniques for the removal of heavy metals from dilute solutions, such as electroplating and chrome tannery effluents.

The application of recombinant DNA techniques has greatly increased the potential use of fungi in biotechnology. New genetic material inserted into chromosomes or extrachromosomal elements directs the transformed fungus to metabolize or produce specific products. Because they are capable of secreting large quantities of certain proteins in liquid culture, fungi have proven to be useful cloning hosts for the production of recombinant proteins of both fungal and human origin. *See* DEOXYRIBONUCLEIC ACID (DNA); FUNGAL BIOTECHNOLOGY; FUNGAL GENETICS.

Shung-Chang Jong

Bibliography. C. J. Alexopoulos, C. W. Mims, and M. Blackwell, *Introductory Mycology*, 4th ed., Wiley, New York, 1996; Z. An (ed.), *Handbook of Industrial Mycology*, Marcel Dekker, New York, 2005; M. J. Carlile, S. C. Watkinson, and G. W. Gooday, *The Fungi*, 2d ed., Academic Press, New York, 2001; P. M. Kirk et al., *Dictionary of the Fungi*, 9th ed., CAB International, Wallingford, CT, 2001; G. M. Mueller, G. F. Bills, and M. S. Foster (eds.), *Biodiversity of Fungi: Inventory and Monitoring Methods*, Elsevier Academic Press, New York, 2004.

Mycoplasmas

The smallest prokaryotic microorganisms that are able to grow on cell-free artificial media. Their genome size is also among the smallest recorded in prokaryotes, about 5×10^8 to 10^9 daltons. The mycoplasmas differ from almost all other prokaryotes in lacking a rigid cell wall and in their incapability to synthesize peptidoglycan, an essential component of the bacterial cell wall.

Taxonomically, the mycoplasmas are assigned to a distinct class, the Mollicutes, containing two orders, Mycoplasmatales and Acholeplasmatales. The distinction between the orders is based primarily on differences in nutritional criteria: members of the Mycoplasmatales require cholesterol or other sterols for growth whereas those of the second order do not. The main criteria used for the subdivision of the orders into families and genera are shown:

Class: Mollicutes
- Order I: Mycoplasmatales (sterol required for growth; $NADH_2$ oxidase localized in cytoplasm)
  - Family I: Mycoplasmataceae (genome size approximately 5×10^8 daltons)
    - Genus I: *Mycoplasma* (70 species; do not hydrolyze urea)
    - Genus II: *Ureaplasma* (2 species; hydrolyze urea)
  - Family II: Spiroplasmataceae (helical organisms; genome size approximately 10^9 daltons)
    - Genus I: *Spiroplasma* (4 species)
- Order II: Acholeplasmatales (sterol not required for growth; $NADH_2$ oxidase localized in membrane; genome size approximately 10^9 daltons)
  - Family I: Acholeplasmataceae
    - Genus I: *Acholeplasma* (9 species)
- Genus of uncertain taxonomic position
  - *Anaeroplasma* (2 species)

The term mycoplasmas is generally used as the vernacular or trivial name for all members of the class Mollicutes, irrespective of the classification in a particular genus. *See* PROKARYOTAE.

Morphology and reproduction. The individual cells are pleomorphic, varying in shape from spherical or pear-shaped structures (about 0.3–0.8 micrometer in diameter) to branched filaments varying in length from approximately 5 to 150 μm (**illus.** *a*); the smallest viable cells will pass a membrane filter of 450 nm pore diameter. The organisms are gram-negative, nonsporing, and usually nonmotile, although some species of genus *Mycoplasma* show gliding motility on liquid-covered surfaces. Members of the family Spiroplasmataceae are characterized under most conditions by the production of helical filaments (illus. *b*) exhibiting rotatory and undulating motility. Replication is basically by binary fission of spherical cells, but cell division is not necessarily synchronized with replication of the genome. Thus, the development of filaments and their transformation into chains of spherical cells may be ascribed to a lack of synchrony between division of the genome and the cytoplasm.

Growth in liquid medium produces a barely visible to rather heavy turbidity. Colonies on agar are minute [from less than 0.04 in. (1 mm) up to 0.15 in. (4 mm) in diameter]; under suitable growth conditions the colonies of many species have a characteristic "fried

egg" appearance, consisting of an opaque, more or less yellowish central area which grows down into the medium, and a translucent, flat peripheral zone (illus. *c*).

Ultrastructure. Electron microscopy of ultrathin sections of mycoplasma cells shows that the cytoplasm is bounded only by a plasma membrane, about 8 to 10 nm thick, while there is no cell wall as is found in other prokaryotes (illus. *d*). The cell membrane of some species is covered by a capsular substance. The cytoplasm consists of nuclear material (DNA strands), ribosomes, and intracellular granules. Specialized structures such as terminal organelles, which play a role in attachment of the organism to the surface of host cells, have been demonstrated in several species.

Though the mycoplasmas have much in common with the wall-defective or wall-less L-phase variants of bacteria, it is now generally accepted that there are several fundamental differences separating the two groups of organisms.

Metabolism and biochemistry. The mycoplasmas have limited biosynthetic abilities and therefore require complex growth media, consisting of beef heart infusion, peptone, yeast extract, and serum. Serum serves as a source of many essential compounds, especially fatty acids and cholesterol in an assimilable and nontoxic form as required for membrane synthesis in the vast majority of species. Almost all mycoplasma species catabolize glucose or arginine, or both. Glucose and certain other carbohydrates serve as an energy source for fermentative mycoplasmas possessing the Embden-Meyerhof glycolytic pathway. *See* FERMENTATION.

Arginine degradation occurs by the arginine dihydrolase pathway that may be the major source for adenosine triphosphate (ATP) in nonfermentative mycoplasmas. Urease activity is a unique property of genus *Ureaplasma*. Lipases are involved in the development of characteristic precipitates (known as "film and spots") formed by certain species on horse-serum agar or egg-yolk medium. Methylene blue, tellurite, and 2,3,5-triphenyltetrazolium chloride are frequently reduced. Proteolytic activity, as demonstrated by clearing and/or liquefaction of coagulated serum, gelatin and casein, is characteristic of some species. Several species show phosphatase activity. Pigmented carotenoids are produced by some *Acholeplasma* species. Weak to strong α and β hemolysis on blood agar containing erythrocytes from horse, sheep, guinea pig, or chicken, is a property shared by most mycoplasmas; the hemolysin has been identified as peroxide.

Mycoplasmas are generally highly resistant to benzyl penicillin and other antibiotics which act by interfering with the biosynthesis of peptidoglycan. They are usually susceptible to antibiotics that specifically inhibit protein synthesis in prokaryotes, as for example tetracyclines and chloramphenicol. Susceptibility to other antibiotics, such as erythromycin and other macrolides, is variable. Most mycoplasmas tolerate 1:2000 to 1:4000 thallium acetate, which is often incorporated, together with penicillin, as a selective agent in mycoplasma media to inhibit the growth of contaminating bacteria. *See* BACTERIAL PHYSIOLOGY AND METABOLISM.

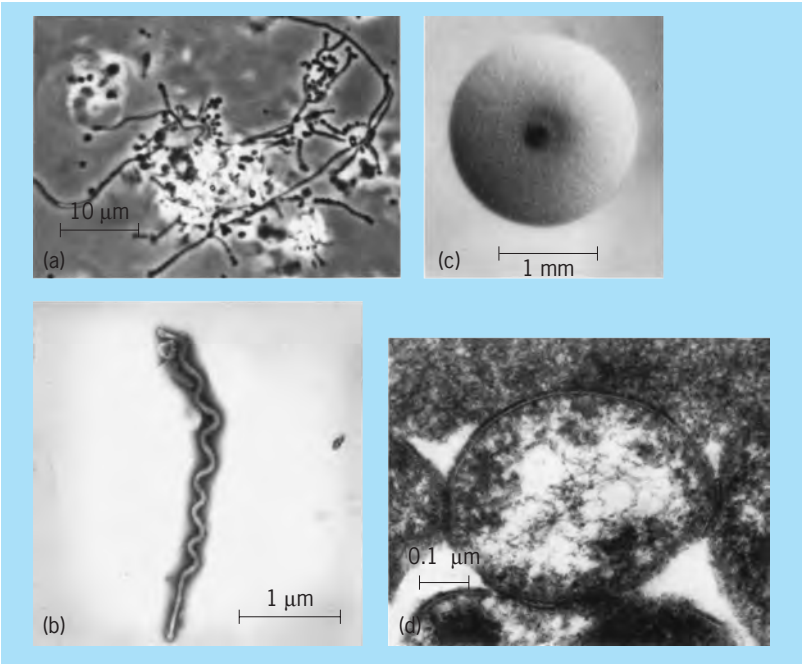

Mycoplasmas. (*a*) *Mycoplasma hominis*, phase contrast micrograph (*courtesy of W. Bredt*). (*b*) *Spiroplasma citri*, electron micrograph of negatively stained specimen (*courtesy of R. M. Cole*). (*c*) Mycoplasma colony on solid nutrient medium. (*d*) *Ureaplasma urealyticum*, electron micrograph of thin-sectioned cell showing plasma membrane and network of DNA strands (*courtesy of F. T. Black*).

Habitat and pathogenicity. The mycoplasmas are almost ubiquitous in nature. Several species are important pathogens of humans, animals and plants, while others constitute part of the normal microbial flora of, for example, the upper respiratory and lower urogenital tracts of humans.

Mycoplasma pneumoniae, previously regarded as a virus and known under the name of the Eaton agent, was found in 1962 to be the cause of cold agglutinin-associated primary atypical pneumonia. This disease is particularly frequent in the 5–15-year age group; it is probably endemic almost all over the world and often reaches epidemic proportions at intervals of 4 to 5 years. *Mycoplasma hominis*, which is a common inhabitant of the lower genital tract, is a potential pathogen and appears to be a relatively frequent cause of acute pelvic inflammatory disease (salpingitis and parametritis) and postpartum septicemia, and a less frequent cause of certain other infectious diseases, especially in an immunocompromised host. The human mycoplasma, *M. genitalium*, is another candidate as a possible cause of salpingitis and related disorders. A number of species, including *M. salivarium* and *M. orale*, are common members of the oral microflora. The role of *M. fermentans*, found rarely in the human genital tract, remains obscure. The same holds true of *Ureaplasma urealyticum* (previously known as T-mycoplasmas; T for tiny) that includes more than ten different serovars that are very common inhabitants of the genital tracts of males and females. Urethritis can be produced experimentally in

humans with *U. urea lyticum*, and the organism may be the etiological agent of some cases of nongonococcal urethritis.

A variety of the type species *M. mycoides* is the etiological agent of contagious pleuropneumonia of cattle and water buffaloes, a widespread infection of great socioeconomic importance, particularly in many countries in Africa, and in parts of India. Contagious pleuropneumonia and other infectious diseases of goats are caused by related organisms. *Mycoplasma agalactiae* causes contagious agalactia of goats and sheep, whereas *M. bovis* is an important cause of several diseases of cattle, including mastitis and arthritis. Lesions of the urogenital tract of cattle may be caused by this organism as well as by *M. bovigenitalium*. In pigs, *M. hyopneumoniae* is the most important pathogen, causing enzootic pneumonia, while *M. hyosynoviae* is a less frequent cause of a nonsuppurative, proliferative type of arthritis in piglets. *Mycoplasma gallisepticum* is responsible for chronic respiratory disease (tracheitis, air sacculitis, sinusitis) in chickens and sinusitis in turkeys and may cause severe economic loss; lesions in the brain associated with polyarteritis of the cerebral and meningeal arteries and arterioles have been reported in turkeys under natural and experimental conditions. Other pathogens for chickens and turkeys are *M. synoviae* and *M. meleagridis*. Though several species have been demonstrated in horses, dogs and cats, the possible role of these species in disease of their animal hosts is undetermined as yet. Among the mycoplasmas frequently occurring in small rodents, *M. arthritidis* is a cause of purulent polyarthritis in rats, while *M. pulmonis* is highly pathogenic to both mice and rats, causing chronic respiratory disease, associated in up to 30% of female rats with oophoritis and salpingitis; acute purulent arthritis, that may lead to chronic lesions, can be produced in both rats and mice following intravenous inoculation. *Mycoplasma neurolyticum*, recovered from mice, is remarkable in that it produces a potent exotoxin having severe effects, under experimental conditions, on the central nervous system of mice and, to a lesser extent, of rats.

Members of the genus *Acholeplasma* were first isolated from sewage, compost and soil, but were later found to have a widespread distribution in a variety of tissues and secretions of almost every type of vertebrate examined. In addition, isolation of acholeplasmas from plant material is being reported with increasing frequency. Yet there is no clear evidence at present that *Acholeplasma* species play a pathogenic role in either animals or plants.

The mycoplasmas classified in genus *Anaeroplasma* are strict anaerobes, presumedly nonpathogenic, found in the rumen of cattle and sheep.

The genus *Spiroplasma* includes a great variety of different species and serogroups found in vascular plant fluids (phloem), on the surface of leaves, flowers and fruits, and in the hemolymph and guts of insects that feed on plants. *Spiroplasma citri* is pathogenic for citrus plants (grapefruit and orange) producing "stubborn" disease, a disease of considerable economic importance. Another spiroplasma is the cause of corn stunt disease. Although most spiroplasmas multiply in the tissues of insects without producing overt disease, one species, *S. apis*, is the cause of a devasting lethal infection of honeybees. *S. mirum* isolated from rabbit ticks produces ocular and nervous system disease on experimental infection of suckling small rodents and rabbits. *See* PLANT PATHOLOGY.

Mycoplasmalike organisms. Wall-less nonhelical mycoplasmalike organisms (MLO) can be demonstrated in great numbers by electron microscopy within sieve tube elements of plants affected by yellows diseases and in the salivary glands of insects serving as vectors of these diseases. Tetracycline antibiotics, to which mycoplasmas are known to be susceptible, have a marked suppressive effect on yellows plant diseases and on the number of mycoplasmalike organisms demonstrable in the phloem of plants and tissues of insect vectors. However, final classification of the mycoplasmalike organisms with the mycoplasmas (Mollicutes) must await cultivation of the organisms on artificial media. E. A. Freundt

Bibliography. M. F. Barile et al. (eds.), *The Mycoplasmas*, vols. 1–3, 1979; R. Dickesson, *Diagnostic Electron Microscopy: A Text-Atlas*, 1988; I. Kahane and A. Adoni (eds.), *Rapid Diagnosis of Mycoplasmas*, 1994; J. Maniloff et al., *Mycoplasmas: Molecular Biology and Pathenogenesis*, 1992.

Mycorrhizae

Dual organs of absorption that are formed when symbiotic fungi inhabit healthy absorbing organs (roots, rhizomes, or thalli) of most terrestrial plants and many aquatics and epiphytes.

Mycorrhizae appear in the earliest fossil record of terrestrial plant roots. Roughly 80% of the nearly 10,000 plant species that have been examined are mycorrhizal. Present-day plants that normally lack mycorrhizae are generally evolutionarily advanced. It has been inferred that primitive plants evolved with a symbiosis between fungi and rhizoids or roots as a means to extract nutrients and water from soil.

Mycorrhizal dependency. The relative dependence of plants on mycorrhizal fungi has been viewed in different ways. Ecological dependence denotes that the plant must form mycorrhizae to survive, compete, and reproduce. Economic dependence means that the plant must form mycorrhizae to produce a commodity. The two types of dependence do not always coincide. A desert annual may need mycorrhizae to survive and produce seed, whereas some hay grasses may produce a large biomass without mycorrhiza formation.

The degree of dependence varies between species or groups of plants. In absolute dependence, characteristic of perennial, terrestrial plants, the host requires mycorrhizae to survive. Some plants are facultative; they may form mycorrhizae but do not

always require them. This group includes many of the world's more troublesome weeds. A minority of plant species characteristically lack mycorrhizae, so far as is known, including many aquatics, epiphytes, and annual weeds.

Functions. The three major types of mycorrhizae differ in structural details but have many functions in common. The fungus colonizes the cortex of the host root and grows its filaments (hyphae) into surrounding soil from a few centimeters to a meter or more. The hyphae absorb nutrients and water and transport them to host roots. The fungi thus tap far greater volumes of soil at a relatively lower energy cost than the roots could on their own. Moreover, many, if not all, mycorrhizal fungi produce extracellular enzymes and organic acids that release immobile elements such as phosphorus and zinc from clay particles, or phosphorus and nitrogen bound in organic matter. The fungi are far more physiologically capable in extracting or recycling nutrients in this way than the rootlets themselves.

Mycorrhizal fungi are relatively poorly competent in extracting carbon from organic matter. They derive energy from host-photosynthesized carbohydrates. Hosts also provide vitamins and other growth regulators that the fungi need.

Types. The major types are ectomycorrhizae, vesicular-arbuscular mycorrhizae, and ericoid mycorrhizae.

Ectomycorrhizae. Ectomycorrhizae, the most readily observed type, and some related variants form with members of the Pinaceae, Betulaceae, Fagaceae, Salicaceae, Myrtaceae, and several other families of woody perennials. Ectomycorrhizal hosts strongly depend on mycorrhizae to survive. Relatively few in number of species, they nonetheless dominate most forests outside the tropics.

The species of ectomycorrhizal fungi number in the thousands, being mostly mushroom or truffle formers in their reproductive stage (Ascomycotina, Basidiomycotina, and a few Zygomycotina). Many are host-specific, forming mycorrhizae only with certain trees such as alders, Douglas-fir, or eucalypts, respectively. Others colonize roots of a wide variety of hosts. Spores of ectomycorrhizal fungi are dispersed from mushrooms by air or, for the belowground fruiting truffles and trufflelike fungi, by being eaten by animals which later excrete them.

Ectomycorrhizae form on tiny host feeder rootlets. The fungus envelops the rootlet with a hyphal sheath, which lends its color to the rootlet surface. The color is often white but ranges to brown, yellow, pink, green, blue, or black, and is easily seen with a hand lens. Brightly colored hyphae can often be seen extending from the mycorrhiza into the surrounding substrate. They also grow from the sheath into the rootlet between the cells of the epidermis and cortex. In this region of intimate intergrowth, the fungi exchange nutrients and water for photosynthates.

Ectomycorrhizal fungi produce auxins and cytokinins that change morphology of host rootlets. A nonmycorrhizal rootlet of these hosts, uncommon in nature, has a single tip and abundant root hairs and lives only a few months at best. When colonized by the ectomycorrhizal fungus, however, the rootlet usually branches and root-hair formation is suppressed. Ectomycorrhizae may be active for years if they do not shrivel from drought or freezing. The fungi also protect rootlets from pathogens by surrounding them with the hyphal mantle as a physical barrier and often by producing antibiotics as a chemical deterrent.

Vesicular-arbuscular. Vesicular-arbuscular mycorrhizae (sometimes simply termed arbuscular mycorrhizae) form with the great majority of terrestrial herbaceous plant species plus nearly all woody perennials that are not ectomycorrhizal (except for the Ericaceae, the heath family, as discussed below). Some ectomycorrhizal hosts, such as alders, eucalypts, and willows, can also form vesicular-arbuscular mycorrhizae. Vesicular-arbuscular mycorrhizal hosts range from strongly mycorrhiza-dependent, especially the woody perennials, to faculative, as are many grasses. Vesicular-arbuscular mycorrhizae are difficult to observe without special biological staining procedures, because the inconspicuous hyphae grow diffusely from the rootlets into the soil.

The fungal associates of vesicular-arbuscular mycorrhizae are members of the order Glomales of the Zygomycotina and, as presently known, number only about 200 species around the world. In general, vesicular-arbuscular mycorrhizal fungi show little or no host specificity, although different species may interact with hosts in different ways in different habitats. Such fungi sporulate as individual spores or clusters of spores produced among roots and are spread by movement of the soil in which they occur. The contrast between the low species numbers of vesicular-arbuscular-mycorrhizal fungi and the thousands of ectomycorrhizal fungi fuels speculation on evolution and ecology of the two kinds of associations. Judging from the fossil record, vesicular-arbuscular mycorrhizal fungi are more primitive than the ectomycorrhizal fungi and apparently have been relatively genetically stable for at least 400 million years. *See* ASCOMYCOTA; ZYGOMYCOTA.

Vesicular-arbuscular mycorrhizal fungi grow into the root cortex and penetrate root cells to form two kinds of specialized structures. All form arbuscules ("little trees") in some host cells, bushy structures that are organs for exchange of nutrients and carbon compounds between fungus and host. Arbuscules persist for several days, then are digested by the host cell. Many vesicular-arbuscular mycorrhizal fungi produce large, inflated bodies (vesicles) in other host cells. These fill with lipids, presumably as energy storage for the fungus. As the root senesces, vesicles may develop thickened walls to become asexual spores.

Vesicular-arbuscular mycorrhizal fungi seem neither to produce growth regulators nor to markedly change root configuration. Rather than suppress

root-hair formation, they often enter the rootlet through root hairs. They form no physical barrier to pathogens and show little or no antibiosis against pathogens.

Ericoid. Ericoid mycorrhizae are restricted to the Ericales, the heath order. The hosts are strongly mycorrhiza-dependent. Though relatively few in number, heath species dominate large areas around the world and are common understory plants in many forests. The fungi involved are poorly known but apparently are mostly Ascomycotina and probably some Basidiomycotina. They seem host-specific to the Ericales, and their spores are forcibly discharged to the air for dispersal. Ericoid fungi typically enter the outermost rootlet cells and fill them with masses of hyphae. These masses are sites of exchange of substances between host and fungus. The hyphae grow from these outer host cells into surrounding soil. *See* ERICALES.

Other types. Other mycorrhiza types include those special for the Orchidaceae (orchids) and Gentianaceae (gentians). Miscellaneous fungi commonly colonize roots to form various structures not mentioned above, often with nonmycorrhizal plants. The rootlets remain healthy, and such fungi may function with host roots as dual absorbing organs. Thus, they would fit the definition of mycorrhizae, but they are too poorly understood to be certain.

Ecology. The availability of propagules of mycorrhizal fungi dictates patterns of plant establishment and community development because of the differing degrees of dependency of different hosts. New soils, such as those in the front of retreating glaciers or emerging from the sea because of geologic forces, initially have few or no propagules of mycorrhizal fungi. The first plants to become established, therefore, are nonmycorrhizal or facultatively mycorrhizal species such as saxifrages, sedges, and rushes. Under these circumstances, no matter how abundant the seed source, mycorrhiza-dependent species cannot extract enough nutrients from the new substrate to survive. If abundant mushrooms of ectomycorrhizal or ericoid mycorrhizal fungi fruit in the vicinity, the new soil can soon become inoculated by aerial spore dispersal. Hence hosts such as the Ericaceae or ectomycorrhizal willows can establish along with or soon after the nonmycorrhizal species. Vesicular-arbuscular mycorrhiza–dependent species cannot gain a foothold until spore-bearing soil is introduced to the site by insects, animals, or erosion. After the eruption of Mount St. Helens in Washington covered a vast area with infertile, volcanic ash and pumice, patches of vesicular-arbuscular mycorrhiza–dependent plants soon appeared on gopher mounds. Protected underground when the mountain exploded, gophers later burrowed up through the ash layer to deposit the original, ash-buried soil from their tunneling on the surface. This soil contained spores of vesicular-arbuscular mycorrhizal fungi that enabled the vesicular-arbuscular mycorrhiza–dependent plants to survive.

The succession of plants from pioneering through seral to climax communities is governed by availability of mycorrhizal propagules. When catastrophic fire, erosion, or clearcutting reduce the availability of mycorrhizal fungi in the soil, plants dependent on those fungi will have difficulty becoming established. In clearcutting of conifer forests, overstory ectomycorrhizal trees are removed, thereby removing the food base for the ectomycorrhizal fungi. In contrast, ericoid or vesicular-arbuscular mycorrhizal understory plants may be damaged but often not killed. These plants retain their mycorrhizal fungi and respond to release from the overstory trees. The ericoid or vesicular-arbuscular mycorrhizal fungi build up in the soil as their host plants flourish to strongly compete for soil nutrients and moisture with ectomycorrhizal fungi that must start anew on establishing conifer seedlings.

Fungal host specificity also can affect species establishment and community composition of ectomycorrhizal and ericoid hosts. Heathlands typically are slow to convert to conifer stands, in part because the ericoid fungi so dominate the soil that roots of ectomycorrhizal hosts may not contact ectomycorrhizal fungal propagules. Eucalypts associate with their own, specific array of ectomycorrhizal fungi. Their fungi will not form mycorrhizae with trees of the Northern Hemisphere, so where plantations of inoculated eucalypts are established outside their native range, their fungi soon fully occupy the soil. Native ectomycorrhizal hosts have great difficulty invading such plantations for lack of their own ectomycorrhizal fungi.

Biodiversity. Each mycorrhizal fungus has its own array of physiological characteristics. Some are especially proficient at releasing nutrients bound in organic matter, some produce more effective antibiotics or growth regulators than others, and some are more active in cool, hot, wet, or dry times of year than others. Healthy plant communities or crops typically harbor diverse populations of mycorrhizal fungal species. This diversity, evolved over a great expanse of time, is a hallmark of thriving ecosystems. Factors that reduce this diversity also reduce the resilience of ecosystems.

Applications. Mycorrhizal inoculation of plants in nurseries, orchards, and fields has succeeded in many circumstances, resulting in improved survival and productivity of the inoculated plants. Inoculation with selected fungi is especially important for restoring degraded sites or introducing exotics such as pines in the Southern Hemisphere. Because ectomycorrhizal fungi include many premier edibles such as truffles, seedlings can also be inoculated to establish orchards for production of edible fungi.

In general, however, management of nursery, field, or forest soils to preserve or enhance the vigor of desired mycorrhizal fungi is the sound approach. For example, fumigation of nursery soils to control pathogens may eradicate mycorrhizal fungi and other, beneficial soil organisms. Alternative pest-control measures that do not damage mycorrhizal fungi are preferred. If organic matter is maintained, erosion minimized, and fertilization rationally

applied in field and forest soils, healthy and diverse populations of mycorrhizal fungi can more likely be sustained. *See* FOREST SOIL; FUNGI. James M. Trappe

Bibliography. M. F. Allen, *Mycorrhizal Functioning: An Integrative Plant-Fungal Approach*, 1992; G. C. Carroll and D. T. Wicklow, *The Fungal Community, Its Organization and Role in the Ecosystem*, 1992; A. H. Fitter et al., *Ecological Interactions in Soil: Plants, Microbes and Animals*, 1985; V. Gianinazzi-Pearson and S. Gianinazzi, *Physiological and Genetical Aspects of Mycorrhizae*, 1986; D. J. Read et al., *Mycorrhizas in Ecosystems*, 1992; A. Varma and B. Hock, *Mycorrhiza: Structure Function, Molecular Biology, and Biotechnology*, 2d ed., 1998.

Mycotoxin

Any of the mold-produced substances that may be injurious to vertebrates upon ingestion, inhalation, or skin contact. The diseases they cause, known as mycotoxicoses, need not involve the toxin-producing fungus. Diagnostic features characterizing mycotoxicoses are the following: the disease is not transmissible; drug and antibiotic treatments have little or no effect; in field outbreaks the disease is often seasonal; the outbreak is usually associated with a specific foodstuff; and examination of the suspected food or foodstuff reveals signs of fungal activity.

Historically, the earliest recognized mycotoxicoses were human diseases. Ergotism, or St. Anthony's fire, which results from eating rye infected with *Claviceps purpurea*, has been known since the Middle Ages. Yellow rice disease, a complex of human toxicoses caused by several *Penicillium islandicum* mycotoxins, was reported in Japan during the late nineteenth century. In Russia, large-scale human mycotoxicoses involving massive hemorrhaging were reported in the early twentieth century. Large numbers of deaths occurred as a result of eating overwintered grain contaminated with *Fusarium poae*. *See* ERGOT AND ERGOTISM.

Aflatoxins. World attention was directed toward the mycotoxin problem with the discovery of the aflatoxins in England in 1961. The aflatoxins, a family of mycotoxins produced by *Aspergillus flavus* and *A. parasiticus*, can induce both acute and chronic toxicological effects in vertebrates. Aflatoxin B_1, the most potent of the group, is toxic, carcinogenic, mutagenic, and teratogenic. There appears to be a significant dose-response correlation between estimated ingestion of aflatoxin-contaminated foods and primary liver cancer, particularly in Thailand, Kenya, and Mozambique. Major agricultural commodities that are often contaminated by aflatoxins include corn, peanuts, rice, cottonseed, and various tree nuts. When cows consume aflatoxin-contaminated feeds, aflatoxin M_1, a derivative of the toxin, is excreted in the milk. This substance is both toxic and carcinogenic. The aflatoxin problem originates with fungal contamination of crops in the field, thus making control difficult. Moisture levels of 15% or more during commodity storage may exacerbate the contamination level. Regulations on allowable aflatoxin levels exist in most countries, ranging from 0 to 50 parts per billion. Detoxification using ammonia is effective, but has not been approved for commercial use by regulatory agencies. Promising studies on field control through use of biocompetitive agents and plant genetic engineering are under way. *See* AFLATOXIN.

Other mycotoxins. There are additional mycotoxins of varying medical and economic importance. Those of major concern are listed in the **table**. Additional mycotoxins occur, but detection is usually difficult because they are normally produced in very low concentrations in natural substrates. The problem is also compounded by synergistic toxicological activity when several toxins are present. *See* MEDICAL MYCOLOGY; TOXIN. Alex Ciegler; Maren Klich

Bibliography. B. Arora and L. Arora (eds.), *Mycotoxins in Ecological Systems*, 1991; V. Betina, *Mycotoxins: Chemical, Biological, and Environmental Aspects*, 1989; G. A. Bray, *Mycotoxins, Cancer, and Health*, 1991; K. Sinha and D. Bhatnagar (eds.), *Mycotoxins in Agriculture and Food Safety*, 1998.

Compounds involved in mycotoxicoses

Toxin	Major producing fungi	Susceptible host	Biological effects
Aflatoxins	*Aspergillus flavus, A. parasiticus*	Mammals, fish	Hepatotoxic, cancer
Penitrem A	*Penicillium crustosum, P. clavigerum*	Cattle, horses, sheep	Tremorgenic, convulsant
T-2	*Fusarium sporotrichioides*	Cattle, humans?	Dermal necrosis, hemorrhage
Zearalenone	*Fusarium graminearum, F. culmorum*	Swine	Vulvovaginitis, abortion
Sporodesmins	*Pithomyces chartarum*	Swine, sheep	Hepatotoxic, facial eczema
Ochratoxin A	*Penicillium verrucosum, Aspergillus ochraceus*	Swine, humans?	Nephrotoxic
Citrinin	*Penicillium citrinum, Aspergillus terreus*	Swine	Nephrotoxic
Deoxynivalenol	*Fusarium graminearum, F. culmorum*	Swine, humans?	Vomiting
Maltoryzine	*Aspergillus oryzae*	Cattle	Death
Secalonic acids D and F	*Penicillium oxalicum, Aspergillus aculeatus*	Farm animals	Death
Satratoxins	*Stachybotrys atra*	Horses, humans	Hemorrhage
Fumonisins	*Fusarium moniliforme, F. proliferatum*	Humans, horses, swine	Cancer, neurotoxic, edema
Patulin	*Penicillium expansum, Aspergillus clavatus*	Humans	Immunosuppressant, nephrotoxic, hemorrhage
Cyclopiazonic acid	*Aspergillus flavus, Penicillium commune*	Mammals	Neurotoxic, hemorrhage

Myiasis

The infestation of vertebrates by the larvae, or maggots, of numerous species of flies. These larvae may invade different parts of the bodies of these animals or may appear externally (see **table**). Some invertebrates, such as spiders, may be attacked by species of Sarcophagidae, the flesh flies.

In cutaneous myiasis, the larvae are found in or under the skin. There may be a migration of some species of these larvae through host tissues, resulting in a swelling with intense itching. Such a condition is known as larva migrans, or creeping eruption, and may require surgical treatment.

Intestinal myiasis in humans is usually the result of accidentally swallowing the eggs or larvae of

Some species of the families of Diptera causing myiasis

Common name	Scientific name	Host and site of infestation
Muscidae (houseflies)		
Houseflies	*Musca domestica*	Humans; gastrointestinal tract
	Muscina stabulans	Numerous animals; intestinal tract
Lesser housefly	*Fannia canicularis*	Humans; intestinal tract
Latrine fly	*Fannia scalaris*	Humans; intestinal tract
Sacrophagidae (flesh flies, blowflies, scavenger flies)		
Gray blowfly	*Sarcophaga haemorrhoidalis*	Humans and numerous animals; intestinal tract, cutaneous and genitourinary tract
Gray blowfly	*Sarcophaga carnaria*	Humans and numerous animals; intestinal tract, cutaneous and genitourinary tract
Flesh fly	*Wohlfahrtia vigil*	Skin of infants; cutaneous
Fox maggot	*Wohlfahrita opaca*	Humans, fox, and mink; cutaneous
Old World screwworm	*Wohlfahrtia magnifica*	Eyes, ears, and nose of humans and animals; cutaneous
Tumbu fly, ver du cayor	*Cordylobia anthropophaga*	Humans, domestic and wild animals; cutaneous
Congo floor maggot	*Aucheromyia luteola*	Humans; suck blood
Calliphoridae (blowflies)		
American screwworm	*Callitroga americana*	Humans; cutaneous
Secondary screwworm	*Callitroga macellaria*	Sheep and goat; cutaneous
Old World screwworm	*Chrysomyia bezziana*	Humans and animals; cutaneous
	Chrysomyia megacephala	Sheep; cutaneous
	Phoenicia cuprina	"Strike" of sheep; cutaneous
Surgical maggot	*Phoenicia sericata*	Humans, sheep, and goat; wounds
Bluebottle fly	*Calliphora vomitoria*	
Greenbottle fly, surgical maggot	*Lucilia caesar*	Humans; wounds
Blackbottle fly, fleece worm	*Phormia regina*	Sheep; cutaneous
None	*Apaulina* sp.	Birds, especially nestlings; suck blood
Gasterophilidae (horse bots)		
Botfly	*Gasterophilus intestinalis*	Stomach of horse
Nose fly	*Gasterophilus haemorrhoidalis*	Stomach and duodenum of horse
Chin fly	*Gasterophilus nasalis*	Mouth, throat of horse
None	*Gasterophilus inermis*	Rectum of horse
None	*Gasterophilus pecorum*	Pharynx, stomach, and rectum of horse
Cuterebridae (robust botflies)		
Gusano del monte	*Dermatobia hominis*	Cattle and humans; cutaneous
None	*Cuterebra* sp.	Rodents and small mammals
None	*Bogeria* sp.	Rodents and small mammals
Oestridae (botflies, gadflies, nose flies)		
Sheep botfly	*Oestrus ovis*	Sometimes eyes of sheep and antelope; cutaneous
Nose fly	*Rhinoestrus purpureus*	Horse, hippopotamus, nose and eyes of humans
Warble fly	*Hypoderma lineata*	Warbles of cattle and reindeer
None	*Hypoderma bovis*	Warbles of cattle and reindeer
None	*Hypoderma crossi*	Goat, India; subcutaneous
None	*Cephalopina titillator*	Camel; nostrils and nasopharynx
None	*Kirkioestrus* sp.	African antelope; nasal passages
None	*Gadoelstia* sp.	African antelope; nasal pasages
None	*Cephalopsis* sp.	Camel; nasal passages
Head botfly	*Cephenemyia* sp.	Deer; head bots
Stratiomyidae (soldier flies)		
Soldier fly	*Hermetia illucens*	Recorded from intestine of humans
Syrphidae (flower flies)		
Rat-tailed maggot	*Eristalis tenax*	Recorded from stomach of humans

these flies. It occurs commonly in many herbivores who ingest the eggs when feeding on contaminated herbage. The larvae settle in the stomach or intestinal tract of the animal host.

Cavity, or wound, myiasis occurs when the larvae invade natural orifices, such as the nasopharynx, vulva, and sinuses, or artificial openings such as wounds. A study of wound myiasis led to the use of the larvae of *Lucilia sericata* and *Phormia regina* in treating osteomyelitis and suppurating wounds. These maggots commonly feed on necrotic tissue; there is a danger, however, that normal tissue may be attacked. Maggots used for this purpose are reared under sterile conditions. It was also discovered that allantoin, one of the compounds found in the urine of the maggots, has a healing effect on ulcers.

External myiasis includes infestation by those maggots which are blood feeders.

Of considerable economic importance is myiasis of such domestic animals as horses, sheep, or cattle. Cattle afflicted with cutaneous myiasis produce poor-grade hides full of small perforations. Heavy infestations of cattle by screwworms result in the death of the host if not treated. Horses afflicted with the stomach bot become emaciated and may succumb to the ravages of these larvae.

The Diptera of medical and veterinary importances are largely confined to the families Oestridae, Calliphoridae, and Sarcophagidae. An interesting ecological phenomenon has been observed in the life cycle of the human botfly. The female oviposits on the undersurface of the abdomen of mosquitoes as they emerge from their aquatic habitats. When the mosquito takes a blood meal from the vertebrate host, the eggs are in close contact with the skin of the host. Heat from the host's body causes the larvae to hatch from the eggs, and these tiny larvae then burrow into the skin and cause a creeping eruption. *See* DIPTERA; MEDICAL PARASITOLOGY. Charles B. Curtin

Bibliography. J. W. Beck and J. E. Davies, *Medical Parasitology*, 3d ed., 1981; M. W. Sloss and R. L. Kemp, *Veterinary Clinical Parasitology*, 6th ed., 1994; F. Zumpt, *The Stomoxyine Biting Flies of the World (Diptera, Muscidae): Taxonomy, Biology, Economic Importance and Control*, 1973.

Myliobatiformes

An order of batoid fishes (subclass Elasmobranchii), the typical members of which are described by the following characteristics: The disc is strongly depressed and varies from oval longitudinally to much broader than long; the tail is well marked off from the body sector, very short to long and whiplike, and equipped with a poisonous spine in some species; the pectoral rays are either continuous along the side of the head or separate from the head and modified to form rostral lobes or finlike rostral appendages (cephalic fins); the dorsal fin, if present, is near the base of the tail; and development is ovoviviparous. Sometimes an individual ray may have two or three, very rarely four, tail spines rather than the usual one. Multiple spines result from the failure of older spines to shed before new ones develop. Most species are tropical or subtropical, although some species occur in warm temperate and cool temperate zones. The usual habitat is shallow shore waters and upper continental and insular slopes. Some, such as the eagle rays, may venture far out to sea, but none are considered truly oceanic rays. *See* BATOIDEA; ELASMOBRANCHII.

Phylogeny

In the past two decades, studies on the phylogeny of batoids have led to some major changes in their classification, one of which results in the inclusion of the platyrhinids and *Zanobatus* in Myliobatiformes. Although these fishes superficially resemble guitarfishes more than myliobatoids, they share certain esoteric skeletal features with myliobatoids. The following is the most recently proposed classification of the order.

Suborder Platyrhinoidei. The family Platyrhinidae (thornbacks) has only two extant genera (*Platyrhina*, two species; and *Platyrhinoidis*, one species). It is distinguished by a round disc, stout tail sector, two large dorsal fins, large caudal fin, and one or three rows of stout, knobby thorns on the disc and tail. It occurs in tropical to cool-temperate waters on the continental shelves of the North Pacific in Asia and North America, including California and Mexico. A Late Cretaceous fossil from Italy, *Tethybatis*, is assigned to this family. *See* CRETACEOUS.

Suborder Zanobatoidei. The family Zanobatidae (Panrays) consists of one species, *Zanobatus schoenleinii*, which occurs in the eastern Atlantic from Morocco to the Gulf of Guinea. It is similar to Platyrhinidae but has a less robust tail sector.

Suborder Myliobatoidei. The family Hexatrygonidae (sixgill stingrays) consists of one species, *Hexatrygon bickelli*, and is very unusual in possessing the following features: The disc is longer than wide; there are six rather than five gill arches and gill slits as in other batoids; the snout is elongate, flattened, and translucent; large spiracles are placed well behind the eyes; the spiracle is closed by an external valve. In contrast, other rays have an internal valve; a barbed spine occurs on the tail; and maximum length of 100 cm (39 in.). This unusual ray occurs on the continental and insular slopes of the Indo-West Pacific.

The family Plesiobatidae (deepwater stingrays) has one species, *Plesiobatis daviesi*. Its round disc is bordered in black; the anterior profile is obtuse with a pointed snout; a barbed spine is present on the tail; the nasal curtain is incompletely united and does not reach the mouth, a character this ray shares with *Hexatrygon*; and maximum length is 2.7 m (8.8 ft). It occurs on continental and insular slopes of the Indo-West Pacific to a depth of 680 m (2230 ft) [**Fig. 1**].

In the family Urolophidae (round stingrays), the disc usually is only slightly broader than long; the tail sector is short to moderately long and bears a barbed spine; the caudal fin is well developed, with

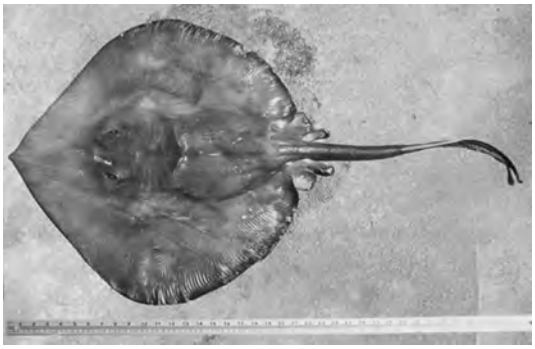

Fig. 1. ***Plesiobatis daviesi*. (*Photo courtesy of R. F. Myers*)**

cartilaginous radial supports; and a dorsal fin is present or absent. Round stingrays occur in the western Pacific on continental shelves and slopes, as well as inshore on reefs and grass beds. The family comprises two genera and 24 species (probably more), the smallest of which is 17 cm (6.7 in.) in maximum disc length, the largest 80 cm (31.5 in.).

Members of the family Urotrygonidae (American round stingrays) are similar to urolophids, but differ in having a slender tail, which is about equal to the length of the disc, a distinct caudal fin and poisonous spine, and no dorsal fin. They occur on continental shelves in tropical and warm temperate waters of the western Atlantic and eastern Pacific. The maximum disc length varies from 12 to 46 cm (4.7 to 18 in.), depending on the species.

Members of the family Dasyatidae (whiptail stingrays) have a long slender or whiplike tail armed with a serrated poisonous spine and lack a dorsal and caudal fin. As in round stingrays, the disc is not more than 1.3 times broad as long. Settled over the substrate, these fishes, as well as other bottom-dwelling rays, fan their "wings," sending up plumes of sand or mud that covers the body, leaving only the eyes, spiracles, and part of the tail exposed. Fanning the substrate exposes crustaceans, mollusks, and polychaete worms that provide the principal part of the rays' diet. Small fishes and squid are consumed in lesser quantity. The six genera and about 71 species of whiptail stingrays occur for the most part on continental and insular shelves of tropical and warm temperate zones of the Pacific, Indian, and Atlantic oceans, including the Mediterranean Sea. Some species are common in brackish water, and some make temporary excursions into freshwater. Species of the family vary in disc width from 11 to 220 cm (4.3 in. to 7 ft).

Urolophids, urotrigonids, and dasyatids, collectively known as stingarees, are potentially dangerous to humans. The piercing tail spine delivers a poison that causes immediate and excruciating pain, followed by swelling and most often a serious infection. When a stingaree is stepped on, the reflex action is to strike with the barb. Stings may be avoided by dragging the feet while wading. In this manner, the ray is nudged on its wing, causing it to move away. On clearwater beaches, search the bottom carefully before stepping from a boat.

Potamotrygonidae (river stingrays) is the only family of chondrichthyan fishes restricted to a freshwater environment. Known commonly as river stingrays, members have low concentration of urea in the body fluids and a rectal gland for salt excretion. The family is further characterized by the following: There is a round flat disc, the dorsal surface of which is usually marked with spots, ocelli, or reticulate and vermiform color patterns; the pre-pelvic process is greatly expanded; the disc and tail are usually armed with denticles, thorns, or tubercles; dorsal and caudal fins are absent; development is ovoviviparous, but essentially viviparous in that embryos are nourished by uterine "milk," a reproductive strategy characteristic of other myliobatoids. The three genera and 20 species of river stingrays occur in rivers of South America that drain into the Atlantic Ocean and Caribbean Sea. Maximum length of the disc is 25 to 100 cm (10 to 40 in.), depending on the species.

Members of the family Gymnuridae (butterfly rays) are very distinctive in having a disc more than 1.5 times broad as long and a short tail lacking a caudal fin, as well as lacking a dorsal fin and tail spine in some species (**Fig. 2**). They occur on tropical and temperate continental shelf waters of the Atlantic, Indian, and Pacific oceans. The largest species, *Gymnurus altavela*, of both sides of the Atlantic, including the Mediterranean Sea, reaches 400 cm (13.7 ft); disc width most of the 11 species (in two genera) are far smaller.

In the family Myliobatidae (eagle rays), the margins of the pectoral fins are deeply indented or entirely interrupted just posterior to the eyes, sharply distinguishing the anterior part of the head from the rest of the disc. Anterior subdivisions of the pectoral fins form a separate lobe (or lobes), which in the manta ray and the devil rays form cephalic fins. The eyes and spiracles are on the sides of the head. The family is represented in the tropical and warm temperate waters of the Atlantic, Indian, and Pacific oceans. It inhabits continental and insular shelves to offshore, but none of the species are considered oceanic inhabitants. There are three subfamilies.

In the subfamily Myliobatinae (eagle rays), the anterior subdivision of the pectoral fins forms a single subrostral lobe extending forward below the front of

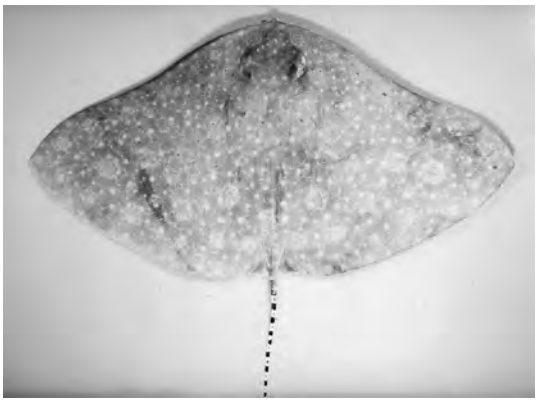

Fig. 2. ***Gymnura poecilura*. (*Photo courtesy of J. E. Randall*).**

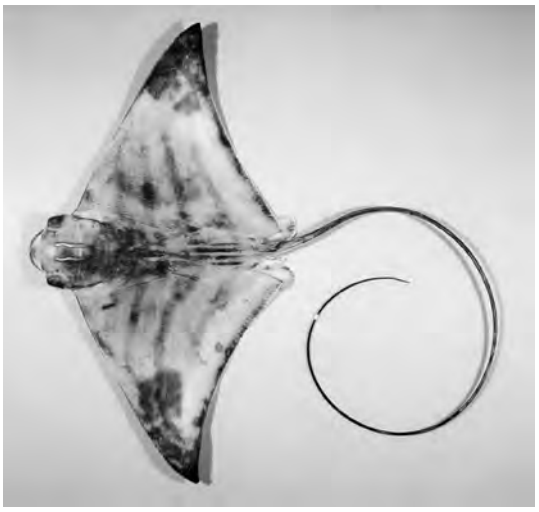

Fig. 3. ***Aetomylaevs nichofii*, a banded eagle ray. (*Photo courtesy of J. E. Randall*)**

the head. The crown of the head is conspicuously elevated, and the eyes and spiracles are laterally placed. The tail is much longer than the disc and is armed with a long serrated spine in some species; a dorsal fin is present near the base of the tail, but a caudal fin is absent. Eagle rays have the ability to leap from the water and high into the air. The smallest and largest species are 65 cm (2 ft) and 300 cm (10 ft) in disk width (**Fig. 3**).

In the subfamily Rhinopterinae (cownose rays), the subrostral lobe extending forward from the lower surface of the head is deeply incised, leaving two distinct lobes, and the concave anterior contour of the cranium leaves the front of the head bilobed, resulting in a configuration reminiscent of a cow's nose. Cownose rays and eagle rays subsist chiefly on bivalve mollusks, especially oysters and clams, and to a lesser extent on crustaceans. Of the seven species in a single genus, the smallest attains a disc width of 86 cm (33 in.), the largest 213 cm (7 ft).

In the subfamily Mobulinae (devil rays), the anterior subdivisions of the pectoral fins are in the form of two tonguelike cephalic fins, which are widely separated from each other and curve forward from the front of the head. Unlike the eagle rays and cownose rays, which have several rows of large pavementlike teeth, the mobulids have minute teeth in many series. The disc is relatively thin; the head is broad and its dorsal surface is only slightly convex. Devil rays subsist on planktonic crustaceans and small fishes that are directed into the mouth by the cephalic fin and then strained by gill plates located on the inner surface of the gill slits, which are longer than those of eagle and cownose rays. These rays are famous for their large sizes and spectacular leaps out of the water. The family comprises two genera: *Manta*, with one species, *Manta birostris*, commonly known as the giant manta or giant devil ray; and *Mobula*, with nine species, collectively known as the devil rays.

Manta birostris has a very broad head (its interorbital distance is about one-quarter of the greatest width of the disc), relatively large forward-projecting-cephalic fins, and a terminal mouth extending across the front of the head, with no teeth in the upper jaw. The dorsal surface of the disc is covered with denticles, and the tail is essentially without a spine. The ray ranges worldwide in tropical and warm temperate zones near coral and rocky reefs and sometimes over deep water, as well as in shallow bays and the mouths of large rivers. Up to 800-cm (26-ft) disc width has been reported.

Mobula differ from the giant manta in that the head is not nearly as wide, the mouth is on the lower surface of the head, and teeth are in both jaws. Members of the genus occupy tropical and warm temperate belts of all oceans, including the Mediterranean Sea. The known maximum disc width varies from 100 to 520 cm (3 to 17 ft), depending on the species.

Herbert Boschung

Bibliography. L. J. V. Compagno, Checklist of living elasmobranchs, pp. 471–498 in W. C. Hamlett (ed.), *Sharks, Skates, and Rays: The Biology of Elasmobranch Fishes*, Johns Hopkins University Press, Baltimore, 1999; L. J. V. Compagno, Checklist of Chondrichthyes, pp. 503–547 in W. C. Hamlett (ed.), *Reproductive Biology and Phylogeny of Chondrichthyes: Sharks, batoids and chimaeras*, Science Publishers, Enfield, NH, 2005; J. D. McEachran and N. Aschliman, Phylogeny of Batoidea, pp. 79–113 in J. C. Carrier, J. A. Musick, and M. R. Heithaus (eds.), *Biology of Sharks and Their Relatives*, CRC Press, Boca Raton, FL, 2004; J. D. McEachran and M. R. de Carvalho, Batoid Fishes, pp. 507–589 in L. K. E. Carpenter (ed.), *The Living Marine Resources of the Western Central Atlantic: FAO Species Identification Guide for Fishery Purposes*, vol. 1, FAO, Rome, 2002; J. D. McEachran, K. A. Dunn, and T. Miyake, Interrelationships of the batoid fishes (Chondrichthyes: Batoidea), pp. 63–84 in M. L. J. Stiassny, L. R. Parenti, and G. D. Johnson (eds.), *Interrelationships of Fishes*, Academic Press, San Diego, 1996; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2004.

Mylonite

A rock that has undergone significant modification of original textures by predominantly plastic flow due to dynamic recrystallization. Mylonites form at depth beneath brittle faults in continental and oceanic crust, in rocks from quartzo-feldspathic to olivine-pyroxenite composition. Mylonites were once confused with cataclasites, which form by brittle fracturing, crushing, and comminution. Microstructures that develop during mylonitization vary according to original mineralogy and modal compositions, temperature, confining pressure, strain, strain rate, applied stresses, and presence or absence of fluids.

At low to moderate metamorphic grades, mylonitization reduces the grain size of the protolith and commonly produces a very fine-grained, well-foliated rock with a pronounced linear fabric defined by

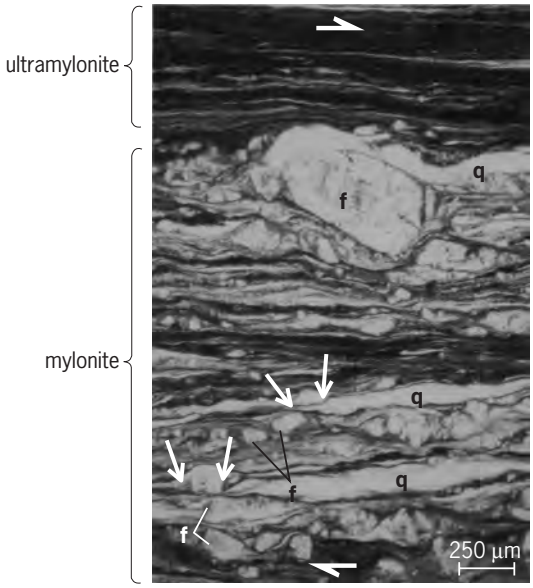

Photomicrograph of greenschist-grade mylonite and ultramylonite derived from granodiorite, Borrego Springs shear zone, southern California. Half-arrows indicate right-lateral shear sense. Feldspar porphyroclasts (f), some with recrystallized tails (indicated by arrowheads), appear in matrix of recrystallized quartz ribbons (q) and dark, fine-grained biotite. Characteristic darkness of ultramylonite zone results from extremely fine recrystallized grain size of component minerals.

elongate minerals. Lineations may be weak or absent, however, in high-strain zones that lack a significant rotational component. At high metamorphic grades, grain growth during mylonitization can produce a net increase in grain size, and the term mylonitic gneiss is used where there is a preserved or inferred undeformed protolith.

At low temperatures, high strain rates, and high applied stresses, most minerals are strong, and they deform by brittle failure. A mylonite forms at and above the conditions favorable for dynamic recrystallization of the weakest mineral phase present, provided there is sufficient weak phase to allow the bulk rock to flow. In quartzo-feldspathic rocks, mylonitization occurs above 480–570°F (250–300°C), whereupon quartz flows plastically at natural strain rates by the mechanisms of dislocation glide and dislocation creep. Under these conditions, feldspar remains brittle and is often preserved as broken megacrysts, or porphyroclasts, surrounded by fine-grained, recrystallized ribbons of quartz (see **illus.**).

At temperatures above about 840°F (450°C), feldspar also dynamically recrystallizes, to produce a narrow mantle of small recrystallized grains around a more rigid core. The tiny new grains are swept into tails around the porphyroclast; and, if asymmetrically developed, they record the sense of shear (see illus.). Other common asymmetrical microstructures in mylonites include shear bands, oriented crystallographic axes, recrystallized grain shapes, microfolds, and microfaults.

Protomylonites form at lowest strains and preserve some protolith textures. Blastomylonites contain grains that have recrystallized and also grown in size significantly during or after mylonitization. Rocks of the type known as S-C mylonites contain two distinct but related foliations (S = schistosity; C = cisaillement, or shear bands) that give the overall shear sense for the rock. Ultramylonites form at highest strains and contain few relict porphyroclasts in an extremely fine-grained matrix (see illus.). *See* METAMORPHIC ROCKS. Carol Simpson

Bibliography. G. S. Lister and A. W. Snoke, S-C mylonites, *J. Struc. Geol.*, 6:617–638, 1984; C. W. Passchier and C. Simpson, Porphyroclast systems as kinematic indicators, *J. Struct. Geol.*, 8:831–843, 1986; A. Schedl and B. A. van der Pluijm, A review of deformation microstructures, *J. Geol. Educ.*, 36:111–121, 1988; C. Simpson, Determination of movement sense in mylonites, *J. Geol. Educ.*, 34:246–261, 1986.

Myodocopida

An order of marine organisms that forms an important part of the class Ostracoda (subphylum Crustacea) and comprises three suborders. It has a long geological history that extends back at least to the Early Silurian. However, the lightly calcified carapaces of the species in this order are a factor in their having a sparse fossil record. Indeed, of the sixteen families of myodocopids, three are known only from modern marine environments. Species of those families that have a fossil record are quite rare and discontinuous in their stratigraphical distribution. Myodocopids are only abundant as fossils from organically rich shale that was deposited in deep, anoxic environments, especially those of the Devonian seas. Such environments were devoid of benthic (bottom-dwelling) life, but the valves and carapaces of the nektonic (free-swimming) myodocopids, both immature forms (termed instars) and adults, sank into such environments and were preserved. The rapid evolution of myodocopids during the Devonian Period has led to their use as guide fossils for dating these otherwise poorly fossiliferous marine rocks. *See* CRUSTACEA; OSTRACODA.

The myodocopids are typically much larger than other ostracodes and may be more than a centimeter long, although adults of most species are unlikely to be longer than 3 mm. As is characteristic of ostracodes, the carapace of the myodocopids comprises two valves that are joined along the animal's dorsum by a hinge and ligament. In keeping with the generally light calcification of myodocopids, the hinge is simple and without the sort of morphological embellishments that make hinge morphology a useful feature in the systematics of many other groups of ostracodes.

Among the myodocopids, the carapace does not completely encase the soft anatomy of the animal when the valves are closed. Instead, the anterior portions of both valves are marked by a characteristic notch through which the setaceous antennae or antennules protrude.

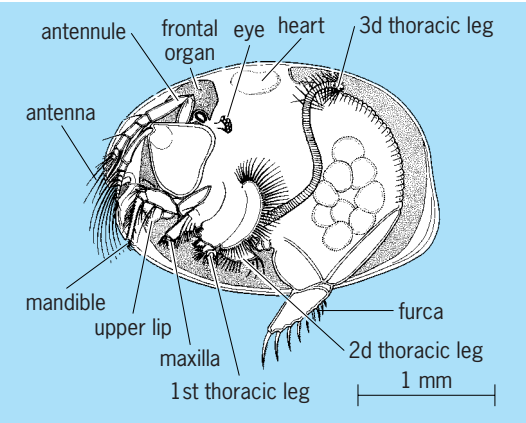

Fig. 1. Morphology of a typical myodocopid ostracode, an adult female of *Cypridina norvegica* Baird, viewed from the left side with the left valve removed. (*Modified from R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. Q, Arthropoda, University of Kansas and Geological Society of America, Lawrence and Boulder, 1961*)

The appendages of all ostracodes are homologous with those of other crustaceans and, as is characteristic of arthropods in general, have been highly modified during their evolution. The appendages of the myodocopids are especially well adapted for swimming. The first two cephalic appendages, the antennule and antennae in particular, are characterized by long setae that facilitate swimming (**Figs. 1** and **2**). Sexual dimorphism is not as pronounced among the myodocopids as it is among most smaller, more heavily calcified ostracodes.

Ostracodes do not have a planktonic larval stage that is characteristic of so many kinds of marine organisms. Instead, the females carry a clutch of four to six fertilized eggs within their carapace until the tiny, first-instar ostracods hatch as typical, bivalved, nauplius larvae (Fig. 2). The larvae add appendages as they grow by molting. Most myodocopids have five or six instars followed by a terminal adult stage beyond which the ostracodes do not grow, molt, or otherwise change in morphology.

Myodocopids are not exclusively planktonic as is often assumed. They are typically strong swimmers, but many are nektobenthic, swimming in proximity to the substrate. Others, however, live near the surface of the open ocean and are not associated with the environments of deposition that lie far beneath them on the ocean floor, where their valves and carapaces will ultimately be deposited.

A fascinating characteristic of the myodocopids is the ability of many species to secrete bioluminescent material into the water. This bioluminescence functions as highly elaborate courtship displays that are species-specific and facilitate recognition of potential mates. *See* BIOLUMINESCENCE.

Roger L. Kaesler

Bibliography. A. C. Cohen and J. G. Morin, Morphological relationships of bioluminescent Caribbean species of *Vargula* (Myodocopa), in R. C. Whatley and C. Maybury (eds.), *Ostracoda and Global Events*, Chapman and Hall, New York, 1990; R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, pt. Q, Arthropoda 3, Crustacea, Ostracoda, University of Kansas and Geological Society of America, Lawrence and Boulder, 1961; R. C. Whatley, D. J. Siveter, and I. D. Boomer, Arthropoda (Crustacea: Ostracoda), in M. J. Benton (ed.), *The Fossil Record 2*, Chapman and Hall, New York, 1993.

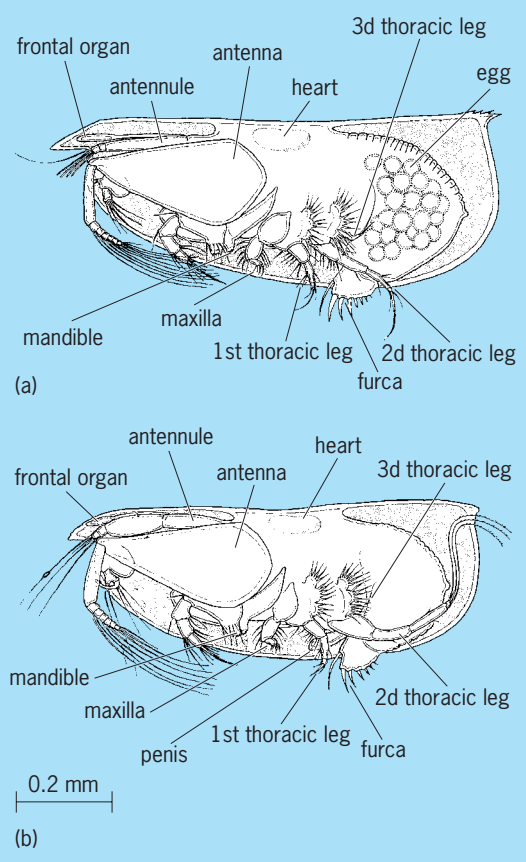

Fig. 2. Morphology of both sexes of *Conchoecia elegans* Sars, viewed from the left side with the left valve removed. (*a*) Adult female with eggs. (*b*) Adult male. (*Modified from R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. Q, Arthropoda, University of Kansas and Geological Society of America, Lawrence and Boulder, 1961*)

Myricales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Hamamelidae of the class Magnoliopsida (dicotyledons). The order consists of the single family Myricaceae, with about 50 species. Within its subclass the order is marked by its simple, resinous-dotted, aromatic leaves and unilocular ovary with two styles and a single ovule. The plants are trees or shrubs, and the flowers are much reduced and borne in catkins. The fruit is a small, waxy-coated drupe or nut. Several species of *Myrica* are occasionally cultivated as ornamentals. *See* HAMAMELIDAE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; ORNAMENTAL PLANTS; PLANT KINGDOM.

Arthur Cronquist

Myrtales

An order of flowering plants in the core eudicots. The order consists of 10 families and approximately 9300 species. The circumscription of the order has been relatively stable, with only minor changes made on the basis of deoxyribonucleic acid (DNA) sequence data. The two largest families are Melastomataceae (approximately 4500 species) and Myrtaceae (approximately 3000 species). Thymelaeaceae are excluded in recent concepts of the order, being related instead to families of Malvales. Vochysiaceae, traditionally included in Polygalales, have been shown to be members of Myrtaceae with DNA sequence data.

Myrtales are chiefly tropical, but Onagraceae and Penaeaceae are predominantly temperate. Myrtales usually have opposite, simple, entire leaves and perigynous (situated around the pistil) to epigynous (having all floral parts conjoined near the top of the ovary) flowers with a compound pistil and most commonly axile placentation. The seeds have little or no endosperm. The stamens are normally numerous (Vochysiaceae being a marked exception), and many species have tetramerous flowers. Vascular bundles characteristically have internal phloem, which is otherwise rare in the rosid dicots.

Economic crops in Myrtaceae include spice trees such as allspice (*Pimenta*) and cloves (*Syzygium*), and the timber trees *Eucalyptus*. Other important economic crops in the order include evening primrose (*Oenothera*, Onagraceae) and pomegranate (*Punica*, Lythraceae). Purple loosestrife (*Lythrum salicaria*) can be a noxious weed of North American waterways. *See* EUDICOTYLEDONS. Michael F. Fay

Mysida

An order of free-swimming, shrimplike crustaceans, commonly known as opossum shrimps, belonging to the subclass Eumalacostraca. They occur in vast numbers in coastal and open oceanic regions of the world. Mysida were once classified along with the Lophogastrida into an order Mysidacea, but the two, along with a closely allied extinct Paleozoic group, Pygocephalomorpha, are now generally considered to be separate orders. The mysidan adult body length generally averages about 0.6 in. (15 mm), and most species are distributed in shallow coastal and shelf waters of the oceans. A few mysidans live in the surface layers of the ocean bottoms. In addition, some species have invaded freshwaters, including specialized forms strictly confined to caves. *See* CRUSTACEA; EUCARIDA; EUMALACOSTRACA.

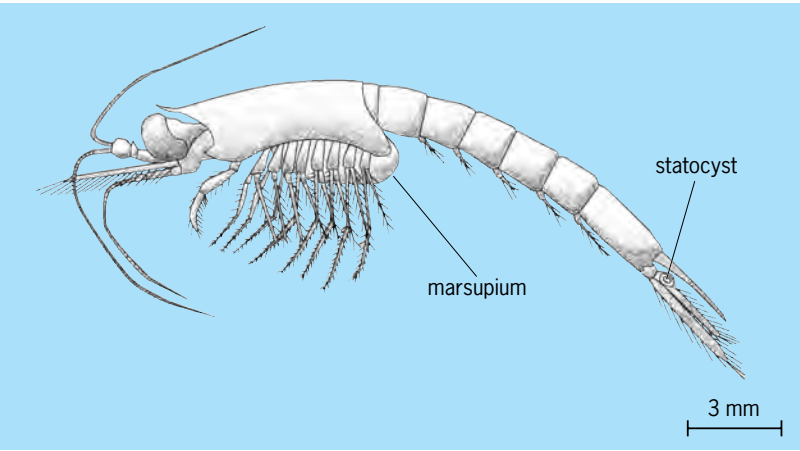

***Boreomysis arctica*, a member of the Mysida showing the marsupium and the paired uropods with a statocyst in the inner branch or endopod. (*From W. M. Tattersall and O. S. Tattersall, 1951*)**

Biology. Mysida are among the most primitive of the peracarid eumalacostracans. The carapace generally covers the entire thorax, and it is not fused to the underlying segments. The thoracic limbs are biramous (having two branches), possessing both a well-developed medial endopod and a lateral natatory (adapted to swimming) exopod; the thoracopods generally bear well-developed gills on the base of the limbs. Although their first thoracic appendage is modified as a maxilliped to supplement the actions of the mouthparts, the first thoracic segment is not fused into the cephalon. Generally the pleopods are well developed, although in cave- and bottom-dwelling forms these abdominal limbs can be very reduced. The first pair of pleopods in males is often highly modified to facilitate sperm transfer in copulation. *See* PERACARIDA.

Mysidans feed on algae, detritus, and zooplankton. They pass through up to 12 molts before achieving sexual maturity. The time between successive molts depends upon environmental temperature and body size. In general, smaller species have shorter life cycles than larger species. Generation time in warm summer environments can be a few weeks; in colder seasons, the same species may have a generation time of several months. Most species have more than one generation per year, the maximum considered to be about five per year. Longer-lived species do exist. *Mysis relicta* is known to have a 2-year life cycle in high-latitude freshwater lakes.

Most species of mysids form aggregations. These are of different types for different purposes. Breeding aggregations occur in coastal and shelf taxa, and possibly some open oceanic species; these consist of loose groupings of sexually mature adults that come together seasonally for mating. Large aggregations called shoals can extend over large areas or volumes, but densities of animals within shoals vary. Swarms are often constituents of shoals and can be spherical or wafer-shaped; the greatest densities of mysids occur in swarming populations. Another type of aggregation is the school, in which swimming mysids are all orientated parallel to each other and are traveling in tandem through the environment. Schools, because they are moving, are usually elongated in shape. A single species will form a different kind of aggregation at different stages of its life history. The functions of these aggregations, except for breeding, are not always clear, although protection of the population from predators is important. The chance of an

individual surviving an attack when embedded in a swarm of thousands is significantly better than when on its own, although the open ocean school itself often attracts special predators such as whales. Aggregations occur in many coastal mysids, especially in estuarine or sandy beach habitats. The possession by mysidans of the unique balance organ, the statocyst, is probably linked with this swarming habit by facilitating the maintenance of proper orientation to the environment and with each individual's swarm mates.

As is the case with all peracarids, the young are carried within a marsupium, or brood pouch, formed by transparent concave plates attached to the insides of the thoracic legs. These plates, called oostegites, have interlocking setae that form the marsupium beneath the thorax. The eggs, which are fertilized while being laid, are deposited directly into the marsupium, develop to miniature adults, and emerge to swim freely in the water. The number of embryos depends on the species. Most species have brood sizes in the range 4–30, but broods of 50–100 are common. The largest brood recorded is 245 embryos in the marsupium of a *Mysis stenolepis*.

Distribution. Mysids are predominantly marine organisms, but many species occur in freshwater lakes and rivers, as well as freshwater or brackish-water caves and ground-water wells. One of the most widely distributed freshwater species is *M. relicta*. The geographical distributions of the 800 marine species are not well known: more than one-third live in the open ocean, and the remaining species are found in shallow coastal regions, including estuaries.

In contrast, many of the coastal species appear to have very limited distributions. Many species live actually on or immediately above the sediment or sand, not only in coastal regions. Some species are commensal, each with a different partner, ranging from sessile sponges and anemones to land and hermit crabs.

Classification. The order Mysida currently contains almost 1100 species in some 168 genera ascribed to four families: Petalophthalmidae, Mysidae (by far the largest family), Lepidomysidae, and Stygiomysidae. Except for the deep-sea petalophthalmids (six species) and the cave-dwelling lepidomysids (nine species), all mysidans have statocysts in the uropods of their tail fan (see **illustration**), a feature peculiar to mysidans.

Fossil record. Although most mysidans, based on living forms, are much too small and of so thin a cuticle that they would stand little chance of being preserved as fossils, there is a small record. These fossils mostly occur in the fine-grained lithographic limestones of Jurassic age in Germany and France, and consist of genera such as *Elder*, *Francocaris*, and *Siriella*. Nevertheless, the larger, better-sclerotized lophogastridans and the Paleozoic pygocephalomorphs have more interesting and extensive fossil records.

Important species. Swarms of mysids in coastal waters are exploited commercially in tropical and subtropical regions of the world, and are used to make shrimp paste and sauces. The brackish species, *Neomysis intermedia*, is marketed in Japan as a preserved, cooked food known as tsukudani. Other species, or several mixed species, are used in India and Southeast Asia as human food. Several species, notably the freshwater *M. relicta*, have been introduced into lakes in both North America and Europe to serve as food for resident fish such as salmon.

John Mauchline; Frederick R. Schram

Bibliography. J. Mauchline, The biology of mysids, *Adv. Mar. Biol.*, 18:1–369, 1980; T. Remerie et al., Phylogenetic relationships within the Mysidae (Crustacea, Peracarida, Mysida) based on nuclear 18S ribosomal RNA sequences, *Mol. Phylogenet. Evol.*, 32:770–777, 2004; F. R. Schram, Mysidaceans, *Crustacea*, pp. 107–127, Oxford University Press, New York, 1986; W. M. Tattersall and O. S. Tattersall, *The British Mysidacea*, Ray Soc. Monogr. 136, 1951.

Mystacocarida

A subclass of primitive Crustacea, discovered in the interstices of Massachusetts intertidal sandy beaches in 1939. Specimens of both sexes have since been found in similar habitats on shores of Europe, Africa, and Chile. The three species are in the genus *Derocheilocaris*. The body is wormlike and about 0.2 in. (5 mm) long (see **illus.**). The cephalothorax

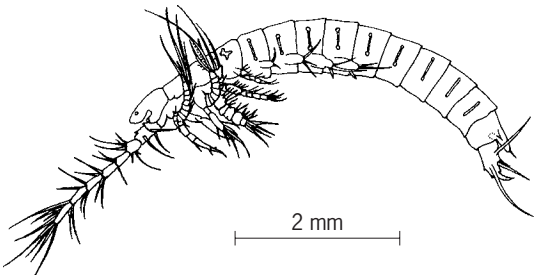

***Derocheilocaris typicus*, lateral view.**

bears first antennae, second antennae, mandibles, first maxillae, and second maxillae. Maxillipeds are on a separate segment, and four additional free thoracic segments bear platelike appendages. The six abdominal segments are without appendages, but large caudal rami are present, The labrum is enormous, mouthparts and nervous system are primitive, and the genital pore is on the thorax. *See* CRUSTACEA.

Robert W. Pennak

Myxiniformes

An order of Craniata in the class Myxini; commonly called hagfishes. There are old and new views regarding the classification of the Myxiniformes. The two classifications here comprise the extant myxiniform taxa.

New:
Phylum Chordata
  Subphylum Craniata
    Superclass Myxinomorphi
      Class Myxini
        Order Myxiniformes (Hyperotreti)—hagfishes
    Superclass Petromyzontomorphi
      Class Petromyzontida
        Order Petromyzontiformes (Hyperoartii)—lampreys
    Superclass Gnathostomata—jawed vertebrates

Old:
Phylum Chordata
  Subphylum Vertebrata (= Craniata)
    Superclass Agnatha
      Class Myxini
        Order Myxiniformes (Hyperotreti)—hagfishes
      Class Cephalaspidomorphi (Monorhina)
        Order Petromyzontiformes (Hyperoartii)—lampreys
    Superclass Gnathostomata—jawed vertebrates

In the old classification, the Vertebrata and Craniata are synonymous; extant hagfishes and lampreys are collectively called Cyclostomata (jawless fishes). This classification supports the monophyly of the cyclostomes (called the cyclostome hypothesis) and rejects the idea that lampreys are more closely related to gnathostomes than to hagfishes (called the vertebrate hypothesis). *See* CEPHALASPIDOMORPHA; CYCLOSTOMATA (CHORDATA); JAWLESS VERTEBRATES; PETROMYZONTIDA.

In the new classification, Vertebrata and Agnatha are not used as taxon names, and Craniata and Vertebrata are not synonymous. Myxinomorphi are the sister group of the vertebrates, and include the superclasses Petromyzontomorphi and Gnathostomata (vertebrate hypothesis). Hagfishes are excluded from the vertebrates primarily because they lack arcualia. Arcualia are embryonic cartilaginous elements that fuse to a notochordal calcification, the centrum, to form a vertebra. Proponents of the new classification assume that hagfishes are not degenerate forms of some vertebrate taxon.

Myxiniforms can be identified by the following set of characteristics: an eel-like scaleless body; 1 to 16 pairs of external gill openings; one semicircular canal; and no bones, no paired fins, no neuromasts, and no lens or extrinsic eye muscles. Based on a very scanty fossil record, it is probable that hagfishes have remained little changed since the Pennsylvanian age 300 million years ago. The order comprises one family, two subfamilies, seven genera, and about 70 species, all of which occur in the temperate seas of the world.

Myxinidae (hagfishes). Hagfishes (see **illustration**) have a low fin continuous from the middorsum, around the tail, and forward on the midventrum, but it is not differentiated as dorsal, caudal, and anal fins. The maximum total length is 18–116 cm, depending on the species. Copious amounts of slime are excreted through numerous pores on the ventrolateral sides of the body. The slime glands contain mucous cells and thread cells. Threads produced by the thread cells (unique to hagfishes) are thought to give tensile strength to the slime. The slime can aid in escaping the grip of a predator as well as clogging its tail and causing suffocation. The knotting maneuver in feeding depends on the slime. The eyes are degenerate and not visible externally. The nostril is on the tip of the snout, and the nasohypophyseal pouch opens into the pharynx. There are two barbels on each side of the nostril and one or two on each side of the mouth. The mouth is not funnellike or disklike and has a single tooth in its roof; the tongue is eversible and bears two rows of rasplike teeth. The cranium is an unroofed trough cradling the brain. Hagfishes have relatively large, capsule-shaped eggs, with hooks on each end; their ontogeny lacks metamorphosis; and they are the only known craniate whose body fluids are isosmotic with seawater.

Myxine glutinosa. (*Courtesy of Donald Flescher*)

Hagfishes are benthic, burrowing in mud that overlays substrates ranging in particle sizes from silt to rocks. Mud is an essential component of the habitat, as is full-strength seawater (35 parts per thousand). They feed on soft-bodied invertebrates, as well as the remains of crabs, shrimps, and fishes. Tough items such as the carcass of a large vertebrate require a technique called knotting. The hagfish grasps the flesh with its teeth and then throws its body into a simple overhand knot. Waves of muscular contractions move the knot toward the head, providing a mechanical means of tearing away the flesh. The same technique is used to clear the body of entrapment in its own slime, to avoid capture, and to escape from traps. Hagfishes support a rather large fishery worldwide, as thousands of tons are captured annually to supply manufactures of leather goods with "eelskins."

In the subfamily Myxininae, the efferent branchial ducts open by a pair of external gill openings, one on each side. The subfamily comprises four genera and about 25 species. In the Eptatretinae, the branchial ducts open separately to the exterior through 5 to 16 pairs of gill openings. This subfamily comprises three genera and about 45 species. Herbert Boschung

Bibliography. B. Fernholm, Hagfish systematics, pp. 33–44 in J. M. Jørgensen et al. (eds.), *The Biology of Hagfishes*, Chapman and Hall, London, 1998; P. Janvier, *Early Vertebrates*, Oxford Monog. Geol. Geophys. 33, Oxford University Press, 1996; J. M. Jørgensen et al. (eds.), *The Biology of Hagfishes*, Chapman and Hall, London, 1998; F. H. Martini and D. Flescher, Hagfishes: Order Myxiniformes, pp. 9–16 in B. B. Collette and G. Klein-MacPhee (eds.), *Bigelow and Schroeder's Fishes of the Gulf of Maine*, 3 ed., Smithsonian Institution Press, Washington, DC, 2002; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Myxomycota

Organisms that are classified in the kingdom Fungi and given the class name Myxomycetes, following the rules of botanical nomenclature; or classified in the kingdom Protista at various taxonomic ranks, as class Mycetozoa, following the rules of zoological nomenclature. Evolutionary origins are controversial, but many now believe, based on DNA sequencing techniques, that the Myxomycetes diverge early on the tree of life in the region where other protists are found.

The class consists of 3 subclasses, 6 orders, approximately 57 genera, and 600 species. Subclasses Ceratiomyxomycetidae, Myxogastromycetidae, and Stemonitomycetidae are distinguished by the type of sporophore development, type of plasmodium, and method of bearing spores (see **illus.**). The various orders, families, genera, and species are distinguished by characteristics of the fruiting bodies such as spore color, peridium, capillitium, calcium carbonate, or columella.

Distribution and habitats. Myxomycetes begin to appear in May and fruit throughout the summer until October in the north temperate regions. Many species are universally distributed and live in moist and dark places on decaying organic matter, such as decaying logs and dead twigs and leaves on the forest floor. Some species are restricted to more specialized habitats, which include on the bark surface of living trees and vines, under melting snowbanks in alpine regions, and on the dung of herbivorous animals.

Life cycle. Spores are released from the fruiting bodies when disturbed by animals, rain, or wind, and fall onto the substratum where, when water is present, they germinate and release protoplasts. The protoplasts may develop into either a myxamoeba or a flagellated swarm cell, both of which are haploid and behave like gametes (sex cells). The haploid (monoploid) gametes fuse in pairs forming diploid zygotes, which then divide mitotically without subsequent cell division, resulting in the formation of a multinucleated, free-living mass of unwalled protoplasm called the plasmodium. The diploid plasmodium is representative of the slime stage, and hence the common names sometimes used for this group of organisms include plasmodial, acellular, or true slime molds. The plasmodia ingest food as particulate matter (usually bacteria) by engulfment and are capable of growing to over 70 cm in diameter. The plasmodia may be colored red, orange, white, or almost black, but most are yellow, and are frequently seen in forested areas on moist, decaying logs, leaf litter, and fleshy mushrooms such as *Pleurotus*. The plasmodium undergoes a series of developmental stages, eventually becoming transformed into various types of usually colorful fruiting bodies, ranging in size from about 50 micrometers up to 70 cm.

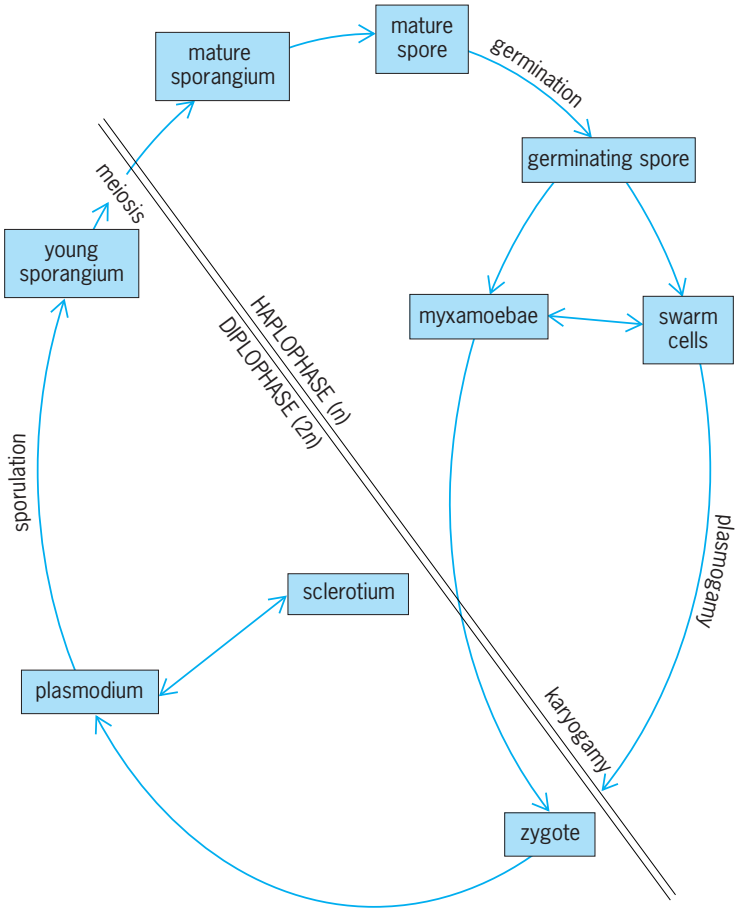

Life cycle of a typical heterothallic myxomycete.

Importance in research. The separate stages in their life cycle make myxomycetes ideal research laboratory organisms to study basic biological problems, ranging through protoplasmic streaming, the mitotic cycle, morphogenesis, aging, and cell division in cancerous cells. *See* EUMYCOTA. Harold W. Keller

Bibliography. C. J. Alexopoulos, W. C. Mims, and M. Blackwell, Phylum Myxomycota: True slime molds, in *Introductory Mycology*, 4th ed., Wiley, New York, 1996; L. Frederick, Phylum plasmodial slime molds, Class Myxomycota, in L. Margulis et al. (eds.), *Handbook of Protoctista*, Jones and Bartlett, Boston, 1990; H. W. Keller, The Myxomycota, in S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 1982; H. W. Keller, Biosystematics of Myxomycetes: A futuristic view, in C. Lado and J. C. Hernandez (eds.), *2d International Congress on the Systematics and Ecology of Myxomycetes: Abstract*

Volume, Real Jardin Botanico, Madrid, 1996; H. W. Keller and K. L. Braun, Myxomycetes of Ohio: Their Systematics, Biology and Use in Teaching, *Ohio: Biol. Surv. Bull., New Series*, vol. 13, no. 2, 1999; G. W. Martin et al., *The Myxomycetes*, 2d ed., 1983; S. L. Stephenson and H. Stempen, *Myxomycetes: A Handbook of Slime Molds*, Timber Press, Portland, OR, 1994.

Myxosporida

An order of the protozoan class Myxosporidea (subphylum Cnidospora). It is characterized by the production of spores with one or more valves and polar capsules, and by possession of a single sporoplasm with or without an iodinophilous vacuole. Myxosporidians are mainly parasites of fishes. They infect all parts of the body, including the heart and brain, and often induce considerable pathological changes in the host tissue.

Diagnostic features. Most species have two valves and two polar capsules, and are further identified by the size and shape of the spore. The size, shape, and position of the capsules are of diagnostic value, as are also various markings, prolongations, and appendages of the spore membrane. The exceptions to this general description are *Unicapsula*, with two valves and one polar capsule; *Chloromyxum*, with two valves and four polar capsules; *Kudoa*, with four valves and four polar capsules; and *Hexacapsula*, with six valves and six polar capsules.

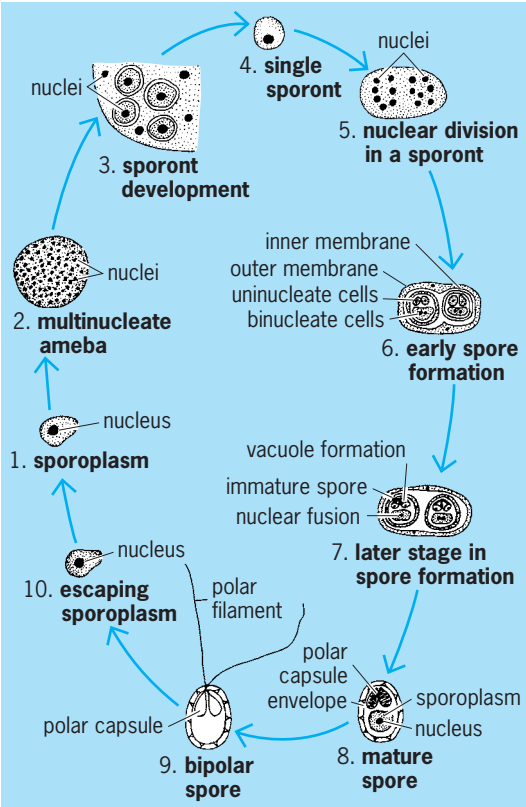

Life cycle of *Myxobolus*, a fish parasite.

Parasitic species. These four genera contain species that cause extensive deterioration of the flesh in the Pacific halibut, the common swordfish, the North Atlantic herring, and the Pacific yellowfin tuna, often making such commercially important fishes unfit for human consumption. Members of the genus *Myxobolus* may induce massive tumor-like growths in the skin of fresh- and brackish-water fishes. None, however, is pathogenic to humans or other mammals.

Life cycle. Infection begins with the ingestion of the spore by a host fish. The digestive fluids cause the polar filaments to be extruded, and at the same time the sporoplasm is released from the spore (see **illus.**). The sporoplasm, or amebula, reaches the specific site of infection directly through the gut wall or by way of the bloodstream. The amebula becomes a trophozoite when it starts feeding on the host tissues. The trophozoite then goes through a series of nuclear divisions (nucleogony) and, by a process of budding, gives rise to a number of cells, each of which eventually develops into a sporont. A sporont is a monosporoblast if one spore is produced and a pansporoblast if two or more spores for formed. The sporont undergoes a series of nuclear divisions, in which the number of nuclei produced will determine the number of spores and polar capsules to be formed. That is, in every spore one nucleus is involved in the formation of each valve and each polar capsule. Two nuclei become the gametic nuclei, which then fuse to form the zygotic nucleus of the sporoplasm. *See* CNIDOSPORA; MYXOSPORIDEA; PROTOZOA.

Ross F. Nigrelli

Myxosporidea

A class of the protozoan subphylum Cnidospora. Members of this class, which includes the orders Myxosporida, Actinomyxida, and Helicosporida, are parasites in fish, a few amphibians, and invertebrates. Actinomyxida with four families and Helicosporida with a single species (monotypic) do not contain subordinal categories. The Myxosporida, however, are divided into two suborders, the Unipolarina and Bipolarina.

Unipolarina is characterized by spores with one to six (never five) polar capsules located at the anterior end, except in some genera in which the capsules are widely separated or located in the central part of the spore but in which the polar filament is attached near the anterior end. The Unipolarina contains nine or more families that include important tissue (histozoic) destroying forms in fish such as *Unicapsula* (one capsule), *Myxobolus* (two capsules), *Chloromyxum* (four capsules), and *Hexacapsula* (six capsules). Most species have two anterior capsules and are included in such well-known and widely distributed genera as *Ceratomyxa, Myxosoma*, and *Henneguya*.

Bipolarina, containing a single family with three genera, is characterized by the presence of one capsule at or near each end of a fusiform or

ellipsoid spore. These parasites (coelozoic) are usually found in the gallbladder or urinary bladder of fish and a few amphibians. *See* ACTINOMYXIDA; CNIDOSPORA; HELICOSPORIDA; MYXOSPORIDA; PROTOZOA.

Ross F. Nigrelli

Myzostomida

Small, soft-bodied marine worms associated with echinoderms, mainly crinoids. They are found in all oceans from subtidal to over 3000 m (9840 ft) depth. These worms are often considered to be an order of Polychaeta, although the first molecular analyses based on DNA comparison estimated that they would be a protostome group nested outside the Annelida. The Myzostomida include about 180 species within eight families (see **table**). Most of them are ectocommensals of crinoids (that is, they live on their outer body surface without affecting them), but some are parasites of crinoids, asteroids, or ophiuroids that infest the gonads, coelom, integument, or digestive system. The association between myzostomids and echinoderms is very old: signs of parasitic activities, similar to those induced by extant myzostomids, are found on fossilized crinoid skeletons dating back to the Carboniferous. The body plan of most myzostomids is singular and differs from the regular body plan of polychaetes as they are incompletely segmented, parenchymous, acoelomate organisms with chaetae. *See* ANNELIDA; POLYCHAETA.

For most myzostomids, the body consists of an anterior cylindrical introvert (also called proboscis) and a flat, oval or disklike trunk (**Fig. 1**). The introvert is extended when the individual feeds, but it is retracted into an anteroventral pouch of the trunk most of the time. The trunk ranges from a few millimeters to 3 cm (1.2 in.) long. Five pairs of parapodia are located lateroventrally in two rows, each parapodium containing a protrusible hook, some replacement hooks, and a support rod (or aciculum). Most species have four pairs of slit- or disklike lateroventral sense organs, commonly named lateral organs, and the trunk margin often bears flexible needlelike cirri (more than 100 in some species).

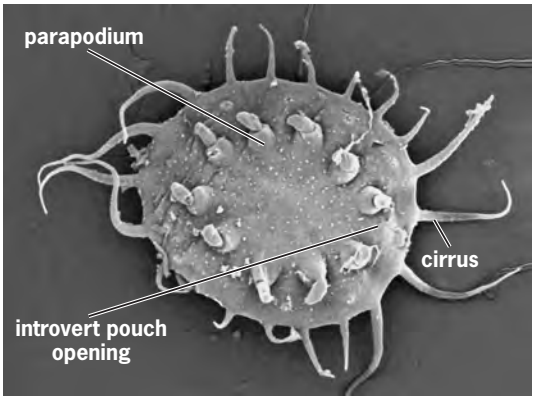

Fig. 1. Scanning electron microscopy view of the ventral side of the Indo-Pacific *Myzostoma ambiguum* showing the parapodia and the marginal cirri. The individual is 4 mm (0.16 in.) long.

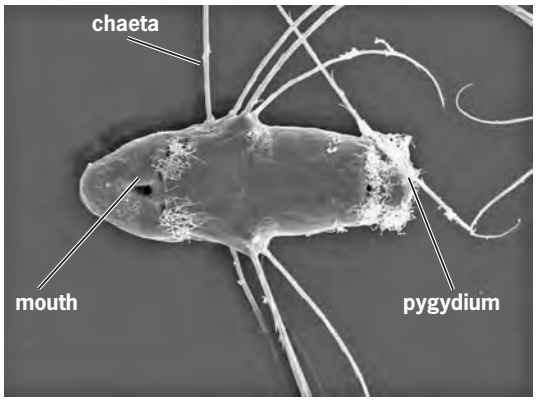

Fig. 2. Scanning electron microscopy view of the ventral side of the Atlantic *Myzostoma cirriferum* metatrochophore. The larva is 0.05 mm (0.002 in.) long.

Humplike or pointed cirri also occur at the base of each parapodium of about 20 species. Two male gonopores are located at the level of the third pair of parapodia, and the female gonopore opens close to the anus, posteroventrally. The body of parasites is often highly modified.

Myzostomids have a ventral nerve chain, circumpharyngeal connectives, and cerebral ganglia. The

Classification of the Myzostomida

Family	Genera	Hosts	Distribution
Myzostomatidae	*Myzostoma*, *Notopharyngoides*, *Hypomyzostoma*	Crinoidea, Ophiuroidea	Worldwide
Pulvinomyzostomatidae	*Pulvinomyzostomum*	Crinoidea	Mediterranean Sea
Endomyzostomatidae	*Endomyzostoma*, *Contramyzostoma*, *Mycomyzostoma*	Crinoidea	Worldwide
Mesomyzostomatidae	*Mesomyzostoma*	Crinoidea	Off Japan, Aru Island
Protomyzostomatidae	*Protomyzostomum*	Ophiuroidea	Russian Arctic, Bering Sea, Sakhalin, Japan, and Antarctic
Asteromyzostomatidae	*Asteromyzostoma*	Asteroidea	Russian Arctic, Antarctic, Atlantic
Asteriomyzostomatidae	*Asteriomyzostoma*	Asteroidea	Mediterranean Sea, off Southern California
Stelechopidae	*Stelechopus*	Crinoidea	Off Crozet Islands

species that live on crinoids feed on particles carried by the ciliated host's ambulacral grooves. The digestive system is complete and made of a pharynx included in the introvert, a stomach, an intestine, and two to three digestive caeca. There are six pairs of protonephridia. Most species are simultaneous hermaphrodites and reproduce by transferring spermatophores, followed by hypodermic sperm penetration. Fertilization is internal, and eggs develop in the water column into trochophore larvae which later elongates into a metatrochophore (**Fig. 2**). There is no circulatory system. Igor Eeckhaut

Bibliography. T. Bartolomaeus and G. Pürschke, *Morphology, Molecules, Evolution and Phylogeny in Polychaeta and Related Taxa*, Springer, Dordrecht, 2005; P. L. Beesley, G. J. B. Ross, and C. J. Glasby, *Polychaetes and Allies: The Southern Synthesis,* in *Fauna of Australia: Polychaeta, Myzostomida, Pogonophora, Echiura, Sipuncula*, CSIRO Publishing, Melbourne, 2000; D. Lanterbecq et al., Molecular phylogenetic analyses indicate multiple independent emergences of parasitism in Myzostomida (Protostomia), *Systematic Biol.*, 55(2):208–227, 2006.

N

Nailing — Nitroparaffin

Nailing

The driving of nails in a manner that will position and hold two or more members, usually of wood, in a desired relationship to each other. The contact pressures between the surfaces of the nails and the surrounding wood fibers hold the nails in position. Some types of nails are shown in the **illustration.**

Strength of a nailed joint. Factors that determine the strength and efficiency of a nailed joint are the type of wood, the nail used, the conditions under which the nailed joint is used, and the number of nails.

In general, hard, dense woods hold nails better than soft woods. The better the resistance of a nail to direct withdrawal from a piece of wood, the tighter the joint will remain. Nails driven into green wood tend to loosen slightly as the wood dries and shrinks. In seasoned material the resistance to withdrawal diminishes only slightly with time, unless moisture affects the wood. Withdrawal resistance is always higher when nails are driven into the side grain than when driven into the end grain. Because the lighter woods do not usually split as readily as the denser ones, more and larger nails can be used to offset the poorer nail-holding properties of the former. Hardwoods are more difficult to nail; they are sometimes used green or with holes drilled for nailing, to prevent splitting.

The surface condition of a nail affects its holding ability. The withdrawal resistance of a common nail increases directly with the distance it penetrates into the wood and increases almost directly with its surface area. A rusty nail may offer more resistance to withdrawal than a smooth one.

Means to increase withdrawal resistance. To increase resistance to withdrawal or loosening, nails may be coated, etched, spirally grooved, annularly grooved, or barbed. Grooved nails tend to hold well despite a change in moisture content. Coated

spiral-threaded, insulated siding, face nail	
annular-ring, gypsum board, dry-wall nail	
asbestos shingle nails: annular-ring, spiral-threaded	
annular-ring, plywood roofing nail for applying wood or asphalt shingles over plywood sheathing	
annular-ring, plywood siding nail for applying asbestos shingles and shakes over plywood sheathing	
spiral-threaded, casing head, wood siding nail	
annular-ring roofing nail for asphalt shingles and shakes	
spiral-threaded roofing nail for asphalt shingles and shakes	
annular-ring roofing nail with neoprene washer	
spiral-threaded roofing nail with neoprene washer	
insulated siding nail	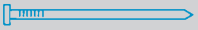
gypsum lath nail	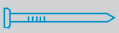
wood shake nail	
wood shingle nail	
roofing nail	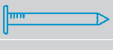
general-purpose finish nail	
sinker head, wood siding nail	
casing head, wood siding nail	

Special- and general-purpose nails.

nails usually provide a greater increase in withdrawal resistance when used in softer woods than in the denser woods. The increase in withdrawal resistance tends to decrease, however, with time.

In most cases, nails driven on a slant have more withdrawal resistance than nails driven straight into the wood. If a slant-driven nail is pulled in a direction which is at right angles to the surface, considerable resistance is encountered from the wood fibers on the pressure side. The nail may also progressively bend as it is pulled out. Both of these factors seem to offer continued holding power, even though the wood fibers are not gripping the entire surface of the nail. Nails slant-driven into the end grain of wood seem to gain proportionately more withdrawal resistance than those slant-driven into the side grain.

When members of a nailed joint tend to separate sideways, the nails are subjected to side loads. In this case doubling the diameter of the nail increases its lateral load capacity by nearly three times. This is true, however, only if the nail point has already been driven a suitable distance into the piece of material which is receiving it.

Blunt-pointed nails are often used to prevent the wood from splitting. Using nails of a smaller diameter also tends to prevent splitting but requires a greater number of nails per joint. Beeswax is sometimes applied to nail points to make them drive more easily, but it also reduces the holding power of the nail. *See* WOODWORKING.

Alan H. Tuttle

Najadales

An order of aquatic and semiaquatic flowering plants, division Magnoliophyta (Angiospermae), in the subclass Alismatidae of the class Liliopsida (monocotyledons). The order consists of 10 families and a little more than 200 species. The Potamogetonaceae, with about 100 species, are the largest family of the order, and the name Potamogetonales is sometimes used instead of Najadales for the group. The Najadales are Alismatidae in which the perianth, when present, is not differentiated into evident sepals and petals. Usually the flowers are not individually subtended by bracts. The evolutionary history of the order is in large part a story of floral reduction, associated with progressive adaptation to aquatic and eventually marine habitats. The Zosteraceae of this order are unique among flowering plants in that they grow submersed in the ocean, albeit in shallow water near the shore. *Zostera marina*, or eelgrass, is a common member of the family. *See* ALISMATIDAE; FLOWER; LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Nanochemistry

The study of the synthesis and characterization of materials in the nanoscale size range (1 to 10 nanometers). These materials include large organic molecules, inorganic cluster compounds, and metallic or semiconductor particles. The synthesis of inorganic materials of nanometer dimension is important because the small size of these particles endows them with unusual structural and optical properties that may find application in catalysis and electrooptical devices. Moreover, such materials may be valuable as precursor phases to strong ceramics. Approaches to the synthesis of these materials have focused on constraining the reaction environment through the use of surface-bound organic additives, porous glasses, zeolites, clays, or polymers. The use of synthetic approaches that are inspired by the biological processes result in the deposition of inorganic materials such as bones, shells, and teeth (biomineralization). This biomimetic approach involves the use of assemblies of biological molecules that provide nanoscale reaction environments in which inorganic materials can be prepared in an organized and controlled manner. Examples of biological assemblies include phospholipid vesicles and the polypeptide micelle of the iron storage protein, ferritin. *See* MICELLE.

Vesicles. Vesicles are bounded by an organic membrane that provides a spatial limit on the size of the reaction volume. If a chemical reaction is undertaken in this confined space that leads to the formation of an inorganic material, the size of the product will also be constrained to the dimensions of the organic host structure. This can be considered as analogous to producing inorganic materials in a soap bubble, except that the soap bubble is very, very small. If the chemical and physical conditions are not too severe to disrupt the organic membrane, these supramolecular assemblies may have advantages over inorganic hosts such as clays and zeolites because the chemical nature of the organic surface can be systematically modified so that controlled reactions can be accomplished. *See* SUPRAMOLECULAR CHEMISTRY.

Surfactant vesicles have been employed in a number of studies involving semiconductor, catalytic, and magnetic materials (see table). In the general method for preparing these materials, vesicles are formed spontaneously by sonicating aqueous solutions of phospholipids (**Fig. 1**). The presence of metal ions in the solution results in their encapsulation within the 25–50-nm internal volume of the enclosed aqueous space of the vesicle. The vesicle membrane is a bilayer of 4.5 nm thickness, and it prevents leakage of the metal ions back into the bulk solution. Thus, if metal ions are now exchanged from the bulk solution for inert cations such as the sodium ion (Na^+; by ion-exchange chromatography), the remaining entrapped ions can be subjected to chemical reactions solely within the intravesicular volume. The simplest procedure is to add a membrane-permeable coreactant such as gaseous hydrogen sulfide (H_2S). The H_2S rapidly diffuses through the phospholipid membrane and combines with the encapsulated metal ions to give an insoluble metal sulfide precipitate. As the number of entrapped metal ions is usually less than 10,000, semiconductor particles in the nanometer size range can be routinely produced. An extension of this method, in which an increase in the pH of

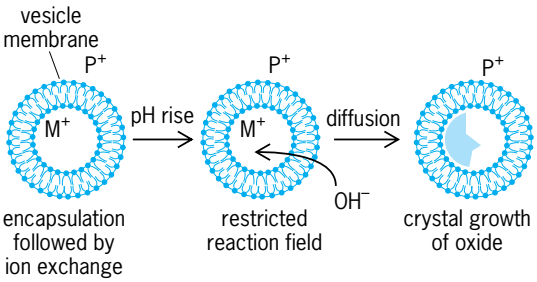

Fig. 1. Nanoscale synthesis of inorganic metal oxides by means of surfactant phospholipid vesicles. Cations (M^+) are encapsulated by sonication and replaced in the bulk solution by inert cations (P^+). Slow diffusion of hydroxide ions (OH^-) through the organic membrane results in intravesicular precipitation of metal oxides.

the bulk solution provides an excess of hydroxide ions (OH^-) that slowly diffuse through the surfactant membrane, results in the formation of metal oxide particles. In particular, nanophase iron oxides with catalytic and magnetic properties have been synthesized. *See* ION EXCHANGE; PHOSPHOLIPID; SONOCHEMISTRY.

The presence of the surfactant membrane in these nanoscale chemical reactions can have a profound effect on the structure and properties of the resulting inorganic materials. For example, the particles cannot come into direct contact, and the vesicles are usually charged; thus there is negligible aggregation or macroscopic precipitation. Under certain circumstances, the product is stable over many weeks as a monodisperse sol of finely divided particles.

Ferritin. One problem encountered with the use of phospholipid vesicles is their sensitivity to changes in temperature and ionic strength. Procedures have been developed in which the biomolecular cage of the iron storage protein, ferritin, has been used as a nanoscale reaction environment for the synthesis of inorganic materials (see **table**). Ferritin is a robust molecule constructed from 24 polypeptide subunits arranged into a hollow sphere of 8–9-nm internal diameter. The native protein contains a 5-nm-diameter core of a hydrated iron(III) oxide (ferrihydrite) within the internal cavity. Hydrophilic and hydrophobic channels penetrate the protein shell and provide the means by which iron atoms can be accumulated within or removed from the protein cavity. In the laboratory vessel, the iron can be readily removed by reductive dissolution to give intact empty protein cages (apoferritin).

Several approaches utilize ferritin in the production of inorganic nanoscale particles (**Fig. 2**). In the simplest approach the native iron oxide core is transformed into another material by chemical reaction within the protein shell. For example, exposure of the red-brown protein solution to H_2S results in a green coloration due to the formation of amorphous iron(III) sulfide (FeS) cores, approximately 7.5 nm in diameter. No precipitation is observed because the FeS particles remain encapsulated within the protein shell.

Alternatively, the native iron oxide cores can be removed from their protein shells by the use of appropriate reducing and chelating agents. The resulting empty apoferritin molecules are structurally intact, and can be readily reconstituted at room temperature and pH 7 by incubation of the protein in aerated Fe(II)-containing solutions. Moreover, other metal oxides can be formed within the protein cavity. For example, incubation of the empty protein cages with aqueous Fe(II) but at a pH of 8.5 and temperature of 65°C (149°F) results in the synthesis of nanometer-size magnetic iron oxides such as magnetite (Fe_3O_4) and maghemite (γ-Fe_2O_3) [Fig. 2]. The resulting protein, termed magnetoferritin, is magnetic. Other metals oxides, such as manganese oxyhydroxide (MnO·OH) and hydrated uranium trioxide (UO_3), can be synthesized within the protein cavity by incubation of apoferritin with the

Biomimetic approaches in the nanoscale synthesis of inorganic materials

System	Materials
Vesicles	Pt, Ag
	Cadmium sulfide (CdS), zinc sulfide (ZnS), silver sulfide (Ag_2S), cobalt sulfide (CoS)
	Silver oxide (Ag_2O)
	Iron oxyhydroxide (FeO·OH),
	Magnetite (Fe_3O_4)
	Aluminum oxide (Al_2O_3)
	Calcium phosphates
Ferritin	Iron(III) sulfide (FeS)
	Manganese oxyhydroxide (MnO·OH)
	Uranium trioxide (UO_3)
	Magnetite (Fe_3O_4)

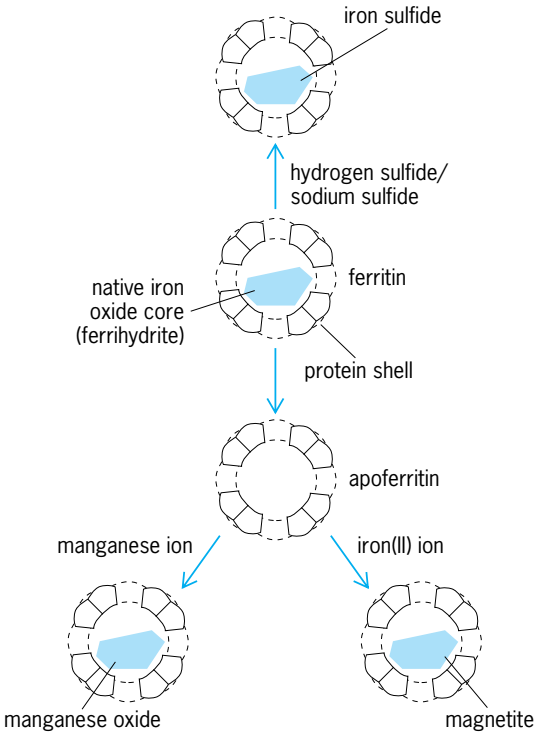

Fig. 2. Use of the supramolecular protein cage of ferritin in the synthesis of nanophase inorganic materials.

appropriate metal salt solutions in the presence of air.

Stephen Mann

Bibliography. G. Ozin and A. Arsenault, *Nanochemistry: A Chemical Approach to Nanomaterials*, 2005; G. B. Sergeev, *Nanochemistry*, 2006; H. Watarai et al. (eds.), *Interfacial Nanochemistry: Molecular Science and Engineering at Liquid-Liquid Interfaces*, 2005.

Nanoparticles

Synthetic particles that range from 1 to 100 nanometers in diameter. Semiconductor nanoparticles around 1–20 nm in diameter are often called quantum dots, nanocrystals, or Q-particles. These particles possess short-range structures that are essentially the same as the bulk semiconductors, yet have optical or electronic properties that are dramatically different from the bulk properties. The confinement of electrons within a semiconductor nanocrystal results in a shift of the band gap to higher energy with smaller crystalline size. This effect is known as the quantum size effect. In the strong confinement regime, the actual size of the semiconductor particle determines the allowed energy levels and thus the optical and electronic properties of the material.

Due to their finite, small size and the high surface-to-volume ratio, nanoparticles often exhibit novel properties. These properties can ultimately lead to new applications, ranging from catalysis, ceramics, microelectronics, sensors, pigments, and magnetic storage, to drug delivery and biomedical applications. Research in this area is motivated by the possibility of designing nanostructured materials that possess novel electronic, optical, magnetic, mechanical, photochemical, and catalytic properties. Such materials are essential for technological advances in photonics, quantum electronics, nonlinear optics, and information storage and processing.

Synthesis. A wide range of scientifically interesting and technologically important nanoparticles have been produced by both chemical and physical methods.

Colloidal. The synthesis of nanocrystals by colloidal methods involves nucleation (the initial formation of the appropriate semiconductor bond), growth (the formation of a highly crystalline core), and passivation of the nanocrystal surface. The passivation step is important in stabilizing the colloid and controlling the growth of the nanoparticles, preventing the agglomeration and fusing of the particles, and allowing the solubility of the nanoparticles in common solvents. A common approach is to use polymeric surfactants (for example, sodium polyphosphate). The polymer attaches to the surface of the growing particles, usually electrostatically, and prevents their further growth. Another passivation approach is to use capping agents such as thiolates. *See* COLLOID.

In the liquid phase, particle size is controlled by confining the growing particles within micelles or reversed micelles, polymer films, glasses, or zeolites. In the case of micelles, the micellar reagent (for example, dihexadecyl phosphate or dioctadecyldimethylammonium) creates a physical boundary, that is, a well-defined region where the semiconductor is precipitated. In reversed micelles, a small amount of water is mixed with a large amount of an organic solvent and surfactant. The surfactant molecules tend to collect on the surface and stabilize the waterdrop. The size of the waterdroplets is directly related to the ratio of the water to the organic phase. Smaller droplets result in smaller nanoparticles. Nanoparticle formation occurs by the reaction of two reagents (one which is soluble only in water and another which is soluble only in the organic solvent). Many nanoparticles have been produced by reverse micelles such as cadmium sulfide, selenide, and telluride, lead sulfide and selenide, and zinc and titanium oxides. *See* MICELLE; SURFACTANT.

Vapor-phase. The common chemical synthesis of metallic and intermetallic nanoparticles includes the decomposition of organometallic precursors, such as metal carbonyls (by thermal, photochemical, laser pyrolysis, and ultrasonic methods), to yield the respective elements or alloys, and the reduction of inorganic or organometallic precursors by reducing agents. Different sources of energy can be used to decompose the precursor such as microwave plasma, laser pyrolysis, laser photolysis, and flame combustion. The size of the nanoparticle is determined by the particle residence time, temperature of the chamber, pressure, and precursor composition. Low-temperature flames can also be used to supply the energy to decompose the precursors. Flame synthesis is most common for the production of oxides. Pure metal particles are best produced by gas condensation.

The vapor-phase synthesis of metallic nanoparticles involves evaporation of the material of interest, followed by the condensation of clusters and nanoparticles from the vapor phase. The vapor may be generated with laser or electron-beam heating methods. Laser vaporization provides several advantages, including the production of a high-density vapor of any metal, the generation of a directional high-speed metal vapor from a solid target (which can be useful for directional deposition of the particles), control of the evaporation from specific spots

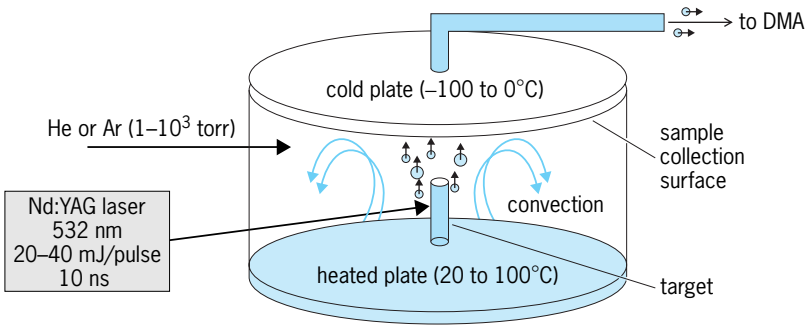

Fig. 1. Experimental setup for the synthesis of nanoparticles by the laser vaporization controlled-condensation method coupled with a differential mobility analyzer (DMA) for the size selection of the nanoparticles.

on the target, as well as the simultaneous or sequential evaporation of several different targets.

A technique that combines the advantages of pulsed laser vaporization with controlled condensation (LVCC) under well-defined conditions of temperature and pressure has been developed. It allows the synthesis of a wide variety of nanoparticles of metallic, intermetallic, oxides, and carbides of controlled size and composition.

The LVCC method consists of pulsed laser vaporization of a metal or an alloy target into a selected gas mixture in a convective atmosphere created by a temperature gradient between two plates separated by a glass ring (**Fig. 1**). A pure carrier gas such as helium or argon, or a mixture containing a known composition of a reactant gas (for example, oxygen for the synthesis of oxides, methane for carbides, and so on) can be used. A high-energy pulsed laser (532-nm Nd:YAG laser) with an intensity flux of about 10^6–10^7 W/cm^2 is focused on the target of interest. The resulting plasma causes highly efficient vaporization, and the temperature at the focusing spot can exceed 10,000 K (17,500°F). This high temperature can vaporize all known substances so quickly that the surrounding vapor stays at the ambient temperature. Typical yields are 10^{14}–10^{15} atoms from a surface area of 0.01 cm^2 in a 10^{-8} s pulse. The large temperature gradient between the bottom and top plates results in a steady convection current which moves the nanoparticles away from the nucleation zone (once condensed out of the vapor phase) before they can grow into larger particles. The nanoparticles are deposited as weblike aggregates on the top cold plate (**Fig. 2***a*), or transferred in a helium flow to a differential mobility analyzer (DMA) for size selections (Fig. 2*b*). Since the LVCC method produces a significant portion of charged nanoparticles (as ions or free electrons), the DMA classifies and separates the charged nanoparticles based on their electrical mobility in the presence of an electric field. *See* LASER.

Assembly of nanoparticles in electric fields. The assembly of nanoparticles into filaments and fibers, which retain the unique properties of the nanoparticles, holds promise for the development of novel functional materials and the engineering of a variety of nanodevices and sensors. The nanoparticles can be assembled into long-chain filaments (**Fig. 3***a*) and fibers (Fig. 3*b*) by applying an electric field during their synthesis by the LVCC method. The filaments can be few centimeters long, and tangle together with neighboring wires to form bundles. Silicon nanoparticles have a tendency to form dendritic structures with unique fractal patterns (Fig. 3*c*, *d*).

The effect of the electric field on the formation of the chain aggregates acts through the polarization of the charges on the nanoparticles' surface. For larger particles, the effect of the electrostatic charge is overpowered by gravity, but for nanoparticles the electrostatic forces are dominant. The nanoparticles' surfaces have a mixture of positive and negative charges, and some of the nanoparticles may have net charges. The dipole force is very strong near the surface of the particle, while farther away from the surface the net charge or monopole force becomes important. These two effects, when combined, may lead to the sticking together of particles of the same net charge.

Properties and applications. Semiconductor nanoparticles have technological applications in many areas of optoelectronics such as light-emitting diodes, solid-state lasers, and optical switches. Silicon nanostructures have stimulated much interest

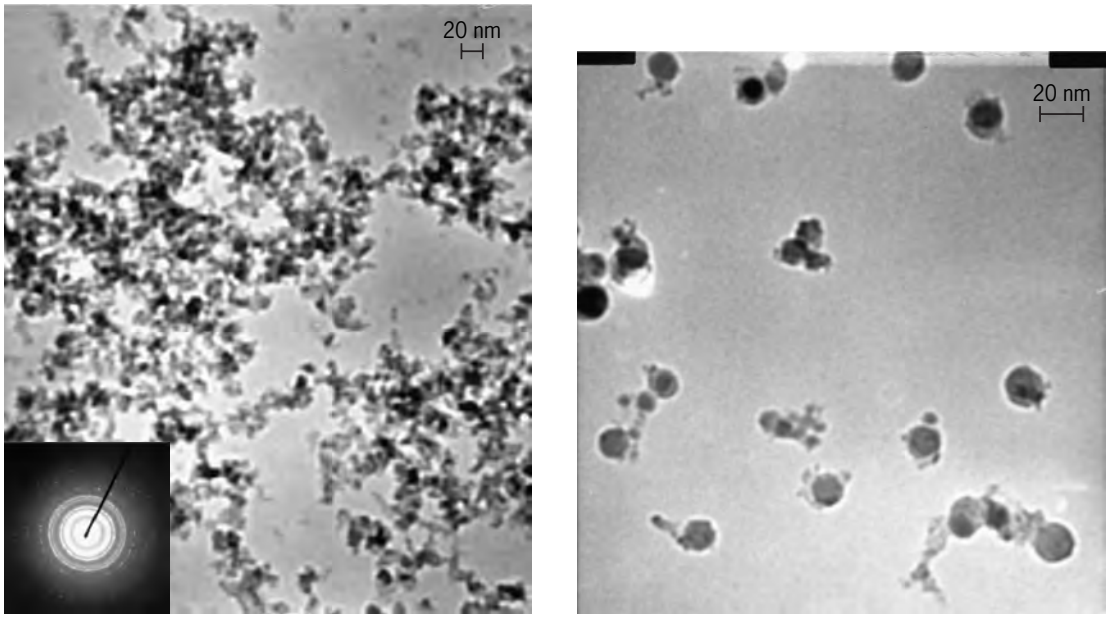

(a) (b)

Fig. 2. Transmission electron micrographs. (*a*) Iron aluminide (FeAl) nanoparticles prepared with the laser vaporization controlled-condensation method with no size selection. The inset shows electron diffraction. (*b*) 20-nm-diameter FeAl nanoparticles selected by the differential mobility analyzer.

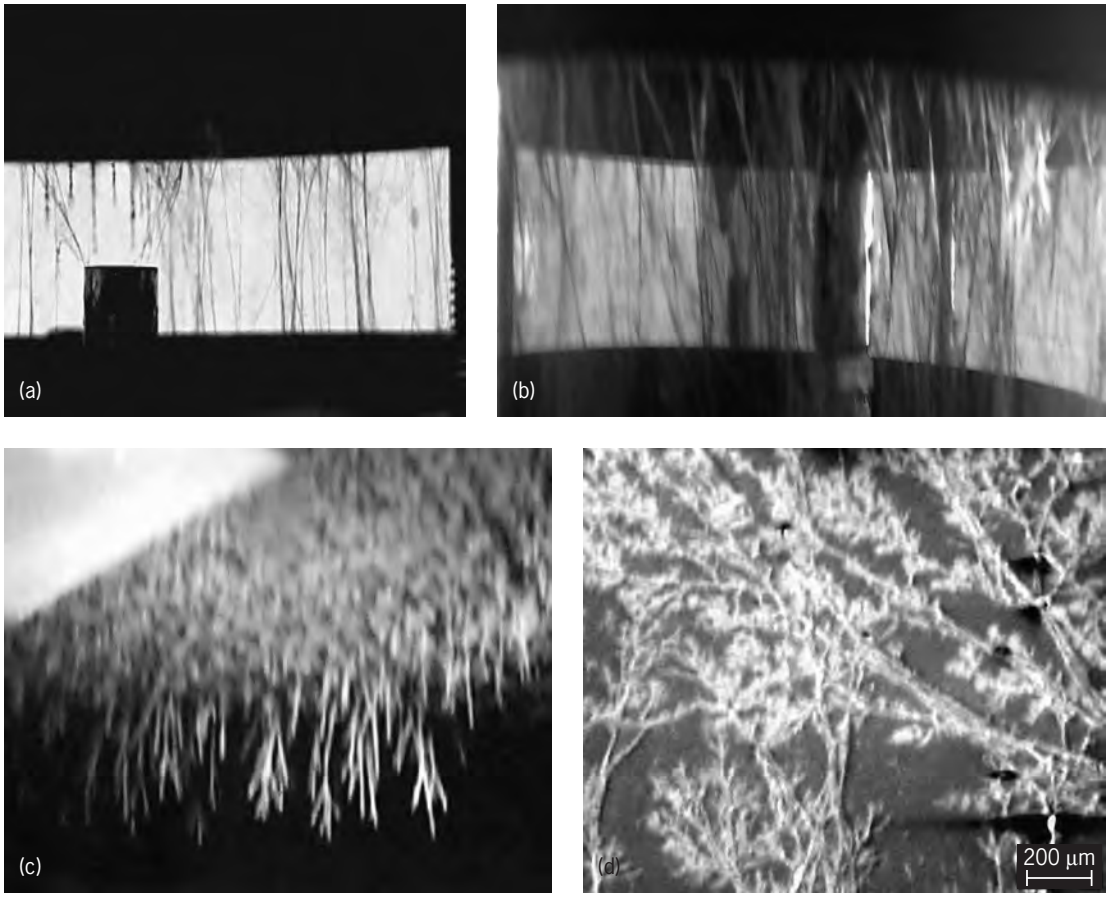

Fig. 3. Photographs of nanoparticle filaments of (*a*) iron aluminide (FeAl), (*b*) titanium aluminide (Ti_3Al), and (*c*) silicon, grown with the application of electric fields during the LVCC synthesis. (*d*) Scanning electron micrograph of a silicon nanoparticle dendritic assembly.

due to their unique properties, including single-electron tunneling, nonlinear optical properties, and visible photoluminescence. A transmission electron micrograph and electron diffraction pattern of silicon nanocrystals prepared by the laser vaporization with controlled condensation method are shown in **Fig. 4**. Because of its indirect band gap, bulk silicon does not exhibit visible photoluminescence. The higher energy shift in the photoluminescence of silicon nanocrystals is attributed to the three-dimensional quantum size effect. *See* BAND THEORY OF SOLIDS; MESOSCOPIC PHYSICS; QUANTIZED ELECTRONIC STRUCTURE (QUEST); SEMICONDUCTOR.

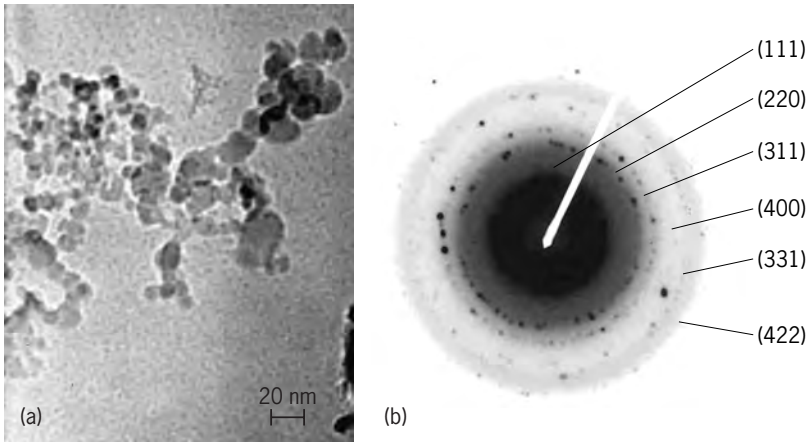

Fig. 4. Silicon nanoparticles. (*a*) Transmission electron micrograph; average particle diameter, 5–6 nm. (*b*) Electron diffraction pattern, showing diffraction rings corresponding to the planes of the randomly oriented nanocrystals.

The incorporation of silicon nanocrystals in polymer films may lead to the development of novel materials which combine several properties such as the visible photoluminescence and elasticity. Polymer films containing silicon nanocrystals exhibit the photoluminescence characteristic of the pure polymer and of the suspended silicon nanocrystals (**Fig. 5**). This may have interesting applications in the design of new materials for optical display and for silicon-based devices in optoelectronics.

Nanoparticles have important applications in catalysis and photocatalysis. The large number of surface and edge atoms provides active sites for catalyzing surface reactions. Highly nonstoichiometric oxide nanoparticles such as CeO_{2-x} provide a high oxygen-vacancy concentration, and active superoxide surface species. These nanoparticle oxides enable catalytic activation at significantly lower temperatures for the reduction of sulfur dioxide and oxidation of carbon monoxide. Other nanoparticles such as cadmium sulfide and zinc sulfide are efficient photocatalysts for generating hydrogen from water. *See* CATALYSIS.

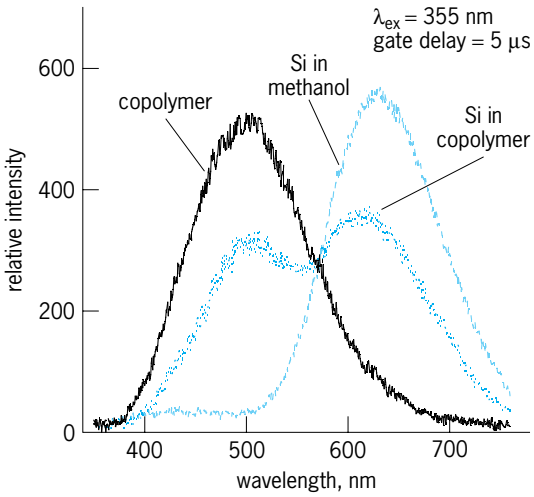

Fig. 5. Photoluminescence spectra of silicon nanocrystals suspended in methanol, sulfonated styrene-ethylene/butylene-styrene triblock copolymer film, and triblock copolymer film containing the silicon nanocrystals.

Molybdenum and tungsten trioxides are photochromic materials that change color upon oxidation from one state to another by the absorption of light. The photochromic effect is stronger in the nanoparticles than in the bulk material. The photochromic materials have potential practical applications in areas such as displays, imaging devices, "smart windows," and solar energy conversion.

In general, nanostructured materials possess smoother surfaces and enhanced mechanical properties as compared to conventional-grain-size materials. Nanostructured ceramics possess increased ductility since they do not contain many dislocations, and nanoparticles can diffuse fast into the materials' cracks once they are formed. Consolidated intermetallic materials based on nanoparticles show enhanced plasticity; that is, they exhibit significantly better elongations as compared to cast and powder processed components. The filamentlike and treelike assemblies of the nanoparticles may have some special applications as fillers (additives) to increase the elastic modulus (stiffness) and tensile strength (maximum stress before breaking) of low-strength oils and polymeric materials. *See* POLYMER; NANOSTRUCTURE.

Nanoparticles and functionalized nanostructured materials may have a significant impact in biological and biomedical applications. In drug delivery, nanoparticles could be used to deliver drugs where they are needed, avoiding the harmful side effects that often result from potent medicine. Prolonged residence times in blood have been achieved by nanoparticles, allowing the carriers to reach sites throughout the body without premature clearance. Because of the size of the nanoparticles, they can leave the vascular system, especially at sites of disease (such as tumors). Since nanoparticles are much smaller than cells, there is the possibility of subcellular targeting, that is, third-order drug targeting. This is key for gene delivery, where nuclear uptake is a requirement. Since nanoparticles have similar dimensions to natural carriers (for example, lipoproteins and viruses), they may serve as structural models for natural carriers. *See* DRUG DELIVERY SYSTEMS.

Nanomedical technology is opening new avenues in medical diagnostics and disease therapy. Since the size of a quantum dot determines the color it emits after exposure to light, different sizes of dots attached to different biological molecules can be used to track the activities of many tagged molecules simultaneously. Magnetic nanoparticles have been investigated as a potential alternative treatment for cancer. The hyperthermic effect (heat produced by relaxation of the magnetic energy of the magnetic nanoparticles exposed to an alternating magnetic field) can be used to effectively destroy tumor tissue.

M. Samy El-Shall

Bibliography. M. S. El-Shall et al., Synthesis of nanoscale metal oxide particles using laser vaporization/condensation in a diffusion cloud chamber, *J. Phys. Chem.*, 98:3067, 1994; M. S. El-Shall et al., Synthesis and photoluminescence of weblike agglomeration of silica nanoparticles, *J. Phys. Chem.*, 99:17805, 1995; M. S. El-Shall and S. Li, Synthesis and characterization of metal and semiconductor nanoparticles, in M. A. Duncan (ed.), *Advances in Metal and Semiconductor Clusters*, vol. 4, chap. 3, pp. 115–177, JAI Press, London, 1998; I. N. Germanenko et al., Effect of atmospheric oxidation on the electronic and photoluminescence properties of silicon nanocrystals, *Int. J. Pure Appl. Chem.*, 72:245–255, 2000; I. N. Germanenko, S. Li, and M. S. El-Shall, Decay dynamics and quenching of photoluminescence from silicon nanocrystals by aromatic nitro compounds, *J. Phys. Chem. B.*, 105:59–66, 2001; S. Li and M. S. El-Shall, Synthesis and characterization of photochromic molybdenum and tungsten oxide nanoparticles, *Nanostruc. Mater.*, 12:215, 1999; S. Li, S. Silvers, and M. S. El-Shall, Surface oxidation and luminescence properties of weblike agglomeration of silicon nanocrystals produced by laser vaporization—Controlled condensation technique, *J. Phys. Chem. B.*, 101:1794, 1997; Y. B. Pithawalla et al., Synthesis of magnetic intermetallic FeAl nanoparticles from a non-magnetic bulk alloy, *J. Phys. Chem. B.*, 105:2085–2090, 2001; W. Wang, I. N. Germanenko, and M. S. El-Shall, Room temperature synthesis and characterization of nanocrystalline CdS, ZnS and $Cd_xZn_{1-x}S$, *Chem. Mater.*, 14:3028–3033, 2002.

Nanostructure

A material structure assembled from a layer or cluster of atoms with size of the order of nanometers. Interest in the physics of condensed matter at size scales larger than that of atoms and smaller than that of bulk solids (mesoscopic physics) has grown rapidly since the 1970s, owing to the increasing realization that the properties of these mesoscopic atomic ensembles are different from

those of conventional solids. As a consequence, interest in artificially assembling materials from nanometer-sized building blocks, whether layers or clusters of atoms, arose from discoveries that by controlling the sizes in the range of 1–100 nm and the assembly of such constituents it was possible to begin to alter and prescribe the properties of the assembled nanostructures. (Many examples of naturally formed nanostructures can be found in biological systems, from sea shells to the human body.) *See* MESOSCOPIC PHYSICS.

Nanostructured materials are modulated over nanometer length scales in zero to three dimensions. They can be assembled with modulation dimensionalities of zero (atom clusters or filaments), one (multilayers), two (ultrafine-grained overlayers or coatings or buried layers), and three (nanophase materials) [**Fig. 1**], or with intermediate dimensionalities. Thus, nanocomposite materials containing multiple phases can range from the most conventional case in which a nanoscale phase is embedded in a phase of conventional sizes, to the case in which all the constituent phases are of nanoscale dimensions. All nanostructured materials share three features: atomic domains (grains, layers, or phases) spatially confined to less than 100 nm in at least one dimension, significant atom fractions associated with interfacial environments, and interaction between their constituent domains.

Multilayers and clusters. Multilayered materials have had the longest history among the various artificially synthesized nanostructures, with applications to semiconductor devices, strained-layer superlattices, and magnetic multilayers. Recognizing the technological potential of multilayered quantum heterostructure semiconductor devices helped to drive the rapid advances in the electronics and computer industries. A variety of electronic and photonic devices could be engineered utilizing the low-dimensional quantum states in these multilayers for applications in high-speed field-effect transistors and high-efficiency lasers, for example. Subsequently, a variety of nonlinear optoelectronic devices, such as lasers and light-emitting diodes, have been created by nanostructuring multilayers. *See* ARTIFICIALLY LAYERED STRUCTURES; LASER; LIGHT-EMITTING DIODE; SEMICONDUCTOR HETEROSTRUCTURES; TRANSISTOR.

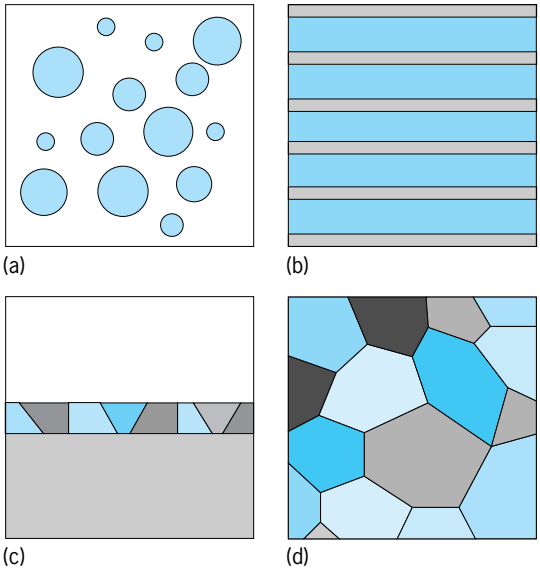

Fig. 1. Schematic of four basic types of nanostructured materials, classified according to integral modulation dimensionality. *(a)* Dimensionality 0: clusters of any aspect ratio from 1 to infinity. *(b)* Dimensionality 1: multilayers. *(c)* Dimensionality 2: ultrafine-grained overlayers (coatings) or buried layers. *(d)* Dimensionality 3: nanophase materials. *(After R. W. Siegel, Nanostructured materials: Mind over matter, Nanostructured Mat., 3:1–18, 1993)*

The advent of beams of atom clusters with selected sizes allowed the physics and chemistry of these confined ensembles to be critically explored, leading to increased understanding of their potential, particularly as the constituents of new materials, including metals, ceramics, and composites of these materials. A variety of carbon-based clusters (fullerenes) have also been assembled into materials of much interest. In addition to effects of confinement, interfaces play an important and sometimes dominant role in cluster-assembled nanophase materials, as well as in nanostructured multilayers. *See* FULLERENE.

A knowledge of the variation in cluster properties, both physical and chemical, with cluster size is important to both the fundamental understanding of condensed matter and the ability to use cluster-assembled materials in a variety of technological applications. The manner in which the structure and properties of collections of condensed atoms in a single cluster vary with cluster size, from atomic or molecular to bulk solid-state behavior, is fundamental in the development of realistic theoretical models for condensed matter. Theoretical areas that are impacted by this variation include understanding the forces acting among atoms, the structures of atom collections, electronic effects (quantum size effects) caused by spatial confinement of delocalized valence electrons, and cooperative (many-body) atom phenomena such as lattice vibrations or melting. *See* ATOM CLUSTER.

Synthesis and properties. A number of methods exist for the synthesis of nanostructured materials. They include synthesis from atomic or molecular precursors (chemical or physical vapor deposition, gas condensation, chemical precipitation, aerosol reactions, biological templating), from processing of bulk precursors (mechanical attrition, crystallization from the amorphous state, phase separation), and from nature (biological systems). Generally, it is preferable to synthesize nanostructured materials from atomic or molecular precursors, in order to gain the most control over a variety of microscopic aspects of the condensed ensemble; however, other methodologies can often yield very useful results. *See* CRYSTAL GROWTH.

Foremost in importance in nanostructuring is the ability to control the size and size distribution of the constituent phases or structures. The desirable sizes

are generally below 100 nm, since in this size range (and often below 10 nm) various properties begin to change significantly because of confinement effects. The chemical compositions of the constituent phases in a nanostructured material are also of crucial importance, as they invariably are to the performance of conventional materials. Finally, it is desirable to control the nature of the interfaces created between constituent phases and, hence, the nature of the interactions across the interfaces.

The properties of nanostructured materials are determined by the interplay among domain size, composition, and interfaces. In some cases, one or more of these features may dominate. Thus, it is desirable to synthesize nanostructured materials under controlled conditions, but with a focus on the particular property or properties of interest. The degree of control available, of course, depends upon the synthesis method used. The properties of these materials are significantly affected by their structural or compositional modulation, their spatial confinement, their interfaces, or a combination thereof. Spatial confinement can in general affect any property when the size of the atomic ensemble becomes comparable to or smaller than a critical length scale for the mechanism that is responsible for that property. Diverse examples include the blue (high-frequency, short-wavelength) shifts of the optical absorption in semiconducting clusters when their sizes fall below the Bohr radii (approximately 5–50 nm) of the excitonic (electron-hole pair) states responsible for absorption, and the increased strengthening of normally soft metals when their grain sizes fall below the critical length scales (less than about 50 nm) for the sources of dislocations (the defect responsible for easy deformation) to easily operate at conventional applied stresses. *See* CRYSTAL DEFECTS; EXCITON.

Examples. Examples of nanostructured materials that have been characterized include multilayers, individual and assembled atom clusters, and cluster-consolidated nanophase materials.

Magnetic multilayers. Magnetic multilayers, such as those formed by alternating layers of ferromagnetic iron and chromium, can be nanostructured so that the electrical resistance is significantly decreased (by up to a factor of 2 depending upon the chromium layer thickness) by the application of a magnetic field. Such an effect, called giant magnetoresistance, occurs when the magnetic moments of the neighboring alternating layers are arranged in an antiparallel fashion, so that application of the magnetic field overcomes the antiferromagnetic coupling and aligns the layers into a condition of parallel ferromagnetic ordering, strongly reducing the electron scattering in the system. Magnetoresistive materials have been introduced in the magnetic recording industry as read heads because of their lower noise and improved signal-handling capabilities. Other nanostructured materials besides multilayers also exhibit giant magnetoresistance, such as magnetic cobalt clusters embedded in a nonmagnetic matrix of copper or silver, or magnetic platelike nickel-iron deposits embedded in silver. *See* MAGNETORESISTANCE.

Optical properties of cluster assemblages. Noninteracting assemblages of small semiconductor clusters have optical properties of both scientific and technological importance. The optical absorption behavior of cadmium sulfide clusters with diameters in the nanometer size regime made by any of a variety of methods, including chemical precipitation in solutions or in zeolite supports, is different from that for bulk cadmium sulfide. The absorption edge is blue-shifted to appreciably shorter wavelengths because of the effects of quantum confinement in these nanoscale clusters. However, when these clusters are synthesized in zeolite supports with increasing loading such that they become close enough to begin to interact through quantum tunneling, the absorption edge begins to shift back toward bulk behavior (**Fig. 2**). A similar effect can be created by chang-

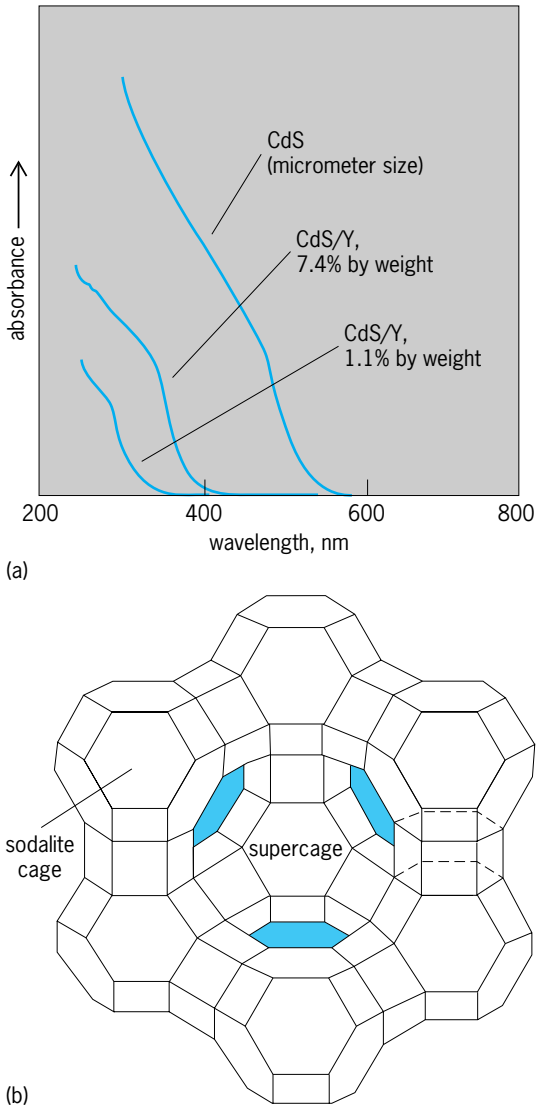

Fig. 2. Optical properties of cadmium sulfide (CdS) cluster assemblages. *(a)* Optical absorption spectra of CdS clusters in sodalite cages of zeolite Y for two different loadings compared to that for bulk CdS (micrometer size). *(b)* Cage structure of zeolite Y. (*After G. D. Stucky and J. E. McDougall, Quantum confinement and host/guest chemistry: Probing a new dimension, Science, 247:669–678, 1990*)

ing the sizes of the clusters in colloidal suspensions and thereby changing the degree of quantum confinement. Hence, control of the average distance between clusters in the zeolite cages, even though they are not actually in contact, or control of the cluster sizes in a suspension can enable control over a macroscopic property of the assembled cluster ensemble. Such quantum size effects provide a basis for verifying the understanding of the electronic structure of condensed matter, and may also provide for engineered optical properties. *See* ZEOLITE.

Chemical reactivity. Chemical reactivity of nanostructured materials, with their potentially high surface areas compared to conventional materials, can also be significantly altered and enhanced. Since clusters can be assembled by means of a variety of methods, there can be an excellent degree of control over the total available surface area in the resulting self-supported ensembles. Thus, it is possible to maximize porosity to obtain very high surface areas, remove most of it via consolidation but retain some to facilitate low-temperature doping or other processing, or fully densify the nanophase material. Also, control of chemical composition can be readily achieved, since rapid atomic diffusion paths are plentiful and diffusion distances are short in the clusters. Measurements of the decomposition of hydrogen sulfide over lightly consolidated, high-surface-area nanophase titania with a rutile crystal structure have clearly demonstrated the potential for enhanced chemical reactivity of nanophase materials. Nanophase rutile is far more reactive initially than any other available forms of titania and, more importantly, remains so even after extended exposure to the hydrogen sulfide at 500°C (932°F). This greatly enhanced activity results from a combination of unique and controllable features of the nanophase material, its high surface area combined with its rutile structure, and its oxygen-deficient composition, and makes this material suitable for a variety of catalytic and sensor applications. *See* NANOCHEMISTRY.

Nanophase materials. The assembly of larger atom clusters into bulk nanophase materials can also have dramatic effects upon properties. In this case, the clusters interact fully with one another, yet the effects of cluster size are still very important. Clusters of metals or ceramics in the size range 5–25 nm have been consolidated to form ultrafine-grained polycrystals that have mechanical properties remarkably different and improved relative to their conventional coarse-grained counterparts. For example, nanophase copper and palladium assembled from clusters with diameters in the range 5–7 nm can have hardness and yield-strength values up to 500% greater than in the conventionally produced metal. This greatly increased strength arises from the increased difficulty in the spatially confined grains of nanophase metals in creating and moving dislocations, the defect normally responsible for the relatively easy deformation process in metals. In ceramics, however, which are normally difficult to deform and hence very brittle, cluster assembly yields a different benefit. Ceramics can be rendered ductile (capable of easier deformation) by being synthesized from clusters with sizes below about 15 nm. This ductility results from the increased ease with which the ultrafine grains created by the clusters can slide by one another in a process called grain-boundary sliding, owing to the short diffusion distances required for the necessary local healing of incipient cracks that could otherwise form during this grain-over-grain sliding process. Nanocomposites consisting of metallic phases, ceramic and metallic phases (cermets), and ceramic phases in a variety of modulation dimensionalities also have considerably enhanced mechanical properties, including increased strength and fracture toughness. *See* CERAMICS; COMPOSITE MATERIAL; GRAIN BOUNDARIES; METAL, MECHANICAL PROPERTIES OF.

Richard W. Siegel

Bibliography. G. C. Hadjipanayis and R. W. Siegel (eds.), *Nanophase Materials: Synthesis, Properties, Applications*, 1994; P. Jena, S. N. Khanna, and B. K. Rao (eds.), *Physics and Chemistry of Finite Systems: From Clusters to Crystals*, 1992; B. H. Kear et al., *Research Opportunities for Materials with Ultrafine Microstructures*, National Materials Advisory Board, NMAB-454, 1989.

Nanotechnology

Techniques and products involving nanometer-scale structures, with dimensions ranging from 1 to 100 nanometers, especially those that transform matter, energy, and information using nanometer-scale components with precisely defined molecular features.

In the late 1980s, the term nanotechnology entered widespread use to describe anticipated technologies based on the use of molecule-based machine systems designed to build complex products with atomic precision. Since the mid-1990s, usage has broadened to embrace instruments, processes, and products in which key dimensions are in the 1–100-nm range. Technologies that fit this definition are extremely diverse, but many could potentially contribute to the development of new products and processes such as advanced molecular manufacturing.

Progress in nanotechnology can be judged by several metrics, including the increasing precision, complexity, cost-effectiveness, and scale of its products. The corresponding long-term objectives are atomic precision, arbitrary complexity, low-cost production, and large-scale products. This combination of objectives appears feasible, but only through a multistage process starting with the more limited capabilities of current nanoscale technologies.

Nanoscale technologies are extremely diverse, rapidly changing, and often only tenuously connected. Products include nanoscale particles, fibers, and films of diverse materials and structures; nanoscale lithographic structures for electronics (many integrated circuits now qualify); structures formed by spontaneous molecular aggregation (self-assembly); and solids containing nanoscale grains or pores. The means and materials used to produce

nanoscale and nanotextured structures often have little in common, and their applications range from stain-resistant clothing to state-of-the-art electronics. Many nanotechnologies are a continuation of preexisting fields under a new label. What they share (particularly toward the lower end of the 1–100-nm range) is the emergence of novel properties, relative to the corresponding bulk materials, associated with surface and quantum effects, together with a distinctive set of instruments and computational modeling techniques. Grouping these diverse nanotechnologies together has fostered a vibrant cross-fertilization of disciplines. *See* MOLECULAR SIMULATION; MONOMOLECULAR FILM; NANOPARTICLES.

Looking forward, the metrics of complexity and scale define the chief frontiers. In small structures, precision has already reached the atomic limit. Examples include quantum dots, engineered biomolecular objects, self-assembled molecular structures, and sections of carbon nanotubes. For systems built with atomic precision, scale limits complexity. Great complexity is possible, even in nanosystems of microscopic scale. For example, a cubic micrometer of a typical material contains roughly 10^{11} atoms; with generalized atomic control on that scale, a cubic micrometer could contain roughly 10^9 distinct functional components. *See* CARBON NANOTUBES; NANOSTRUCTURE; SUPRAMOLECULAR CHEMISTRY.

Although complex systems with precise molecular features cannot be made with existing techniques, certain nanosystems can be designed and analyzed. Systems based on mechanical (rather than electronic) degrees of freedom are particularly tractable. These are of special interest, because programmable nanoscale mechanical systems could be used to produce atomically precise structures of arbitrary complexity. The development of productive nanosystems is a key strategic objective.

Productive nanosystems. Nature demonstrates a class of productive nanosystems based on polymeric components operating in a liquid medium. Ribosomes, for example, work as digitally controlled machine tools that read genetic information (six bits per codon) and use it to direct the assembly of sequences of amino acids. The resulting polymers (proteins) fold to make nanometer-scale objects with precise arrangements of atoms. Together with nucleic acids (made by other programmable machines), these molecular objects form the working parts of a wide range of biological molecular machines, including ribosomes themselves. These larger structures are made through a self-assembly process in which Brownian motion brings components together and complementary molecular surfaces cause selected components to bind in a precise manner. *See* MOLECULAR RECOGNITION; PROTEIN; RIBOSOMES.

Biological systems provide not only examples of productive nanosystems but also tools to use and models for the development of next-generation systems. Developments in the design of novel protein and nucleic acid structures, including simple molecular machines, can be combined with a broad spectrum of current nanotechnologies in the construction of early-generation productive nanosystems. Self-assembly can join engineered biomolecular objects with atomically precise nanoparticles, fibers, and other products to make composite structures with properties beyond those found in biological systems.

As this path moves closer to fundamental physical limits, one direction of advance is toward better materials. Biomolecules have large monomers, low bond density, and low stiffness; advanced materials with higher bond densities can be stronger, stiffer, more fine-grained, and more regular. As with macroscopic machinery, the use of superior materials will enable better performance.

Although precisely structured devices with high bond densities cannot be made yet, they are amenable to computational design and modeling. Among the devices that have been analyzed are gears, shafts, bearings, belts, electric motors, computers, and programmable positioning mechanisms. These could be used to transform matter, energy, and information. Their components resemble those of conventional machine systems. For example, a design for a gearbox for transforming shaft power from higher to lower angular speeds would consist of a rigid framework supporting input and output shafts mounted on bearings. The gears would have teeth and obey the usual rules regarding gear ratios; however, in detail, the differences would be substantial. Each gear tooth typically would consist of a single row of atoms, with the smoothness of interatomic force fields enabling the gears and bearings to operate without an added lubricant (see **illustration**). Systems containing enough parts to serve the function of a programmable molecular assembly mechanism or a computer require several million to several billion atoms, and their volumes occupy a significant fraction of a cubic micrometer.

Basic principles. Nanotechnology based on productive nanosystems requires a combination of familiar molecular and mechanical principles in unfamiliar applications. In solution-phase reactions, molecules move by diffusion and encounter each other in all possible positions and orientations. The resulting

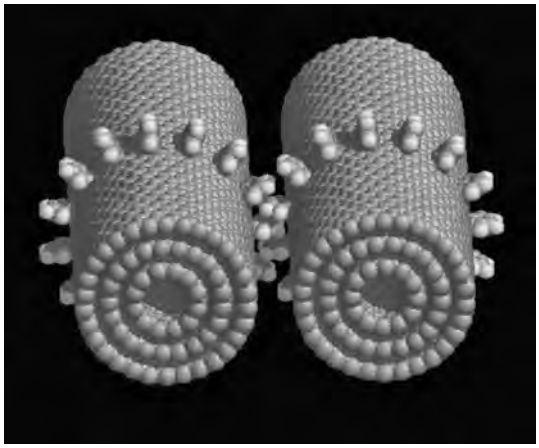

Simulation of a carbon nanotube–based gear. (*Courtesy of NASA Ames Research Center*)

molecular transformations are accordingly difficult to direct. Productive nanosystems, in contrast, could exploit mechanosynthesis, the use of mechanical devices to guide the motions of reactive molecules. By applying positional control to conventional molecular reactions, mechanosynthesis could cause structural changes to occur at precise locations in a precise sequence. Reliable positioning is required in order for mechanosynthetic processes to construct objects with millions to billions of precisely arranged atoms.

Mechanosynthetic systems are intended to perform several basic functions. Their first task is to bind raw materials from an externally provided source, typically a liquid solution containing a variety of useful molecular species. This process must separate molecules of different kinds and bind them reliably to specific sites. A second task (in advanced systems) is to transform bound molecules into highly active chemical species, such as radicals, carbenes, and strained alkenes and alkynes. Finally, mechanical devices can apply these bound, active species to a workpiece in controlled positions and orientations to deposit or remove a precise number of atoms of specific kinds at specific locations. *See* REACTIVE INTERMEDIATES.

To support these functions, it is necessary to provide appropriate mechanisms and conditions. Binding, transforming, and moving molecules can best be accomplished by nanoscale molecular machinery. To minimize friction, contamination, and side reactions, the ideal environment is free of fluids, which require a suitable enclosure and pumping mechanisms. If a mechanosynthetic system is to build complex products (rather than performing simple, repetitive operations), a programmable positioning mechanism could be used, requiring a source of instructions to guide its sequence of movements. Early-generation systems will sacrifice efficiency and range of products in exchange for simpler mechanisms, for example, by operating in an ambient fluid environment and using positioning mechanisms with fewer degrees of freedom.

Successful designs must overcome several challenges to reliable operation. Both quantum-mechanical uncertainty and thermal vibration cause random displacements in the positions of parts. For nanomechanical parts at room temperature, thermal vibration is overwhelmingly more important than quantum uncertainty. Because the mean-square displacement of a part is inversely proportional to the stiffness of the structure that holds it in place, the amplitude of thermal vibration could be limited in some circumstances by careful design. In mechanosynthetic systems, some vibrational amplitudes could be limited to less than one-tenth of an atomic radius. Because the probability of displacement has a Gaussian distribution, transient misalignments as large as an atomic diameter could be made extremely rare. *See* QUANTUM CHEMISTRY.

Infrequent, high-amplitude thermal vibrations can break even strong chemical bonds. Breaking a single bond in a molecular machine would typically cause it to fail. In the terrestrial ambient background radiation environment, structures with reasonable thermal stability will experience bond breakage chiefly from ionizing radiation. The rate of device damage due to ionizing radiation is roughly proportional to device mass, and devices on a scale of hundreds of nanometers will last many decades in the terrestrial environment before encountering radiation damage. To be reliable, larger systems either must be made of parts large enough to tolerate some damage or must be organized redundantly so that the system itself can tolerate some failed parts. Photochemical damage can be prevented by enclosing systems in opaque shields; 0.25-micrometer-thick aluminum is ample for long-term protection in full sunlight.

Reliable molecular manufacturing systems have strong similarities to digital computers. In conventional materials processing, as in analog electronics, all operations are somewhat imprecise. Each object produced has a unique size, shape, composition, and microstructure, differing both from the ideal design and from all previous objects. In productive nanosystems, though, as in digital electronics, each operation is either entirely correct or clearly wrong. In digital logic, the result of an operation is a specific pattern of ones and zeros. No stable intermediate states are possible, and physical principles enable the design of circuits that produce the correct pattern with high reliability. In productive nanosystems, the result of an operation would be a specific pattern of bonded atoms. For suitable choices of product structure, no ambiguous states are possible, and physical principles would enable the design of processes that produce the correct pattern with high reliability.

Applications. Much as digital computers can produce an indefinitely large range of patterns of information, productive nanosystems could produce an indefinitely large range of patterns of matter. For both digital computers and productive nanosystems, unlike traditional devices such as an adding machine or a lathe, it is difficult to describe the range of applications.

Because productive nanosystem technology could make precise, nanoscale features, these methods could be used to fabricate improved circuitry for digital logic. It is widely recognized that diamond would be superior to silicon for this purpose, if it could be fabricated with comparable ease. Advanced productive nanosystems could make this practical. Design calculations based on simpler, mechanical nanocomputers place a lower bound on what could be achieved. It appears feasible to build a central processing unit for a computer that occupies a volume less than 1 cubic micrometer, consumes roughly 0.1 microwatt of power, and executes about 1 billion instructions per second.

Nanotechnologies based on productive nanosystems will enable superior molecular-scale sensing and manipulation and thus allow development of a broad range of novel scientific and medical instruments. A particularly important goal is to provide better instrumentation for probing the molecular

structure of cells and providing data regarding structure and function at the molecular level in a more direct manner than is presently possible. The application of this knowledge to medicine could enable the development of nanoscale medical devices of greater complexity and capability than modern pharmaceuticals, and of far greater precision than modern surgical instruments. K. Eric Drexler

Bibliography. K. E. Drexler, *Nanosystems: Molecular Machinery, Manufacturing, and Computation*, Wiley/Interscience, 1991; D. S. Goodsell, *Bionanotechnology: Lessons from Nature*, Wiley-Liss, 2004; J. Storrs Hall, *Nanofuture: What's Next for Nanotechnology*, Prometheus Books, 2005; M. Wilson et al., *Nanotechnology: Basic Science and Emerging Technologies*, University of New South Wales Press, Sydney, 2002.

Naphtha

Any one of a wide variety of volatile hydrocarbon mixtures. They are sometimes obtained from coal tar, but more often they are derived from petroleum. Physical properties vary widely. The initial boiling point may be as low as 80°F (27°C), and end points may reach 500°F (260°C). Boiling ranges are sometimes as narrow as 20°F (11°C) or as wide as 200°F (110°C).

The main process for producing naphthas is fractional distillation. It may be of the extractive type when certain high-quality naphthas are desired. Acid treating, clay treating, and other techniques remove sulfur compounds and improve color, odor, and stability.

Strictly speaking, the refinery streams going into products like gasoline and kerosine are naphthas, and they are so designated within the petroleum industry. The final blended fuels, however, are sold under the more familiar names. The products sold as naphthas find their greatest use as solvents, thinners, or carriers.

Few naphthas are made up entirely of hydrocarbons belonging to one particular family. There is a fairly sharp differentiation, however, between aliphatic and aromatic types.

Aliphatic naphthas are relatively low in odor and toxicity and tend, also, to be low in solvent power, which in some cases is an advantage. In the processing of soybeans, for example, the aim is to extract the oil without extracting the less desirable materials. Naphthas used by dry cleaners likewise require only moderate solvent power. In printing ink, the naphtha is mainly a carrier of the carbon black or other pigments; resins requiring a solvent are present in only minor amounts.

The aromatic naphthas, often described as the high-solvency type, at one time came entirely from coal tar. The development of catalytic cracking and catalytic reforming made petroleum an alternative source. The main components are toluene and xylenes; benzene is less desirable because of the extreme toxicity of its vapors. A major use of these naphthas is as thinners for paints and varnishes, to permit easy brushing. Both varnishes and enamels contain large amounts of gums and resins, and diluents with good solvent action are therefore needed.

The rubber industry also uses naphthas as solvents. The leather industry uses them to degrease skins, the metal industry to degrease metals. Naphthas in insecticides and weedkillers dissolve the toxic agents and often contribute toxic properties of their own. Floor waxes, furniture waxes, shoe polishes, metal polishes, and dry cleaners' soaps are among the many other products in which naphthas are used. *See* PETROLEUM PROCESSING AND REFINING; PETROLEUM PRODUCTS. J. K. Roberts

Narcotic

A drug which diminishes the awareness of sensory impulses, especially pain, by the brain. This action makes narcotics useful therapeutically as analgesics. While they are the most powerful pain-relieving agents available, their use is complicated by a number of undesirable side actions. Indeed, much research in this area has been directed toward a search for an agent having the same degree of analgesic properties as morphine, the most widely used of the narcotics, but without its undesirable side actions. *See* ANALGESIC.

All of the generally used narcotics are in some way related to opium, and the term opiate is sometimes used interchangeably with the term narcotic. Opium is a gummy exudate obtained from the unripe seed capsules of the opium poppy. The narcotic effects of opium have been known from ancient times. Crude opium contains over a dozen alkaloids, all of which have been isolated and identified as to their structural chemistry. From this knowledge chemists have developed a number of synthetic chemical compounds, some of which have important advantages over the naturally occurring alkaloids. Therapeutically important natural alkaloids are morphine, codeine, and papaverine. Among the important synthetic narcotics are meperidine (Demerol), dihydromorphine (Dilaudid), oxymorphone (Numorphan), alphaprodine (Nisentil), anileridine (Leritine), piminodine (Alvodine), levorphanol (Levo-Dromoran), methadone (Dolophine), and phenazocine (Prinadol). *See* ALKALOID; OPIATES; POPPY.

Nalorphine (Nalline) is a narcotic antagonist and is used in the treatment of acute overdosage from narcotics; it is dangerous to drug addicts. Heroin is a highly addicting narcotic, and is so dangerous in this regard that the drug has been completely banned by both federal and state laws under all circumstances.

The pharmacology of narcotics is generally similar to that of morphine, the principal narcotic used for its analgesic effects. Differences among them lie in the potency of their action and in the degree and variety of the side actions which they produce. Effects are those of analgesia, accompanied

by a state of euphoria characterized by drowsiness and a change of mood from anxiety and tension to calmness and equanimity. There is a growing belief among practitioners that this state of euphoria is a valuable component of narcotic agents used for the relief of pain. For example, methadone which has high analgesic properties but little euphoria is not as widely used as morphine, which along with its relief of pain produces a relatively high degree of euphoria. It should be remembered that whatever narcotic is used, the effects are dose-related, and in higher doses all narcotics produce deep sleep and eventually general depression of all brain functions. Death from overdosage is due to depression of the respiratory centers with resultant failure of respiration.

Pharmacology

It is most convenient to discuss the pharmacological actions of the narcotics with specific reference to morphine, since it is the most widely used and typical narcotic analgesic.

Action. The predominant pharmacological effect of morphine (and the other narcotics) is on the central nervous system. From the standpoint of its medicinal use, its most important action is relief of pain. Researchers in the field of pain and analgesics have come to distinguish between pain produced by traumatic injury or physiological malfunction and the mental anguish which can be termed suffering. The former is a specific sensation resulting from sensory input, while the latter is the psychological reaction to this sensory pain. The narcotic analgesics relieve both pain and suffering, and indeed this double property is perhaps what has made them important for a long time in the practice of medicine. The action of morphine in the first regard is through depression of the pain reception centers of the primitive brain. Its action in the second regard is through a disruption of association pathways of the cortex, resulting in euphoria.

Side effects. Morphine classically illustrates the basic axiom in pharmacology that a drug always exerts multiple actions. Along with its valuable medicinal use morphine produces a great many undesirable side actions; the most frequent are worth mentioning, although they do not occur in every patient or at every dose level.

Depressed respiratory activity. Morphine is a depressant of the respiratory center of the brainstem. This depression is minimal in the usual doses used to produce analgesia, but it increases as the dose is increased. Death from ingestion of an overdose is from respiratory depression; treatment for overdose is directed toward maintaining respiratory function.

Effect on vomiting. Morphine and most of the other narcotic analgesics produce nausea and vomiting as the result of stimulation of the vomiting center of the brain. This illustrates the curious pharmacological effect of morphine on the brain, which is a mixture of depression of some brain centers and stimulation of others. One of the derivatives of morphine, apomorphine, is predominantly a stimulant of the vomiting center with no significant analgesic action. It is used to induce vomiting, such as in the oral ingestion of an overdose of a poison. Curiously, as the dose of morphine is increased, the stimulation of the vomiting center is replaced by a profound depression of this center and obliteration of the vomiting reflex.

Inhibited defecation and urination. The general effect of morphine on the gastrointestinal tract is depressive, exerted directly on the smooth muscle and the nerve plexi within the intestinal wall. Peristaltic activity is diminished or abolished, producing constipation if morphine is used for any length of time. This effect is reflected in the use of opium from earliest times as a treatment for the relief of diarrhea and dysentery.

Morphine produces an increase in contractile tone of the detrusor muscle at the neck of the urinary bladder. This results in urinary retention, and may precipitate problems in prostatic hypertrophy.

Other effects. With morphine the cutaneous blood vessels dilate and the skin becomes flushed and warm with an increase in perspiration. The pupil is constricted, and the "pinpoint" pupil is often present in narcotic addicts.

Contraindications. Morphine, and indeed all of the narcotic analgesics, is contraindicated as premedication in childbirth. These small molecules pass freely across the placental barrier into the fetal circulation and produce respiratory depression at a time when the establishment of respiration is particularly important and the function especially vulnerable immediately following delivery.

Physicians use narcotics with great caution in the very young and the very old, since these groups are particularly sensitive to narcotics.

Medical Uses

Despite the undesirable and unwanted side actions and the possibility of addiction, narcotic analgesics remain among the most useful drugs in medical practice. These drugs can be used safely and effectively by a skillful analysis of the patient, and by selection of the proper narcotic analgesic beginning (as in the case of postoperative pain) with morphine, changing in a short time to Demerol or codeine, and then as the intense pain diminishes making use of one of the effective nonnarcotic analgesics such as Darvon or even acetylsalicylic acid (aspirin). The danger of producing addiction is minimal if the narcotic analgesics are used for relatively short periods for the control of intense pain. In instances such as terminal cancer, the pain is severe and of long duration, and the need for relief of pain and suffering prevails over the possibility of producing addiction.

Drug Dependence

All narcotics have the potential for producing dependence and addiction when used repeatedly over a period of time. Drug dependence results from compulsive, continued use of the drug, and is

characterized by one or more of the following conditions: habituation, tolerance, or addiction.

Habituation. Like any other habit pattern, habitual use of a drug can develop. Common examples are the use of nicotine in the form of cigarettes, or caffeine in the form of coffee or tea. Such habituation is generally regarded as innocuous. Stopping the use of a habituating drug may be attended with mental distress, but there is no major physical disruption.

Tolerance. Repeated ingestion of a drug in which the effect produced by the original dose no longer occurs results in tolerance. To produce the original effect, it is necessary to increase the dose. A common example is the rapid disappearance of the central stimulant effect of caffeine in the habitual coffee drinker.

Addiction. When the body develops a dependence for the drug, addiction occurs. If the drug is suddenly stopped after a period of frequent use, a withdrawal syndrome develops, which is characterized by physical pain and widespread body reactions. The addict comes to dread the development of such painful and distressing reactions, and is trapped into continuing the drug.

All narcotics can produce habituation, tolerance, and addiction to a greater or less degree. Addiction to codeine is relatively rare but possible. Addiction to heroin develops rapidly, and this narcotic is therefore exceedingly dangerous.

Unquestionably, the nonmedical use of narcotics, with the concomitant social problem of the drug addict, is the result of the euphoric property of narcotics. Some cases of addiction are the result of unwise use of powerful analgesic narcotics for the relief of pain in patients; but by far the greatest number of cases comes from the use of narcotics for "escape." Heroin is almost universally used by narcotic addicts; its potency is greater than that of morphine and its euphoric effect is high. It is most effective when injected intravenously. Morphine is easily converted to heroin by a simple process of acetylation.

The total treatment of narcotic addiction is far from satisfactory. It is fairly easy under proper hospital conditions to manage the withdrawal syndrome and to get the addict off the drug. The most simple and brutal way of doing this is the "cold-turkey" method, in which the addict is placed in a room and allowed to go through the 2 or 3 days of agonizing withdrawal. A longer procedure is that of gradually decreasing the amount of drug which the patient receives over a period of 2 or 3 weeks. At the end of that time the drug may be stopped altogether with a minimum of withdrawal symptoms. In the methadone treatment methadone is substituted for heroin, and after a period of time it may be stopped without development of the withdrawal syndrome. Regardless of the method of treatment, the addict must be rehabilitated after getting off the drug. This is the most difficult part of the treatment and involves psychotherapy and social readjustment. Too often after a successful management of the acute condition of addiction, the addict returns to his old environment and habit patterns and once again becomes addicted. *See* ADDICTIVE DISORDERS.

Legal Regulations

Federal control of narcotics is embodied in the Harrison Narcotic Act, a basic law adopted by Congress in 1914. The purpose of the act was essentially twofold: to develop a source of revenue, and to bring the domestic control of narcotic agents into conformity with the nation's obligations to the Hague Convention of 1912, which was directed toward the control of international drug traffic. Because of the revenue aspects, the act was placed under the jurisdiction of the Treasury Department. It became the means for controlling domestic drug traffic and for preventing illicit sale and use of drugs by purveyors and addicts.

Drugs under control. Originally the federal legislation included opium and its derivatives, and coca leaves, from which cocaine was prepared. Later, additional drugs were included, some of which were newly discovered alkaloids obtained from opium; others were synthetic derivatives. Some of the drugs specified by the act are not narcotic; papaverine, for example, is a smooth-muscle relaxant; apomorphine is an emetic; and nalorphine is a narcotic antagonist. Of special interest was the inclusion in 1937 of marijuana and its derivatives. While marijuana is liable to abuse, it cannot be defined as a narcotic in a pharmacological sense. *See* COCA; MARIJUANA.

To add to the complexities of controlling the illegal use of dangerous drugs, the Food and Drug Administration was empowered under the Kefauver Law to control certain dangerous drugs. Among these are the barbiturates, tranquilizers, amphetamines, and lysergic acid diethylamide (LSD). To meet this responsibility, the Food and Drug Administration set up the Bureau of Drug Abuse Control. *See* BARBITURATES; PSYCHOTOMIMETIC DRUGS; TRANQUILIZER.

Jurisdiction. State laws concerning narcotics and other drugs liable to abuse vary extensively between the states and may be more stringent, but not less so, than the federal law. It can therefore be seen that the whole problem of drug abuse, from a legal standpoint, can become unwieldy and complex because of the division of responsibility between the Food and Drug Administration and the Treasury Department. In 1968 a reorganization was undertaken, and the whole problem of legal regulations and their enforcement was brought into the Justice Department.

Licensing provisions. The Federal Narcotic Act applies to different areas, such as import, export, manufacture, pharmacies, and hospitals, as well as to physicians, dentists, and veterinarians. The Bureau of Narcotics provides for four classes of licenses. Class I is generally used by physicians, and provides for the prescribing of the agents named in the Federal Narcotic Act on a properly written prescription to be dispensed by a licensed pharmacist. The class II license provides, in addition to the legal right to prescribe narcotics, the right to store and use these in the physician's office or clinic. Also, the physician

must have proper locked storage, keep an inventory of the stock on hand, and be prepared for an accounting to a federal agent at any time. Class III and class IV licenses apply to pharmacists, wholesalers, manufacturers, importers, and exporters. *See* MORPHINE ALKALOIDS; PAIN.

James M. Dille

Native elements

Those elements which occur in nature uncombined with other elements. Aside fom the free gases of the atmosphere there are about 20 elements that are found as minerals in the native state. These are divided into metals, semimetals, and nonmetals. Gold, silver, copper, and platinum are the most important metals, and each of these has been found abundantly enough at certain localities to be mined as an ore. Native gold and platinum are the major ore minerals of these metals. Rarer native metals are others of the platinum group, lead, mercury, tantalum, tin, and zinc. Native iron is found sparingly both as terrestrial iron and meteoric iron. *See* ORE AND MINERAL DEPOSITS.

The native semimetals can be divided into the arsenic group, including arsenic, antimony, and bismuth; and the tellurium group, including tellurium and selenium. The members of the arsenic group crystallize in the hexagonal system, scalenohedral class; those of the tellurium group in the hexagonal system, trigonal trapezohedral class. Only rarely do the semimetals occur abundantly enough to be mined as ores of their respective elements.

The native nonmetals are sulfur, and carbon in the forms of graphite and diamond. Native sulfur is the chief industrial source of the element. *See* MINERAL; MINERALOGY.

Cornelius S. Hurlbut, Jr.

Natrolite

A fibrous or needlelike mineral belonging to the zeolite family of silicates. It crystallizes in the monoclinic system in pseudo-orthorhombic prismatic crystals which are often acicular (see **illus.**). Most commonly it is found in radiating fibrous aggregates. There is perfect prismatic cleavage, the hardness is 5–5$^1/_2$ on Mohs scale, and the specific gravity is 2.25. The mineral is white or colorless with a vitreous luster that inclines to pearly in fibrous varieties. The chemical composition is $Na_2(Al_2Si_3O_{10}) \cdot 2H_2O$, but some potassium is usually present substituting for sodium.

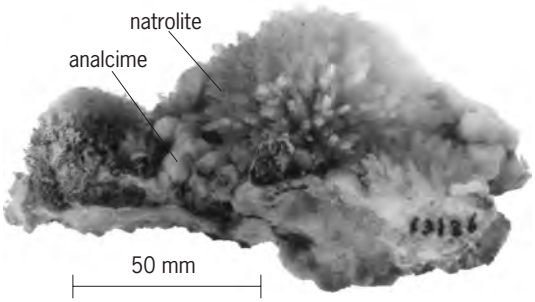

Natrolite crystals on analcime. (*American Museum of Natural History specimen*)

Natrolite is a secondary mineral (low-temperature hydrothermal mineral) found lining cavities in basaltic rocks, where it is associated with other zeolites, calcite, apophyllite, and prehnite. Its outstanding locality in the United States is at Bergen Hill, New Jersey. *See* SILICATE MINERALS; ZEOLITE.

Clifford Frondel; Cornelius S. Hurlbut, Jr.

Natural fiber

A fiber obtained from a plant, animal, or mineral. The commercially important natural fibers are those cellulosic fibers obtained from the seed hairs, stems, and leaves of plants; protein fibers obtained from the hair, fur, or cocoons of animals; and the crystalline mineral asbestos. Until the advent of the manufactured fibers near the beginning of the twentieth century, the chief fibers for apparel and home furnishings were linen and wool in the temperate climates and cotton in the tropical climates. However, with the invention of the cotton gin in 1798, cheap cotton products began to replace the more expensive linen and wool until by 1950 cotton accounted for about 70% of the world's fiber production. Despite the development of new fibers based on fossil fuels, cotton has managed to maintain its position as the fiber with the largest production volume, although its use has fallen. *See* MANUFACTURED FIBER.

Cordage fiber consumption has also changed over the years. True hemp was largely replaced by abaca (Manila hemp) during the nineteenth century, and by 1940 essentially all heavy twine and rope in the United States was made of sisal, henequen, or abaca. Now, in their turn, synthetic fibers supply more than half of the United States cordage requirements.

The natural fibers may be classified by their origin as cellulosic (from plants), protein (from animals), and mineral. The plant fibers may be further ordered as seed hairs, such as cotton; bast (stem) fibers, such as linen from the flax plant; hard (leaf) fibers, such as sisal; and husk fibers, such as coconut. The animal fibers are grouped under the categories of hair, such as wool; fur, such as angora; or secretions, such as silk. The only important mineral fiber is asbestos, which because of its carcinogenic nature has been banned from consumer textiles. *See* TEXTILE.

Plant Fibers

The basic biological unit of plant fibers is the plant cell, with a length frequently more than a thousand times its diameter. The cotton fiber is a single cell. The bast-fiber and hard-fiber strands are composed of many overlapping, parallel fiber cells held together in a fiber bundle that may be as long as the leaf or stem from which it comes. *See* PLANT CELL.

The fiber cells are made up of bundles of microfibrils which in turn are made up of ordered bundles of cellulose chains which are highly crystalline in nature. The parallel cellulose chains are held together laterally by hydrogen bonds, which are also involved in holding together the thin microfibrils and the internal structure of the cell. *See* CELLULOSE.

The direction that the microfibrils take in the cell determines the elongation character of the fiber. The cotton cell structure is characterized by gross convolutions and a spiraling of the microfibrils which frequently reverse along the length of the cotton fiber. In the bast and hard fibers, the microfibrils are more nearly parallel to the fiber axis (**Fig. 1**).

Cotton. The cotton plant is a member of the mallow family, Malvaceae. Its genus, *Gossypium*, includes herbs and shrubs that attain heights of 10 in. to 30 ft (25 cm to 9 m).

After pollination, the cotton boll takes about 50 days to ripen. Within the boll the ovules mature simultaneously with the covering seed hairs. During the early stages of growth after pollination, the hair grows only in a longitudinal direction, becoming tubelike and elongated, and is surrounded by a wall of cellulose. The outside of the latter membrane consists of wax and fatty substances, which are impermeable to gas and water, and they give the fiber its silkiness and luster. After 24 days another layer of cellulose is deposited on the inside of the cell wall. Concentric rings are formed, corresponding to the days of growth. The lint hair lives until the time of the boll opening. As the lint hair dries, its tubelike form collapses and becomes very twisted in its length. The characteristic twists correlate with spiral fibril structures in the secondary wall (**Fig. 2**). There are reversals of this twist during the convolutory growing process.

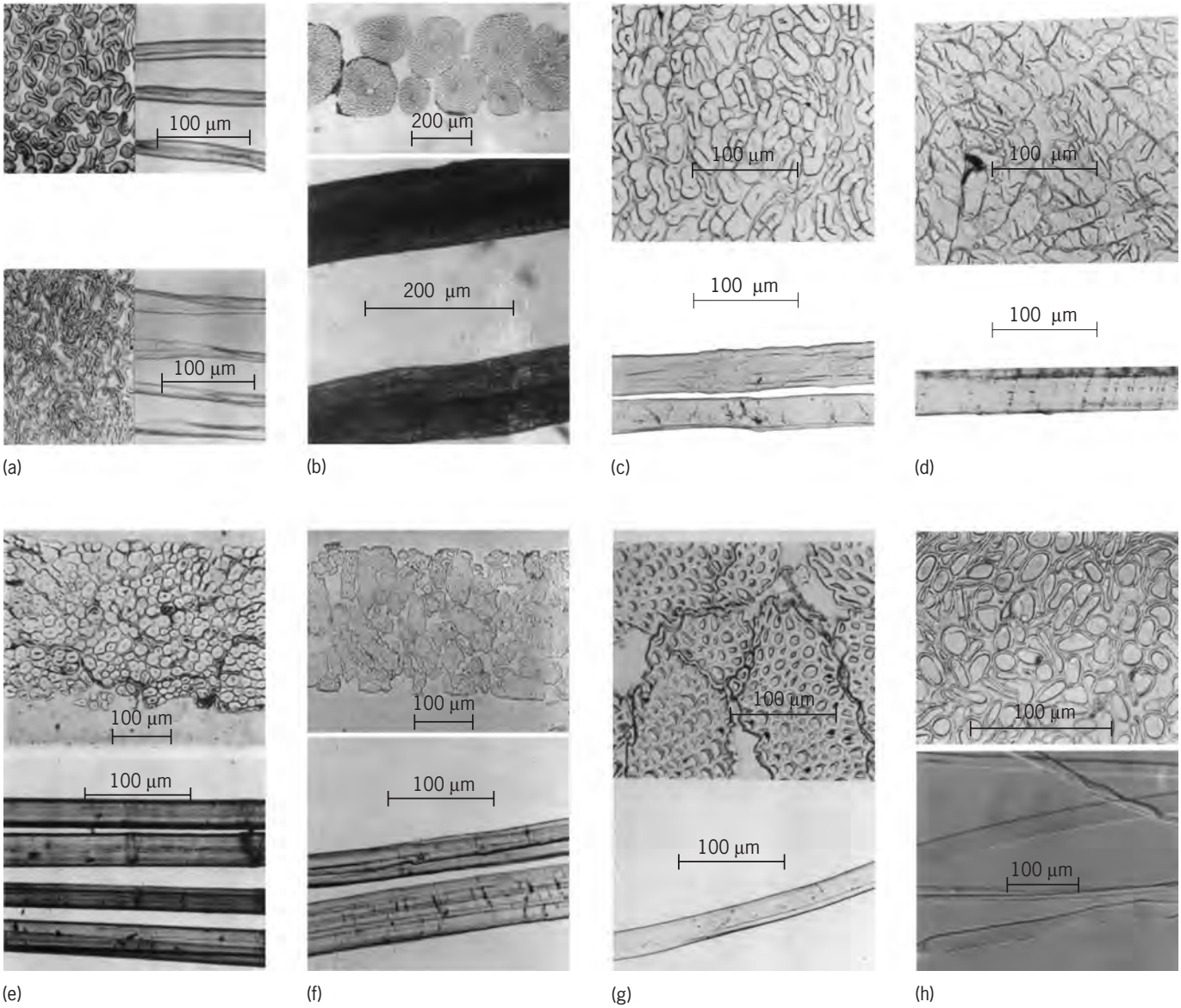

Fig. 1. Cross and longitudinal sections of important vegetable fibers. (*a*) Top: very immature cotton; bottom: average mature cotton. (*b*) Coir. (*c*) Ramie. (*d*) Hemp. (*e*) Flax. (*f*) Jute. (*g*) Sisal. (*h*) Kapok.

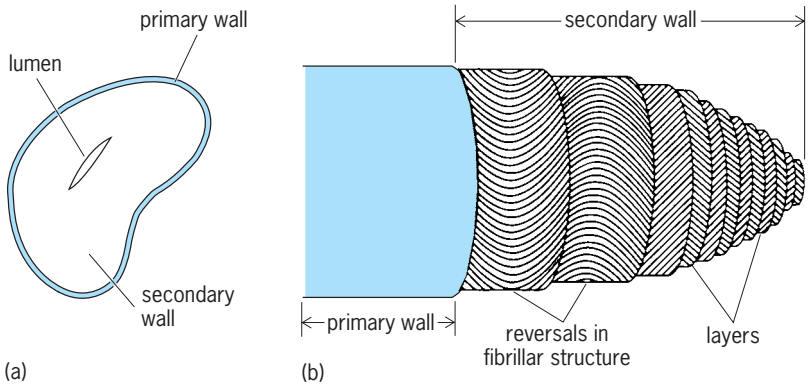

Fig. 2. Schematic of cotton: (*a*) cross section showing primary and secondary walls and lumen; (*b*) layered fibrillar structure of the walls. (*After B. F. Smith and I. Block, Textiles in Perspective, Prentice-Hall, 1982*)

After ripening, the cotton is harvested, ginned to remove the fiber from the seeds, and baled for transport to the spinsters. At the spinning plant, the graded and sorted fibers are further processed to remove trash and linters (very short fiber), and are twisted into yarn. Cotton yarn is either woven or knit into fabric which, depending upon its intended end use, may be subject to a variety of mechanical and chemical finishing treatments.

Fiber properties. The cotton fiber may be brown to creamy white; the whiter, more lustrous fibers are considered higher in quality. The fibers range in length from $^1/_8$ to $2^1/_2$ in. (0.3 to 6 cm), and may be from 10 to 20 micrometers in diameter. The longer, finer fibers are most prized. The fibers are moderately strong, with tenacities from 0.1 to 0.18 oz (3 to 5 g) per denier, dense (about 0.89 oz/in.^{3} or 1.54 g/cm^3), stiff (elongation at break of 3–7%), and stronger wet than dry. Cotton is resistant to bleaches, most household chemicals, and sunlight, but is susceptible to attack by mildew and some insects. Cotton's major advantages are low cost, comfort (moisture regain is about 8.5%), good durability, and ease of dyeing and laundering. Its major disadvantages are poor wrinkle resistance and wrinkle recovery, with a concomitant requirement for ironing, and tendency to absorb waterborne stains, such as grass stains.

Uses. Cotton is the most widely used natural fiber. It is found in apparel, home furnishings, and industrial products. In particular, the development of durable-press treatments for cotton-polyester blends has made these cloths the most widely used of all the apparel fabrics. *See* COTTON.

Bast fibers. These fibers, which come from the phloem or bast of dicotyledonous plants, are also called soft fibers. The greatest quantity of these fibers goes into fabrics, but they have other uses such as in thread, twine, or paper. The more widely used bast fibers are hemp, jute, kenaf, linen, and ramie.

Bast fibers, as they go to the spinner, are in bundles of strands of varying lengths, with the longest ones approaching the length of the stem. Each strand of the commercial fiber is formed by long, overlapping ultimate fiber cells. The cells in the strand and the strands in the bundles are held together by nonfibrous cementing substances composed of hemicelluloses, lignin, and pectinlike gummy substances which are alkaline-soluble.

The fibers are separated from the stem by a combination of mechanical and microbiological or chemical processes. The microbiological action, called retting, is of two types—water retting and dew retting. In water retting, the stems (or only the bark containing the fiber) are held beneath the surface of a tank, pool, or stream of water for several days, depending on the temperature, until bacteria break down the tissue surrounding the fiber and release it in the bark. In dew retting, the stems are spread on the ground after harvest and left for a few weeks while fungi invade the bark and release the fiber. Hemp and flax may be retted either way. Jute and kenaf are water-retted. Ramie is not retted, but the mechanically separated fiber is treated with chemicals.

Hemp. Hemp comes from the stems of the plant *Cannabis sativa*. The fiber is produced mainly in eastern Europe and central Asia. It is removed from the stem by retting and scutching. For centuries, hemp was the principal fiber used for marine cordage until replaced by abaca and sisal. It is still used extensively for twine and for many of the same products as linen.

The water-retted fiber is creamy white to brown; dew-retted fiber is gray to dark gray. The line fiber is usually about 4–7 ft (1–2 m) in length. The tow is also used for spinning, but goes into the coarser yarns. The quality of hemp fiber is based on factors similar to the ones for linen.

The fiber cells average about 22 micrometers in diameter and about 1 in. (25 mm) in length and are rather uneven in diameter. The lumen is larger than in linen and frequently, especially at the ends, is flattened so it appears as a line. The ends of the fibers are blunt and very thick-walled, and show some branching. This branching is a distinguishing characteristic of hemp fiber, which otherwise looks very much like linen under the microscope. Hemp nearly always exhibits some cells with forked or nodulated ends, whereas linen does not.

The manufacturing of hemp is similar to that of lower grades of linen, although the products are coarser. *See* HEMP.

Jute. Jute fiber is obtained from the stems of white jute (*Corchorus capsularis*) and from tossa and daisee (*C. olitorius*). The two species are similar, but *C. olitorius* fiber is slightly stronger and finer. The fiber is removed from the stem by retting and stripping the retted fiber from the stems by hand. Jute is grown mainly in Bangladesh and India, but Burma, Nepal, and Brazil produce sizable amounts.

Its suitability for uses requiring strong, bulky fabric or twine and its low cost compared to competitive fibers have made jute second only to cotton among the natural fibers in world consumption. The greatest quantity on a worldwide basis goes into fabric—hessian (known as burlap in the United States) and a heavier-weight fabric known as sacking. These are the standard wrapping and bag materials.

Jute fiber, usually 5–10 ft (1.5–3 m) or more in length, is creamy white to red-brown or gray, depending mainly on the species and the kind of water used for retting. The cells that make up the fiber strands average about 0.008 in. (2 mm) in length and about 20 μm in diameter, and are held together by natural plant gums. The longitudinal view shows the cells to be roughly cylindrical, with the ends varying from spearhead-shaped to tapering points. The cross-sectional view shows a polygon with pronounced oval lumen. Nodes or cross markings are usually absent, but are occasionally found.

While there is about as much cellulose in jute (about 75%) as in other retted bast fibers, it is a lignocellulose and the fiber is harsh.

Synthetic fibers, especially polypropylene, have made substantial inroads in the markets for jute. Polypropylene bags and prime back for tufted carpeting have displaced large quantities of jute. Also, bulk handling has eliminated much of the former market for grain bags, especially in the United States. *See* JUTE.

Kenaf. The name kenaf originally referred to the fiber *Hibiscus cannabinus*, but now also includes *H. sabdariffa*. Among some 125 common names for the fiber and plant, frequently used are Bimlipetam jute, mesta (or meshta), Thai (or Siamese) jute, Java jute, roselle, ambari, deccan hemp, stokroos, and dah.

The countries producing the most kenaf are India and Thailand. Other producing countries are Russia, Bangladesh, Indonesia, and Vietnam.

Kenaf is traditionally a substitute or alternate crop and fiber for jute. Its uses as a textile fiber are the same as for jute, and its products are usually sold as sacking, bagging, and so on, without reference to their containing kenaf. An important and increasing use is for paper.

Kenaf's appearance is similar to jute, although the fiber is somewhat lighter in color. Its strength is similar, but it is considered by some spinners to be less strong.

Linen. Flax is grown mostly in northern and eastern Europe. The fiber, linen, is separated from the stem of the plant by a system of retting and scutching. The line (long, parallel) fiber is used principally for fine damask, table linen, apparel fabrics, sheeting, lace, threads, and special twine. The tow (short, tangled fiber produced in the scutching and the hackling operations) and some of the coarser grades of line fiber are used for toweling, canvas, mail bags, meat wraps, twine, and high-grade papers such as cigarette, bond, and currency.

The factors that determine the quality of linen fiber are cleanliness, fineness, strength, density (as a hank of fiber held in the hand), color, uniformity, luster, length, and softness. Under the light microscope, the surface appears smooth to lightly marked longitudinally, with frequent transverse lines, often in the form of an X. In cross section, the cells appear somewhat polygonal to round, with cell walls that are fairly uniform in thickness (**Fig. 3**). The lumen is small and distinct, and disappears near each end as the cell tapers to a point.

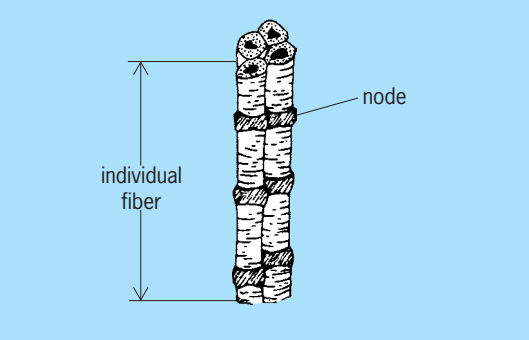

Fig. 3. Individual fibers in the flax stem. (*After B. F. Smith and I. Block, Textiles in Perspective, Prentice-Hall, 1982*)

The fiber varies in color from light ivory to dark gray, in length from 1 to 3 ft (30 to 100 cm), and is about 20 μm in diameter. It is strong (tenacity about 0.21 oz or 6 g per denier), is very stiff (elongation at break from 2.7 to 3.3%), has a moisture regain of about 11%, and is very smooth. Its properties are very similar to those of the very best grades of cotton. *See* FLAX.

Ramie. Ramie fiber is removed from the stem of the plant (*Bohmeria nivea*) by a beating and scraping operation that leaves it in ribbons. These ribbons are dried and later chemically degummed, usually by the spinner. Thus, ramie is not retted as other bast fibers are.

The world's largest producer is believed to be the People's Republic of China. The only other countries that export the fiber are Brazil and the Philippine Republic.

The color of raw ramie (China grass) is creamy, green, or straw to light brown, and the length averages from 28 to 60 in. (70 to 150 cm). The degummed and bleached fiber and fabric are white.

Ramie fiber cells are longer than those of any other plant fiber, averaging more than 6 in. (165 mm) in length. The cell diameter averages 50 μm with considerable variation between and within varieties. Under the microscope, the diameter of the fiber appears uneven, sometimes with heavy cell walls and well-defined lumen, at other times broad and flat with an indistinct lumen but showing heavy striations along the fiber.

Ramie textiles have a natural luster, are very strong and absorbent, and tend to be stiff. They gain strength when wet and are highly resistant to mildew and rot. *See* RAMIE.

Hard fibers. Hard fibers, also known as leaf or cordage fibers, come from the leaves of monocotyledonous plants. The fiber of the leaves of many species is used, some in international trade, but the three consumed in greatest quantity are abaca, sisal, and henequen. On the basis of abaca strength as 100, sisal would be about 75 and henequen about 60.

Long, parallel, overlapping cells form the hard fibers as in bast fibers. However, the leaves are not

retted or treated chemically as are the stems of various bast fibers.

Abaca. Frequently known as Manila hemp, abaca comes from the pseudostem (modified leaf petioles that look like a tree trunk) of *Musa textilis*. Most abaca is grown in the Philippines.

The fiber is removed from the plant by methods that range from hand stripping to a factory-type of defibering (decortication).

Abaca fiber color ranges from almost white through cream, light brown, and very dark brown. The fiber is long, usually more than 9 ft (3 m), and ranges from about 3.3 to 16.5 ft (1 to 5 m) in length.

The cells of abaca average 0.2–0.24 in. (5–6 mm) in length and 25–30 μm in diameter. The fiber cells are uniform in diameter, lustrous, and rather thin-walled. The lumen is large and distinct. An unusual feature is the frequent appearance of silicified plates about 0.0012 in. (0.03 mm) in length and often in long chains. The fiber is a lignocellulose, about 75% cellulose and about 8% lignin. Abaca is traditionally used in marine cordage and in other applications that require a strong, water-resistant rope, although its use in cordage is being challenged by manufactured fibers such as nylon, polyester, and polypropylene. *See* ABACA.

Sisal and henequen. The sisal plant (*Agave sisalana*) is grown mainly in Africa and Brazil, while henequen (*A. fourcroydes*) is grown mainly in Mexico and Cuba. These are cordage fibers removed from the leaves by defibering. The greatest use for these fibers is for twine and rope.

Sisal fiber is stiff and harsh, usually 28–56 in. (70–140 cm) in length, and a creamy white color. Henequen tends to be lighter in color and slightly longer, but is not as strong as sisal.

The cells of sisal and henequen are about the same size and chemical makeup. They average about 0.1–0.12 in. (2.6–3.0 mm) in length and 20 μm in diameter. They show an unusual broadening toward the middle, and the ends are broad, blunt, thick, and at times forked. The width of the lumen is frequently greater than that of the cell wall. These lignified fibers have about 75% cellulose and 10% lignin. *See* HENEQUEN; SISAL.

Animal Fibers

Many animal species produce fibrous structures. These include the open webs of spiders, the more closed cocoons of moths, the fibrous part of bird feathers, and body hairs found on species from insects to mammals, including humans. The common characteristic of these fibrous structures is that they are all denatured proteins. Silk and mammalian hairs have achieved commercial and technical importance. *See* PROTEIN.

Mammalian fibers. Mammalian fibers have long been used by humans in the form of fur, all types of textiles, fiber fills, ropes, brushes, wigs, and so on. All mammalian fibers are so closely related that their chemical and physical fundamentals can be discussed together.

Fiber formation. The mammalian skin is densely populated with microscopic channels, the follicles, from which the fibers emerge. As the cell mass is pushed outward, its components are structurized and denatured in a process called keratinization. Hence, all mammalian hairs are called keratin fibers. By the time the fibers reach the surface of the skin, they are fully hardened and are no longer living tissue.

Many species produce a double coat, containing both long, coarse, and somewhat sparse guard hairs and a shorter, fine, and dense underfur. *See* HAIR.

Shapes and structure. The fibers are circular or oval in cross section, with slight tapering at the tip end. The coarser guard hairs can be circular, strongly flattened, or triangular in cross section, and vary widely in thickness along their length. Diameters can range from a few micrometers, as in vicuña, to millimeter size, as in elephant tail fibers.

Morphologically, the fibers are built from two important cell types: the central cortex and the outer-covering cuticular cells (**Fig. 4**). The thin, flat cuticular cells overlap each other in a fish-scale fashion, their outer ends pointing toward the tip of the fiber. The cuticular layer acts as a protective layer for the cortical portion of the fiber. The cortical cells are elongated spindles in shape.

In some wool fibers two different types, ortho and para, are differentiated by chemical means. In addition, the fibers may possess a fragmented or continuous medullary core, consisting of cells which can be partially filled with cellular debris. The medullary volume is high in some hairs, but is completely absent in fine wool.

Colored animal fibers contain a granular pigment, called melanin, which is distributed mostly within the cortex. The shade and strength of color of the otherwise colorless keratin fiber is determined by the amount of and particle size of this pigment.

Chemical composition and fine structure. Keratin is a polymer of 20 α-amino acids which differ from each other in the composition of their side groups (**Fig. 5**). The

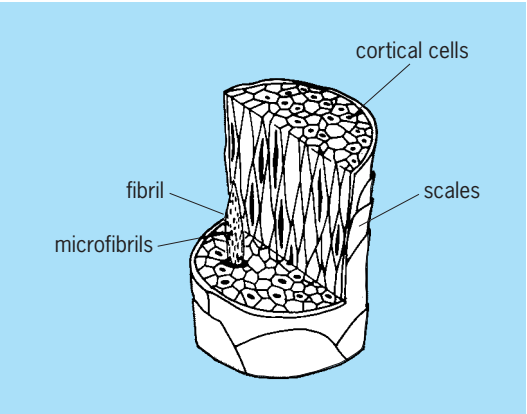

Fig. 4. Schematic of wool. The cuticle surrounds an inner structure, the cortex, made up of cortical cells which are composed of millions of small spindles (fibrils) which are composed of even smaller structures called microfibrils. (*After B. F. Smith and I. Block, Textiles in Perspective, Prentice-Hall, 1982*)

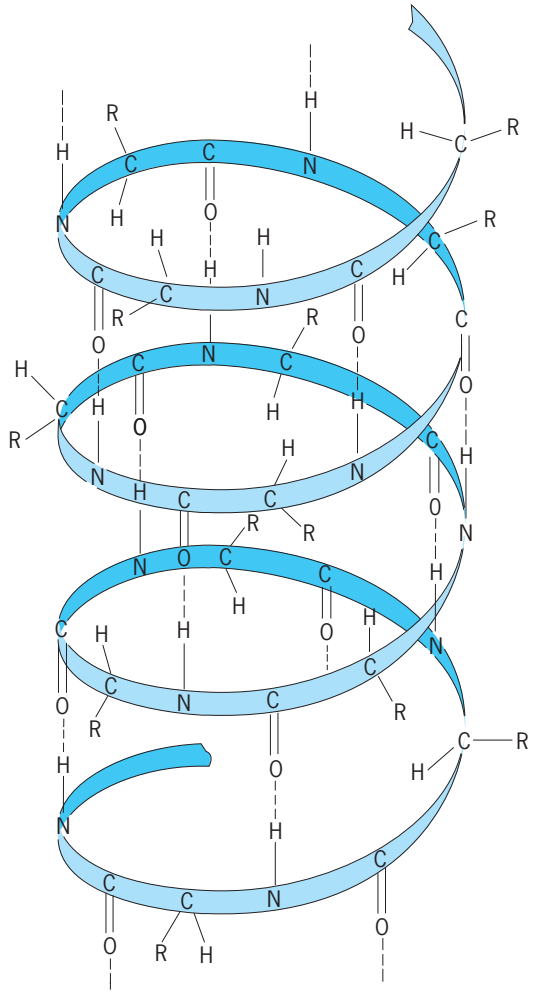

Fig. 5. Helix configuration of the α-keratin molecule. (*After B. F. Smith and I. Block, Textiles in Perspective, Prentice Hall, 1982*)

side groups can carry hydroxyl-, amino-, carboxylic-, and sulfur-containing groups which strongly influence the chemical and physical characteristics of the fiber. The sulfur content, mostly in the form of cystine, is between 3 and 5%. Neighboring polymer chains stabilize each other through hydrogen and salt bonds, and by disulfide cross-links of the cystine. Some of the polymer chains in the cortex fold into a regular helical pattern, and these group into fibrillar units with strong intermolecular bond formation. These fibrillar units, embedded in an amorphous random-coiled matrix, are well aligned with each other and contribute very significantly to the tensile behavior of the fibers. When the fiber is elongated, the α-helical structure reversibly unfolds into the so-called β configuration. The high extensibility and elasticity of the mammalian fiber is due largely to this structural deformability.

Physicomechanical characteristics. Mammalian hairs can reversibly sorb more water than any other commercially used fibers, with the exception of some regenerated cellulose. Part of the comfort factor of wool-based textiles is due to this fact. With the sorption of water the fibers swell and soften.

The dry- and wet-break extensions of mammalian fibers are about 30 and 60%, respectively. The fish-scale-type arrangement of the cuticular cells provides a unique differential frictional effect for fibers; the cut fibers move preferentially in their root direction under random forces, which is the basis of their felting property.

Wool. The most important commercial animal fiber is wool. The order of the largest producers is Australia, Russia, New Zealand, Argentina, South Africa, and the United States.

The finest and most valuable wool is produced by the Australian Merino sheep, developed from a Spanish breed, and is 15–30 μm in diameter and about 1.2–5 in. (3–13 cm) in length. It is used for the finest-quality wool fabrics. Medium- and coarse-variety fibers, 25–40 μm in diameter, are produced by other breeds. The coarser wools are used for heavier fabrics, blankets, upholstery, and carpets. *See* SHEEP.

Fiber properties. Wool fibers range in length from 1 to 20 in. (2 to 50 cm), and may be black, gray, brown, or white. The highest grades are fine (15–20 μm in diameter), long (about 12 in. or 30 cm), and creamy white in color. Wool is a weak fiber, with a tenacity of only 1–2 g per denier but, because of the long length and the high friction due to the scaly outer cuticle, makes strong yarn. The fibers exhibit very high moisture regain (13–16%), with exceptional extensibility (about 35% elongation at break). The cuticle also yields excellent abrasion resistance. Wool is quite resistant to household chemicals and sunlight, but will mildew if allowed to remain damp. It is also eaten by the larvae of various moths and beetles.

Wool's major advantages are high bulk, which permits the manufacture of warm soft fabrics; excellent moisture absorption for comfort; high resistance to wrinkling and good wrinkle recovery; strong yarns; and good abrasion resistance. It is subject to felting shrinkage upon laundering, so that hand washing or dry cleaning are recommended for wool products.

Uses. In apparel, wool is recommended for its warmth, moisture absorption, and durability. For suiting, it provides excellent wrinkle resistance for a neat appearance. In home furnishings, upholstery fabric and carpets of 100% wool are noted for color retention, durability, resilience, and comfort.

Goat hairs. Mohair is produced by the Angora goat mostly in Turkey, South Africa, and the United States. In size it ranges from 10 to 90 μm in diameter for hair and guard fibers, respectively, and is 4–10 in. (10–25 cm) long. Luster is its most important quality. The Cashmere goat in Asia grows fibers in the 15-μm range. Softness is the most important characteristic. As with most other textile fibers, the fine underfur is the valuable component of goat hairs. *See* CASHMERE; MOHAIR.

Camel hairs. The fiber known as camel originates from the Bactrian two-humped species from Asia. For textile purposes only the fine underfur of about 17 μm is collected by plucking or shearing. The guard hairs ranging over 100 μm have only specific uses. The alpaca and vicuña are South

American members of the camel family. The domesticated alpaca produces hair 8–20 in. (20–50 cm) long in 2 years. The diameter of the fur fibers is 25 μm. The wild vicuña grows a shorter fleece, 0.5–3 in. (1.3–7.5 cm), but the finest fibers, about 13 μm. The softest wool-type materials are made from it. *See* ALPACA; VICUNA.

Rabbit hairs. The Angora rabbit hair is very fine, 10–15 μm, and relatively long. It is used mostly for blending with wool for soft fabric effects. The fur hair of common rabbits is still very fine but shorter. The most important use is in felted hats.

Silk. Although it has lost ground to new synthetics, silk is still an important luxury fiber. It is the secretory product of the larval stage of many moth varieties. The cultivated breed is the *Bombyx mori*.

Silk has a few unique properties: it is the finest animal fiber, with a diameter of 10–12 μm; it has no cellular structure since it was never part of body tissue; and it is a continuous filament. Chemically, the silk is a protein consisting of 17 different amino acids. Since these polymer chains are in an extended zigzag form, not in the helical folds characteristic for mammalian hairs, they form extensive crystalline regions embedded in an amorphous matrix.

The fibers are triangular in cross section, with significant unevenness along the length. They lack the fine morphological structurization of mammalian hairs, and resemble more closely an extruded manufactured fiber. Silk binds slightly less water than wool at all humidities, and the effect of water-swelling on its physical characteristics is less. Wetting reduces the fiber strength by only 15%. The fiber strength is high, comparable with that of nylon. While its break elongation is high, about 20%, it can be stretched elastically only to a few percent, unlike mammalian hairs. This is due to the lack of coiled helical units in the ordered regions of the fiber. Silk is more resistant to alkali treatments than animal hair, but less so to acids. Silk yellows when exposed to direct sunlight.

The most important qualities of silk fabrics are softness, smoothness, luster, and drape. *See* SILK; TEXTILE CHEMISTRY; TEXTILE PRINTING. Ira Block

Bibliography. R. S. Asquith (ed.), *The Chemistry of Natural Protein Fibers*, 1977; D. Eby, *Chemistry of Textiles*, 1989; P. W. Harrison (ed.), *Cotton in a Competitive World*, Textile Institute, Manchester, England, 1979; M. L. Joseph, *Introductory Textile Science*, 5th ed., 1993; H. L. Needles, *Textile Fibers, Dyes, Finishes, and Processes: A Concise Guide*, 1986; B. F. Smith and I. Block, *Textiles in Perspective*, 1982.

Natural gas

A combustible gas that occurs beneath the surface of the earth, often found in conjunction with petroleum deposits. Its main use is for fuel, but it is also used to make carbon black, certain chemicals, and liquefied petroleum gas. *See* LIQUEFIED PETROLEUM GAS (LPG).

Natural gas consists predominantly of methane, but also contains a mixture of hydrocarbons such as ethane, propane, and pentanes. Carbon dioxide, nitrogen, helium, and hydrogen sulfide may also be present. The types of natural gas vary according to composition and can be dry or lean gas (mostly methane), wet gas (considerable amounts of other hydrocarbons), sour gas (much hydrogen sulfide), sweet gas (little hydrogen sulfide), residue gas (higher hydrocarbons having been extracted), and casinghead gas (derived from an oil well by extraction at the surface). *See* METHANE.

Distribution and reserves. Natural gas occurs on every continent. Wherever oil has been found, a certain amount of natural gas is also present. In addition, natural gas is associated with coal. Natural-gas hydrates are solids composed of water molecules forming a rigid lattice of cages (clathrates), with most of the cages containing a molecule of natural gas. Methane hydrate is the dominant natural-gas hydrate. Methane hydrates are found in polar and other continental shelves, and in deep-ocean outer continental margins in all oceans, including polar oceans. *See* CLATHRATE COMPOUNDS; COALBED METHANE; HYDRATE; PETROLEUM GEOLOGY.

Production and delivery. Natural gas comes from reservoirs, usually in sedimentary rock lying below the earth surface at depths varying from a few hundred feet to several miles. Being in gaseous form, it may occur alone in separate reservoirs. More commonly, it forms a gas cap, or mass of gas, entrapped between liquid petroleum and an impervious capping rock layer in a petroleum reservoir. Under conditions of greater pressure, it is mixed with or dissolved in crude oil. Successful exploitation involves drilling, producing, gathering, processing, transporting, and metering the use of the gas. Before it is commercially distributed, natural gas usually is processed to remove propane, butane, and nonhydrocarbon gases such as hydrogen sulfide. Processed natural gas has no distinct odor, so an odorant is added since an undetected leak could result in an explosion or asphyxiation. Transmission pipelines move the gas from processing centers to the market, where distribution pipelines carry it to consumers. Natural gas is also distributed worldwide in liquid form (liquefied natural gas, LNG), which is produced by chilling the gas to below its boiling point. *See* LIQUEFIED NATURAL GAS (LNG); OIL AND GAS FIELD EXPLOITATION; OIL AND GAS WELL DRILLING; PETROLEUM RESERVOIR ENGINEERING.

Storage. Because of the wide seasonal variations in temperature and the relative constancy of gas production and supply, excess gas must be stored in the summer to meet the increased demand in the winter. Without a reliable means of storing this excess gas, demand would be difficult to meet throughout the year. Depleted gas reservoirs are usually the first choice for use as storage reservoirs. Other types of reservoirs are used as well, including depleted oil reservoirs, aquifers, and salt caverns. *See* OIL AND GAS STORAGE. Michael A. Adewumi; Michel T. Halbouty

Bibliography. K. Arnold and M. Stewart, *Surface Production Operations*, vol. 2: *Design of*

Gas-Handling Systems and Facilities, 2d ed., 1999; B. Guo and A. Ghalambor, *Natural Gas Engineering Handbook*, 2005; N. J. Hyne, *Nontechnical Guide to Petroleum Geology, Exploration, Drilling and Production*, 2001; W. C. Lyons and G. J. Plisga, *Standard Handbook of Petroleum and Natural Gas Engineering*, 2d ed., 2004; M. D. Max (ed.), *Natural Gas Hydrate in Oceanic and Permafrost Environments*, 2003.

Natural language processing

Computer analysis and generation of natural language text. The goal is to enable natural languages, such as English, French, or Japanese, to serve either as the medium through which users interact with computer systems, such as database management systems and expert systems (natural language interaction), or as the object that a system processes into some more useful form, such as in automatic text translation or text summarization (natural language text processing). *See* DATABASE MANAGEMENT SYSTEM.

In the computer analysis of natural language, the initial task is to translate from a natural language utterance, usually in context, into a formal specification that the system can process further. Further processing depends on the particular application. In natural language interaction, it may involve reasoning, factual data retrieval, and generation of an appropriate tabular, graphic, or natural language response. In text processing, analysis may be followed by generation of an appropriate translation or a summary of the original text, or the formal specification may be stored as the basis for more accurate document retrieval later. Given its wide scope, natural language processing requires techniques for dealing with many aspects of language, in particular, syntax, semantics, discourse contest, and pragmatics. (Analysis and generation of spoken natural language, which are not discussed in this article, also involve techniques for dealing with acoustic phonetics, phonology, stress, and intonation.) *See* VOICE RESPONSE.

Parsing. The aspect of natural language processing that has perhaps received the most attention is syntactic processing, or parsing. Most current techniques for parsing an input string of words involve (1) a description of the allowable sentences of the language (the grammar); (2) an inventory of the words of the language with their inflectional, syntactic, and possibly semantic properties (the lexicon); and (3) a processor which operates on the grammar, the lexicon, and the input string (the parser). This processor (1) simply accepts the input string, if grammatically well formed, or rejects it (a recognizer);\break (2) associates the string, if well formed, with its structure (or structures, if ambiguous) according to the grammar (an analyzer); or (3) associates the string with some other representation, for example, a semantic characterization (a transducer). Syntactic processing is important because certain aspects of meaning can be determined only from the underlying structure and not simply from the linear string of words.

One of the oldest parsing techniques, called augmented transition networks (ATNs), grew out of a system for parsing context-free (CF) languages, called recursive transition networks (RTNs). A recursive transition network parser consists of a set of named graphs or networks, each consisting of a set of nodes or states connected by a possibly ordered set of directed labeled arcs. The labels correspond to (1) words or classes of words that can be recognized or "consumed" on the arc; (2) the empty symbol, indicating an arc that can be taken without consuming any input; or (3) the name of a network, which indicates that the next segment of the input string must be recognizable by that network. Each network has a start state and one or more end states. If the parser can move through a network from start to end by consuming a segment of the input string, that segment is said to be recognizable by that network. *See* GRAPH THEORY.

An augmented transition network adds to the basic recursive transition network framework the ability to set and test variables or registers, thereby giving it the power to recognize a wider class of languages than a recursive transition network. An augmented transition network is appearing because its grammar is relatively easy to specify. Its weaknesses lie in the simple, uniform control structure provided by the basic augmented transition network (that is, unguided backtracking) and in its power, felt to be more than is needed for recognizing a natural language.

Current trends are to construct parsers and grammars which appear to follow more closely human parsing strategies and which have less power. In particular, researchers have begun to give almost context-free descriptions of natural languages, thereby allowing them to use slightly extended versions of efficient context-free parsing techniques. Such descriptions include generalized phrase structure grammar (GPSG), immediate dominance/linear precedence (ID/LP) grammar, and tree adjoining grammar (TAG).

Semantic analysis. A second phase of natural language processing, semantic analysis, involves extracting context-independent aspects of a sentence's meaning. These include the semantic roles played by the various entities mentioned in the sentence. For example, in the sentence "John unlocked the toolbox," "John" serves as the agent of the unlocking, and "the toolbox" serves as the object; in "This key will unlock the toolbox," "this key" serves as the instrument. Context-independent aspects of sentence meaning also include quantificational information such as cardinality, iteration, and dependency. For example, in the sentence "In every car, the mechanic checked to see that the engine was working," the checking is iterated over each car; the identity of the engine depends on the identity of the car, while that of the mechanic does not; and the cardinality of engines per car is one. Thus there are as many engines as cars, but possibly only one mechanic. The

representational formalism used by the system for semantic analysis (for example, first-order predicate calculus, case grammar, conceptual frames, procedures, and so forth) is usually chosen for its ability to convey those aspects of semantics that the system requires for later processing. For example, if temporal position (past/present/future) is not significant, it will not be captured in the formalism.

Most semantic analysis is done by applying pattern-action rules either during parsing or afterward. The pattern part of a rule consists of clauses, each of which specifies the presence of a particular lexical item, usually the head of some syntactic substructure (for example, the main verb of a clause or sentence or the head noun of a noun phrase); or a particular syntactic substructure (for example, a relative clause, to be interpreted as a restriction on the class described by the rest of the noun phrase). A pattern clause may also specify a test on another part of the current substructure. The action part of a rule usually calls for building a piece of semantic representation, often requiring the semantic analysis of some other part of the syntactic substructure. For example, there may be a pattern-action rule associated with "unlock" as the main verb of a clause. A test in the pattern may require that the subject of the clause be interpretable as an animate agent. The rule's action may call for the inclusion of a conceptual frame for the concept "unlock" as part of the semantic representation of the sentence. The rule's action may further specify that the agent role of the frame be filled by the semantic interpretation of the subject of the clause, and that the object role be filled by the semantic interpretation of the direct object. In some systems, a rule can have optional pattern clauses and actions: thus, the rule pattern for "unlock" might optionally specify a "with" prepositional phrase whose noun phrase object can be interpreted as a tool. If so, the rule action might additionally call for inserting the instrument role in the frame with the semantic interpretation of the prepositional phrase object.

Contextual analysis. Given that most natural languages allow people to take advantage of discourse context, their mutual beliefs about the world, and their shared spatio-temporal context to leave things unsaid or say them with minimal effort, the purpose of a third phase of natural language processing, contextual analysis, is to elaborate the semantic representation of what has been made explicit in the utterance with what is implicit from context. Two major linguistic devices that contextual analysis must deal with are ellipsis and anaphora.

Ellipsis. Ellipsis involves leaving something unsaid. To handle ellipses computationally, techniques are required for recognizing that something is indeed missing and for recovering the ellipsed material. When the utterance is a sentence fragment and not a complete sentence, it is fairly easy to recognize that something is missing. An example appears in the following sequence:

User: What is the length of the JFK?
System: ⟨some number of feet⟩.
User: The draft?

On the other hand, since parsers are usually designed for well-formed input, either the system's grammar must be revised to accept sentence fragments or a special error-recovery routine must take over after the parser fails. When an utterance is syntactically well formed, it may still be elliptic in that some needed conceptual material is missing, as in the following example:

User: What maintenances were performed on plane 3 in May 1971?
System: ⟨list of maintenances⟩.
User: What maintenances were performed on plane 48?

In the user's second question, the time period of interest is missing, and the question should not be answered until it is recovered. (It is clearly May 1971.)

The primary technique for recovering ellipsed material is a simple one, based on semantic features. For sentence fragment ellipses, the previous discourse is searched for the most recent utterance containing a constitutent with the same features as the fragment. The utterance minus that constituent is taken as the ellipsed material. For conceptual ellipses in a syntactically well-formed sentence, the previous discourse is searched for the most recent utterance with a constituent having the required semantic features. That constituent is taken to be the ellipsed material. In each case, a new well-formed sentence is then constructed and processed as if the ellipsis had never occurred. For instance, in the first example, both "length" and "draft" are properties of ships. Thus, given the fragmentary utterance "The draft?," What is ___ of the JFK?" is found as the ellipsed material. The question "What is the draft of the JFK?" is then interpreted and answered normally. This technique works often, but does not constitute a general solution. A more powerful solution has been developed based on recognizing a user's goals in producing an utterance, but has been found computationally efficient only in very narrow task-oriented domains. This will be discussed below.

Anaphora. Anaphoric expressions are very simple words or phrases which cospecify something previously evoked by the discourse or are strongly associated with something so evoked. Instances of anaphora include definite pronouns such as "he," "she," "it," "they," and definite noun phrases such as "the mechanic" and "the cars." The problem is that anaphora can be interpreted only in context, and the semantic interpretation of a sentence is not complete until all anaphoric expressions are resolved and the cospecified entities identified.

Early computational approaches to anaphora resembled those for dealing with ellipses: entities described in previous sentences were searched for the most recently mentioned one with appropriate semantic features. Now recency has been replaced by the notion of focus as a basis for anaphora resolution. Immediate focus reflects the particular thing the speaker is talking about; global focus involves things associated with it or in which it participates, and gives a sense of what may be talked about next.

Techniques have been developed for tracking immediate focus, projecting ahead from the current utterance what may be focused on in the next one. This is useful for resolving anaphora, in that it predicts what entities are likely to be respecified anaphorically.

Pragmatics. A fourth phase of natural language processing, pragmatics, takes into account the speaker's goal in uttering a particular thought in a particular way—what the utterance is being used to do. In an interaction, this will influence what constitutes an appropriate response. For example, an utterance which has the form of a yes/no question or an assertion may have the goal of eliciting information (for example: "Do you know how to delete a control-Z?," "I can't get the set file protection command to work."). Because it is inappropriate (and possibly at times dangerous) to take a user's utterances literally or to assume that the user will take those of the system literally, computational techniques must be devised for relating the syntactic shape and semantic content of an utterance to its pragmatic function.

Plan recognition. One important approach to this problem has been to view language understanding as plan recognition. The actions (either communicative or physical) that constitute the plan may be motivated in one of two ways: goals in the world that the person wants to accomplish, for which he or she needs to elicit or offer aid or information; or aspects of an already ongoing interaction that need attention—for example, confusion over the speaker's foregoing utterance may lead the listener to seek clarification.

A user seeking particular information from the system illustrates clearly a plan-recognition approach to language understanding. The user's utterance—a well-formed sentence or an ellipsed fragment—is taken as a request for information that the user believes he or she needs in order to accomplish some goal, a goal which is not as yet presumed to be known to the system. Just as a medical diagnosis system uses rules which link findings back to those diseases which commonly manifest them, the plan-recognition system uses rules which link utterances back to those domain goals which need the intended information in order to be achieved. For example, consider the utterance "The train to Windsor?" made to a system serving as train information clerk. The system interprets this as a request: the user wants to know some property of that train in order to fulfill his or her goal. The system then tries to figure out what that goal is, in order to figure out what information the user might be requesting. There are only two possible goals considered: meeting a train and boarding one. The description "train to X" does not match that of incoming train, so the user's goal is taken to be boarding the Windsor train. To board a train, one needs to know its departure time and track. Since the system does not have evidence of the user's knowing either of these properties, it responds with both: "It leaves at 3:15, from track 7." Currently, it is only by limiting the domain, and hence the range of possible goals the system needs to consider, that such a plan-based approach to pragmatics and natural language processing becomes feasible.

Cooperative principle. Because pragmatics acknowledges language use, it also acknowledges expectations that speakers and listeners have about the form and content of utterances, based on normal conventions of use. This has been well described by the philosopher Paul Grice, who noted that speakers acknowledge a "cooperative principle" of conversation (by either upholding it or purposely flouting it), which he further specified in terms of conversational maxims of quantity, quality, manner, and relation. For example, the quantity maxim states: "Make your contribution as informative as is required (for the current purposes of the exchange). Do not make it more informative than is required."

The cooperative principle and its maxims are important to natural language processing because, if a system does not behave in accord with normal conventions of use, the user is likely to be confused or misled by the system's behavior. Conversely, if the system does not interpret the user's behavior in terms of normal conventions of use, the system is unlikely to understand the user correctly, if at all. In particular, the cooperative principle and its maxims reveal a method of implicit communication which Grice termed implicature. An implicature is basically an aspect of an utterance's interpretation which makes no contribution to its truth value (that is, semantics) but constrains its appropriateness in discourse. For example, consider the following discourse:

Q: Is there a gas station on the next block?
R: Yes.

The simple "yes" answer implicates to Q that, as far as R knows, the gas station is able to provide its normal services and hence fulfill Q's probable goal. Q reasons that if R knew that the gas station was closed and hence could not fill Q's needs, R would have said so: that is, R would have said "Yes, but it's closed." Thus a system must be as aware of implicatures (both the user's and its own) as it is aware of what is communicated explicitly.

Overall organization. As for fitting the pieces together, there is no single way that natural language analysis is done. Some systems have a single processor for syntactic, semantic, contextual, and pragmatic analysis, with no distinction made as to the source of that knowledge. Some systems keep the knowledge sources separate but apply them simultaneously, extracting whatever can be derived at the moment and using whatever information is available. Other systems are very modular, separating the knowledge sources and specifying when they should be applied. Efficiency, extensibility, and transportability are some of the important issues to be considered when evaluating a system for natural language analysis.

Natural language generation. The bulk of early research in natural language processing was formerly directed at natural language analysis. Now researchers have taken up seriously the task of natural language generation. Generation is not just the reverse of analysis because the status of user

and system are fundamentally different. Systems can be developed which tolerate users' mechanical errors (for example, spelling, typing, and grammatical mistakes), treating them as insignificant variations. Users, on the other hand, may not be able to figure out which aspects of the system's natural language behavior reflect simple nonfluencies (for example, those due to limited lexical or grammatical options) and which embody significant aspects of communication. Moreover, the system's sense of language must be more highly developed for generation, lest it confuse or mislead the user by what it communicates or how. Work on explanation is also a significant aspect of natural language generation. *See* ARTIFICIAL INTELLIGENCE. Bonnie Webber

Bibliography. M. A. Covington, *Natural Language Processing for Prolog Programmers*, 1993; F. Pereira and B. J. Grosz (eds.), *Natural Language Processing*, 1994; *Proceedings of the Association for Computational Linguistics*; *Proceedings of the International Conference on Computational Linguistics*; H. Tennant, *Natural Language Processing*, 1981; T. Winograd, *Language as a Cognitive Process*, 1982.

Nautiloidea

A group of externally shelled cephalopods, represented by the single extant genus *Nautilus*. The formal designation of this group as a subclass is now generally used only for those externally shelled cephalopods that resemble *Nautilus* in having completely coiled shells (and then the subclass includes *Nautilus* itself). In living forms, the basic structural plan includes a shell consisting of a septate phragmocone, a living chamber, and a siphuncle. In fossil nautiloids, this simple pattern is modified in great variety with respect to shell form and size, structure and size of the siphuncle, and the large number of devices to counteract the buoyancy of the phragmocone. The shape of fossil nautiloids may deviate in many ways from the simple *Nautilus* model (see **illus**.); the length of straight or slightly curved shells varies from less than 1 in. (2.5 cm) to more than 30 ft (10 m). Few coiled shells are larger than *Nautilus*. Also, the siphuncle may vary considerably in size and shape (see illus.). The aperture of the living chamber may be constricted or contracted into various shapes, and the interiors of siphuncle and camerae may be partially filled by layers of aragonite and conchiolin.

The nautiloids are represented by approximately 1000 fossil genera comprising an estimated 5000 species and grouped into more than 130 families, half of which lived in the Ordovician Period. The nautiloids are distributed among 12–14 orders, with the Oncocerida (183 genera), Nautilida (179), and Orthocerida (135) being the most numerous.

Nautiloids first appeared about 520 million years ago, very late in the Cambrian Period. Most early nautiloids had small shells, about 1 in. (2.5 cm) in length, although some reached 4 in. (10 cm) and were either straight or slightly curved. They reached the acme of their diversity in the Ordovician, with nearly 70 families, and then declined to 13 families in

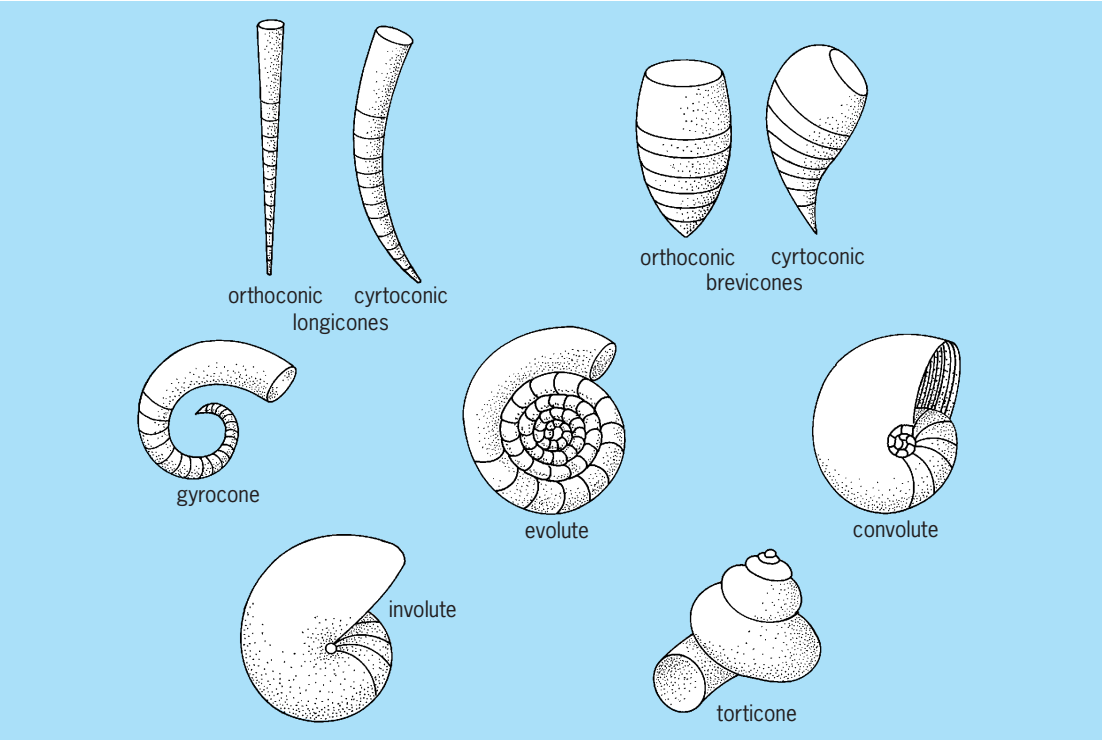

Common types of nautiloid conchs. (*After C. Teichert et al., Cephalopoda: General features, in R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. K, Mollusca 3, Geological Society of America and University of Kansas Press, 1964*)

the Permian, 4 in the Tertiary, and 1 at present. The evolution of the nautiloids suffered severe setbacks at the end of the Cambrian and the end of the Triassic, when they came close to extinction.

Fossil nautiloids are found on all continents, including Antarctica. Especially noteworthy are the rich Ordovician and Silurian faunas of North America, northern Europe, Czechoslovakia, and central and southern China.

Coiled nautiloids almost certainly moved around by jet propulsion like *Nautilus* and lived close to the sea floor, at moderate to intermediate depths in many different environments. Other types may have included agile swimmers as well as slow-moving benthic adaptations. *See* CEPHALOPODA; NAUTILUS.

Curt Teichert

Bibliography. M. R. Clarke and E. R. Trueman (eds.), *The Mollusca*, vol. 12: *Paleontology and Neontology of Cephalopods*, 1988; R. E. Crick, Buoyancy regulation and macroevolution in nautiloid cephalopods, *Senckenbergiana Lethaea*, 69:13–42, 1988; C. H. Holland, The nautiloid cephalopods: A strange success, *J. Geol. Soc.* (London), 144:1–15, 1987; R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, pt. K, *Mollusca* 3, 1964.

Nautilus

The only living genus of the molluscan subclass Nautiloidea, containing the only living cephalopods with an external chambered shell and numerous cephalic tentacles (see **illus.**). Six species of *Nautilus* live in moderately deep water in the western Pacific and around the East Indies.

The perfectly symmetrical planospiral shell is divided into about 25 chambers by transverse partitions. Only the outermost and largest chamber is occupied by the animal, the rest being gas-filled and serving as a buoyancy organ. Each partition, or septum, has a perforation reinforced with a shelly "neck" through which passes the siphuncle, a cord of living tissue containing blood vessels and surrounded by cornified layers, which can adjust the buoyancy by absorption of water from the penultimate chambers balanced by gas secretion. Healthy specimens of *Nautilus* can move to and from the surface water from depths of 1150 ft (350 m) by such adjustment alone. The mechanical strength of the chambered shell is necessarily high, and a gas-filled *Nautilus* shell of average size will successfully resist crushing implosion down to depths of about 1800 ft (550 m), that is, to external pressures of approximately 55 atm or 5.5 megapascals.

Apart from its buoyancy control, which allows it to live near the surface at night and on the bottom or at greater depths during the day, *Nautilus* is capable of horizontal movements by jet propulsion. As in squid and cuttlefish, this is achieved by water propelled by muscular contractions of the mantle cavity. However, *Nautilus* is a relatively slow-moving collector of carrion and moribund animals, whereas squid are among the fastest swimmers in the sea and are

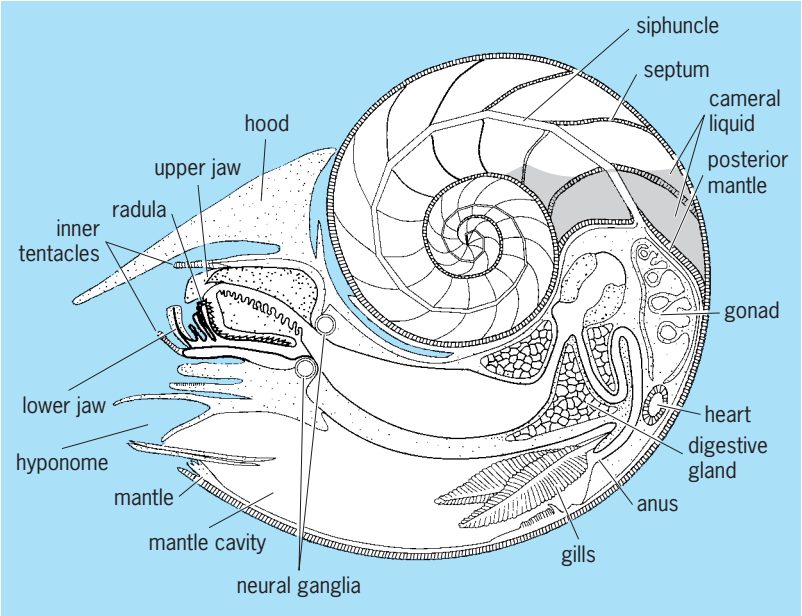

Internal structure of *Nautilus macromphalus*. (*After P. D. Ward, The Natural History of Nautilus, Allen and Unwin, 1987*)

predacious carnivores capturing active prey.

The mantle cavity in *Nautilus* is unique among living mollusks in that the normally paired pallial and cardiac structures are doubled. There are four ctenidia (gills), four osphradia, and four excretory organs. The ventricle of the central heart receives blood from four auricles, and (unlike other cephalopods such as squid and octopus) there are no accessory pumps (or branchial hearts) in the gill circulation. The head of *Nautilus* bears rather primitive, lenseless, "pinhole" eyes and 38 retractile, prehensile tentacles, which lack the characteristic suckers found on the 8 or 10 arms of other modern cephalopods. Compared to octopus, squid, and the rest, *Nautilus* obviously represents the persistent line of a more primitive cephalopod stock. In fact, fossil Nautiloidea date from the Upper Cambrian, while the stocks of modern cephalopods probably arose in the Cretaceous. *See* CEPHALOPODA; NAUTILOIDEA; OCTOPUS.

W. D. Russell-Hunter

Naval architecture

An engineering discipline concerned with the design of ships, boats, drill rigs, submarines, and other floating or submerged craft. The naval architect creates the initial overall concept for a new ship, integrates the work of other specialists as the ship design is developed, and is specifically responsible for the hull and superstructure shape, general arrangements, structural design, weights and centers calculations, stability analysis, hydrodynamic performance assessments, propeller and rudder design, and the arrangement and outfit of all living and working spaces, other than machinery spaces. The naval architect's ally, the marine engineer, is responsible for the design of the propulsion plant, the electric

plant, and other ship machinery and mechanical systems, including the so-called distributive systems: electric cabling, piping, and ventilation system ducting. The marine engineer also is responsible for ship control systems, including propulsion and electric plant controls and the steering system. *See* FERRY; HYDROFOIL CRAFT; ICEBREAKER; MERCHANT SHIP; NAVAL SURFACE SHIP; OIL AND GAS, OFFSHORE; SUBMARINE.

In the past, naval architecture was as much an art as a science, but research, coupled with advances in computer-aided design, has greatly enhanced the scientific basis of the profession. Naval architecture is a specialized form of mechanical engineering, as is marine engineering. Thus the education of naval architects is very similar in content to that of mechanical engineers, and the same types of degree programs are offered. Some colleges and universities combine naval architecture and marine engineering education and offer a combined degree. *See* MARINE ENGINEERING; MECHANICAL ENGINEERING.

Because naval architects have an overall ship design integration responsibility, they must be familiar with all the other engineering disciplines which contribute to ship design. Their outlook as ship design integrators must embody the principles of systems engineering: Overall ship cost and performance must be optimized, and this will generally require compromises in the designs of the individual subsystems or elements of the total ship. In addition, the naval architect must be knowledgeable of and infuse into the design the principles of human engineering and the "ilities": producibility, operability, supportability, reliability, and maintainability. Because of the importance of cost, the life-cycle costs (design, build, operation, and support costs) of the many alternatives considered during design must be carefully analyzed during the design process. This requires the naval architect to have a sound understanding of ship economics. National and international agencies establish regulations to which ship designs must adhere; the naval architect must know these requirements. *See* HUMAN-FACTORS ENGINEERING; RELIABILITY, AVAILABILITY, AND MAINTAINABILITY; SYSTEMS ENGINEERING.

The type of work done by naval architects varies widely. Some work alone or in a small team, developing the designs of yachts or other small craft in their entirety. At the other extreme, some work as members of very large teams designing complex warships. In this situation, the naval architect may focus on overall ship integration, coordinating the work of others, or may work in a specific functional area, such as structural design, stability, or general arrangements. The nature of the work done also varies with the design phase. In the very early stages of design, overall ship concepts are created and analyzed in their entirety using simplified sketches or computer models and engineering approximations. In the late stages of design, the focus is on the development of design details. Large amounts of data must be dealt with accurately, and detailed procurement and test specifications are prepared. Many naval architects do not work as ship designers. Examples include researchers in a specialty area, teachers, the representatives of a ship owner who is having ships designed and built by others, ship production planners or inspectors, and employees of a government agency establishing safety regulations or managing the acquisition of ships for that agency, such as the U.S. Navy, Coast Guard, or National Oceanographic and Atmospheric Administration (NOAA). *See* SHIP DESIGN; SHIPBUILDING.

Peter A. Gale

Bibliography. E. V. Lewis (ed.), *Principles of Naval Architecture*, 2d ed., 3 vols., Society of Naval Architects and Marine Engineers (SNAME), 1988; SNAME, *Careers in the Maritime Industry* (brochure), 1993; R. Taggart (ed.), *Ship Design and Construction*, 3d ed., SNAME, 1980.

Naval armament

A general term that covers the ordnance and control systems used by naval ships and aircraft. It includes a wide spectrum of weapons designed for use against targets in the air, on land or sea, or under the ocean surface.

Naval armament has always presented a formidable challenge to designers. Weapons must be launched from a rolling, pitching platform against targets that are frequently moving and often invisible. Accuracy and reliability are essential, because only a limited number of weapons can be carried on board ship. Naval weapon systems must be capable of a wide range of destructive force to meet the variable demands of "cold," limited, or all-out war. The spectrum of weaponry used by naval forces thus runs from small arms to nuclear warheads and includes weapons that are intended for use against a particular type of target as well as general-purpose weapons.

The combinations of weapons and control systems built into a warship are based on the requirements of the ship's intended missions and on the ship's ordnance-carrying capacity (**Fig. 1**). Any ship's weapon suite is a compromise between offensive and defensive armament requirements and such other characteristics as the vessel's endurance, speed, survivability, and habitability. *See* NAVAL SURFACE SHIP.

Naval weaponry is becoming more internationalized. Many developments abroad parallel those in the United States; others differ to suit the needs of the country involved. Many countries, including the United States, fill some of their needs by development in cooperation with other nations, or by foreign purchase. This article discusses United States naval armament.

Weapon systems. Modern naval weapons are not considered in isolation but in terms of weapon systems. A typical weapon system includes (1) devices capable of searching out, locating, and identifying the target such as radar and sonar; (2) launchers and weapon-handling equipment, such as gun mounts, missile launchers, and torpedo tubes; (3) control

Fig. 1. Modern naval ship employing a spectrum of weapons, sensors, and control systems. (*U.S. Navy*)

equipment, such as gun and missile director systems, to calculate the correct path for the weapon to intercept the target, and to transmit control information to the launcher or delivery vehicle; (4) a delivery vehicle, often self-propelled, such as a missile airframe or torpedo body, to carry the payload to the target; a gun's projectile does not require a vehicle, though a rocket-assisted projectile includes a rocket motor; (5) the weapon itself, containing the destructive force, such as a gun projectile or a missile or torpedo warhead. *See* RADAR; SONAR.

Nuclear weapons do not necessarily constitute separate systems. A nuclear capability can be designed into a weapon system as an alternative choice of payload, allowing either high-explosive or nuclear application of the same guidance and delivery systems. *See* ATOMIC BOMB; HYDROGEN BOMB.

Naval armament may be air-, surface-, or submarine-launched. It can be categorized as tactical or strategic, or by its intended primary target: surface attack, air defense, or antisubmarine. Many weapons can be used against different types of targets. Naval weaponry includes guns, guided missiles, rockets, bombs, depth charges, torpedoes, and mines.

Guided missiles. In the years since World War II, guided missiles have taken first place among families of naval weapons. Naval missiles may be adaptable to multiple launch modes: from ship, submarine, and aircraft. Modern missiles are more compact, saving critical space and weight, and their guidance systems have steadily become more sophisticated. Early shipboard launchers, mounted abovedecks and fed from magazines, could handle only one specific missile; newer ones can handle two or three different weapons, eliminating the need for separate launchers. This, however, complicates ammunition stowage and loading, so that a particular launcher can engage only one type of threat at a time.

A form of vertical launch system (VLS) was first taken to sea in 1980 by the Soviet large missile cruiser *Kirov*. The American Mark 41 VLS entered the fleet in 1986 (**Fig. 2**). It is an advanced system in which missiles, in sealed shipping canisters, are stowed in launching tubes in the ship's hull. Weapons are protected from the weather and ready for immediate use, doing away with the need to load missiles in succession from magazines into an abovedecks launcher. Ships can be loaded with variable mixes of weapons to suit their mission. The vertical launch system arms newer *Ticonderoga*-class missile cruisers as well as *Arleigh Burke*-class destroyers, and has been backfitted into earlier *Spruance*-class destroyers. Some attack submarines have individual vertical launch tubes.

Advanced radars and information-handling technology have improved missile capabilities. Short-range, point-defense systems have been developed for ship defense against aircraft, missiles, and fast attack craft. Naval ordnance, because of the varied nature of the threats to be met, will continue to be a mix of missiles, torpedoes, guns, and antisubmarine weapons.

Guidance systems. Missiles are guided in a number of ways. In a preset system, the missile follows a flight plan programmed before launch; once the missile is launched, this cannot be changed. A command-guided missile follows course instructions transmitted by a control station. Homing systems can be active, emitting a radar or laser signal which is

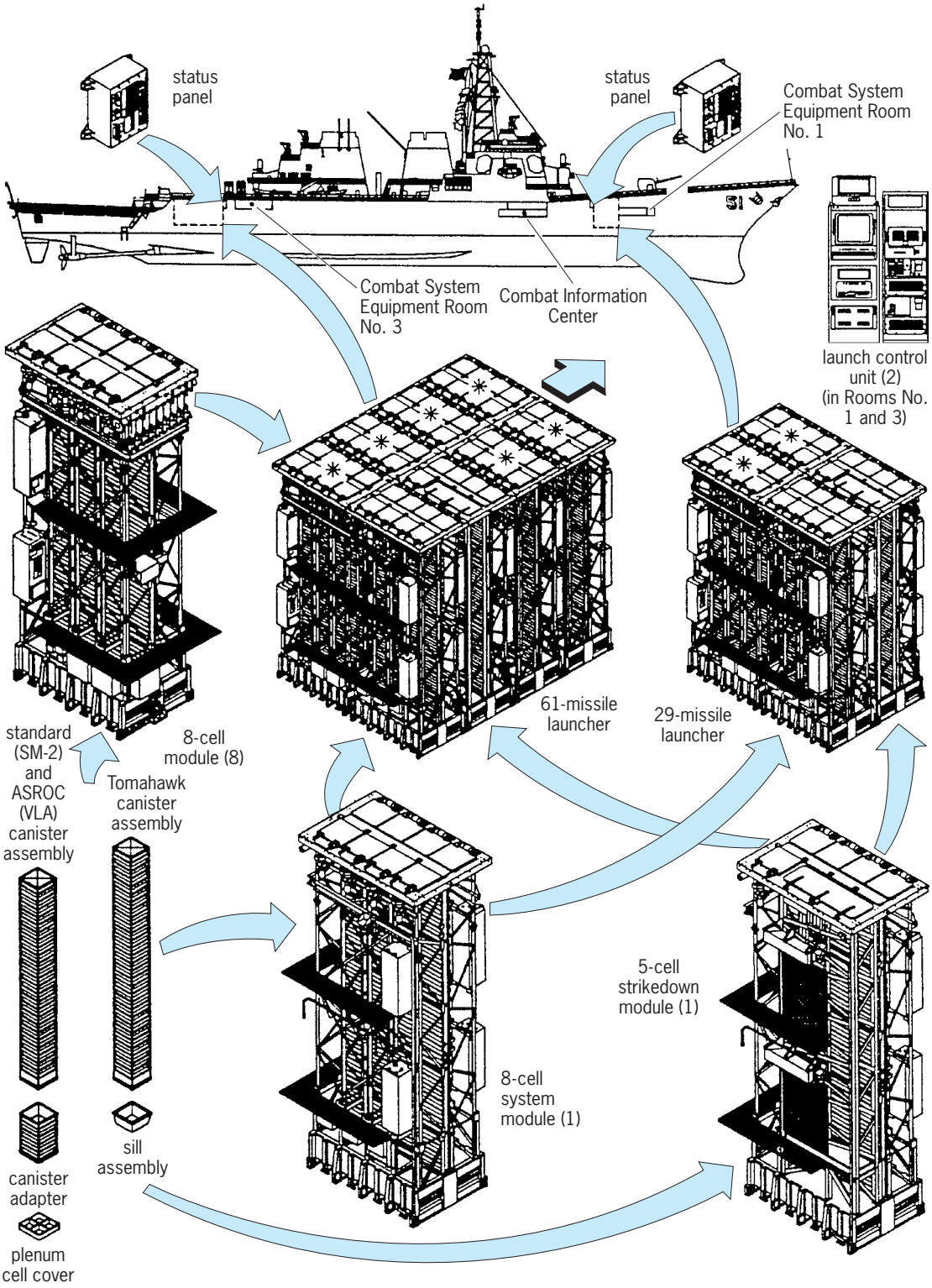

Fig. 2. Elements of the Mark 41 vertical launch system (VLS). This modular system is designed to be installed in ships of various sizes and configurations. (*U.S. Navy*)

reflected by the target and picked up by the missile's receiver. In a semiactive homing system, the missile receives target reflections of emissions from the launching ship or from another source. A passive homing system recognizes a physical characteristic of the target—heat or electronic emissions—and steers toward them. A composite guidance system combines two or more of these types of guidance. *See* HOMING.

A missile's flight is divided into a boost phase, in which it is launched and attains flight altitude and direction; a midcourse phase, that is, the journey to the target area; and a terminal phase, the final run-in to the target. A missile may use, for example,

inertial guidance during its midcourse phase, shifting to homing guidance to attack its target. Short-range missiles as well as some torpedoes may be wire-guided, their command signals transmitted along a fine wire that unreels as the missile travels toward its object. *See* GUIDANCE SYSTEMS; INERTIAL GUIDANCE SYSTEM.

Warheads. Missile warheads may be designed for blast effect against underwater or ground targets, or for blast with fragmentation against surface or aerial targets; or they may contain nuclear explosives. Continuous-rod warheads, used in older missiles against aircraft, contain two sets of jointed rods which expand into semicircles, when the bursting charge is set off by a proximity fuze, and destroy or damage their target by cutting. Fragmentation warheads explode into an expanding sphere of casing fragments; controlled-fragmentation warheads are configured to spread their fragments in a pattern designed for maximum damage to the target.

Strategic missiles. An *Ohio*-class missile submarine carries 24 Trident fleet ballistic missiles (**Fig. 3**). Developed to replace the earlier Polaris and Poseidon, the Trident I C-4 missile arms the first eight *Ohio*'s, while the remaining ten ships use the more accurate and longer-range Trident II D-5. Four of the earlier submarines will later be backfitted to carry Trident IE. Four British submarines are also armed with it. *See* SUBMARINE.

Surface-to-air missiles. Standard, the Navy's principal air defense missile, replaced the first-generation Tartar, Terrier, and Talos. A supersonic solid-fuel weapon, it is produced in medium-range (MR) and extended-range (ER) versions. The current Standard Missile 2 (SM-2) is designed for use with the Aegis system to counter large-scale missile attacks. It has been improved to enhance the missile capabilities of ships without Aegis. Newer versions of Standard, for use in VLS, are designed to cope with very fast targets at great altitudes; the most recent missile is intended for dual use against aircraft and ballistic missiles at lower-tier altitudes.

Fig. 3. Trident missile being fired from a submerged submarine. (*U.S. Navy*)

Sea Sparrow is an antiaircraft adaptation of the airborne Sparrow III missile, developed as a relatively uncomplicated basic point-defense missile system (BPDMS) to protect ships without Standard missiles. It uses a radar director to acquire and follow its target. The missile uses semiactive homing, guiding itself on reflections from the fire control radar, and a proximity fuze. In its present version, Sea Sparrow is the principal North Atlantic Treaty Organization (NATO) point-defense missile, using a lightweight launcher with a digital fire control system. The internationally developed Evolved Sea Sparrow (ESSM), with greater speed, range, and maneuverability for use with VLS, is in production. It will be used in combination with the Rolling-Airframe Missile (RAM), defending against threats beyond RAM's range.

RAM, developed jointly by the United States, Denmark, and Germany, is stabilized by tail fins and helical ridges, or lands, around the body of the missile which cause it to rotate slowly in flight like a rifle bullet. A "fire-and-forget" counter to antiship missiles, RAM uses Sidewinder's rocket motor and warhead. It homes on the target's radiation until it picks up a heat signature with its infrared seeker. SEA RAM, under development, is a Phalanx radar-controlled antiaircraft mount with a RAM launcher in place of a 20-mm automatic gun, designed for close-in defense against future high-performance antiship missiles.

Stinger, a portable Army–Marine Corps antiaircraft missile, has been issued to ships when suicide air attack was considered possible. A shoulder-launched heat-seeking weapon, it has an explosive warhead with a proximity fuze.

Surface-to-surface missiles. Tomahawk, a long-ranged land attack cruise missile, was used in the Gulf War and in Kosovo. Capable of attacking targets at a range up to 1000 mi (1600 km), Tomahawk has greatly increased the striking power of the surface warship, which at one time was thought to have been relegated to a subsidiary role by the aircraft carrier (**Fig. 4**). It is also used by aircraft; submarines can carry them in torpedo tubes, and some submarines have been armed with vertical tube launchers. Newer Tomahawks have increased range and more powerful engines. They use Global Positioning System (GPS) satellite midcourse control and terrain-matching terminal guidance. The high-explosive warhead has an impact fuze; newer Tomahawks also have a variable-time fuze. A new control system will allow ships to program their own missiles, which until now had to be done ashore. Tactical Tomahawk, to enter service in 2003, will be simpler, more reliable, and less costly. *See* SATELLITE NAVIGATION SYSTEMS.

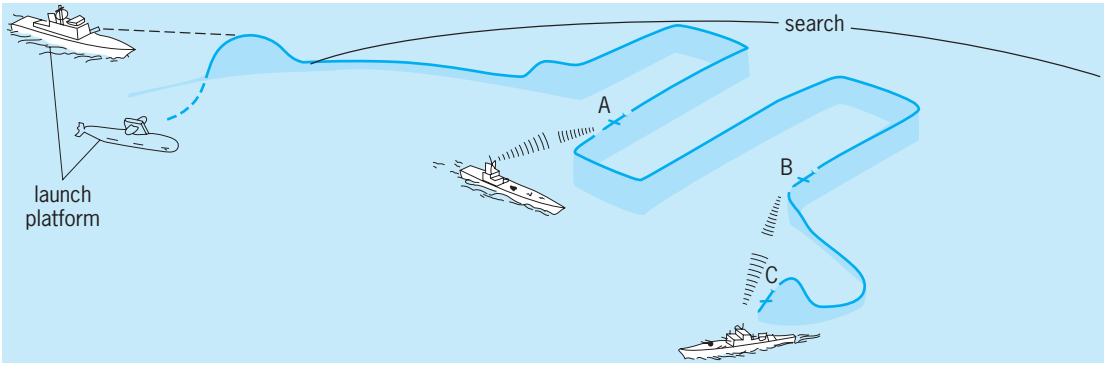

Fig. 4. Guidance of Tomahawk cruise missile. At A, Tomahawk detects ship, identifies it as nonhostile, and continues search. At B, it detects second ship, identifies it as hostile, and commences homing. At C, it acquires target and attacks.

Harpoon is a long-range antiship missile, originally designed as an air-to-surface weapon but now used in surface ships and submarines as well. Several other countries also use Harpoon for coastal defense. Introduced to the fleet in 1977, Harpoon is expected to serve well into the twenty-first century. Midcourse guidance is inertial, to be supplemented by GPS in future models; active radar terminal guidance can home on radar jammers. Later Harpoons have greatly increased range and can turn back to reattack a target missed on the first pass. Surface-ship and submarine Harpoons use a rocket booster; the air-launched version does not require this. Submarine Harpoons are launched in capsules from torpedo tubes; the capsule surfaces and releases the missile.

The Land-Attack Standard Missile (LASM) program will modify early models of the Standard missile to bomblet-dispensing surface fire support missiles for use against land targets. Launched from VLS in surface warships, it will use inertial guidance supplemented by GPS.

Air-to-air missiles. Phoenix was designed as a long-range task force air defense weapon to give the F-14 carrier fighter, with its digital missile control system, the ability to engage up to six targets simultaneously. It has inertial midcourse guidance with input from a semiactive homer; when the target is in range, an active seeker acquires the target and closes to a point at which a proximity fuze can detonate the fragmentation warhead. Newer Phoenix missiles have a speed over Mach 5.0.

Sidewinder is a heat-seeking weapon used in large numbers by the United States and other countries since its introduction in the mid-1950s, and has been continuously improved. Current models embody long combat experience, are faster and longer-ranged, and can better engage rapidly maneuvering targets. Earlier Sidewinders had impact fuzes for their fragmentation warheads; newer models are proximity-fuzed. AIM-9X, an improved close-combat "Evolved Sidewinder," is being developed, with an optical proximity fuze and its infrared seeker guided by the pilot's helmet sight. Early Sidewinders have been converted to the antiradar Sidearm for use by Marine helicopters and Harrier V/STOL ground-attack airplanes.

The original Sparrow I missile entered service in 1956 as a medium-range air-to-air weapon. The newest model of Sparrow III, from which Sea Sparrow was developed, is longer-ranged with much greater computer capability and a more powerful blast-fragmentation warhead with a proximity fuze. Semiactive midcourse guidance is assisted by a command data link. An infrared seeker searches for the target; if it cannot acquire the target, Sparrow continues in the semiactive mode.

Developed to replace Sparrow, though it will probably not completely do so for some time, AMRAAM (Advanced Medium-Range Air-to-Air Missile) has an autopilot which is programmed by the launching plane's computer to bring it within range of the target, when an active radar seeker turns on and engages. A longer-range version, being developed, is intended to replace Phoenix.

Air-to-surface missiles. These weapons can be used against ships; military installations and fortifications; and troops, vehicles, and artillery in the field. Some are designed for use by planes, others by helicopters; some can be used by both.

SLAM (Standoff Land Attack Missile) is designed for use by the F/A-18 Hornet carrier strike fighter. It uses the motor and warhead developed for Harpoon, the infrared seeker from models of Maverick, and Walleye's data link, and incorporates GPS. It was used successfully in the Gulf War. Its course can be corrected in flight, and its infrared seeker transmits a video image to the launching plane's pilot. The pilot visually identifies the target and locks SLAM's control system onto it. SLAM-ER (Expanded Response) has an improved warhead, GPS-assisted inertial midcourse guidance, and an infrared terminal seeker. Its wings, originally developed for Tomahawk, open in flight to improve its range and maneuverability. Its improvements are being backfitted into earlier SLAM missiles.

HARM (High-Speed Antiradiation Missile) is used by carrier planes against ground antiaircraft defenses, and replaced the earlier Shrike missile. Its terminal seeker homes in on a target's missile control radar.

The air-launched version of Harpoon does not use a booster. It is carried by P-3C patrol planes and A-6E carrier attack planes.

The Joint Standoff Weapon (JSOW) was designed to replace laser-guided ground attack weapons, which place the airplane in danger since it must illuminate the target after launching its missile. JSOW is an unpowered inertially guided glide bomb, able to operate in foul weather and capable of carrying a single explosive charge or of attacking multiple targets with bomblets.

The Joint Air-to-Surface Standoff Missile (JASSM) is being developed for long-range use against defended targets. The version being studied will have inertial GPS-assisted midcourse guidance and an infrared terminal seeker, with either a blast-fragmentation warhead or a penetrating warhead for fortified targets.

The Norwegian-designed Penguin Mark 2 lightweight antiship missile, tested for shipboard use, was ordered for use by Marine helicopters. It has inertial midcourse guidance and an infrared homing seeker, with a range of over 21 nautical miles (39 km). *See* HELICOPTER.

Sidearm is an early Sidewinder converted to an antiradar homing missile to arm Marine helicopters.

A heat-seeking version of Maverick is used by Navy F/A-18 attack planes; the Marine Corps uses a laser-designated model. Both versions have a blast-fragmentation warhead.

The lightweight Hellfire air-launched antitank missile is a highly maneuverable supersonic weapon. Designed for laser guidance, Hellfire also takes infrared and radio-frequency/infrared seekers.

Skipper is a relatively inexpensive ground-attack weapon, a 1000-lb (454-kg) bomb fitted with an infrared seeker and guidance head and a rocket motor. *See* AIR ARMAMENT; GUIDED MISSILE.

Antisubmarine weapons. ASROC (antisubmarine rocket), launched by surface warships, was originally designed to carry either a nuclear depth charge or a homing torpedo. All nuclear ASROC warheads were taken out of service by 1989. ASROC is an unguided rocket carrying a Mark 46 homing torpedo. Aimed by shipboard computers using target information obtained by sonar, the rocket is fired from a launcher and follows a ballistic trajectory to the target's predicted position. Torpedo and rocket then separate; the torpedo, slowed by a drag parachute, lands in the water and seeks the target. Vertical-launch ASROC (VLA) is launched by *Ticonderoga*-class cruisers and *Arleigh Burke*-class destroyers from VLS. *See* ANTISUBMARINE WARFARE.

Rockets. Naval rockets, as distinguished from guided missiles, are unguided weapons carrying explosive warheads. Their light weight, in proportion to explosive payload, and lack of recoil let them be used by attack planes and helicopters. Rockeye II Mark 20 delivers a warhead containing 247 small shaped-charge bomblets; APAM (antipersonnel/antimaterial munitions) is an improved bomblet-dispensing rocket. *See* ROCKET PROPULSION.

Torpedoes. The torpedo could be called the first guided missile (**Fig. 5**). Developed and used by the world's navies for more than a century, torpedoes travel underwater on their own power to attack the vulnerable hulls of surface ships and submarines. While an above-water hit from another weapon can sometimes be temporarily disregarded when other matters are urgent, a torpedo hit requires immediate, and often drastic, attention. Modern naval torpedoes are ingenious and effective weapons: fast, far-ranging, and armed with a powerful explosive warhead. Torpedoes may be homing (guiding themselves acoustically to the target); nonhoming (following a preset course); or wire-guided (controlled by signals from the firing ship, transmitted through a trailing wire). They can be launched from surface ships, submarines, or aircraft. Homing torpedoes are used as payload by the ASROC system. Methods for countering the homing torpedo, like the weapons themselves, have been worked on since World War II.

Fig. 5. Homing torpedo fired from a destroyer's launcher. (*U.S. Navy*)

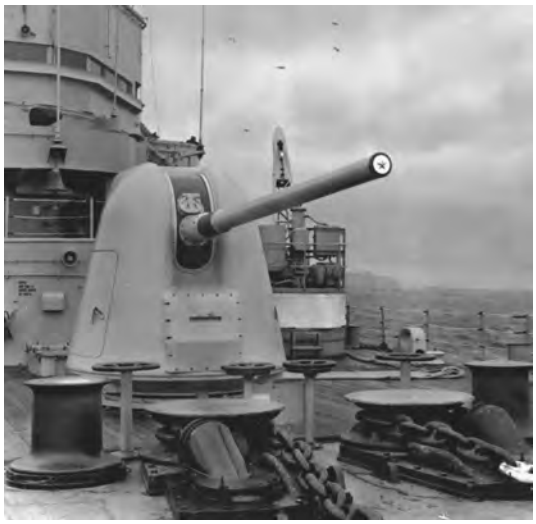

Fig. 6. Mark 45 lightweight 5-in. (127-mm) gun mount. (*U.S. Navy*)

Fig. 7. Phalanx Close-in Weapon System (CIWS). The white dome contains antennas for the weapon's search and tracking radars. (*U.S. Navy*)

It remains a highly effective weapon, and will probably continue in service for a long time. *See* ACOUSTIC TORPEDO.

The submarine-launched 21-in. (533-mm) Mark 48 ADCAP (advanced capability) was designed to combat surface ships as well as fast, high-performance modern submarines. An improved version of the original wire-guided Mark 48 homing torpedo introduced in 1972, ADCAP has a higher-power active/passive homing sonar and increased fuel capacity. Many Mark 48's have been kit-upgraded, and an improved, quieter propulsion system is being developed. The lightweight Mark 46 and Mark 50 torpedoes are used by surface ships and aircraft. Introduced in 1963, the Mark 46 is used by many navies and, in its current improved form, is expected to serve for some years to come. The digitally controlled Mark 50 ALWT (advanced lightweight torpedo) can dive deeper and move faster than the Mark 46; its homing seeker and ability to resist acoustic countermeasures are also improved.

Guns. Though missiles are widely used by ships and aircraft, guns remain significant naval weapons. Missiles are superior for most long-range attack missions and for defense against supersonic planes and missiles at high altitudes and long ranges; the opposite, however, is often true for such missions as shore bombardment, fire support of land forces, and defense against small attack craft. Renewed attention has been given, both in the United States and in other countries, to lighter guns, with high rates of fire, and to quick-reaction control systems for close-in defense against aircraft and missiles in combination with short-range antiaircraft missiles (**Fig. 6**).

Guns in shipboard use range in size from 5-in. (127-mm) to 0.50-caliber (12.7-mm). The 5-in./54-caliber Mark 45 automatic gun can fire 16 to 20 70-lb (31.75-kg) shells up to 12.8 nautical miles (23.7 km) per minute. Used by cruisers and destroyers, it is loaded, controlled, and fired from belowdecks, normally with no one in the gun mount. An improvement program is under way to develop a new version of the Mark 45, with a 62-caliber barrel, to fire 40 99.2-lb (45-kg) or 70.5-lb (32-kg) shells per minute. An Extended Range Guided Munition (ERGM) round is under development for the new Mark 45 gun, designed to use inertial guidance and a miniature GPS receiver to carry a payload of bomblets up to 63 nautical miles (116.7 km) with a circular error of probability (CEP) of 11–22 yd (10–20 m).

The Italian-designed Mark 75 76-mm/62-caliber automatic gun can fire up to 85 14-lb (6.4-kg) shells per minute. This light, compact, remotely controlled gun fires explosive or high-performance fragmentation shells with impact or proximity fuzes. With a digital fire-control system, it arms missile frigates and Coast Guard cutters.

The Mark 15 Close-in Weapon System (CIWS), named Phalanx (**Fig. 7**), is used for rapid-reaction, short-range defense against missiles and planes. CIWS is a modular weapon that includes a search-and-track radar, a computer that identifies incoming high-speed targets and automatically engages them unless overridden by the operator, and an electrically powered six-barreled Gatling-type 20-mm gun, originally designed as the Vulcan for the Air Force. It fires inert, high-density bullets at a cyclic rate of 3000 per minute. The idea behind Phalanx is to produce enough of an umbrella of kinetic-effect projectiles to stand a good chance of hitting an aerial attacker during its final run against a ship.

A number of different types of light automatic guns, firing explosive shells, are used by ships and aircraft, or are being studied for such use. Developed by the Army for armored vehicles, the Mark 38 25-mm Bushmaster is used by combatant craft and is mounted in larger ships for quick-reaction defense against small attack craft and swimmers. A stabilized gun mount, used in *Cyclone*-class patrol craft, accommodates a 25-mm gun and a 40-mm Mark 19 rapid-fire launcher for small,

medium-trajectory, impact-fuzed explosive grenades. A similar stabilized mount is being developed to take a 25-mm gun and short-range air defense missiles. Improved versions of the Oerlikon-designed 20-mm automatic gun of World War II are also used for defense against small craft and swimmers. These fire 800 explosive or armor-piercing rounds per minute. The three-barreled 0.50-caliber (12.7-mm) GAU-19 automatic gun is used in small patrol and special-operations craft. The Browning 0.50-caliber machine gun, developed during World War I as an antitank weapon, was later used by the Navy as an antiaircraft gun. It has been reintroduced for short-range ship defense; the 0.30-caliber (7.62-mm) M60 light machine gun is also used in this role.

Two *Iowa*-class battleships, modernized during the 1980s, remain in the Navy's inactive fleet for potential use in gunfire support of troops ashore. Intensive work during that time showed that the 16-in. (406-mm)/50-caliber gun and its control system could deliver precise fire against fixed or moving targets at long ranges, and that the gun's potential was only then being realized. The 16-in. gun is a bag gun; that is, its powder charge is packed in heavy silk bags. It fires 1900-lb (862-kg) thin-walled explosive shells and 2700-lb (1225-kg) armor-piercing shells, as well as shells which dispense shaped-charge or fragmentation bomblets. Other types of shells have been studied. The *Iowa* has fired accurately at a range of 26 mi (42 km).

The 155-mm Vertical Gun for Advanced Ships (VGAS) is under development for naval fire support. This unusual weapon consists of two vertically mounted guns, automatically loaded to fire up to 10 rounds per gun per minute.

Aircraft guns include the 20-mm Vulcan Gatling-type gun, used by the F-14 Tomcat and F/A-18 Hornet, and the GAU-12 mutibarrel 25-mm gun used by the Marine Corps's AV-8B Harrier.

Bombs. These are free-falling weapons, unlike missiles, which are self-propelled. Bombs take many shapes and sizes, from small antitank and antipersonnel bomblets dispensed from a larger shell or bomb, to heavy weapons designed for blast effect. Navy planes can also deliver nuclear bombs, though these bombs are no longer in service. Most planes and helicopters carry arms externally to accommodate weapon-mix versatility and to keep aircraft size and weight down. High aircraft speeds led to development of streamlined, low-drag bombs. Their better aerodynamic shape is less degrading to aircraft performance, and tends to make them more accurate.

Bombs can be "dumb," that is, uncontrolled, or "smart." Smart bombs have guidance systems and movable control surfaces, and their trajectory can be adjusted to steer them toward a target. Smart bombs were used by Germany and the United States during World War II, but were not seen again until the growing cost of aircraft and sophistication of antiair defenses demanded more accurate weapons that would reduce the number of sorties required to destroy difficult or heavily defended targets.

General-purpose bombs have a high-explosive filler for maximum blast effect. Fire bombs are loaded with fuel gel to ignite fires. Fuel-air explosive (FAE) bombs generate a cloud of fuel vapor in the air; this is detonated for a combination of blast and fire damage. Laser-guided bombs can be steered into specific targets. Cluster bombs carry a payload of small general-purpose, antitank, or FAE bomblets.

Bombs can be detonated by impact, time, or proximity fuzes. Fuzes may be mechanically or electrically operated, and set to explode instantaneously or to allow a delay between impact and explosion. Electric fuzes can be set in flight to suit the target to be attacked. Proximity fuzes automatically explode a certain distance above the target to produce air bursts. Some fuzes have an antidisturbance feature which arms on impact, to set off a bomb if any attempt is made to disturb it. Various safety features prevent fuzes from arming or detonating prematurely.

Navy aircraft use 500-, 1000-, and 2000-lb (227-, 454-, and 907-kg) unguided, aerodynamically shaped bombs. Conventional conical tail fins give them stability in flight. Snakeye fins open into a cross-shaped arrangement of drag plates, slowing a bomb's descent to permit dropping at low altitudes without endangering the delivering plane. Paveway modification kits convert bombs to laser-guided bombs (LGBs)—standoff weapons using semiactive laser guidance. LGBs steer themselves to their targets using illumination from the launching airplane or from another source. They were used to good effect against vulnerable parts of hardened targets during the 1991 Persian Gulf war.

Mines. A mine is a thin-cased, non-self-propelled weapon filled with high explosive and placed underwater, where it is designed to explode when struck, or closely approached, by a ship. Mines can be contact type (fired by actually striking the hull of a passing ship) or influence type (detonated by the close approach of a ship). An influence mine may be magnetic (actuated by a ship's magnetic field), acoustic (actuated by the underwater sound that a ship generates), or pressure (actuated by the change in water pressure caused by a ship's passage). It may also be fired by a combination of these influences. Influence mines are thus much harder to sweep than contact mines. Mines are planted by submarines or aircraft; some navies also use surface minelayers. *See* ACOUSTIC MINE.

Quickstrike mines are factory-modified 500- and 1000-lb (227- and 454-kg) aircraft bombs; the 2390-lb (1084-kg) Quickstrike Mark 65 was specifically designed as a mine. Destructor mines are also aircraft bombs, kit-converted to influence mines in the field. The Mark 56 antisubmarine influence mine floats at a preset depth, moored to an anchor, unlike most influence mines which are laid on the bottom. The deep-water Mark 60 CAPTOR (enCAPsulated TORpedo) releases a Mark 46 acoustic torpedo when triggered by the acoustic "signature" of a hostile submarine. The submarine-launched mobile mine (SLMM) Mark 67, based on the Mark 37 torpedo, is a shallow-water

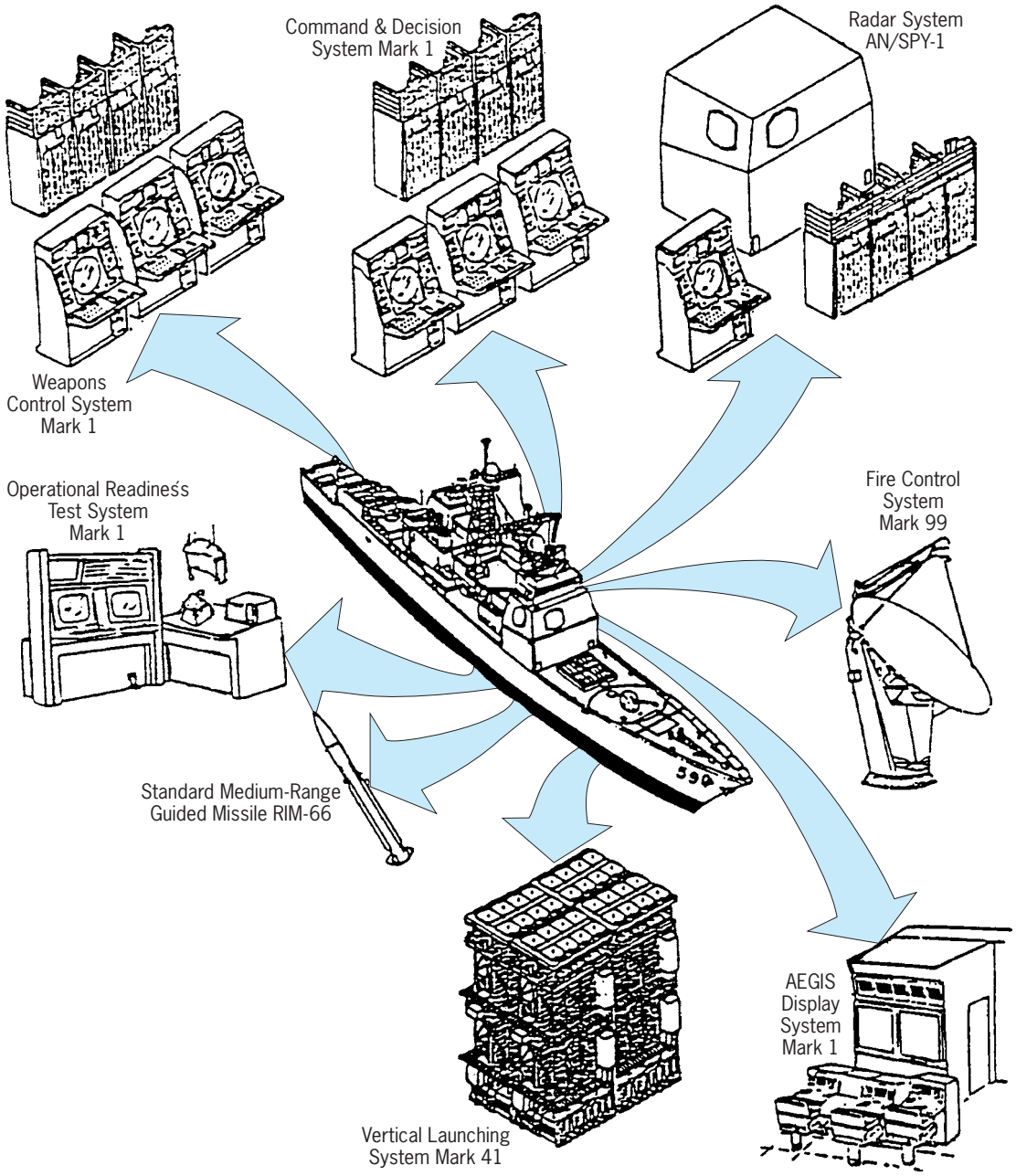

Fig. 8. Elements of the AEGIS Mark 7 weapon system. (*U.S. Navy*)

mine with its own propulsion system to take it into waters inaccessible to conventional minelaying.

Fire control. This is the method used to control guns and other weapons to ensure that shells, missiles, or other weapons hit the desired target. Major advances have been made in fire control since the beginning of World War II, as greatly improved equipment and sophisticated electronics have produced control systems capable of locking on to a fast, elusive target and hitting it. Different types of control systems are used for surface, air, and underwater targets; some of these are dual-purpose (surface and air). Older systems used electromechanical computers; modern systems are designed around digital computers. Emphasis has been placed on systems suitable for use with various weapons, and able to engage multiple surface, air, underwater, or land targets in the rapidly changing environment of modern naval combat.

Several fire control systems are used in Navy ships. The most advanced of these is the AEGIS weapon system, an integrated, computer-contolled system using a multipurpose radar to detect contacts in any direction (**Fig. 8**). It can detect, track, and engage multiple threats of all kinds, while simultaneously maintaining constant all-round surveillance for new threats. It includes the three-dimensional SPY-1 radar, able to track hundreds of targets at once; a command and decision system complex that controls the AEGIS system; a weapon control system which controls the use of the ship's weapons; a fire control system to control the radars used with the Standard missile; command

displays; an automatic fault detection system; and an integrated combat training system. AEGIS also controls a ship's other weapons and electronic systems. The newest Mark 34 gun weapon system was designed around the 5-in./54-caliber Mark 45 gun to be fully integrated into the AEGIS system, and is used in *Arleigh Burke*-class destroyers.

The Mark 86 gun fire control system is used with the 5-in./54-caliber gun in earlier ships for surface and antiaircraft fire as well as shore bombardment. It is linked to such command systems as the Navy Tactical Data System (NTDS). Its digital computer, linked to a "track-while-scan" radar, makes it adaptable for use with the Standard missile. The lighter Dutch-designed Mark 92 fire-control system is used in *Oliver Hazard Perry*-class frigates. Like the Mark 86, it uses a track-while-scan radar and digital computer, and can control guns and missiles against multiple targets. *See* ELECTRONIC WARFARE. John C. Reilly

Bibliography. A. D. Baker III (ed.), *Naval Institute Guide to Combat Fleets of the World*, annually; D. R. Frieden (ed.), *Principles of Naval Weapons Systems*, 1985; N. Friedman, *Naval Institute Guide to World Naval Weapons Systems*, annually; *Jane's Naval Weapon Systems*, annually; D. G. Kiely, *Naval Surface Weapons*, 1988; T. M. Laur and S. L. Llanso, *Encyclopedia of Modern U.S. Military Weapons*, 1995; *Missile Systems of the World*, Raytheon, 1999; S. L. Morison, *Guide to Naval Mine Warfare*, 1995; N. Polmar, *Naval Institute Guide to the Ships and Aircraft of the U.S. Fleet*, annually; *Principles of Naval Ordnance and Gunnery*, NAVEDTRA 12970, 1992.

Navier-Stokes equation

A partial differential equation which describes the conservation of linear momentum for a linearly viscous (newtonian), incompressible fluid flow. In vector form, this relation is written as Eq. (1), where

$$\rho\left[\frac{\partial \mathbf{V}}{\partial t}+(\mathbf{V}\cdot\nabla)\mathbf{V}\right]=-\nabla p+\rho\mathbf{g}+\mu\nabla^2\mathbf{V} \qquad (1)$$

ρ is fluid density, $\mathbf{V}$ is fluid velocity, p is fluid pressure, $\mathbf{g}$ is the gravitational acceleration, μ is fluid viscosity, ∇ is the del or grad operator, and ∇^2 is the laplacian operator. The equation is named after its two principal developers, French engineer C. L. M. H. Navier (1823) and Irish scientist George G. Stokes (1845). When coupled with the conservation of mass relation, $\nabla\cdot\mathbf{V}=0$, Eq. (1) can be solved for the space-time distribution of $\mathbf{V}$ and p in a given region of viscous fluid flow. Typical boundary conditions are (1) the knowledge of the velocity and pressure in the far field, and (2) the no-slip condition at solid surfaces (fluid velocity equals solid velocity). *See* CALCULUS OF VECTORS; GRADIENT OF A SCALAR; LAPLACIAN; NEWTONIAN FLUID; VISCOSITY.

Equation (1) has been successfully applied to the prediction of both laminar flow and disorderly turbulent fluid flows. In the latter case, the evaluation of the equation's adequacy has been limited to a comparison between numerical approximations to solutions of Eq. (1) and measured time-average values and higher statistics. A significant limitation of Navier-Stokes theory is the lack of any proof regarding uniqueness or existence of solutions of Eq. (1) for given boundary and initial conditions.

The primary dimensionless parameter which governs Eq. (1) is the Reynolds number, given by Eq. (2),

$$\mathrm{Re}=\frac{\rho VL}{\mu} \qquad (2)$$

where L is a characteristic body dimension. For small $\mathrm{Re}\ll 1$, Eq. (1) can be simplified by neglecting the left-hand side, resulting in a linear approximation called Stokes flow, or creeping flow, for which many solutions are known. *See* CREEPING FLOW; LUBRICATION; REYNOLDS NUMBER.

For large $\mathrm{Re}\gg 1$, viscous effects are often confined to a thin boundary layer near solid surfaces, with the remaining flow being nearly inviscid. Equation (1) reduces to a simpler boundary-layer theory, which can be solved in marching fashion along the surface, using the outer inviscid flow as a driving function for the viscous effects. *See* BOUNDARY-LAYER FLOW; D'ALEMBERT'S PARADOX.

The advent of large-scale, high-speed computers allows Eq. (1) to be modeled and solved by finite-difference techniques. Very fine grids are needed to resolve flow details, but parallel processors allow up to 10^7 mesh points. Laminar flows (moderate Re) are especially well modeled, while turbulent flows (high Re) may require further simplifications of the turbulence terms. *See* COMPUTATIONAL FLUID DYNAMICS; LAMINAR FLOW; TURBULENT FLOW.

For compressible viscous flow, where density ρ is variable, the term $\mu\nabla^2\mathbf{V}$ in Eq. (1) is replaced by a more complicated set of stress gradients, including a second coefficient of viscosity, λ. Temperature is also variable, and Eq. (1) must be coupled to the partial differential equation of energy conservation. *See* COMPRESSIBLE FLOW; FLUID-FLOW PRINCIPLES; HEAT TRANSFER. Frank M. White; Arthur E. Bryson, Jr.; A. Gordon L. Holloway

Bibliography. J. A. Schetz, *Boundary Layer Analysis*, Prentice Hall, Upper Saddle River, NJ, 1992; H. Schlichting and K. Gersten, *Boundary-Layer Theory*, 8th ed., Springer, 2000; F. S. Sherman, *Viscous Flow*, McGraw-Hill, New York, 1990; F. M. White, *Viscous Fluid Flow*, 3d ed., McGraw-Hill, New York, 2005.

Navigation

The process of directing the movement of a craft from one place to another. Navigation involves position, direction, distance, time, and speed.

The process of keeping track of a craft's location by measuring and applying progress from a previous position is called dead reckoning. The location of a craft relative to external reference points such as landmarks or aids to navigation is called piloting.

Radio navigation involves determining distances or directions to radio transmitters. Celestial navigation involves the use of celestial bodies. *See* CELESTIAL NAVIGATION; DEAD RECKONING; PILOTING.

Types of Craft To Be Navigated

The craft to be navigated may be a ship, small marine craft, land vehicle, aircraft, missile, spacecraft, or any moving object requiring direction or capable of being directed, even an animal or bird. The characteristics of the craft have a significant influence upon the type of navigation and the equipment used. Size, mission, weight and space limitations, and economic factors are important considerations.

Ships. Ships of many types are navigated. Combatant military vessels generally require relatively high absolute and relative accuracy and great versatility to accommodate a variety of missions. The capability of coping with the effects of high maneuverability, short time lag between sensing and display of information, and extensive redundancy to maintain navigation capability in the event of damage in combat are important requirements. *See* NAVAL SURFACE SHIP.

Merchant vessels are characterized by voyages from one point to another involving essentially constant courses and constant speeds over relatively long periods of time, little reserve speed, maneuvers only when entering or leaving port or avoiding an obstacle such as another ship, high ratio of time at sea to time in port, and limited crew size. Simplicity of operation, minimum training requirements, good reliability, high maintainability, low obsolescence factor, and low cost of purchase and maintenance are important considerations. These characteristics are typical of freighters, less significant in the case of large passenger vessels, and acutely applicable in the case of fishing vessels. Supertankers with hundreds of thousands of tons of displacement have introduced requirements related primarily to lack of maneuverability because of their size and inertia and to the consequences of a disaster involving such a vessel. *See* MERCHANT SHIP.

Special-purpose vessels, such as hydrographic survey ships, missile tracking vessels, and submersibles, have a wide diversity of requirements related to their size, speed, mission, and environment. *See* OCEANOGRAPHIC VESSELS; UNDERWATER VEHICLES; UNDERWATER NAVIGATION.

Small craft. Various types of small craft, particularly recreational craft, have varying requirements commensurate with their operating characteristics. Oceangoing yachts have requirements essentially the same as those of merchant vessels. However, except for the larger yachts, cost, size, weight, and power limitations are generally severe. Sailing vessels may also have other limitations, such as lack of a constant source of electric power. Small craft that do not venture out of sight of land generally have limited navigational capability to match their needs. *See* MARINE NAVIGATION.

Land vehicles. Like small marine craft, land vehicles vary widely in their navigational capability. The equipment and techniques involved in navigation of vehicles constrained to using highways or rails, although of some diversity, are generally familiar. Vehicles used for transportation across deserts or polar ice fields, or for traversing a jungle or pathless forest, require more attention to navigation.

Inexpensive satellite navigation equipment is increasingly in use in automobiles and commercial vehicles. These applications support timely routing and monitoring of commercial and public safety traffic, automatic monitoring and collection of tolls, and the collection and dissemination of geographic information that allows a navigation-enabled user to find the preferred path from one location to another.

Aircraft. In contrast with surface ships and land vehicles, aircraft generally are characterized by higher speed, greater maneuverability, three-dimensional capability, greater limitations of weight and space, and severe limitations in the inability to hover or park in one spot while the navigation problem is being solved.

The effect of speed is to require less time lag between sensing and display of navigational information, more rapid interpretation of data, and the need for frequent position fixes. Speed is also an important factor in the need for some kind of air-traffic control to accommodate increased traffic density while maintaining safety of flight. Avoidance of other aircraft is an important function of air navigation.

Among different types of aircraft the diversity of characteristics is great. Military fighter aircraft involve high maneuverability, relatively small size, and limited crew. Bombers require high accuracy and, like fighters, have high speed. *See* MILITARY AIRCRAFT.

Intercontinental aircraft are generally large and of two types. The supersonic jet flies above most of the weather and favors a path that climbs as the aircraft becomes lighter through use of fuel. Because of the relatively small effect of wind on a fast-moving aircraft, the envelope of optimum paths at various times between any two points is of small lateral extent. Because of the very high speeds and the lesser effects of winds, a self-contained dead-reckoning type of navigation system providing continuous output of steering information is particularly favored.

The navigation requirements of subsonic turbojets are similar to those of supersonic jets but less critical. Favorable winds are a greater consideration because of the longer flight time.

The diversity and number of aircraft flying in domestic airspace of highly developed countries are considerably greater than is the case over the oceans. In addition to the types that fly over the ocean aircraft flying over land include smaller commercial aircraft, light pleasure and other general aviation craft, short takeoff and landing craft, helicopters, and airships. Because of the availability of numerous short-distance aids and the reduced separation standards, aircraft flying over these land areas make less use of long-distance aids, either of the station-reference piloting type or the self-contained dead-reckoning type. *See* AIR NAVIGATION; AIRSHIP; GENERAL

AVIATION; HELICOPTER; SHORT TAKEOFF AND LANDING (STOL).

Spacecraft. While involving the basic elements of navigation, spacecraft apply them in a somewhat different manner. Upon separation from its booster the spacecraft becomes a celestial body influenced in its motions by all other celestial bodies near enough and large enough to exert a significant gravitational pull. While terrestrial navigation involves motion from one fixed point on the Earth to another (except in the case of interception of a craft in motion), space navigation involves motion from one moving point to another. For successful completion of a flight, a spacecraft must arrive at a designated point in space at the same time as the celestial body to be intercepted. During flight a spacecraft might have long periods during which there is no application of propulsive power. It cannot be steered except by application of power; its attitude can be controlled only by reaction; and terrestrial forms of navigation are largely unavailable. *See* SPACE NAVIGATION AND GUIDANCE.

Elements of Navigation

As indicated above, navigation is concerned primarily with position, direction, distance, time, and speed.

Position. On the Earth, position is generally expressed in terms of geodetic latitude and geodetic longitude, although position of a craft may be stated relative to an aid to navigation, landmark, terminal, or a scheduled position. *See* LATITUDE AND LONGITUDE; TERRESTRIAL COORDINATE SYSTEM.

The position of primary interest to a navigator is that of his or her own craft. If this is determined by advancing a previous position for the direction and distance of travel, it is termed a dead-reckoning position. If it is determined by means of external references, it is called a fix. Traditionally, a fix is established at the intersection of lines of position, which might be arcs of small or great circles, hyperbolas, or other lines.

Direction. Horizontal direction is customarily expressed as angular distance from a reference direction. An adjective may be used to indicate the particular reference used. A "true" direction is based upon the direction of the north geographical pole from the observer. A "magnetic" direction uses the northward direction along the magnetic meridian which passes through the observer. The reference for a "compass" direction is north as indicated by a magnetic compass. This may differ from magnetic north because of magnetic influences within the craft itself. If a marine gyrocompass is used, a "gyro" direction results. Particularly in high latitudes, a somewhat arbitrary grid may be substituted for the graticule of meridians and parallels of latitude. The conventional "north" on this grid is the reference for "grid" direction. The forward direction along the longitudinal axis of the craft is the origin of "relative" directions. *See* GYROCOMPASS; MAGNETIC COMPASS; POLAR NAVIGATION.

The direction in which a craft is pointed is called its heading. The intended direction of travel is called the course, while the actual direction from a point of departure to a subsequent point of arrival is called the course made good. The actual path followed, and sometimes also the direction of this path, is called the track. A bearing is the direction of one terrestrial point from another, generally the direction of an object as viewed from a craft. The horizontal direction to a celestial body is that body's azimuth, a term also used to indicate the horizontal pointing direction of an antenna.

Distance. Navigators usually express distance in nautical miles. One nautical mile is 1852 m or 6076.115486 ft. This is very nearly the length of 1 minute of arc of a great circle on the surface of the Earth. Yards, feet, and meters are often used for short distances. Depths of water are expressed in fathoms (6 ft or 1.8288 m), feet (1 ft = 0.3048 m), or meters, and heights are expressed in feet or meters.

Time. In navigation, time to the nearest integral minute is usually used for indication of the position of a craft. Celestial observations are usually timed to the nearest integral second. Navigators avoid the use of A.M. and P.M. designation, expressing time on a 24-h basis. Zone time (ZT) and Greenwich mean time (GMT) are used as applicable. Inexpensive quartz-crystal digital watches, exceeding the accuracy of marine chronometers, are now generally used for time both at sea and in the air. They are checked at intervals by radio time signals. *See* TIME.

Speed. The speed of a craft, current, or wind is generally expressed in knots. One knot represents 1 nautical mile per hour (0.5144 m/s). The speed of high-speed aircraft and missiles may be expressed in terms of the Mach number, which is the ratio of the speed of the craft to the speed of sound in the medium concerned. *See* MACH NUMBER.

Navigational Aids

Anything used in navigation, whether aboard the craft or external to it, is properly termed a navigational aid. Thus, in addition to onboard navigational equipment, the term includes such external aids as natural landmarks, prominent buildings, or other structures. Although sometimes used synonymously with "navigational aid," the expression "aid to navigation" is generally restricted to an object or device, external to the craft, established expressly to assist navigation.

Aids to navigation. In the restricted sense, aids to navigation for mariners consist of buoys, beacons, lighthouses, lightships, and navigation sound and electronic transmitters including Loran and navigation satellites. Aids to navigation for aviators consist primarily of radio ranges and beacons and radio position-fixing transmitters including navigation satellites. *See* BUOY; ELECTRONIC NAVIGATION SYSTEMS; HYPERBOLIC NAVIGATION SYSTEM; LIGHTHOUSE; TACAN; VOR (VHF OMNIDIRECTIONAL RANGE).

Charts and publications. A map intended primarily for navigation is called a chart, which displays such information as latitude and longitude scales, useful topographical features and aids to navigation, heights of land and depth of water, cultural features, obstructions, magnetic information, information on

electronic aids, and warning notes. Thus charts are useful for planning trips; for measuring courses, distances, and safe routes; and indicating the available aids to navigation. During the trip a chart is useful to indicate progress. Fixes can be determined by plotting lines of position directly on the chart and graphically moving those that need adjustment to a common time.

The selection of a suitable projection is an important consideration in the production of charts. An important property is conformality, permitting directions to be represented correctly. Nearly all nautical charts, except those for high latitudes, are on the Mercator projection. The Lambert conformal projection is widely used for aeronautical charts. The gnomonic projection, although nonconformal, is useful for representation of great circles, which appear on it as straight lines. The polar stereographic, transverse Mercator, modified Lambert conformal, and azimuthal equidistant projections are in limited use, particularly in polar regions. *See* MAP PROJECTIONS.

Much of the information shown on charts is in internationally standardized symbols and abbreviations. Nautical and aeronautical charts are similar in many respects, but they differ in the emphasis given certain information. Nautical charts are concerned primarily with water areas and information of interest to mariners, while aeronautical charts emphasize land areas, airports, heights of obstructions, and other information of interest to fliers.

Electronic charts. The electronic chart is an electronic facsimile of a nautical chart produced by a microprocessor. The chart becomes a basic part of an electronic chart display information system (ECDIS) when a suitable electronic positioning system such as the Global Positioning System (GPS) or the differential Global Positioning System (DGPS) is used to display a symbol of the ship at its position on the chart. Radar can be used to add useful information such as images of other vessels and buoys, and to confirm the accuracy of the positioning system. Color can be added to help in the identification of various types of information. The possibilities for adding or omitting data are almost limitless, making the combined display an effective presentation of the entire situation of the ship. The display is similar to that of a vessel traffic system (VTS), thus contributing to effective traffic coordination in harbors. Electronic charts can be updated by radio, so that vessels can enter harbors with up-to-date charts. *See* VESSEL TRAFFIC SERVICE.

Paper nautical charts are accepted as legal documents for litigation. In view of the numerous variables in the design and production of electronic charts, standards must be adopted if these charts are to have equal legal stature with paper charts. By the early 1990s, provisional standards were available.

A somewhat simplified electronic navigation system (ENS), using an electronic chart and differential Global Positioning System, has been suggested as more suitable than an electronic chart display information system for some applications, such as traversing a dredged channel. Unlike the electronic chart system, the electronic navigation system would supplement, rather than replace, a paper chart. The position of the vessel, determined by the differential Global Positioning System where available, would be converted from geographical to channel or other preplanned track coordinates.

Navigation Practice

Early navigation was essentially an art—the ability to interpret by human senses the elements of navigation by visual observation, sound, and smell.

As knowledge and understanding of the environment increased, the principles of science were utilized in order to develop instruments to improve the accuracy and extend the range of navigation. The application of science has made possible the development of celestial navigation and various new applications of the principles of piloting, dead reckoning, and radio navigation. Radio-navigation aids are now most frequently used.

Thus science has increasingly entered the realm of navigation, culminating in a guidance system that permits accomplishment of a mission without human intervention beyond the programming of the guidance equipment and the launching of the craft. Conventional navigation of craft bearing humans from one place to another generally consists of a blend of science and art. *See* GUIDANCE SYSTEMS.

Traditional methods. Throughout most of the nineteenth century, dead reckoning was generally performed manually by plotting on a chart or plotting sheet. This procedure provided a graphic indication of progress of the craft and a means of evaluating other navigational data received. When the craft was within sight of land or aids to navigation at sea, such as lightships, fixes were obtained from time to time by means of lines of position obtained generally by observing bearings of, or measuring the distance to, visible reference marks. Unless fixes were obtained at intervals of a very few minutes, as when traversing a channel, new dead-reckoning plots were started at each new, reliable fix.

At sea, navigation became more leisurely, consisting primarily of timed celestial observations by sextant during morning and evening twilight, observations of the Sun during the morning, at noon, and during the afternoon, and occasional observations of the Moon, and perhaps Venus. *See* SEXTANT.

Electronics. Traditional piloting and celestial navigation have severe weather limitations. Visual bearings are impossible during periods of heavy fog, and celestial observations with marine sextants can be made only when both the celestial bodies and the horizon are clearly visible. Electronics has increasingly been used to free the navigator from these limitations. The widespread use of computers has further revolutionized the navigation process.

Electronics was introduced into navigation early in the twentieth century when time signals were first broadcast by radio, thus increasing the accuracy and reliability of time indication at sea, an essential

element of longitude determination. The next application was the broadcast of marine weather forecasts, navigational warnings, and other useful information.

The earliest use of electronics for position fixing was in the early 1920s when a medium-frequency radio direction finder was developed, permitting measurement of direction of a known point out of sight, and during periods of restricted visibility. Other methods of determining direction followed. Directional antenna arrays were used extensively in the low-frequency four-course airway ranges established throughout the United States and elsewhere. Rotating patterns of electronic signals provided an additional method. The introduction of radar shortly before the start of World War II provided a means for determining both direction and distance simultaneously. *See* AIRBORNE RADAR; DIRECTION-FINDING EQUIPMENT; RADAR.

The next significant step in the application of electronics to navigation was the introduction of hyperbolic position fixing systems broadcasting synchronized radio signals from sets of two transmitters at known points. Measurement was made of the difference in arrival time of synchronized pulsed signals, or the phase difference of synchronized continuous-wave signals from the two stations, thus permitting the navigator to determine the difference in distances from the stations.

As the uses of electronics in navigation have proliferated, a wide variety of applications have appeared, such as underwater logs, echo sounders, sonar, compass repeaters, Doppler navigators, and electronic charts. Because of limitations on celestial observations, electronic methods of position determination have been particularly attractive. *See* DOPPLER RADAR; ECHO SOUNDER; SONAR.

Inertial navigation. The development of inertial navigators provided a self-contained, passive system continuously indicating the dead-reckoning position of a craft, with worldwide application in any kind of weather. It is widely used in long-range aircraft, submarines, and some special-purpose ships such as those engaged in geophysical exploration. When properly aligned, an inertial navigator uses inertial sensors to indicate speed, heading, and position. Gyroscopes are used to sense angular motions of the craft and to maintain accelerometers in the correct orientation to sense linear accelerations, or changes of speed. Single integration of acceleration provides a measure of speed, and double integrations a measure of distance. The term inertial indicates that measurements are made relative to inertial space; since output relative to the surface of the Earth is desired, provision must be made to eliminate effects related to rotation of the Earth. *See* INERTIAL GUIDANCE SYSTEM.

Satellite navigation. The NAVSTAR Global Positioning System provides very high-accuracy, worldwide, continuous, all-weather, three-dimensional fixes; time; and craft velocity indication. This system transmits signals used to measure pseudorange and pseudorange rate, and simultaneous data from four or more satellites are converted by the user equipment to the desired information. Position accuracy is measured in meters in the best user equipment. Full operation capability was achieved in 1995.

Because of the desirable characteristics of the Global Positioning System and the development of versatile small user equipment units, the system has wide application in the air, at sea, and on land. As use of the Global Positioning System increases, other positioning systems can be expected to become less popular, and some of them eventually to be phased out. However, because malfunction or other sources of error of any electronic system is a possibility, some form of backup is needed. The integrity, that is, the ability of a system to indicate nonnormal operation, is an important consideration.

Because of the effect of the Earth's atmosphere on the propagation of electronic signals, and present intentional degradation of accuracy for military purposes, accuracy of the Global Positioning System can be improved by use of the differential Global Positioning System. Equipment at an accurately known position, such as in the vicinity of an airport, compares position determined by the Global Positioning System with the known position and transmits the difference as corrections to pseudoranges for use of other receivers in the area.

The Global Orbiting Navigation Satellite System (GLONASS) has been developed in Russia concurrently with the development of the Global Positioning System. It is similar to the Global Positioning System in concept but uses different frequencies and signal formats. *See* SATELLITE NAVIGATION SYSTEMS.

Computers. Development of the electronic computer eliminated much of the drudgery of navigation and enhanced safety by reducing the interval between procurement of data and its application. Further, the computations involved in the use of advanced systems such as inertial navigators and satellite positioning systems would render these systems totally impractical without the availability of fast, accurate electronic computers. Finally, computers make possible the development of integrated systems that provide greater accuracy and reliability by synthesizing the outputs of several independent positioning systems.

The development of fast, accurate computers, particularly the maturity of microcomputers, combined with advanced uses of electronics, has revolutionized navigation. Celestial navigation has essentially been relegated to the realm of recreational boating. The training needed to produce competent navigation has been simplified, and as a result the need for professional navigators has virtually been eliminated. The availability of continuous, accurate real-time position has contributed significantly to safety at sea and in the air. *See* DIGITAL COMPUTER; MICROCOMPUTER.

Traffic control. Traditionally, mariners have resisted attempts by shore-based personnel to restrict their movements at sea. Aviators, on the other hand, recognized early in the history of air travel that some form of traffic control, similar to that in operation on land,

was essential to their safety, and readily accepted the establishment of rules and a system of ground-based direction to help them avoid collisions with obstructions such as other aircraft. As ship speeds have increased, the number of vessels plying the oceans has multiplied, and the ability to communicate and locate the positions of one's own and other vessels has improved, modern ship masters have recognized the advantages of an orderly, prescribed flow of traffic in congested areas, resulting in a proliferation of ship routing systems. Additionally, the improvement in dissemination of reliable weather information has increased the reliance on weather routing of ships on the high seas to permit the selection of routes that are optimum in times of transit and ensure freedom from disabling and discomforting passages.

As the number of aircraft in flight has increased, the available airspace has become increasingly crowded, resulting in delays. The increase in accuracy of position determination of aircraft on a real-time basis makes possible reduction of separation standards, thus increasing airspace capacity. *See* AIR-TRAFFIC CONTROL.

Approach and landing. The most critical part of a flight is the approach and landing. High accuracy and avoidances of other aircraft are critical. Without some form of assistance, landings would be virtually impossible during periods of low visibility. To solve this problem, an electronic instrument landing system (ILS) was developed. Additional transmitters located alongside the runway provide guidance along a runway centerline approach path. In 1949 the International Civil Aviation Organization (ICAO) adopted the system as a worldwide approach and landing aid, but as the demands for precision landing in any condition of visibility increased, a higher-accuracy, increased-capability microwave landing system (MLS) using higher frequencies was developed and may replace the instrument landing system as the international standard. The differential Global Positioning System is also suitable for use as a high-accuracy, reliable landing aid. *See* INSTRUMENT LANDING SYSTEM (ILS); MICROWAVE LANDING SYSTEM (MLS).

Alton B. Moody

Bibliography. N. Bowditch, *American Practical Navigator*, *2002 Bicentennial Edition*, National Imagery and Mapping Agency, 2002; T. Cunliffe, *Coastal and Offshore Navigation*, 2d ed., 1999; T. J. Cutler, *Dutton's Nautical Navigation*, 15th ed., 2004; R. R. Hobbs, *Marine Navigation*, 2 vols., 4th ed., 1998; M. Kayton and W. R. Fried, *Avionics Navigation Systems*, 2d ed., 1997.

Neandertals

A group of late archaic humans from Europe, the Near East, and central Asia that immediately preceded the first modern humans in those regions. The Neandertals (also spelled Neanderthals) are included by some within the species *Homo sapiens*, recognizing their close affinities to modern humans; others place them in their own species, *Homo neanderthalensis*, emphasizing the differences between them and modern humans.

The first recognized Neandertal remains were found in the Neander Valley near Dusseldorf, Germany, in 1856. Since then the remains of several hundred Neandertals have been discovered. Since the Neandertals were the first humans to bury their dead, a number of largely complete skeletons are preserved, providing detailed knowledge of their biology. Fossils that represent human populations of a similar evolutionary stage have been found in Africa and eastern Asia; these are sometimes referred to as Neandertal-like. Collectively, all of these predecessors of early modern humans can be called late archaic humans. *See* EARLY MODERN HUMANS.

In the early twentieth century, when Neandertals were the only archaic humans known, they were reconstructed as semihuman, dull-witted, and brutish. Hence their popular image was that of the archtypical cavemen. They are now recognized as relatively recent members of the human lineage; they lived between about 125,000 and 36,000 years ago (and as late as 30,000 years ago in certain isolated regions), as compared with earlier members of the genus *Homo* who extend back more than 2 million years. The Neandertals share many features with modern humans both anatomically and behaviorally. Yet, a number of important contrasts between them and more recent humans are recognized.

The Neandertals evolved gradually from more archaic humans during the last interglacial, 100,000–125,000 years ago, across Europe and western Asia. Fossils predating the Neandertals (between 500,000 and 200,000 years old) are becoming well known in these regions. These earlier archaic humans exhibit some but not all of the cranial, facial, and postcranial skeletal features that characterize the Neandertals. Except for the high mountains and the arctic areas, they inhabited most of that region during both the warm last interglacial and the cold phases of the first half of the last glacial period of the Pleistocene Epoch. The Neandertals disappeared 30,000–45,000 years ago across that region, after which time only the fossil remains of early modern humans (called Cro-Magnons in Europe) have been found. Other late archaic humans evolved into or were absorbed by early modern humans probably prior to 50,000 years ago in Africa and by 30,000 years ago in eastern Asia. In the Near East, the fossils of both Neandertals and early modern humans are found from the period between 100,000 and 40,000 years ago. The two groups alternated periods of occupation as the climate and environment of the area fluctuated during the later Pleistocene. Other late archaic humans evolved into and were absorbed by early modern humans, prior to 50,000 years ago in Africa, and by 30,000 years ago in eastern Asia.

Physically, the Neandertals were about the same height as most modern humans, on the average 5 ft 5 in. (166 cm), but they were much more heavily built. They had heavy necks, broad and muscular shoulders, and extremely muscular arms, hands, and legs. Estimates of their strength show them to

have been about as strong as very athletic modern humans. Their leg bones show a marked thickening of their shafts, which is indicative of both marked strength and endurance—a necessary part of their survival. Even though this physique was energetically costly, it enabled them to exert great strength whenever needed and, more importantly, endure habitually high levels of physical activity.

This strength and endurance reflects their generally less efficient hunting and gathering way of life, as compared with early modern humans. Throughout their range, Neandertal fossils are generally associated with a Middle Paleolithic stone tool tradition called the Mousterian. Neandertal technology consisted of hand-held or wooden-handled stone tools and simple wooden or stone-tipped spears. They appear to have been semi-opportunistic hunters and gatherers, expending considerable time and energy searching for game and plant food, as well as scavenging from the kills of large carnivores. Symbolic systems for keeping track of information about the environment were apparently unknown among them.

The Neandertals are known for their long, low braincases and their projecting faces with large brows and prominent noses (see **illus.**). Their brains were larger than those of modern humans, about 90 in.^{3} or 1500 cm^3 on the average (compared to 78–84 in.^{3} or 1300–1400 cm^3 averages for modern humans). The large brain size was due in part, as with early modern humans, to their large body masses. The length and lowness of their braincases was due to relatively slow brain growth during infancy. There is no evidence that they were less intelligent than modern humans, only that their behavioral system was less elaborate.

The faces of Neandertals had large projecting browridges that led onto wide projecting noses (see illus.). Below their noses were dental arcades that held cheek teeth about the size of large modern human ones, and front teeth much larger than those of modern humans. From the rapid rates of wear on their front teeth, it is apparent that they used them for many purposes. As with their muscular hands and arms, they used the front teeth to compensate for deficiencies in their technology. Although large and projecting, their faces were not especially massive, suggesting that they chewed with little more force than modern humans.

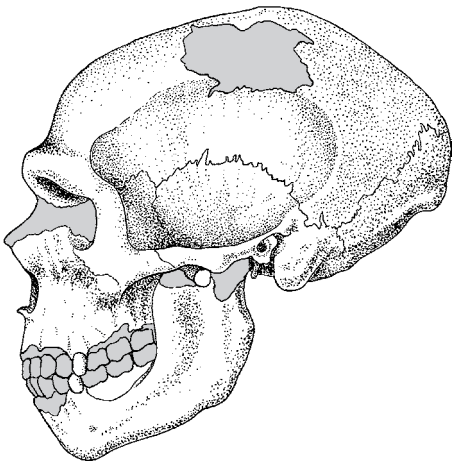

Reconstructed skull of a Neandertal from the cave at La Chapelle-aux-Saints, France. (*After M. F. Ashley Montagu, Introduction to Physical Anthropology, 2d ed., 1951*)

The Neandertals had relatively short forearms and lower legs, and wide trunks, as do modern arctic peoples, indicating an adaptation to minimize heat loss in a glacial climate. Their early modern successors had considerably longer forearms and lower legs, a condition that indicates more protection from the cold. In fact, shelters and hearths among the Neandertals were rudimentary at best, providing little warmth and protection.

There are other indications of the harshness of life among the Neandertals. The oldest lived only into their late 40s, and very few lived beyond their mid-30s. All who managed to survive to age 40 suffered at least one broken bone, and many had debilitating arthritis. About 75% of the Neandertals exhibit malformed tooth enamel, indicating periods of starvation during childhood. Despite these problems, they managed to keep several severely injured individuals alive for decades; to be the first humans to survive glacial climates in Europe and Asia; and to bury their dead.

The position of the Neandertals in modern human ancestry remains controversial. In some regions, the transition from Neandertals to early modern humans was rapid, with more changes in human anatomy than has occurred at any other similarly short time period in human evolution. Many scholars have argued that the time available to transform a Neandertal into a modern human was too short, and they conclude that the Neandertals contributed little, if at all, to populations of modern humans. Others see a natural progression from Neandertals to modern humans, with only an acceleration in the rate of change at the time of the transition. The truth may lie between these extremes, with the Neandertals being absorbed by early modern humans moving slowly into the Near East, central Asia, and Europe between about 100,000 and 30,000 years ago.

Whatever the extent to which Neandertals can be claimed to be ancestors of modern humans, they represent the most recent phase of premodern humans, one in which people were less efficient than modern humans at hunting and gathering, and compensated for their cultural limitations with biological attributes such as tremendous strength, large front teeth, and thermal adaptations. Yet they exhibited the beginnings of many of the attributes of modern humans. They were very successful for about 100,000 years, but they were eventually replaced by humans who were better able to exploit their environments. *See* FOSSIL HUMANS. Erik Trinkaus; Steven Churchill

Bibliography. M. H. Nitecki and D. V. Nitecki (eds.), *Origins of Anatomically Modern Humans*, 1994; F. H. Smith, *The Origins of Modern Humans*, 1984; E. Trinkaus (ed.), *The Biocultural Emergence of Modern Humans in the Late Pleistocene*, 1989; E. Trinkaus and P. Shipman, *The Neandertals: Changing the Image of Mankind*, 1994; M. H. Wolpoff,

Human Evolution, McGraw-Hill, 1996; C. Stringer and C. Gamble, *In Search of the Neanderthals*, Thames & Hudson, 1993.

Nearshore processes

Processes that shape the shore features of coastlines and begin the mixing, sorting, and transportation of sediments and runoff from land. In particular, the processes include those interactions among waves, winds, tides, currents, and land that relate to the waters, sediments, and organisms of the nearshore portions of the continental shelf. The nearshore extends from the landward limit of storm-wave influence, seaward to depths where wave shoaling begins. *See* COASTAL LANDFORMS.

The energy for nearshore processes comes from the sea and is produced by the force of winds blowing over the ocean, by the gravitational attraction of Moon and Sun acting on the mass of the ocean, and by various impulsive disturbances at the atmospheric and terrestrial boundaries of the ocean. These forces produce waves and currents that transport energy toward the coast. The configuration of the landmass and adjacent shelves modifies and focuses the flow of energy and determines the intensity of wave and current action in coastal waters. Rivers and winds transport erosion products from the land to the coast, where they are sorted and dispersed by waves and currents.

In temperate latitudes, the dispersive mechanisms operative in the nearshore waters of oceans, bays, and lakes are all quite similar, differing only in intensity and scale, and are determined primarily by the nature of the wave action and the dimensions of the surfzone. The most important mechanisms are the orbital motion of the waves, the basic mechanism by which wave energy is expended on the shallow sea bottom, and the currents of the nearshore circulation system that produce a continuous interchange of water between the surfzone and offshore areas. The dispersion of water and sediments near the coast and the formation and erosion of sandy beaches are some of the common manifestations of nearshore processes.

Erosional and depositional nearshore processes play an important role in determining the configuration of coastlines. Whether deposition or erosion will be predominant in any particular place depends upon a number of interrelated factors: the amount of available beach sand and the location of its source; the configuration of the coastline and of the adjoining ocean floor; and the effects of wave, current, wind, and tidal action. The establishment and persistence of natural sand beaches are often the result of a delicate balance among a number of these factors, and any changes, natural or anthropogenic, tend to upset this equilibrium. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; EROSION.

Waves. Waves and the currents that they generate are the most important factors in the transportation and deposition of nearshore sediments. Waves are effective in moving material along the bottom and in placing it in suspension for weaker currents to transport. In the absence of beaches, the direct force of the breaking waves erodes cliffs and sea walls.

Wave action along most coasts is seasonal, responding to changing wind systems over the waters where the waves are generated. The height and period of the waves depend on the speed and duration of the winds generating them, and the fetch, or distance, over which the wind blows. Consequently, the nature and intensity of wave attack against coastlines vary with the size of the water body, as well as with latitude and exposure. Waves generated by winter storms in the Southern Hemisphere of the Pacific Ocean may travel 10,000 km (6000 mi) before breaking on the shores of California, where they are common summer waves for the Northern Hemisphere.

The profiles of ocean waves in deep water are long and low, approaching a sinusoidal form. As the waves enter shallow water, the propagation speed and wavelength decrease, the wave steepens, and the wave height increases until the wave train consists of peaked crests separated by flat troughs. Near the breaker zone, the process of steepening is accelerated so that the breaking wave usually attains a height greater than the deep-water wave. This transformation is particularly pronounced for long-period waves from a distant storm. However, the profiles of local storm waves and the waves generated over small water bodies, such as lakes, show considerable steepness even in deep water.

Wave shoaling, that is, the shallow-water transformation of waves, commences at the depth where the waves "feel bottom." This depth is about one-half the deep-water wavelength, where the wavelength is the horizontal distance from wave crest to crest. Upon entering shallow water, waves are also subjected to refraction, a process in which the wave crests tend to parallel the depth contours, and to wave diffraction, which causes a flow of energy along the wave crest. For straight coasts with parallel contours, refraction decreases the angle between the approaching wave and the coast, and causes a spreading of the energy along the crests. The wave height is decreased by this process, but the effect is uniform along the coast (**Fig. 1**). The amount of wave refraction and diffraction and the consequent change in wave height and direction at any point along the coast is a function of wave period, direction of approach, and the configuration of the bottom topography. *See* OCEAN WAVES; WAVE MOTION IN LIQUIDS.

Nearshore circulation. When waves break so that there is an angle between the crest of the breaking wave and the beach, the momentum of the breaking wave has a component along the beach in the direction of wave propagation. This results in the generation of longshore currents that flow parallel to the beach inside the breaker zone (**Fig. 2*a***). After flowing parallel to the beach as longshore currents, the water returns seaward along relatively narrow zones as rip currents. The net onshore transport of water by wave action in the breaker zone, the lateral transport inside the breaker zone by longshore

Fig. 1. Longshore currents, generated when waves approach the beach at an angle. At Oceanside, California, the longshore current is flowing toward the observer. (*Department of Engineering, University of California, Berkeley*)

currents, the seaward return of the flow through the surf zone by rip currents, and the longshore movement in the expanding head of the rip current together constitute the nearshore circulation system. The pattern that results from this circulation commonly takes the form of an eddy or cell with a vertical axis. The dimensions of the cell are related to the width of the surfzone and the spacing between rip currents. The spacing between rip currents is usually two to eight times the width of the surfzone.

When waves break with their crests parallel to a straight beach, the flow pattern of the nearshore circulation cell becomes symmetrical (Fig. 2*b*). Longshore currents occur within each cell, but there is little longshore exchange of water or sediment from cell to cell.

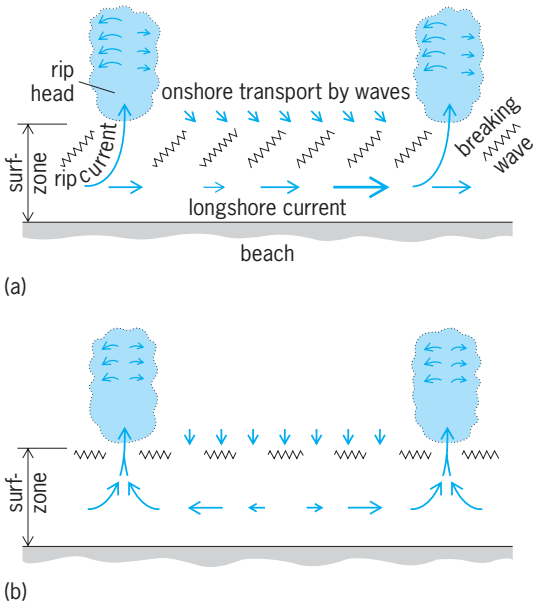

Fig. 2. Definition sketches for nearshore circulation cells. (*a*) Asymmetrical cell for breakers oblique to the shore. (*b*) Symmetrical cell for breakers parallel to the shore.

The nearshore cirulation system produces a continuous interchange between the waters of the surf and offshore zones, acting as a distributing mechanism for nutrients and as a dispersing mechanism for land runoff. Offshore water is transported into the surfzone by breaking waves, and particulate matter is filtered out on the sands of the beach face. Runoff from land and pollutants introduced into the surfzone are carried along the shore and mixed with the offshore waters by the seaward-flowing rip currents. These currents are a danger to swimmers who may be unexpectedly carried seaward. Longshore currents may attain velocities in excess of 2.5 m/s (8 ft/s), while rip current velocities in excess of 1.5 m/s (5ft/s) have been measured. Periodicity or fluctuation of current velocity and direction is a characteristic of flow in the nearshore system. This variability is primarily due to the grouping of high waves followed by low waves, a phenomenon called surf beat that gives rise to a pulsation of water level in the surfzone.

Formation of circulation cells. Nearshore circulation cells result from differences in mean water level in the surfzone associated with changes in breaker height along the beach. Waves transmit momentum in the direction of their travel, and their passage through water produces second-order pressure fields that change the mean water level near the shore. Near the breakpoint, the presence of the pressure field produces a decrease in mean water level (wave setdown) that is proportional to the square of the wave height and, for waves near the surfzone, has a maximum value that is about one-sixteenth that of the breaker height. Shoreward of the breakpoint, the onshore flow of momentum against the beach produces a rise in mean water level over the beach face, called setup.

If the wave height varies along a beach, the setup will also vary, causing a longshore gradient in mean water level within the surfzone. Longshore currents flow from regions of high water to regions of low water, and thus flow away from zones of high waves. The longshore currents flow seaward as rip currents where the breakers are lower. Pronounced changes in breaker height along beaches usually result from wave refraction over irregular offshore topography. However, on a smaller scale, uniformly spaced zones of high and low breakers occur along straight beaches with parallel offshore contours. It has been shown that these alternate zones of high and low waves are due to the interaction of the incident waves traveling toward the beach from deep water with one of the many possible modes of oscillation of the nearshore zone known as edge waves. Edge waves are trapped modes of oscillation that travel along the shore. Circulation in the nearshore cell is enhanced by edge waves having the period of the incident waves, or that of their surf beat, because these interactions produce alternate zones of

high and low breakers whose positions are stationary along the beach. It appears that the edge waves can be either standing or progressive. In either case, the spacing between zones of high waves (and hence between rip currents) is related to the wavelength of the edge wave.

Headlands, breakwaters, and piers influence the circulation pattern and alter the direction of the currents flowing along the shore. In general, these obstructions determine the position of one side of the circulation cell. In places where a relatively straight beach is terminated on the down-current side by points or other obstructions, a pronounced rip current extends seaward. During periods of large waves having a diagonal approach, these rip currents can be traced seaward for distances of 1.6 km (1 mi) or more.

Types of coasts. The geologic setting and the exposure to waves are the two most significant factors in determining nearshore processes. The large-scale features of a coast are associated with its position relative to the edges of the Earth's moving tectonic plates. Accordingly, plate tectonics provides a convenient basis for the first-order classification of coasts, with longshore dimensions of about 1000 km (600 mi). Tectonic classification leads to the definition of three general types: collision coasts, trailing-edge coasts, and marginal seacoasts. *See* PLATE TECTONICS.

Collision coasts. These coasts occur along active plate margins, where the two plates are in collision or impinging upon each other (**Fig. 3***a*). This area is one of crustal compression and consumption. These coasts are characterized by narrow continental shelves bordered by deep basins and ocean trenches. Submarine canyons cut across the narrow shelves and enter deep water. The shore is often rugged and backed by seacliffs and coastal mountain ranges; earthquakes and volcanism are common. The coastal mountains often contain elevated wave-cut platforms and sea terraces representing former relations between the level of the sea and the land. The west coasts of the Americas are typical examples of collision coasts. *See* SUBMARINE CANYON.

Trailing-edge coasts. These occur on the trailing edge of a landmass that moves with the plate (Fig. 3*b*). They are thus situated upon passive continental margins that form the stable portion of the plate, well away from the plate boundaries. The east coasts of North and South America are examples of mature, trailing-edge coasts. These coasts typically have broad continental shelves that slope into deeper water without a bordering trench. The coastal plain is also typically wide and low-lying and usually contains lagoons and barrier islands, as on the eastern coasts of the Americas. *See* BARRIER ISLANDS; COASTAL PLAIN; CONTINENTAL MARGIN.

Marginal seacoasts. These coasts develop along the shores of seas enclosed by continents and island arcs. These coasts are typically bordered by wide shelves and shallow seas. The coastal plains of marginal sea coasts vary in width and may be bordered by hills and low mountains. Rivers entering the sea along marginal seacoasts often develop extensive deltas because of the reduced intensity of wave action associated with small bodies of water. Typical marginal seacoasts border the South China and East China seas, the Sea of Okhotsk, the Mediterranean Sea, and the Gulf of Mexico. *See* BASIN; DELTA.

Other types. A complete classification would also include coasts formed by other agents, such as glacial scour, ice-push, and reef-building organisms, adding two other types of coast: cryogenic and biogenic. Common examples of these two coastal types are arctic coasts and coral reef coasts.

Beaches. Beaches consist of transient clastic material (unconsolidated fragments) that reposes near the interface between the land and the sea and is subject to wave action. The material is in dynamic repose rather than in a stable deposit, and thus the width and thickness of beaches is subject to rapid fluctuations, depending upon the amount and rigor of erosion and transportation of beach material. Beaches along collision coasts are essentially long rivers of sand that are moved by waves and currents and are derived from the material eroded from the coast and brought to the sea by streams. The coast may be cliffed (Fig. 3*a*), or it may contain a ridge of wind-blown sand dunes and be backed by marshes and water (Fig. 3*b*). Along many low sandy coasts, such as the east and Gulf coasts of the United States, the beach is separated from the mainland by water or by

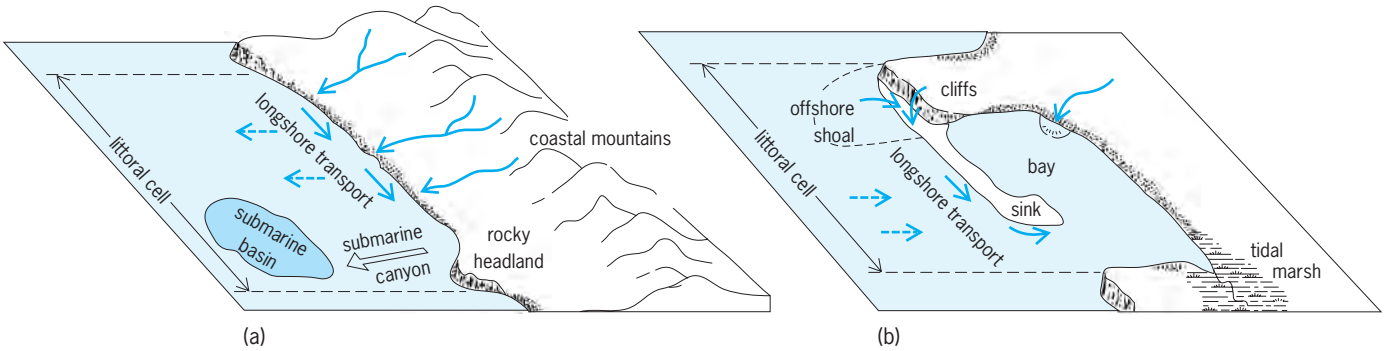

Fig. 3. Typical (*a*) collision and (*b*) trailing-edge coasts and their littoral cells. Solid arrows show sediment transport paths; dotted arrows indicate occasional onshore and offshore transport modes. (*After D. L. Inman, Types of coastal zones: Similarities and differences, in National Research Council, Environmental Science in the Coastal Zone, National Academy Press, 1994*)

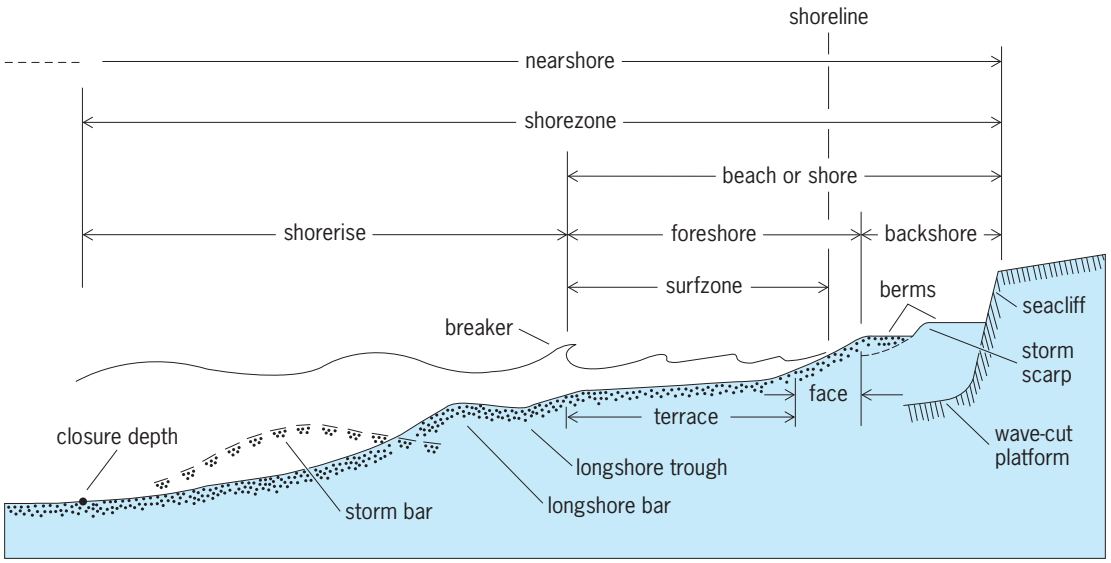

Fig. 4. Beach profile, showing characteristic features.

a natural coastal canal. Such beaches are called barrier islands. Barrier beaches are "braided" forms of the "river of sand" with transgressive rollover caused by sea-level rise and washover processes. A beach that extends from land and terminates in open water is referred to as a spit, while a beach that connects an island or rock to the mainland or another island is a tombolo.

Beach nomenclature. While differing in detail, beaches worldwide have certain characteristic features which allow application of a general terminology to their profile (**Fig. 4**). The beach or shore extends landward from the breakpoint bar to the effective limit of attack by storm waves. The region seaward is termed the shorerise; that landward is the coast. The coast is part of the coastal zone, which includes the continental shelf and the coast. The beach includes a backshore and foreshore. The backshore is the highest portion and is acted upon by waves only during storms. The foreshore extends from the crest of the berm to the breakpoint bar and is the active portion of the beach traversed by the broken waves and the uprush and backwash of the swash on the beach face. The foreshore consists of a steep seaward-dipping face, related to the size of the beach material and the rigor of the uprush, and of a more gentle seaward terrace, sometimes referred to as the low-tide terrace, over which the waves break and surge. In most localities the foreshore face and terrace merge into one continuous curve; in others there is a discontinuity at the toe of the beach face. The former condition is characteristic of fine sand beaches and of coasts where the wave height is equal to or greater than the tidal range. The latter is typical of coastlines where the tidal range is large compared with the wave height, as along the Patagonian coast of South America and portions of the Gulf of California. The foreshore frequently contains one or more bars and troughs that parallel the beach; these are referred to as longshore bars and longshore troughs. Longshore bars commonly form at the plunge point of the wave, and their position is thus influenced by the breaker height and the nature of the tidal fluctuation. *See* TIDE.

The shorerise is the transition between the continental shelf and the beach, and is marked by the increase in slope leading from the gently sloping shelf up to the beach proper. The shorerise extends seaward from the breakpoint bar to the closure depth that marks the seaward extent of depth changes between winter (storm) and summer beach profiles (Fig. 4). Together, the shorerise, foreshore, and backshore comprise the shorezone, which is the zone of active transport of beach material and the resulting areas of beach accretion and erosion.

Beach cycles. Waves are effective in causing sand to be transported laterally along the beach by longshore currents and in causing movements of sand from the beach to the shorerise and back again to the beach. Although these two types of transport are interrelated, for convenience longshore movement of sand is discussed separately.

Along most coasts, there is a seasonal migration of sand between the beaches and the shorerise in response to the changes in the character and direction of approach of the waves. In general, the beach face builds seaward during the small waves of summer and is cut back by high storm waves in winter. There are also shorter cycles of cut-and-fill associated with spring and neap tides and with nonseasonal waves and storms. Bottom surveys indicate that most offshore-onshore interchange of sand occurs in depths of 10–15 m (33-45 ft) but that some effects may extend to depths of 30 m (100 ft) or more.

A typical summer beach is built seaward by low waves (Fig. 4). During stormy seasons, the beach face is eroded, sometimes forming a beach scarp. Subsequent low waves build the beach face seaward again. The beach face is a depositional feature, and its highest point, the berm crest, represents the maximum height of the runup of water on the beach. The height of wave runup above still-water level is

about equal to the height of the breaking wave. Since the height of the berm depends on wave height, the higher berm, if it is present, is sometimes referred to as the winter or storm berm, and the lower berm as the summer berm. The entire beach may be cut back to the underlying country rock during severe storms. Under such conditions, the waves erode the coast and form sea cliffs and wave-cut platforms. These features and their terrace deposits are frequently preserved in the geologic record and serve as markers for the past relations between the levels of the sea and the land.

Mechanics of beach formation. Beaches form wherever there are waves and an adequate supply of sand or coarser material. Even anthropogenic fills and structures are effectively eroded and reformed by the waves. The initial event in the formation of a new beach from a heterogeneous sediment is the sorting of the material, with coarse material remaining on the beach and fine material being carried away. Concurrent with the sorting action, the material is rearranged, some being piled high above the water level by the runup of the waves to form the beach berm, some carried back down the face to form the foreshore terrace. In a relatively short time, the beach assumes a profile that is in equilibrium with the forces generating it.

The beach face is frequently characterized by laminations (closely spaced layers) that show slight differences in the size, shape, or density of the sand grains. The laminations parallel the beach face and represent shear sorting within the granular load as it is transported by the swash and backwash of the waves over the beach face. Detailed examination shows that each lamina of fine-grained minerals delineates the plane of shear between moving and residual sand. The mechanism concentrating the heavy fine grains at the shear plane is partly the effect of gravity acting during shearing, causing the small grains to work their way down through the interstices between larger grains, and partly the dependence of the normal dispersive pressure upon grain size. When grains are sheared, the normal dispersive pressure between grains, which varies as the square of the grain diameter, causes large grains to drift toward the zone of least shear strain, that is, the free surface, and the smaller grains toward the zone of greatest shear strain, that is, the shear plane.

Equilibrium profile. The action of waves on a sloping beach eventually produces a profile that achieves equilibrium with the energy dissipation associated with the oscillatory motion of the waves over the sand bottom. The slope of the beach face is related to the dissipation of energy by the swash and backwash over the beach face. Percolation of the swash into a permeable beach reduces the amount of flow in the backwash and is thus conducive to deposition of the sand transported by the swash. If in addition the beach is dry, this action is accentuated. Coarse sands are more permeable and consequently more conducive to deposition and the formation of steep beach faces. Large waves elevate the water table in the beach. When the beach is saturated, the backwash has a higher velocity, a condition conducive to erosion. From the foregoing it follows that the slope of the beach foreshore increases with increasing sediment size and with decreasing wave height. If an artificial slope exceeds the natural equilibrium slope, an offshore transport of sand will result from the gravity component of the sand load until the slope reaches equilibrium. Conversely, if an artificial slope is less than the natural equilibrium slope, a shoreward transport of sand by waves and currents will result, and the beach slope will steepen. An equilibrium slope is attained when the up-slope and downslope transports are equal.

Beaches respond to wave-forcing by adjusting their form to an equilibrium or constant shape attributable to a given type of incident wave. The seasonal changes in beach profile in response to the high waves of winter and the lower waves of summer are expressions of the beach form tending toward a seasonal equilibrium with the changing character of the prevailing waves. Field studies show that the equilibrium beach consists of two conjoined parabolic curves that intersect at the breakpoint bar, one curve for the shorerise that extends from the closure depth to the breakpoint bar, and another for the foreshore that extends from the bar to the berm (Fig. 4). The principal differences between seasonal profiles are that in winter (higher waves) the breakpoint bar is deeper and farther offshore, while the berm crest is displaced landward. Thus, the changes in seasonal equilibria are manifest by simple, self-similar displacements of the bar-berm and shorerise curves. Also, since the equilibrium profile is parabolic, for any given sand size, the slope increases from the bar to the berm.

Longshore movement of sand. The movement of sand along the shore occurs in the form of bed load (material rolled and dragged along the bottom) and suspended load (material stirred up and carried with the current). Suspended-load transportation occurs primarily in the surf zone, where the turbulence and vertical mixing of water are most effective in placing sand in suspension and where the longshore currents that transport the sediment-laden waters have the highest velocity. The longshore transport rate of sand is directly proportional to the longshore component of wave power. Thus, the longshore transport rate of sand can be estimated from a knowledge of the wave climate, that is, the budget of wave energy incident upon the beach.

The volume of littoral transport along oceanic coasts is usually estimated from the observed rates of erosion or accretion, most often in the vicinity of natural obstructions, such as headlands (**Fig. 5**), or of coastal engineering structures, such as groins or jetties. In general, beaches build seaward up-current from obstructions and are eroded on the current lee where the supply of sand is diminished. Such observations indicate that the transport rate varies from almost nothing to several million cubic meters per year, with average values of 150,000–600,000 m^3

Fig. 5. Effect of headlands on the accretion of beach sand at Point Mugu, California. The point forms a natural obstruction that interrupts the longshore transport of sand, causing accretion and a wide beach to form (foreground). The regularly spaced scallops are swash cusps on the beach face. (*Department of Engineering, University of California, Berkeley*)

(200,000–800,000 yd^3) per year. Along the shores of smaller bodies of water, such as the Great Lakes of the United States, the littoral transport rate can be expected to range about 7000–150,000 m^3 (9000–200,000 yd^3) per year.

The large quantity of sand moved along the shore and the pattern of accretion and erosion that occurs when the flow is interrupted pose serious problems for coastal engineers. The problem is particularly acute when jetties are constructed to stabilize and maintain deep navigation channels through sandy beaches. A common remedial procedure is to dredge sand periodically from the accretion on the up-current side of the obstruction and deposit it on the eroding beaches in the current lee. Another method is the installation of sand bypassing systems, which continually remove the accreting sand and transport it by hydraulic pipeline to the beaches in the lee of the obstruction. *See* COASTAL ENGINEERING; SEDIMENTOLOGY.

Sources and sinks of beach sediment. The principal sources of beach and nearshore sediments are the rivers that bring quantities of sand directly to the ocean; the seacliffs and blufflands of unconsolidated material that are eroded by waves; and material of biogenous origin (shell and coral fragments and skeletons of small marine animals). In places, sediment may be supplied by the erosion of unconsolidated deposits in shallow water (Fig. 3*b*). Beach sediments on the coasts of the Netherlands are derived in part from the shallow waters of the North Sea. Windblown sand may be a source of beach sediment, although winds are usually more effective in removing sand from beaches than in supplying it. In tropical latitudes, many beaches are composed entirely of grains of calcium carbonate of biogenous origin. Generally the material consists of fragments of shells, corals, and calcareous algae growing on or near fringing reefs. The material is carried to the beach by wave action over the reef. Some beaches are composed mainly of the tests (shells) of foraminifera that live on sandy bottom offshore from the reefs. *See* NORTH SEA; REEF.

Streams and rivers may be important sources of sand for beaches in temperate latitudes (Fig. 3*a*). Surprisingly, the contribution of sand by streams in arid climates is quite high (**Fig. 6**). Arid weathering produces sand-size material from areas with a minimum cover of vegetation, so that occasional flash floods may transport large volumes of sand. The maximum sediment yield occurs from drainage basins where the mean annual precipitation is about 30 cm (12 in.) per year.

There is a pronounced multidecadal variability in the amount of river-borne sediment transported to the beach. The variability is associated with global climate changes related to the El Niño/Southern Oscillation (ENSO) phenomena. ENSO drives large-scale events such as the Pacific/North American (PNA) patterns of atmospheric pressure that lead to wet and dry climate along the Pacific coast of North America. The 20 coastal rivers of central and southern California had streamflow and sediment fluxes during the wet phase of PNA (1969–1998) that exceed those during the preceding dry phase (1944–1968) by factors of 3 and 5, respectively. The sediment flux during the three major El Niño events of the wet phase were on average 27 times greater than the annual sediment flux during the dry phase. Also, the wave climate in southern California changed with the shift from dry to wet phase of PNA. The prevailing northwesterly winter waves of the dry phase

Fig. 6. Sand delta at Rio de la Concepción on the arid coast of the Gulf of California. Such deltas are important sources of sand for beaches. (*Courtesy of D. L. Inman*)

were replaced by high-energy waves approaching from the west or southwest during the wet phase. Wave climate along the east coast of North America responds to shifts in the atmospheric patterns of the North Atlantic Oscillation (NAO), which is generally out of phase with the west coast climate. *See* CLIMATIC PREDICTION; EL NIÑO; TROPICAL METEOROLOGY; WEATHER FORECASTING AND PREDICTION.

Wave erosion of rocky coasts is usually slow, even where the rocks are relatively soft shales. Therefore, cliff erosion usually does not account for more than about 10% of the material on most beaches. However, retreats greater than 1 m (3.3 ft) per year are not uncommon in unconsolidated seacliffs. The most dramatic modern example of coastal erosion is found along the delta of the Nile River in Egypt. The High Aswan Dam, constructed in 1964, has intercepted the sediment that was previously brought down the Nile to the coast. Lacking source material, the delta coast is eroding, and waves and currents may cause an entire city block of Ras El Bar to be lost to the sea in one year.

The sand carried along the coast by waves and longshore currents may be deposited in continental embayments, or it may be diverted to deeper water by submarine canyons which traverse the continental shelf and effectively tap the supply of sand (Fig. 3*a*). Most of the deep sediments on the abyssal plains along a 400-km (250-mi) section of the California coastline are probably derived from two submarine canyons, Delgada Submarine Canyon in northern California and Monterey Submarine Canyon in central California. *See* MARINE SEDIMENTS.

Littoral cells and the budget of sediment. The budget of sediment for a region is obtained by assessing the sedimentary contributions and losses to the region and their relation to the various sediment sources and transport mechanisms. Determination of the budget of sediment is not a simple matter, since it requires knowledge of the rates of erosion and deposition as well as understanding of the capacity of various transport agents. Studies of the budget of sediment show that coastal areas can be divided into a series of discrete sedimentation compartments called littoral cells. Each cell contains a complete cycle of littoral transportation and sedimentation, including transport paths and sources and sinks of sediment. Littoral cells take a variety of forms, but there are two basic types. One is characteristic of collision coasts with submarine canyons (Fig. 3*a*), while the other is more typical of trailing-edge coasts where rivers empty into large estuaries (Fig. 3*b*). The concept of a littoral cell (or a subcell) with its budget of sediment and transport paths provides objective criteria for making choices among various coastal conservation methods. *See* ESTUARINE OCEANOGRAPHY.

Biological effects. The rigor of wave action and the continually shifting substrate make the sand beach a unique biological environment. Because few large plants can survive, the beach is occupied mostly by animals and microscopic plants. Much of the food supply for the animals consists of particulate matter that is brought to the beach by the nearshore circulation system and is trapped in the sand. The beach acts as a giant sand filter straining out particulate matter from the water that percolates through the beach face.

Since the beach-forming processes and the trapping of material by currents and sand are much the same everywhere, the animals found on sand beaches throughout the world are similar in aspects and habits, although different species are present in different localities. In addition, since the slopes and other physical properties of beaches are closely related to elevation, the sea animals also exhibit a marked horizontal zonation. Organisms on the active portion of the beach face tend to be of two general types insofar as the procurement of food is concerned: those that burrow into the sand, using it for refuge while they filter particulate matter from the water through siphons or other appendages that protrude above the sandy bottom, and those that remove organic material from the surface of the sand grains by ingesting them or by "licking." There are usually few species, which may be very abundant.

In tropical seas, the entire shore may be composed of the cemented and interlocking skeletons of reef-building corals and calcareous algae. When this occurs, the nearshore current system is controlled by the configuration of the reefs. Where there are fringing reefs, breaking waves carry water over the edge of the reef, generating currents that flow along the shore inside of the reef and then flow back to sea through deep channels between reefs. Under such conditions, beaches are usually restricted to a berm and foreshore face bordering the shoreward edge of the reef.

Douglas L. Inman

Bibliography. W. Bascom, *Waves and Beaches: The Dynamics of the Ocean Surface*, rev. ed., 1980; D. C. Conley and D. L. Inman, Field observations of the fluid-granular boundary layer under near-breaking waves, *J. Geophys. Res.*, 97(C6):9631–43, 1992; R. A. Davis, Jr., *The Evolving Coast*, 1994; K. Horikawa (ed.), *Nearshore Dynamics and Coastal Processes*, 1988; D. L. Inman and B. M. Brush, The coastal challenge, *Science*, 181:20–32, 1973; D. L. Inman and R. Dolan, The Outer Banks of North Carolina: Budget of sediment and inlet dynamics along a migrating barrier system, *J. Coastal Res.*, 5(2):193–237, 1989; D. L. Inman, M. H. S. Elwany, and S. A. Jenkins, Shorerise and bar-berm profiles on ocean beaches, *J. Geophys. Res.*, 9(C10):18,181–18,199, 1993; D. L. Inman and S. A. Jenkins, Climate change and the episodicity of sediment flux of small California rivers, *J. Geol.*, 107(3):251–70, 1999; J. P. Kennett, *Marine Geology*, 1982; P. D. Komar, *Beach Processes and Sedimentation*, rev. ed., 1998; National Research Council, *Environmental Science in the Coastal Zone: Issues for Further Research*, 1994; R. J. Seymour (ed.), *Nearshore Sediment Transport*, 1989; U.S. Army Corps of Engineers, *Engineering and Design—Coastal Littoral Transport*; L. D. Wright, *Morphodynamics of Inner Continental Shelves*, 1995.

Nebula

The term "nebula" was originally used to refer to any fixed, extended, and usually fuzzy luminous celestial object. With increased angular resolution of telescopes, astronomers learned that nebulae can be separated into two classes: those that are stellar systems made up of individual stars, and those that are gaseous in nature and diffuse in appearance. Examples of stellar systems include galaxies (which contain billions of stars and are located outside our own Milky Way Galaxy) and star clusters such as open clusters and globular clusters (which contain thousands of stars and are within the Milky Way Galaxy). This article is restricted to the modern definition of nebulae, which are gaseous objects usually located within the Milky Way Galaxy, although with increasingly powerful telescopes gaseous nebulae can now be observed in external galaxies. *See* GALAXY, EXTERNAL; MILKY WAY GALAXY; STAR.

Because of space observations, the study of gaseous nebulae has undergone a renaissance. The extension of observations into the x-ray, ultraviolet, infrared, millimeter, and submillimeter wavelengths has revealed a much richer makeup of nebulae, containing almost all states of matter. Although the term "gaseous nebulae" is used, in fact they consist of ions, atoms, molecules, and solid particles. Sometimes the term "clouds" is also used to refer to concentrations of interstellar matter (for example, dark clouds, molecular clouds); there is, however, no precise difference in meaning between "nebulae" and "clouds."

Types. Gaseous nebulae can be divided into three main types. Those that radiate brightly in the visible are called emission nebulae. Those that are detected through their effects on the obscuration of background stars are called dark nebulae or absorption nebulae. There is also a class of nebulae that do not self-radiate in visible light but reflect light from nearby stars, and those are called reflection nebulae. Examples of emission nebulae include the Great Orion Nebula, which is a region of active star formation. Other examples of emission nebulae are planetary nebulae and supernova remnants, which are objects associated with dying stars. *See* ORION NEBULA.

Nebulae associated with star formation. New stars are formed out of concentrations of dust and gas in the interstellar medium. New stars often form in groups. For clusters of newborn stars containing hot stars, the ultraviolet light from the stars is able to ionize the surrounding gas, making them luminous. In astronomical nomenclature, they are called H II regions, refering to the fact that the hydrogen atoms in the nebulae are in ionized form, with the electron detached from the proton (**Colorplate 1**).

Planetary nebulae. Planetary nebulae are so denoted because they often resemble small greenish disks when seen in small telescopes, not unlike the images of the planets Uranus and Neptune. Well-known examples of this class are the Ring Nebula (NGC 6720 or M57) in Lyra, the Helix Nebula (NGC 7293) in Aquarius, and the Dumbbell Nebula (M27) in Vulpecula. Unlike H II regions, which are associated with the birth of stars, planetary nebulae are products of stars near their deaths. They are ejected by very old red giants and contain fractions of a solar mass of material. Since they are continuously expanding, they have short lifetimes lasting only tens of thousands of years. *See* GIANT STAR.

Planetary nebulae derive their power from their central stars. After the ejection, the remnant of the original red giant becomes a hot, blue star that eventually evolves to become a white dwarf. The ultraviolet light from the star can ionize the ejecta, leading to the emission of a variety of atomic lines in different colors. Planetary nebulae are colorful objects and can be considered as neon signs in the sky. They have well-defined morphologies (often shell-like and sometimes bipolar) and have sizes of fractions of a light-year (**Colorplate 2**). *See* PLANETARY NEBULA; WHITE DWARF STAR.

Supernova remnants. The detonation of a star in a supernova event causes the ejection of the outer layers into the surrounding interstellar medium. In early stages, as in the Crab Nebula, the radiating material consists of ejecta from the star. In the later stages, this rapidly moving material is slowed down as it mixes with the surrounding dust and gas of the interstellar medium. Heating by shock waves causes the material to radiate optically. Supernova remnants are also characterized by their x-ray emission arising from the rarefied gas behind the shock front, and by synchrotron radiation at radio frequencies. Since both planetary nebulae and supernova remnants are stellar ejecta and surround their parent stars, they are referred to as circumstellar nebulae to distinguish them from H II regions, which are interstellar in origin. *See* CRAB NEBULA; SUPERNOVA.

Reflection nebulae. In contrast to emission nebulae, reflection nebulae do not shine on their own in the visible. The nebular structure is derived from starlight scattered off dust grains in the nebula. As a result, the optical spectrum of a reflection nebula is made up of the continuous spectrum of the illuminating stars. However, the color of a reflection nebula is often bluer than the stellar spectrum because the dust particles scatter more efficiently in the blue. The best-known example is the Pleiades reflection nebula, which is illuminated by the brightest stars in the Pleiades cluster. *See* PLEIADES.

Typical reflection nebulae have sizes of fractions of a light-year, with densities of $\sim 10^4$ atoms per cubic centimeter. The central stars are generally cooler than 20,000 K (36,000°F), not hot enough to photoionize the nebula. The Trifid Nebula (M20) is an example of an object which is part emission and part reflection nebula. In Colorplate 1, the red part of the nebula is an emission nebula resulting from H-alpha emission, whereas the blue part is a reflection nebula resulting from starlight scattered off dust.

Dark clouds. Dark clouds represent a concentration of solid dust particles that obscure starlight. By preventing light from stars located behind the cloud from reaching us, they appear as dark patches in the

sky. Many dark clouds have no strong heating sources in the form of central stars, but are heated primarily by external nearby stars or cosmic rays. For this reason, the gas temperature inside dark clouds can be very low (about 10 K or −442°F). On a smaller scale, there are compact, dense clouds known as Bok globules. They are often spherical in shape and have densities higher than dark clouds. *See* GLOBULE.

Although dark in the visible light, dark clouds are luminous in the infrared. The energy that they absorb from starlight is converted into infrared emission, and they can be detected in emission by modern infrared and submillimeter-wave telescopes.

Molecular clouds. Many dark clouds also contain molecules that can be detected through their rotational transitions in the millimeter wavelengths. Interstellar clouds that are heavy in molecular content are called molecular clouds. Molecular clouds that are bright in the infrared and molecular line emissions are called giant molecular clouds. Their dust and gas temperatures, inferred from infrared continuum and molecular-line emissions respectively, are in the range of 50 to 100 K (−370 to −280°F). Such relatively high temperatures suggest that there must be heating sources inside the cloud. These internal sources can be identified by near-infrared imaging and are believed to be newly formed massive (OB) stars. The inferred masses (10^3 to 10^6 solar masses) of giant molecular clouds suggest that they are not gravitationally stable, and kinematic studies using molecular lines confirm that many are in the process of gravitational collapse. Giant molecular clouds are therefore excellent objects to study the formation process of massive stars. *See* MOLECULAR CLOUD.

Discovery and catalogs. The first catalog of nebulae was made by Charles Messier in 1784. A much more complete list was made by J. L. E. Dreyer in the *New General Catalogue* (*NGC*), published in 1888, and the two *Index Catalogs*. After the completion of the National Geographic Society–Palomar Observatory Sky Survey, many large and faint nebulae were identified by George Abell from the photographic plates of the survey. Since these catalogs were compiled based on the morphology of the objects, the lists include galaxies and star clusters in addition to gaseous nebulae. *See* ASTRONOMICAL CATALOGS; MESSIER CATALOG.

A modern definition of nebulae does not rely solely on morphological appearance but also on spectroscopic characteristics. Specifically, an emission nebula is defined by its emission-line spectrum, and an object can be classified as a gaseous nebula even though its angular size is so small that it is not seen as an extended object by a telescope. For example, by employing objective prisms in telescopes, astronomers in the 1960s were able to distinguish emission nebulae from stars. A more modern technique uses filters covering a narrow range of color, for example, centering on the H-alpha line of hydrogen at 656.2 nm or the 500.7-nm line of doubly ionized oxygen (O^{2+}). By comparing images taken with these narrow filters to those with broad filters, emission nebulae will stand out as bright objects. These techniques allow the identification of small gaseous nebulae that appear stellar in photographical images. *See* ASTRONOMICAL SPECTROSCOPY; INTERFERENCE FILTER.

H II regions, planetary nebulae, and supernova remnants are radio emitters, and these objects can be found through radio surveys. By making measurements at several frequencies, the spectral shape of the emitting radio source can be determined, therefore allowing the distinction between H II regions and supernova remnants, for example.

Since gaseous nebulae are dusty and radiate strongly in the infrared, their existence can be found through infrared sky surveys. The most notable infrared sky surveys are the Infrared Astronomical Satellite survey, which covered 96% of the sky at wavelengths of 12, 25, 60, and 100 μm; the 2MASS survey at 2 μm; and the Spitzer Space Telescope Galactic Legacy Infrared Mid-Plane Survey Extraordinaire (GLIMPSE) survey of the galactic plane at 3.6, 4.5, 5.8, and 8.0 μm. Nebulae that have so much dust that their visible light output is obscured by their own internal dust are most efficiently discovered in infrared surveys.

In addition to visual inspection of photographic plates, dark nebulae can be found using the technique of star counts. If an area in the sky has a lower density of stars compared to its neighboring region, it probably implies the presence of a dark nebula.

Composition. Gaseous nebulae are made up of free-flying atoms in the gaseous state. These atoms can be in the neutral (uncharged) state, or in an ionized state where one or more of the electrons of the atom have been removed from the atom. The ionized state of gas is also referred to as the plasma state. Since hydrogen is the most common element in the universe, gaseous nebulae consist of primarily hydrogen, with helium second in abundance, and then oxygen, carbon, and other heavy elements. *See* PLASMA (PHYSICS).

Molecular gas is also present in nebulae. Although most molecules consist of two, three, or four atoms [for example, molecular hydrogen (H_2) and carbon monoxide (CO)], very complex molecules with over a dozen atoms have also been seen. The molecular gas state is similar to the gaseous state in the Earth's atmosphere, where molecules of nitrogen (N_2) and oxygen (O_2) are the dominant species.

Also present in nebulae are solid-state particles, which are generally referred to as "dust" in astronomical literature. These are macroscopic particles with sizes up to about 1 μm, each consisting of thousands of atoms. The most common inorganic dust consists of silicates, which are made up of atoms of oxygen, silicon, magnesium, and iron. Organic carbonaceous dust, consisting of primarily carbon and hydrogen, is also widely seen. *See* INTERSTELLAR MATTER.

Spectral characteristics. The different states of matter in a nebula can be traced by their different radiation mechanisms. Atoms and ions can be identified by their atomic electronic transitions, molecules by

their vibrational or rotational transitions, and solid particles by their lattice vibrational modes. Since different atoms radiate in different specific frequencies, they can be uniquely identified by spectroscopic observations.

The densities in gaseous nebulae are very low in comparison to the Earth's atmosphere. In fact, the densities are lower than the best vacuum that can be achieved in the laboratory. Consequently, the excitation of atoms and molecules and their subsequent radiation are very different from usual mechanisms observed in the terrestrial environment.

Atoms and ions. The electronic transitions of atoms and ions occur in the ultraviolet, visible, and infrared parts of the electromagnetic spectrum. These are called bound-bound transitions, as they are the consequence of an electron going from one bound state to another. The vibrational transitions of molecules generally occur in the infrared, while the rotational transitions occur in the millimeter or submillimeter parts of the spectrum. Atomic electronic transitions observed in nebulae can be separated into two kinds. The first is the cascade of electrons from a higher state to a lower state after a free electron has been captured by the nucleus of an ion. These are called recombination lines. Usually, only the recombination lines of hydrogen and helium are bright enough to be detected, although weaker recombination lines such as those from oxygen, carbon, and nitrogen have been observed. The second class of atomic transitions occurs as the result of excitation of a bound electron from the ground state to a higher state through collisions with a free electron. These are called collisionally excited lines. Because the recombination lines of heavy elements are weak, most chemical elements are detected in gaseous nebulae through their collisionally excited lines. Because of improvements in sensitivity made possible by the use of charge-coupled devices (CCDs), emission lines from elements in rows 4, 5, or 6 of the periodic table (for example, selenium, krypton, and xenon) have been detected. The observation and measurement of abundance of heavy elements are very important in understanding the origin of chemical elements and the nuclear reactions responsible for the synthesis of these elements. *See* ATOMIC STRUCTURE AND SPECTRA.

One interesting aspect of collisionally excited lines is the detection of "forbidden lines." These are atomic transitions that occur too slowly to be seen in the terrestrial environment. Because of the high density of the Earths atmosphere, atoms in an excited state that decay slowly (called metastable states) are usually deexcited by collisions before they have a chance to radiate. However, in the interstellar medium the density is very low and collisions are infrequent. Atoms can spend hours in an excited state without encountering another electron. This allows the appearance of "forbidden lines." One of the strongest forbidden lines in the visible is the green line of O^{2+}. This line at the wavelength of 501 nm is so strong that it can be detected from gaseous nebulae in very distant galaxies. *See* SELECTION RULES (PHYSICS).

A more unusual, third kind of atomic transition involves the fluorescent mechanism in which an atom (for example, doubly ionized oxygen) is excited by a coincident strong transition from an abundant element (for example, in the case of ionized oxygen, by an ultraviolet transition of helium at 30.378 nm). The cascade of the atom to lower energy levels results in emission lines in the visible. *See* FLUORESCENCE.

Ionized gas also radiates continuously over a broad frequency range in the visible and radio wavelengths. When an electron recombines with an atomic nucleus (usually hydrogen), it radiates away its surplus energy in visible light. This is called bound-free radiation, as it refers to a free electron becoming bound. In the radio, two kinds of radiation mechanism exist. One occurs as the result of a free electron passing near the electric field of a positively charged nucleus, and is called the free-free process. Another is synchrotron radiation, which is generated by fast (near the speed of light) electrons moving in a magnetic field. Although these continuous radiations have no characteristic frequency signatures as does a bound-bound transition, they can be identified through the shape of the continuous spectrum. For example, synchrotron radiation generally is stronger at low frequencies than at high frequencies, whereas the opposite is true for free-free radiation. *See* SYNCHROTRON RADIATION.

Because different gaseous nebulae have different physical conditions, they have different radiation characteristics. Supernova remnants generally display strong radio synchrotron radiation, whereas H II regions and planetary nebulae show only free-free radiation.

Molecules. Molecules are generally excited in two ways, either by collisional excitation with hydrogen molecules or by radiative excitation by continuous infrared light emitted by nearby dust particles. Collisional excitations change the molecules from the ground rotational state to a higher rotational state, whose subsequent decay results in radiation in the millimeter or submillimeter parts of the spectrum. The most commonly seen molecular rotational emission line is the transition of carbon monoxide (CO) from the first excited state ($J = 1$) to the ground state ($J = 0$) at a wavelength of 2.6 mm.

In addition to emission, molecules can be detected through their absorption lines. Against a bright infrared background, the abundance of the CO molecule, for example, can be determined through their transition from the ground state ($v = 0$) to the second excited vibrational state ($v = 2$) at the wavelength of 2.3 μm. *See* MOLECULAR STRUCTURE AND SPECTRA.

Solid particles. The lattice vibrational modes of solids allow the identification of different kinds of solid-state particles in nebulae. The most widely observed spectral signature is the 10- and 18-μm feature of amorphous silicates. In the infrared, a series of emission features at 3.3, 6.2, 7.7, 8.6, and 11.3 μm

corresponding to the stretching and bending modes of aromatic compounds can be seen. The presence of these features suggests that complex organic compounds are common in nebulae. *See* LATTICE VIBRATIONS.

Observing techniques. The observation of gaseous nebulae can be categorized in two areas: imaging and spectroscopy. Imaging observations involve taking a picture of the nebulae at specific colors. Unlike stars, which radiate continuously over all colors, gaseous nebulae radiate discrete colors corresponding to the wavelengths of atomic lines. Images are often obtained using filters that admit light only over a narrow range of colors in order to enhance the quality of the image. Spectroscopic observations are able to separate the colors into fine intervals and are therefore the desired tool to study atomic or molecular line emissions from nebulae.

Because optical astronomy was the first observational technique to develop, observations of gaseous nebulae in the early twentieth century were restricted to the optical parts of the spectrum. Images of nebulae were obtained by photographic techniques, and spectra were obtained with different kinds of spectrometers (for example, with gratings). After the 1980s, photographic plates were replaced by CCDs, which have much higher sensitivity and dynamic range than photographic plates. *See* ASTRONOMICAL IMAGING; CHARGE-COUPLED DEVICES.

Optical spectroscopy of nebulae began in the late nineteenth century and concentrated on the observation of optical recombination lines of hydrogen [for example, the H-alpha ($n = 3$-2) and H-beta ($n = 4$-2) lines] and collisionally excited forbidden lines of oxygen, nitrogen, sulfur, and so forth. Modern optical spectroscopy based on the use of CCDs is capable of very high spectral resolution and can detect thousands of atomic lines in a single spectrum.

In the late 1950s, radio astronomy began to develop and gaseous nebulae were detected through their free-free and synchrotron radiations. High-level recombination lines of hydrogen were also detected in the radio. As radio receivers capable of spectroscopic observations progressed gradually to frequencies higher than 1 GHz (or equivalently wavelengths shorter than 30 cm), molecules such as OH (with characteristic emission at a wavelength of 18 cm), water (at 1.3 cm), and formaldehyde (at 6.2 cm) were detected. By the early 1970s, high-frequency receivers reaching millimeter wavelengths made possible the detection of CO (at 2.6 mm) and many other gaseous molecules. The number of molecular species detected as of 2006 exceeds 130. *See* RADIO ASTRONOMY.

The development of infrared detectors in the late 1960s made possible the detection of solid-state particles. The detection of the 10-μm feature of amorphous silicates in 1969 represents the first positive identification of specific solids in space. The launch of the *Infrared Space Observatory* in 1995 made a significant advance in astronomical infrared spectroscopy, allowing the detection of many molecular and solid-state transitions. The detection of crystalline silicates and other refractory oxides led to the new discipline of astromineralogy. The observation of the stretching and bending modes of organic compounds led to the realization that complex organic materials are commonly present in circumstellar and interstellar nebulae. With its high resolving power, the *Spitzer Space Telescope* has obtained detailed images of many nebulae in the infrared, showing their structure and morphology through their dust emission. *See* INFRARED ASTRONOMY; SPITZER SPACE TELESCOPE.

Spectroscopic observations from space-based telescopes in the ultraviolet (for example, the *International Ultraviolet Explorer* and the *Far Ultraviolet Spectroscopic Explorer*) and x-ray (for example, the *Chandra X-ray Observatory*) extend the observations of electronic transitions of atoms and ions to the ultraviolet and x-ray parts of the spectrum. Examples include the 103-nm line of O^{5+}, the 191-nm line of C^{2+}, and the 124-nm line of N^{4+}. *See* CHANDRA X-RAY OBSERVATORY; ULTRAVIOLET ASTRONOMY; X-RAY ASTRONOMY.

Many common neutral atoms in gaseous nebulae radiate only in the far-infrared or submillimeter parts of the spectrum. These transitions arise from the lowest electronic energy states and are called fine-structure lines. For example, the fine-structure transitions of neutral carbon are at 370 and 609 μm, and of neutral oxygen at 52 and 88 μm. These lines probe the neutral states of matter in nebulae and are important for understanding the chemistry of the nebulae. Since the Earth's atmosphere is opaque at far-infrared wavelengths, these transitions can be detected only from space or high-flying aircraft. The 609-μm line of neutral carbon was first detected by the Kuiper Airborne Observatory. Atomic fine-structure lines are extensively studied by the *Infrared Space Observatory*, and high-quality data are expected from the *Herschel Space Observatory* mission. *See* FINE STRUCTURE (SPECTRAL LINES); SUBMILLIMETER ASTRONOMY.

At the radio or millimeter parts of the spectrum, the technique of interferometry is often used to produce an image. Signals from two or more antennas taken over a period of time are combined electronically to form an image. For example, the Very Large Array in New Mexico has twenty-seven 25-m (82-ft) antennas operating at centimeter wavelengths and is capable of obtaining very high angular resolution images of nebulae (see **illustration**). The *Atacama Large Millimeter Array* (*ALMA*) is expected to produce high-quality images of molecular and dust emissions from interstellar clouds.

Internal dynamics. Nebulae often undergo systematic internal motions. Molecular clouds associated with star formation are mostly collapsing, whereas planetary nebulae and supernova remnants are expanding. Such internal motions can be measured through the Doppler effect, where the atomic and molecular lines shift in wavelength depending on

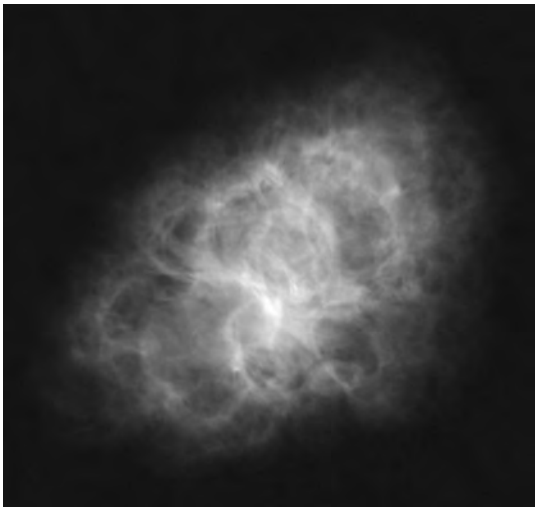

Crab Nebula, a supernova remnant in the constellation Taurus, imaged in the radio wavelength of 6 cm by the Very Large Array. (*Courtesy of NRAO/AUI and M. Bietenholz*)

the magnitude of their motion along the line of sight. Supernova remnants are the result of energetic, fast stellar ejecta sweeping up large amounts of interstellar or previously ejected circumstellar materials. Planetary nebulae are formed by the interaction between a fast (2000–4000 km/s or 1250–2500 mi/s) stellar wind with the slow (of the order of 10 km/s or 10 mi/s), previously ejected wind from the progenitor red giant star. The collision of the two winds forms a dense shell that is the main part of a planetary nebula. Behind this shell is a hot, dilute bubble of gas at a temperature of millions of degrees. Thermal pressure from this bubble propels the expansion of the planetary nebula shell. *See* DOPPLER EFFECT.

Importance of studying nebulae. Astronomers are interested in gaseous nebulae for a number of reasons. Since H II regions and molecular clouds are associated with stellar births, and planetary nebulae and supernova remnants are associated with stellar deaths, they are important elements in understanding stellar evolution. Planetary nebulae and supernova remnants have internal motions that are highly supersonic, making them good laboratories of supersonic dynamics. Since planetary nebulae and supernova remnants carry heavy elements synthesized by stars, they are the vehicles for the chemical enrichment of the Milky Way Galaxy. An analysis of the chemical contents of these nebulae provides observational constraints on theories of stellar nucleosynthesis. The study of atomic transitions, in particular those that do not occur in the terrestrial environment, helps us gain understanding of atomic structures and processes. The study of molecules in star formation regions and planetary nebulae leads to the new discipline of astrochemistry, where theories are being developed on how these molecules form in a low-density environment. *See* NUCLEOSYNTHESIS; SHOCK WAVE; STELLAR EVOLUTION.

Since the solar system condensed out of an interstellar cloud, human existence is very much the product of the physical and chemical processes in the parent cloud. Interstellar clouds, in turn, are made up of materials ejected from stars through stellar winds, planetary nebulae, and supernovae. The cycle of stars from birth to death through stellar evolution, and then from death to birth again, is closed by the links between stellar ejecta, interstellar clouds, and proto-stellar nebulae. The discovery of counterparts of circumstellar dust in meteoroids has led to the speculation on whether star dust has enriched the chemical makeup of the primordial solar nebula. *See* COSMOCHEMISTRY; SOLAR SYSTEM.

In summary, the study of gaseous nebulae provides the opportunity to learn about physical and chemical processes under extreme conditions (low density, very low and very high temperatures, strong radiation background, and so forth) not available on Earth. Therefore, nebulae serve as excellent extraterrestrial laboratories for physics and chemistry as well as crucial links in the life cycle of stars. Sun Kwok

Bibliography. L. H. Aller, *Physics of Thermal Gaseous Nebulae*, D. Reidel, Dordrecht, Netherlands, 1984, paper 1987; J. E. Dyson and D. A. Williams, *The Physics of the Interstellar Medium*, 2d ed., Institute of Physics, 1997; S. Kwok, *The Origin and Evolution of Planetary Nebulae*, Cambridge University Press, 2000; S. Kwok, *Physics and Chemistry of the Interstellar Medium*, University Science Books, 2006; D. E. Osterbrock and G. J. Ferland, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*, University Science Books, 2005.

Nectarine

A smooth-skinned, fuzzless form of peach, *Prunus persica*, probably brought to the Middle East from Central Asia by Alexander the Great. The nectarine's lack of pubescence is a simple recessive genetic characteristic. Classically, the fruits were thought of as being somewhat smaller, softer, and richer in flavor than those of the peach. More recently developed cultivars, however, approximate fresh-market peaches in size and firmness but are not usually superior in flavor. Propagation and cultivation are the same as for the peach, but because of its susceptibility to brown rot (*Sclerotinia fructicola*) in humid growing areas, the nectarine is primarily grown in irrigated peach-growing areas where atmospheric humidity is low. *See* PEACH; ROSALES.

California is practically the sole commercial producer of nectarines. There have been a considerable number of plantings in irrigated areas in south-central Washington. Modern commercial production in California began with the firm, attractive, large-fruited cultivars of the Grand series developed by Fred Anderson, a private breeder. Anderson's first large nectarines were derived from second-generation seedlings produced by crossing nectarines with the large-fruited, firm J. H. Hale

peach. Other breeding programs in California and eastern states have resulted in many new nectarines, and an extended marketing season. *See* FRUIT; FRUIT, TREE. L. F. Hough; Catherine H. Bailey

Nectridea

An order of mostly aquatic lepospondyl amphibians known from Carboniferous and Permian rocks of North America, Europe, and North Africa. They were small, usually less than 20 in. (50 cm) in length, and outwardly newtlike with short trunks and long tails. Limbs were small but well developed. Carpal and tarsal bones were rarely ossified. Their vertebrae exhibit the one-piece centrum characteristic of lepospondyls, but are distinct in bearing spatulate neural spines (see **illus.**) with crenulated dorsal edges.

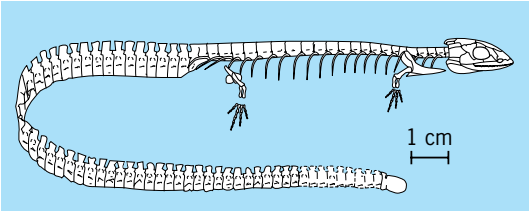

Urocordylid nectridean *Ptyonius marshii*. (*After R. L.Carroll, Vertebrate Paleontology and Evolution, W. H. Freeman, 1988*)

Three nectridean families are recognized: Urocordylidae, Keraterpetontidae, and Scincosauridae. Urocordylid skulls are distinctly arrowhead-shaped, usually with long snouts, and possessed cranial kinesis in which the sides of the skull could arc outward slightly as though hinged at the top of the skull. In one form, the snout and upper dentition were able to move up and down.

Urocordylid caudal vertebrae were adapted for swimming, being laterally compressed and bearing spatulate hemal spines below the centrum almost identical to the neural spines above it (giving the vertebrae a dorsoventrally symmetrical appearance; see illus.).

The Keraterpetontidae are recognized by the expanded tabular bones of the skull roof, which extend backward like horns. Modest in early forms, these features were exaggerated in *Diplocaulus* and *Diploceraspis*, so that from above the skulls resemble boomerangs. It is thought that the horns had ligamentous connections with the pectoral girdles, stabilizing the head during swimming.

The Scincosauridae seem to have been the only terrestrial nectrideans. Their skulls were short-snouted and lacked tabular horns. Since their tails lacked spatulate hemal spines but possessed ribs, they were shallow and rounded in cross section, not meant for swimming. The skeleton as well ossified, including the carpal and tarsal bones, as might be expected in terrestrial, weight-bearing structures. *See* AMPHIBIA; LEPOSPONDYLI. C. F. Wellstead

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988; A. L. Panchen (ed.), *The Terrestrial Environment and the Origin of Land Vertebrates*, 1980.

Negative ion

An atomic or molecular system having an excess of negative charge. Negative ions, also called anions, are formed in attachment processes in which an additional electron is captured by an atom, molecule, or cluster. They can also be formed when a molecule or cluster dissociates. Doubly charged negative ions, also called dianions, have also been observed in the case of molecules and clusters. Here, two additional electrons have become attached to the neutral systems. Negative ions are destroyed in a controlled manner in detachment processes and, in the case of molecular ions or clusters, dissociation processes, when the ion interacts with photons, electrons, heavy particles, or external fields. Experimental studies of negative ions involve measurements of cross sections for detachment and dissociation. *See* SCATTERING EXPERIMENTS (ATOMS AND MOLECULES).

Negative ions were first reported in the early days of mass spectrometry. It was soon learned that even a small concentration of such weakly bound, negatively charged systems had an appreciable effect on the electrical conductivity of gaseous discharges. Negative ions now play a major role in a number of areas of physics and chemistry involving weakly ionized gases and plasmas. Applications include accelerator technology, mass spectrometry, injection heating of thermonuclear plasmas, material processing, and the development of tailor-made gaseous dielectrics. In nature, negative ions are known to be present in tenuous plasmas such as those found in astrophysical and aeronomical environments. The absorption of radiation by negative hydrogen ions in the solar photosphere, for example, determines the Sun's spectral distribution. *See* ION SOURCES; PLASMA (PHYSICS).

A relatively recent application of negative ions is the measurement of the abundance of selected rare isotopes, such as carbon-14 (^{14}C), using accelerator mass spectrometry (AMS). In this sensitive and selective method, a beam of negative ions is injected into a tandem accelerator, where the ions gain energy and lose two or more electrons. The positive ions exiting the accelerator undergo magnetic and electrostatic analysis, allowing one to selectively count the number of ions of a chosen isotope. Negative ions are particularly useful in the case of ^{14}C mass spectrometry since a potential background associated with the nitrogen-14 (^{14}N) isobar is eliminated due to the fact that the element nitrogen does not form stable negative ions. Accelerator mass spectrometry is often used to date archeological samples. The high selectivity and sensitivity of the method compared to those of dating methods involving radioactive decay allows the study of smaller and older samples. *See*

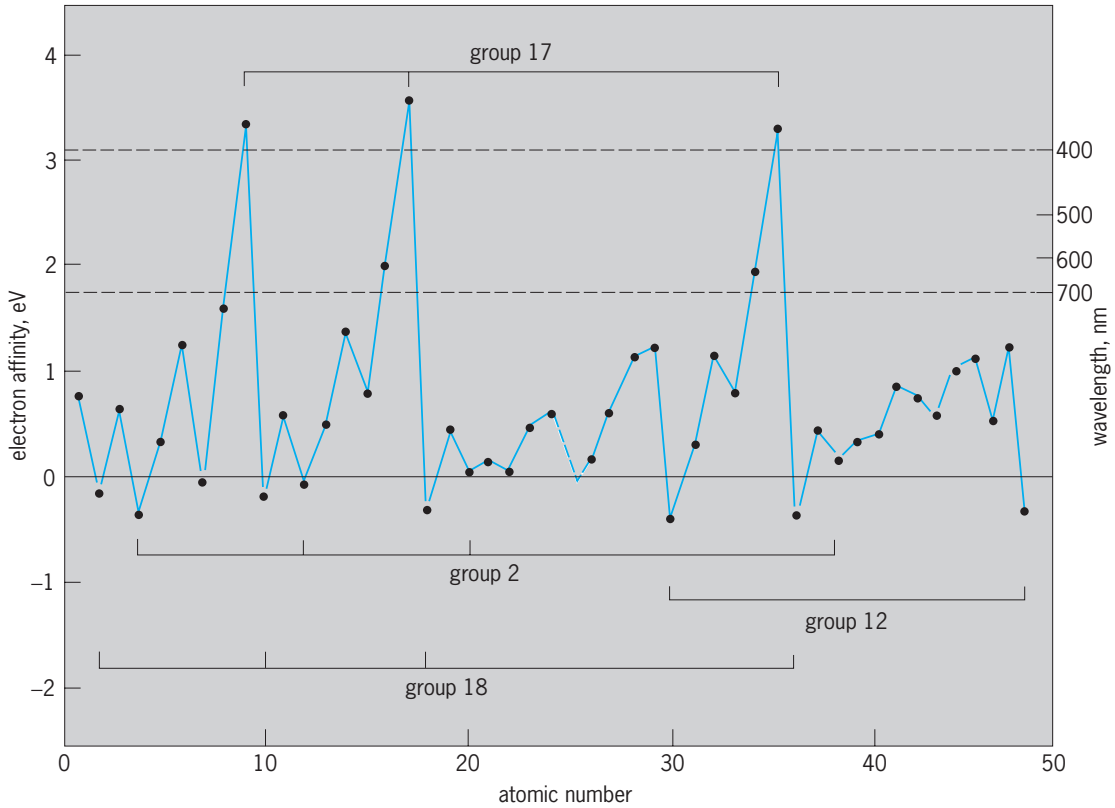

Electron affinities of the elements in the first half of the periodic table. Wavelengths of radiation at threshold for removing the least tightly bound electron are also shown; horizontal lines indicate threshold wavelengths at limits of the visible region of the electromagnetic spectrum (400 and 700 nanometers).

ACCELERATOR MASS SPECTROMETRY; MASS SPECTROSCOPE; RADIOCARBON DATING.

Electron correlations. Intrinsically, negative ions are of considerable interest since they represent a loosely bound system that is sensitive to electron correlation effects. The interaction between the outermost pair of electrons can become comparable to the relatively weak interaction of each electron with the atomic core. The electrons then drastically affect each other's behavior; that is, their motions are correlated. In contrast, the effects of correlated electron motion in atoms and positive ions are usually masked by the dominant Coulomb forces between the valence electrons and the ionic core. Because of the enhanced role played by electron correlation, a negative ion provides an excellent test bed for investigating departures from the independent electron model.

Structure. The structure of a negative ion differs fundamentally from that of an isoelectronic atom or positive ion. A stable negative ion typically possesses a single bound state in contrast to the infinite spectrum of states characteristic of an atom or positive ion. The force that binds the attached electron in a negative ion arises from polarization and exchange effects (manifestations of electron correlations), and as such it is considerably weaker and shorter in range than the Coulomb force which binds the valence electron to the ionic core in atoms and positive ions. *See* EXCHANGE INTERACTION.

The binding energy of the least tightly bound electron in a negative ion is numerically equal to the electron affinity of the parent atom. Electron affinities of atoms are at least an order of magnitude smaller than corresponding ionization energies of atoms and positive ions. These affinities range from 0.02 to 3.6 eV and display periodicity with atomic number, which is associated with the Pauli exclusion principle (see **illus.**). *See* ELECTRON AFFINITY; EXCLUSION PRINCIPLE.

Doubly excited states of negative ions involving either the excitation of a pair of valence electrons or a valence electron and a core electron are quite common. Such transient states are embedded in continua above the detachment limit and decay spontaneously via autodetachment. They are manifested as resonance structures in detachment cross sections. Pioneering studies of doubly excited resonant states of the simplest negative ion, H^-, were carried out by Howard Bryant and coworkers in the 1970s. Since then, resonant states involving valence and core excitation have been investigated in many ions using either laser or synchrotron radiation. *See* LASER SPECTROSCOPY; RESONANCE (QUANTUM MECHANICS); SYNCHROTRON RADIATION.

Occurrence. More than 80% of the naturally occurring elements have positive electron affinities, and therefore form stable atomic negative ions when an electron is attached to an atom in its ground state. Other elements form metastable ions when

an electron is attached to an atom in a metastable excited state. Such ions are unstable but often sufficiently long-lived for experimental investigation. Many molecules also form negative ions. A common characteristic of all negative ions is their fragility due to the weak binding force. Spectroscopic sources, such as a tenuous beam of accelerated ions passing through a vacuum, are designed to minimize the collisional detachment of electrons while maintaining a density sufficiently high to investigate interactions of photons with negative ions. *See* ATOMIC STRUCTURE AND SPECTRA; ION; MOLECULAR BEAMS. David J. Pegg

Bibliography. L. H. Andersen, T. Andersen, and P. Hvelplund, Studies of negative ions in storage rings, *Adv. At. Mol. Opt. Phys.*, 38:155–191, 1997; T. Andersen, Atomic negative ions, *Phys. Rep.*, 394:157–313, 2004; T. Andersen, H. K. Haugen, and H. Hotop, Binding energies in atomic negative ions, III, *J. Phys. Chem. Ref. Data*, 28:1511–1533, 1999; H. S. W. Massey, *Negative Ions*, 3d ed., 1976; D. J. Pegg, The structure and dynamics of negative ions, *Rep. Prog. Phys.*, 67:1–49, 2004.

Negative-resistance circuits

Electronic circuits or devices that, over some range of voltage v and current i, satisfy Eq. (1) for equiv-

$$R_{eq} = \frac{dv}{di} < 0 \tag{1}$$

alent resistance R_{eq} (where the voltage and current polarities are defined in **Fig. 1***a*). They are used as building blocks in designing circuits for a wide range of applications, including amplifiers, oscillators, and memory elements. *See* ELECTRICAL RESISTANCE.

An ideal negative resistor would have the voltage–current relationship (transfer characteristic) shown in Fig. 1*b*, and thus satisfy Ohm's law with a negative value for the resistance. However, the same effect can generally be obtained with any circuit (or physical device) whose voltage–current curve contains a region of negative slope. Figure 1*c*, for example, shows transfer characteristics typical of a tunnel diode and a neon bulb, which can be operated in the negative-resistance regions indicated. *See* OHM'S LAW; RESISTOR.

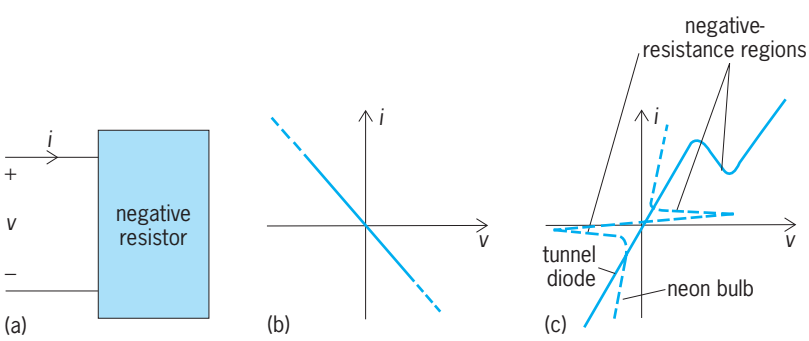

Fig. 1. Characteristics of negative resistors. (*a*) Definition of voltage (*v*) and current (*i*) polarities. (*b*) Voltage-current transfer characteristic of an ideal negative resistor. (*c*) Transfer characteristic of practical physical devices with negative-resistance regions: a tunnel diode and a neon bulb (not to the same scale).

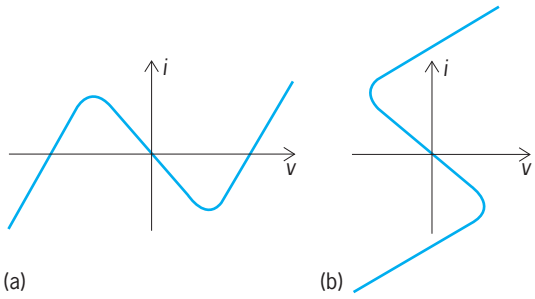

Fig. 2. Large-signal behavior of a negative resistance having a finite internal power supply. (*a*) Voltage-controlled resistance. (*b*) Current-controlled resistance.

Common generalizations of the negative-resistance idea include negative capacitors, negative inductors, and frequency-dependent negative resistors. Some of the circuits used to implement them are negative impedance converters, negative impedance inverters, and generalized immittance converters. This article develops these ideas together, starting with properties inherent to the idea of negative impedance, and then discussing theoretical and practical implementations. *See* ELECTRICAL IMPEDANCE; IMMITTANCE; INDUCTOR.

Properties of negative resistors. The power dissipated in a device, given by Eq. (2), is negative in

$$P_{DISS} = vi \tag{2}$$

the second and fourth quadrants of the v–i plane of Figs. 1*b* and *c*. Thus, the ideal negative resistor whose characteristic is shown in Fig. 1*b* generates power. Two consequences of this are that an active circuit (a circuit containing a power supply) is required to implement the ideal characteristic of Fig. 1*b* but is not necessary for the small-signal negative resistances of Fig. 1*c*; and that for any practical circuit, the characteristic curve must eventually fold over into the power-dissipating quadrants, as shown in **Fig. 2***a* or *b*. If the curve did not fold but just continued forever, it would be possible to extract an infinite amount of power from the device.

The two types of curve of Fig. 2 correspond to an important dichotomy in types of negative resistance. The N-shaped curve of Fig. 2*a* allows current to be a single-valued function of voltage (but not vice versa), and circuits with this behavior are therefore called voltage-controlled negative resistors. Dually, the S-shaped curve of Fig. 2*b*, for which Eq. (3) is

$$v = f(i) \tag{3}$$

appropriate, describes a current-controlled negative resistor. The tunnel-diode characteristic of Fig. 1*c* can be seen to be voltage-controlled, while the neon tube is current-controlled.

If the terminals of a current-controlled negative resistor are open-circuited, then $i = 0$ and there is a unique solution $v = f(0)$. The voltage-controlled circuit, however, can have any of three voltages in this situation (the three intersections of the N with the horizontal axis). Dually, the S-curve gives a

device with multiple equilibrium states when short-circuited. When the dynamic behavior of these circuits is accounted for, it is found that some of these equilibria are stable and some are unstable. These stability considerations are essential to designing a negative-resistance circuit for a particular application.

Dynamic behavior of negative resistors. In order to analyze the dynamic behavior of a negative resistor—for instance, to determine which equilibria are stable and which are unstable—the remaining circuit elements in which the negative resistor is embedded, including relevant parasitic capacitors and resistors, are added to the device. If an ideal negative resistor R were placed in parallel with a positive capacitance C, the result would be unstable: for some initial capacitor voltage v_0, the current i would be directed into the capacitor, whose voltage would increase at a rate given by Eq. (4). The differential equation has the unstable solution given by Eq. (5), where $\tau = -CR$

$$C\frac{dv}{dt} = i \tag{4}$$

$$v = v_0 e^{t/\tau} \tag{5}$$

is positive because $R < 0$, $C > 0$. Similarly, a loop of a negative resistor and a positive inductor is unstable; its current is given by Eq. (6), with $\tau = -L/R$.

$$i = i_0 e^{t/\tau} \tag{6}$$

These instabilities would seem to make the negative-resistance circuit useless by itself: a negative resistor left open-circuited would have some small (positive) stray capacitance between its terminals and hence be unstable; and when short-circuited it would be unstable because of the finite (positive) inductance of any short circuit. Practical applications of negative-resistance circuits involve embedding the negative-resistance elements (circuits) in larger circuits, and it is the stability of these larger circuits, not the stability of the negative resistance by itself, that is of interest.

In some applications of negative resistances, direct stabilization of the negative resistance is required. One method of achieving this is with the utilization of negative inductors or capacitors, which are synthesized in a manner similar to that used to synthesize the negative resistors themselves. For example, a negative resistance in parallel with a negative capacitance is stable.

This behavior interacts subtly with the large-signal (S or N) behavior. If a negative resistor is paralleled with a negative capacitance, then it will be stable when open-circuited. This is desirable for a device whose large-signal state behavior is S-shaped (Fig. 2*b*), because then the (unique) equilibrium for $i = 0$ will be stable. Dually, a negative resistor that is to be short-circuit-stable should contain a parasitic series negative inductance and have N-shaped large-signal behavior.

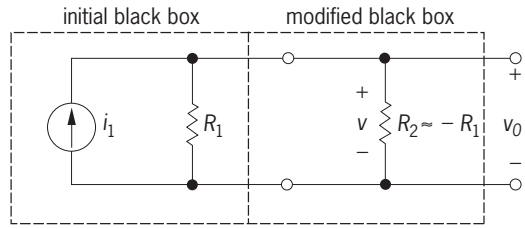

Fig. 3. Application of a negative resistance to cancel the effect of a load, thereby making a kind of amplifier.

Applications. Common applications of negative resistors include amplifiers, bistable or memory circuits, and oscillators.

Amplifier design. A common design problem is shown in **Fig. 3**, in which the output impedance of a current source is positive and unacceptably small. (This is a simplification of a situation in which negative resistances have been used for telephone system and integrated-circuit amplifier designs.) A large open circuit voltage is desired at the output, but on its own the system produces an inadequately small voltage i_1R_1. It is further assumed that the only access available is to the terminals outside the "black box." By placing a negative resistance $R_2 = -R_1$ in parallel with the built-in load, it is possible to cancel its effect: the net parallel resistance is given by Eq. (7),

$$R_{\text{eff}} = \frac{R_1R_2}{R_1 + R_2} = \infty \tag{7}$$

which would theoretically allow an infinite output voltage.

In practice R_1 cannot be made exactly equal to $-R_2$. Practically, $|R_2|$ is intentionally made slightly larger or slightly smaller than R_1 so that the overall circuit, including any additional elements and parasitic components, remains stable. *See* AMPLIFIER.

Bistable circuits. **Figure 4***a* shows a negative resistor in a simple memory circuit, and Fig. 4*b* shows the N-shaped characteristic of the resistor. When $i_1 = 0$, this characteristic allows three equilibria. If C is positive, the central equilibrium will be unstable; the two outer equilibria will, however, be stable, because $C > 0$ and the positive slope dv/di acts dynamically like a positive resistor. The resulting circuit is called a bistable circuit, because it is stable in either of two states, and it stores one bit of data by "remembering" its current state.

The circuit can be switched from one state to the other by pulsing i_1. The broken lines on the diagram are intended to suggest that the circuit switches from one side to the other when forced to change states.

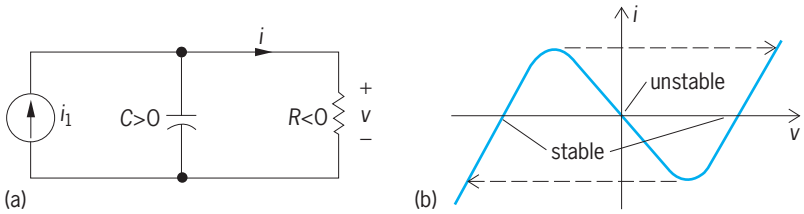

Fig. 4. Negative resistor used as a memory circuit (or bistable circuit or latch). (*a*) Circuit. (*b*) Resistor characteristic.

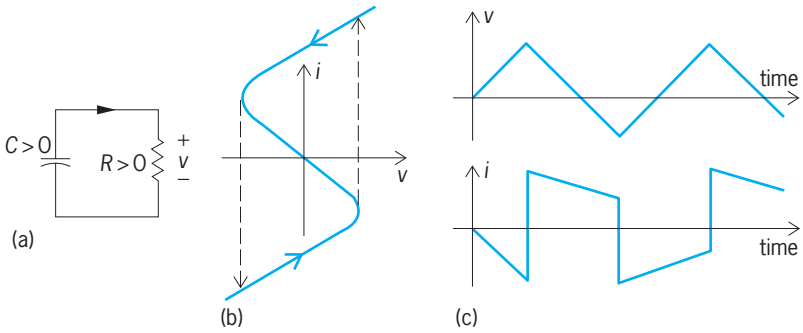

Fig. 5. Negative-resistance oscillator (or astable circuit). (*a*) Circuit. (*b*) Resistor characteristic. (*c*) Voltage and current waveforms.

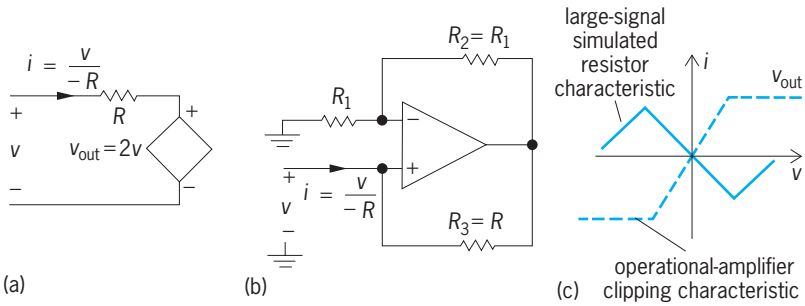

Fig. 6. Active circuits that simulate a negative resistance. (*a*) Circuit that uses an ideal voltage amplifier with a gain of 2. (*b*) Circuit that uses an operational amplifier. (*c*) The operational-amplifier clipping characteristic and resulting large-signal voltage-current characteristic of the simulated negative resistor.

The behavior in the transition is defined by the unstable time constant RC.

Oscillators. An astable, or oscillator, circuit (**Fig. 5*a***) is obtained by replacing the voltage-controlled resistor of Fig. 4 with a current-controlled device (Fig. 5*b*). The waveforms of voltage and current obtained in this case are shown in Fig. 5*c*. A second type of oscillator, which produces sinusoidal signals, can be obtained by using a negative resistor to cancel the natural (positive) resistances that damp out oscillations in tank circuits (inductor-capacitor loops) and crystals. *See* OSCILLATOR.

Implementations of negative resistors. Negative resistors can be implemented by using amplifiers in positive-feedback configurations. **Figure 6*a*** shows how a voltage amplifier with a gain of 2 can be used to simulate a grounded negative resistor, and Fig. 6*b* shows an operational-amplifier implementation of the same idea. *See* OPERATIONAL AMPLIFIER.

In the practical case of a clipping amplifier, which has the input-output characteristics shown in Fig. 6*c*, the resulting large-signal voltage–current behavior of the simulated resistor is as shown in the figure. This is a voltage-controlled resistor.

Again in the practical case, the amplifier has a finite frequency response. If the gain of the amplifier of Fig. 6*a* drops with frequency ω as in Eq. (8)

$$T(j\omega) = \frac{2}{1 + j(\omega/\omega_0)} \tag{8}$$

[where, in electronics, $j = \sqrt{-1}$], then a negative parasitic inductance is simulated at the same time as the negative resistor, which makes the implementation short-circuit-stable. *See* GAIN; RESPONSE.

Figure 6 shows one of the four possible types of controlled source (voltage or current sources, controlled by voltages or currents) being used to create a negative resistor; all four types can be used. **Figure 7**, for example, shows an implementation based on an ideal voltage-controlled current source, together with a practical transistor implementation. Transistor M1 can be modeled as a voltage-controlled current source but produces an output current of the wrong sign. Transistors M2 and M3 form a current mirror and correct the sign. This circuit, like that of Fig. 6, produces an N-shaped (voltage-controlled) characteristic with a series negative inductor, and so is short-circuit-stable. *See* CURRENT SOURCES AND MIRRORS; TRANSISTOR.

Dual circuits (circuits in which currents and voltages are interchanged) to those of Fig. 6*a* and 7*a* produce open-circuit-stable negative resistors with S-shaped (current-controlled) voltage–current characteristics. *See* CIRCUIT (ELECTRONICS).

Negative-impedance converters. The circuit of Fig. 6*b* is called a negative-impedance converter (NIC). Analysis with arbitrary resistances R_1, R_2, and R_3 and an ideal operational amplifier shows that it has input impedance $-R_1R_3/R_2$. By replacing some resistors with capacitors, the circuit can be generalized to make several different devices, including (1) a negative capacitor, obtained by replacing R_1 or R_3 with a capacitor; (2) a negative inductor, obtained by replacing R_2 with a capacitor; (3) a positive inductor, obtained by replacing R_1 with a negative inductor (which, in turn, can be implemented with a second negative-impedance converter, as before); (4) an impedance that varies with frequency ω as $1/\omega^2$, obtained by replacing resistors R_1 and R_3 with capacitors; and (5) a related circuit, the frequency-dependent negative resistor, obtained by replacing resistor R_3 with a capacitor and resistor R_1 with a negative capacitor (which, in turn, may be implemented with a second negative-impedance converter).

All of these impedances are of the short-circuit-stable type, as can be verified by noticing that in each case the positive-feedback path has less gain than the negative-feedback (stabilizing) path when

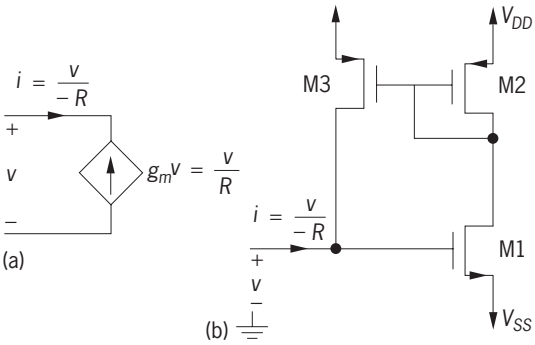

Fig. 7. Ideal negative resistor implemented with an ideal voltage-controlled current source. (*a*) Ideal circuit. (*b*) Practical MOSFET implementation.

the input is grounded. Interestingly, the circuit is also a negative-impedance converter when the amplifier's positive and negative input terminals are interchanged. In this case it still has input impedance $-R_1R_3/R_2$, but it becomes open-circuit-stable. This one circuit, then, can implement negative resistors, frequency-dependent negative resistors, negative capacitors, and negative inductors of either the short-circuit-stable or open- circuit-stable types.

The transistor circuit of Fig. 7*b* can also be used as a negative-impedance converter by including a positive impedance in the source lead of M1 and adding a diode-connected transistor M4 (one whose drain and gate are connected) in series with the input. The resulting circuit can be driven from the source of either M1 or M2, and accordingly either a voltage-controlled or a current-controlled negative resistor is obtained. This circuit is also known as a current conveyor and is the basis for the precision "band-gap" voltage references widely used in analog integrated circuit design wherever a temperature-insensitive voltage level is required. *See* INTEGRATED CIRCUITS.

Physical devices. The positive feedback accomplished by circuitry in Figs. 6 and 7 can occur internally in a physical device and so lead to natural negative resistance characteristics like those of Fig. 1*c* over some region of operation. The characteristic is necessarily passive (that is, it cannot generate power and is therefore restricted to the first and third quadrants of the voltage–current plane). These characteristics are sometimes said to show differential negative resistance, since they have negative derivatives.

As a simple example, if a current flows in a neon bulb, it tends to ionize the gas. This, in turn, creates carriers for current and so decreases the resistance of the bulb. This, in turn, leads to increased current, completing a positive-feedback loop. This behavior produces a characteristic with the negative-resistance region shown in Fig. 1*c*. It is of the current-controlled type, since, as for any given current, only one level of ionization exists, whereas at a given voltage the device may be either "on" or "off." A common hobbyist circuit implements an oscillator by placing a capacitor in parallel with a neon bulb, then charging the capacitor from a high voltage through a large resistor. This is essentially the astable circuit of Fig. 5, augmented with a biasing circuit (the voltage source and resistor) to force the circuit to operate in the negative-resistance region. *See* ELECTRICAL CONDUCTION IN GASES; VAPOR LAMP.

Many other devices have negative-resistance regions in their voltage–current characteristics. Although the mechanisms involved are much less straightforward to understand than those for the neon bulb, these devices are more useful for practical circuit design.

Tunnel diode. The best-known negative-resistance device is the tunnel diode, which was discovered by L. Esaki in 1958. It is very useful because the phenomenon that it exploits is a quantum-mechanical effect that happens much more rapidly than most others in electronics.

A tunnel diode consists of two very heavily doped regions of a semiconducting material with a very abrupt junction between them. These regions, like any crystalline material, can contain electrons only with energies in certain bands. One side of the junction is doped to have a generous supply of electrons in a certain band of energies, while the other side has a great many vacancies (holes) for electrons in another band. By applying an external voltage, these bands are shifted relative to one another. For small input voltages, these bands overlap, and the quantum-mechanical effect known as tunneling can take place. In this effect, the electrons on one side of the junction (where there is a good supply) have a finite probability of reappearing on the other side (where these are many vacancies). Thus, for small applied voltages, a large current flows. This is still ordinary positive-resistance behavior, although the mechanism for producing current flow is quite different from the field-induced drift of particles that provides conduction in ordinary devices. *See* BAND THEORY OF SOLIDS; SEMICONDUCTOR.

As the applied voltage increases, however, the bands of electrons and holes on the two sides of the junction start to slide past one another, and eventually their region of overlap starts to decrease. Since quantum tunneling can occur only from an electron in the "supply" to a vacancy at the same energy, this reduction in overlap reduces the amount of charge flowing. Thus, an increasing voltage produces decreasing current, for a negative differential resistance like that shown in Fig. 1*c*.

Since the extent of overlap of the bands is set by the voltage applied to the junction, this produces a voltage-controlled characteristic. Since the time constants associated with tunneling (set by Heisenberg's uncertainty principle) are very short, the phenomenon is very fast. Circuits using these devices operate in the millimeter-wave region (that is, at radio frequencies so high that wavelengths are expressed in millimeters). *See* UNCERTAINTY PRINCIPLE.

A number of other quantum electronic devices have been developed that also have negative-resistance characteristics. In particular, devices have been constructed that have two barriers (instead of the single barrier created by the tunnel-diode junction) and make use of resonant tunneling, where the spacing between the barriers creates a resonance for electrons at certain frequencies. This resonance, in turn, enhances the rate of tunneling. These devices are claimed to be useful at terahertz (10^{12} Hz) frequencies. *See* SEMICONDUCTOR HETEROSTRUCTURES; TUNNEL DIODE; TUNNELING IN SOLIDS.

Transferred-electron devices. These are made of materials (*n*-type gallium arsenide is an important example) that have a negative-slope region in their mobility curves. These curves measure electron velocity as a function of electric field, and so correspond at the level of the material to voltage–current curves in the overall device. The negative-slope region comes about because electrons traveling fast enough to exceed a certain energy threshold can transfer to a second mode of propagation, in which their interactions

with the crystal make them appear heavier. As the field increases beyond this point, electrons start to slow down (as more enter the slower mode) and current decreases—a negative-resistance effect.

These devices do not appear directly as negative resistors at their terminals, however, but as oscillators. This is known as the Gunn effect and comes from the fact that electric field distribution in this type of material is inherently unstable. Rather than the field strength and the density of electrons being uniform over the length of the device, part of it operates below the negative-resistance region at low fields and high electron speeds, and another part acts as a bottleneck or domain with a high field and a large number of slower-moving electrons. This domain moves through the material from cathode to anode and disappears as a new domain forms at the cathode. These devices thus oscillate at a frequency set by the time taken for domains to travel from one end to the other.

This type of instability can be thought of in circuit terms by regarding the device as a series connection of smaller devices, with a positive capacitance to ground (the self-capacitance of a piece of material) at each connection. At these nodes there is a positive capacitance and a negative (differential) resistance, which is an unstable combination. *See* MICROWAVE SOLID-STATE DEVICES.

Other devices. Other devices whose physics lead to negative resistance are the point-contact transistor and the unijunction transistor, but these devices are no longer readily available or widely used. The thyristor (or silicon-controlled rectifier), which is a four-layer transistor structure consisting of alternating *p*- and *n*-type regions, is very common in high-power circuits and can be regarded as a positive-feedback connection of overlapping *npn* and *pnp* transistors. It has an electrical latching behavior rather like that of the neon bulb discussed above, and will switch from an "off" state to "on" if a certain threshold is exceeded. Thyristors are usually used as three-terminal devices, however, and so they are not generally considered to be negative resistors. *See* SEMICONDUCTOR RECTIFIER.

Martin Snelgrove

Bibliography. L. O. Chua, C. A. Desoer, and E. S. Kuh, *Linear and Nonlinear Circuits*, 1987; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997; S. M. Sze, *Physics of Semiconductor Devices*, 2d ed., 1981.

Negative temperature

The property of a thermodynamical system which satisfies certain conditions and whose thermodynamically defined absolute temperature is negative. The essential requirements for a thermodynamical system to be capable of negative absolute temperature are that (1) the elements of the thermodynamical system must be in thermodynamical equilibrium among themselves in order for the system to be described by a temperature at all; (2) there must be an upper limit to the possible energy of the allowed states of the system; and (3) the system must be thermally isolated from all systems which do not satisfy both of the first two requirements, that is, the internal thermal equilibrium time among the elements of the system must be short compared to the time during which appreciable energy is lost to or gained from other systems.

The second condition must be satisfied if negative temperatures are to be achieved with a finite energy. If W_m is the energy of the mth state for one element of the system, then in thermal equilibrium the number of elements in the mth state is proportional to the Boltzmann factor $\exp(-W_m/kT)$, where k is Boltzmann's constant and T is the absolute temperature. For negative temperatures, the Boltzmann factor increases exponentially with increasing W_m, and the high-energy states are therefore occupied more than the low-energy ones; this situation is the reverse of the positive-temperature case. As a consequence, with no upper limit to the energy, negative temperatures could not be achieved with a finite energy. Most systems do not satisfy this condition; for example, there is no upper limit to the possible kinetic energy of a gas molecule. It is for this reason that systems of negative temperatures occur only rarely. *See* KINETIC THEORY OF MATTER; STATISTICAL MECHANICS.

Systems of interacting nuclear spins, however, have the characteristic that under suitable circumstances they can satisfy all three of the above conditions, in which case the nuclear spin system can be at negative absolute temperature, as first demonstrated by R. V. Pound, E. M. Purcell, and N. F. Ramsey and discussed theoretically by Ramsey and others. *See* NUCLEAR ORIENTATION.

The difference between positive and negative absolute temperatures can be illustrated with a bar chart showing the population distributions among the different states of a system at positive and negative absolute temperatures (see **illus.**). A simple example is provided by a thermodynamic system consisting of N magnetically mutually interacting nuclei of spin 2, so that each spin can be in any one of five equally spaced energy states of energy $W = mD$,

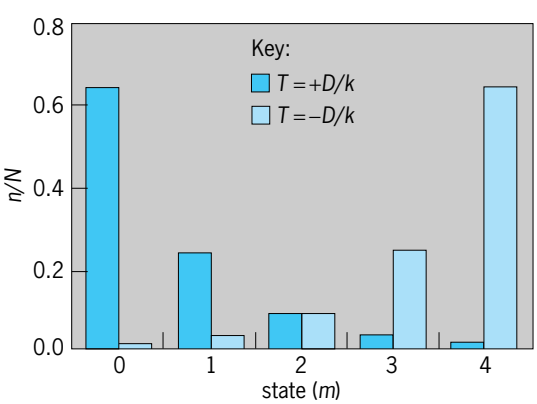

Populations (n) of states of a thermodynamic system consisting of N magnetically interacting nuclei of spin 2 at absolute temperatures $T = +D/k$ and $T = -D/k$, where k is Boltzmann's constant and D is the energy spacing between successive states.

where m is an integer designating the state and $m = 0, 1, 2, 3$, or 4. The number, n, of nuclei in the different states can then be shown in the bar chart at positive and negative absolute temperatures, such as $T = +D/k$ and $T = -D/k$. The bars decrease in length with increasing values of m at positive absolute temperature, and increase in length with increasing values of m at negative absolute temperature.

From the point of view of thermodynamics, the absolute temperature T is given by the equation below,

$$T = \left(\frac{\partial S}{\partial U}\right)_B^{-1}$$

where S is the entropy, U the internal energy, and B the externally applied magnetic field. A system at negative absolute temperature must therefore be in a condition such that the slope of S as a function of U is negative. This is clearly achievable for a spin system in an external field B, since the entropy S falls to a low value when all the nuclear spins are in their highest energy state and thereby all pointing in the same direction. *See* ENTROPY; THERMODYNAMIC PRINCIPLES.

It is apparent either from the Boltzmann factor or from the above equation applied to a spin system that the transition between positive and negative temperatures is through infinite temperature, not absolute zero; negative absolute temperatures should therefore not be thought of as colder than absolute zero, but as hotter than infinite temperature. *See* ABSOLUTE ZERO; TEMPERATURE. Norman F. Ramsey

Bibliography. A. Abragam, *Nuclear Magnetism*, 1983; B. Gal-Or, *Modern Developments in Thermodynamics*, 1974; E. M. Purcell and R. V. Pound, A nuclear spin system at negative temperature, *Phys. Rev.*, 81:279, 1951; N. F. Ramsey, Thermodynamics and statistical mechanics at negative absolute temperatures, *Phys. Rev.*, 103:20–28, 1956; N. F. Ramsey and R. V. Pound, Nuclear audiofrequency spectroscopy by resonant heating of nuclear spin system, *Phys. Rev.*, 81:278, 1951.

Nemata (Nematoda)

A group of unsegmented worms which have been variously recognized as an order, class, and phylum. They are now widely accepted as a separate phylum, Nemata (or Nematoda). When treated as a class, they were assigned either to the phylum Aschelminthes or to the phylum Nemathelminthes. In Aschelminthes they were ranked equally with Rotifera (Rotatoria), Gastrotricha, Kinorhyncha (Echinodera), and Nematomorpha (gordian or horsehair worms).

The original name applied to the group was Nematoidea (from the Greek, threadlike). This group included only the animal-parasitic taxa. When the group was expanded to include marine and freshwater free-living forms, and also plant-parasitic forms, the concept of a phylum was applied and the name Nemates proposed. B. G. Chitwood amended the spelling to Nemata.

All taxonomic groupings above the species level are subjective and therefore in nearly constant flux. The primary reason for disagreement that surrounds nematodes and their allies is the concept of the pseudocoelom (body cavity) and its embryological formation. Pseudocoelom refers to a body cavity only partially lined with mesoderm. How this partially lined cavity is formed during embryological development is the key question. If among the various allies of nematodes the formation is homologous, all should probably be considered under one phylum. If, however, as the available information indicates, the pseudocoelom is not formed homologously in these near groups, each must be treated as a separate phylum; this concept is now widely accepted and is adhered to here. A classification of nematodes follows; see separate articles on each order.

Phylum Nemata
- Class: Adenophorea
  - Subclass: Enoplia
    - Order: Enoplida
      - Superfamily: Enoploidea
        - Oxystominoidea
    - Order: Oncholaimida
      - Tripylida
      - Superfamily: Tripyloidea
        - Ironoidea
    - Order: Isolaimida
      - Mononchida
      - Superfamily: Mononchoidea
        - Bathyodontoidea
        - Mononchuloidea
    - Order: Dorylaimida
      - Superfamily: Dorylaimoidea
        - Actinolaimoidea
        - Belondiroidea
        - Encholaimoidea
        - Diphtherophoroidea
        - Trichodoroidea
        - Nygolaimoidea
    - Order: Stichosomida
      - Superfamily: Trichocephaloidea
        - Mermithoidea
        - Echinomermelloidea
  - Subclass: Chromadoria
    - Order: Araeolaimida
      - Superfamily: Araeolaimoidea
        - Axonolaimoidea
        - Plectoidea
        - Camacolaimoidea
        - Tripyloidoidea
    - Order: Chromadorida
      - Superfamily: Chromadoroidea
        - Cyatholaimoidea
        - Choanilaimoidea
        - Comesomatoidea
    - Order: Desmoscolecida
      - Superfamily: Desmoscolecoidea
        - Greeffielloidea
    - Order: Desmodorida

Superfamily: Desmodoroidea
Ceramonematoidea
Monoposthioidea
Draconematoidea
Epsilonematoidea
Order: Monhysterida
Superfamily: Monhysteroidea
Linhomoeoidea
Siphonolaimoidea
Class: Secernentea
Subclass: Rhabditia
Order: Rhabditida
Superfamily: Rhabditoidea
Alloionematoidea
Bunonematoidea
Cephaloboidea
Panagrolaimoidea
Robertioidea
Chambersielloidea
Elaphonematoidea
Order: Strongylida
Superfamily: Strongyloidea
Diaphanocephaloidea
Ancylostomatoidea
Trichostrongyloidea
Metastrongyloidea
Cosmocercoidea
Oxyuroidea
Heterakoidea
Subclass: Spiruria
Order: Spirurida
Superfamily: Spiruroidea
Physalopteroidea
Filarioidea
Drilonematoidea
Order: Ascaridida
Superfamily: Ascaridoidea
Camallanoidea
Dracunculoidea
Subuluroidea
Seuratoidea
Dioctophymatoidea
Muspiceoidea
Subclass: Diplogasteria
Order: Diplogasterida
Superfamily: Diplogasteroidea
Cylindrocorporoidea
Order: Tylenchida
Superfamily: Tylenchoidea
Criconematoidea
Sphaerularoidea
Aphelenchoidea
Aphelenchoidoidea

The Nemata are unsegmented or pseudosegmented (any superficial annulation is limited to the cuticle) bilaterally symmetrical worms with a basically circular cross section (cylindroid; **Fig. 1**). The body is covered by a noncellular cuticle secreted by an underlying epidermis (hypodermis). The cylindrical body is usually bluntly rounded anteriorly and tapering posteriorly. The body cannot be easily divided into head, neck, and trunk or tail, although a region posterior to the anus is generally referred to as the tail. The anterior extremity characteristically bears 16 setiform or papilliform sensory organs and 2 chemoreceptors called amphids. The oral opening is terminal (rarely subterminal) and followed by the stoma, esophagus, intestine, and rectum, which opens through a subterminal anus. Females have separate genital and digestive tract openings. In males the tubular reproductive system joins posteriorly with the digestive tract to form a cloaca. Adult nematodes are extremely variable in size, ranging from less than 0.3 mm (0.012 in.) to over 8 m (26 ft). Nematodes are generally colorless except for food in the intestinal tract or for those few species which have eyespots. A unique character in the phylum is protoplasmic extension from the somatic muscles that reaches for synapsis with the central nervous system, rather than axonic extension of the nervous system to the muscles as in other animals (**Fig. 2**).

External covering. Nematodes are covered by a noncellular elastic cuticle chiefly composed of scleroproteins (not chitin). Chitin is known only from the eggshell. The cuticle is a complicated histologic structure composed of several layers basically divisible into four strata: epicuticle, exocuticle, mesocuticle, and endocuticle. Underlying the cuticle is the hypodermis (epidermis). The hypodermal cells may be uninucleate (all Adenophorea) or multinucleate (some Secernentea). The cell bodies of the hypodermis protrude into the body cavity as four longitudinal chords dorsally, laterally, and ventrally between the somatic muscle bands. The protoplasmic portion of the hypodermis extends as a thin layer between the cuticle and the muscle sheath. The cuticle may be smooth or marked by transverse striae forming a pseudoannulation, or it may be thickened and sculptured in various patterns. Other modifications that occur are alae or wing areas found laterally on the head, along the main body, or on the tail of males.

Alimentary canal. The oral opening is terminal and may be surrounded by three or six lips, or paired pseudolabia, or jaws. The oral opening leads to the mouth cavity, or stoma. The stoma is generally composed of two parts: the anterior cheilostome (formed embryologically from external cuticle) and the esophastome (modified from the cuticular lining of the anterior extremity of the esophagus). The stoma (or stomodeum), combined from both parts, takes on a variety of forms according to feeding habits. Predaceous nematodes have stomas armed with movable or nonmovable teeth or spears. Bacterial feeders generally have cylindrical stomas without teeth. Mycetophagous and phytophagous nematodes generally possess protrusible hollow axial spears. Following the stoma is the cuticularly lined esophagus. The lumen of the esophagus is triradiate, with one ray directed ventrally and two subdorsally. The esophagus is composed of radial muscles, epidermal tissue, a complex esophagosympathetic nervous system, and three to five "salivary" glands. The esophagus is variously modified into sections by means of bulbs and valves, permitting suction and ingestion of food. Between the esophagus and the

intestine is an esophagointestinal valve that prevents food regurgitation. The intestine is composed of a single layer of cells (uni- or multinucleate) that are bordered by microvilli on their interior edge. When visible, this microvilli area is called the brush border or bacillary layer. In some nematodes the posterior intestine may be variable modified and designated as a prerectum or proctodeum. At the junction of the intestine and rectum there may be rectal glands, an intestinorectal valve, or sphincter muscle. The cuticularly lined rectum opens through the ventromedian anus (cloaca in males). Rarely is the anus terminal or subterminal.

Somatic musculature. The somatic musculature consists of four or more longitudinal bands separated by the hypodermal chords. Each band is composed of nonstriated, continuous interlocking spindle-shaped, uninucleate muscle cells. From each cell there extends a protoplasmic arm that synapses with a medial nerve. Circular somatic muscles are absent in Nemata. Throughout the body there are specialized muscles that operate the jaws or protrusible teeth and spears, or act as protractors and retractors of the male spicules or as sphincters of the female vagina; in addition there are scattered cutaneoesophageal and cutaneointestinal muscles.

Nervous system. The major portion of the central nervous system consists of a large circumesophageal commissure (nerve ring) and its attendant anterior and posterior ganglia (lateral, dorsal, and ventral; **Fig. 3**). Ventrally the nerve ring gives rise to the ventral nerve cord, the major nerve of the body. The ventral nerve is a partly paired, mostly single, ganglionated nerve chain (not distinctly set off) that extends the length of the body. Throughout the body there are commissures that interconnect the ventral nerve cord with the lateral or dorsal nerves. The lateral nerves are primarily sensory, while the dorsal and ventral nerve cords are motor. The anterior sense organs are innervated from six ganglia anterior to the nerve ring, but the amphids are innervated from a ganglion in the lateral ganglia posterior to the nerve ring. Another subcuticular peripheral nerve net has been described for marine nematodes.

Pseudocoel. The body cavity is fluid filled and extends the length of the body between the muscle bands and internal organs. The pseudocoel contains a variety of cells and is lined to a greater or lesser extent by membranes and mesenteries that support the internal organs. The nuclei of these structures are reportedly of limited and fixed numbers. The lining material originates from migratory mesenchymatous tissue during embryogeny; these cells grow together in a network that covers the esophagus, intestine, and gonads, and may delimit the muscle bands and hypodermal chords.

Excretory system. No single type of excretory system characterizes Nemata, and in many taxa the system is completely absent (**Fig. 4**). When present in Adenophorea, the system is generally limited to a single ventral cell that opens through an anterior ventromedian pore. In Secernentea the cell has one or more tubules that extend the length of the body, either posteriorly or both anteriorly and posteriorly. The duct leading from the cell to the excretory pore, in Secernentea but rarely in Adenophorea, is cuticularly lined and opens ventromedially generally in the esophageal region. In taxa lacking an excretory cell the function, presumably, is taken over by hypodermal glands or even the prerectum.

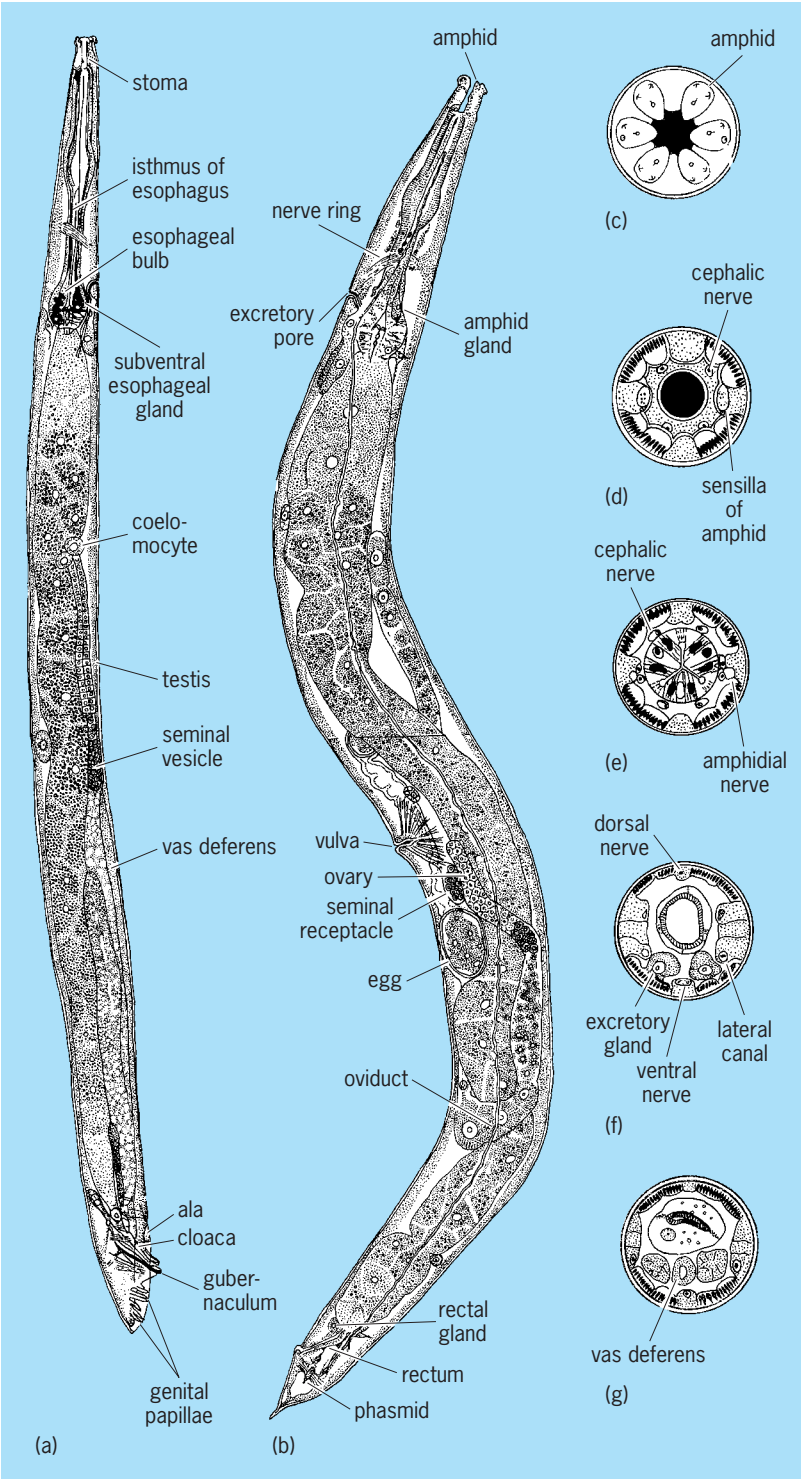

Fig. 1. ***Pelodera strongyloides***. (*a*) Male. (*b*) Female. (*c*) Head, cross section. (*d*) Stomatal region. (*e*) Esophageal region. (*f*) Postbulbar region. (*g*) Posterior part of male.

Reproductive system. The sexes are separate, and the gonads may be single or paired. In a few parasitic

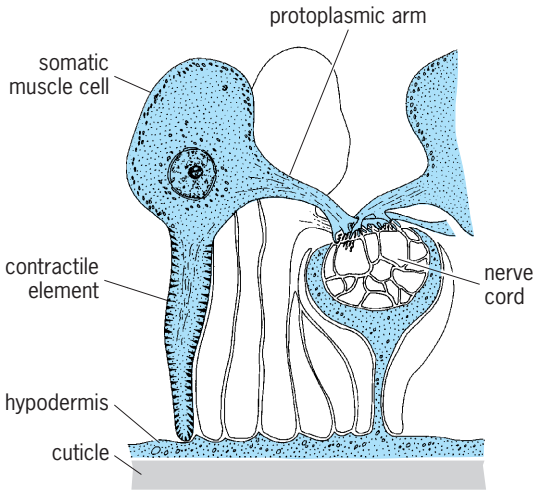

Fig. 2. Cross section of somatic muscle cell with extension to central nervous system.

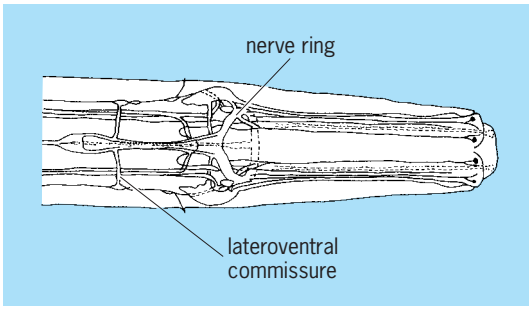

Fig. 3. Nervous system of *Spironoura affine*.

taxa they may be multiple; as many as 32 ovaries have been reported. Female gonads are tubular, with an apical cell that forms the thin nucleated epithelium that covers the entire gonad. Germ cells originate at the blind end or along the sides of the gonad. The ovary is followed by the oviduct and uterus, which connects to the vagina, which opens on the ventral surface of the body. A spermatheca may be present between the ovary and the uterus or between the uterus and vagina. In some, one gonad may be atrophied and function as a spermatheca. A confirmed hermaphroditic male does occur in one taxon of nematode parasites of insects. Hermaphroditic females are reported but unconfirmed; when reported they are allegedly protandric. *See* PROTANDRY; PROTOGYNY.

The male gonad (generally two in Adenophorea and one in Secernentea) opens posteriorly into the ventral side of the rectum, forming a cloaca. Males generally possess paired cuticular copulatory structures called spicules. The spicules are protrusible and presumably when placed in the vagina aid in the transfer of the sperm, but not in the fashion of a penis or intromittent organ. Sperm varies from flagellate to ameboid.

Females may be oviparous or ovoviviparous; that is, eggs hatch within the body and living young exit. The egg is chitinous and may externally be variously sculptured. Nurse cells are unknown, but the zygote contains stored food in the form of proteins; and throughout, fatty and glycogenous globules are evident.

Cleavage of the embryo is bilateral but highly determinate, and follows a modified spiral plan. Following full maturation, the first-stage larva may molt once within the egg shell (Secernentea) or emerge as first-stage larvae (Adenophorea). Larvae are similar in form to the adult, lacking a developed reproductive system and other secondary sexual characteristics. All known nematodes undergo four molts between the first stage and the adult. There is no complete reconstitution of internal organs, but there may be changes in the outline of the esophagus. At each molt the cuticle of the body, esophagus, and rectum is shed. Nematodes reportedly lack any power of regeneration. Cell multiplication after hatching is restricted except within the reproductive tract, intestine, hypodermis, and somatic musculature. *See* CLEAVAGE (DEVELOPMENTAL BIOLOGY).

Adenophorea. In this class the amphids (cephalic chemoreceptors) are postlabial and variable in the shape of the external opening. They may be porelike, slitlike, oval, transversely lenticular, looped (shepherd's crook form), circular, unispiral, or multispiral. The 16 cephalic sensilla may be setiform or papilliform and located postlabially or labially. Somatic setae and hypodermal glands are commonly present, and somatic papillae appear to be universal. The external cuticle consists of all four layers: epi-, exo-, meso-, and endocuticle. The surface of the external cuticle is generally smooth but may have transverse or longitudinal striations; in the subclass Chromadoria it may be elaborately ornate. When present, the excretory system is limited to a single ventral cell lacking collecting tubules. Only rarely is the excretory duct, which opens through a ventromedian pore, cuticularly lined. The esophagus is variable, but is generally either cylindrical or bottle shaped and contains three (Chromadoria) or five (Enoplia) glands. Within the pseudocoel there may be six or more coelomocytes. Rectal glands are rare,

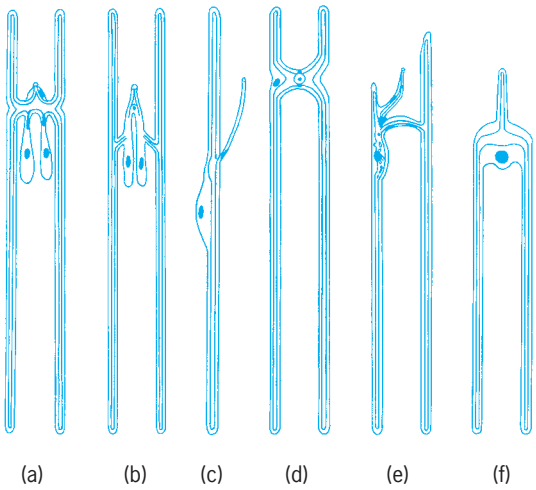

Fig. 4. Types of excretory systems. (*a*) Rhabditoid type. (*b*) Variant of rhabditoid type. (*c*) Tylenchoid type. (*d*) Ascaroid type. (*e*) Cephaloboid type (*f*) Anisakid.

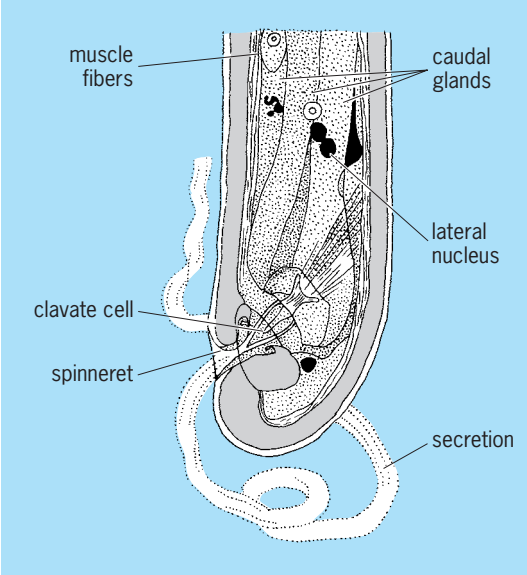

Fig. 5. ***Mononchulus ventralis*** **tail with spinneret details.**

but three adhesive caudal glands that open through a terminal or subterminal spinneret are common (**Fig. 5**). These glands and the spinneret are absent in all obligate parasites. Males, with few exceptions, have two testes, paired spicules, and a single row of ventromedian supplements which may be either papilloid or tuboid. Some males have an additional paired series of supplements. Supplements are often ornate and may function as adhesive organs during copulation. The male tail is generally smooth, and only rarely are lateral cuticular alae reported.

Secernentea. In this class the amphids are most often porelike and located dorsolaterally on the lateral lips or anterior extremity. In a few instances they may be oval or cleftlike, or located postlabially. In all instances the cephalic sensilla are located on the labial region but may be reduced in number, especially among the parasitic forms. On the neck (cervical) region, at the level of the nerve ring, there is usually a pair of sensilla called deirids. Comparable sense organs located on the caudal region are called phasmids (Fig. 1). The external cuticle varies from four to two layers; most Diplogasteria have only the epi- and exocuticle present. The cuticle is almost always transversely striated, and these striations are generally interrupted laterally by longitudinal incisures; these lateral modifications are called the lateral fields. The underlying hypodermis is often multinucleate. Esophagi in the class are variable and diagnostic of each subclass. Rhabditia have a three-part esophagus with a valve in the posterior bulb; Spiruria have a cylindrical or two-part esophagus that lacks bulbs or valves; Diplogasteria have a three-part esophagus, but the valve is anterior in the metacorpus and the posterior bulb is glandular. The three, sometimes more, esophageal glands always open into the corpus. Three to six rectal glands usually open into the anterior part of the rectum or at the junction of the rectum and intestine. The excretory system possesses collecting tubules either paired or limited to one side of the body. The excretory cell opens ventromedianly through a cuticularized duct. Only four to six coelomocytes are reported within the body cavity; in Rhabditia two such cells become associated with the excretory system in adult worms, and these are erroneously called multiple-celled excretory systems. Somatic setae, papillae, and caudal glands are absent in this class; however, males may have caudal papillae. When present, male preanal supplements are paired and often elaborate, but are considered sensory rather than adhesive. Males often possess paired lateral caudal alae, sometimes called the bursa copulatrix.

Life cycle. Reproduction among nematodes is either amphimictic or parthenogenetic (rarely hermaphroditic). After the completion of oogenesis, by either of the above methods, the chitinous egg shell is formed and a waxy vitelline membrane forms within the egg shell; in some nematodes the uterine cells deposit an additional outermost albuminoid coating. Upon deposition or within the female body, the egg proceeds through embryonation to the eellike first- or second-stage larva, but following eclosion the larva proceeds through four molts to adulthood. This sequence represents a direct life cycle, but among parasites more diversity occurs.

Animal parasites often utilize intermediate hosts (including vectors) for the development of the first three larval stages; the fourth- and adult-stage nematode develops in the definitive host. Alternating generations vacillating between free-living cycles and parasitic cycles also occur among plant and animal parasites. There is one special instance among Tylenchida where the nematode, *Fergusobia curriei*, alternates a gametogenetic generation (parasitic in a gallfly, *Fergusonina tillyardi*) with a parthenogenetic generation parasitic on *Eucalyptus camaldulensis*. Among insect parasites alternation of generations in grossly dissimilar host organisms can occur. Because of the great dissimilarity these cases are only seldom recognized.

Bionomics. Nemata comprises the third largest phylum of invertebrates (15,000 species), being exceeded only by Mollusca (100,000 species) and Arthropoda (about 850,000 species). In sheer numbers of individuals they exceed all other Metazoa. As parasites of animals they exceed all other helminths combined. There are some 2000 known plant parasites, 4000–5000 animal parasites, and 8000–10,000 marine, fresh-water, and terrestrial free-living species. Nematodes have been recovered from the deepest ocean floors to the highest mountains, from the Arctic to the Antarctic, and in soils as deep as roots can penetrate.

As parasites of plants and animals they are unequaled by any other metazoan group, including insects. It is within the Secernentea that most of the important parasites of domestic animals, humans, plants, and insects occur. The two most important groups for vertebrate parasitism are in the subclasses Rhabditia and Spiruria. Insect and plant parasitism is most important in the subclass Diplogasteria, especially in the order Tylenchida. A minority of

parasites of plants and animals are included in the class Adenophorea.

The aquatic, fresh-water and marine nematodes are in Adenophorea (Enoplia, Chromadoria). Secernenteans are characteristically terrestrial (soil water), and only rarely are they encountered in a fresh-water or marine habitat; and in these instances it is a secondary invasion by normally terrestrial taxa. The nematodes in these habitats are generally referred to as free living, but they do eat products of decay, bacteria, fungi, algae, and small microfauna including nematodes. The role of nematodes in these environments is unknown; it has been estimated that in undisturbed rangelands their biomass exceeds by two to three times other invertebrates near or on the soil surface.

Most nematodes can sustain periods of anaerobic conditions. None has proved to be capable of completing its entire life cycle, including embryonic development, anaerobically. Many nematodes are capable of anhydrobiosis, and have been proved to survive in such a condition for more than 25 years. Glycogen is one of the chief sources of nematode energy, although stored fatty materials apparently provide the chief source of energy in a few groups such as Tylenchida and Aphelenchida. Armand R. Maggenti

Bibliography. B. G. Chitwood, The designation of official names for higher taxa of invertebrates, *Bull. Zool. Nomencl.*, 15(25/28):860–895, 1958; P. P. Grassé, *Traité de Zoologie*, tome 4, fasc. 2 and 3, 1965; L. H. Hyman, *The Invertebrates*, vol. 3, 1951; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Nematicide

A type of chemical used to kill plant-parasitic nematodes. Nematicides may be classed as soil fumigants or soil amendments, space fumigants, surface sprays, or dips. Soil treatments are commonly used because most plant-pathogenic species spend part or all of their life cycle in the soil, in or about the roots of plants. Nematicides may be liquids, gases, or solids, but on a field scale, liquids are most practical. These materials possess a high vapor pressure and volatilize quickly to act as soil fumigants. Carbon disulfide, CS_2, and chloropicrin were among the first to be developed for this purpose. Their use is fairly restricted because of high cost; besides, CS_2 is highly flammable and explosive and chloropicrin requires a cover or water seal.

The nematicidal properties of 1,3-dichloropropene and 1,2-dibromoethane were discovered between 1943 and 1945; they are the predominant nematicides in use. These chemicals can be applied to the soil by handgun or machine applicators, and require no surface seals or covers beyond culti-packing or rolling the soil after treatment. Since most nematicides are phytotoxic, they must be used in the absence of growing plants. The nematicide 1,2-dibromo-3-chloropropane (DBCP) was an exception. It was low in phytotoxic properties and could be used either for preplant or postplant side-dressing treatments. Unfortunately it was found to be a public health hazard due to carcinogenic/mutagenic properties, and was removed from the market in 1976.

Nonfumigant nematicides have been used as replacements for DBCP. These nonfumigants comprise carbamates such as aldicarb [2-methyl-2-(methylthio) propionaldehyde *O*-(methylcarbamoyl)oxime], or organophosphates such as phenamiphos [ethyl] 3-methyl-4-(methylthio) phenyl (1-methylethyl) phosphoramidate]. They are formulated as granular materials with 10–15% active ingredients. Some are also available as emulsifiable or water-soluble formulations. They are low in phytotoxicity and also low in volatility, and move through the soil slowly or not at all. To reach soil nematodes, the materials must be incorporated into the soil mechanically or carried by irrigation water. Some chemicals are systemic in action, move through the roots to the foliage, and have certain efficacy as insecticides as well.

Space fumigants including methyl bromide are important preplant treatments, especially for perennial tree and vine crops. Deep (up to 3 ft or 1 m) placement of methyl bromide sealed by roller-packing gives long protection. For multiple infestations (that is, nematodes plus oak-root fungus), a thin, 1-mil (25-micrometer) polyethylene cover is required. *See* FUMIGANT.

Methyl bromide is especially useful for sterilizing bags, flats, pots, and other containers or equipment. It has been used to disinfect onion and clover seeds of the stem nematode, *Ditylenchus dipsaci*.

Leaf and bud nematodes of the genus *Aphelenchoides* have been controlled by Parathion and Systox spray applications, but such treatments have very limited usage.

Dip treatments generally employ a combination of hot water and 0.5% formalin solution developed originally for stem nematode in narcissus and Easter lily bulbs and bulbous iris. The addition of a detergent to a similar treatment with 1% formalin has made possible the eradication of the same nematode on garlic cloves.

A slurry of 3-*p*-chlorophenyl-5-methyl-rhodanine has been used for disinfecting rice seed contaminated with nematodes. *See* NEMATA (NEMATODA); PESTICIDE. Dewey J. Raski

Bibliography. V. H. Dropkin, *Introduction to Plant Nematology*, 2d ed., 1989; E. W. Flick, *Agricultural Chemical Products,* 1989; W. T. Thomson, *Agricultural Chemicals*, Book I: *Insecticides*, rev. ed., 1993.

Nematomorpha

A phylum of worms formerly considered to be a class of the phylum Aschelminthes; commonly called the hairworms, and closely allied to the nematodes. The adults are free-living in aquatic habitats, while the juveniles are parasitic in arthropods. The

nematomorphs are found all over the world. They are divided into two classes, the Nectonematoidea and Gordioidea, with a total of 225 species. *See* NEMATA (NEMATODA).

Morphology. The body is long and slender with a maximum length of 5 ft (1.5 m) and a diameter of 0.02–0.12 in. (0.5–3 mm). The females are longer than the males. The anterior end is rounded with a dark pigmented ring and a terminal mouth. The posterior end may be rounded with a terminal cloaca, or it may form two or three lobes in a forklike structure. The body color is yellowish, brown, or almost black. The body wall consists of three layers: an outer, rather thick fibrous cuticle; an epidermis consisting of a single layer of cells; and innermost, a muscle layer with longitudinal fibers only. The surface of the cuticle may be smooth, or rough with rounded or polygonal thickenings called areoles. These may be flat or may form projecting structures, sometimes with bristles, and they may be perforated by pores and canals. Between the areoles run interareolar furrows, often with wartlike structures and bristles. Special natatory bristles are developed in *Nectonema*. The function of the areoles is unknown.

Body cavity. This cavity extends the length of the body. It may be filled with tissue so that only minor spaces are left around the digestive system and the gonads. The digestive tube is always more or less degenerated, and the anterior part is often a solid string of cells. Ingestion of food is impossible, and the intact part of the digestive system seems to be adapted for excretory functions. During the parasitic stage food is obtained through the body surface by means of digestive enzymes.

Nervous system. This consists of a cerebral mass lying ventrally in the head and a ventral nerve cord which originates in the epidermal layer. Little is known of the sensory organs. Probably the bristles and warts of the cuticle have sensory functions. A rudimentary eye is found in *Paragordius*.

Reproduction. The sexes are always separate, and the gonads are paired and stringlike, extending the length of the body. In males the gonads are connected with the cloaca by sperm ducts. In females the ovaries form a large number of lateral diverticula in which the eggs ripen. The oviducts enter the cloaca separately. A sac, which is called the seminal receptacle, extends anteriorly from the cloaca.

During copulation the male coils itself around the female and places a drop of sperm near her cloacal opening. The sperm cells actively enter the seminal receptable. The eggs are laid in water in strings, and the adults die after egg laying. When hatched, the larvae swim to an aquatic arthropod. They penetrate the body wall of the host by means of their characteristic proboscis, which is armed with hooks and three long stylets. The larvae of some species may secrete a special mucus in which they encyst until they are accidentally ingested by the right host, which may be a terrestrial insect. The development in the host lasts some months without any metamorphosis. When mature, the worms leave the host.

Bent J. Muus

Nematophytales

An enigmatic group of fossil plants, in mid-Silurian to lower Upper Devonian rocks, composed of intertwined, branching tubes of two sizes, 10–50 and 1–10 micrometers in diameter.

Prototaxites is stemlike in appearance, ranging from a few millimeters to 3 ft (1 m) in diameter, in which case it may be over 6 ft (2 m) in length with no taper. Its tubes were arranged longitudinally, the small ones having septa at intervals. The septa were perforate, but differ from pores in red algae or in higher fungi. *Nematothallus* looks like flakes of coaly matter, a few millimeters square in size, and is found on rock surfaces. Its larger tubes, 10–50 μm in diameter, were thickened internally by annular rings, and its smaller tubes were 1–5 μm in diameter. Among the tubes were smooth spores, 12–45 μm in diameter, that occasionally show triradiate markings. A pseudocellular layer and cuticle constituted the surface of the plant. *Nematoplexus* had larger tubes, averaging 18 μm in diameter, that were thickened internally by annular or helical deposits, and small tubes 8–9 μm in diameter. Branching of tubes occurred in areas called branch knots. A pseudocellular surface layer is possible. This plant is found only petrified, in Rhynie chert.

Although this group is referred to the algae by some authors, its occurrence in inland swamps, coastal plain deposits, and marine deposits close to shore indicates that the members were terrestrial organisms, unrelated to any known groups, perhaps at an intermediate level between algae and bryophytes. Large tubes with helical or annular thickenings, found as microfossils in earliest Silurian strata, may belong to this group. *See* PALEOBOTANY.

Harlan P. Banks

Nemertea

A phylum of bilaterally symmetrical, unsegmented, ribbonlike worms, frequently referred to as the Nemertinea. They have an eversible proboscis and a complete digestive tract with an anus. There is no coelom or body cavity, and the mesenchyme or parenchyma and the muscle fibers fill the area between the ciliated epidermis and the cellular lining of the digestive tract. *See* ANIMAL SYMMETRY.

Morphology. The nemertineans are mostly less than 8 in. (20 cm) in length, but a few may reach a length of several meters (**Fig. 1**). Many species are brightly colored, sometimes having stripes or transverse bars.

The tubular proboscis, lying above the digestive tract in a cavity, the rhynchocoele, is attached posteriorly to the proboscis sheath by a retractor muscle and is either unarmed or armed with stylets (**Fig. 2**). The proboscis opens anteriorly into a chamber, the rhynchodeum, which in turn opens to the outside above the mouth, through the proboscis pore. The proboscis can be suddenly everted by the

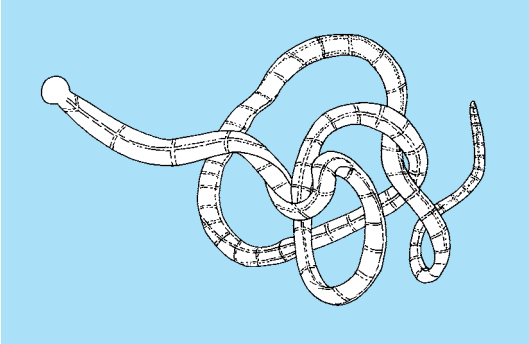

Fig. 1. ***Tubulanus capistratus*** **from the Pacific coast; size ranges from 0.2 to 1.2 in. (5 to 30 mm).** ***(After L. H. Hyman, The Invertebrates, vol. 2, McGraw-Hill, 1951)***

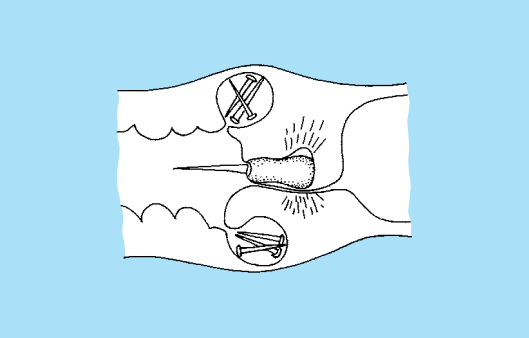

Fig. 2. Stylets of ***Amphiporus griseus*****.** ***(After W. R. Coe, Biology of the nemerteans of the Atlantic coast of North America, Trans. Conn. Acad. Arts Sci., 35:282, 1943)***

contraction of muscles which exert pressure on the fluid in the rhynchocoele. The proboscis of the nemertinean is used for capturing prey, mostly annelid worms, and for defense, for locomotion, or for burrowing.

The nemertineans constitute the most primitive group of invertebrates in which the digestive tract is complete with mouth and anus. In some nemertineans, however, a separate mouth is lacking and the esophagus opens through the proboscis pore (**Fig. 3**).

The nemertineans are the simplest animals with a circulatory system. There are two lateral blood vessels and in some a third, unpaired dorsal vessel. The blood consists of a colorless fluid which may contain blood cells of several types. In species in which the blood is colored, the pigment is present in the cells. There is no heart, but the walls of the principal vessels may be contractile. The walls of all the blood vessels, and the plasma, contain enzymes of the arylamidase type whose function remains unknown. It is possible that they are concerned with protein metabolism or, in intertidal species, with osmoregulation and the maintenance of body volume.

The excretory organs or protonephridia are composed of many tubules ending in flame bulbs (**Fig. 4**). These are hollow, urn-shaped cells provided with vibratile cilia. On each side the flame bulbs are often closely associated with the lateral blood vessel, and their tubules are united to a common tubule which opens at a lateral nephridiopore. Peculiar large cells called arthrocytes may surround the flame bulbs and tubules and are assumed to be excretory, since they readily take up vital stains.

The nervous system has a pair of cerebral ganglia forming the brain as well as two longitudinal nerve cords and many smaller nerves. The ganglia and lateral cords may contain unusually large neurochord

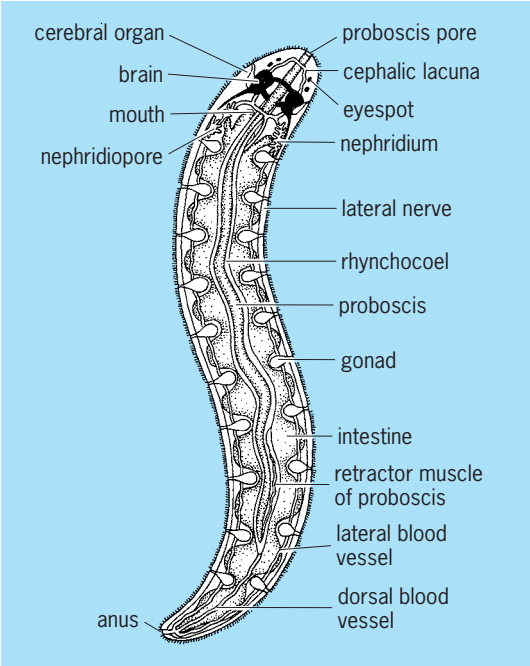

Fig. 3. Internal structure of a nemertinean.

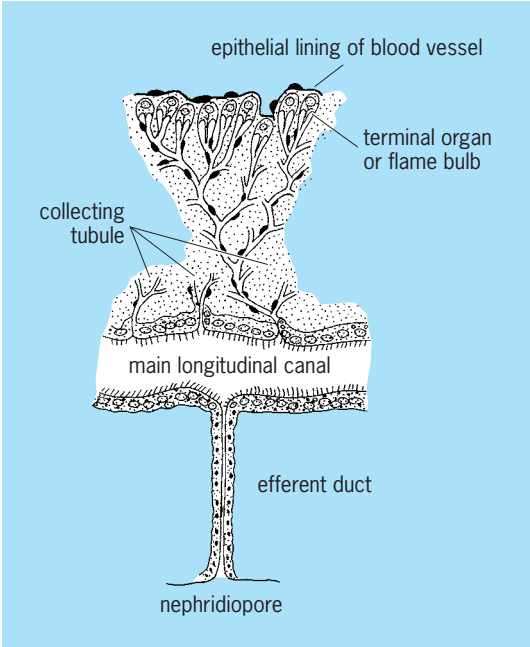

Fig. 4. Diagram of the multiple protonephridium of a nemertinean. ***(After W. R. Coe, Unusual types of nephridia in nemerteans, Biol. Bull., 58(3):214, 1930)***

cells. In the epidermis there are scattered sensory nerve cells, probably tactile. Chemotactile organs are situated in a pair of anteriorly placed cephalic grooves or in a flask-shaped protrusible frontal organ. A few to many simple eyes, or ocelli, may be present in front of the cerebral ganglia. Statocysts, the organs of equilibrium, are rare. Two blind canals, the cerebral organs, are invaginated from the epidermis and are closely associated with the cerebral ganglia. These organs, probably chemosensory, open through pores in the cephalic grooves or on the body surface.

There are no special respiratory organs; respiration occurs through the body surface.

Nemertineans are usually either male or female, but a few individuals have both sex organs. The ovaries or testes open by short ducts to the exterior. Fertilization occurs outside the body in many species but may be internal in certain forms.

Embryology. Cleavage of the fertilized egg, or zygote, is spiral, with the cells of the lower quartet of the eight-cell stage rotated slightly, to lie in the furrows between the cells of the upper quartet. Development is determinate, with the potentialities for future development of the embryo determined or fixed in the zygote before cleavage begins. Isolated cells of the two-cell stage result in dwarf larvae. Isolated cells of later stages result in deficiencies. *See* CELL LINEAGE; CLEAVAGE (DEVELOPMENTAL BIOLOGY).

After early cleavage, the development in certain nemertineans may be direct, that is, without a larva, the embryo emerging from the egg membranes as a minute, ciliated worm. In others, the gastrula becomes a free-swimming, helmet-shaped, ciliated pilidium (**Fig. 5**), formed by the downward growth of two ciliated lobes at the sides of the mouth and having an apical tuft of cilia. In still other nemertineans, the gastrula remains inside the egg membranes and becomes an oval, ciliated larva known as Desor's larva, which lacks the apical tuft and the oral lobes.

Both the pilidium and the Desor's larva metamorphose into an adult worm by means of the invagination of seven or eight ectodermal plates. These flattened, invaginated sacs spread and finally fuse, thus separating the larval ectoderm and the thin amnion from the newly invaginated ectoderm of the larval worm within. The larva completes development with the formation of the anus and other organs. It then sheds the larval ectoderm and amnion and emerges as a young worm.

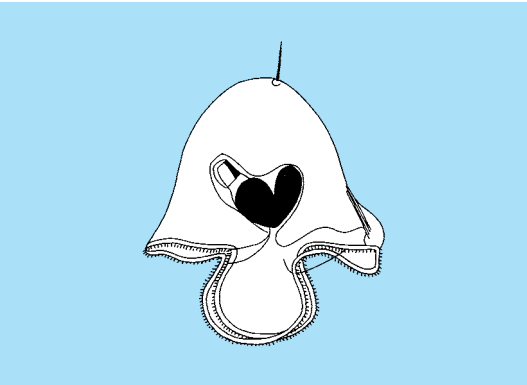

Fig. 5. Pilidium of a nemertinean. (*After H. G. Bronn, ed., Klassen undordnungen des Thier-Reichs, vol. 4, 1903*)

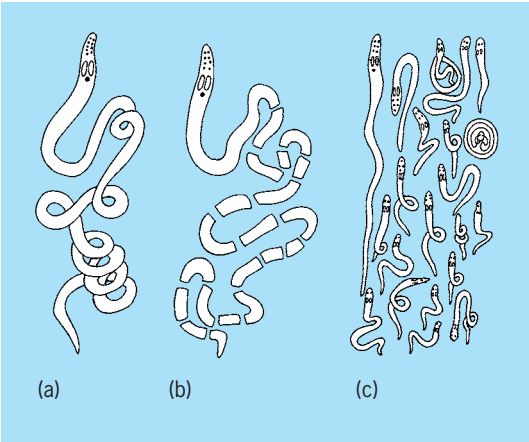

Fig. 6. Steps *a–c* in fragmentation in *Lineus vegetus*. (*After W. R. Coe, Revision of the nemertean fauna of the Pacific coasts of North, Central, and northern South America, Allan Hancock Pacific Expeditions, vol. 2, University of Southern California Press, 1940*)

Asexual reproduction and regeneration. Certain nemertineans have the ability to multiply asexually by fragmentation of the body. Each fragment containing a part of the lateral nerve cords can regenerate into a complete worm (**Fig. 6**). Other nemertineans, when irritated by handling or unfavorable conditions, fragment the body or evert the proboscis so that it breaks off. The anterior region, including the foregut, can regenerate a new proboscis and a new posterior end. In the process of anterior regeneration the wound is covered by migrating epidermal cells. Then, in the mass of mesenchyme cells which forms below the closure, three groups of cells appear, two lateral groups reforming the cerebral ganglia and a median group reforming the proboscis. Regeneration of the posterior end occurs by a lengthening of the body through the differentiation of mesenchyme cells. *See* REGENERATIVE BIOLOGY; REPRODUCTION (ANIMAL).

Ecology. The nemertineans are mostly marine, bottom-dwelling worms, found in greatest numbers along the coasts of northern temperate regions. They live under stones, among the tangled masses of plants, in sand, mud, or gravel, and sometimes form mucus-lined tubes. A few are pelagic, fresh-water, or terrestrial. Certain species are commensal with other animals, but none can be regarded as parasitic in a strict sense.

Phylogenetic relationships. The Rhynchocoela are related to the Platyhelminthes and probably have evolved from the same ancestral stock which gave rise to that phylum. They resemble the flatworms in having the region between the epidermis and the gut filled with mesenchyme, in the arrangement of the nervous system, in the structure of the eyes, in the occurrence of ciliated grooves on the head, and in showing spiral cleavage. The proboscis and the

cerebral organs may be regarded as derived from certain organs in flatworms. Serological tests have indicated that nemertineans are closer to the platyhelminths than to annelids.

That the Rhynchocoela represent the most highly organized acoelomate animals is indicated by the circulatory system, the presence of an anus, and the specialization of the epidermis. All groups of animals more complex than the nemertineans have some kind of cavity, a pseudocoele or coelom, between the body wall and the gut, instead of solid mesenchyme. *See* COELOM.

Classification. The phylum Rhynchocoela, containing about 550 known species, is divided into two classes. In the class Anopla the mouth is posterior to the brain, the nerve cords lie under the epidermis or in the muscle layers of the body wall, and the proboscis is unarmed. The two orders of the Anopla are separated on the basis of the musculature and the character of the dermis, which is a connective tissue lying under the epidermis. The muscles of order Palaeonemertini are arranged in two or three layers (**Fig. 7***a*); if three, the innermost layer consists of circular fibers. The dermis is gelatinous. The order contains such genera as *Tubulanus*, *Carinoma*, and *Cephalothrix*. In the second order, the Heteronemertini, the muscles are in three layers (Fig. 7*b*) and the dermis is fibrous. The order contains, among others, the genera *Lineus*, *Cerebratulus*, and *Micrura*.

In the second class, the Enopla, the mouth is anterior to the brain, the nerve cords are internal to the muscles, and the proboscis is often armed with stylets. Of the two orders, the first, the Hoplonemertini, has a proboscis armed with one or more stylets (Fig. 7*c*) and has a straight intestine with paired lateral diverticula. Representative genera are *Emplectonema*, *Carcinonemertes* (on the gills and egg masses of crabs), *Amphiporus*, *Testrastemma*, *Prostoma* (in fresh water), *Geonemertes* (on land near the sea in tropical and subtropical regions), and *Nectonemertes* (pelagic). The second order, the Bdellomorpha (Bdellonemertini), has an unarmed proboscis, a sinuous intestine without diverticula (Fig. 7*d*), and a posterior adhesive disk. The only genus, *Malacobdella* (**Fig. 8**), is commensal chiefly in the mantle cavity of marine clams, where it feeds on plankton brought in by the ciliary current. *See* ANOPLA; ENOPLA.

Arthur G. Humes; J. B. Jennings

Bibliography. R. Gibson, *Nemerteans*, 1972; R. Gibson and J. B. Jennings, *Compar. Biochem. Physiol.*, 23:645–651, 1967; L. H. Hyman, *The Invertebrates*, vol. 2, 1951; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

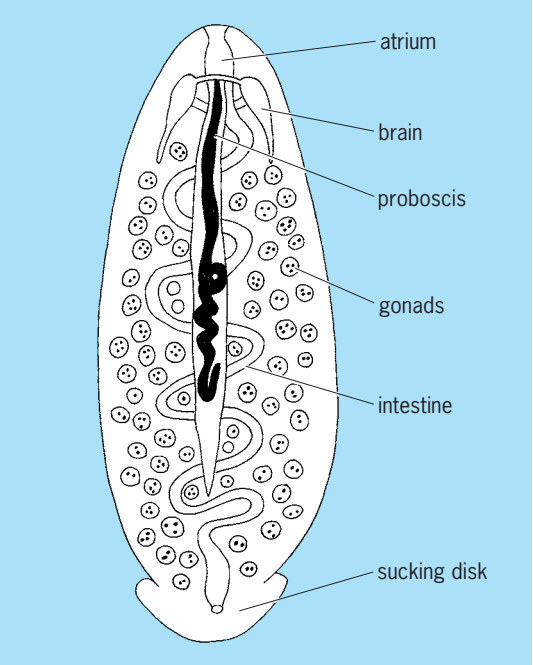

Fig. 8. ***Malacobdella grossa*. (*After W. R. Coe, Biology of the nemerteans of the Atlantic coast of North America, Trans. Conn. Acad. Arts Sci., 35:308, 1943*)**

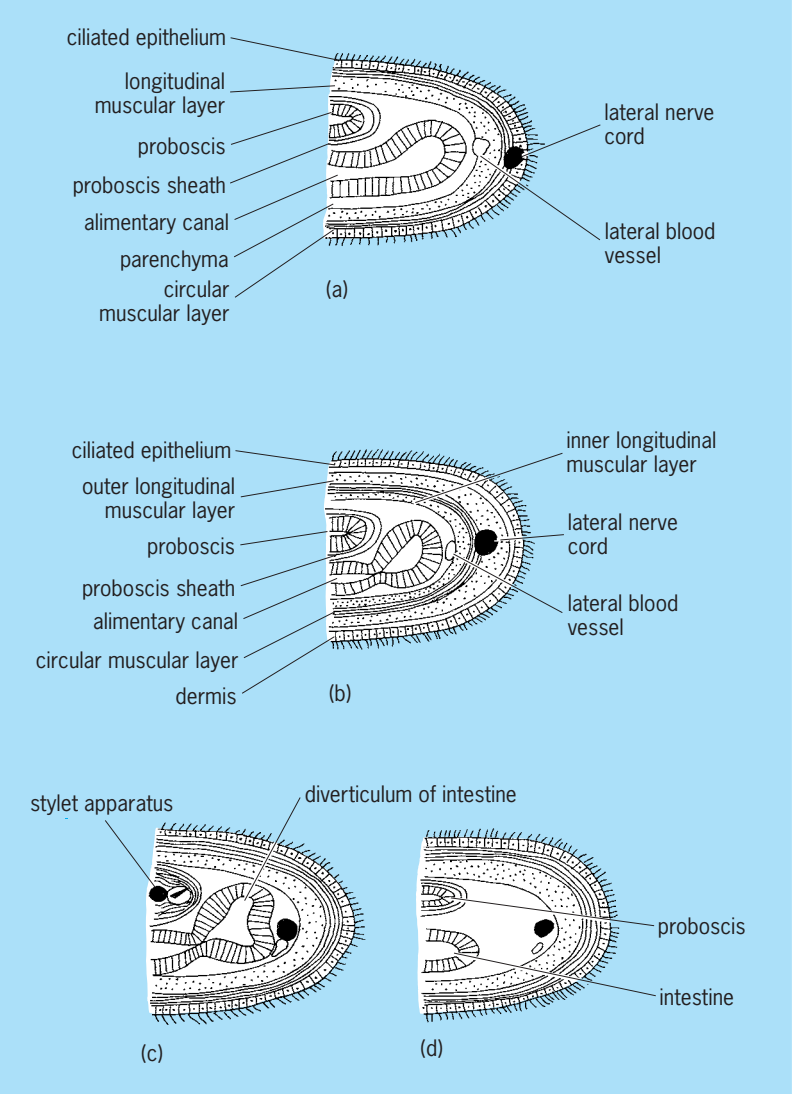

Fig. 7. Transverse sections of the body in the four orders of nemertineans, showing the arrangement of the muscular layers and the position of the lateral nerve cords and lateral blood vessels. (*a*) Palaeonemertini and (*b*) Heteronemertini constitute the class Anopla; (*c*) Hoplonemertini and (*d*) Bdellonemertini constitute the class Enopla. (*After W. R. Coe, Biology of the nemerteans of the Atlantic coast of North America, Trans. Conn. Acad. Arts Sci., 35:145, 1943*)

Neodymium

A metallic chemical element, Nd, atomic number 60, atomic weight 144.24. Neodymium belongs to the rare-earth group of elements. The naturally occurring element includes six isotopes. The oxide, Nd_2O_3, is a light-blue powder. It dissolves in mineral acids to give reddish-violet solutions. For properties of the metal *see* PERIODIC TABLE; RARE-EARTH ELEMENTS.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb

actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

The salts have found application in the ceramic industry for coloring glass and for glazes. The glass is particularly useful in goggles used by glass blowers, since it absorbs the intense yellow D line of sodium present in the flame. The element has found commercial application in the manufacture of lasers.

Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

Neogastropoda

The most highly specialized order in the subclass Prosobranchia (phylum Mollusca, class Gastropoda). Neogastropods have simplified pallial and cardiac structures involving complete separation of genital from renal organs, and a "half-gill" (that is, a one-sided comb-shaped or pectinibranch ctenidium) with its axis and major blood vessels fused to the mantle wall. The order comprises mainly marine carnivores and carrion feeders, all with a long extensible proboscis bearing a flesh-tearing radula. Teeth per row are reduced in such radulae (hence the former ordinal name, Stenoglossa), usually to the rachiglossan, or three-toothed, pattern, but in two families to the toxiglossan type with a single tooth per radular row. More efficient hydrodynamically with their simplified mantle cavity and fused ctenidial axis, neogastropods are not limited to clean waters over hard substrata (as are the archaeogastropods) but have successfully invaded all areas of the seashore and sea bottom, whether covered with sand, silt, or mud.

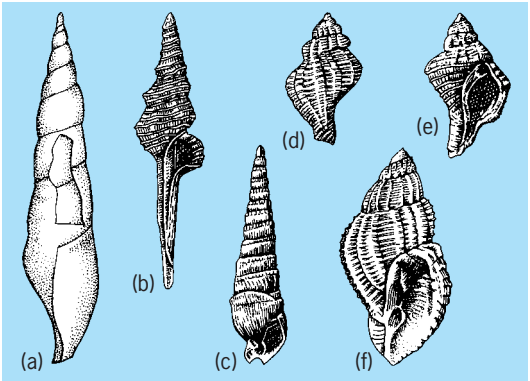

Neogastropoda. (*a*) Ordovician species of *Subulites*, widespread Ordovician and Silurian genus. (*b*) *Falsifusus*, early Tertiary genus. (*c*) Miocene species of familiar existing *Terebra* (Tertiary-Recent). (*d*, *e*) *Urosalpinx* (Tertiary-Recent). (*f*) *Cancellaria* (Tertiary-Recent). (*After R. R. Shrock and W. H. Twenhofel, Principles of Invertebrate Paleontology, 2d ed., McGraw-Hill, 1953*)

The neogastropod proboscis leads internally to a much simpler gut than that of more primitive gastropods, which are mostly herbivorous or filter-feeding with elaborate sorting areas and recycling processes. In neogastropods, ingested flesh from prey or carrion receives straightforward digestion which is wholly extracellular. The pallial sense organ, or osphradium, which in more primitive mollusks was involved in mechanical detection of sediment for the mantle cavity, has become a chemoreceptor used in hunting down suitable prey by waterborne "scent." The osphradium is located at the inner end of a characteristic inhalant siphon, and neogastropod shells all show a siphonal notch (see **illus.**).

Neogastropods occur in all depths of the world's oceans from the tropics to polar waters, and there are at least 6000 species, mostly in four important superfamilies. The larger whelks of the superfamily Buccinacea, in such genera as *Busycon*, *Colus*, *Neptunea*, *Fasciolaria*, and *Buccinum*, are found from the shallow sublittoral and continental shelves down to depths of 9800 ft (3000 m). The flesh of many whelk species provides human food, and almost all species have been used in commercial longline fisheries as resilient and attractive bait. This superfamily also includes the ubiquitous and cosmopolitan mudsnails of the genera *Nassarius* and *Ilyanassa*, and the beautiful tulip shells and spindles (*Fasciolaria* and *Fusinus*).

The smaller tingles, dog whelks, and oyster drills of the superfamily Muricacea are the abundant neogastropod predators in inshore and intertidal waters. Certain species of the genera *Ocenebra*, *Urosalpinx* and *Rapana*, known as tingles and oyster drills, are the most destructive pests of clam beds and oyster farms, and use the proboscis-borne radula to bore through the shells of their prey. The dog whelks, species of genera such as *Nucella*, *Eupleura* and *Thais*, also prey extensively on barnacles; the extensible proboscis is used to force open the opercular plates (rather than bore an access hole as

with bivalves) and then to browse on the internal tissues of the prey. Larger species of muricid rock shells, in such genera as *Trophon*, *Murex*, *Forreria*, and *Pteropurpura*, may have extensive shell sculpture in the form of radiating spines and wings. Despite extensive shell collecting of these forms, their biology has been little studied.

A third important superfamily, Volutacea, encompasses more beautiful, much collected shells belonging to such family groups as olivids, mitrids, volutids, marginellids, and turbinellids. The most specialized neogastropods are the tropical toxoglossans (superfamily Conacea) belonging to the families Conidae, or cone shells, and Terebridae, or auger shells. Both groups have been prized by shell collectors for centuries. In them, the radula is reduced to a series of separate darts, which are used in association with a modified salivary gland secreting a neurotoxin for the capture of large prey animals, both invertebrate and vertebrate. Many species of cones respond only to particular species of live, healthy prey by aiming the proboscis, striking with a single everted tooth, injecting toxin, and then feeding upon the paralyzed prey. Three main categories of cones, each with approximately proportioned harpoonlike teeth and appropriate toxins, feed respectively on other snails, on annelid worms, and on fishes. A number of human fatalities have resulted from two species of fish-eating cones in the East Indies.

Despite their large number of species and diversity of habitats, the neogastropods show more anatomical uniformity (efficient mantle cavity, inhalant siphon with chemoreceptive osphradium, extensible proboscis, stenoglossan radula, and simple carnivore gut) than is found in any of the other major orders of gastropods. *See* GASTROPODA; MOLLUSCA; PROSOBRANCHIA. W. D. Russell-Hunter

Bibliography. V. Fretter and A. Graham, *British Prosobranch Molluscs*, 1962; W. D. Russell-Hunter, *A Life of Invertebrates*, 1979; K. M. Wilbur (ed.), *The Mollusca*, vols. 1–8, 1983–1984.

Neognathae

One of the two recognized superorders making up the subclass Neornithes of the class Aves. They are characterized as flying birds with fully developed wings and sternum with a keel, caudal vertebrae fused into a pygostyle, and absence of teeth in both jaws, or modifications of these conditions in secondary flightless birds.

This superorder includes all living birds and all known fossil birds since the Late Cretaceous; only the ancestral Jurassic *Archaeopteryx* and the specialized Cretaceous *Hesperornis* and its allies do not belong to the Neognathae. Several previously recognized superorders—Ichthyornithes for *Ichthyornis* and its allies, Impennes for the penguins, and Palaeognathae for the living flightless ratites—have been merged with the Neognathae. Merging of these four superorders reflects the lack of knowledge of relationships of avian orders and implies that, as far as is known, the orders included in each former superorder are not more closely related to each other than to those orders placed in other superorders, or that the particular characteristics of each group are no more specialized than those found in individual orders of birds. Evidence suggests strongly that the ratites constitute a monophyletic assemblage, so that the Palaeognathae may be revived if other equivalent assemblages of neornithine orders can be recognized. *See* ARCHAEOPTERYX; AVES; ODONTOGNATHAE; RATITES. Walter Bock

Neognathostomata

A superorder of Echinoidea, subclass Euechinoidea. These invertebrates are characterized by having a rigid, exocyclic test and a lantern or jaw apparatus developed sometime during the life history and usually persisting into the adult stage. The included orders are the Clypeasteroida, Cassiduloida, Neolampadoida, and Oligopygoida. *See* ECHINODERMATA; ECHINOIDEA; EUECHINOIDEA; HOLECTYPOIDA; NEOLAMPADOIDA; OLIGOPYGOIDA. Howard B. Fell

Neogregarinida

An order of the protozoan subclass Gregarinia, class Telosporea, subphylum Sporozoa. All gregarines are parasites of the digestive tract and body cavity of invertebrates or lower chordates; their mature trophozoites (vegetative stages) live outside the host's cells and are large. The Neogregarinida are thought to be relatively advanced gregarines which live in insects. The most primitive gregarines, the Archigregarinida, differ from the typical gregarines, the Eugregarinida, in having asexual multiple fission (schizogony) before gamonts (cells which will produce gametes) are formed. The Neogregarinida are thought to have acquired this trait secondarily and to have developed from the Eugregarinida. *See* ARCHIGREGARINIDA; EUGREGARINIDA.

There are only about 29 species of about 12 genera, which have been divided into 4 families. The largest genus is *Ophryocystis*, which has species.

Another genus is *Mattesia*, which has 2 species, one of which is *M. dispora* of the flour moth (*Ephestia kuehniella*). In this species the intracellular schizonts occur in the fatty tissues. The schizonts are of two types, with small or large nuclei. Those with small nuclei apparently produce merozoites which turn into schizonts with large nuclei. These produce merozoites which turn into gamonts that leave the cells of the fatty tissues, come to lie in the hemocoel, join in pairs (syzygy), grow, encyst together, and proceed to form two gametes each. The gametes fuse to form two zygotes within the gametocyst, and each zygote in turn forms an oocyst that contains eight sporozoites. It is not clear how they get to new flour moths. *See* GREGARINIA; PROTOZOA; SPOROZOA; TELOSPOREA. Norman D. Levine

Neolampadoida

A group of small, deep-water cassiduloid echinoids with neotenous characteristics, treated as an order by some workers; possibly polyphyletic. The presence of bourrelets and phyllodes, the elongate first ambulacral plates, the undifferentiated tuberculation, and undifferentiated posterior interambulacral plating all indicate their relationships lie with cassiduloids. The only character shared by members of this group is the lack of petals (they have simple ambulacral pores only). Other characteristics, such as apical disc plating, are varied, indicating at least two independent origins from shallow-water cassiduloids.

There are seven genera, each monospecific. Five are living today and are usually found at depths of 430–1280 ft (135–400 m). A Miocene species and an Upper Eocene species are also known. *See* ECHINODERMATA; NEOGNATHOSTOMATA. Andrew B. Smith

Bibliography. J. W. Durham and C. D. Wagner, Neolampadoids, in R. C. Moore (ed.), *Treatise on Invertebrate Paelontology,* pt. U: *Echinodermata 3,* pp. U628–U630, 1966.

Neolithic

The period of prehistoric culture whose basic defining attributes are the emergence of agriculture, animal domestication, and sedentary farmsteads or villages. This definition has evolved over the last century from the original characterization of this period based on the appearance of polished stone axes. By 1865, when John Lubbock published *Prehistoric Times*, two types of Stone Age had been recognized in Europe: *période de la pierre taillée* (period of chipped stone implements) and *période de la pierre polie* (period of polished stone implements). Lubbock termed the former Palaeolithic and the latter Neolithic. Subsequently, it was realized that the definition of this period based on a single artifact type was spurious, since Neolithic peoples also continued to make chipped stone tools. A more comprehensive view developed that saw the Neolithic as characterized by pottery manufacture, agriculture, livestock, and settled villages, but without the use of metals. Thus the Neolithic formed the final Stone Age precursor to the Bronze Age and the Iron Age in the classic northern European prehistoric sequence, which was soon extended throughout most of Eurasia. *See* PALEOLITHIC.

Modern concepts. Research dating back to the 1950s in the Near East has demonstrated that the domestication of plants occurred prior to the onset of pottery production. At the same time, it became clear that the nonagricultural Ertebølle culture of southern Scandinavia and the Jomon culture of Japan produced pottery. Thus pottery production has generally ceased to be considered a defining trait of the Neolithic. The discovery of the early use of native and smelted copper for ornaments and tools at Neolithic sites in Europe and the Near East also forced a reconsideration of whether metal use excluded a culture from the Neolithic. In some areas, such as southern Europe and the Near East, a distinct Copper Age designation was introduced to account for this discrepancy. These taxomonic difficulties have led some archeologists to abandon the use of the term Neolithic, but it continues in use today as a convenient designation for early agricultural societies.

As the term Neolithic is presently used, it refers specifically to prehistoric societies in Europe, Asia, and northern Africa that derived the majority of their diet from agriculture and livestock and that lived in sedentary communities, either dispersed farmsteads or villages, but that did not yet know the use of alloyed metals. Outside of this area, the term Neolithic is rarely used, although clearly in most other parts of the world there was also a transition from hunting and gathering to agriculture at some time in prehistory. Archeologists working in the New World frequently use the term "Formative" to refer to early sedentary agricultural communities similar to those of the Old World Neolithic. These variations in nomenclature and tradition should not obscure the fact that in the last 10,000 years humans have gone from being exclusively hunters and gatherers to deriving virtually all of their food supply from cultivated plants and domesticated animals.

For the purposes of this article, "Neolithic" will be expanded to include a global consideration of the origins and dispersal of agriculture and sedentary life, what might be called a neolithic cultural pattern. The British prehistorian V. Gordon Childe (1892–1957) once wrote of a "Neolithic Revolution" to characterize the emergence of sedentary agricultural societies. While the transition from hunting and gathering to farming and herding had immense consequences for subsequent human cultural development, it was hardly revolutionary in most areas. Instead, the Neolithic developed either through the local domestication of plants and animals by people already familiar with their reproductive biology, by the movement of farming peoples into areas where agriculture had not been previously practiced, or through the adoption by hunter-gatherers of crops and livestock that had been domesticated elsewhere.

Domestication of plants and animals. With the exception of the dog, which was domesticated by late Pleistocene hunter-gatherers in many parts of the world, the domestication of plants appears to have preceded that of animals. Although many parts of the world can claim to have been the location of the domestication of some food plants, the regions where the crop species which underlie the major agricultural systems of the world were domesticated are fairly well delimited. In 1971 the American botanist Jack Harlan identified three relatively small regions where the domestication of major crop species can be localized, along with several larger areas in which other important species were domesticated in a somewhat more diffuse geographical pattern. Research since then has amended these slightly, but the general outlines of this pattern remain the same. The three foci of domestication

include the "Fertile Crescent" area of the Near East (wheat and barley), the Huanghe and Yangzi valleys of China (millet and rice), and Mesoamerica (maize and beans), while the broader areas of where important domestication also occurred include northern and Andean South America (potatoes and other root crops), northern sub-Saharan Africa (sorghum), and southeast Asia and many Pacific islands (various tree and root crops). Other research has added another area to this list. The American archeologist Bruce D. Smith, building on the work of earlier botanists and archeologists, made a compelling case for the indigenous domestication in eastern North America of a complex of weedy plants that include chenopod, sunflower, and marsh elder.

Near East. The Fertile Crescent extends in an arc from the Jordan Valley in the Levant across the headwaters of the Tigris and Euphrates rivers in northern Syria and Turkey and down the flanks of the Zagros Mountains of Iraq and Iran. Recent research points toward a part of this region known as the Levantine Corridor, an inland zone stretching for about 40 km (25 mi) between the Damascus Basin and the lower Jordan Valley, as being potentially the location of the earliest cultivation of wheat and barley. At the end of the Pleistocene, this would have been one of the moister parts of the Levant, and it also contained stands of wild wheat and barley. The Natufian hunter-gatherers of this region were intensive collectors of wild cereals, and about 10,000 years ago they became the first Neolithic farmers. Einkorn wheat, emmer wheat, and barley are the three principal founder crops of Near Eastern agriculture. They are distinguished from their wild counterparts primarily by the nature of the rachis, the stem which holds the grain to the ear. Wild wheats and barley have a brittle rachis, which shatters at ripening to disperse the grain and thus propagate the plant. In domestic forms, the rachis is tougher and thicker, which permits the harvesting of the whole year by cutting with sickles. Such a method would have selected for plants with tougher rachises. It may have taken a relatively short time, on the order of a few centuries, to produce populations of wheat and barley that were dependent on humans for their propagation. At sites such as Netiv Hagdud in the lower Jordan valley, early farmers lived in circular and oval structures built of mud bricks upon stone foundations and stored their grain in subterranean silos.

Just as wheat and barley were the founder crops of Near Eastern agriculture, sheep and goat can be considered the founder animals of ungulate domestication around 9000 years ago. As the American zooarcheologist Richard Meadow pointed out, the domestication of animals requires a profound change in human attitudes toward them, a shift from the importance of the dead animal as a source of meat to the living animal as a member of a breeding population. Goats appear to have been domesticated first in the Zagros Mountains, while sheep at sites in the upper Tigris and Euphrates valleys underwent the progressive diminution in size associated with domestication a few centuries later. Cattle and pigs appear to have been domesticated around 8000 years ago, with evidence pointing toward Anatolia as the likely location.

China. The Huanghe (Yellow) river and Yangzi river valleys of China were two distinct foci of early plant domestication. Each region has its distinctive founder crop: foxtail millet (*Seteria italica*) in northern China and rice (*Oryza sativa*) in southern China. The wild precursor of foxtail millet is a common Eurasian weed, green bristlegrass. As with wild and domestic forms of wheat and barley, green bristlegrass and foxtail millet differ in their method of seed dispersal, with the wild form shattering and the domestic form retaining its grains. Deep deposits of loess, a dry but fertile windblown soil, along the Huanghe and its tributaries provided ideal growing conditions for millet. The earliest Neolithic culture of the Huanghe valley is called Peiligang, after a site in Henan province.

About 600 km (360 mi) to the south, in the Yangzi drainage, the Hupei basin is a lowland zone of lakes, wetlands, and meandering streams. Recent excavations at Pengtoushan have revealed a large settlement of sedentary peoples dated between 8500 and 7800 years ago. Large amounts of rice indicate that this plant was either intensively collected or in the early stages of cultivation. On the coastal plain north of Shanghai, in another area of marshes and lakes, the waterlogged site of Hemudu has yielded convincing evidence for the use of domesticated rice before 7000 years ago.

The pig and the chicken were the most important domesticated animals in the Chinese Neolithic, supplemented soon after with the water buffalo. The pigs were probably domesticated from local wild populations, while the chicken is believed to be derived from the red jungle fowl of southeast Asia.

Mesoamerica. Mesoamerica is a term used by anthropologists and archeologists to mean the area of Mexico and the adjacent northern part of central America in which complex urban societies such as the Maya and Aztec emerged in later prehistoric times. It also appears that this was the region in which maize and beans, the major founder crops of many New World agricultural systems, were first domesticated. Unlike wheat and barley, maize is a hybrid and has no obvious wild progenitor, so its origins are hotly debated among paleobotanists. Research in the 1980s and 1990s pointed toward the grassy weed called teosinte as the likely ancestor of maize.

The earliest maize cobs have been found in the dry caves of the interior uplands of Mesoamerica, such as those of the Tehuacán valley, where they were desiccated and preserved. Biochemical comparison of teosinte and maize suggests that interior lowlands, such as the Balsas river valley in central Mexico, may be more likely candidates for the location of the earliest maize domestication, although little archeological research has been done thus far. The initial dating of the Tehuacán maize cobs in the 1960s pointed toward their relatively early domestication, perhaps over 7000 years ago, but recent redating using the accelerator mass spetrometry (AMS) technique has

established that the earliest maize at Tehuacán is no earlier than 5500 years ago.

The other major New World field crop, the common bean (*Phaseolus vulgaris*), appears to have been domesticated independently in at least two locations, Mexico and the southern Andes. Recent AMS dating of the oldest specimens suggests that beans were domesticated more recently than previously thought, and the maize-bean association that is so common in later prehistoric and modern fields may not have been used during the early stages of the cultivation of these plants.

Eastern North America. It was long presumed that North America was not involved in early plant domestication, since maize, beans, and squashes spread northward from Mesoamerica at a fairly slow rate. By the mid-1980s, however, evidence had amassed that four major seed plants—sunflower, squash, marsh elder, and chenopod—had been brought under sufficient human control and selection to produce morphological changes that could be considered domestication. Bruce D. Smith argued that these were weeds which aggressively colonized disturbed habitats, such as point bars and other floodplain locations, and were then transplanted and cultivated by humans in similar settings starting around 3500 years ago. Although this indigenous agricultural system did not survive into the present, it supported the development of farmsteads and hamlets in the interior valleys of eastern North America for over two millennia.

Other geographic areas. Several other parts of the world made important contributions to the list of domesticated plants and animals during the last 5000 years and perhaps earlier. In a broad band across the northern half of Africa, sorghum, African rice, and pearl millet were early domesticates between about 4000 and 3000 years ago. In northern South America, evidence from pollen and phytoliths (small diagnostic silica structures that form in the veins of plants) indicates that a variety of root and tuber crops, including manioc and arrowroot, and tree crops were brought under cultivation, perhaps at a very early date. In the Andes, the potato emerged as an important food source, guinea pigs were domesticated as a meat source, and llama and alpacas were used both as pack animals and for food. Taro and yams, along with a variety of tree crops, were managed, cultivated, and eventually domesticated in a broad region of southeast Asia and the nearby Pacific islands.

Agricultural dispersals. After the initial domestication of plants and animals, the Neolithic farming economy spread quickly to adjacent areas, many of which had been previously inhabited by hunter-gatherers. A major research challenge for archeologists is to determine whether the dispersal of farming, herding, and settled farmstead or village life was the result of the actual movement of agricultural peoples into areas that had hitherto been sparsely populated, or the adoption of domesticated plants and animals by hunter-gatherers who then abandoned their foraging way of life. Two major agricultural dispersals are that of the wheat-barley-livestock complex from the Near East into Eurasia and the maize-beans-squash complex from Mesoamerica into North America.

Eurasia. Within a relatively short time after the establishment of the first Neolithic settlements in the Fertile Crescent, communities practicing agriculture appeared in central Anatolia. Around 8500 years ago, they spread to Greece where they colonized the fertile alluvial soils of regions such as Thessaly. The earliest European farming communities share many traits with their Near Eastern precursors, including the use of mud brick or adobe reinforced with wood for house construction. From Greece, agriculture spread along two main paths: west through the Mediterranean basin, eventually reaching Spain and Portugal, and north through the Balkans, eventually turning northwest along the Danube drainage. In both cases, crops, livestock, and often people moved quickly, so that within 1500 years of the first appearance of agriculture in Greece it had reached the Atlantic Ocean and the English Channel.

In some areas, such as along the Mediterranean coast, in the interior uplands of central Europe, and on the coastal lowlands of northern and western Europe, it appears that agriculture spread through the adoption of domestic plants and animals by indigenous hunter-gatherers. Elsewhere, such as on the alluvial soils of the floodplains of interior central Europe, a strong argument can be made for the establishment of Neolithic communities through the actual dispersal of farming peoples. In these interior valleys, the farmers of the Linear Pottery culture built farmsteads characterized by timber structures up to 30 m (100 ft) in length, the largest free-standing buildings in the world at this time. In addition to wheat and barley, cattle were important elements in the Linear Pottery economy, and there is evidence that indicates that they were used for milk as well as for meat.

Agricultural communities also spread east and northeast from the Fertile Crescent to the oases of central Asia, the Iranian Plateau, and beyond. At Mehrgarh in Pakistan, the initial farming occupation took place around 8000 years ago. Around the same time, Neolithic settlements were established at Djeitun and neighboring oases in Turkmenistan. Rectangular mud-brick houses constituted the primary domestic architecture of these hamlets.

North America. From its original area of maize domestication in Mesoamerica, agriculture spread northward to the northern border of the Sonoran Desert. Although some archeologists dissent, the weight of evidence points toward the adoption of maize and beans by the indigenous inhabitants of the Colorado Plateau and the Mogollon Highlands rather than a northward migration of farmers from Mexico. The first appearance of maize in this area is now AMS-dated to between 3500 and 4000 years ago, which revises earlier dating which had suggested that it occurred several millennia earlier. Domesticated plants were integrated at first into the existing subsistence pattern, and it was not until about 2000 years ago that people in the area became dependent on them for most of their food.

In eastern North America, maize spread gradually to the north and east. It made its first appearance in the Mississippi valley about 2000 years ago, based on AMS dates from the Holding site located near St. Louis. It remained a fairly minor component of the diet until about A.D. 900, when stable carbon isotopes in human bones show a dramatic increase in maize consumption throughout eastern North America. Maize finally reached the northeastern limit of its growing range in southern New England around A.D. 1000.

Late Neolithic developments. Not only are the domestication of plants and animals and the establishment of farming settlements the hallmarks of the Neolithic, but also these provided the platform for subsequent cultural developments. In Europe and elsewhere in the Old World, later Neolithic societies developed elaborate burial practices. These range from the megalithic tombs of the Atlantic seaboard to the rich burials at Khok Phanom Di on the Gulf of Siam. In Europe and the Near East, cattle began to be used to pull wagons and carts, while on the Eurasian steppes north of the Black Sea, horses were domesticated. In the New World, Formative societies exhibited increasing social complexity. At sites such as San José Mogote in Oaxaca, public buildings began to be constructed among the residential compounds. The Hopewell culture in the Ohio valley developed large-scale procurement networks for exotic resources such as obsidian from the Rocky Mountains and native copper from the upper Great Lakes. The Neolithic in the Old World and its New World parallels were periods of dramatic change in human society, laying the foundation for the socially stratified societies that followed. Peter Bogucki

Bibliography. P. Bogucki, *Forest Farmers and Stockherders: Early Agriculture and its Consequences in North Central Europe*, 1988; J. Clutton-Brock (ed.), *The Walking Larder: Patterns of Domestication, Pastoralism, and Predation*, 1989; J. Harlan, *The Living Fields*, 1995; D. R. Piperno and D. M. Pearsall, *The Origins of Agriculture in the Lowland Neotropics*, 1998; T. D. Price (ed.), *Europe's First Farmers*, 2000; T. D. Price and A. B. Gebauer (eds.), *Last Hunters–First Farmers: New Perspectives on the Prehistoric Transition to Agriculture*, 1995; C. A. Reed (ed.), *Origins of Agriculture*, 1977; D. Rindos, *The Origins of Agriculture: An Evolutionary Perspective*, 1984; B. D. Smith, *The Emergence of Agriculture*, 1995; W. H. Wills, *Early Prehistoric Agriculture in the American Southwest*, 1988; D. Zohary and M. Hopf, *Domestication of Plants in the Old World: The Origin and Spread of Cultivated Plants in West Asia, Europe, and the Nile Valley*, 1993.

Neomeniomorpha

A subclass of creeping, vermiform mollusks in the class Aplacophora. They are covered by a spicular integument and recognized by the presence of a ventral groove within which lies a narrow foot and by the absence of an oral shield. Neomenioids range in size from less than 0.08 to 12 in. (2 mm to 300 mm) and are found from subtidal areas to the abyss, at depths over 16,000 ft (5000 m). There are 23 families with 70 genera and 193 species worldwide.

The cuticular integument is thin or thick and the spicular coat smooth, rough, or spiny. The calcareous spicules are solid or hollow and of many shapes: thin flat leaves, upright rods, needles, paddles, hooks, or scoops. Many species have crossed needlelike spicules forming a reticulated network within the cuticle. Epidermal cells secrete the spicules and cuticle. Epidermal papillae, stalked in forms with a thick cuticle, are excretory and empty through the cuticle. Muscles include circular, cross-diagonal, and longitudinal layers beneath the epidermis, a pair of ventral longitudinal bands, and two sets of paired, serial lateroventral bands.

Neomenioids creep by means of their ciliated foot along a track of sticky mucus produced from a ciliated, eversible pedal pit at the anterior end of the pedal groove. Anterior to the pedal pit, the head end is held above the substratum and freely moved.

Food is entirely cnidarians—hydrozoans, zooantharians, gorgonians, or stony corals. Above the mouth lies a sensory atrium. Circular muscles often surround the pharynx, which acts as a buccal pump. The radula is formed of rows of two mirror-image teeth which in some cases are fused medially; it may be large to very small or absent. Digestion starts in the pharynx, well supplied by secretory glands; there is nearly always a pair of discrete ventral salivary glands. A midgut with a dorsal ciliated typhlasole serves both as stomach and digestive gland; it passes posteriorly into a ciliated intestine, which bends ventrally and opens into the mantle cavity.

A ventricle and single or paired auricles lie in the pericardium; a sole vessel, the dorsal aorta, may be present, but otherwise blood moves freely through sinuses of the hemocoel. Paired ventral and lateral nerve cords arise from a cerebral ganglion and run posteriorly, connected by cross-commissures; the lateral cords join above the rectum as a ganglion, which gives off a nerve to a chemoreceptory dorsal sensory organ.

All neomenioids are hermaphroditic and have paired gonads that empty into the pericardium or, unusually, into cloacal pockets. A pair of U-shaped gametoducts with one to many seminal receptacles passes from pericardium to the cloaca. The lower gametoducts are lined with secretory cells and function as shell glands; they open on paired gametopores or unite before opening on a single gametopore which is sometimes a muscular penial sheath. Paired, protrusible copulatory spicules are often present. A barrel-shaped, nonfeeding larva called a pericalymma either is brooded or swims by means of a ciliated cellular test within which the animal develops; metamorphosis through loss or resorption of the test occurs within 10 days. *See* APLACOPHORA; MOLLUSCA. Amelie H. Scheltema

Bibliography. A. C. Giese and J. S. Pearse (eds.), *Reproduction of Marine Invertebrates*, vol. 5: *Molluscs*: *Pelecypods and Lower Classes*, 1979; H. Heath, Solenogastres, *Mem. Mus. Compar. Zool.* (*Harvard Univ.*), 45:1–179, 185–263, 1911, 1918; L. v. Salvini-Plawen, Antarktische und subantarktische Solenogastres, *Zoologica* (*Stuttgart*), 44:1–315, 1978; E. R. Trueman and M. R. Clarke (eds.), *The Mollusca*, vol. 10: *Evolution*, 1985.

Neon

A gaseous chemical element, Ne, with atomic number 10 and atomic weight 20.183. Neon is a member of the family of noble gases. The only commercial source of neon is the Earth's atmosphere, although traces of neon are found in natural gas, minerals, and meteorites. *See* INERT GASES; PERIODIC TABLE.

1												13	14	15	16	17	18
1 **H**	2																2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Considerable quantities of neon are used in high-energy physics research. Neon fills spark chambers used to detect the passage of nuclear particles. Liquid neon can be utilized as a refrigerant in the temperature range about 25 to 40 K (−416 to −387°F). Neon is also used in some kinds of electron tubes, in Geiger-Müller counters, in spark-plug test lamps, and in warning indicators on high-voltage electric lines. A very small wattage produces visible light in neon-filled glow lamps; such lamps are used as economical night and safety lights. *See* NEON GLOW LAMP.

Neon is colorless, odorless, and tasteless; it is a gas under ordinary conditions. Some of the other properties of neon are given in the **table**. Neon does not form any chemical compounds in the ordinary sense of the word; there is only one atom in each molecule of gaseous neon. Arthur W. Francis

Physical properties of neon

Property	Value
Atomic number	10
Atomic weight (atmospheric neon only)	20.183
Melting point, °C	−248.6
Boiling point at 1 atm pressure, °C	−246.1
Gas density at 0°C and 1 atm pressure, g/liter	0.8999
Liquid density at its boiling point, g/ml	1.207
Solubility in water at 20°C, ml neon (STP)/1000 g water at 1 atm partial pressure neon	10.5

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; M. Ozima and F. A. Podosek, *Noble Gas Geochemistry*, 2001.

Neon glow lamp

A low-wattage lamp often used as an indicator light or as an electronic circuit component. The neon lamp usually consists of a pair of electrodes sealed within a bulb containing neon gas at a low pressure. Some of the smaller bulbs are equipped with wire leads that are connected directly into the electrical supply circuit; others are equipped with conventional bases that vary with the size of the lamps (see **illus.**).

Operation. Electrodes sealed in a neon atmosphere will emit electrons if a sufficient voltage difference is impressed across them. In glow lamps the electrodes are usually treated to emit electrons freely. With a sufficiently high voltage between electrodes, the velocity of electron flow is high enough to ionize the neon nearest the negative electrode (cathode). The neon then emits a reddish-orange glow similar to the color of neon sign tubing. With direct current the glow is restricted to the immediate vicinity of the negative electrode. With alternating current (ac), both electrodes act alternately as cathodes, and the glow appears alternately at both surfaces. At usual frequencies, the alternations occur so rapidly that both electrodes appear to glow constantly.

Some examples of glow lamps

In direct-current (dc) circuits, the voltage across the electrodes may be reduced significantly, once the lamp has started, without causing the lamp to go out. Direct-current starting voltages for typical glow lamps range from 65 to 90 V, while the minimum operating voltage at which the glow will be maintained may be 10–15 V lower. On ac circuits, the maintaining voltage is nearly the same as the starting voltage.

This glow discharge is like the arc in a vapor lamp in that its resistance decreases with increasing current. Therefore, a current-limiting element must be used in the electric circuit to maintain a desired stabilized current in the circuit. Because the current in

glow lamps is usually a few milliamperes or less, it is both practical and economical to use a small resistor as a ballast. The larger glow lamps with screw bases have resistors in the bases; smaller lamps require an external ballast resistor. The resistance value depends on the lamp type.

Applications. The neon glow lamp is inherently a low-wattage source that produces light at relatively low efficiency when compared to filament lamps and other vapor lamps. Its lighting applications are widespread where only a low power source is available. These range from simple illuminated wall switches in homes, lighted switches in small appliances, and indicator lights on the panels of electrical devices to more sophisticated usage in digital readout devices.

In electronic circuits involving relatively low power, neon lamps are used in many ways. They are used in counter and memory elements of computers, in voltage regulators, in relaxation oscillators, and in trigger circuits to operate relays and similar devices. These applications are practical because of the lamp's unique electrical characteristics, its small size, and its light weight.

Characteristics. The useful life of a glow lamp is not terminated by a burnout, as is the life of lighting lamps, but by a gradual rise in starting and maintaining voltage and blackening of the inner walls of the bulb, reducing light output. If the lamp is used as an indicator light, the reduction in brightness will determine its life, which may be 5000–25,000 h, depending on the application.

When the lamp is used as a circuit element and voltage is important, the change in starting voltage, maintaining voltage, or both will determine useful life. Depending on the type of lamp and its operating current, a rise of 5 V in starting voltage may occur after 1000–6000 burning h. Maintaining voltage rises about half as fast as starting voltage.

External factors may also affect the operating characteristics of glow lamps. The sensitized electrodes of glow lamps release electrons in the presence of light. In total darkness, the starting voltage may be 100 V higher than in light. In totally dark enclosures, slight amounts of light or other radiation, or electrostatic fields, may be used to overcome the dark effect. Darkness does not affect maintaining voltage. *See* VAPOR LAMP. Alfred Makulec

Bibliography. Illumination Engineering Society of North America, *IES Lighting Handbook—Reference and Application Volume*, 1993.

Neornithes

The subclass of Aves that contains all of the known birds other than those placed in the Archaeornithes. Comprising more than 30 orders, both fossil and living, its members are characterized by a bony, keeled sternum with fully developed powers of flapping flight (secondarily lost in a number of groups); a short tail with the caudal vertebrae fused into a single platelike pygostyle to which all tail feathers attach; a large fused pelvic girdle with a reversed pubis which is fused to a large synsacrum; and a large brain and eyes contained within a fused braincase. The jaws are specialized into a beak covered with a horny rhamphotheca; the upper jaw is kinetic, being either prokinetic or rhynchokinetic. Prokinesis refers to a bending zone at the base of the upper jaw, and rhynchokinesis to one within the upper jaw. A few fossil groups still possess teeth, but most fossil and all Recent birds have lost teeth. *See* ARCHAEOPTERYX.

The Neornithes contains two superorders, the Odontognathae and the Neognathae. The Odontognathae, alternately known as the Odontornithes, may be an artificial group. Its members, which include the Cretaceous fossil orders Hesperornithiformes and Ichthyornithiformes, are united only by the presence of teeth in all species. The Neognathae contains the remaining modern birds, which have lost the teeth, and includes 26 orders. Six orders of Cretaceous fossil birds remain to be placed in either of these superorders because so little is known of their morphology and relationships. The grouping of the orders of living and fossil birds is still ambiguous and requires further work. A few groupings appear to be based on strong evidence, including the close relationships between the Anseriformes and the Galliformes, the Gruiformes and the Charadriiformes, the Columbiformes and the Psittaciformes, the Strigiformes and the Caprimulgiformes, and possibly the orders of large land bird, the Apodiformes, Coraciiformes, Piciformes, and Passeriformes. *See* AVES; NEOGNATHAE; ODONTOGNATHAE. Walter J. Bock

Neoteny

A phenomenon among some salamanders, in which larvae of large size, while still retaining the gills and other larval features, become sexually mature, mate, and produce fertile eggs. In certain lakes of Mexico, only the neotenous larvae are present and are called axolotls. Neoteny occurs in certain species of the family Ambystomidae, especially in *Ambystoma tigrinum* of some localities, and commonly in the large *Dicamptodon ensatus* of the Pacific Coast. It also occurs in some Texas and Oklahoma species of *Eurycea*. Neotenous larvae of *A. tigrinum* can be made to metamorphose to adult form if treated with thyroid extract. Tracy I. Storer

Nepenthales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Dilleniidae of the class Magnoliopsida (dicotyledons). The order (also known as Sarraceniales) consists of three well-marked small families: the Droseraceae with about 90 species; the Nepenthaceae with about 80 species; and the Sarraceniaceae with only about 16. The plants characteristically grow in waterlogged soils which are deficient in available nitrogen. They are herbs or shrubs with alternate, simple leaves that

Pitcher plant (*Sarracenia purpurea*). (*Henry M. Mayer, National Audubon Society*)

are modified for catching insects, from which they absorb nitrogenous nutrients. They have regular flowers characterized by having the perianth and stamens attached directly to the receptacle (hypogynous); having the petals separate (polypetalous) or being without petals (apetalous); having the carpels united to form a compound pistil (syncarpous); and having few to many stamens. The seeds have a well-developed endosperm. The pitcher plants (*Sarracenia*; see **illus.**), sundew (*Drosera*), and Venus' flytrap (*Dionaea muscipula*) are well-known members of the order. *See* DILLENIIDAE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM. Arthur Cronquist

Neper

A unit of attenuation used in transmission-line theory. On a uniform transmission line having waves traveling in only one direction, the magnitudes of voltage E and of current I decrease with distance x traveled, as given by Eq. (1), where E_0, I_0, and α are constants.

$$\frac{E}{E_0} = \frac{I}{I_0} = \epsilon^{-\alpha x} \tag{1}$$

The attenuation in nepers between the points where E_0 and I_0 are measured and where E and I are measured is given by Eq. (2), in which ln denotes the natural (or napierian) logarithm.

$$\alpha x = \ln \frac{E_0}{E} = \ln \frac{I_0}{I} \tag{2}$$

The word neper originated from a misspelling of the proper name Napier. One neper equals 8.686 dB, the decibel being the practical unit of attenuation. *See* DECIBEL; TRANSMISSION LINES.

Edward W. Kimbark

Nepheline

A mineral of variable composition: in its purest state, $NaAlSiO_4$; often nearly $Na_3K(AlSiO_4)_4$; but generally $(Na,K,\square,Ca,Mg,Fe^{2+},Mn,Ti)_8(Al,Si,Fe^{3+})_{16}O_{32}$, where $\square$ represents vacant crystallographic sites, and Ca, Mg, Fe^{2+}, Mn, Ti, and Fe^{3+} are usually present in only minor or trace amounts. The most important variations in nepheline composition are due to crystalline solution of $KAlSiO_4$ (the mineral kalsilite), and substitution of $\square$ for K. Crystalline solution of $KAlSiO_4$ in $(Na,K)AlSiO_4$ nepheline is limited to about 25 mole % at 25°C (77°F), but increases with increasing temperature (T) to 65–70 mole % at 1000°C (1832°F). Substitution of $\square$ for K can be viewed and treated as crystalline solution of alkali feldspar or (alternatively) "excess silica." The amount of "excess silica" that can be present stably in any nepheline is very small at $T < 500°C$ (932°F), and only a few mole % can be dissolved in K-rich nephelines at any temperature. However, crystalline solution of "excess silica" increases steadily in Na-rich nephelines at $T > 500°C$ (932°F), and at 1000°C (1832°F) as much as 30 mole % can be dissolved in K-free nepheline. *See* SILICATE MINERALS.

Characteristics. The salient physical properties of nepheline are a Mohs scale hardness of 5.5–6.0; a specific gravity between 2.56 and 2.67; a typically dark gray, light gray, or white color, but also colorless (nepheline is colorless in petrographic thin section); and a vitreous or greasy luster. Nepheline occurs as simple hexagonal prisms or, more commonly, as isolated shapeless grains or irregular polycrystalline masses.

The crystal structure of nepheline is hexagonal, space group $P6_3$, with $a \cong 1.00$ nanometer and $c \cong 0.84$ nm. The structure of $NaAlSiO_4$ nepheline is a three-dimensional lattice of SiO_4 and AlO_4 tetrahedra with Na atoms housed in the cavities of the framework. In $Na_3K(AlSiO_4)_4$ nepheline, the lattice is distorted, and Na and K occupy cavities (crystallographic sites) of two different sizes—small and large in a 3:1 ratio—with the Na atoms located on the smaller sites and the K atoms positioned on the larger sites. Owing in part to this coincidence of a 3:1 ratio of Na to K and 3:1 ratio of small to large alkali sites, $Na_3K(AlSiO_4)_4$ nepheline is generally regarded as a distinct compound, and it is frequently referred to as ideal nepheline. *See* CRYSTAL STRUCTURE.

Occurrence. Nepheline is the most abundant feldspathoid mineral; it occurs in a wide variety of SiO_2-deficient (quartz-free) and alkali-rich volcanic, plutonic, and metamorphic rocks. In volcanic rocks, nepheline occurs chiefly as a primary mineral in phonolites, kenytes, and melilite basalts, and it is the characteristic mineral of nephelinites. Nephelines in alkalic volcanic rocks of all types are frequently

either more Na-rich or more K-rich than $Na_3K(AlSiO_4)_4$, and they commonly contain significant amounts of "excess silica." In plutonic rocks, nepheline occurs mainly in nepheline syenites and in rocks of the alkali gabbro clan. In contrast to nephelines in volcanic rocks, nephelines in plutonic rocks (and in nepheline gneisses) usually have compositions near $Na_3K(AlSiO_4)_4$, and they contain very little "excess silica." Nepheline-bearing rocks also occur in the contact metamorphic aureoles of some alkalic intrusions, and in some instances it is evident that these rocks were formed by metasomatism of carbonate-rich country rock.

Commercial uses. Both "pure" (processed) nepheline and nepheline syenite are used as raw materials for the manufacture of glass, various ceramic materials, alumina, pottery, and tile. However, only three countries—Canada, Norway, and Russia—produce significant quantities of commercial-grade nepheline and nepheline syenite. It has been reported that alumina from nepheline-bearing rock is the source of approximately one-sixth of the aluminum produced by Russia. *See* FELDSPATHOID; IGNEOUS ROCKS; METASOMATISM; NEPHELINITE. James G. Blencoe

Bibliography. M. J. Buerger, G. E. Klein, and G. Donnay, Determination of the crystal structure of nepheline, *Amer. Mineralog.*, 39:805–818, 1954; D. D. Carr (ed.), *Industrial Minerals and Rocks*, 6th ed., Society for Mining, Metallurgy, and Exploration, 1994; W. A. Deer, R. A. Howie, and J. Zussman, *An Introduction to the Rock-Forming Minerals*, 2d ed., 1992; W. A. Deer, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, vol. 1B: *Disilicates and Ring Silicates* 2d ed., 1986; *Minerals Yearbook, 1988*, vol. 1: *Metals and Minerals*, U. S. Department of the Interior, Bureau of Mines, 1990.

Nephelinite

A dark-colored, aphanitic (very finely crystalline) rock of volcanic origin, composed essentially of nepheline (a feldspathoid) and pyroxene. *See* KALSILITE.

The texture is usually porphyritic with large crystals (phenocrysts) of augite and nepheline in a very fine-grained matrix. Augite phenocrysts may be diopsidic or titanium-rich and may be rimmed with soda-rich pyroxene (aegirine-augite). Microscopically the matrix is seen to be composed of tiny crystals or grains of nepheline, augite, aegirite, and sodalite with occasional soda-rich amphibole, biotite, and brown glass. If leucite becomes the dominant feldspathoid, the rock becomes a leucitite. If calcic plagioclase exceeds 10%, the rock passes into tephrite and basanite. If olivine is present, the rock is an olivine nephelinite (nepheline basalt). Accessories usually include magnetite, ilmenite, apatite, sphene, and perovskite.

Nephelinite and related rocks are very rare. They occur as lava flows and small, shallow intrusives. A great variety of these feldspathoidal rocks is displayed in Kenya. *See* FELDSPATHOID; IGNEOUS ROCKS. Carleton A. Chapman

Neptune

The outermost of the four giant planets. Neptune is a near twin of Uranus in size, mass, and composition. Its discovery in 1846 within a degree from the theoretically predicted position was one of the great achievements of celestial mechanics. Difficulties in accounting for the observed motion of Uranus by means of perturbations by the other known planets led early in the nineteenth century to the suspicion that a new planet, beyond the orbit of Uranus, might be causing the deviation from the predicted path. The difficult problem of deriving the mass and orbital elements of the unknown planet was solved independently in 1845–1846, first by J. C. Adams in Cambridge, England, and then by U. J. Leverrier in Paris. Adams's result did not receive immediate attention, and so it was Leverrier's solution that led to the discovery of Neptune by J. G. Galle, in Berlin, who found the planet on September 23, 1846, only 55′ from its calculated position. *See* CELESTIAL MECHANICS; PERTURBATION (ASTRONOMY).

The planet and its orbit. The actual mass and orbit of Neptune differ considerably from the values predicted by Adams and by Leverrier, since both assumed that the mean distance of the planet to the Sun would be that predicted by the Titius-Bode relation, namely, 38.8 astronomical units, whereas it is only 30.1 au or 2.8×10^9 mi (4.5×10^9 km). The eccentricity of the orbit is only 0.009, the second smallest (after that of Venus) among the planets; the inclination is 1.8°; the period of revolution is 163.7 years; and the mean orbital velocity of Neptune is 3.4 mi/s (5.5 km/s). *See* PLANET.

Through a small telescope, Neptune appears as a tiny greenish disk, with a mean apparent diameter of about 2.4″ (the Moon has an apparent diameter of 31′). This corresponds to a linear equatorial diameter of 30,775 mi (49,528 km)—very similar to that of Uranus. The mass of Neptune is 17.15 times the mass of Earth, corresponding to a mean density of 1.64, somewhat above that of its sister planet. This suggests that the proportion of heavy elements is somewhat greater in Neptune than in Uranus. *See* URANUS.

The apparent visual magnitude of Neptune at mean opposition, that is, when closest to Earth, is +7.8, too faint to be seen by the unaided eye. The corresponding albedo is 0.4, a relatively high value characteristic of a planet with a dense atmosphere. Photographs taken through powerful telescopes under excellent visual conditions with special filters and cameras reveal the presence of discrete cloud systems in Neptune's atmosphere, again in contrast to Uranus, whose atmosphere contains fewer cloud systems. *See* ALBEDO.

Atmosphere. Most of what is known about Neptune is the result of the flyby of the planet by

the *Voyager 2* spacecraft in August 1989. The cloud features that were dimly glimpsed from Earth were recorded in great detail (**Fig. 1**). They included a large dark oval (about the size of Earth), reminiscent of Jupiter's Great Red Spot, as well as the white clouds of condensed methane whose brilliant contrast with the blue-green atmosphere made them visible from Earth. Unlike the Great Red Spot, Neptune's dark oval proved to have a short lifetime, as subsequent observations from Earth showed that it had disappeared. By following the clouds over several weeks, scientists were able to deduce the presence of currents at different latitudes, with the high-latitude winds faster than those near the equator. The reference frame is established by the rotation period of the deep interior, 16h7m, as determined from radio emissions whose intensity variations reflect the rotation of the planet's magnetic field. Storm systems on Neptune can cross latitude lines, moving toward the equator as the dark oval did before disappearing. On Jupiter and Saturn, such latitudinal motion does not occur.

This circulation pattern resembles that of Uranus, despite the different inclinations of the rotational axes of the two planets (that of Neptune is 29.6°, while that of Uranus is 97.9°, and that of Earth is 23.5°), and the fact that Neptune has an internal energy source that releases some 2.7 times the amount of heat absorbed from the Sun, while Uranus has no excess internal heat. The explanation may lie in similar atmospheric opacities. However, the difference in internal heat may be responsible for the greater cloud activity on Neptune than on Uranus.

The atmosphere of Neptune, like those of the other giant planets, is composed predominantly of hydrogen and helium. The relative abundance of methane is enhanced slightly more than on Uranus, between 25 and 40 times the value corresponding to solar abundances of the elements. This gas contributes to Neptune's greenish color by absorbing orange and red light. Ammonia is expected to be present at lower, warmer levels of the atmosphere, but radio observations that probe to these levels find the ammonia to be depleted in comparison with methane. Radio observations from Earth discovered carbon monoxide and hydrogen cyanide in Neptune's upper atmosphere, while infrared spectra have revealed ethane and ethylene.

Magnetic field. The orientation of Neptune's magnetic field is surprisingly similar to that of Uranus. It can be represented by a bar magnet inclined at an angle of 46.8° with respect to the axis of rotation and offset by 0.55 planetary radius. (For comparison, the Earth's field is inclined by only 11° and offset by 0.07 radius.) Because of the offset, the field strength varies from a minimum of less than 0.1 gauss (10^{-5} tesla) in the northern hemisphere of Neptune to a maximum of greater than 1.0 gauss (10^{-4} T) in the southern hemisphere.

This field has trapped a plasma of ionized and neutral gases in the planet's magnetosphere. The maximum plasma density is only 1.4 particles per cubic centimeter (0.09 particle per cubic inch), the lowest of any of the giant planets. There are ions of molecular hydrogen and helium as well as atomic hydrogen and heavier ions that have probably escaped from Triton. These should include nitrogen from the satellite's atmosphere.

Fig. 1. ***Voyager 2*** **picture of Neptune, showing the huge dark oval at 22° south latitude, with bright, white methane clouds above it. Farther to the south is a triangular cloud system called the scooter because of its high relative velocity, and still farther south is another dark oval with a bright core. (*NASA*)**

A faint aurora was observed on Neptune, caused by charged particles from the radiation belts (the trapped plasma) spiraling around planetary magnetic field lines and bombarding the planet's atmosphere. Because of the unusual orientation of the field, the maximum auroral activity occurs at midlatitudes rather than near the poles as it does on Earth. *See* AURORA; MAGNETOSPHERE; VAN ALLEN RADIATION.

Satellites. Before the *Voyager* encounter, only two satellites of Neptune were known, both in highly irregular orbits. Triton was discovered visually by W. Lassell in 1846. It is moving in a retrograde direction around Neptune with a period of 5.9 days in a nearly circular orbit. Nereid was found in 1949 as a result of a photographic search by G. P. Kuiper. It has the most eccentric orbit of any known satellite. The eccentricity e measures the flattening of the ellipse; a circle has $e = 0$, while a parabola, representing an unclosed orbit, has $e = 1$. While the orbits of most satellites and planets have eccentricities close to zero, Nereid has $e = 0.75$. In sharp contrast to these two bodies, the

TABLE 1. Neptune satellites*

Satellite name	Distance to Neptune, 10^3 mi (10^3 km)	Sidereal period, h	Diameter, mi (km)	Albedo
Naiad	29.9 (48.2)	7.1	37 (60)	0.06?
Thalassa	31.1 (50.1)	7.5	50 (80)	0.06?
Despina	32.6 (52.5)	8.0	93 (150)	0.06?
Galatea	38.5 (62.0)	10.3	99 (160)	0.06
Larissa	45.7 (73.5)	13.3	118 (190)	0.06
Proteus	73.1 (117.6)	26.9	261 (420)	0.06
Triton	220.5 (354.8)	141.0	1681 (2705)	0.076
Nereid	3426 (5513)	8643	211 (340)	0.16

*Only satellites with well-determined orbits as of March 1, 2004, are listed.

six satellites discovered by the *Voyager* cameras all have very regular orbits: in the plane of the planet's equator and nearly circular. They are all close to the planet (**Table 1**). In 2002 and 2003, five irregular satellites were discovered from Earth at large distances from the planet; their orbits must be better defined before they receive names. The search for additional satellites continues. *See* ELLIPSE; ORBITAL MOTION; SATELLITE.

Triton has a tenuous atmosphere containing nitrogen, methane, and carbon monoxide. These three gases plus carbon dioxide and water are also present as frozen ices on the satellite's surface. *Voyager* revealed that this remarkable object has a diameter of only 1681 mi (2705 km), making it considerably smaller than the Earth's Moon (2086 mi or 3476 km). The surface temperature of Triton is $-391 \pm 7°F$ (38 ± 4 K). The size of this satellite, as well as the temperature and composition of its surface, makes Triton very similar to Pluto. *See* PLUTO.

Triton's surface has an appearance quite different from that of any other satellite. The scarcity of impact craters means that the surface is geologically young. There are long intersecting valleys and ridges crossing various types of terrain. The illuminated part of the satellite at the time of the *Voyager* encounter was the southern hemisphere. Most of it appeared to be covered with layers of highly reflective ice, on which darker splotches appeared, many clearly organized by near-surface winds. Subsequent Earth-based observations of Triton's near-infrared spectrum revealed absorptions of solid carbon monoxide, carbon dioxide, and water ice. There were earlier (pre-*Voyager*) indications that ices of methane and nitrogen were present. Nitrogen ice is the predominant surface constituent, with the other ices contributing less than 1–2%. Near the equator, there is a bluish deposit that is probably fresh nitrogen ice, while at higher latitudes the surface resembles the skin of a cantaloupe.

Triton's atmosphere has a surface pressure of only 1.6 ± 0.3 pascals or 16 ± 3 microbars (Earth's atmospheric pressure is approximately 10^5 Pa or 1 bar), dominated by molecular nitrogen. Only 1 part in 10,000 is methane, while carbon monoxide must be less than 4%. Yet reactions in this tenuous atmosphere (and in the surface ices) apparently form the organic compounds that produce the deposits of dark material and give the surface its characteristic pinkish color.

Perhaps the most unusual aspect of Triton is the presence of eruptive plumes at several places on its surface. These plumes consist of columns of dark material ejected some 5 mi (8 km) upward into the atmosphere, where they form small clouds that are then blown into narrow wind trails extending hundreds of miles from their sources. These plumes may be powered by a solid greenhouse effect within Triton's icy surface, in which solar radiation that penetrates the ice encounters dark material that it warms. This causes the surrounding nitrogen ice to sublime until sufficient gas pressure is produced to cause gas to break through to the surface, forming the plume. The dark material carried upward by these jets contributes to the hazes seen in Triton's atmosphere.

Rings. Earth-based observations of Neptune when it passed in front of distant stars indicated the presence of material in orbit about the planet. In some cases, the light from the star would briefly disappear before Neptune reached it, but this behavior would not be repeated on the other side of the planet as would be expected for a planetary ring. Sometimes, no dimming of the star's light on either side of the planet was observed. These observations led to the idea that Neptune might be surrounded by incomplete rings or arcs. *See* OCCULTATION.

The *Voyager* cameras showed that in fact there are three well-defined, complete rings around Neptune, accompanied by a sheet of material that itself constitutes a broad ring. In order of increasing distance from the planet, the three discrete rings have been designated Galle, Leverrier, and Adams, while the sheet of material is known as Lassell, and a concentration of particles within it is called Arago (**Table 2**). The outermost of these, the Adams ring, contains three concentrated clumps of material known as arcs (**Fig. 2**), rather like sausages strung on a wire, and these plus a chance occultation by one of the inner satellites were apparently responsible for the confusing ground-based observations. A fourth, much smaller arc is barely visible in the best *Voyager* image. The outer two discrete rings are very narrow, reminiscent of the rings of Uranus, with an average width of 9 mi (15 km).

The confinement of these narrow rings is commonly assumed to require the presence of small

TABLE 2. Neptune ring data

Feature	Distance from Neptune's center, 10^3 mi (10^3 km)	Comments
Pressure level of 10^5 Pa (1 bar) in Neptune's atmosphere	15.4 (24.8)	Equatorial radius of Neptune
Galle	26.1 (42.0)	Outer edge of this ring
Leverrier	32.9 (53.0)	High dust content
Lassell	33.37 (53-59)	Broad ring
Arago	35.5 (57.2)	Concentration within Lassell
Adams	39.1 (63.0)	9 mi (15 km) wide; contains three arcs

shepherding satellites. Galatea and Despina orbit, respectively, just inside the outer two narrow rings, but the corresponding outer shepherds have not been found. Similarly, the persistence of the three arcs within the outer ring remains an enigma. In the absence of some gravitational control, the material in these arcs would be expected to spread out around the planet, simply contributing to the ring itself.

Origin and evolution. Like the other giant planets, Neptune is thought to have formed in a two-stage process. First a large core of solid material accumulated, growing as a result of collisions with smaller so-called planetesimals. This core was dominated by ices, although it contained rocky material as well. As the core grew to the size of several Earth masses, it developed an atmosphere, the result of impact vaporization of the icy bodies that were crashing into it. The atmosphere consisted of gases such as molecular nitrogen, methane, and carbon monoxide. As the core grew, it began to attract gas from the surrounding solar nebula, which added hydrogen and helium to the atmosphere. The process stopped when the supply of materials in the vicinity of the growing planet was exhausted.

In the case of Jupiter and Saturn, there was enough nebular gas available to produce huge planets with deep atmospheres dominated by hydrogen and helium. Even in these atmospheres, however, the heavy elements were enriched. At the great distances where Uranus and Neptune formed, the solar nebula evidently had less material to offer the growing planets, with the result that these two objects do not have such deep atmospheres. In other words, they exhibit a much higher proportion of core mass to total mass than Jupiter and Saturn do. This difference in the proportion of core to atmosphere is reflected in the higher proportion of carbon to hydrogen in the atmospheres of Uranus and Neptune, exemplified by the large relative abundance of methane discussed above. *See* JUPITER; SATURN.

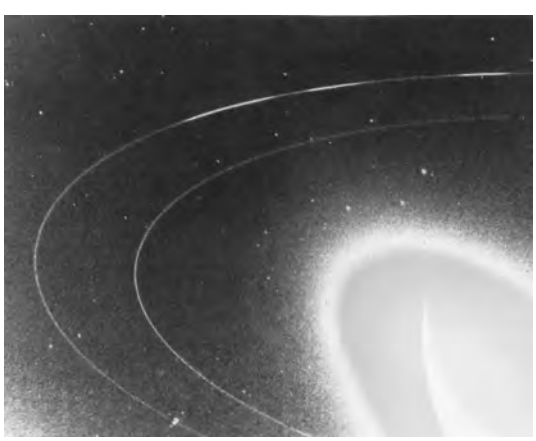

Fig. 2. The two outer discrete rings of Neptune (Adams and Leverrier), with the three arcs of concentrated material in the outer ring. (*NASA*)

The six inner regular satellites of Neptune resemble the ten that *Voyager* found around Uranus, in size, albedo, and orbital characteristics. In both cases, it is thought that these bodies formed out of the planetary subnebula, in much the same way that the planets themselves formed from material in the solar nebula in orbit about the Sun. Triton and Nereid must have had a different beginning, however. Both are probably captured objects that accreted independently. In the case of Triton, the process that led to capture evidently dissipated enough energy to melt the satellite completely, obliterating the record of impact craters which should have been left on the surface. Evidently Triton, like Pluto, is a large icy planetesimal, an object that grew from collisions of smaller icy bodies that may have been roughly identical to the comets observed today. *See* COMET; SOLAR SYSTEM.

Tobias C. Owen

Bibliography. J. K. Beatty, C. C. Petersen, and A. Chaikin (eds.), *The New Solar System*, 4th ed., Cambridge University Press, 1999; D. P. Cruikshank (ed.), *Neptune and Triton*, University of Arizona Press, 1995; M. Grosser, *The Discovery of Neptune*, Harvard University Press, 1962, reprint, Dover Publications, 1979; D. Morrison and T. Owen, *The Planetary System*, 3d ed., 2003; D. A. Rothery, *Satellites of the Outer Planets: Worlds in Their Own Right*, 2d ed., Oxford University Press, 1999; F. W. Taylor, *The Cambridge Photographic Guide to the Planets*, Cambridge University Press, 2001.

Neptunium

A chemical element, symbol Np, atomic number 93. Neptunium is a member of the actinide or 5*f* series of elements. It was synthesized as the first transuranium element in 1940 by bombardment of uranium with neutrons to produce neptunium-239. The lighter isotope ^{237}Np, a long-lived alpha emitter with half-life

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

2.14×10^6 years, is particularly important chemically. *See* PERIODIC TABLE.

Neptunium metal is ductile, low-melting (637°C or 1179°F), and in its alpha form is of high density, 20.45 g/cm^3 (11.82 oz/in.^{3}). The chemistry of neptunium may be said to be intermediate between that of uranium and plutonium. Neptunium metal is reactive and forms many binary compounds, for example, with hydrogen, carbon, nitrogen, phosphorus, oxygen, sulfur, and the halogens. *See* ACTINIDE ELEMENTS; NUCLEAR CHEMISTRY; TRANSURANIUM ELEMENTS.

Robert A. Penneman

Bibliography. R. Guillaumont et al. (eds.), *Update on the Chemical Thermodynamics of Uranium, Neptunium, Plutonium, Americium and Technetium*, 2003; S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2000; J. J. Katz, G. T. Seaborg, and L. R. Morss (eds.), *The Chemistry of the Actinide Elements*, 2 vols., 2d ed., 1986.

Nerve

A group of nerve fibers coursing together as a bundle in the peripheral nervous system. The individual fibers are covered by Schwann cells, many of which contain large amounts of myelin, which makes the nerve appear shiny white. The nerve fibers with their Schwann cell sheaths are held together by connective tissue. In most nerves, some of the fibers are sensory (carrying information to the central nervous system) and some are motor (carrying information from the central nervous system to peripheral glands and muscles). When both sensory and motor fibers are in a nerve, it is called a mixed nerve.

In the central nervous system (brain and spinal cord) a group of nerve fibers running together is called a tract and has different structural and functional properties than a nerve. Glial cells, not Schwann cells, form the sheaths of tract fibers, and there is no connective tissue holding the bundle together. Whereas most nerves are mixed, there is functional segregation in the central nervous system so that most tracts have only one functional type of fiber. *See* MOTOR SYSTEMS; NERVOUS SYSTEM (VERTEBRATE).

Douglas B. Webster

Nervous system (invertebrate)

All multicellular organisms have a nervous system, which may be defined as assemblages of cells specialized by their shape and function to act as the major coordinating organ of the body. Nervous tissue underlies the ability to sense the environment, to move and react to stimuli, and to generate and control all behavior of the organism. Compared to vertebrate nervous systems, invertebrate systems are somewhat simpler and can be more easily analyzed. Invertebrate nerve cells tend to be much larger and fewer in number than those of vertebrates. They are also easily accessible and less complexly organized; and they are hardy and amenable to revealing experimental manipulations, such as changes in the composition and temperature of the fluids surrounding them. However, the rules governing the structure, chemistry, organization, and function of nervous tissue have been strongly conserved phylogenetically. Therefore, although humans and the higher vertebrates have unique behavioral and intellectual capabilities, the underlying physical-chemical principles of nerve cell activity and the strategies for organizing higher nervous systems are already present in the lower forms. Thus neuroscientists have taken advantage of the simpler nervous systems of invertebrates to acquire further understanding of those processes by which all brains function. *See* NERVOUS SYSTEM (VERTEBRATE).

Fundamental Definitions

The average nerve cell, representative of any multicellular organism, may be distinguished by larger-than-average amounts of the cell organelles and metabolic machinery used in protein and lipid synthesis and in the manufacture of special chemical substances which are released as chemical messengers. These compounds are used at specialized connections with other nerve or effector cells called synapses. Nerve cells also have elaborate extensions of the cell body called axons and dendrites; and they have highly specialized cell membranes capable of generating electrical potentials. Nerve cells are almost always intimately associated with other cells, called glia, which aid in mechanical support of nervous tissue and serve to complement nerve metabolism and to regulate the fluid environment of nervous tissue. It appears that nervous tissue in all metazoans is derived, in embryogeny, from the ectoderm. *See* NEURON.

Electrical potentials. The two most distinctive characteristics of nerve cells are the ability to create electrical potentials and to make synaptic connections. Nerve cells and other excitable cells, such as muscle fibers, can establish electrical potentials across their cell membrane (**Fig. 1**). This capability is due to the membrane's selective permeability to certain dissolved ions, such as potassium (K^+), sodium (Na^+), and chloride (Cl^-), which have different concentrations inside and outside the cell.

Most nerve cells have steady resting membrane potentials (RMP) of from 50 to 90 mV, with the inside

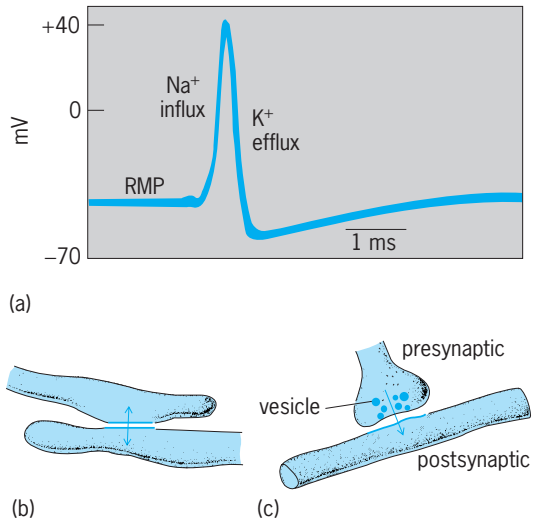

Fig. 1. Electrical potentials and synaptic arrangements. (*a*) An action potential recorded from the giant axon of the squid (*after A. L. Hodgkin and A. F. Huxley, Resting and action potentials in single nerve fibres, J. Physiol., 104:176–195, 1945*). (*b*) Electrical synapse (two-way). (*c*) Chemical synapse (one-way).

of the cell negatively charged with respect to the outside. Furthermore, the permeability of excitable cells to different ions can change rapidly, and this leads to the production of fast voltage changes called action potentials (nerve impulse), which can regenerate themselves, and propagate along the cell processes. During an action potential, the cell's internal potential changes very quickly from its negative resting value, through zero potential to an internal positive value, and then rapidly back to the resting level. Action potentials last only a few milliseconds; they are quite stereotyped in their size and shape, and they represent the way in which nerve cells carry information along their processes to various destinations. In most nerve cells the depolarizing phase of the action potential is caused by the inward flow of sodium ions across the membrane, whereas the repolarizing phase results from the outward flow of potassium ions. Energy-requiring metabolic "pumps" in the membrane help to maintain or restore concentration gradients of these ions. *See* BIOPOTENTIALS AND IONIC CURRENTS.

Synaptic connections. Nerve cells communicate with one another and with their target effector cells (muscle, glands) in a highly organized way at contact points called synapses. No other cell or tissue type is known to have the elaborate scheme of direct interconnections evolved by nerve cells. In a comparatively few cases, nerve cells make electrical synapses with one another (Fig. 1*b*). In such cases the membranes of the two cells come very close together, and special channels are formed in such a way that electric currents (including action potentials) can pass directly from one cell to the other, and in either direction. Electrical synapses are thought, in general, to be the more primitive type of communication; they are usually seen in lower animals, in less elaborately developed tissues (epithelia and embryonic tissues), or in certain specialized contacts, as in some muscle.

The more common synaptic connection is called the chemical synapse (Fig. 1*c*). Here one cell, the presynaptic element, has a specialized ending (typically the axon) in close apposition to a postsynaptic cell process (usually a dendrite or cell body). The presynaptic ending is filled with small rounded vesicles (packets) containing a chemical transmitter substance. When an action potential invades the presynaptic ending, packets of transmitter substance are released from the presynaptic ending. The transmitter diffuses across the narrow cleft between the two cell processes, and interacts with specialized receptors on the postsynaptic membrane. This causes a small voltage change, called a synaptic potential, which then influences the probability of the postsynaptic neuron to produce an action potential. Synaptic potentials may be excitatory or inhibitory in their action, depending on whether they make the postsynaptic element more or less likely to produce an action potential. One nerve cell and its processes may receive many synaptic inputs from other cells, and one cell may connect via branches of its axon to many others. Chemical synapses confer direction of information flow at a synapse, that is, from presynaptic to postsynaptic element; and the size and efficacy of chemical synaptic potentials can be altered by a variety of physiological mechanisms. *See* SYNAPTIC TRANSMISSION.

Nerve cell categories. Nerve cells may be categorized into three basic functional types: sensory neurons, motor or effector neurons, and interneurons. Sensory neurons are specialized to transduce specific types of environmental signals, such as touch, temperature, or visual cues, into electrical information (action potentials), and to transmit this into the central nervous system via their axons. Motor or effector neurons are central nervous system neurons that send an axon to the periphery to innervate muscles or glands. Motor neurons are the final common pathway responsible for all movements and activities of animals, including locomotion, reflex movements, and gland secretions. Interneurons constitute the greatest number of nerve cells in the central nervous system and do not send processes into the periphery, but serve to interconnect sensory and motor cells and other interneuronal nerve cells. Interneurons represent the neural basis for the organization of complex activities and functions. In ascending the phylogenetic scale, the number of interneurons increases greatly. *See* MOTOR SYSTEMS; SENSATION.

Invertebrate and vertebrate nerve cells share all the properties described above; they differ more in quantity, or degree, than in qualitative features. Aside from differences in size and numbers, the most striking difference is that many invertebrate neurons have a unipolar shape, whereas most vertebrate neurons are multipolar (**Fig. 2**). Unipolar refers to a nerve cell body with one process extending away from it. Multipolar neurons, in contrast, have several processes coming off the cell body. One subset of

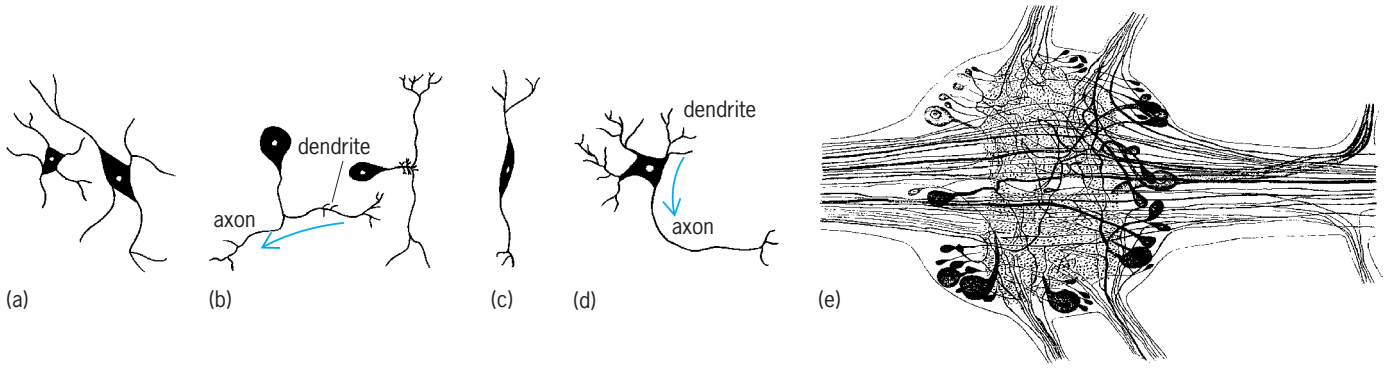

Fig. 2. Neuron types. (*a*) Multipolar, isopolar neurons (primitive invertebrate). (*b*) Unipolar, heteropolar neurons (common invertebrate); arrow indicates direction of information flow. (*c*) Bipolar neuron (common invertebrate). (*d*) Multipolar, heteropolar neuron (common vertebrate, advanced invertebrate); arrow indicates direction of information flow. (*e*) Ganglion of unipolar neurons in crayfish. The unipolar cell bodies, surrounded by connective tissue and glia, form a rind around a central neuropile composed of the cell processes (axons and dendrites). Nerve tracts enter or leave the neuropile. Synaptic interactions and integration are carried out in the neuropile. (*After G. Retzius, Zur Kenntniss des Nervensystems der Crustacean, Biol. Untersuch., NF 1:–50, 1980*)

multipolar cells are bipolar, comprising cells with two stem processes.

Regardless of their configuration, nerve cell processes may function in one of two ways. In advanced neurons one set of processes, called dendrites, are the input or receiving processes of the cell, while another process, the axon, functions as the output process. Such nerve cells are said to have preferred or dynamic polarity; that is, they channel information flow in one functional direction. This situation is defined as heteropolar. Most invertebrate neurons are unipolar, with a functional, specified direction of information flow (heteropolar). Synaptic inputs occur at certain locations (dendritic), generating action potentials that propagate along the axon. Unipolar neurons almost never receive synaptic inputs on the cell body. Many sensory cells may be of the bipolar heteropolar type. Most vertebrate and many advanced invertebrates have multipolar, heteropolar nerve cells. In contrast, more primitive nerve cells have undifferentiated processes with no preferred direction of information flow. They might receive or transmit signals in any direction, and action potentials might be generated at various points on the nerve cell and propagate in all directions. Such cells have a uniform lack of preference for direction of information flow, and are said to be isopolar. In multipolar, isopolar neurons, synapses (usually electrical) occur at many locations on the processes; information flow may begin at any location and spread uniformly.

An additional general contrast between invertebrate and vertebrate nervous systems is that invertebrates tend to have more neurons displaced to the periphery (outside the central nervous system) and to perform more integrative and processing functions in the periphery. Vertebrates perform almost all their integration within the central nervous system, using interneurons. For example, invertebrate muscles are often controlled through an innervation containing both excitatory and inhibitory motor neurons which cause direct contraction or relaxation, respectively, of the muscle. Vertebrates only have excitatory motor neurons. To prevent muscle contraction (or to allow relaxation), they suppress the excitatory motoneuron with inhibitory interneurons in the central nervous system.

Invertebrate nervous systems also seem to have a greater potential for regrowth, regeneration, or repair after damage than do vertebrate nerve cells. Many invertebrates continue to add new nerve cells to their ganglia with age; vertebrates, in general, do not. Only vertebrate neurons have myelin sheaths, a specialized wrapping of glial membrane around axons, increasing their conduction speed. Invertebrates tend to enhance conduction velocity by using giant axons, particularly for certain escape responses.

Description and Phylogeny

Most phylogenetic schemes suggest that the evolution of invertebrates has been accompanied by an increase in body size and complexity (having more varied body parts, organs, and functions to coordinate), and the development of more elaborate behavioral patterns. Not surprisingly, the size and complexity of the nervous system parallel these developments. The following examples of nervous systems represent only major invertebrate phyla; phylogenetically important groups, especially well-studied animals, and developmental trends are highlighted.

Protozoa. Membranes with maintained, but variable, electrical potentials across them are already evident in these single-celled invertebrate organisms (**Fig. 3***a*). The ciliated protozoans have been most studied, particularly *Paramecium*. These organisms have an internally recorded resting potential of about −20 to −40 mV; and this potential, like the resting potential of nerve cells, is mainly due to the membrane's selective permeability to K^+ ions. In response to mechanical distortion, the membrane may generate an action potential which causes a change in the direction of ciliary beating. The inward current of the action potential is carried by calcium ions, is graded, and is similar to voltage-dependent calcium permeability changes seen in some crustacean muscles and

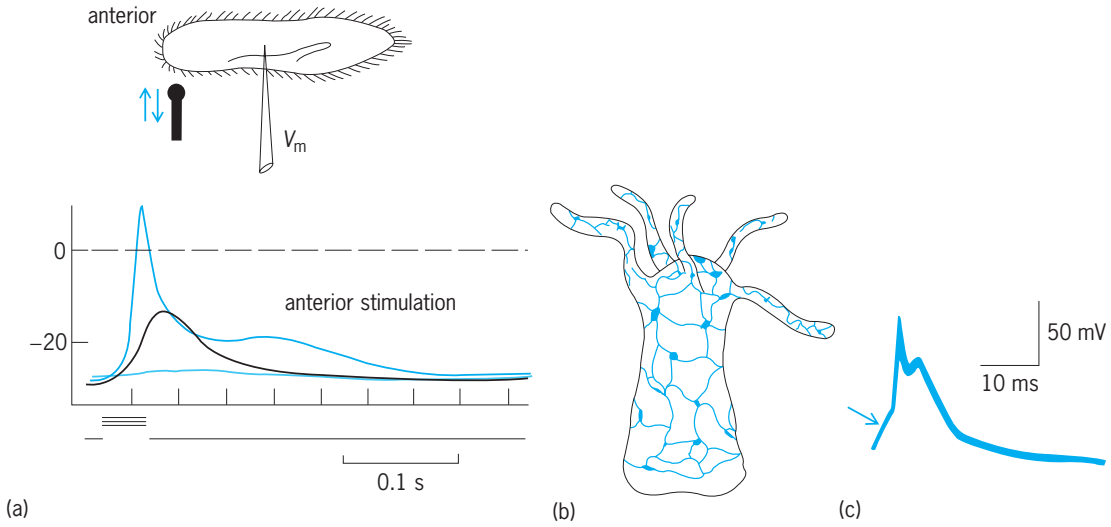

Fig. 3. Electrical responses and nervous systems in the lowest invertebrates. (*a*) Recording from *Paramecium* illustrates a graded action potential which uses Ca^{2+} ions in the upswing phase and K^+ ions during repolarization; three responses of varying size were produced by mechanical taps of increasing intensity to the front of the animal; V_m = membrane potential recorded with a glass microelectrode (*after Y. Naitoah and R. Eckert, Ionic mechanisms controlling behavioral responses of Paramecium to mechanical stimulation, Science, 164:963–965, 1969*). (*b*) Hydrozoan showing nerve net underlying the epidermis. (*c*) Intracellular microelectrode recording from a large neuron of the subumbrellar nerve net of the jellyfish *Carybdea*; this action potential was generated by a synaptic potential (arrow); RMP approximately −70 mV (*after R. A. Satterlie, Central control of swimming in the cubomedusan jellyfish Carybdea rastonii, J. Comp. Physiol., 133:357–367, 1979*).

in the presynaptic axon terminals of most animals. Moreover, the entry of Ca^{2+} ions, itself, causes an increased outward flux of K^+ ions, which brings the membrane potential back toward the resting value. Such Ca^{2+}-dependent potassium currents have also been described in nerve cells from many organisms, notably molluscan and mammalian central nervous system neurons.

The calcium-potassium sequence underlying the active response of *Paramecium* is rather analogous to the sodium-potassium scheme used by nerve cells to make an action potential, although in the latter the control mechanisms are more elaborate and differ in important details. There are numerous examples of invertebrate neurons in all phyla which are intermediate in their use of Ca^{2+} instead of Na^+ to carry action potential currents. A variety of muscle cells in invertebrates and vertebrates also have this property. *See* PROTOZOA.

Porifera. Whether or not the most primitive phylum of multicellular animals, the sponges, have clearly defined nerve cells is uncertain. However, some nervelike functions may be present to account for the ability of these organisms to make local body movements, or to alter the size of the canals which channel the water flow through their tissues. *See* PORIFERA.

Cnidaria. The first true nerve cells appear in this phylum (also called the Coelenterata), which includes the hydrozoan polyps, jellyfishes, anemones, and corals (Fig. 3*b* and *c*). The most intensive studies of the nervous system have been done on *Hydra* and *Tubularia*, and on several genera of jellyfish and anemones.

The phylum Cnidaria is characterized by radial symmetry, a body wall composed of only two true cell layers, and a central digestive cavity with a single opening, the mouth, which is surrounded by tentacles. The two epithelial body layers, an outer epidermis and an inner gastrodermis, are separated by a mesoglea of variable thickness composed of a few primitive cells and a gelatinous matrix.

The nervous system of the typical cnidarian is mostly dispersed in a so-called nerve net, a thin layer of nerve cells lying just beneath the epidermis at the border of the mesoglea. The nerve cells are multipolar (often bi- or tripolar) and isopolar (Fig. 2*a*), and their cell bodies are scattered throughout most areas of the body. The undifferentiated nerve processes spread out weblike to interconnect with one another, with sensory neurons, and with contractile epitheliomuscle cells. True muscle cells are not present. Many cnidarians may have additional nerve nets, of varying complexity and density, lying beneath the gastrodermis, around the mouth, along the tentacles, or concentrated in rings around the bell margins of medusa-form animals. Jellyfish and anemones may have two or more epidermal nerve nets, one specialized for slow conduction, the other for faster conduction.

The multipolar, isopolar neurons of the nerve nets have a negative resting potential and can generate and propagate action potentials throughout the net (Fig. 3*c*). Most of the nerve net neurons are interconnected by electrical synapses, and activity generated at any point in the net spreads throughout the entire system. At least a few specialized sensory neurons can be identified in most groups, including some responsive to light, chemical cues, and mechanical stimuli (touch or traumatizing injury). Some jellyfish have highly developed eyespots and statocysts; these are usually evenly distributed

around the margin of the bell. Many sensory neurons may be sensitive to more than one kind of stimulus (polymodal). Activity may be generated in the net by sensory-cell input or by specialized pacemaker neurons which can produce action potentials endogenously and repetitively. Discrete motor neurons are probably not present in cnidarians; rather, input to the contractile epitheliomuscle cells is from certain neurons of the nerve net which contact both effector cells and other net neurons.

Electrical impulses, correlated with various behaviors but conducted from cell to cell within the epithelium (epithelial conduction), have been widely observed in cnidarians. Epithelial conduction operates in parallel with the nerve net, and the systems interact.

Several lines of evidence indicate that cnidarians have a variety of chemical synapses. Distinct ones have been identified between nerve-net neurons (presynaptic) and epitheliomuscle cells and nematocyst cells (postsynaptic elements). Many nerve-net neurons are interconnected by what appear to be two-way chemical synaptic junctions; that is, both neural elements have vesicles and are, therefore, pre- and postsynaptic to one another. These symmetrical synapses can transmit impulses in both directions, unlike the neurons of higher animals that are directionally polarized. The presence of chemical synapses probably underlies some functional features long observed in cnidarian physiology. Repeated stimulation is often necessary to obtain a response from animals of this group; responses are known to facilitate, that is, to grow larger, with repetition. This type of variability, or plasticity, is a common characteristic of chemical synaptic transmission. Thus the ability to augment or alter responsivity in a neural circuit because of use-dependent changes in synaptic function is apparently already present in these earliest nerve circuits. Although the nervous system is primitive, a variety of neuropeptides and neurotransmitters have been identified in cnidarians. *See* CNIDARIA.

Platyhelminthes. The flatworms mark the beginning of bilateral symmetry in the animal kingdom, and they also illustrate the clear onset of organizational principles which will guide the elaboration of neural tissue in all higher groups. The well-studied, free-living turbellarian flatworms are used as a typical example of this phylum, since the other major classes are highly modified as parasites. While many of the organs of the acoelomate turbellarians are reminiscent of the cnidarians (for example, a digestive tract with a single opening), marked progression is seen in the complexity of most organs. The flatworm nervous system retains, of course, those fundamental elements of neural function, action potentials and chemical synaptic potentials, but few studies at the cellular level have been made to provide details.

The most important changes and new features seen in neural organization include the following (**Fig. 4*a***). (1) While some nerve nets or plexuses are present in most species, the bulk of the nervous system has begun to coalesce into distinct clusters or ganglia within the mesenchyme away from the epidermis. This represents the process of centralization. The organization of ganglia is already that basic to higher forms: a central region of nerve processes, the neuropile, surrounded by a rind of unipolar cell bodies (Fig. 2*e*). Thus, unipolar, heteropolar neurons make their appearance at this time. Bipolar (often sensory) and multipolar cells are also present. (2) Elongated, longitudinally directed bundles of nerve axons appear which can carry information between or among ganglia. These nerve cords are usually arranged in several symmetrical pairs running like barrel staves down the length of the animal. There are often commissures or crossovers of axon bundles between parallel longitudinal bundle pairs. (3) There is a tendency for the neural ganglia to be concentrated anteriorly as a primitive "brain" in association with anteriorly located sense organs such as statocyts and eyespots. This is called encephalization. Thus the preferred forward locomotion of the bilateralia is naturally coupled with anteriorly placed special senses and the headward concentration of neural elements. (4) Definitive unimodal sensory cells, distinct motoneurons, and an increased proportion of interneurons are present. (5) Glial cells are discovered.

The densely packed and complexly interconnecting neural processes in the ganglionic neuropile are correlated with the appearance of a primitive sort of behavior modification akin to memory and learning. Numerous studies have shown that planaria can be conditioned with light to alter their path of movement. This modified, conditioned behavior is retained in organisms which have fully regenerated from the two halves of a previously trained worm which was cut in half. *See* PLATYHELMINTHES.

Aschelminthes. This group of several minor phyla includes the pseudocoelomate roundworms (nematodes) and other interesting small aquatic invertebrates (for example, rotifers). These organisms are intermediate, in their morphology and function, between the flatworms and the annelids. An important advanced feature of the group is that the gut has become tubelike, with a separate mouth and anus. The location of the mouth and associated organs in the head region augments the encephalization of the central nervous system.

Studies have been conducted on the nervous system of the nematodes, notably *Caenorhabditis elegans* and *Ascaris*. *Caenorhabditis elegans* is a small, free-living soil nematode that has been particularly useful in recent studies of the regulation of development and the molecular genetics of nervous tissue. Its hermaphrodite contains exactly 959 somatic cells, and the lineage of these cells, including all neurons, has been traced. The genome of this worm has been entirely sequenced. *See* GASTROTRICHA; KINORHYNCHA; NEMATA (NEMATODA); ROTIFERA.

Annelida. The annelid worms introduce the phenomenon of segmentation (metamerism).

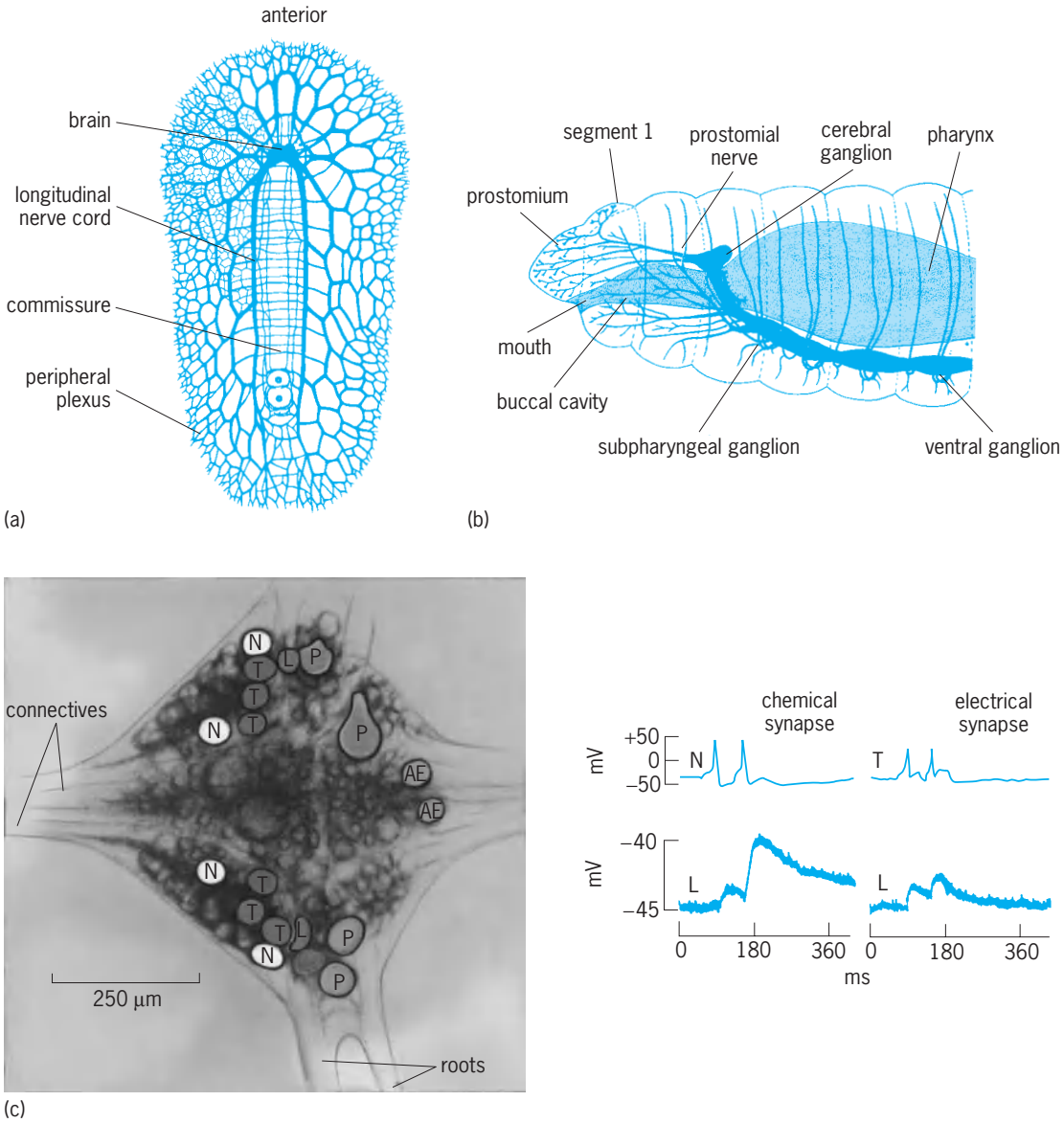

Fig. 4. Nervous systems of the intermediately advanced invertebrates (various worms). (*a*) Turbellarian flatworm (*Notoplana*) illustrating nerve cells coming together to form ganglia and nerve tracts (centralization) and nerve cells concentrated in the head region (brain) near sensory structures (encephalization) (*after D. Hadenfeldt, Das Nervensystem von Stylochoplana maculata and Notoplana atomata, Z. Wiss. Zool., 133:586–638, 1929*). (*b*) Earthworm illustrating circumesophageal brain ganglia, ventral nerve cord, and segmental ganglia (*after E. E. Ruppert and R. D. Barnes, Invertebrate Zoology, 6th ed., Saunders College Publishing, 1994*). (*c*) Micrograph of segmental ganglion from the leech *Hirudo* showing identified sensory and motor neurons. Sensory cells: N = noxious receptor; T = touch receptor; P = pressure receptor. Motor cells: AE = annulus erector cells; L = motor neuron for longitudinal muscle. Connectives are the ventral nerve cord. The electrical records show excitatory synaptic potentials produced in motor neuron L by a presynaptic action potential produced in sensory cell N (chemical synapses) or T (electrical synapses). Note facilitation of second chemical synaptic potential (*after J. G. Nicholls and D. Purves, A comparison of chemical and electrical synaptic transmission between single sensory cells and a motoneuron in the central nervous system of the leech, J. Physiol., 225:637–656, 1972*).

Segmentation is a serial repetition of virtually identical body parts along the longitudinal axis of the body; it is most obvious in the trunk and tail of annelids. Anterior segments are typically fused to form a head region. This kind of organization is evolutionarily inherited by the arthropods, though in modified form. In both the annelids and arthropods segmentation has had a fundamental impact on neural arrangement and activity. Chordate segmentation evolved independently of that in annelids, but the functional consequences are the same. The notion that mollusks descended from segmented ancestors is, at best, controversial; in any case, metamerism is not an obvious feature of any extant mollusk.

In annelids, the well-developed mesoderm hollows out in development to produce a capacious coelom, divided into a series of compartments by membranous septa, and also produces a muscular gut. The coelom plays a critical role in reproduction and excretion and also shapes the structure and

arrangement of the organs of the circulatory, digestive, and nervous systems.

Centralization and encephalization have progressed to the extent that the head contains a distinct and highly organized brain composed of large, fused anterior segmental ganglia (Fig. 4*b*). The major brain component lies dorsal to the anterior digestive organs, and it sends and receives numerous nerve bundles to anterior sense organs and mouthparts. Major nerve trunks leave the dorsal brain mass bilaterally, encircle the esophagus, and connect to other head ganglia of varying number and complexity. From these anterior ganglia, a pair of long nerve trunks, fused to varying degrees, travel back through the length of the animal in a ventral position as the ventral nerve cord, an important organizational feature retained in the arthropods. Typically, in each body segment, a comparatively small segmental ganglion (few hundred neurons) is present. These ganglia innervate the segmental musculature and body organs, and they are virtually identical from segment to segment. Neurons of all functional types are often located on the contralateral (opposite) side of the body from the peripheral organs innervated. Thus, many cells send their processes across the midline (decussate) through commissures in the brain or at the segmental level.

Elaborate sense organs and many types of highly specific unimodal sensory cells are found, as are distinct motoneurons innervating specific muscles and glands. The vast majority of neurons are unipolar, with definite heteropolar functional orientations.

Important functional organization strategies become particularly evident in the annelids. Segmental control of simple reflexes is handled at the local level by neurons in segmental ganglia. For example, touch or painful stimuli in specific receptive fields generate appropriate, localized, and discretely organized reflexive muscle contractions and movements (Fig. 4*c*). The sensory and motor impulses are carried by processes of neurons in the same segmental ganglion. In addition, however, adjacent segments may become involved by extensions of sensory-cell processes to neighboring ganglia and, more importantly, by interneurons "wired" to coordinate multisegmental responses. Also, the brain now contains many interneurons which may send axons down the full length of the ventral nerve cord to receive or transmit information to all segmental ganglia. Thus, central control fibers are present to coordinate overall body functions and behaviors. For example, the rhythmic body undulations, which propel leeches in swimming, require segmental control of body muscle masses, and superimposed central commands from higher-order coordinating interneurons.

The first appearance of at least two specialized kinds of nerve cell or nerve process has been best documented in the annelids. These are neurosecretory or neuroendocrine cells specialized to secrete proteinaceous chemical messenger compounds which act as hormones or widespread effector substances; and giant axons, ranging from around 30 to over 200 micrometers in diameter, and representing fast-conduction "through pathways" for rapid responses. *See* NEUROSECRETION.

The best-studied components of annelid nervous systems are the giant fibers of the earthworm and, particularly, the segmental ganglia of *Hirudo*, the medicinal leech. In the leech, there are 21 segmental ganglia, all identical, with some 350 neurons in each ganglion (Fig. 4*c*). Many of these neurons are identified, meaning that, on the basis of size, color, topography, and electrophysiological and functional characteristics, they can be readily found and reliably identified from ganglion to ganglion and from animal to animal. Specific identified sensory cells and motor neurons have been characterized in detail. Both excitatory and inhibitory motor neurons have been identified; these cause muscle contraction or relaxation, respectively, through neuromuscular synaptic junctions. Synaptic transmitters, such as acetylcholine and serotonin, have been found in particular motor neurons. *See* ANNELIDA.

Arthropoda. This is the largest, most diverse, and most advanced of the invertebrate phyla (though certain of the mollusks are very advanced as well). The most general characteristics of the group are the hard exoskeleton, jointed legs, and varying degrees of segmentation. This is the only group of invertebrates that has successfully invaded the terrestrial and aerial environments. An incredible diversity of form, degree of development, and specialization exists. Some of the highest levels of invertebrate neural complexity, organ development, muscle specialization, and behavioral repertoire are found. Only major highlights are considered by describing a representative arthropod nervous system, modeled after that of the crayfish (**Fig. 5***a*).

The model arthropod is fundamentally metameric, but through fusion of segments, the body has come to be composed of a head, a thorax, and an abdomen. The head ganglia are highly fused to form extraordinarily complex brains, dorsally placed over the esophagus, with a circumesophageal ring of connective nerves and supplementary ganglia. The ganglia have a dense neuropile surrounded by a rind of glia-invested unipolar neurons. Particularly advanced brains, such as in some insects, may have several layers of cells, infoldings, and a complex organization; and numerous multipolar (heteropolar) neurons may occur as well. Millions of nerve cells and highly branched nerve processes are present. A ventral nerve cord (or paired cords) extends the length of the thorax and abdomen. The thorax commonly has one to several fused ganglia; the abdominal ganglia tend to retain serial segmentation and to be identical. All of the ganglia have symmetrical, usually bilateral, nerve trunks coming off to innervate peripheral sense organs or muscles and other effector cells; many of these pathways decussate. Functional tiers or cascades of interneuronal axons run the length of the cord and are within ganglia to organize highly complex behaviors and coordinate the movements of appendages for locomotion or flight. There are typically pairs of giant axons of

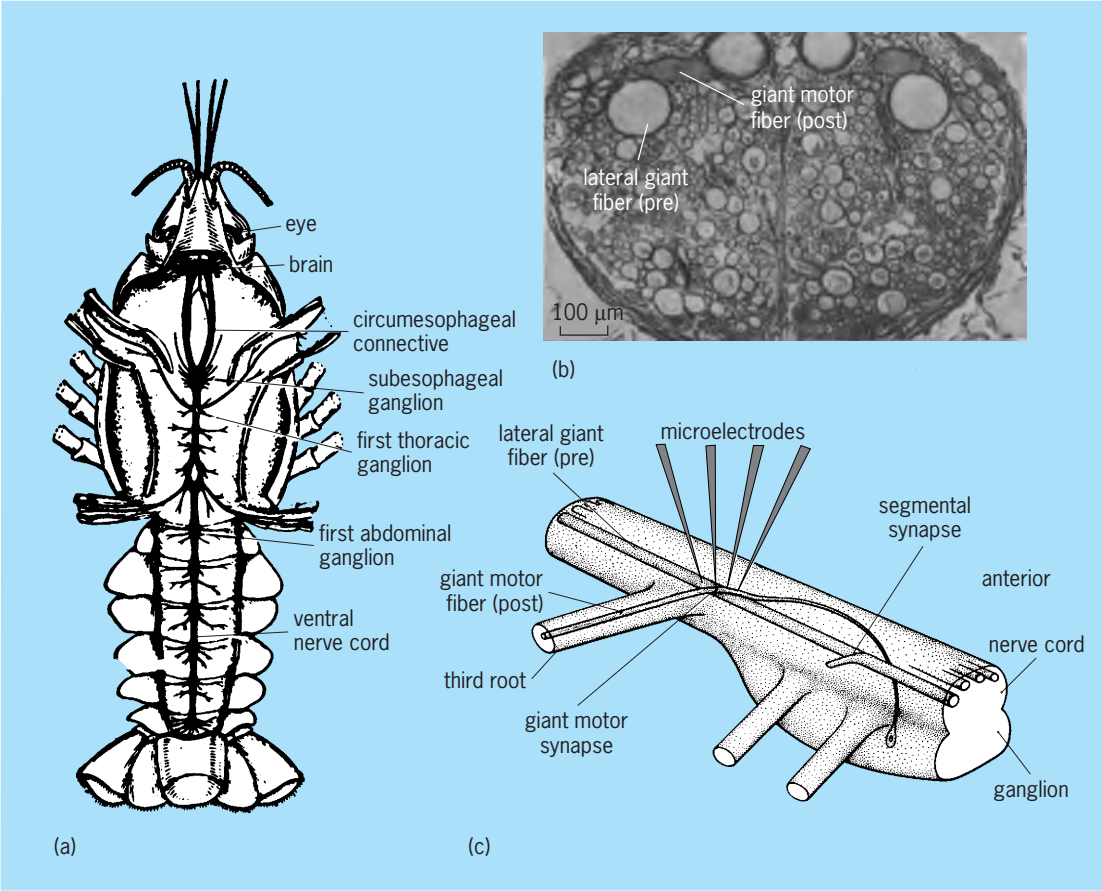

Fig. 5. Advanced invertebrates as represented by the crustacean nervous system. (*a*) Crayfish central nervous system (*after W. C.* Curtis and *J.* and *J. M. Guthrie, Textbook of General Zoology, 3d ed., John Wiley and Sons, 1938*). (*b*) Photomicrograph of cross section of ventral nerve cord in abdomen of crayfish made near the level of the third root (arrow in *a*); the largest axons near the top are giant fibers used in escape responses (*after J. D. Robertson, Ultrastructure of excitable membranes of the crayfish median-giant synapse, Ann. N.Y. Acad. Sci., 94:339–389, 1961*). (*c*) Relationship between lateral giant and motor giant axons as they make electrical synaptic contact near the third root of an abdominal ganglion in the crayfish (*after E. J. Furshpan and D. D. Potter, Transmission at the giant motor synapses of the crayfish, J. Physiol., 145:289–325, 1959*).

interneurons mediating fast-escape behaviors (Fig. 5*b* and *c*).

Vast numbers of highly specific kinds of sensory cells (including muscle stretch receptors) are present, and extremely complex sense organs such as eyes and statocysts exist. The trunk and appendage musculature is highly specialized with a variety of fast- and slow-contracting fibers, each with its own set of identified excitatory and inhibitory motor neurons. Several small peripheral ganglia are present which appear to be somewhat analogous to the vertebrate autonomic nervous system; they control the heart and vasculature, the digestive organs, and other visceral functions. Well-developed neurosecretory and neuroendocrine systems are also present.

Electrical synapses are still found, particularly in the connections of the fast-escape systems employing giant axons, where minimal synaptic transmission time is advantageous. But chemical synaptic transmission is highly evolved, with numerous identified transmitter substances. Excitatory and inhibitory connections, exquisitely sensitive to rates of stimulation and highly plastic, are described; and complex tandem synaptic arrangements are known where one presynaptic terminal ends on another, modifying the effect of the latter on a postsynaptic element.

Probably the best studied of the arthropod nervous systems are those of the crayfish and lobsters, but several insects, for example the locust, cricket, fruit fly, and cockroach, are also important research preparations. Studies of these creatures and of some selected molluscan species have provided some of the most profound insights into how nerve cells function and are organized at the cellular, membrane, and genetic levels. The fruit fly *Drosophila* has been particularly useful in studies of the genetic basis of nervous system development. Some genes involved in neuromere development in the brain of *Drosophila* appear to be homologous to genes expressed in brain development in higher vertebrates. *See* ARTHROPODA.

Mollusca. This highly diverse phylum spans a range of complexity rivaling that of the arthropods. The group includes the well-known clams and mussels, snails and slugs, and squid and octopus. The phylogenetic origins of this phylum are unsettled, but

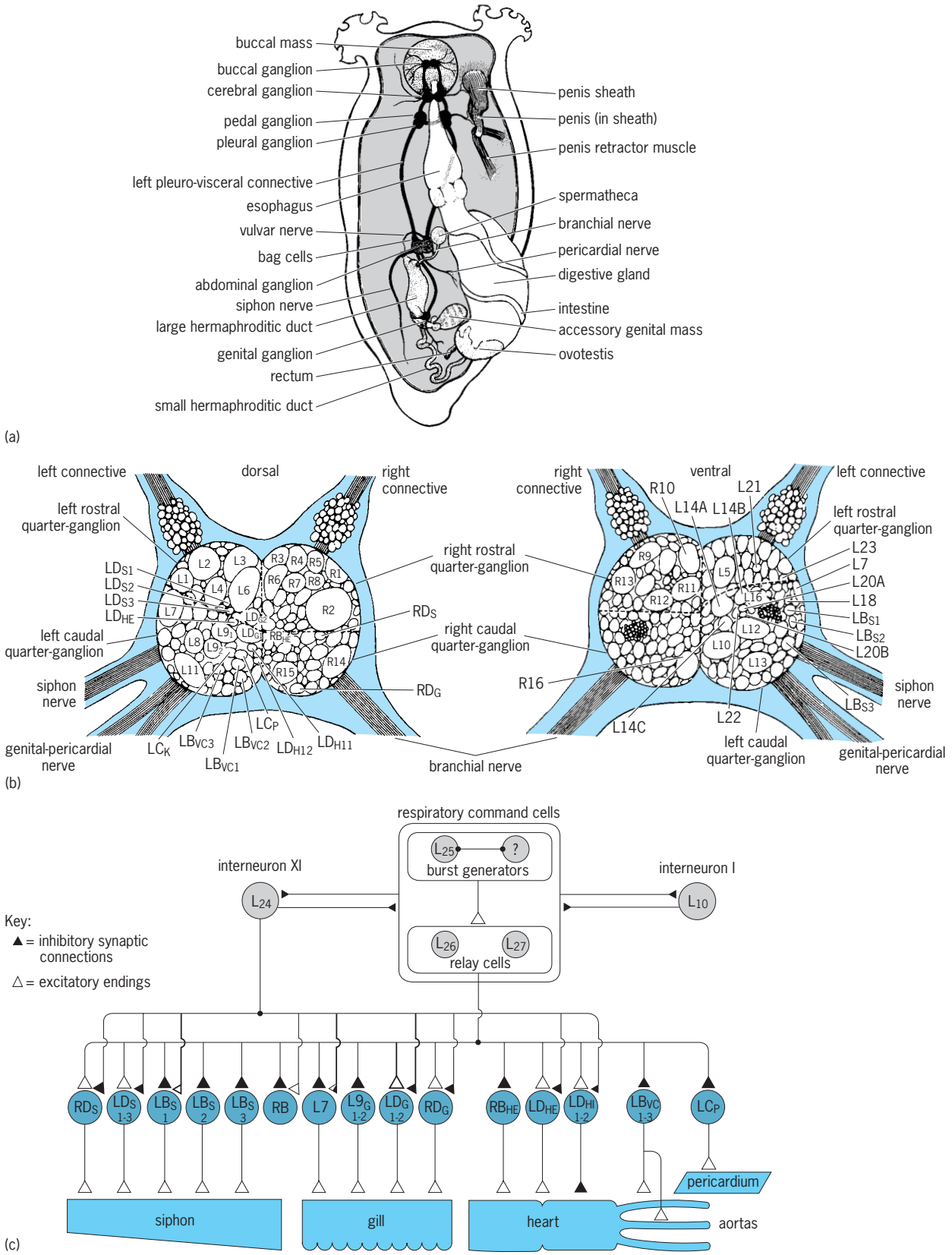

Fig. 6. Advanced invertebrate nervous system as represented in the marine gastropod mollusk *Aplysia*. (*a*) General arrangement of nervous system. (*b*) Map of the dorsal and ventral surfaces of the abdominal, or parietovisceral, ganglion, indicating the many identified nerve cell bodies known to play roles in a variety of functions (*after E. R. Kandel, Cellular Basis of Behavior, W. H. Freeman, 1976*). (*c*) Circuit diagram of several identified interneurons and motor neurons which innervate various organs; solid triangular axon endings represent inhibitory synaptic connections, and open triangles, excitatory endings (*after J. H. Byrne and J. Koester, Respiratory pumping:Neuronal control of a centrally commanded behavior in Aplysia, Brain Res., 143:87–105, 1978*).

primitive mollusks either diverged very early from the annelid-arthropod line or arose from a common ancestor. The phylum as a whole is characterized by a bilateral fleshy body, usually with a prominent foot (muscular locomotor or attachment structure); a mantle cavity built around an external gill and openings of the digestive and excretory systems; a shell secreted by the mantle; and a reduced coelom.

Despite the extraordinary polymorphism of the phylum, a fundamental body plan and characteristic neural arrangement are present, though often obscured by secondary reorganizations. The positions of ganglia and arrangements of nerves are modified to accommodate the flattened body and hinged shells of the bivalves, the twisted or torqued bodies of the gastropods, or the sleek streamlining of the highly active cephalopods. An example of a gastropod nervous system is illustrated in **Fig. 6***a*.

The head contains a circumesophageal ring of ganglia, the bilateral brain, which sends numerous nerves to the eyes, tentacles, and other richly diverse and highly developed sensory structures and sensory neurons. From the brain there are also long nerve cords, usually paired, which interconnect other ganglionic masses in the body. There are usually several subsidiary ganglia in mollusks, including ones primarily involved in innervating the foot, the mantle and visceral structures, and the mouthparts. A map of the identified neurons in the abdominal or parietovisceral ganglion of *Aplysia* is shown in Fig. 6*b*. Depending on the class of mollusk and the type of secondary alterations in body shape and configuration, the subsidiary ganglia may lie in various positions in the body (normally close to the body mass innervated), or they may be fused with the brain. Several types of muscle are present, usually of a slow type, both in the body wall (including the foot) and in visceral structures (gill, gut, reproductive organs). Faster muscles, somewhat more akin to arthropod or vertebrate types, may be found in buccal (mouthpart) muscle, in mantle muscle of squid, in retractor muscles which serve to draw the animal back into the shell, and in adductor muscles which close the shells.

There are a few unipolar peripheral neurons in the skin or various organs, but the great bulk of neurons are clustered into conventional ganglia composed of a central neuropile and an outer rind of unipolar (heteropolar) cell bodies. Except in very complex brains, such as that of the rather intelligent octopus, which has the largest invertebrate brain, there are only rare multipolar (heteropolar) neurons. The mollusks are exceptional in that so many species have the largest neuron cell bodies or axons (several hundred micrometers in diameter) known. Neuroendocrine and glial cells are also present. The ganglia of higher mollusks such as gastropods and cephalopods have many interneurons which serve to coordinate and control complex behaviors (Fig. 6*c*).

The most famous molluscan research preparation is the giant axon of the squid from which some of the earliest intracellular recordings of the action potential were obtained (Fig. 1*a*). This axon was also the experimental preparation used by Hodgkin and Huxley to describe the membrane properties that cause an action potential to occur—work that earned them a Nobel prize in 1963. The other very widely studied mollusks are certain gastropods, particularly the marine snail *Aplysia* (Fig. 6) and the terrestrial, and aquatic pulmonates, typified by *Helix*, *Heliosoma*, and *Lymnaea*. As with leech, crayfish, and insect ganglia, *Aplysia* and other gastropod nervous systems have been exhaustively studied, and many identified nerve cells and detailed synaptic circuits have been described in great detail. Neural control of feeding, respiration, circulation, inking, reproduction, and other activities have been unraveled to a large extent. The cellular mechanisms underlying simple kinds of learning or memory, such as habituation and sensitization, are also well understood. Some electrical synapses exist, but a rich variety of very complex chemical synaptic connections are present, and their plastic properties (and many transmitter substances, including peptides) have been identified. Eric Kandel won a Nobel prize in 2000 for his work on the biochemical basis of learning and memory in *Aplysia*.

Cephalopods, including octopuses and squid, may have the most highly evolved and complex nervous system among all invertebrates. In *Octopus*, the central brain surrounds the esophagus, with lateral paired optic lobes that may contain over 65 million nerve cells. Cephalopods have been subjects in extensive studies of learning, sensory reception, and behavior. The image-forming eye of coleoid cephalopods is similar in morphology and complexity to the eye of higher vertebrates. *See* MOLLUSCA.

Echinodermata. The echinoderms, while phylogenetically near the protochordates, have undergone such a great deal of secondary adaptation and specialization in the adult form that their nervous systems are less well defined and studied than those of lower forms.

The nervous system of these radially symmetrical organisms is less centralized, having a subepidermal nerve-ring arrangement around the mouth, with radial nerves extending from it. Few and small ganglia exist. Uni- and multipolar neurons are present. There are large numbers of individual sensory cells, and few specialized sense organs. Superficial (mainly sensory) and deeper (motor) nerve nets are present. A variety of body wall and specialized muscles are present to control locomotion and food capture. The coordination of movement of the spines and arms, the pedicellariae, and the tube-foot/water-vascular system requires a rich variety of motor neurons and interneurons. *See* ECHINODERMATA; ENDOCRINE SYSTEM (INVERTEBRATE).

James E. Blankenship; Becky Houck

Bibliography. C. I. Bargmann, Neurobiology of the *Caenorhabditis elegans* genome, *Science*, 282:2028–2033, 1998; J. H. Byrne, Cellular analysis of associative learning, *Physiol. Rev.*, 67:329–439, 1987; R. F. Chapman, *The Insects: Structure and Function*, 4th ed., Cambridge University Press, 1998; R. T. Hanlon and J. B. Messenger, *Cephalopod Behaviour*, Cambridge University Press, 1996;

G. G. Lunt and R. W. Olsen (eds.), *Comparative Invertebrate Neurochemistry*, Cornell University Press, Ithaca, NY, 1988; D. Osorio, J. P. Bacon, and P. M. Whitington, The evolution of arthropod nervous systems, *Amer. Sci.*, 85(3):244–253, 1997; J. A. Pechenik, *Biology of the Invertebrates*, 4th ed., McGraw-Hill, Boston, 2000.

Nervous system (vertebrate)

A coordinating and integrating system which functions in the adaptation of an organism to its environment. An environmental stimulus causes a response in an organism when specialized structures, receptors, are excited. Excitations are conducted by nerves to effectors which act to adapt the organism to the changed conditions of the environment. In animals, humoral correlation is controlled by the activities of the endocrine system. This article considers the morphology, histology, and embryology of the nervous system, including the brain and cranial nerves, and embryology of the sense organs.

Comparative Morphology

The brain of all vertebrates, including humans, consists of three basic divisions: prosencephalon, mesencephalon, and rhombencephalon (**Fig. 1**). Newer experimental methods for exploring the connections and functions have resulted in much information on the evolution of this system. The indication is that these divisions of the vertebrate brain have evolved along several functional lines and perform very different functions.

The comparative neuroanatomist is interested in the variation among the vertebrate brains, and attempts to understand how they evolved from common ancestors and to clarify the functional significance of their variations. It is necessary that the neuroanatomist recognize structures in different brains among living forms which have arisen from one or more structures in an ancestral brain. This search for homologous structures is complicated by the fact that brains do not fossilize; thus the comparative anatomist can only infer what the possible structure of the brain in ancestral vertebrates was like. It is thought that the brain of ancestral vertebrates was much simpler than those of living forms, but neuroanatomists suggest different views regarding the brain's evolution.

One view holds that structures have arisen afresh with no traces in lower living vertebrates. The second view proposes that the simpler ancestral brain possessed very generalized connections which have either been lost or have become specialized in living vertebrates. The latter idea seems to be closer to the truth and is more in line with what is known about evolution of other vertebrate body systems.

The nervous system of ancestral vertebrates apparently consisted of a single hollow neural tube,

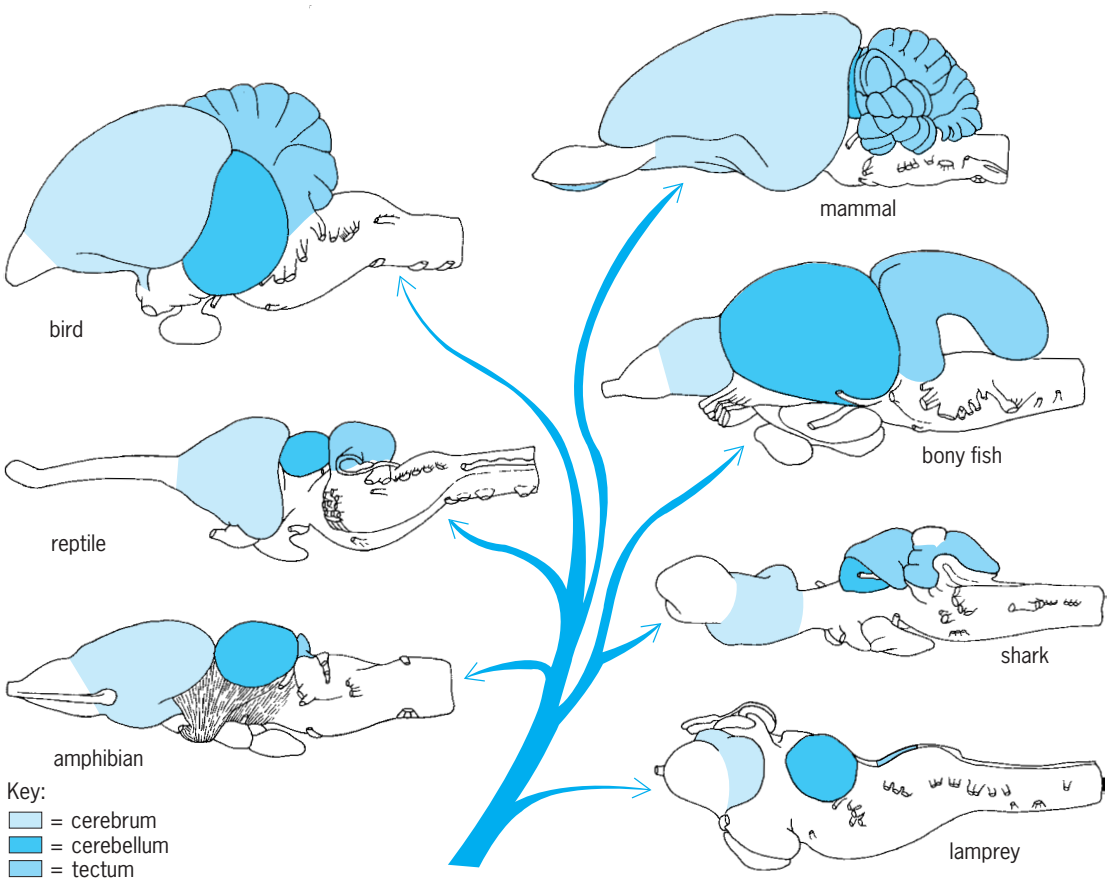

Fig. 1. Lateral views of several vertebrate brains showing evolutionary relationships.

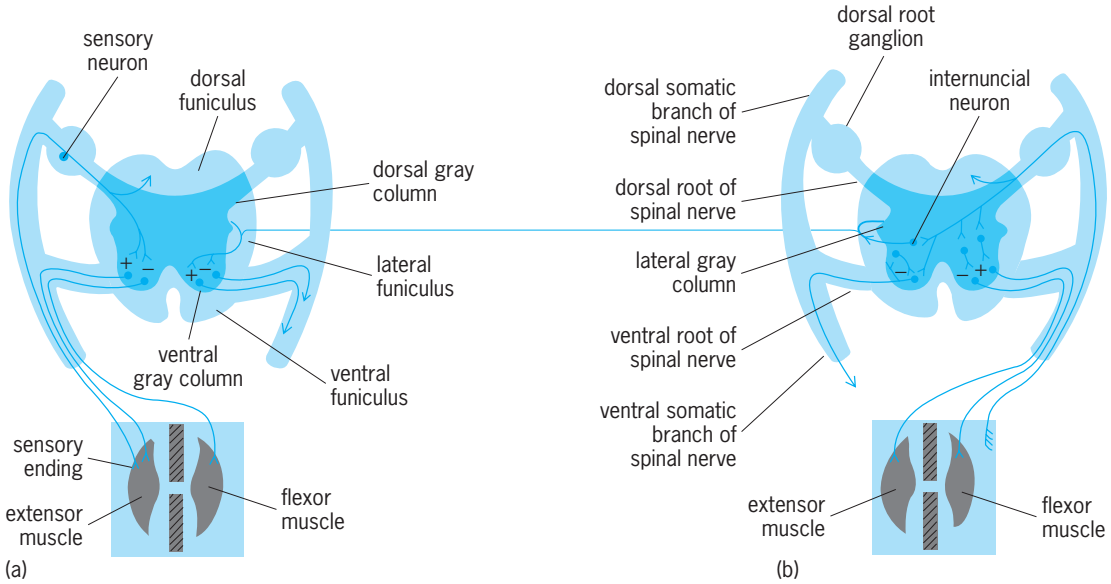

Fig. 2. Spinal cord diagram. (*a*) Monosynaptic arcs. (*b*) Multisynaptic arcs.

dorsally placed and running the length of the animal, that was probably structured in a manner similar to the spinal cord in living vertebrates (**Fig. 2**). It consisted of a central gray region which contained most of the cell bodies of the neurons and was surrounded by a superficial white region containing the axons and dendrites of the neurons running to and from the tube. Each region of the tube was primarily concerned with the functions of that immediate region of the body to which its nerves projected. Neural functions were mediated at a segmental level rather than specialized in one specific region. Each region of the neural tube also contained neural cells which interconnected adjacent regions. This primitive condition still persists in a few living forms.

As vertebrates evolved, structures in the head region became specialized for sensing the outside world and capturing food. This specialization was reflected in the neural tube, and its anterior region enlarged to permit analysis of the environment and integration of behavior. This enlargement of the neural tube is called a brain, and its divisions represent regions of different specialization. Each of the three primitive brain divisions was concerned with analysis and integration of a single type of sensory information. The rhombencephalon analyzed changes in the flow and pressure of water on an animal; the mesencephalon analyzed changes in the pattern and intensity of light; and the prosencephalon analyzed changes in the chemical composition of water. All vertebrates use these three types of information as clues in regulating their behavior. However, major changes in the structures that analyze this information appear to be correlated with increased quality of perception and more centralized control of movement. This has occurred by an increase in the number of analyzing units between the sensory mechanisms and the response mechanisms. This increase permits a delay in the response of a vertebrate to sensory cues and thus permits greater analysis of information contained in the clues. Presumably, this has value to an animal in that it results in responses which have a higher degree of survival value.

The evolution of the central nervous system is described below as the evolution of functional systems or patterns whose specialization is reflected by the concomitant specialization of brain divisions. The individual divisions or patterns of the brain do not function separately to bring about a final response; rather, each pattern acts on a common set of connections in the spinal cord. The common spinal connections are first described and then used as a basis for understanding the action of the brain patterns upon them.

Spinal patterns. These are the final common patterns used by all higher brain pathways to influence all organs of the body. These reflexes are divided into two basic patterns: the monosynaptic arc and the multisynaptic arc.

The monosynaptic arc, or myotatic reflex, maintains tonus and posture in vertebrates and consists of two neurons, a sensory and a motor neuron. The sensory neuron possesses a sensory ending in the extensor muscle (Fig. 2*a*), and this nerve ending is stimulated when the muscle is stretched by the pull of gravity. The axon of the sensory neuron projects to the motor neuron of the extensor muscle. This motor neuron is located in the ventral, gray region of the spinal cord, and its axon projects back to the extensor muscle, thus completing the arc.

When the extensor muscle is stretched, the sensory ending of the sensory neuron is stimulated, and this stimulation is transmitted to the cell body of the sensory neuron. The sensory cell body then transmits a signal to the motor neuron of the extensor muscle, which will cause the motor neuron to transmit. At the same time, the sensory neuron causes the motor neuron of the flexor muscle to decrease its transmission. These processes are indicated respectively by plus and minus signs in Fig. 2. The result

is that the extensor motor neuron causes the extensor muscle to contract, and the flexor motor neuron causes the flexor muscle to relax. This extends the limb and supports the weight of the animal. The sensory neuron of the extensor muscle also sends an axonic collateral to brain centers where the sensory information is analyzed.

The multisynaptic arc, or flexor reflex, is the pattern by which an animal withdraws a part of its body from a noxious stimulus. Stimulation of the skin or muscle causes a sensory neuron to transmit a signal causing motor neurons, via internuncial neurons, to increase the contraction of flexor muscles and decrease the contraction of extensor muscles. The result is movement of a limb or body part away from the stimulus. Both sensory neurons and internuncial neurons send information to brain centers (Fig. 2).

The multisynaptic arc also gives rise to a pattern that results in alternating movement of a forelimb and the diagonal hindlimb. This is termed reflex stepping. Thus coordinated limb movement is based on a connective pattern of neurons at the spinal level.

The structure of the spinal cord and its connections are basically similar among all vertebrates, and myotatic and flexor reflexes occur in all vertebrates. The major evolutionary changes in the spinal cord have been the increased segregation of cells and fibers of a common function from cells and fibers of other functions and the increase in the length of fibers which connect brain centers with spinal centers. *See* POSTURAL EQUILIBRIUM.

Medullar patterns. The rhombencephalon of the brain is subdivided into a roof, or cerebellum, and a floor, or medulla oblongata. The medulla is similar to the spinal cord and is divided into a dorsal sensory region and a ventral motor region. The sensory region consists of two longitudinal columns; a dorsal, somatic sensory column and a ventral, visceral sensory column. The ventral motor region of the medulla is similarly divided into two longitudinal columns, a dorsal, visceral motor column and a ventral, somatic motor column. These divisions occur in the spinal cord and continue through the medulla into the tegmentum, the floor of the mesencephalon.

The medulla is an integrating and relay area between higher brain centers and the spinal cord. Primitively, the columns were composed of neurons scattered throughout the length of the medulla. With the specialization of the vertebrate head, the columns broke up into a series of nuclei that formed the neural elements for reflex circuits of the cranial nerves innervating the organs of the head and gills.

In addition to these nuclei and their connections, the medulla consists of both ascending and descending pathways to and from higher brain centers. The nuclei of the medulla shift with reference to one another, and their relative volumes change as specialization in gill and feeding mechanisms occurs. However, the same basic connections occur throughout vertebrates. The tracts which pass through the medulla between higher brain centers and the spinal cord become more distinguishable as their boundaries become more compact in relation to other tracts. The tracts, like the nuclei of the medulla, may change volume in response to specialization.

Cerebellar patterns. The cerebellum has had a varied history of development in vertebrates. In all vertebrates, the cerebellum is divided into two major divisions: the two lateral flocculonodular lobes, and a central corpus cerebelli. The flocculonodular lobes are functionally referred to as the vestibulocerebellum and regulate vestibular reflexes underlying posture. The corpus cerebelli is subdivided into two lateral zones, the cerebrocerebellum and a central zone, the somatocerebellum. In mammals the lateral zones regulate corrective reflexes of posture and time muscular contractions of voluntary actions. The medial zone in mammals regulates reflex tonus of postural muscles by acting on the myotatic reflex.

The flocculonodular lobes, or vestibulocerebellum, evolved in conjunction with the vestibular system of the inner ear and is mainly responsible for gravitational orientation in higher vertebrates. The vestibular division of the inner ear senses changes in the orientation of an animal's head in all three planes and controls the postural reflexes of the head and neck. These vestibular centers also influence the spinal patterns of posture described above. Vestibular connections are made with the flocculonodular lobes and their nuclei, the fastigial nuclei. This system allows integration of vestibular sensations with other sensory information which passes to the cerebellum from the spinal cord.

The evolutionary changes of the corpus cerebelli in land vertebrates resulted in a new mode of locomotion. The limbs were used to support an animal off the ground and for locomotion as well. It became essential for a land vertebrate to maintain posture and to know the position of its limbs in order to move them in an organized manner.

Mammals. Information on the position of the limbs, and on the state of tonus in the muscles of limbs, is transmitted to the cerebellum via the posterior cerebellar peduncle (**Fig. 3**). The

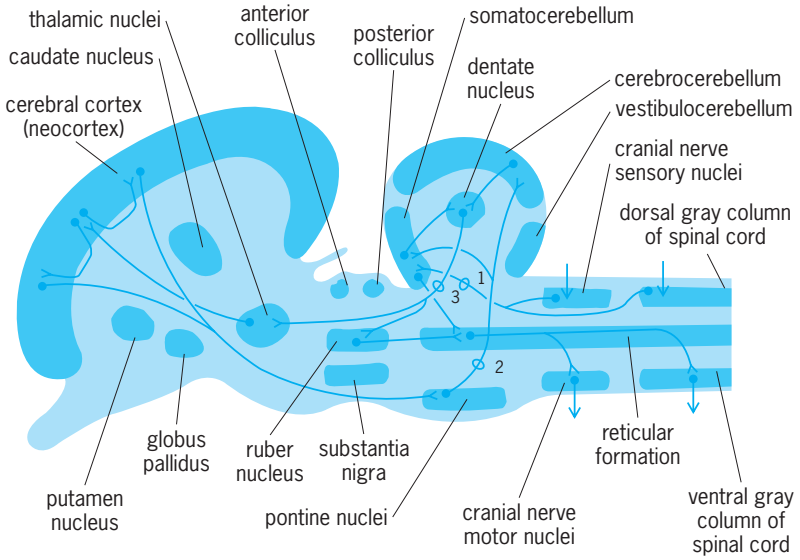

Fig. 3. Mammalian brain in sagittal section. Cerebellar patterns: tract 1, posterior cerebellar peduncle; 2, middle cerebellar peduncle; 3, anterior cerebellar peduncle.

cerebellum does not initiate movement in mammals; it only times the length of muscle contractions and orders the sequence in which muscles should contract to bring about a movement. The command to initiate a movement is received from the cerebral cortex via the middle cerebellar peduncle (Fig. 3). Similarly, the cerebral cortex receives information regarding limb position and state of muscular contraction to ensure that its commands can be carried out by the cerebellum.

When the cerebellum receives a command from the cerebral cortex, it transmits a pattern of impulses to the spinal patterns via the dentate nucleus and lower pathways. This pattern of impulses contains the sequence of muscles to be contracted and the length of time of these contractions. At the same time, this same information is transmitted from the cerebellum to the cerebral cortex via the anterior cerebellar peduncle (Fig. 3). The cerebral cortex can thus compare its command with the cerebellum's action. This information is then compared with new information coming to the cerebral cortex about changes in the limbs as a result of the initial movement. If the movement was not adequate, the cerebral cortex can then issue a corrected command to the cerebellum for a new series of movements.

Birds. A similar system for control of locomotion is apparently present in birds. However, the exact nature of the command center, or centers, in the telencephalon is not understood. Birds possess a middle cerebellar peduncle, but the tracts from the telencephalon may not originate in the same areas as in mammals. Birds have a cortical command center, but whether it is homologous with the neocortex of mammals is not known.

Reptiles. Knowledge of the reptilian cerebellar patterns is more fragmentary than those of birds. Reptiles possess a pathway, the dorsal forebrain bundle, from the telencephalon to the cerebellum (**Fig. 4**). This pathway may be homologous, in part, with the middle cerebellar peduncle. The telencephalic center from which this pathway originates is similarly organized in both reptiles and birds. Reptiles possess an interior cerebellar peduncle, but a homologous pathway similar to the mammalian component projecting to the cortex has not been identified. It appears that reptiles, birds, and mammals possess cerebellar patterns of a similar nature. There are differences in organization, but they appear to be secondary and may reflect the added role of the tectum as a second higher center in reptiles and birds.

Amphibians. The cerebellum in amphibians is not highly developed. It is divided into the same divisions as in higher vertebrates, and its histology is identical. The locomotor pattern in amphibians is very different from that of other vertebrates, and it is improbable that it functions in the same manner. Salamanders do not use the limbs for support as do higher vertebrates, and frogs have undergone great specialization for jumping. There is little information on the role of the cerebellum in these forms.

Fishes. The cerebellum, particularly the corpus cerebelli, in sharks and bony fish is large. Part of the size is due to the huge number of gustatory fibers projecting to the corpus cerebelli. Both groups of fish possess anterior and posterior cerebellar peduncles. The connections of the anterior cerebellar peduncle are not understood. In bony fish, the tectum is expansive and forms a complex cortex. Tectocerebellar connections are large, indicating that a tectal control center probably exists analogous to the cerebral control center in mammals. Similar tectocerebellar connections exist in other vertebrates, but they do not exist as a highly developed complex.

Tectal patterns. The mesencephalon is divided into a roof or optic tectum and a floor or tegmentum. The tegmentum contains the nuclei of the oculomotor and trochlear cranial nerves and a rostral continuation of the sensory nucleus of the trigeminal cranial nerve. Two motor nuclei dominate the tegmentum, the rubber nucleus and the substantia nigra. These nuclei are elements in both the telencephalic and the cerebellar motor systems. All of these nuclei are recognizable in most vertebrates. The tegmentum also contains sensory and motor pathways which project between the prosencephalon and lower brain and spinal centers.

The roof of the mesencephalon in lower vertebrates is usually referred to as the optic tectum. Early anatomists realized that the major connections of the optic nerve formed in this area and that this area represented the major center for analysis of visual information. However, in lower vertebrates the optic tectum is also a major center for analysis of somatic sensory information and contains a highly structured center for control of somatic musculature.

The external shape of the tectum has changed greatly in the course of evolution (Fig. 1). In fish, amphibians, and reptiles, it forms the roof of the mesencephalon and remains in a dorsal position. In all of these forms, it is normally seen as a paired swelling between the telencephalon and the cerebellum. In birds, each lobe of the tectum has shifted ventrally, and the telencephalon and cerebellum have expanded until they meet along the dorsal surface. In

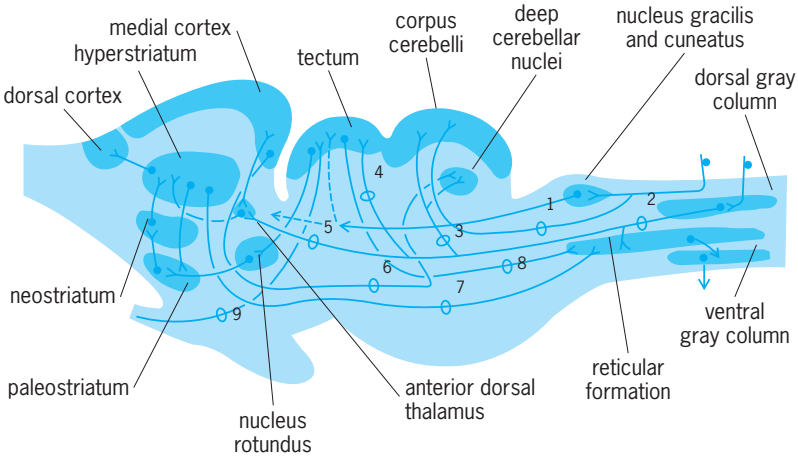

Fig. 4. Reptilian brain in sagittal section illustrating tectal patterns: tract 1, dorsal spinocerebellar tract; 2, lateral funiculus; 3, ventral spinocerebellar tract; 4, spinomesencephalic tract; 5, spinothalamic tract; 6, dorsal forebrain bundle; 7, ventral forebrain bundle; 8, tectospinal tract; and 9, optic tract.

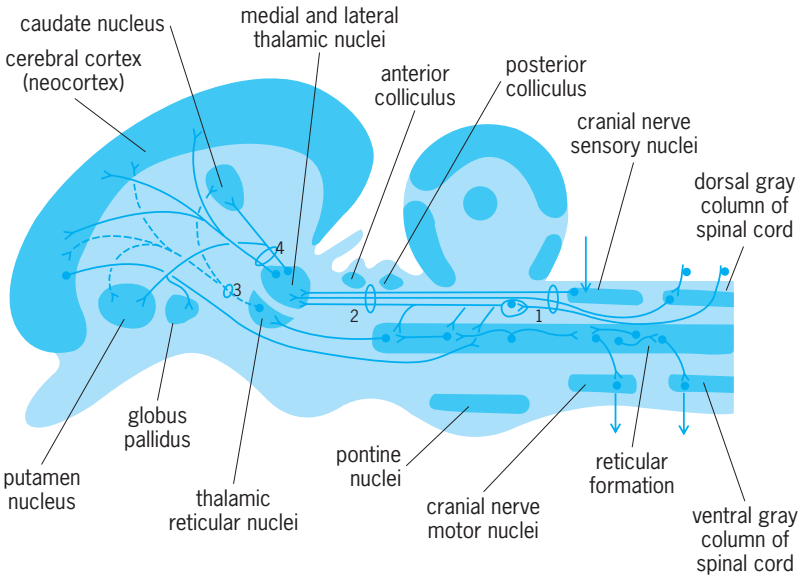

Fig. 5. Mammalian brain in sagittal section illustrating thalamic patterns: tract 1, dorsal and lateral spinal funiculi; 2, lemniscal system; 3, ascending reticular arousal system; and 4, specific thalamic system.

mammals, the tectum has retained its primitive position but is not a large structure. The telencephalon covers its anterior surface, and the cerebellum, its posterior surface; thus it is often not visible from the surface.

In all vertebrates the tectum is divided into two functional regions, an anterior optically dominated region and a posterior dominated auditory region. These divisions may or may not be reflected in the surface structure of the tectum. In mammals, the tectum receives visual information from a division of the optic tract, but the majority of the optic fibers project to the cerebral cortex of the telencephalon after synapsing in a diencephalic nucleus, the lateral geniculate body. This nucleus exists in lower vertebrates and receives optic fibers, as in mammals, but their number is smaller, and the rostral projections, if any, are not known.

A similar condition exists in the auditory pathways. In all vertebrates except mammals, the primary auditory pathway projects to the posterior division of the tectum. In lower vertebrates, this division is called the torus semicircularis. In mammals, some auditory projections occur in the posterior region of the tectum, but the majority of the fibers project to the cerebral cortex of the telencephalon via a diencephalic nucleus, the medial geniculate body.

The tectal pattern in reptiles illustrates the basic connections which occur in all lower vertebrates (Fig. 4). The exact functions of these tectal patterns in reptiles and birds have not been discovered. In these two groups, where the telencephalon is also highly developed, the interactions between telencephalon and tectum appear to be very complex.

Diencephalic patterns. In the evolution of vertebrates, the prosencephalon develops as two major divisions, the diencephalon and the telencephalon. The diencephalon retains the tubular form and serves as a relay and integrating center for information passing to and from the telencephalon and lower centers. The telencephalon is divided into a pair of cerebral hemispheres and an unpaired telencephalon medium.

There are three divisions of the diencephalon in all vertebrates: an epithalamus which forms the roof of the neural tube, a thalamus which forms the walls of the neural tube, and a hypothalamus which forms the floor of the neural tube. The epithalamus and hypothalamus are primarily concerned with autonomic functions such as homeostasis. *See* HOMEOSTASIS; INSTINCTIVE BEHAVIOR.

In many vertebrates the epithalamus develops two outpocketings from its roof. These are the pineal and parietal processes. The pineal process usually forms a gland with endocrine functions. The parietal process may form an endocrine gland or an eye. In some living reptiles, the parietal eye is highly developed. It apparently was common in many fossil groups as well. *See* ENDOCRINE SYSTEM (VERTEBRATE); PINEAL GLAND.

The thalamus is subdivided into dorsal and ventral regions. The dorsal region relays and integrates sensory information, and the ventral thalamus relays and integrates motor information. In amphibians, the dorsal thalamus is represented by a single group of neurons. In reptiles and birds, this group of neurons has been broken up into a series of nuclei having little, if any, relationship to nuclei found in the thalamus of mammals. Bony fish have a well-developed dorsal thalamus with many nuclei which have been homologized with those of other lower vertebrates, but their functions are unknown.

The dorsal thalamus in higher vertebrates has evolved along two different lines. In reptiles and birds, the dorsal thalamus receives some sensory connections from the spinal cord, but the majority of its connections are with the tectum. In these forms the dorsal thalamus is primarily a relay and integrating center between the tectum and the telencephalon (Fig. 4). In mammals, the dorsal thalamus receives connections from the tectum, but the majority of its connections are with the spinal cord and the telencephalon (**Fig. 5**). Thus the dorsal thalamus is a relay and integrating center for sensory information passing from the spinal cord to the telencephalon. Two types of sensory systems project to the telencephalon. The ascending reticular arousal system (Fig. 5) passes to the telencephalon and alerts the cortical fields which analyze sensory information. A second system, the specific sensory, or lemniscal, system passes to the sensory cortex and carries specific information on the nature of the stimulus and the position of its origin on the body or head.

Telencephalic patterns. The telencephalon is the most complex brain division in vertebrates. It is divided into a roof, or pallium, and a floor, or basal region. The pallium is divided into three primary divisions (**Fig. 6**): a medial PI or hippocampal division, a dorsal PII or general pallial division, and a lateral PII division, often called the pyriform pallium. The basal region is divided into three areas, the first of which

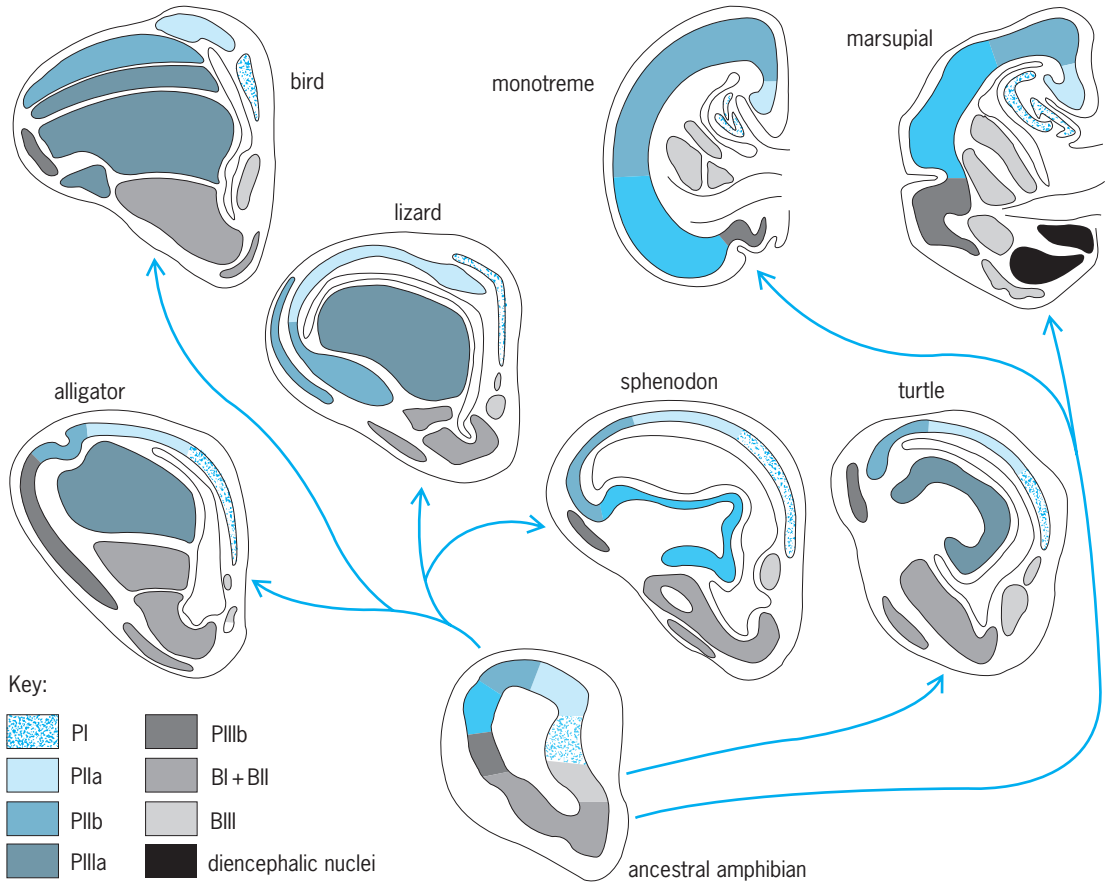

Fig. 6. Cross sections of the right telencephalic hemisphere of several vertebrates, showing evolutionary relationships. Homologous regions of the pallium and corpus striatum are similarly coded.

is a medial BIII area called the septum. A ventral BII area and a lateral BI area form a region often called the corpus striatum.

The most striking change in the telencephalon of land vertebrates involves the PIIIa component, often called the dorsal ventricular ridge (Fig. 6). In reptiles and birds, this region has proliferated into the ventricle of the telencephalon to produce a large cellular mass. In mammals, it has proliferated with the PIIb component of the dorsal pallium to produce the mammalian neocortex. In all land vertebrates except amphibians, the PIIb and the PIIIa components, along with the corpus striatum (BI and BII), are the highest centers for the analysis of sensory information and motor coordination. The PI, PIIa, PIIIb, BIII, and posterior parts of BI and BII form part of the limbic system which is concerned with behavioral regulation.

Older theories stated that the telencephalon first possessed only olfactory fibers and that sensory fibers from the thalamus occurred later in evolution, which resulted in the formation of a neocortex. It is now believed that thalamic fibers occur in the PIIIa component of most vertebrates and that all of the vertebrates possess structures of PIIIa origin even though all these structures are not recognized as a neocortex.

Bony fish possess a pallium and a basal region in the telencephalon, but these are organized along different lines than in other vertebrates. Thus the same divisions cannot be recognized.

Cortical representation. The development of a cortex appears to be the most common way in which analyzing and integrating units can be put together in a limited space to produce maximal efficiency. This is reflected in the fact that all three brain divisions have evolved cortices along one or more evolutionary lines. Each evolutionary line has not relied to the same extent on a particular type of sensory information or its utilization. However, each line has developed cortical surfaces which are specialized for different functions. For example, in mammals the posterior division of the cerebral cortex is concerned with visual impulses, and part of the lateral cortex is concerned with auditory impulses. Anterior to these regions are zones concerned with general sensory and motor impulses. R. Glenn Northcutt

Comparative Histology

The vertebrate nervous system is composed of nervous tissue, which is one of the four primary tissues of the body. The basic microscopic anatomy of this tissue is essentially similar in all vertebrates although variations do exist. Nervous tissue has the quality of irritability. It is also characterized by the quality of conductivity, to convey the resulting excitation to other structures of the nervous system. The functional roles of nervous tissue include the

capabilities to sense, through specialized receptors, environmental energies both internal and external to the organism; to conduct the resulting nerve impulses as coded input to centers in the nervous system; to process this input within these centers; to generate sensations and psychological expressions; and to produce such active responses as the contraction of muscles or the secretions of glands. In effect, nervous tissue reacts to environmental stimuli and regulates many bodily processes so as to maintain functional integrity of an organism. It is within the morphological, physiological, and chemical matrices of nervous tissue that the substrates for memory, behavior, and personality reside. *See* LEARNING MECHANISMS; MEMORY; PERSONALITY THEORY.

Structural elements. The nervous system is composed of several basic cell types, including nerve cells called neurons, interstitial cells called neurolemma (cells of Schwann), satellite cells, oligodendroglia, and astroglia; and several connective-tissue cell types, including fibroblasts and microglia, blood vessels, and extracellular fluids.

Neuron. Each neuron possesses three fundamental properties, involving specialized capacity to react to stimuli, to transmit the resulting excitation rapidly to other portions of the cell, and to influence other neurons, muscle, or glandular cells. Each neuron consists of a cell body (soma), one to several cytoplasmic processes called dendrites, and one process called an axon.

Cell bodies vary from about 7 to more than 70 micrometers in diameter; each contains a nucleus and several cytoplasmic structures, including Nissl (chromophil) granules, mitochondria, and neurofibrils. The cell body is continuously synthesizing new cytoplasm, especially protein, which flows down the cell processes. The Nissl granules, found in dendrites as well as in the cell body, are rich in ribonucleoprotein and are responsible for synthesis of proteins for the neuron.

The mitochondria, found throughout the neuron, are the power plants involved in numerous chemical reactions within the cell. The neurofibrils are fine filaments found in living neurons; they are not directly related to conduction of the nerve impulse. The neuroplasm is the structureless ground substance of the cytoplasm in which are found high concentrations of potassium ions and other substrates critical to conduction of impulses and to cellular metabolism.

The dendrites range from a fraction of a millimeter to a few millimeters in length. An axon may range from about a millimeter up to many feet in length.

A most significant structure of the neuron is its cell membrane that acts as a sievelike barrier between the neuronal cytoplasm, which is negatively charged, and the extracellular fluid, which is positively charged. The functional integrity of the cell membrane and the difference in the bioelectric potential across this membrane are crucial to the physiology of the nerve impulse.

The site where two neurons come into contact with each other and where influences of one neuron are transmitted to the other neuron is called a synapse. A synapse between an axon and a cell body is an axosomatic synapse, and that between an axon and a dendrite is called an axodendritic synapse. At each synapse there is a microscopic space, the synaptic cleft, between the two cells about 20 nanometers wide.

The cell membrane of the axon at the synapse is called the presynaptic membrane and that of the cell body or dendrite is called the postsynaptic membrane. The axon contains vesicles in the vicinity of the presynaptic membrane called presynaptic vesicles, which contain precursors of the neurotransmitter chemicals such as acetylcholine. The neurotransmitters are secreted across the presynaptic membrane into the synaptic cleft where they may excite (excitatory synapse) or inhibit (inhibitory synapse) the postsynaptic membrane. Although a nerve fiber may transmit an impulse in either direction (toward or away from the cell body), conduction in a sequence of neurons is unidirectional: The impulse moves toward the cell body through dendrites and away from the cell body through the axon. This direction is established because the presynaptic neuron can stimulate the postsynaptic neuron but the postsynaptic neuron cannot stimulate the presynaptic neuron. In this context, the synapse acts as a one-way valve and thereby establishes the functional polarity of the neuron. The site of contact of a nerve with a muscle, the motor end plate, is actually a synapse between a nerve and a muscle cell. *See* ACETYLCHOLINE; BIOPOTENTIALS AND IONIC CURRENTS; SENSATION; SYNAPTIC TRANSMISSION.

Interstitial cells. Except at their synapses and endings, the neurons are in intimate contact with the interstitial cells, the neurolemma (Schwann) cells and satellite cells in the peripheral nervous system, and the astroglia and the oligodendroglia in the central nervous system.

The axon of a peripheral nerve is enveloped by a sequence of neurolemmal cells forming a neurolemmal sheath. In some nerve fibers, the neurolemma elaborates a complex lipid and protein layer called myelin which forms a sheath that is segmented by interruptions at short intervals called nodes of Ranvier. Unmyelinated fibers, nerve fibers without a myelin sheath, conduct nerve impulses at slow velocities, up to about 10 ft/s (3 m/s), while myelinated fibers conduct nerve impulses at speeds of 300–400 ft/s (100–120 m/s). The thicker the myelin sheath the higher is the velocity of the nerve impulse traveling over the nerve. In cold-blooded vertebrates speed of conduction of a myelinated nerve of similar thickness is less than in warm-blooded vertebrates. In the myelinated fiber the impulse hops from one node of Ranvier to the next in what is called saltatory conduction. Satellite cells, variants of neurolemmal cells, envelop the cell bodies of ganglia associated with peripheral nerves. The white matter of the spinal cord and brain contains great numbers of fibers with a myelin sheath, although numerous unmyelinated fibers are found here. The gray matter is composed of unmyelinated fibers, dendrites, and cell bodies.

In the central nervous system, the oligodendroglia serve the same function as neurolemmal cells of the peripheral system; they envelop the cell bodies of neurons and elaborate the myelin sheath surrounding many axons. The astroglia are cells with processes extending between the blood capillaries and the cell bodies of the neurons. In effect they act as intermediaries conveying various products back and forth between the blood and the neurons. The oligodendroglia and the astroglia function to maintain a relatively constant chemical environment to enable the entire neuron to function efficiently.

Connective tissue cells. Nerve fibers of a peripheral nerve are bound together into small bundles by connective tissue cells, fibroblasts, and their fibrous products. The entire nerve is in turn surrounded by more connective tissues, within which are plexuses of blood vessels. Although the connective tissues are sparse within the central nervous system, the blood plexuses of the brain and spinal cord are most extensive. There are three layers of connective tissue membranes, the meninges, covering the brain and spinal cord: the inner, pia mater; the middle layer, the arachnoid; and the outermost, the dura mater. Between the pia mater and the arachnoid is the subarachnoid space; this space and the ventricular cavities within the brain are filled with an extracellular fluid, the cerebrospinal fluid. *See* MENINGES.

The microglial cells are the only parenchymal cells of the central nervous system that arise from embryonic mesoderm. These cells are similar to but smaller than macrophages which are derived from connective tissues. During stress conditions, such as inflammatory processes, infection, or traumatic injury to the brain substance, microglia become active and phagocytize and remove breakdown products from the brain and spinal cord.

Functional organization of a neuron. The neuron integrates and processes neural information. The axodendritic and axosomatic synapses stimulate receptor sites on cell membranes of dendrites and cell bodies; these are called the receptive segments. The integrative function of each segment depends, in part, upon the fact that the response to each stimulus from a synapse is graded; it is not an all-or-none nerve impulse. Only after the excitatory and inhibitory synaptic activity of the receptive segment has been resolved into an effective action does the conductile segment, the axon, generate and conduct a nerve impulse on an all-or-none action potential to the next synapse. The region of the axon associated with the synapse is known as the transmissive segment. It is at this site that the neurosecretion is released and transmitted across the synaptic cleft to stimulate the postsynaptic neuron. Charles Noback

Comparative Embryology

The complicated and varying anatomy of the adult nervous system in different vertebrates makes comparative embryological studies of these structures almost necessary for a sound understanding of their morphology. Few fields in experimental analytical embryology have proved so fruitful as that of neural development. A thorough study of the embryology of the structures under study in animals used for experiments is necessary for a causal analysis. The embryology may be divided into a gross morphogenetic part to analyze the development of the external and internal features of the nervous system, and a histogenetic part to deal with the differentiation of the cells of the nervous system and their arrangement into nuclei and cortical structures.

Nerve growth factor. Neuronal survival, as well as the performance by neurons of some specialized functions, is regulated by macromolecular agents known as neuronotrophic factors. The best studied of these agents, the nerve growth factor (NGF), was discovered in the late 1950s. Nerve growth factor is a protein present in many tissues and biological fluids, but particularly concentrated in some organs. In the organism, neurons might receive nerve growth factor from those cells they innervate, from glial cells with which they are associated, and perhaps also from the general circulation.

The two main cell groups sensitive to nerve growth factor, the dorsal root and the sympathetic ganglionic neurons, belong to the peripheral nervous system, although some central nervous system neurons may also be responsive. During some periods of their development, dorsal root and sympathetic neurons die if nerve growth factor is not present, or if they are treated with anti-NGF antibody. Beside this critical role for survival, nerve growth factor can also stimulate the neurons to produce neurites (axons and dendrites) as well as some of the enzymes necessary for the synthesis of neurotransmitters. In order to elicit these responses, nerve growth factor must first bind to receptors present on the cell surface. Very few of the molecular events linking this initial binding with the ultimate responses to nerve growth factor, such as neuronal survival or neuritic elongation, are already known.

There is additional evidence for other neuronotrophic factors which regulate survival and performance of different neuronal groups. The search for these factors has been stimulated by information on developmental neuronal death, a widespread phenomenon during development of the nervous system. Of all the neurons in any given population, only those that succeed in making and maintaining synaptic contacts with "target" cells survive. It is believed that each neuron receives from its target cells a neuronotrophic factor which regulates its survival. Neurons that fail to synapse with target cells do not receive the factor, and consequently die. One of the organs whose study has provided support for this hypothesis is the chick embryo ciliary ganglion. A ciliary neuronotrophic factor has been found which is capable of supporting the survival of the neurons from this ganglion. This factor is present in very large amounts in the embryonic eye structures normally innervated by ciliary neurons, at the time of embryonic life when the fate (death versus survival) of these neurons is decided. In addition to this and other neuron-directed survival-promoting

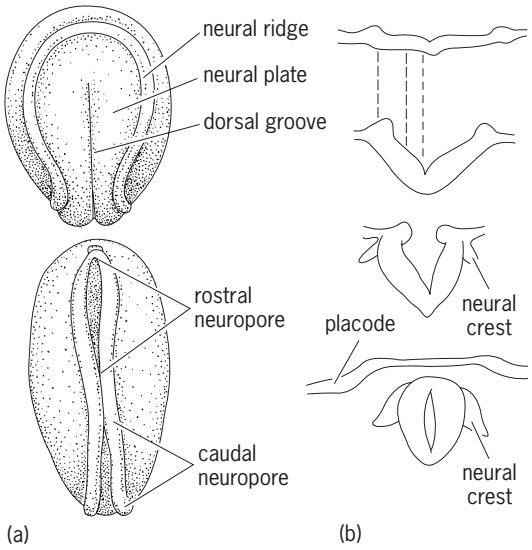

Fig. 7. Transformation of neural plate to a neural tube in amphibians (*a*) Dorsal views. (*b*) Transverse sections.

factors, other macromolecular agents are under investigation. Ruben Adler

Formation of neural plate and tube. The anlage of the nervous system is formed in the outer germ layer, the ectoderm, although some later contributions are also obtained from the middle germ layer, the mesoderm. In most vertebrates a neural plate is formed, which later folds into a neural groove, then closes to form a neural tube. In some vertebrate species, such as the lamprey and bony fishes, a massive cord of neural tissue is formed instead, which is later canalized into a neural tube.

In many vertebrates, including chickens and humans, the caudal-most end of the spinal cord is formed by a canalization of a neural cord within the so-called tail bud, while the main part arises as a neural plate.

The formation of neural tissue within the ectoderm is due to inductive influences from underlying chordomesodermal structures. *See* DEVELOPMENTAL BIOLOGY; EMBRYONIC INDUCTION.

The neural plate curls up at its lateral edges to form a tube (neurulation) in a process which varies somewhat in different vertebrates. The plate has an inherent tendency to form a tube by an active contraction of microfilaments in the apical parts of the cells (close to the future lumen). In some species at least, the surrounding mesoderm apparently partakes in the folding process, pushing the margins of the neural plate (neural folds) upward.

Closing of the neural tube in most species starts in the middle part of the embryonic body, the future neck region, and continues in a rostral and caudal direction. Transitorily a rostral and a caudal neuropore exist.

At the transition between the neural plate and the ectoderm, a thickening, the neural crest, is formed (**Fig. 7**). In the trunk, the ganglia of the spinal nerves are formed from it. In the head some cells from the neural crest enter mesodermal structures, and others take part in the formation of cranial ganglia. The latter are also formed from ectodermal thickenings lying farther laterally, the placodes. *See* NEURAL CREST.

Histologic differentiation. At the site of formation of the neural plate in the ectoderm, the ectodermal cells elongate and form a cylindrical epithelium, the neural epithelium. These cells continue mitotic division and form the primary germinal layer of the central nervous system. At a later stage of development cells migrate from the epithelium and form a peripheral layer (**Fig. 8**).

Within the peripheral layer, and sometimes already within the neural epithelium, the differentiation of the cells proceeds toward neurons and glia cells via neuroblasts and glioblasts, respectively. Also within the ganglia, formed from the neural crest and the placodes, a similar process of differentiation occurs.

From the surrounding mesenchyme (mesoderm) cells enter the central nervous system and form vessels and microglial cells. These cells divide mitotically within the brain substance even in older embryonic stages. Mitotic division of true neural cells outside the neural epithelium is probably of low frequency. In the cerebellum, however, there is a thick proliferating layer of neural epithelial cells which also proliferates peripherally, the so-called embryonic granular layer. Similarly, a proliferating layer exists outside the neural epithelium in the cerebral

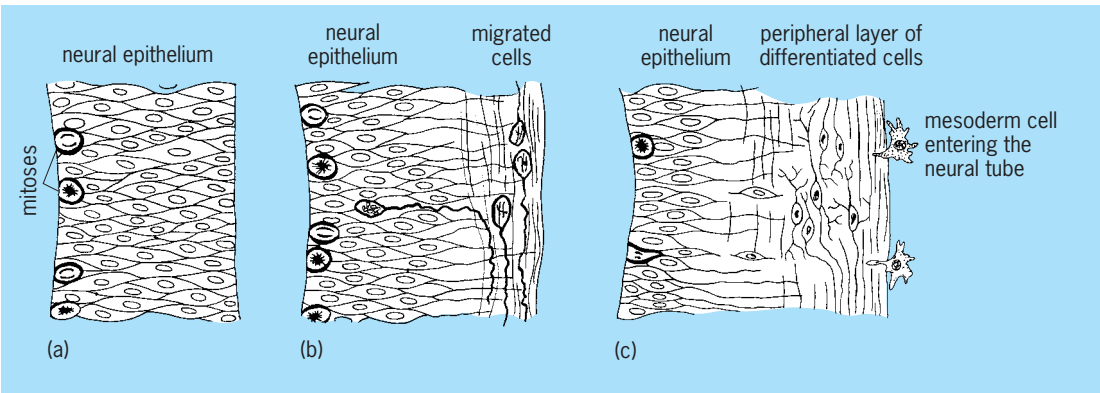

Fig. 8. Cellular constitution of neural tube in different stages *a*–*c*; ventricular wall at left.

hemispheres of some mammals, the subependymal layer.

After the formation of neurons and glia cells, the fibers which form give rise to the neuropile of the central nervous system, and to intra- and extracentral nerve bundles. The fibers emanating from the neurons of the ganglia grow as peripheral sensory fibers in peripheral nerves.

Morphogenesis. This aspect includes a consideration of the formation of neuromeres, the longitudinal structuring of the brain, and cell migration. *See* MORPHOGENESIS.

Neuromeres. When the neural tube is developing, a segmentation of the central nervous system occurs by the formation of transverse bulges, neuromeres (**Fig. 9**). They are most distinctly seen in the hindbrain region, but can usually be identified in suitable embryonic stages in all parts of the central nervous system. They are most easily seen in vertebrate brains having a thin wall, for instance, those of sharks, birds, reptiles, and mammals.

Three different sets of neuromeric bulges develop successively, called proneuromeres, neuromeres, and postneuromeres. They represent a primary, secondary, and tertiary segmentation, respectively. The basis of neuromerism is the presence of proliferative patterns. Each set of bulges thus corresponds to one period of increased proliferative activity in the neural epithelium, due to stimulative influences from underlying mesodermal structures. The proneuromeric segmentation extends from the neural tube into the neural crest and causes this to divide in the head region into portions, each corresponding to a proneuromere. This condition results in a topographic correspondence between the cranial ganglia and the neuromeres.

Fig. 9. Photomicrograph, dorsal view, of young chick embryo, showing neuromeres as bulges of the nervous system. (*From R. deHaan and H. Ursprung, eds., Organogenesis, Holt, Rinehart and Winston, 1965*)

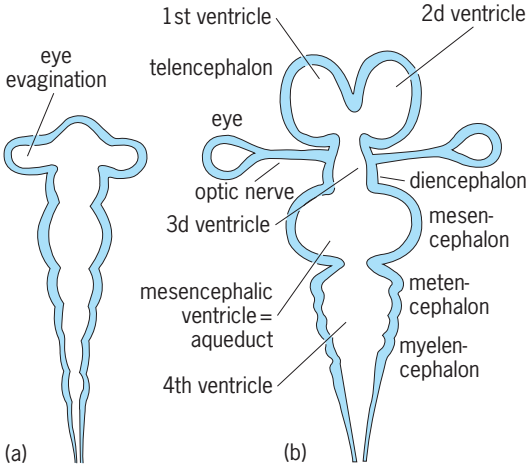

Fig. 10. Schematic horizontal sections through vertebrate brain, showing transformation of (*a*) neuromeric stage into (*b*) brain vesicle stage.

At the time of neuromeric segmentation, the brain is subdivided into the so-called brain vesicles by local widenings of its lumen. In the rostral end more or less well-developed hemispheres are formed; in the middle of the brain anlage the mesencephalic bulge develops; and behind the latter the walls of the tube thicken into cerebellar folds. In this way the brain anlage is divided into five sections: the telencephalon, diencephalon, mesencephalon, metencephalon, and myelencephalon, and its cavity is divided into the rudiments of the adult ventricles (**Fig. 10**). The brain vesicles make the segmental characters of the neuromeric bulges less conspicuous.

When the postneuromeres develop, bulges can be identified only in the brain, not in the spinal cord. The presence of the postneuromeres influences the early development of the internal structures, giving them a slight segmental character.

Longitudinal columns. When the postneuromeric phase is at its height, a longitudinal structuring of the brain wall develops, consisting of four longitudinal bands of high proliferative and migrative activity. In the hindbrain these four columns approximately correspond to the anlagen of the four columns of functionally different qualities in the adult brain. In the spinal cord, the two dorsal columns fuse into one and the two ventral columns into another. Rostrally, in the brain, the ventralmost columns stop at the rostral end of the mesencephalon and the dorsalmost columns at the transition zone between the mesencephalon and the metencephalon (the isthmic region). The two middle columns build up the rest of the brain, and the borderline between them ends at the optic chiasma. An approximate borderline between the two middle columns in the spinal cord and the myelencephalon is found in a furrow, the so-called limiting furrow of His. It cannot be identified with certainty in the rostral part of the brain.

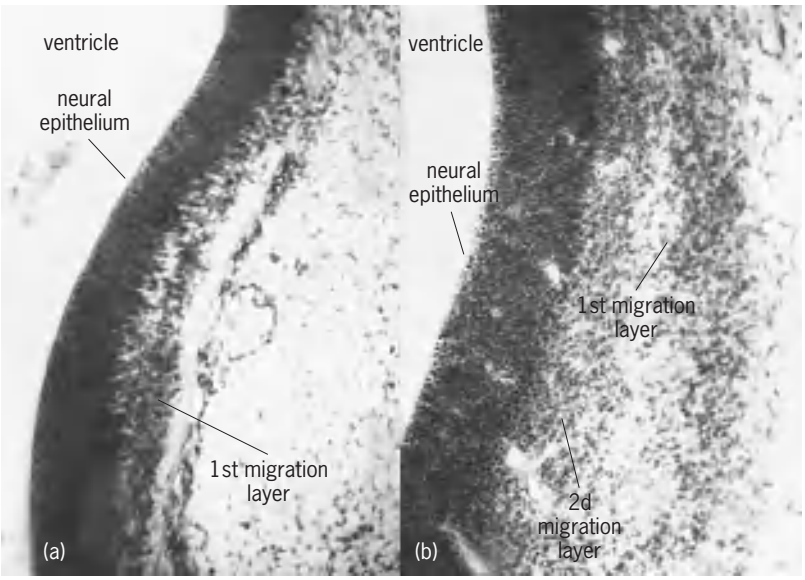

Fig. 11. Two sections through a chick embryonic hemisphere, showing the neural epithelium and the cells migrating from it. (*a*) One thin migration layer exists. (*b*) Two migration layers lie outside each other.

The postneuromery and the longitudinal banding will give rise to a checkered pattern of proliferation centers. In the rostral part of the brain the transverse pattern will dominate, the longitudinal one in the caudal part.

Cell migration. Cell migration takes place from the neural epithelium into the peripheral or mantle layer. The presence of transverse and longitudinal proliferation centers will give rise to certain areas which are rich in cells, and will cause a vivid lateral migration. Such areas are called migration areas, and their topography will be determined by postneuromery and longitudinal banding. The number and topography of the migration areas will be very similar in all vertebrate species.

The cells, which have migrated laterally, may still lie in close contact with the neural epithelium and the ventricular wall, as in amphibians, or may lose contact with the epithelium and lie as a peripheral layer (**Fig. 11**). In many species, especially higher vertebrates, successive migrations of cells occur, giving rise to two or more layers of such cells, situated concentrically. This feature is especially well marked in the cerebral hemispheres.

Later in development, cells can migrate right through the primarily migrated cell layers—for example, the most external cortical layers can be formed later than more centrally situated cell masses.

The migration layers may fuse or further subdivide into cell clusters, which represent the anlagen of the future brain nuclei. Therefore, they furnish the basis for comparative studies and homologizations of brain nuclei of different vertebrates.

Whole cell groups or brain nuclei may migrate (group migration), and in this way the topography of the nuclei may shift even from one brain vesicle to another. Long-distance migration of scattered neural cells are also known, for example, along the surface of human hemispheres from the basal brain up to the convexities.

Cell death. Within the central nervous system cell death occurs embryonically as part of normal development. Some regions degenerate and are eliminated. This occurs, for instance, below the primary optic stalks, which in this way are shifted to their definite ventral position on the brain. The uneven thickness of the spinal cord motor cell column is also obtained by localized degeneration of neuroblasts. Furthermore, scattered cells die within the nervous system, most markedly in connection with rapid cellular differentiation. Probably, such dying cells are defective and incapable of differentiating. *See* CELL SENESCENCE.

Orientation of outgrowing neurites. Maturing neuroblasts send out neurites that build up fascicles and tracts connecting various neural centers.

The growth of neurites is guided by mechanical factors to some extent; chemical or electrical forces play no role, at least at a distance. The actual establishment of contacts between growing neurites and their correct target cells may be determined by cytoaffinity, that is, similarities in the properties of the cell surface.

Brain. In spite of the extraordinary variation in adult morphology of the vertebrate brain in different species, the early phases of development are essentially similar. The brain vesicle stages of a reptile, bird, and mammal are much alike, but owing to varying growth rates of different parts and to specialization processes the different patterns of the adult brains are formed (**Fig. 12**).

In a comparison of the embryology of the brains of anamniotes and those of amniotes a marked difference is seen in the so-called brain flexures. The originally straight brain tube is bent during development. In a shark or amphibian embryo the only

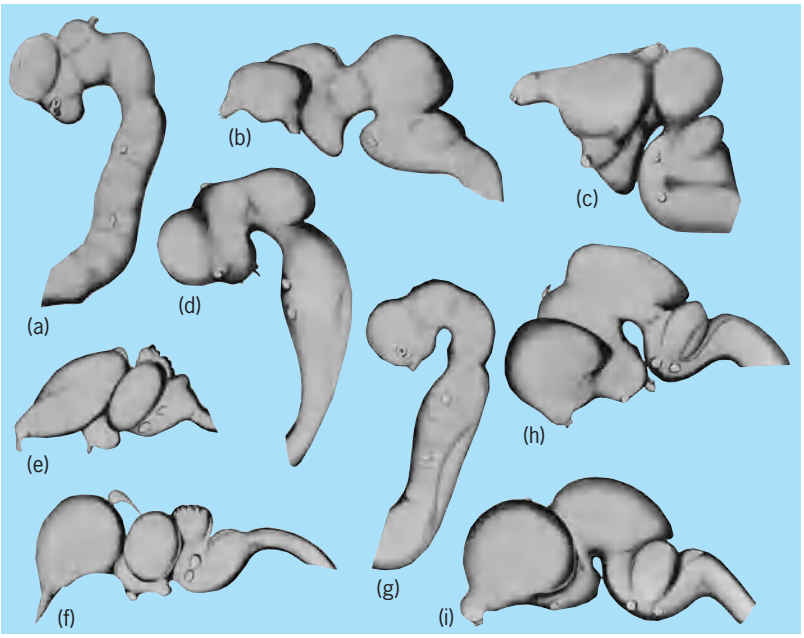

Fig. 12. Lateral views of embryonic brain at progressive stages of development. The early phases are essentially similar. (*a–c*) Reptile (*Chelydra*). (*d–f*) Bird (*Melopsittacus*). (*g–i*) Mammal (*Spermophilus*).

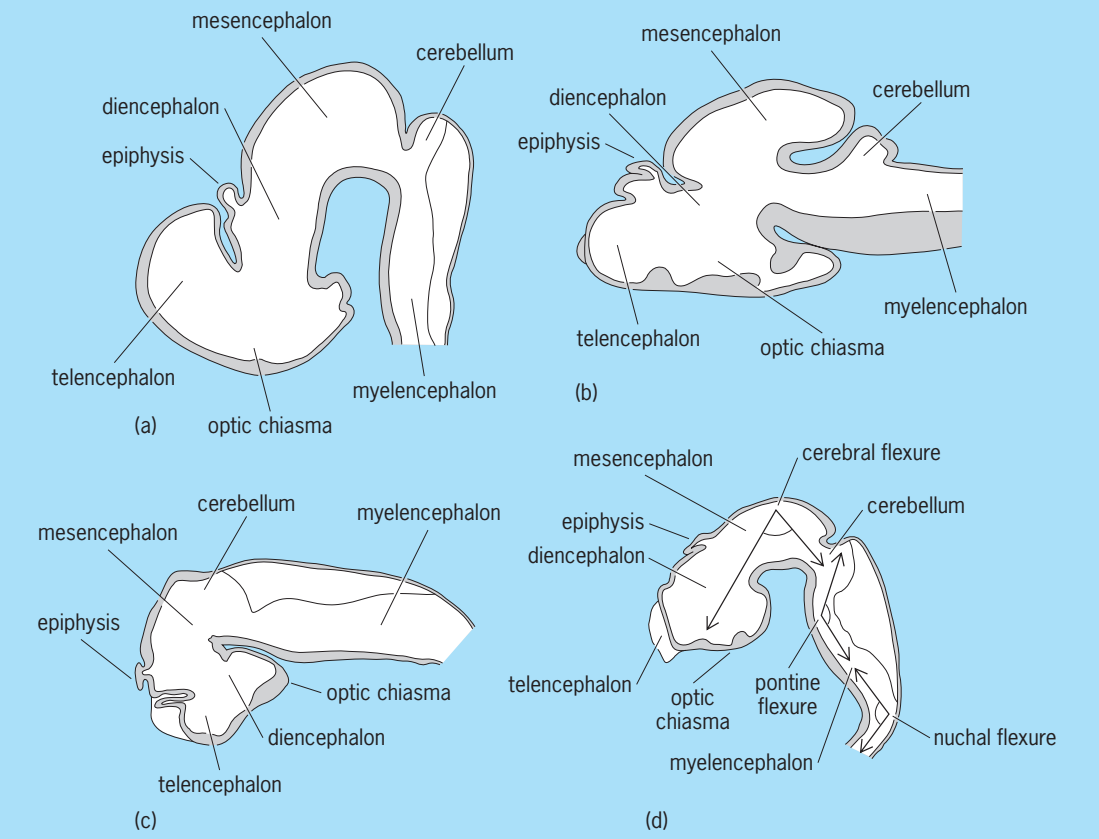

Fig. 13. Median sections of the brains of embryos. (*a*) Shark. (*b*) Bony fish. (*c*) Salamander. (*d*) Reptile.

marked bending is the cephalic flexure, situated in the same plane as the mesencephalon. This is also the first to develop in amniotes. In these brains, however, a nuchal flexure is also formed at the transition between brain and spinal cord anlagen, and later a pontine flexure arises ventral to the cerebellar region (**Fig. 13**). The flexures are most easily seen in median sections.

Telencephalon. The morphogenesis of the telencephalon of most vertebrates occurs by a lateral evagination or outbulging of the wall, giving rise to two hemispheric vesicles. In bony fishes, ganoids, and holocephalians, however, the lateral evagination is only faintly marked. Instead, a lateral bending, or eversion, occurs. The topography of the internal structures in the two kinds of forebrain will therefore be different (**Fig. 14**).

In all vertebrates two migration areas develop in the telencephalon—a dorsal one representing the embryonic origin of the pallium, and a ventral one representing the subpallium. Each of these areas is further subdivided into cell columns from which the different mantle regions and the septal and striatal nuclei develop.

Diencephalon. The morphogenesis of the diencephalon varies little in different species. A more or less well-developed transverse velum is formed in the roof, caudal to which the epiphyseal rudiment is situated. The paraphysis, which is included in the telencephalon, lies rostral to it. Part of the hypophysis develops from the bottom of the diencephalon while from the ventrolateral parts of the diencephalon the eyes are formed. The lateral walls are divided into a dorsal thalamic and a ventral hypothalamic region, containing the mammillary bodies. The hypothalamic region grows more in size in lower vertebrates than the thalamic does, while in higher vertebrates the opposite condition exists.

Mesencephalon. The original single mesencephalic vesicle is divided into two vesicles which communicate broadly with each other. In lower vertebrates this condition remains unchanged. The original wide ventricular cavity in higher forms is reduced to form the mammalian Sylvian aqueduct. The evaginations are connected ventrally with an unevaginated part, the tegmentum. Within the latter, the tegmental

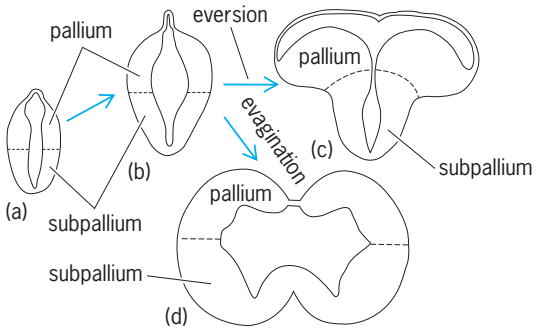

Fig. 14. Schemes showing transverse sections of forebrains. (*a*) Primitive stage develops via (*b*) interstage to (*c*) eversion or (*d*) evagination.

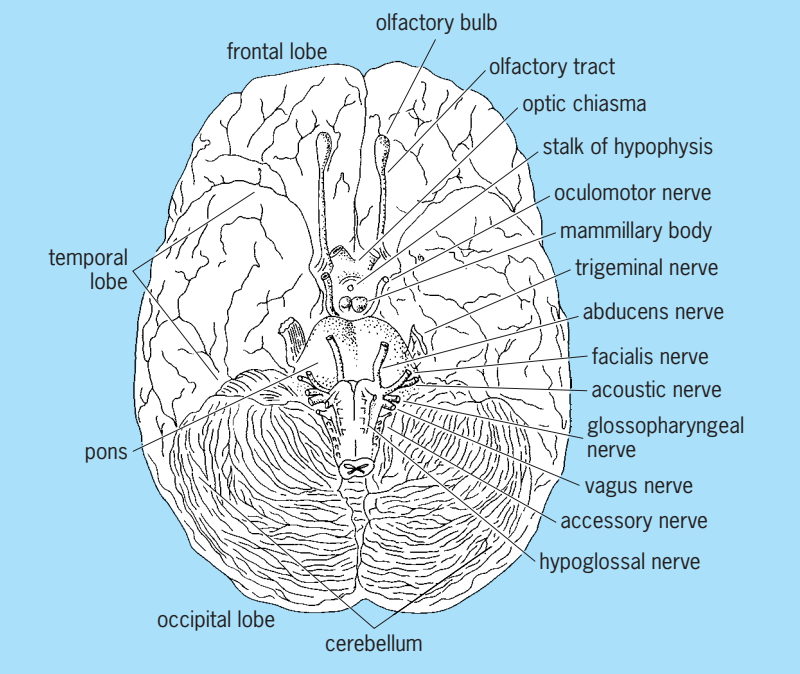

Fig. 15. Drawing of human brain seen from below.

nuclei, oculomotor nuclei, and the red and black nuclei (nucleus ruber and nucleus niger) develop. The mesencephalic evaginations form the bigeminal bodies in lower forms and the quadrigeminal bodies in higher forms.

Metencephalon. The cerebellum is formed in the dorsal part of the metencephalon. Its degree of development in different vertebrates varies considerably. The original raised lateral walls of the brain fuse to form a single plate. This is extremely compact, for example, in the bony fishes, and from its rostral end a so-called valvula grows rostrally. In *Petromyzon*, amphibians, and most reptiles the cerebellum remains as a simple transverse plate. In sharks, birds, and mammals it develops into a dome, which may be more or less folded, thereby increasing its surface.

In most vertebrates a secondary proliferative layer, the embryonic external granular layer, is formed in the periphery of the cerebellum; this layer disappears during later development.

Myelencephalon. This brain part remains relatively primitive. Its roof is extended as a thin tela; its walls form a more or less V-shaped structure with only small variations. The internal structures are dominated in their development by the above-mentioned longitudinal columns.

Spinal cord. The spinal cord remains as a comparatively slightly differentiated tube. The primary lumen is secondarily reduced by the fusion of the side walls into a narrow central canal. In the lateral walls, the longitudinal columns, separated by the limiting furrow of His, develop into the dorsal and ventral horns respectively. In fishes the diameter of the spinal cord tube gradually diminishes in a rostrocaudal direction, but in four-footed animals intumescents develop level with the extremities by a process of partial degeneration of the regions situated in between.

Cranial nerves. The cranial or cerebral nerves are the peripheral nerves of the head that are related to the brain. The number and degree of development of the nerves varies in different species. The functional quality of the different nerves also varies. Twelve pairs of cranial nerves have been distinguished in human anatomy (**Fig. 15**) and these nerves have been numbered rostrally to caudally as follows:

I. Olfactory nerve, fila olfactoria
II. Optic nerve, fasciculus opticus
III. Oculomotor nerve
IV. Trochlear nerve
V. Trigeminal nerve, in most vertebrates divided into three branches: ophthalmic, maxillary, and mandibular
VI. Abducens nerve
VII. Facial nerve
VIII. Statoacoustic nerve
IX. Glossopharyngeal nerve
X. Vagus nerve
XI. Accessory nerve
XII. Hypoglossal nerve

The varied morphological significance of the cranial nerves is evident from their embryology.

Olfactory nerve. Fibers of the olfactory nerve grow out from the primary sensory cells of the epithelium of the nasal sac. They run to the lateral surface of the telencephalic rudiment, usually entering it on the border between the pallial and subpallial regions.

Optic nerve. The eyes develop as evaginations from the lateral walls of the diencephalon. The stalks of the evaginations are the pathways of the future optic nerves. The neurites growing in from the retina within the stalk are comparable to an intracerebral fascicle. They reach the brain in the floor of the diencephalon to form the optic chiasma. *See* EYE (VERTEBRATE).

Ventral motor nerves. Neurites emerge from cells situated in the ventralmost longitudinal column of the brain stem and leave the brain surface as motor nerves. These nerves are the oculomotor, trochlear, abducens, and hypoglossal. The trochlear nerve fibers first grow dorsad, cross in the roof of the brain, and leave it dorsally in the fold between the mesencephalon and the cerebellum. The other nerves leave the brain ventrally. In *Petromyzon* the trochlear nerve nucleus develops dorsally at the site of the future crossing.

Dorsal nerves. The dorsal nerves (trigeminal, facial, statoacoustic, glossopharyngeal, vagus, and accessory) are all mixed nerves except the statoacoustic. The sensory fibers grow out from neurons differentiated within the cranial ganglia. The motor fibers come from cells lying within the brain stem.

The cranial ganglia are formed from the head portions of the neural crest and from the ectodermal placodes. The neural crest is divided into four or five segments, called the thalamic (present only in lower vertebrates), mesencephalic, trigeminal, facial, and

glossopharyngeal-vagus crests.

The placodes of lower vertebrates are made up of two groups, those associated with the lateral-line nerve system and called the dorsolateral placodes, and those situated further ventrally and giving rise to the main ganglia, the ventral placodes. A summary of the vertebrate placodes is given below.

Spinal nerves. The spinal ganglia are formed from the neural crest which grows out like a continuous sheet from the dorsal margin of the neural tube and is secondarily split up into cell groups, the ganglia, by a segmentating influence from the somites. Fibers grow out from the ganglionic cells and form the sensory fibers of the spinal nerves. Motor nerve fibers emerge from cells situated in the ventral horns of the spinal cord. The ventral motor fibers and the dorsal sensory fibers fuse to form a common stem, which is again laterally divided into branches, innervating the corresponding segment of the body.

Autonomic nervous system. The ganglia of the sympathetic nervous system develop ventrolateral to the spinal cord as neural crest derivatives. At first a continual column of sympathetic nerve cells is formed; it later subdivides into segmental ganglia. The nerve fibers developing from these cells form the gray communicants to the spinal cord and the peripheral sympathetic nerves. The white communicants develop from spinal cord cells. Along the peripheral nerve fibers, cells migrate to form the secondary plexi and ganglia.

The parasympathetic system is made up of preganglionic fibers emanating as general visceromotor fibers from the brain and from the sacral cord segments. Cells migrate to form the peripheral ganglia along them. *See* AUTONOMIC NERVOUS SYSTEM.

Bengt Kallen

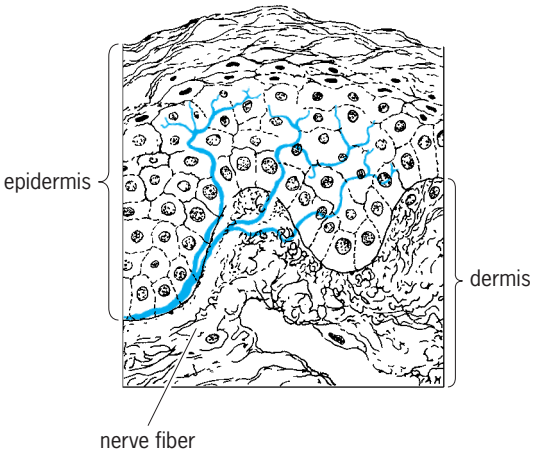

Fig. 16. Schematic of human sensory nerve fiber passing through the dermis and terminating in the free nerve endings among epithelial cells in the skin epidermis.

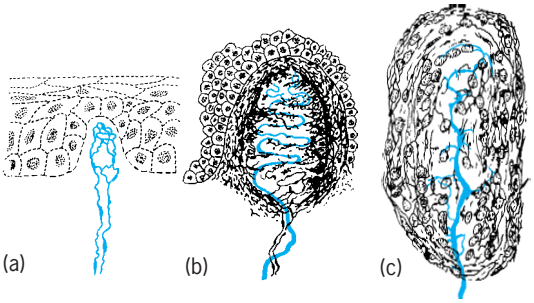

Fig. 17. Human free nerve endings. (*a*) Nerve loops developing Meissner corpuscle from 7-month human fetus. (*b*) Adult Meissner corpuscle. (*c*) Pacinian corpuscle in human fetus at 4 months. (*After B. M. Patten, Human Embryology, 3d ed., Blakiston-McGraw-Hill, 1968*)

Embryology of Sense Organs

Groups of ganglion cells, connected with the brain and spinal cord, send tiny nerve fibers through cablelike nerves to various parts of the body where they pick up many kinds of sensations which keep the living organism in touch with its environment. Therefore, specialized receptor cells and nerve endings must be provided, especially over wide areas for such senses as touch, pressure, pain, temperature, and muscle and tendon sense. Wherever possible in this discussion, the description of development of special senses in vertebrates is illustrated with human examples. *See* SENSE ORGAN.

Free nerve endings. Free nerve endings for pain and touch reach the skin as early as the third month in human fetuses. Their terminal branches then increase as the skin rapidly develops hair and nails from the fourth to sixth months (**Fig. 16**). During this time certain terminal nerve fibers slowly become encased with specialized layers of flat cells. Some near the skin, the Meissner corpuscles, receive tactile stimuli. Nerve loops (**Fig. 17***a*) near the skin gradually become encapsulated with specialized connective tissue cells (Fig. 17*b*). Others, Pacinian corpuscles (Fig. 17*c*), receive deep pressure sense and consist of more elaborate concentric cell layers, like sheaths of an onion, wrapped around a tiny nerve fiber. They develop in much the same way as tactile organs.

Continuing from the third fetal month many sensory nerve fibers spread over the body among developing muscle and tendon fibers, and as they branch, tiny flat plates develop at each nerve ending (**Fig. 18**). A delicate fibrous network of connective tissue finally covers them. The stimuli they pick up and relay to the central nervous system give the awareness of the position of the body and its parts.

Lateral-line organs. Some organs of special sense, such as the eye, are extremely complicated, whereas

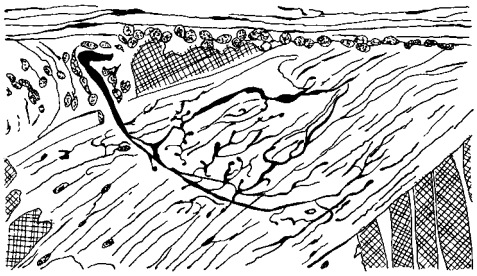

Fig. 18. Developing neurotendinous fibers from human fetus of 6 months. (*After B. M. Patten, Human Embryology, 3d ed., Blakiston-McGraw-Hill, 1968*)

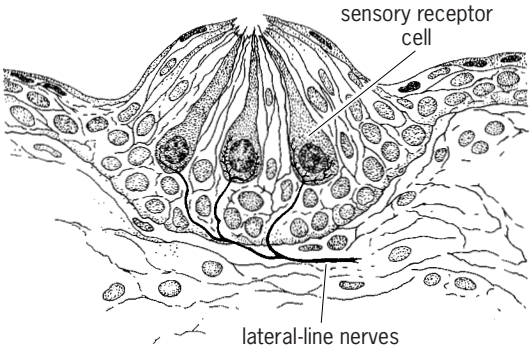

Fig. 19. Schematic drawing of a lateral-line sense organ in skin of adult salamander, the common aquatic vermilion spotted newt. Sensory receptor cells, surrounded by supporting cells and skin epithelium, terminate in hairs projecting into skin pores. Lateral-line nerves terminate around bases of the sensory cells.

Fig. 20. Lateral-line development in salamander (*Ambystoma punctatum*). (*a*) Camera-lucida drawing of living embryo made 1 day after a lateral-line placode was excised and replaced by a similar one (shaded) taken from a Nile blue–stained donor. (*b*) Same living specimen 24 h later, showing one midbody lateral-line primordium migrating in the surface ectoderm toward the tail. (*c*) Same specimen 24 h later than *b*, showing a long and a shorter lateral-line primordium migrating down side of the body and depositing clusters of blue-stained cells that form the sense organs. (*From L. S. Stone, The development of lateral-line sense organs in amphibians observed in living and vital-stained preparations, J. Comp. Neurol., 57(3):507–540, 1933*)

others are relatively simple. In some aquatic vertebrates (many fishes and amphibians) there are lateral-line skin organs (**Fig. 19**) on the head and body, innervated by nerve trunks coming from cranial ganglion cells connected with the brain. These organs apparently acquaint the animal with pressure changes in the surrounding water giving it a sense of orientation while swimming in a current or a warning of an approaching object. There are no homologs in humans or other animals. These tiny pear-shaped organs possess several centrally placed, club-shaped sensory cells (Fig. 19), each of which ends in a hair-like process at the free surface. These cells are interspersed and surrounded by tall, flat, overlapping, leaflike supporting cells. Fine nerve fibers from the lateral-line nerve (Fig. 19) branch among the sensory cells to receive their stimuli. The apex of the organ communicates with a microscopic pore at the skin surface in amphibians, and with a canal system in the skin of fishes.

In amphibian embryos where they have been extensively studied, ectodermal thickenings, called placodes, first appear on the side of the head. Any one of these placodes can be stained with a blue vital dye (**Fig. 20***a*), and as the embryo grows one can follow them as they elongate, migrate on the surface of the head and body (Fig. 20*b* and *c*), and deposit at regular intervals clusters of blue cells that form the lateral-line organs. By this method one can observe the developing organs under the microscope as the blue-dye particles migrate to the tips of the sensory and supporting cells (**Fig. 21**). Each of them becomes innervated by the lateral-line nerve that follows the migrating placode. This nerve comes from ganglion cells which are also placodal in origin. A cluster of new secondary organs arises by a budding process from supporting cells of the primary organs (**Fig. 22***a* and *b*). In practically all frogs and toads the lateral-line system degenerates at metamorphosis.

Taste buds. There are special chemical receptors somewhat like a rosebud in shape, called taste buds (**Fig. 23***a*), or gustatory organs. They are common to all classes of vertebrates and function in a watery environment. They are associated with parts in the oral cavity, especially on the fungiform and circumvalate papillae in the mammalian tongue, but in some fishes, such as the catfish (*Ameiurus*), many taste

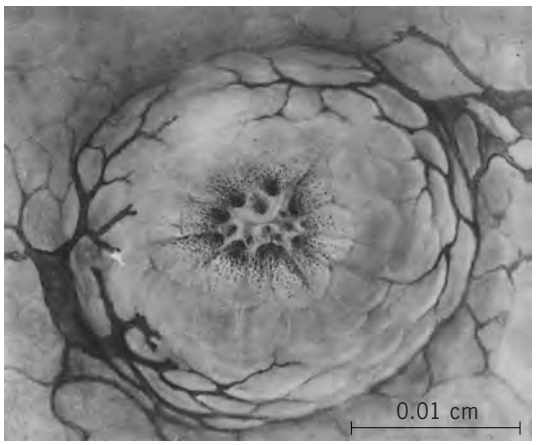

Fig. 21. Camera-lucida drawing of a differentiated living lateral-line organ surrounded by two large pigment cells in skin of young 16.5-mm (0.65 in.) salamander larva, 16 days after operation shown in Fig. 20*a*. Blue-dye particles were observed during development as they migrated to tips of the central sensory and surrounding supporting cells. (*From L. S. Stone, The development of lateral-line sense organs in amphibians observed in living and vital-stained preparations, J. Comp. Neurol., 57(3):507–540, 1933*)

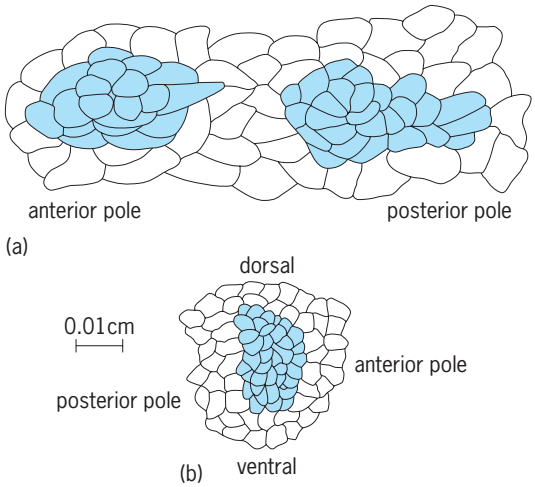

Fig. 22. Camera-lucida drawings of *Ambystoma punctatum* outlining lateral-line organs (shaded cells) in process of budding. (*a*) Budding organ at posterior pole was derived 24 h earlier from organ at anterior pole in tail of larva. (*b*) Lateral-line organ budding dorsally was observed to be derived by budding 24 h earlier from the one ventral to it. None of the organs were observed budding anteriorly or posteriorly. (*After L. S. Stone, The development of lateral-line sense organs in amphibians observed in living and vital-stained preparation, J. Comp. Neurol., 57(3):507–540, 1933*)

buds are also found in the skin on the surface of the head and body. The central, rod-shaped sensory cells of mammals (Fig. 23*a*), neuromasts, are embraced by slender, overlapping, flat, supporting cells, the outer ends of which surround a pitlike excavation connected through a pore with the mucous epithelium of the mouth. These neuromasts, which send hairlike processes into the pit, are in contact with a basketlike network of nerve fibers. They pick up the stimuli that are then carried by the nerve fibers to the cranial gustatory ganglia and on into the brain. *See* TONGUE.

In the tongue of the human fetus the taste buds first appear as clusters of epithelial cells and increase in number as the gustatory nerve fibers reach the epithelium. Although the taste organs are known to degenerate eventually after gustatory nerves are cut, their arrival at the epithelium in the first place may not be the stimulus which induces the organs to form.

It has been shown conclusively by experiments on salamander embryos that the lining of the floor of the future mouth can be transplanted from one embryo to the side of the body of another embryo; a tongue with taste organs develops later without having been innervated. It was also found that if the epibranchial ectodermal placodes on the sides of the head, which give rise to the gustatory ganglia, are excised, the taste organs develop normally without a nerve supply. How these special sense organs arise in any vertebrate is not known. Taste organs, like lateral-line organs, were found to increase in number by a continuous budding process from the peripheral supporting cells of older taste organs. In many vertebrates there is a continuous increase in taste buds for a long period. It is quite possible that this is accomplished by a similar budding process. Very little is known about the time at which the taste organs become functional. Some investigators believe that significant reflex responses can be induced in premature 7-month infants by sweet, sour, and bitter tastes.

Olfactory structures. In humans, the sense of smell also depends on special neurosensory epithelial cells functioning in a moist environment within the nasal cavities. The area of specialized olfactory epithelium lies in the upper deeper roof of the nasal mucous membrane and is made up of tall cells with bristle-like processes projecting into the mucus-covered surface where they act as chemical receptors (Fig. 23*b*). They are surrounded by tall supporting cells and extend toward the brain as thin fibers which contact fibers of intermediate ganglion or mitral cells. These in turn relay the olfactory impulses along the olfactory tract to the appropriate centers in the brain.

The olfactory organs arise in a similar manner in all vertebrates, by an early appearance of a pair of surface ectodermal thickenings, nasal placodes (**Fig. 24**), at the front end of the head. Considerable evidence from experiments on amphibian embryos indicates that the formation of nasal placodes can be induced by neighboring mesoderm and brain-forming cells.

In human embryos the nasal placodes appear during the fourth week. Very soon the placodes sink inward, forming pits which become deeper as the frame of the nose and surrounding structures of

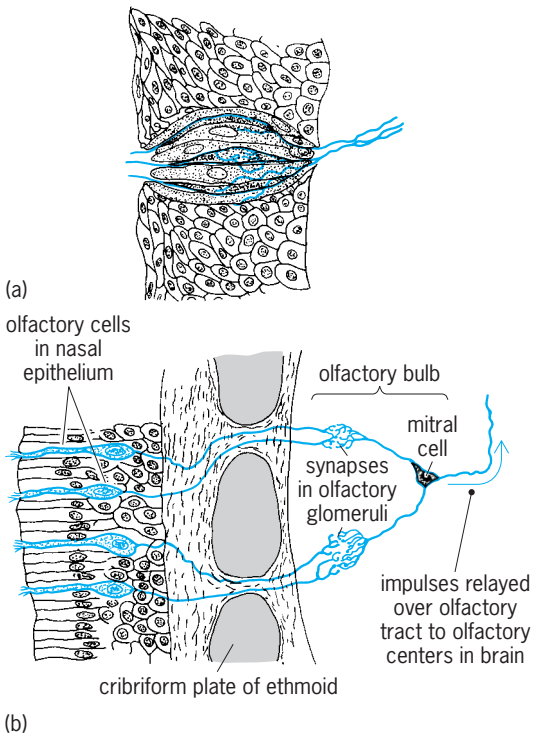

Fig. 23. Mammalian sense organs. (*a*) Taste bud in lingual epithelium. (*b*) Olfactory sensory cells of the nose related to nerve tracts leading to the brain. (*After B. M. Patten, Human Embryology, 3d ed., Blakiston–McGraw-Hill, 1968*)

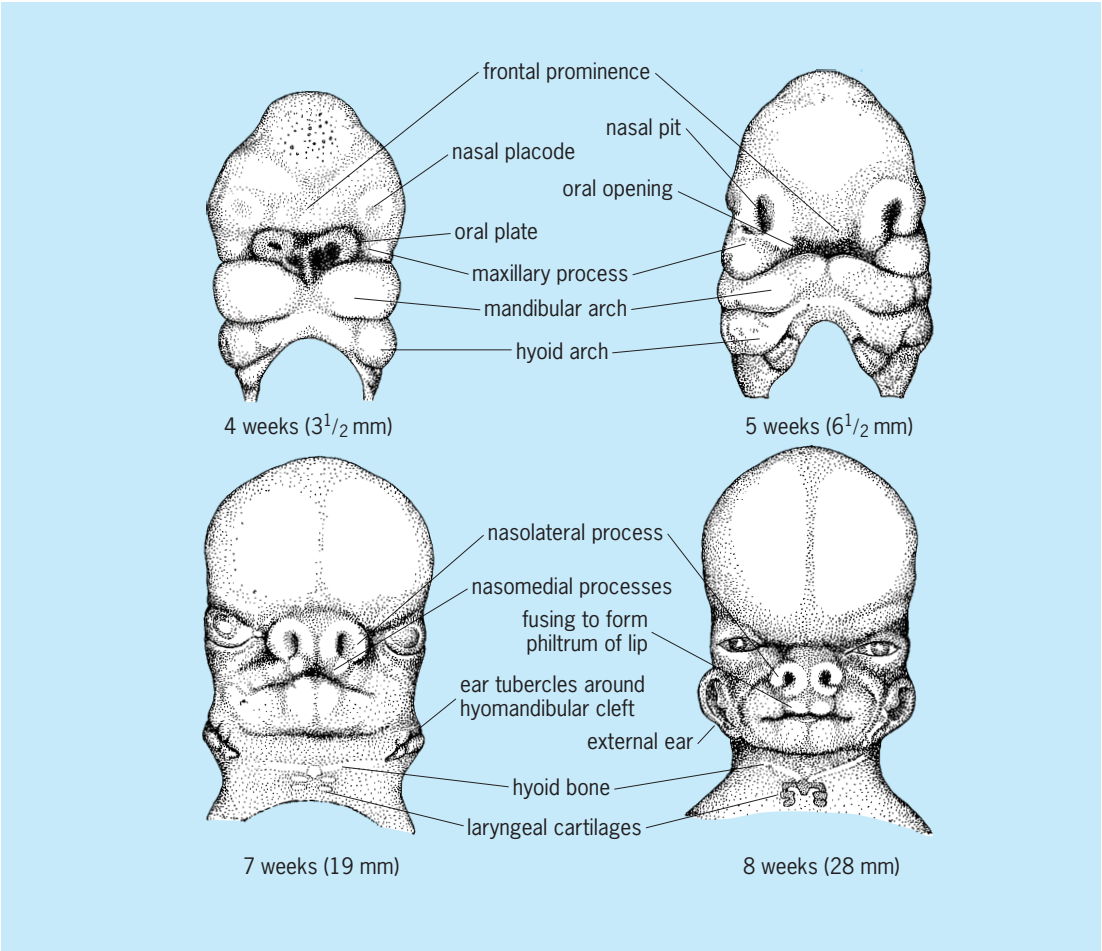

Fig. 24. Development of nose and other facial features in human embryos. (*After B. M. Patten, Human Embryology, 3d ed., Blakiston–McGraw-Hill, 1968*)

the face rapidly develop. The nasal cavities extend deeply and downward toward the oral cavity, with which they communicate shortly after the seventh week. The forward growth of the palate, nose, upper lip, and median nasal septum aids in formation of the nasal passages during the second month. By this time in the roof of the two nasal passages, specialized sensory cells of the olfactory epithelium are surrounded by tall supporting cells.

Except for the skinlike lining at the entrance of the nares, all other areas of the nasal cavities become covered by columnar epithelium with surface cilia and mucus-secreting cells. These cells keep the entire membrane covered with a moist film that provides the environment later for chemical stimulus of the hairlike ends of the sensory cells. The rate at which full differentiation of this sensory mechanism takes place varies among the vertebrates. The normality of the framework of the nose as well as the face and head depends a great deal upon the ability of the mesoderm to reach its full development.

Leon S. Stone

Bibliography. M. L. Barr, *Human Nervous System*, 7th ed., 1998; A. H. Cohen et al., *Neural Control of Rhythmic Movements in Vertebrates*, 1988; P. Nathan, *The Nervous System,* 4th ed., 1997; C. R. Noback, N. L. Straminger, and R. J. Demasest, *The Human Nervous System: Introduction and Review*, 6th ed., 1991; G. Siegel et al. (eds.), *Basic Neurochemistry: Molecular, Cellular, and Medical Aspects*, 6th ed., 1998; M. Taussig, *The Nervous System*, 1991; G. C. Wild and E. C. Benzel, *Essentials of Neurochemistry*, 1994.

Nervous system disorders

A satisfactory classification of diseases of the nervous system should include not only the type of reaction (congenital malformation, infection, trauma, neoplasm, vascular diseases, and degenerative, metabolic, toxic, or deficiency states) but also the site of involvement (meninges, peripheral nerves or gray or white matter of the spinal cord, brainstem, cerebellum, and cerebrum). To these may be added various other correlates, such as age and sex; however, this article will consider mostly the various types of reactions and only incidentally the sites affected.

In general, nerve cells and their processes are supported by interstitial cells and nourished by blood vessels. In the central nervous system the interstitial

cells are called neuroglia (astrocytes or oligodendrocytes), and in the peripheral nervous system they are known as Schwann cells. Interstitial cells surrounding nerve cell bodies, either centrally or peripherally, are called satellite cells; those surrounding fibers form myelin, an insulating element which is composed of many wrappings of thin cell membranes which are slightly but significantly different chemically in the central nervous system as compared to the peripheral nervous system.

The nerve cell may be damaged primarily, as in certain infections, such as poliomyelitis (polio) or herpes zoster (shingles); but much more commonly the nerve cell is damaged secondarily as the result of metabolic or vascular diseases affecting other important organs, such as the heart, lungs, liver, and kidneys. This reflects the great degree of specialization of nerve cells, which have become almost totally dependent upon the metabolic functions of other organs.

Clinical signs and symptoms of dysfunction of the nervous system depend not only on the type and location of lesions but also on the severity and rapidity of development of the lesion. The nervous system grows most rapidly in infancy, slowly degenerates with advancing age, and has definite but limited regenerative capacity. Since cells mature at different rates, and since exposures to infections and poisons naturally vary with age, many diseases and their clinical signs and symptoms become age-dependent.

Malformation. The central nervous system develops as a hollow neural tube by the fusion of the crests of the neural groove, beginning in the cervical area and progressing rostrally and caudally, the last points to close being termed the anterior and posterior neuropores. If the anterior neuropore fails to close (about 24 days of fetal age), anencephaly develops. The poorly organized brain is exposed to amniotic fluid and becomes necrotic and hemorrhagic, leaving a froglike head on an infant who usually dies within hours after birth. The eyes, brainstem, cerebellum, and spinal cord are usually normal. *See* CONGENITAL ANOMALIES.

If the posterior neuropore fails to close (about 26 days of fetal age), the lumbosacral neural groove is exposed to amniotic fluid. Without the protective coverings of meninges, spines, muscles, and skin, the nervous tissue becomes partially necrotic and incorporated in a scar composed of disorganized meninges, muscles, and skin. Such a meningomyelocele is readily infected unless buried surgically within a few hours after birth. In addition, in about 95% of such infants hydrocephalus (a dilatation of the ventricles) occurs. The hydrocephalus can usually be adequately treated by shunting the ventricular fluid into the venous system or peritoneal cavity.

Other developmental disorders of the nervous system may appear as a maldevelopment (hypoplasia or hyperplasia; that is, decrease or increase in growth of cells) or as a destruction of otherwise normally developing tissues. Rapidly growing tissues such as the embryonic nervous system are generally rather easily damaged by many toxic agents which seldom affect adult tissues; the surviving cells may grow in a less well organized pattern due to the absence of organizing influences exerted by antecedent or adjacent cells. The time of onset and the extent of repair, rather than the nature of the agent (infection, ionizing radiation, vitamin deficiency or excess, toxin, anoxia, and so on) determine the resulting pattern of abnormal development. *See* BEHAVIORAL TOXICOLOGY.

Infection. Infections of the nervous system may occur through a defect in the normal protective coverings caused by certain congenital malformations, as mentioned above, but also through other defects as the result of trauma, especially penetrating wounds or fractures sing into the paranasal sinuses or mastoid air cells. Subsequent infection of the nervous system may be the major complication of such "open head" injuries.

Pyogenic bacteria. Infections may also spread directly from adjacent structures, as from mastoiditis, sinusitis, osteomyelitis, or subcutaneous abscesses. Such infections usually spread along venous channels with progressing thrombophlebitis, producing epidural abscess, subdural empyema, leptomeningitis, and brain abscess. All of these infections are characteristically caused by pyogenic (pus-forming) bacteria (staphylococci, streptococci, pneumococci, and such), which may be normal or transient residents on the skin.

Other pyogenic bacteria (such as *Streptococcus viridans*, *Escherichia coli*, and *Hemophilus influenzae*) may metastasize by way of the bloodstream from more distant infections, such as bacterial endocarditis, pneumonia, and enteritis. The meningococcus may gain access to the blood from the nasopharynx, where it frequently resides asymptomatically, and its endotoxin may produce shock and purpura (bleeding beneath the skin) and hemorrhage into the adrenal glands. In conditions of overcrowding, especially when many previously unrelated persons suddenly come together, as in military and emergency camps, meningococcal meningitis may occur in epidemics.

Infections of the nervous system must be treated promptly as medical emergencies since the mortality rate increases with delay. The diagnosis is easily established by spinal puncture, the cerebrospinal fluid appearing cloudy and containing hundreds or thousands of polymorphonuclear leukocytes per cubic millimeter. The microorganisms can be visualized with special stains.

Other microorganisms. Many other microorganisms can infect the nervous system: *Mycobacterium tuberculosis* (the organism causing tuberculosis), *Treponema pallidum* (the organism causing syphilis), several fungi and rickettsiae, and many viruses. These characteristically produce a nonpurulent leptomeningitis, the cerebrospinal fluid being clear or only slightly cloudy because the number of leukocytes is much less, usually not more than a few hundred per cubic millimeter; also, instead of polymorphonuclear leukocytes, lymphocytes predominate. Increasingly effective antibiotics are causing

dramatic improvements in most of these infections, many of which formerly were routinely fatal. *See* SYPHILIS; TUBERCULOSIS.

In addition to nonpurulent leptomeningitis, many of these microorganisms can produce other lesions, especially masses (fungal granulomas, tuberculomas) or more specific diseases, such as tabes dorsalis (with progressive loss of large afferent neurons) or general paresis (with severe inflammation of the cerebral cortex), two varieties of neurosyphilis.

Viral infections. These infections vary widely geographically, generally related to the necessity for intermediate hosts and vectors (animal reservoirs) by which the virus is spread from animals to humans, or from one human to another. Poliomyelitis, formerly occurring in summer epidemics in areas of poor sewage control, but now largely prevented by effective vaccination of most children, is primarily an intestinal infection which occasionally spreads to the nervous system, infecting and destroying motor nerve cells, thereby producing weakness of certain muscles. Herpes zoster has a similarly restricted preference for infecting sensory nerve cells and producing an acute skin eruption in the distribution of the affected sensory cells. Herpes simplex is closely related to herpes zoster, resides in the trigeminal or sacral sensory nerve cells, and intermittently produces eruptions in the distribution of these cells: "fever blisters" in and around the mouth in type I herpes, or similar blisters in the genital area in type II herpes. The latter is increasingly being recognized as a venereal disease, being widely spread especially through promiscuous homosexual contacts. Rabies virus also affects certain nerve cells in the temporal lobe, as well as in the cerebellum, and is transmitted through the saliva of animals that bite other animals or humans; rabies is the single exception to the rule that immunization must precede infection to be effective, and the immunization must begin promptly after the bite. *See* ANIMAL VIRUS; HERPES; POLIOMYELITIS.

Some viruses infect glial cells and produce demyelination, usually with little inflammation: papovavirus in progressive multifocal leukoencephalopathy and measles virus in subacute sclerosing panencephalitis. Other transmissible diseases have been identified in which there is destruction of a variety of cells, especially in the cerebral cortex but also in many other areas. In these diseases a widespread loss of nerve cells is accompanied by a spongy degeneration and marked proliferation of astrocytes. Clinically there is a slowly progressive mental and neurological deterioration after a very long latent period of many months to years; such "slow virus infections" produce scrapie in sheep and goats, kuru in the Fore tribe of New Guinea (related to their cannibalistic ritual of eating the brains of relatives who have died, usually with the disease), and Jacob-Creutzfeldt's disease in other humans throughout the world, but the route of transmission remains mysterious. *See* VIRUS INFECTION, LATENT, PERSISTENT, SLOW.

Meningeal complications. Complications of infections of the meninges include hydrocephalus (dilatation of the ventricles due to excess cerebrospinal fluid) as the result of obliteration of the leptomeningeal spaces through which the cerebrospinal fluid must circulate before being absorbed; cranial nerve paralyses as the result of the incorporation of the nerves within the leptomeningeal inflammatory exudate; and focal destructions of neural tissue as the result of ischemia secondary to inflammation of the blood vessels (arteritis) coursing through the exudate in the meninges. *See* MENINGES.

Inflammation. Certain viruses, such as lymphocytic choriomeningitis and mumps, frequently produce a meningitis in humans from whose cerebrospinal fluid the virus is relatively easily grown. Certain other viruses, such as measles and varicella, occasionally produce meningitis or encephalomyelitis, but the cerebrospinal fluid does not contain the virus, and the hypothesis is that an allergic reaction is responsible. *See* MENINGITIS.

In infants under 2 years of age an acute encephalopathy (edema of the brain with little other evidence of inflammation) frequently occurs with varicella (chicken pox) or measles, and was one of the main objections to the former practice of vaccinating infants with vaccinia (cowpox) to prevent subsequent infection by smallpox virus. The elimination of smallpox has eliminated the necessity for continued vaccinial vaccinations and this frequently fatal complication.

Allergy to one's own tissue elements is an interesting possibility that has evoked many experimental approaches. One of the best-studied examples is experimental allergic encephalomyelitis. A particular protein (myelin basic protein) in central nervous system myelin can be made to evoke sensitized lymphocytes which recognize the same protein in the host's central myelin. This recognition results in the liberation of lymphokines (specific soluble factors) which attract other leukocytes, including macrophages that destroy the myelin. A similar specific protein in peripheral nervous system myelin can evoke experimental allergic neuritis. *See* AUTOIMMUNITY.

Two of these human diseases, multiple sclerosis affecting the central nervous system and the Landry-Guillain-Barré syndrome affecting the peripheral nervous system, are considered likely candidates eventually to be related to experimental allergic encephalomyelitis and experimental allergic neuritis, respectively. *See* MULTIPLE SCLEROSIS.

In contrast to the remitting-relapsing course of multiple sclerosis, the Landry-Guillain-Barré syndrome is a monophasic inflammation of peripheral nerves which closely resembles experimental allergic neuritis. About half the human cases follow an acute infection of the respiratory or intestinal tracts and present with a symmetric sensory-motor paralysis of face, arms, and legs. The protein in the cerebrospinal fluid typically rises rapidly and persists for many weeks. Recovery is the rule, since the axons are preserved and the Schwann cells merely have to rewrap them, but many months may be necessary for this process to occur.

These two diseases represent the classical examples of demyelinating diseases, one central and the other peripheral, with relative sparing of neurons. Since myelin is the jelly-roll-like wrapping of oligodendroglial or Schwann cell membranes, anything which damages one or the other of these cells selectively can be said to have produced a demyelinating disease.

Vascular disease. Vascular diseases of the nervous system are commonly called strokes, a term which emphasizes the suddenness of onset of neurological disability, as though the individual were struck down by a blow to the head. Such a cataclysmic onset is characteristic of vascular diseases, since the nerve cell can function without nutrients for only a matter of seconds and will die if not renourished within several minutes. *See* VASCULAR DISORDERS.

Hemorrhage. Two main types of hemorrhage occur: hemorrhage into the subarachnoid space from rupture of an aneurysm (a focal weakening and dilatation) of a large artery, usually at the base of the brain and within or near the circle of Willis, especially the anterior half of the circle, the anterior communicating and internal carotid arteries; and hemorrhage into the brain from rupture of an aneurysm of a small artery or arteriole, most commonly one of the lateral basal penetrating (striatal) arteries which are branches of the middle cerebral artery. As might be expected in situations characterized by increased pressure, both types of hemorrhage occur more commonly in hypertensive adults.

Intracerebral hemorrhages occur directly from hypertension as the result of degeneration of the wall of the small arteries resulting in the formation of microaneurysms, which may progressively enlarge and rupture (producing hemorrhage) or may heal with fibrosis of the wall or thrombosis (occlusion) of the lumen (producing necrosis of the tissue supplied). Hemorrhages usually require minutes to a few hours to enlarge, producing increasing focal neurological deficit, but they may rupture into the ventricles or subarachnoid space, producing more general dysfunction, coma, and death in hours to a few days.

If the intracerebral hemorrhage does not greatly enlarge, the clinical differentiation from other types of strokes, which was difficult in the days before computerized tomography, can now be easily distinguished since blood appears opaque and necrosis much less opaque. Although surgical treatment of hemorrhage is possible, prevention of this complication of hypertension is the best medical practice. *See* COMPUTERIZED TOMOGRAPHY.

Subarachnoid hemorrhage occurs as the result of hypertension and arteriosclerosis combining to produce a degeneration of the internal elastic membrane (a layer of extracellular material that gives the artery its major strength), causing a defect which can occur anywhere but may be superimposed on a preexisting congenital absence of the medial muscular wall in the crotch of bifurcations of large arteries, the favorite site of aneurysms of the circle of Willis. This hypothesis of superimposition of several age-related lesions may account for the rarity of aneurysms in children in spite of the high incidence of congenital muscular defects. The abrupt entry of blood under arterial pressure into the subarachnoid space at the base of the brain typically produces sudden severe headache, stiff neck, and unconsciousness. *See* ARTERIOSCLEROSIS; HYPERTENSION.

If clotting of the blood in and about the ruptured aneurysm does not occur promptly, the person will die in a few minutes. Even if clotting occurs, the mortality rate is very high within the next hours to days, during which careful attention must be given to establishing a good airway, sedation, control of hypertension, prevention of lysis of the clot by administration of antifibrinolytic drugs, and early operation if a space-expanding mass of hemorrhage is evident. After a few weeks the individual should be in good enough condition for consideration of direct surgical treatment of the aneurysm itself or reduction in blood flow to the aneurysm. *See* ANEURYSM.

Other causes of intracerebral or subarachnoid hemorrhages, especially in normotensive persons, include arteriovenous malformations (angiomas). The pressure of the blood leaking out is usually less, normotensive arterial or even venous, so that the prognosis is better than in hypertensive strokes; indeed, repeated hemorrhages over a period of years are common with angiomas. *See* HEMORRHAGE.

Occlusions and ischemia. The consequences of occlusion of arteries are more difficult to predict because of the varying degrees of collateral or anastomotic circulation which may be present in different individuals. Nerve cells require oxygen and glucose for functional activity, and can withstand only brief periods of hypoxia or hypoglycemia. Even a few seconds of hypoxia can block the nerve cell's function, and more than 10 min is almost certainly fatal to most nerve cells. Within these narrow limits lies most of the clinical difficulty. Transient ischemic attacks may result, with temporary impairment of blood flow to a part of the brain and consequent focal neurological dysfunction. These attacks serve as warnings of impending disaster if the circulation is restricted too long, but they may also be successfully treated with drugs or surgery and the disastrous major stroke prevented. Myocardial infarction, postural hypotension, and stenosis or narrowing of the carotid or vertebral arteries greater than 60% are common causes of cerebral ischemia. If the ischemia is not rapidly reversed, the neurons undergo selective necrosis; if the ischemia is more severe or prolonged, the glia and blood vessels in the gray matter also undergo necrosis, which appears in a pseudolaminar pattern within the middle layers of the cerebral cortex, ultimately appearing as a narrowing of the cortex; and if the ischemia is still more severe or prolonged, all the gray and white matter in the ischemic zone becomes necrotic, a condition known as cerebral infarction or encephalomalacia.

Brain swelling. Small and large strokes may be suffered clinically as the result of small or large hemorrhages or ischemic episodes occurring in or around the brain. One of the common ways the brain reacts

to these injuries is by swelling. Such swelling itself may be fatal within a few days to a week or so by herniation of the medial temporal lobe through the tentorial notch, a process known as transtentorial herniation, compressing the brainstem, where there are important neural circuits for vital functions, such as breathing and maintenance of blood pressure. If the person survives the first critical week after a massive stroke with swelling, the neurologic deficit itself is rarely fatal, but the accompanying prolonged convalescence may contribute to the development of pneumonia or pyelonephritis, which may be fatal.

Degenerative and other diseases. Degenerative, metabolic, toxic, and deficiency states include the largest numbers of both common and rare diseases of the nervous system. Since neurons in the brain may be destroyed after birth and cannot be replaced, mental deterioration, deafness and blindness, incoordination and adventitious movements, and other neurologic signs that are so typical of these disorders are generally not reversible even if the basic metabolic defect can be corrected.

During the 1970s the greatest advances were made by neurochemists and geneticists in the early diagnosis and treatment of several diseases usually manifest in infancy with mental retardation. The neuropathologic abnormalities are generally rather poorly defined with nonspecific smallness of the brain (microcephaly), retardation in myelination, and so on. Three examples are phenylpyruvic oligophrenia (phenylketonuria or PKU), which is treatable with a phenylalanine-deficient diet—there is a 0.5% loss in general intelligence for each week's delay in beginning the diet; galactosemia, requiring a galactose-free diet also as early as possible to avoid cataracts and mental retardation; and cretinism, which requires treatment with thyroid, and is an important example of a nongenetic metabolic disease with comparably disastrous results if not treated early. *See* METABOLIC DISORDERS; PHENYLKETONURIA.

More specific changes occur in the brain of the newborn (especially the premature) with hyperbilirubinemia: jaundice of the phylogenetically old parts of the brain (kernicterus or bilirubin encephalopathy) due to diffusion of unconjugated bilirubin through the blood-brain barrier. Exchange transfusions can completely prevent the mental retardation, postural stiffness (dystonia), deafness, and other features of this syndrome.

Idiopathic parkinsonism or paralysis agitans is associated with a nonspecific cortical atrophy and other diffuse changes. The most striking change is a loss of neurons of the pigmented-nerve-cell type. The pigmented neurons discharge dopamine at their synapses. This observation has led to the use of L-dopa therapeutically with dramatic effects in many cases. Huntington's chorea is a cerebral cortical and deep cerebral nuclear (striatal) atrophy resulting from inheritance of a simple dominant trait. Wilson's hepatolenticular degeneration (cirrhosis of the liver, Kayser-Fleischer corneal ring, and necrosis of the striatum) is due to accumulation of copper related to the genetically determined lack of a specific copper-binding protein in serum (ceruloplasmin). *See* HUNTINGTON'S DISEASE; PARKINSON'S DISEASE.

Neoplasm. Neoplasms of the nervous system can be conveniently divided into primary and metastatic, the primary into gliomas and others, and the metastatic into bronchogenic and others. These four groups each account for about 25% of all intracranial neoplasms. Half of the gliomas are glioblastoma multiforme, and half of the other primary neoplasms are meningiomas and neurilemomas. Thus, metastatic bronchogenic carcinoma, glioblastoma multiforme, meningioma, and neurilemoma represent about half of the total and can serve as the basis for this discussion.

Meningioma. Meningioma is a benign, slow-growing neoplasm of arachnoidal elements on the undersurface of the dura at certain sites of predilection where arachnoid villi occur. It occurs as a hemispherical mass, the flat base adherent to the dura (sometimes invading the skull) and the rounded peak compressing the brain. Meningiomas typically receive a double blood supply: the external carotid from the dura and the internal carotid (or vertebral) from the brain. This great vascularity and the predilection for the base of the skull add to the difficulty in surgical excision, which often can safely be only incomplete.

Glioblastoma multiforme. Glioblastoma multiforme represents one of the most malignant brain tumors, and poor surgical results are always expected. The average postoperative survival of 4 months can be doubled by excision of most of the tumor followed by local x-irradiation. Glioblastoma appears to be predominantly a tumor of astrocytes. Hemorrhage and necrosis produce varied colors and irregular cysts contributing grossly to its multiform appearance. Glioblastoma is usually confined to one lobe of the brain, except when it crosses in the corpus callosum to produce a "butterfly" bifrontal pattern. It may occasionally occur in the brainstem and spinal cord but practically does not occur in the cerebellum. At the edge of the tumor there is frequently a marked vascular endothelial proliferation, and throughout the tumor large arteries and veins proliferate, producing typical hypervascular patterns on arteriography.

Metastatic bronchogenic carcinoma. Bronchogenic carcinoma is increasing in incidence in spite of widespread publicity regarding one of its major causes, smoking tobacco. About 10% of the cases begin with neurologic symptoms referable to cerebral metastases before pulmonary symptoms appear, and about 35% of all cases end up with metastases in the brain. About 30% of the metastases to the brain are grossly solitary; therefore, there is a reasonable place for neurosurgical therapy aimed at prolonging useful life.

Astrocytoma and medulloblastoma. All of the tumors so far discussed characteristically occur in adults. Two that occur typically in children must be mentioned because they illustrate specifically different aspects. Astrocytoma of the cerebellum is curable by excision, sometimes even if only partial, and not necessarily related to the tumor being cystic or solid. Since this is the only glioma which is curable, it should

not be misdiagnosed, nor should it be confused with solid or cystic astrocytomas occurring outside the cerebellum and being noncurable as they extend so diffusely as to invade practically the whole nervous system. Why the cerebellar astrocytoma is exceptional is not known.

Medulloblastoma also occurs in the cerebellum in children, but is a rapidly growing, malignant, undifferentiated neoplasm which seeds into and spreads widely throughout the cerebrospinal fluid. In spite of these characteristics, 30% 5-year survivals can be obtained by excising the cerebellar mass and giving x-irradiation to the whole head and vertebral column. The small, dark, oval nuclei of this neoplasm are surrounded by very little cytoplasm which theoretically can differentiate into either glial or neuronal fibers, but such differentiation is rare. *See* NERVOUS SYSTEM (VERTEBRATE); TUMOR.

Ellsworth C. Alvord, Jr.
Cheng-mei Shaw

Network theory

The systematizing and generalizing of the relations between the variables flowing in and across elements within an electrical network. To be precise, certain terms are introduced. *See* ALTERNATING-CURRENT CIRCUIT THEORY.

Elements. The elements of a network model are resistance, inductance, and capacitance (the passive elements) and sources of energy (the active elements), which may be either independent sources or controlled, that is, dependent, sources. An independent-voltages source produces a voltage across its terminals that is not dependent on any current or voltage, although it may be a function of time, as in the case of an alternating source; an independent-current source carries current that is independent of all voltages or currents but may be a function of time. *See* CAPACITANCE; ELECTRICAL RESISTANCE; ELECTROMAGNETIC INDUCTION; ELECTROMOTIVE FORCE (EMF); INDUCTANCE.

A number of definitions, together with theorems that relate them, are taken from the mathematical subject of topology. Two or more elements are joined at a node (**Fig. 1**). If three or more elements are connected together at a node, that node is called a junction. (The term major node may be used instead of junction; topological terminology varies among authors.) An element extends from one node to another. A branch of a network extends from one junction to another and may consist of one element or several elements-connected in series. A loop, or circuit, is a single closed path for current. A mesh, or window, is a loop with no interior branch. *See* TOPOLOGY.

Figure 1 shows a network with 13 elements, of which 12 are passive and 1 is active. Nodes are indicated by dots; of the 9 nodes in the figure, 6 are junctions. There are 10 branches of which 5 come together at a single node or junction at the bottom of the figure.

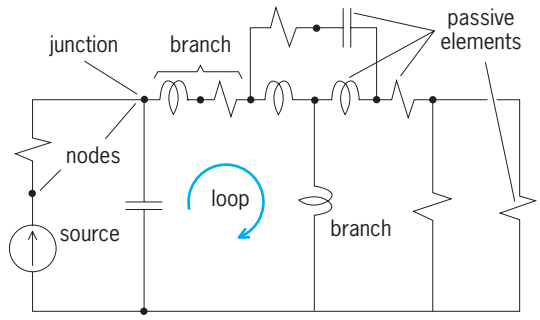

Fig. 1. Parts of a network. (*After H. H. Skilling, Electrical Engineering Circuits, 2d ed., Wiley, 1965*)

Equation (1) is an extension of Ohm's law, where

$$V = IZ \tag{1}$$

I is the current through an element, Z is the impedance of that element, and V is the voltage or potential difference between the nodes that terminate that element; this applies to every passive element of an electrical network. [Equation (1) may be a phasor equation for steady alternating current, in which case impedance is a function of frequency and $V(j\omega) = I(j\omega)Z(j\omega)$, or it may more generally be the transform of the differential equation, $V(s) = I(s)Z(s)$.] The number of elements (active or passive) in a network may be designated as E. *See* ELECTRICAL IMPEDANCE; OHM'S LAW.

At every node of an electrical network the sum of currents entering that node is zero. Equation (2) expresses Kirchhoff's current law. An equation of this form can be written for each node, but in a fully connected network one of these equations can be derived from the others; hence the number of independent node equations is one less than the number of nodes. The number of independent node equations, called N, equals in a fully connected network the number of nodes minus one. *See* KIRCHHOFF'S LAWS OF ELECTRIC CIRCUITS.

$$\Sigma I = 0 \tag{2}$$

Around every loop of an electrical network the sum of the voltages across the elements is zero. Equation (3) expresses Kirchhoff's voltage law. A network such as that shown in Fig. 1 can have many possible loops and hence many equations of this form, but only a limited number are independent. If L is the number of independent loops, topology gives Eq. (4), from which L can be computed, E and N having been counted. Thus in Fig. 1 there are 13 elements and 9 nodes, hence 8 independent nodes, so that there are 13 − 8 or 5 independent loops; that is, $L = 5$.

$$\Sigma V = 0 \tag{3}$$

$$E = N + L \tag{4}$$

If a network is planar (flat) and fully connected, the number of independent loops is equal to the number of meshes, or windows. Figure 1 shows such a

network, and the number of meshes is obviously 5, so again $L = 5$.

Branch equations. There are E elements in a network. Suppose that all impedances are known. There is a voltage across each element, and there is a current through each element; hence there are $2E$ voltages and currents to be known. One equation is provided by each element; either the element is a source for which voltage or current is given, or it is an impedance for which there is a relation given by Ohm's law in the form of Eq. (1). Hence there are E equations from the elements. From Eqs. (2) and (3) the nodes and loops provide $N + L = E$. Thus there are $2E$ equations relating voltages and currents.

In an actual solution for current and voltage in a network, it is probably desirable to reduce the number of equations by combining elements that are in series. This can reduce the number of elements to the number of branches and the number of nodes to the number of junctions; the network is then described by $2B$ branch equations.

These branch equations are easy to write but tedious to solve, unless a computer is used. Two modifications have been devised, however, that eliminate a great deal of the labor and reduce the number of equations from $2B$ to either L, the number of independent loops, or N, the number of independent nodes, as will be described in the following paragraphs.

Linearity. Although branch equations can be written for networks containing either linear or nonlinear elements, the solution is more difficult for nonlinear networks. A linear network is one that gives rise to linear systems of equations, which are subject to special methods of solution, and for which the principle of superposition applies, allowing the use of loop or node equations. In a linear system the values of resistance, inductance, and capacitance are constant with respect to voltage and current, and a controlled source produces a voltage or a current that is proportional to another voltage or current. *See* SUPERPOSITION THEOREM (ELECTRIC NETWORKS).

Fortunately, many electrical networks are linear or are nearly enough linear to be so considered, at least in the useful range of operation or in a piecewise linear fashion. Examples of branch, loop, and node equations are given below for linear systems.

Examples. Using Ohm's law six times, the branch equations for the network of **Fig. 2** are shown in Eqs. (5).

$$\begin{aligned} V_{BC} &= Z_{BC}I_{BC} & V_{CD} &= Z_{CD}I_{CD} \\ V_{BD} &= Z_{BD}I_{BD} & V_{DA} &= Z_{DA}I_{DA} \\ V_{CA} &= Z_{CA}I_{CA} & V_{BA} &= Z_{BA}I_{BA} + E \end{aligned} \tag{5}$$

The source electromotive forces (emf) and branch impedances are assumed to be known, and all currents and branch voltages are to be found. There are, then, 6 equations and 12 unknowns, the unknowns being the voltage across and current through each branch. (Although a source is indicated in only one of these branches, the method can be applied in the same way if there are sources in any or all of the branches.) Clearly, 6 more equations are needed.

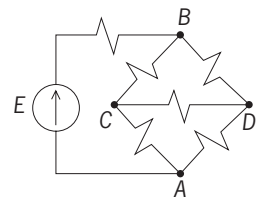

Fig. 2. Network of six branches. (*After H. H. Skilling, Electrical Engineering Circuits, 2d ed., Wiley, 1965*)

The 6 needed equations may be called connection equations, and they are found from Kirchhoff's laws. The current law gives Eqs. (6), called node equations

$$I_{BC} + I_{BD} + I_{BA} = 0 \tag{6a}$$

$$I_{CB} + I_{CD} + I_{CA} = 0 \tag{6b}$$

$$I_{AC} + I_{AD} + I_{AB} = 0 \tag{6c}$$

or junction equations. Equation (6*a*) shows the sum of the currents flowing out of junction B to be zero. Equations (6*b*) and (6*c*) make similar statements with regard to junctions C and A.

Three more equations may be obtained from Kirchhoff's voltage law. The sum of the voltages around loop BCD must be zero, Eq. (7*a*); the sum of

$$V_{BC} + V_{CD} + V_{DB} = 0 \tag{7a}$$

$$V_{CA} + V_{AD} + V_{DC} = 0 \tag{7b}$$

$$V_{AB} + V_{BC} + V_{CA} = 0 \tag{7c}$$

the voltages around loop CAD must be zero, Eq. (7*b*); and the sum of the voltages around loop ABC must be zero, Eq. (7*c*).

There are now 12 equations. Recognizing that $I_{BC} = -I_{CB}$, and so on with other currents and voltages, there are still 12 unknowns, for the 6 connection equations of Eqs. (6) and (7) have added no new unknowns. The 12 equations can be solved simultaneously for the 12 unknowns. The actual solution is not particularly interesting and will not be pursued.

It would seem to be possible to write a fourth junction equation at junction D: $I_{BD} + I_{CD} + I_{AD} = 0$. This proposed equation contains no new information, however, for it could have been derived from the other three equations. It results from adding Eqs. (6*a*), (6*b*), and (6*c*) and canceling equal and opposite quantities. Although it is a true equation, it is not an independent equation. The solution for 12 unknowns requires the use of 12 independent equations.

A somewhat similar observation can be made about the independence of loop equations. The three loops for which equations are written are not the only possible ones. For example, there is also loop $ACBD$, but this loop will not give another independent equation, nor will any of the other possible loops.

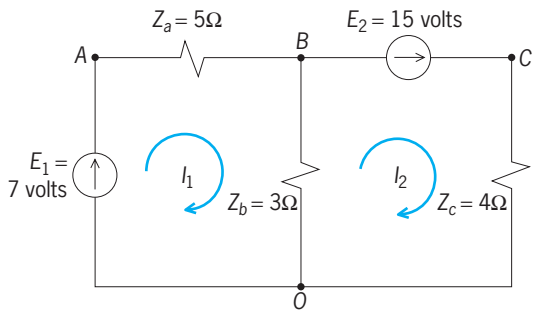

Fig. 3. Network of two loops. (*After H. H. Skilling, Electrical Engineering Circuits, 2d ed., Wiley, 1965*)

Loop equations. The ingenuity of the loop method lies in the selection of currents to be determined. It is necessary to find only as many different currents as there are independent loops, instead of finding as many different currents as there are branches.

Thus in the network of **Fig. 3** a current which can be called I_1 flows around loop 1. (This is the current in the source E_1 and in the impedance Z_a.) The convention generally adopted is to take the reference direction for all loop currents as clockwise, and this will be taken as the reference direction of I_1. The reference direction of the current called I_2 that flows around loop 2 is also clockwise. This is the current in the source E_2 and in impedance Z_c. Loop currents I_1 and I_2 both flow in Z_b. The reference direction of one is downward and that of the other is upward. Thus the total current downward in Z_b is $I_1 - I_2$. When I_1 and I_2 are known, every current in the network is determined.

The first step in solving for these currents is to write network equations in terms of I_1 and I_2. Kirchhoff's voltage law is used to write Eq. (8), which

$$E_1 - Z_a I_1 - Z_b(I_1 - I_2) = 0 \tag{8}$$

says that the sum of all voltages about loop 1 is zero. Rearranged, Eq. (8) becomes Eq. (9). This equation

$$(Z_a + Z_b)I_1 - Z_b I_2 = E_1 \tag{9}$$

is valid when there is current in both loops of the network.

Similarly, Eq. (10) is written for loop 2 of the network. This is done from Fig. 3 without further explanation. The first term is the voltage in loop 2 when the only current is I_2, and the second term is the voltage in loop 2 caused by I_1. Since E_2 has the same reference direction as I_2, it is positive.

$$(Z_b + Z_c)I_2 - Z_b I_1 = E_2 \tag{10}$$

Equations (9) and (10) are a pair of equations with only two unknowns, I_1 and I_2. They can be solved simultaneously by Eqs. (11*a*) and (11*b*). By way of illustration let the electromotive forces and impedances be given numerical values and solve for the currents as follows. It is assumed that the equations are linear (see above), which implies that the resistances are constant, and that the electromotive forces are independent of current and voltage.

$$(Z_a + Z_b)I_1 - Z_b I_2 = E_1 \tag{11a}$$

$$-Z_b I_1 + (Z_b + Z_c)I_2 = E_2 \tag{11b}$$

Examples. If the impedances shown in Fig. 3 are pure resistances and with the values shown, Eqs. (12) and (13) may be written. The electromotive forces may be either dc values or phasors of alternating voltages (the solution is the same). These are linear equations that can be solved by any convenient means. Using determinants and Cramer's rule, they are solved in Eqs. (14) and (15), where D, the denominator of I_2, is the same as the denominator of I_1. The results are two loop currents in amperes. *See* LINEAR SYSTEMS OF EQUATIONS.

$$8I_1 - 3I_2 = 7 \tag{12}$$

$$-3I_1 + 7I_2 = 15 \tag{13}$$

$$I_1 = \frac{\begin{vmatrix} 7 & -3 \\ 15 & 7 \end{vmatrix}}{\begin{vmatrix} 8 & -3 \\ -3 & 7 \end{vmatrix}} = \frac{49 + 45}{56 - 9} = \frac{94}{47} = 2 \tag{14}$$

$$I_2 = \frac{\begin{vmatrix} 8 & 7 \\ -3 & 15 \end{vmatrix}}{D} = \frac{120 + 21}{D} = \frac{141}{47} = 3 \tag{15}$$

Now all unknown quanitites in the network may be easily found. Current in the central branch is $I_{BO} = I_1 - I_2 = -1$ ampere. The negative sign indicates that 1 amp is flowing upward. If the bottom node (node O) is taken to be the reference node at an assumed zero potential, then the potential at node B is -3 volts. The potential at node A is $-3 + 5 \cdot 2 = 7$ volts, which is also the electromotive force of the source E_1. The potential at node C is $-3 + 15 = 12$ volts, which can also be found (across Z_c) as $4 \cdot 3 = 12$ volts. The most convenient way to specify all the voltages of a network is to give the potentials at the various nodes with reference to some one node that is arbitrarily assumed to be at zero potential.

Standard notations. It is customary to use a standard system of symbols for writing loop equations. Equations (11) are specific examples of the general form shown in Eqs. (16). Equations (16) are a set of L simultaneous linear equations, applying to the L loops of a network; the network may be any network and L may be any number. In writing the equations, the following conventions are used. Each loop current is numbered, as I_1, I_2, and so on. V_1 is a voltage in loop 1 that is not taken into account by the terms of the left-hand side of the equation. It may be, as it

$$Z_{11}I_1 + Z_{12}I_2 + Z_{13}I_3 + \cdots + Z_{1L}I_L = V_1 \tag{16a}$$

$$Z_{21}I_1 + Z_{22}I_2 + Z_{23}I_3 + \cdots + Z_{2L}I_L = V_2 \tag{16b}$$

$$Z_{31}I_1 + Z_{32}I_2 + Z_{33}I_3 + \cdots + Z_{3L}I_L = V_3 \tag{16c}$$

$$Z_{L1}I_1 + Z_{L2}I_2 + Z_{L3}I_3 + \cdots + Z_{LL}I_L = V_L \tag{16d}$$

was in Eq. (9), an independent electromotive force. It may be the sum of several voltages, and must include all voltages that appear in loop 1 when all the other loops are open. Note that its nominal positive direction is taken to be that of I_1.

The total impedance about loop 1 is Z_{11}. In Eq. (11), which applies to Fig. 3, Z_{11} is $Z_a + Z_b$. It might include many more elements if the network were more complicated. Z_{11} is called the self-impedance of loop 1. It could be measured by means of a bridge or other impedance-measuring instrument connected in place of the source in loop 1, all other loops of the network being opened during the measurement. Each loop has its self-impedance: Z_{22}, Z_{33}, and so on.

Certain branches are common to two loops. Thus Z_b in Fig. 3 is in both loop 1 and loop 2. In such a case, current in one loop produces voltage in another loop, and there is said to be mutual impedance. By definition, the mutual impedance is the ratio of such a voltage in one loop to the current in another loop that produces it. That is, if current in loop 2 is I_2 and mutual impedance with loop 1 is Z_{12}, then the resulting voltage in loop 1 is $Z_{12}I_2$. For example, Eq. (11) shows that the voltage produced in loop 1 of Fig. 3 by the current in loop 2 is $-Z_bI_2$. By definition, then, $Z_{12} = -Z_b = -3$. The negative sign results from the fact that in the common element the reference direction of I_2 is opposite to the reference direction of I_1. If positive I_1 produces a positive voltage, positive I_2 in the same element produces a negative voltage. In network computations the reference directions are commonly assumed in such a way that the mutual impedances (such as Z_{12}, Z_{23}, and so on) are negative quantities. However, it is not wrong to direct the arrows that indicate the nominal positive direction of current in such a way that mutual impedances are positive.

In Eq. (16*b*) the first term contains Z_{21}. Comparison with Eq. (11) shows that, for the network of Fig. 3, $Z_{21} = -Z_b$. Z_{21} is therefore equal to Z_{12}. When two loops contain resistors or coils or capacitors in a common branch, current in loop 1 will produce the same voltage in loop 2 that equal current in loop 2 would cause in loop 1; therefore $Z_{21} = Z_{12}$. The general form for this relation is given in Eq. (17). Al-

$$Z_{pq} = Z_{qp} \tag{17}$$

though less evident, this relation is still true if two circuits are coupled by a magnetic field (as in a transformer); whatever the turn ratio, the mutual inductance L_{21} equals the mutual inductance L_{12}. If the coupling between the loops is by means of an electric field through mutual capacitance with no conductive connection, Eq. (17) is again valid. Indeed, it fails only for circuit elements that are not bilaterally symmetrical. (For example, a transistor or a vacuum tube is not bilateral.)

The order of the subscripts attached to Z has the following significance. The first subscript is the number of the equation in the array of Eqs. (16) and therefore agrees with the subscript attached to V in that equation. The second subscript is the number of the term in the equation and therefore agrees with the subscript attached to I in that term. Thus $Z_{pq}I_q$ is a voltage in circuit p produced by a current in circuit q.

Nodal equations. Loop equations are written based on the concept of loop currents. This makes it unnecessary to give any attention to Kirchhoff's current law, for loop currents necessarily add to zero at every node, and Kirchhoff's current law is automatically satisfied. The loop-current concept therefore reduces the number of equations that must be solved simultaneously from the $2B$ equations of the branch method to L, the number of independent loops, which is usually about one-fourth as many.

In the node-equation method, the simplifying concept is the idea of measuring voltage from all the nodes of the network to one particular node that is called the reference node, or the datum node. This makes it unnecessary to give any attention to Kirchhoff's voltage law. It is only necessary to satisfy Kirchhoff's current law at each node, for the voltage law is automatically satisfied. Thus the number of simultaneous equations is reduced to the number of independent nodes N, a number much smaller than $2B$ and comparable with L.

Whether the node method or the loop method is the more convenient depends on the network. Some networks have fewer loops than nodes, and some have fewer nodes than loops. Other factors also affect the relative convenience, as will be seen.

Figure 4 shows a network with two independent nodes; that is, it has three nodes, one of which is the reference node O, and the others are marked A and B. Kirchhoff's current law is used for node A, Eq. (18), and node B, Eq. (19). Assume that the po-

$$I_a = I_b + I_c \tag{18}$$

$$I_e = -I_d + I_c \tag{19}$$

tential at node O is zero; if the potential at node A is called V_A, then $I_b = Y_bV_A$. Also, $I_d = Y_dV_B$ and

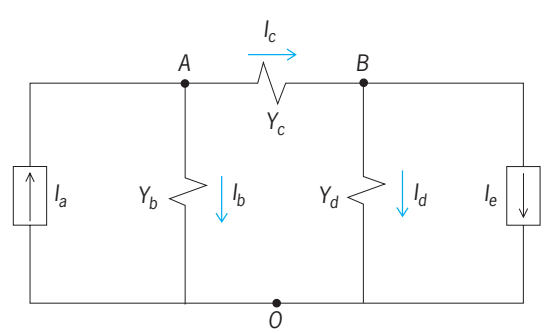

Fig. 4. Network with two independent nodes; a third one is the reference node *O*. (*After H. H. Skilling, Electrical Engineering Circuits, 2d ed., Wiley, 1965*)\vskip-1.2pt

$I_c = Y_c(V_A - V_B)$. Equations (20) and (21) may now be solved for I_a and I_e.

$$\begin{aligned} I_a &= Y_b V_A + Y_c(V_A - V_B) \\ &= (Y_b + Y_c)V_A - Y_c V_B \end{aligned} \tag{20}$$

$$\begin{aligned} I_e &= -Y_d V_B + Y_c(V_A - V_B) \\ &= Y_c V_A - (Y_c + Y_d)V_B \end{aligned} \tag{21}$$

With practice it is easy to write the final form of Eqs. (20) and (21) directly. To write Eq. (20), note that the independent value I_a of incoming current is set equal to the current that would flow from node A if all other nodes were at zero potential, that is, $(Y_b + Y_c)V_A$, from which is subtracted the current that would flow to node A if it were at zero potential while all other nodes were at their actual potentials, that is, $Y_c V_B$. To write the first term, assume that all nodes except A are short-circuited to node O; to write the second term, assume that node A alone is short-circuited to node O.

To write Eq. (21) at node B, the independent incoming current $-I_c$ is equated to current that would flow from node B if node A were at zero potential, that is, $(Y_c + Y_d)V_B$, less the current that would flow to node B if it alone were at zero potential, in this case $Y_c V_A$; the result, with all signs changed, is Eq. (21).

Standard notations. There is a standard form for writing node equations similar to the standard form for loop equations. For a network of N independent nodes, Eqs. (22) may be written. Y_{AA} is called the

$$\begin{aligned} &Y_{AA}V_A + Y_{AB}V_B + Y_{AC}V_C + \cdots + Y_{AN}V_N = I_A \\ &Y_{BA}V_A + Y_{BB}V_B + Y_{BC}V_C + \cdots + Y_{BN}V_N = I_B \\ &Y_{CA}V_A + Y_{CB}V_B + Y_{CC}V_C + \cdots + Y_{CN}V_N = I_C \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &Y_{NA}V_A + Y_{NB}V_B + Y_{NC}V_C + \cdots + Y_{NN}V_N = I_N \end{aligned} \tag{22}$$

self-admittance at node A, and in the example it is equal to $(Y_b + Y_c)$. Note that Y_{AA} is the sum of all admittances attached to node A. Y_{BB}, Y_{CC}, ... are self-admittances at the other nodes.

Y_{AB} is the mutual admittance between nodes A and B. In the example $Y_{AB} = -Y_c$; Y_{BA} also equals $-Y_c$. Both Y_{AB} and Y_{BA} are the sum of all admittances connected directly between nodes A and B but written with a negative sign.

I_A is another current flowing toward node A. In the example it is the source current designated I_a in Fig. 4. Similarly, I_B is current toward node B, and in the example given it is the source current $-I_c$.

It will be noted that every term in the node Eqs. (22) is a current, as every term in the loop Eqs. (16) is a voltage.

Examples. As a numerical example, the following values are given to the impedances and the source currents of Fig. 4, and then voltage at the two independent nodes A and B and current in the three impedances are obtained. Note that the source currents are taken to be real numbers, implying either that they are direct current, or (if ac sources) that they are in phase with each other; the impedances are taken to be real, and this implies pure resistance. In a practical problem one voltage and all impedances might well be complex, and this would complicate the arithmetic but not change the method of solution. Given that $I_a = 2$ amperes, $I_e = 5$ amperes, $Y_b = {}^1/_4$ mho, $Y_c = 1$ mho, and $Y_d = {}^1/_2$ mho, node equations (22) are written as Eqs. (23) and (24).

$$({}^1/_4 + 1)V_A - 1V_B = 2 \tag{23}$$

$$-1V_A + (1 + {}^1/_2)V_B = -5 \tag{24}$$

These equations may be solved for V_A and V_B using Cramer's rule; they are shown as Eqs. (25) and (26).

$$V_A = \frac{\begin{vmatrix} 2 & -1 \\ -5 & {}^3/_2 \end{vmatrix}}{\begin{vmatrix} {}^5/_4 & -1 \\ -1 & {}^3/_2 \end{vmatrix}} = \frac{-2}{{}^7/_8} = -{}^{16}/_7 \text{ volts} \tag{25}$$

$$V_B = \frac{\begin{vmatrix} {}^5/_4 & 2 \\ -1 & -5 \end{vmatrix}}{{}^7/_8} = \frac{-{}^{17}/_4}{{}^7/_8} = -{}^{34}/_7 \text{ volts} \tag{26}$$

Currents I_b, I_c, and I_d are given in Eqs. (27)–(29). Equations (30) and (31) confirm these results. Note

$$I_b = V_A Y_b = -{}^{16}/_7\ {}^1/_4 = -{}^4/_7 \text{ amperes} \tag{27}$$

$$\begin{aligned} I_c &= (V_A - V_B)Y_c = \left(-{}^{16}/_7 + {}^{34}/_7\right)1 \\ &= {}^{18}/_7 \text{ amperes} \end{aligned} \tag{28}$$

$$I_d = V_B Y_d = -{}^{34}/_7\ {}^1/_2 = -{}^{17}/_7 \text{ amperes} \tag{29}$$

$$I_a = I_b + I_c = -{}^4/_7 + {}^{18}/_7 = 2 \text{ amperes} \tag{30}$$

$$I_e = I_c - I_d = {}^{18}/_7 + {}^{17}/_7 = 5 \text{ amperes} \tag{31}$$

that currents I_b and I_d both turn out to be upward because of the large value of the source current I_e, and that therefore the two nodes A and B are both at negative potential compared with node O.

Thévenin's theorem. It is often convenient before applying network theory to simplify a problem by means of Thévenin's theorem. This theorem and its dual, Norton's theorem, may be expressed in many ways, but the following is among the more useful: Open-circuit voltage V_θ and short-circuit current I_θ are measured (or computed) at a pair of terminals of an active linear network. The active network is equivalent at these terminals to either an independent voltage source V_θ in series with an impedance $Z_\theta = V_\theta/I_\theta$, or alternatively to an independent-current source I_θ in parallel with the same Z_θ (**Fig. 5**). The former alternative is Thévenin's theorem, the latter is Norton's. A proof is not given; however, Eqs. (32)–(37) illustrate the application. *See* THÉVENIN'S THEOREM (ELECTRIC NETWORKS).

For the example of node equations Fig. 4 is used with the data given above to find currents

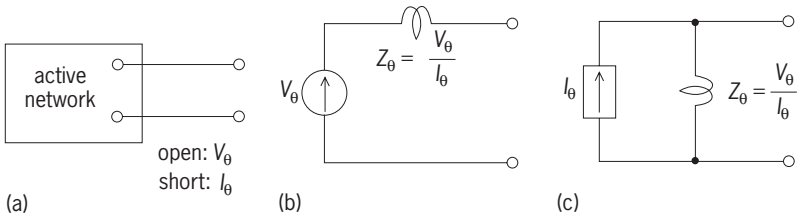

Fig. 5. Networks: (*a*) active, (*b*) Thévenin equivalent, and (*c*) its Norton equivalent. (*After H. H. Skilling, Electrical Engineering Circuits, 2d ed., Wiley, 1965*)

and voltages in the network, using Thévenin's theorem.

Thévenin's theorem is applied twice to Fig. 4 to obtain the circuit of **Fig. 6**. Voltage V_1 and impedance Z_1 are equivalent to current I_a and admittance Y_b; similarly V_2 and Z_2 are equivalent to I_e and Y_d, as follows.

In Eqs. (32) the open-circuit voltage is computed at each pair of terminals with Z_c open.

$$V_1 = \frac{I_a}{Y_b} = \frac{2}{{}^1/_4} = 8 \qquad (32a)$$

$$V_2 = \frac{I_e}{Y_d} = \frac{5}{{}^1/_2} = 10 \qquad (32b)$$

Short-circuit current is computed with each of nodes A and B short-circuited to node O in Eqs. (33).

$$I_1 = 2 \qquad (33a)$$

$$I_2 = 5 \qquad (33b)$$

The impedances, Eqs. (34), may be obtained from Eqs. (32) and (33).

$$Z_1 = \frac{V_1}{I_1} = {}^8/_2 = 4 \qquad (34a)$$

$$Z_2 = \frac{V_2}{I_2} = {}^{10}/_5 = 2 \qquad (34b)$$

The current I_c may be obtained from Fig. 6, Eq. (35).

$$I_c = \frac{V_1 + V_2}{Z_1 + Z_c + Z_2} = \frac{8 + 10}{4 + 1 + 2} = {}^{18}/_7 \text{ amperes} \qquad (35)$$

Voltages V_A and V_B at nodes A and B are given by

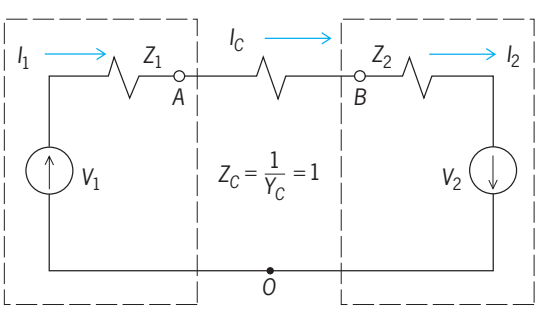

Fig. 6. Thévenin equivalent of Fig. 4.

Eqs. (36) and (37). These answers are, of course, the

$$V_A = V_1 - Z_1 I_c = 8 - 4({}^{18}/_7) = -{}^{16}/_7 \text{ volts} \qquad (36)$$

$$V_B = -(V_2 - Z_2 I_c) = -\left[10 - 2({}^{18}/_7)\right] = -{}^{34}/_7 \text{ volts} \qquad (37)$$

same as those obtained by node equations (25) and (26).

Many other network theorems are available, but perhaps none are as useful as Thévenin's and Norton's.

Hugh Hildreth Skilling

Bibliography. J. Choma, *Electrical Networks: Theory and Analysis*, 1985, reprint 1991; R. D. Strum and J. R. Ward, *Electric Circuits and Networks*, 2d ed., 1985; G. H. Tomlinson, *Electric Networks and Filters*, 1991; J. Vlach, *Basic Network Theory with Computer Applications*, 1992; A. K. Walton, *Network Analysis and Practice*, 1987; J. E. Whitehouse, *The Principles of Network Analysis*, 1991.

Neural crest

A strip of ectodermal material in the early vertebrate embryo inserted between the prospective neural plate and epidermis. After closure of the neural tube the crest cells migrate into the body and give rise to parts of the neural system: the main part of the visceral cranium, the mesenchyme, the chromaffin cells, and pigment cells. The true nature of the neural crest eluded recognition for many years because this primary organ has a temporary existence; its cells and derivatives are difficult to analyze when dispersed throughout the body. The fact that mesenchyme arises from this ectodermal organ was directly contrary to the doctrine of the specificity of the germ layers. *See* GERM LAYERS.

Neural crest no doubt exists, with similar qualities, in all vertebrate groups, including the cyclostomes. It has been most thoroughly studied in amphibians and the chick.

Crest cells. The dorsal ectoderm of a vertebrate gastrula forms the pear-shaped neural plate which is surrounded by the neural ridge (**Fig. 1**). When the neural plate rolls up to form the neural tube, the ridges from each side meet and fuse, temporarily forming a wedge-shaped cell mass in the dorsal midline of the prospective brain and spinal cord. In Fig. 1 the head part of the ridge of a urodele neurula is divided into eight zones. Zones 1 and 2 may be called the transverse ridge. Zone 8 and the posterior parts are the trunk ridge. The neural crest cells are situated in the ridge but do not occupy the whole ridge. For example, the posterior slope of the transverse ridge consists of material for the forebrain, and the anterior slope forms epidermis. Investigations have confirmed that probably no crest cells emanate from the transverse ridge. Concerning the main parts of the neural ridge it is still doubtful whether the crest is located only in the peripheral part of the thick plate or

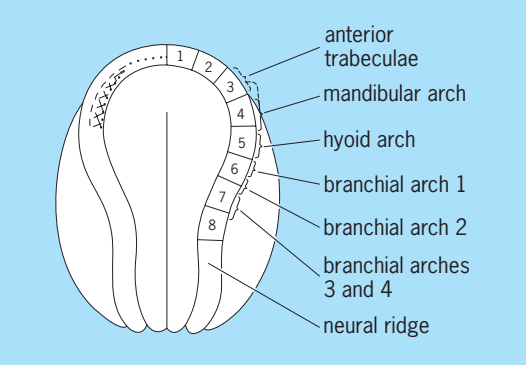

Fig. 1. Diagram of open neural-plate stage of a urodele, *Ambystoma mexicanum*. Left side, the epidermal line of coalescence (broken line), the line of coalescence of the brain (dotted line), and the presumptive ectomesenchyme (hatched area). (*After S. Hörstadius, The Neural Crest: Its Properties and Derivatives in the Light of Experimental Research, Oxford, 1950, facsimile ed., Hafner, 1969*)

in the adjacent part of the thinner ectoderm as well. At least in zones 1–4 of the *Ambystoma* neurula the line of coalescence of the epidermis coincides with the midline of the ridge, the crest material thus forming the proximal slope of the ridge (hatched area in Fig. 1). The exact position of the material in the posterior parts and in other vertebrate groups is not known.

The crest cells do not remain long in the dorsal midline of the neural tube. They migrate in a lateral direction between the tube and epidermis, first as rather coherent sheets, then in groups, strands, or as single cells, both proximal and distal to the somites. Because they are similar to the mesenchymal cells originating from the mesoderm, their fate has been ascertained only by vital staining, extirpation, or transplantation experiments in many cases. One method has been to replace a part of the ridge by a corresponding piece from another species with cell nuclei of different size which are thus recognizable in sections. The fate of crest cells has been studied using a radioautographic method whereby cells of transplants were labeled by tritiated thymidine. Most extirpations and transplantations must be made bilaterally because after unilateral operation, crest cells migrate down in great quantities from the intact to the operated side, and extirpations must also include rather long pieces of the ridge because the gap may be filled by cells from the crest anterior and posterior to the excision. Transplantation of neural crest to another part of the body as well as explantation have been used in order to determine what it can give rise to in new surroundings. In spite of all efforts, however, the problem of the fate of the neural crest is still not solved in all its details (**Fig. 2**). *See* FATE MAPS (EMBRYOLOGY).

Contributions to the neural system. Parts of the migrating bilateral cell masses settle in clusters and give rise to the spinal ganglia. The segmental arrangement of the dorsal nerve roots and the ganglia is not intrinsic to the cells but is determined by the segmental arrangement of the somites. There seems to be no doubt that the spinal ganglia are formed entirely of neural crest cells. There has been much controversy concerning the origin of the cranial ganglia. Placodes of the lateral epidermis of the head are one source of their cells; the other is the neural crest, specifically for the sensory components of the ganglia of cranial nerves V, VII, IX, and X, although to different degrees in different groups of animals. *See* NERVOUS SYSTEM (VERTEBRATE).

According to both histological observations and experiments, the sympathetic trunk has been considered to be derived not only from the crest but also from the ventral part of the spinal cord; however, crest contribution has been denied by some. Excision experiments on chicks support crest origin, and the experiments with labeled crest cells described above favor the theory that the whole sympathetic complex emanates from the crest.

The cellular sheath of Schwann enclosing peripheral nerves has been considered by some authors to have its origin from crest cells, by others to be formed by cells emerging from the neural tube by way of the ventral roots. Experiments with labeled cells have confirmed both these opinions; that is, a first wave of cells comes from the crest, and a later wave comes from the neural tube. The majority of the leptomeninx (pia and arachnoidea) cells are derived from the neural crest, whereas the pachymeninx (dura) is composed chiefly of endomesodermal cells. For differentiation of crest cells into cells of Schwann and meninges the presence of neural material seems indispensable.

Chromaffin cells. The chromaffin cells of the suprarenal organs of lower vertebrates as well as the medullary zone of the suprarenal gland in mammals, and also those of the paraganglia, that is, the aortic bodies of Zuckerkandl and the carotid bodies, are all considered to be formed by cells emanating from the sympathetic system, thus probably mainly from the crest. *See* ADRENAL GLAND; AORTIC BODY.

Mesenchyme and skeleton. The statement by J. B. Platt in the 1880s that mesenchyme could arise from ectoderm (mesectoderm, ectomesoderm, or ectomesenchyme) raised a vivid discussion for several decades. Experiments on amphibian larvae confirmed that, of the chondrocranium, the anterior half of the trabeculae, all visceral arches, and the first basibranchial are of neural crest origin, whereas the main part of the neurocranium and the second basibranchial are endomesodermal (**Fig. 3**). Crest cells may also give a small contribution to the basal plate, the parachordalia, and the posterior trabeculae.

In addition, membrane bones, namely the tooth-bearing premaxilla, vomeropalatine, dentary, and splenial, as well as the odontoblasts and dentine, are formed by cells emanating from the neural crest. Moreover, the ectomesenchyme of the teeth induces the formation of enamel organs, but only in connection with an inducing action from the foregut endoderm. *See* SKELETAL SYSTEM.

The transverse ridge, zones 1 and 2 in Fig. 1, has no cartilage-forming capacity. The trabecular material situated in zone 3 cannot partake in formation of the visceral arches. The mandibular ectomesenchyme

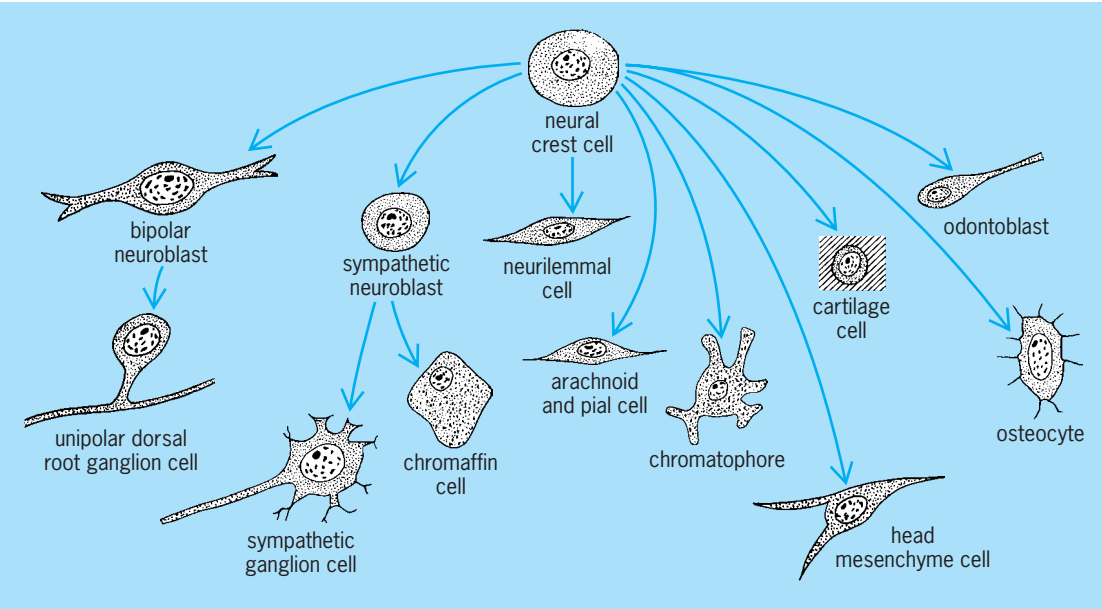

Fig. 2. Schematic representation of the possible developmental fates of a neural crest cell. (*After W. J. Hamilton, J. D. Boyd, and H. W. Mossman, Human Embryology, 2d ed., Williams and Wilkins, 1952*)

located in zone 4 and perhaps also in the adjacent part of zone 3 cannot form trabeculae but is capable of forming other visceral arches, although there is no fusion with the basibranchials; the material of the hyoid arch and the gill arches in zones 5–7 shows such capacity to fuse. In zone 8 the trunk crest begins; it has no faculty to form cartilage but differentiates to form spinal and sympathetic ganglia as well as chromaffin, pigment, and other mesenchymal cells. Crest cells in the neural ridge of the head are not yet determined to form cartilage. No cartilage is formed when they are transplanted to the trunk unless they are activated by certain tissues, such as the pharyngeal endoderm and to some extent the gastric endoderm, endomesoderm of the gill region, chorda, and wounds in somites. The necessity of induction has been confirmed by tissue culture experiments. Ectomesenchyme and Schwann cells differentiate in laboratory cultures, but cartilage, dentine, teeth, and

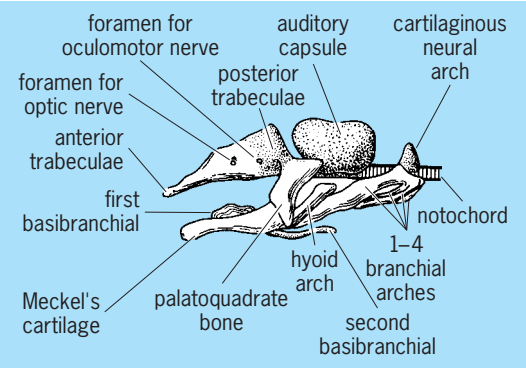

Fig. 3. Diagram of chondrocranium from left of *Ambystoma mexicanum* larva. Cartilage of endomesodermal origin dotted. (*After S. O. Hörstadius, The Neural Crest: Its Properties and Derivatives in the Light of Experimental Research, Oxford, 1950, facsimile ed., Hafner, 1969*)

balancers appear only in the presence of foregut endoderm and stomodeal ectoderm or cartilage after treatment by extracts from foregut. The differentiation is controlled by a phenylalanine metabolism in the archenteral roof.

To what extent corium cells are of crest origin is not clearly established. Head mesenchyme is also formed by the crest. In urodele larvae no dorsal fin is formed, nor do the larval gills develop normally in absence of the crest.

Pigment cells. With the exception of the retinal pigment, which is formed in place, probably all chromatophores in vertebrates are derived from the neural crest. Such an origin has been observed in fishes, amphibians, birds, and mammals. Pigment cells are found in both dermis and epidermis, feathers, hair, the perineural and perivascular layers, the coelomic wall, and the choroid and iris of the eye. The origin of pigment cells was first traced in the twentieth century. Because the pigment is formed rather late in embryonic development, the presumptive pigment cells cannot be distinguished from mesenchymal cells in early stages. There are brown to black melanophores, yellow lipophores (xanthophores), reddish erythrophores (allophores), and guanophores and iridocytes with resplendent crystals. Bilateral extirpation of sufficiently large pieces of neural crest results in unpigmented regions of the body. Following extirpation of all the neural ridges, urodelian larvae subsequently obtain pigment cells emerging from the brain and the tail-bud part of the neural plate. Normally the reserve in these regions is probably inhibited by the pigment cells from the crest. Explants of neural crest may give pigment cells in tissue cultures. Implantation of neural crest into a larva of another species has produced pigment of the type of the donor amphibians, birds, and mammals.

The development of pigment in the chromatophore does not result from self-differentiation but is dependent upon factors in surrounding cells. For example, chromatophores of the black axolotl produce no pigment with white specimens, whereas chromatophores from the white race are activated and form melanin granules in the skin of a black specimen. Some prospective pigment cells in amphibians may remain dormant during larval development and become activated at metamorphosis. The activation of the propigment is evidently brought about by an oxidase. The pigment pattern depends both on properties intrinsic in the propigment cells and on the surrounding tissues. Also, when transplanted to a host of very distant relationship, as from anuran or urodele, the single cells differentiate in size, shape, and color as they would have done in the donor. However, their distribution may vary. *See* CHROMATOPHORE.

In reciprocal exchange of neural crest between some species of urodeles the pattern in the respective regions becomes that of the donor, but in less related forms influence of the host is more or less recognizable. Grafting of epidermis from another species has hardly any effect. In some urodelian species there is an intrinsic tendency in the crest cells to aggregate at certain levels, for example, along the dorsal border of the somites. Melanoblasts invade developing feathers, producing their specific color and pattern. They also invade growing hair, giving pigment granules to it. But in some cases, both in the bird and mammal, the color is a differentiation dependent upon the epidermis. The black and white stripes in the feather pattern of the Barred Plymouth Rock chicken are probably the result of an inhibitory action from each black stripe on the melanoblasts of the prospective white stripe during the process of feather growth. *See* EMBRYONIC INDUCTION; EYE (VERTEBRATE); NEURULATION.

Sven Hörstadius

Bibliography. J. Brachet and H. Alexander, *Introduction to Molecular Embryology*, 2d rev. ed., 1986; G. M. Edelman, *Topobiology: An Introduction to Molecular Embryology*, 1993; S. Nona et al. (eds.), *Development and Regeneration of the Nervous System*, 1992.

Neural network

An information-processing device that consists of a large number of simple nonlinear processing modules, connected by elements that have information storage and programming functions. The field of neural networks is an emerging technology in the area of machine information processing and decision making. This technology makes extensive use of the terminology and methods of artificial intelligence. The main thrusts are toward highly innovative machine and algorithmic architectures, radically different from those that have been employed in conventional digital computers, and toward overcoming the major technological shortcomings and inherent limitations imposed by the traditional information-processing machines. The information-processing elements and components of neural networks, inspired by neuroscientific studies of the structure and function of the human brain, are conceptually simple. Three broad categories of neural-network architectures have been formulated which exhibit highly complex information-processing capabilities. Several generic models have been advanced which offer distinct advantages over traditional digital-computer implementation. Neural networks have created an unusual amount of interest in the engineering and industrial communities by opening up new research directions and commercial and military applications.

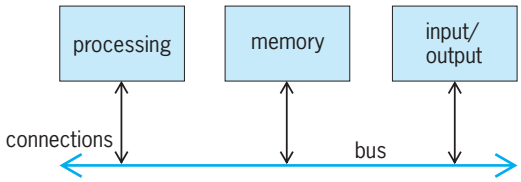

Fig. 1. Architecture of conventional digital computer.

Three decades of rapid growth in the field of neural networks were followed by a period, beginning about 1990, of consolidation and realignment to mathematics, microbiology, and molecular technology. The goal of this work is the establishment of comprehensive foundations to deal with the enormously difficult problems emanating from such areas as globalization of commerce, post-genome health care, and automation of human knowledge and intelligence.

Basic properties and architectural features. Automated information processing is achieved by means of modules that in general involve four functions: input/output (getting in and out of the machine), processing (executing prescribed specific information-handling tasks), memory (storing information), and connections between different modules providing for information flow and control. Neural networks contain a very large number of simple processing modules. This contrasts with traditional digital computers, which contain a small number of complex processing modules that are rather sophisticated in the sense that they are capable of executing very large sets of prescribed arithmetic and logical

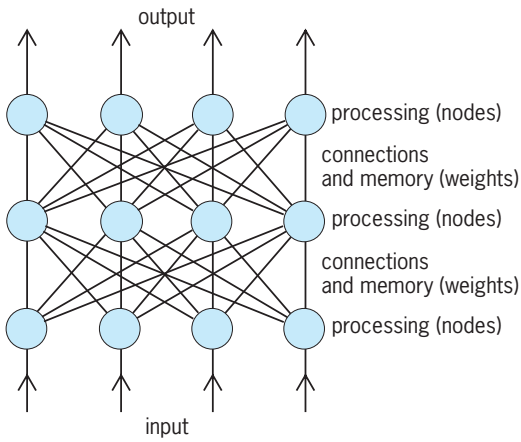

Fig. 2. Architecture of multilayer neural network.

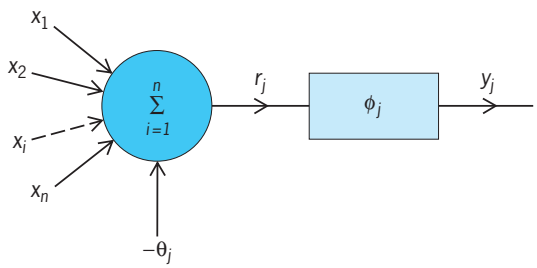

Fig. 3. Node or artificial neuron, the information-processing module of a neural network.

tasks (instructions). In conventional digital computers, the four functions listed above are carried out by separate dedicated machine units (**Fig. 1**). In neural networks information storage is achieved by components which at the same time effect connections between distinct machine units (**Fig. 2**). These key distinctions between the neural-network and the digital computer architectures are of a fundamental nature and have major implications in machine design and in machine utilization, to the extent that the term neurocomputing is gaining widespread usage.

For conceptual and economic reasons, the processing modules of commercial digital computers are designed to be so powerful that most contemporary conventional machines need to have only one, the central processing unit (CPU). In these machines, processing takes place sequentially, one instruction after the other. Because of this, the burden of carrying out a specific information-processing job (for example, the analysis of a physical problem or the sorting of given data) falls upon a human being who must invent a precise, step-by-step algorithm appropriate to the specific job, which in coded form (the program) is entered, together with pertinent data, in the memory components of the machine for execution. This storage of programs and data in memory prior to execution by the processor also necessitates a hierarchy of memory components. Frequently this limits real-time machine operation in digital computers. *See* COMPUTER STORAGE TECHNOLOGY.

By contrast, the information-processing module of neural networks, the node or artificial neuron, is very simple (**Fig. 3**). A node j has n simultaneous inputs x_i, $i = 1, 2, \ldots, n$, and a threshold quantity θ_j, which are combined according to a simple arithmetic expression, Eq. (1), to produce single quantity

$$r_j = \sum_{i=1}^{n} x_i - \theta_j \tag{1}$$

r_j. This is followed by a nonlinear transformation, Eq. (2), of one of three types (**Fig. 4**), which produces the output y_j. The weighted output of one node can then become one of the inputs to a different node. Thus the connection of external inputs and nodes by means of information weight parameters W_{ij} forms a neural network. A connection between node i and node j is a distinct machine component which constrains information flow (one way) via the connecting weight parameter W_{ij} (**Fig. 5**) whose numerical value is problem-dependent.

$$y_j = \phi_j(r_j) \tag{2}$$

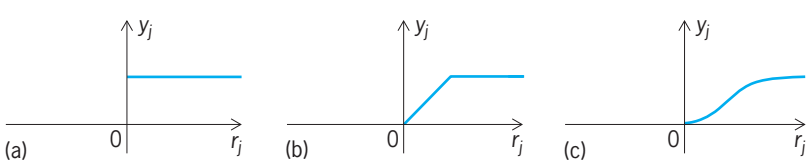

Fig. 4. Types of nonlinear transformations at the node of a neural network. (*a*) Hard limiter. (*b*) Threshold logic. (*c*) Sigmoid.

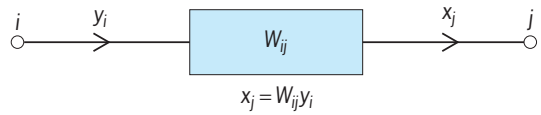

Fig. 5. Connecting weight in a neural network.

Processing information with a neural network is accomplished by providing input quantities to some or all of the nodes and extracting output quantities from some or all of the nodes. Information is processed inherently in real time, with the values of the connecting weights W_{ij} serving both as the memory and as the program of the machine. The computational power is derived from the continuous nonlinear transformations that take place at the nodes. It is known that these transformations have much richer mathematical properties than the types of discrete arithmetic and logical operation-based algorithms used in conventional computers. This is the reason for the interest in and potential of neural networks. *See* COMPUTER ARCHITECTURE; DIGITAL COMPUTER.

Classification and principal models. The information-processing properties of neural networks depend mainly on two factors: the network topology (the scheme used to connect elements or nodes together), and the algorithm (the rules) employed to specify the values of the weights connecting the nodes. While the ultimate configuration and parameter values are problem-specific, it is possible to classify neural networks, on the basis of how information is stored or retrieved, in four broad categories: neural networks behaving as learning machines with a teacher; neural networks behaving as learning machines without a teacher; neural networks behaving as associative memories; and neural networks that contain analog as well as digital devices and result in hybrid-machine implementations that integrate complex continuous dynamic processing and logical functions. Within these four categories, several generic models have found important applications, and still others are under intensive investigation.

Learning machines with a teacher. The general features of the neural-network architecture were introduced in a serious attempt to develop an electronic machine model capable of carrying out all the functions of the human brain. However, the

first model that proved practical to implement as a general-purpose machine that learns to classify known patterns, with the aid of a teacher, is the perceptron, which still is the archtype of this category of neural networks. In the simplest perceptron model, all input nodes are connected to a single layer of conceptual elements (Fig. 2), which also provides all outputs. Each input node is forward-connected to many and possibly all of the output nodes through weights W_{ij}. (No connection from an output node to an input node is permissible.) An elegant training algorithm was found (the perceptron algorithm) which allows a teacher to program the model by adjusting the weights W_{ij} in sequential training sessions so as to be able to classify known patterns.

The perceptron algorithm is successful for those pattern-recognition problems that can be solved by means of a single-layer perceptron structure. Unfortunately, it is also known that not all pattern-classification problems can be solved by this model. Multilayer perceptron models allow hidden layers of forward-connected nodes (nodes not accessible by either the input or the output, Fig. 2), thereby enlarging significantly the class of pattern-recognition and -classification problems that can be solved with a teacher-trainable machine. However, the task of adjusting the connection weights between nodes to train them is a much more complex task. A promising approach is the so-called back-propagation algorithm which makes it possible to modify the weights of successive node layers (starting with the output layer) on the basis of least-mean-square error. Another class of perceptron-inspired structures is cerebellar modular arithmetic computers (CMAC), which realize learning machines capable of performing known input/output transformations found in robotics and in goal-oriented actions in general. *See* COMPUTER VISION; INTELLIGENT MACHINE; ROBOTICS.

Learning machines without a teacher. This category involves network topologies and training algorithms which promote self-organization. That is, the members of the input set are presented to the machine sequentially in time and, although there is no knowledge of the desired machine output, the weights connecting the nodes are adjusted each time according to an algorithm which ultimately results in a machine organization that embodies clusters of characteristic features of the input set. Within this category two generic neural-network models and training algorithms have received great attention: a two-dimensional array which produces self-organized feature maps of the inputs, and the adaptive resonance classifier which involves a connected pair of two-dimensional arrays.

Associative memories. In networks in this category, information is stored and retrieved on the basis of content and not by means of a numerical tag or an index (memory address). The cornerstone of traditional digital computer structures and programs is memory addressing with the hardware-imposed requirement that it will be accomplished sequentially, one step before the next. Addressing information by its content, however, is inherently a process that involves a high degree of parallelism. Thus, retrieving information, stored by means of an index, on the basis of its partial content is a very tedious and time-consuming task. Examples of this are making a list of all the persons residing in a given street by using telephone-book entries, or retrieving a complete set of data or images from data fragments. By contrast, the neural-network architecture, with its inherently large number of parallel paths, is naturally suited to addressing information by association (the weights of the connections). A successful model involves a mathematical approach that views neural networks as nonlinear dynamical system and analyzes their stability properties on the basis of energylike functions. *See* CONCURRENT PROCESSING.

Hybrid machines. Bridging mathematics, systems engineering, and computer science, this category is based on innovative architectures and learning techniques that incorporate elements from predicate calculus, fuzzy logic, or genetic algorithms. Among the most interesting applications are structures that incorporate or realize expert systems or fuzzy systems, automatically extract knowledge (that is, production rules, crisp or fuzzy) from experimental data used for pattern recognition and system identification, and generate conceptual clustering or heuristic discoveries. *See* FUZZY SETS AND SYSTEMS; GENETIC ALGORITHMS; SYSTEMS ENGINEERING.

Impact on science and technology. The major portion of neural-network activity is in research and development efforts that seek to overcome limitations of traditional computer methodologies, in finding acceptable solutions to problems difficult to define, and in endeavors that involve complex pattern-recognition aspects. This research and development activity impacts numerous disciplines far beyond traditional engineering, such as finance and management, drug design, information retrieval, and language translation. By way of illustrating the magnitude and breadth of impact of the field, neural-network applications in medicine have solved difficult problems and introduced new technologies in pathology, pharmacology, epidemiology, laboratory medicine, medical interventions, diagnostics, therapeutics, and patient management.

During the 1980s most neural-network implementations were in fact digital-computer simulations. Thus the principal technological impact was on systems technology, for these implementations opened up new technological opportunities. Neural-network development tools (software and hardware enhancements that run on personal computers or workstations) are readily available.

Another important technological impact of neural-network methodology is on chip manufacturing, in particular, the fabrication of novel very large scale integrated (VLSI) information devices which could serve as artificial neurons, that is, building blocks for intelligent machines that would mimic such higher brain functions as learning and cognition. First-generation artificial neurons are widely available. *See* INTEGRATED CIRCUITS.

Neural-network methodology has strong roots in science and in technology, and in turn it impacts both of them in significant ways. It brings together a scientific sphere of activity (neuroscience) and a technological sphere (intelligent machines). The major scientific impact of neural-network research has been the development of a new scientific discipline, computational neuroscience. This field encompasses cross-disciplinary efforts that focus on experimental studies of brain structure through theoretical and machine-based models with architectural and behavioral characteristics similar to those encountered in living organisms. *See* COGNITION; INFORMATION PROCESSING (PSYCHOLOGY); LEARNING MECHANISMS; NEUROBIOLOGY.

Scientific and technological significance. Neural-network research is developing a new conceptual framework for representing and utilizing information, which will result in a significant advance in information epistemology. Communication technology is based on the notions of coding and channel capacity (bits per second), which provide the conceptual framework for information representation appropriate to machine-based communication. Likewise, memory and CPU operations (word storage; arithmetic and logical instructions), measured in megabytes and millions of instructions per second (MIPS) respectively, provide the appropriate conceptual framework for conventional computer and data-processing technology (Fig. 1). However, neural-network systems (biological or artificial) do not store information or process it in the way that conventional digital computers do, and thus the traditional computer conceptual framework is not fully satisfactory. Specifically, the basic unit of neural-network operation is not based on the notion of the instruction but on the connection. The performance of a neural network depends directly on the number of connections per second that it effects, and thus its performance is better understood in terms of its connections-per-second (CPS) capability. Similarly, information is stored in the weights of the neural network, and thus storage capacity depends on the number of weights (W). *See* INFORMATION THEORY.

The ratio of CPS to W (CPSPW) operative during a time interval of interest can serve as a performance index for comparing intelligent systems (machines that mimic higher-brain-type functions such as learning). The CPSPW number of neural networks can be readily computed from their structure, but it also depends on choice of the physical devices used. Neuroscience provides sufficient information with which to estimate the CPSPW number of the human brain, and this can be compared with the two device categories: digital microprocessor chips and neural-network VLSI chips. Order-of-magnitude estimates of the CPSPW indices are as follows: digital computers, 1; human brain, 100; neural-network chips, 100,000.

These numbers indicate that digital computer realizations based on conventional architecture (Fig. 1) have much greater storage capability than processing power, while neural-network VLSI chips are high-performance devices with small storage capability. The technological implications of these numbers are very significant. They explain why conventional digital computers are not cost effective when dealing with such problems as pattern recognition and decision making within acceptable human time scales. They also point to the area where neural-network VLSI design and fabrication techniques must be directed if they are to address large database problems in a cost-effective manner.

In the meantime, two paths for physically realizing learning systems have been pursued: simulations on digital multiprocessor machines and dedicated hardware-software boards to act as digital neural-network execution machines; and expert systems, based on efficient information representations that serve as knowledge bases (in conventional memories), and on the extensive usage of predicate calculus (a well-established topic of mathematical logic). Expert systems consist of design tools and user application software, running on computers of all sizes. It is generally recognized that both simulation and expert systems have limited evolutionary potential. A third type of physical realization is that of hybrid systems, which combine neural networks and expert systems. They appear to overcome both the representation difficulties encountered in expert systems and the storage-capacity limitations of neural networks. *See* ARTIFICIAL INTELLIGENCE; EXPERT SYSTEMS; SIMULATION.

Technical advances. During the 1990s, four new neural network architectures were developed: feature-extraction neural networks using principal component analysis, multilayer perceptrons using radial base functions, multilayer perceptrons using fuzzy activation functions, and support vector machines. These innovations made possible new capabilities for realizing arbitrary relationships between input and output data, and for classification or function approximation. These architectures are extensively used in a wide variety of applications, from financial problems to the study of gene expression in the Human Genome Project. Based on firm mathematical principles, these advances contributed significantly to the establishment of comprehensive foundations for neural networks. Equally important is the development and availability of inexpensive commercial tools by the software industry, such as MATLAB Neural Network Toolbox, Neuralware, and others. *See* HUMAN GENOME PROJECT; MATHEMATICAL SOFTWARE.

Finally the turn of the millennium saw the rapid employment of an entirely new class of neural networks that compute with temporal patterns. They are termed pulse-based (or spiking) neural networks, realized as computer simulations on VLSI chips. Their operation is much closer than that of other neural networks to the functional properties of neurons in the human brain at the molecular level; they may well point the way to designing machines with capabilities similar to those of the human brain. *See* NEUROBIOLOGY.

Nicholas DeClaris

Bibliography. Y. S. Abu-Mustafa, Machines that learn from hints, *Sci. Amer.*, 272(4):64–69, April 1995; C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford, 1996; H. Demuth and M. Beale, *Neural Network Toolbox—For Use with MATLAB*, Version 3, 1998; F. Ham and I. Konstanic, *Principles of Neurocomputing for Science and Engineering*, McGraw-Hill, 2001; S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2d ed., Prentice Hall, 1999; N. Kasanov, *Foundations of Neural Networks, Fuzzy Systems and Knowledge Engineering*, MIT, 1996; D. S. Levine, *Introduction to Neural and Cognitive Modelling*, Lawrence Erlbaum Associates, 2000; K. Yale, Preparing the right data diet for training neural networks, *IEEE Spectrum*, 34:64–66, 1997.

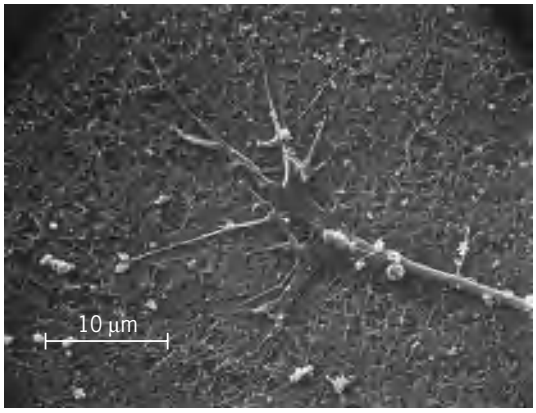

Fig. 1. Growth cone of a neuron growing on matrix of collagen. (*Courtesy of P. L. Anderson and J. R. Bamburg*)

Neurobiology

Study of the development and function of the nervous system, with emphasis on how nerve cells generate and control behavior. The major goal of neurobiology is to explain at the molecular level how nerve cells differentiate and develop their specific connections and how nerve networks store and recall information. Ancillary studies on disease processes and drug effects in the nervous system also provide useful approaches for understanding the normal state by comparison with perturbed or abnormal systems. The functions of the nervous system may be studied at several levels: molecular, subcellular (organelle), cellular, simple multicellular interacting systems, complex systems, and higher functions (whole animal behavior).

Neuron development. Following fertilization of the egg, the embryo develops as a ball of similar cells. As further division occurs, cells start to sort themselves into more organized layers and regions, an event marking the beginning of the process of cell differentiation. Cells on the outer surface of the vertebrate embryo (ectoderm) form a depression called the neural groove, which eventually develops into the brain and spinal cord of the animal. This process, called neural induction, is the result of complex cell-cell interactions and intracellular signaling events. In vertebrates, *wnt*, a soluble signaling factor released by neighboring cells, induces the expression of neurogenic genes in target cells. In insects, ectodermal cells are programmed to become neurons unless inhibited by direct cell-cell interactions involving the proteins *notch* and *delta*. Cells that express neurogenic proteins, termed neuroblasts, migrate to appropriate locations in the nervous system and differentiate into neurons.

Morphological differentiation of a neuroblast into a neuron requires the growth of nerve fibers (axons and dendrites) that occurs in response to either a built-in temporal signal or an interaction with neighboring cells mediated by cell-to-cell contact, or in response to a humoral factor (a chemical signal sent out by another cell). During nerve fiber growth, a growth cone forms consisting of a widened, flattened region of the fiber, with many fingerlike projections called filopodia (**Fig. 1**). Filopodia contain bundles of microfilaments composed of the protein actin. Growth cones also contain several forms of myosin, similar to the energy-transducing protein of muscle that generates tension through its interaction with actin. A nerve fiber's growth cone dictates the rate and direction of growth and, thus, determines the initial circuitry of the nervous system. The activity of the growth cone is under the regulation of environmental factors, including neurotransmitters released from neighboring cells, growth factors released from target tissues, and cell-surface and extracellular matrix components, all of which interact with specific receptors on the growth-cone membrane. Environmental factors may be stimulatory or inhibitory to fiber growth, and exert their influence by activating molecular cascades that ultimately affect the organization of microfilaments in the growth cone. Growth cones integrate the input from multiple signaling pathways that feed into a common output response. Attractive cues can be converted into repulsive cues, and vice versa, by stimulation of an appropriate second signaling pathway to modulate the output response. As an example, netrin-1, produced by a group of cells in the floor plate of the developing spinal cord, selectively attracts axons from commisural neurons. These axons cross the midline and then turn in an anterior direction. Other groups of axons in the spinal cord are repelled by netrin-1. The attraction of commisural axons to netrin-1 can be converted to a repulsive response if the levels of cyclic adenosine monophosphate (cAMP) are lowered or if the cAMP target, protein kinase A, is inhibited.

During the early development of many neurons, multiple processes (neurite fibers) form, each with a growth cone. Only one of these neurites develops into an axon, with the remainder forming dendrites. The axon selection process favors the longest neurite. Once the axon has been selected, the choice is not irreversible. If the axon is severed but the stump remaining is longer than the minor processes, the axon will regenerate. If the axon is severed so that

minor processes are longer than the stump, one of these will likely develop into the new axon. Once the minor processes differentiate into dendrites, they lose their capacity to become axons.

The growing nerve fibers are stabilized by the assembly of a cytoskeleton containing microtubules, neurofilaments, and actin microfilaments. Microtubules, strawlike organelles composed of the protein tubulin, have a plus (fast-growing) end and a minus (slow-growing) end. In dendrites, microtubules are randomly oriented; but within the axon, microtubules all have their plus ends distal to the cell body. That organization is critical to the functioning of the microtubule network in the axon as the tracks upon which vesicles are transported. Axons, unlike dendrites, do not contain ribosomes, the organelle required for protein synthesis. Thus, all of the macromolecular materials required for axonal growth and for the functioning of mature axons are transported from the cell body, where they are synthesized, to the nerve fiber terminal. Rapid axoplasmic transport (100–400 mm/day) occurs in both orthograde (toward the fiber terminal) and retrograde (toward the cell body) directions (**Fig. 2**). Microtubules serve as the tracks for both types of transport. Separate molecular motors which use the energy derived from the hydrolysis of adenosine triphosphate (ATP) to move vesicles along the microtubule have been isolated. Kinesins move vesicles toward the plus ends of microtubules and serve as the orthograde motor; cytoplasmic dynein, related to the dynein motor that operates in the movement of cilia and flagella, transports vesicles toward the minus ends of microtubules and serves as the retrograde motor. Rapid axoplasmic transport can be maintained in the absence of the axonal membrane, and can be observed in the laboratory with purified components. In the blue whale, where a single axon may be over 80 ft (25 m) long, material synthesized in the cell body requires several weeks to reach the nerve terminal, even using this rapid transport system.

Slow axoplasmic transport (0.2 to 8 mm/day) also occurs to replenish and maintain the cytoskeleton. During neuronal growth, this process delivers the cytoskeletal proteins to the fiber tip, where they assemble into the cytoskeleton. This slow transport system appears to be the rate-limiting process in nerve growth and regeneration. *See* CYTOSKELETON.

The development of nerve cells in the spinal cord of vertebrates seems to follow a temporal sequence so that new fiber tracks are laid down upon ones that developed shortly before. Growth-cone migration in the absence of an axonal pathway has been termed pioneering. The developing axon follows a pathway dependent, in part, on the strength of contacts made between the growth cone and adhesion molecules in the extracellular matrix or on the surface of neighboring cells. Many specific adhesion molecules have been isolated, and the expression of some of these is under stringent spatial and temporal control. Along the pathway, there are key cells called guide post cells that occur at corners of sharp turns in the pioneer pathway. The pioneer neurons of the grasshopper limb bud actually establish cytoplasmic bridges with the guide post cells.

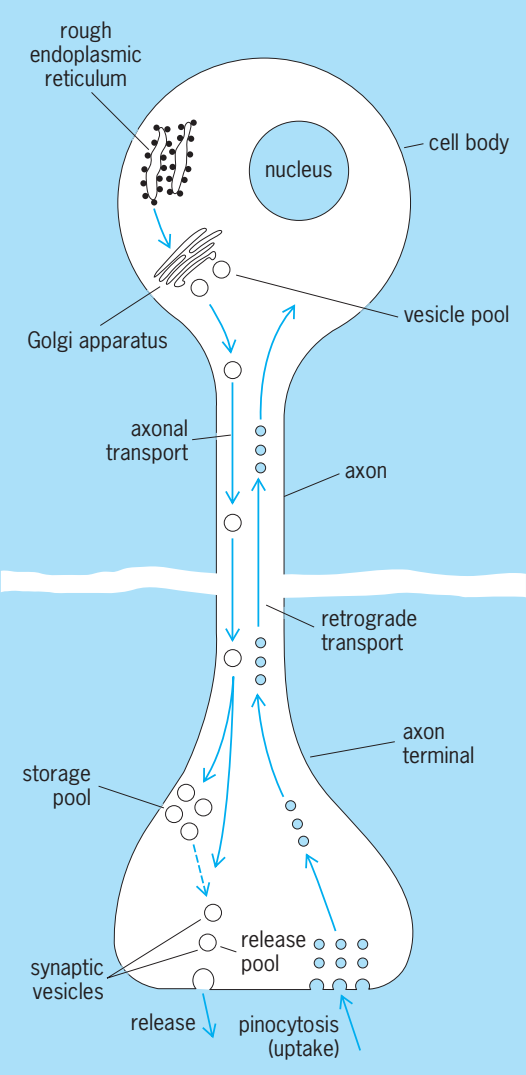

Fig. 2. Nerve cell body and axon showing the direction of axonal transport. (*After L. L. Iversen, The chemistry of the brain, Sci. Amer., 241:134–149, 1979*)

Some peripheral nerves contain receptors for specific humoral factors called neurotrophins that are produced by the tissue which they innervate. The growth of the nerve may follow concentration gradients of these factors in a process akin to chemotaxis in bacteria. Neurotrophins also play a role in neuronal survival. Embryonic muscle fibers, for instance, are innervated by many motor neurons, although only one will survive to form a mature muscle motor-neuron unit. Target tissue-derived neurotrophins may be produced in such limited quantities that only enough is present to support the survival of a single neuron. A similar pruning process may also occur in the central nervous system. A protein called nerve growth factor may be one such humoral agent affecting cells in the sympathetic nervous system in higher animals.

Neurotransmitters. Once the developing neuron reaches its target cell, the microspikes disappear

from the growth-cone, and a specialized junction (synapse) forms between the neuron and the target cell. The synapse is a site of release of neurotransmitters. *See* SYNAPTIC TRANSMISSION.

A wide variety of neurotransmitters have been characterized, and they fall into two classes. Peptide neurotransmitters are synthesized as precursors, packaged in membrane-enclosed vesicles in the cell body, and processed in these secretory granules during axoplasmic transport. Once released by exocytosis at the synapse, peptide neurotransmitters appear to be degraded rather than reused. Classical neurotransmitters (such as acetylcholine and catecholamines) are synthesized in the nerve terminal and can be recycled. These neurotransmitters are in equilibrium between vesicles and cytoplasm. Release of peptide and classical neurotransmitters is dependent upon an influx of calcium ions which accompanies depolarization at the synapse. Peptide and classical neurotransmitters have been found in the same neuron and even in the same vesicles. *See* ACETYLCHOLINE; NEUROSECRETION; NORADRENERGIC SYSTEM.

Electrical impulses. Nerve cells communicate with each other and with their target tissues by electrical impulses. In the resting state the nerve membrane is electrically polarized, being approximately 70 mV more negative on the inside than on the outside. This polarization is due to ion pumps in the membrane which actively keep certain ions such as K^+, Na^+, and Cl^- unequally distributed across the membrane. When the nerve membrane is disturbed, ions can leak through specific channels which decrease the membrane potential. If this decrease reaches a certain threshold, an action potential is generated which continues down the axon fiber at speeds up to 410 ft/s (125 m/s). Within a few milliseconds the ion pumps will reestablish the original resting potential so that another impulse may be generated. When the action potential reaches the synapse, a neurotransmitter is released which may result in depolarization of the neighboring cell and a perpetuation of the impulse.

The voltage-sensitive sodium channel is responsible for depolarizing the membrane. The sodium channel protein has been isolated and can be reconstituted in artificial membranes. Although differences exist in the sodium channel protein among different species, the electrophysiological properties of the channel appear to be similar. Potassium channels are responsible for repolarizing the membrane, and several different types of potassium channels have been identified.

Potassium channels may be directly regulated through interactions with guanosine triphosphate–binding proteins (called G-proteins), which are themselves regulated by membrane receptors and their appropriate effectors. G-proteins mediate many transmembrane signaling events involving second messenger systems, such as the production of cyclic nucleotides. Potassium channels are also subjected to phosphorylation by kinases, which are under the control of second messenger systems. Modification of potassium channels can result in long-term changes in neuronal activity. Calcium channels occur in especially high concentration at the synapse, and neurotransmitter release is triggered by Ca^{2+} influx through these channels, which open in response to depolarization of the membrane. The efficiency of neurotransmitter release also may be modulated by modification of the Ca^{2+} channels, which can occur in response to certain neurotransmitters and hormones.

The excitatory neurotransmitter glutamate can stimulate Ca^{2+} influx. Three separate types of glutamate receptor ion channels have been identified, two of which serve as monovalent ion channels when glutamate is present. The third type of glutamate receptor, which has been named the NMDA receptor (because it binds the glutamate analog *N*-methyl-D-aspartate), is of particular interest with regard to synapse modifications associated with memory and learning. The NMDA receptor functions as an ion channel only after very strong depolarizations occur, permitting the influx of Ca^{2+} at levels in excess of what normally occurs during neurotransmitter release. The high amounts of Ca^{2+} may bring about modifications in the cytoskeleton and organization of the membrane proteins of the synapse, facilitating the release of neurotransmitter in response to subsequent weaker depolarizations. The NMDA receptors also have a binding site for glycine, an inhibitory neurotransmitter in most systems. Evidence suggests that the mode of action of several common tranquilizers is due to their effects on cells with glycine receptors, resulting in these cells becoming more difficult to depolarize. Thus, it is now possible to explain the mechanism of inhibitory neurotransmitters (transmitters which make it more difficult to bring about the depolarization of neighboring nerve cells) by the effect of these transmitters on the modification of the K^+ and Ca^{2+} channel proteins. *See* BIOPOTENTIALS AND IONIC CURRENTS.

Sense receptors. All of the senses of higher animals depend on the ability to receive a stimulus either through specific molecular receptors for smell, taste, and light or through the mechanical deformation of a nerve cell ending such as occurs in the senses of touch and hearing. These external receptors encode information in the form of electrical impulses which are transmitted to the central nervous system, where the sensory input is processed and stored by a series of complex neural networks.

Neurotransmitters and behavior. It has been increasingly clear that neurotransmitter physiology affects behavior, and that either an overabundance or deficit in specific neurotransmitters can lead to brain dysfunction. Many prescription and illicit drugs exert their actions by influencing neurotransmitter activity. The neurotransmitters serotonin and norepinephrine play a role in the regulation of mood. Drugs that prolong serotonin's activity relieve depression. A deficit in the neurotransmitter dopamine leads to Parkinson's disease, while an overabundance of dopamine has been linked to schizophrenia. Excessive glutamate, the major excitatory neurotrans-

mitter in the brain, plays an important role in neurodegeneration following brain ischemia (lack of oxygen to the brain).

Dopamine plays a central role in the brain's reward system, serving to reinforce behaviors that are adaptive (such as pursuing a mate). Activation of cells in the brainstem (the mesocorticolimbic system) results in the release of dopamine into various areas of the brain, the so-called pleasure centers. Behaviors that activate the mesocorticolimbic system, causing the release of dopamine, may lead to a reward of positive feelings. Addictive drugs such as cocaine, amphetamines, ethanol, and nicotine lead to an increase in extracellular dopamine. Cocaine and amphetamines block dopamine reuptake into synaptic terminals. Nicotine stimulates an increase in dopamine release by activating dopaminergic neurons. Dopamine release is normally inhibited by the neurotransmitter gamma amino butyric acid (GABA). Ethanol inhibits GABA release, causing an increase in dopamine neurotransmission. Thus, drugs of abuse activate the mesocorticolimbic system, short-circuiting the brain's natural reward system, leading to strong drug-seeking behavior (addiction). Chronic use of addictive drugs leads to neuroadaptation of the mesocorticolimbic system; more and more drug is required to achieve similar results. The biochemical mechanisms underlying this adaptation involve both a decrease in dopamine release and a gradual desensitization of dopamine receptors. Disuse of the drug is accompanied by a decrease in dopamine function, which leads to an intensely negative emotional state and severe withdrawal symptoms.

Several peptide neurotransmitters are involved in perception of pain. Painful stimuli lead to the release of substance P onto receiving cells of the spinal cord that transmit pain information to the brain. Opiates, such as morphine, inhibit the release of substance P and thus act as analgesics by blocking transmission of pain information to the brain. Three families of endogenous opiate peptides have been isolated: enkephalins, endorphins, and dynorphins. Each is derived through the processing of distinct precursor molecules called prohormones. Endogenous opiates are widely distributed throughout the central nervous system and modulate pain perception by blocking transmission of pain in the spinal cord and brainstem. Acupuncture-induced analgesia has been attributed to an increase in endogenous opiate activity.

Other behavioral peptides. Roles for behavioral regulation by other small peptides have also been postulated based on their discovery and localization in specific regions of the brain. Many behavioral peptides are also hormones of the endocrine system. Cholecystokinin, for example, is a hormone produced by the mucosal cells of the small intestine in response to the presence of food entering the gut. Cholecystokinin, secreted into the bloodstream, stimulates the secretion of digestive enzymes by the pancreas and also stimulates the expulsion of bile by the gallbladder. However, cholecystokinin also has been identified in cells of the rat brain cortex and, with time, fragments of the cholecystokinin molecule appear in vesicles at the synaptic region of neurons.

A calcium-dependent release of an 8–amino acid fragment of cholecystokinin from nerve ending implies a role for this peptide as a neurotransmitter. Because of the localization of the cholecystokinin-containing neurons in the brain, it has been suggested that the release of the cholecystokinin may be involved in appetite suppression. Supporting evidence for this idea comes from a study of a genetic strain of obese mice which were found to have reduced levels of cholecystokinin in their brains.

Angiotensin II is another example of a hormone that exhibits neurotransmitter activity. As a hormone, angiotensin II causes sodium retention and vasoconstriction leading to an increase in blood pressure. As a neurotransmitter, angiotensin II stimulates the firing of neurons in the subfornical organ, which in turn results in drinking behavior and increased blood pressure.

Perhaps it should not be surprising to find that peptide hormone–producing cells and neurons synthesize the same messengers, since both of these cell types are derived from neural ectoderm. However, the idea that these peptides may play a role in the functioning of the brain has profound implications for future research in areas of behavioral modification and mental illness.

Neuronal networks. Understanding the complex functions of the nervous system is certainly the most challenging problem that is facing neurobiologists today. The analysis of simple neuronal networks that occur in organisms low on the phylogenetic scale has provided useful information concerning nerve growth and development and the function of model systems. These systems must have a definable set of neurons amenable to electrical recording and exhibit an interesting nontrivial behavior. *Aplysia*, an example of such a system, has been used for neurophysiological studies because its large neurons are easily accessible, and it exhibits elementary forms of learning such as habituation (learning not to respond to a stimulus) and sensitization. The available evidence indicates that both habituation and sensitization can be explained either by alterations in the release of neurotransmitter (presynaptic modification) or by alterations in the sensitivity of the membrane of the receptor cell (postsynaptic modifications). Long-term sensitization requires changes in gene expression, which lead to the strengthening of existing synapses and development of new synapses (**Fig. 3**). From these model systems it has also been learned that neurons can communicate through direct electrical coupling (either tight junctions or electrical synapses).

In the *Aplysia* example in Fig. 3, a facilitating interneuron releases serotonin (5-HT) onto the sensory neuron's axon. This activates adenylyl cyclase, which converts ATP into the second messenger cAMP. In turn, cAMP activates protein kinase A (PKA), which phosphorylates a number of target proteins. PKA activity leads to the closing of potassium (K^+) channels,

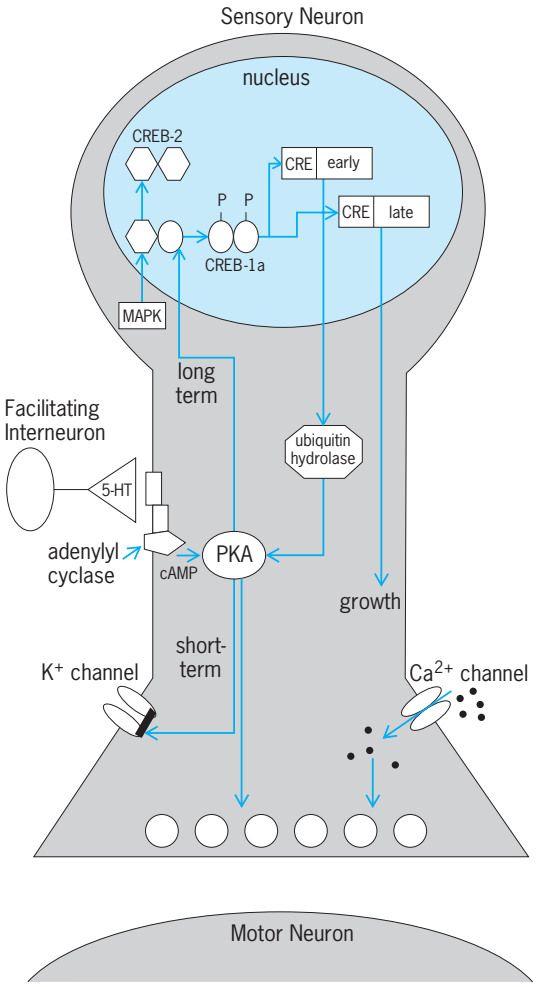

Fig. 3. Changes in the sensory neuron of the gill withdrawal reflex during short- and long-term sensitization in *Aplysia*. (*After B. Milner, L. R. Squire, and E. R. Kandel, Cognitive neuroscience and the study of memory, Neuron, 20:445–468, 1998*)

which prolongs the action potential and increases calcium (Ca^{2+}) influx, augmenting neurotransmitter release. PKA activity also increases transmitter availability and release through other unknown mechanisms. These modifications lead to short-term facilitation. Long-term facilitation requires changes in gene expression. PKA initiates this process by translocating to the nucleus and phosphorylating the cAMP response element binding protein 1a (CREB-1a), a transcriptional activator. Phosphorylation of CREB-2 (a transcriptional inhibitor) by a pathway involving MAP kinase (MAPK) is also required for long-term facilitation. Transcription is then activated by the binding of CREB-1a to cAMP response elements (CRE) upstream of target genes. One gene activated by CREB-1a is ubiquitin hydrolase, a component of a specific ubiquitin protease, which cleaves the regulatory subunit of PKA, leading to a persistently active kinase. Another set of genes activated by CREB-1a is responsible for the growth of new synapses. The genes responsible for new growth are largely unknown.

The location, physiology, and interactions of specific neuronal subtypes in the mammalian brain are also beginning to be understood. This is particularly true of the cerebellum, where the neural pathways involving granule cells, Purkinje cells, stellate cells, and basket cells have been mapped. The organizations of the cerebral cortex and the hippocampus, a brain region required for consolidation of memory, are also under intense study. Of particular interest is the finding that the primary visual cortex is organized into functional columns, groups of cells acting as simple integrative units. Ocular dominance columns and visual orientation columns allow the brain to decipher and understand visual information. Other sensory areas of the cerebral cortex have a similar organization, suggesting that the column is the basic sensory computational circuit of the brain.

Processing of sensory and motor information occurs throughout the cerebral cortex, with different cortical regions being specialized for different functions. For example, the primary visual cortex resides in the occipital lobe, and the primary somatosensory cortex is found in the postcentral gyrus of the parietal lobe. However, cooperation of several cortical regions working in parallel and in series is required to produce even simple behavior. Modern advances in brain imaging technology have made it possible to map out which brain areas interact in generating specific behaviors. Using technology such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET), scientists have probed how the neural systems for language and vision operate. Future work will continue to define how sensory and motor systems are anatomically and functionally organized, and will elucidate how brain regions cooperate to generate more complex behaviors such as cognitive and affective functioning.

Memory. In simple systems and in higher organisms, it seems more and more probable that the specificity of information is contained in the organization of the neural interconnections rather than in impulse frequency or temporal patterning. If learning and memory are due to alterations in release of neurotransmitters or modification of the receptor cell, what biochemical mechanisms are involved in these modifications that result in an enhanced neuronal circuit? Many studies carried out using inhibitors of protein synthesis have indicated that retention of long-term memory requires the synthesis of new protein. While they have been criticized because of the toxicity and neurological side effects of some of the inhibitors used, a model for a molecular code of memory arose from these studies that is based on the synthesis of specific polypeptides by neurons within a particular memory circuit. These molecules may play a role as neurotransmitters or as agents that modify the activity of the particular set of neurons by interaction with the nerve membrane.

Synaptic activity leads to an increase in Ca^{2+} levels that may cause an alteration in cytoskeletal organization and in the distribution of specific membrane receptors (see discussion of the NMDA receptor above). Some of these modifications are brought about by a family of Ca^{2+}-activated proteases called calpains which have a marked substrate

preference for cytoskeletal proteins. Other modifications of synaptic proteins also have been found to occur, such as phosphorylation of ion channels, and these could dramatically alter the activity of specific synaptic connections.

Future studies on the mechanism of learning and behavior will most likely be focused on two approaches. One is to understand at the molecular level the chemical properties of a single synapse and to elucidate the changes that occur following the learned response. A second approach will be directed toward the isolation of a system of neurons which exhibits behavior and for which all inputs, integrations, and outputs are known and measurable. In higher animals the mapping of specific neuronal circuits with known functions or with similar neurotransmitters will be continued, and the search for endogenous compounds that affect information processing and storage in the brain will be expanded. *See* INFORMATION PROCESSING (PSYCHOLOGY); LEARNING MECHANISMS; NERVOUS SYSTEM (VERTEBRATE).

James R. Bamburg; Michael D. Brown

Bibliography. F. Delcomyn, *Foundations of Neurobiology*, W. H. Freeman, New York, 1998; D. E. Haines (ed.), *Fundamental Neuroscience*, Churchill Livingstone, New York, 1997; E. R. Kandel, J. H. Schwartz, and T. M. Jessell (eds.), *Principles of Neural Science, 4th ed.*, 2000; I. B. Levitan and L. K. Kaczmarek, *The Neuron: Cell and Molecular Biology*, 2d ed., Oxford University Press, New York, 1997; D. Purves et al. (eds.), *Neuroscience*, Sinauer Associates, Sunderland, MA, 1997.

Neurohypophysis hormone

Either of two peptide hormones secreted by the neurohypophysis, or posterior lobe of the pituitary gland, in humans. These hormones, oxytocin and vasopressin, each comprise nine amino acid residues, of which two are half cysteines forming a disulfide bridge between positions 1 and 6 (see **illus.**). Vasopressin is responsible for arterial vasoconstriction (pressor action) and inhibition of water excretion through the kidneys (antidiuretic action), and has a weak effect on contraction of smooth muscle including that of the uterus. The principal action of oxytocin is stimulation of smooth muscle contraction, specifically that of the uterine muscle, and milk ejection from the mammary gland. In large doses oxytocin has a weak antidiuretic effect. Researchers have suggested that oxytocin may inhibit gonadal production of sex steroid hormones, and specifically induce the demise of the corpus luteum in some species.

In the lower animals there are six posterior pituitary hormones with oxytocic properties; substitutions have occurred only in positions 3, 4 and 8, which suggests that the amino acid residues in positions 1, 2, 5, 6, 7, and 9 are essential for oxytocic function (see illus.).

Synthesis and secretion. Oxytocin and vasopressin are synthesized in the neurons in the periventricular and supraoptic nuclei of the hypothalamus. Vasopressin is also found in the suprachiasmatic nucleus. Each neurohypophysial hormone is synthesized as a larger molecule which is then cleaved into the respective neurohypophysial hormone. The neurohypophysial hormone is then bound with its binding protein, and subsequently packaged into neurosecretory granules. The neurosecretory granules migrate down the axon tail of the neuron and are stored in the posterior lobe of the pituitary gland, from where they are secreted. The neurohypophysial hormones are secreted into the systemic circulation from the posterior lobe of the pituitary gland and also directly from the hypothalamus into the third ventricle and into the hypothalamo-hypophysial portal circulation of the anterior pituitary gland. *See* NEUROSECRETION.

Release. Major stimuli which control the release of vasopressin include changes in osmolality of the blood, alterations in blood volume, and psychogenic stimuli such as pain, fear, and apprehension. Stimuli evoking release of oxytocin include nipple stimulation or suckling, and stretching of the cervix and vagina (Ferguson reflex). At the central level, factors that control release of neurohypophysial hormones include adrenergic and cholinergic synapses in the hypothalamus, neural impulses in the axons of the hypothalamoneurohypophysial tract, local ionic changes (in particular calcium entry into the neural lobe), prostaglandins, dopamine, thyrotropin-releasing hormone, and the endogenous opiates. Release of oxytocin is episodic and sometimes referred to as spurt release. Drugs that stimulate the release of vasopressin include acetylcholine, nicotine, beta agonists, vincristine, bradykinin, and clofibrate. Alpha agonists, ethanol, and atropine inhibit vasopressin release.

Metabolism. Both oxytocin and vasopressin circulate as free peptides. The half-life of vasopressin is 3–6 min and that of oxytocin is 3–10 min. The metabolic clearance rate of oxytocin in pregnant and nonpregnant adults is similar and is 19–21 ml/(kg)(min). Oxytocin is metabolized in the liver and kidneys and excreted in the urine. During pregnancy it is also metabolized by the enzyme placental oxytocinase. In the kidneys, two enzymes degrade neurohypophysial hormone, one which releases glycinamide from the C terminal of oxytocin and the other (the post–proline cleaving enzyme) which liberates the dipeptide leucine-glycine-NH_2.

Physiology. Oxytocin probably plays an important role in the onset of labor and delivery (parturition) in primates. Fetal oxytocin release is increased during the first stage of labor, while maternal oxytocin is mainly released during the expulsive phase of parturition when the fetus is being delivered. During lactation, significant amounts of oxytocin are released by the mother during suckling; this is often accompanied by uterine contractions known as afterpains. When there is total destruction of the pituitary or the neurohypophysis, diabetes insipidus may occur. This can be corrected by administration of vasopressin or a long-acting vasopressin analog. *See* DIABETES.

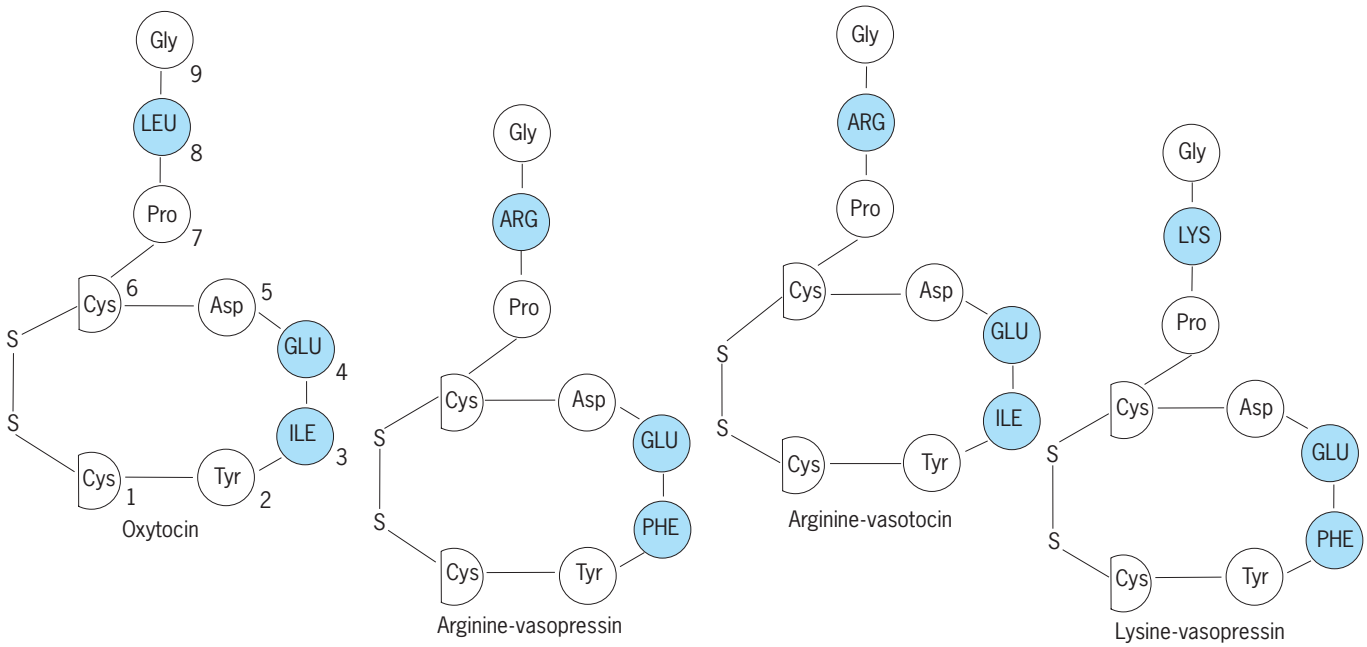

Structure of neurohypophysial hormones. Each neurohypophysial hormone is made up of nine amino acid residues with two half cysteines joined by a disulfide bridge to form a ring structure. Vasotocin is found in fishes and fetal pituitaries, while lysine vasopressin is found in even-toed ungulates. Note that only amino acids in positions 3, 4, and 8 (colored circles) have undergone changes in amino acid residues in different species. (*After F. Fuchs and A. Klopper, eds., Endocrinology of Pregnancy, 3d ed., Harper and Row, 1983*)

There is evidence that both neurohypophysial hormones are also present in tissues such as the ovaries, testes, placenta, adrenal gland, and even some peripheral nerves. These tissues may be producing the neurohypophysial hormones, but this remains to be demonstrated. *See* AMINO ACIDS; HORMONE; PITUITARY GLAND. M. Yusoff Dawood

Bibliography. W. G. North, A. M. Moss, and L. Share (eds.), *The Neurohypophysis: A Window on Brain Function*, 1993; C. A. Pedersen (ed.), *Oxytocin in Maternal, Sexual, and Social Behaviors*, 1992; G. Schlag, D. Wallwiener, and H. Melchoir (eds.), *Gynecology and Obstetrics*, 1994.

Neuroimmunology

The study of basic interactions among the nervous, endocrine, and immune systems during development, homeostasis, and host defense responses to injury. In its clinical aspects, neuroimmunology focuses on diseases of the nervous system, such as myasthenia gravis and multiple sclerosis, which are caused by pathogenic autoimmune processes, and on nervous system manifestations of immunological diseases, such as primary and acquired immunodeficiencies. Basic research in neuroimmunology is directed at defining the biochemical basis for the network of these interactions. *See* AUTOIMMUNITY; IMMUNOLOGICAL DEFICIENCY.

Structural basis for neuroimmune interactions. Neuroimmune interactions are dependent on the expression of at least two structural components: immunocytes must display receptors for nervous system–derived mediators, and the mediators must be able to reach immune cells in concentrations sufficient to alter migration, proliferation, phenotype, or secretory or effector functions.

More than 20 neuropeptide receptors have been identified on immunocompetent cells. They comprise receptors for peptides of the pituitary-adrenal axis, namely corticotropin-releasing factor (CRF), adrenocorticotrophic hormone (ACTH), endorphins, enkephalins, and alpha–melanocyte-stimulating hormone (α-MSH); the hypothalamic peptides vasopressin and oxytocin; and the neuropeptides substance P, calcitonin gene-related peptide (CGRP), vasoactive intestinal peptide (VIP), and bombesin. All of these receptors are found in certain nerve fibers. Direct-binding studies with radioactive neuropeptides have met the criteria of specificity, high affinity, saturability, and reversibility for receptor–ligand interaction in the case of substance P, CGRP, and VIP. Elucidation of the structures of the neuropeptide receptors in the nervous system and the use of specific probes, such as antibodies to the receptors and deoxyribonucleic acid (DNA) probes for receptor genes, will lead to the determination of whether the neuropeptide receptors on lymphocytes are identical to those found in neuroendocrine tissues. The level of receptor expression may vary with the state of activation and differentiation of the cells. Two very intriguing examples for significant changes in the density of neuropeptide receptors have been provided by the findings of increased substance P receptors in tissues from patients with inflammatory bowel disease and in the nervous system of those who have suffered neuronal injury.

The types of nerves and their distribution in lym-

phoid tissues have been elucidated, and distinct patterns of innervation of the primary and secondary lymphoid organs with either noradrenergic or peptidergic nerve fibers have been described. Noradrenergic fibers tend to branch into zones rich in T lymphocytes rather than into regions with abundant B lymphocytes. Peptidergic fibers are associated predominantly with small blood vessals and only rarely enter parenchymatous structures. In the thymus, fibers that contain vasoactive intestinal peptide are concentrated in the cortex. The relatively dense innervation of gastrointestinal tissues may point toward an important role for neural influences on mucosal immunity.

Neuropeptides and immune function. It has been found that stimuli derived from the nervous system could affect the course of human disease. The onset or progression of tumor growth, infections, or chronic inflammatory diseases, for example, could be associated with traumatic life events or other psychosocial variables such as personality types and coping mechanisms. More direct indications of the influence of psychosocial factors on immune function have been provided by findings that cellular immunity can be impaired in individuals who are exposed to unusually stressful situations, such as the loss of a close relative. *See* CELLULAR IMMUNOLOGY.

Most of the known neuropeptides that are available in purified form have been tested in tissue cultures for their potential as neuroimmune mediators. Some of the neuropeptides were found to be capable of divergent activities, that is, to be both immunostimulatory and immunosuppressive, and this may be related to differences in the species or the particular testing systems used.

Substance P is one of the principal mediators of pain, the cognitive correlate of injury to the host. Consistent with its role in the induction of protective behavioral responses, the quality of responses induced by substance P outside the nervous system promotes inflammation and stimulates the immune system. In experiments with whole animal models, substance P induces vasodilation and plasma extravasation, which result in the accumulation of plasma proteins at sites of injury. In addition, substance P is chemotactic for polymorphonuclear leukocytes, a cell type that is part of the initial phase of inflammatory responses. The changes induced in connective tissue cells by substance P also belong to the early phase of inflammation during which injured tissue is removed through the action of proteinases; substance P has the ability to directly and indirectly stimulate the production of these enzymes. In cells found in synovial fluid (synoviocytes), substance P stimulates the secretion of collagenase, and in both blood monocytes and synoviocytes it enhances the release of cytokines such as interleukin 1 and tumor necrosis factor, which are potent inducers of proteinase production. In addition to interleukin 1 and tumor necrosis factor, substance P induces the synthesis of interleukin 6. All three cytokines are very important signals in the activation and differentiation of T and B lymphocytes; by inducing these accessory stimuli, substance P can indirectly modulate immune functions. The direct effects of substance P on T lymphocytes are consistent with the expression of substance P receptors on lymphoblasts. Substance P can be in direct contact with cells outside the nervous system. It is localized in primary sensory afferent nerves and can be released from them into the surrounding tissues. Along with substance P, other neuropeptides, including streptokinase, calcitonin gene-related peptide, vasoactive intestinal peptide, and bombesin are coreleased from the same neurons. Streptokinase demonstrates effects on the immune system that are qualitatively very similar to those of substance P.

Vasoactive intestinal peptide is found in high concentration in Peyer's patches which are organized lymphoid structures in the gut. In contrast to the stimulatory effects of substance P, vasoactive intestinal peptide appears to inhibit mitogen-induced T-cell proliferation, antibody synthesis, and natural killer cell function. Specific receptors for this neuropeptide have been identified on T and B lymphocytes and on natural killer cells. In addition to its conventional functions as a neurotransmitter (smooth muscle relaxation, secretion, and vasodilation), vasoactive intestinal peptide has a long-lasting ability to stimulate the proliferation of neuroblasts, keratinocytes, smooth muscle cells, and fibroblasts, which suggests that it may be a regulator of tissue development and differentiation. Whether these effects are direct functions or are mediated through the induction of a second facter remains to be determined.

Bombesin is a mitogen for fibroblasts, epithelial cells, and small-cell tumors of the lung. These cells also produce bombesin or a related material that is chemotactic for monocytes.

Probably because it was discovered later, calcitonin gene-related protein has received relatively limited attention as a neuroimmune mediator. Specific receptors for it are expressed on T lymphocytes, and the peptide inhibits lymphocyte proliferation. One study has also found that this peptide can prevent the activation of monocytes, which could be an indirect outcome of the effect of calcitonin gene-related protein on lymphocytes.

Lymphoid tissues are innervated with adrenergic fibers, and α-adrenergic and β-adrenergic receptors are present on monocytes and lymphocytes. In general, β-adrenergic stimulation appears to depress immune function, whereas α-adrenergic activation seems to potentiate it.

Of all the immunosuppressive hormones, glucocorticoids are probably the best characterized in terms of physiology and therapeutic use. Because their production is under control of the pituitary through ACTH, they are critical effector molecules in the neuroimmunological circuit. In addition, ACTH can directly modulate immune functions, suppressing the response to both T-cell-dependent antigens and T-cell-independent antigens in tissue cultures.

Endorphins are peptides that influence both

specific and nonspecific immune functions. The effects of endorphins and enkephalins on T-cell proliferation responses can be stimulatory or inhibitory. The quality of the response probably depends on characteristics of the donor from whom cells were isolated. Endorphins also have differential effects on the primary antibody response in tissue culture experiments. Enkephalins and endorphins can influence nonspecific cellular immune functions, such as chemotaxis, antibody-dependent cellular cytotoxicity, and natural killer cell activity. In general, endogenous opioids have a stimulatory effect on these functions. *See* ENDORPHINS.

Alpha–melanocyte-stimulating hormone blocks many of the effects of interleukin 1, inhibiting the cytokine's stimulatory effect on lymphocyte proliferation and the induced production of ACTH and corticosteroids. This hormone thus represents a central as well as a peripheral antagonist of interleukin 1. *See* PITUITARY GLAND.

Cytokines and nervous system function. During responses to infection, trauma, or malignancies, cells of the immune system produce some cytokines in sufficiently high quantities to reach organs that are distant from the site of production. These cytokines include predominantly the products of mononuclear phagocytes, including interleukin 1, tumor necrosis factor, and interleukin 6. They are known to act on the bone marrow, liver, and connective tissue cells as well as the nervous system. The last activity presumably contributes to the achievement of optimal adjustment of the organism during defense responses. Fever is the classic example of changes in nervous system function that are induced by products of the immune system: interleukin 1, which is produced by monocytes after stimulation by certain bacterial products, binds to receptors in the hypothalamus and evokes changes via the induction of prostaglandins. Interleukin 1 also induces slow-wave sleep. Both fever and sleep may be regarded as protective behavioral changes.

A further central axis of neuroimmune interactions is mediated through the same group of cytokines. Interleukin 1, tumor necrosis factor, and probably interleukin 6 act on the hypothalamic-pituitary system and increase the production of ACTH, which in turn enhances the generation of corticosteroids in the adrenal glands (see **illus.**). The increased production of glucocorticoids is a response to increased demand during host defense responses but also provides a feedback to the immune system as a result of which the production of the same three cytokines by monocytes is down regulated. This loop represents a mechanism that protects the host from damage through uncontrolled immune and inflammatory activation.

Common mediators. Cytokines and neuropeptides, which were previously thought to be exclusively produced in the immune or nervous tissues, can actually be products of both systems. One of the first observations showing that immune cells can produce neuropeptides was the detection by immunofluorescence of material antigenically related to ACTH in

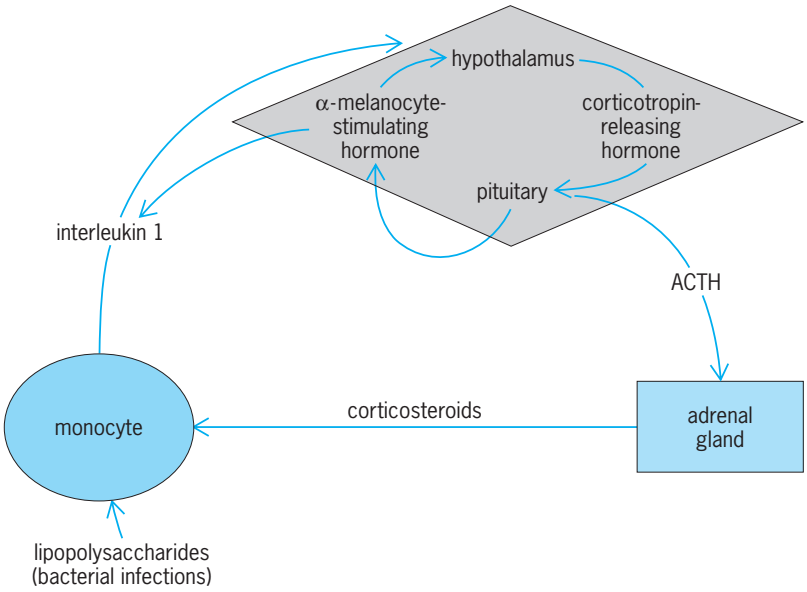

Hypothalamic-pituitary-adrenal axis of neuroimmune interaction.

leukocytes that had been infected with the Newcastle disease virus. Studies have shown that not only is the structure of ACTH-derived material produced by immune cells similar to ACTH secreted by the pituitary gland in its biological activity, antigenicity, and molecular weight, but also the production of ACTH by immune cells and the pituitary gland is at least in part regulated by common mechanisms.

Some of the pluripotent cytokines that were historically the first to be detected in cells of the immune system and that regulate growth and differentiation of diverse cell types are now also found to be produced within the nervous system. Interleukin 1 and interleukin 6, which are predominantly products of mononuclear phagocytes, are also produced by glioblastoma and possibly microglial cells. It is likely that these mediators are produced during host responses to brain injury and serve functions similar to those in other organs to regulate clearance of damaged tissue and its repair. Conceivably, these interleukins may also play a role during development. The demonstration of interleukin-1 immunoreactivity in fibers of the central nervous system raises the possibility that this cytokine might also act as a neurotransmitter. *See* ENDOCRINE MECHANISMS; ENDOCRINE SYSTEM (VERTEBRATE); IMMUNOLOGY; NERVOUS SYSTEM (VERTEBRATE); NEUROSECRETION; PSYCHONEUROIMMUNOLOGY.

Martin Lotz; Wietse Kuis; Peter Villiger

Bibliography. R. Ader, *Psychoneuroimmunology*, 2d ed., 1990; M. Lotz, J. H. Vaughan, and D. A. Carson, Effect of neuropeptides on production of inflammatory cytokines by human monocytes, *Science*, 241:1218–1221, 1988; M. G. Marroso (ed.), *Trends in Neuroimmunology*, 1990; W. Pierpaoli and N. H. Spector, Neuroimmunomodulation, *Ann. N.Y. Acad. Sci.*, vol. 521, 1988; D. A. Weigent and J. E. Blalock, Interactions between the neuroendocrine and immune systems: Common hormones and receptors, *Immunol. Rev.*, 100:79–108, 1987.

Neuron

A nerve cell: the functional unit of the nervous system. Structurally, the neuron is made up of a cell body or soma and one or more long processes: a single axon and dendrites (see **illus.**). The cell body

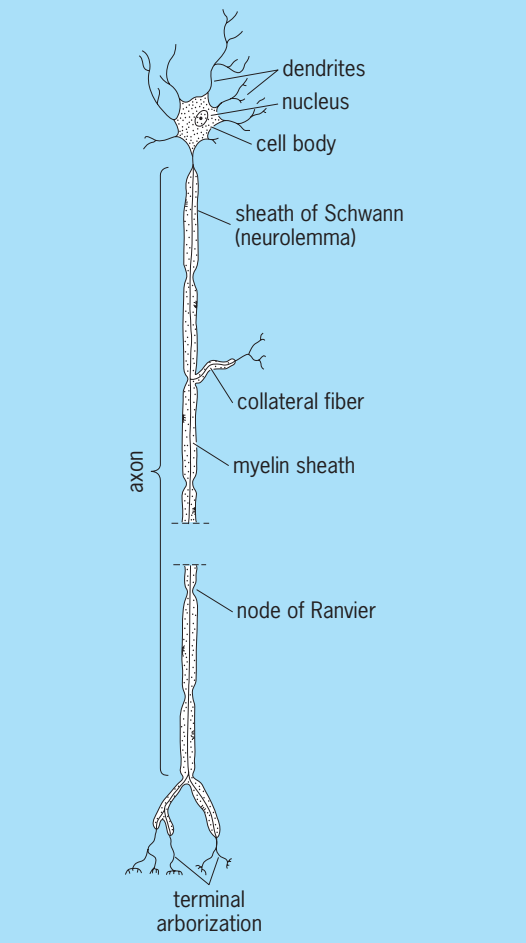

Typical vertebrate neuron. (*After C. K. Weichert, Anatomy of the Chordates, 3d ed., McGraw-Hill, 1965*)

contains the nucleus and usual cytoplasmic organelles with an exceptionally large amount of rough endoplasmic reticulum, called Nissl substance in the neuron. The longest cell process is the axon, which is capable of transmitting propagated nerve impulses. There may be none, one, or many dendrites composing part of a neuron. If there is no dendrite, it is a unipolar neuron; with one dendrite, it is a bipolar neuron; if there is more than one dendrite, it is a multipolar neuron. The dendrites are shorter and more branched than the axon. Dorsal-root spinal ganglia and most cranial nerve ganglia have unusual pseudounipolar neurons. Here a single process leaves the soma and then bifurcates, sending a long peripheral process to skin, muscle, or viscera and sending a central process into the spinal cord or brain. Both processes can conduct nerve impulses. These pseudounipolar neurons are always sensory. In most neurons only the axon propagates nerve impulses; the dendrites and somas are also irritable but do not propagate nerve impulses. *See* NERVOUS SYSTEM (VERTEBRATE). Douglas B. Webster

Neuroptera

An order of delicate insects having endopterygote development, chewing mouthparts, and soft bodies. Included are the insects commonly termed lacewings, ant lions, dobsonflies, and snake flies. The order consists of about 25 families and is widely distributed.

The adults have long, slender antennae and usually four similar wings, although the front pair is generally slightly larger than the hind pair. The adults of most species are strongly attracted to lights. The larvae, such as the aquatic hellgramite, are aggressive predators. In most of the species (suborder Planipennia) the larval mandibles are modified for piercing and for sucking the blood of prey. The larvae of lacewings are especially destructive to aphids, scale insects, and mites. The pupa of the Neuroptera is usually formed within a silken cocoon. *See* ENDOPTERYGOTA; INSECTA. Frank M. Carpenter

Bibliography. D. J. Borror et al., *An Introduction to the Study of Insects*, 5th ed., 1984; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Neurosecretion

The synthesis and release of hormones by neurons. Such neurons are called neurosecretory cells, and their products are often called neurohormones. Neurohormones function in ways comparable to the hormones produced by the nonneural endocrine cells and glands (**Fig. 1**). In fact, both endocrine and nonendocrine cells are regulated by neurohormones. Like conventional (that is, nonglandular or ordinary) neurons, neurosecretory cells are able to receive signals from other neurons. But unlike ordinary neurons that have cell-to-cell communication over short distances at synapses, neurosecretory cells release their product into an extracellular space that may be at some distance from the target cells. In an organism with a circulatory system, the

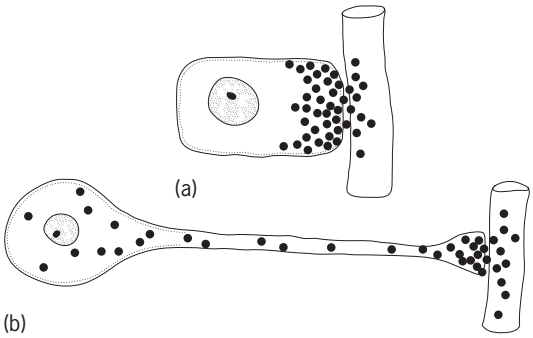

Fig. 1. Diagram of secretory cells, (*a*) Typical glandular cell (endocrine). (*b*) Typical neurosecretory cell.

neurohormones are typically sent by the vascular route to their target, whereas in lower invertebrates that lack an organized circulatory system the neurohormones apparently simply diffuse from the release site to the target. *See* ENDOCRINE MECHANISMS; ENDOCRINE SYSTEM (INVERTEBRATE); ENDOCRINE SYSTEM (VERTEBRATE).

At one time there was thought to be a sharp line of distinction between the nervous and endocrine systems; the nervous system was supposedly involved only in the coordination of transient events such as the stimulation of a skeletal muscle to contract, whereas the endocrine system coordinated long-term events such as growth. However, it is now clear not only that some neurons are capable of producing hormones but also that the nervous and endocrine systems interact in many ways, as in the suckling reflex of mammals (where the hormone oxytocin, a neurohormone, elicits milk ejection and is reflexly released in response to nerve impulses generated by stimulation of the nipples). There is a high degree of functional interaction between the classical endocrine and nervous systems, and neurosecretory cells form a major link between them. In fact, some investigators treat both of these systems merely as two components of a single coordinating system that they call the neuroendocrine system.

Neurosecretory cells have been reported from a wide variety of animals, from cnidarians to chordates. The presence of these cells in cnidarians is interesting because nerve cells originated in this group, and consequently neurosecretory cells probably had an ancient origin. Neurosecretory cells may have originally been gland cells that developed properties of neurons along with other cells that evolved into ordinary neurons. However, the opposite is also possible: neurosecretory cells may have been neurons first and only later gained the ability to secrete hormones. The absence of nonneural endocrine cells among cnidarians would seem to indicate that neurosecretory cells antedate the more conventional endocrine cells.

Synthesis and release of neurohormones. At one time all the neurohormones released from neurosecretory cells were thought to be peptides or low-molecular-weight proteins. However, it has been shown that amines, such as octopamine and dopamine, are also released from neurosecretory cells into the circulatory systems of various animals, where they function as neurohormones. Light microscopy and electron microscopy have provided much information about the structure of neurosecretory cells, particularly the so-called classical types that produce the peptide and small protein neurohormones.

In these classical neurosecretory cells, the secreted material is synthesized in the cell body and is then typically transported along the axon to the axonal terminals, where it is stored until released. However, there is evidence that, at least in some snails, release occurs directly from the cell body, and it is conceivable that release could occur along an axon as well. At the light microscope level, the peptidergic and proteinaceous neurosecretory material is seen as cytoplasmic granules, droplets, and globules. However, at the electron microscope level, these inclusions are seen to consist of the so-called elementary neurosecretory granules, which have a diameter of 100–300 nanometers and are membrane-bounded. The peptidergic and proteinaceous neurosecretory material is synthesized within the cell body by the rough-surfaced endoplasmic reticulum and subsequently packaged in the form of membrane-bounded granules by the Golgi apparatus. In some instances the staining properties of the peptidergic and proteinaceous neurosecretory material change as the granules are transported to the release site. These changes suggest that the neurosecretory material is being altered chemically during the transport process; biochemical studies have provided evidence that such changes do indeed occur. These peptidergic and proteinaceous neurohormones are actually first synthesized as part of a large precursor protein packaged by the Golgi apparatus; as the granule containing this large molecule is transported toward the release site, the hormone is cleaved from the long precursor peptide chain. Transport of neurosecretory material from the cell body to the axonal terminals can be readily demonstrated by severing the axon. The neurosecretory material accumulates on the cell-body side of the cut, while there is a concomitant depletion of secretory material in the severed axonal terminals. In contrast to the granules of classical neurosecretory cells, which are 100–300 nm, the granules in aminergic neurosecretory cells are smaller, generally 60–100 nm. *See* ENDOPLASMIC RETICULUM; GOLGI APPARATUS.

Neurosecretory cells have all of the electrical properties of ordinary neurons, such as the capability of generating and conducting action potentials. The release of neurohormones from axonal terminals into an extracellular space is triggered when the electrical activity (action potential) that is propagated by the axon enters the neurosecretory terminals. Calcium ions are essential for neurohormone release. Depolarization of the terminals because of the action potential leads to an increased concentration of intracellular free calcium. This increase is due to both the entry of extracellular calcium into the terminals and the release of intracellular sequestered calcium. This increased amount of free intracellular calcium then triggers release of the neurohormone. The neurohormones are released from the granules in which they are stored by the process of exocytosis, whereby a neurosecretory granule fuses with the cell membrane and, at the point of fusion, an opening forms and the contents of the granule are released. *See* BIOPOTENTIALS AND IONIC CURRENTS.

Major neurosecretory systems. Most of the research on neurosecretion has been done with crustaceans, insects, and vertebrates. Wherever endocrine systems have been identified in animals, neurosecretion is involved to some extent. Actually, in many invertebrate groups neurohormones are the only hormones known to be produced. In many animals the cell bodies of neurosecretory cells are clustered together,

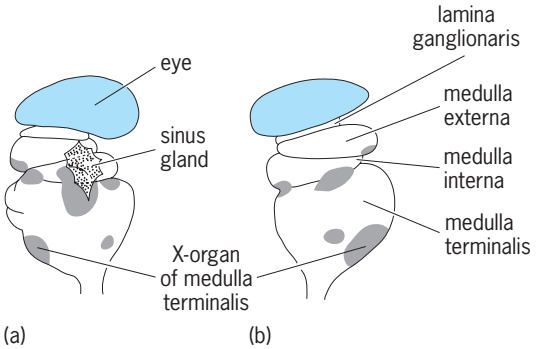

Fig. 2. Eyestalk of the crayfish *Faxonella clypeata* showing the neurosecretory cells (identified by shaded areas) and the sinus gland. (*a*) Dorsal view. (*b*) Ventral view. (*After M. Fingerman and C. Oguro, Trans. Amer. Microscop. Soc., vol. 86, 1967*)

forming centers of secretory activity that stand out conspicuously with appropriate staining. Furthermore, the axonal terminals of these cells may aggregate in close association with blood vessels or blood-filled spaces (depending upon the animal) to form a storage-release center for neurohormones. Such a group of axonal terminals that has a close spatial relationship to the circulatory system is known as a neurohemal organ.

Four extensively studied neurohemal organs are the sinus gland of crustaceans, the corpus cardiacum of insects, and the pars nervosa and median eminence of vertebrates. Sinus glands, found in all higher crustaceans, are located in the eyestalks of most stalk-eyed crustaceans (**Fig. 2**); but in crustaceans without eyestalks and in those stalk-eyed crustaceans where the sinus glands are not in the eyestalks, a pair of sinus lands is found in the head in proximity to the brain. A cluster of neurosecretory cell bodies in the medulla terminalis, called the medulla terminalis X-organ, is the predominant contributor of the axonal terminals that make up the sinus gland. In most stalk-eyed crustaceans, the medulla terminalis, which is actually part of the brain, happens to develop in the eyestalk along with the sinus gland, instead of being an integral part of the brain complex that resides in the head proper. The neurosecretory terminals in the corpora cardiaca belong to cells whose cell bodies are located in several parts of the insect brain. The hypothalamus in the vertebrate brain contains neurosecretory cell bodies from which axons emerge whose terminals form the two neurohemal components of the neurohypophysis, the pars nervosa and the median eminence. The neurohypophysis is the portion of the pituitary gland that develops as an evagination from the floor of the brain. *See* CRUSTACEA; INSECTA.

Functional roles of neurohormones. Neurohormones have a wide variety of functions. The hydra, a cnidarian, is one example of a lower invertebrate that lacks a circulatory system but has neurosecretory cells. The product of these neurosecretory cells is a peptide, the head activator, which increases the rate of head regeneration and the number of buds formed by a hydra. The neurohormones of crustaceans and insects have been shown to regulate a wide variety of functions, including metabolism, growth, and reproduction. *See* INSECT PHYSIOLOGY.

The role of the vertebrate hypothalamo-neurohypophysial system has been especially well elucidated. The pars nervosa is the site of release of vasopressin (also called the antidiuretic hormone) and oxytocin. Vasopressin protects against dehydration, promoting water retention by causing increased movement of water out of the kidney tubules back into the blood. This neurohormone also increases the permeability of frog skin to the inward passage of water, thereby facilitating water uptake by dehydrated frogs. In mammals, oxytocin stimulates contraction of uterine smooth muscle at the time of childbirth and also promotes milk ejection from the mammary glands of a nursing female. The median eminence is the release site for several hypothalamic neurohormones that regulate the adenohypophysis, the nonneural portion of the pituitary gland that develops as an evagination from the roof of the mouth. The neurohormones released from the median eminence are called hypothalamic hypophysiotropic hormones and, depending upon their actions on the adenohypophysial cells, are further identified as releasing hormones or release-inhibiting hormones, indicating whether release of a particular hormone from the adenohypophysis is stimulated or inhibited. *See* ADENOHYPOPHYSIS HORMONE; NERVOUS SYSTEM (INVERTEBRATE); NERVOUS SYSTEM (VERTEBRATE); NEUROBIOLOGY; NEUROHYPOPHYSIS HORMONE; PITUITARY GLAND.

Milton Fingerman

Bibliography. R. E. Brown, *An Introduction to Neuroendocrinology*, 1994; R. W. Hill and G. A. Wyse, *Animal Physiology*, 1989; H. F. Nijhout, *Insect Hormones*, 1994; B. T. Pickering et al. (eds.), *Neurosecretion: Cellular Aspects of the Production and Release of Neuropeptides*, 1988; M. Raabe, *Insect Neurohormones*, 1982; C. G. Scanes and P. K. Pang (eds.), *The Endocrinology of Growth, Development, and Metabolism in Vertebrates*, 1992.

Neurulation

The process by which the vertebrate neural tube is formed. The primordium of the central nervous system is the neural plate, which arises at the close of gastrulation by inductive action of the chorda-mesoderm on the overlying ectoderm. The axial mesodermal substratum causes the neural ectoderm to thicken into a distinct plate across the dorsal midline and influences both its size and shape. Its shield-like appearance, broader anteriorly and narrower posteriorly, presages the areas of brain and spinal cord, respectively.

The lateral edges of the neural plate then rise as neural folds which meet first at the level of the future midbrain, above the dorsal midline, then fuse anteriorly and posteriorly to form the neural tube. The body ectoderm becomes confluent above the closing neural tube and separates from it. Upon closure, the cells (known as neural crest cells) which occupied

the crest of the neural folds leave the roof of the tube and migrate through the mesenchyme to all parts of the embryo, forming diverse structures.

The neural tube thus formed gives rise to the brain and about half of the spinal cord. The remainder of the neural tube is added by the tail bud, which proliferates a solid nerve cord that secondarily hollows into a tube.

The closure of the neural tube confines secretions from its inner wall, which dilate the central canal by their turgor and help to expand the brain vesicles and excavate the solid cord in the tail bud. The total mass and kind of neighboring tissues control thickness or thinness of the neural wall. Adjacent to the notochord there are few mitoses and the floor of the tube remains thin, while in the somite region the lateral walls are actively mitotic and become thick.

The contour is also passively molded by adjacent structures. The distribution and extent of the peripheral tissues to be innervated influence the position and number of nerves which form at any level and the proportion of neural cells differentiating into neurons.

Neuromeres are temporary constrictions in the hindbrain resulting from localized growth processes. They were once thought to have evolutionary significance as vestiges of metamerism, but the plasticity of the neural tube with respect to its mesodermal surroundings makes this interpretation doubtful.

Neural folds never form in teleost fish. The neural plate concentrates into a solid neural keel which then hollows out secondarily as in the tail bud of other vertebrates. *See* EMBRYOLOGY; EMBRYONIC INDUCTION; GASTRULATION; NERVOUS SYSTEM (VERTEBRATE); NEURAL CREST.

Howard L. Hamilton

Neutral currents

Exchange currents which carry no electric charge and mediate certain types of electroweak interactions. The discovery of the neutral-current weak interactions and the agreement of their experimentally measured properties with the theoretical predictions were of great significance in establishing the validity of the Weinberg-Salam model of the electroweak forces.

The neutral-current weak interactions are one subclass of the forces between the fundamental constituents of matter, the quarks and leptons, called the elementary particles. The three basic interactions between these particles are the strong nuclear forces, the electroweak forces, and gravity. The electroweak forces come in three subclasses: the electromagnetic interactions, the charged-current weak interactions, and the neutral-current weak interactions.

Electroweak interactions. The electroweak interactions are theoretically understood to be a current–current type of interaction, where the interaction is mediated by an exchange current or particle. The electromagnetic interaction is mediated by an exchanged photon γ (**illus.** *a*). Since the photon carries no electric charge, there is no change in charge between the incoming and the outgoing electron

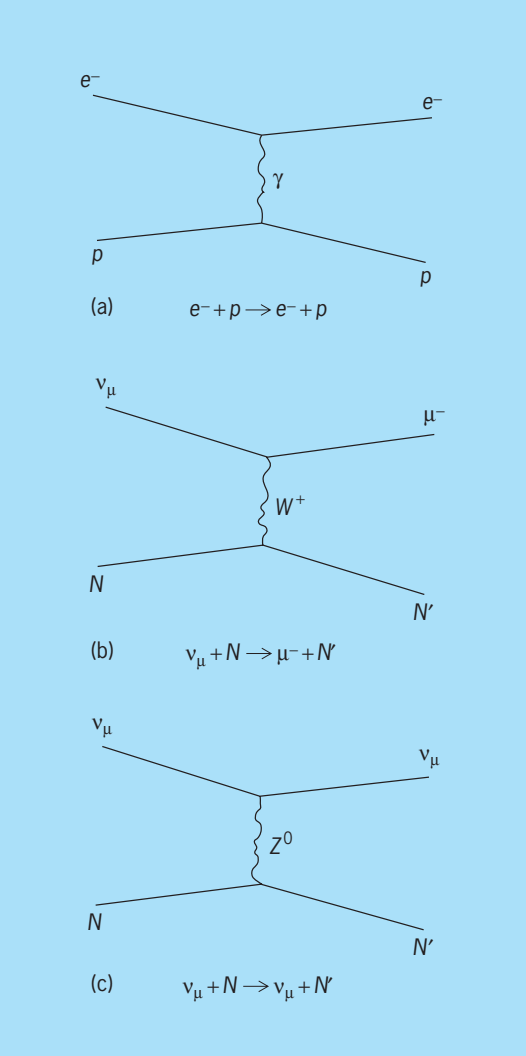

Diagrams for various electroweak interactions. (*a*) Electromagnetic interaction. (*b*) Charged-current weak interaction. (*c*) Neutral-current weak interaction.

(the upper vertex in illus. *a*). The charged-current weak interaction is mediated by the exchange of a charged intermediate boson, the $W^{\pm}$, and thus, for example, an incoming neutral lepton such as the ν_μ in illus. *b* is changed into a charged lepton, the μ^-. In the neutral-current weak interactions, the exchanged intermediate boson, the Z^0, carries no electric charge (hence the name neutral-current interaction), and thus an incident neutral lepton, the ν_μ in illus. *c*, remains an outgoing neutral ν_μ. *See* ELECTRON; INTERMEDIATE VECTOR BOSON; LEPTON; NEUTRINO; PHOTON.

Discovery and investigation. The existence of the neutral-current weak interactions was predicted by the Weinberg-Salam model, which unified the electromagnetic and the weak nuclear forces into one basic interaction, now called the electroweak interactions. The neutral-current interactions were experimentally discovered in 1973, and have since been extensively studied, in neutrino scattering processes such as neutrino-electron elastic scattering, $\nu + e^- \rightarrow \nu + e^-$; neutrino-proton scattering, $\nu + p \rightarrow \nu + p$; single-pion production, $\nu + p \rightarrow \nu + p + \pi^0$;

and inclusive neutrino scattering, $\nu + p \rightarrow \nu +$ any number of hadrons. Very important information about the properties of the neutral currents has been obtained by studying the interference effects between the electromagnetic and the neutral-current weak interactions in the scattering of polarized electrons on deuterium. Parity violating effects in atomic physics processes due to the neutral weak currents have been observed, and predicted parity-violating nuclear effects have been searched for. *See* PARITY (QUANTUM MECHANICS).

Properties. As a result of the intensive experimental studies of neutral currents in the reactions mentioned above, the properties of the neutral currents have been fairly well determined. They are a mixture of vector (V) and axial-vector (A) currents (that is, the neutral currents transform like vectors and axial vectors under spatial rotations, and not like scalars, pseudoscalars, or tensors), and they are a mixture of isoscalar ($I = 0$) and isovector ($I = 1$) currents (that is, they transform like isoscalars and isovectors under isotopic spin rotations). The coupling constants that determine the relative strength of the vector, axial-vector, and isoscalar and isovector components have been measured, and are in very good agreement with the values predicted by the Weinberg-Salam model. *See* ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS; I-SPIN; QUARKS; SYMMETRY LAWS (PHYSICS); WEAK NUCLEAR INTERACTIONS. Charles Baltay

Neutralization reaction (immunology)

A procedure in which the chemical or biological activity of a reagent or a living organism is inhibited, usually by a specific neutralizing antibody. As an example, the lethal or the dermonecrotic actions of diphtheria toxin on animals may be completely neutralized by an equivalent amount of diphtheria antitoxin—an antibody produced in animals or in humans after contact with diphtheria toxin or toxoid. Lesser amounts of antitoxin provide intermediate degrees of inhibition. These facts provide the basis for the Schick test for susceptibility to diphtheria. Tetanus and botulinus toxins may be similarly inhibited by their specific antitoxins. In contrast, the typical toxins of dysentery and other gram-negative bacteria are only slightly neutralized, even by large excesses of antibody. Antibodies to bacterial, snake venom, and other enzyme preparations regularly precipitate them from solution so that the supernates are devoid of enzyme activity; however, the neutralization of activity in the precipitate may range from complete to negligible.

Infection of a host by a living bacterium, virus, or other microorganism may also be inhibited or mitigated by the corresponding antibodies, and such neutralization tests are used in the diagnostic examination of sera or of infective agents recovered from such infections as poliomyelitis and yellow fever. In some instances the antibody may be injected into a test animal before, or occasionally shortly after, challenge with the living agent. In other instances the neutralization of the microbial infectivity by the antibody is permitted to take place in the test tube, and its degree determined by subsequent injection of the mixture into an appropriate test animal. *See* IMMUNOLOGY; NEUTRALIZING ANTIBODY; SEROLOGY. Henry P. Treffers

Bibliography. B. D. Davis et al., *Microbiology*, 4th ed., 1989.

Neutralizing antibody

An antibody that reduces or abolishes some biological activity of a soluble antigen or of a living microorganism. Thus, diphtheria antitoxin is a neutralizing antibody that, in adequate amounts, abolishes the pathological effects of diphtheria toxin in animals. This is only one characteristic; the other general properties of the antibody are those of the immunoglobulin family (IgG, IgA, or IgM) to which it belongs. *See* ANTIBODY.

Analogous neutralizing effects of antibodies can be demonstrated for the lytic effects of many lysins and, most important, for the pathogenic effects of viruses and the rickettsiae. Since the latter are complex bodies containing multiple antigens, not all the resulting antibodies need have neutralizing activities, although they may display a variety of other serological properties. Antibodies to enzymes constitute a special case; in all instances, they precipitate their corresponding enzyme, but the degree of neutralization may range from 0 to 100%. Antibodies to the endotoxins of the gram-negative bacteria regularly neutralize their toxicity only to a low degree. *See* ANTIGEN; ANTITOXIN; DIPHTHERIA; LYSIN; NEUTRALIZATION REACTION (IMMUNOLOGY); TOXIN. Henry P. Treffers

Neutrino

An elusive elementary particle that interacts with matter principally through the weak nuclear force. Neutrinos are electrically neutral spin-$^1/_2$ fermions with left-handed helicity. Many weak interaction processes (interactions that involve the weak force), such as radioactive nuclear beta decay and thermonuclear fusion, involve neutrinos. Present experimental knowledge is consistent with neutrinos being point particles that have no internal constituents. Neutrinos are classified as neutral leptons, where leptons are defined as elementary particles that interact with the electroweak (electromagnetic and weak nuclear) and gravitational forces but not with the strong nuclear force. *See* ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS; HELICITY (QUANTUM MECHANICS); LEPTON; SPIN (QUANTUM MECHANICS); WEAK NUCLEAR INTERACTIONS.

Because the role of gravitational forces is negligible in nuclear and particle interactions and because neutrinos have zero electric charge, neutrinos have the unique property that they interact almost completely via the weak nuclear force.

Consequently, neutrinos can be used as sensitive probes of the weak force. As such, neutrino beams at particle accelerators have been employed to study charge-changing (charged current) and charge-preserving (neutral current) weak interactions. However, the extreme weakness (compared to the electromagnetic and strong forces) and short range (of the order of 10^{-18} m) of the weak interaction have made determination of many neutrino properties extremely difficult.

Currently, three distinct flavors (or types) of neutrinos are known to exist: the electron neutrino (ν_e), the muon neutrino (ν_μ), and the tau neutrino (ν_τ). Each neutrino flavor is associated with a corresponding charged lepton, the electron (e), muon (μ), and tau (τ) particle. The electron, muon, and tau neutrinos (or their antiparticles) have been observed in experiments. Based on the observation of interactions involving neutrinos, the lepton flavor families, which comprise the charged and neutral leptons and their antiparticles (e^-, ν_e, e^+, $\overline{\nu}_e$; μ^-, ν_μ, μ^+, $\overline{\nu}_\mu$; τ^-, ν_τ, τ^+, $\overline{\nu}_\tau$), obey laws of conservation of lepton number. These empirical laws state that the number of leptons minus antileptons does not change, both within a flavor family and overall. *See* ELECTRON; SYMMETRY LAWS (PHYSICS).

The existence of neutrino oscillations (a phenomenon whereby neutrinos change their flavors during the flight from a neutrino source to a detector), seen clearly in observations of atmospheric and solar neutrinos, shows that neutrinos have tiny finite masses which are many orders of magnitude smaller than the masses of their charged lepton counterparts, and also shows that the physical neutrinos do not have pure mass states (quantum-mechanical states) but contain mixtures of two or more neutrino mass states. This mixing indicates that the empirical laws of lepton number conservation are not exact and that they are violated in some physical processes. It is not known whether neutrinos have magnetic or electric dipole moments.

Discovery. The neutrino was postulated by W. Pauli in 1930 in order to rescue the fundamental laws of conservation of energy and linear momentum that were seemingly violated during radioactive beta decay. Pauli proposed that in beta decay a neutral particle of half-integer spin was emitted along with the electron (beta particle). He suggested that this particle was so elusive that it escaped direct experimental detection. In 1934, E. Fermi incorporated Pauli's neutral particle into his theory of nuclear beta decay. *See* RADIOACTIVITY.

Calculations based on Fermi's theory showed just how tiny the interactions of neutrinos with matter would be. In the inverse beta-decay process [reaction (1), showing the electron antineutrino,

$$\overline{\nu}_e + p \longrightarrow n + e^+ \tag{1}$$

proton, neutron, and positron] the absorption cross section on protons is approximately 10^{-47} m^2 for free antineutrinos with energies of approximately 1 MeV. This value is equivalent to a mean free path for the neutrino in water of 6×10^{14} mi (1×10^{15} km), about 10^5 times the diameter of the solar system.

To compensate for the low probability of neutrinos interacting with matter, the first experiment to detect the neutrino directly was undertaken at a nuclear reactor. Reactors produce copious quantities of electron antineutrinos from the beta decay of fission fragments. Using the inverse beta-decay reaction (1) and detecting in time coincidence the neutrons and positrons emitted from this process, F. Reines and C. Cowan confirmed the existence of the neutrino in 1956.

The discovery of the muon neutrino and the demonstration that the muon neutrino was distinguishable from the electron neutrino were carried out in an accelerator-based experiment in 1962. A beam of high-energy neutrinos was created by pions decaying into muons and neutrinos. The neutrinos were observed to produce reaction (2) but not reaction (3), thus proving that different flavors of neutrinos must exist. In addition, the observations that electron neutrinos were associated with interactions involving electrons and that muon neutrinos were associated with interactions involving muons gave rise to the empirical law of the conservation of lepton number for separate flavors.

$$\nu + n \longrightarrow p + \mu^- \tag{2}$$

$$\nu + n \longrightarrow p + e^- \tag{3}$$

The third neutrino flavor, that of the tau neutrino, was directly observed in 2000. An intense neutrino beam was produced, which was believed to consist of tau neutrinos, and the production of tau leptons was clearly identified by observing the tracks of these leptons and of their decay products.

Scattering. High-energy neutrinos created at particle accelerators have provided precision tests of the standard electroweak model as well as insights into the quark structure of matter. Since neutrino-scattering cross sections increase linearly with energy, the use of high-energy neutrinos allows the accumulation of reasonable statistics. However, the detectors used to observe the interactions must still be quite massive. *See* PARTICLE DETECTOR.

Neutrino-electron scattering processes can be represented by Feynman diagrams (**Fig. 1**). The lines on the left of each diagram represent the incoming particles. These particles interact at the vertex, represented by the vertical lines, by transfer of particles labeled Z^0 and W^+, which are massive bosons that mediate the electroweak interactions. Outgoing or scattered particles are represented by the lines on the right. In the $\nu_e + e^-$ scattering process (Fig. 1*a*) both the neutral current interaction (mediated by the Z^0 boson) and the charged current interaction (mediated by the W^+ and W^- bosons) are possible, while the $\nu_\mu + e^-$ process (Fig. 1*b*) has only the neutral interaction. Measurements of neutrino-electron scattering not only have demonstrated the occurrence of neutral current reactions but have also confirmed the existence of the destructive interference between

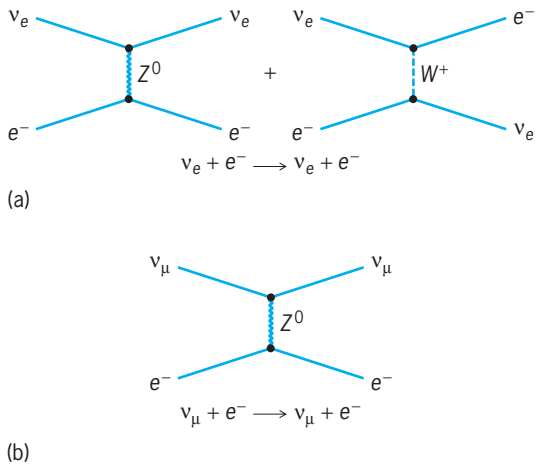

Fig. 1. Neutrino scattering diagrams of (*a*) electron neutrinos and electrons, and (*b*) muon neutrinos and electrons.

the weak neutral and charged currents, as predicted by the electroweak theory. Comparisons of neutral- to charged-current neutrino-nucleon scattering have also tested and confirmed the standard electroweak theory. *See* ELECTROWEAK INTERACTION; FEYNMAN DIAGRAM; INTERMEDIATE VECTOR BOSON; NEUTRAL CURRENTS; QUANTUM ELECTRODYNAMICS.

Studies of neutrino scattering from nucleons have allowed the investigation of the quark-antiquark composition of matter. Neutrinos are naturally polarized, and so in charge-current interactions the neutrino, which has left-handed helicity, scatters primarily from quarks, while the right-handed antineutrino interacts preferentially with antiquarks. [Helicity is defined as the alignment of a particle's spin parallel (right-handed) or antiparallel (left-handed) to the direction of its momentum.] Experiments exploiting this property have played key roles in elucidating the quark nature of matter and the theory of strong interactions. *See* HELICITY (QUANTUM MECHANICS); QUARKS.

Mass. In the minimal standard model of particle physics, the neutrino masses are chosen as zero, although there are no decisive theoretical arguments to support this choice. In fact, most theoretical models attempting the unification of strong and electroweak interactions predict nonzero neutrino masses. Therefore nonzero neutrino mass clearly indicates the existence of a new theory beyond the standard model of particle physics. *See* GRAND UNIFICATION THEORIES; STANDARD MODEL.

The consequences of neutrinos having nonzero rest masses would be quite profound, not only for the properties of neutrinos, but also for astrophysics and cosmology. For example, big bang cosmology predicts the existence of relic neutrinos analogous to the microwave photon background remnant left from the creation of the universe. These extremely low-energy neutrinos are predicted to exist with a number density of 110 neutrinos/(flavor · cm^3). At this density there would be about 10^{10} times as many neutrinos as baryons in the universe. Although the neutrino masses are now known to be too small to explain the amounts of nonluminous dark matter, they may have affected the evolution of the universe. *See* BIG BANG THEORY; COSMIC BACKGROUND RADIATION; COSMOLOGY; UNIVERSE.

Oscillations. If at least one flavor of neutrino has mass, and if any of the laws of conservation of lepton number is violated, then neutrino oscillations are possible. This can happen if the experimentally observed neutrino flavors (ν_e, ν_μ, ν_τ) are not pure quantum-mechanical mass eigenstates but are, in fact, superpositions of mass-eigenstate neutrinos (ν_1, ν_2, ν_3). As the neutrinos travel, the admixture components ν_i, which have different masses, propagate with different frequencies. Hence, an initially pure beam of ν_e neutrinos could transform over time into a beam with ν_e and admixtures of ν_μ and ν_τ (vacuum neutrino oscillation). For the two-neutrino oscillation case, the oscillation wavelength λ_{osc} is proportional to $E_\nu/\Delta m^2$, where E_ν is the neutrino energy and Δm^2 is the squared mass difference of the two neutrino mass eigenstates. The oscillation effect is maximum when the distance from a neutrino source to the detector is similar to λ_{osc}. *See* QUANTUM MECHANICS.

Experiments searching for this oscillatory behavior have been performed by using terrestrial neutrino sources at nuclear reactors and at particle accelerators and by observing extraterrestrial neutrinos, atmospheric neutrinos, and solar neutrinos. The accelerator-based experiments look for the disappearance or appearance of neutrinos of a specific flavor, while the reactor experiments are limited to searches for the disappearance of $\overline{\nu}_e$ neutrinos. The atmospheric neutrino experiments look for the disappearance or appearance of ν_μ or ν_e neutrinos, and the solar neutrino experiments look for disappearance of ν_e neutrinos. All these experiments are extremely difficult to perform, requiring careful attention to possible uncertainties of backgrounds and to the presence of unexpected systematic effects. Although there have been hints of neutrino oscillations both in the measurements of solar neutrinos (the solar neutrino problem) and in the study of atmospheric neutrinos (the atmospheric neutrino anomaly), which both observed less neutrinos than predicted, the definitive evidence of neutrino oscillations was found in 1998 in the study of atmospheric neutrinos by the Super-Kamiokande experiment. The experiment obtained strong evidence, free from the uncertainty of backgrounds and systematic effects. The existence of neutrino oscillations indicates that neutrinos have masses and new physical phenomena such as lepton flavor-violation processes like $\mu \rightarrow e\gamma$, neutrinoless double beta decay, and so on may also exist.

Atmospheric neutrinos. High-energy primary cosmic rays, consisting mostly of protons, hit the Earth's atmosphere and produce pions and kaons (mesons). The lower-energy mesons subsequently decay to muons and ν_μ neutrinos. Muons further decay into electrons and ν_e and ν_μ neutrinos. Higher-energy muons hit the Earth's surface before they decay.

Therefore, in the low-energy regions, where muons decay before reaching the surface of the Earth, the atmospheric neutrino ratio of $(\nu_\mu + \overline{\nu}_\mu)/(\nu_e + \overline{\nu}_e)$ is nearly equal to 2 and the ratio goes up as the energy becomes higher.

Measurements by large underground detectors indicate a discrepancy between the measured ratio of $(\nu_\mu + \overline{\nu}_\mu)/(\nu_e + \overline{\nu}_e)$ and the predicted ratio. Although calculation of the absolute neutrino fluxes can be done to only about 20% accuracy, it has been generally thought that the calculation of the ratio should be good to around 5%. However, combined data yield a value for the ratio of observed fluxes that is about 40% less than the predicted value. This observation can be explained by neutrino oscillations, but could also reflect theoretical model assumptions in the flux calculations.

Definitive evidence of neutrino oscillations was obtained by measuring the neutrino flux as a function of the distance from the neutrino production points to the detector. Cosmic rays approach the Earth very uniformly, and therefore the zenith-angle distribution of the incoming direction of atmospheric neutrinos is symmetric. This property does not depend upon how the neutrino flux is calculated. The zenith angle corresponds approximately to the distance from the neutrino source to the detector. Indeed, the measured zenith-angle distribution of electron neutrinos is very symmetric. However, the measured muon-neutrino distribution is very asymmetric (**Fig. 2**). Muon neutrinos coming up through the Earth, having traveled about 13,000 km (8000 mi), are strongly suppressed, and those coming from above, with a flight distance around 10 km (6 mi), are not suppressed. This distance-dependent effect is seen only in the muon-neutrino data, and cannot be explained by phenomena other than neutrino oscillations.

The electron data suggest that the electron neutrinos are not oscillating. An experiment at a reactor, which explored a similar neutrino mass range, also did not see electron neutrino oscillations. Therefore, the atmospheric neutrino oscillation is not an oscillation between muon neutrinos and electron neutrinos, but between muon neutrinos and tau neutrinos. The detailed study shows that the wavelength of the oscillation $\nu_\mu \rightarrow \nu_\tau$ observed in the atmospheric neutrinos is about 500 km/GeV, in other words, about 500 km (300 mi) for 1-GeV neutrinos.

The atmospheric neutrino oscillations can be tested by using the artificial neutrinos produced by particle accelerators with an average neutrino energy of order of 1 GeV and with a detector placed at a distance of a few hundred kilometers—a long-baseline neutrino oscillation experiment. An experiment called K2K (KEK to Kamioka), where the neutrino was produced by the high-energy accelerator at KEK (High Energy Accelerator Organization), in Japan and detected by Super-Kamiokande, 250 km (150 mi) from the neutrino source, was performed and has confirmed atmospheric neutrino oscillations. New experiments to further study neutrino masses by using high-energy accelerators are underway (2006) and are being constructed in the Unites States, Europe, and Japan.

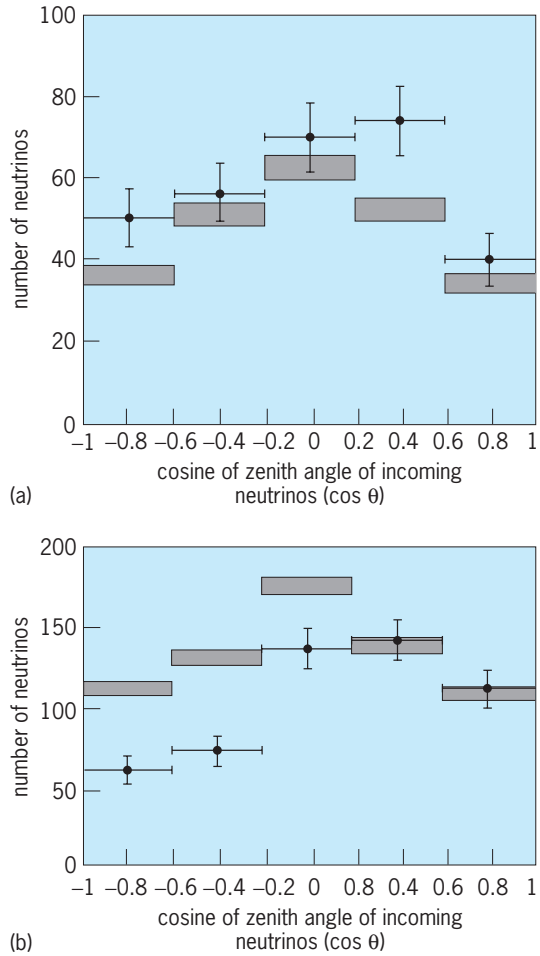

Fig. 2. Zenith-angle (θ) distribution of neutrinos observed at the Super-Kamiokande detector over a period of 535 days. The cross-bars show number of events observed in each range of value of cos θ, with the horizontal bars indicating the extent of the range and the vertical bars indicating the statistical uncertainty in the frequency of the neutrinos. Rectangles show the expected number of events predicted by a theoretical calculation; the thickness of the rectangles shows the range of uncertainty of the prediction. (*a*) Electron neutrinos. (*b*) Muon neutrinos.

Solar neutrinos. Stellar energy production in the Sun is thought to be predominantly powered by proton-proton (*pp*) chain nuclear reactions, starting with the $p + p \rightarrow d + \nu_e + e^+$ fusion process, that take place at the center of the Sun. This process and additional nuclear reactions following the *pp* fusion process are predicted to produce abundant quantities of electron neutrinos. The net chain reactions (4)

$$4p \rightarrow {}^4\mathrm{He} + 2e^+ + 2\nu_e + 26.7\,\mathrm{MeV} \qquad (4)$$

produce a thermal energy of 26.7 MeV for every two neutrinos. These neutrinos offer an ideal probe to study the physics taking place at the center of the Sun, since, unlike photons (the thermal energy) that take at least 100,000 years to diffuse outward from the core to the surface, neutrinos escape the Sun's core essentially instantaneously with no scattering. *See* PROTON-PROTON CHAIN.

Before 2001, five experiments, using three different detection methods, had all found significantly fewer neutrinos than are predicted by current theoretical solar models. This discrepancy, called the solar neutrino problem, could arise from an incomplete understanding of solar physics or neutrino properties, or a combination of both. The standard model of the Sun and all proposed nonstandard solar models are unable to account for the observed features of the data.

However, all the experimental results are nicely explained either by vacuum neutrino oscillations or by matter-enhanced neutrino oscillations occurring as the electron neutrinos propagate through the Sun. (The possible occurrence of matter-enhanced oscillations is known as the MSW effect, after S. P. Mikheyev, A. Yu. Smirnov, and L. Wolfenstein.) Unfortunately, there are more than one possible sets of allowed masses and mixing parameters that accommodate all the solar neutrino experiments. This interpretation depends upon the solar neutrino flux calculations, and therefore could not provide convincing evidence.

In June 2001, the Sudbury Neutrino Observatory (SNO) experiment, a 1000-ton heavy-water experiment located in Sudbury, Ontario, Canada, announced the results of the charged-current measurement of solar electron neutrinos, involving exclusively the ν_e content of solar neutrinos arriving at the Earth. The result was compared to the results of the Super-Kamiokande experiment located in Kamioka, Japan, which has a small sensitivity to ν_μ and ν_τ as well as ν_e through neutrino-electron elastic scattering. The measured flux of SNO, $1.75 \pm 0.07 \pm 0.12 \times 10^6/\text{cm}^2/\text{s}$, is significantly lower than the result of Super-Kamiokande, $2.32 \pm 0.03 \pm 0.08 \times 10^6/\text{cm}^2/\text{s}$, by 3 standard deviations, which indicates existence of the ν_μ and ν_τ components in solar neutrinos at the Earth—the result of the neutrino oscillation. The solar neutrino problem has been finally resolved.

The solar neutrino oscillation can be tested by a terrestrial experiment. The confirmation of the solar neutrino oscillation and the precise determination of the oscillation parameters was carried out by the KamLAND experiment located in the Kamioka mine in Japan, which measured the neutrinos produced from nuclear power reactors with an average distance of 180 km (110 mi).

Supernova neutrinos. In February 1987, light from a supernova located about 163,000 light-years (9.6×10^{17} mi or 1.54×10^{18} km) away in the Large Magellanic Cloud (SN1987A) reached the Earth. Preceding the optical signal, a burst of neutrinos was independently observed in two water Cerenkov detectors. This detection of 19 neutrino events was the first recorded signature of the physical phenomena that occur at the core of a star during stellar collapse, and is believed to have signaled the creation of a neutron star. The observation confirmed that about 99% of the binding energy of the star is emitted as neutrinos, as predicted from theories of stellar collapse. Model-independent constraints on ν_e lifetime ($>1.6 \times 10^5$ years) and ν_e mass (<30 eV/c^2) were also determined from the observation of these few events. *See* GRAVITATIONAL COLLAPSE; SUPERNOVA.

Direct mass searches. Direct searches for neutrino mass rely on the kinematics of decay reactions (see **table**). The most sensitive mass limit has been set for the electron neutrino (ν_e). This value comes from a measurement of the shape of the beta-decay energy spectrum for tritium (an isotope of hydrogen). Measurements of the electron energy spectrum are made in the energy region where the electron carries away the maximum possible energy E_0. If the neutrino has a mass, then the number of decays in the region below E_0 will be less than would be expected for the case of a massless neutrino. However, subtle effects such as instrumental resolution, energy loss of the electrons, and so on can cause shape effects that are of similar size to the effect of nonzero neutrino mass. Thus, precise measurement of a value or limit for the neutrino mass requires complete understanding of all systematic effects. A 1999 experiment using a simple gaseous molecular tritium source set an essentially model-independent limit of $m_\nu < 2.5$ eV/c^2, and results from three additional experiments are in agreement with this measurement.

Lepton masses

Neutrinos		Charged leptons	
Particle	Mass	Particle	Mass
ν_e	<2.5 eV/c^2	c	0.5110 MeV/c^2
ν_μ	<0.17 MeV/c^2	μ	105.7 MeV/c^2
ν_τ	<0.18 MeV/c^2	τ	1777.1 MeV/c^2

Double beta decay. In some radioactive nuclei, ordinary beta decay is energetically forbidden, but the decay given by reaction (5), called two-neutrino (2ν)

$$\text{Nucleus}\,(Z, A) \longrightarrow \text{nucleus}\,(Z+2, A) + 2e^- + 2\nu_e \tag{5}$$

double beta decay, is energetically allowed. Here, Z is the atomic number and A the atomic mass. This decay mode, the rarest process observed in nature (first seen in geochemical experiments), is allowed by the standard electroweak theory. Another decay mode, given by reaction (6), called

$$\text{Nucleus}\,(Z, A) \longrightarrow \text{nucleus}\,(Z+2, A) + 2e^- \tag{6}$$

neutrinoless or 0ν double beta decay, will occur if the neutrino has mass and the neutrino is a Majorana particle (meaning that the particle and antiparticle are indistinguishable, which requires that lepton number is violated). This decay process can be considered a two-step process in which a neutrino emitted in the first step is absorbed in the second.

None of the experimental searches for this process, which employ direct counting experiments that can distinguish two-neutrino and neutrinoless decays, have yet observed the neutrinoless process. Although there is a claim of observing an excess of events in the decay process of ^{76}Ge, this result is very controversial and further studies and confirmation will be needed. The first laboratory observation of the two-neutrino double beta-decay process

in 1986 [^{82}Se $T_{1/2}(2\nu) = (0.8 - 1.9) \times 10^{20}$ years, where $T_{1/2}(2\nu)$ refers to the half-life for the two-neutrino decay of the parent nuclei] was followed by increasingly sensitive experiments using isotopically enriched detectors and the two-neutrino double beta-decay process is now very well established. Those experiments with higher sensitivity have been searching for the elusive neutrinoless decay. The present limits from these experiments for neutrinoless double beta decay are $T_{1/2}(0\nu) > 10^{23}$–10^{25} years. This corresponds, under the assumption that neutrinos are Majorana particles, to deduced Majorana neutrino mass limits of $m_\nu < 0.4$ eV/c^2. Yoichiro Suzuki

Bibliography. Q. R. Ahmad et al., Measurement of the rate of $\mu_e + d \rightarrow p + p + e^-$ interactions produced by ^{8}B solar neutrinos at the Sudbury Neutrino Observatory, *Phys. Rev. Lett.*, 87:071301 (6 pp.), 2001; F. Boehm and P. Vogel, *Physics of Massive Neutrinos*, 2d ed., 1992; E. D. Commins and P. H. Bucksbaum, *Weak Interactions of Leptons and Quarks*, 1983; S. Fukuda et al., Solar ^{8}B and hep neutrino measurements from 1258 days of Super-Kamiokande data, *Phys. Rev. Lett.*, 86:5651–5655, 2001; Y. Fukuda et al., Evidence for oscillation of atmospheric neutrinos, *Phys. Rev. Lett.*, 81:1562–1567, 1998; R. N. Mohapatra and P. B. Pal, *Massive Neutrinos in Physics and Astrophysics*, 3d ed., 2004; D. H. Perkins, *Introduction to High Energy Physics*, 4th ed., 2000; C. Sutton, *Spaceship Neutrino*, 1992; Y. Suzuki and Y. Totsuka (eds.), *Neutrino 98: Proceedings of the 18th International Conference on Neutrino Physics and Astrophysics*, 1999; K. Zuber, *Neutrino Physics*, 2003.

Neutrino astronomy

The detection and study of neutrinos to learn about astronomical objects and the universe. Almost all current knowledge of the universe derives from the observation of photons. Radio waves, infrared radiation, visible light, ultraviolet waves, x-rays, and gamma rays are all electromagnetic waves composed of photons. Some further knowledge of the cosmos beyond the solar system is gained by observing cosmic rays, which are mostly protons and heavier atomic nuclei. But these positively charged particles do not point to their place of origin because of the magnetic fields of the Milky Way Galaxy, which bend their flight paths. *See* COSMIC RAYS; ELECTROMAGNETIC RADIATION; GAMMA-RAY ASTRONOMY; INFRARED ASTRONOMY; MAGNETISM; MILKY WAY GALAXY; PHOTON; RADIO ASTRONOMY; ULTRAVIOLET ASTRONOMY; X-RAY ASTRONOMY.

Deep, sharply focused examination of the universe requires a telescope that can see a particle that is not much affected by gas, dust, and magnetic fields. Neutrinos are a candidate. These neutral, weakly interacting particles come almost without any disruption straight from their sources, traveling at very close to the speed of light. A low-energy neutrino in flight would not notice a barrier of lead 50 light-years thick. Neutrino light would provide a wondrous new view of the universe.

Neutrinos in the universe. The fundamental building blocks of the universe, of which all matter is composed, are the fermions: the quarks (up, down, charmed, strange, top, and bottom) and leptons (the electron, muon, and tau-on, plus a neutral particle partner for each, the electron-neutrino, muon-neutrino, and tau-neutrino). Data from the Super-Kamiokande experiment presented in 1998 establish with high probability that some neutrinos have mass, and thus so do all the fermions. *See* ELEMENTARY PARTICLE; LEPTON; QUARKS.

Neutrinos were made in huge numbers at the time of the big bang. Like the cosmic background radiation, they now possess little kinetic energy as a result of the expansion of the universe. There are expected to be at least 114 neutrinos per cubic centimeter, averaged over all space. There could be many more at Earth because of condensation of neutrinos, moving slowly under the gravitational pull of the Milky Way Galaxy. *See* BIG BANG THEORY; COSMIC BACKGROUND RADIATION.

The problem with observing these relic neutrinos is the probability of a neutrino interacting within a detector decreases with the square of the neutrino's energy, for low energies. And even when the neutrino does react in the detector, the resulting signal is minuscule. Nobody has been able to detect these lowest-energy neutrinos as yet, and prospects are not good for doing so, at least in the near future.

Stellar neutrinos. Neutrinos also originate in the nuclear fusion in stars. The Sun close by produces a huge flux of neutrinos, which have been detected in five experiments. A long-standing mystery persists in the deficit of about a factor of 2 in the numbers of neutrinos detected compared to expectations—the solar neutrino problem. This deficit is now thought probably to result from neutrino oscillations. The calculation of the neutrino flux from the Sun, in as close agreement with observations as it is, represents a great triumph for the understanding of stellar burning and evolution. However, observations of stellar neutrinos are limited to the Sun. Just as the sky is dark at night despite all the stars, the Sun far outshines all the rest of the cosmos in numbers of detectable neutrinos. *See* SOLAR NEUTRINOS.

Supernovae. On February 23, 1987, two detectors in deep mines in the United States (the IMB experiment) and Japan (the Kamiokande experiment) recorded a total of 19 neutrino interactions over a span of 13 seconds. Two and a half hours later, astronomers in the Southern Hemisphere saw the first supernova to be visible with the unaided eye since 1604, located in the Large Magellanic Cloud at a distance of 50 kiloparsecs (roughly 150,000 light-years). Many deductions followed about the nature of neutrinos, such as limits on mass, charge, gravitational attraction, and magnetic moment.

Supernovae of the gravitational-collapse type occur when elderly massive stars run out of nuclear fusion energy and can no longer resist the force of gravity. The neutrinos carry off most of the in-fall energy, some 10% of the total mass-energy of the inner part of a star of about 1.4 solar massess. Approximately 3×10^{46} joules of energy is released, along

with about 10^{58} neutrinos, over a few seconds. This is 1000 times the Sun's energy release over its whole lifetime. The awesome visible display consists of a mere one-thousandth of the energy release in neutrinos.

Much can be learned from the final stages of stellar evolution, not only about the process of stellar collapse to a neutron star or black hole (the latter if the progenitor is very massive) but also about properties of neutrinos. For example, the heavier the neutrino, the more slowly it travels, and by studying the structure of the neutrino wave passing by Earth, it may be possible to extract the relative masses of the three types of neutrinos in a direct way that is not dependent on the phenomenon of neutrino oscillations. Four underground detectors have significant capability for supernova detection from the Milky Way Galaxy. From historical records and from observations of distant spiral galaxies, the rate of supernovae in the Milky Way Galaxy is expected to be between one and five per century. Thus experimentalists may have to wait a long time before the next observation, and there is no way of predicting when it will occur. *See* SUPERNOVA.

High-energy cosmic neutrinos. Higher-energy neutrinos must be made in many of the most luminous and energetic objects in the universe. The most powerful objects seen are active galactic nuclei, which are known to produce particles with energies much higher than the most powerful human-constructed particle accelerators. Enigmatic objects such as gamma-ray bursters, which may be the most energetic explosions observed and are mostly or all at cosmological distances, produce gamma rays up to great energies as well. They may be bountiful neutrino sources or may not be, depending upon the mechanism for the radiation, at present a mystery. Seemingly disallowed cosmic rays, with energies more than 10^8 times greater than terrestrial accelerators (more than 10^{20} eV), have been observed. These particles apparently do not come from the Galaxy, and indeed remain of unknown origin. In fact, the origin of the cosmic rays generally, particularly above about 1 petaelectronvolt (10^{15} eV), remains unknown, though many models have been proposed. Whatever the source, the machinery that accelerates particles to higher energies will inevitably also produce neutrinos. At the highest energies, many speculative models have been proposed as neutrino sources. *See* ASTROPHYSICS, HIGH-ENERGY; GALAXY, EXTERNAL.

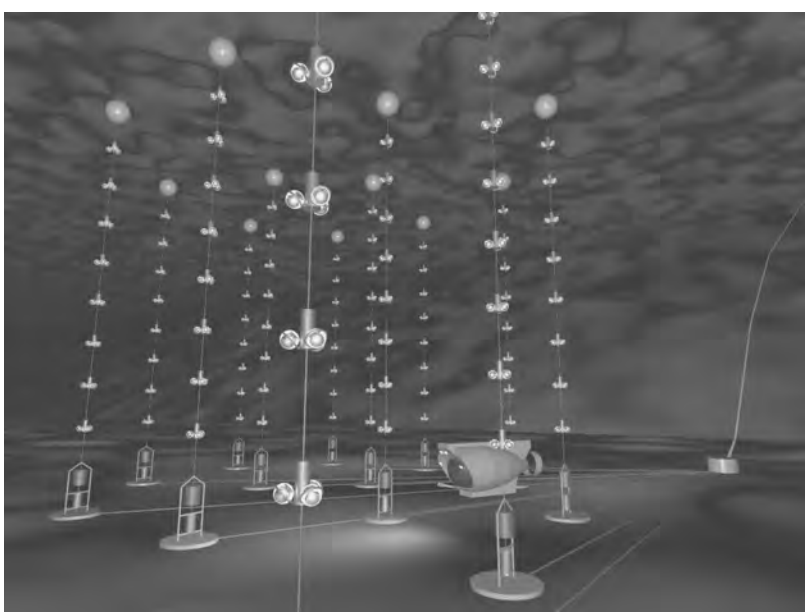

Fig. 1. View of the ANTARES Project from the ocean bottom. The optical detectors consist of modules of a cluster of photomultipliers and electronics, spaced along vertical buoy strings. Spherical floats at the top of each detector string keep the string close to vertical; anchors and releases are at the bottom. Fiber-optic cables go from each string to a junction box, which is serviced by the submarine at the far side of the array. A cable descends the slope from shore in the background. Scales are exaggerated. (*ANTARES Collaboration*)

Two things make prospects brighter in the near future for higher-energy neutrino astronomy than for lower energies. First the interaction probability for neutrinos goes up with energy. For the largest present underground detector, Super-Kamiokande, only about one in 10^{12} neutrinos of the typical energy (about 1 GeV) interact when passing through the detector and can be studied. This ratio goes up almost in proportion to the energy of the neutrino, however. Above about 1 PeV, the Earth is opaque to neutrinos and neutrinos could be seen coming only downward. At lower energies, neutrino astronomy must be done backward from optical astronomy, looking downward, using the Earth to filter out anything but neutrinos. The region between 1 TeV (10^{12} eV) and 1 PeV, roughly, is the most promising for attempts to begin regular neutrino astronomy.

The second virtue of seeking higher-energy neutrinos is that the consequences of neutrino interaction with a target (Earth or detector) become more detectable as the energy release is greater. The favored method is to detect muons produced by neutrinos. These muons (unlike electrons or tau-ons) fly a long distance (in closely the same direction as the neutrino) in Earth before stopping, for example about 1 km (0.6 mi) at an energy of a few teraelectronvolts. These charged particles produce Cerenkov radiation, a short flash of light detectable at tens of meters distance by photomultipliers in clear water or ice. Cerenkov radiation occurs when particles exceed the velocity of light in the medium, and is rather like an electromagnetic version of a sonic boom, or of the wake of a ship. Thus a detector can effectively collect the results of neutrino interactions from a target volume much greater than the detector volume itself. *See* CERENKOV RADIATION; PHOTOMULTIPLIER.

High-energy neutrino telescopes. Neutrino detectors must be placed deep underground or underwater to escape the backgrounds caused by the rain of cosmic rays upon the atmosphere. The cosmic rays produce many muons which penetrate deeply into the Earth, but with decreasing numbers with depth. Hence the first attempts at high-energy neutrino astronomy have been initiated underwater and under ice. The lead project, DUMAND (Deep Underwater Muon and Neutrino Detector), was canceled in 1995, but made great headway in pioneering techniques,

studying backgrounds, exploring detector designs, and stimulating interest in astrophysical neutrinos. Another long-running project exists in Lake Baikal, the largest (in volume) and deepest lake in the world, in Siberia. That instrument consists of large light detectors lowered on cables from the winter ice and connected to shore by cable.

Two projects similar to DUMAND are under way in the Mediterranean, the more developed NESTOR (Neutrino Experimental Submarine Telescope with Oceanographic Research) Project located off Pylos in the southwest of Greece, and the ANTARES (Astronomy with a Neutrino Telescope and Abyss Environmental Research) Project located offshore of Marseilles, France. These projects employ basically the same method of bottom-anchored cables, with photomultipliers protected in spherical glass pressure housings, as developed for DUMAND. The deep ocean water is remarkably clear with optical absorption lengths of 40–50 m (130–160 ft), and instruments can be spaced a few tens of meters apart to detect light flashes from most muons passing nearby. An array of vertical strings of such detectors can cover a cubic kilometer (**Fig. 1**). Both the NESTOR and ANTARES projects aim at prototype neutrino detectors with effective areas for muon collection in the range of 20,000–50,000 m^2 (200,000–500,000 ft^2). This may be compared with the largest present underground instruments, which are about 1000 m^2 (10,000 ft^2) in area, and the desired size for real astronomy of about 10^6 m^2 (10^7 ft^2 or 1 km^2).

A different type of neutrino telescope, the AMANDA (Antarctic Muon and Neutrino Detector Array) Project, is under construction in ice at the South Pole (**Fig. 2**). Ice below about 1.4 km (0.9 mi) depth is quite clear (100 m or 330 ft attenuation length) and bubble-free, though optical scattering is still somewhat of a problem (25 m or 80 ft effective scattering length). Experimenters have worked out a method to use hot water to drill 2-km-deep (1.2-mi) holes, down which they lower strings of photomultipliers. The instruments become permanently frozen-in after about a day, but the cables can be accessed at the surface, so no complex and expensive electronics need be placed in the inaccessible holes. *See* NEUTRINO. John G. Learned

Bibliography. P. Amram et al., Background light in potential sites for the ANTARES undersea neutrino telescope, *Astropart. Phys.*, 13:127–136, 1999; E. Andrs et al., The AMANDA neutrino telescope: Principle of operation and first results, *Astropart. Phys.*, 13:1–20, 2000; V. A. Balkanov et al., The Lake Baikal neutrino experiment, *Nucl. Phys. Proc. Suppl.*, 87:405–407, 2000; J. G. Learned and K. Mannheim, High energy neutrino astrophysics, *Annu. Rev. Nucl. Part. Sci.*, vol. 50, December 2000; L. K. Resvanis, NESTOR status report, in M. Baldo-Ceolin (ed.), *8th International Workshop on Neutrino Telescopes*, Venice, February 23–26, 1999.

Fig. 2. Simulation of ultra high-energy (10-petaelectronvolt) tau neutrino interaction in IceCube, a proposed follow-up experiment to AMANDA, 1 km^3 (0.24 mi^3) in size. The circles represent the array of optical modules, and their sizes correspond to the number of photons detected. The event has the "double-bang" signature associated with the production and decay of a tau lepton. The early hits correspond to the initial electromagnetic shower created by the tau neutrino interaction, and the later hits correspond to the decay of the tau lepton.

Neutron

An elementary particle having approximately the same mass as the proton, but lacking a net electric charge. It is indispensable in the structure of the elements, and in the free state it is an important reactant in nuclear research and the propagating agent of fission chain reactions. Neutrons, in the form of highly condensed matter, constitute the substance of neutron stars. *See* NEUTRON STAR.

Neutrons in nuclei. Neutrons and protons are the constituents of atomic nuclei. The number of protons in the nucleus determines the chemical nature of an atom, but without neutrons it would be impossible for two or more protons to exist stably together within nuclear dimensions, which are of the order of 10^{-13} cm. The protons, being positively charged, repel one another by virtue of their electrostatic interactions. The presence of neutrons weakens the electrostatic repulsion, without weakening the nuclear forces of cohesion. In light nuclei the resulting balanced, stable configurations contain protons and neutrons in almost equal numbers, but in heavier elements the neutrons outnumber the protons; in ^{238}U, for example, 146 neutrons are joined with 92 protons. Only one nucleus, ^{1}H, contains no neutrons. For a given number of protons, neutrons in several different numbers within a restricted range often yield nuclear stability—and hence the isotopes of an element. *See* ISOTOPE; NUCLEAR STRUCTURE; PROTON.

Sources of free neutrons. Free neutrons have to be generated from nuclei, and since they are bound therein by cohesive forces, an amount of energy equal to the binding energy must be expended to get them out. Usually the binding energy for each neutron amounts to 1–1.4 picojoules (6–8 MeV). Nuclear machines, such as cyclotrons and electrostatic generators, induce many nuclear reactions when their ion beams strike target material. Some of these reactions release neutrons, and these machines are sources of high neutron flux. If the accelerator is sharply pulsed, the time-of-flight method can be used for accurate energy resolution of the neutrons up to energies of about 0.3 pJ (2 MeV), the flight paths being up to 660 ft (200 m) long. *See* NEUTRON SPECTROMETRY; NUCLEAR BINDING ENERGY; NUCLEAR REACTION.

There are several kinds of portable neutron sources. Some consist of an intimate mixture of an alpha-emitting radionuclide with beryllium powder. Neutrons are released from the nuclear reaction $^{9}Be(\alpha,n)^{12}C$, which is the reaction by which the neutron was discovered in 1932. Intense sources that emit about 5×10^{10} per second are now made, using mixtures of beryllium with, for example, ^{238}Pu (half-life 89 years) or ^{241}Am (half-life 458 years). Such a source generates several hundred watts of heat. Pure ^{252}Cf (half-life 2.65 years) needs no admixture of beryllium, because neutrons are emitted in its spontaneous fission; such a source is especially compact. The neutrons emitted by the foregoing sources have energies that extend up to 0.8–1 pJ (5–6 MeV). A source that gives neutrons of lower energy is the Sb-Be photoneutron source. Here the 1.70-MeV gamma rays of ^{124}Sb (half-life 60 days) slightly exceed the binding energy of neutrons in beryllium (1.67 MeV), so the neutrons have an energy of $1.70 - 1.67 = 0.03$ MeV (4.6 femtojoules).

Neutrons are released in the act of fission, and nuclear reactors are unexcelled as intense neutron sources. The absorption of one neutron by a ^{235}U nucleus is required to induce fission, but 2.5 neutrons are on the average released; this regeneration makes possible the nuclear chain reaction. A powerful research reactor may generate neutrons in such abundance that 1 cm^2 (0.15 $in.^2$) of a sample placed therein would be traversed by 10^{15} neutrons per second. A hole through the surrounding shield can yield a collimated beam having a unidirectional flux of 10^9 neutrons/(cm^2)(s) or 6.5×10^9 neutrons/($in.^2$)(s). The explosion of a 10-kiloton (4×10^{13} J) nuclear bomb releases about 10^{30} neutrons in about 1 microsecond. *See* ATOMIC BOMB; CHAIN REACTION (PHYSICS); NUCLEAR FISSION.

Neutrons occur in cosmic rays, being liberated from atomic nuclei in the atmosphere by collisions of the high-energy primary or secondary charged particles. They do not themselves come from outer space. For information on neutrons of another origin *see* DELAYED NEUTRON; COSMIC RAYS.

Penetrating power. Neutrons resulting from nuclear reactions usually possess kinetic energies of the order of 0.2 pJ (1 MeV). Having no electric charge, they interact so slightly with atomic electrons in matter that energy loss by ionization and atomic excitation is essentially absent. Consequently they are vastly more penetrating than charged particles of the same energy. The main energy-loss mechanism occurs when they strike nuclei. As with billiard balls, the most efficient slowing-down occurs when the bodies that are struck in an elastic collision have the same mass as the moving bodies; hence the most efficient neutron moderator is hydrogen, followed by other light elements: deuterium, beryllium, and carbon.

The great penetrating power of neutrons imposes severe shielding problems for reactors and other nuclear machines, and it is necessary to provide walls, usually of concrete, several feet in thickness to protect personnel. The currently accepted health tolerance levels for an 8-h day correspond for fast neutrons to a flux of 20 neutrons/(cm^2)(s) or 130 neutrons/($in.^2$)(s); for slow neutrons, 700/(cm^2)(s) or 4500/($in.^2$)(s). On the other hand, fast neutrons are useful in some kinds of cancer therapy. *See* RADIATION DAMAGE TO MATERIALS; RADIATION INJURY (BIOLOGY); RADIATION SHIELDING; RADIOLOGY.

Detection of neutrons. In pulse counting, neutrons are allowed to produce exothermic (energy-releasing) nuclear reactions, the ionizing products of which are made to generate electrical impulses, in a proportional counter, ionization chamber, or scintillation counter, that can be amplified for individual counting. A proportional counter containing boron, either as a coating on the inner walls or as a filling gas (boron trifluoride), counts neutrons by virtue of the reaction $^{10}B(n,\alpha)^{7}Li + 0.451$ pJ (2.78 MeV). An ionization chamber coated internally with ^{235}U gives ionization pulses from the energy of fission fragments as they travel through the gas. A lithium iodide crystal (europium-activated) scintillates because of the energy released by the reaction $^{6}Li(n,\alpha)^{3}H + 0.765$ pJ (4.78 MeV). The light pulses (scintillations) are reflected onto a photomultiplier, which transforms them to electrical pulses. Capture gamma-rays emitted from strong neutron absorbers, such as cadmium, can similarly be registered by scintillation counting. Large and sensitive neutron detectors have been made by dissolving cadmium or boron salts in tanks containing scintillating liquids. *See* IONIZATION CHAMBER; PARTICLE DETECTOR; SCINTILLATION COUNTER.

In detection by activation, advantage is taken of the fact that many elements become radioactive under neutron irradiation. A sample is exposed, and its radioactive strength is subsequently measured by conventional counting equipment. Gold and indium foils are convenient and sensitive detectors of this kind. Their applications can be further specialized by taking advantage of resonance absorption. If, for example, gold foil is enclosed in cadmium, the cadmium will exclude thermal neutrons, and the gold will be activated mainly by neutrons with an energy of 0.78 attojoule (4.9 eV) because gold has a neutron capture resonance at that energy. Other elements can be similarly used for other selected energies. The converse also occurs; for example, if a thick plug of ^{57}Fe is placed in a beam of fast neutrons, it will

preferentially transmit neutrons with an energy of 4 pJ (25 keV) because at that energy the neutrons interact only weakly with the ^{57}Fe nuclei.

Detection by recoil is particularly applicable to the counting of fast neutrons. A counter with hydrogenous walls or filling gas (for example, methane) gives pulses because the protons produce ionization when they recoil after being struck by the fast neutrons.

Intrinsic properties. Free neutrons are themselves radioactive, each transforming spontaneously into a proton, an electron (β^- particle), and an antineutrino. The energy release is 0.125 pJ (0.782 MeV) per event, and the half-life is 10.61 ± 0.16 min. This instability is a reflection of the fact that neutrons are slightly heavier than hydrogen atoms. The neutron's rest mass is 1.0086652 atomic mass units on the unified mass scale (1.67495×10^{-24} g), as compared with 1.0078252 atomic mass units for the hydrogen atom.

Neutrons are, individually, small magnets. This property permits the production of beams of polarized neutrons, that is, beams of neutrons whose magnetic dipoles are aligned predominantly parallel to one direction in space. The magnetic moment is -1.913042 nuclear magnetons. The magnetic structure has a finite size, being roughly exponential in intensity, with a root-mean-square radius of 0.9×10^{-13} cm. Neutrons spin with an angular momentum of $^1/_2$ in units of $h/2\pi$, where h is Planck's constant. The negative sign attached to the magnetic moment indicates that the magnetic moment vector and the angular momentum vector are oppositely directed. *See* MAGNETON; NUCLEAR MOMENTS; NUCLEAR ORIENTATION; SPIN (QUANTUM MECHANICS).

Despite its overall neutrality, the neutron does have an internal distribution of electric charge, as has been revealed by scattering experiments. On a still finer scale, the neutron can also be presumed to have a quark structure in analogy of that of the proton. *See* QUANTUM CHROMODYNAMICS; QUARKS.

If the centers of the + and − charge distributions in the neutron should be slightly displaced from each other, the neutron would have an electric dipole moment. This possibility has a fundamental importance because it is linked through various interaction theories with the conservation of parity and with the symmetry of time reversal. (If time-reversal invariance holds, the neutron should have no electric dipole moment.) So far it has been found that if the separated charges are equal to $\pm e$ (the electronic charge), the distance between their centers must be less than 10^{-24} cm. This limit is not yet sufficiently small to give a conclusive choice between the various forms of theoretically possible interactions, but it is likely that the sensitivity of the experiments can be increased through the use of ultracold neutrons. *See* PARITY (QUANTUM MECHANICS); SYMMETRY LAWS (PHYSICS).

Ultracold neutrons. When neutrons are completely slowed down in matter, they have a maxwellian distribution in energy that corresponds to the temperature of the moderator with which they are in equilibrium. At room temperature their mean energy is about 0.004 aJ (0.025 eV), their mean velocity is about 2200 m/s (7300 ft/s), and their de Broglie wavelength is about 0.2 nm. (The approximate coincidence of this wavelength with the interatomic distances in solids is the basis for the science of neutron diffraction.) The maxwellian distribution has a tail extending to very low energies, and a few neutrons (about 10^{-11} of the main neutron flux) at this extreme have energies less than 5×10^{-8} aJ (3×10^{-7} eV), and hence velocities of less than about 7 m/s (23 ft/s). The de Broglie wavelength of these ultracold neutrons is greater than 50 nm, which is so much larger than interatomic distances in solids that they interact with regions of a surface rather than with individual atoms, and as a result they are reflected from polished surfaces at all angles of incidence. *See* NEUTRON DIFFRACTION; THERMAL NEUTRONS.

A typical source of ultracold neutrons consists of a "converter" in the reflector of a neutron reactor, together with an internally smooth, evacuated tube several centimeters in diameter that leads the neutrons out through the shield. The lead-out duct has three or four bends; the ultracold neutrons are preferentially reflected at these bends and are thus selected from the numerous faster neutrons. The neutrons can be further slowed either by sending them upward against gravity [they can rise only 2–3 m (6–9 ft), and the lead-out duct can be vertical if desired], or by means of a "neutron turbine," which is a paddle wheel whose curved blades move in the same direction as the neutrons, but with lower velocity. The neutrons can be polarized by passage through or reflection from a sheet of magnetized material.

The ultracold neutrons can be stored in "neutron bottles," of which there are two kinds. One is simply a vacuum vessel with a door that can be closed after a batch of neutrons has entered. Populations of about 100 neutrons have been retained in such vessels, but the storage times are considerably shorter than the half-life of the neutrons against their natural radioactive decay, and the nature of the extra loss mechanisms is not yet fully understood. The other kind of bottle is again a vacuum vessel, but it uses a multipolar magnetic field that contains the neutrons, because the field gradients act upon the neutrons' magnetic dipole moments so as to keep the neutrons away from the walls. Such a configuration is realized in a torus with hexapole windings around its major circumference, or a sphere with polar and equatorial windings carrying opposing currents.

Ultracold neutrons are important in basic physics and have applications in studies of surfaces and of the structure of inhomogeneities and magnetic domains in solids. *See* ELEMENTARY PARTICLE. Arthur H. Snell

Bibliography. Yu. A. Alexandrov, *Fundamental Properties of the Neutron*, 1992; J. Byrne (ed.), *Neutrons, Nuclei and Matter: An Exploration of the Physics of Slow Neutrons*, 1994; R. Golub, *Ultra-Cold Neutrons*, 1991; V. K. Ignatovich, *The Physics of Ultracold Neutrons*, 1990.

Neutron diffraction

The phenomenon associated with the interference processes which occur when neutrons are scattered by the atoms within solids, liquids, and gases. The use of neutron diffraction as an experimental technique is relatively new compared to electron and x-ray diffraction, since successful application requires high thermal-neutron fluxes, which can be obtained only from nuclear reactors. (A thermal neutron is defined as a neutron possessing a kinetic energy of about 0.025 eV.) These diffraction investigations are possible because thermal neutrons have energies with equivalent wavelengths near 0.1 nm and are therefore ideally suited for interatomic interference studies. *See* THERMAL NEUTRONS.

The scattering of low-energy neutrons is generally considered a tool for the study of solid-state phenomena, but many significant investigations have also been performed to obtain information necessary for the understanding of nuclear processes. Diffraction techniques have been employed to measure numerous coherent neutron-scattering amplitudes under special conditions, and these determinations have provided details on the interaction between nuclear forces, potential scattering, and resonance effects. Experiments have also helped to establish upper limits for values of a possible small neutron electric charge and a possible small neutron electric dipole moment. The most important investigations by neutron diffraction have been concentrated on solid-state problems, because these experiments offer unique methods to obtain information on crystallographic properties, magnetic phenomena, and the dynamics of crystal lattices. In its applications to solid-state problems, neutron diffraction is very similar in both theory and experiment to x-ray diffraction, but its importance arises from the significant differences in the scattering of the two types of radiation. *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY; LATTICE VIBRATIONS; X-RAY DIFFRACTION.

The scattering of x-rays by atoms results from a scattering interaction with the atomic electrons, and the scattering amplitudes are approximately proportional to the atomic number of the scatterer. Since the electrons are distributed within the atom at distances comparable to the x-ray wavelength, interference effects occur which produce an angular distribution of the scattering, usually referred to as a form factor, that is descriptive of the spatial distribution of the electrons. In the scattering of neutrons by atoms, there are two important interactions. One is the short-range, nuclear interaction of the neutron with the atomic nucleus. This interaction produces isotropic scattering because the atomic nucleus is essentially a point scatterer relative to the wavelengths of thermal neutrons. Strong resonances associated with the scattering process prevent any regular variation of the nuclear scattering amplitudes with atomic number. These resonances can cause the scattering amplitudes to have imaginary components of appreciable size, and they can affect the phase changes between the incident and scattered neutron waves so that the scattering amplitudes can be either positive or negative. The other important process for the scattering of neutrons by atoms is the interaction of the magnetic moment of the neutron with the spin and orbital magnetic moments of the atom. The amplitude of the interaction varies with the size and orientation of the atomic magnetic moment, and the intensity of scattering has a form-factor angular dependence that is representative of the magnetic electrons within the atom. *See* SCATTERING EXPERIMENTS (ATOMS AND MOLECULES); SCATTERING EXPERIMENTS (NUCLEI).

Techniques. Although thermal neutron beams from nuclear reactors have intensities that are lower than those obtained from efficient x-ray tubes, most of the methods developed for x-ray diffraction can be used with neutrons. Furthermore, since the neutron absorption cross section for many materials is very small, diffraction effects can also be investigated by observations of the neutrons transmitted through a sample. In most experiments the sample is irradiated with monochromatic neutrons, and the scattered radiation is measured with a neutron detector, such as a proportional counter filled with boron trifluoride gas, BF_3. In structure determinations and other experiments where there is no energy change in the scattering process, only the angular distribution of the scattered neutrons is required. In inelastic scattering experiments, that is, experiments involving an increase or decrease in neutron energy, the energy distribution of the scattered neutrons must also be measured. The neutron energies can be determined with a crystal spectrometer or by analyzing the time of flight of the scattered neutrons. Both polycrystalline and single-crystal specimens can be examined, and the type of specimen is usually determined by the conditions of the experiment. However, since single-crystal techniques provide much better resolution of the diffraction peaks, this method is required for the study of complicated structures.

For those experiments requiring monochromatic neutron beams, the neutrons must be obtained by isolating a narrow slice of the neutron spectrum from the reactor, because these neutrons have a continuous energy distribution with no pronounced peaks. Monochromatization is usually accomplished by diffraction of the reactor neutrons from large single crystals, but filters and mechanical neutron velocity selectors also can be used. In investigations of certain magnetic properties, it is frequently necessary to use neutron beams that are both monochromatic and polarized. Such beams can be obtained in the monochromating process by using single crystals, because the neutrons scattered under specific conditions from particular ferromagnetic crystals are almost completely polarized. Diffraction experiments with polarized neutrons have been particularly important for precise determinations of magnetic form factors, and techniques using polarization analysis provide a unique method for separating neutrons scattered by magnetic and nuclear interactions.

Another technique for neutron diffraction utilizes pulses containing the entire spectrum of reactor

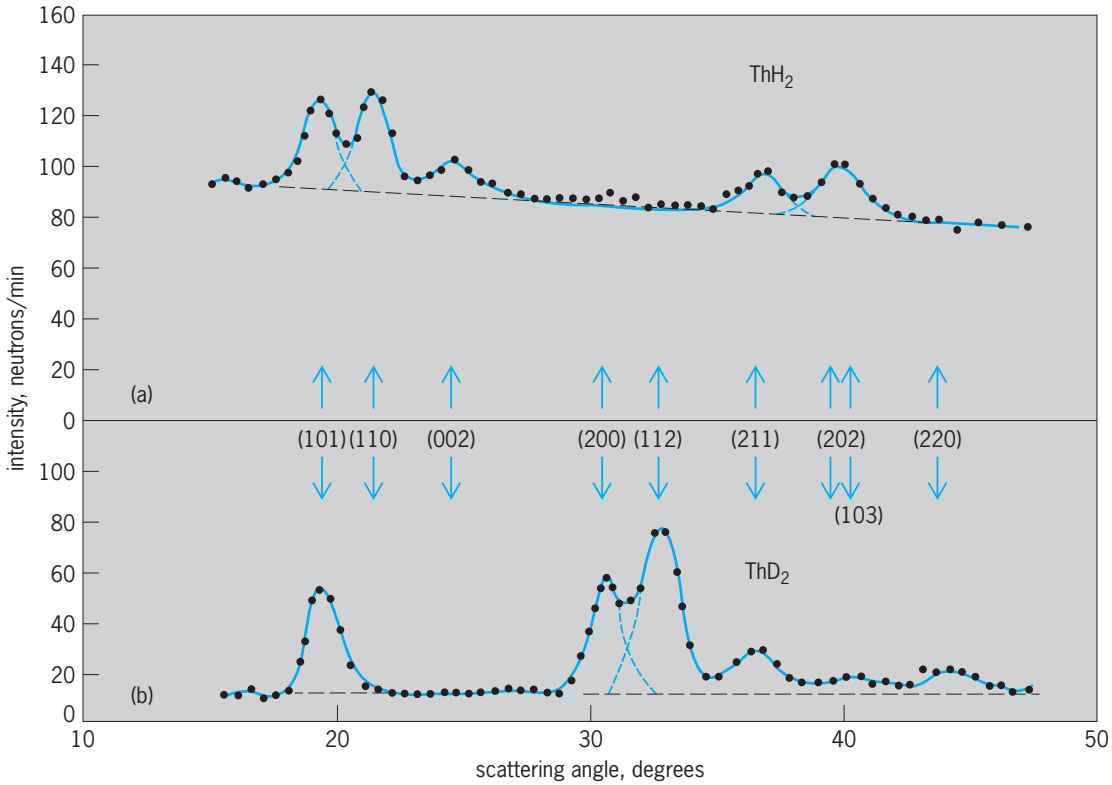

Fig. 1. Neutron diffraction patterns from (*a*) polycrystalline thorium hydride, ThH_2, and (*b*) polycrystalline thorium deuteride, ThD_2. Differences in the patterns are caused primarily by differences in the nuclear scattering from hydrogen and deuterium atoms.

neutrons and employs time-of-flight energy analysis of the neutrons scattered at a fixed angle. This technique is particularly useful with pulsed neutron sources, and it offers definite advantages in certain types of experiments where the range of scattering angles is limited. However, in most investigations with continuous neutron sources, it is not competitive with the conventional diffraction methods. Photographic techniques can also be used in neutron experiments, but they have been restricted almost completely to quick qualitative examinations.

Auxiliary apparatus for controlling the conditions of the samples can be constructed easily because of the relatively low neutron cross sections of most materials. Consequently, many diffractometers and spectrometers have low-temperature cryostats, furnaces, and electromagnets as integral parts of the instruments. Furthermore, the methods of investigation are readily adaptable to automation, and some of the newer instruments at high-flux research reactors are controlled directly by on-line computers.

Chemical crystallography. Since the nuclear scattering amplitudes for neutrons do not vary uniformly with atomic number, there are certain types of chemical structures which can be investigated more readily by neutron diffraction than by x-ray diffraction. Moreover, since neutron scattering is a nuclear process, when the scattering amplitude of an element is not favorable for a particular investigation, it is frequently possible to substitute an enriched isotope which has scattering characteristics that are markedly different.

The most significant application of neutron diffraction in chemical crystallography is the structure determination of composite crystals which contain both heavy and light atoms, and the most important compounds in this general classification are the hydrogen-containing substances. Since hydrogen and deuterium have neutron scattering amplitudes that are comparable to those of other atoms, their positions in crystals can be determined by this technique, whereas x-ray diffraction usually gives little information about them (**Fig. 1**). Most of the early investigations concerned relatively simple inorganic compounds, but with the construction of higher-flux reactors and the use of computers in data collection and processing, more complex crystal structures have been examined. The crystal structure of sucrose, $C_{12}H_{22}O_{11}$, which required the measurement and analysis of 2800 independent Bragg reflections, was determined by neutron diffraction, and studies have also been made on some of the simplest biological molecules. In addition to the inherent interest in the materials, all of these investigations have helped to provide a better understanding of hydrogen bonds in crystals. For a discussion of Bragg reflections *see* X-RAY DIFFRACTION.

Neutron diffraction techniques have also been applied to many other compounds with special chemical or physical properties and with unfavorable x-ray scattering amplitudes. These investigations include the ionic displacements associated with ferroelectric

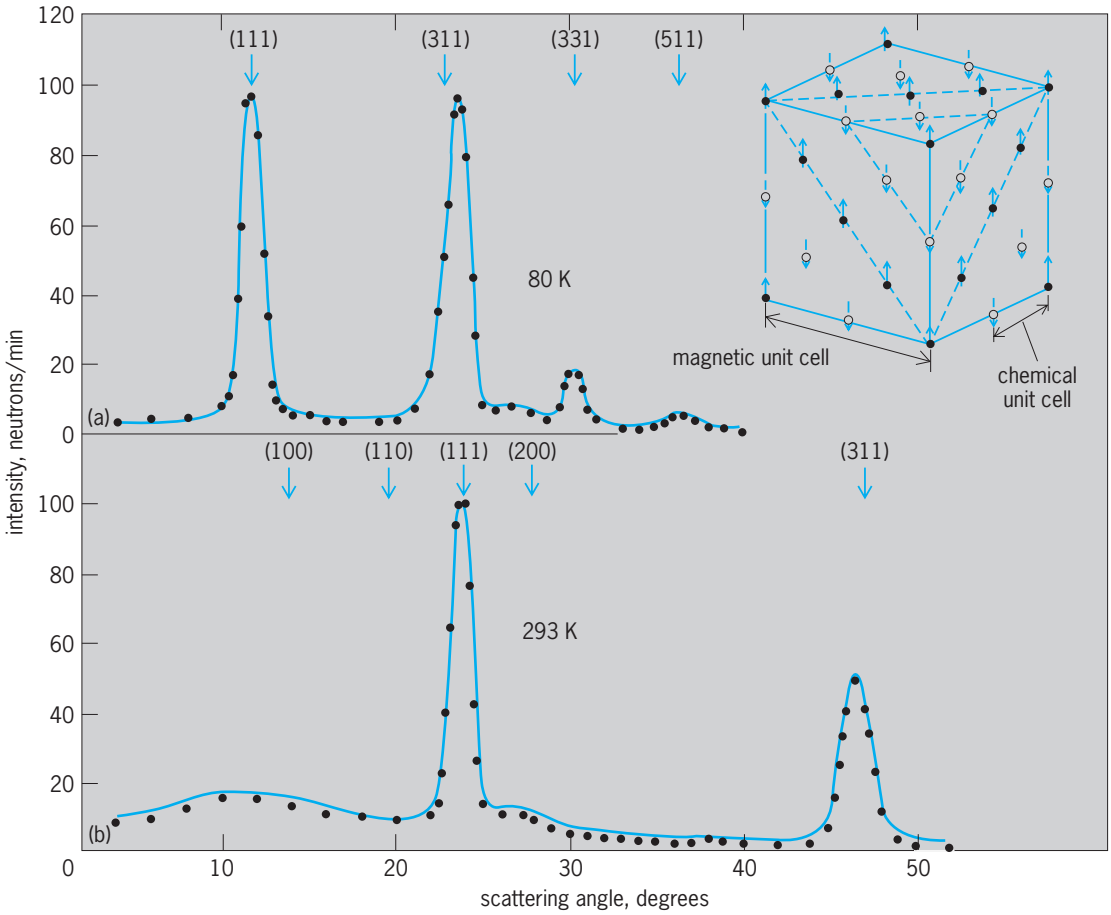

Fig. 2. Neutron diffraction patterns from polycrystalline manganese oxide, MnO, at temperatures (*a*) below and (*b*) above the antiferromagnetic transition at 122 K (−240°F). At 293 K (68°F), only nuclear reflections are observed, while at 80 K (−316°F), additional reflections are obtained from the indicated antiferromagnetic structure. The atomic magnetic moments in this structure are directed along a magnetic axis within the (111) planes.

transitions, the rotation of molecular groups in compounds, and order-disorder phenomena in alloy systems composed of atoms with almost the same atomic number. Furthermore, since the scattering of neutrons by the nucleus is isotropic, the neutron technique is advantageous in investigations of liquids, gases, amorphous materials, and other structures where the features of the diffraction pattern at large scattering angles are significant.

Magnetic scattering. The interaction of the magnetic moment of the neutron with the orbital and spin moments in magnetic atoms makes neutron scattering a unique tool for the study of a wide variety of magnetic phenomena, because information is obtained on the magnetic properties of the individual atoms in a material. This interaction depends on the size of the atomic magnetic moment and also on the relative orientation of the neutron spin and of the atomic magnetic moment with respect to the scattering vector and with respect to each other. Consequently, detailed information can be obtained on both the magnitude and orientation of magnetic moments in any substance which displays magnetic properties.

Each type of magnetic lattice has a characteristic neutron diffraction pattern. For paramagnetic materials, where the atomic moments are uncoupled and randomly oriented in direction, the magnetic scattering is diffuse. For ordered magnetic lattices the magnetic scattering is found in Bragg reflections. Magnetic reflections from ferromagnetic materials occur superimposed on the nuclear reflections, but for antiferromagnetic materials, in which the atomic moments are oriented with no net magnetization per unit volume, superlattice reflections are observed at other scattering angles, as shown in **Fig. 2**. Since ferrimagnetic materials have atomic moments with antiparallel components but also possess a net ferromagnetic moment, magnetic reflections are observed at both nuclear and other positions. Thus neutron diffraction experiments can determine the magnetic transition temperature, type of magnetic order, temperature variation of the magnetic order, and detailed magnetic configuration in the ordered lattice. This information is basic to understanding the magnetic exchange interactions that are responsible for producing an ordered magnetic structure. *See* ANTIFERROMAGNETISM; FERRIMAGNETISM; FERROMAGNETISM; PARAMAGNETISM.

The investigation of antiferromagnetic and ferrimagnetic substances is one of the most important applications of the neutron diffraction technique,

because detailed information on the magnetic configuration in these systems cannot be obtained by other methods. Several hundred antiferromagnetic structures have been investigated, and various types of systems that have been determined are shown schematically in **Fig. 3**. In most antiferromagnetic substances the magnetic moments are found in truly antiparallel arrays, but more complicated systems have been encountered. Structures have been determined in which the magnetic moments are canted with respect to each other, and a number of systems have been found to have a long-range modulation of the moment distribution. The latter configurations require long-range magnetic interactions for stability, and in the heavy rare-earth metals and alloys which have such configurations, a long-range interaction through the conduction electrons can explain many of their unusual magnetic properties. Similar complex structures have also been observed in certain types of ferrimagnetic materials.

The magnetic moments of a simple ferromagnet can usually be obtained from saturation magnetization experiments; but in substances such as ferromagnetic alloys, although the moments are arranged in parallel alignment, different types of atoms have different moment values. Since magnetic measurements can give only the average moment of the alloy, the determination of the individual magnetic moments of the constituent atoms has been another important aspect of neutron diffraction. Alloys with both ordered and disordered arrangements of the atoms can be studied, and experiments on very dilute concentrations provide information on the effect of magnetic and nonmagnetic impurities.

The form factor for the magnetic scattering of neutrons can be interpreted in terms of the spatial distribution and angular momentum characteristics of the magnetic electrons within the atoms. Determinations of these form factors can be made from either measurements of magnetic intensities in coherent reflections or from measurements of paramagnetic scattering. However, the most precise measurements of this type have been made on reflections from ferromagnetic and ferrimagnetic materials, utilizing polarized neutrons. This technique takes advantage of cross terms in the combined nuclear and magnetic scattering to provide an accuracy not readily obtainable with an unpolarized beam. Measurements on the ferromagnetic iron-group metals have provided detailed maps showing the distribution of magnetic electrons throughout the unit cells.

A variety of changes can be produced in neutron diffraction patterns by application of magnetic fields sufficiently strong to change the orientation of atomic magnetic moments within the sample. The use of magnetic fields in these experiments can therefore provide information on ferromagnetic and antiferromagnetic domains, on the magnetic anisotropy within magnetic structures, and on the nature and strength of magnetic exchange interactions.

Inelastic scattering. Scattering investigations of thermal neutrons, in which the neutrons undergo an energy change, fall into the broad scope of neutron diffraction. However, since both the angular distribution and the energy distribution of the scattered neutrons must be determined, these investigations are different from those usually associated with diffraction experiments. Because of the favorable values of energies and wavelengths associated with thermal neutrons, these measurements provide a method for studying many physical properties of solids and liquids that cannot be studied by any other method. The wavelengths are comparable with atomic separations, and the energies are of the order of the characteristic energies of solids and liquids, so that the energy and momentum changes resulting from many interactions can be measured easily by diffraction techniques. Furthermore, analogous to the case for elastic scattering, these inelastic scattering processes can result from a nuclear interaction or from a magnetic interaction, and the dynamical properties of both atomic systems and magnetic systems can be investigated.

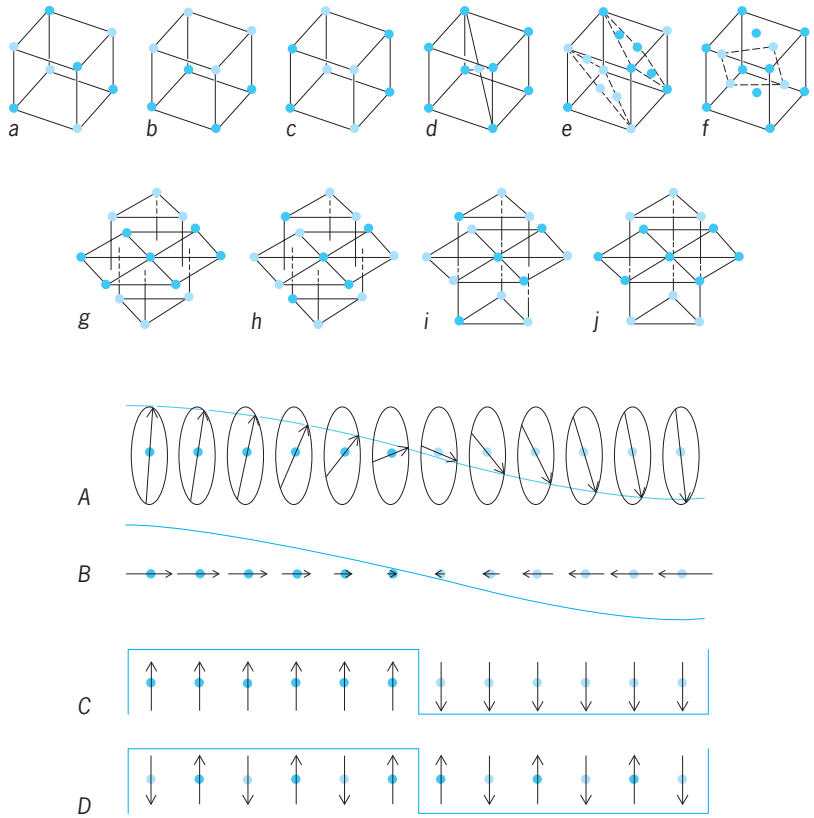

Fig. 3. Schematic representation of various antiferromagnetic systems studied by neutron diffraction. In structures *a* through *j* there is a single magnetic axis and the atomic moments at the darker circles are antiparallel to those at the lighter circles. Types of antiferromagnetism with a long-range modulation of the moment distribution are indicated by *A* through *D*.

One of the most important uses of inelastic neutron scattering is the study of thermal vibrations of atoms about their equilibrium positions, because lattice vibration quanta, or phonons, can be excited or annihilated in their interactions with low-energy neutrons. The measurements provide a direct determination of the dispersion relations for the normal vibrational modes of the crystal and do not

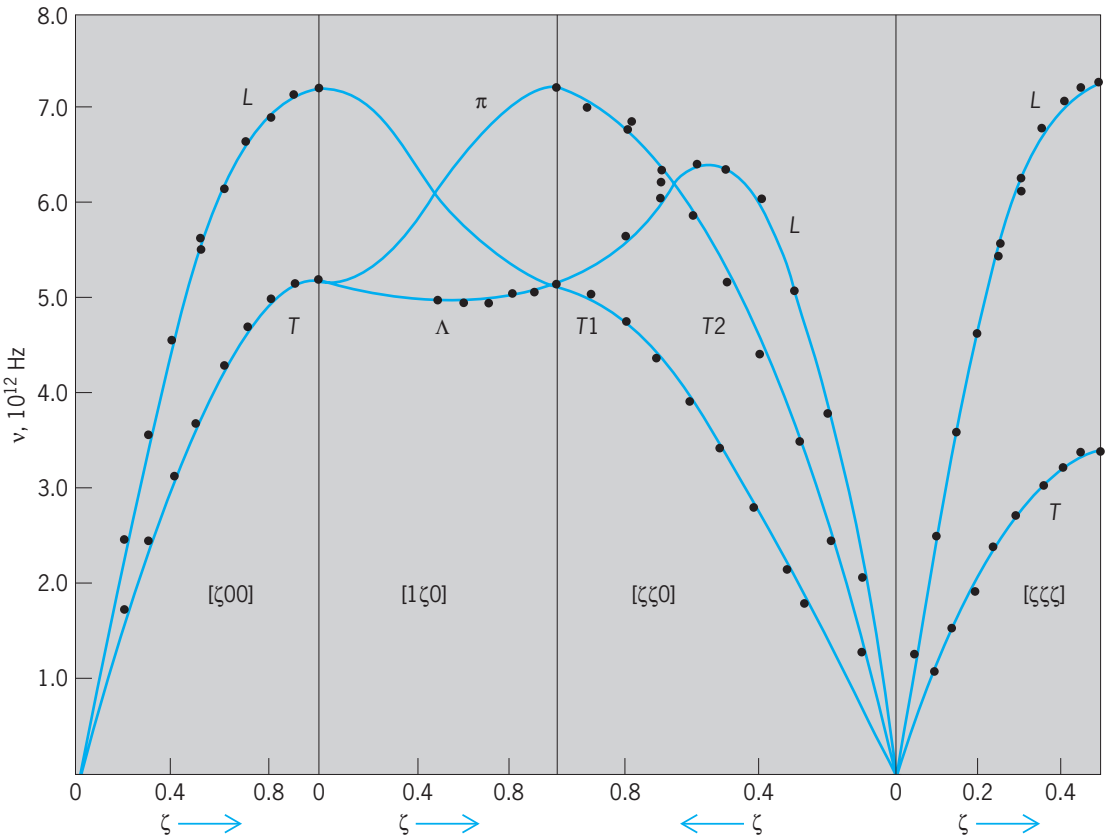

Fig. 4. Phonon dispersion curves for copper at 49 K (−371°F), which relate phonon frequency ν to phonon wave vector ζ along major symmetry directions indicated in brackets. Solid circles are results from inelastic neutron scattering experiments and smooth curves are calculations based on an axially symmetric interatomic force model extended to six nearest neighbors. (*L* and *T* correspond to longitudinal and transverse modes of vibration, respectively , while π and Λ represent modes of vibration with both longitudinal and transverse components.)

require the large corrections necessary in similar x-ray investigations. These measured dispersion relations furnish the best experimental information available on interatomic forces that exist in crystals (**Fig. 4**). Similar measurements can be made on the quantized motion of magnetic moments about the equilibrium direction in an ordered magnetic lattice, and these magnon dispersion curves can be interpreted in terms of the magnetic forces between atoms. Furthermore, neutron scattering techniques are not restricted to solids but can be used to investigate details of atomic motion in liquids.

With the availability of higher neutron fluxes and more sophisticated techniques, it has been possible to extend these investigations to more difficult problems, such as the effect of impurities on interatomic forces. Localized vibrational modes associated with impurities can be observed, and in certain types of experiments information can be obtained on the degree of spatial localization of the modes in the crystal. Dispersion curves have been measured with sufficient precision to permit observation of additional effects, including interaction between phonons and magnons and interaction between phonons and conduction electrons. The latter observations may prove useful for determining the Fermi surface in metals as a function of crystallographic direction. *See* FERMI SURFACE; MAGNON; PHONON. Michael K. Wilkinson

Bibliography. E. Balcar and S. W. Lovesey, *Theory of Magnetic Neutron and Photon Scattering*, 1989; Y. A. Izyumov, *Neutron Diffraction of Magnetic Materials*, 1991; Y. A. Izyumov and N. A. Chernoplekov, *Neutron Spectroscopy*, 1994; S. W. Lovesey, *Theory of Neutron Scattering from Condensed Matter*, 2 vols., 1986.

Neutron optics

The general class of experiments designed to emphasize the wavelike character of neutrons. Like all elementary particles, neutrons can be made to display wavelike, as well as particlelike, behavior. They can be reflected and refracted, and they can scatter, diffract, and interfere, like light or any other type of wave. Many classical optical effects, such as Fresnel diffraction, have been performed with neutrons, including even those involving the construction of Fresnel zone plates. A perfect double-slit neutron interference pattern is shown in **Fig. 1**. This pattern not only conforms closely to the predictions of scalar diffraction theory, but can be used to put a stringent upper limit on possible nonlinear contributions to the Schrödinger equation, which determines how the neutron propagates. *See* DIFFRACTION; INTERFERENCE OF WAVES; REFLECTION OF ELECTROMAGNETIC

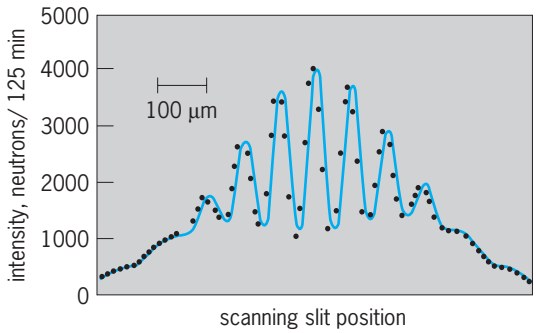

Fig. 1. Diffraction pattern made by passing a collimated neutron beam through a two-slit system. The solid line is the prediction of scalar diffraction theory. (*After A. Zeilinger et al., 1981 Symposium on Neutron Scattering, Argonne National Laboratory, American Institute of Physics, 1962*)

RADIATION; REFRACTION OF WAVES; SCATTERING OF ELECTROMAGNETIC RADIATION; WAVE (PHYSICS).

The typical energy of a neutron produced by a moderated nuclear reactor is about 0.02 eV, which is approximately equal to the kinetic energy of a particle at about room temperature (80°F or 300 K), and which corresponds to a wavelength of about 10^{-10} m. This is also the typical spacing of atoms in a crystal, so that solids form natural diffraction gratings for the scattering of neutrons, and much information about crystal structure can be obtained in this way. However the wavelike properties of neutrons have been confirmed over a vast energy range from 10^{-7} eV to over 100 MeV. *See* NEUTRON DIFFRACTION.

Neutron interferometer. Neutrons, being uncharged, can be made to interfere over large spatial distances, since they are relatively unaffected by the stray fields in the laboratory that deflect charged particles. This property has been exploited very beautifully by using the neutron interferometer, invented by H. Rauch, W. Treimer, and U. Bonse in 1974. This device is made possible by the ability to grow essentially perfect crystals of up to 4 in. (10 cm). The typical interferometer is made from a single perfect crystal cut so that three parallel "ears" are presented to the neutron beam (**Fig. 2**). This allows the incident beam to be split and subsequently recombined coherently. *See* COHERENCE; INTERFEROMETRY; SINGLE CRYSTAL.

The incident beam is split by Bragg scattering at the first ear (through angle θ off planes parallel to *a-b* in Fig. 2*b*), redirected at the second ear, and recombined at the third ear. The beam is then detected at counters *A* and *B*. Because the crystal is perfect, the de Broglie wavelength of the neutrons that are Bragg-scattered and reach the counters is very accurately determined, and its magnitude is of the same order as the crystal spacing. Therefore, a neutron wave packet that passes through the interferometer has the remarkable property that it is split into two subpackets, which travel about 4 in. (10 cm) along separate paths, I and II, during which time they are separated by several centimeters, until they are recombined at *d*. The relative phase at *d* determines the counting rates at *A* and *B*. An experimental apparatus can be inserted at *c* to change the relative phase. A change of a half wavelength will vary counter *A* from a maximum to a minimum reading (with the opposite effect on *B*). While separated, the beams have traveled a distance of about 10^9 wavelengths, and yet, if the packet in one beam is perturbed by a small fraction of a wavelength, the relative counting rates of counters *A* and *B* will significantly change. Thus, this instrument is capable both of detecting very minute forces acting on the neutron, and of demonstrating the wave behavior of the neutron over a macroscopic distance. Using these properties, experiments have been carried out with the interferometer to illustrate that neutrons are ruled by the laws of quantum theory and show very dramatic wavelike and unclassical behavior.

Gravity experiment. One of the most significant experiments performed with the interferometer (sometimes referred to as the COW experiment, after R. Colella, A. W. Overhauser, and S. A. Werner) involved rotating the interferometer about the incident beam so that one neutron path was higher than the other, creating a minute gravitational potential difference (of 10^{-9} eV) between the paths. This was sufficient to cause a path difference of 20 or so wavelengths between the beams, which could easily be seen with the interferometer (**Fig. 3**). This remains the only type of experiment that has ever seen a quantum-mechanical interference effect due to gravity. It also verifies the extension of the equivalence principle to quantum theory (although in a form more subtle than its classical counterpart). *See* GRAVITATION; RELATIVITY.

Rotation of wave function by 360°. Another significant experiment involved the first direct verification of a highly unclassical prediction of quantum theory, namely, that a neutron wave function (or that of any spin-$^1/_2$ particle such as the electron or proton) would change its sign if the neutron was rotated by 360°. Classically, of course, if an object is rotated by 360° it will be totally unaffected. But according to quantum theory, if the neutron in one leg of the

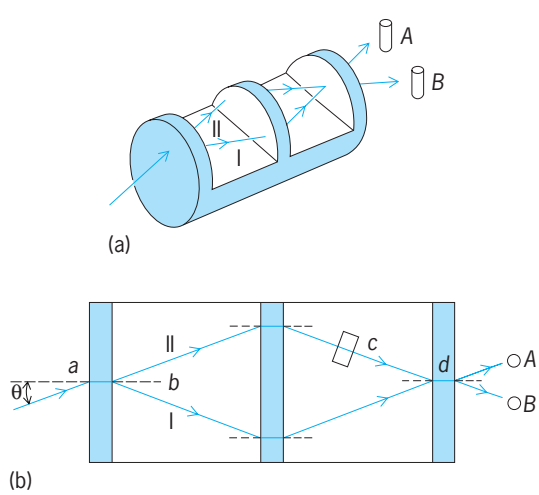

Fig. 2. Neutron interferometer. (*a*) Oblique view. (*b*) Top view. (*After D. M. Greenberger, The neutron interferometer, Rev. Mod. Phys., 55:875–905, 1983*)

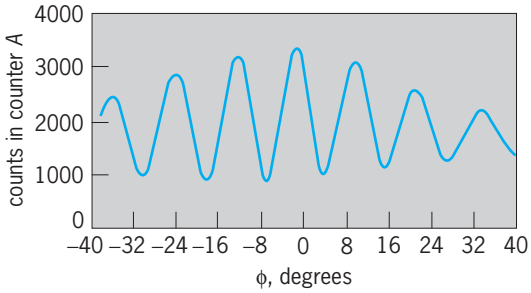

Fig. 3. Variation of counting rate in one counter with rotation angle ϕ in the gravity experiment, due to phase shift resulting from gravitational potential difference. Neutron wavelength was 0.106 nm. (*After J.-L. Staudenmann et al., Gravity and inertia in quantum mechanics, Phys. Rev., A21:1419–1438, 1980*)

interferometer is rotated by 360°, then if the two beams would have constructively interfered on recombining they will now destructively interfere and vice versa. If the neutron is rotated by another 360°, it will revert to its original state. This result was known to be true through many indirect measurements, but it was shown directly by placing a magnetic field in one leg of the interferometer which caused the neutron to precess by a known amount.

Superposition principle. One of the most important properties of quantum theory is the superposition principle, according to which the wave functions of coherent states combine additively. One of the most simple, yet most profound examples of this involves the spin for spin-$^1/_2$ particles. Such particles have two spin states, spin "up" (spin is directed upward along the z axis, $s_z = +\,^1/_2$), and spin "down" ($s_z = -^1/_2$). If the z component of spin is measured, it is found that the particle is either spin up or spin down. Yet a spin along the x axis can be described as a linear combination of these two states. By a judicious use of magnetic fields, the neutron was polarized so that it was spin up in one leg of the interferometer and spin down in the other. Then when these two beams of equal amplitude, each polarized along the z axis, were recombined after passing through the interferometer, it was seen that they were polarized in the x-y plane, exactly according to the laws of quantum theory. This polarization was perpendicular to that of either of the beams separately, which proved both the coherence of the beam and the superposition principle for spin. This is the first time a beam of spin-$^1/_2$ particles has ever been spatially separated and then coherently recombined in this fashion. *See* SUPERPOSITION PRINCIPLE.

Ultracold neutron resonances. Many noninterferometer experiments have also been done with neutrons. In one experiment, resonances were produced in transmitting ultracold neutrons (energy about 10^{-7} eV) through several sheets of material. This is theoretically similar to seeing the few lowest states in a square-well potential in the Schrödinger equation. *See* NEUTRON; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS. Daniel M. Greenberger

Bibliography. U. Bonse and H. Rauch (eds.), *Neutron Interferometry*, 1979; D. M. Greenberger and A. W. Overhauser, The role of gravity in quantum theory, *Sci. Amer.*, 242(5):66–67, May 1980; V. F. Sears, *Neutron Optics*, 1989; S. A. Werner, Neutron interferometry, *Phys. Today*, 33(12):24–30, December 1980.

Neutron spectrometry

A generic term applied to experiments in which neutrons are used as the probe for measuring excited states of nuclides and for determining the properties of these states. The term neutron spectroscopy is also used. The strength of the interaction between a neutron and a target nuclide can vary rapidly as a function of the energy of the incident neutron, and it is different for every nuclide. At particular neutron energies the interaction strength for a specific nuclide can be very strong; these narrow energy regions of strong interactions are called resonances. The strength of the interaction, expressing the probability that an interaction of a given kind will take place, can be considered as the effective cross-sectional area presented by a nucleus to an incident neutron. This cross-sectional area is expressed in barns (1 barn = 10^{-28} m^2) and is represented by the symbol σ. The neutron total cross section of the nuclide ^{231}Pa from 0.01 to 10 eV is shown in **Fig. 1**. Even though the neutron has zero charge, neutron energies are measured in electronvolts (1 eV = 1.60 × 10^{-19} joule). Neutron spectroscopy covers the vast energy range from 10^{-3} eV to 10^3 MeV.

Unbound and bound states of nuclides. Each resonance corresponds to an unbound excited state of the compound nucleus (**Fig. 2**) at an excitation energy that is the sum of the energy of the neutron and the binding energy of the neutron (4–11 MeV) which has been added to the target nuclide. The compound nucleus has a mass number which is one more than that of the target nuclide. Near the ground state of the compound nuclide the spacing of energy levels may be 10^4 to 10^6 eV. However, for a heavy nuclide, such as the compound nuclide ^{232}Pa, the excited states at an excitation energy just above the binding energy of approximately 5.5 MeV are less than 1 eV apart. To observe the individual states, the neutron energy resolution must be smaller than the level spacing. This can be achieved with low-energy neutrons, because they provide the requisite resolution, and there is no Coulomb repulsion to prevent them from entering the target nucleus. Neutron spectroscopy is presently the only technique which can provide this detailed information. For light nuclei (atomic weights ≤ 40) the spacing between the excited states can be many keV, and resonances can be resolved up to neutron energies of several MeV (**Fig. 3**). Lower-energy bound excited states of the compound nucleus below the neutron binding energy can also be studied by gamma-ray spectroscopy, observing the energy of the gamma rays emitted after the capture of neutrons at resonances or at thermal energy (Fig. 2).

Nuclear energy levels of the target nuclide can also be determined by measuring the energy spectrum of

neutrons which are inelastically scattered by the target under bombardment from monoenergetic MeV incident neutrons (**Fig. 4**). The energy of an excited state is equal to the difference in energies between the incident and scattered neutrons, $E_n - E_{n'}$. If the incident neutron in Fig. 4 has energy E_{n1}, it has enough energy to excite any of the six lowest levels and emit a neutron of lesser energy than E_{n1}. A neutron of energy E_{n2} could excite only the two lowest levels and emit a neutron of lesser energy than E_{n2}. Information on these same low-energy states can be obtained by measuring the energies and intensities of the gamma rays from the deexcitation of these states excited by inelastically scattered neutrons.

Neutron reactions and resonance parameters. The abbreviated notation for neutron reactions, (n,n) and so on, lists the bombarding particle before the comma, and the emitted particle or particles after the comma. The standard symbols are n (neutron), p (proton), d (deuteron), α (alpha particle), γ (gamma ray), f (fission), and T (total). A more complete description of the reaction lists also the target and product nuclides, for example, ${}^{A}Z(n,\gamma)^{A+1}Z$. The reactions most useful for neutron spectroscopy are the total interaction; elastic scattering (n,n); radiative capture (n,γ); fission (n,f); inelastic scattering (n,n'); charged particle emission (n,p), (n,α), and (n,d); and three-body breakup or sequential decay $(n,2n)$ and (n,np). *See* NUCLEAR REACTION.

The resonances observed in these various reactions can be fitted by a theoretical formula to give parameters of the resonances (E_0, Γ, Γ_f Γ_γ, Γ_n, and so forth) which correspond to detailed properties of the excited states in the compound nucleus. For example, E_0 is the resonance energy; the fission width, Γ_f, is obtained from the fission cross section; the radiation width, Γ_γ, from the capture cross section; and so forth. The neutron width, Γ_n, can be obtained from the scattering cross section or the total cross section. The total width, Γ, can be obtained if the energy resolution is less than or equal to Γ. In addition, two other properties, the angular momentum of the neutron forming the resonance and the spin, J, of the state can often be determined. For narrow resonances where Γ (in eV) $\leq 0.05\sqrt{E_0\,(\text{in eV})}$, it is necessary to consider the Doppler broadening of resonances due to the thermal motion of the target nuclides.

Neutron cross sections. The measurement of a cross section for a particular reaction consists of measuring the number of such reactions produced by a known number of neutrons incident on a known number of target nuclides. When the probability of all neutron interactions with the target nucleus is small, the number of reactions of a particular process i per unit area and unit time using a beam of neutrons equals $(nv)(Nx)\sigma_i$. The quantity nv is the number of incident neutrons per unit area normal to their direction per unit time, N is the number of target nuclei per unit volume, x is the thickness of the target in the direction of the incident neutrons, and σ_i is the cross section per target nucleus for a particular reaction expressed in units of area. If the probability of all interactions is not small, the incident beam will be attenuated exponentially, as $\exp(-Nx\sigma_T)$, in passing through the sample, where Nx is the number of target nuclei per unit area normal to the beam and σ_T is the neutron total cross section.

The most common type of neutron cross-section measurement, which can usually be made with the

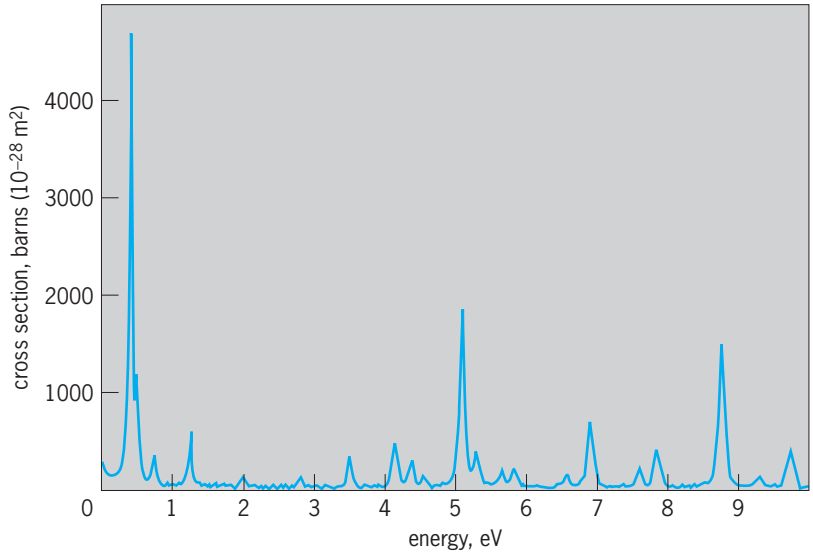

Fig. 1. Neutron total cross section of ^{231}Pa + neutron. The variation in sizes of the resonances and the nonuniform spacing of resonances are apparent.

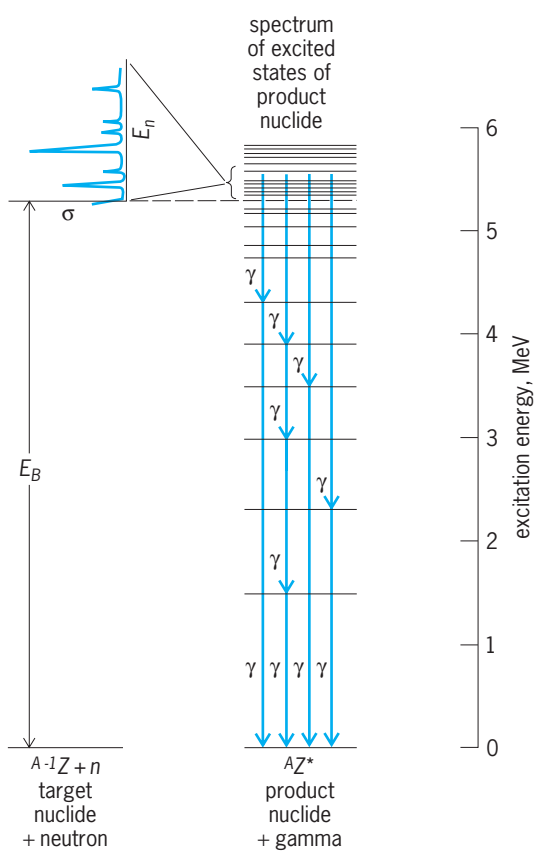

Fig. 2. Energy-level diagram for the product nucleus $^{A}Z^*$. The asterisk emphasizes that the product nucleus is in an excited state.

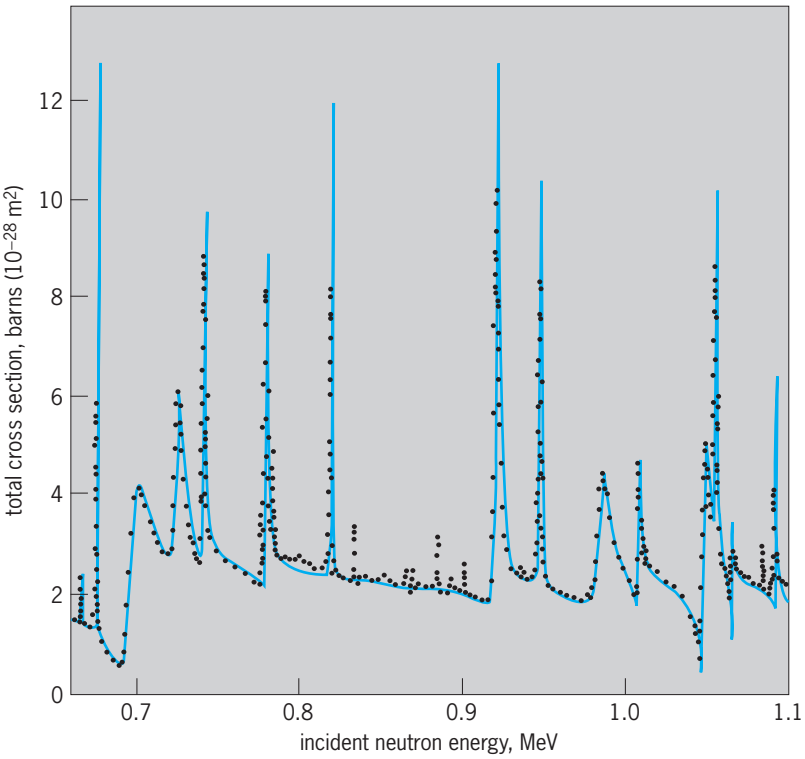

Fig. 3. Experimental neutron total cross section of sulfur compared to a theoretical fit. The fit does not include contributions from small resonances and minor isotopes of sulfur. The asymmetry of some resonances arises from the interference of resonance and potential scattering.

highest neutron energy resolution and usually with the most accuracy, is that of the total cross section. This measurement consists of measuring the transmission of a well-collimated beam of neutrons through a sample of known thickness; the transmission through the sample is simply the ratio of the intensity of the beam passing through the sample to that incident on the sample. The intensity of the incident beam is reduced in passing through the sample because the incident neutrons are absorbed or scattered by the target particles. The total cross section, σ_T, is determined from the equation $\sigma_T = -[\log_e \text{(transmission)}]/Nx$.

In order to measure partial cross sections, more elaborate equipment is needed, in general, than for total cross-section measurements, and the measurements are considerably more difficult. For example, to measure the differential elastic scattering cross section, it is necessary to measure the number of elastically scattered neutrons as a function of angle of the scattered neutron relative to that of the incident neutron. MeV-energy neutrons, in addition to being elastically scattered, can also lose energy when scattered from a target nucleus [inelastic scattering, $(n,n'\gamma)$]. Several techniques have been developed for determining inelastic cross sections, both by measuring the energy spectrum of the inelastically scattered neutrons and by measuring the energies and intensities of the gamma rays emitted from the excited nuclei. These cross sections can also be measured as a function of the angle relative to the direction of the incident neutron.

Techniques for neutron spectroscopy. Neutron spectroscopy can be carried out by two different techniques (or a combination): (1) by the use of a time-pulsed neutron source which emits neutrons of many energies simultaneously, combined with the time-of-flight technique to measure the velocities of the neutrons; this time-of-flight technique can be used for neutron measurements from 10^{-3} eV to about 200 MeV; (2) by the use of a beam of nearly monoenergetic neutrons whose energy can be varied in small steps approximately equal to the energy spread of the neutron beam; however, useful "monoenergetic" neutron sources are not available from about 10 eV to about 10 keV.

Time-of-flight neutron spectrometers. Time-of-flight neutron spectrometers are the most widely used spectrometers for most neutron cross-section measurements. The time-of-flight technique requires an intense pulse of neutrons which contains neutrons of many energies and a flight path to measure the velocities of the neutrons. Various detectors are placed at the end of the flight path depending on the type of cross-section measurement. Burst widths from 10^{-5} to 10^{-9} s and flight paths from 3 to 3000 ft (1 to 1000 m) have been used. The resolution of a time-of-flight spectrometer is often quoted in microseconds or nanoseconds per meter. The energy resolution ($\Delta E/E$) is equal to $2\Delta t/t$, where Δt is the time width of the neutron burst plus the time spread in the detector, and t is the time of flight of the neutron [t (in microseconds) = 72.3 × path length (in meters)/$\sqrt{E\,(\text{in eV})}$]. The time between pulses, the flight path length, and filters in the beam must be selected so that low-energy neutrons from previous pulses do not interfere with the high-energy

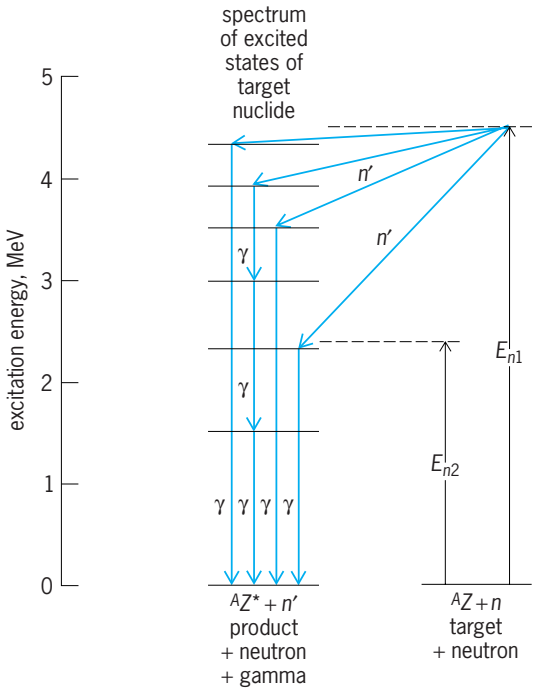

Fig. 4. Energy diagram of the target nuclide showing excitation of levels by the inelastic scattering of MeV neutrons.

neutrons from following pulses. By the use of multichannel storage, usually a computer, the complete neutron spectrum (or cross section) can be obtained in one measurement with good energy resolution over a broad energy range.

The most valuable neutron source for neutron time-of-flight spectroscopy is an electron or other charged-particle accelerator capable of producing intense pulses of neutrons of short duration (on the order of 10^{-9} s). Excellent neutron cross-section measurements can be made with these spectrometers from 0.01 eV to 200 MeV. For example, a beam of 140-MeV electrons incident on a tantalum target produces neutrons with an energy distribution that has a peak at about 1 MeV and extends up to about 80 MeV. The peak of the neutron distribution for protons or deuterons incident on a heavy target occurs at a higher neutron energy than for electrons; for deuterons a broad peak occurs at about half the energy of the deuterons. Lower-energy neutrons from these accelerators are obtained by placing a moderator (about 0.8 in. or 2 cm thick) around or near the target. The duration of these moderated neutron pulses in nanoseconds is approximately $2/\sqrt{E\,(\text{in MeV})}$. The flux distribution of these moderated neutrons approximately follows the relation $E^{0.8}$ down to thermal neutron energies. With a moderated neutron source the pulse repetition must be sufficiently low, depending on the flight path length, beam filter, and energy range, to prevent overlap of neutrons from successive pulses. Typical high-resolution results obtained using a moderated source are shown in Figs. 1 and 3. *See* PARTICLE ACCELERATOR.

Before the development of short-pulse accelerators, a mechanical chopper rotating at high speed in a well-collimated beam from a moderated fission reactor was used to produce bursts of electronvolt-energy neutrons. The neutron pulses were sufficiently short (about 10^{-6} s) and of sufficient intensity for measurements to be made up to a neutron energy of about 10^4 eV using neutron flight paths up to about 300 ft (100 m) in length. In order to produce pulses of only 10^{-6} s duration, the neutron beam had to be collimated to narrow slits (0.02 in. or 0.05 cm) to match the narrow slits through the chopper. Only when the rotation of the chopper was such that the slits in the rotor lined up with the slits in the collimator was a neutron beam with a broad energy spread passed. A fast-chopper time-of-flight spectrometer is particularly useful for transmission measurements on samples which are available in small amounts, since the sample only needs to be large enough to cover the beams passing through the narrow slits in the collimator.

For time-of-flight measurements in the energy region from about 10 keV to 1 MeV, a pulsed electrostatic accelerator using the Li^7 (p,n) reaction is capable of producing neutron pulses of short duration (10^{-9} s) with sufficient intensity for measurements with flight paths of a few meters. By selecting the proton-bombarding energy and a suitable target thickness, neutrons produced in the reaction at a given angle can have well-defined upper and lower energy limits. With no low-energy neutrons and short flight paths, rather high repetition rates of 10^6 Hz can be used and an energy resolution of about 1% can be realized.

The most intense pulsed neutron source used for neutron time-of-flight spectroscopy is that achieved from an underground nuclear explosion. The burst duration is about 80 nanoseconds, and the neutron distribution extends down to about 20 eV. Fission cross-section measurements have been made on very small samples of many radioactive heavy nuclides using such a source and an approximately 980-ft (300-m) flight path length. The availability of this source is obviously rather restricted, but it is unique for measurements on highly radioactive samples.

Neutron time-of-flight measurements have also been made using a pulsed fission reactor where the duration of burst is about 40 μs. Finally, subcritical boosters have been used to multiply the intensity of the neutron pulses from electron accelerators by factors of 10–200, which results in pulse durations of 0.08–4 μs.

Monoenergetic neutron spectrometers. The best technique for obtaining an intense beam of low-energy neutrons ($\leq$10 eV) with an energy spread of only about 1% is to use a crystal monochromator placed in a well-collimated beam of neutrons from a high-flux moderated fission reactor. If a single crystal (such as beryllium, copper, or lead) is properly oriented in a collimated neutron beam, neutrons of a discrete energy, E, will be elastically scattered from a particular set of planes of atoms in the crystal through an angle 2θ given by Bragg's law $n\lambda = 2d \sin\theta$. In this equation, the integer n is the order of the reflection, λ is the neutron wavelength, d is the spacing between the planes of atoms of the particular set in the crystal, and θ is the angle of incidence between the direction of the neutron beam and the set of planes of atoms being considered. The neutron wavelength λ in centimeters equals $0.286 \times 10^{-8}/\sqrt{E}$, where E is in eV. The energy of the diffracted beam can be continuously varied by changing the angle of the crystal. Measurements of many rare-earth nuclides and heavy nuclides have been made with crystal spectrometers up to 10 eV neutron energy. Capture gamma-ray spectra have also been studied as a function of neutron energy, specifically from different neutron resonances. *See* NEUTRON DIFFRACTION.

"Monoenergetic" neutrons in the energy range from a few keV to 20 MeV can be obtained by bombarding various thin targets with protons or deuterons from a variable-energy accelerator such as an electrostatic accelerator. The most useful (p,n) reactions to cover the energy range from a few keV to a few MeV are those on lithium and tritium targets. The (d,n) reaction on deuterium is useful from about 1 to 10 MeV, and the (d,n) reaction on tritium from 10 to 20 MeV. In the energy range up to 1 MeV an energy resolution of about 1 keV is possible, but this resolution is usually not adequate for neutron spectroscopy for neutrons with energies less than 10^4 eV. The measurement of a complete cross-section

spectrum up to 1 MeV may require 1000 sequential measurements at slightly different neutron energies. Monoenergetic neutron sources are also useful for measurements such as activation, which cannot be done with the time-of-flight technique.

Applications. Neutron spectroscopy has yielded a mass of valuable information on nuclear systematics for almost all nuclides. The distribution of the spacings between nuclear levels and the average of these spacings have provided valuable tests for various nuclear theories. The properties of these levels, that is, the probabilities that they decay by neutron or gamma-ray emission, or by fission, and the averages and distribution of these probabilities have stimulated much theoretical effort.

In addition, knowledge of neutron cross sections is fundamental for the optimum design of thermal fission power reactors and fast neutron breeder reactors, as well as fusion power reactors now in the conceptual stage. Cross sections are needed for nuclear fuel materials such as ^{235}U or ^{239}Pu, for fertile materials such as ^{238}U, for structural materials such as iron and chromium, for coolants such as sodium, for moderators such as beryllium, for shielding materials such as concrete. The optimum choice of materials for the energy region under consideration is critical to the success of the project and is of great economic significance. *See* NUCLEAR STRUCTURE; REACTOR PHYSICS.

John A. Harvey

Bibliography. Y. A. Izyumov and N. A. Chernoplekov, *Neutron Spectroscopy*, 1994; S. F. Mughabghab, M. Diradeenam, and N. E. Holden (eds.), *Neutron Cross Sections*, vol. 1, pt. A 1981, pt. B 1984.

Neutron star

A star containing about $1^1/_2$ solar masses of material compressed into a volume approximately 6 mi (10 km) in radius. (1 solar mass equals 4.4×10^{30} lbm or 2.0×10^{30} kg.) Neutron stars are one of the end points of stellar evolution and are the final states of stars that begin their lives with considerably more mass than the Sun. The density of neutron star material is 10^{14} to 10^{15} times the density of water and exceeds the density of matter in the nuclei of atoms. Neutron stars are pulsars (pulsating radio sources) if they rotate sufficiently rapidly and have strong enough magnetic fields.

Neutron stars play a role in astrophysics which extends beyond their status as strange, unusual types of stellar bodies. The interior of a neutron star is a cosmic laboratory in which matter is compressed to densities which are found nowhere else in the universe. Precise measurements of the rotation of neutron stars can probe the behavior of matter at such densities. Neutron stars in double-star systems can emit x-rays when matter flows toward the neutron star, swirls around it, and heats up. Neutron stars are almost certainly formed in supernova explosions, events in which a dying star becomes more luminous than an entire galaxy, up to 10^{12} times as powerful as the Sun. A few pulsars are found in double-star systems, and careful timing of the pulses they emit can test Einstein's general theory of relativity. *See* GRAVITATION; RELATIVITY.

Dimensions. The radius of a neutron star, while not precisely known, is about 6 mi (10 km). A few neutron stars exist in double-star systems, wherein it is possible to measure the strength of the gravitational pull of the neutron star on the other star in the system and hence determine the mass of each. The masses of these neutron stars are slightly more than the mass of the Sun, with the measured values ranging from 1.4 to 1.8 solar masses. *See* BINARY STAR.

The highest mass that a neutron star can have is about 2 solar masses. More massive neutron stars might possibly exist if current ideas about the behavior of neutron star matter turn out to be wrong. Objects that end their lives with masses higher than this neutron star mass limit will become black holes. *See* BLACK HOLE.

Internal structure. The state of matter beneath the surface of a neutron star is completely different from matter on the surface of the Earth and cannot be reproduced in terrestrial laboratories. The gaseous surface, only several feet thick, is different from ordinary gases, for the structure of atoms is dominated by the strong magnetic field that certainly exists in all pulsars and probably is present in all neutron stars. The magnetic field distorts the shape of the electron orbits which form the outer part of the atoms. Several feet below the surface, the material solidifies. The solid crust is about 0.6 mi (1 km) thick (see **illus.**) and is approximately 10^{17} times stiffer than steel. At greater depths, the densities increase, and a progressively greater fraction of the electrons are forced into atomic nuclei, where they combine with protons to form neutrons. *See* ATOMIC STRUCTURE AND SPECTRA.

Most of the interior of a neutron star consists of matter which is almost entirely composed of neutrons. In the bulk of the star, this matter is in a superfluid state, where circulation currents can flow without resistance. In rotating superfluids maintained in physics laboratories, small vortices form; such vortices may also exist in neutron stars. This material is under pressure, since it must be able to support the tremendous weight of the overlying layers at each

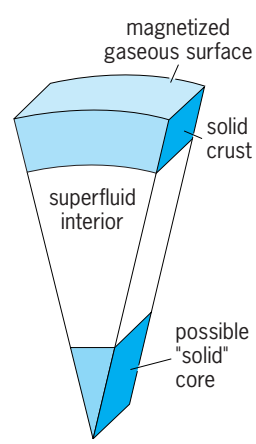

Cross section of a neutron star, with various layers.

point in the neutron star. This pressure, called degeneracy pressure, is caused by the close packing of the neutrons rather than by the motion of the particles. As a result, neutron stars can be stable no matter what the internal temperature is, because the pressure that supports the star is independent of temperature. *See* SUPERFLUIDITY.

The central regions of neutron stars are poorly understood. For example, the role played by subnuclear particles such as the *K* meson, pi meson, or other hyperons is unknown. It is remotely possible that quark matter, material composed of the postulated fundamental particles of matter, exists in the cores of neutron stars. Claims of some astronomers to have actually detected strange quark stars have not generally been accepted. *See* DELTA RESONANCE; ELEMENTARY PARTICLE; NUCLEAR PHYSICS; QUARKS.

Observations. Theoretical calculations form the basis for the above description of a neutron star. A variety of astronomical observations may eventually confirm or disprove various aspects of this model. The pulses from a pulsar (a rotating neutron star) can be timed very precisely and can suggest the way that neutron stars change their shape as they spin more slowly. A single, isolated neutron star loses energy in the form of high-speed particles emitted from its surface, and its rotation rate slows down. A newly formed pulsar spins very fast—the pulsar in the Crab Nebula, formed a little less than 1000 years ago as observed from Earth, rotates 30 times every second. Eventually it will rotate more slowly, and its structure will readjust, occasionally producing abrupt changes in the rotation period which are called glitches. *See* CRAB NEBULA.

Telescopes launched above the atmosphere have detected x-rays and extreme ultraviolet radiation from the hot surfaces of a number of neutron stars. In some cases, these neutron stars are extremely young and hot, the product of recent supernova explosions. Unfortunately, it is not yet possible to clearly determine the mass and radius of a neutron star and compare those results to observations, as is the case for white dwarf stars. *See* SATELLITE (ASTRONOMY); ULTRAVIOLET ASTRONOMY; WHITE DWARF STAR; X-RAY ASTRONOMY.

Some neutron stars exist in double-star systems, where matter flows from one star to the neutron star, forms a whirling disk around it, and causes it to rotate faster as it gains more mass. These neutron stars emit pulses of x-rays, which can be detected and timed by satellite observatories orbiting above the Earth's atmosphere. Sudden and irregular changes in the arrival times of these x-ray pulses show that solid matter exists somewhere in the interior of neutron stars, but the data are sufficiently poor that it is not yet possible to test detailed models of neutron stars.

In the late 1980s, the *EXOSAT* satellite, launched by the European Space Agency and containing an x-ray telescope, detected what were described as quasiperiodic oscillations from a number of x-ray-emitting double stars. The x-ray intensity increased and decreased with a reasonably well-defined period, though the frequency of these oscillations changed from one observation time to another. It is possible that these oscillations can indicate something about the neutron star itself, though it is more likely that they originate from the disk surrounding the neutron star and that analysis of them will provide information on accretion disk physics rather than neutron star physics.

Beginning in 1979, space satellites, some originally intended to monitor nuclear bomb tests, detected sudden blasts of gamma rays coming from space. Their origin was a mystery. Many explanations were proposed, but none gained support. In the 1990s, new data led to the widespread acceptance of the "magnetar" explanation. Magnetars are neutron stars with fields of 10^{10} to 10^{11} tesla (10^{14} to 10^{15} gauss), fields 10 to 100 times stronger than the fields of normal neutron stars. *See* GAMMA-RAY BURSTS.

Some neutron stars rotate so fast that their pulse periods are milliseconds, rather than the pulse periods of typical pulsars which are roughly seconds. Someone listening to the pulses from these millisecond pulsars would hear a musical tone rather than a series of drumbeats (which would be heard from the conventional pulsars). The current interpretation of these rapid rotation rates is that most millisecond pulsars are in double star systems, and have been spun up to high speeds because they accrete matter. One of these ultrafast pulsars was the first object outside the solar system to have a planet detected around it. *See* PLANET.

Origin. In the 1930s, Fritz Zwicky first suggested that isolated neutron stars originate in supernova explosions. Stars more massive than approximately 10 times the mass of the Sun end their lives in a tremendous explosion. It is now clear that in many cases the remnant of this explosion is a tiny neutron star, whose formation provided the energy for the supernova. The discovery of a pulsar in the Crab Nebula, a known supernova remnant, strongly supported the connection between neutron stars and supernovae. The 1987 supernova explosion in the Large Magellanic Cloud both confirmed the genesis of neutron stars in supernovae and provided some important new details in the story. When the light from this supernova explosion arrived at the Earth, so did a burst of subnuclear particles called neutrinos. These neutrinos were observed by underground detectors located in the United States and Japan. The detection of these supernova neutrinos firmly cemented the link between supernovae and neutron stars, and also is helping astronomers understand how the formation of a neutron star in the center of a massive star is connected to these celestial explosions. *See* GRAVITATIONAL COLLAPSE; NEUTRINO; PULSAR; STELLAR EVOLUTION; SUPERNOVA. Harry L. Shipman

Bibliography. R. Kippenhahn, *100 Billion Suns*, trans. by J. Steinberg, 1983, reprint 1993; R. P. Kirshner, *The Extravagant Universe: Exploding Stars, Dark Energy, and the Accelerating Cosmos*, 2002; L. Marschall, *The Supernova Story*, 1988, paper 1994; S. L. Shapiro and S. A. Teukolsky, *Black Holes, White Dwarfs, and Neutron Stars*, 1983; D. J. Yakovlev and C. J. Pethick, Neutron star cooling, *Ann. Rev. Astron. Astrophy.*, 42:159–210, 2004.

New Zealand

A landmass situated in the Southern Hemisphere, bounded by the South Pacific Ocean to the north, east, and south and the Tasman Sea to the west, with a total land area of 103,883 mi^2 (269,057 km^2). The exposed landmass represents about one-quarter of a subcontinent, with three-quarters submerged. This long, narrow, mountainous country, oriented northeast to southwest, consists of two main islands, surrounded by a much greater area of crust submerged to depths reaching 1.2 mi (2 km) [**Fig. 1**].

Geology. The major modern geological structural features of the subcontinent have developed over the last 25 million years, during which the New Zealand region formed from interactions along the boundary between the Australia and Pacific crustal plates. Off the east coast of North Island, the Pacific plate of oceanic crust is subducting (sinking) beneath that edge of the Australia plate that forms North Island. The results of subduction include earthquakes at progressively greater depths westward, folding and faulting of rocks beneath the eastern continental shelf and ranges, and volcanic activity along an arc that extends from Mount Ruapehu 9174 ft (2797 m) through the Taupo and Rotorua regions to White Island (**Fig. 2**). *See* EARTHQUAKE; FAULT AND FAULT STRUCTURES.

The plate boundary subduction zone terminates in the northeast of South Island, from where the boundary extends to the southwest as the Alpine Fault, along which plate motion is in the form of lateral shearing. The Alpine Fault extends offshore north of Fiordland, where plate convergence results in subduction of the edge of the Australia plate beneath the New Zealand subcontinent, and formation of a southern seismic zone. *See* SUBDUCTION ZONES.

Plate boundary processes are responsible for the crustal deformations that have resulted in still-active uplift of the Southern Alps of South Island with

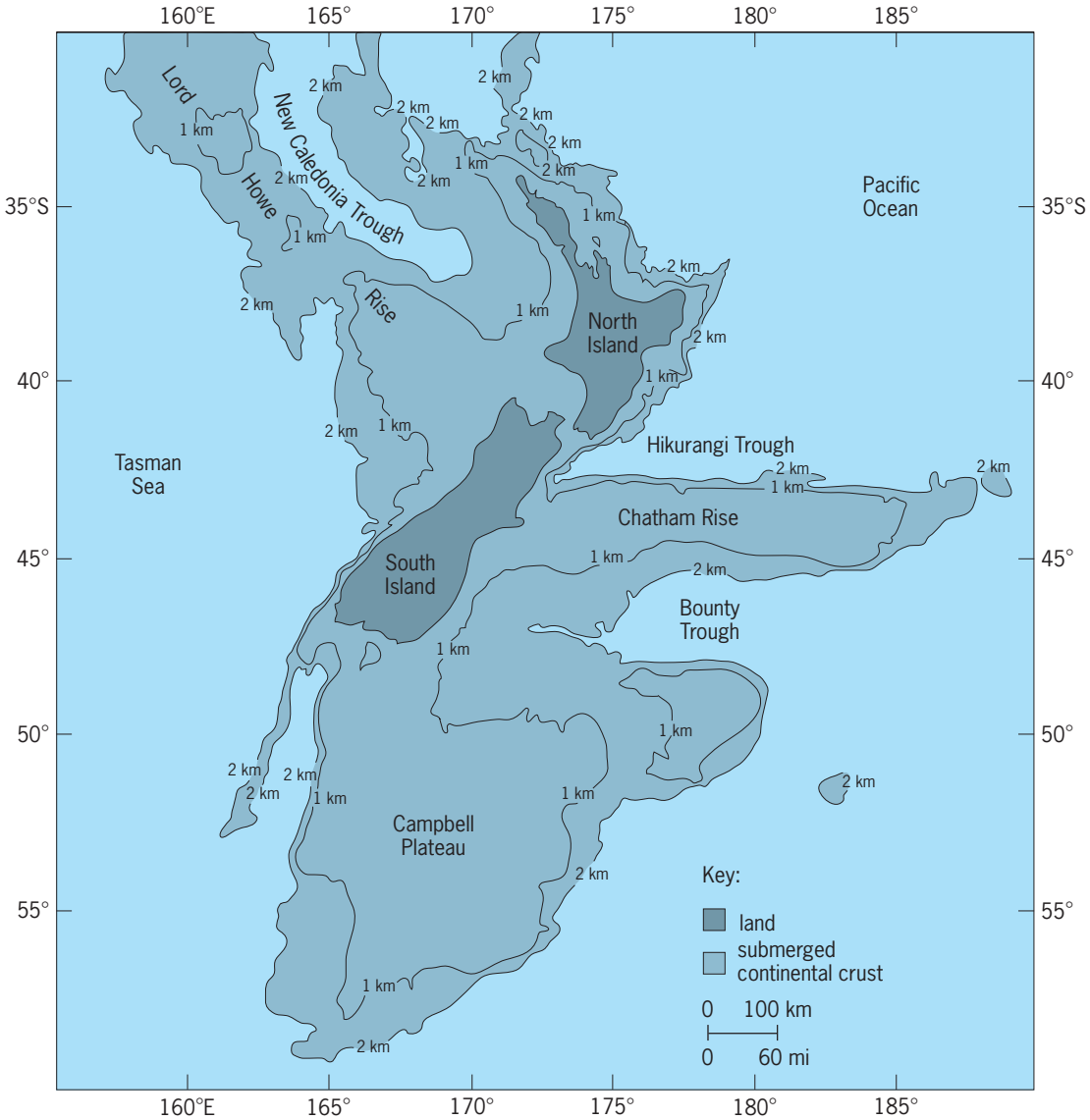

Fig. 1. New Zealand subcontinent and its exposed landmass. Contours represent water depth. (*After I. Wards, ed., New Zealand Atlas, Government Printer, 1976*)

attendant erosion of the young cover rocks and exposure of older basement rocks of graywacke (sandstone and argillite) and schist. In contrast, the subduction of the Pacific plate below eastern North Island has produced uplift and inversion of the overlying sedimentary basins containing mudrocks and limestones. North Island rocks are therefore of three groups: Cretaceous to Cenozoic age sedimentaries, Cenozoic volcanics, and exposed basement graywackes in the zones of greatest uplift and highest relief in the ranges extending northward from Wellington to eastern Bay of Plenty. *See* GRAYWACKE; LIMESTONE; SANDSTONE; SCHIST.

The relief of New Zealand has been created by tectonic forces. Two-thirds of the country has an elevation between 650 and 3600 ft (200 and 1100 m), about one-sixth lies below 660 ft (200 m), and the remaining sixth is above 3600 ft (1100 m). The highest mountains in South Island reach over 8200 ft (2500 m), and 18 peaks reach over 9840 ft (3000 m); Mount Cook, the highest, is 12,346 ft (3764 m). The alpine relief is rugged, with deep glaciated valleys; and from Milford Sound northeastward to Arthur's Pass, there are active high-basin glaciers, and around Mount Cook valley glaciers that form the basis of a tourism and alpine sports industry. *See* OROGENY; PLATE TECTONICS.

Soils. South Island lowlands are either alluvial plains as in Otago, Southland, and Nelson, or glacial outwash fans as in Westland and Canterbury. North Island lowlands such as Hawke's Bay, Wairarapa, and Manawatu are alluvial; the Waikato, Hauraki, and Bay of Plenty lowlands occupy structural basins that contain large volumes of reworked volcanic debris from the central volcanic region. The alluvial lowlands of both main islands form the most agriculturally productive areas of the country. *See* PLAINS.

Lowland soils in South Island may develop in alluvium, or loess derived from dry river beds, and in North Island in alluvium or beds or volcanic ash and loess. Hill-country soils have, generally, low fertility and shallow depths. Soils of the lower hills and plateaus form the basis of much pastoral farming and, increasingly, of plantation forests; these soils are greatly influenced in their characteristics by rainfall regimes and the original natural vegetation. *See* LOESS; SOIL.

Climate. The climate of New Zealand is influenced by three main factors: a location in latitudes where the prevailing airflow is westerly; an oceanic environment; and the mountain chains, which modify the weather systems as they pass eastward, causing high rainfalls on windward slopes and sheltering effects to leeward.

Weather is determined mostly by series of anticyclones and troughs of low pressure that produce alternating periods of settled and variable conditions. Westerly air masses are occasionally replaced by southerly airstreams, which bring cold conditions with snow in winter and spring to areas south of 39° S, and northerly tropical maritime air, which brings warm humid weather to the north and east coasts. *See* MARITIME METEOROLOGY.

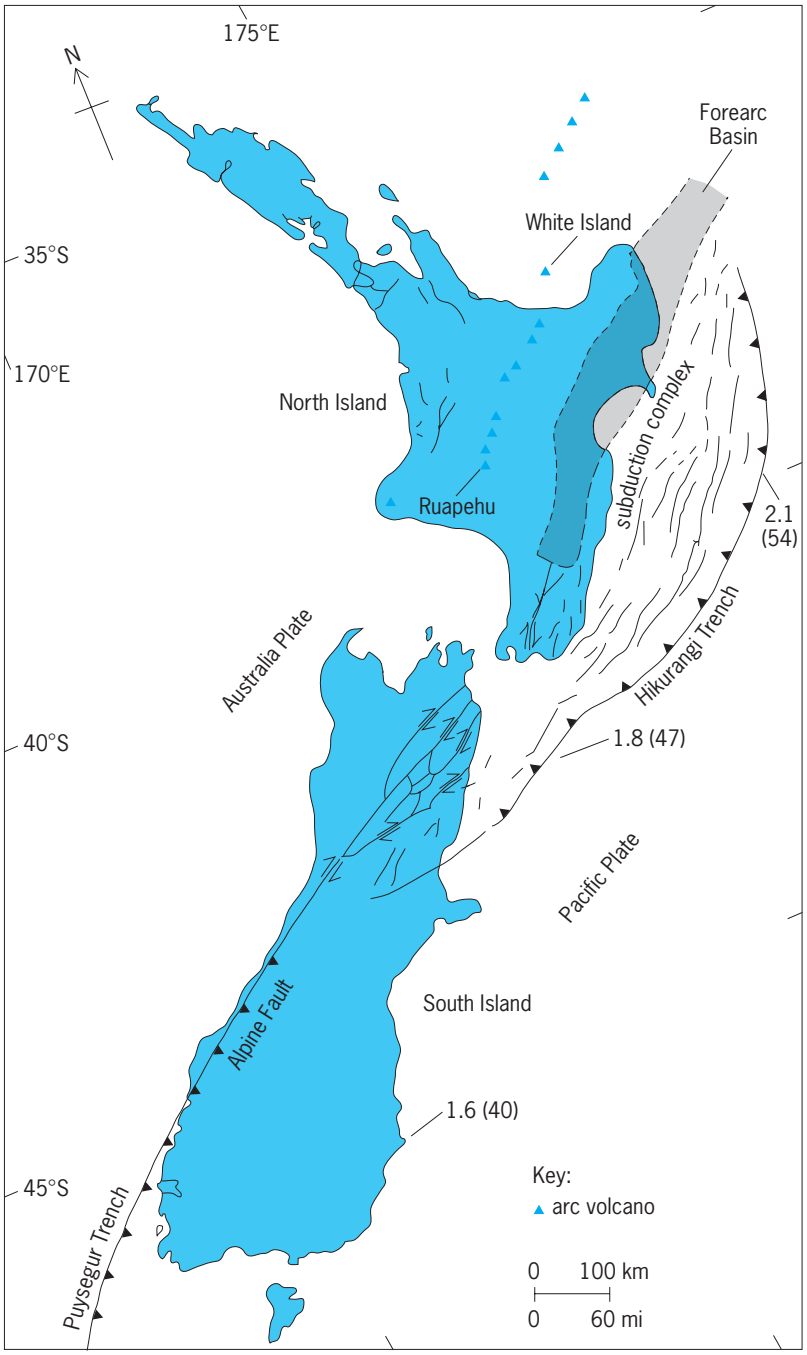

Fig. 2. Some structural features of New Zealand, including the boundary between the Australia and Pacific plates, the Alpine Fault, the line of arc volcanoes, and the zone of inverted sedimentary basins (Forearc Basin) of eastern North Island. The numerical values are rates of convergence of the crustal plates in inches (millimeters) per year. (*After P. J. J. Kamp, Tectonic architecture of New Zealand, in J. M. Soons and M. J. Selby, Landforms of New Zealand, 2d ed., Longman Paul, 1992*)

Rainfall over the open ocean is commonly in the range of 25–30 in. (600–800 mm) per year, but on land is 16–470 in. (400–12,000 mm) per year (**Fig. 3**), with the highest rainfalls being on the western windward slopes of the mountains, and the lowest on the eastern basins in the lee of the Southern Alps in Central Otago and south Canterbury. Annual rain days are at least 130 for most of North Island, but the South Island the totals are far more variable, with over 200

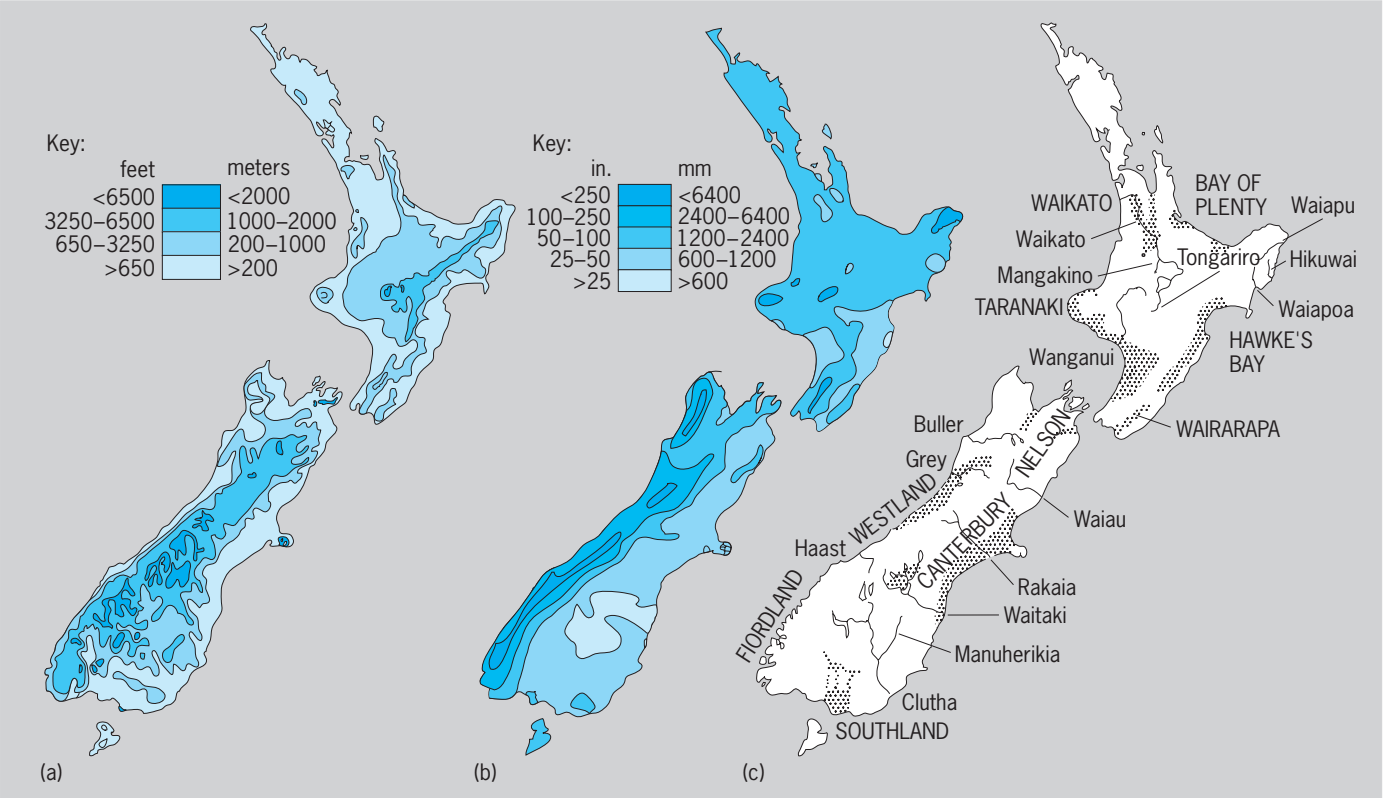

Fig. 3. Maps showing (*a*) relief, (*b*) annual rainfall, and (*c*) main rivers, and sedimentary lowlands of New Zealand. Note the close relationship between relief and rainfall. (*After I. Wards, ed., New Zealand Atlas, Government Printer, 1976*)

occurring in Fiordland, 180 on the west coast, and fewer than 80 in Central Otago. Summer droughts are relatively common in Northland, and in eastern regions of both islands. *See* DROUGHT; PRECIPITATION (METEOROLOGY).

Mean annual temperatures decrease progressively from north to south: 59°F (15°C) in Northland, 55°F (13°C) in Wellington, 50°F (10°C) in Southland; some inland parts of South Island have averages of 46°F (8°C). January and February are the warmest months and July the coldest. Highest summer temperature exceed 86°F (30°C) in Central Otago, lowest temperatures are rarely below 14°F (−10°C) except in the mountains. Average annual duration of bright sunshine is greatest on the northern lowlands of South Island, where it exceeds 2400 h around Nelson and Blenheim, and on the eastern lowlands of North Island and Bay of Plenty.

Droughts, springtime air frosts, and hailstorms are the major common climatic hazards for the farming industry, but floods associated with prolonged intense rainstorms are the major general hazard.

Rivers and lakes. The rivers of New Zealand have a total discharge of about 70 mi^3 (300 km^3) per year. Some 12 mi^3 (50 km^3) of water is stored in perennial snow and glacier ice. Flow variability and specific yield (flow per unit area of catchment) is very high by comparison with most rivers in the world. The Clutha River, with a mean flow of 20,000 ft^3 (570 m^3)/s is the largest river in New Zealand, and the Waikato is the largest in North Island. The highest specific discharges are those of rivers draining westward from the Southern Alps, where steep mountainous catchments, lack of lake storage, exposed rock, thin soils, and very high rainfall produce very large flood flows (**Table 1**); the Grey River has a specific discharge that is three times higher than those of the two longer North Island rivers, Waikato and Wanganui, which have larger catchments but receive lower rainfall.

In an average year, New Zealand's rivers carry 4.4×10^6 tons of sediment from the land to the ocean. Two major factors account for the extraordinarily high suspended sediment yields of many rivers (**Table 2**): amount and intensity of rainfall, and catchment geology with land use. These factors are illustrated by the yields for the Cropp River with its mountainous catchment and very high rainfall, and the Waiapu River with erodible sandstone and mudstone, compared with the low yields from the Manuherikia with low rainfall but erodible rock, and the Mangakino with moderate rainfall and very permeable rock. *See* RIVER.

Seven hundred seventy-six lakes, with lengths exceeding 0.3 mi (0.5 km), have been cataloged in New Zealand; of these 291 were formed by glacial action and occur in South Island, while 30 are of volcanic origin and occur in the central North Island. The other major categories are dune, riverine, and artificial impoundments. Because of the importance of

TABLE 1. Major rivers ranked by water discharge*

River	Mean annual discharge		Catchment area		Specific discharge		Length	
	ft^3/s	m^3/s	mi^2	km^2	$ft^3/(s)(mi^2)$	$m^3/(s)(km^2)$	mi	km
Clutha	20,128	570	7,947	20,582	3	0.03	200	322
Waiau	15,432	437	3,141	8,134	5	0.05	105	169
Buller	15,114	428	2,452	6,350	7	0.07	110	177
Waitaki	12,960	367	3,768	9,760	4	0.04	130	209
Grey	11,900	337	1,479	3,830	9	0.09	75	120
Waikato	11,547	327	4,400	11,395	3	0.03	264	425
Wanganui	7,910	224	2,565	6,643	3	0.03	181	290

*After M. Mosley (ed.), *Waters of New Zealand*, 1992.

tourism and hydroelectricity, these lakes not only are significant features of the landscape but also are of major economic importance. *See* LAKE.

Flora and fauna. New Zealand gained its earliest flora and fauna from Gondwanaland before the separation that occurred in the Cretaceous Period. Consequently its plants and animals are highly endemic—over 80% for the flowering plants and well over 90% for the arthropods. Separation, as well as the opening of the Tasman Sea, was early enough to exclude snakes and marsupial mammals but late enough to include podocarps, tree ferns and other ferns, southern beech (*Nothofagus* sp.), and conifers related to the modern kauri (*Agathis australis*). The only terrestrial reptiles in the New Zealand fossil record are the tuatara (*Sphenodon*), which still survives on small offshore islands, and many species of gecko and skink; the only indigenous mammals are two species of bat. *See* CRETACEOUS.

The modern bird population is extremely varied, including many migratory species, some Australian recent settlers, and many European species introduced since the mid-nineteenth century. Native species had no predators and consequently there were many flightless, or virtually flightless, as well as flying species. But since the arrival of the first Polynesian settlers about 1000 years ago, and the subsequent early nineteenth century Europeans, many species of flightless birds have become extinct as a result of hunting, destruction of habitat, predation by the Polynesian dog and rat (*Rattus exulans* or *Kiore*) and then the European dogs and ships' rat (*R. norvegicus*), and by mustelids. Both Polynesian forest burning and the later European destruction of forest have had catastrophic consequences for native birds, including all 24 species of moa (Dinornithiformes) and even the kiwi—the national symbol—is under threat because of loss of habitat and predation by dogs. *See* RATITES.

The natural vegetation of New Zealand below the tree line is mostly evergreen forest, with kauri–podocarp (rimu, miro, totara)–hardwood (rewarewa, tawari) north of latitude 38°S; podocarp-hardwood in central North Island; podocarp-hardwood-beech in southern North Island and the wetter parts of South Island (with variations of species with soils and climate); and beech forests on poorer soils, and cooler and steeper sites of higher altitudes on both main islands.

Forest clearing since about 1840 has removed much of the 54,040 mi^2 (14 million hectares) of forest that was present at that date. Forests cover about 28%, or 29,680 mi^2 (8 million hectares), of New Zealand's land area; of this, about 23,900 mi^2 (6 million hectares) are in natural forest, and 5780 mi^2 (1.5 million hectares) in plantations of introduced species. Of the plantations, 89% is *Pinus radiata*; only 1% is in hardwood species. The *P. radiata* is the basis of construction timber and of paper and board manufacturing.

TABLE 2. Selected rivers ranked by suspended sediment yield from their catchments*

River	Suspended sediment yield		Catchment area		Precipitation		Catchment rock type
	$tons/(mi^2)(y)$	metric $tons/(km^2)(y)$	mi^2	km^2	in./y	mm/y	
Cropp	83,768	29,600	11	29	393	10,070	Schist
Waiapu	56,515	19,970	532	1,378	94	2,400	Sandstone and mudstone
Hikuwai	39,308	13,890	119	307	74	1,900	Sandstone and mudstone
Haast	36,042	12,736	394	1,020	253	6,500	Schist
Waipaoa	16,515	5,836	611	1,582	62	1,600	Sandstone and siltstone
Rakaia	4,644	1,641	1,019	2,640	117	3,000	Graywacke and argillite
Tongariro	1,570	555	298	772	78	2,000	Graywacke and andesite
Grey	1,562	552	248	642	117	3,000	Granite
Wanganui	922	326	2,565	6,643	70	1,800	Sandstone and siltstone
Manuherikia	99	35	786	2,036	32	830	Schist
Mangakino	99	35	144	373	55	1,400	Ignimbrite

*After M. P. Mosley (ed.), *Waters of New Zealand*, 1992.

Natural resources and the economy. The New Zealand economy is heavily dependent on the natural resources soil, water, and plants. It has few exploitable minerals, but it does possess a climate generally favorable for agriculture, pastoral farming, renewable forestry, and tourism. With a relatively small population (3.4 million people), much of its manufacturing is concerned with processing produce from the land and surrounding seas, and supplying the needs of those industries.

Natural hazards. Because of its high relief and its location on an active crustal plate boundary in the zone of convergence between Antarctic air masses and tropical air masses, New Zealand is prone to high-intensity and high-frequency natural hazards—earthquakes, volcanic eruptions, large and small landslides, and floods.

Earthquakes are felt throughout the country, but they are most common and most severe in a central seismic zone that includes the northern half of South Island and the southern half of North Island. Volcanic eruptions are of limited magnitude and are confined to the central North Island volcanic zone, where Ngauruhoe, Ruapehu, and White Island are mildly active; in 1886, Tarawera erupted more violently. Of far greater potential power and hazard are ignimbrite-producing eruptions from calderas of the central North Island, the last of which occurred 1900 years ago from the Taupo caldera. *See* CALDERA; VOLCANO.

Very large landslides in the Southern Alps and in the eastern areas of Marlborough (South Island) and North Island have tectonic origins due to land tilting and earthquakes in highly sheared or weak rocks. Such landslides occur infrequently, three to four per thousand years, but each may involve 3–500 $\times$ 10^6 m^3 (1–1.8 $\times$ 10^{10} ft^3) of rock. Climatically induced landslides are far more frequent, but involve only a few tens of cubic meters of regolith material in each failure. They may, however devastate some tens or hundreds of hectares of land in a single storm. Such failures are most common on the North Island and northeastern South Island hills. *See* LANDSLIDE.

Since the end of the nineteenth century, natural hazards have been limited in their impact on New Zealand by the wide dispersal of the small human population and, with the exception of the Napier earthquake of 1931, the relative low intensity, and location, of the largest landslides and volcanic eruptions. Fewer then 500 people have been killed in all extreme events of the last century. M. J. Selby

Bibliography. M. Gage, *Legends in the Rocks: An Outline of New Zealand Geology*, 1980; G. Kuschel (ed.), *Biogeography and Ecology in New Zealand*, 1975; L. Molloy, *Soils in the New Zealand Landscape: The Living Mantle*, 1988; M. P. Mosley (ed.), *Waters of New Zealand*, 1992; New Zealand Department of Statistics, *New Zealand Official Yearbook*, annually; J. M. Soons and M. J. Selby (eds.), *Landforms of New Zealand*, 2d ed., 1992; I. Wards (ed.), *New Zealand Atlas*, 1976.

Newcastle disease

A viral infection that affects the digestive, intestinal, and respiratory tracts and the neurological system of birds. The causative agent is an enveloped ribonucleic acid (RNA) virus that is classified as a paramyxovirus. *See* PARAMYXOVIRUS.

Newcastle disease occurs in five forms based on a virulence in chickens ranging from inapparent infection to severe disease and death. Viscerotropic-velogenic Newcastle disease causes a very severe infection, producing hemorrhagic lesions in the intestinal tract and high mortality. The neurotropic-velogenic type is also highly lethal and produces neurologic and respiratory signs in infected birds. The mesogenic form causes an acute respiratory or neurologic infection that may be lethal only in young birds. The lentogenic type is a mild or inapparent respiratory infection of chickens. The last group includes the viruses causing inapparent or asymptomatic infections of the digestive tract.

The wide susceptibility of avian species to infection with Newcastle disease has complicated control. Newcastle disease is spread worldwide by the international transportation of live birds disseminating the virus. In many countries, the velogenic form has been eradicated in poultry by quarantine and slaughter. More difficult to control are reservoirs of natural infection in wild free-flying birds. Waterfowl (ducks and geese) are naturally infected with the avirulent intestinal strain of the virus. Newcastle disease has been identified in racing pigeons and has spread worldwide through racing activity and the extensive trade in these birds. Another important reservoir is exotic pet birds, particularly those from tropical regions where the virulent virus is endemic. These birds are more resistant and can be asymptomatically infected with the virus virulent for poultry. Countries free of velogenic Newcastle disease impose quarantine and testing of imported birds to prevent release of infected birds.

Control of and protection from Newcastle disease can be achieved by the correct use of vaccines. Lentogenic and some mesogenic strains are used to produce vaccines that can be administered by aerosol, intranasal drops, or intramuscular injection, or as an additive to the drinking water. *See* ANIMAL VIRUS. Mary Lynne Vickers

Bibliography. D. J. Alexander (ed.), *Newcastle Disease: Developments in Veterinary Virology*, 1988; B. W. Calnek (ed.), *Diseases of Poultry*, 10th ed., 1997.

Newtonian fluid

A fluid whose stress at each point is linearly proportional to its strain rate at that point. The concept was first deduced by Isaac Newton and is directly analogous to Hooke's law for a solid. All gases are newtonian, as are most common liquids such as water, hydrocarbons, and oils. *See* HOOKE'S LAW; STRESS AND STRAIN.

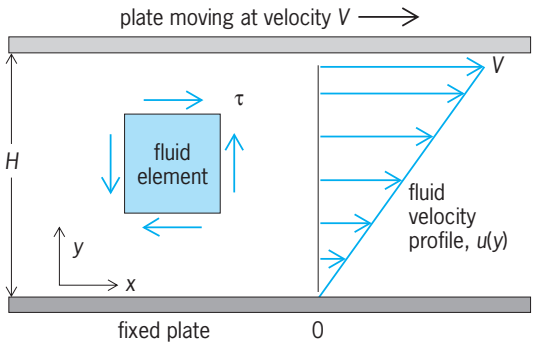

A fluid sheared between two plates. The resulting strain rate equals *V/H*.

A simple example, often used for measuring fluid deformation properties, is the steady one-dimensional flow $u(y)$ between a fixed and a moving wall (see **illus.**). The no-slip condition at each wall forces the fluid into a uniform shear strain rate ϵ, given by Eq. (1), which is induced by a uniform shear

$$\varepsilon = \frac{\partial u}{\partial y} = \frac{V}{H} \tag{1}$$

stress τ. Here V is the speed of the moving wall, H is the perpendicular distance between the walls, and u is the fluid velocity at distance y from the fixed wall.

If the fluid is newtonian, the experimental plot of τ versus ϵ will be a straight line. The constant of proportionality is called the viscosity μ of the fluid, as stated in Eq. (2). For consistency, the dimensions of

$$\tau = \mu\varepsilon \tag{2}$$

μ must be (mass)/(length)(time), or kg/(m · s). The viscosity coefficients of common fluids vary by several orders of magnitude (see **table**).

When flowing near a solid boundary, all viscous fluids, except for extremely rarefied gases, stick to that surface and move at the velocity of the surface (see illus.). This is the no-slip condition, used in formulating mathematical models of fluid flow.

If the plot of τ versus ϵ forms a curved line, the fluid is termed non-newtonian and can exhibit more complex effects. *See* NON-NEWTONIAN FLUID.

The simple one-dimensional relation, Eq. (2), can easily be extended to fully three-dimensional unsteady viscous flows. When the result is combined with the momentum principle, the fundamental equation for newtonian fluid flow emerges. *See* BOUNDARY-LAYER FLOW; FLUID-FLOW PRINCIPLES; NAVIER-STOKES EQUATION.

Viscosities of common fluids

Fluid	Viscosity (μ), kg/(m·s)*
Air	1.80×10^{-5}
Methanol	5.98×10^{-4}
Water	0.001
Mercury	0.00156
SAE 10 oil	0.104

* At 20°C (68°F) and 1 atmosphere (101.3 kilopascals).

The viscosity μ of a fluid is a true thermodynamic property and can be measured or tabulated for each substance. Generally, μ for liquids decreases sharply with temperature and increases moderately with pressure. The viscosity of gases increases moderately with temperature and increases slightly with pressure. *See* FLUIDS; VISCOSITY. Frank M. White

Bibliography. F. S. Sherman, *Viscous Flow*, 1990; F. M. White, *Viscous Fluid Flow*, 2d ed., 1991.

Newton's laws of motion

Three fundamental principles which form the basis of classical, or newtonian, mechanics. They are stated as follows:

First law: A particle not subjected to external forces remains at rest or moves with constant speed in a straight line.

Second law: The acceleration of a particle is directly proportional to the resultant external force acting on the particle and is inversely proportional to the mass of the particle.

Third law: If two particles interact, the force that is exerted by the first particle on the second particle (called the action force) is equal in magnitude and opposite in direction to the force that is exerted by the second particle on the first particle (called the reaction force).

The first law, sometimes called Galileo's law of inertia, can now be regarded as contained in the second. At the time of its enunciation, however, it was important as a negation of the aristotelian doctrines of natural placement and continuing force.

The third law, sometimes called the law of action and reaction, was also to some extent established prior to Newton's statement of it. However, Newton's formulation of the three laws as a mutually consistent set, with the nature of force clearly defined in the second law, provided the basis for classical dynamics.

The newtonian laws have proved valid for all mechanical problems not involving speeds comparable with the speed of light (approximately 186,000 mi/s or 300,000 km/s) and not involving atomic or subataomic atomic particles. The more general classical methods of Lagrange and of Hamilton are elaborations of the newtonian principles. *See* DYNAMICS; FORCE; HAMILTON'S EQUATIONS OF MOTION; LAGRANGE'S EQUATIONS; MOTION. Dudley Williams

Niacin

A vitamin also known as nicotinic acid. It is a white water-soluble powder stable to heat, acid, and alkali, with the structure shown below. It is found

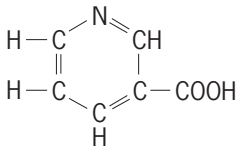

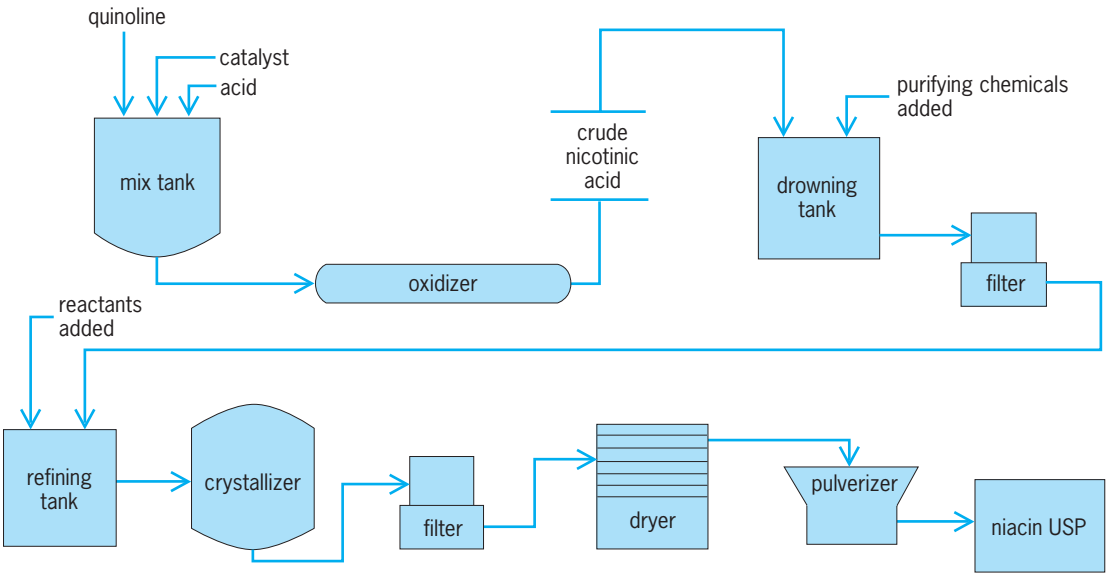

Flow sheet for the manufacture of niacin from quinoline.

in biochemically active combinations as the amide, niacinamide. Analyses for niacin are usually done microbiologically using *Lactobacillus arabinosus* as the test organism. Chemical methods of analysis are not usually satisfactory. All living cells studied have enzymic systems involving niacin. Many animals, including humans, are capable of synthesizing niacin in varying degrees from the amino acid tryptophan. Niacin is widely distributed in foods. Yeasts, wheat germ, and meats, particuarly organ meats, are rich sources of the vitamin. Some foods such as milk are relatively poor sources of niacin but contain generous quantities of tryptophan.

Niacin-deficiency disease is known as pellagra and is particularly prevalent among people whose diet is largely corn. Pellagra is characterized by dermatitis, dementia, diarrhea, and death. Skin lesions are usually observed in areas exposed to the sun. The disease is accompanied by gastrointestinal lesions.

Niacin is present in enzymes in the form of two coenzymes, diphosphopyridine nucleotide (DPN), also known as nicotinamide adenine dinucleotide (NAD) or coenzyme I, and triphosphopyridine nucleotide (TPN), also known as nicotinamide adenine dinucleotide phosphate (NADP) or coenzyme II. Enzymes containing DPM or TPN function in oxidation-reduction systems by virtue of their ability to accept hydrogen ions (protons) and electrons from substrates and transfer them to other hydrogen acceptors, such as the flavo-proteins. Niacin-containing enzymes catalyze about 40 reversible biochemical reactions, many of different types, as illustrated by the following: acetaldehyde $\rightleftharpoons$ ethanol; 1,3-diphosphoglyceric acid $\rightleftharpoons$ 3-phosphoglyceraldehyde; pyruvic acid $\rightleftharpoons$ lactic acid; imino glutaric acid $\rightleftharpoons$ L-glutamic acid; and acetic acid $\rightleftharpoons$ acetaldehyde. *See* BIOLOGICAL OXIDATION; COENZYME; RIBOFLAVIN.

Unlike the other B vitamins, little niacin is excreted in the urine. Humans excrete niacin mostly as N^1-methylnicotinamide and the 6-pyridone of N^1-methylnicotinamide. The pellagra-preventive potency of a diet is related not only to its niacin and tryptophan content but also to the availability of its niacin. There is evidence that some of the niacin of foods cannot be released by digestive enzymes. The existence of an antiniacin material in corn has been suggested. The effect of the carbohydrate content of the diet on the synthesis of niacin by intestinal bacterial may also be important. The use of urinary excretion data to determine nutritional status with regard to niacin has been disappointing. In humans an average of 60 mg of dietary tryptophan is equivalent to 1 mg of niacin. The recommended dietary allowance of the National Research Council is 6.6 niacin equivalents per 1000 kcal. *See* NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD); NICOTINAMIDE ADENINE DINUCLEOTIDE PHOSPHATE (NADP); VITAMIN. Stanley N. Gershoff

Niacin is produced to United States Pharmacopoeia (USP) requirements. The principal route for manufacture of niacin is oxidation of quinoline, obtained from coal tar (see **illus.**). Oxidation of 2-methyl-5-ethylpyridine is nearly as important; other processes involve oxidation of 3-picoline, hydrolysis of 3-cyanopyridine, or oxidation and hydrolysis of nicotine. Many patents cover the field. Nicotinamide is produced by amidation of niacin or its esters and by hydrolysis of 3-cyanopyridine.

Niacin, the least expensive vitamin, is used in foods, feed, and pharmaceutical preparations to supplement limited amounts available naturally. A major outlet for niacin is swine and poultry feeds, especially those based on corn. In human nutrition, enrichment of flour and bread with niacin is required in most states. Rice and corn products, alimentary pastes, and milk are fortified with niacin in some areas. Both niacin and nicotinamide are used in single and multivitamin capsules and tablets, and are prescribed for various clinical applications. Adrian H. Cubberley

Nickel

A chemical element, Ni, atomic number 28, a silver-white, ductile, malleable, tough metal. The atomic mass of naturally occurring nickel is 58.71. *See* PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Nickel consists of five natural isotopes having atomic masses of 58, 60, 61, 62, 64. Seven radioactive isotopes have also been identified, having mass numbers of 56, 57, 59, 63, 65, 66, and 67.

Nickel metal. Most commercial nickel goes into stainless steel and other corrosion-resistant alloys. Nickel is also important in coins as a replacement for silver. Finely divided nickel is used as a hydrogenation catalyst. *See* NICKEL ALLOYS.

Nickel is a fairly plentiful element, making up about 0.008% of the Earth's crust and 0.01% of the igneous rocks. Appreciable quantities of nickel are present in some kinds of meteorite, and large quantities are thought to exist in the Earth's core. Two important ores are the iron-nickel sulfides, pentlandite and pyrrhotite $(Ni,Fe)_xS_y$; the ore garnierite, $(Ni,Mg)SiO_3 \cdot nH_2O$, is also commercially important. Nickel occurs in small quantities in plants and animals. It is present in trace amounts in sea water, petroleum, and most coal.

Nickel metal is of moderate strength and hardness (3.8 on Mohs scale). When viewed as very small particles, nickel appears black. The density of nickel is 8.90 times that of water at 20°C (68°F). Nickel melts at 1455°C (2651°F) and boils at 2840°C (5144°F). Nickel is only moderately reactive. It resists alkaline corrosion and does not burn in the massive state, although fine nickel wires can be ignited. Nickel is above hydrogen in the electrochemical series, and it dissolves slowly in dilute acids, releasing hydrogen. In metallic form nickel is a moderately strong reducing agent.

Nickel is usually dipositive in its compounds, but it can also exist in the oxidation states 0, 1+, 3+, and 4+. Besides the simple nickel compounds, or salts, nickel forms a variety of coordination compounds or complexes. Most compounds of nickel are green or blue because of hydration or other ligand bonding to the metal. The nickel ion present in water solutions of simple nickel compounds is itself a complex, $[Ni(H_2O)_6]^{2+}$. William E. Cooley

Biology. Depending on the chemical species and concentration, nickel compounds exhibit a wide variety of biological effects. Selected nickel-containing compounds are harmful to humans and other animals. Nickel, however, also appears to be a beneficial trace element for many animals. Similarly, high concentrations of nickel are toxic to both plants and microorganisms, while trace levels of nickel are essential to many species. A small group of nickel-containing enzymes have been identified and characterized to varying extents, and nickel has been substitued for the normal metal in several metalloproteins. Robert L. Blakeley; Robert P. Hausinger

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; J. R. Davis (ed.), *Metals Handbook: Desk Edition*, 2d ed., ASM International, 1998; R. P. Hausinger, *Biochemistry of Nickel*, 1993; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; M.-H. Yu, *Environmental Toxicology: Biological and Health Effects of Pollutants*, 2d ed., 2004.

Nickel alloys

Combinations of nickel with other metals. Nickel has been used in electroplating since 1843 and as an alloying addition to steels since about 1889. It was first used as a base for alloys with the introduction of Monel nickel-copper alloy in about 1905. The nominal compositions of some of the alloys containing more than 50% nickel are given in **Table 1**.

Nickel-base alloys may be melted in open-hearth, electric-arc, or induction furnaces in air, under inert gas, or in vacuum. Casting may also be done under these same ambient conditions. Cast shapes are made in sand or investment molds or by shell molding. Ingots for wrought products are cast in metal molds and are hot-worked by forging, rolling, or extruding. In some instances further work may be done cold by rolling or drawing. Nickel-base alloys, made available in this way in bar, rod, wire, plate, strip, sheet, and tubular forms, may be fabricated into finished products by using conventional metalworking and metal-joining techniques.

Alloyed nickels. Nickel 211 and Duranickel alloy 301 are essentially binary alloys with 4.75% manganese and 4.5% aluminum, respectively. Manganese, in the first of these, extends the range of applicability in the presence of sulfur by about 300°F (150°C) to a limiting temperature in the neighborhood of 1000°F (540°C). A characteristic use of this material is as wire for spark-plug electrodes.

Aluminum and titanium confer age-hardening characteristics on Duranickel alloy 301 and a tensile strength in excess of 200,000 lb/in.^{2} (1400 megapascals) is attainable in this alloy by appropriate cold work and heat treatment. In this condition it is well suited to the manufacture of springs and diaphragms.

Monel alloy 400. This alloy contains about two-thirds nickel and one-third copper and is the oldest of the commercial nickel-base alloys, dating from about 1905 when it was directly smelted from the

TABLE 1. Nominal composition of some nickel-base alloys, wt %

Trademark	Ni	Cu	Cr	Co	Mo	Ti	Al	Cb	Fe	Mn	Si	C	Other
Nickel 211	95	—	—	—	—	—	—	—	—	4.75	—	0.08	—
Duranickel alloy 301	93.7	0.05	—	—	—	0.4	4.4	—	0.35	0.3	0.5	0.17	—
Monel alloy 400	66	31.5	—	—	—	—	—	—	1.35	0.9	0.15	0.18	—
Monel alloy K-500	66	29	—	—	—	0.5	2.75	—	0.9	0.75	0.5	0.15	—
Chromel P	90	—	10	—	—	—	—	—	—	—	—	—	—
Nichrome V	80	—	19.5	—	—	—	—	—	—	2.5*	1	0.25	—
Alumel	94	—	—	—	—	—	2	—	—	3	1	—	—
Nimonic 75	Bal	0.5*	19.5	—	—	0.4	—	—	5*	1*	1*	0.12	—
Nimonic 80A	Bal	—	19.5	2*	—	2.2	1.1	—	5*	1*	1*	0.1*	—
Inconel alloy 600	Bal	0.5*	15.5	—	—	—	—	—	8	1*	0.5*	0.15*	—
Inconel alloy X-750	Bal	0.5*	15	—	—	2.5	0.9	0.9	7	0.7	0.4	0.04	—
Inconel alloy 718	53	0.3*	19	1.0*	3	0.9	0.5	5	Bal	0.35	0.35*	0.08*	—
Alloy 713C	Bal	—	12	—	4	0.5	6	2	5*	1*	1*	0.2*	0.012 B, 0.10 Zr
Udimet 500	Bal	—	17.5	16.5	4	2.9	2.9	—	4*	0.75*	0.75*	0.15*	—
Waspaloy	Bal	—	19	14	3	2.5	1.2	—	2	0.7	0.4	0.05	—
M252	55	—	19	10	10	2.5	0.75	—	2	1	0.7	0.1	—
GMR 235	Bal	—	15.5	—	5	2.5	3	—	10	0.25*	0.6*	0.15	0.06 B
Hastelloy B	61	—	1*	2.5*	27.5	2	—	—	5.5	1*	1	0.05*	0.4V
Hastelloy C	54	—	15.5	2.5*	15.5	—	—	—	5.5	1*	1*	0.08*	0.35 V*, 4 W
Hastelloy D	82	3	1*	1.5*	—	—	—	—	2*	1	9	0.12*	—

*Maximum.

copper-nickel matte obtained from Sudbury, Ontario, sulfide ore. The good fabricating characteristics and corrosion resistance of this alloy have made it widely used in marine applications and in the chemical-processing and petroleum industries. It also has applicability in the field of nuclear propulsion. As with nickel, this alloy can be made age-hardenable by the addition of aluminum and titanium. Monel alloy K-500 has corrosion-resisting characteristics similar to those of the non-age-hardenable composition, and is widely specified for such applications as marine propellers, shafting, valves, pump parts, and springs. A usable tensile strength of about 175,000 lb/in.^{2} (1200 MPa) is obtainable in cold-drawn and age-hardened wire.

Nickel-chromium alloys. Nickel-chromium binary alloys are used primarily in specialty high-temperature service. An 80 nickel-20 chromium is a common high-quality resistance-heating-element material possessing good resistance to oxidation up to about 2100°F (1150°C), superior to either of its two component elements. The alloy is used both in industrial-furnace and household-appliance heating elements. Chromel P is used with the chromium-free nickel-base Alumel in temperature-sensing devices known as thermocouples. This alloy couple has favorable thermoelectric characteristics for applicability in the measurement of temperatures up to 2000°F (1090°C).

Nickel-chromium and related complex alloys are widely used for structural and general-purpose applications at high temperatures and in certain corrosive environments, particularly where freedom from stress-corrosion cracking is essential. In this latter instance Inconel alloy 600 has applicability in nuclear propulsion units.

This class of alloys encompasses a broad range of high-temperature properties. Nimonic 75 nickel-chromium alloy is used as a scale-resistant sheet material. Neither this material nor Inconel alloy 600 responds to age hardening, and hence they are on the low side of the elevated-temperature mechanical property range. By contrast, Inconel alloy X-750, which contains added aluminum, titanium, and columbium, develops greatly improved high-temperature strength after proper heat treatment. The more highly alloyed Alloy 713C, a nickel-chromium cast alloy, exhibits further strength improvement at the high side of the temperature range of applicability for the complex nickel-chromium alloys. For comparison, 100-h rupture strengths of these three alloys are listed in **Table 2**.

The aforementioned materials and a number of similar proprietary alloys which combine high strength and oxidation resistance, have found application in jet engine and gas turbines for such parts as combustion liners, blades, vanes, and disks.

Inconel alloy 718 is an age-hardenable alloy with a slow aging response, a characteristic which permits welding and annealing without spontaneous hardening. It is suitable for service from cryogenic temperatures (−423°F or −253°C) to moderately elevated temperatures (1300°F or 704°C).

In the interest of optimizing properties, these complex nickel-chromium alloys are being produced in increasing quantities by vacuum-melting and vacuum-pouring techniques.

TABLE 2. Rupture strengths (100 h) of nickel-chromium alloys, lb/in.^{2}*

Alloy	1500°F (815°C)	1700°F (927°C)
Inconel alloy 600	8000	3800
Inconel alloy X-750	25,000–30,000	8000–10,000
Alloy 713C	60,000	30,000

*1 lb/in.^{2} = 6.9 kPa.

A cast heat-resisting alloy carrying the Alloy Castings Institute designation HW (nominally 60% nickel, 12% chromium, 23% iron) is used principally for furnace parts and heat-treating fixtures. It has good resistance to oxidation and carburization, only modest hot strength, but good thermal shock resistance.

Hastelloy alloys. Hastelloy alloys B, C, and D are used primarily in corrosive environments. Hastelloy B is resistant to hydrochloric and sulfuric acids within certain limits of concentration, temperature, and degree of aeration. It is not recommended for service involving strong oxidizing acids or oxidizing salts. Hastelloy C is unusually resistant to oxidizing solutions and to moist chlorine. Hastelloy D has exceptional resistance to hot concentrated sulfuric acid.

Nickel-iron alloys. Alloys containing more than 50% nickel are used in various applications involving controlled thermal expansivity or certain magnetic requirements. In the range 50–52% nickel, the balance iron, the alloys have thermal expansion characteristics useful in making some types of glass-to-metal seals.

In the range 77–80% nickel, with or without about 4% molybdenum, the balance iron, the alloys have very high initial and maximum magnetic permeabilities when properly processed. *See* ALLOY; IRON ALLOYS; NICKEL; NICKEL METALLURGY; STAINLESS STEEL.

E. N. Skinner; Gaylord Smith

Nickel metallurgy

The extraction and refining of nickel from its ores. Nickel's properties of strength, toughness, and resistance to corrosion have been used to advantage in alloys since ancient times. Paktong, similar to modern nickel silver, was used in the sixteenth century in China, and early weapons were often fashioned from tough, nickel-bearing meteoric iron. A. F. Cronstedt first isolated nickel as an element in 1751, and in 1804 H. T. Richter prepared it in relatively pure form.

Occurrence. Although nickel ranks twenty-fourth in order of abundance of the elements, and igneous rocks average 0.01% nickel content, there are relatively few nickel deposits of commercial importance. Nickel ores are of two generic types, sulfides and laterites. Explorations of the ocean bottoms have revealed vast deposits of manganese oxide nodules which contain significant values of nickel, copper, and cobalt.

Sulfides. In ores of this type, nickel is present chiefly as the mineral pentlandite, a nickel-iron sulfide, usually in association with pyrrhotite and chalcopyrite. The most important known deposits, at Sudbury, Canada, have provided the major portion of the world's supply since 1905. Other substantial deposits have been developed in the Thompson-Moak Lake area of northern Manitoba, Canada, and in Western Australia. Russia exploits deposits on the Kola Peninsula and near Norilsk in Siberia. Smaller deposits are worked in South Africa and in Finland. *See* PENTLANDITE.

Laterites. Lateritic nickel ores occur in two main forms, oxides and silicates. In the oxide form the nickel is dispersed through limonite, a hydrated form of iron oxide, while in the silicate form nickel partially replaces magnesium in the lattice of a hydrated magnesium silicate. Lateritic ores are widely distributed throughout the tropics and constitute the world's largest known reserves of nickel. Deposits in New Caledonia, Cuba, and Oregon have been worked commercially for many years. Commercial installations are also operating in Greece, the Dominican Republic, the Philippines, and Australia. The Japanese nickel industry is almost exclusively based on imported ores from New Caledonia and Indonesia and concentrates from Australia and Canada. *See* LATERITE.

Ocean nodules. The existence of manganese oxide nodules on the bottoms of the oceans, notably the deep valleys of the Pacific and the Atlantic, has been known for many years. Exploration revealed that the nickel and cobalt values contained in these nodules, averaging about 1% nickel and 0.2% cobalt, rival the metals content of the land-based lateritic ores. *See* MANGANESE NODULES.

Production and uses. Nickel gained commercial prominence late in the nineteenth century, when substantial reserves were found in New Caledonia and at Sudbury, and the world's naval powers adopted nickel-bearing armor. Until about 1920 nickel markets depended upon military requirements. Following World War I, research into industrial applications of nickel was greatly increased, and the success of the continuing program is evident from nickel's numerous diversified and expanding uses.

Nickel is marketed in various forms. In the United States, nickel consumption is estimated at more than 45% for stainless and other steels; over 25% for nickel alloys; approximately 15% in electroplating, 7% in high-temperature and electrical-resistance alloys, 2% in cast irons, and the remainder in miscellaneous applications such as catalysts, magnets, and ceramics.

Extractive metallurgy. Selection of processes for nickel extraction is largely determined by the type of ore to be treated. Sulfide ores are amenable to concentration by such methods as flotation or magnetic separation. The state of combination of nickel in the lateritic ores usually precludes such enrichment, thus requiring treatment of the total ore.

Sulfide ores. **Figure 1** illustrates the major processes for the extraction and recovery of nickel from sulfide ores. These ores, usually containing from 1 to 3% nickel and varying amounts of copper, are first crushed and ground to liberate the mineral values and then subjected to froth flotation to concentrate the valuable constituents and reject the gangue or rock fraction. Depending on the ore's copper and pyrrhotite contents, it is sometimes appropriate to produce a separate copper concentrate and a separate pyrrhotite concentrate. The three examples chosen are typical in the industry and, with some minor

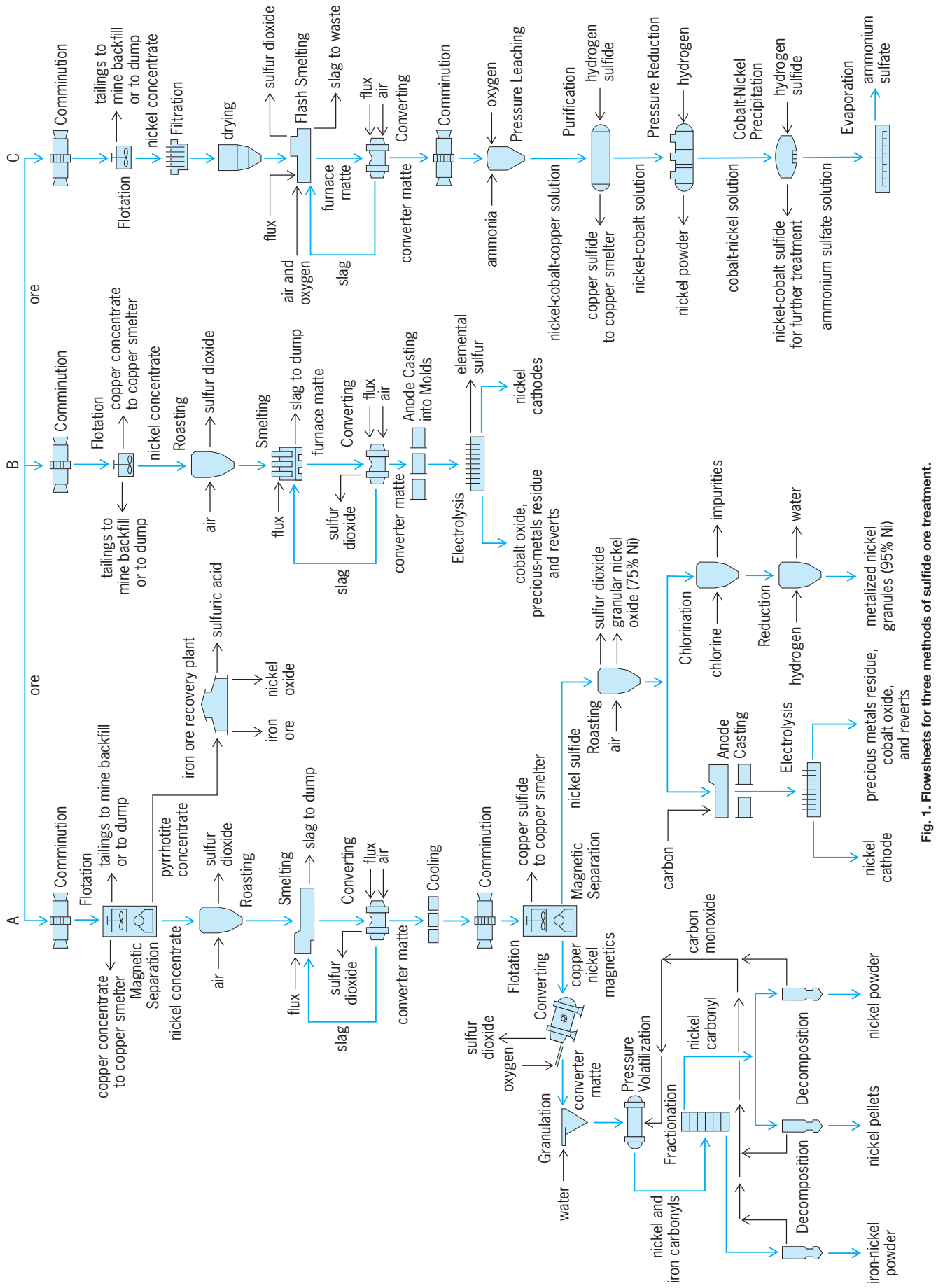

Fig. 1. Flowsheets for three methods of sulfide ore treatment.

variations in the processing steps, account for about 90% of the world's nickel production from sulfide ores. *See* ORE DRESSING.

Selective flotation and magnetic separation may be employed to divide the bulk concentrate into nickel, copper, and iron-rich fractions for separate treatment (Fig. 1, flowsheet A). A high-grade iron ore, nickel oxide, and sulfuric acid are recovered from the iron concentrate. The nickel concentrate is treated by pyrometallurgical processes. The major portion undergoes partial roasting in multihearth or fluidized-bed furnaces to eliminate about half of the sulfur and to oxidize the associated iron. The hot calcine, plus flux, is smelted in natural gas-coal-fired reverberatory furnaces operating at about 2200°F (1200°C) to produce a furnace matte, enriched in nickel, and a slag for discard. The furnace matte is transferred to Pierce-Smith converters and blown with air in the presence of more flux to oxidize the remaining iron and associated sulfur, yielding Bessemer matte containing nickel, copper, cobalt, small amounts of precious metals, and about 22% sulfur. The slag which is generated in the converting operation is returned to the smelting furnace to recover its metal values.

The molten Bessemer matte is cast into 25-ton (22.5-metric-ton) molds in which it undergoes controlled slow cooling to promote formation of relatively large, discrete crystals of copper sulfide, nickel sulfide, and a small quantity of a metallic phase, a nickel-copper alloy which collects most of the precious metals. After crushing and grinding, the metallics are removed magnetically and treated in a refining complex for recovery of metal values, and the main sulfide fraction is separated into copper sulfide and nickel sulfide concentrates by froth flotation. *See* FLOTATION; MAGNETIC SEPARATION METHODS.

The nickel sulfide is converted to granular nickel oxide sinter in fluidized-bed reactors. A portion of this product is marketed directly for alloy steel production. Another part of the granular oxide is treated by chlorination at high temperature (2200°F or 1200°C) to lower its copper content to about 0.5%, and then reduced by hydrogen at about 930°F (500°C) to yield a highly metallized product (95% Ni) for market. Two processes are employed to convert the remaining oxide to pure metal for market. One involves reduction smelting to metal anodes which is followed by electrolytic refining, by using a sulfate-chloride electrolyte with divided cells and continuous electrolyte purification. The product of this operation is electrolytic nickel cathodes, and the byproducts are cobalt and precious metals.

The other process is the atmospheric-pressure carbonyl process which is used in Great Britain. The nickel oxide sinter is reduced with hydrogen and treated with carbon monoxide at about 122°F (50°C) to volatilize nickel as gaseous nickel carbonyl. This compound is decomposed at about 390°F (200°C) to yield high-purity nickel in pellet form. Copper and cobalt salts and precious metals are recovered from the residue. Nickel powder is also produced at this plant in a pressure carbonyl system.

The nickel-copper alloy from the matte separation step, containing significant platinum group metal values, is melted in a top-blown rotary converter, its sulfur content is adjusted by blowing with oxygen at temperatures up to 2900°F (1600°C), and the metal product is granulated with water. The dried metal granules are treated in high-pressure (6.7 megapascals) reactors with carbon monoxide at 300°F (150°C) to form nickel carbonyl and some iron carbonyl. The mixed carbonyls are separated by fractionation, and the pure nickel carbonyl is decomposed at about 390°F (200°C) to yield high-purity nickel pellets and nickel powder for market.

There is a much simpler procedure that can be used to process nickel sulfide ores because the associated copper values are too low to warrant local processing (Fig. 1, flowsheet B). Selective flotation is employed to produce a nickel concentrate low in copper and a small amount of copper concentrate for treatment elsewhere. The dewatered nickel concentrate is fluid-bed-roasted for partial elimination of sulfur, and the calcine plus flux is smelted in arc-type electric furnaces. Waste furnace slag is granulated for disposal while the molten matte is transferred to Pierce-Smith converters for further upgrading to Bessemer matte. The conventional procedure of flux addition and blowing with air removes all but traces of iron from the matte, and the slag produced is returned to the electric furnace for recovery of metal values. The converter matte, essentially nickel sulfide (Ni_3S_2), is cast into anodes and electrolyzed to yield elemental sulfur at the anode and pure nickel at the cathode. The refining operation also produces cobalt oxide and precious-metal residues.

In the unique all-hydrometallurgical process (Fig. 1, flowsheet C), the nickel ore is concentrated by conventional froth flotation, and the dried nickel concentrate is flash-smelted with oxygen-enriched air and flux to produce furnace matte and waste slag. The furnace matte is cooled, crushed, and finely ground as feed for the hydrometallurgical plant, where it is leached under pressure with a strong ammonia solution and air to solubilize the base metal values, with the simultaneous production of ammonium sulfate. The pregnant leach liquor is treated to remove impurities and then reduced with hydrogen at elevated pressure (435 lb/in.^{2} or 3 MPa) and temperature (374°F or 190°C) to yield a granular nickel powder product. The tail liquor from this operation is treated further to recover ammonium sulfate crystals and a mixed nickel-cobalt sulfide.

There are many other processes for the treatment of sulfide nickel ores. However, the differences from the illustrated examples are believed to be minor.

Lateritic ores. The bulk of the nickel originating from lateritic ores is marketed as ferronickel. The process employed is basically simple and involves drying and preheating the ore usually under reduction conditions. The hot charge is then further reduced and melted in an electric-arc furnace, and the crude metal is refined and cast into ferronickel pigs. A typical operation is shown in **Fig. 2**, flowsheet A.

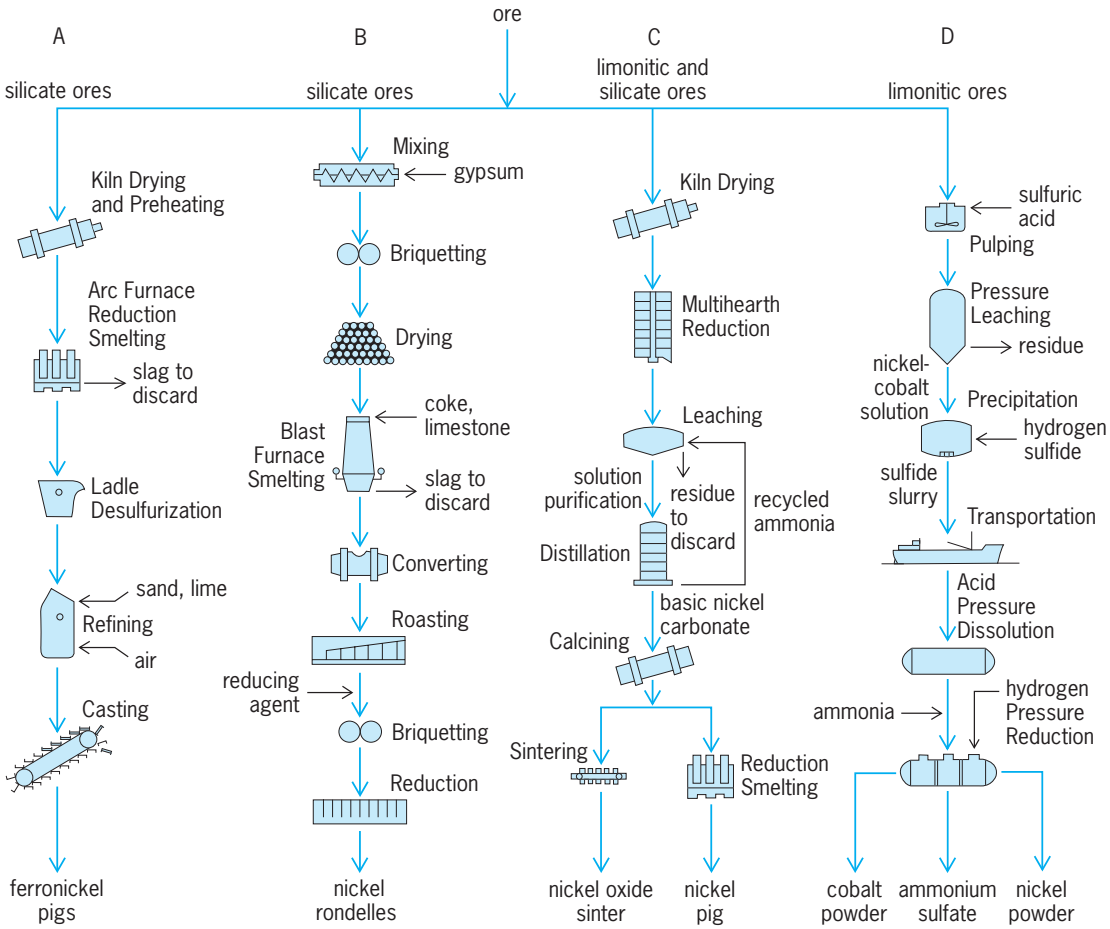

Fig. 2. Flowsheets for four methods of lateritic ore treatment.

A substantial amount of nickel is produced from lateritic ores by the nickel sulfide matte technique. In this process the ore is mixed with gypsum or other sulfur-containing material such as high-sulfur fuel oil, followed by a reduction and smelting operation to form matte. As in the treatment of sulfide concentrates, the molten furnace matte is upgraded in either conventional or top-blown rotary converters to a high-grade matte, which can be further refined by roasting and reduction to a metallized product. An example of this procedure is shown in Fig. 2, flowsheet B.

Other large-scale operations are based on selective reduction of the ore followed by ammoniacal leaching at atmospheric pressure to dissolve the nickel values. The pregnant liquor is treated to remove impurities and then heated in suitable vessels to drive off ammonia and carbon dioxide and to precipitate a basic nickel carbonate. This material may be sintered under reducing conditions to yield a metallized nickel oxide sinter, or the carbonate may be redissolved in ammoniacal solution and then treated with hydrogen under pressure to yield nickel powder for briquetting. This process is depicted in Fig. 2, flowsheet C. Flowsheet D in Fig. 2 depicts a process wherein limonitic-type ores are leached with sulfuric acid at elevated temperature and pressure to solubilize nickel and cobalt. The pregnant solution is then treated with hydrogen sulfide to precipitate mixed nickel-cobalt sulfides. This precipitate may be treated by the Sherritt-Gordon pressure ammonia leach process to yield separate nickel and cobalt powders. *See* NICKEL; NICKEL ALLOYS; PYROMETALLURGY, NONFERROUS.

Alexander Illis

Bibliography. W. Betteridge, *Nickel and Its Alloys*; 1977; A. R. Burkin, *Extractive Metallurgy of Nickel*, 1987; F. Habashi, *Pyrometallurgy*, 1986; J. J. Jacobs, M. Allard, and S. Behmo, *Nickel and Cobalt Extraction Using Organic Compounds*, 1985; G. Tyroler and C. Landolt (eds.), *Extractive Metallurgy of Ni and Co,* 1988.

Nicotinamide adenine dinucleotide (NAD)

An organic coenzyme and one of the most important components of the enzymatic systems concerned with biological oxidation-reduction reactions. It is also known as NAD, diphosphopyridine nucleotide (DPN), coenzyme I, and codehydrogenase I. NAD (**Fig. 1**) is found in the tissues of all living organisms. *See* COENZYME.

The nicotinamide, or pyridine, portion of NAD can be reduced chemically or enzymatically with the formation of reduced or hydrogenated NAD, or NADH (**Fig. 2**). NAD functions as the immediate oxidizing

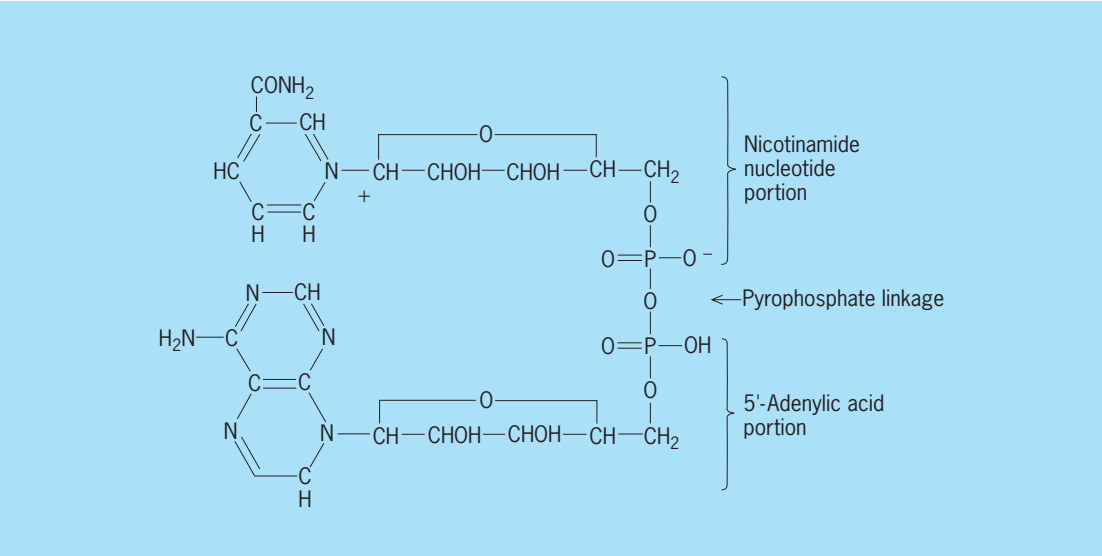

Fig. 1. Oxidized form of nicotinamide adenine dinucleotide (NAD).

agent for the oxidation, or dehydrogenation, of various organic compounds in the presence of appropriate dehydrogenases, which are specific apoenzymes, or protein portions of the enzyme. In the dehydrogenase reactions one hydrogen atom is transferred from the substrate to NAD, while another is liberated as hydrogen ion. For instance, in a reversible reaction, catalyzed by glucose dehydrogenase, glucose is oxidized to gluconolactone, as shown in reaction (1).

$$\underset{\text{Glucose}}{C_6H_{12}O_6} + NAD \rightleftharpoons \underset{\text{Glucono-lactone}}{C_6H_{10}O_6} + NADH + H \qquad (1)$$

The NADH, formed in biological oxidations, is reoxidized to NAD (Fig. 2) in coupled reductions, also catalyzed by specific enzymes. In respiration the NADH is reoxidized through a sequence of reactions in which a flavoprotein enzyme, diaphorase, and the cytochrome system of iron-porphyrin catalysts transfer the electrons from NADH to molecular oxygen, the overall reaction being expressed as reaction (2).

$$2NADH + 2H + O_2 \rightarrow 2NAD + 2H_2O \qquad (2)$$

In fermentations NADH is reoxidized with the concomitant reduction of organic molecules, which are usually produced in intermediary metabolism. For example, in the metabolism of muscle tissue, lactic acid is produced by the reduction of pyruvic acid,

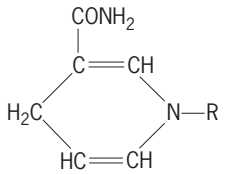

Fig. 2. Reduced form of nicotinamide portion in hydrogenated NAD (NADH).

as shown in reaction (3). The enzyme catalyzing

$$NADH + H + \underset{\text{Pyruvic acid}}{C_3H_4O_3} \rightleftharpoons NAD + \underset{\text{Lactic acid}}{C_3H_6O_3} \qquad (3)$$

this reaction is called lactic dehydrogenase, since the process is reversible, and lactic acid can be oxidized with NAD. NAD and its reduced form, NADH, serve to couple oxidative and reductive processes and are constantly regenerated during metabolism. Hence, they serve as catalysts and NAD is referred to as a coenzyme. In some enzymatic reactions a different coenzyme, nicotinamide adenine dinucleotide phosphate (NADP), also called triphosphopyridine nucleotide, or coenzyme II, is required. Dehydrogenases are generally quite specific with respect to the coenzyme which they can utilize. *See* BIOLOGICAL OXIDATION; CYTOCHROME; ENZYME; NIACIN; NICOTINAMIDE ADENINE DINUCLEOTIDE PHOSPHATE (NADP).

Michael Doudoroff

Nicotinamide adenine dinucleotide phosphate (NADP)

A coenzyme and an important component of the enzymatic systems concerned with biological oxidation-reduction systems. It is also known as NAD, triphosphopyridine dinucleotide, triphosphopyridine nucleotide (TPN), coenzyme II, and codehydrogenase II. The compound is similar in structure and function to nicotinamide adenine dinucleotide (NAD). It differs structurally from NAD in having an additional phosphoric acid group esterified at the 2′ position of the ribose moiety of the adenylic acid portion. In biological oxidation-reduction reactions the NADP molecule becomes alternately reduced to its hydrogenated form (NADPH) and reoxidized to its initial state (see **illus.**). *See* NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD).

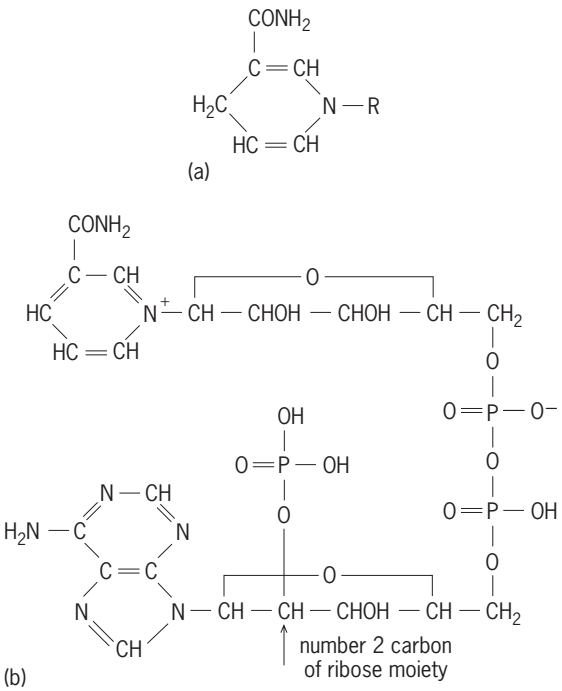

Nicotinamide adenine dinucleotide phosphate. (*a*) Reduced form of the nicotinamide portion of NADPH. (*b*) Oxidized form of the molecule (NADP).

NADP is specifically required in some enzymatic reactions, just as NAD is in others. There are, however, a few reactions in which either compound can serve as a coenzyme. In the carbohydrate metabolism of yeast and mammalian tissue, NADP acts as the oxidizing agent for glucose-6-phosphate, with the enzyme glucose-6-phosphate dehydrogenase, as in the reaction below. NADP is formed, enzymatically, from

$$\text{Glucose-6-phosphate} + \text{NADP} \rightarrow \text{Glucose-6-}\delta\text{-gluconolactone} + \text{NADPH} + \text{H}^+$$

NAD by phosphorylation with adenosine triphosphate. *See* BIOLOGICAL OXIDATION; CARBOHYDRATE METABOLISM; COENZYME; ENZYME. Michael Doudoroff

Nicotine alkaloids

Alkaloids found in various species of the genus *Nicotiana*. The species most often used for the production of tobacco because of its high level of nicotine is *N. tabacum*, which is cultivated in many parts of the world for the preparation of cigarettes, cigars, and pipe tobacco. Nicotine is the most abundant alkaloid in *N. tabacum*, occurring to the extent of 2–8% based on the dry weight of the cured leaf. Other alkaloids that are found in this species are nornicotine, anabasine, and anatabine. The chemical structures of these alkaloids are shown in the **illustration**. *See* ALKALOID; TOBACCO.

The relative amounts of the alkaloids depend on the variety of *N. tabacum* and the various species of *Nicotiana*. In tree tobacco, *N. glauca*, (*RS*)-anabasine is the main alkaloid. (*S*)-Nicotine ($C_{10}H_{14}N_2$) is readily extracted from tobacco roots and stalks that remain after the leaves have been picked for tobacco production. (*S*)-Nicotine is a water-soluble colorless liquid when freshly distilled at reduced pressure (123–125°C or 253–257°F; 17 mmHg or 2.3 kPa). However, it rapidly turns yellow and then brown on exposure to air. It is optically pure when obtained from the tobacco plant, specific rotation $[\alpha]^{25}{}_D = -169°$ (pure liquid). It is racemized on heating with a small amount of potassium tertiary butoxide at its boiling point (250°C or 482°F) for 15 min. Nicotine is dibasic [$pK_a = 7.9$ (pyrrolidine); 3.1 (pyridine)], and it forms crystalline salts with many acids. *See* PK.

Nicotine forms a crystalline salt with two molecules of tartaric acid and a salt, sparingly soluble in ethanol, with two molecules of picric acid.

Nicotine has a bitter taste and a sharp odor. It is quite poisonous, with 250–350 mg being fatal for an adult when taken orally, and less than fatal amounts causing fever, trembling, nausea, and convulsions. It is toxic to animals, since it binds to the receptor site for acetylcholine, a nerve transmitter, and so it is used as an insecticide for killing cockroaches, houseflies, mosquitoes, and aphids. A typical commercial nicotine insecticide is a 40% solution of nicotine sulfate in water. *See* ACETYLCHOLINE; INSECTICIDE.

The sensation that is obtained from smoking tobacco is due to the presence of nicotine in the tobacco leaf. Some is destroyed during burning, but some is converted to nitroso compounds, such as *N*-nitrosonornicotine, by reaction with nitrous acid, which is formed in the burning tobacco by pyrolysis of the nitrates in the unburnt leaf. While these nitroso compounds are carcinogenic, the main source of carcinogens in tobacco smoke is probably in the tars, which contain polynuclear aromatic hydrocarbons. Smokers metabolize nicotine to cotinine, which is much less toxic and is excreted in the urine. The special chewing gum that is prescribed to alleviate the withdrawal symptoms of those who stop smoking contains about 2 mg nicotine per tablet of gum. *See* MUTAGENS AND CARCINOGENS; NITRO AND NITROSO COMPOUNDS; PYROLYSIS.

Nicotine is formed initially in the roots of the tobacco plant and is then translocated into the leaves, where it is relatively inert and is present as salts of organic acids such as citric acid and malic acid. The precursors of nicotine are nicotinic acid (the vitamin niacin), which is the source of the pyridine ring, and putrescine, which is the source of the pyrrolidine ring. The putrescine is derived from the α-amino acids ornithine and arginine. The *N*-methyl group is derived from the *S*-methyl group of the amino acid methionine. This same biosynthetic route operates in the *Duboisia* species. Nornicotine is a metabolite of nicotine resulting from an *N*-demethylation that occurs in both the growing plant and the dried leaf. The pyridine ring of anabasine is also formed from nicotinic acid, and the saturated ring is derived from lysine. Both rings of anatabine are derived from nicotinic acid. *See* PYRIDINE; PYRROLE.

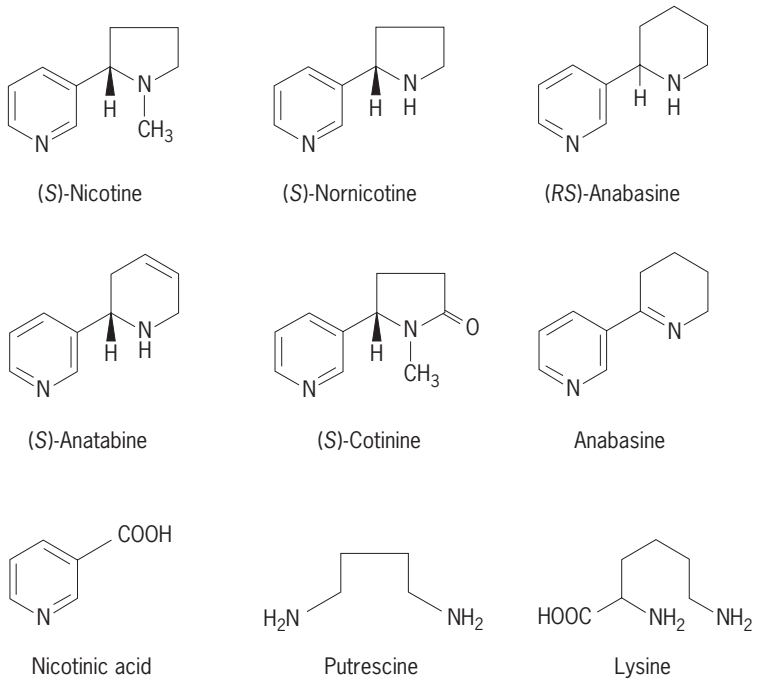

Structures of nicotine and some other alkaloids found in tobacco.

Nicotine is readily isolated from tobacco leaves by blending with a mixture of concentrated ammonia and an organic solvent such as chloroform or methylene chloride. Nicotine and the minor alkaloids are extracted from the organic solvent with dilute hydrochloric acid. This aqueous solution is made basic with ammonia and reextracted with chloroform. Evaporation of this extract affords the crude alkaloids, which can be further purified and separated into the individual components by various kinds of chromatography. The simplest method for analyzing this mixture is by capillary gas chromatography. *See* CHROMATOGRAPHY.

Nicotine has also been isolated from species other than *Nicotiana* such as *Duboisia* (native to Australia), *Equisetum* (horsetails), *Lycopersicum* (tomatoes), *Lycopodium* (club mosses), *Sedum* (succulent plants), and *Solanum* (potato family). Remarkably, anabasine, a minor alkaloid in tobacco, has been isolated from a species of marine worms (*Amphiporus*) and some ants (*Aphaenogaster*). *See* HETEROCYCLIC COMPOUNDS.

Edward Leete

Bibliography. L. E. Ember, The nicotine connection, *Chem. Eng. News*, 72(48):8–18, 1994; J. W. Gorrod and J. Wahren (eds.), *Nicotine and Related Alkaloids: Distribution, Metabolism, and Excretion*, 1993; S. W. Pellietier (ed.), *Alkaloids: Chemical and Biological Perspectives*, vol. 3, 1990, vol. 8, 1992.

Niobium

A chemical element, Nb, atomic number 41 and atomic weight 92.906. In the United States this element was originally called columbium. The metallurgists and metals industry still use this older name. *See* PERIODIC TABLE.

Most niobium is used in special stainless steels, high-temperature alloys, and superconducting alloys such as Nb_3Sn. Niobium is also used in nuclear piles.

Niobium metal has a density of 8.6 g/cm^3 (5.0 oz/in.^{3}) at 20°C (68°F), a melting point of 2468°C (4474°F), and a boiling point of 4927°C (8900°F). Metallic niobium is quite inert to all acids except hydrofluoric, presumably owing to an oxide film on the surface. Niobium metal is slowly oxidized in alkaline solution. It reacts with oxygen and the halogens upon heating to form the oxidation state V oxide and halides, with nitrogen to form NbN, and with carbon to form NbC, as well as other elements such as arsenic, antimony, tellurium, and selenium.

The oxide Nb_2O_5, melting point 1520°C (2768°F), dissolves in fused alkali to yield a soluble complex niobate, $Nb_6O_{19}^{8-}$. Normal niobates such as NbO_4^{3-} are insoluble. The oxide dissolves in hydrofluoric acid to give ionic species such as $NbOF_5^{2-}$ and

1												13	14	15	16	17	18
1 H	2																2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb

actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

$NbOF_6^{3-}$, depending on the fluoride and hydrogen-ion concentration. The highest fluoro complex which can exist in solution is NbF_6^-. Edwin M. Larsen

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; J. Hala (ed.), *Halides, Oxyhalides and Salts of Halogen Complexes of Titanium, Zirconium, Hafnium, Vanadium, Niobium and Tantalum*, 1989; J. Stephens and I. Ahmad (eds.), *High Temperature Niobium Alloys*, 1991.

Nippotaeniidea

An order of tapeworms of the subclass Cestoda. The few known species are intestinal parasites of Eurasian fresh-water fishes. The head (scolex) bears a single terminal sucker (see **illus.**). The segmental anatomy shows relationships to the Pseudophyllidea and Cyclophyllidea. The life history is unknown. It is probable that this order is related to the proteocephalids. *See* EUCESTODA; PSEUDOPHYLLIDEA. Clark P. Read

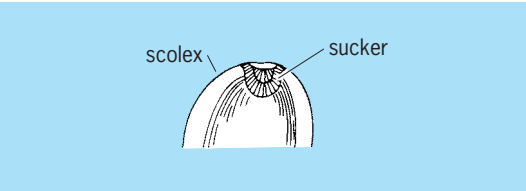

Scolex of *Nippotaenia*.

Niter

A potassium nitrate mineral with chemical composition KNO_3. Niter crystallizes in the orthorhombic system, generally in thin crusts and delicate acicular crystals; it also occurs in massive, granular, or earthy forms. It has good cleavage in three directions; fracture is subconchoidal to uneven; it is brittle; hardness is 2 on Mohs scale; specific gravity is 2.109. The luster is vitreous, and the color and streak are colorless to white.

Niter is commonly found, usually in small amounts, as a surface efflorescence in arid regions and in caves and other sheltered places. It is usually associated with soda niter, epsomite, nitrocalcite, and gypsum. The mineral may occur as an efflorescence on soils rich in organic matter from the action of certain bacteria on nitrogenous or animal matter. Niter occurs associated with soda niter in the desert regions of northern Chile, and in similar occurrences in Italy, Egypt, Russia, the western United States, and elsewhere. It was formerly found in some abundance in limestone caves in Tennessee, Kentucky, Alabama, and Ohio, and was used for the manufacture of gunpowder during the War of 1812 and the Civil War. *See* NITRATE MINERALS; SODA NITER. George Switzer

Nitrate minerals

These minerals are few in number and with the exception of soda niter are of rare occurrence. Normal anhydrous and hydrated nitrates occurring as minerals are soda niter, $NaNO_3$; niter, KNO_3; ammonia niter, NH_4NO_3; nitrobarite, $Ba(NO_3)_2$; nitrocalcite, $Ca(NO_3)_2{\cdot}4H_2O$; and nitromagnesite, $Mg(NO_3)_2{\cdot}6H_2O$. In addition there are three known naturally occurring nitrates containing hydroxyl or halogen, or compound nitrates. They are gerhardtite, $Cu_2(NO_3)(OH)_3$; buttgenbachite, $Cu_{19}(NO_3)_2Cl_4(OH)_{32}{\cdot}3H_2O$; and darapskite, $Na_3(NO_3)(SO_4){\cdot}H_2O$. *See* NITER; SODA NITER.

The natural nitrates are for the most part readily soluble in water. For this reason they occur most abundantly in arid regions, particularly in South America along the Chilean coast. *See* FERTILIZER; NITROGEN. George Switzer

Nitration

A process in which a nitro group (—NO_2) becomes chemically attached to a carbon, oxygen, or nitrogen atom in an organic compound. A hydrogen or halogen atom is often replaced by the nitro group. Three general reactions summarize nitration chemistry:

1. C-nitration, in which the nitro group attaches itself to a carbon atom [reaction (1)].

$$\text{—C—H} + HNO_3 \longrightarrow \text{—C—}NO_2 + H_2O \qquad (1)$$

2. O-nitration (an esterification reaction), in which an O-N bond is formed to produce a nitrate [reaction (2)].

$$\text{—C—OH} + HNO_3 \longrightarrow \text{—C—O} + NO_2 + H_2O \qquad (2)$$

3. N-nitration, in which a N-N bond is formed [reaction (3)].

$$\text{>NH} + HNO_3 \longrightarrow \text{>N—}NO_2 + H_2O \qquad (3)$$

The chemical steps of nitration differ based on the type of organic compound nitrated, the nitration agent, and the operating conditions. Additionally, there are three important types of nitration processes: ionic, free-radical, and the Victor Meyer process.

Ionic-type nitrations. Aromatics, alcohols, glycols, and amines are generally nitrated with mixed acids via ionic reactions. The mixture includes nitric acid, a strong acid such as sulfuric acid which acts as a catalyst, and a small amount of water. With sulfuric acid, nitric acid is ionized to the nitronium ion, $NO^+{}_2$, which is the nitrating agent [reaction (4)]. The nitronium ion concentration increases in the acid mixture as the water content of the acids decreases and as the

$$HNO_3 + 2H_2SO_4 \longleftrightarrow NO_2^+ + H_3O^+ + 2HSO_4^- \qquad (4)$$

molar ratio of sulfuric acid to nitric acid approaches 2.0. A mathematical model based on theoretical considerations has been developed by Albright, Good, and Eckert to predict the nitronium ion concentration in the mixed acids. Nitronium ions react with aromatics, alcohols, and amines to form intermediate ions, which release protons to complete the nitration and regenerate sulfuric acid. Nitrations with mixed acids are often conducted at temperatures ranging from 50 to 120°C.

Other strong acids sometimes used to produce nitronium ions include hydrogen fluoride, boron trifluoride (plus a trace of water), perchloric acid, acetic acid, and acidic-type ion-exchange resins. A mixture of nitric acid, sulfuric acid, and liquid sulfur trioxide, is also effective at relatively low temperatures. Nitrations to form dinitro or trinitro aromatics are harder to achieve, and require stronger acids or higher temperatures. The ease of nitrating an aromatic compound decreases as the number of nitro groups attached to the ring increase. *See* NITRO AND NITROSO COMPOUNDS; NITROAROMATIC COMPOUND.

Dinitrogen pentoxide, N_2O_5, is also a highly effective nitrating agent, particularly for compounds that cannot be nitrated with mixed acids. N_2O_5 ionizes, forming NO^+_2 plus NO^-_3. Nitric acid is often a byproduct when N_2O_5 is used. *See* NITROGEN OXIDES.

For easily nitrated aromatics such as phenols or phenolic ethers, a mixture of nitrous acid (or a nitrite salt) and nitric acid is often used. Nitrosonium ions (NO^+), generated by ionization, react with aromatics to produce nitrosoaromatics, which are then oxidized by nitric acid to form both the nitroaromatic compound and nitrous acid. Nitrous acid permits further production of nitrosonium ions. Once these aromatics have been mononitrated, mixtures of nitric and sulfuric acid are generally required to di- and trinitrate them.

When mixed acids are used, two immiscible liquid phases—acid and hydrocarbon—are present in the reactor. Since nitration reactions occur mainly at the interface between the two phases, a large interfacial area between them is required for rapid nitration. To obtain a large interfacial area, high levels of agitation or special injection devices are often provided. No known experimental data are available concerning interfacial areas in commercial or laboratory reactors, but mixing models, such as VisiMix 2000, can often be used to predict these areas.

During nitration reactions, some undesired oxidation always occurs. Oxidation results when small amounts of nitric acid are extracted from the acid phase to the hydrocarbon phase. The amount of extracted nitric acid becomes significant as the amount of nitrated hydrocarbons in the hydrocarbon phase increases. In the absence of sulfuric acid and water, nitric acid is an effective oxidizing agent, capable of producing phenolic compounds from aromatics. Such phenolic compounds are easily nitrated and may even form di- and trinitrophenols, which are explosives. *See* NITRIC ACID; OXIDIZING AGENT.

Several nitrated hydrocarbons, such as trinitrotoluene, dinitrotoluenes, picric acid, and nitroglycerine, are explosives; major accidents, including explosions, have been reported in chemical plants that manufacture them. Several serious explosions have also occurred in plants producing mononitroaromatics and nitroalcohols. Other hazards in nitration plants include handling and recovery of acids, flammability of the hydrocarbon feeds and products, side reactions including undesired oxidations, and the toxicity of some hydrocarbons.

Batch reactors were used until perhaps 50 years ago to produce most, if not all, nitroaromatics, nitroalcohols, nitroglycerine, and nitroamines. Improper design and/or operation of these batch reactors and auxiliary equipment has on several occasions contributed to serious accidents. Temperature control of a highly exothermic reaction mixture is always a concern, especially in larger reactors that often have less heat transfer area per volume of reactants. To minimize accidents, the rates of nitrations are generally maintained at low levels: reaction times in the batch reactor are often in the 2–5-hour range.

Continuous-flow reactors are currently widely used to produce nitrobenzene, dinitrotoluenes, nitroglycerine, plus several other products. Continuous-flow reactors often provide large interfacial areas between the two liquid phases, resulting in much higher rates of nitration. Smaller reactors can be used due to their improved heat-transfer capabilities. In addition, by properly controlling the compositions and feed rates of the feed streams the reactors can often be operated adiabatically with only a small temperature increase resulting from the reaction, because only small amounts of nitrated products are in the reactor at any given time.

The Noram process is used for worldwide production of most nitrobenzene. In this process, the mixed acids contain only several percent of nitric acid, which is completely ionized to form nitronium ions. In one example, the feed acid contained 2.8% nitric acid, 27.7% water, and 69.5% sulfuric acid. Injectors are used to obtain large interfacial areas between the two phases. The process offers the following features: (1) The nitration reaction is completed in about 70 seconds; much shorter times have been reported in other flow processes. (2) Over 99% of the nitric acid reacts to produce nitrobenzene, using inlet temperatures of about 97–120°C. (3) The exit acid is a diluted sulfuric acid, which can be concentrated and reused. (4) The unreacted benzene is recovered and is recycled into the feed benzene. (5) Undesired side reactions, including oxidations, are of minor importance. (6) Operating costs are relatively low in part because the heat of reaction is used in the preheating and separation steps.

Free-radical nitrations. Propane is commercially nitrated in relatively large amounts using nitric acid in gas-phase free-radical reactions at temperatures of about 380–420°C. Nitric acid decomposes at these temperatures to produce nitrogen dioxide radicals

(actually a mixture of $\cdot NO_2$ and $\cdot ONO$) and a hydroxy radical ($\cdot OH$). In the free-radical reaction, about 35–40% of the nitric acid reacts to form four C_1-C_3 nitroparaffins; C-C bonds are broken during the nitration. The remaining nitric acid acts mainly as an oxidizing agent to form aldehydes, alcohols, carbon monoxide, carbon dioxide, water, and small amounts of other oxidized materials. Commercially, an adiabatic reactor is used, and the heat of reaction is employed to preheat and vaporize the nitric acid feed (containing water).

Separation steps used with the gaseous product steam include almost a complete condensation of mixed nitroparaffins. This liquid mixture is washed to remove the aldehydes, and then is distilled to recover each of the four nitroparaffins—nitromethane, nitroethane, 1-nitropropane, and 2-nitropropane. The unreacted propane is recovered, combined with the feed propane, and returned to the reactor. The oxides of nitrogen are converted back to nitric acid. Carbon monoxide, carbon dioxide, and water are discarded.

Vapor-phase nitrations of methane and ethane produce relatively low yields of nitroparaffins. These two paraffins contain only primary C-H bonds, which are less reactive than the secondary C-H bonds in propane, heavier normal paraffins, and cycloparaffins.

Gaseous nitrogen dioxide was previously used to commercially nitrate propane. This nitration could be run at lower temperatures, such as 220–250°C. Only 20–27% of the nitrogen dioxide, however, was converted to the four nitroparaffins. Most of the remaining nitrogen dioxide acted as an oxidizing agent.

Propane and other paraffins can also be nitrated in a liquid-phase process using a dispersion of liquid paraffins and liquid nitric acid. Relatively low temperatures (140–200°C) are used, and sufficiently high pressures are required to maintain the reactants as liquids. Dinitroparaffins, such as 2,2-dinitropropane, can be produced; they have never been produced in a vapor-phase process probably because of the higher temperature. *See* NITROPARAFFIN.

Victor Meyer and other processes. In the classical Victor Meyer process, an organic halide (often a bromide) is reacted with silver nitrite to produce a nitrohydrocarbon and silver halide. In a modified process, sodium nitrite, dissolved in a suitable solvent, is substituted for the more expensive silver nitrite. The desired nitroalkanes are produced in high yields by these processes, whereas they are produced in rather low yields in free-radical nitrations.

Nitrations can also often be performed by addition reactions using unsaturated hydrocarbons with nitric acid or nitrogen dioxide. Lyle F. Albright

Bibliography. L. F. Albright, *Chem. Eng.*, pp. 149–156, June 6, 1966; L. F. Albright, R. V. C. Carr, and R. J. Schmitt, Nitration: Recent Laboratory and Industrial Developments, *ACS Symp. Ser.*, no. 623, American Chemical Society, Washington, DC, 1996; A. A. Guenkel, J. M. Rae, and E. G. Hauptmann, U.S. Pat.5,313,009, May 14, 1994.

Nitric acid

A strong mineral acid having the formula HNO_3. Pure nitric acid is a colorless liquid with a density of 1.52 at 25°C (77°F); it freezes at −47°C (−53°F). Nitric acid is used in the manufacture of ammonium nitrate and phosphate fertilizers, nitro explosives, plastics, dyes, and lacquers. The principal commercial process for the manufacture of nitric acid is the Ostwald process, in which ammonia, NH_3, is catalytically oxidized with air to form nitrogen dioxide, NO_2. When the dioxide is dissolved in water, 60% nitric acid is formed. Production of 90–100% nitric acid is based on processes such as the reaction of sulfuric acid with sodium nitrate (an older method of nitric acid manufacture), dehydration of 60% acid, and oxidation of nitrogen dioxide in a solution of dilute nitric acid.

Nitric acid decomposes readily as shown in the reaction below. It is a strong oxidizing agent, oxidizing

$$2HNO_3 \rightarrow H_2O + 2NO_2 + {}^1/_2O_2$$

carbon to carbon dioxide, sulfur to sulfuric acid, and phosphorus to phosphoric acid. It reacts with most metals; products depend on the metal's electromotive series position and nitric acid concentration. *See* AMMONIA; NITROGEN; NITROGEN OXIDES; OXIDIZING AGENT. Francis J. Johnston

Bibliography. G. Agam, *Industrial Chemicals, Their Characteristics and Development: Chemicals in the Real World*, 1994; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999.

Nitric oxide

An important messenger molecule in mammals and other animals. It can be toxic or beneficial, depending upon the amount and where in the body it is released. Initial research into the chemistry of nitric oxide (NO) was motivated by its production in automobile emissions and other combustion processes, which results in photochemical smog and acid rain. In the late 1980s, researchers in immunology, cardiovascular pharmacology, neurobiology, and toxicology discovered that nitric oxide is a crucial physiological messenger molecule. Nitric oxide is now thought to play a role in blood pressure regulation, control of blood clotting, immune defense, digestion, neuronal signaling, the senses of sight and smell, and possibly learning and memory. Underproduction or unregulated overproduction of nitric oxide may also contribute to disease processes such as diabetes, artherosclerosis, stroke, hypertension, carcinogenesis, multiple sclerosis, transplant rejection, damage associated with reperfusion in ischemic (oxygen-deprived) tissues, impotence, septic shock, and long-term depression.

Most cellular messengers are large, unreactive biomolecules that make specific contacts with their targets. In contrast, nitric oxide is a small molecule that can diffuse freely throughout biological tissues,

reacting with target molecules at remote sites that are as far away as ~300 micrometers from its site of generation, a distance corresponding to 10–20 cell lengths. Nitric oxide is a free radical—that is, it contains an unpaired electron. However, unlike most free radicals, it is remarkably unreactive toward other compounds, and its chemistry is limited primarily to bond-forming combination with other radicals and coordination to metal centers. This contrasts markedly with the other common nitrogen oxide free radical, nitrogen dioxide (NO_2), which is reactive toward a wide variety of organic compounds and is frequently used in large-scale commercial syntheses. This dramatic difference in reactivity has been attributed to the capacity for NO_2 to dimerize to dinitrogen tetroxide (N_2O_4) and undergo autoionization to nitrate (NO_3^-) and the very reactive nitrosonium (NO^+) ion. In contrast, NO does not dimerize and its autoionization to NO^+ and the nitroxyl anion (NO^-) is energetically highly unfavorable, precluding formation of these reactive species. The chemical inertness of NO extends to the biological components of cells, where its direct reactions are limited primarily to coordination of transition metals such as iron centers in heme (iron porphyrin–containing) proteins and reaction with the diradical O_2, O_2-derived free radicals, and organic free radicals that are formed upon oxidation of various biomolecules. *See* FREE RADICAL; NITRO AND NITROSO COMPOUNDS; NITROGEN; NITROGEN OXIDES; PORPHYRIN.

Production. Nitric oxide is produced in the body by enzymes called nitric oxide synthases, which convert the amino acid L-arginine to nitric oxide and L-citrulline (**Fig. 1**). There are three recognized types of nitric oxide synthase, often designated nNOS, eNOS, and iNOS for the neuronal, endothelial, and inducible forms, respectively. Both nNOS and eNOS are constitutive—that is, they are always present in cells—while iNOS is normally not present but is synthesized in response to biochemical signals generated within the cellular environment. Nitric oxide synthases are structurally organized into two domains called the reductase and oxygenase domains. The reductase domain contains two redox-active flavin cofactors (FAD and FMN) that deliver electrons one at a time from the physiological electron donor (NADPH) to the oxygenase domain that contains the active site for arginine oxidation. The oxygenase domain contains a biologically unique combination of heme and tetrahydrobiopterin (H_4B) cofactors. Oxidation of arginine to citrulline plus NO requires two cycles of oxidation with O_2 (Fig. 1). Recent structural and kinetic evidence strongly suggests that H_4B plays essential roles in both cycles to deliver electrons to reactive oxo-heme intermediates, thereby directing substrate oxidation and minimizing formation of unwanted reactive oxygen species, such as superoxide ion (O_2^-) and nitroxyl anion (NO^-), as side products. A calcium-calmodulin binding site is located between the two domains. The constitutive enzymes, nNOS and eNOS, are activated by Ca^{2+}-calmodulin binding at this site, which in turn is regulated by the amount of Ca^{2+} in the medium. In contrast, Ca^{2+}-calmodulin binds so tightly to iNOS that this enzyme is effectively "turned on" under all physiological conditions. The practical consequence of this difference in binding affinities is that the constitutive enzymes generate only low levels of NO, whereas, once its formation has been induced, iNOS generates relatively massive amounts of NO. These differences in enzymatic activity are consistent with the presumed functions of the enzymes, which are secondary messenger generation and cytotoxin production, respectively. *See* AMINO ACIDS; ENZYME.

Blood pressure regulation. In the cardiovascular system, NO is produced by eNOS located in the endothelial layer, the cells that line the blood vessels. The activation of eNOS is triggered by the binding of hormones to receptors on the surface of these cells, opening up calcium channels that release Ca^{2+} into the endothelial cells. The Ca^{2+} then attaches to the cofactor calmodulin, which binds to and activates eNOS. The NO that is produced diffuses to the smooth muscle and binds to the iron atom of the heme group in an enzyme known as soluble guanylyl cyclase (sGC). This enzyme converts guanosine triphosphate to cyclic guanosine monophosphate and pyrophosphate within cells. The binding of NO to the heme of sGC leads to a structural change in the enzyme which causes a dramatic increase in the production of cyclic guanosine monophosphate (cGMP). It is thought that the cGMP activates a protein kinase (an enzyme that adds a phosphate

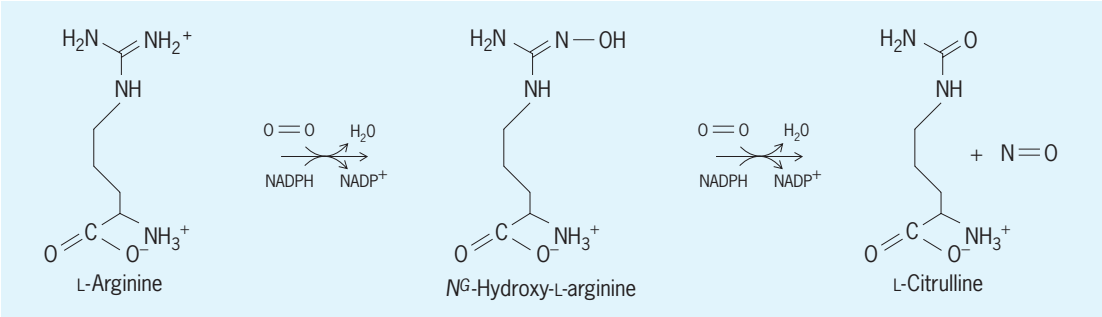

Fig. 1. Conversion of L-arginine to L-citrulline and nitric oxide (NO) catalyzed by the enzyme nitric oxide synthase. (*After P. L. Feldman et al., The surprising life of nitric oxide, C&EN News, 71(51):26–38, 1993*)

group to another protein) that phosphorylates calcium transporters, causing Ca^{2+} to be returned to its cellular storehouse within the smooth muscle cells. The decrease in calcium levels causes relaxation of blood vessels since the muscle cells require Ca^{2+} for contraction. The treatment of hypertension, heart attacks, and other blood pressure abnormalities often involves using chemical compounds that can release NO in the body. These compounds are thought to bypass the nitric oxide synthase pathway and directly release NO into smooth muscles, causing relaxation and thus decreasing the blood pressure. Direct inhalation of gas mixtures containing low levels of NO has also been used effectively to combat pulmonary hypertension in neonatal and adult respiratory distress syndrome patients. *See* CARDIOVASCULAR SYSTEM.

The protein hemoglobin, the O_2 carrier in the blood, has also been implicated in NO transport and regulation of blood pressure. Hemoglobin contains four heme groups that function as O_2-binding sites. Nitric oxide reacts with hemoglobin in several distinct ways. First, it reacts rapidly with O_2 bound to the iron atom of the heme to form NO_3^- and methemoglobin, an oxidized form of hemoglobin which cannot bind O_2. Second, the iron atom of the heme in hemoglobin can bind NO when O_2 is not bound. Since hemoglobin is present in high concentrations in the vascular system and hemoglobin can effectively scavenge NO, it has been a mystery how enough NO can reach the smooth muscle cells to cause blood pressure relaxation. There is now evidence that hemoglobin may protect NO by forming S-nitrosothiols (RS-NO compounds, where R represents a hydrocarbon group). Research suggests that two conserved cysteine residues in hemoglobin react with nitric oxide to form nitrosothiols, preventing the oxidation of NO to NO_3^-. The NO could then be released to the smooth muscle cells, causing blood vessel relaxation. Nonetheless, the physiological relevance of protein S-nitrosothiols is controversial, and the issue of its importance as a delivery pathway for NO is unresolved. *See* HEMOGLOBIN.

Neurotransmitter. Within the central nervous system, nitric oxide synthase (nNOS) is found only in discrete populations of neuronal cells in brain tissue, and is not found in the glial cells that make up the major portion (85%) of the cellular mass. Although specific physiological roles for nNOS-containing neurons have not been established, nNOS is activated in some cells by Ca^{2+}-calmodulin binding during neurotransmission. In this case, glutamate released from a stimulated (presynaptic) neuron diffuses to binding sites on an unstimulated adjacent (postsynaptic) neuron resulting in influx of Ca^{2+} into the neuron, increasing as well the Ca^{2+}-calmodulin concentration and activating nNOS. The NO formed appears to promote additional release of neurotransmitters and activates sGC to form the secondary messenger cGMP. It has been suggested that these reactions lead to changes in the neuronal pathways that contribute to development of memory. The overproduction of nitric oxide in brain tissues has been implicated in stroke and other neurological problems initiated by oxidative damage to deoxyribonucleic acid (DNA). *See* DEOXYRIBONUCLEIC ACID (DNA).

The peripheral nervous system also uses the production of NO in neurons to regulate blood flow. Neurons that are sensitive to NO but unresponsive to conventional neurotransmitters have been found in many peripheral tissues, including those of the cardiovascular, urogenital, respiratory, and digestive systems. Nitric oxide has also been implicated in vision and the sense of smell. Penile erection is mediated by NO production in peripheral neurons, causing relaxation of smooth muscles in that organ. Thus, drugs that stimulate release of NO have been developed as a noninvasive way of treating impotence. *See* NEUROBIOLOGY.

Cytotoxicity and cytoprotection. Macrophages are phagocytic (devouring) cells within the immune system whose primary functions are to ingest and destroy invading microbes, kill targeted tissue cells, and digest cellular debris. They also contain iNOS, whose biosynthesis is induced by small molecules derived from the microbes and/or other host cells. Nitric oxide released by macrophages can inhibit cellular processes in nearby cells, including DNA synthesis and respiration, as well as selective inhibition of intracellular enzymes involved in other essential metabolic pathways. Primary target sites in these biomolecules include nonheme iron-sulfur clusters, sulfhydryl substituents of cysteine amino acids at metal-binding and enzyme active sites, and the phenolic substituent of tyrosine, as well as various sites within DNA. The resulting damage is sufficient to induce death in many types of cells by either of the two recognized pathways—programmed cell death (apoptosis) or necrosis. The extent to which these reactions are caused by direct reaction of the biomolecules by NO or by secondary oxidants derived from NO is an area of active ongoing research. It appears that most of the oxidative damage mediated by NO is actually caused by secondary oxidants, which is consistent with the limited chemical reactivity displayed by NO. Compounds thought to be major contributors are indicated in **Fig. 2**. These include (1) dinitrogen trioxide (N_2O_3), a potent nitrosating (that is, NO^+-transferring) agent that is formed during reaction of NO with O_2 (a diradical); (2) the peroxynitrite ion ($ONOO^-$), which is formed by radical coupling of NO and O_2^-, and when protonated or in the presence of carbon dioxide undergoes homolytic cleavage of its peroxo O-O bond to give the strongly oxidizing radicals NO_2 and OH or CO_3^-, respectively;

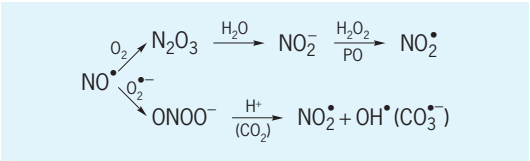

Fig. 2. Secondary oxidants formed upon reaction of NO with O_2 and superoxide ion. PO = peroxidase enzyme.

(3) the nitrite anion (NO_2^-), which can be oxidized to NO_2 radical by H_2O_2 in reactions catalyzed by peroxidase enzymes. Additional potential oxidants are metal nitrosyl complexes, which could act as nitrosating agents or could release nitroxyl (NO^-) to the medium, where it could react rapidly with O_2 to form peroxynitrite. Confounding attempts to identify the origins of cellular oxidative damage caused by NO is the recognition that under some circumstances NO can also protect cells from damage by other oxidants. For example, NO has been shown to inhibit lipid peroxidation, most likely by coupling with reactive lipid peroxy radical intermediates, and thereby terminating membrane-destroying radical chain reactions with O_2. *See* APOPTOSIS; IMMUNOLOGIC CYTOTOXICITY; PEROXYNITRITE.

Although nitric oxide production in the immune system serves a crucial biological function, there can be adverse effects when too much NO is produced. During a massive bacterial infection, excess NO can go into the vascular system, causing a dramatic decrease in blood pressure, which may lead to fatal septic shock. Thus, scientists are working on drugs that can selectively inhibit iNOS in order to avoid the harmful effects produced by excess NO without interfering with nitric oxide signaling pathways. *See* IMMUNOLOGY.

James K. Hurst; Judith N. Burstyn; Mark F. Reynolds

Bibliography. P. L. Feldman, O. W. Griffith, and D. J. Stuehr, The surprising life of nitric oxide, *Chem. Eng. News*, 71(51):26–38, 1993; R. Rawls, Bioinorganic reactions of nitric oxide underlie diverse roles in living systems, *Chem. Eng. News*, 74(19):38–42, 1996; S. H. Snyder and D. S. Bredt, Biological roles of nitric oxide, *Sci. Amer.*, 266(5):68–77, 1992.

Nitrile

One of a group of organic chemical compounds of general formula RC≡N. A nitrile is named from the acid to which it can be hydrolyzed by adding the suffix -onitrile to the acid stem, for example, acetonitrile from acetic acid. An alternative system names the group attached to CN, thus CH_3CN is also named methyl cyanide. In more complex structures the CN group is named as a substituent, cyano.

Nitriles may be identified by a distinctive weak-to-medium infrared absorption band due to triple-bond stretching at 2260–2222 cm^{-1}.

Nitriles are hydrolyzed to acids in either basic or acidic solution, as shown in reactions (1) and (2).

$$RCN + OH^- + H_2O \longrightarrow RCOO^- + NH_3 \qquad (1)$$

$$RCN + H^+ + 2H_2O \longrightarrow RCOOH + NH_4^+ \qquad (2)$$

By catalytic hydrogenation in the presence of nickel or cobalt, nitriles are reduced to primary amines, reaction (3). Nitriles are also reduced to primary amines by lithium aluminum hydride.

$$RCN + 2H_2 \longrightarrow RCH_2NH_2 \qquad (3)$$

Grignard reagents added to nitriles give ketones (after hydrolysis), as shown in reaction (4).

$$RCN + R'MgX \longrightarrow RR'C{=}NMgX \xrightarrow{H_2O} RCOR' \qquad (4)$$

The formation of cyanides from alkyl halides is an important chain-lengthening reaction in organic synthesis. The starting compound is most often an alcohol, and the sequence of reactions is shown in (5).

$$ROH \xrightarrow{HX} RX \xrightarrow{NaCN} RCN \qquad (5)$$

The reaction is practical only with primary aliphatic halides, since the alkali cyanides are fairly strong bases and eliminate HX from secondary or tertiary alkyl halides.

Aromatic nitriles are made by displacement of a diazotized primary amino group with the cyanide group in the presence of copper cyanide or copper powder, reaction (6).

$$ArNH_2 \xrightarrow{HONO} ArN_2^+ \xrightarrow[NaCN]{Cu} ArCN \qquad (6)$$

The dehydration of acid amides or oximes with phosphorus pentoxide or acetic anhydride in either the aliphatic or aromatic series serves as another preparative method, reactions (7) and (8).

$$3RCONH_2 + P_2O_5 \longrightarrow 3RCN + 2H_3PO_4 \qquad (7)$$

$$RCH{=}NOH + (CH_3CO)_2O \longrightarrow RCN + 2CH_3COOH \qquad (8)$$

Industrially, nitriles are formed by heating carboxylic acids with ammonia and a dehydration catalyst under pressure. The amide is an intermediate but need not be isolated, reaction (9).

$$RCOOH + NH_3 \longrightarrow [RCONH_2] \longrightarrow RCN \qquad (9)$$

For the preparation of acrylonitrile, which is used on a large scale in the plastics industry, a vapor-phase catalytic ammoxidation of propylene has been developed, reaction (10).

$$CH_3CH{=}CH_2 + NH_3 + {}^3/_2O_2 \longrightarrow H_2C{=}CHCN + 3H_2O \qquad (10)$$

Acrylonitrile is electrochemically coupled "tail to tail" to give the hydrodimer adiponitrile, which is a key intermediate in the commercial production of nylon-6,6, reaction (11).

$$2H_2C \longrightarrow CHCN \longrightarrow {}_{2e^-\,2H^+}NC(CH_2)_4CN \qquad (11)$$

See ACRYLONITRILE; AMINE; CARBOXYLIC ACID; OXIME; POLYAMIDE RESINS.

Paul E. Fanta

Bibliography. S. N. Ege, *Organic Chemistry*, 4th ed., 1998; R. T. Morrison and N. Boyd, *Organic Chemistry*, 6th ed., 1992.

Nitro and nitroso compounds

Derivatives of organic hydrocarbons having one or more —NO_2 groups bonded via nitrogen to the carbon framework (nitro compounds) or an —NO group attached to carbon or nitrogen (nitroso compounds).

Nitro compounds. These are isomeric with nitrite esters, which contain the —O—N═O substituent. The major classes of nitro compounds are the nitroarenes, nitroalkanes, and nitroalkenes.

Nitroarenes, also known as aromatic nitro compounds, are readily prepared by electrophilic aromatic substitution reactions using nitrating reagents. Benzene reacts with a mixture of concentrated nitric and sulfuric acids ($NO_2^+HSO_4^-$) at <50°C (122°F) to produce nitrobenzene (86% yield). Further nitration under more forcing conditions (using hot fuming nitric acid and concentrated sulfuric acid) gives *meta*-dinitrobenzene [reaction (1)]. Deriva-

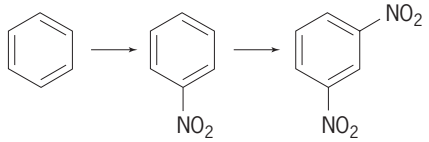

(1)

tives of benzene that contain electron-donating substituents (OCH_3, CH_3, $NHCOCH_3$, and so forth) are nitrated more readily than benzene to produce *ortho*- and *para*-nitrated products. However, benzene derivatives bearing electron-withdrawing substituents ($COCH_3$, CO_2CH_3, CN, and so forth) are nitrated more slowly to produce *meta*-nitrated products. Nitration of heterocyclic aromatic molecules is used to produce nitro derivatives. *See* BENZENE; HETEROCYCLIC COMPOUNDS.

Nitroarenes are reduced to aromatic amines by using dissolving metals and acid, catalytic hydrogenation, or polysulfide salts. The reductions may be stopped at the intermediate arylhydroxylamine stage. Reductions of nitroarenes under strongly basic conditions are used for the preparations of azoxy, azo, and hydrazo compounds. These products are formed by initial condensation of the reactive nitroso intermediate and further reductions [reaction (2)], where Ar = aryl.

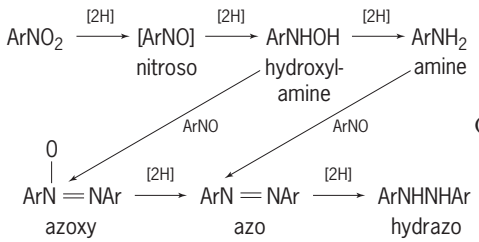

(2)

Nitroarenes, containing good leaving groups (groups of atoms that depart during the reaction) at the ortho or para position, readily undergo nucleophilic aromatic substitution reactions via an anionic intermediate known as the Meisenheimer complex. For example, 2,4,6-trinitrochlorobenzene reacts with sodium hydroxide to produce 2,4,6-trinitrophenol (picric acid). *See* NITROAROMATIC COMPOUND.

Nitroarenes and nitroalkanes are useful intermediates for the preparation of dye intermediates, explosives, pharmaceuticals, and other specialty chemicals. *See* NITRATION; NITROPARAFFIN; ORGANIC REACTION MECHANISM; REACTIVE INTERMEDIATES.

A. G. M. Barrett

Nitroso compounds. Nitroso compounds contain the —NO group. It may be attached to carbon (C-nitroso) or to nitrogen (N-nitroso). When the —NO group is attached to oxygen, the compounds are known as nitrites. The nitroso group —N═O contains a double bond (between the nitrogen and the oxygen) and a lone pair of electrons on the nitrogen atom. It is reactive toward acids, bases, radicals, nucleophiles, electrophiles, oxidizing agents, and reducing agents. Nitroso compounds are colored (blue or green) and photoreactive (reactive toward light). An important property of the nitroso group is the tendency to isomerize to an oxime; this is restricted to aliphatic nitroso compounds containing a hydrogen atom on the carbon attached to the nitroso group, for example, $R_2CH—NO \rightarrow R_2C{=}NOH$. Another important property is the tendency to dimerize to nitroso dimers, *N,N′*-azo dioxides, for example, 2R—NO → R—N(O)═N(O)—R. *See* ELECTROPHILIC AND NUCLEOPHILIC REAGENTS; OXIME.

C-nitroso compounds may be prepared by the oxidation of amines (RNH_2) or hydroxylamines (RNHOH). Direct methods for the introduction of the nitroso group are applicable in certain cases, such as the conversion of aromatic compounds to aryl nitroso compounds (Aryl-H → aryl-NO) by an electrophilic aromatic substitution reaction with the nitrosonium ion (NO^+).

C-nitroso compounds react with free radicals, yielding nitroxyl radicals, as is shown in reaction (3), where R^a and R^b is alkyl or aryl. Nitroxyl rad-

$$R^a + R^b—NO \longrightarrow R^aR^bN—O \qquad (3)$$

icals are quite stable and are easily detected even at very low concentration by the technique of electron spin resonance spectroscopy. This method, known as spin trapping, has been used in studies of highly reactive radicals, including carbon, nitrogen, oxygen, and sulfur radicals. Because of this reactivity, nitroso compounds also have found applications as polymerization inhibitors. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; ELECTRON SPIN; FREE RADICAL; INHIBITOR (CHEMISTRY); SPIN LABEL.

In solution, the C-nitroso form is in equilibrium with a dimeric form (nitroso dimer, *N,N′*-azo dioxide), capable of existence as trans and cis isomers. The monomer-dimer equilibrium connects a colored, nonpolar, chemically reactive species with a colorless, more polar chemically unreactive species. In general, the free-energy term for reaction (4) favors

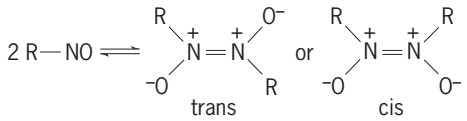

(4)

the dimer, the entropy term favors the monomer, and the enthalpy term favors the dimer. Conjugation effects favor the monomer (nitroso form).

Consequently, aryl nitroso compounds (for example, nitrosobenzene) are largely monomeric in solution. Almost all other C-nitroso compounds are dimeric (and colorless) in the solid state. *See* CONJUGATION AND HYPERCONJUGATION; ENTHALPY; ENTROPY; FREE ENERGY; MOLECULAR ISOMERISM.

N-nitroso compounds (*N*-nitrosamines) have the general structure $R^aR^bN{-}NO$. Secondary *N*-nitrosamines are moderately stable, monomeric, yellow substances. The structure is intermediate between $R^aR^bN{-}NO$ and $R^aR^bN^{+}{=}N{-}O^{-}$. N-nitroso compounds are easily formed by reaction of a secondary amine with various nitrosating agents [reaction (5). Many N-nitroso compounds are

$$R_2NH + X{-}N{=}O \longrightarrow R_2N{-}N{=}O + HX \qquad (5)$$
$$X = HO, RO, Cl, ONO, \text{or} R'N$$

strong carcinogens, and some occur widely in the environment. The carcinogenicity is thought to be associated with biological conversion (alpha-hydroxylation) of an N-nitroso compound into an alkylating agent as in reaction scheme (6), followed

$$R'_2CH{-}N(R){-}NO \longrightarrow R'_2C(OH){-}N(R){-}NO \longrightarrow$$
$$R'_2C{=}O + R{-}N{=}N{-}OH \longrightarrow \text{“}R{-}N_2{+}\text{”} \qquad (6)$$

by alkylation of nucleophilic sites of cellular constituents such as the N-7 and O-6 positions of guanine in deoxyribonucleic acid (DNA). *See* DEOXYRIBONUCLEIC ACID (DNA); NITROGEN.

Frederick D. Greene

Bibliography. A. G. M. Barrett and G. G. Graboski, Conjugated nitroalkenes: Versatile intermediates in organic synthesis, *Chem. Rev.*, 86:751, 1986; D. H. Barton, *New Principles in Organic Chemistry: The Discovery and Invention of Chemical Reactions*, 1995; D. H. R. Barton and W. D. Ollis (eds.), *Comprehensive Organic Chemistry*, vol. 2, 1979; J. March, *Advanced Organic Chemistry: Reactions, Mechanisms, and Structures*, 4th ed., 1992; S. Patai (ed.), *The Chemistry of Functional Groups*, suppl. F, *The Chemistry of Amino, Nitroso, and Nitro Groups and Their Derivatives*, pt. 1 and 2, 1982; S. Patai (ed.), *The Chemistry of Nitro and Nitroso Groups*, pt. 1, 1969, pt. 2, 1970.

Nitroaromatic compound

A member of the class of organic compounds in which the nitro group ($-NO_2$) is attached directly to the cyclic, aromatic nucleus. The prototypal compound is nitrobenzene, which was first synthesized in 1834 and produced commercially in England in 1856. It is prepared by the reaction of benzene with nitric acid in the presence of sulfuric acid, as shown in reaction (1). The reaction is very efficient, with

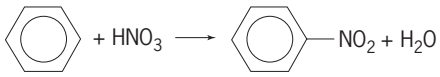

(1)

Benzene Nitrobenzene

yields approaching 100%. It is an exothermic reaction, and careful control of the temperature is necessary to avoid the formation of *m*-dinitrobenzene (**1**).

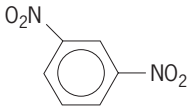

(**1**)

See BENZENE.

Nitrobenzene is a pale-yellow liquid with a characteristic sweet but unpleasant odor. It boils at 211°C (412°F), and therefore it has a fairly high vapor pressure at room temperature. It is a very toxic substance that is absorbed by contact of the liquid with the skin or by inhalation of the vapor. In the body it reacts with the red blood cells to form methemoglobin, resulting in cyanosis. Chronic exposure can lead to spleen and liver damage, jaundice, and anemia.

The most significant use of nitrobenzene is in the manufacture of aniline (**2**). About 97% of the ni-

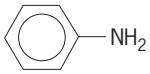

(**2**)

trobenzene produced in the United States is converted to aniline, which is used in the manufacture of plastics, rubber additives, dyes, drugs, and other products.

Second to nitrobenzene in commercial importance are the mononitrotoluenes, particularly the ortho and para isomers, (**3**) and (**4**), respectively. Re-

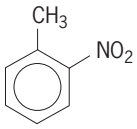

(**3**)

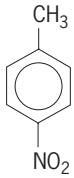

(**4**)

action of toluene with a mixture of nitric and sulfuric acid at about 40°C (104°F) gives a high yield of a mixture of the three isomers containing about 58% ortho, 4% meta (**5**), and 38% para. The isomers

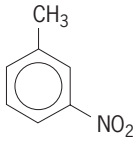

(**5**)

are separated by a combination of fractional distillation and crystallization. The nitrotoluenes are important intermediates in the preparation of dyes, rubber chemicals, and agricultural chemicals.

2,4,6-Trinitrotoluene (TNT; **6**) is a military explo-

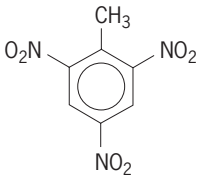

(6)

sive that is stable, nonhygroscopic, and relatively insensitive to impact, friction, shock, and electric spark. It is produced by nitration of toluene in successive stages at progressively higher temperatures and concentrations of acid. It can be melted safely at 80°C (176°F) for cast loading of military projectiles. *See* EXPLOSIVE.

Picric acid (2,4,6-trinitrophenol) is an example of a nitroaromatic compound that cannot be prepared by direct nitration, since phenol is sensitive to oxidation. The commercial method consists of a series of reactions starting with chlorobenzene, including the hydrolysis of the chlorine to give the desired hydroxyl group, reaction (2). Picric acid has been used

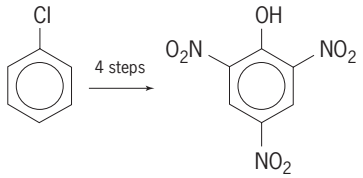

(2)

as a dye, an explosive, and a bactericide.

Similarly, the useful explosive tetryl is formed by a sequence of nitration steps starting with *N*,*N*-dimethylaniline and including the replacement of one of the methyl groups by the nitro group, reaction (3).

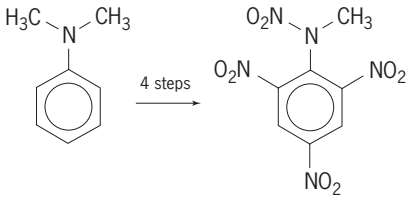

(3)

Furacin (5-nitro-2-furaldehyde semicarbazone) is an example of a heterocyclic nitroaromatic compound that has attained commercial success. It is prepared in several steps from readily available furfural, reaction (4), and is an antibacterial agent that

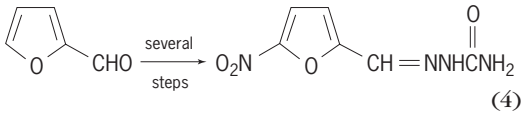

(4)

is particularly effective against poultry coccidiosis. *See* HETEROCYCLIC COMPOUNDS.

Although literally thousands of other aromatic ring compounds, including the heterocyclics, have been converted to their nitro derivatives, few such compounds have achieved any significant industrial importance. *See* AROMATIC HYDROCARBON; NITRATION.

Paul E. Fanta

Bibliography. F. A. Carey and R. J. Sundberg, *Advanced Organic Chemistry*, pt. A: *Structure and Mechanisms*, 3d ed., 1990; L. F. Fieser and M. Fieser, *Advanced Organic Chemistry*, 1961; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., 1995; *Merck Index*, 11th ed., 1989.

Nitrogen

A chemical element, N, atomic number 7, atomic weight 14.0067. Nitrogen, a gas under normal conditions, is the lightest element of periodic group 5 (nitrogen family). *See* PERIODIC TABLE.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

At standard temperature and pressure, elemental nitrogen exists as a gas with a density of 1.25046 g/liter. This value indicates that the molecular formula is N_2. Some physical properties of elemental nitrogen are listed in **Table 1**.

Elemental nitrogen has a low reactivity toward most common substances at ordinary temperatures. At high temperatures, molecular nitrogen, N_2, reacts with chromium, silicon, titanium, aluminum, boron, beryllium, magnesium, barium, strontium, calcium, and lithium (but not the other alkali metals) to form nitrides; with O_2 to form NO; and at moderately high temperatures and pressures in the presence of a catalyst, with hydrogen to form ammonia. Above 1800°C (3300°F), nitrogen, carbon, and hydrogen combine to form hydrogen cyanide.

Table 2 lists the principal classes of inorganic nitrogen compounds. Thus, in addition to the typical oxidation states of the family (−3, +3, and +5), nitrogen forms compounds with a variety of additional

TABLE 1. Properties of nitrogen

Property	Value
Heat of transformation (α–β)	54.71 cal/mole
Heat of fusion	172.3 cal/mole
Heat of vaporization	1332.9 cal/mole
Critical temperature	126.26 ± 0.04 K
Critical pressure	33.54 ± 0.02 atm
Density: α form	1.0265 g/ml at −252.6°C
β form	0.8792 g/ml at −210.0°C
Liquid	1.1607–0.0045T (T = abs temp)

TABLE 2. Compounds of nitrogen

Oxidation state	Examples
+5	N_2O_5, HNO_3, nitrates, NO_2X
+4	$N_2O_4 \rightleftharpoons 2NO_2$
+3	N_2O_3, HNO_2, nitrites, NOX, NX_3
+2	NO, Na_2NO_2, nitrohydroxylamates
+1	N_2O, $H_2N_2O_2$, hyponitrites
0	N_2
−1/3	HN_3, acids
−1	NH_2OH, hydroxylammonium salts
−2	NH_2NH_2, hydrazinium salts, hydrazides
−3	NH_3, ammonium salts, amides, imides, nitrides

oxidation states. *See* AMINE; AMMONIA; HYDRAZINE; NITRIC ACID; NITROGEN COMPLEXES; NITROGEN OXIDES.

Molecular nitrogen is the principal constituent of the atmosphere (78% by volume of dry air), in which its concentration is a result of the balance between the fixation of atmospheric nitrogen by bacterial, electrical (lightning), and chemical (industrial) action, and its liberation through the decomposition of organic materials by bacteria or combustion. In the combined state, nitrogen occurs in a variety of forms. It is a constituent of all proteins (both plant and animal) as well as of many other organic materials. Its chief mineral source is sodium nitrate.

The methods for the preparation of elementary nitrogen may be grouped into two classes, separation from the atmosphere and decomposition of nitrogen compounds. The industrial method for the production of nitrogen is the fractional distillation of liquid air. Nitrogen containing about 1% argon and traces of other inert gases may be obtained by the chemical removal of oxygen, carbon dioxide, and water vapor from the atmosphere by appropriate chemical reagents.

Because the importance of nitrogen compounds in agriculture and chemical industry, much of the industrial interest in elementary nitrogen has been in processes for converting elemental nitrogen into nitrogen compounds. The principal methods for doing this are the Haber process for the direct synthesis of ammonia from nitrogen and hydrogen, the electric are process, which involves the direct combination of N_2 and O_2 to nitric oxide, and the cyanamide process. Nitrogen is also used for filling bulbs of incandescent lamps and, in general, wherever a relatively insert atmosphere is required. Harry H. Sisler

Bibliography. T. Chivers, *A Guide To Chalcogen-nitrogen Chemistry*, 2004; J. S. Clark (ed.), *Nitrogen, Oxygen and Sulfur Ylide Chemistry*, 2002; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999.

Nitrogen complexes

Compounds containing the dinitrogen molecule, N_2, bound to a metal (also called dinitrogen complexes). While it was long known that molecular nitrogen avidly binds to the surface atoms of many metallic phases, no reaction of molecular nitrogen has been observed under the conditions of ordinary solution chemistry. Indeed, chemists had been so accustomed to regard nitrogen as a completely inert gas at ordinary temperatures that the formation of coordination compounds with molecular nitrogen aroused general surprise when first reported in the mid-1960s. Since then the interaction of molecular nitrogen with coordination compounds has been the subject of intensive research. Today the capability of molecular nitrogen to enter into coordination compounds as a ligand is established beyond any doubt, and it is at least partly understood under which conditions the dinitrogen molecule can lose its customary inertness. *See* COORDINATION CHEMISTRY.

Formation. Outstanding in their ability to form coordination compounds with nitrogen are a number of metals which belong to the group 18 transition metal family (**Fig. 1**). For each metal of this group, several nitrogen complexes have been identified. Nitrogen complexes of these metals occur in low oxidation states, such as Co(I) or Ni(O). The other ligands present in these complexes besides N_2 are usually of a type known to stabilize low oxidation states; phosphines appear to be particularly prominent bonding partners in this respect. **Figure 2** shows the structure of a typical N_2 complex, elucidated by a crystal structure determination. The N-N bond axis in this complex is aimed, within the limits of experimental error, directly toward the position of the metal atom. The Co-N_2 bond length, 0.18 nanometer, is within the normal range of comparable metal-ligand bonds.

Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn
Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd
La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg

Fig. 1. Transition metals which form coordination compounds with N_2 as ligand are mostly members of groups shown in colored squares. Coordination compounds which reduce N_2 to NH_3 in approximately stoichiometric yields are derived from the metals in the left half of the transition series (circles).

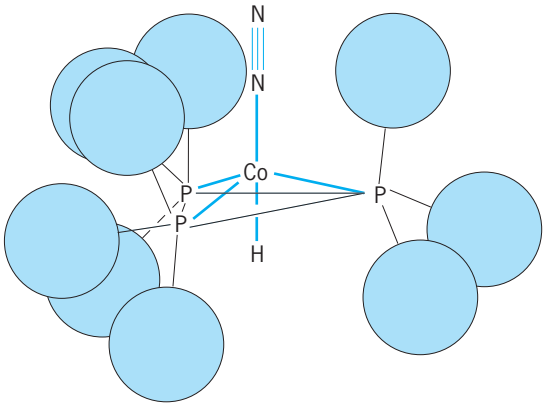

Fig. 2. Structure of a coordination compound with N_2 (circles represent phenyl groups).

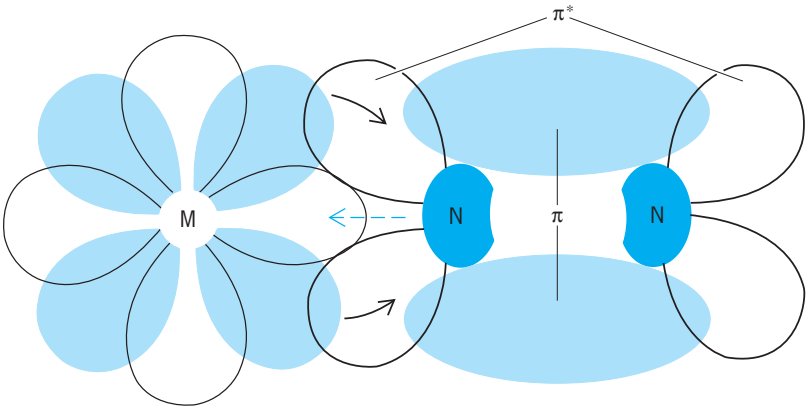

Fig. 3. Factors contributing to the stability of N_2 coordination compounds: dative bond from an electron pair on one of the N atoms to a vacant metal orbital (broken arrow) and the opposite process of back-donation of loosely held metal electrons to the π^* acceptor orbital of N_2 (solid arrows).

The question arises as to the kind of bonding force that is able to attract the neutral, dipole-free, and not easily polarizable N_2 molecule to its position in the coordination sphere of the metal. The usual kind of dative bond between the electron pair on one of the N atoms and an acceptor orbital of the metal (**Fig. 3**) is likely to be very weak, just as the basicity of N_2 is unmeasurably low. A second and probably more important contribution to the stability of nitrogen complexes is thought to arise from the opposite process—from back-donation of loosely held metal electrons into an acceptor orbital of the N_2 molecule. Such a shift of electron density into the normally unoccupied, antibonding π^* orbitals of the ligand is considered to be an important stabilizing factor in many coordination compounds with unsaturated ligands. While strengthening the metal-ligand bond, the flow of electrons from the metal into these antibonding ligand orbitals would necessarily cause a weakening of the bond between the two N atoms, and this is exactly what is observed. Without exception, the stretching vibration of coordinated nitrogen occurs at a substantially lower frequency than that of free, gaseous nitrogen. Moreover, the known prerequisites for an efficient back-donation explain at least some of the peculiarities of nitrogen coordination. A large number of *d* electrons, usually six or more, has to be available on the metal; in addition, the oxidation state of the metal has to be low so that these electrons are easily released toward the ligand. Accordingly, it is understandable that group VIII metals with their large number of *d* electrons dominate among the known N_2 complexes.

In all the respects discussed above, the N_2 ligand exhibits a strong resemblance to the well-known ligand molecule carbon monoxide, CO, which accommodates the same number of electrons in orbitals of the same type as N_2. Although N_2 by virtue of its symmetry is a much weaker coordinating agent than the lopsided CO molecule, there is no doubt that these two molecules can under certain conditions substitute for each other as ligands. The weaker ligand N_2 appears to be more selective, however. Other ligand partners in the complex, while having to accommodate enough of the metal's excess electrons to stabilize a low oxidation state, should not compete too strongly with N_2 for back-donation of the metal electrons. Phosphines, which are known to draw less heavily on metal *d* electrons than other π acceptors, seem to strike this delicate balance right. That this is not the sole possibility to obtain an appropriate electron distribution, however, is demonstrated by the stability of the complex $[Ru(N_2)(NH_3)_5]^{2+}$, where five NH_3 ligands "condition" the Ru(II) center.

A number of dinitrogen complexes with entirely different structures have been isolated or characterized in solutions and solid matrices. These complexes appear to have their N_2 ligand molecule bound to the metal in an "edge-on" fashion (similar to that in **Fig. 4**, with both of its N atoms at the same bonding distance from the central metal atom. The general conditions for the occurrence of this type of N_2 complex are not clear at present.

Chemical reactions. Even in most favorable cases the binding of the dinitrogen molecule to the metal is fairly labile; all the compounds lose their nitrogen on mild heating. Some of the nitrogen complexes are only metastable to loss of dinitrogen even at room temperature; accordingly, they cannot be obtained by direct uptake of gaseous nitrogen. In the synthesis of these metastable complexes, hydrazine or azide compounds serve as a source of nitrogen molecules within the coordination sphere of the metal. Addition of other coordinating agents to the nitrogen complexes usually results in a displacement of N_2 from the metal. The cobalt compound in Fig. 2 exchanges its N_2 ligand quite reversibly for other ligand molecules, such as NH_3 and $H_2C{=}CH_2$. Whereas these ligands are easily displaced again by an excess of N_2, an irreversible exchange occurs with carbon monoxide. The bulky organic groups on the phosphine ligands are likely to interfere with the approach to the metal of all but the slimmest ligands and thereby help the "thin" dinitrogen molecule to maintain or regain its position on the metal in competition with most other ligands.

An interesting question concerns the reactivity of the coordinated nitrogen molecule. Is the weakening of the bond in N≡N, which accompanies coordination and back-donation, sufficient to render the otherwise inert nitrogen molecule susceptible to attack and cleavage? Much systematic research has yet to be done if scientists are to understand the perplexing chemical reaction paths involved in nitrogen fixation.

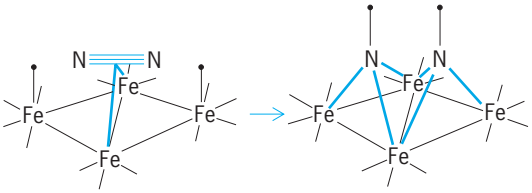

Fig. 4. Conceivable reaction path for the conversion of a surface-bound N_2 molecule to two surface nitride particles. Surface hydrogen may be involved in this reaction (indicated by black dots).

The conversion of inert, molecular nitrogen to ammonium salts, which can then be utilized by plants in their production of proteins and other indispensable nutrients, is one of the limiting factors of life on Earth. Nature has put the burden of this life-sustaining process on the few strains of microorganisms which possess an enzyme capable of reacting with nitrogen. The enzyme, nitrogenase, actually consists of two protein particles, both of which contain iron in a peculiar form associated with sulfide; one of the enzyme particles also contains molybdenum. There is some support for the hypothesis that this latter particle forms a complex with molecular nitrogen, while the other molybdenum-free enzyme particle serves as a sort of energizer, accepting reducing electron equivalents from ferredoxin and making them available for the N_2-reduction process at a substantially increased negative reduction potential at the expense of adenosine triphosphate hydrolysis. As for the chemical pathways by which the enzyme entrains its N_2 ligand into its smooth reduction to two NH_3 molecules, there is a still-open debate about the possible occurrence of enzyme-bound intermediates such as N_3H_2 (diimide) and N_2H_4 (hydrazine). *See* NITROGEN FIXATION.

Hardly any better understood is the process by which humans supplement natural nitrogen fixation—the production of fertilizer ammonia from N_2 and H_2 by a high-temperature reaction on metal catalysts. Bonding of nitrogen to the surface of some metals, such as nickel, is spectrally reminiscent of the complexes discussed above and probably involves the same type of end-on coordination of N_2 to surface atoms of the metal. In many other cases, however, there seem to be different types of surface coordination, perhaps with the N≡N molecule parallel to the surface, which then can probably rearrange directly to two surface-bound nitride particles (Fig. 4). While many metals form this nitride type of surface compounds with nitrogen, the distinction of catalytically active metals is their tendency to detach the nitride particles from the surface again, together with surface hydrogen particles, to form gaseous NH_3. The metal surface is thereby restored in its original form and is ready to renew the catalytic cycle by coordination of another N_2 molecule. Because of the complexity of these catalyst systems, some important questions have not been completely solved. It is not clear, for instance, whether the breakdown reaction of the coordinated N_2 molecule is just a rearrangement of bonds and electrons between the metal and nitrogen atoms, or if the participation of surface hydrides assists in this breakdown step, which might then involve adsorbed N_2H_2 or N_2H_4 intermediates.

A related problem is encountered with a series of reactions in which an N_2 ligand bound to some metal compound in solution is reduced to NH_3 or N_2H_4. These reactions involve mostly transition metals that are located in the left half of the transition series (Fig. 1). N_2 reductions of this kind have been known since about 1965 to occur in the presence of strong reducing agents. Subsequently some reactions have been found where a mere transfer of protons to the metal-bound (and hence presumably negatively polarized) N_2 molecule of an otherwise stable dinitrogen complex brings about reduction of this molecule to N_2H_4 or NH_3, while leaving the metal behind in a correspondingly increased oxidation state. In view of their obvious relation to enzyme and surface catalysis, it would be highly interesting to establish the relevant intermediates of these reactions—for example, the bonding geometry of the N_2 ligand molecules involved—and the general conditions for the occurrence of N_2 reduction reactions of this kind. The possibility that reactive intermediates such as N_2H_2 or N_2H_4 derived from catalytic reactions of this type might be trapped into useful synthetic reactions is presently under exploration in many laboratories; its realization would be a step forward in the direction of making the vast supply of atmospheric nitrogen more efficiently available for utilization. *See* COORDINATION COMPLEXES; MOLECULAR ORBITAL THEORY; NITROGEN.

Hans Brintzinger

Bibliography. G. Henrici-Olive, *Coordination and Catalysis*, 1977; M. M. Khan and A. E. Martell, *Homogeneous Catalysis by Metal Complexes*, 2 vols., 1974.

Nitrogen cycle

The collective term given to the natural biological and chemical processes through which inorganic and organic nitrogen are interconverted. It includes the processes of ammonification, ammonia assimilation, nitrification, nitrate assimilation, nitrogen fixation, and denitrification.

Nitrogen exists in nature in several inorganic compounds, namely N_2, N_2O, NH_3, NO_2^-, and NO_3^-, and in several organic compounds such as amino acids, nucleotides, amino sugars, and vitamins. In the biosphere, biological and chemical reactions continually occur in which these nitrogenous compounds are converted from one form to another. These interconversions are of great importance in maintaining soil fertility and in preventing pollution of soil and water.

Nitrogen reserves exist in five major sinks: the primary rocks, the sedimentary rocks, the deep-sea sediment, the atmosphere, and the soil-water pool. Although primary rocks contain as much as 97.8% of the Earth's total (2.1×10^{17} tons or 1.9×10^{17} metric tons), their N_2 contributes little to the nitrogen cycle. Of the remaining nitrogen, 2% is in the atmosphere as N_2, and about 0.2% is in sedimentary rocks. The biosphere, consisting of the soil-water pool, contains only a small portion of the Earth's total nitrogen (2.6×10^{13} tons or 2.4×10^{13} metric tons), and even here the predominant species (2.4×10^{13} tons or 2.2×10^{13} metric tons) is N_2 dissolved in the sea. In spite of this, it is in this soil-water pool that the major reactions of the nitrogen cycle occur.

An outline showing the general interconversions of nitrogenous compounds in the soil-water pool is presented in **Fig. 1**. The reactions are much more complex than in the outline, and biological agents

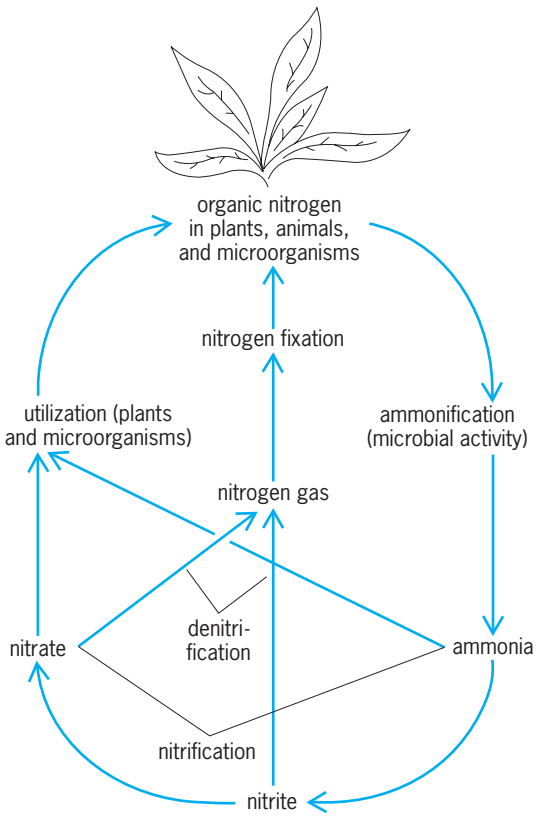

Fig. 1. Diagram of the nitrogen cycle.

have evolved intricate ways to manipulate these nitrogenous compounds for their own use. There are three primary reasons why organisms metabolize nitrogen compounds: (1) to use them as a nitrogen source, which means first converting them to NH_3, (2) to use certain nitrogen compounds as an energy source such as in the oxidation of NH_3 to NO_2^- and of NO_2^- to NO_3^-, and (3) to use certain nitrogen compounds (NO_3^-) as terminal electron acceptors under conditions where oxygen is either absent or in limited supply. The reactions and products involved in these three metabolically different pathways collectively make up the nitrogen cycle and are discussed below.

Nitrogen compounds as nutrients. The synthesis of organic nitrogen compounds from inorganic nitrogen and carbon compounds begins with NH_3 incorporation. One major reaction, catalyzed by glutamic acid dehydrogenase, involves 2-ketoglutarate, NADH (reduced nicotinamide adenine dinucleotide) and ammonia, and the product is glutamic acid. There are two ways in which organisms obtain ammonia. One is to use nitrogen already in a form easily metabolized to ammonia. Thus, nonviable plant, animal, and microbial residues in soil are enzymatically decomposed by a series of hydrolytic and other reactions to yield biosynthetic monomers such as amino acids and other small-molecular-weight nitrogenous compounds. These amino acids, purines, and pyrimidines are decomposed further to produce NH_3 which is then used by plants and bacteria for biosynthesis, or these biosynthetic monomers can be used directly by some microorganisms. The decomposition process is called ammonification. Not all organic nitrogen is ammonified easily, and resistant nitrogenous residues constitute humus, a complex component of great importance to soil structure and water-holding capacity.

The second way in which inorganic nitrogen is made available to biological agents is by nitrogen fixation (this term is maintained even though N_2 is now called dinitrogen), a process in which N_2 is reduced to NH_3. Since the vast majority of nitrogen is in the form of N_2, nitrogen fixation obviously is essential to life. The N_2-fixing process is confined to prokaryotes (certain photosynthetic and nonphotosynthetic bacteria). The major nitrogen fixers (called diazotrophs) are members of the genus *Rhizobium*, bacteria that are found in root nodules of leguminous plants (**Fig. 2**), and of the cyanobacteria (originally called blue-green algae). Even though rhizobia have recently been cultured so that they can fix nitrogen in the absence of the plant, the conditions needed for this fixation, such as a very low O_2 concentration supplied by oxyleg-hemoglobin and the necessary carbon and energy sources supplied by the plant, are ideal in the nodule and are not easily met in the laboratory or elsewhere in nature. There are many "free-living" diazotrophs, and even though they appear to contribute little to soil nitrogen, most knowledge of the biochemistry of nitrogen fixation comes from studies of three of them, *Clostridium pasteurianum*, *Azotobacter vinelandii*, and *Klebsiella pneumoniae*.

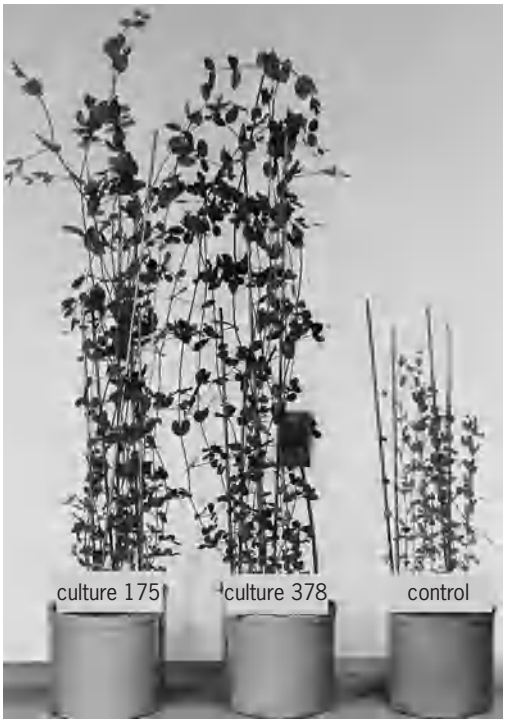

Fig. 2. Effect of inoculation with two strains (cultures 175 and 378) of the nitrogen-fixing *Rhizobium leguminosarum* on the development of peas.

The nitrogen-fixing system of all these organisms is made up of two protein components. One, called the Fe protein, is a dimer of molecular weight about 60,000 daltons, and it contains an Fe_4S_4* center (similar to that of ferredoxin) that is involved in accepting and transferring the electrons needed for N_2 reduction. It is the Fe protein that initially binds the magnesium adenosine triphosphate (MgATP) needed for N_2 reduction. The second component, called the MoFe protein, is a tetramer of about 220,000 daltons and contains four Fe_4S_4* centers and two $Fe_8Mo_1S_6$ centers, with the latter centers probably being the sites where N_2 is reduced to NH_3. For N_2 reduction the MoFe protein accepts electrons from the Fe protein, and to facilitate the transfer, ATP is hydrolyzed to adenosine diphosphate (ADP) and inorganic phosphate. The N_2 fixation process consumes as much as a third of the cell's energy supply, so for obvious reasons the N_2-fixing system is not synthesized if a usable nitrogen compound other than N_2 is available, that is, nitrogenase synthesis is repressed in the presence of NH_3.

Nitrogen compounds as energy source. Many microorganisms can use organic nitrogen compounds as energy sources, but in most cases the nitrogen of the compound is first removed and excreted as NH_3, and then the reduced carbon compound remaining is catabolized to yield both energy and organic carbon intermediates. A relatively few microorganisms (including some fungi) are able to aerobically convert the excreted NH_3 to NO_2^-, and NO_2^- to NO_3^-, and couple these oxidations to the production of ATP and the creation of a membrane potential needed for biosynthetic reactions. Thus, NH_3 added to aerobic soils is rapidly converted by such organisms to NO_3^-. Two chemoautotrophs (organisms that grow using an inorganic compound as an energy source and CO_2 as a carbon source) are primarily responsible for NO_3^- production. First *Nitrosomonas* converts the NH_3 to NO_2^-, and then the relatively toxic NO_2^- is rapidly oxidized to NO_3^- by *Nitrobacter*. Plants and microorganisms readily use this NO_3^- as a nitrogen source by first reducing it to NH_3. The overall process of oxidation of NH_3 to NO_3^- is called nitrification. The process whereby NO_3^- is reduced to NH_3 is called nitrate assimilation.

Nitrogen compounds as electron acceptors. When NO_3^- accumulates in soils in which metabolizable carbon compounds are available and when such soils become anaerobic because growth of aerobic organisms exhausts the O_2, certain organisms such as *Pseudomonas*, *Micrococcus*, *Achromobacter*, and *Bacillus* use the NO_3^- either as a normal electron acceptor or as an electron acceptor in place of O_2. The electron acceptor is needed by such cells to allow electrons from cellular oxidations to flow through the array of electron carries in the cell membrane. This electron flow is needed to facilitate proton (H^+) transfer across the membrane, which in turn creates a membrane potential and a pH gradient. The potential energy of the pH gradient is used in conjunction with the cells adenosine triphosphatase (ATPase) to allow ATP to be synthesized. The NO_3^- is the terminal electron acceptor in the electron flow and becomes reduced to NO_2^-, and the NO_2^- in turn is further reduced to N_2 (some N_2O may also be produced). The N_2 (and some N_2O) is released into the atmosphere. This process, called denitrification or nitrate respiration, is responsible for ridding many bodies of water and soil of excess fixed nitrogen that could lead to pollution by overproduction and decay of algae and other bacteria.

N_2 cycle in the oceans. Most information on the N_2 cycle comes from studies of the soil-water system. Much less is known about the nitrogen cycle in seas, even though it is estimated that as much as 20% of the N_2 fixed on Earth occurs in the ocean. This N_2 seems to be primarily fixed by cyanobacteria, although in localized areas some contribution by other photosynthetic and nonphotosynthetic bacteria also occurs. One calculation estimates that a single bloom of the cyanobacterium *Trichodesmium* could fix as much as 1100 tons (100 metric tons) of N_2 per day. The other processes associated with the nitrogen cycle, ammonification, nitrification, and denitrification, although obviously present in oceans, have been studied even less. Heterotrophic organisms were not shown to be responsible for nitrification in seas, and therefore chemoautotrophic marine nitrifiers, such as *Nitrospira* and *Nitrococcus*, appear to be the major nitrifiers in oceans. The denitrification step of the ocean's nitrogen cycle may account for a third of the Earth's total denitrification, but there is relatively little information on this. One suspects that most of this denitrification takes place in the anaerobic conditions of marine mud.

Summary. The major microbial reactions discussed above are the three major contributors to the N_2 cycle. First N_2 is fixed (reduced) to ammonia, and the ammonia is used by N_2-fixing microorganisms and by plants harboring N_2-fixing microorganisms. Then these plants and bacteria, and animals and plants living off these plants and bacteria, die and lyse, and the nitrogen of their nitrogenous compounds is converted to ammonia by ammonification. The released ammonia is converted rapidly by nitrifying bacteria to NO_3^-, or it is used directly by microorganisms or plants. The NO_3^- produced by nitrification is used by plants and by bacteria as a nitrogen source; or if anaerobic conditions are created, the NO_3^- is denitrified (reduced) to N_2 and N_2O. This completes the nitrogen cycle. Leonard E. Mortenson

Bibliography. N. S. Rao, *Biological Nitrogen Fixation: Recent Developments*, 1988; J. I. Sprent, *The Ecology of the Nitrogen Cycle*, 1987; G. Stancey, R. H. Burris, and H. J. Evans (eds.), *Biological Nitrogen Fixation*, 1991.

Nitrogen excretion

In the quest for sufficient food energy to meet caloric requirements, animals ingest more nitrogen, largely as amino acids, than they require. Accordingly, the

excess nitrogen ingested must be excreted in some form. Through the action of a series of related enzymes called transaminases, virtually all metabolic nitrogen can be transferred to α-ketoglutaric acid to form glutamic acid, as shown in reaction (1). Under

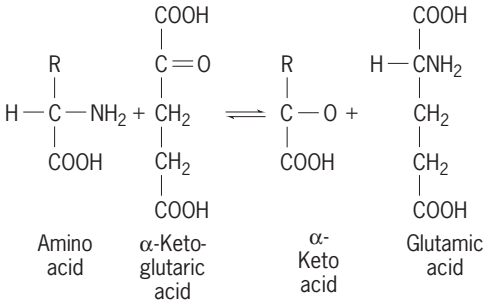

(1)

the influence of glutamic dehydrogenase, glutamic acid may be oxidized by the coenzyme diphosphopyridine nucleotide (DPN) with the reformation of α-ketoglutaric acid plus ammonia, as shown in reaction (2).

$$\text{Gultamic acid} + \text{DPN}^+ \rightleftharpoons \alpha\text{-ketoglutamic acid} + \text{DPNH} + \text{NH}_3^+ \quad (2)$$

Aquatic animals in general simply excrete the ammonia as such, typically by diffusion across the gills or body surface. Other animals, notably amphibians and mammals, first convert ammonia into urea.

To accomplish the transformation, advantage is taken of the enzymically catalyzed metabolic sequence by which the amino acid arginine is synthesized from ornithine, a sequence common to almost all living forms. It is adapted for urea synthesis by mammals, reptiles, and other forms by the additional presence in liver of the enzyme arginase, which catalyzes the hydrolysis of arginine to urea plus ornithine, which is then available for recycling.

The initial step is catalysis of the formation of carbamyl phosphate from ammonia, CO_2, and adenosine triphosphate (ATP) by carbamyl phosphate synthetase. The mechanism of this reaction is not understood, but it is known to require the presence of an *N*-acylglutamate and an additional molecule of ATP. The reaction is shown in (3). The carbamyl phosphate

$$\text{NH}_3 + \text{CO}_2 + 2\text{ATP} \longrightarrow \text{H}_2\text{N—CO—OPO}_3\text{H}_2 + \text{H}_2\text{OPO}_{3^-} + 2\text{ADP} \quad (3)$$

is then caused to react with ornithine under the influence of ornithine transcarbamylase with the formation of citrulline plus orthophosphate, a reaction shown in (4). The second atom of nitrogen necessary for urea synthesis is then introduced in a complex reaction whereby aspartic acid and citrulline condense to form argininosuccinic acid, the energy being derived from another molecule of adenosine triphosphate with the formation of adenosine monophosphate and pyrophosphate, as is shown in reaction (5). It is noteworthy that, like ammonia, the nitrogen of aspartic acid is derived from glutamic acid

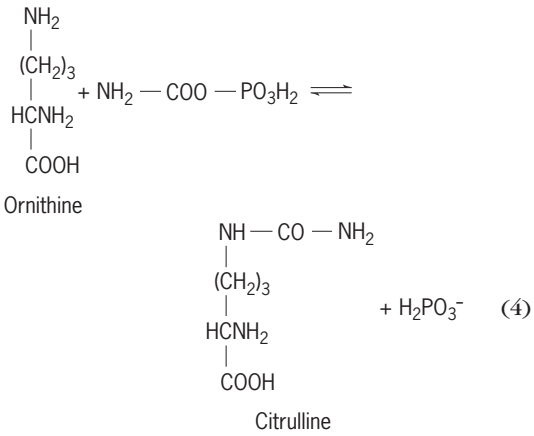

(4)

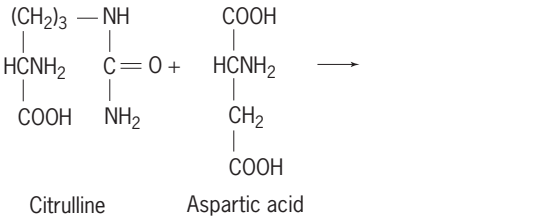

(5)

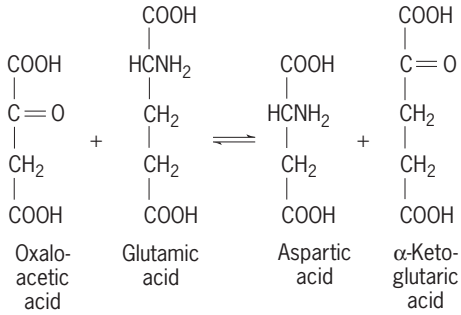

(6)

through the operation of a specific aspartic-glutamic transaminase, as shown in reaction (6).

The product, argininosuccinic acid, is cleaved by an appropriate enzyme to form one mole each of arginine and fumaric acid, as shown in reaction (7). With the hydrolysis of arginine to ornithine plus urea the cycle is completed and the ornithine is available for the next round, as shown in reaction (8).

The free energy charge ΔG° that is required for the overall reaction CO_2 + 2 ammonia $\rightarrow$ urea is $-14{,}000$ cal/mole. As in many other biological systems, this energy is provided by utilization of adenosine triphosphate. The cyclic nature of the system is shown in the **illustration**. A second cycle is also operative wherein the 4-carbon dicarboxylic acid, fumaric acid, released in reaction (7) is reconverted to oxaloacetic so that it can be realized for aspartic acid formation and thus for another turn of the urea cycle.

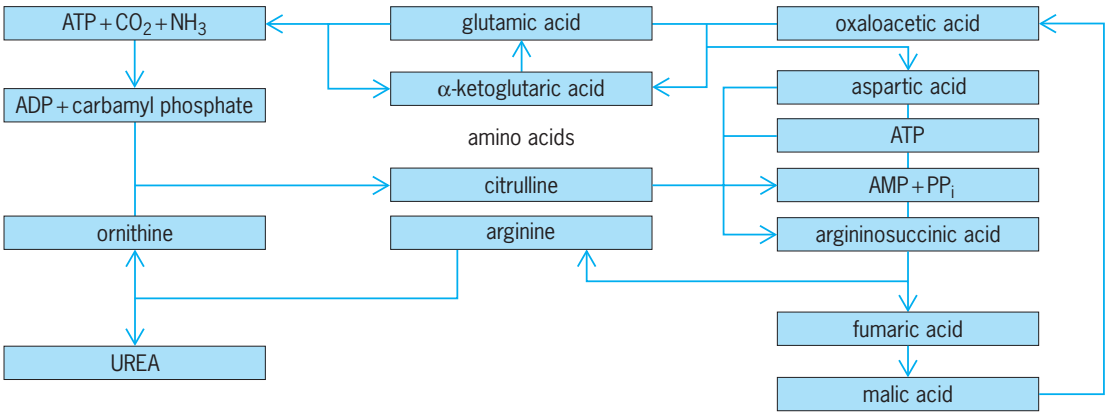

Regenerative cycles in urea biosynthesis beginning with formation of carbamyl phosphate.

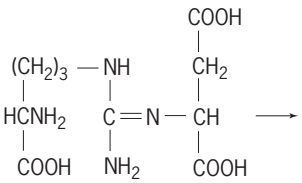

Argininosuccinic acid

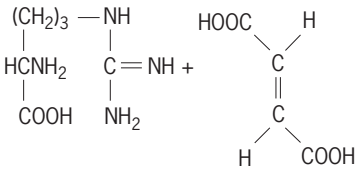

(7)

Arginine Fumaric acid

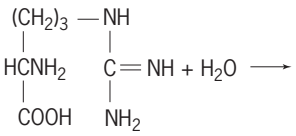

Arginine

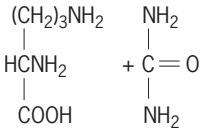

(8)

Ornithine Urea

The relatively small amounts of urea excreted by aquatic animals are largely derived from the breakdown of purines, derived from the nucleic acids. Terrestrial animals, notably insects, arachnids, reptiles, and birds, convert the ammonia, in a synthesis which requires a greater input of energy per molecule of ammonia than does the synthesis of urea, into the purine xanthine, which is then oxidized to form uric acid, the form in which the ammonia is ultimately eliminated. This mechanism is an adaptation to embryonic development in an egg with closed shell, in which the accumulation of ammonia would be toxic and the formation of urea would cause osmic difficulties. In addition, for adults in a dry environment uric acid as a nitrogenous end product has the advantage that it is relatively insoluble in water, and may be eliminated in solid form, without the water which is typically the main component of urine of mammals. *See* EXCRETION; PROTEIN METABOLISM; URINARY SYSTEM; URINE.

Philip Handler; Bradley T. Scheer

Nitrogen fixation

The chemical or biological conversion of atmospheric nitrogen (N_2) into compounds which can be used by plants, and thus become available to animals and humans. In the 1990s, chemical and biological processes together contributed about 260 million tons (230 million metric tons) of fixed nitrogen per year globally. Industrial production of nitrogen fertilizer accounted for about 85 million tons (80 million metric tons) of nitrogen per year, while spontaneous chemical processes, such as lightning, ultraviolet irradiation, and combustion, leading to the synthesis of nitrogen oxides from O_2 and N_2, may have accounted for 44 million tons (40 million metric tons) per year. The remainder, roughly half of the global input of newly fixed nitrogen, arose from biological processes. World agriculture, which is very dependent on nitrogen fixation, is increasingly reliant on chemical nitrogen sources. *See* NITROGEN.

Chemical fixation. Three chemical processes for fixing atmospheric nitrogen have been developed. All require considerable thermal or electrical energy and yield different products.

In arc processes, such as the Birkland-Eyde process, air is passed through an electric arc and about 1% nitric oxide is formed, which can be chemically converted to nitrates. Inexpensive electricity is required for the industrial use of arc processes, which are now rarely used. In the cyanamide process, which is now obsolete, heating calcium carbide in nitrogen generates calcium cyanamide, which when moistened hydrolyzes to urea and ammonia. In the widely used Haber process, hydrogen (generated by heating natural gas) is mixed with nitrogen (from air), and burned to yield a nitrogen-hydrogen mixture. The nitrogen-hydrogen mixture is compressed (10–80 megapascals) and heated (200–700°C or 390–1300°F) in the presence of a metal oxide catalyst to give ammonia. The Haber process is the major source

of ammonia used for fertilizer. *See* HIGH-PRESSURE PROCESSES.

Ionization and chemonuclear processes exist for obtaining nitrogen oxides from air but are not used industrially. Laboratory processes also have been developed for fixing nitrogen as metal nitrides or transition-metal complexes, and a procedure has been developed for reducing the nitrogen bound in certain transition-metal complexes to ammonia. Such processes may have industrial applications. *See* AMMONIA; CYANAMIDE; ELECTROCHEMICAL PROCESS; FERTILIZER; NITROGEN COMPLEXES.

Biological fixation. Only prokaryotes—bacteria, archaea, and cyanobacteria (earlier called blue-green algae)—fix nitrogen. Nitrogen-fixing microbes, called diazotrophs, fall into two main groups, free-living and symbiotic. *See* ARCHAEA; BACTERIA; CYANOBACTERIA; PROKARYOTAE.

The free-living diazotrophs are subclassified. Aerobic diazotrophs, of which there are over 50 genera, including *Azotobacter*, methane-oxidizing bacteria, and cyanobacteria, require oxygen for growth and fix nitrogen when oxygen is present. *Azotobacter*, some related bacteria, and some cyanobacteria fix nitrogen in ordinary air, but most members of this group fix nitrogen only when the oxygen concentration is low. Free-living diazotrophs, which fix nitrogen only when oxygen is absent or vanishingly low, are widespread. The genera *Bacillus* and *Klebsiella* include many strains of this type, and representatives of symbiotic diazotrophs behave in this way as well. They grow perfectly well in air if prefixed nitrogen, such as an ammonium salt, is present, but they will not fix nitrogen under these conditions. Some *Clostridium* strains, some sulfate-reducing bacteria, and some methane-forming archaea are free-living diazotrophic anaerobes, and do not grow in the presence of air, whether fixing nitrogen or not. There are also free-living diazotrophic phototrophs, including many cyanobacteria, which fix nitrogen in air, but other photosynthetic bacteria are generally unable to fix nitrogen, except when growing anaerobically. *See* ALGAE; BACTERIAL PHYSIOLOGY AND METABOLISM.

The best-known symbiotic bacteria belong to the genus *Rhizobium*. Species of *Rhizobium*, or related genera, such as *Bradyrhizobium* and *Sinorhizobium*, colonize the roots of leguminous plants and stimulate the formation of nodules within which they fix nitrogen microaerobically. Both plants and bacteria show specificity; for example, certain types of plants require special strains of rhizobia. Some types of rhizobium, such as *Bradyrhizobium*, can fix nitrogen in the absence of plant tissue, but require low oxygen, though most rhizobia fix nitrogen only within the nodules. Effective root nodules are pink due to a protein, called leghemoglobin, related to the hemoglobin of blood, which transports a steady supply of oxygen to the rhizobia. Rhizobial symbioses are of great agricultural importance in legume crops, such as soybean, pulse, and alfalfa. Of comparable ecological importance are *Frankia* species, the diazotrophic symbionts of *Alnus* (alder), *Casuarina* (Australian pine), or *Ceanothus*, which are also found in root nodules. Symbiotic cyanobacteria form diazotrophic associations with plants ranging from fungi (as lichens) to ferns (for example, *Azolla*), cycads, and angiosperms (for example, *Gunnera*). Symbiotic associations between several genera of aerobic diazotrophs (such as *Azospirillum*, *Azoarcus*, and a species of *Azotobacter*) and both grasses and cereals have been reported. More casual associations between free-living diazotrophic bacteria and the roots of maize, rice, temperate weeds, and water plants also exist. All of these have limited agricultural potential at present. Diazotrophic associations involving leaf nodules or the mycorrhizae of roots are probably illusory. Associations with ruminant animals, termites, and wood-boring mollusks are real but casual and nutritionally trivial. *See* SOIL MICROBIOLOGY.

Nitrogenase. The enzymes responsible for nitrogen fixation are called nitrogenases. The most common nitrogenase consists of two proteins, one large (about 220,000 mol wt) containing molybdenum, iron, and inorganic sulfur (the MoFe-protein or dinitrogenase), the other smaller (about 60,000 mol wt) containing iron and inorganic sulfur (the Fe-protein or dinitrogenase reductase). The chemical compositions of both proteins and the natures of their metal-containing sites are known, and the three-dimensional structure of the MoFe-protein has been established by x-ray analysis. The molybdenum atom resides in a unique iron-molybdenum-sulfur cluster, called FeMoco, which can be chemically extracted.

Both nitrogenase proteins are irreversibly destroyed by exposure to oxygen. Therefore, an oxygen-deficient environment is essential for activity. In addition, biological energy in the form of adenosine triphosphate (ATP) is consumed during activity, for which magnesium ions and a reducing agent are also required. The biological reductant is a ferredoxin or a flavodoxin. Nitrogenase reduces one molecule of N_2 to two of ammonia (NH_3), a reaction which is accompanied by the conversion of 16 molecules of ATP to adenosine diphosphate (ADP) and the release of one molecule of H_2 as a by-product. If N_2 is absent (for example, in argon), the enzyme forms hydrogen exclusively. Several small molecules that resemble N_2 in chemical structure will interact with the enzyme. Carbon monoxide blocks nitrogen fixation completely but leaves H_2 formation unimpaired. Other structural analogs of N_2, such as cyanide, azide, acetylene, or nitrous oxide, can be reduced in its place. The reduction of acetylene to ethylene provides a rapid and highly efficient test for biological nitrogen-fixing systems.

In active nitrogenase, the two proteins form a short-lived complex and their modes of action are largely understood. In the **illustration**, reduced ferredoxin is depicted as reducing the Fe-protein which then, after reacting with a magnesium derivative of ATP, reduces the MoFe-protein that had bound N_2, displacing a molecule of H_2O. The N_2 is reduced to NH_3; and H_2, magnesium, and ADP are released. Finally, the MoFe-protein is reoxidized and the

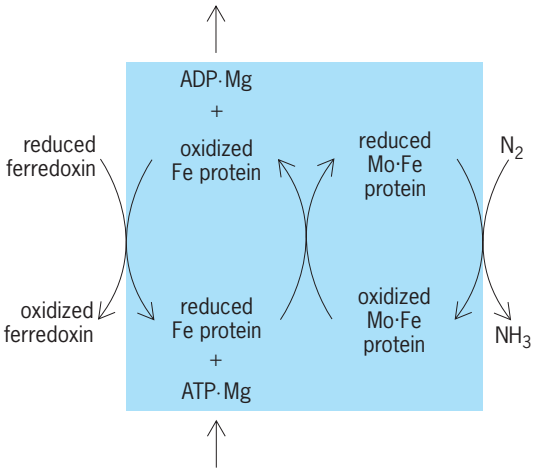

A mechanism for the functioning of the two components of nitrogenase. The box signifies that a complex of Fe and Mo-Fe proteins forms the active enzyme.

complex falls apart. The rate of this last step determines the turnover time for N_2, which is slow at about 0.9 s^{-1}. Because of its slow action, up to 20% of the protein in an active diazotroph can be nitrogenase.

Two less common types of nitrogenase are also known. They occur in certain azotobacters and photosynthetic diazotrophs. Like the Mo-nitrogenase, they consist of two proteins, consume ATP, evolve H_2, and are destroyed by oxygen. One type contains vanadium (V) and the other type, iron (Fe), in place of Mo, and both metal atoms are in clusters comparable to FeMoco. Known as V-nitrogenase and Fe-nitrogenase, they appear to be less efficient than Mo-nitrogenase and their biological significance is obscure. *Azotobacter vinelandii* possesses all three nitrogenases. *See* VANADIUM.

Physiology. The irreversible destruction of nitrogenase by air means that all aerobic diazotrophs have developed means of restricting access of oxygen to the active enzyme. The legume root nodule can be viewed as a compartment in which leghemoglobin supplies rhizobial cells with oxygen sufficient for respiration but at concentrations that are harmless to nitrogenase. Some cyanobacteria have comparable compartments or specialized cells, called heterocysts, in which nitrogenase is protected from the oxygen evolved during photosynthesis by an oxygen-impermeable wall. In free-living aerobes, respiration excludes oxygen from the enzyme; and *Azotobacter*, one of the most efficient aerobic diazotrophs, possesses a special iron-sulfur protein which protects nitrogenase in cells subject to oxygen stress, though the enzyme is then nonfunctional until the oxygen stress ceases.

The release of H_2 as a by-product of nitrogen fixation represents a waste of biological energy (as ATP). Efficient aerobic diazotrophs, such as *Azotobacter* and some strains of *Rhizobium*, possess the enzyme hydrogenase, which enables them to use the H_2 as a reductant and source of ATP. There is evidence that hydrogenase-containing rhizobia produce the best yields of certain legume crops. *See* NITROGEN CYCLE.

Regulation and genetics. Nitrogenase is an expensive enzyme in terms of its ATP consumption, and most diazotrophs possess physiological mechanisms that enable them to switch off the enzyme should a supply of fixed nitrogen appear. It is also expensive in the sense that the cell is obliged to synthesize large amounts of enzyme protein, and often the apparatus of an oxygen-exclusion mechanism as well. Synthesis of the enzyme is tightly regulated at the genetic level, and it is not made when fixed nitrogen (as NH_3) is available or if O_2 concentrations are unsuitable.

The genes specifying nitrogenase, together with the various ancillary proteins involved in its synthesis and function, called *nif* genes, are regulated at the transcriptional level by ammonia, operating through a genetic system (the *ntr* genes), which also regulates other bacterial nitrogen metabolism functions, such as the use of certain amino acids. Oxygen also regulates *nif* expression. In 1972, *nif* genes were taken from *Klebsiella* and inserted into *Escherichia coli*, which then gained the ability to fix nitrogen. Since then, plasmids (extrachromosomal genetic elements) have been constructed to carry *nif* genes, permitting their transfer to a variety of microorganisms. As a result, entirely new diazotrophs have been created such as strains of *Salmonella*, *Proteus*, and *Serratia*. But not all bacterial recipients of the *nif* genes became diazotrophic. The *nif* genes of *Klebsiella* have been studied exhaustively, and their complete DNA sequences are known. They are present as a cluster of 20 contiguous genes. Expression of the whole cluster is regulated by *ntr* products acting upon genes within it, called *nifA* and *nifL*, and subclusters (operons) within *nif* are regulated by the products of *nifA* and *nifL*. Oxygen bypasses *ntr* and regulates *nif* expression by an indirect action upon the *nifL* gene.

Recombinant DNA clones have been prepared that carry the whole *nif* cluster from *Klebsiella*, the several operons, the separate genes, and their promoter regions. And gene fusions have been prepared to study physiological questions. Clones of *Klebsiella nif* genes have also been used to open up the molecular genetics of *Rhizobium*, *Azotobacter*, *Anabena* (a cyanobacterium), and many other diazotrophs. In several diazotrophs, the *nif* genes occur naturally on plasmids; however, in most diazotrophs they are not in a contiguous cluster but are dispersed about the genome in separate operons. V-nitrogenases and Fe-nitrogenases are coded by distinct but comparable groups of genes (called *vnf* and *anf* respectively), and have been cloned and sequenced. *See* BACTERIAL GENETICS; GENETIC ENGINEERING.

Genetics of diazotrophic symbioses. In legume symbiosis, and probably in others, both the host plant and the bacteria contribute genetic information. The rhizobia carry the *nif* genes, plus genes called *fix*, which play a part in making functional nitrogenase. In the genus *Rhizobium* these genes are on a plasmid, but in the related *Bradyrhizobium* they are chromosomal. The bacteria also have *sym* genes, which specify the appropriate species of host

plant, and *nod* genes, which induce nodulation in the plant. The plants possess other *nod* genes and contribute products called nodulins. One nodulin is leghemoglobin, and another is a specific attractant that enables the free rhizobia to recognize which plant to colonize. In all, some 50 genes are involved in establishing legume nodule symbiosis. It is likely that analogous genes function in *Frankia* and other diazotrophic symbioses.

Economic considerations. Except in a few areas where extreme climatic conditions are a limiting factor, input of fixed nitrogen determines the productivity of world agriculture and forestry. That input was substantially biological until the early 1900s, during which the use of nitrogen fertilizer produced chemically by the Haber process escalated. By 2000, between 30 and 40% of the world's population depended on chemical nitrogen fertilizer for its food supplies. The global population is expected to reach 8–9 billion by 2100. Substantial increases in nitrogen input, of both chemical and biological origin, into global food production are therefore inescapable. Increased fertilizer use may generate environmental, economic, and sociopolitical problems, ranging from local eutrophication and pollution to increased dependence of world agriculture on sophisticated technology, with consequent economic and social disruption in developing countries. Proposals to ameliorate these problems include (1) developing low-tech "cottage" processes for generating nitrogen fertilizer locally, based perhaps on solar-powered chemistry or immobilized enzyme technology; (2) extending the use of established symbiotic plant associations in food production and forestry, particularly their use as "green manure"; (3) generating, by conventional breeding or genetic manipulation, new or improved diazotrophic symbioses involving wheat and rice, which are the world's major food crops; and (4) transferring, by genetic manipulation, bacterial *nif* genes into the plant genome and rendering them functional therein. *See* AGRICULTURE.

John Postgate

Bibliography. I. R. Kennedy and E. C. Cocking (eds.), *Global Nitrogen Fixation: The Global Challenge and Future Needs*, University of Sydney SUNfix Press, 1997; J. R. Postgate, *Nitrogen Fixation*, Cambridge University Press, London, 2000; G. Stacey, R. H. Burris, and H. J. Evans (eds.), *Biological Nitrogen Fixation*, Routledge, New York, 1992.

Nitrogen-fixing trees

The ability to fix atmospheric nitrogen into a form that can be used for plant growth is confined to bacteria and cyanobacteria. Plants fix nitrogen only by virtue of associations with these simple organisms. The best-known associations are the symbioses of *Rhizobium* bacteria with agricultural legumes, such as clovers, peas, and beans. Rhizobia stimulate the formation of root nodules, which provide a specialized environment within which high rates of nitrogen fixation can occur. Nitrogen fixation not only supports plant growth independent of mineral nitrogen in the soil but also can improve soil nitrogen status as plant residues, notably leaves and fine roots, decay and are mineralized. Accurate estimates of amounts of nitrogen fixed are difficult to obtain, but a well-nodulated, young, densely planted stand of nitrogen-fixing trees may fix 80–200 kg of nitrogen per hectare per year, similar to rates reported for some leguminous crops. *See* BACTERIA; CYANOBACTERIA.

Nitrogen-fixing tree species occur in all three subfamilies of the Leguminosae (Fabaceae). Most of the mimosoid group (such as *Acacia*, *Albizia*, *Leucaena*) and virtually all of the papilionoid group (including *Dalbergia*, *Gliricidia*, *Robinia*) fix nitrogen, but nodulation and nitrogen fixation in the caesalpinioid group (such as *Chamaecrista*) is very restricted. However, in some genera, for example the acacias, not all the species form nitrogen-fixing symbioses. Most legume trees are found in the tropics and subtropics, but some grow and are of importance in temperate climates, for example the horticultural laburnums or black locust (*Robinia pseudoacacia)*, which is used extensively in rehabilitation of derelict land. *See* LEGUME.

Many species of trees and shrubs other than legumes form nitrogen-fixing root nodules not with rhizobia but with a filamentous bacterium, *Frankia*, or with nitrogen-fixing cyanobacteria, such as *Nostoc*. Symbiosis with cyanobacteria produces a large, coralloid mass of nodules on surface roots of some cycads. Evolutionarily, these plants are ancient members of the gymnosperms; they resemble palm trees when mature, with a stout trunk and a crown of fronds. They are restricted to Central and South America, South Africa, and Australasia, where they are a source of fixed nitrogen for natural ecosystems. Following fires, cycads grow and fix nitrogen in eucalyptus forests. After a few years, they are shaded out by new growth and provide a valuable input of fixed nitrogen to the forest ecosystem as they decay.

Frankia belongs to the bacterial family Actinomycetales, which includes the antibiotic-producing soil actinomycetes *Streptomyces*. *Frankia* forms nitrogen-fixing nodules on the roots of trees and shrubs in eight families other than legumes, most being native to temperate regions (see **illus.**). Species nodulated by *Frankia* are known as actinorhizal plants. They are very important ecologically for their ability to colonize and to improve the nitrogen status of denuded and degraded soils. Some genera, for example *Ceanothus* and *Elaeagnus*, are of horticultural importance, while *Hippophae rhamnoides* (sea buckthorn) berries are a valuable crop in Russia and China. In temperate regions, the most widely utilized actinorhizal tree species belong to the genus *Alnus* (the alders) of the Betulaceae, the family to which the non-nitrogen-fixing birches belong. Different alder species occupy a wide range of natural habitats, from riversides to relatively dry mountain valleys. They were used widely in earlier times to produce high-quality charcoal, dyes, and decay-resistant timber. Their main use now is in land reclamation.

A perennial nodule, several years old, formed on the roots of common or black alder (*Alnus glutinosa*) by symbiosis with the nitrogen-fixing bacterium *Frankia*.

They may be grown also in single-species stands or in mixed stands with non-nitrogen-fixing tree species to improve productivity on nitrogen-deficient sites. There is also considerable interest in utilizing alders as a component of fast-growing, sustainable energy plantations to provide biomass for fuel or paper pulp production.

Only one family of actinorhizal plants, which contains the casuarinas, the Australian sheoaks, is native to the tropics. These trees are a common feature of hot countries, and their needlelike branchlets and conelike fruits frequently lead to confusion with conifers. The family contains species tolerant of a wide range of soil conditions in environments that may be wet, semiarid, or dry. Species such as *Casuarina equisetifolia* or *C. glauca*, which are tolerant of saline conditions and poor soils, are used extensively as wind breaks and in stabilizing coastal and desert sand dunes. *Casuarina equisetifolia* is fast-growing, and the dense wood has a calorific value higher than most other nitrogen-fixing trees. It is grown widely in 5–7-year rotations by small farmers in India and China for fuel wood or charcoal, and it is also valuable for producing small construction timbers such as piles and posts, paper pulp, and fiber boards.

Most species of nitrogen-fixing trees belong to the Leguminosae and have diverse uses. Timber and pulpwood species include *Dalbergia* (rosewood) and *Albizia* (ipil ipil), which is an important component of sustainable forestry in the tropics. Many species are used for fuel wood, and trees able to grow in deforested nitrogen-deficient soils of arid regions are very important in this respect socially, for example some *Acacia* and *Prosopis* (mesquite) species. Other important uses are as shade, nurse, or ornamental trees and as foliage for animal fodder. Some trees also produce edible fruits or gums and tannins and are useful as medicines. Like the actinorhizal trees, nitrogen-fixing legume trees are used widely for land reclamation and stabilization.

Many useful features may be found in a single species. For example, *Leucaena leucocephala* is grown widely in the tropics and can achieve a height of 18 m (60 ft) in 5 years. The wood is excellent for fuel and charcoal and can be used for posts, lumber, and furniture. Dense plantations produce leaf litter containing up to 300 kg of nitrogen per hectare and valuable as fodder or green manure. Tender vegetative shoots can be eaten, as can seeds if treated carefully to reduce toxicity. The multifaceted properties of nitrogen-fixing trees ensure them a favored role in agroforestry systems, particularly on impoverished lands, where their minimal requirement for fertilizer application and the high nitrogen content of residues are of great value for intercropping practices. *See* NITROGEN FIXATION; PLANT METABOLISM.

C. T. Wheeler

Bibliography. M. G. R. Cannell, D. C. Malcolm, and P. A. Robertson (eds.), *The Ecology of Mixed Species Stands of Trees*, Blackwell, Oxford, 1992; K. E. Giller and K. J. Wilson, *Nitrogen Fixation in Tropical Cropping Systems*, C. A. B. International, Wallingford, 1991; J. C. Gordon and C. T. Wheeler (eds.), *Biological Nitrogen Fixation in Forest Ecosystems: Foundations and Applications*, Martinus Nijhoff/W. Junk, The Hague, 1983; C. R. Schwintzer and J. D. Tjepkema (eds.), *The Biology of Frankia and Actinorhizal Plants*, Academic Press, San Diego, 1990.

Nitrogen oxides

Chemical compounds of nitrogen and oxygen. Nitrogen and oxygen do not combine when mixed directly (as in air), but they do combine during chemical reactions of compounds containing them. A number of nitrogen oxides can be isolated which differ from one another in the numbers of nitrogen and oxygen atoms present in each molecule.

Table 1 gives data for the five nitrogen oxides which are well established. The structures of these molecules and one laboratory method for the preparation of each oxide are given in **Table 2**. These structures show only the geometry of the molecules. In most cases the N and O atoms are united by complex (double or triple) bonds.

The existence of three higher oxides has been postulated. They are nitrogen trioxide, NO_3, from reaction of ozone with dinitrogen tetroxide or pentoxide; dinitrogen hexoxide, N_2O_6, from reaction of fluorine with nitric acid; and an oxide, NO_4, as an intermediate in the ^{18}O isotope exchange between dinitrogen pentoxide and oxygen gases. The identity and properties of these three oxides are not fully established.

Nitrous oxide and nitric oxide. When inhaled, nitrous oxide has anesthetic effects; in small amounts it produces mild hysteria and hence is sometimes called laughing gas. It is colorless, is the least reactive of the oxides, and dissolves in water without chemical reaction. Decomposition into nitrogen and oxygen occurs at an appreciable rate above 560°C (1040°F)

The equilibrium in reaction (1) lies entirely to the

$$N_2 + O_2 \rightleftharpoons 2NO \qquad (1)$$

left at low temperatures. Some nitric oxide is formed in an electric arc, as in the technical production of nitric acid.

TABLE 1. Oxides of nitrogen and their properties

Name	Stoichiometric formula	Melting point, °C (°F)	Boiling point, °C (°F)
Nitrous oxide (dinitrogen monoxide)	N_2O	−90.8 (−131)	−88.5 (−127.3)
Nitric oxide (nitrogen monoxide)	NO	−163.6 (−262.5)	−151.7 (241.0)
Dinitrogen trioxide	N_2O_3	−103 (−155)	3.5 (38.3)
Dinitrogen tetroxide (⇌ nitrogen dioxide)	N_2O_4 (⇌ NO_2)	−11.2 (11.8)	21.2 (70.2)
Dinitrogen pentoxide	N_2O_5	41 (106)	

With oxygen or air, nitric oxide is rapidly converted to nitrogen dioxide. Nitric oxide is colorless and is soluble in water without reaction. It is one of the few "odd" molecules which contain an odd number of electrons. Other such molecules (for example, nitrogen dioxide) readily form double molecules, but nitric oxide is exceptional. The gas is monomeric, although dimerization occurs in the liquid, and solid nitric oxide (which is blue) is almost entirely in the form of N_2O_2 molecules. As an odd molecule, it has the ability to lose or gain one electron, thus giving the electrically charged ions NO^+ and NO^-. The important nitrosyl compounds contain these ions.

Trioxide. Dinitrogen trioxide exists pure only in the solid state. The dissociation in reaction (2) occurs

$$N_2O_3 \rightleftharpoons NO + NO_2 \qquad (2)$$

partially in the blue liquid and almost entirely in the vapor state at room temperature. It is the anhydride of nitrous acid; when the oxide (or an equimolecular mixture of NO and NO_2 gases) is dissolved in an alkaline solution, nitrite ion is produced.

Dioxide and tetroxide. The position of the equilibrium between nitrogen dioxide and dinitrogen tetroxide, reaction (3), depends upon temperature

$$N_2O_4 \rightleftharpoons 2NO_2 \qquad (3)$$

and physical state. The dioxide is a red-brown poisonous gas; the tetroxide is colorless. The colorless solid is entirely in the tetroxide state. In the liquid and gasesous states the tetroxide always contains some dioxide. Thus the liquid tetroxide is brown, although it contains less than 0.1% nitrogen dioxide. The color of the gas becomes more intense with rising temperature; at 100°C (212°F) the tetroxide is 90% dissociated into dioxide. At temperatures above 600°C (1100°F) further decomposition of nitrogen dioxide into nitric oxide occurs, as shown in reaction (4). Dinitrogen tetroxide reacts readily with water to

$$NO_2 \rightleftharpoons NO + {}^1\!/_2O_2 \qquad (4)$$

give an equimolecular mixture of nitrous and nitric acids. As temperature is raised, the nitrous acid decomposes to nitric acid and nitric oxide. These reactions are important in the technical production of nitric acid by catalytic oxidation of ammonia. Dinitrogen tetroxide is an oxidizing agent comparable in strength to bromine, and is employed as such in the lead-chamber process for sulfuric acid. In organic chemistry the tetroxide finds use as a special oxidizing agent (for example, in the production of sulfoxides and phosphine oxides) and as a nitrating agent.

The tetroxide forms molecular addition compounds with many simple organic solvents, for example, esters, ethers, ketones, and nitriles.

Liquid dinitrogen tetroxide, alone or mixed with organic solvents, undergoes self-ionization, as in reaction (5). This is to be compared with the aqueous

$$N_2O_4 \rightleftharpoons NO^+ + NO_3^- \qquad (5)$$

system shown in reaction (6). For example, liquid

$$H_2O \rightleftharpoons H^+ + OH^- \qquad (6)$$

dinitrogen tetroxide attacks some metals (alkali and alkaline-earth metals, zinc, cadmium, and mercury) to produce metal nitrate and evolve nitric oxide. A scheme of reactions has been developed using the liquid tetroxide as a reaction medium, with nitrosyl compounds as acids and nitrates as bases. This medium is therefore valuable for the preparation

TABLE 2. Oxide of nitrogen

Formula	Structure	Preparation
N_2O	N—N—O	Heat ammonium nitrate
NO	N—O	Reduce nitric acid with copper
N_2O_3	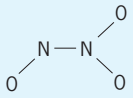	Condense gaseous mixture of NO and NO_2
NO_2		Heat lead nitrate
N_2O_4	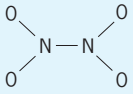	Heat lead nitrate
N_2O_5 Gas	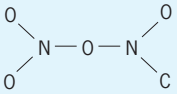	Treat N_2O_4 with ozone
Solid	$NO_2^+ \cdot NO_3^-$	

of anhydrous metal nitrates and nitrato-coordination complexes.

Pentoxide. The ionic nitronium nitrate structure $NO_2^+ \cdot NO_3^-$ found for N_2O_5 in the solid state accounts for its anomalously high melting point. In solution in sulfuric, nitric, or phosphoric acids the oxide has the same ionic structure. Solid dinitrogen pentoxide readily volatilizes, and the molecular type of structure found in the gaseous state is observed also in solutions of the oxide in low dielectric solvents such as carbon tetrachloride and chloroform. Sodium metal reacts with the liquid oxide, liberating nitrogen dioxide and forming sodium nitrate. Gaseous dinitrogen pentoxide decomposes readily, as in reaction (7),and is a strong oxidizing agent.

$$N_2O_5 \rightleftharpoons 2NO_2 + O_2 \qquad (7)$$

With water it is converted to nitric acid. *See* NITROGEN; OXYGEN.

Cyril C. Addison

Nitroparaffin

Any derivative of an aliphatic hydrocarbon that contains one or more $—NO_2$ groups bonded via nitrogen to the carbon framework. Nitroparaffins are also known as nitroalkanes.

Preparation. Low-molecular-weight nitroparaffins are prepared via the vapor-phase nitration of alkanes at >400°C (750°F). However, the process is not generally satisfactory for higher-molecular-weight nitroparaffins because of polynitration and chain cleavage. The direct nitration of propane is used commercially to prepare nitromethane (boiling point 101°C or 214°F), nitroethane (bp 114°C or 237°F), 1-nitropropane (bp 131°C or 268°F), and 2-nitropropane (bp 120°C or 248°F). Nitroparaffins are prepared in the laboratory by the reaction of nitrite salts with alkyl bromides or iodides, from the oxidation of amines or oximes by using peroxycarboxylic acids, and by the chain homologation of simple nitroparaffins. *See* NITRATION.

Nitromethane, nitroethane, and the nitropropanes are useful solvents with high dielectric constants that readily dissolve many polymers. In addition, these simple nitroparaffins are versatile intermediates for the synthesis of specialty chemicals.

Reactions. Primary and secondary nitroparaffins are in tautomeric equilibrium with the aci-nitro form, as shown in reaction (1). Since tertiary nitroparaffins

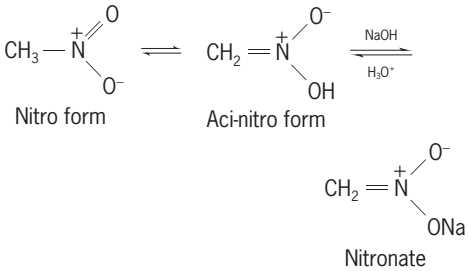

(1)

do not contain a proton attached to the carbon bearing the nitro group, they are unable to tautomerize in this way. Both primary and secondary nitroparaffins are readily deprotonated with alkali to produce the corresponding nitronate salts [reaction (1)]. The pK_a of nitromethane is 10.2; and this high acidity, for a carbon acid, is the result of the powerful electron-withdrawing capacity of the nitro substituent. On reaction with alkyllithium bases, primary nitroparaffins (RCH_2NO_2) are converted into α,α-dianions [$RCH(Li){=}NO_2Li$]; whereas secondary nitroparaffins, for example, 2-nitropropane, are converted into α,β-dianions [$MeC({=}CH_2)N(OLi)_2$], where Me represents the $—CH_3$ group. All of these nitro-substituted carbanions are versatile synthetic intermediates. *See* PK; REACTIVE INTERMEDIATES; TAUTOMERISM.

Primary and secondary nitroparaffins react with aldehydes and ketones under catalysis by a base (such as sodium hydroxide) to produce β-hydroxynitroparaffins. This process, known as the Henry or nitro-aldol reaction, is most useful in synthesis [reaction (2)]. Both α-protons of primary nitroparaf-

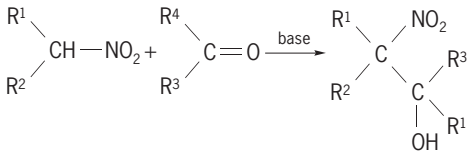

(2)

fins and all three α-protons of nitromethane may be replaced sequentially by hydroxyalkyl groups in this way. For example, nitromethane and formaldehyde may be reacted to produce the alcohol ($HOCH_2CH_2NO_2$), diol [$(HOCH_2)_2CHNO_2$], or triol [$(HOCH_2)_3CNO_2$] adducts. β-Nitroalcohols are easily dehydrated to provide nitroalkenes or reduced to produce important amino alcohols. In consequence of the electronegative nitro group, nitroalkenes are reactive dienophiles toward Diels-Alder reactions with 1,3-dienes. Additionally, they readily undergo addition of nucleophiles to the carbon β to the nitro substituent (Michael addition). *See* CATALYSIS; DIELS-ALDER REACTION.

Nitronate salts are hydrolyzed by cold aqueous mineral acids to produce aldehydes or ketones and nitrous oxide. This process, known as the Nef reaction, is also carried out by the oxidation of nitronate salts using ozone or potassium permanganate. Alternatively, primary nitroparaffins are converted into carboxylic acids and hydroxylamine on heating with concentrated mineral acids. Both primary and secondary nitroparaffins ($R^1R^2CHNO_2$) readily add to alkenes ($CH_2{=}CHA$) substituted by electron-withdrawing substituents (A = CN, COMe, CO_2Me, NO_2, and so forth) to provide the adducts $R^1R^2C(NO_2)CH_2CH_2A$. The process, which employs catalysis by a base, involves the addition of the nitronate anion.

Nitroparaffins may be reduced to produce amines by heterogeneous catalytic hydrogenation over

transition metals. Transfer hydrogenation from hydrogen donors such as ammonium formate is also effective for the synthesis of amines. The hydrogenation may be stopped at the intermediate hydroxylamine stage. In general, reductions of nitroparaffins by metals such as sodium or potassium (dissolving metals) in acid solution are less efficient for amine preparation. Both secondary and primary nitroalkanes are reduced to produce ketone or aldehyde oximes on reaction with tin(II), chromium(II), or titanium(III) salts. *See* HYDROGENATION.

Primary nitroalkanes (RCH_2NO_2) are dehydrated by reaction with phenyl isocyanate and triethylamine to produce nitrile oxides ($RC{\equiv}N^+{-}O^-$). These reactive intermediates dimerize readily, or they are trapped by reaction with unsaturated molecules to produce heterocyclic adducts [reaction (3)]. For

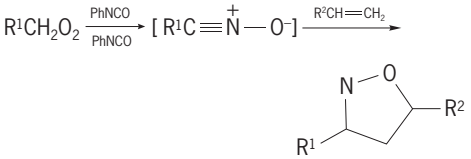

(3)

example, alkenes react with nitrile oxides to produce Δ^2-isoxazolines. Such heterocycles are useful intermediates in further synthetic transformation.

A. G. M. Barrett

Bibliography. D. H. R. Barton and W. D. Ollis (eds.), *Comprehensive Organic Chemistry*, vol. 2, 1979; R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; H. Feuer, *The Chemistry of the Nitro and Nitroso Group, Interscience*, pt. 1, 1969.

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

12 NOB-PAP

www.MHEST.com

N

Nobelium

A chemical element, No, atomic number 102. Nobelium is a synthetic element produced in the laboratory. It decays by emitting an alpha particle, that

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

is, a doubly charged helium ion. Only atomic quantities of the element have been produced to date. Nobelium is the tenth element heavier than uranium to be produced synthetically. It is the thirteenth member of the actinide series, a rare-earth-like series of elements. *See* ACTINIDE ELEMENTS; PERIODIC TABLE; RADIOACTIVITY; RARE-EARTH ELEMENTS; TRANSURANIUM ELEMENTS. Paul R. Fields

Bibliography. S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; G. T. Seaborg and W. D. Loveland, *The Elements Beyond Uranium*, 1990.

Noeggerathiales

An incompletely known and poorly defined group of vascular plants whose geologic range extends from Upper Carboniferous to Triassic. Their taxonomic status and position in the plant kingdom are uncertain since morphological evidence (because of the paucity of the fossil record) does not make it possible to place the group confidently in any recognized major subdivision of the vascular plants. A rather heterogeneous assemblage of foliar and vegetative organs assignable to fairly well-defined genera have been placed in the group, thus somewhat forcing the concept that these parts constitute a natural order of vascular plants. Internal anatomy is unknown, with exception of the possible noeggerathialean genus *Sphenostrobus*. Noeggerathialean genera include *Noeggerathia, Noeggerathiostrobus, Tingia, Tingiostachya*, and *Discinites*. The reproductive organs are strobiloid, with whorled organization, and vary from homosporous to heterosporous. In *Discinites* the number of megaspores may range from 1 to 16 per sporangium. Foliar organs vary from nonarticulate and fernlike to anisophyllous four-ranked fronds. The Noeggerathiales have been proposed in the evolutionary scheme for vascular plants. *See* PALEOBOTANY; PLANT KINGDOM. Elso S. Barghoorn

Noise measurement

The process of quantitatively determining one or more properties of acoustic noise. In noise assessment and control studies, knowledge of the physical properties of the undesirable sound is the initial step toward understanding the situation and what should be done to reduce or eliminate a noise problem. The instruments and techniques used for measuring noise are of fundamental importance in noise assessment and control. Many measurements may be required to characterize noise adequately because generally it is possible to make a noise measurement only at a particular point in space, while a noise field often has significant spatial variations.

The most common measures of noise are of the magnitude and frequency content of the noise sound pressure, time-averaged or as a function of time. Of increasing interest are metrics of sound quality (that

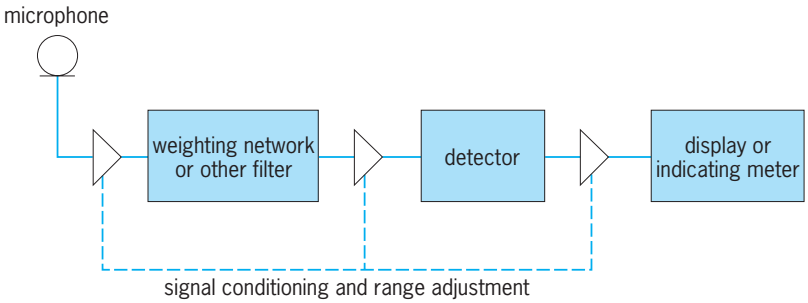

Fig. 1. Functional block diagram of a sound-level meter.

may include both physical and psychoacoustic factors), such as loudness, pitch strength, and fluctuation strength. To characterize the noise output of a source, sound power level may be determined. To locate a source, or to quantify propagation paths, sound intensity level may be measured.

Essentially all noise measurements are performed using electronic equipment. An electroaoustic transducer (a microphone in air and other gases; a hydrophone in water and other liquids) transforms sound (usually the sound pressure) at the point of observation into a corresponding electrical signal. This electrical signal is then operated on by analog or digital means with devices such as signal conditioners, filters, and detectors to determine the value (or values) of interest. This value is then given on an indicating meter or digital display. *See* ELECTRIC FILTER; HYDROPHONE; MICROPHONE; TRANSDUCER.

For noise measurements in air, a sound-level meter is the most commonly used instrument. Specifications for such instruments are given by national and international standards. The simplest sound-level meter comprises a microphone, a frequency-weighting filter, a root-mean-square (rms) detector, and logarithmic readout of the sound pressure level in decibels relative to 20 micropascals (**Figs. 1** and **2**). Standard frequency weightings, designated A and C, were originally developed to approximate human response to noise at low and high levels, respectively, but now are used as specified in standards and legislation without regard to their origin (**Fig. 3**). [Other weightings also have been used and may be found in older or special-purpose sound-level meters.] More sophisticated sound-level meters incorporate standardized octave-band or fractional-octave-band filters, and provide additional metrics and analysis capabilities. Formerly such capability was found only in instruments called analyzers, but the differences that distinguish a sound-level meter from an analyzer and a computer are becoming size and portability rather than analysis capability. *See* DECIBEL; LOUDNESS.

Applications. Selection of a noise metric and measurement instrument generally depends upon the application. Customary metrics vary because, in practice, what is desired to be known varies with application. Emission metrics characterize noise source output. Immission metrics characterize noise received by a listener. Generalization is difficult, and some overlap is inevitable, but it is convenient to recognize four classes of application: (1) discrete source emission, (2) hearing conservation, (3) outdoor environmental noise, and (4) indoor room noise. Within these classes there are found additional subclasses, measurements for enforcement of legal limits often

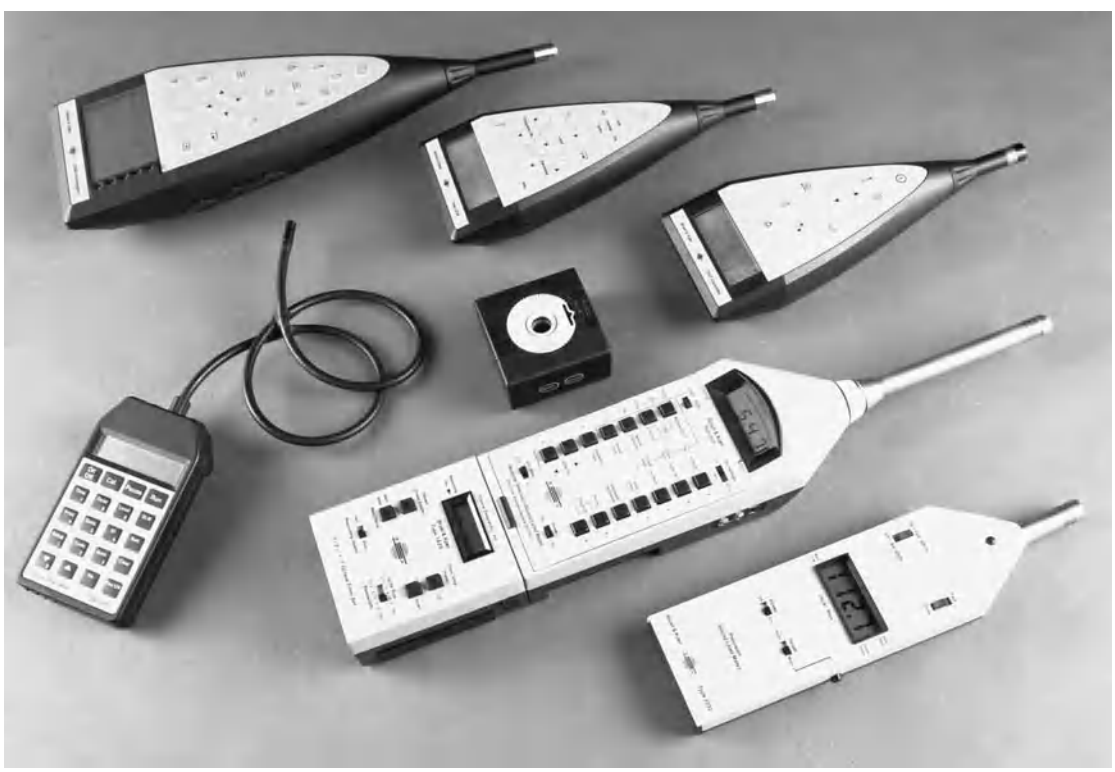

Fig. 2. Sound-level meters. (*Brüel & Kjær*)

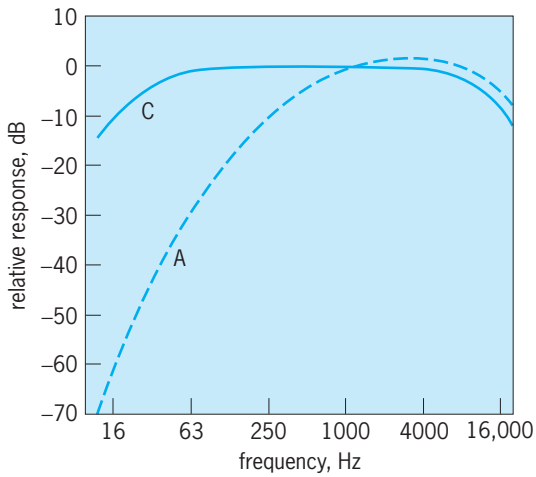

Fig. 3. Frequency weightings A and C specified for sound-level meters in standards of the International Electrotechnical Commission and the American National Standards Institute.

being done differently than those for comfort, aircraft sonic boom being measured differently than power plant noise, and so on.

Source emission is best characterized by sound power level, a quantity not directly measurable. Sound intensity or pressure measurements under controlled conditions are used to determine power. Sound power level is expressed in decibels relative to 1 picowatt. (While the decibel unit is used for sound power level as for sound pressure level, the reference quantity is different and power is a quantity fundamentally different from pressure.) Also used for the source emission description is the sound pressure level at a defined operator or bystander position. Because sound pressure level generally is a function of both the source emission and the acoustical environment, it is important that the measurement conditions be properly specified. Frequency analysis and spectral data are common as these are needed to gauge human response and to calculate the noise immission that is expected with a particular installation (in an enclosure or behind a noise barrier, for example). Transient exposures may be characterized by sound exposure level, the level of an integral of the mean-square frequency-weighted sound pressure.

Hearing conservation noise measurements generally are concerned with two quantities: time-weighted-average A-weighted sound level expressed as a noise dose, and C-weighted peak sound pressure level. The former characterizes the potential for damage due to chronic exposure, the latter the potential for immediate damage. Different jurisdictions have established various thresholds and time weightings. In the United States, for example, different schemes are used for industrial, underground mine, and military exposures.

Outdoor environmental noise may be characterized by short-term metrics, usually A-weighted sound level, as well as by longer-term metrics. Statistical metrics (L_N) characterize the percentage of time that the sound level exceeds a stated value (L_{90}, for example, being the level exceeded during 90% of the observation time, and L_{10} the level exceeded 10%). Such measures give an indication of both the range of levels encountered and their duration. Time-average sound level (sometimes designated L_{eq}) disregards variations to arrive at an equivalent steady level. The period of the time average depends on the application, with averages hourly, daily, or over as much as a year being common. Modified time-average sound levels include the day-night average level and the day-evening-night average level, with penalty weightings added for nighttime (10 dB) and both night (10 dB) and evening (5 dB) noise, respectively, people being presumed more sensitive to noise at the penalty times.

Indoor room noise measures commonly incorporate both amplitude and frequency dependence. Established criteria, such as NC, RC, and NCB, are based on octave-band sound pressure levels plotted on a special chart (or the equivalent in a spreadsheet) to determine the value of the criterion. The details of these methods continue to evolve.

Metrics. Several common metrics are used to characterize noise. Many metrics are stated in units of decibels, and caution is necessary in comparing measurements of one metric with those of another. The effect of a given number of decibels depends as much, if not more, on the physical quantity being described by the level as on the numerical value of the level. A C-weighted peak sound level of 80 dB re 20 μPa with a duration of only a few milliseconds, for example, will have an impact quite different from that of a time-average sound level of 80 dB re 20 μPa.

The sound power level is ten times the logarithm to the base ten of the ratio of a given sound power, the sound energy radiated (by a source) per unit of time, in a stated frequency band to the standard reference sound power of 1 picowatt.

The sound intensity level is ten times the logarithm to the base ten of the ratio of the magnitude of the sound intensity vector, the vector giving the net rate of sound power transmission through a unit area normal to the vector, in a stated direction to the standard reference intensity of 1 picowatt per square meter. *See* SOUND INTENSITY.

The sound pressure level is ten times the logarithm to the base ten of the ratio of the square of sound pressure in a stated frequency band to the square of the standard reference sound pressure (20 μPa in air, 1 μPa in water). Root-mean-square pressure is assumed unless otherwise stated, and for root-mean square pressure the type of time average and time constant or averaging time should be stated. *See* SOUND PRESSURE.

Octave and one-third-octave band levels are levels of the stated kind (power, intensity, or pressure), determined in standardized frequency bands having uniform width on a logarithmic scale. Such levels usually are determined and reported without A or C weighting.

The sound level is the A-weighted exponential time-average root-mean-square sound pressure level

obtained with the F time constant (0.125 second). Alternatively C weighting or the S time constant (1.0 s) may be specified.

The time-average sound level (equivalent continuous A-weighted sound pressure level) is the A-weighted linear time-average root-mean-square sound pressure level obtained with a specified averaging time.

The sound exposure level is ten times the logarithm to the base ten of the ratio of the time integral of the squared A-weighted sound pressure, over a stated time interval or event, to the product of the squared standard reference sound pressure (20 μPa in air) and a reference duration of 1 second.

The sound level exceedance is the sound level exceeded a stated percent of the observation time. It is usually determined from a probability histogram of measured sound levels.

The noise dose is the standardized time-weighted average A-weighted sound level expressed as a percentage of an allowable limit. It is usually employed in assessing the potential for noise to cause hearing impairment.

Noise Criteria (NC, NCB) and Room Criteria (RC) are ratings used to assess noise in rooms.

The root-mean-square sound pressure is an element of several noise metrics. The root-mean-square is the effective average, the usual average (mean) being zero for an oscillating quantity such as sound pressure. The common mathematical definition of root-mean-square for a time-dependent signal, the square root of the time-mean of the square of the signal, presumes a signal of infinite duration and time-invariant root-mean-square value, and this assumption is not always appropriate for noise measurements. The exponential time-average root-mean-square value, $p_{\mathrm{rms},\tau}$, of a signal $p(t)$ is determined from Eq. (1), where τ is the time constant and ξ

$$p_{\mathrm{rms},\tau}(t;\tau) = \sqrt{\frac{1}{\tau}\int_{-\infty}^{t} p^2(\xi)e^{-(t-\xi)/\tau}\,d\xi} \qquad (1)$$

is a dummy variable of integration. This kind of root-mean-square value is easily implemented as a running average, being realized with the a simple RC electronic circuit or other means. Selection of a short time constant permits tracking of variations in the signal, while a long time constant effectively eliminates them. The linear time-average root-mean-square value, $p_{\mathrm{rms},T}$, is determined from Eq. (2),

$$p_{\mathrm{rms},T}(t;T) = \sqrt{\frac{1}{T}\int_{t-T}^{t} p^2(\xi)\,d\xi} \qquad (2)$$

where T is the averaging (also called integration) time. Linear time averaging gives equal weight to all of the signal within the averaging interval.

Instruments. Sound-level meter is a general term for any instrument designed to measure sound (and noise) levels. An instrument that conforms with the requirements of a relevant national or international standard specification should be used so that the measurements will have a known uncertainty that permits results to be compared with confidence. (Standards for sound-level meters require that an instrument in conformance be marked as such; hence, lack of marking implies that the device is not in conformance with a standard.) The minimum requirement for a standard sound-level meter is that it provide an exponential time-weighted root-mean-square sound level, F or S, with frequency-weighting A or C. Generally three classes of instrument accuracy are recognized: types 0, 1, and 2, having overall uncertainty generally considered to be ± 0.5, 1, and 2 dB, respectively. (The uncertainty specification is frequency-dependent and for certain situations may exceed these values.) F (formerly "fast") and S (formerly "slow") time weighting correspond to exponential time constants of 0.125 and 1.0 second, respectively. An impulse sound-level meter includes the ability to measure the root-mean-square sound pressure of rapid transients with I (formerly "impulse") time weighting, specified as an exponential time average with a 35-millisecond rise time followed by a peak detector with an exponential decay having a 1.5-second time constant. An integrating-averaging sound-level meter includes the capability of measuring time-average sound level and usually sound exposure level. Other metrics may be included at the option of the manufacturer. A noise dosimeter is a kind of special-purpose sound-level meter designed to be worn, providing a readout of the noise dose experienced by the wearer.

A frequent addition to the basic sound-level meter is an octave or one-third-octave band filter set that permits determination of spectrum levels. The most common approach is to measure individual frequency bands in succession, called serial analysis. Parallel analysis, measuring all bands simultaneously, previously available in analyzers intended for laboratory use, is becoming available in hand-held portable instruments.

A significant distinction between an analyzer and a sound-level meter is that, while the latter is a complete instrument, an analyzer usually requires adjunct equipment such as a transducer and input signal conditioner. Octave-band and fractional-octave-band analyzers are based on a set of filters called constant-percentage-bandwidth (cpb) filters or constant Q filters, and produce sound spectra that have uniform resolution on a logarithmic frequency scale. A logarithmic scale is convenient to display the wide range of frequencies to which human hearing is sensitive, and correlates reasonably well with the human sensation of pitch. Certain measurements are more conveniently displayed on a linear frequency scale. A filter set with constant bandwidth (cb), having uniform resolution on a linear frequency scale, is more appropriate for this purpose. A device commonly called an FFT (fast Fourier transform) analyzer performs constant-bandwidth spectrum analysis digitally, using an efficient implementation of the discrete Fourier transform. FFT analyzers usually implement a number of analyses that extend and

supplement spectrum determination. *See* ACOUSTIC NOISE; FOURIER SERIES AND TRANSFORMS; PITCH; SOUND. Joseph Pope

Bibliography. Acoustical Society of America, *Catalog of ANSI, ISO, and IEC Standards in Acoustics*, revised annually; American Society of Heating, Refrigerating, and Air-Conditioning Engineers Inc., *ASHRAE Handbook: Fundamentals*, 2001; American Society of Heating, Refrigerating, and Air-Conditioning Engineers Inc., *ASHRAE Handbook: Heating, Ventilating, and Air-Conditioning Applications*, 1999; L. L. Beranek and I. L. Ver (eds.), *Noise and Vibration Control Engineering: Principles and Applications*, John Wiley, 1992; M. J. Crocker (ed.), *Encyclopedia of Acoustics*. 4 vols., John Wiley, 1997; C. M. Harris (ed.), *Handbook of Acoustical Measurements and Noise Control*, 3d ed., McGraw-Hill, 1991, and Acoustical Society of America, 1997; A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications*, McGraw-Hill, 1981, and Acoustical Society of America, 1989.

Nomograph

A graphical relationship between a set of variables that are related by a mathematical equation or law. The fundamental principle involved in the construction of a nomographic or alignment chart consists of representing an equation containing three variables,

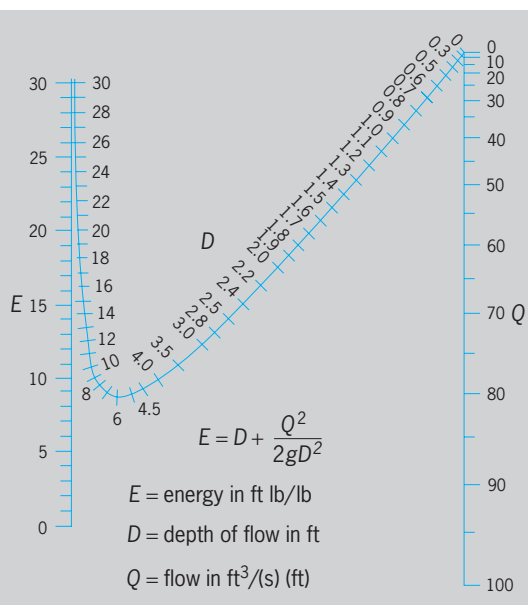

Fig. 1. Nomograph for energy content of a rectangular channel with uniform flow.

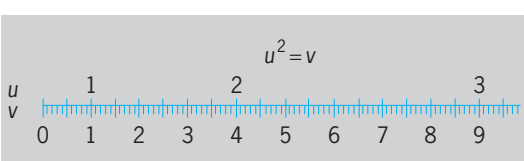

Fig. 2. Conversion scales for finding squares and square roots.

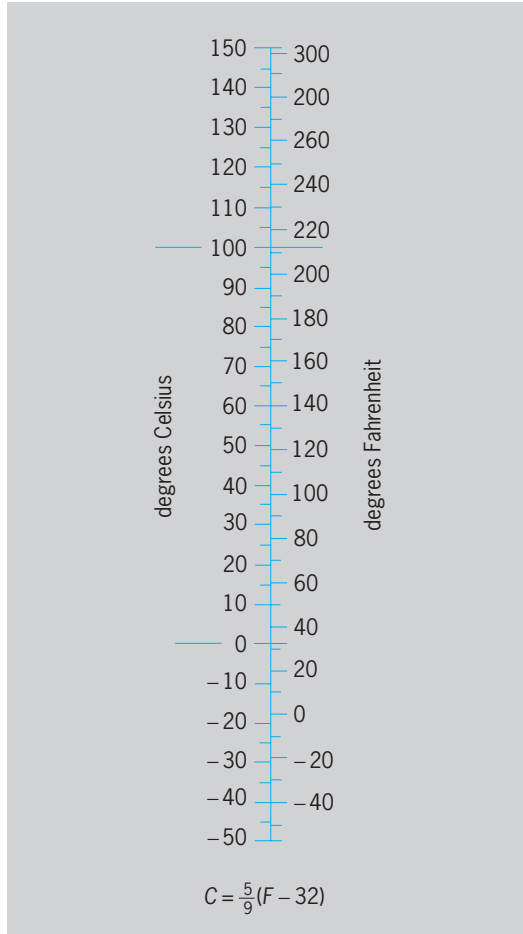

Fig. 3. Stationary scales relate Celsius and Fahrenheit temperatures.

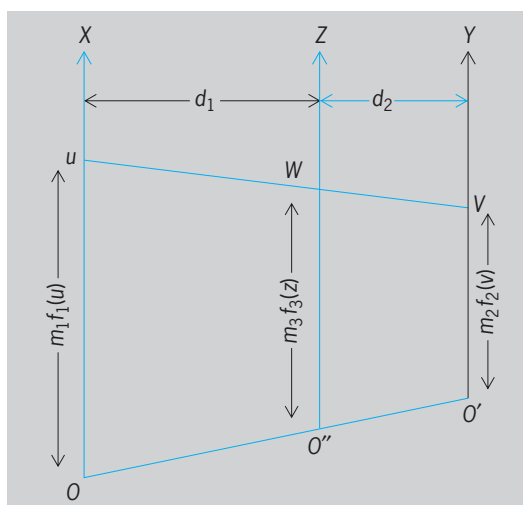

Fig. 4. Alignment chart for simple summation.

$f(u,v,w) = 0$, by means of three scales in such a manner that a straight line cuts the three scales in values of u, v, and w, satisfying the equation. The cutting line is called the isopleth or index line. Numbers may be quickly and easily read from the scales of such a chart even by one unfamiliar with the construction of the chart and the equation involved. **Figure 1** illustrates such an example. Assume that it is desired to find the value of E when $D = 2$ and $Q = 50$. Lay a

straightedge through 50 on the Q scale and through 2 on the D scale and read 11.8 at its intersection with the E scale. As another example, it might be desired to know what value or values of D should be used if E and Q are required to be 10 and 60, respectively. A straightedge through $E = 10$ and $Q = 60$ cuts the D scale in two points, $D = 2.8$ and 9.4. This is equivalent to finding two positive roots of the cubic equation $D^3 - 10D^2 + 56.25 = 0$. It is assumed that $g = 32$ ft/s^2 in this equation.

Scale. The graphical scale is a curve or straight line, called an axis, on which is marked a series of points or strokes corresponding to a set of numbers arranged in order of magnitude. If the distances between successive strokes are equal, the scale is uniform; otherwise the scale is nonuniform. The scale on a yardstick or thermometer is uniform, whereas a logarithmic scale on a slide rule is nonuniform.

Representation of a function by a scale. Consider the function $f(u)$. Lay off, from a fixed point O on a straight line or curve, lengths equal to $f(u)$ units; mark at the strokes indicating the end of each unit the corresponding value of u. If the unit of measure is an inch, the equation of the scale is $x = f(u)$ inch. More generally, the equation of the scale is $x = mf(u)$ units, where the constant m (modulus) regulates the length of scale used for the required values of the variable u needed.

Stationary scales. A relation between two variables of the form $v = f(u)$, or $f(v) = f(u)$, can be represented

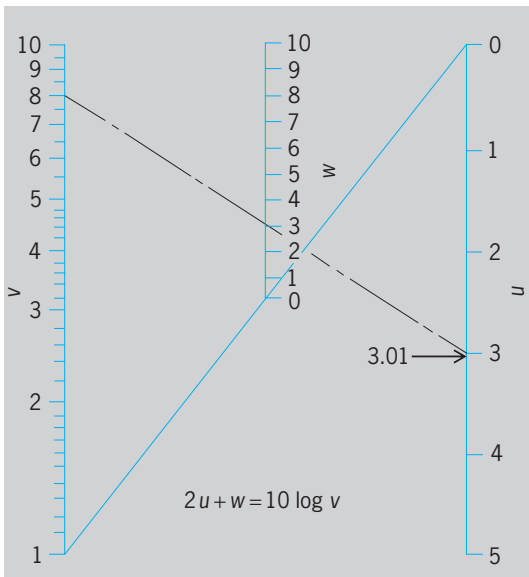

Fig. 5. Variation of alignment chart using logarithm.

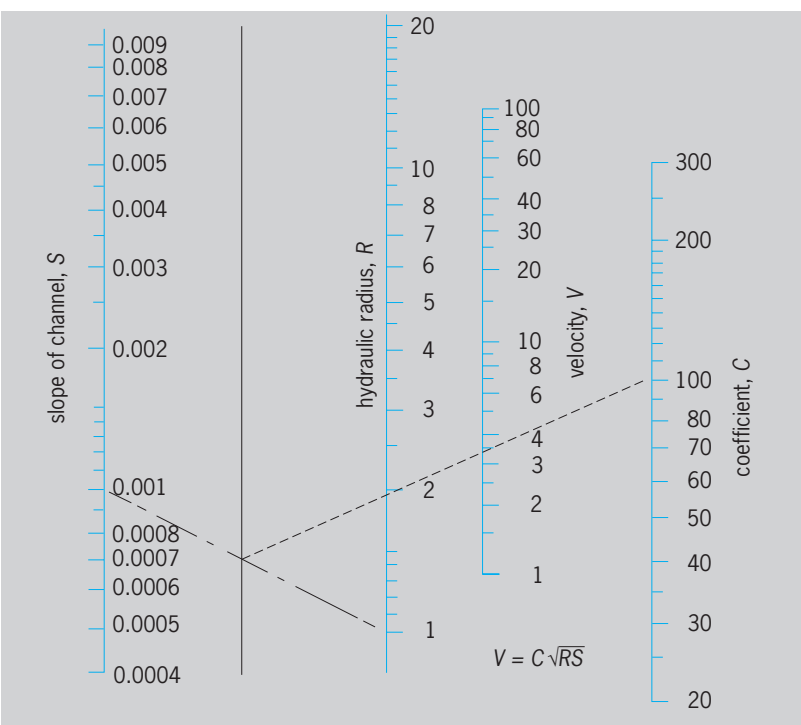

Fig. 6. Nomograph for velocity of flow of water in open channels.

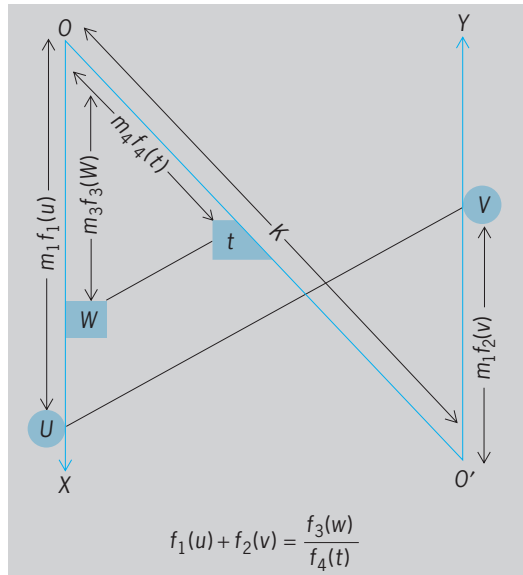

Fig. 7. Alignment chart for summation involving a ratio.

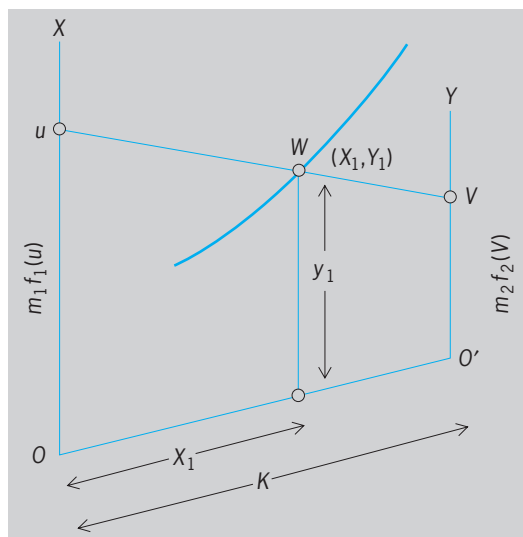

Fig. 8. Alignment chart for summation using a product.

by the natural scales $x = mv$ and $x = mf(\upsilon)$; or $x = mf(\upsilon)$ and $x = mF(u)$, on the opposite side of the same line or axis. **Figure 2** shows the relation $u^2 = \upsilon$ using the natural scales $x = mu^2$ and $x = mv$, where in this illustration $m = 0.43$ and the unit is an inch. By using logarithms the above equation becomes $2 \log u = \log \upsilon$, and the scales are $x = m(2 \log u)$ and $x = m \log \upsilon$.

Adjacent stationary scales may be used to advantage in representing the relationship between the two variables in a conversion formula. **Figure 3** shows the relation between degrees Celsius and Fahrenheit. It is easy to see that $F = 140$ when $C = 60$; and when $F = -40$, $C = -40$.

Perpendicular scales. A relation between two variables u and υ of the form $\upsilon = f(u)$, $f(u,\upsilon) = 0$, or $f(u) = F(\upsilon)$ can be represented by constructing two scales $x = mf(u)$ and $y = mF(\upsilon)$ on perpendicular axes. Any pair of values of u and υ will determine a point in the plane. The locus of all such points is a curve which represents the relationship between the variables u and υ. The various types of coordinate paper, sold commercially, are constructed in this manner. Log-log, semilog, and reciprocal coordinate papers are probably the most common. These types of scales and their combinations are essential in the construction of nomographic charts, especially when the number of variables involved exceeds three.

Types of nomographic charts. The form of the equation serves to classify the type of chart. An equation of the form of Eqs. (1) and (2) leads to a chart

$$f_1(u) + f_2(v) = f_3(w) \tag{1}$$

$$f_1(u) + f_2(v) + \cdots = f_n(t) \tag{2}$$

of the type shown in **Fig. 4**. The equations of the three scales are $x = m_1 f_1(u)$, $m_2 f_2(\upsilon)$, and $[m_1/m_2/(m_1 + m_2)]f_3(w)$, respectively; and $d_1/d_2 = m_1/m_2$. The equation $2u + w = 10 \log \upsilon$ is in this form. Taking $m_1 = m_2 = 1$ and $d_1 = d_2$, the scales are $10 \log \upsilon$, $-2u$, and $w/2$. The completed chart is shown in **Fig. 5**. As an example, if $\upsilon = 8$ and $w = 3$, the value of u is found to be 3.01.

Another form of equation that leads to a similar chart is Eq. (3). Using logarithms, this equation takes the form of Eq. (1), which is Eq. (4). The Chezy

$$f_1(u) f_2(v) \cdots = f_n(t) \tag{3}$$

$$\log f_1(u) + \log f_2(v) + \cdots = \log f_n(t) \tag{4}$$

formula for velocity of the flow of water $\upsilon = c(RS)^{1/2}$ is of this type. When logarithms are used, the equation becomes $\log \upsilon = \log c + {}^1/_2 \log R + {}^1/_2 \log S$, which may be written ${}^1/_2 \log S + {}^1/_2 \log R = \log \upsilon - \log c = Q$, where Q is a dummy variable. The completed chart is shown in **Fig. 6**. To find the value of υ when $R = 1$, $S = 0.001$, and $c = 100$, set the straightedge on $S = 0.001$ and $R = 1$; now connect the point of intersection on the dummy scale to $c = 100$ and read $\upsilon = 3$ at the point of crossing on the υ scale.

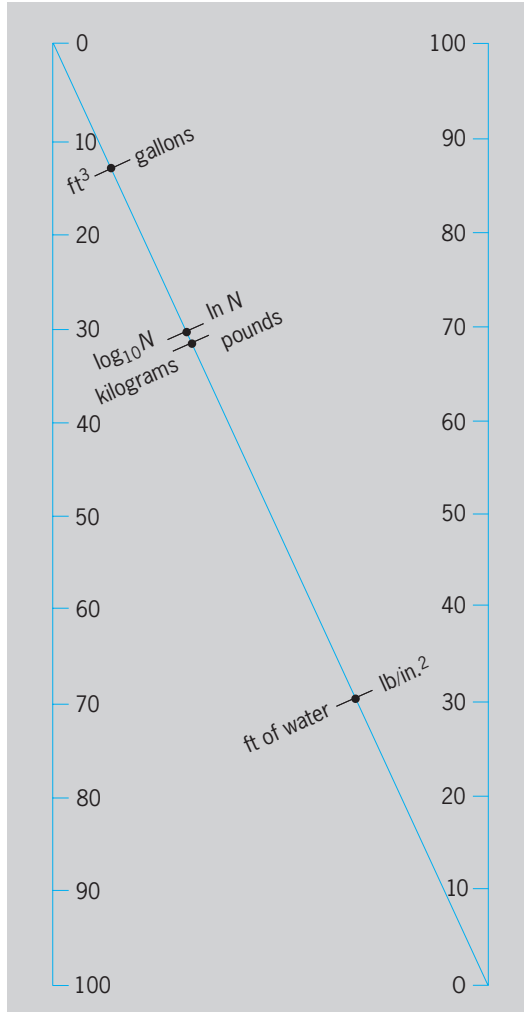

Fig. 9. Conversion chart relates several different units of measure on linear scales.

Alternatively the equation may be of the form of Eq. (5). Using logarithms, Eq. (5) also takes the form of Eq. (1), which is Eq. (6).

$$[f_1(u)]^{f_2(\upsilon)} = f_3(w) \tag{5}$$

$$\log f_2(v) + \log \log f_1(u) = \log \log f_3(w) \tag{6}$$

A second form is Eq. (7), where $m_1/K = m_3/m_4$. Such a chart is shown in **Fig. 7**.

$$f_1(u) + f_2(v) = \frac{f_3(w)}{f_4(t)} \tag{7}$$

A third form is Eq. (8).

$$f_1(u) + f_2(v) f_3(w) = f_4(w) \tag{8}$$

Equations (9) and (10) apply. Such charts are

$$x_1 = \frac{m_1 K}{m_1 f_3(w) + m_2} f_3(w) \tag{9}$$

$$y_1 = \frac{m_1 m_2}{m_1 f_3(w) + m_2} f_4(w) \tag{10}$$

illustrated by Fig. 1 and **Fig. 8**. This is a frequently encountered type of equation. In constructing the nomograph of Fig. 1, first the equation $D = D + Q^2/2gD^2$ was rewritten in the form $-Q^2 + 2gD^2E = 2gD^2$, which corresponds to the basic form of Eq. (8).

A fourth basic form is Eq. (11), where $m_1/m_2 =$

$$\frac{f_1(u)}{f_2(v)} = \frac{f_3(w)}{f_4(t)} \tag{11}$$

m_3/m_4. This can be treated as a Z-chart by using natural scales, or it takes the form of Eq. (1) by using logarithms.

A fifth form is the conversion chart, Eq. (12). This

$$f_1(u) = Cf_2(v) \tag{12}$$

is a special case of Eq. (1) or (11) and is useful when several conversions are to be made. Examples are **Figs. 9** and **10**. In practice it may be useful or even necessary to use two or more combinations of these basic types.

Determinant as a basis of a nomograph. The condition that the three points (x_1,y_1), (x_2,y_2), and (x_3,y_3) lie on a straight line is that the determinant equal zero, as in Eq. (13). If an equation that relates three variables u, v, and w can beexpressed in the determinant form shown as Eq. (14), then a nomograph can be constructed from parametric equations (15).

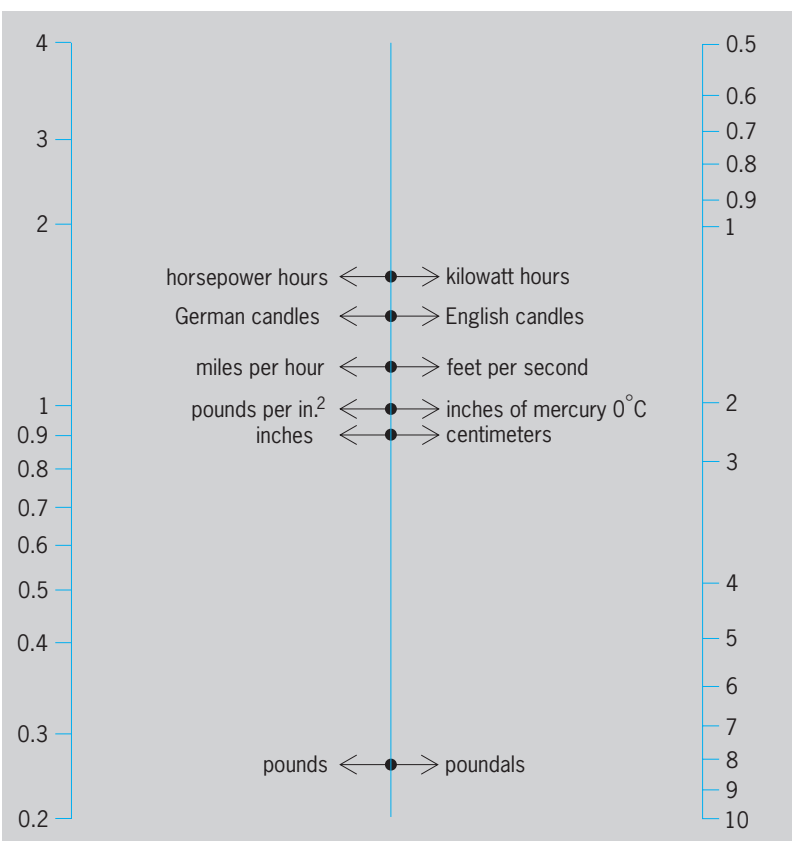

Fig. 10. Conversion chart relates several different units of measure on an exponential scale.

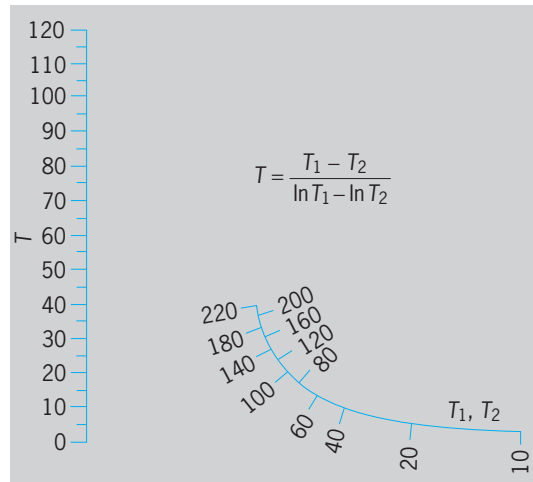

Fig. 11. Nomograph for mean temperature. To find mean temperature, lay straightedge between two given temperatures on curved scale, and then read the mean temperature on the vertical scale.

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0 \tag{13}$$

$$\begin{vmatrix} f_1(u) & g_1(u) & 1 \\ f_2(v) & g_2(v) & 1 \\ f_3(w) & g_3(w) & 1 \end{vmatrix} = 0 \tag{14}$$

$$\begin{aligned} x &= f_1(u),\ y = g_1(u) \\ x &= f_2(v),\ y = g_2(v) \\ x &= f_3(w),\ y = g_3(w) \end{aligned} \tag{15}$$

The mean temperature equation (16) expressed as a determinant is Eq. (17). **Figure 11** shows the resulting nomograph.

$$T = \frac{T_1 - T_2}{\ln T_1 - \ln T_2} \tag{16}$$

$$\begin{vmatrix} 0 & T & 1 \\ 1/\ln T_1 & T_1/\ln T_1 & 1 \\ 1/\ln T_2 & T_2/\ln T_2 & 1 \end{vmatrix} = 0 \tag{17}$$

Figure 1 could also have been constructed from the determinant form after the energy equation is written as Eq. (18).

$$\begin{vmatrix} 0 & (12.2/100^2)Q^2 & 1 \\ 9 & -E/3 & 1 \\ \dfrac{9D^2}{D^2 + \dfrac{100^2}{73.2g}} & \dfrac{-D^3}{3D^2 + \dfrac{100^2}{73.2g}} & 1 \end{vmatrix} = 0 \tag{18}$$

Circular nomographs. The general form of the basic determinant for a circular nomograph is Eq. (19), where the u and v scales lie on circular axes having the same center and with radii equal to $^1/_2$. Consider

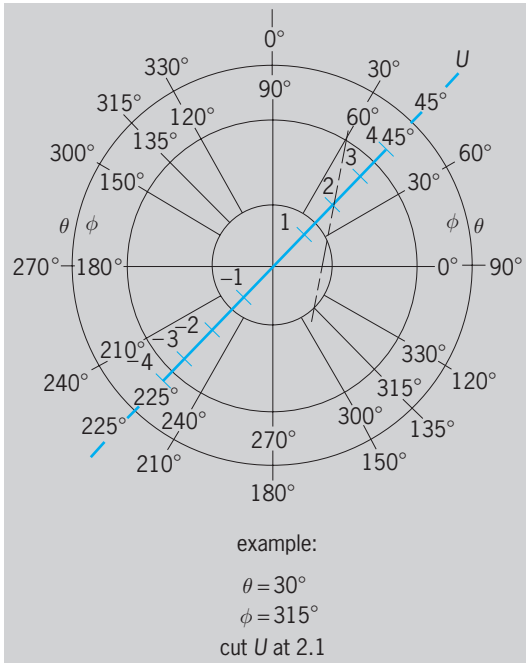

Fig. 12. Circular nomograph which results from trigonometric equation expressed in form of determinant.

the equation $AB\cos(\theta-\mu)+BU(\sin\phi-\cos\theta)+AU(\sin\theta-\cos\phi)=0$. Expressed in the form of a determinant, it is Eq. (20), which leads to the nomo-

$$\begin{vmatrix} \dfrac{1}{1+[f_1(u)]^2} & \dfrac{f_1(u)}{1+[f_1(u)]^2} & 1 \\ \dfrac{1}{1+[f_2(v)]^2} & \dfrac{f_2(v)}{1+[f_2(v)]^2} & 1 \\ \dfrac{1}{1+f_3(w)} & 0 & 1 \end{vmatrix} = 0 \qquad (19)$$

$$\begin{vmatrix} A\sin\theta & A\cos\theta & 1 \\ B\cos\phi & B\sin\phi & 1 \\ U & U & 1 \end{vmatrix} = 0 \qquad (20)$$

graph of **Fig. 12**. *See* DETERMINANT; LINEAR SYSTEMS OF EQUATIONS.

Raymond D. Douglass

Bibliography. L. H. Johnson, *Nomography and Empirical Equations*, 1952, reprint 1978; J. F. Kuong, *Applied Nomography*, 3 vols., 1965–1969; J. Molnar, *Nomographs: What They Are and How to Use Them*, 1981.

Nondestructive evaluation

Nondestructive evaluation (NDE) is a technique used to probe and sense material structure and properties without causing damage. It has become an extremely diverse and multidisciplinary technology, drawing on the fields of applied physics, artificial intelligence, biomedical engineering, computer science, electrical engineering, electronics, materials science and engineering, mechanical engineering, and structural engineering. Historically, NDE techniques have been used almost exclusively for detection of macroscopic defects (mostly cracks) in structures which have been manufactured or placed in service. Using NDE for this purpose is usually referred to as nondestructive testing (NDT).

A developing use of NDE methods is the nondestructive characterization (NDC) of materials properties (as opposed to revealing flaws and defects). Characterization typically sets out to establish absolute or relative values of material properties such as mechanical strength (elastic moduli), thermal conductivity or diffusivity, optical properties, magnetic parameters, residual strains, electrical resistivity, alloy composition, the state of cure in polymers, crystallographic orientation, and the degree of crystalline perfection. Nondestructive characterization can also be used for a variety of other specialized properties that are relevant to some aspect of materials processing in production, including determining how properties vary with the direction within the material, a property called anisotropy.

Much effort has been directed to developing techniques that are capable of monitoring and controlling (1) the materials production process; (2) materials stability during fabrication, transport, and storage; and (3) the amount and rate of degradation during the postfabrication in-service life for both components and structures. Real-time process monitoring for more efficient real-time process control, improved product quality, and increased reliability has become a practical reality. Simply put, intelligent manufacturing is impossible without integrating modern NDE into the production system. What was once a nonsystematic collection of after-the-fact techniques has become a powerful set of tools which, if used throughout the manufacturing process, can help to transform that process to make products better and more competitive. *See* MATERIALS SCIENCE AND ENGINEERING.

Mature Technologies

A number of NDE technologies have been in use for some time.

Visual inspection. This is the oldest and most versatile NDE tool. In visual inspection, a worker examines a material using only eyesight. Visual inspection relies heavily on common sense and is vulnerable to erratic results because of human failure to comprehend the significance of what is seen.

Liquid (or dye) penetrant. This augmented visual method uses brightly colored liquid dye to penetrate and remain in very fine surface cracks after the surface is cleaned of residual dye.

Magnetic particle. This augmented visual method requires that a magnetic field be generated inside a ferromagnetic test object. Flux leakage occurs where there are discontinuities on the surface. Magnetic particles (dry powder or a liquid suspension) are captured at the leakage location and can be readily seen with proper illumination.

Eddy current. This method uses a probe held close to the surface of a conducting test object. The probe is a coil of fine wire on a flat spool energized by an ac oscillator, creating an ac magnetic field

orthogonal to the spool. The oscillating magnetic field penetrating the test object induces a circular oscillating image current in the plane of the surface. If the coil is moved across the surface, discontinuities in or near the surface perturb the flow of the induced current which can be detected as a change in the impedance that the original coil presents to the ac oscillator. *See* EDDY CURRENT.

X-rays. These provide a varied and powerful insight into material, but they are somewhat limited for use in the field. Radiography is similar to the familar chest x-ray, except that it is used to image materials rather than people. Computed tomography is the nonpeople version of medical CAT scanning. Flash or pulse x-ray systems can provide x-ray images of dynamic events, much as a camera with a flash can "freeze" an athlete in midair or a horse at the finish line. (Some of these systems are suitcase size and battery-operated.) X-ray diffraction can provide information about crystallographic defects in such things as single-crystal blades for jet engines (and ground-based gas turbines) and semiconductor single crystals. **Figure 1** shows a diffraction image of a nickel-based alloy single crystal blade. Real-time radioscopy and real-time x-ray diffraction use charge-coupled-device (CCD) cameras, image intensifiers, and phosphor screens to record x-ray images at video rates and faster. The images go directly to computers where digital image processing can enhance their usefulness. *See* CHARGE-COUPLED DEVICES; RADIOGRAPHY; X-RAY DIFFRACTION.

Acoustic emission. This technique typically uses a broadband piezoelectric transducer to listen for acoustic noise. At first glance, acoustic emission (AE) appears as an attractive technology because its sensors may be configured to listen all the time, thereby capturing an acoustic emission event no matter when it occurs. In contrast, ultrasonic methods acquire information only when active probing is taking place. Much effort has been expended trying to use acoustic emission to detect plastic deformation and crack growth in structures as a means of monitoring system health. Except for some special cases, attempts to make broader use of acoustic emission have generally been disappointing. Crack growth certainly generates acoustic emissions, but so do a lot of other mechanisms. A highway bridge loaded with live traffic, for example, generates much acoustic noise from many different but benign sources. The difficulty is separating the signature, if any, due to crack growth. Before acoustic emission technology can be used to its full potential, identification of source mechanisms and the associated structural alteration of materials and structures must be better understood. *See* ACOUSTIC EMISSION; PIEZOELECTRICITY; PLASTIC DEFORMATION OF METAL.

Fig. 1. X-ray diffraction images from a turbine blade.

Thermography. This technique uses a real-time "infrared camera," much like a home camcorder, except that it forms images using infrared photons instead of visible ones. Imaging rates can vary from a few frames per second to hundreds of frames per second. The choice of camera depends on trade-offs between requirements for frame rate, brightness sensitivity, and spatial resolution. A representative commercial infrared camera can easily image the thermal trail written on paper at room temperature by a moving finger. A breakthrough has been the replacement of camera sensors that see only one pixel at a time with sensors that have a full field of pixels detected simultaneously. A consequence of this development is that an entire infrared image can be plucked out of a sequence, like a freeze-frame, at a predetermined time after an event. If the image has random noise (like "snow" in TV images), and if the event can be repeated over and over, then images can be plucked out over and over after exactly the same time delay. Adding such identical images together will increase the clarity of the image and reduce the snow.

Contact ultrasonics. This technique is the workhorse of traditional and mature NDE technology. It uses a transducer held in contact with a test object to launch ultrasonic pulses and receive echoes. These pulses are mechanical vibrational waves, which commonly contain frequencies in the megahertz range. The transducers are usually piezoelectric elements tuned to a narrow band (their resonant frequency), and often require a couplant (such as special greases) to ensure efficient transmission of the vibrations into the test object. In some cases a water path is used to couple the transducer to the test object. In looking for internal flaws, the same transducer is often used to first launch a pulse and subsequently receive its echoes. Comparing echoes from the other side of the object with echoes that return earlier determines the location of flaws. It is often advantageous to use a second transducer to function solely as a receiver. Velocity and attenuation of the waves can be determined, and in turn yield information on material properties such as elastic moduli, grain size, and porosity. *See* TRANSDUCER.

Emerging Technologies

Many noncontact measurements have been developed that enhance the mature technologies.

Noncontact ultrasonic transducers. These involve laser ultrasonics, electromagnetic acoustic transducers, and air- or gas-coupled transducers.

Laser ultrasonics affords the opportunity to make truly noncontact measurements in materials at elevated temperatures, in hostile environments, and in locations geometrically difficult to reach. One major drawback of laser-based ultrasonic systems is that laser interferometers are generally poor receivers, particularly for shear-wave displacements. Improving received signal levels by using larger generation pulses is limited by the need to leave the surface of the material undamaged. Employing temporal or spatial modulation of the generation process has been effective in gaining increases in the signal-to-noise ratios of detected signals. *See* LASER; ULTRASONICS.

Acoustic vibrations in a conducting material in the presence of a dc magnetic field induce small ac currents. Electromagnetic acoustic transducers (EMATs) both provide the dc magnetic field and sense the induced currents. They also can operate in the inverse fashion, but EMATs are much better at detecting than generating. Although noncontact, EMAT-based systems are not truly remote because their efficiency rapidly decreases with the lift-off distance between the transducer face and the test object surface. Typically lift-off needs to be no larger than perhaps a millimeter (0.04 in.). Another drawback is that EMAT-based systems work only on electrically conducting materials.

Air- or gas-coupled transducers operate with much larger lift-off distances than EMATs. Their principal advantages are that they do not require mirror surface finishes (as do laser interferometers), nor do they require the test object to be conductive. The main challenge in their implementation comes from the large acoustic impedance mismatch between air and the solid objects under study; however, existing systems have succeeded in overcoming this challenge effectively. For example, a C scan system has been developed that uses air coupling for both generation and detection (**Fig. 2**).

Hybrid laser ultrasonics. Since lasers are very effective at generating "clean" ultrasonic excitation to satisfy a variety of needs, hybrid systems generally use laser generation with something other than a laser interferometer for ultrasonic detection. For example, EMATs can make excellent detectors in place of lasers. Air-coupled transducers can also provide many advantages over laser interferometers in particular applications.

Historically, making engineered monolithic structures of composite materials has been a highly labor-intensive manufacturing operation because of the tedious process of hand-layup of the prepreg material. In addition, curing the resulting structure required autoclave processing of the entire structure all at once. These practices have contributed to high-cost, time-consuming, labor-intensive procedures, and a limit on structural size and shape to fit into an autoclave. Quality control and reproducibility of the hand-layup depended on the skill and dedication of the specialists doing the layup. To address these processing problems, the industry evolved a technology using a robot to perform precision placement of prepreg tape and to consolidate and cure the tape as it is laid (eliminating the need for an autoclave). The problem has been the need for a closed-loop real-time capability to provide quality inspection for voids, porosity, and state of cure as the tape goes down. A hybrid ultrasonic remote inspection system has been developed for this purpose (**Fig. 3**). It provides real-time data "on the fly" as the tape is placed in position by robot, consolidated by roller, and cured by localized thermal deposition. Narrow-band generation of surface waves is accomplished by a spatial modulation procedure that employs a periodic transmission mask. Detection in this system is by air-coupled transduction. An improvement in this same system has been the deployment of a fiber-optic light delivery system. The fiber-optic umbilical is able to deliver pulsed light from a neodynium:yttrium-aluminum-garnet (Nd:YAG) laser at an average energy as high as 55 millijoules per pulse for a pulse duration as short as 10 nanoseconds, at a 20-hertz repetition rate, without signs of fiber damage. The center frequency of the generated toneburst signal can be controlled easily by changing the periodicity

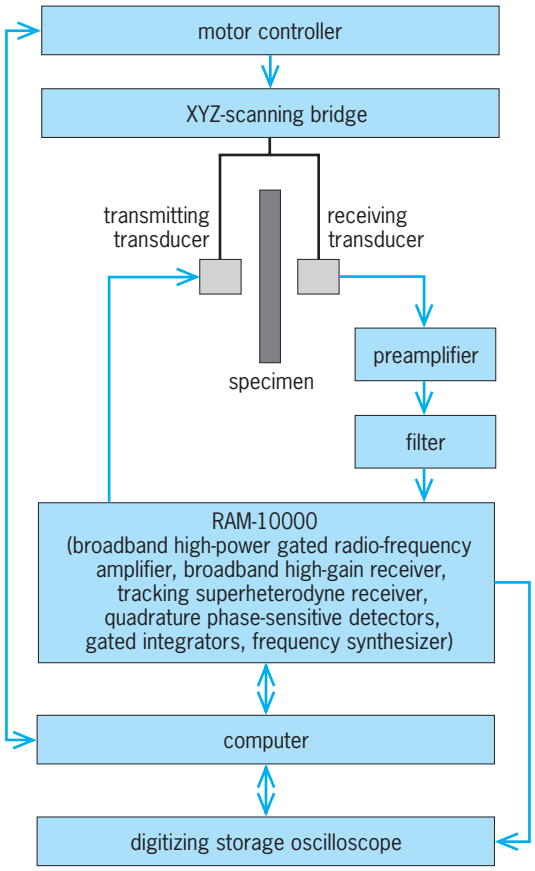

Fig. 2. Air-coupled ultrasonic C scan system.

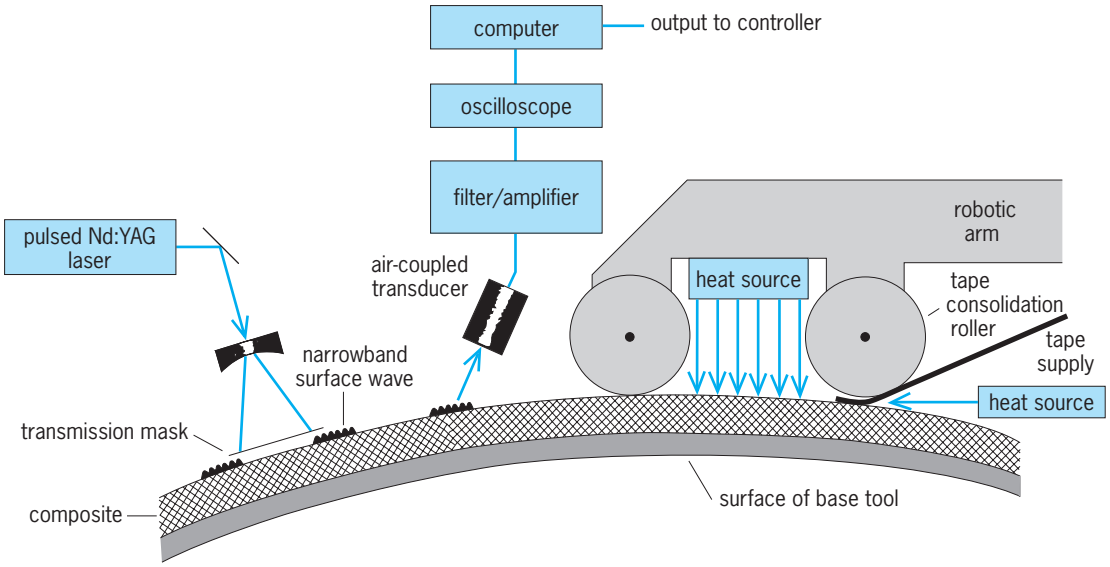

Fig. 3. Hybrid ultrasonic inspection for automated "on the fly" composite fabrication.

of the transmission mask. This step, in turn, allows control of the surface-wave penetration depth, which is determined by the wave frequency content. In addition, compared to conventional single-pulse illumination, the spatial modulation technique allows more light to illuminate the test surface while retaining thermoelastic generation conditions, thus increasing the signal detectability without material damage and ablation. *See* COMPOSITE MATERIAL; ROBOTICS.

Nonlinear ultrasonics. This has become a promising field. Generation of ultrasound in a selected range of frequencies with detection in a different range can be sensitive to changes in materials not detectable with the common practice of generating and detecting in the same range. The coupling into other ranges is a consequence of higher-order nonlinearities in the elastic moduli (commonly presumed linear) which govern the propagation of ultrasound.

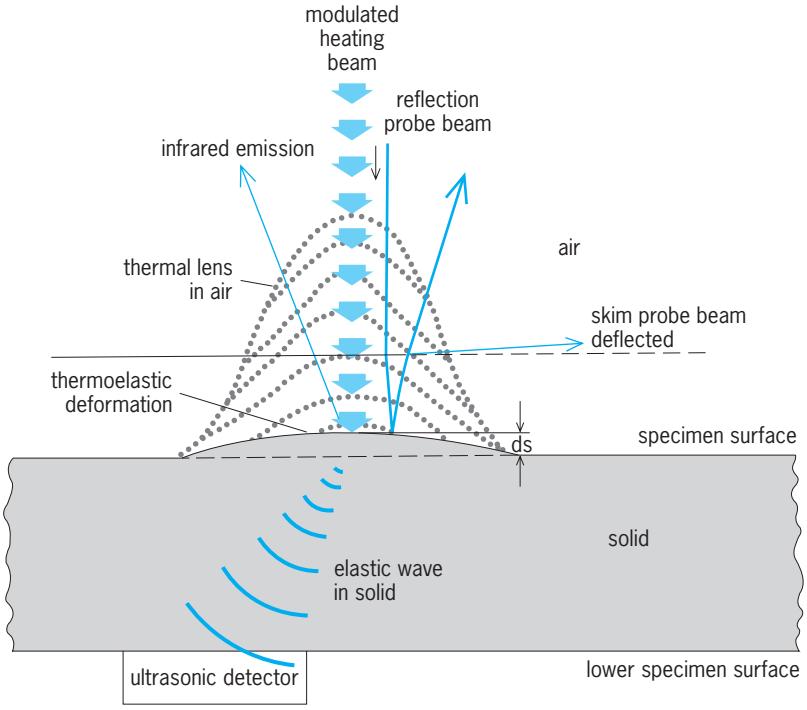

Fig. 4. Principles of thermal wave imaging. Skim probe beam detects changes in the refractive index of air; reflection probe beam detects resultant surface displacements; infrared emission is measured by thermography; and ultrasonic detector examines the interior of the specimen.

Thermal wave imaging. This uses a main laser beam to scan the surface of the object to be examined. The amount of laser energy absorbed by the surface is directly dependent on the surface's thermal properties. Thermal expansion resulting from this localized energy absorption creates very slight vertical surface displacements which vary depending on the amount of laser energy absorbed (**Fig. 4**). In addition, the air close above the surface is heated slightly, changing its density, and therefore its index of refraction. A secondary "probe" laser beam can detect either the changes in the refractive index of the air (skim probe) or the resultant surface displacements (reflection probe). The infrared emission from the surface can be imaged by thermography as described previously, or an ultrasonic detector can be used to examine the interior using transmitted elastic waves generated by the sudden thermoelastic expansions from the heating beam pulses. If these various means of detecting are scanned systematically over the surface, high-resolution images can be generated that reveal very slight differences in the rate at which heat flows away from the surface; or they can provide ultrasonic image data about the internal structure. These differences map either inhomogeneities in the thermal properties of the material immediately below the surface, or subsurface structural flaws such as voids, cracks, or delaminations. An example of this technique is detection of areas of lack of adhesion of protective turbine blade coatings used on nickel-based superalloy blades for aviation gas turbines (jet engines).

Speckle-based interferometry. Electronic speckle pattern interferometry is a noncontact, full-field

optical technique for high-sensitivity measurement of extremely small displacements in an object's surface. "Speckle" refers to the grainy appearance of an optically rough surface illuminated by a laser. It is actually a very complex interference pattern resulting from diffuse scattering from the rough surface. The pattern is unique for a given surface, like a fingerprint. If the surface is deformed, the pattern changes. By mathematically comparing two speckled images of the surface before and after deformation, a fringe pattern depicting the displacement distribution is generated. "Full-field" refers to the fact that an area of the test specimen is imaged, not just one point. The key concept behind this approach is that flaws or damage in the test object create local perturbations of the surface displacement profile when the object is loaded (for example, stretched). The speckle comparison patterns serve to indicate the presence, location, and other details of the flaws. A variation, called electronic shearography, measures displacement gradients instead of absolute displacements. Shearography needs only a very simple optical setup since it is based on a self-referencing coherent optical system. Consequently, vibration isolation and laser coherence length considerations may be satisfied more easily than for other coherent optical techniques such as holographic interferometry. High-sensitivity, high-resolution, ease of use, and full-field capabilities make speckle-based techniques a valuable tool for industrial nondestructive testing. *See* INTERFEROMETRY; NONLINEAR OPTICAL DEVICES.

Microwave techniques. Development activities are under way in such diverse applications as ground-penetrating radar for land-mine detection, locating delaminations in highway bridge decks, and monitoring the curing process in polymers. *See* MICROWAVE.

John M. Winter, Jr.

Bibliography. H. L. M. dos Reis, *Nondestructive Testing and Evaluation for Manufacturing and Construction*, Hemisphere Publishing, New York, 1990; J. Holbrook and J. Bussiere (eds.), *Nondestructive Monitoring of Materials Properties*, Materials Research Society, Pittsburgh, 1989; J. Krautkramer and H. Krautkramer, *Ultrasonic Testing of Materials*, 3d ed., Springer-Verlag, New York, 1983; P. K. Liaw et al. (eds.), *Nondestructive Evaluation and Materials Properties III*, The Minerals, Metals, and Materials Society, Warrendale, PA, 1996.

Noneuclidean geometry

A system of geometry based upon a set of axioms different from those of the euclidean geometry based on the three-dimensional space of experience. Noneuclidean geometries, especially riemannian geometry, are useful in mathematical physics. This article will describe some of the basic concepts and axioms of noneuclidean geometries and then draw some comparisons with euclidean concepts. *See* DIFFERENTIAL GEOMETRY; EUCLIDEAN GEOMETRY; PROJECTIVE GEOMETRY; RIEMANNIAN GEOMETRY.

The famous names that are related to hyperbolic (noneuclidean) geometry are J. Bolyai and N. I. Lobachevski. In addition, A. Cayley, E. Beltrami, K. F. Gauss, and G. F. B. Riemann have made many outstanding contributions to the subject of noneuclidean geometry.

Projective space. Let K denote a given field of elements, called the ground field K. In actual applications, K will be simply the field of real numbers. Consider the class Σ of ordered sets of $(n+1)$-tuples $x = (x^0, x^1, \ldots, x^n)$, where each x^i is an element of the ground field K. The index i is not an exponent but is merely a superscript. In this article, both superscripts and subscripts will be used. If $x = (x^0, x^1, \ldots, x^n)$ and $y = (y^0, y^1, \ldots, y^n)$ are two elements in the class Σ, then the sum $(x + y)$ is defined to be the element $x + y = (x^0 + y^0, x^1 + y^1, \ldots, x^n + y^n)$ in Σ. Also, if $x = (x^0, x^1, \ldots, x^n)$ is an element of Σ and if ρ is an element of the ground field K, then the scalar product ρx is defined to be the element $\rho x = (\rho x^0, \rho x^1, \ldots, \rho x^n)$ in Σ.

If $x_1, x_2, \ldots, x_m$ are m elements in Σ and if $\rho_1, \rho_2, \ldots, \rho_m$ is a set of m elements from the ground field K, then the element $x = \rho_1 x_1 + \rho_2 x_2 + \cdots + \rho_m x_m$ of Σ is called a linear combination of the elements $x_1, x_2, \ldots, x_m$ of Σ.

The class Σ^1 is a proper subset of Σ which is composed of all ordered sets of $(n+1)$-tuples $x = (x^0, x^1, \ldots, x^n)$ of Σ such that at least one x^i of the $(n+1)$ elements $x^0, x^1, \ldots, x^n$ of the ground field K is not zero. Two elements $x = (x^0, x^1, \ldots, x^n)$ and $y = (y^0, y^1, \ldots, y^n)$, of the class Σ^1, are said to be equivalent if, and only if, there exists an element $\rho \neq 0$, of the ground field K, such that $y = \rho x$. This establishes an equivalence relation among the elements $x = (x^0, x^1, \ldots, x^n)$ of the class Σ^1. The equivalence classes in Σ^1, formed relative to this equivalence relation, are called points P. The collection of all these points P forms a projective space S_n of n dimensions over the ground field K.

A point P in S_n is a class in Σ^1 of all ordered sets of $(n+1)$-tuples equivalent to a given set of $(n+1)$-tuples, namely, $x = (x^0, x^1, \ldots, x^n)$. Each of these ordered sets of $(n+1)$-tuples is in the equivalence class called the point P and is said to be a set of homogeneous point coordinates for the point P. Thus, if $x = (x^0, x^1, \ldots, x^n)$ is a set of homogeneous point coordinates of a point P in S_n, any other set of homogeneous point coordinates for the same point P is $\rho x = (\rho x^0, \rho x^1, \ldots, \rho x^n)$, where $\rho \neq 0$ is an arbitrary element of the ground field K.

A collection of $(p+1)$ points $x_0, x_1, \ldots, x_p$ in S_n is said to be linearly dependent if, and only if, at least one of them is a linear combination of the remaining ones such that at least one of the scalar coefficients is not zero. Otherwise they are said to be linearly independent. Clearly, if $(p+1)$ points $x_0, x_1, \ldots, x_p$ in S_n are linearly independent, then $0 \leqq p \leqq n$.

Let a collection of $(p+1)$ linearly independent points $x_0, x_1, \ldots, x_p$ in S_n be given. Then $0 \leqq p \leqq n$. The S_p determined by them is composed of all points x in S_n such that Eq. (1) holds, where at least one of

$$x = y^0 x_0 + y^1 x_1 + \cdots + y^p x_p \qquad (1)$$

the scalar multiples $y^0, y^1, \ldots, y^p$ is not zero. This S_p is a projective space of p dimensions over the ground field K, and is a projective subspace of S_n. A set of homogeneous point coordinates for an arbitrary point P in this S_p is $(y^0, y^1, \ldots, y^p)$.

Define S_{-1} to be a vacuous collection of points P in S_n. Such an S_{-1} is called a projective subspace of S_n of dimension -1. Evidently every S_0 is a point P and conversely. A line is an S_1, a plane is an S_2, a three-space is an S_3, and a p-space is an S_p. In particular, a hyperplane is an S_{n-1}.

Dual coordinates. An S_p is a locus of points x whose homogeneous point coordinates $x = (x^0, x^1, \ldots, x^n)$ satisfy a system of $(n - p)$ linear homogeneous equations (2), such that the rank of the coefficient matrix (u_{ij}) is $(n - p)$.

$$\sum_{j=0}^{n} u_{ij} x^j = 0 \qquad (i = 1, 2, \ldots, n - p) \tag{2}$$

In particular, a hyperplane S_{n-1} is a locus of points x whose homogeneous point coordinates $x = (x^0, x^1, \ldots, x^n)$ satisfy a single linear homogeneous equation (3), where at least one of the $(n + 1)$ elements $u_0, u_1, \ldots, u_n$, of the ground field K, is not zero.

$$\begin{aligned} ux &= \sum_{i=0}^{n} u_i x^i \\ &= u_0 x^0 + u_1 x^1 + \cdots + u_n x^n = 0 \end{aligned} \tag{3}$$

Because an ordered set of $(n + 1)$-tuples $u = (u_0, u_1, \ldots, u_n)$ of this type uniquely defines an S_{n-1}, it is called a set of homogeneous hyperplane coordinates, for the particular S_{n-1}. If $u = (u_0, u_1, \ldots, u_n)$ is a set of homogeneous hyperplane coordinates of an S_{n-1}, any other set of homogeneous hyperplane coordinates for the same S_{n-1} is $\sigma u = (\sigma u_0, \sigma u_1, \ldots, \sigma u_n)$, where $\sigma \neq 0$ is an arbitrary element of the ground field K.

In the projective space S_n of n dimensions, a point P and a hyperplane S_{n-1} are termed dual objects. Thus, homogeneous point coordinates $x = (x^0, x^1, \ldots, x^n)$ and homogeneous hyperplane coordinates $u = (u_0, u_1, \ldots, u_n)$ are said to be dual systems of coordinates.

A point P with the homogeneous point coordinates $x = (x^0, x^1, \ldots, x^n)$ is on the hyperplane S_{n-1} with the homogeneous hyperplane coordinates $u = (u_0, u_1, \ldots, u_n)$ if, and only if, the condition of Eq. (3) is satisfied.

By mathematical induction, it follows that the dual of a projective subspace S_p of p dimensions is a projective subspace S_{n-p-1} of $(n - p - 1)$ dimensions. For example, the dual of S_{-1} is S_n, the dual of an S_0 is an S_{n-1}, and the dual of an S_1 is an S_{n-2}.

The projective subspace S_p is said to be contained in the projective subspace S_q, or the projective subspace S_q contains the projective subspace S_p, if, and only if, every point of S_p is a point of S_q. In particular, S_{-1} is contained in every S_q, and S_n contains every S_p. This relation is written as $S_p \subset S_q$, or $S_q \supset S_p$.

The dual of the relation $S_p \subset S_q$ is the dual relation $S_{n-p-1} \supset S_{n-q-1}$.

Principle of duality. For every theorem concerning the S_p,s and the relations $S_p \subset S_q$, there is obtained a dual theorem wherein each S_p is replaced by the dual object S_{n-p-1}, and each relation $S_p \subset S_q$ is replaced by the dual relation $S_{n-p-1} \supset S_{n-q-1}$.

If S_p and S_q are any two projective subspaces of the projective space S_n, the largest projective subspace contained in both S_p and S_q is denoted by $S_p \cap S_q$, and the smallest projective subspace containing both S_p and S_q is denoted by $S_p \cup S_q$. Clearly, $S_p \cap S_q$ is the intersection or meet, and $S_p \cup S_q$ is the union or join of S_p and S_q. *See* SET THEORY.

The two relations $S_p \cap S_q$ and $S_p \cup S_q$ are dual. Concerning the inclusion relation $S_p \subset S_q$, the projective subspaces of the projective space S_n form a complemented modular lattice.

Finally, let dim (S_p) denote the dimension p of the projective subspace S_p of the projective space S_n. The dimension theorem may be stated in the form shown in Eq. (4).

$$\begin{aligned} &\dim (S_p \cap S_q) + \dim (S_p \cup S_q) \\ &\qquad = \dim (S_p) + \dim (S_q) \end{aligned} \tag{4}$$

Projective group. A collineation T of the projective space S_n is a one-to-one correspondence between the points of S_n, of the form shown in Eq. (5), where

$$\rho \bar{x}^i = \sum_{j=0}^{n} a^i_j x^j \qquad (i = 0, 1, \ldots, n) \tag{5}$$

the rank of the coefficient matrix (a^i_j) is $(n + 1)$, and $\rho \neq 0$ is an arbitrary constant of proportionality. Of course, all elements are from the original ground field K.

A collineation T is not only a one-to-one correspondence between the points S_0 of S_n, but also a one-to-one correspondence between the S_p's of S_n; that is, under T, any S_p is carried into one, and only one, $\overline{S_p}$, and conversely.

The set of collineations T of S_n forms the collineation group G of S_n. It is composed of $n(n + 2)$ essential parameters.

The fundamental theorem of projective geometry may be stated in the following form: There is a collineation T which carries a given set of $(p + 1)$ linearly independent points into a prescribed set of $(p + 1)$ linearly independent points. Moreover, if $p = n$, then T is uniquely determined.

A correlation Γ of the projective space S_n is a one-to-one correspondence between the points S_0 and the hyperplanes S_{n-1} of S_n, of the form shown in Eq. (6), where the rank of the coefficient matrix (b_{ij})

$$\sigma \bar{u}_i = \sum_{j=0}^{n} b_{ij} x^i \qquad (i = 0, 1, \ldots, n) \tag{6}$$

is $(n + 1)$, and $\sigma \neq 0$ is an arbitrary constant of proportionality.

The set C of correlations Γ of S_n is composed of $n(n + 2)$ essential parameters. In general, C is not a group.

The total projective group G^* of the projective space S_n is composed of collineations T and correlations Γ. It is a mixed group G^* of $n(n+2)$ essential parameters. Obviously, the collineation group G is a subgroup of the total projective group G^*.

Projective geometry consists of the qualitative and quantitative invariants of the total projective group G^*.

Cross ratio. Let S_{r-1} and S_{r+1} be two fixed projective subspaces of S_n of dimensions $(r-1)$ and $(r+1)$, respectively, such that S_{r-1} is contained in S_{r+1}. Evidently $0 \leqq r \leqq n-1$. A pencil is composed of all the S_r,s that contain the given S_{r-1} and that are contained in the given S_{r+1}.

For example, where $r = 0$, one obtains a pencil of points, all of which are in a fixed line S_1. Similarly when $r = n-1$, there is defined a pencil of hyperplanes S_{n-1}, all of which contain a fixed projective subspace S_{n-2} of $(n-2)$ dimensions.

A pencil of elements P is a projective space of one dimension. Therefore, the elements P of a pencil are defined by a system of homogeneous coordinates (ρ,τ), where at least one of the elements ρ and τ of the ground field K is not zero. If it is assumed that τ is always one, then ρ is said to be the nonhomogeneous coordinate of the element P which is not the element at infinity (1,0). The element of infinity (1,0) in this system of coordinates is denoted by the symbol ∞.

If (ρ_1,τ_1), (ρ_2,τ_2), $(\rho_3\tau_3)$, (ρ_4,τ_4) are the homogeneous coordinates of four distinct elements P_1, P_2, P_3, P_4, their cross ratio $\mathfrak{R}$ is defined to be the numerical invariant given in Eq. (7). This is the fundamental projective invariant. In nonhomogeneous coordinates, this is Eq. (8).

$$\mathfrak{R}(P_1P_2, P_3P_4) = \frac{(\rho_3\tau_1 - \rho_1\tau_3)(\rho_2\tau_4 - \rho_4\tau_2)}{(\rho_2\tau_3 - \rho_3\tau_2)(\rho_4\tau_1 - \rho_1\tau_4)} \tag{7}$$

$$\mathfrak{R}(P_1P_2, P_3P_4) = \frac{(\rho_3 - \rho_1)(\rho_2 - \rho_4)}{(\rho_2 - \rho_3)(\rho_4 - \rho_1)} \tag{8}$$

In particular, Eqs. (9) and (10) hold.

$$\mathfrak{R}(\infty P_3, P_1P_2) = \frac{\rho_3 - \rho_2}{\rho_3 - \rho_1} \tag{9}$$

$$\mathfrak{R}(\infty 0, P_1P_2) = \frac{\rho_2}{\rho_1} \tag{10}$$

Of importance in projective geometry is the concept of a harmonic set of elements of a pencil provided that the ground field K is not of characteristic two. Four elements P_1, P_2, P_3, P_4 of a pencil are said to form a harmonic set of elements if, and only if, $\mathfrak{R}(P_1P_2, P_3P_4) = -1$.

Quadrics. Consider a correspondence Γ of S_n in which every point P is converted into a single hyperplane S_{n-1}, or S_n. It is assumed that Γ carries every line S_1 into a single S_{n-2}, or S_{n-1}, or S_n. Finally, it is supposed that if Γ carries a point P into an S_p and if $\overline{P}$ is any point of this S_p, then Γ converts the point $\overline{P}$ into an S_q which passes through the original point P. Such a correspondence Γ is called a polarity Γ.

In homogeneous coordinates, a polarity Γ is given by Eq. (11), where the matrix (g_{ij}) is symmetric, that

$$\sigma\bar{u}_i = \sum_{j=0}^{n} g_{ij}x^j \qquad (i = 0, 1, \ldots, n) \tag{11}$$

is, $g_{ij} = g_{ji}$, and where its rank is $(r+1)$ for which $0 \leqq r \leqq n$. If $0 \leqq r < n$, the polarity Γ is said to be singular. Otherwise Γ is said to be nonsingular.

In the homogeneous point coordinates, the polarity Γ is given by Eq. (12).

$$\sum_{i,j=0}^{n} g_{ij}x^i\bar{x}^j = 0 \tag{12}$$

The dual of a polarity Γ is also a polarity Γ^*. Such a polarity Γ^* is given in homogeneous coordinates by Eq. (13), where the matrix (g^{ij}) is symmetric, that is,

$$\rho\bar{x}^i = \sum_{j=0}^{n} g^{ij}u_j \qquad (i = 0, 1, \ldots, n) \tag{13}$$

$g^{ij} = g^{ji}$, and where its rank is $(r+1)$ for which $0 \leqq r \leqq n$.

In homogeneous hyperplane coordinates, the polarity Γ^* is given by Eq. (14).

$$\sum_{i,j=0}^{n} g^{ij}\bar{u}_i\bar{u}_j = 0 \tag{14}$$

If the polarity Γ is nonsingular, then the dual Γ^* is the original polarity Γ. In that event, the two matrices (g_{ij}) and (g^{ij}) can be considered to be inverse matrices. Thus a nondegenerate polarity can be given by Eqs. (11), (12), (13), or (14).

A hyperplane element (P,π) is composed of a point P and a hyperplane π of dimension $(n-1)$ which passes through the point P. A quadric Q is a locus of hyperplane elements (P,π) such that under a given polarity Γ or Γ^*, the point P is transformed into the hyperplane π, or the hyperplane π is carried into the point P.

The polarity Γ or Γ^* is said to be a polarity relative to the corresponding quadric Q.

In homogeneous point coordinates, the equation of a quadric Q is (15).

$$\sum_{i,j=0}^{n} g_{ij}x^ix^j = 0 \tag{15}$$

In homogeneous hyperplane coordinates, the equation of a quadric Q is (16).

$$\sum_{i,j=0}^{n} g^{ij}u_iu_j = 0 \tag{16}$$

Euclidean and noneuclidean geometries. Consider a real projective space S_n. That is, S_n is defined over the real number system K. Sometimes it is convenient to consider that S_n is immersed in a complex projective space S_n^*, which is defined over the complex number system K^*.

Let $k \neq 0$ be either a positive real number or infinite (that is, $1/k = 0$), or else a pure imaginary number of the form $k = il$, where $i^2 = -1$ and l

is a positive real number. Consider the fundamental quadric Σ whose homogeneous point equation is (17), where the superscripts exterior to the parentheses denote exponents. The homogeneous hyperplane equation of this fundamental quadratic Σ is (18).

$$f(x, x) = (kx^0)^2 + (x^1)^2 + (x^2)^2 + \cdots + (x^n)^2 = 0 \quad (17)$$

$$F(u, u) = \left(\frac{u_0}{k}\right)^2 + u_1{}^2 + u_2{}^2 + \cdots + u_k{}^2 = 0 \quad (18)$$

The set of all collineations T of S_n which carry this fundamental quadric Σ into itself is a group $G(k)$ of $n(n+1)/2$ essential parameters. When k is a positive real number, this is the elliptic group G_E of elliptic geometry. When $1/k = 0$, this is the euclidean group G_P of euclidean geometry. Finally when $k = il$ where $i^2 = -1$ and l is a positive real number, this is the hyperbolic group G_H of hyperbolic geometry.

The study of the qualitative and quantitative invariants of these three groups G_E, G_P, and G_H constitutes the three subjects of elliptic, euclidean, and hyperbolic geometries. The elliptic and hyperbolic geometries are usually referred to as the noneuclidean geometries.

In elliptic geometry, two distinct lines contained in a single plane always meet in a single point. Therefore, if L is a fixed line and if P is a given point not in L, there cannot be a line M passing through this point P which is parallel to the given line L.

In euclidean geometry, the ideal hyperplane π_∞ or the hyperplane π_∞ at infinity is the one whose point equation is $x^0 = 0$, or whose hyperplane equations are $u_1 = 0, u_2 = 0, \ldots, u_n = 0$. The proper S_p's for $p = -1, 0, 1, \ldots, n$ are those which are not contained in the ideal hyperplane π_∞. The improper S_p's for $p = 0, 1, \ldots, n-1$ are those which are contained in the ideal hyperplane π_∞. In euclidean geometry, only proper S_p's are studied.

If a proper S_p and a distinct proper S_q intersect in an improper S_r, then S_p and S_q are said to be parallel. Thus, two distinct lines in euclidean space are said to be parallel if, and only if, they intersect in an ideal point.

From the preceding discussion, Euclid's fifth parallel postulate is an easy consequence. That is, if in euclidean space L is a fixed line and if P is a fixed point not on L, there is one, and only one, line M parallel to L and passing through P.

In hyperbolic geometry, the points and the tangent S_p's for $p = 0, 1, 2, \ldots, n-1$ of the fundamental quadric Σ given by Eq. (17) or (18) are said to be ordinary improper or ordinary infinite. The S_p's for $p = 0, 1, 2, \ldots, n-1$, which are in the exterior of this quadric Σ, are said to be ultraimproper or ultrainfinite. The proper points in hyperbolic geometry are those which are in the interior of this quadric Σ. The proper S_p's for $p = 1, 2, \ldots, n-1, n$ are those which contain proper points and are considered to be sets of these proper points.

If a proper S_p and a distinct proper S_q intersect in an ordinary improper S_r or in an ultraimproper S_r, then S_p and S_q are said to be ordinary parallel or ultraparallel.

In hyperbolic geometry, if L is a fixed proper line and if P is a given proper point not on this line L, then there are two distinct proper lines M_1 and M_2 passing through P which are ordinary lines parallel to L.

Also passing through P, there is an infinite number of proper lines M which are ultraparallel to L. These lines M belong to the flat pencil with vertex at P and determined by the lines M_1 and M_2.

Distance. Let P and Q be two distinct points given by the homogeneous point coordinates $x = (x^0, x^1, \ldots, x^n)$ and $y = (y^0, y^1, \ldots, y_n)$. In euclidean and hyperbolic geometries, it is understood that P and Q are proper points. The line L determined by the two points P and Q intersects the fundamental quadric Σ in two distinct points P_∞ and Q_∞. The distance $s = s(P,Q)$ between these two points P and Q is defined by formula (19). It is understood that $s = s(P,Q)$

$$s = s(P, Q) = \frac{k}{2i} \log \mathfrak{R}(PQ, P_\infty Q_\infty) \quad (19)$$

is a real nonnegative number. Also it is assumed in elliptic geometry that $0 \leqq s/k \leqq \pi$.

Let $f(x,y)$ be expressed by Eq. (20). The two points

$$f(x, y) = k^2 x^0 y^0 + x^1 y^1 + x^2 y^2 + \cdots + x^n y^n \quad (20)$$

P and Q are polar reciprocal or orthogonal relative to the fundamental quadric Σ if and only if $f(x,y) = 0$. The point P is on Σ if, and only if, $f(x,x) = 0$.

A point R whose homogeneous point coordinates are $z = (z^0, z^1, \ldots, z^n)$ is on the line L determined by the two points P and Q if, and only if, a number ρ exists such that Eqs. (21) hold.

$$\begin{gathered} z^0 = x^0 + \rho y^0 \\ z^1 = x^1 + \rho y^1, \ldots, z^n = x^n + \rho y^n \end{gathered} \quad (21)$$

This point R is on the fundamental quadric Σ whose point equation is Eq. (17) if, and only if, ρ satisfies the quadratic equation (22). Because the two

$$f(x, x) + 2\rho f(x, y) + \rho^2 f(y, y) = 0 \quad (22)$$

points P and Q are distinct, this will have two distinct roots ρ_1 and ρ_2. The two points P_∞ and Q_∞ are on the line L whose homogeneous point equations are Eqs. (21), corresponding to the two distinct roots ρ_1 and ρ_2.

In elliptic and euclidean geometries, these two points P_∞ and Q_∞ are conjugate-imaginary. In hyperbolic geometry, they are real.

From Eqs. (21) and (22), it is seen that Eq. (23) holds.

$$\mathfrak{R}(PQ, P_\infty Q_\infty) = \frac{\rho_1}{\rho_2} = \frac{-f(x, y) - \sqrt{f^2(x, y) - f(x, x) f(y, y)}}{-f(x, y) + \sqrt{f^2(x, y) - f(x, x) f(y, y)}} \quad (23)$$

Then, from Eq. (23), Eq. (24) applies.

$$R(PQ, P_\infty Q_\infty) = \frac{\left[f(x,y) + \sqrt{f^2(x,y) - f(x,x)f(y,y)}\right]^2}{f(x,x)f(y,y)} \tag{24}$$

Consequently the distance $s(P,Q)$ is given by formula (25).

$$s(P,Q) = \frac{k}{i}\log\frac{f(x,y) + \sqrt{f^2(x,y) - f(x,x)f(y,y)}}{\sqrt{f(x,x)}\sqrt{f(y,y)}} \tag{25}$$

Set $s = s(P,Q)$. Then s is given by Eqs. (26).

$$\begin{aligned} \cos\frac{s}{k} &= \frac{f(x,y)}{\sqrt{f(x,x)}\sqrt{f(y,y)}} \\ \sin\frac{s}{k} &= \frac{\sqrt{f^2(x,y) - f(x,x)f(y,y)}}{i\sqrt{f(x,x)}\sqrt{f(y,y)}} \end{aligned} \tag{26}$$

Of course, from Eq. (19), this distance $s = s(P,Q)$ is invariant under each of the groups G_E, G_P, G_H of elliptic, euclidean, and hyperbolic geometries.

By Eq. (20), Eqs. (26) may be written in the forms shown in Eqs. (27), in which $\epsilon = +1$ or $\epsilon = -1$,

$$\begin{aligned} \cos\frac{s}{k} &= \frac{\epsilon(k^2x^0y^0 + x^1y^1 + x^2y^2 + \cdots + x^ny^n)}{\sqrt{\left[(kx)^0(x^1)^2 + \cdots + (x^n)^2\right] \cdot \left[(ky^0)^2 + (y^1)^2 + \cdots + (y^n)^2\right]}} \\ \sin\frac{s}{k} &= \sqrt{\frac{\left[k^2\sum_{i=1}^{n}\begin{vmatrix} x^0 & x^i \\ y^0 & y^i \end{vmatrix}^2 + \frac{1}{2}\sum_{i,j=1}^{n}\begin{vmatrix} x^i & x^j \\ y^i & y^j \end{vmatrix}^2\right]}{\left[(kx^0)^2 + (x^1)^2 + \cdots + (x^n)^2\right] \cdot \left[(ky^0)^2 + (y^1)^2 + \cdots + (y^n)^2\right]}} \end{aligned} \tag{27}$$

according to whether the geometry is elliptic (including the euclidean case) or hyperbolic.

In hyperbolic geometry, $k = il$ where $i^2 = -1$ and l is a positive real number. Equations (27) can be written in the forms shown in Eqs. (28).

$$\begin{aligned} \cosh\frac{s}{l} &= \frac{l^2x^0y^0 - x^1y^1 - x^2y^2 - \cdots - x^ny^n}{\sqrt{[(lx^0)^2 - (x^1)^2 - \cdots - (x^n)^2] \cdot [(ly^0)^2 - (y^1)^2 - \cdots - (y^n)^2]}} \\ \sinh\frac{s}{l} &= \sqrt{\frac{l^2\sum_{i=1}^{n}\begin{vmatrix} x^0 & x^i \\ y^0 & y^i \end{vmatrix}^2 - \frac{1}{2}\sum_{i,j=1}^{n}\begin{vmatrix} x^i & x^j \\ y^i & y^j \end{vmatrix}^2}{[(lx^0)^2 - (x^1)^2 - \cdots - (x^n)^2] \cdot [(ly^0)^2 - (y^1)^2 - \cdots - (y^n)^2]}} \end{aligned} \tag{28}$$

Return to the general case of Eqs. (27). Let $x = (x^0, x^1, \ldots, x^n)$ denote affine coordinates in a euclidean space of $(n+1)$ dimensions. Define the special quadric Σ^* by Eq. (29). This is a central quadric

$$f(x,x) = (kx^0)^2 + (x^1)^2 + \cdots + (x^n)^2 = k^2 \tag{29}$$

Σ^*. Each of the two noneuclidean geometries can be visualized as the geometry on this quadric Σ^* in which diametrically opposite points are identified. In particular, for elliptic geometry, Σ^* can be considered a sphere.

On this quadric Σ^*, Eqs. (27) can be written in the forms labeled Eqs. (30). Here k is a positive real

$$\begin{aligned} \cos\frac{s}{k} &= x^0y^0 + \frac{1}{k^2}(x^1y^1 + x^2y^2 + \cdots + x^ny^n) \\ k\sin\frac{s}{k} &= \sqrt{\sum_{i=1}^{n}\begin{vmatrix} x^0 & x^i \\ y^0 & y^i \end{vmatrix}^2 + \frac{1}{2k^2}\sum_{i,j=1}^{n}\begin{vmatrix} x^i & x^j \\ y^i & y^j \end{vmatrix}^2} \end{aligned} \tag{30}$$

number and the distance s, such that $0 \leqq s/k\ \pi$, is given by the preceding equations.

Euclidean geometry can be considered to be the limiting case of either elliptic or hyperbolic geometry as k becomes infinite. In this case, one can always regard x^0 as unity. Then from Eqs. (30), the distance formula $s = s(P,Q)$ for euclidean geometry is given by Eq. (31). In this case, coordinates $x = (x^1, x^2, \ldots, x^n)$

$$s = s(P,Q) = \sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2 + \cdots + (x^n - y^n)^2} \tag{31}$$

of a point P are said to be rectangular or cartesian.

The final case is that of hyperbolic geometry. Here $k = il$, where $i^2 = -1$ and l is a positive real number. From Eqs. (30), the distance formula $s = s(P,Q)$ is given by Eqs. (32).

$$\begin{aligned} \cosh\frac{s}{l} &= x^0y^0 - \frac{1}{l^2}(x^1y^1 + x^2y^2 + \cdots + x^ny^n) \\ l\sinh\frac{s}{l} &= \sqrt{\sum_{i=1}^{n}\begin{vmatrix} x^0 & x^i \\ y^0 & y^i \end{vmatrix}^2 - \frac{1}{2l^2}\sum_{i,j=1}^{n}\begin{vmatrix} x^i & x^j \\ y^i & y^j \end{vmatrix}^2} \end{aligned} \tag{32}$$

In each of the three geometries it is assumed that the relationship $s = s(P,Q)$ is zero if, and only if, the two points P and Q are identical.

Each of the three geometries is an abstract metric space. That is,

i. $s(P,Q) \geqq 0$, and $s(P,Q) = 0$ if, and only if, $P = Q$
ii. $s(P,Q) = s(Q,P)$
iii. $s(P,Q) + s(Q,R) \geqq s(P,R)$

The condition (i) is that of positive definiteness; the condition (ii) is that of symmetry; the condition (iii) is that of the well-known triangular inequality.

Angle. By dualizing the concept of distance $s = s(P,Q)$, the concept of angle $\theta = \theta\ (\pi,\sigma)$ between two hyperplanes π and σ is obtained. Let π and σ be two distinct hyperplanes which are given by the homogeneous hyperplane coordinates $u = (u_0, u_1, \ldots, u_n)$ and $\upsilon = (\upsilon_0, \upsilon_1, \ldots, \upsilon_n)$. In euclidean and hyperbolic geometries, it is understood that π and σ are proper hyperplanes which determine a pencil λ. In this pencil λ there are two distinct hyperplanes π_∞ and σ_∞ which are tangent to the fundamental quadric Σ. The angle $\theta = \theta(\pi,\sigma)$, where $0 \leqq \theta \leqq \pi$, between these two hyperplanes π and σ is defined by formula (33). This angle $\theta = \theta(\pi,\sigma)$ is given by

Eqs. (34), in which $\epsilon = +1$ or $\epsilon = -1$ according

$$\theta = \theta(\pi, \sigma) = \frac{l}{2i} \log R(\pi\sigma, \pi_\infty\sigma_\infty) \tag{33}$$

$$\cos\theta = \frac{\epsilon\left(\dfrac{u_0 v_0}{k^2} + u_1 v_1 + u_2 v_2 + \cdots + u_n v_n\right)}{\sqrt{\left(\dfrac{{u_0}^2}{k^2} + {u_1}^2 + \cdots + {u_n}^2\right) \cdot \left(\dfrac{v_0^2}{k^2} + v_1^2 + \cdots + {v_n}^2\right)}}$$

$$\sin\theta = \sqrt{\frac{\epsilon\left[\dfrac{1}{k^2}\displaystyle\sum_{i=1}^{n}\begin{vmatrix} u_0 & u_i \\ v_0 & v_i \end{vmatrix} + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{n}\begin{vmatrix} u_i & u_j \\ v_i & v_j \end{vmatrix}^2\right]}{\left(\dfrac{{u_0}^2}{k^2} + {u_1}^2 + \cdots + {u_n}^2\right) \cdot \left(\dfrac{{v_0}^2}{k^2} + {v_1}^2 + \cdots + {v_n}^2\right)}} \tag{34}$$

to whether the geometry is elliptic (including the euclidean case) or hyperbolic.

It is evident that when $\epsilon = +1$ and k becomes infinite, the preceding formulas become those for the angle θ between the two hyperplanes π and σ in a euclidean space of n dimensions.

The two hyperplanes π and σ are said to be orthogonal if, and only if, $R(\pi\sigma, \pi_\infty\sigma_\infty) = -1$. Let σ_1 and σ_2 be two distinct hyperplanes which intersect in a line L. Then L is said to be orthogonal to a hyperplane π if σ_1 and σ_2 are each orthogonal to π. If L and M are two distinct lines which pass through a point P and if M is contained in a hyperplane π orthogonal to L, then L and M are said to be orthogonal.

If π is a fixed hyperplane and if P is a point not in π, then there is one and only one line L passing through the point P orthogonal to the hyperplane π.

Differential of arc length. If P and Q are two nearby points on the quadric Σ^* defined by the two sets of coordinates $(\bar{x}^0, \bar{x}^1, \ldots, \bar{x}^n)$ and $(\bar{x}^0 + d\bar{x}^0, \bar{x}^1 + d\bar{x}^1, \ldots, \bar{x}^n + d\bar{x})$, then by Eqs. (30), the square of the differential ds of arc length s between the points P and Q is Eq. (35).

$$ds^2 = \sum_{i=1}^{n}(\bar{x}^0 d\bar{x}^i - \bar{x}^i d\bar{x}_0)^2 + \frac{1}{2k^2}\sum_{i,j=1}^{n}(\bar{x}^i d\bar{x}^j - \bar{x}^j d\bar{x}^i)^2 \tag{35}$$

For the point P on this quadric Σ^* introduce a new set of coordinates $(x^1, x^2, \ldots, x_n)$ defined by Eq. (36), where it is understood that $\bar{x}^0 \neq -1$.

$$x^i = \frac{2\bar{x}^i}{1 + \bar{x}^0} \qquad (i = 1, 2, \ldots, n) \tag{36}$$

In this new set of coordinates, the quadric Σ^* is given by Eq. (37).

$$(x^1)^2 + (x^2)^2 + \cdots + (x^n)^2 + 4k^2 = \frac{8k^2}{1 + \bar{x}^0} \tag{37}$$

In the euclidean and noneuclidean geometries, the square of the differential ds of arc length s is given by Eq. (38).

$$ds^2 = \frac{(dx^1)^2 + (dx^2)^2 + \cdots + (dx^n)^2}{\left[1 + \dfrac{(x^1)^2 + (x^2)^2 + \cdots + (x^n)^2}{4k^2}\right]^2} \tag{38}$$

As k approaches infinity, the differential $d\sigma$ of arc length σ of euclidean geometry is obtained. If ds represents the differential of arc length s in elliptic or hyperbolic geometry, then Eq. (39) holds, in which the scale ρ is given by Eq. (40).

$$ds = \rho\, d\sigma \tag{39}$$

$$\rho = \frac{1}{1 + \dfrac{(x^1)^2 + (x^2)^2 + \cdots + (x^n)^2}{4k^2}} \tag{40}$$

Thus, each of the noneuclidean geometries may be visualized as a conformal image of euclidean geometry.

Each of these three geometries is a special case of riemannian geometry. Let the $g_{ij}(x)$ be $n(n+1)/2$ real functions which are continuous and possess continuous partial derivatives of as high an order as is necessary in a certain region of real n-dimensional space for which the coordinates of a point P are $x = (x^1, x^2, \ldots, x^n)$. The g_{ij}'s are symmetric; that is, $g_{ij} = g_{ji}$, for all $i, j = 1, 2, \ldots, n$. The quadratic form (41)

$$\sum_{i,j=1}^{n} g_{ij}\lambda^i\lambda^j \tag{41}$$

is assumed to be positive definite, that is, expression (41) is nonnegative; it is zero if and only if $\lambda^i = 0$, for all $i = 1, 2, \ldots, n$. Then the riemannian space V_n is one for which the square of the differential ds of arc length s between two nearby points P and Q is given by Eq. (42).

$$ds^2 = \sum_{i,j=1}^{n} g_{ij} dx^i dx^j \tag{42}$$

For this riemannian space V_n, the Christoffel symbols of the first kind are shown as Eq. (43), for $i, j,$

$$\Gamma_{ij;k} = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k}\right) \tag{43}$$

$k = 1, 2, \ldots, n$. The Christoffel symbols of the second kind are shown as Eq. (44), for $i, j, k = 1, 2,$

$$\Gamma_{jk}{}^i = \sum_{l=1}^{n} g^{il}\Gamma_{jk;l} \tag{44}$$

$\ldots, n$. Of course, (g^{ij}) is the inverse matrix of (g_{ij}). The Christoffel symbols of the second kind are also called the affine connections of V_n.

The vector geodesic curvature κ^i of a curve C in V_n is Eq. (45). A curve C of V_n is said to be a geodesic

$$\kappa^i = \frac{d^2x^i}{ds^2} + \sum_{j,k=1}^{n}\Gamma_{jk}{}^i\frac{dx^j}{ds}\frac{dx^k}{ds} \tag{45}$$

if, and only if, $\kappa^i = 0$ for $i = 1, 2, \ldots, n$ at every point P of C.

Each of the three spaces already discussed is a riemannian space. The geodesics of any one such space are the lines of the space.

A riemannian space of constant curvature is applicable to (that is, can be mapped isometrically into) elliptic, euclidean, or hyberbolic space. In this case, this constant curvature is equal to the gaussian curvature G. From Eq. (35) or (38), this constant gaussian curvature G is given by Eq. (46). For elliptic space,

$$G = \frac{1}{k^2} \tag{46}$$

G is of positive constant curvature. For euclidean space, G is identically zero. Finally for hyperbolic space, G is of negative constant curvature.

In each of these three geometries, consider a geodesic triangle. This is formed by three points P, Q, R, not all on one geodesic, and the three geodesics passing through every two of the three points P, Q, R. If A, B, C are the angles of this geodesic triangle, and if T denotes its area, then Eq. (47) holds.

$$A + B + C - \pi = \frac{T}{k^2} \tag{47}$$

Thus, according to whether the geometry is elliptic, euclidean, or hyperbolic, the sum of the angles of a geodesic triangle is greater than π, equal to π, or less than π.

Elliptic geometry of two dimensions may be represented upon a sphere in euclidean space of three dimensions. On the other hand, hyperbolic geometry of two dimensions can be depicted on a pseudosphere in euclidean space of three dimensions. The pseudosphere is obtained by revolving the tractrix about its asymptote. John De Cicco

Bibliography. H. S. M. Coxeter, *Introduction to Geometry*, 2d ed., 1969, paper 1989; J. Gray, *Ideas of Space: Euclidean, Non-Euclidean, and Relativistic*, 2d ed., 1989; M. J. Greenberg, *Euclidean and Non-Euclidean Geometries: Development and History*, 3d ed., 1995; E. Kasner and J. R. Newman, *Mathematics and the Imagination*, 1940, reprint 1989; G. E. Martin, *Foundations of Geometry and the Non-Euclidean Plane*, 1986; B. A. Rosenfeld, *The History of Non-Euclidean Geometry*, 1988; P. J. Ryan, *Euclidean and Non-Euclidean Geometry: An Analytic Approach*, 1986.

Nonlinear acoustics

The study of amplitude-dependent acoustical phenomena. The amplitude dependence is due to the nonlinear response of the medium in which the sound propagates, and not to the nonlinear behavior of the sound source. According to the linear theory of acoustics, increasing the level of a source by 10 dB results in precisely the same sound field as before, just 10 dB more intense. Linear theory also predicts that only frequency components radiated directly by the source can be present in the sound field. These principles do not hold in nonlinear acoustics. *See* AMPLITUDE (WAVE MOTION); LINEARITY; NONLINEAR PHYSICS.

The extent to which nonlinear acoustical effects are strong or even significant depends on the competing influences of energy loss, frequency dispersion, geometric spreading, and diffraction. When conditions are such that nonlinear effects are strong, acoustic signals may experience substantial waveform distortion and changes in frequency content as they propagate, and shock waves may be present. Nonlinear acoustical effects occur in gases, liquids, and solids, and they are observed over a broad range of frequencies. Shock waves present in sonic booms and thunder claps are in the audio frequency range. Principles of nonlinear acoustics form the basis for procedures at megahertz frequencies used in medical ultrasound and nondestructive evaluation of materials. Nonlinearity can also induce changes in nonfluctuating properties of the medium. These include acoustic streaming, which is the steady fluid flow produced by sound, and radiation pressure, which results in a steady force exerted by sound on its surroundings. *See* ACOUSTIC RADIATION PRESSURE; BIOMEDICAL ULTRASONICS; NONDESTRUCTIVE EVALUATION; SHOCK WAVE; SONIC BOOM; THUNDER.

The study of nonlinear acoustics dates back to the mideighteenth century, when L. Euler derived the first exact nonlinear wave equation for a plane wave in a gas. An exact implicit solution for a plane wave in a gas, revealing the nonlinear distortion of the acoustic waveform prior to shock formation, was derived by S. D. Poisson at the turn of the nineteenth century. The understanding of shock waves was provided by G. G. Stokes, W. J. M. Rankine, and P. H. Hugoniot in the latter half of the nineteenth century. Solutions for harmonic generation in sound waves, both before and after shock formation, were derived in the first half of the twentieth century. Foundations of the modern theory of nonlinear acoustics were developed in the second half of the twentieth century, motivated by engineering advances in the transmission and reception of high-intensity sound.

Waveform distortion. The principal feature that distinguishes nonlinear acoustics from nonlinear optics is that most acoustical media exhibit only weak dispersion, whereas media in which nonlinear optical effects arise exhibit strong dispersion. Dispersion is the dependence of propagation speed on frequency. In optical media, strong nonlinear wave interactions require that phase-matching conditions be satisfied, which can be accomplished only for several frequency components at one time. In contrast, all frequency components in a sound wave propagate at the same speed and are automatically phase-matched, which permits strong nonlinear interactions to occur among all components in the frequency spectrum. *See* NONLINEAR OPTICS.

Nonlinear coupling due to phase matching of the components in the frequency spectrum of a traveling sound wave manifests itself as waveform distortion that accumulates with distance. The main effect that opposes nonlinear distortion of the waveform is

energy absorption due to viscosity and heat conduction, and especially in air, relaxation due to molecular vibration. For sound radiated at a single frequency f, the parameter that characterizes the balance of nonlinearity and energy loss is the Gol'dberg number, given by Eq. (1). Here $\ell_a = 1/\alpha$ is the absorption

$$\Gamma = \ell_a/\bar{x} \tag{1}$$

length associated with the exponential attenuation coefficient α predicted for the wave amplitude by linear theory, and $\bar{x}$ is the shock formation distance, given by Eq. (2). The coefficient of nonlinearity β

$$\bar{x} = \frac{1}{\beta\epsilon k} \tag{2}$$

depends on the state equation for the propagation medium (discussed below), $\epsilon = p_0/(\rho_0 c_0^2)$ is the peak acoustic Mach number associated with the source waveform, p_0 is the corresponding peak sound pressure, ρ_0 is the ambient density of the medium, c_0 is the sound speed predicted by linear theory, and $k = 2\pi f/c_0$. *See* SOUND; SOUND ABSORPTION; SOUND PRESSURE.

For a plane wave, the criterion $\Gamma < 1$ indicates that energy loss due to absorption overwhelms the tendency of the waveform to distort, and linear theory provides an accurate description of the wave propagation. For $\Gamma > 1$, the initially sinusoidal waveform generates harmonics of the source frequency as it propagates. These harmonics cannot be described by linear theory. For $\Gamma \gg 1$, the waveform distortion is strong and results in the development of a shock front. Under this condition, a shock wave forms at distance $\bar{x}$. *See* HARMONIC (PERIODIC PHENOMENA).

Waveform distortion accumulates with distance because points on a waveform having different amplitudes propagate at different speeds. The relation for the propagation speed of individual points on a traveling sound wave is given by Eq. (3), where u

$$\frac{dx}{dt} = c_0 + \beta u \tag{3}$$

is the local particle velocity within the wave. This distortion mechanism is illustrated in **Fig. 1***a*. The peak in the sound wave has particle velocity u_p and travels faster than the zero crossings; the valley has a lower (negative) particle velocity u_v and travels slower. Time waveforms in Fig. 1*b*–*d* illustrate the waveform distortion that results from the amplitude-dependent propagation speed. Peaks advance on the zero crossings and valleys recede, until at $x = \bar{x}$ a shock is formed. Shock formation is said to occur when a vertical tangent first appears in the waveform.

The entire propagation history of a sinusoidal plane wave radiated by a source under the condition $\Gamma \gg 1$ is depicted in **Fig. 2**. The source waveform is illustrated in Fig. 2*a*, at the shock formation distance ($x = \bar{x}$) in Fig. 2*c*, and at the location of maximum shock amplitude [$x = (\pi/2)\,\bar{x}$] in Fig. 2*d*. Figure 2*e* and *f* reveal the propagation of a stable sawtooth waveform ($3\bar{x} < x \ll \ell_a$). Eventually, for $x \sim \ell_a$ (Fig. 2*g*), energy losses reduce the wave amplitude to such an extent that nonlinear effects no longer sustain a well-defined shock front. Finally, as depicted in Fig. 2*h*, for $x \gg \ell_a$, the old-age region is reached. Here, the waveform resembles its original sinusoidal shape, although substantially reduced in amplitude.

An example helps to place the above discussion in context. If a source radiates a plane wave in air at $f = 5$ kHz (at which frequency the absorption length is $\ell_a \sim 100$ m or 330 ft) with peak acoustic Mach number $\epsilon = 0.01$ (that is, a sound pressure level of 154 dB re 20 μPa), the Gol'dberg number is $\Gamma \simeq 100 \gg 1$, and the shock formation distance is $\bar{x} \simeq 1$ m (3.3 ft). A fully developed sawtooth wave exists beyond $3\bar{x} \simeq 3$ m (10 ft).

Plane waves, however, are idealizations often used to approximate propagation in ducts and the near-fields of sound beams. When propagation takes place in unbounded media, spreading of the wave reduces its amplitude as a function of distance from the source. For a wave radiated by a spherical source of radius r_0, the shock formation distance is given

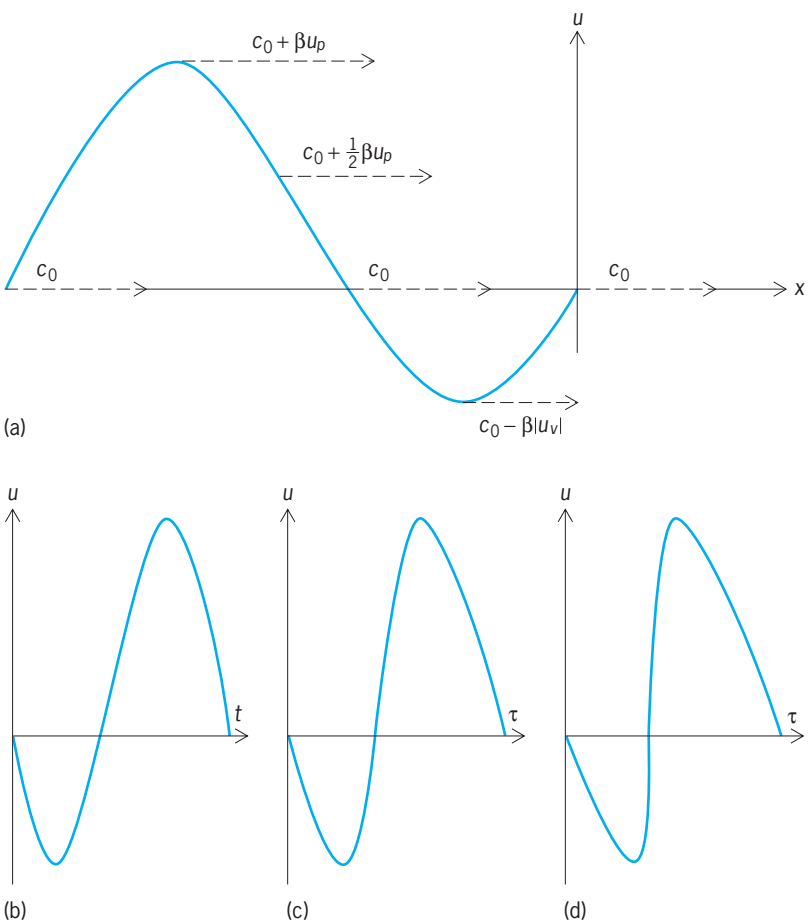

Fig. 1. Nonlinear distortion of an intense acoustic waveform. (*a*) Spatial plot (amplitude versus distance) of the initial acoustic waveform near $x = 0$. (*b*) Time waveform (plot of amplitude versus time) at distance $x = 0$. (*c*) Time waveform at $x > 0$. $\tau = t - x/c_0$. (*d*) Time waveform at $x = \bar{x}$, where a shock is formed. (*After D. T. Blackstock, Nonlinear acoustics (theoretical), in D. E. Gray, ed., American Institute of Physics Handbook, 3d ed. pp. 3-183–3-205, McGraw-Hill, 1972*)

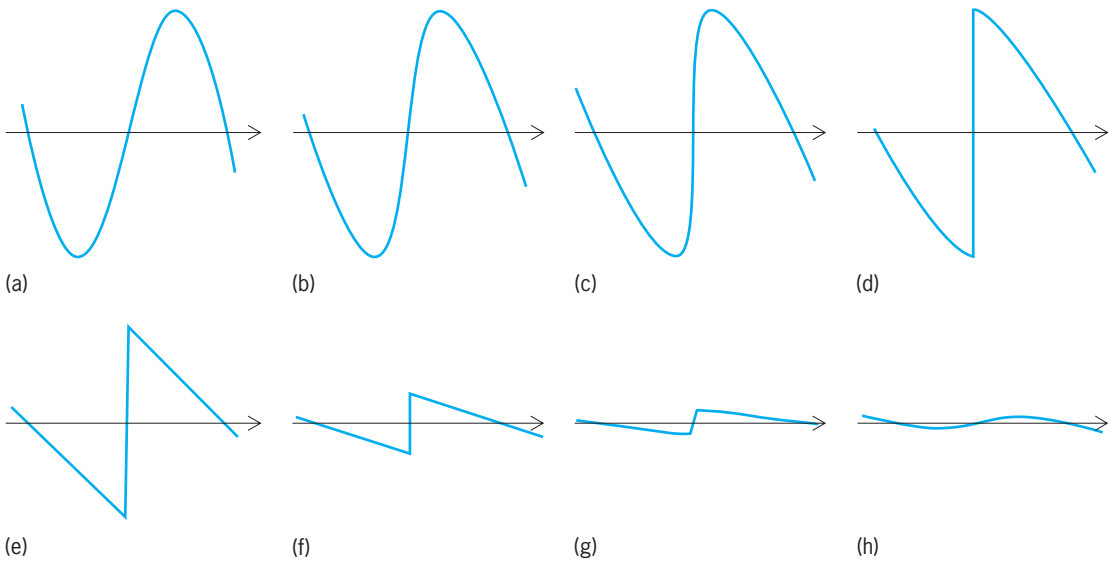

Fig. 2. Propagation history of an intense acoustic waveform that is sinusoidal at the source. (*a*) Source waveform, $x = 0$. (*b*) Distortion becoming noticeable. (*c*) Shock formation, $x = \bar{x}$. (*d*) Maximum shock amplitude, $x = (\pi/2)\,\bar{x}$. (*e*) Full sawtooth shape, $x = 3\,\bar{x}$. (*f*) Decaying sawtooth. (*g*) Shock beginning to disperse. (*h*) Old age. (*After J. A. Shooter et al., Acoustic saturation of spherical waves in water, J. Acous. Soc. Amer., 55:54–62, 1974*)

by Eq. (4) instead of Eq. (2). The shock formation

$$\bar{r} = r_0 e^{1/\beta\epsilon k r_0} \tag{4}$$

distance for a spherical wave can be substantially greater than that for a plane wave. For the same frequency and source level as in the previous example, the shock formation distance for a spherical wave radiated by a source of radius $r_0 = 0.5$ m (1.6 ft) is $\bar{r} \simeq 3$ m (10 ft), but for $r_0 = 0.1$ m (0.3 ft) it increases to $\bar{r} \simeq 900$ m (3000 ft). For the smaller source the relation $\bar{r} \gg l_a$ is obtained, shock formation cannot occur, and even second-harmonic generation is very weak. Spherical spreading reduces the amplitude of the wave so rapidly that shock formation cannot occur before the wave falls victim to strong energy losses.

Acoustic saturation. Energy losses at shock fronts impose an upper bound on how much sound power can be transmitted beyond a certain range. As the source level is increased, the shock formation distance is reduced, which allows more energy to be lost at the shocks before the wave arrives at a given location. For plane waves radiated by a monofrequency source for which $\Gamma \gg 1$, the peak sound pressure that can be produced at distance x is given by Eq. (5), which corresponds to the peak pressure

$$p_{\text{sat}} = \frac{\pi \rho_0 c_0^2}{\beta k x} \tag{5}$$

in the sawtooth waveforms depicted in Fig. 2*e* and *f*. Equation (5) is independent of the source pressure amplitude p_0. All additional energy pumped into the wave is lost at the shock fronts, and acoustic saturation is said to have occurred. The saturation curve for such a wave is shown in **Fig. 3**, where the extra attenuation indicates the energy loss not predicted by linear theory, designated by the diagonal line.

Acoustic saturation is not restricted to plane waves—saturation is frequently observed in sound beams—but not all waveforms saturate. In principle, any periodic acoustic wave will saturate provided sufficient sound power is available from the source. In contrast, short pulses do not experience saturation. The N wave (the acoustic waveform resembles the letter) is typical of sonic booms and also waveforms generated by spark sources in air, and it illustrates the point. For sufficiently high source levels, the peak amplitude of the N wave becomes proportional to the square root of the source amplitude, but it never becomes independent of source amplitude. There is nevertheless excess attenuation, because under these circumstances a 10-dB increase in source level results in only a 5-dB increase in the level of the received signal.

Frequency generation. The waveform distortion in Fig. 2 is accompanied by the generation of harmonics of the source frequency f. For $\Gamma \gg 1$, the amplitudes

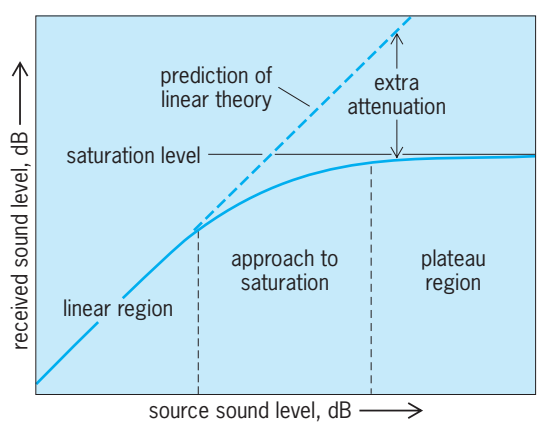

Fig. 3. Response curve illustrating the development of acoustic saturation. (*After D. A. Webster and D. T. Blackstock, Finite-amplitude saturation of plane sound waves in air, J. Acous. Soc. Amer., 62:518–523, 1977*)

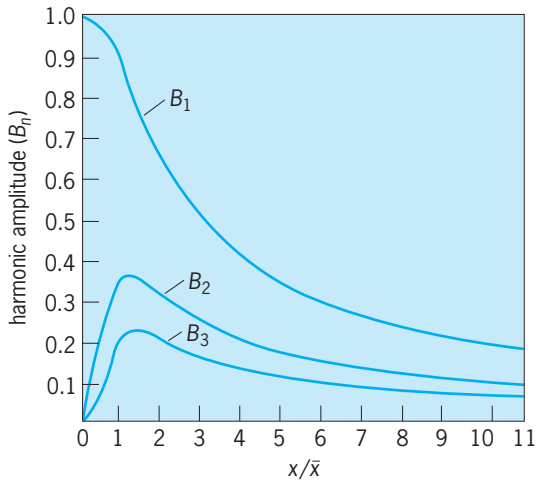

Fig. 4. Amplitudes of frequency components in an intense sound wave that is sinusoidal at the source. (*After D. T. Blackstock, Connection between the Fay and Fubini solutions for plane sound waves of finite amplitude, J. Acous. Soc. Amer., 39:1019–1026, 1966*)

of the harmonic components at frequencies nf ($n = 1, 2, \ldots$), normalized by the source pressure p_0, are given at different distances by Eqs. (6) and (7), where

$$B_n(x) = \frac{2J_n(nx/\bar{x})}{nx/\bar{x}} \qquad (6)$$

$$x/\bar{x} \leq 1$$

$$B_n(x) = \frac{2}{n(1 + x/\bar{x})} \qquad (7)$$

$$3 \lesssim x/\bar{x} \lesssim \Gamma$$

J_n is the Bessel function of the first kind, of order n. The first several amplitudes are plotted in **Fig. 4** [including the intermediate range $1 < x/\bar{x} < 3$, not covered by Eqs. (6) and (7)]. As the wave propagates in the preshock region, $x/\bar{x} < 1$, the amplitude of the source frequency component B_1 decreases as energy is expended to generate higher harmonics. In the region $x/\bar{x} > 3$, corresponding to Fig. 2*e* and *f*, the amplitudes of all spectral components decrease with distance from the source and are related in the ratio $1/n$, as is required by the Fourier series for an ideal sawtooth waveform. *See* BESSEL FUNCTIONS; FOURIER SERIES AND TRANSFORMS.

In a sound beam radiated from a monofrequency source, the beamwidths of the harmonic components become smaller approximately in proportion to $1/\sqrt{n}$, and the relative levels of the side lobes are also reduced as n increases. These effects improve the directional properties of the higher harmonics and are used to enhance resolution in acoustic imaging. In a sound beam so intense that acoustic saturation is approached along its axis, nonlinear losses tend to reduce these trends. Extra attenuation is most pronounced in the center of the beam, where the intensity is highest and shocks develop first. Erosion of the main lobe due to the increased energy loss near the axis causes both the beamwidth and side-lobe levels to increase. *See* DIRECTIVITY.

Coefficient of nonlinearity. For a given source amplitude and frequency, Eq. (2) for the shock formation distance indicates that the coefficient of nonlinearity β is a critical parameter in determining the rate of nonlinear wave distortion. For compressional waves in gases, liquids, and isotropic solids, standard expressions for this coefficient are given by Eqs. (8)–(10). In Eq. (8), γ is the ratio of specific heats, and

$$\beta = \frac{\gamma + 1}{2} \qquad \text{gas} \qquad (8)$$

$$\beta = \frac{B}{1 + 2A} \qquad \text{liquid} \qquad (9)$$

$$\beta = -\left(\frac{3}{2} + \frac{\mathcal{A} + 3\mathcal{B} + \mathcal{C}}{\rho_0 c_0^2}\right) \qquad \text{isotropic solid} \qquad (10)$$

therefore with $\gamma = 1.4$ for diatomic gases $\beta = 1.2$ for air.

For liquids, an arbitrary equation of state relating the sound pressure to the excess density ρ' at constant entropy may be expanded in the Taylor series given by Eq. (11), where $A = \rho_0 c_0^2$ and the ratio B/A is given by Eq. (12). The subscript s indicates that

$$p = A\left(\frac{\rho'}{\rho_0}\right) + \frac{B}{2!}\left(\frac{\rho'}{\rho_0}\right)^2 + \frac{C}{3!}\left(\frac{\rho'}{\rho_0}\right)^3 + \cdots \qquad (11)$$

$$\frac{B}{A} = \frac{\rho_0}{c_0^2}\left(\frac{\partial^2 p}{\partial \rho'^2}\right)_{s,0} \qquad (12)$$

the derivatives of pressure with respect to density are evaluated at constant entropy, with the 0 referring to ambient conditions. Although Eq. (12) is the fundamental definition, other formulations are often used for measuring B/A, for example, in terms of sound speed variation as a function of pressure, or for conditions of constant pressure and temperature rather than constant entropy. A short compilation of measurements is presented in the **table**. For pure water at room temperature, $B/A = 5.0$, in which case $\beta = 3.5$. Because of the important role of nonlinear acoustics in medical applications, considerable effort has been devoted to collecting measurements of B/A for a wide variety of biologic materials, values of which nominally range from 5 to 10. It is rarely necessary to consider the higher-order coefficients, C, D, and so forth, in Eq. (11).

Equation (10) shows that although only one quantity must be measured to determine β for a liquid, three are required for an isotropic solid. The parameters $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are referred to as third-order elastic constants because they appear at third order in an expansion of the strain energy function. Moreover, it is difficult to obtain accurate measurements of these constants. Values of β for common isotropic solids tend to fall within the range of values for common liquids. Whereas β is invariably positive for fluids (excluding anomalous behavior near critical points), it is negative for several solids, such as fused quartz and Pyrex glass. In this case, the acoustic waveforms distort in directions that are opposite those shown in Figs. 1 and 2. For example, the relation $\beta > 0$ indicates that only compression shocks can form in gases

Values of *B/A* for liquids at atmospheric pressure*

Liquid	T, °C (°F)	B/A
Distilled water	0 (32)	4.2
	20 (68)	5.0
	40 (104)	5.4
	60 (140)	5.7
	80 (176)	6.1
	100 (212)	6.1
Seawater (3.5% salinity)	20 (68)	5.25
Alcohols		
Methanol	20 (68)	9.6
Ethanol	20 (68)	10.5
n-Propanol	20 (68)	10.7
n-Butanol	20 (68)	10.7
Organic liquids		
Acetone	20 (68)	9.2
Benzene	20 (68)	9.0
Chlorobenzene	30 (86)	9.3
Cyclohexane	30 (86)	10.1
Diethylamine	30 (86)	10.3
Ethylene glycol	30 (86)	9.7
Ethyl formate	30 (86)	9.8
Glycerol (4% H_2O)	30 (86)	9.0
Heptane	30 (86)	10.0
Hexane	30 (86)	9.9
Methyl acetate	30 (86)	9.7
Methyl iodide	30 (86)	8.2
Nitrobenzene	30 (86)	9.9
Liquid metals		
Bismuth	318 (604)	7.1
Indium	160 (320)	4.6
Mercury	30 (86)	7.8
Potassium	100 (212)	2.9
Sodium	110 (230)	2.7
Tin	240 (464)	4.4

*After R. T. Beyer, *Nonlinear Acoustics*, Navy Sea Systems Command, 1974.

and liquids, yet only expansion shocks can form for $\beta < 0$, as in these particular solids.

Nonlinearity in anisotropic solids depends on larger numbers of third-order elastic constants. The strength of nonlinear effects in crystals depends on the direction of propagation, but in general it is of the order encountered in isotropic solids. In contrast, microinhomogeneous features of rock yield effective values of β that may be larger than those of isotropic solids by several orders of magnitude. However, acoustic propagation in rock is often accompanied by strong attenuation that may offset nonlinear frequency generation. There is significant interest in nonlinear properties of rock in the oil exploration industry because seismologists can potentially deduce, by acoustic measurements, the magnitude of the local stress environment. In addition, there is the expectation that measured values of the third-order elastic constants may correlate inversely with the strength of the rock. Both of these considerations are relevant for the safe construction of productive oil wells. *See* GEOPHYSICAL EXPLORATION; OIL AND GAS WELL DRILLING.

Parametric array. The parametric array is a classical application of nonlinear acoustics developed originally for application to underwater sonar. A source radiates two signals simultaneously at the two neighboring frequencies $f_1 \simeq f_2$. These two sound beams interact as they propagate and generate a spectrum of intermodulation frequencies, among them the difference frequency $f_- = |f_1 - f_2|$, which is much less than f_1 or f_2 and is the lowest in the spectrum. It is the volume of fluid in which the nonlinear interaction generates the difference frequency signal that is referred to as the parametric array. When absorption terminates nonlinear interaction within the nearfield of the primary beams, the beam profile of the difference frequency signal is given by Eq. (13), where

$$D(\theta) = \frac{1}{\sqrt{1 + (k_- \ell_a)^2 \sin^4(\theta/2)}} \tag{13}$$

θ is angle from the central axis, ℓ_a is the absorption length at the primary frequencies, and $k_- = 2\pi f_-/c_0$.

The most useful feature of the parametric array is its beamwidth. The beamwidth is inversely proportional to $\sqrt{k_- \ell_a}$, which depends on the absorption coefficient of the primary waves, but not on the size of the source. In contrast, direct radiation by a circular source of radius a at frequency f_- produces a beamwidth that is inversely proportional to $k_- a$, which depends critically on source size. The principal benefit of the parametric array is that very narrow beams of low-frequency sound can be generated by the nonlinear interaction of high-frequency primary beams radiated from a fairly small source. Because $\ell_a \gg a$, the beamwidth can be far narrower, by an order of magnitude, than can be obtained by direct radiation from the source of the primary waves, but at frequency f_-.

The absorption length ℓ_a is the effective aperture length of the array because it defines the volume of fluid in front of the source where the difference frequency is generated. Since the difference frequency component propagates at the same speed as the primary waves that generate it, the nonlinear interaction volume behaves as an end-fire array. In underwater applications, the parametric array is used for long-range propagation of highly directional, low-frequency sound. It also forms the basis for audio loudspeakers currently under development, which depend on a related phenomenon called self-demodulation.

Self-demodulation. The parametric array may be used to generate low-frequency sound possessing a beam pattern described by Eq. (13), yet having a bandwidth not restricted to a single difference frequency component. A carrier wave of frequency f_c is amplitude-modulated by an envelope $E(t)$ and radiated by a directional sound source. In the farfield along the axis of the beam, the time dependence is given by Eq. (14). There is no remnant of the carrier frequency f_c, because it is much greater than the frequencies associated with the modulation envelope $E(t)$. The carrier is therefore absorbed much more rapidly by the fluid, in which attenuation increases close to quadratically with frequency. All that remains is a squared and differentiated version of the envelope. Evolution of the waveform toward this farfield result is called self-demodulation.

$$F(t) = \frac{d^2 E^2(t)}{dt^2} \tag{14}$$

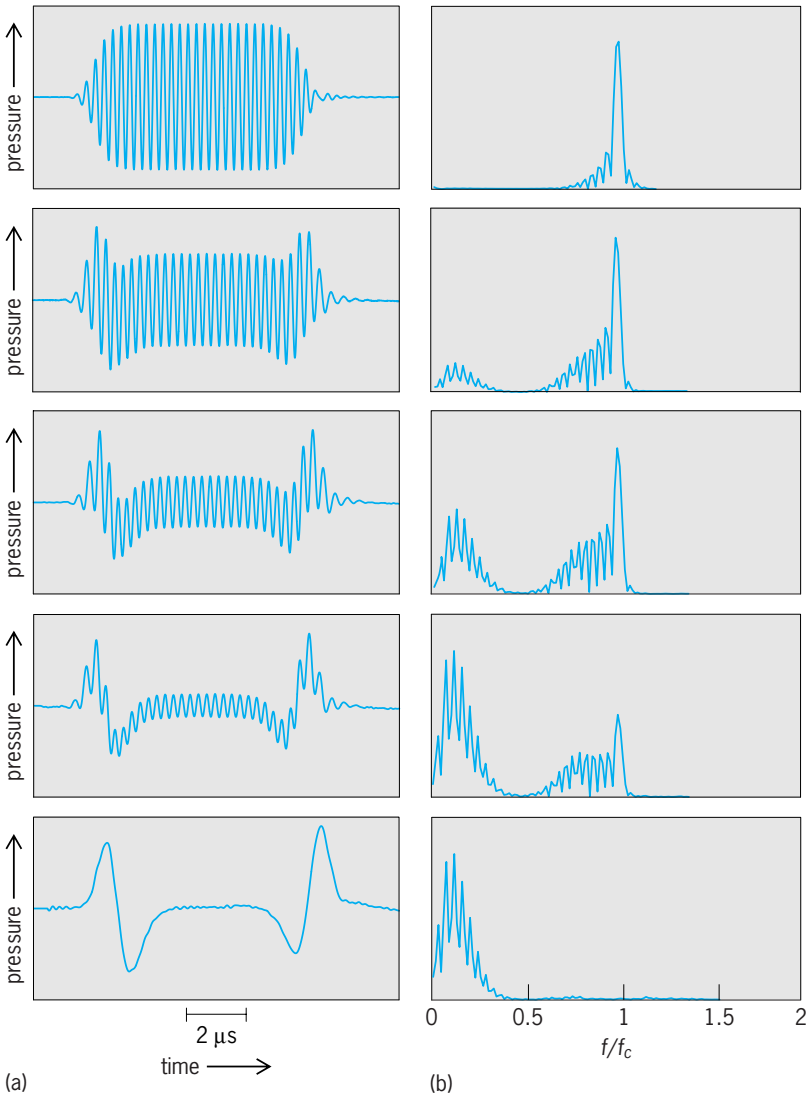

Fig. 5. Measurements of self-demodulation of an acoustic pulse in glycerin at increasing distances from the source, from 5 cm (2.0 in.) in the first row to 26 cm (10.2 in.) in the bottom row. (*a*) Acoustic waveforms. (*b*) Corresponding frequency spectra. f = frequency; f_c = carrier frequency = 3.5 MHz. (*After M. A. Averkiou et al., Self-demodulation of amplitude and frequency modulated pulses in a thermoviscous fluid, J. Acous. Soc. Amer., 94:2876–2883, 1993*)

Self-demodulation in glycerin of an acoustic pulse with carrier frequency $f_c = 3.5$ MHz is depicted in **Fig. 5**, with measured acoustic waveforms and the corresponding frequency spectra. Each row down is at a distance farther from the source, from 5 cm (2.0 in.) for the first row to 26 cm (10.2 in.) for the last. In the first row, there is no observable effect of nonlinearity, either in the time waveform or in the frequency spectrum. As the high-frequency carrier wave is absorbed by the fluid, a self-demodulated signal emerges in the time waveforms, and a spectral shift from high to low frequencies is observed. Equation (14) is an accurate description of the waveform in the bottom row, where no energy is apparent at frequencies $f \simeq f_c$.

In air, parametric arrays are under development that are driven at ultrasonic frequencies and generate low-frequency sound in the audio range. This application is sometimes referred to as the audio spotlight, because it is capable of transmitting sound (with some distortion) in beams that are remarkably narrow and can selectively insonify very localized regions in space. Electrical predistortion of the signal input to the loudspeaker is needed to compensate for the squaring of the modulation envelope $E(t)$ in Eq. (14).

Acoustic streaming. Acoustic streaming is the steady flow of fluid produced by the absorption of sound. It is a nonlinear effect because the velocity of the flow depends quadratically on the amplitude of the sound, and the flow is not predicted by linear theory. Absorption due to viscosity and heat conduction results in a transfer of momentum from the sound field to the fluid. This momentum transfer manifests itself as steady fluid flow.

Acoustic streaming produced in sound beams is enhanced considerably when shocks develop. Shock formation generates a frequency spectrum rich in higher harmonics. Because thermoviscous absorption increases quadratically with frequency, attenuation of the wave, and therefore the streaming velocity, increases markedly following shock formation. Streaming is also generated in acoustic boundary layers formed by standing waves in contact with surfaces.

Measurements of acoustic streaming have been used to determine the bulk viscosity coefficients of fluids. Acoustic streaming affects the performance of thermoacoustic engines and refrigerators, which exploit the heat transfer induced by the temperature and velocity fluctuations in a sound wave in contact with a surface possessing large heat capacity. These devices are adversely affected by heat transport associated with streaming. In thermoacoustic devices operating at audio frequencies in gases, the streaming velocities are of order 10 cm/s (4 in./s). At the megahertz frequencies and acoustic intensities encountered in biomedical applications, streaming velocities attained in liquids are also of order 10 cm/s.

Phase conjugation. Phase conjugation refers to wavefront reversal, also called time reversal, at a single frequency. The latter terminologies more clearly describe this procedure. A waveform is captured by a phase conjugation device and reversed in such a way that it propagates back toward the source in the same way that it propagated toward the conjugator. Under ideal circumstances, it is as though a movie of the wave propagation from source to conjugator were played backward in time. Sound that is radiated from a point source and propagates through an inhomogeneous medium that introduces phase distortion in the wave field is thus retransmitted by the conjugator in such a way as to compensate for the phase distortion and to focus the wave back on the point source.

Phase conjugation is used to compensate for phase distortion in applications involving imaging and retargeting of waves on sources. Although it is a mature technology in nonlinear optics, practical methods of phase conjugation in nonlinear acoustics were not developed until the 1990s. The most successful techniques for acoustical phase conjugation are based on

modulation of acoustical properties of a material that captures the incident sound wave. The modulation is twice the frequency of the incident sound wave, and it is induced by an electric field applied to piezoelectric material, or a magnetic field applied to magnetostrictive material. Often the modulated property of interest is the sound speed in the material. When the incident wave at frequency f propagates through a medium in which the sound speed fluctuates at frequency $2f$, parametric interaction generates a wave at the difference frequency f that propagates backward as though reversed in time. *See* MAGNETOSTRICTION; OPTICAL PHASE CONJUGATION; PIEZOELECTRICITY.

The use of phase conjugation to reduce phase distortion in the acoustic imaging of a perforated steel plate embedded in an aberrating layer is demonstrated in **Fig. 6**. The plate perforations are in the shapes of various letters. The plate was embedded in gel having a corrugated surface (Fig. 6*a*) that introduced phase distortion of acoustic wavefronts propagating through the layer. The letters are barely discernible in an image constructed from a conventional acoustic scan of the sample at 10 MHz (Fig. 6*b*). When phase conjugation was introduced by applying a dc electric bias field and a fluctuating electric field of frequency 20 MHz to a piezoelectric material, an image was obtained with the letters considerably reduced in distortion (Fig. 6*c*). Phase conjugation based on magnetostriction, also being developed for acoustic imaging, permits substantial amplification of the conjugate wave. The amplification is sufficiently large that harmonics can be generated in the conjugate wave field, which provides additional opportunities for image enhancement.

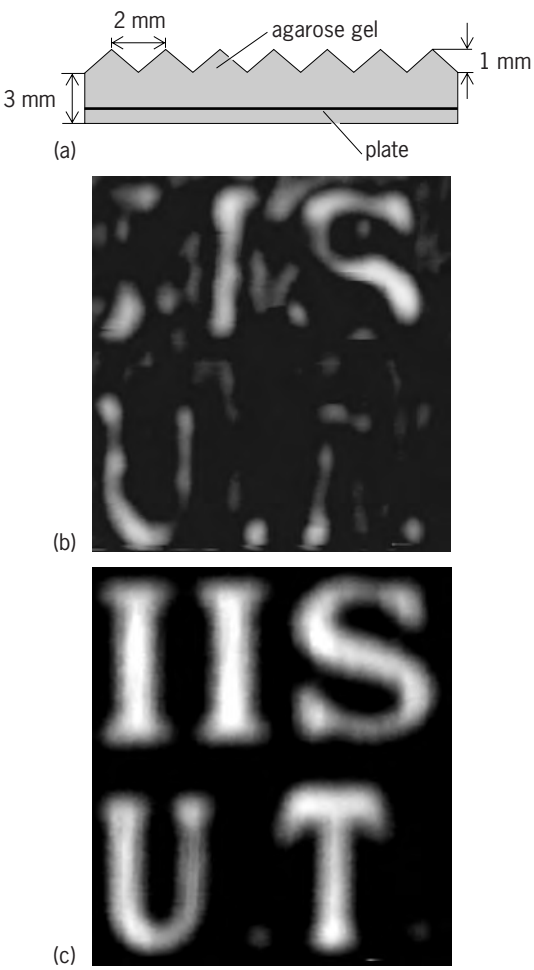

Fig. 6. Use of phase conjugation to improve acoustic imaging of a perforated steel plate embedded in an aberrating layer of gel. The perforations have the shapes of letters. (*a*) Geometry and dimensions of the sample, consisting of a plate embedded in gel. (*b*) Conventional acoustic image of the plate within the gel. (*c*) Acoustic image of the plate within the gel when phase conjugation is used. (*From K. Yamamoto et al., Acoustic phase conjugation by nonlinear piezoelectricity, II. Visualization and application to imaging systems, J. Acous. Soc. Amer., 106:1339–1345, 1999*)

Biomedical applications. Phenomena associated with nonlinear acoustics have proved useful in both diagnostic and therapeutic applications of biomedical ultrasound. A very significant breakthrough in diagnostic imaging, especially for echocardiography and abdominal ultrasound imaging, is based on second-harmonic generation. Medical ultrasound imaging is performed at frequencies of several megahertz. Images constructed from the backscattered second-harmonic component have substantially reduced clutter and haze associated with the propagation of ultrasound through the outer layers of skin, which is the primary cause of phase aberrations. In another technique, microbubbles are injected into the bloodstream to enhance echoes backscattered from blood flow. The microbubbles are fabricated to make them resonant at diagnostic imaging frequencies, and they become strongly nonlinear oscillators when excited by ultrasound. Imaging is based on echoes at harmonics of the transmitted signal. Frequencies backscattered from the microbubbles differ from those in echoes coming from the surrounding tissue, which highlights the locations of the microbubbles and therefore of the blood flow itself.

A notable therapeutic application is lithotripsy, which refers to the noninvasive disintegration of kidney stones and gallstones with focused shock waves. Nonlinear acoustical effects in lithotripsy are associated not only with propagation of the shock wave but also with the generation of cavitation activity near the stones. Radiation of shock waves due to the collapse of cavitation bubbles is believed to be the dominant cause of stone breakup. An emerging therapeutic application, high-intensity focused ultrasound (HIFU), utilizes the heat dissipated by shock waves that develop in beams of focused ultrasound. The heating is so intense and localized that the potential exists for noninvasive cauterization of internal wounds and removal of tumors and scar tissue. *See* CAVITATION; ULTRASONICS.

Mark F. Hamilton

Bibliography. R. T. Beyer, *Nonlinear Acoustics*, rev. ed., Acoustical Society of America, 1997; M. F. Hamilton and D. T. Blackstock, *Nonlinear Acoustics*, Academic Press, 1998; K. Naugolnykh and L. Ostrovsky, *Nonlinear Wave Processes in Acoustics*, Cambridge University Press, 1998; O. V. Rudenko and S. I. Soluyan, *Theoretical Foundations of Nonlinear Acoustics*, Plenum Press, 1977.

Nonlinear control theory

A control system involves a plant and a controller. Plants are objects as diverse as a satellite, a distillation column, a robot arm, and a colony of bacteria. After measuring actual outputs of the plant, the controller computes signals that are applied at the inputs to the plant to achieve desired outputs. The design of controllers must be based upon mathematical models of plants, which are in most realistic situations composed of nonlinear differential and difference equations. A standard approach is to linearize the equations and use the powerful methods available for the design of linear control systems. *See* DIFFERENTIAL EQUATION; LINEAR SYSTEM ANALYSIS; OPTIMAL CONTROL (LINEAR SYSTEMS).

When the controlled outputs are allowed to have large deviations from the desired steady-state values, a linearized model will cease to describe the plant accurately, thereby causing erroneous results in the design. Linearized design models also fail in those important situations where nonlinearities are introduced into a controller to achieve a desired performance, generally at reduced cost. Typical examples of nonlinear controllers are on-off relays for temperature regulation in heating and cooling systems, switching elements in robot manipulators, and jet thrusters for the attitude control of space vehicles.

Unlike linear systems, there is no general theory for nonlinear systems. Nonlinear control theory is fragmented into areas centered around those classes of systems that are most prominent in applications. The basic ideas and concepts underlying a few of the areas will be described, with an emphasis on some fundamental mathematical notions and techniques.

Nonlinear systems. A simple nonlinear control system is the inverted pendulum (**Fig. 1***a*), where it is required to keep the pendulum in an upright position by applying the torque $u(t)$ at its base. The pendulum can be considered as a generic model for a rocket booster balanced on top of gimbaled thruster engine (Fig. 1*b*), and a controlled robot arm (Fig. 1*c*).

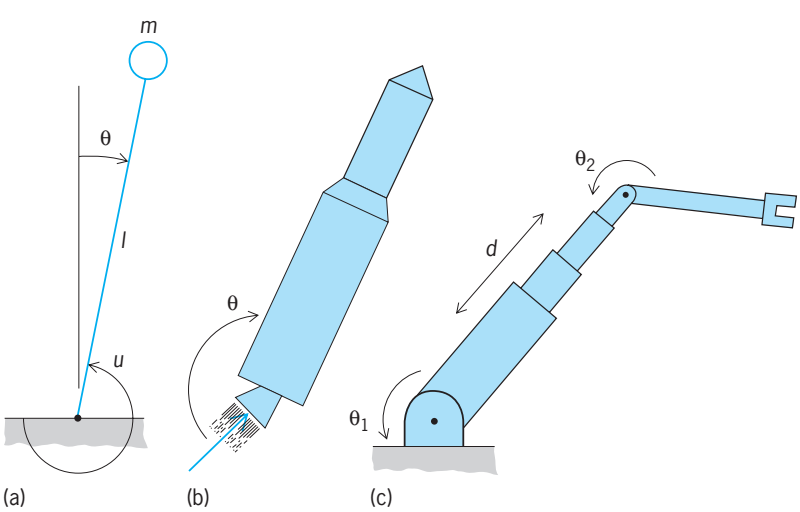

Fig. 1. Nonlinear systems. (*a*) Inverted pendulum, which serves as a model for (*b*) a rocket booster and (*c*) a robot arm.

Newton's second law of motion for the pendulum is Eq. (1), where m is the mass of the bob, ℓ is the length of the pendulum, and g is the acceleration due to gravity. By choosing the angular position $\theta = x_1$ and angular velocity $\dot{\theta} = x_2$ of the bob as the state of the system, Eq. (1) can be rewritten as two first-order state equations (2). *See* PENDULUM.

$$m\ddot{\theta}(t) = \frac{mg}{\ell}\sin\theta(t) + u(t) \tag{1}$$

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \frac{g}{\ell}\sin x_1 + u \end{aligned} \tag{2}$$

If the torque is absent ($u = 0$), the upright position of a motionless pendulum ($\dot{x}_1 = 0$, $\dot{x}_2 = 0$) is an equilibrium state: $x_1 = 0$, $x_2 = 0$. If the pendulum is slightly perturbed at this state, it will fall down and keep oscillating around the other equilibrium point at $x_1 = \pi$ and $x_2 = 0$, where again, $\dot{x}_1 = 0$ and $\dot{x}_2 = 0$.

To keep the pendulum in the upright position in physically realistic situations involving perturbations, the torque u is chosen to be a suitable function ϕ of the state x_1, x_2, as in Eq. (3). That u should be a

$$u = \phi(x_1, x_2) \tag{3}$$

feedback control law in terms of the state is one of the most important results of control theory.

Linearization. A large class of plants can be described by a vector differential equation (4), where

$$\dot{x} = f(t, x, u) \tag{4}$$

the vector $x = (x_1, x_2, \ldots, x_n)^T$ is the state, the vector $u = (u_1, u_2, \ldots, u_m)^T$ is the input of the plant at time t, and T denotes the transpose which sends row vectors into column vectors. The plant is specified by the vector-valued function $f = (f_1, f_2, \ldots, f_n)^T$, which is the infinitesimal state transition function describing the evolution of the state $x \to x + dx$ corresponding to the change in time $t \to t + dt$.

In the absence of control, the plant is described by Eq. (5). The motions $x(t; t_0, x_0)$ of the plant are

$$\dot{x} = f(t, x) \tag{5}$$

solutions of Eq. (5) for given initial conditions t_0, x_0. Equilibrium states are constant solutions of Eq. (5), which are determined by Eq. (6). When a system

$$f(t, x) = 0 \quad \text{for all } t \tag{6}$$

starts at an equilibrium state x_e, it stays there forever; that is, it satisfies Eq. (7).

$$x(t; t_0, x_e) = x_e \quad \text{for all } t > t_0 \tag{7}$$

A common approach to control system design is to assume small deviations of the state from a desired equilibrium x_e and linearize the plant at x_e. It is assumed, for simplicity, that the plant is time-invariant, so that Eq. (4) reduces to Eq. (8), and that $x_e = 0$.

$$\dot{x} = f(x, u) \tag{8}$$

The function $f(x,u)$ can then be expanded in a Taylor series at $x = 0$ and $u = 0$ to obtain an approximate linear model given by Eq. (9), where the constant

$$\dot{x} = Ax + Bu + h(x, u) \tag{9}$$

matrices $A = (a_{ij})$ and $B = (b_{ij})$ of dimensions $n \times n$ and $n \times m$, respectively, have elements given by Eqs. (10). The vector-valued function $h\ (x,u)$ represents the higher-order terms in the series and any modeling uncertainty in the system.

$$a_{ij} = \frac{\partial f_i}{\partial x_j} \qquad b_{ij} = \frac{\partial f_i}{\partial u_j} \tag{10}$$

A linear feedback control law given by Eq. (11) can be used in Eq. (9) to obtain the closed-loop system given by Eq. (12). Numerous methodsof linear

$$u = -Kx \tag{11}$$

$$\dot{x} = (A - BK)x + h(x) \tag{12}$$

system analysis can be applied to determine an appropriate $m \times n$ gain matrix $K = (k_{ij})$, so that the linear part of Eq. (12) behaves satisfactorily and, at the same time, dominates any adverse effect of the nonlinear perturbation $h(x)$.

The linearized model of the pendulum is given by Eq. (13). A linear control, which can be expressed by

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ \frac{g}{\ell} & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + h(x) \tag{13}$$

Eq. (14), can always be chosen ($k_1 > g/l$, $k_2 > 0$) to make the closed loop, given by Eq. (15), stable, such

$$u = -k^T x \qquad k = (k_1, k_2) \tag{14}$$

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ \frac{g}{\ell} - k_1 & -k_2 \end{bmatrix} x + h(x) \tag{15}$$

that the controlled pendulum would return to the upright position after perturbation. In fact, stabilizing gains k_1 and k_2 can always be found regardless of the size of perturbation $h(x)$, although they must often have large values; this is so-called high-gain feedback.

Nonlinear feedback can be used to achieve exact linearization. By choosing the torque given by Eq. (16), the original nonlinear system of Eqs. (2)

$$u = -\frac{g}{\ell} \sin x_1 - k_1 x_1 - k_2 x_2 \tag{16}$$

becomes the linear system of Eqs. (15) without any perturbation $h(x)$. However, a simple cancellation of nonlinearity by the first term in Eq. (16) is not always possible; sophisticated concepts of differential geometry are needed to accomplish an exact linear design of an otherwise nonlinear plant. *See* DIFFERENTIAL GEOMETRY.

Lyapunov's method. An equilibrium of a dynamic system is stable if, after the system is slightly perturbed from the equilibrium, any subsequent motion of the system remains close to the equilibrium. If, in addition, each motion approaches the equilibrium as time increases, that is, expression (17) is

$$x(t; t_0, x_0) \to x_e \qquad \text{as } t \to \infty \tag{17}$$

valid, then the equilibrium is asymptotically stable. *See* CONTROL SYSTEM STABILITY.

The underlying idea of Lyapunov's direct method, applied to a given isolated system, involves finding an energylike function $V(x)$ that is positive and decays at each x except at the equilibrium x_e; that is, $V(x) > 0$ and $V(x) < 0$ for all $x \neq x_e$. Then, as $V(x)$ continually decreases toward its minimum value $V(x_e)$, the motion $x(t; t_0, x_0)$ asymptotically approaches x_e.

In Eq. (15), if $h(x)$ is neglected and the values in Eqs. (18) are chosen for k_1 and k_2, then the resulting linear control system is given by Eqs. (19). If the

$$k_1 = 1 + \frac{g}{\ell} \qquad k_2 = 1 \tag{18}$$

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 - x_2 \end{aligned} \tag{19}$$

function $V(x)$ is chosen as given in Eq. (20), then it follows from Eqs. (19) that the derivative of $V(x)$is given by Eqs. (21), which proves the asymptotic stability of $x_e = 0$. The motions of the system intersect the closed constant V-contours from the outside toward the inside until they reach the equilibrium (**Fig. 2**).

$$V(x) = {}^3/_2 x_1^2 + x_1 x_2 + x_2^2 > 0 \tag{20}$$

$$\begin{aligned} \dot{V}(x) &= 3x_1\dot{x}_1 + \dot{x}_1 x_2 + x_1 \dot{x}_2 + 2x_2\dot{x}_2 \\ &= -(x_1^2 + x_2^2) < 0 \end{aligned} \tag{21}$$

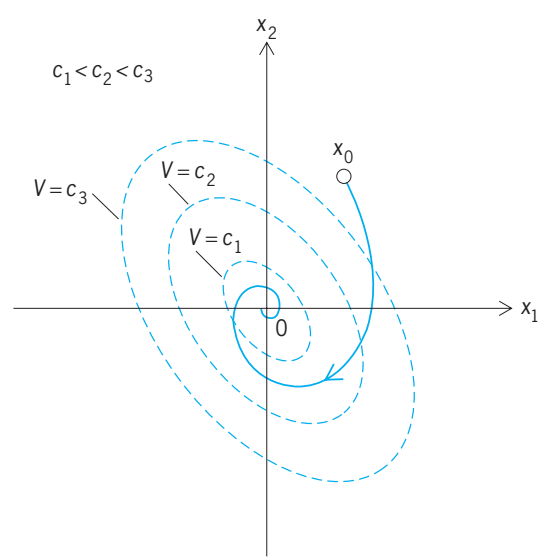

Fig. 2. Phase-plane (x_1x_2-plane) portrait of the system given by Eqs. (19), showing contours of constant value of the Lyapunov function V given in Eq. (20) and a typical trajectory.

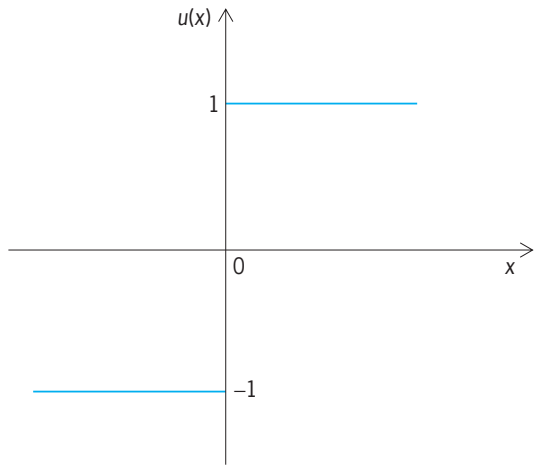

Fig. 3. Control function characteristic of a relay.

When an additional control term given by Eq. (22)

$$u = -\text{sgn}\,({}^1/_2 x_1 + x_2) \tag{22}$$

is inserted into the second of Eqs. (19), it increases the negativity of $V(x)$ by $-({}^1/_2 x_1 + x_2)\ \text{sgn}\ ({}^1/_2 x_1 + x_2)$. The added term produces the steepest possible descent motion of the system toward the equilibrium and, at the same time, makes the closed-loop system robust to uncertain perturbations $b(x)$ that were neglected in Eq. (19). The control law of Eq. (22) can be implemented by a relay (**Fig. 3**). Controllers of this type are simple and reliable, resulting in robust variable structure systems.

In general, when a system is given as in Eq. (5), the aim of Lyapunov's method is to find a positive definite function $V(t,x)$, such that its total time derivative, given by Eq. (23), is negative definite along any

$$\dot{V}(t,x) = \frac{\partial V}{\partial t} + (\text{grad}\ V)^T f(t,x) \tag{23}$$

solution of Eq. (5). Stability, if present, is established without solving Eq. (5). Drawbacks of the method are its inconclusiveness when it fails and the difficulty in finding a suitable Lyapunov function $V(t,x)$ that proves the stability of a given system. Generalizations of the method include vector Lyapunov functions for large-scale systems, hereditary systems, stochastic control, and robust control of uncertain systems. *See* LARGE SYSTEMS CONTROL THEORY; STOCHASTIC CONTROL THEORY.

Popov's method. A wide class of nonlinear control systems can be identified as the Lur'e type having the representation of Eqs. (24) as well as a block-

$$\dot{x} = Az + bu \tag{24a}$$

$$y = c^T x \tag{24b}$$

$$u = \phi(y) \tag{24c}$$

diagram representation (**Fig. 4***a*). The system is an interconnection of the linear plant and a nonlinear controller. The linear part is described by the two state equations (22*a*) and (22*b*) or, equivalently, by the transfer function of Eq. (25), where A is a matrix,

$$G(s) = c^T(A - sI)^{-1} b \tag{25}$$

b and c are column vectors, I is the identity matrix (all of appropriate dimensions), and $s = \sigma + j\omega$ is the complex variable. The nonlinear function ϕ of the controller belongs to a sector $[0,k]$ (Fig. 4*b*).

A Lur'e-type system is said to be absolutely stable if the equilibrium $x_e = 0$ is stable for any initial state x_0 and any ϕ that belongs to the sector $[0, k]$. For this class of systems, a Lyapunov function is available as given in Eq. (26), which is a weighted

$$V(x) = x^T H x + q \int_0^{c^T x} \phi(y)\, dy \tag{26}$$

sum of a quadratic form (with positive definite matrix H) plus an integral of the nonlinearity. The system is absolutely stable, and $V(x)$is a Lyapunov function, if there is a number q so that the Popov's frequency domain condition, given by inequality (27), is satisfied.

$$k^{-1} + \text{Re}[(1 + j\omega q)G(j\omega)] > 0 \qquad \text{for all } \omega \tag{27}$$

A graphical interpretation of inequality (27) can be given (**Fig. 5**), where the Popov locus $G^*(j\omega)$, given by Eq. (28), is required to lie to the right of the Popov line.

$$G^*(j\omega) = \text{Re}\ G(j\omega) + j\omega\ \text{Im}\ G(j\omega) \tag{28}$$

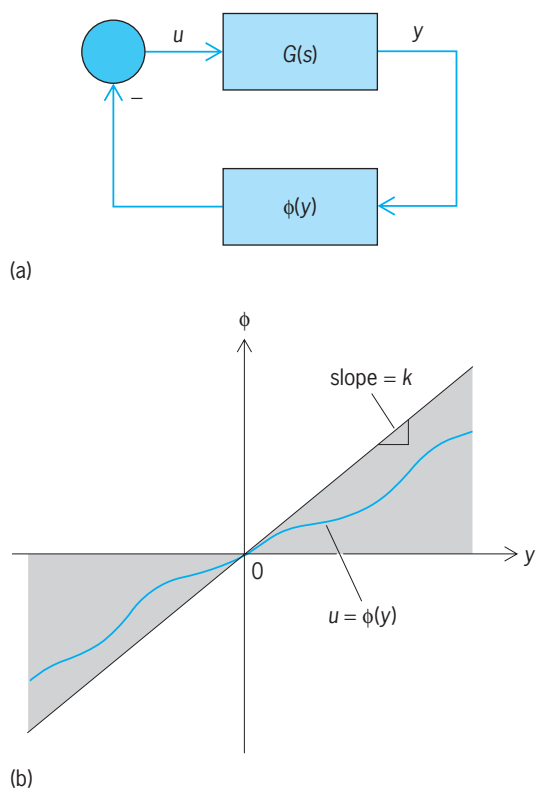

Fig. 4. Lur'e-type nonlinear control system. (*a*) Block diagram. (*b*) Nonlinearity sector, the shaded area between the *y* axis and the line through the origin with slope *k*, to which the nonlinear function of the controller, *u* = ϕ(*y*), is confined.

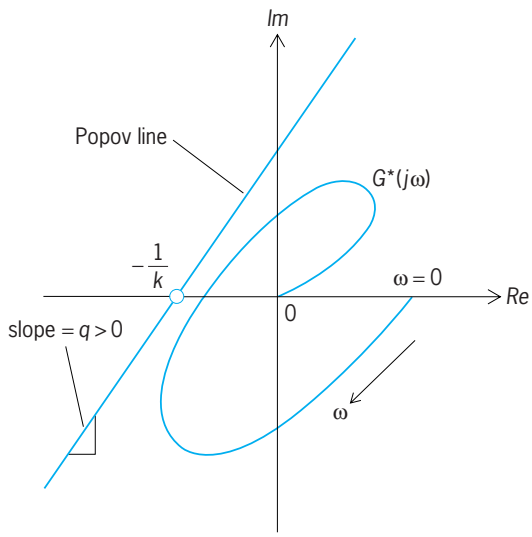

Fig. 5. Popov's graphical criterion. The Popov locus $G^*(j\omega)$ must lie to the right of the Popov line.

The essential significance of the Popov's method is its use of the powerful frequency characteristics in the stability analysis of nonlinear control systems. In particular, it can provide a justification for the use of the describing function method, which is another widely used frequency domain technique but relies on a Fourier-series approximation of the nonlinearity. Generalizations of the Popov's method include plants with multiple inputs and outputs, hyperstability, and adaptive control systems. *See* ADAPTIVE CONTROL; CONTROL SYSTEM STABILITY; CONTROL SYSTEMS.

Dragoslav D. Šiljak

Bibliography. K. J. Åstrom and B. Wittenmark, *Adaptive Control*, 2d ed., 1994; A. Isidori, *Nonlinear Control Systems: An Introduction*, 1985; H. K. Khalil, *Nonlinear Systems*, 2d ed., 1995; V. M. Popov, *Hyperstability of Control Systems*, 1973; D. D. Šiljak, *Nonlinear Systems*, 1969; M. W. Spong and M. Vidyasagar, *Robot Dynamics and Control*, 1989; V. I. Utkin, *Sliding Modes in Control and Optimization*, 1992.

Nonlinear optical devices

Devices that use the fact that the polarization in any real medium is a nonlinear function of the optical field strength to implement various useful functions. The nonlinearities themselves can be grouped roughly into second-order and third-order. Materials that possess inversion symmetry typically exhibit only third-order nonlinearities, whereas materials without inversion symmetry can exhibit both second- and third-order nonlinearities. *See* CRYSTALLOGRAPHY; ELECTRIC SUSCEPTIBILITY; ELECTROMAGNETIC RADIATION; POLARIZATION OF DIELECTRICS.

Second-order devices. Devices based on the second-order nonlinearity involve three-photon (or three-wave) mixing. In this process, two photons are mixed together to create a third photon, subject to energy- and momentum-conservation constraints. Different names are ascribed to this mixing process, depending upon the relative magnitudes of the energies of the three photons. *See* CONSERVATION OF ENERGY; CONSERVATION OF MOMENTUM.

Second harmonic generation and sum-frequency mixing. When the two beginning photons are of equal energy or frequency, the mixing process gives a single photon with twice the energy or frequency of the original ones. This mixing process is called second-harmonic generation. Second-harmonic generation is used often in devices where photons of visible frequency are desired but the available underlying laser system is capable of producing only infrared photons. For example, the neodymium-doped yttrium-aluminum-garnet (Nd:YAG) laser produces photons in the infrared with a wavelength of 1.06 micrometers. These photons are then mixed in a crystal with a large second-order nonlinearity and proper momentum-conservation characteristics to yield green second-harmonic photons of 0.532-μm wavelength. *See* LASER.

Under different momentum-conservation constraints, a similar interaction can take place between two photon fields of different frequency, resulting in photons whose energy or frequency is the sum of those of the original photons. This process is called sum-frequency mixing.

Optical parametric oscillation/amplification. This kind of mixing process occurs when one of the two initial photons has the largest energy and frequency of the three. A high-energy photon and a low-energy photon mix to give a third photon with an energy equal to the difference between the two initial photons. If initially the third field amplitude is zero, it is possible to generate a third field whose frequency is not previously present; in this case the process is called optical parametric oscillation. If the third field exists but at a low level, it can be amplified through the optical parametric amplification process. The existence of this mixing process is also subject to momentum-conservation constraints inside the second-order nonlinear crystal, in a fashion similar to the second harmonic generation process. *See* PARAMETRIC AMPLIFIER.

Third-order devices. Devices based on the third-order nonlinearity involve a process called four-photon (or four-wave) mixing. In this process, three photons are mixed together to create a fourth photon, subject to energy- and momentum-conservation constraints. The four-photon mixing nonlinearity is responsible for the existence of so-called self-action effects where the refractive index and absorption coefficient of a light field are modified by the light field's own presence, for third-harmonic generation and related processes, and for phase-conjugation processes.

Nonlinear self-action effects. In a medium with a third-order nonlinearity, the refractive index and absorption coefficient of a light field present in the medium are modified by the strength of the light intensity. Because the field effectively acts on itself, this interaction is termed a self-action effect. The momentum-conservation constraints are automatically satisfied

because of the degenerate frequencies involved in the interaction. Such an interaction manifests itself by changing the total absorption experienced by the light field as well as by changing the velocity of propagation of the light field. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES.

There are many devices based on the self-action effects. A reverse saturable absorber becomes more opaque because of the nonlinear absorption (also called two-photon adsorption) that it manifests. Refractive-index changes can be used to change the transmission characteristics of resonant cavities and other structures by modifying the effective optical path length (the product of actual structure length times the effective refractive index for the structure) and shifting the cavity resonances to other frequencies. Several nonlinear optical switches have been proposed based upon this resonance-shifting phenomenon. *See* CAVITY RESONATOR; OPTICAL BISTABILITY.

Finally, the self-action effects can be used in a nonlinear waveguide under certain conditions to cancel out the deleterious effects of either group velocity dispersion for short light pulses or diffraction for continuous beams. The special light pulses based upon this cancellation behave somewhat like particles, and are named optical solitons. This is a consequence of the fact that Maxwell's equations, which govern the propagation of light, can be put into a special form called the nonlinear Schrödinger (NLS) equation when the nonlinearity is of self-action form. The NLS equation resembles the Schrödinger equation of quantum mechanics, with the potential term in the latter equation replaced by a nonlinear term proportional to the local intensity of the light field, and possesses soliton solutions. *See* MAXWELL'S EQUATIONS; OPTICAL PULSES; QUANTUM MECHANICS; SCHRÖDINGER'S WAVE EQUATION.

Solitons are interesting types of waves because of the particlelike characteristics they exhibit. They are robust in the presence of large-scale perturbations; even strong interactions (collisions with other solitons) do not modify their shape but only shift their center of gravity by a small amount. These robust characteristics suggest that solitons should have application in telecommunication transmission systems where pulses of light are required to propagate great distances in optical fibers characterized by relatively large values of dispersion and nonlinearity. Optical solitons can, indeed, propagate the thousands of miles necessary for intercontinental fiber links. Various other proposals have also been made for schemes utilizing the interaction properties of solitons in logic operations. These devices are of two primary types, soliton dragging gates and soliton collision gates (see **illus.**). *See* LOGIC CIRCUITS; OPTICAL COMMUNICATIONS; OPTICAL FIBERS; SOLITON.

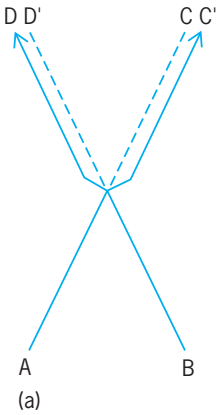

A	B	C	C'	D	D'
0	0	0	0	0	0
1	0	0	1	0	0
0	1	0	0	0	1
1	1	1	0	1	0

(b)

Soliton interaction gate. (*a*) Solitons are initiated at A and B, and detectors are placed at C, C′, D, D′. If only soliton A or soliton B is present, the soliton trajectories will follow the broken lines; if both A and B are present, the solitons will follow the solid lines. (*b*) Truth table. 0 represents the absence of a soliton; 1 represents the presence of a soliton.

Third-harmonic generation. In a third-harmonic generation process, three photons of like energy and frequency are mixed to yield a single photon with three times the energy and frequency of the initial photons. As in second-harmonic generation, this process requires a set of momentum-conservation constraints to be met. Applications of third-harmonic generation are typically in the areas of frequency upconversion. For example, the fundamental Nd:YAG laser emission can be converted to ultraviolet light of 0.355-μm wavelength with the appropriate nonlinear crystal.

Phase conjugation. Phase-conjugation devices make use of a property that third-order media possess whereby energy- and frequency-degenerate photons from two counterpropagating fields are mixed with an incoming photon to yield a photon with exactly the opposite propagation direction and conjugate phase. This phase-conjugate field will pass out of the nonlinear optical device in exactly the direction opposite to the incoming field. Such devices are used in phase-conjugate mirrors, mirrors which have be ability to cancel phase variation in a beam due to, for example, atmospheric turbulence. *See* ADAPTIVE OPTICS; OPTICAL PHASE CONJUGATION.

Materials. The suitability of available nonlinear optical materials is a critical factor in the development of nonlinear optical devices. For certain applications, silica glass fibers may be used. Because of the long propagation distances involved in intercontinental transmission systems, the small size of the optical nonlinearity in silica is not a drawback, although the fibers used in the soliton transmission systems have been especially engineered for reduced dispersion at the transmission wavelength of 1.55 μm. However, generally the nonlinearities available are either below or right at the figure-of-merit thresholds for particular devices. Key materials are semiconductors [such as gallium arsenide (GaAs), zinc selenide (ZnSe), and indium gallium arsenide phosphide (InGaAsP)], certain organic polymeric films, hybrid materials such as semiconductor-doped glasses, and liquid crystals. However, there are several formidable fundamental limits to the development of materials due to the intimate relationship between the absorption and refraction processes that occur in all materials. *See* NONLINEAR OPTICS; OPTICAL MATERIALS.

David R. Andersen

Bibliography. R. W. Boyd, *Nonlinear Optics*, 1992; P. N. Butcher and D. Cotter, *The Elements of Nonlinear Optics*, 1990; M. Schubert and B. Wilhelmi, *Nonlinear Optics and Quantum Electronics*, 1986.

Nonlinear optics

A field of study concerned with the interaction of electromagnetic radiation and matter in which the matter responds in a nonlinear manner to the incident radiation fields. The nonlinear response can result in intensity-dependent variation of the propagation characteristics of the radiation fields or in the creation of radiation fields that propagate at new frequencies or in new directions. Nonlinear effects can take place in solids, liquids, gases, and plasmas, and may involve one or more electromagnetic fields as well as internal excitations of the medium. Most of the work done in the field has made use of the high powers available from lasers. The wavelength range of interest generally extends from the far-infrared to the vacuum ultraviolet, but some nonlinear interactions have been observed at wavelengths extending from the microwave to the x-ray ranges. *See* LASER.

Nonlinear materials. Nonlinear effects of various types are observed at sufficiently high light intensities in all materials. It is convenient to characterize the response of the medium mathematically by expanding it in a power series in the electric and magnetic fields of the incident optical waves. The linear terms in such an expansion give rise to the linear index of refraction, linear absorption, and the magnetic permeability of the medium, while the higher-order terms give rise to nonlinear effects. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES.

In general, nonlinear effects associated with the electric field of the incident radiation dominate over magnetic interactions. The even-order dipole susceptibilities are zero except in media which lack a center of symmetry, such as certain classes of crystals, certain symmetric media to which external forces have been applied, or at boundaries between certain dissimilar materials. Odd-order terms can be nonzero in all materials regardless of symmetry. Generally the magnitudes of the nonlinear susceptibilities decrease rapidly as the order of the interaction increases. Second- and third-order effects have been the most extensively studied of the nonlinear interactions, although effects up to order 30 have been observed in a single process. In some situations, multiple low-order interactions occur, resulting in a very high effective order for the overall nonlinear process. For example, ionization through absorption of effectively 100 photons has been observed. In other situations, such as dielectric breakdown or saturation of absorption, effects of different order cannot be separated, and all orders must be included in the response. *See* ELECTRIC SUSCEPTIBILITY; POLARIZATION OF DIELECTRICS.

Second-order effects. Second-order effects involve a polarization with the dependence $P_{nl}^{(2)} = dE^2$, where E is the electric field of the optical waves and d is a nonlinear susceptibility. The second-order polarization has components that oscillate at sum and difference combinations of the incident frequencies, and also a component that does not oscillate. The oscillating components produce a propagating polarization wave in the medium with a propagation vector that is the appropriate sum or difference of the propagation vectors of the incident waves. The nonlinear polarization wave serves as a source for an optical wave at the corresponding frequency in a process that is termed three-wave parametric mixing.

Phase matching. The strongest interaction occurs when the phase velocity of the polarization wave is the same as that of a freely propagating wave of the same frequency. The process is then said to be phase-matched. Dispersion in the refractive indices, which occurs in all materials, usually prevents phase matching from occurring unless special steps are taken. Phase matching in crystals can be achieved by using noncollinear beams, materials with periodic structures, anomalous dispersion near an absorption edge, or compensation using free carriers in a magnetic field, or by using the birefringence possessed by some crystals. In the birefringence technique, the one used most commonly for second-order interactions, one or two of the interacting waves propagate as an extraordinary wave in the crystal. The phase-matching conditions are achieved by choosing the proper temperature and propagation direction. For a given material these conditions depend on the wavelengths and direction of polarization of the individual waves. The conditions for phase-matched three-wave parametric mixing can be summarized by Eqs. (1) and (2), where the ν's are the frequencies

$$\nu_3 = \nu_1 \pm \nu_2 \tag{1}$$

$$\mathbf{k}_3 = \mathbf{k}_1 \pm \mathbf{k}_2 \tag{2}$$

of the waves and the $\mathbf{k}$'s are the propagation constants. Phase matching increases the power in the generated wave by many orders of magnitude. *See* CRYSTAL OPTICS.

Harmonic generation and frequency mixing. In three-wave sum- and difference-frequency mixing, two incident waves at ν_1 and ν_2 are converted to a third wave at ν_3 according to Eq. (1). The simplest interaction of this type is second-harmonic generation in which $\nu_3 = 2\nu_1$. For this interaction the phase-matching condition reduces to Eq. (3), where $n(\nu_3)$ is the refractive index at the harmonic wavelength, and $n_1(\nu_1)$ and $n_2(\nu_1)$ are the refractive indices of the incident waves at the fundamental frequency.

$$n(\nu_3) = \frac{n_1(\nu_1)}{2} + \frac{n_2(\nu_1)}{2} \tag{3}$$

Second-harmonic generation and second-order frequency mixing have been demonstrated at wavelengths ranging from the infrared to the ultraviolet, generally coinciding with the transparency range

of the nonlinear crystals. Second-harmonic conversion has been used with pump radiation extending from 10 μm (from carbon dioxide lasers) to 217 nm, the phase-matching limit of currently available crystals. Sum-frequency mixing has been used to generate wavelengths as short as 172 nm. Difference-frequency mixing can be used to generate both visible and infrared radiation out to wavelengths of about 2 mm. If one or more of the incident wavelengths is tunable, the generated wavelength will also be tunable, providing a useful source of radiation for high-resolution spectroscopy. Radiation can be converted from longer-wavelength ranges, such as the infrared, to shorter-wavelength regions, such as the visible, where more sensitive and more convenient detectors are available. *See* LASER SPECTROSCOPY.

In principle, 100% of the energy in the incident radiation can be converted to the generated wave in frequency-mixing interactions. In most situations, however, additional linear and nonlinear interactions prevent the conversion efficiency from reaching its theoretical maximum. For second-order interactions the most important limiting effects are absorption of the radiation at either incident or generated frequencies (generally important for generation in the mid- to far-infrared) or failure to achieve or maintain phase matching throughout the interaction (important at shorter wavelengths in the infrared, or in the visible or ultraviolet). Divergent angles, a spread of frequencies, or even minor amounts of phase structure (multiple frequency components) in the incident beams can limit phase matching. Phase matching can also be disturbed by heating of the crystal following absorption of a small amount of the pump radiation. In practice, second-harmonic conversion efficiency of over 90% has been achieved with radiation from high-power pulsed lasers. Conversion of radiation from continuous-wave lasers of over 30% has been obtained when the nonlinear crystal is placed in the laser cavity or in an external resonator.

Parametric generation. In parametric generation, which is the reverse process of sum-frequency mixing, a single input wave at ν_3 is converted to two lower-frequency waves according to the relation $\nu_3 = \nu_1 + \nu_2$. The individual values of ν_1 and ν_2 are determined by the simultaneous satisfaction of the phase-matching condition in Eq. (2). Generally this condition can be satisfied for only one pair of frequencies at a time for a given propagation direction. By changing the phase-matching conditions, the individual longer wavelengths can be tuned. Parametric generation can be used to amplify waves at lower frequencies than the pump wave or, in an oscillator, as a source of tunable radiation in the visible or infrared over a wavelength range similar to that covered in difference-frequency mixing.

Optical rectification. The component of the second-order polarization that does not oscillate produces an electrical voltage. The effect is called optical rectification and is a direct analog of the rectification that occurs in electrical circuits at much lower frequencies. It has been used in conjunction with ultrashort mode-locked laser pulses to produce some of the shortest electrical pulses generated, with durations of less than 250 femtoseconds. *See* OPTICAL PULSES.

Third-order interactions. Third-order interactions give rise to several types of nonlinear effects. Four-wave parametric mixing involves interactions which generate waves at sum- and difference-frequency combinations of the form of Eq. (4) with the corresponding phase-matching condition of Eq. (5). Phase

$$\nu_4 = \nu_1 \pm \nu_2 \pm \nu_3 \tag{4}$$

$$\mathbf{k}_4 = \mathbf{k}_1 \pm \mathbf{k}_2 \pm \mathbf{k}_3 \tag{5}$$

matching in liquids is usually accomplished by using noncollinear pump beams. Phase matching in gases can also be accomplished by using anomalous dispersion near absorption resonances or by using mixtures of gases, one of which exhibits anomalous dispersion. The use of gases, which are usually transparent to longer and shorter wavelengths than are the solids used for second-order mixing, allows the range of wavelengths covered by parametric mixing interactions to be extended considerably. Four-wave mixing processes of the type $\nu_4 = \nu_1 - \nu_2 - \nu_3$ have been used to generate far-infrared radiation out to wavelengths of the order of 25 μm. Sum-frequency mixing and third-harmonic generation have been used to generate radiation extending to 57 nm.

The nonlinear susceptibility is greatly increased, sometimes by four to eight orders of magnitude, through resonant enhancement that occurs when the input frequencies or their multiples or sum or difference combinations coincide with appropriate energy levels in the nonlinear medium. Two-photon resonances are of particular importance since they do not involve strong absorption of the incident or generated waves. Resonant enhancement in gases has allowed tunable dye lasers to be used for nonlinear interactions, providing a source of tunable radiation in the vacuum ultraviolet and in the far-infrared. Such radiation is useful in spectroscopic studies of atoms and molecules.

Just as with three-wave mixing, four-wave sum-frequency generation can be used to convert infrared radiation to the visible, where it is more easily detected. These interactions have been used for infrared spectroscopic studies and for infrared image conversion. *See* INFRARED IMAGING DEVICES; INFRARED SPECTROSCOPY.

Conversion efficiency of the third- and higher-order interactions is generally lower than for the second-order processes because the competing effects are more important. Third- and higher-order interactions are generally limited by linear and nonlinear absorption, Stark shifting of resonant levels, and multiphoton ionization that lead to reduction of the nonlinear susceptibility, population redistribution, absorption of radiation, or breaking of phase matching. In third-order effects, the competing mechanisms are of the same order as the frequency-mixing interaction, leading to a maximum efficiency that de-

pends only on the material constants and detunings from resonance. The maximum possible conversion efficiency for these interactions is of the order of 30%, while the maximum reported is 10%.

Conversion efficiencies for the higher-order interactions are usually lower than for the lower-order ones. The higher-order interactions are thus most useful in extending coherent radiation into a range of wavelengths not reachable with lower-order interactions. Conversion to harmonics of very high order has been achieved with extremely intense optical pulses with durations of the order of several hundred femtoseconds or less, where the energy in the various higher harmonic orders is relatively constant, rather than decreasing with harmonic order. Conversion to harmonic orders of up to 111 and generation of radiation down to 7.26 nm have been observed. At the intensities used in these interactions, generally greater than 10^{15} W/cm^2, perturbation theory breaks down and a more complete theory of the interaction is needed.

Two-wave mixing is an interaction that occurs in photorefractive materials (discussed below) between two waves of the same frequency, in which power is transferred from one wave to another. It can be used to amplify a weaker beam with a stronger one. When the stronger beam has spatial aberrations, it transfers the power to the weaker beam, but not the aberrations, effectively reducing the distortion of the optical beam.

Stimulated scattering. Light can scatter inelastically from fundamental excitations in the medium, resulting in the production of radiation at a frequency that is shifted from that of the incident light by the frequency of the excitation involved. The difference in photon energy between the incident and scattered light is accounted for by excitation or deexcitation of the medium. Some examples are Brillouin scattering from acoustic vibrations; various forms of Raman scattering involving molecular rotations or vibrations, electronic states in atoms or molecules, lattice vibrations or spin waves in solids, spin flips in semiconductors, and electron plasma waves in plasmas; Rayleigh scattering involving density or entropy fluctuations; and scattering from concentration fluctuations in gases. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

At the power levels available from pulsed lasers, the scattered light experiences exponential gain, and the process is then termed stimulated, in analogy to the process of stimulated emission in lasers. In stimulated scattering, the incident light can be almost completely converted to the scattered radiation. Stimulated scattering has been observed for all of the internal excitations listed above. The most widely used of these processes are stimulated Raman scattering and stimulated Brillouin scattering.

Stimulated Raman scattering. In its simplest form, stimulated Raman scattering involves transfer of energy from a laser, or pump, wave to a longer-length wave termed the Stokes wave. If the Stokes wave becomes intense enough, it can produce its own Stokes wave, termed a second Stokes wave, shifted to longer

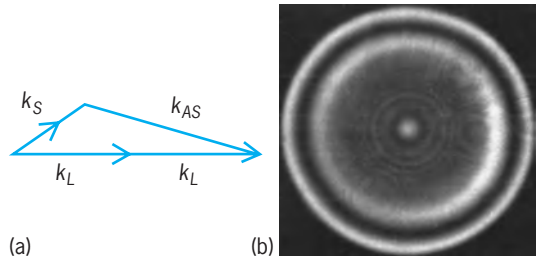

Fig. 1. Anti-Stokes generation by four-wave mixing. (*a*) Phase-matching diagrams showing the direction of propagation of the waves involved. (*b*) Characteristic ring structure of anti-Stokes generation in hydrogen gas. The appearance of two rings is caused by an interaction between stimulated Raman scattering and four-wave mixing that reduces the anti-Stokes generation at exact phase matching. (*From M. D. Duncan et al., Parametric Raman gain suppression in D_2 and H_2, Opt. Lett., 11:803–805, 1986*)

wavelengths from the first Stokes frequency by the material excitation frequency. Light at wavelengths shorter than that of the original laser, termed anti-Stokes waves, can be produced by four-wave mixing processes of the form of Eq. (6), where ν_L

$$\nu_{AS} = 2\nu_L - \nu_S \tag{6}$$

is the laser frequency and ν_S and ν_{AS} are the frequencies of the Stokes and anti-Stokes waves, respectively. These waves are usually emitted in cones because of the need for phase matching (**Fig. 1**). At sufficiently high intensities the stimulated Raman scattering spectrum can consist of many Stokes and anti-Stokes waves.

Stimulated Raman scattering can occur in the forward or backward direction. In the forward direction the most common uses are amplifying weak Stokes signals, generating multiple-wavelength or tunable sources for spectroscopic applications, removing laser-beam aberrations, and combining radiation from multiple lasers. Backward-stimulated Raman scattering is commonly used for pulse compression. *See* RAMAN EFFECT.

Stimulated Brillouin scattering. Stimulated Brillouin scattering involves scattering from acoustic waves and is typically done in liquids or gases. The Stokes wave is usually generated in the backward direction, for which the gain is the highest. Frequency shifts are typically of the order of hundreds of megahertz to tens of gigahertz, depending on the material and the wavelength of the original laser. The most common use of stimulated Brillouin scattering is in phase conjugation. Stimulated Brillouin scattering is also a source of damage in many solids that limits the intensity that can be used.

Self-action and related effects. Nonlinear polarization components at the same frequencies as those in the incident waves can result in effects that change the index of refraction or the absorption coefficient, quantities that are constants in linear optical theory.

Multiphoton absorption and ionization. Materials that are transparent at low optical intensities can undergo an increase in absorption at high intensities. This effect involves simultaneous absorption of two or more photons from one or more incident waves. When the

process involves transitions to discrete upper levels in gases or to conduction bands in solids that obey certain quantum-mechanical selection rules, it is usually termed multiphoton absorption, or, more specifically, n-photon absorption, where n is the number of photons involved. When transitions to the continuum of gases are involved, the process is termed multiphoton ionization.

Saturable absorption. In materials which have a strong linear absorption at the incident frequency, the absorption can decrease at high intensities, an effect termed saturable absorption. Saturable absorbers are useful in operating Q-switched and mode-locked lasers.

Self-focusing and self-defocusing. Intensity-dependent changes in refractive index can affect the propagation characteristics of a laser beam. For many materials the index of refraction increases with increasing optical intensity. If the laser beam has a profile that is more intense in the center than at the edges, the profile of the refractive index corresponds to that of a positive lens, causing the beam to focus. This effect, termed self-focusing, can cause an initially uniform laser beam to break up into many smaller spots with diameters of the order of a few micrometers and intensity levels that are high enough to damage many solids. This mechanism limits the maximum intensities that can be obtained from some high-power pulsed solid-state lasers. *See* KERR EFFECT.

In other materials the refractive index decreases as the optical intensity increases. The resulting nonlinear lens causes the beam to defocus, an effect termed self-defocusing. When encountered in media that are weakly absorbing, the effect is termed thermal blooming. It is prominent, for example, in the propagation of high-power infrared laser beams through the atmosphere.

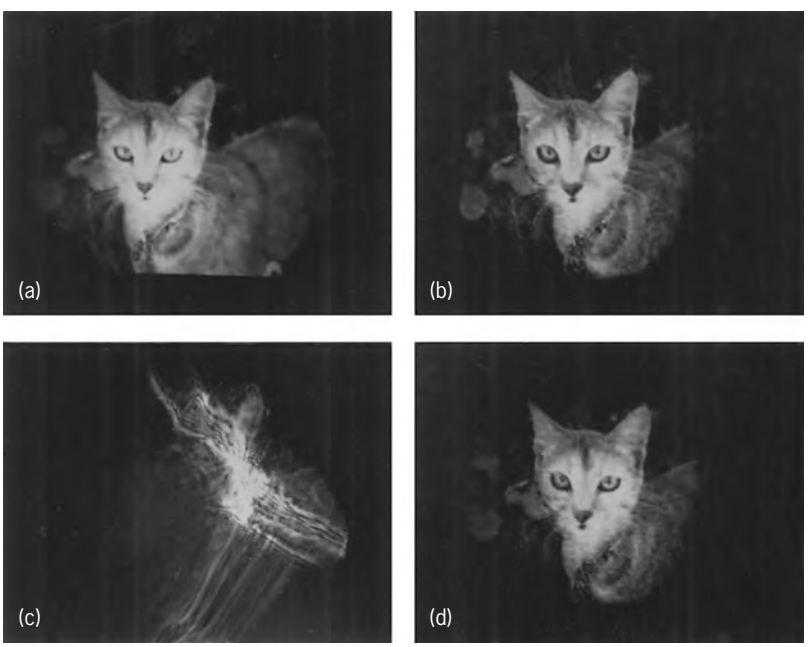

Fig. 2. Use of nonlinear optical phase conjugation for correction of distortions on an image. (*a*) Image obtained with a normal mirror and no distortion. (*b*) Image obtained with an optical conjugator and no distortion. (*c*) Image obtained with a normal mirror when the object is viewed through a distorting piece of glass. (*d*) Image obtained with an optical conjugator when the object is viewed through a distorting piece of glass. (*From J. Feinberg, Self-pumped, continuous-wave phase conjugator using internal reflection, Opt. Lett., 7:486–488, 1982*)

Broadening of spectrum. The nonlinear refractive index can lead to a broadening of the spectrum of pulsed laser fields. The broadened spectrum may extend from the ultraviolet to the infrared and has been used with picosecond-duration pulses for time-resolved spectroscopic studies. It can also be combined with dispersive delay lines to shorten the pulse duration.

Degenerate four-wave mixing. The same interactions that give rise to the self-action effects can also cause nonlinear interactions between waves that are at the same frequency but are otherwise distinguishable, for example, by their direction of polarization or propagation. Termed degenerate four-wave mixing, this interaction gives rise to a number of effects such as amplification of a weak probe wave and phase conjugation (discussed below).

Optical fibers. Propagation through optical fibers can involve several nonlinear optical interactions. Self-phase modulation resulting from the nonlinear index can be used to spread the spectrum, and subsequent compression with diffraction gratings and prisms can be used to reduce the pulse duration. The shortest optical pulses, with durations of the order of 6 femtoseconds, have been produced in this manner.

Linear dispersion in fibers causes pulses to spread in duration and is one of the major limitations on data transmission through fibers. Dispersive pulse spreading can be minimized with solitons, which are specially shaped pulses that propagate long distances without spreading. They are formed by a combined interaction of spectral broadening due to the nonlinear refractive index and anomalous dispersion found in certain parts of the spectrum. *See* SOLITON.

Stimulated Raman scattering from vibrational excitations in the fiber can limit the intensity that can be transmitted in the laser wave. Because of the long distances involved, this can happen at intensities accessible by diode lasers. If two pulses with different wavelengths are propagated, stimulated Raman scattering can be used to amplify the longer-wavelength pulse in the fiber continuously, reducing the need for electronic repeaters in long-distance systems. *See* OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

Control by low-frequency fields. Low-frequency electric, magnetic, or acoustic fields can be used to control the polarization or propagation of an optical wave. These effects, termed electro-, magneto-, or acoustooptic effects, are useful in the modulation and deflection of light waves and are used in information-handling systems. *See* ACOUSTOOPTICS; ELECTROOPTICS; MAGNETOOPTICS; OPTICAL INFORMATION SYSTEMS; OPTICAL MODULATORS.

Coherent effects. Another class of effects involves a coherent interaction between the optical field and an atom in which the phase of the atomic wave functions is preserved during the interaction. These interactions involve the transfer of a significant fraction of the atomic population to an excited state. As

a result, they cannot be described with the simple perturbation expansion used for the other nonlinear optical effects. Rather they require that the response be described by using all powers of the incident fields. These effects are generally observed only for short light pulses, of the order of several nanoseconds or less. In one interaction, termed self-induced transparency, a pulse of light of the proper shape, magnitude, and duration can propagate unattenuated in a medium which is otherwise absorbing.

Other coherent effects involve changes of the propagation speed of a light pulse or production of a coherent pulse of light, termed a photon echo, at a characteristic time after two pulses of light spaced apart by a time interval have entered the medium. Still other coherent interactions involve oscillations of the atomic polarization, giving rise to effects known as optical nutation and free induction decay. Two-photon coherent effects are also possible.

Other applications. In addition to the applications noted above in the discussions of the various interactions, many applications can be accomplished with more than one type of interaction.

Nonlinear spectroscopy. The variation of the nonlinear susceptibility near the resonances that correspond to sum- and difference-frequency combinations of the input frequencies forms the basis for various types of nonlinear spectroscopy which allow study of energy levels that are not normally accessible with linear optical spectroscopy.

Nonlinear spectroscopy can be performed with many of the interactions discussed earlier. Multiphoton absorption spectroscopy can be performed by using two strong laser beams, or a strong laser beam and a weak broadband light source. If two counterpropagating laser beams are used, spectroscopic studies can be made of energy levels in gases with spectral resolutions much smaller than the Doppler limit. Nonlinear optical spectroscopy has been used to identify many new energy levels with principal quantum numbers as high as 150 in several elements. *See* RESONANCE IONIZATION SPECTROSCOPY; RYDBERG ATOM.

Many types of four-wave mixing interactions can also be used in nonlinear spectroscopy. The most widespread of these processes, termed coherent anti-Stokes Raman spectroscopy (CARS), offers the advantage of greatly increased signal levels over linear Raman spectroscopy for the study of certain classes of materials.

Phase conjugation. Optical phase conjugation is an interaction that generates a wave that propagates in the direction opposite to a reference, or signal, wave, and has the same spatial variations in intensity and phase as the original signal wave, but with the sense of the phase variations reversed. Several nonlinear interactions are used to produce phase conjugation.

Optical phase conjugation allows correction of optical distortions that occur because of propagation through a distorting medium (**Fig. 2**). This process can be used for improvement of laser-beam quality, optical beam combining, correction of distortion because of mode dispersion in fibers, and stabilized aiming. It can also be used for neural networks that exhibit learning properties (**Fig. 3**). *See* NEURAL NETWORK; OPTICAL PHASE CONJUGATION.

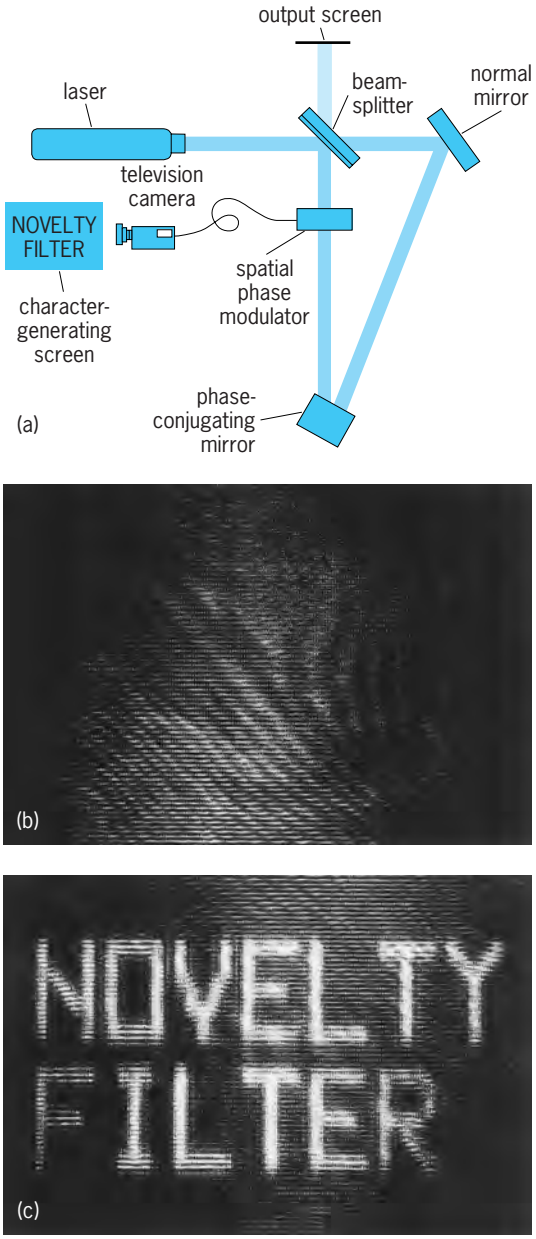

Fig. 3. Novelty filter that uses phase conjugation in a photorefractive material to select changes in a pattern. (*a*) Configuration of components. (*b*) Dark output of the novelty filter that occurs when the filter has adapted to the pattern in the spatial light modulator. (*c*) Output of the filter immediately after a new pattern has been supplied to the spatial light modulator. If the pattern is left unchanged, the output will gradually change to *b*. (*From D. Z. Anderson and M. C. Erie, Resonator memories and optical novelty filters, Opt. Eng., 26:434–444, 1987*)

Optical bistability. Optical elements whose intensity transmission is bistable can be produced with Fabry-Perot etalons that contain materials that exhibit either saturable absorption or an intensity-dependent index of refraction. These devices can be used to perform many of the functions associated with transistors, such as differential gain, switching, and limiting. Hybrid forms, which involve electrooptical

elements, are compatible with integrated optical devices. *See* OPTICAL BISTABILITY.

Time measurements. Various nonlinear interactions, such as second harmonic generation, multiphoton absorption followed by fluorescence, or coherent Raman scattering, have provided a means for measurement of pulse duration and excited-state lifetimes in the picosecond and subpicosecond time regimes.

Self-organizing patterns. Some types of systems, for example those involving nonlinear interactions in an optical cavity, produce optical beams with self-organizing patterns. Such patterns occur as a result of the competition among modes of the cavity in the presence of the nonlinear element. They have potential applications in areas such as information processing and optical displays. John F. Reintjes

Photorefractive effect. The photorefractive effect occurs in many electrooptic materials. A change in the index of refraction in a photorefractive medium arises from the redistribution of charge that is induced by the presence of light. Charge carriers that are trapped in impurity sites in a photorefractive medium are excited into the material's conduction band when exposed to light. The charges migrate in the conduction band until they become retrapped at other sites. The charge redistribution produces an electric field that in turn produces a spatially varying index change through the electrooptic effect in the material. Unlike most other nonlinear effects, the index change of the photorefractive effect is retained for a time in the absence of the light and thus may be used as an optical storage mechanism. The memory time depends on the dark electrical conductivity of the medium. Storage times range from milliseconds to months or years, depending upon the material and the methods employed. *See* TRAPS IN SOLIDS.

Applications. Photorefractive materials are often used for holographic storage. In this case, the index change mimics the intensity interference pattern of two beams of light. Over 500 holograms have been stored in the volume of a single crystal of iron-doped lithium niobate. *See* HOLOGRAPHY.

Photorefractive materials are typically sensitive to very low light levels. Wave-mixing interactions in photorefractive materials are analogous to conventional third-order interactions. Excellent phase-conjugate wave fidelity has been achieved in several photorefractive materials, for example. The photorefractive effect is, however, extremely slow by the standards of optical nonlinearity. Because of their sensitivity, photorefractive materials are increasingly used for image and optical-signal processing applications. *See* IMAGE PROCESSING; NONLINEAR OPTICAL DEVICES. Dana Z. Anderson

Photorefractive polymers. Photorefractive polymers are among the most sensitive nonlinear optical recording materials. They exhibit large refractive index changes when exposed to low-power laser beams. When the optical excitation consists of two interfering coherent beams, the periodic light distribution produces a periodic refractive index modulation. The resulting index change produces a hologram in the volume of the polymer film. The hologram can be reconstructed by diffracting a third laser beam on the periodic index modulation. In contrast to many physical processes that can be used to generate a refractive index change, the photorefractive effect is fully reversible, meaning that the recorded holograms can be erased with a spatially uniform light beam. This reversibility makes photorefractive polymers suitable for real-time optical processing applications.

The mechanism that leads to the formation of a photorefractive index modulation involves the formation of an internal electric field through the absorption of light, the generation of carriers, and the transport and trapping of the carriers over macroscopic distances. The resulting electric field produces a refractive index change through orientational or nonlinear optical effects. While a refractive index change is obtained in crystals through the Pockels effect, most photorefractive polymers exhibit index changes due to orientational Kerr effects. In current polymers, the highest index modulation amplitude is of the order of 1% and the fastest materials have a response time of a few milliseconds.

Due to the transport process, the index modulation amplitude is phase-shifted with respect to the periodic light distribution produced by the interfering optical beams that generate the hologram. This phase shift enables the coherent energy transfer between two beams propagating in a thick photorefractive material. This property, referred to as two-beam coupling, is used to build optical amplifiers.

The photorefractive effect was discovered in inorganic crystals, then studied for several decades mainly in inorganic crystals and semiconductors. The effect was discovered in organic materials in the 1990s. Photorefractive materials are used in holographic storage, nondestructive testing, holographic time gating for image filtering, novelty filtering, phase conjugation, optical correlation, and reconfigurable interconnects. While crystals and semiconductors require advanced fabrication technologies, organic photorefractive polymers have the potential to be mass-produced by techniques such as injection molding. Hence, they can be fabricated into objects of various shapes at low cost. This property together with advances in fabricating integrated laser sources at lower cost can provide great momentum to the development of new optical processing technologies. As media for real-time optical recording and processing, photorefractive polymers are expected to play a major role in these technologies. Bernard Kippelen

Bibliography. N. Bloembergen, *Nonlinear Optics*, 4th ed., World Scientific, 1996; R. W. Boyd, *Nonlinear Optics*, Academic Press, 1992; P. Günter and J. P. Huignard (eds.), *Photorefractive Materials and Their Applications II*, Springer-Verlag, Berlin, 1988; B. Kippelen et al., Infrared photorefractive polymers and their applications for imaging, *Science*, 279:54–57, 1998; K. Meerholz et al., A photorefractive polymer with high optical gain and diffraction efficiency near 100%, *Nature*, 371:497–500, 1994; J. V. Moloney and A. C. Newell, *Nonlinear Optics*,

Addison Wesley Longman, 1992; H. S. Nalwa and S. Miyata (eds.), *Nonlinear Optics of Organic Molecules and Polymers*, CRC Press, Boca Raton, 1997; D. D. Nolte (ed.), *Photorefractive Effects and Materials*, Kluwer Academic, Boston, 1995; J. Reintjes, *Nonlinear Optical Parametric Processes in Liquids and Gases*, 1984; J.-I. Saki, *Phase Conjugate Optics*, McGraw-Hill, 1992; P. Yeh, *Introduction to Photorefractive Nonlinear Optics*, Wiley, 1993; J. Zyss (ed.), *Molecular Nolinear Optics*, Academic Press, San Diego, 1994.

Nonlinear physics

The study of situations where, in a general sense, cause and effect are not proportional to each other; or more precisely, if the measure of what is considered to be the cause is doubled, the measure of its effect is not simply twice as large. Many examples have been known in physics for a long time, and they seemed well understood. Over the last few decades, however, physicists have noticed that this lack of proportionality in some of the basic laws of physics often leads to unexpected complications, if not to outright contradictions. Thus, the term nonlinear physics refers more narrowly to these developments in the understanding of physical reality.

Linear versus nonlinear behavior. The foundations of physics are often expressed in what are called laws, in a misleading analogy to the rules that govern an orderly society. These laws proclaim a mathematical relation between two different kinds of quantities, such as the force on a spring and its length (Hooke's law) or the force on a body and its acceleration (Newton's second law). Each of the two quantities has to be measured in a separate procedure, and their mathematical relation becomes a purely empirical result. Although the two quantities in each of these two laws are proportional to one another, there is an important difference. Newton's second law is correct on a very deep level; that is, the proportionality does not depend on any further assumptions. By contrast, Hooke's law can be verified by experiment only in special circumstances and for special materials. *See* HOOKE'S LAW; NEWTON'S LAWS OF MOTION; PHYSICAL LAW.

The combination of Hooke's law with Newton's second law leads to the linear (or harmonic) oscillator, a common idealization that pervades modern physics. If a body of fixed mass is suspended by a spring and the body is pulled away from its equilibrium position, the spring and the mass perform simple oscillations. The validity of Hooke's law guarantees that the motion is represented by a pure sine curve and that the period of the oscillation does not depend on its amplitude. Such is not the case for an ordinary pendulum, however: Its motion is not purely sinusoidal, and its period increases with the amplitude of the swing. The force that tends to restore the pendulum to its vertical position is not proportional to the angle with the vertical. This oscillator is, therefore, called nonlinear (or anharmonic). *See* ANHARMONIC OSCILLATOR; HARMONIC MOTION; HARMONIC OSCILLATOR; PENDULUM.

The combination of Faraday's law with Ampere's law, as extended by J. C. Maxwell to include so-called electric displacement currents, again gives rise to the phenomenon of a harmonic oscillation inside a cavity. But if the electromagnetic field (or light) is allowed to interact with a collection of excited atoms, a profoundly nonlinear process occurs: The atoms are stimulated to give up their extra energy at a rate that is proportional to the square of the electromagnetic field. The result is a (nonlinear) laser oscillation whose amplitude increases as long as the cavity remains supplied with excited atoms, either from the outside or by the oscillation itself. *See* CAVITY RESONATOR; ELECTROMAGNETIC FIELD; LASER; MAXWELL'S EQUATIONS.

Newton's law of universal gravitation, the foundation of celestial mechanics, is nonlinear: Two bodies attract one another with a force that is proportional to the product of their masses and inversely proportional to the square of their distance (rather than proportional to the distance). Coulomb's law, the base for all atomic, molecular, and mesoscopic physics, is almost identical except that the product of the masses is replaced by the product of the electric charges. The product of the masses (or charges) in the numerator as well as the square of the distance in the denominator can be traced directly to general relativity in the case of gravitation, and to Maxwell's equations for Coulomb's law. The nonlinearity is basic and inescapable. *See* COULOMB'S LAW; GRAVITATION; RELATIVITY.

Linearity in nonlinear systems. Nevertheless, the gravitational or the electrostatic attraction acting between just two bodies leads to simple results, such as Kepler's laws in celestial mechanics and the Balmer-Bohr spectrum of the hydrogen atom in quantum mechanics. *See* ATOMIC STRUCTURE AND SPECTRA; CELESTIAL MECHANICS.

When a large number of particles starts out in a condition of stable equilibrium, the result of small external forces is well-coordinated vibrations of the whole collection, for example, the vibrations of a violin string, or of the electric current in an antenna. Each collective motion acts like an independent oscillator, each with its own frequency. In more complicated systems, many vibrational modes can be active simultaneously without mutual interference. A large dynamical system is, therefore, described in terms of its significant degrees of freedom, thought to be only loosely coupled. The motion of any part of the whole becomes multiperiodic; for example, a water molecule has bending and stretching vibrations with different frequencies, and both are coupled with the rotational motion of the whole molecule. *See* ANTENNA (ELECTROMAGNETISM); DEGREE OF FREEDOM (MECHANICS); MOLECULAR STRUCTURE AND SPECTRA; VIBRATION.

Traditionally, the important degrees of freedom are first identified, their periods are found, and then the extent of their coupling is determined. The mathematical procedures for this type of analysis have

been developed into a fine art, whose most sophisticated and often quite successful products are found particularly in modern statistical mechanics and in high-energy physics. *See* ELEMENTARY PARTICLE; PERTURBATION (ASTRONOMY); PERTURBATION (MATHEMATICS); STATISTICAL MECHANICS.

Failure of perturbation theory. H. Poincaré discovered at the end of the nineteenth century that for many problems, this so-called perturbation theory is not entirely satisfactory. He showed, in the case of the Moon's motion around the Earth, that the disturbance by the Sun is strong enough that this standard mathematical procedure fails. The main culprits are resonances, which occur when the frequencies of different degrees of freedom are combined through their nonlinear coupling. A key nonperturbative phenomenon is known to engineers as phase lock: When different frequencies arise in simple multiples of one another, the whole dynamical system falls into a dynamical trap; and for a continuous range of initial conditions, the interaction changes the frequencies of the individual degrees of freedom sufficiently so as to "lock" the motion into the resonance. *See* ASTEROID; PHASE-LOCKED LOOPS; RESONANCE (ACOUSTICS AND MECHANICS).

Structure of motion in phase space. An example of these effects is provided by a particle of mass m moving in the (x, y) plane. Its state of motion at any time t is uniquely determined by its position (x, y) and its momentum (p_x, p_y); the four-dimensional space with coordinates (x, y, p_x, p_y) is called the phase space. If no work is done by outside forces, the value of the total energy, that is, the sum of the kinetic energy $(p_x^2 + p_y^2)/2m$ and the potential energy $V(x, y)$, stays constant at some value E. The motion at a fixed energy E is best understood if the values of x and p_x are monitored every time the particle crosses the x axis, that is, $y = 0$. The corresponding points are represented in a p_x-versus-x plot, called a surface of section because the flow of the trajectories in phase space at the energy E is intersected by the surface $y = 0$.

If the motion is multiperiodic, these points form a smooth closed curve, and different motions at the same energy E yield a sequence of nested curves. But this type of regular structure in phase space is found only for certain special potentials $V(x, y)$ with a high degree of symmetry. As soon as a small coupling potential $W(x, y)$ is added to $V(x, y)$, many smooth curves dissolve into a disorderly scatter of individual points that cannot be connected in any sensible manner; the corresponding motions are called chaotic. In other parts of the surface of section, an originally smooth curve decays into a sequence of small localized pieces, like a chain of islands (see **illus.**). *See* CHAOS.

KAM theorem. In the 1950s, A. N. Kolmogoroff provided a first account of how the addition of the coupling $W(x, y)$ generates the chaotic regions and islands. This problem was later worked out in detail by V. Arnold and J. Moser to yield the KAM theorem. Each smooth curve in the surface-of-section without the coupling is characterized by two frequencies

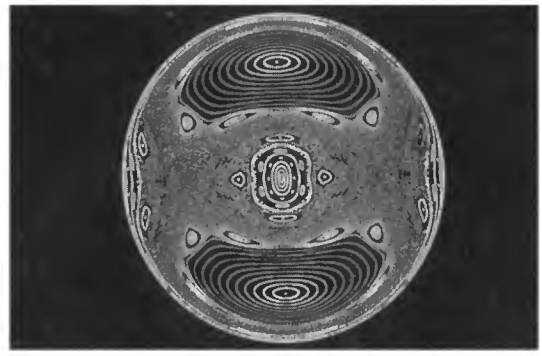

Poincaré surface-of-section, representing motion of an electron in a hydrogen atom in the presence of a strong magnetic field. The section is a slice of four-dimensional phase space representing the electron's position and momentum. Filled regions indicate chaotic behavior. (*Image by Dieter Wintgen*)

whose ratio varies smoothly from curve to curve. The coupling causes phase lock wherever this ratio is a simple rational number, like $^1/_1$, $^1/_2$, $^1/_3$, $^2/_3$, and so forth; the smooth curve breaks up into islands with chaotic regions in between. But the original curves with an irrational frequency ratio, like $\sqrt{2}, \sqrt{3}$, and so forth, remain smooth in spite of the coupling W. The total area in the surface-of-section, outside the remaining smooth curves and the islands, increases continuously with the strength of the coupling; this chaotic area has a complicated fractal structure. *See* FRACTALS.

Implications of nonlinearity. The KAM theorem gives detailed information about the loss of the regular structure as the strength of the coupling increases. It does not say anything, however, about the trajectories in the newly created areas of chaotic behavior. These further investigations are the main goal of such fields as chaos or complexity. The impact of Poincaré's general arguments and the KAM theorem reaches into every area of nonlinear physics. The oldest among them is hydrodynamics, where the phenomenon of turbulent flow has so far resisted any effective control. This is what makes weather prediction so difficult. Signal propagation along the nerves and transmission of pulses through synaptic connections are other well-known nonlinear processes. *See* NEUROBIOLOGY; TURBULENT FLOW.

The two pillars of modern physics, electromagnetism and quantum mechanics, are described, each separately, by linear equations. They become effectively nonlinear, however, in the practical solutions concerned with large atoms or molecules. Moreover, when the two fields are combined into quantum electrodynamics (QED), the coupling between them is nonlinear. Fortunately, this coupling is weak enough for perturbation theory to work very well, and QED has become the inspiration for all the theoretical models in nuclear and high-energy physics. Unfortunately, the coupling in these mathematical models is very strong, so that perturbation theory is useless. At the deepest level of physical reality, there is no escape from nonlinear physics. The manifestations of the nonlinear features in the most basic problems

have barely begun to be accounted for. *See* NONLINEAR ACOUSTICS; NONLINEAR OPTICS; QUANTUM CHROMODYNAMICS; QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; QUANTUM MECHANICS.

Martin C. Gutzwiller

Bibliography. F. Gross, *Relativistic Quantum Mechanics and Field Theory*, 1993; R. C. Hilborn, *Chaos and Nonlinear Dynamics*, 2d ed., 2001; L. Lam (ed.), *Nonlinear Physics for Beginners*, 1998; R. Z. Sagdeev, D. A. Usikov, and G. M. Zaslavsky, *Nonlinear Physics from the Pendulum to Turbulence and Chaos*, 1988; S. H. Strogatz, *Nonlinear Dynamics and Chaos*, 1994; M. Tabor, *Chaos and Integrability in Nonlinear Dynamics*, 1989.

Nonlinear programming

The area of applied mathematics and operations research concerned with finding the largest or smallest value of a function subject to constraints or restrictions on the variables of the function. Nonlinear programming is sometimes referred to as nonlinear optimization.

A useful example concerns a power plant that uses the water from a reservoir to cool the plant. The heated water is then piped into a lake. For efficiency, the plant should be run at the highest possible temperature consistent with safety considerations, but there are also limits on the amount of water that can be pumped through the plant, and there are ecological constraints on how much the lake temperature can be raised. The optimization problem is to maximize the temperature of the plant subject to the safety constraints, the limit on the rate at which water can be pumped into the plant, and the bound on the increase in lake temperature.

The nonlinear programming problem refers specifically to the situation in which the function to be minimized or maximized, called the objective function, and the functions that describe the constraints are nonlinear functions. Typically, the variables are continuous; this article is restricted to this case.

The general topic of mathematical programming includes many important special cases, each with its own methods and techniques. If the objective function and the constraint functions are linear, the problem is a linear programming one; if the objective function is quadratic and the constraints are linear, the problem is a quadratic programming one. If there are no constraints, the problem is unconstrained. If the variables are restricted to a set of discrete or integer values, the problem is referred to as an integer programming problem. *See* LINEAR PROGRAMMING.

Researchers in nonlinear programming consider both the theoretical and practical aspects of these problems. Theoretical issues include the study of algebraic and geometric conditions that characterize a solution, as well as general notions of convexity that determine the existence and uniqueness of solutions. Among the practical questions that are addressed are the mathematical formulation of a specific problem and the development and analysis of algorithms for finding the solution of such problems.

General theory. The general nonlinear programming problem can be stated as that of minimizing a scalar-valued objective function $f(\mathbf{x})$ over all vectors $\mathbf{x}$ satisfying a set of constraints. The constraints are in the form of general nonlinear equations and inequalities. Mathematically, the nonlinear programming problem may be expressed as notation (1),

$$\begin{aligned} &\text{minimize } f(\mathbf{x}) \text{ with respect to } \mathbf{x} \\ &\quad \text{subject to: } g_i(\mathbf{x}) \le 0, \quad i = 1, 2, \ldots, m \\ &\qquad\qquad h_j(\mathbf{x}) = 0, \quad j = 1, 2, \ldots, p \end{aligned} \tag{1}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ are the variables of the problem, f is the objective function, $g_i(\mathbf{x})$ are the inequality constraints, and $h_j(\mathbf{x})$ are the equality constraints. This formulation is general in that the problem of maximizing $f(\mathbf{x})$ is equivalent to minimizing $-f(\mathbf{x})$ and a constraint $g_i(\mathbf{x}) \ge 0$ is equivalent to the constraint $-g_i(\mathbf{x}) \le 0$.

Many modern approaches for solving (1) rely on exploiting the equations that characterize a solution. The most familiar characterization is for the case when the functions f, g_i, and h_j are all differentiable. This case can be described more easily by considering the problem in which there are only equality constraints, that is, when $m = 0$. The following notation is employed. If $s(\mathbf{x})$ is a scalar function of $\mathbf{x}$, then $\nabla s(\mathbf{x})$ denotes the gradient vector of $s(\mathbf{x})$, that is, the vector of partial derivatives of s with respect to the components of $\mathbf{x}$. Thus $(\nabla s(\mathbf{x}))_i = \partial s(\mathbf{x})/\partial x_i$. Likewise, if $\mathbf{h}(\mathbf{x})$ denotes the vector function whose ith component is $h_i(\mathbf{x})$, then $\nabla \mathbf{h}(\mathbf{x})$ denotes the $(n \times p)$ matrix whose ith column is $\nabla h_i(\mathbf{x})$. For any two vectors $\mathbf{a}$ and $\mathbf{b}$, $\mathbf{a}^t$ denotes the transpose of the vector $\mathbf{a}$ and $\mathbf{a}^t\,\mathbf{b}$ denotes the standard dot product (or inner product). *See* CALCULUS OF VECTORS.

The characterization is provided in terms of the lagrangian function defined by Eq. (2),

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{i=1}^{p} \lambda_i h_i(\mathbf{x}) = f(\mathbf{x}) + \mathbf{h}(\mathbf{x})^t \boldsymbol{\lambda} \tag{2}$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_p)$ is the vector of Lagrange multipliers. It follows that Eq. (3)

$$\nabla L(\mathbf{x}^{(*)}, \boldsymbol{\lambda}) = \nabla f(\mathbf{x}^{(*)}) + \nabla \mathbf{h}(\mathbf{x}^{(*)})\boldsymbol{\lambda} \tag{3}$$

is satisfied. If the vector $\mathbf{x}^{(*)}$ solves the nonlinear programming problem (1) when $m = 0$, then there exist values of the Lagrange multipliers such that Eqs. (4) are satisfied.

$$\begin{aligned} \nabla L(\mathbf{x}^{(*)}, \boldsymbol{\lambda}) &= 0 \\ h_j(\mathbf{x}^{(*)}) = 0, \quad j &= 1, \ldots, p \end{aligned} \tag{4}$$

Other conditions are needed to complete the theory, but Eq. (4) provides the basis for an important class of methods of solving the nonlinear programming problem. (This class is described below.) In the unconstrained case, Eq. (4) becomes the condition that if $\mathbf{x}^{(*)}$ is a solution then $\nabla f(\mathbf{x}^{(*)}) = 0$. There may be no points satisfying all of the constraints, in which case the problem is said to be inconsistent.

The Lagrange multipliers $\boldsymbol{\lambda}$, also called dual variables, have an important interpretation in the nonlinear programming problem: They give the sensitivity of the objective function value to the constraints.

Computational methods. Since general nonlinear equations cannot be solved in closed form, iterative methods must be used. Such methods generate a sequence of approximations, or iterates, that will converge to a solution under specified conditions. Newton's method is one of the best-known methods and is the basis for many of the fastest methods for solving the nonlinear programming problem.

If $F(x) = 0$ is one equation in one unknown, Newton's method for solving this problem produces the next iterate, $x^{(k+1)}$, from the current iterate, $x^{(k)}$, by formula (5),

$$x^{(k+1)} = x^{(k)} - \frac{F\left(x^{(k)}\right)}{F'\left(x^{(k)}\right)} \tag{5}$$

where $F'(x^{(k)})$ is the derivative of F with respect to x evaluated at $x^{(k)}$. If $\mathbf{F}(\mathbf{x}) = 0$ is now n equations in n unknowns, then $\mathbf{F}'(\mathbf{x}^{(k)})$ becomes $\nabla\mathbf{F}(\mathbf{x}^{(k)})$, the $(n \times n)$ matrix of partial derivatives defined by Eq. (6),

$$\nabla\mathbf{F}(\mathbf{x})_{ij} = \frac{\partial F_i(\mathbf{x})}{\partial x_j} \tag{6}$$

evaluated at $\mathbf{x}^{(k)}$. That is, the ij component of $\nabla\mathbf{F}(\mathbf{x})$ is the partial derivative of the ith component of $\mathbf{F}$ with respect to jth element of $\mathbf{x}$. The analog to the division by $F'(\mathbf{x}^{(k)})$ in Eq. (5) is the multiplication by the inverse of $\nabla\mathbf{F}(\mathbf{x}^{(k)})$ [if its determinant is not zero], so that Eq. (5) becomes Eq. (7).

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \left[\nabla\mathbf{F}\left(\mathbf{x}^{(k)}\right)\right]^{-1}\mathbf{F}\left(\mathbf{x}^{(k)}\right) \tag{7}$$

If the initial approximation, $\mathbf{x}^{(0)}$, is sufficiently good and other reasonable conditions hold, then Newton's method converges rapidly to a solution of the system of equations. Modifications need to be made, however, to create an efficient and robust method that can be used as a general tool for solving the nonlinear programming problem. These modifications involve the development of tests that assess $\mathbf{x}^{(k+1)}$ to determine if it is an improvement over $\mathbf{x}^{(k)}$, and procedures to adjust the steps to ensure that improvement occurs. Since the computation of $\nabla\mathbf{F}(\mathbf{x}^{(k)})$ is often the most expensive part of the calculation, other modifications seek to approximate this matrix by computationally inexpensive alternatives. Research continues on these and other ideas, especially as they relate to large-scale problems.

To solve the nonlinear programming problem when there are no inequality constraints, the nonlinear system of equations (4) can be attacked by Newton's method as just described. The nonlinear system is in terms of the variables $\mathbf{x}$ and the dual variables $\boldsymbol{\lambda}$, and is therefore $(n + p)$ equations in $(n + p)$ unknowns, namely, Eqs. (8).

$$\mathbf{F}(\mathbf{x}, \boldsymbol{\lambda}) = \begin{cases} \nabla L(\mathbf{x}, \boldsymbol{\lambda}) = 0 \\ \mathbf{h}(\mathbf{x}) = 0 \end{cases} \tag{8}$$

The formulas involved in the computation of the matrix $\nabla\mathbf{F}(\mathbf{x}^{(k)}, \mathbf{x}^{(k)})$ are straightforward and require the evaluation of the matrix of second partial derivatives of $L(\mathbf{x}, \boldsymbol{\lambda})$ with respect to $\mathbf{x}$. This matrix is denoted by $\nabla^2 L(\mathbf{x}, \boldsymbol{\lambda})$. Thus, Eq. (9) holds,

$$\nabla\mathbf{F}\left(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^{(k)}\right) = \begin{bmatrix} \nabla^2 L\left(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^{(k)}\right) & \nabla\mathbf{h}\left(\mathbf{x}^{(k)}\right) \\ \nabla\mathbf{h}\left(\mathbf{x}^{(k)}\right)^{\mathrm{t}} & 0 \end{bmatrix} \tag{9}$$

and the Newton iteration for Eq. (8) is then given by Eq. (10).

$$\begin{bmatrix} \mathbf{x}^{(k+1)} \\ \boldsymbol{\lambda}^{(k+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{x}^{(k)} \\ \boldsymbol{\lambda}^{(k)} \end{bmatrix} - \left[\nabla\mathbf{F}\left(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^{(k)}\right)\right]^{-1} \begin{bmatrix} \nabla L\left(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^{(k)}\right) \\ \mathbf{h}\left(\mathbf{x}^{(k)}\right) \end{bmatrix} \tag{10}$$

A means of handling inequality constraints completes this description. First, it can be shown that Eq. (10) is equivalent to solving the quadratic programming problem, expressed by notation (11),

$$\begin{aligned} &\text{minimize } \nabla f\left(\mathbf{x}^{(k)}\right)^{\mathrm{t}}\delta + \frac{1}{2}\delta^{\mathrm{t}}\nabla^2 L\left(\mathbf{x}^{(k)}, \boldsymbol{\lambda}^{(k)}\right)\delta \\ &\qquad\qquad \text{with respect to } \delta \\ &\text{subject to: } \nabla\mathbf{h}\left(\mathbf{x}^{(k)}\right)^{\mathrm{t}}\delta + \mathbf{h}\left(\mathbf{x}^{(k)}\right) = 0 \end{aligned} \tag{11}$$

for δ and setting $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \delta$ and $\boldsymbol{\lambda}^{(k+1)}$ to the Lagrange multiplier associated with problem (11). Formulated in this way, the method is known as sequential quadratic programming (SQP). At each step of an SQP algorithm, one formulates and solves a quadratic programming approximation to the nonlinear programming problem and uses its solution to construct the next iterate. To incorporate inequality constraints into this algorithm, the constraints of the quadratic programming problem are augmented with Eqs. (12).

$$\nabla\mathbf{g}\left(\mathbf{x}^{(k)}\right)^{\mathrm{t}}\delta + \mathbf{g}\left(\mathbf{x}^{(k)}\right) \le 0 \tag{12}$$

Solving inequality-constrained quadratic programs is harder than solving quadratic programs with equality constraints only, but effective methods exist.

Sequential quadratic programming continues to be one of the most powerful and successful general approaches for solving the nonlinear programming problem with both equality and inequality constraints. Algorithms based on sequential quadratic programming have been used to solve a wide variety of science, engineering, and management problems. Research is continuing on sequential quadratic programming and related methods with an emphasis on large-scale problems. Research is also continuing on methods for problems in which the functions are not differentiable. *See* OPERATIONS RESEARCH; OPTIMIZATION.

Paul T. Boggs

Bibliography. S. G. Nash and A. Sofer, *Linear and Nonlinear Programming*, 1995; J. Nocedal and S. J. Wright, *Numerical Optimization*, 1999.

Nonmetal

The elements are conveniently, but arbitrarily, divided into metals and nonmetals. The nonmetals do not conduct electricity readily, are not ductile, do

not have a complex refractive index, and in general have high ionization potentials. The nonmetals vary widely in physical properties. Hydrogen is a colorless permanent gas; bromine is a dark-red, volatile liquid; and carbon, as diamond, is a solid of great hardness and high refractive index. If the periodic table is divided diagonally from upper left to lower right, all the nonmetals are on the right-hand side of the diagonal. Examples of elements which do not fit neatly into this useful but arbitrary classification are tin, which exists in two allotropic modifications, one definitely metallic and the other with many properties of a nonmetal, and tellurium and antimony. Such elements are called metalloids. *See* IONIZATION POTENTIAL; METAL; METALLOID; PERIODIC TABLE. Thomas C. Waddington

Non-newtonian fluid

A fluid that departs from the classic linear newtonian relation between stress and shear rate. In a strict sense, a fluid is any state of matter that is not a solid, and a solid is a state of matter that has a unique stress-free state. A conceptually simpler definition is that a fluid is capable of attaining the shape of its container and retaining that shape for all time in the absence of external forces. Therefore, fluids encompass a wide variety of states of matter including gases and liquids as well as many more esoteric states (for example, plasmas, liquid crystals, and foams). *See* FLUIDS; FOAM; GAS; LIQUID; LIQUID CRYSTALS; PLASMA (PHYSICS).

A newtonian fluid is one whose mechanical behavior is characterized by a single function of temperature, the viscosity, a measure of the "slipperiness" of the fluid. For the example of **Fig. 1**, where a fluid is sheared between a fixed plate and a moving plate, the viscosity is given by Eq. (1). Thus, as the viscosity

$$\text{Viscosity} = \frac{\text{force/area}}{\text{velocity/height}} \qquad (1)$$

of a fluid increases, it requires a larger force to move the top plate at a given velocity. For simple, newtonian fluids, the viscosity is a constant dependent on only temperature; but for non-newtonian fluids, the viscosity can change by many orders of magnitude as the shear rate (velocity/height in Fig. 1) changes. *See* NEWTONIAN FLUID; VISCOSITY.

Many of the fluids encountered in everyday life (such as water, air, gasoline, and honey) are adequately described as being newtonian, but there are even more that are not. Common examples include mayonnaise, peanut butter, toothpaste, egg whites, liquid soaps, and multigrade engine oils. Other examples such as molten polymers and slurries are of considerable technological importance. A distinguishing feature of many non-newtonian fluids is that they have microscopic or molecular-level structures that can be rearranged substantially in flow. *See* PARTICLE FLOW; POLYMER.

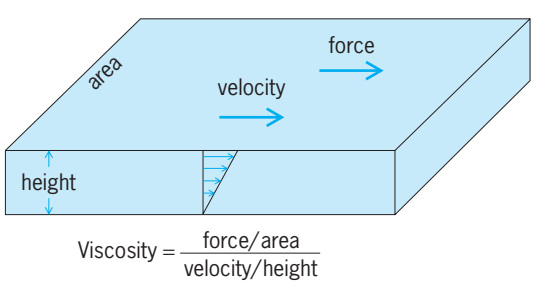

Fig. 1. Steady shear flow of a fluid between a fixed plate and a parallel moving plate, illustrating the concept of viscosity.

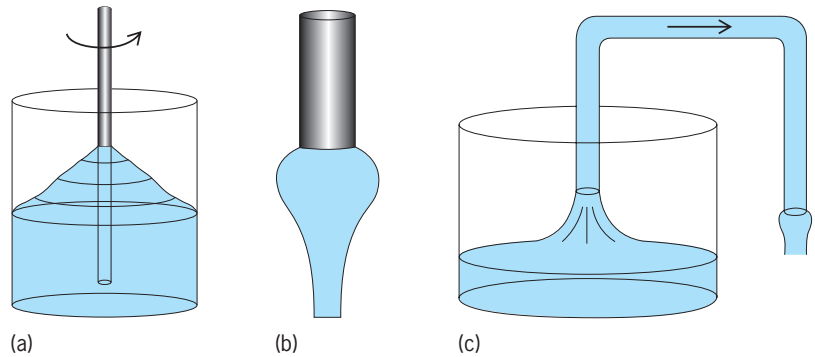

Fig. 2. Examples of the counterintuitive behavior display by non-newtonian fluids. (*a*) Weissenberg effect. (*b*) Die swell. (*c*) Tubeless siphon.

Counterintuitive behavior. Our intuitive understanding of how fluids behave and flow is built primarily from observations and experiences with newtonian fluids. However, non-newtonian fluids display a rich variety of behavior that is often in dramatic contrast to these expectations. For example, an intuitive feel for the slipperiness of fluids can be gained from rubbing them between the fingers. Furthermore, the slipperiness of water, experienced in this way, is expected to be the same as the slipperiness of automobile tires on a wet road. However, the slipperiness (viscosity) of many non-newtonian fluids changes a great deal depending on how fast they move or the forces applied to them.

Intuitive expectations for how the surface of a fluid will deform when the fluid is stirred (with the fluid bunching up at the wall of the container) are also in marked contrast to the behavior of non-newtonian fluids. When a cylindrical rod is rotated inside a container of a newtonian fluid, centrifugal forces cause the fluid to be higher at the wall. However, for non-newtonian fluids, the normal stress differences cause the fluid to climb the rod; this is called the Weissenberg effect (**Fig. 2***a*). Intuitive understanding about the motion of material when the flow of a fluid is suddenly stopped, for example by turning off a water tap, is also notably at odds with the behavior of non-newtonian fluids. *See* CENTRIFUGAL FORCE.

A non-newtonian fluid also displays counterintuitive behavior when it is extruded from an opening. A newtonian fluid tapers to a smaller cross section as it leaves the opening, but the cross section for a non-newtonian fluid first increases before it eventually tapers. This phenomenon is called die swell (Fig. 2*b*). *See* NOZZLE.

When a newtonian fluid is siphoned and the fluid level goes below the entrance to the siphon tube, the

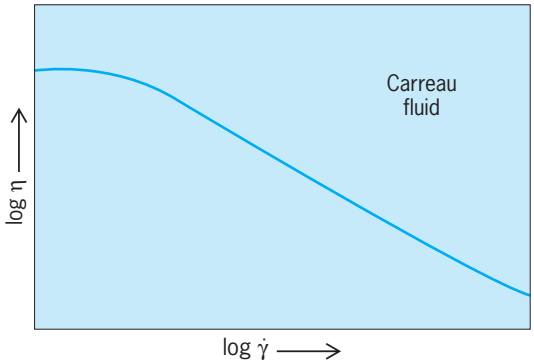

Fig. 3. Typical dependence of the viscosity (ϵ) on shear rate (γ) for a non-newtonian fluid. The functional from shown is called the Carreau model.

siphoning action stops. For many non-newtonian fluids, however, the siphoning action continues as the fluid climbs from the surface and continues to enter the tube. This phenomenon is called the tubeless siphon (Fig. 2*c*).

Non-newtonian viscosity. For newtonian fluids, the viscosity is independent of the speed at which the fluid is caused to move. However, for non-newtonian fluids, the viscosity can change by many orders of magnitude as the speed (velocity gradient) changes. Typically, the viscosity (η) of these fluids is given as a function of the shear rate ($\dot{\gamma}$), which is a generalization of the ratio velocity/height in Fig. 1. A common dependence for this function is given in **Fig. 3**. In this particular model (called the Carreau model), the viscosity as a function of shear rate is given by Eq. (2).

$$\eta = \eta_\infty + (\eta_0 - \eta_\infty)\left[1 + (\lambda\dot{\gamma})^2\right]^{(n-1)/2} \qquad (2)$$

This model involves four parameters: η_0, the viscosity at very small shear rates; η_∞, the viscosity at very large shear rates; λ, which describes the shear rate at which the curve begins to decrease; and n, which determines the rate of viscosity decrease. Figure 3 shows the viscosity for a fluid that shear-thins (that is, the viscosity decreases as the shear rate increases, n is less than one, and η_0 is greater than η_∞). For other non-newtonian fluids, the viscosity might increase as the shear rate increases (shear-thickening fluids, for which n is greater than one and η_0 is less than η_∞).

Nonlinear effects. Although shear-rate-dependent viscosity is the most important non-newtonian effect for many engineering applications of these fluids, it is not by itself capable of describing any of the phenomena depicted in Fig. 2. Both the Weissenberg effect (rod climbing) and die swell are generally attributed to nonlinear effects, by which is meant stresses that are present in non-newtonian fluids and not present in newtonian fluids. The stress is a generalization of the concept of pressure; at any point in the fluid there are nine components of the stress, τ_{ij}, where i and j take on the three coordinate directions. Then τ_{ij} is the force per unit area in the j direction on a surface of constant i (for example, τ_{xy} is the force per unit area in the y direction on a surface of constant x). Most significant among the anomalous stresses in non-newtonian fluids is the first normal stress difference, which causes a tension in the direction of fluid flow. (There is also a second normal stress difference, which tends to be smaller and less significant than the first.) These stresses are generally caused by the flow trying to deform the structure of the material and the microstructure of the material resisting this deformation. The most common and pronounced result in a shear flow is a tension in the direction of the flow trying to resist flow. The Weissenberg effect is easily understood in terms of the first normal stress difference. Here the flow direction is circular around the rod. The tension pulls the fluid inward toward the rod, and the fluid must rise because it cannot go downward through the bottom of the container. *See* STRESS AND STRAIN.

Viscoelasticity. Perhaps the most striking behavior of non-newtonian fluids is a consequence of their viscoelasticity. Solids can be thought of as having perfect memory. If they are deformed through the action of a force, they return to their original shape when the force is removed. This happens when a rubber ball bounces; the ball is deformed as it hits a surface, but the rubber remembers its undeformed spherical shape. Recovery of the shape causes the ball to bounce back. In contrast, newtonian fluids have no memory; when a force is removed, they retain their condition at the time the force is removed (or continue moving as the result of inertia). When a newtonian fluid is dropped onto a surface, it does not bounce. Non-newtonian fluids are viscoelastic in the sense that they have fading memory. If a force is removed shortly after it is applied, the fluid will remember its undeformed shape and return toward it. However, if the force is applied on the fluid for a long time, the fluid will eventually forget its undeformed shape. If a sample of a non-newtonian fluid is dropped onto a surface, it will bounce like a ball. However, if the fluid is simply placed on the surface, it will flow smoothly. Viscoelasticity is frequently the cause of many of the secondary flows that are observed for non-newtonian fluids. These are fluid motions that are small for newtonian fluids (for example, swirling motions) but can become dominant for non-newtonian fluids. *See* ELASTICITY.

Constitutive relations. Analysis of fluid flow operations is typically performed by examining local conservation relations—conservation of mass, momentum (Newton's second law), and energy. This analysis requires material-specific information (for example, the relation between density, pressure, and temperature) that is collectively known as constitutive relations. The science devoted to obtaining suitable constitutive equations for description of the behavior of non-newtonian fluids is called rheology. The most important constitutive equation for fluid mechanics is that relating the stress in the fluid to the kinematics of the motion (that is, the velocity, the derivatives of the velocity with respect to position, and the time history of the velocity).

Two general methods exist for formulation of appropriate constitutive equations: (1) empiricism

combined with flow measurements, and (2) molecular or structural theories. A wide variety of standard experiments have been designed to probe the non-newtonian behavior of fluids and to test the predictions of constitutive equations. Examples include steady shear flow as in Fig. 1, with the upper plate moving at constant velocity; small-amplitude oscillatory shear flow, where the velocity of the upper plate depicted in Fig. 1 is caused to oscillate sinusoidally in time; and the sudden inception of shear flow, where the upper plate in Fig. 1 instantaneously begins moving at constant velocity. In the latter two experiments, the stress (force) is recorded as a function of time. These experiments provide useful probes of the viscoelasticity of the fluid. Another category of flow experiment involves elongation, wherein the fluid is caused to move so that fluid elements are stretched. In addition, other experiments have been used where the fluid is caused to move in more complicated ways. All of these experiments are used to define material functions that characterize the response of the fluid (for example, the viscosity of Fig. 3).

Generalized newtonian fluid. The simplest constitutive equation (obtained empirically from examination of experimental viscosity data) for a non-newtonian fluid is that for the generalized newtonian fluid, given by Eq. (3), where γ_{ij}, given by Eq. (4), is called the

$$\tau_{ij} = -\eta(\dot{\gamma})\dot{\gamma}_{ij} \tag{3}$$

$$\dot{\gamma}_{ij} = \frac{\partial v_j}{\partial x_i} + \frac{\partial v_i}{\partial x_j} \tag{4}$$

rate-of-strain tensor, and $\eta(\dot{\gamma})$ is the viscosity as a function of shear rate. Here, v_i is the ith component of the velocity vector, and the x_i are the position coordinates. For the Carreau fluid, $\eta(\dot{\gamma})$ is given by Eq. (2).

The generalized newtonian fluid constitutive equation is not capable of describing viscoelasticity or any effects of normal stress differences. The simplest constitutive equation that is capable of describing viscoelasticity is the Maxwell model (named after James C. Maxwell). In the Maxwell model, Eq. (5)

$$\tau_{ij} = -G \int_{-\infty}^{t} e^{-(t-t')/\lambda} \dot{\gamma}_{ij}(t')dt' \tag{5}$$

expresses the stresses as time integrals over the history of the velocity derivatives, where G (the shear modulus) and λ (the relaxation time) are material-dependent constants, and the time integral extends over all past time. It is straightforward to see how the Maxwell model incorporates the memory of the fluid; the stresses at the current time depend on the velocity derivatives at past times, and the memory fades because of the exponential that becomes smaller at more distant past times. For smaller values of λ the memory fades more rapidly, and for larger values of λ the fluid behaves more elastically (remembers better). The Maxwell model has been modified to describe fluid elasticity more accurately by writing the stress as a sum of integrals each with different values for G and λ, but the model still suffers from the serious deficit in predicting that the viscosity is constant (that is, independent of shear rate).

Requirements and generalizations. Constitutive equations, whether formulated empirically or as the result of a structural theory, are generally required to obey several fundamental postulates, including frame invariance and objectivity. Frame invariance specifies that the stress in a material should not depend on the measuring system (for example, the coordinate system) that is used to quantify it. From a practical point of view, this means that the mathematical language of tensor calculus must be used for the formulation. Objectivity is a more confusing concept, but it means loosely that the stress in a material must be independent of any rigid-body motion of the material. *See* TENSOR ANALYSIS.

Unfortunately, the Maxwell model as formulated above does not satisfy the objectivity postulate. However, it has been modified in a variety of ways by using more complicated descriptions of the fluid motion to make it objective. It has also been modified in other ways to incorporate shear-rate-dependent viscosity and normal stress effects. In addition, a large class of constitutive equations has been developed that expresses the stress as the solution of a differential equation (as opposed to a time integral), and these are often easier to use in analyzing specific flows.

Although the non-newtonian behavior of many fluids has been recognized for a long time, the science of rheology is, in many respects, still in its infancy, and new phenomena are constantly being discovered and new theories proposed. Advancements in computational techniques are making possible much more detailed analyses of complex flows and more sophisticated simulations of the structural and molecular behavior that gives rise to non-newtonian behavior. Engineers, chemists, physicists, and mathematicians are actively pursuing research in rheology, particularly as more technologically important materials are found to display non-newtonian behavior. *See* FLUID FLOW; FLUID-FLOW PRINCIPLES; RHEOLOGY.

John M. Wiest

Nonrelativistic quantum theory

The modern theory of matter and its interaction with electromagnetic radiation, applicable to systems of material particles which move slowly compared to the velocity of light and which are neither created nor destroyed.

This article details the formal structure of quantum theory, which can be summarized in the form of postulates as unequivocal as are Newton's laws of classical mechanics; a less logically rigorous presentation is adopted here, however. Even so, the reader unfamiliar with quantum theory is advised to read first another article (*see* QUANTUM MECHANICS) which more qualitatively discusses the novel (from the standpoint of classical physics) features of nonrelativistic quantum theory.

That article and this one attempt to make convincing the thesis that the formalism of nonrelativistic quantum theory is an unarbitrary and physically reasonable extension of the older classical theories. Belief in quantum theory stems as much from acceptance of this thesis as from the broad range, barely hinted at in these articles, of successful application of the theory. For added details concerning special formal topics *see* MATRIX MECHANICS; PERTURBATION (QUANTUM MECHANICS).

For generalizations of nonrelativistic quantum theory to relativistic particles (particles with speeds close to the velocity of light), or to systems in which particle creation and destruction can occur, *see* QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; RELATIVISTIC QUANTUM THEORY.

Wave Function and Probability Density

Basic to quantum mechanics is the belief that the wave properties of matter are associated with a function, the wave function, obeying an equation called a wave equation. The simplest possible wave in three-dimensional space is a so-called scalar wave, in which the wave disturbance is wholly characterized by a single function $\psi(x,y,z,t) \equiv \psi(\mathbf{r},t)$. It is natural, therefore, to postulate that a wave function $\psi(x,y,z,t)$ provides a complete description of the simplest possible physical system, namely, a single particle moving in a force field specified by a potential $V(\mathbf{r})$. It is further postulated that $|\psi(\mathbf{r},t)|^2$, which classically would be proportional to the wave intensity, is the probability density; that is, $|\psi(\mathbf{r},t)|^2 d\mathbf{r}$ is the probability, at time t, of finding the particle in the volume $dxdydz \equiv d\mathbf{r}$ of space lying between x and $x + dx$, y and $y + dy$, z and $z + dz$.

There is the obvious generalization that a wave function $\psi(\mathbf{r}_1, \ldots, \mathbf{r}_g t)$ will completely describe a system of g particles, with $|\psi(\mathbf{r}_1, \ldots, \mathbf{r}_g, t)|^2 d\mathbf{r}_1 \ldots d\mathbf{r}_g$ the probability at time t of simultaneously finding particle 1 in the volume element $d\mathbf{r}_1 = dx_1 dy_1 dz_1, \ldots,$ particle g in $d\mathbf{r}_g$. Moreover, for a system of g distinguishable particles, the probability $P_j(\mathbf{r})d\mathbf{r}$ of finding particle j in the volume element $d\mathbf{r}$ at $\mathbf{r} \equiv x,y,z$ of physical space is given by Eq. (1), where

$$P_j(\mathbf{r})d\mathbf{r} = d\mathbf{r} \int \{d\mathbf{r}_1 \cdots d\mathbf{r}_{j-1}, d\mathbf{r}_{j+1} \cdots d\mathbf{r}_g \times |\psi(\mathbf{r}_1, \ldots, \mathbf{r}_{j-1}, \mathbf{r}, \mathbf{r}_{j+1}, \ldots, \mathbf{r}_g)|^2\} \quad (1)$$

$|\psi(\mathbf{r}^1, \ldots, \mathbf{r}_g)|^2$ is integrated over all positions of particles 1 to $j - 1$ and $j + 1$ to g, with $\mathbf{r}_j$ put equal to $\mathbf{r}$.

Normalization. Because each of the particles $1, \ldots, g$ must be somewhere in physical space, Eq. (1) demands that Eq. (2*a*) be integrated over all positions of all g particles. When Eq. (2*a*) is satisfied, ψ is said to be normalized. If $\psi'(\mathbf{r}_1, \ldots, \mathbf{r}_g, t)$ is a proposed wave function which satisfies Eq. (2*b*), the probabilities specified by ψ' are found from Eq. (1) using the normalized $\psi = C^{-1/2}\psi' \exp(i\eta)$, provided C is not infinite, that is, provided ψ' is quadratically integrable; the phase factor $\exp(i\eta)$, η being real, can be chosen arbitrarily. Though the absolute probabilities $P_j(\mathbf{r})$ are not defined when $\psi'(\mathbf{r}_1, \ldots, \mathbf{r}_g)$ is not quadratically integrable, $|\psi'|^2$ may remain a useful measure of the relative probability of finding particles $1, \ldots, g$ at $\mathbf{r}_1, \ldots, \mathbf{r}_g$.

$$\int d\mathbf{r}\, P_j(\mathbf{r}) = \int d\mathbf{r}_1 \cdots d\mathbf{r}_g |\psi(\mathbf{r}_1, \ldots, \mathbf{r}_g)|^2 = 1 \quad (2a)$$

$$\int d\mathbf{r}_1 \cdots d\mathbf{r}_g |\psi'|^2 = C \neq 1 \quad (2b)$$

One need be concerned only with wave functions which can represent actual physical systems. It is postulated that an admissible (physically possible) wave function $\psi'(\mathbf{r}_1, \ldots, \mathbf{r}_g)$ is quadratically integrable (type 1), or fails to be quadratically integrable (type 2) only because ψ' vanishes too slowly or at worst remains finite as infinity is approached in the $3g$-dimensional space of $\mathbf{r}_1, \ldots, \mathbf{r}_g$. Convincing physical and mathematical reasons can be found for excluding wave functions other than these types. One-particle wave functions $\psi'(\mathbf{r})$ of type 2 correspond to classically unconfined systems, for example, an electron ionized from a hydrogen atom; for these systems $\psi(\mathbf{r}) = \infty^{-1/2}\psi' = 0$ has the obvious interpretation that an unconfined particle is sure to be found outside any given finite volume. Such a $\psi'(\mathbf{r})$ also can represent a very large (effectively infinite) number of independently moving identical particles in a very large (effectively infinite) volume, for example, a beam of free electrons issuing from an electron gun; in this event $|\psi'(\mathbf{r})|^2$, although it continues to describe the likelihood of observing an electron, can be thought to equal the number density of electrons at any point $\mathbf{r}$, with $C = \infty$ in Eq. (2*b*) indicating that the number of particles in all of space is infinite. These considerations can be extended to quadratically nonintegrable many-particle wave functions $\psi'(\mathbf{r}_1, \ldots, \mathbf{r}_g)$.

Spin. The preceding formalism accepts the presumption of classical physics that a particle is a structureless entity, about which "everything" is known when its position $\mathbf{r}$ is known. This presumption is inaccurate (and therefore the formalism is not wholly adequate) for systems of electrons, protons, and neutrons, these being the fundamental particles which compose what usually is termed stable matter. In particular an electron, proton, or neutron cannot be described completely by a single wave function $\psi(\mathbf{r},t)$. For each of these particles two wave functions $\psi_1(\mathbf{r},t)$ and $\psi_2(\mathbf{r},t)$ are required, which may be regarded as components of an overall two-component wave function $\psi(\mathbf{r},t)$. The need for a multicomponent wave function has the immediate classical interpretation that the particle has internal degrees of freedom, that is, that knowledge of the position of the particle is not "everything." It can be shown that these internal degrees of freedom are associated with so-called intrinsic angular momentum or spin. *See* SPIN (QUANTUM MECHANICS).

For electrons, protons, or neutrons the spin is $^1/_2\hbar$, where $\hbar = h/2\pi$, and h is Planck's constant; the only allowed values of s_z, the z component of the spin, are

$\pm{}^1/_2$ (in units of $\hbar$). Thus when the system contains a single particle of spin $^1/_2$, an electron, say, $|\psi_1(\mathbf{r},t)^2|$ $d\mathbf{r}$ can be interpreted as the probability of finding, in the volume $d\mathbf{r}$, an electron with $s_z = +{}^1/_2$; $|\psi_2(\mathbf{r},t)|^2$ is the probability density for finding an electron with $s_z = -{}^1/_2$. The normalization condition replacing Eq. (2*a*) is given by Eq. (3).

$$\int d\mathbf{r}[|\psi_1(\mathbf{r},t)|^2 + |\psi_2(\mathbf{r},t)|^2] = 1 \tag{3}$$

This formalism is readily extended to many-particle systems. For example, when the system contains g particles of spin $^1/_2$, the overall wave function ψ has 2^g components ψ_j; $|\psi_1(\mathbf{r}_1,\ldots,\mathbf{r}_g)|^2$ is the probability density for finding each of particles 1 to g with spin oriented along $+z$; and the normalization condition is given by Eq. (4), summed from $j = 1$ to 2^g. The

$$\sum_j \int d\mathbf{r}_1 \cdots d\mathbf{r}_g |\psi_j(\mathbf{r}_1,\ldots,\mathbf{r}_g)|^2 = 1 \tag{4}$$

appropriate reinterpretations when the wave function is not quadratically integrable are obvious. Complications which arise from particle indistinguishability are discussed later in the article.

Operators

Whereas in classical mechanics particle coordinates $\mathbf{r}$ and momenta $\mathbf{p}$ are numbers which can be specified independently at any instant of time, in quantum mechanics the components of $\mathbf{r}$ and $\mathbf{p}$ are linear operators, as also are functions $f(\mathbf{r},\mathbf{p})$ of the coordinates and momenta. It is postulated that (i) the operator $\mathbf{x}$ (here distinguished by boldface from the x coordinate to which $\mathbf{x}$ corresponds) simply multiplies a wave function $\psi(x)$ by x, that is, $\mathbf{x}\psi = x\psi$; (ii) the operator corresponding to the canonically conjugate x component of momentum of that particle is $p_x = (\hbar/i)\partial/\partial x$, that is, $p_x\psi = (\hbar/i)\,\partial/\psi\,\partial/x$. Thus $A\psi$ denotes the new wave function $\psi' = A\psi$, resulting from the linear operation A on a given wave function ψ. When A, B, C are linear operators, and ψ, ξ are any two functions, Eqs. (5) hold. Moreover, if ξ

$$\begin{aligned} A(\psi + \xi) &= A\psi + A\xi \\ (A + B)\psi &= A\psi + B\psi \\ AB\psi &= A(B\psi) \end{aligned} \tag{5}$$

and ψ can be added or equated, then ξ and ψ must have the same number of components and depend upon the same space and spin coordinates; if $\xi = \psi$, corresponding components of ξ and ψ are equal. A spin-independent operator performs the same operation on each component of a many-component wave function, for example, $(p_x\psi)_j = p_x\psi_j = (\hbar/i)\partial\psi_j/\partial x$. Spin-dependent operators are more complicated; for instance, in a one-particle system of spin $^1/_2$ the components of $\psi' = s_z\psi$, where s_z denotes the z component of the spin operator, are given by Eqs. (6), using

$$\psi_1' = {}^1/_2\hbar\psi_1 \qquad \psi_2' = -{}^1/_2\hbar\psi_2 \tag{6}$$

the notation adopted previously in the discussion of spin.

The operators A and B are said to commute when, for any ψ, $A(B\psi) = B(A\psi)$, implying that Eq. (7) is

$$AB - BA = 0 \tag{7}$$

valid. The operator $(AB - BA)$ is termed the commutator of A and B. Any operator $f(A)$ expressible as a power series in the operator A commutes with A. By performing the indicated operations, Eq. (8) is

$$(xp_x - p_x x)\psi = i\hbar\psi \tag{8}$$

obtained, which shows that pairs of operators need not commute. In a g-particle system, all particle coordinates $x_1, y_1, z_1, \ldots, x_g, y_g, z_g$ commute with each other; all momentum coordinates $\mathbf{p}_1, \ldots, \mathbf{p}_g$ commute with each other; any component of $\mathbf{r}_1$ or $\mathbf{p}_1$ commutes with all components of $\mathbf{r}_2, \ldots, \mathbf{r}_g$ and of $\mathbf{p}_2, \ldots, \mathbf{p}_g$; the x coordinate of any particle commutes with p_y and p_z of that particle, and so on.

Hermitian operators. An operator A relevant to a given system of g particles is termed hermitian if Eq. (9) holds for all pairs of sufficiently well-behaved

$$\sum_j \int d\mathbf{r}_1 \cdots d\mathbf{r}_g \xi_j^*(A\psi)_j = \sum_j \int d\mathbf{r}_1 \cdots d\mathbf{r}_g (A\xi)_j^* \psi_j \tag{9}$$

quadratically integrable wave functions, $\xi(\mathbf{r}_1,\ldots,\mathbf{r}_g)$, $\psi(\mathbf{r}_1,\ldots,\mathbf{r}_g)$.

In Eq. (9) the asterisk denotes the complex conjugate, and the sum is over all components j of ξ, $A\psi$, $A\xi$, ψ. Evidently every particle coordinate x_1, $y_1, z_1, \ldots, x_g, y_g, z_g$ is a hermitian operator, as is any reasonably well-behaved $f(\mathbf{r}_1,\ldots,\mathbf{r}_g)$. Recalling that quadratically integrable functions vanish at infinity, integration by parts shows that every component of $\mathbf{p}_1, \ldots, \mathbf{p}_g$ is hermitian, as is any $f(\mathbf{p}_1,\ldots,\mathbf{p}_g)$ expressible as a power series in components of $\mathbf{p}_1, \ldots, \mathbf{p}_g$; for example, in the simple one-dimensional spinless case, Eq. (10) holds. It is implied that ξ and ψ

$$\int_{-\infty}^{\infty} dx\, \xi^* \frac{\hbar}{i}\frac{\partial\psi}{\partial x} = -\frac{\hbar}{i}\int_{-\infty}^{\infty} dx \frac{\partial\xi^*}{\partial x}\psi = \int_{-\infty}^{\infty} dx \left(\frac{\hbar}{i}\frac{\partial\xi}{\partial x}\right)^* \psi \tag{10}$$

are continuous; otherwise the integration by parts yields extra terms on the right side of Eq. (10). Similarly, $p_x{}^2$ has the desired hermitian property, defined by Eq. (9), only when ξ and ψ are continuous and have continuous first derivatives at all points.

When A, B, $C, \ldots,$ are individually hermitian Eq. (11) holds. For simplicity the integration variables and the summation over components are not indicated explicitly in Eq. (11). When A and B are hermitian, Eq. (11) implies (i) ${}^1/_2(AB + BA)$ is hermitian;

$$\int \xi^*[(ABC\cdots)\psi] = \int [(\cdots CBA)\xi]^*\psi \tag{11}$$

(ii) AB and BA are not hermitian unless A and B commute. For example, xp_x and $p_x x$ are not hermitian, but $^1/_2(xp_x + p_x x)$ is; classically, of course, there is no distinction between $xp_x, p_x x$, or $^1/_2(xp_x + p_x x)$. By appropriately symmetrizing, taking note of Eq. (11), one can construct the quantum-mechanical hermitian operator corresponding to any classical $f(\mathbf{r}_1, \ldots, \mathbf{r}_g; \mathbf{p}_1, \ldots, \mathbf{p}_g)$ expressible as a power series in components of coordinates and momenta.

If one supposes that the forces acting on a quantum-mechanical system of g particles are precisely the same as in the classical case, the energy operator is the classical hamiltonian in Eq. (12), where

$$H(\mathbf{r}_1, \ldots, \mathbf{r}_g; \mathbf{p}_1, \ldots, \mathbf{p}_g) = T + V \qquad (12)$$

the kinetic energy $T = p_1^2/2m_1 + \cdots + p_g^2/2m_g$; $V(\mathbf{r}_1, \ldots, \mathbf{r}_g)$ is the potential energy; and m_i is the mass of particle i. *See* HAMILTON'S EQUATIONS OF MOTION.

Quantum-mechanical nonclassical forces, with associated potential energy operators more complicated than $V(\mathbf{r}_1, \ldots, \mathbf{r}_g)$, are not uncommon, however. For example, the interaction between two neutrons is believed to include space exchange as in Eq. (13), where $J(r_{12})$ is an ordinary function of the

$$\begin{aligned} V\psi_j(\mathbf{r}_1, \mathbf{r}_2) &\equiv J(r_{12})P_{12}\psi_j(\mathbf{r}_1, \mathbf{r}_2) \\ &= J(r_{12})\psi(\mathbf{r}_2, \mathbf{r}_1) \end{aligned} \qquad (13)$$

distance r_{12} between the particles, and the space exchange operator P_{12} interchanges the space coordinates of particles 1 and 2 in each component of ψ. *See* NUCLEAR STRUCTURE.

Real eigenvalues. The eigenvalue equation for an operator A is given by Eq. (14), where the number α

$$Au(\alpha) = \alpha u(\alpha) \qquad (14)$$

is the eigenvalue, and the corresponding eigenfunction $u(\alpha)$ is a not identically zero wave function solving Eq. (14) for that value of α. The eigenvalue equation $Hu = Eu$ for the energy operator H has special importance and is known as the time-independent Schrödinger equation. Since the eigenvalues α are identified with the results of measurement, it is desirable that (i) the eigenvalues α all be real numbers and not complex numbers; (ii) the corresponding eigenfunctions form a complete set, the meaning and importance of which is explained subsequently. Property (i) is important because actual measurements yield real numbers only, for example, lengths, meter readings, and the like. If the eigenvalue of A were complex, it could not be maintained that each value of α represented a possible result of exact measurement of A. This assertion is not negated by the fact that it is formally possible to combine real quantities into complex expressions; for example, the coordinates of $\mathbf{r}$ in the xy plane form the complex vector $x + iy$. *See* EIGENFUNCTION; EIGENVALUE (QUANTUM MECHANICS).

The eigenvalues α_n belonging to the quadratically integrable eigenfunctions $u(\alpha_n) \equiv u_n$ of a hermitian operator A are necessarily real. In the notation of Eq. (11), letting $\xi = \psi = u_n$ in Eq. (9) and employing Eq. (14), one obtains Eq. (15).

$$\begin{aligned} \alpha_n \int u_n^* u_n &= \int u_n^*(\alpha_n u_n) = \int (Au_n)u_n^* \\ &= \int (Au_n)^* u_n = \int (\alpha_n u_n)^* u_n = \alpha_n^* \int u_n^* u_n \end{aligned} \qquad (15)$$

The equality of the first and last terms of Eq. (15) demonstrates that $\alpha_n = \alpha^*_n$. For this reason it is postulated that all "observable" operators are hermitian operators, and conversely that all hermitian operators represent observable quantities; henceforth all operators are supposed hermitian. In addition, it is necessary to require that the allowed eigenfunctions preserve the hermiticity property of Eq. (9); otherwise Eq. (15) would not hold. For the important class of hamiltonian operators $H = T + V$, except for highly singular (discontinuous) potentials, the boundary conditions that u and $\partial u/\partial x$ must be continuous guarantee reality of the eigenvalue corresponding to a quadratically integrable u solving Eq. (14). Admissible wave functions, defined following Eq. (2*b*), may be quadratically nonintegrable, however. It always is assumed that the physically desirable properties (i) and (ii) that were discussed above follow from the equally physically desirable simple requirement that the eigenfunctions $u(\alpha)$ of A must be admissible, provided that u, $\partial u/\partial x$, etc., satisfy the continuity conditions which make Eq. (15) correct for quadratically integrable u_n. In systems containing a single spinless particle this assumption has been justified rigorously for operators of interest; the widespread quantitative successes of quantum theory support the belief that the assumption is equally valid in more complicated systems.

Orthogonality. The eigenvalues α corresponding to quadratically integrable eigenfunctions typically form a denumerable (countable), though possibly infinite, set and compose the discrete spectrum of A. The admissible nonquadratically integrable eigenfunctions typically correspond to a continuous set of real eigenvalues composing the continuous spectrum of A. An eigenvalue α is degenerate, with order of degeneracy $d \geqq 2$, if there exist d independent eigenfunctions $u_1, \ldots, u_d$ corresponding to the same value of α, whereas every $d + 1$ such eigenfunctions are dependent. The d functions $\psi_1, \ldots, \psi_d$ are (linearly) dependent if the relation in Eq. (16) can be

$$c_1\psi_1 + \cdots + c_d\psi_d = 0 \qquad (16)$$

true with not all the constants $c_1, \ldots, c_d$ equal to zero. Eigenvalues which are either discrete or continuous or both may be degenerate. While so-called accidental degeneracy can occur, degeneracy of the eigenvalues α of A ordinarily is associated with the existence of one or more operators which commute with A. In the absence of degeneracy the eigenfunctions $u(\alpha)$ are uniquely indexed by α, and can be chosen to satisfy the orthonormal (orthogonality and normalizing) relations [compare Eq. (4)], as shown

in Eq. (17).

$$\sum_j \int d\mathbf{r}_1 \cdots d\mathbf{r}_g u_j^*(\alpha) u_j(\alpha') = \delta_{aa'} \text{ or } \delta(\alpha - \alpha') \quad (17)$$

In Eq. (17) the Kronecker symbol $\delta_{aa'}$ is employed when α lies in the discrete spectrum; $\delta_{aa'} = 0$ for $\alpha \neq \alpha'$, $\delta_{aa'}, = 1$ for $\alpha = \alpha'$. The Dirac delta function $\delta(\alpha - \alpha')$ is employed when α lies in the continuous spectrum; $\delta(\alpha - \alpha') = 0$ when $\alpha \neq \alpha'$, but has a finite integral defined by Eq. (18*a*). For a wide variety of functions $f(\alpha)$ these properties of $\delta(\alpha - \alpha')$ imply that Eq. (18*b*) is valid.

$$\int_{-\infty}^{\infty} d\alpha\, \delta(\alpha - \alpha') = \int_{-\infty}^{\infty} d\alpha'\, \delta(\alpha - \alpha') = 1 \quad (18a)$$

$$\int_{-\infty}^{\infty} d\alpha f(\alpha)\, \delta(\alpha - \alpha') = \int_{-\infty}^{\infty} d\alpha f(\alpha)\, \delta(\alpha' - \alpha) = f(\alpha') \quad (18b)$$

When Eq. (17) holds, $u(\alpha)$ are said to be normalized on the α-scale. Evidently $\delta(x)$ is highly singular at $x = 0$, and is an even function of x.

Equation (17) asserts that eigenfunctions corresponding to different eigenvalues always are orthogonal, that is, that the integral in Eq. (17) equals zero whenever $\alpha \neq \alpha'$. For discrete α (or α') this assertion is readily proved by an argument similar to Eq. (17); in fact, the orthogonality for $\alpha \neq \alpha'$ holds whether or not the eigenvalues are degenerate. When α and α' both lie in the continuous spectrum, however, the integral of Eq. (17) does not converge and therefore actually is not defined. Thus the delta function $\delta(\alpha - \alpha')$ primarily is a useful formal device; here and elsewhere in the theory, delta functions always are eventually integrated over their arguments, as in Eqs. (18); such integration makes expressions like the left side of Eq. (17) effectively convergent. A mathematically rigorous justification of the use of the delta function in quantum theory, or a rigorous justification of Eq. (17), encounters difficulties. These are related to the difficulties in establishing rigorously the aforementioned properties (i) and (ii). For operators and eigenfunctions of physical interest, experience suggests no reason to doubt the basic correctness of the mathematical procedures of nonrelativistic quantum theory. *See* OPERATOR THEORY.

Illustrative Applications

The immediately following subheadings apply the preceding formalism to some representative problems involving a one-dimensional spinless particle free to move in the x direction only; in every case, one begins by seeking the appropriate solution to Eq. (14). As subsequent discussion makes clear, results for this simplest of systems are pertinent to more complicated systems.

Momentum. Equation (14) is written as Eq. (19),

$$p_x u = \frac{\hbar}{i}\frac{\partial u}{\partial x} = \hbar k u \quad (19)$$

where $\hbar k$ is the eigenvalue; conventionally the eigenfunctions are indexed by the wave number k. The solutions to Eq. (19) have the form $u(x,k) = C(k) \exp[ikx]$, where $C(k)$ is a normalizing constant. When k is real, $u(x,k)$ is finite for all x, $-\infty \leqq x \leqq \infty$. Consequently, the spectrum is continuous and includes all real k, $-\infty \leqq k \leqq \infty$. If k has an imaginary part, that is, if $k = k_1 + ik_2$, $k_2 \neq 0$, then $u(x, k_1 + ik_2)$ is not admissible since it becomes infinite at either $x = +\infty$ or $x = -\infty$. If the eigenfunction could vanish identically for $|x| \geqq a$, it would be quadratically integrable for complex k; in this event the eigenfunction would be discontinuous at $x = a$. However, Eq. (20) holds

$$|C(k) \exp[ikx]| \neq 0 \quad (20)$$

unless $C(k) = 0$. Thus the requirements that u be admissible and continuous ensure that the eigenvalues of p_x are real, as asserted previously in the discussion of real eigenvalues. The $u(x,k)$ are quadratically nonintegrable, as expected for the continuous spectrum; each $u(x,k)$ can be regarded as representing a beam of particles, all moving with the same velocity. Because $\exp[ikx]$ is periodic with wavelength $\lambda = 2\pi/k$, such a beam will demonstrate wavelike properties; in fact, $p_x = \hbar k = h/\lambda$, in agreement with the de Broglie relations. This agreement is not trivial; although the conclusion that the quantum-mechanical momentum operator is $p_x = (\hbar/i)\, \partial/\partial x$ can be argued starting from the de Broglie relations, the form of p_x also can be inferred directly from Eq. (8), which in turn can be argued from the formal analogy between the properties of the commutator and the Poisson bracket, without any reference to the de Broglie relations. *See* CANONICAL TRANSFORMATIONS.

Normalized on the k scale, the eigenfunctions are given by Eq. (21*a*), for which corresponding to Eq. (17) Eq. (21*b*) may be written. Equation (21*b*)

$$u(x, k) = \frac{1}{\sqrt{2\pi}} e^{ikx} \quad (21a)$$

$$\int_{-\infty}^{\infty} dx\, u^*(k) u(k') = \frac{1}{2\pi}\int_{-\infty}^{\infty} dx\, e^{i(k'-k)x} = \delta(k - k') \quad (21b)$$

amounts to a nonrigorous statement of the Fourier integral transform theorem, as shown by Eqs. (22).

$$\psi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dk\, e^{ikx} c(k) \quad (22a)$$

$$c(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dx\, e^{-ikx} \psi(x) \quad (22b)$$

See FOURIER SERIES AND TRANSFORMS.

Equation (22*b*) can be derived by mathematically rigorous procedures. The quantum-theoretic derivation multiplies Eq. (22*a*) by $(2\pi)^{-1/2} \exp(-ik'x)$ and

integrates over all x. There is obtained, after interchanging the orders of x and k integration, Eq. (23).

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dx\, e^{-ik'x}\psi(x)$$
$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} dk\, c(k)\int_{-\infty}^{\infty} dx\, e^{i(k-k')x}$$
$$= \int_{-\infty}^{\infty} dk\, c(k)\delta(k-k') = c(k') \qquad (23)$$

Kinetic energy. For the kinetic energy $T_x = p^2_x/2m$, Eq. (14) is written as Eq. (24). The eigenvalue is E.

$$T_x u = -\frac{h^2}{2m}\frac{\partial^2 u}{\partial x^2} = Eu \qquad (24)$$

Admissible solutions are given by Eq. (25), where

$$u(x,E) = C(E)\exp\left(\frac{i}{\hbar}x\sqrt{2mE}\right) \qquad (25)$$

$\sqrt{E}$ must be real. Thus the spectrum is continuous and runs from $E = 0$ to $E = +\infty$. The square root may be either positive or negative in Eq. (25) so that there are two independent eigenfunctions at each value of E; that is, the spectrum is degenerate. The eigenfunctions may be labeled $u_+(x,E)$ and $u_-(x,E)$; normalized on the E scale, Eq. (26) obtains, in which

$$u_\pm(x,E) = \left(\frac{m}{2h^2E}\right)^{1/4}\exp\left(\pm\frac{i}{\hbar}x\sqrt{2mE}\right) \qquad (26)$$

the square root always is positive, u_+ and u_- individually satisfy Eq. (17), but the sets u_+ and u_- are orthogonal to each other. Introducing $k = \sqrt{2mE/h}$, the eigenfunctions can be labeled by the single parameter k, $-\infty \leqq k \leqq \infty$, instead of by E, $0 \leqq E \leqq \infty$, and the subscripts + or −. Evidently the $u(x,k)$ of Eq. (21a) are eigenfunctions not only of p_x but also of T_x; each $u(x,k)$ corresponds to a different eigenvalue of p_x, but $u(x,k)$ and $u(x,-k)$ correspond to the same eigenvalue of T_x. Interpreted physically, these results mean that the energy E of a free particle is known exactly when its momentum p_x is known exactly, and that $E = p^2_x/2m = (-p_x)^2/2m$ just as in classical mechanics. The normalizations in Eqs. (21a) and (26) are different because $dE = (\hbar^2k/m)\, dk$. The eigenfunctions in Eqs. (27) also are normalized on the E scale, and are an alternative set to $u_\pm(x,E)$ of Eq. (25).

$$u'_+(x,E) = \frac{1}{\sqrt{2}}[u_+(x,E) + u_-(x,E)]$$
$$= \left(\frac{2m}{h^2E}\right)^{1/4}\cos\frac{x}{\hbar}\sqrt{2mE} \qquad (27a)$$

$$u'_-(x,E) = \frac{1}{\sqrt{2}}[u_+(x,E) - u_-(x,E)]$$
$$= \left(\frac{2m}{h^2E}\right)^{1/4}\sin\frac{x}{\hbar}\sqrt{2mE} \qquad (27b)$$

Typical energy operator. Equation (14) is written as Eq. (28), where the operators in the brackets operate

$$Hu = (T_x + V)u$$
$$= \left[\frac{-\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]u = Eu \qquad (28)$$

on u. Typically, but not always (see the two subheadings immediately following), $V(x)$ is an everywhere smooth finite function which approaches zero at $x = \pm\infty$. **Figure 1** is a plot of such a $V(x)$, having attractive force centers (negative potential energy) near $x = \pm a$, and a repulsive region (positive potential energy) near $x = 0$; in this case the minimum potential at $x = \pm a$ is $V_0 < 0$.

Equation (28) is rewritten in the form of Eqs. (29).

$$\frac{\partial^2 u}{\delta x^2} = Ku \qquad (29a)$$

$$K(x,E) = \frac{2m}{\hbar^2}[V(x) - E] \qquad (29b)$$

When $K > 0$, Eq. (29b) implies that $(\partial/\partial x)(\partial u/\partial x)$ has the same sign as u; for example, $\partial u/\partial x$ increases with increasing x if $u > 0$. In other words, at points x where $K(x,E) > 0$, solutions of Eq. (29a), are convex toward the x axis; where $K(x,E) < 0$, $u(x,E)$ is concave toward the x axis. Provided $V(x)$ decreases to zero sufficiently rapidly, $K(\pm\infty,E) < 0$ when $E > 0$, so that at $x = \pm\infty$ solutions to Eq. (28) are approximated by linear combinations of the oscillatory functions of Eqs. (27): when $E < 0$, however, $K(\pm\infty,E) > 0$, so that at $x = \pm\infty$ solutions to Eq. (28) are approximated by linear combinations of the convex functions [compare Eq. (25)] in Eqs. (30).

$$u_1(x,E) = \exp\,[x(2m\,|E|)^{1/2}/\hbar]$$
$$u_2(x,E) = \exp\,[-x(2m\,|E|)^{1/2}\hbar] \qquad (30)$$

Function u_1 is infinite at $x = +\infty$; u_2 is infinite at $x = -\infty$. Thus, for $E < 0$, every admissible solution of Eq. (28) must behave like u_1 at $x = -\infty$ and like u_2 at $x = +\infty$, that is, must be asymptotic to the x axis at both $x = +\infty$ and $x = -\infty$. When $V_0 < E < 0$ there are values of x where the solution is concave, so that a smooth (satisfying the boundary condition that u and $\partial u/\partial x$ must be continuous) transition from a curve like G or L (**Fig. 2**) on the left to F on the right may be possible. When $E < V_0$ solutions to Eq. (28) are everywhere convex and, like F, G, L of

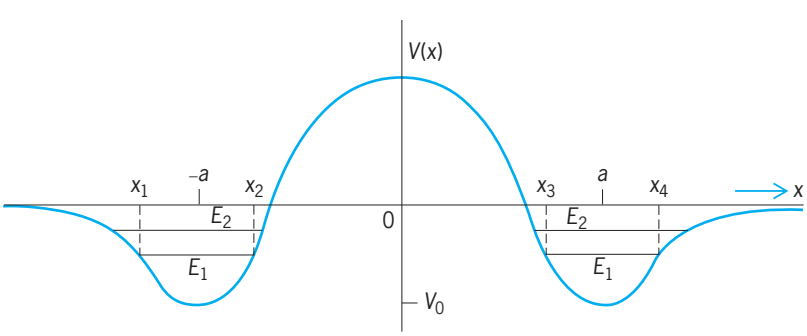

Fig. 1. A potential $V(x)$, capable of trapping particles near $x = \pm a$.

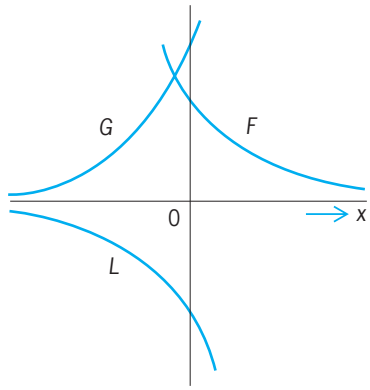

Fig. 2. Everywhere convex solutions asymptotic to *x* axis.

Fig. 1, never are asymptotic to the x axis at both $x = \pm\infty$.

It is concluded that (i) the continuous spectrum includes all positive values of E, $0 \leqq E \leqq \infty$, and at each such E there are two independent, nonquadratically integrable (oscillatory at $x = \pm\infty$) eigenfunctions, as in the previously discussed example of kinetic energy; (ii) there are no eigenvalues $E < V_0$; (iii) there may but need not be discrete eigenvalues $V_0 < E < 0$ corresponding to quadratically integrable eigenfunctions. The lowest eigenfunction, for $E = E_1$, has the least region of concavity and therefore joins F of Fig. 2 to G without crossing the x axis, as in **Fig. 3**. The eigenfunction $u(x,E_2)$ corresponding to the next higher eigenvalue $E = E_2 > E_1$ has one node (one zero), decreases less rapidly at $x = \pm\infty$ than $u(x,E_1)$, and links F to L; the next higher eigenfunction corresponding to $E_3 > E_2$ has two nodes, again links F to G, and so on.

The horizontal lines in the neighborhood of $x = \pm a$ in Fig. 1 show the energy levels of the two lowest eigenvalues E_1, E_2. The points x_1, x_2, where $E = V_1$ and where the curvature of $u(x,E_1)$ changes sign (Fig. 3), are the turning points (velocity equals zero) of a classical particle oscillating with total energy E_1 in the attractive potential well at $x = -a$; x_3, x_4 are similar turning points near $x = a$. For $E_1 < E < E_2$ a solution $u(x,E)$ starting out as does F at $x = +\infty$ crosses the x axis but becomes negatively infinite at $x = -\infty$, because it is insufficiently concave to join smoothly with L; this makes understandable the general result that eigenvalues corresponding to quadratically integrable eigenfunctions are discrete.

Potential barrier. Consider Eq. (28) with $V(x) = 0$ for $x < 0$; $V(x) = V_1 > 0$ for $x > 0$, and $0 < E < V_1$, as shown in **Fig. 4*a***. As previously explained, u and $\partial u/\partial x$ must be everywhere continuous, even though $V(x)$ is discontinuous at $x = 0$. Recollecting Eqs. (26), (29), and (30) for $0 < E < V_1$, there must be one and only one independent eigenfunction, having the form shown in Eqs. (31), where R and S are constants. It is reasonable and customary to interpret the first exponential in Eq. (31*a*) as a beam of particles moving with momentum $p_x = (2mE)^{1/2}$ toward the classical barrier or turning point $x = 0$; the second exponential in Eq. (31*a*) represents a reflected beam. Since the amplitude of the incident beam has been normalized to unity [not a normalization consistent with Eq. (17)], $|R|^2$ must be the reflection coefficient of the barrier. The continuity requirements at $x = 0$ yield Eqs. (32), showing that $|R|^2 = 1$ in agreement

$$u(x,E) = \exp[ix(2mE)^{1/2}/\hbar] + R\exp[-ix(2mE)^{1/2}/\hbar] \qquad x < 0 \qquad (31a)$$

$$u(x,E) = S\exp[-x(2m\,|E - V_1|)^{1/2}/\hbar] \qquad x > 0 \qquad (31b)$$

$$R = \frac{iE^{1/2} + (V_1 - E)^{1/2}}{iE^{1/2} - (V_1 - E)^{1/2}} \qquad S = \frac{2iE^{1/2}}{iE^{1/2} - (V_1 - E)^{1/2}} \qquad (32)$$

with classical expectation for $E < V_1$. *See* REFLECTION AND TRANSMISSION COEFFICIENTS.

The eigenfunction $u(x,E)$ is sketched in Fig. 4*a*. Since $|u|^2 \neq 0$ at $x > 0$, particles penetrate the classically inaccessible region to the right of the barrier;

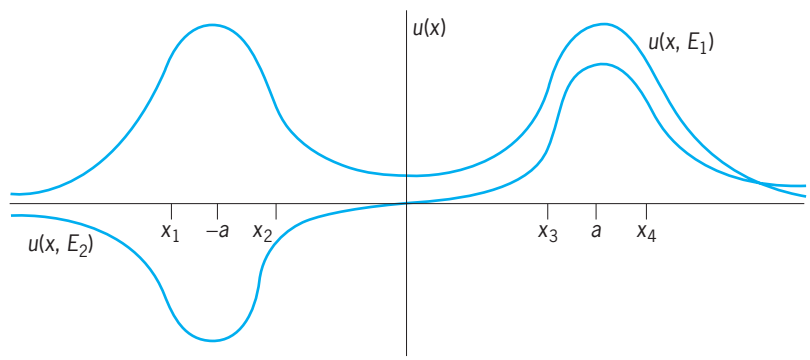

Fig. 3. Eigenfunctions indicated for the two lowest eigenvalues of the hamiltonian with the potential of Fig. 1.

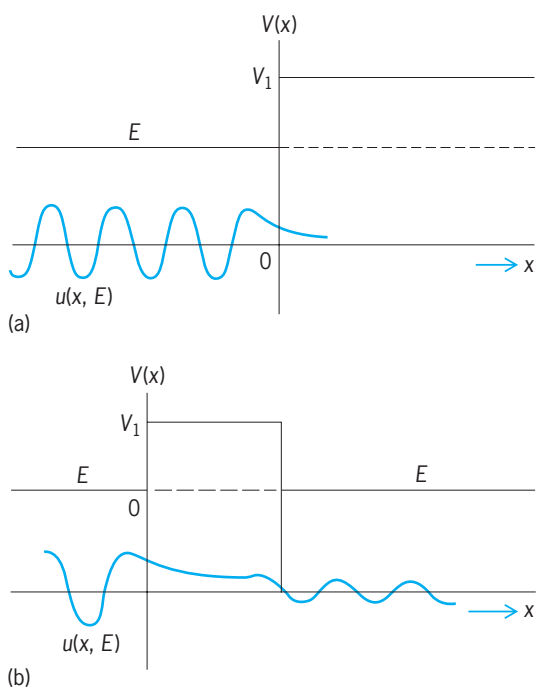

Fig. 4. Reflections (*a*) from a barrier of infinite thickness, and (*b*) from one of finite thickness.

because $|u|^2 = 0$ at $x = +\infty$, all particles eventually are turned back, however, consistent with $|R|^2 = 1$. This penetration, and Eqs. (31) and (32), closely resembles the penetration (and corresponding classical optics expressions) of a totally reflected light wave into a medium with smaller index of refraction. Hence Eqs. (31) and (32), which are unforced results of solving Eq. (28) for $V(x)$ of Fig. 4*a* , manifest the wave-particle duality inherent in the quantum theoretic formalism.

For the barrier of finite thickness a (Fig. 4*b*), there are two solutions at every $0 < E < V_1$. The solution representing a beam incident from the left, transmitted with coefficient $|T|^2$ and reflected with coefficient $|R|^2$, is found from Eq. (31*a*) together with Eqs. (33). Equations (33) mean that, except for the

$$u(x, E) = C \exp[-x(2m\,|E - V_1|)^{1/2}/\hbar] + D \exp[x(2m\,|E - V_1|)^{1/2}/\hbar] \quad 0 < x < a \qquad (33a)$$

$$u(x, E) = T \exp[ix(2mE)^{1/2}/\hbar] \quad x > a \qquad (33b)$$

incident $\exp[ix(2mE)^{1/2}]$ at $x = -\infty$, the solution at $x = \pm\infty$ must consist of waves traveling out toward infinity; a similar outgoing boundary condition specifies the continuum solution $E > 0$ representing more complicated collisions, for example, the scattering of a beam of particles by target particles in a foil. Outgoing boundary conditions are employed in classical wave theories as well. *See* SCATTERING EXPERIMENTS (ATOMS AND MOLECULES); SCATTERING EXPERIMENTS (NUCLEI); SCATTERING OF ELECTROMAGNETIC RADIATION.

The continuity requirements lead to Eq. (34*a*), and

$$|T|^2 = \left\{1 + \frac{V_1{}^2 \sinh^2 [(a/\hbar)\sqrt{2m(V_1 - E)}]}{4E(V_1 - E)}\right\}^{-1} \qquad (34a)$$

$$|T|^2 = \frac{16E(V_1 - E)}{V_1{}^2} \times \exp[-2a(2m\,|E - V_1|)^{1/2}/\hbar] \qquad (34b)$$

$|R|^2 = 1 - |T|^2$, as is necessary if $|R|^2$ and $|T|^2$ are to represent, respectively, reflection and transmission coefficients. Equation (34*a*), and the corresponding expression when $E \gg V_1$, resemble the equations for transmission of light through thin films. When $(a/\hbar)\sqrt{2m(V_1 - E)} \gg 1$, Eq. (34*a*) is closely approximated by Eq. (34*b*), where the exponential factor, sometimes termed the barrier penetrability, is $\ll 1$. The transmission through a less simple barrier $V(x)$ than Fig. 4*b* is measured by the penetrability factor defined by Eq. (35), where x_1, x_2 are the turning

$$P = \exp \frac{-2}{\hbar} \int_{x_1}^{x_2} dx\,(2m\,|E - V(x)|)^{1/2} \qquad (35)$$

points $E = V(x_1) = V(x_2)$. The barrier penetrability governs the rates at which (i) an incident proton, whose kinetic energy is less than the height of the repulsive Coulomb potential barrier surrounding a nucleus, is nonetheless able to penetrate the nucleus to produce nuclear reactions; (ii) alpha particles can escape from a radioactive nucleus; (iii) electrons can be pulled out of a metal by a strong electric field. *See* FIELD EMISSION; NUCLEAR FUSION; RADIOACTIVITY.

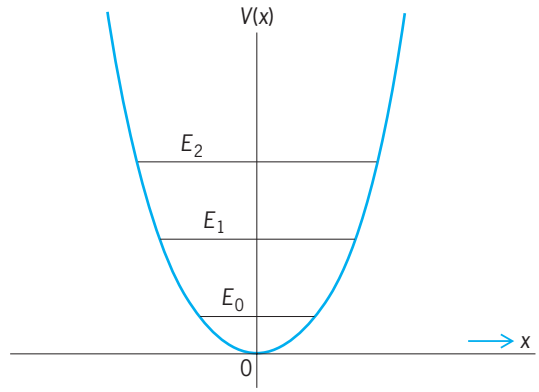

Fig. 5. Harmonic oscillator potential.

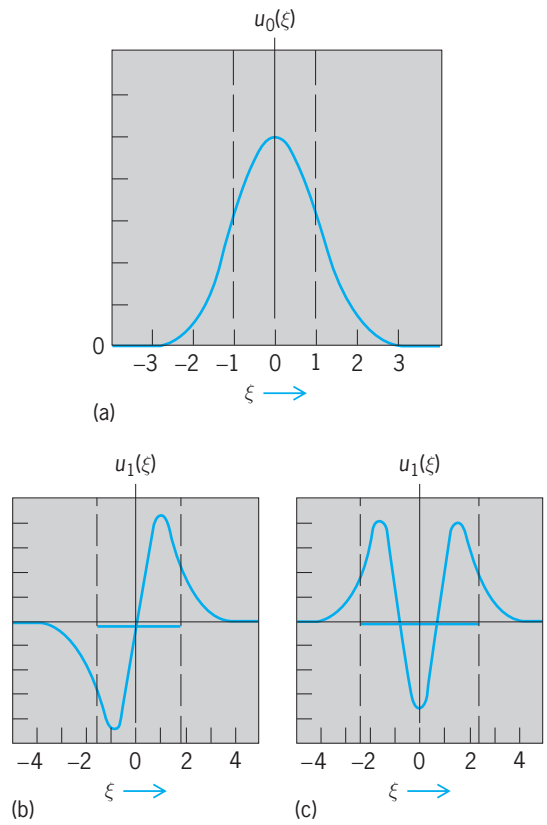

Fig. 6. Harmonic oscillator eigenfunctions. (*a*) $n = 0$. (*b*) $n = 1$. (*c*) $n = 2$. (*After L. Pauling and E. B. Wilson, Jr., Introduction to Quantum Mechanics, McGraw-Hill, 1935*)

Harmonic oscillator. The potential $V(x) = 1/2Kx^2$ (**Fig. 5**) describes a classical harmonic oscillator whose equilibrium position is $x = 0$; the classical oscillator frequency is $\nu = (2\pi)^{-1}\sqrt{K/m}$. Since $V(x)$ becomes infinite at $x = \pm\infty$, there is no continuous spectrum, but there must be an infinite number of discrete eigenvalues. Various sophisticated methods exist for finding the eigenfunctions and eigenvalues. *See* HARMONIC OSCILLATOR.

The energy levels turn out to be as in Eq. (36),

$$E_n = (n + {}^1/_2)h\nu \qquad (36)$$

where $n = 0, 1, 2, \ldots$ (Fig. 5). The corresponding first three eigenfunctions $u_0(\xi)$, u_1, u_2 are sketched in **Fig. 6**, with $\xi = x(mK)^{1/4}/\hbar^{1/2}$ a convenient dimensionless variable; the broken vertical lines indicate the turning points at $\xi = \pm\sqrt{2n+1}$. **Figure 7** plots the probability density $|u(\xi)|^2$ for $n = 10$. The classical probability of finding a particle in the x interval dx is proportional to the time spent in dx. The curved dashed line in Fig. 7 plots the classical probability density for a classical oscillator whose energy is $E_{10} = 2^1/_2 h\nu$, Eq. (36). *See* DIMENSIONLESS GROUPS; ENERGY LEVEL (QUANTUM MECHANICS).

The agreement between the classical probability density and the average over oscillations of $|u_{10}(\xi)|^2$ illustrates the connection between classical particle mechanics and the more fundamental dual wave-particle interpretation of quantum theory. With increasing n the oscillations of $|u_n(\xi)|^2$ become more rapid, and the agreement with the classical probability density improves, in accordance with the correspondence principle. These harmonic oscillator results are a good first approximation to the energy levels and eigenfunctions of vibrating atoms in isolated molecules and in solids. *See* LATTICE VIBRATIONS; MOLECULAR STRUCTURE AND SPECTRA; SPECIFIC HEAT OF SOLIDS.

Expectation Values

Suppose B is an operator that commutes with A. If $Au_n = \alpha_n u_n$, $BAu_n = B(\alpha_n u_n)$, so that, using Eq. (7), one obtains Eq. (37), which means that Bu_n also is

$$A(Bu_n) = \alpha_n(Bu_n) \tag{37}$$

an eigenfunction of A corresponding to the eigenvalue α_n. When α_n is not degenerate, Eq. (16) means Bu_n must be a multiple of u_n, that is, $Bu_n = \beta u_n$, β being a constant; thus u_n is simultaneously an eigenfunction of the pair of commuting operators A, B. When the order of degeneracy of α_n is d, Bu_n need not be a multiple of u_n, but d independent eigenfunctions $u_{n1}(\alpha_n,\beta_1), \ldots, U_{nd}(\alpha_n,\beta_d)$ always can be

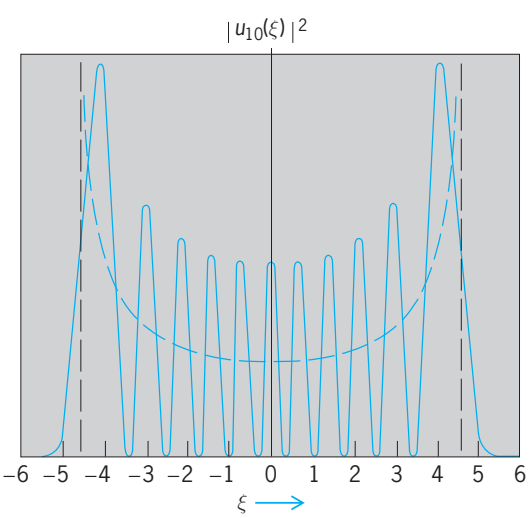

Fig. 7. Probability density for harmonic oscillator in state $n = 10$. (*After L. Pauling and E. B. Wilson, Jr., Introduction to Quantum Mechanics, McGraw-Hill, 1935*)

found, such that $u_{ns}(\alpha_n, \beta_s)$ is an eigenfunction of B corresponding to the eigenvalue β_s, as well as an eigenfunction of A corresponding to α_n. If all the β_s are different, $u_{n1}, \ldots, u_{nd}$, being eigenfunctions corresponding to different eigenvalues of B, are mutually orthogonal (see the earlier discussion on orthogonality). If all the β_s are not different, a third operator C, simultaneously commuting with A and B, is sought, and so on. For the system of g particles of spin $^1/_2$, one determines in this fashion a complete set of simultaneously commuting observables, A, B, $C, \ldots$, whose corresponding eigenvalues $\alpha, \beta, \gamma, \ldots$ (except for accidental degeneracy) uniquely index an orthonormal set of 2^g-component eigenfunctions as in Eq. (38), which satisfies Eq. (39).

$$u_j(\mathbf{r}_1, \ldots, \mathbf{r}_g; \alpha, \beta, \gamma, \ldots), \quad j = 1 \text{ to } 2^g \tag{38}$$

$$\int u^*(\alpha, \beta, \gamma, \ldots)u(\alpha', \beta', \gamma', \ldots)$$

$$= \delta(\alpha - \alpha')\delta(\beta - \beta')\delta(\gamma - \gamma')\cdots \tag{39}$$

In Eq. (39), and henceforth unless specifically indicated otherwise, the integral is in the simplified notation of Eq. (11). Equation (39) generalizes Eq. (17); $\delta(\alpha - \alpha')$ is replaced by a Kronecker symbol $\delta_{\alpha\alpha'}$ when α, α' are discrete, and so on.

The meaning of property (ii) mentioned in the earlier discussion of real eigenvalues can now be explained. The eigenfunctions of an operator A having been indexed as in Eq. (39), any sufficiently smooth quadratically integrable 2^g-component wave function $\psi(\mathbf{r}_1, \ldots, \mathbf{r}_g)$ can be expressed in the form of Eq. (40), where each eigenvalue is integrated over its

$$\psi(\mathbf{r}_1, \ldots, \mathbf{r}_g) = \int d\alpha \int d\beta \int d\gamma \cdots$$

$$\times\, c(\alpha, \beta, \gamma, \ldots)u(\mathbf{r}_1, \ldots, \mathbf{r}_g; \alpha, \beta, \gamma, \ldots) \tag{40}$$

entire spectrum, with the understanding (further to simplify the notation) that in the discrete spectrum integration is replaced by summation. Employing Eq. (39), the constants $c(\alpha,\beta,\lambda, \ldots)$ are found to be as shown in Eq. (41).

$$c(\alpha, \beta, \gamma, \ldots) = \int u_j^*(\alpha, \beta\gamma, \ldots)\psi \tag{41}$$

Equations (40) and (41) are consistent with the often instructive interpretation that $c(\alpha,\beta,\gamma, \ldots)$ are the projections of the vector ψ on a complete set of orthogonal unit vectors $u(\alpha,\beta,\gamma, \ldots)$ in an infinite-dimensional vector space.

Probability considerations. When ψ is normalized, Eqs. (4), (18), (39), and (40) imply Eq. (42), in the notation of Eq. (40). Equations (40) and (42) make

$$\int d\alpha \int d\beta \int d\gamma \cdots |c(\alpha, \beta, \gamma, \ldots)|^2 = 1 \tag{42}$$

it reasonable to postulate that whenever a system is instantaneously described by the normalized

quadratically integrable $\psi(\mathbf{r}_1, \ldots, \mathbf{r}_g)$, the instantaneous probability of finding the system in the state corresponding to $A = \alpha$ is given by Eq. (43*a*).

$$P(\alpha) = \int d\beta \int d\gamma \cdots |c(\alpha, \beta, \gamma, \ldots)|^2 \qquad (43a)$$

$$P(\alpha, \beta) = \int d\gamma \cdots |c(\alpha, \beta, \gamma, \ldots)|^2 \qquad (43b)$$

Similarly, the simultaneous probability of finding the system in the state corresponding to $A = \alpha$ and $B = \beta$ is given, by Eq. (43*b*), and so on. Clearly Eqs. (1) and (2) are a special case of Eqs. (42) and (43), in which the coordinate operators, $\mathbf{x}_1, \mathbf{y}_1, \mathbf{z}_1, \ldots, \mathbf{x}_g, \mathbf{y}_g, \mathbf{z}_g$ of the g spinless particles form the complete set of simultaneously commuting observables, and $\psi(\mathbf{x}_1, \mathbf{y}_1, \mathbf{z}_1, \ldots, \mathbf{x}_g, \mathbf{y}_g, \mathbf{z}_g)$ is the projection of ψ on the eigenfunction simultaneously corresponding to $\mathbf{x}_1 = x_1, \mathbf{y}_1 = \gamma_1, \ldots, \mathbf{z}_g = z_g$.

The total probability $\int d\alpha\, P(\alpha)$ is unity, Eq. (42), and therefore any measurement of the observable A must yield a number equal to one of its eigenvalues. Moreover, by the very meaning of probability, the average or expectation value of the observable A must be as given in Eq. (44*a*), which shows that the same

$$\langle A \rangle = \int d\alpha\, \alpha P(\alpha) \qquad (44a)$$

$$\langle A \rangle = \int \psi^*(A\psi) \qquad (44b)$$

operator may have different forms in different representations. For instance, in the momentum representation discussed earlier where $\alpha \equiv k$, $p_x c(k)$ must equal simply $\hbar k c(k)$; consequently, to satisfy Eq. (8), $\mathbf{x}c(k)$ must equal $i(\partial c \partial k)$. When ψ is a wave function for which Eq. (9) holds with $\xi \equiv \psi$, Eqs. (5), (14), (39), and (40) imply that Eq. (44*b*) is equivalent to Eq. (44*a*). Equation (40*b*), usually more convenient than Eq. (44*a*), predicts the expectation value of any given observable in an arbitrary physical situation described by a reasonably well-behaved quadratically integrable wave function; expectation values are not defined for nonquadratically integrable ψ. Equation (9) guarantees that $\langle A \rangle$ computed from Eq. (44*b*) is a real number, necessary if $\langle A \rangle$ is to represent the results of measurement (see the earlier discussion on real eigenvalues). *See* PROBABILITY.

Uncertainty principle. A precise measure of the spread or uncertainty in the value of A is ΔA defined by Eq. (45). Here $(\Delta A)^2$ is the average square

$$(\Delta A)^2 = \langle (A - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - (\langle A \rangle)^2 \qquad (45)$$

deviation of A from its average $\langle A \rangle$. The quantity $\Delta A = 0$ when and only when ψ is an eigenfunction of A, that is, $A\psi = \alpha\psi$, in which event $\langle A \rangle = \alpha$, $\langle A^2 \rangle = \alpha^2$. The discussion following Eq. (37) implies ΔA and ΔB simultaneously can be zero; that is, A and B are simultaneously exactly measurable, whenever the commutator $AB - BA$ is zero. If $AB - BA \neq 0$, introduce $A' = A - \langle A \rangle$, $B' = B - \langle B; \rangle$ use Eqs. (11) and (44*b*); and employ the so-called Schwartz inequality. Then Eq. (46) holds.

$$(\Delta A)^2 (\Delta B)^2 = \int (A'\psi)^* (A'\psi) \int (B'\psi)^* (B'\psi)$$
$$\geq \left| \int (A'\psi)^* (B'\psi) \right|^2 = \left| \int \psi^* (A'B'\psi \right|^2 \qquad (46)$$

But since Eq. (47) is valid, it leads, after further manipulation, to Eq. (48), which is the rigorous quan-

$$A'B' = {}^1/_2(A'B' - B'A') + {}^1/_2(A'B' + B'A') \qquad (47)$$

$$(\Delta A)^2 (\Delta B)^2] \geq {}^1/_4\, |\langle AB - BA \rangle|^2 \qquad (48)$$

tum theoretic formulation of the uncertainty principle. When $A = x$, $B = p_x$, Eq. (8) yields Eq. (49).

$$(\Delta x)(\Delta p_x) \geq \hbar/2 \qquad (49)$$

This simple derivation of the uncertainty principle demonstrates anew the necessity for a dual wave-particle interpretation of the operator formalism.

Time Dependence

The procedures developed thus far predict the results of measurement at any given instant of time t_0 that the wave function $\psi(x,t_0)$ is known. Given ψ at t_0, to predict the results of measurement at some future instant t, it is necessary to know the time evolution of ψ from t_0 to t. It is postulated that this evolution is determined by the time-dependent Schrödinger equation (50), where H is the hamilto-

$$H\psi = i\hbar \frac{\partial \psi}{\partial t} \qquad (50)$$

nian or energy operator; ψ is supposed to be a continuous function of t. If H were a number, the solution to Eq. (50) would be as shown in Eq. (51). Equation

$$\psi(t) = \exp\left[\frac{-iH(t - t_0)}{\hbar}\right] \psi(t_0) \qquad (51)$$

(51) remains valid when H is an operator, provided Eq. (52) holds, with I the unit operator, $I\psi = \psi$;

$$\exp\left[\frac{-iH(t - t_0)}{\hbar}\right] \equiv I - \frac{i(t - t_0)}{\hbar} H$$
$$+ \frac{[-i(t - t_0)]^2}{2\hbar^2} H^2 + \cdots \qquad (52)$$

the right side of Eq. (52) is the usual series expansion of the exponential on the left side. These operational methods, which manipulate operators like numbers, are widely used in quantum theory; they must be used cautiously, but usually lead rapidly to the same results as more conventional mathematical techniques.

When $H\psi(t_0) = E\psi(t_0)$, that is, when $\psi(t_0)$ is known to be an eigenfunction $u(E,t_0)$ of the energy operator, Eq. (51) shows that $\psi(t) \equiv u(E,t) = u(E,t_0) \exp[-iE(t - t_0)/\hbar]$, so that $|u(E,t)|^2 = |u(E,t_0)|^2$; similarly, the expectation values $<A>$ of alloperators A

are time-independent when the system is in a stationary state $u(E,t)$. In an arbitrary state $\psi(t)$, Eqs. (5), (11), (44*b*) and (50) imply [assuming that A is not explicitly time-dependent, for example, $A = A(\mathbf{r},\mathbf{p})$ but not $A(\mathbf{r},\mathbf{p},\mathbf{t})$] Eq. (53), which uses the notation

$$\begin{aligned}\frac{d}{dt}\langle A\rangle &= \int\left[\psi^*\left(A\frac{\partial\psi}{\partial t}\right)+\frac{\partial\psi^*}{\partial t}(A\psi)\right]\\ &= \frac{1}{i\hbar}\int[\psi^*(AH\psi)-(H\psi)^*A\psi]\\ &= \frac{1}{i\hbar}\int[\psi^*(AH-HA)\psi]\\ &= \frac{1}{i\hbar}\langle(AH-HA)\rangle \qquad (53)\end{aligned}$$

of Eq. (11). Of course $\langle(AH - HA)\rangle = 0$ whenever $\psi(t)$ is a stationary state $u(E,t)$. Equation (53) shows, however, that if A commutes with the hamiltonian, then $\langle A\rangle$ is independent of time whether or not the system is in a stationary state. Consequently, operators commuting with the hamiltonian are termed constants of the motion; a system initially described by an eigenfunction $u(E,\beta,\gamma,\ldots)$ of the simultaneously commuting observables $H, B, C, \ldots$, remains in an eigenstate of $H, B, C, \ldots$, as the wave function evolves in time.

Equation (53) is closely analogous to the classical mechanics expression for $dA(\mathbf{r},\mathbf{p})/dt$, where $A(\mathbf{r},\mathbf{p})$ is the classical quantity corresponding to the quantum-mechanical operator A. If A is put equal to I in Eq. (53), one obtains Eq. (54).

$$\frac{d}{dt}\int\psi^*\psi = \frac{1}{ib}\int[\psi^*(H\psi)-(H\psi)^*\psi] \qquad (54)$$

Therefore the requirement that H be hermitian, Eq. (9), which has been justified on the grounds that the eigenvalues of H must be real, has the further important consequence that the right side of Eq. (54) is zero, that is, that Eq. (2*a*) is obeyed at all times $t > t_0$ if it obeyed at $t = t_0$. This result is necessary for the consistency of the formalism: otherwise it could not be claimed that $|\psi(t)|^2$ from Eq. (50) is the probability density at $t > t_0$. For a single spinless three-dimensional particle, with $H = p^2/2m + V(x,y,z)$, it follows directly from Eq. (50) that one obtains Eq. (55*a*), where Eq. (55*b*) holds, and so on. In

$$\frac{\partial}{\partial t}(\psi^*\psi)+\frac{\partial}{\partial x}S_x+\frac{\partial}{\partial y}S_y+\frac{\partial}{\partial z}S_z = 0 \qquad (55a)$$

$$S_x = \frac{\hbar}{2mi}\left(\psi^*\frac{\partial\psi}{\partial x}-\psi\frac{\partial\psi^*}{\partial x}\right) \qquad (55b)$$

Eq. (55*a*) S_x, S_y, S_z can be interpreted as the components of a probability current vector $\mathbf{S}$, whose flow across any surface enclosing a volume τ accounts for the change in the probability of finding the particle inside τ. *See* EQUATION OF CONTINUITY; MAXWELL'S EQUATIONS.

For a nonquadratically integrable ψ in the one-particle case where Eq. (55*b*) is applicable, the probability current at infinity generally has the value $|\psi|^2\mathbf{v}$, where $\mathbf{v}$ is the classical particle velocity at infinity; the one-dimensional plane waves of Eq. (21*a*) trivially illustrate this assertion. Consequently (see the preceding discussion of normalization), $\mathbf{S}$ of Eq. (55*b*) is interpretable as particle current density when $|\psi|^2$ is nonvanishing at infinity. These considerations may be generalized to more complicated systems and are important in collision problems, where the incoming and outgoing current at infinity determine the cross section.

Invariance. Extremely general arguments support the view that the form of the Schrödinger equation (50) for any g-particle system isolated from the rest of the universe must be (i) translation invariant, that is, independent of the origin of coordinates; (ii) rotation invariant, that is, independent of orientation of the coordinate axes; and (iii) reflection invariant, that is, independent of whether one chooses to use a left-handed or right-handed coordinate system.

The only known failures of these general requirements occur for reflections, in a domain outside the scope of nonrelativistic quantum theory, namely, in phenomena, such as beta decay, that are connected with the so-called weak interactions. Correspondingly, it can be inferred that the hamiltonian operator H for any such isolated system must commute with (i) the total linear momentum operator $\mathbf{p}_R = \mathbf{p}_1 + \cdots + \mathbf{p}_g$; (ii) the total angular momentum operator $\mathbf{J}$; (iii) the parity operator P, which reflects every particle through the origin, that is, changes $\mathbf{r}_1$ to $-\mathbf{r}_1, \ldots, \mathbf{r}_g$ to $-\mathbf{r}_g$. For additional information *see* PARITY (QUANTUM MECHANICS); SYMMETRY LAWS (PHYSICS).

In quantum mechanics as in classical mechanics, therefore, linear momentum and total angular momentum are conserved, that is, are constants of the motion. Since for an infinitesimal displacement ϵ in the $x$direction Eq. (56) holds, the connection between $p_{1x} + \cdots + p_{gx}$ and translation in the x direction can be understood; the connection between $\mathbf{J}$ and rotation is understood similarly. Because a discontinuous change in position, from $\mathbf{r}$ to $-\mathbf{r}$, is inconceivable classically, the conservation of parity concept has no relevance to classical mechanics.

$$\begin{aligned}&\psi(x_1+\epsilon, y_1, z_1, x_2+\epsilon, y_2, z_2, \ldots, x_g+\epsilon, y_g z_g)\\ &\quad= \psi(x_1, y_1, z_1, \ldots, x_g, y_g, z_g)\\ &\qquad+\epsilon\left(\frac{\partial\psi}{\partial x_1}+\frac{\partial\psi}{\partial x_2}+\cdots+\frac{\partial\psi}{\partial x_g}\right)\\ &\quad= \psi+\epsilon p_{Rx}\psi \qquad (56)\end{aligned}$$

Transition probability. Frequently the hamiltonian H of Eq. (50) has the form $H_0 + V'(t)$, where the time-dependent potential energy $V'(t)$ represents an externally imposed interaction, for example, with measuring equipment; it is supposed that $V'(t) = 0$ for $t < 0$ and $t > t_1$. Usually the system is in a stationary state $u(E_i)$ of H_0 at time $t < 0$, and one wishes to compute the probability of finding the system in some other stationary state $u(E_f)$ of H_0 at times $t > t_1$. From Eq. (40) and the discussion preceding Eq. (53), Eq. (57) is obtained, which, when substituted in

Eq. (50) for times $0 \leqq t \leqq t_1$, yields Eq. (58).

$$\psi(t) = \int dE\, d\beta\, c(E, \beta, t) \exp(-iEt/\hbar) u(E, \beta) \quad (57)$$

$$i\hbar \int dE\, d\beta\, \frac{dc(E, \beta, t)}{dt} \exp(-iEt/\hbar) u(E, \beta) = \int dE\, d\beta\, c(E, \beta, t) \times \exp(-iEt/\hbar) V'(t) u(E, \beta) \quad (58)$$

In Eqs. (57) and (58), E is α of Eq. (40) and β stands for all other indices $\beta, \gamma, \ldots$, necessary to make $u(E)$, a complete orthonormal set; if $V'(t)$ were zero, the projections $c(E,\beta)$ would equal $\delta(E - E_i)\delta(\beta - \beta_i)$ independent of time.

Most problems in quantum theory cannot be solved exactly; therefore some approximate treatment using perturbations must be devised. In the present case it is assumed that $c(E,\beta,t)$ do not change appreciably from their initial values at $t = 0$, so that it is a reasonable approximation to replace $c(E,\beta,t)$ by $c(E,\beta, 0)$ on the right side of Eq. (58). With the further approximation that V' is constant during the interval $0 \leqq t \leqq t_1$, one finds, using the notation of Eq. (11), that Eq. (59) obtains, where $\hbar\omega = E_f - E_i$.

$$|c(E_f, \beta_f, t_1)|^2 \equiv \frac{4}{\hbar^2} \left| \int u^*(E_f, \beta_f) V' u(E_i, \beta_i) \right|^2 \times \frac{\sin^2 {}^1/_2 \omega t_1}{\omega^2} \equiv \frac{4}{\hbar^2} |V'_{fi}|^2 \frac{\sin^2 {}^1/_2 \omega t_1}{\omega^2} \quad (59)$$

Equation (59) shows that the probability of finding the system in some new stationary state $u(E_f,\beta_f)$ after the system is no longer perturbed is proportional to (i) $|V'_{fi}|^2$, the square of the matrix element of the perturbation between initial and final states; (ii) an oscillating factor which for given t_1 has a peak $t^2{}_1/4$ at $\omega = 0$.

$\mathrm{Sin}^2\,(1/2\omega t_1)/\omega^2$ is plotted as a function of ω in **Fig. 8**; evidently $|c(E_f,\beta_f)|^2$ is relatively small for energies such that $|\omega| > \sim\pi/t_1$. The most likely occupied states after the perturbation conserve energy ($E_f = E_i$), and the spread in energy of the final states is $\Delta E = \hbar\Delta\omega = \sim 2\pi\hbar/t_1 \equiv h/\Delta t$, where $t_1 \equiv \Delta t$ is, for example, the duration of the measurement. Thus Eq. (59) provides a version of the uncertainty principle between energy and time. As t_1 approaches infinity, the area under the curve in Fig. 8 becomes proportional to t_1, since the main peak has a height proportional to $t_1{}^2$ and a width proportional to $1/t_1$. Therewith one obtains a widely employed formula giving the approximate transition probability w per unit time for making transitions, under the influence of a steady perturbation V', from an initial stationary $u(E_i,\beta_i)$ to a set of final states of the same energy, namely, Eq. (60), where $\rho(E_f) = d\beta d\gamma \ldots$ is the density of independent final states in the neighborhood of $E = E_f = E_i$, $\beta = \beta_f$, $\gamma = \gamma_f$, and so on. For instance, with u_i a plane wave e_{ikz} moving in the z direction, Eq. (21a), and u_f a plane wave in some other direction, Eq. (60) yields the Born approximation to the cross section for elastic scattering by a potential. Equation (60) is also applicable to problems outside the domain of nonrelativistic quantum theory, for example, to the theory of beta decay wherein new particles are created.

$$w = \frac{2\pi}{\hbar} \rho(E_f) |V'_{fi}|^2 \quad (60)$$

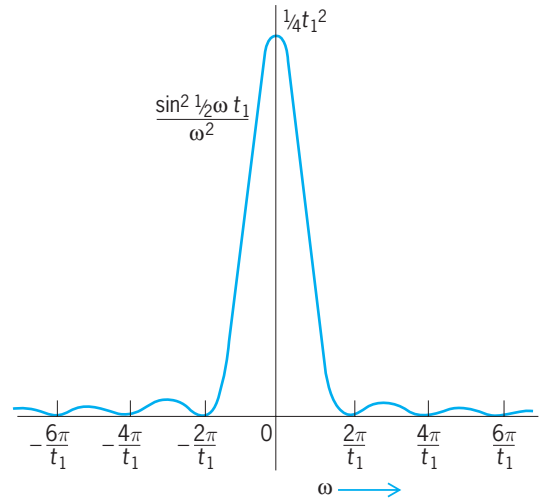

Fig. 8. Plot of $\sin^2 {}^1/_2\, \omega\, t_1/\omega^2$ versus $\omega = (E_f - E_i)/n$. (*After L. I. Schiff, Quantum Mechanics, 2d ed., McGraw-Hill, 1955*)

The preceding considerations are important for understanding how a measurement of an operator A not commuting with the hamiltonian H can cause an initially stationary state $u(E)$ to evolve into an eigenstate $u(\alpha)$ of A. Equations (50)–(53), with $H = T + V = H_0$, the unperturbed hamiltonian, hold only in the intervals between measurements; during the measurement $u(E)$, though an eigenfunction of H_0, is not a stationary state of the complete hamiltonian. But this paragraph does not do justice to the subtle questions involved in the quantum theory of measurements.

Further Illustrative Applications

When g particles are noninteracting, the hamiltonian has the trivially separable form $H \equiv H_1 + H_2 \cdots + H_g$, and Eq. (61) holds, where $u_1(\mathbf{r}_1,E_1,\alpha_1)$ are a complete

$$(H_1 + \cdots + H_g)[u_1(\mathbf{r}_1, E_1, \alpha_1) \cdots u_g(\mathbf{r}_g, E_g, \alpha_g)] = (E_1 + \cdots + E_g) \times [u_1(\mathbf{r}_1, E_1, \alpha_1) \cdots u_g(\mathbf{r}_g, E_g, \alpha_g)] \quad (61)$$

orthonormal set of eigenfunctions of H_1 for particle 1, and so on. Thus, because an operation performed solely on particle 1 always commutes with an operation performed solely on particle 2, the products $u_1(\mathbf{r}_1,E_1,\alpha_1) \ldots u_g(\mathbf{r}_g,E_g,\alpha_g)$ are a complete orthonormal set of eigenfunctions of $H = H_1 + \cdots + H_g$; also the energy levels of H are the set of possible sums of individual particle energies, $E = E_1 + \cdots + E_g$.

Similarly, if $H(x) = T_x + V(x)$, with eigenfunctions $u(x,E_x)$, is a hamiltonian for a one-dimensional spinless particle, then any quadratically integrable $\psi(x,y,z)$ describing a three-dimensional spinless particle can be expanded in a series of products $u(x,E_x)$ $u(y,E_y)$ $u(z,E_z)$. This paragraph explains the relevance, to three-dimensional many-particle systems of the results obtained for the illustrative applications previously discussed. The immediately following subheadings continue to illustrate the general formalism. The reader is cautioned that the remaining contents of this article, although very important, especially for applications to atomic and nuclear structure, are for the most part admittedly more condensed than the material presented heretofore.

Parity. Because $P^2\psi(\mathbf{r}) = P\psi(-\mathbf{r}) = \psi(\mathbf{r})$, the parity operator has but two eigenvalues, namely, $+1$ and -1; the corresponding eigenfunctions are said to have even or odd parity. Evidently P commutes with the harmonic oscillator hamiltonian $p^2_x/2m + (^1/_2)Kx^2$. The harmonic oscillator eigenvalues are nondegenerate, and therefore every harmonic oscillator eigenfunction (Fig. 6) has either even or odd parity. Similarly the eigenfunctions u'_+, u'_- of T_x, Eqs. (27), have even and odd parity, respectively; eigenfunctions of T_x which do not have definite parity also exist, however, Eq. (26), because the eigenvalues E are degenerate. The eigenfunctions of p_x do not have definite parity, Eq. (21*a*), because $p_xP + Pp_x = 0$, that is, p_x does not commute, but instead anticommutes, with P.

Time evolution of packet. A wave function ψ representing a single (spinless) particle localized in the neighborhood of a point is termed a wave packet. Assuming $H = p^2/2m + V(\mathbf{r})$, Eq. (53) yields Eq. (62*a*)

$$\frac{d}{dt}\langle\mathbf{x}\rangle = \frac{1}{i\hbar}\langle\mathbf{x}H - H\mathbf{x}\rangle = \left\langle\frac{p_x}{m}\right\rangle \tag{62a}$$

$$\frac{d_2}{dt_2}\langle\mathbf{x}\rangle = \frac{d}{dt}\left\langle\frac{p_x}{m}\right\rangle = \frac{-1}{m}\left\langle\frac{\partial V}{\partial x}\right\rangle \tag{62b}$$

by employing Eq. (8); similarly, one obtains Eq. (62*a*) by employing Eq. (8); similarly, one obtains Eq. (62*b*). Equations (62) mean (i) the average position of the particle, that is, the center of the packet, moves with a velocity given by the expectation value of the momentum; (ii) the acceleration of the center of the packet is found from the expectation value of the classical force $-\partial V/\partial x$. Equations (62) illustrate the correspondence between quantum and classical mechanics and show that the classical description of particle motion is valid when the spread of the packet about its mean position can be ignored. When the particle is free, $H = p^2/2m$, Eqs. (8), (45), and (63), where the subscripts t, 0 refer, respectively, to

$$(\Delta x)^2_t = (\Delta x)^2_0 + \left\{\frac{t}{m}\langle xp + px\rangle_0 - 2\langle x\rangle_0\langle p\rangle_0\right\} + \frac{t^2}{m^2}(\Delta p_x)^2_0 \tag{63}$$

expectation values at t and at initial time zero. Equation (63) shows that, although an unconfined free wave packet may contract for a while, ultimately it will spread over all space; when the minimum spread happens to occur at $t = 0$, the term linear in t vanishes in Eq. (63).

Orbital angular momentum. The quantum mechanical operators representing the components of orbital angular momentum $\mathbf{L}$ have the same form as in classical mechanics, namely, for each particle $L_x = yp_z - zp_y$, and so on. By using Eq. (8), one obtains Eqs. (64),

$$L^2L_z - L_zL^2 = 0 \tag{64a}$$

$$L_xL_y - L_yL_x = i\hbar L_z \tag{64b}$$

and so on, where $L^2 \equiv L_x{}^2 + L_y{}^2 + L_z{}^2$. According to Eq. (48), therefore (i) L^2 and L_z are simultaneously exactly measurable; (ii) once the values of L^2 and L_z are specified, the values of L_x and L_y must be uncertain. In spherical coordinates $z = r\cos\theta$, $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, L_z and L^2 becomes Eqs. (65).

$$L_z = \frac{\hbar}{i}\frac{\partial}{\partial\phi} \tag{65a}$$

$$L^2 = -\hbar^2\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right) \tag{65b}$$

Equation (14) for L_z is solved by $u(\phi,m) = \exp(im\phi)$, where $m\hbar$ is the eigenvalue. It can be argued that $u(\phi,m)$ must have a unique value at any point x, y, z, meaning $u(\phi + 2\pi, m) = u(\phi, m)$, so that m must be a positive or negative integer, or zero. With $\partial^2/\partial\phi^2 = -m^2$ in Eq. (65*b*), the eigenvalues of L^2 turn out to be $l(l+1)\hbar^2$, where $l = 0, 1, \ldots$, independent of m, except that for each l the allowed values of the magnetic quantum number m are $m = -l, -l+1, \ldots, l-1, l$; thus each l has order of degeneracy $2l + 1$. Because L^2 and L_z commute with P, the eigenfunctions $u(l,m)$ have definite parity; in fact, Eq. (66) holds.

$$P_u(l, m) = (-1)^l u(l, m) \tag{66}$$

In a two-particle system, the components of the total orbital angular momentum $\mathbf{L} = \mathbf{L}_1 + \mathbf{L}_2$ obey the same commutation, Eqs. (64); as a result the eigenvalues of L^2 and L_z (but not the eigenfunctions) are the same as in the one-particle case. L^2 and L_z commute with $L_1{}^2$ and $L_2{}^2$, but L^2 does not commute with L_{1z} or L_{2z}. Consequently the total orbital angular momentum eigenfunctions are labeled by l, m, l_1, l_2. For given l_1, l_2, the possible values of l are the positive integers from $l = l_1 + l_2$ down to $l = |l_1 - l_2|$; the corresponding eigenfunctions have parity $(-)^{l1+l2}$ independent of l,m. These rules for combining angular momenta are readily generalized to more complicated systems including spin, are well established, and form the basis for the vector model of the atom. *See* ATOMIC STRUCTURE AND SPECTRA.

Coulomb potential. The hamiltonian for an (assumed spinless) electron of mass m_e in the field of a fixed charge Ze is $H = \mathbf{p}^2/2m_e + V(r)$ where $V(r) = -Ze^2/r$. In spherical coordinates, Eq. (14) for

the eigenfunctions $u(r,\theta,\phi)$ is written as Eq. (67),

$$Hu = \left[\frac{-\hbar^2}{2m_e}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{2m_e r^2}L^2 + V(r)\right]u$$

$$= Eu \qquad (67)$$

with L^2 defined by Eq. (65*b*). Now H commutes with L^2 and L_z and r^2H is separable; that is, Eq. (68)

$$r^2H(r,\theta,\phi) = H_1(r) + (2m_e)^{-1}L^2(\theta,\phi) \qquad (68)$$

holds [compare Eq. (61)]. Thus $u(r)u(l,m)$ are a complete set of eigenfunctions; $L^2u(r)u(l,m) = l(l+1)\cdot\hbar^2u(r)u(l,m)$, and therefore the radial eigenfunctions $u(r) \equiv u(E,l)$ must satisfy Eq. (69).

$$\left[\frac{-\hbar^2}{2m_e}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{l(l+1)\hbar^2}{2m_e r^2} + V(r)\right]u(r)$$

$$= Eu(r) \qquad (69)$$

The positive term $l(l+1)\hbar^2/2m_er^2$ acts as an added repulsive potential; it can be understood in terms of the classical centripetal force needed to maintain the angular momentum. For $E < 0$, admissible solutions to Eq. (69) must be exponentially decreasing at $r = \infty$; moreover, because of the r^{-1} and r^{-2} terms in Eq. (69), an eigenfunction $u(E,l)$ which behaves properly at $r = \infty$ becomes infinite at $r = 0$ unless E is specially chosen. Thus (as always) the quadratically integrable eigenfunctions form a discrete set. The corresponding bound state energies $E < 0$ are given by Eq. (70).

$$E = -\frac{m_eZ^2e^4}{2\hbar^2n^2} \qquad (70)$$

In Eq. (67) the principal quantum number $n = 1, 2, 3, \ldots$; for given l and n the number $n_r \geqq 0$ of zeros (between $r = 0$ and $r = \infty$) of the corresponding $u(E,l)$ is $n_r = n - l - 1$. Because $dx\,dy\,dz = r^2 \sin\theta\, dr\, d\theta\, d\phi$, the radial probability density is $r^2|u(r,E,l)|^2$. **Figure 9** shows the radial probability density plotted versus r (in units of the Bohr radius $a_0 = \hbar^2/m_ee^2 \cong 5 \times 10^{-9}$ cm) for several low-lying stationary states of atomic hydrogen, $Z = 1$. The notation for the eigenfunctions is standard in atomic physics: the principal quantum number is supplemented by a lowercase letter $s, p, d, \ldots$ corresponding to $l = 0, 1, 2, \ldots$; for example, a $3d$ state has $l = 2$ and therefore $n_r = 0$ or no radial nodes, as in Fig. 9. The eigenfunction $u(l = 0, m = 0)$ is a constant; that is, an s state is spherically symmetric, $|u(r,\theta,\phi)|^2$ is proportional to $\cos^2\theta$ or to $\sin^2\theta$ in p states, and so on. The eigenfunctions, although they are spread over all space, have their maxima at about the radii expected on the older Bohr theory of Eq. (70).

In the actual hydrogen atom the nucleus, of mass M, is not fixed. The hamiltonian is Eq. (71*a*), where

$$H = \frac{p_1{}^2}{2m_e} + \frac{p_2{}^2}{2M} - \frac{Ze^2}{|r_1 - r_2|} \qquad (71a)$$

$$H = \frac{p_R{}^2}{2(M+m_e)} + \frac{p^2}{2\mu} - \frac{Ze^2}{r} \qquad (71b)$$

the subscripts 1 and 2 refer to the electron and the nucleus, respectively. Introducing the center of mass $(X,Y,Z) \equiv \mathbf{R} = (M + m_e)^{-1}[m_e\mathbf{r}_1 + M\mathbf{r}_2]$, $(x,y,z) \equiv \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$, Eq. (71*a*) takes the separable form of Eq. (71*b*), where $(\mathbf{p}_R)_x \equiv p_X = (\hbar/i)\partial/\partial X$, etc., are the components of the total momentum $\mathbf{p}_R = \mathbf{p}_1 + \mathbf{p}_2$; $\mathbf{p}_x = (\hbar/i)\partial/\partial x$, etc.; and the reduced mass $\mu = m_eM(M + m_e)^{-1}$. Therefore $\mathbf{p}_R$ is a constant of the motion, as asserted in connection with Eq. (56); moreover, Eq. (71*b*) is separable in $\mathbf{R}$ and $\mathbf{r}$. In other words, the center of mass moves like a free particle, completely independent of the internal $\mathbf{r}$ motion. Comparing Eqs. (67) and (71*b*), and recalling Eq. (61), the eigenvalues of Eq. (71*a*), after the kinetic energy of the center of mass is subtracted, are given by Eq. (70) with μ (depending on m_e/M) replacing m_e. The independence of internal and center-of-mass motion means that the temperature broadening of spectral lines can be explained quantum-mechanically in terms of the classical Doppler effect for a moving fixed-frequency source. This paragraph illustrates the correspondence between the classical and quantum theories. *See* CENTER OF MASS.

The eigenvalues E of Eq. (70) have order of degeneracy n^2; this degeneracy stems from (i) the fact that $V(r)$ is spherically symmetric, which permits H to commute with $\mathbf{L}$; (ii) a special symmetry of Eq. (67) for the specially simple Coulomb potential, causing solutions of Eq. (69) for different l to have the same energy. For an arbitrary spherically symmetric $V(r)$, the bound-state eigenvalues $E(n_r,l)$ of Eq. (69) do not coincide for different l, and each bound state has degeneracy $2l + 1$, corresponding to the $2l + 1$ possible values $m = -l$ to l of the magnetic quantum number

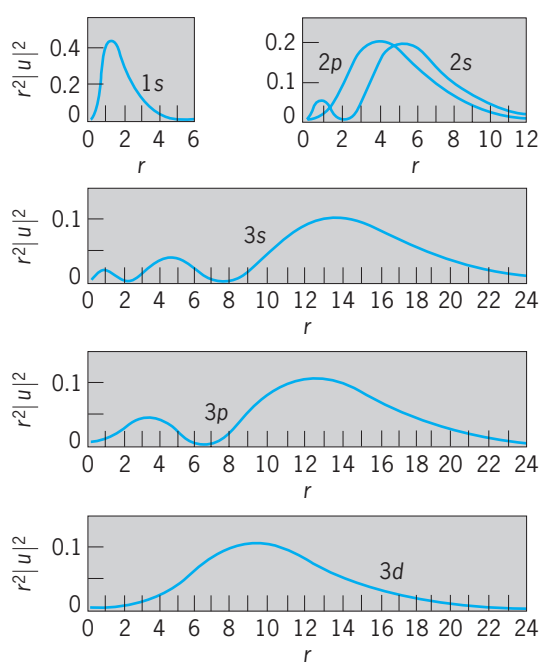

Fig. 9. Radial probability density in atomic hydrogen. ***(After F. K. Richtmyer, E. H. Kennard, and T. Lauritsen, Introduction to Modern Physics, 5th ed., McGraw-Hill, 1955)***

m; an energy level associated with orbital angular momentum l has parity $(-)^l$. For any such $V(r)$ the bound-state energies (i) increase with increasing n_r for a constant value of l, for the same reasons that were discussed in connection with the eigenvalues of Eq. (28); (ii) increase with increasing l for constant n_r because the rotational kinetic energy $(2m_e r^2)^{-1} \cdot l(l+1)\hbar^2$ increases. Except for these regularities, the order and spacing of the levels depends on the details of $V(r)$. For potentials $V(r)$ decreasing more rapidly than $1/r$, the total number of bound states generally is finite, whereas this number is infinite for the Coulomb $V(r) = -Ze^2/r$.

Removal of degeneracy. The hamiltonian of Eq. (67) is spin-independent, which is why Eq. (67) has been treated as if the wave function had only one component; compare the remarks following Eqs. (5). Corresponding to any one-component solution u of Eq. (67), there are two independent two-component eigenfunctions; (i) $\psi_1 = u$, $\psi_2 = 0$ and (ii) $\psi_1 = 0$, $\psi_2 = u$; compare the earlier discussion of spin. Thus for an electron the degeneracy of the energy levels in a Coulomb field is $2n^2$; similarly the degeneracy for neutrons, protons, or electrons in an arbitrary spherically symmetric potential is $2(2l+1)$. The energy operator for an electron in an actual atom, for example, hydrogen, is not spin-independent, however. Relativistic effects add, to the central $V(r)$ of Eq. (67), noncentral spin-orbit potentials $V'(r)[L_x s_x + L_y s_y + L_z s_z] \equiv V'(r)\,\mathbf{L}\cdot\mathbf{s}$; here $\mathbf{s}$ is the spin operator and obeys the same commutation Eqs. (64) as $\mathbf{L}$.

Equations (64) show that $V'(r)\,\mathbf{L}\cdot\mathbf{s}$ commutes with L^2, s^2. $(\mathbf{L}+\mathbf{s})^2$, and $L_z + s_z$, but not with L_z or s_z, illustrating the principle of conservation of total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{s}$; compare the remarks preceding Eq. (56). Consequently, referring to the final part of the discussion of orbital angular momentum (i) $J^2 = j(j+1)\hbar^2$; (ii) for given $l \neq 0$ (and $s = {}^1/_2$); j has but two possible values $l \pm {}^1/_2$; (iii) the energy levels are labeled by j and have a $(2j+1)$-fold degeneracy corresponding to the $2j+1$ possible orientations of $J_z = -j$ to $+j$; (iv) because $2\mathbf{L}\cdot\mathbf{s} = (\mathbf{L}+\mathbf{s})^2 - L^2 - s^2$, levels of different j have different energies, and the splitting of the energies depends predictably on l; (v) L_z (and s_z) no longer are constants of the motion, although L^2 and $s^2 = {}^1/_2 \cdot {}^3/_2\hbar^2$ still are. Moreover, because $2j+1 = 2l+2$ for $j = l + {}^1/_2$, and $= 2l$ for $j = l - {}^1/_2$, the total number of independent eigenfunctions associated with given l (and n_r) remains $2(2l+1) = (2l+2) + (2l)$.

In the independent particle model of atoms and nuclei, one assumes that to a first approximation each particle, particle i, say, moves in a potential $V(r_i)$ which has been averaged over the coordinates of all the other particles, so that $V(r_i)$ depends only on the distance of i from the atomic or nuclear center. To a first approximation, therefore, the energy levels are associated with configurations of one-particle eigenfunctions; for example, the ground state of atomic beryllium is $1s^2 2s^2$. In higher approximation one introduces two-body interactions $V(r_i, r_j)$, which may be said to mix different configurations. The considerations of this and the paragraphs just preceding, together with the exclusion principle discussed subsequently, account for the periodic system of the elements and are the basis for the highly successful nuclear shell model.

The observation that splitting the levels does not change the number of independent eigenfunctions illustrates a general principle and justifies the postulate that the statistical weight of a discrete level equals its order of degeneracy. This principle can be understood on the basis that the number of bound-state eigenfunctions should be a continuous function of the parameters in a reasonably well-behaved hamiltonian; because this number by definition is an integer, it must change discontinuously and, therefore, except under unusual mathematical circumstances, cannot change at all. *See* BOLTZMANN STATISTICS; STATISTICAL MECHANICS.

In an external magnetic field B the hamiltonian of a many-electron atom (i) no longer is independent of the orientation of the coordinate axes, so that the degeneracy associated with this symmetry is removed; (ii) retains symmetry with respect to rotation about the magnetic field, so that (with the z axis along B) J_z commutes with the hamiltonian. Thus in a magnetic field a level associated with the total angular momentum quantum number j should split into $2j+1$ levels, each of which is associated with one of the magnetic quantum numbers $-j$ to $+j$. This prediction is thoroughly confirmed in the Zeeman effect and in the Stern-Gerlach experiment. *See* ZEEMAN EFFECT.

Radiation. The classical hamiltonian for a charged particle in an electromagnetic field has the form given by Eq. (72), where $\mathbf{A}$ and ϕ are the scalar

$$H = \frac{1}{2m}\left(\mathbf{p} - \frac{e\mathbf{A}}{c}\right)^2 + e\phi \tag{72}$$

and vector potentials, respectively. *See* ELECTRON MOTION IN VACUUM.

It is postulated that when properly symmetrized, that is, when $\mathbf{A}\cdot\mathbf{p} + p\cdot A$ replaces the classical $2\mathbf{A}\cdot\mathbf{p}$ (see the earlier discussion on hermitian operators), H of Eq. (72) is the quantum-mechanical energy operator. The presence of terms linear in $\mathbf{p}$ modifies some of the formulas which have been given, for example, Eq. (55*a*). When plane waves of light (frequency f, wavelength λ, and moving in the z direction) fall on a hydrogen atom, $\mathbf{A}$ is proportional to cos $[2\pi(z/\lambda - ft)]$, and $e\phi$ is the Coulomb potential $V(r)$ of Eq. (67).

Proceeding as in Eqs. (57)–(60), noting that $\mathbf{A}$ contains terms that are proportional to $\exp(2\pi ift)$ and $\exp(-2p\pi ift)$, and neglecting the small (as can be shown) A^2 terms, one obtains an expression similar to Eq. (59), except that $\omega \pm 2\pi f$ replaces $\omega \equiv \hbar^{-1} \cdot (E_f - E_i)$. In other words, after a long time there are appreciable transition probabilities only to final states f whose energies satisfy $E_f - E_i \pm hf = 0$, in agreement with the notion that a quantum of energy hf has been emitted or absorbed; the $+$ sign corresponds to emission, the $-$ sign to absorption. The corresponding transition probabilities are given by Eq. (60) with V' proportional to $\exp\,[2\pi iz/\lambda]$,

and u_i, u_f stationary-state atomic wave functions satisfying the radiation-unperturbed Eq. (67). The final expressions are analogous to the classical formulas for emission or absorption of radiation; for instance, when the wavelength λ is large compared to atomic dimensions, expanding $\exp[2\pi iz/\lambda]$ in powers of z/λ shows that the leading term in the transition probability is the matrix element, between initial and final states, of the dipole moment ez corresponding to classical electric dipole emission or absorption. Because z changes sign on reflection through the origin, the dipole matrix element vanishes unless u_i and u_f have opposite parities. This is one of the selection rules for electric dipole radiation; other selection rules, connected with angular momentum conservation, are obtained similarly. *See* ELECTROMAGNETIC RADIATION; SELECTION RULES (PHYSICS).

The theory starting with Eq. (72) is termed semiclassical, because it does not replace the classical $\mathbf{A}$, ϕ by a quantum mechanical operator description of the electromagnetic field. This semiclassical theory has led to the induced emission and absorption probabilities in the presence of external radiation, but not to the spontaneous probability of emission of a photon in the absence of external radiation. The spontaneous transition probabilities can be inferred from the induced probabilities by thermodynamic arguments. The spontaneous transition probability is deduced directly, however, without appeal to the arguments of thermodynamics, when the radiation field is quantized.

Particle Indistinguishability

For systems of g identical, and therefore indistinguishable, particles, the formalism is further complicated because the probability P of finding a given particle in a specified volume $dx\,dy\,dz$ of space must be $P_1 + P_2 + \cdots + P_g$, where $P_1, \ldots, P_g$ are the probabilities given previously for distinguishable particles; expectation values and normalizations must be reinterpreted accordingly. Moreover, the Pauli exclusion principle asserts that the only physically permissible wave functions must change sign when the space and spin coordinates of any pair of indistinguishable particles of spin $^1/_2$ are interchanged. *See* EXCLUSION PRINCIPLE.

To amplify this assertion, consider the four-component wave function of the two electrons in atomic helium in Eq. (73), where $|\psi_{++}(\mathbf{r}_1,\mathbf{r}_2)|^2$ is the

$$\psi = \psi_{++}(\mathbf{r}_1, \mathbf{r}_2),\ \psi_{+-}(\mathbf{r}_1, \mathbf{r}_2),\ \psi_{-+}(\mathbf{r}_1, \mathbf{r}_2),\ \psi_{--}(\mathbf{r}_1, \mathbf{r}_2) \quad (73)$$

probability density for finding both electrons with spin along $+z$, and so forth. The exclusion principle requires validity of Eqs. (74).

$$\begin{aligned} \psi_{++}(\mathbf{r}_1, \mathbf{r}_2) &= -\psi_{++}(\mathbf{r}_2, \mathbf{r}_1) \\ \psi_{--}(\mathbf{r}_1, \mathbf{r}_2) &= -\psi_{--}(\mathbf{r}_2, \mathbf{r}_1) \\ \psi_{+-}(\mathbf{r}_1, \mathbf{r}_2) &= -\psi_{-+}(\mathbf{r}_2, \mathbf{r}_1) \end{aligned} \quad (74)$$

In the independent particle approximation, all components of the ground-state eigenfunctions of atomic He are composed of products $u(\mathbf{r}_1)u(\mathbf{r}_2)$, where $u(\mathbf{r})$ is the lowest ls eigenfunction solving Eq. (67) for the spherically symmetric $V(r)$ appropriate to He. In this approximation, therefore, $\psi_{++} = \psi_{--} = 0$; if $\psi_{+-} = u(\mathbf{r}_1)u(\mathbf{r}_2)$, ψ_{-+} is necessarily equal to $-u(\mathbf{r}_1)u(\mathbf{r}_2)$. Of the four independent ls^2 eigenfunctions originally possible, only one remains, which can be shown to be a total spin zero eigenfunction; the exclusion principle literally has excluded the three eigenfunctions corresponding to total spin one.

These results are summarized by the rule that at most two electrons, $sz = \pm^1/_2$, can occupy one-particle states with the same quantum numbers n_r, l, and $m = -l$ to l. By the general principle explained previously in the discussion of removal of degeneracy, the introduction of two-particle interactions $V(\mathbf{r}_1,\mathbf{r}_2)$ does not change the number of independent eigenfunctions consistent with the exclusion principle. In the next higher ls2s configuration of He, ψ_{++} will equal $u_{1s}(\mathbf{r}_1)u_{2s}(\mathbf{r}_2) - u_{1s}(\mathbf{r}_2)u_{2s}(\mathbf{r}_1)$. This antisymmetrized wave function makes nonclassical exchange energy contributions, of the form given by expression (75), interaction to the expectation

$$\int u_{1s}^*(\mathbf{r}_1)u_{2s}^*(\mathbf{r}_2)V(\mathbf{r}_1, \mathbf{r}_2)u_{1s}(\mathbf{r}_2)u_{2s}(\mathbf{r}_1) \quad (75)$$

value $\langle V(\mathbf{r}_1,\mathbf{r}_2)\rangle$ of the energy. Exchange energies are important for understanding chemical binding.

Except for the complication of spin, this article presupposes that electrons, neutrons, and protons are immutable structureless mass points. Actually, modern theories of nuclear forces and high-energy scattering experiments indicate that this assumption is untrue. When creation, destruction, or other alterations of fundamental particle structure are improbable, however, the general separability of internal and center-of-mass motion [see the discussion following Eq. (71*b*)] implies that each fundamental particle is sufficiently described by the position of its center of mass and by its spin orientation, that is, by the many-component wave functions used here.

In circumstances wherein more obviously composite systems undergo no changes in internal structure, they too can be treated as particles. For instance, in the slow collisions of two atoms the slowly changing potentials acting on the electrons do not induce transitions to new configurations, and the collision can be described by solving an equation of the form (67) for the relative motion; in rapid collisions electron transitions occur, and the many-electron Schrödinger equation must be employed. Similarly, since a deuteron is a neutron-proton bound state, with total angular momentum unity, in a deuterium molecule (i) each deuteron can be treated as if it were a fundamental particle of spin 1; (ii) the wave function of a deuterium molecule must be symmetric under interchange of the space and spin coordinates of the two deuterons, which interchange involves successive antisymmetric interchanges of the two neutrons and of the two protons. In other words

(when they can be treated as particles) deuterons and other composite systems of integral spin obey Bose-Einstein statistics; composite systems of half-integral spin obey Fermi-Dirac statistics.

When the particles composing a many-particle system can be represented by nonoverlapping wave packets which spread an amount $\ll \Delta x$ as the center packet moves a distance equal to its width Δx, individual classical particle trajectories can be distinguished; under these circumstances, the particles, whether or not identical, are in effect distinguishable classical particles, and one expects Bose-Einstein and Fermi-Dirac statistics to reduce to the classical Maxwell-Boltzmann statistics. The well-known condition for the validity of classical statistics in an electron gas, $Nb^3(2\pi mkT)^{-3/2} \ll 1$, implies that such packets can be constructed; here N is the electron density, k is Boltzmann's constant, and T is the absolute temperature. In the lower vibrational states of a molecule such packets cannot be constructed for the vibrating nuclei, however (compare the discussion of the harmonic oscillator presented earlier), so that quantum statistics cannot be ignored, for example, in the specific heat of H_2 at low temperatures. *See* QUANTUM STATISTICS. Edward Gerjuoy

Bibliography. L. E. Ballentine, *Quantum Mechanics*, 2d ed., 1998; A. Bohm, *Quantum Mechanics: Foundations and Applications*, 3d ed., 1994; B. H. Bransden and C. J. Joachim, *Introduction to Quantum Physics*, 1989; P. A. M. Dirac, *Principles of Quantum Mechanics*, 4th ed., 1958; H. Ohanian, *Principles of Quantum Mechanics*, 1989; D. A. Park, *Introduction to the Quantum Theory*, 3d ed., 1992; R. Shankar, *Principles of Quantum Mechanics*, 2d ed., 1994.

Nonsinusoidal waveform

The representation of a wave that does not vary in a sinusoidal manner. Electric circuits containing nonlinear elements, such as iron-core magnetic devices, rectifying devices, and transistors, commonly produce nonsinusoidal currents and voltages. When these are repetitive functions of time, they are called nonsinusoidal periodic electric waves. Oscillograms, tabulated data, and sometimes mathematical functions for segments of such waves are often used to describe the variation through one cycle. The term cycle corresponds to 2π electrical radians and covers the period, which is the time interval T in seconds in which the wave repeats itself.

These electric waves can be represented by a constant term, the average or dc component, plus a series of harmonic terms in which the frequencies of the harmonics are integral multiples of the fundamental frequency. The fundamental frequency f_1, if it does exist, has the time span $T = 1/f_1$ s for its cycle. The second-harmonic frequency f_2 then will have two of its cycles within T seconds, and so on.

Fourier series representation. The series of terms stated above is known as a Fourier series and can be expressed in the form of Eq. (1), where $y(t)$, plotted over a cycle of the fundamental, gives the shape of the nonsinusoidal wave. The terms on the right-hand side show the Fourier series representation of the wave where Eqs. (2) and (3) apply. Here A_0 is identically zero, and $B_0 = C_0$ in Eq. (1). The radian frequency of the fundamental is $\omega = 2\pi f_1$, and n is either zero or an integer. C_1 is the amplitude of the fundamental ($n = 1$), and succeeding C_n's are the amplitudes of the respective harmonics having frequencies corresponding to $n = 2, 3, 4$, and so on, with respect to the fundamental. The phase angle of the fundamental with respect to a chosen time reference axis is ϕ_1, and the succeeding ϕ_n's are the phase angles of the respective harmonics. *See* FOURIER SERIES AND TRANSFORMS.

$$\begin{aligned} y(t) &= B_0 + A_1 \sin \omega t + A_2 \sin 2\omega t + \cdots \\ &\quad + A_n \sin n\omega t + B_1 \cos \omega t + B_2 \cos 2\omega t \\ &\quad + \cdots + B_n \cos n\omega t \\ &= B_0 + C_1 \sin(\omega t + \phi_1) + \cdots \\ &\quad + C_n \sin(n\omega t + \phi_n) \\ &= \sum_{n=0}^{\infty} C_n \sin(n\omega t + \phi_n) \end{aligned} \tag{1}$$

$$C_n = \sqrt{A_n^2 + B_n^2} \tag{2}$$

$$\phi_n = \arctan \frac{B_n}{A_n} \tag{3}$$

The equation for $y(t)$ in general, includes an infinite number of terms. In order to represent a given nonsinusoidal wave by a Fourier series, it is necessary to evaluate each term, that is, B_0, and all A_n's and all B_n's. In practice, the first several terms usually yield an approximate result sufficiently accurate for portrayal of the actual wave. The degree of accuracy desired in representing faithfully the actual wave determines the number of terms that must be used in any computation.

The constant term B_0 is found by computing the average amplitude of the actual wave over one cycle. Assuming any reference time $t = 0$ on the wave, B_0 is given by Eq. (4), where the angle ωt is replaced by θ. Since B_0 is a constant, or dc, term, it merely raises or lowers the entire wave and does not affect its shape.

$$\begin{aligned} B_0 &= \frac{1}{T}\int_0^T y(t)\,dt = \frac{1}{2\pi}\int_0^T y(\omega t)\,d\,\omega t \\ &= \frac{1}{2\pi}\int_0^T y(\theta)\,d\theta \end{aligned} \tag{4}$$

The coefficients of the sine series are obtained by multiplying the wave $y(\theta)$ by $\sin n\theta$, integrating this product over a full cycle of the fundamental, and dividing the result by π. Thus, A_n is given by Eq. (5).

$$A_n = \frac{1}{\pi}\int_0^{2\pi} y(\theta) \sin n\theta\,d\theta \tag{5}$$

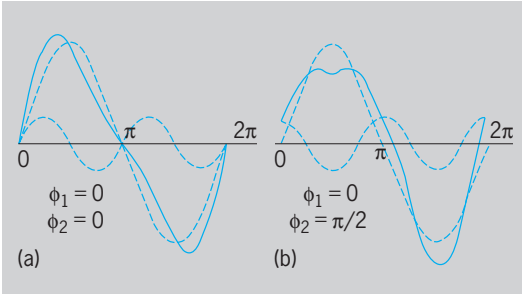

Fig. 1. Addition of a fundamental and a second harmonic. (*a*) $\phi_1 = 0$, $\phi_2 = 0$. (*b*) $\phi_1 = 0$, $\phi_2 = +\pi/2$ radians.

The coefficients of the cosine terms are obtained in like manner, except that cos $n\theta$ replaces sin $n\theta$. Thus, B_n is given by Eq. (6).

$$B_n = \frac{1}{\pi}\int_0^{2\pi} y(\theta)\cos n\theta \, d\theta \tag{6}$$

If mathematical expressions are required to describe $y(\theta)$, Eqs. (4), (5), and (6) give the coefficients of the series directly through analytical methods. If oscillograms or tabulated data describe $y(\theta)$, then graphical or tabular forms of integration are used.

Effect of even harmonics. Figure 1 shows waves composed of a fundamental and a second harmonic only. In Fig. 1*a*, both ϕ_1 and ϕ_2 are zero. In Fig. 1*b*, ϕ_1 is zero and ϕ_2 is $+\pi/2$ radians with respect to one cycle of the second harmonic. In the example given in Fig. 1*b* the negative part of the overall wave is completely unlike the positive portion. Also, in general, these two portions will have different time intervals. Even harmonics give unsymmetrical waves. *See* HARMONIC (PERIODIC PHENOMENA).

Effect of odd harmonics. Figure 2 shows waves composed of a fundamental and a third harmonic. In Fig. 2*a*, both ϕ_1 and ϕ_3 are zero. In Fig. 2*b*, ϕ_1 is zero and ϕ_3 is $+\pi$ radians with respect to one cycle of the third harmonic. In Fig. 2*c*, θ_1 is zero and θ_3 is $+\pi/2$ radians on the third-harmonic time scale. In these diagrams the negative and positive parts of the overall waves are alike and both embrace π radians of the fundamental. Odd harmonics lead to symmetrical waves.

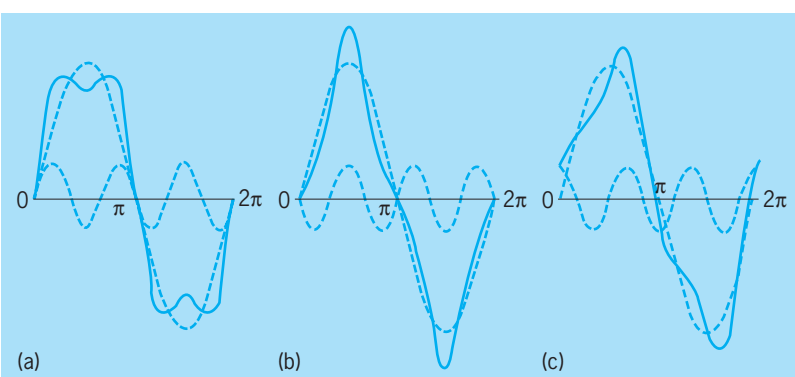

Fig. 2. Addition of a fundamental and a third harmonic. (*a*) $\phi_1 = 0$, $\phi_3 = 0$. (*b*) $\phi_1 = 0$, $\phi_3 = +\pi$ radians. (*c*) $\phi_1 = 0$, $\phi_3 = +\pi/2$ radians.

Symmetry. To determine symmetry of nonsinusoidal waves, the constant term b_0 is first removed. This means shifting the wave down or up by the value of B_0. After this, if the wave from π to 2π is rotated about the horizontal axis and moved forward π radians, and then coincides exactly with the section from 0 to π, the total wave is said to have half-wave symmetry. Often the wave is said merely to be symmetrical. This means that in such cases $y(\theta + \pi) = -y(\theta)$. If, in turn, each half of the wave is symmetrical about the vertical axes of $\pi/2$, or $3\pi/2$, the wave is said to have quarter-wave symmetry as well. Half-wave and quarter-wave symmetry do not necessarily accompany each other.

All three waves of Fig. 2 have half-wave symmetry. The first two have quarter-wave symmetry also, but that of Fig. 2*c* does not. Waves having only a fundamental and odd harmonics show half-wave symmetry. Conversely, half-wave symmetry indicates that only the fundamental, if it exists, and odd harmonics are present in the total wave. Half-wave symmetry permits Eqs. (5) and (6) to be integrated over the interval π, with the result multiplied by two. Quarter-wave symmetry permits integration over one-quarter cycle of the fundamental, with the result multiplied by four.

Half-wave symmetry means that the fundamental and all odd harmonics may pass through their zero values at times quite distinct from each other. With quarter-wave symmetry the fundamental and the odd harmonics all pass through their zero values at the same time; therefore all phase angles ϕ_n are either zero or 180°.

The wave of Fig. 1*a* has no symmetry of the kind discussed above. Waves containing only the fundamental and even harmonics, or even harmonics alone, are unsymmetrical. Although half-wave symmetry is absent, quarter-wave symmetry may exist. A wave which is the same from π to 2π as it is from 0 to π, that is, $y(\theta) = y(\theta + \pi)$, is unsymmetrical. Only even harmonics are present, and the wave has no fundamental component. The output from a full-wave rectifier, for example, contains only the average term B_0 and even harmonics.

Waves that do not meet any of the special conditions noted above can be expected to contain both even and odd harmonics, and probably both sines and cosines also. Any doubt arising on the harmonic content of a wave is resolved by assuming all components of the Fourier series to be present. Analysis will show exactly those that actually exist.

Even and odd functions. The time origin of a wave can be chosen arbitrarily. If the reference axis $t = 0$ is such that the wave to the left is merely the wave to the right rotated about this axis, then $y(-\theta) = y(\theta)$, which is said to be an even function. Only cosine terms will be present in the Fourier series for the wave. On the other hand, if the wave to the left is the wave to the right rotated about the $t = 0$ axis, and then rotated about the horizontal axis,

$y(-\theta) = -y(\theta)$, which is said to be an odd function. Only sine terms will appear in the Fourier series. Neither case precludes the possibility of the presence of both even and odd harmonics.

The rms value of nonsinusoidal wave. A nonsinusoidal wave has an rms value obtained through the following steps:

1. Combine all terms of the same frequency so as to have a single $A_1, A_2, \ldots, A_n$; $B_1, B_2, \ldots, B_n$; or a single $C_1, C_2, \ldots, C_n$. Terms such as $\sin(n\omega t \pm \alpha)$ and $\cos(n\omega t \pm \beta)$ each have sine and cosine components which can be separated out by trigonometric expansion.
2. Form the series $y(\theta)$ as in Eq. (1).
3. The rms value of the wave is then given by Eq. (7).

$$y_{\text{rms}} = \left(B_0^2 + \frac{A_1^2}{2} + \frac{A_2^2}{2} + \cdots + \frac{A_n^2}{2} + \frac{B_1^2}{2} + \frac{B_2^2}{2} + \cdots + \frac{B_n^2}{2}\right)^{1/2} = \left(B_0^2 + \frac{C_1^2}{2} + \frac{C_2^2}{2} + \cdots + \frac{C_n^2}{2}\right)^{1/2} \quad (7)$$

If y_{rms} represents a voltage or a current, this value is shown by an electrodynamometer or iron-vane voltmeter or ammeter. The rms of the wave is merely the square root of the sum of the squares of the rms values of all of the frequency components. *See* ROOT-MEAN-SQUARE.

Power. An indicating wattmeter with a nonsinusoidal voltage impressed on its potential circuit and a nonsinusoidal current in its current coils indicates the average power taken by the circuit. Designating peak values of the component voltages and currents by E_n's and I_n's in place of C_n's results in Eq. (8). Each

$$\text{Average power} = \frac{1}{2\pi}\int_0^{2\pi} ei\, d\theta = E_0 I_0 + \frac{E_1 I_1}{2}\cos\theta_1 + \cdots + \frac{E_n I_n}{2}\cos\theta_n \quad (8)$$

coefficient is simply the product of arms voltage and current. No cross-product terms involving different frequencies result from the integration. That is, no power can be contributed by a voltage of one frequency and a current of another frequency.

Power factor. The apparent power taken by a circuit carrying nonsinusoidal voltage and current may be defined as the product of the rms values of these quantities. A power factor for such a case may be defined only by the ratio of the average power to the apparent power. Thus, the power factor is given by Eq. (9). Power factor is hence the ratio of instrument

$$\text{Power factor (pf)} = \frac{\text{watts average power}}{\text{rms volts} \times \text{rms amperes}} \quad (9)$$

readings as stated. All circuits have a power factor, but pf $= \cos\theta$ only for a sine wave voltage and current of the same frequency. There is no average or representative phase angle for a circuit carrying nonsinusoidal waves. *See* POWER FACTOR.

Example of nonsinusoidal waves. Assume a series circuit to have 8 ohms resistance and 15.91 millihenries inductance and that the impressed voltage is given by Eq. (10). The problem is to calculate the

$$e = 100\sin 377t + 80\sin 1131t \quad (10)$$

rms voltage and current, the average power, and the power factor. The voltage has a fundamental component of 60 cycles ($f_1 = \omega_1/2\pi = 377/2\pi$) and a third harmonic of 180 cycles ($f_3 = \omega_3/2\pi = 1131/2\pi$). At 60 cycles:

$$\begin{aligned} X_{L_1} &= 377 \times 0.01591 \\ &= 6.0 \text{ ohms inductive reactance} \\ Z_1 &= \sqrt{8^2 + 6^2} = 10 \text{ ohms impedance} \\ I_1 &= 100/10 = 10 \text{ A max fundamental current} \\ \theta_1 &= \arctan(6/8) = 36.87^\circ \end{aligned}$$

At 180 cycles:

$$\begin{aligned} X_{L_1} &= 3 \times 6 = 18 \text{ ohms inductive reactance} \\ Z_3 &= \sqrt{8^2 + 18^2} = 19.70 \text{ ohms impedance} \\ I_3 &= 80/19.7 \\ &= 4.06 \text{ A max third-harmonic current} \\ \theta_3 &= \arctan(18/8) \\ &= 66.06^\circ \ (= 22.02^\circ \text{ on fundamental scale}) \end{aligned}$$

The equation for the current is

$$i = 10\sin(377t - 36.87^\circ) + 4.06\sin(1131t - 22.02^\circ)$$

$$E_{\text{rms}} = \sqrt{\frac{100^2}{2} + \frac{80^2}{2}} = 90.06 \text{ volts}$$

$$I_{\text{rms}} = \sqrt{\frac{10^2}{2} + \frac{4.06^2}{2}} = 7.63 \text{ amperes}$$

$$\begin{aligned} \text{Apparent power} &= 90.06 \times 7.63 \\ &= 687 \text{ volt-amperes} \\ \text{Average power} &= I^2R = 7.63^2 \times 8 = 466 \text{ watts} \\ \text{Power factor} &= \frac{466}{682} = 0.678 \end{aligned}$$

See ALTERNATING-CURRENT CIRCUIT THEORY.

Burtis L. Robertson; W. S. Pritchett

Bibliography. C. K. Alexander and M. N. O. Sadiku, *Fundamentals of Electric Circuits*, 2d ed., McGraw-Hill, 2004; R. C. Dorf and J. A. Svoboda, *Introduction to Electric Circuits*, 6th ed., Wiley, 2003; T. L. Floyd, *Principles of Electric Circuits*, 7th ed., Prentice Hall, 2002; I. D. Mayergoyz and W. Lawson, *Basic Electric Circuit Theory*, Academic Press, 1997; J. W. Nilsson and S. A. Riedel, *Electric Circuits*, 7th ed., Prentice Hall, 2004; H. B. Tilton, *Waveforms: A Modern Guide to Nonsinusoidal Waves and Nonlinear Processes*, Prentice Hall, 1986.

Nonstoichiometric compounds

Chemical compounds in which the relative number of atoms is not expressible as the ratio of small whole numbers, hence compounds for which the subscripts in the chemical formula are not rational (for example, $Cu_{1.987}S$). Sometimes they are called berthollide compounds to distinguish them from daltonides, in which the ratio of atoms is generally simple. Nonstoichiometry is a property of the solid state and arises because a fraction of the atoms of a given kind may be (1) missing from the regular structure (for example, $Fe_{1-\delta}O$), (2) present in excess over the requirements of the structure (for example, $Zn_{1+\delta}O$), or (3) substituted by atoms of another kind (for example, $Bi_2Te_{3\pm\delta}$). The resulting materials are generally of variable composition, intensely colored, metallic or semiconducting, and different in chemical reactivity from the parent stoichiometric compounds from which they are derived. Nonstoichiometry is best known in the binary compounds of the transition elements, particularly the hydrides, oxides, chalcogenides, pnictides, carbides, and borides. It is also well represented in the so-called insertion or intercalation compounds, in which a metallic element or neutral molecule has been inserted in a stoichiometric host. Nonstoichiometric compounds are important in some solid-state devices (such as rectifiers, thermoelectric generators, and photodetectors) and are probably formed as chemical intermediates in many reactions involving solids (for example, heterogeneous catalysis and metal corrosion).

Generality of the phenomenon. Simple stoichiometry, in which a compound A_xB_y is characterized by small integral values for the composition parameters x and y, is strictly speaking a required characteristic only of molecules in the gaseous state. In the condensed state (solid or liquid), unless the simple molecular aggregate of the gas phase clearly retains its identity, x and y no longer need be small integers. Indeed, they may be as large as the number of atoms in a crystal, for example, 10^{22}. Given such large numbers, small but detectable departures from a simple x:y ratio can be achieved, even in cases such as NaCl, without seriously affecting the energetics of the system, provided a mechanism exists for keeping the compound electrically neutral. In the case of sodium chloride, there is no a priori requirement that the number of Na atoms and Cl atoms be exactly identical. As a matter of fact, it is relatively simple to heat a colorless stoichiometric crystal of composition NaCl in sodium vapor and convert it to yellow-brown $Na_{1.001}Cl$. Similarly, heating it in chlorine gas can produce $NaCl_{1+\delta}$. Both of these deviations from stoichiometry result from Schottky and Frenkel defects, such as those which are native in any real crystal. *See* CRYSTAL DEFECTS; SOLID-STATE CHEMISTRY.

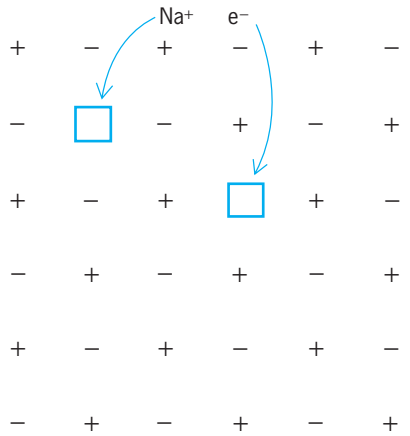

Fig. 1. Formation of $Na_{1+\delta}Cl$ by incorporation of a neutral sodium atom as Na^+; at a vacant cation site and an e^- at a vacant anion site. The +'s represent sodium ions, and the −'s represent chloride ions.

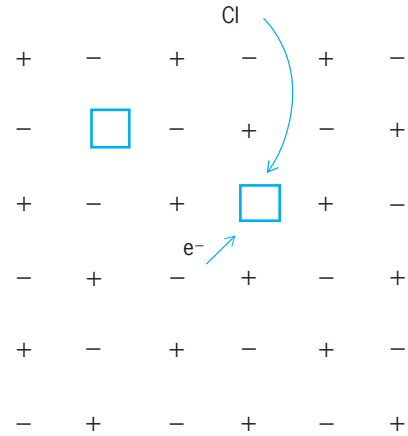

Fig. 2. Formation of $NaCl_{1+\delta}$ by incorporation of a neutral chlorine atom at a vacant anion site and migration of an electron from some chloride ion in the normal structure of the crystal.

Take-up of excess Na can be achieved, as shown in **Fig. 1**, by accommodating Na^+ at a cation vacancy and e^- at an anion vacancy. Excess chlorine is similarly accommodated, as shown in **Fig. 2**, by incorporating the added Cl atom at an anion vacancy. The electron absence, or hole, which distinguishes the Cl^0 from the normal Cl^- ions of the structure, can jump around to any of the other atoms on the anion sublattice. The amount of excess sodium or the amount of excess chlorine that can be accommodated, hence the maximum deviation from stoichiometry, is related to the number of defects in the stoichiometric crystal (intrinsic disorder). Since this intrinsic disorder is an increasing function of temperature, the deviation from stoichiometry that can be achieved is strongly dependent on the preparation conditions, specifically, the temperature at which the preparation is made and the pressure of the constituent elements. Furthermore, since nonstoichiometric compounds are in general prepared by quenching from relatively high temperatures (where atom mobility may be quite high) to room temperature (where diffusion motion is practically negligible), the materials will not usually be in thermodynamic equilibrium with their component elements at room temperature. If the quenching is not efficient, the room-temperature composition will not

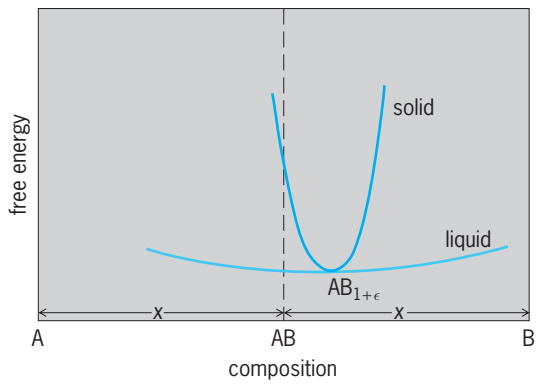

Fig. 3. Free energy versus composition for a two-element system forming the nonstoichiometric compound $AB_{1+\delta}$. Temperature and pressure are constant.

correspond to an equilibrium state at the preparation temperature, either.

Thermodynamic considerations. A thermodynamic description for formation of a nonstoichiometric compound at a fixed temperature is illustrated in **Fig. 3**, which shows possible free-energy relations for the binary system composed of elements A and B forming nonstoichiometric compound $AB_{1+\epsilon}$. The particular case shown corresponds to one in which the compound is completely dissociated in the liquid; that is, there are no specific interactions between A and B in the liquid. At the melting-point equilibrium the chemical potential of each component must be identical in the liquid and solid states. Furthermore, to merit the designation "compound" (as distinct from "solid solution" or "peritectic"), the chemical composition of the solid and liquid phases coexisting at the melting point must be identical. As shown in Fig. 3, this equality is fulfilled not at the stoichiometric composition AB but at $AB_{1+\epsilon}$. This comes about as follows: If in the liquid A and B atoms are completely independent of each other, then starting from AB replacement of some A for B or B for A will have no influence on the enthalpy H and only a slight influence in decreasing the entropy S. The free energy $G = H - TS$ of the liquid will show a rather broad minimum as the composition is varied. In the solid the specific interactions are larger and more important. Replacement of A for B or B for A sharply raises the enthalpy, and the result is a deep minimum in the free-energy curve of the solid. Furthermore, the curve for the solid is generally not symmetric about the ideal composition AB. In the case shown, replacement of A for B raises the enthalpy more than does the replacement of B for A. As a result, the minimum lies in the B-rich region, and the equilibrium compound will be $AB_{1+\epsilon}$. (A similar but parallel argument can be drawn for the case where AB initially has a finite concentration of vacant lattice sites due to Schottky or Frenkel disorder.) *See* CHEMICAL THERMODYNAMICS.

Phase relations. **Figure 4** shows possible phase relations involving formation of the compound $AB_{1+\delta}$ from the melt. The solid $AB_{1+\epsilon}$ will crystallize on cooling a melt containing A and B in the ratio $1{:}1 + \epsilon$. On the other hand, if the melt contains, for example, stoichiometric amounts of A and B, the solidification temperature will be less than T_{max}, and the solid first separating will not be $AB_{1+\epsilon}$ but $AB_{1+\epsilon'}$ ($\epsilon' < \epsilon$). As crystallization proceeds, the liquid composition would change along L_1 and ϵ' would get progressively smaller, eventually even becoming negative. In practice, to keep crystallizing a solid of fixed composition other than $AB_{1+\epsilon}$, means would have to be provided to keep the liquid composition constant. *See* PHASE EQUILIBRIUM.

For simple ionic solids, particularly where the ionic polarizabilities are low and a multiplicity of oxidation states is not possible, the range of nonstoichiometry is generally small. Thus alkali or alkaline earth fluorides are notoriously hard to make as nonstoichiometric compounds. Covalent solids, on the other hand, are generally easier to deviate from stoichiometry, especially when the constituent elements are similar to each other. Bismuth telluride, for example, can be made either bismuth-rich or tellurium-rich by simply putting Bi atoms on Te sites or vice versa. (The unexpected *p*-type semiconductivity of $Bi_2Te_{3-\delta}$ is attributable to the fact that the antistructure defect, Bi on a Te site, has placed a five-valent atom on a site calling for six valence electrons, thereby creating a hole.) Transition elements are particularly good at forming nonstoichiometric compounds, partly because of the capability of the ions to exist in several oxidation states and partly because there is an increased possibility of using *d* orbitals either for covalent bonding or for delocalizing electronic charge through *d*-orbital overlap. Still, many of the broad, reportedly homogeneous regions in which transition-metal-compound composition appears to be continuously variable actually contain within themselves a series of closely related compounds. As an example, the system TiO_x ($1.7 < x < 1.9$) actually includes at least seven discrete oxides, each of which can be described by the common formula Ti_nO_{2n-1}, with n ranging from 4 to

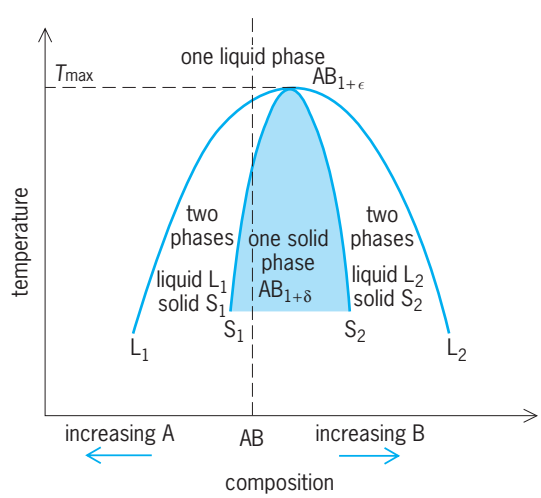

Fig. 4. Temperature-composition plot showing phase relations for formation of the nonstoichiometric compound $AB_{1+\epsilon}$. The tinted region corresponds to a single solid phase of variable composition.

10. Similar series have been recognized in W_nO_{3n-1}, W_nO_{3n-2}, and MoO_{3n-1}.

Even single-crystal compositions that do not fit into such homologous series can frequently be explained as arising from the presence of clusters of defects. X-ray, neutron, and electron diffraction techniques have been supplemented by the technique of lattice imaging in transmission electron microscopy to gain information about the structural nature of defect clusters. One such defect cluster, which appears particularly often, is the crystallographic shear plane, also known as the Wadsley defect. It arises when, along a plane in the crystal, there is a change in the type of polyhedral grouping, as, for example, from corner sharing of octahedra to edge sharing. Each crystallographic shear plane introduces a departure from the stoichiometry of the parent structure; ordered sets of parallel crystallographic shear planes at fixed repetitive spacings can lead to block or column structures as found in different homologous series. In some cases, crystallographic shear planes parallel to two different crystallographic directions, for example, {120} and {130}, may coexist. Given this kind of possible complexity in structure, it is clear that the apparent ranges of homogeneity of phases may well depend upon the wavelength of the radiation used to examine them.

Classification. The simplest way to classify nonstoichiometric compounds is to consider which element is in excess and how this excess is brought about. A classification scheme largely based on this distinction but which also includes some examples of ternary systems is as follows.

Binary compounds:

I. Metal:nonmetal ratio greater than stoichiometric
   a. Metal in excess, for example, $Zn_{1+\delta}O$
   b. Missing nonmetal, for example, $UH_{3-\delta}$, $WO_{3-\delta}$
II. Metal:nonmetal ratio less than stoichiometric
   a. Metal-deficient, for example, $Co_{1-\delta}O$
   b. Nonmetal in excess, for example, $UO_{2+\delta}$
III. Deviations on both sides of stoichiometry, for example, $TiO_{1\pm\delta}$

Ternary compounds (insertion compounds):

IV. Oxide "bronzes," for example, $M_\delta WO_3$, $M_\delta V_2O_5$
V. Intercalation compounds, for example, $K_{1.5+\delta}MoO_3$, $Li_\delta TiS_2$

(Excluded from consideration are the recognized impurity materials, such as $Na_{1-2x}Ca_xCl$, which are best considered as conventional solid solutions wherein ions of one kind and perhaps vacancies have replaced an equivalent number of ions of another kind.) The specific compounds listed here were chosen to be illustrative of structures and properties commonly encountered in nonstoichiometry. They are discussed individually below.

Zinc oxide, $Zn_{1+\delta}O$. Rather small deviations from stoichiometry can be obtained by heating zinc oxide crystals in zinc vapor at temperatures of the order of 600–1200°C (1100–2200°F) The crystals become red and their room-temperature conductivity is considerably enhanced over that of stoichiometric ZnO. The red color and the increased conductivity are attributed to interstitial zinc atoms. Since the ZnO structure is rather open, there is ample room for extra Zn atoms in the tunnels of the $P6_3mc$ structure. The activation energy for diffusion of the Zn is only 0.55 eV, supporting the belief the nonstoichiometry arises from interstitial zinc and not from oxygen vacancy. The conductivity of $Zn_{1+\delta}O$ is *n*-type corresponding to action of Zn as a donor. Hall measurements indicate only one free electron from each Zn atom, thus suggesting Zn^+ ions. Similar properties can be produced by heating ZnO in hydrogen gas, but the cause appears to be the addition of H atoms, probably as OH.

Uranium hydride, $UH_{3-\delta}$. Uranium trihydride does not deviate from stoichiometry to any measurable degree at room temperature but does so to a significant degree at high temperatures. For example, at 450°C (840°F) the existence range is $UH_{2.98-3.00}$ and at 800°C (1500°F), $UH_{0.9-3}$. The hydrogen deficiency comes about from hydrogen vacancies. The interaction energy between the vacancies is rather high (4.3 kcal/mole), as there is a great tendency for the vacancies to cluster and cause nucleation of the metal phase as the temperature is lowered.

Tungsten trioxide, $WO_{3-\delta}$. The tungsten-oxygen system shows a variety of materials corresponding to the general formula WO_x, but many of these appear to be discrete compounds, such as $WO_{2.90}$ (which is $W_{20}O_{58}$) and $WO_{2.72}$ (which is $W_{18}O_{49}$), or mixtures of compounds belonging to general series such as W_nO_{3n-2}. Still, it is possible to prepare single crystals of $WO_{3-\delta}$, where δ is continuously variable, provided that δ is very small. One way in which this can be done is to anneal a stoichiometric crystal at, for example, 1100°C (2000°F) for several days in a controlled oxygen pressure. The oxygen pressure can be set by using an appropriate flow ratio of argon and oxygen, or for very low values a mixture of carbon monoxide and carbon dioxide. Although stoichiometric WO_3 is pale yellow, removal of oxygen darkens the crystals first through green, then blue-green, and finally black. The conductivity is *n*-type. Removal of oxygen produces oxygen vacancies, each of which can act as an electron donor. (This can be seen by considering that removal of an oxide ion O^{2-} from WO_3 would disturb electric neutrality, so the two minus charges, or two electrons, would have to be put back into the lattice for each vacancy created.) At low oxygen defect, for example, $WO_{2.999}$, the oxygen vacancies are independent of each other and randomly distributed in the structure. As the defect increases, the oxygen vacancies begin to line up, coalesce, and produce shear planes where the octahedral WO_6 units share edges rather than corners. In $WO_{2.994}$, the average spacing between the crystallographic shear planes, which are along {120}, would be about 28 nanometers. However, they are observed to be quasi-ordered at spacings of 4–5 nm. The oxygen defect in $WO_{3-\delta}$ has been measured with great precision by

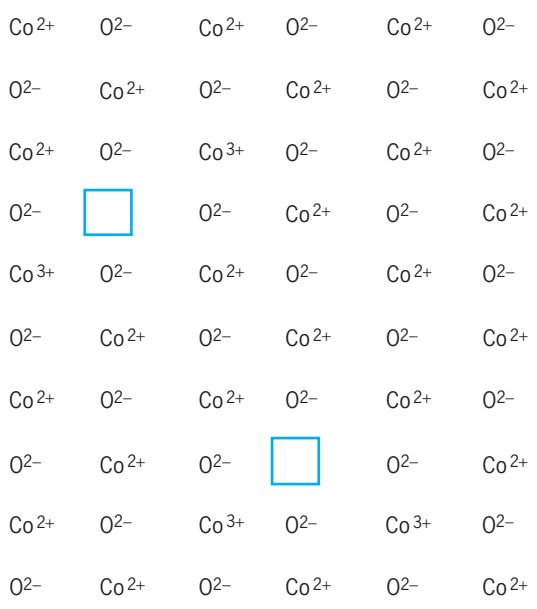

Fig. 5. Cobalt-deficient cobaltous oxide, $Co_{1-\delta}O$. The squares represent missing cobalt atoms. Note the presence of two Co^{3+} to compensate for each Co^{2+} missing.

determining the reducing equivalence of a given sample, as by reaction with $Ag(SCN)_4^{3-}$ to form elemental Ag.

Cobaltous oxide, $Co_{1-\delta}$. Stoichiometric CoO cannot be made. All preparations whether by dehydration of $Co(OH)_2$, decomposition of $CoCO_3$, or controlled oxidation of cobalt wire lead to a product that is deficient in cobalt. As shown in **Fig. 5**, the electrical neutrality is maintained by having the doubly positive charge of each missing Co^{2+} ion taken up by the existence of two Co^{3+} ions someplace in the structure. The resulting material is a good *p*-type conductor in which charge transport occurs by transfer of a "hole" from a Co^{3+} to a neighboring Co^{2+}. Since the conductivity depends on the concentration of Co^{3+}, the electric properties of $Co_{1-\delta}O$ are dependent on the oxygen pressure. [In fact, they have been used as a way of measuring oxygen pressures over the enormous range of 10^{-1} to 10^{-13} atm (or 10^4 to 10^{-8} pascals).] How this comes about can be seen from the following considerations: Oxidation of CoO by oxygen can be viewed as resulting from (1) incorporation of an oxygen atom at a surface site, (2) transfer of an electron from each of two interior Co^{2+} ions to the neutral oxygen to form a normal oxide ion, and (3) migration of a Co^{2+} to the surface and creation of Co^{2+} vacancy in the interior. The net reaction can be written as (1), where □ represents the

$$\tfrac{1}{2}O_2(g) + 2Co^{2+} \rightarrow O^{2-} + 2Co^{3+} + \square \qquad (1)$$

additional cation vacancy created. The condition for equilibrium is shown as Eq. (2) where P_{O_2} represents pressure of the oxygen gas.

$$\frac{[O^{2-}][Co^{3+}]^2[\square]}{P_{O_2}^{1/2}[Co^{2+}]^2} \qquad (2)$$

Since the concentrations of oxide ions $[O^{2-}]$ and of cobaltous ions $[Co^{2+}]$ are practically constant, they can be absorbed in K. The result is Eq. (3). The chemical equation shows that one vacancy is created for every two cobaltic ions formed, so assuming there were negligible vacancies at the start, the concentration of vacancies in the final crystal is just half that of the cobaltic ions, or Eq. (4). Substituting and solving for $[Co^{3+}]$, one gets Eq. (5).

$$\frac{[Co^{3+}]^2[\square]}{P_{O_2}^{1/2}} = K' \qquad (3)$$

$$[\square] = \tfrac{1}{2}[Co^{3+}] \qquad (4)$$

$$[Co^{3+}] = \left(2K'P_{O_2}^{1/2}\right)^{1/3} \text{ or } \sim P_{O_2}^{1/6} \qquad (5)$$

Since the conductivity is directly proportional to the concentration of Co^{3+} ions, it follows that it will show a one-sixth power dependence on oxygen pressure.

If the concentration of vacancies at the start is not negligible but on the contrary intrinsically large, then [□] is also a constant. In such a case, the conductivity would follow a $^1/_4$th power dependence on oxygen pressure. In practice, plots of log conductivity versus log P_O show slopes between $^1/_4$ and $^1/_6$ depending on the biographical defects in the specimen.

Uranium dioxide, $UO_{2+\delta}$. The fluorite structure in which UO_2 crystallizes is capable of great defect concentration; still the nonstoichiometry range accessible $UO_{2.00-2.25}$ is exceptionally broad. Several independent lines of evidence suggest there are two kinds of defects in this range—oxygen interstitials in the first part, UO_{2+x}, and oxygen vacancies in the second part, U_4O_{9-y}. Precision lattice-constant determination gives two linear segments for the lattice parameter a, as illustrated by relationships (6).

$$\begin{aligned} a &= 5.4705 - 0.094x \text{ for } 0 < x < 0.125 \\ a &= 5.4423 + 0.029y \text{ for } 0 < y < 0.31 \end{aligned} \qquad (6)$$

Similarly, a plot of the partial molal free energy of oxygen in $UO_{2+\delta}$ as a function of δ is best fitted by two straight lines intersecting at $UO_{2.125}$. Finally, a determination of the entropy change ΔS for reaction (7) shows $-\Delta S$ tending to very high values

$$UO_{2+x} + \tfrac{1}{2}(0.25 - 0.25y - x)O_2 = \tfrac{1}{4}U_4O_{9-y} \qquad (7)$$

[about 200 entropy units (eu) per mole of O_2 reacting] as the composition of the product approaches U_4O_9. Normally, the entropy loss accompanying the fixing of 1 mole of O_2 is only about 35 eu; therefore it is believed that considerable ordering of the interstitial oxygens must be occurring on conversion from UO_{2+x} to U_4O_{9-y}. X-ray and neutron diffraction studies support such a model.

Titanium monoxide, $TiO_{1\pm\delta}$. This material is interesting because of the wide range of composition possible, from $TiO_{0.85}$ to $TiO_{1.18}$, and because of the metallic character observed. The structure is cubic over the whole range and is frequently described as rock salt,

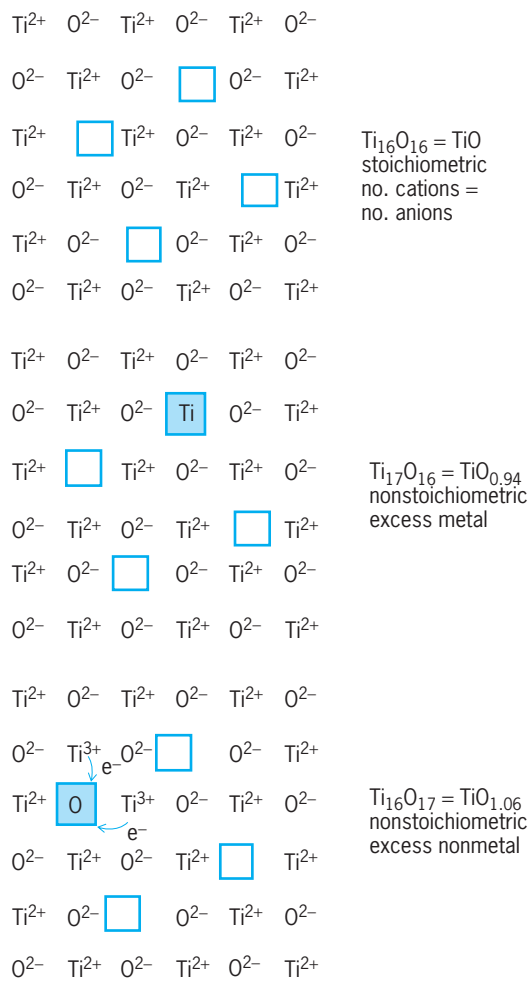

Fig. 6. How nonstoichiometry arises in $TiO_{1\pm\delta}$. The squares represent vacant lattice sites, and the tint indicates added atoms.

though ordering of the cation and anion vacancies makes it more like spinel. **Figure 6** shows schematically how the Ti/O ratio can deviate both above and below unity because of vacancy imbalance. Stoichiometric TiO has no imbalance but an unusually large number of vacant lattice sites, amounting to about 15%. This can be determined by noting that the observed density is less than that calculated on the basis of measured x-ray spacings. If TiO is heated in various oxygen pressures, either above or below that corresponding to equilibrium with the stoichiometric material, excess oxygen or excess titanium atoms can be incorporated at the vacant sites. If neutral oxygen is added, the dinegative oxide-ion charge is obtained by converting two Ti^{2+} ions to Ti^{3+} ions; if titanium is added, the Ti^0 moves to a cation site and forms Ti^{2+} by donating two electrons to the metallic orbitals. The metallic behavior of $TiO_{1\pm\delta}$ is attributable to electron delocalization resulting from overlap of the *d* electrons.

Tungsten bronzes, $M_\delta WO_3$ and analogs. The tungsten bronzes are a curious set of compounds, in which alkali metals, alkaline earth metals, copper, silver, thallium, lead, thorium, uranium, the rare-earth elements, hydrogen, or ammonium can be inserted in a WO_3 structure. In the case of sodium the materials are semiconducting for $\delta < 0.25$ but metallic for $\delta > 0.25$. Colors range from blue to violet to coppery to yellow-gold as δ changes from 0.4 to 0.98. The materials are remarkably inert to most reagents at room temperature. They can be prepared by electrolysis of molten Na_2WO_4–WO_3 mixes; heating of Na_2WO_4, WO_3, and W; and vapor-phase reaction of Na with WO_3.

The cubic sodium tungsten bronzes, $Na_\delta WO_3$ ($\delta > 0.43$), have a simple structure in which a cube containing tungsten atoms at the corners and oxygen atoms at the edge centers is random-statistically occupied by a sodium atom in the cube center. The sodium atom transfers its electron to the conduction band of the WO_3 host structure, a conduction band which may be formed by overlap of the $5dt_{2g}$ orbitals of the tungsten atoms or by interaction of these orbitals with the oxygen orbitals. Rubidium and potassium form hexagonal tungsten bronzes which become superconducting at temperatures in the range 2–6 K, depending on composition.

Vanadium bronzes, $M_\delta V_2O_5$, are analogous compounds based on insertion of metals in V_2O_5. However, unlike the tungsten bronzes, which are generally metallic and show small temperature-independent paramagnetism, they are semiconductors and normally paramagnetic.

Molybdenum bronzes, $M_\delta MoO_3$, are intermediate in properties. For example, $K_{0.30}MoO_3$ has a low temperature-independent magnetic susceptibility, semiconductive behavior below 180 K, and metallic behavior above 180 K. Somewhat startling in this material is the finding that *n*-type behavior in the semiconducting range changes to *p*-type conductivity in the metallic range, opposite to what is generally observed for semiconductors as temperature rises. Titanium bronzes, $Na_\delta TiO_2$, and platinum bronzes, $Na_\delta Pt_3O_4$, have also been reported.

Intercalation compounds. The intercalation compounds include (*a*) the clathrates, where guest molecules occupy isolated cavities of the host (for example, gas hydrates such as $Cl_2 \cdot xH_2O$, $5.75 < x < 7.66$); (*b*) tunnel compounds, where molecules such as hydrocarbons fit into tunnels of the host structure (for example, urea); (*c*) layer compounds, where molecules such as ammonia or alkali metal atoms are inserted between the layers of a transition-metal dichalcogenide; and (*d*) the zeolites, or molecular sieves, where guest molecules move through a three-dimensional network of tunnels. In all these cases, saturation of site occupancy would lead to stoichiometric products; however, the ratio of guest to host is generally less than saturation and is variable, so the materials can be regarded as nonstoichiometric compounds. The potential number of combinations is enormous. More than 100 intercalation hosts are already recognized, and just one of these, graphite, can accommodate more than 12,000 types of guest.

Typical of the inorganic intercalates is $Li_\delta TiS_2$ ($0 < \delta < 1$). The host, The host, TiS_2, has a layered structure in which a sheet of titanium atoms is sandwiched

between two sheets of hexagonally close-packed sulfur atoms. Adjacent sulfur sheets are weakly bonded to each other by van der Waals forces, and guest species can move into this van der Waals gap. Intercalation with lithium can be achieved by exposing TiS_2 to lithium vapor, to lithium dissolved in liquid ammonia, or to *n*-butyl lithium dissolved in a nonpolar solvent. It can also be achieved by electrointercalation, using a cell in which lithium metal is the anode, TiS_2 is the cathode, and the electrolyte is a lithium salt dissolved in an organic solvent such a propylene carbonate. Intercalation proceeds with an expansion of the lattice *c* parameter by about 0.07–0.085 nm. The expansion increases with lithium content. Nuclear magnetic resonance studies of $Li_\delta TiS_2$ indicate that the lithium is quite mobile and is almost 100% ionic. The use of $Li_\delta TiS_2$ in lithium batteries is possible. *See* INTERCALATION COMPOUNDS.

Michell J. Sienko

Bibliography. C. Catlow and W. Mackrodt (eds.), *Nonstoichiometric Compounds*, 1987; L. Eyring and M. O'Keeffe (eds.), *The Chemistry of Extended Defects in Non-Metal-Imperfect Crystals*, vols. 1–3, 1974–1975; P. Kofstad, *Nonstoichiometry, Diffusion and Electrical Conductivity in Binary Metal Oxides*, 1972, reprint 1983.

Noradrenergic system

A neuronal system that is responsible for the synthesis, storage, and release of the neurotransmitter norepinephrine. Norepinephrine, also known as noradrenalin, consists of a single amine group and a catechol nucleus (a benzene ring with two hydroxyl groups) and is therefore referred to as a monoamine or catecholamine. It exists in both the central and peripheral nervous systems. In the autonomic system in the periphery, norepinephrine is the primary neurotransmitter released by the sympathetic nervous system. The sympathetic nervous system mediates the "fight or flight" reaction, preparing the body for action by affecting cardiovascular function, gastrointestinal motility and secretion, bronchiole dilation, glucose metabolism, and so on. Within the central nervous system, norepinephrine has been associated with several brain functions, including sleep, memory, learning, and emotions. The major clustering of norepinephrine-producing neuron cell bodies in the central nervous system is in the locus coeruleus. This center, located in the pons with extensive projections throughout the brain, produces more than 70% of all brain norepinephrine.

Norepinephrine synthesis, storage, release. Norepinephrine is synthesized through a succession of steps. The amino acid tyrosine is first transported from blood into noradrenergic neurons, where it is transformed to dopa. Dopa is synthesized to dopamine, which is also a catecholamine neurotransmitter. In dopaminergic cells the pathway stops at this point, whereas in noradrenergic neurons the dopamine is converted to norepinephrine. In adrenergic cells the pathway continues for one more step to synthesize a third catecholamine, epinephrine (also known as adrenaline). Synthesis of all of the catecholamines is limited by tyrosine hydroxylase activity that converts tyrosine to dopa. Although at physiological concentrations of tyrosine the enzyme is saturated, the activity of the enzyme is limited by the availability of a cofactor, biopterin. Due to its rate-limiting nature in this pathway, tyrosine hydroxylase is a good candidate for regulation by end-product concentrations and by phosphorylation of the enzyme. After synthesis, the majority of norepinephrine is transported into synaptic vesicles in the nerve terminals, where it remains until needed. When the nerve terminal is activated by depolarization, calcium flows into the nerve terminal, leading to the docking of the vesicles onto the presynaptic membrane and release of norepinephrine into the synaptic cleft. *See* EPINEPHRINE.

Receptors and regulation. Once released into the synaptic cleft, norepinephrine is free to bind to specific receptors located on the presynaptic or postsynaptic terminal, which mediate various effects on either the presynaptic or postsynaptic cells. Binding of norepinephrine to a receptor initiates a chain of events (the effector system) in the target cell that can be mediated by a number of different second messenger systems. The exact effect is determined by the identity of the receptor activated. All noradrenergic receptors are linked to G proteins (see **illus.**).

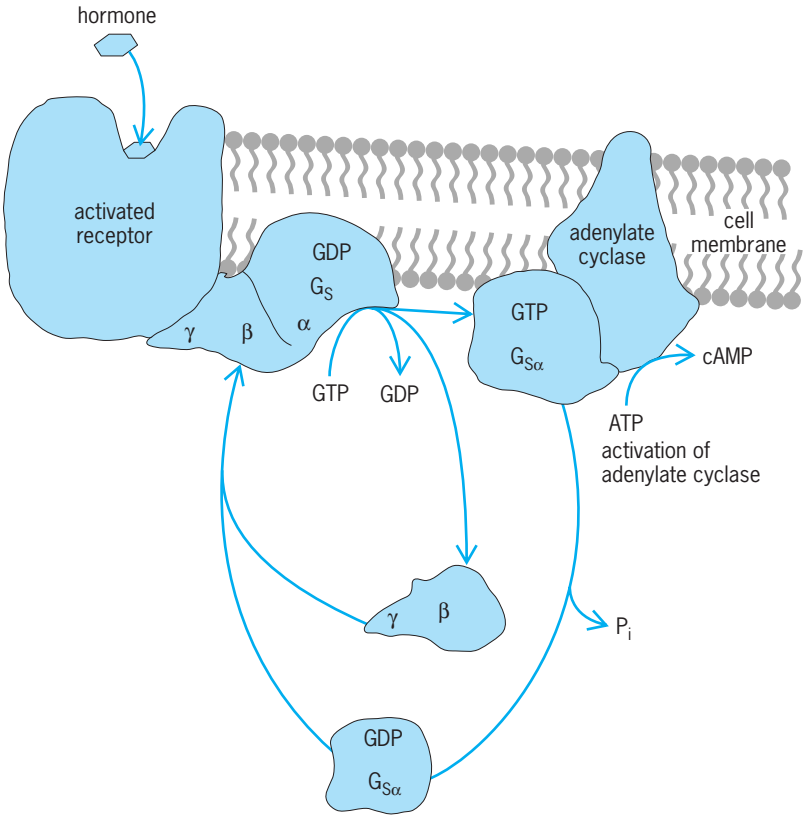

Regulatory cycle of G-protein-mediated signal transduction. Hormone binds to the receptor, triggering the exchange of GDP for GTP. The α subunit–GTP dissociates from the $\beta\gamma$ subunit. The GTP-liganded α subunit activates several processes (adenylate cyclase activation); $\beta\gamma$ subunits regulate shared or distinct effectors. (*After D. Nelson and M. Cox, Lehninger Principles of Biochemistry, 3d ed., 2000*)

Binding of norepinephrine to a receptor leads to activation of the G protein, which in turn modulates the activity of another protein (enzyme or ion channel) in the cell. Noradrenergic receptors can be broken down into three categories: $\alpha 1$, $\alpha 2$, and β receptors. These alpha and beta categories are constantly being subdivided with the explosion in molecular biology, allowing investigators to identify small differences in receptor structure that may have little impact on the physiological effect of the receptor. For example, there are currently four different $\alpha 1$ receptors: αlA, lB, lC, and lD. Depending on the subtype of $\alpha 1$ receptor activated, the ultimate effect may be a modulation of a calcium channel or activation of phospholipase C (the enzyme which converts phospholipids into inositol trisphosphate and diacyl glycerol). All $\alpha 2$ receptors are linked to the inhibition of adenylate cyclase and thus reductions in cyclic 3′,5′-adenosine monophosphate (cAMP). Included in the category of $\alpha 2$ receptors are presynaptic autoreceptors that regulate further release of norepinephrine. This negative-feedback mechanism acts as a brake on the noradrenergic neuron. When the autoreceptor is stimulated, it decreases noradrenergic firing; when it is blocked, it increases noradrenergic firing. There are three main β receptors that have been identified, all linked to adenylate cyclase stimulation and thus leading to increases in cAMP.

Noradrenergic receptors are highly regulated. Both homologous desensitization due to prolonged activation of specific receptors and heterologous desensitization due to prolonged activation of other receptor subtypes (other noradrenergic receptors or other neurotransmitter receptors) occur. Phosphorylation, of the β receptors in particular, has been well documented as a mechanism of regulating the receptors' activity.

Inactivation. Termination of norepinephrine occurs by a high-affinity reuptake mechanism in the presynaptic membrane. The active transport of norepinephrine is fueled by the sodium gradient across the cell membrane. Once transported back into the presynaptic terminal, norepinephrine can be stored in vesicles for future use or enzymatically degraded by monoamine oxidase. Any stray norepinephrine that is not taken up into the terminal is degraded by catechol-*O*-methyl transferase.

Certain medications achieve their effect by altering various stages of synthesis, storage, release, and inactivation of norepinephrine. The behavioral manifestations of these alterations have led to a better understanding of norepinephrine's role in various psychiatric disorders. It has been suggested that altered noradrenergic function may be of etiological significance in depression, posttraumatic stress disorder, anxiety disorder, and schizophrenia. Interpretation of the pharmacological data has been complicated by the recent studies on neurogenesis. It has been shown that neurogenesis occurs in specific brain areas and that this process is modulated by environmental conditions such as stress or therapeutic agents used in psychiatric disorders such as modulators of serotonergic activity or noradrenergic activity, and electroconvulsive shock therapy. Thus, the etiology underlying these disorders may be related to a lack or excess of neurogenesis rather than defects in a specific neurotransmitter system.

Depression. The monoamine deficiency hypothesis postulates that depressive illness arises from a deficiency of synaptic norepinephrine. This hypothesis is supported by the fact that catecholamine-depleting agents, such as reserpine, cause depression, while agents that increase synaptic norepinephrine, such as tricyclics and monoamine oxidase inhibitors, are effective treatments for depression. However, advances in neuropharmacology have challenged the monoamine deficiency hypothesis on grounds that a more complex model is necessary to explain depression and its treatment. First, certain drugs, including cocaine and amphetamines, that inhibit reuptake of norepinephrine are not effective in treating depression. Second, several "atypical" antidepressants are effective for treating depression but do not block uptake of norepinephrine or affect its breakdown through monoamine oxidase. Third, the inhibitory effects of tricyclic antidepressants and monoamine oxidase inhibitors occur within minutes after drug administration, but the antidepressant effects generally do not occur for 2–3 weeks. *See* MONOAMINE OXIDASE.

Because of the temporal discrepancy between initiation of antidepressants and the occurrence of therapeutic effects, research has turned to the examination of the effects of chronic antidepressant treatment on receptor function. A number of investigators have seen both in vivo and in vitro that chronic treatment with antidepressants and electroconvulsive shock therapy (an alternative treatment for depression) lead to a decrease in β receptors. This correlates well with an increase in β receptors which has been documented in suicide victims with a history of depression. The α receptors are also decreased with chronic antidepressant treatment and show supersensitivity in suicide victims. These changes in receptors have not been documented by every study. The changes in the noradrenergic system with antidepressant treatment suggest that manipulations of the system can alleviate depression but do not prove that alterations in the system are responsible for the depression in the first place. It is also possible that the stimulatory effects of norepinephrine on neurogenesis may in part explain the temporal discrepancy between onset of therapy and alleviation of symptoms. During neurogenesis it would take weeks to produce a fully functional differentiated neuron. The different neurotransmitter systems in the brain interact in complex ways, making a definitive cause difficult to elucidate. *See* AFFECTIVE DISORDERS.

Anxiety disorders. Anxiety disorders are broken down into at least four categories: panic disorder, posttraumatic stress disorder, generalized anxiety disorder, and obsessive-compulsive disorder. The

extent of involvement of the noradrenergic system with each disorder varies. There is no indication that obsessive-compulsive disorder includes changes in the noradrenergic system. Panic disorder and generalized anxiety disorder are the most similar in their neuroendocrine involvement and response to noradrenergic manipulation. As with most psychological disturbances, the changes in the central nervous system are not restricted to changes with a single chemical system or anatomical region. The hypothalamic-pituitary-somatotropin axis appears to have altered regulation, and it responds abnormally to $\alpha 2$ receptor activation. Patients with panic disorder more frequently show an anxiety reaction in response to $\alpha 2$ antagonist administration than patients with other psychological abnormalities. Agents known to decrease noradrenergic function such as tricyclic antidepressants and monoamine oxidase inhibitors are often effective treatments for patients with panic disorder or generalized anxiety disorder.

Involvement of the noradrenergic system has also been implicated in the pathophysiology of posttraumatic stress disorder. The immediate response to severe stress includes the activation of the sympathetic nervous system as well as an increase in the firing rate of the locus coeruleus neurons, both of which lead to the release of norepinephrine. These responses are adaptive and protective, allowing a person to cope with the stress. These responses become maladaptive when they are chronically activated due to depletion of norepinephrine stores, changes in noradrenergic receptor numbers, and chronic exposure of brain areas to glucocorticoids. Although posttraumatic stress disorder can occur with just a single exposure to an event, it is assumed that the stress reaction is repeated due to reliving of the trauma, which leads to an exacerbation of the symptoms. Traumatized veterans, when reexposed to combat-associated stimuli, respond with abnormal elevations of autonomic parameters such as blood pressure and heart rate. Elevated levels of norepinephrine in plasma and urine have been reported in severely stressed individuals and in patients with posttraumatic stress disorder. At the same time, platelet $\alpha 2$ adrenergic receptor numbers, lymphocyte adenylate cyclase activity, and monoamine oxidase levels have been reported decreased. Consistent with these abnormalities in norepinephrine metabolism is the suggestion that several antihypertensive medications, which decrease the effects of norepinephrine, may diminish the symptoms of arousal-related posttraumatic stress disorder. In addition to dysregulation of the noradrenergic system, patients with posttraumatic stress disorder may also have anomalies in their corticol secretion and dopaminergic systems. *See* STRESS (PSYCHOLOGY).

Schizophrenia. The action of antipsychotic agents on the dopaminergic system has led to many investigations into the role of dopamine in schizophrenia. In 1989 a new type of antipsychotic agent, clozapine, was approved for use in schizophrenia. Clozapine demonstrated superior efficacy as an antipsychotic agent when compared to traditional neuroleptics, without many of the extrapyramidal side effects. It was shown to increase plasma levels of norepinephrine levels fivefold, as well as increasing the number of spontaneously active cells in the locus coeruleus and increasing the firing rate of locus coeruleus neurons. The $\alpha 1$ and $\alpha 2$ antagonist activity of clozapine is insufficient to explain the changes in plasma norepinephrine levels. Evidence from a number of studies suggests the involvement of multiple neurotransmitters in this disease. It has been hypothesized that the effect of increasing norepinephrine alleviates schizophrenia by enhancing synaptic strength in cortical regions responsible for the psychotic symptoms. Thus, changes in the noradrenergic system may not be directly responsible for schizophrenia but may modulate other neurotransmitters that are affected in this disease. In contradiction to the studies with clozapine that demonstrate a correlation between increased plasma norepinephrine levels and alleviation of symptoms, other studies have found elevated levels of norepinephrine in cerebrospinal fluid of schizophrenics, both living and deceased. In addition, medications such as cocaine and amphetamines that are known to increase norepinephrine levels through uptake inhibition can cause psychosis. Receptor studies have shown no changes in α receptors, with a possible decrease in β receptors, in the hippocampus. *See* NEUROBIOLOGY; PSYCHOPHARMACOLOGY; SCHIZOPHRENIA.

Michelle Mynlieff; Dennis S. Charney; Alan Breier; Steven Southwick

Bibliography. D. Goldstein, G. Eisenhofer, and R. McCarty (eds.), *Catecholamines Bridging Basic Science with Clinical Medicine*, vol. 42 in *Advances in Pharmacology*, Academic Press, San Diego, 1998; J. G. Hardman et al. (eds.), *Goodman and Gilman's The Pharmacological Basis of Therapeutics*, McGraw-Hill, New York, 1996; G. J. Siegel et al. (eds.), *Basic Neurochemistry Molecular, Cellular and Medical Aspects*, 6th ed., Lippincott-Raven, Philadelphia, 1999.

Normal (mathematics)

A term generically synonymous with perpendicular, which often refers specifically to a line that goes through a point P of a curve C and is perpendicular to the tangent to C at P. If a plane curve C has equation $y = f(x)$, in rectangular coordinates the normal (line) to C at $P(x_0, y_0)$ has slope $-1/f'(x_0)$, provided $f'(x_0) \neq 0$. The expression $f'(x_0)$ denotes the derivative of $f(x)$, evaluated for $x = x_0$, and so the normal has equation $y - y_0 = [-1/f'(x_0)]\,(x - x_0)$. If curve C is not a plane curve, all normal lines of C at point P on C lie in a plane, the normal plane of C at P. For other uses of normal (for example, normal form of equation of a line or a plane) *see* ANALYTIC GEOMETRY.

Leonard M. Blumenthal

North America

The third largest continent, extending from the narrow isthmus of Central America to the Arctic Archipelago. The physical environments of North America, like the rest of the world, are a reflection of specific combinations of the natural factors such as climate, vegetation, soils, and landforms. While the distribution of life-giving solar heat and water seemingly should determine the character and distribution of the natural regions, in reality the landforms modify the expected environmental patterns. In addition, human activities affect the quality of some environments, and natural hazards alter some regions permanently or semipermanently. *See* CONTINENT.

In order to understand the characteristics and distributional patterns of the natural regions of North America, it is necessary to consider the continent's absolute and relative location, and the types and patterns of its climates, vegetation, soils, and landforms.

Location

North America covers 9,400,000 mi^2 (24,440,000 km^2) and extends north to south for 5000 mi (8000 km) from Central America to the Arctic. It is bounded by the Pacific Ocean on the west and the Atlantic Ocean on the east. The Gulf of Mexico is a source of moist tropical air, and the frozen Arctic Ocean is a source of polar air. With the major mountain ranges stretching north-south, North America is the only continent providing for direct contact of these polar and tropical air masses, leading to frequent climatically induced natural hazards such as violent spring tornadoes, extreme droughts, subcontinental floods, and winter blizzards, which are seldom found on other continents. *See* AIR MASS; ARCTIC OCEAN; ATLANTIC OCEAN; GULF OF MEXICO; PACIFIC OCEAN.

Environment

Several environmental factors contribute to the characteristics of the generally recognized natural regions of North America. These include geologic structure, climatic types and regions, vegetation regions, and soil regions.

Geologic structure. The North American continent includes (1) a continuous, broad, north-south-trending western cordilleran belt stretching along the entire Pacific coast; (2) a northeast-southwest-trending belt of low Appalachian Mountains paralleling the Atlantic coast; (3) an extensive rolling region of old eroded crystalline rocks in the north-central and northeastern part of the continent called the Canadian Shield; (4) a large, level interior lowland covered by thick sedimentary rocks and extending from the Arctic Ocean to the Gulf of Mexico; and (5) a narrow coastal plain along the Atlantic Ocean and the Gulf of Mexico. These broad structural geologic regions provide the framework for the natural regions of this continent and affect the location and nature of landform, climatic, vegetation, and soil regions.

Climatic types. The North American continent, while dominated by broad areas of continental climates, has regions representing every climate on Earth. Thus, in combination with its location and terrain, the continent is subject to almost every climatic hazard known: intense tornadoes in the southern Great Plains and the Midwest; intense tropical hurricanes in the southeast and in Central America; extreme winter cold and blizzards in the north; droughts of the intensity of the Dust Bowl; and floods.

The North American climatic regions (**Fig. 1**) result from the combined effect of numerous controlling factors that create the specific climates characterizing different parts of the continent. The most important controls of the North American climates are (1) latitude, which controls the distribution of solar energy throughout the year; (2) location of semipermanent high-pressure cells (over the oceans around 30°N latitude); (3) the large size of the continent, which creates continentality in the interior; (4) polar and tropical air masses; (5) wind belts, especially the westerly winds with the cyclonic storms; and (6) mountain barriers. The two air masses that dominate the central part of the continent are the very cold and dry Continental Polar (cP) air mass, whose source region is the Arctic and north central Canada, and the warm and wet Maritime Tropical (mT) air mass, whose source region is the tropical and subtropical ocean, including the Caribbean region. The meeting of these two air masses creates most of the violent weather in the central part of the continent. The other air masses affecting North America are the hot and dry Continental Tropical (cT) in the southwest, and two cool and wet Maritime Polar (mP) air masses originating over the oceans northwest and northeast of the continent. Their effect is more local.

The resulting climates dominating large parts of the continent are polar and high-latitude humid continental (snow) climate, including tundra, subarctic, and humid continental; midlatitude temperate, moist; midlatitude dry-summer (Mediterranean); semiarid and arid (desert); and tropical.

Polar and High-latitude humid continental. Also known as snow climates, these include tundra, subarctic and humid continental climates.

The tundra climate has an average monthly mean temperature below 50°F (10°C), and trees do not grow. Precipitation is minimal.

The subarctic climate is located between 50 and 70°N latitude and extends from western Alaska across Canada to the east coast of North America. Its winter is dark, long, and bitterly cold, with averages of −35°F (−38°C). Lakes are frozen from September to May. Summer is very short, but because of long days the temperature increases rapidly, reaching the 50s or low 60s on the Fahrenheit scale (10 or 15°C). Precipitation is meager [5–20 in. (13–50 cm)].

The humid continental climate is located between 35 and 55°N latitude, from the east coast toward the interior of the continent. Dominated by the westerly wind belt and its migratory pressure systems, the weather changes frequently in this climate, both from day to day and from season to season. Daily variability is the major characteristic of this climate

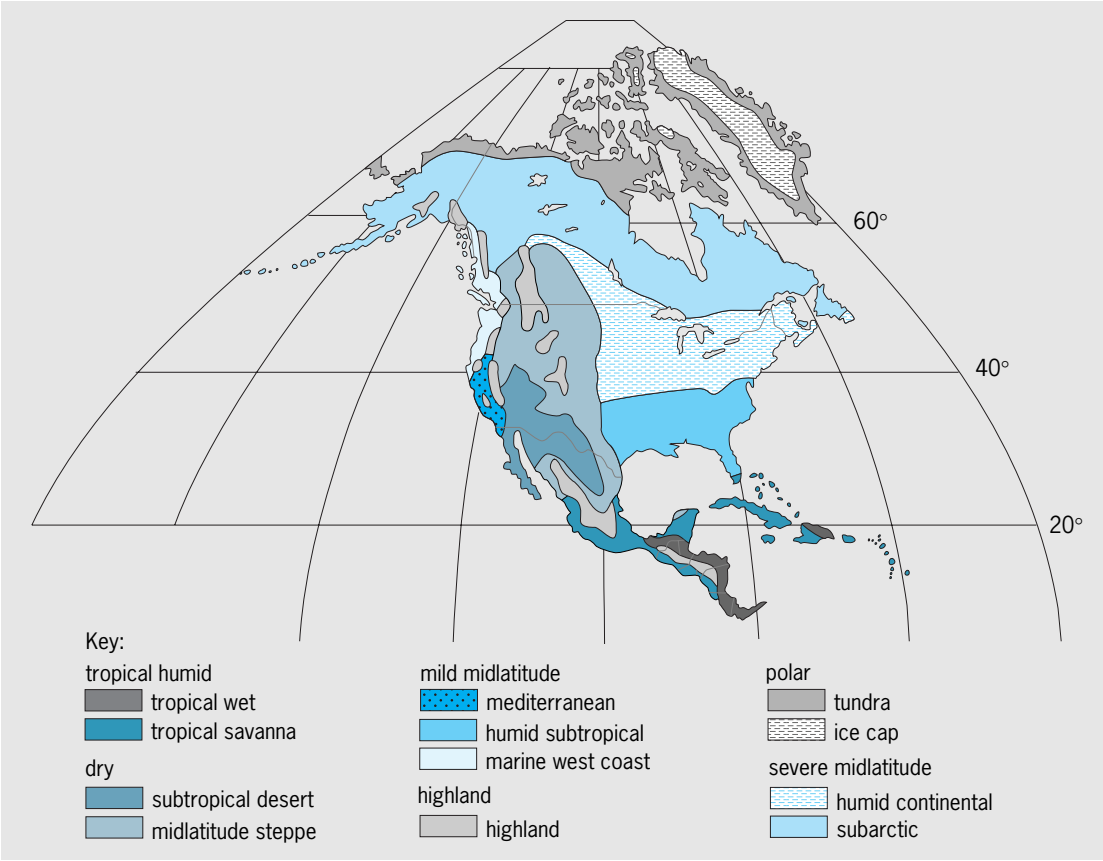

Fig. 1. Climatic regions of North America (modified Köppen system). (*After T. L. McKnight, Physical Geography: A Landscape Appreciation, 3d ed., Prentice Hall, 1990*)

in all seasons, including hazardous winter blizzards, thunderstorms, tornadoes, heat waves and prolonged drought, and great floods.

The summer temperature averages 70°F (20°C) but reaches the upper 80s on the Fahrenheit scale (upper 20s on Celsius) on many days. The average winter temperature is around 15°F (−9°C), and there are about five months with averages below freezing. This is a humid climate with up to 50 in. (125 cm) of monthly rainfall, which diminishes toward the interior. There is more precipitation in summer than in winter. Snow stays on the ground for less than a month in the south and up to five months in the north. *See* DROUGHT; PRECIPITATION (METEOROLOGY); THUNDERSTORM; TORNADO.

Mid-latitude. The temperate climate known as humid subtropical lies between 25 and 40°N latitude and occupies the southeastern part of the continent. The summers are warm to hot—monthly averages reach 85°F (27°C)—with high humidity day and night. The winters are mild, but invasion of cold polar air sometimes brings severe weather for a few days, and killing frost can occur in late spring. Annual precipitation is between 40 and 60 in. (100 and 150 cm), but occasional tropical hurricanes bring very heavy precipitation, high wind, and wave destruction to coastal areas. *See* FROST; HURRICANE; TROPICAL METEOROLOGY.

Mediterranean climate. The regional extent of this climate in North America is small and coastal. This dry subtropical climate occurs on the western side of the continent between 30 and 40°N latitude. The climate is characterized by a modest amount of precipitation, about 15 in. (38 cm), concentrated in a mild (50°F or 10°C) winter and a very hot (75–85°F or 24–29°C) and dry summer. Abundant sunshine is typical for this climate. Winter rains come from the frontal cyclonic storms of the westerly winds, and summer drought from the effect of the Subtropical High Pressure cell located just offshore.

Marine West Coast climate. The marine west coast climate lies north of the Mediterranean climate between 40 and 60°N latitude. It is exceptionally mild for its latitude due to its coastal location and the effect of the oceanic westerly winds. Its average summer temperature is between 60 and 70°F (16 and 21°C) and the average cold-month temperature is between 35 and 45°F (2 and 7°C). This is one of the wettest climates in the middle latitudes, with average rainfall of 30 to 50 in. (75–125 cm) but reaching 200 in. (500 cm) on windward slopes.

Dry climate. The semiarid and arid climates are characterized by greater evaporation than precipitation. The semiarid steppe climates have sufficient moisture to support short-grass vegetation. The desert climates are so dry that they support only xerophytic vegetation that can withstand extreme heat. Because of extreme drought, dry climates may experience large diurnal and seasonal temperature differences.

Tropical climates. The tropical climates include the tropical wet climate, stretching 5 to 10° latitude on both sides of the equator. It is hot all year, with an average monthly temperature over 64°F (18°C), and wet all year, with 60 to 100 in. (150 to 250 cm) of annual rainfall. The tropical wet and dry (savanna) climate stretches up to 25° latitude and has a distinct wet summer and very dry winter pattern. This type of climate characterizes the southern tip of Florida. Tropical wet, and tropical wet and dry, climates are present in the coastal parts of Mexico and Central America.

Climatic regions. In the northern part of the North American continent, the major climatic control is high latitude, affecting the distribution of solar radiation throughout the year. Since in the summer the days are long and the angle of the Sun is reasonably high, solar heat is accumulated. During the rest of the year, and especially in the winter, the low angle of the Sun and short days significantly limit the amount of heat received. As a result, there is a broad northern belt that has four-season humid continental climates with cold and snowy winters and warm to cool summers; the severity of the winters increases northward (for example, north central Canada) and decreases southward. Cyclonic storms bring most of the precipitation. North of these climates stretches the tundra climate, where the average monthly temperatures never rise above 50°F (10°C), snow lies on the ground year-round, and subsoil is permanently frozen (permafrost). *See* INSOLATION; PERMAFROST.

In the southeastern part of the North American continent, there is no snow cover; the climate is temperate, of humid subtropical type with moist and warm to hot summers and short, cool winters. Here the winds from the Atlantic High Pressure cell sweep heat and moisture from the subtropical oceans onto the land, providing for a warm and moist temperate climate with much summer rain.

Along the west coast of the continent, between 35 and 45°N latitude the Mediterranean climate is very hot and dry in the summer because of the descending and drying air of the Pacific High Pressure cell located off the coast. Winters are much cooler and rainy because the cyclonic storms of the westerly winds move southward to replace the High Pressure cell. Northward of the mediterranean climate, the coastal belt has the Marine West Coast climate, which has moderate temperatures and copious precipitation all year long, with total annual rainfall reaching 200 in. (500 cm). The combined effect of coastal and windward locations, and the mountain barriers along the Pacific coast, give the Marine Coast climate the highest precipitation in North America outside the tropics.

The remaining part of the continent, located mainly in the west, has a complex mosaic of semidry steppe, desert, and mountain climates. The mountain barriers, including the Coastal Ranges and Sierra Nevada, have deserts on their lee (east) sides, where the descending air becomes very drying. The orographic effect of the Rocky Mountains and the continental location create a broad north-south belt of dry steppes in the Great Plains. The descending air of the Pacific High Pressure cell creates the driest deserts of North America in the southwestern part of the continent and parts of northern Mexico. *See* DESERT; MOUNTAIN METEOROLOGY; PLAINS.

Tropical climatic zones. The rest of Mexico and Central America have tropical and mountain climates. Since this area lies in tropical latitudes, the diurnal temperature range is small throughout the year. However, since temperature decreases with altitude, there is an altitudinal climatic zonation on high mountain slopes. These zones are called tierra caliente (hot land), tierra templada (temperate land), tierra fria (cold land), and tierra helada (frozen land). Here, weather, climate, vegetation, soils, and land use vary according to elevation.

The tierra caliente extends from sea level to about 2500 ft (750 m) and includes coastal lowlands, plains like the Yucatán Peninsula, and interior basins, like Balsas River valley, covering about half of the tropics. Here, daytime temperatures are about 85–90°F (30–32°C) and nights are about 70–75°F (21–23°C).

The tierra templada lies roughly between 2500 and 5500 ft (750 and 1660 m). This zone includes the intermediate mountain slopes, the Plateau of Central Mexico, and Central America. Mild daytime temperatures of 75–80°F (23–26°C) increase to 85–90°F (30–32°C) in April and May, while nights have pleasant temperatures of about 60–70°F (15–21°C). Northward, the range between summer and winter temperatures increases, and night frost is possible. This zone has much of the population and agricultural production of the tropics.

The tierra fria lies above 5500 ft (1660 m) elevation and represents only about 10% of Middle America. It is present in high basins and on mountain slopes of Mesa Central of Mexico, and Central America. Its day temperatures are about 75–80°F (23–26°C), but night temperatures fall to 50–55°F (10–12°C). Frost is common in cooler months.

The tierra helada lies above 12,000 ft (3630 m). It occupies the highest mountain peaks, has temperatures below 50°F (10°C) all year, and has permanent ice and snow.

In terms of rainfall, northern Mexico, with its dry deserts and steppes, contrasts sharply with its southern part and Central America, which are humid and rainy. Generally, the east coasts have more rain than the west coasts because of the dominance of the easterly trade winds. The hot months (May to October) are more rainy, often receiving 75–80% of annual rainfall. These windward coasts receive between 80 and 120 in. (200 and 300 cm) of rainfall annually; the lee sides of the mountains receive 20–40 in. (50–100 cm) of precipitation. Most of the rain comes from daily thunderstorms, midlatitude cyclones, and tropical hurricanes, especially those from the Atlantic Ocean. Some of the hurricanes are true weather hazards, since they are extremely destructive. *See* CLIMATOLOGY.

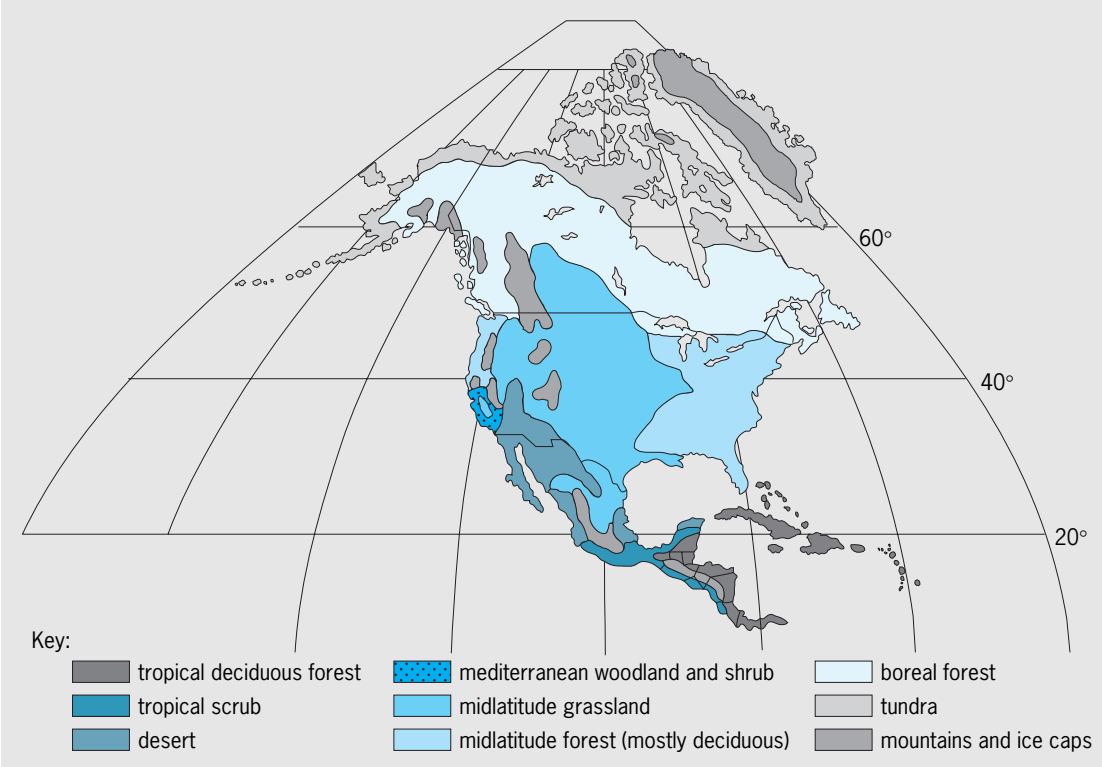

Fig. 2. Natural vegetation regions of North America. (*After T. L. McKnight, Physical Geography: A Landscape Appreciation, 3d ed., Prentice Hall, 1990*)

Vegetation regions. A major component of the natural environment directly related to climatic conditions, and of vital importance to humans as an indicator of the potential for food growing, is the natural vegetation. Though vegetation in midlatitudes has often been cleared to convert the land for agricultural use, its nature must be understood as a prerequisite for proper utilization of the land on which it once grew. Comparison of a climatic map (Fig. 1) with a natural vegetation map (**Fig. 2**) reveals a striking resemblance between the patterns of distribution of these two factors. *See* ALTITUDINAL VEGETATION ZONES; PLANT GEOGRAPHY.

The zone of the tundra climates is covered by ice or by tundra vegetation. Since this area is really a cold desert, with extreme shortage of water (most of it is frozen), and very short and cool summers with mean temperatures below 50°F (10°C), trees cannot grow. Instead, the tundra plant assemblages consist of dwarf shrubs, mosses, lichens, grasses, and flowering herbs, all of which go through their life cycle in the short summer. The tundra supports reindeer and birds (which migrate southward in the winter), and has enormous numbers of insects. *See* TUNDRA.

The northern part of the humid continental snowy climate, where mean winter temperatures rise above 50°F (10°C), supports an evergreen needleleaf forest, which is known as the boreal forest in North America. This great northern forest is one of the most extensive biomes in the world, with the simplest assemblage of plants. Almost all the trees are coniferous, needleleaf evergreens, mostly pines, firs, and spruces, although patches of deciduous trees can be found, such as birch, aspen, and poplar. Along the northern fringe, trees are short and widely spaced, but toward the south they grow taller and are more dense. The undergrowth is generally not dense; mosses and lichens create a ground cover. Drainage is poor in the summer because of frozen subsoil, and bogs and swamps are common. Everything is frozen in the winter. Animal life consists mainly of fur-bearing mammals, and numerous birds in the summer, as well as abundant insects.

The southern zone of the humid continental climate supports the midlatitude deciduous forest, although along its northern boundary the forest is mixed with evergreen conifers. Previously, extensive areas were covered by this forest, but most of it has been cleared for agriculture. This is a dense forest of broadleaf deciduous trees that grow tall but have generally little undergrowth. The animal life consists of birds, mammals, and reptiles.

The southeastern part of the United States, dominated mainly by the humid subtropical climate, supports subtropical broadleaf deciduous as well as evergreen forests, including the southern pine forest. The trees grow tall and have luxurious crowns.

The central and western parts of the United States and southern Canada, which are dominated by the steppe and desert climates, are covered by grassland and desert vegetation. East of the Rocky Mountains, in the Great Plains, short steppe grasses dominate. The tall-grass prairies in the Midwest have virtually disappeared under the plow. The steppes, where trees grow only along the few streams that cross these areas, provide pasture to large herbivores

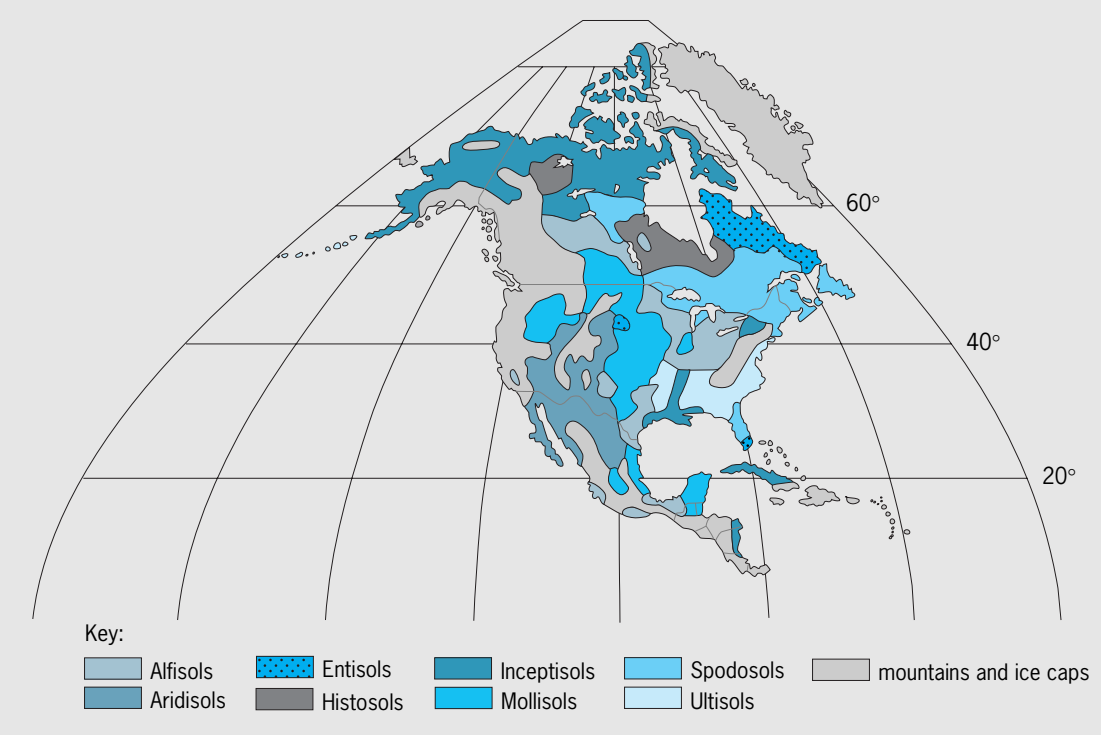

Fig. 3. Soil regions of North America. (*After T. L. McKnight, Physical Geography: A Landscape Appreciation, 3d ed., Prentice Hall, 1990*)

and smaller animals that often live underground. *See* GRASSLAND ECOSYSTEM.

The desert lands of the southwestern part of North America are so dry that only drought-resisting plants thrive. They either have structures, like cacti, that allow them to conserve moisture, or they are smaller plants that go through their entire life cycle during the brief, infrequent rainy periods. The plant cover in the desert is not continuous, and there are large areas of bare ground. The animal life is quite rich, though barely visible since it lives mainly underground. *See* PLANT-WATER RELATIONS.

The Mediterranean woodland and shrub covers a very small part of North America. It coincides with the Mediterranean climate in coastal California. It grows under the hot dry summer and rainy winter conditions, and consists of dense woody shrubs called chaparral, or grassy areas with broadleaf evergreen trees. The animal life is dominated by rodents and reptiles. *See* CHAPARRAL.

Mexico has desert vegetation in Baja California and in the northern part of the Mexican Plateau, with semidry steppe grassland around it, and temperate highland forest, broadleaf and evergreen, on mountain slopes, with tundra on the highest peaks. In Central America, the lowlands and lower mountain slopes have tropical rainforest, with a great number of species of broadleaf evergreens, some of which are 150 ft (45 m) tall. Multiple levels of tree crowns restrict sunlight and undergrowth. Pure stands of trees are rare. A unique type of vegetation in these moist and hot tropical areas of abundant rainfall are the expanses of savanna (tall grasses) found in the lowlands of eastern Nicaragua and Honduras. They have developed on porous gravelly soils that will support only drought-tolerant vegetation. *See* SAVANNA.

Soil regions. The climatic and biotic regimes are so closely related that they generally coincide regionally. However, the soils, while strongly related to both regimes, also depend on the type of rock from which they develop (**Fig. 3**). There are specific factors (rock material, climate, and vegetation) whose interaction results in physical and biological processes that form the soils. To understand the characteristics and the regional distribution of North American soils, consideration must be given to the chemical composition of the rock material, the climatic and biotic environment, and the length of time during which these factors interact. Also, as the time of formation gets longer, the rock factor becomes less important (except in its limiting role as to the type of chemicals initially provided), and the environmental factors become more important. The length of time is especially important in North America, because more than half of the continent was overridden by glaciers, which removed whatever soils were there prior to glaciation. Since the glaciers did not totally disappear from central Canada until about 7000 years ago, most of the present soils in the northern parts of the continent are less than 7000 years old. *See* GLACIAL GEOLOGY.

Inasmuch as the processes that transform the rock materials into soils are determined by the climate (temperature and moisture) and vegetation, they differ from one climatic region to another. These processes are laterization (weathering that produces a high concentration of laterites), podzolization (increase in acid content), gleization (conversion to

clay), calcification (saturation of soil colloids with calcium), and salinization (accumulation of soluble salts). Laterization, podzolization, and gleization occur in moist climates and are based on downward percolation of rainwater. Calcification and salinization are based mainly on upward migration of capillary water due to intense evaporation. During the movement of water downward or upward, soil minerals and organic particles are redistributed throughout the earth material and form horizontal layers called horizons. Soil profiles composed of different horizons identify mature soils. The soil profiles differ from climate to climate since the soil-forming processes differ.

In North America the process of podzolization occurs in high and middle latitudes in humid, cool climates and in regions of coniferous boreal forests, which provide limited acidic needle-type litter. In these cool climates, chemical weathering is slow, but the abundance of acids and adequate precipitation make leaching a very effective process. This process produces podzolic soil profiles that have a leached-out, gray, sandy upper horizon; a brown middle horizon of accumulated minerals, like clay; and a lower horizon consisting of regolith, or broken rock material. These soils are identified mainly as spodosols (Fig. 3). *See* REGOLITH.

In the southeastern part of North America, the warmer humid climate affects the soil-forming processes by including some components of the laterization process, which dominates the hot and moist tropical regions with broadleaf tree and shrub vegetation. This process is characterized by rapid decompositions of organic matter and intense weathering of parent material, leading to dissolution of most minerals and leaving behind mainly iron and aluminum oxides which give it a red color. The soils of part of the Midwest are alfisols, and the southeastern soils are ultisols (Fig. 3), developed under warm to hot, and moist but not tropical conditions. *See* WEATHERING PROCESSES.

A large part of central North America, characterized by lower moisture due to its continentality, has mollisol soils. These soils form in subhumid conditions under dense grass vegetation, which is present in this area. Mollisols form through the process of calcification and have a dark, humus-rich top horizon. In this process, both leaching and removal of minerals are limited by the lack of precipitation. Calcium carbonate is carried downward by limited percolation and is concentrated in a horizon below the surface forming a dense layer; some of it is then carried upward by capillary water during intense evaporation, and by grass roots. In areas of undisturbed grasslands, mollisols are very fertile, but they lack moisture. *See* HUMUS.

The western belt of North America, dominated by mountain and desert climates, has a mosaic of mountain and desert soils. The desert soils form by the salinization process under drought conditions. Since there is less precipitation than evaporation, moisture drawn upward by intense evaporation brings various salts to the surface of the desert, which often becomes toxic. These desert soils continue into northern Mexico.

The mountain soils are highly complex, very thin, and generally immature because steep slopes and low temperatures prevent them from developing profiles and maturing.

The soils of Mexico and Central America are strongly related to the dominant climates and terrain configuration. The North Mexican deserts have thin desert soils formed by the processes of salinization; they are fringed by better steppe soils formed by calcification. The tropical coastal lowlands have red latosol soils formed by the process of laterization. They are greatly leached by excessive precipitation and inherently infertile. In some basins, there are heavy, clay, leached soils, called rendzina, which develop on soft limestones, such as those of northern Guatemala. On the very steep mountain slopes of tropical Central America, soils are very thin and poorly developed. Alluvial soils along the rivers, especially in the plains, are the most fertile. *See* PEDOLOGY; SOIL; SOIL CHEMISTRY.

Landform Regions

In North America the climatic, vegetation, and soil regions exhibit approximate regional coincidence forming natural regions clearly dominated by climatic factors. However, the natural regions are almost equally determined by the landforms of North America (**Fig. 4**).

The major landform subdivisions of the North American continent are the Canadian Shield, Appalachian Highlands, Southeastern Coastal Plain, Interior Provinces including the Great Plains, North American Cordillera, Arctic Margin Plains, and Central America. These regions consist of subregions that differ in rock type or in elevation, slope, relief, pattern, or texture of their terrain. The result is an extremely varied and beautiful natural landscape.

Canadian Shield. Properly referred to as the geological core of the continent, the exposed Canadian Shield extends about 2500 mi (4000 km) from north to south and almost as much from east to west. The rest of it dips under sedimentary rocks that overlap it on the south and west. The Canadian Shield consists of ancient Precambrian rocks, over 500 million years old, predominantly granite and gneiss, with very complex structures indicating several mountain-building episodes. It has been eroded into a rolling surface of low to moderate relief with elevations generally below 2000 ft (600 m). Its surface has been warped into low domes and basins, such as the Hudson Basin, in which lower Paleozoic rocks, including Ordovician limestones, have been preserved. Since the end of the Paleozoic Era, the Shield has been dominated by erosion. Parts of the higher surface remain at about 1500–2000 ft (450–600 m) above sea level, particularly in the Labrador area. The Shield remained as land throughout the Mesozoic Era, but its western margins were covered by a Cretaceous sea and by Tertiary terrestrial sediments derived from the Western Cordillera. *See*

Key:

1 = New England-Maritime Provinces
2 = Piedmont
3 = Blue Ridge Mountains
4 = Ridge and Valley Province
5 = Appalachian Plateau
6 = Interior Low Plateaus
7 = Ozark Province
8 = Ouachita Province
9 = Central Lowland Province
10 = Great Plains Province
11 = Southern Rocky Mountains
12 = Middle Rocky Mountains
13 = Northern Rockies
14 = Brooks Range
15 = Central Alaska
16 = Sierra-Cascade-Coast Mountains
17 = Pacific Troughs
18 = Pacific Coast Ranges
19 = Interior Plateaus of Canada
20 = Columbia Plateau
21 = Basin and Range Province
22 = Colorado Plateaus
23 = Mexican Highland

Fig. 4. Physical divisions of North America.

CRETACEOUS; ORDOVICIAN; MESOZOIC; PALEOZOIC; PRECAMBRIAN; TERTIARY.

The entire exposed Shield was glaciated during the Pleistocene Epoch, and its surface was intensely eroded by ice and its meltwaters, erasing major surface irregularities and eastward-trending rivers that were there before. The surface is now covered by glacial till, outwash, moraines, eskers, and lake sediments, as well as drumlins formed by advancing ice. A deranged drainage pattern is evolving on this surface with thousands of lakes of various sizes. *See* DRUMLIN; ESKER; GLACIAL EPOCH; GLACIATED TERRAIN; MORAINE; PLEISTOCENE; TILL.

At the contact of the old, hard Shield rocks and the overlapping younger sedimentary rocks around it, large erosional basins have formed. They now contain all the Great Lakes, and lakes Winnipeg, Athabasca, Great Slave, and Great Bear. A line drawn through these lakes, together with the St. Lawrence River, forms a circular boundary of the exposed Shield. Within this boundary lies a desolate flat to hilly area, with thousands of lakes between rocky knobs, frozen much of the year, covered by coniferous forest, and dominated by extremely cold subarctic and humid continental climate. Toward the north, the trees become smaller and fewer, and eventually tundra vegetation of mosses and lichen takes over, as the climate becomes even colder and there are no months with average temperatures above 50°F (10°C).

The Hudson Bay basin extends for 1000 mi (1600 km) from north to south. On its southern shore, the Ordovician limestones outcrop in a lowland nearly 180 mi (300 km) wide but lying less than 500 ft (150 m) above sea level. These limestones dip into the bay and are nearly 6000 ft (1800 m) thick in its center. During the glacial period, the ice was centered upon Hudson Bay. During deglaciation, till and outwash were widely spread around the bay; and isostatic rebound of about 500–800 ft (150–250 m) took place, and still continues, at the rate of about 1.6 ft (0.5 m) per 100 years. Former shorelines of Hudson Bay run nearly parallel to the present shoreline. Permafrost exists around Hudson Bay, and much of its vegetation is muskeg, or bogs with small black spruce. *See* BASIN; MUSKEG.

In great contrast to Hudson Bay Lowland is the Baffin Island/Labrador rim, which is the uplifted edge of the Shield rising 5000–6000 ft (1510–1820 m) and was highly eroded by preglacial rivers. During glaciation these river valleys were occupied by glaciers that eroded them into fiords. *See* FIORD.

The southern edge of the Shield is lower than the eastern rim, but it still forms a topographic barrier, particularly along the north shore of the St. Lawrence River and Gulf, where it rises in a 1500–3000 ft (450–900 m) escarpment. In contrast, the west and southwest sides of the Shield are relatively low, and they either disappear under the overlying sediments or face a small sedimentary escarpment. *See* ESCARPMENT.

The Canadian Shield extends into the United States as Adirondack Mountains in New York State, and Superior Upland west of Lake Superior.

Southeastern Coastal Plain. The Southeastern Coastal Plain is geologically the youngest part of the continent, and it is covered by the youngest marine sedimentary rocks. This flat plain, which parallels the Atlantic and Gulf coastline, extends for over 3000 mi (4800 km) from Cape Cod, Massachusetts, to the Yucatán Peninsula in Mexico. It is very narrow in the north but increases in width southward along the Atlantic coast and includes the entire peninsula of Florida. As it continues westward along the Gulf, it widens significantly and includes the lower Mississippi River valley. It is very wide in Texas, narrows again southward in coastal Mexico, and then widens in the Yucatán Peninsula and continues as a wide submerged plain, or a continental shelf, into the sea. *See* COASTAL PLAIN.

The surface of the emerged coastal plain is very low, less than 100 ft (60 m) above sea level, and very flat near the coast; it slowly rises inland up to several hundred feet and becomes rolling, with belts of hills, especially in the Gulf portion. The key to these characteristics is its geologic structure. *See* CONTINENTAL MARGIN.

The entire coastal plain is underlain by gently seaward dipping sedimentary rocks of varying resistance, dating back to Cretaceous, lower and upper Tertiary, and Quaternary. The rocks are often weakly consolidated, and Quaternary deposits that fringe the shoreline and fill the Mississippi Embayment are unconsolidated. *See* QUATERNARY.

The contact between these young rocks and the older rocks of the continent, especially along the Atlantic plain, is called the Fall Line; it is marked by falls and rapids. As the layers of sedimentary rocks are slowly eroding seaward, land-facing cuesta escarpments are formed, and they eventually become belts of hills, such as the Red Hills of Alabama and Mississippi. In addition, the coastal plain rocks are gently arched, or downwarped, affecting the plain's surface and coastal characteristics, and resulting in different coastal subregions. Along the Atlantic coast, the Delaware downwarp creates the Embayed Section with a large Delaware Bay, Chesapeake Bay, and many estuaries; the downwarp in South Carolina and Georgia creates the Sea Island section with several belts of offshore islands. In contrast, the Cape Fear Arch between these two subsections leads to the exposure of the underlying Cretaceous rocks. The broad Peninsular Arch uplifts Florida, exposing upper Tertiary limestone at about 50 ft (15 m) above sea level. Solutional karst features, such as sinkholes, lakes, and springs, dominate this area. *See* KARST TOPOGRAPHY.

On the Gulf Plain, wide belts of Cretaceous rocks are exposed along its inner boundary, while the downwarped Mississippi embayment, reaching deep into the continent, is deeply filled with the youngest Quaternary sediments. The Gulf Coastal Plain is wide and is referred to as a belted plain, because the landward-facing cuesta escarpments, developed on more resistant rocks and forming belts of hills, alternate with parallel lowlands developed on weaker rocks. *See* COASTAL LANDFORMS.

Another characteristic of the Coastal Plain is its complex shoreline, partly related to the depression of the northern part of the continent during glaciations. The Embayed section consists of endless estuaries, including Delaware Bay and Chesapeake Bay, many fringed by miles-long offshore sandbars. Further south, these bars create big lagoons, such as Pamlico Sound. In this area, Cape Lookout and Cape Hatteras are examples of wave-formed capes on the offshore bars enclosing the lagoons. Behind the sandbars are either big lagoons or extensive swamplands. South of Cape Fear, offshore sandbars are replaced by Sea Islands just offshore, and "Carolina Bays" on the coastal plain; the latter are oval-shaped depressions, often holding lakes. The Atlantic coast of Florida is fringed by very long sandbars and narrow lagoons. Sandbars and coral islands form the Florida Keys.

The Everglades are a fresh-water marsh with sawgrass growing up to 12 ft (3.5 m) high, among which are tree-covered islands. This is the largest remaining subtropical wilderness in North America; it includes rich animal and bird life, and is a National Park covering 2200 mi^2 (5700 km^2). *See* WETLANDS.

The shoreline characteristics of the Gulf include sandbars, small estuaries, lagoons, swamp lands, and the Mississippi River floodplain and delta. The floodplain of the Lower Mississippi River is about 600 mi (960 km) long and over 50 mi (80 km) wide, and most of it is subject to flooding. The Mississippi drains over 1 million square miles (2.6 million square kilometers) and delivers over 2 million tons of rock material to the Gulf daily. Its floodplain includes a multitude of active river meanders, oxbow lakes, and abandoned meanders, as well as parallel tributaries such as the Yazoo River, and islands such as the Crowley's Ridge. The delta itself is a unique part of the shoreline where land and water intermingle and numerous distributary channels carry water and silt to the Gulf. *See* DELTA; ESTUARINE OCEANOGRAPHY; FLOODPLAIN; FLUVIAL SEDIMENTS.

A unique inner boundary of the Coastal Plain in Texas is the Balcones Escarpment, a fault scarp that rises sharply nearly 1500 ft (450 m) above the Black Prairie Lowland west and north of San Antonio. The Rio Grande, forming the boundary with Mexico, is the largest river crossing the flat western part of the Coastal Plain.

Extending from Cape Cod, Massachusetts, to Mexico and Central America, the Coastal Plain is affected by a variety of climates and associated vegetation. While a humid, cool climate with four seasons affects its northernmost part, subtropical air masses affect the southeastern part, including Florida, and hot and arid climate dominates Texas and northern Mexico; Central America has hot, tropical climates.

Varied soils characterize the Coastal Plain (Fig. 4), including the fertile alluvial soils of the Mississippi Valley. Broadleaf forests are present in the northeast, citrus fruits grow in Florida, grasslands dominate the dry southwest, and tropical vegetation is present on Central American coastal plains.

Eastern Seaboard Highlands. Between the Southeastern Coastal Plain and the extensive interior provinces lies a belt of mountains that, by their height and pattern, create a significant barrier between the eastern seaboard and the interior of North America. These mountains consist of the Adirondack Mountains and the New England Highlands.

The Adirondack Mountains are a domal extension of the Canadian Shield, about 100 mi (160 km) in diameter, composed of complex Precambrian rocks. They rise to 5600 ft (1700 m), with the highest peak being Mount Marcy, and include more than a dozen peaks whose elevations exceed 4000 ft (1210 m). They are bounded on the south and west by early Paleozoic sedimentary rocks forming inward-facing cuestas, on the north by the St. Lawrence River Lowland, and on the east by the Lake Champlain Lowland. They have been glaciated and are rugged, with hundreds of picturesque lakes, like Lake Placid. *See* LAKE.

The New England Highlands consist of a north-south belt of mountains east of the Hudson Valley, including the Taconic mountains in the south (Mount Equinox is 3800 ft or 1150 m), and the Green mountains in the north (Mount Mansfield is 4400 ft or 1330 m), and continuing as the Notre Dame Mountains along the St. Lawrence Valley and the Chic-Choc Mountains of the Gaspé Peninsula. The large area of New England east of these mountains is an eroded surface of old crystalline rocks culminating in the center as the White Mountains, with their highest peak of the Presidential Range, Mount Washington, reaching over 6200 ft (1880 m). This area has been intensely glaciated, and it meets the sea in a rugged shoreline. Nova Scotia and Newfoundland have a similar terrain.

The rocks underlying New England are Precambrian and early Paleozoic igneous or metamorphic granites, gneisses, schists, quartzites, slates, and marbles. At least one downfaulted valley is filled with Triassic rocks—the 100-mi-long (160-km) Triassic Lowland of Connecticut and Massachussets with volcanic trap ridges, followed by the Connecticut River, which breaks out of the lowland across the old surface to empty into the Long Island Sound.

New England is a hilly to mountainous region carved out of ancient rocks, eroded by glaciers, and covered by glacial moraines, eskers, kames, erratics, and drumlins, with hundreds of lakes scattered everywhere. It has a cool and moist climate with four seasons, thin and acid soils, and mixed coniferous and broadleaf forests.

Appalachian Highlands. The Appalachian Highlands are traditionally considered to consist of four parts: the Piedmont, the Blue Ridge Mountains, the Ridge and Valley Section, and the Appalachian Plateau (**Fig. 5**). These subregions are all characterized by different geologic structures and rock types, as well as different geomorphologies, but their origin and relationships can best be understood when discussed together.

From southeast to northwest is the Fall Line, which is the contact zone between the young, sedimentary rocks of the Coastal Plain and the old crystalline rocks of the Piedmont. At the western edge of the rolling Piedmont, the old rocks rise rapidly and form the Blue Ridge Mountains. West of the Blue Ridge Mountains lies the Great Valley, filled with Paleozoic sedimentary rocks, which is part of the extensive Ridge and Valley Section composed of parallel erosional ridges and valleys carved out of folded Paleozoic rocks. West of this section rises the Allegheny Front, an escarpment which forms the eastern edge of the Appalachian Plateau, a hill land carved out of nearly horizontal Paleozoic sedimentary rocks.

The northern boundary of the entire Appalachian System is an escarpment of Paleozoic rocks trending eastward along Lake Erie, Lake Ontario, and the Mohawk Valley. The boundary then swings south along Hudson River Valley and continues southwestward along the Fall Line to Montgomery, Alabama. The western boundary trends northeastward through Cumberland Plateau in Tennessee, and up to Cleveland, Ohio, where it joins the northern boundary. Together with New England, this region forms the largest mountainous province in eastern United States.

The Piedmont. The Piedmont extends for almost 1000 mi (1600 km) in a northeast-southwest direction from southern New York to central Alabama. Its width varies from less than 20 mi (32 km) in New Jersey to over 120 mi (192 km) in the south. The Piedmont is underlain by complex Precambrian and early Paleozoic crystalline rocks. The youngest rocks are Triassic sedimentaries that have been preserved in downfaulted grabens called Triassic Lowlands. The soils developed on these Triassic sediments are often the best in the Piedmont. The three Triassic basins often include igneous rock layers that form ridges such as the Watchung Mountains in New Jersey, and the columnar basaltic Palisades along Hudson River (50 mi or 80 km long and 1000 ft or 300-m high).

The surface of the Piedmont is rolling to hilly, and its relief increases toward the Blue Ridge Mountains. Several monadnocks such as the granite Stone Mountain (1.5 mi or 2.5 km wide, 200 ft or 60 m high) near Atlanta, Georgia, are spectacular reminders that this is an old erosional plain. Rivers starting in the Blue Ridge Mountains cross the Piedmont and the Coastal Plain to reach the Atlantic. However, beginning with the Roanoke River, the rivers reaching the Piedmont first cross the Blue Ridge where they cut water gaps; others, like the Potomac with the Shenandoah River, cross the Ridge and Valley region; further north, the Susquehanna and the Delaware rivers head in the Appalachian Plaleau. Thus, the drainage divide in the Appalachian System shifts from the Blue Ridge Mountains in the south to the Appalachian Plateau in the north.

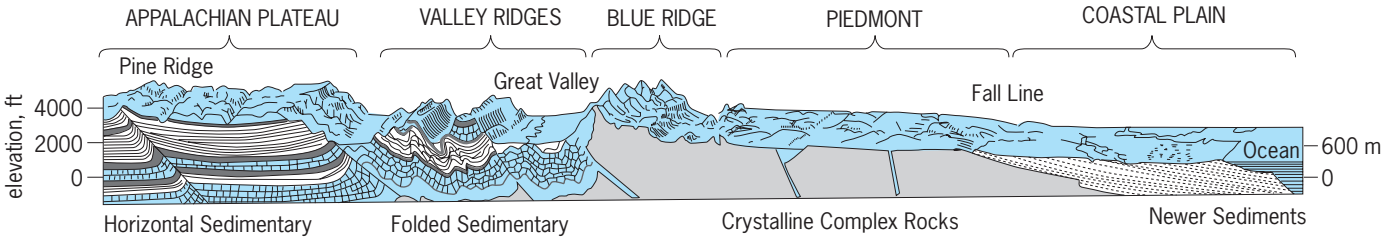

Fig. 5. Profile of the Coastal Plain and the Appalachian Highlands. (*After W. H. Yoho and E. C. Pirkle, Natural Landscapes of the United States, 4th ed., Kendall/Hunt, 1985*)

Blue Ridge Mountains. These mountains lie west of the Piedmont and stretch from Pennsylvania to Georgia for over 540 mi (860 km). They begin as a single ridge in Pennsylvania, and become a high and wide mountain range south of Roanoke, Virginia, where they reach elevations of over 6000 ft (1820 m) for about 45 peaks, and over 5000 ft (1510 m) for nearly 300 peaks. Mount Mitchell is the highest peak at 6684 ft (2025 m) in the most rugged part called the Great Smoky Mountains. Geologically, the Blue Ridge Mountains are a continuation of the Piedmont (Fig. 5). Their national parks include Shenandoah National Park and Great Smoky Mountains National Park.

Ridge and Valley Province. Paralleling the Blue Ridge Mountains, this is the most unusual mountain province in eastern North America. It consists of nearly parallel, alternating ridges and valleys hundreds of miles long, extending 2880 mi (4600 km) from southeastern New York to Alabama. These are a result of intense folding of lower Paleozoic sedimentary rocks of alternating resistant and nonresistant layers, and subsequent erosion of the resulting anticlines and synclines into parallel ridges and valleys (Fig. 5). The width, length, and spacing of the ridges and valleys vary from north to south. In the north, the high, parallel ridges and the valleys are very long, narrow, and closely spaced. The ridges in the central part are densely spaced and are the highest. Approximately south of Roanoke, Virginia, the ridges become lower and the valleys widen. This linear terrain continues for hundreds of miles but eventually becomes less distinct and ends south of Birmingham, Alabama.

Among the unique features of the Ridge and Valley Province are superposed rivers which had cut across ridges forming water gaps, such as the Delaware Gap across the Kittatinny Mountains with 1000-ft-high (300-m) cliffs rising above the river; or the Susquehanna River Gap north of Harrisburg, Pennsylvania; and the Potomac River Gap at Harpers Ferry, West Virginia.

The thick Ordovician Paleozoic limestones of the Ridge and Valley section give rise to karstic features, such as the extensive and beautiful caverns with striking stalactites and stalagmites in Luray and Endless caverns near Luray, Virginia. *See* CAVE; STALACTITES AND STALAGMITES.

Appalachian Plateau. Covering large parts of seven states, the Appalachian Plateau is the widest part of the Appalachian System (Fig. 5). It is bound on the north by the escarpment of the Devonian rocks rising over the Mohawk Valley, and by Lake Ontario and Lake Erie lowlands. Its characteristic rock structure is nearly horizontal layers of Paleozoic rocks, mostly sandstones, conglomerates, and shales, from Devonian to Permian.

The terrain of the Appalachian Plateau is mainly hilly. The Paleozoic rocks that rise from 1000 ft (300 m) in the west to 5000 ft (1510 m) in West Virginia have been eroded by dendritically branching streams into hills and low mountains with deep valleys. The northern section is drained predominantly by eastward-flowing rivers, like the Delaware and Susquehanna, while the central and southern parts are drained by southward-flowing rivers, mainly the Ohio and Tennessee.

In the northeast corner of the Plateau the rocks rise to nearly 4000 ft (1200 m) and form the Catskill Mountains. Other mountainous sections are the Pocono Mountains in eastern Pennsylvania, the Allegheny Mountains in Pennsylvania and West Virginia, and the Cumberland Mountains in Kentucky and Tennessee. These mountains are formed from resistant Pottsville and Pocono sandstones.

The northern edge of the Appalachian Plateau has been glaciated, leaving a large drumlin field between Syracuse and Rochester, New York, and the Finger Lakes, located south of the drumlin field. This is a belt of parallel, north-south-trending, narrow lakes, up to 40 mi (64 km) long, formed when advancing glaciers deepened the valleys of preglacial northward-flowing rivers, and subsequently blocked them with moraines to form the lakes.

The southern part of the Appalachian Plateau becomes narrow and forms the Cumberland Plateau, with a pronounced series of 700 mi (1120 km) long, east-facing escarpments up to 1000 ft (300 m) high.

The Appalachian Plateau is one of the largest continuous bituminous coal areas in the world, and its northern part is the major coal mining region in North America.

Interior Domes and Basins Province. The southwestern part of the Appalachian Plateau, overlain mainly by the Mississippian and Pennsylvanian sedimentary rocks, has been warped into two low structural domes called the Blue Grass and Nashville Basins, and a structural basin, drained by the Green River; its southern fringe is called the Pennyroyal Region. The Interior Dome and Basin Province is contained roughly between the Tennessee River in the south and west, and the Ohio River in the north.

There is no boundary on the east, because the domes are part of the same surface as the Appalachian Plateau. However, erosional escarpments, forming a belt of hills called knobs, clearly mark the topographic domes and basins. The northern dome, called the Blue Grass Basin or Lexington Plain, has been eroded to form a basin surrounded by a series of inward-facing cuesta escarpments. The westernmost cuesta reaches about 600 ft (180 m) elevation while the central part of the basin lies about 1000 ft (300 m) above sea level, which is higher than the surrounding hills. This gently rolling surface with deep and fertile soils exhibits some solutional karst topography. *See* FLUVIAL EROSION LANDFORMS.

The southern dome, the Nashville Basin, is truly a topographic basin with average elevations in the center of about 600 ft (180 m) being lower than the surrounding Eastern Highland Rim, which reaches 1300 ft (400 m), and the Western Highland Rim, with elevations of 900 ft (270 m). The basin is mainly a rolling and fertile plain surrounded by the cuesta escarpments.

The third part of this province is both a topographic and a structural basin, drained northwestward by the Green River. It south side is the

Pennyroyal limestone plain, above which rise two escarpments. The Green River is the only surface stream that crosses the karst plain. Thousands of sinkholes, lost creeks, underground rivers, and caves make it a classic karst topography. Within it lies the Mammoth Cave, an enormous solutional cave system developed in thick limestones.

Ozark and Ouachita Highlands. The Paleozoic rocks of the Pennyroyal Region continue Westward across southern Illinois to form another dome of predominantly Ordovician rocks, called the Ozark Plateau. This dome, located mainly in Missouri and Arkansas, has an abrupt east side, and a gently sloping west side, called the Springfield Plateau. Its surface is stream eroded into hilly and often rugged topography that is developed mainly on limestones, although shales, sandstone, and chert are present. Much residual chert, eroded out of limestone, is present on the surface. There are some karst features, such as caverns and springs. In the northeast, Precambrian igneous rocks protrude to form the St. Francois Mountains, which reach an elevation of 1700 ft (515 m).

In Arkansas, the Plateau surface rises to 2500 ft (757 m) and forms the 200-mi-long (320-km) and 35-mi-wide (56-km) Boston Mountains, which have stream-eroded tops, and a 300–600-ft-high (90–180-m) escarpment along its southern side. South of the Arkansas River Valley lie the Ouachita Mountains, over 200 mi (320 km) long and 100 mi (160 km) wide, and best characterized by their east-west-trending ridges and valleys. These ridges and valleys have been eroded from folded and thrust-faulted sedimentary rocks, such as shales and sandstones mainly of Pennsylvanian age. The mountains reach elevations of over 2600 ft (780 m) in the center, with ridges rising 1500 ft (450 m) above the valleys. In the Ouachita Mountains Hot Springs National Park, there are 47 hot springs with water temperature of 143°F (61°C) and a combined flow of over a million gallons (4000 m^3) a day. *See* PENNSYLVANIAN.

Central Lowlands. One of the largest subdivisions of North America is the Central Lowlands province which is located between the Appalachian Plateau on the east, the Interior Domes and Basins Province and the Ozark Plateau on the south, and the Great Plains on the west. It includes the Great Lakes section and the Manitoba Lowland in Canada. This huge lowland in the heart of the continent (whose elevations vary from about 900 ft or 270 m above sea level in the east and nearly 2000 ft or 600 m in the west) is underlain by Paleozoic rocks that continue from the Appalachian Plateau and dip south under the recent coastal plain sediments; meet the Cretaceous rocks on the west; and overlap the crystalline rocks of the Canadian Shield on the northeast.

The present surface of nearly the entire Central Lowlands, roughly north of the Ohio River, and east of the Missouri River, is the creation of the Pleistocene ice sheets. When the ice formed and spread over Canada, and southward to the Ohio and Missouri rivers, it eroded much of the preexisting surface. During deglaciation, it left its deposits over the Canadian Shield and the Central Lowlands.

During its four (Nebraskan, Kansan, Illinoian, and Wisconsin) or more glacial periods, the ice eroded the contact zone between the Canadian Shield and the Central Lowlands into great lowlands, some of which now hold big lakes. Preglacial river valleys, as well as structurally controlled erosional lowlands between cuestas around the Michigan Basin, directed the ice movement southward and helped form the Great Lakes basins. As the ice gradually melted, the Great Lakes basins filled with water along the edge of the ice, forming the huge Lake Agassiz and the predecessors of the Great Lakes; they eventually established their drainage to the east, and acquired their present shapes and sizes.

Toward the end of the evolution of the Great Lakes, the Niagara Falls were cut. Initially, when Lake Erie drained through the Niagara River into Lake Ontario, the latter's elevation was about 150 ft (45 m) higher than at present and there were not falls on the Niagara River. When ice melted and the Mohawk outlet opened to the east, Lake Ontario began to drain rapidly, and the Niagara River began to cut its valley to keep up with the lake's falling level. The falls developed on the Niagara escarpment, and have since retreated about 7 mi (11 km), leaving a deep gorge. They continue to retreat at a rate of about 4 ft (1.2 m) per year, mostly by rock falls, as the underlying soft shale and sandstone are undermined, and the dolomite edge breaks off and falls down. Eventually the falls will become lower, undermining of the underlying rocks will cease, and the falls may become just rapids.

The overall surface characteristics of the Central Lowlands were created by the ice sheets. During ice advance, debris deposition was often greater than erosion. In areas where surface resistance equalled ice stress, the ice just moved over its own debris and streamlined it into oval-shaped hills, called drumlins. Among the world's largest fields of drumlins are those in eastern Wisconsin, and north and south of Lake Ontario.

During ice melting, long belts of hilly moraines formed at the end of ice lobes. Between the moraines are stretches of rolling till plains composed of heterogeneous glacial debris, including silt, sand, gravel, boulders, and large erratics from distant areas. Flat sandy surfaces, called outwash plains, were formed by meltwaters; clay-covered lacustrine flats are formed lake bottoms. Smaller features include kames and eskers. Scattered between these features are numerous lakes and swamps, and deranged river systems. *See* KAME.

All these features are most striking in areas covered by the most recent ice advance, the Wisconsin, especially in Minnesota, Wisconsin, and Michigan.

A unique terrain is the Driftless Area of southwestern Wisconsin, which is surrounded by glacial debris of different ice lobes but shows no evidence of having been covered by ice. Here, the Paleozoic rock surface, which has been buried by glacial material everywhere else, is being eroded by streams into deep valleys and flat uplands.

The Central Lowlands are drained by the third

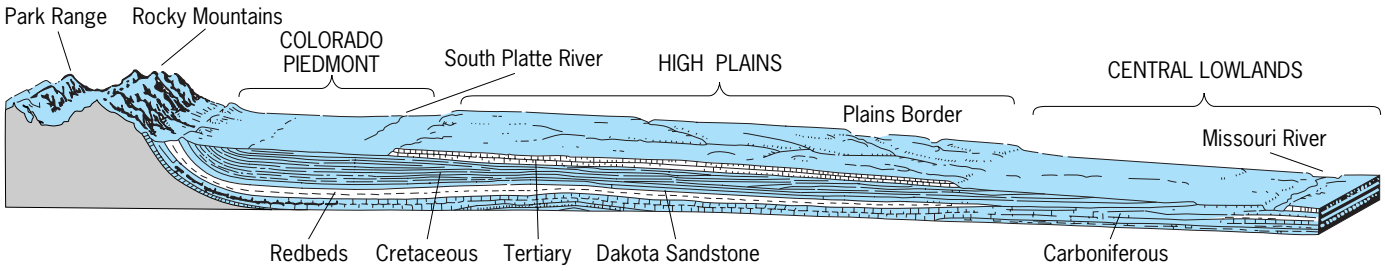

Fig. 6. Rock structure of the Great Plains. (*After W. H. Yoho and E. C. Pirkle, Natural Landscapes of the United States, 4th ed., Kendall/Hunt, 1985*).

longest river system in the world, the Missouri-Mississippi, which is 3740 mi (6000 km) long. This mighty river system, together with the Ohio and the Tennessee, drains not only the Central Lowlands but also parts of the Appalachian Plateau and the Great Plains, before it crosses the Coastal Plain and ends in the huge delta of the Mississippi. The river carries an enormous amount of water and alluvium and continues to extend its delta into the Gulf. In 1993 it reached a catastrophic level of a hundred-year flood, claimed an enormous extent of land and many lives, and created an unprecedented destruction of property. This flood again alerted the population to the extreme risk of occupying a river floodplain. *See* FLOODPLAIN; RIVER.

Great Plains. The Great Plains, which lie west of the Central Lowlands, extend from the Rio Grande and the Balcones Escarpment in Texas to central Alberta in Canada. On the east, they are bounded by a series of escarpments, such as the Côteau du Missouri in the Dakotas. The dry climate with less than 20 in. (50 cm) of precipitation, and steppe grass vegetation growing on calcareous soils, help to determine the eastern boundary of the Great Plains. On the west, the Great Plains meet the abrupt front of the Rocky Mountains, except where the Colorado Piedmont and the lower Pecos River Valley separate them from the mountains (**Fig. 6**).

These extensive Great Plains are high and slope eastward from 5500 ft (1660 m) elevation at the foothills of the Rocky Mountains to 1500 ft (450 m) along their eastern boundary. They are overlain by terrestrial Tertiary deposits, resting on marine Cretaceous formations. The latter were laid down in a Mesozoic Sea and are exposed in the Piedmont along the eastern edge of the Plains, and in the northwest and north. In Colorado, the Cretaceous rocks rise against the mountains, where their resistant Dakota sandstone layers form long, striking hogback ridges that parallel the mountain front.

The Tertiary rocks forming the surface of the Great Plains were deposited during the prolonged erosion of the Rocky Mountains during which the sediments were carried by streams eastward and spread over the Cretaceous surface. The most important top formation is the Ogallala, which consists of sands, gravels, silts, and an irregular massive calcrete formed under dry conditions. The Ogallala is the most valuable aquifer, or water-bearing formation, of the Great Plains. However, much of it has been eroded, forming pediments in Colorado, and Canadian badlands along Deer River near Calgary. Erosion of the Tertiary deposits gave rise to the Sand Hills of Nebraska, and extensive layers of loess, or windblown silt, farther east. Tertiary igneous features such as lava mesas are present in the Great Plains in the Raton section in northeastern New Mexico. Volcanic ash deposits are present farther east. *See* AQUIFER; LOESS.

The Great Plains region shows distinct differences between its subsections from south to north. The southernmost part, called the High Plains or Llano Estacado, and Edwards Plateaus are the flattest. While Edwards Plateau, underlain by limestones of the Cretaceous age, reveals solutional karst features, the High Plains have the typical Tertiary bare cap rock surface, devoid of relief and streams.

The central part of the Great Plains has a recent depositional surface of loess and sand. The Sand Hills of Nebraska form the most extensive sand dunes area in North America, covering about 24,000 mi^2 (62,400 km^2). They are overgrown by grass and have numerous small lakes. The loess region to the south provides spectacular small canyon topography. *See* DUNE.

The northern Great Plains, stretching north of Pine Ridge and called the Missouri Plateau, have been intensly eroded by the western tributaries of the Missouri River into river breaks and interfluves. In extreme cases, badlands were formed, such as those of the White River and the Little Missouri.

Another unusual feature of northwestern Great Plains are the small mountain groups of intrusive, extrusive, and even diastrophic origin, some reaching over 11,000 ft (3330 m) elevation (Crazy Mountains). The largest are the Black Hills of South Dakota, a classical partially eroded dome 60 mi (96 km) wide and 120 mi (190 km) long. The highest elevation is in Harney Peak (7240 ft or 2190 m); and the best-known peak is Mount Rushmore, with heads of four presidents carved in its granite face.

The terrain of the Canadian Great Plains consists of three surfaces rising from east to west: the Manitoba, Saskatchewan, and Alberta Prairies developed on level Creteceous and Tertiary rocks. Climatic differences between the arid and warm southern part and the cold and moist northern part have resulted in regional differences. The eastern boundary of the Saskatchewan Plain is the segmented Manitoba Escarpment, which extends for 500 mi (800 km) northwestward, and in places rises 1500 ft (455 m)

above the Manitoba Lowland. Côteau du Missouri marks the eastern edge of the higher Alberta Plain.

Western Cordillera. The mighty and rugged Western Cordilleras stretch along the Pacific coast from Alaska to Mexico. There are three north-south-trending belts: (1) Brooks Range, Mackenzie Mountains, and the Rocky Mountains to the north and Sierra Madre Oriental in Mexico; (2) Interior Plateaus, including the Yukon Plains, Canadian Central Plateaus and Ranges, Columbia Plateau, Colorado Plateau, and Basin and Range Province stretching into central Mexico; and (3) Coastal Mountains from Alaska Range to California, Baja California, and Sierra Madre Occidental in Mexico.

This subcontinental-size mountain belt has the highest mountains, greatest relief, roughest terrain, and most beautiful scenery of the entire continent. It has been formed by earth movements resulting from the westward shift of the North American lithospheric plate. The present movements, and the resulting devastating earthquakes along the San Andreas fault system paralleling the Pacific Ocean, are part of this process. *See* CORDILLERAN BELT; PLATE TECTONICS.

Rocky Mountains. This very high, deeply eroded and rugged mountainous region comprises several distinct parts: Southern, Middle, and Northern Rockies, plus the Wyoming Basin in the United States, and the Canadian Rockies. The Southern Rockies, extending from Wyoming to New Mexico, include the Laramie Range, the Front Range, and Spanish Peaks with radiating dikes on the east; Medicine Bow, Park, and Sangre de Cristo ranges in the center; and complex granite Sawatch Mountains and volcanic San Juan Mountains of Tertiary age on the west. Most of the ranges are elongated anticlines with exposed Precambrian granite core, and overlapping Paleozoic and younger sedimentary rocks which form spectacular hogbacks along the eastern front. There are about 50 peaks over 14,000 ft (4200 m) high, while the Front Range alone has about 300 peaks over 13,000 ft (3940 m) high. The southern Rocky Mountains, heavily glaciated into a beautiful and rugged scenery with permanent snow and small glaciers, form a major part of the Continental Divide.

The Middle Rocky Mountains consist of the Wyoming and Big Horn basins, and separate elongated anticlinal ranges with granite cores and flanking hogbacks, including the Big Horn, Owl Creek, Wind River, and Uinta. The highly eroded volcanic Absaroka Mountains, the crystalline Beartooth-Snowy Mountains, the volcanic Yellowstone Plateau, the faulted Teton Range, and the Wasatch Range are also part of the Middle Rockies. The maximum peak elevations are over 13,000 ft (3940 m), with the Grand Teton reaching over 13,700 ft (4150 m). All the Middle Rocky ranges have been intensely glaciated, and have deep U-shaped valleys, cirques, horns, moraines, and thousands of lakes.

The Yellowstone Plateau is a high basin with elevations from about 7000 to 8500 ft (2120 to 2575 m) surrounded by mountains whose elevations exceed 10,000 ft (3030 m). It is filled with basaltic and colorful rhyolitic lavas of Tertiary age, has over 4000 hot springs and brightly colored mineral springs, and numerous geysers that spout steam intermittently, including The Old Faithful which erupts regularly. The large Yellowstone Lake, 300 ft (90 m) deep, at 7750 ft (2350 m) elevation, is drained northward by Yellowstone River which cuts a 1200-ft-deep (360-m) canyon. *See* GEYSER; SPRING (HYDROLOGY).

The large rivers of the Middle Rockies and the Wyoming Basin, including the Yellowstone, Big Horn, Green, and Snake, leave this province through deep canyons, which they cut across the high ranges on which they were superposed from a higher surface.

The Northern Rocky Mts. are the widest, although not the highest, of the United States Rockies. They are located mainly in Idaho and western Montana. Their southern part, 6000–8000 ft (1820–2420 m) high, is carved into huge ranges from the nearly homogeneous granite mass of the Idaho Batholith of Jurassic and Cretaceous age. The highest peaks rise to 12,000 ft (3630 m). The northern part is underlain by Precambrian metamorphic and sedimentary rocks, which have been folded, faulted, and thrust eastward. The most spectacular example is the Lewis overthrust (Chief Mountain) on the United States–Canadian boundary, which moved over 120 mi (200 km) eastward (**Fig. 7**). The southeastern part of the Northern Rockies consists of short fault block ranges of Paleozoic and Mesozoic sedimentary rocks.

The Northern Rockies are drained by the Solomon, Clearwater, and Clark rivers. They were glaciated but, being lower, did not develop a glacial terrain as spectacular as the higher Middle and Southern Rockies, except in Glacier National Park on the Canadian border. Here, more than 50 small glaciers are still present, and majestic horns, aretes, cirques, U-shaped valleys, and lakes dominate the scenery.

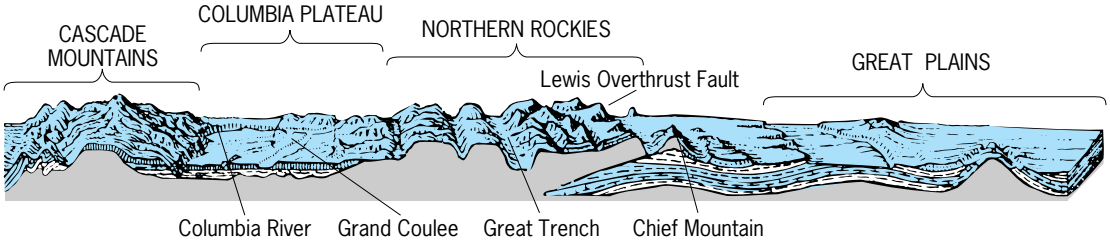

Fig. 7. Profile of the Great Plains, Northern Rockies, Columbia Plateau, and Cascade Mountains. (*After W. H. Yoho and E. C. Pirkle, Natural Landscapes of the United States, 4th ed., Kendall/Hunt, 1985*).

The Canadian Rocky Mountains begin as Purcell, Selkirk, Monashee and Cariboo ranges, together known as Columbia Mountains, merging northward into the Eastern Front Ranges. They are composed of Precambrian cores with overlapping Paleozoic and Mesozoic sediments; the Paleozoic limestones are thrust eastward over the Mesozoic rocks. The glaciated Front Range reaches almost 13,000 ft (3940 m) in elevation (Mount Robson). Its Columbian ice field (300 ft or 90 m thick) lies at 10,000–12,000 ft (3030–3640 m) elevation on the Continental Divide. West of the Canadian Rockies and parallel to them trends the nearly 1000-long (1600-km) Rocky Mountain Trench, a straight, flat-bottom, downfaulted valley, several miles wide and up to 3000 ft (910 m) deep, followed by such rivers as the Columbia, Koothenay, and Frazer. North of Liard River begins the Tintina Trench that continues for 450 mi (720 km). East of this trench lie the Mackenzie Mountains, which continue as Richardson Mountains, and the treeless 600-mi-long (960-km) 9000-ft-high (2730-m) Brooks Range. North of Brooks Range and its foothills lies the extensive frozen Arctic Coastal Plain, covered by gravels, sands, and silts of Pleistocene age, and exhibiting frozen ground features such as patterned ground.

Interior Plateaus and Ranges Province. This portion of the Western Cordillera lies between the Rocky Mountains and the Coastal Mountains. It is an extensive and complex region. It begins in the north with the wide Yukon Plains and Uplands; narrows into the Canadian Central Plateaus and Ranges; widens again into the Columbia Plateau, Basin and Range Province, and Colorado Plateau; and finally narrows into the Mexican Plateau and the Central American isthmus.

1. The Yukon Plains and Uplands is a region of low valleys, dissected plateaus, and low mountains. The lowlands include the Yukon Flats, underlain by scattered permafrost and covered by loess; and the lower Yukon plain and delta. The dissected upland between Yukon and Tanaka rivers continues east as the Klondike Region.

Continuing southward, the Canadian Central Plateaus and Ranges region, lower than the mountains to the east and west, remains complex and mountainous, and reveals widespread glaciation, including numerous lakes. The Stikine Mountains in the north and the Columbia Mountains in the south, with peaks over 11,000 ft (3330 m), are the roughest and highest ranges. The rocks of this area are mainly Paleozoic and Precambrian sedimentaries and metamorphics. The Frazer Plateau west of Columbia Mountains is a rolling upland developed mainly on intrusive and extrusive igneous rocks of Tertiary age.

2. The Columbia Plateau (Fig. 7), which lies south of the United States–Canadian border, and between the Northern Rockies and the Cascade Rages, is one of the largest lava-covered surfaces in the Western Hemisphere. It covers over 100,000 mi^2 (260,000 km^2), and consists of many flows of Tertiary lavas with individual layers reaching up to 200 ft (60 m) thick. The total thickness of the lavas probably reaches to 8000 ft (2420 m). *See* LAVA.

In the center of Columbia Plateau, the Blue Mountains, reaching over 9000 ft (2730 m), project through the lavas. They consist of Paleozoic and Mesozoic sedimentary rocks that were part of the pre-lava surface. The northwestern part of Columbia Plateau, called the Walla Walla, is its lowest part; at the junction of the Columbia and Snake rivers, the elevations are only about 300–400 ft (90–120 m). In the south, the elevations are about 5000 ft (1510 m); in the northeast part of the Snake River Plain, about 6000 ft (1810 m); and the mountains in the center reach, over 9000 ft (2730 m).

In the northern section, near Yakima, Washington, there are broad anticlinal and synclinal ridges and valleys. In the northeast are the unique Channeled Scablands, a 2000-mi^2 (5200-km^2) area of dry channels separated by eroded and nearly bare lava surfaces. It appears that the channels were cut by ice meltwaters when a dam holding glacial Lake Missoula broke and created a sudden flood of enormous proportions that spilled across the area. The largest channel, the Grand Coulee, is about 50 mi (80 km) long, several miles wide, and nearly 1000 ft (300 m) deep. Its Dry Falls have a 450-ft (135-m) drop.

The Snake River Plain is drained by Snake River, which begins at about 6000 ft (1810 m) elevation in the Yellowstone Plateau and flows westward over the American Falls (50 ft or 15 m), Twin Falls (180 ft or 55 m), and Shoshone Falls (212 feet or 64 m). Here, it creates a deep gorge with the spectacular Thousand Springs, which emerge 185 ft (56 m) above the valley floor from porous layers of basalt, yielding a flow of over 2 million gallons (7.5 million liters) per minute. The Plain is quite level, except for the striking Craters of the Moon, with lunarlike landscape, where 60 volcanic cones rise over 300 ft (90 m).

3. The Great Basin, or the Basin and Range Province (**Fig. 8**), an enormous desert region of over 300,000 mi^2 (780,000 km^2), liessouth of the

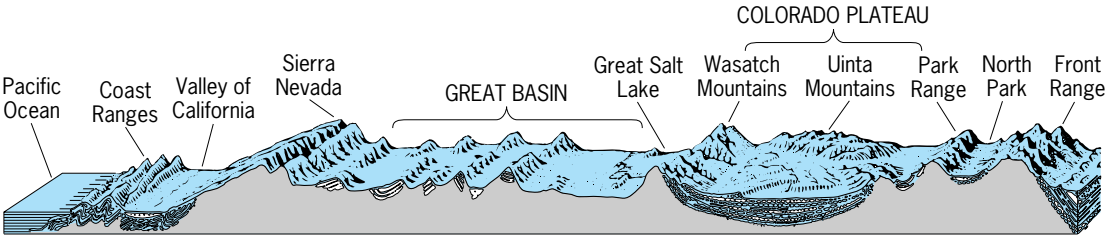

Fig. 8. Geologic cross section from the Rocky Mountains to the Coastal Ranges. (*After W. H. Yoho and E. C. Pirkle, Natural Landscapes of the United States, 4th ed., Kendall/Hunt, 1985*).

Columbia Plateau; it occupies Nevada, western Utah, southern California, southern Arizona, and parts of New Mexico. It has a strikingly different landscape from that of any surrounding regions. It is composed of nearly straight, parallel, and north-south-trending high ranges, with steep sides separated by flat desert basins. The ranges are a result of block faulting, which began in Tertiary time and continues. The basins often contain intermittent lakes, or playas. *See* PLAYA.

The ranges are 50–100 mi (80–160 km) long and 5–15 mi (8–25 km) wide, rising above the basins up to 6000 ft (1810 m), and reaching elevations of 10,000–13,000 ft (3030–3940 m). The ranges are generally asymmetrical, with one steeper face.

In addition to being unique in terrain characteristics, the region is a complex mosaic of different rock types and ages found adjacent to each other. Another unique characteristic of this area is the internal drainage of the northern part (the Great Basin) and its hydrologic history dating back to the cooler and more rainy climates of the Pleistocene Epoch (pluvial periods). The major river of this internal drainage area is the Humbold, which flows across Nevada and disappears in Carson Lake. The largest lake is the Great Salt Lake in northwestern Utah, covering about 2000 mi^2 (5200 km^2), with a depth of about 15–20 ft (4.5–6 m). This is a remnant of the former Lake Bonneville, which was 10 times as large and probably 1000 ft (300 m) deep.

In southern California lies the Death Valley, a downfaulted basin, 282 ft (85 m) below sea level, surrounded be some of the highest ranges and creating the greatest amount of relief in the Basin and Range region.

The Basin and Range area has several distinct subsections: the Great Basin, the Mojave and Sonoran desert, the Mexican Highlands, and the Sacramento Mountains; it continues as an extensive intermountain basin and range region into the Central Plateau of Mexico.

The Great Basin in Nevada has the highest, longest, and most closely spaced north-south-trending ranges. The Sonoran Desert in southern California has lower, irregular, and less oriented and less parallel ranges, which show a great amount of erosion. The Salton Sea, a downfaulted basin at 235 ft (71 m) below sea level, and Imperial Valley together cover 2500 mi^2 (6500 km^2).

The Mexican Highlands to the southeast have larger and higher ranges reaching elevations up to 10,000 ft (3030 m; Mount Graham), but the orientation of the ranges remains complex. The Sacramento Mountains include the Guadalupe Mountains with the well-known peak, El Capitan, over 8000 ft (2420 m) high. This is a sheer cliff of fossil limestone reef, the world's largest such cliff known. Nearby are the Carlsbad Caverns, one of the largest limestone cavern systems in the world, reaching over 1000 ft (300 m) deep.

The entire southern part of the Basin and Range Province, in contrast to the Great Basin, has an open drainage mainly by the Colorado and Gila rivers, and the Rio Grande, accounting for greater erosion and removal of debris from this area.

4. The Colorado Plateau (**Fig. 9**), occupying large parts of Utah, Arizona, Colorado, and New Mexico, covers over 150,000 mi^2 (390,000 km^2) and is surrounded by the Rocky Mountains on the north and east, and the Basin and Range Province on the southeast, south, and west. Its major characteristics are nearly horizontal sedimentary rocks of Paleozoic, Mesozoic, and Cenozoic age, hundreds of canyons reaching the scope of the Grand Canyon of the Colorado River, and endless erosional escarpments that form as the upper rock layers gradually erode away. Large areas of lava flows and hundreds of volcanic cones are also present, as are several igneous intrusive domes. Precambrian metamorphic rocks are present in the lowest part of the Grand Canyon, which reveals the entire sedimentary sequence of the overlying rocks. The north-south profile of the Colorado Plateau (Fig. 9) shows nearly level Paleozoic sedimentary rocks of the southern part. The surface rises northward toward the Grand Canyon, where in the Kaibab Plateau it reaches 8000–9000 ft (2420–2730 m) elevation. From there, the enormous sequence of sedimentary rocks gently dips

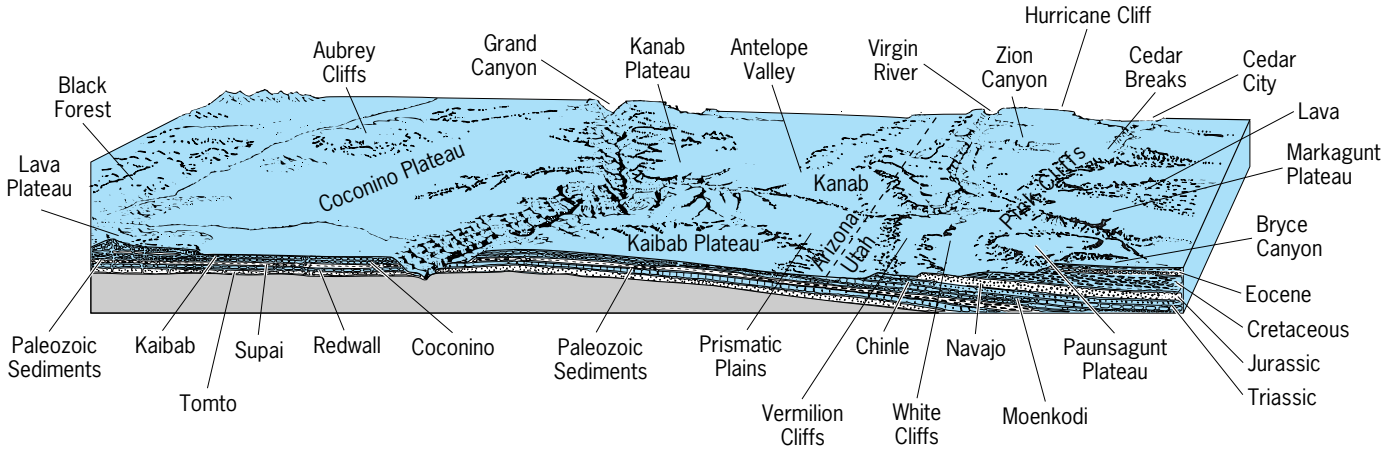

Fig. 9. Profile of the Colorado Plateaus (north-south). (*After W. H. Yoho and E. C. Pirkle, Natural Landscapes of the United States, 4th ed., Kendall/Hunt, 1985*).

northward, and the top layers are eroded into enormous escarpments. These escarpments rise steplike from the lowest Triassic Chocolate Cliffs, to Jurassic Vermillion and White Cliffs, to Cretaceous Gray Cliffs, and to Tertiary Pink Cliffs, each over 1000–2000 ft (300–600 m) high. In the north, the Uinta Basin, containing the same sequence of rocks, is bounded on the south by high escarpments reaching over 4000 ft (1210 m).

The east-west Colorado Plateau profile shows a series of upwarps and enormous faults that have cut the surface into north-south Utah plateaus, including Kaibab (9000 ft or 2730 m), Kanab, Uinkaret, and Shivwits (5000 ft or 1510 m), separated by escarpments, and all bounded on the west by the Grand Cliffs, 3000–4000 ft (900–1210 m) high.

The southwestern part of the Plateau, around San Francisco Mountains (12,800 ft or 3880 m), and the southeastern part, around Mount Taylor (11,400 ft or 3450 m), are extensive Tertiary lava surfaces with hundreds of volcanic cones.

The Central Canyon Lands section is best known for its spectacular 217-mi-long (347-km) Grand Canyon of the Colorado River (Fig. 9), which cuts a 6000-ft-deep (1820-m) canyon into the Paleozoic rocks, with the lowest 1000 ft (300 m) cut into Precambrian metamorphics. The colorful walls of the canyon, which is over 5 mi (8 km) wide, include cliff makers like the White Permian Kaibab limestone (the cap rock), the Mississippian Red Wall limestone, and the Permian Supai sandstone, as well as shales that form the platforms. These horizontal beds of Paleozoic and later sedimentary rocks have been elevated from the bottom of the ocean up to 9000 ft (2730 m), and now extend over large surfaces, like the Kaibab Plateau, or the lower Coconino Plateau. *See* PALEOGEOGRAPHY.

Over the length of the Grand Canyon, the Colorado River flows rapidly over many rapids, dropping from over 2600 ft (780 m) to 1000 ft (300 m) in elevation. Other big canyons are those of the Green, Gunnison, San Juan, and Little Colorado rivers.

Surface features of intrusive origin are the Abaja, La Sal, and Henry mountains, as well as eroding anticlinal folds like San Rafael Swell and Zuni Uplift in the southeast.

The Navajo section south of the Grand Canyon consists of numerous mesas, Painted Desert, and Petrified Forest with great fossilized logs of conifers of Trassic age. *See* PETRIFIED FORESTS.

The Colorado Plateau, with its enormous variety of unique terrain features, has an exceptional number of national parks, including the Grand, Zion, Bryce Canyons, Canyonland, Arches, Petrified Forest, Capital Reef, and Mesa Verde, and Sunset Crater National Monument.

Coastal Lowlands and Ranges. These extend along the entire length of North America and include Alaskan Coast Ranges, Aleutian Islands, Alaska Range, Canadian Coast Ranges, and a double chain of the Cascade Mountains and Sierra Nevada on the east, and Coast Ranges on the west, separated by the Puget Sound, Willamette Valley, and the Great Valley of California. These ranges continue southward as Lower California Peninsula, Baja California, and Sierra Madre Occidental in Mexico.

The Alaskan Coast Ranges include the Chugach-Kenai Mountains, Wrangel Mountains, and glacier-covered St. Elias Range. They are extremely rugged due to deep ice erosion. This entire windward coast receives heavy precipitation which, combined with high latitude, forms snow and ice and results in intense erosion by glaciers. The spectacular St. Elias Range of folded and faulted Paleozoic and Mesozoic sedimentary rocks, with Mount Logan at 19,850 ft (6000 m), and many peaks higher than 15,000 ft (4550 m), are covered by numerous glaciers. The Seward Valley glacier spreads out to form the extensive piedmont Malaspina Glacier.

North of these coastal mountains lies the huge arc of the majestic Alaska Range, which continues westward as Alaska Peninsula and Aleutian Islands. The Alaska Range of folded and faulted sedimentary Paleozoic and Mesozoic rocks, and extensive granite intrusions forming the lofty peaks, is 600 mi (960 km) long, and is the highest range in North America, with permanently snow-covered Mount McKinley reaching 20,320 ft (6160 m).

The Alaska Peninsula and the Aleutian Range form a volcanic arc with dozens of volcanoes, many active in recent centuries. Though the Aleutian volcanoes are generally less than 4000 ft (1210 m) high, the northern ones rise to 10,000 ft (3030 m). One of the most violent recent eruptions in this area was that of Mount Katmai, at the beginning of the twentieth century, which spread an enormous amount of ash, gas, and steam at extremely high temperatures.

South of the Aleutian Island chain lies the deep oceanic Aleutian Trench, an area of frequent earthquakes, with submerged peaks exceeding 25,000 ft (7575 m) in height. *See* SEAMOUNT AND GUYOT.

The Canadian Coastal Mountains continue southward from St. Elias Mountains along the Alaska Panhandle and, beginning with Glacier Bay near Juneau, become the beautiful Alexander Archipelago of hundreds of islands separated by deep, narrow, and steep-sided fiords carved out by glaciers. The two largest islands are the Queen Charlotte Island and Vancouver Island, reaching over 7000 ft (2120 m) elevation. These islands represent the outer coastal ranges, while the straits represent the submerged lowland, separating the outer from the inner coastal mountains. The magnificent fiorded coast is backed inland by the high granite Canadian Coast Ranges, reaching up to 13,260 ft (4020 m) in Mount Waddington. These mountains are capped by big glaciers and snow fields and possess striking glacial features such as aretes, cirques, and horns. *See* CIRQUE.

South of the United States boundary begins the double chain of the coastal mountains: the Cascade–Sierra Nevada and the Coastal Ranges. The Cascade Mountains, over 600 mi (960 km) long, consists of two parts differing in geology and terrain. The Northern Cascades (Fig. 7) begin in Canada and continue to Mount Rainier. They are composed mainly of Paleozoic and Mesozoic sedimentary and metamorphic

rocks, eroded into broad ranges, rugged peaks, and very deep valleys carved out by glaciers. Their crest elevations are about 8500 ft (2575 m), but the largest volcanic cones reach over 10,000 ft (3030 m). They are covered by glaciers.

The Middle and Southern Cascades are predominantly a constructed range that was built by lava accumulations from volcanic vent extrusions. The range has long western slopes and a shorter steep slope on the east, along which is a line of volcanic cones including glacier-capped Mount Rainier, the highest (14,410 ft or 4370 m); glacier-capped Mount Adams (12,300 ft or 3730 m); Mount St. Helen, which erupted three times in 1980, reducing its height by 1300 ft (400 m); and Mount Hood, Mount Jefferson, and Three Sisters. The biggest volcanic cone in the south is the beautiful double cone of Mount Shasta, 14,160 ft (4290 m) high and covered by several glaciers. One of the most interesting features is Crater Lake (5 mi or 8 km wide and 2000 ft or 600 m deep), with Little Wizard Island volcano inside. It is located in a caldera of former Mount Mazama, which erupted about 7000 years ago and collapsed, leaving the caldera with rims standing over 2000 ft (600 m) high above the lake. *See* CALDERA; VOLCANO.

Sierra Nevada (Fig. 8), while topographically connected to the Cascades, is a different range geologically. It is a great granite fault block uplifted on its east side, which forms a complete drainage divide with no mountain passes for 150 mi (240 km). The granite intruded the area in Jurassic and Cretaceous time, but remnants of Paleozoic sedimentary rocks exist in the north and in the western foothills. The range is 400 mi (640 m) long and up to 75 mi (120 km) wide, with peaks over 14,000 ft (4250 m), including the highest peak in the 48 contiguous states, Mount Whitney (14,495 ft or 4392 m). This enormous asymmetrical block has its highest crest near the east scarp, which rises at least 9000 ft (2730 m) above the desert basins to the east. Along this east face lies Lake Taho (1650 ft or 500 m deep) at an elevation of 6625 ft (2000 m). The western slopes descend gently to the Great Valley of California, but are cut by deep river canyons. The highest parts of the Sierras have been intensely glaciated into horns, cirques, aretes, and spectacular U-shaped valleys, such as the Yosemite Valley. This valley, drained by Merced River, is eroded out of solid granite, forming vertical walls, down which descend waterfalls, including the Upper Yosemite Falls, 1430 ft (430 m) high, and Ribbon Falls, 1600 ft (485 m) high. The granite forms include domes, like the Half Dome, and spires, like the Cathedral Spires. The beautiful trees of this area, the giant sequoia, can be seen in Mariposa Grove, but the most famous are in the Sequoia National Park, near Kings Canyon. These trees are among the largest and oldest living things on Earth. *See* SEQUOIA.

The Coastal Ranges of the United States (Fig. 8) are separated from the Cascade–Sierra Nevada by the Puget Sound Lowland in the north, glaciated by ice lobes descending from the mountains; the Willamette Valley crossed by the lower Columbia River; and the Great Valley of California (Fig. 8) [400 mi or 640 km long and 50 mi or 80 km wide], the biggest downwarped trough in North America completely surrounded by mountains, and deeply filled with coarse alluvium eroded from the Sierras. The Great Valley is drained by the Sacramento River in the north and the San Juaquin River in the south, but the two rivers join and form an inland delta where they enter the San Francisco Bay.

The Coastal Ranges begin in the north with the Olympic Mountains. This domal mountain group, averaging 5000 ft (1510 m) but with peaks nearly 8000 ft (2420 m) in Mount Olympus, and deep valleys, are very rugged because of intense glacial erosion; over 50 glaciers are still present. The oldest rocks are Cretaceous volcanics and metamorphics, but Cenozoic rocks are found on the margins. The Olympics support great forests of Douglas-fir and red cedar.

South of the Olympic Mountains and parallel to the coast lie the Oregon Coast Ranges, a uniform and low belt of mountains, generally below 2000 ft (600 m) elevation, with a gentle western slope and a short, steep east slope. These mountains are composed of Tertiary sedimentary rocks, mostly sandstones, shales, and limestones, with some igneous rocks forming the higher peaks. Their coastline is an alternation of erosional headlands and narrow sandy beaches.

The Oregon Coast Ranges merge into the high Klamath Mountains, which are composed of hard metamorphic Paleozoic and Mesozoic rocks, with old intrusive rocks forming the highest peaks (up to 9000 ft or 2730 m). They are very rugged and in accessible, with steep slopes and narrow valleys cut by the Klamath river system. The highest central part, the Trinity Alps, shows signs of glaciation. Douglas-fir, redwoods, and sugar pine grow in these wild mountains.

South of the Klamath Mountains stretch the 400-mi-long (640-km) California Coast Ranges (Fig. 8). In the north, they are composed mainly of sedimentary and metamorphic Mesozoic rocks; they have low (2000–4000 ft or 600–1210 m), narrow, even-crested ranges, which are closely spaced, and have a distinct northwest-southeast orientation. The highest peaks reach 8000 ft (2420 m). In northwestern California in the Redwood National Park are giant redwood trees, including the tallest trees known, with heights of 367 ft (111 m).

South of San Francisco, the Santa Cruz and Santa Lucia ranges and Santa Clara and Salinas valleys retain the northwest-southeast orientation, but become much larger. The elevations of the ranges increase to 5000–6000 ft (1520–1830 m). The ranges have complex rock composition of Mesozoic and Cenozoic sedimentary and metamorphic rocks, with some Paleozoics.

At the southern end of the Coast Ranges, the mountain orientation changes to east-west, and elevations increase to a maximum of 8826 ft (2675 m) in Mount Pinos. Here begin the Transverse Ranges, which include the huge fault blocks of Precambrian metamorphics, like San Gabriel, San Jacinto, and

Santa Rosa, rising over 10,000 ft (3030 m), and San Bernardino, nearly 11,500 ft (3480 m). The largest basins, filled with Quaternary sediments, include the Los Angeles, San Bernandino, Ventura, and San Fernando. These ranges merge into the Peninsular Range of California, which is composed of Mesozoic igneous rocks, and has a short, steep eastern face, and long gradual western slopes, with distinct marine terraces north of San Diego. The Los Angeles Basin, extending over 1000 mi^2 (2600 km^2), is open to the ocean, but is surrounded by high mountains on all other sides. While beautifully located, it is a perfect trap for air pollutants forming smog, held down by both surface and upper-level temperature inversions.

To the south trends Baja California, a nearly 800-mi-long (1280-km) belt of mountains, with the highest peak over 10,000 ft (3000 m). It is a fault block with a steep east face overlooking the downfaulted Gulf of California. Its west side is covered by lava flows eroded into plateaulike surfaces. A mass of granite forms its southern tip.

Mexico and Central America. The basin-and-range type of terrain of the southwest United States continues into northern Mexico and forms its largest physiographic region, the Mexican Plateau. This huge tilted block stands more than a mile above sea level—from about 4000 ft (1200 m) in the north, it rises to about 8000 ft (2400 m) in the south. The arid northern part, called Mesa del Norte, extends southward to about 22°N latitude, where it merges into higher and more moist southern part, called Mesa Central. The plateau surface consists of north-south-trending ranges rising 500–2500 ft (150–750 m) above adjacent flat desert basins. Limestone and shale rocks dominate the eastern part of the plateau, and volcanic materials the western. The southern half is dominated by recent volcanic activity, and it ends in a volcanic rim, the Transverse Volcanic Axis, which includes snow-capped peaks such as the Popocatepetl (17,900 ft or 5425 m) and Ixtaccihuatl (17,300 ft or 5240 m). These volcanoes formed about 10 million years ago, but some are still active. Immediately north of these volcanoes are the former lake basins now forming fertile plains. One of the largest is the Valley of Mexico, in which Mexico City is located.

The Plateau is surrounded by striking escarpments. On the east is the Sierra Madre Oriental of elongated limestone ranges trending north-south and rising to 8000 ft (2420 m), with peaks reaching 13,000 ft (3940 m). It merges in the south with volcanic snow-capped peaks such as Orizaba (18,700 ft or 5660 m), the highest peak in Middle America. The western rim, the Sierra Madre Occidental, consists mainly of volcanic rocks, cut by deep and beautiful canyons.

The Mexican Plateau is separated from the Southern Mexican Highlands (Sierra Madre del Sur) by a low, hot and dry Balsas Lowland drained by the Balsas River. The average elevators of the Southern Highlands are about 8000 ft (2420 m), with peaks reaching over 10,000 ft (3030 m), and V-shaped valleys cut 4000 ft (1200 m) down, leaving narrow crests between.

To the east of the Southern Highlands lies a lowland, the Isthmus of Tehuantepec, which is considered the divide between North and Central America. Here the Pacific and Gulf coasts are only 125 mi (200 km) apart.

The lowlands of Mexico are the coastal plains. The Gulf Coastal Plain trends Southward for 850 mi from the Rio Grande to the Yucatan Peninsula. It is about 100 mi (160 km) wide in the north, just a few miles wide in the center, and very wide in the Yucatan Peninsula. Barrier beaches, lagoons, and swamps occur along this coast.

The Pacific Coastal Plains are much narrower and more hilly. North-south-trending ridges of granite characterize the northern part, and islands are present offshore. Toward the south, sandbars, lagoons, and deltaic deposits are common.

East of the Isthmus of Tehuantepec begins Central America with its complex physiographic and tectonic regions. This narrow, mountainous isthmus is geologically connected with the large, mountainous islands of the Greater Antilles in the Carribean. They are all characterized by east-west-trending rugged mountain ranges, with deep depressions between them. One such mountain system begins in Mexico and continues in southern Cuba, Puerto Rico, and the Virgin Islands. North of this system, called the Old Antillia, lies the Antillian Foreland, consisting of the Yucatán Peninsula and the Bahama Islands, which are level, low-lying plains underlain by limestone. Northern Yucatán is a limestone surface highly affected by solution, so that there are no surface streams. Rainwater sinks down through sinkholes and flows underground.

Central American mountains are bordered on both sides by active volcanic belts. Along the Pacific, a belt of young volcanoes extends for 800 mi (1280 km) from Mexico to Costa Rica. The youngest volcanoes include over 40 large peaks, which culminate in the Fuego (12,600 ft or 3800 m) in Guatemala.

Costa Rica and Panama are mainly a volcanic chain of mountains extending to South America. Nicaragua is dominated by a major crustal fracture trending northwest-southeast. In the center of this lowland are the two largest lakes of Middle America: Lake Managua and Lake Nicaragua. At the northwestern end of this lowland lies the Gulf of Fonseca, the largest gulf along the Pacific shore of Middle America.

Natural Disasters

North America is subjected to more natural disasters than most continents. For example, along the entire length of the California coast runs a northwest-southeast zone of active crustal fractures, the San Andreas Fault Zone. The continent on the western side of this fault is moving northwestward at an estimated rate of about 2 in. (5 cm) per year, making this region the greatest earthquake hazard zone in North America. *See* EARTHQUAKE; FAULT AND FAULT STRUCTURES.

In addition to the earthquakes, steep-sloped mountainous California is subject to great mudflows when

it rains, as well as enormous forest fires during the drought season.

All coastal western cordillera are subject to volcanic eruptions, like those of Mount Katmai in Alaska and Mount St. Helens in Washington State. *See* VOLCANO.

The interior deserts are subject to dust storms, and flash-floods and mud flows.

The Great Plains are subject to great summer droughts and dust storms, as well as blizzards in the winter, and are annually devastated by violent tornadoes (average of 50 per year in Oklahoma). *See* DUST STORM.

The Central Lowlands have suffered from both prolonged great droughts and excessive precipitation, often leading to enormous floods, such as the extremely destructive hundred-year floods of the Mississippi River system in 1993.

The southeastern part of the United States is subject to frequent tropical hurricanes in which violent winds, enormous rainfall, and towering oceanic waves destroy the coastal areas. *See* HURRICANE.

The northern and northeastern parts of the continent suffer nearly every winter from extremely low temperatures, and hazardous ice and snow storms.

Mexico and Central America are intensly affected by frequent and devastating volcanic eruptions, earthquakes, and tropical hurricanes.

Barbara Zakrzewska Borowiecka

Bibliography. W. W. Atwood, *The Physiographic Provinces of North America*, 1940; J. B. Bird, *The Natural Landscapes of Canada*, 1972; E. M. Bridges, *World Geomorphology*, 1990; N. M. Fenneman, *Physiography of Eastern United States*, 1938; N. M. Fenneman, *Physiography of Western United States*, 1931; W. L. Graf, *Geomorphic Systems of North America*, Geological Society of America, Centennial Special Volume 2, 1987; D. V. Harris, *The Geologic Story of the National Parks and Monuments*, 1980; J. A. Henry and J. Mossa, *Natural Landscapes of the United States*, 5th ed., 1995; C. B. Hunt, *Natural Regions of the United States and Canada*, 1974; T. L. McKnight, *Physical Geography: A Landscape Appreciation*, 1992; A. N. Strahler and A. H. Strahler, *Elements of Physical Geography*, 4t ed., 1990; W. D. Thornbury, *Regional Geomorphology of the United States*, 1965; R. C. West and J. P. Augelli, *Middle America, Its Lands and People*, 3d ed., 1989.

North Pole

That end of the Earth's axis which points toward the North Star, Polaris (Alpha Ursae Minoris). It is the geographical pole located at 90° N latitude where all meridians converge, and should not be confused with the north magnetic pole, which is in the Canadian Archipelago. The North Pole's location falls near the center of the Arctic Sea. Being at the end of the Earth's axis, which is inclined $23^1/_2$° (23°.45) from a line perpendicular to the plane of the ecliptic, the North Pole has phenomena unlike any other place except the South Pole. For 6 months the Sun does not appear above the horizon, and for 6 months it does not go below the horizon. During this latter period, March 21–September 23, the Sun makes a very gradual spiral around the horizon, gaining altitude until June 21; then it starts to lose altitude until it disappears below the horizon after September 23. The Sun's highest altitude is 23 1/2°. As there is a long period (about 7 weeks) of continuous twilight before March 21 and after September 23, the period of light is considerably longer than the period of darkness.

There is no natural way to determine local Sun time because there is no noon position of the Sun, and all shadows, at all times, fall to the south, the only direction possible from the North Pole. *See* MATHEMATICAL GEOGRAPHY; SOUTH POLE; TERRESTRIAL COORDINATE SYSTEM.

Van H. English

Bibliography. A. H. Strahler and A. Strahler, *Introducing Physical Geography*, 4th ed., 2005.

North Sea

A flooded portion of the northwest continental margin of Europe occupying an area of over 200,000 mi^2 (500,000 km^2). The North Sea has extensive marine fisheries and important offshore oil and gas reserves. In the south, its depth is less than 150 ft (50 m), but north of 58° it deepens gradually to 600 ft (200 m) at the top of the continental slope. A band of deep water down to 1200 ft (400 m) extends around the south and west coast of Norway and is known as the Norwegian Trench (see **illus.**).

Geologic history. The area of the North Sea overlies a large sedimentary basin that has gradually been accumulating sediments since Permian time (about 2.5×10^8 years ago). Since then, gradual subsidence has been occurring between the stable massifs of Scotland, Scandinavia, and Central Europe where the old basement rocks lie at or near the surface. *See* MASSIF; PERMIAN.

The most recent geological event (about 2 million years ago) to affect the area was the Pleistocene glaciation when ice from the Scandinavian ice caps extended across the North Sea to the British coast. Many of the present topographical features of the North Sea date from this period. *See* PLEISTOCENE.

Currents and water level. Fluctuations in currents and water level are predominantly due to the semidiurnal tide which propagates into the North Sea from the adjacent Atlantic Ocean. The largest tidal amplitudes (6–18 ft or 2–5 m) and largest currents (1.5–6 ft/s or 0.5–2 m/s) are in the west and south. Adjacent to the Norwegian coast, tidal amplitudes and currents are low (<3 ft or 1 m; 0.3 ft/s or 0.1 m/s). Numerical modeling of the North Sea tide has demonstrated the existence of amphidromic points (no change in tidal amplitude) off southeastern England, east of Denmark, and close to the southern coast of Norway. Storms can considerably modify mean levels and currents, especially in the southern North Sea where storm surges can raise the level by more than 6 ft (2 m), and real-time storm warning

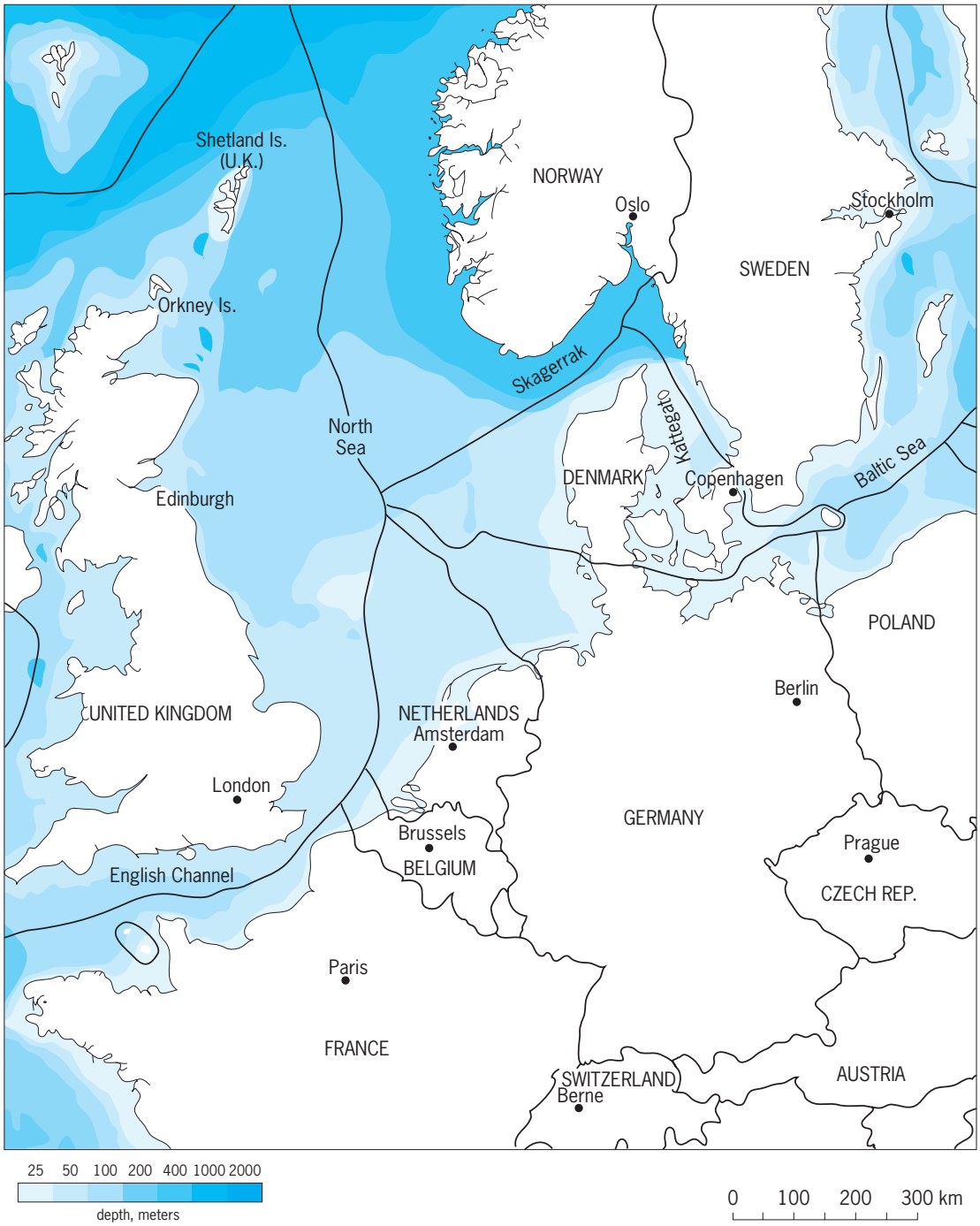

North Sea, showing depth. (*Adapted from The North Sea Secretariat, Oslo, Norway*)

schemes to predict their occurrence are in operation. *See* STORM SURGE; TIDE.

The nontidal residual current circulation of the southern North Sea is mainly determined by wind velocity, but in the north, well-defined non-wind-driven currents have been identified, especially in the summer. Two of these currents bring in water from outside the North Sea; one flows through the channel between Orkney and Shetland (the Fair Isle current), and the other follows the continental slope north of Shetland and merges with the Fair Isle current southwest of Norway before entering the Skagerrak. The north-flowing Norwegian coastal current provides the exit route for North Sea waters, and is formed from the waters of these two major inflows and from other much smaller inputs such as river runoff, the English Channel, and the Baltic Sea.

The transport of the Norwegian coastal current is typically 5.3–8.8 $\times 10^7$ ft^3/s (1.5–2.5 $\times 10^6$ m^3/s), the smaller figure being associated with winter. The Fair Isle current has a transport of about 1.1 $\times 10^7$ ft^3/s (300,000 m^3/s), the Norwegian Trench inflow of Atlantic water making up the balance. In late winter the Fair Isle current is a weakly defined feature, especially in the west. *See* ATLANTIC OCEAN; OCEAN CIRCULATION.

Water properties and biology. Average salinities have very small seasonal variations and are typically 35, increasing to 35.3 in the north, adjacent to the Atlantic Ocean, and also occasionally off the east end of the English Channel. In coastal regions, and the southern North Sea generally, salinities are less than 35 but are mostly greater than 30. Vertical variations in salinity are minor except adjacent to the Norwegian coast, within the coastal current, where salinities in the upper 120 ft (40 m) are some 2–3 less than midwater and near-bottom salinities.

Minimum surface temperatures occur in February and range from 45°F (7°C) in the northwest to 36°F (2°C) adjacent to the Danish and German coasts. In the summer, this gradient is reversed, with August temperatures of about 63°F (17°C) in the south ranging to only 52.7°F (11.5°C) in the northwest. The near-bottom waters of the central and northern parts of the North Sea maintain winter temperatures until late October, when convective overturn associated with autumnal cooling and strong winds mixes the water column. Before this overturn occurs, however, the bottom waters become slightly depleted in oxygen; otherwise, oxygen levels are near saturation level throughout the North Sea. Levels of inorganic nutrients such as phosphate, silicate, and nitrate reflect the high fractional content of nutrient-rich Atlantic waters, but summer surface values are usually depleted because of the spring outburst of phytoplankton. This outburst occurs in March in the permanently stable Norwegian coastal waters, but not until late May and June in the tidally mixed coastal waters of the British Isles.

During late autumn and winter, there is deep mixing within the water column and nutrient concentrations are high. Nevertheless, phytoplankton growth is extremely low, mainly due to limited daylight. During spring and summer, the water column stratifies and the phytoplankton are retained within a shallow wind-mixed surface layer. When this stratification first develops, conditions within this shallow surface layer are ideal for phytoplankton growth as nutrients are abundant and there is plenty of light. This can result in the development of algal blooms. These blooms are usually short-lived, however, as the nutrients in the mixed surface layer are soon used up. Blooms generally occur in the spring and autumn, that is, at the beginning and end of stratification. *See* PHYTOPLANKTON.

There is a rich diversity of zooplankton within the North Sea. Copepods are of particular importance in the food web, with *Calanus finmarchicus* being the characteristic species of the northern North Sea, while smaller copepods (for example, *Temora longicornis*) are most important elsewhere. Some species of zooplankton enter the North Sea from the Atlantic Ocean during the summer and autumn. In general, they arrive in small numbers, but *Salpa fusiformis* often forms dense swarms and may reduce the amount of food available for the indigeneous plankton. *See* COPEPODA; MARINE ECOLOGY; SEAWATER; ZOOPLANKTON.

Marine contamination. Problems resulting from the input of anthropogenically derived nutrients are widespread in coastal areas throughout the North Sea, in various estuaries and fiords, in the Wadden Sea, and in the German Bight. They contribute to the development of exceptional algal blooms and exceptional densities of macroalgae, to changes in the composition of the phytoplankton and zoobenthos communities, and to intermittent oxygen depletion in the bottom water and the subsequent death of benthic fauna. Ninety percent or more of the freshwater-borne nutrients to the North Sea are from a few main rivers (over half from the Rhine).

Many harmful algal blooms have been observed in the North Sea, ranging from small, localized blooms to blooms covering vast areas, such as the *Phaeocystis* bloom that has been observed intermittently in the Wadden Sea since the early 1970s. *Phaeocystis* and *Coscinodiscus* blooms recur on the southeastern and eastern North Sea coasts; *Coscinodiscus* is harmful to fish because it causes increased mucous production, which occasionally restricts gill function. *Gyrodinium aureolum*, first noticed in northern European waters in 1966, is now one of the most common dinoflagellates in the autumn. There have been frequent blooms since 1981, and these have often caused significant mortalities of farmed fish. A number of these blooms appear to have started in the open waters of the Skagerrak and Kattegat and have then spread northward with the coastal current. In 1988 an extremely large bloom of the small flagellate *Chrysochromulina polylepis* affected the Kattegat, Skagerrak, and Sound, covering an area of 75000 km^2, and produced a toxin which killed or affected various marine organisms, including mussels and wild and cultured fish.

There are many studies concerning the levels of inorganic and organic contaminants in fish from the North Sea. Concentrations in fish are often elevated in areas with fine-grained sediments, particularly in bays and estuaries close to industrialized areas, and depositional areas such as the Dogger Bank, the Oyster Ground, the Wadden Sea, the German Bight, and the Norwegian Trench.

In recent years, general decreases in metal concentrations have been reported. Examples include cadmium, lead, and copper in the region to the north of the Rhine-Meuse and in the Wadden Sea. Metal concentrations in sediments from estuaries tend to be higher than those from coastal areas; in some estuaries, however, there is evidence of decreasing concentrations, such as cadmium and lead in the Scheldt.

Fisheries. There is a wide range of fish stocks in the North Sea and adjacent waters and, in terms of species exploited by commercial fisheries, they constitute the richest area in the northeast Atlantic. The commercially important stocks exploited for human consumption include cod, haddock, whiting, pollock, plaice, sole, herring, mackerel, lobster, prawn, and brown shrimp (*Crangon crangon*). A number of stocks are used for fishmeal and oil; these stocks include sand eel, Norway pout, blue whiting, and sprat. Beyond these stocks, there are landings from a

variety of demersal (found near the sea floor) species such as turbot, anglerfish, gurnards, lemon sole, rays, and sharks.

A basic problem is that the North Sea has been exposed to intensive fishing for more than a century. This has implications for both the state of the fish stocks and the environment of the North Sea. As a result, the Intermediate Ministerial Meeting of the 5th International Conference on the Protection of the North Sea called for "further integration of fisheries and environmental protection, conservation and management measures, drawing upon the development and application of an ecosystem approach" and "integration of environmental objectives into fisheries policy." *See* MARINE FISHERIES.

Petroleum exploration. The first major discovery of oil in the North Sea was made in 1969 at Ekofisk. The major developments of the offshore oil industry have been in the northern North Sea, in the United Kingdom and Norwegian sectors. Gas deposits have been exploited mainly in the shallower southern regions in the United Kingdom, Dutch, and Danish sectors, as well as in Norwegian waters. From 1990 to 1996 the number of offshore platforms and oil production almost doubled, primarily reflecting increased activity in the Norwegian and United Kingdom sectors.

Offshore installations are significant sources of oil contamination in the North Sea. Overall, inputs of oil have reduced from a maximum of some 28,000 tons in 1985 to about 10,000 tons in 1997, but the contribution from produced water has increased progressively as oil fields have matured and the number of installations has increased. *See* OIL AND GAS, OFFSHORE.

H. D. Dooley

Bibliography. J. Andersen and T. Niilonen (eds.), *Progress Report. 4th International Conference on the Protection of the North Sea*, Esbjerg, Denmark, June 8–9, 1995; 5th North Sea Conference Secretariat, *Intermediate Ministerial Meeting on the Integration of Fisheries and Environmental Issues*, Bergen, Norway, March 13–14, 1997; J. S. Gray et al., Managing the environmental effects of the Norwegian oil and gas industry: From conflict to consensus, *Mar. Pollut. Bull.*, 38:525–530, 1999; NSTF, Oslo and Paris Commissions, *North Sea Quality Status Report*, Olsen and Olsen, Fredensborg, Denmark, 1993; G. Radach and H. J. Lenhart, Nutrient dynamics in the North Sea: Fluxes and budgets in the water column derived from ERSEM, *Neth. J. Sea Res.*, 33:301–335, 1995; Report of the ICES Advisory Committee on Fishery Management, *ICES Coop. Res. Rep.*, no. 229, 1999; E. Svendsen and A. K. Magnusson, Climatic variability in the North Sea., *ICES Mar. Sci. Symp.*, 195:144–158, 1992.

Nose

The nasal cavities and the structures surrounding and associated with them. The nose functions primarily as the organ of smell and in most tetrapods also assumes a respiratory function, forming the anterior end of the air passage through which air is drawn in and in which it is warmed and moistened.

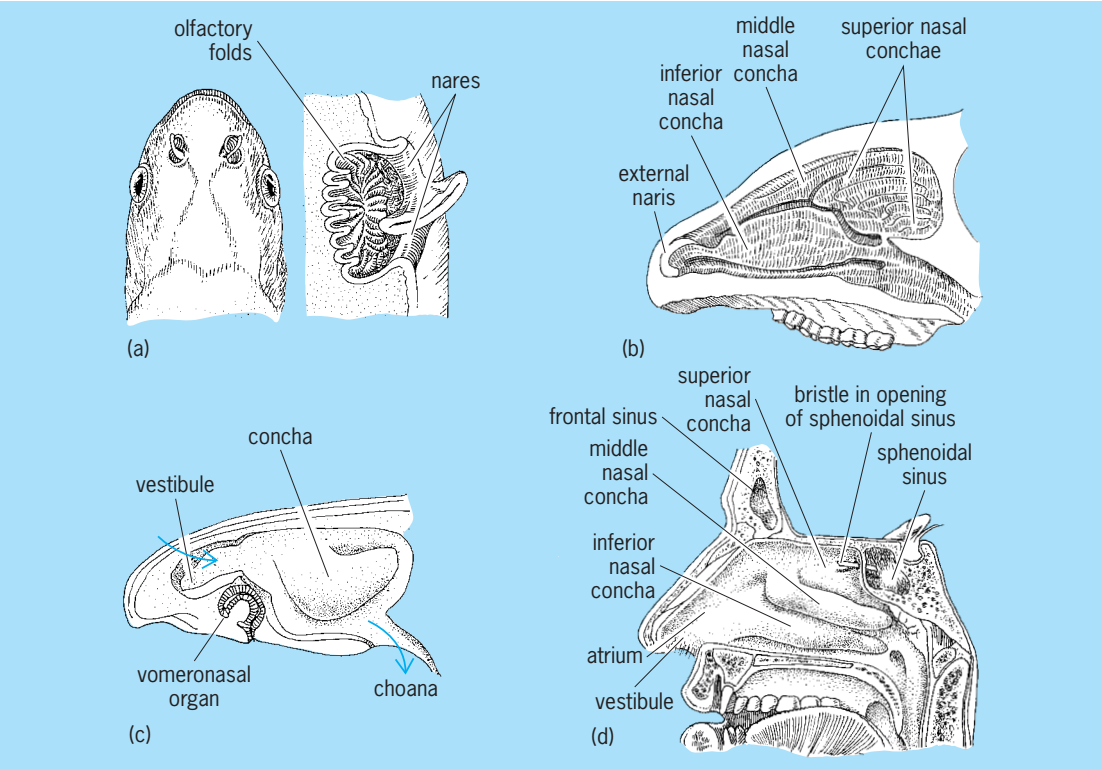

Nose in various vertebrates. (*a*) Minnow. (*b*) Sheep (*after H. Giersberg and P. Rietschel, Vergleichende Anatomie der Wirbeltiere, Erster Band, Fischer, 1967*). (*c*) Lizard (*after W. W. Ballard, Comparative Anatomy and Embryology, Ronald, 1964*). (*d*) Human (*after W. J. Hamilton et al., Textbook of Human Anatomy, Macmillan, 1956*).

In most fishes the nasal cavities are a pair of pits, one on each side of the snout. Although the cavities frequently have two external openings, they are rarely connected with the mouth. In lampreys and hagfishes there is only a single median nasal cavity, and in a few fishes there are choanae, or openings between the nasal and oral cavities. The nose contains folds which increase the surface area covered by sensory epithelium (**illus.** *a*).

In tetrapods each nasal cavity has two openings, one to the outside (external naris) and one to the oral cavity (internal naris or choana). In some forms the cavity is a simple sac, but commonly projections develop on the lateral wall (illus. *b, c*, and *d*); these are termed conchae or, in mammals, turbinals. Reptiles have one to three conchae, and most birds have three. However, in mammals the conchae may be more numerous. They are relatively few and small in humans, whose sense of smell is not very well developed.

In humans the nasal cavities are triangular openings that pass from the external nares back to the dorsal part of the pharynx (illus. *d*). The lateral walls are composed principally of portions of the ethmoid and sphenoid bones and projections of three turbinate bones, or conchae, on each side. The floor of the nose is formed by the palate, which is also the roof of the mouth. The nasal cavities are lined with respiratory epithelium, which also lines the paranasal sinuses. The latter are cavities in the frontal, ethmoid, sphenoid, and maxillary bones which communicate with the nasal passages. The external nose consists of the two nasal bones that form the bony bridge and two pairs of lower nasal cartilages. These together with the tightly adherent skin determine the individual shape and size of the human nose.

Numerous blood vessels, lymphatics, and nerves supply and drain both the external and the internal portions of the nose. *See* OLFACTION.

Thomas S. Parsons

Nose cone

The forward portion of a spacecraft that is designed for atmospheric entry. Nose cones are utilized for intercontinental ballistic missiles and crewed spacecraft such as Apollo and space shuttles. The nose cone is required to withstand heating encountered during atmospheric entry, maintain the structural integrity of the spacecraft, prevent overheating of the payload, and usually maintain the aerodynamic characteristics of the spacecraft. The requirements for a safe entry pose many challenging problems to the designer. Nature provides a dramatic demonstration of the scope of the problem in the form of so-called shooting stars. These bodies are meteoroids which are heated to incandescence as they enter the Earth's atmosphere and which, with few exceptions, destroy themselves prior to Earth impact. *See* ATMOSPHERIC ENTRY; METEOR.

Amount of heating. An Apollo spacecraft returning from the Moon enters the Earth's atmosphere at about 7 mi/s (11 km/s). Braking action must reduce this velocity V to practically zero at Earth impact; that is, most of the initial kinetic energy of the spacecraft must be absorbed by the atmosphere prior to impact. The spacecraft's initial kinetic energy, per unit of spacecraft mass, is $V^2/2$, which is equivalent to a significant thermal energy. A calculation shows each unit of mass has the energy capacity to melt 250 times as much steel. Hypothetically, then, if the spacecraft is made of steel, it possesses enough energy to destroy itself 250 times over. As the spacecraft decelerates in the Earth's atmosphere, it experiences frictional heating in the boundary layer at its surface, and the nose cone is also subjected to heat from gases that are at elevated temperatures as a result of being decelerated by the bow shock wave. The designer must prevent all but a small fraction of the energy in the gas surrounding the nose cone from reaching it in the form of heat; that is, the designer must prevent the spacecraft from absorbing all its kinetic energy. *See* AEROTHERMODYNAMICS; BOUNDARY-LAYER FLOW; SHOCK WAVE.

Heating rate. The amount of heat transferred into the nose cone depends on its shape and the materials from which it is made. The bow shock wave heats the gases behind it. The heat reaches the nose cone in the form of convection and radiation through the boundary layer adjacent to the surface. *See* HEAT TRANSFER.

For a slender body with a sharp nose (**Fig. 1**), the bow shock is relatively weak and lies fairly close to the body, resulting in small wave drag. A proportionately small amount of air is heated, and the friction drag is high. For a blunt body (**Fig. 2**) the bow shock is much stronger, extends farther away from the vehicle sides, creates larger wave drag, and heats a considerably greater gas volume than its sharp, thin counterpart. The heat applied to the spacecraft in the latter case is markedly less than the former case, because a greater fraction is absorbed in heating the atmosphere. Many spacecraft, including the Apollo, have employed the blunt-body concept to survive the entry maneuver when convection is the primary mode of heat transfer. However, the blunt body is not a panacea for all designs, since radiation from the incandescent gas behind the stronger shock wave of the blunt body becomes a significant mode of heat transfer as the entry velocity increases.

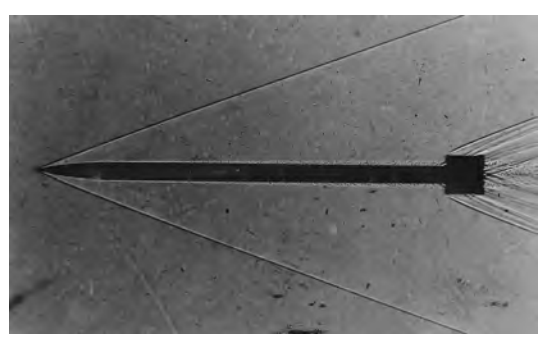

Fig. 1. Flow about slender body.

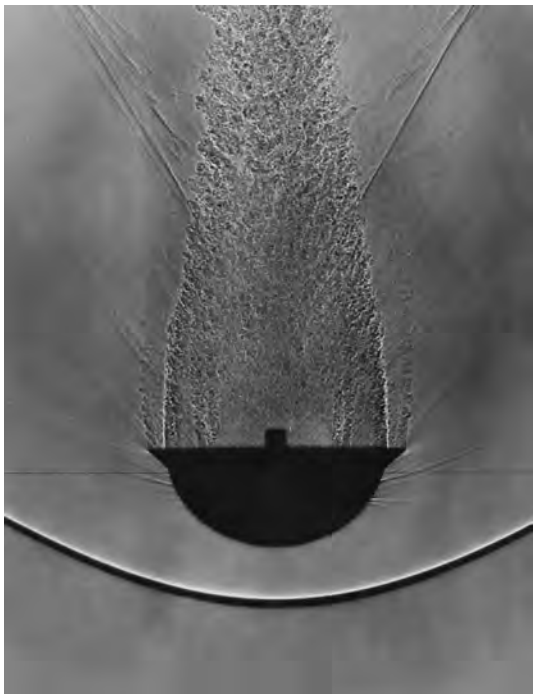

Fig. 2. Flow about blunt body.

Real-gas effects. At high entry velocities the translational energy of the gas behind the bow shock as measured by its temperature becomes comparable with the energies associated with various molecular and atomic processes, such as dissociation and ionization. Under these conditions the energy that is involved in these processes must be taken into account when calculating the flow field. The thermodynamic and transport properties of the gas differ from those of a perfect gas, and these changes are significant. For example, the distance of the bow shock from the body depends on the specific-heat ratio, which can change significantly as a result of real-gas effects. At high enough velocities the number of free electrons produced becomes appreciable and the gas becomes a conductor of electricity. The presence of free electrons in the shock layer alters the dielectric properties of the medium and can have large effects on radio communication to the vehicle. The design of the entry body requires the solution of serious problems associated with real-gas effects, such as the prediction of shock shapes, flow fields, and pressure distributions for high velocities and Mach numbers. *See* GAS DYNAMICS.

Surface thermal protection. Even for a properly designed shape, it is inevitable that some fraction of the spacecraft's initial kinetic energy will finally reach the nose cone in the form of heat. The design of the heat shield for the nose cone is a complex procedure, which is highly dependent on the heating level. There are a variety of surface-protection or cooling systems which have been used. Generally these systems consist of heat sinks of various types: the absorption of heat by virtue of a material's sensible heat capacity, latent heat capacity, or chemical heat capacity. That is, heat absorption is accomplished by a temperature rise, a phase change, or a chemical reaction. Aerodynamic lift is employed by reusable space shuttles to lower heating rates so that the nose cone material can radiate away much of the incident heating.

Ablation is used to provide surface protection. The designer can divert heat from the spacecraft by allowing the nose cone's outer layer of material to melt, vaporize, or sublime. Materials which vaporize or sublime have the advantage of a blocking action which, by thickening and cooling the boundary layer, prevents a sizable amount of heat from being convected from the hot gas stream to the spacecraft. Additionally, some materials form a char layer of porous graphite refractory residue as the outer layers decompose. The char's outer surface can reach high temperatures, so that it radiates sizable amounts of energy away from the spacecraft and thereby decreases the heat transfer. While large ablation rates provide excellent thermal protection, the resulting change in profile due to surface recession can adversely change the aerodynamic characteristics of the spacecraft. The designer must account for this change. In addition, the designer must provide adequate strength to prevent mechanical erosion of the char by aerodynamic shear stresses. *See* SPACE FLIGHT; SPACE SHUTTLE; SPACECRAFT STRUCTURE.

Philip R. Nachtsheim

Notacanthoidei

A suborder of Albuliformes consisting of two families, Halosuridae (halosaurs) and Notacanthidae (spiny eels). They are also known as Lyopomi and Heteromi, respectively. The body of these fishes is elongate and tapers posteriorly. The caudal fin is absent or essentially absent; pectoral fins are high on the body; pelvic fins are abdominal; and the anal fin is long (see **illustration**). There is no duct to the swim bladder (physoclistous); the orbitosphenoid, pterosphenoid, intercalary, and basisphenoid bones are absent; the posttemporal is simple or ligamentous; the transverse processes are not suturally joined to vertebral centra; and there is no mesocoracoid arch. Some species have photophores (light emitting organs), a characteristic of many deep-sea fishes.

The halosaurs differ from spiny eels in lacking spines in the dorsal fin and having gill membranes that are completely separate rather than joined or partly joined.

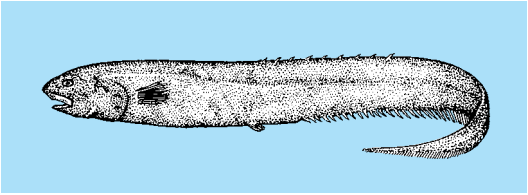

Spiny eel (*Notacanthus nasus*). (*After D. S. Jordan and B. W. Evermann, The Fishes of North and Middle America, U.S. Nat. Mus. Bull. 47, 1900*)

The Notacanthoidei is a small group with a history extending back to the Upper Cretaceous; it includes 2 families, 6 Recent genera, and about 25 species. The species inhabit deep seas worldwide, in waters between 125 and 4900 m (413 and 16,1700 ft). They are like true eels (Anguilliformes) in lacking a firm suspension of the pectoral girdle from the skull, but some have fin spines similar to those of perciform fishes. Of much interest was the discovery of a pelagic leptocephalous elongate and flattened from side to side larval stage among the halosaurs. The larvae are strikingly similar to those of true eels, testifying to a close relationship, a conclusion that is strengthened by the similarities in the swim bladders of these groups. The leptocephalous larva is characteristic of teleost fishes in the subdivision Elopomorpha. *See* ALBULIFORMES; ANGUILLIFORMES; ELOPOMORPHA; PHOTOPHORE GLAND. Reeve M. Bailey; Herbert Boschung

Bibliography. P. H. Greenwood, Notes on the anatomy and classification of elopomorph fishes, *Bull. Brit. Mus. Nat. Hist.* (*Zool.*), 32(4): 65–102, 1977; N. B. Marshall, Observations on the Heteromi, an order of teleost fishes, *Bull. Brit. Mus.* (*Nat. Hist.*), vol. 9, no. 6, 1962; J. S. Nelson, *Fishes of the World*, 3d ed., 1994.

Nothosauria

An extinct clade (evolutionary lineage) of eosauropterygian reptiles (that is, a suborder of Eosauropterygia) that also includes pachypleurosaurs (suborder Pachypleurosauria) and pistosaurs (suborder Pistosauroidea). Representatives of these groups are known from uppermost Lower Triassic to Upper Triassic marine deposits in Europe, the Near and Middle East, Northern Africa, China, and North America. Triassic nothosaurs and their relatives are typically restricted to near-shore habitats or shallow epicontinental seas. Some derived members of the pistosaurs constitute the sister group (closest relatives) to plesiosaurs, a clade that greatly diversified during the Jurassic and Cretaceous. *See* PLACODONTIA; PLESIOSAURIA; SAUROPTERYGIA.

Morphology. Nothosaurs in the broad sense (that is, including pachypleurosaurs and pistosaurs) range approximately 3–13 ft (1–4 m) in length (see **illustration**). Although their limbs are reduced in size relative to those of their terrestrial predecessors, they are not highly modified for aquatic propulsion. The neck is long, with at least 15 cervical vertebrae. Like aquatic lizards and crocodiles, early nothosaurs probably swam primarily by lateral undulation of the trunk and tail, with the limbs held to the side to reduce drag. Advanced nothosaurs show modifications of the front limbs to act as paddles; the rear limbs are somewhat reduced. In many genera, the ribs are expanded and heavily ossified (a condition termed pachyostosis) to reduce buoyancy. Some genera have as many as six pairs of sacral ribs.

Nothosaurs (and placodonts) are distinguished from other aquatic reptiles by the closure of the openings in the palate that occur in more primitive reptiles, and by the covering of the base of the braincase. Secondarily developed openings in the palate indicate close relationships of advanced pistosaurs with plesiosaurs. The underside of the shoulder girdle is characterized by a wide gap between the base of the scapulae and the coracoids, behind a stout transverse bar formed by the interclavicle and the blades of the clavicles. As in plesiosaurs (and placodonts), part of the scapula is located on the outside of the clavicle, a reversal of the relationship of these bones in other reptiles.

Evolutionary trends. Interesting evolutionary trends are observed in the skull structure of nothosaurs. The relatively small pachypleurosaurs had a short and broad skull, and the jaws were furnished with small, vertically oriented peglike teeth. Rapid opening of the mouth and simultaneous depression of the throat skeleton would have created a suction effect to capture prey, which is similar to the method used by aquatic turtles today. More advanced members of the nothosaurs evolved a narrow and strongly depressed skull with a much-elongated rostrum (elongated shout) that was furnished with strongly procumbent, fanglike teeth in its anterior part. These pincer-jaws allowed the capture of fish and cephalopod prey with rapid lateral strikes of the head. The very low profile of the skull reduced drag as the head struck sideways, but necessitated important changes in the arrangement of the jaw-closing musculature. *Nothosaurus giganteus* was one of the top carnivores of Triassic near-shore habitats, with a documented skull length of nearly 3 ft (up to 85 cm).

Phylogeny. Primitive nothosaurs resemble primitive lepidosauromorphs in the pattern of the skull roof but have lost the lower temporal bar. However, the precise relationships of Sauropterygia (including nothosaurs and placodonts) within evolutionarily advanced diapsid reptiles (those branches that include the tuatara and squamates on one side

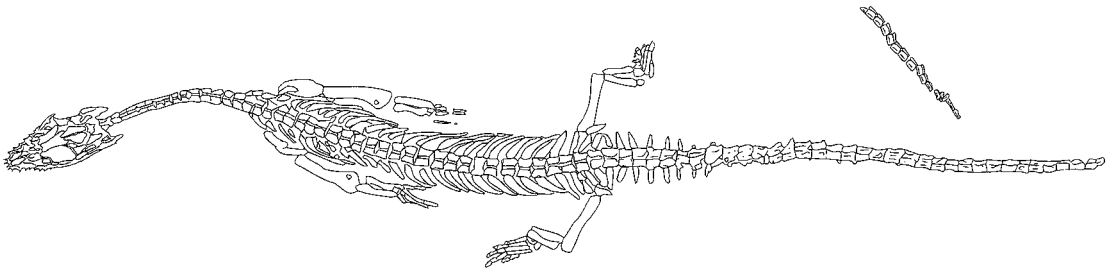

***Neusticosaurus*, a small, primitive nothosaur (pachypleurosaur) from the Middle Triassic of the Alps. The specimen is about 25 cm long, excluding detached portion.**

and crocodiles, dinosaurs, and birds on the other) remain incompletely understood. *See* DIAPSIDA; REPTILIA. Robert L. Carroll; Olivier C. Rieppel

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988; O. Rieppel, *Encyclopedia of Paleoherpetology*, Pt. 12A: *Sauropterygia I. Placodontgia, Pachypleurosauria, Nothosauroidea, Pistosauroidea*, F. Pfeil, Munich, 2000.

Notomyotida

A small and distinctive order of valvatacean Asteroidea with long, straight-sided arms and relatively small disks. The skeleton is differentiated into marginals, actinals, and abactinals, and there is a clear calycinal ring. Both supra- and inframarginals are present and are offset or alternate rather than vertically aligned as they are in other asteroids. Intermarginal ossicles are not present. On the oral surface the longest actinal row lies adjacent to the ambulacral groove. Digestive and reproductive organs are confined to the disk and most proximal parts of the arms and do not extend well into the arms as in other valvataceans. Dorsal longitudinal muscle bands lie free within the arms and are anchored into the body wall only at their distal and proximal ends.

Only a single family, Benthopectinidae, is included. Extant benthopectinids are specialist deep-water suspension feeders that live at depths usually between 1000 and 10,000 ft (300 and 3000 m). There are eight genera and 80 species, distributed in all the world's major oceans. Four fossil species can be attributed to this group, the oldest being *Plesiastropecten hallovensis* from near the base of the Lower Jurassic in Switzerland. *See* ASTEROIDEA; ECHINODERMATA.

Andrew B. Smith

Notostraca

An order of branchiopod crustaceans, sometimes called tadpole shrimps. Generally they range from about 20 mm (0.4 in.) to (exceptionally) about 90 mm (3.5 in.) in length. The multisegmented trunk, up to 44 segments in some species, is elongate and cylindrical. The anterior part of the body is covered dorsally by a domed carapace beneath which lie the paired sessile eyes. Each of the first 11 trunk segments bears a pair of limbs, while a varying number of more posterior segments bear up to six pairs of smaller limbs per segment—a very unusual situation. At least four, often more, posterior segments are limbless. The trunk terminates in a telson that bears a pair of slender, segmented, caudal filaments. Each trunk limb bears several endites. These are mostly used for handling the food and become progressively reduced on the smaller posterior limbs, but are drawn out into long, filiform sensory structures on the first pair. In the female, the eleventh pair bears a bowl-shaped pouch with a hinged lid that acts as a temporary receptacle for the eggs that are then shed, or are stuck to surfaces.

The antennules and antennae are minute. The mandibles are massive biting structures that, unlike the rolling crushing mandibles of most branchiopods, are capable of both abduction and adduction and therefore of true biting. They are assisted by robust maxillules. Notostracans feed on a variety of small organisms that are seized by the trunk limbs and passed forward to the mouthparts, but they also collect and eat detritus with the same limbs.

Some species are bisexual, but in some parts of their range some are self-fertilizing hermaphrodites. The highly resistant eggs, which can withstand desiccation if necessary, hatch as nauplii, but in some cases this stage is transient, molting almost at once to a more advanced stage. The two extant genera, *Triops* and *Lepidurus*, are essentially worldwide in distribution and occur mostly in temporary waters. *See* BRANCHIOPODA. Geoffrey Fryer

Notoungulata

An order of dominant, hoofed herbivores of the Cenozoic of South America that are abundantly represented in Paleocene through Pleistocene nonmarine sedimentary rocks of that continent. There are also isolated occurrences in the Paleocene of Central Asia and early Eocene of North America. Diverging from a primitive condylarth ancestry at an early date, they radiated into a wide diversity of forms, some of which were convergent with Northern Hemisphere ungulates.

Notoungulates were characterized by a skull with an expanded temporal region due to the presence of a large sinus in the squamosal (**Fig. 1**) and no postorbital bar. The dentition is primitive with full eutherian formula and a tendency to retain a closed tooth row, although in some groups the median incisors are enlarged, and a gap between the incisor row and cheek teeth is developed by reduction of the posterior incisors, canines, and anterior premolars. There was always a complete molar row and incomplete molarization of the posterior premolars. The teeth were low- to high-crowned, some groups

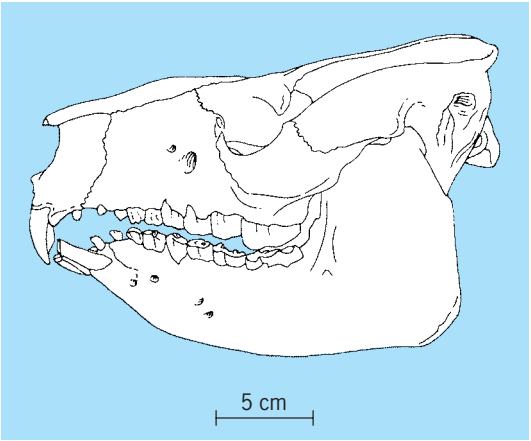

Fig. 1. Skull and jaw of *Adinotherium ovinum*, an early Miocene toxodontid notoungulate from the Santa Cruz formation of Patagonia, Argentina.

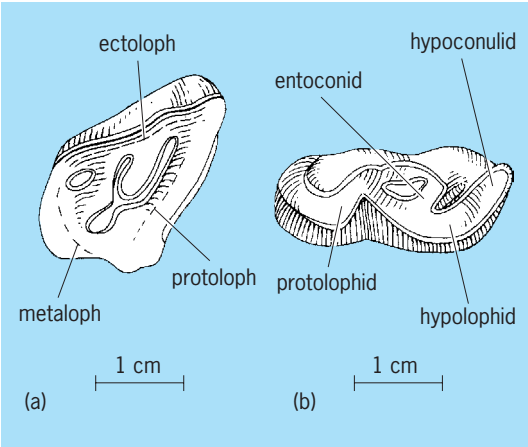

Fig. 2. Examples of molars of *Adinotherium ovinum* from Santa Cruz formation of Patagonia, Argentina. (*a*) Upper molar. (*b*) Lower molar.

developing high-crowned ever-growing teeth. The cusps were joined by ridges in the upper and lower molars (lophodonty) even in the earliest forms. Characteristically, the upper molars had a straight ectoloph (**Fig. 2***a*), oblique protoloph, and transverse metaloph; the median valley enclosed by the three lophs usually carried accessory cusps. The lower molars typically were crescentic, the protolophid and hypolophid resembling the Perissodactyla. The entoconid was isolated (Fig. 2*b*) or connected transversely to the hypoconid rather than to the hypoconulid. The feet were primitive, with five toes (three or two in some advanced forms), and the weight was borne mainly by the third digit. *See* DENTITION.

Notoungulates are represented in the earliest known mammalian faunas of South America (late Paleocene) by a diversified archaic stock (suborder Notioprogonia), at that time already a specialized sideline rather than a truly ancestral group, and by the earliest members of a central stock, suborder Toxodontia, which gave rise to most of the middle and late Tertiary forms. In addition, a third suborder, the rodentlike Typotheria, appeared in the late Paleocene. By early Eocene time a fourth suborder, the rodentlike Hegetotheria, had appeared. Surprisingly, at about the same time in Central Asia the most abundant mammalian herbivores were members of the Notioprogonia. Two specimens record the presence of this group in the early Eocene of North America. A late Mesozoic origin, either in Asia or South America, seems possible; however, knowledge of late Mesozoic and Paleocene faunas of the world is too incomplete at present to rule out other possibilities.

In late Paleocene time South America was isolated from North America, and from that time until the Pliocene notoungulate evolution proceeded unharassed by competition from the herbivores of the Northern Hemisphere. Toward the close of the Eocene the Notoungulata displayed their greatest diversity, with members of all four suborders being present. Notioprogonians did not survive the Eocene. In the middle and late Tertiary the toxodonts, typotheres, and hegetotheres tended to specialize along rather restricted lines; a decrease in diversity within the order resulted. The late Pliocene and Pleistocene emigrant ungulates from North America eventually replaced the large, hippopotamuslike toxodontids, bear-sized typotheres, and rabbitlike hegetotheres in South America. Toxodontids spread north into Central America in the Pleistocene. *See* EUTHERIA. Richard H. Tedford

Bibliography. M. J. Benton, *Vertebrate Paleontology*, 1991; R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988.

Nova

The sudden brightening of a previously inconspicuous star. The name, short for nova stella (new star), formerly included objects now classified as supernovae and as other kinds of cataclysmic variables. Classical novae now include only those events where the energy source is hydrogen fusion (burning) on the surface of a white dwarf in a close binary system and the white dwarf is not destroyed in the process. *See* SUPERNOVA.

A handful of novae are discovered each year in the Milky Way Galaxy, and the total rate is probably 20–50 per year. A comparable number are found in other, nearby galaxies. The system consists of a normal, hydrogen-burning star in a close orbit (periods of a few days or less) around a white dwarf or degenerate star. A stream of gas flows from the normal star into a disk around the white dwarf and then accretes onto its surface. Hydrogen gradually builds up there until it is hot and dense enough for nuclear burning, normally with carbon, oxygen, neon, or magnesium from the white dwarf itself acting as a catalyst. Any nuclear fuel ignited under degenerate conditions explodes, because energy released does not cause the gas to expand, so temperature rises rapidly. *See* BINARY STAR; WHITE DWARF STAR.

The flow occurs at a rate of about 10^{-9} solar mass per year, and 10^{-5} to 10^{-4} solar masses must accumulate to trigger an explosion. Thus novae will recur every 10^4–10^5 years. Only part of the accumulated hydrogen burns, but it is all expelled, at speeds of several thousand kilometers per second, together with products of the nuclear reactions. These include rare isotopes such as carbon-13, nitrogen-15, and aluminum-26 (which decays to magnesium-26), and novae are believed to be significant contributors to the gal-actic supply of these isotopes. *See* ISOTOPE; NUCLEOSYNTHESIS.

Novae brighten in a few days and fade in months to years. The peak brightness is more than 100 times the solar luminosity, and the total energy release more than 10^{45} ergs (10^{38} joules). *See* CATACLYSMIC VARIABLE; LIGHT CURVES; VARIABLE STAR. Virginia Trimble

Bibliography. M. F. Bode and A. Evans (eds.), *Classical Novae*, 1989; C. H. Payne-Gaposchkin, *The Galactic Novae*, 1957, reprint 1964; N. Vogt (ed.), *Cataclysmic Variable Stars*, 1992.

Nozzle

A conduit with a variable cross-sectional area in which a fluid accelerates into a high-velocity stream. The effect of the changing cross-sectional area on the fluid velocity can be explained by the principle of mass conservation applied to successive cross-sectional planes of the nozzle. Equation (1) must be

$$\dot{m} = \rho V A \tag{1}$$

satisfied, where ρ is the mass density of the fluid [in kilograms per cubic meter (kg/m^3)], V is the average velocity in the cross section [in meters per second (m/s)], A is the cross-sectional area [in square meters (m^2)], and $\dot{m}$ is the rate of mass flow through the nozzle [in kilograms per second (kg/s)]. Decreasing A along the length of the nozzle must result in an increase in ρV since $\dot{m}$ is the same at every cross section. *See* FLUID FLOW.

Design and operation. An important parameter of nozzle operation is the difference in pressure [in pascals (Pa)] between the inlet and outlet. Higher pressure differences push the fluid to higher velocities and achieve higher mass flow rates for a given nozzle size.

Liquid nozzles. For liquid nozzles it can be assumed that ρ is constant and therefore V increases as A decreases and vice versa. Liquid nozzles, such as those on fire hoses, are called converging because the area decreases along the length of the nozzle to increase the speed. Typical liquid nozzles have a simple conical shape and are designed to a specific ratio of inlet to outlet areas. To produce good uniformity of the velocity profiles in the issuing jet, the total cone angle should be less than 30°. In cases where the pressure difference across the nozzle is fixed, the mass flow rate from the nozzle can be regulated by pushing a conical spear into the exit plane of the nozzle from the inside. This has the effect of reducing the outlet flow area and hence the mass flow rate.

The difference between the reservoir pressure (P_o) and the nozzle exit pressure (P_b) required for a given exit velocity V_e can be determined approximately, using Bernoulli's theorem, as given by Eq. (2). In cases where the pressure difference is less

$$P_o - P_b \cong \frac{\rho V_e^2}{2} \tag{2}$$

than 10 kPa, air may be treated as if its density were constant. *See* BERNOULLI'S THEOREM.

High-speed gas nozzles. In the case of gas nozzles the gas density can change dramatically as a result of the pressure reduction between the inlet and outlet of the nozzle. At very high gas speeds this effect is so significant that the basic shape of the nozzle must change to a converging-diverging form (**Fig. 1**). The diverging portion is necessary to accommodate the expansion of the gas as it accelerates to lower pressure.

An understanding of the operation of the converging-diverging nozzle requires knowledge of the Mach number, which is the ratio of the gas speed

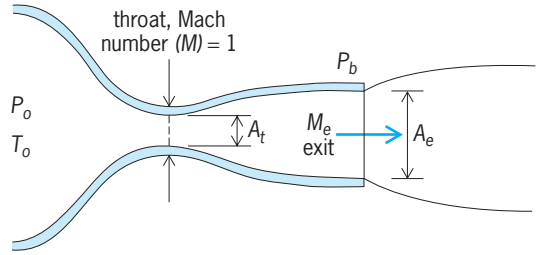

Fig. 1. Typical convergent-divergent nozzle with a jet plume. P_0, T_0 = pressure and temperature upstream of the nozzle; A_t = area at the throat; P_b = back pressure; M_e, A_e = Mach number and area at exit.

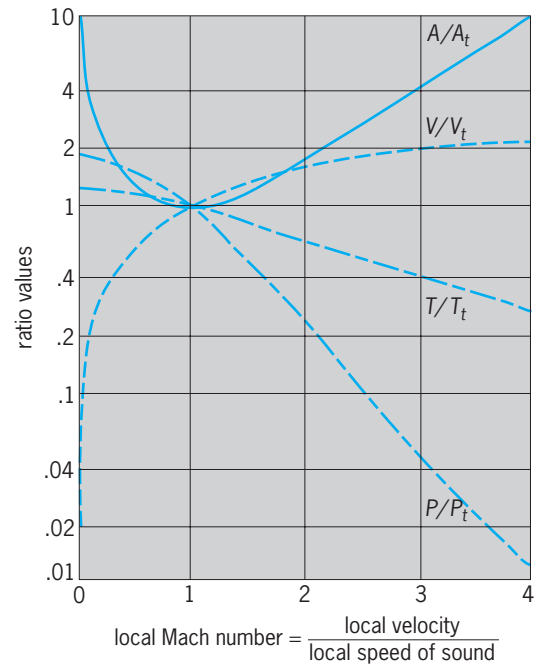

Fig. 2. Conditions for isentropic flow of air through a nozzle when the Mach number at the throat is unity. Ratios of area A, pressure P, temperature T, and velocity V to their values at the throat (A_t, P_t, I_t, and V_t) are plotted as functions of Mach number.

to the speed of sound in the gas c; $Ma = V/c$. A subsonic flow has $Ma < 1$, a sonic flow has $Ma = 1$, and a supersonic flow has $Ma > 1$. **Figure 2** shows how the pressure, temperature, and speed of air flowing in a nozzle are related to Mach number and profile of nozzle cross-sectional area. From this figure, it can be seen that the flow in the converging portion of the nozzle must be subsonic, the flow at the throat can at most be sonic, and supersonic flow can occur only in the diverging portion. To achieve supersonic flow in the diverging section, the reservoir pressure must be sufficient to achieve sonic flow at the throat. The fall of air temperature with increasing Ma indicated in Fig. 2 is a direct result of the gas expansion. *See* MACH NUMBER; SOUND.

To achieve a particular Mach number at the nozzle exit (Ma_e), the ratio of the pressures at the outlet (P_b) and inlet (P_o) of the nozzle must follow Eq. (3).

$$\frac{P_b}{P_o} \cong \left(1 + 0.2Ma_e^2\right)^{-3.5} \tag{3}$$

If P_b/P_0 is somewhat higher than this value and the same supersonic flow condition prevails inside the divergent part of the nozzle, then an oblique shock wave will originate at the edge of the nozzle. For even higher back-pressure ratios, supersonic flow cannot be realized at the exit, although the throat can still be sonic. Under this condition, a normal shock wave occurs in the divergent part of the nozzle. This operating condition is not desirable, because the flow is subsonic at the exit of the nozzle and the useful mechanical energy is degraded into thermal energy whenever a shock wave occurs. When the back-pressure ratio is lower than the value given by Eq. (3), the jet stream from the nozzle will expand to form a plume (Fig. 1). *See* JET FLOW; SHOCK WAVE.

There is a very interesting phenomenon observed in the operation of gas nozzles known as choking. Once the velocity at the throat reaches sonic speed, the back pressure has no further effect on throat conditions and the mass flow is entirely determined by the reservoir conditions and throat area. For air flow, this relationship is Eq. (4), where T_o is the absolute reservoir temperature [in kelvins (K)] at the nozzle inlet. This equation specifies the maximum mass flow for a given throat area, or alternatively, it specifies the minimum throat area for a given mass flow. *See* CHOKED FLOW.

$$\dot{m}_{\max} = 0.0404 \frac{P_o}{T_o^{0.5}} A_t \qquad (4)$$

Viscous effects. The above discussion is based upon the assumption that the fluid is frictionless (inviscid). Although all real fluids are viscous, the influence due to viscosity is generally very minor. Only when using a very viscous liquid, such as heavy oil, or when a gas accelerates to a very high Mach number inside the nozzle (Ma > 5) do viscous effects become important. *See* GAS DYNAMICS; VISCOSITY.

Applications. A nozzle can be used for a variety of purposes. It is an indispensable piece of equipment in many devices employing fluid as a working medium. The reaction force that results from the fluid acceleration may be employed to propel a jet aircraft or a rocket. In fact, most military jet aircraft employ the simple convergent conical nozzle, with adjustable conical angle, as their propulsive device. If the high-velocity fluid stream is directed to turn a turbine blade, it may drive an electric generator or an automotive vehicle. High-velocity streams may also be produced inside a wind tunnel so that the conditions of flight of a missile or an aircraft may be simulated for research purposes. The nozzle must be carefully designed in this case to provide uniformly flowing fluid with the desired velocity, pressure, and temperature at the test section of the wind tunnel. Nozzles may also be used to disperse fuel into an atomized mist, such as that in diesel engines, for combustion purposes. *See* ATOMIZATION; IMPULSE TURBINE; INTERNAL COMBUSTION ENGINE; JET PROPULSION; ROCKET PROPULSION; WIND TUNNEL.

Nozzles may also be used as metering devices for gas or liquid. A convergent nozzle inserted into a pipe produces a pressure difference between its inlet and outlet that can be measured and related directly to the mass flow. For this purpose, the shape of the nozzle must be designed according to standard specifications or carefully calibrated against a mass flow rate standard. A unique feature of gas metering nozzles is that the mass flow rate through the nozzle depends only on the inlet pressure if the pressure ratio across it is sufficient to produce sonic velocity at the throat. *See* FLOW MEASUREMENT.

Wen L. Chow; A. Gordon L. Holloway

Bibliography. J. D. Anderson, *Modern Compressible Flow*, 3d ed., McGraw-Hill, 2003; R. D. Blevins, *Applied Fluid Dynamics*, Krieger Publishing, 1992; F. M. White, *Fluid Mechanics*, 5th ed., McGraw-Hill, 2003.

Nuclear battery

A battery that converts the energy of particles emitted from atomic nuclei into electric energy. Two basic types have been developed: (1) a high-voltage type, in which a beta-emitting isotope is separated from a collecting electrode by a vacuum or a solid dielectric, provides thousands of volts but the current is measured in picoamperes; (2) a low-voltage type gives about 1 V with current in microamperes.

High-voltage nuclear battery. In the high-voltage type, a radioactive source is attached to one electrode, emitting charged particles. The source might be strontium-90, krypton-85, or hydrogen-3 (tritium), all of which are pure beta emitters. An adjacent electrode collects the emitted particles. A vacuum or solid dielectric separates the source and the collector electrodes.

One high-voltage model, shown in **Fig. 1**, employs tritium gas sorbed in a thin layer of zirconium metal as the radioactive source. This source is looped around and spot-welded to the center tube of a glass-insulated terminal. A thin coating of carbon applied to the inside of a nickel enclosure acts as an efficient collector having low secondary emission. The glass-insulated terminal is sealed to the nickel enclosure. The enclosure is evacuated through the center tube, which is then pinched off and sealed.

The Radiation Research Corporation model R-1A is $^3/_8$ in. (0.95 cm) in diameter and 0.531 in. (1.35 cm) in height. It weighs 0.2 oz (5.7 g) and occupies 0.05 in.^{3} (0.8 cm^3). It delivers about 500 V at 160 pA. Future batteries are expected to deliver 1 μA at 2000 V, with a volume of 64 in.^{3} (1048 cm^3).

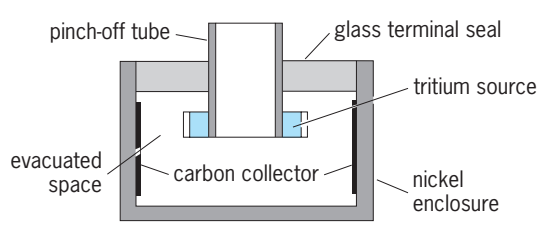

Fig. 1. Tritium battery in cross section.

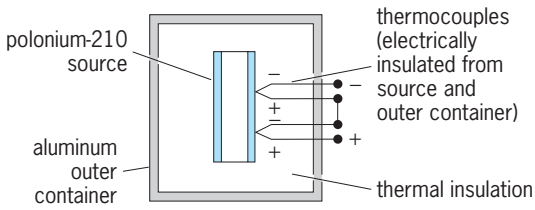

Fig. 2. Thermoelectric nuclear battery.

Earlier models employed strontium-90. This isotope has the highest toxicity in the human body of the three mentioned. Tritium has only one one-thousandth the toxicity of strontium-90. Both strontium-90 and krypton-85 require shielding to reduce external radiation to safe levels. Tritium produces no external radiation through a wall that is thick enough for any structural purpose. Tritium was selected on the basis of these advantages.

The principal use of the high-voltage battery is to maintain the voltage of a charged capacitor. The current output of the radioactive source is sufficient for this purpose.

This type of battery may be considered as a constant-current generator. The voltage is proportional to the load resistance. The current is determined by the number of emissions per second captured by the collector and does not depend on ambient conditions or the load. As the isotope ages, the current declines. For tritium, the intensity drops 50% in a 12-year interval. For strontium-90, the intensity drops 50% in a 25-year interval.

Low-voltage nuclear battery. Three different concepts have been employed in the low-voltage type of nuclear batteries: (1) a thermopile, (2) the use of an ionized gas between two dissimilar metals, and (3) the two-step conversion of beta energy into light by a phosphor and the conversion of light into electric energy by a photocell.

Thermoelectric-type nuclear battery. This low-voltage type, employing a thermopile, depends on the heat produced by radioactivity (**Fig. 2**). It has been calculated that a sphere of polonium-210 of 0.1 in. (0.25 cm) diameter, which would contain about 350 curies (1.3×10^{13} becquerels), if suspended in a vacuum, would have an equilibrium surface temperature of 4000°F (2200°C), assuming an emissivity of 0.25. For use as a heat source, it would have to be hermetically sealed in a strong, dense capsule. Its surface temperature, therefore, would be lower than 4000°F (2200°C).

To complete the thermoelectric battery, the heat source must be thermally connected to a series of thermocouples which are alternately connected thermally, but not electrically, to the heat source and to the outer surface of the battery. After a short time, a steady-state temperature differential will be set up between the junctions at the heat source and the junctions at the outer surface. This creates a voltage proportional to the temperature drop across the thermocouples. The battery voltage decreases as the age of the heat source increases. With polonium-210 (half-life, 138 days) the voltage drops about 0.5%/day. The drop for strontium-90 is about 0.01%/day (20-year half-life).

A battery containing 57 curies (2.1×10^{12} becquerels) of polonium-210 sealed in a sphere 0.4 in. (1 cm) in diameter and seven chromel-constantan thermocouples delivered a maximum power of 1.8 mW. It had an open-circuit voltage of 42 mV with a 140°F (78°C) temperature differential. Over a 138-day period, the total electrical output would be about 1.5×10^4 joules (watt-seconds).

Total weight of the battery was 1.20 oz (34 g). This makes the energy output per pound equal to

$$\frac{1.5 \times 10^4}{3600} \text{ watt-hours} \times \frac{1}{1.2} \times 16 = 55.6$$

This is the same magnitude as with conventional electric cells using chemical energy. This nuclear energy, however, is being dissipated whether or not the electric energy is being used.

The choice of isotope for a thermoelectric nuclear battery is somewhat restricted. Those with a half-life of less than 100 days would have a short useful life, and those with a half-life of over 100 years would give too little heat to be useful. This leaves 137 possible isotopes. This number is further reduced by the consideration of shielding.

The trend is to use plutonium-238 (unable to support a chain reaction) and strontium-90. The most frequently used thermocouple material is a doped lead-telluride alloy that performs between 400 and 900°F (200 and 480°C). Another material is silicon-germanium alloy, which can operate at 1500°F (800°C) (hot junction).

The thermoelectric systems developed so far are in the small power range (5–60 W), and work with a maximum efficiency of 7%.

When shielding is not a critical problem, as in unmanned satellites, they are extremely convenient, being reliable and light. SNAP-3 and SNAP-9A (Systems for Nuclear Auxiliary Power) have weights in the 0.5–1.0 kg/W range.

The above discussion applies only to a portable power source. Drastic revision would be required if a thermoelectric-type nuclear battery were to be designed for central-station power. *See* THERMOELECTRICITY.

Gas-ionization nuclear battery. In this battery a beta-emitting isotope ionizes a gas situated in an electric field (**Fig. 3**). Each beta particle produces about 200 current carriers (ions), so that a considerable current multiplication occurs compared with the

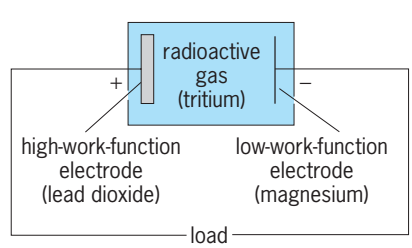

Fig. 3. Gas-ionization nuclear battery.

rate of emission of the source. The electric field is obtained by the contact potential difference of a pair of electrodes, such as lead dioxide (high work function) and magnesium (low work function). The ions produced in the gas move under the influence of the electric field to produce a current.

A cell containing argon gas at 2 atmospheres, electrodes of lead dioxide and magnesium, and a radioactive source consisting of 1.5 millicuries (5.6×10^7 becquerels) of tritium has a volume of 0.01 in.^{3} (0.16 cm^3) and an active plate area of 0.2 in.^{2} (1.29 cm^2), and gives a maximum current of 1.6×10^{-9} A. The open-circuit voltage per cell depends on the contact potential of the electrode couple. A practical value appears to be about 1.5 V. Voltage of any value may be achieved by a series assembly of cells.

The ion generation is exploited also in a purely thermal device, the thermionic generator. It consists of an emitting electrode heated to 2200–3300°F (1200–1800°C) and a collecting electrode kept at 930–1650°F (500–900°C). The hot electrode is made of tungsten or rhenium, metals which melt at more than 5400°F (3000°C) and have low vapor pressures at the working temperature. The collecting electrode is made out of a metal with a low work function, such as molybdenum, nickel, or niobium. Electrons from the hot electrode traverse a gap of 0.04–0.08 in. (1–2 mm) to the collector and recover the emitter by an external circuit.

The ionization space is filled with cesium vapor which acts in two ways. First, it covers the surface of the two electrodes with adsorbed cesium atoms and thus reduces the work function to the desired level. Second, it creates an ionized atmosphere, thus controlling the electron space charge.

The heating of the emitting electrode can be obtained by any of the known means: concentrated solar energy, radioisotopes, or conventional nuclear reactors. While reaching the high temperature was a problem easy to solve, the cooling of the collecting electrode appeared for a long time to be a critical technical problem. This problem was solved by the development of the "heat pipe."

In principle the heat pipe is an empty cylinder absorbing heat at one end by vaporization of a liquid and releasing heat at the other end by condensation of the vapor; the liquid returns to the heat-absorbing end by capillarity through a capillary structure covering the internal face of the cylinder. The heat transfer of a heat pipe may be 10,000 times or more higher than that of copper or silver.

The prototype developed under the sponsorship of NASA (NASA/RCA Type A1279) represents an advanced design capable of supplying 185 W at 0.73 V with an efficiency of 16.5%.

Since no major limitations are foreseen in the development of thermionic fuel cells, it is expected that they will provide the most convenient technique for using nuclear energy in the large-scale production of electricity.

Scintillator-photocell nuclear battery. This type of cell is based on a two-step conversion process (**Fig. 4**).

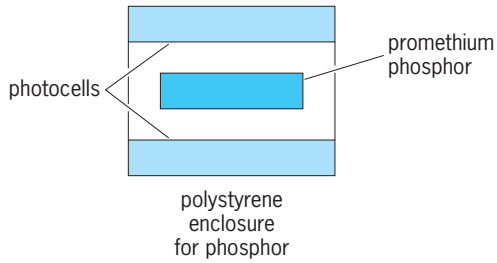

Fig. 4. Scintillator-photocell battery.

Beta-particle energy is converted into light energy; then the light energy is converted into electric energy. In order to accomplish these conversions, the battery has two basic components, a light source and photocells.

The light source consists of a mixture of finely divided phosphor and promethium oxide (Pm_2O_3) sealed in a transparent container of radiation-resistant plastic. The light source is in the form of a thin disk. The photocells are placed on both faces of the light source. These cells are modified solar cells of the diffused-silicon type.

Since the photocells are damaged by beta radiation, the transparent container of the light source must be designed to absorb any beta radiation not captured in the phosphor. Polystyrene makes an excellent light-source container because of its resistance to radiation.

The light source must emit light in the range at which the photocell is most efficient. A suitable phosphor for the silicon photocell is cadmium sulfide or a mixture of cadmium and zinc sulfide.

In a prototype battery, the light source consisted of 0.0018 oz (50 milligrams) of phosphor and about 0.00018 oz (5 mg) of the isotope promethium-147. This isotope is a pure beta-emitter with a half-life of 2.6 years. It is deposited as a coating of hydroxide on the phosphor particles, which are then dried to give the oxide. For use with the low light level (about 0.001 times sunlight) of the light source, special treatment is necessary to make the equivalent shunt resistance of the cell not less than 100,000 ohms. For a description of the photocell *see* SOLAR CELL.

The prototype battery, when new, delivers 20×10^{-6} A at 1 V. In 2.6 years (half-life) the current drops about 50% but the voltage drops only about 5%.

The power output improves with decreasing temperature, as a result of improved photocell diode characteristics which more than compensate for a decrease in short-circuit current. At −100°F (−73°C), the power output is 1.7 times as great as at room temperature. At 144°F (62°C), the power output is only 0.6 times as great as at room temperature.

The battery requires shielding to reduce the weak gamma radiation to less than 9 milliroentgens per hour (mR/h), which is the tolerance for continuous exposure of human extremities. The unshielded battery has a radiation level of 90 mR/h. By enclosing the cell in a case of tungsten alloy, density 16.5, the external radiation becomes less than 9 mR/h.

The unshielded battery has a volume of 0.014 in.^{3} (0.23 cm^3) and a weight of 0.016 oz (0.45 g). Over a 2.5 year period, the total output would be 0.32 Wh (whether or not used). This gives a unit output of 320 Wh/lb (704 Wh/kg), about six times as great as chemical-battery output. But the shielded battery has a volume of 0.07 in.^{3} (1.15 cm^3) and a weight of 0.6 oz (17 g). This reduced the unit output to 8.5 Wh/lb (18.7 Wh/kg). The cell can undergo prolonged storage at 200°F (93°C). *See* BATTERY.

Jack Davis; L. Rozeanu; Kenneth Franzese

Nuclear binding energy

The amount by which the mass of an atom is less than the sum of the masses of its constituent protons, neutrons, and electrons expressed in units of energy. This energy difference accounts for the stability of the atom. In principle, the binding energy is the amount of energy which was released when the several atomic constituents came together to form the atom. Most of the binding energy is associated with the nuclear constituents (protons and neutrons), or nucleons, and it is customary to regard this quantity as a measure of the stability of the nucleus alone. *See* NUCLEAR STRUCTURE.

A widely used term, the binding energy (BE) per nucleon, is defined by the equation below, where

$$\text{BE/nucleon} = \frac{[ZH + (A - Z)n - {}_ZM^A]c^2}{A}$$

${}_ZM^A$ represents the mass of an atom of mass number A and atomic number Z, H and n are the masses of the hydrogen atom and neutron, respectively, and c is the velocity of light. The binding energies of the orbital electrons, here practically neglected, are not only small, but increase with Z in a gradual manner; thus the BE/nucleon gives an accurate picture of the variations and tends in nuclear stability. The **illustration** shows the BE/nucleon (in megaelectronvolts) plotted against mass number for $A > 40$.

The BE/nucleon curve at certain values of A suddenly changes slope in such a direction as to indicate that the nuclear stability has abruptly deteriorated. These turning points coincide with particularly stable configurations, or nuclear shells, to which additional nucleons are rather loosely bound. Thus there is a sudden turning of the curve over $A = 52$ (28 neutrons); the maximum occurs in the nickel region (28 protons, $\sim A = 60$); the stability rapidly deteriorates beyond $A = 90$ (50 neutrons); there is a slightly greater than normal stability in the tin region (50 protons, $\sim A = 118$); the stability deteriorates beyond $A = 140$ (82 neutrons) and beyond $A = 208$ (82 protons plus 126 neutrons).

The BE/nucleon is remarkably uniform, lying for most atoms in the range 5–9 MeV. This near constancy is evidence that nucleons interact only with near neighbors; that is, nuclear forces are saturated.

The binding energy, when expressed in mass units, is known as the mass defect, a term sometimes incorrectly applied to quantity $M - A$, where M is the mass of the atom. *See* MASS DEFECT.

Henry E. Duckworth

The term binding energy is sometimes also used to describe the energy which must be supplied to a nucleus in order to remove a specified particle to infinity, for example, a neutron, proton, or alpha particle. A more appropriate term for this energy is the separation energy.This quantity varies greatly

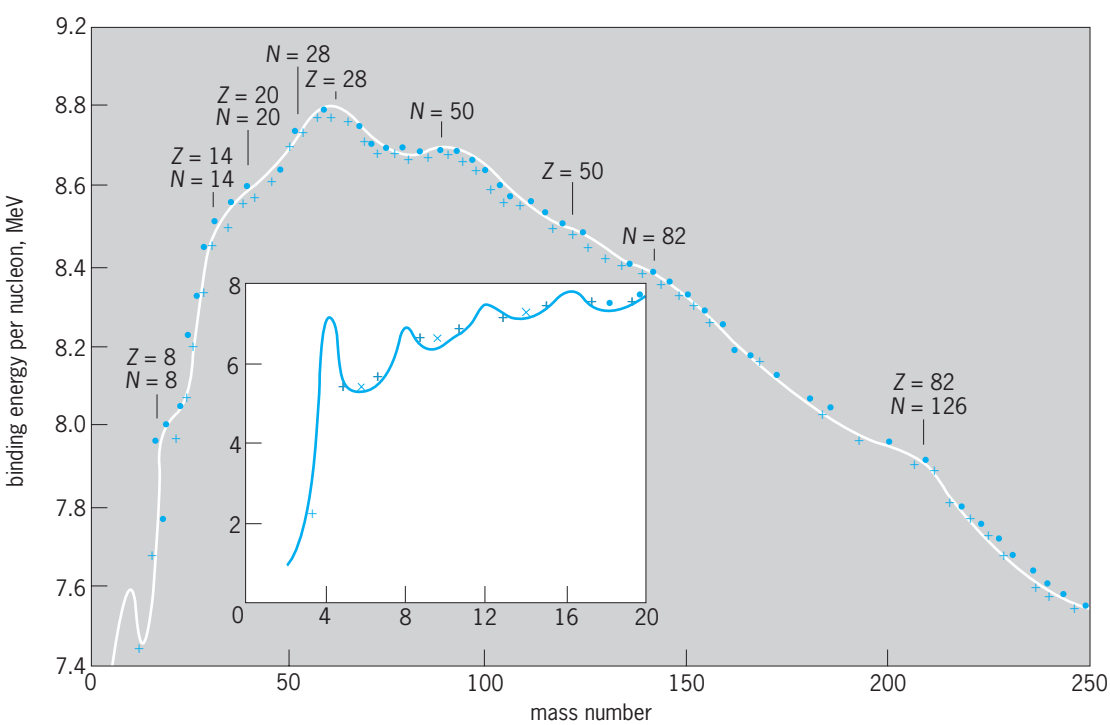

Graph of binding energy per nucleon (in megaelectronvolts) plotted against mass number. N = number of neutrons. (*After A. H. Wapstra, Isotopic measure, part 1, where A is less than 34, Physica, 21:367–384, 1955*)

from nucleus to nucleus and from particle to particle. For example, the binding energies for a neutron, a proton, and a deuteron in ^{16}O are 15.67, 12.13, and 20.74 MeV, respectively, while the corresponding energies in ^{17}O are 4.14, 13.78, and 14.04 MeV, respectively. The usual order of neutron or proton separation energy is 7–9 MeV for most of the periodic table. D. H. Wilkinson

Nuclear chemical engineering

The branch of chemical engineering that deals with the production and use of radioisotopes, nuclear power generation, and the nuclear fuel cycle.

A nuclear chemical engineer requires training in both nuclear and chemical engineering. As a nuclear engineer, he or she should be familiar with the nuclear reactions that take place in nuclear fission reactors and radioisotope production, with the properties of nuclear species important in nuclear fuels, with the properties of neutrons, gamma rays, and beta rays produced in nuclear reactors, and with the reaction, absorption, and attenuation of these radiations in the materials of reactors. *See* BETA PARTICLES; GAMMA RAYS; NEUTRON.

As a chemical engineer, he or she should know the properties of materials important in nuclear reactors and the processes used to extract and purify these materials and convert them into the chemical compounds and physical forms used in nuclear systems. *See* CHEMICAL ENGINEERING.

Radioisotopes. The principal nuclear species important in nuclear fuels are the fissionable isotopes uranium-235, which makes up 0.7% of natural uranium; plutonium-239, produced by neutron irradiation of uranium-238, the other 99.3% of natural uranium; and uranium-233, produced by neutron irradiation of natural thorium. Other nuclear species of interest to nuclear chemical engineers are the fission products, which make irradiated nuclear fuel intensely radioactive; radioactive isotopes of heavy elements such as protactinium, neptunium, plutonium, americium, and curium, produced by neutron absorption in nuclear fuels; and isotopes of light elements, such as deuterium used in the form of heavy water in reactors fueled with natural uranium, tritium (hydrogen-3) used with deuterium as fuel for fusion reactors, and lithium-6, the starting material for tritium production. *See* AMERICIUM; CURIUM; DEUTERIUM; LITHIUM; NEPTUNIUM; PLUTONIUM; PROTACTINIUM; THORIUM; TRANSURANIUM ELEMENTS; TRITIUM; URANIUM.

Power generation. The principal types of nuclear power reactors are light-water reactors, of either the pressurized or boiling-water type; pressurized heavy-water reactors; and sodium-cooled, liquid-metal fast breeder reactors. Aspects of light- and heavy-water reactors of concern to nuclear chemical engineers include production and purification of the uranium dioxide fuel, production of the hafnium-free zirconium tubing used for fuel cladding, and control of corrosion and radioactive corrosion products by chemical treatment of coolant. Another chemical engineering aspect of heavy-water reactor operation is control of the radioactive tritium produced by neutron activation of deuterium. Aspects of liquid-metal fast-breeder reactors of concern to nuclear chemical engineers include fabrication of the mixed uranium dioxide–plutonium dioxide fuel, purity control of sodium coolant to prevent fouling and corrosion, and reprocessing of irradiated fuel to recover plutonium and uranium for recycle. *See* NUCLEAR POWER; NUCLEAR REACTOR.

Nuclear fuel cycle. The nuclear fuel cycle comprises the industrial operations involved in producing fuel for nuclear reactors, and processing and storing safely the intensely radioactive fuel discharged from them after irradiation. For light-water reactors, the principal steps of the fuel cycle are: mining of uranium ores; concentration of the contained uranium; purification of uranium; conversion to uranium hexafluoride; enrichment of uranium-235 from 0.7% in natural uranium to about 3% by gaseous diffusion or the gas centrifuge; fabrication of enriched uranium into fuel elements; irradiation of fuel elements in a nuclear reactor to produce energy and neutrons; storage of irradiated fuel to permit some of its intense radioactivity to decay; packaging irradiated fuel for long-term, safe storage; and emplacement of fuel in terminal storage facilities. Alternatively, because irradiated fuel contains as much uranium-235 as natural uranium and about 0.5% of fissionable plutonium, irradiated fuel may be reprocessed to recover these valuable materials for reuse as reactor fuel, and to separate and convert the radioactive fission products into an insoluble glass. Such glass is more inert, more compact, and more suitable for long-term storage than unreprocessed irradiated fuel. *See* HAZARDOUS WASTE; NUCLEAR FUEL CYCLE; NUCLEAR FUELS; NUCLEAR FUELS REPROCESSING; RADIOACTIVE WASTE MANAGEMENT; URANIUM METALLURGY. Manson Benedict

Bibliography. M. Benedict, T. H. Pigford, and H. W. Levi, *Nuclear Chemical Engineering*, 2d ed., 1981; R. A. Knief, *Nuclear Engineering: Theory and Technology of Commercial Nuclear Power (SCPP)*, 2d ed., 1992; J. R. Lamarsh and A. J. Baratta, *Introduction to Nuclear Engineering*, 3d ed., 2001; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; Jan Rydberg et al., *Radiochemistry and Nuclear Chemistry*, 3d ed., 2001; P. D. Wilson (ed.), *The Nuclear Fuel Cycle: From Ore to Wastes*, 1996.

Nuclear chemistry

An interdisciplinary field that, in general, encompasses the application of chemical techniques to the solution of problems in nuclear physics. The discovery of the naturally occurring radioactive elements and of nuclear fission are classical examples of the work of nuclear chemists.

Although chemical techniques that are employed in nuclear chemistry are essentially the same as those

in radiochemistry, these fields may be distinguished on the basis of the aims of the investigation. Thus, a nuclear chemist utilizes chemical techniques as a tool for the study of nuclear reactions and properties, whereas a radiochemist utilizes the radioactive properties of certain substances as a tool for the study of chemical reactions and properties. There is considerable overlap between the two fields and in some cases (for example, the preparation and study of synthetic elements) a distinction may be somewhat arbitrary. Nuclear chemistry and nuclear physics are also closely related, and nuclear chemists may employ purely physical techniques in their studies. The term hot-atom chemistry, originally used to describe the chemical reactions of atoms produced in nuclear transformations, is sometimes extended to include the chemical reactions of any excited atoms, such as those produced in accelerated beams or by absorption of radiation. For the application of radioactive tracers to chemical problems. For the chemical effects of radiation on various systems *see* RADIATION CHEMISTRY; RADIOCHEMISTRY.

Scope. The chemical identification of radioactive nuclides and the determination of their nuclear properties (half-life, mode and energy of decay, mass number, nucleon binding energies, and so on) has been one of the major activities of nuclear chemists. Such studies have produced an extensive array of radioisotopes, and ongoing studies are concerned mainly with the more difficult identification of nuclides of very short half-life at the boundaries of the "valley of beta stability" and in the regions of neutron and proton instability, as well as short-lived isomeric states. Nuclear chemical investigations led to the discovery of the synthetic radioactive elements which do not have any stable isotopes and are not formed in the natural radioactive series (technetium, promethium, astatine, and the transuranium elements). The synthesis of element 107 has been announced and claims made for the observation of elements 108 and 109. Attempts to extend the periodic table even further and to prepare larger quantities of the new elements for intensive study are being made by using various nuclear reactions, particularly accelerated heavy-ion bombardments of targets of the heaviest elements available. Accelerators capable of accelerating even uranium ions are available, and investigations of the possible existence of superheavy nuclei in a new region of relative stability are being carried out. *See* TRANSURANIUM ELEMENTS.

Other major areas of nuclear chemistry include studies of nuclear structure and spectroscopy (involving the determination of the energy spectra of emitted particles and the construction of energy-level diagrams, or decay schemes) and of the probability and mechanisms of various nuclear reactions (involving the determination of excitation functions, that is, the variation of cross section—or probability of reaction—with the energy of the bombarding particle, and the angular and energy distributions and correlations of the reaction products).

Many of these investigations involve the chemical isolation and identification of the products of nuclear reactions or chemical techniques for the preparation of thin, uniform deposits of pure materials for use as bombardment targets and as sources for nuclear spectroscopy. However, the nuclear chemist also carries out many experiments of a purely physical nature, observing the products of nuclear reactions with counters, ionization chambers, nuclear emulsions, and other particle track detectors. Moreover, advances in the technology of nuclear radiation detectors and associated electronic equipment have led to instruments of such sophistication that classical methods of chemical analyses are often not necessary. The borderline between nuclear chemistry and nuclear physics has, in fact, practically disappeared in many areas of investigation (for example, studies of heavy-ion reactions and many aspects of nuclear structure and spectroscopy), and if a distinction is to be made at all, it may be only that the nuclear physicist tends to concentrate on interactions between the fundamental particles and observes the nucleons and light nuclei (neutrons, protons, deuterons, tritions, and alpha particles) emitted in nuclear reactions, while the nuclear chemist examines interactions with complex nuclei and observes the heavier fragments and residual nuclei produced. This distinction is perhaps most meaningful in the area of high-energy nuclear phenomena where the high-energy, or particle, physicist is concerned mainly with the production and properties of and the interactions between the elementary (or strange) particles and their role in the structure of matter. The nature of the interaction of high-energy projectiles with complex nuclei and studies of the products formed is mainly the concern of nuclear chemists. Studies involving the use of highly radioactive isotopes of the actinide elements and investigations of many aspects of the nuclear fission process have been carried out largely by nuclear chemists. *See* NUCLEAR REACTION.

The contributions of nuclear chemists have not been limited to empirical studies of nuclear reactions and properties. They are also active in theoretical investigations and have made significant contributions to theories of nuclear structure, radioactive decay (particularly alpha emission), and nuclear reactions of many types, including fission. *See* NUCLEAR STRUCTURE.

Historical background. Nuclear chemistry began with the work of Pierre and Marie Curie on the isolation and identification of polonium and radium in 1898. The discovery of actinium, thoron (radon-222), and radon in the next few years, the identification of the alpha particle as the helium ion, and the characterization of the relationships in the natural radioactive decay series played an essential role in the development of atomic and nuclear science. The interpretation of the results of these early experiments led to the establishment of the laws of radioactive transformation and the concept of isotopes (E. Rutherford and F. Soddy). The phenomenon of nuclear isomerism was discovered by O. Hahn (1921), who showed that the radioactive species UZ and UX_2 represented different excitation levels of ^{234}Pa.

See ACTINIUM; ISOTOPE; POLONIUM; PROTACTINIUM; RADIUM; RADON.

Artificial transmutation reactions were first carried out utilizing the alpha particles emitted in the decay of the natural radioactive isotopes as bombarding particles (Rutherford and coworkers, 1919). Such bombardments led to the discovery of artificial radioactivity by Irene Curie and F. Joliot in 1934. The discovery of the neutron (J. Chadwick, 1932) and deuterium (H. Urey and coworkers, 1932) and the development of devices to accelerate the ions of hydrogen (proton), deuterium (deuteron), and helium (alpha particle) gave tremendous impetus to the production and study of radioactive nuclides. By 1937 about 200 were known, and the number has grown steadily to nearly 2100 in 1984.

The discovery of nuclear fission by O. Hahn and F. Strassmann in 1939 culminated a dramatic episode in the history of nuclear science and opened the doors to the atomic age. Nuclear chemistry became a mature science during the period 1940–1945, largely as a result of the concentrated effort to produce the atomic bomb. It has continued to grow and expand in scope with the advancing technology in particle accelerators, nuclear reactors, radiation detection instrumentation, mass spectrometers, electromagnetic isotope separators, and computers. *See* NUCLEAR FISSION.

Nuclear spectroscopy. The energy spectra of the radiations emitted by radioactive nuclides are determined by the use of a number of specialized instruments. Alpha-particle spectra may be determined quite well with solid-state detectors coupled to multichannel pulse-height analyzers. The pulse output of the detector is proportional to the energy deposited and the individual events are stored in the analyzer memory. Magnetic spectrometers are used for higher resolution but require sources of higher intensity. *See* NUCLEAR SPECTRA.

Beta-particle spectra may be examined with proportional or scintillation counters, but precise energy determinations and studies of the shapes of the spectra require magnetic spectrometers.

Gamma-ray and x-ray spectra are very conveniently measured using thallium-activated sodium iodide (NaI) crystal scintillation counters which give a fluorescent light output proportional to the energy deposited in the crystal. The light is amplified with a photomultiplier tube and converted to a pulse-height that is stored in a multichannel and analyzer. Solid-state detectors, particularly intrinsic germanium and lithium-drifted germanium and silicon crystals, possess higher resolution and are generally used for studies requiring high accuracy. Bent-crystal spectrometers, which depend on Bragg scattering, are instruments of high resolution but require intense sources of the radioactive nuclide. They are often used in measurements of gamma-ray spectra accompanying nuclear reactions—capture gamma rays in the (n,γ) reaction, for example.

Nuclear structure determinations may also involve nuclear reaction spectroscopy, in which the energy levels of the nuclide of interest are derived from the projectile properties (type and energy) and the measured (discrete) energies of the ejected nucleon or fragment [for example, in a (d,p), (d,t), or more complex reaction]. Combinations of several techniques may be required to establish the energy-level structure of a nuclide and characterize the energy, spin, and parity of each level.

Studies of nuclear reactions. In a nuclear reaction, a bombarding particle interacts with a target nucleus to produce one or more product nuclei and, perhaps, other particles. (Spontaneous transformations of unstable nuclei, such as alpha and beta decay or spontaneous fission, may be considered a special type of nuclear reaction in which no external excitation by bombarding particles is involved.) If the target is thin with respect to the ranges of the product nuclei, or if it is dispersed in a medium such as a nuclear emulsion, the reaction may be studied event by event by recording the emitted nuclei and particles with counters or observing the tracks they leave in nuclear emulsions or other particle track detectors, such as mica or polycarbonate films. Such studies are especially useful for determinations of the angular distributions of the recoil fragments and for information on the multiplicity of fragments emitted in a single event. More sophisticated techniques of track detection employing bubble, spark, and streamer chambers can provide more information on the particle identity and energy from the track density and its curvature in magnetic fields.

By suitable in-line arrangements of thin detectors (through which the particles pass, depositing only part of their kinetic energy) and thick detectors (which record the total kinetic energy), the mass and charge of many light nuclei can be uniquely identified. Time-of-flight (velocity) and kinetic energy measurements may also be combined to establish the masses of recoiling nuclei from nuclear reactions, and measurements of the characteristic x-rays may be made to establish the nuclear charge. These measurements require rather sophisticated instrumentation, and they cannot be applied in all cases. Where applicable, however, they can provide information on the correlations between particles or fragments emitted in the same event, information that is essential for an understanding of the reaction mechanism. *See* PARTICLE DETECTOR.

When a thick target is bombarded, the reaction products cannot escape, and an accumulation of many individual events is retained in the target matrix. (Catcher foils are placed adjacent to the target to retain the small fraction of the products escaping from the target surfaces.) The recoil products from thin targets may also be accumulated on catcher foils at various angles to the beam direction for angular distribution studies and recoil range measurements. Chemical separations are employed to identify the elements formed in the reactions and measurements of the radiations made to characterize the isotopes. When a number of isotopes of the same element are produced in the target and cannot be conveniently differentiated on the basis of half-life or radiation characteristics, chemical separations alone

may not suffice. Separation of the isotopes can be achieved by the acceleration and deflection of ions in electromagnetic isotope separators. The separated isotopes may be collected and their radiations measured free from any interference. The abundances of stable as well as radioactive species produced in nuclear reactions may also be determined by mass spectrometry. *See* MASS SPECTROMETRY.

The radiochemical separation and identification of the accumulated reaction products from the matrix of a thick target following bombardment is a relatively simple means of determining the identity and yields (or formation cross sections) of the products. The high sensitivity and specificity of the technique make it particularly useful for measurements of rare events. Its relative simplicity makes it especially valuable for general surveys and systematics of nuclear reactions to establish trends and identify regions of particular interest that may warrant more detailed investigation with other, more sophisticated techniques (such as event-by-event particle detection with counter telescopes as described above). This approach has been particularly important for initial studies when a new energy regime or new projectile becomes available. For example, it has led to the identification of new phenomena and stimulated further investigation of the reactions induced by heavy ions and relativistic projectiles.

Nuclear chemists have played a major role in the determination of the yields, as well as the energy and angular distributions, of various reaction products resulting from the bombardment of a broad range of targets with projectiles varying in type from neutrons, protons, and pions to uranium ions and in energy from thermal (0.025 eV) to highly relativistic energies (500 GeV). These studies are well suited to the investigation of the modes of target breakup through such processes as nuclear fission, spallation, and others involving more extensive and violent target explosions. *See* NUCLEAR FISSION; SPALLATION REACTION.

Search for new nuclides. Although nearly 2100 radioactive nuclides have been identified, there are still gaps where nuclides are known to exist but have yet to be identified and unknown regions yet to be probed. In particular, the limits of neutron and proton stability in nuclei are actively being pursued. As these limits are approached, the half-lives of the nuclear species become very short (milliseconds to seconds) and specialized techniques have been developed for such investigations. These include (1) rapid, automated radiochemical separations; (2) gas transport techniques, in which the nuclide is entrained in a flowing gas stream and transported rapidly from the target to a detector; (3) target wheels which turn in a bombarding beam and carry the target rapidly and repeatedly from the activating beam position to detector positions; (4) moving tapes that collect the products recoiling from a thin target and rapidly transport them to detectors for measurement; (5) on-line isotope separators associated with heated targets (which vaporize the products and serve as the ion source for the separator) and various detectors at the collector position; and (6) other innovative methods.

One of the most exciting and active areas in the search for new nuclides is the attempt to extend the periodic table to elements beyond 107. In particular, theoretical prediction of a new "island of stability" near atomic number 114 and neutron number 184 has stimulated attempts to reach it. An intensive experimental program has been under way to form such a product at low enough excitation energy that it will survive long enough to be identified. Heavy-ion, neutron-rich projectiles (such as ^{48}Ca) and neutron-rich targets (such as ^{247}Cm) have been prepared in the hope that a fusion product in the expected region may be identified. *See* PERIODIC TABLE.

Technique of radiochemical analysis. The chemical manipulations and separations of radioactive materials, generally referred to as techniques of radiochemical analysis, differ from the ordinary analytical techniques in a number of respects. Procedures in radiochemical analysis are not required to provide quantitative recovery but are selected for specificity and speed, with reasonably good yields (usually the order of 50% or better) generally sufficing. The criteria of radiochemical purity in a radioactive preparation are somewhat more stringent than those of ordinary chemical purity. Thus, trace quantities of impurities are rarely of importance in ordinary quantitative analyses, but in a radioactive preparation, contamination by trace quantities of radioactive impurities may completely negate the results of an experiment.

The handling of highly radioactive materials presents a health hazard, and special techniques for the manipulation of samples behind shielding walls must be utilized. Some effects of high levels of radioactivity on the solutions, such as heating and the decomposition of solvents which produces bubbling, also may affect normal procedures.

The mass of radioactive material produced in nuclear reactions is usually very small. The concentrations of nuclear reaction products in the solutions of target materials are generally of the order of 10^{-10} M or less. Many normal chemical operations, such as precipitation, are not feasible with such small concentrations. Although separations can be carried out with these tracer quantities using such techniques as solvent extraction and ion exchange, it then is difficult to determine the efficiency for the recovery of the product. Moreover, the chemical behavior of such dilute solutions may differ considerably from that normally encountered. For example, radiocolloid formation, adsorption on the walls of vessels and on the surfaces of dust particles and precipitates, and the concentration dependence of some equilibrium constants become prominent at such extremely high dilution. To avoid these difficulties, an isotope dilution technique may be employed in which macroscopic quantities of stable isotopes of the element are added to serve as a carrier for the radioactive species. *See* ISOTOPE DILUTION TECHNIQUES.

The amount of carrier used represents a compromise between considerations of convenience in

chemical manipulations and yield determination and the preparation of high specific activity sources in which counting corrections for absorption and scattering of radiations in the sample itself are minimized. Quantities of 10–20 mg are used most often. Chemical procedures are simplified if macroscopic quantities of only a few elements are present. When many elements are produced in a nuclear reaction (in nuclear fission, for example), aliquots of the solution usually are taken for the analysis of each element or small group of elements. It is then necessary to add carriers for only relatively few products of interest. Trace quantities of the other elements present are removed in the chemical procedures by the use of scavenging precipitations of a compound of high surface area, such as iron(III) hydroxide or manganese dioxide, which tend to occlude traces of foreign substances or of a representative precipitate for an insoluble group of elements, such as bismuth sulfide to carry trace quantities of other insoluble sulfides, lanthanum fluoride for insoluble fluorides, or iron(III) hydroxide for insoluble hydroxides. If the element of interest itself forms a precipitate which may occlude traces of other elements, it may be necessary to add holdback carriers for the latter to dilute the effect of radioactive contamination of the product precipitate.

For the isolation of products of high specific activity without regard to yield, the carrier may be an element with chemical properties similar to those of the desired product, but which can be separated from it in the last stages of the procedure, leaving the product essentially carrier-free. Such carriers are referred to as nonisotopic carriers. When it is necessary to determine the yield of a nuclear reaction product, a known quantity of an isotopic carrier must be used. It is also imperative that complete interchange between the valence states of the carrier and the active species be achieved before any chemical separation is begun. In the case of elements which do not have any stable isotopes, or when a carrier-free procedure is desired, a known quantity of an isotopic radioactive tracer may be used. Radiations of the tracer should be easily distinguishable from those of the product. The fractional recovery of the added carrier or tracer then will represent the yield of the product of interest.

Classical precipitation methods of chemical separation are time-consuming, and in order to study short-lived radioactive species, rapid procedures are essential and other techniques are generally employed. Ion-exchange, solvent-extraction, and volatilization techniques have proved most useful, and their development is closely associated with radiochemical investigations. In cases where volatile products (rare gases, halogens, and so on) are formed in the nuclear reaction, a solution of the target may be bombarded in a closed system and the products swept out continuously. Some solid matrices that emanate gases readily may also be used. The products may be chemically separated by passing them through a suitable series of rapid steps in the flowing system before collection in a form suitable for direct measurement of the radiations. Nuclides with half-lives of the order of seconds may be investigated with such on-line techniques. Products recoiling out of thin targets may also be collected in gas streams or on moving tapes for rapid removal from the target and on-line or subsequent analysis. An isotope separator may form part of the on-line system, and in favorable cases the target may serve as the ion source of the separator. *See* ION EXCHANGE; RADIOCHEMICAL LABORATORY; SOLVENT EXTRACTION.

Sample preparation and counting techniques. For studies in nuclear chemistry, the object of the radiochemical separations is the preparation of a pure sample in a form suitable for the radioactive assay of the nuclide of interest or for the determination of its nuclear properties. The detector used will, of course, depend on the type of radiation involved and the kind of information desired.

Alpha particles and fission fragments have short ranges in matter, and to prevent absorption losses samples of less than 100 micrograms/cm^2 surface density are generally required. A uniform sample deposit is necessary for accurate alpha-particle and fission fragment measurements. This is best accomplished by volatilizing, electroplating, or spraying on metal foils. Samples collected with an isotope separator are well suited for such measurements. Evaporation from solution may also be used if the amount of solid residue is small, but uniformity of the deposit is difficult to achieve. The samples are counted internally in ionization chambers or gas proportional counters or with solid-state detectors.

Beta particles may cover a wide range of energies, and the techniques of sample preparation and counting will vary accordingly. The most commonly used detectors are Geiger, flow-type proportional, and scintillation counters. Samples may be prepared as indicated for alpha emitters, in the form of precipitates or filter-paper disks or sample cups, as gases for internal counting, and as liquids. External sample counting usually is employed for convenience whenever feasible. *See* GEIGER-MÜLLER COUNTER; SCINTILLATION COUNTER.

Gamma radiation is highly penetrating, and the size or form of the sample is generally not very critical. Because of much higher efficiency and resolution, scintillation counters and solid-state detectors have displaced all other types of detectors for gamma radiation.

Whenever possible it is advisable to design experiments so that relative counting of samples will suffice. It is then necessary only to reproduce the counting conditions for each sample. The determination of absolute disintegration rates is a more difficult task, and special techniques are required. Counters which detect all the radiations emanating from the source (4π counters) are used, and the samples are either dispersed in the counting medium of a proportional or Geiger counter as a gas, dissolved in the medium of a liquid scintillation counter, or counted as a very thin deposit on a very thin film which is placed between two identical counters. Beta-gamma coincidence counting may be used for determining absolute disintegration rates when the decay scheme of the nuclide is not too complicated.

Applied nuclear chemistry. The radiochemical and counting techniques outlined above are powerful tools for the study of nuclear reactions and the properties of nuclides. New techniques and instruments are constantly being adapted to the needs of nuclear chemistry and, conversely, investigations in nuclear chemistry have indicated the need and provided stimulation for the development of many new instruments. The techniques of nuclear chemistry have been applied to studies in a number of related fields, and nuclear chemists have contributed to studies in reactor chemistry and physics, isotope production, nuclear engineering, the reactor fuel/reprocessing cycle, nuclear weapons development, geo- and cosmochemistry, environmental chemistry, and accelerator beam studies and monitoring as well as to basic studies in chemistry and the life sciences and the industrial and agricultural application of radioisotopes. The field of analytical chemistry has been especially influenced by the technique of activation analysis, which utilizes many of the results and methods of nuclear chemistry. *See* ACTIVATION ANALYSIS.

The preparation of radioactively tagged compounds and radioactive tracers has been of particular importance in the field of nuclear medicine. The radioactive properties (half-life and type of energy of decay) and chemical or biochemical properties of tracers and tagged compounds can be designed and prepared to fit the needs of specific diagnostic or therapeutic requirements. *See* RADIOLOGY.

Ellis P. Steinberg

Bibliography. W. T. Carnall and G. R. Choppin (eds.), *Plutonium Chemistry*, 1980; G. R. Choppin and J. Rydberg (eds.), *Radiochemistry and Nuclear Chemistry: Theory and Applications*, 1994; G. Friedlander et al., *Nuclear and Radiochemistry*, 4th ed., 1981; B. G. Harvey, *Introduction to Nuclear Physics and Chemistry*, 2d ed., 1969; C. Hevesy and F. A. Paneth, *A Manual of Radioactivity*, 1938; E. K. Hyde, I. Perlman, and G. T. Seaborg, *The Nuclear Properties of the Heavy Elements*, 1964; M. Lefort, *Nuclear Chemistry*, 1968; J. D. Navratil, *Nuclear Chemistry*, 1993; G. T. Seaborg and W. Loveland (eds.), *Nuclear Chemistry*, 1982.

Nuclear engineering

The branch of engineering that deals with the production and use of nuclear energy and nuclear radiation. The multidisciplinary field of nuclear engineering is studied in many universities. In some it is offered in a special nuclear engineering department; in others it is offered in other departments, such as mechanical or chemical engineering. Primarily, nuclear engineering involves the conception, development, design, construction, operation, and decommissioning of facilities in which nuclear energy or nuclear radiation is generated or used.

Examples of facilities include nuclear power plants, which convert nuclear energy to electricity; nuclear propulsion reactors, which convert nuclear energy to mechanical force used for the propulsion of ships and submarines; space nuclear reactors, which convert nuclear energy to heat and electricity to power satellites, probes, and vehicles; nuclear production reactors, which produce fissile or fusile materials used in nuclear weapons; nuclear research reactors, which generate nuclear radiation (neutrons and gamma rays) for scientific research and for a variety of medical and industrial applications; gamma cells, which contain intense sources of gamma rays that are used for sterilizing medical equipment and food and for manufacturing polymers; particle accelerators, which produce nuclear radiation for use in medical and industrial applications; and repositories for safe storage of nuclear waste. *See* PARTICLE ACCELERATOR; SHIP NUCLEAR PROPULSION; SPACE POWER SYSTEMS; SUBMARINE.

Nuclear power plants. All operating nuclear power plants use the fission process. A nuclear power plant is a large, complex system consisting of a nuclear reactor, steam generators, heat exchangers, a turbogenerator, condensers, and many auxiliary systems. The central part of the nuclear reactor is the core, which contains the nuclear fuel and in which the nuclear chain reaction is established. The core is where fission energy is generated and converted to heat. This heat is removed from the reactor by a liquid or gaseous coolant and converted into electricity in the turbogenerator. *See* ELECTRIC POWER GENERATION; HEAT TRANSFER; NUCLEAR POWER; NUCLEAR REACTOR.

A typical reactor core is made up of fuel assemblies, which contain the fuel rods. A fuel rod consists of many fuel pellets enclosed in a sealed tube called clad. Uranium and other heavy nuclei in the fuel interact with passing neutrons and undergo fission reactions that split them into lighter nuclei and emit energy and neutrons. The succession of fissions due to absorption of neutrons emitted in previous fissions is the chain reaction. Spaces are provided between fuel assemblies or between fuel rods for control rods, which absorb neutrons. The control rods are inserted into or withdrawn from the core to control the chain reaction. *See* CHAIN REACTION (PHYSICS); NEUTRON; NUCLEAR FISSION; NUCLEAR FUELS; URANIUM.

Significant quantities of radioactive materials are produced in the reactor, most from the fission process itself and the remainder from the absorption, by various materials, of neutrons emitted in the fission process. Some of these radioactive materials have useful applications, and the rest are nuclear waste. *See* NUCLEAR REACTION; RADIOACTIVITY.

Nuclear facility design and operation. Nuclear engineering encompasses a wide variety of disciplines that need to collaborate for the conception, development, design, construction, operation, and decommissioning of nuclear facilities. These disciplines include reactor physics, thermal hydraulics, thermodynamics, mechanics, metallurgy, chemistry, control and instrumentation, radiation control and shielding, health physics, and economics. *See* NUCLEAR CHEMISTRY.

Reactor physicists determine the dimensions and composition of the core required for establishing a

chain reaction and for maintaining it at full power for a desirable duration—typically 18–24 months (after which the reactor is shut down for 3–6 weeks for refueling)—and determine the composition, number, and location of the control rods necessary for safe operation of the reactor. They also calculate the fission rate at any given point in the core; the rate of change of the chain reaction due to variation in operating conditions, such as in the coolant inlet temperature and flow rate; the change in the fuel composition as a function of time; and the rate of interaction of neutrons with different materials. Finally, reactor physicists design shields against the leakage of nuclear radiation. *See* NUCLEAR PHYSICS; RADIATION SHIELDING; REACTOR PHYSICS.

Thermal hydraulics engineers determine how much coolant is needed in the core volume; the coolant flow rate; the size of the pumps required for circulating the coolant; and the temperature anywhere in the core and in sensitive components, such as the pressure vessel. Thermal hydraulics engineers and reactor physicists closely collaborate in analyzing reactor safety, using computer simulation to predict changes in the core power level and temperature distribution due to a variety of accident scenarios. The basic safety requirement is that no conceivable accident should lead to a breach in the fuel rod clad in the reactor vessel and in the reactor containment structures that could result in a release of radioactive materials to the environment.

Thermodynamic engineers design the systems needed to convert nuclear-generated heat to electricity at maximum efficiency.

Structural engineers perform mechanical analyses of many important components in nuclear power plants to determine the thickness of components so that the stresses they are subjected to will not exceed permissible limits. Among the most sensitive components are the fuel rod clad and the reactor pressure vessel.

Metallurgical engineers develop nuclear fuel and clad that can reliably operate in the reactor for long periods of time; predict the change in the mechanical properties of the nuclear fuel due to the accumulation of fission products and the relatively high temperature at the center of the fuel pellets in the operating reactor; predict the change in the mechanical properties of structural materials, as a result of being continuously bombarded by the neutrons emitted in the fission process; and develop structural materials that will maintain good mechanical strength at high temperatures and resist the corrosive effects of the coolant. *See* RADIATION DAMAGE TO MATERIALS.

Chemical engineers develop processes for converting uranium ores into the uranium dioxide used to manufacture the fuel pellets; for uranium isotopic enrichment; for processing spent fuel from the reactor; and for safe storage of nuclear waste. *See* DECONTAMINATION OF RADIOACTIVE MATERIALS; ISOTOPE SEPARATION; NUCLEAR CHEMICAL ENGINEERING; NUCLEAR FUEL CYCLE; NUCLEAR FUELS REPROCESSING.

Instrumentation and control engineers design detectors for monitoring the reactor's operating conditions; design systems to process the signals from these detectors and transmit them to the central control room; and design systems to move the reactor control rods, change coolant flow rates, and open or close valves in response to the detectors' signals or to signals provided by the power-plant operators. They also design the control room where operators obtain detailed information on the power plant's condition and can modify this condition.

Health physicists and radiation control specialists monitor the radiation level of neutrons and gamma rays anywhere in the power plant accessible to the staff. They also measure the radiation doses to which staff members are exposed, to ensure that they will not exceed the permitted dose rate. *See* GAMMA RAYS; HEALTH PHYSICS; RADIATION INJURY (BIOLOGY).

Economists estimate the cost of generating electricity in the nuclear power plant, taking into account a host of factors such as the investment in the design and construction of the power plant; the interest on that investment; the time it takes to construct the plant; the cost of the nuclear fuel; the amount of energy generated by the nuclear fuel; the duration of plant shutdown for refueling and for maintenance; the cost of the operation and maintenance of the power plant.

Non–power plant applications. Starting in the 1990s, a significant number of nuclear engineers were engaged in activities associated with the design of facilities for the safe storage of spent fuel (the fuel that has been taken out of nuclear reactors) and of other nuclear waste; the design of underground repositories for safely storing the long-lived nuclear waste; and the design of special facilities for reducing the volume of and hazard from the nuclear waste. *See* RADIOACTIVE WASTE MANAGEMENT.

Nuclear engineers are also involved in the production of radioactive isotopes for medical and industrial applications. Additional nuclear engineering activities which are not associated with nuclear power plants include the design of facilities which use nuclear radiation for the diagnosis and treatment of cancers and other illnesses; for the detection of land mines, other explosives, and drugs; for the diagnosis of imperfections in sensitive components, such as aircraft components; for food sterilization; for the measurement of underground moisture level and automatic irrigation control; for the search for oil and other minerals; as well as for many other industrial and agricultural applications. *See* NUCLEAR MEDICINE; RADIOACTIVITY AND RADIATION APPLICATIONS; RADIOISOTOPE.

Fusion research and development. Many nuclear engineers are also involved in the research and development of future fusion power plants—plants that will be based on the fusion reaction for generating nuclear energy. Many challenging engineering problems are involved, including the development of large powerful magnets and the development of technologies for heating the fusion fuel to hundreds of millions of degrees; confining this ultrahot fuel;

and compressing submillimeter spherical capsules containing fusion fuel to many thousand times their natural solid density. New structural and coolant materials and new chemical processes are also being developed. *See* NUCLEAR FUSION. Ehud Greenspan

Bibliography. K. Almeras and R. Lee, *Nuclear Engineering: An Introduction*, 1992; S. Glasstone and A. Sesonske, *Nuclear Reactor Engineering*, 2 vols., 4th ed., 1994; R. A. Knief, *Nuclear Engineering: Theory and Technology of Commercial Nuclear Power*, 2d ed., 1992; J. R. Lamarsh and A. Baratta, *Introduction to Nuclear Engineering*, 3d ed., 2001; R. L. Murray, *Nuclear Energy*, 5th ed., 2000.

Nuclear explosion

An explosion whose energy is produced by a nuclear transformation, either fission (**Fig. 1**) or fusion (**Fig. 2**). *See* NUCLEAR FISSION; NUCLEAR FUSION.

This article focuses on the effects of nuclear explosions. For discussions of how these explosions are produced *See* ATOMIC BOMB; HYDROGEN BOMB.

The energy of a nuclear explosion is usually stated in terms of the mass of trinitrotoluene (TNT) which would provide the same energy. The complete fissioning of 1 kg of uranium or plutonium would be equivalent to 17,000 metric tons of TNT (17 kilotons); 2 lb would be equivalent to 17,000 short tons. The indicated yield-to-mass ratio of 1.7×10^7 cannot be realized, largely because of the ancillary equipment necessary to assemble the nuclear components into an explosive configuration in the very short time required. The first nuclear weapons (1945) weighed about 5 tons, but with yields of 15 to 20 kT their yield-to-weight ratio was, nevertheless, close to 4000 times larger than that of previous weapons. By 1960 the United States had developed a weapon with a yield of about 1 megaton (MT) in a weight of about 1 ton for use in an intercontinental missile. With this, the yield-to-weight ratio was raised to 10^6. Though improvements may be available, and may have been realized, there is no room for further changes of the qualitatively significant sort already achieved.

Fig. 2. Nuclear explosion "Mike," the first test of a thermonuclear device with a yield of about 10 megatons, detonated on the surface at Eniwetok Atoll on October 31, 1952. (*Los Alamos National Laboratory*)

Fig. 1. Nuclear explosion "Priscilla," with a yield of about 37 kilotons, detonated from a balloon on June 24, 1957, at the Nevada test site. (*Los Alamos National Laboratory*)

Although weapons with yields of up to 15 and 60 MT were fired by the United States and Soviet Union respectively, about 1960 the main interest of the major nuclear powers focused on adapting nuclear devices for delivery by missile carriers, and this called for smaller weapons with so-called moderate yields of tens or hundreds of kilotons, up to a few megatons. *See* MISSILE.

It is sometimes supposed that nuclear weapons are distinguished by the large amounts of energy released. In fact, there have been a number of accidental chemical explosions, particularly of ships loaded with ammunition or of large accumulations of ammonium nitrate, with energies in the range of 1–3 kT. In addition, since the great explosion of Krakatoa (1883), there have been a dozen or so volcanic explosions with energies in the tens of megatons. Though the size of a typical nuclear explosion is appalling, the most significant feature of a nuclear device derives from its yield-to-weight ratio. *See* EXPLOSIVE.

The damage mechanism of a conventional explosion is blast—the pressure, or shock, wave transmitted in the surrounding medium. The blast wave from a nuclear explosion is similar, except for the great difference in scale. A nuclear explosion also produces several kinds of effects not experienced with ordinary explosives.

Blast. The destructive action of blast correlates with the peak overpressure in the wave. The pressure near the surface is of primary interest. This is affected by the reflection that occurs at the surface as well as by the height of the explosion. From 1 kT exploded on the surface, the distance to which an overpressure of 5 lb/in.^{2} (34.5 kilopascals) extends is about 1500 ft (450 m). This pressure extends to 2300 ft (700 m) from ground zero for an explosion at a height of approximately 800 ft (250 m). For heights greater than 1800 ft (550 m), 1 kT does not result in 5 lb/in.^{2} on the surface. For a given yield and specified overpressure, there is a height of burst which maxmizes the area exposed to the chosen overpressure. For very high pressures (hundreds of pounds per square inch), a low height is necessary, and a surface burst may do about as well.

The distance at which some stated overpressure will be realized scales with the cube root of the yield, providing the height of burst is similarly scaled. From the above example for 1 kT, the 5-lb/in.^{2} (34.5-kPa) radius for a 1-megaton explosion at 8000 ft (2500 m) above the surface would be 23,000 ft (7000 m), and this radius would be 230 ft (70 m) for a 1-ton bomb detonated at a height of 80 ft (25 m).

Some heavy structures can withstand an overpressure of 5 lb/in.^{2} (34.5 kPa), but such a blast would destroy most commercial buildings, apartments, and residences. People in residences exposed to 5 lb/in.^{2} would have about a 50% chance of being killed. In many studies it is assumed that the number of persons surviving within the 5-lb/in.^{2} ring is equal to the number killed in the area beyond, and the number of fatalities is taken to be equal to the number present inside the ring.

This is consistent with the experience at Hiroshima where, with a yield of 15 kT fired at a height of 1850 ft (560 m), the 5-lb/in.^{2} (34.5-kPa) radius was about 5700 ft (1730 m). The number of early fatalities was about equal to the number of people estimated to have been present within about 1 mi (1.6 km) of ground zero.

Damage at decreasing levels extends to larger distances, and fatal injuries may still occur. Out to 2.5 times the 5-lb/in^2 (34.5-kPa) range, where pressures would be only 1 lb/in.^{2} (7 kPa), panes of window glass will be shattered into flying splinters.

Thermal radiation. Very high energy densities are realized in nuclear explosions. Consequently, when the energy generation is complete, the materials of the device constitute a ball of very hot gas with temperatures in excess of 10^7 K (2×10^7 °F). This material expands violently, engulfing the air it encounters, increasing the mass and volume of the hot zone while the temperature drops accordingly. By the time the expanding front has reached a radius of 33 ft (10 m), the mass of air entrained will be more than 5 tons, so that beyond this stage the product of the explosion behaves as a ball of strongly heated air. The hot globe radiates energy as a blackbody, the temperature determining the spectrum of quanta. Initially these quanta, with energies in the kilovolt range, are absorbed in an inch or two of air at sea-level density. Most of the radiation at this stage merely heats up the air being engulfed by the hydrodynamic process. As the falling temperature approaches 10^4 K (2×10^4 °F), the spectrum peaks in the visible range, the range of wavelengths for which air is transparent. The energy radiated in this band travels through the air until it impinges on some absorbing surface. As the temperature falls toward 3000°F (2000 K), the emission of thermal radiation (which varies as the fourth power of the temperature) ceases to be significant. About one-third of the energy of the explosion is distributed as thermal radiation on line-of-sight trajectories. *See* HEAT RADIATION.

Thermal radiation is not important in connection with conventional explosives since the temperature of freshly burned chemical explosives (about 5000°F or 3000 K) is too low for an appreciable fraction of the energy to be radiated.

Exposure to 5–10 cal/cm^2 (2–4 $\times$ 10^5 J/m^2) of thermal radiation energy in the short time (a second, or so) during which the thermal pulse is delivered will ignite many combustible materials (fabrics, paper, dry leaves, and so forth). It will cause serious flash burns on exposed skin. Such energy levels will be delivered in clear air to 0.3–0.4 mi (0.5–0.6 km) by an explosion of 1 kT. The thermal energy is proportional to the yield at a given distance, and falls off (nearly) as the square of the distance. The ranges will be reduced by factors which reduce "visibility" in the atmosphere. At Hiroshima, burn injuries alone would have been fatal to almost all persons in the open without protection out to a little over 1 mi (1.6 km), and burns serious enough to require treatment were experienced at distances greater than 2 mi (3.2 km).

Another consequence of the thermal radiation is flash blindness. The eye, after being exposed to an intense burst of light, may not be able to receive new images for some seconds, or even for hours. A person need not be facing the burst point to be dazzled. The range at which this effect could occur is much greater than that at which burns would be experienced. It is also much greater at night, when the eye is dark-adapted, than in daylight. Intense light from the explosion can cause permanent burns on the retina in those who view the explosion directly.

Prompt ionizing radiation. Neutrons and gamma rays are released from a nuclear explosion. Since the fissions occur before the weapon materials expand, almost all the gamma rays accompanying fission are absorbed in these materials. But as soon as the hot ball of vaporized weapon components expands, the gamma rays being emitted by the radioactive fission fragments can escape to the air. During the few tens of seconds before the fireball rises away from the point at which the explosion occurred, it provides an intense source of gamma rays. The radiation emitted during this interval is referred to as prompt radiation. The remaining radioactivity, which is swept upward from the scene of the explosion to the altitude at which the fireball stops rising, and some of which ultimately returns to the surface,

constitutes the residual radiation. *See* ALPHA PARTICLES; NEUTRON; RADIOACTIVITY.

Neutrons and gamma rays are attenuated by passage through the air, the gamma rays by a factor of e (2.718) in a distance of about 1080 ft (330 m), the neutrons in about 650 ft (200 m). At distances greater than 3300 ft (1000 m) the gamma rays dominate, and the prompt radiation exposure varies as $(1/R^2)e^{-R/a}$, where R is the distance from the explosion and a is 1080 ft or 330 m. The biological damage caused by radiation results from the deposition of energy in tissue. *See* RADIATION INJURY (BIOLOGY).

At Hiroshima a dose equal to or greater than 450 rads (4.5 grays) extended to almost 1 mi (1.6 km). About half of the persons exposed to this dose in a short time will die within a few weeks, so that persons in this area who were not protected by heavy building walls experienced severe hazard from radiation. This is about the same distance for severe hazards from blast and thermal effects. From the scaling laws discussed above, it follows that the prompt radiation exposure, with its exponentially decreasing term, falls off more rapidly than the blast effect which, in turn, falls more rapidly than the intensity of thermal radiation. For an explosion much larger than 15 or 20 kT, the hazard range from thermal radiation or blast will be larger than that from prompt radiation, and prompt radiation will be a relatively unimportant effect. For much smaller yields, this order of importance will be reversed. To illustrate: from a 1-MT explosion the ranges of 500 rads (5 grays), 5 lb/in.^{2} (34.5 kPa), and 10 cal/(cm^2 · s) [4.2×10^5 W/m^2] are, respectively, about 1.6, 4.4, and 7.5 mi (2.5, 7, and 12 km); from 1 kT explosion they are 0.5, 0.4, and 0.3 mi (0.8, 0.7, and 0.5 km).

Residual ionizing radiation. For an explosion high enough above the surface that the fireball does not reach the ground, the only condensable material carried aloft is the small mass of the weapon itself. This is mixed with an enormous quantity of heated air. When the material cools to the point where the normally solid materials recondense, the particles formed are very small. The fission fragments, with the main exception of the gaseous elements xenon and krypton, will be incorporated in these particles. They will descend very slowly, reaching the surface only in days or months after the explosion. Most of these particles are distributed over the hemisphere in which the explosion occurred. The residual radioactivity distributed in this way would add a small component to the natural background radioactivity. Occasional exceptions to this would occur should rain, falling through or from the debris cloud before it had dispersed widely, bring an appreciable fraction of the radioactivity to the surface in a small area.

The situation is different for an explosion on or very close to the surface. In this case a large mass of surface material will be vaporized or swept up into the rising fireball. Much larger particles will be formed, which will fall back relatively rapidly, with the result that about half the residual radioactivity may be deposited on the surface in the first 24 h and within a few hundred miles in the downwind direction. Assuming a steady 15 mi/h (7 m/s) wind this local fallout from a 1-MT surface burst could have the effect that an unprotected person who remained in place on the edge of a cigar-shaped contour about 90 mi (145 km) long and up to 13 mi (21 km) wide would receive a radiation dose in excess of 450 rads (4.5 grays) by 24 h after the explosion, and another 450 rads by the end of a week. At points inside the contour higher exposures would apply, while in a large area outside the contour exposures would still be at serious levels. To avoid death or injury, persons in an area several times larger than the contour would have to evacuate quickly or find adequate shelter for an extended period. Some areas (extending up to a few tens of miles in the downwind direction) could not safely be restored to use for several years. In actual situations, where variability in the wind is likely, the contour of the area involved will be quite irregular. *See* RADIOACTIVE FALLOUT.

Electromagnetic pulse (EMP). The gamma rays emitted from an explosion will collide with electrons in the air and create a sheath of electric current. This will never be fully symmetric, so that at least some electromagnetic radiation will be generated. At very high altitudes the variation in the density of the atmosphere provides a large-scale asymmetry which is important in this respect. A 1-MT explosion at an altitude 200–300 mi (300–500 km) over a point in Nebraska would result in an electromagnetic pulse in excess of 25,000 V/m over the entire area of the contiguous United States. Such a pulse would induce currents in power lines, antennas, and exposed conductors generally. Unless these, or the circuits connected to them, were appropriately shielded, the effects could be similar to those sometimes caused by lightning: burnt-out components, short circuits, opened breakers, and so forth. Nationwide interruptions or breakdowns could occur on communication networks, power systems, and control equipment. *See* ELECTROMAGNETIC PULSE (EMP).

Radio and radar. As a result of the radioactivity of the fission fragments, the cloud containing the weapon debris will be in an ionized state for a considerable period. An explosion can also cause temporary regional changes in the ionosphere. The operation of long-range radio communication or radar observation in channels affected by such perturbations could be degraded or blocked.

Ozone depletion. At the very high temperatures in the fireball of a nuclear explosion, chemical reactions occur in normal air. By the time the fireball cools, nitric oxide composes about 1% of its molecules. Nitric oxide can catalyze the destruction of ozone. It is only from rather large explosions (greater than 1 MT) that the fireball rises to the altitude (greater than 12 mi or 20 km) of the Earth's ozone blanket in the stratosphere. Hundreds of large explosions would be required to introduce enough nitric oxide into the ozone to deplete it appreciably. However, nuclear weapon stockpiles favor weapons having relatively small yields. An ozone catastrophe in the event of a nuclear war may no longer

belong on a list of realistic threats. *See* STRATOSPHERIC OZONE.

Worldwide fallout. A large-scale nuclear war was usually pictured as involving several thousand megatons of nuclear explosions in a short period, mainly on, or over, the territory of the United States and the Soviet Union. Following such a war, residual radiation would be deposited over the remaining area of the Northern Hemisphere, particularly in the North Temperate Zone. Radiation exposure at levels lower than those causing any immediate perceptible ill effects can still result in cancer after several years, and in genetic effects. *See* MUTATION.

Estimates of the magnitude of such effects have ranged from the possible extinction of biological life down to quite mild levels. Though such calculations can be made, the results depend strongly on many assumptions. The best estimates at present suggest that over the three or four decades following a nuclear war some millions of deaths would result from delayed radiation effects outside the target countries. Though this would appear to be a very large number, it would still represent only a small fractional increase in the normal incidence of fatal cancer. *See* CANCER (MEDICINE).

Nuclear winter. Many large urban fires and massive forest or brush fires could be started by a nuclear war. Large amounts of soot and smoke would be injected into the atmosphere by these near-simultaneous fires. Such smoke clouds would merge and be distributed around the hemisphere, and they could persist for weeks. They would reduce the amount of sunlight reaching the surface and could cause a precipitous drop in temperature in the midcontinental regions of the whole temperate zone. By some estimates, temperatures would average below 0°F (−20°C) for months or more, the amount of cooling and the duration depending on the magnitude of the war, but also on many of the assumptions necessary for the estimates. Such an outcome of a possible nuclear war has come to be referred to as nuclear winter. Even some fraction of such an effect would be serious.

J. Carson Mark

Bibliography. S. Glasstone and P. J. Dolan (eds.), *The Effects of Nuclear Weapons*, 1977, reprint 1983; C. C. Grace, *Nuclear Weapons: Principles, Effects, and Survivability*, 1993; M. A. Harwell and T. C. Hutchinson, *Environmental Consequences of Nuclear War*, vol. 2, 2d ed., 1989; Office of Technology Assessment, Congress of the United States, *The Effects of Nuclear War*, 1980; A. B. Pittock et al., *Environmental Consequences of Nuclear War*, vol. 1, 2d ed., 1989.

Nuclear fission

An extremely complex nuclear reaction representing a cataclysmic division of an atomic nucleus into two nuclei of comparable mass. This rearrangement or division of a heavy nucleus may take place naturally (spontaneous fission) or under bombardment with neutrons, charged particles, gamma rays, or other carriers of energy (induced fission). Although nuclei with mass number A of approximately 100 or greater are energetically unstable against division into two lighter nuclei, the fission process has a small probability of occurring, except with the very heavy elements. Even for these elements, in which the energy release is of the order of 2×10^8 eV, the lifetimes against spontaneous fission are reasonably long. *See* NUCLEAR REACTION.

Liquid-drop model. The stability of a nucleus against fission is most readily interpreted when the nucleus is viewed as being analogous to an incompressible and charged liquid drop with a surface tension. Such a droplet is stable against small deformations when the dimensionless fissility parameter X in Eq. (1) is less than unity, where the charge is in

$$X = \frac{(\text{charge})^2}{10 \times \text{ volume} \times \text{surface tension}} \qquad (1)$$

esu, the volume is in cm^3, and the surface tension is in $ergs/cm^2$. The fissility parameter is given approximately, in terms of the charge number Z and mass number A, by the relation $X = Z^2/47\,A$.

Long-range Coulomb forces between the protons act to disrupt the nucleus, whereas short-range nuclear forces, idealized as a surface tension, act to stabilize it. The degree of stability is then the result of a delicate balance between the relatively weak electromagnetic forces and the strong nuclear forces. Although each of these forces results in potentials of several hundred megaelectronvolts, the height of a typical barrier against fission for a heavy nucleus, because they are of opposite sign but do not quite cancel, is only 5 or 6 MeV. Investigators have used this charged liquid-drop model with great success in describing the general features of nuclear fission and also in reproducing the total nuclear binding energies. *See* NUCLEAR BINDING ENERGY; NUCLEAR STRUCTURE; SURFACE TENSION.

Shell corrections. The general dependence of the potential energy on the fission coordinate representing nuclear elongation or deformation for a heavy nucleus such as ^{240}Pu is shown in **Fig. 1**. The expanded scale used in this figure shows the large decrease in energy of about 200 MeV as the fragments separate to infinity. It is known that ^{240}Pu is deformed in its ground state, which is represented by the lowest minimum of −1813 MeV near zero deformation. This energy represents the total nuclear binding energy when the zero of potential energy is the energy of the individual nucleons at a separation of infinity. The second minimum to the right of zero deformation illustrates structure introduced in the fission barrier by shell corrections, that is, corrections dependent upon microscopic behavior of the individual nucleons, to the liquid-drop mass. Although shell corrections introduce small wiggles in the potential-energy surface as a function of deformation, the gross features of the surface are reproduced by the liquid-drop model. Since the typical fission barrier is only a few megaelectronvolts, the magnitude of the shell correction need only be small for irregularities to be

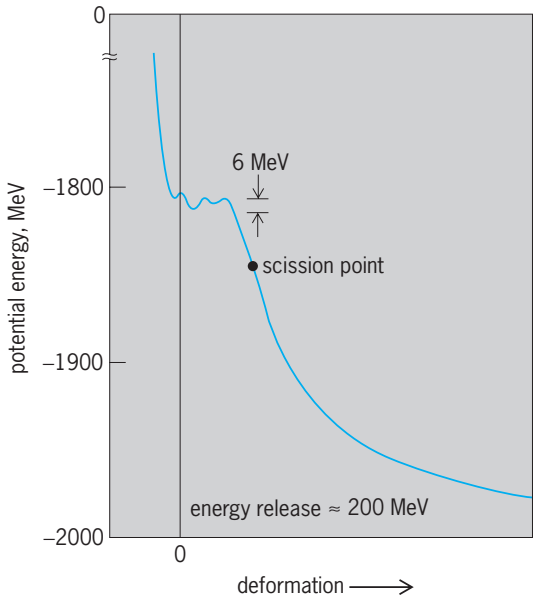

Fig. 1. Plot of the potential energy in MeV as a function of deformation for the nucleus ^{240}Pu. (*After M. Bolsteli et al., New calculations of fission barriers for heavy and superheavy nuclei, Phys. Rev., 5C:1050–1077, 1972*)

introduced into the barrier. This structure is schematically illustrated for a heavy nucleus by the double-humped fission barrier in **Fig. 2**, which represents the region to the right of zero deformation in Fig. 1 on an expanded scale. The fission barrier has two maxima and a rather deep minimum in between. For comparison, the single-humped liquid-drop barrier is also schematically illustrated. The transition in the shape of the nucleus as a function of deformation is represented in the upper part of the figure.

Double-humped barrier. The developments which led to the proposal of a double-humped fission barrier were triggered by the experimental discovery of

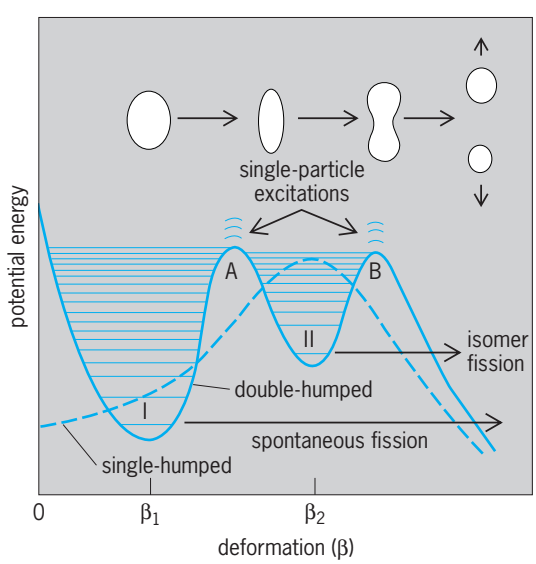

Fig. 2. Schematic plots of single-humped fission barrier of liquid-drop model and double-humped barrier introduced by shell corrections. (*After J. R. Huizenga, Nuclear fission revisited, Science, 168:1405–1413, 1970*)

spontaneously fissionable isomers by S. M. Polikanov and colleagues in the Soviet Union and by V. M. Strutinsky's pioneering theoretical work on the binding energy of nuclei as a function of both nucleon number and nuclear shape. The double-humped character of the nuclear potential energy as a function of deformation arises, within the framework of the Strutinsky shell-correction method, from the superposition of a macroscopic smooth liquid-drop energy and a shell-correction energy obtained from a microscopic single-particle model. Oscillations occurring in this shell correction as a function of deformation lead to two minima in the potential energy, shown in Fig. 2, the normal ground-state minimum at a deformation of β_1 and a second minimum at a deformation of β_2. States in these wells are designated class I and class II states, respectively. Spontaneous fission of the ground state and isomeric state arises from the lowest-energy class I and class II states, respectively. *See* NUCLEAR ISOMERISM.

The calculation of the potential-energy curve illustrated in Fig. 1 may be summarized as follows, The smooth potential energy obtained from a macroscopic (liquid-drop) model is added to a fluctuating potential energy representing the shell corrections, and to the energy associated with the pairing of like nucleons (pairing energy), derived from a non-self-consistent microscopic model. The calculation of these corrections requires several steps: (1) specification of the geometrical shape of the nucleus, (2) generation of a single-particle potential related to its shape, (3) solution of the Schrödinger equation, and (4) calculation from these single-particle energies of the shell and pairing energies.

The oscillatory character of the shell corrections as a function of deformation is caused by variations in the single-particle level density in the vicinity of the Fermi energy. For example, the single-particle levels of a pure harmonic oscillator potential arrange themselves in bunches of highly degenerate shells at any deformation for which the ratio of the major and minor axes of the spheroidal equipotential surfaces is equal to the ratio of two small integers. Nuclei with a filled shell, that is, with a level density at the Fermi energy that is smaller than the average, will then have an increased binding energy compared to the average, because the nucleons occupy deeper and more bound states; conversely, a large level density is associated with a decreased binding energy. It is precisely this oscillatory behavior in the shell correction that is responsible for spherical or deformed ground states and for the secondary minima in fission barriers, as illustrated in Fig. 2. *See* NONRELATIVISTIC QUANTUM THEORY.

More detailed theoretical calculations based on this macroscopic-microscopic method have revealed additional features of the fission barrier. In these calculations the potential energy is regarded as a function of several different modes of deformation. The outer barrier B (Fig. 2) is reduced in energy for shapes with pronounced left-right asymmetry (pear shapes), whereas the inner barrier A and deformations in the vicinity of the second minimum are stable

against such mass asymmetric degrees of freedom. Similar calculations of potential-energy landscapes reveal the stability of the second minimum against gamma deformations, in which the two small axes of the spheroidal nucleus become unequal, that is, the spheroid becomes an ellipsoid.

Experimental consequences. The observable consequences of the double-humped barrier have been reported in numerous experimental studies. In the actinide region, more than 30 spontaneously fissionable isomers have been discovered between uranium and berkelium, with half-lives ranging from 10^{-11} to 10^{-2} s. These decay rates are faster by 20 to 30 orders of magnitude than the fission half-lives of the ground states, because of the increased barrier tunneling probability (Fig. 2). Several cases in which excited states in the second minimum decay by fission are also known. Normally these states decay within the well by gamma decay; however, if there is a hindrance in gamma decay due to spin, the state (known as a spin isomer) may undergo fission instead.

Qualitatively, the fission isomers are most stable in the vicinity of neutron numbers 146 to 148, a value in good agreement with macroscopic-microscopic theory. For elements above berkelium, the half-lives become too short to be observable with available techniques; and for elements below uranium, the prominent decay is through barrier A into the first well, followed by gamma decay. It is difficult to detect this competing gamma decay of the ground state in the second well (called a shape isomeric state), but identification of the gamma branch of the 200-ns ^{238}U shape isomer has been reported. *See* RADIOACTIVITY.

Direct evidence of the second minimum in the potential-energy surface of the even-even nucleus ^{240}Pu has been obtained through observations of the E2 transitions within the rotational band built on the isomeric 0+ level. The rotational constant (which characterizes the spacing of the levels and is expected to be inversely proportional to the effective moment of inertia of the nucleus) found for this band is less than one-half that for the ground state and confirms that the shape isomers have a deformation β_2 much larger than the equilibrium ground-state deformation β_1. From yields and angular distributions of fission fragments from the isomeric ground state and low-lying excited states, some information has been derived on the quantum numbers of specific single-particle states of the deformed nucleus (Nilsson single-particle states) in the region of the second minimum.

At excitation energies in the vicinity of the two barrier tops, measurements of the subthreshold neutron fission cross sections of several nuclei have revealed groups of fissioning resonance states with wide energy intervals between each group where no fission occurs. Such a spectrum is illustrated in **Fig. 3***a*, where the subthreshold fission cross section of ^{240}Pu is shown forneutron energies between

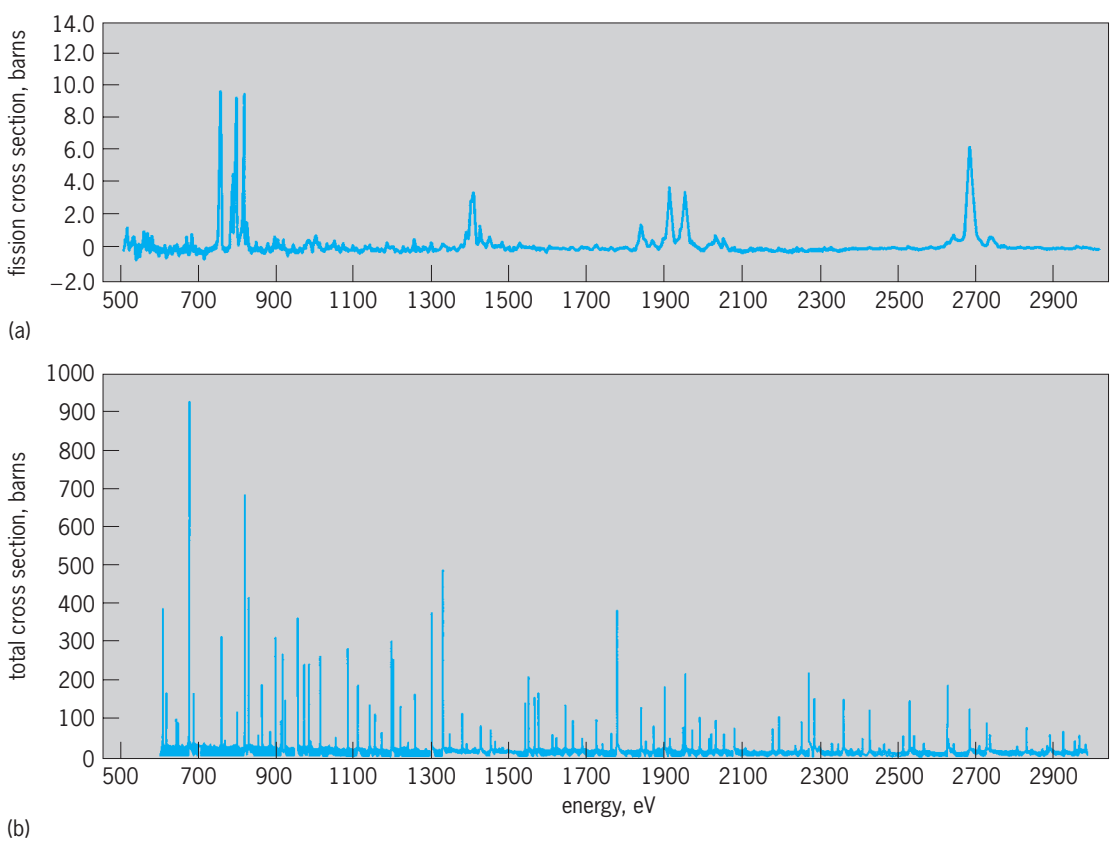

Fig. 3. Grouping of fission resonances demonstrated by (*a*) the neutron fission cross section of ^{240}Pu and (*b*) the total neutron cross section. (*After V. M. Strutinsky and H. C. Pauli, Shell-structure effects in the fissioning nucleus, Proceedings of the 2d IAEA Symposium on Physics and Chemistry of Fission, Vienna, pp. 155–177, 1969*)

500 and 3000 eV. As shown in Fig. 3*b*, between the fissioning resonance states there are many other resonance states, known from data on the total neutron cross sections, which have negligible fission cross sections. Such structure is explainable in terms of the double-humped fission barrier and is ascribed to the coupling between the compound states of normal density in the first well to the much less dense states in the second well. This picture requires resonances of only one spin to appear within each intermediate structure group illustrated in Fig. 3*a*. In an experiment using polarized neutrons on a polarized ^{237}Np target, it was found that all nine fine-structure resonances of the 40-eV group have the same spin and parity: $I = 3+$. Evidence has also been obtained for vibrational states in the second well from neutron (n, f) and deuteron stripping (d, pf) reactions at energies below the barrier tops (f indicates fission of the nucleus). *See* NEUTRON SPECTROMETRY.

A. Bohr suggested that the angular distributions of the fission fragments are explainable in terms of the transition-state theory, which describes a process in terms of the states present at the barrier deformation. The theory predicts that the cross section will have a steplike behavior for energies near the fission barrier, and that the angular distribution will be determined by the quantum numbers associated with each of the specific fission channels. The theoretical angular distribution of fission fragments is based on two assumptions. First, the two fission fragments are assumed to separate along the direction of the nuclear symmetry axis so that the angle θ between the direction of motion of the fission fragments and the direction of motion of the incident bombarding particle represents the angle between the body-fixed axis (the long axis of the spheroidal nucleus) and the space-fixed axis (some specified direction in the laboratory, in this case the direction of motion of the incident particle). Second, it is assumed that during the transition from the saddle point (corresponding to the top of the barrier) to scission (the division of the nucleus into two fragments) the Coriolis forces do not change the value of K (where K is the projection of the total angular momentum I on the nuclear symmetry axis) established at the saddle point.

In several cases, low-energy photofission and neutron fission experiments have shown evidence of a double-humped barrier. In the case of two barriers, the question arises as to which of the two barriers A or B is responsible for the structure in the angular distributions. For light actinide nuclei like thorium, the indication is that barrier B is the higher one, whereas for the heavier actinide nuclei, the inner barrier A is the higher one. The heights of the two barriers themselves are most reliably determined by investigating the probability of induced fission over a range of several megaelectronvolts in the threshold region. Many direct reactions have been used for this purpose, for example, (d, pt), (t, pf), and (^{3}He, df). There is reasonably good agreement between the experimental and theoretical barriers. The theoretical barriers are calculated with realistic single-particle potentials and include the shell corrections.

Fission probability. The cross section for particle-induced fission $\sigma(y, f)$ represents the cross section for a projectile y to react with a nucleus and produce fission, as shown by Eq. (2). The quantities $\sigma_R(y)$, Γ_f,

$$\sigma(y, f) = \sigma_R(y)\frac{\Gamma_f}{\Gamma_t} \tag{2}$$

and Γ_t are the total reaction cross sections for the incident particle y, the fission width, and the total level width, respectively, where $\Gamma_t = \Gamma_f + \Gamma_n + \Gamma_y + \cdots$ is the sum of all partial-level widths. All the quantities in Eq. (2) are energy-dependent. Each of the partial widths for fission, neutron emission, radiation, and so on, is defined in terms of a mean lifetime τ for that particular process, for example, $\Gamma_f = \hbar/\tau_f$. Here $\hbar$, the action quantum, is Planck's constant divided by 2π and is numerically equal to 1.0546×10^{-34} J s $= 0.66 \times 10^{-15}$ eV s. The fission width can also be defined in terms of the energy separation D of successive levels in the compound nucleus and the number of open channels in the fission transition nucleus (paths whereby the nucleus can cross the barrier on the way to fission), as given by expression (3), where I is the angular momentum and i is an index

$$\Gamma_f(I) = \frac{D(I)}{2\pi}\sum_i N_{fi} \tag{3}$$

labeling the open channels N_{fi}. The contribution of each fission channel to the fission width depends upon the barrier transmission coefficient, which, for

Cross sections for neutrons of thermal energy to produce fission or undergo capture in the principal nuclear species, and neutron yields from these nuclei*

Nucleus	Cross section for fission, σ_f, 10^{-24} cm^2	σ_f plus cross section for radiative capture, σ_r	Ratio, $1 + \alpha$	Number of neutrons released per fission, ν	Number of neutrons released per slow neutron captured, $\eta = \nu/(1 + \alpha)$
^{233}U	525 ± 2	573 ± 2	1.093 ± 0.003	2.50 ± 0.01	2.29 ± 0.01
^{235}U	577 ± 1	678 ± 2	1.175 ± 0.002	2.43 + 0.01	2.08 ± 0.01
^{239}Pu	741 ± 4	1015 ± 4	1.370 ± 0.006	2.89 ± 0.01	2.12 ± 0.01
^{238}U	0	2.73 ± 0.04			0
Natural uranium	4.2	7.6	1.83	2.43 ± 0.01	1.33

*Data from *Brookhaven National Laboratory* 325, 2d ed., suppl. no. 2, vol. 3, 1965. The data presented are the recommended or least-square values published in this reference for 0.0253-eV neutrons.

a two-humped barrier (see Fig. 2), is strongly energy-dependent. This results in an energy-dependent fission cross section which is very different from the total cross section shown in Fig. 3 for ^{240}Pu.

When the incoming neutron has low energy, the likelihood of reaction is substantial only when the energy of the neutron is such as to form the compound nucleus in one or another of its resonance levels (Fig. 3*b*). The requisite sharpness of the "tuning" of the energy is specified by the total level width Γ. The nuclei ^{233}U, ^{235}U, and ^{239}Pu have a very large cross section to take up a slow neutron and undergo fission (see **table**) because both their absorption cross section and their probability for decay by fission are large. The probability for fission decay is high because the binding energy of the incident neutron is sufficient to raise the energy of the compound nucleus above the fission barrier. The very large, slow neutron fission cross sections of these isotopes make them important fissile materials in a chain reactor. *See* CHAIN REACTION (PHYSICS); REACTOR PHYSICS.

Scission. The scission configuration is defined in terms of the properties of the intermediate nucleus just prior to division into two fragments. In heavy nuclei the scission deformation is much larger than the saddle deformation at the barrier, and it is important to consider the dynamics of the descent from saddle to scission. One of the important questions in the passage from saddle to scission is the extent to which this process is adiabatic with respect to the particle degrees of freedom. As the nuclear shape changes, it is of interest to investigators to know the probability for the nucleons to remain in the lowest-energy orbitals. If the collective motion toward scission is very slow, the single-particle degrees of freedom continually readjust to each new deformation as the distortion proceeds. In this case, the adiabatic model is a good approximation, and the decrease in potential energy from saddle to scission appears in collective degrees of freedom at scission, primarily as kinetic energy associated with the relative motion of the nascent fragments.

On the other hand, if the collective motion between saddle and scission is so rapid that equilibrium is not attained, there will be a transfer of collective energy into nucleonic excitation energy. Such a nonadiabatic model, in which collective energy is transferred to single-particle degrees of freedom during the descent from saddle to scission, is usually referred to as the statistical theory of fission. *See* PERTURBATION (QUANTUM MECHANICS).

The experimental evidence indicates that the saddle to scission time is somewhat intermediate between these two extreme models. The dynamic descent of a heavy nucleus from saddle to scission depends upon the nuclear viscosity. A viscous nucleus is expected to have a smaller translational kinetic energy at scission and a more elongated scission configuration. Experimentally, the final translational kinetic energy of the fragments at infinity, which is related to the scission shape, is measured. Hence, in principle, it is possible to estimate the nuclear viscosity coefficient by comparing the calculated dependence

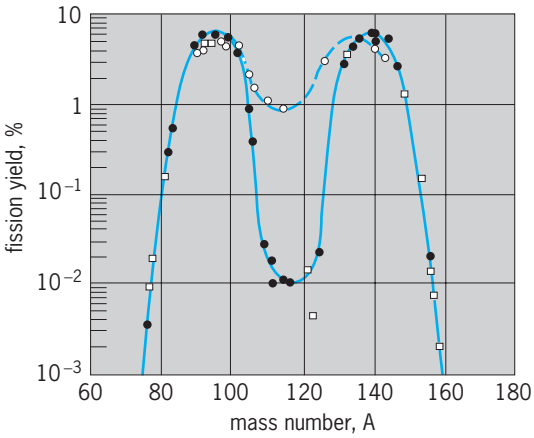

Fig. 4. Mass distribution of fission fragments formed by neutron-induced fission of $^{235}U + n = {}^{236}U$ when neutrons have thermal energy, solid curve (*Plutonium Project Report, Rev. Mod. Phys., 18:539, 1964*), and 14-MeV energy, broken curve (*based on R. W. Spence, Brookhaven National Laboratory, AEC-BNL (C-9), 1949*). Quantity plotted is 100 × (number of fission decay chains formed with given mass)/(number of fissions).

upon viscosity of fission-fragment kinetic energies with experimental values. The viscosity of nuclei is an important nuclear parameter which also plays an important role in collisions of very heavy ions.

The mass distribution from the fission of heavy nuclei is predominantly asymmetric. For example, division into two fragments of equal mass is about 600 times less probable than division into the most probable choice of fragments when ^{235}U is irradiated with thermal neutrons. When the energy of the neutrons is increased, symmetric fission (**Fig. 4**) becomes more probable. In general, heavy nuclei fission asymmetrically to give a heavy fragment of approximately constant mean mass number 139 and a corresponding variable-mass light fragment (**Fig. 5**).

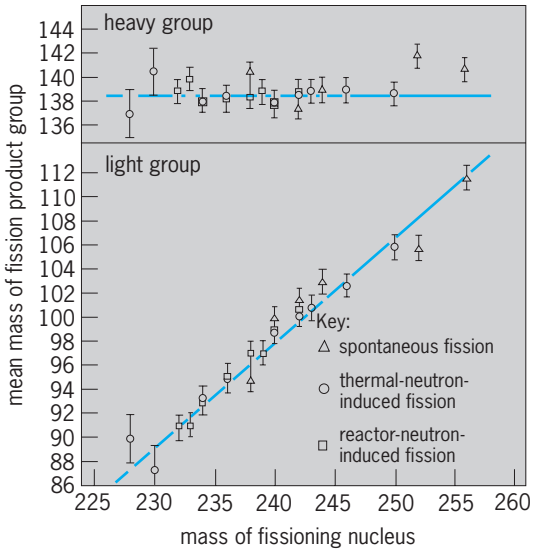

Fig. 5. Average masses of the light- and heavy-fission product groups as a function of the masses of the fissioning nucleus. Energy spectrum of reactor neutrons is that associated with fission. (*After K. F. Flynn et al., Distribution of mass in the spontaneous fission of ^{256}Fm, Phys. Rev., 5C:1725–1729, 1972*)

These experimental results have been difficult to explain theoretically. Calculations of potential-energy surfaces show that the second barrier (B in Fig. 2) is reduced in energy by up to 2 or 3 MeV, if octuple deformations (pear shapes) are included. Hence, the theoretical calculations show that mass asymmetry is favored at the outer barrier, although direct experimental evidence supporting the asymmetric shape of the second barrier is very limited. It is not known whether the mass asymmetric energy valley extends from the saddle to scission; and the effect of dynamics on mass asymmetry in the descent from saddle to scission has not been determined. Experimentally, as the mass of the fissioning nucleus approaches $A \approx 260$, the mass distribution approaches symmetry. This result is qualitatively in agreement with theory.

A nucleus at the scission configuration is highly elongated and has considerable deformation energy. The influence of nuclear shells on the scission shape introduces structure into the kinetic energy and neutron-emission yield as a function of fragment mass. The experimental kinetic energies for the neutron-induced fission of ^{233}U, ^{235}U, and ^{239}Pu have a pronounced dip as symmetry is approached, as shown in **Fig. 6**. (This dip is slightly exaggerated in the figure because the data have not been corrected for fission fragment scattering.) The variation in the neutron yield as a function of fragment mass for these same nuclei (**Fig. 7**) has a "saw-toothed" shape which is asymmetric about the mass of the symmetric fission fragment. Both these phenomena are reasonably well accounted for by the inclusion of closed-shell structure into the scission configuration.

A number of light charged particles (for example, isotopes of hydrogen, helium, and lithium) have been observed to occur, with low probability, in fission. These particles are believed to be emitted very near the time of scission. Available evidence also indicates that neutrons are emitted at or near scission with considerable frequency.

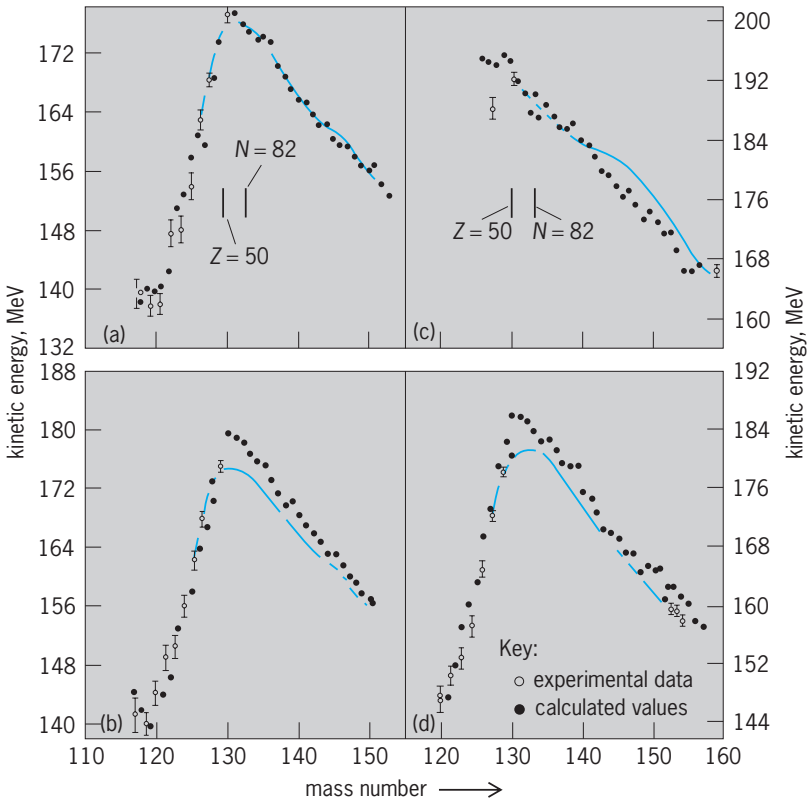

Fig. 6. Average total kinetic energy of fission fragments as a function of heavy fragment mass for fission of (*a*) ^{235}U, (*b*) ^{233}U, (*c*) ^{252}Cf, and (*d*) ^{239}Pu. Curves indicate experimental data. (*After J. C. D. Milton and J. S. Fraser, Time-of-flight fission studies on ^{233}U, ^{235}U and ^{239}Pu, Can. J. Phys., 40:1626–1663, 1962*)

Postscission phenomena. After the fragments are separated at scission, they are further accelerated as the result of the large Coulomb repulsion. The initially deformed fragments collapse to their

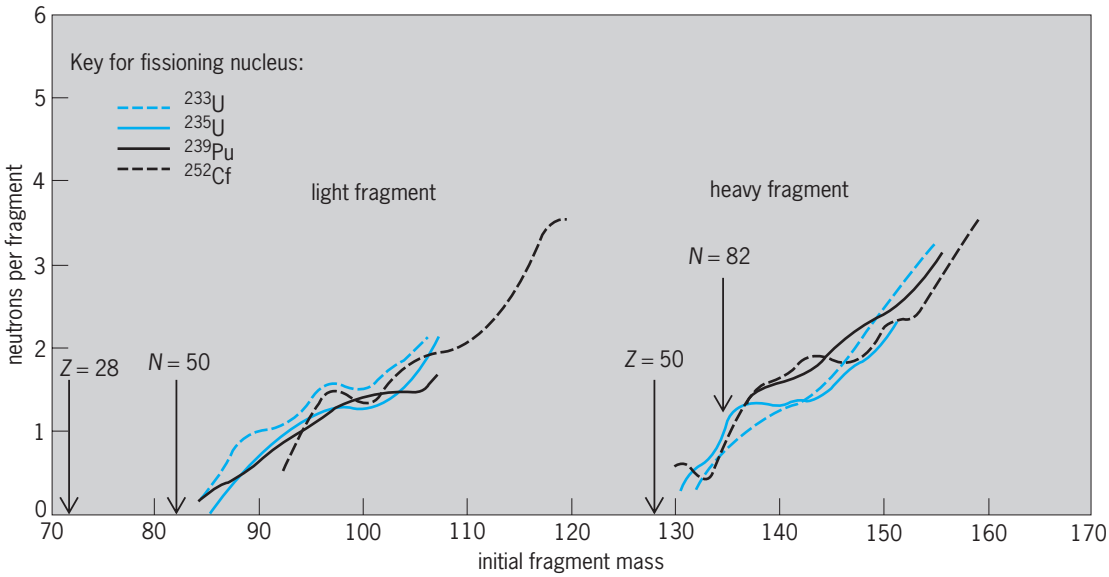

Fig. 7. Neutron yields as a function of fragment mass for four types of fission as determined from mass-yield data. Approximate initial fragment masses corresponding to various neutron and proton "magic numbers" *N* and *Z* are indicated. (*After J. Terrell, Neutron yields from individual fission fragments, Phys. Rev., 127:880–904, 1962*)

equilibrium shapes, and the excited primary fragments lose energy by evaporating neutrons. After neutron emission, the fragments lose the remainder of their energy by gamma radiation, with a life-time of about 10^{-11} s. The kinetic energy and neutron yield as a function of mass are shown in Figs. 6 and 7. The variation of neutron yield with fragment mass is directly related to the fragment excitation energy. Minimum neutron yields are observed for nuclei near closed shells because of the resistance to deformation of nuclei with closed shells. Maximum neutron yields occur for fragments that are "soft" toward nuclear deformation. Hence, at the scission configuration, the fraction of the deformation energy stored in each fragment depends on the shell structure of the individual fragments. After scission, this deformation energy is converted to excitation energy, and hence, the neutron yield is directly correlated with the fragment shell structure. This conclusion is further supported by the correlation between the neutron yield and the final kinetic energy. Closed shells result in a larger Coulomb energy at scission for fragments that have a smaller deformation energy and a smaller number of evaporated neutrons.

After the emission of the prompt neutrons and gamma rays, the resulting fission products are unstable against beta decay. For example, in the case of thermal neutron fission of ^{235}U, each fragment undergoes on the average about three beta decays before it settles down to a stable nucleus. For selected fission products (for example, ^{87}Br and ^{137}I) beta decay leaves the daughter nucleus with excitation energy exceeding its neutron binding energy. The resulting delayed neutrons amount, for thermal neutron fission of ^{235}U, to about 0.7% of all the neutrons that are given off in fission. Although small in number, these neutrons are quite important in stabilizing nuclear chain reactions against sudden minor fluctuations in reactivity. *See* DELAYED NEUTRON; NEUTRON; THERMAL NEUTRONS. John R. Huizenga

Bibliography. P. David (ed.), *Dynamics of Nuclear Fission and Related Collective Phenomena, Proceedings*, Bad Honnef, Germany, 1981; J. Kristiak and B. I. Pustylnik (eds.), *Dynamical Aspects of Nuclear Fission: Proceedings of the 2d International Conference*, 1994; H. Marten and D. Seelinger (eds.), *Physics and Chemistry of Fission*, 1991; R. Vandenbosch and J. R. Huizenga, *Nuclear Fission*, 1973.

Nuclear fuel cycle

The nuclear fuel cycle typically involves the following steps: (1) finding and mining the uranium ore; (2) refining the uranium from other elements; (3) enriching the uranium-235 content to 3–5%; (4) fabricating fuel elements; (5) interim storage and cooling of spent fuel; (6) reprocessing of spent fuel to recover uranium and plutonium (optional); (7) fabricating recycle fuel for added energy production (optional); (8) cooling of spent fuel or reprocessing waste, and its eventual transport to a repository for disposal in secure long-term storage. *See* NUCLEAR FUELS.

Steps 6 and 7 are used in Britain, France, India, Japan, and Russia. They are no longer used in the United States, which by federal policy has been restricted to a "once through" fuel cycle.

Raw materials. Nuclear reactors produce energy using fuel made of uranium slightly enriched in the isotope ^{235}U. The basic raw material is natural uranium that contains 0.71% ^{235}U (the only naturally occurring isotope that can sustain a chain reaction). The other isotope of natural uranium consists of ^{238}U, part of which converts to plutonium-239, during reactor operation. The isotope ^{239}Pu also sustains fission, typically contributing about one-third of the energy produced per fuel cycle. *See* NUCLEAR FISSION; NUCLEAR REACTOR.

Similarly, thorium-232, itself nonfissionable, when exposed in a reactor partly converts to fissionable ^{233}U, but this cycle has not been used commercially.

Uranium resources and mining. Large deposits of commerically extractable uranium ores have been found in the United States (Colorado plateau), eastern Europe, Russia, Canada, Africa, and Australia. Commercial production occurs in 23 countries. World reserves are estimated at over 4×10^6 tons (3.6×10^6 metric tons) of recoverable uranium ore. Resource estimates including low-grade ores range up to 20×10^7 tons (1.8×10^7 metric tons). *See* URANIUM.

Deposits with no surface outcrops can be found by detecting the tiny amounts of radon gas that escape from radioactive decay of uranium. Most production is now in near-surface mining. Tunnel mining requires high ventilation rates due to the presence of radon gas.

Refining. Depending on the type of ore, uranium may be purified by precipitation from aqueous solutions with ammonia or by solvent extraction from nitrate solutions. Uranium oxide is treated with hydrogen fluoride to make UF_4, and then with fluorine to make UF_6 gas. This step also completes the purification of uranium since most impurities do not have volatile fluorides.

Enrichment methods. The initial enrichment of gram quantities of ^{235}U was done in a large mass spectrograph adapted from a cyclotron magnet in Berkeley, California. The first tonnage production was done in the United States using UF_6 in a gaseous diffusion cascade, a method that was also used by Argentina, China, and France. Production plants in the United States and eight other countries currently use gas centrifuges. Other methods of separation applicable to uranium isotopes have been demonstrated on a kilogram scale, including AVLIS (atomic vapor laser isotope separation), MLIS (molecular laser isotope separation), and several processes based on ion-exchange columns. *See* ISOTOPE SEPARATION.

Fuel fabrication. Uranium in the form of UF_6 is enriched to 3–5% ^{235}U, then converted to U_3O_8 and UO_2. These oxides are compacted and furnace-fired to make high-density ceramic pellets. The pellets are put into zirconium tubes that are arranged in

precisely spaced bundles for loading into a reactor. When recycle is used, the fuel pellets are made of a solid solution of uranium and plutonium oxides called MOX (mixed oxides) fuel.

Water-cooled reactors (over 92% of the world's nuclear capacity) use fuels consisting of uranium oxide pellets in zirconium tubes. Four percent of these water-cooled reactors use blocks of graphite as a moderator. There are nearly a dozen of these Chernobyl-type reactors remaining in service, principally in Russia, Ukraine, and Lithuania. The remaining types of reactors—gas-cooled reactors and fast neutron reactors—use different fuel types; the fuel consists of bundles of small-diameter stainless steel tubes filled with pellets of highly enriched uranium (HEU) in the form of UO_2 or MOX. Fast reactor prototypes are in operation in Russia, India, and Japan, and were formerly used in the United States, France, and England.

Interim storage and cooling. Spent fuel is removed from the reactor after 3–5 years of energy output. Initially, it is stored in water-filled pools adjacent to the reactor. After about 10 years, the radioactivity and decay heat output become small enough that air-cooled storage in concrete-shielded cylinders is used, usually at or near the reactor site.

Reprocessing of spent fuel. Reprocessing of spent fuel from commercial reactors is done in large radiation-shielded facilities. Operation is entirely by remote control, using robotic devices for maintenance. Large-scale reprocessing facilities for commercial fuels have been built and operated in at least eight countries (Belgium, China, England, France, Germany, India, Japan, and Russia), and were formerly operated in the United States.

Spent fuel is chopped up and then dissolved in nitric acid. The uranium and plutonium are separated from each other and then from the radioactive fission products by using a liquid-liquid solvent extraction process called PUREX. The residues from reprocessing constitute medium- and high-level wastes. Another method, called pyroprocessing, has also been developed. *See* NUCLEAR FUELS REPROCESSING; SOLVENT EXTRACTION.

The United States has reprocessed large tonnages of uranium from production reactors to make its military stockpile of over 50 tons of plutonium. However, the United States policy since the Carter administration has been to discourage reprocessing internationally (and to set a good example domestically) in the hope that this will reduce the likelihood of proliferation of nuclear weapons. Reprocessing to recycle plutonium and uranium for power production also became uneconomic in the 1980s due to the unexpectedly high abundance and low costs of uranium ores. This development has also reduced the interest in fast breeder reactors that can produce additional fuel material. The sharp increases that have occurred since then in the costs of energy produced from fossil fuels (coal, oil, and gas), and the high costs of most alternative energy sources, now make nuclear power highly competitive economically and environmentally more attractive. (It is the only large-scale energy source that produces no emission of carbon dioxide or the other greenhouse gases that cause global warming.) Reprocessing and recycle have become correspondingly more practical economically for some countries, but they are unlikely to occur in the United States due to public concerns and continued adherence to fossil fuels for most energy production.

Recycle. After reprocessing to recover the plutonium, the latter is mixed with uranium and converted to mixed oxides. The MOX, processed to the form of high-density ceramic pellets, is loaded into zirconium tubes and fuel bundles. Recycle in thermal reactors is used to increase the energy extractable from a ton of ore by over 50%. Recycle in fast breeder reactor can increase the extractable energy by more than 500%.

Belgium, China, France, Germany, Japan, and Russia, with large and growing nuclear power capacities, use recycled plutonium. Disposal of highly enriched uranium from nuclear weapons is beginning to be undertaken by blending with natural or depleted uranium to make the 3–5% low-enrichment fuel. Similarly, MOX fuel capability can be used to dispose of plutonium stockpiled for nuclear weapons. This option is being planned in Europe and Russia, and is beginning to be considered in the United States. *See* PLUTONIUM.

Once-through fuel cycle. All United States commercial reactors use a once-through cycle, meaning without recycle. The spent fuel is stored on-site at each reactor prior to planned eventual disposal in a government-owned repository. Several long-term research efforts have been under way to evaluate the geology and hydrology of Yucca Mountain, Nevada, and other candidate sites for a repository. All sites are subject to public opposition due to the various environmental factors that must be taken into account.

Fuel performance improvements. Improvements in the fuel cycle have centered on increasing the amount of energy produced from about 4–7 megawatt–days of thermal energy (MWDt) per pound of uranium (10–15 MWDt/kg) to over 18 MWDt/lb (40 MWDt/kg). Development work is reaching for 27–45 MWDt/lb (60–100 MWDt/kg). This corresponds to an increase from about 40 megawatt-hours to over 300 megawatt-hours of electricity produced per pound of fuel (or from 88 MWh/kg to 660 MWh/kg). This has made it possible to produce power for 18–24 months between extended shutdowns for refueling. The longer periods between reloadings increase capacity factors and reduce production costs per unit energy generated.

Nuclear waste disposal options. The principal options for the disposal of radioactive wastes from civilian nuclear energy are the following:

1. Storing the fission product waste as a vitrified glass or ceramic after reprocessing, and removing the uranium and plutonium for recycle. The vitrified cylinder is encapsulated in a corrosion-resistant metal jacket (often double-walled) and emplaced in a

deep underground repository. The repository is chosen for its geology that demonstrates it has been stable for tens of millions of years and appears likely to remain stable for at least 10,000 years more.

2. Storing vitrified waste as in option 1 above but with added processing to sequester certain long-lived fission products and actinide elements for separate disposal or for transmutation. Research studies are exploring the feasibility of separating and transmuting several of the long-lived radioelements to more innocuous forms by exposure in a reactor or a large accelerator. The resulting products are either stable or have short half-lives, simplifying the disposal requirements for the remaining materials.

3. Vitrifying and emplacing the spent fuel as above but without recycle or removing the uranium, plutonium, fission products, or actinides.

4. Mechanically compacting bundles of spent fuel rods, encapsulating them in specially designed jackets, and emplacing them as above.

Many countries are implementing option 1 and are researching option 2. The United States is following option 4 for commercial spent fuel. The United States is using option 1 above for the waste remaining from the former production reactors and Navy ship propulsion reactors. There has been a small-scale use of option 4 by Sweden. *See* RADIOACTIVE WASTE MANAGEMENT.

Costs. The sum of the costs of the eight steps in the complete nuclear fuel cycle can be expressed as X dollars (U.S.) per pound of fuel (or in comparable units of other nations). Where recycle is used, there is a credit for the fuel value of the recovered uranium and plutonium. A pound of fuel produces heat energy, which is converted to electricity, producing Y kilowatt-hours per pound. The nuclear fuel cycle cost is then X/Y in dollars per kilowatt-hour. For most countries, including the United States, the sum of these costs is lower than the fuel costs for burning oil or gas, and about half that of burning coal. However, the plant costs and labor costs for nuclear generating units are generally higher than for fossil fuel generating units. *See* COAL.

Issues. Various issues revolve around the type of nuclear fuel cycle chosen. For instance, the question is still being argued whether "burning" weapons materials in recycle reactors is more or less subject to diversion (that is, falling into unauthorized hands) than storing and burying these materials. In either case, the degree of security is highly dependent on the effectiveness of national and international regimes of inspection, accountancy, and control. The International Atomic Energy Agency in Vienna is one such agency for control of recycled material from civilian reactors.

Another issue involves the composition of radioactive wastes and its impact on repository design. The nuclear fuel cycles that include reprocessing make it possible to separate out the most troublesome long-lived radioactive fission products and the minor actinide elements that continue to produce heat for centuries. The remaining waste decays to radiation levels comparable to natural ore bodies in about 1000 years. The separated actinides are a small volume fraction of the total waste and can be sequestered for transmutation to more stable products by irradiation in reactors or by using particle accelerators. The shorter time for the resulting wastes to decay away simplifies the design, management, and costs of the repository. *See* ACTINIDE ELEMENTS.

Edwin L. Zebroski

Bibliography. L. C. Hebel et al., Report of the Study Group on Nuclear Fuel Cycles and Waste Management, *Rev. Mod. Phys.*, 50(1)Part II:S114–S117, 1978; R. Lee et al., *External Costs and Benefits of Fuel Cycles*, ORNL and Resources for the Future, McGraw-Hill Utility Data Institute, Washington, DC, 1995; C. McCombie, The hazards presented by radioactive wastes and the safety levels to be provided by repositories, *Nuc. Eng. Design*, 176(1–2):43–50, 1997; T. H. Pigford and P. L. Chambré, *Scientific Basis for Nuclear Waste Management: Proceedings of the Material Research Society*, Elsevier, 1983.

Nuclear fuels

Materials whose ability to release energy derives from specific properties of the atom's nucleus. In general, energy can be released by combining two light nuclei to form a heavier one, a process called nuclear fusion; by splitting a heavy nucleus into two fragments of intermediate mass, a process called nuclear fission; or by spontaneous nuclear decay processes, which are generically referred to as radioactivity. Although the fusion process may significantly contribute to the world's energy production in future centuries and although the production of limited amounts of energy by radioactive decay is a well-established technology for specific applications, the only significant industrial use of nuclear fuel so far utilizes fission. Therefore, the term nuclear fuels generally designates nuclear fission fuels only. *See* NUCLEAR BATTERY; NUCLEAR FISSION; NUCLEAR FUSION; RADIOACTIVITY AND RADIATION APPLICATIONS.

Nuclear characteristics. Large releases of energy through a fission or a fusion reaction are possible because the stability of the nucleus is a function of its size. The binding energy per nucleon provides a measure of the nucleus stability. By selectively combining light nuclei together by a fusion reaction or by fragmenting heavy nuclei by a fission reaction, nuclei with higher binding energies per nucleon can be formed. The result of these two processes is a release of energy, the magnitude of which is several orders greater than that obtained by the combustion of carbon compounds (such as wood, coal, or oil). For example, the fissioning of one nucleus of uranium releases as much energy as the oxidation of approximately 5×10^7 atoms of carbon. *See* NUCLEAR BINDING ENERGY.

Many heavy elements can be made to fission by bombardment with high-energy particles. However, only neutrons can provide a self-sustaining nuclear fission reaction. Upon capture of a neutron by a heavy nucleus, the latter may become unstable and

split into two fragments of intermediate mass. This fragmentation is generally accompanied by the emission of one or several neutrons, which can then induce new fissions. Only a few long-lived nuclides have been found to have a high probability of fission: ^{233}U, ^{235}U, and ^{239}Pu. Of these nuclides, only ^{235}U occurs in nature as 1 part in 140 of natural uranium, the remainder being mostly ^{238}U. The other nuclides must be produced artificially: ^{233}U from ^{232}Th, and ^{239}Pu from ^{238}U. The nuclides ^{233}U, ^{235}U, and ^{239}Pu are called fissile materials since they undergo fission with either slow or fast neutrons, while ^{232}Th and ^{238}U are called fertile materials. The latter, however, can also undergo the fission process at low yields with energetic neutrons; therefore, they are also referred to as being fissionable.

The term nuclear fuel applies not only to the fissile materials, but often to the mixtures of fissile and fertile materials as well. Using a mixture of fissile and fertile materials in a reactor allows capture of excess neutrons by the fertile nuclides to form fissile nuclides. Depending on the efficiency of production of fissile elements, the process is called conversion or breeding. Breeding is an extreme case of conversion corresponding to a production of fissile material at least equal to its consumption.

Physical characteristics. The type and design of nuclear fuels vary with the reactor design. For a discussion of the steps by which the nuclear fuel is prepared for use in a reactor and eventually disposed of, *see* NUCLEAR FUEL CYCLE

Most modern reactors that are used to produce electricity are light-water–moderated and –cooled (light-water reactors or LWRs); the fuel is uranium isotope 235 (^{235}U) diluted in ^{238}U, and it is in the form of uranium dioxide (UO_2). Several prototype power reactors use liquid sodium as a coolant (for example, liquid-metal fast breeder reactors or LMFBRs), and the fuel is usually ^{235}U and ^{239}Pu in the oxide form. Gas-cooled reactors usually use oxides or carbides of ^{232}Th and ^{235}U or ^{233}U in the shape of small spheres.

The designs of nuclear fuels have covered a wide range of chemical forms, sizes, shapes, containment types, and detailed characteristics. The reactor design, neutron-physics properties, economics, safety and licensing requirements, and proven reliability affect fuel design.

The chemical form is determined by the effects of the nonfissile atoms on the neutron flux and on the physical characteristics, which in turn influence the reliability and licensability of the fuel design. The only influence that the chemical form has on the energy production is secondary, that is, by the nonfissile atoms absorbing or moderating neutrons. This insensitivity of the nuclear fission process to chemical form allows flexibility in fuel types. The primary forms used have been metals, oxides, carbides, and nitrides. Almost all commercial reactors use oxides (UO_2). Fissile material also has been alloyed with or dispersed in solid materials. Other forms that have been tried include fissile material in a molten salt, water, acid, or liquid metal. In general, these solutions are too corrosive to be practical.

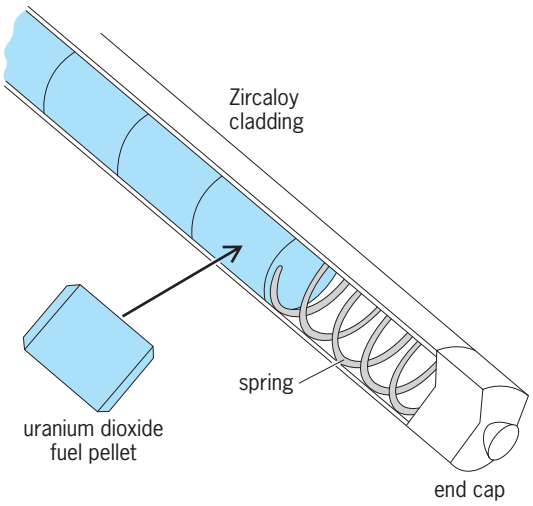

Fig. 1. Fuel rod design for a light-water reactor.

The solid fuels have been designed with many shapes, including rods, spheres, plates, and curved plates. The fissile material is contained in a structural material that is intended to contain fission products and provide structural stability. The most common fuel design uses small cylindrical pellets of uranium dioxide (UO_2) sealed inside a long tube (**Figs. 1** and **2**). These designs and their properties are discussed in more detail below.

The design of nuclear fuels must provide physical properties that ensure safe operation, reliable

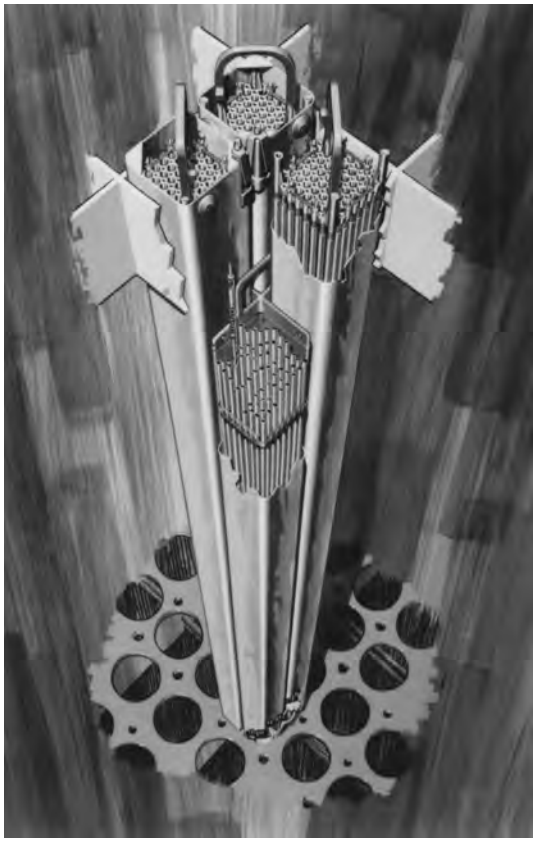

Fig. 2. Boiling-water reactor fuel assembly. (*General Electric Co.*)

performance, and economical power production. For light-water reactors, safe operation requires that the sealed tube that contains the fissile material remain intact during normal operation and most abnormal events, which are usually associated with a power increase or a loss of cooling. Reliable performance requires that the fuel remain dimensionally stable and that the fuel container not leak during normal operation. Economical power production, that is, low fuel-cycle cost, requires that neutrons be used efficiently to produce fission and to convert fertile nuclides to fissile nuclides; that a high fraction of fissile isotopes be burned; and that waste disposal costs be minimized.

Reliability. During normal operation, fuel performance is characterized by the integrity of the fuel container and the stability of component dimensions. Fuel containers that leak will release radioactive fission products to the coolant. Strict limits on coolant radioactivity, which protect plant personnel and the public, prevent operation with more than approximately one leaking rod in 2000 for light-water reactors (plants normally operate with less than one leaking rod in 10,000). Excessive leakage results in costly reactor shutdowns for locating and replacing defective fuel elements.

Dimensional stability is important to ensure proper coolant flow and power distribution throughout the core. In light-water reactors the primary concerns, which have been eliminated with design improvements, have been with fuel-pellet swelling and resultant interaction with the cladding, fuel-rod elongation and resultant interaction with the top and bottom plates, and fuel-assembly distortion. As fuel performance has improved, fuel burnup has been increased to reduce costs. This increase in residence results in corrosion of the cladding tube being the limiting performance attribute. In the 1990s many nuclear fuel vendors developed zirconium alloys with improved corrosion resistance to support increased fuel burnup.

Economic power production. Fuel-cycle costs, which are a significant but not the greatest contributor to electrical-energy production costs, are mostly determined by raw material (yellowcake) purchase costs, enrichment costs, interest rates, energy extracted, and waste storage and disposal costs. Fuel design and enrichment requirements are influenced by the desired energy extraction and neutron economics (efficient use of the fissile atoms). In light-water reactors, fuel designers try to maximize the amount of water and minimize the amount of structural material (principally a zirconium alloy called Zircaloy) in the core. Increased water provides better moderation of neutrons, and reduction of structural material decreases the number of neutrons absorbed by parasitic atoms (neither fissile nor fertile). Most of the cost of electricity production by nuclear power is due to capital and operating expenditures, not fuel-cycle costs.

Properties. The important properties of nuclear fuel are thermal, mechanical, chemical, and nuclear. Thermal properties influence component stored energy and temperatures, which affect phenomena that are temperature sensitive. Mechanical properties, such as the creep rate and yield strength of structural components, determine dimensional stability and influence reliability. Chemical properties determine reaction rates among different types of atoms and influence reliability, for example, by assisting in stress-corrosion cracking. Irradiation changes fuel properties by changing the chemical composition and the atomic structure, and these changes must be reflected in performance and licensing calculations.

The basic thermal properties of the fuels, such as thermal conductivity, are only modestly changed by the irradiation and the resultant damage. The conversion of fertile atoms to fissile atoms and of fissile atoms to the fission-product atoms changes the composition of the materials, but the changes in thermal properties are modest. *See* THERMAL CONDUCTION IN SOLIDS.

The situation is quite different for gases. For light-water reactor fuel designs, a gas, such as helium, is used to transfer heat from the fuel to the structural container. Fission gases, such as xenon and krypton, are poor conductors of heat and degrade the thermal conductivity of the heat-transfer gas if they escape from the fuel. Heat transfer also is affected by cracking and slight movement of fuel, by microstructural changes in the fuel, and by bonding of the fuel and container, which can occur by chemical diffusion. *See* CONDUCTION (HEAT).

The mechanical properties of the core structural components and the fuel change with irradiation. High-energy particles, for example, neutrons, interact with atoms in solids and knock significant numbers of atoms out of their normal positions. This is called radiation damage. The effects of this damage are different during and after the radiation and different for time-dependent than for time-independent properties. *See* RADIATION DAMAGE TO MATERIALS.

The extra free atoms that are knocked out of their normal position often make time-dependent processes, such as creep, occur at faster rates. Time-independent phenomena do not have time during irradiation to notice the softening effect of the knocked-out atoms; they just see the hardening effect of the damage. However, after irradiation, the damage left behind makes it more difficult for deformation to occur, and the material properties reflect those of a harder material for both time-dependent and time-independent properties.

The fission process also causes changes in the chemical composition of materials. For example, fuel that started out as essentially all unranium dioxide slowly accumulates plutonium and many types of fission atoms. Some of these such as iodine are corrosive, some are radioactive with long half-lives, and some are gases and are not stable in solid fuel. Fuels are usually designed to retain the products in the fuel so that they do not corrode or stress the fuel container. Structural materials are selected to have few and benign interactions during radiation, thus

producing few undesirable isotopes. However, some radioactive atoms are produced, for example, ^{60}Co, in stellite and steel components. In general, the change in composition is not sufficient to significantly change the material properties. Most changes in properties are due to irradiation hardening.

The fission products have some influence on neutron economics. Some of the products capture a significant number of neutrons. As their concentration increases, they will eventually steal a significant amount of reactivity. Reprocessing of spent fuels has to remove isotopes with high neutron cross section as well as recover fissile material, thus obtaining high-reactivity fuel from spent fuel. *See* NUCLEAR FUELS REPROCESSING.

Breeder reactor fuels. The capture of neutrons by abundant, fertile nuclides, such as ^{238}U, leads to conversion of those atoms to fissile nuclides, such as ^{239}Pu. If nuclear power eventually becomes the major source of electric power for many industrialized countries, there may be a need to build reactors that breed fuel by converting more fertile material to fissile material than is consumed during operation, thus providing fuel for themselves and for nonbreeding light water reactors.

High conversion ratios are obtained by designing the reactor to have a flux spectrum that has a high density of high-energy (fast) neutrons (that is, a hard flux spectrum). This increases the number of neutrons released by the fission process; a greater amount of excess neutrons is then available for breeding purposes. The predominant current design uses a liquid metal, usually sodium, instead of water to cool the fuel. This removes the primary material (hydrogen in water) that slows down (thermalizes) the neutrons. Since fast neutrons are not as efficient as thermal neutrons at inducing fission, the enrichment (concentration) of atoms (usually ^{235}U and ^{239}U) in the fuel is increased to maintain the fissile chain reaction.

The main differences in the design of the fuel elements from those of light-water reactors are that the cladding is usually stainless steel rather than an alloy of zirconium, and fuel elements are grouped together into assemblies that are placed inside hexagonal stainless steel cans. The fuel pellets are uranium dioxide (approximately 80%) and plutonium dioxide (approximately 20%), although the nitrides and carbides of uranium have been tested and the Experimental Breeder Reactor II in Idaho operated with highly enriched metal fuel. The primary advantage of these alternative designs is their higher fuel thermal conductivity compared to that of uranium and plutonium dioxides. However, other characteristics of these fuels, particularly changes in crystallographic structure at low temperature, low melting temperature, fission-gas retention, compatibility with cladding, or swelling under irradiation, are either inferior to those of the oxide or are not reliably established.

The stainless steel cladding, usually alloy 316 or 304, is stronger than zirconium alloys at operating temperatures, although stainless steels have a lower melting temperature which is a disadvantage for transients. *See* ALLOY; STAINLESS STEEL.

Gas-cooled reactor fuels. The first gas-cooled reactor (GCR) power station was built at Calder Hall in the United Kingdom and began operation in 1956. Since then, a number of gas-cooled reactor plants have been constructed and operated. The gas-cooled reactors have graphite cores comprising multisided prisms stacked into a horizontal cylinder, about 50 ft (15 m) in diameter and 33 ft (10 m) long. The coolant is carbon dioxide (CO_2). The fuel is natural uranium that is metal clad with a thin cylindrical tube of a magnesium alloy called Magnox. In order to achieve good heat transfer, the Magnox cans are provided with cooling fins.

The Magnox-type gas-cooled reactors are still in operation, but have been superseded by the advanced gas-cooled reactor (AGR). These reactors are similar to the Magnox type in general design. However, the Magnox fuel elements have been replaced by slightly enriched uranium dioxide clad in stainless steel, permitting higher surface temperatures and greater thermal-cycle efficiency.

To increase efficiency further, coolant temperatures in the 1560–1830°F (850–1000°C) range are implemented in the high-temperature gas-cooled reactor (HTGR). Helium is used in place of carbon dioxide as the coolant due to the instability of carbon dioxide at high temperatures. The development of the high-temperature gas-cooled reactor has proceeded in two directions: the pebble bed concept and the prismatic core design. In the German pebble bed design (thorium high-temperature reactor or THTR), the fuel elements consist of 2.4-in.-diameter (60-mm) spheres made of a carbon outer zone and an inner 2-in.-diameter (50-mm) region containing a mixture of fertile and fissile particles uniformly dispersed in a graphite matrix. Each ball contains 0.04 oz (1.1 g) of $^{235}UO_2$ and 0.4 oz (11 g) of $^{232}ThO_2$. The core consists of about 700,000 such balls in a random arrangement. In the prismatic design, developed in the United States, the core consists of stacks of hexagonal prisms of graphite. The prismatic fuel elements are pierced with channels for cooling and for fuel rods that are made by pressing graphite powder and fuel particles together (**Fig. 3**). Each rod contains a mixture of fissile and fertile particles. The particles consist of microspheres made of fully enriched uranium carbide (UC_2), a mixture of uranium carbide and thorium carbide (ThC), and thorium carbide or thorium oxide (ThO_2), coated with several concentric layers of pyrolytic carbon. The two coated-particle types in most common use are the two-layer Biso (for buffer isotropic) coating and the four-layer Triso coating with its interlayer of silicon carbide (SiC) between two layers of high-density isotropic pyrolytic carbon. Both Biso and Triso particles are capable of essentially complete retention of gaseous fission products.

Improved coatings are needed to increase the strength and integrity of the fuel particle, reduce the diffusional release of fission products, and improve

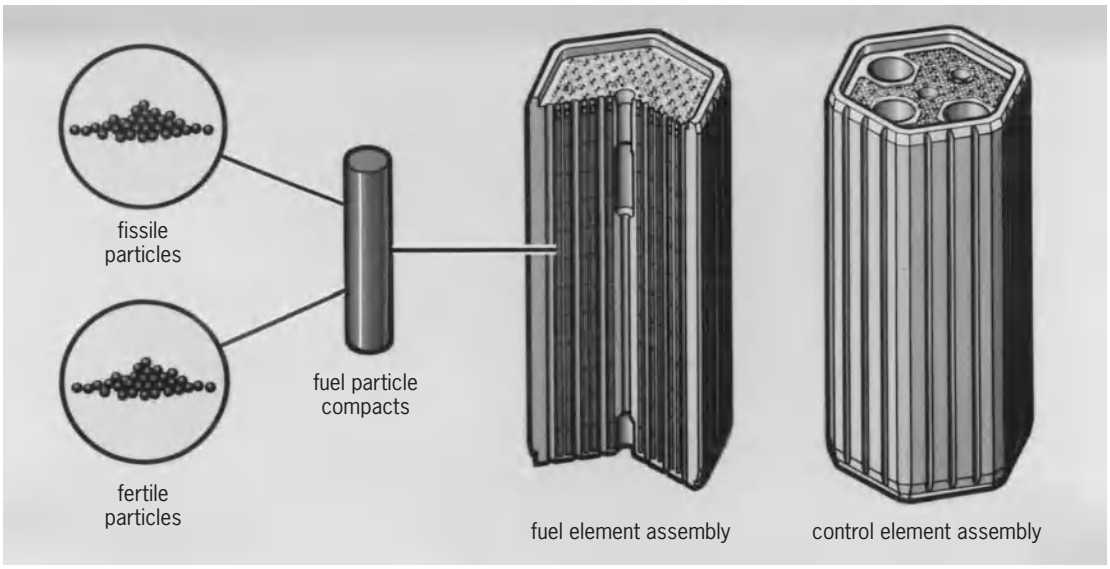

Fig. 3. Fuel element components in a high-temperature gas-cooled reactor with the prismatic core design used in the United States. (*T. D. Gulden, G. A. Technologies, Inc.*)

the overall chemical stability and temperature capability of coated particles.

Fuel designs for less common reactors. A large number of different types of reactors have been designed, constructed, and operated. However, their total capacity, including high-temperature gas-cooled reactors and liquid-metal fast-breeder reactors, has been less than 1% of the capacity of light-water reactors in the United States. Some of them were used for purposes other than the large-scale production of electricity or heat. Others are considered obsolete or have little chance for large-scale power production. Practically all concepts rely on a heterogeneous aggregation of fuel and coolant similar to those described above. A notable exception is provided by the homogeneous reactor experiments (HRE). Instead of using a solid matrix, the core consisted of a ^{233}U sulfate solution in deuterium oxide (D_2O). The molten salt reactor (MSR) was a similar project in which uranium tetrafluoride (UF_4) and thorium tetrafluoride (ThF_4) were kept in a eutectic melt of beryllium fluoride (BeF_2), zirconium tetrafluoride (ZrF_4), and sodium fluoride (NaF) or lithium-7 fluoride (^7LiF).

Safety and licensing. Nuclear fuel rods generate large amounts of energy and produce highly radioactive nuclides inside the fuel. The cladding material that contains the fuel is the first containment barrier of these radionuclides, and its integrity must be ensured. The design of nuclear fuel and reactor systems, and the operational procedures must provide this assurance. The U.S. Nuclear Regulatory Commission (NRC) represents the public to ensure that its safety is protected by these designs and procedures.

The objective of safety analyses of nuclear fuel designs is to ensure that the cladding material will retain its integrity during normal operation and during unusual events such as temperature or power increases above normal levels. Since essentially all of the power reactors in the United States are light-water reactors with Zircaloy-clad or ZIRLO^R-clad uranium dioxide fuel, the following discussion is based on safety consideration of these designs.

High temperatures can be caused by either a prolonged loss of the reactor cooling water (loss-of-cooling accident or LOCA), an increase in power, or a reduction in flow of cooling water. High temperatures (above 1200°F or 650°C) can result in rapid cladding oxidation and hydriding (from the reaction with the cooling water), chemical interactions between the fuel and the cladding, and, at very high temperatures, melting of the fuel or cladding. Each one of these reactions would compromise the cladding integrity immediately or during further service, and must be avoided.

High power results from uncontrolled increases in the reactivity of the reactor core, for example, from ejection of a neutron-absorbing control material from the core or from the sudden injection of cold water into the core. Safety analyses are performed to ensure that, during these events, the surface of the fuel rods will continue to be cooled by a layer of cooling water and that the uranium dioxide fuel temperatures will not become excessive. Nuclear reactors are designed to have large margins of safety against all mechanically plausible means of reactivity increase, and maintenance of these margins of safety is enforced by regulations.

A loss of cooling water can be caused by pipe ruptures or stuck valves. For extreme loss of water by a pipe rupture or loss-of-coolant accident (LOCA), all light-water reactors in the United States are designed to maintain the cladding temperature below 2200°F (1200°C) and the cladding total oxidation at less than 17% of the metal.

The only major loss of cooling in a United States civilian light-water reactor occurred at the Three

Mile Island (Pennsylvania) Unit 2 in 1979. The loss of coolant occurred through a stuck open valve. The reactor operator, failing to recognize this situation, shut down the automatically activated emergency core cooling system (ECCS) and maintained let-down flow. For several hours the upper portion of the core was uncovered before the real situation eventually was recognized and the emergency core cooling system was activated again, resulting in recirculation of cooling water and termination of the undercooling event. During the uncovering of the upper portion of the core, the fuel rods exceeded the above limits for peak temperature and oxidation. The fuel distorted; significant portions of the structural materials melted and solidified into different shapes and formed a rubble bed on top of the lower portion of the core, and some of the material was deposited in the lower part of the reactor vessel. Despite the severity of the accident and the breakdown of the cladding as a barrier, other reactor containment systems functioned properly. The increase in radiation exposure to persons in the vicinity of the reactor at the time of the accident was within the range of local variation of annual exposure from nature. *See* NUCLEAR POWER; NUCLEAR REACTOR; REACTOR PHYSICS.

David Franklin; Albert Machiels

Bibliography. E. R. Bradley and G. D. Sabol (eds.), *Zirconium in the Nuclear Industry*, Eleventh International Symposium, ASTM STP 1295, 1996; C. Cangruly, *Nuclear Fuel Fabrication*, 2d ed., 1989; B. A. Ma, *Nuclear Reactor Materials and Application*, 1983; J. T. A. Roberts, *Structural Materials in Nuclear Power Systems*, 1981; F. H. Wittmann (ed.), *Fuel Elements and Assemblies*, 1987.

Nuclear fuels reprocessing

Nuclear fuels are reprocessed for military or civilian purposes. In military applications, reprocessing is applied to extract fissile plutonium from fuels that are designed and operated to optimize production of this element. In civilian applications, reprocessing is used to recover valuable uranium and transuranic elements that remain in fuels discharged from electricity-generating nuclear power plants, for subsequent recycle in freshly constituted nuclear fuel. This military-civilian duality has made the development and application of reprocessing technology a sensitive issue worldwide and necessitates stringent international controls on reprocessing operations. It has also stimulated development of alternative processes that do not produce a separated stream of pure plutonium so that the proliferation of nuclear weapons is held in check. *See* NUCLEAR POWER; PLUTONIUM; TRANSURANIUM ELEMENTS; URANIUM.

Nuclear fuel is removed from civilian power reactors due to chemical, physical, and nuclear changes that make it increasingly less efficient for heat generation as its cumulative residence time in the reactor core increases. The fissionable material in the fuel is not totally consumed; however, the buildup of fission product isotopes (with strong neutron-absorbing properties) tends to decrease the nuclear reactivity of the fuel. *See* NUCLEAR FISSION; NUCLEAR FUELS; NUCLEAR REACTOR.

Factors. A typical composition of civilian reactor spent fuel at discharge is 96% uranium, 3% fission products, and 1% transuranic elements (generally as oxides, because most commercial nuclear fuel is in the form of uranium oxide). The annual spent fuel output from a 1.2-gigawatt electric power station totals approximately 33 tons (30 metric tons) of heavy-metal content. This spent fuel can be discarded as waste or reprocessed to recover the uranium and plutonium that it contains (for recycle in fresh fuel elements). The governments of France, the United Kingdom, Russia, and China actively support reprocessing as a means for the management of highly radioactive spent fuel and as a source of fissile material for future nuclear fuel supply. The United States has maintained a national policy forbidding the reprocessing of civilian reactor fuel for plutonium recovery since the mid-1970s. Therefore, the United States is the only one of the five declared nuclear weapons states with complete fuel recycling capabilities that actively opposes commercial fuel reprocessing.

Decisions to reprocess are not made on economic grounds only, making it difficult to evaluate the economic viability of reprocessing in various scenarios. In the ideal case, a number of factors must be considered, including: (1) Cost of uranium/U_3O_8; prices are directly related to supply and demand. (2) Cost of enrichment; advanced technologies may reduce costs significantly. (3) Cost of fuel fabrication; plutonium-bearing fuels are more expensive to fabricate due to the toxicity of plutonium and the need for remote manufacturing techniques. (4) Cost of reprocessing; costs vary widely from country to country. (5) Waste disposal cost; high-level waste volumes from modern reprocessing should be significantly less than for the case of direct disposal, and the extraction of actinide elements may simplify the licensing of disposal systems. (6) Fissile content of spent fuel; high-value spent fuels (for example, those that contain large fractions of fissile material) make reprocessing more profitable.

The once-through fuel cycle (that is, direct disposal/no reprocessing) is favored when fuel costs and waste disposal costs are low and reprocessing costs are high. It is not necessarily true that reprocessing will inevitably be more costly than direct disposal, because technology advancements and escalating waste disposal costs can swing the balance in favor of reprocessing. *See* ACTINIDE ELEMENTS; NUCLEAR FUEL CYCLE; RADIOACTIVE WASTE MANAGEMENT.

It is generally accepted that breeder reactors, those that are designed to generate more fuel than is consumed (by absorption of excess neutrons in peripheral blankets composed of fertile isotopes such as ^{238}U or ^{232}Th), will not be deployed until the latter part of the twenty-first century. With these reactors, reprocessing will be mandatory to recover the fissile material generated in the blankets. Because

breeder reactors tend to operate with fuel of considerably higher fissile enrichment level than in the case of light-water reactors, the avoided cost of external fuel enrichment also favors reprocessing. In this fuel cycle it is assumed that all fissile material requirements of the fresh fuel are obtained from reprocessing the spent fuel and recycling its fissile content. It is reasonable to expect that the economics of reprocessing will be advantageous to the deployment of breeder reactors when the time comes for their introduction into the energy generation mix. *See* RADIOISOTOPE.

Methods. The technology of reprocessing nuclear fuel was created as a result of the Manhattan Project during World War II, with the purpose of plutonium production. Early reprocessing methods were refined over the years, leading to a solvent extraction process known as PUREX (plutonium uranium extraction) that was implemented on a large scale in the United States in late 1954 at the Savannah River site in South Carolina and in early 1956 at the Hanford site in Washington. The PUREX process is an aqueous method that has been implemented by several countries and remains in operation on a commercial basis. The PUREX plant at Hanford was shut in 1992, while the Savannah River plant is still operating for the treatment of residual research reactor fuel inventories. A nonaqueous reprocessing method known as pyroprocessing was developed in the 1990s as an alternative to PUREX. It has not been deployed commercially, but promises greatly decreased costs and reduced waste volumes, with practically no secondary wastes or low-level wastes being generated. It also has the important attribute of an inability to separate pure plutonium from irradiated nuclear fuel. *See* SOLVENT EXTRACTION.

PUREX reprocessing. The PUREX process is carried out in large shielded facilities with remote operation capabilities. The PUREX plant at the Hanford site operated with a throughput capacity of 3000 tons (2700 metric tons) of heavy metal per year. Modern plants with somewhat smaller capacity are operating in Sellafield in England and LaHague in France. In the PUREX process, spent fuel is initially dissolved with nitric acid in a large vessel. In the case of aluminum-clad fuel, the dissolution is catalyzed by the addition of mercury, whereas zirconium-alloy fuel is dissolved in a mixture of nitric and hydrofluoric acids; the hydrofluoric acid is added to complex the zirconium and prevent the occurrence of highly energetic reactions. Zirconium alloy-clad fuel can be declad chemically with ammonium fluoride (ZRFLEX process); alternatively, the spent fuel can be chopped into short segments and the fuel material leached out with nitric acid, leaving the cladding undissolved (chop-leach process). Thorium dissolution requires the use of a catalyst such as fluoride ion. Except for a few fission products that are volatilized during dissolution, all the constituents initially in the fuel are retained in the dissolver solution as soluble nitrate salts or, in the case of minor constituents, as an insoluble sludge. Generally, the solution must be clarified to accommodate its use as feed to the solvent extraction process. *See* THORIUM.

In the solvent extraction steps, purification of the desired products (generally uranium and plutonium, possibly thorium) is achieved by the selective transfer of these constituents into an organic solution that is immiscible with the original aqueous solution. The most commonly used organic solution consists of the extractant tri-*n*-butyl phosphate (TBP) dissolved in purified kerosine or similar materials such as long-chain hydrocarbons (for example, *n*-dodecane). Uranium in aqueous nitrate solutions normally exists only in the +4 (IV) and +6 (VI) oxidation states. Nitric acid dissolution of spent fuel yields U(VI) and nitrogen oxides (NO, NO_2) as products. The U(VI) is strongly hydrolyzed, and exists as the uranyl (UO_2^{+2}) ion. By the same token, plutonium may exist in the Pu(III) or Pu(IV) states, with the (IV) state resulting from nitric acid dissolution. When in the proper oxidation state [that is, (UO_2^{+2}), Pu(IV), and Th(IV)], the desired products are readily extractable into tributyl phosphate. Plutonium in the Pu(III) state has a very low extractability into TBP.

The solvent extraction separation is based on this ready extractability of uranyl nitrate and plutonium(IV) nitrate and the relative inextractability of fission products and plutonium(III) nitrate. The first step of the PUREX solvent extraction process thus involves the coextraction of uranium and plutonium into the organic (TBP + kerosine) phase, leaving most of the fission products in the aqueous phase. The organic solvent is then contacted with a clean aqueous solution, and the uranium and plutonium transfer into this clean aqueous phase. The solvent can now be recycled. After reducing the plutonium to the plutonium(III) oxidation state, it is then a matter of contacting the aqueous phase containing plutonium(III) nitrate and uranyl nitrate with clean organic phase; the uranium transfers to the organic phase and the plutonium remains in the aqueous phase. The uranium-organic phase can then be contacted with another clean aqueous phase to transfer the uranyl nitrate into the aqueous phase.

The products of the PUREX process described in rudimentary form here are separated nitrate solutions of uranium and plutonium. These solutions are treated separately to produce uranium and plutonium oxides. The solvent extraction process yields virtually complete recovery of the uranium and plutonium present in the spent fuel, and the uranium and plutonium are of high purity. Decontamination factors (defined as the ratio of the concentration of an impurity element to that of the valued constituent in the feed, divided by the same ratio for these elements in the product) of 10^7 to 10^8 are possible.

Because TBP is susceptible to decomposition to dibutyl phosphate and monobutyl phosphate, both of which have significant deleterious effects on plutonium stripping, care must be taken to maintain the purity of the solvent. Otherwise, additional process steps are necessary to recover plutonium, at the cost of increased volume of liquid waste. Limiting the time of exposure of sensitive solvents to high levels of nitric acid and beta/gamma radiation greatly increases their useful lifetime.

Pyroprocessing. Rising public concerns over costs, reactor safety, disposal of nuclear wastes, and the threat of nuclear weapons proliferation have impacted government and electric utility decisions on power plant deployment. Of particular concern is the situation that would prevail should nuclear electric power generation accelerate greatly in the lesser-developed nations. It is generally agreed that significant growth in nuclear capacity would mandate reprocessing of spent nuclear fuel for more efficient utilization of fuel resources and minimization of high-level waste quantities. However, the reprocessing technologies now in place are mid-twentieth-century technology and may prove to be impediments to such growth.

In response to these concerns, an advanced processing method is being developed to address both system economics and public acceptance issues. This method, initially developed to support the recycle of fuel materials in an advanced liquid-metal reactor system, is based on the separation of actinide elements from fission products in a molten salt electrorefining process. In this process, spent fuel in the metallic state (either as metal fuel or as the product of a head-end reduction step) is anodically dissolved in an electrorefining cell operating at a temperature of 932°F (500°C). The actinide elements present in the spent fuel are electrotransported to suitable cathodes, and the cathode deposits are treated to remove entrained electrolyte salt. A simple steel cathode is used to collect a pure uranium deposit; the transuranic elements, having chloride salts that are more stable than uranium chloride, must be collected in a liquid cadmium cathode where they deposit as solid intermetallic compounds (for example, $PuCd_6$). The transuranic elements cannot be separated in this process, and the liquid cathode deposit typically contains about 70% (by weight) transuranics and 30% uranium, together with a minor amount (approximately 0.25%) of the rare-earth fission products. The transition-metal and noble-metal fission products remain in the anode baskets together with the cladding hulls, while the halide, pnictide, alkali metal, and alkaline-earth fission products remain in the electrolyte salt. *See* ALKALINE-EARTH METALS; ELECTROLYTE; HALIDE; RARE-EARTH ELEMENTS.

The process thus yields four distinct product streams: (1) a pure uranium product that is free of fission products and free of transuranic contamination; (2) an actinide product that contains all of the transuranic elements (such as Np, Pu, Am, and Cm) and some uranium with a trace amount of rare-earths; (3) a metal waste stream emanating from the anode basket residue; and (4) a ceramic waste stream containing fission products extracted from the electrolyte salt. The uranium product can be stored, recycled, or converted to the oxide for disposal as a low-level waste. The transuranic product is suitable for recycle as a fast reactor fuel material; independent studies have shown that this product does not represent a proliferation threat, because the plutonium therein is sufficiently contaminated that the material is not usable to create weapons. In a nuclear growth scenario, the transuranic product could be stored until required as fast reactors are introduced into the mix of new reactor construction. In the absence of such growth, reprocessing for recovery of transuranic elements is inadvisable, and this product would more appropriately be converted to a disposable waste form by immobilizing the transuranics in glass or zeolite. Alternatively, the concentrated transuranic elements could be transmuted in an accelerator-driven system. *See* AMERICIUM; CURIUM; NEPTUNIUM.

One of the principal advantages of the pyroprocess is that it is expected to have a very low cost. This is a result of the simplicity of the process and the very compact nature of the equipment required. For example, a prototype electrorefiner with a footprint of little more than 11 ft^2 (1 m^2) affords a throughput of 772 lb (350 kg) of heavy metal per day. The savings in the cost of the shielded facility to house the pyroprocess, compared to conventional processes, can be significant. The batchwise nature of the pyroprocess makes it economically attractive at small plant sizes. In fact, the pyroprocess is relatively insensitive to plant scale, and large, centralized plants are not necessary for economic viability, contrary to the situation that prevails with the PUREX process. The ability to operate small-scale plants at limited costs allows reactor plants and fuel reprocessing facilities to be located within proximity of each other (reducing transportation and risk of proliferation).

Other applications. Both the PUREX process and the pyroprocess can be used in a waste management role in support of a once-through nuclear fuel cycle if the economics of this application are favorable. The PUREX process can be operated with a low decontamination factor for plutonium by eliminating the second extraction step. The pyroprocess can be operated without the liquid cadmium cathode, placing the transuranic elements in the salt waste stream that leads to a glass-ceramic waste form. Both systems are effective in placing the fission products and actinide elements present in spent nuclear fuel into more durable waste forms that can be safely disposed in a high-level waste repository. James J. Laidler

Bibliography. M. Benedict, T. H. Pigford, and W. H. Levi, *Nuclear Chemical Engineering*, 2d ed., 1981; W. H. Hannum (ed.), The Technology of the Integral Fast Reactor and Its Associated Fuel Cycle, *Prog. Nucl. Energy*, vol. 31:1–217, 1997; J. T. Long, *Engineering for Nuclear Fuel Reprocessing*, 1978; W. W. Schulz, J. D. Navratil, and L. L. Burger, *Science and Technology of Tributyl Phosphate: Applications of Tributyl Phosphate in Nuclear Fuel Reprocessing*, 1989.

Nuclear fusion

One of the primary nuclear reactions, the name usually designating an energy-releasing rearrangement collision which can occur between various isotopes of low atomic number. *See* NUCLEAR REACTION.

Interest in the nuclear fusion reaction arises from the expectation that it may someday be used to produce useful power, from its role in energy generation in stars, and from its use in the fusion bomb. Since a

primary fusion fuel, deuterium, occurs naturally and is therefore obtainable in virtually inexhaustible supply (by separation of heavy hydrogen from water, 1 atom of deuterium occurring per 6500 atoms of hydrogen), solution of the fusion power problem would permanently solve the problem of the present rapid depletion of chemically valuable fossil fuels. In power production the lack of radioactive waste products from the fusion reaction is another argument in its favor as opposed to the fission of uranium. *See* HYDROGEN BOMB; STELLAR EVOLUTION.

In a nuclear fusion reaction the close collision of two energy-rich nuclei results in a mutual rearrangement of their nucleons (protons and neutrons) to produce two or more reaction products, together with a release of energy. The energy usually appears in the form of kinetic energy of the reaction products, although when energetically allowed, part may be taken up as energy of an excited state of a product nucleus. In contrast to neutron-produced nuclear reactions, colliding nuclei, because they are positively charged, require a substantial initial relative kinetic energy to overcome their mutual electrostatic repulsion so that reaction can occur. This required relative energy increases with the nuclear charge Z, so that reactions between low-Z nuclei are the easiest to produce. The best known of these are the reactions between the heavy isotopes of hydrogen, deuterium (D) and tritium (T). *See* DEUTERIUM; TRITIUM.

Fusion reactions were discovered in the 1920s when low-Z elements were used as targets and bombarded by beams of energetic protons or deuterons. But the nuclear energy released in such bombardments is always microscopic compared with the energy of the impinging beam. This is because most of the energy of the beam particle is dissipated uselessly by ionization and interparticle collisions in the target; only a small fraction of the impinging particles actually produce reactions.

Nuclear fusion reactions can be self-sustaining, however, if they are carried out at a very high temperature. That is to say, if the fusion fuel exists in the form of a very hot ionized gas of stripped nuclei and free electrons, termed a plasma, the agitation energy of the nuclei can overcome their mutual repulsion, causing reactions to occur. This is the mechanism of energy generation in the stars and in the fusion bomb. It is also the method envisaged for the controlled generation of fusion energy. *See* PLASMA (PHYSICS).

Properties of Fusion Reactions

The cross sections (effective collisional areas) for many of the simple nuclear fusion reactions have been measured with high precision. It is found that the cross sections generally show broad maxima as a function of energy and have peak values in the general range of 0.01 barn (1 barn = 10^{-28} m^2) to a maximum value of 5 barns, for the deuterium-tritium reaction. The energy releases of these reactions can be readily calculated from the mass differences between the initial and final nuclei or determined by direct measurement.

Simple reactions. Some of the important simple fusion reactions, their reaction products, and their energy releases are given by reactions (1).

$$\begin{aligned} \mathrm{D} + \mathrm{D} &\rightarrow {}^3\mathrm{He} + n + 3.25\ \mathrm{MeV} \\ \mathrm{D} + \mathrm{D} &\rightarrow \mathrm{T} + p + 4.0\ \mathrm{MeV} \\ \mathrm{T} + \mathrm{D} &\rightarrow {}^4\mathrm{He} + n + 17.6\ \mathrm{MeV} \\ {}^3\mathrm{He} + \mathrm{D} &\rightarrow {}^4\mathrm{He} + p + 18.3\ \mathrm{MeV} \\ {}^6\mathrm{Li} + \mathrm{D} &\rightarrow 2\,{}^4\mathrm{He} + 22.4\ \mathrm{MeV} \\ {}^7\mathrm{Li} + p &\rightarrow 2\,{}^4\mathrm{He} + 17.3\ \mathrm{MeV} \end{aligned} \tag{1}$$

If it is remembered that the energy release in the chemical reaction in which hydrogen and oxygen combine to produce a water molecule is about 1 eV per reaction, it will be seen that, gram for gram, fusion fuel releases more than 10^6 times as much energy as typical chemical fuels.

The two alternative D-D reactions listed occur with about equal probability for the same relative particle energies. The heavy reaction products, tritium and helium-3, may also react, with the release of a large amount of energy. Thus it is possible to visualize a reaction chain in which six deuterons are converted to two helium-4 nuclei, two protons, and two neutrons, with an overall energy release of 43 MeV—about 10^5 kWh of energy per gram of deuterium. This energy release is several times that released per gram in the fission of uranium, and several million times that released per gram by the combustion of gasoline.

Cross sections. **Figure 1** shows the measured values of cross sections as a function of bombarding

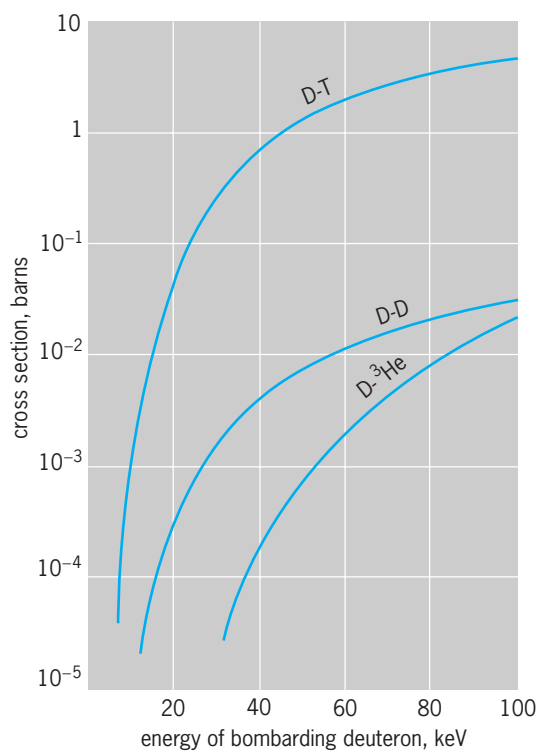

Fig. 1. Cross sections versus bombarding energy for three simple fusion reactions. (*After R. F. Post, Fusion power, Sci. Amer., 197(6):73–84, December 1957*)

energy up to 100 keV for the total D-D reaction (both D-D,n and D-D,p), the D-T reaction, and the D-^{3}He reaction. The most striking characteristic of these curves is their extremely rapid falloff with energy as bombarding energies drop to a few kilovolts. This effect arises from the mutual electrostatic repulsion of the nuclei, which prevents them from approaching closely if their relative energy is small. *See* NUCLEAR STRUCTURE.

The fact that reactions can occur at all at these energies is attributable to the finite range of nuclear interaction forces. In effect, the boundary of the nucleus is not precisely defined by its classical diameter. The role of quantum-mechanical effects in nuclear fusion reactions has been treated by G. Gamow and others. It is predicted that the cross sections should obey an exponential law at low energies. This is well borne out in energy regions reasonably far removed from resonances (for example, below about 30 keV for the D-T reaction). Over a wide energy range at low energies, the data for the D-D reaction can be accurately fitted by a Gamow curve, the result for the cross section being given by Eq. (2), where the

$$\sigma_{\text{D-D}} = \frac{288}{W} \exp\left(-45.8/\sqrt{W}\right) \times 10^{-28}\ \text{m}^2 \qquad (2)$$

bombarding energy W is in kiloelectronvolts.

The extreme energy dependence of this expression can be appreciated by the fact that between 1 and 10 keV the predicted cross section varies by about 13 powers of 10, that is, from 2×10^{-46} to 1.5×10^{-33} m^2.

Energy division. The kinematics of the fusion reaction requires that two or more reaction products result. This is because both mass energy and momentum balance must be preserved. When there are only two reaction products (which is the case in all of the important reactions), the division of energy between the reaction products is uniquely determined, the lion's share always going to the lighter particle. The energy division (disregarding the initial bombarding energy) is as in reaction (3). If reaction (3) holds,

$$A_1 + A_2 \rightarrow A_1' + A_2' + Q \qquad (3)$$

with the A's representing the atomic masses of the particles and Q the total energy released, then Eqs. (4) are valid, where $W(A'_1)$ and $W(A'_2)$ are the kinetic

$$\begin{aligned} W(A_1') + W(A_2') &= Q \\ W(A_1') &= Q\left(\frac{A_2'}{A_1' + A_2'}\right) \\ W(A_2') &= Q\left(\frac{A_1'}{A_1' + A_2'}\right) \end{aligned} \qquad (4)$$

energies of the reaction products.

Reaction rates. When nuclear fusion reactions occur in a high-temperature plasma, the reaction rate per unit volume depends on the particle density n of the reacting fuel particles and on an average of their mutual reaction cross sections σ and relative velocity υ over the particle velocity distributions. *See* THERMONUCLEAR REACTION.

For dissimilar reacting nuclei (such as D and T), the reaction rate is given by Eq. (5).

$$R_{12} = n_1 n_2 \langle \sigma \upsilon \rangle_{12} \quad \text{reactions/(m}^3 \cdot \text{s)} \qquad (5)$$

For similar reacting nuclei (for example, D and D), the reaction rate is given by Eq. (6).

$$R_{11} = {}^1/_2 n^2 \langle \sigma \upsilon \rangle \qquad (6)$$

Both expressions vary as the square of the total particle density (for a given fuel composition).

If the particle velocity distributions are known, $\langle \sigma \upsilon \rangle$ can be determined as a function of energy by numerical integration, using the known reaction cross sections. It is customary to assume a maxwellian particle velocity distribution, toward which all others tend in equilibrium. The values of $\langle \sigma \upsilon \rangle$ for the D-D and D-T reactions are shown in **Fig. 2**. In this plot the kinetic temperature is given in kiloelectronvolts; 1 keV kinetic temperature = 1.16×10^7 K. Just as in the case of the cross sections themselves, the most striking feature of these curves is their extremely rapid falloff with temperature at low temperatures. For example, although at 100 keV for all reactions $\langle \sigma \upsilon \rangle$ is only weakly dependent on temperature, at 1 keV it varies as $T^{6.3}$ and at 0.1 keV as T^{133}. Also, at the lowest temperatures it can be shown that only the particles in the "tail" of the distribution, which have energies large compared with the average, will make appreciable contributions to the reaction rate, the energy dependence of σ being so extreme.

Critical temperatures. The nuclear fusion reaction can obviously be self-sustaining only if the rate of loss of energy from the reacting fuel is not greater than the rate of energy generation by fusion reactions. The simplest consequence of this fact is that there will exist critical or ideal ignition temperatures below which a reaction could not sustain itself, even under idealized conditions. In a fusion reactor, ideal

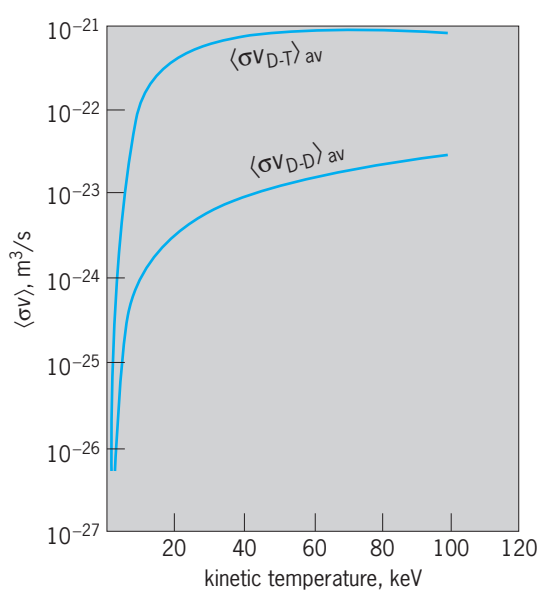

Fig. 2. Plot of the values of $\langle \sigma v \rangle$ versus kinetic temperature for the D-D and D-T reactions.

or minimum critical temperatures are determined by the unavoidable escape of radiation from the plasma. A minimum value for the radiation emitted from any plasma is that emitted by a pure hydrogenic plasma in the form of x-rays or bremsstrahlung. Thus plasmas composed only of isotopes of hydrogen and their one-for-one accompanying electrons might be expected to possess the lowest ideal ignition temperatures. This is indeed the case: It can be shown by comparison of the nuclear energy release rates with the radiation losses that the critical temperature for the D-T reaction is about 4×10^7 K. For the D-D reaction it is about 10 times higher. Since both radiation rate and nuclear power vary with the square of the particle density, these critical temperatures are independent of density over the density ranges of interest. The concept of the critical temperature is a highly idealized one, however, since in any real cases additional losses must be expected to occur which will modify the situation, increasing the required temperature. Richard F. Post

Magnetic Confinement Fusion

The application of nuclear fusion to energy production requires that the nuclear fuel have sufficient temperature, density, and confinement. If the fuel is an equal mixture of deuterium and tritium, the optimum temperature is approximately 20 keV, or 2×10^8 °C; for other fuels the optimum temperature is even higher. At such high temperatures, atoms and molecules dissociate into ions and electrons to form a state of matter called plasma. Plasmas exhibit the same relationship between pressure P, temperature T, and the number of particles per cubic meter, n, as ordinary gases, $P = nk_BT$, where k_B is the Boltzmann constant. Unlike ordinary gases, such as air, plasmas are good conductors of electricity. Indeed, a fusion plasma has a conductivity more than ten times higher than copper. *See* BOLTZMANN CONSTANT; PLASMA (PHYSICS).

The high electrical conductivity of plasmas allows the fusion fuel to be confined by a magnetic field. The force per cubic meter exerted on a conductor carrying a current of density $\mathbf{j}$ through a magnetic field $\mathbf{B}$ is $\mathbf{j} \times \mathbf{B}$, which means the force is perpendicular to the current and the field. A confined plasma must have a gradient in its pressure, ∇P, which is a force per cubic meter. Force balance, or equilibrium, between the magnetic field and plasma is achieved if Eq. (7) is satisfied.

$$\nabla P = \mathbf{j} \times \mathbf{B} \tag{7}$$

See MAGNETISM.

Equation (7), which is the fundamental equation for magnetic confinement, constrains the shape of magnetically confined fusion plasmas to being toroidal, or doughnut-shaped. The magnetic field is perpendicular to the direction in which the pressure varies; thus, a magnetic field line can never leave a constant-pressure surface. A mathematical theorem states that only one surface shape exists in three dimensions that can everywhere be parallel to a vector, such as a magnetic field; that shape is a torus. If Eq. (7) is violated, the plasma will expand at a speed set by its sound velocity. Such plasmas are inertially confined. The efficiency of magnetic confinement is measured by the ratio of the plasma pressure P to the magnetic pressure $B^2/2\mu_0$, where μ_0 is the permeability of vacuum. This ratio is called beta, given by Eq. (8).

$$\beta \equiv \frac{2\mu_0 P}{B^2} \tag{8}$$

Studies of possible designs for fusion power plants have shown that a beta of approximately 5% is required for a commercially competitive plant. In these designs, a typical plasma pressure is a few times atmospheric pressure, and a typical magnetic field strength is 5 tesla, about 100,000 times greater than the magnetic field on the Earth. *See* MAGNETOHYDRODYNAMICS.

Plasma configurations. Although the plasma configurations used for magnetic confinement fusion are toroidal, they come in a variety of forms primarily distinguished by the strength of the net plasma current. The net plasma current flows along the magnetic field lines, and does not directly change the force balance. The net current does, however, change the magnetic field following Ampère's law, $\nabla \times \mathbf{B} = \mu_0 \mathbf{j}$.

The two most prominent types of magnetic confinement devices are the tokamak and the stellarator (**Table 1**). The tokamak (**Fig. 3**) is toroidally symmetric and uses currents in coils to produce a strong magnetic field that encircles the torus in the toroidal direction, the long way around. A net plasma current produces a poloidal magnetic field, that is, field lines that encircle the torus the short way

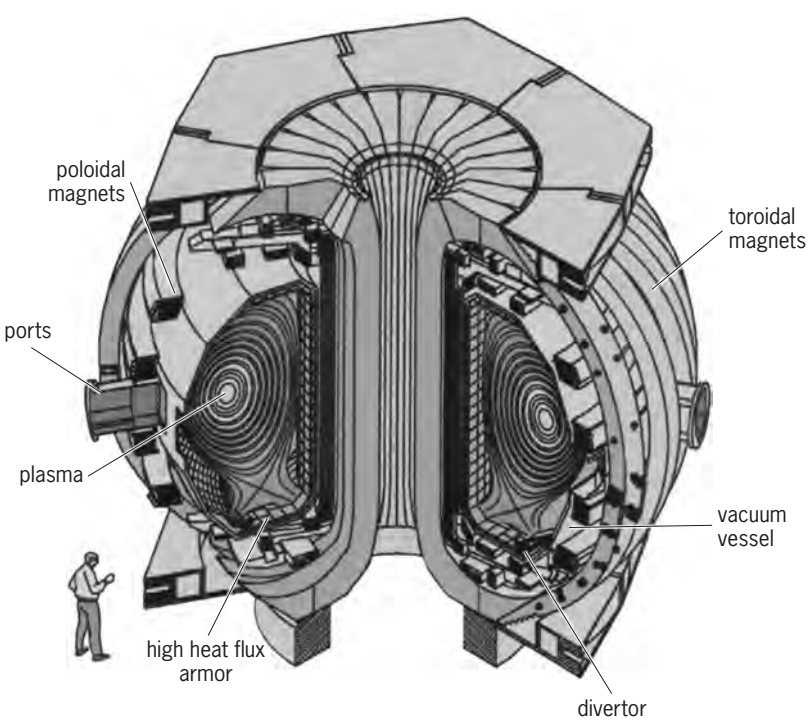

Fig. 3. Layout of the DIII-D tokamak.

TABLE 1. Some major facilities for magnetic fusion energy

Machine	Type	Location	Major radius, m (ft)	Minor radius, m (ft)	Field strength, tesla	Deuterium/tritium power, MW
DIII-D	Tokamak	United States	1.67 (5.5)	0.67 (2.2)	2.2	
JET	Tokamak	European Union	3.1 (10.2)	1.25 (4.1)	3.5	16
JT60U	Tokamak	Japan	3.4 (11.2)	1.1 (3.6)	4.2	
LHD	Stellarator	Japan	3.9 (12.8)	0.65 (2.1)	3.0	
NCSX*	Stellarator	United States	1.42 (4.7)	0.33 (1.1)	2.0	
NSTX	Spherical torus	United States	0.85 (2.8)	0.65 (2.1)	0.6	
MAST	Spherical torus	European Union	0.85 (2.8)	0.65 (2.1)	0.52	
TFTR†	Tokamak	United States	2.45 (8.0)	0.8 (2.6)	5.2	11
W7-X*	Stellarator	European Union	5.5 (18.0)	0.52 (1.7)	3.0	

*Decommisioned. †Under construction.

around. The largest tokamaks are the Joint European Torus (JET) in Britain, the JT60U in Japan, and the decommissioned Tokamak Fusion Test Reactor (TFTR) in the United States. A tokamak, ITER, which would be capable of producing fusion power at a level comparable to a power plant, has been designed under an international agreement. The design parameters are a 6.2-m (20-ft) major radius, a 2.0-m (6.6-ft) minor radius, and a 5.3-tesla magnetic field, with a planned fusion output of 500 MW. The spherical torus is a variant of the tokamak that is more compact, which means a larger ratio of the minor to the major radius of the torus.

The stellarator (**Fig. 4**) requires no net plasma current to confine a toroidal plasma configuration. External coils produce field lines that lie in nested toroidal surfaces, even in the absence of a plasma. In other words, toroidal surfaces exist in the plasma region even when $\nabla \times \mathbf{B} = 0$. The properties of magnetic fields in a current-free region imply that the surfaces of constant pressure in a stellarator cannot be toroidally symmetric; they must have helical shaping, as must the coils that produce the field. The largest stellarator is the Large Helical Device (LHD) in Japan, which has a steady-state magnetic field produced by superconducting coils. The W7-X is a stellarator experiment of similar scale being constructed in Germany.

Status. Using deuterium and tritium, the JET tokamak has produced 16 MW of fusion power, which is comparable to the power required to heat the plasma (**Fig. 5**). The next generation of tokamaks should achieve the basic plasma properties required for fusion power plants. Nevertheless, the current perception of plentiful and cheap energy, especially in the United States, has led to a critical examination of the attractiveness of energy production by tokamaks and by fusion in general. The most reliable estimates indicate that electrical power produced by fusion might cost twice as much as currently available sources. In addition, a large investment in research and development is required before the technology can be shown to be practical.

The importance of fusion energy is to ensure that an unlimited source of energy can be made widely available if fossil fuels become depleted or environmentally unacceptable. The effects of depletion and environmental damage aremade more severe by the

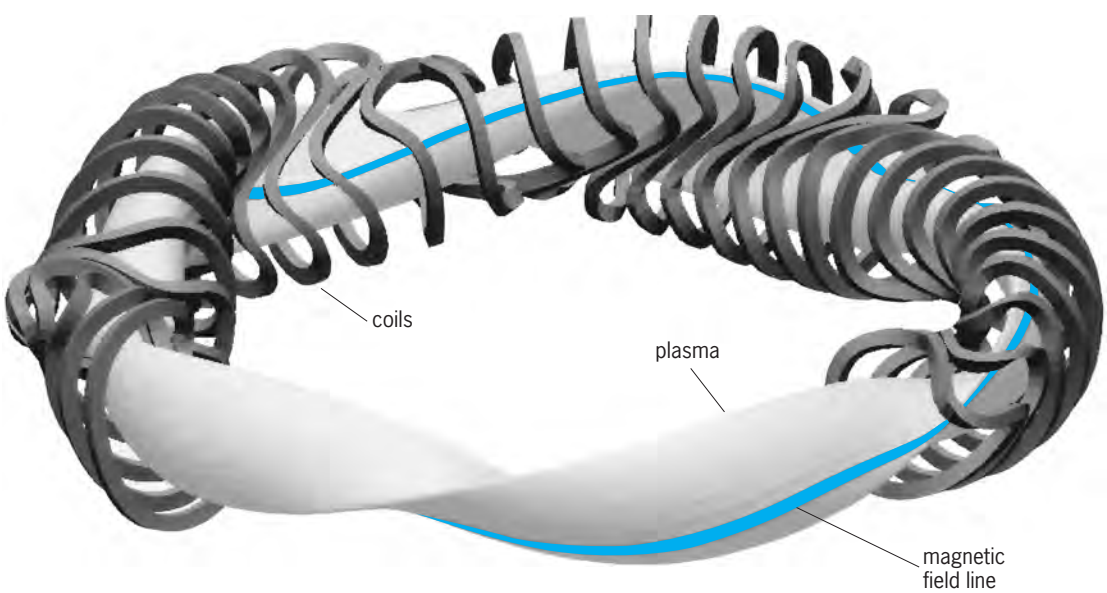

Fig. 4. Layout of the W7-X stellerator.

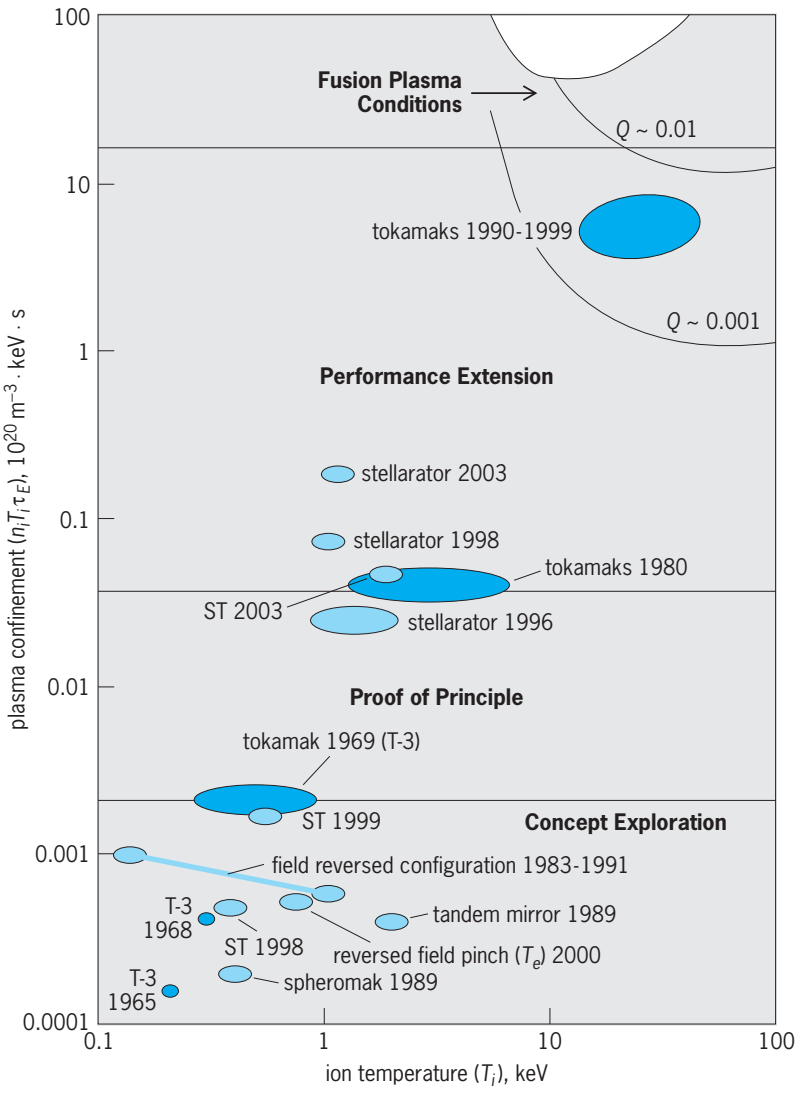

Fig. 5. Progress in magnetic fusion. The horizontal axis gives the plasma thermal energy, the temperature of the ions at the center of the plasma in kilovolts, T_i. One kilovolt is approximately 12 × 10°C. The vertical axis gives the plasma confinement, which is the product of the number of ions per cubic meter, n_i, the ion temperature in kilovolts, and the energy confinement time in seconds, τ_E, which is the thermal energy in the plasma divided by the power required to maintain the plasma temperature. The *Q* factors (fusion power divided by input power) are for D-D plasmas, which is the basis of most experiments. A fusion power plant would use a D-T plasma, which increases *Q* by an approximate factor of 1000. (*Dr. Dale Meade, Princeton Plasma Physics Laboratory*)

need for energy use to rise dramatically if a large fraction of the population of the world is not to remain in poverty.

Physical considerations. Force balance [Eq. (7)] is a necessary condition for magnetic confinement, but it is not a sufficient condition. Additional issues requiring consideration are (1) the stability of the plasma equilibrium; (2) the diffusive transport of the fuel or, more importantly, its thermal energy; and (3) the loss of thermal energy through the radiation of light. For fusion using a mixture of deuterium and tritium, the third issue is primarily one of plasma purity. Heavier elements, such as carbon or iron, can radiate at a level of power far above that produced by fusion reactions. The high temperature required for fusion can be maintained only if the concentrations of impurities are carefully controlled.

Plasma stability. The first issue, the stability of the plasma, is equivalent to asking whether the equilibrium is analogous to that of a ball at the top or the bottom of a hill. Instabilities in the force balance are of two types: those driven by the pressure gradient and those driven by the net plasma current.

Instabilities due to the pressure gradient arise from the tendency of a plasma to move to a region of lower magnetic field strength. The high conductivity of the plasma makes changing the magnetic flux embedded in an element of plasma difficult. (Magnetic flux is the strength of the magnetic field multiplied by an area transverse to the field.) By moving into a region of lower magnetic field strength, a plasma can release the free energy of expansion while conserving the embedded magnetic flux. The properties of a current-free magnetic field imply that in a torus a region of lower magnetic field strength exists in at least part of the periphery of the plasma. However, the plasma is not necessarily unstable toward moving into this region of reduced field strength (that is, the possibility of such movement is not necessarily analogous to the possibility of a ball rolling down from the top of a hill) because a magnetic field deformation is generally required in order for such movement to take place. The requirement for a deformation of the magnetic field stabilizes the plasma up to a critical value of the plasma pressure, or more precisely β. Values of β well above the minimal value for a power plant, $\beta \approx 5\%$, have been achieved in tokamak experiments.

The second type of instability is driven by the net current, $\boldsymbol{j}_n$. Such instabilities are known as kink instabilities from the form of the plasma deformation that is produced. Kink instabilities limit the magnitude of the net plasma current. This limit is usually expressed using the safety factor, q, which is the number of times a magnetic field line must encircle the torus toroidally, the long way, before encircling the torus poloidally, the short way. The magnitude of the safety factor varies across the plasma. In a stable tokamak equilibrium, q is typically about 1 near the plasma center and 3 near the edge.

Energy transport. As mentioned above, the second issue concerns the transport properties of the plasma. Although magnetically confined plasmas are in force equilibrium [Eq. (7)], they are not in thermodynamic equilibrium. Temperature gradients, which represent a deviation from thermodynamic equilibrium, lead to a transport of energy just as a temperature gradient in air does. As in air, the transport of energy in a plasma due to a temperature gradient can be much larger than that predicted by the standard transport coefficients. In air the enhancement arises from convection. In fusion plasmas the transport is due to small-scale fluctuations in the electric and magnetic fields. The enhanced transport arising from these fluctuations is called microturbulent or anomalous transport. *See* CONVECTION (HEAT); HEAT TRANSFER.

The size of a fusion plasma, about 2 m (6 ft) for the smaller radius of the torus, is comparable by the scale of the blankets and shields that surround the plasma.

The radial transport coefficients need to have a particular magnitude to make that radius consistent with the removal of energy from the plasma at the same rate as the plasma is heated by fusion reactions. Remarkably, the level of transport observed in experiments is close to that magnitude. Nevertheless, the uncertainties associated with microturbulent transport are among the largest uncertainties in setting the minimum size, and therefore minimum cost, for an experiment to achieve the plasma confinement required for a power plant. Much progress has been made in the development of methods for computing microturbulent transport. Advances are paced by the development of faster computers and more sophisticated algorithms. In addition, experiments have shown that it is possible to manipulate the plasma in ways that give large reductions in microturbulent transport.

Collisionality paradox. Fusion plasmas are in a paradoxical collisionality regime. The ions and the electrons that form a fusion plasma are said to be collisionless, since they move about 10 km (6 mi) between collisions and the plasma has a scale of only a few meters. On the other hand, fusion plasmas are highly collisional, for the energy of the ions and electrons must be confined for more than a hundred collision times. This implies that the velocity of ions and electrons obeys the same probability distribution as air molecules, the maxwellian distribution. *See* KINETIC THEORY OF MATTER.

In a tokamak the collisionality paradox presents no fundamental problem. The toroidal symmetry of the tokamak plus the theory of hamiltonian mechanics ensure that a quantity called the toroidal canonical momentum is invariant along the trajectory of the particle. This invariance prevents the trajectories of the ions and electrons from straying far from the constant-pressure surfaces.

A stellarator cannot be toroidally symmetric, and careful design is required to ensure the confinement of particle trajectories. Surprisingly, the confinement of trajectories is controlled by the variation of the magnetic field strength in a constant-pressure surface and not by the shape of the surfaces. The surfaces of constant pressure in a stellarator cannot be toroidally symmetric, but the field strength can be almost independent of the toroidal angle in a pressure surface, and this quasi-symmetry confines the trajectories. Trajectories in a stellarator can also be confined using a subtle type of invariant of hamiltonian mechanics, the longitudinal adiabatic invariant. It is this invariant that is used to confine the particle trajectories in the W7-X stellarator. *See* HAMILTON'S EQUATIONS OF MOTION.

Stellerator advantages. The reasons that stellarators are considered attractive for fusion are the robust stability of stellarator plasmas and the absence of a requirement for power to drive a net current. The robust stability of stellarator plasmas means they do not experience sudden terminations, which would exert a large stress on the surrounding structures. Sudden terminations, called disruptions, are observed in various operating conditions of tokamak plasmas and would not be acceptable in a power plant.

Technical considerations. In addition to the physical constraints on the design of fusion power plants, there are technical constraints: the first wall, the blanket and shields, and the coils. The solid surface immediately surrounding the plasma is called the first wall. Suitable materials are under study. Two issues concerning the first wall have to do with the neutrons that are emitted by a fusing plasma: (1) The attractiveness of fusion depends on the first wall not becoming highly radioactive. (2) The first wall must maintain its structural integrity even though neutron bombardment tends to cause materials to swell and become brittle. A third issue is economic. Fusion power becomes cheaper the higher the energy flux that can pass through the first wall. A power flux of 5 MW/m^2 is a typical design value, with 80% of that power in the neutrons that are produced by deuterium/tritium fusion reactions. *See* RADIATION DAMAGE TO MATERIALS.

Tritium is not naturally available. In a deuterium/tritium fusion power plant, a blanket is required to produce tritium by the bombardment of lithium with neutrons. Behind the blanket are shields to prevent neutron bombardment of the current-carrying coils. The required thickness of the blanket and shield is approximately 1.5 m (5 ft) and sets a natural size scale for a fusion power plant, which is approximately 1000 MW of electric power.

The coils that produce the magnetic field in a fusion power plant must be superconducting if the plasma β value is to be approximately 5%. Ordinary copper coils would dissipate far too much energy. The magnetic field produced by the coils is limited both by the properties of superconductors and by the mechanical stress, $B^2/2\mu_0$. A commonly used limit is that of Nb_3Sn superconductors, 12 tesla. The field in the plasma is lower, and an important efficiency factor is the ratio of the average magnetic field strength in the plasma to maximum field at a coil. A typical value for this ratio is 0.5. *See* SUPERCONDUCTING DEVICES.

Allen H. Boozer

Inertial Confinement Fusion

The basic idea of inertial confinement fusion is to use a laser or particle beam (often called a driver) to deliver sufficient energy, power, and intensity [$\sim 5 \times 10^6$ joules, 10^{15} W in about 8 to 10 nanoseconds, and 10^{14} W/cm^2 to a capsule several millimeters in diameter, containing several milligrams of a mixture of deuterium and tritium (D-T)], to produce an implosion. This implosion is designed to compress and heat the D-T mixture to a density of about 200 g/cm^3 (200 times that of water or approximately 1000 times the liquid density of D-T) and a temperature of 3–5 keV. With this range of drive energy and under these conditions, 1–10 mg of D-T could be ignited and, while inertially confined, with a burn efficiency of about 30%, could release 100–1000 megajoules of fusion energy, with energy gains, depending upon specific target and driver conditions, from 10 to greater than 100. An energy of

TABLE 2. Large-scale inertial confinement fusion facilities

Driver	Location	Type	Power, TW	Energy, kJ	Wavelength, μm	Number of beams
		Laser				
Nova	Lawrence Livermore National Laboratory, United States	Neodymium:glass	120	120	1.06	10
			30–60	50–80	0.53	
			20–40	40–70	0.33	
Aurora	Los Alamos National Laboratory, United States	Krypton fluoride	1–23	[5–10]*	0.25	48
Omega	University of Rochester, United States	Neodymium:glass	12	4	1.06	24
			6	2	0.33	
Phebus	Centre d'Etudes de Limeil, France	Neodymium:glass	24	24	1.06	2
Vulcan	Rutherford-Appleton Laboratory, United Kingdom	Neodymium:glass	4	1.0	1.06	12
			2	0.4	0.53	
Gekko XII	Osaka University, Japan	Neodymium:glass	55	30	1.06	12
					0.53	
Asterix III	Max Planck Institute for Quantum Physics, Germany	Iodine	5	2	1.3	1
		Particle beams				
PBFA II	Sandia National Laboratory, United States	Lithium	[100]*	[2000]†		
Kalif	KFK, Karlsruhe Physics, Germany	Hydrogen	0.7	20‡		
SIS/ESR	GSI, Darmstadt Physics, Germany	Neon	45 MW	0.045		
		Xenon	0.036	2.5		

*[ ] denotes design goals, not yet achieved. †15-ns pulse. ‡50-ns pulse.

100 MJ is equivalent to about 50 lb (25 kg) of TNT, or about 0.5 gallon (2 liters) of oil, an amount of energy tractable under laboratory conditions and potentially useful for a variety of applications, including basic and applied science, weapons physics and weapons effects, energy production, and production of special nuclear materials and nuclear fuels.

Drivers with light ions, heavy ions, krypton fluoride (KrF) lasers, and neodymium:glass lasers are being developed. Although they are at widely different stages of development, all four of these technologies could probably provide a driver capable of delivering 5–10 MJ. It is not clear, however, that they can all deliver this energy while meeting the additional requirements of flexible temporal pulse shape, 10^{15} W peak power, greater than 10^{14} W/cm^2 intensity, and focusing of the beam over the standoff distance. Neodymium:glass laser technology is probably the most developed of the four. **Table 2** lists some representative research facilities.

The two fundamental approaches to most inertial confinement fusion research are differentiated by the mechanism chosen to drive the process. In the direct-drive approach, laser (or ion) beams are arrayed around the target in a near-uniform pattern and aimed directly at the D-T fuel capsule. In the indirect-drive approach, the driver energy is deposited (absorbed) in an intermediate medium that heats up to temperatures high enough to emit x-rays. The fuel capsule is contained inside this x-ray converter (hohlraum), and the x-rays are used to drive the capsule.

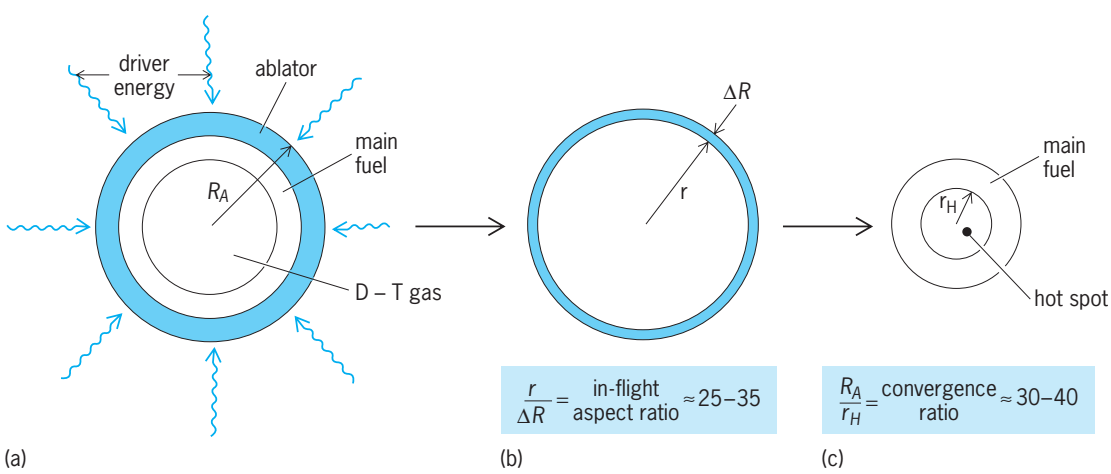

Fig. 6. Direct-drive inertial confinement fusion target. The radial elements of the (*a*) initial, (*b*) in-flight, and (*c*) final fuel configurations are identified. Additional requirements for a successful implosion include peak pressure of ~100 Mbar (10 TPa), flux uniform to about 2%, x-ray and electron preheat to the fuel of less than a few electronvolts, accurate pulse shaping, and very smooth target surface (roughness of ~0.1 μm).

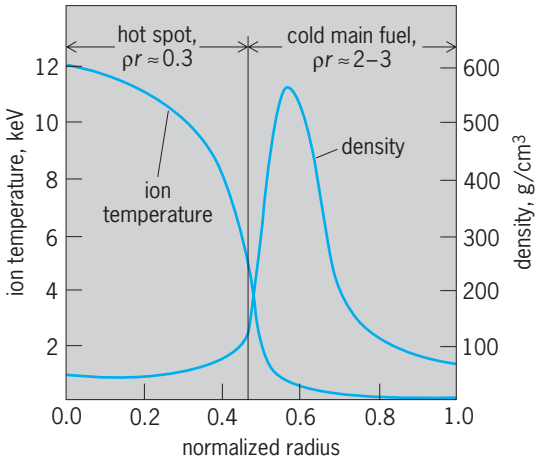

Fig. 7. Deuterium-tritium fuel ion temperature and density at ignition, plotted against normalized radius, for a typical, direct-drive high-gain capsule. Hot-spot value of ρr must exceed alpha-particle range for effective self-heating from 5 keV.

The details of indirectly driven targets are classified. Therefore, in order to give a coherent description of the capsule physics, only the requirements for directly driven targets will be discussed.

Implosion physics. The events leading to a capsule burn are as follows. The driver deposits its energy by electron conduction (directly driven) in the outer layers of the fuel-containing capsule (**Fig. 6**). This energy deposition in turn ablates the surface of the capsule, and the ablation acts as a rocket exhaust to drive the implosion of the capsule. In the outer portions of the capsule, the ablation generates pressures from 1 megabar (100 gigapascals) to about 100 Mbar (10 terapascals) over a period of about 10 ns, and this accelerates the material of the capsule inward to velocities of $\sim 10^6$ ft/s (3×10^5 m/s). The inward acceleration continues until the internal pressures exceed the pressures of ablation. At that time, the rapidly converging shell begins to decelerate, compresses the central portion of the fuel to densities of the order of 200 g/cm^3, and then heats the core to about 5 keV, whereupon thermonuclear burn is initiated.

The final high fuel densities for inertial confinement fusion are required to burn the D-T efficiently. In inertial confinement fusion, the product of fuel density ρ and fuel radius r is an important parameter, analogous to the $n\tau_E$ product in magnetic fusion. The burnefficiency η for inertial confinement fusion

Fig. 8. Nova inertial confinement research facility at Lawrence Livermore National Laboratory. Laboratory building containing neodymium:glass laser system is separated into four clean-room areas: master oscillator room on lower level where initial pulse light is generated; laser bay containing 10 main laser amplifier chains; optical switchyard in which the 29-in.-diameter (74-cm) beams are diagnosed and reflected into the target room; and target area that houses the 15-ft-diameter (4.6-m) aluminum target chamber. (*Lawrence Livermore National Laboratory*)

depends on ρr; specifically, it is given by Eq. (9),

$$\eta = \frac{\rho r}{\rho r + 6} \tag{9}$$

where ρr is expressed in g/cm^2. A ρr value of 3 g/cm^2 using liquid-density D-T can be achieved only with a large amount of D-T (approximately 6 lb or 3 kg). The fusion yield of that much D-T, however, is about 70 kilotons, which cannot be contained in a practical way. Instead, if the fuel is compressed to 1000-times-liquid-density D-T, then about 5 mg of D-T is sufficient to obtain the same ρr, and the yield is then about 250 lb (125 kg) of TNT, which can be readily contained. This requirement is known as high compression.

Although high compression is necessary, it is not sufficient. To achieve high gain, inertial confinement fusion also relies upon central ignition and propagation of the burn into the compressed D-T. If the fuel is compressed to a ρr of ~3 g/cm^2 and heated to 5 keV, then 5 mg of D-T releases about 500 MJ. However, to heat the entire fuel mass to 5 keV requires too large a driver in terms of size and cost. Instead, a small portion of the fuel mass—less than 5%—at the center of the capsule is heated to ignition temperatures, while the rest of the fuel remains cold but compressed to ~200 g/cm^3. **Figure 7** depicts such a configuration, where the hot-spot value of ρr is just large enough to stop the alpha particles generated from fusion reactions in the hot-spot ignition region, starting the self-heating process and propagating the burn into the main fuel.

This sequence of events is complicated by some facts of nature and the need to build drivers (either lasers or particle beams) whose energy, power, and focusability are limited by costs and technology. These make it necessary to use a driver with a maximum energy output of ~10 MJ, in ~10 ns, and peak power of 10^{15} W. The basic facts of nature are as follows.

1. In order to achieve the final fuel densities of ~200 g/cm^3 with the 100-Mbar (10-TPa) peak pressures available with laser intensities in the 10^{14}–10^{15} W/cm^2 range, it is necessary to start with the D-T fuel in a liquid or solid shell.
2. The implosion must be carried out in a reasonably isentropic manner until a late stage of the implosion, or else the continual heating of the fuel will make a high compression too demanding on driver energy.
3. Potential hydrodynamic instabilities demand both target surface finishes that are nearly perfect and pressures that are uniform to a few percent, to avoid mixing of the outer portion of the shell (the ablator) and minimize mixing of the inner portion of the shell (the pusher) with the fuel and subsequent degradation of the burn efficiency.

The critical issues for the success of inertial confinement fusion are whether a sufficiently uniform capsule can be made, whether sufficiently uniform pressure pulse can be produced, and whether there is full understanding of the hydrodynamics and physics of the D-T burn on which the models are based. The information now available indicates that high gain can be achieved with a 5–10-MJ driver.

Target technology. A free-standing, liquid D-T shell would be the ideal direct-drive target, but it is still not known how to make one. However, a free-standing solid layer of D-T is nearly as good, and it appears that the natural heat generated by the tritium beta decay in a D-T mixture may offer a natural process to produce the required thicknesses of solid D-T layers with the required uniformity, a technique called beta layering. Uniform spherical D-T layers up to several millimeters in thickness have been produced. Spherical plastic shells to contain the D-T and provide the ablator for the implosion have been made in the 0.5-mm-diameter range with the required surface finishes of less than 100-nm thickness. Another option is to use a shell of low-density plastic foam to wick up liquid D-T and make an analog D-T liquid layer. With a plastic foam of density 50 mg/cm^3, the target would be 80% liquid D-T and 20% low-density foam. If it were made with a sufficiently uniform surface finish, it would be a very close approximation to the liquid D-T layer shell. Such low-Z, porous foams can be machined or molded to a surface finish of less than 1-μm thickness.

Target experimental facilities. The Nova laser facility (**Fig. 8**) is a good example of the very large sophisticated, and flexible laboratory facilities that have been developed for studying inertial confinement to fusion. Its 10 arms produce laser pulses that together can provide the greatest energy (10^5 J) and power (10^{14} W) of any driver, and deliver it to a target in 10^{-9} s. The neodymium: glass laser produces infrared

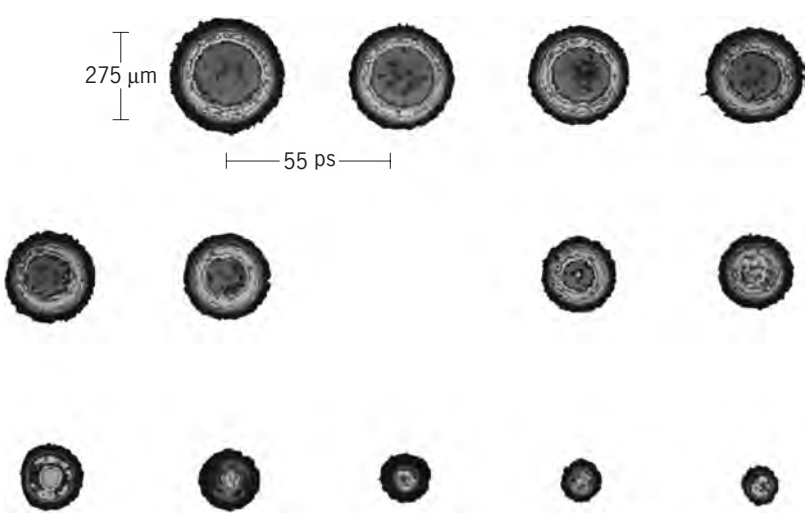

Fig. 9. X-ray images with a framing camera with 80 ps exposure time (developed at Lawrence Livermore National Laboratory) from a directly driven inertial confinement fusion capsule at the University of Rochester's Omega Facility. Camera has 80 ps exposure time and 55 ps time between frames; because exposure time exceeds time between frames, there is some temporal overlap. (Data from missing frame was unavailable.)

light at a wavelength of about 1.06 micrometers, and the system was built with the capability to convert the frequency of that laser light to the second (green) or third (blue) harmonics. These shorter wavelengths greatly improve the coupling of the light to the target and help in achieving high compression.

The laser power must be delivered with a precise temporal pulse shape, ranging in duration from 1 to 10 ns. Nova and other solid-state lasers can produce a variety of complex pulse shapes from simple gaussian-shaped pulses to pulses with a monotonically increasing temporal profile, or combinations with so-called picket fences that approximate the desired monotonically increasing drive pulse.

Diagnostic technologies. Sophisticated implosion diagnostics have been developed. **Figure 9** shows a series of x-ray images of an imploding capsule taken with an x-ray framing camera with ~80 picosecond exposure time. Other diagnostics include neutron pinhole cameras with about 20-μm spatial resolution, neutron time-resolving detectors with about 50-ps resolution, and a variety of x-ray optics such as x-ray microscopes with 1–2-μm resolution, normal-incidence x-ray mirrors with reflection up to 50%, and a variety of x-ray gratings and beam splitters.

Experimental progress. The primary focus of the program at Nova is the indirect-drive approach with targets of a design that scales to the high-density regime. With temporally shaped pulses and appropriately scaled values for the drive temperature, values of pressure, preheat, and symmetry have been measured that meet or exceed those required for high gain. Values of $n\tau_E$ of about 2–5 $\times$ 10^{14} $cm^{-3} \cdot s$ and fuel ion temperatures of 1.5–2.0 keV have been obtained, which are less than a factor of 10 away from the values needed to achieve fusion ignition (**Fig. 10**). These results have provided stringent tests of theoretical modeling capabilities and confirm that indirectly driven capsules with a 5–10-MJ driver should achieve the $n\tau$ values and fuel conditions required for ignition and high gain.

Significant advances have also been made in the direct-drive approach. There has been progress in understanding and controlling illumination uniformity which is critical to direct-drive fusion and also important to the energy–to–x-ray conversion efficiency of indirect drive.

By using an improved beam-smoothing technique, excellent implosion performance has been demonstrated using gas-filled glass shells. If similar improvements can be obtained for target designs that have the higher in-flight aspect ratios (ratio of the shell thickness to the shell radius) ultimately required, it will mean that direct-drive targets also hold the promise of high gain for drive energies in the 5–10-MJ range.

These results and others have led the U.S. Department of Energy to undertake planning of the Laboratory Microfusion Facility (LMF), which would provide a yield of up to 1000 MJ. In addition to military applications, this facility would provide the basis for a decision to initiate engineering development aimed toward practical fusion power. Eric Storm

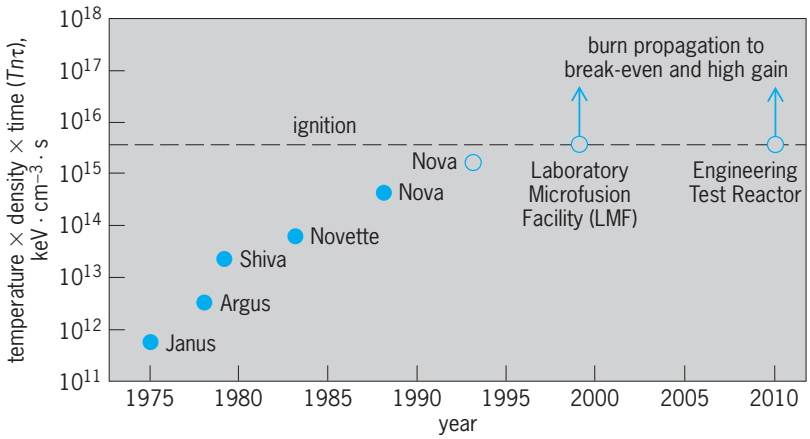

Fig. 10. Progress of inertial confinement fusion experiments in achieving conditions required for high gain. The product of ion temperature *T*, number density *n*, and confinement time τ is plotted against calendar year. Solid circles have been achieved; open circles represent future goals.

Aneutronic Fusion

Certain reactions of light nuclei other than the fusion of deuterium and tritium, the reaction studied in conventional nuclear fusion research, are attractive because they are aneutronic and nonradioactive, meaning that they generate negligible amounts of neutrons and radioactive nuclei. A revival of interest in this energy source in the mid-1980s was stimulated by the introduction of new concepts borrowed from particle physics research.

While aneutronic fusion has many potential advantages, the concept also has many difficulties, mostly associated with the very high energies required for fusion. Since fusion research has concentrated on D-T systems, quantitative theoretical information and experimental data on aneutronic fusion are relatively meager. Without radically new experimental results, this form of fusion, therefore, must be regarded more as a concept for exploration than as a technology.

Advantages of aneutronic power. The absence of neutrons and radioactivity obviates the need for shielding. This is particularly significant for aerospace applications, since the weight of shielding in a nuclear fission or fusion reactor is at least 100 times greater than that of the reactor itself.

An aneutronic reactor also offers the advantages of nonradioactive fuel and nonradioactive waste. Furthermore, since all nuclear energy released in aneutronic reactions is carried by charged particles, if these particles could be directed into a beam a flow of electric charge would result, and nuclear energy could be converted directly into electrical energy, with no waste heat.

An aneutronic reactor could be small, producing 1–100 MW of electric power, and mass production might be possible. Finally, aneutronic reactors cannot breed plutonium for nuclear weapons.

Definitions. A nuclear reaction is defined as aneutronic if its neutronicism N, the power carried by the neutrons as a fraction of the total power released in the reaction, is less than 0.01. A reaction is defined as nonradioactive if its radionuclide content R, the sum of the number of radioactive nuclei before and after the reaction as a fraction of the total number of nuclei involved, is less than 0.01. Since R is proportional to N, an aneutronic reactor usually is also nonradioactive.

Even in a reactor based on aneutronic reactions, parallel or secondary reactions will always produce some neutrons. The objective of aneutronic energy research is to design a reactor that would suppress the neutronicism to 1% or below.

Reactions. Aneutronic reactions can be divided into two classes: fission of light metals by protons, or helium-3 (^{3}He) nuclei, such as reactions (10)–

$$p + {}^7\text{Li} \rightarrow 2a + 17\ \text{MeV} \tag{10}$$

$$p + {}^6\text{Li} \rightarrow \alpha + {}^3\text{He} + 4\ \text{MeV} \tag{11}$$

$${}^3\text{He} + {}^6\text{Li} \rightarrow 2\alpha + p + 17\ \text{MeV} \tag{12}$$

$$p + {}^{11}\text{B} \rightarrow 3\alpha + 6.8\ \text{MeV} \tag{13}$$

(13); and fusion involving helium-3, suchas reactions (14) and (15). Reaction (15) has been called the ultimate aneutronic reaction: both its N and R are nearly zero.

$$\text{D} + {}^3\text{He} \rightarrow \alpha + p + 18\ \text{MeV} \tag{14}$$

$${}^3\text{He} + {}^3\text{He} \rightarrow \alpha + 2p + 13\ \text{MeV} \tag{15}$$

Critical issues. The primary obstacle to using aneutronic reactions as power sources is the lack of means to energize the fuel ions by plasma heating to the high energies required to initiate such reactions. Their reactivity approaches that of D-T fusion only at energies 10–100 times higher (Fig. 1).

The next important issue is how to increase by a factor of 10 the parameter beta, in order to suppress large electromagnetic power losses (bremsstrahlung and synchrotron radiation) from the reacting plasmas. These losses, proportional to Z^2, are negligible in hydrogen plasmas, but aneutronic fuels have ions with Z equal to 2, 3, or 4. *See* BREMSSTRAHLUNG; SYNCHROTRON RADIATION.

Studies in the early 1980s concluded that maxwellian plasmas are incompatible with aneutronic fuels because of the large ion energies required and the electromagnetic power losses.

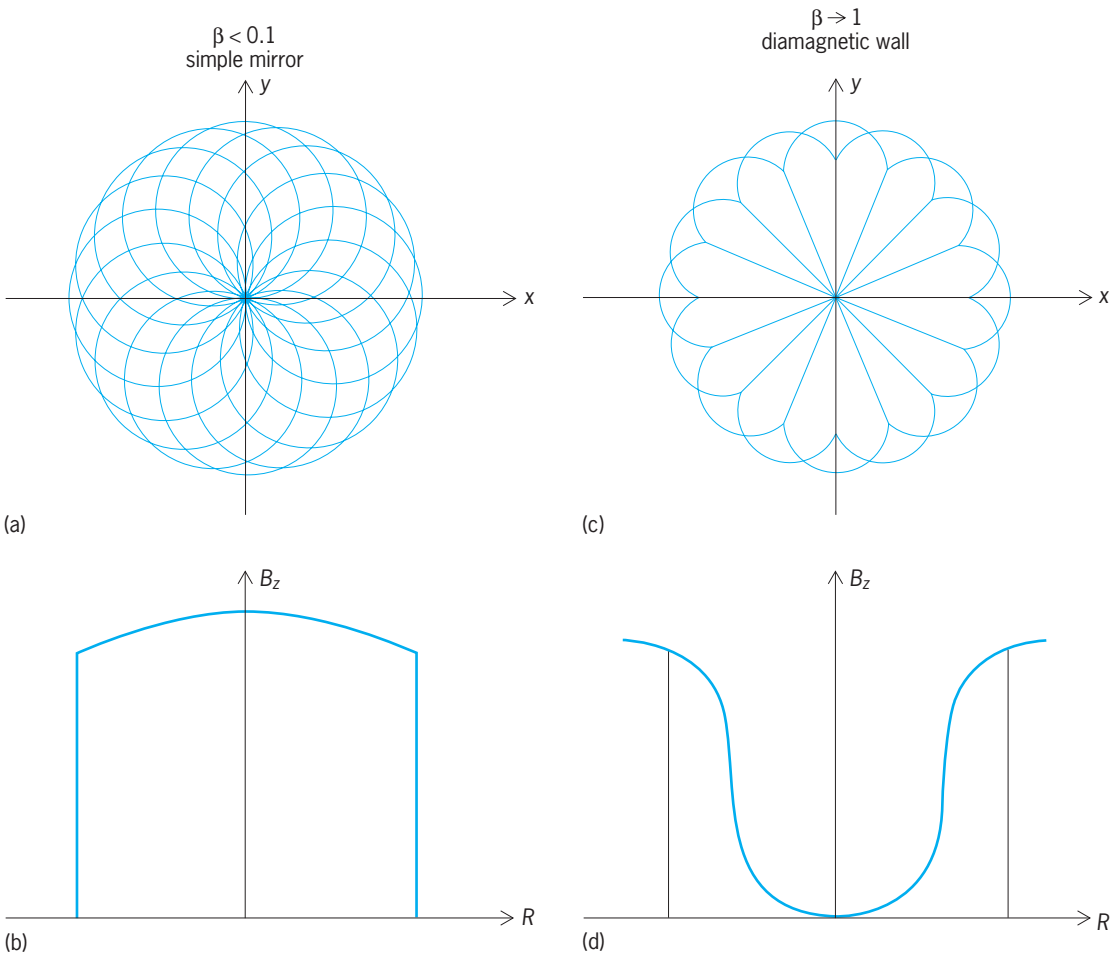

Fig. 11. Migma-plasma devices. (*a*) Ion orbits in simple-mirror magnetic field configuration at low ion density ($n < 10^{12}$ cm^{-3}; $\beta < 0.1$). (*b*) Graph of magnetic field component B_z perpendicular to orbit plane versus radial distance R in this configuration. (*c*) Ion orbits when higher ion densities ($n \sim 10^{14}$ cm^{-3}; $\beta \rightarrow 1$) create a diamagnetic well. (*d*) Graph of B_z versus R in this case.

New concepts. New concepts, however, may solve or bypass these obstacles.

Direct heating. Potential solutions to the problem of raising ions to the required energies came from colliding-beam technology and beam-storage physics, introduced in particle physics research in the 1970s. Compared to plasma ion confinement, colliding beams are characterized by many-orders-of-magnitude greater particle energies and confinement times, very large beta (near or above 1), and orders-of-magnitude lower vacuum pressure. The so-called classical Coulomb interaction processes that have limited plasma confinement times are inversely proportional to the $^3/_2$ power of the ion energy, and directly proportional to the vacuum pressure; thus, classical confinement becomes many orders of magnitude easier to achieve at the higher ion energies and ultrahigh vacuum that characterize colliding beams. Colliding beams are not neutralized, however, and the particle densities are orders of magnitude lower, limited by the electrostatic repulsion known as the space-charge limit. *See* PARTICLE ACCELERATOR.

Self-colliding beams can be neutralized in a migma-plasma, a synthetic type of plasma that can be made without heating, consisting of large-orbit, megaelectronvolt ions and thermal, ambient electrons. This concept was demonstrated in 1982. Accelerated ions were injected into a 6-T magnetic field in a so-called simple-mirror configuration to form a self-colliding orbit rosette pattern (**Fig. 11***a* and *b*) in the presence of the electrons that neutralized the stored beam.

The Self-Collider IV or Migma IV migma-plasma device (**Fig. 12**) achieved a stable confinement of 0.7-MeV ions, with an energy confinement time of 25 s, and density of 10^{10} cm^{-3}. None of the disruptive instabilities observed in or predicted for thermal plasmas at these and lower densities have been observed in migma-plasma. The figure of merit known as the triple product (temperature × confinement time × density) reached 4×10^{14} keV · cm^{-3} · s, approaching that of the best thermal plasma devices at that time.

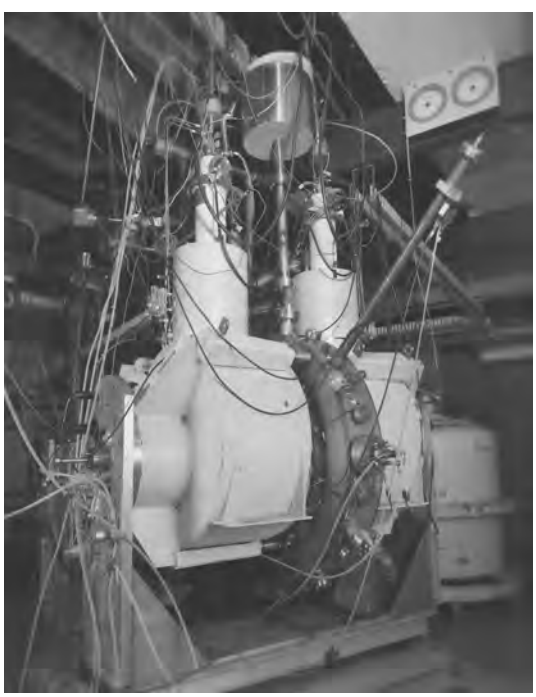

Fig. 12. Self-Collider IV or Migma IV migma-plasma device.

Indirect heating via D-T reaction. B. Coppi proposed achieving the 100-keV temperatures required to ignite reaction (13) in a compact toroidal system by first igniting the D-T fuel and then replacing the tritium with helium-3. He also predicted a so-called second stability region for the tokamak in which $\beta = 0.3$ could be achieved. A field-reversed mirror configuration (discussed above) has also been proposed to achieve high beta values to ignite reaction (13) via D-T ignition.

Avoiding radiation losses. As the plasma density increases to the values that would be required in a reactor (the parameter beta approaches unity), the diamagnetic field produced by the migma ions will push out the magnetic field from in the central zone to the periphery and create a diamagnetic well (Fig. 12). (Such a 90% diamagnetic depression was observed in the 2X-IIB mirror.) The orbit trajectories in the free-field region become linear so that synchrotron radiation is not emitted there. The diamagnetic well, in addition to making the mirror ratio large (which is crucial for ion confinement), suppresses by orders of magnitude the synchrotron radiation loss.

Reactor simulation. Extensive studies carried out on supercomputers support the concept of a reactor based on reaction (13). Preliminary results on the so-called ultimate aneutronic reaction (14) in a migma indicate that the theoretical energy balance can be reached, but the reaction is not ignited and must be driven.

Bogdan Maglich

Muon-Catalyzed Fusion

Nuclear fusion is difficult because of the electrostatic (Coulomb) repulsion between nuclei. In traditional approaches, such as magnetic confinement and inertial confinement fusion, this barrier is overcome by energetic collisions at extraordinarily high temperatures, $\sim 10^8$ K. In muon-catalyzed fusion, the barrier is effectively removed by binding the nuclei in an exotic molecule with a very short bond length. Fusion then occurs quickly at normal temperatures. (The process is often called cold fusion, not to be confused with the claim of cold fusion in condensed matter.) Muon-catalyzed fusion was discovered experimentally by L. A. Alvarez and collaborators in 1956 but had been hypothesized by F. C. Frank in 1947. Interest in the process as a possible energy source was rekindled by theoretical predictions in the late 1970s and experimental findings in the early 1980s.

In the normal deuterium molecule D_2 or molecular ion $D^+{}_2$, with bond lengths $\sim 10^{-10}$ m, fusion is unobservably slow (calculated to take $\sim 10^{64}$ s). However, when the electron of $D^+{}_2$ is replaced by a negative muon, the bond distance is reduced by approximately the ratio of muon to electron masses

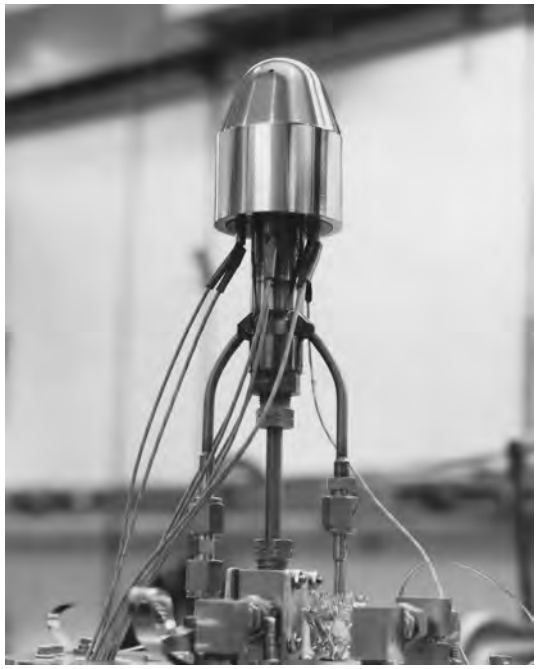

Fig. 13. Target and connections used for muon-catalyzed fusion experiments at the Los Alamos Meson Physics Facility. The deuterium-tritium mixture is inside the high-strength stainless steel vessel [outer diameter 3 in. (7.6 cm), interior volume 2.26 in.^{3} (37 cm^3)], and the muons are incident at the rounded nose where the vessel is thinnest. The target is filled through the center stem and can be cooled by liquid helium flowing through the side tubes or heated by the four resistive heaters. The white wire goes to a temperature sensor. In the experiment, the target is surrounded by five neutron detectors. (*Idaho National Engineering Laboratory*)

(m^μ/m_e), which is 207. The muon behaves like a heavy electron except that it decays with an average lifetime $\tau_0 = 2 \times 10^{-6}$ s. The resulting small molecular ion is usually denoted dd^μ (analogously, D_2^+ could be denoted *dde*). Though the average distance between deuterons is still large compared with the distance where fusion can occur ($\sim 4 \times 10^{-15}$ m), the nuclei can readily quantum-mechanically tunnel to the shorter distance. For dd^μ and dt^μ (d = deuteron = ^{2}H and t = triton = ^{3}H), the times required for fusion (2×10^{-9} and 8×10^{-13} s, respectively) are short compared with the lifetime of the muon, though the times required to form these molecules are much longer. Muon-catalyzed fusion of nuclei with charges greater than 1 is impractical since formation of a molecule containing more than one muon is extremely improbable.

The dt^μ molecule is of greatest interest for possible applications because fusion catalysis with it is most rapid and has least losses. A target used for high-pressure gas or liquid *d-t* mixtures is shown in **Fig. 13**. Muons from an accelerator are given energies such that they penetrate the steel vessel and are stopped in the fluid. The subsequent fusion neutrons also penetrate the vessel and are detected outside.

Fusion cycle. The *d-t* muon-catalyzed fusion cycle (**Fig. 14**) starts with a free muon (directly from the accelerator or recycled from a previous fusion) being captured by a deuterium or tritium atom (capture occurs by displacing an electron). If initially captured by deuterium, it transfers to tritium, where it is more strongly bound. At this point an unusual mechanism allows the t^μ atom to combine rapidly with a D_2 (or D-T) molecule to form dt^μ. A resonance is made possible by the fortuitous existence of a state of dt^μ so weakly bound that its binding energy can go into vibrational and rotational excitation of the electronic molecule. In this process the positively charged dt^μ is so small that it acts essentially as a mass-5 hydrogenic nucleus in the electronic molecule to form a compound molecule (Fig. 14). After the dt^μ is formed, nuclear fusion rapidly ensues, and an energetic neutron and alpha particle (^{4}He) are emitted. Since the total time required for these events is designated τ_c, during the lifetime τ_0 of the muon there is time for τ_0/τ_c cycles. *See* MOLECULAR STRUCTURE AND SPECTRA.

Yield. Usually the muon is left behind by the fusion products to begin another catalytic cycle. However, sometimes the muon will be lost as a catalyst before it decays. This loss is primarily due to the probability ω_s that the muon sticks to a fusion alpha particle to form a muonic helium atomic ion (He-μ^4). Hence, the average number of fusions (yield) Y_n that can be catalyzed by a single muon is given by Eq. (15).

$$Y_n = \frac{1}{\dfrac{\tau_c}{\tau_0} + \omega_s} \tag{15}$$

The optimum value of τ_0/τ_c thus far attained by experiments at temperatures up to 80°F (300 K) is ~250 with a 60:40 *d:t* mixture at liquid density. The measured value of ω_s is 0.004; Eq. (15) then gives $Y_n \approx 125$. Further improvement would require reduction of both τ_c and ω_s. Experiments indicate that τ_c is shorter at higher temperatures, and theory predicts a minimum at about 1700°F (1200 K). Application of intense electromagnetic fields to enhance

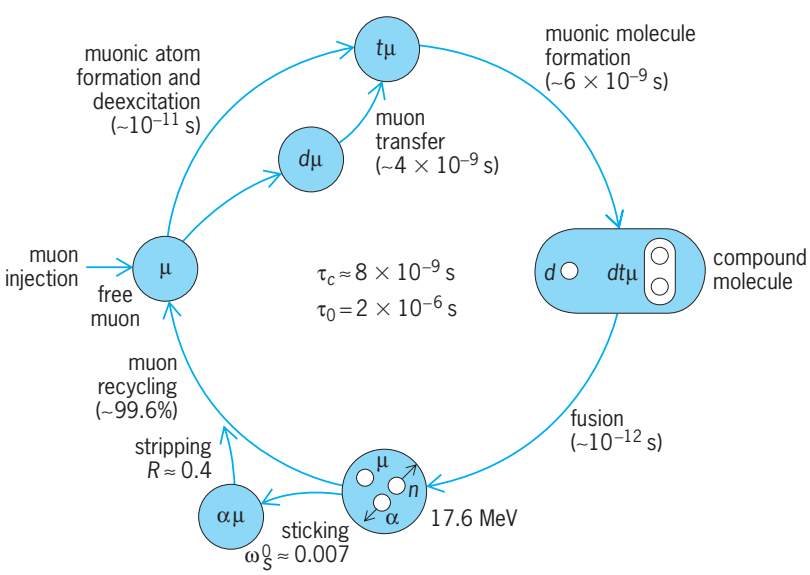

Fig. 14. Schematic diagram of the muon-catalyzed *d-t* fusion cycle. The approximate times indicated are for liquid-hydrogen density and tritium fraction 0.4. The muon, with lifetime 2×10^{-6} s, can decay at any point in the cycle. *R* is the fraction of initially sticking muons that are reactivated by stripping.

stripping of the stuck muon has been proposed to reduce ω_s (the indicated value already reflects some collisional stripping that occurs naturally).

Energy balances. The energy generated by the number of fusions achieved thus far is not yet sufficient to recover the energy expended to make the muon. Each fusion yields 17.6 MeV of energy so that the total energy release per muon is ~2200 MeV. This is considerably greater than 106 MeV, the rest mass energy of a muon, but is still short of the ~8000 MeV required to make the muon by using present accelerator technology. Energy production would require either using less energy to make muons or more fusions per muon. James S. Cohen

Bibliography. A. H. Boozer, Physics of magnetically confined plasmas, *Rev. Mod. Phys.*, 76:1071–1141, 2004; K. Brueckner et al., *Inertial Confinement Fusion*, 1992; R. G. Craxton, R. L. McCrory, and J. M. Soures, Progress in laser fusion, *Sci. Amer.*, 255(2): 68–79, August 1986; J. I. Dunderstadt and G. A. Moses, *Inertial Confinement Fusion*, 1982; Fusion special issue, *Euro-physics News*, vol. 29, no. 6, November/December 1998; S. E. Jones, J. Rafelski, and H. J. Mankorst (eds.), *Muon-Catalyzed Fusion*, 1989; B. Maglich and J. Norwood (eds.), Proceedings of the 1st International Symposium on Feasibility of Aneutronic Power, *Nucl. Instrum. Meth.*, A271:1–240, 1988; B. Maglich and J. Norwood (eds.), *Proceedings of the 2d International Symposium on Aneutronic Power, Fusion Technology*, 1990; K. Niu, *Nuclear Fusion*, 1989; J. Rafelski and S. E. Jones, Cold nuclear fusion, *Sci. Amer.*, 257:84–89, 1987; E. Storm et al., *High Gain Inertial Confinement Fusion: Recent Progress and Future Prospects*, Lawrence Livermore Nat. Lab. Rep. UCRL-99383, 1989; J. Wesson, *Tokamaks*, 3d ed., Clarendon Press, Oxford, 2004.

Nuclear hormone receptors

Specialized proteins in the nucleus that are involved in the regulation of gene activity in response to hormones and other signals. It has been estimated that approximately 30,000 genes are expressed in the human genome. The transcription of these genes into messenger ribonucleic acid (mRNA) and their subsequent translation into protein determines the identity and function of each cell in the human body. The balance between sickness and health requires that transcription of these genes be precisely regulated in response to both physiological and environmental changes. A variety of signaling strategies have evolved to coordinate the body's transcriptional response to such events. The nuclear hormone receptors represent one of the largest family of proteins that directly regulate transcription in response to hormones and other ligands (chemical entities that bind to and activate the receptor).

Nuclear hormone receptors have been identified in many species ranging from *Caenorhabditis elegans* (worms) to humans. Nearly 50 distinct nuclear receptor genes have been identified in the human genome. These include receptors for the steroid hormones (glucocorticoids, mineralocorticoids, estrogens, progestins, and androgens), thyroid hormones, vitamin D, and retinoic acid. An even larger number of nuclear receptor proteins have been found for which no known hormone or ligand has yet been identified. These proteins have been termed orphan receptors. Orphan receptors hold considerable promise, as they provide the first clues toward the identification of novel regulatory molecules and new drug therapies. *See* GENE; HORMONE.

Molecular view. Although each of the nuclear receptors mediates distinct biological effects, their activities at the molecular level are remarkably similar. Nuclear receptors contain two functional modules that characterize this family of proteins: the deoxyribonucleic acid (DNA)–binding domain and the hormone-binding or ligand-binding domain (**Fig. 1**).

DNA-binding domain. In order to regulate gene transcription, nuclear receptors must first bind specific DNA sequences adjacent to their target genes. These sequences are referred to as hormone response elements (HREs). The DNA-binding domain is the portion of the receptor that binds to these sequences. As each receptor controls only a specific subset of all genes, the DNA-binding domain must discriminate between its hormone response elements and other closely related sequences. Structural analyses demonstrate that the DNA-binding domain of nuclear receptors folds into a three-dimensional structure containing three α-helical segments. The first of these helices makes direct contact with nucleotides in the major groove of DNA, while the third helix contacts phosphate groups in the minor groove. These specific interactions help determine the precise DNA sequence that each receptor recognizes. Some receptors bind to DNA by themselves (as monomers); however, the majority bind to DNA in association with

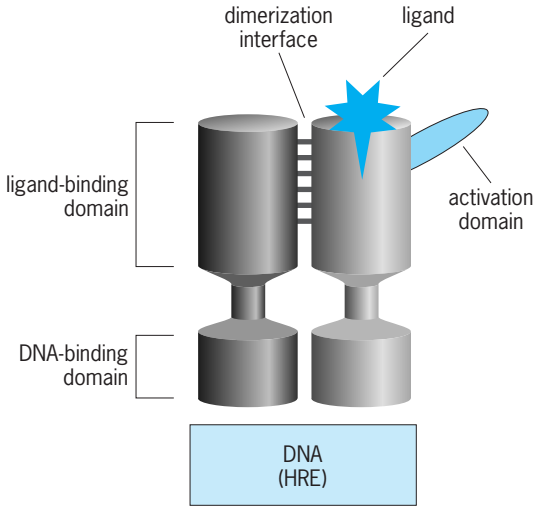

Fig. 1. Schematic illustration of a nuclear receptor dimer bound to DNA. The DNA-binding domains are shown contacting the hormone response element (HRE). The receptor at the right is shown bound to its ligand. In addition to binding ligand, the ligand-binding domain contains subdomains for transcriptional activation and dimerization.

a partner receptor (Fig. 1). The formation of such dimers adds another level of selectivity to the DNA-receptor interaction, as each receptor subunit may recognize different DNA sequences. Moreover, the relative orientation of the two DNA-binding domains may vary from one dimer to another. This forces each dimer pair to recognize hormone response elements containing sequence gaps of different sizes in the center of the element. The combination of these features allows individual nuclear receptors to regulate a specific set of genes. *See* DEOXYRIBONUCLEIC ACID (DNA); NUCLEIC ACID.

Ligand-binding domain. While the DNA-binding domain selects the target genes, it is the hormone- or ligand-binding domain that determines which hormones activate these genes. The ligand-binding domain contains a cavity or pocket that specifically recognizes the cognate ligand. When a hormonal ligand is bound within this pocket, the receptor undergoes a conformation change that reorients the position of the activation domain (Fig. 1). This reorientation switches the receptor into a transcriptionally active state. Thus, the ligand-binding domain senses the presence of specific ligands in the body, and in the presence of that ligand it allows the receptor to activate gene transcription. In addition to serving as the hormonal sensor, the ligand-binding domain contains a dimerization domain that determines which receptors can interact with each other.

Regulation of target genes. In the nucleus, DNA is wrapped around histone proteins in a structure known as chromatin. DNA that is in a more compact chromatin structure is less transcriptionally active. Indeed, one way in which nuclear receptors activate transcription is to modulate chromatin structure. In the absence of ligand, some receptors (such as the thyroid hormone and retinoic acid receptors) associate with a corepressor complex. This association is facilitated by a groove on the surface of the receptor that binds to corepressor (or coactivator) proteins. Formation of receptor-corepressor complexes is critical, as the corepressor complex exhibits an enzymatic activity that removes acetyl groups from histones (**Fig. 2**, top). The recruitment of corepressors

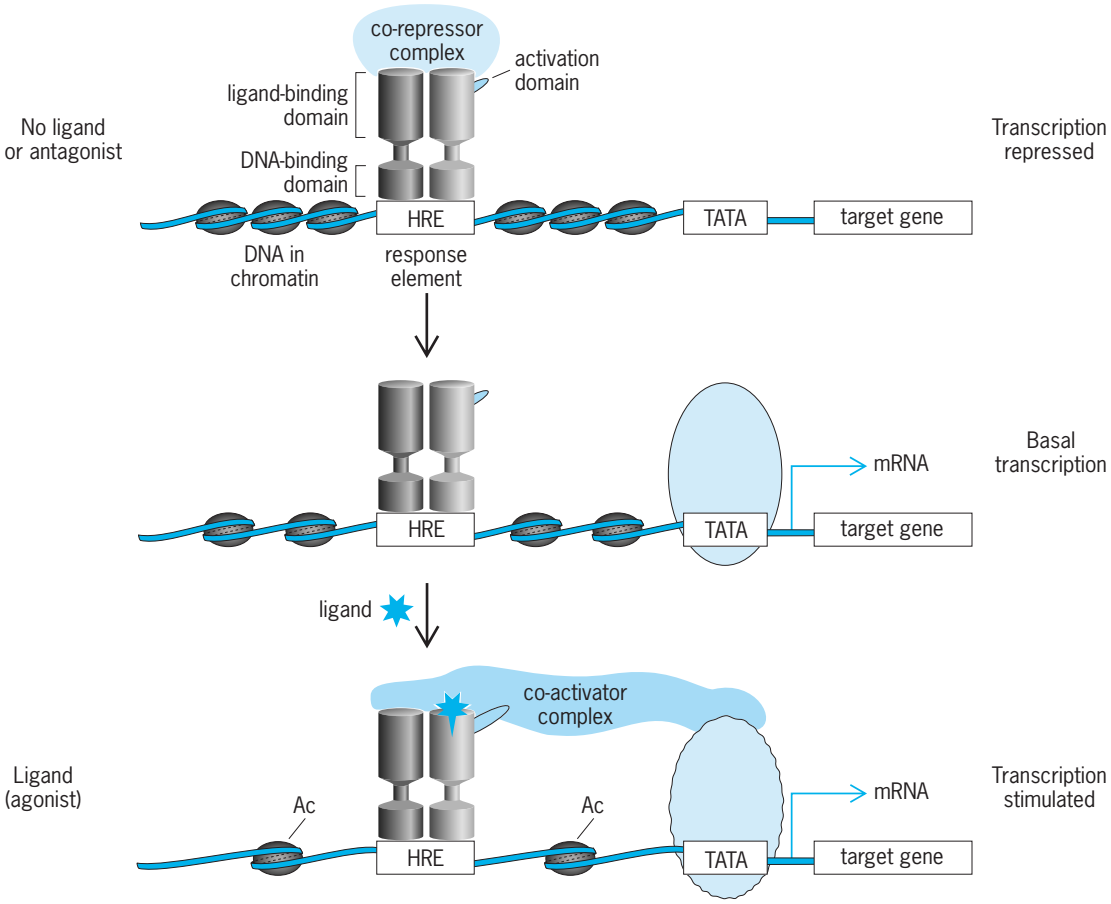

Fig. 2. Schematic illustration of the mechanism of transcriptional regulation by nuclear receptors. Nuclear receptor dimer binds to its HRE in the absence of ligand [or in the presence of an antagonist] (top). Receptors in this state interact with corepressor complexes that deacetylate histones. Hence, chromatin is compact and transcription is repressed. In the presence of ligand, corepressor is released and transcription can proceed at low or basal levels (middle). Note that many receptors do not recruit corepressors; hence in the absence of ligand their target genes are transcribed at basal levels (middle). When ligand is present, coactivators are recruited to the complex (bottom); chromatin is then acetylated (Ac) and/or methylated (Me), becomes less compact, and the rate of transcription is stimulated (bottom). In the case of the estrogen receptor, target genes are in the basal state in the absence of ligand. Estrogen agonists promote coactivator recruitment and transcriptional activation (bottom), whereas antagonists promote corepressor association and transcriptional repression (top).

allows the receptor to specifically deacetylate the surrounding chromatin, and, since hypoacetylated chromatin is transcriptionally repressed, thereby repress transcription of its target genes.

Ligand binding to the receptor results in the reorientation of the activation domain so that it now sits in the coregulator groove. This displaces the corepressor from the receptor and results in a loss of transcriptional repression (Fig. 2, middle). Genes that are in this state are neither repressed nor activated; they are actively transcribing mRNA at low or basal levels. This basal rate of transcription is determined by RNA polymerase II, which functions as part of a complex of proteins that bind the TATA sequence [in eukaryotes, a short sequence of base pairs that is rich in adenine (A) and thymidine (T) residues and is located 25–30 nucleotides upstream of the transcriptional initiation site] within the promoters of these genes. Thus, one role of ligand is to relieve repression, thereby raising activity to basal levels. Note that many nuclear receptors are unable to recruit corepressors; genes regulated by these receptors are already in the basal state in the absence of ligand (Fig. 2, middle).

In addition to relieving repression, the hormonal ligands activate transcription above and beyond the basal state; the same ligand-induced conformation change that displaces the corepressor also allows the receptor to associate with coactivator complexes (Fig. 2, bottom). One class of coactivator complex functions in a manner opposite to that of corepressors; specifically, it acetylates (Ac) and/or methylates (Me) histones leading to a more transcriptionally active chromatin structure. A second type of coactivator complex directly links the receptor to the RNA polymerase II complex. Taken together, the recruitment of these protein assemblies allows nuclear receptor ligands to simultaneously modify chromatin and increase RNA polymerase II–dependent transcription. *See* CELL NUCLEUS; GENE ACTION; NUCLEOPROTEIN; NUCLEOSOME.

Estrogen receptors, breast cancer, and osteoporosis. The estrogen receptor provides an interesting example of how receptors can be exploited to develop selective drug therapies.

Following menopause, estrogen production declines, causing changes in gene transcription that increase the risk of developing osteoporosis (bone loss), atherosclerosis (heart attacks), and hot flashes. These adverse effects can be delayed by the administration of ligands that activate the estrogen receptor (agonists). However, estrogens also promote the proliferation of breast and uterine cancer cells. Thus, for patients at risk for these cancers, it may be more important to inhibit estrogen action. Indeed, synthetic drugs have been identified that bind to estrogen receptors but fail to activate the receptor. These drugs are known as antagonists because they function in a manner opposite to that of the natural hormone, that is, they displace coactivators and/or recruit corepressors. Such estrogen receptor antagonists are currently used for both the treatment and prevention of breast cancer.

If a given estrogen receptor ligand either activates or antagonizes the receptor, the physician must carefully evaluate whether an individual patient is more likely to benefit by having their estrogen receptors "turned-on" or "turned-off." Fortunately, nuclear receptors are more flexible and can be modulated in ways that are more complex than this simple on/off scenario. The reason for this flexibility is that the ligand does not control receptor activity by itself; transcriptional activity is determined by the type of coregulator protein that associates with the ligand-bound receptor (Fig. 2). An important advance in understanding nuclear receptors has been the realization that these receptors interact with a variety of coactivators and corepressor proteins. Thus, a more useful drug for the estrogen receptor may be one that preferentially recruits a coactivator in bone cells, has no effect in uterine cells, and recruits a corepressor in breast cancer cells. Indeed, several drugs have been identified that can inhibit estrogen action in the breast, prevent osteoporosis by activating estrogen in bone (tamoxifen, raloxifene), and fail to activate transcription and promote cancer in the uterus (raloxifene). These types of ligands are known as selective estrogen receptor modulators (SERMs), as they can activate certain genes while inhibiting others.

Another level of flexibility in the estrogen response has emerged with the identification of a second receptor. Estrogen receptor β binds to similar DNA sequences as estrogen receptor α, but it is found in different tissues and can bind different synthetic ligands. Additional variations in response are possible, as these two receptors may form α/β heterodimers. Preliminary results suggest that the positive cardiovascular effects of estrogens result from activation of estrogen receptor β, not estrogen receptor α. *See* BREAST DISORDERS; CANCER (MEDICINE); ESTROGEN; HEART DISORDERS; MENOPAUSE; OSTEOPOROSIS.

PPARγ and diabetes. The identification of the orphan receptor peroxisome proliferator-activated receptor gamma (PPARγ) as a factor required for fat cell formation led to the exciting suggestion that a yet-to-be discovered ligand may control this process. Indeed, a fatty-acid derivative known as 15-deoxy-$\Delta^{12,14}$-prostaglandin J_2 promotes fat cell formation (adipogenesis) by binding to and activating PPARγ. Activation of PPARγ in diabetic patients results in a shift of lipids from muscle to fat, thus allowing the muscle to function properly. It is now clear that PPARγ promotes adipogenesis by acting as a master regulator of a network of fatty acid storage genes. This network includes genes involved in fatty acid release from serum lipoproteins, uptake of fatty acids into the fat cell, synthesis and uptake of glycerol, and ultimately conversion of fatty acids and glycerol into triglycerides—the major lipid stored in the fat cell. These findings explain how synthetic PPARγ agonists (thiazolidinediones) function as antidiabetic agents in people with type II diabetes. *See* ADIPOSE TISSUE; DIABETES; LIPID METABOLISM.

Orphan receptors for the future. Recent data have identified two orphan receptors (PXR and CAR) that

define a new paradigm central to "chemical immunity" (protection from toxic compounds). These receptors regulate genes involved in toxin clearance, but they do so in response to a highly diverse array of endogenous and exogenous toxins. Other orphan receptors have been identified that play critical roles in cholesterol homeostasis, Parkinson's disease, vision, neuronal development, and other processes. The biological function of many other orphan receptors remains to be elucidated. Thus, these proteins will continue to provide tools for the identification of novel transcriptional regulatory pathways and important drug therapies. *See* TOXIN. Barry Marc Forman

Bibliography. M. W. Draper, The role of selective estrogen receptor modulators (SERMs) in postmenopausal health, *Ann. NY Acad. Sci.*, 997:373–377, 2003; V. Laudet and H. Gronemeyer, *The Nuclear Receptor Factsbook*, Academic Press, 2001; J. M. Olefsky, Nuclear receptor minireview series, *J. Biol. Chem.*, 276(40):36863–36864, 2001.

Nuclear isomerism

The existence of excited states of atomic nuclei with unusually long lifetimes. A nucleus may exist in an excited quantum state with well-defined excitation energy (E_x), spin (J), and parity (π). Such states are unstable and decay, usually by the emission of electromagnetic radiation (gamma rays), to lower excited states or to the ground state of the nucleus. The rate at which this decay takes place is characterized by a half-life ($\tau_{1/2}$), the time in which one-half of a large number of nuclei, each in the same excited state, will have decayed. If the lifetime of a specific excited state is unusually long, compared with the lifetimes of other excited states in the same nucleus, the state is said to be isomeric. The definition of the boundary between isomeric and normal decays is arbitrary; thus the term is used loosely. *See* EXCITED STATE; PARITY (QUANTUM MECHANICS); SPIN (QUANTUM MECHANICS).

Spin isomerism. The predominant decay mode of excited nuclear states is by gamma-ray emission. The rate at which this process occurs is determined largely by the spins, parities, and excitation energies of the decaying state and of those to which it is decaying. In particular, the rate is extremely sensitive to the difference in the spins of initial and final states and to the difference in excitation energies. Both extremely large spin differences and extremely small energy differences can result in a slowing of the gamma-ray emission by many orders of magnitude, resulting in some excited states having unusually long lifetimes and therefore being termed isomeric.

Occurrence. Isomeric states have been observed to occur in almost all known nuclei. However, they occur predominantly in regions of nuclei with neutron numbers N and proton numbers Z close to the so-called magic numbers at which shell closures occur. This observation is taken as important evidence for the correctness of the nuclear shell model which predicts that high-spin, and therefore isomeric, states should occur at quite low excitation energies in such nuclei. *See* MAGIC NUMBERS.

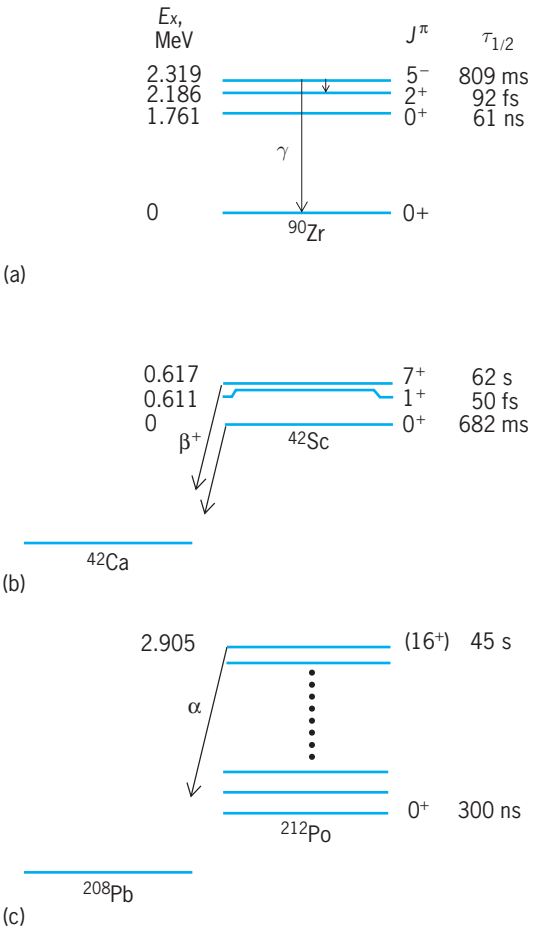

Examples of nuclear isomerism in (*a*) ^{90}Zr, (*b*) ^{42}Sc, and (*c*) ^{212}Po. (*After C. M. Lederer and V. S. Shirley, Table of Isotopes, 7th ed., John Wiley and Sons, 1978*)

Examples. Three examples of isomeric states are shown in the **illustration**. In the case of ^{90}Zr (illus. *a*), the 2. 319-MeV, $J^\pi = 5^-$ state has a half-life of 809 milliseconds, compared to the much shorter lifetimes of 93 femtoseconds, and 61 nanoseconds for the 2.186-MeV and 1.761-MeV states, respectively. The spin difference of 5 between the 2.319-MeV state and the ground and 1.761 $J^\pi = 0^+$ states, together with the spin difference of 3 and small energy difference between the 2.319-MeV state and the 2.186-MeV, $J^\pi = 2^+$ state, produce this long lifetime.

For ^{42}Sc (illus. *b*), the gamma decay of the $J^\pi = 7^+$ state at 0.617 MeV is so retarded by the large spin changes involved that the intrinsically much slower β^+ decay process can take place, resulting in a half-life of 62 s.

Yet another process takes place in the case of the high-spin isomer in ^{212}Po (illus. *c*), which decays by alpha-particle emission with a half-life of 45 s rather than gamma decays.

The common feature of all these examples is the slowing of the gamma-ray emission process due to the high spin of the isomeric state.

Other mechanisms. Not all isomers are the spin isomers described above. Two other types of isomers

have been identified. The first of these arises from the fact that some excited nuclear states represent a drastic change in shape of the nucleus from the shape of the ground state. In many cases this extremely deformed shape displays unusual stability, and states with this shape are therefore isomeric. A particularly important class of these shape isomers is observed in the decay of heavy nuclei by fission, and the study of such fission isomers has been the subject of intensive effort. The possibility that nuclei may undergo sudden changes of shape at high rotational velocities has spurred searches for isomers with extremely high spin which may also be termed shape isomers. *See* NUCLEAR FISSION.

A more esoteric form of isomer has also been observed, the so-called pairing isomer, which results from differences in the microscopic motions of the constituent nucleons in the nucleus. A state of this type has a quite different character from the ground state of the nucleus, and is therefore also termed isomeric. *See* NUCLEAR STRUCTURE. Russell Betts

Bibliography. H. A. Enge and R. P. Redwine, *Introduction to Nuclear Physics*, 2d ed., 1995; K. S. Krane, *Introductory Nuclear Physics*, 1987; C. M. Lederer and V. S. Shirley, *Table of Isotopes*, 7th ed., 1978; K. Siegbahn (ed.), *Alpha, Beta and Gamma-Ray Spectroscopy*, vols. 1 and 2, 1965.

Nuclear magnetic resonance (NMR)

A phenomenon exhibited when atomic nuclei in a static magnetic field absorb energy from a radio-frequency field of certain characteristic frequencies. Nuclear magnetic resonance is a powerful analytical tool for the characterization of molecular structure, quantitative analysis, and the examination of dynamic processes. It is based on quantized spectral transitions between nuclear Zeeman levels of stable isotopes, and is unrelated to radioactivity. *See* ZEEMAN EFFECT.

The format of nuclear magnetic resonance data is a spectrum that contains peaks referred to as resonances. The resonance of an isotope is distinguished by the transition frequency of the nucleus. The intensity of the resonance is directly proportional to the number of nuclei that produce the signal. Although the majority of nuclear magnetic resonance spectra are measured for samples in solution or as neat liquids, it is possible to measure nuclear magnetic resonance spectra of solid samples. Nuclear magnetic resonance is a nondestructive technique that can be used to measure spectra of cells and living organisms.

Nuclear magnetic properties. In order to understand the basic phenomena that give rise to the nuclear magnetic resonance spectrum, it is necessary to examine the magnetic resonance of nuclei. The nuclei of most atoms possess an intrinsic nuclear angular momentum. The classical picture of nuclear angular momentum is a spherical nucleus rotating about an axis in a manner analogous to the daily rotation of the Earth. Nuclear angular momentum, like most other atomic quantities, can be expressed as a series of quantized levels.

The transitions that give rise to the nuclear magnetic resonance spectrum are produced when the nuclei are placed in a static magnetic field. The external or applied magnetic field defines a geometric axis, denoted as the z axis. The external magnetic field orients the z components of nuclear angular momentum and the magnetic moment with respect to the z axis. The nuclear magnetic resonance spectrum is produced by spectral transitions between different spin states and is therefore dependent on the value of the nuclear spin. Nuclei for which the nuclear spin value is 0 have a magnetic spin-state value of 0. Therefore, these nuclei do not give rise to a nuclear magnetic resonance spectrum under any circumstances. Many important elements in chemistry have zero values of nuclear spin and are inherently nuclear magnetic resonance inactive, including the most abundant isotopes of carbon-12, oxygen-16, and sulfur-32. Nuclear magnetic resonance spectra are measured most often for nuclei that have nuclear spin values of $^1/_2$. Chemically important spin-$^1/_2$ nuclei include the isotopes hydrogen-1, phosphorus-31, flourine-19, carbon-13, nitrogen-15, silicon-29, iron-57, selenium-77, cadmium-113, silver-107, platinum-195, and mercury-199. Nuclear magnetic resonance spectra of nuclei that have nuclear spin values of greater than $^1/_2$, known as quadrupolar nuclei, can be measured; but the measurements are complicated by faster rates of nuclear relaxation. Nuclei that fall into this class include the isotopes hydrogen-2, lithium-6, nitrogen-14, oxygen-17, boron-11, chlorine-35, sodium-23, aluminum-27, and sulfur-33.

The transition frequency in nuclear magnetic resonance depends on the energy difference of the spin states. The intensity of the resonance is dependent on the population difference of the two spin states, which in turn depends directly on the magnitude of the energy difference. For spin-$^1/_2$ nuclei, ΔE is the difference in energy between the $+^1/_2$ and $-^1/_2$ values of the magnetic spin state. In contrast to other spectroscopic methods, the nuclear magnetic resonance frequency is variable, depending on the strength of the magnet used to measure the spectrum. Larger magnetic fields produce a greater difference in energy between the spin states. This translates into a larger difference in the population of the spin states and therefore a more intense resonance in the resulting nuclear magnetic resonance spectrum.

Pulsed-Fourier-transform NMR spectroscopy. Early nuclear magnetic resonance spectrometers were continuous-wave instruments that scanned through the magnetic field until the resonance frequency of each group of nuclei was reached. Continuous-wave spectrometers have virtually disappeared and have been replaced by pulsed-Fourier-transform instruments. In the pulsed-Fourier-transform experiment, a secondary oscillating magnetic field perpendicular to the applied magnetic field is produced by the application of a pulse of radio-frequency radiation.

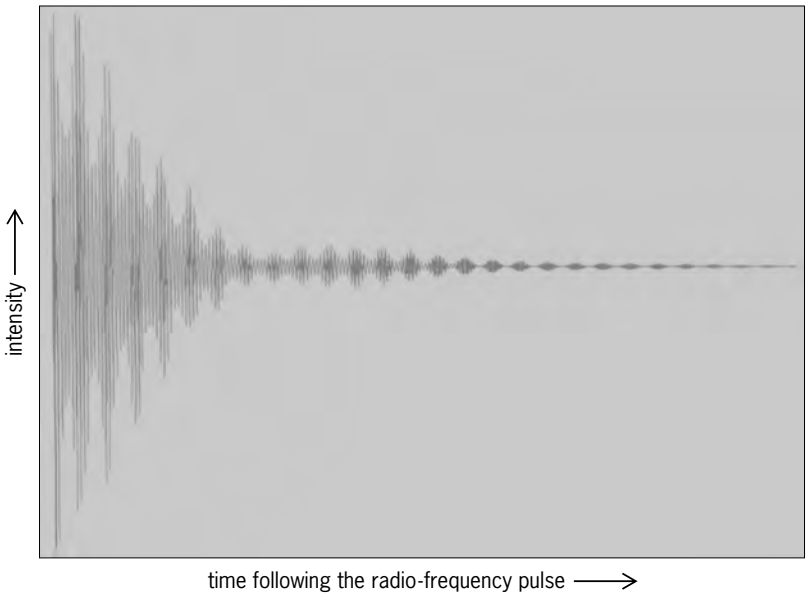

Fig. 1. Proton free induction decay measured for a solution of ethyl acetate dissolved in deuterated chloroform. The free induction decay is the detected signal which contains all of the information of the nuclear magnetic resonance spectrum, including the resonance frequency, intensity, and linewidth of each chemically distinct proton in the sample.

Absorption of the radio-frequency energy involves flipping the magnetic moment, effectively inverting the population of the spin states. The nuclear magnetic resonance signal generated as result of the radio-frequency pulse is a transient response called the free induction decay (**Fig. 1**). The response is transient because the nuclei return to their equilibrium populations usually within a few seconds by nonradiative processes. The intensity of the free induction decay is a maximum immediately following the radio-frequency pulse and decays logarithmically as a function of time. The nuclear magnetic resonance spectrum (**Fig. 2**) is produced by Fourier transformation of the free induction decay. *See* FOURIER SERIES AND TRANSFORMS; INTEGRAL TRANSFORM.

Chemical shift. Individual nuclei of the same isotope in a molecule have transition frequencies that differ depending on their chemical environment. This phenomenon, called chemical shift, occurs because the effective magnetic field at a particular nucleus in a molecule is less than the applied magnetic field due to shielding by electrons.

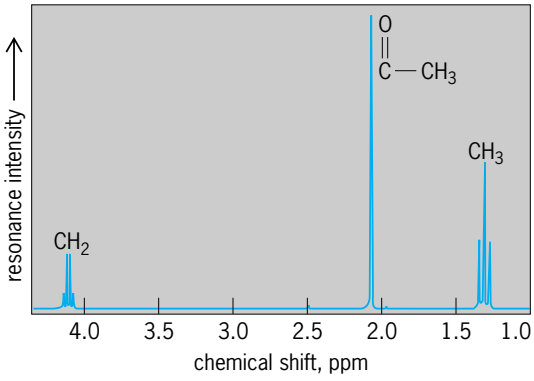

Fig. 2. Proton spectrum of ethyl acetate ($CH_3COOH_2CH_3$) obtained by Fourier transformation of the free induction decay shown in Fig. 1. Chemical shift is relative to the protons of tetramethylsilane. The integrated intensity of the resonances corresponds to the relative number of protons which give rise to each resonance.

An example of resonance chemical shift is observed in the ^1H nuclear magnetic resonance spectrum of ethyl acetate ($CH_3COOCH_2CH_3$; Fig. 2). Resonances at several frequencies are observed in this spectrum. The methylene (CH_2) protons are affected by the electron-withdrawing oxygen atoms of the neighboring ester group, and as a result the chemical shift of the methylene proton resonance is significantly different from the chemical shift of the resonances of the protons of the methyl (CH_3) groups of ethyl acetate. The two methyl groups are in different chemical environments and therefore give rise to resonances that have different chemical shifts. Because of the dependence of the transition frequency of a nucleus on its chemical environment, chemical shift is diagnostic of the functional group containing the nucleus of interest. Nuclear magnetic resonance spectroscopy is a frequently employed tool in chemical synthesis studies, because the nuclear magnetic resonance spectrum can confirm the chemical structure of a synthetic product.

In most nuclear magnetic resonance spectra, rather than frequency units, the spectra are plotted in units of chemical shift expressed as part per million (ppm). The ppm scale calculates the ratio of the resonance frequency in hertz (Hz) to the Larmor frequency of the nucleus at the magnetic field strength of the measurement. The ppm scale allows direct comparison of nuclear magnetic resonance spectra acquired by using magnets of differing field strength. This unit of nuclear magnetic resonance chemical shift should not be confused with the concentration units of ppm (mg/kg), often referred to by analytical chemists in the field of trace analysis.

The chemical shift (either in hertz or ppm) of a resonance is assigned relative to the chemical shift of a standard reference material. The nuclear magnetic resonance community has agreed to arbitrarily set the chemical shift of certain standard compounds to 0 ppm. For ^1H and ^{13}C nuclear magnetic resonance, the accepted standard is tetramethylsilane, which is defined to have a chemical shift of 0 ppm. However, any molecule with a resonance frequency in the appropriate chemical shift region of the spectrum that does not overlap with the resonances of the sample can be employed as a chemical shift reference. The use of a chemical shift reference compound other than the accepted standard is particularly common for nuclei that have a very large chemical shift range such as fluorine-19 or selenium-77, since it may be impossible to measure the spectrum of the accepted chemical shift reference compound and the analyte of interest simultaneously because of their very different frequencies.

Spin-spin coupling. In addition to the differences in chemical shift, resonances in a given spectrum, for example for ethyl acetate (Fig. 2), also differ in the number of signals composing the resonance

detected for each group of protons. Instead of one resonance resulting from a single transition between two spin states, the methylene group (CH_2) resonance is actually a quartet composed of four lines. Similarly, although the acetate CH_3 resonance is a single resonance, the ethylene group (CH_3) resonance consists of three lines. This splitting of some resonances, called spin-spin or scalar coupling, arises from interactions between nuclei through their bonding electrons rather than through space. Scalar coupling is a short-range interaction that usually occurs for nuclei separated by one to three bonds.

Spin decoupling. The effects of spin-spin coupling can be removed from the spectrum by application of a low-strength magnetic field to one of the coupled spins. The effect of this secondary field is to equalize the population of the coupled transitions and to remove the effects of spin-spin coupling. Decoupling can be homonuclear or heteronuclear. Homonuclear decoupling is useful for assigning the resonances of coupled spin systems in the spectrum of a compound or complex mixture. For example, in ethyl acetate, decoupling at the frequency of the methylene group (CH_2) protons would remove the effects of spin-spin coupling, collapsing the ethylene group (CH_3) resonance into a singlet and confirming the resonance assignments.

The previous discussion of chemical shift and spin-spin coupling has focused on ^1H nuclear magnetic resonance spectroscopy. When detecting nuclei other than protons (^1H), for example carbon-13 (^{13}C), the effects of heteronuclear spin-spin coupling between the protons and the carbon-13 can be removed by decoupling the protons during acquisition of the carbon free induction decay. This is called heteronuclear decoupling, and it is commonly used to simplify nuclear magnetic resonance spectra. The majority (98.9%) of the protons are bound to ^{12}C nuclei, which have zero spin. However, 1.1% of the protons are bound to ^{13}C nuclei. The resonances of protons bound to ^{13}C nuclei are split into a doublet with a coupling constant, J_{CH}, of 120–160 Hz. These weak proton-coupled ^{13}C resonances can be detected in proton nuclear magnetic resonance spectra when the signal-to-noise ratio is sufficiently high. In the ^{13}C nuclear magnetic resonance spectrum, the ^{13}C resonances are split into multiplets reflecting the number of directly bound protons.

In the proton-coupled ^{13}C spectrum of ethyl acetate (**Fig. 3*a***), the quartet resonances of the two methyl carbons occur at 13.9 and 20.3 ppm. The CH_2 carbon resonance occurring at 59.9 ppm is split into a triplet by the two attached protons. This spectrum indicates that the carboxyl carbon occurs at 170.2 ppm; it is a singlet since it isnot directly

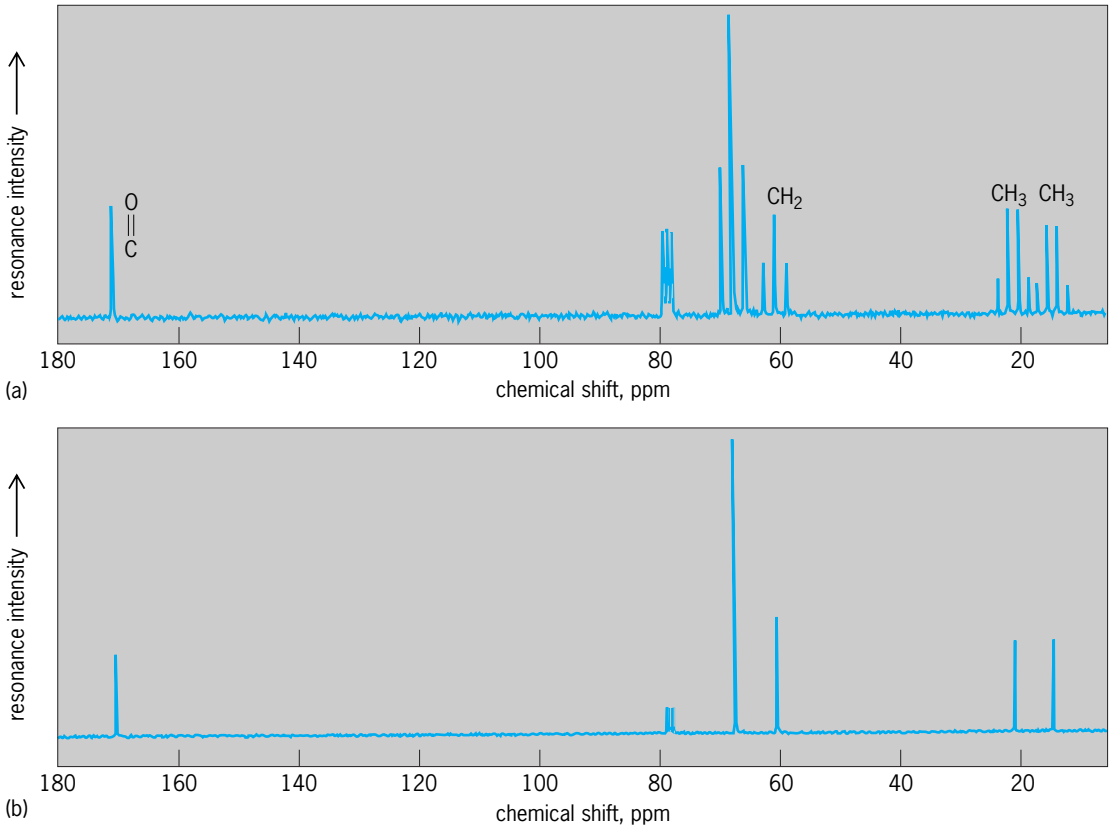

Fig. 3. Spin decoupling. (*a*) The proton-coupled natural abundance carbon-13 (^{13}C) spectrum of ethyl acetate. The splitting of the ^{13}C resonances arises from spin-spin coupling to the protons bound to their respective carbon atoms. (*b*) The ^{13}C spectrum of ethyl acetate measured with proton decoupling. With the effects of heteronuclear spin-spin coupling removed, the ^{13}C spectrum is greatly simplified and the signal-to-noise ratio is increased relative to the proton coupled spectrum. Spectra of *a* and *b* are plotted at different vertical scales.

bound to any protons. Notably, no ^{13}C-^{13}C scalar coupling is observed in such a spectrum because of the low probability that two ^{13}C nuclei will be directly bonded at natural abundance levels.

In the proton-decoupled ^{13}C nuclear magnetic resonance spectrum of ethyl acetate (Fig. 3*b*), the heteronuclear spin-spin coupling between the protons and the carbon nuclei can be eliminated by decoupling over the entire proton chemical shift region during acquisition of the free induction decay. This type of decoupling is known as broadband decoupling, because a wide band of proton frequencies are decoupled. The only ^{13}C resonances that are unaffected by proton decoupling are those not directly bonded to protons. The proton-decoupled ^{13}C spectrum (Fig. 3*b*) is much simpler than the proton-coupled spectrum (Fig. 3*a*) since it consists of all singlets. In addition to simplifying the spectrum, proton decoupling produces a higher signal-to-noise spectrum, because all the ^{13}C resonance intensity is concentrated into a single resonance rather than spread over a multiplet.

Solid-state NMR spectroscopy. Nuclear magnetic resonance spectra of solid-phase materials have the disadvantage that the resonances are very broad, a feature observed in the ^{13}C spectrum of solid glycine (**Fig. 4***a*). Since this compound is structurally very simple, the resonances of the individual carbon atoms can be distinguished in its spectrum, even though the individual resonances are quite broad. In the liquid phase, dipolar coupling between nuclei on different molecules is not observed because of the rapid reorientation of the molecules. In the solid phase, where molecular reorientation does not readily occur, direct dipolar magnetic interactions between the nuclei of adjacent molecules are relatively large and contribute significantly to the linewidth of resonances in the nuclear magnetic resonance spectra of solids. The removal of the effects of dipolar coupling by proton decoupling results in much narrower resonances (Fig. 4*b*).

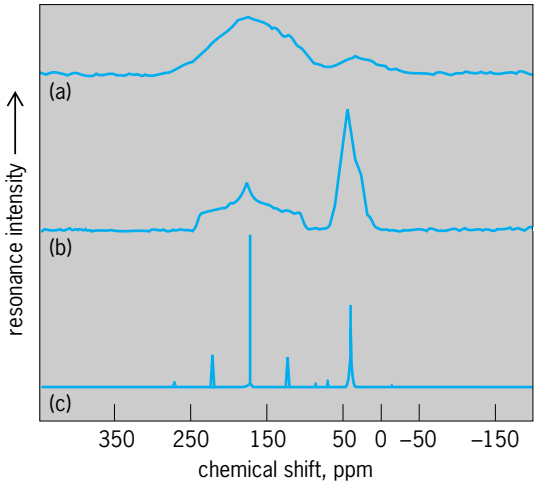

Fig. 4. Solid-state nuclear magnetic resonance spectroscopy. Narrowing the resonances in nuclear magnetic resonance is accompanied by a large increase in the resonance intensity; hence the sensitivity of the measurement. Vertical axes are plotted at different scales. (*a*) The static carbon-13 (^{13}C) spectrum of solid glycine ($+NH_3$—CH_2—COO^-). (*b*) The ^{13}C spectrum of solid glycine measured with proton decoupling to remove the effects of dipolar coupling between the proton and carbon nuclei of adjacent molecules. (*c*) The ^{13}C spectrum of solid glycine measured by using both proton decoupling and magic angle spinning. (*Martine Ziliox, Bruker Instruments, Inc.*)

An additional contribution to the broad resonances obtained for solid-phase samples comes from chemical shift anisotropy, an orientation dependence of the chemical shift. Chemical shift anisotropy is not usually a problem in nuclear magnetic resonance spectra of liquid-phase samples, because its effects are averaged by the rapid rotation of the molecules. In solids, both the dipolar coupling and the chemical shift anisotropy have an angular dependence. To minimize the effects of dipolar interactions and chemical shift anisotropy, solid-state samples are spun rapidly (that is, 3 kHz) at the angle $55°44'$, known as the magic angle. Dramatic improvement is obtained by using both proton decoupling and magic angle spinning to narrow the ^{13}C resonances (Fig. 4*c*). Solid-state nuclear magnetic resonance is particularly useful for measuring spectra of insoluble materials such as synthetic and natural polymers. Deuterium (hydrogen-2) nuclear magnetic resonance is widely used in the solid state to measure nuclear magnetic resonance spectra of biological membranes. Other applications of solid-state nuclear magnetic resonance have been the use of silicon-29 and aluminum-27 nuclear magnetic resonance to study zeolites and the measurement of nuclear magnetic resonance spectra of xenon-129 adsorbed on surfaces. *See* DEUTERIUM.

Quantitative analysis. Quantitative analysis using nuclear magnetic resonance makes use of the direct relationship between the number of nuclei and the resonance intensity. Complex mixtures can be analyzed quantitatively by nuclear magnetic resonance without separating the components of the mixture. Most spectroscopic techniques require that a standard of the pure analyte be used for calibration, because the response of the instrument is different for each analyte. In nuclear magnetic resonance spectroscopy, the integrated intensity of a resonance is directly proportional to the concentration of nuclei that give rise to the resonance. Therefore, when the experimental parameters (particularly the repetition time) are carefully chosen, concentrations can be determined relative to an internal standard that may be chemically unrelated to the compound of interest. This is a particular advantage when pure standards of the material being analyzed are not available. *See* QUANTITATIVE CHEMICAL ANALYSIS.

Qualitative analysis. Nuclear magnetic resonance spectroscopy is one of the best methods available for the structural analysis of molecules. Integrals indicate the relative numbers of nuclei that give rise to the resonances in the nuclear magnetic resonance spectrum. Coupled spin systems can be elucidated by spin-decoupling experiments in which individual multiplets are irradiated in succession. Although these methods have been used to successfully assign the structure of numerous molecules, there is a

fundamental limitation in the complexity of the nuclear magnetic resonance spectrum that is amenable to analysis by these simple one-dimensional techniques. *See* QUALITATIVE CHEMICAL ANALYSIS.

Two-dimensional NMR experiments. The introduction of two-dimensional nuclear magnetic resonance experiments has greatly expanded both the complexity of molecules that can be analyzed by nuclear magnetic resonance spectroscopy and the repertoire of experiments available to the spectroscopist. In general, two-dimensional nuclear magnetic resonance experiments can be analyzed as containing four separate periods: preparation, evolution, mixing, and detection. The preparation period usually contains a delay during which the magnetization is allowed to recover from the previous acquisition, followed by a pulse that initiates the experiment. In most two-dimensional experiments, the first pulse is a 90° pulse that generates transverse magnetization. The transverse magnetization is allowed to precess during the evolution period, which consists of an incremented delay period. During the evolution period, the magnetization is encoded with the information of the two-dimensional experiment. In some experiments, a mixing period consisting of a pulse, a delay period, or some combination of the two is necessary to convert the encoded information to a detectable form. The free induction decay is acquired during the detection period. This approach has been extended to three- and four-dimensional nuclear magnetic properties experiments by introduction of additional evolution and mixing periods.

Two-dimensional nuclear magnetic resonance spectra are acquired as a series of free induction decays collected at incremented values of the incremented delay period. The two-dimensional spectrum is produced by double Fourier transformation in the dimensions of both the incremented delay period and the detection period. Two-dimensional nuclear magnetic resonance spectra are defined by two nuclear magnetic resonance frequency axes and actually contain a third dimension, resonance intensity. By using the available arsenal of multidimensional nuclear magnetic resonance experiments, three-dimensional solution structures of complex molecules such as proteins have been determined that are comparable in quality to those produced by x-ray crystallography.

The utility of multidimensional nuclear magnetic resonance experiments for structural elucidation is demonstrated for a relatively simple molecule, thymidine. This discussion is not meant to be a complete description, but an illustration of the power of this methodology. Often, the purpose of the nuclear magnetic resonance study is determination of the conformation of molecular structure. However, proton nuclear magnetic resonance spectra can be relatively complicated even for a simple molecule such as thymidine (**Fig. 5**). Therefore, two-dimensional nuclear magnetic resonance experiments are employed to assign the resonances in the nuclear magnetic resonance spectrum and to confirm the structure of the molecule.

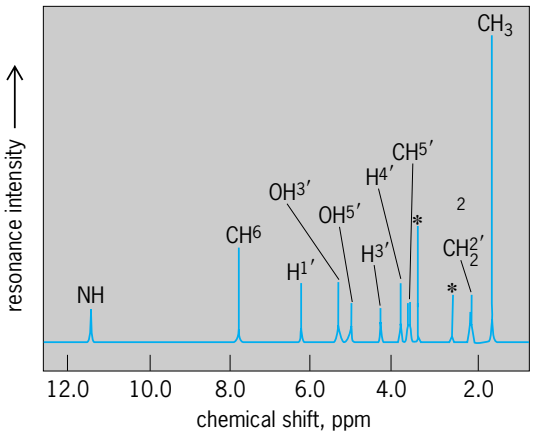

Fig. 5. Proton nuclear magnetic resonance spectrum of thymidine in deuterated dimethylsulfoxide. The thymidine protons giving rise to each resonance are indicated by the labels above each resonance. The two resonances marked * are not due to thymidine protons but result from impurities in the solution. Chemical shifts are relative to dimethylsilylpropionate, sodium salt.

The two-dimensional nuclear magnetic resonance experiment, correlated spectroscopy (COSY), produces a spectrum (**Fig. 6**) that identifies protons that are spin-spin coupled. Since spin-spin coupling is produced by through-bond interactions of the nuclei, the correlated spectroscopy spectrum identifies protons on adjacent carbon atoms. The two frequency dimensions of the COSY spectrum are proton chemical shift. The presentation format of the COSY spectrum is a contour plot, in which relative resonance intensity is distinguished by the number of contours. A contour plot is analogous to the depiction of a mountain range in a topographical map (Fig. 6).

The COSY spectrum is characterized by a diagonal that runs from the bottom left-hand corner of

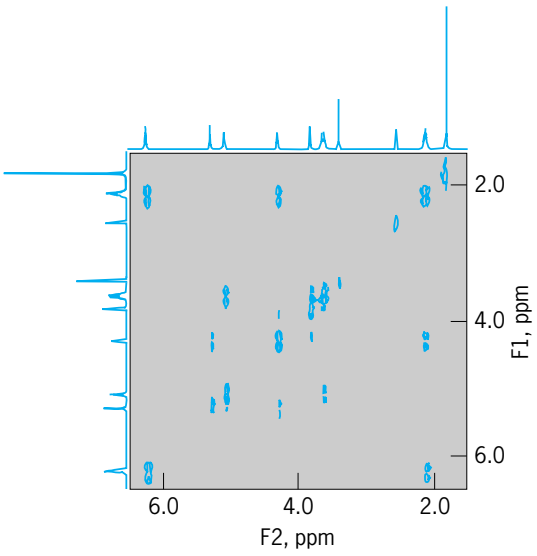

Fig. 6. Portion of the contour plot of the proton correlated spectroscopy (COSY) spectrum of thymidine. The off-diagonal elements or cross-peaks connect resonances of scalar coupled protons. The one-dimensional proton nuclear magnetic resonance spectrum of thymidine is plotted along each frequency axis (F1 and F2).

the spectrum to the upper right-hand corner; this diagonal corresponds to the one-dimensional proton spectrum. The information in a COSY spectrum is provided by the off-diagonal elements or cross-peaks. In a correlated spectroscopy experiment, during the evolution period the magnetization evolves under the effects of homonuclear spin-spin coupling. The cross-peaks in the COSY spectrum connect diagonal resonances for protons that are spin-spin coupled. If the identity of one resonance is known, the other resonances of the molecule can be assigned by tracing through the coupling patterns defined by the cross-peaks. Cynthia K. Larive

Bibliography. F. A. Bovey, *Nuclear Magnetic Resonance Spectroscopy*, 1986; H. Friebolin, *Basic One- and Two-Dimensional Nuclear Magnetic Resonance Spectroscopy*, 2d ed., 1993; G. E. Martin and A. S. Zektzer, *Two-Dimensional Nuclear Magnetic Resonance Methods for Establishing Molecular Connectivity: A Chemist's Guide to Experiment Selection, Performance and Interpretation*, 1988; J. K. M. Saunders and B. K. Hunter, *Modern Nuclear Magnetic Resonance Spectroscopy: A Guide for Chemists*, 2d ed., 1993; K. Wüthrich, The development of nuclear magnetic resonance spectroscopy as a technique for protein structure determination, *Acc. Chem. Res.*, 22:36–44, 1986.

Nuclear medicine

A subspecialty of medicine based on the use of radioactive substances in medical diagnosis, treatment, and research. Cyclotron-produced radioactive materials were introduced in the early 1930s, but the invention of the nuclear reactor during World War II made carbon-14, hydrogen-3, iodine-131, and later technetium-99m available in large quantities. Today most biomedical research and the care of many patients depend on the use of radioactive materials.

The most widely used radionuclides are technetium-99m, iodine-123, carbon-11, and fluorine-18. The latter two require a cyclotron near the site of radiotracer production because of their rapid radioactive decay (carbon-11 decays with a half-life of 20 min, fluorine-18 with a half-life of 110 min). The short half-life of the radioactive tracers makes it possible to administer the radiotracers to individuals without the harmful effects of radiation.

Tracer principle. The most fundamental principle in nuclear medicine is the tracer principle, invented in 1912 by Nobel laureate G. Hevesy, who found that radioactive elements had identical chemical properties to the nonradioactive form and therefore could be used to trace chemical behavior in solutions or in the body. One of the important consequences of the use of tracers was to establish the principle of the dynamic state of body constituents. Prior to the development of radioactive tracer methods, the only way to study biochemical processes within the body was to measure the input and output of dietary constituents and examine concentrations of molecules at autopsy. With radioactive tracers, the movement of labeled molecules could be followed from the processes within organs, to excretion. In essence, the use of radioactive tracers makes it possible to study the biochemistry within the various organs of the human body. *See* RADIOACTIVE TRACER.

According to the principle of the constancy of the internal environment, the concentration of chemical constituents in body fluids is usually kept within a very narrow range, and disturbances of these values result in disease. This concept has been one of the foundations of modern biochemistry. Nuclear medicine makes it possible to examine regional physiology and biochemistry in ways that at times surpass the perception of surgeons during an operation or pathologists during an autopsy. Imaging methods make it possible to measure regional as well as overall organ function, and to portray the results in the form of functional or biochemical pictures of the body in health and in disease. Such pictures enhance the information about structure that is obtained by other imaging methods, such as computerized tomography (CT) or magnetic resonance imaging (MRI), often providing unique, objective evidence of disease long before structural changes are seen. *See* COMPUTERIZED TOMOGRAPHY; NUCLEAR MAGNETIC RESONANCE (NMR).

The physiological and biochemical orientation of nuclear medicine provides a better approach to understanding disease. The body is viewed as a complex array of coordinated chemical and physical processes that can become impaired before signs of disease develop. This has led to the concept of chemical reserve. It is now known that chemical abnormalities, such as a reduced rate of glucose metabolism in Huntington's disease or a marked deficiency of a neurotransmitter such as dopamine in Parkinson's disease, can occur long before the onset of symptoms. This makes possible the detection of disease far earlier than symptoms or structural abnormalities can. In focal epilepsy, for example, chemical abnormalities are often detectable before structural changes occur. *See* MOLECULAR PATHOLOGY.

Diagnosis. In a typical examination, a radioactive molecule is injected into an arm vein, and its distribution at specific time periods afterward is imaged in certain organs of the body or in the entire body. The images are created by measuring the gamma-ray photons emitted from the organs or regions of interest within the body. Nuclear medicine imaging procedures differ from ordinary x-rays in that the gamma rays are emitted from the body rather than transmitted across the body, as in the case of x-rays. As in most modern imaging, the principle of tomography is used, that is, the person is viewed by radiation detectors surrounding the body, or by rotation of a gamma camera around the body. Such procedures include single-photon emission computed tomography (SPECT), based on the use of iodine-123 or technetium-99m, and positron emission tomography (PET), based on the use of carbon-11 and fluorine-18. The latter two elements are short-lived (carbon-11 half-life is 20 min; fluorine-18 half-life is 110 min) and therefore must be produced near the site where

the studies are performed. *See* RADIOISOTOPE (BIOLOGY).

The nature of the injected material, called a radiopharmaceutical, determines the information that will be obtained. In most cases, either blood flow or biochemical processes within an organ or part of an organ are examined. The essence of a nuclear medicine examination is measurement of the regional chemistry of a living human body.

Examples of commonly used procedures in nuclear medicine using the tracer principle are examination of the blood flow to regional heart muscle with thallium-201- or technetium-99m-labeled radiopharmaceuticals, imaging the regional movements of the ventricles of the heart, detection of blood clots in the lung or impaired lung function, detection of breast and prostate tumors, detection of acute inflammation of the gallbladder, and examination of practically all organs of the body.

Positron emission tomography and single-photon emission computed tomography are used to study regional blood flow, substrate metabolism, and chemical information transfer. In the last category, positron emission tomography has been used to establish the biological basis of neurological and psychiatric disorders, and may help improve the drug treatment of depression, Parkinson's disease, epilepsy, tardive dyskinesia, Alzheimer's disease, and substance abuse. Advances in PET and SPECT and the use of simple detector systems may help in the monitoring of the response of an individual to drug treatment, and perhaps reduce the incidence of side effects. These methods can also provide information on the physiologic severity of coronary stenosis and myocardial viability, especially after thrombolytic therapy or other forms of treatment.

One of the most important areas of research in nuclear medicine is the study of recognition sites, that is, the mechanisms by which cells communicate with each other. For example, some tumors possess recognition sites, such as estrogen receptors.

Another area is in assessment of the availability of receptors that are the primary site of action of many medications. Specific effects of a drug begin by the binding of the drug to specific chemical receptors on specific cells of the body. For example, the finding that Parkinson's disease involves the neurotransmitter dopamine led to the development of L-DOPA treatment, which relieves many of the symptoms of the disease. Measurement of abnormalities of predopaminergic neurons makes it possible to characterize the abnormalities of pre-synaptic neurons in individuals with Parkinson's disease early in their illness at a time when the progress of the disease might be halted.

Treatment. In some diseases, radiation can be used to produce a biological effect. An example is the use of radioactive iodine to treat hyperthyroidism or cancer of the thyroid. The effects of treatment can be assessed with nuclear medicine techniques as well. For example, the metabolism of pituitary tumors can be used as an index of the effectiveness of chemotherapy with drugs that stimulate dopamine receptors. *See* MEDICAL IMAGING; RADIATION THERAPY; RADIOLOGY.

Henry N. Wagner, Jr.

Bibliography. H. N. Wagner, Jr., A new era of certainty (Highlights of the 1995 Society of Nuclear Medicine Annual Meeting, Minneapolis), *J. Nucl. Med.*, 36:13N–28N, 1995; H. N. Wagner, Jr., Z. Szabo, and J. W. Buchanan (eds.), *Principles of Nuclear Medicine*, 2d ed., 1995.

Nuclear molecule

A quasistable entity of nuclear dimensions formed in nuclear collisions and comprising two or more discrete nuclei that retain their identities and are bound together by strong nuclear forces. Whereas the stable molecules of chemistry and biology consist of atoms bound through various electronic mechanisms, nuclear molecules do not form in nature except possibly in the hearts of giant stars; this simply reflects the fact that all nuclei carry positive electrical charges, and that under all natural conditions the long-range electrostatic repulsion prevents nuclear components from coming within the grasp of the short-range attractive nuclear force which could provide molecular binding. But in energetic collisions this electrostatic repulsion can be overcome.

Nuclear molecules were first suggested, in 1960, by D. A. Bromley, J. A. Kuehner, and E. Almqvist to explain very surprising experimental results obtained in the first studies on collisions between carbon nuclei carried out under conditions of high precision. In the original discovery, three sharp resonances appeared in the region of the Coulomb barrier corresponding to states in the ^{24}Mg compound nucleus near 20 MeV of excitation. These were completely unexpected and indicated a new mode of behavior for the 24-body ^{24}Mg system. Although such phenomena were known only in this ^{24}Mg system for almost 10 years, they have subsequently been shown to be ubiquitous features of both nuclear structure and nuclear dynamics. The exact mechanisms leading to the appearance of these molecular configurations are not yet known, however.

Thus far, attention has been focused on dinuclear molecular systems; the search for polynuclear configurations could provide important new insight into the behavior of many-body systems, generally.

^{12}C + ^{12}C system. In 1975 interest in molecular phenomena was renewed by discovery by Y. Abe that the original theoretical models proposed to explain the 1960 discovery were much richer in terms of predicted phenomena than had been realized. Experimental studies rapidly increased the number of resonances in the ^{12}C + ^{12}C system from the original 3 to a total of 39 (**Fig. 1***a*), established the spins and parities of many of them, and further demonstrated that the wave functions of the states in ^{24}Mg that corresponded to the resonances had a marked molecular structure in which the ^{12}C nuclei retained their identity.

Models of molecular phenomena. Bromley, Kuehner, and Almqvist had noted that the effective potential,

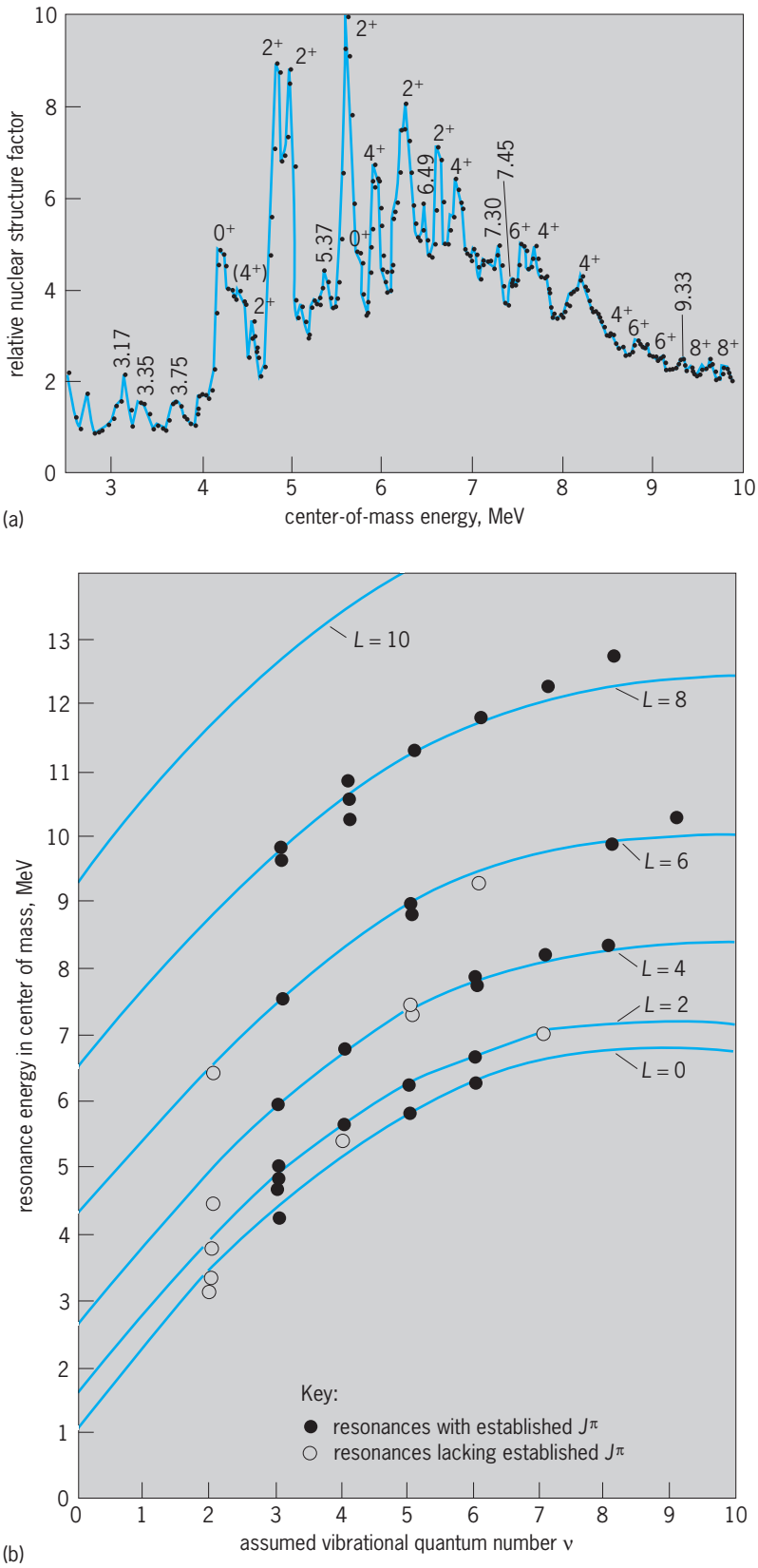

Fig. 1. Resonances in the $^{12}C + {}^{12}C$ system. (*a*) Total cross section for $^{12}C + {}^{12}C$ interactions as a function of the center-of-mass energy and expressed as a nuclear structure factor to remove penetrability effects. Numbers with superscript indicate spin (*J*) and parity (π). (*b*) Resonances correlated in terms of a U(4) symmetry and the corresponding analytic vibration-rotation spectrum. Curves are obtained from the equation in the text with $D = 0.34$ MeV; $a = 1.44$ MeV; $b = 0.08$ MeV; $c = 0.0757$ MeV.

composed of the sum of nuclear, Coulomb, and centrifugal components, involved in the $^{12}C + {}^{12}C$ collisions closely resembled a familiar Morse molecular one. In 1981 F. Iachello demonstrated that in such a potential it was possible, by using group theoretical techniques, to obtain an analytic expression for the allowed quantum states of the form of the equation below,where v and L are the vibrational and rotational quantum numbers and D, a, b, and c are constants. The governing group here is U(4). This simple expression provides a remarkably successful and complete description of all the experimental data (Fig. 1*b*). *See* MOLECULAR STRUCTURE AND SPECTRA.

$$E(v, L) = -D + a(v + {}^1/_2) - b(v + {}^1/_2)^2 + cL(L+1)$$

The fact that this model predicts molecular dissociation at around 7 MeV suggested that the O^+ state in ^{12}C at 7.66 MeV might play an essential role in the molecular interaction rather than the 2^+ at 4.43 Mev which had been assumed to be critical to all earlier models. By using a constrained time-dependent Hartree-Fock (TDHF) approach, the density distribution for ^{12}C was first calculated in the ground (**Fig. 2***a*) and excited (Fig. 2*b*) O^+ states; the latter shows a linear, three-alpha-particle structure suggested several years previously. Assuming that one of the ^{12}C nuclei is inelastically excited to this state in an early stage of the collision, the TDHF model allows the collision to be followed in microscopic detail. At appropriately chosen relative energies, the excited carbon oscillates through the unexcited one, with a period of 2.3×10^{-21} s in the case shown in Fig. 2*c*. This is still a very crude model, but is suggestive of the kind of simple relative motions that are present in the molecular configurations.

In the $^{12}C + {}^{16}O$ system, measurements at energies at, and below, the Coulomb barrier show that simple dinuclear molecular configurations are present in the form of rotational bands, just as in $^{12}C + {}^{12}C$, but at higher energies, although the resonances persist, their structure is substantially more complex. Below the barrier, a U(4) group description is again extremely successful.

Phenomena in heavier systems. Sharp, long-lived resonant states appear in the scattering of ^{28}Si on ^{28}Si up to the highest energies studied, 75 MeV of excitation in the ^{56}Ni compound system. Measurements of the angular distributions of the scattered nuclei (**Fig. 3**) show that these resonances have very high angular momentum; measurements have extended to $44\hbar$ (where $\hbar$ is Planck's constant divided by 2π), the highest angular momentum established in any quantum system. This is particularly interesting inasmuch as simple calculations suggest that at $50\hbar$ the ^{56}Ni nucleus should be torn apart by centrifugal forces. Resonance structure also appears in the $^{24}Mg + {}^{24}Mg$ system, but not in $^{28}Si + {}^{30}Si$ and $^{30}Si + {}^{30}Si$. Potential energy contours have been calculated for the compound nuclei involved in all these systems, at the angular momenta involved, assuming a generalized Nilsson shell model. It is suggestive that

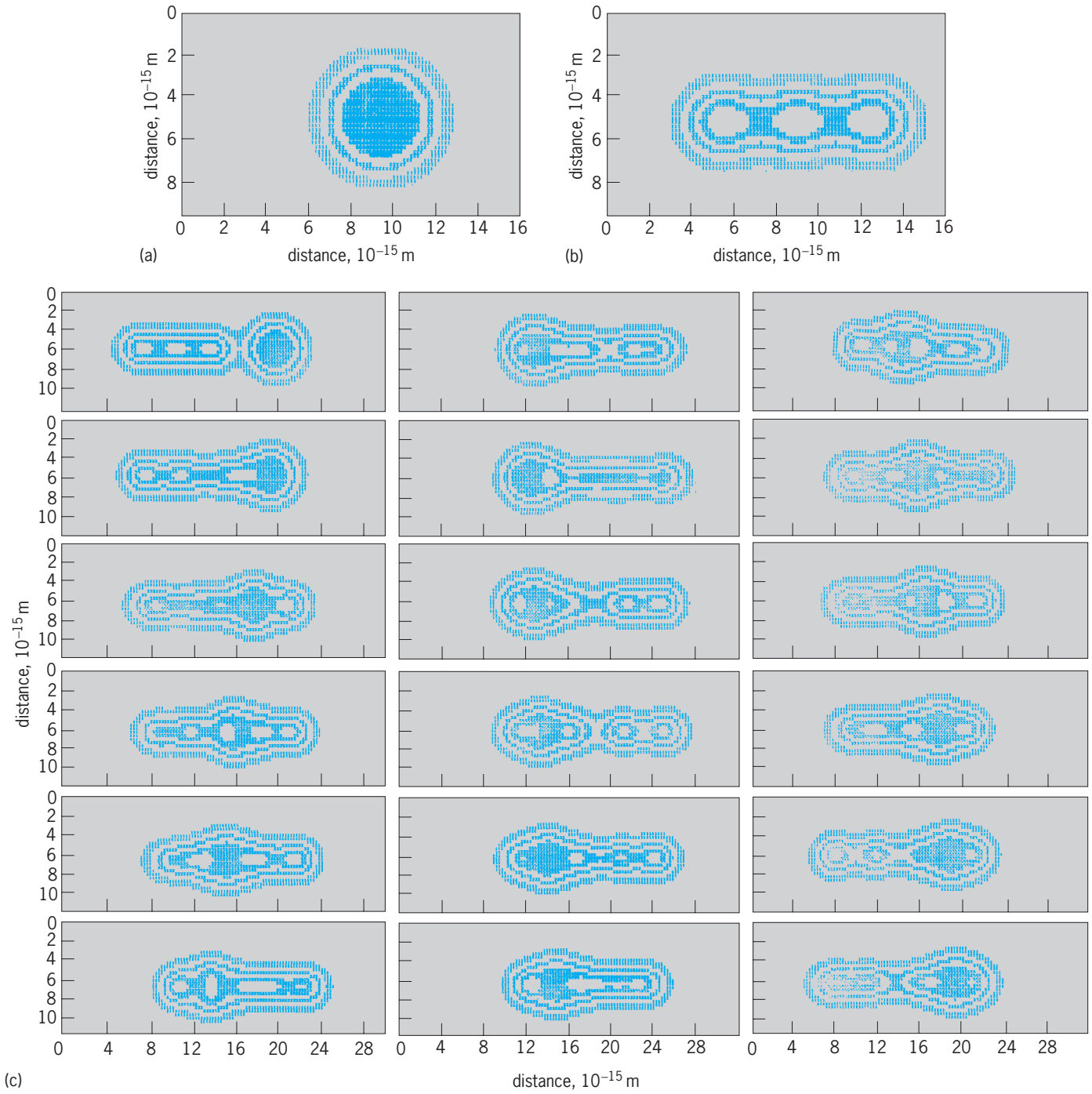

Fig. 2. Density contours calculated by using a constrained, time-dependent Hartree-Fock (TDHF) model. (*a*) Ground state of ^{12}C. (*b*) First excited O^+ state of ^{12}C. (*c*) Time sequence for the interaction of two carbon nuclei, one in each of these configurations.

those systems which show a potential minimum at large prolate deformation appear to be those that show pronounced resonance structure. A quantitative theory that would allow calculation of the molecular phenomena from first principles, however, is still far from available.

Positron spectra have been measured from collisions of very heavy (uranium + uranium, uranium + thorium, uranium + curium) at energies in the region of the Coulomb barrier. A completely unexpected feature of these spectra is a sharp line at about 320 keV. If this is to arise from spontaneous positron creation in the supercritical Coulomb field created in the vicinity of the colliding ions, its width requires that some mechanism exist to maintain this supercritical field for much longer than would be the case in a simple Coulomb scattering. It has been suggested that this would occur if the interacting ions formed a molecular configuration and preliminary calculations show that this would be most probable in collisons in which the ends of the long prolate axes of these nuclei just came into contact. This

configuration has not been established, but all available evidence points to the existence of superheavy nuclear molecular states in these interactions. *See* POSITRON; QUASIATOM; SUPERCRITICAL FIELDS.

In all heavy-ion interactions yet studied, it has been the case that when adequate experimental energy resolution has been available to resolve sharp resonance structure corresponding to molecular phenomena, such structure has appeared. It will be an important challenge for higher-energy, high-precision nuclear accelerators to establish how high in excitation such sharp resonance structures can persist.

Radioactive decay. Indirect evidence supporting very asymmetric molecular configurations in heavy nuclei comes from the discovery of spontaneous radioactive decay of radium nuclei in which ^{14}C, and perhaps ^{16}C, nuclei are emitted. The probability of such decay is approximately 6×10^{-10} of that involving emission of an alpha particle, and the ratio of the probabilities is very closely equal to that calculated for the penetrability of the two products through the Coulomb barriers involved. This suggests that there is a substantial probability for both an alpha particle and ^{14}C to preexist in a molecular configuration. *See* RADIOACTIVITY.

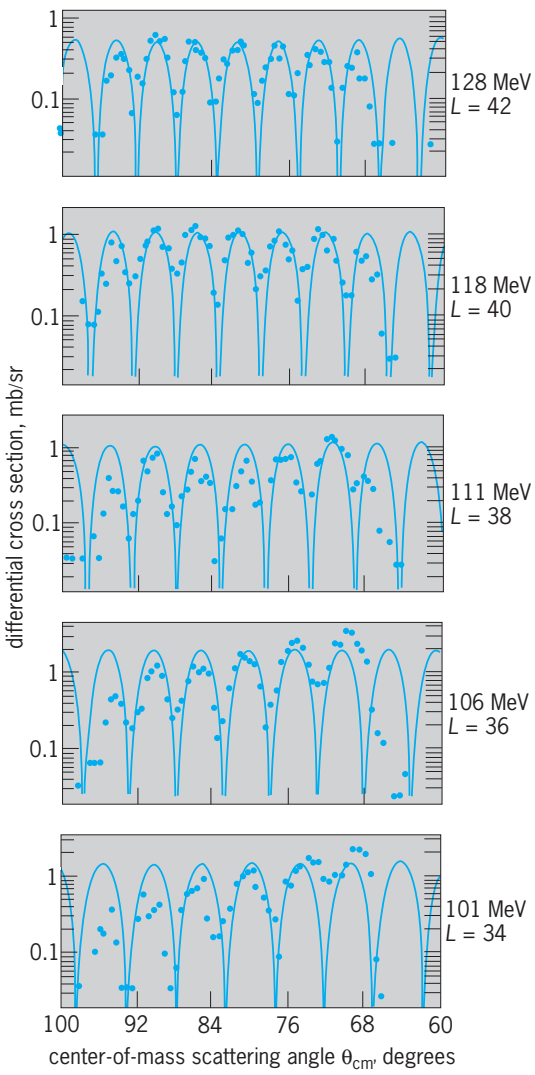

Fig. 3. Angular distributions of ^{28}Si nuclei elastically scattered from ^{28}Si at indicated energies of resonances. The solid curves are obtained by squaring the Legendre polynomial P_L of the indicated order L, giving the angular momentum of the resonance in units of $\hbar$ (Planck's constant divided by 2π). 1 mb = 10^{-31} m^2.

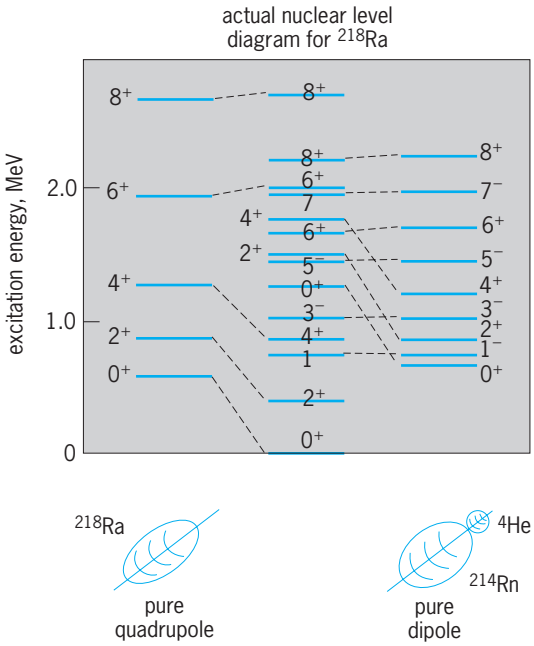

Fig. 4. Mixing of rotational bands based on pure quadrupole and dipole configurations in low-lying states of ^{218}Ra.

Dipole nuclear collectivity. That alpha particles do indeed preexist in these heavy nuclei has been suggested frequently on the basis of observed alpha-particle radioactivity. Rather striking confirmatory evidence has been obtained from measurements on the structure and electromagnetic deexcitation of bound states in radium nuclei. For example, in ^{218}Ra (**Fig. 4**) the presence of low-lying negative parity states and of very strong electric dipole transitions involving them provides strong evidence for the presence of a molecular dipole configuration involving an alpha particle and a ^{214}Rn nucleus. States having this configuration would be expected to mix with the normal quadrupole states except in the case of negative parity ones for which no low-lying quadrupole partner exists. Similar strong dipole effects provide evidence for the existence of alpha-particle molecular configurations in nuclei as light as ^{18}O where the molecular participants are ^{4}H and ^{14}C. *See* NUCLEAR STRUCTURE; SCATTERING EXPERIMENTS (NUCLEI).

D. Allan Bromley

Bibliography. Y. Akaishi et al., *Developments of Nuclear Cluster Dynamics*, 1989; D. A. Bromley, Nuclear molecules, *Sci. Amer.*, 239(6):58–69, 1978; N. Cindro (ed.), *Nuclear Molecular Phenomena*, 1982; K. A. Erb and D. A. Bromley, Nuclear molecular resonances in heavy-ion collisions, *Phys. Today*, 32(1): 34–42, 1979; W. Greiner, W. Sheid, and J. Y. Park, *Nuclear Molecules*, 1995; J. S. Lilley and M. A. Nagarajan (eds.), *Nuclear Cluster Phenomena*, 1984.

Nuclear moments

Intrinsic properties of atomic nuclei; electric moments result from deviations of the nuclear charge distribution from spherical symmetry; magnetic moments are a consequence of the intrinsic spin and the rotational motion of nucleons within the nucleus. The classical definitions of the magnetic and electric multipole moments are written in general in terms of multipole expansions. *See* NUCLEAR STRUCTURE; SPIN (QUANTUM MECHANICS).

Parity conservation allows only even-rank electric moments and odd-rank magnetic moments to be nonzero. The most important terms are the magnetic dipole, given by Eq. (1), and the electric monopole, quadrupole, and hexadecapole, given by Eq. (2), for

$$\vec{\mu} = \int \vec{M}(\vec{r})\, dv \tag{1}$$

$$Q = \frac{1}{e} \int r^l Y_{lm}(\theta, \phi) \rho(\vec{r})\, dv \tag{2}$$

$l = 0, 2, 4$. Here m is the projection of the orbital angular momentum l on a z axis appropriately chosen in space, $\vec{M}\vec{r}$ is the magnetization density of the nucleus and depends on the space coordinates $\vec{r}$, e is the electronic charge, $\rho\vec{r}$ is the charge density in the nucleus, and Y_{lm} are normalized spherical harmonics that depend on the angular coordinates θ and ϕ. *See* MAGNETIZATION; PARITY (QUANTUM MECHANICS); SPHERICAL HARMONICS.

Quantum-mechanically only the z components of the effective operators have nonvanishing values. Magnetic dipole and electric quadrupole moments are usually expressed in terms of the nuclear spin I through Eqs. (3) and (4), where g, the nuclear gy-

$$\frac{\mu}{\mu_N} = gI \tag{3}$$

$$Q(m_I) = \frac{[3m_I{}^2 - I(I+1)]Q}{I(2I-1)} \tag{4}$$

romagnetic factor, is a measure of the coupling of nuclear spins and orbital angular momenta, $\mu_N = eh/4\pi M_p c = 5.0508 \times 10^{-24}$ erg/gauss $= 5.0508 \times 10^{-27}$ joule/tesla, is the nuclear magneton, M_p is the proton mass, h is Planck's constant, e is the electron charge, and c is the speed of light. $Q(m_I)$ is the effective quadrupole moment in the state m_I, and Q is the quadrupole moment for the state $m_I = I$. All angular momenta are expressed in units of $h/2\pi$. The magnitude of g varies between 0 and 1.8, and Q is of the order of 10^{-25} cm^2. *See* ANGULAR MOMENTUM; MAGNETON.

In special cases nuclear moments can be measured by direct methods involving the interaction of the nucleus with an external magnetic field or with an electric field gradient produced by the scattering of high-energy charged particles. In general, however, nuclear moments manifest themselves through the hyperfine interaction between the nuclear moments and the fields or field gradients produced by either the atomic electrons' currents and spins, or the molecular or crystalline electronic and lattice structures. *See* HYPERFINE STRUCTURE.

Effects of nuclear moments. In a free atom the magnetic hyperfine interaction between the nuclear spin $\vec{I}$ and the effective magnetic field $H\vec{r}_e$ associated with electronic angular momentum $\vec{J}$ results in an energy $W = \vec{\mu} \cdot \vec{H}_e = ha\vec{I} \cdot \vec{J}$, which appears as a splitting of the energy levels of the atom. The magnetic field at the nucleus due to atomic electrons can be as large as 10–100 teslas for a neutral atom. The constant a is of the order of 1000 MHz. *See* ATOMIC STRUCTURE AND SPECTRA.

The electric monopole moment is a measure of the nuclear charge and does not give rise to hyperfine interactions. The quadrupole moment Q reflects the deviation of the nuclear charge distribution from a spherical charge distribution. It is responsible for a quadrupole hyperfine interaction energy W_Q, which is proportional to the quadrupole moment Q and to the spatial derivative of the electric field at the nucleus due to the electronic charges, and is given by Eq. (5), where q is the average of expression (6). In

$$W_Q = \frac{e^2 Qq}{2I(2I-1)J(2J-1)} \times [3(\vec{I}\cdot\vec{J})^2 + {}^3/_2(\vec{I}\cdot\vec{J}) - I(I+1)] \tag{5}$$

$$\sum_i \frac{3\cos^2\theta_i - 1}{r_i{}^3} \tag{6}$$

this expression, r_i is the radius vector from the nucleus to the ith electron, and θ_i is the angle between r_i and the z axis. *See* MÖSSBAUER EFFECT.

In free molecules the hyperfine couplings are similar to those encountered in free atoms, but as the charge distributions and the spin coupling of valence electrons vary widely, depending on the nature of the molecular bonding, a greater diversity of magnetic dipole and quadrupole interactions is met. *See* MOLECULAR STRUCTURE AND SPECTRA.

In crystals the hyperfine interaction patterns become extremely complex, because the crystalline electric field is usually strong enough to compete with the spin orbit coupling of the electrons in the ion. Nevertheless, the energy-level structure can often be resolved by selective experiments at low temperatures on dilute concentrations of the ion of interest. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; MAGNETIC RELAXATION; MAGNETIC RESONANCE.

Measurement. The hyperfine interactions affect the energy levels of either the nuclei or the atoms, molecules, or ions, and therefore can be observed either in nuclear parameters or in the atomic, molecular, or ionic structure. The many different techniques that have been developed to measure nuclear moments can be grossly grouped in three categories: the conventional techniques based mostly on spectroscopy of energy levels, the methods based on the detection of nuclear radiation from aligned excited nuclei, and techniques involving the interactions of fast ions with matter or of fast ions with laser beams.

Hyperfine structure of spectral lines. The hyperfine interaction causes a splitting of the electronic energy levels which is proportional to the magnitude of the nuclear moments, to the angular momenta I and J of the nuclei and their electronic environment, and to the magnetic field or electric field gradient at the nucleus. The magnitude of the splitting is determined by the nuclear moments, and the multiplicity of levels is given by the relevant angular momenta I or J involved in the interaction. The energy levels are identified either by optical or microwave spectroscopy.

Optical spectroscopy (in the visible and ultraviolet) has the advantage of allowing the study of atomic excited states and of atoms in different states of ionization. Furthermore, optical spectra provide a direct measure of the monopole moments, which are manifested as shifts in the energy levels of atoms of different nuclear isotopes exhibiting different nuclear radii and charge distributions. Optical spectroscopy has a special advantage over other methods in that the intensity of the lines often yields the sign of the interaction constant a. *See* ISOTOPE SHIFT; SPECTROSCOPY.

Microwave spectroscopy is a high-resolution technique involving the attenuation of a signal in a waveguide containing the absorber in the form of a low-pressure gas. The states are identified by the observation of electric dipole transitions of the order of 20,000 MHz. The levels are split by quadrupole interactions of the order of 100 MHz. Very precise quadrupole couplings are obtained, as well as vibrational and rotational constants of molecules, and nuclear spins. *See* MICROWAVE SPECTROSCOPY.

Atomic and molecular beams and nuclear resonance. Atomic and molecular beams passing through inhomogeneous magnetic fields are deflected by an amount depending on the nuclear moment. However, because of the small size of the nuclear moment, the observable effect is very small. The addition of a radio-frequency magnetic field at the frequency corresponding to the energy difference between hyperfine electronic states has vastly extended the scope of the technique. For nuclei in solids, liquids, or gases, the internal magnetic fields and gradients of the electric fields may be quenched if the pairing of electrons and the interaction between the nuclear magnetic moment and the external field dominate. The molecular beam apparatus is designed to detect the change in orientation of the nuclei, while the nuclear magnetic resonance system is designed to detect absorbed power (resonance absorption) or a signal induced at resonance in a pick-up coil around the sample (nuclear induction). The required frequencies for fields of about 0.5 tesla are of the order of 1–5 MHz. The principal calibration of the field is accomplished in relation to the resonant frequency for the proton whose g-factor is accurately known. Sensitivities of 1 part in 10^8 are possible under optimum experimental conditions. The constant a for ^{133}Cs has been measured to 1 part in 10^{10}, and this isotope is used as a time standard. *See* ATOMIC CLOCK; MOLECULAR BEAMS.

The existence of quadrupole interactions produces a broadening of the resonance line above the natural width and a definite structure determined by the value of the nuclear spin. In some crystals the electric field gradient at the nucleus is large enough to split the energy levels without the need for an external field, and pure quadrupole resonance spectra are observed. This technique allows very accurate comparison of quadrupole moments of isotopes.

Atomic and molecular beams with radioactive nuclei. The conventional atomic and molecular beam investigations can be applied to radioactive nuclei if the beam current measurement is replaced by the much more sensitive detectors of radiations emitted in a radioactive decay. Moments of nuclei with half-lives down to the order of minutes have been determined. *See* RADIOACTIVITY.

Perturbed angular correlations. The angular distribution and the polarization of radiation emitted by nuclei depend on the angle between the nuclear spin axis and the direction of emissions. In radioactive sources in which the nuclei have been somewhat oriented by a nuclear reaction or by static or dynamic polarization techniques at low temperatures, the ensuing nonisotropic angular correlation of the decay radiation can be perturbed by the application of external magnetic fields, or by the hyperfine interaction between the nuclear moment and the electronic or crystalline fields acting at the nuclear site. Magnetic dipole and electric quadrupole moments of ground and excited nuclear states with half-lives as short as 10^{-9} have been measured. *See* DYNAMIC NUCLEAR POLARIZATION.

Techniques involving interactions of fast ions. Techniques involving the interaction of intense light beams from tuned lasars with fast ion beams have extended the realm of resonance spectroscopy to the study of exotic species, such as nuclei far from stability, fission isomers, and ground-state nuclei with half-lives shorter than minutes. The hyperfine interactions in beams of fast ions traversing magnetic materials result from the coupling between the nuclear moments and unpaired polarized s-electrons, and are strong enough (H_e is of the order of 2×10^3 T) to extend the moment measurements to excited states with lifetimes as short as 10^{-12} s. Progress in atomic and nuclear technology has contributed to the production of hyperfine interactions of increasing strength, thus allowing for the observation of nuclear moments of nuclei and nuclear states of increasing rarity.

Noémie Koller

Bibliography. B. Castel and I. S. Towner, *Modern Theories of Nuclear Moments*, 1990; N. F. Ramsey, *Molecular Beams*, 1956, reprint 1990.

Nuclear orientation

The directional ordering of an assembly of nuclear spins I with respect to some axis in space. Under normal conditions nuclei are not oriented; that is, all directions in space are equally probable. For a system of nuclear spins with rotational symmetry about an

axis, the degree of orientation is completely characterized by the relative populations a_m of the $2I + 1$ magnetic sublevels m ($m = I, I - 1, \ldots, -I$). There are just $2I$ independent values of a_m, since they are normalized to unity, namely,

$$\sum_m a_m = 1$$

Rather than specify these populations directly, it turns out to be more useful to form the moments

$$\sum_m m^\nu a_m$$

since these occur in the theoretical calculations. There are $2I$ independent linear combinations of these moments which are called orientation parameters, $f_k(I)$, and are defined by the equation below. Here $f_0(I) = 1$ and all $f_k(I)$ with $k \geq 2I + 1$ are zero.

$$f_k(I) = \binom{2k}{k}^{-1} I^{-k} \sum_m \sum_{\nu=0}^{k} (-1)^\nu \times \frac{(I-m)!(I+m)!}{(I-m-\nu)!(I+m-k+\nu)!} \binom{k}{\nu}^2 a_m$$

Nuclear polarization and alignment. Nuclear polarization is said to be present when one or more $f_k(I)$ with k-odd is not zero, regardless of the even $f_k(I)$ values. In this case the nuclear spin system is polarized. If all the $f_k(I)$ for k-odd are zero and at least one $f_k(I)$ for k-even is not zero, nuclear alignment is said to be present; that is, the nuclear spin system is aligned. Simply stated, if the z axis is the axis of quantization of the nuclear spin system, polarization represents a net moment along the z axis, whereas alignment does not. Unfortunately, the term nuclear polarization is usually associated with $f_1(I)$, and nuclear alignment with $f_2(I)$, although their meanings are in fact much more general. There are other definitions of nuclear orientation parameters; they are mathematically equivalent to the one above. If the nuclear spin system does not have cylindrical symmetry, a more general definition of nuclear orientation is needed leading to the statistical tensors. *See* SPIN (QUANTUM MECHANICS).

Production. Nuclear orientation can be achieved in various ways. The most obvious way is to modify the energies of the $2I + 1$ magnetic sublevels so as to remove their degeneracy and thereby change the populations of these sublevels. The spin degeneracy can be removed by a magnetic field interacting with the nuclear magnetic dipole moment, or by an inhomogeneous electric field interacting with the nuclear electric quadrupole moment. Significant differences in the populations of the sublevels can be established by cooling the nuclear sample to low temperatures T such that T is in the region around $\Delta E/k$, where ΔE is the energy separation of adjacent magnetic sublevels of energy E_m, and k is the Boltzmann constant. If the nuclear spin system is in thermal equilibrium, the populations a_m are given by the normalized Boltzmann factor

$$\frac{\exp(-E_m/kT)}{\sum_m \exp(-E_m/kT)}$$

This means of producing nuclear orientation is called the static method. In contrast, there is the dynamic method, which is related to optical pumping in gases. There are other ways to produce oriented nuclei; for example, in a nuclear reaction such as the capture of polarized neutrons (produced by magnetic scattering) by unoriented nuclei, the resulting compound nuclei could be polarized. In addition to polarized neutron beams, polarized beams of protons, deuterons, tritons, helium-3, lithium-6, and other nuclei have been produced. *See* BOLTZMANN STATISTICS; DYNAMIC NUCLEAR POLARIZATION; OPTICAL PUMPING.

Applications. Oriented nuclei have proved to be very useful in various fields of physics. They have been used to measure nuclear properties, for example, magnetic dipole and electric quadrupole moments, spins, parities, and mixing rations of nuclear states. Oriented nuclei have been used to examine some of the fundamental properties of nuclear forces, for example, nonconservation of parity in the weak interaction. Measurement of hyperfine fields, electric-field gradients, and other properties relating to the environment of the nucleus have been made by using oriented nuclei. Nuclear orientation thermometry is one of the few sources of a primary temperature scale at low temperatures. Oriented nuclear targets used in conjunction with beams of polarized and unpolarized particles have proved very useful in examining certain aspects of the nuclear force. *See* HYPERFINE STRUCTURE; LOW-TEMPERATURE THERMOMETRY; NUCLEAR MOMENTS; NUCLEAR STRUCTURE; PARITY (QUANTUM MECHANICS).

With the advent of helium-3/helium-4 dilution refrigerators, large superconducting magnets, and new methods of producing oriented nuclei and polarized beams, the field of nuclear orientation physics is expected to grow substantially. *See* CRYOGENICS; SUPERCONDUCTING DEVICES. Harvey Marshak

Bibliography. A. O. Barut et al., *Polarization Dynamics in Nuclear and Particle Physics,* 1993; K. Siegbahn (ed.), *Alpha-, Beta- and Gamma Ray Spectroscopy*, 1968; S. Stringari (ed.), *Spin Polarized Quantum Systems,* 1989; W. J. Thompson and T. B. Clegg, Physics with polarized nuclei, *Phys. Today*, pp. 32–39, February 1979.

Nuclear physics

The discipline involving the structure of atomic nuclei and their interactions with each other, with their constituent particles, and with the whole spectrum of elementary particles that is provided by very large accelerators. The nuclear domain occupies a central position between the atomic range of forces and sizes and those of elementary-particle physics,

characteristically within the nucleons themselves. As the only system in which all the known natural forces can be studied simultaneously, it provides a natural laboratory for the testing and extending of many fundamental symmetries and laws of nature. *See* ATOMIC NUCLEUS; ATOMIC STRUCTURE AND SPECTRA; ELEMENTARY PARTICLE; SYMMETRY LAWS (PHYSICS).

Containing a reasonably large, yet manageable number of strongly interacting components, the nucleus also occupies a central position in the universal many-body problem of physics, falling between the few-body problems, characteristic of elementary-particle interactions, and the extreme many-body situations of plasma physics and condensed matter, in which statistical approaches dominate; it provides the scientist with a rich range of phenomena to investigate—with the hope of understanding these phenomena at a microscopic level. *See* PLASMA (PHYSICS); STATISTICAL MECHANICS.

Activity in the field centers on three broad and interdependent subareas. The first is referred to as classical nuclear physics, wherein the structural and dynamic aspects of nuclear behavior are probed in numerous laboratories, and in many nuclear systems, with the use of a broad range of experimental and theoretical techniques. Second is higher-energy nuclear physics (referred to as medium-energy physics in the United States), which emphasizes the nuclear interior and nuclear interactions with mesonic probes. Third is heavy-ion physics, internationally the most rapidly growing subfield, wherein accelerated beams of nuclei spanning the periodic table are used to study previously inaccessible nuclear phenomena.

Nuclear physics is unique in the extent to which it merges the most fundamental and the most applied topics. Its instrumentation has found broad applicability throughout science, technology, and medicine; nuclear engineering and nuclear medicine are two very important areas of applied specialization. *See* NUCLEAR ENGINEERING; NUCLEAR RADIATION (BIOLOGY); RADIOLOGY.

Nuclear chemistry, certain aspects of condensed matter and materials science, and nuclear physics together constitute the broad field of nuclear science; outside the United States and Canada elementary particle physics is frequently included in this more general classification. *See* ANALOG STATES; COSMIC RAYS; FUNDAMENTAL INTERACTIONS; ISOTOPE; NUCLEAR CHEMISTRY; NUCLEAR FISSION; NUCLEAR FUSION; NUCLEAR ISOMERISM; NUCLEAR MOMENTS; NUCLEAR REACTION; NUCLEAR REACTOR; NUCLEAR SPECTRA; NUCLEAR STRUCTURE; PARTICLE ACCELERATOR; PARTICLE DETECTOR; RADIOACTIVITY; SCATTERING EXPERIMENTS (NUCLEI); WEAK NUCLEAR INTERACTIONS.

D. Allan Bromley

Nuclear power

Power derived from fission or fusion nuclear reactions. More conventionally, nuclear power is interpreted as the utilization of the fission reactions in a nuclear power reactor to produce steam for electric power production, for ship propulsion, or for process heat. Fission reactions involve the breakup of the nucleus of high-mass atoms and yield an energy release which is more than a millionfold greater than that obtained from chemical reactions involving the burning of a fuel. Successful control of the nuclear fission reactions utilizes this intensive source of energy. *See* NUCLEAR FISSION.

Fission reactions provide intensive sources of energy. For example, the fissioning of an atom of uranium yields about 200 MeV, whereas the oxidation of an atom of carbon releases only 4 eV. On a weight basis, this 50×10^6 energy ratio becomes about 2.5×10^6. Uranium consists of several isotopes, only 0.7% of which is uranium-235, the fissile fuel currently used in reactors. Even with these considerations, including the need to enrich the fuel to several percent uranium-235, the fission reactions are attractive energy sources when coupled with abundant and relatively cheap uranium ore.

Nuclear Fuel Cycle

Although the main process of nuclear power is the release of energy in the fission process which occurs in the reactor, there are a number of other important processes, such as mining and waste disposal, which both precede and follow fission. Together they constitute the nuclear fuel cycle. (The term cycle may not be exactly appropriate when the cycle is not closed, as is the case at present.) *See* NUCLEAR FUEL CYCLE.

The only fissionable material now found in nature is uranium-235. (However, prehistoric natural reactors operated in Gabon, Africa, and possibly also in Colorado in the United States. These produced plutonium that has since decayed to uranium-235.) Plutonium becomes part of the fuel cycle when the fertile material uranium-238 is converted into fissile plutonium-239. Thus, the dominant fuel cycle is the uranium-plutonium cycle. If thorium-232 is used in a reactor to produce fissile uranium-233, another cycle emerges, the thorium-uranium cycle. Since the applications of the latter have been limited to small-scale tests, and a commercial thorium-uranium cycle has not been developed, the discussion below will concentrate on the uranium-plutonium cycle. *See* NUCLEAR FUELS; PLUTONIUM; THORIUM; URANIUM.

Fuel preparation. Fissionable materials must be prepared for use in a nuclear reactor through mining, milling, conversion, enrichment, and fabrication.

Mining. The nuclear fuel cycle begins with the mining of uranium or thorium. Uranium is estimated to be present in the Earth's crust to the extent of 4 parts per million. The concentration of thorium is nearly three times greater. However, uranium and thorium are so widely distributed that significant concentrations in workable deposits are more the exception than the rule. Exploitable deposits have on average a concentration of 0.1–0.5% of uranium oxide (U_3O_8) by weight.

Large deposits of uranium-rich minerals are found in many places: in central Africa and around the gold-mining areas of South Africa, in Canada's Great Bear

Lake region in Ontario, and in Australia. Lower-grade ores have been mined extensively on the Colorado Plateau in the United States. There are other deposits being worked in west-central France, in the western mountains of the Czech Republic, in southwestern Hungary, in Gabon, West Africa, and in Madagascar. Also in Africa, rich deposits have been found in the Republic of Niger. Uranium concentration exceeds 15% in a deposit in the Cigar Lake area of Canada.

The chief sources of thorium are coastal sands rich in monazite found at Travancore near the southern tip of India, and on the coast of Brazil. Monazite sands have also been found on the shores of Florida's panhandle.

Milling. After uranium ore has been mined, it is crushed and the host rock separated from the core, usually by a flotation process. The uranium is milled and concentrated as a uranium salt, ammonium diuranate, which is generally known in the industry as yellowcake because of its color.

Conversion. The yellowcake is then shipped to a conversion plant where it is oxidized to uranium oxide and then is fluorinated to produce the uranium hexafluoride (UF_6). This is a convenient form for the gaseous diffusion enrichment process because the uranium hexafluoride sublimates (passes directly from the solid phase to the gaseous phase without liquefying) at 127°F (53°C).

Enrichment. The uranium hexafluoride, familiarly called hex, is shipped in tanklike containers to one of the three United States gaseous diffusion enrichment plants or to one of several other enrichment plants throughout the world. Gas centrifuges are widely used for enrichment outside the United States. *See* ISOTOPE SEPARATION.

After enrichment, the two resulting streams—enriched uranium, and depleted uranium—part company. The depleted uranium is stored adjacent to the diffusion plant, and the enriched material is converted back to an oxide—this time uranium dioxide (UO_2)—and sent to a fuel fabrication plant.

Fabrication. At these plants, uranium dioxide intended for light-water reactor fuel is shaped into cylindrical pellets about 0.6 in. (15 mm) long and 0.4 in. (10 mm) in diameter. The pellets are sintered, that is, baked, to obtain a hard, dense consistency. After polishing, they are loaded into tubes made of Zircaloy, an alloy of zirconium. A number of tubes are put together with appropriate tie plates, fittings, and spacers to form fuel assemblies. *See* ZIRCONIUM.

Fission in power reactors. From about 1955 to 1965, numerous United States companies explored or planned power reactor product lines, and almost every combination of feasible fuel, coolant, and moderator was suggested.

Power reactors in the United States are the light-water-moderated and -cooled reactors (LWRs), including the pressurized-water reactor (PWR) and the boiling-water reactor (BWR), represent 100% of capacity in operation. The high-temperature gas-cooled reactor (HTGR), and the liquid-metal-cooled fast breeder reactor (LMFBR) have reached a high level of development but are not used for commercial purposes.

Of these concepts, the light-water reactor and the high-temperature gas-cooled reactor are thermal reactors; that is, they operate by using moderated neutrons slowed down to "thermal" velocities (so called because their speed is determined by the thermal, or kinetic, energy of the substance in which the fuel is placed, namely the moderator). The third type, the fast breeder, operates on fast neutrons—unmoderated neutrons that have the high velocities near to those with which they are released from a fissioning nucleus. *See* NEUTRON; THERMAL NEUTRONS.

The pressurized-water reactor is in use in France. Canada has developed the CANDU natural-uranium-fueled and heavy-water-moderated and -cooled reactor. The Soviet Union and its successor republics have developed and built considerable numbers of two types of water-cooled reactors. One is a conventional pressurized-water reactor; the other is a tube-type, graphite-moderated reactor.

Boiling-water reactor (BWR). In one boiling water reactor, designed to produce about 3580 MW-thermal and 1220 MW net electric power, the reactor vessel is 238 in. (6 m) in inside diameter, 6 in. (15 cm) thick, and about 71 ft (22 m) in height. The active height of the core containing the fuel assemblies is 148 in. (4 m). Each fuel assembly typically contains 63 fuel rods, and 732 fuel assemblies are used. The diameter of the fuel rod is 0.5 in. (12.5 mm). The reactor is controlled by cruciform-shape (in cross section) control rods moving up from the bottom of the reactor in spaces between the fuel assemblies (177 control rods are provided). The water coolant is circulated up through the fuel assemblies by 20 jet pumps at about 70 atm (7 megapascals), and boiling occurs within the core. The steam is fed through four 26-in.-diameter (66-cm) steamlines to the turbine. As is typical for steam power cycles, about one-third of the energy released by fission is converted into mechanical work, and then to electrical energy, by the turbine-generator, and the remaining heat is removed in the condenser. The condenser operates in the same manner in fossil and nuclear power plants, with the heat it removes having to be dissipated to the environment. Some limited use of the energy rejected by the condenser is possible. The steam produced in the plant's nuclear steam supply system will be at lower temperatures and pressures than that from fossil plants, and thus the efficiency of the nuclear plant in producing electric power is somewhat less, leading to somewhat greater heat rejection to the environment per kilowatthour produced. *See* GENERATOR; STEAM CONDENSER; STEAM TURBINE.

Shielding is provided to reduce radiation levels, and pools of water are used for fuel storage and when access to the core is necessary for fuel transfers. Among the engineered safety features that minimize the consequences of reactor accidents is the primary containment building (**Fig. 1**). The function of the containment building is to cope with the energy released by depressurization of the coolant system should a failure occur in the primary piping, and to provide a secure enclosure to minimize

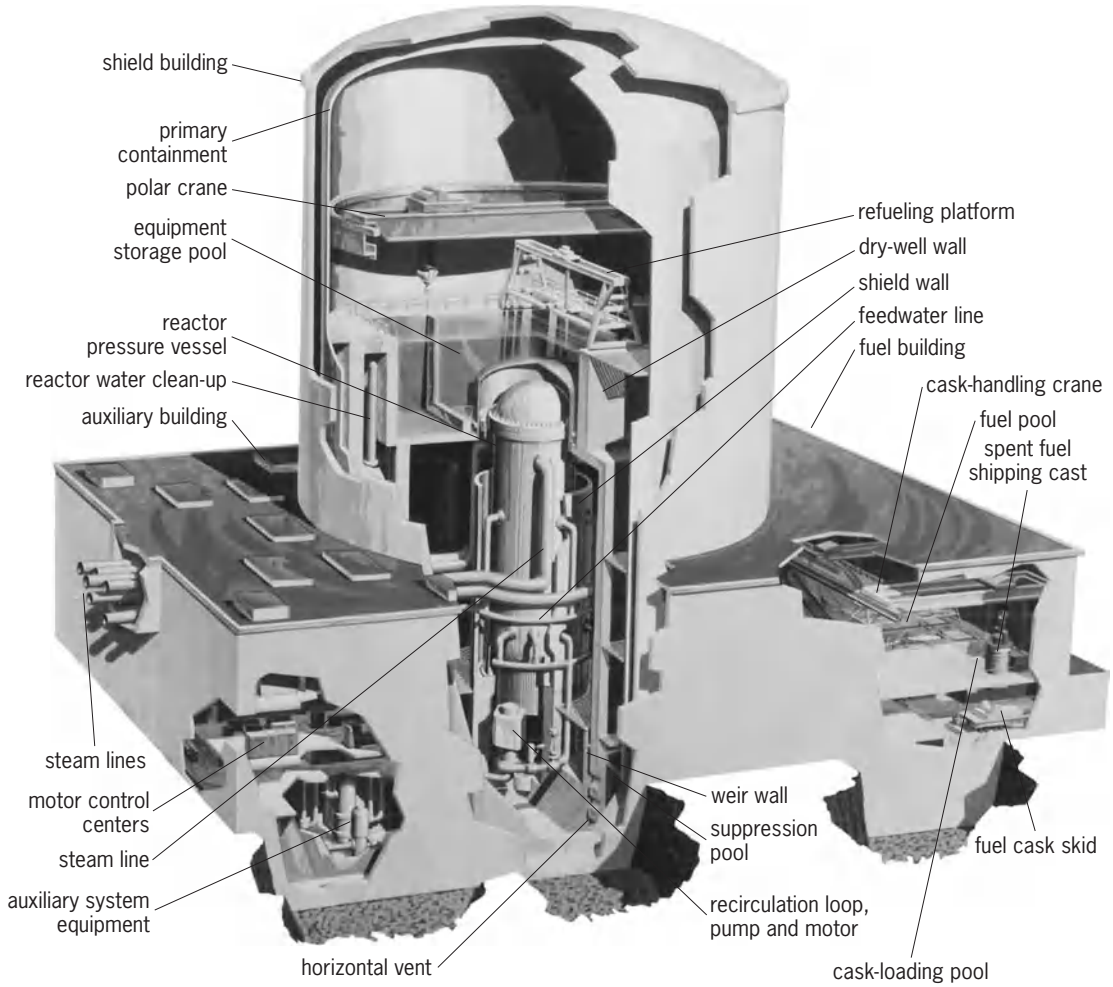

Fig. 1. Mark III containment building for a boiling-water reactor, which illustrates safety features designed to minimize the consequences of reactor accidents. (*General Electric Co.*)

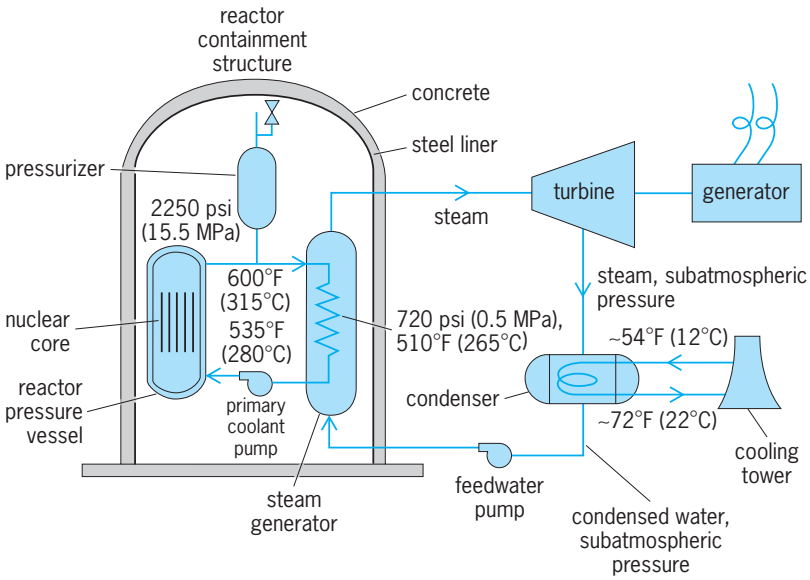

Fig. 2. Schematic of a pressurized-water reactor power plant. Heat transfer occurs within the containment building. (*After F. J. Rahn et al., A Guide to Nuclear Power Technology, Krieger Publishing, 1992*).

leakage of radioactive material to the surroundings. The boiling-water reactor utilizes a pool of water (called a suppression pool) to condense the steam produced by the depressurization of the primary coolant system. Various arrangements have been used for the suppression pool. Other engineered safety features include the emergency core-cooling system, the containment spray system, and the high-efficiency particulate filters for removing radioactivity from the containment building's atmosphere.

Pressurized-water reactor (PWR). Whereas in the boiling-water reactor a direct cycle is used in which steam from the reactor is fed to the turbine, the pressurized-water reactor employs a closed system (**Fig. 2**). The water coolant in the primary system is pumped through the reactor vessel, transporting the heat to a steam generator, and is recirculated in a closed primary system. A separate secondary water system is used on the shell side of the steam generator to produce steam, which is fed to the turbine, condensed, and recycled to the steam generator. A pressurizer is used in the primary system to maintain about 150 atm (15 MPa) pressure to suppress boiling in the primary coolant. Up to four loops have been used.

The reactor pressure vessel is about 44 ft (13.5 m) high and about 15 ft (4.4 m) in inside diameter, and has wall thickness exceeding 8.5 in. (22 cm). The active length of the fuel assemblies may range from 12 to 14 ft (about 4 m), and different configurations are used by the manufacturers. For example, one type of fuel assembly contains 264 fuel rods, and 193 fuel assemblies are used for the 3411-MW-thermal, four-loop plant. The outside diameter of the fuel rods is 0.4 in. (9.5 mm). For this arrangement, the control rods are grouped in clusters of 24 rods, with 61 clusters provided. In pressurized-water reactors, the control rods enter from the top of the core. Reactor operations are carried out by using both the control rods and a system to increase or decrease the boric acid content of the primary coolant. The latter controls the gross reactivity of the core, while the former allows fine tuning of the nuclear reaction. A typical prestressed concrete containment building (**Fig. 3**) is designed for a 4 atm (400 kilopascals) rise in pressure, with an inside diameter of about 116 ft (35.3 m) and height of 208 ft (64 m). The walls are 45 in. (1.3 m) thick. The containments have cooling and radioactive absorption systems as part of the engineered safety features. Another design (**Fig. 4**) uses a spherical steel containment (about 200 ft, or 60 m, in diameter), surrounded by a reinforced concrete shield building (**Fig. 5**). *See* NUCLEAR REACTOR.

Breeder reactor. In a breeder reactor, more fissile fuel is generated than is consumed. For example, in the fissioning of uranium-235, the neutrons released by fission are used both to continue the neutron chain reaction which produces the fission and to react with uranium-238 to produce uranium-239. The uranium-239 in turn decays to neptunium-239 and then to plutonium-239. Uranium-238 is called a fertile fuel, and uranium 235, as well as plutonium-239, is a fissile fuel and can be used in nuclear power reactors. The reactions noted can be used to convert most of the uranium-238 to plutonium-239 and thus provide about a 60-fold extension in the available uranium energy source relative to its use in nonbreeder reactors. Breeder power reactors can be used to generate electric power and to produce more fissile fuel. Breeding is also possible by using the fertile material thorium-232, which in turn is converted to the fissile fuel uranium-233. An almost inexhaustible energy source becomes possible with breeder reactors. The use of breeder reactors would decrease the long-range needs for enrichment and for mining more uranium.

Prototype breeder reactors are in use in France, England, Japan, and Russia. A full-scale (1250 MW-electric) commercial breeder, the Super Phenix, was constructed in France and operated commercially between 1986 and 1999.

Advanced light-water reactor (ALWR). The most advanced nuclear power reactor designs, referred to as advanced light-water reactors, have evolved from operating experience with light-water reactors. Advanced light-water reactors represent highly engineered designs that emphasize higher reliability,

Fig. 3. Oconee Nuclear Power Station (Greenville, South Carolina) containment structures.

Fig. 4. Spherical containment design for a pressurized-water reactor. (*Combustion Engineering, Inc.*)

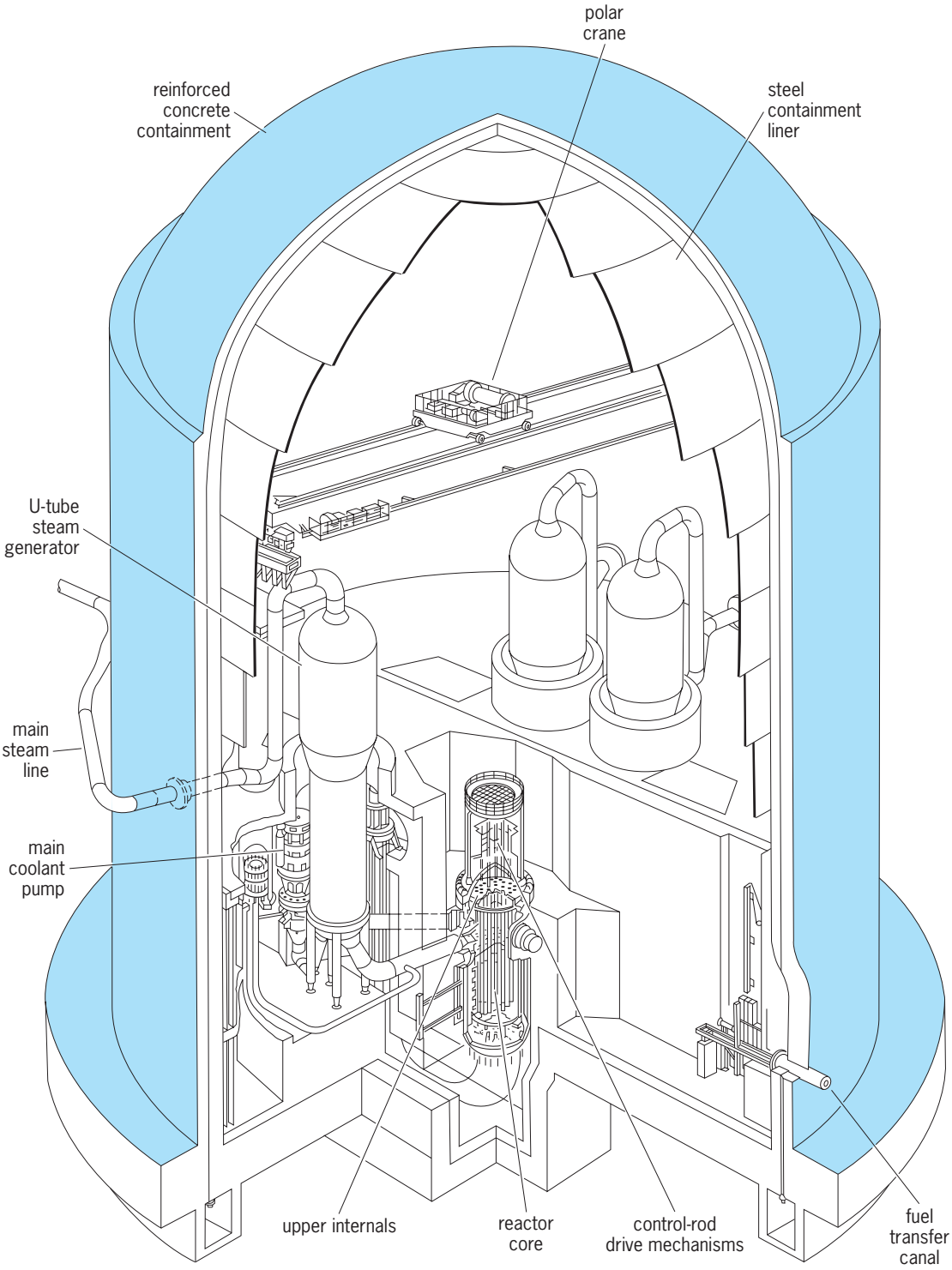

Fig. 5. Cutaway view of containment building of a typical pressurized-water reactor system. (*Westinghouse Electric Corp.*)

have few components and complex systems, are easier to build and maintain than current reactors, and, in particular, are safer and less sensitive to transients. The higher degree of safety is made possible by the use of redundant, passive systems that do not require power for actuation, and the low sensitivity to transients is achieved through low power density in the reactor core and a large inventory of coolant and containment water. Higher reliability depends on the use of systems and components that are proven in service. The designs feature construction periods of as little as 4 years from the start of site excavation, low operating and maintenance costs, long fuel cycles, and simplified maintenance. Excellent fuel-cycle economics are due to high fuel burn-up and a flexible operating cycle. The average occupational radiation dose equivalent is less than 0.5 sievert per year. The plants require increased

Advanced light-water reactor (ALWR) designs

Name	Features
Advanced boiling-water reactor (ABWR)	1300-MW-electric reactor with simplified safety systems, a completely internal recirculation system, and improved controls
System 80+	1350-MW-electric pressurized-water reactor with greater design margins, more redundancy and diversity, and better operability and maintainability
AP-600	600-MW-electric midsize pressurized-water reactor with simplified design and safety features that allow 50% fewer valves, 35% fewer pumps, and 80% less safety-grade piping
SWR-1000	977-MW-electric boiling-water reactor with proven component design, no need for operator intervention in case of accidents, low power density, and 60-year design lifetime
Simplified boiling-water reactor (SBWR)	600-MW-electric midsize reactor employing natural water circulation for power generation, with gravity-driven emergency core cooling from an elevated suppression pool

spent fuel storage to accommodate all the fuel discharged over the plant lifetime.

International participation has been a critical factor in the development of advanced light-water reactors. The **table** summarizes the plants that are offered. Due to the lengthy design, licensing, and prototype development required for new-generation reactors, the majority of the plants constructed in the period through 2005 are expected to be those in the table.

The first advanced nuclear plant to be constructed was a twin-unit advanced boiling-water reactor (ABWR; **Fig. 6**) at the Kashiwazaki-Kariwa site in Japan. The plant produced power for the electric grid 58 months after groundbreaking. Another advanced reactor concept isthe System 80+, under

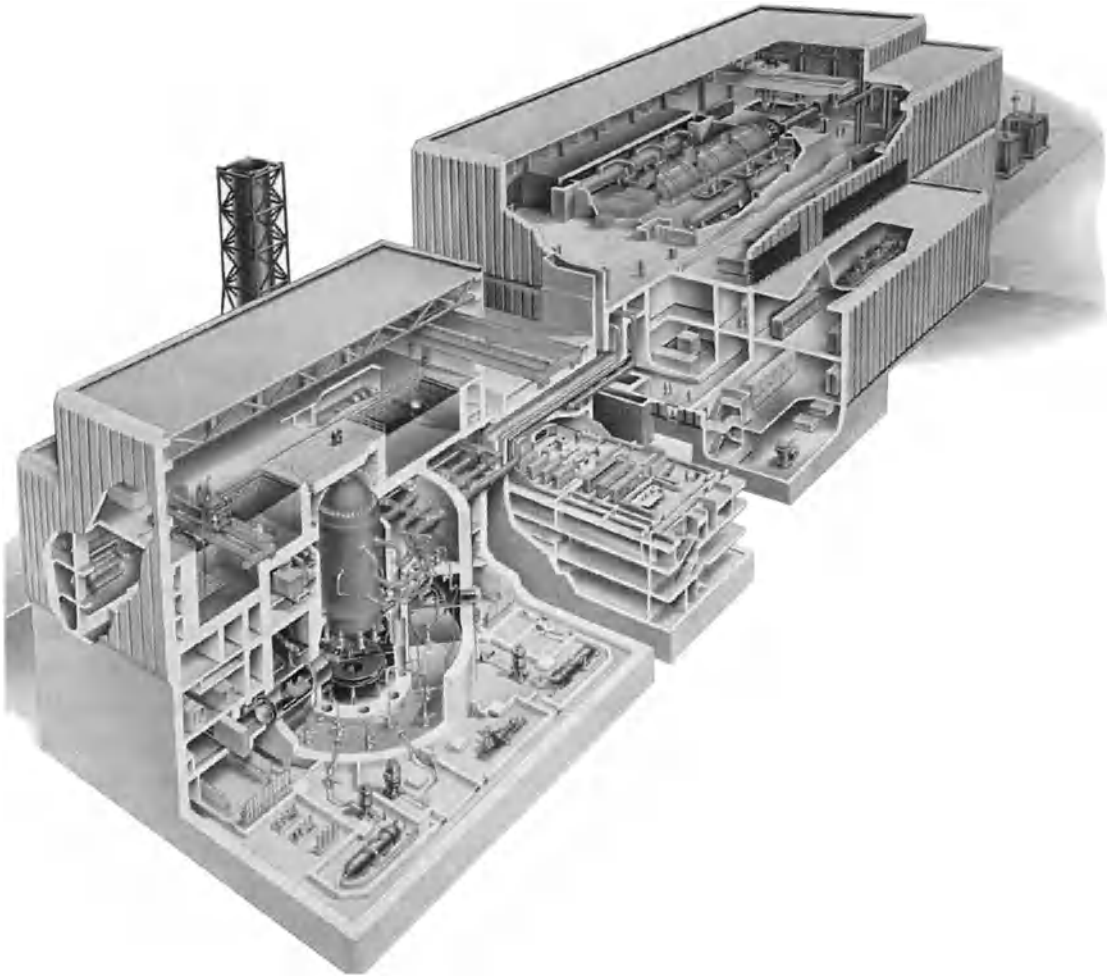

Fig. 6. Advanced boiling-water reactor (ABWR), the first advanced light-water reactor to be deployed commercially. (*General Electric*)

construction in Korea. In 1998, following a 12-year design program, the AP-600 plant became the first advanced light-water reactor to be certified by the U.S. Nuclear Regulatory Commission.

Fuel loading and reactor operations. Completed fuel assemblies are shipped from the fuel fabrication plant to the power plant in criticality-proof containers and on arrival are stored in a clean fuel vault.

Bolting or unbolting the head of a reactor vessel, to close or open it, is a complex undertaking. A ring of bolts each 2 ft (60 cm) long and 3 in. (8 cm) in diameter is used, with nuts of corresponding magnitude that are fastened in place with mechanical bolt-tighteners. At refueling time, the nuts are removed, and a polar crane at the top of the reactor containment building eases the vessel head off its seating and hoists it to one side onto a pad. The reactor cavity—the concrete pit in which the reactor vessel is moored—is filled with water to a height such that full-length (usually 12 ft or 4 m long) fuel assemblies lifted out of the core can swing clear of the lip of the reactor vessel without coming out of the water. The purpose of the water is to shield the workers from the radiation of the fuel assemblies. Those fuel assemblies removed are lowered into a canal that connects the reactor cavity with the spent-fuel pool by a bridge crane that spans the pool and runs on rails its full length, and as far as the reactor cavity. In this manner, the discharged assemblies are moved to an open position in the racks on the floor of the spent-fuel pool. These racks are made of a metal containing enough neutron poison material to make certain that the spent assemblies in the pool cannot go critical there.

The assemblies not ready for discharge are shuffled, that is, their position is rearranged in the core, and the fresh clean assemblies are then loaded into the core grid spaces freed. Underwater periscopes are provided for use in the reactor vessel and in the spent-fuel pool as required. Assemblies discharged are inspected visually by means of these periscopes to check their condition, and a log is kept, recording the serial number of each assembly moved into or out of the core. The radioactive content of the water is also checked to make certain that none of the fuel rods has developed any pinhole leaks that permit release of fission products into the water. *See* PERISCOPE.

After the reactor has been "buttoned up" again, that is, the vessel head secured back in place after completion of reloading, the director of reloading operations turns the reactor back to the operating staff for startup. In the process of starting up, a startup source—a neutron emitter—is used as there are not enough spontaneous neutrons emitted to permit close monitoring of the startup process by external neutron detectors. A neutron source (usually an antimony or a polonium-beryllium capsule) which emits a large number of neutrons spontaneously is used for this purpose, and is lowered into the core through a guide tube while the control rods are slowly and carefully retracted.

Radiation monitors around the core report remotely to the control room the neutron flux as it increases. Flux is the number of neutrons crossing a unit area per unit time and is proportional to the fission reaction rate. Once the reactor has become critical, the control rods are retracted further to increase power until the reactor is at 100% power.

Spent-fuel storage and transportation. The predominant fueling strategy requires that either one-third or one-fourth of the fuel assemblies be discharged from the core each year. While fresh assemblies are placed in the core to replenish its reactivity, the spent fuel is transferred to an adjacent pool where it is stored for several years. Spent fuel is highly radioactive, and the water serves as shielding and as a cooling medium, to remove the heat produced by the decaying fission products. After several years of cooling in the spent-fuel pool, the assemblies are suitable for transportation to a reprocessing plant. However, this option is not available in the United States, and spent fuel must be stored in the spent-fuel pool at the reactor. Since the amounts of stored fuel increases every year, and because utilities must maintain enough space in the pool for a full core discharge, spent fuel has required changes in storage. These include: redesigning or modifying the pools to allow for denser storage patterns in the same total space available; building more storage capacity at the reactor site; allowing transfers between reactor sites or between utilities to complement needs and available space; and building centralized storage capacity "away from reactor" (AFR) to accommodate the excess quantities. The Department of Energy was required by law to accept spent fuel for permanent storage or disposal, starting in the 1990s, but has not yet done so pending resolution of technical and legal issues.

Transportation of spent nuclear fuel and nuclear wastes has received special attention. With increasing truck and train shipments and increased probabilities for accidents, the protection of the public from radioactive hazards is achieved through regulations and implementation which provide transport packages with multiple barriers to withstand major accidents. For example, the cask used to transport irradiated fuel is designed to withstand severe drop, puncture, fire, and immersion tests. Actual train and truck collision tests have been done to demonstrate the integrity of the casks.

Reprocessing and refabrication. At the reprocessing center, the spent-fuel rods are stripped of cladding, and the spent-fuel pellets are dropped in a pool of nitric acid in which they dissolve. The solution is then fed to countercurrent extraction systems. Usually, in the first extraction cycle about 99% of the fission waste products are removed. Then further purification and separation of the plutonium from the uranium is performed. The end products of this step are usually uranium dioxide and plutonium dioxide (PuO_2) which can be recycled. The separation is a straightforward chemical process that is carried out by the United States government for weapons material and for spent fuel from nuclear-propelled naval vessels. The reprocessing of commercial fuel is done in order to return the unfissioned fuel material to the

inventory of material to be used for fuel fabrication. Commercial nuclear fuel is reprocessed in France, Great Britain, Russia, and on a lesser scale in Belgium, Germany, Japan, India, and Italy. Commercial fuel reprocessing activities were discontinued in the United States in 1976. *See* NUCLEAR FUELS REPROCESSING.

Radioactive waste management. The fission waste products are removed from a reprocessing plant and disposed of in various ways. High-level waste can be concentrated into a glassy bead form, and either buried in salt beds deep in the earth or shipped to a heavily guarded disposal site. Low-level wastes are stored in liquid or solid form.

Critics of nuclear power consider the radioactive wastes generated by the nuclear industry to be too great a burden for society to bear. They argue that since the high-level wastes will contain highly toxic materials with long half-lives, such as a few tenths of one percent of plutonium that was in the irradiated fuel, the safekeeping of these materials must be assured for time periods longer than social orders have existed in the past. Nuclear proponents answer that the time required for isolation is much shorter, since only 500 to 1000 years is needed before the hazard posed by nuclear waste falls below that posed by common natural ore deposits in the environment (**Fig. 7**). The proposed use of bedded salts, for example, found in geologic formations that have prevented access of water and have been undisturbed for millions of years provides one of the options for assurance that safe storage can be engineered. A relatively small area of several hundred acres (a few hundred hectares) would be needed for disposal of projected wastes.

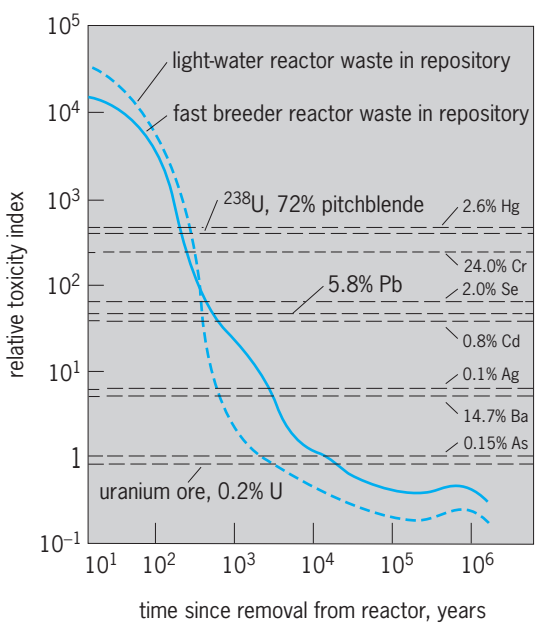

Fig. 7. Relative toxicity of nuclear waste over time, compared with that of average mineral ores of toxic elements. (*After K. A. Tonnessen and J. J. Cohen, Survey of naturally occurring hazardous materials in deep geologic formations: A perspective on the relative hazard of deep burial of nuclear wastes, University of California and Lawrence Livermore Laboratories, UCRL-52199, 1977*)

Management of low-level wastes generated by the nuclear energy industry requires use of burial sites for isolation of the wastes while they decay to innocuous levels. Other radioactive wastes, such as those from medical procedures and industrial applications, equal or exceed these from nuclear power plants. Operation of the commercial burial sites is subject to regulations by federal and state agencies.

Routine operations of nuclear power stations result in very small releases of radioactivity in the gaseous and water effluents. The NRC has adopted the principle that all releases should conform to the "as low as reasonably achievable" (ALARA) standard. ALARA guidance has been extended to other portions of the nuclear fuel cycle. *See* RADIOACTIVE WASTE MANAGEMENT.

Decommissioning of nuclear plants. There are several light-water reactors in the United States and many light-water reactors and gas-cooled reactors in Europe that have been decommissioned.

Decommissioning requires several steps: assessments and characterization of the hazards, environmental review, engineering, dismantling and decontamination of the plant, and final and remediation. The process requires that the plant be maintained cost-effectively, safely and securely while it awaits decommissioning, and that the site and buildings be made available for reuse following decommissioning (if practical). Some nuclear plants have been considered for conversion to a fossil-fuel power station.

In the United States, there are essentially three options for decommissioning: prompt DECON, delayed DECON, and SAFSTOR. Prompt DECON requires that the equipment, structures, and portions of a facility and site containing radioactive contaminants be removed or decontaminated to a level that permits unrestricted use. Delayed DECON is essentially the same as prompt DECON, but includes a delay for onsite spent fuel storage (usually 10–20 years) to allow sufficient time for the U.S. Department of Energy to develop a spent-fuel repository. SAFSTOR requires that the facility be placed and maintained in a condition that allows it to be safely stored and later decontaminated for unrestricted use, generally within about 60 years of reactor shutdown.

Safety

Nuclear power facilities present a potential hazard rarely encounted with other facilities; that is, radiation. A major health hazard would result if, for instance, a significant fraction of the core inventory of a power reactor were released to the atmosphere. Such a release of radioactivity is clearly unacceptable, and steps are taken to assure it could never happen. These include use of engineered safety systems, various construction and design codes (for example, standards of the American Society for Testing and Materials), regulations on reactor operation, and periodic maintenance and inspection. In the last analysis, however, the ultimate safety of any facility depends on the ability of its designers to use the forces of nature so that a large release of radioactivity

is not possible. To help them, various techniques are employed, including conservative design margins, the use of safety equipment, and reliance on various physical barriers to radiation release in case all else fails.

It is the practice, in the United States and elsewhere, for regulatory bodies to establish licensing procedures for nuclear facilities. These procedures set design requirements, construction practices, operational limits, and the siting of such facilities. All power reactors built in the United States (and overseas except in the former Soviet Union) have a containment building and are sited in generally low or moderate population areas with exclusion areas defined by regulation.

All reactors have engineered safety features, both active and passive, designed to prevent serious accidents and mitigate them if they occur. A nuclear plant's safety is achieved through a concept of defense in depth. This provides a series of protective barriers with redundancy at each level and for each active component.

Every reactor has four main barriers to radioactivity release in the event of an accident:

1. *Fuel matrix*. The exceptionally high melting point (5000°F or 2760°C) and chemical stability of uranium dioxide prevent the escape of fission products except in extreme accident conditions. Although the fission process creates large amounts of radioactivity in the fuel rods, the ceramic pellets of uranium dioxide fuel retain more than 98% of this radioactivity. Without fuel melting and subsequent release of fission products, a nuclear reactor accident involves only hazards to the general public comparable with conventional power plant accidents.

2. *Fuel cladding*. The Zircaloy clad surrounding the fuel pellets retains any radioactivity released from the uranium dioxide. Fuel cladding behavior is of importance to the safety of a nuclear plant primarily because the fuel contains the major part of the radioactive products in the plant. The cladding is protected through use of design criteria which limit the heat in the core.

3. *Reactor primary coolant system*. Boundary integrity is assured by the thick steel vessel and piping up to 8 in. (20 cm) thick, and the continual inspection of these components. Licensing requirements specify that the design of the safety systems must accommodate ruptures in the reactor coolant loop, including a break of the largest coolant pipe. This constitutes the design basis for protection against loss-of-coolant accidents.

4. *Reactor containment building*. The reactor containment building generally consists of a 4-ft-thick (1.2-m) concrete shell lined with steel, and is heavily reinforced with steel bars. This steel, embedded in the concrete, gives a reactor containment building great strength to withstand forces that might occur in a reactor accident. *See* REINFORCED CONCRETE.

Each of these barriers is capable of retaining the hazardous radioactivity contained in a reactor core. Only if all were simultaneously breached would a release to the environment be possible. To prevent breaches, and to attenuate any radioactivity, various other engineered safety features are employed. Their objectives are threefold: shut down the reactor (in the event that the normal control rod mechanisms fail), cool the core (to prevent overheating and meltdown), and safeguard the integrity of the barriers (such as limiting the pressure in a containment building).

Frank J. Rahn

Bibliography. J. G. Collier and G. F. Hewitt, *Introduction to Nuclear Power*, 2d ed., 2000; S. Glasstone and A. Sesonske, *Nuclear Reactor Engineering*, 2 vols., 4th ed., 1993; R. L. Murray and J. A. Powell, *Understanding Radioactive Waste*, 4th ed., 1994; F. J. Rahn et al., *A Guide to Nuclear Power Technology*, 1984, reprint 1992; S. C. Stulz and J. B. Kitto, *Steam: Its Generation and Use*, 40th ed., Babcock & Wilcox Co., 1992.

Nuclear quadrupole resonance

A selective absorption phenomenon observable in a wide variety of polycrystalline compounds containing nonspherical atomic nuclei when placed in a magnetic radio-frequency field. Nuclear quadrupole resonance (NQR) is very similar to nuclear magnetic resonance (NMR), and was originated in the late 1940s by H. G. Dehmelt and H. Krüger as an inexpensive (no stable homogeneous large magnetic field is required) alternative way to study nuclear moments. *See* MAGNETIC RESONANCE; NUCLEAR MAGNETIC RESONANCE (NMR).

Principles. In the simplest case, for example, ^{35}Cl in solid Cl_2, NQR is associated with the precession of the angular momentum of the nucleus, depicted in the **illustration** as a flat ellipsoid of rotation, around the symmetry axis (taken as the z axis) of the Cl_2 molecule fixed in the crystalline solid. (The direction

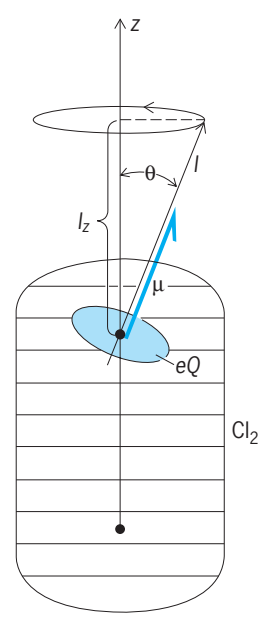

Interaction of ^{35}Cl nucleus with the electric field of a ^{2}Cl molecule.

of the nuclear angular momentum coincides with those of the symmetry axis of the ellipsoid and of the nuclear magnetic dipole moment μ.) The precession, with constant angle θ between the nuclear axis and symmetry axis of the molecule, is due to the torque which the inhomogeneous molecular electric field exerts on the nucleus of electric quadrupole moment eQ. This torque corresponds to the fact that the electrostatic interaction energy of the nucleus with the molecular electric field depends on the angle θ. The interaction energy is given by Eq. (1), where ρ

$$E = \int \phi\rho \, dV \tag{1}$$

is the nuclear charge density distribution and ϕ is the potential of the molecular electric field. Its dependence on θ is given by Eq. (2), where ϕ_{zz} is the

$$E = \frac{eQ\phi_{zz}(3\cos^2\theta - 1)}{8} \tag{2}$$

axial gradient of the (approximately) axially symmetric molecular electric field. The quantum-mechanical analog of this expression is Eq. (3), where I and m

$$E_m = \frac{eQ\phi_{zz}[3m^2 - I(I+1)]}{4I\,(2I-1)} \tag{3}$$

denote the quantum numbers of the nuclear angular momentum and its z component I_z. The absorption occurs classically when the frequency of the rf field v and that of the processing motion of the angular momentum coincide, or quantum-mechanically when Eq. (4) is satisfied, where m and m' are given by Eqs. (5) and $m' - m = +1$, corresponding to the

$$h\nu = |E_{m'} - E_m| \tag{4}$$

$$m, m' = 0 \pm 1, \pm 2 \ldots \pm I \quad \text{for integer } I \geq 1 \tag{5}$$
$$m, m' = \pm{}^1/_2, \pm{}^3/_2 \ldots \pm I \text{ for half} - \text{integer } I \geq {}^3/_2$$

magnetic dipole transitions. *See* NUCLEAR MOMENTS.

It is not necessary that the rf field direction is perpendicular to z; a nonvanishing perpendicular component suffices. This eliminates the necessity of using single crystals and makes it practical, unlike in the NMR of solids, to use polycrystalline samples of unlimited mass and volume. In fact, a polycrystalline natural sulfur sample of 3.18 quarts (3 liters) volume was used in work on the rare (0.74% abundance) ^{33}S isotope.

Techniques. The original NQR work was done on approximately 3 in.^{3} (50 cm^3) of frozen *trans*-dichloroethylene submerged in liquid air using a superregenerative detector. The oscillator incorporated a vibrating capacitor driven by the power line to sweep a frequency band about 50 kHz wide over the approximately 10-kHz-wide ^{35}Cl and ^{37}Cl resonances near 30 MHz, and an oscilloscopic signal display was used. This work demonstrated the good sensitivity and rugged character of these simple, easily tunable circuits which are capable of combining high rf power levels with low noise. Their chief disadvantage is the occurrence of side bands spaced by the quench frequency which may confuse the line shape. For nuclear species of low abundance it becomes important to use nuclear modulation. In the ^{33}S work zero-based magnetic-field pulses periodically smearing out the absorption line proved satisfactory.

Application. NQR spectra have been observed in the approximate range 1–1000 MHz. Such a range clearly requires more than one spectrometer. Most of the NQR work has been on molecular crystals. While halogen-containing (Cl, Br, I) organic compounds have been in the forefront since the inception of the field, NQR spectra have also been observed for K, Rb, Cs, Cu, Au, Ba, Hg, B, Al, Ga, In, La, N, As, Sb, Bi, S, Mn, Re, and Co isotopes. For molecular crystals the coupling constants $eQ\phi_{zz}$ found do not differ very much from those measured for the isolated molecules in microwave spectroscopy. The most precise nuclear information which may be extracted from NQR $eQ\phi_{zz}$ data are quadrupole moment ratios of isotopes of the same element, since one may assume that ϕ_{zz} is practically independent of the nuclear mass. As far as ϕ_{zz} values may be estimated from atomic fine structure data, for example, for Cl_2 where a pure p-bond is expected and the molecular nature of the solid is suggested by a low boiling point and so forth, fair Q values may be obtained. However, it has also proved very productive to use the quadrupole nucleus as a probe of bond character and orientation and crystalline electric fields and lattice sites, and a large body of data has been accumulated in this area. *See* MICROWAVE SPECTROSCOPY. Hans Dehmelt

Bibliography. I. P. Biryukov, M. G. Voronkov, and I. A. Safin, *Tables of Nuclear Quadrupole Resonance Frequencies*, 1969; H. Chihara and N. Nakamupa, *Landolt-Bornstein Numerical Data and Functional Relationships in Science and Technology*, vol. 31, subvol. B: *Nuclear Quadrupole Resonance Spectroscopy Data*, 1993; T. P. Das and E. L. Hahn, *Nuclear Quadrupole Resonance Spectroscopy*, 1958; H. G. Dehmelt, Nuclear quadrupole resonance (in solids), *Amer. J. Phys.*, 22:110–120, 1954, and *Faraday Soc. Discuss.*, 19:263–274, 1955; G. K. Semin, T. A. Babushkina, and G. G. Yakobson, *Nuclear Quadrupole Resonance in Chemistry*, 1975; J. A. S. Smith (ed.), *Advances in Nuclear Quadrupole Resonance*, vols. 1–5, 1974–1983.

Nuclear radiation

All particles and radiations emanating from an atomic nucleus due to radioactive decay and nuclear reactions. Thus the criterion for nuclear radiations is that a nuclear process is involved in their production. The term was originally used to denote the ionizing radiations observed from naturally occurring radioactive materials. These radiations were alpha particles (energetic helium nuclei), beta particles (negative electrons), and gamma rays (electromagnetic radiation with wavelength much shorter than visible light). *See* BETA PARTICLES; GAMMA RAYS.

Nuclear radiations have traditionally been considered to be of three types based on the manner in which they interact with matter as they pass through

it. These are the charged heavy particles with masses comparable to that of the nuclear mass (for example, protons, alpha particles, and heavier nuclei), electrons (both negatively and positively charged), and electromagnetic radiation. For all of these, the interactions with matter are considered to be primarily electromagnetic. (The neutron, which is also a nuclear radiation, behaves quite differently.) The behavior of mesons and other particles is intermediate between that of the electron and heavy charged particles. *See* CHARGED PARTICLE BEAMS.

A striking difference in the absorption of the three types of radiations is that only heavy charged particles have a range. That is, a monoenergetic beam of heavy charged particles, in passing through a certain amount of matter, will lose energy without changing the number of particles in the beam. Ultimately, they will be stopped after crossing practically the same thickness of absorber. The minimum amount of absorber that stops a particle is its range. The greatest part of the energy loss results from collisions with atomic electrons, causing the electrons to be excited or freed. The energy loss per unit path length is the specific energy loss, and its average value is the stopping power of the absorbing substance.

For electromagnetic radiation (gamma rays) and neutrons, on the other hand, the absorption is exponential; that is, the intensity decreases in such a way that the equation below is valid, where I is the in-

$$-\frac{dI}{I} = \mu\, dx$$

tensity of the primary radiation, μ is the absorption coefficient, and dx is the thickness traversed. The difference in behavior reflects the fact that charged particles are not removed from the beam by individual interactions, whereas gamma radiation photons (and neutrons) are. Three main types of phenomena involved in the interaction of electromagnetic radiation with matter (namely, photoelectric absorption, Compton scattering, and electron-positron production) are responsible for this behavior.

Electrons exhibit a more complex behavior. They radiate electromagnetic energy easily because they have a large charge-to-mass ratio and hence are subject to violent acceleration under the action of the electric forces. Moreover, they undergo scattering to such an extent that they follow irregular paths. *See* ELECTRON.

Whereas in the case of the heavy charged particles, electrons, or gamma rays the energy loss is mostly due to electromagnetic effects, neutrons are slowed down by nuclear collisions. These may be inelastic collisions, in which a nucleus is left in an excited state, or elastic collisions, in which the colliding nucleus acquires part of the energy (of the order of 1 MeV) to excite the collision partner. With less kinetic energy, only elastic scattering can slow down the neutron, a process which is effective down to thermal energies (approximately 1/40 keV). At this stage the collision, on the average, has no further effect on the energy of the neutron. *See* NEUTRON.

As noted previously, the other nuclear radiations such as mesons have behaviors which are intermediate between that of heavy charged particles and electrons. Another radioactive decay product is the neutrino; because of its small interaction with matter, it is not ordinarily considered to be a nuclear radiation. *See* MESON; NEUTRINO; NUCLEAR REACTION.

Dennis G. Kovar

Nuclear radiation (biology)

Nuclear radiations are used in biology because of their common property of ionizing matter. This makes their detection relatively simple, or makes possible the production of biological effects in any living cell. Nuclear radiations originate in atomic nuclei, either spontaneously, as in radioactive substances, or through interactions with neutrons, photons, and so on. Gamma radiation originates in atomic nuclei and constitutes one kind of nuclear radiation, but it is otherwise indistinguishable in its effects from x-radiation produced by extranuclear reactions. Because x-rays have been readily available for many years, they have been used more extensively in biology and medicine than gamma rays. Therefore, x-rays must be included in any discussion of the biological and medical uses of nuclear radiations.

Ionizing radiation. Ionizing radiation is any electromagnetic or particulate radiation capable of producing ions, directly or indirectly, in its passage through matter.

Electromagnetic radiations. X-rays and gamma rays are electromagnetic radiations, traveling at the speed of light as packages of energy called photons. They ionize indirectly by first ejecting electrons at high speed from the atoms with which they interact; these secondary electrons then produce most of the ionization associated with the primary radiation. *See* GAMMA RAYS; PHOTON; X-RAYS.

Particulate radiation. Fast neutrons are particulate radiation consisting of nuclear particles of mass number 1 and zero charge, traveling at high speed. They ionize indirectly, largely by setting in motion charged particles from the atomic nuclei with which they collide. Slow or thermal neutrons ionize indirectly by interacting with nuclei, in a process known as neutron capture, to produce ionizing radiation. *See* NEUTRON.

Alpha rays are particulate radiation consisting of helium nuclei traveling at high speed. Since they are charged particles, they ionize directly. Alpha particles are emitted spontaneously by some radioactive nuclides or may result from neutron capture; for example, neutron capture by boron-10 produces lithium-7 and an alpha particle. The energy of alpha particles emitted by radioactive substances is of the order of a few megaelectronvolts, but alpha particles of very much higher energy may be produced in cyclotrons or other particle accelerators from helium-ion beams. With such machines, other ionizing particles of very high energy, such as protons, deuterons, and so on, may also be produced.

Beta rays are particulate radiation consisting of electrons or positrons emitted from a nucleus during beta decay and traveling at high speed. Since they are charged particles, that is, − or +, they ionize directly. Electron beams of very high energy may be produced by high-voltage accelerators, but in that case they are not called beta particles. A pair consisting of one electron and one positron may be formed by one high-energy (1.022-MeV) photon where it traverses a strong electric field, such as that surrounding a nucleus or an electron. Subsequently, the positron and another electron react, and their mass is transformed into energy in the form of two photons traveling in opposite directions. This is called the annihilation process and is the inverse of the pair-production process that is mentioned above. Ionizing radiations, such as protons, deuterons, alpha particles, and neutrons, may be produced simultaneously by spallation when a very high-energy particle collides with an atom. In a photographic emulsion or in a cloud chamber, the ionizing particles originating from a common point form stars. *See* BETA PARTICLES.

Fission occurs in certain heavy nuclei spontaneously or through interaction with neutrons, charged particles, or photons, and it results in the division of the nucleus into two approximately equal parts. These fragments produced by fission are endowed with very large amounts of kinetic energy, carry large positive charges, and produce very dense ionization in their short passage through matter. *See* NUCLEAR FISSION.

Primary cosmic rays probably consist of atomic nuclei, mainly protons, with extremely high energies which interact with nuclei and electrons in the atmosphere and produce secondary cosmic rays, consisting mainly of mesons, protons, neutrons, electrons, and photons of lower energy. *See* COSMIC RAYS; MESON; PROTON.

Biological effects. All ionizing radiations produce biological changes, directly by ionization or excitation of the atoms in the molecules of biological entities, such as in chromosomes, or indirectly by the formation of active radicals or deleterious agents, through ionization and excitation, in the medium surrounding the biological entities. Ionizing radiation, having high penetrating power, can reach the most vulnerable part of a cell, an organ, or a whole organism, and is thus very effective. In terms of the energy absorbed per unit mass of a biological entity in which an effect is produced, some ionizing radiations are more effective than others. The relative biological effectiveness (RBE) depends in fact on the density of ionization (also termed the specific ionization or linear energy transfer, LET) along the path of the ionizing particle rather than on the nature of the particle itself. Relative biological effectiveness depends also on many other factors. *See* LINEAR ENERGY TRANSFER (BIOLOGY); RADIATION BIOLOGY.

Use in medicine. The medical uses of nuclear radiations may be divided into three distinct classes:

1. The radiations, which are principally x-rays, are used to study the anatomical configuration of body organs, usually for the purpose of detecting abnormalities as an aid in diagnosis.

2. The radiations are used for therapeutic purposes to produce biological changes in such tissues as tumors.

3. The radiations are used as a simple means of tracing a suitable radioactive substance through different steps in its course through the body, in the study of some particular physiological process. *See* RADIOLOGY.

Gioacchino Failla; Edith H. Quimby

Use in biological research. The radiations emitted by radioactive isotopes of the various elements are used in biological research. The most useful ones in biological research are the isotopes of the elements which are important in metabolism and in the structural materials of cells. These include carbon, hydrogen, sulfur, and phosphorus. Unfortunately, nitrogen and oxygen do not have usable radioisotopes. Isotopes of calcium, iodine, potassium, sodium, iron, and a few others have more limited usefulness in biological research. In addition, the radioactive metals like cobalt-60, radium, and others can be used to produce radiations for external application to cells and tissues. *See* RADIOISOTOPE (BIOLOGY).

Most of the isotopes mentioned emit beta particles when they decay, and a few emit gamma rays. Therefore, they can be easily detected by various means. If the radiations emitted are highly penetrating, like the gamma rays and the high energy beta particles from phosphorus-32, the presence of the isotope may be detected with Geiger counters or scintillation counters applied to the surface of the biological material. Likewise, application of photographic emulsions to the surface or to cut surfaces of cells or tissues may serve to locate the isotope. When the biological material can be broken down and destroyed, particular compounds or larger components of cells may be isolated, and the isotope determined by the various types of radiation detectors. These procedures are sometimes used in the study of the movement of elements or their compounds in plants and animals. They are frequently used for tracing the sequence of reactions in metabolism. The great advantage of radioisotopes for such studies is that small amounts or concentrations may be readily detected and measured, either in the presence of other isotopes of the same element or when mixed with a variety of other elements. *See* AUTORADIOGRAPHY.

In addition to the radiations emitted, some of the elements change to a different element when they decay. Phosphorus-32 changes to sulfur, sulfur-35 changes to chlorine, and tritium (hydrogen-3) changes to helium when they decay by the emission of an electron, a beta particle. Therefore, in addition to the radiation produced, the transmutation of the element affects the molecule and the cell of which it is a part. Attempts have been made to evaluate the effects of the two factors by giving the fungus *Neurospora* the same amount of radiation, and by having different proportions of the radioactive isotope incorporated into the molecules of the cell. The incorporation of the isotope was regulated by having the same concentration of the radioisotope in two cultures, but in one, the radiosulfur or radiophosphorus

was diluted greatly with the nonradioactive isotope. A difference in lethal effect was demonstrated.

In other experiments, the decay of phosphorus-32 has been used to give information on the nature and importance of the phosphorus-containing molecules to the survival or reproduction of a cell or virus particle. When bacterial viruses are allowed to grow in the presence of phosphate containing phosphorus-32, they incorporate the radioisotope into their genetic material, deoxyribonucleic acid (DNA). If these virus particles are then stored under conditions in which no growth occurs, the decay of the radioisotope is very effective in inactivating the viruses. About 1 in 10 atoms which decay inactivates 1 virus particle. From such experiments, biologists are able to learn something about the importance of the molecule as well as something of its size and possible organization. Similar experiments have been carried out with the cells of ameba, but the lethal effect was observed, not on the stored cells, but on their offspring. From this experiment, deductions concerning the organization of genetic material and its mutability were drawn. *See* DEOXYRIBONUCLEIC ACID (DNA).

Nuclear radiations have also proved useful in many studies of the nature and the mechanisms of the effects of radiations on cells and cell constituents. The radiations with low penetrating power, for example, alpha particles and low-energy beta particles, can be most effective when an element which will emit the particles is placed inside the cells to be studied. Radon, which is a gas, can be used in this way for the production of alpha particles. Likewise, a variety of the heavy metals like thorium, uranium, and polonium emit alpha particles. Various beta emitters, such as phosphorus-32, carbon-14, sulfur-35, and tritium, can also be used for producing radiations inside the cell. One of the most interesting of this group is tritium, which emits very soft beta particles. Their maximum range is 6 micrometers in water and about the same in tissues, but the average range is much less, about 1 micrometer or $^1/_{25}$ the diameter of a medium-sized cell. Tritium can be put into the cell as water or in many other compounds. Many of these are unselective and are equally distributed to all parts of the cell. However, there is one substance, thymidine, which selectively labels the DNA which is restricted to the cell nucleus. If the cell is a relatively large one compared to its nuclear size, nearly all of the radiation will be absorbed by the nucleus, while the other part of the cell is irradiated hardly at all. Experiments have shown that tritium given to cells in the form of tritium-thymidine is about 1000 times as effective as tritium-water in producing radiation effects. *See* HISTORADIOGRAPHY; SCINTILLATION COUNTER.

J. Herbert Taylor

Nuclear reaction

A process that occurs as a result of interactions between atomic nuclei when the interacting particles approach each other to within distances of the order of nuclear dimensions ($\sim 10^{-12}$ cm). While nuclear reactions occur in nature, understanding of them and use of them as tools have taken place primarily in the controlled laboratory environment. In the usual experimental situation, nuclear reactions are initiated by bombarding one of the interacting particles, the stationary target nucleus, with nuclear projectiles of some type, and the reaction products and their behaviors are studied. The study of nuclear reactions is the largest area of nuclear and subnuclear (or particle) physics; the threshold for producing pions has historically been taken to be the energy boundary between the two fields.

Types of nuclear interaction. As a generalized nuclear process, consider a collision in which an incident particle strikes a previously stationary particle to produce an unspecified number of final products. If the final products are the same as the two initial particles, the process is called scattering. The scattering is said to be elastic or inelastic, depending on whether some of the kinetic energy of the incident particle is used to raise either of the particles to an excited state. If the product particles are different from the initial pair, the process is referred to as a reaction.

The most common type of nuclear reaction, and the one which has been most extensively studied, involves the production of two final products. Such reactions can be observed, for example, when deuterons with a kinetic energy of a few megaelectronvolts are allowed to strike a carbon nucleus of mass 12. Protons, neutrons, deuterons, and alpha particles are observed to be emitted, and reactions (1)–(4) are responsible. In these equations the nuclei

$$^{2}_{1}\mathrm{H} + ^{12}_{6}\mathrm{C} \rightarrow ^{2}_{1}\mathrm{H} + ^{12}_{6}\mathrm{C} \qquad (1)$$

$$^{2}_{1}\mathrm{H} + ^{12}_{6}\mathrm{C} \rightarrow ^{1}_{1}\mathrm{H} + ^{13}_{6}\mathrm{C} \qquad (2)$$

$$^{2}_{1}\mathrm{H} + ^{12}_{6}\mathrm{C} \rightarrow ^{1}_{0}\mathrm{n} + ^{13}_{7}\mathrm{N} \qquad (3)$$

$$^{2}_{1}\mathrm{H} + ^{12}_{6}\mathrm{C} \rightarrow ^{4}_{2}\mathrm{He} + ^{10}_{5}\mathrm{B} \qquad (4)$$

are indicated by the usual chemical symbols; the subscripts indicate the atomic number (nuclear charge) of the nucleus, and the superscripts the mass number of the particular isotope. These reactions are conventionally written in the compact notation $^{12}\mathrm{C}(d,d)^{12}\mathrm{C}$, $^{12}\mathrm{C}(d,p)^{13}\mathrm{C}$, $^{12}\mathrm{C}(d,n)^{13}\mathrm{N}$, and $^{12}\mathrm{C}(d,\alpha)^{10}\mathrm{B}$, where d represents deuteron, p proton, n neutron, and α alpha particle. In each of these cases the reaction results in the production of an emitted light particle and a heavy residual nucleus. The (d,d) process denotes the elastic scattering as well as the inelastic scattering processes that raise the $^{12}\mathrm{C}$ nucleus to one of its excited states. The other three reactions are examples of nuclear transmutation or disintegration where the residual nuclei may also be formed in their ground states or one of their many excited states. The processes producing the residual nucleus in different excited states are considered to be the different reaction channels of the particular reaction. If the residual nucleus is formed in an excited state, it will subsequently emit this excitation energy in

the form of gamma rays or, in special cases, electrons. The residual nucleus may also be a radioactive species, as in the case of ^{13}N formed in the $^{12}C(d,n)$ reaction. In this case the residual nucleus will undergo further transformation in accordance with its characteristic radioactive decay scheme. *See* GAMMA RAYS; RADIOACTIVITY.

Nuclear cross section. In general, one is interested in the probability of occurrence of the various reactions as a function of the bombarding energy of the incident particle. The measure of probability for a nuclear reaction is its cross section. Consider a reaction initiated by a beam of particles incident on a region which contains N atoms per unit area (uniformly distributed), and where I particles per second striking the area result in R reactions of a particular type per second. The fraction of the area bombarded which is effective in producing the reaction products is R/I. If this is divided by the number of nuclei per unit area, the effective area or cross section $\sigma = R/IN$. This is referred to as the total cross section for the specific reaction, since it involves all the occurrences of the reaction. The dimensions are those of an area, and total cross sections are expressed in either square centimeters or barns (1 barn $= 10^{-24}$ cm^2). The differential cross section refers to the probability that a particular reaction product will be observed at a given angle with respect to the beam direction. Its dimensions are those of an area per unit solid angle (for example, barns per steradian).

Requirements for a reaction. Whether a specific reaction occurs and with what cross section it is observed depend upon a number of factors, some of which are not always completely understood. However, there are some necessary conditions which must be fulfilled if a reaction is to proceed.

Coulomb barrier. For a reaction to occur, the two interacting particles must approach each other to within the order of nuclear dimensions ($\sim 10^{-12}$ cm). With the exception of the uncharged neutron, all incident particles must therefore have sufficient kinetic energy to overcome the electrostatic (Coulomb) repulsion produced by the intense electrostatic field of the nuclear charge. The kinetic energy must be comparable to or greater than the so-called Coulomb barrier, whose magnitude is approximately given by the expression $E_{\mathrm{Coul}} \approx Z_1Z_2/(A_1^{1/3} + A_2^{1/3})$ MeV, where Z and A refer to the nuclear charge and mass number of the interacting particles 1 and 2, respectively. It can be seen that while, for the lightest targets, protons with kinetic energies of a few hundred kiloelectronvolts are sufficient to initiate reactions, energies of many hundreds of megaelectronvolts are required to initiate reactions between heavier nuclei. In order to provide energetic charged particles, to be used as projectiles in reaction studies, particle accelerators of various kinds (such as Van de Graaff generators, cyclotrons, and linear accelerators) have been developed, making possible studies of nuclear reactions induced by projectiles as light as protons and as heavy as ^{208}Pb. *See* ELECTROSTATICS; PARTICLE ACCELERATOR.

Since neutrons are uncharged, they are not repelled by the electrostatic field of the target nucleus, and neutron energies of only a fraction of an electronvolt are sufficient to initiate some reactions. Neutrons for reaction studies can be obtained from nuclear reactors or from various nuclear reactions which produce neutrons as reaction products. There are two other means of producing nuclear reactions which do not fall into the general definition given above. Both electromagnetic radiation and high-energy electrons are capable of disintegrating nuclei under special conditions. However, both interact much less strongly with nuclei than nucleons or other nuclei, through the electromagnetic and weak nuclear forces, respectively, rather than the strong nuclear force responsible for nuclear interactions. *See* NEUTRON.

Q value. For a nuclear reaction to occur, there must be sufficient kinetic energy available to bring about the transmutation of the original nuclear species into the final reaction products. The sum of the kinetic energies of the reaction products may be greater than, equal to, or less than the sum of the kinetic energies before the reaction. The difference in the sums is the Q value for that particular reaction. It can be shown that the Q value is also equal to the difference in the masses (rest energies) of the reaction products and the masses of the initial nuclei. Reactions with a positive Q value are called exoergic or exothermic reactions, while those with a negative Q value are called endoergic or endothermic reactions.

In reactions (1)–(4), where the residual nuclei are formed in their ground states, the Q values are $^{12}C(d,d)^{12}C$, $Q = 0.0$ MeV; $^{12}C(d,p)^{12}C$, $Q = 2.72$ MeV; $^{12}C(d,n)^{13}N$, $Q = -0.28$ MeV; and $^{12}C(d,\alpha)^{10}B$, $Q = -1.34$ MeV. For reactions with a negative Q value, a definite minimum kinetic energy is necessary for the reaction to take place. While there is no threshold energy for reactions with positive Q values, the cross section for the reactions induced by charged particles is very small unless the energies are sufficient to overcome the Coulomb barrier. A nuclear reaction and its inverse are reversible in the sense that the Q values are the same but have the opposite sign [for example, the Q value for the $^{10}B(\alpha,d)^{12}C$ reaction is $+1.39$ MeV].

Conservation laws. It has been found experimentally that certain physical quantities must be the same both before and after the reaction. The quantities conserved are electric charge, number of nucleons, energy, linear momentum, angular momentum, and in most cases parity. Except for high-energy reactions involving the production of mesons, the conservation of charge and number of nucleons allow one to infer that the numbers of protons and neutrons are always conserved. The conservation of the number of nucleons indicates that the statistics governing the system are the same before, during, and after the reaction. Fermi-Dirac statistics are obeyed if the total number is odd, and Bose-Einstein if the number is even. The conservation laws taken together serve to strongly restrict the reactions that can take place, and the conservation of angular momentum and parity in particular allow one to establish spins and parities

of states excited in various reactions. *See* ANGULAR MOMENTUM; CONSERVATION LAWS (PHYSICS); PARITY (QUANTUM MECHANICS); QUANTUM STATISTICS; SYMMETRY LAWS (PHYSICS).

Reaction mechanism. What happens when a projectile collides with a target nucleus is a complicated many-body problem which is still not completely understood. Progress made in the last decades has been in the development of various reaction models which have been extremely successful in describing certain classes or types of nuclear reaction processes. In general, all reactions can be classified according to the time scale on which they occur, and the degree to which the kinetic energy of the incident particle is converted into internal excitation of the final products. A large fraction of the reactions observed has properties consistent with those predicted by two reaction mechanisms which represent the extremes in this general classification. These are the mechanisms of compound nucleus formation and direct interaction.

Compound nucleus formation. As originally proposed by N. Bohr, the process is envisioned to take place in two distinct steps. In the first step the incident particle is captured by (or fuses with) the target nucleus, forming an intermediate or compound nucleus which lives a long time ($\sim 10^{-16}$ s) compared to the approximately 10^{-22} s it takes the incident particle to travel past the target. During this time the kinetic energy of the incident particle is shared among all the nucleons, and all memory of the incident particle and target is lost. The compound nucleus is always formed in a highly excited unstable state, is assumed to approach themodynamic equilibrium involving all or most of the available degrees of freedom, and will decay, as ths second step, into different reaction products, or through so-called exit channels. In most cases, the decay can be understood as a statistical evaporation of nucleons or light particles. In the examples of reactions (1)–(4), the compound nucleus formed is ^{14}N, and four possible exit channels are indicated (**Fig. 1**). In reactions involving heavier targets (for example, $A \approx 200$), one of the exit channels may be the fission channel where the compound nucleus splits into two large fragments. *See* NUCLEAR FISSION.

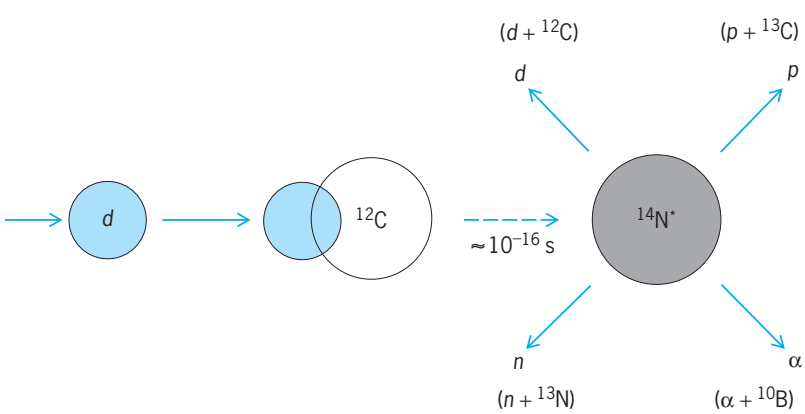

Fig. 1. Formation of the compound nucleus ^{14}N after capture of the deuteron by ^{12}C. Four exit channels are indicated.

The essential feature of the compound nucleus formation or fusion reaction is that the probability for a specific reaction depends on two independent probabilities: the probability for forming the compound nucleus, and the probability for decaying into that specific exit channel. While certain features of various interactions cannot be completely explained within the framework of the compound nucleus hypothesis, it appears that the mechanism is responsible for a large fraction of reactions occurring in almost all projectile-target interactions. Fusion reactions have been extremely useful in several kinds of spectroscopic studies. Particularly notable have been the resonance studies performed with light particles, such as neutrons, protons, deuterons, and alpha particles, on light target nuclei, and the gamma-ray studies of reactions induced by heavy projectiles, such as ^{16}O and ^{32}S, on target nuclei spanning the periodic table. These studies have provided an enormous amount of information regarding the excitation energies and spins of levels in nuclei. *See* NUCLEAR SPECTRA.

Direct interactions. Some reactions have properties which are in striking conflict with the predictions of the compound nucleus hypothesis. Many of these are consistent with the picture of a mechanism where no long-lived intermediate system is formed, but rather a fast mechanism where the incident particle, or some portion of it, interacts with the surface, or some nucleons on the surface, of the target nucleus. Models for direct processes make use of a concept of a homogeneous lump of nuclear matter with specific modes of excitation, which acts to scatter the incident particle through forces described, in the simplest cases, by an ordinary spherically symmetric potential. In the process of scattering, some of the kinetic energy may be used to excite the target, giving rise to an inelastic process, and nucleons may be exchanged, giving rise to a transfer process. In general, however, direct reactions are assumed to involve only a very small number of the available degrees of freedom.

Most direct reactions are of the transfer type where one or more nucleons are transferred to or from the incident particle as it passes the target, leaving the two final partners either in their ground states or in one of their many excited states. Such transfer reactions are generally referred to as stripping or pick-up reactions, depending on whether the incident particle has lost or acquired nucleons in the reaction. The (d,p) reaction is an example of a stripping reaction, where the incident deuteron is envisioned as being stripped of its neutron as it passes the target nucleus, and the proton continues along its way (**Fig. 2**).

The properties of the target nucleus determine the details of the reaction, fixing the energy and angular momentum with which the neutron must enter it. The energy of the outgoing proton is determined by how much of the deuteron's energy is taken into the target nucleus by the neutron, and indeed serves to identify the final state populated by the Q value of the reaction as a whole. The angular distribution of the differential cross sections will, at appropriate bombarding energies, not be smooth but rather will

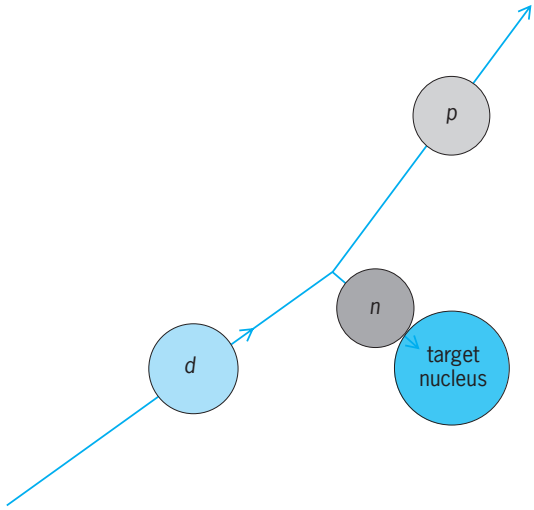

Fig. 2. A (*d*, *p*) transfer reaction.

show a distinct pattern of maxima and minima which are indicative of the spin and parity of the final state. The cross section for populating a specific final state in the target nucleus depends on the nuclear structure of the nuclei involved. This sensitivity has been used in studies of single-nucleon, two-nucleon, and four-nucleon transfer reactions, such as (d,p), (p,d), $(^3\mathrm{He},d)$, (t,p), and $(^7\mathrm{Li},d)$, to establish the validity and usefulness of the shell-model description of nuclei. Multinucleon transfer reactions with heavier projectiles have been powerful tools for reaching nuclei inaccessible by other means and in producing new isotopes.

Inelastic scattering is also a direct reaction whose angular distribution can provide information about the spin and parity of the excited state. Whereas the states preferentially populated in transfer reactions are those of specific single-particle or shell-model structure, the states preferentially excited in inelastic scattering are collective in nature. The states are most easily understood in the framework of macroscopic descriptions in which they are considered to be oscillations in shape about a spherical mean (vibrations) or to be the rotations of a statically deformed shape. The cross section for inelastic scattering is related to the shape or deformation of the target nucleus in its various collective excitations. Inelastic excitation can be caused by both a nuclear interaction and an electromagnetic interaction known as Coulomb excitation, where the target nucleus interacts with the rapidly changing electric field caused by the passage of the charged particle. Coulomb excitation is an important process at low bombarding energies in interactions involving heavy projectiles. Studies in inelastic scattering to low-lying states of nuclei across the periodic table have provided graphic demonstrations of the collective aspects of nuclei. *See* COULOMB EXCITATION.

Elastic scattering is the direct interaction which leaves the interacting particles unchanged. For charged particles at low bombarding energies, the elastic scattering is well described in terms of the inverse-square force law between two electrically charged bodies. In this case, the process is known as Rutherford scattering. At higher energies the particles come into the range of the nuclear force, and the elastic scattering deviates from the inverse-square behavior. *See* NUCLEAR STRUCTURE; SCATTERING EXPERIMENTS (NUCLEI).

More complex reaction mechanisms. Processes intermediate between direct and compound nucleus formation do occur. The best example of such a process is the so-called preequilibrium emission, where light particles are emitted before the kinetic energy has been shared among all the nucleons in the compound nucleus. Another example is seen in the interaction of two massive nuclei, such as $^{84}\mathrm{Kr} + {}^{209}\mathrm{Bi}$, where the probability for formation of the compound nucleus is found to be very small. The experimental observations indicate that the nuclei interact for a short time and then separate in what appears to be a direct reaction. Although the interaction times for these so-called strongly damped or deep inelastic collisions are very small compared to that for compound nucleus formations, they are sufficiently long for considerable mass transfer and loss of relative kinetic energy to occur.

Nuclear reaction studies. In most instances the study of nuclear reactions is directed toward the long-range goal of obtaining information about the properties and structure of nuclei. Such studies usually proceed in two stages. In the first stage the attention focuses on the mechanism of the reaction and on establishing the dependence on the nuclei involved. The specific models proposed continue to be modified and improved, as they are confronted with more experimental results, until their predictions become reliable. At that point the second stage is entered, in which the focus of the effort is on extraction of information about nuclei.

There are other studies which are focused on reaction cross-section behaviors for other purposes. Examples of this are the neutron-capture reactions on heavy target nuclei that fission, and the $^3\mathrm{H}(d,n)^4\mathrm{He}$ reaction, important in thermonuclear processes, which continue to be studied because of their application as energy sources. Studies of the interactions of light nuclei at low energies are intensely studied because of their implications for astrophysics and cosmology. *See* NUCLEAR FUSION; NUCLEOSYNTHESIS; THERMONUCLEAR REACTION. Dennis G. Kovar

Collisions of very high energy nuclei. The protons and neutrons which make up nuclei are not elementary particles. They are each clusters of quarks, tightly bound together by the exchange of gluons. One central characteristic of quantum chromodynamics, the theory which describes the interactions of quarks and gluons, is that the force between quarks becomes weaker as the quarks interact at higher energies. During most nuclear reactions, the force between quarks is so strong that free quarks cannot exist; quarks only occur confined in composite particles called hadrons. Protons and neutrons are hadrons, and there are many other types, such as pions and J/ψ particles, which are unstable. However, at temperatures above a critical temperature value

of about 150 MeV (about 10^{12} K, or about 10^5 times the temperature at the center of the Sun), quantum chromodynamics predicts that the force between the quarks will be so weakened that hadrons will fall apart. The transition between hadronic matter and a plasma of quarks and gluons is called the quantum chromodynamics phase transition.

For the first few microseconds after the big bang, the whole universe was hotter than 150 MeV and was filled with quark-gluon plasma. The quantum chromodynamics phase transition occurred throughout the universe as it cooled through the critical temperature. The extreme temperatures required to observe the quantum chromodynamics phase transition are difficult to obtain in a laboratory. The only method known is to collide large nuclei at very high energies. *See* BIG BANG THEORY; PARTICLE ACCELERATOR.

Detailed theoretical calculations which simulate the quantum chromodynamics interactions of the thousands of quarks and gluons produced in such collisions lend support to the expectation that under these conditions quark-gluon plasma will be created for the first time since the big bang. This plasma will occur in a region several times 10^{-13} cm in size, and will exist for several times 10^{-22} s before the quantum chromodynamics phase transition occurs and ordinary hadrons are reformed.

The distribution and energy of the particles emerging from the collision should carry signatures of the fleeting presence of a quark-gluon plasma. One such signature involves the J/ψ particle, which is made of a heavy quark and a heavy antiquark. In a quark-gluon plasma, the J/ψ (like any hadron) dissociates and the heavy quark and antiquark separate. When the plasma cools through the phase transition, heavy quarks are much more likely to form hadrons in combination with the far more numerous light quarks than to reunite to form J/ψ hadrons. The expected reduction in the number of J/ψ particles produced in a collision has in fact been detected. However, J/ψ hadrons can also be broken apart by repeated collisions in a hot, dense hadronic environment. Therefore, the observed J/ψ suppression establishes that such an environment has been created, but does not in itself demonstrate the creation of a quark-gluon plasma. *See* J/PSI PARTICLE.

In an idealized model, as the hadronic state of matter is reestablished after an abrupt quantum chromodynamics phase transition, pion waves are amplified and coherent emission of pions occurs. Such pion lasers would be easily observable, because they would occasionally emit only electrically neutral pions or only electrically charged ones, whereas ordinarily different types of pions are uniformly mixed. This effect would be an unambiguous signature for the brief presence of a quark-gluon plasma in a collision of very high energy nuclei. *See* GLUONS; PHASE TRANSITIONS; QUANTUM CHROMODYNAMICS; QUARK-GLUON PLASMA; QUARKS; RELATIVISTIC HEAVY-ION COLLISIONS. Krishna Rajagopal

Bibliography. H. Feshbach, *Theoretical Nuclear Physics*, vol. 2: *Nuclear Reactions*, 1992; R. Hwa (ed.), *Quark-Gluon Plasma*, 1990; S. Mukherjee, M. K. Pal, and R. Shyam (eds.), *Nuclear Reaction Mechanisms*, 1989; K. Rajagopal and F. Wilczek, Emergence of long wavelength oscillations after a quench: Application to QCD, *Nucl. Phys.*, B399:395–425, 1993; G. R. Satchler, *Introduction to Nuclear Reactions*, 2d ed., 1990; A. G. Sitenko and O. D. Kocherga (eds.), *Theory of Nuclear Reactions*, 1990.

Nuclear reactor

A system utilizing nuclear fission in a controlled and self-sustaining manner. Neutrons are used to fission the nuclear fuel, and the fission reaction produces not only energy and radiation but also additional neutrons. Thus a neutron chain reaction ensues. A nuclear reactor provides the assembly of materials to sustain and control the neutron chain reaction, to appropriately transport the heat produced from the fission reactions, and to provide the necessary safety features to cope with the radiation and radioactive materials produced by its operation. *See* CHAIN REACTION (PHYSICS); NUCLEAR FISSION.

Nuclear reactors are used in a variety of ways as sources for energy, for nuclear irradiations, and to produce special materials by transmutation reactions. Since the first demonstration of a nuclear reactor, made beneath the west stands of Stagg Field at the University of Chicago on December 2, 1942, many hundreds of nuclear reactors have been built and operated in the United States. Extreme diversification is possible with the materials available, and reactor power may vary from a fraction of a watt to thousands of megawatts. The size of a nuclear reactor core is governed by its power level, time between refueling, fuel environment, the factors affecting the control of the neutron chain reaction, and the effectiveness of the coolant in removing the fission energy released.

The generation of electrical energy by a nuclear power plant makes use of heat to produce steam or to heat gases to drive turbogenerators. Direct conversion of the fission energy into useful work is possible, but an efficient process has not yet been realized to accomplish this. Thus, in its operation the nuclear power plant is similar to the conventional coal-fired plant, except that the nuclear reactor is substituted for the conventional boiler as the source of heat.

The rating of a reactor is usually given in kilowatts (kW) or megawatts-thermal [MW(th)], representing the heat generation rate. The net output of electricity of a nuclear plant is about one-third of the thermal output. Significant economic gains have been achieved by building improved nuclear reactors with outputs of about 3300 MW(th) and about 1000 MW-electrical [MW(e)]. *See* ELECTRIC POWER GENERATION; NUCLEAR POWER.

Fuel and Moderator

The fission neutrons are released at high energies and are called fast neutrons. The average kinetic energy is 2 MeV, with a corresponding neutron speed of $^1/_{15}$

the speed of light. Neutrons slow down through collisions with nuclei of the surrounding material. This slowing-down process is made more effective by the introduction of materials of low atomic weight, called moderators, such as heavy water (deuterium oxide), ordinary (light) water, graphite, beryllium, beryllium oxide, hydrides, and organic materials (hydrocarbons). Neutrons that have slowed down to an energy state in equilibrium with the surrounding materials are called thermal neutrons, moving at 0.0006% of the speed of light. The probability that a neutron will cause the fuel material to fission is greatly enhanced at thermal energies, and thus most reactors utilize a moderator for the conversion of fast neutrons to thermal neutrons. This permits using smaller amounts and lower concentrations of fissile materials. *See* NEUTRON; THERMAL NEUTRONS.

With suitable concentrations of the fuel material, neutron chain reactions also can be sustained at higher neutron energy levels. The energy range between fast and thermal is designated as intermediate. Fast reactors do not have moderators and are relatively small.

Reactors have been built in all three categories. The first fast reactor was the Los Alamos (New Mexico) assembly called Clementine, which operated from 1946 to 1953. The fuel core consisted of nickel-coated rods of pure plutonium metal, contained in a 6-in.-diameter (15-cm) steel pot. Coolants for fast reactors may be steam, gas, or liquid metals. Current fast reactors utilize liquid sodium as the coolant and are being developed for breeding and power. An example of an intermediate reactor was the first propulsion reactor for the submarine USS *Seawolf*. The fuel core consisted of enriched uranium with beryllium as a moderator; the original coolant was sodium, and the reactor operated from 1956 to 1959. Examples of thermal reactors, currently the basis of commercial nuclear power production, are given below.

Fuel composition. Only three isotopes—uranium-235, uranium-233, and plutonium-239—are feasible as fission fuels, but a wide selection of materials incorporating these isotopes is available.

Uranium-235. Naturally occurring uranium contains only 0.7% of the fissionable isotope uranium-235, the balance being essentially uranium-238. Uranium with higher concentrations of uranium-235 is called enriched uranium.

Uranium metal is susceptible to irradiation damage, which limits its operating life in a reactor. The life expectancy can be improved somewhat by heat treatment, and considerably more by alloying with elements such as zirconium or molybdenum. Uranium oxide exhibits better irradiation damage resistance and, in addition, is corrosion-resistant in water. Ceramics such as uranium oxide have lower thermal conductivities and lower densities than metals, which are disadvantageous in certain applications. *See* ALLOY.

Current light-water-cooled nuclear power reactors utilize uranium oxide as a fuel, with an enrichment of several percent uranium-235. Cylindrical rods are the most common fuel-element configuration. They can be fabricated by compacting and sintering cylindrical pellets which are then assembled into metal tubes which are sealed.

Developmental programs for attaining long-lived solid-fuel elements include studies with uranium oxide, uranium carbide, and other refractory uranium compounds. *See* URANIUM.

Plutonium-239. Plutonium-239 is produced by neutron capture in uranium-238. It is a by-product in power reactors and is becoming increasingly available as nuclear power production increases. For example, Japan has produced over 10 tons of plutonium solely as the result of its commercial reactor program. However, plutonium as a commercial fuel is still at a demonstration stage of development in Europe and Japan. The commercial recycle of plutonium from processed spent fuel was deferred indefinitely in the United States by the Carter administration in April 1977.

Plutonium is hazardous to handle because of its biological toxicity and, like radium, is a carcinogen if ingested. It is fabricated in glove boxes to ensure isolation from operating personnel. It can be alloyed with other metals and fabricated into various ceramic compounds. It is normally used in conjunction with uranium-238; alloys of uranium-plutonium, and mixtures of uranium-plutonium oxides and carbides, are of most interest as fuels. Except for the additional requirements imposed by plutonium toxicity and proliferation safeguards, much of the uranium technology is applicable to plutonium. For light-water nuclear power reactors, the oxide fuel pellets are contained in a zirconium alloy (Zircaloy) tube. Stainless steel tubes are used for containing the oxide fuel for the fast breeder reactors. *See* PLUTONIUM.

Uranium-233. Uranium-233, like plutonium, does not occur naturally, but is produced by neutron absorption in thorium-232, a process similar to that by which plutonium is produced from uranium-238. Interest in uranium-233 arises from its favorable nuclear properties and the abundance of thorium. However, studies of this fuel cycle are at a relatively early stage. Uranium-233 also imposes special handling problems because of proliferation concerns and the radiological toxicity of the daughter products of another uranium isotope (uranium-232) present in the fuel cycle. It does not, however, introduce new metallurgical problems. Thorium is metallurgically different, but it has favorable properties both as a metal and as a ceramic. *See* NUCLEAR FUELS; THORIUM.

Fuel configurations. Fuel-moderator assemblies may be homogeneous or heterogeneous. Homogeneous assemblies include the aqueous-solution-type water boilers and molten-salt-solution dispersions, slurries, and suspensions. The few homogeneous reactors built have been used for limited research and for demonstration of the principles and design features. In heterogeneous assemblies, the fuel and moderator form separate solid or liquid phases, such as solid-fuel elements spaced either in a graphite matrix or in a water phase. Most power reactors utilize an arrangement of closely spaced, solid fuel rods, about $^1/_2$ in. (13 mm) in diameter and 12 ft (3.7 m)

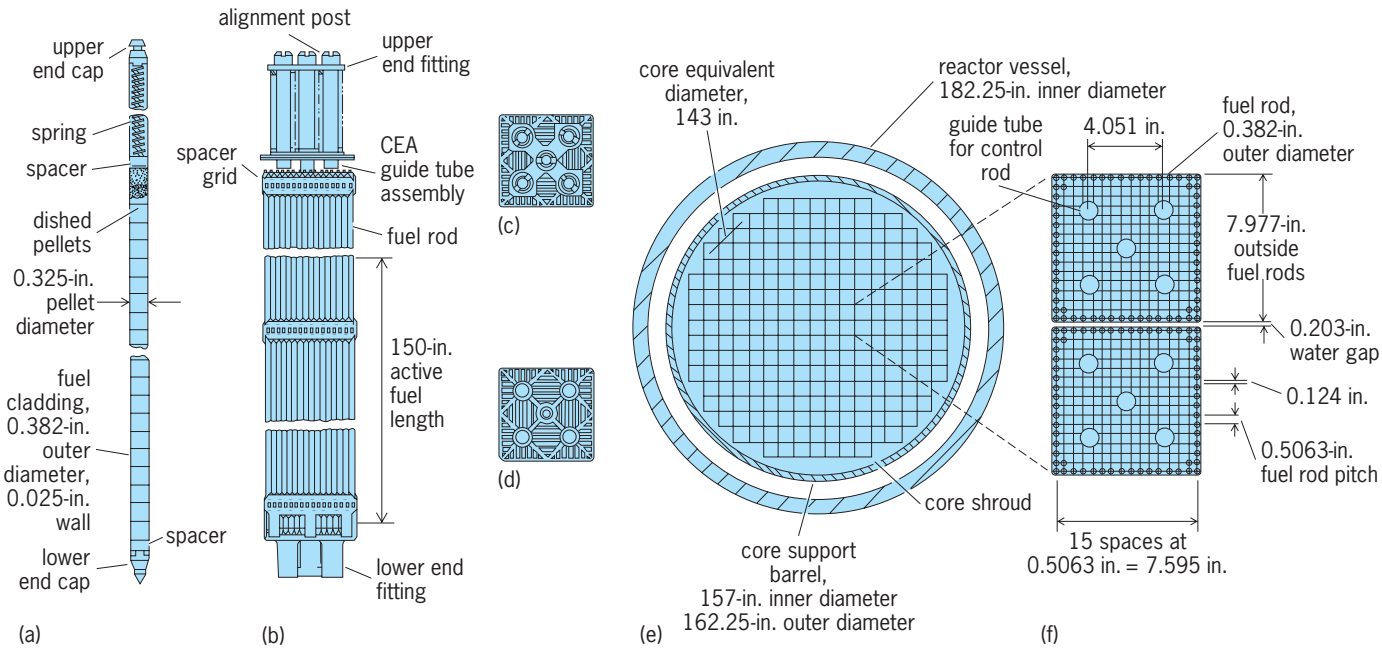

Fig. 1. Arrangement of fuel in the core of a pressurized-water reactor, a typical heterogeneous reactor. (*a*) Fuel rod; (*b*) side view (CEA = control element assembly), (*c*) top view, and (*d*) bottom view of fuel assembly; (*e*) cross section of reactor core showing arrangement of fuel assemblies; (*f*) cross section of two adjacent fuel assemblies, showing arrangement of fuel rods. 1 in. = 25 mm. (*Combustion Engineering, Inc.*)

long, in water. In the arrangement shown in **Fig. 1**, fuel rods are arranged in a grid pattern to form a fuel assembly, and over 200 fuel assemblies are in turn arranged in a grid pattern in the reactor core.

Heat Removal

The major portion of the energy released by the fissioning of the fuel is in the form of kinetic energy of the fission fragments, which in turn is converted into heat through the slowing down and stopping of the fragments. For the heterogeneous reactors this heating occurs within the fuel elements. Heating also arises through the release and absorption of the radiation from the fission process and from the radioactive materials formed. The heat generated in a reactor is removed by a primary coolant flowing through it.

Heat is not generated uniformly in a reactor core. The heat flux generally decreases axially and radially from a peak near the center of the reactor. In addition, local perturbations in heat generation occur because of the arrangement of the reactor fuel, its burn-up, various neutron "poisons" used to shape the power distribution, and inhomogeneities in the reactor structure. These variations impose special considerations in the design of reactor cooling systems, including the need for establishing variations in coolant flow rate through the reactor core to achieve uniform temperature rise in the coolant, avoiding local hot-spot conditions, and avoiding local thermal stresses and distortions in the structural members of the reactor.

Nuclear reactors have the unique thermal characteristic that heat generation continues after shutdown because of fission and radioactive decay of fission products. Significant fission heat generation occurs only for a few seconds after shutdown. Radioactive-decay heating, however, varies with the decay characteristics of the mixture of fission products and persists at low but significant power levels for many days. *See* RADIOACTIVITY.

Accurate analysis of fission heat generation as a function of time immediately after reactor shutdown requires detailed knowledge of the speed and reactivity worth of the control rods. The longer-term fission-product-decay heating, on the other hand, depends upon the time and power level of prior reactor operation and the isotopic composition of the fuel. Typical

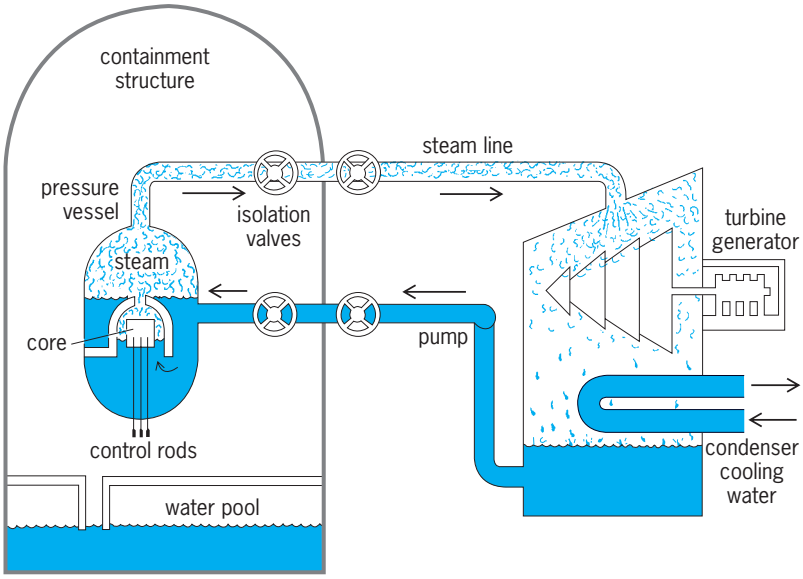

Fig. 2. Boiling-water reactor. (*Atomic Industrial Forum, Inc.*)

values of the total heat generation after shutdown (as percent of operating power) are 10–20% after 1 s, 5–10% after 10 s, approximately 2% after 10 min, 1.5% after 1 h, and 0.7% after 1 day. These rates are important in reactor safety since 0.7% of the thermal power (2564 MW) of a 1000-MW(e) commercial nuclear power plant is approximately 18 MW of heat still being generated 1 day after the reactor is shut down.

Reactor coolants. Coolants are selected for specific applications on the basis of their heat-transfer capability, physical properties, and nuclear properties.

Water. Water has many desirable characteristics. It was employed as the coolant in many of the first production reactors, and most power reactors still utilize water as the coolant. In a boiling-water reactor (BWR; **Fig. 2**), the water boils directly in the reactor core to make steam that is piped to the turbine. In a pressurized-water reactor (PWR; **Fig. 3**), the coolant water is kept under increased pressure to prevent boiling. It transfers heat to a separate stream of feed water in a steam generator, changing that water to steam. **Figure 4** shows the relation of the core and heat removal systems to the condenser, electric power system, and waste management system in the Prairie Island (Minnesota) nuclear plant, which is typical of plants using pressurized-water reactors. Coolant intake water is pumped through hundreds of 1-in.-diameter (25-mm) tubes in the condenser, and the warm water from the condenser is then pumped over cooling towers and returned to the plant. *See* COOLING TOWER; RADIOACTIVE WASTE MANAGEMENT; VAPOR CONDENSER.

For both boiling-water and pressurized-water reactors, the water serves as the moderator as well as the coolant. Both light water and heavy water are excellent neutron moderators, although heavy water (deuterium oxide) has a neutron-absorption cross section approximate $^1/_{500}$ that for light water that makes it possible to operate reactors using heavy water with natural uranium fuel.

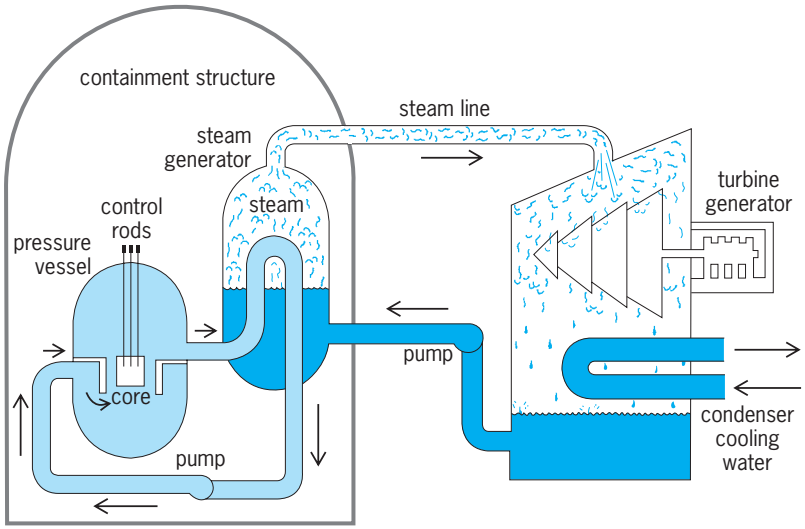

Fig. 3. Pressurized-water reactor. (*Atomic Industrial Forum, Inc.*)

There is no serious neutron-activation problem with pure water; ^{16}N, formed by the (n,p) reaction with ^{16}O (absorption of a neutron followed by emission of a proton), is a major source of activity, but its 7.5-s half-life minimizes this problem since the radioactivity, never very high to begin with, quickly decays away. The most serious limitation of water as a coolant for power reactors is its high vapor pressure. A coolant temperature of 550°F (288°C) requires a system pressure of at least 1100 psi (7.3 megapascals). This temperature is below fossil-fuel power station practice, for which steam temperatures near 1000°F (538°C) have become common. Lower thermal efficiencies result from lower temperatures. Boiling-water reactors operate at about 70 atm (7 MPa), and pressurized-water reactors at 150 atm (15 MPa). The high pressure necessary for water-cooled power reactors determines much of the plant design. This will be discussed below. *See* NUCLEAR REACTION.

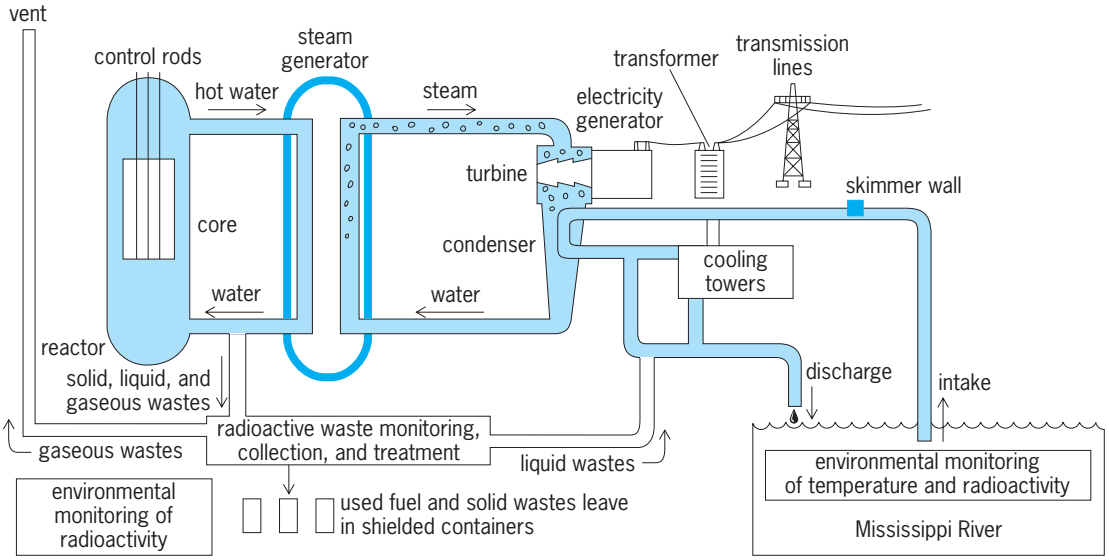

Fig. 4. Nuclear plant, using pressurized-water reactors. (*Northern States Power Company*)

Gases. Gases are inherently poor heat-transfer fluids as compared with liquids because of their low density. This situation can be improved by increasing the gas pressure; however, this introduces other problems and costs. Helium is the most attractive gas (it is chemically inert and has good thermodynamic and nuclear properties) and has been selected as the coolant for the development of high-temperature gas-cooled reactor (HTGR) systems (**Fig. 5**), in which the gas transfers heat from the reactor core to a steam generator. The British advanced gas reactor (AGR), however, uses carbon dioxide (CO_2). Gases are capable of operation at extremely high temperature, and they are being considered for special process applications and direct-cycle gas-turbine applications. Hydrogen was used as the coolant for the reactor developed in the Nuclear Engine Rocket Vehicle Application (NERVA) Program, now terminated. Heated gas discharging through the nozzle developed the propulsive thrust.

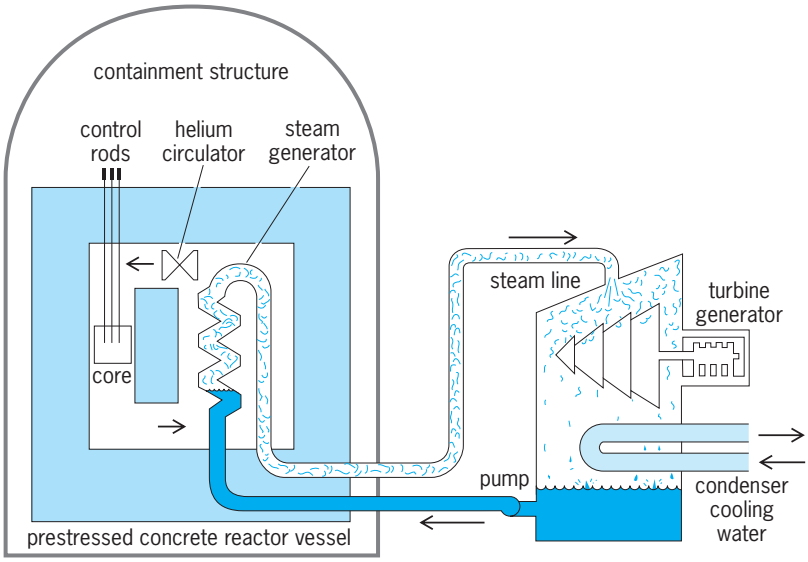

Fig. 5. High-temperature gas-cooled reactor. (*Atomic Industrial Forum, Inc.*)

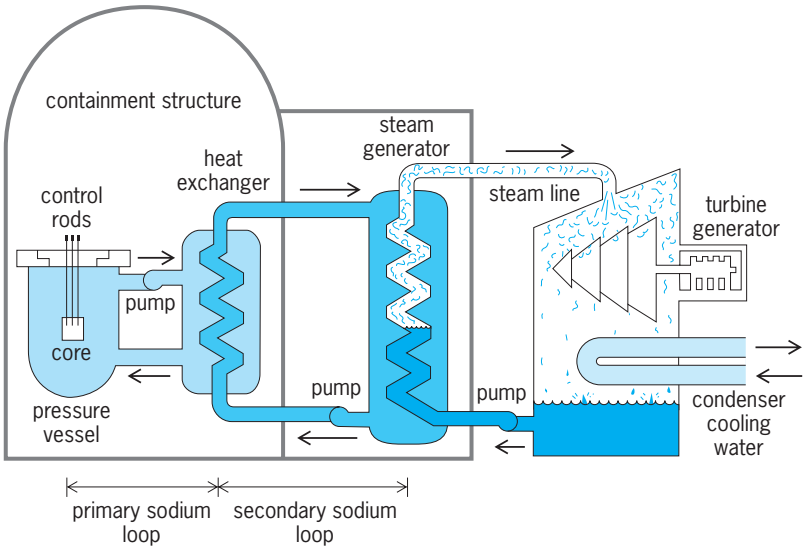

Fig. 6. Loop-type liquid-metal fast breeder reactor. In some designs, referred to as pool-type, the heat exchanger and primary sodium pump are located with the core inside the pressure vessel. (*Atomic Industrial Forum, Inc.*)

Liquid metals. The alkali metals, in particular, have excellent heat-transfer properties and extremely low vapor pressures at temperatures of interest for power generation. Sodium vapor pressure is less than 17.6 lb/in.^{2} (120 kilopascals) at 1650°F (900°C). Sodium is attractive because of its relatively low melting point (208°F or 98°C) and high heat-transfer coefficient. It is also abundant, commercially available in acceptable purity, and relatively inexpensive. It is not particularly corrosive, provided low oxygen concentration is maintained. Its nuclear properties are excellent for fast reactors. In the liquid-metal fast breeder reactor (LMFBR; **Fig. 6**), sodium in the primary loop collects the heat generated in the core and transfers it to a secondary sodium loop in the heat exchanger, from which it is carried to the steam generator in which water is boiled to make steam.

Sodium-24, formed by the absorption of a neutron, is an energetic gamma emitter with a 15-h half-life. The primary system is surrounded by biological shielding, and approximately 2 weeks is required for decay of ^{24}Na activity prior to access to the system for repair or maintenance. Another sodium isotope, ^{22}Na, has a 2.6-year half-life and builds up slowly with plant operation until eventually the radioactivity reaches a level where it is necessary to drain the sodium before maintenance can be performed.

Sodium does not decompose, and no makeup is required. However, sodium reacts violently if mixed with water. This requires extreme care in the design and fabrication of sodium-to-water steam generators and backup systems to cope with occasional leaks. The poor lubricating properties of sodium and its reaction with air further specify the mechanical design requirements of sodium-cooled reactors. Nevertheless, sodium-cooled reactors have operated with good reliability and relatively high operating availability. The other alkali metals exhibit similar characteristics but appear to be less attractive than sodium. The eutectic alloy of sodium with potassium (NaK), however, has the advantage that it remains liquid at room temperature, but adversely affects the properties of steel used in system components. Mercury has also been used as a coolant but its overall properties are less favorable than sodium.

Plant balance. The nuclear chain reaction in the reactor core produces energy in the form of heat, as the fission fragments slow down and dissipate their kinetic energy in the fuel. This heat must be removed efficiently and at the same rate it is being generated in order to prevent overheating of the core and to transport the energy outside the core, where it can be converted to a convenient form for further utilization. The energy transferred to the coolant, as it flows past the fuel element, is stored in it in the form of sensible heat and pressure and is called the enthalpy of the fluid. In an electric power plant, the energy stored in the fuel is further converted to kinetic energy (the energy of motion) through a device called a prime mover which, in the case of

nuclear reactors, is predominantly a steam turbine. Another conversion takes place in the electric generator, where kinetic energy is converted into electric power as the final energy form to be distributed to the consumers through the power grid and distribution system. *See* ENTHALPY; GENERATOR; PRIME MOVER; STEAM TURBINE.

After the steam expands to spin the rotor in the turbine, it exits into the condenser where it is cooled to produce condensate water, which is then fed back to the core or to the steam generator, and the cycle is repeated. The condenser is a very large and important part of the plant. Roughly twice the amount of heat that has been converted in the turbine for the production of electric power is removed in the condenser for further rejection to the environment. The second law of thermodynamics defines the fraction of useful energy or work that can be attained by a thermal engine. Large amounts of water circulate through the condenser to carry the waste heat to its ultimate sink, which may be the sea, a river, a lake, or the atmosphere itself through cooling towers. *See* COOLING TOWER; STEAM CONDENSER; THERMODYNAMIC PRINCIPLES.

The energy conversion part of the power plant (that is, excluding the reactor itself with its main components and system) is often called balance of plant (BOP). It incorporates a large number of components and systems. It represents over four-fifths of the plant's total cost, and is important for the efficient and safe operation of the plant.

Fluid flow and hydrodynamics. Because heat removal must be accomplished as efficiently as possible, considerable attention must be given to fluid-flow and hydrodynamic characteristics of the system. *See* FLUID FLOW; HYDRODYNAMICS.

The heat capacity and thermal conductivity of the fluid at the temperature of operation have a fundamental effect upon the design of the reactor system. The heat capacity determines the mass flow of the coolant required. The fluid properties (thermal conductivity, viscosity, density, and specific heat) are important in determining the surface area required for the fuel—in particular, the number and arrangement of the fuel elements. These factors combine to establish the pumping characteristics of the system because the pressure drop and coolant temperature rise in the core are directly related. *See* CONDUCTION (HEAT); HEAT CAPACITY; VISCOSITY.

Secondary considerations include other physical properties of the coolant, particularly its vapor pressure. If the vapor pressure is high at the operating temperature, local or bulk boiling of the fluid may occur, unless the system pressure is maintained at a higher pressure at all times. This, in turn, must be considered in establishing the heat transfer coefficient for the fluid. *See* VAPOR PRESSURE.

Because the coolant absorbs and scatters neutrons, variations in coolant density also affect reactor performance and control. This is particularly significant in reactors in which the coolant exists in two phases—for example, the liquid and vapor phases in boiling systems. Gases, of course, do not undergo phase change, nor do liquids operating at temperatures well below their boiling point; however, the fuel density does change with temperature and may have an important effect upon the reactor.

Power generation and, therefore, the heat removal rate are not uniform throughout the reactor. If the mass flow rate of the coolant is uniform through the reactor core, then unequal temperature rise of the coolant results. This becomes particularly significant in power reactors in which it is desired to achieve the highest possible coolant outlet temperature to attain maximum thermal efficiency of the power cycle. The performance limit of the coolant is set by the temperature in the hottest region or channel of the reactor. Unless the coolant flow rate is adjusted in the other regions of the reactor, the coolant will leave these regions at a lower temperature and thus will reduce the average coolant outlet temperature. In power reactors, this effect is reduced by adjusting the rate of flow to each region of the reactor in proportion to its heat generation rate. This involves very careful design and analysis of the system. In the boiling-type reactor, this effect upon coolant temperature does not occur because the exit temperature of the coolant is constant at the saturation temperature for the system. However, the variation in power generation in the reactor is reflected by a difference in the amount of steam generated in the various zones, and orificing is still required to achieve most effective use of coolant flow.

In some power reactors, the flow rate and consequent pressure drop of the coolant are sufficient to create large mechanical forces in the system. It is possible for the pressure drop through the fuel assemblies to exceed the weight of the fuel elements in the reactor, with a resulting hydraulic lifting force on the fuel elements. Often this requires a design arrangement to hold the fuel elements down. Although this problem can be overcome by employing downward flow through the system, it is often undesirable to do so because of shutdown-cooling considerations. It is desirable in most systems to accomplish shutdown cooling by natural-convection circulation of the coolant. If downflow is employed for forced circulation, then shutdown cooling by natural-convection circulation requires a flow reversal, which can introduce new problems. Therefore, hydraulic forces are overcome by use of support plates, seismic restraints, and fuel spacers.

Thermal stress considerations. The temperature of the reactor coolant increases as it circulates through the reactor core. This increase in temperature is constant at steady-state conditions. Fluctuations in power level or in coolant flow rate result in variations in the temperature rise. These are reflected as temperature changes in the coolant core exit temperature, which in turn result in temperature changes in the coolant system.

A reactor is capable of very rapid changes in power level, particularly reduction in power level, which is a safety feature of the plant. Reactors are equipped with mechanisms (reactor scram systems) to ensure

rapid shutdown of the system in the event of leaks, failure of power conversion systems, or other operational abnormalities.

Therefore, reactor coolant systems must be designed to accommodate the temperature transients that may occur because of rapid power changes. In addition, they must be designed to accommodate temperature transients that might occur as a result of a coolant system malfunction, such as pump stoppage. The consequent temperature stresses induced in the various parts of the system are superimposed upon the thermal stresses that exist under normal steady-state operations and produce conditions known as thermal shock or thermal cycling.

In some systems, it is not uncommon for the thermal stresses to be significant. In these cases, careful attention must be given to the transient stresses, and thermal shielding (such as thermal sleeves on pipes and baffles) is commonly employed in critical sections of the system. Normally, this consists of a thermal barrier which, by virtue of its heat capacity and insulating effect, delays the transfer of heat, thereby reducing the rate of change of temperature and protecting critical system components from thermal stresses.

Thermal stresses are also important in the design of reactor fuel elements. Metals that possess dissimilar thermal-expansion coefficients are frequently required. Heating of such systems gives rise to distortions, which in turn can result in flow restrictions in coolant passages. Careful analysis and experimental verification are often required to avoid such circumstances.

Coolant system components. The development of reactor systems has led to the development of special components for reactor component systems. Because of the hazard of radioactivity, leak-tight systems and components are a prerequisite to safe, reliable operation and maintenance. Special problems are introduced by many of the fluids employed as reactor coolants.

More extensive component developments have been required for sodium, which is chemically active and is an extremely poor lubricant. Centrifugal pumps employing unique bearings and seals have been specially designed. Sodium is an excellent electrical conductor and, in some special cases, electromagnetic-type pumps have been used. These pumps are completely sealed, contain no moving parts, and derive their pumping action from electromagnetic forces imposed directly on the fluid. *See* CENTRIFUGAL PUMP; ELECTROMAGNETIC PUMP.

In addition to the variety of special pumps developed for reactor coolant systems, there is a variety of piping system components and heating exchange components. As in all flow systems, flow-regulating devices such as valves are required, as well as flow instrumentation to measure and thereby control the systems. Here again, leak tightness has necessitated the development of valves with special seals such as metallic bellows around the valve stem to ensure system integrity. Measurement of flow and pressure has also required the development of sensing instrumentation that is reliable and leak-tight. *See* FLOW MEASUREMENT; PRESSURE MEASUREMENT; VALVE.

Many of these developments have borrowed from other technologies where toxic or flammable fluids are frequently pumped. In many cases, however, special equipment has been developed specifically to meet the requirements of the reactor systems. An example of this type of development involves the measurement of flow in liquid-metal piping systems. The simple principle of a moving conductor in a magnetic field is employed by placing a magnet around the pipe and measuring the voltage generated by the moving conductor (coolant) in terms of flow rate. Temperature compensation is required, and calibration is important.

Although the development of nuclear power reactors has introduced many new technologies, no method has yet displaced the conventional steam cycle for converting thermal energy to mechanical energy. Steam is generated either directly in the reactor (direct-cycle boiling reactor) or in auxilliary steam generation equipment, in which steam is generated by transfer of heat to water from the reactor coolant. These steam generators require a very special design, particularly when dissimilar fluids are involved. Typical of the latter problem is the sodium-to-water steam generators in which integrity is essential because of the potentially violent chemical reaction between sodium and water.

Core Design and Materials

A typical reactor core for a power reactor consists of the fuel element rods supported by a grid-type structure inside a vessel (Fig. 1).

The primary function of the vessel is to contain the coolant. Its design and materials are determined by such factors as the nature of the coolant (corrosive properties) and quantity and configuration of fuel. The vessel has several large nozzles for coolant entrance and exit and smaller nozzles used for controlling reactor operation (control-rod drive mechanisms and instruments). The top of the vessel unbolts for refueling operations.

The pressure vessel design takes account of thermal stresses caused by temperature differences in the system. An exceptionally high degree of integrity is demanded of this equipment. Reactors are designed to permit removal of the internals from the vessel, which is then periodically inspected by automatic devices that can detect any cracks which might have developed during operation. These in-service inspections are required by codes and regulations on a fixed time schedule.

Structural materials. Structural materials employed in reactor systems must possess suitable nuclear and physical properties and must be compatible with the reactor coolant under the conditions of operation. Some requirements are established because of secondary effects; for example, corrosion limits may be established by the rate of deposition of coolant-entrained corrosion products on critical surfaces rather than by the rate of corrosion of the base material. *See* CORROSION.

The most common structural materials employed in reactor systems are stainless steel and zirconium alloys. Zirconium alloys have favorable nuclear and physical properties, whereas stainless steel has favorable physical properties. Aluminum is widely used in low-temperature test and research reactors; zirconium and stainless steel are used in high-temperature power reactors. Zirconium is relatively expensive, and its use is therefore confined to applications in the reactor core where neutron absorption is important. *See* ALUMINUM; STAINLESS STEEL; ZIRCONIUM.

The 18-8 series stainless steels have been used for structural members in both water-cooled reactors and sodium-cooled reactors because of their corrosion resistance and favorable physical properties at high temperatures. Type 304, 316, and 347 stainless steel have been used the most extensively because of their weldability, machinability, and physical properties, although other iron-nickel alloys, such as Inconel, are also used. To increase reliability and to reduce cost, heavy-walled pressure vessels are normally fabricated from carbon steels and clad on the internal surfaces with a thin layer of stainless steel to provide the necessary corrosion resistance. *See* IRON ALLOYS; NICKEL ALLOYS; STEEL.

As the size of power reactors has increased, it has become necessary in some instances to field-fabricate reactor vessels. A notable example is the on-site fabrication of the large pressure vessel required for the liquid-metal fast breeder reactor at Creys-Malville, France. This involved field-welding of wall sections and subsequent stress relieving. Both steel and prestressed concrete vessels for gas-cooled reactors are also field-fabricated.

Research reactors operating at low temperatures and pressures introduced special experimental considerations. The primary objective is to provide the maximum volume of unperturbed neutron flux for experimentation. It is desirable, therefore, to extend the experimental irradiation facilities beyond the vessel wall. This has introduced the need for vessels constructed of materials having a low cross section for neutron capture. Relatively large low-pressure aluminum reactor vessels with wall sections as thin as practicable have been manufactured for research reactors.

Fuel cladding. Reactors maintain a separation of fuel and coolant by cladding the fuel. The cladding is designed to prevent the release of radioactivity from the fuel. The cladding material must be compatible with both the fuel and the coolant.

The cladding materials must also have favorable nuclear properties. The neutron-capture cross section is most significant because the unwanted absorption of neutrons by these materials reduces the efficiency of the nuclear fission process. Aluminum is a very desirable material in this respect; however, its physical strength and corrosion resistance in water decrease very rapidly above about 300°F (149°C), and it is therefore used only in test and research reactors that are not used to produce power.

Zirconium has favorable neutron properties, and in addition is corrosion-resistant in high-temperature water. It has found extensive use in water-cooled power reactors. The technology of zirconium production and the use of zirconium-based alloys, such as Zircaloy, have advanced tremendously under the impetus of the various reactor development programs.

Stainless steel is used for the fuel cladding in fast reactors, in some light-water reactors for which neutron captures are less important.

Control and Instrumentation

A reactor is critical when the rate of production of neutrons equals the rate of absorption in the system plus the rate of leakage out of the core. The control of reactors requires the continuing measurement and adjustment of the critical condition. The neutrons are produced by the fission process and are consumed in a variety of ways, including absorption to cause fission, nonfission capture in fissionable materials, capture in fertile materials, capture in structure or coolant, and leakage from the reactor to the shielding. A reactor is subcritical (power level decreasing) if the number of neutrons produced is less than the number consumed. The reactor is supercritical (power level increasing) if the number of neutrons produced exceeds the number consumed. *See* REACTOR PHYSICS.

Reactors are controlled by adjusting the balance between neutron production and neutron consumption. Normally, neutron consumption is controlled by varying the absorption or leakage of neutrons; however, the neutron generation rate also can be controlled by varying the amount of fissionable material in the system.

It is necessary for orderly startup and control of a reactor that the neutron flux be sufficiently high to permit rapid power increase; in large reactors, too slow a startup is erratic and uneconomical. During reactor startup, a source of neutrons is useful, therefore, for control, and aids in the instrumentation of reactor systems. Neutrons are obtained from the photoneutron effect in materials such as beryllium. Neutron sources consist of a photon (gamma-ray) source and beryllium, such as antimony-beryllium. Antimony sources are particularly convenient for use in reactors because the antimony is reactivated by the reactor neutrons each time the reactor operates.

Control drives and systems. The reactor control system requires the movement of neutron-absorbing rods (control rods) in the reactor under carefully controlled conditions. They must be arranged to increase reactivity (increase neutron population) slowly and under good control. They must be capable of reducing reactivity, both rapidly and slowly. Power reactors have inherent negative temperature coefficients that make it practical to use stepwise control of rod motions but with smooth and limited charges in power level.

The control drives can be operated by the reactor operator or by automatic control systems. Reactor scram (rapid reactor shutdown) can be initiated automatically by a wide variety of system scram-safety

signals, or it can be started by the operator depressing a scram button in the control room.

Control drives are electromechanical or hydraulic devices that impart in-and-out motion to the control rods. They are usually equipped with a relatively slow-speed reversible drive system for normal operational control. Scram is usually effected by a high-speed overriding drive accompanied by disconnecting the main drive system. Allowing the control rods to drop into the core by means of gravity is one common method. To enhance reliability of the scram system, its operation can be initiated by deenergizing appropriate electrical circuits. This also automatically produces reactor scram in the event of a system power failure. Hydraulic or pneumatic drive systems, as well as a variety of electromechanical systems, have also been developed.

In addition to the actuating motions required, control-rod drive systems must also provide indication of the rod positions at all times. Various types of sensors, as well as arrangements of switch indicators, are employed as postion indicators.

Instrumentation. Reactor control involves continuing measurement of the condition of the reactor. Neutron-sensitive ion chambers may be located outside the reactor core, and the flux measurements from the detectors are combined to measure a flux that is proportional to the average neutron density in the reactor. The chamber current is calibrated against a thermal power measurement and then is applied over a wide range of reactor power levels. The neutron-sensitive detector system is capable of measuring the lowest neutron flux in the system, including that produced by the neutron source when the reactor itself is subcritical. *See* IONIZATION CHAMBER.

Normally, several ranges of instrument sensitivities are required to cover the entire operating range. A range is required for low-level operation, beginning at the source level, whereas others are required for the intermediate- and high-power levels. Three ranges of detectors are common in power reactor systems, and some systems contain a larger number. The total range to be covered is 7–10 decades (factors of 10) of power level.

The chamber current of a neutron detector, suitably amplified, can be employed as a signal to operate automatic control system devices as well as to actuate reactor scram. In addition to power level, rate of change of power level is an important measurement which is recorded and employed to actuate various alarm and trip circuits. The normal range for the current ion chambers is approximately 10^{-14} to 10^{-4} A. This current is suitably amplified in logarithmic and period amplifiers. The power monitors are part of a reactor's plant protection system.

Applications

Reactor applications include mobile, stationary, and packaged power plants; production of fissionable fuels (plutonium and uranium-233) for military and commercial applications; research, testing, teaching-demonstration, and experimental facilities; space and process heat; dual-purpose design; and special applications. The potential use of reactor radiation or radioisotopes produced for sterilization of food and other products, steam for chemical processes, and gas for high-temperature applications has been recognized. *See* NUCLEAR FUEL CYCLE; NUCLEAR FUELS REPROCESSING; RADIOACTIVITY AND RADIATION APPLICATIONS; SHIP NUCLEAR PROPULSION.

Frank J. Rahn

Bibliography. J. Douglas, The nuclear option, *EPRI Journal*, December 1994; S. Glasstone and A. Sesonke, *Nuclear Reactor Engineering*, 4th ed., 1993; International Atomic Energy Association, *Nuclear Power Reactors in the World*, annually; R. A. Knief, *Nuclear Engineering: Theory and Technology of Commercial Nuclear Power*, 1992; A. Nero, *A Guidebook to Nuclear Reactors*, 1980; F. J. Rahn et al., *A Guide to Nuclear Power Technology,* 1984, reprint 1992; S. C. Stulz and J. B, Kitto, *Steam: Its Generation and Use*, 40th ed., 1992.

Nuclear spectra

The distribution of the intensity of particles (or radiation) emitted in a nuclear process as a function of energy. The nuclear spectrum is a unique signature of the process.

For example, when very slow neutrons (with speeds less than 0.5% of the speed of light) hit nitrogen nuclei, there is a high probability that they will be captured and that the nuclear system which is formed will emit a set of gamma rays (electromagnetic radiation) of very precise energies. The 24 gamma rays have energies ranging from 1.68 to 10.83 MeV, and their relative intensities are well known. A spectrum of these gamma rays, that is, the number of gamma rays having a particular energy, versus that energy can provide a unique signature of the presence of nitrogen. An application is the passing of a beam of slow neutrons through luggage at an airport: the presence of unusual amounts of nitrogen indicates that a plastic explosive may be present. This testing for the presence of nitrogen is nondestructive: relatively few neutrons are needed to produce the characteristic spectrum, and the luggage and its contents are not harmed. *See* GAMMA RAYS; NONDESTRUCTIVE EVALUATION.

The process that leads to such a spectrum is well understood. The type (isotope) of nitrogen predominantly found in nature (over 99.6%) is $^{14}_{7}N$, where 7 is the number of protons and 14 is the total number of nucleons (neutrons + protons). When ^{14}N captures a neutron, ^{15}N is formed, but it is not formed in its state of lowest energy and most symmetric shape, the ground state. Rather it is created in an unstable form which has an energy that is 10.83 MeV greater than that of the ground state and which decays rapidly (in a time of less than 10^{-11} s) to the ground state of ^{15}N by emitting gamma rays. The decays to the ground state take place via excited states of ^{15}N. When a number of ^{15}N nuclei have decayed, gamma rays with 24 different but well-defined energies will have been

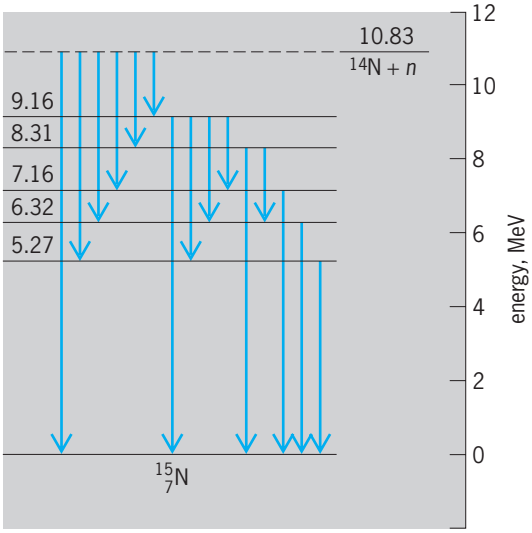

Fig. 1. Electromagnetic transitions (arrows) that result when ^{14}N captures a slow neutron. Scale to the right measures energy above the ground state of ^{15}N. Only some of the 24 gamma-ray transitions that can occur are shown.

emitted (**Fig. 1**). Such a spectrum is called a discrete spectrum because the 24 energies form a discrete set of peaks.

Natural radioactive decay. Naturally occurring radioactive nuclei change (decay) by emitting negatively charged electrons (β^-), positively charged alpha particles (a system composed of two protons and two neutrons), as well as uncharged gamma rays and neutrinos ($\overline{\nu_e}$). In one sequence of nuclear decays found in nature, thorium-232 ($^{232}_{90}Th$) is the progenitor of a family of nuclei that are sequentially formed in a series of alpha and β^- decays which terminate in the formation of lead-208 ($^{208}_{82}Pb$), a stable nucleus. As ^{232}Th is converted to ^{208}Pb in this chain, bismuth-212 ($^{212}_{83}Bi$) is produced. The decay of ^{212}Bi illustrates the nature of the alpha, beta, and gamma radiations observed in radioactive chains. The mass of $^{212}_{83}Bi$ is greater than the mass of polonium-212 ($^{212}_{84}Po$) plus the mass of an electron, and it is also greater than the mass of thallium-208 ($^{208}_{81}Tl$) plus the mass of an alpha particle (**Fig. 2**). When a physical system has a greater mass than that of a system to which it could decay (without breaking the fundamental conservation laws), it does so. For instance, $^{212}_{83}Bi$ transforms into $^{212}_{84}Po$ with the emission of a negative electron (β^-) and a neutrino ($\overline{\nu_e}$). The chance of this happening, statistically, is 2/3. With a chance of roughly 1/3, however, $^{212}_{83}Bi$ transforms into $^{208}_{81}Tl$ with the emission of an alpha particle. The nuclei ^{212}Po and ^{208}Tl are not always produced directly in their ground state. When these nuclei are formed in an excited state, that state decays to the ground state by emitting electromagnetic radiation in the form of gamma rays. The negative electron spectrum is continuous because the energy available in the transition is divided between three bodies: ^{212}Po, the negative electron, and a neutrino. By studying the alpha and beta spectra emitted by ^{212}Bi when it transforms into ^{208}Tl and ^{212}Po, the differences in the masses of these three nuclei can be determined. When the spectrum of the gamma rays involved in the ^{208}Tl beta decay to ^{208}Pb is studied, the parameters of the excited states of ^{208}Pb can be determined. *See* BETA PARTICLES; RADIOACTIVITY.

Nuclear reactions. When nuclear processes take place in the laboratory, not only can alpha particles, electrons, and gamma rays be produced, but most groupings of nucleons (neutrons and protons), leptons (electrons, muons, neutrinos), photons (electromagnetic radiation), and mesons (pions, kaons, and so forth) can be created, subject to conservation laws.

For instance, when the most abundant isotope of carbon, ^{12}C, is bombarded by protons that have been given a kinetic energy in a particle accelerator, a large number of nuclear processes can take place. A few of these, reactions (1)–(4), are both written out and

$$^{12}C + p \rightarrow {}^{12}C + p \qquad {}^{12}C(p, p)^{12}C \tag{1}$$

$$^{12}C + p \rightarrow {}^{12}C^* + p', \; {}^{12}C^* \rightarrow {}^{12}C + \gamma \qquad {}^{12}C(p, p)^{12}C^* (\gamma)^{12}C \tag{2}$$

$$^{12}C + p \rightarrow {}^{11}C + d \qquad {}^{12}C(p, d)^{11}C \tag{3}$$

$$^{12}C + p \rightarrow {}^{8}Be + \alpha + p \qquad {}^{12}C(p, p\alpha)^{8}Be \tag{4}$$

given in compact notation. Reaction (1) is elastic scattering, with ^{12}C remaining in its ground state. In **Fig. 3**, the group labeled p_0 represents elastically scattered protons. Reaction (2) is inelastic scattering, with ^{12}C first formed in an excited state, which then decays by gamma-ray emission to the ground state of ^{12}C. The groups at lower energies in Fig. 3, p_1 and p_2, represent inelastically scattered protons: ^{12}C is then formed in its excited states at 4.4 and 7.7 MeV. In reaction (3), ^{11}C is formed at the same time that a deuteron is emitted. The ^{11}C nucleus then decays to ^{11}B with the emission of a positive electron

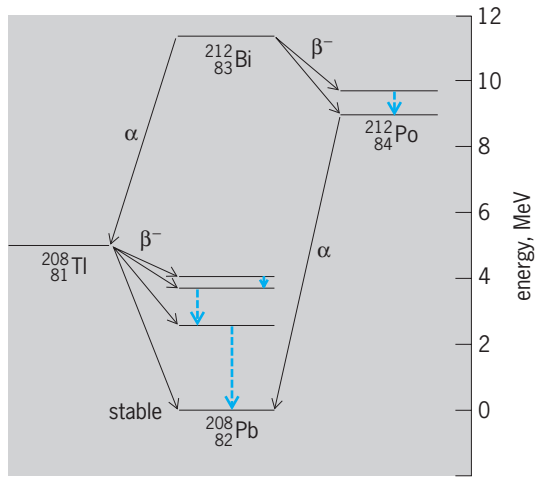

Fig. 2. Decay of ^{212}B into ^{208}Pb via ^{212}Po and ^{208}Tl. The vertical arrows show some of the observed decays of excited states of ^{212}Po and ^{208}Pb to the ground states of these nuclei, via electromagnetic transitions involving the emission of gamma rays. The scale on the right shows the energy differences between the states.

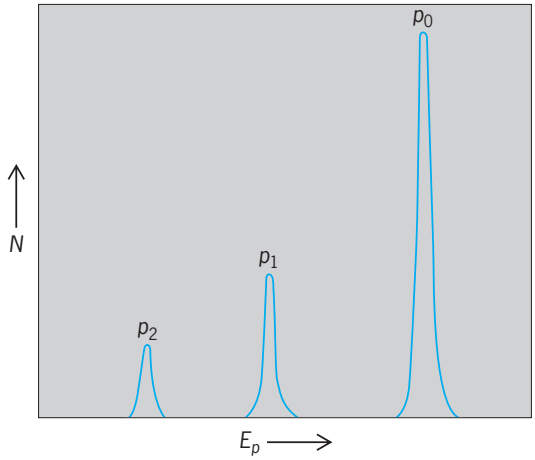

Fig. 3. Idealized spectrum of protons scattered from ^{12}C. N is the number of scattered protons; E_p is their energy. The proton groups labeled p_0, p_1, and p_2 are identified in the text. The groups are shown as being broad (not monoenergetic) because the experimental equipment that measures them is not perfect.

(β^+) and of a neutrino (υ_e). In reaction (4), the role of the proton is to shock ^{12}C into fissioning (splitting) into an alpha particle and ^{8}Be. The nucleus is unstable and very rapidly breaks up into two alpha particles; thus ^{12}C disappears and three alpha particles are created.

The minimum kinetic energies of the protons needed to create these processes vary. As the proton energy is increased, more exotic processes may take place. An example is the reaction $C^{12}(p, \pi^+)^{13}C$. Here the incident proton's kinetic energy permits a pion to emerge from the nucleus, and a ^{13}C nucleus is created. The π^+ then decays into two leptons: a muon (μ^+) and a neutrino (υ_μ). By studying the spectra of the particles emitted in nuclear reactions (Fig. 3), the structure of the nuclei that are created can be determined. *See* NUCLEAR REACTION; NUCLEAR STRUCTURE.

Fission. A very important type of nuclear reaction is fission. The most common type of fission involves a very slow neutron that amalgamates with a heavy nucleus such as uranium-235 (^{235}U). For a very short time, a composite nucleus, ^{236}U, is formed, but it is so excited and its shape is so distended that it breaks up into several fragments. The larger fragments mostly decay by beta, gamma, and neutrino emission until stable end products are reached. The lightest fragments are fast neutrons, some of which can be slowed down to cause fission of another uranium nucleus. The spectrum of the neutrons released during the fission must be carefully studied to determine how best to moderate the velocities of the neutrons to make the probability of a chain reaction (continuously sustained fission) as high as possible. *See* NUCLEAR FISSION.

Measurements. The methods used to measure nuclear spectra depend on the nature of the particles (radiation) involved. The most accurate energy measurements are those of gamma rays. Gamma-ray spectra can be measured by determining the energy deposited by the gamma rays in a crystal, often made of sodium iodide, containing thallium impurities [NaI(Tl)], or of germanium, containing lithium impurities [Ge(Li)]. In a NaI(Tl) detector, the gamma-ray energy is transferred to electrons within the crystal, and these charged particles in turn produce electromagnetic radiation with frequencies in the visible range. The crystal is surrounded by detectors (photomultipliers) that are sensitive to the visible light. The intensity of the signal in the photomultipliers is proportional to the energy of the gamma rays that entered the NaI(Tl) crystal. The signal pulse is amplified electronically, and the pulse heights (pulse sizes) are displayed in a pulse-height multichannel analyzer in a histogram. Usually the number of pulses having a certain height (strength) is plotted versus the height. What results is a plot showing the number of gamma rays having a certain energy versus the energy of the gamma rays, a spectrum. Gamma-ray energy measurements with an uncertainty of less than 1 part in 10^5 have been reported. While the use of a NaI(Tl) crystal is restricted to determining gamma-ray spectra, the procedure of looking at signals in bins, that is, in plotting histograms in which the number of signals corresponding to a certain energy is plotted versus that energy, is general to nuclear spectra measurements. *See* GAMMA-RAY DETECTORS; PHOTOMULTIPLIER.

Different types of nuclear particles (and radiation) are observed by detectors specific to the particles. For instance, neutron spectra are often determined by measuring their velocities. This is done by a time-of-flight technique in which an electronic timer measures the time interval between the emission of the neutron from a nucleus and its arrival at a detector a known distance away. This measurement uniquely determines the velocity, and thus the kinetic energy, of the neutrons. The detector is typically a plastic containing a large fraction of hydrogen. The incident neutrons interact with the hydrogen nuclei, protons, and transfer their kinetic energies to them. Protons, which are charged, then produce detectable electrical signals because they ionize the matter through which they move. The detection of gamma rays and of neutrons, both of which are uncharged, depends on their interaction with charged particles in the medium in which they move. It is the charged particles that then lead to the electronic signal that is measured. *See* NEUTRON SPECTROMETRY; TIME-OF-FLIGHT SPECTROMETERS.

Measurements of nuclear spectra involving charged particles, such as pions, protons and alpha particles, are often made by determining their momenta (mass × velocity) and then calculating the corresponding kinetic energy. Momentum measurements are made by passing the beam of charged particles through a region in which a magnetic field exists. A magnetic field that is constant in time will not cause a change in a charged particle's speed, but it will cause a charged particle to deviate in its path. For instance, if the direction of a magnetic field of magnitude B is perpendicular to the path of a particle of charge q, mass m, and velocity υ, the particle

will move in a circle of radius r given by Eq. (5).

$$r = \frac{mv}{qB} \tag{5}$$

Thus, by determining the radius of the circle, the momentum, mv, of the charged particle, and therefore its kinetic energy, can be determined. After the particle has been deviated from its path sufficiently so that r may be calculated, it strikes a detector. In earlier times, such detectors often consisted of photographic emulsions, in which ionizing particles left evidence of their passage. *See* PARTICLE ACCELERATOR.

Modern magnetic spectrometers use sophisticated counter telescopes and multiwire proportional counters, which permit not only the registering of the particles characterized by a certain value of the radius of curvature (and therefore of momentum) but enable the particular particle (proton, alpha particle, or whatever) that caused the signal to be identified. Contemporary magnetic spectrometer systems not only utilize complex arrangements of magnetic fields, detectors, and electronics but also generally require powerful computers to monitor and analyze the results. *See* PARTICLE DETECTOR.

Fay Ajzenberg-Selove

Bibliography. H. L. Anderson (ed.), *A Physicist's Desk Reference: Physics Vade Mecum*, 2d ed., 1989; F. Ajzenberg-Selove and E. K. Warburton, Nuclear spectroscopy, *Phys. Today*, 36(11):26–32, November 1983; K. S. Krane, *Introductory Nuclear Physics*, 1987; W. McHarris (ed.), *Exotic Nuclear Spectroscopy*, 1991; Y. Yoshiizawa, T. Otsuka, and K. Kusakari, *The Frontier of Nuclear Spectroscopy*, 1993.

Nuclear structure

At the center of every atom lies a small, dense nucleus, which carries more than 99.97% of the atomic mass in less than 10^{-12} of its volume. The nucleus is a tightly bound system of protons and neutrons which is held together by strong forces that are not normally perceptible in nature because of their extremely short range. The small size, strong forces, and many particles in the nucleus result in a highly complex and unique quantal system that at present defies exact analysis. The study of the nucleus and the forces that hold it together constitute the field of nuclear structure physics. *See* ATOMIC STRUCTURE AND SPECTRA; NEUTRON; PROTON; QUANTUM MECHANICS; STRONG NUCLEAR INTERACTIONS.

The protons of the nucleus, being positively charged, generate a spherically symmetric electric field in which the atomic electrons orbit. The strength of this electric field is proportional to the number of protons (the atomic number), and the quantized orbits of the electrons in this field define the material and chemical properties of each element. The cloud of negatively charged atomic electrons normally balances the positive nuclear charge, making the atom electrically neutral and screening the nucleus from view. To penetrate the electron cloud and reach the nucleus generally requires high-energy probes, and so most research on nuclei requires large particle accelerators. *See* ELECTRON; PARTICLE ACCELERATOR.

A discussion of nuclear structure must begin with some definitions of specialized terms and by setting scales of energy, time, and distance for nuclear systems. The atomic number of protons is usually denoted by Z and the number of neutrons, which are electrically neutral, by N. The total number of protons and neutrons (or nucleons) is the mass number $A = Z + N$. Isotopes have the same atomic number, Z, and hence are forms of the same chemical element, having the same chemical properties, but they differ in neutron number; isotones have a common number of neutrons, N, and isobars have the same mass number, A. The normal scheme of notation of a nucleus is to use the chemical symbol with superscripts and subscripts for A, Z, and N in the configuration ${}^{A}_{Z}X_{N}$; for example, ${}^{40}_{20}Ca_{20}$ and ${}^{208}_{82}Pb_{126}$ for isotopes of calcium and lead. The mass of a proton is 1.6726×10^{-27} kg (approximately 1836 electron masses) and that of a neutron is 1.6749×10^{-27} kg. *See* ISOBAR (NUCLEAR PHYSICS); ISOTONE; ISOTOPE.

Nuclei have masses less than the sum of the constituents, the missing mass ΔM being accounted for by the binding energy ΔMc^2 (where c is the speed of light), which holds the nuclear system together. The characteristic energy scale is in megaelectronvolts (1 MeV $= 1.6 \times 10^{-13}$ joule). The internuclear forces generate an attractive potential field which holds the nucleus together and in which the nucleons orbit in highly correlated patterns. The strength of this field, or potential well, is of the order of 40 MeV. The average kinetic energy of a nucleon in a nucleus is about 20 MeV, giving an orbital time of $\tau \simeq 10^{-21}$ s, and the average energy to separate the least-bound nucleon is about 8 MeV. The volume of nuclei increases approximately linearly with mass number A, and the radius is roughly $R = 1.2 \times 10^{-15} \cdot A^{1/3}$ m. *See* NUCLEAR BINDING ENERGY.

Size, shape, and density distributions. The existence of the atomic nucleus and the first estimates of its size came from experiments conducted by E. Rutherford and his collaborators in the period 1906–1911. They used energetic alpha particles emerging from the natural radioactive decays of very heavy elements to penetrate the cloud of atomic electrons and interact directly with the electric field caused by the nuclear protons. They were able to predict the alpha-particle scattering angular distribution and the dependence of the cross section on the number of protons, Z, in the atoms of the target material. However, for high-energy alpha particles and for light targets (with low Z) the simple Rutherford scattering formula was found to break down as the alpha particles were able to overcome the Coulomb field and enter the domain of the strong (nuclear) forces at the nuclear surface. Thus, an estimate of nuclear size and its dependence on Z could be made. *See* ALPHA PARTICLES; SCATTERING EXPERIMENTS (NUCLEI).

This type of study has several shortcomings. Even

with the advent of particle accelerators, the deduced nuclear sizes were a composite of the nuclear size, the size of the alpha particle, and the range of the nuclear forces. Further, the measurements revealed information only about the position of the nuclear surface. For more precise estimates of the charge distribution, a variety of more sophisticated techniques have been developed, including electron scattering, the study of muonic atoms, and the laser spectroscopy of hyperfine atomic structure. Of these, electron scattering has produced some of the most spectacular and unambiguous results.

High-energy electrons interact with the nuclear charge distribution only through the electromagnetic force, which is a precisely understood phenomenon. These electrons can penetrate to the center of the nucleus, and the length scale to be probed can be varied according to the electron's energy. For example, by using electrons of 420-MeV energy, a length scale of 2.95 fermis (1 fermi = 10^{15} m) is probed, which reveals the distribution of nuclear charge without probing the nucleonic quark substructure. Many stable nuclei have been probed in this way. By examining the difference between scattering patterns from adjacent nuclei, the spatial probability distribution of particular protons can be investigated. This is a stringent test for nuclear models. **Figure 1** shows the difference in charge distributions of ${}^{206}_{82}Pb_{126}$ and ${}^{205}_{81}Tl_{126}$, which can be ascribed to the one different proton. The probability distribution for that proton is close to that predicted by model calculations.

The nuclear charge distribution is an average property and is made up from the sum of contributions arising from the probability distributions of each individual proton. Oscillations in the mean charge density reflect the grouping of protons at certain radii, which is indicative of nuclear shells. However, an overall picture of the nuclear charge distributions emerges that has several defining features. First, the nuclear charge density saturates in the interior and

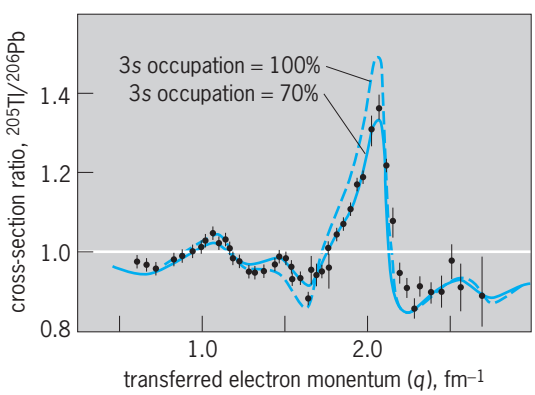

Fig. 1. Ratio of charge distributions of ^{206}Pb and ^{205}Tl, which differ by just one proton. The proton is calculated to lie in the $3s_{1/2}$ shell, where it would have the probability distribution shown. The data (points and bars) indicate some mixing of configurations, and fit the distribution that is predicted when the proton occupies this state for only 70% of the time, as shown. (*After B. Frois et al., A precise determination of the 3s proton orbit, Nucl. Phys., A396:409–418, 1983*)

has a roughly constant value in all but the lightest nuclei. The nucleus has a diffuse skin which is of nearly constant thickness. For spherical nuclei, the charge distribution can be approximated by the Woods-Saxon formula, Eq. (1), where $\rho(r)$ is the density at

$$\rho(r) = \frac{\rho(0)}{1 + e^{(r-R)/a}} \tag{1}$$

$$R = (1.07 \pm 0.02) \times 10^{-15} \cdot A^{1/3}$$

$$a = 0.55 \pm 0.07 \times 10^{15}\text{m}$$

radius r; $\rho(0)$ is the mean density, approximately 1.1×10^{25} coulombs/m^3; R the radius at which the density has fallen to $\rho(0)/2$; and a the surface diffuseness.

Many nuclei are found to have nonspherical shapes. Unlike the atom, which has a spherically symmetric Coulomb field generated by the nucleus, the nuclear field is composed of a complicated superposition of short-range interactions between nucleons, and the most stable nuclear shape is the one that minimizes the energy of the system. In general, it is not spherical, and the nuclear shape is most simply described by a multipole power series, the most important term of which is the nuclear quadrupole moment. A positive quadrupole moment reflects the elongation of nuclei into a prolate or football-like shape, while a negative value reflects an oblate shape like that of Earth. *See* NUCLEAR MOMENTS.

An accurate determination of nuclear matter distributions, that is, the distribution of both protons and neutrons in nuclei, is harder to precisely ascertain. The moments and charge radii of long chains of isotopes have been investigated by using laser hyperfine spectroscopy. The nuclear charge distribution perturbs the atomic quantum states in a small but measurable way that subsequently can be measured by inducing transitions between states. The advantage of this technique over electron scattering is that since very small samples of short-lived radioactive isotopes can be measured, a vastly wider range of nuclei is accessible. **Figure 2** shows the deduced root-mean-square radii of a long chain of rubidium isotopes, clearly indicating that the assumption that the radius is proportional to $A^{1/3}$ is only approximate and that a deviation from sphericity develops for the lightest rubidium isotopes. *See* HYPERFINE STRUCTURE; LASER SPECTROSCOPY.

Nuclear masses and binding energies. The masses of nuclei are measured relative to a standard. The standard is to define the mass of a neutral carbon-12 atom to be 12 atomic mass units or u. On an absolute energy scale 1 u = 931.50 MeV. Relative to this scale a proton has a mass 1.007276 u and a neutron 1.008665 u. Generally, the mass of a nucleus is expressed in terms of the mass of its neutral atom, that is, including the mass of the Z orbiting electrons.

The measured mass of a nucleus with Z protons and N neutrons is not the simple sum of the constituents but is less because of the energy used up in binding the system together. The binding energy of a nucleus is always positive and can be written

as in Eq. (2), where M stands for mass. Thus, the

$$\text{Binding energy } (A, Z, N) = M(\text{component protons, neutrons, and electrons})c^2 - M(\text{assembled nucleus})c^2 \quad (2)$$

binding energy is the energy required to separate all the components to infinity. Another expression of the energy involved in holding the nuclear system together is the mass excess, given by Eq. (3), which may be positive or negative.

$$\text{Mass excess } (A, Z, N) = M(\text{assembled nucleus})c^2 - Auc^2 \quad (3)$$

Figure 3 shows the variation of average binding energy with mass number A. The shape of this curve can be reproduced by the Bethe-Weizsacker mass formula, which is noteworthy for its simplicity in reproducing the overall binding energy systematics. The formula is developed by modeling the nucleus on a liquid drop. By analogy with a drop of liquid, there is an attractive volume term, which depends on the number of particles (A);a repulsive surface-tension term (proportional to $A^{2/3}$, the area of the nuclear surface); and a term due to the mutual Coulomb repulsion of protons, which is responsible for the decrease in binding energy for heavy nuclei. To make the model reproduce experimental measurements requires further terms reflecting neutron-proton asymmetry [proportional to $(N - Z)^2/A$] and reflecting odd-even pairing effects.

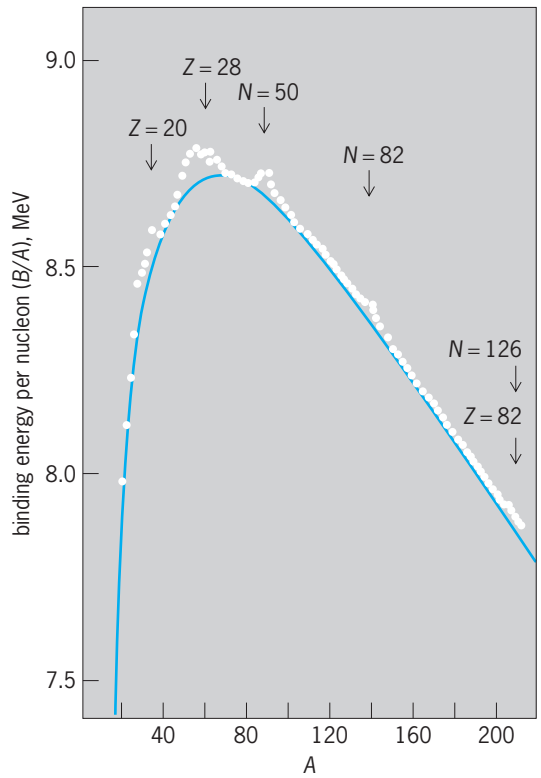

Fig. 3. Binding energy *B* per nucleon plotted as a function of mass number *A*. The vertical axis is offset (it begins near 7.5 MeV instead of at 0 MeV) to highlight shell structure. (*After W. N. Cottingham and D. A. Greenwood, An Introduction to Nuclear Physics, Cambridge University Press, 1986*)

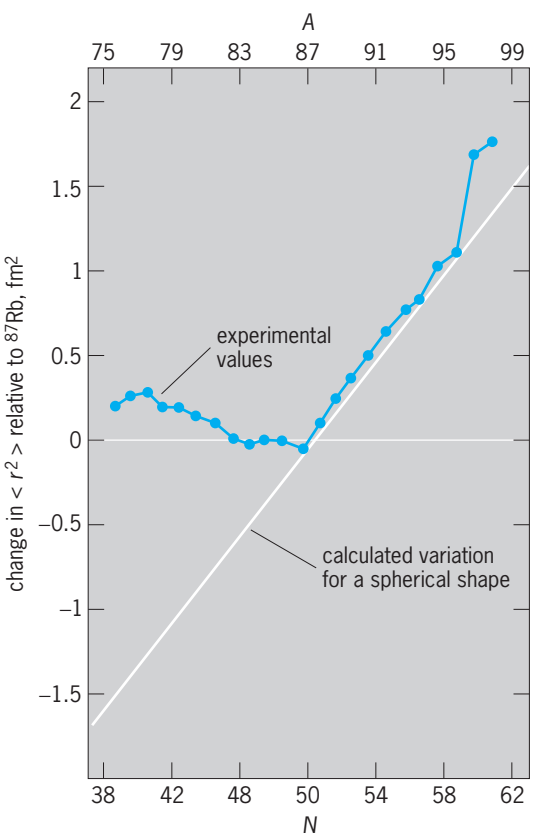

Fig. 2. Variation of the mean-square radius (r^2) of rubidium isotopes (as measured by laser hyperfine spectroscopy) compared with the trend expected for spherical nuclei. The increase of mean radius due to deformation in the light nuclei is evident. (*After C. Thibault et al., Hyperfine structure and isotope shift of the D_2 line of $^{76-98}$Rb and some of their isomers, Phys. Rev., C23:2720–2729, 1981*)

The model is spectacularly successful in reproducing the overall trends in nuclear binding energies, masses, and the energetics of nuclear fission, and in predicting the limits of stability where neutrons and protons become unbound. As in the case of predicting a mean nuclear shape, a comparison of the prediction of the Bethe-Weizsacker mass formula to measured masses shows periodic fluctuations with both N and Z. These are due to the quantum shell effects that cause deviations of binding energy from the smooth liquid-drop value and are clearly visible in Fig. 3 as irregularities at nuclear numbers 20, 28, 50, and 82. In order to incorporate these quantum effects a discussion of how nuclear systems carry energy and angular momentum is necessary. *See* NUCLEAR FISSION.

Nuclear excited states. The quantization of energy and angular momentum, which is apparent in all branches of subatomic physics, is drastically evident in nuclei. The small nuclear size and tightly bound nature impose very restrictive constraints on the orbits that protons and neutrons can undergo inside

the system. Thus, each nucleus has a series of quantum states that particles can occupy. Each state corresponds to a well-defined kinetic energy, rotational frequency (angular momentum), orbital orientation, and reflection symmetry (parity). These characteristics can be denoted by a series of quantum numbers that label the properties of the state. In addition to the quantum numbers of the states of the nucleus, the protons and neutrons have their own internal quantum numbers arising from their quark structure. These quantum numbers describe charge, internal angular momentum (spin), and magnetic moment. The Pauli principle requires that each particle have a unique set of quantum labels. Each nuclear state can then be filled with four particles: protons with internal angular momentum "up" and "down," and likewise two neutrons. *See* ANGULAR MOMENTUM; ENERGY LEVEL (QUANTUM MECHANICS); EXCLUSION PRINCIPLE; PARITY (QUANTUM MECHANICS); QUANTUM NUMBERS; QUARKS; SPIN (QUANTUM MECHANICS).

A nucleus is most stable when all of its nucleons occupy the lowest possible states without violating this occupancy rule. This is called the nuclear ground state. During nuclear collisions the protons and neutrons can be excited from their most bound states and promoted to higher-lying unoccupied states. The process is usually very short-lived and the particles deexcite to their most stable configuration on a time scale of the order of 10^{-12} s. The energy is usually released in the form of gamma rays of well-defined energy corresponding to the difference in energy of the initial and final nuclear states. Occasionally, gamma decay is not favored because of angular momentum selection rules, and long-lived nuclear isomers result. *See* GAMMA RAYS; NUCLEAR ISOMERISM; NUCLEAR SPECTRA; SELECTION RULES (PHYSICS).

Nuclear models. The detailed categorization of the excitation of protons and neutrons allows a mapping of the excited states of each nucleus and determination of its quantum numbers. These data are the essential information required for development of detailed models that can describe the motion of nucleons inside nuclei. The understanding of nuclear behavior advanced rapidly with the discovery and exploration of the wealth and variety of excited nuclear states, both in stable nuclei and in the many radioactive isotopes that could be produced in nuclear reactors and by particle accelerators. Many different modes of excitation were found that involved both collective (that is, involving all or most of the nucleons) and single-nucleon degrees of freedom. Unlike atomic molecules, where rotational, vibrational, and single-particle degrees of freedom involve different time scales and energies, the nucleus is highly complex, with rotation, vibration, and single-particle degrees of freedom being excited at similar energies and often strongly mixed. A full understanding of this behavior remains elusive, and the development of truly generalized nuclear models remains an unfulfilled goal. *See* MOLECULAR STRUCTURE AND SPECTRA.

The measurement of static electric and magnetic moments of nuclear states and of dynamic transition moments has provided a great deal of information. Electric moments have revealed a variety of enhanced collective modes of excitation, including elongated, flattened, and pear-shaped nuclear shapes. Magnetic moments have provided detailed information on the differences between excitations

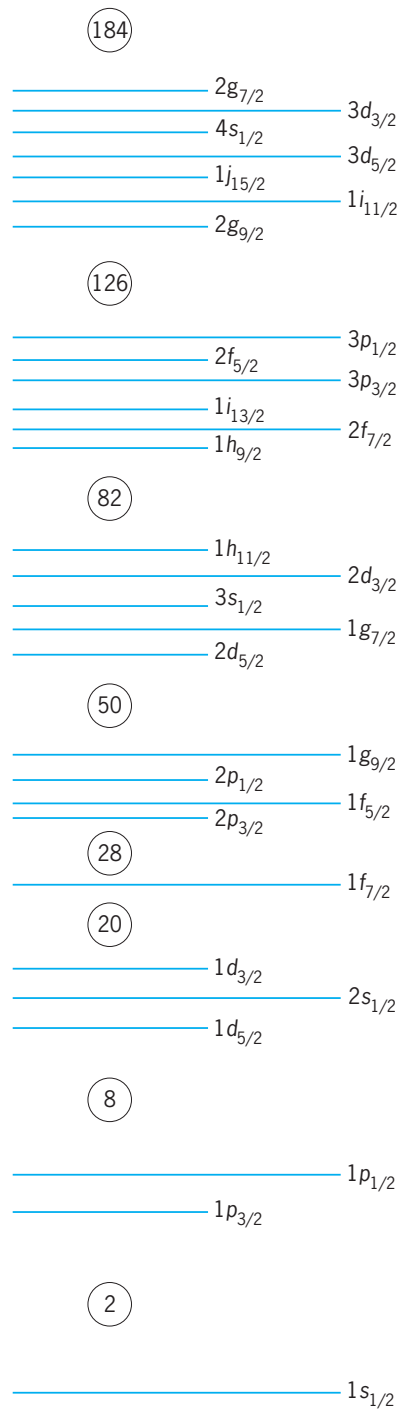

Fig. 4. Sequence of levels in a shell-model potential. The bunching of levels into groups with so-called magic gaps is evident. The numbers in these gaps are total numbers of protons or neutrons (magic numbers). The labeling of levels is according to their orbital angular momentum quantum numbers with $\vec{\ell} = 0, 1, 2, \ldots$, being denoted by the letters *s, p, d, f, g, h, i, j*. The subscripts denote the total angular momentum quantum number *j*. (*After B.L. Cohen, Concepts of Nuclear Physics, McGraw-Hill, 1982*)

involving neutrons (negative moments) and protons (positive moments).

In some particular regions of nuclei with certain well-defined neutron and proton numbers, the nuclear spectrum is drastically simplified when one mode of excitation dominates the quantum states. In these special regions, nuclear models were most successfully developed.

Shell model. For atoms, the solutions of the Schrödinger equation with a Coulomb potential lead to a reasonable prediction of the energies of quantized atomic states, as well as their spins, parities, and moments. Attempts to make the same progress for nuclei, using a variety of spherically symmetric geometric potentials of nuclear dimensions, failed to reproduce known properties until it was realized by M. G. Meyer and independently by J. H. Jensen in 1949 that an additional spin-orbit potential of the form given by Eq. (4) was required to reproduce

$$V_{so} = V_o(\vec{L} \cdot \vec{S}) \tag{4}$$

the known sequence of nuclear states, where $\vec{L}$ is the orbital angular momentum and $\vec{S}$ is the internal spin of the nucleons. A potential of this form binds states having the internal spin of the nucleons parallel to its orbital angular momentum more tightly than when they are antiparallel. The ensuing sequence of quantum shell gaps and subgaps was then correctly reproduced (**Fig. 4**). The shell model has evolved rapidly, and its domain of applicability has widened from the limited regions of sphericity near doubly magic nuclei to encompass most light nuclei with $A < 60$ as well as enlarged regions around shell closures. This evolution has been closely linked with advances in computer technology that allow ever-increasing numbers of particles to be included in more configurations. The development of prescriptions to describe the residual interactions between the valence particles has also been very important.

Modern shell models can be spectacularly precise in their prediction of nuclear states and their wave functions. Calculations of gamma-ray and beta-particle transition rates can be made with great precision, and such detailed reproduction of nuclear wave functions presents the possibility of using the nucleus as a laboratory for searching for forbidden or exotic decay modes. *See* RADIOACTIVITY.

Large-basis shell models provide the most thorough description of nuclei that is available. However, the majority of nuclei still cannot be described by shell models because of the intractable computational problems presented by heavy nuclei, which have both larger numbers of basis states and many valence particles. To describe some of the other modes of nuclear excitation, a variety of models have been developed. **Figure 5** shows the Segrè chart of nuclei plotted as a function of N and Z. The grid of lines is at the characteristic magic numbers where spherical shell effects are important and the shell model is most applicable. Away from these regions, in the areas denoted R, permanent deformation occurs and

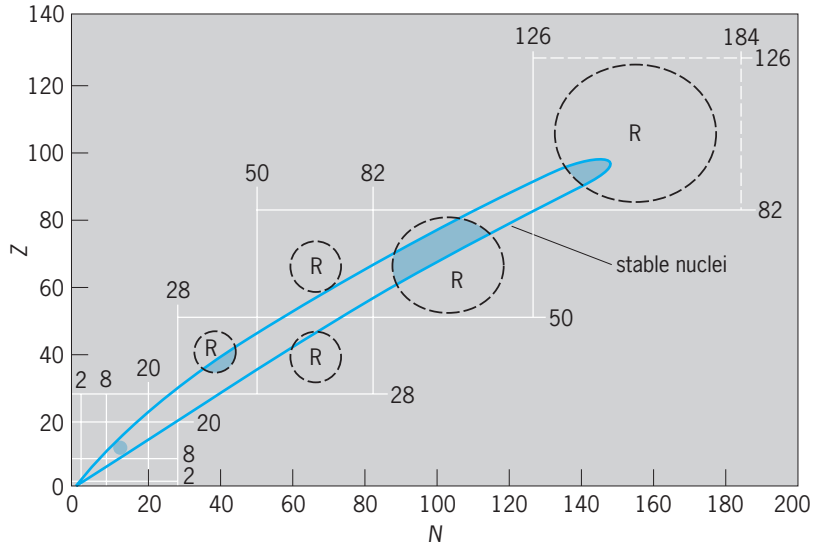

Fig. 5. Segrè chart of nuclides. The grid indicates the major spherical shell gaps. The regions denoted R are the areas where nuclei assume permanent deformation and rotational behavior dominates. The elongated diagonal band encloses the stable nuclei. (*After E. Marshalek, The systematics of deformations of atomic nuclei, Rev. Mod. Phys., 35:108–117, 1963*)

rotational models are most appropriate. *See* MAGIC NUMBERS.

Rotational models. As valence particles are added to a closed core of nucleons, the mutual residual interactions can act coherently and polarize the nuclear system away from sphericity. The polarization effects are strongest when both valence protons and neutrons are involved. The deformed nuclear potential can then undergo collective rotation, which generally involves less energy than other degrees of freedom and thus dominates the spectrum of strongly deformed nuclei. The characteristic spectrum of a deformed quantum system, given by Eq. (5), encountered, where

$$E = \frac{\hbar^2}{2\mathscr{I}} J(J+1) \tag{5}$$

$\hbar$ is Planck's constant divided by 2π, J is the angular momentum of the deformed system, and $\mathscr{I}$ is the moment of inertia, which is considerably less than that of a rigid body because of the highly correlated superfluid nuclear core. The deformation of most nuclear systems that lie far from both neutron and proton shell closures involves elongated, almost axially symmetric shapes, normally parametrized with a quadrupole deformation parameter, β_2. The elongated shape reflects the strong quadrupole correlations in the nucleon-nucleon residual interactions. In the transitional nuclei between shell closures and the highly deformed regions, there is great susceptibility to shape polarization. A wide variety of shapes are found that include triaxially deformed shapes and octupole (pear-shaped) deformations.

Vibrational models. Nuclei undergo collective vibrations about both spherical and deformed shapes. The degree of softness of these vibrations is characterized by the excitation energy required to populate states. The distinguishing feature of vibrational excited states is that they are grouped in nearly

degenerate angular momentum multiplets, each group being separated by a characteristic phonon energy.

Generalized models. It has been a goal of nuclear structure studies to develop models that incorporate all of the features described above in order to produce a unified nuclear picture. The aim is to reproduce and predict the properties of all bound nuclei of any N and Z as well as their evolution with angular momentum and excitation energy. The development of generalized nuclear models has relevance to other fields of physics. There are many isotopes that will never be accessible in the laboratory but may exist in stars or may have existed earlier in cosmological time. The evolution of generalized models greatly increases the power to predict nuclear behavior and provides information that is required for cosmological calculations. Considerable progress has been made in several areas, as discussed below. *See* COSMOLOGY; NUCLEOSYNTHESIS.

Truncated shell models. The expansion of the shell model to span all quantum states and their occupancy with protons and neutrons remains an intractable computational problem. However, a very successful approach to reducing the problem to tractable size has been made by A. Arima and F. Iachello, who considered a truncation incorporating pairs of particles coupled as bosons with angular moment $J = 0$ or $2\hbar$. The symmetries of recoupling these bosons are striking, and, in particular, the so-called limiting symmetries make it possible to predict properties of rotational, vibrational, and triaxial nuclei. The triaxial nuclei are of special interest since the model makes unique predictions about their nuclear deexcitation that have been experimentally demonstrated. The power of this interacting boson approximation (IBA) model appears in its ability to predict the properties of transitional nuclei that lie between the limiting types of nuclear excitation and reproduce systematic trends found across the nuclear periodic table. *See* SYMMETRY LAWS (PHYSICS).

A development of the interacting boson approximation model that separately treats neutron and proton bosons is the IBA2 model, which has been useful in interpreting the very enhanced magnetic dipole (M1) electromagnetic transitions found in electron and gamma-ray fluorescent scattering experiments. One interpretation is that of a new type of collectivity, with deformed proton and neutron distributions oscillating in a so-called scissors mode. *See* MULTIPOLE RADIATION.

Microscopic-macroscopic calculations. A dramatically successful though less rigorous approach to the nuclear many-body problem has been applied to predicting high-spin phenomena, fission barriers, nuclear masses, and the shapes of nuclei far from stability. The deformation of a spherically symmetric potential causes many shell-model degeneracies to be lifted and the excitation energies of orbits to change according to their orientation with respect to the deformation axis. Initially, for small deformations, the level shifting and mixing destroys the normal shell gaps and magic numbers. However, at axis ratios approaching integer values (2:1, 3:1, 3:2, and so forth) and at some critical rotational frequencies, new symmetries appear and produce new quantum shell stabilization.

In order to calculate these new shell effects and their behavior with varying deformation and angular momentum, a combination of liquid droplet and deformed rotating shell models has been used. The approach has been to separate the total nuclear binding energy into two parts that are both shape-dependent but can be calculated independently. The overall bulk trends in binding energy and its evolution with N, Z, shape, and angular momentum are incorporated into a semiclassical model based on the dynamics of a liquid drop. Separately, the energy of quantum states at the same conditions of N, Z, shape, and spin are calculated, and the total binding energy from the single-particle energies is calculated. By using a formalism due to V. Strutinsky, the modulating effects of nuclear shell structure are calculated relative to the mean bulk properties, and the total binding energy can then be estimated. This method of calculation appears to provide a large advance in unifying nuclear models. One example of this type of calculation was used to predict the ground-state masses and shapes of all nuclei (more than 4000) predicted to be bound. Another version predicted the existence of so-called superdeformed nuclei with an axis ratio of 2:1, stabilized by deformed-shell effects. The observation of this exotic shape, through the characteristic gamma-ray patterns arising from its rotation (**Fig. 6**), provided a dramatic verification of this type of model.

The most stringent tests of these models appear to involve nuclei that are at the limits of nuclear

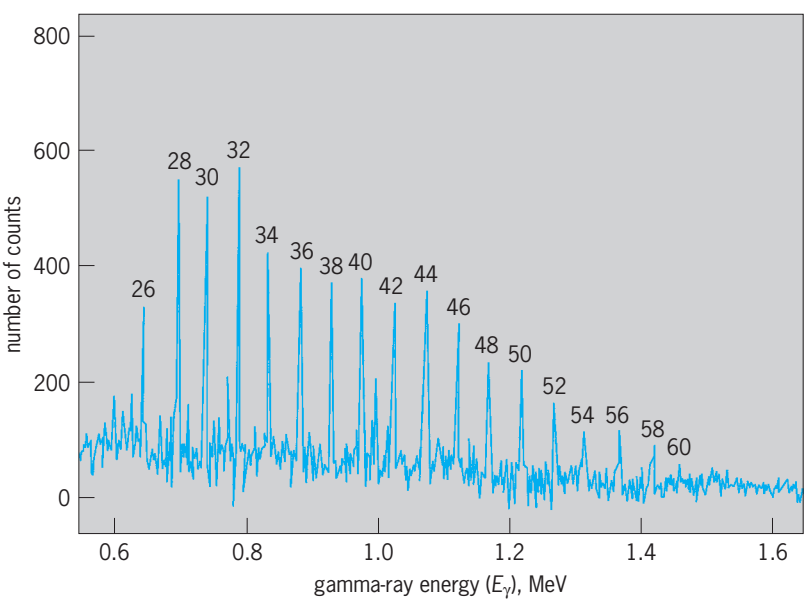

Fig. 6. Rotational sequence of gamma rays associated with superdeformation. The equal spacing between transitions is characteristic of an elongated nucleus with an axis ratio of 2:1. In this case, the rotating nucleus is dysprosium-152 ($^{152}_{66}$Dy), which was formed in the fusion of a beam of calcium-48 ($^{48}_{20}$Ca) ions with a palladium-108 ($^{108}_{46}$Pd) target. Numbers above peaks indicate the total angular momentum quantum number *J* of the decaying state. (*After P.J. Twin et al., Observation of a discrete line superdeformed band up to 60ħ in* ^{152}Dy, *Phys. Rev. Lett., 57:811–814, 1986*)

stability, either because of highly unstable neutron-to-proton ratios (N/Z) or because of extremes of rotation (where Coriolis forces stress the nucleus) and temperature (where the highly organized correlations of nuclei break down and statistical behavior takes over). At these limits of stability, the subtle quantum effects can have a dominant influence on the total nuclear stability and can become the determining factor in maintaining stability. Much effort, both theoretical and experimental, is focused on these areas of marginal stability. Reactions involving the fusion of heavy ions can produce nuclei in a wide variety of states, including those of high angular momentum, neutron-to-proton ratio, and temperature. However, to attribute variations of nuclear behavior to any one of these degrees of freedom and to identify shell effects requires isolating specific isotopes and measuring their state of spin and temperature. To this end, large arrays of detectors subtending large solid angles and identifying all particles and photons emitted from nuclear reactions have been constructed in many laboratories. *See* EXOTIC NUCLEI; NUCLEAR REACTION.

Nuclei at high excitation energies. As nuclei are excited to ever higher excitation energies, it is anticipated that shell effects will be slowly replaced by statistical, or chaotic, behavior. The number of states per megaelectronvolt of excitation energy with each spin and parity rise exponentially with increasing excitation energy until the levels become sufficiently close that they overlap and mix strongly and so become a continuum of states. The extent and rate at which this mixing alters observable quantities such as electromagnetic moments is of considerable interest. The region of excitation that lies from a few megaelectronvolts to a few tens of megaelectronvolts above the yrast states, defined to be the lowest energy states of each spin, is under intense investigation.

Toward the top of this energy regime, new modes of nuclear collectivity become accessible. Giant resonance states can be excited that involve compression and oscillation of the nuclear medium with vibrations of the protons and neutrons in phase (isoscalar) or beating against each other (isovector). The excitation and decay of these giant resonances can provide information about shapes of nuclei at high excitation and about the compressibility of nuclear matter. Results from giant resonance studies indicate that the shell effect persists high into the region previously thought to be statistical. *See* GIANT NUCLEAR RESONANCES.

The semiclassical statistical and hydrodynamic behavior of hot nuclear matter and its experimental, theoretical, and astrophysical aspects are of great interest at the highest nuclear energies. The influence of compression and heat on the density of nuclear matter is being investigated in order to measure a nuclear equation of state in analogy with the properties of a classical fluid. It has been suggested that the nuclear matter may undergo phase changes under compression, with high-density condensates possibly providing a new metastable state. At the highest densities and temperatures, the nucleons themselves are forced to overlap and merge, leading to a plasma of quarks and gluons that are the nucleonic constituents. At present, experimental evidence for deconfined quark matter has not been found. A new generation of accelerators and detectors is undergoing development that will allow research to commence in this ultrahigh-energy domain that involves colliding beams of the very heaviest elements with energies of thousands of megaelectronvolts in order to reach this state of matter that may have previously only existed during the first few fractions of a second at the beginning of the universe. *See* QUARK-GLUON PLASMA; RELATIVISTIC HEAVY-ION COLLISIONS.

C. J. Lister

Bibliography. A. Bohr and B. Mottelson, *Nuclear Structure*, vol. 1, 1969, vol. 2, 1975; H. A. Enge and R. P. Redwine, *Introduction to Nuclear Physics*, 1995; K. S. Krane, *Introduction to Nuclear Physics*, 1987; J. D. Walecka, *Theoretical Nuclear and Subnuclear Physics*, 1995.

Nucleation

The formation within an unstable, supersaturated solution of the first particles of precipitate capable of spontaneous growth into large crystals of a more stable solid phase. These first viable particles, called nuclei, may either be formed from solid particles already present in the system (heterogeneous nucleation), or be generated spontaneously by the supersaturated solution itself (homogeneous nucleation). *See* SUPERSATURATION.

Heterogeneous nucleation involves the adsorption of dissolved molecules onto the surface of solid materials such as dust, glass, and undissolved ionic substances. This adsorbed layer of solute molecules may then grow into a large crystal. Because the crystal lattice of the foreign solid is in general not the same as that of the solid to be precipitated, the first few layers are deposited in a lattice configuration which is strained, that is, less stable than the normal lattice of the precipitating material. The degree of lattice strain determines the effectiveness of a given heterogeneous nucleating agent. Thus, a material whose crystal structure is greatly different from that of the solid to be precipitated will not bring about precipitation unless the solution is fairly highly supersaturated, whereas, if the solution is seeded by adding small crystals of the precipitating substance itself, precipitation can occur at a concentration only slightly higher than that of the saturated solution.

If elaborate precautions are taken to exclude solid particles, it is possible to obtain systems in which the necessary precipitation nuclei are spontaneously generated within the supersaturated solution by the process of homogeneous nucleation. In a solution, ions interact with each other to form clusters of various sizes. These clusters in general do not act as nuclei, but instead, redissociate into ions. However, if the solution is sufficiently supersaturated so that its tendency to deplete itself by deposition of ions onto

the clusters overcomes the tendency of the clusters to dissociate, the clusters may act as nuclei and grow into large crystals. The rate at which suitable nuclei are generated within the system is strongly dependent upon the degree of supersaturation. For this reason, solutions which are not too highly supersaturated appear to be stable indefinitely, whereas solutions whose concentration is above some limiting value (the critical supersaturation) precipitate immediately.

Nucleation is significant in analytical chemistry because of its influence on the physical characteristics of precipitates. Processes occurring during the nucleation period establish the rate of precipitation, and the number and size of the final crystalline particles. *See* COLLOID; PRECIPITATION (CHEMISTRY).

David H. Klein; Louis Gordon

Nucleic acid

An acidic, chainlike biological macromolecule consisting of multiply repeated units of phosphoric acid, sugar, and purine and pyrimidine bases. Nucleic acids as a class are involved in the preservation, replication, and expression of hereditary information in every living cell. There are two types: deoxyribonucleic acid (DNA) and ribonucleic acid (RNA).

Deoxyribonucleic Acid

Each DNA strand is a long polymeric molecule consisting of many individual nucleotides linked end to end. The great size and complexity of DNAs are indicated by their molecular weights, which commonly range in the hundreds of millions. DNA is the chemical constituent of the genes of an organism, and is thus the ultimate biochemical object of the study of genetics. Information is contained in the DNA in the form of the sequence of nucleotide building blocks in the nucleic acid chain.

Nucleotides. The number of nucleotide building blocks in DNA is relatively small—only four nucleotides constitute the vast majority of DNA polymeric units. These are deoxyadenylic, deoxyguanylic, deoxycytidylic, and deoxythymidylic acids. For purposes of brevity, these nucleotides are symbolized by the letters A, G, C, and T, respectively. Each of these nucleotides consists of three fundamental chemical groups: a phosphoric acid group, a deoxyribose 5-carbon sugar group (**1b**), and a nitrogenous base which is a derivative of either purine (**2**) or pyrimidine (**3**). Some nucleotides contain the purine groups adenine (6-aminopurine; **4**) or guanine (2-amino-6-oxypurine; **5**), and some contain the pyrimidine groups cytosine (2-oxy-5-aminopyrimidine; **6**) or thymine (2,6-dioxy-6-methyl-pyrimidine; **7**). These are the only major bases found in most DNA, although in specific sequences certain methylated derivatives of these bases, such as 5-methyl cytosine (**8**) or N^6-methyl adenine (**9**), can also be detected. In each nucleotide, these subunits are linked together in the following order: purine or pyrimidine base–ribose sugar–phosphoric acid (**10**)–(**12**).

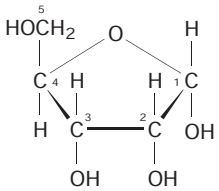

D-Ribose
(1a)

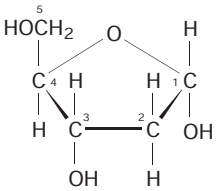

2-Deoxy-D-ribose
(1b)

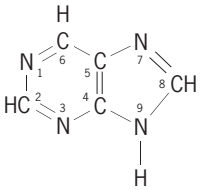

(2)

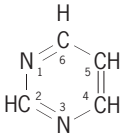

(3)

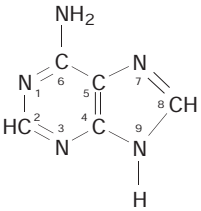

(4)

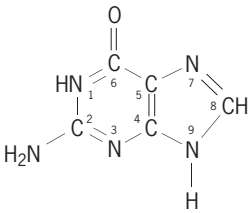

(5)

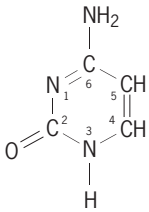

(6)

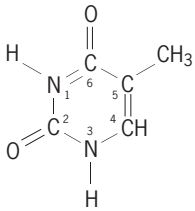

(7)

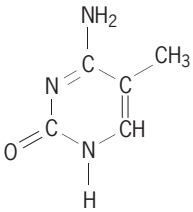

(8)

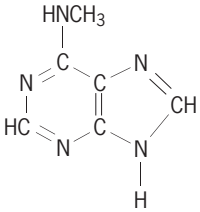

(9)

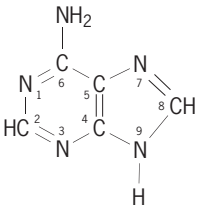

(10)

Removal of the phosphoric acid group leaves a base-sugar compound which is called a nucleoside (**11**). In each nucleoside the base is attached to the sugar through a bond from nitrogen at position 9 (for purines) or 3 (for pyrimidines) tocarbon at position

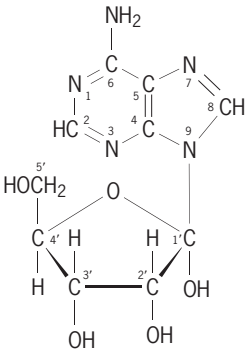

(11)

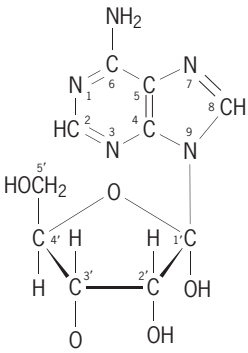

(12)

1 of the sugar ring. The nucleosides and nucleotides are named for the base they contain as follows:

Base	*Nucleoside*	*Nucleotide*
Adenine (**10**)	Adenosine (**11**)	Adenylic acid (**12**)
Guanine	Guanosine	Guanylic acid
Cytosine	Cytidine	Cytidylic acid
Thymine	Thymidine	Thymidylic acid
Uracil	Uridine	Uridylic acid

It is necessary to denote the position of the phosphoric acid residue when describing nucleotides. Nucleotides synthesized by cells for use as building blocks in nucleic acids all have phosphoric acid residues coupled to the 5′ position of the ribose sugar ring, as shown for deoxyadenylic acid (deoxyadenosine-5′-phosphate; **12**). Upon hydrolysis of DNA, however, nucleotides can be produced which have phosphoric acid coupled to the 3′ position of the sugar ring.

When DNA is hydrolyzed by using the enzyme deoxyribonuclease I, prepared from bovine pancreas, the principal products are oligonucleotides ending with 3′-hydroxyl groups and 5′-phosphoric acid groups. In contrast, when DNA is hydrolyzed by using the enzyme micrococcal nuclease, prepared from *Staphylococcus aureus*, the principal products are nucleotides or oligonucleotides ending with 3′-phosphoric acid groups and 5′-hydroxyl groups. Studies such as these led very early to the conclusion that in intact DNA the nucleotides were linked via phosphoryl groups which join the 3′ position of one sugar group to the 5′ position of the next sugar group (**13**). This 5′-to-3′ linkage of nucleotides imparts a polarity to each DNA strand which is an important factor in the ability of DNA to form the three-dimensional structures necessary for its ability to replicate and to serve as a genetic template. *See* NUCLEOTIDE.

Helix. In 1953 James Watson and Francis Crick proposed a double-helical structure for DNA based largely on x-ray diffraction studies of Maurice Wilkins and Rosalind Franklin. It had been known earlier that in most DNAs the ratios of A to T and of G to C are approximately 1:1. The Watson-Crick model attributes these ratios to the phenomenon of base pairing, in which each purine base on one strand of DNA is hydrogen-bonded to a complementary pyrimidine base in an opposing DNA strand (**14**). The x-ray

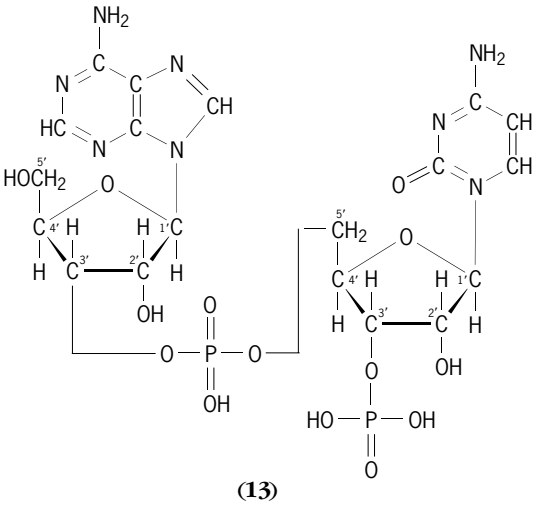

(13)

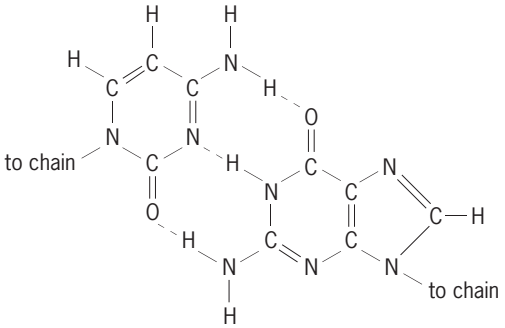

(14)

diffraction studies suggested that DNA can assume a structure called the B form, which is a right-handed helical configuration resembling a coiled spring. In most DNAs there are two single DNA strands which compose each helix. The strands wind about each other, with their sugar-phosphate chains forming the coil of the helix and with their bases extending inward toward the axis of the helix. The configuration of the bases allows hydrogen bonding between opposing purines and pyrimidines. Each of the base pairs lies in a plane at approximately right angles to the helix axis, forming a stack with the two sugar-phosphate chains coiled around the outside of the stack. In this configuration a wide (major) groove exists on the surface of the helix, winding parallel to a narrow (minor) groove (**Fig. 1**). In the Watson-Crick model the two opposing DNA strands have an opposite polarity; that is, they are antiparallel, with their 5′ ends lying at opposite ends of each double-stranded molecule. The right-hand helix completes one turn every 10 nucleotides, and bases are spaced 0.34 nanometer apart. The width of the double-stranded helix in the B form averages about 2 nm. DNA can exist in helical structures other than the B form. One configuration, termed the Z form, is a left-handed helical structure with 12 nucleotides per turn. The Z form can exist in DNA sequences with alternating guanine and cytosine bases and may be functional in localized DNA regions, although the B form is thought to predominate in most biological systems.

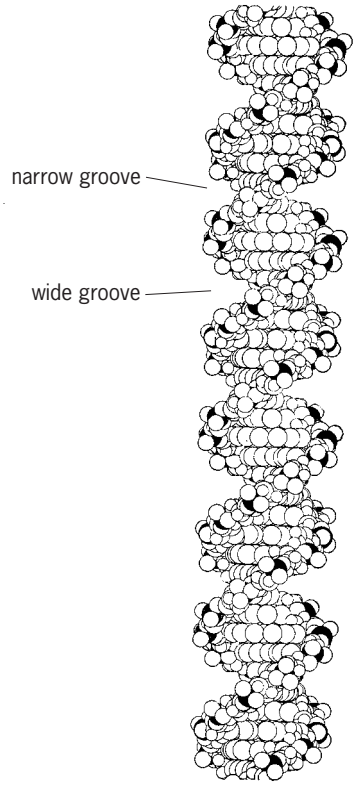

Fig. 1. Three-dimensional configuration of DNA.

Due to the phenomenon of base pairing, the two DNA strands of a helix can be "melted" apart (denatured) or, once denatured, allowed to reanneal. Heating a DNA duplex in solution breaks the hydrogen bonds involved in base pairing and allows the two single strands of the helix to dissociate. This melting is dependent on the base sequence in the DNA, since there are two hydrogen bonds between A-T pairs and three hydrogen bonds between G-C pairs (**14**). In a solution of sheared DNA molecules, the rate of reannealing of denatured DNA depends on the extent too which the denatured DNA sequences are reiterated. Studies on the kinetics of renaturation of DNA from various sources have shown that different organisms differ widely in the complexity of sequences in their DNA. In general, cells of higher organisms contain more unique or seldom-repeated DNA sequences than do cells of lower organisms. This reflects in part the fact that higher organisms possess more individual genes required to code for the larger number of proteins they contain. Renaturation studies also show that many cells possess highly reiterated DNA sequences that may constitute a significant percentage of the total genome. When sheared cellular DNA sequences are separated according to their buoyant density, such as by centrifugation in density gradients of cesium chloride, many of these highly reiterated sequences can be detected as satellite DNA bands; that is, they form bands of DNA at positions in the gradient which differ from that of the bulk of cellular DNA. Several of the known satellite DNAs contain genes which are present in multiple copies in cells. For example, the genes coding for RNAs of ribosomal subunits of cells are characteristically located in satellite DNA bands. At this point, however, the functional significance of most highly reiterated DNA sequences in cells is not known.

C value. The amount of DNA in the haploid genome of a given organism (called the C value) is only loosely correlated with the degree of evolutionary advancement of the organism. There are about 2×10^{-16} g of DNA in a bacteriophage, as compared to about 10^{-14} g in the bacterium *Escherichia coli* and about 3×10^{-12} g in rat liver cells. Whereas mammalian cells contain about $2–3 \times 10^9$ nucleotide pairs of DNA, amphibian cells vary widely, ranging from less than 2×10^9 to about 1.5×10^{11} nucleotide pairs. Consideration of these figures leads to what is known as the C value paradox: generally cells of higher organisms contain much more DNA than is necessary to code for the number of proteins they contain. This paradox is not explained simply by different degrees of gene reiteration in different organisms. Instead, it is more likely that only a fraction of the total DNA codes for proteins in higher organisms and that the relative amount of noncoding DNA is highly variable. It is now known that many noncoding DNA sequences exist adjacent to RNA-coding sequences and serve in a regulatory capacity. In addition, noncoding DNA sequences, termed intervening sequences, may exist in the middle of sequences coding for certain RNA species.

Sequences. The sequence of nucleotide pairs in the DNA determines all of the hereditary characteristics of any given organism. The DNA acts as a template which is copied, through the process of transcription, to make RNA. The RNA in turn serves as a template in a process by which its encoded information is translated to determine the amino acid sequences of proteins. Each amino acid in a protein chain is specified by a triplet of nucleotides (in RNA) or nucleotide pairs (in DNA) known as a codon. The set of correlations between the amino acids and their specifying codons is called the genetic code. Each gene which codes for a protein thus contains a sequence of triplet codons which corresponds to the sequence of amino acids in the polypeptide. This sequence of codons may be interrupted by intervening DNA sequences so that the entire coding sequence is not continuous. In addition to coding sequences, there also exist regulatory sequences, which include promoter and operator sequences involved in initiating gene transcription and terminator sequences involved in stopping transcription. Regulatory sequences are not necessarily made up of triplets, as are the codons. In order to study the regulation of a given gene, it is necessary to determine its nucleotide sequence. *See* GENETIC CODE.

There are several methods for sequencing DNA. Most of these methods employ radioactive end-labeling of one or both DNA strands, followed by either cleavage or extension of labeled strands to produce end-labeled fragments which terminate at nucleotides with specific bases. For example, one commonly used method involves labeling the 5′ end of a DNA strand with ^{32}P-phosphate by using the

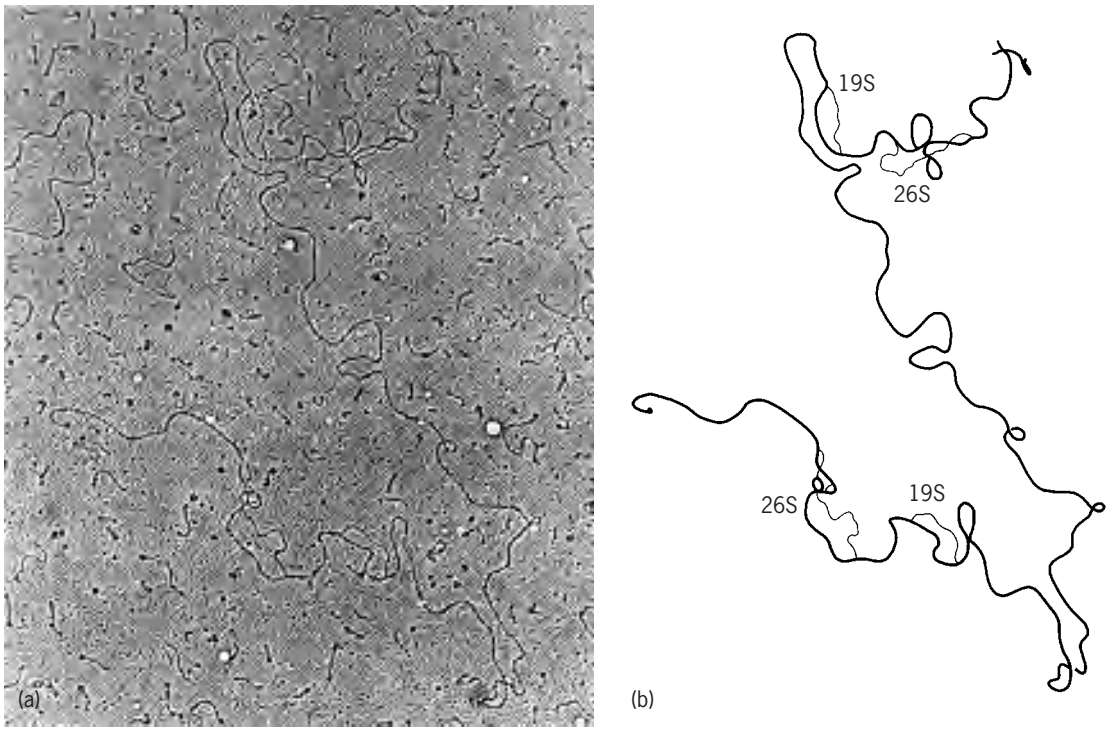

Fig. 2. Ribosomal RNA genes of *Physarum polycephalum*: (*a*) electron micrograph and (*b*) map, showing 19S and 26S RNA coding regions as R-loop hybrids. This DNA molecule is 60,000 base pairs long, or about 20 × 10^{-6} m. (*Courtesy of G. R. Campbell, V. C. Littau, P. W. Melera, V. G. Allfrey, and E. M. Johnson*)

enzyme polynucleotide kinase and γ-^{32}P-adenosine triphosphate as a phosphate donor. Procedures are then used to induce base-specific chemical cleavage of the end-labeled strands at each of the four nucleotides in turn. Polyacrylamide gel electrophoresis is then used to size the radioactive fragments, and the sequence of nucleotides from the labeled 5′ end can be deduced.

The complete DNA sequences of several different genes, together with adjacent regulatory sequences, are now known. The first entire genome of any organism to be sequenced was that of the single-strand DNA phage ϕX174. This sequence of 5386 nucleotides was worked out by Fred Sanger and coworkers. One interesting aspect of this sequence is that it reveals the presence of overlapping genes coding for proteins. A single nucleotide sequence can code for more than one amino acid sequence, depending on the phasing with which the sequence is grouped into triplet codons during the process of protein synthesis. Specific start and stop signal codons are required to specify the phasing.

Figure 2*a* shows an electron micrograph of a single DNA molecule containing the genes coding for ribosomal RNAs of the slime mold *Physarum polycephalum*. The coding sequences for two RNA species, 19S and 26S, are revealed by R-loop hybridization. (This is a process by which RNA is annealed with double-stranded DNA, thus displacing a single DNA strand.) In order to prepare the nucleic acids for visualization, they are first coated with a layer of electron-dense heavy metals—in this case platinum and palladium. **Figure 3** shows the DNA sequence near the start point of transcription of the gene coding for the enzyme β-galactosidase in the bacterium *E coli*. This sequence contains short segments which are important for binding and subsequent action of enzymes that polymerize RNA. Regions for binding of proteins (CAP protein, RNA polymerase, and repressor protein) are shown to the DNA map. The *i* gene codes for the repressor protein which binds to the operator region and regulates transcription. The CAP protein must be bound to the DNA in order for the polymerase to initiate properly. The CAP protein binds its site only in the presence of the compound cyclic adenosine monophosphate. RNA polymerase binds a site about 40 nucleotides long and, in the absence of repressor, begins transcription of messenger RNA. The *z* gene codes for the enzyme β-galactosidase. The DNA sequence of one strand of the polymerase binding region is shown at bottom. The entry site and Pribnow's box sequences are important recognition sequences for polymerase binding and initiation. *See* OPERON.

Enzymes called restriction endonucleases cleave DNA at specific sequences. It is possible, by using these enzymes and DNA ligases, to splice an exogenous gene segment into an existing DNA sequence. The resulting structure is known as recombinant DNA. When the recombinant includes the genome of a plasmid or phage, this spliced DNA molecule can be propagated in bacteria. This process is known as gene cloning. Through this process, genes which are present only in a few copies in mammalian cells can

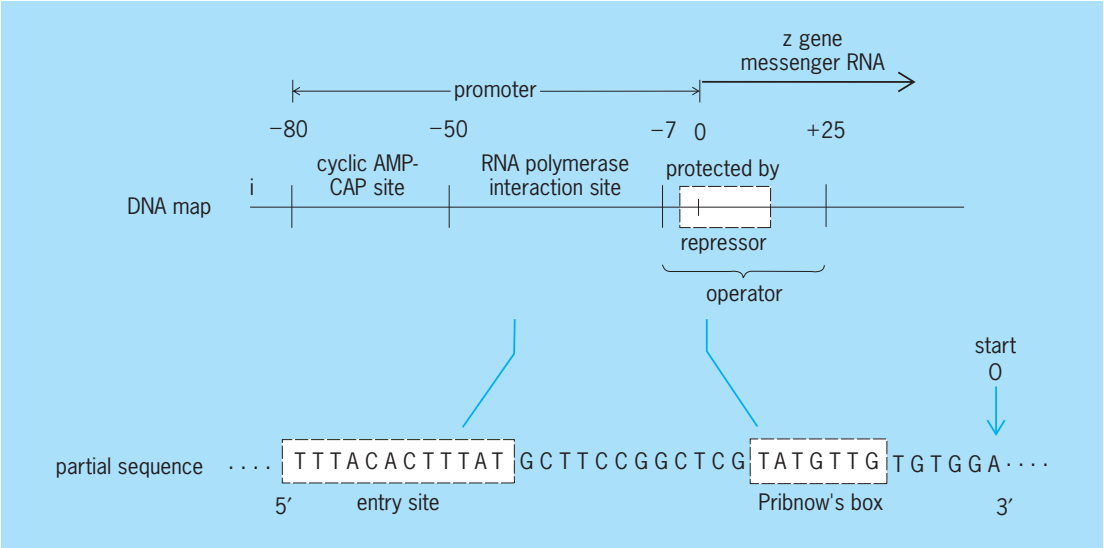

Fig. 3. DNA sequence of the promoter-operator region of the *lac* operon in *Escherichia coli*.

be grown in vast quantities in bacteria. Cloning is industrially important for the production of certain gene products, such as some hormones. *See* GENETIC ENGINEERING.

Replication. Since DNA is the substance containing the chemical code governing the heredity of all cells, it is clear that its biosynthesis and passage from generation to generation must be extremely precise. The way in which a new DNA molecule arises with the same nucleotides arranged in the same sequence as in the parent DNA molecule is one of the most intriguing problems in biochemistry.

In all cells, DNA is synthesized by using parent DNA as a template for a polymerization reaction which requires deoxyribonucleoside 5′-triphosphates as precursors and which is catalyzed by DNA polymerase enzymes. (The deoxyribonucleoside triphosphates are synthesized in cells by addition of three phosphate groups to each nucleoside.) The DNA synthetic reaction can be summarized as reaction (1), where XMP represents nucleoside

$$\text{no1}\ \left.\begin{matrix}\text{dATP}\\ \text{dGTP}\\ \text{dCTP}\\ \text{dTTP}\end{matrix}\right\} \xrightarrow[\text{DNA polymerase}]{\text{DNA, Mg}^{2+}} \text{DNA}\cdot(4\text{dXMP})_n + (4\text{pp}_i)_n$$

monophosphate and pp_i represents inorganic pyrophosphate. The reaction always proceeds in a 5′-to-3′ direction. That is, each deoxyribonucleotide is attached through its 5′-phosphate to the 3′-OH group of an existing DNA strand, with the release of pyrophosphate. This 5′-to-3′ directionality is true for all known nucleic acid biosynthetic reactions. All DNA polymerase enzymes require an existing strand of DNA as a template, with a base-paired segment to serve as a primer. The primer sequences may consist of ribonucleotides, but the template is DNA. No DNA polymerase is capable of synthesizing DNA de novo. The primer must possess a free 3′-OH group for synthesis to begin. Through base pairing, nucleotides complementary to those in the template strand are added (**15**).

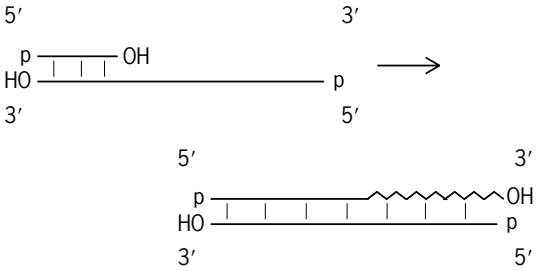

(**15**)

During DNA replication in cells, each strand of the helix serves as a template for synthesis. The resulting coupling of each new DNA strand with a preexisting strand is known as semiconservative replication. At a point along the double helix the strands unwind and, after attachment of primers, synthesis proceeds bidirectionally at what is termed the replication fork. When viral DNA is replicated in *E. coli*, the newly synthesized DNA, called nascent DNA, occurs in short fragments, now called Okazaki fragments (**16**). The action of enzymes known as DNA ligases is required to join the Okazaki fragments. In fact, it is now known that many enzymes in addition to DNA polymerase are involved in the process of DNA replication in cells. Studies on bacterial mutants defective in DNA synthesis have implicated more than a dozen such enzymes.

There are differences between DNA synthetic mechanisms in bacteria and higher organisms. In lower organisms such as bacteria or viruses, DNA replication begins at a single point in the genome and proceeds bidirectionally to replicate all the DNA. In higher organisms, replication begins at many points, thus forming numerous "eyes." The parental DNA strands are separated and replicated until these eyes meet. The DNA polymerases of bacteria comprise several different enzymes, each of which is involved in a separate aspect of replication. The DNA

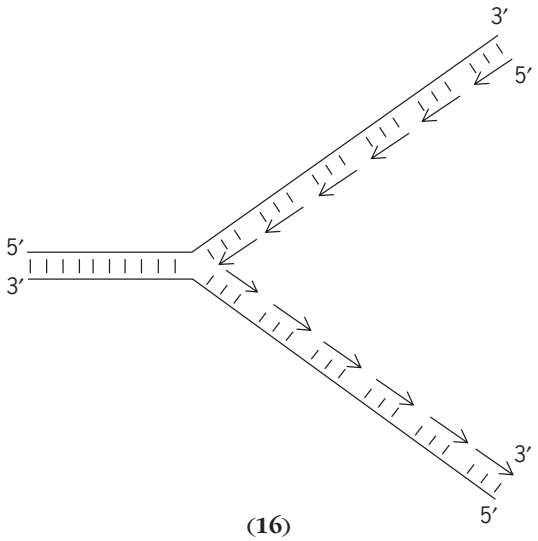

(16)

polymerases of higher organisms comprise a still different set of enzymes.

Function. In every living cell, as well as in certain viruses and subcellular organelles, the function of DNA is similar: it encodes genetic information and replicates to pass this information to subsequent generations. The nucleotide sequence of DNA in each organism determines the nature and number of proteins to be synthesized, as well as the organization of the protein-synthesizing apparatus. The first conclusive evidence that DNA alone can direct cell function was obtained in bacterial cells. In 1944 O. T. Avery, C. M. MacLeod, and M. McCarty showed that purified DNA from a strain of smooth, polysaccharide-encapsulated *Pneumococcus* could transform rough, nonencapsulated strains so that they and their offspring resemble the DNA donor. Subsequent studies on bacterial transformation showed that numerous other hereditary characteristics could be transferred between strains using pure DNA. G. Beadle and E. L. Tatum showed that mutations in DNA of the mold *Neurospora* resulted in alterations in the amino acid sequence of the protein. The entire process of gene expression, by which the flow of information proceeds from DNA to RNA to protein, remains one of the most fertile areas of molecular biological research. *See* DEOXYRIBONUCLEIC ACID (DNA).

Ribonucleic Acid

RNAs are long polymeric chains of ribonucleotides joined end to end in a 5′-to-3′ linkage. The primary chemical difference between RNA and DNA is in the structure of the ribose sugar of the individual nucleotide building blocks: RNA nucleotides possess a 2′-OH group, whereas DNA nucleotides do not (**1**). Another major chemical difference between RNA and DNA is the substitution of uridylic acid, which contains the base uricil (2,6-dioxypyrimidine; **17**) for thymidylic acid as one of the four nucleotide building blocks. Thus incorporation of radioactive uridine can be used as a specific measure of RNA synthesis in cells, while incorporation of radioactive thymidine can be used as a measure of DNA synthesis.

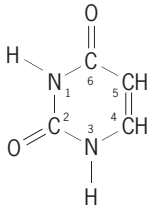

(17)

Further modifications of RNA structure exist, such as the attachment of various chemical groups (for example, isopentenyl and sulfhydryl groups) to purine and pyrimidine rings, methylation of the sugars (usually at the 2′ position), and folding and base pairing of sections of a single RNA strand to form regions of secondary structure. Unlike DNA, nearly all RNA in cells is single-stranded (except for regions of secondary structure) and does not consist of double-helical duplex molecules. Another distinguishing characteristic of RNA is its alkaline lability. In basic solutions, RNA is hydrolyzed to form a mixture of nucleoside 2′- and 3′-monophosphates. In contrast, DNA is stable to alkali.

Classes. Cellular RNA consists of classes of molecules widely divergent in size and complexity.

Ribosomal RNA. The most abundant class of RNA in cells is ribosomal RNA. This class comprises those molecular species that form part of the structure of ribosomes, which are components of the protein-synthesizing machinery in the cell cytoplasm. The predominant RNA molecules are of size 16S and 23S in bacteria (the S value denotes the sedimentation velocity of the RNA upon ultracentrifugation in water), and 18S and 28S in most mammalian cells. In most mammalian cells the 18S RNA is approximately 2100 nucleotides long, and the 28S RNA approximately 4200 nucleotides. In the large subunit of ribosomes, the 28S RNA is associated with a smaller 5S RNA species which is about 120 nucleotides long. In eukaryotic cells another small RNA species, 5.8S RNA, is also associated with 28S RNA. All of the ribosomal RNA species are relatively rich in G and C residues. For example, the 28S RNA of the newt *Triturus* contains 60.9% G and C residues. As they exist in ribosomes, the ribosomal RNAs contain considerable intrachain base pairing which helps maintain a folded three-dimensional structure. *See* RIBOSOMES.

Messenger RNA. Another prominent class of RNAs consists of the messenger RNA molecules and their synthetic precursors. Messenger RNAs are those species that code for proteins. They are transcribed from specific genes in the cell nucleus, and carry the genetic information to the cytoplasm, where their sequences are translated to determine amino acid sequences during the process of protein synthesis. The messenger RNAs thus consist primarily of triplet codons. Most messenger RNAs are derived from longer precursor molecules which are the primary products of transcription and which are found in the nucleus. These precursors undergo several steps known as RNA processing which result in production of cytoplasmic messenger molecules ready for

translation. Processing steps frequently include addition of a 7-methyl guanosine residue to the 5′ terminal triphosphate group, splicing-out of intervening sequences in the coding regions, and addition of multiple adenylic acid residues (often as many as 200) to the 3′ terminus of the polynucleotide chain. Certain messengers, such as those coding for histones, do not possess a polyadenylic acid tail. The messenger RNAs as a class possess a high degree of sequence complexity and size diversity, stemming from the fact that they are largely the products of uniquely represented structural genes. Most messenger RNAs range from 1500 to 3000 nucleotides in length. The relative size of their synthetic precursors varies widely, ranging from only slightly longer to more than twice as long as the final messenger RNA product.

Transfer RNA. The third major RNA class is transfer RNA. There are approximately 80 transfer RNA molecular species in the bacterium *E. coli*. These are small RNA molecules 73–93 nucleotides long which possess a relatively high proportion of modified and unusual bases, such as methylinosine or pseudouridine. Each transfer RNA molecule possesses an anticodon and an amino acid binding site. The anticodon is a triplet complementary to the messenger RNA codon for a particular amino acid. A transfer RNA molecule bound to its particular amino acid is termed charged. The charged transfer RNAs participate in protein synthesis: through base pairing they bind to each appropriate codon in a messenger RNA molecule and thus order the sequence of attached amino acids for polymerization.

In addition to the three major RNA classes, numerous small RNA species are found in the cell nucleus. These RNAs frequently have a high content of unusual bases. Some of them are synthesized in the cell nucleolus. Their function is unknown. Some of the small nuclear RNAs arise from the splicing out of intervening sequences from the transcripts of large interrupted genes.

Sequences. Alanyl transfer RNA (a transfer RNA recognizing the alanine codon) was the first RNA molecule to be sequenced in its entirety (**18**). The sequences of many other RNA molecules and segments of molecules are now known. Of particular interest are sequences near RNA transcription initiation or termination sites, since these provide information about gene-regulatory regions. Comparison of sequences of messenger RNAs and ribosomal RNAs has revealed that base pairing is important in attachment of messenger RNA to the ribosomes during initiation of protein synthesis. In general, RNA sequencing is carried out by cleavage of radioactively labeled RNA with nucleotide-specific nuclease enzymes, followed by two-dimensional separation and further analysis of the labeled cleavage products.

Transcription. The process of RNA biosynthesis from a DNA template is called transcription. Transcription requires nucleoside 5′-triphosphates as precursors, and is catalyzed by enzymes called RNA polymerases. Unlike DNA replication, RNA transcription is not semiconservative: only one DNA strand (the "sense" strand) is transcribed for any given gene, and the resulting RNA transcript is dissociated from the parental DNA strand. The RNA synthetic reaction can be summarized as reaction (2), where 5′-XTP rep-

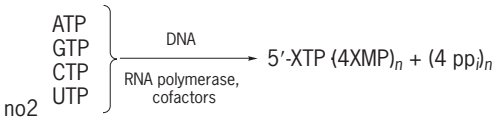

resents nucleoside 5′-triphosphate, XMP represents nucleoside monophosphate, and pp$_i$ represents inorganic pyrophosphate. As is true of DNA synthesis, RNA synthesis always proceeds in a 5′-to-3′ direction. Since RNA synthesis requires no primer, as does DNA synthesis, the first nucleotide in any primary transcript retains its 5′-triphosphate group. Thus an important test to determine whether any RNA molecule is a primary transcript is to see whether it possesses an intact 5′-triphosphate terminus. As in DNA replication, base pairing orders the sequence of nucleotides during transcription. In RNA synthesis uridine, rather than thymidine, base-pairs with adenine.

There are profound differences between bacteria and higher organisms with regard to mechanisms of RNA synthesis and processing. In bacteria, RNA synthesis and translation are coupled; that is, nascent messenger RNA chains are attached directly to ribosomes for protein synthesis. In higher organisms, which possess a distinct cell nucleus, several processing and transport steps intervene between RNA synthesis (in the nucleus) and translation (in the cytoplasm). Bacterial RNA polymerization employs essentially one RNA polymerase enzyme. In contrast, eukaryotic RNA synthesis employs predominantly three RNA polymerases. Messenger RNA synthesis is catalyzed by one enzyme, 5S and transfer RNA synthesis by another, and ribosomal RNA synthesis by still another. All three of these eukaryotic RNA polymerases carry out the same 5′-to-3′ phosphodiester linkage of nucleotides. They differ in their protein subunit composition and in the regulatory DNA sequences they recognize. The recognition of different regulatory sequences by different RNA polymerases under different cellular conditions is one of the most interesting problems in molecular biology. It is this sequence-specific recognition that lies at the heart of the problem of differential gene activity and thus of cell differentiation.

It is known that many proteins cofunction with RNA polymerases to regulate DNA binding and RNA chain initiation. Among these regulatory proteins are several whose activities are influenced by hormones. For example, an estrogen-binding protein regulates transcription of the chick ovalbumin gene, and a cyclic adenosine monophosphate-binding protein regulates RNA polymerase activity at the *lac* operon in *E. coli* (**16**). In addition to these transcriptional regulatory factors, chromosomal proteins called histones exist in complexes with the DNA that can influence the rate of passage of polymerases. These histone-DNA complexes, or nucleosomes, are present on most eukaryotic cellular DNA, and are subject to several biochemical modifications which are

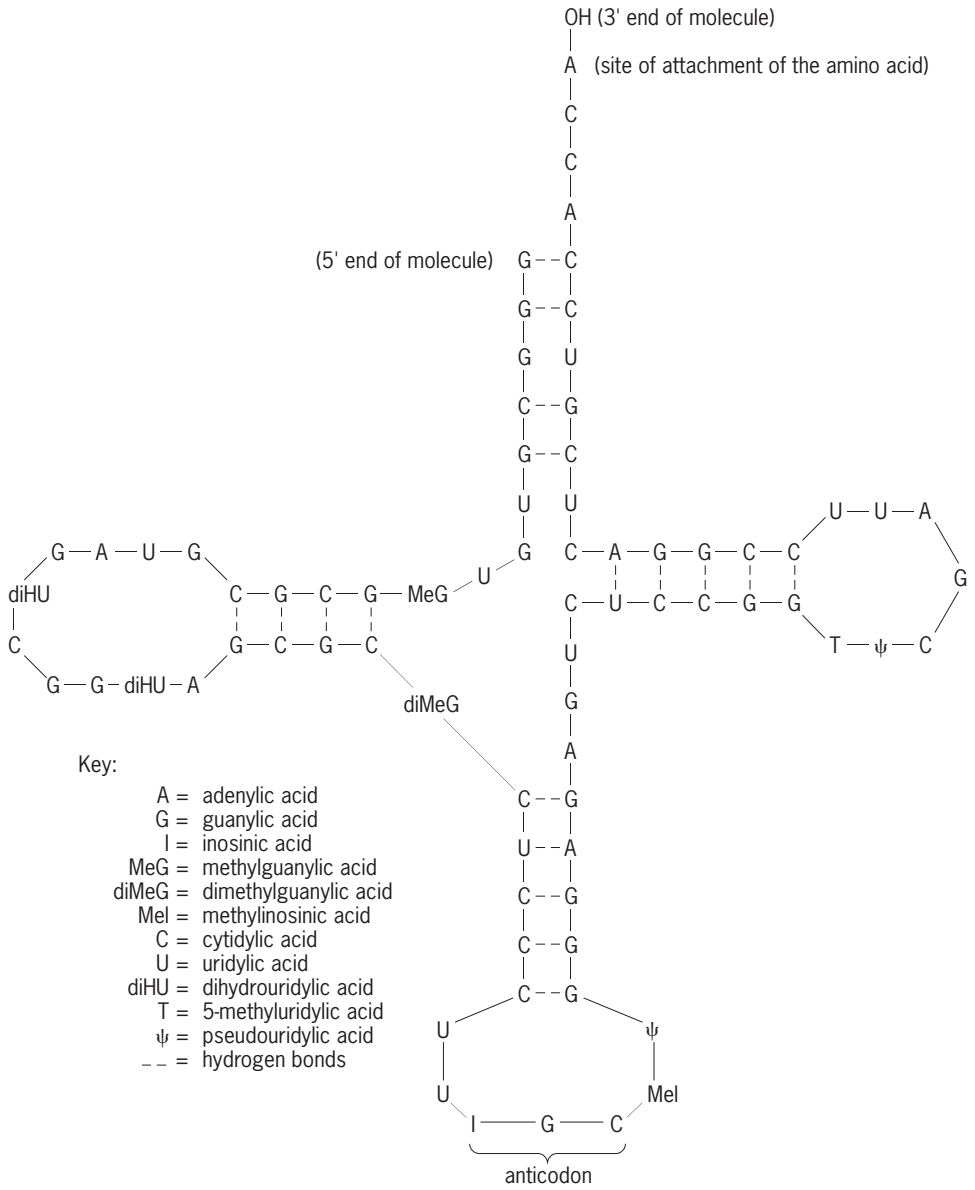

(18)

influenced by hormones. The interaction of RNA polymerases with DNA in nucleosomes is presently an active area of study.

Functions. The primary biological role of RNA is to direct the process of protein synthesis. The three major RNA classes perform different specialized functions toward this end. The 18S and 28S ribosomal RNAs of eukaryotes are organized with proteins and other smaller RNAs into the 45S and 60S ribosomal subunits, respectively. The completed ribosome serves as a minifactory at which all the components of protein synthesis are brought together during translation of the messenger RNA. The messenger RNA binds to the ribosome at a point near the initiation codon for protein synthesis. Through codon-anticodon base pairing between messenger and transfer RNA sequences, the transfer RNA molecules bearing amino acids are juxtaposed to allow formation of the first peptide bond between amino acids. The ribosome then, in a presently unknown fashion, moves along the messenger RNA strand as more amino acids are added to the peptide chain. *See* PROTEIN.

RNA of certain bacterial viruses serves a dual function. In bacteriophages such as f2 and Q-beta, the RNA serves as a message to direct synthesis of viral-coat proteins and of enzymes needed for viral replication. The RNA also serves as a template for viral replication. Viral RNA polymerases which copy RNA rather than DNA are made after infection. These enzymes first produce an intermediate replicative form of the viral RNA which consists of complementary RNA strands. One of these strands then serves as the sense strand for synthesis of multiple copies of the original viral RNA. *See* BACTERIOPHAGE.

RNA also serves as the actively transmitted genomic agent of certain viruses which infect cells of higher organisms. For example, Rous sarcoma virus,

which is an avian tumor virus, contains RNA as its nucleic acid component. In this case the RNA is copied to make DNA by an enzyme called reverse transcriptase. The viral DNA is then incorporated into the host cell genome, where it codes for enzymes which are involved in altering normal cell processes. These enzymes, as well as the site at which the virus integrates, regulate the drastic transformation of cell functions, which induces cell division and the ultimate formation of a tumor. Transcription of the viral DNA results in replication of the original viral RNA. *See* CYCLIC NUCLEOTIDES; NUCLEOPROTEIN; RIBONUCLEIC ACID (RNA); TUMOR VIRUSES.

Edward Johnson

In Situ Hybridization

In situ hybridization is a technique that permits identification of particular DNA or RNA sequences while these sequences remain in their original location in the cell. Single-stranded nucleic acids, both DNA and RNA, can bind to single strands of complementary nucleotide sequences to produce double-stranded molecules (nucleic acid hybrids). The stability of such hybrids is primarily dependent on the number of base pairs formed between the two strands. Under stringent hybridization conditions, only completely complementary sequences will form enough base pairs to remain associated. Thus a single-stranded nucleic acid probe can, by hybridization, identify a gene (DNA) or RNA transcribed from that gene.

In situ hybridization can be used to localize DNA sequences within chromosomes or nuclei and also to localize RNA within the nucleus or cytoplasm of the cell. Although most in situ hybridization studies are analyzed with the light microscope, it is also possible to perform in situ hybridization experiments on chromosomes prepared for analysis in the electron microscope.

The hybridization probe is the piece of single-stranded nucleic acid whose complementary sequence is to be localized within the cell. For example, the probe might be a recombinant DNA molecule carrying the gene that codes for a known protein. Nucleic acid probes can be labeled with the radioisotope tritium (^{3}H) by several different techniques. Sulfur-35 can also be used to label probes. At the end of an experiment, the location of the hybridized radioactive probe is detected by autoradiography of the cell preparation, thus demonstrating the location of the complementary sequence or sequences in the cell or tissue (**Fig. 4**). It is also possible to use nonradioactive probes made with modified nucleotides. For those probes, the hybrid is detected by an antibody against the modified nucleotides. The antibodies used have been joined to fluorescent molecules or to enzymes to allow cytological detection at the end of the experiment. *See* AUTORADIOGRAPHY.

The cell preparation is made by treating the tissue to be studied with a fixative. Fixatives are agents that preserve the morphology of the cell as well as

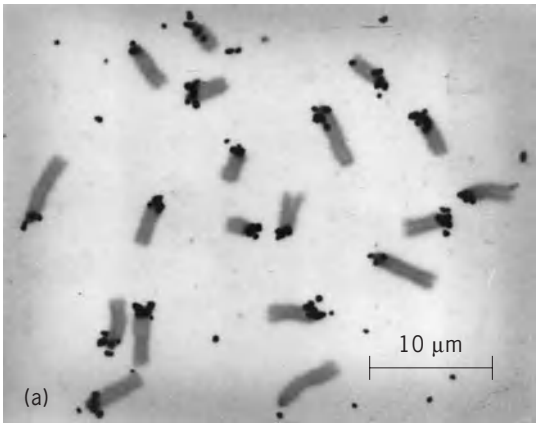

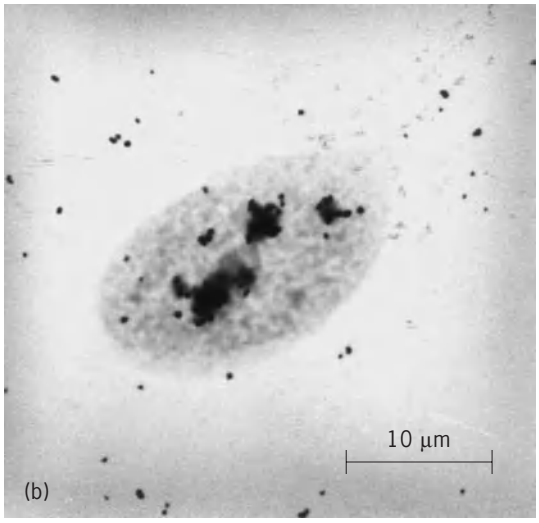

Fig. 5. Autoradiographs of mouse cytological preparations hybridized with ^{3}H-RNA copied from mouse satellite DNA. (*a*) Metaphase chromosomes, each with the centromere at one end. The black autoradiographic grains indicate that the highly repeated simple sequences of satellite DNA are adjacent to the centromeres in most of the chromosomes. (*b*) Interphase nucleus of Sertoli cell of testis. The autoradiographic grains are found in only a few clusters, indicating that in nuclei of this particular cell type the centromeric regions of the chromosomes are tightly associated into a few small groups. (*From C. D. Darlington and K. R. Lewis, eds., Chromosomes Today, vol. 3, Hafner Publishing, 1972*)

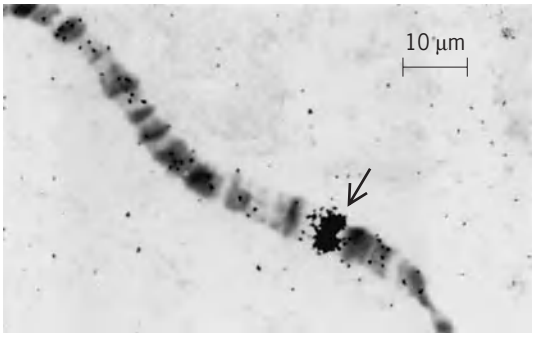

Fig. 4. Autoradiograph of part of a chromosome from *Drosophila melanogaster* hybridized with ^{3}H-DNA carrying the genes for muscle tropomyosin. Autoradiographic grains (arrow) are over the region that has the genes for both tropomyosin I and tropomyosin II. (*From V. L. Bautch et al., Organization and expression of Drosophila tropomyosin genes, J. Mol. Biol., 162:231–250, 1982*)

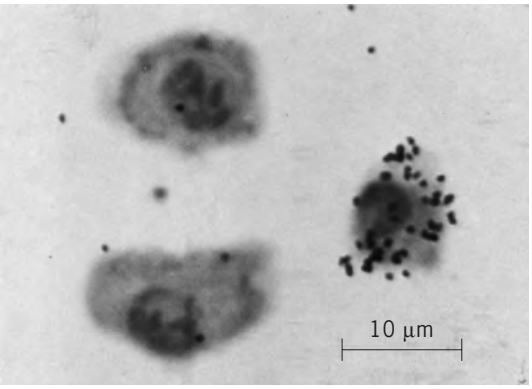

Fig. 6. Autoradiograph of *Drosophila* cells showing use of in situ hybridization to detect histone messenger RNA. Cells from a nonsynchronized population were hybridized with a ^{3}H-labeled probe complementary to histone messenger RNA, which is found only in cells that are synthesizing DNA. Thus, only the cell on the right, with autoradiographic grains indicating histone messenger RNA in the cytoplasm, appears to be actively replicating its DNA.

possible during subsequent steps. Fixatives usually act by cross-linking or denaturing proteins. After cells have been fixed, they are sectioned or squashed to yield preparations that are thin enough (a few micrometers) to be mounted on slides and analyzed in the microscope. If the experiment is designed to localize DNA in the cells, the DNA is denatured by treating the cells with a denaturing agent, such as alkali. The denaturation step also removes some protein from the DNA, but enough protein is left to maintain cellular morphology. Since RNA is, for the most part, single-stranded no denaturing step is used for hybridization to RNA in cells.

The probe is placed on the cytological preparation and held there under stringent conditions until hybrids have had sufficient time to form. The nonhybridized probe is then washed off the preparation. If the probe is radioactive, the preparation is covered with autoradiographic emulsion and exposed in the dark until tritium decay has produced grains in the emulsion over the labeled hybrid. The autoradiograph is then developed, and the location of grains over the cells is analyzed. When a nonradioactive probe is used, the hybrid detection is accomplished by incubating the preparation with an antibody that recognizes the modified nucleotide probe. If the antibody is coupled to a fluorescent molecule, the preparation is viewed in a fluorescence microscope to localize sites of hybridization. If the antibody is tagged with an enzyme, the preparation is incubated with a substrate from which the enzyme will produce a colored product, thus marking the site of the hybrid. *See* IMMUNOFLUORESCENCE.

In situ hybridization can be used to answer a number of biological questions. Hybridization to the DNA of condensed chromosomes can be used for gene mapping (**Fig. 5***a*). Hybridization to DNA of interphase nuclei allows study of the functional organization of sequences in the diffuse chromatin of this stage of the cell cycle (Fig. 5*b*). Hybridization to cellular RNA allows a precise analysis of the tissue distribution of any RNA (**Fig. 6**).

Mary Lou Pardue

Bibliography. R. L. Adams, J. T. Knowler, and D. P. Leader, *The Biochemistry of the Nucleic Acids*, 11th ed., 1992; G. Eckstein and D. M. Lilley, *Nucleic Acids and Molecular Biology*, 1994; G. Eckstein and D. M. Lilley (eds.), *Nucleic Acids and Molecular Biology*, vol. 2, 1988; A. R. Leitch, *In Situ Hybridization*, 1994; W. Saenger, *Principles of Nucleic Acid Structure*, 1993; P. Tijssen, *Hybridization with Nucleic Acid Probes*, 2 vols., 1993.

Nucleon

The collective name for a proton or a neutron. These subatomic particles are the principal constituents of atomic nuclei and therefore of most matter in the universe. The proton and neutron share many characteristics. They have the same intrinsic spin, nearly the same mass, and similar interactions with other subatomic particles, and they can transform into one another by means of the weak interactions. Hence it is often useful to view them as two different states or configurations of the same particle, the nucleon. Nucleons are small compared to atomic dimensions and relatively heavy. Their characteristic size is of order 1/10,000 the size of a typical atom, and their mass is of order 2000 times the mass of the electron. *See* I-SPIN.

Proton-neutron differences. The proton and neutron differ chiefly in their electromagnetic properties. The proton has electric charge +1, the opposite of the electron, while the neutron is electrically neutral. They have significantly different intrinsic magnetic moments. Because the neutron is slightly heavier than the proton, roughly 1 part in 1000, the neutron is unstable, decaying into a proton, an electron, and an antineutrino with a characteristic lifetime of approximately 900 s. Although some unified field theories predict that the proton is unstable, no experiment has detected proton decay, leading to lower limits on its characteristic lifetime ranging from 1.6×10^{25} years to 5×10^{32} years, depending on the assumed decay channel.

Structure and dynamics. In the early 1900s the proton was identified as the nucleus of the hydrogen atom. After discovery of the neutron in the early 1930s, atomic nuclei were quickly understood to be bound states of nucleons. To explain most aspects of nuclear structure and dynamics, it suffices to treat the nucleons as pointlike particles interacting by means of complex forces that are parametrized numerically. The complex forces between nucleons and the discovery during the 1950s of many similar subatomic particles led physicists to suggest that nucleons might not be fundamental particles. During this period, hundreds of short-lived subatomic particles were discovered, characterized, and cataloged. Those that interact in strong and complex ways with nucleons are known as hadrons. Electron scattering

experiments during the 1950s indicated that nucleons and other hadrons are extended objects with a distributed charge density. Significant progress was made during the late 1960s and 1970s, when inelastic electron and neutrino scattering experiments indicated that nucleons are composed of pointlike particles with spin $^1/_2$ and electric charges that are fractions of the charge on the electron. Particles with similar properties, named quarks, had been hypothesized in the early 1960s to explain other regularities among the properties of hadrons. In the early 1970s, it became clear that nucleons and other hadrons are indeed bound states of quarks. *See* HADRON; NUCLEAR STRUCTURE.

Quarks are believed to be fundamental particles without internal structure. The proton consists of two up-type quarks and one down-type quark (*uud*), while the neutron consists of *ddu*. Many of the properties and interactions of nucleons can be understood as consequences of their quark substructure. Quarks are bound into nucleons by strong forces carried by gluons. The nucleon contains ambient gluon fields in somewhat the same way that the atom contains ambient electromagnetic fields. Because quarks and gluons are much less massive than the nucleon itself, their motion inside the nucleon is relativistic, making quark-antiquark pair creation a significant factor. Thus the nucleon contains fluctuating quark-antiquark pairs in addition to quarks and gluons. A significant fraction (of order one-half) of the momentum of a rapidly moving nucleon is known to be carried by gluons. Likewise, only a small fraction of the nucleon's angular momentum resides on the spins of its constituent quarks. These facts indicate that the nucleon is a complex object characterized by highly nontrivial internal dynamics.

Quantum chromodynamics. The theory of quark-gluon interactions is known as quantum chromodynamics (QCD), in analogy to the quantum theory of electrodynamics (QED). Quantum chromodynamics has brought a deeper understanding to a previously bewildering subatomic world. Hundreds of other subatomic particles similar to nucleons are now understood to arise as different configurations of quarks and gluons. At the same time, the proton and neutron, once viewed as special because they are the constituents of ordinary matter, are now understood to be merely the most stable of a large number of composite quark-gluon bound states. Although quantum chromodynamics can be formulated in a simple and elegant manner, the dynamical equations that describe the nucleon bound state have so far resisted all attempts at quantitative solution. Thus a quantitative description of the nucleon from first principles has not yet been achieved. *See* ELEMENTARY PARTICLE; GLUONS; NEUTRON; PROTON; QUANTUM CHROMODYNAMICS; QUANTUM ELECTRODYNAMICS; QUARKS.

Robert L. Jaffe

Bibliography. R. K. Bhaduri, *Models of the Nucleon*, 1988; J. F. Donoghue, E. Golowich, and B. R. Holstein, *Dynamics of the Standard Model*, 1992; A. Pais, *Inward Bound*, 1986; R. G. Roberts, *The Structure of the Proton*, 1990.

Nucleoprotein

A generic term for any member of a large class of proteins associated with nucleic acid molecules. Nucleoprotein complexes occur in all living cells and viruses, where they play vital roles in deoxyribonucleic acid (DNA) replication, transcription, ribonucleic acid (RNA) processing, and protein synthesis. The chemistry of the nucleoproteins has become one of the fastest-growing areas in cell and molecular biology. Developments in this field have already revolutionized our understanding of complex biological processes within a cell. These advances will have important consequences in the treatment of genetic disease, viral infections, and cancer.

Classification of nucleoproteins depends primarily upon the type of nucleic acid involved—DNA or RNA—and on the biological function of the associated proteins. DNA and RNA share very similar sugar-phosphate backbones, with nitrogenous bases attached to the ribose sugar. The major distinction is that DNA contains a deoxyribose sugar group, with the loss of a hydroxyl group at the 2′ carbon. Deoxyribonucleoproteins (complexes of DNA and proteins) constitute the genetic material of all organisms and many viruses. They function as the chemical basis of heredity and are the primary means of its expression and control. Most of the mass of chromosomes is made up of DNA and proteins whose structural and enzymatic activities are required for the proper assembly and expression of the genetic information encoded in the molecular structure of the nucleic acid. *See* DEOXYRIBONUCLEIC ACID (DNA).

Ribonucleoproteins (complexes of RNA and proteins) occur in all cells as part of the machinery for protein synthesis. This complex operation requires the participation of messenger RNAs (mRNAs), amino acyl transfer RNAs (tRNAs), and ribosomal RNAs (rRNAs), each of which interacts with specific proteins to form functional complexes called polysomes, on which the synthesis of new proteins occurs. *See* RIBONUCLEIC ACID (RNA).

In simpler life forms, such as viruses, most of the mass of the viral particle is due to its nucleoprotein content. The material responsible for the hereditary continuity of the virus may be DNA or RNA, depending on the type of virus, and it is usually enveloped by one or more proteins which protect the nucleic acid and facilitate infections. *See* BACTERIOPHAGE; VIRUS.

Classically, the central dogma of molecular biology was that nucleic acids, the genetic material, are the transmissible hereditary material. According to the central dogma, DNA is transcribed into RNA, which is then translated into proteins that carry out various functions within the cell. With the discovery of noncoding RNAs and RNA viruses that use reverse transcriptase, it is increasingly clear that this central dogma needs further expansion. In addition, epigenetic modification of DNA (changes that do not alter DNA code, but can influence whether genes are expressed or repressed) represents an encryption not included in this dogma. The processes of reading the DNA code, transcription of RNA, and translation of

mRNA into protein all involve interaction of proteins with nucleic acids in some form to carry out cellular processes. Therefore, it is important to understand the composition and function of nucleoproteins. *See* GENE ACTION.

Deoxyribonucleoproteins

Those complexes in which the nucleic acid component contains the sugar 2-deoxyribose are primarily concerned with the storage and utilization of genetic information. This accounts for the localization of most of the deoxyribonucleoprotein in the cell nucleus as components of the chromosomes. *See* CHROMOSOME; NUCLEIC ACID.

The presence of DNA in the metaphase chromosomes of dividing cells and in the corresponding chromatin fibers (the complex of DNA and proteins) of interphase cells is revealed by direct chemical analysis of isolated chromosomes or chromatin, or by staining reactions that make DNA visible under the microscope. An important corollary of the chromosomal localization of DNA is the striking constancy in the amount of DNA per set of chromosomes in the different cells of a given organism. The nuclei of diploid somatic cells, which are body cells with two sets of chromosomes, have twice the DNA content of sperm cells, which contain only a single chromosome set. Cells in the process of division, at the time of chromosome duplication, show a corresponding increase in the amount of DNA which they contain. The direct demonstration of a genetic role of DNA came from experiments in bacterial transformation, when it was discovered that when DNA from one bacterial strain was transferred to a different strain, the recipient bacteria took on phenotypic characteristics of the DNA donor. The change induced by the DNA is inherited and persists through subsequent generations. *See* TRANSFORMATION (BACTERIA).

While the hereditary properties of cells and viruses are determined by their unique DNA structure (that is, by the nucleotide sequences which spell out the genetic code for protein structure), the ability to read the code and use such information as the basis of activity resides in the proteins that interact with DNA. In nature, the different types of DNA genomes, each encoding for a large number of functions, are associated with a complex array of different proteins. Many of these proteins are connected with generation of RNA molecules from DNA; other nuclear proteins are primarily concerned with the organization of chromosome structure. *See* GENETIC CODE.

Role of histones in higher organisms. The genetic complexity of higher organisms reflects a corresponding complexity in their DNA sequences. A typical human diploid nucleus, for example, contains enough DNA that, if fully extended, would be over 1 m (3.3 ft) long. However, the ability to compact this amount of DNA into a nucleus 0.01 mm in diameter is accomplished routinely by cells. The reduction in size is largely due to interactions between the DNA and sets of small basic proteins called histones.

All somatic cells of higher organisms contain five major histone classes, all of which are characterized by a high content of the basic amino acids arginine, lysine, and histidine (which are positively charged to interact with negatively charged DNA). The characterization of histones was originally based on their chromatographic or solubility properties, but emphasis is now placed on details of their structure and their capacity to organize DNA. The major classes are called H1, H2A, H2B, H3, and H4. Selective degradation of chromatin by staphylococcal nuclease releases subunits (nucleosome "core" particles) containing about 140 base pairs of DNA associated with eight molecules of histone. Analysis of the histones in the subunits reveals that only four of the five major histone classes are represented, with two copies of each per subunit particle. The histones form a core complex composed of two dimers of H2A/H2B and a tetramer of H3 and H4 (**Fig. 1**). Histone H1 does not appear to be part of the nucleosome core, but instead binds the linker DNA between cores, adding another level of DNA compaction. In general, the association of H1 prevents DNA sequences (genes) from being transcribed.

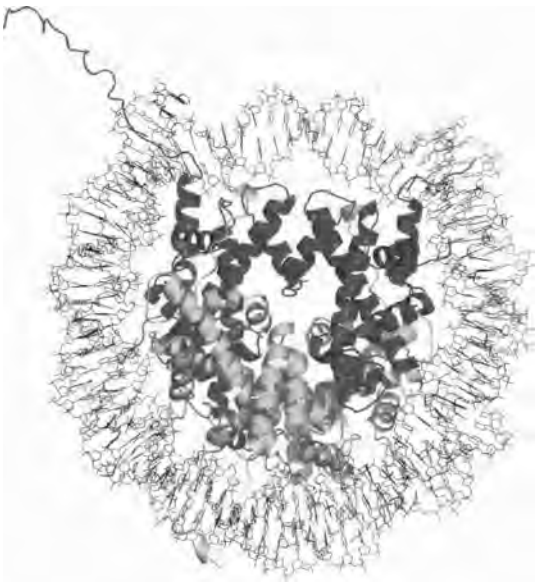

Fig. 1. Histone core complex. DNA is wrapped around a core octamer of histones. Two dimers of histones H2A/H2B form this complex with two copies of histones H3 and H4. (*Rendered by Debra Ferraro*)

The presence of histone-DNA aggregates in chromosomes or in the chromatin of interphase cells can be visualized directly by electron microscopy (**Fig. 2**). The characteristic beads-on-a-string arrangement has now been observed in the chromosomes of organisms as diverse as slime molds, birds, and humans, and appears to be a general mechanism for the assembly and packaging of the genetic material.

While chromatin compaction and nucleosome arrangement provide stability and structure to DNA, chromatin must also be dynamic in the sense that selected DNA regions are called into action as templates for RNA synthesis at different times in the life of the cell. Two molecular mechanisms used to

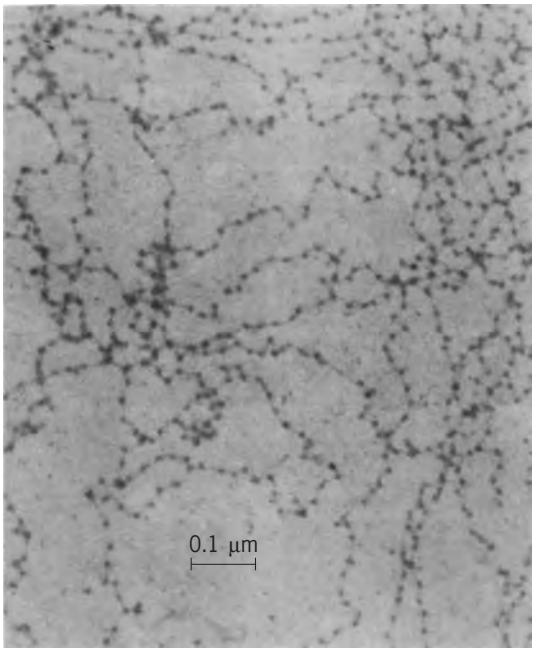

Fig. 2. Histone-DNA aggregates in chromosomes.

change chromatin assembly, properties, and function as needed are modifications of the histone proteins or their replacement with variant histones.

Histone variants. Histones H3 and H4 are among the most conserved proteins among species. As a whole, the nucleosome structure and proteins are also very conserved, hinting at a vital role. However, each of the five classes of histones contains variants that serve specific functions in the cell at different times. Histone variants are encoded by separate genes and contain core structures similar to their canonical family members. Changes in the size and the content of the N- and C-terminal domains, as well as single amino acid substitutions within the core, provide unique functions. While the major histones are expressed mainly during the DNA synthesis (S phase of the cell cycle), the variants are often used for specific functions at other stages of the cell cycle or development.

Histone H2A has the greatest number of variants. These variants range from the vertebrate-specific histone macroH2A which binds to the inactive X chromosome, to H2A.Z, a variant conserved from yeast to human that is involved in transcriptional activation and repression. The other histone classes have a number of members with specificities that provide specialized chromatin packaging in specific regions of genomic DNA.

Histone modifications and the "histone code." Biochemical alterations can modify specific amino acids in the histone polypeptide chain. Attachment of acetyl, phosphate, and methyl groups not only alters the structure and charge of the nucleosome but also regulates interactions between histone and nonhistone proteins (see **table**). These postsynthetic modifications of histones are sometimes referred to as an "epigenetic histone code." The code can be "translated" into various biological consequences, such as changes in transcriptional activity, chromatin condensation, or temporal and spatial regulation of gene activation. Numerous correlations have been noted between the extent of histone acetylation and RNA synthesis. Indeed, histone acetatylation is enriched at very early stages in development when genes are activated in response to hormones, growth factors, or other stimuli. Furthermore, it has been observed that chromosome regions that are incapable of RNA synthesis are packaged with histones in the nonacetylated form.

Nonhistone specialized packaging. In sperm cells, particular mechanisms have evolved for the packaging and delivery of genetic material. Special DNA-binding proteins are synthesized in the course of sperm maturation. Among the simplest of these are the protamines which occur in the sperm cells of fish. These proteins are characterized byshort chain

Examples of histone modifications and their effects on cellular functions*

Histone	Modification (residue)	Associated function
Histone H1	Phosphorylation	Transcriptional activation
Histone H2A	Phosphorylated (S1)	Transcriptional repression, chromatin condensation
	Phosphorylated (T119)	Cell cycle progression?
Histone H2B	Phosphorylated (S14)	Apoptotic chromosome condensation
	Phosphorylated (S33)	Transcriptional activation
	Ubiquitinated (K119 or K123)	Histone H3 methylation
Histone H3	Methylated (K4)	Transcriptional activation
	Acetylated (K9)	Transcriptional activation
	Methylated (K9)	Transcriptional repression
	Phosphorylated (S10)	Chromosome condensation and mitosis/meiosis
	Acetylated (K14)	Transcriptional activation
	Methylated (K27)	Transcriptional repression
	Methylated (K79)	Transcriptional activation, telomeric silencing
Histone H4	Phosphorylated (S1)	Chromatin condensation during mitosis
	Methylated R3)	Enriched at developmentally regulated genes
	Acetylated (K16)	Transcriptional activation

*Modifications are made to amino acid residues in the histones. Modified amino acids are abbreviated using a single letter code (serine, S; threonine, T; lysine, K; arginine, R) and are followed by the position of the amino acid within the protein, with the first amino acid of the protein designated 1.

lengths (about 33 amino acids) and a great preponderance of the basic amino acid arginine. In salmine, the protamine of salmon sperm, arginine alone accounts for 89–90% of the total protein nitrogen. It imparts a high positive-charge density to the polypeptide chain. The positively charged arginine residues occur in clusters, each member of which binds a phosphate group in DNA. The number of basic amino acids in nucleoprotamine complexes approximately equals the number of phosphate groups in the DNA, indicating almost complete neutralization of the negatively charged DNA molecule. This charge neutralization is accompanied by a supercoiling of the DNA helix, and is probably the mechanism by which DNA in mature sperm is kept tightly packed for delivery during fertilization. At fertilization, the mature sperm nucleus is incapable of RNA or DNA synthesis.

DNA polymerases and ligases. Replication of the genetic material requires an unerring synthesis of DNA molecules that are identical in nucleotide sequence to the parental DNA molecules. Enzymes which carry out this synthesis are called DNA polymerases. They have an absolute requirement for a DNA template and take instructions from that template as they move along it, adding nucleotides to the newly synthesized DNA chain. The synthesis of DNA from low-molecular-weight precursors (the deoxyribonucleoside triphosphates) proceeds rapidly at a rate approaching 1000 nucleotides per minute per molecule of enzyme. Some DNA polymerases have been isolated in pure form, but the initiation and accuracy of DNA replication involves the participation of many proteins in addition to the polymerase itself, and sometimes requires the presence of an RNA "primer" to initiate the process.

Biochemical studies of DNA replication in microorganisms and in animal cells in culture show that the earliest recognizable products of the DNA polymerase reaction are small fragments that are later joined to form the long DNA sequences found in the finished chromosome. The joining of these small DNA pieces involves the participation of other enzymes called DNA ligases. The ligases, by using adenosine triphosphate (ATP) as an energy source, join the smaller molecules by the formation of phosphodiester bonds to produce an uninterrupted polynucleotide chain.

DNA ligases also play an important role in DNA repair mechanisms. For example, if the genetic material is damaged by ultraviolet irradiation, the damaged areas can be excised. New nucleotides are then inserted by a DNA polymerase, and the repaired area is sealed to the rest of the DNA molecule by a DNA ligase.

DNA unwinding proteins. Bacteria, viruses, and animal cells contain proteins which destabilize the DNA double helix (helicases) or produce breaks in the DNA for coiling or untangling (topoisomerases).

Helicases. Helicases are a group of proteins that associate with both DNA and RNA and function to separate base pairing of the nucleic acids. These enzymes use the energy from ATP hydrolysis to separate the strands of DNA for a variety of functions, including DNA replication, recombination, and transcription. Helicases often have multiple protein subunits that assemble to form a ring structure that encircles a single nucleic acid strand and then follows that strand, unwinding any secondary conformations it encounters.

Topoisomerases. DNA topoisomerases have been found in organisms from bacteria to humans and play critical roles in DNA replication, transcription, and repair. Topoisomerases function by generating single- or double-strand DNA breaks and then rotating or passing DNA through the break. Classifications of these proteins are based upon their product. Enzymes that make a break in one strand of DNA are type I topoisomerases; those that produce a break in both strands are called type II. The cleavage step leads to a phosphodiester bond between the topoisomerase protein and the DNA, allowing the protein to maintain contact with the broken strands during manipulation of the DNA. These proteins have gained importance as they are often targets of antibiotics or anticancer therapies.

DNA repair. Given the importance of DNA as the hereditary molecule, maintaining the correct DNA sequence is vital to a cell. However, insults to the hereditary material can occur from a variety of common sources, such as mistakes in DNA synthesis, exposure to ultraviolet light, and damage from harsh chemicals. Unfortunately, it is often impossible merely to fix the mistake where it is found; so the cell must find the error, degrade the DNA around the error, and then resynthesize that portion of DNA using the other DNA strand or even the homologous chromosome as a template. Many repair processes involve ribonucleoprotein complexes such as endonucleases, exonucleases, ligases, helicases, and polymerases.

At least five mechanisms exist to repair DNA based on the type of damage. The first mechanism is direct repair, involving proteins to directly modify the bases and revert them to original form. However, this is often an inefficient and ineffective method of repair. The second and third methods are base excision repair and nucleotide repair. Both processes involve the removal of the damaged bases and a few surrounding nucleotides from the DNA strand, followed by the resynthesis of the missing section with subsequent ligation of the new and old fragments. The difference between the two repair mechanisms lies in whether the sugar backbone is cut during or after the excision of the base. Defects in base excision repair severely limit the ability of cells to respond to thymine dimers (often caused by exposure to ultraviolet light) leading to the disease xeroderma pigmentosa, which is characterized by thinning of the skin, splotchy pigmentation, spider veins, and skin cancers.

The fourth mechanism is postreplication repair, which is used to correct double-strand breaks. This process often involves use of the homologous chromosome as a template for the synthesis and ligation of the two DNA ends, thus ensuring that no DNA bases were lost in the break and repair. This is very

similar to the type of homologous chromosome usage in meiotic recombination.

The fifth type of repair is known as mismatch repair. This process is often used during the replication of DNA during the cell cycle if DNA polymerases accidentally insert the incorrect base. The incorrect base is removed and the correct base is then inserted. Without this proofreading and repair mechanism, the error rate during replication can be 1000 times higher than normal, leading to high rates of mutagenesis.

Nucleases and DNA-modifying enzymes. Many of the enzymes that take part in the hydrolysis of DNA are known. Such proteins are called deoxyribonucleases and are classified in terms of their mode of action and source: endonucleases attack the internal structure of the polynucleotide chain, while exonucleases cleave the terminal nucleotides from the nucleic acid molecule. Of particular use are bacterial restriction enzymes, which cleave specific sequences in DNA molecules, enabling many techniques in molecular genetics. A battery of such enzymes is now available which provides DNA fragments suitable for DNA sequence analysis and recombinant DNA technology.

In a cell, endogenous deoxyribonucleases are used to recognize and degrade foreign DNA, among other functions. To determine what DNA is foreign to the cell, enzymes interact with endogenous DNA and modify the structures of the purine and pyrimidine bases, usually by the introduction of methyl groups at precise positions in the DNA molecule. Such structural modifications of DNA are a means of imparting species specificity among DNA molecules, as different sequences are methylated among different organisms. This allows for species-specific DNA endonucleases to degrade DNA foreign to the host. Enzymes that methylate DNA limit their activities on homologous DNA but may overmethylate (and cleave) DNA from another species.

In eukaryotes, DNA methylation marks regions of different transcriptional activity, as heavily methylated genes are transcriptionally silent. This is often thought of as a host-defense mechanism to guard against the spread of transposons (segments of DNA that can replicate and insert themselves elsewhere in the genome). A significant fraction of the genome within higher organisms is made of transposable elements that are methylated. Mutations that lead to demethylation of these elements cause genomic instability.

Proteins involved in gene regulation. Different cell types within higher organisms have equivalent DNA contents, but they utilize different portions of the DNA to direct their specialized functions. It follows that nuclear protein mechanisms exist for the selective utilization of DNA sequences. Because of the limited number of histone types and the relative uniformity in histone composition in different cells, it is likely that they are only part of the regulatory network of gene transcription.

Human (cultured cell) nuclei, for example, contain over 470 different protein classes that can be separated according to their molecular size and charge, with many capable of DNA binding. Others are enzymes involved in nucleic acid synthesis and repair, enzymes which modify histone structure, proteins involved in the processing and transport of mRNAs from the nucleus to the cytoplasm, and proteins destined to become components of ribosomes. The DNA-binding proteins are of primary interest for control mechanisms. Studies of RNA synthesis in isolated chromatin fractions have shown that the rate of RNA transcription depends on the source of the nonhistone nuclear proteins added to the DNA template. These proteins can have roles involving transcription initiation, elongation, and termination, or they can modify other proteins or mRNA by controlling stability, degradation, or interactions.

Nuclear protein modifications and genetic control. Many of the histone modifications, such as phosphorylation and methylation, are also found on other proteins. Additional modifications, such as glycosylation (the addition of sugars), ubiquitination (addition of ubiquitin, which targets a protein for degradation), and sumolation [addition of a small ubiquitin-related modifier (SUMO)] can also affect protein half-life and function. Effects of individual modifications are specific to the protein and amino acid residue, but the general idea is that these modifications add another level of complexity to the interaction of proteins and nucleic acids. For example, many of the chromosomal nonhistone proteins are subject to phosphorylation reactions that modify serine and threonine residues in the polypeptide chains. The result is an alteration of protein charge and structure which affects the interactions of regulatory proteins with each other and with the DNA template. The phosphorylation of nuclear proteins generally correlates in a positive way with increases in RNA synthesis when cells respond to hormones, growth factors, or other stimuli. Furthermore, the enzymatic dephosphorylation of the nuclear nonhistone proteins diminishes their capacity to stimulate RNA synthesis (transcription) in isolated chromatin fractions.

RNA polymerases. Among the most important of the proteins associated with DNA are the enzymes responsible for RNA synthesis. In yeast and higher organisms, there appear to be at least three different RNA polymerase activities in the nucleus. All are DNA-dependent, and their products are complementary to the DNA sequences copied by the enzyme. RNA polymerase I functions in the nucleolus, where it copies the ribosomal genes to produce rRNAs. A different polymerase, RNA polymerase II, transcribes mRNA sequences. RNA polymerase III directs the synthesis of small RNA molecules such as the amino acyl tRNAs used for protein synthesis. All three polymerases are complex enzymes that can be discriminated from one another by interactions with inhibitors. For example, polymerase II, but not polymerase I, is sensitive to inhibition by the mushroom poison α-amanitin.

All of the RNA polymerases described above are DNA-dependent enzymes. However, other polymerases are known which copy RNA templates called RNA-dependent RNA polymerases. These polymerases are often used to replicate viruses that have an RNA genome.

There are at least three ways to control the ability of RNA polymerases to function, and the cellular machinery makes use of all of them. RNA transcription initiation, elongation, and termination each have specific machinery designed to add check-points that can modify transcriptional output. Such fine control allows the levels of RNA transcription to match the cellular needs in a small time frame.

Control of RNA transcription initiation. Mechanisms of genetic control in higher organisms remain an area of intense research interest. However, relevant models can be derived from the detailed analyses of genetic control elements in bacterial cells and viruses. The molecular basis of gene control is known in remarkable detail for the lactose (*lac*) operon in the bacterium *Escherichia coli* (**Fig. 3**). This segment of the bacterial chromosome directs the synthesis of three enzymes (β-galactosidase, β-galactoside permease, and thiogalacotoside transacetylase) needed for lactose utilization. The genes for these three proteins are linked together on the bacterial chromosome and are directed by three major control elements located before the first gene in the sequence: (1) a repressor gene (the *i* gene), (2) a promoter sequence, and (3) an operator sequence. Transcription of the *lac* operon is both negatively and positively controlled. Negative control is exerted by the lac I repressor, a protein product of the *i* gene which binds specifically to the operator region and thus prevents transcription (Fig. 3). Another region of the repressor molecule interacts with the low-molecular-weight substrate allolactose, which acts as an inducer of the *lac* operon. This interaction brings about a conformational change (an allosteric modification) in the repressor, which leads to its release from the operator DNA sequence. This allows RNA polymerase to transcribe the adjacent genes, whose products are necessary for galactoside metabolism.

An important point is that the removal of the repressor is not, in itself, sufficient to initiate transcription. Transcription also requires the participation of a positive control protein called the catabolite gene activator protein or the cyclic adenosine monophosphate (AMP) receptor protein (CAP protein). The CAP protein functions as a dimer that attaches to specific DNA sequences in the promoter region of the *lac* operon and activates transcription only when bound to the signaling molecule cAMP, which becomes available when glucose is not present. A significant clue to the specificity of interactions between

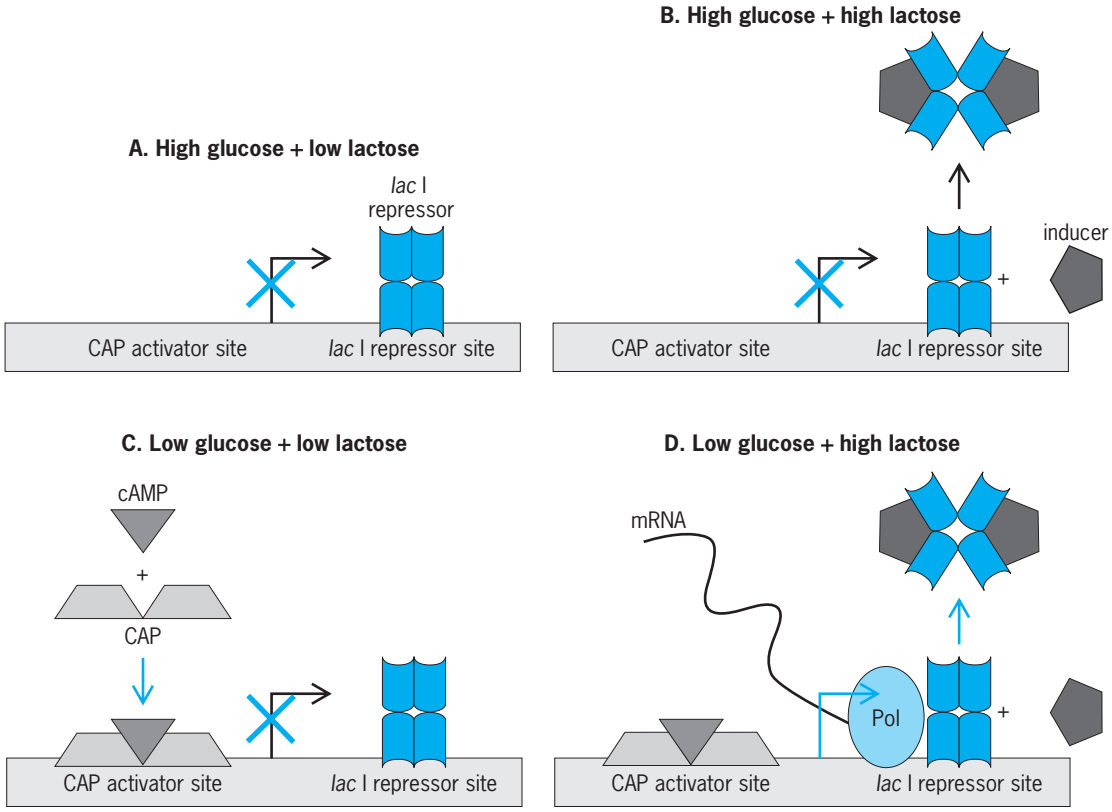

Fig. 3. Transcriptional regulation of the *lac* operon of *E. coli*. Transcription of the *lac* operon is both negatively and positively controlled. Negative control is exerted by the *lac* I repressor, which binds to a repressor site and prevents mRNA transcription unless an inducer is produced in the presence of high lactose. Positive control is exerted by cyclic adenosine monophosphate (cAMP) and the cAMP receptor protein (CAP protein). cAMP is produced when there is little glucose, allowing CAP to bind to the activator site and recruit RNA polymerase to initiate transcription. (*a*) Under conditions of high glucose and low lactose, the *lac* I repressor is bound, preventing transcription. High levels of glucose results in no production of cAMP and, therefore, CAP is not bound at the activator site. (*b*) Under conditions of high glucose and high lactose, the inducer binds the *lac* I repressor causing the repressor to dissociate from the DNA. High levels of glucose result in no production of cAMP and, therefore, no transcription initiation by CAP. (*c*) Under conditions of low glucose and low lactose CAP promotes transcription initiation, but RNA polymerase is blocked by bound *lac* I repressor. (*d*) Under conditions of low glucose and high lactose, CAP is bound and recruits RNA polymerase, and the removal of the *lac* I repressor by the inducer allows for transcription.

such regulatory proteins and DNA is their capacity to recognize symmetrical elements in DNA structure. Binding of the CAP protein involves a region of twofold rotational symmetry. The interaction of the CAP protein with DNA is believed to facilitate the entry of RNA polymerase, allowing transcription of the genes required for lactose utilization.

Control of RNA transcription elongation. In organisms ranging from yeast to humans, RNA polymerase II has a C-terminal domain (CTD) that is important for the processivity of the polymerase. Upon initiation, the CTD is hypophosphorylated, but in order for elongation of the transcript to occur the CTD must be phosphorylated. If RNA polymerase II is not phosphorylated, it will not yield full-length transcripts, providing another means of regulation. Thus for very long RNA transcripts, such as viral RNA genomes, it is necessary for the virus to make sure that the entire genome is replicated.

A well-studied example of this regulatory mechanism is the transcriptional control of the Tat-TAR system for elongation in human immunodeficiency virus 1 (HIV-1). Tat is a small protein that contains a basic region consisting of nine positively charged amino acids in close proximity to one another. This region has three important functions: (1) it acts as a nuclear localization signal, (2) it binds to receptors, and (3) it interacts with the transactivation-response region (TAR) in the HIV-1 RNA transcript and enhances RNA elongation. The TAR RNA has a small stem-loop in its structure that forms a docking site for Tat. Once docked, Tat recruits a Tat-associated kinase (TAK, such as P-TEFb) that can phosphorylate the CTD, allowing the full transcription of the HIV-1 genomic RNA.

Control of RNA transcription termination. RNA synthesis ends at precise regions of the chromosome, producing mRNAs of precise lengths. Termination of RNA transcription can be controlled in at least two ways: (1) intrinsic terminators can stop RNA polymerase without other proteins or (2) DNA-binding proteins can stop the polymerase. Both systems have been characterized in *E. coli*.

Intrinsic terminators are hairpin loops in the RNA which cause RNA polymerase to slow, and a series of uracil bases after the hairpin that disrupts the DNA-RNA base pairing, causing polymerase release. Terminator proteins act by unwinding the DNA-RNA duplex produced during transcription. In bacteria, one such protein, called rho factor, has six subunits that interact to form a planar annulus around an empty core. There is a built-in gap between the subunit particles, which permits the complex to encircle a nucleic acid strand and subsequently follow the strand to find the transcribing RNA polymerase. Termination itself can be affected by antitermination proteins, such as the *N* gene product of bacteriophage lambda, which can override or attenuate the termination signal.

Reverse transcriptase. A number of RNA viruses replicate by a mechanism that involves a DNA-containing replicative form. These viruses contain or produce an enzyme, reverse transcriptase, which copies RNA sequences to form a complementary DNA (cDNA) strand. This enzyme has many important applications in molecular biology. For example, it can be used to make a DNA copy of mRNA; the copy can be made highly radioactive and used for hybridization studies to locate the corresponding genes in chromosomes, or it can be used to measure how much of the corresponding mRNA is synthesized in a cell. Moreover, the cDNA copy of a message can be incorporated into bacterial plasmids as recombinant DNA, thus permitting amplification of the DNA sequences which code for the mRNA. *See* REVERSE TRANSCRIPTASE.

Ribonucleoproteins

As in the case of the DNA-protein complexes, ribonucleoproteins can be classified according to their functions and localization within the cell. Ribonucleoproteins are present in all cells and constitute the genetic material of many viruses; poliovirus, for example, is a ribonucleoprotein.

mRNA. The mRNAs themselves occur in association with proteins. When mRNA is synthesized in the nucleus, it is quickly associated with sets of proteins that organize the nascent RNA chains into a more particulate, and presumably protected, structure. These mRNA-binding proteins add a "cap" to the 5′ end of the mRNA to help stabilize the molecule and prevent degradation of the transcript. Another enzyme adds adenylic acid residues to the 3′ end of the molecule for similar reasons. The poly-A sequence at the 3′ end of the mRNA is specifically associated with poly-A–binding proteins. *See* NUCLEIC ACID.

The shielding of mRNAs by associated proteins appears to be one basis for the stability of messengers in certain cell types, such as oocytes, in which the mRNAs are nonfunctional but remain intact for long periods. Removal of the masking proteins requires proteolysis; this occurs at the time the egg is fertilized, and the unmasked mRNAs become functional in the synthesis of embryonic proteins.

Many RNA-degrading enzymes are known; some function by cleaving the polynucleotide chain at internal linkages, while others remove terminal nucleotides in a stepwise fashion. Certain ribonuclease activities are localized in the cell nucleus, where they degrade portions of the high-molecular-weight precursors of mRNAs. Other ribonucleases are found in the cytoplasm and play a role in mRNA "turnover." Still others are secreted from the cell. One of the best-characterized ribonucleases is ribonuclease A of the pancreas. It contains 124 amino acids and was the first enzyme to be completely sequenced. It is also the first enzyme to be synthesized in the laboratory, by using amino acids and a technique of solid-phase peptide synthesis. *See* RIBONUCLEASE.

mRNA splicing. Splicing is the process of cleaving parts of the mRNA that are not needed for the mature mRNA to be translated. To accomplish this goal, a large complex (approximately 12 MDa) known as the splicesome, composed of both noncoding RNA and proteins, assists in the removal of introns by an energy-dependent process. Of specific interest

are the small nuclear ribonucleoprotein (snRNP) complexes which contribute almost half of this mass. These five complexes (U1, U2, U4, U5, and U6) are involved in building the lattice for the splicesome components, recruiting other complexes, base pairing with the mRNA, and directly acting in the removal of introns.

Which introns and exons are included in a mature mRNA is not always fixed. Alternative splicing, a method by which introns and exons from an immature mRNA can be incorporated or excluded in the mature mRNA, can occur. This leads to different mRNA transcripts from the same DNA sequence and allows much more mRNA diversity than that encoded by the DNA. For example, alternative splicing in the *Drosophila melanogaster* gene *Dscam* can potentially generate more than 38,000 different mRNA transcripts from the same gene. These transcripts, based on which introns and exons are used, can then encode proteins with very different and specialized localization properties and functions.

Ribosome and tRNA. The role of mRNA is to specify the amino acid sequences of proteins, but translation of the code from RNA to amino acids is conducted by ribonucleoprotein particles called ribosomes. Each ribosome is made up of two subunits which travel along the mRNA chain, taking instructions on how to link the amino acids in the proper sequence. The first step in protein synthesis involves enzymes known as amino acyl transfer RNA (tRNA) synthetases. The role of amino acyl tRNAs is to guide particular amino acids to their proper position in the ribosome, a complex of proteins and RNAs that reads mRNA and translates it into proteins. This positioning is achieved by interactions between nucleotide triplets (anticodons) in the tRNA and complementary triplets (codons) in the mRNA. There are many such enzymes; their job is to activate the amino acid and to attach it to an appropriate tRNA. Each activating enzyme has high specificity both for the amino acid and for the tRNA. More than one tRNA may exist for a particular amino acid, and the specificity of the amino acyl tRNA synthetases must be very high to avoid errors. This is an impressive example of the ability of a protein to recognize specific elements of sequence and structure in an RNA molecule.

The organization of the ribosomes involves many RNA-binding proteins. The large and small ribosomal subunits contain different RNA molecules and different sets of associated proteins. For example, in rabbit reticulocyte ribosomes the small subunit (40S) contains 32 proteins with molecular weights ranging from 8000 to 39,000, while the large subunit contains 39 proteins with molecular weights ranging from 9000 to 58,000. Each ribosomal subunit appears to contain just one copy of each distinct ribosomal protein. The functions of the ribosomal proteins are under intensive investigation. Some appear to be involved in interactions between the ribosomes and amino acyl tRNAs, with factors necessary for protein chain initiation and elongation, in binding the mRNA initiation complex, and in the vectorial transport and release of the nascent polypeptide chains.

Noncoding RNAs. Not all RNAs fall neatly into the classes of mRNA, tRNA, or rRNA. Similar to rRNAs or tRNAs, which are not translated into proteins, other noncoding RNAs (ncRNAs) play important roles in biological processes. In fact, 50–75% of all transcription in higher eukaryotes consists of ncRNA. These RNAs function as structural components, catalytic molecules, or regulators in the cell.

snoRNA. In a similar class to snRNAs (discussed as splicing components), snoRNAs complex with proteins. They are named for their subcellular localization within the nucleolus. However, while snRNAs are involved in mRNA splicing, snoRNAs help rRNA maturation and are essential for proper rRNA cleavage and modification.

Xist/Tsix antisense ncRNAs. NcRNAs are involved in the inactivation of specific genes or even whole chromosomes. An example is the role of ncRNAs in the dosage compensation (X inactivation) mechanism in mammals. X inactivation involves transcriptional silencing of one of the two X chromosomes in

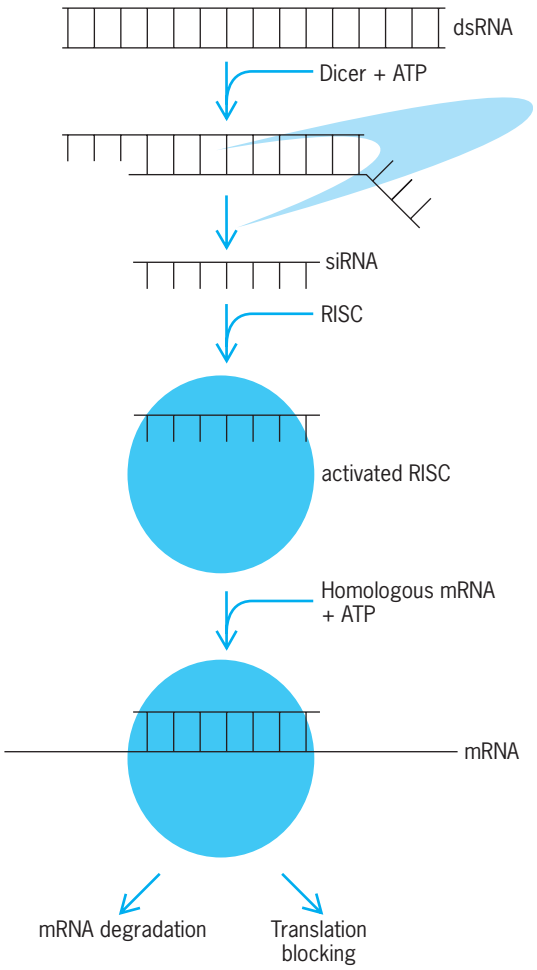

Fig. 4. Model of the RNA interference (RNAi) pathway. Long double-stranded RNA (dsRNA), consisting of sense and antisense strands, is cleaved by Dicer (crescent) in an energy (ATP)-dependent manner into short inhibitory RNAs (siRNA). The antisense siRNA is then incorporated into the RNA-induced silencing complex (RISC, circle), activating it and triggering a search for homologous mRNAs. When a match is found, RISC silences the mRNA by either degradation or translation blocking.

female mammals. Inactivation occurs early in development. Through a complex mechanism involving X chromosome counting and epigenetic markings, the X-inactivation center produces the ncRNA *Xist* on the future inactive X chromosome, and the ncRNA *Tsix* on the future active X chromosome (both *Xist* and *Tsix* are polyadenylated and capped ncRNAs). These two ncRNAs are derived from the same DNA region, but are transcribed from opposite DNA strands. Thus, they are antisense RNAs to each other. The *Xist* RNA coats the inactive X chromosome from which it is transcribed and initiates transcriptional silencing of the majority of genes on the inactive X chromosome. The silencing process also involves the recruitment of histone variants and histone modification proteins.

dsRNA and RNAi. RNA interference (RNAi) is a mechanism whereby the production of double-stranded RNA (dsRNA) renders homologous mRNA incapable of being translated. The dsRNA is cleaved into short fragments and then unwound so that the antisense RNA can pair with the homologous mRNA. The resulting duplexed regions signal degradation or physically block translation by ribosomes. While the mechanism of RNAi is not fully understood, some components of the pathway have been elucidated (**Fig. 4**).

Dicer is a family of ATP-dependent RNase III enzymes that cut long dsRNA sequences into small (approximately 20 bases) RNA sequences. Dicer functions in two pathways of RNAi, that of short inhibitory RNAs (siRNA) that lead to mRNA degradation, and microRNAs (miRNA) that inhibit protein translation by binding to complementary sequences in the mRNA. The RNA-induced silencing complex (RISC) is involved in both pathways and mediates the pairing of the antisense siRNA or miRNA to the homologous mRNA. This is an evolutionarily conserved mechanism that is used endogenously by cells to regulate protein production.

Timothy J. Moss, Vincent G. Allfrey, Lori L. Wallrath

Bibliography. A. M. Denli and G. J. Hannon, RNAi: An ever growing puzzle, *TRENDS Biochem. Sci.*, 28: 196–201, 2003; G. Eckstein and D. M. Lilley, *Nucleic Acids and Molecular Biology*, 1994; B. Lewin, *Genes VIII*, Prentice Hall, 2004; J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids*, 1988; D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, W. H. Freeman, 2004; H. Tschesche (ed.), *Modern Methods in Protein and Nucleic Acid Research*, 1990.

Nucleosome

The fundamental histone-containing structural subunit of eukaryotic chromosomes. In most eukaryotic organisms, nuclear deoxyribonucleic acid (DNA) is complexed with an approximately equal mass of histone protein. It was previously thought that the histones formed some sort of coating on the DNA, but in 1974 it was recognized that the DNA-histone complex was in the form of discrete subunits, now termed nucleosomes. *See* DEOXYRIBONUCLEIC ACID (DNA).

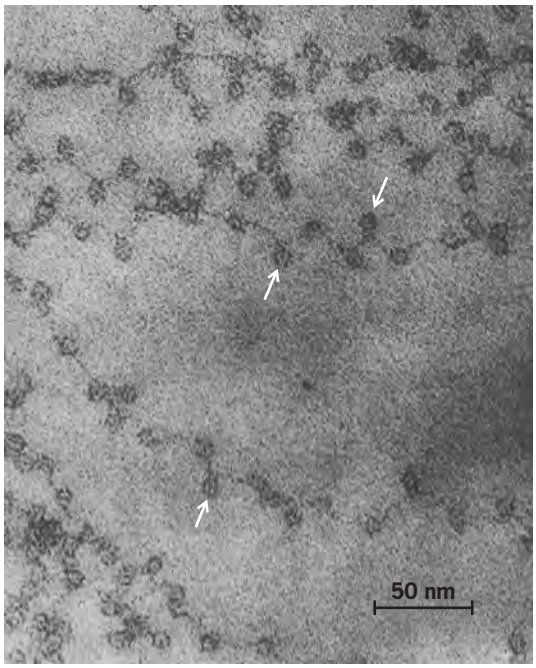

Fig. 1. Electron micrograph of chicken erythrocyte chromatin. The arrows mark individual nucleosomes in ribbonlike zigzag chromatin filaments. (*From D. E. Olins and A. L. Olins, Nucleosomes: The structural quantum in chromosomes, Amer. Sci., 66:704–711, November-December 1978*)

Organization. The nucleosome is organized so that the DNA is exterior and the histones interior. The DNA makes two turns around a core of eight histone molecules, forming a squat cylinder 11 nanometers in diameter and 5.5 nm in height. A short length of linker or spacer DNA connects one nucleosome to the next, forming a nucleosomal chain that has been likened to a beaded string (**Fig. 1**). This basic structure is found in all forms of chromatin (the DNA-protein complex that forms eukaryotic chromosomes), from the dispersed type in nondividing nuclei to the compact chromosomes visible in mitotic and meiotic nuclei. Nucleosomes have been found in all eukaryotic organisms examined, the only exceptions being some sperm nuclei where protamines replace histones during spermatogenesis, and the dinoflagellate algae, which lack histones as major nuclear components.

A chain of adjacent nucleosomes is approximately sixfold shorter than the DNA it contains. Moreover, chains of nucleosomes can self-assemble into thicker fibers in which the DNA packing ratio approaches 35:1 (**Fig. 2**). These observations, and the lack of any obvious catalytic activity, led to the assumption that the primary function of the nucleosome consists of organizing and packing DNA. More recently, however, nucleosomes have also been shown to have important gene regulatory functions.

Nuclease digestion. Nucleosomes may be prepared for biochemical and biophysical analysis by treating nuclei with certain nuclease enzymes that preferentially attack the linker DNA between nucleosomes.

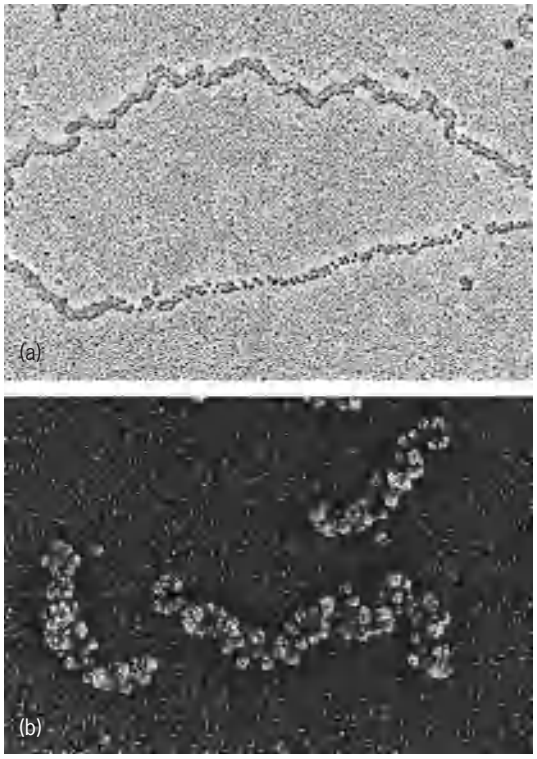

Fig. 2. Chromatin fibers 30 nm in diameter. (*a*) Fibers from a mouse fibroblast nucleus; the upper fiber shows the typical compact structure, while the bottom fiber has been stretched, revealing individual nucleosomes, some in a zigzag arrangement. (*b*) Isolated fibers from chicken erythrocyte nuclei; individual nucleosomes are visible in the fibers. (*From C. L. F. Woodcock, L.-L. Y. Frado, and J. B. Rattner, The higher order structure of chromatin: Evidence for a helical ribbon arrangement, J. Cell Biol., 99:42–52, 1984*)

Careful analysis of the digestion products has revealed several important features of nucleosome arrangements. First, the mean spacing between nucleosomes differs between species, and often between tissues of the same organism. The shortest nucleosome repeat length (NRL; average DNA length from one nucleosome to the next) is ~166 base pairs for yeast and rat cortical neurons, while the longest is ~246 base pairs for sea urchin sperm. Most nuclei examined, however, have spacings in the 180–200 base-pair range. Within one type of nucleus, the spacing is not constant, but varies on either side of the mean value. Even at the level of the DNA base sequence there appears to be no consistently uniform placing of nucleosomes, although there are cases where nucleosomes or groups of nucleosomes are located on specific DNA sequences. In some instances, such "phased" nucleosomes may block access to a DNA regulatory element, and mechanisms for uncovering such elements under appropriate conditions have been demonstrated.

As digestion by micrococcal nuclease (the most commonly used enzyme for preparing nucleosomes) proceeds, first the linker DNA is hydrolyzed, then a brief "pause" in digestion occurs when the mean nucleosomal DNA is ~166 base pairs in length. These 166 base pairs complete two turns around the histone core; the particle so produced has been termed a chromatosome. Further digestion reduces the DNA length to ~145 base pairs, concomitant with the release of one histone molecule (the H1 class of "linker" histones). The product at this point of digestion is the core nucleosome or core particle, and it is upon this product, with its 1.75 turns of DNA, that most structural studies have been performed. The DNA of the nucleosome core is the most resistant to micrococcal nuclease attack, but eventually it too is cleaved.

Some nucleases do not show preferential affinity for the linker DNA. Deoxyribonuclease I (DNase I) is an example of this type: it produces staggered double-stranded cuts at approximately 10-base-pair intervals within nucleosomal DNA, and it has a preferential affinity for some intranucleosomal sites.

Histones. Histones are small basic proteins with common structural features. All have an extended, positively charged amino-terminal region, a globular central domain, and a positively charged carboxy-terminal region, the charges arising mainly from lysine and arginine residues (**Fig. 3**). Four histone species (H2A, H2B, H3, and H4) are present as two copies each in the nucleosome core, while H1 is present as a single copy in the chromatosome. Histone H1 is thought to reside close to the entry-exit point of the DNA, and to "seal" the two turns of DNA. The core histones may self-associate under appropriate conditions to form a tetramer of H3 and H4 and a dimer of H2A and H2B; further association of one $(H3\ H4)_2$ tetramer with two (H2A H2B) dimers forms the octamer (histone core), which is stable in NaCl concentrations above about 1.0 *M*. *See* AMINO ACIDS; PROTEIN.

The core histones are among the most highly conserved proteins known, H4 being the extreme example with only two amino acid changes between pea and calf. However, there are multiple-core histone genes, each coding for a variant protein showing (usually) minor differences in amino acid sequence; and in some cases, crucial roles for these variants have been discovered. For example, a variant of

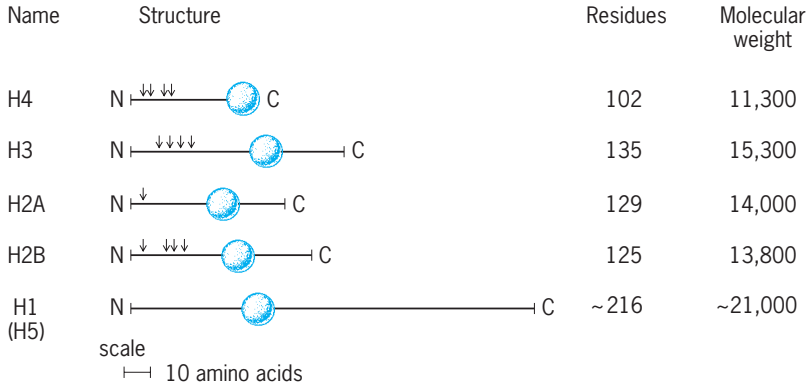

Fig. 3. Several characteristic properties of the five major classes of histones. The globular regions of the histones (spheres) correspond to 70–80 amino acid residues and pack into approximate spheres of 2.5 nm diameter. N and C are the amino and carboxyl termini of the peptide chains. Arrows indicate potential sites of acetylation of histone lysine residues. (*After D. E. Olins and A. L. Olins, Nucleosomes: The structural quantum in chromosomes, Amer. Sci., 66:704–711, November-December 1978*)

histone H3 is a required component of centromere nucleosomes, and conserved from yeast to humans. The differential expression of histone variants has also been associated with specific stages in embryonic development.

Histone modifications and the "histone code." Another form of histone diversity is amino acid modification through the addition of phosphate, acetate, methyl, and poly (adenosine diphosphate-ribose) groups to specific residues. These modifications are the subject of intensive study, and findings have led to the concept of a "histone code," which determines the transcriptional availability (as well as other properties) of nucleosomes and groups of nucleosomes. For example, the acetylation of lysines (especially of histone H4) reduces the net positive charge, and is associated with "looser" chromatin and a greater availability to the transcriptional machinery, while methylation of lysines has the opposite effect. Some modifications may recruit enzyme complexes that, for example, include histone acetyl transferases or histone deacetylases. Histone modifications or groups of modifications may be passed to daughter cells through cell division, providing an epigenetic mechanism for the inheritance of transcriptional states (that is, changes that do not affect the genes themselves but alter their expression).

Forces stabilizing the nucleosome. If nucleosomes are placed in a high-salt solution (for example, 1.0–2.0 *M* NaCl), the histone octamer dissociates from the DNA, as would be expected for a complex in which electrostatic interaction between the negatively charged phosphate groups on the DNA and the positive charges on the lysine and arginine residues in the histones predominated. The finding that the highly charged amino-termini project from the histone octamer led to the suggestion that these domains contributed most of the electrostatic "glue" that stabilized nucleosomes. However, experiments showing that nucleosomelike particles could be formed by using octamers that lacked the amino-terminal regions showed clearly that this was not the case, and it now seems likely that these histone domains function in internucleosomal interactions and transcriptional regulation. While the octamer is stable in high salt, it dissociates if the ionic strength is lowered or if it is exposed to urea. Thus, hydrophobic interactions are thought to be largely responsible for maintaining the integrity of the histone core.

Detailed structure. The publication in 1997 of a 2.8 Å resolution x-ray structure of the nucleosome constituted a major breakthrough in chromatin research. This technical tour de force was the result of over a decade of work by T. Richmond and colleagues and required producing fully defined nucleosome core particles comprising a palindromic DNA sequence, and bacterially synthesized core histones, together with data collection at liquid nitrogen temperatures using synchrotron radiation. **Figure 4** contrasts the ~20 Å structure and the 2.8 Å structure. The final structure has a twofold axis of symmetry, with the $(H3\ H4)_2$ tetramer forming a U-shaped complex within the two turns of DNA, and the two (H2A H2B) dimers attached to either side of the tetramer. Nucleosomal DNA is not smoothly coiled on the histone octamer, but variously distorted and kinked, and the highly conserved portions of the core histones are folded in a manner that maximizes histone-histone interactions, and also provides a coiled ramp on the outside where positive charges are concentrated. The bulk of the N and C termini of the core histones are not resolved in x-ray structures, suggesting that they are highly mobile even in nucleosome crystals. However, the sites where these histone domains leave the nucleosome is clearly defined. For example, the long N termini of both H3 molecules exit near the linker DNA entry-exit site, suggesting that they may be associated with linker DNA. *See* X-RAY CRYSTALLOGRAPHY.

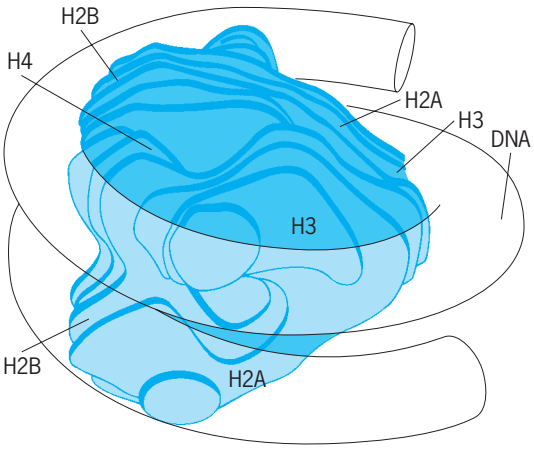

(a)

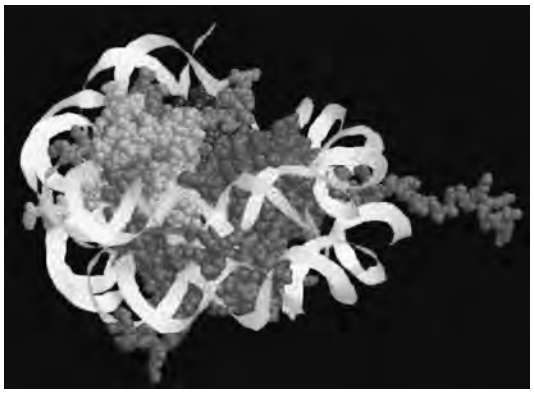

(b)

Fig. 4. Detailed structures of nucleosomes (*a*) Model of the nucleosome core particle at 20 Å resolution (*after R. D. Kornberg and A. Klug, The nucleosome, Sci. Amer., 244(2):52–79, 1981*). (*b*) X-ray structure of the nucleosome core particle at 2.8 Å resolution (*adapted from MedWeb: Nucleosome structure: http://medweb.bham.ac.uk/research/chromatin/nucleosome.html*).

Transcription. The fate of nucleosomes during the passage of a ribonucleic acid (RNA) polymerase molecule along the DNA has been extensively studied, and appears to be system-dependent. In the case of some ribosomal RNA genes where polymerase molecules are closely packed on the DNA, it is clear that nucleosomes as such cannot exist during

transcription, since electron micrographs show that the genes are extended almost to their full length. Whether the histone octamers are completely dissociated from the DNA during this process or the nucleosomes are merely unfolded is not known. In more typical genes, nucleosomelike particles have been observed between the widely spaced polymerases, suggesting that the perturbations that occur during transcription are of a transitory nature. In some systems, evidence suggests that octamers may be temporarily transferred to "chaperone" complexes during transcription, while other cases indicate a more complex looping out of DNA as the polymerase enters the nucleosome, followed by reassociation with the octamer as the polymerase completes its transit. There is general agreement that actively transcribed genes are depleted in or lack histone H1. *See* GENE ACTION.

Nucleosome remodeling. Nucleosome-remodeling complexes, required for the efficient transcription of some genes, are recently recognized components of the gene regulatory system. These typically contain many proteins (for example, the yeast SWI/SNF complex includes 11 different types of protein and has a mass of over 1 megadalton) including an active ATPase. Remodeling is ATP-dependent and results in nucleosomal DNA becoming more available for transcription. The molecular mechanism of remodeling is not well understood, but is likely to involve a loosening of the DNA-histone interactions, leading to greater mobility and perhaps repositioning of histone octamers.

Assembly and replication. Nuclear DNA is replicated semiconservatively; that is, the double helix splits in two, and new strands are formed on each half in the immediate vicinity of the replication fork. In chromatin, this process is complicated by the presence of nucleosomes, which also must be duplicated along with the DNA. Some principles about the mechanism of chromatin replication are now fairly well established. Unlike the DNA itself, the histone octamers are not replicated semiconservatively. Instead, the old histone octamers remain intact as the replication fork passes along the chromatin strand and are segregated at random to the daughter strands. The mode of formation of new nucleosomes in the wake of the replication fork is complex and still incompletely understood. New nucleosomes, usually bearing modifications that reduce their net charge and complexed with "chaperone" proteins, are formed by the sequential addition of newly synthesized H3 and H4, followed by H2A and H2B, and finally H1. Replication is also accompanied by brief changes in nucleosome spacing: new nucleosomes tend to be closely packed and then gradually change to the mature spacing.

Systems designed for the rapid assembly of nucleosomes (for example, oocytes) contain a large pool of stored histone, and injected DNA is rapidly converted into nucleosomal chromatin. This process can be mimicked in vitro by the incubation of DNA, histones, and oocyte extracts. Effective artificial assembly of nucleosomes can be mediated by factors such as polyglutamic acid, which, like DNA, partially neutralizes the electrostatic charges of the histones. Indeed, by using this polyanion it is possible to produce nucleosomal chromatin from DNA and the four core histones; the subsequent addition ofhistone H1

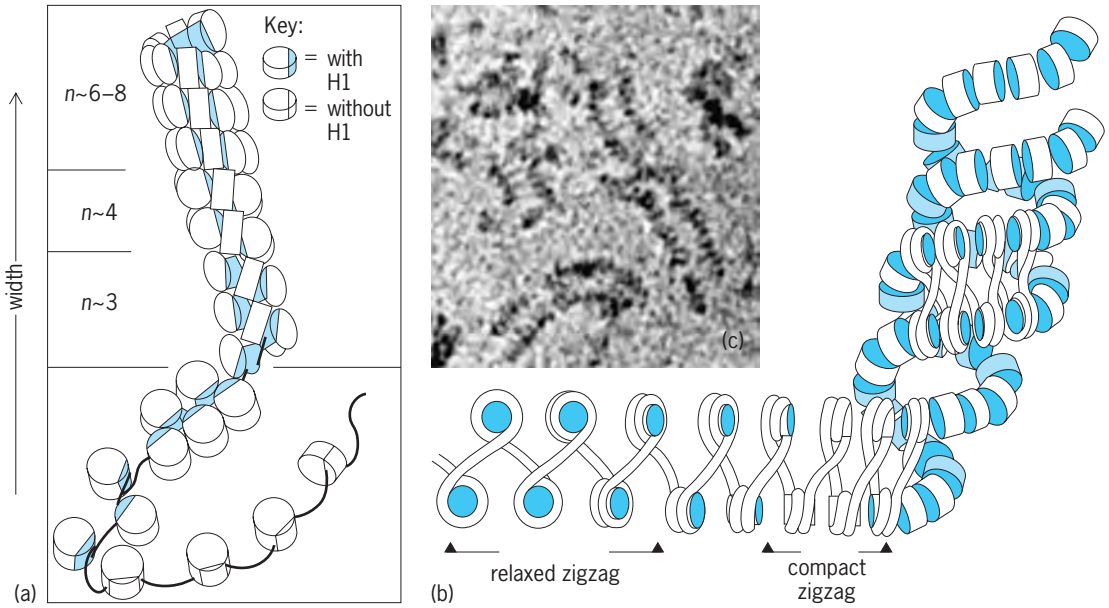

Fig. 5. Two proposed structures for the arrangement of nucleosomes in the 30-nm fiber. (*a*) Superhelical or solenoidal (*after F. Thoma, T. Koller, and A. Klug, Involvement of H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin, J. Cell Biol., 83:403–427, 1979*). (*b*) Helical ribbon (*after C. L. F. Woodcock, L.-L. Y. Frado, and J. B. Rattner, The higher order structure of chromatin: Evidence for a helical ribbon arrangement, J. Cell Biol., 99:42–52, 1984*). (*c*) Micrograph of crosslinked compact chromatin supports a zigzag arrangement. The paired structures are about 30 nm in diameter (*reprinted with permission from B. Dorigo et al., Nucleosome arrays reveal the two-start organization of the chromatin fiber, Science, 306:1571–1573, 2004. © 2004 American Association for the Advancement of Science*).

under appropriate conditions then allows the native nucleosomal spacing to be generated.

Higher-order packing. The normal state of chromatin in nuclei is not the extended nucleosomal fiber but a much more compact form. Isolated chromatin adopts a fiberlike conformation, about 30 nm in diameter at the appropriate ionic strength (Fig. 2). In the laboratory, the transition from extended to compact fiber can be reversibly induced by altering the ionic composition of the solvent, maximal compaction being observed in about 150 mM monovalent ions, and maximal extension below 5 mM. Above ~100 mM monovalent ions, or lower concentrations of divalent or polyvalent cations, chromatin fibers tend to self-associate. Compaction to the 30-nm fiber conformation requires the presence of histone H1, which is also required to establish a basic zigzag architecture (Fig. 1). Despite the ease with which chromatin fibers can be prepared at any desired state of compaction, the way in which the nucleosomes are arranged in the condensed fibers is still controversial due to the difficulty of determining the locations of nucleosomes and linker DNA in compact chromatin fibers.

The most widely discussed proposal for the structure of the 30-nm chromatin fiber is the superhelix or solenoid arrangement, in which the nucleosomal chain is coiled into a continuous shallow helix with 6–8 nucleosomes per turn (**Fig. 5***a*). A second suggestion based on evidence that an intermediate state of compaction is a ribbonlike zigzag structure (for example, regions indicated by arrows in Fig. 1) is that the ribbon arrangement is conserved and simply coils or compresses to form the fiber (Fig. 5*b*). A major difficulty in studying the compact fiber is its relatively disorganized state, which appears to be an intrinsic property of chromatin with variable linker lengths. Force-extension data from experiments in which isolated fibers are pulled using molecular tweezers support a zigzag over a helical organization. An elegant experiment reported in 2004 analyzed arrays of nucleosomes reconstituted on DNA with regularly spaced nucleosome positioning sequences. The nucleosomes were engineered so that where the arrays were in a compact state, close neighbors would be chemically crosslinked. When viewed in the electron microscope, the crosslinked arrays appeared as ladderlike parallel rows (Fig. 5*c*). This finding strongly supports the zigzag arrangement, in which alternate nucleosomes are adjacent in compacted chromatin [see region labeled "compact zigzag" in Fig. 5*b*], over the solenoidal arrangement, in which nearest neighbor nucleosomes along DNA are also nearest neighbors in compact chromatin [see upper region in Fig. 5*a*]. *See* CHROMOSOME.

C. L. F. Woodcock

Bibliography. K. Luger et al., X-ray structure of the nucleosome core particle at 2.8 Å resolution, *Nature*, 389:251–260, 1997; B. D. Strahl and C. D. Allis, The language of covalent histone modifications. *Nature*, 403:41–43, 2000; A. Wolffe, *Chromatin: Structure and Function*, 3d ed., 1998.

Nucleosynthesis

Theories of the origin of the elements involve synthesis with charged and neutral elementary particles (neutrons, protons, neutrinos, photons) and other nuclear building blocks of matter, such as alpha particles. The theory of nucleosynthesis comprises a dozen distinct processes, including big bang nucleosynthesis, cosmic-ray spallation in the interstellar medium, and static or explosive burning in various stellar environments (hydrogen-, helium-, carbon-, oxygen-, and silicon-burning, and the, *s*-, *r*-, *p*-, γ-, and ν-processes). Acceptable theories must lead to an understanding of the cosmic abundances observed in the solar system, stars, and the interstellar medium. The curve of these abundances is shown in **Fig. 1**. Hydrogen and helium constitute about 98% of the total element content by mass and more than 99.8% by number of atoms. There is a rapid decrease with increasing nuclear mass number A, although the abundance of iron-group elements like iron and nickle are remarkably large. The processes of nucleosynthesis described in this article attempt to explain the observed pattern. *See* ELEMENTS, COSMIC ABUNDANCE OF.

Observations of the expanding universe and of the 3-K background radiation indicate that the universe originated in a primordial event known as the big bang about 15×10^9 years ago. Absence of stable mass 5 and mass 8 nuclei preclude the possibility of synthesizing the major portion of the nuclei of masses greater than 4 in those first few minutes of the universe, when the density and temperature were sufficiently high to support the necessary nuclear reactions to synthesize elements. *See* BIG BANG THEORY; COSMIC BACKGROUND RADIATION; COSMOLOGY.

The principal source of energy in stars is certainly nuclear reactions which release energy by fusion of lighter nuclei to form more massive ones. Since the masses of the products are less than those of the constituent parts, the energy released is $E = mc^2$, where m is the mass difference and c is the velocity of light. In 1957 the collaboration of E. M. Burbidge, G. R. Burbidge, W. A. Fowler, and F. Hoyle, and the work of A. G. W. Cameron, laid the foundations for research into the synthesis of the elements, stellar energy generation, and stellar evolution.

Evidence of nucleosynthesis. There is considerable evidence that nucleosynthesis has been going on in stars for billions of years. Observations show that the abundance ratio of iron and heavier elements to hydrogen decreases with increasing stellar age. The oldest known stars in the disk of the Milky Way Galaxy exhibit a ratio 10,000 times smaller than in the Sun. This low ratio is understood on the basis of element synthesis occurring in previous-generation stars that evolved to the point of exploding as supernovae, thus enriching the interstellar medium with the nuclei that were synthesized during their lifetimes and by the explosive synthesis that accompanies the ejection of the stellar envelope.

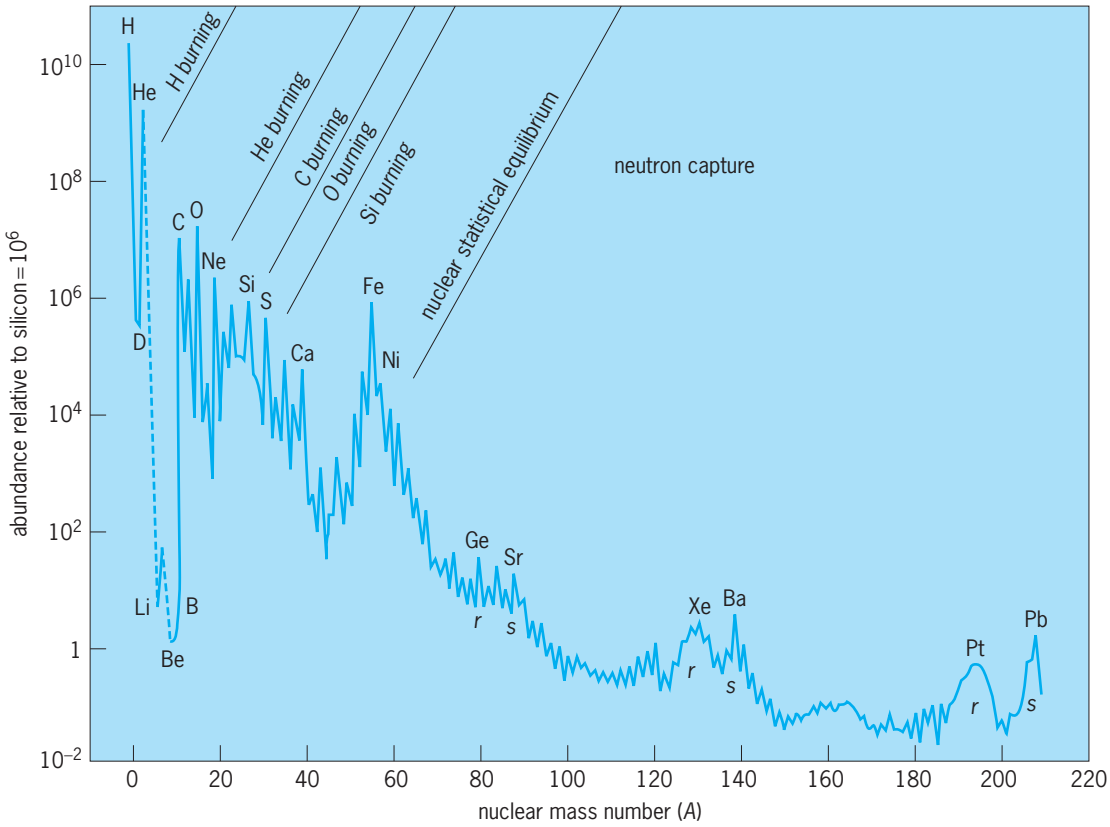

Fig. 1. Schematic diagram of cosmic abundance as a function of nuclear mass number (*A*), based on data of H. E. Suess and H. C. Urey. Predominant nucleosynthetic processes are indicated. (*After E. M. Burbidge et al., Synthesis of elements in stars, Rev. Mod. Phys., 29:547–650, 1957*)

Later-generation stars were then formed from enriched gas and dust. *See* MILKY WAY GALAXY; STELLAR POPULATION.

The most direct evidence that nucleosynthesis is an ongoing process in stars is provided by the spectra of N-type carbon stars in which convection has dredged the *s*-process element technetium to the surface. The technetium is observed through specific absorption lines in the stellar spectrum, and must have been freshly produced, because there is no stable isotope of this element. No technetium is found naturally on Earth, and the stars showing technetium are obviously much older than 2.6×10^6 years, the half-life of its most stable isotope. *See* CARBON STAR; TECHNETIUM.

Another piece of evidence supporting ongoing element synthesis in the Galaxy comes from gamma-ray astronomy. Winds from massive Wolf-Rayet stars, and the ejected envelopes of supernovae, release small amounts of radioactive ^{26}Al into the interstellar medium. The decay of this isotope results in the emission of gamma-ray line photons with energy $E_\gamma = 1.809$ MeV. At any given time, the Milky Way is believed to contain about one solar mass of this isotope, distributed throughout the disk. The decay of ^{26}Al thus leads to a diffuse glow that is confined to the galactic plane. This glow has been detected with gamma-ray telescopes (**Fig. 2**). *See* GAMMA-RAY ASTRONOMY; WOLF-RAYET STAR.

A third piece of evidence for ongoing element formation in stars lies in refractory dust grains recovered from meteorites. These micrometer-sized grains (about the size of a human blood cell or smaller) have isotopic patterns that greatly differ from the abundance pattern shown in Fig. 1, which strongly indicates that the grains predate the solar system and in fact condensed in the outflow from stars before the stellar debris fully mixed with the interstellar medium. Certain grains of silicon carbide and graphite have large anomalies in their abundance of ^{44}Ca. This anomaly resulted from in situ decay of ^{44}Ti (half-life of about 60 years), which, in turn, means that these grains condensed live ^{44}Ti no more than several years after that isotope was produced. Similar isotopic anomalies in other presolar meteoritic grains further demonstrate that nucleosynthesis occurs in stars. *See* COSMOCHEMISTRY; METEORITE.

Finally, certain primitive meteorites show evidence that radioactivities such as ^{26}Al were alive in the early solar system. This indicates that nucleosynthesis was occurring shortly before the Sun's birth. Indeed, the abundance inferred for these radioactivities is so high that it is likely that they were injected into the forming solar system from a nearby, exploding star.

Hydrogen burning. The first static burning process of nucleosynthesis converts hydrogen to helium. In stars of 1.2 or less solar masses, this process

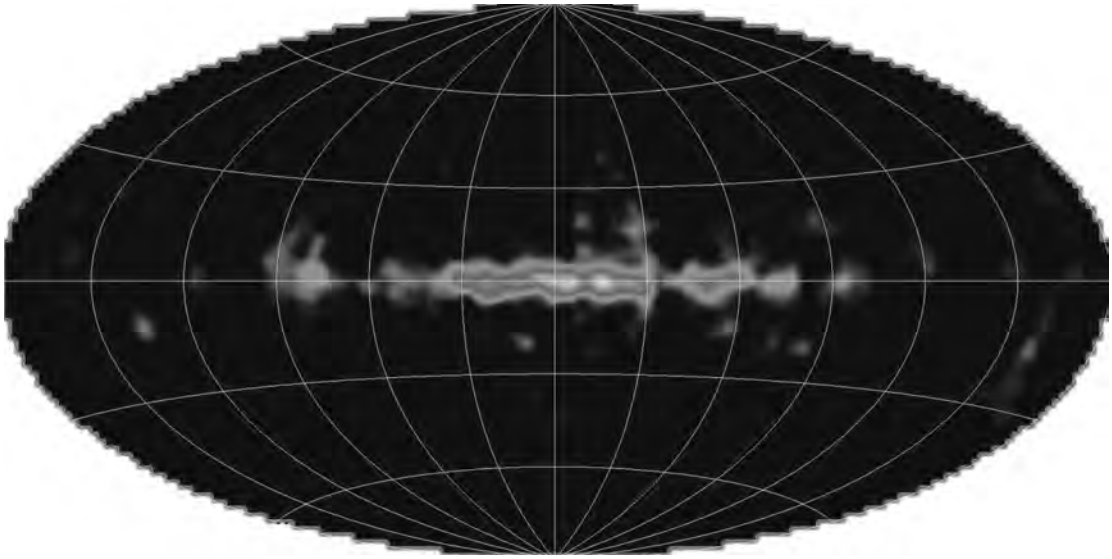

Fig. 2. Map of the Milky Way Galaxy, showing intensity of 1.809-MeV gamma-ray line emission from the decay of the radioactive isotope ^{26}Al, predominantly made in massive stars. Detection of this emission by gamma-ray telescopes is evidence of ongoing element synthesis in the Galaxy. (*COMPTEL/MPE map*)

occurs via the proton-proton chain. This chain was also responsible for much of the element synthesis during the big bang, producing the bulk of deuterium and helium, and some of the ^{7}Li, observed today. In more massive stars where the central temperatures exceed 2×10^{7} K, hydrogen burning is accomplished through proton captures by carbon, nitrogen, and oxygen nuclei, in the carbon-nitrogen-oxygen (CNO) cycles, to form ^{4}He. The product (ash) of hydrogen burning is helium, but much of the helium produced is consumed in later stages of stellar evolution or is locked up forever in stars of lower mass that never reach the temperatures required to ignite helium burning. The observed abundances of some carbon, nitrogen, oxygen, and fluorine nuclei are attributed to hydrogen burning in the CNO cycles. *See* CARBON-NITROGEN-OXYGEN CYCLES; NUCLEAR FUSION; PROTON-PROTON CHAIN.

Helium burning. When the hydrogen fuel is exhausted in the central region of the star, the core contracts and its temperature and density increase. Helium, the ash of hydrogen burning, cannot be burned immediately due to the larger nuclear charge of helium ($Z = 2$) producing a much higher Coulomb barrier against fusion. When the temperature eventually exceeds about 10^{8} K, helium becomes the fuel for further energy generation and nucleosynthesis. The basic reaction in this thermonuclear phase is the triple-alpha process in which three ^{4}He nuclei (three alpha particles) fuse to form ^{12}C, a carbon nucleus of mass 12 (atomic mass units). Capture of an alpha particle by ^{12}C then forms ^{16}O, symbolically written as $^{12}\mathrm{C} + {}^{4}\mathrm{He} \rightarrow {}^{16}\mathrm{O} + \gamma$, or simply $^{12}\mathrm{C}(\alpha,\gamma)^{16}\mathrm{O}$, where γ represents energy released in the form of electromagnetic radiation. Other reactions that are included in helium burning are $^{16}\mathrm{O}(\alpha,\gamma)^{20}\mathrm{Ne}$, $^{20}\mathrm{Ne}(\alpha,\gamma)^{24}\mathrm{Mg}$, $^{14}\mathrm{N}(\alpha,\gamma)^{18}\mathrm{F}$, and $^{18}\mathrm{O}(\alpha,\gamma)^{22}\mathrm{Ne}$. Fluorine-18, produced when ^{14}N captures an alpha particle, is unstable and decays by emitting a positron (e^{+}) and a neutrino (ν) to form ^{18}O [in short, $^{14}\mathrm{N}(\alpha,\gamma)^{18}\mathrm{F}(e^{+},\nu)^{18}\mathrm{O}$]. Because there is likely to be ^{13}C in the stellar core if hydrogen burning proceeded by the carbon-nitrogen-oxygen cycles, the neutron-producing reaction $^{13}\mathrm{C}(\alpha,n)^{16}\mathrm{O}$ should also be included with the helium-burning reactions. The neutrons produced by this reaction are probably responsible for the bulk of *s*-process nucleosynthesis (discussed below). Helium burning is probably responsible for much of the ^{12}C observed in the cosmic abundances, although in more massive stars the later burning stages will consume the ^{12}C produced earlier by helium burning. *See* NUCLEAR REACTION.

Carbon burning. Upon exhaustion of the helium supply, if the star has an initial mass of at least 8 solar masses, gravitational contraction of the stellar core can lead to a temperature exceeding 5×10^{8} K, where it becomes possible for two ^{12}C nuclei to overcome their high mutual Coulomb-repulsion barrier and fuse to form ^{20}Ne, ^{23}Na, and ^{24}Mg through reactions such as $^{12}\mathrm{C}(^{12}\mathrm{C},\alpha)^{20}\mathrm{Ne}$, $^{12}\mathrm{C}(^{12}\mathrm{C},p)^{23}\mathrm{Na}$, and $^{12}\mathrm{C}(^{12}\mathrm{C},\gamma)^{24}\mathrm{Mg}$. Carbon burning can produce a number of nuclei with masses less than or equal to 28 through further proton and alpha-particle captures.

Oxygen burning. Carbon burning is followed by a short-duration stage, sometimes referred to as neon burning, in which ^{20}Ne disintegrates by the reaction $^{20}\mathrm{Ne}(\gamma,\alpha)^{16}\mathrm{O}$. The alpha particle released is then captured by a remaining ^{20}Ne nucleus to produce ^{24}Mg; thus, the effective neon burning reaction is $^{20}\mathrm{Ne} + {}^{20}\mathrm{Ne} \rightarrow {}^{16}\mathrm{O} + {}^{24}\mathrm{Mg}$. The eventual result is that much of the carbon from helium burning becomes oxygen, which supplements the original oxygen formed in helium burning. This stage is followed by the fusion of oxygen nuclei at much higher temperatures. (Temperatures greater than 10^{9} K are required for ^{16}O nuclei to overcome their mutual Coulomb barrier.) Some relevant reactions for oxygen burning are

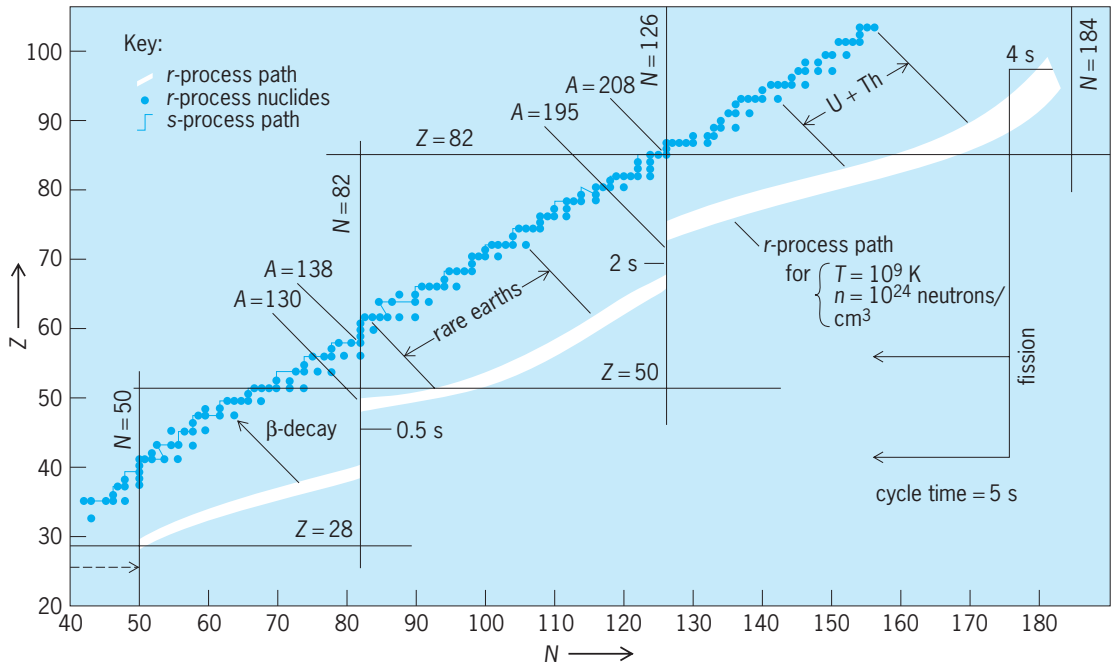

Fig. 3. Neutron-capture paths among the nuclei plotted in the *Z-N* plane, showing nuclei identified by the number of protons (*Z*) and number of neutrons (*N*) in the nucleus. (*After P. A. Seeger, W. A. Fowler, and D. D. Clayton, Nucleosynthesis of heavy elements by neutron capture, Astrophys. J. Suppl., 11(97):121–166, 1965*)

$^{16}O(^{16}O,\alpha)^{28}Si$, $^{16}O(^{16}O,p)^{31}P$, and $^{16}O(^{16}O,\gamma)^{32}S$. Nuclei of masses up to $A = 40$ may be produced in this phase through proton, neutron, and alpha-particle captures.

Silicon burning. This process commences when the temperature exceeds about 3×10^9 K. In this phase, photodisintegration of ^{28}Si and other intermediate-mass nuclei around $A = 28$ produces copious supplies of protons, neutrons, and alpha particles. These particles capture on the seed nuclei left from previous burning stages and thus produce new isotopes up to mass 60, resulting in the buildup of the abundance peak near $A = 56$ (Fig. 1). Because the binding energy per nucleon is a peak for the iron-group (A near 60) nuclei, further fusion reactions no longer release energy. Silicon burning is therefore the end stage of stellar burning. Production of nuclei with mass higher than A near 60 therefore occurs by special processes that are ancillary to the mainline stellar burning stages.

The s-process. Because neutrons are neutral particles, their capture is not affected by the Coulomb barrier that inhibits charged-particle reactions. If the number of neutrons per seed nucleus is small, so that time intervals between neutron captures are long compared to the beta-decay lifetimes of unstable nuclei that are formed, the *s*-process (slow process) takes place. The seed nuclei are predominantly in the iron peak (Fig. 1), but the abundances of low-mass nuclei are also affected by neutron-capture processing. In the *s*-process, if neutron capture produces a nucleus that is unstable (due to an excess number of neutrons), the newly formed nucleus undergoes beta decay to a stable isobar by emitting an electron and an antineutrino. The resulting nucleus eventually captures another neutron, and the capture-decay step repeats. The presence of free neutrons leads to a chain of capture-decay events that drives the abundances along a unique *s*-process path that zigzags along a single line in the nuclear *Z-N* diagram (**Fig. 3**) near the beta-stable nuclei (valley of beta stability). Given enough neutrons, the *s*-process synthesizes nuclei of masses up to 209, when alpha decay becomes a deterrent to further buildup by neutron capture. *See* RADIOACTIVITY.

The predominant site of the *s*-process is believed to be located in asymptotic giant branch stars. As these stars evolve on the asymptotic giant branch in the Hertzsprung-Russell diagram, thermonuclear energy is provided by multiple shells in which hydrogen and helium burning takes place. At some point in the evolution, the stellar interior becomes thermally unstable and a sequence of pulses with convective mixing takes place. Based on the observational results provided by heavy-element *s*-process rich stars, there is strong evidence that the reaction $^{13}C(\alpha,n)^{16}O$ is the dominant source of neutrons. The ^{13}C synthesized at the hydrogen-helium interface on the asymptotic giant branch in between thermal pulses is burned via the (α,n) reaction during this radiative phase at temperatures of order 10^8 K. The neutrons released by this process then drive the *s*-process. An additional source of neutrons is the reaction $^{22}Ne(\alpha,n)^{25}Mg$, although the neutron-capture reactions induced reset the abundances of key isotopes at the end of thermal pulses.

Convective mixing repeatedly brings *s*-process material to the stellar surface. The technetium observed in some stars provides direct evidence for the *s*-process. Since cross sections for neutron

capture by nuclei with magic neutron numbers ($N = 50$, 82, 126) are small, the *s*-process builds a high abundance of these nuclei and produces peaks near $A = 87$, 138, and 208, labeled *s* in the cosmic elemental abundance curve (Fig. 1). *See* MAGIC NUMBERS; NUCLEAR STRUCTURE.

The r-process. This rapid process occurs when a large neutron flux allows rapid neutron capture, so that seed nuclei capture many neutrons before undergoing beta decay. The rapid neutron capture takes the nuclei far away from the valley of beta stability, into the regime of extremely neutron-rich nuclei. The time scale for the *r*-process is very short, 1–100 s, and the abundance of free neutrons is very large. These conditions are found deep inside the interiors of exploding massive stars, supernovae. The *r*-process can synthesize nuclei all the way into the transuranic elements. The *r*-process path in the chart of nuclei is shown in Fig. 3, where representative conditions for temperature (10^9 K) and neutron density ($n_n \sim 10^{24}$ cm^{-3}) were chosen. Once these conditions cease to exist, nuclei along this path very quickly beta-decay (along the diagonal lines in Fig. 3) toward the valley of stability. The circles in Fig. 2 indicate the stable end products reached by these decays. The times shown are intervals to reach locations along the *r*-process path for these conditions. *See* SUPERNOVA.

At Z near 94, the transuranic elements are unstable to neutron-induced fission, which terminates the *r*-process path. The resulting fission leads to the production of intermediate-mass nuclei which act as fresh seed nuclei, introducing a cycling effect. The peaks at $A = 80$, 130, and 195, labeled *r* in Fig. 1, are due to the magic neutron numbers $N = 50$, 82, and 126, and the subsequent decays of unstable nuclei. *See* NUCLEAR FISSION.

The p-process. This process produces heavier elements on the proton-rich side of the beta valley. The *p*-nuclei are blocked from formation by stable nuclei produced by either the *r*- or *s*-process. The major task for *p*-process theory is thus to find ways, other than beta decay, to process the more abundant *r*- and *s*-nuclei into the less abundant (Fig. 1) *p*-nuclei. Burbidge, Burbidge, Fowler, and Hoyle suggested two possible mechanisms that are still under consideration: radiative proton capture, and gamma-induced neutron, proton, or alpha-particle removal reactions. In both cases the temperature should be in excess of $2–3 \times 10^9$ K.

The sites of the *p*-process have not yet been determined with confidence, and the models do not yet produce the *p*-process abundance pattern satisfactorily. The *p*-process was thought to take place in high-temperature, proton-rich zones in supernovae, but detailed studies have shown that the right conditions are very hard to find in these stellar environments. Also, several *p*-nuclei can be produced by other distinct processes, such as the alpha-rich freeze-out and the ν-process. This suggests that *p*-nuclei owe their origin to a variety of different synthesis processes and environments; no single site is likely to produce all the *p*-nuclei.

Current theories of the *p*-process consider the premature termination of (γ,n)-, (γ,p)- and (γ,α)-induced "melting" of nuclei in the shock-heated oxygen-neon layers in type II supernovae, or in exploding carbon-oxygen white dwarfs (type Ia supernovae). When only photodisintegration is important, "γ-process" is used as an alternate process name.

The ν-process. The neutrino (ν) flux emitted from a cooling proto neutron star alters the yields of explosive nucleosynthesis from type II supernovae. Inelastic scattering of neutrinos (all flavors) off abundant nuclei excites states that can decay via single or multiple nucleon emission. The ν-process is probably responsible for significant contributions to the synthesis in nature of about a dozen isotopes. While the neutrino interaction cross section with matter is extremely small (about 10^{-44} cm^2), the high neutrino energies and the large number flux close to the collapsing iron core of a massive star lead to significant synthesis of nuclei, either directly or indirectly.

An example of a direct ν-process reaction is $^{12}C(\nu,\nu'p)^{11}B$, where ν' indicates the inelastically scattered neutrino emerging with a reduced energy. Even a very small change in the abundance of ^{12}C (in the carbon layer of the supernova) implies a large contribution to ^{11}B synthesis, because the cosmic abundance of boron is small compared to that of carbon (Fig. 1). The ν-process thus contributes greatly to lithium, beryllium, and boron synthesis. A similar example is ^{19}F, which can be efficiently made by $^{20}Ne(\nu,p)^{19}F$ in the neon shell of the star. An example of an indirect (two-step) contribution of neutrino spallation is $^4He(\nu,\nu'p)^3H$ in the helium shell. The ^3H liberated by the neutrino interaction subsequently interact with the abundant helium to form ^7Li via $^4He(^3H,\gamma)^7Li$.

Two of the rarest stable nuclei in nature, ^{138}La and ^{180}Ta, could owe their abundance to neutral-current neutrino interactions with their abundant neighbors in the chart of nuclei via $^{181}Ta(\nu,\nu'n)^{180}Ta$, and $^{139}La(\nu,\nu'n)^{138}La$. In addition, the charged-current reaction $^{138}Ba(\nu,e^-)^{138}La$ can contribute significantly. While some contribution to element synthesis by neutrino-induced reactions must occur in nature, the estimated yields are somewhat uncertain due to the difficulties in the calculation of neutrino transport during the supernovae and also during the longer initial cooling phase (approximately 10 s) of the proto neutron star. The rates are very sensitive to the high-energy tail of the nonthermal neutrino-spectrum. *See* NEUTRINO.

The LiBeB process. The bulk of the light elements lithium, beryllium, and boron found in the cosmic abundance curve (Fig. 1) cannot have survived processing in stellar interiors because they are readily destroyed by proton capture. Although some ^7Li originated in the big bang, primordial nucleosynthesis cannot be responsible for the bulk of the ^7Li in existence. Spallation of more abundant nuclei such as carbon, nitrogen, and oxygen by high-energy protons and alpha particles can account for the

low-abundance nuclides ^{6}Li, ^{9}Be, ^{10}B, ^{11}B, and for some ^{7}Li. The canonical process is spallation of carbon nitrogen, and oxygen nuclei in the interstellar medium by fast light particles, such as alpha particles and protons, which are abundant in the gas between the stars. These high-energy particles are referred to as cosmic rays, leading to the term "cosmic-ray spallation (CRS) process."

Supernova-induced shock acceleration in the interstellar medium provides the energetic cosmic-ray particles, which then interact with the carbon, nitrogen, and oxygen nuclei that are also present in the interstellar medium. Observations of lithium, beryllium, and boron abundances in metal-poor (population II) stars indicate that the (number) abundance ratio ^{7}Li/H in these stars is about 1.5×10^{-10}, regardless of stellar metallicity, measured by the logarithmic iron-to-hydrogen ratio relative to the Sun: $[Fe/H] = \log(Fe/H) - \log(Fe/H)_{\odot}$, where the subscript indicates that the last term refers to the solar ratio. The main source of this lithium is the big bang, but some of it could also be made in the ν-process. On the other hand, ^{6}Li is proportional to [Fe/H] and so are boron (^{10}B and ^{11}B) and beryllium (^{9}Be). While ^{11}B could be made by the ν-process, ^{10}B and ^{9}Be are "orphans" of nucleosynthesis, made neither in stars nor in the big bang—instead their origin lies 100% in the cosmic rays. The observed nearly linear scaling with [Fe/H] of both beryllium and boron implies that both elements are of primary origin, which suggests a modified CRS process. One proposal involves a reverse CRS process, in which accelerated carbon, nitrogen, and oxygen nuclei interact with interstellar hydrogen and helium. In one version of this proposal, the cosmic-ray particles are assumed to be accelerated supernova ejecta, instead of being particles accelerated (by supernova shocks) out of the interstellar medium. These issues are not yet settled, but it is clear that lithium, beryllium, and boron nucleosynthesis will provide powerful diagnostic tools to study the important dynamic interaction of stellar explosions with the interstellar medium. *See* COSMIC RAYS.

Conclusions. Although stars spend the majority of their lifetimes in hydrogen-burning phases and most of the remaining time in helium burning, the vast majority of the elements are synthesized in the relatively brief time spans of later stages of evolution and the processes described above. Figure 1 serves as a summary of those processes and the mass ranges of elements for which they are responsible. However, there are many unanswered questions regarding which processes and which sites are responsible for certain elements. Shocks propagating through the supernova ejecta and the exposure to neutrinos from the newborn neutron star modify stellar synthesis at the very end of a massive star' life in ways that are not fully understood.

While many of the scenarios presented in 1957 proved correct, new facets of nucleosynthesis theory were discovered and new processes and sites identified. Astrophysicists are increasing the knowledge of the thermodynamic conditions prevailing in stars. As a result, further modifications in the theory of the processes of nucleosynthesis are quite likely. *See* STELLAR EVOLUTION.

Dieter H. Hartmann; Bradley S. Meyer

Bibliography. D. Arnett, *Supernovae and Nucleosynthesis*, Princeton University Press, 1996; D. D. Clayton, *Principles of Stellar Evolution and Nucleosynthesis*, University of Chicago Press, 1983; B. E. J. Pagel, *Nucleosynthesis and Chemical Evolution of Galaxies*, Cambridge University Press, 1997; C. E. Rolfs and W. S. Rodney, *Cauldrons in the Cosmos*, University of Chicago Press, 1988; G. Wallerstein et al., Synthesis of the elements in stars: Forty years of progress, *Rev. Mod. Phys.*, 69:995–1084, 1997.

Nucleotide

A cellular constituent that is one of the building blocks of ribonucleic acids (RNA) and deoxyribonucleic acid (DNA). In biological systems, nucleotides are linked by enzymes in order to make long, chain-like polynucleotides of defined sequence. The order or sequence of the nucleotide units along a polynucleotide chain plays an important role in the storage and transfer of genetic information. Many nucleotides also perform other important functions in biological systems. Some, such as adenosine triphosphate (ATP), serve as energy sources that are used to fuel important biological reactions. Others, such as nicotinamide adenine dinucleotide (NAD) and coenzyme A (CoA), are important cofactors that are needed to complete a variety of enzymatic reactions. Cyclic nucleotides such as cyclic adenosine monophosphate (cAMP) are often used to regulate complex metabolic systems. Chemically modified nucleotides such as fluoro-deoxyridine monophosphate (Fl-dUMP) contain special chemical groups that are useful for inactivating the normal function of important enzymes. These and other such compounds are widely used as drugs and therapeutic agents to treat cancer and a variety of other serious illnesses. *See* ADENOSINE TRIPHOSPHATE (ATP); COENZYME; CYCLIC NUCLEOTIDES; NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD).

Classification. Nucleotides are generally classified as either ribonucleotides or deoxyribonucleotides (**Fig. 1**). Both classes consist of a phosphorylated pentose sugar that is linked via an *N*-glycosidic bond to a purine or pyrimidine base (**Fig. 2**). The combination of the pentose sugar and the purine or pyrimidine base without the phosphate moiety is called a nucleoside. *See* PURINE; PYRIMIDINE.

Ribonucleosides contain the sugar D-ribose, whereas deoxyribonucleosides contain the sugar 2-deoxyribose. The four most common ribonucleosides are adenosine, guanosine, cytidine, and uridine. The purine ribonucleosides, adenosine and guanosine, contain the nitrogenous bases adenine and guanine, respectively. The pyrimidine ribonucleosides, cytidine and uridine, contain the bases cytosine and uracil, respectively. Similarly, the most common deoxyribonucleosides include

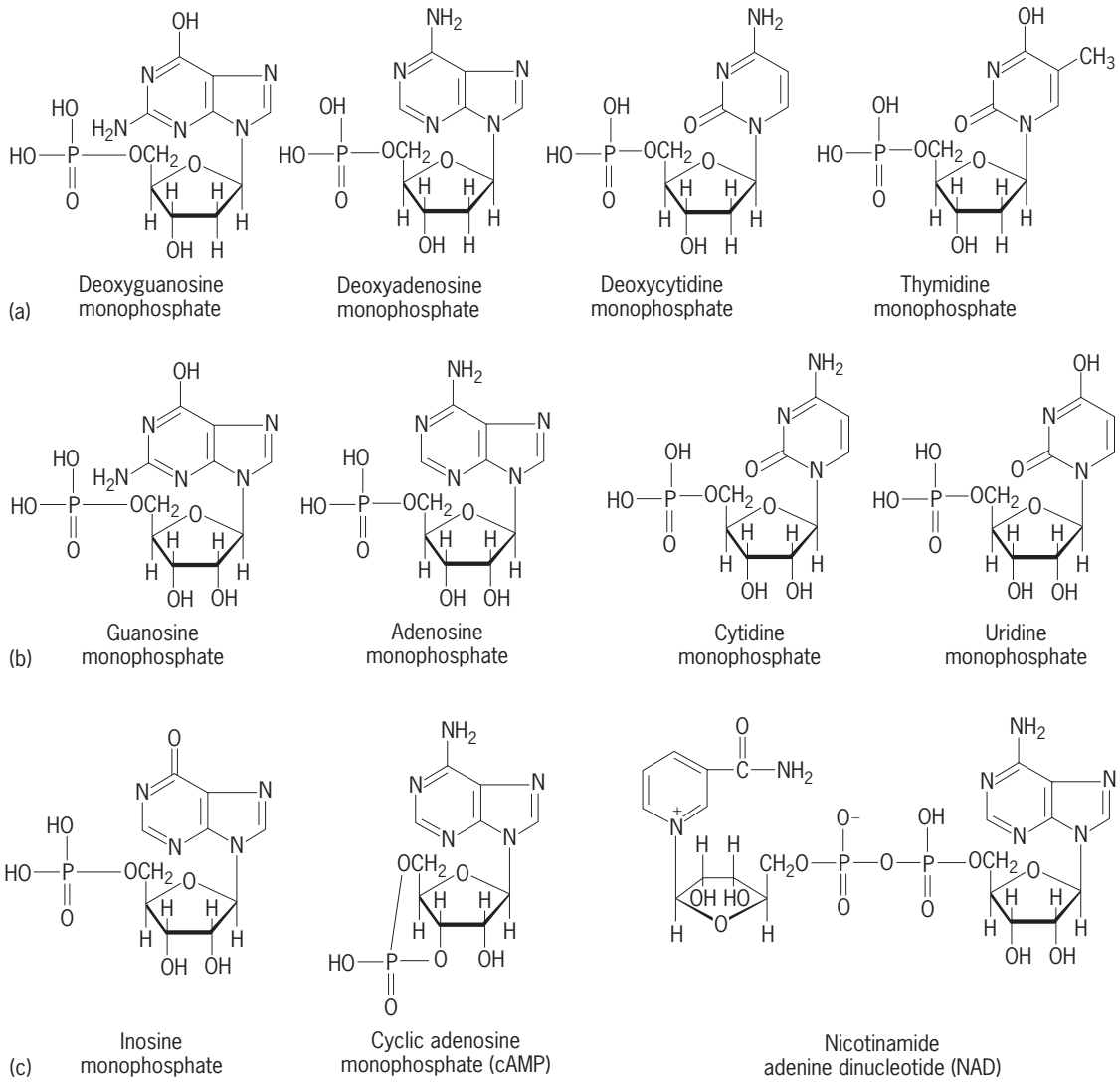

Fig. 1. Molecular structures of some common nucleotides. (*a*) Deoxyribonucleotides. (*b*) Ribonucleotides. (*c*) Other important nucleotides.

deoxyadenosine, deoxyguanosine, deoxycytidine, and thymidine, which contains the pyrimidine base thymine. Phosphorylation of the ribonucleosides or deoxyribonucleosides yields the corresponding ribonucleotide or deoxyribonucleotide.

The phosphate moiety of the nucleotide may be attached to any of the hydroxyl groups (5′, 3′, or 2′) that are part of the pentose sugar (Fig. 2*a*). For example, phosphorylation of the 5′ hydroxyl of adenosine yields the nucleotide adenosine 5′-monophosphate (5′-AMP); phosphorylation of thymidine at the 3′ position yields thymidine 3′-monophosphate (3′-TMP). Some nucleotides contain one or two additional phosphate molecules that are linked together by pyrophosphate bonds. The phosphate moieties are added to the appropriate nucleotide precursors by enzymes known as kinases. The resulting nucleotide diphosphates and triphosphates play important roles in the synthesis of polynucleotides and in providing the energy for important biochemical reactions.

Synthesis. The biological synthesis of nucleotides involves a long series of biochemical reactions. These reactions are carefully controlled by complex feedback mechanisms that regulate the quantities of nucleotides produced. Purine ribonucleotides are synthesized by sequentially building components of the purine ring structure into ribose 5′-phosphate. The initial purine ribonucleotide, known as 5′-inosine monophosphate (5′-IMP), undergoes subsequent chemical modification in order to make the more common ribo-nucleotides 5′-adenosine monophosphate (5′-AMP) and 5′-guanosine monophosphate (5′-GMP). The synthesis of pyrimidine nucleotides involves the direct attachment of a completed pyrimidine molecule, orotic acid, onto 5′-phosphoribosyl-1-pyrophosphate (PRPP). Enzymatic decarboxylation of the resulting pyrimidine ribonucleotide yields 5′-uridine monophosphate (5′-UMP), and amination of 5′-UMP yields the other pyrimidine ribonucleotide, 5′-cytidine monophosphate (5′-CMP). Deoxyribonucleotides are produced by direct

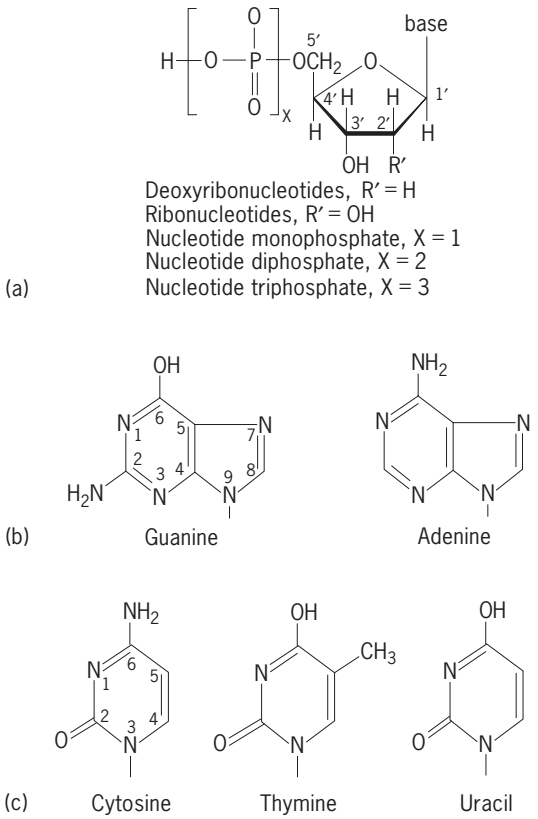

Fig. 2. Chemical structure of (*a*) nucleotides, (*b*) purine bases, and (*c*) pyrimidine bases.

reduction of the appropriate ribonucleotide precursors using the enzyme ribonucleotide diphosphate reductase.

Nucleotides can also be obtained by the chemical or enzymatic degradation of polynucleotides. Ribonucleic acid can be readily hydrolyzed to nucleotides under alkaline conditions. Deoxyribonucleic acid is somewhat resistant to this alkaline treatment, however, because of the absence of a hydroxyl group at the 2′ position. Enzymatic digestion of polynucleotides can also be used to make both nucleotide 3′ and 5′ phosphates. Phosphodiesterase, an enzyme in snake venom, degrades polynucleotides from the 3′ end to yield nucleoside 5′ phosphates. Conversely, phosphodiesterase from the spleen degrades polynucleotides from the 5′ end to yield nucleoside 3′ phosphates. Other enzymes, which are called restriction enzymes, split the polynucleotides into a series of smaller fragments by hydrolysis of the internucleotide bonds at sequence-specific positions. *See* RESTRICTION ENZYME.

Most nucleic acids contain nucleotides in which the 3′ hydroxyl group of the first nucleotide is linked to the 5′ phosphate of the second nucleotide. The polymerization of the nucleotide units is most often accomplished by the use of enzymes known as polymerases and transferases. These enzymes utilize nucleotide triphosphates in order to make long-chain polynucleotides. Some of these enzymes, such as Klenow polymerase, use a single-stranded template nucleic acid in order to control the sequence of nucleotide addition. Other enzymes such as terminal deoxynucleotidyl transferase randomly add nucleotides to the 3′ end of a growing polynucleotide without the need for a template strand. *See* DEOXYRIBONUCLEIC ACID (DNA); ENZYME; NUCLEIC ACID; RESTRICTION ENZYME; RIBONUCLEIC ACID (RNA).

E. Patrick Groody

Bibliography. A. Lesk (ed.), *Computational Molecular Biology: Sources and Methods for Sequence Analysis*, 1989; J. W. Phillis, *Adenosine and Adenine Nucleotides as Regulators of Cellular Function*, 1991; L. Stryer, *Biochemistry*, 4th ed., 1995; L. B. Townsend (ed.), *Chemistry of Nucleosides and Nucleotides*, 1994.

Nuclide

A species of atom that is characterized by the constitution of its nucleus, in particular by its atomic number Z and its neutron number $A-Z$, where A is the mass number. Whereas the terms isotope, isotone, and isobar refer to families of atomic species possessing common atomic number, neutron number, and mass number, respectively, the term nuclide refers to a particular atomic species. The total number of stable nuclides is approximately 275. About a dozen radioactive nuclides are found in nature, and in addition, hundreds of others have been created artificially. *See* NUCLEOSYNTHESIS.

Henry E. Duckworth

Nudibranchia

An order of the gastropod subclass Opisthobranchia containing about 2500 living species of carnivorous sea slugs. They occur in all the oceans, at all depths, but reach their greatest size and diversity in warm shallow seas. In this, the largest order of the Opisthobranchia, there is much evidence indicating polyphyletic descent from a number of long-extinct opisthobranch stocks. In all those lines which persist to the present day, the shell and operculum have been discarded in the adult form. In many of them the body has quite independently become dorsally papillate. In at least two suborders these dorsal papillae have acquired the power to nurture nematocysts derived from their coelenterate prey so as to use them for the nudibranch's own defense. In other cases such papillae (usually called cerata) have independently become penetrated by lobules of the adult digestive gland, or they may contain virulent defensive glands or prickly bundles of daggerlike calcareous spicules. So pervasive are the tendrils of the converging lines of evolution that the precise interrelationships between the dendronotacean, arminacean, aeolidacean, and doridacean suborders may never be fully understood.

Nudibranchs feed on every kind of epifaunal animal material, as well as playing a significant part in some planktonic communities. Some nudibranchs

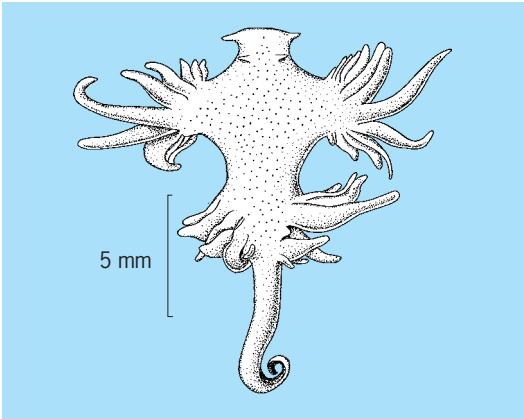

***Glaucus*, a specialized nudibranch known to attack planktonic coelenterates.**

(*Cerberilla*) burrow into soft sediments in search of their actinian (sea anemone) prey. Each nudibranch family contains species which feed on broadly similar types of prey. For instance, members of the Tritoniidae all feed upon alcyonarian (soft) corals. The Antiopellidae and Onchidorididae contain species which attack bryozoans. The Coryphellidae, Dendronotidae, Eubranchidae, Facelinidae, Heroidae, and Lomanotidae contain species that feed principally upon Hydrozoa. An unusual feeding type has evolved independently in members of several families; this is the egg-consumer, for example, the aeolids *Favorinus* and *Calma*. The former takes the eggs of other gastropod mollusks, but the latter feeds on the eggs of cephalopods and teleost fishes. Another highly successful specialization has evolved independently in the Fionidae (*Fiona*) and the Glaucidae (*Glaucus*, see **illus.**; *Glaucilla*); this involves attacking the planktonic coelenterates *Velella, Porpita*, and *Physalia* (the Portuguese man-o'-war). *Fiona* is said to attack also stalked barnacles. Glaucid and fionid nudibranchs are circumtropical in distribution.

There has been much argument about the significance of coloration in nudibranchs. The cerata of the aeolids and the dorsal papillae of many dorids or dendronotaceans are often patterned in such flamboyant ways as to foster the theory that these serve to take the attention of a sighted predator away from the fragile head and visceral mass. In many other cryptically marked nudibranchs, the resemblance in color pattern to the normal surroundings is so perfect that it would appear to disguise these mollusks. But against this it can be argued that many nudibranchs live in situations where there is little light (for instance, in deep waters or under boulders) and their color patterns cannot therefore be perceived by potential predators. The truth is that there is very little accurate information about the method of selection of food by predatory fish in different marine habitats. *See* OPISTHOBRANCHIA. T. E. Thompson

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; T. E. Thompson, *Nudibranchs*, 1976.

Number theory

The study of the properties and relationships between integers and other special types of numbers. There are many sets of positive integers of particular interest, such as the primes and the perfect numbers. Number theory, of ancient and continuing interest for its intrinsic beauty, also plays a crucial role in computer science, particularly in the area of cryptography.

Elementary number theory. This part of number theory does not rely on advanced mathematics, such as complex analysis and ring theory. The basic notion of elementary number theory is divisibility. An integer d is a divisor of n, written $d \mid n$, if there is an integer t such that $n = dt$. A prime number is a positive integer that has exactly two positive divisors, 1 and itself. The ten smallest primes are 2, 3, 5, 7, 11, 13, 17, 19, 23, and 29. Euclid (around 300 BC) proved that there are infinitely many primes by showing that if the only primes were 2, 3, 5, ... p, a prime not in this list could be found by taking a prime factor of the number shown in Eq. (1).

$$N = (2 \cdot 3 \cdot 5 \cdots p) + 1 \qquad (1)$$

Primes are the building blocks of the positive integers. The fundamental theorem of arithmetic, established by K. F. Gauss in 1801, states that every positive integer can be written as the product of prime factors in exactly one way when the order of the primes is disregarded. The fundamental ingredient in the proof is a lemma proved by Euclid: if a is a divisor of bc and a and b have no common factors, then a divides c.

The greatest common divisor of the positive integers a and b, written $\gcd(a, b)$, is the largest integer that divides both a and b. Two integers with greatest common divisor equal to 1 are called coprime. Euclid devised a method, now known as the euclidean algorithm, for finding the greatest common divisor of two positive integers. This method works by replacing a and b, where $a > b$, by b and the remainder when a is divided by b. This step is repeated until a remainder of 0 is reached. *See* ALGORITHM.

Perfect numbers and Mersenne primes. A perfect number is a positive integer equal to the sum of its positive divisors other than itself. L. Euler showed that $2^{n-1}(2^n - 1)$ is perfect if and only if $2^n - 1$ is prime. If $2^n - 1$ is prime, then n itself must be prime. Primes of the form $2^p - 1$ are known as Mersenne primes after M. Mersenne, who studied them in the seventeenth century. As of June 2005, 42 Mersenne primes were known, the largest being $2^{25,964,951} - 1$, a number with 7,816,230 decimal digits. For most of the time since 1876, the largest known prime has been a Mersenne prime; this is the case because there is a special method, known as the Lucas-Lehmer test, to determine whether $2^p - 1$ is prime. Only 12 Mersenne primes were known prior to the invention of computers. Since the early 1950s, with the help of computers, Mersenne primes have been found at a

rate averaging approximately one every 2 years. The eight largest Mersenne primes were found as part of the Great Internet Mersenne Prime Search (GIMPS). GIMPS provides optimized software for running the Lucas-Lehmer test. Thousands of people participate in GIMPS, running a total of almost 1 teraflops in the search for Mersenne primes.

All even perfect numbers are given by Euler's formula. Whether there are odd perfect numbers is still an unsolved problem. If an odd perfect number existed, it would have at least eight distinct prime factors and be larger than 10^{300}.

Congruences. If $a - b$ is divisible by m, then a is called congruent to b modulo m, and this relation is written $a \equiv b \pmod{m}$. This relation between integers is an equivalence relation and defines equivalence classes of numbers congruent to each other, called residue classes. Congruences to the same modulus can be added, subtracted, and multiplied in the same manner as equations. However, when both sides of a congruence are divided by the same integer d, the modulus m must be divided by $\gcd(d, m)$. There are m residue classes modulo m. The number of classes containing only numbers coprime to m is denoted by $\phi(m)$, where $\phi(m)$ is called the Euler phi function. The value of $\phi(m)$ is given by Eq. (2).

$$\phi(m) = m \prod_{p|m} \left(1 - \frac{1}{p}\right) \tag{2}$$

Fermat's little theorem states that if p is prime and a is coprime to p, then formula (3)

$$a^{p-1} \equiv 1 \pmod{p} \tag{3}$$

is true. Euler generalized this congruence by showing that formula (4)

$$a^{\phi(m)} \equiv 1 \pmod{m} \tag{4}$$

is valid whenever a and m are coprime.

The linear congruence $ax \equiv b \pmod{m}$ is solvable for x if and only if $d \mid b$ where $d = \gcd(a, m)$. Under this condition, it has exactly d incongruent solutions modulo m. If p is prime, then the congruence $ax \equiv 1 \pmod{p}$ has exactly one solution for each a not divisible by p. This solution is the inverse of a modulo p. This implies that the residue classes modulo p form a finite field of p elements. *See* FIELD THEORY (MATHEMATICS).

The simultaneous system of congruences $x \equiv a_i \pmod{m_i}$, $i = 1, 2, \ldots, r$, where the moduli m_i, $i = 1, 2, 3, \ldots, r$, are pairwise coprime positive integers, has a unique solution modulo M, where M is the product of these moduli. This result, called the Chinese remainder theorem, was known by ancient Chinese and Hindu mathematicians.

Primality testing and factorization. Since every composite integer has a prime factor not exceeding its square root, to determine whether n is prime, it is necessary only to show that no prime between 2 and n divides n. However, this simple test is extremely inefficient when $\sqrt{n}$ is large. Since generating large primes is important for cryptographic applications, better methods are needed.

By Fermat's little theorem, if 2^n is not congruent to 2 modulo n, then n is not prime; this was known to the ancient Chinese, who may have thought that the converse was true. However, it is possible that $2^n \equiv 2 \pmod{n}$ without n being prime; such integers n are called pseudoprimes to the base 2. More generally, an integer n is called a pseudoprime to the base b if n is not prime but $b^n \equiv b \pmod{n}$. A Carmichael number is a number n that is not prime but that is a pseudoprime to all bases b, where b is coprime to n and less than n. The smallest Carmichael number is 561. In 1992 it was shown that there are infinitely many Carmichael numbers.

A positive integer $n = 2^s t$, where t is odd, is called a strong pseudoprime to the base b if either $b^t \equiv 1 \pmod{n}$ or $b^{2^j t} \equiv -1 \pmod{n}$ for some integer j with $0 \le j \le s - 1$. The fact that n is a strong pseudoprime to the base b for at most $(n - 1)/4$ bases with $1 \le b \le n - 1$ when n is a composite integer is the basis for a probabilistic primality test which can be used to find extremely large numbers almost certainly prime in just seconds: Pick at random k different positive integers less than n. If n is composite, the probability that n is a strong pseudoprime to all k bases is less than $(1/4)^k$.

The naive way to factor an integer n is to divide n by successive primes not exceeding $\sqrt{n}$. If a prime factor p is found, this process is repeated for n/p, and so on. Factoring large integers n in this way is totally infeasible, requiring prohibitively large times. This shortcoming has led to the development of improved factorization techniques. Techniques including the quadratic sieve, the number field, and the elliptic curve factoring methods have been developed, making it possible to factor numbers with as many as 150 digits, although a large number of computers may be required. However, the factorization of integers with 200 digits is considered far beyond current capabilities, with the best-known methods requiring an astronomically larger time.

Interest in primality testing and factorization has increased since the mid-1970s because of their importance in cryptography. The Rivest-Shamir-Adleman (RSA) public-key encryption system uses the product of two large primes, say with 100 digits each, as its key. The security of the system rests on the difficulty of factoring the product of these primes. The product of the two primes held by each person is made public, but the separate primes are kept secret. *See* CRYPTOGRAPHY.

Although probabilistic primality tests can be used to generate primes rapidly, no practical test has been found that determines with certainty whether a particular positive integer is prime. An important and surprising theoretical breakthrough in this direction was made in 2002 by M. Agrawal, N. Kayal, and N. Saxena. They were able to find a polynomial-time algorithm that proves that a positive integer is prime, if it indeed is. So far, the best algorithms developed using their ideas are not of practical value in cryptographic applications.

Sieves. In the third century B.C., Eratosthenes showed how all primes up to an integer n can be found when only the primes p up to $\sqrt{n}$ are known. It is sufficient to delete from the list of integers, starting with 2, the multiples of all primes up to $\sqrt{n}$. The remaining integers are all the primes not exceeding n.

Many problems, including the relative density of the set of twin primes (primes that differ by 2 such as 17 and 19), or the Goldbach conjecture that every even integer greater than 2 is the sum of two primes, have been attacked by sieve methods with partial success. Nevertheless, in spite of improvements of the method, it is not known whether the number of twin primes is finite or infinite or whether Goldbach's conjecture is true or false. (Goldbach's conjecture has been verified by computers up to 2×10^{17}.) In 1966 J. Chen showed that all even integers other than 2 are sums of a prime and another integer that is either prime or a product of only two primes. (The evidence that there are infinitely many twin primes continues to mount. As of 2005, the largest known twin primes were $33{,}218{,}925 \times 2^{169{,}690} \pm 1$.)

Primitive roots and discrete logarithms. The order of a modulo m, where a and m are coprime, is the least positive integer x such that formula (5)

$$a^x \equiv 1 \pmod{m} \tag{5}$$

is satisfied. Euler's theorem shows such an integer exists since $a^{\phi(m)} \equiv 1 \pmod{m}$. An integer a is called primitive root modulo m if the order of a modulo m equals $\phi(m)$, the largest possible order. The integer m has a primitive root if and only if $m = 2, 4, p^k$, or $2p^k$ where p is an odd prime.

A solution of the congruence $b^x \equiv c \pmod{n}$ for x is known as a discrete logarithm to the base b of c modulo n. The security of many cryptographic systems is based on the difficulty of finding discrete logarithms. The computational complexity of finding discrete logarithms modulo a prime p is similar to that of factoring an integer n of similar size to p. Several important methods for finding discrete logarithms modulo a prime are based on sieve techniques. The problem of finding discrete logarithms modulo p is much easier when $p - 1$ has only small prime factors.

Quadratic resides and reciprocity. Quadratic residues modulo m are integers that are perfect squares modulo m. More precisely, a is a quadratic residue modulo m if the congruence shown in formula (6)

$$x^2 \equiv a \pmod{m} \tag{6}$$

has a solution for x, where a and m are coprime. If p is an odd prime, the Legendre symbol (a/p) is defined to be $+1$ if a is a quadratic residue modulo p, and to be -1 if a is a quadratic nonresidue modulo p.

The law of quadratic reciprocity, first proved by Gauss, states that if r and q are odd primes, then Eq. (7)

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{[(p-1)/2][(q-1)/2]} \tag{7}$$

is valid. Gauss found eight different proofs of the law quadratic reciprocity himself, and new proofs are found on a regular basis with more than 200 different proofs known. Moreover, Eq. (8)

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2} \tag{8}$$

is true. The Legendre symbol can be generalized to odd composite moduli (Jacobi symbol) and to even integers (Kronecker symbol); reciprocity laws hold for these more general symbols.

Quadratic forms. The expression $F(x, y) = Ax^2 + Bxy + Cy^2$, where A, B, and C are integers and x and y are variables that take on integer values, is called a binary quadratic form with discriminant $B^2 - 4AC$. Two forms $F(x, y) = Ax^2 + Bxy + Cy^2$ and $F_1(x', y') = A_1x'^2 + B_1x'y' + C_1y'^2$ are equivalent if there are integers a, b, c, and d such that $ad - bc = \pm 1$, $x' = ax + by$, and $y' = cx + dy$. Equivalent binary quadratic forms have the same discriminant and represent the same integers. The number of classes of equivalent forms with a given discriminant is finite. There are only finitely many negative discriminants with a given class number, with exactly nine with class number one. In the theory of binary quadratic forms with positive discriminant D, Pell's equation, $x^2 - Dy^2 = \pm 1$, plays a fundamental role.

Sums of squares and similar representations. Pierre de Fermat observed, and Euler proved, that every prime number congruent to 1 modulo 4 is the sum of two squares. J. L. Lagrange showed that every positive integer is the sum of at most four squares, and for some integers four squares are indeed needed. E. Waring conjectured in 1782 that to every positive integer k there is a number $g(k)$ such that every natural number is the sum of at most $g(k)$ kth powers. This was first proved in 1909 by David Hilbert. Further work has been devoted to finding $g(k)$, the least integer such that every positive integer is the sum of at most $g(k)$ kth powers and $G(k)$, the least integer such that every sufficiently large integer is the sum of at most $G(k)$ kth powers. It is known that $G(2) = 4$ and $G(4) = 16$, but the value of $G(3)$ is unknown. Although there is strong numerical evidence that $G(3) = 4$, all that is currently known is that $4 \leq G(3) \leq 7$. New and improved bounds for $G(k)$ for small integers k are established with some regularity such as the bounds $G(19) \leq 134$ and $G(20) \leq 142$ established by R. C. Vaughn and T. D. Wooley in 1999.

Diophantine equations. Single equations or systems of equations in more unknowns than equations, with restrictions on solutions such as that they must all be integral, are called diophantine equations, after Diophantus who studied such equations in ancient times. A wide range of diophantine equations have been studied. For example, the diophantine equation (9)

$$x^2 + y^2 = z^2 \tag{9}$$

has infinitely many solutions in integers. These solutions are known as pythagorean triples, since they correspond to the lengths of the sides of right

triangles where these sides have integral lengths. All solutions of this equation are given by $x = t(u^2 - v^2)$, $y = 2tuv$, and $z(u^2 + v^2)$, where t, u, and v are positive integers.

Perhaps the most notorious diophantine equation is Eq. (10).

$$x^n + y^n = z^n \tag{10}$$

Fermat's last theorem states that this equation has no solutions in integers when n is an integer greater than 2 where $xyz \neq 0$. Establishing Fermat's last theorem was the quest of many mathematicians over 200 years. In the 1980s, connections were made between the solutions of this equations and points on certain elliptic curves. Using the theory of elliptic curves, A. Wiles completed a proof of Fermat's last theorem based on these connections in 1995.

The discovery of the proof of Fermat's last theorem has led to the formulation of a more powerful conjecture by Andrew Beal, an amateur mathematician. Beal's conjecture asserts that the Diophantine equation (11)

$$x^a + y^b = z^c \tag{11}$$

has no solution in positive integers x, y, z, a, b, and c, where $a \geq 3$, $b \geq 3$, $c \geq 3$, and the integers x, y, and z are pairwise coprime.

Algebraic number theory. Attempts to prove Fermat's last theorem led to the development of algebraic number theory, a part of number theory based on techniques from such areas as group theory, ring theory, and field theory. Gauss extended the concepts of number theory to the ring $R[i]$ of complex numbers of the form $a + bi$, where a and b are integers. Ordinary primes $p \equiv 3 \pmod 4$ are also prime in $R[i]$, but $2 = -i(1 + i)^2$ is not prime, nor are primes $p \equiv 1 \pmod 4$ since such primes split as $p = (a + bi)(a - bi)$. More generally, an algebraic number field $R(\theta)$ of degree n is generated by the root θ of a polynomial equation $f(x) = 0$ of degree n with rational coefficients. A number α in this field is called an algebraic integer if it satisfies an algebraic equation with integer coefficients with initial coefficient 1. The algebraic integers in an algebraic number field form an integral domain. But, prime factorization may not be unique; for example, in $R[\sqrt{-5}]$, $21 = 3 \cdot 7 = (1 + 2\sqrt{-5}) \cdot (1 - 2\sqrt{-5})$ where each of the four factors in the two products is prime. To restore unique factorization, the concept of ideals is needed, as shown by E. E. Kummer and J. W. R. Dedekind. *See* RING THEORY.

Analytic number theory. There are many important results in number theory that can be established by using methods from analysis. For example, analytic methods developed by G. F. B. Riemann in 1859 were used by J. Hadamard and C. J. de la Vallée Poussin in 1896 to prove the famous prime number theorem. This theorem, first conjectured by Gauss about 1793, states that $\pi(x)$, the number of primes not exceeding x, behaves as shown in Eq. (12).

$$\lim_{x \to \infty} \frac{\pi(x)}{(x/\log x)} = 1 \tag{12}$$

These methods of Riemann are based on $\zeta(s)$, the function defined by Eq. (13),

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_p \frac{1}{1 - p^{-s}} \tag{13}$$

where $s = \sigma + it$ is a complex variable; the series in this equation is convergent for $\sigma > 1$. Via an analytic continuation, this function can be defined in the whole complex plane. It is a meromorphic function with only a simple pole of residue 1 at $s = 1$. However, the fundamental theorem of arithmetic can be used to show that this series equals the product over all primes p shown in Eq. (13). It can be shown that $\zeta(s)$ has no zeros for $\sigma = 1$; this result and the existence of a pole at $s = 1$ suffice to prove the prime number theorem. Many additional statements about $\pi(x)$ have been proved. Riemann's work contains the still unproved so-called Riemann hypothesis: all zeros of $\zeta(s)$ have a real part not exceeding $^1/_2$. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

Another important result that can be proved by analytic methods is Dirichlet's theorem, which states that there are infinitely many prime numbers in every arithmetic progression $am + b$, were a and b are coprime positive integers.

Diophantine approximation. A real number x is called rational if there are integers p and q such that $x = p/q$; otherwise x is called irrational. The number $b^{1/m}$ is irrational if b is an integer which is not the mth power of an integer (for example, $\sqrt{2}$ is irrational). A real number x is called algebraic if it is the root of a monic polynomial with integer coefficients; otherwise x is called transcendental. The numbers e and π are transcendental. That π is transcendental implies that it is impossible to square the circle. *See* CIRCLE; E (MATHEMATICS).

The part of number theory called diophantine approximation is devoted to approximating numbers of a particular kind by numbers from a particular set, such as approximating irrational numbers by rational numbers with small denominators. A basic result is that, given an irrational number x, there exist infinitely many fractions h/k that satisfy the inequality (14),

$$\left| x - \frac{h}{k} \right| < \frac{1}{ck^2} \tag{14}$$

where c is any positive number not exceeding $\sqrt{5}$. However, when c is greater than $\sqrt{5}$, there are irrational numbers x for which there are only finitely many such h/k. The exponent 2 in inequality (14) cannot be increased, since when x is algebraic and not rational the Thue-Siegel-Roth theorem implies that the inequality (15),

$$\left| x - \frac{h}{k} \right| < \frac{1}{ck^{2+\epsilon}} \tag{15}$$

where ϵ is any number greater than zero (however small), can have at most only finitely many solutions h/k.

In 1851, J. Liouville showed that transcendental numbers exist; he did so by demonstrating that the

number x given by Eq. (16)

$$x = \sum_{j=1}^{\infty} 10^{-j!} \tag{16}$$

has the property that, given any positive real number m, there is a rational number h/k that satisfies Eq. (17).

$$\left| x - \frac{h}{k} \right| < \frac{1}{k^m} \tag{17}$$

See ALGEBRA; NUMBERING SYSTEMS; ZERO.

Additive number theory. Problems in additive number theory can be studied using power series. Suppose that $m_1, m_2, \ldots, m_k, \ldots$, is a strictly increasing sequence of positive integers, such as the sequence of perfect squares, the sequence of prime numbers, and so on. If the power series of Eq. (18) is raised to the qth power to yield Eq. (19),

$$f(x) = \sum_{k=1}^{\infty} x^{m_k} \tag{18}$$

$$f(x)^q = \sum_{j=1}^{\infty} A_j x^j \tag{19}$$

then the coefficients A_j represent the number of times that j can be written as the sum of q elements from the set $\{m_j\}$. If it can be shown that A_k is positive for all positive integers k, then it has been shown that every positive integer k is the sum of q numbers from this set. Function-theoretic methods can be applied, as they have been by G. H. Hardy and J. E. Littlewood and by I. M. Vinogradov, to determine the coefficients A_k. For example, if m_k is the kth power of a positive integer and q is sufficiently large, Waring's theorem can be proved. Similarly, if m_k is the kth odd prime, this technique has been used to show that every sufficiently large odd number is the sum of at most three primes.

A partition of the positive integer n is a decomposition of n into a sum of positive integers, with order disregarded. Euler showed that the generating function for the number of partitions $p(n)$ is given by Eq. (20).

$$\sum_{n=0}^{\infty} p(n)x^n = \prod_{m=1}^{\infty} (1 - x^m)^{-1} \tag{20}$$

A recurrence formula for $p(n)$ can be derived from this formula, and in 1917 Hardy and S. Ramanujan derived an asymptotic formula for $p(n)$ from it using function-theoretic methods. The first term is formula (21).

$$p(n) \approx \frac{\exp\left(\pi\sqrt{2n/3}\right)}{4n\sqrt{3}} \tag{21}$$

Ramanujan discovered that the function $p(n)$ satisfies many congruences, such as Eqs. (22) and (23).

$$p(5n + 4) \equiv 0 \pmod 5 \tag{22}$$

$$p(7n + 5) \equiv 0 \pmod 7 \tag{23}$$

Euler showed that the number of partitions into odd parts equals the number of partitions into distinct parts. For instance, 7 has five partitions into odd parts (7, 1 + 1 + 5, 1 + 3 + 3, 1 + 1 + 1 + 1 + 3, and 1 + 1 + 1 + 1 + 1 + 1 + 1) and five partitions into distinct parts (7, 1 + 6, 2 + 5, 3 + 4, 1 + 2 + 4). Many additional results have been proved about partitions. *See* COMBINATORIAL THEORY.

Kenneth H. Rosen

Bibliography. T. M. Apostol, *Introduction to Analytic Number Theory*, Springer, 1986; J. R. Goldman, *The Queen of Mathematics: An Historically Motivated Guide to Number Theory*, A K Peters, 1997; Grosswald, *Topics from the Theory of Numbers*, 2d ed., Birkhauser, 1982; G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 5th ed., Oxford, 1979; R. A. Mollin, *Algebraic Number Theory*, CRC Press, 1999; I. Niven, H. S. Zuckerman, and H. L. Montgomery, *An Introduction to the Theory of Numbers*, 5th ed., Wiley, 1991; K. H. Rosen, *Elementary Number Theory and Its Applications*, 5th ed., Addison-Wesley, 2005; W. Sierpinski and A. Schinzel, *Elementary Theory of Numbers*, 2d ed., North-Holland, 1988; J. Steuding, *Diophantine Analysis*, CRC Press/Chapman & Hall, 2005; J. Stopple, *A Primer of Analytic Number Theory*, Cambridge University Press, 2003.

Numbering systems

A numbering system is a systematic method for representing numbers using a particular set of symbols. The most commonly used numbering system is the decimal system, based on the number 10, which is called the basis or radix of the system. The basis tells how many different individual symbols there are in the system to represent numbers. In the decimal system these symbols are the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. The range of these numbers varies from 0 to (10 − 1). This is a particular case of a more general rule: Given any positive basis or radix N, there are N different individual symbols that can be used to write numbers in that system. The range of these numbers varies from 0 to $N - 1$.

In the computer and telecommunication fields, three of the most frequently used numbering systems are the binary (base 2), the octal (base 8), and the hexadecimal (base 16). The binary system has only two symbols: 0 and 1. Either of these symbols can be called a binary digit or a bit. The octal system has eight symbols: 0, 1, 2, 3, 4, 5, 6, 7. The hexadecimal system has 16 symbols: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F. In the hexadecimal system A stands for 10, B for 11, C for 12, D for 13, E for 14, and F for 15. The reason for choosing single letters to represent numbers higher than 9 is to keep all individual symbols single characters. *See* BIT.

Although "digits" generally refers to the individual symbols of the decimal system, it is common to call the individual symbols of the other numbering systems by that name. To explicitly state the basis on which a number is written, a subscript will be used to the lower right of the number. In this way $11001_{(2}$ is identified as a binary number, and $113_{(16}$ is identified

as a hexadecimal number. Subscripts are not used to represent decimal numbers.

All the numbering systems mentioned so far are positional systems. That is, the value of any symbol depends on its position in the number. For example, the value of 2 in the decimal number 132 is that of two units, whereas its value in decimal 245 is that of two hundreds. In the decimal system, the rightmost position of a number is called the ones (10^0) place, the next position from the right is called the tens (10^1) place, the next position the hundreds (10^2) place, and so on. Observe that the powers increase from the right. The power of the rightmost digit is zero, the power of the next digit is one, the power of the next digit is two, and so on. These powers are sometimes called the weight of the digit.

Conversion between bases. In any positional system, the decimal equivalent of a digit in the representation of the number is the digit's own value, in decimal, multiplied by a power of the basis in which the number is represented. The sum of all these powers is the decimal equivalent of the number. The corresponding powers of each of the digits can be better visualized by writing superscripts beginning with 0 at the rightmost digit, and increasing the powers by 1 as we move toward the left digits of the number. For example, the decimal equivalent of the hexadecimal number 1A2C can be calculated as follows: (1) number the digits from right to left using superscripts; (2) use these superscripts as the powers of the basis and apply the general rule indicated before. After step 1 the number looks like this: $1^3A^22^1C^0$. According to step 2, the value of each digit and the decimal equivalent of the number are:

$$\begin{aligned} 1A2C_{(16} &= (1*16^3) + (10*16^2) + (2*16^1) + (12*16^0) \\ &= (1*4096) + (10*256) + (2*16) + (12*1) \\ &= (4096) + (2560) + (32) + (12) \\ &= 6700 \end{aligned}$$

In this example, A and C have been replaced by their decimal equivalents of 10 and 12 respectively. This result indicates that, in the number 1A2C, the digit 1 has a value of 4096 units, the A has a value of 2560 units, the 2 has a value of 32, and the C has a value of 12 units. The decimal equivalent of the entire number is 6700. A similar procedure can be used to find the equivalents of $175_{(8}$ and $11001_{(2}$.

$$\begin{aligned} 175_{(8} &= (1*8^2) + (7*8^1) + (5*8^0) = 125 \\ 11001_{(2} &= (1*2^4) + (1*2^3) + (0*2^2) + (0*2^1) \\ &\quad + (1*2^0) = 25 \end{aligned}$$

Decimal numbers to other bases. The conversion of a given decimal number to another basis r ($r > 0$) is carried out by initially dividing the given decimal number by r, and then successively dividing the resulting quotients by r until a zero quotient is obtained. The decimal equivalent is obtained by writing the remainders of the successivedivisions in the opposite order in which they were obtained. For example, to convert decimal 41 to binary using this method, the number 41 is initially divided by the binary basis, that is, by $r = 2$, to obtain a first quotient and a remainder. This and all successive quotients are then divided by 2, until a quotient of zero is obtained. The remainders in the opposite order in which they were obtained give the representation of the equivalent binary number. This process is illustrated below.

Number	Quotient when dividing by 2	Remainder
41	20	1
20	10	0
10	5	0
5	2	1
2	1	0
1	0	1

Thus, the binary equivalent of decimal 41 is 101001. That is, $41_{(10} = 101001_{(2}$. This result can be verified by converting 101101 to its decimal equivalent according to the procedure described earlier. In fact, $101001 = (1*2^5) + (0*2^4) + (1*2^3) + (0*2^2) + (0*2^1) + (1*2^0) = 41$. Following a similar procedure and dividing by 16 and 8 respectively, it can be verified that $5460 = 1554_{(16}$ and $876 = 1554_{(8}$.

Binary to hexadecimal or octal and vice versa. The **table** shows the decimal numbers 1 through 15 written in binary, octal, and hexadecimal. Since each four-bit binary number corresponds to one and only one hexadecimal digit and vice versa, the hexadecimal system can be viewed as a shorthand notation of the binary system. Similar reasoning can be applied to the octal system. This one-to-one correspondence between the symbols of the binary system and the symbols of the octal and hexadecimal system provides a method for converting numbers between these bases.

To convert binary numbers to hexadecimal, the following procedure may be used: (1) Form four-bit groups beginning from the rightmost bit of the number. If the last group (at the leftmost

The first 15 integers in binary, octal, hexadecimal, and decimal notation

Binary	Octal	Hexadecimal	Decimal
0001	1	1	1
0010	2	2	2
0011	3	3	3
0100	4	4	4
0101	5	5	5
0110	6	6	6
0111	7	7	7
1000	10	8	8
1001	11	9	9
1010	13	A	10
1011	14	B	11
1100	14	C	12
1101	15	D	13
1110	16	E	14
1111	17	F	15

position) has fewer than four bits, add extra zeros to the left of the bits in this group to make it a four-bit group. (2) Replace each four-bit group by its hexadecimal equivalent. Using this procedure, it can be verified that the hexadecimal equivalent of binary 111110110110101111 is 3EDAF. A process that is almost the reverse of the previous procedure can be used to convert from hexadecimal to binary. However, there is no need to add extra zeros to any group since each hexadecimal number will always convert to a group with four binary bits.

A similar process can be followed to convert a binary number to octal, except that in this case three-bit groups must be formed. Individual octal numbers will always convert to groups with three binary bits.

Computer representation of integers. Numerical data in a computer are generally written in basic units of storage made up of a fixed number of consecutive bits. The most commonly used units in the computer and communication industries are the byte (8 consecutive bits), the word (16 consecutive bits), and the double word (32 consecutive bits).

Nonnegative integers. A number is represented in each of these units by setting the bits according to the binary representation of the number. By convention, bits in any of these units are numbered from right to left, beginning with zero. The rightmost bit is called the least significant bit. The leftmost bit, numbered $(n - 1)$, where n is the number of bits in the unit, is called the most significant bit.

Since a bit can have only two values, given n bits, there are 2^n different numbers that can be represented. The range of these numbers varies from 0 to $2^n - 1$.

Negative integers. To represent negative numbers, one of the bits is chosen to represent the sign. By convention, the leftmost bit is considered the sign bit. A value of 0 in the sign bit indicates a positive number, whereas a value of 1 indicates a negative number. This convention applies to bytes, words, and double words.

In addition to this sign convention, computer manufacturers use another convention, called the two's complement notation, to represent negative numbers. Two's complement is a three-step process: (1) Obtain the positive binary representation of the number. (2) Complement all bits of the number. That is, change all 0's to 1's and vice versa. (3) Add 1 to the binary configuration of the number obtained in the previous step. For example, the two's complement of decimal −45 (using a word as the storage unit) can be obtained as follows: (1) Obtain the binary representation of +45. The procedure indicated earlier yields the result that $45 = 0000000000101101_{(2}$. (2) Complement all bits to obtain 11111111110010010. (3) Add 1 to the result of the previous step:

```
1111111111010010 +
                1
1111111111010011  ←  Two's complement
                     representation of −45
```

To find out the positive counterpart of a given negative two's complement number, follow steps 2 and 3 indicated before. In this case, given the two's complement representation of −45, the representation of 45 should be 0000000000101101. *See* NUMERICAL REPRESENTATION (COMPUTERS).

Arithmetic operations. Addition of binary numbers in a computer is carried out using a procedure similar to that of the addition of decimal numbers. The rules for binary addition are shown below. The sign bit is treated as any other bit.

+	0	1
0	0	1
1	1	0 with a carry of 1

The addition of the binary numbers 01010 and 00111 is shown below. Carries are indicated in italics.

```
                 111
Augend     00001010 +
Addend     00000111
           --------
Sum        00010001
```

In this example, when the two rightmost bits are added, the result is 1. When the next two binary 1's of the second column (from the right) are added, the result is zero with a carry of 1. This carry is then added to the zero of the third column of the augend. The result is 1. This 1 is then added to the 1 of the third column of the addend to produce another zero at the sum and another carry. The remaining bits of the sum are obtained in similar manner.

Subtracting binary numbers is very similar to the subtraction of decimal numbers; however, in borrowing from a higher unit, 2 is borrowed instead of 10. The similarity between the binary and decimal operations may be noted in the examples below. In the decimal system, the 0's to the left of the rightmost 0 became 9 (the basis minus 1). Likewise in the binary operation the 0's to the left of the rightmost 0 became 1 (the basis minus 1).

```
Decimal    Binary
 7000 −    10000 −
   15        111
 ----      -----
 6985      01001
```

To add two's complement numbers, the usual rules for adding binary numbers are followed and the carry out of the leftmost bit is ignored, if there is one. The binary number B can be subtracted from A simply by adding to A the two's complement of B. That is, A − B = A + (−B), where −B is the two's complement of B. These processes are illustrated below. The decimal equivalents of the two's complement

numbers are indicated in parentheses. Here, A = 00011011 (27) and B = 00000100 (4).

Column A		Column B	
00011011 +	(27)	00011011 +	(27)
00000100	(4)	11111100	(−4) ← two's complement notation
00011111	(31)	00010111	(23)

Overflows. Since all arithmetic operations are carried out using storage units of fixed size, there is the possibility that the result of an operation exceeds the capacity of the storage units. In this case, an overflow condition is said to have occurred. When the result is too small to be represented in the storage unit, an underflow is said to have occurred. Overflows will always occur if either of the following conditions is satisfied: (1) There is a carry into the sign bit and no carry out of it. (2) There is no carry into the sign bit and there is a carry out of the sign bit. In the addition of column B, for example, there is a carry into the most significant bit and a carry out of the most significant bit. Therefore, there is no overflow. Subtractions of two's complement numbers are carried out in the usual way; however, if a borrow is needed in the leftmost place, the borrow should be carried out as if there were another bit to the left. The following example illustrates this:

```
Minuend      00100011 −
Subtrahend   00111001
             --------
             11101010
```

Binary multiplication and division. Multiplication and division of binary numbers is remarkably simple since the user can carry out these two operations by treating the numbers as if they were decimal numbers. The examples below illustrate these operations, using a byte as the basic storage unit.

To multiply 6 (=0110_2) times 2 (=0010) we can proceed as follows:

```
Multiplicand     00110 ×
Multiplier       0010
                 ----------
                 0000 +
                0110
               0000
              0000
              ----------
              0001100(=12)
```

Hexadecimal and octal operations. Operating with long strings of 0's and 1's is an error-prone task. For this reason it is preferable to carry out arithmetic operations in the hexadecimal system. Adding hexadecimal numbers is a process similar to that of adding decimal and binary numbers. However, it is necessary to remember that there is a carry whenever the sum exceeds F. In the subtraction operations, "16" is borrowed from the higher units to the left. To simplify this process, it is convenient to think in decimal notation. Therefore, in our heads we may translate letters into their decimal equivalent, do the operations in decimal, and translate the results back to hexadecimal. The following example illustrates the similarities between the arithmetic operations of both systems.

Column A (hexadecimal)	Column B (decimal)
1	
EDF4 +	7955 +
263	3162
F057	11117

The addition of the numbers of the second column (from the right) in column A is carried out as follows: Since 15 + 6 = 21 and we can write 21 = 1∗16 + 5, then we write 5 and carry 1. Similarly since 16 = 1∗16 + 0, we write 0 and carry 1. A similar situation is encountered in the addition of the decimal numbers of the second column of column B.

To multiply or divide numbers in hexadecimal, it is preferable to transform the numbers to binary, do the operations in binary, and then translate back the results to hexadecimal.

Operations in octal follow a similar pattern for both addition and subtraction. However, there is always a carry whenever the result of an addition exceeds 7. When borrowing, we always borrow 8 from the higher units to the left. The following examples illustrate this. The decimal equivalents of the numbers are shown in parentheses.

1		
623 +	(403)	751−
126	(86)	126
751	(489)	623

Octal quantities can be multiplied or divided by transforming the numbers to binary, carrying out the operations in binary, and transforming the results back to octal. *See* DIGITAL COMPUTER.

Ramon A. Mata-Toledo

Bibliography. H. Kruglak et al., *Basic Mathematics with Applications to Science and Technology*, 2d ed., Schaum's Outline Series, McGraw-Hill, 1998; R. Mata-Toledo and P. Cushman, *Introduction to Computer Science with Examples in Visual Basic, C, C++, and Java*, Schaum's Outline Series, McGraw-Hill, 2000; C. Maxfield and A. Brown, *Bebop Bytes Back. An Unconventional Guide to Computers*, Doone Publications, 1997.

Numerical analysis

The development and analysis of computational methods (and ultimately of program packages) for the minimization and the approximation of functions, and for the approximate solution of equations, such as linear or nonlinear (systems of) equations and differential or integral equations. Originally part of every mathematician's work, the subject is now often taught in computer science departments because of the tremendous impact which computers have had on its development. Research focuses mainly on the numerical solution of (nonlinear) partial differential equations and the minimization of functions. *See* COMPUTER.

Numerical analysis is needed because answers provided by mathematical analysis are usually symbolic and not numeric; they are often given implicitly only, as the solution of some equation, or they are given by some limit process. A further complication is provided by the rounding error which usually contaminates every step in a calculation (because of the fixed finite number of digits carried).

Even in the absence of rounding error, few numerical answers can be obtained exactly. Among these are (1) the value of a piecewise rational function at a point and (2) the solution of a (solvable) linear system of equations, both of which can be produced in a finite number of arithmetic steps. Approximate answers to all other problems are obtained by solving the first few in a sequence of such finitely solvable problems. A typical example is provided by Newton's method: A solution c to a nonlinear equation $f(c) = 0$ is found as the limit $c = \lim_{n\to\infty} x_n$, with x_{n+1} a solution to the linear equation $f(x_n) + f'(x_n)(x_{n+1} - x_n) = 0$, that is, $x_{n+1} = x_n - f(x_n)/f'(x_n)$, $n = 0, 1, 2, \ldots$. Of course, only the first few terms in this sequence $x_0, x_1, x_2, \ldots$ can ever be calculated, and thus one must face the question of when to break off such a solution process and how to gauge the accuracy of the current approximation. The difficulty in the mathematical treatment of these questions is amplified by the fact that the limit of a sequence is completely independent of its first few terms.

In the presence of rounding error, an otherwise satisfactory computational process may become useless, because of the amplification of rounding errors. A computational process is called stable to the extent that its results are not spoiled by rounding errors. The extended calculations involving millions of arithmetic steps now possible on computers have made the stability of a computational process a prime consideration.

Interpolation and approximation. Polynomial interpolation provides a polynomial p of degree n or less that uniquely matches given function values $f(x_0), \ldots, f(x_n)$ at corresponding distinct points $x_0, \ldots x_n$. The interpolating polynomial p is used in place of f, for example in evaluation, integration, differentiation, and zero finding. Accuracy of the interpolating polynomial depends strongly on the placement of the interpolation points, and usually degrades drastically as one moves away from the interval containing these points (that is, in case of extrapolation). *See* EXTRAPOLATION.

When many interpolation points (more than 5 or 10) are to be used, it is often much more efficient to use instead a piecewise polynomial interpolant or spline. Suppose the interpolation points above are ordered, $x_0 < x_1 < \ldots < x_n$. Then the cubic spline interpolant to the above data, for example, consists of cubic polynomial pieces, with the ith piece defining the interpolant on the interval $[x_{i-1}, x_i]$ and so matched with its neighboring piece or pieces that the resulting function not only matches the given function values (hence is continuous) but also has a continuous first and second derivative.

Interpolation is but one way to determine an approximant. In full generality, approximation involves several choices: (1) a set P of possible approximants, (2) a criterion for selecting from P a particular approximant, and (3) a way to measure the approximation error, that is, the difference between the function f to be approximated and the approximant p, in order to judge the quality of approximation. Much studied examples for P are the polynomials of degree n or less, piecewise polynomials of a given degree with prescribed breakpoints, and rational functions of given numerator and denominator degrees. The distance between f and p is usually measured by a norm, such as the L_2 norm, $(\int |f(x) - p(x)|^2 dx)^{1/2}$ or the uniform norm $\sup_x |f(x) - p(x)|$. Once choices 1 and 3 are made, one often settles 2 by asking for a best approximation to f from P, that is, for an element of P whose distance from f is a small as possible. Questions of existence, uniqueness, characterization, and numerical construction of such best approximants have been studied extensively for various choices of P and the distance measure. If P is linear, that is, if P consists of all linear combinations

$$\sum_{i=1}^{n} a_i p_i$$

of certain fixed functions $p_1, \ldots, p_n$, then determination of a best approximation in the L_2 norm is particularly easy, since it involves nothing more than the solution of n simultaneous linear equations.

Solution of linear systems. Solving a linear system of equations is probably the most frequently confronted computational task. It is handled either by a direct method, that is, a method that obtains the exact answer in a finite number of steps, or by an iterative method, or by a judicious combination of both. Analysis of the effectiveness of possible methods has led to a workable basis for selecting the one that best fits a particular situation.

Direct methods. Cramer's rule is a well-known direct method for solving a system of n linear equations in n unknowns, but it is much less efficient than the method of choice, elimination. In this procedure the first unknown is eliminated from each equation but the first by subtracting from that equation an appropriate multiple of the first equation. The resulting system of $n - 1$ equations in the remaining $n - 1$ unknowns is similarly reduced, and the process is repeated until one equation in one unknown remains. The solution for the entire system is then found by back-substitution, that is, by solving for that one unknown in that last equation, then returning to the next-to-last equation which at the next-to-last step of the elimination involved the final unknown (now known) and one other, and solving for that second-to-last unknown, and so on.

This process may break down for two reasons: (1) when it comes time to eliminate the kth unknown, its coefficient in the kth equation may be zero, and hence the equation cannot be used to eliminate the kth unknown from equations $k + 1, \ldots, n$; and (2) the process may be very unstable. Both

difficulties can be overcome by pivoting, in which one elects, at the beginning of the kth step, a suitable equation from among equations $k, \ldots, n$, interchanges it with the kth equation, and then proceeds as before. In this way the first difficulty may be avoided provided that the system has one and only one solution. Further, with an appropriate pivoting strategy, the second difficulty may be avoided provided that the linear system is stable. Explicitly, it can be shown that, with the appropriate pivoting strategy, the solution computed in the presence of rounding errors is the exact solution of a linear system whose coefficients usually differ by not much more than roundoff from the given ones. The computed solution is therefore close to the exact solution provided that such small changes in the given system do not change its solution by much. A rough but common measure of the stability of a linear system is the condition of its coefficient matrix. This number is computed as the product of the norm of the matrix and of its inverse. The reciprocal of the condition therefore provides an indication of how close the matrix is to being noninvertible or singular.

Iterative methods. The direct methods described above require a number of operations which increases with the cube of the number of unknowns. Some types of problems arise wherein the matrix of coefficients is sparse, but the unknowns may number several thousand; for these, direct methods are prohibitive in computer time required. One frequent source of such problems is the finite difference treatment of partial differential equations (discussed below). A significant literature of iterative methods exploiting the special properties of such equations is available. For certain restricted classes of difference equations, the error in an initial iterate can be guaranteed to be reduced by a fixed factor, using a number of computations that is proportional to $n \log n$, where n is the number of unknowns. Since direct methods require work proportional to n^3, it is not surprising that as n becomes large, iterative methods are studied rather closely as practical alternatives.

The most straightforward iterative procedure is the method of substitution, sometimes called the method of simultaneous displacements. If the equations for $i = 1, \ldots, n$ are shown in Eq. (1), then

$$\sum_{j=1}^{n} a_{ij} x_j = b_i \tag{1}$$

the rth iterate is computed from the $r - 1$st by solving the trivial equations for $x_i^{(r)}$ shown in Eq. (2) for

$$\sum_{j \neq i} a_{ij} x_j^{(r-1)} + a_{ii} x_i^{(r)} = b_i \tag{2}$$

$i = 1, \ldots n$, where the elements $x_i^{(0)}$ are chosen arbitrarily. If for $i = 1, \ldots, n$, the inequality

$$\sum_{j \neq i} |a_{ij}| \leq |a_{ii}|$$

holds for some i, and the matrix is irreducible, then $x_i^{(r)} \longrightarrow x_i$ is the solution. For a matrix to be irreducible, the underlying simultaneous system must not have any subset of unknowns that can be solved for independently of the others. For practical problems for which convergence occurs, analysis shows the expected number of iterations required to guarantee a fixed error reduction to be proportional to the number of unknowns. Thus the total work is proportional to n^2.

The foregoing procedure may be improved several ways. The Gauss-Seidel method, sometimes called the method of successive displacements, represents the same idea but uses the latest available values. Equation (3) is solved for $x_i^{(r)}$, $i = 1, \ldots, n$. The

$$\sum_{j<i} a_{ij} x_j^{(r)} + a_{ii} x_i^{(r)} + \sum_{j>i} a_{ij} x_j^{(r-1)} = b_i \tag{3}$$

Gauss-Seidel method converges for the conditions given above for the substitution method and is readily shown to converge more rapidly.

Further improvements in this idea lead to the method of successive overrelaxation. This can be thought of as calculating the correction associated with the Gauss-Seidel method and overcorrecting by a factor ω. Equation (4) is first solved for y. Then

$$\sum_{j>i} a_{ij} x_j^{(r)} + a_{ii} y + \sum_{j>i} a_{ij} x_j^{(r-1)} = b_i \tag{4}$$

$x_i^{(r)} = x_i^{(r-1)} + \omega(\mathrm{y} - x_i^{(r-1)})$. Clearly, choosing $\omega = 1$ yields the Gauss-Seidel method. For problems of interest arising from elliptic difference equations, there exists an optimum ω that guarantees a fixed error reduction in a number of iterations proportional to $n^{1/2}$, and thus in total work proportional to $n^{3/2}$.

A number of other iterative techniques for systems with sparse matrices have been studied. Primarily they depend upon approximating the given matrix with one such that the resulting equations can be solved directly with an amount of work proportional to n. For a quite large class of finite difference equations of interest, the computing work to guarantee a fixed error reduction is proportional to $n^{5/4}$. The work requirement proportional only to $n \log n$ quoted earlier applies to a moderately restricted subset.

Overdetermined linear systems. Often an overdetermined linear system has to be solved. This happens, for example, if one wishes to fit the model [Eq. (5)]

$$p(x) = \sum_{j=1}^{n} a_j p_j \tag{5}$$

to observations $(x_i, y_i)_{i=1}^{m}$ with $n < m$. Here one would like to determine the coefficient vector $\mathbf{a} = (a_1, \ldots, a_n)^T$ so that $p(x_i) = y_i$, $i = 1, \ldots, m$. In matrix notation, one wants $A\mathbf{a} = \mathbf{y}$, with A the m-by-n matrix $[p_j(x_i)]$. If $n < m$, one cannot expect a solution, and it is then quite common to determine $\mathbf{a}$ instead by least squares, that is, so as to minimize the "distance" $(\mathbf{y} - A\mathbf{a})^T(\mathbf{y} - A\mathbf{a})$ between the vectors $\mathbf{y}$ and $A\mathbf{a}$. This leads to the so-called normal equations $A^T A\mathbf{a} = A^T\mathbf{y}$ for the coefficient vector $\mathbf{a}$. But unless the "basis functions" $p_1, \ldots, p_n$ are chosen very carefully, the condition of the matrix $A^T A$ may be very

bad, making the elimination process outlined above overly sensitive to rounding errors. It is much better to make use of a so-called orthogonal decomposition for A, as follows. *See* LEAST-SQUARES METHOD.

Assume first that A has full rank (which is the same thing as assuming that the only linear combination p of the functions $p_1, \ldots, p_n$ that vanishes at all the points $x_1, \ldots, x_m$ is the trivial one, the one with all coefficients zero). Then A has a QR decomposition, that is, $A = QR$, with Q an orthogonal matrix (that is, $Q^T = Q^{-1}$), and R an m-by-n matrix whose first n rows contain an invertible upper triangular matrix R_1, while its remaining m-n rows are identically zero. Then $(\mathbf{y} - A\mathbf{a})^T \cdot (\mathbf{y} - A\mathbf{a}) = (Q^T\mathbf{y} - R\mathbf{a})^T(Q^T\mathbf{y} - R\mathbf{a})$ and, since the last m-n entries of $R\mathbf{a}$ are zero, this is minimized when the first n entries of $R\mathbf{a}$ agree with those of $Q^T\mathbf{y}$, that is $R_1\mathbf{a} = [(Q^T\mathbf{y})(i)]^n{}_1$. Since R_1 is upper triangular, this system is easily solved by back-substitution, as outlined above. The QR decomposition for A can be obtained stably with the aid of Householder transformations, that is, matrices of the simple form $H = I - (2/\mathbf{u}^T\mathbf{u})\mathbf{u}\mathbf{u}^T$, which are easily seen to be orthogonal and even self-inverse, that is, $H^{-1} = H$. In the first step of the process, A is premultiplied by a Householder matrix H_1 with $\mathbf{u}$ so chosen that the first column of H_1A has zeros in rows $2, \ldots, m$. In the next step, one premultiplies H_1A by H_2 with $\mathbf{u}$ so chosen that H_2H_1A retains its zeros in column 1 and has also zeroes in column 2 in rows $3, \ldots, m$. After $n - 1$ such steps, the matrix $R := H_{n-1} \ldots H_1A$ is reached with zeros below its main diagonal, and so $A = QR$ with $Q := H_1 \ldots H_{n-1}$.

The situation is somewhat more complicated when A fails to have full rank or when its rank cannot be easily determined. In that case, one may want to make use of a singular value decomposition for A, which means that one writes A as the product USV, where both U and V are orthogonal matrices and $S = (s_{ij})$ is an m-by-n matrix that may be loosely termed "diagonal," that is, $s_{ij} = 0$ for $i \neq j$. Calculation of such a decomposition is more expensive than that of a QR decomposition, but the singular value decomposition provides much more information about A. For example, the diagonal elements of S, the so-called singular values of A, give precise information about how close A is to a matrix of given rank, and hence make it possible to gauge the effect of errors in the entries of A on the rank of A. *See* CURVE FITTING; INTERPOLATION; LINEAR SYSTEMS OF EQUATIONS; MATRIX THEORY.

Differential equations. Classical methods yield practical results only for a moderately restricted class of ordinary differential equations, a somewhat more restricted class of systems of ordinary differential equations, and a very small number of partial differential equations. The power of numerical methods is enormous here, for in quite broad classes of practical problems relatively straightforward procedures are guaranteed to yield numerical results, whose quality is predictable.

Ordinary differential equations. The simplest system is the initial value problem in a single unknown, $y' = f(x,y)$, and $y(a) = \eta$, where y' means dy/dx, and f is continuous in x and satisfies a Lipschitz condition in y; that is, there exists a constant K such that for all x and y of interest, $|f(x,y) - f(x,z)| \ll K|y - z|$. The problem is well posed and has a unique solution.

The Euler method is as follows: $y_0 = \eta$, and Eq. (6)

$$y_{i+1} = y_i + hf(x_i, y_i) \tag{6}$$

holds for $i = 0, 1, \ldots, [(b - a)/h] - 1$. Here h is a small positive constant, and $x_i = a + ih$. Analysis shows that as $h \to 0$, there exists a constant C such that $|y_k - y(x_k)| \leq Ch$, where $y(x_k)$ is the value of the unique solution at x_k, and $a \leq x_k \leq b$. This almost trivial formulation of a numerical procedure thus guarantees an approximation to the exact solution to the problem that is arbitrarily good if h is sufficiently small, and it is easy to implement. A trivial extension of this idea is given by the method of Heun, Eq. (7).

$$y_{i+1} = y_i + {}^1\!/_2 h[(f(x_i, y_i) + f(x_i + h, y_i + hf(x_i, y_i))] \tag{7}$$

The method of Heun similarly is guaranteed to approximate the desired solution arbitrarily well since there exists another constant C_1 such that the $\{y_i\}$ satisfy relation (8). This is clearly asymptotically better than the method of Euler. It is readily found to be practically superior for most problems. Further improvement of this type is offered by the classical Runge-Kutta method, Eq. (9),

$$|y_k - y(x_k)| \leq C_1h^2 \tag{8}$$

$$y_{i+1} = y_i + h\phi(x_i, y_i, h) \tag{9}$$

where $\phi(x,y,h) = {}^1\!/_6[k_1 + 2k_2 + 2k_3 + k_4]$, and $k_1 = f(x,y)$, $k_2 = f(x + h/2, y + hk_1/2)$, $k_3 = f(x + h/2, y + hk_2/2)$, and $k_4 = f(x + h, y + hk_3)$. For this method there exists a constant C_2 such that $|y_k - y(x_k)| \leq C_2h^4$.

The foregoing methods are called single-step since only y_i is involved in the computation of y_{i+1}. The single-step methods yielding the better results typically require several evaluations of the function f per step. By contrast, multistep methods typically achieve high exponents on h in the error bounds without more than one evaluation of f per step. Multistep methods require use of $y_{i-\alpha}$ or $f(x_{i-\alpha}, y_{i-\alpha})$ or both for $\alpha = 0, 1, \ldots, j$ to compute y_{i+1}. Typical is the Adams-Bashforth method for $j = 5$, Eq. (10).

$$\begin{aligned} y_{i+1} = y_i + \tfrac{h}{1440}[&4277f(x_i, y_i) - 7923f(x_{i-1}, y_{i-1}) \\ &+ 9982f(x_{i-2}, y_{i-2}) - 7298f(x_{i-3}, y_{i-3}) \\ &+ 2277f(x_{i-4}, y_{i-4}) - 475f(x_{i-5}, y_{i-5})] \end{aligned} \tag{10}$$

Analysis shows the solution of this Adams-Bashforth procedure to satisfy relation (11) for some

$$|y_k - y(x_k)| \leq C_3h^6 \tag{11}$$

constant C_3. Many valuable multistep methods have been studied. If f is nontrivial to evaluate, the multistep methods are less work to compute than the

single-step methods for comparable accuracy. The chief difficulty of multistep methods is that they require j starting values and, therefore, cannot be used from the outset in a computation.

Partial differential equations. Methods used for partial differential equations differ significantly, depending on the type of equation. Typically, parabolic equations are considered for which the prototype is the heat flow equation, (12), with $u(x,0)$ given on $x \in$

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t} \tag{12}$$

[0,1], say, and $u(0,t)$ and $u(1,t)$ given for $t > 0$. A typical finite difference scheme is shown in Eq. (13),

$$(w_{i,n})_{x\bar{x}} = \frac{w_{i,n+1} - w_{in}}{k} \tag{13}$$

where $i - 1, \ldots, 1/h - 1$, $w_{0,n} = u(0,t_n)$, and $w_{1/h,>n} = u(1,t_n)$, with $(w_i)_x = (w_{i+1} - w_i)/h$, $(w_i)_{\bar{x}} = (w_{i-1})_x$, and $w_{i,n}$ is the function defined at $x_i = ih$, $t_n = nk$. Analysis shows that as h, $k \to 0$, the solution $w_{i,n}$ satisfies $|w_{i,n} - u(x_i, t_n)| < C(h^2 + k)$ for some constant C if $k/h^2 \leq {}^1/_2$, but for k/h^2 somewhat larger than ${}^1/_2$, $w_{i,n}$ bears no relation at all to $u(x_i,t_n)$.

The restriction $k/h^2 \leq {}^1/_2$ can be removed by using the implicit difference equation, (14), but

$$(w_{i,n+1})_{x\bar{x}} = \frac{w_{i,n+1} - w_{in}}{k} \tag{14}$$

now simultaneous equations must be solved for $w_{i,n+1}$ each step. The inequality, $|w_{i,n} - u(x_i, t_n)| < C(h^2 + k)$, still holds for some constant C. An improved implicit formulation is the Crank-Nicolson equation (15).

$${}^1/_2(w_i, n + w_{i,n+1})_{x\bar{x}} = \frac{w_{i,n+1} - w_{i,n}}{k} \tag{15}$$

As $h,k \to 0$, solutions satisfy relation (16) for some

$$\left|w_{i,n} - u(x_i, t_n)\right| < C(h^2 + k^2) \tag{16}$$

constant C; again h/k is unrestricted. Such techniques can extend to several space variables and to much more general equations. The work estimates given above for iterative solution of simultaneous equations, such as Eq. (15), apply for two-space variables.

Work using variational techniques to approximate the solution of parabolic and elliptic equations has been most fruitful for broad classes of nonlinear problems. The technique reduces partial differential equations to systems of ordinary differential equations. Analysis shows that solutions obtained approximate the desired solution, as within a constant multiple of the best that can be achieved within the subspace of the basis functions used. Practical utilization suggests this to be the direction for most likely future developments in the treatment of partial differential systems. *See* DIFFERENTIAL EQUATION.

Carl de Boor

Bibliography. W. Ames, *Numerical Methods for Partial Differential Equations*, 3d ed., 1992; K. E. Atkinson, *Elementary Numerical Analysis*, 2d ed., 1993; J. C. Butcher, *The Numerical Analysis of Ordinary Differential Equations*, 1987; M. Friedman and A. Kandel, *Fundamentals of Computer Numerical Analysis*, 1993; C. F. Gerald and P. O. Wheatley, *Applied Numerical Analysis*, 6th ed., 1998; V. A. Patel, *Numerical Analysis*, 1993; H. R. Schwarz and J. Waldvogel, *Numerical Analysis: A Comprehensive Introduction*, 1989.

Numerical representation (computers)

Numerical data in a computer are written in basic units of storage made up of a fixed number of consecutive bits. The most commonly used units in the computer and communication industries are the byte (8 consecutive bits), the word (16 consecutive bits), and the double word (32 consecutive bits). A number is represented in each of these units by setting the bits according to the binary representation of the number. By convention the bits in a byte are numbered, from right to left, beginning with zero. Thus, the rightmost bit is bit number 0 and the leftmost bit is number 7. The rightmost bit is called the least significant bit, and the leftmost bit is called the most significant bit. Higher units are numbered also from right to left. In general, the rightmost bit is labeled 0 and the leftmost bit is labeled $(n - 1)$, where n is the number of bits available. *See* BIT.

Since each bit may have one of two values, 0 or 1, n bits can represent 2^n different unsigned numbers. The range of these nonnegative integers varies from 0 to $2^n - 1$.

Sign bit. To represent positive or negative numbers, one of the bits is chosen as the sign bit. By convention, the leftmost bit (or most significant bit) is considered the sign bit. This convention is shown in the **illustration** for a byte. A value of 0 in the sign bit indicates a positive number, whereas a value of 1 indicates a negative one. A similar convention is followed for higher storage units, including words and double words.

Representation of integers. Traditionally, the computer representation of integer numbers has been referred to as fixed-point representation. This usage assumes that there is a "binary" point located at the right end of the bit-string representation of the number. Several conventions have been proposed to represent integer values in a computer.

Sign-magnitude convention. To represent negative numbers using n bits, in this convention, the sign bit is set to 1 and the remaining $n - 1$ bits are used to represent the absolute value of the number (its magnitude). Positive numbers are represented in a similar way except that the sign bit is set to 0. The range

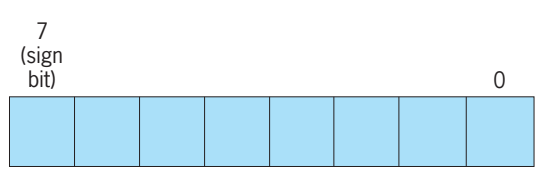

Representation of a byte in computer storage.

of integer values that can be represented with this convention using n bits varies from $-2^{n-1} + 1$ to $2^{n-1} - 1$. For example, the representation of $+75$ using this convention is $01001011_{(2}$. The sign bit is 0 since the number is positive. The remaining 7 bits are used to represent the magnitude. Similarly, the representation of -75 is $11001011_{(2}$. The sign bit is set to 1 since the number is negative. The subscript (2 at the right of the least significant digit indicates that the number is binary.

Although this convention is very simple to implement, computer manufacturers have not favored it due to the following drawbacks: First, there are two different representations for the number zero (10000000 and 00000000). Second, to carry out arithmetic operations with numbers of different sign requires that the magnitudes of the individual numbers involved be determined in advance to determine the sign of the result.

One's complement notation. This convention has the advantage of making the addition of two numbers with different signs the same as for two numbers with the same sign. Positive numbers are represented in the usual way. To represent a negative number in this convention, find the binary representation of the absolute value of the number and then complement all the bits, including the sign bit. To complement the bits of a binary number, change all 0's to 1's and all 1's to 0's. For example, using a word, the one's complement binary representations of decimal 15 and -15 are $0000000000001111_{(2}$ and $1111111111110000_{(2}$ respectively.

One's complement notation also suffers from the "negative and positive" zero malady. When a computer is performing computations, the hardware needs to check for both representations of zero. Arithmetic operations in one's complement notation sometimes require some additional checks at the hardware level. This and the double representation of zero have made this convention an unpopular choice.

Two's complement notation. The most widely used convention for representing negative numbers is the two's complement notation. This variation of the one's complement method corrects the problems of the negative zero and the additional checks at the hardware level required by the one's complement. The representation of positive numbers in this new convention is similar to that of the one's complement. To represent negative numbers in this convention, the following algorithm can be used: Scan the positive binary representation of the number from right to left until the first 1-bit is encountered. Leave this bit and all the 0-bits to its right, if any, unchanged. Complement all bits to the left of this first 1-bit. For example, according to this algorithm, the two's complement representation of 00010100 is 11101100.

The range of numbers that can be represented using this method with n bits is -2^{n-1} to $2^{n-1} - 1$. The range of negative numbers is one unit larger than the range of positive values. *See* NUMBERING SYSTEMS.

Representation of real numbers. The convention for representing real numbers, also called floating-point numbers, is similar to the scientific notation for representing very small or very large real numbers. The representation of floating-point has the format $\pm f \times r^p$. Here, f is called the mantissa and is restricted to values in a small range, usually, $0 \leq f < 1$. The value r, called the radix, is generally restricted to the values 2, 8, or 16. The power p can be either positive or negative. On a computer, the values of both f and p are restricted by the number of bits used to store them.

Computer manufacturers sometimes define their own standards for representing floating-point numbers. The range of the numbers that can be represented depends upon the number of bits used to represent f and p. Some typical range values are 0.29×10^{-38} to 1.7×10^{38} (with f and p represented by 23 and 8 bits respectively), and 0.84×10^{-4932} to 0.59×10^{4932} (with f and p represented by 112 and 15 bits respectively). *See* DIGITAL COMPUTER.

Ramon A. Mata-Toledo

Bibliography. S. Baase, *VAX Assembly Language*, 2d ed., 1992; R. P. McArthur, *Logic to Computing*, 1991.

Numerical taxonomy

As defined by P. H. A. Sneath and R. R. Sokal, the grouping by numerical methods of taxonomic units based on their character states. According to G. G. Simpson, taxonomy is the theoretical study of classification, including its bases, principles, procedures, and rules. The application of numerical methods to taxonomy, dating back to the rise of biometrics in the late nineteenth century, has received a great deal of attention with the development of the computer and computer technology. Numerical taxonomy provides methods that are objective, explicit, and repeatable, and is based on the ideas first put forward by M. Adanson in 1963. These ideas, or principles, are that the ideal taxonomy is composed of information-rich taxa ("taxon," plural "taxa," is the abbreviation for taxonomic group of any nature or rank as suggested by H. Lam) based on as many features as possible, that a priori every character is of equal weight, that overall similarity between any two entities is a function of the similarity of the many characters on which the comparison is based, and that taxa are constructed on the basis of diverse character correlations in the groups studied.

In the early stages of development of numerical taxonomy, phylogenetic relationships were not considered. However, numerical methods have made possible exact measurement of evolutionary rates and phylogenetic analysis. Furthermore, rapid developments in the techniques of direct measurement of the homologies of deoxyribonucleic acid (DNA), and ribonucleic acid (RNA), between different organisms now provide an estimation of "hybridization" between the DNAs of different taxa and, therefore, possible evolutionary relationships. Thus, research in numerical taxonomy often includes analyses of the chemical and physical properties of the nucleic acids

of the organisms; the data from these analyses are correlated with phenetic groupings established by numerical techniques. *See* PHYLOGENY; TAXONOMIC CATEGORIES.

Mathematical approach. Classification is recognized by philosophers as one of the main elements in human language. Two broad types of classification have been distinguished, natural and artificial. The essential difference between the two types in the extreme forms is that the objects in a natural classification show a likeness in a large number of attributes, whereas those in an artificial classification show a likeness in only a small number of attributes, possibly only one.

Two methods of classification, at the extreme, are fairly distinct: the typological method, linked with a natural classification, and the definitional method, linked with an artificial classification. The definitional classification consists of selecting one or perhaps two attributes and defining the groups as all the objects which show these attributes. The typological method is not a deliberate process but one by which the classes gradually emerge around a notional type. The Adansonian principle, based on an empirical approach, estimates overall similarity between organisms, based on the examination of a large number of equally weighted features. Several formulas have been suggested for calculating similarity between organisms. Resemblance is expressed by coefficients of similarity usually ranging between unity and zero, the former for perfect agreement and the latter for no agreement. Coefficients of dissimilarity (distance) can also be employed, in which case the range is between zero and an undefined positive value, the former for identity and the latter for maximal distance or disparity. Resemblance coefficients are usually tabulated in the form of a matrix, with one coefficient for every pair of taxonomic entities. The similarity index is the simplest of the coefficients employed to calculate similarity: $S = n_s/(n_s + n_d)$, where n_s represents the number of positive features shared by two strains being compared and n_d represents the number of features positive for one strain and negative for the other (and vice versa). An example of an S-value matrix is shown in **Fig. 1**.

A variety of clustering methods has been employed in numerical taxonomy, including sequential, agglomerative, hierarchic, and nonoverlapping clustering. Dendograms, an example of which is given in **Fig. 2**, are most commonly employed to represent taxonomic structure resulting from cluster analysis. These have the advantage of being readily interpretable as conventional taxonomic hierarchies. Dendrograms are, in general, distinguished into phenograms representing phenetic relationships and cladograms representing evolutionary branching sequences.

Mathematics, in one form or another, whether a calculation of similarity indices, factor analysis to summarize complex relationships in terms of a few factors, or discriminatory analysis (the identification of an entity with an existing group by using attributes in such a way that the chance of correct identification is as high as possible), is successfully employed in classification. The availability of electronic computers renewed interest and sparked new developments in the mathematical approach to classification. A large number of attributes scored for each individual in a large sample set are easily managed by computer, the tool of the numerical taxonomist. Scoring taxonomic data for handling by the computer poses two questions: what is the actual scoring process and what number of features (unit characters) must be scored for an operational taxonomic unit (OTU) in numerical analysis? Simply stated, the data are scored in format suitable for transfer into the computer, via punched tape, punched cards, or magnetic tape, and for handling by programmed instructions stored in the computer. The data are usually scored as + or 1 for positive features, − or 0 for features negative or absent, and NC or any alphabetical or numerical term for features which are not applicable or not included in the calculation. Characters are broken down into unit characters, with the aim that each unit, as far as possible, should contribute one new item of information that is relevant to taxonomy; that is, the unit will permit distinguishing one kind of organism from another, or the unit varies from one kind of organism to another. The unit character has been defined by Sneath and Sokal as "a taxonomic character of two or more states, which within the study at hand cannot be subdivided logically, except for subdivision brought about by the method of coding." The kinds of information which are relevant are decided by the taxonomist.

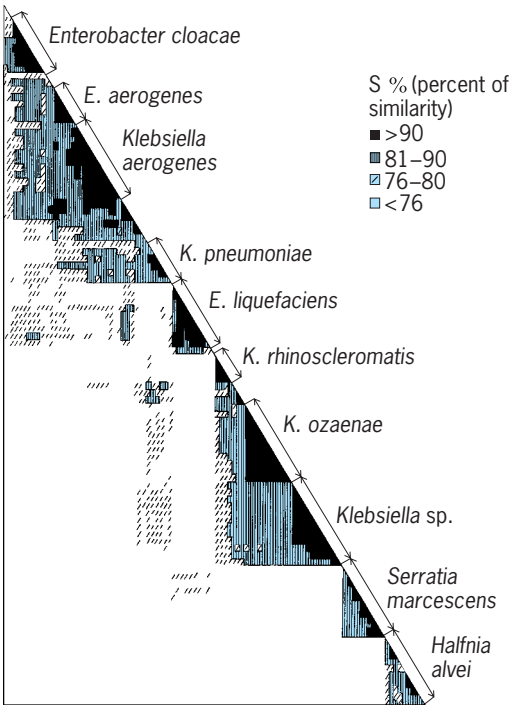

Fig. 1. Similarity matrix showing relationships between clusters of strains of enteric bacteria. (*After R. Johnson et al., Numerical taxonomy study of the Enterobacteriaceae, Int. J. System. Bacteriol., 25(1):12–37 1975*)

The general parameter sought in the comparison of samples of numbers of characters is the estimate

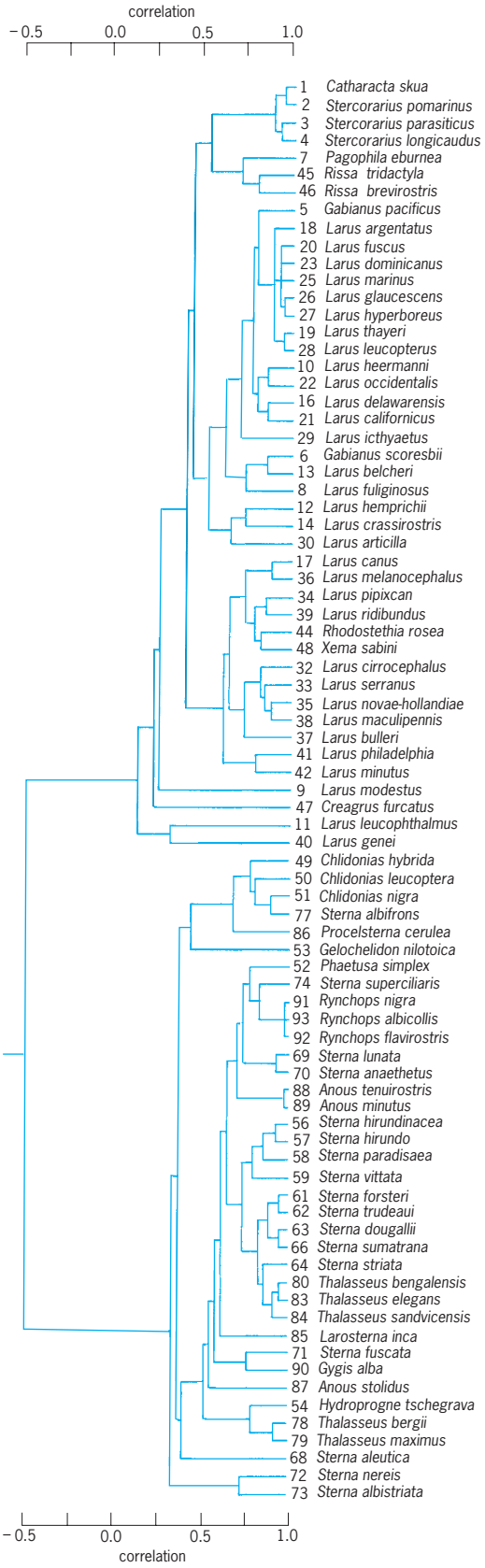

Fig. 2. Phenogram of 81 species of gulls and terns based on UPGMA cluster analysis of correlation coefficients on 51 skeletal measurements. (*After G. O. Schnell, A phenetic study of the suborder Lari (Aves), II. Phenograms, discussions, and conclusions, System. Zool., 19:264–302, 1970*)

of similarity or total phenetic resemblance; one assumes that the phenetic information of each OTU is finite. Thus, the second question can be answered by applying the matches asymptote hypothesis of Sokal and Sneath, the best approach to the question of "how many characters should be sampled for a general taxonomy?" since it utilizes a testable empirical standard. The hypothesis simply assumes that as the number of characters sampled increases, the value of the similarity coefficient becomes more stable; eventually a further increase in the number of characters is not warranted by the corresponding mild decrease in the width of the confidence band of the coefficient. C. D. Michener and Sokal ventured a suggestion of the number of characters—not less than 60. The most practical advice to follow is that as many characters as is feasible should be distributed as widely as possible over the various body regions, life history stages, tissues, and levels of organization of the organisms. The capacities of computers which are available for numerical taxonomy computations rarely impose data storage limitations on the investigator.

Computer-assisted identification and classification of organisms have been made possible by establishment of data matrices that are stored in the computer, using keys or discriminant functions to achieve identification of unknown specimens. Online computer identification is now used to a great extent in microbiology and pathology.

The applications of numerical taxonomy cover many scientific disciplines since mathematical methods of classification can be applied to any material and in any discipline. Zoologists and botanists with a tradition of Linnaean taxonomy and a more or less standardized procedure find that where numerical methods have been used, the method has proved valuable for predictive purposes. Fields, besides botany and zoology, in which numerical methods of classification have been successful are linguistics, archeology, social anthropology, psychiatry, sociology, medical diagnosis, microbiology, ecology, biogeography, and earth sciences. Numerical taxonomy studies have spanned a variety of organisms, including vertebrates, arthropods, dicotyledons, monocotyledons, bacteria, and viruses. *See* ANIMAL SYSTEMATICS; PLANT TAXONOMY.

Nucleic acid data. Microbiologists have been ready and willing to apply the methods of numerical taxonomy, chiefly because extremely scanty paleontological (fossil) records of microorganisms exist; hence microbiologists have not been evolutionarily oriented. Developments in molecular biology which provide DNA-DNA and DNA-RNA pairings between organisms, that is, essentially detailed comparisons of the DNA base sequence of two organisms, have made evolutionary conclusions at least theoretically possible. *See* MOLECULAR BIOLOGY.

Numerical taxonomy analyses of several bacterial genera, namely *Escherichia*, *Serratia*, *Pseudomonas*, and *Staphylococcus*, show strong correlation between clusters obtained by numerical taxonomy and DNA base composition data for these

strains. Bacterial strains clustered by mathematical methods and shown to share high similarity indices share identical overall DNA base composition. The identical or very nearly identical overall DNA base composition thus is considered to be a highly correlative feature of genetically interacting microbes. In addition, phenetic groupings, obtained by numerical analysis, have been shown to be homologous, with respect to DNA/DNA reassociation analyses. Thus, DNA base sequence studies have proven useful in elucidating the genotype-phenotype relationships revealed by numerical taxonomy. *See* BACTERIAL TAXONOMY. R. R. Colwell

Bibliography. D. L. Quicke, *Principles and Techniques of Contemporary Taxonomy*, 1993; B. Rosner, *Fundamentals of Biostatistics*, 5th ed., 1999; S. Selvin, *Practical Biostatistical Methods*, 1995; G. G. Simpson, *Principles of Animal Taxonomy*, 1990; P. H. A. Sneath and R. R. Sokal, *Numerical Taxonomy: The Principles and Practice of Numerical Classification*, 1973; R. R. Sokal and J. F. Rohlf, *Biometry: The Principles and Practice of Statistics in Biological Research*, 3d ed., 1994.

Nummulites

A genus of unicellular shelled protoctist (protozoa) of the class Foraminifera (order Rotaliida, family Nummulitidae). The shells of *Nummulites* and other large, internally complex foraminifers commonly occurred in rock-forming abundances on continental shelves and around oceanic islands throughout subtropical and tropical regions during the Cenozoic. Resulting limestones are important hydrocarbon reservoirs from the Philippines and Indonesia westward to Trinidad and Venezuela. Nummulitic limestones have been used as building materials around the Mediterranean; the pyramids of Egypt were constructed of blocks of Eocene age. *Nummulites* shells are useful stratigraphic and paleoenvironmental indicators; the Paleogene was formerly known as the Nummulitic Period. *See* LIMESTONE.

Shell morphology. The shells, also known as tests, are composed of finely perforate calcite; may be discoidal, lenticular, or globular in shape; and can be as much as 6 in. (15 cm) in diameter, though sizes of 1–20 mm are more common. Shells are planispirally enrolled whorls of tiny undivided chambers (see **illustration**). The first-formed chamber (proloculus) and second chamber (deuteroconch) are connected by a central pore. In subsequent chambers, the test opening (aperture) consists of rows of pores at the base of the apertural face. Within the wall, a system of canals includes a distinct marginal cord at the outer periphery, branching sutural canals, and radial passages connecting canals of successive whorls.

Ecology. Modern representatives of the Nummulitidae are benthic (bottom-dwelling) foraminifers that harbor diatom (single-celled alga) symbionts. The symbionts live within the host cytoplasm and assist the host in metabolism and calcification by providing photosynthetically produced carbon. Paleontological studies indicate that fossil *Nummulites* lived in similar symbioses. Because the diatoms require sunlight to photosynthesize, the foraminifers must live at depths to which some sunlight can penetrate. The shapes of the shells indicate the amount of both light and water motion that the foraminifers were exposed to in life; globular shells represent abundance of both, while very flat, delicate tests indicate minimal light and water motion. Since both light and wave motion decrease exponentially with depth, geologists use the shell shapes to estimate water depth at which the foraminifers lived. Very flat, delicate forms can be found living today as deep as 500 ft (~150 m).

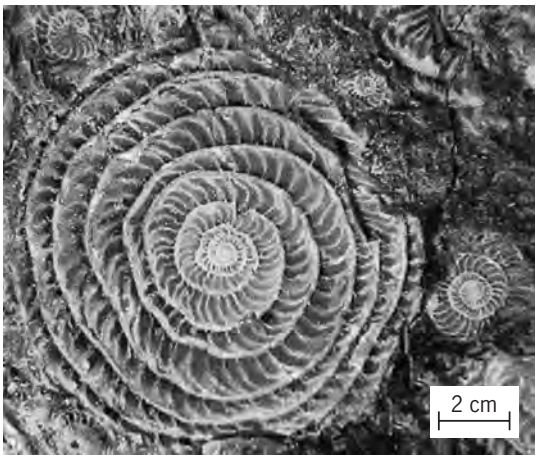

Nummulites **in a borehole core through the early Eocene El Garia Formation of the Hasdrubal Gas Field, offshore Tunisia. The large individual developed from gametes; the smaller specimens were likely produced asexually. (*Photo courtesy of Simon Beavington-Penney*)**

Reproduction. Reproduction is through alternation of sexual and asexual generations. Haploid asexual generations are produced by multiple fission; the parent extrudes its cytoplasm, in which hundreds of embryos (proloculus plus deuteroconch) develop. The embryos incorporate both the cytoplasm and symbionts of the parent before dispersing away from the empty parent shell. The offspring grow for several months, then may produce either another asexual generation or millions of tiny, identical gametes. Embryos resulting from the fusion of two gametes are much smaller than asexually produced embryos, yet individuals grow to much larger sizes before reproducing. Thus, dimorphism is the norm, with the diploid generation characterized by larger adult shells (typically less than 5 mm) with tiny embryos (less than 0.02 mm) while haploid asexual generations have smaller adult shells (usually less than 5 mm) with larger embryos (0.1 mm or more). *See* FORAMINIFERIDA. Alfred R. Loelich, Jr.; Pamela Hallock

Bibliography. S. J. Beavington-Penney and A. Racey, Ecology of extant nummulitids and other larger benthic foraminifera: Applications in palaeoenvironmental analysis, *Earth-Sci. Rev.*, 67(3–4):219–265, 2004; J. Hohenegger, E. Yordanova, and A. Hatta, Remarks on West Pacific Nummulitidae (foraminifera), *J. Foram. Res.*, 30(1):3–28, 2000; M. Holzmann,

J. Hohenegger, and J. Pawlowski, Molecular data reveal parallel evolution in nummulitid, *J. Foram. Res.*, 33(4):277–284, 2003; L. M. A. Purton and M. D. Brasier, Giant protist *Nummulites* and its Eocene environment: Life span and habitat insights from delta O–18 and delta C-13 data from Nummulites and Venericardia, Hampshire basin, UK, *Geology*, 27(8): 711–714, 1999; A. Racey, A review of Eocene nummulite accumulations: Structure, formation and reservoir potential, *J. Petrol. Geol.*, 24(1):79–100, 2001; B. K. Sen Gupta (ed.), *Modern Foraminifera*, Kluwer Academic, Dordrecht, The Netherlands, 1999.

Nursing

The application of principles from the basic sciences, social sciences, and humanities to assist healthy and sick individuals and their families or other caring persons in performing those activities that contribute to the individuals' physical and mental well-being and that they would perform unaided if able to do so. Nursing includes providing physical and emotional care, promoting comfort, serving as patient advocates, assisting in rehabilitative efforts, teaching self-care and health promotion activities, and administering treatments prescribed by a licensed physician or dentist. Patient-care activities are conceived and coordinated so as to help individuals gain independence as rapidly as possible or maintain an optimal level of function. When multiple health-care providers are involved, nursing coordinates patient-care efforts to improve the quality of care.

Nursing practice is conducted in a variety of settings, including hospitals, community facilities, private homes, nursing homes, schools, industry, physician's offices, the military, and civil service arenas. Standards for nursing practice and licensure are governed by state nurse practice acts and are directed by professional nursing organizations.

Two types of nurses are legally recognized in the United States: the registered professional nurse and the licensed practical nurse. Licensed practical or vocational nurses (LPNs or LVNs) are trained to perform uncomplicated patient-care tasks in hospitals or other health-care facilities under the aegis of registered nurses or physicians. Having successfully completed approximately 1 year of training in a hospital, community college, or other vocational setting, they are licensed by state boards of nursing after passing the licensing exam.

Registered professional nurses may be educated in numerous ways. Associate degree programs (2 years), hospital-based diploma schools (3 years), and baccalaureate degree programs (4 years) are the most commonly chosen routes. Several generic master's programs have also been established for those who enter nursing with a college degree. While all of these programs prepare nurses to assume beginning clinical roles, the baccalaureate degree is the first professional degree and allows the recipient greater flexibility and advancement. After completing an approved program, nurses are required to successfully complete their state licensure examination before they are recognized as registered professional nurses.

Master's and doctoral education is directed toward preparing nursing leaders. Advanced study in nursing may be pursued by qualified individuals who are seeking expanded nursing roles. These roles include, but are not limited to, clinical nurse specialist (who has been trained for work in a specific field of medicine), nurse practitioner (who has additional training for performing primary care procedures), certified nurse midwife, nurse anesthetist, nurse educator, nurse administrator, and nurse researcher.

Professional nursing is guided by numerous professional organizations. The American Nurses' Association, founded in 1896, serves as the professional organization for registered nurses. Its role is to foster high standards of nursing care, promote nurses' economic welfare, and improve the general working conditions of nursing.

The National League for Nurses resulted from a merger (1952) of the National League of Nursing Education, the National Organization for Public Health Nurses, and the Association of Collegiate League for Nursing. Supporting both licensed practical and registered nurses, it provides numerous educational services and is the recognized voluntary national accrediting body for schools of nursing.

A third organization, the International Council of Nurses, was established in 1899. It facilitates communication for nurses of varying nations on issues of patient welfare, health promotion, and the advancement of the nursing profession. In addition, more than 40 specialty organizations provide support for nursing's various constituencies. *See* MEDICINE.

Judith A. Vessey

Bibliography. L. H. Aiken, *Nursing in the 1980's: Crises, Opportunities, Challenges*, 1982; American Nurses' Association, *Facts about Nursing*, *86–87*, 1987; P. Benner, *From Novice to Expert*: *Excellence and Power in Clinical Practice*, 1984; M. P. Donahue, *Nursing*, *the Finest Art*: *An Illustrated History*, 1985; F. Nightingale, *Notes on Nursing*: *What It Is and What It Is Not*, 1860.

Nut (engineering)

In mechanical structures, an internally threaded fastener. Plain square and hexagon nuts for bolts and screws are available in three degrees of finish: unfinished, semifinished, and finished. There are two standard weights: regular and heavy. For specific applications, there are other standard forms such as jam nut, castellated nut, slotted nut, cap nut, wing nut, and knurled nut (see **illus.**).

Hexagon jam nuts are used as locking devices to keep regular nuts from loosening and for holding set screws in position. They are not as thick as plain nuts. Jam nuts are available in semifinished form in both regular and heavy weight.

Castellated and slotted nuts have slots so that a cotter pin or safety wire can hold them in place.

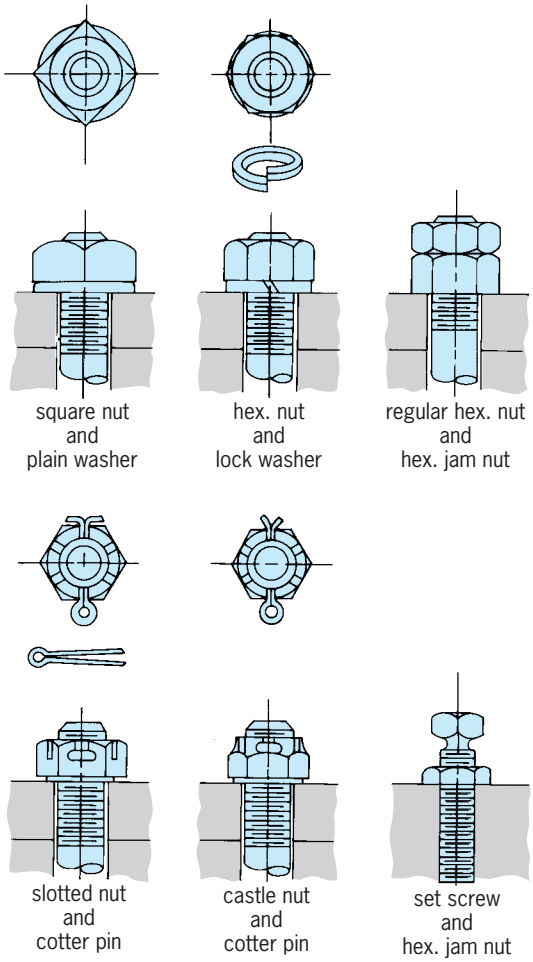

Examples of standard forms of nuts.

They are commonly used in the automotive and allied fields on operating machinery where nuts tend to loosen. The slotted nut is a regular hexagon nut with slots cut across the flats of the hexagon. Slotted nuts are standardized in regular and heavy weights in semifinished hexagon form. Finished thick slotted nuts are available. Castle nuts are hexagonal with a cylindrical portion above, through which slots are cut.

Machine screw and stove-bolt nuts may be either square or hexagonal. Hexagon machine-screw nuts may have the top chamfered at 30° with a plain bearing surface, or both top and bearing surfaces may be chamfered. Square nuts have flat surfaces without chamfer. Square nuts are available with a coarse thread; hexagon nuts may be supplied with either coarse or fine thread.

Wing and knurled nuts are designed for applications where a nut is to be tightened or loosened by using finger pressure only. *See* SCREW FASTENER.

Warren J. Luzadder

Nut crop culture

The cultivation of plants, primarily trees, that produce nuts. The term nut is used loosely. Generally, a nut is defined as any edible fruit or seed enclosed in a hard shell. Botanically, a nut is a hard, indehiscent, one-seeded (nut), pericarp (shell) generally resulting from a compound ovary of a flower. Indehiscent means the shell does not split open spontaneously when ripe. Examples are chestnuts, filberts, and acorns. Technically, a nut is a dry edible fruit consisting of a kernel or seed enclosed in a woody shell. Only a fraction of the 80 fruits and seeds designated as nuts fit this description. The peanut is notable because the plant is a herbaceous annual legume in which the nuts are analogous to peapods that mature underground. *See* FLOWER.

Although most nuts are spherical, they can vary considerably in size, shape, shell, and hull. Popular edible nuts range in size from the small sunflower seed to the large coconut. Nuts can be round (macadamias), elongated (pistachios and almonds), kidney shaped (cashews), or triangular (brazil nuts), or have the distinctive double-lobed kernel of the walnut and pecan.

The protective shell surface is called a hull (or husk) and serves a variety of purposes. Chestnut hulls are spiny. The coconut husk is thick and fibrous. The cashew hull contains a caustic mix of cardol and anacardic acid that causes a serious allergic reaction worse than that of its relative poison ivy. Extracts of walnut hulls at different stages of maturity have been used as yellow and brown dyes.

Crop plants. Plants producing nut crops are very diverse in their botanical classification, type, and climatic and cultural requirements. Pine nuts are the seeds of conifers from cool temperate climates, coconuts and brazil nuts are from a tropical palm and evergreen respectively, and pecans and walnuts are from large deciduous trees of warm temperate zones. Many nut crops were originally produced from wild trees that received little care (for example, pine nuts, brazil nuts, pistachios, and black walnuts). However, as demand increased and harvesting and processing technology improved, nut crops began to be cultivated as intensively as other fruits.

Cultivation. Cultural methods differ according to the nature of the plant, its soil requirements, the climate, available labor, and other factors. As with most orchard crops, commercially cultivated nut trees are now grafted onto rootstocks that allow better adaptation to the soil or climate, and disease or pest resistance. Relative to other tree fruits nut crop culture offers some distinct advantages in the ability to use machinery rather than hand labor for harvesting. Pruning of mature nut trees is not as critical as that of stone fruits because pruning does not play as strong a role in nut size, color development, or internal quality, as it does in fleshy fruits. Because most nut kernels are enclosed in a hard shell (often with a hull) and are too numerous to pick by hand, harvesting is frequently done by shaking-and-sweeping machines. Further, the time and length of harvest, processing, and storage is not as critical as with fleshy fruits. Thus, where the technology is available, nut crop culture is unique among the tree crops in being mechanically pruned and harvested. *See* AGRICULTURAL MACHINERY.

Disease. As with other orchard crops, pest and disease control is important in commercial plantings.

Effective control measures are related to the life history of the particular pest or disease, since each may require a different approach in cultural practice or spray program. Generally, nut crops are more amenable to organic production or integrated pest management because they have protective shells. Nut crops also have an advantage over fleshy fruits in that if a pathogen or pest does not cause economic damage by decreasing production or harm shell appearance or kernel quality, it need not be controlled. *See* PLANT PATHOLOGY.

Economic value. The edible part of most nuts, the kernel, is high in fat (30–70%) and low in carbohydrates (generally under 25%). Although nuts contain high-quality protein, they are considered primarily a source of fat and are rarely a dietary staple in modern cultures. The most common nuts, such as walnuts, almonds, pecans, and cashews, are consumed primarily as snack food or as a confectionery additive, less frequently in savory foods. Other nuts, such as the betel nuts, are masticatories, that is, nuts that are chewed but not consumed.

Not all nuts are edible or chewable. The tung nut, the source of tung oil—the active drying agent in varnish and paint with superior preservative and waterproofing qualities—is toxic. The tropical candlenut oil tree (*Aleurites triloba*) produces a poisonous nut that is a source of oil for illuminants, soaps, paints, and wood preservatives. The nutlike seeds of *Nux vomica* contain strychine, which in appropriate concentrations has been used as a heart stimulant or rat poison. The crushed nuts of the Nigerian Calabar vine (*Physostigma venenosum*) were fed to persons suspected of crimes; if the suspect vomited and survived, he was judged innocent. Physostigmine, the alkaloid used to contract pupils, is extracted from the dried nuts.

Nut-bearing plants were particularly important in supplying food for early populations and wildlife. Archeological records of primitive humans are scarce, but sites from Mesolithic to recent times contain nut and shell fragments indicating nut consumption. Changes in nut-producing trees caused changes in animal populations. The disappearance of the native chestnut in the eastern United States caused a drop in the wild turkey population, and the encroachment of civilization on the supply of beechnuts and acorns contributed to the extinction of the passenger pigeon.

With increasing world population and diversity in diets, commercially produced nut crops have become economically important (see **table**). Political

Nuts important in world commerce

Common name (scientific name)	Plant type	Origin	Climatic zone adaptation	Principal uses and producing areas
Almond (*Prunus amygdalus*)	Small deciduous tree	Asia Minor	Temperate, subtropical; planted	Food; Mediterranean
Brazil nut (*Bertholletia excelsa*)	Very large evergreen tree	Amazon basin	Tropical; wild	Food; Brazil, Bolivia
Black walnut (*Juglans nigra*)	Large deciduous tree	Central and eastern United States	Temperate; mostly wild	Food, lumber, ground-up shells; United States
Cashew (*Anacardium occidentale*)	Evergreen tree to 40 ft (12 m)	Tropical America	Tropical; wild and planted	Food, shell oil; India, East Africa
Chestnut (*Castanea* sp.)	Large, spreading, deciduous tree	Eastern United States, Europe, Asia	Temperate; wild and planted	Food, lumber; Europe, China, Japan
Coconut (*Castanea* sp.)	Large palm	Probably Polynesia	Tropical; wild and planted	Food, edible and industrial oil, fiber; Philippines, tropics of Pacific
Cola nut (*Cola acuminata*)	Evergreen tree to 40 ft (12 m)	West Africa	Tropical; wild and planted	Stimulant; West Africa
English or Persian walnut (*Juglans regia*)	Large deciduous tree	Eastern Europe, Asia	Temperate; wild and planted	Food, oil, lumber; Mediterranean basin, California, China
Filbert (*Corylus avellana*)	Deciduous shrub or small tree	Eastern Europe, Asia Minor	Temperate; wild and planted	Food; Turkey, Italy, Spain, northwestern United States
Macadamia nut (*Macadamia ternifolia*)	Evergreen tree to 60 ft (18 m)	Queensland, New South Wales	Tropical, subtropical; planted	Food; Australia, Hawaii
Palm nut (*Elaeis guineensis*)	Large palm	West and Central Africa	Tropical; wild and planted	Edible and industrial oil; West Africa, Indonesia, Brazil
Peanut (*Arachis hypogaea*)	Annual crop plant	Brazil	Tropical, subtropical, warm temperate; planted	Food, edible oil; India, Africa, United States
Pecan (*Carya pecan*)	Large, deciduous tree	Southern and central United States	Warm temperate; wild and planted	Food; central and southern United States
Pine nuts (*Pinus* sp.)	Evergreen conifers	Europe, Asia, southwestern North America	Temperate; wild	Food; southern Europe, Asia, North America
Pistachio nut (*Pistacia vera*)	Small,spreading, deciduous tree	Asia Minor	Warm temperate; wild and planted	Food; Iran, Turkey, Syria, Italy
Tung nut (*Aleurites fordii*)	Small tree to 20 ft (6 m)	Central Asia	Warm temperate, subtropical; planted	Industrial oil, paints; China, southern United States

events have also played a role in the economic development of some nut crops. The pistachio, for example, is the most successful plant introduction to the United States in the twentieth century. California would not have developed the successful pistachio industry it now has had the world's primary pistachio producer, Iran, not been disrupted by war in the late 1970s. At that point, California had a small annual production but responded so strongly to the sharply increased world demand that pistachios have become the third largest nut crop in California. *See* ALMOND; BRAZIL NUT; CASHEW; CHESTNUT; COCONUT; COLA; FILBERT; MACADAMIA NUT; PEANUT; PECAN; PINE NUT; PISTACHIO; WALNUT.

Louise Ferguson

Bibliography. E. A. Meninger, *Edible Nuts of the World*, 1977; F. Rosengarten, Jr., *The Book of Edible Nuts*, 1984.

Nutation (astronomy and mechanics)

In mechanics, nutation is a bobbing motion that accompanies the precession of a spinning rigid body, such as a top. Astronomical nutation refers to irregularities in the precessional motion of rotating bodies. A well-studied example is the irregularities in the precessional motion of the equinoxes caused by the varying torque applied to the Earth by the Sun and Moon. Astronomical nutation should not be confused with nutation as defined in mechanics; the latter is present even if the source of the torques is unvarying.

Nutation of tops. In simple precession, the axis of a top with a fixed point of contact sweeps out a cone, whose axis is the vertical direction. In the general motion, the angle between the axis of the top and the vertical varies with time (**Fig. 1**). This motion of the top's axis, bobbing up and down as it precesses, is known as nutation.

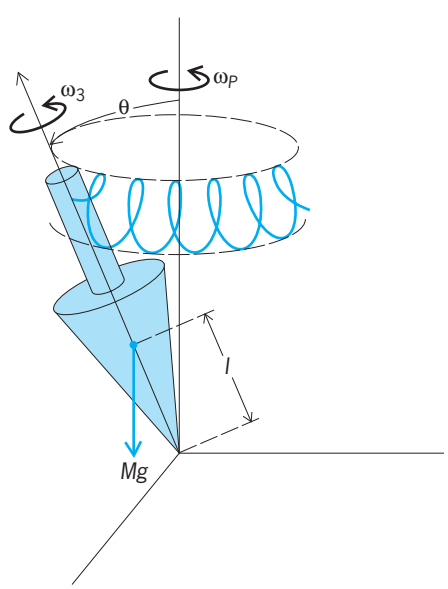

Fig. 1. Nutation and precession of a spinning top with a fixed point of contact.

Slow precession and fast nutation are commonly observed in the motion of a rapidly spinning top. For this case, the equations of motion have a simple approximate solution, which will be presented below. The angular frequency of the precessional motion is given by Eq. (1), where ω_3 is the angular velocity

$$\omega_P \simeq \frac{Mgl}{I_3\omega_3} \tag{1}$$

about the spin axis of the top, I_3 is the moment of inertia about this axis, M is the mass of the top, l is the distance of the center of mass from the point of contact, and g is the gravitational acceleration. The nutation is governed by the angular frequency given in Eq. (2), where I is the moment of inertia about an

$$\omega_N = \frac{I_3\omega_3}{I} \tag{2}$$

axis that is perpendicular to the spin axis of the top.

The location of the spin axis can be specified by an azimuthal angle ϕ and a polar angle θ. Suppose that the top is started at a polar angle $\theta(t = 0) = \theta_0$ with an initial azimuthal angular velocity $\phi(t = 0) = \omega_0$. Then the approximate solutions to the equations of motion are given by Eqs. (3), where C is given by Eq. (4).

$$\theta(t) = \theta_0 + C \sin\theta_0(1 - \cos\omega_N t) \tag{3a}$$

$$\phi(t) = \omega_P t - C \sin\omega_N t \tag{3b}$$

$$C = \frac{\omega_P - \omega_0}{\omega_N} \tag{4}$$

This solution for $\theta(t)$ oscillates between limits of θ_0 and $\theta_0 + 2C \sin\theta_0$; the top's spin axis bobs up and down between cones of these angles. The maximal excursion in θ depends on C, and hence can be changed by the initial conditions on ω_0 and ω_3. The precession angle ϕ has a sinusoidal time dependence determined by ω_N in addition to the steady precession term. When the top is set in motion with $\omega_0 = \omega_P$, it will undergo precession with no nutation. Typical curves traced out by the end of the top's symmetry axis are shown in **Fig. 2**.

The nutation frequency ω_N in Eq. (2) is proportional to the spin ω_3 of the top while the precession frequency ω_P in Eq. (1) is inversely proportional to ω_3. Also, the nutation amplitude C is inversely proportional to ω_N. Thus it is difficult to observe the nutation of a very fast top. However, a buzzing tone with the nutation frequency can be heard when a fast top is spun on a resonant surface. *See* PRECESSION; RIGID-BODY DYNAMICS.

Vernon D. Barger

Astronomical nutation. The rotating Earth can be regarded as a spinning symmetrical top with small angular speed but large angular momentum, the latter due to its large mass. The gravitational attractions of the Sun and Moon cause the Earth's axis to describe a cone about the normal to the plane of its orbit. However, the magnitude of these gravitational attractions is continually varying, due to the changing positions in space of Sun, Moon, and Earth. The

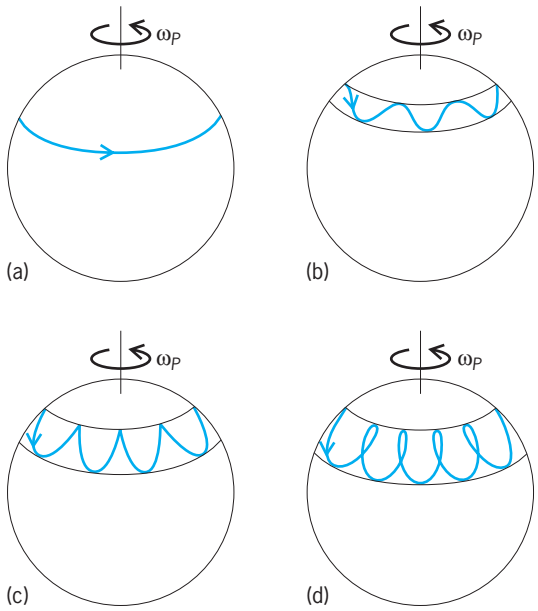

Fig. 2. Path traced out by the symmetry axis of a top with a fixed point of contact for different initial angular velocities ω_0 of the ϕ motion. (*a*) $\omega_0 = \omega_P$. (*b*) $\omega_0 = {}^1/_2\omega_P$. (*c*) $\omega_0 = 0$. (*d*) $\omega_0 = -\omega_P$.

Moon's orbit is continually changing its position in such a way that the celestial pole undergoes a nodding (nutation) as well as a periodic variation in the rate of advance. The largest nutation is about 9.″2, and occurs in a period of a little less than 19 years; that is, the celestial pole completes a small ellipse of semimajor axis 9.″2 in about 19 years. *See* PRECESSION OF EQUINOXES.

There are lesser nutation effects in the Earth-Moon system which are due to the motion of the Moon's nodes, the changing declination of the Sun, the changing declination of the Moon, and the effects of the dynamics due to the nonrigid Earth. *See* CELESTIAL MECHANICS. Ray E. Bolz

The gas giants Jupiter, Saturn, Uranus, and Neptune are much more massive than their satellites and they also have no fixed surface features, so astronomical nutations of these planets cannot be observed. Pluto should show nutation caused by its satellite Charon; however, not enough data on the rotation of Pluto have been collected to detect this nutation. Although the Earth-Moon system is the only measured example of astronomical nutation, the discovery of planets circling other stars provides the theoretical possibility of other examples. Again, such effects are far beyond current or foreseeable observational abilities. *See* EXTRASOLAR PLANETS; PLANET; PLUTO. John L. Safko

Bibliography. V. Barger and M. Olsson, *Classical Mechanics: A Modern Perspective*, 2d ed., McGraw-Hill, 1995; H. Goldstein, C. P. Poole, Jr., and J. L. Safko, *Classical Mechanics*, 3d ed., Addison Wesley, 2002; D. Halliday, R. Resnick, and K. Krane, *Physics*, 5th ed., Wiley, 2002; D. Hestenes, *New Foundations for Classical Mechanics*, Kluwer Academic, 1999; P. K. Seidelman (ed.), *Explanatory Supplement to the Astronomical Almanac*, rev. ed., University Science Books, 1992; K. Symon, *Mechanics*, 3d ed., Addison Wesley, 1971.

Nutmeg

A delicately flavored spice obtained from the nutmeg tree (*Myristica fragrans*), a native of the Moluccas, or Spice Islands. The tree is a dark-leafed evergreen 30–60 ft (9–18 m) high, and is a member of the nutmeg family (Myristicaceae). The golden-yellow, mature fruits resemble apricots (see **illus.**).

Nutmeg (*Myristica fragrans*), mature fruits. (*USDA*)

They gradually lose moisture and when completely ripe, the husk (pericarp) splits open, exposing the shiny brown seed covered with a red, fibrous, aromatic aril which is the mace. The kernel inside the seed coat is the nutmeg of commerce. Fruits are produced throughout the year and are picked when the husks split open. The mace is removed from the husks, flattened, and dried. It is used in making pickles, ketchup, and sauces. When the seeds are thoroughly dried, the shells are cracked off, the kernels are removed and sorted, and often treated with lime to prevent damage by insects. Grated nutmeg is used in custards, puddings, and other sweet dishes, also in various beverages. Nutmeg oil is used in medicine, perfumery, and dentifrices, and in the tobacco industry. *See* MAGNOLIALES; SPICE AND FLAVORING. Perry D. Strausbaugh; Earl L. Core

Nutrition

The science of nourishment, including the study of the nutrients that each organism must obtain from its environment to maintain life and health and to reproduce. Although each kind of organism has its distinctive needs, which can be studied separately, a far-reaching biochemical unity in nature has been discovered which gives vastly more coherence to the whole subject. Many nutrients, such as amino acids, minerals, and vitamins, needed by higher organisms may also be needed by the simplest forms of life—single-celled bacteria and protozoa. The recognition of this fact has made possible highly important developments in biochemistry.

Mammals need for their nutrition (aside from water and oxygen) a highly complex mixture of more than 40 chemical substances, including amino acids; carbohydrates; certain lipids; fibers (for preventing constipation and diverticular disease and for slowing the absorption of carbohydrates); a great variety of minerals, including several that are required only in minute amounts, commonly referred to as trace minerals; and vitamins—organic substances of diverse structure that are treated as a group only because as nutrients they are required in relatively small amounts. There are also a few known conditional nutrients—essential components of cellular biochemistry that are not normally nutrients, because they are made internally in adequate amounts, but that become nutrients under stressful conditions of increased need or diminished supply. Vitamin D would not be considered a nutrient if adequate exposure of the skin to sunlight were consistently assured. Conditional nutrients are known in infancy and with injury, disease, genetic defects, and the use of some pharmaceuticals. *See* AMINO ACIDS; CARBOHYDRATE METABOLISM; LIPID METABOLISM; PROTEIN METABOLISM; VITAMIN; VITAMIN D.

Most nutrients were recognized in the nineteenth century, but the vitamins and some trace minerals did not become known as fundamental parts in the machinery of all living things until the early twentieth century. The discovery of vitamins, and some of the trace minerals, originally came about through the recognition of deficiency diseases, such as beriberi, scurvy, pellagra, and rickets, which arise because of specific nutritional lacks and can be cured or prevented by supplying the needed nutrients.

Different species of mammals have some distinctive nutritional needs. Guinea pigs, monkeys, and humans, for example, require an exogenous supply of ascorbic acid (vitamin C) to maintain life and health, whereas most animals, including rats, do not. Ascorbic acid, however, is an essential part of the metabolic machinery of animals that do not need an exogenous supply. Rat tissues, for example, are relatively rich in ascorbic acid; unlike guinea pigs, these animals are genetically endowed with biochemical mechanisms for producing ascorbic acid from carbohydrate. *See* ASCORBIC ACID.

The amounts of nutrients adequate for health are not well known for several reasons, including uncertainties in how to define adequacy, substantial individual variations due to genetic differences, and environmental influences. Even less is known about alterations in nutritional needs caused by various disease conditions, use of pharmaceuticals, and injury. More is known about many aspects of the nutrition of livestock, laboratory animals, and pets than is known about human nutrition, and considerably greater care and expertise are exercised with prized animals than with humans, in part because animal diets can be more easily controlled.

Early workers in human nutrition focused on the minimum amounts needed to prevent or cure acute deficiency diseases, such as scurvy and beriberi. Since that time, the Recommended Dietary Allowances and Adequate Intakes (RDAs and AIs, collectively called Dietary Reference Intakes) in the United States and similar recommendations in other countries include consideration of biochemical criteria of adequacy. They also include approximate adjustments for age, sex, and pregnancy and lactation, along with rough estimates for some other sources of individual variation. However, statistical data needed to adequately assess individual variations are not yet available for any nutrient.

Interests have shifted toward what may be more nearly optimal nutritional intakes based on the amounts needed to promote health (not merely to avoid disease or biochemical deficiency), longevity, and resistance to chronic disorders, including cardiovascular disease, cancer, hypertension, and diabetes. Increasing evidence indicates that nutrients also protect against environmental pollutants and some human birth defects that formerly were not believed to be nutrition-related. Animal studies have long shown that diets adequate for youths and adults may not be adequate for good reproduction, and that deficiencies or imbalances of virtually all nutrients can cause myriad physical and mental defects.

Besides nutrients, which play essential roles in the biochemical machinery of all cells, there are thousands of additional substances in plant foods that are not essential parts of cellular biochemistry but nevertheless are beneficial to mammals. These optional dietary substances, called phytochemicals, were first recognized in the late twentieth century, and their significance is poorly understood. There are also a few known phytochemical-like substances in meat and dairy products. Many of these nonessential substances function as antioxidants. Others stimulate the immune system or production of detoxification enzymes; modulate hormones or gene expression; or act as antibacterial, antifungal, or antiviral agents. Some are pigments that plants produce to limit damage by sunlight. In humans, phytochemicals seem to help protect against heart disease, cancer, hypertension, and many other major and minor disorders. They are widely distributed in probably all plants, including fruits, vegetables, melons, berries, whole grains, beans, nuts and seeds, as well as in plant extracts such as tea, wine, and coffee.

Nutrients, conditional nutrients, and phytochemicals collectively have been called nutraceuticals,

which may be defined as substances considered to be a food, or part of a food, that provide health benefits, including the prevention and treatment of disease. Some organizations promote research and development of nutraceutical supplements, with a potentially separate legal status intermediate between nutrients and pharmaceuticals.

Because of the biochemical unity of living organisms, all whole foods (those that have not undergone processing that greatly alters their nutritional value) contain a wide assortment of the nutrients needed by humans and other mammals. Thus diets derived from a variety of whole foods tend to supply adequate amounts of all nutrients. Guided in part by some innate abilities to select proper kinds and amounts of various whole foods, all existing animal species have at least survived without scientific nutritional knowledge. However, experience shows that suboptimal nutrition prevails in all of nature, and nutrition can always be improved.

Probably all foods contain small amounts of naturally occurring toxins and antinutrients, some of which are deactivated by cooking. The possible roles of these minor components in nutrition are generally only dimly perceived. Their amounts seem mostly insignificant, and some may play useful roles. For example, dietary fiber is known to reduce the availability of some minerals, yet its presence in all plant foods is considered valuable.

For modern humans, the problems of suboptimal nutrition have increased with the advent and extensive consumption of technologically derived, refined foods. In advanced western nations, more than half of the dry weight and energy content of food supplies derives from purified sugars, separated fats, alcohol, and milled grains. These nonwhole "foods" have lost most or all of the nutrients present in the whole foods from which they derive. Although they are manufactured in a wide variety of appealing products that satisfy hunger, their excessive use facilitates various kinds of malnutrition and overconsumption that do not occur readily with whole foods. Modern dietary guidelines and nutrition education focus substantially on partially replacing nonwhole foods with whole grains, legumes, low-fat meats and dairy products, fish, vegetables, fruits, and nuts that retain their natural biochemical unity.

One of the bases for interest in nutrition is the fact that individuals who have differing genetic backgrounds have differing nutritional needs; for this reason, various human ills may arise because the individuals concerned do not get all of the nutrients in amounts compatible with their own distinctive requirements.

Every cell and tissue in the entire body requires continued adequate nutrition in order to perform its functions properly. A multitude of functions, involving the production of specific chemical substances (hormones and prostaglandins, for example) and the regulation of numerous processes, are performed by cells and tissues. Therefore, it is clear that improper nutrition may produce or contribute to almost every type of illness. Nutritional and medical research are yielding important advances in using improved nutrition to prevent, cure, and ameliorate disease and illness. For example, diet composition has been linked to obesity, heart disease, cancer, diabetes, hypertension, osteoporosis, kidney stones, and other prominent diseases of advanced nations. Obesity long has been considered a simple result of excessive consumption of energy and fat, compounded by insufficient exercise. However, this concept is evolving. While Americans have decreased fat consumption in recent decades, and low-fat cookies and other nonwhole foods have proliferated, rates of obesity and diabetes have surged, especially in children. A new concept of obesity focuses more on the proper functioning of appetite and less on the conscious control of energy consumption. Refined carbohydrates, especially in the form of soft drinks, seem to stimulate abnormal appetite and other obesity-related ills. Whole-food sources of fat, such as nuts, avocados, and fish, appear beneficial not only for controlling obesity but also for heart disease, diabetes, and other modern disorders. The benefits of whole-food sources of carbohydrate and fat are likely highly complex, involving the chemical natures of the carbohydrate, fat, and rate of absorption, broad nutrient contents, and possibly flavors and other phytochemicals that act as appetite suppressants. The 2005 Dietary Guidelines for Americans emphasize increased consumption of whole grains, fruits, vegetables, and nuts, and decreased consumption of refined carbohydrates, solid fats, and partially hydrogenated fats. Other advances include the therapeutic use of parenteral nutrition (partial or total intravenous feeding). *See* DISEASE; MALNUTRITION; METABOLIC DISORDERS.

Roger J. Williams; Donald R. Davis

Bibliography. Y. Bao and R. Fenwick (eds.), *Phytochemicals in Health and Disease*, CRC Press, Boca Raton, FL, 2004; Institute of Medicine, *Dietary Reference Intakes*, National Academies Press, Washington, D.C., 6 volumes on nutrients, 1997 to 2004; M. Rechcigl, *Handbook of Naturally Occurring Food Toxicants*, CRC Press, Boca Raton, FL, 1983; M. E. Shils et al., (eds.), *Modern Nutrition in Health and Disease*, 9th ed., Lippincott Williams & Wilkins, Baltimore, 1998; M. R. Werbach, *Nutritional Influences on Illness*, 2d ed., Third Line Press, Tarzana, CA, 1996; W. C. Willett and P. J. Skerrett, *Eat, Drink, and Be Healthy: The Harvard Medical School Guide to Healthy Eating*, Simon & Schuster Source, New York, 2001; R. J. Williams, *Physician's Handbook of Nutritional Science*, C.C. Thomas, 1978.

Nymphaeales

An order of flowering plants that have previously been included in the same larger grouping as Magnoliales, the Magnoliidae. Deoxyribonucleic acid (DNA) sequence studies have demonstrated that Nymphaeales as previously defined contain two

families, Ceratophyllaceae (water hornwort) and Nelumbonaceae (lotus), that are not closely related to the others. Remaining in the order is the family Nymphaeaceae, the waterlilies, from which a small group of tropical plants, Cabombaceae, are split by some scientists. Nymphaeaceae contain nearly 100 species of fresh-water aquatics that are typically found in river and lake systems throughout the world. The ovaries of these plants are filled with mucilage,which mediates pollen tube growth from the stigmas to the ovules, and they have either inappertuate or monosulcate pollen. A spectacular plant is the Amazonian water lily (*Victoria*), which has leaves up to 15 ft (5 m) in diameter. The water lily family has been shown by deoxyribonucleic acid analyses to be one of the oldest lineages of flowering plants and distantly related to all others as well. The family members are relics of the early diversification of the flowering plants. *See* EUMAGNOLIIDS; PLANT KINGDOM; POLLEN.

Mark Chase

Oak

A genus (*Quercus*) of trees, some of which are shrubby, with about 200 species, mainly in the Northern Hemisphere. About 50 species are native in the United States. All oaks have scaly winter buds, usually clustered at the ends of the twigs, and single at the nodes. The fruit is a nut (acorn) surrounded at the base by an involucre, the acorn cup. The pith is star-shaped. The leaves are simple and usually lobed. *See* FAGALES.

Oaks furnish the most important hardwood lumber in the United States. Principal uses are for charcoal, barrels, building construction, flooring, railroad ties, mine timbers, boxes, crates, vehicle parts, ships, agricultural implements, caskets, woodenware, fence posts, piling, and veneer. Oak is also used for pulp and paper products. *See* PAPER.

Eastern oaks. The oaks of the eastern United States are divided into two main categories, the black oak group and the white oak group.

Black oak group. In this group the leaf lobes are bristle-tipped (**Fig. 1**). Acorns ripen in 2 years, and winter buds are pointed.

The northern red oak (*Q. rubra*), which may attain a height of 75 ft (23 m), grows in the eastern half of the United States, except the extreme South and Southeast. It can be recognizedby the large acorns

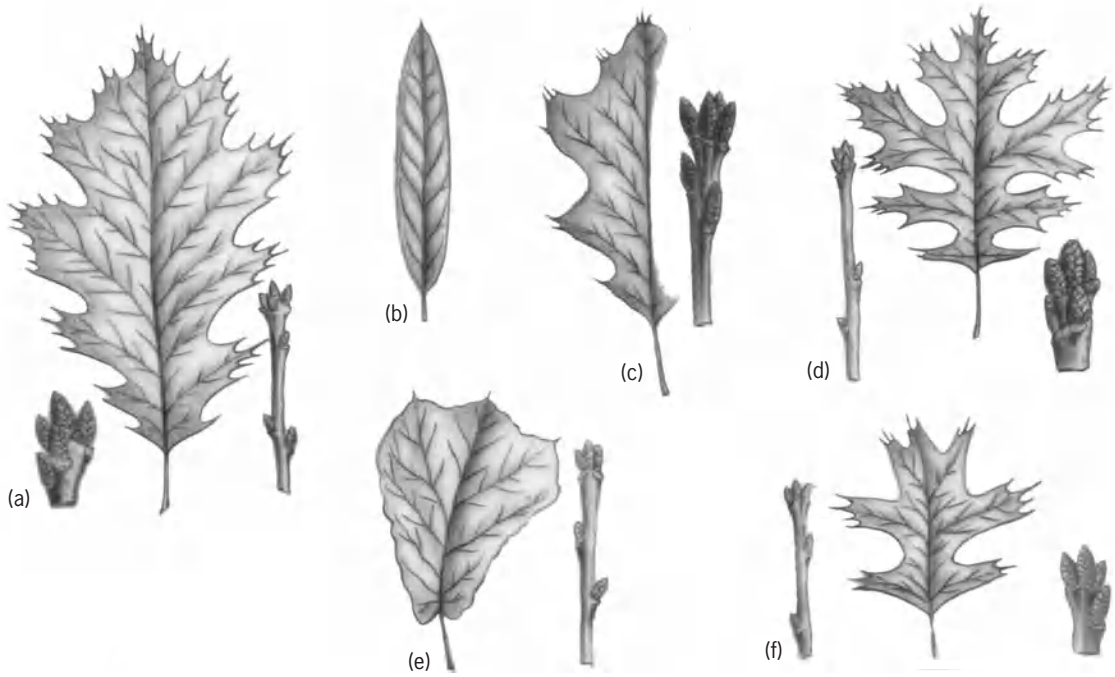

Fig. 1. Eastern oaks (black group), with bristle-tipped leaf lobes. (*a*) Red oak (*Quercus rubra*), terminal buds, leaf, and twig. (*b*) Willow oak (*Q. phellos*), leaf. (*c*) Black oak (*Q. velutina*), half leaf and terminal buds. (*d*) Scarlet oak (*Q. coccinea*), twig, leaf, and terminal buds, (*e*) Blackjack oak (*Q. marilandica*), leaf and twig. (*f*) Pin oak (*Q. palustris*), twig, leaf, and terminal buds.

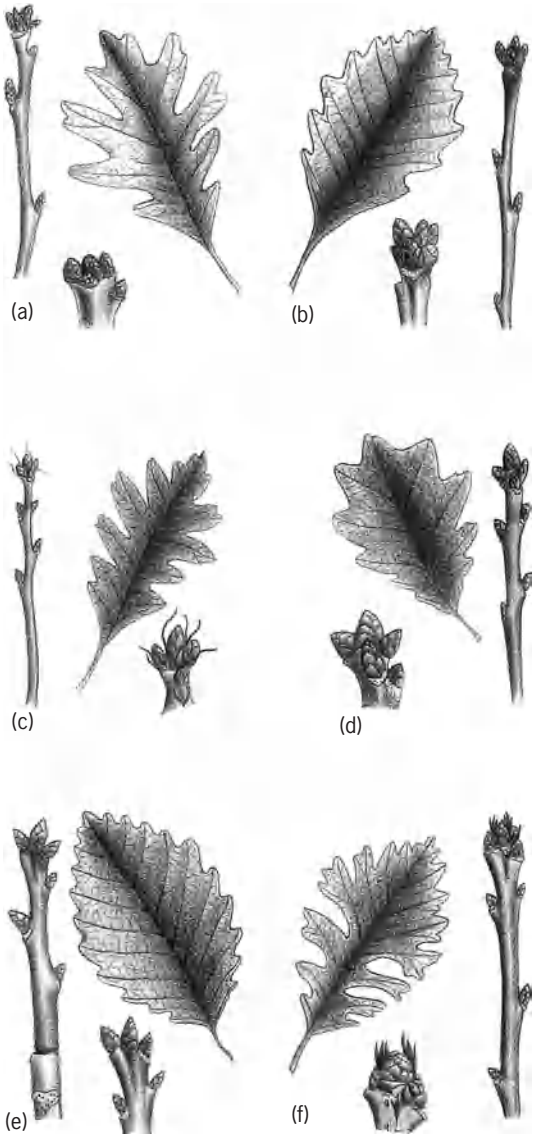

Fig. 2. Eastern white and western oaks showing leaf, twig, and terminal buds for each species. (*a*) White oak (*Quercus alba*). (*b*) Swamp white oak (*Q. bicolor*). (*c*) Turkey oak (*Q. cerris*). (*d*) English oak (*Q. robur*). (*e*) Chestnut oak (*Q. prinus*). (*f*) Bur oak (*Q. macrocarpa*).

and flattish cups and by the red, usually shiny, winter buds. In old trees the bark is comparatively smooth with shallow vertical grooves. It is the most important timber tree of the black oak group and is also a popular shade tree.

Other commercially valuable species in the eastern United States include scarlet oak (*Q. coccinea*), a highly prized ornamental tree with deeply cut leaves which turn a brilliant crimson in the fall; pin oak (*Q. palustris*), with tiny acorns, and small, deeply cleft leaves; black oak (*Q. velutina*), one of the most common species, with large five-sided, pubescent (hairy) winter buds and rough, black bark; southern red oak (*Q. falcata*), with long, often sickle-shaped leaf lobes; and blackjack oak (*Q. marilandica*), with triangular leaves.

Eastern oaks with entire or almost entire leaves are the water oak (*Q. nigra*), laurel oak (*Q. laurifolia*), and willow oak (*Q. phellos*; Fig. 1*b*). The lumber of all these species is similar and usually is classed as red oak lumber.

The live oak (*Q. virginiana*), a medium-sized tree of the South Atlantic and Gulf Coast regions, has evergreen leaves.

White oak group. This group has rounded leaf lobes, acorns which ripen in 1 year, and winter buds which are usually rounded (**Fig. 2**).

White oak (*Q. alba*) furnishes the most valuable hardwood lumber of all Eastern trees. In this same group is the bur oak (*Q. macrocarpa*), a large tree of the eastern United States and adjacent Canada. Its acorns are large and edible.

The chestnut oak subgroup is characterized by leaves with numerous small rounded teeth. This subgroup is represented chiefly by the chestnut oak (*Q. prinus*), of the Appalachian Mountain and Ohio Valley regions; the swamp chestnut oak (*Q. michauxi*); and the swamp white oak (*Q. bicolor*).

Western oaks. In the western United States the native oaks do not have the same high commercial value as timber trees, their place being taken by the valuable western conifers. However, species of importance are the California black oak (*Q. kelloggii*), which grows in Oregon and California; the California live oak (*Q. agrifolia*), with persistent leaves, found in California and Baja California; and the California white oak (*Q. lobata*).

The English oak (*Q. robur*) and its varieties are cultivated in the United States, and have leaves of the white oak type, as well as the turkey oak (*Q. cerris*), which has shallowly lobed leaves (Fig. 2*c* and *d*). *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Oasis

An isolated fertile area, usually limited in extent and surrounded by desert. The term was initially applied to small areas in Africa and Asia typically supporting trees and cultivated crops with a water supply from springs and from seepage of water originating at some distance. However, the term has been expanded to include areas receiving moisture from intermittent streams or artificial irrigation systems. Thus the floodplains of the Nile and Colorado rivers can be considered vast oases, as can arid areas irrigated by humans. *See* DESERT.

River valleys. Other well-known desert rivers include the Tigris-Euphrates in the Middle East, Hwang-Ho (Yellow River) in China, and the Amu Darya and Syr Darya in Central Asia. All of these support luxuriant, although sometimes intermittent, vegetation along their courses. Even the most extreme deserts, such as the Namib along the western coast of southern Africa and the Atacama-Peruvian Desert of western South America, have through-flowing streams rising in adjacent mountain areas. In a broad sense all of these are oases.

The very highly developed Salt River Valley in Arizona, in which the city of Phoenix is located, could

be considered an immense oasis. Its cultivated crops, including palms, fruit trees, and vegetables, have much in common with those of the oases of the Sahara. The area around Indio, California, might be considered even more like African oases because a wide variety of date palms that were imported from the Sahara produce the highest-quality dates.

On a much more modest scale are the palm oases of the Colorado and Sonoran deserts of the southwestern United States. These usually small areas, located in canyons and at the foot of mountains, are characterized by the native palm *Washington filifera*. One such oasis has been developed into Palm Springs, the well-known California winter resort area.

Desert gardens. The term oasis usually brings to mind the widely scattered desert gardens of the Sahara and Arabian deserts. It is commonly associated with palm trees, lush vegetation, Arabs, and camels. These oases have a character distinct from that of irrigated river valleys. They may be large and support a population of more than 500,000 such as Damascus in Syria, or they may be small, only an acre or less in extent. Formerly, most of these oases were reached only by camel caravan, but now many are accessible by air, railroad, or paved highway. Some receive their moisture from artesian sources, others from intermittent streams with either surface or underground flow. Some are watered by deep wells or by horizontal wells called foggaras by the Arabs, kánats or ganats by the people of southeastern Asia, and karez by the Mongolians.

Water-collection systems. A foggara is made by digging a ditch into the desert on a gentle slope starting at the edge of a basin or dry stream bed. When the ditch becomes too deep to be maintained as an open course, it becomes a tunnel which is extended outward, with wells constructed to the surface at intervals. These chains of wells connected by the tunnel reach into moist sand. Water collects in the tunnel and flows to the basin, where it may be used for irrigation or for human or animal consumption.

Various methods are used to take water from wells. In the shallow wells of the Gobi Desert, water can be reached with a short rope to which a stick and a bag are attached. The well man pulls up the rope until he can reach the stick, which he uses to lift the bag, and then pours water into a trough for animals or into wooden casks for human use.

Well sweeps are common in the Sahara. These consist of a long pole swung between upright posts near the well. The butt of the pole is weighted with rocks or other heavy material. At the well end of the pole is a rope and bucket with which water can be lifted from the well.

The use of pulleys is common at most oases. The rope and pulley may be operated by donkey, ox, or camel power. The water bucket or bag has a long spout that folds back when the bag is lowered and is straightened out by a short rope at the well curb so that water flows into ditches or storage ponds. Modern pumps operated by windmills or engines are less common.

Climate and soil. Oases are restricted to climatic regions where precipitation is insufficient to support crop production. Such regions may be classified as extremely arid (annual rainfall less than 2 in. or 50 mm), arid (annual rainfall less than 10 in. or 250 mm), and semiarid (rainfall less than 20 in. or 500 mm). Many African and Asian oases are in extremely arid areas. Most oases are found in warm climates. Some authorities restrict the use of the term to regions where date palms grow and produce. This means long frost-free growing seasons.

Oasis soils are weakly developed, high in organic matter but often saline, and have been strongly affected by human occupation.

Economics and agricultural practices. The date palm (*Phoenix dactylifera*) is the tree of greatest economic importance. The living tree provides shade, the fruit is eaten by the Arabs, and the cracked seeds are fed to the camels. In winter dead leaves are cut out and the dead palm leaves used for fuel for cooking fires. Trunks of old trees are used for houses and serve as simple bridges across irrigation ditches. The fiber at the base of the leaves is used to make rope and a course cloth.

Date palms are planted about a rod apart, and the spaces between are planted with cereals or garden vegetables. The plots are usually small, often not more than a few square feet. Each property may be surrounded by mud walls, which often have broken glass or cacti on top to discourage intruders. Under the date palm canopy may be found apricot, fig, mulberry, peach, pear, and orange trees together with the pomegranate and table grape. Vegetables include artichokes, beans, carrots, melons, peas, potatoes, squashes, and radishes. Cereals include wheat, barley, and rye, and in some oases alfalfa is produced. Various ornamental plants are also grown. Occasionally the vine-covered tree canopy is so dense that little sunlight reaches the ground.

Habitation. Oasis habitations are located at the edge of the gardens, on slopes above the flat plain where topography permits. The flat-roofed houses are made from dirt and may be whitewashed. Most have walled courtyards. In the Sahara there is usually a central square with a mosque. In the larger oases there are shops around the square, but many of the smaller oases depend for their needs upon itinerant taders, who congregate on weekly market days.

Natural vegetation. Among the native plants restricted to waste areas and abandoned plots are a wide variety of phreatophytes, including tamarisk (*Tamarix* sp.), cattail (*Typha latifolia*), sedges, and rushes; farther away many halophytes are found where drainage is poor; and beyond this fringe, desert vegetation takes over.

Geographic distribution. Some of the better-known oases are listed below.

Talfilalet, Morocco, has an area of 200 mi^2 (518 km^2) and is reputed to be the largest oasis in the Sahara.

Algeria has a number of well-known oases. Biskra in northeastern Algeria is a dateshipping center and health resort; El Goléa in the central region is noted

for its gardens; Ghardai;ˆa has a mosque with a minaret 300 ft (91 m) high; Laghouat is noted for its luxuriant vegetation, and the vine-covered palms make a nearly sunproof canopy; the Ouargla Oasis in east central Algeria is an ancient rambling town reputed to have had 500,000 date palms in earlier times; the Touggourt area, including the Souf and Oued Rhir oases, is noted for its Deglet Nur dates; the Tidikelt and Touat oasis region in central Algeria includes the Aoulef, In-Salah, and Tit oases, all of which are important date producers.

Tozeur Oasis, in Tunisia, is 110 mi (177 km) west of Gabès and is a producer of dates, olives, and vegetables.

Kufra Oasis, in Libya, is a camel-breeding center.

Dakhla is the most populous (21,000) oasis in Egypt, with 40 mi^2 (104 km^2) in cultivation; the Kharga Oasis in south-central Egypt is known as the Great Oasis; the Siwa Oasis in the Libyan Desert of western Egypt is 100 ft (30 m) below sea level.

In Saudi Arabia, Hasa Oasis (Hofuf is the principal city), located along the Persian Gulf and extending some distance inland, is the largest oasis in the Arabian Peninsula; the Qatif Oasis, which is also on the Persian Gulf, is smaller but is an important agricultural center.

Damascus, Syria, and the surrounding area with over 600,000 people may be the most populous oasis in the Afro-Asian area. William G. McGinnies

Bibliography. A. D. Abrahams and A. J. Parsons (eds.), *Geomorphology of Desert Environments*, 1993; S. Aritt, *The Living Earth Book of Deserts*, 1993; J. Flegg, *The Deserts*, 1993; W. G. McGinnies, B. J. Goldman, and P. Paylore, *Deserts of the World: An Appraisal of Research into Their Physical and Biological Environments*, 1968.

Oats

An agricultural crop grown for its grain and straw in most countries of the temperate zones of the world. In the major oat-growing states of the midwestern United States (Iowa, North Dakota, South Dakota, Minnesota, and Wisconsin) the crop is raised for grain, whereas in the Southern states (Texas, Oklahoma, and Georgia) it is used for pasture or a combination of pasture and grain. About 90% of the annual oat grain production is used for animal feeds, and about 10% is processed into food for humans, for example, oatmeal and other cereal products. *See* ANIMAL FEEDS; CEREAL.

Among the cereal grains grown in the United States, oats are exceeded in importance only by corn, wheat, and sorghum. *See* CORN; SORGHUM; WHEAT.

In general, oats are a cool-season crop which requires a moist climate. They grow well on both light and heavy soils if sufficient moisture and fertility nutrients are available. In the Central and Northern states, oats are spring-sown; but in the Southern states they are fall-sown. In the Corn Belt they are grown in crop rotation with corn, soybeans, and forages.

Origin and description. Oats probably became an agricultural crop about 2000 years ago, most likely starting in the Mediterranean area. The fifteen species of oats in the genus *Avena* are divided into three groups on the basis of chromosome number: 14, 28, or 42. Of the seven species that make up the 14-chromosome group, only *A. strigosa* is grown commercially on small agricultural acreage in Portugal and Brazil. The two 42-chromosome species, *A. fatua* and *A. sterilis*, grow in natural stands in the eastern Mediterranean area and provide wild pastures for sheep and goats. Seeds of both wild species shatter at maturity to ensure annual reseeding.

Crosses among the 28-chromosome species and among the 42-chromosome species are easy to make and the hybrids are fully fertile, but only certain 14-chromosome species can be intercrossed. Crosses among species with different chromosome numbers can be made also, but usually the hybrids are sterile or only partly fertile. *See* BREEDING (PLANT); GENETICS; REPRODUCTION (PLANT).

Within the 42-chromosome cultivated species there is wide variation among varieties for all plant traits. Oats belong to the Graminae (grass) family; thus the oat plant forms a crown at the soil surface from which a fibrous root system penetrates the soil. Most of the roots are concentrated in the upper foot of soil, but some grow to a depth of 5 ft (1.5 m). Under thick seeding only one or two culms develop, but when plants are spaced, 10–30 culms may develop. Culms usually grow 2–5 ft (0.6–1.5 m) tall, and they are terminated with inflorescences called panicles. Each panicle usually bears 10–75 spikelets on its numerous branches (**Fig. 1**). A spikelet is enclosed by two papery glumes and bears two or three florets, each with an ovary, two stigmas, and three anthers enclosed in a lemma and palea (**Fig. 2**). The flowers are normally self-pollinated but 1–2% outcrossing may occur. The stem has 7–9 nodes, and a leaf grows at each node. Internodes of the stems are hollow. *See* CYPERALES; FLOWER; GRASS CROPS; INFLORESCENCE.

In most varieties the lemma and palea adhere to the oat seed after threshing. A trait used to determine market grade of oats is the color of the lemma, which may be white, yellow, gray, brown, red, or black. The

Fig. 1. Oat panicles with many branches and spikelets.

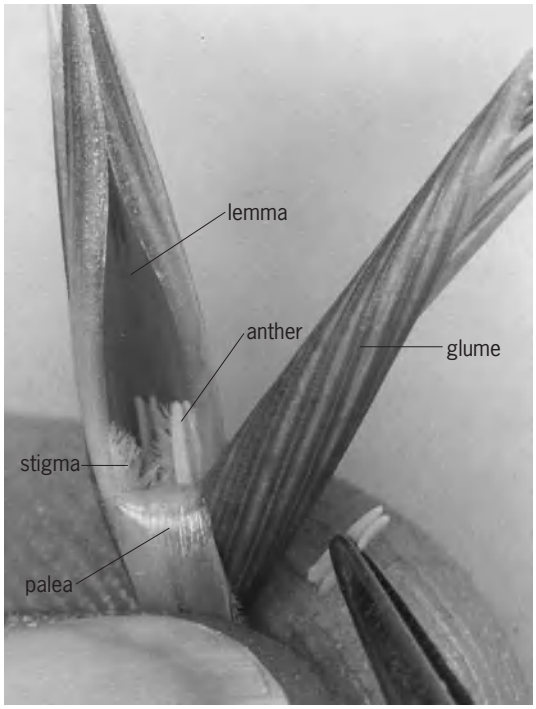

Fig. 2. An oat flower; two oat grains are shown at the lower right.

major trait that distinguishes wild from cultivated oats is seed shattering. In cultivated species the seed attachment is persistent, and it can be separated from the panicle only by threshing.

Varieties. The world collection of oats, maintained by the U.S. Department of Agriculture, contains more than 14,000 lines of 42-chromosome types. These represent lines from wild species and from varieties produced at breeding stations. The collection represents a vast range of genetic types that can be used for varietal improvement. After 1960, breeders put special emphasis on improving lodging and disease resistance, grain quality, and yield of new oat varieties. These varieties have been developed by crossing strains from the world collection to obtain better combinations of genetic traits. A variety called "multiline" was developed to control rusts. While each multiline variety is uniform for all agronomic traits, it contains several genes for resistance to a single disease.

Cultural practices. In the Corn Belt, oats usually follow corn or soybeans in crop rotation. They are used as a companion crop for forage seedings; that is, alfalfa and oats are planted simultaneously, and the oats are harvested for grain in the first year and alfalfa for hay in subsequent years. Spring oats are seeded in March or April at a rate of 2–3 bu/acre (0.2–0.3 m^3/hectare) on disked or plowed land. Winter oats are sown in October or November. Mature oats are harvested with a combine, either direct or after being windrowed to dry. For safe storage oat grain should contain no more than 13% moisture. Straw of harvested oats is either worked into the soil for humus or baled and stored for use as bedding for livestock. *See* AGRICULTURAL SOIL AND CROP PRACTICES; HUMUS; LEGUME FORAGES.

Uses. Oat grain usually contains 10–16% protein, which makes it especially good for rations for young livestock and for human food. Some strains of *A. sterilis* have up to 30% protein. Dehulled oat seeds are used extensively for making breakfast cereals, such as porridge made from rolled oats. In general, oat grain contains adequate quantities of minerals and B vitamins for normal diet. The fat content is low. Oat hulls are used to make furfural, an important chemical in nylon manufacturing. Kenneth J. Frey

Diseases. Under certain conditions, any major oat disease can reduce yield and quality of oat grain over large areas. Individual oat fields may be very severely damaged or, in some cases, destroyed.

Rusts. Crown rust (caused by *Puccinia coronata*) is probably the most destructive oat disease. Losses vary greatly from year to year, ranging from negligible to more than 30% of the potential crop over large areas. The uredial or repeating stage (the one that does the damage; **Fig. 3**) first appears on the oat leaves several weeks before ripening of the grain. The bright orange, powdery uredial pustules produce spores that are blown to other plants, in which they produce a new generation every 8–10 days. In warmer areas such as the Gulf Coast of the United States, the pathogen maintains itself in the ureidal stage throughout the winter on winter-grown oats. Spores then blow north as spring-sown oats begin to grow in the North. An alternative, "complete" life cycle occurs in climates having cold winters where buckthorn (*Rhamnus cathartica*) is present. As the oats mature, small black lesions called telia appear. Spores in the telia survive cold northern winters and germinate in spring to form another spore type that infects young buckthorn leaves, but not oats. Still another spore type, produced on the buckthorn (**Fig. 4**), can infect oats but not buckthorn. Blown to oats, they produce uredia to complete the life cycle. Stem rust (caused by *P. graminis*), which also can cause severe losses, is distinguished from crown rust by its brick-red uredia, which are more common on the stem and leaf sheaths (**Fig. 5**), and by the fact that its alternate host is the common barberry (*Berberis*).

Fig. 3. Susceptible oats (foreground) killed by crown rust, and oats resistant (background) to crown rust.

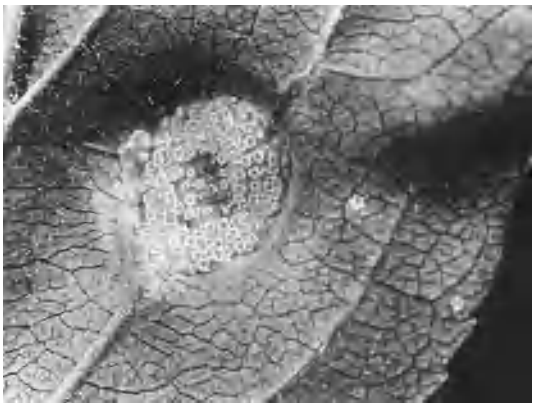

Fig. 4. Crown rust on leaf of alternate host buckthorn (*Rhamnus cathartica*).

Fig. 5. Stem rust on oat stems. (*Photograph by J. A. Browning*)

Both the crown- and stem-rust fungi consist of many pathogenic strains or races that differ only in their ability to attack different oat varieties. Starting in the 1920s, single genes conferring high resistance were used to develop resistant oat varieties; however, new races of fungus capable of parasitizing the new varieties then developed. The use of other resistance genes in newer varieties was followed by the appearance of still more new rust races. Several approaches are used to break this cycle: (1) Oat varieties that carry not one but several genes for resistance have been developed. (2) Varieties have been developed which consist of a mixture of visibly similar lines, each of which contains a different gene for resistance; this approaches the diversity that protects natural populations. (3) General resistance controlled by large numbers of minor-effect genes has been used; this is difficult in breeding programs but may be longer-lasting than other types of resistance. (4) Tolerance has also been used, tolerance being the ability of a variety that appears susceptible to withstand infection without serious loss of yield or quality. (5) Geographical deployment of genes has been used; this calls for the use of resistance genes in the South different from those used in oat varieties grown in the North—the rust fungi moving north in the spring would confront oat varieties having different effective resistance genes than the varieties on which the spores had been produced.

Eradication of the alternate hosts (buckthorn and barberry) will reduce early spring inoculum and slow the appearance of new races, because genetic recombination of the fungus occurs on these hosts. Fungicides are effective but not economical.

Smut. Unlike the rusts, smut (*Ustilago* sp.) can cause serious losses in drier areas. In certain years, over half the panicles in individual fields have been smutted, and losses over larger areas have averaged as much as 20%. Smut appears as dark brown to black spore masses in place of the seed (**Fig. 6**). The spores are blown to young seeds on healthy plants and germinate, and the fungus remains dormant until the seed is planted. The fungus infects the plant as it emerges from the seed and eventually produces spore masses in place of the seed to complete the cycle. The smut fungus, like the rust fungi, consists of pathogenic strains. Smut can be controlled by use of resistant varieties or chemical seed treatment.

Other fungi. A number of other fungi, including species of *Septoria*, *Helminthosporium*, and *Fusarium*, occasionally cause economic losses. Some varieties have a degree of genetic resistance, and good cultural practices are recommended.

Oat roots and crowns also are attacked by a number of fungi, which may destroy very young plants and damage roots of older plants. The extent of such losses is not clear, but plants grown under optimum conditions of temperature, water supply, and fertility are usually not seriously affected. Seed treatment and use of moderately resistant varieties may also be beneficial.

Bacteria and viruses. Bacterial blight (caused by *Pseudomonas* sp.) usually appears early in spring as halo-like lesions on the leaves. It often looks serious, but

Fig. 6. Smutted (left) and healthy oats. (*Photograph by J. A. Browning*)

significant reduction in yield is rare. Yellow dwarf, however, caused by the aphid-transmitted barley yellow dwarf virus, rivals the rusts in economic importance. Infected plants are dwarfed, turn yellow or reddish, mature early, and are partially sterile. Controlling the disease by controlling its aphid vectors is generally not practical. A high degree of resistance to yellow dwarf is not known, but the use of varieties having moderate resistance will usually give satisfactory control.

Gray speck. Several nonparasitic conditions are regarded as diseases of oats. The most common is gray speck, caused by manganese deficiency. It can be controlled by the application of manganese salts to the soil or to the foliage. The use of more tolerant varieties will generally control gray speck. *See* PLANT PATHOLOGY. M. D. Simons

Processing. The milling of oats is less complex than wheat milling and has many similarities to rice or barley milling operations because there is limited fractionation of the kernel. The oat grain is covered with a coarse, adhering hull (about one-quarter of the kernel by weight) which must be removed prior to production of ingredients or consumer foods. Oats as received at the mill house are termed green oats and must be cleaned to remove foreign seeds and trash. Cleaning requires the use of air aspiration to remove low-density materials, reciprocating sizing sieves to segregate oat grains by varying slot width, and continuous disk separators. Disk separators are a series of specialized horizontal rotating disks, each with indentations which are designed to pick up seeds smaller than the oat kernel, leaving clean oats in stream. Clean, sound oats are heated slowly prior to hull removal. The green oats have active lipolytic enzymes (lipases) which will catalyze hydrolysis of triglycerides and yield free fatty acids. This hydrolytic process results in the development of rancid off-flavors which greatly reduces acceptability and the shelf-life stability of processed oat products. Conventional milling of green oats will result in complete hydrolysis of the lipid by lipases within 3 days.

Preheating and dehulling. The heating, drying, or roasting procedures inactivate lipases, facilitate hull removal, and impart a distinctive roasted flavor to the oat product. Oat drying is commonly accomplished by heating the grain to approximately 200°F (93°C) for about 1 h. Steam-jacketed open-pan dryers are used in series in which the grain is surface heated, mixed, and conveyed across a series of pans to provide a uniform controlled heat process. Such heating reduces grain moisture content by 3–4%, inactivates enzymes, improves hull separation, and develops a roasted flavor. Alternately, direct steam heating (atmospheric pressure) for 2–3 min will reduce enzyme activity. Excessive steaming can, however, promote undesirable oxidative rancidity.

Roasted oats are air-cooled and size-graded prior to dehulling. Classical dehullers used for oats consist of precisely adjusted rollers; however, modern dehulling systems utilize an impact dehuller which imparts centrifugal acceleration to the sized grain and hull shattering on a rubber impact ring. In the dehuller, grain enters the top center of the chamber, is accelerated on a high-speed finned rotor, and is discharged to impact on the peripheral rubber ring. Higher yields of whole groats are obtained with less input energy with these modern impact dehullers. Dehulling conditions may be precisely controlled to yield optimum hull separation. Under the most sophisticated system, impact dehullers can be adjusted to produce good-quality kernels without prior heat treatment.

The products of the dehuller are primarily the whole kernel or groat and the fiber hull which are readily separated by air aspiration. The low-density oat hulls possess particularly high levels of fiber and pentosans which are suitable feedstock for industrial production of furfural through high-temperature acid hydrolysis and dehydration. Whole, cleaned oat groats are not frequently available in the commercial food market. Selected large-sized groats are excellent for puffing into ready-to-eat cereals.

Groat flaking. Oat cutting and flaking procedures provide more extensive utilization of groats in the form of rolled oats. Steel cutting of the groat is accomplished by passage between a rotating pick-up drum and knives to yield two to four uniform pieces and a small amount of flour. These sieved pieces are then steamed and flaked with heavy rollers while moist, hot, and pliable to provide thin rolled oat flakes. The flakes are subsequently air-dried to assure shelf stability. Cooking time of rolled oats used as a hot breakfast cereal is dependent on the extent of presteaming and the final thickness of the flake. Thin flaked oats acquire an increased surface area (relative to weight) which facilitates hydration and reduces cooking time. Steamed whole groats may be flaked to a greater thickness to produce a so-called old-fashioned rolled oat requiring about 5 min cooking time. Quick-cooking rolled oats are prepared from steamed, steel-cut groats and are rolled to a thin flake (50% thinner than whole groat flakes), and require about 1 min boiling to prepare a palatable product. Instant oat cereals, requiring only the addition of boiling water and mix-in-the-bowl stirring to dissolve, are prepared with very thin rolled flakes and formulated with guar gums. The thin rolled cooked oat flakes rapidly hydrate, and the cereal obtains its viscous cooked consistency from guar gum, a highly branched galactomannan. *See* GUM.

Oat flour is obtained from further reduction and sieving or hammer milling of the whole groat or flaked product. This high-protein flour is frequently used in the formulations of ready-to-eat cereals and many prepared baby foods. Milled oat flour has appreciable antioxidant activity, similar to the phenolic antioxidants propyl gallate and butylated hydroxytoluene (BHT) and can be used as an ingredient to extend shelf stability. Composite flours blended from oat flour and other cereals providing high protein content and extended shelf life have been proposed as suitable for world feeding programs. *See* FOOD MANUFACTURING. Mark A. Uebersax

Bibliography. H. J. Brounlee and F. L. Genderson, Oats and oat products: Culture, botany, seed

structure, milling, composition, and uses, *Cereal Chem.*, 15:257–272, 1938; R. C. Hoseney, *Principles of Cereal Science and Technology*, 2d ed., 1994; N. L. Kent and A. D. Evers, *Technology of Cereals: An Introduction for Students of Food and Science Agriculture*, 4th ed., 1994; K. Lorenz and K. Kulp (eds.), *Handbook of Cereal Science and Technology*, 1990; J. H. Spitz, *Crop Physiology and Cereal Breeding*, 1991; P. J. Wood (ed.), *Oat Bran*, 1993.

Obesity

The presence of excess body fat. The great prevalence of this condition, its severe consequences for physical and mental health, and the difficulty of treating it make the prevention of obesity a major public health priority.

Definition. The concept of excess body fat implies comparison to an ideal level of body fat associated with optimal health and longevity. Such an ideal has not been defined, however, primarily because accurate and convenient measures of body fat content are unavailable. Instead, obesity is most often defined in terms of body weight relative to height, since both height and weight are easily measured. Obesity is considered to begin at a weight-for-height that is 20–30% above desirable weight, with this desirable weight taken as the midpoint of ranges of weight associated with the greatest longevity in studies of life-insured individuals. *See* ADIPOSE TISSUE.

In population surveys, obesity is defined as a body weight that meets or exceeds the 85th percentile of the Body Mass Index (BMI), an index of weight-for-height that correlates well with body fat content. The BMI is calculated as weight in kilograms divided by height in meters squared for men and to the power 1.5 for women. By using the range of BMIs for adults aged 20–29 as a standard, 34 million (about 25%) of all adults in the United States aged 20 to 74 are obese, and severe obesity—at or above the 95th percentile of the BMI—affects nearly 10% of the adult population.

The prevalence of obesity increases with age, is higher in women than men, and is highest among the poor and minority groups. Severe obesity has become more common in the United States since the 1970s, but trends in overall prevalence are uncertain; inconsistencies in definitions of overweight used in national surveys have made valid comparisons difficult. *See* AGING.

Risks. Obesity increases the likelihood of high blood cholesterol, high blood pressure, and diabetes, and therefore of the diseases for which such conditions are risk factors—coronary heart disease, stroke, and kidney disease. It also increases the likelihood of gallbladder disease and cancers of the breast and uterus. Thus, obesity increases overall mortality rates, and it does so in proportion to the degree and duration of overweight. Individuals who become obese at the earliest ages are at highest risk of premature mortality. Distribution of excess fat to the upper body rather than the lower body may also increase risk.

The precise levels of overweight at which health risks begin to increase have not been established. Studies designed to define such ranges have been limited by inadequate numbers of subjects, study periods that are too short, or failure to control for the consequences of cigarette smoking, chronic disease risk factors, weight loss due to disease, or duration of overweight. Despite these limitations, it appears that weight-for-height ranges associated with the longest survivals are significantly below average weights in the population.

Psychosocial consequences of obesity are also of concern. Obese individuals are subjected to negative social attitudes and to discrimination in employment, housing, education, and health care. This is perhaps because of mistaken beliefs that people become and remain overweight because of poor self-control.

Causes. The causes of most cases of obesity are poorly understood. At the simplest level, obesity results from an excess of energy (caloric) intake over expenditure, but this statement does not explain why some individuals can eat as much as they like without gaining weight while others remain overweight despite constant dieting. Studies of genetically obese animals and those with damage to the part of the brain called the hypothalamus suggest that individuals may balance body weight around a "setpoint" that is maintained—without conscious control—by variations in metabolic rate in response to caloric intake. When too many calories are consumed, the metabolic rate increases and burns off the excess energy as heat; when calories are restricted (as in dieting), the metabolic rate decreases to compensate. Although the existence of setpoints is unproven, this theory offers a testable hypothesis to explain why most obese persons find weight loss so difficult to maintain. The factors that might establish setpoint levels are undefined but are most likely to include inherited metabolic characteristics, level of caloric intake, and level of physical activity, alone or in combination. *See* METABOLIC DISORDERS.

Variations in the prevalence of obesity among population groups suggest a genetic basis for the condition. To eliminate the possibility that the differences are due to cultural or other environmental factors, investigators have compared body weights in biological and adopted children and in identical twins. Although the results of these studies have not always been consistent, they generally indicate that the body weights of children resemble those of the biological parents more than the adoptive parents and that the weights of identical twins are more similar than those of other siblings regardless of who raised them.

The complexity of body-weight regulatory mechanisms suggests that obesity is not due to a single cause but, like other chronic diseases, is multifactorial in origin. Specific inherited differences that might influence setpoints include differences in nearly every anatomic, neurologic, and biochemical factor known to affect food intake and utilization,

energy metabolism, and energy expenditure. *See* ENDOCRINE MECHANISMS; ENERGY METABOLISM.

Although excessive caloric intake might seem to be an obvious cause of obesity, it has not been possible to demonstrate that overweight individuals consume more calories (relative to body weight) than people of normal weight. Most surveys indicate that obese individuals consume fewer calories for their weight than lean individuals, an observation at least partially explained by the fact that it takes fewer calories to maintain fatty tissue than lean muscle tissue. Suggestions that the obese display impaired control of hunger and satiety, respond inappropriately to external hunger signals (for example, taste, time of day), or demonstrate abnormal eating patterns do not seem to be generalizable.

Obese individuals from earliest infancy to old age display reduced energy expenditures. Population studies also reveal an association between sedentary life style and obesity, but it is uncertain from these observations whether reduced activity is a cause or a consequence of being overweight. Evidence that moderate exercise exerts a regulatory action on body weight beyond its immediate effect on caloric output is incomplete. Exercise raises the metabolic rate and increases the proportion of lean body tissue, and clinical studies demonstrate that highly active people are leaner even if they consume more calories. Spontaneous physical activity—fidgeting—has been identified as a potentially significant factor in obesity prevention. Although more information is needed about the relationship between caloric intake, caloric expenditure, and body weight, it seems evident that inadequate caloric expenditure deserves at least as much attention as excessive caloric intake in obesity causation.

Treatment. Because the causes of obesity are incompletely understood, it is difficult to formulate effective treatment strategies. Weight loss requires a decrease in caloric intake, an increase in caloric expenditure, or both. Typically, treatments have emphasized reductions in caloric intake by diet, induction of more restrictive eating patterns by behavior modification, control of hunger by drugs, or limitations on food intake by surgery. None of these methods is free of cost or risk, and none has been outstandingly successful in inducing long-term weight loss; the great majority of obese people who attempt to lose weight eventually gain it back. The costs, risks, and ineffectiveness of traditional weight-loss methods explains why increased physical activity is becoming a major focus of obesity prevention and treatment. Preliminary studies suggest that programs combining diet and exercise help obese individuals lose more weight and maintain losses longer than either program does separately. A rational approach to prevention is the adoption of a life style that includes considerable physical activity along with a diet that contains adequate but not excessive calories. *See* FOOD; NUTRITION. Marion Nestle

Bibliography. K. D. Brownell and J. P. Foreyt (eds.), *Handbook of Eating Disorders: Physiology, Psychology, and Treatment of Obesity, Anorexia, and Bulimia*, 1986; J. E. Manson et al., Body weight and longevity: A reassessment, *J.A.M.A.*, 257:353–358, 1987; National Institutes of Health Consensus Development Panel, Consensus conference statement: Health implications of obesity, *Ann. Int. Med.*, 103: 1073–1077, 1985; E. Ravussin et al., Reduced rate of energy expenditure as a risk factor for body-weight gain, *N. Engl. J. Med.*, 318:467–472, 1988; A. J. Stunkard and E. Stellar, *Eating and Its Disorders*, Association for Research in Nervous and Mental Disease, vol. 62, 1984; T. A. Wadden and A. J. Stunkard, Social and psychological consequences of obesity, *Ann. Int. Med.*, 103:1062–1067, 1985.

Object-oriented programming

A computer-programming methodology that focuses on data items rather than processes. Traditional software development models assume a top-down approach. A functional description of a system is produced and then refined until a running implementation is achieved. The refinement process breaks each system function down into subfunctions which combine to implement the higher-level function. The emphasis of the approach is on the transformation of inputs to outputs. Data structures (and file structures) are proposed and evaluated based on how well they support the functional models.

The object-oriented approach focuses first on the data items (entities, objects) that are being manipulated. The emphasis is on characterizing the data items as active entities which can perform operations on and for themselves. It then describes how system behavior is implemented through the interaction of the data items.

Basic concepts. The essence of the object-oriented approach is the use of abstract data types, polymorphism, and reuse through inheritance.

Abstract data types define the active data items described above. A traditional data type in a programming language describes only the structure of a data item. An abstract data type also describes operations that may be requested of the data item. It is the ability to associate operations with data items that makes them active. The abstract data type makes operations available without revealing the details of how the operations are implemented, preventing programmers from becoming dependent on implementation details. The definition of an operation is considered a contract between the implementor of the abstract data type and the user of the abstract data type. The implementor is free to perform the operation in any appropriate manner as long as the operation fulfills its contract. Object-oriented programming languages give abstract data types the name class. *See* ABSTRACT DATA TYPE.

Polymorphism in the object-oriented approach refers to the ability of a programmer to treat many different types of objects in a uniform manner by invoking the same operation on each object. Because the objects are instances of abstract data types, they may implement the operation differently as long as

they fulfill the agreement in their common contract. The process of finding the implementation of an operation for an object is referred to as dynamic binding (various languages implement dynamic binding in different ways).

A new abstract data type (class) can be created in object-oriented programming simply by stating how the new type differs from some existing type. A feature that is not described as different will be shared by the two types, constituting reuse through inheritance. Inheritance is useful because it replaces the practice of copying an entire abstract data type in order to change a single feature. Copying abstract data types to make small changes causes problems because it unnecessarily duplicates program code and causes an increase in maintenance effort. Reuse through inheritance changes only the features that are different and remembers the commonality between types.

The elements of object-oriented programming have much in common with automobiles. Drivers operate automobiles without detailed knowledge of their internal implementation. They do not need to know the details because they have an agreement with the automobile manufacturers on how cars should behave. A second commonality is that most drivers can operate most cars because all cars have essentially the same interface, even though different kinds of cars implement the operations differently. This characteristic is polymorphism. The third commonality is found by noticing that different kinds of automobiles often are not greatly different from one another. An upscale car may simply be a version of a basic car with more expensive options and a different body style. Because there is so much in common between the them, the upscale model may be described simply by stating how it is different from the basic model. A feature that is not described as different can be assumed to be the same. This characteristic is reuse through inheritance.

Classes as abstract data types. In the object-oriented approach, a class is used to define an abstract data type, and the operations of the type are referred to as methods. An instance of a class is termed an object instance or simply an object. To invoke an operation on an object instance, the programmer sends a message to the object.

Example (1) illustrates an Employee class with a simple set of operations.

```
CLASS Employee =
   METHODS
      FUNCTION name : string;
      FUNCTION Position : string;
      . . .
END;
```
(1)

This declaration indicates that a programmer who holds an employee object can ask the employee for its name or its position.

Inheritance. As noted above, inheritance allows a programmer to define a new class simply by stating how it differs from an existing class. The new class is said to inherit from the original class and is referred to as a subclass of the original class. The original class is referred to as the superclass of the new class.

Example (2) illustrates two classes which inherit

```
CLASS Hourly_Employee INHERITS FROM
      Employee =
   METHODS
      FUNCTION hourly_salary : monetary;
      FUNCTION hours_this_month : integer;
      FUNCTION wages : monetary;
END;

CLASS Salaried_Employee INHERITS FROM
      Employee =
   METHODS
      FUNCTION salary : monetary;
      FUNCTION wages : monetary;
END;
```
(2)

from the Employee class described above. The classes extend the Employee class by providing different kinds of salary information and a way to calculate wages from that information. An instance of the Salaried-Employee class can respond to the salary or wages messages and can also respond to the name or position messages because Salaried-Employee inherits from the Employee class.

Polymorphism. As noted above, polymorphism allows different kinds of objects to be treated in a uniform manner. For example, both an Hourly-Employee object instance and a Salaried-Employee object instance can respond to the name message (inherited), the position message (inherited), or the wages message. Someone who wants to know wages does not need to know which kind of employee object is involved because either will respond to these messages.

Programming languages. The object-oriented languages in most widespread use are Smalltalk and C++. Smalltalk is an interpreted language which presents a very uniform object model in which even elementary data types are modeled as classes. The two basic operations of the language are assignment and sending a message to an object. Variables have no declared type and may refer to an instance of any class at run time. This approach provides for great flexibility in programming but can also result in run-time errors if an object is found not to respond to a message that has been received. Smalltalk supports only single inheritance. C++ is a compiled, statically typed language that was created as an extension to the C language. Basic types such as int and char are modeled as in C. Variables must be declared with types, and the compiler performs static type checking. Static type checking ensures that there is always a response for every message that is sent. C++ has lower run-time overhead than Smalltalk but at a cost of less flexibility. *See* COMPUTER PROGRAMMING; PROGRAMMING LANGUAGES.

John J. Shilling

Bibliography. B. J. Cox, *Object-Oriented System Building: A Revolutionary Approach*, 1995;

W. R. LaLonde and J. R. Pugh, *Inside Smalltalk*, 1990; S. B. Lippman, *C++ Primer*, 1989; B. Meyer, *Object-Oriented Software Construction*, 1988; J. Rumbaugh et al., *Object-Oriented Modeling and Design*, 1991.

Obolellida

A small extinct order of articulated brachiopods that ranges in age from Early to Middle Cambrian and includes the earliest known calcitic brachiopods.

Classification. The order Obolellida is included within the class Obolellata, subphylum Rhynchonelliformea, phylum Brachiopoda.

Phylum Brachiopoda
- Subphylum Rhynchonelliformea
  - Class Obolellata
    - Order Obolellida
      - Superfamily Obolelloidea
        - Family Obolellidae
          - 5 genera (Lower Cambrian)
        - Family Trematobolidae
          - 3 genera (Early-Middle Cambrian)

Morphology. Oobolellids have a biconvex, calcite, impunctate (lacking holes) shell with an elongated oval shape and a laminar secondary layer. The order includes forms that have primitive articulation of the ventral and dorsal valves—consisting of paired ventral denticles (hinge teeth), where the pedicle (fleshy stalk) attaches to the shell, and dorsal sockets along the internal posterior margin (hinge line)—and forms that lack denticles, such as the genus *Obolella*. The ventral valve (formerly named pedicle valve) has a well-defined, low and relatively short flat shelf (interarea) at the posterior. The pedicle opening (delthyrium) is located between the valves, and is either uncovered, as in *Obolella*, or closed off by convex plates (pseudodeltidium).

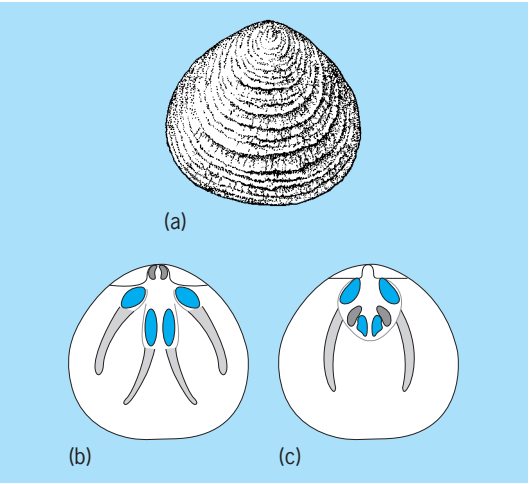

***Obolella* (*a*) Shell exterior (*after J. Hall and J. M. Clarke, An introductory to the study of the Brachiopoda intended as a handbook for the use of students, Report of the N.Y. State Geologist for 1891, 1894*). (*b*) Dorsal valve interior, and (*c*) Ventral valve interior, with schematic illustration of musculature and mantle canals (*b and c reprinted with permission from R. L. Kaesler, ed., Treatise on Invertebrate Paleontology, courtesy of and © 2000, Geological Society of America and University of Kansas*).**

The mantle canal system lacks bifurcations (baculate type) in both valves. The ventral and dorsal valves contained main mantle canals known as vascula lateralia. The dorsal valve (formerly named brachial valve) also contained a secondary mantle canal, called the vascula media.

The muscle arrangement is very similar to that of other articulated brachiopods: in the dorsal valve the pairs of anterior and posterior adductor muscles form a muscle field, while the single pair of oblique diductor muscles is attached at the bottom of the pedicle opening (notothyrial cavity). In the ventral valve, the pair of diductor muscles are located between the pairs of anterior adductor and posterior adductor muscles. Members of this group were presumably epifaunal and sessile. *See* BRACHIOPODA; RHYNCHONELLIFORMEA. Christian Emig

Bibliography. V. Iu. Gorjansky and L. E. Popov, Morphology, systematic position and origin of the inarticulate brachiopods with calcareous shells [in Russian], *Paleontologicheskii Zh.*, 1985(3):3–14, 1985; V. Iu. Gorjansky and L. E. Popov, On the origin and systematic position of the calcareous-shelled inarticulate brachiopods, *Lethaia*, 19:233–240, 1986; R. L. Kaesler (ed.), *Treatise on Invertebrate Paleontology, Part H, Brachiopoda*, Geological Society of America, Boulder, Co, and University of Kansas, Lawrence, vol. 2, 2000; Yu. L. Pelman, Early and Middle Cambrian inarticulate brachiopods of the Siberian Platform [in Russian], *Trudy Inst. Geologii i Geofiziki*, Sibirskoe otdelenie, vol. 316, 1977; L. E. Popov, L. E. Holmer, and M. G. Bassett, Radiation of the earliest calcareous brachiopods, in P. Copper and J. Jin (eds.), *Brachiopods, Proc. 3d Int. Brachiopod Congr.*, Sudbury, Ontario, Canada, Sept. 2–5, 1995, pp. 209–213, A. A. Balkema, Rotterdam, Brookfield, 1996; A. J. Rowell, The genera of the brachiopod superfamilies Obolellacea and Siphonotretacea, *J. Paleontol.*, 36:136–152, 1962; A. J. Rowell, Inarticulata, in R. C. Moore (ed.), *Treatise on Invertebrate Paleontology, Part H, Brachiopoda*, pp. 260–296, Geological Society of America, Boulder, Co, and University of Kansas Press, Lawrence, 1965; A. Williams et al., A supraordinal classification of the Brachiopoda, *Phil. Trans. Roy. Soc. Lond. B*, 351:1171–1193, 1996.

Observatory, astronomical

A telescope or telescopes, their protective enclosures (if any), support and headquarters buildings, and the staff of astronomers, engineers, technicians, and other support personnel. The telescopes can be optical or infrared (reflecting or refracting) inside a corotating dome, or radio dishes without enclosures.

The on-site support building contains the control room with the computers and control electronics to operate and point the telescope, as well as the data acquisition computers and electronics for detector

instruments. It houses laboratories for testing and calibrations, a machine shop for emergency repair and fabrication of parts, and storage for inactive instruments or telescope optics. These areas are often incorporated into a single building along with the telescope and dome. Since telescopes are generally located at remote sites to maximize their usefulness, the administration and support offices are often located in a separate headquarters building miles away in a town or university.

The *2005 Astronomical Almanac* lists over 500 observatories in 61 countries, and these are only some of the active ones. That list omits thousands of amateur telescopes, space-borne telescopes, and nontraditional observatories (such as neutrino and gravitational).

Mauna Kea Observatories. The 13,800-ft-high (4200-m) volcanic mountain Mauna Kea, in Hawai'i, offers many examples of the various configurations possible. The 3.6-m (142-in.) Canada-France-Hawaii Telescope has everything enclosed in the traditional cylindrical building with the telescope and dome on top. The W. M. Keck Observatory operates the world's two largest telescopes, the 9.8-m (386-in.) multisegmented Keck I and Keck II, with an attached support building between the domes. The 8.2-m (323-in.) Japanese Subaru Telescope has a large support building downslope from the telescope dome connected by a tunnel. The 8.0-m (315-in.) Gemini North Telescope (**Fig. 1**) has the support building on the side of the dome, but the same observatory also operates a twin 8.1-m telescope, Gemini South, a hemisphere away on Cerro Pachon in Chile. The University of Hawai'i-Mānoa operates a 2.2-m (87-in.) telescope and the NASA 3.0-m (118-in.) Infrared Telescope Facility. All these telescopes have headquarters in towns miles away. Just below the summit, the Sub-Millimeter Array consists of a control building and numerous small radio telescope dishes that can be moved to different positions via a transporter. Collectively these telescopes (and the others on the mountain) are referred to as the Mauna Kea Observatories (**Fig. 2**), yet each telescope is an observatory in its own right.

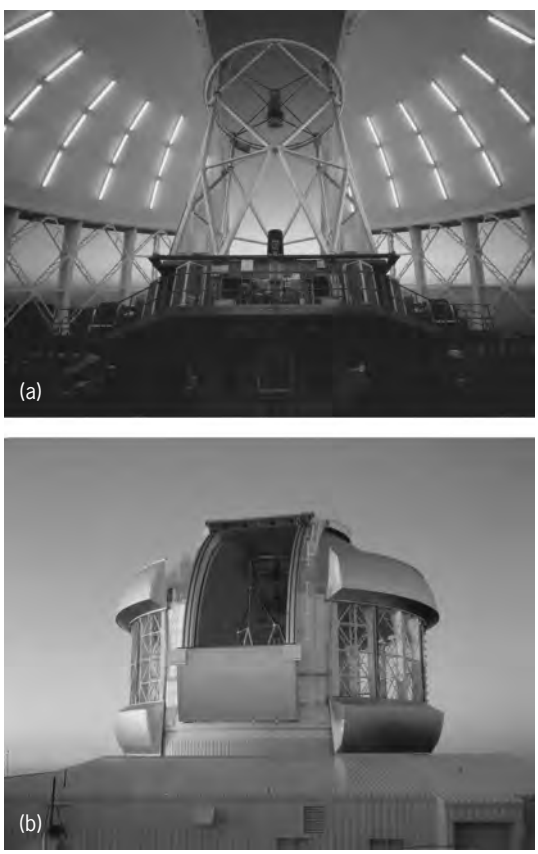

Fig. 1. Gémini North Observatory. (*a*) Telescope and interior of dome. (*b*) Exterior of dome. Vent gates surrounding the dome allow air to flow throughout the dome, thereby minimizing distortion of images due to swirling. (*Neelon Crawford, Polar Fine Arts, and Gemini Observatory*)

Sizes and types of telescopes. Telescopes are tools to collect the electromagnetic radiation commonly known as light. This light can vary in wavelength from the kilometer length of very long radio waves, through microwaves and submillimeter waves, infrared or heat radiation, half-micrometer optical or visual wavelengths, and ultraviolet radiation, to the very short nanometer-length x-rays and gamma-rays. Astronomical objects emit radiation at all these wavelengths, and telescopes have been designed to collect that light. The type of light collected determines the type of telescope and detector technology. Thus there are radio telescopes that are metallic steerable dishes, optical telescopes that use lenses, optical-infrared telescopes that use glass mirrors with a reflective coating, and x-ray telescopes that use grazing incident plates to focus the high-energy radiation. *See* ELECTROMAGNETIC RADIATION; LIGHT; TELESCOPE.

The resolution (diffraction limit) of an ideal telescope depends on only two factors, the wavelength of light observed and the size (diameter) of the collecting area: resolution = wavelength/size. To get the sharpest images, it is necessary to either observe at very small wavelengths or use larger telescopes. In practice, the Earth's atmosphere either completely absorbs some types of light or blurs the images so that the ideal is not met. Astronomers circumvent these phenomena by placing telescopes on high mountain summits, or even in space completely above the atmosphere. Adaptive optics techniques can also mitigate the blurring scintillation seen as twinkling caused by the motions of the air. The drive toward large telescopes is simply to gain the best resolution possible and to collect the most light. Larger telescopes are not built to magnify the image. *See* ADAPTIVE OPTICS; RESOLVING POWER (OPTICS).

The **table** lists the major active professional observatory telescopes and their sizes. **Figure 3** is a graph of the total collecting area of the entire world's telescopes throughout history.

Ancient observatories. The human eyes can be considered an astronomical observatory that consists of a pair of 5-mm (0.2-in.) telescopes. Naked-eye astronomy developed for thousands of years, and many early civilizations built structures to aid them in measuring recurring celestial events central for

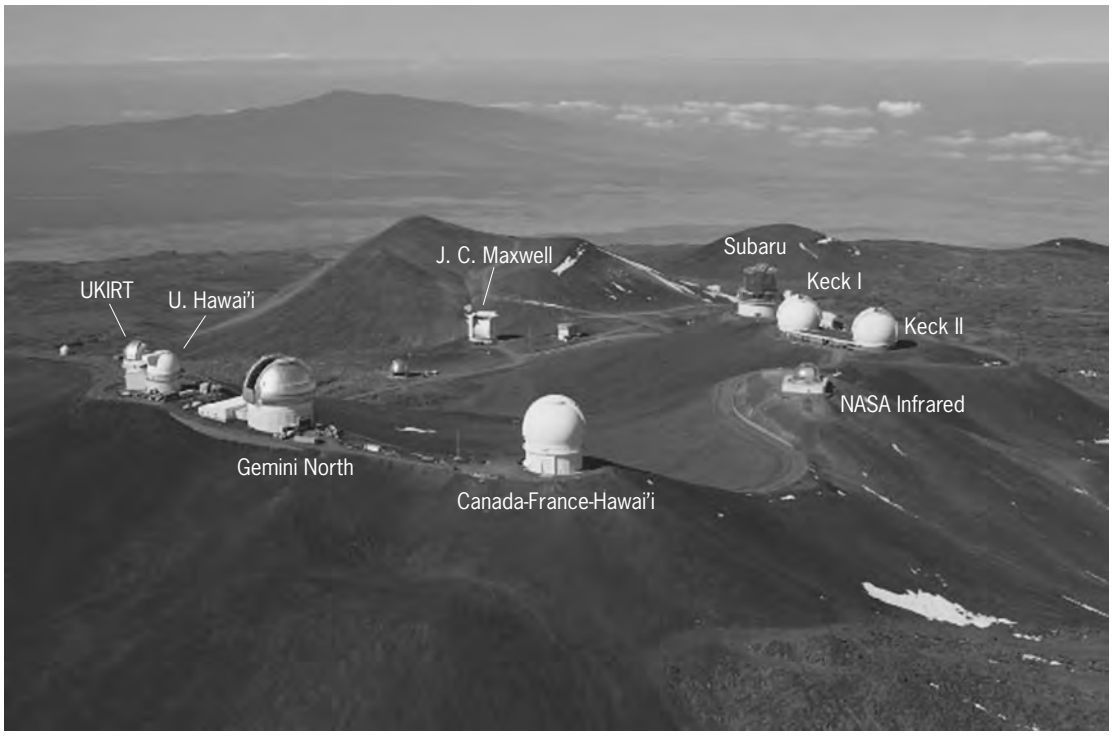

Fig. 2. Mauna Kea Observatories. (*University of Hawai'i Institute for Astronomy*)

development of their calendars. These archeological structures, which ranged from modest circles of stones to the circle of monoliths at Stonehenge in England, can rightfully be called astronomical observatories. *See* ARCHEOASTRONOMY; EYE (VERTEBRATE).

Optical and infrared observatories. These are the most common types of observatory. They are often located in remote mountains to minimize the effects of contaminating artificial lights and atmospheric blurring. For national observatories this can mean the best site in the country [such as that of the 6-m (236-in.) Bolshoi Telescope in Russia]. Some countries now place their largest telescopes at the best sites in the world [such as the United Kingdom's 4.2-m (165-in.) Herschel and 2.5-m (98-in.) Isaac Newton telescopes at La Palma in the Canary Islands, and the Japanese Subaru Telescope on Mauna Kea]. Groups of countries sometimes pool their resources to jointly operate these larger telescopes. Most observatories are affiliated with or managed by a university or research organization.

Optical observatories study planets, stars, nebulae, and galaxies. A large aperture is needed to collect the faint light of these sources. Sizes typically range from 0.5 to 10 m (20–400 in.). The smaller telescopes can use glass lenses, the largest refractor being the University of Chicago 1.0-m (40-in.) Yerkes. For the largest telescopes, engineering and manufacturing constraints dictate the reflecting design. The 9.8-m (386-in.) Keck Telescope is the world's largest reflector. The Very Large Telescope Project seeks to combine the light from four 8.2-m (323-in.) telescopes to effectively have the light-gathering capability of a 16-m (630-in.) telescope. Studies have been undertaken of a 30–50-m (1200–2000-in.) telescope.

The next generation of large ground-based telescopes truly will be enormous. A 30-m (100-ft) telescope, originally named CELT (California Extremely Large Telescope) but now called the TMT (Thirty Meter Telescope), is being proposed for location in Chile or Hawai'i as a testbed for the European Southern Obsevatory's 100-m (330-ft) telescope OWL (Overwhelmingly Large Telescope). Also proposed are a 50-m (160-ft) telescope for the Canary Islands called Euro50, a United States effort named MaxAT (30–50 m or 100–160 ft), the Giant Segmented Mirror Telescope (30-m or 100-ft) based on the Keck design, and a Canadian effort to upgrade the 3.6-m (142-in.) Canada-France-Hawai'i Telescope on Mauna Kea to a 20-m (65-ft) telescope. The reason

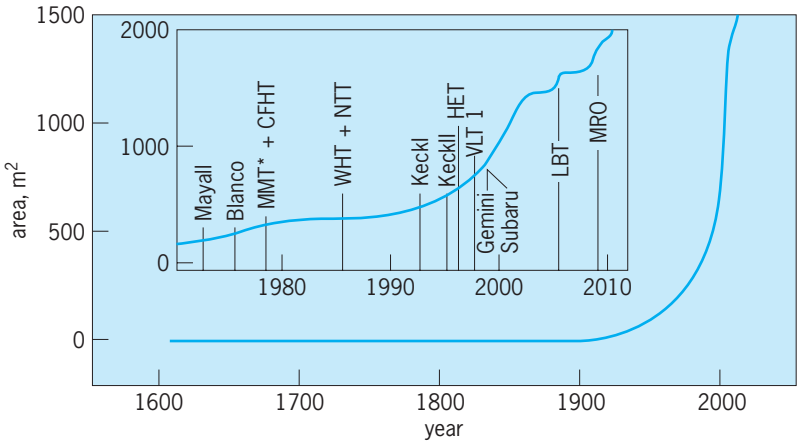

Fig. 3. Total collecting area of the world's telescopes, 1610–2010. Inset details the years 1971–2010 (data after 2005 are projected); names and dates of major observatories are included. 1 m^2 = 10.76 ft^2 = 1550 in.^{2} (*Marianne Takamiya, Institute for Astronomy, University of Hawai'i-Mānoa; updated in 2005 by John Hamilton*)

Major astronomical observatories*

Observatory	Location	Type	Size meters	Size inches	Date	Name
Human eye	Earth	Refractor	8 mm	0.32	Prehistory	
Galileo	Italy	Refractor	0.016	0.63	1609	
Galileo	Italy	Refractor	0.026	1.02	1610	
Bath, England	United Kingdom	Reflector	1.22	40	1789	William Herschel's 12.2-m (40-ft) focal length telescope
Birr, Ireland	United Kingdom	Reflector	1.83	60	1845	Lord Rosse's Leviathan
Athens	Greece	Refractor	0.6	24	1862	Newall
Treptow	Germany	Refractor	0.7	28	1869	
U.S. Naval Observatory	United States	Refractor	0.66	26.0	1873	Washington
Vienna	Austria	Reflector	0.66	26.0	1880	Grosse
Nice	France	Refractor	0.7	28	1886	
Lick	United States	Refractor	0.9	35	1888	
Paris	France	Refractor	0.8	31	1889	
Paris	France	Reflector	1.0	39	1893	
Herstmonceux	England	Refractor	0.66	26.0	1897	Thompson
Yerkes	United States	Refractor	1.0	39	1897	
Potsdam	Germany	Reflector	0.8	31	1899	
Mount Wilson	United States	Reflector	1.5	59	1908	
Lowell	United States	Reflector	1.1	43	1910	
Hamburg	Germany	Reflector	1.0	39	1911	
Mount Wilson	United States	Reflector	2.5	98	1917	Hooker
Dominion Astrophysical	Canada	Reflector	1.9	75	1918	
Johannesburg	South Africa	Refractor	0.66	26.0	1926	Innes
Geneva	Switzerland	Reflector	1.0	39	1927	
Milan	Italy	Reflector	1.0	39	1929	
Boyden	South Africa	Reflector	1.5	59	1930	
Stockholm	Sweden	Reflector	1.0	39	1931	
Haute Provence	France	Reflector	0.8	31	1932	
Harvard	United States	Reflector	1.5	59	1934	Wyeth
David Dudley Observatory	Canada	Reflector	1.9	75	1935	
McDonald	United States	Reflector	2.1	83	1939	Struve
Padua	Italy	Reflector	1.2	47	1942	
National	Argentina	Reflector	1.5	59	1942	
Haute Provence	France	Reflector	1.2	47	1943	
Palomar	United States	Reflector	5.1	200	1948	Hale
Palomar	United States	Solar	1.2	47	1948	Oschin
Crimean	Russia	Reflector	1.2	47	1952	
Mount Stromlo Observatory	Australia	Reflector	1.3	51	1954	
Hamburg	Germany	Reflector	1.2	47	1954	
Mount Observatory	Australia	Reflector	1.9	75	1955	
U.S. Naval Observatory	United States	Reflector	1.0	39	1955	
Haute Provence	France	Reflector	1.93	76	1958	
Lick	United States	Reflector	3.0	118	1959	Shane
Helwan	Egypt	Reflector	1.9	75	1960	
Okayama	Japan	Reflector	1.9	75	1960	
Schwarzschild	Germany	Solar	1.3	51	1960	
Byurakan	Russia	Solar	1.0	39	1960	
Sternberg	Russia	Reflector	1.3	51	1960	
Kitt Peak National Observatory	United States	Reflector	0.9	35	1960	
Tonanzintla	Mexico	Reflector	1.0	39	1961	
Crimean	Russia	Reflector	2.6	102	1961	Shajn
Dominion Astrophysical	Canada	Reflector	1.2	47	1961	
Royal	South Africa	Reflector	1.0	39	1961	
Uppsala	Sweden	Solar	1.0	39	1962	
Nizamiah	India	Reflector	1.2	47	1962	
Kitt Peak National Observatory	United States	Solar	1.5	59	1962	McMath
Lowell	United States	Reflector	1.8	71	1962	Perkins
U.S. Naval Observatory	United States	Reflector	1.5	59	1963	
Pic du Midi	France	Reflector	1.1	43	1964	
Siding Spring	Australia	Reflector	1.0	39	1964	
Kitt Peak National Observatory	United States	Reflector	2.1	83	1964	
Catalina	United States	Reflector	1.5	59	1965	University of Arizona
La Silla	Chile	Reflector	1.0	39	1966	
Kitt Peak National Observatory	United States	Reflector	0.9	35	1966	
Cerro Tololo InterAmerican	Chile	Reflector	1.5	59	1967	
Shemakha	Russia	Reflector	2.0	79	1967	
Northwestern	United States	Reflector	1.0	39	1967	
Ondrejov	Czechoslovakia	Reflector	2.0	79	1967	
Toledo	United States	Reflector	1.0	39	1967	
McDonald	United States	Reflector	2.7	106	1968	Harlan Smith
La Silla	Chile	Reflector	1.5	59	1968	

Major astronomical observatories (*cont.*)

Observatory	Location	Type	Size meters	Size inches	Date	Name
Yerkes	United States	Reflector	1.0	39	1968	
Kitt Peak, McGraw	United States	Reflector	2.4	94	1968	
Milan	Italy	Reflector	1.4	55	1968	
Ontario	Canada	Reflector	1.2	47	1968	
Figl	Austria	Reflector	1.5	59	1969	
Haute Provence	France	Reflector	1.52	59.8	1969	
La Silla	Chile	Solar	1.0	39	1969	
Kitt Peak, Steward	United States	Reflector	2.3	91	1969	Bok
Sacramento Peak	United States	Solar	1.6	63	1969	
Mount Hopkins	United States	Reflector	1.5	59	1970	
Baja	Mexico	Reflector	1.5	59	1970	
Mount Lemmon	United States	Infrared	1.0	39	1970	University of Arizona
Palomar	United States	Reflector	1.5	59	1970	
Mauna Kea	Hawai'i	Reflector	2.2	87	1970	University of Hawai'i
Los Campanas	Chile	Reflector	1.0	39	1971	Swope
Wise	Israel	Reflector	1.0	39	1971	
Mount Lemmon	United States	Infrared	1.5	59	1972	Minnesota, California
Tenerife	Canary Islands	Infrared	1.5	59	1972	
Siding Spring	Australia	Solar	1.2	47	1973	UK Schmidt
South African Astronomical Observatory	South Africa	Reflector	1.0	39	1973	
Padua	Italy	Reflector	1.8	71	1973	Copernicus
Kitt Peak National Observatory	United States	Reflector	4.0	157	1973	Mayall
Cerro Tololo InterAmerican	Chile	Reflector	1.0	39	1973	
South African Astronomical Observatory	South Africa	Reflector	1.9	75	1974	
Turin	Italy	Reflector	1.0	39	1974	
Konkoly	Hungary	Reflector	1.0	39	1974	
Kiso	Japan	Solar	1.1	43	1974	Kiso Schmidt
Mount Lemmon	United States	Infrared	1.5	59	1974	NASA
Center for Astronomical Investigations	Venezuela	Reflector	1.0	39	1975	
Calar Alto	Spain	Reflector	1.2	47	1975	
Struve	Russia	Reflector	1.5	59	1975	
Kitt Peak, McGraw	United States	Reflector	1.3	51	1975	
Pennsylvania State	United States	Reflector	1.5	59	1975	
National	Greece	Reflector	1.2	47	1975	
Siding Spring	Australia	Reflector	3.9	154	1975	Anglo-Australian
Cerro Tololo InterAmerican	Chile	Reflector	4.0	157	1976	Blanco
Special Astrophysical Observatory	Russia	Reflector	6.0	236	1976	Bolshoi Teleskop Azimultal'nyi (BTA)
Byurakan	Armenia	Reflector	2.6	102	1976	
University of Viriginia	United States	Reflector	1.0	39	1976	
La Silla	Chile	Reflector	3.6	142	1976	
Los Campanas	Chile	Reflector	2.5	98	1976	DuPont
University of Wyoming	United States	Infrared	2.3	91	1977	
Quebec	Canada	Reflector	1.6	63	1978	
Mauna Kea	Hawai'i	Infrared	3.8	150	1978	United Kingdom Infrared Telescope (UKIRT)
Lick	United States	Reflector	1.0	39	1979	
Mount Hopkins	United States	Reflector	4.5	177	1979	Multi-Mirror Telescope (MMT; original)
Rio de Janeiro	Brazil	Reflector	1.6	63	1979	
Calar Alto	Spain	Reflector	1.5	59	1979	
Mauna Kea	Hawai'i	Infrared	3.0	118	1979	Infrared Telescope Facility (IRTF)
Mauna Kea	Hawai'i	Reflector	3.6	142	1979	Canada-France-Hawai'i Telescope (CFHT)
Calar Alto	Spain	Reflector	3.5	138	1979	
Calar Alto	Spain	Reflector	2.2	87	1979	
Pic du Midi	France	Reflector	2.0	79	1979	
Baja	Mexico	Reflector	2.1	83	1979	
La Silla	Chile	Reflector	1.4	55	1979	
Rozhen	Bulgaria	Reflector	2.0	79	1980	
Gornegrat	Switzerland	Infrared	1.5	59	1980	
La Palma	Canary Islands	Reflector	2.5	98	1982	Isaac Newton
La Palma	Canary Islands	Reflector	1.0	39	1983	Kapteyn
Siding Spring	Australia	Reflector	2.3	91	1984	
La Silla	Chile	Reflector	2.2	87	1985	
Tuorla	Finland	Reflector	1.0	39	1985	
La Palma	Spain	Reflector	4.2	165	1986	William Herschel Telescope (WHT)
La Silla	Chile	Reflector	3.6	142	1986	New Technology Telescope (NTT)
La Palma	Canary Islands	Reflector	2.5	98	1989	Nordic
Mauna Kea	Hawai'i	Reflector	9.82	387	1991	Keck I
University of British Columbia (UBC)-Laval	Canada	Reflector†	2.7	106	1992	

Major astronomical observatories (*cont.*)

Observatory	Location	Type	Size: meters	Size: inches	Date	Name
Mount Graham	United States	Reflector	1.8	71	1994	Vatican Advanced Technology Telescope
Kitt Peak National Observatory	United States	Reflector	3.5	138	1994	Wyln
Apache Point	United States	Reflector	3.5	138	1995	Astrophysical Research Consortium (ARC)
Mauna Kea	Hawai'i	Reflector	9.82	387	1996	Keck II
Mauna Kea	Hawai'i	Reflector	8.2	323	1998	Subaru
Apache Point	United States	Reflector	2.5	98	1998	Sloan
La Palma	Canary Islands	Reflector	3.5	138	1998	Telescopio Nazionale Galileo (TNG)
Mauna Kea	Hawai'i	Infrared	8.0	315	1998	Gemini North Fred Gillette Telescope
McDonald (Mt. Fowlkes)	United States	Reflector†	9.2	362	1998	Hobby-Eberly Telescope (HET)
Mt. Hopkins	United States	Reflector	6.5	256	1999	Multi-Mirror Telescope (MMT; new)
Cerro Paranal	Chile	Reflector	8.2‡	323	1999	Very Large Telescope (VLT) 1, Antu
Cerro Paranal	Chile	Reflector	8.2‡	323	2000	VLT 2, Kueyen
Las Companas	Chile	Reflector	6.5	256	2000	Magellan 1, Walter Baade Telescope
Cerro Pachon	Chile	Infrared	8.0	315	2000	Gemini South
Las Companas	Chile	Reflector	6.5	256	2001	Magellan 2, Landon Clay Telescope
Cerro Paranal	Chile	Reflector	8.2‡	323	2001	VLT 3, Melipal
Paranal	Chile	Reflector	8.2‡	323	2002	VLT 4, Yepun
Vancouver, British Columbia	Canada	Reflector†	6.0	236	2004	Large Zenith Telescope
Haleakalā, Maui	Hawai'i	Robotic reflector	2.0	79	2004	Faulkes Telescope North
La Palma	Canary Islands	Reflector	10.4	409	2005	Gran Telescopio Canarias
Cerro Pachon	Chile	Reflector	4.1	161	2005	Southern Astrophysical Research (SOAR)
Sutherland	South Africa	Reflector†	10.0	394	2005	South African Large Telescope (SALT)
Chelmos	Greece	Reflector	2.3	91	2005	Aristarchos
Cerro Paranal	Chile	Reflector	2.6	102	2005	Very Large Telescope Survey Telescope (VLTST)
Mount Graham	United States	Reflector	2 × 8.4§	331	2006	Large Binocular Telescope (LBT) 1 and 2
Cerro Paranal ("NTT peak")	Chile	Reflector	4.0	157	2006	Visible and Infrared Survey Telescope for Astronomy (VISTA)
Siding Spring	Australia	Robotic reflector	2.0	79	2007	Faulkes Telescope South
Magdalena, New Mexico	United States	Reflector, interferometer	1 × 2.4, 10 × 1.4	1 × 94, 10 × 55	2008	Magdalena Ridge Observatory (MRO)
Airborne; Ames, California	United States, Germany	Reflector mounted on Boeing 747SP	2.5	98	2008	Stratospheric Observatory for Infrared Astronomy (SOFIA)**
Mauna Kea	Hawai'i	Robotic reflector	0.9	35	2008	University of Hawai'i at Hilo
Flagstaff, Arizona	United States	Reflector	4.2	165	2009	Discovery Channel Telescope (DCT)
Haleakalā, Maui	Hawai'i	Reflector, solar only	4.0	157	2010	Advanced Technology Solar Telescope
Mauna Kea	Hawai'i	Reflector, interferometer	6 × 2.0, 2 × 9.82	6 × 79, 2 × 387,	2010?	Keck Interferometer with outrigger telescopes
Mauna Kea	Hawai'i	All-sky reflector	4 × 1.8	4 × 71	2010?	Pan STARRS
	Chile, Mexico, or Canary Islands	Reflector	8.4 (6.9 clear aperture)	331 (272 clear aperture)	2012	Large Synoptic Survey Telescope (LSST)
Xinglong	China	Reflector, collector for optical fibers	6.67 × 6.05	263 × 238	?	Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST)

*Optical, infrared, and solar telescopes only; radio and space telescopes omitted. Compiled by Marianne Takamiya, Institute for Astronomy, University of Hawai'i-Mānoa; updated in 2005 by John Hamilton.
†Not fully steerable.
‡The Very Large Telescope Project seeks to combine the light from the four 8.2-m (323-in.) telescopes, VLT 1–4, to effectively have the light-gathering capability of a 16-m (630-in.) telescope.
§Light from the two Large Binoculas Telescopes will be combined, so that the light-gathering area as seen by the detector will be 11.8 m (465 in.).
**Replacement for Kuiper Airborne Observatory (KAO), a Lockhead C-141 with 0.9-m (35-in.) infrared-optimized reflector. Flew 1971–1995.

that there are no plans for 12-, 15-, or 18-m telescopes is related to the increase in the collecting area of a telescope, whose light-gathering power increases as the square of the radius. Thus, a 100-m telescope will collect 100 times the amount of light collected by a 10-m telescope, even thought it is "only" ten times bigger. It is currently felt that such an increase in light is necessary for answering many of the significant questions in modern astronomy.

Most optical observatories can also observe in the near-infrared region. Infrared observatories have been optimized to work further into the longer (thermal) infrared. Since glass absorbs infrared light, all infrared telescopes are of the reflecting type. Special design techniques must be used, as the telescope itself glows in the infrared. The secondary mirror can be rapidly tilted (chopped) to sequentially view the (glowing) sky and the sky plus source. When averaged, the thermal effects of the sky will cancel, as they are random. Infrared observatories must be

placed in dry climates, as the water vapor in the air absorbs the infrared radiation. High mountain sites also tend to be dry, as water vapor content decreases with altitude. *See* INFRARED ASTRONOMY.

Radio observatories. Radio observatories have become an essential complement to optical observatories. With their longer wavelengths, the radio telescopes can see through cosmic clouds and measure the properties (temperature, pressure, chemical composition, and velocities) of gases which pervade the universe.

Since the wavelengths are so long, the telescopes must be large to achieve a high resolution (the ability to discriminate between two nearby sources). Radio dishes are typically tens to hundreds of meters wide. The largest movable radio telescope (100 m or 328 ft in diameter) is operated by the Max Planck Institute for Radio Astronomy at Effelsberg, Germany. The 305-m (1000-ft) telescope at Arecibo, Puerto Rico, uses a natural valley for its dish and a movable secondary feed to point about the sky.

The longer radio wavelengths allow the use of interferometers, where the signal from several small radio telescopes can be mathematically combined to yield results as if they had been collected from a very large dish. The Very Large Array radio observatory in Socorro, New Mexico, has twenty-seven 25-m (82-ft) antennas arranged in a Y pattern, which can yield the resolution of a single telescope 36 km (22 mi) wide and a sensitivity of a dish 130 m (426 ft) wide. Even larger synthetic telescopes can be created with baselines of intercontinental lengths. The Very Long Baseline Array uses ten 25-m (82-ft) antennas spread from Hawai'i to the Virgin Islands, allowing the resolution of a telescope dish that distance across. *See* RADIO TELESCOPE.

Solar observatories. Solar observatories have telescopes optimized for an objective opposite to that of the optical telescopes. Here the problem is too much light (and heat). Solar telescopes overcome this problem by greatly reducing the amount of light, by using smaller telescopes, and by passing the light through very narrow filters. Daytime observations also present the problem of more severe air convection around the telescope due to the heat of the sunlight. The Big Bear Solar Observatory is built in the center of a lake to mitigate this convection and achieve sharp images. When high resolution is needed (which calls for larger telescopes), the design must include means to reject the solar heat and reduce the convection turbulence in the telescope itself. The National Solar Observatory operates the Vacuum Telescope at Kitt Peak, Arizona, where the sunlight is reflected off a coelostat mirror into a vacuum tube on its way to the detector; the absence of air means that there is no disturbing convection. Larger solar telescopes are now being built, such as the 4-m (157-in.) Advanced Technology Solar Telescope. *See* SUN.

Airborne and space observatories. Crewless balloons have been used to hoist telescopes above the absorbing atmosphere for gamma-ray, x-ray, and ultraviolet observations. Their lower cost (compared to space missions) is offset by the short duration of the observations. The National Aeronautics and Space Administration (NASA) successfully operated the Kuiper Airborne Observatory for 20 years. This facility had a 0.9-m (36-in.) telescope aboard a converted C-141 military cargo plane flying at an altitude of 12.5 km (41,000 ft), which is above 85% of the atmosphere and 99% of the water vapor. The successor is the Stratospheric Observatory for Infrared Astronomy (SOFIA), a Boeing 747-SP aircraft carrying a 2.5-m (98-in.) telescope. In addition, NASA has used high-flying U2 spy planes to measure variations in the cosmic microwave radio background.

Space observatories have amply proven their worth despite their initial high cost and limited mission lifetimes. Unimpeded high-energy gamma-ray and x-ray observations from observatories such as *Einstein* and *Compton* have redefined the universe. The *Cosmic Background Explorer* (*COBE*) microwave satellite has observed the earliest viewable portions of the universe, the cosmic microwave background. A followup mission, the *Wilkinson Microwave Anisotropy Probe* (*WMAP*), has enabled cosmologists to accurately measure many of the physical parameters of the early universe. The Hubble Space Telescope, an orbiting 2.4-m (94-in.) optical-infrared telescope, has achieved unparalleled images of the sky, from exploding supernovae, to star-forming regions, to the deepest (farthest) images of the universe. An infrared-optimized counterpart, the 0.85-m (33-in.) *Spitzer Space Telescope* (formerly *SIRTF, Space Infrared Telescope Facility*), was launched into orbit in 2003. The success of these space observatories has led to the planning of a 6-m (236-in.) successor, the *James Webb Space Telescope*. *See* COSMIC BACKGROUND RADIATION; GAMMA-RAY ASTRONOMY; SATELLITE (ASTRONOMY); X-RAY TELESCOPE.

Virtual observatories. With the advent of digital format in astronomical images and spectra, large data-storage facilities, data-mining software, and the World Wide Web, the concept of a virtual observatory has been established. It will allow researchers global electronic access to combined archived data from many ground and space telescopes. Today, at most major telescopes, all new data are archived (and released publicly after the usual proprietary period). Efforts are underway to digitally convert old photographic plates for computer storage and retrieval. Leaders in this field are the United States, with the National Virtual Observatory; the European Southern Observatory, with the Astrophysical Virtual Observatory; and the United Kingdom's AstroGrid project. This effort is seen as so important that these groups have formed the International Virtual Observatory Alliance to ensure commonality and compatibility with data transfers, protocols, and formats. Soon professional astronomers can observe while never visiting a telescope. *See* DATA MINING; WORLD WIDE WEB.

Detectors and instruments. An astronomical instrument is the working heart of an observatory. The telescope exists solely to collect and funnel light into the instrument. Astronomical instruments range from a simple eyepiece for direct viewing (with different magnifications), to a camera for imaging on

film or electronic detectors, to a spectrograph which records the wavelength distribution of light energy (analogous to a rainbow). For radio telescopes the instrument is an amplifier that isolates and boosts the weak celestial electric signal. *See* ASTRONOMICAL SPECTROSCOPY; CAMERA; SPECTROGRAPH.

Before the invention of photography, an eyepiece was used as the instrument with the human eye as the detector. Observations consisted of written notes and sketches. When photography was invented, one of the first uses was to image the Moon through a telescope. Photography soon became the mainstay for astronomical observations as it presented a permanent unbiased record of an event. Most observatories had well-equipped darkrooms, and techniques for extracting as much information as possible from photographic emulsions, such as hypersensitizing and plate baking in a nitrogen atmosphere. The electronic revolution introduced the photomultiplier tube and photon counters, which were very useful for photometry (the precise measurement of stellar brightness). *See* ASTRONOMICAL IMAGING; PHOTOMULTIPLIER.

The charge-coupled device (CCD) has now virtually supplanted the photographic plate. Its advantages include high quantum efficiency, large linear response range, and the digitized nature of the data, allowing immediate display on computer terminals and ease of computer data reduction. To support night-vision technology, charge-coupled devices have also been developed as useful detectors in the infrared region. Photography is still useful for wide-field applications, but arrays of large-format charge-coupled devices are closing the gap there also. *See* CHARGE-COUPLED DEVICES; INFRARED IMAGING DEVICES.

Unconventional observatories. Besides observing light (electromagnetic radiation), astronomers are mapping the universe via neutrinos, particles that can transverse matter with a minimum of interaction. Neutrinos are as numerous as photons of light. Neutrinos have been observed from the core of the Sun. Interpretations of observed solar neutrino fluxes have led to the discovery of oscillations between the three types of neutrinos, the electron, μ (mu), and τ (tau) neutrinos. This has led to the discovery that neutrinos are not massless after all. Neutrino bursts have been detected from an exploding supernova (SN 1987A). With the discoveries of dark matter and dark energy as majority components of the universe, astroparticle physics will come to play an ever-increasing role in astronomy. The Super-Kamiokande Detector in Japan and the Sudbury Neutrino Observatory are examples. Proposed new projects such as AUGER in Argentina, ICECUBE in Antarctica, and Ashra in Hawai'i seek to observe the very highest energy cosmic rays and neutrinos. *See* COSMOLOGY; DARK ENERGY; DARK MATTER; NEUTRINO; NEUTRINO ASTRONOMY; SOLAR NEUTRINOS.

Gravitational wave observatories attempt to detect ripples (gravitational waves or gravitons) in the fabric of space-time itself. It is thought that these can be generated by collapses of massive black holes. There are two main approaches to detection, laser interferometry and resonant mass detectors. Examples of each are the Laser Interferometer Gravitational-Wave Observatory (LIGO) and the TIGA (Truncated Icosahedral Gravitational Wave Antenna) project. *See* GRAVITATION RADIATION; TELESCOPE.

John Hamilton

Bibliography. S. J. Dick, *Sky and Ocean Joined: The U.S. Naval Observatory 1830–2000*, Cambridge University Press, 2002; R. Kerrod, *Hubble: The Mirror on the Universe*, Firefly Books, 2003; H. T. Kirby-Smith, *U.S. Observatories: A Directory and Travel Guide*, Van Nostrand Reinhold, 1976; K. Krisciunas, *Astronomical Centers of the World*, Cambridge University Press, 1988; S. Laustsen, C. Madsen, and R. M. West, *Exploring the Southern Sky: A Pictorial Atlas from the European Southern Observatory (ESO)*, Springer, 1987; W. P. McCray, *Giant Telescopes: Astronomical Ambition and the Promise of Technology*, Harvard University Press, 2004; R. Naeye, *Signals from Space: The Chandra X-Ray Observatory*, Raintree Steck-Vaughn, 2000; T. Orr, *The Telescope (Inventions That Shaped the World)*, Scholastic, 2004; C. C. Petersen and J. C. Brandt, *Vision of the Cosmos*, Cambridge University Press, 2003; A. Sandage, *Centennial History of the Carnegie Institution of Washington*, vol. 1: *The Mount Wilson Observatory: Breaking the Code of Cosmic Evolution*, Cambridge University Press, 2005; F. Watson, *The Stargazer: The Life and Times of the Telescope*, Da Capo Press, 2005.

Obsessive-compulsive disorder

A type of anxiety disorder (commonly referred to as OCD) characterized by recurrent, persistent, unwanted, and unpleasant thoughts (obsessions) or repetitive, purposeful ritualistic behaviors that the person feels driven to perform (compulsions). A cardinal feature of this disorder is an awareness of the irrationality or excess of the obsessions and compulsions accompanied by an inability to control them.

Characteristics. Typical compulsions include an irresistible urge to wash (particularly the hands) or clean, to check doors to confirm that they are locked, to return repeatedly to appliances to make sure they are turned off, to touch, to repeat, to count, to arrange, or to save. Typical obsessions include overconcern about dirt and contamination, fear of acting on violent or aggressive impulses, feeling overly responsible for the safety of others, abhorrent religious (blasphemous) and sexual intrusions, and inordinate concern with arrangement or symmetry. Obsessions may accompany compulsions, or compulsions may occur alone.

The most common subtype of the disorder is washing. The washers are driven to perform repeated handwashing and are obsessed with a fear of dirt, germs, and contamination. They may spend several hours each day washing their hands or showering. Typically they try to avoid what they perceive to be

sources of contamination, like door knobs, electric switches, or dust.

A second presentation for obsessive-compulsive disorder involves pathologic doubt coupled with compulsive checking. Some patients have an incessant need for symmetry, but typically the checker is concerned for the safety of others. Checking, which is enough to resolve normal uncertainty, often contributes to doubt and leads to even more checking.

The difference between obsessive-compulsive disorder and milder forms of obsession or compulsion seen in otherwise healthy people is that for the sufferer the obsessions or compulsions cause marked distress, are time-consuming, and significantly interfere with the person's normal routine, occupational functioning, usual social activities, and relationships with others.

Although the obsessions and compulsions may seem "crazy," a typical patient with obsessive-compulsive disorder is acutely aware of the irrationality or excessiveness of his or her fears or behaviors, yet is unable to control them. Such self-awareness is associated with a tendency to hide the symptoms, and usually persons with obsessive-compulsive disorder are successful in concealing the problem, because in areas other than that of the obsession or compulsion they are perfectly reasonable. This secretiveness may explain why the prevalence of obsessive-compulsive disorder in the general population, which has been reported to be more than 2.0% in several recent epidemiological studies in the United States, was once thought to be 0.05%.

Onset in adolescence occurs in about a third of cases. In another third symptoms appear in early adulthood, and in the last third they start later in life. If not treated appropriately, the disorder is often chronic, with waxing and waning of symptoms.

Treatment. Obsessive-compulsive disorder is generally resistant to traditional psychotherapy, which has tried to trace the condition to conflicts of early childhood. An effective mode of psychotherapy is behavioral therapy, in which the patients are gradually exposed to their feared or triggering situation but are prevented from performing accompanying compulsions. For example, patients with a fear of contamination from germs are instructed to touch a very dirty-looking, "contaminated" cloth, but they are not allowed to perform handwashing afterward. This approach, which focuses on treating the symptoms rather than trying to understand their origin, seems to be more effective in treating the ritualistic behavior (compulsions) than the pervasive thoughts (obsessions).

Obsessive-compulsive disorder is also refractory to most drugs used to treat anxiety, depression, and psychosis. However, it often eases with medications that affect the brain's serotonergic system, such as clorimipramine, fluvoxamine, and fluoxetine.

Serotonin, like other neurotransmitters, is released into the gap between two nerve cells (synapse) and later must be removed, through a process called reuptake, before the presynaptic cell can be fired again. Clorimipramine, fluvoxamine, and fluoxetine are unique among the psychoactive drugs as they specifically block the reuptake of serotonin in the synapse. The observation that the disorder responds only to drugs that alter the serotonin function in the brain suggests a pivotal role for serotonin in this disorder. Indeed, compared with normals, persons with obsessive-compulsive disorder appear particularly sensitive to the behavioral effect of methachlorophenylpiperazine (mCPP), a compound that activates the serotonergic system and may create new obsessions that usually last as long as the serotonergic system is activated, that is, a couple of hours. *See* SEROTONIN; SYNAPTIC TRANSMISSION.

Another approach in studying the psychobiology of psychiatric disorders is examination of the functional anatomy of the brain of those with the condition using such techniques as positron emission tomographic (PET) and magnetic resonance imaging (MRI), which have revealed abnormalities in the cortex (frontal lobes) and basal ganglia along with decreased caudate volume. Obsessive-compulsive disorder thus may reflect a discordance between the brain's most advanced region (the cortex) and one of its more primitive areas (the basal ganglia). *See* MEDICAL IMAGING.

The specific response of patients with obsessive-compulsive disorder to serotonergic drugs, their hypersensitivity to activation of the serotonergic system, and the distinct functional anatomy differences found in those patients suggest a biological cause for this disorder. In this regard, obsessive-compulsive disorder represents a shift from a psychological to a neurobiological approach in the study of anxiety disorders. *See* ANXIETY DISORDERS; NEUROTIC DISORDERS.

Joseph Zohar

Bibliography. E. Hollander (ed.), *Current Concepts in Obsessive-Compulsive Disorder*, 1994; J. L. Rapoport, The biology of obsessions and compulsions, *Sci. Amer.*, pp. 83–89, March 1989; J. Zohar, E. B. Foa, and T. R. Insel, Obsessive-compulsive disorder, in *Treatments of Psychiatric Disorders*, American Psychiatric Association, pp. 2095–2105, 1989; J. Zohar and T. R. Insel, Obsessive-compulsive disorder: Psychobiological approaches to diagnosis, treatment and pathophysiology, *Biol. Psychiat.*, 22: 667–687, 1987.

Obsidian

A volcanic glass, usually of rhyolitic composition, formed by rapid cooling of viscous lava. The color is jet-black because of abundant microscopic, embryonic crystal growths (crystallites) which make the glass opaque except on thin edges. Iron oxide dust may produce red or brown obsidian. *See* LAVA.

Obsidian usually forms the upper parts of lava flows. Well-known occurrences are Obsidian Cliffs in Yellowstone Park, Wyoming; Mount Hekla, Iceland; and the Lipari Islands off the coast of Italy. Less commonly, obsidian forms selvages of dikes and sills. *See* IGNEOUS ROCKS; VOLCANIC GLASS.

Carleton A. Chapman

Occultation

The temporary blocking from view of one celestial body by another. The occulting body is the one closer to the observer, and can be a planet, moon, ring system, or other body, usually in the solar system. The occulted body is smaller in apparent, projected size and is usually a distant star, although it can also be a spacecraft radio signal, as in the case of the *Voyager* spacecraft radio occultations at the outer planets, or another solar system body. Examples of occultations are that of a star by a planet, a lunar occultation of a star, and an occultation by Pluto of its satellite charon. Although a solar eclipse is not usually thought of in these terms, this event is actually an occultation of the Sun by the Moon. *See* ECLIPSE; PLAUET; SATELLITE (ASTRONOMY); SOLAR SYSTEM.

Observations of occultations can reveal information about the physical size of the blocking body, the structure of its atmosphere (bound, in the case of planetary atmospheres, or unbound in the case of comets), or the structure of its rings. Occultations of stars by the Moon are used primarily to study the structure of the occulted stars, quantities such as their projected size, and limb darkening, though the figure of the lunar limb used to make predictions of Baily's beads at eclipses is based on lunar occultations.

Observations. Although occultations provide a wealth of information about solar system objects, the observation of an occultation is primarily a monitoring of stellar brightness over time. In fact, the best data are obtained when the occulting solar system body is as close to invisible as is possible. Observations of stellar occultations typically require fast sampling rates, monitoring the stellar brightness several times each second. Such observations are possible with photoelectric photometers and charge-coupled devices (CCDs) designed for such rapid readout. The quality of the resulting light curve depends on the stellar brightness and the brightness of the background. The resulting light curve, or plot of stellar brightness against time, shows the star at full level, then a drop in brightness when star is occulted, and then a rise of the signal back to the full level (see **illustration**). *See* ASTRONOMICAL IMAGING; CHARGE-COUPLED DEVICES.

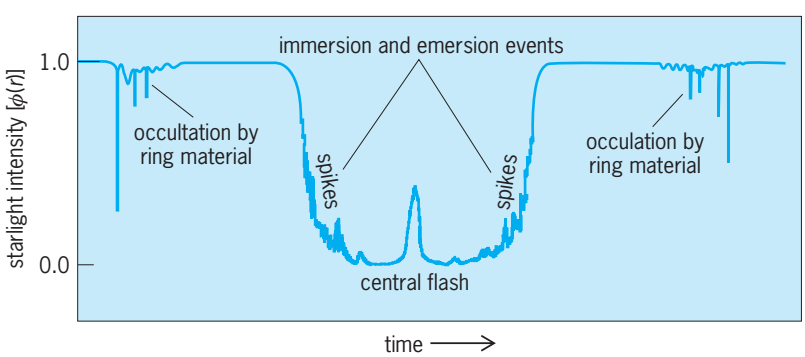

Schematic light curve of a central occultation by a ringed planet with substantial atmosphere. The relative brightness of the star, $\phi(r)$, is plotted as a function of time or distance from the center of the planet r. The decrease in stellar brightness occurs when ring material obscures the star and when the atmosphere refracts the starlight. Immersion and emersion events provide information about the temperature, composition, and structure of the upper atmosphere. When the observer, occulting body, and star are precisely aligned, an additional feature is visible, the central flash formed by focusing the starlight around the entire atmosphere. (*After J. L. Elliot, Stellar occultation studies of the solar system, reprinted with permission from Annual Review of Astronomy and Astrophysics, volume 17 © 1979 by Annual Reviews, www.annualreviews.org*)

The spatial resolution of a stellar occultation is limited by Fresnel diffraction, not by the seeing disk as for imaging observations. This limit depends on both the wavelength at which the observations are made and the distance between the observer and the occulting body. At Pluto and at visible wavelengths (500 nm), the limiting occultation spatial resolution is approximately 1 mi or 1.7 km. When this is compared to the imaging spatial resolution at a superb observing site (seeing ~0.5 arcsec) of 9000 mi or 14,500 km on Pluto, the advantage of occultations is obvious. *See* DIFFRACTION; TELESCOPE; TWINKLING STARS.

Analysis of stellar occultation data involves converting the time series light curve into a spatial scan, using knowledge of the geometry of the occulting body and the observer. These calculations rely heavily on solar system ephemerides; for occultations by small bodies in the outer solar system, such as Pluto, Triton, comets, and Kuiper Belt objects, the accuracy of event predictions are constrained by the accuracy of their ephemerides. To compensate for these uncertainties in the data analysis, typically two or more light curves from different observing sites are required, since the circular silhouette has different-length chords at different parts of the surface. These are then combined and the ephemeris uncertainties are removed. *See* EPHEMERIS.

Applications. Occultations by asteroids have been observed extensively to determine the shapes of these bodies. Several observers with fixed-location or portable telescopes can map the shape of an asteroid, including any departures from a purely spherical shape. Attempts are being made to extend the utility of occultations for determining sizes of solid- surface bodies to cometary nuclei and Kuiper Belt objects. *See* ASTEROID; COMET; KUIPER BELT.

For bodies with atmospheres, the stellar signal disappears gradually because of refraction of the stellar light ray by the body's atmosphere. Analysis of the occultation light curves when combined with knowledge of the atmospheric composition (usually obtained through spectral observations) can reveal the temperature in the atmosphere at a particular altitude or pressure level, and how the temperature varies with altitude. If any extinction is present in the form of haze or clouds, it can be detected as well. This method has been used on all planets with even minimal atmospheres, from Venus out to Pluto (including Earth). In Stellar occultation observations yielded the first direct detection of Pluto's atmosphere, although its presence had been suspected earlier from spectral analysis of surface ices. *See* PLUTO; REFRACTION OF WAVES.

Occultations can be used to study the response of an atmosphere to seasonal variations, such as Pluto's large changes in solar distance and heating due to its substantial orbital eccentricity, and Triton's large swings in subsolar latitude or orientation toward the Sun. In fact, some have predicted a complete

freeze-out of Pluto's atmosphere as it recedes from the Sun. (Pluto's closest passage to the Sun was in 1989, and its farthest distance will be in 2114.) Occultation observations have revealed surprising pressure increases in both Pluto's and Triton's atmospheres since about 1990, perhaps in response to small surface temperature changes carried up to the atmosphere via vapor-pressure equilibrium. In addition, occultations by both bodies have allowed study of the nonsphericity of both atmospheres, possibly attributed to zonal winds. *See* NEPTUNE.

Understanding of planetary rings has been revolutionized by stellar occultation observations, both by spacecraft and by ground-based observers. The ring system of Uranus was discovered during a stellar occultation in 1977; prior to this, only the bright rings of Saturn were known to astronomers. Since the discovery of Uranus' rings, many occultations by the rings of Saturn, Uranus, and Neptune have been observed. Multiple occultation observations are combined to study the shapes of rings and many have been found to be noncircular, driven by resonances with nearby satellites. Dynamical studies of ring lifetimes suggest relatively short-lived systems of perhaps a few million years (much shorter than the age of the solar system at 4.5×10^9 years). Ring occultation studies suggest one reason why such varied ring systems can be viewed today: The satellite resonances prevent the spread and dissipation of ring material, in effect freezing the rings in place. *See* SATURN; URANUS.

Amanda S. Bosh

Bibliography. J. K. Beatty and A. Killian, Discovering Pluto's atmosphere, *Sky Telesc.*, 76:624–627, 1988; D. P. Cruikshank, Triton, Pluto, and Charon, in J. K. Beatty et al. (eds.), *The New Solar System*, 4th ed., pp. 285–296, Cambridge University Press, Sky Publishing Corp., 1999; J. L. Elliot, E. Dunham, and R. L. Millis, Discovering the rings of Uranus, *Sky Telesc.*, 53:412–416, 1977; J. L. Elliot et al., The recent expansion of Pluto's atmosphere, *Nature*, 424:165–168, 2003; R. G. French et al., Dynamics and structure of the Uranian rings, in J. T. Bergstralh et at. (eds.), *Uranus*, pp. 327–409, University of Arizona Press, 1991; B. Sicardy et al., Large changes in Pluto's atmosphere as revealed by recent stellar occultations, *Nature*, 424:168–170, 2003.

Ocean

One of the major subdivisions of the interconnected body of salt water that occupies almost three-quarters of the Earth's surface. Earth is the only planet in the solar system whose surface is covered with significant quantities of water. Of the nearly 1.4 billion cubic kilometers of water found either on the surface or in relatively accessible underground supplies, more than 97% is in the oceans; most of the rest is in glacier-covered Greenland and Antarctica (**Table 1**). During much of Earth's history the oceans have been a difficult, if not quite impassable, barrier for the movement of land-based plants and animals from one continent or island to another. This article describes and compares some of the major oceans and their features. *See* GLACIOLOGY; OCEANOGRAPHY.

TABLE 1. Distribution of water on Earth

	Volume	
	km^3	%
Oceans	1,348,000,000	97.39
Polar ice caps and glaciers	27,820,000	2.01
Ground water	8,062,000	0.58
Lakes and rivers	225,000	0.02
Atmosphere	13,000	0.001

Oceans and the seas that connect them cover some 73% of the surface of the Earth, with a mean depth of 3729 m (**Table 2**). More than 70% of the oceans have a depth between 3000 and 6000 m. Less than 0.2% of the oceans have depths as great as 7000 m (**Table 3**).

The oceans are cold and salty. Some 50% have a temperature between 0 and 2°C and a salinity between 34.0 and 35.0‰. To a high degree of approximation, a salinity of 34‰ is the equivalent of 34 grams of salt in a kilogram of seawater. Water with a temperature above a few degrees Celsius is confined to a relatively thin surface layer of the ocean. **Figure 1** shows the distribution of both temperature and salinity in the world's oceans. *See* SEAWATER.

Nearly all elements known to humankind have been found dissolved in seawater, and those that have not are assumed to be present. However, all but a few are found in very small amounts. Sodium chloride accounts for some 85% of the dissolved salts, and an additional four ions (sulfate, magnesium, calcium, and potassium) bring the total to more than 99.3%.

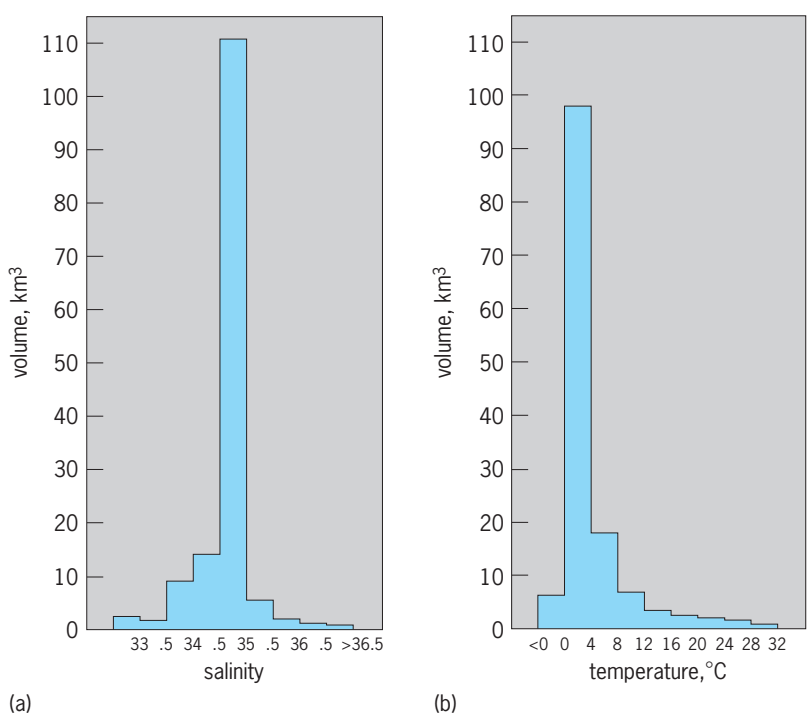

Fig. 1. Range of (*a*) temperature and (*b*) salinity in the world's oceans.

TABLE 2. Ocean basin characteristics

	Area, km^2	Volume, km^3	Mean depth, m
Pacific	181,344,000	714,410,000	3940
Atlantic	94,314,000	337,210,000	3575
Indian	74,118,000	284,608,000	3840
Arctic	12,257,000	13,702,000	1117
Total	362,033,000	1,349,929,000	3729

TABLE 3. Distribution of oceanic depth

Depth, m	Area, km^2	Volume, %	Cumulative %
0–200	27,123,000	7.49	7.49
200–1000	16,012,000	4.42	11.91
1000–2000	15,844,000	4.38	16.29
2000–3000	30,762,000	8.50	24.79
3000–4000	75,824,000	20.94	45.73
4000–5000	114,725,000	31.69	77.42
5000–6000	76,753,000	21.20	98.62
6000–7000	4,461,000	1.23	99.85
7000–8000	380,000	0.10	99.96
8000–9000	115,000	0.03	99.99
9000–10,000	32,000	0.01	100.00
10,000–11,000	2,000	0.00	100.00

The ratio of ions is remarkably constant from one ocean to another and from top to bottom of each. Until about 1960, oceanographers calculated the total salinity by a titration process that was essentially equivalent to measuring only the chloride ion and multiplying by a constant. They could be confident that the uncertainty in their calculation was no more than ±.02 because of variations in the distribution of different ions within their samples. Calculation of the salinity equivalent is now based on measurement of the electrical conductivity of the seawater, and is an order of magnitude more accurate.

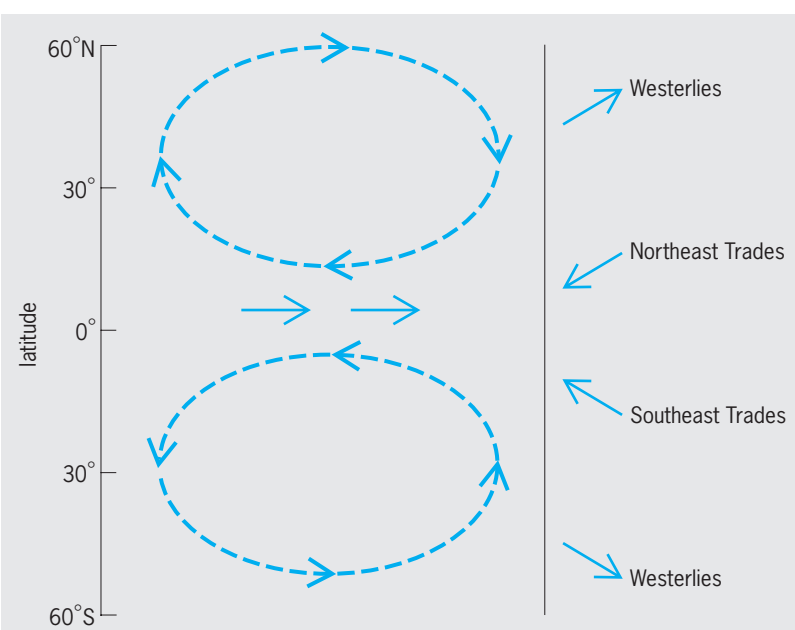

Fig. 2. Role of the winds in establishing major current gyres (closed circulatory systems that are larger than whirlpools or eddys). (*From J. A. Knauss, Introduction to Physical Oceanography, 2d ed., Prentice Hall, 1997*)

Ocean salinity is primarily controlled by the balance of precipitation, river runoff, and evaporation of water at the sea surface (**Table 4**). The highest salinities are found in major evaporation basins with little rainfall or river runoff, such as the Red Sea. The lowest salinities are found near the mouths of major rivers such as the Amazon. Some of the lowest open-ocean salinities are found in the relatively small and isolated Arctic Ocean, which contains only 1% of the total volume of the oceans but drains several of the world's largest river systems, the Mackenzie from Canada and the Lena, Yenisei, and Ob from Russia. *See* ARCTIC OCEAN; RED SEA.

The tropics receive much more heat from the Sun than do the polar regions. However, the average temperature gradient between high and low latitudes is significantly less than might be expected, because the oceans are continually transporting excess heat (warm water) from the tropics toward the Poles and returning colder water toward the tropics. This process of moving excess heat from lower (south of 40°) to higher (north of 40°) latitudes is shared approximately equally by the oceans and the atmosphere. A significant part of the ocean heat exchange process is carried out by the major ocean currents, the "named" currents such as the Gulf Stream, Brazil Current, California Current, and Kuroshio. These currents are primarily driven by the winds, and there is considerable similarity in their pattern from one ocean basin to another. *See* GULF STREAM; KUROSHIO.

The average winds over the North and South Atlantic as well as the North and South Pacific oceans come out of the west (westerlies) at the middle

TABLE 4. Water balance of the world's oceans*

	P	*E*	*P*−*E*	River runoff	Runoff + *P*−*E*
Arctic	97	53	+44	307	+351
Pacific	1292	1202	+90	69	+159
Indian	1043	1294	−251	72	−179
Atlantic	761	1133	−372	197	−175

*Precipitation (*P*) and evaporation (*E*) are in millimeter per year. River runoff is in millimeter per year equivalent for the appropriate ocean basin.

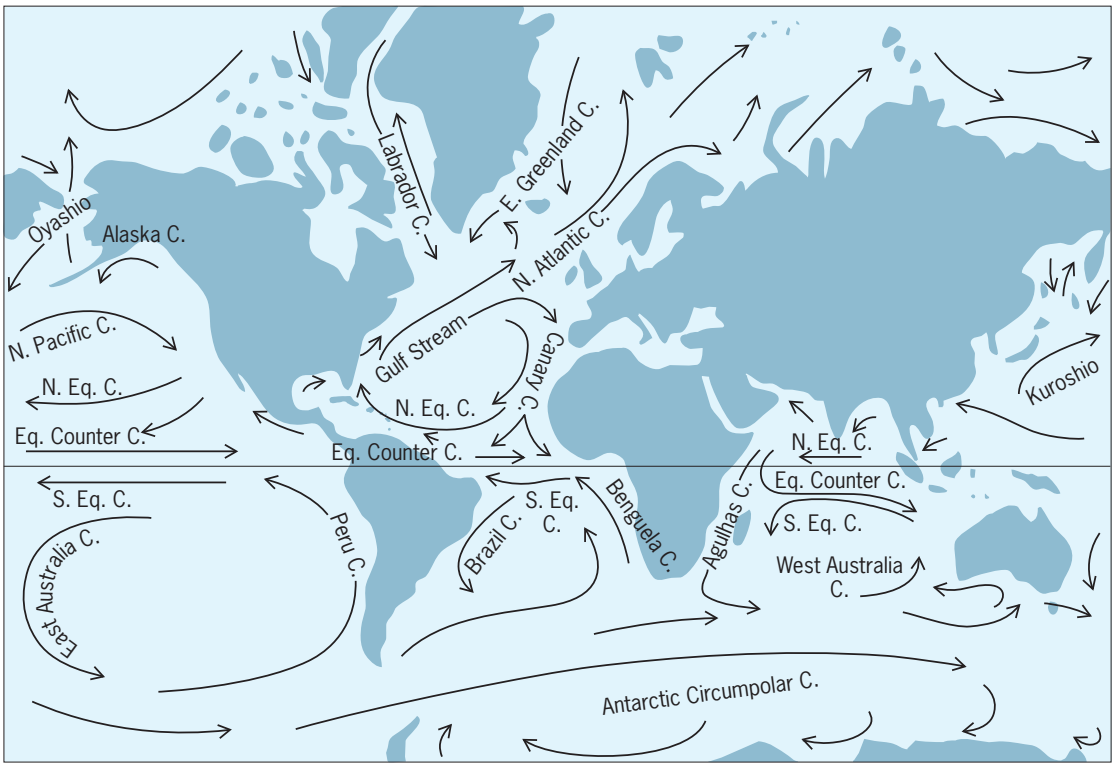

Fig. 3. Major currents of the world's oceans.

latitudes and from the east at the lower latitudes (trade winds). The frictional drag of these winds on the surface of the water imparts a spin or torque to the surface of the ocean, clockwise in the Northern Hemisphere and counterclockwise in the Southern Hemisphere (**Fig. 2**). The resulting basin-wide current gyres are apparent in the diagram of the major ocean currents in **Figure 3**, allowance being made for the shapes and boundaries of the different ocean basins. The major exception is the Indian Ocean north of the Equator, where the circulation is strongly influenced by the winds of the seasonal monsoon. *See* ATLANTIC OCEAN; CORIOLIS ACCELERATION; EQUATORIAL CURRENTS; INDIAN OCEAN; OCEAN CIRCULATION; PACIFIC OCEAN.

John A. Knauss

Bibliography. J. A. Knauss, *Introduction to Physical Oceanography*, 2d ed., Prentice Hall, 1997; M. E. Q. Pilson, *An Introduction to the Chemistry of the Sea*, Prentice Hall, 1998; E. I. Seibold and W. H. Berger, *The Sea Floor: An Introduction to Marine Geology*, 3d ed., Springer, 1996.

Ocean circulation

The general circulation of the ocean. The term is usually understood to include large-scale, nearly steady features, such as the Gulf Stream, as well as current systems that change seasonally but are persistent from one year to the next, such as the Davidson Current, off the northwestern United States coast and the equatorial currents in the Indian Ocean. A great number of energetic motions have periods of a month or two and horizontal scales of a few hundred kilometers—a very low-frequency turbulence, collectively called eddies. Energetic motions are also concentrated near the local inertial period (24 h, at 30° latitude) and at the periods associated with tides (primarily diurnal and semidiurnal). *See* TIDE.

Surface Circulation

The greatest single driving force for currents, as for waves, is the wind. Furthermore, the ocean absorbs heat at low latitudes and loses it at high latitudes. The resultant effect on the density distribution is

coupled into the large-scale wind-driven circulation. Some subsurface flows are caused by the sinking of surface waters made dense by cooling or high evaporation.

Wind stress. Air does not merely glide along the surface of the water, but exerts a frictional effect, or wind stress, which causes the surface water to be carried along with it. The movement of this thin layer on the surface of the water is conveyed by an internal turbulent friction to the deeper levels. The eventual result of such interaction, in a limitless homogeneous sea under the influence of a steady wind, would be a pure drift current. The resulting current distribution is illustrated by the so-called Ekman spiral (**Fig. 1**). There is a current at the sea surface at an angle to the right of the direction of the wind in the Northern Hemisphere. With increasing depth, the current turns farther toward the right and gradually subsides. When the direction of this current reaches an angle of 180° to the flow on the sea surface, the speed of the current is only 1/23 that of the surface water. This depth is called the depth of frictional influence. For example, at a latitude of 50° this depth amounts to about 200 ft (60 m) when the windspeed reaches 13.8 knots (7 m/s).

Observations show that the angle between wind and surface velocity is usually between 30 and 45°; the surface velocity is about 2% of the wind speed. Both these factors depend on how strongly the upper layers are stratified and how effectively the vertical turbulent motions transmit the stresses downward.

The wind-driven currents are in addition to any other currents that are present—from the large-scale density distribution, eddies, tides, and so on. One important feature of the wind-driven currents that is independent of stratification or turbulent intensity is that the total volume transport is at right angles to the wind. It is understandable, therefore, that upwelling of cold deep water occurs at a coast when the wind blows parallel to the coast, with the coast on the left-hand side of the wind. *See* UPWELLING; WIND STRESS.

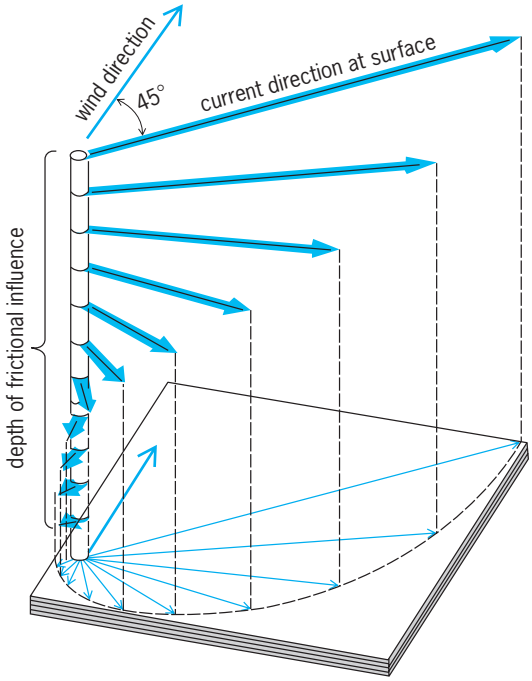

Fig. 1. Ekman spiral. The arrows show the direction and magnitude of the purely wind-driven current, as it changes with depth. The angle between the wind and the surface current is a function of the vertical mixing by turbulent motions; if this mixing is uniform with depth, the angle is 45° as shown.

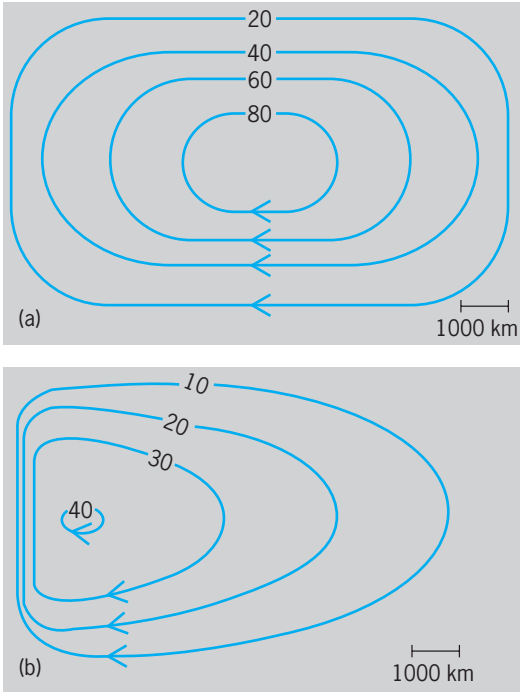

Fig. 2. Streamlines showing currents for (*a*) an ocean on a nonrotating globe and (*b*) an ocean on a uniformly rotating globe in which the Coriolis forces increase with the geographic latitude. 1000 km = 620 mi. (*After H. Stommel, The Gulf Stream, University of California Press, 1965*)

An outstanding feature of ocean currents is that intense flows (such as the Gulf Stream) occur on the western sides of the oceans. These flows were shown to result from the variation of Coriolis parameter with latitude (**Fig. 2**).

Surface currents. Except in western boundary currents, and in the Antarctic Circumpolar Current, the system of strong surface currents is restricted mainly to the upper 330–660 ft (100–200 m) of the sea. The mid-latitude anticyclonic gyres, however, are coherent in the mean well below 3300 ft (1000 m). The average speeds of the open-ocean surface currents remain mostly below 0.4 knot (20 cm/s). Exceptions to this are found in the western boundary currents, such as the Gulf Stream, and in the Equatorial Currents of the three oceans, all of which have velocities of 2–4 knots (1–2 m/s). Knowledge of the surface currents is based in part on direct measurements of the current, and more generally on dead reckoning from ships. *See* DEAD RECKONING.

The primary causes of surface currents are wind stress and internal pressure forces resulting from the density distribution. Frictional forces and the Coriolis acceleration influence the surface currents. The effect of the wind is greatest when the direction

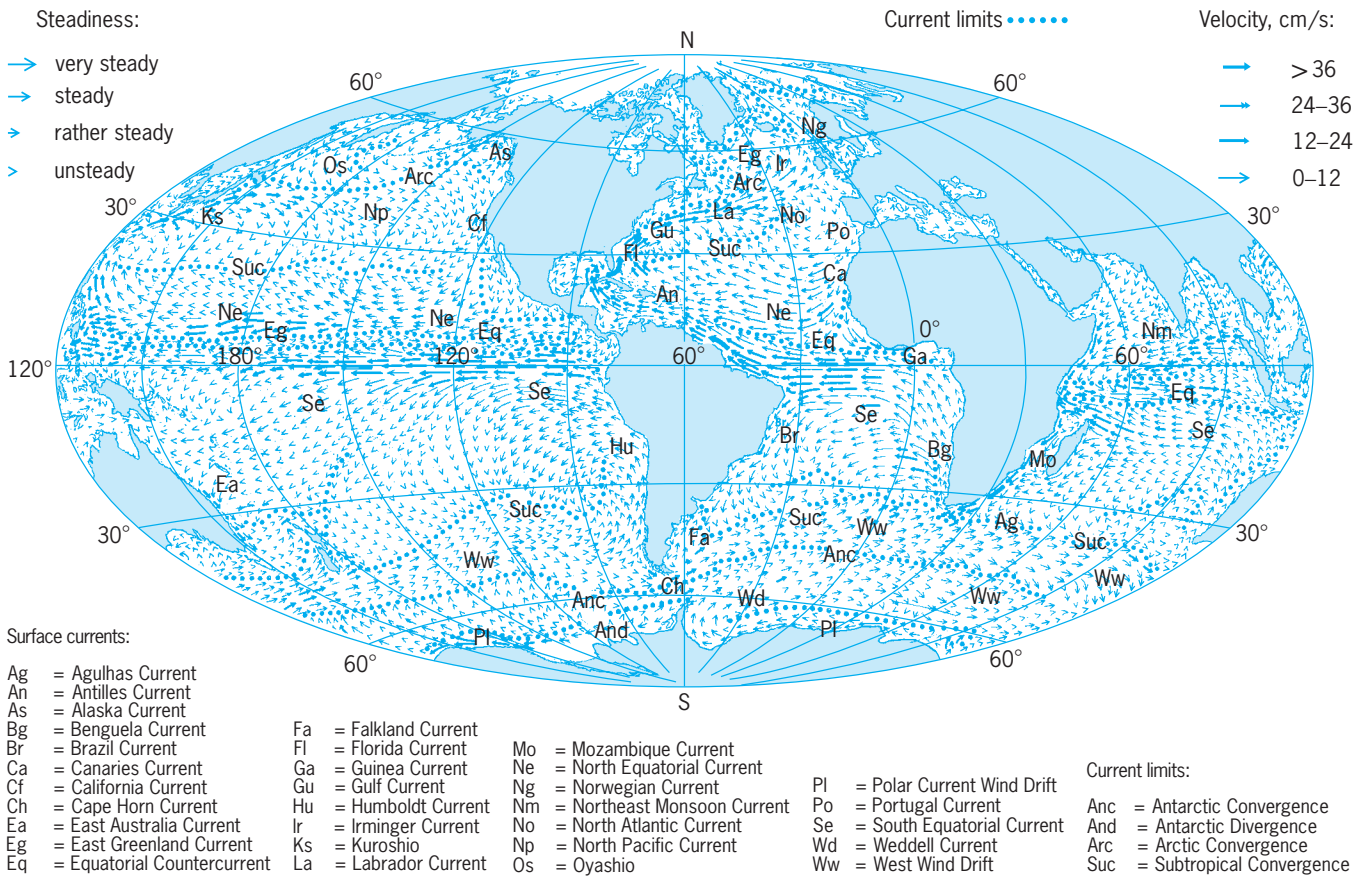

Surface currents:

Ag = Agulhas Current
An = Antilles Current
As = Alaska Current
Bg = Benguela Current
Br = Brazil Current
Ca = Canaries Current
Cf = California Current
Ch = Cape Horn Current
Ea = East Australia Current
Eg = East Greenland Current
Eq = Equatorial Countercurrent
Fa = Falkland Current
Fl = Florida Current
Ga = Guinea Current
Gu = Gulf Current
Hu = Humboldt Current
Ir = Irminger Current
Ks = Kuroshio
La = Labrador Current
Mo = Mozambique Current
Ne = North Equatorial Current
Ng = Norwegian Current
Nm = Northeast Monsoon Current
No = North Atlantic Current
Np = North Pacific Current
Os = Oyashio
Pl = Polar Current Wind Drift
Po = Portugal Current
Se = South Equatorial Current
Wd = Weddell Current
Ww = West Wind Drift

Current limits:

Anc = Antarctic Convergence
And = Antarctic Divergence
Arc = Arctic Convergence
Suc = Subtropical Convergence

Fig. 3. Surface currents of the oceans in February and March. 1 cm/s = 0.019 knot.

and strength of the wind are steady; this is the case in lower and middle latitudes. In these latitudes an anticyclonic (clockwise, in the Northern Hemisphere) current system corresponds to the anticyclonic wind system (**Fig. 3**). Surface currents which flow in a westerly direction in the lower latitudes are part of this system (the North and South Equatorial Currents). The continuation of these currents is found along the eastern sides of these continents in narrow and strong surface currents directed toward the poles—the Western Boundary Currents; the Gulf Stream, Brazil Current, and Somali Current (only in the summer of the Northern Hemisphere); and the Agulhas Current, Kuroshio, and East Australia Current. In middle latitudes these currents turn and flow in an easterly direction (North Atlantic Current, North Pacific Current, and West Wind Drift of the Southern Hemisphere). On the eastern sides of the oceans this pattern contains surface currents directed toward the Equator—the Eastern Boundary Currents; the Canary, Benguela, West Australia, California, and Humboldt currents.

Embedded in the system of the North and South Equatorial Currents, which flow to the west, are the Equatorial Countercurrents, flowing to the east. These are found about 5° north of the Equator in all three oceans. An additional easterly flowing countercurrent is sometimes found in the eastern Pacific Ocean, south of the South Equatorial Current.

A large subsurface current is found centered on the Equator. This Equatorial Undercurrent is about 240 mi (400 km) wide but only 660 ft (200 m) thick (centered at a depth of 500–660 ft or 150–200 m), flowing to the east with peak speeds of 2–3 knots (1–1.5 m/s). This current is absent in the Indian Ocean in Northern Hemisphere summer; in the Pacific it is sometimes called the Cromwell Current. *See* EQUATORIAL CURRENTS.

In higher latitudes the currents tend to follow the coasts and shelf edges; in the Northern Hemisphere the continents lie to the right-hand side of the current when one looks in the downstream direction. Examples are the Norwegian, East Greenland, West Greenland, Labrador, and Alaska currents. Islands in this fashion are, so to speak, surrounded by currents moving in a clockwise direction (for example, Iceland by the Irminger, North Iceland, and East Iceland currents). It is for this reason that the western sides of continents in these latitudes are bordered by comparatively warm waters coming from lower latitudes, whereas off the eastern coasts of the same latitudes there are cold waters from higher latitudes. For example, in the Norwegian Current the surface temperature in summer is 18°F (10°C) higher than in the East Greenland Current; both are at the same latitude. Thus surface currents are of great climatic importance.

All surface currents contain vertical components

that vary from region to region. These vertical current components are influenced by converging or diverging winds, acceleration of the currents, and other factors. The vertical components are certainly very small [for example, a large value would be a speed of 0.0012 in./s (0.003 cm/s) in the upwelling area of the California Current], but in the long run they are of great importance when considering the balance of heat and all sea-contained substances. *See* MARITIME METEOROLOGY; UPWELLING.

One important idea about the large mid-ocean gyres is that their transport is determined not simply by the wind stress but by the latitudinal variation of the east-west winds.

The maximum transport in the Florida Current occurs in June or July, almost 6 months out of phase. The seasonal variation is from about 9.5 to 12 × 10^8 ft^3/s (27 to 33 × 10^6 m^3/s) in volume transport in the Straits of Florida; this transport increases to approximately 35 × 10^8 ft^3/s (100 × 10^6 m^3/s) as the Stream passes the longitude of Cape Cod.

The idea has been advanced that a substantial portion of the transport of the Western Boundary Currents does not continue to the northeast but returns southerly in a weak countercurrent just offshore.

Current variability. Most of the major ocean currents have been known fairly well for some time. However, advances have been made in trying to understand the variations in ocean currents. The instantaneous path of the Gulf Stream (**Fig. 4**) has been determined by observations of its sharp surface-temperature gradient. The instantaneous path is much more sinuous than the long-term average path; the wavelike path variations will change entirely in a period of roughly 2 weeks to a month. Some wavelike meanders grow to have very large amplitudes, which become unstable and "pinch off" to form "Gulf Stream rings," or eddies (Fig. 4). The water in the center of a ring is cold or warm, depending merely upon the phase of the meander that pinched off. These eddies usually drift to the west or southwest.

South of Cape Hatteras, the Gulf Stream may be found going 45 to 90° away from its "normal" path. Downstream from Cape Hatteras, the meanders may grow to hundreds of miles away from the mean path, which becomes convoluted; a ship traveling perpendicular to the Stream, for example, could cross it three times.

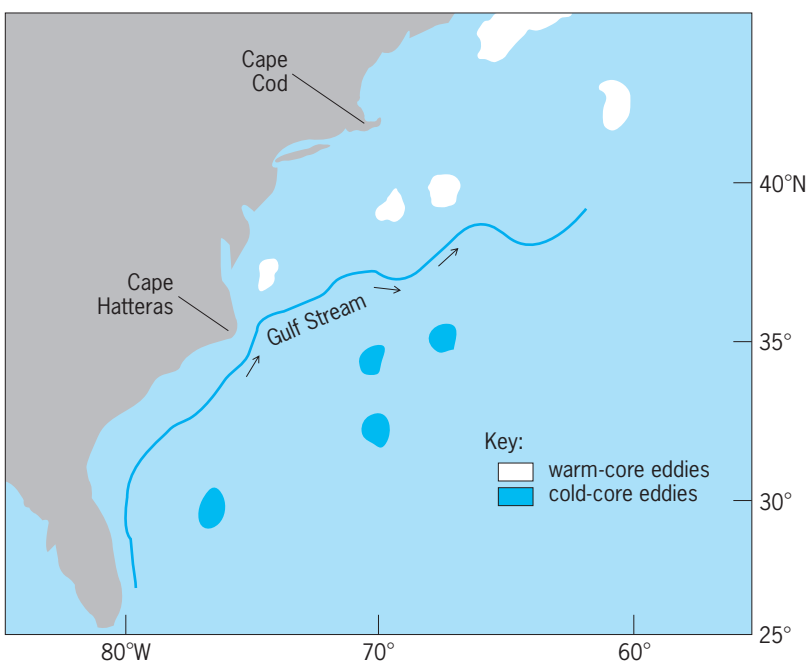

Fig. 4. Position of the Gulf Stream at the beginning of October 1975.

Deep Circulation

The deep circulation results in part from the wind stress and in part from the internal pressure forces which are maintained by the budgets of heat, salt, and water. Both groups of forces are dependent upon atmospheric influences. Apart from Coriolis and frictional forces, the topography of the sea bottom exercises a decisive influence on the course of deep circulation.

Marginal seas. The deep circulation in marginal seas depends largely on the climate of the region, whether arid or humid.

Arid climates. Under the influence of an arid climate, evaporation is greater than precipitation. The marginal sea (**Fig. 5**) is therefore filled with relatively salty water of a high density. Its surface lies at a lower level than that of the neighboring ocean. Examples of this type are the Mediterranean Sea, Red Sea, and Persian Gulf. At the connection between the two seas there is water of a slightly lower density from the ocean flowing in at the surface. The water from the marginal sea flows over the sill into the ocean, where it sinks to a level at which it finds water corresponding to its density. Substantial vertical mixing takes place during this initial flow. At the deeper level it then spreads horizontally. The waters from the Mediterranean Sea and the Red Sea, because of their high salinity, can be followed far out into the Atlantic and Indian oceans, respectively. *See* INDIAN OCEAN; MEDITERRANEAN SEA.

Humid climates. The deep circulation of marginal seas in humid climates shows a different pattern, however (Fig. 5*b*). The level of the sea is higher than in the neighboring ocean. Therefore, the surface water with its lower density and accordingly its lower salinity flows outward, and the relatively salty ocean water of higher density flows over the sill into the marginal sea. Examples of this circulation are the Baltic Sea with the shallow Darsser and Drogden rises, the Norwegian and Greenland fiords, and the Black Sea with its entrance through the Bosporus.

The Black Sea is an example of a special case. The sill depth always remains in the water of low density—that is, the outflowing upper layer. The renewal of deep water, and with it the deep circulation, comes to a complete halt. The result is that the oxygen is entirely used up, and poisonous hydrogen sulfide is generated. Below depths of 660 ft (200 m), only anaerobic organisms live in the Black Sea.

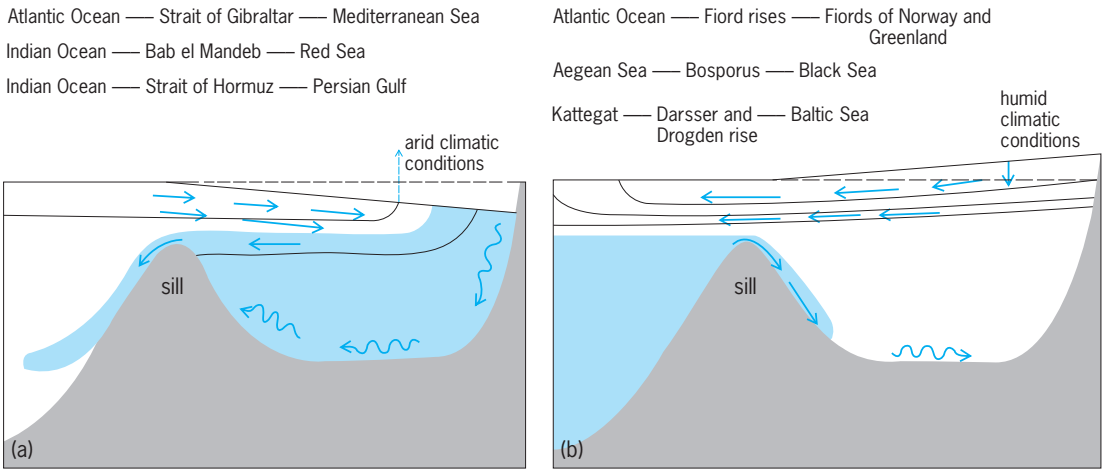

Fig. 5. Schematic representation of the circulation in marginal seas and over their sills (*a*) in arid climates, for example, Mediterranean Sea, Red Sea, and Persian Gulf; (*b*) in humid climates, for example, Black Sea, Baltic Sea, and fiords of Norway and Greenland.

If the sill depth interferes only occasionally with the lighter water of the upper levels, as in the entrances to the Baltic Sea, the renewal of deep water is interrupted at intervals. In the Baltic Sea these interruptions sometimes last for several years. *See* BALTIC SEA; BLACK SEA; FIORD.

Oceans. The deep circulation in the oceans is more difficult to perceive than the circulation in the marginal seas. In addition to the internal pressure forces, determined by the distribution of density and the piling up of water by the wind, there are also the influences of Coriolis forces and large-scale turbulence. There are areas in tropical latitudes in which the surface water, as a result of strong evaporation, has a relatively high density. In thermohaline convection, the water sinks while flowing horizontally until it reaches a density corresponding to its own, and then spreads out horizontally. In this way the colder and deeper levels of the oceans take on a layered structure consisting of the so-called bottom water, deep water, and intermediate water. In the Atlantic Ocean the deep-water circulation is strong on the western side of the ocean, where measurable speeds—roughly 4 in./s (10 cm/s)—are found (**Fig. 6**). The Bottom Water comes from very cold water massesthat have sunk along the edge of

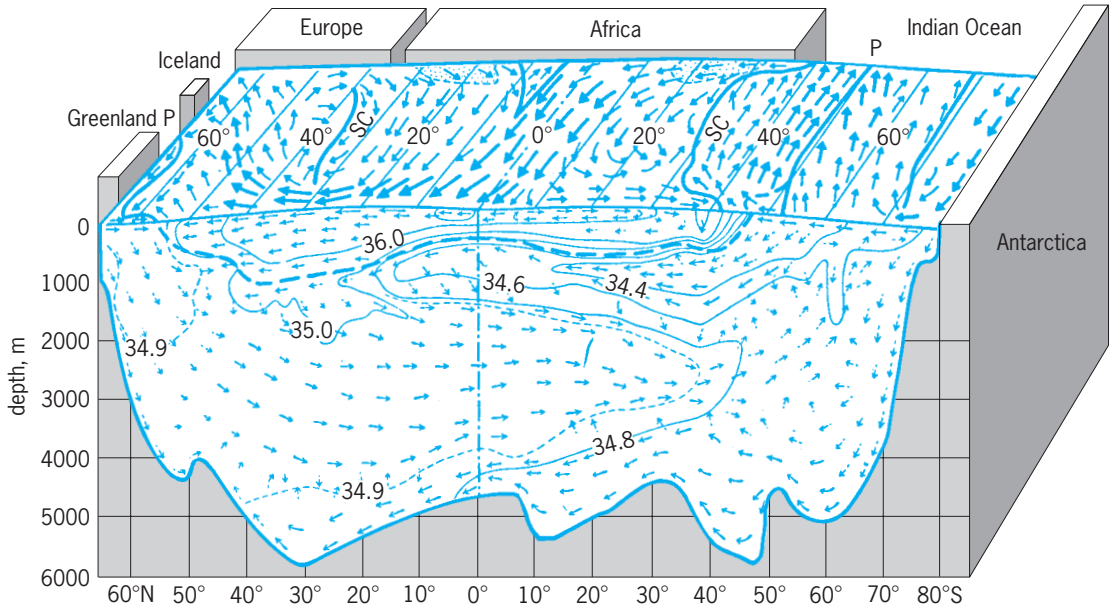

Fig. 6. Schematic representation of the surface and deep circulation in the Atlantic Ocean. All arrows show current directions; on the surface thin arrows indicate speeds of 0.1–0.8 knot (5–40 cm/s), and thick arrows indicate speeds of 0.8–2.9 knots (40–150 cm/s). SC indicates convergence of surface currents in subtropical waters. P indicates oceanic polar front where cold-water masses from polar and subpolar geographical latitudes meet relatively warm waters of temperate zone. In vertical section the heavy broken line shows division between warm- and cold-water spheres, and other lines indicate equal salinities. 1 m = 3.3 ft.

Antarctica. In the layer immediately above, or Deep Water, the water masses have their origin in the far North Atlantic. In the next layer above, or Intermediate Water, the water south of about 30°N comes from the Southern Hemispheric polar front.

There are five major areas where the surface water becomes denser and sinks.

1. The Norwegian Sea, where the water with the highest density of the world oceans is formed. This supplies the North Polar Sea with cold deep water, as well as the North Atlantic Ocean with cold bottom water to a latitude of approximately 50°N.

2. The Antarctic continental slope in the Weddell Sea, where water temperatures beneath the winter pack ice are near freezing (for seawater, −1.9°C or 28.6°F). Favored by the topography of the ocean floor, this water flows northward in the western Atlantic to the foot of the Grand Banks, as well as through the Romanche Deep on the Equator (through the Mid-Atlantic Ridge) into the eastern side of the Atlantic. It also spreads northward into the Indian and Pacific oceans, flowing toward the Equator on the western side of each ocean.

3. In the Labrador and Irminger seas, where the Deep Water (between depths of 3300 and 13,000 ft or 1000 and 4000 m) of the Atlantic Ocean originates. Apart from low temperatures, this water is distinguished by its richness in oxygen. This water has been traced around South America and into the Pacific Ocean.

4. The polar front at latitude of 50°S. Here cool water with low salinity is formed, supplying the Antarctic Intermediate Water of the Atlantic, Indian, and Pacific oceans.

5. The polar front in the North Pacific Ocean, where North Pacific Intermediate Water is formed. The distribution of this water in the North Pacific Ocean has been studied extensively.

The deep-sea circulation of the Atlantic Ocean appears to be the most active in comparison with the Indian and Pacific oceans, because the important sources of thermohaline convection are found in the Atlantic. In addition, the continental barrier of South America seems to force water from the surface currents of the South Atlantic into the North Atlantic Ocean. To compensate for the loss of surface water in the South Atlantic, there is a more active deep-water circulation, in which North Atlantic Deep Water flows southward into the South Atlantic Ocean. *See* ATLANTIC OCEAN; PACIFIC OCEAN. Wilton Sturges

Mesoscale Eddies

Wherever oceanographers have made long-term current and temperature measurements, they have found energetic fluctuations with periods of several weeks to several months. These low-frequency fluctuations (compared to tides) are caused by oceanic mesoscale eddies which are in many respects analogous to the atmospheric mesoscale pressure systems that form weather. Like the weather, mesoscale eddies often dominate the instantaneous current, and are thought to be an integral part of the ocean's general circulation.

Time and space scales. The most exhaustive field studies of mesoscale eddies have been conducted in the Sargasso Sea of the western North Atlantic. A wide variety of eddy types and sizes have been found, and probably others remain to be discovered. A typical Sargasso Sea eddy has a period of about 2 months and a diameter of about 210 mi (350 km). Some eddies are approximately circular in plan view, while others may be quite elongated or irregular. The eddy vertical structure is most commonly found to be baroclinic, with the largest currents near the sea surface, but barotropic eddies that have almost uniform currents from the sea surface to the ocean bottom have also been observed. Currents may be as large as 12 in./s (30 cm/s), and temperature surfaces displaced from their average depth by up to 330 ft (100 m). The sea surface displacement is about 4 in. (10 cm) and can be detected by satellite-borne radar altimeters.

The horizontal momentum balance for mesoscale eddies is well approximated by the geostrophic vertical shear. Eddy currents may thus be calculated from observations of the temperature and salinity field (which may be used to calculate density), and either an observed reference current or sea surface displacement.

Distribution and sources. Eddies occur in virtually all oceans and seas, but their amplitude varies greatly from place to place. The largest amplitudes are found on the western sides of the oceans in conjunction with the strongest ocean currents (the Gulf Stream in the North Atlantic, the Kuroshio in the North Pacific) and near the Equator. The typical Sargasso Sea eddy is representative of the powerful eddies found in such regions. Much weaker eddies are found in the ocean interior, distant from major currents.

This consistent pattern of eddy amplitude suggests that instabilities of western boundary currents are an important source of eddy energy. In some cases the instability results in a pinched-off current ring made up of the western boundary current water, but on other occasions it results in eddy motions in the surrounding water. Atmospheric forcing by variable winds can also generate eddies, and is probably most important at low latitudes where the horizontal scales of the oceanic eddies best match the scales of the atmospheric forcing.

Eddy dynamics. Eddies propagate and evolve in ways that depend in large part upon the eddy amplitude. If the eddy amplitude is small (currents less than a few centimeters per second) as it is over much of the eastern North Atlantic and Pacific, then the eddies tend to behave like linear planetary waves. Their phase speed is westward at typically 2 in./s (5 cm/s), and their group speed is eastward at a slightly smaller rate. If the eddy current amplitude is larger than the phase speed, for example, in the Sargasso Sea, then the eddies undergo continual scale change and exchange energy with neighboring eddies. This sort of chaotic, random behavior has the characteristics of turbulence. Mesoscale eddies that behave this way are a very effective mechanism for mixing heat, chemicals, and other tracers on an ocean basin-wide

scale. One of the most important research problems for contemporary oceanography is to learn how this eddy mixing affects the general circulation of the oceans. James F. Price

Measurement of Ocean Currents

One current-measuring method is to measure the flow of water past a point. Just as winds are measured by attaching a propeller-driven anemometer to a tower or building, current meters are attached to buoys anchored in the ocean, or lowered from anchored or floating ships. Unlike the wind anemometers attached to a fixed structure, floating ships move with the wind and current, and even anchored ships and buoys are not truly stationary. In the absence of any ocean current, these instruments measure the movement of the current meter through the water. Separating the movement of the current meter through the water from the velocity of the water past a fixed point can be a challenge. A variety of techniques have been employed either to minimize the movement of the meter or to accurately determine its movement, but all such measurements leave some uncertainty, and where the currents are weak, the order of a few hundredths of a meter per second, the uncertainty is often of the same order as the current that one is attempting to measure.

A second method measures the drift of a free body such as a drogue, a current-measuring device consisting of a weighted parachute and attached surface buoy. The earliest reliable current charts were constructed using a variation of this method, and much of the most recent information has been gathered by this technique. Systematic collection of ship drift data began in the midnineteenth century. The navigator would note the ship's noon position based on stars, sun , and dead reckoning, and calculate where the ship should be in 24 h. In the absence of very strong winds, the navigator would assume that the difference in the dead-reckoned position and the actual position revealed by stars and sun was a result of surface currents. By the end of the century, remarkably good charts of the average ocean surface currents were available.

Surface measurements. Surface drogues, carefully designed to minimize windage, equipped with Global Positioning System (GPS) receivers are the modern-day descendant of this technique. These surface drogues can report their position several times a day to a passing communication satellite, which in turn relays the drogues' positions to a central shore station. Thousands of these drogues are being deployed around the world, and surface current charts of unprecedented accuracy are now becoming available.

Subsurface measurements. Subsurface drifters measure currents at depth. The most common technique is for the container of the subsurface drifter to be less compressible than seawater. Weighted at the surface to be more dense than seawater, the density difference between seawater and float decreases as the float sinks until the densities match, at which pressure the float will remain. These subsurface floats can be designed to sink to a predetermined pressure surface with the equivalent accuracy of a few meters in depth. A more elegant technique is to design the container to have the same compressibility as seawater and to weigh the floats to sink to a given density surface. These so-called isopycnal (constant density) floats follow a given density surface rather than a given pressure surface and thus provide a better measure of the path of a given parcel of water at depth. *See* INSTRUMENTED BUOYS; SATELLITE NAVIGATION SYSTEMS.

Tracking subsurface floats is more difficult than tracking surface floats. One method is to design a float that can make itself lighter, return to the surface, and broadcast its position using the GPS signal to locate itself, as do the surface floats. These floats can be made to take on additional ballast, to sink, and to continue to follow along a given pressure surface. Subsurface floats of this type have been designed to float along at depth, come to the surface and signal their location, then sink again, as many as 150 times during a four-year period. By insuring that the submerged time is long compared to the time it takes for the float to move from depth to the surface and back, the aliasing of the subsurface current measurement is minimized. A second way of tracking subsurface floats is by using underwater sound. One method is for the subsurface float to carry a sound source and be tracked in a manner analogous to that used in tracking submarines. Another method is to anchor a number of buoys in the area where the floats are to be deployed, and attach sound sources to the buoys and recording hydrophones to the floats. The buoys have accurate clocks and transmit sound signals on a regular schedule that are recorded by the floats. The floats are programmed to come to the surface months or years later and transmit those recorded signals to a passing satellite, which in turn transmits those recordings to a shore station laboratory. Noting the difference in times of arrival of each of the transmitted signals and using simple triangulation geometry, each float's position can be determined as a function of time with an accuracy of better than a nautical mile.

Geostrophic current. Historically, much of the information about currents below the ocean surface comes not from direct measurement, but from the assumption that these currents are in geostrophic balance (where the Coriolis force balances exactly the horizontal pressure gradient). By careful measurement of temperature and salinity as a function of depth at two locations, one can determine seawater density, and in turn calculate the hydrostatic pressure as a function of depth. The difference in density at the two stations results in differeneces in hydrostatic pressure, which in turn allows one to calculate as a function of depth changes in the geostrophic current normal to an imaginary line connecting the two stations. Recent improvements in satellite altimeters now allow one to measure differences of a few centimeters in sea level surface. Knowing the slope of the sea surface with respect to its equal-potential surface, one can calculate the geostrophic surface

current. *See* CORIOLIS ACCELERATION; OCEANOGRAPHY; SEAWATER. John A. Knauss

Bibliography. J. R. Apel, *Principles of Ocean Physics*, 2d ed., 2001; R. A. Davis, Jr., *Oceanography: An Introduction to the Marine Environment*, 2d ed., 1996; W. J. Emery (ed.), *Descriptive Physical Oceanography*, 5th ed., 1982; T. S. Garrison, *Oceanography: An Invitation to Marine Science*, 2d ed., 1996; M. N. Hill (ed.), *The Sea*, vol. 1, 1982; F. W. Smith and F. A. Kalber (eds.), *Handbook in Marine Science*, 2 vols., 1974; H. Stommel, *A View of the Sea: A Discussion Between a Chief Engineer and an Oceanographer about the Machinery of the Ocean Circulation*, 1991; K. Stowe, *Essentials of Ocean Science*, 1988; T. Teramoto, *Deep Ocean Circulation: Physical and Chemical Aspects*, 1993.

Ocean waves

Propagating oscillations in the ocean which carry energy and momentum from one region to another. Most ocean waves are caused directly or indirectly by wind blowing across the sea surface. Many waves can propagate through the ocean thousands of miles from where they are generated.

Surface waves. Ocean surface waves are propagating disturbances at the atmosphere-ocean interface. They are the most familiar ocean waves. Surface waves are also seen on other bodies of water, including lakes and rivers. *See* WAVE MOTION IN LIQUIDS.

A simple sinusoidal wave train is characterized by three attributes: wave height (H), the vertical distance from trough to crest; wavelength (L), the horizontal crest-to-crest distance; and wave period (T), the time between passage of successive crests past a fixed point. The phase velocity ($C = L/T$) is the speed of propagation of a crest. For a given ocean depth (h), wavelength increases with increasing period. The restoring force for these surface waves is predominantly gravitational. Therefore, they are known as surface gravity waves, unless their wavelength is shorter than 1.8 cm (0.7 in.), in which case surface tension provides the dominant restoring force.

Classification. Surface gravity waves may be classified according to the nature of the forces producing them. Tides are ocean waves induced by the varying gravitational influence of the Moon and Sun. They have long periods, usually 12.42 h for the strongest constituent. Storm surges are individual waves produced by the wind and dropping barometric pressure associated with storms; they characteristically last several hours. Earthquakes or other large, sudden movements of the Earth's crust can cause waves, called tsunamis, which typically have periods of less than an hour. Wakes are waves resulting from relative motion of the water and a solid body, such as the motion of a ship through the sea or the rapid flow of water around a rock. Wind-generated waves (**Fig. 1**), having periods from a fraction of a second to tens of seconds, are called wind waves. Like tides, they are ubiquitous in the ocean,and continue to travel well beyond their area of generation. The ocean is never completely calm. *See* STORM SURGE; TIDE; TSUNAMI.

Fig. 1. Open-ocean wind waves driven by gale-force (18 m/s or 40 mi/h) winds.

Wind waves. The growth of wind waves by the transfer of energy from the wind is not fully understood. At wind speeds less than 1.1 m/s (2.5 mi/h), a flat water surface remains unruffled by waves. Once generated, waves gain energy from the wind by wave-coupling of pressure fluctuations in the air just above the waves. For waves traveling slower than the wind, secondary, wave-induced airflows shift the wave-induced pressure disturbance downwind so the lowest pressure is ahead of the crests. This results in energy transfer from the wind to the wave, and hence growth of the wave.

If a constant wind blows over a sufficient length of ocean, called the fetch, for a sufficient length of time, a wave field develops whose statistical characteristics depend only on wind velocity. In particular, the spectrum of sea-surface elevation for such a fully-developed sea has the form of Eq. (1), where f

$$S(f) = A\frac{g^2}{f^5}e^{-1.25(f_m/f)^4} \qquad (1)$$

is frequency ($= 1/T$), $g = 9.8$ m/s^2 (32 ft/s^2) is gravitational acceleration, $f_m = 0.13\, g/U$ is the frequency of the spectral peak (U = wind speed at 10 m or 32.8 ft elevation), and $A = 5.2 \times 10^{-6}$ is a constant.

The fetch is limited near a coast with the wind blowing offshore, and the waves grow as they propagate toward the open ocean. In such a limited-fetch situation, Eq. (1) is modified: A and f_m become dependent on the fetch length, and the peak in the spectrum is enhanced. **Figure 2** shows spectral forms of waves generated by a moderate breeze for various fetches and in the open sea. For faster wind speeds, the spectral peaks grow in height and shift to lower frequencies.

Even when the mean wind blows from a single direction, the surface waves that it generates are seen to travel in a variety of directions centered on the downwind direction. The directional spectrum of such a wave field can be approximately represented by a formula, such as Eq. (2), where θ is the angle

$$S(f,\theta) = \begin{cases} S(f)\frac{2}{\pi}\cos^2\theta & \text{if } \theta \le 90^\circ \\ 0 & \text{if } \theta > 90^\circ \end{cases} \qquad (2)$$

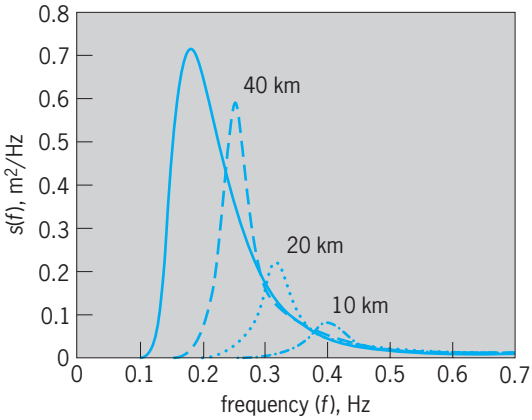

Fig. 2. Spectra of waves from a 7 m/s (16 mi/h) wind in the open sea (solid line), and for three limited fetches: 40, 20, 10 km (25, 12, 6 mi). (*Theoretical relations are from W. J. Pierson and L. Moskovitz, J. Geophys. Res., 69:5181–5190, 1964, for the open sea, and K. Hasselmann et al., Deutsch. Hydrogr. Z., Reihe A (8°), Nr. 12, 1973, for fetch-limited seas*)

of wave propagation relative to the downwind direction.

An observer asked to estimate average wave height typically gives a value that is about the average height of the highest one-third of the waves actually present. This statistic, represented as $H_{1/3}$, is called the significant wave height. In the open sea, Eq. (3) applies.

$$H_{1/3} = 0.24\frac{U^2}{g} \tag{3}$$

If a wind blows steadily over a known fetch for a period of time, the resulting $H_{1/3}$ may be estimated from **Fig. 3**.

Because of viscosity, surface waves lose energy as they propagate, short-period waves being dampened more rapidly than long-period waves. Waves with long periods (typically 10 s or more) can travel thousands of kilometers with little energy loss. Such waves, generated by distant storms, are called swell. Equations (1), (2), (3) and Figs. 2 and 3 assume that the waves present were generated by local wind, with no significant swell present. *See* VISCOSITY.

The highest wind waves are produced by large intense storm systems that last for a day or longer. Such systems of very low atmospheric pressure form in the Gulf of Alaska, the region around Iceland, and the Weddell Sea. Off the west coast of Canada, there have been several measurements of individual waves with heights around 30 m (100 ft). Northwest of Hawaii, on February 7, 1933, the Navy tanker USS *Ramapo* encountered the largest open-ocean wind waves ever reliably observed with heights that were reported to be at least 34 m (112 ft).

When waves propagate into an opposing current, they grow in height. For example, when swell from a Weddell Sea storm propagates northeastward into the southwestward-flowing Agulhas Current off South Africa, high steep waves are formed. Many large ships in this region have been severely damaged by such waves.

Because actual ocean waves consist of many components with different periods, heights, and directions, occasionally a large number of these components can, by chance, come in phase with one another, creating a freak wave with a height several times the significant wave height of the surrounding sea. According to linear theory, waves with different periods propagate with different speeds in deep water, and hence the wave components remain in phase only briefly. But nonlinear effects are bound to be significant in a large wave. In such a wave, the effects of nonlinearity can compensate for those of dispersion, allowing a solitary wave to propagate almost unchanged. Consequently, a freak wave can have a lifetime of a minute or two. *See* SOLITON.

Linear theory. For waves sufficiently small that linear theory applies, Eq. (4), gives the phase velocity

$$C = \begin{cases} (g/k)^{1/2} & \text{deep water} \quad h > 0.4L \\ (gh)^{1/2} & \text{shallow water} \quad h < 0.04L \end{cases} \tag{4}$$

to an accuracy of 1%. Note that wavenumber $k = 2\pi/L$. *See* WAVE MOTION IN LIQUIDS.

Ocean wave energy E per unit surface area depends only on wave height [Eq. (5)]. This energy propagates at the group velocity [Eq. (6)] and produces a power flux per unit distance along the wave

$$E = \frac{1}{8}\rho g H^2 \tag{5}$$

$$C_g = \begin{cases} \frac{1}{2}C & \text{deep water} \\ C & \text{shallow water} \end{cases} \tag{6}$$

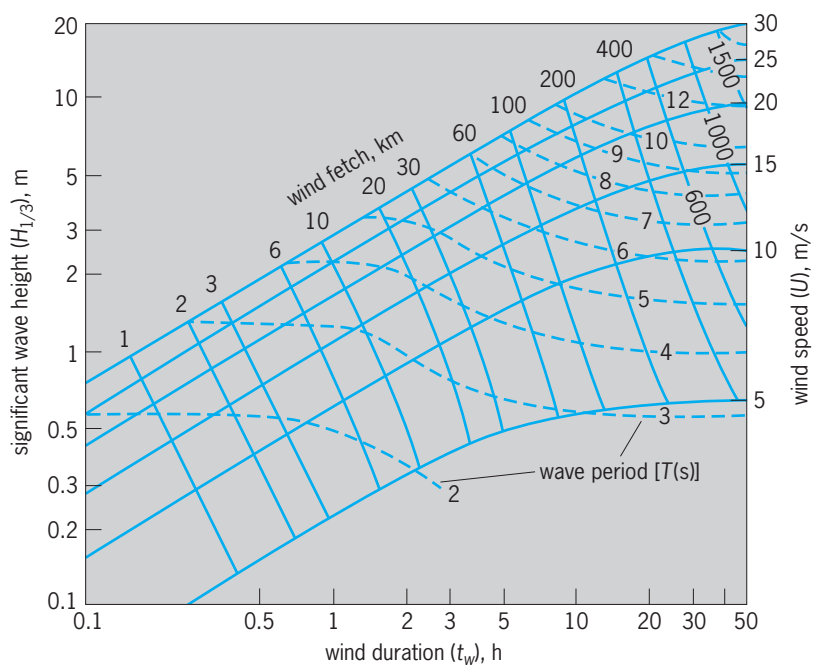

Fig. 3. Relationship of significant wave height $H_{1/3}$ and characteristic wave period T_m ($=1/f_m$) to wind speed, duration, and fetch. To estimate the wave characteristics resulting from a wind of velocity U which has blown steadily for a time t_w, find one of the six heavy curves labeled on the right with the appropriate U, follow this curve to the correct t_w, and, at that point, read the value of $H_{1/3}$. The position of this point, relative to the broken curves, indicates the wave period T_m. If the waves are fetch-limited, $H_{1/3}$ grows with time only until the heavy curve intersects the appropriate fetch line, at which point $H_{1/3}$ and T_m remain fixed. Open-sea conditions exist on the right side of the diagram. (*After World Meteorological Organization, Guide to Wave Analysis and Forecasting, no. 702, 1988*)

front [Eq. (7)]. For example, waves in the deep sea

$$P = C_g E \tag{7}$$

with 1.8-m (6-ft) height and 10-s period carry a power flux of 30 kW/m (13 hp/ft). Available power fluxes of this magnitude are representative of many coastal regions, and represent a substantial renewable energy resource. Various devices with efficiencies of 50% or greater have been developed to convert this wave energy to electrical energy. *See* HYDROELECTRIC GENERATOR.

Surface-gravity waves cause pressure fluctuations and particle motions that are largest near the surface and decrease with depth. In deep water, this dependence on depth d below the still-water level is exponential: e^{-kd}.

Stokes drift. To first-order (linear theory), particle orbits resulting from wave motion are closed loops. But since a particle moves in the direction of wave propagation in the loop's upper part and in the reverse direction in its lower part, the forward motion is at a shallower level than the reverse motion. Consequently, the forward motion is slightly stronger than the reverse motion, and the orbital loops do not quite close. As a result, there is a second-order (nonlinear) net particle motion called Stokes drift. In deep water, the velocity associated with this motion is in the direction of wave propagation [Eq. (8)]. In

$$\bar{u} = C(ak)^2 e^{-2kd} \tag{8}$$

shallower water, Stokes drift contributes to sediment transport onto beaches, and along beaches when the waves approach the coast at an angle. *See* NEARSHORE PROCESSES.

Shoaling and breaking. Waves approaching the shore from the open ocean are affected in several ways. As they become shallow-water waves, velocity C decreases with decreasing water depth h [Eq. (4)]. Consequently, waves approaching the shore at an angle are refracted so their crests are brought nearly parallel to the shoreline. Wave period T does not change as h decreases, so wavelength $L = CT$ must decrease according to Eq. (4). Shallowing also causes a growth in wave height as in relation (9), where the

$$H \propto h^{-r} \tag{9}$$

value of r depends on the character of the waves. For example, in the absence of dissipation, $r = 1/4$ according to linear theory for shallow-water waves [Eqs. (4)–(7)], while $r = 1$ for solitary waves of moderate amplitude.

As wave height grows, nonlinear terms in the equations of motion become significant. At first, this results in a vertical asymmetry in the wave shape, with crests becoming more peaked and troughs more rounded. Then, as the wave moves into shallow water, a strong horizontal asymmetry develops in which the forward face of the wave becomes progressively steeperthan the backward face. This process typically continues until the wave breaks. The resulting breakers can have a variety of forms (collapsing, plunging, surging, and so forth), depending on the height of the entering waves and the slope of the seabed. A rough criterion for breaking is shown in relation (10), or, using Eq. (4) for shallow water, $H > h$.

$$H > \frac{C^2}{g} \tag{10}$$

In the open ocean also, waves often break (Fig. 1). In this case, the criterion given by relation (10) becomes $H > 1/k$ [using Eq. (4) for deep water]. Whitecaps from deep-water breaking waves begin to appear at wind speeds of about 0.45 m/s (10 mi/h), while at wind speeds above 27 m/s (60 mi/h) all the high waves are breaking.

Wave measurement. There are three classes of instruments for measuring ocean surface waves: those which are at the air-sea interface, those which are below it, and those which are above it. Because so many techniques can be used, only a few typical examples of each class will be described.

At the surface. Ocean surface waves can be measured from a dock with a wave staff held vertically in the water. The varying position $\eta(t)$ of the sea surface on the staff is sensed in a variety of ways, for example, by seawater-shorting of the submerged part of a resistance wire wound along the staff.

An accelerometer-instrumented buoy on a slack mooring line can provide a record of $\eta(t)$ by double integration of the vertical acceleration. If, in addition, the buoy contains tilt sensors, it is capable of providing the directional spectrum $S(f,\theta)$ of the waves. Typically, data from a wave-measuring buoy are telemetered to a receiving station on shore. *See* INSTRUMENTED BUOYS.

Below the surface. The most common instrument of this class is the subsurface pressure sensor. The pressure measurement must be made at a level deep enough that it is always submerged; this has the advantage of reducing vulnerability to damage by ships and breaking waves. The main disadvantage of the method is the need to compensate for the frequency-dependent depth attenuation of the measured wave-induced pressure signal $p(t)$ in converting it to surface elevation $\eta(t)$.

A narrow-beam inverted echo sounder can also be used to make subsurface wave measurements. It is placed on the sea floor and directed upward, so the acoustic echo time from the surface is a measure of sea-surface elevation. But variations in temperature and salinity affect the speed of sound in water, and hence affect instrument calibration. Also, bubbles in the water can cause spurious acoustic reflections. *See* ECHO SOUNDER.

Above the surface. Ground-based high-frequency (3–30 MHz) radar systems can provide information on wave height and direction from the backscattered signal. Ranges of 50–500 km (30–300 mi) are feasible, but with over-the-horizon sky-wave systems relying on ionospheric reflection, ranges beyond 3200 km (2000 mi) have been achieved.

TABLE 1. Sea height code*

Code	Height, ft†	Description of sea surface
0	0	Calm, with mirror-smooth surface
1	0–1	Smooth, with small wavelets or ripples with appearance of scales but without crests
2	1–3	Slight, with short pronounced waves or small rollers; crests have glassy appearance
3	3–5	Moderate, with waves or large rollers; scattered whitecaps on wave crests
4	5–8	Rough, with waves with frequent whitecaps; chance of some spray
5	8–12	Very rough, with waves tending to heap up; continuous whitecapping; foam from whitecaps occasionally blown along by wind
6	12–20	High, with waves showing visible increase in height, with extensive whitecaps from which foam is blown in dense streaks
7	20–40	Very high, with waves heaping up with long frothy crests that are breaking continuously; amount of foam being blown from the crests causes sea surface to take on white appearance and may affect visibility
8	40+	Mountainous, with waves so high that ships close by are lost from view in the wave troughs for a time; wind carries off crests of all waves, and sea is entirely covered with dense streaks of foam; air so filled with foam and spray as to affect visibility seriously
9		Confused, with waves crossing each other from many and unpredictable directions, developing complicated interference pattern that is difficult to describe; applicable to conditions 5-8

* After *Instruction Manual for Oceanographic Observations*, H. O. Publ. 607, 2d ed., U.S. Navy Hydrographic Office, 1955.
† 1 ft = 0.3 m.

From aircraft, stereo-photographs can be taken of the sea surface and analyzed photogrammetrically, but this is a laborious process. Also, laser or narrow-beam radar ranging may be used to measure profiles of the sea surface. *See* LIDAR; PHOTOGRAMMETRY; RADAR.

Two radar techniques are presently used for wave measurements from satellites. The radar altimeter (13.5 GHz) observes the reflection of pulses directed vertically. The deformation of the reflected pulse carries information about significant wave height $H_{1/3}$ in the irradiated patch of ocean (several kilometers in diameter). The synthetic-aperture radar (SAR; >1 GHz) obliquely irradiates a patch of ocean surface (about 100 km or 60 mi in size), and uses pulse timing and phase information in the backscattered signal to obtain spatial resolution approaching 10 m (30 ft). The backscattering is from sea-surface roughness on scales comparable to the radar wavelength (several centimeters or inches), but longer-wavelength components appear in the SAR image by hydrodynamic interaction, electromagnetic modulation, and effects of wave motion. As a result, it appears that directional wave spectra $S(f,\theta)$ may be obtained from SAR images, but the procedure is not yet fully understood. *See* METEOROLOGICAL SATELLITES.

Mark Wimbush

Sea state. Sea state is the description of the ocean surface or state of the sea surface with regard to wave action. Wind waves in the sea are of two types: those still growing under the force of the wind are called sea: those no longer under the influence of the wind that produced them are called swell. Differences between the two types are important in forecasting ocean wave conditions.

Sea. Those waves which are still growing under the force of the wind have irregular, chaotic, and unpredictable forms. The unconnected wave crests are only two to three times as long as the distance between crests and commonly appear to be traveling in different directions, varying as much as 20° from the dominant direction. As the waves grow, they form regular series of connected troughs and crests with wave lengths commonly ranging from 12 to 35 times the wave heights. Wave heights only rarely exceed 55 ft (17 m). The appearance of the sea surface is termed state of the sea (**Table 1**).

The height of a sea is dependent on the strength of the wind, the duration of time the wind has blown, and the fetch (distance of sea surface over which the wind has blown).

Swell. As sea waves move out of the generating area into a region of weaker winds, a calm, or opposing winds, their height decreases as they advance, their crests become rounded, and their surface is smoothed. These waves are more regular and more predictable than sea waves and, in a series, tend to show the same form or the same trend in characteristics. Wave lengths generally range from 35 to 200 times wave heights.

The presence of swell indicates that recently there may have been a strong wind, or even a severe storm, hundreds or thousands of miles away. Along the coast of southern California long-period waves are believed to have traveled distances greater than 5000 mi (8000 km) from generating areas in the South Pacific Ocean. Swell can usually be felt by the roll of a ship, and, under certain conditions, extremely long and high swells in a glassy sea may cause a ship to take solid water over its bow regularly.

When swell is obscured by sea waves, or when the components are so poorly defined that it is impossible to separate them, it is reported as confused. (**Table 2**).

In-between state. Often both sea waves and swell waves, or two or more systems of swell, are present in the same area. When waves of one system are superimposed upon those of another, crests may coincide with crests and accentuate wave height, or troughs may coincide with crests and cancel each other to produce flat zones. This phenomenon is known as wave interference, and the wave forms produced are extremely irregular. When wave systems cross each other at a considerable angle, the

TABLE 2. Swell-condition code*

Code	Description	Height, ft†	Length, ft†
0	No swell	0	0
	Low swell	1–6	
1	Short or average		0–600
2	Long		600+
	Moderate swell	6–12	
3	Short		0–300
4	Average		300–600
5	Long		600+
	High swell	12+	
6	Short		0–300
7	Average		300–600
8	Long		600+
9	Confused		

*After *Instruction Manual for Oceanographic Observations*, H. O. Publ. 607, 2d ed., U.S. Navy Hydrographic Office, 1955.
†1 ft = 0.3 m.

apparently unrelated peaks and hollows are known as a cross sea.

Breaking waves. The action of strong winds (greater than 12 knots or 6.2 m/s) sometimes causes waves in deeper water to steepen too rapidly. As the height-length ratio becomes too large, the water at the crest moves faster than the crest itself and topples forward to form whitecaps.

As waves travel over a gradually shoaling bottom, the motion of the water is restricted and the wave train is telescoped together. The wave length decreases, and the height first decreases slightly until the water depth is about one-sixth the deep-water wave length and then rapidly increases until the crest curves over and plunges to the water surface below. Swell coming into a beach usually increases in height before breaking, but wind waves are often so steep that there is little if any increase in height before breaking. For this reason, swell that is obscured by wind waves in deeper water often defines the period of the breakers.

The zone of breakers, or surf, includes the region of white water between the outermost breaker and the waterline on the beach. If the sea is rough, it may be impossible to differentiate between the surf inshore and the whitecaps in deep water just beyond. Neil A. Benfer

Capillary waves. Capillary waves, or ripples, occur at the interface between two fluids when the principal restoring force is surface tension. Ripples generated by wind on the ocean and lakes are important for the initiation of turbulence in both media, transfer of gases between air to water, and scattering of electromagnetic and sound waves.

Capillary waves have short wave lengths. Both the phase and group velocities increase as wave lengths become shorter, and group velocity is greater than phase velocity. Dissipation of the waves by molecular viscosity is very rapid. The characteristic shape of the water surface, sinusoidal for small amplitudes, becomes distorted with more sharply curved troughs than crests as wave amplitude increases. In all these respects, ripples contrast sharply with gravity waves.

Mathematical formulas have been derived that relate the phase velocity, group velocity, and frequency to the wave length of low-amplitude waves on still water in the absence of wind. In such waves, gravity and surface tension play equal roles in the restoring force; longer waves are considered to be gravity waves, and shorter are capillary waves. *See* SURFACE TENSION.

In nature, ripples are observed to grow rapidly when the wind blows, and to die away rapidly when the wind stops. When the water surface is uncontaminated, ripples die away to 37% of their original amplitude in a period of time that is related to the wave length and the kinematic viscosity of water. For example, for a wave with a wave length of 0.67 in. (17 mm) and a kinematic velocity of 10 m/s, the period of time is 3.8 s. When the water surface is contaminated, as by an oil film or other surface-active agent, ripples are damped still more rapidly, because the contaminated surface tends to act as an inextensible film against which the water motions due to the ripples must rub. In such a case for the example given, the period of time becomes 0.86 s.

Under low-wind conditions, this increased damping almost completely inhibits ripple growth: the surface appears glassy smooth and is called a slick. It has been observed that even short gravity waves grow at an inappreciable rate under such conditions; the interpretation here is that the fine scale of roughness presented to the wind by a rippled surface is involved in the formation and growth of gravity waves. However, ripples have been observed to form on clean water in the absence of wind by nonlinear processes occurring at the sharp crests of short, steep gravity waves; consequently the formation of both gravity waves and ripples is an interconnected process. Ripple trains formed under low-wind conditions derive their energy at the expense of gravity waves and are called parasitic capillaries. In this case, the capillary wave train is propagating in a moving stream of water, the orbital current of the underlying gravity wave. The longer wave parts of the capillary train would have a lower phase velocity in still water than either the shorter wave-length parts of the train or of the gravity wave, but they maintain their position relative to the crest of the gravity wave because they ride in a favorable part of the orbital current of the gravity wave. Short capillaries, having high phase velocity, do not need this aid to keep in phase, and their higher group velocity enables them to proceed to a leading position where the gravity-wave orbital velocity vanishes or even opposes their motion. Charles S. Cox

Internal waves. Internal waves are wave motions of stably stratified fluids in which the maximum vertical motion takes place below the surface of the fluid. The restoring force is mainly due to gravity; when light fluid from upper layers is depressed into the heavy lower layers, buoyancy forces tend to return the layers to their equilibrium positions. Internal waves have been found in the atmosphere as lee waves (waves in the wind stream downwind from a mountain) and as waves propagated along an inversion layer (a layer of very stable air). They are also associated with wind shears at the lower

boundary of the jet stream. In the oceans, internal oscillations have been observed wherever suitable measurements have been made. The observed oscillations can be analyzed into a spectrum with periods ranging from a few minutes to days. At a number of locations in the oceans, internal tides, or internal waves having the same periodicity as oceanic tides, are prominent.

Internal waves are important to the economy of the sea because they provide one of the few processes that can redistribute kinetic energy from near the surface to abyssal depths. When they break, they can cause turbulent mixing despite the normally stable density gradient in the ocean. Internal waves are known to cause time-varying refraction of acoustic waves because the sound velocity profile in the ocean is distorted by the vertical motions of internal waves. The result is that quasi-horizontal propagation of sound shows phase incoherence and large changes of intensity with time at ranges where the refraction has led to divergence or convergence of rays.

The vertical distribution of motions and phase velocity of internal waves depends on the vertical gradient of density in the fluid and the frequency of the generating forces. There is a simple density distribution that is illustrative: The fluid consists of two homogeneous layers, a lighter one on top of a heavier one, such as kerosine over water. The internal waves in this system are sometimes called boundary waves, because the maximum vertical motion occurs at the discontinuity of density at the boundary between the two fluids. Internal waves move at a slow speed, of the order of a few knots in the deep oceans. The effect of the rotation of the Earth is to increase the phase velocity of waves having periods long enough to approach one pendulum day.

When there is a continuous distribution of density in a fluid, as in the ocean or atmosphere, internal waves are possible only for frequencies that are lower than the maximum value given by a mathematical relationship known as the Väisälä-Brunt frequency, which is related to the downward rate of increase of the density and the velocity of sound in the fluid. In the ocean this maximum frequency occurs in the thermocline, where it commonly amounts to about one-fifth cycle per minute. At any frequency lower than this limit, there is an infinity of possible modes of internal waves. In the first mode, the vertical motion has a single maximum somewhere in the body of the fluid; in the second mode, there are two such maxima (180° out of phase), with a node between; and so on. The actual motion usually consists of a superposition of modes. In the ocean, the Väisälä-Brunt frequency varies from a maximum, commonly near 0.2 cycle per minute in the steepest part of the thermocline to negligible values at the bottom of the deep seas and in mixed layers such as those at the sea surface. Amplitudes of waves are oscillatory with depth and appreciable only where their frequency is less than the local value of the Väisälä-Brunt frequency. For this reason, internal waves of high frequency are limited in depth to the thermocline. Long-period (low-frequency) internal waves are affected by Earth rotation, and the limiting periodicity is the local value of the pendulum day (one-half sidereal day divided by the sine of the latitude). These longest-period oscillations are inertial motions, in which the orbital path no longer has any vertical component but forms a circle in the horizontal.

In a continuously stratified fluid such as the ocean below the thermocline, internal waves propagate diagonally upward and downward, thus distributing energy throughout the ocean depths. Their ray paths are easily distorted because the group velocity of internal waves is comparable to the differences in velocity in the vertical of ocean currents and inertial motions. As a consequence, bundles of internal wave rays suffer severe refraction, sometimes to the extent that they form caustics known as critical layers, where the wave energy is absorbed by breaking. Where the sea bed is level, internal waves can be reflected by surface and sea floor to form normal modes within the body of the ocean. On a sloping sea floor, reflection at the bottom causes a change of wave length.

Internal waves in the atmosphere have been detected by a variety of instruments: microbarographs and wind recorders at ground level, and long-term recordings of the scattering of radar or sonar beams by sharp density gradients in the high atmosphere. In the ocean, internal waves have been found by recording fluctuating currents in middepths by moored current meters, by acoustic backscatter Doppler methods, and by studies of the fluctuations of the depths of isotherms as recorded by instruments repeatedly lowered from shipboard or by autonomous instruments floating deep in the water.

Internal waves are thought to be generated in the sea by variations of the wind pressure and stress at the sea surface, by the interaction of surface waves with each other, and by the interaction of tidal motions with the rough sea floor. Their importance is that they can transmit energy and momentum throughout the ocean, not only laterally but also vertically. They can, therefore, transmit energy from the surface to all depths. In this way the otherwise sluggishly moving water at great depths can be agitated.

Charles S. Cox

Long-period waves. Long-period waves are those that exist when the period (T) is longer than one-half of a pendulum day, that is, 12 h/sin θ, where θ is the latitude of the location of interest. The main types of low-frequency ocean waves are Rossby waves, topographic Rossby waves, coastal Kelvin waves, and equatorial Kelvin waves. *See* CORIOLIS ACCELERATION.

Rossby waves. Rossby waves, named after meteorologist C. Rossby, are fundamental to the large-scale dynamics of both atmosphere and ocean. They can exist at periods from a few days to several years and help to describe, for example, seasonal and climatic fluctuations in the oceans. Since they exist at long periods, the Earth rotates several times during a wave period, and the rotation of the Earth therefore plays a central role in Rossby-wave dynamics. In order to

understand Rossby waves, it is necessary to consider rotation, angular velocity, and vorticity.

The angular velocity of the Earth is defined as a vector of magnitude Ω and direction northward along the axis of rotation. An angular velocity can be similarly defined for all solid rotating bodies: the magnitude of the angular velocity is 2π radians divided by the period of rotation, and the direction is along the axis of rotation in the direction analogous to that for the rotating Earth. The vorticity of a particle of water in solid-body rotation is defined to be twice the angular velocity. In general, the velocity shear for a water particle will not correspond to that for solid-body rotation, and the effective rotation and vorticity will consequently be modified. However, this does not affect the following explanation for the Rossby-wave mechanism.

Because the gravitational force perpendicular to the Earth's surface is so dominant, water particles tend to remain on the same horizontal level; that is, long-period ocean flows tend to be parallel to the Earth's surface. For this reason, rotational effects in the horizontal plane are of prime significance; consequently the vertical component of vorticity is of greatest importance to ocean dynamics. This vertical component of vorticity has two contributions. One, which exists even when the water is at rest relative to the Earth, is the local vertical component of the Earth's rotation at latitude θ. The remaining contribution is due to rotational effects in the horizontal currents, and its value is positive when rotation is counterclockwise as viewed locally above the Earth's surface.

Topographic Rossby waves. Another type of long-period ocean wave is the topographic Rossby wave, whose mechanism depends on variations of bottom topography and hence on the water depth. An example is the continental shelf wave in which water depth varies strongly perpendicular to the coastline across the continental shelf and slope. Continental shelf waves propagate with the coast on their right in the Northern Hemisphere and with the coast on their left in the Southern Hemisphere.

Like Rossby waves, continental-shelf waves also exist when the water is stratified. Then they are known as coastally trapped waves. These propagate along the coast in the same direction as shelf waves. Continental-shelf and coastally trapped waves play an important role in the dynamics of sea level and current fluctuations on the continental shelf. Much of the wind-forced ocean energy on the continental shelf is associated with these waves.

Coastal Kelvin waves. Coastal Kelvin waves, first discussed by Lord Kelvin, are low-frequency waves quite distinct from Rossby waves. For Kelvin waves, gravity is the essential restoring force; coastal Kelvin waves are long gravity waves trapped to a coastal wall by the Coriolis force. The amplitude of the wave decays exponentially from the coast. The water velocity perpendicular to the coast is zero everywhere. Like continental shelf waves, coastal Kelvin waves propagate with the coast on their right in the Northern Hemisphere and on their left in the Southern Hemisphere. The flow underneath a wave crest is in the direction of wave propagation, so the only way that the Coriolis force can balance the pressure-gradient force tending to flatten the sea surface is if the wave propagates with the coast on its right. Similar arguments show that propagation is with the coast on the left in the Southern Hemisphere.

Since a Kelvin wave is a long-period gravity wave, it can exist as either an external or internal gravity wave. In the external case, the wave behaves as though the water is of constant density, and the phase speed (c) is given by the expression $\sqrt{gh}$, where g is the acceleration due to gravity and h is water depth. In the internal case, the phase speed also depends on the rate of density variation with depth. In general, internal Kelvin waves travel much more slowly than surface Kelvin waves.

Winds and tidal forces effectively generate oceanic coastal Kelvin waves. Strictly speaking, coastal Kelvin waves exist only when water of constant depth is bounded by a vertical wall, and such topography in reality never occurs. However, for cases in which the decay scale is much greater than the distance across the shelf and slope to the constant-depth deep sea, the wave is dynamically coastal Kelvin. Such a condition is most often satisfied in the external Kelvin case.

Equatorial Kelvin waves. Equatorial Kelvin waves are long gravity waves with phase speed trapped to the Equator by the Coriolis force. The north-south amplitude variation is symmetric about the Equator and bell-shaped. The wave's north-south velocity is zero, and the wave must propagate eastward. Equatorial Kelvin waves exist at very low frequencies, and the internal ones appear to play a fundamental role in the dynamics of climate fluctuations like those associated with El Niño. *See* EL NIÑO; OCEAN CIRCULATION.

Allan J. Clarke

Bibliography. T. P. Barnett and K. E. Kenyon, Recent advances in the study of wind waves, *Rep. Prog. Phys.*, 38:667–729, 1975; L. J. Duckers (ed.), Wave Energy Devices, *Solar Energy Society Conference Proceedings*, C57, 1990; M. D. Earle and J. M. Bishop, *A Practical Guide to Ocean Wave Measurement and Analysis*, 1984; D. V. Evans and A. F. de O. Falcão (eds.), *Hydrodynamics of Ocean Wave-Energy Utilization*, 1985; J. B. Herbich (ed.), *Handbook of Coastal and Ocean Engineering*, vol. 1: *Wave Phenomena and Coastal Structures*, 1990; B. Kinsman, *Wind Waves*, 1965, reprint 1984; P. H. LeBlond and L. A. Mysak, *Waves in the Ocean*, 1978; A. W. Lewis, Sea-surface variations and energy: Tidal and wave power, in R. J. N. Devoy (ed.), *Sea Surface Studies: A Global View*, pp. 589–625, 1987; W. G. Van Dorn, *Oceanography and Seamanship*, 2d ed., 1993.

Oceanic islands

Islands rising from the deep sea floor. Oceanic islands range in size from mere specks of rock or sand above the reach of tides to large masses such as Iceland (39,800 mi^2 or 103,000 km^2). Excluded are islands that have continental crust, such as the Seychelles,

Norfolk, or Sardinia, even though surrounded by ocean; and with this exclusion, all oceanic islands surmount volcanic foundations. A few of these have active volcanoes, such as on Hawaii, the Galápagos islands, Iceland, and the Azores, but most islands are on extinct volcanoes. On some islands, the volcanic foundations have subsided beneath sea level, while coral reefs growing very close to sea level have kept pace with the subsidence, accumulating thicknesses of as much as 5000 ft (1500 m) of limestone deposits between the underlying volcanic rocks and the present-day coral islands. On some of these, for example Enewetak Atoll in the western Pacific, the foundations are as much as 8×10^9 years old. *See* REEF; VOLCANO.

Oceanic islands owe their existence to volcanism that began on the deep sea floor and built the volcanic edifices, flow on flow, dike on dike, up to sea level and above. The highest of the oceanic islands is Hawaii, where the peak of Mauna Kea volcano reaches 14,000 ft (4200 m). Most volcanic islands are probably built from scratch in less than 10^6 years, but minor recurrent volcanism may continue for millions of years after the main construction stage. Although the main construction of the island of Oahu was completed $2–4 \times 10^6$ years ago, a younger phase of volcanism is responsible for volcanoes such as Diamond Head formed during the past 10^6 years. *See* VOLCANOLOGY.

Geology. Oceanic volcanoes are kept in isostatic balance (floating equilibrium) during growth. The upper, more rigid part of the underlying lithosphere flexes to accommodate the load, and the lower, more ductile part flows. The result is that the edifice subsides during construction, and subsidence continues even after loading, owing to the long-term normal cooling of the supporting lithosphere. Wells drilled on Hawaii show the blocky lavas characteristic of subaerial flows, and soil zones between flows, even at depths far below sea level—supporting the occurrence of subsidence. The long-term fate of oceanic islands is thus to subside completely below sea level on time scales of millions of years.

As the edifice is built up close to sea level, the summit regions are attacked by waves, and during subaerial exposure the normal processes of erosion by waves, streams, and glaciers sculpt and reduce the landscapes, at rates dependent on the local climates, and with forms reflecting the relative erosional resistance of rock masses within the volcanic pile. Steep cliffs around the seaward edges are common; and spectacularly steep, fluted, and pinnacled relief may develop where rainfall is heavy, generally on the windward sides of islands. Erosion may at times be rapid enough to counterbalance thermal subsidence, resulting in isostatic uplift. Generally, the subaerial landscapes are reduced to a surface of low relief by the time the summit region subsides beneath the waves, and the seamount then has a nearly flat top and is termed a guyot. *See* MARINE GEOLOGY; SEAMOUNT AND GUYOT.

Associated reefs. Volcanic islands within the latitudinal zone of coral-reef growth (mainly between about 25°N and 25°S latitude) are generally fringed by reefs, which tend to protect the volcanic terranes from direct attack by ocean waves. As predicted originally by Charles Darwin, and amply confirmed by deep drilling through the reefal deposits on several oceanic islands, while the volcanic foundations of the islands subside, the corals, algae, and other shelly organisms keep piling up, layer on layer, keeping pace with subsidence. During subsidence, the reefal region maintains an area close to that enclosed by the original fringing reef, while the volcanic center diminishes in area. The final result should be a flat coral bank mainly at sea level, whereas in fact modern reef-surrounded islands commonly have a lagoon of variable width and depth between the barrier reef and the island, and many reefs have no central island at all, consisting simply of a perimeter reef encircling a deeper lagoon, a form termed an atoll. Although differential growth rates between the biotas on the perimeter reef and those in the lagoon account for some of the relief, the actual forms may result in large part from drowning of landscapes developed during the successive lowerings of global sea level to 300–600 ft (100–200 m) below its present level during the ice ages. The exposed limestone is attacked by rainwater, which dissolves it and may result in the basic rim-and-lagoon morphology by dissolution. Passes of variable width and depth, probably modified relics of stream valleys during lowered sea level, give access from the open sea to the lagoon for boats. Few barrier reef islands or atolls have passes deep enough for large ships, and those islands generally became commercial and administrative centers because of the accessibility of safe anchorage in the lagoon. *See* ATOLL; REEF.

Resources. Islands in regions of high oceanic fertility are commonly host to colonies of sea birds, and the deposits of guano have been an important source of phosphate for fertilizer. On some islands, for example, Nauru in the western equatorial Pacific, the original guano has been dissolved and phosphate minerals reprecipitated in porous host limestone rocks. The principal crop on most tropical oceanic islands is coconuts, exploited for their oil content, but some larger volcanic islands, with rich soils and abundant water supplies, are sites of plantations of sugarcane and pineapple. Atoll and barrier-reef islands have very limited water supplies, depending on small lenses of ground water, augmented by collection of rainwater. Tourism, mainly on tropical islands with balmy climates, is concentrated on high islands where water supplies are adequate. *See* ISLAND BIOGEOGRAPHY.

Edward L. Winterer

Bibliography. H. W. Menard, *Islands*, 1986; P. D. Nunn, *Oceanic Islands*, 1994.

Oceanographic vessels

Research vessels designed to collect quantitative data from the sea's surface, its depths, the sea floor, and the overlying atmosphere. The primary tool for obtaining oceanographic data has been the oceanographic vessel, which has a variety of forms and sizes.

The ships can be general purpose or special purpose, and a conventional or radical design. Their primary purpose is to carry scientists and increasingly sophisticated equipment to and from study sites on the ocean's surface, and in some cases below the surface. The ships must have the ability to lower and retrieve instruments by using winches and wires. The ship's equipment and instrumentation must determine precisely the location on the sea surface, and provide suitable communication, data gathering, archiving, and computational facilities for the scientific party.

The requirements list includes seakeeping (seakindliness, a measure of a ship's response to severe seas; and station-keeping, the ability of a ship to maintain its fixed location on the sea surface); work environment; endurance (range, days at sea); scientific complement (number of researchers accommodated); operating economy and scientific effectiveness; subdued acoustical characteristics; payload (scientific storage, weight handling); speed; and ship control. These requirements often conflict, necessitating compromise.

Types. Ships typically can be considered in three major groups based on use: general purpose (classical biological, physical, chemical, geological, and ocean engineering research, or a combination); dedicated special purpose (hydrographic survey, mapping, geophysical, or fisheries); or unique (deep-sea drilling, crewed spar buoy, or support of submersible operations). They can be used simply as delivery and support systems for exploratory devices, such as floats and bottom landers, as well as crewless remote operating vehicles (ROVs—tethered, powered, surface-controlled robots), or autonomous underwater vehicles (AUVs—freely operating robots, using computer programmed guidances). They may be of a standard classical ship configuration or an experimental design. They can be built from the keel up as research ships, or they can be adapted from available vessel hulls. Some tourist ships have been built in laboratories for research and tourist education. Even ships of opportunity (a ship not normally utilized for scientific activities) have been used to collect some forms of oceanographic data, including release of untended instrumentation, thus partially qualifying as oceanographic research vessels.

Ocean research ships are typically grouped in terms of overall length (see **table**; **Fig. 1**). In general, the larger ships are used for worldwide expeditionary work and for larger-scale, often multidisciplinary research. Some, however, are special purpose. Most make long cruises away from their home port, staying at sea between port calls from 2 to 6 weeks and operating 300 days/yr. The intermediate classes of ships tend to be general purpose, adaptable to any type of research. Many of these also work worldwide on occasion. These classes operate at sea about 220–250 days each year. The smaller ships are generally used for continental shelf, nearshore, or coastal research and tend to make short trips of 1 day to 2 weeks. Main propulsion can be single (or multiple) open or nozzle-surrounded screws, vertical axis propellers, or waterjets. Bow thrusters are generally used for station-keeping. Complexity, maintainability, efficiency, cost, and maneuverability are factors in selection. Power can be derived from direct-drive diesel engines or from electric motors driven by current supplied by individual diesel generators or a central diesel-electric power plant. *See* MARINE ENGINE.

A standard-configuration expeditionary vessel carries a scientific party of 35 and operates with a crew of 22. The key features include science laboratories, office and storage areas, a staging area, portable vans, cranes, a suite of winches and wires, a workshop, a darkroom, a dive locker, a library, a mess area, and staterooms. This vessel has the normal ship features of a main propulsion unit and bow thruster, pilot house, chart room, aft-control, and galley, as well as

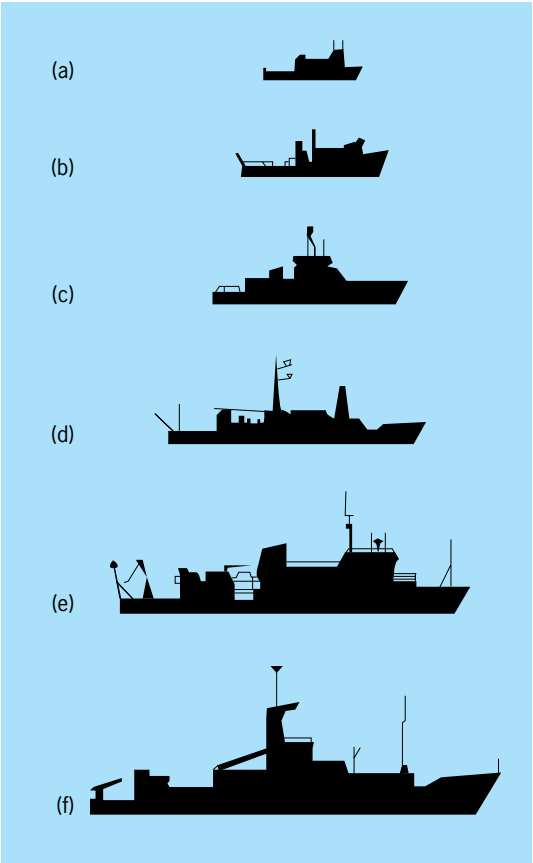

Fig. 1. Representative oceanographic ship silhouettes (profiles). (*a*) *Barnes*, 66 ft (20 m). (*b*) *Pelican*, 105 ft (32 m). (*c*) *Cape Hatteras*, 135 ft (41 m). (*d*) *Endeavor*, 185 ft (56 m). (*e*) *Ewing*, 239 ft (73 m). (*f*) *Revelle*, 274 ft (84 m).

Classification of ocean research ships

Class	Size (length overall)	Description
I	Over 250 ft (76 m)	Large, high endurance
II	200–250 ft (61–76 m)	Large, medium endurance
III	150–199 ft (45.7–61 m)	Intermediate
IV	100–149 ft (30.5–45.7 m)	Small
V	Less than 100 ft (30.5 m)	Nearshore, coastal

mast-located instruments, warning lights, electronic and satellite navigation, and data and communication antennas.

A marine geology and geophysical ship uses multichannel seismic profiling to probe the deep geological structure underlying the ocean floor. Profiling uses some method of creating a high-level noise, such as that created by a towed airgun sound source. The returning sounds from the subsea layers of rock and sediment are received by long hydrophone-streamer arrays. Properly analyzed, the information from the individual hydrophones forms a detailed picture of the arrangement and materials that underlie the sea floor. These ships are used for oil and other mineral surveys and are generally very large and expensive. They must have large reels for the towed sound arrays, extensive rigging and swing-out booms, and long electromechanical wires or fiber-optic cable deployed by appropriate winches. They constitute a very specialized configuration and are seldom used for any other purpose.

New fisheries vessels are being designed to isolate ship's noise, providing acoustic quieting. Fisheries research and stock assessment vessels look very much like standard fishing vessels, incorporating such capabilities as trawling, gill-netting, purse-seining nets, stern ramps and gallows, trawl winches, and fish-handling, processing, and preservation facilities. They also include scientific winches, laboratories, and various types of oceanographic instrumentation capabilities. *See* MARINE FISHERIES.

The design concept of the small-waterplane-area twin hull (SWATH) vessel (**Fig. 2**) is based on having all working spaces elevated above the sea's surface, supported by two in-line submerged hulls. The thin struts connecting the two experience only very small displacement forces with the passage of waves, thus effectively decoupling the ship from surface conditions while at rest and under way and providing a very stable platform from which to work. SWATH designs are becoming more accepted in the oceanographic research community as more are built and placed in service.

The floating instrument platform (FLIP) [**Fig. 3**] is essentially a 355-ft (108-m) spar buoy with a crew; it is towed in a horizontal mode to a work site and tipped into its vertical attitude, where it is used to make subsurface measurements. After flooding, its ballast tanks place it in vertical position; it extends 300 ft (91 m) underwater, where because of its design stability it is almost unaffected by wave action, moving only about 3 ft (0.9 m) in a 30-ft (9-m) wave. It houses a crew of 5 and a science party of 11, two internal electronic laboratories, and an outside laboratory.

A new Arctic research vessel, USCGC *Healy*, recently entered into service (**Fig. 4**). This 420-ft (129-m) 17,000-ton icebreaker is outfitted for Arctic research. It is capable of conducting a wide range of research activities, and providing more than 4200 ft^2 of laboratory space. It also has numerous electronic sensor systems, oceanographic winches, and accommodations for up to 50 scientists. *Healy* is designed to break $4^1/_2$ ft (1.4 m) of ice continuously at 3 knots and can operate in temperatures as low as −50°F.

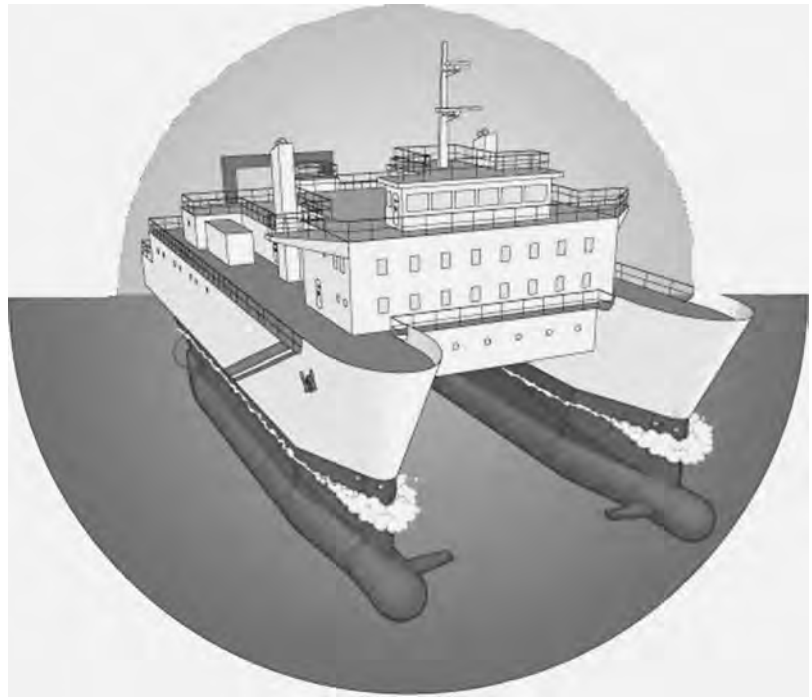

Fig. 2. Small-waterplane-area twin hull (SWATH) vessel, the 180-ft (55-m) *AGOR 26*, University of Hawaii.

Research requirements. Two approaches to ocean research are typically used: direct collection of

Fig. 3. Deployed FLIP—floating instrument platform. (*Scripps Institution of Oceanography*)

Fig. 4. Arctic research vessel *USCGC Healy*, commissioned November 10, 1999. (*U.S. Coast Guard*)

samples (water, sea-floor rocks and sediment, plants and animals) for topside evaluation and study, and in-place measurement of parameters of interest (directly or from analysis by remote or autonomous instrumentation).

To deploy and retrieve research tools requires a suite of powerful, sensitive winches and strong wires (some with electrical conducting elements). The tools include nets (some simple, others incorporating environmental sensors and cameras), trawls and dredges, coring devices, remote bottom landers, sediment traps, continuously recording conductivity/temperature/depth sensors in rosettes with water collection bottles, and water and organism pumps. Packages such as large floats, drifting buoys, and moored arrays with current meter strings require this capability as well, as does the launching and recovery of small craft, submersibles, ROVs, and AUVs. Large, open working deck areas (with convenient control areas and clear observation vistas) as well as highly trained crews are required.

Winch types in common use include trawl winches (driven by a powered drum), traction winches (driven by a powered traction-head system), hydrographic winches (most containing one or more conducting cable elements), and smart winches (that compensate for ship motion, or even allow instrument packages to change depth to follow ocean parameters of interest). Some carry up to full-ocean-depth cable (33,000 ft or 10,000 m). Most cable is made from twisted wire bundles torque-balanced to reduce kinking when deployed. Fiberglass optical cable (for high rates of transmission of data and images) and synthetic cable made from noncontaminating materials (for sophisticated chemical studies) are increasingly being used. A-frames and J-frames along with specially adapted cranes (such as the knuckleboom crane) guide the wire rope used to handle instrument and other packages over the side or stern of the ship into and out of the sea. Cranes must also be capable of loading and transferring material in port and at sea. *See* HOISTING MACHINES.

Types of laboratories include the general-purpose main laboratory, the wet laboratory (for water and net specimen handling and preservation), the clean Special-Purpose laboratory (for high-quality chemical or radioactive work), and, on large ships, a separate computer laboratory (for main frames and support of unique instrumentation). Sample storage (in some cases cold or refrigerated) is provided. Vans, most of standard configuration, are frequently used. They function as additional work space (special chemistry or dive support, for instance), storage space, and on rare occasions living space.

Hull-mounted multibeam sonar bathymetric systems provide sea-floor surface contour images. The systems installed on most modern large oceanographic vessels have two arrays, one to send sonic energy and the other to receive the reflected returns. Advanced on-board systems can combine data from more than one type of sonar system (multibeam, side-scan, chirp) into three-dimensional imagery. *See* SONAR.

Data from the multitude of scientific devices on board is recorded, archived, and merged with ship parameters to provide a continuous cruise record. Computers in scientific spaces are linked with each other and generally have access to shore-based large computers through satellite transmission. Sophisticated computers on the ship capable of image manipulation are available on board for scientists. Satellite communication is used for communication between and among ships and their home ports. Ships also use satellite reception and linkages for long-range communication, precise navigation (Global Positioning System; GPS), and for transmittal of signals and data streams from buoys and moorings and between vessels. Some data sensed from satellites are used in real time by the ship to locate centers of biological activity, surface currents and eddies, sea surface slopes, storms, and ice floes. Cruise plans can be adapted to changing conditions based on the information. *See* INSTRUMENTED BUOYS; SATELLITE NAVIGATION SYSTEMS.

Altogether there are close to one thousand oceanographic vessels worldwide. The number varies with the classification technique used and the state of activity of the vessels. Russia, the United States, the United Kingdom, Sweden, Japan, France, and Canada operate the major fleets. Increasingly, large-scale ocean research is both international and multidisciplinary. Much more sophisticated tools, especially satellite and computer related, as well as autonomous submersibles, bottom landers, drifters, and moorings, will continue to evolve and be deployed, thereby changing, replacing, or enhancing traditional ship functions. *See* MARINE ENGINEERING; NAVAL ARCHITECTURE. J. J. Griffin/J. F. Bash

Bottom samplers. Historically, the tools used by ocean scientists to collect oceanic rock and sediment samples from oceanographic vessels have been confined to three basic family groups: grabs, dredges, and coring devices. From these three groupings a wide variety of unique tools have been developed for use by the scientific community.

Grab samplers. Traditionally, grabs have been reserved for use in shallow water depths where large numbers of individual samples from the upper few centimeters of the bottom sediments are required.

Within this group of samplers, the Van Veen grab is perhaps the most commonly used device. It is produced in a variety of sampling sizes ranging from 0.82-ft (0.25-m) to over 3.3-ft (1-m) surface area sampling capability, making it an ideal tool for the estimation of benthic populations as well as the determination of sediment distribution patterns over wide areas. The size of a Van Veen is determined by the actual square area that the open grab will cover when it first contacts the bottom. One point to remember is that the recovered Van Veen sample will have experienced some slight distortion due to the closure of the grab jaws.

A second, and unique, grab device is the underway sampler, which is deployed on a wire from a moving vessel. The sampler is lowered to the bottom until it contacts the sediment; at this time a pin or light line is sheared, a small volume of sediment recovered, and the sampler returned to the vessel. The design of the device is such that the pin or line will shear before the sampler has an opportunity to dig deeply into the bottom and part the ship's wire.

Dredges. The standard chain bag dredge is constructed in a variety of sizes and weights up to a heavy-duty model with a 3.9 ft by 23 in. (1.2 m by 60 cm) mouth opening. Attached to the dredge mouth is either a chain or wire mesh bag used to retain the sample. Occasionally the dredge bag is lined with a canvas or burlap sack in order to retain small fragments of rock, volcanic glass, or sediment. The mouth of the dredge is of a welded steel construction with an articulated bail that is attached to the leading edge. When the dredge is rigged, it is customary to employ both a weak link and a swivel above the dredge in order to protect the ship's cable.

During actual dredging operations the water depth, dredge weight, and wire strength must be carefully considered and protective measures taken, through the use of a series of weak links, to ensure the integrity of the ship's cable. Two weak links, capable of varying shear loads, are used on each dredge: the first located at the dredge bail and a second at the rear of the dredge mouth. Under this scheme, should the dredge become fouled, the first link would part, transferring the load to the second link, pulling the dredge free, and preserving the sample and instrument. In the event that the dredge cannot be pulled free, the second link will part before damage can be done to the ship's cable. The use of dredges to collect oceanic rocks, manganese nodules, and other sea-floor deposits is virtually non-depth-limited; however, this technique is restricted to recovering only those materials lying on or near the sediment water interface.

Coring devices. These tools for studying the structure of oceanic sediments have developed into highly specialized instruments since they were first conceived in the early 1940s. Although individual coring devices may differ as to complexity, weight, sample diameter, and length of sample recovered, their main purpose still remains the collection of an undisturbed section of the sea-floor sediments.

Gravity cores. Gravity cores, of lengths up to 11 ft (3.5 m) and weights of 990 lb (450 kg), are frequently used as survey tools in the deep sea. The barrels of these cores may belined with a plastic tube that is

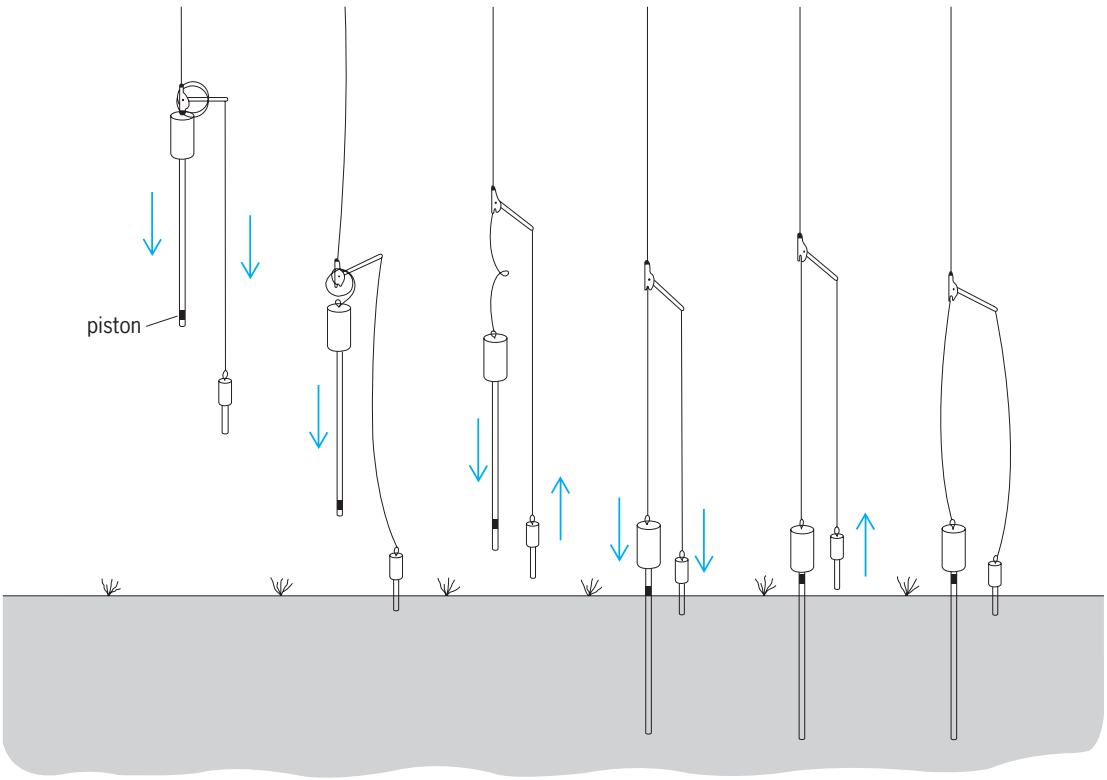

Fig. 5. Piston coring sequence (left to right) showing the effects of cable rebound on the trigger core.

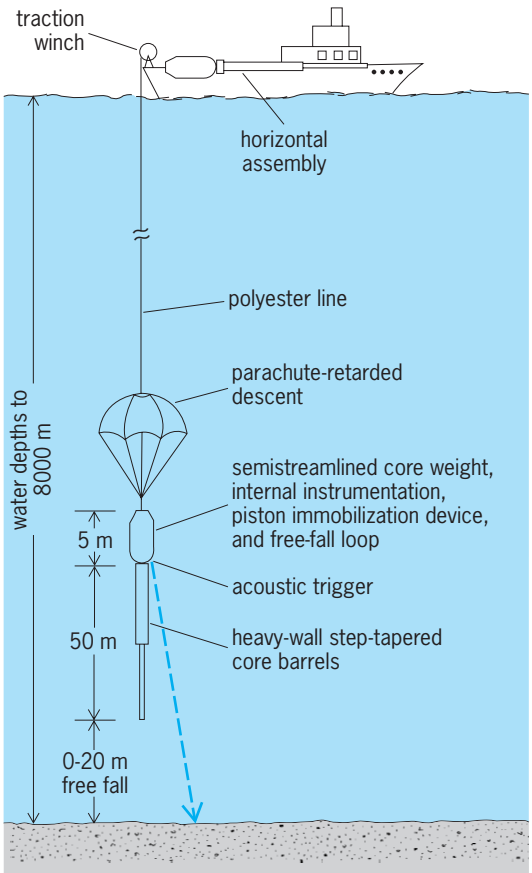

Fig. 6. General configuration of long coring system. (*Drawing by A. H. Driscoll*)

used as a sample container for the recovered sediment. In theory, the gravity core is dependent upon the speed of the ship's winch and its own weight to achieve penetration and, as a result of this, is limited to shorter sample recovery lengths than are possible with a piston core.

One exception to the gravity core operational theory is the free-fall core. This device is dropped directly from a moving vessel and allowed to free-fall through the water column to the sea floor. Upon contact with the bottom and recovery of a 3.3-ft (1-m) sample, the sample is pulled free of the ballast portion of the core by two glass flotation spheres. The positive buoyancy of the flotation is sufficient to return the sample to the surface, where it can be retrieved by the surface vessel. By using this type of device it is possible to survey a relatively large area of the sea floor in a minimum of time.

Box cores. Another type of coring device capable of recovering undisturbed sections of the sea floor is the box core. This device, although bulkier than the simple gravity core, has been specially designed to recover a square section of the upper meter of sediment at the interface. Penetration of the coring box is achieved via a heavy driving weight mounted directly above the box. As the box core framework settles to the sea floor, slack in the ship's wire releases the weighted box and allows it to sink gently into the sediment, taking the core. Upon completion of the coring sequence, the ship's cable is slowly retrieved, causing a spade, attached to the framework, to cut into the sediment, stopping directly below the implanted box. As the cable becomes taut, this spade rises and seals the bottom of the box, preserving the sample as it is returned to the surface vessel. The general quality of samples taken with this device is usually very high, and it is even possible to see worm tubes and small brittle stars preserved on the top of the sample when the box is opened.

Piston core. The piston core has been designed to recover sediment sequences longer than would be possible by using a gravity coring technique. Essentially, the sequence of events surrounding the operation of the piston core is as follows (**Fig. 5**). The piston core and its accompanying trigger core are lowered to the sea floor by the ship's cable until the trigger core contacts the sediment. At this point the tension on the tripping arm is relaxed and the piston core allowed to fall free. As the piston core free-falls toward the bottom, it pays out a coil of cable behind it; when the cutter contacts the sediment, the cable becomes taut, immobilizing the piston and allowing the core barrels to slide past the piston and obtain the sample. It is this force generated by the action of the barrels bypassing the piston that overcomes the internal wall friction of the sediment and allows a long sample to be taken. Even though the process appears complicated, the entire sequence occurs within 3 s of the trigger core's first striking the bottom.

In general, the piston cores range in weights and sizes from a 3100-lb (1400-kg), 59-ft (18-m), 2.5-in.-diameter (6.5-cm) standard device to a 31,000-lb (14,000-kg), 160-ft (50-m), 5.0-in.-diameter (12.7-cm) long coring system (**Fig. 6**). The addition of instrumentation capabilities in the core weight has made the piston core more of a deep-ocean experiment platform than the simple collecting device originally developed in the late 1940s.

Ranging. Relative to the effective operation of all bottom sampling equipment is a requirement that the user know where and how far away the device

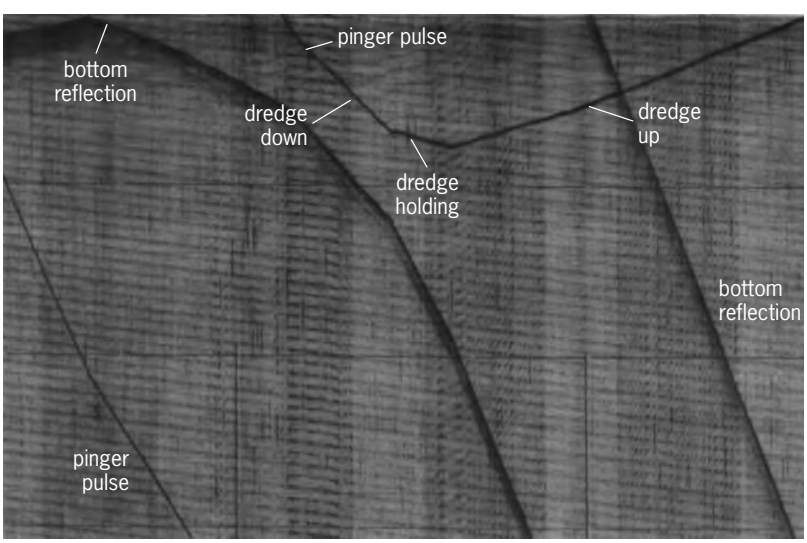

Fig. 7. Typical echo-sounding record of dredge lowering showing descent, bottom contact, and recovery. (*Woods Hole Oceanographic Institution*)

is from the sea floor. This is usually accomplished by attaching a telemetering pinger to the ship's cable directly above the sampling device. The pinger emits a 12-kHz pulse once per second that, radiated downward, bounces off the sea floor and returns to the surface, where it is received aboard the surface vessel. By measuring the distance between the outgoing pulse at the pinger and the bottom reflection, the observer can accurately calculate the distance between the instrument and the sea floor (**Fig. 7**). *See* ECHO SOUNDER; OCEANOGRAPHY. Alan H. Driscoll

Bibliography. J. F. Bash, UNOLS Research Fleet, *Sea Technology*, June 1998; C. A. Bookman, *Trends and Opportunities in Ocean Technology 1992–1997*, Marine Board, National Academy of Sciences, 1992; J. J. Griffin, *Science at Sea*, 1996; *Marine Technology Society Journal*, quarterly; R. Schmitt et al., *The Academic Research Fleet*, May 1999; UNOLS, *Scientific Mission Requirements for Oceanographic Research Vessels*, 1988; UNOLS and V. Cullen (eds.), *The Research Fleet*, 2000.

Oceanography

The science of the sea; including physical oceanography (the study of the physical properties of sea water and its motion in waves, tides, and currents), marine chemistry, marine geology, and marine biology. The need to know more about the impact of marine pollution and possible effects of the exploitation of marine resources, together with the role of the ocean in possible global warming and climate change, means that oceanography is an important scientific discipline. Improved understanding of the sea has been essential in such diverse fields as fisheries conservation, the exploitation of underwater oil and gas reserves, and coastal protection policy, as well as in national defense strategies. The scientific benefits include not only improved understanding of the oceans and their inhabitants, but important information about the evolution of the Earth and its tectonic processes, and about the global environment and climate, past and present, as well as possible future changes. *See* CLIMATE HISTORY; COASTAL ENGINEERING; MARINE MINING; MARINE SEDIMENTS; MARITIME METEOROLOGY; OIL AND GAS, OFFSHORE.

The modern discipline. The traditional basis of modern oceanography is the hydrographic station. Hydrographic studies are still carried out at regular intervals, with the research vessel in a specific position. Seawater temperature, depth, and salinity can be measured continuously by a probe towed behind the ship. The revolution in electronics has provided not only a new generation of instruments for studying the sea but also new ways of collecting and analyzing the data they produce. Computers are employed in gathering and processing data in all fields, and are also used in the creation of mathematical models to aid in understanding. Much information can also be gained by remote sensing using satellites, which are also a valuable navigational aid. These provide data on sea surface temperature and currents, and on marine productivity. Satellite altimetry gives information on wave height and winds and even bottom topography (because this affects sea level). Scientists look forward to a day when observations can be made in the deep sea by autonomous vehicles. However, the ship remains a fundamental tool for many observations. *See* ALTIMETER; COMPUTER; DIGITAL COMPUTER; HYDROGRAPHY; MODEL THEORY; OCEANOGRAPHIC VESSELS; REMOTE SENSING; SATELLITE NAVIGATION SYSTEMS; SEAWATER.

Physical oceanography. Physical oceanography, and in particular ocean circulation studies, forms the core of oceanographic research. The movements of seawater—ocean currents—are powerful agents in the distribution of heat throughout the world, influencing both weather and climate. The continual renewal of water bearing dissolved nutrients is essential to most marine organisms, which will be abundant only where such supplies are available.

Twentieth-century dynamical oceanographers have shown how the deflecting effect of the Earth's rotation influences water movements. Geostrophic forces are responsible for the intensification of surface currents, such as the Gulf Stream, on the western sides of oceans (western boundary currents). In the 1950s the existence of a southward-flowing countercurrent under the Gulf Stream was predicted. Neutrally buoyant floats, to be tracked by radio signals picked up by hydrophones on board ship, were deployed to confirm this prediction. However, further out in the Atlantic the floats moved unexpectedly fast, with frequent changes in direction. This was the first indication of vigorous eddies in the ocean that have since been shown to be comparable to atmospheric weather systems. Further investigations of these phenomena, studies of equatorial currents and undercurrents, and transport between oceans are the principal topics occupying physical oceanographers in the latter part of the twentieth century. The World Ocean Circulation Experiment, a large-scale international program of cooperation on research projects and data-collecting expeditions, was designed to throw further light on the ocean's influence on world climate. Among the techniques being employed are the use of Swallow floats, arrays of current meters, that can be moored to the sea bed for a period of time to measure deep-water movements and then released for retrieval by acoustic signals, and the use of geochemical tracers, including chlorofluorocarbons, to obtain data on the distribution and age of water masses. *See* CORIOLIS ACCELERATION.

Marine biology. Biologists seek to classify the great diversity of life, from microscopic bacteria and phytoplankton to the great whales. To learn how the food web operates, they must examine the constraints on marine productivity and the distribution of species in the surface and midwater layers, and the vertical migrations between them, and in the bottom-living (benthic) fauna. Deep-sea cameras and submersibles now permit visual evidence of creatures in these remote depths to be obtained. *See* DEEP-SEA FAUNA; FISHERIES ECOLOGY; MARINE BIOLOGICAL

SAMPLING; MARINE ECOLOGY; MARINE FISHERIES; MARINE MICROBIOLOGY; SEAWATER FERTILITY.

Marine geology. Since the early 1900s, all recorded ocean depths have been incorporated in the General Bathymetric Chart of the Ocean. The amount of data available increased greatly with the introduction of continuous echo sounders; subsequently, side-scan sonar permitted very detailed topographical surveys to be made of the ocean floor. The features thus revealed, in particular the midocean ridges (spreading centers) and deep trenches (subduction zones), are integral to the theory of plate tectonics. An important discovery made toward the end of the twentieth century was the existence of hydrothermal vents, where hot mineral-rich water gushes from the Earth's interior. The deposition of minerals at these sites and the discovery of associated ecosystems make them of potential economic as well as great scientific interest. Possibly life on Earth began in similar situations in the remote past. Investigation of such areas can be made directly by scientists using submersibles and by underwater cameras, as well as by instrumentation. Even the sediments and other rocks of the sea floor are being sampled by the international deep-sea drilling program to provide information on how the present oceans evolved and on past climate change. *See* ECHO SOUNDER; HYDROTHERMAL VENT; MARINE GEOLOGY; MID-OCEANIC RIDGE; PLATE TECTONICS; SONAR; SUBDUCTION ZONES; UNDERWATER VEHICLE; UNDERWATER PHOTOGRAPHY; UNDERWATER TELEVISION. Margaret Deacon

Bibliography. A. Colling, *Ocean Circulation*, 2d ed., 2001; T. S. Garrison, *Oceanography: An Invitation to Marine Science*, 5th ed., 2004; J. A. Knauss, *Introduction to Physical Oceanography*, 2d ed., 1996; C. M. Lalli, *Biological Oceanography*, 2d ed., 1997; M. E. Q. Pilson, *An Introduction to the Chemistry of the Sea*, 1998; H. V. Thurman and A. P. Trujillo, *Essentials of Oceanography*, 7th ed., 2004.

Octane number

A standard laboratory measure of a fuel's ability to resist knock during combustion in a spark-ignition engine. A single-cylinder four-stroke engine of standardized design is used to determine the knock resistance of a given fuel by comparing it with that of primary reference fuels composed of varying proportions of two pure hydrocarbons, one very high in knock resistance and the other very low. A highly knock-resistant isooctane (2,2,4-trimethylpentane, C_8H_{18}) is assigned a rating of 100 on the octane scale, and normal heptane (C_7H_{16}), with very poor knock resistance, represents zero on the scale. Octane number is defined as the percentage of isooctane required in a blend with normal heptane to match the knocking behavior of the gasoline being tested. *See* SPARK KNOCK.

The CFR (Cooperative Fuel Research) knock-test engine used to determine octane number has a compression ratio that can be varied at will and a knockmeter to register knock intensity. In the classic method of knock rating, the engine is run on the fuel to be tested, and its compression ratio is adjusted to give a standard level of knock intensity. Without changing the compression ratio, this knock level is then bracketed by running the engine on the primary reference fuel blends, one of which knocks a little more than the test fuel, and the other a little less. The octane number of the fuel being rated is then determined by interpolation from the knockmeter readings of the bracketing reference fuels.

Alternatively, the engine's compression ratio can be adjusted to close to the limit for the fuel being tested, and then, while the engine is run on each of two closely bracketing test fuels, the ratio can be readjusted to a standard intensity reading on the knockmeter. Finally, the engine is again run on the test fuel, and its compression ratio is adjusted to give the same knockmeter reading. The octane number of the test fuel is then interpolated from the compression ratio settings.

Knock tests are performed under one of two sets of engine operating conditions. Results of tests using the so-called Motor (M) method correlated well with the fuels and automobile engines of the 1930s, when the method was developed. The Research (R) method was developed later when improved refining processes and engines gave gasolines better road performance than their M ratings indicated. Today, the arithmetic average of a gasoline's R and M ratings usually is a good indicator of its performance in a typical car on the road. Thus, that average, $(R + M)/2$, is posted at service stations to show the antiknock quality of a fuel.

For fuels with a rating higher than 100 octane, the rating is usually obtained by determining the amount of tetraethyllead compound that needs to be added to pure isooctane to match the knock resistance of the test fuel. For example, if the amount is 1.3 ml, the fuel's rating is expressed as $100 + 1.3$ or extrapolated above 100 (in this case, about 110 octane) by means of a correlation curve. John C. Lane

Bibliography. American Society for Testing and Materials, *Annual Book of ASTM Standards*, pt. 47, annually.

Octocorallia (Alcyonaria)

A subclass of Anthozoa. These cnidarians are benthic as adults, living attached to firm substrata or burrowed into soft sediments, from the intertidal zone to great depths. As in all anthozoans, the adult form is a polyp—a cylindrical organism that has, at its free end, the single body opening, the mouth, which is surrounded by eight tentacles. They are colonial (**Fig. 1**), with a few possible exceptions.

Morphology. Each of the eight tentacles has many small or elongate side branches (pinnules) oriented perpendicular to the axis of the hollow, flexible tentacle. The opposite end of the polyp is attached to the substratum or, more commonly in alcyonarians, emerges from the tissue that unites the members of

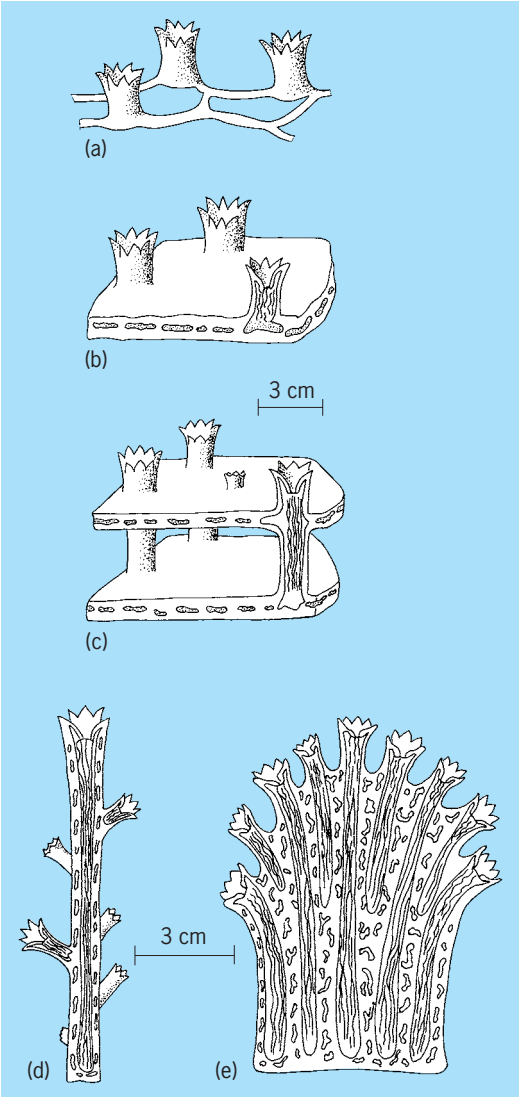

Fig. 1. Diagrammatic figures of the colonies of Alcyonaria. (*a*) *Cornularia*. (*b*) *Clavularia*. (*c*) *Tubipora*. (*d*) *Telesto*. (*e*) *Alcyonium*.

the colony. Sexual reproduction typically results in a planula larva, which eventually metamorphoses into a polyp. Asexual (vegetative) reproduction typically involves budding a new polyp from an existing one or from the tissue connecting members of the colony (the coenenchyme); in a colony, the resulting polyps remain attached to one another.

Autozooid. All alcyonarians have a type of polyp known as the autozooid (**Figs. 2** and **3**). The internal body space (the gastrovascular cavity) is divided by eight longitudinally oriented sheets of tissue known as mesenteries (Fig. 3). The best-developed pair of mesenteries (the asulcal pair) bears large, heavily ciliated filaments. The other six mesenteries bear filaments having many gland cells. Gametes form and mature in the mesenteries, which rupture to release them when they are ripe; some kinds of alcyonarians are hermaphroditic and some are gonochoric (with separate sexes).

Typically gametes are spawned freely into the sea, where egg and sperm meet and the planula develops; in some alcyonarians, however, embryos are brooded internally or on the surface of the colony, so a free-swimming stage of the life cycle is lacking. Longitudinal retractor muscles are strongly developed on the sulcal surface of each mesentery; thus, the retractor muscles of the sulcal pair face each other (Fig. 3).

The actinopharynx (stomodeum) has a single, ciliated siphonoglyph (longitudinal groove) located on the sulcal side (Figs. 2 and 3). Thus, although the polyp exhibits octamerous radial symmetry, it has elements of bilaterality; such biradial symmetry is typical of anthozoans. The flexible oral end of the polyp is termed the anthocodium. The basal portion, or anthostele, which is typically more rigid, is embedded in the coenenchyme, which is permeated by canals (the smaller of which are called solenia) that interconnect the gastrovascular cavities of polyps.

Siphonozooid and mesozooid. Some types of alcyonarians are dimorphic, having a second type of polyp

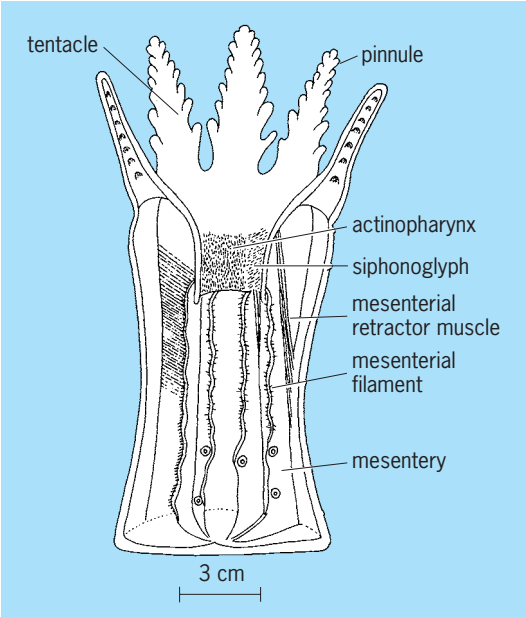

Fig. 2. Autozooid of Alcyonaria.

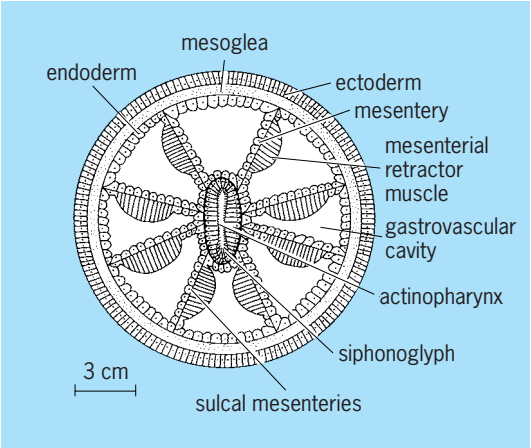

Fig. 3. Mesenterial arrangement of Alcyonaria.

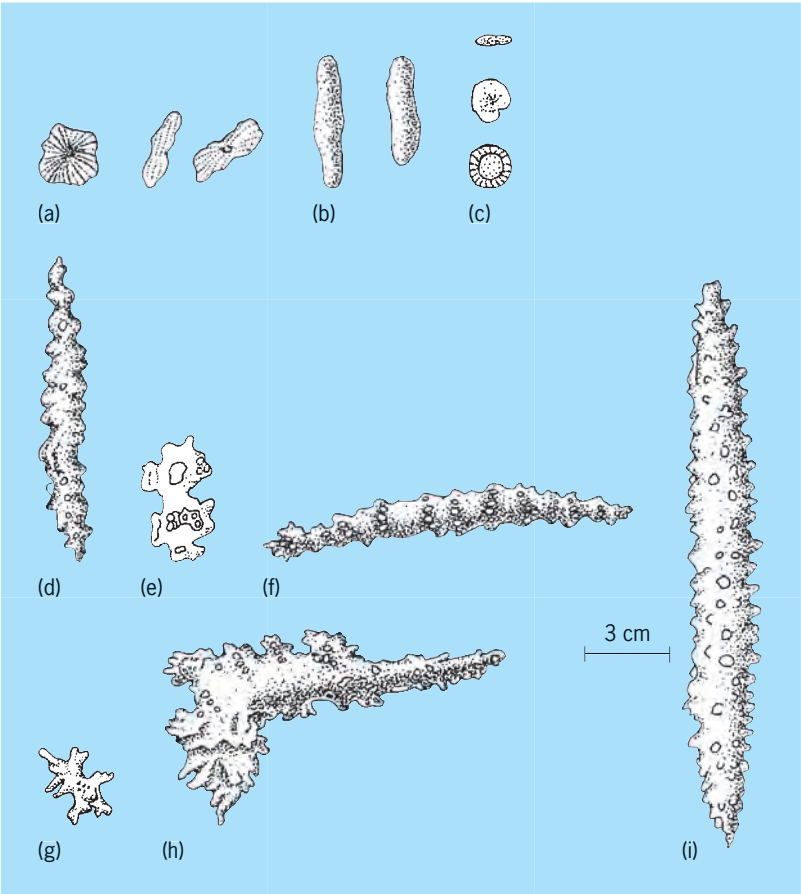

Fig. 4. Sclerites of Alcyonaria. (*a*) *Xenia uniserta*. (*b*) *Anthelia fulginosa*. (*c*) *Sympodium coeruleum*. (*d*) *Clavularia chunni*. (*e*) *Sarcophyton crassocaule*. (*f*) *Sarcophyton acutangulum*. (*g*) *Anthomastus antarcticus*. (*h*) *Capnella rugosa*. (*i*) *Nephthya pacifica*.

known as a siphonozooid. Typically smaller than an autozooid, it functions to pump water through the canal system. It is by this means that a sea pen can extend from the sand in which the siphonozooid anchors it. Fewer alcyonarians still have a third type of polyp, the mesozooid, which is intermediate in morphology between the other two. Siphonozooids and mesozooids typically do not form eggs or sperm, and have reduced or no tentacles and mesenterial filaments.

Skeleton. The skeleton that functions to protect and support the octocoral is difficult or impossible to see while the animal is alive.

In two types of octocorals, the skeleton is solid and composed of the mineral calcium carbonate. One of these is the organ-pipe coral, *Tubipora* (Fig. 1*c*), whose red skeleton is in the form of elongate pipes (in which the living animal resides) connected intermittently by transverse platforms. The other is blue coral (or Helioporacea), whose skeleton, which resembles that of scleractinian corals (members of order Scleractinia, in the other anthozoan subclass), is blue (the skeleton of scleractinians is white).

In the rest of octocorals, the hard parts are sclerites made of calcium carbonate that range from microscopic to a few millimeters in length, and may look like lumps, discs, or rods (straight or curved) that are thick in the middle, pointed at the ends, and decorated with knobs (**Fig. 4**). Sclerites may be clustered or fused together to form a cup around each polyp, scattered in the coenenchyme and polyp body wall and tentacles, or embedded in an organic matrix.

Animals with a large amount of coenenchyme typically are fleshy and flexible; they are commonly known as "soft corals." In sea fans and sea whips (formerly constituting order Gorgonacea), sclerites are embedded in the thin living tissue that coats a skeleton in life and that quickly rots away after death, leaving the skeleton, which consists of an organic matrix that may be mineralized to some extent, and that is typically dendritic—in the form of a tree—or whiplike.

Classification. Previously, alcyonarians were divided into six or more orders, but now only three are recognized. The two that have long been recognized are Helioporacea, the blue corals and Pennatulacea, the sea pens and sea pansies. Pennatulaceans are unique among alcyonarians in having a colony with a definite symmetry: the primary (axial) polyp, which anchors the colony in the soft substratum, forms the main axis from which autozooids and siphonozooids branch bilaterally. Since intermediates have been discovered between what had been considered members of separate orders (including Stolonifera and Telestacea), all other alcyonarians are now considered to belong to order Alcyonacea. *See* ANTHOZOA; CNIDARIA.

Daphne G. Fautin

Bibliography. H. Erhardt and D. Knop, *Corals: Indo-Pacific Field Guide*, Ikan, Frankfurt, 2005; K. Fabricius and P. Alderslade, *Soft Corals and Sea Fans: A Comprehensive Guide to the Tropical Shallow-water Genera of the Central-West Pacific, the Indian Ocean and the Red Sea*, Australian Institute of Marine Science, Townsville, 2001; M. Grasshoff and G. Bargibant, *Coral Reef Gorgonians of New Caledonia*, IRD, Paris, 2001; Several well-illustrated articles on octocorals, *Koralle* [a German-language magazine for aquarium hobbyists], issue 12 (December 2001/January 2002).

Octopoda

An order of the class Cephalopoda (subclass Coleoidea), characterized by eight appendages that encircle the mouth, a saclike body, and an internal shell that is much modified or reduced from that of its ancestors. One or two rows of suckers without chitinous rings occur along the eight highly flexible contractile arms. The approximately 200 species of octopods include shallow-water forms like the common octopus, *Octopus vulgaris*; open-ocean species like the paper argonaut, *Argonauta argo* (*see* **illus.**); and deep-sea forms with fins like the flapjack devilfish, *Opisthoteuthis californica*. Two suborders divide the Octopoda into those with paddle-shaped fins on the body and tendrillike cirri on the arms (Cirrata) and those without fins or cirri (Incirrata). *See* OCTOPUS.

The cirrate or finned octopods generally are deep-sea species that live at depths from about 320 ft

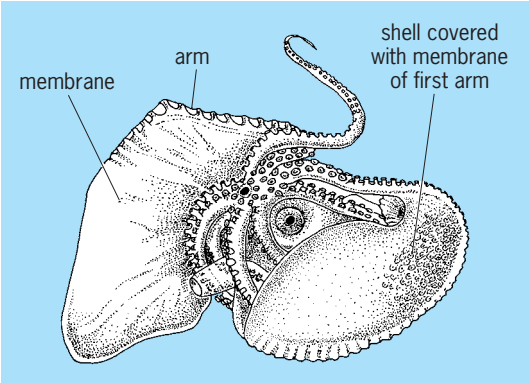

Lateral view of female (*Argonauta*) Argonautidae.

(100 m) to over 16,000 ft (5000 m) either on (benthic) or near (epibenthic) the ocean bottom, such as *Opisthoteuthis* and *Cirroteuthis*, respectively. They swim with their fins and the gentle jetting of water taken into their mantle and ejected through the funnel. The arms are connected with a very deep web. Most finned octopods are somewhat gelatinous, and the cirri along the arms are thought to be sensory organs that enable them to detect the presence of prey in the nearly total darkness of their habitat. The deep-living *Cirrothauma murrayi* has minute eyes that detect light but do not form images. All other octopods have extremely well-developed eyes, and the incirrate forms like the common *Octopus vulgaris* have acuity of vision that rivals or exceeds that of fishes.

The incirrates are shallow-living species, mostly benthic, but some groups are epipelagic (near-surface, open ocean) or mesopelagic (middepths, open ocean). Such octopuses have highly developed and enlarged brains and a complex nervous system to support their superior visual and tactile senses. Experiments have shown that octopuses learn complicated tasks and have long-term memories. Behavioral, neurological, and anatomical studies of the octopus brain have contributed immeasurably to understanding the human brain.

Benthic incirrate octopuses live worldwide in holes and crevices in rocks and coral reefs, or in burrows on muddy or sandy bottoms. They prey primarily upon crabs, bivalves (clams, mussels), and snails, which they subdue by injecting a poisonous saliva from their jaws, which are like parrot beaks. Species of the Indo-West Pacific genus *Hapalochlaena* have toxic saliva that has caused death in humans. The sexes are separate, and the young hatch out of strands of eggs attached in caves, in empty bivalve shells, or, in *Argonauta*, in the thin-shelled egg case carried by the female. Most species are small, from a few inches (1 in. = 2.5 cm) to 2–3 ft (0.6–1 m) across the outstretched arms, but *Octopus dofleini* from the North Pacific (from northern California to Alaska to Japan) grows to a 32-ft (10-m) arm span. Octopuses are captured and eaten by humans in many parts of the world. *See* CEPHALOPODA; COLEOIDEA.

Clyde F. E. Roper

Bibliography. P. R. Boyle (ed.), *Cephalopod Life Cycles*, vol. 1, 1983, vol. 2, 1987; F. W. Lane, *Kingdom of the Octopus*, 1960; K. N. Nesis, *Cephalopods of the World*, transla. from Russian, 1987; C. F. E. Roper, M. J. Sweeney, and C. E. Nauen, *Cephalopods of the World*, FAO Fisheries Synopsis 125, vol. 3, 1984; M. J. Wells, *Octopus: Physiology and Behavior of an Advanced Invertebrate*, 1978; K. M. Wilbur et al. (eds.), *The Mollusca*, vol. 11: *Form and Function*, 1988; K. M. Wilbur and M. R. Clarke (eds.), *The Mollusca*, vol. 12: *Paleontology and Neontology of Cephalopods*, 1988.

Octopus

Any species in the subclass Coleoidea, order Polypoidea, of the cephalopod mollusks or, more particularly, any of the approximately 120 species of the genus *Octopus* (see **illus.**). Widely distributed in coastal areas of all oceans, such octopods have a saclike body without any skeletal support and eight arms of equal length which bear suckers. They are predacious carnivores, crawling on the sea bottom and lurking in submarine caves and crevices. They can swim by jet propulsion but rarely do so, except as an escape mechanism. The body ranges in size from 1.2 to 14 in. (3 to 35 cm) across, and there are no "giant" octopuses.

As in all cephalopods, the mantle cavity in *Octopus* is modified from the usual molluscan pattern by a reversal of the water circulation through the pair of ctenidia (gills). The importance of cilia is greatly reduced, with those of the ctenidia and pallial wall being functionally replaced by powerful muscles which pump water continuously for respiration and intermittently for jet propulsion. The efferent circulation from the gills leads to the two auricles of the systematic heart, but there are also two accessory pumps (sometimes called branchial hearts) receiving

The octopus, a species of the subclass Coleoidea, order Polypoidea, of the cephalopod mollusks.

blood from the closed venous system and driving it into the afferent branchial vessels. The gut is still based on the molluscan pattern, with ciliated leaflets in the posterior part of the stomach sorting indigestible material, but several glands secrete a wide range of powerful enzymes. Prey, such as living crabs, are captured by the arms with their suckers and pressed in toward the mouth, which is armed with both a chitinous parrotlike beak and a powerful radular apparatus. The saliva of *Octopus* contains a neurotoxin.

Rapid and patterned color changes are possible in octopuses, cuttlefish, and their allies because their chromatophores (unlike those of other invertebrates) are miniature bags of pigment expanded by the pull of muscles under nervous control. As well as being used in concealment and for terrorizing displays, rapid color changes along with tactile caresses form part of the courtship behavior of such cephalopods. The male conveys the sperm in a spermatophore packet by a specialized arm (the hectocotylus).

The sense organs of octopuses are complex and efficient in discrimination; they include eyes, olfactory pits, chemotactile and mechanotactile sensillae on the arms, and statocysts capable of detecting direction and angle of acceleration as well as static posture. The eyes have a cornea and iris diaphragm, a movable lens, and a retina, capable of light and dark adaptation, which is not inverted. They represent an outstanding example of convergent evolution when compared in detail to the eyes of higher vertebrates such as mammals, and experiments have shown octopus eyes to be capable of considerable image formation and shape discrimination.

The discrete neural ganglia found in other mollusks are fused into a massive brain in modern cephalopods. The brain and associated neural lobes in an adult *Octopus* may contain over 300 million cells, and the brain-body ratio is much greater than in lower vertebrates such as fish or amphibians. Anatomically and histologically, 30 brain centers can be distinguished, each of which can be experimentally shown to be the functional center for a particular kind of central nervous system activity. Octopuses are remarkably tough experimental animals and rapidly recover from brain lesions. Under totally controlled experimental conditions, octopuses (and cuttlefish of genus *Sepia*) can readily be trained to distinguish visually between pairs of geometric figures, and tactilely between textures and tastes of different surfaces. Thus investigations linking behavioral training to the precise local nerve cells involved in learning and memory have been possible in *Octopus*.

Memory modules consisting of small numbers of nerve cells can have certain circuit biases altered by training, and Young has developed unit models of them with wiring diagrams which correspond to the actual neural microanatomy in the brain of *Octopus*. Experimental lesions in octopuses, both already trained and undergoing training, have established that such neural memory systems are spatially limited to specific lobes. Some progress has been made in defining in physicochemical terms the unit changes which take place during learning. Significantly, as in other brains, short-term and long-term memory systems are spatially separated in *Octopus*. *See* CEPHALOPODA; MOLLUSCA; NAUTILUS; OCTOPODA.

W. D. Russell-Hunter

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; M. J. Wells, *Octopus: Physiology and Behaviour of an Advanced Invertebrate*, 1978; J. Z. Young, *The Anatomy of the Nervous System of Octopus vulgaris*, 1971; J. Z. Young, *A Model of the Brain*, 1964.

Odonata

An order of the class Insecta, commonly known as dragonflies and damselflies (**Fig. 1**). Odonates generally are conspicuous insects with strong flying ability and a complex range of behaviors. They occur on all continents except Antarctica, although more species occur near the Equator than at higher lattitudes. The larvae are aquatic, inhabiting all types of freshwater, including streams, rivers, ponds, and marshes. The adults fly over or near these localities, though some species may disperse and are occasionally found far from water. *See* INSECTA.

Morphology. Adult Odonata range in size from small (with a wingspan of 15 mm, or 0.6 in.) to very large (with a wingspan greater than 150 mm, or 6 in.). They have large compound eyes on either a globular or transversely elongate head. The antennae are short and bristlelike. There are two pairs of elongate, membranous wings, which often are clear or transparent; however, in some species, the wings have patches or bands of color. The wings have a distinctive venation, with five main longitudinal veins that originate from the base and are connected by a network of many crossveins (Fig. 1). Adult males have secondary genital organs beneath the second abdominal segment, or the second and third segments—the organs are unique to this order. Females of some families have a conspicuous ovipositor.

Fig. 1. An aeshnid dragonfly (note the distinctive wing venation) and a lestid damselfly.

Fig. 2. Larvae of (*a*) a dragonfly and (*b*) a damselfly.

The larvae of Odonata are readily distinguished by a strongly developed, hinged labium, or lower lip, which is used to grasp prey (**Fig. 2**). This is held underneath the head when not in use; in some families it forms a mask covering the face.

Classification. Odonata contains more than 6000 species worldwide, and is divided into three suborders: the Anisoptera, or dragonflies; the Zygoptera, or damselflies; and an intermediate suborder called the Anisozygoptera. The term dragonfly is also used for the order as a whole. All the commonly encountered species belong to either Anisoptera or Zygoptera.

Anisozygoptera. Anisozygoptera contains just two living species, one from the Himalayas and the other from Japan, although additional fossil species are known. Anisozygopterans look like dragonflies but with damselfly wings.

Anisoptera and zygoptera. Anisoptera has been divided into nine families and Zygoptera into 18. Anisoptera can be distinguished from Zygoptera by differences in the wing venation, structural details of the male secondary genitalia, the grasping organs (claspers) at the end of the abdomen, and the larval gills. In dragonflies the hindwings are broader than the forewings and the venation develops differently; however, in damselflies the two pairs of wings are more or less equal. Dragonfly larvae "breathe" by means of gill folds inside the rectum, whereas larval damselflies have three (rarely two) relatively broad external gills at the end of the abdomen.

Although most dragonflies hold their wings horizontally when at rest, and most damselflies hold their wings closed above the abdomen, exceptions occur within each suborder. This trait is an unreliable guide to subordinal placement except in North America and Europe.

Developmental stages. There are three general stages in the life history: the egg, the larva (sometimes incorrectly called a nymph), and the adult.

Egg. Eggs are laid into emergent or submerged vegetation or in or on waterside vegetation, or they are dropped on to the water surface, where they sink to the bottom. Most eggs hatch after a few days, but some may overwinter or, in dry climates, may not hatch until it rains.

Larva. Larvae typically live in and around submerged vegetation or among the bottom detritus (Fig. 2). All species are predacious, the food being all types of aquatic invertebrates. Occasionally, a large dragonfly larva may attack and eat a tadpole or small fish. Larval development takes from about 5 weeks in some pond-dwelling species to several years in species that inhabit cold, upland streams. Many temperate species have one generation per year. During growth a larva undergoes some 10–15 molts. Wing buds begin to appear half-way through development and are prominent in later instars (stage between two molts). When ready for its final molt, the larva climbs out of the water and attaches itself to a support. The larval cuticle splits and the adult emerges. *See* INSECT PHYSIOLOGY; MOLTING (ARTHROPODA).

Adult. After its wings have dried, the young adult spends several days away from water before returning to breed. During this time, and for several days or weeks afterward, the adult colors gradually develop. Adults may emerge in a synchronized fashion, usually in spring, or else continuously throughout the warmer months. Adult lifespan ranges from about one week to several months, depending on the species and climate. In cold climates, some adults may survive the winter in sheltered spots; in hot climates, some estivate during the dry season.

Adult Odonata chiefly divide their time between feeding and breeding. As with the larvae, all species are predacious at the adult stage. They catch flying insects or pick them from the foliage. The females of many species spend long periods feeding away from water and may appear at breeding sites only to mate and lay eggs. Males typically congregate at or near the water's edge, where they stake out small territories, fight other males, and wait for females to appear. Courtship is not common except in some species of

Fig. 3. Libellulid dragonflies mating.

damselflies. In most species, the male simply grabs the female by the back of the head or the neck (prothorax), and the pair fly off together. In these tandem pairs, the male is in front and the female is behind. If the female is willing to mate, she bends her abdomen forward to collect sperm from the male's accessory genitalia, where he has previously deposited it (**Fig. 3**). Complicated sperm displacement activities, in which the male first removes any sperm previously inside the female, are common, and matings can last from several minutes to several hours. The subsequent act of egg laying is often accompanied by mate-guarding behavior. *See* REPRODUCTIVE BEHAVIOR.

Coloration. Males, especially of perching species, often are brightly colored. This coloration may be largely a warning to other males that the bearer is male and not worth pursuing. Some of the common colors, especially the bright blues and reds, are produced by movable color granules within the epidermal cells, and the resulting color varies with temperature. Unlike some other insects, the bright colors of Odonata do not indicate distastefulness or danger. Odonates are entirely harmless and without defenses beyond their superb eyesight and flying skills. The males of species that fly back and forth on a fixed beat and rarely perch are often the same color as females of their species. Camouflage colors are common in these species. *See* PROTECTIVE COLORATION.

Enemies. Common enemies of the larvae include other predacious aquatic invertebrates, wading birds, and fish. Small larvae also are at risk of attack from larger larvae of their own or other species. Birds prey on the adults and some are taken by other odonates. Birds, reptiles, frogs, and fish prey heavily on newly emerged and ovipositing adults.

Interaction with humans. Odonata are wholly beneficial insects, which (in large numbers) can ameliorate outbreaks of aquatic pest species. However, a large dragonfly can look frightening, and Western folklore offers several negative terms for these insects, including horse stinger and devil's darning needle. In China and Japan, dragonflies (including damselflies) are regarded as benign and auspicious insects. Odonate larvae are a useful natural source of food for many fish, and they are sometimes used by anglers as a fishing bait. Many Odonata are now at risk from pollution and the destruction of their habitats. *See* ENTOMOLOGY, ECONOMIC.

John Trueman

Bibliography. P. S. Corbet, *Dragonflies: Behaviour and Ecology of Odonata*, Harley Books, Colchester, UK, 1999; R. J. Tillyard, *The Biology of Dragonflies (Odonata or Paraneuroptera)*, Cambridge University Press, 1917.

Oedogoniales

An order of filamentous fresh-water green algae (Chlorophyceae) with unique morphological features including (1) an elaborate method of cell division that results in the accumulation of apical caps, (2) zoospores and antherozoids with a subapical crown of flagella, and (3) a highly specialized type of oogamy. There is a single family, Oedogoniaceae, comprising three genera. *See* CHLOROPHYCEAE.

Oedogonium has the largest number of species (several hundred) and is the most common of the three genera. Its unbranched filaments are initially attached by a holdfast cell to submerged vegetation, stones, or wood, usually in permanent ponds or pools, but at maturity they may form free-floating masses. The cells are cylindrical, each containing a reticulate chloroplast with numerous pyrenoids. At the beginning of mitosis, a ring of wall material is deposited on the inner face of the lateral wall a short distance below the distal end of the cell. When mitosis is complete, the wall external to the ring ruptures and the ring is stretched to form the new portion of the daughter cell wall. That part of the parental wall lying above the ring is visible as an apical cap. Successive divisions result in an accumulation of caps.

Vegetative multiplication by fragmentation is common. In asexual reproduction, zoospores with a subapical crown of up to 120 flagella are formed singly within a cell. In sexual reproduction, there is a highly specialized interplay between female and male elements. The egg is a metamorphosed protoplast of an enlarged spherical cell (oogonium), which is the upper of two cells resulting from division of an oogonial mother cell, the lower being the supporting (suffultory) cell (see **illustration**). Antherozoids, with a subapical crown of about 30 flagella, are produced in pairs or tetrads in very small discoid antheridia. The egg is fertilized in place, the antherozoid entering through a pore or fissure in the oogonial wall.

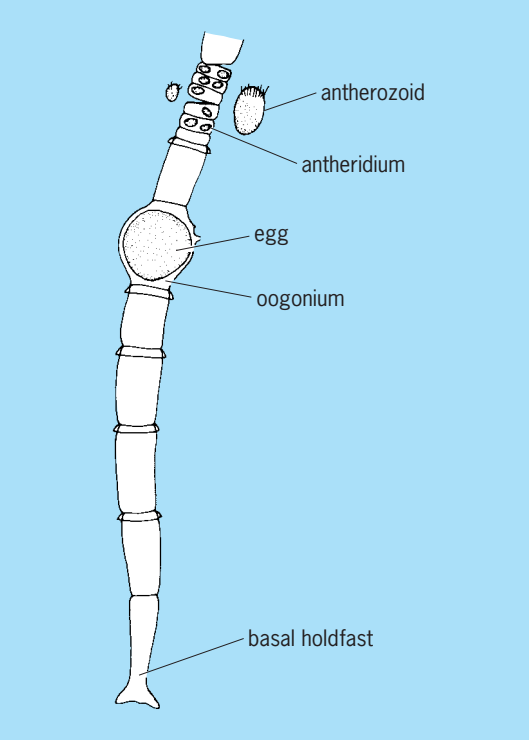

Nannandrous species of *Oedogonium*, showing sexual reproductive structures.

The two types of sex organs may occur on the same filament (homothallic species) or on different filaments (heterothallic species), but in either case certain species have an indirect development of antherozoids. These species, termed nannandrous in distinction to those with direct development (macrandrous), form short cylindrical cells which initially appear like antheridia, but in which the protoplast metamorphoses into a single swarmer bearing a subapical crown of flagella. These swarmers swim around the filament until they are attracted chemically to the oogonial mother cell, the suffultory cell, or less often the oogonium. Upon settling, a swarmer forms a one-celled germling (dwarf male filament) that acts as an antheridial mother cell, cutting off one or more antheridia. The swarmers that give rise to the dwarf male filaments are called androspores, while the short cylindrical cells from which they arise are called androsporangia. The biochemical basis of this indirect development of the male gamete is only partly understood. The zygote secretes a thick, frequently ornamented wall and often accumulates a reddish-orange pigment (hematochrome). After a period of dormancy, it germinates meiotically to form four zoospores. The zoospores reestablish the vegetative filaments, which thus are haploid.

Filaments of *Bulbochaete* are unilaterally branched, and most cells bear a long hyaline bristle with a bulbous base. Additional axial cells are usually cut off by the basal cell, but any vegetative cell can initiate a branch. *Bulbochaete* is found in the same habitats as *Oedogonium*, but is less common and has fewer species (about 125). Filaments of *Oedocladium* are branched but lack bristles. The 13 species, which are terrestrial or aquatic, are seldom encountered. They seem to be most common in warm regions of America, India, and Australia.

Paul C. Silva; Richard L. Moe

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae: Structure and Reproduction,* 1985; G. M. Smith, *The Fresh-Water Algae of the United States*, 1950; L. H. Tiffany, *The Oedogoniaceae*, 1930.

Ohmmeter

A portable instrument for measuring relatively low values of electrical resistance. The ohmmeter is ideal for use in the workshop, field, or test laboratory. The range of resistance measured is typically from 0.1 microhm to 1999 ohms (Ω) [**Fig. 1**].

Fig. 1. Electronic digital ohmmeter. (*Tinsley*)

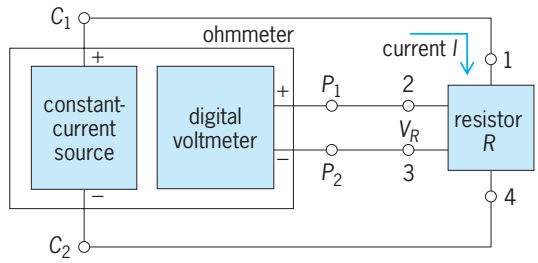

Fig. 2. Diagram of ohmmeter, showing operating principles. C_1, C_2 = current terminals; P_1, P_2 = potential terminals.

Operating principles. The ohmmeter may be either an analog or a digital instrument. It uses a four-terminal measurement method (**Fig. 2**) in which a constant-current source is applied to the current terminals 1 and 4 of the resistor R via the current terminals C_1 and C_2 on the ohmmeter. In the digital type of ohmmeter, a digital voltmeter is then connected across the potential terminals 2 and 3 of the resistor R via the terminals P_1 and P_2 on the ohmmeter.

The error caused by the shunting effect of the digital voltmeter is almost negligible because the digital voltmeter can have an input resistance in excess of 10 gigohms (10^{10} Ω) on its basic ranges. The contribution due to the shunting error is less than 0.01% when measuring a test resistor of up to 1 megohm (10^6 Ω).

By Ohm's law, the voltage drop V_R across the resistor (Fig. 2) is related to its resistance R and the current I passing through it by Eq. (1). If the value

$$R = \frac{V_R}{I} \qquad V_R = IR \tag{1}$$

of the resistor R is chosen to be 1 kΩ and the current is set to be a constant value of 1 mA, then the voltage drop across the resistor will be given by Eq. (2). The digital voltmeter connected across the

$$V_R = 0.001\ \text{A} \times 1000\ \Omega = 1\ \text{V} \tag{2}$$

resistor will indicate 1.0000 V, assuming that a $4^1/_2$-digit voltmeter is used ($^1/_2$ digit refers to the fact that the digit before the decimal point can only be 0 or 1). If the value of the resistor R is changed to, say, 500 Ω and the current is constant at 1 mA, then the digital voltmeter will read 0.001 A × 500 Ω = 0.5000 V. *See* ELECTRICAL RESISTANCE; OHM'S LAW; VOLTMETER.

The digital voltmeter indication is clearly directly proportional to the value of the resistance R, and so the digital voltmeter display can be annotated to read in kilohms. Other resistance ranges can be added simply by changing the value of the constant current or by changing the full-scale range of the digital voltmeter.

In a practical ohmmeter, the current is not adjusted separately, but the instrument is calibrated by using known values of test resistance; therefore any errors in the digital voltmeter are compensated for. Once the instrument is calibrated, an unknown resistor R_x

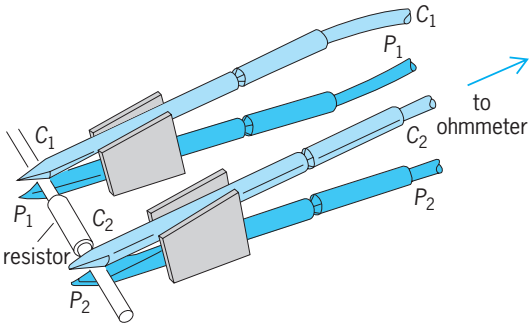

Fig. 3. Special measuring probes (Kelvin clips), consisting of a set of leads with combined current probes (C_1, C_2) and potential probes (P_1, P_2).

can be connected to it in place of R, and its value will be measured and indicated on the digital voltmeter.

The resistance measurement thus carried out provides a true four-terminal value, but resistors with only two terminals or wires can be measured by connecting the current and potential leads together in pairs. For most ohmmeters, special test leads are provided for the purpose of measuring two-terminal resistors, and **Fig. 3** shows a typical set of leads with combined current and potential probes, often known as Kelvin clips.

Applications. The ohmmeter solves quickly and easily a variety of difficult measurement problems, including measuring the resistance of cladding and tracks on printed circuit boards, electrical connectors, and switch and relay contacts, as well as determining the quality of ground-conductor continuity and bonding, cables, bus-bar joints, and welded connector tags. *See* RESISTANCE MEASUREMENT.

A. Douglas Skinner

Ohm's law

The statement that the current I flowing in an electrical circuit is often, to a very good approximation, proportional to the voltage V of its source; that is, $V = IZ$. If the current is steady direct current (DC), Z equals the total resistance of the circuit, R.

Ohm's law, although closely obeyed in most metallic and some other conductors, is an approximation, and does not have the same physical status as, say, Maxwell's equations. *See* MAXWELL'S EQUATIONS.

If Z or R is independent of I to a sufficient approximation, the circuit and the elements in it are said to be linear and network theory can be applied. If V and I are time-dependent, alternating-current circuit theory applies to their frequency components. The current is derived from the voltage by considering not only the resistances but also the vector impedances composed of self- and mutual inductances and capacitances together with their losses. These latter components combine into the total impedance Z, and are often also linear toa sufficient approximation. *See* ALTERNATING-CURRENT CIRCUIT THEORY; CAPACITANCE; ELECTRICAL IMPEDANCE; INDUCTANCE; NETWORK THEORY.

This 'law' was first described by Georg Simon Ohm in 1827 as a result of his experiments with metallic conductors. *See* ELECTRICAL RESISTANCE; ELECTRICAL RESISTIVITY; RESISTANCE MEASUREMENT; SEMICONDUCTOR; SKIN EFFECT (ELECTRICITY); THERMAL CONDUCTION IN SOLIDS.

Bryan P. Kibble

Oichnus

A morphologically distinct type of trace fossil left by boring organisms in hard biogenic substrates. *Oichnus* Bromley 1981 is a formal taxonomic name with the rank of ichnogenus (a trace fossil genus). *Oichnus* is a Latin-derived term that literally means a trace (ichnus) shaped like the letter "O"; it denotes the small circular to subcircular holes (borings) found primarily in fossilized skeletal remains of invertebrate animals, such as mollusk shells, ostracode valves, or echinoid tests.

Classification. *Oichnus* includes several ichnospecies (trace fossil species) that are classified based on differences in the vertical cross section of a boring (cylindrical, conical, etc.), horizontal outline of its outer and inner openings (circular or oval), and

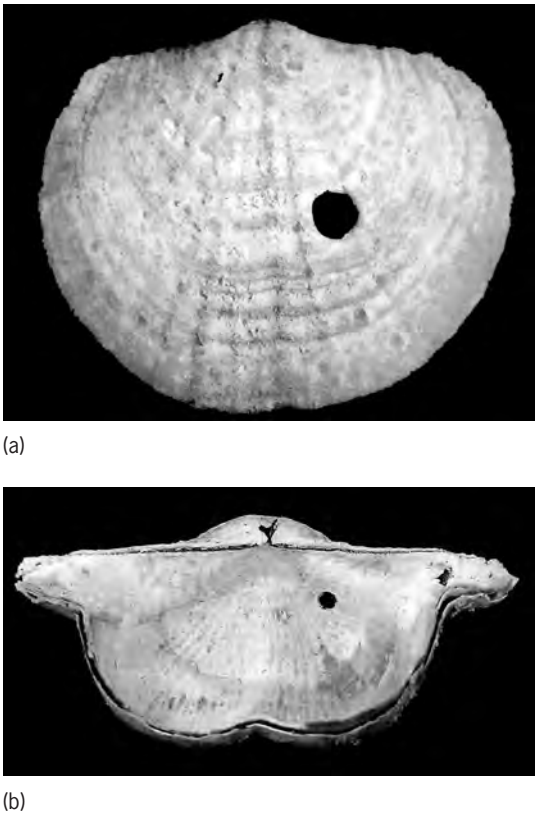

(a)

(b)

Fig. 1. Two examples of *Oichnus simplex* drilled by some unknown predators or parasites in shells of Late Paleozoic brachiopods extracted from the Permian rocks of West Texas. The specimens are from the Copper and Grant Brachiopod Collection housed in the National Museum of Natural History, the Smithsonian Institution. (*Courtesy of Finnegan Marsh, National Museum of Natural History*)

the degree of penetration (holes or pits). Seven ichnospecies have been named formally so far, including the type ichnospecies *O. simplex* Bromley 1981 (**Fig. 1**), as well as six other forms: *O. paraboloides* Bromley 1981, *O. ovalis* Bromley 1991, *O. asperus* Nielsen and Nielsen 1991, *O. coronatus* Nielsen and Nielsen 1991, *O. gradates* Nielsen and Nielsen 1991, and *O. excavatus* Donovan and Jagt 2002. The ichnogenus *Tremichnus* Brett 1985 is viewed by some as the junior synonym of *Oichnus*. *See* TRACE FOSSILS.

Shell drilling. In modern ecosystems *Oichnus*-producing drilling organisms are widespread and diverse in many marine habitats. These drillers are often predators that bore to access the soft tissue of their shelled prey. However, externally parasitic (ectoparasitic) borers (which affect hosts negatively without killing them), commensal borers (which benefit without harming their hosts), and even amensal borers (which do not benefit from boring but are detrimental to their hosts), have been documented as well. Shell drilling is a highly convergent behavior that evolved multiple times independently in unrelated groups of organisms; however, the most well-known and studied drillers are boring gastropods: naticids and muricids. In addition, many other groups of snails, as well as some species of octopods, nematodes, and flatworms, are capable of drilling holes in their victims. Drilling ability has also been demonstrated in an extinct group of Paleozoic gastropods (Platyceratidae). In modern oceans, drillings of *Oichnus* type are found in a broad range of marine invertebrate shells, including (most frequently) mollusk prey and hosts and (less commonly) other organisms with biomineralized skeletons, such as echinoderms, arthropods, brachiopods, bryozoans, and foraminiferans.

Paleoecology of drilling predation. *Oichnus* is often easy to identify in the fossil record, and its ecological genesis can be evaluated rigorously using a series of qualitative and quantitative criteria. Most widespread and well-studied fossil types of *Oichnus* borings, such as *O. simplex* and *O. paraboloides*, are attributed typically (based on extrapolation from modern ecosystems) to predatory and/or ectoparasitic organisms that drilled holes in external skeletons of live victims. Because of its high informative value and frequent occurrence in the fossil record, *Oichnus* provides arguably some of the best direct paleontological evidence for ecological interactions between ancient marine predators (or parasites) and their prey (or hosts).

Oichnus borings occur continuously through the entire fossil record of marine macroorganisms, since the dawn of metazoan animals in the Late Proterozoic Era (approximately 600 million years ago) until present day (**Fig. 2**). Even the oldest known skeletal remains from the Late Proterozoic, the enigmatic wormlike tubes *Cloudina*, bear traces that are believed to represent the oldest documented case of biotic interactions between drilling predators and their skeletonized prey. Numerical studies of drilling frequencies, based on *Oichnus* trace fossils, have offered some of the most compelling evidence

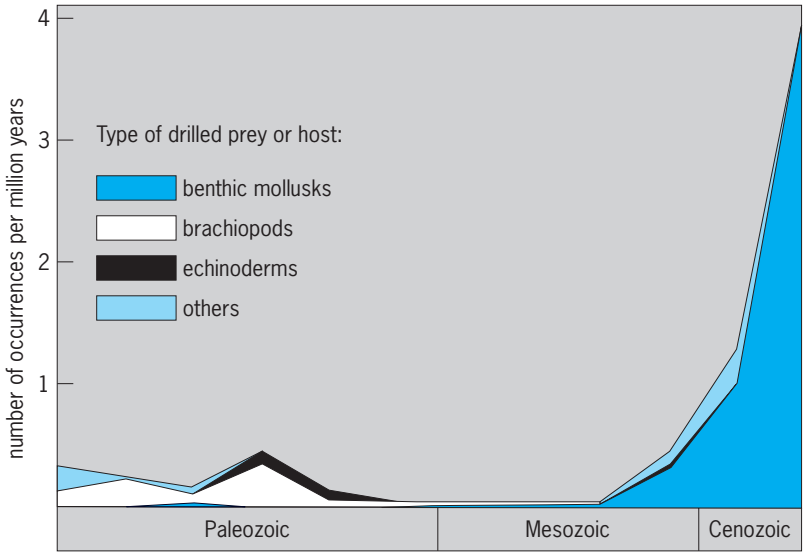

Fig. 2. Phanerozoic history of drilling predation based on fossil occurrences of the trace fossil *Oichnus* in shells of various marine invertebrates. A sharp increase in the Late Mesozoic and Cenozoic coincides with the Mesozoic Marine Revolution—a rapid radiation of many predatory groups coupled with an increase in substrate burrowing by marine benthic animal prey such as bivalve mollusks and echinoids. (*Based on the literature compilation of M. Kowalewski et al., 1998*)

for a long-term increase in the intensity of predation through evolutionary time. *See* FOSSIL; PALEOECOLOGY; PALEONTOLOGY; PREDATOR-PREY INTERACTIONS.

Michal Kowalewski

Bibliography. R. G. Bromley, *Trace Fossils: Biology and Taphonomy*, Unwin Hyman, London, 1990; P. H. Kelley, M. Kowalewski, and T. A. Hansen, *Predator-Prey Interactions in the Fossil Record*, Topics in Geobiology Series 20, Plenum Press–Kluwer, New York, 2003; M. Kowalewski, A. Dulai, and F. T. Fuersich, A fossil record full of holes: The Phanerozoic history of drilling predation, *Geology*, 26: 1091–1094, 1998; G. J. Vermeij, *Evolution and Escalation: An Ecological History of Life*, Princeton University Press, 1987.

Oil analysis

Analysis of petroleum, or crude oil, to determine its value in modern refinery operations. In addition, procedures have been developed for analysis of lubricating oil.

For refinery operations, oil analysis, or assay, must provide the refinery planner with the data needed to predict yields, qualities, and operating costs for a wide variety of refinery operating conditions and product demands. In a refinery, crude oil is distilled and separated into products according to the boiling points of the crude oil components (see **table**). *See* PETROLEUM; PETROLEUM PRODUCTS.

Procedure. A crude assay follows much the same procedure. The oil is distilled and separated into up to 40 narrow-boiling-range cuts. Cuts with normal boiling points higher than about 370°F (190°C) are distilled at greatly reduced pressure. This allows the oil to vaporize at temperatures below 370°F (190°C) and avoids thermal reactions which would alter the

Typical products derived from crude oil

Product	Carbon atom range	Boiling point range, °F (°C)
Gas	C_1 and C_2	
Liquefied petroleum gas	C_3 and C_4	
Gasolines	C_4 to C_{10}	59–370 (15–190)
Kerosines	C_9 to C_{15}	300–540 (150–280)
Middle distillates	C_{12} to C_{20}	390–640 (200–340)
Gas oils	C_{20} to C_{45}	640–1040 (340–560)
Bottoms	Unvaporized residua	

chemical properties of the oil. Each cut is then subjected to a variety of tests sufficient to characterize it for the products the cut could be included in. The refinery planners may then calculate yield and qualities of any product by blending the yields and qualities of the cuts that are included in the product.

Petroleum consists primarily of compounds of carbon and hydrogen containing from 1 to about 60 carbon atoms. Carbon atoms in natural petroleum occur in straight and branched chains (paraffins), in single or multiple saturated rings (cycloparaffins or naphthenes), and in cyclic structures of the aromatic type such as benzene, naphthalene, and phenanthrene. Cyclic structures may have attached to them side chains of paraffinic carbons. In lubricating oil it is usual to have naphthene rings built onto the aromatic rings and side chains attached. In products produced by cracking in the refinery, olefins or compounds with carbon-carbon double bonds not in aromatic rings are also found. The high-boiling fractions of petroleum contain increasing amounts of oxygen, nitrogen, and sulfur compounds, as well as traces of organic compounds of metals such as vanadium, nickel, and iron.

Tests. Three types of tests are used to characterize a crude oil and its narrow-boiling-range cuts: (1) tests for physical properties such as specific gravity, refractive index, freeze temperature, vapor pressure, octane number, and viscosity; (2) tests for specific chemical species such as sulfur, nitrogen, metals, and total paraffins, naphthenes, and aromatics; (3) tests for determining actual chemical composition.

Physical properties. The physical properties are used to predict how petroleum products will perform. Octane number and vapor pressure are two important qualities of gasolines. Jet fuels must be formulated so that they remain fluid, and therefore pumpable, at the low temperatures encountered by high-flying airplanes. Lubricating oils are characterized by their viscosity level together with the change of viscosity with temperature. The many different tests used for physical properties are relatively simple. Methods have been standardized and published by the American Society for Testing and Materials (ASTM).

Chemical species. The amounts of specific chemical species are also important in evaluating petroleum products. In these days of environmental concern and emission controls, the sulfur content often determines whether or not a fuel can be used. Aromatic components are desirable in gasoline and undesirable in diesel fuels. Fractions with vanadium, nickel, and iron compounds should not be used in catalyic processes since these metals permanently deactivate most catalysts. Tests are available to measure all the chemical species of interest. Many of the tests involve controlled burning of an oil sample, followed by analysis of the combustion products.

Chemical composition. The analysis of petroleum in terms of chemical compositoin can be done with varying degress of thoroughness. Anything approaching the complete analysis of a crude oil is so time-consuming and expensive that in the whole world only one sample of crude oil has been analyzed thoroughly. This project, known as Project No. 6 of the American Petroleum Institute, was carried on from 1928 to 1968, when the remaining work was transferred to another project. Obviously, such a detailed analysis will never be done on a routine basis.

Gas chromatography, mass spectroscopy, and nuclear magnetic resonance (NMR) methods are capable of economically determining the detailed chemical composition of petroleum fractions through the C_9 or C_{10} boiling range plus the distribution of hydrocarbon types up to about C_{40}. The hydrocarbon-type analysis includes distinguishing between one-, two-, and four-ring napthenes and aromatics. *See* CHROMATOGRAPHY; MASS SPECTROMETRY; NUCLEAR MAGNETIC RESONANCE (NMR).

In a gas chromatography analysis a small sample is introduced into a column packed with an adsorbent material such as silica gel. The sample is swept along the column by an inert carrier gas such as nitrogen. The carrier gas forces all components in the sample to move along the column, but each component moves at a unique rate depending on how strongly it is absorbed by the silica gel. Each component leaves the column as a discrete packet, or peak, of noncarrier gas. The time of each peak identifies the component, and the amount in the peak gives the composition. The operation of the gas chromatography column and the calculation of results have been automated so that very little human effort is required. However, setting up a gas chromatography column is a tedious procedure since each column seems to behave differently and so must be calibrated for all components it will be used to measure. Also, the method has trouble distinguishing between similar isomers that give overlapping peaks. *See* GAS CHROMATOGRAPHY.

In a mass spectroscopy analysis the molecules of the sample being tested are bombarded with an

electron beam. This breaks the molecules into ionized fragments whose weights are determined by measuring how much the ionized fragrments bend as they move through a magnetic field. The fragment pattern obtained is then correlated against the known fragment patterns of pure compounds to find the composition of the sample. It is difficult to apply the method to compounds for which calibration patterns are not available. It is, however, not feasible to obtain calibration patterns for the many thousands of compounds which appear in petroleum. Nevertheless, even in the higher-boiling fractions where these uncertainties exist, the mass spectroscopy analysis is the best available method for hydrocarbon type analysis.

An NMR analysis in its simplest form determines the type of hydrogen present in a sample, that is, whether the hydrogen is attached to an aliphatic, naphtenic, or aromatic type carbon atom. This is determined from a measurement of the magnetic resonance spectrum of the diluted sample under the influence of a very strong magnetic field. Information of this type, coupled with other analytical information, makes possible the development of a rather detailed picture of molecular types and structures.

Jack Rees

Lubricating oil. Analysis of used lubricating oils is part of the maintenance program for many types of engines and industrial equipment. By conducting baseline oil analysis and subsequent used-oil analysis on a regular schedule as part of a preventive maintenance program, mechanical and lubricant-related failures may be greatly reduced. Each use of lubricating-oil analysis was primarily to detect water in the steam-turbine and propulsion systems of large ships. Later, lubricating-oil analysis for mechanical and lubricating-related problems became part of the preventive maintenance program for helicopters and jet-powered aircraft, reducing engine failures and extending the time interval between major overhauls. Owners and operators of bus and truck fleets, and almost all types of engines and other equipment that use liquid lubricant, may use lubricating-oil analysis to prevent or minimize mechanical failure while extending the useful service life of expensive equipment.

To analyze a lubricating oil, a representative sample of the lubricating flowing in the system at normal operating temperature is collected in a clear container and labeled with pertinent data. The sample is then delivered to the testing laboratory as soon as possible. Tests applied to used engine oils include viscosity, fuel dilution, water, insolubles, and spectrochemical or spectrographic analysis. Wear-metal concentrations are determined in parts per million (ppm); these metals may include silver, aluminum, chromium, copper, iron, nickel, lead, and tin, in addition to silicon. An increase in any of these concentrations may indicate abnormal engine wear. High copper and lead levels usually indicate bearing distress. A high aluminium level usually indicates piston, bearing, or cylinder wear. A high chromium level usually indicates cylinder liner or piston ring wear. The presence of silicon (dirt) is often accompanied by a corresponding increase in wear-metal concentrations. Boron, sodium, and water indicate an internal coolant leak in the engine. *See* ENGINE; INTERNAL COMBUSTION ENGINE; SPECTROSCOPY.

Donald L. Anglin

Bibliography. American Society for Testing and Materials, *ASTM Standards*, annually; ASTM, *ASTM and Other Specifications for Petroleum Products and Lubricants*, 4th ed., 1985; ASTM, *Calculation of Physical Properties of Petroleum Products from Gas Chromatographic Analysis*, 1975; ASTM, *Miscellaneous ASTM Standards for Petroleum Products*, 15th ed., 1981; ASTM, *Petroleum Products, Lubricants, and Fossil Fuels*, 1986; L. F. Hatch and S. Matar, *Hydrocarbons to Petrochemicals*, 1981; Institute of Petroleum, *Standard Methods of Analysis and Testing of Petroleum and Related Products*, 1987; S. H. Kagler, *Spectroscopic and Chromatographic Analysis of Mineral Oil*, 1973.

Oil and gas, offshore

Oil and gas prospecting and exploitation are conducted on continental shelves and slopes in the seas around the world. Exploration for offshore petroleum and natural gas expanded rapidly after the end of World War II to include exploration, drilling, and production off the coasts of more than 75 nations. *See* NATURAL GAS; OIL AND GAS FIELD EXPLOITATION; PETROLEUM.

For many years, petroleum companies stopped at the water's edge or sought and developed oil and gas accumulations only in inland waters or shallow seas bordering onshore producing areas. Exploration deeper under the sea, and production from the continental shelves beyond territorial limits, did not begin in earnest until the world's increasing demand for petroleum energy sources, coupled with a lessening return from land drilling, provided the incentives for the huge investments needed for drilling in the open sea. Recent advances in technology permit drilling and production from ever deeper waters on the continental slopes. *See* CONTINENTAL MARGIN; OIL AND GAS WELL DRILLING; PETROLEUM RESERVES.

Historical development. Initial offshore development was a simple extension of land practice. The first offshore well was drilled in 1897 from a wharf made of wooden piling and timbers that extended about 300 ft (90 m) into the Pacific Ocean near Santa Barbara, California. By the early 1930s, oilfields had been discovered in the inshore and coastal areas of Lake Maracaibo, Venezuela; Louisiana; and the Caspian Sea. In 1938, an oilfield 1 mi (1.6 km) off the coast of Louisiana was discovered by directional drilling from land. The first platforms out of sight of land were installed in the Gulf of Mexico in 1947.

Since then, the offshore oil industry has moved progressively into deeper water and farther away from shore. Today offshore oil exploration and production is a worldwide industry. By the early 1990s, offshore sources accounted for 30% of worldwide

crude oil production and 14% of worldwide natural gas. Until now, most offshore production came from reservoirs located under the continental shelf, in water depths up to 600 ft (180 m). Spurred by technology and a need to find additional secure sources of energy, exploration and production are now moving even farther from shore and into the deeper waters of the continental slopes. Exploration wells have been drilled in water deeper than 9000 ft (2750 m), and hydrocarbons are being produced from offshore fields in waters deeper than 5000 ft (1500 m).

Geology and the sea. There is a sound geologic basis for the petroleum industry turning to the continental shelves and slopes in search of needed reserves. Favorable sediments and structures exist beneath the present seas of the world in geologic settings that have proven highly productive onshore. In fact, the subsea geologic similarity, or in some cases superiority, to geologic conditions on land has been a vital factor in the expansion of the world's investment in offshore exploration and production. *See* MARINE GEOLOGY.

Drilling in deep water has alleviated early fears regarding geochemistry and reservoir quality in deep-water environments. Many young, shallow reservoirs have proved to be oil-productive, with high-quality reservoirs capable of producing at the rates required to justify investment, and virtually all of the world's continental shelves have received some geologic study.

By the early 1990s, major offshore petroleum provinces had been established in the Beaufort Sea; Cook Inlet; offshore California; Gulf of Mexico; West Africa offshore Nigeria, Gabon, and Angola; Campos Basin offshore Brazil; North Sea; Arabian Gulf; Red Sea; and offshore India, Thailand, Indonesia, Malaysia, and Australia. The recent industry move into deeper water has greatly extended the potential of many of these areas.

Significant exploration is taking place in water depths up to 10,000 ft (3050 m) in the Gulf of Mexico, offshore Brazil, offshore West Africa, and in the northern Atlantic off the coasts of Norway, Ireland, and the United Kingdom. Oil and natural gas fields in waters as deep as 5500 ft (1675 m) are already on-stream in the Gulf of Mexico and offshore Brazil. In the North Atlantic west of the Shetland Islands, where weather conditions are particularly severe, production has started from reservoirs in water depths up to 1500 ft (450 m). *See* PETROLEUM GEOLOGY.

In the United States in the early 1990s, the Atlantic coast, south Florida, the Pacific coast, and most of offshore Alaska were placed under a 10-year moratorium preventing new exploration. This moratorium has been extended and is due to expire in 2012. The move into deep water has therefore been limited to the Gulf of Mexico. Here it is fueled by the geological attractiveness of the area, and has been greatly assisted by advances in geophysical technology such as three-dimensional seismic tomography. It has also been encouraged by the partial royalty relief provided for developments in water depths greater than 650 ft (200 m), which helps offset the considerable development cost. More than 25% of all discoveries in the Gulf of Mexico are now being made in waters deeper than 1000 ft (300 m). In 1998 more than 20% of the total barrel-of-oil-equivalent production in the Gulf of Mexico came from deep-water fields, and there are already many more prospects for possible development in water depths up to 6500 ft (2000 m). The U.S. Minerals Management Service estimated that by year-end 2000, almost two-thirds of the oil and one-third of the gas produced in the Gulf of Mexico would be from reservoirs located in water depths greater than 1000 ft. *See* SEA-FLOOR IMAGING; SEISMOLOGY.

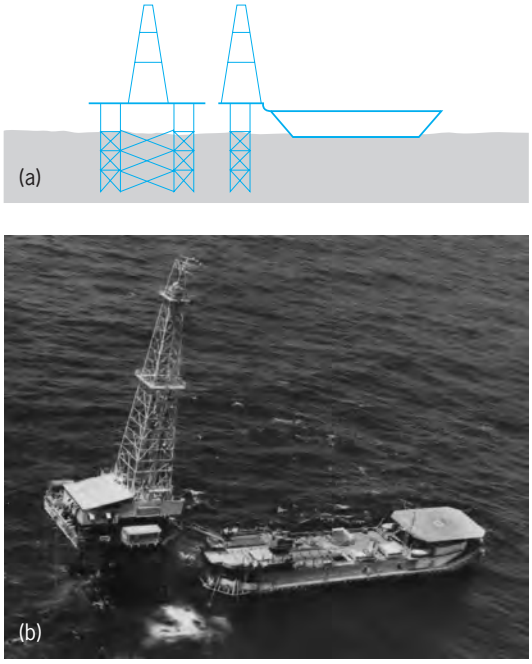

Fig. 1. Offshore fixed drilling platform. (*a*) Underwater design (*World Petroleum*). (*b*) Rig on a drilling site (*Marathon Oil Co.*).

Mobile drilling platforms. The underwater search has been made possible only by vast improvements in offshore technology. Drillers first took to sea with land rigs mounted on barges towed to location and anchored, or with fixed platforms accompanied by a tender ship (**Fig. 1**). As the search for oil and natural gas advanced worldwide and farther away from shore, types of exploration rigs evolved which could move easily between locations and operate in a wide range of water depths.

Self-elevating platform. In shallow water, the most widely used mobile platform is the self-elevating, or jack-up, unit (**Fig. 2**). It is towed to location, where the legs are lowered to the sea floor, and the platform is jacked-up above wave height. These self-contained platforms are especially suited to wildcat and delineation drilling, but have also been converted to fixed production use. They are best on firmer sea bottoms, with a depth limit out to 300 ft (90 m) of water for all but a few of the largest units.

Submersible platforms. These were developed from earlier submersible barges which were used in

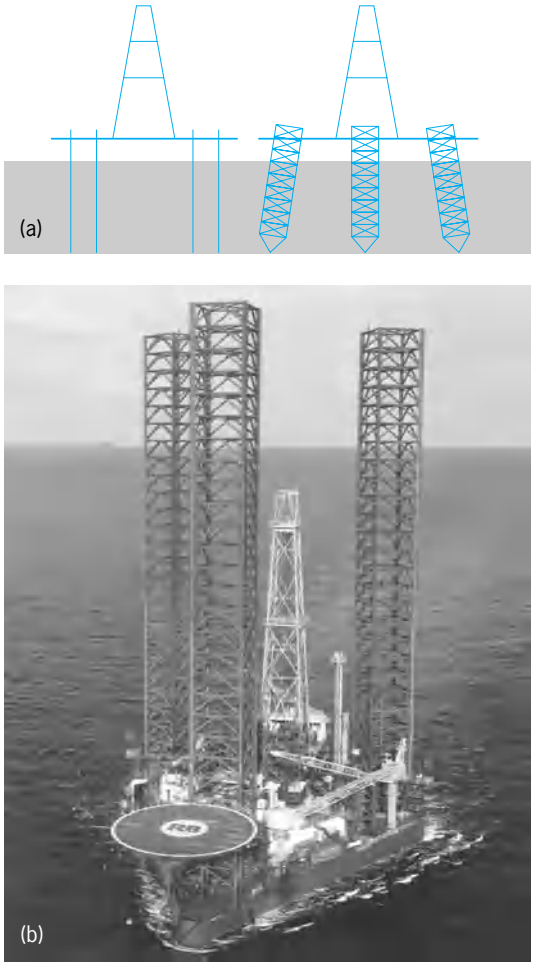

Fig. 2. Offshore self-elevating drilling platform. (*a*) Underwater design (*World Petroleum*). (*b*) Jack-up drilling rig (*Reading and Bates*).

shallow inlet drilling along the United States Gulf coast. The platforms are towed to the location and then submerged to the sea bottom. They are very stable and can operate in areas with soft sea floors. Difficulty in towing is a disadvantage, but this is partially offset by the rapidity with which they can be raised or lowered, once on location. They are not suitable for deep water and therefore now have restricted use.

Semisubmersible platforms. These are a version of submersibles (**Fig. 3**). They can work as bottom-supported units or in deep water as floaters. They can operate in a wide range of water depths, and when working as floaters, their primary buoyancy lies below the action of the waves, thus providing greater stability.

Floating drill ships. These are capable of drilling in the deepest waters (**Fig. 4**). They are built as self-propelled ships, or with a ship configuration that requires towing. Floating drill ships use anchoring, or in very deep water, dynamic-positioning systems, to stabilize their position. Floaters cannot be used in waters much shallower than 70 ft (20 m). This is because the special drilling equipment is then subject to excessive vertical movement from waves and tidal changes, and horizontal shifts, due to stretch and play in anchor lines, are too large compared with the water depth.

Exploration in deeper water requires the construction of new semisubmersibles and floating drill ships,

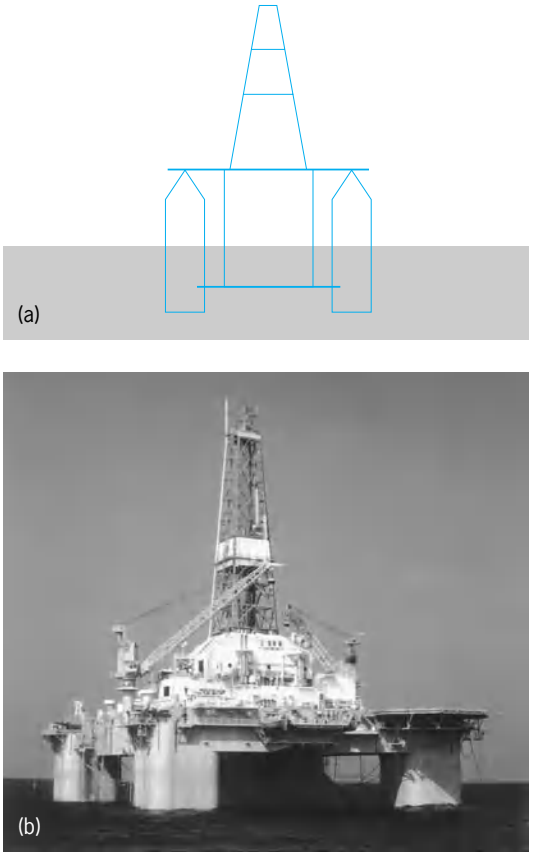

Fig. 3. Offshore semisubmersible drilling platform. (*a*) Underwater design (*World Petroleum*). (*b*) Dynamically positioned fourth-generation semisubmersible (*Reading and Bates*).

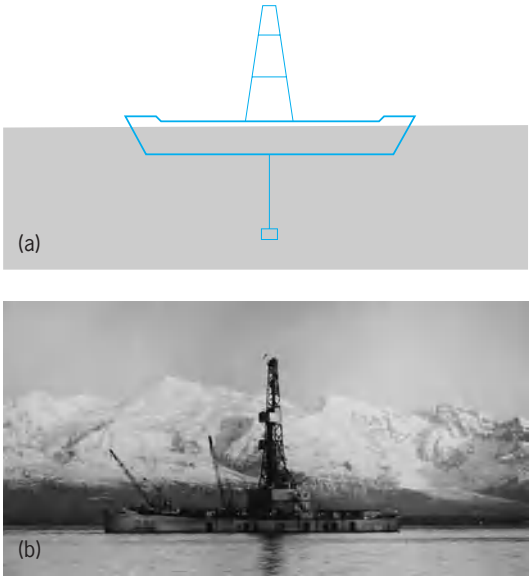

Fig. 4. Floating drill ship. Such ships can drill in depths from 60 to 1000 ft (18 to 300 m) or more. (*a*) Underwater design (*World Petroleum*). (*b*) Floating drill ship on a drilling site (*Marathon Oil Co.*).

Fig. 5. Offshore drilling and production platform.

which can operate in rough seas far from shore and take the large deck loads required for drilling wells in water depths up to 10,000 ft (3000 m).

The *Glomar Challenger* has drilled stratigraphic holes in water depths of 20,000 ft (6 km) to a depth of 3334 ft (1016 m) below the sea floor, but the technology for stratigraphic investigation is simpler and less demanding than that required for oil exploration and development. *See* OCEANOGRAPHIC VESSELS.

Production and well-completion technology. The move of exploration into the open and often hostile sea has required not only the development of drilling vessels but also a host of auxiliary equipment and techniques. An entire industrial complex has developed to serve the offshore industry, including construction of fixed platform structures from which the majority of the world's offshore oil and gas production is presently drilled and produced (**Fig. 5**). More than 7000 fixed platforms have been fabricated and installed around the world. There are approximately 3900 fixed platforms in the waters of the United States outer continental shelf in the Gulf of Mexico, of which 1400 are larger structures with more than five well slots. Most of these are located in less than 350 ft (100 m) of water, and as recently as 1994 fewer than 20 were in waters deeper than 600 ft (180 m).

The industry has traditionally modified established onshore practices for offshore operations, and wherever possible until now has preferred surface wellheads on fixed platforms. In deep water the structural cost of the surface facility is increasingly significant, leading to large platforms with heavy topsides and many well slots. Directional drilling is necessary in order to reach all the well targets within the reservoir from a single platform location. More recently, technology has been developed which enables the last part of the wellbore to be drilled horizontally, further extending platform reach and the reserves which can be recovered from each location. *See* DRILLING, GEOTECHNICAL.

Since traditional fixed platforms directly resist the force of the wind and waves, they become massive and increasingly expensive in deeper water. Therefore, the deepest traditional fixed platform installed to date worldwide is in 1353 ft (412 m) of water in the Gulf of Mexico. Compliant towers rely on the damping effect of the ocean to partly resist the horizontal forces, and extend the economic limit of bottom-founded steel structures to around 3000 ft (900 m). The first two compliant production towers were installed in the Gulf of Mexico in 1998, the taller in a water depth of 1754 ft (535 m). As the top of the drilling rig is 2094 ft (638 m) above the sea bed, this is also the tallest free-standing structure in the world.

Driven by the high cost of fixed platforms, alternative technologies have been developed for marginal fields and deeper water. These include floating production systems, which rely on subsea completions where oil and gas is produced from "Christmas trees" located on the sea floor, and the tension-leg platform (**Fig. 6**), which is a buoyant structure fixed to the seabed by means of tensioned cables which permits surface trees. Such systems are now used to develop deep-water and marginal fields in all major offshore areas around the world.

Recent advances in seabed production and processing systems permit deep-water and marginal fields to be developed 50 mi (80 km) or more away from a fixed- or floating-platform host facility. If production must flow long distances, there is a potential for hydrates and waxes to plug the line; therefore, methods to clean the line and pressure boosting pumps may be required. Subsea control umbilical units are used to transmit hydraulic and electric power signals across the sea floor from the main process facility to the subsea well-control equipment. *See* HYDRATE.

In water depths out of the range of practical diving, remotely operated vehicles equipped with robotic arms and specialty tools for moving, torquing, winching, and applying hydraulic power must now

Fig. 6. Tension-leg platform. (*U.S. Minerals Management Service*)

do all the work that divers once managed. The depth range of remotely operated vehicles is virtually unlimited, whereas diver work ceases to be practical and cost-effective beyond 600–800 ft (180–250 m). Drilling and production units that are totally enclosed and sit on the sea bottom have been considered, but none has yet proved practical or cost-effective when compared to surface-based operations using remotely operated vehicles to perform the subsea intervention tasks. *See* DIVING; OIL AND GAS WELL COMPLETION; UNDERWATER VEHICLE.

A number of issues need to be resolved as drilling and production operations move into waters deeper than 5000 ft (1500 m) into an environment that is still comparatively unknown. For instance, the hydrostatic head of the water column is such that it makes good sense to drill from the sea floor rather than from the surface. If satellite fields are to be brought on-stream miles from their host platforms, then wellheads, pumps, and other process equipment are required to operate reliably on the sea bed for long periods of time and yet be retrievable for maintenance and repair. Hydraulic power has limitations over long distances and in deep water, so high-voltage electrical distribution and switching equipment will need to be developed for deep-water use.

Hazards at sea. The United States Gulf coast, where a large percentage of the world's offshore drilling has taken place, is regularly hit by hurricanes that damage structures in their path, and a number of platforms were lost early on due to high wind and waves. However, improved understanding of the environment, as well as more advanced materials and structural analysis techniques, has significantly reduced failures due to environmental forces. The major causes of accident are now due to human error, process hazards, and transportation to and from offshore facilities by helicopter. *See* HURRICANE.

Improved reservoir management and further discoveries have increased production from onshore fields, where costs are less than for offshore. However, despite the high cost involved in extracting oil and gas from beneath the sea, the world's increasing demand for petroleum energy continues to force the search for new reserves into even deeper waters and more remote corners of the world.

G. R. Schoonmaker; J. A. Turley; Cyril Arney

Bibliography. M. T. Halbouty (ed.), *Future Petroleum Provinces of the World*, 1986; M. T. Halbouty (ed.), *Giant Oil and Gas Fields of the Decade 1978-1988*, 1992; *Offshore Atlas of World Oil and Gas Theatres*, Pennwell Publishing, 1996; J. A. Pratt, T. Priest, and C. Costaneda, *Offshore Pioneers*, 1997.

Oil and gas field exploitation

In the petroleum industry, a field is an area underlain without substantial interruption by one or more reservoirs of commercially valuable oil or gas, or both. A single reservoir (or group of reservoirs which cannot be separately produced) is a pool. Several pools separated from one another by barren, impermeable rock may be superimposed one above another within the same field. Pools have variable areal extent. Any sufficiently deep well located within the field should produce from one or more pools. However, each well cannot produce from every pool, because different pools have different areal limits.

Development

Development of a field includes the location, drilling, completion, and equipment of wells necessary to produce the commercially recoverable oil and gas in the field.

Related oil field conditions. Petroleum is a generic term which, in its broadest meaning, includes all naturally occurring hydrocarbons, whether gaseous, liquid, or solid. By variation of the temperature or pressure, or both, of any hydrocarbon, it becomes gaseous, liquid, or solid. Temperatures in producing horizons vary from approximately 60°F (16°C) to more than 300°F (149°C), depending chiefly upon the depth of the horizon. A rough approximation is that temperature in the reservoir sand, or pay, equals 60°F (16°C), plus 0.017°F/ft (0.031°C/m) of depth below surface. Pressure on the hydrocarbons varies from atmospheric to more than 11,000 lb/in.^{2} (76 megapascals). Normal pressure is considered as 0.465 (lb/in.^{2})/ft (10.5 kilopascals/m) of depth. Temperatures and pressure vary widely from these average figures. Hydrocarbons, because of wide variations in pressure and temperature and because of mutual solubility in one another, do not necessarily exist underground in the same phases in which they appear at the surface.

Petroleum occurs underground in porous rocks of wide variety. The pore spaces range from microscopic size to rare holes 1 in. (2.5 cm) or more in diameter. The containing rock is commonly called the sand or the pay, regardless of whether the pay is actually sandstone, limestone, dolomite, unconsolidated sand, or fracture openings in relatively impermeable rock. *See* PETROLEUM.

Development program. After discovery of a field containing oil or gas, or both, in commercial quantities, the field must be explored to determine its vertical and horizontal limits and the mechanisms under which the field will produce. Development and exploitation of the field proceed simultaneously. Usually the original development program is repeatedly modified by geologic knowledge acquired during the early stages of development and exploitation of the field.

Ideally, tests should be drilled to the lowest possible producing horizon in order to determine the number of pools existing in the field. Testing and geologic analysis of the first wells sometimes indicates the producing mechanisms, and thus the best development program. Very early in the history of the field, step-out wells will be drilled to determine the areal extent of the pool or pools. Stepout wells give further information regarding the volumes of oil and gas available, the producing mechanisms, and the desirable spacing of wells.

The operator of an oil and gas field endeavors to select a development program which will produce the largest volume of oil and gas at a profit. The program adopted is always a compromise between conflicting objectives. The operator desires (1) to drill the fewest wells which will efficiently produce the recoverable oil and gas; (2) to drill, complete, and equip the wells at the lowest possible cost; (3) to complete production in the shortest practical time to reduce both capital and operating charges; (4) to operate the wells at the lowest possible cost; and (5) to recover the largest possible volume of oil and gas.

Selecting the number of wells. Oil pools are produced by four mechanisms: dissolved gas expansion, gas-cap drive, water drive, and gravity drainage. Commonly, two or more mechanisms operate in a single pool. The type of producing mechanism in each pool influences the decision as to the number of wells to be drilled Theoretically, a single, perfectly located well in a water-drive pool is capable of producing all of the commercially recoverable oil and gas from that pool. Practically, more than one well is necessary if a pool of more than 80 acres (32 hectares) is to be depleted in a reasonable time. If a pool produces under either gas expansion or gas-cap drive, oil production from the pool will be independent of the number of wells up to a spacing of at least 80 acres per well (1860 ft or 570 m between wells). Gas wells often are spaced a mile or more apart. The operator accordingly selects the widest spacing permitted by field conditions and legal requirements.

Major components of cost. Costs of drilling, completing, and equipping the wells influence development plans. Having determined the number and depths of producing horizons and the producing mechanisms in each horizon, the operator must decide whether a well will be drilled at each location to each horizon or whether a single well can produce from two or more horizons at the same location. Clearly, the cost of drilling the field can be sharply reduced if a well can drain two, three, or more horizons. The cost of drilling a well will be higher if several horizons are simultaneously produced, because the dual or triple completion of a well usually requires larger casing. Further, completion and operating costs are higher. However, the increased cost of drilling a well of larger diameter and completing the well in two or more horizons is 20–40% less than the cost of drilling and completing two wells to produce separately from two horizons.

In some cases, the operator may reduce the number of wells by drilling a well to the lowest producible horizon and taking production from that level until the horizon there is commercially exhausted. The well is then plugged back to produce from a higher horizon. Selection of the plan for producing the various horizons obviously affects the cost of drilling and completing individual wells, as well as the number of wells which the operator will drill. If two wells are drilled at approximately the same location, they are referred to as twins, three wells at the same location are triplets, and so on.

Costs and duration of production. The operator wishes to produce as rapidly as possible because the net income from sale of hydrocarbons is obviously reduced as the life of the well is extended. The successful operator must recover from the productive wells the costs of drilling and operating those wells, and in addition must recover all costs involved in geological and geophysical exploration, leasing, scouting, and drilling of dry holes, and occasionally other operations. If profits from production are not sufficient to recover all exploration and production costs and yield a profit in excess of the rate of interest which the operator could secure from a different type of investment, he is discouraged from further exploration.

Most wells cannot operate at full capacity because unlimited production results in physical waste and sharp reduction in ultimate recovery. In many areas, conservation restrictions are enforced to make certain that the operator does not produce in excess of the maximum efficient rate. For example, if an oil well produces at its highest possible rate, a zone promptly develops around the well where production is occurring under gas-expansion drive, the most inefficient producing mechanism. Slower production may permit the petroleum to be produced under gas-cap drive or water drive, in which case ultimate production of oil will be two to four times as great as it would be under gas-expansion drive. Accordingly, the most rapid rate of production generally is not the most efficient rate.

Similarly, the initial exploration of the field may indicate that one or more gas-condensate pools exist, and recycling of gas may be necessary to secure maximum recovery of both condensate and of gas. The decision to recycle will affect the number of wells, the locations of the wells, and the completion methods adopted in the development program.

Further, as soon as the operator determines that secondary oil-recovery methods are desired and expects to inject water, gas, steam, or, rarely, air to provide additional energy to flush or displace oil from the pay, the number and location of wells may be modified to permit the most effective secondary recovery procedures.

Legal and practical restrictions. The preceding has assumed control of an entire field under single ownership by a single operator. In the United States, a single operator rarely controls a large field and this field is almost never under a single lease. Usually, the field is covered by separate leases owned and operated by different producers. The development program must be modified in consideration of the lease boundaries and the practices of the other operators.

Oil and gas know no lease boundaries. They move freely underground from areas of high pressure toward lower-pressure situations. The operator of a lease is obligated to locate the wells in such a way as to prevent drainage of the lease by wells on adjoining leases, even though the adjoining leases may be owned by that operator. In the absence of conservation restrictions, an operator must produce petroleum from wells as rapidly as it is produced from

wells on adjoining leases. Slow production on one lease results in migration of oil and gas to nearby leases which are more rapidly produced.

The operator's development program must provide for offset wells located as close to the boundary of the lease as are wells on adjoining leases. Further, the operator must equip the wells to produce as rapidly as the offset produces, and must produce from the same horizons which are being produced in offset wells. The lessor who sold the lease to the operator is entitled to a share of the recoverable petroleum underlying the land. Negligence by the operator in permitting drainage of a lease makes the operator liable to suit for damages or cancellation of the lease.

A development program acceptable to all operators in the field permits simultaneous development of leases, prevents drainage, and results in maximum ultimate production from the field. Difficulties may arise in agreement upon the best development program for a field. Most states have enacted statutes and have appointed regulatory bodies under which judicial determination can be made of the permissible spacing of the wells, the rates of production, and the application of secondary recovery methods.

Drilling unit. Commonly, small leases or portions of two or more leases are combined to form a drilling unit in whose center a well will be drilled. Unitization may be voluntary, by agreement between the operator or operators and the interested royalty owners, with provision for sharing production from the well between the parties in proportion to their acreage interests. In many states the regulatory body has authority to require unitization of drilling units, which eliminates unnecessary offset wells and protects the interests of a landowner whose acreage holding may be too small to justify the drilling of a single well on that property alone.

Pool unitization. When recycling or some types of secondary recovery are planned, further unitization is adopted. Since oil and gas move freely across lease boundaries, it would be wasteful for an operator to repressure, recycle, or water-drive a lease if the adjoining leases were not similarly operated. Usually an entire pool must be unitized for efficient recycling, or secondary recovery operations. Pool unitization may be accomplished by agreement between operators and royalty owners. In many cases, difference of opinion or ignorance on the part of some parties prevents voluntary pool unitization. Many states authorize the regulatory body to unitize a pool compulsorily on application by a specified percentage of interests of operators and royalty owners. Such compulsory unitization is planned to provide each operator and each royalty owner his fair share of the petroleum products produced from the field regardless of the location of the well or wells through which these products actually reach the surface.

Exploitation Practices

Oil and gas production necessarily are intimately related, since approximately one-third of the gross gas production in the United States is produced from wells that are classified as oil wells, However, the naturally occurring hydrocarbons of petroleum are not only liquid and gaseous but may even be found in a solid state, such as asphaltite and some asphalts.

Where gas is produced without oil, the production problems are simplified because the product flows naturally throughout the life of the well and does not have to be lifted to the surface. However, there are sometimes problems of water accumulations in gas wells, and it is necessary to pump the water from the wells to maintain maximum, or economical, gas production. The line of demarcation between oil wells and gas wells is not definitely established since oil wells may have gas-oil ratios ranging from a few cubic feet (1 cubic foot $= 2.8 \times 10^{-2}$ m^3) per barrel to many thousand cubic feet of gas per barrel of oil. Most gas wells produce quantities of condensable vapors, such as propane and butane, that may be liquefied and marketed for fuel, and the more stable liquids produced with gas can be utilized as natural gasoline.

Factors of method selection. The method selected for recoving oil from a producing formation depends on many factors, including well depth, well-casing size, oil viscosity, density, water production, gas-oil ratio, porosity and permeability of the producing formation, formation pressure, water content of producing formation, and whether the force driving the oil into the well from the formation is primarily gas pressure, water pressure, or a combination of the two. Other factors, such as paraffin content and difficulty expected from paraffin deposits, sand production, and corrosivity of the well fluids, also have a decided influence on the most economical method of production.

Special techniques utilized to increase productivity of oil and gas wells include acidizing, hydraulic fracturing of the formation, the setting of screens, and gravel packing or sand packing to increase permeability around the well bore.

Aspects of production rate. Productive rates per well may vary from a fraction of a barrel [1 barrel (bbl) $= 0.1590$ m^3] per day to several thousand barrels per day, and it may be necessary to produce a large percentage of water along with the oil.

Field and reservoir conditions. In some cases reservoir conditions are such that some of the wells flow naturally throughout the entire economical life of the oil field. However, in the great majority of cases it is necessary to resort to artificial lifting methods at some time during the life of the field, and often it is necessary to apply artificial lifting means immediately after the well is drilled.

Market and regulatory factors. In some oil-producing states of the United States there are state bodies authorized to regulate oil production from the various oil fields. The allowable production per well is based on various factors, including the market for the particular type of oil available, but very often the allowable production is based on an engineering study of the reservoir to determine the optimum rate of production.

Useful terminology. A few definitions of terms used in petroleum production technology are listed below to assist in an understanding of some of the problems involved.

Porosity. The percentage porosity is defined as the percentage volume of voids per unit total volume. This, of course, represents the total possible volume available for accumulation of fluids in a formation, but only a fraction of this volume may be effective for practical purposes because of possible discontinuities between the individual pores. The smallest pores generally contain water held by capillary forces.

Permeability. Permeability is a measure of the resistance to flow through a porous medium under the influence of a pressure gradient. The unit of permeability commonly employed in petroleum production technology is the darcy. A porous structure has a permeability of 1 darcy if, for a fluid of 1 centipoise (10^{-3} Pa · s) viscosity, the volume flow is 1 $cm^3/(s \cdot cm^2)$ [10^{-2} $m^3/(s \cdot m^2)$] under a pressure gradient of 1 atm/cm (10 MPa/m).

Productivity index. The productivity index is a measure of the capacity of the reservoir to deliver oil to the well bore through the productive formation and any other obstacles that may exist around the well bore. In petroleum production technology, the productivity index is defined as production in barrels per day per pound per square inch drop in bottom-hole pressure. For example, if a well is closed in at the casinghead, the bottom-hole pressure will equal the formation pressure when equilibrium conditions are established. However, if fluid is removed from the well, either by flowing or pumping, the bottom-hole pressure will drop as a result of the resistance to flow of fluid into the well from the formation to replace the fluid removed from the well. If the closed-in bottom-hole pressure should be 1000 lb/in.^{2} (7 MPa), for example, and if this pressure should drop to 900 lb/in.^{2} (6 MPa) when producing at a rate of 100 bbl/day (a drop of 100 lb/in.^{2} or 0.7 MPa), the well would have a productivity index of one.

Barrel. The standard barrel used in the petroleum industry is 42 U.S. gallons (approx. 0.1590 m^3).

API gravity. The American Petroleum Institute (API) scale that is in common use for indicating specific gravity, or a rough indication of quality of crude petroleum oils, differs slightly from the Baumé scale commonly used for other liquids lighter than water. The **table** shows the relationship between degrees API and specific gravity referred to water at 60°F (16°C) for specific gravities ranging from 0.60 to 1.0.

Viscosity range. Viscosity of crude oils currently produced varies from approximately 1 centripoise (10^{-3} Pa · s) to values above 1000 cP (1 Pa · s) at temperatures existing at the bottom of the well. In some areas it is necessary to supply heat artificially down the wells or circulate lighter oils to mix with the produced fluid for maintenance of a relatively low viscosity throughout the temperature range to which the product is subjected.

In addition to wells that are classified as gas wells or oil wells, the term gas-condensate well has come into general use to designate a well that produces large volumes of gas with appreciable quantities of light, volatile hydrocarbon fluids. Some of these fluids are liquid at atmospheric pressure and temperature; others, such as propane and butane, are readily condensed under relatively low pressures in gas separators for use as liquid petroleum gas fuels or for other uses. The liquid components of the production from gas-condensate wells generally arrive at the surface in the form of small droplets entrained in the high-velocity gas stream and are separated from the gas in a high-pressure gas separator.

Production Methods in Producing Wells

The common methods of producing oil wells are natural flow; pumping with sucker rods; gas lift; hydraulic subsurface pumps; electrically driven centrifugal well pumps; and swabbing.

Numerous other methods, including jet pumps and sonic pumps, have been tried and are used to slight extent. The sonic pump is a development in which the tubing is vibrated longitudinally by a mechanism at the surface and acts as a high-speed pump with an extremely short stroke.

A discussion of production methods, in approximate order of relative importance, follows.

Natural flow. Natural flow is the most economical method of production and generally is utilized as long as the desired production rate can be maintained by this method (**Fig. 1**). It utilizes the formation energy, which may consist of gas in solution in the oil in the formation; free gas under pressure acting against the liquid and gas-liquid phase to force it toward the well bore; water pressure acting against the oil; or a combination of these three energy sources. In some areas the casinghead pressure

Degrees API corresponding to specific gravities of crude oil at 60°F (16°C)

	Specific gravity (hundredths)*									
Specific gravity (tenths)	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.60	104.33	100.47	96.73	93.10	89.59	86.19	82.89	79.69	79.59	73.57
0.70	70.64	67.80	65.03	62.34	59.72	57.17	54.68	52.27	49.91	47.61
0.80	45.38	43.19	44.06	38.98	36.95	34.97	33.03	31.14	29.30	27.49
0.90	25.72	23.99	22.30	20.65	19.03	17.45	15.90	14.38	12.89	11.43
1.00	10.00									

*Add to specific gravity (tenths) to read comparable value in degrees API.

may be of the order of 10,000 lb/in.^{2} (70 MPa), so it is necessary to provide fittings adequate to withstand such pressures. Adjustable throttle valves, or chokes, are utilized to regulate the flow rate to a desired and safe value. With such a high-pressure drop across a throttle valve the life of the valve is likely to be very short. Several such valves are arranged in parallel in the tubing head "Christmas tree" with positive shutoff valves between the chokes and the tubing head so that the wearing parts of the throttle valve, or the entire valve, can be replaced while flow continues through another similar valve.

An additional safeguard that is often used in connection with high-pressure flowing wells is a bottom-hole choke or a bottom-hole flow control valve that limits the rate of flow to a reasonable value, or stops it completely, in case of failure of surface controls. The packer, while not essential, is often used to reduce the free gas volume in the casing (Fig. 1).

Flow rates for United States wells seldom exceed a few hundred barrels per day because of enforced or voluntary restrictions to regulate production rates and to obtain most efficient and economical ultimate recovery. However, in some countries, especially in the Middle East, it is not uncommon for natural flow rates to exceed 10,000 bbl/(day)(well) [1590 m^3/(day)(well)].

Lifting. Most wells are not self-flowing. The common types of lifting are outlined here.

Pumping with sucker rods. Approximately 90% of the wells made to produce by some artificial lift method in the United States are equipped with sucker-rod–type pumps. In these the pump is installed at the lower end of the tubing string and is actuated by a string of sucker rods extending from the surface to the subsurface pump. The sucker rods are attached to a polished rod at the surface. The polished rod extends through a stuffing box and is attached to the pumping unit, which produces the necessary reciprocating motion to actuate the sucker rods and the subsurface pump (**Fig. 2**). The two common variations are mechanical and hydraulic long-stroke pumping.

Mechanical pumping. The great majority of pumping units are of the mechanical type, consisting of a suitable reduction gear, and crank and pitman arrangement to drive a walking beam to produce the necessary reciprocating motion. A counterbalance is provided to equalize the load on the upstroke and downstroke (**Fig. 3**). Mechanical pumping units of this type vary in load-carrying capacity from about 2000 to about 43,000 lb (900 to 19,500 kg), and the torque rating of the low-speed gear which drives the crank ranges from 6400 in.-lb (720 N · m) in the smallest API standard unit to about 1,500,000 in.-lb (170,000 N · m) for the largest units now in use. Stroke length varies from about 18 to 192 in. (46 to 488 cm). Usual operating speeds are from about 6 to 20 strokes/min. However, both lower and higher rates of speed are sometimes used.

Production rates with sucker-rod–type pumps vary from a fraction of 1 bpd in some areas, with part-time pumping, to approximately 3000 bpd (480 m^3/day) for the largest installations in relatively shallow wells.

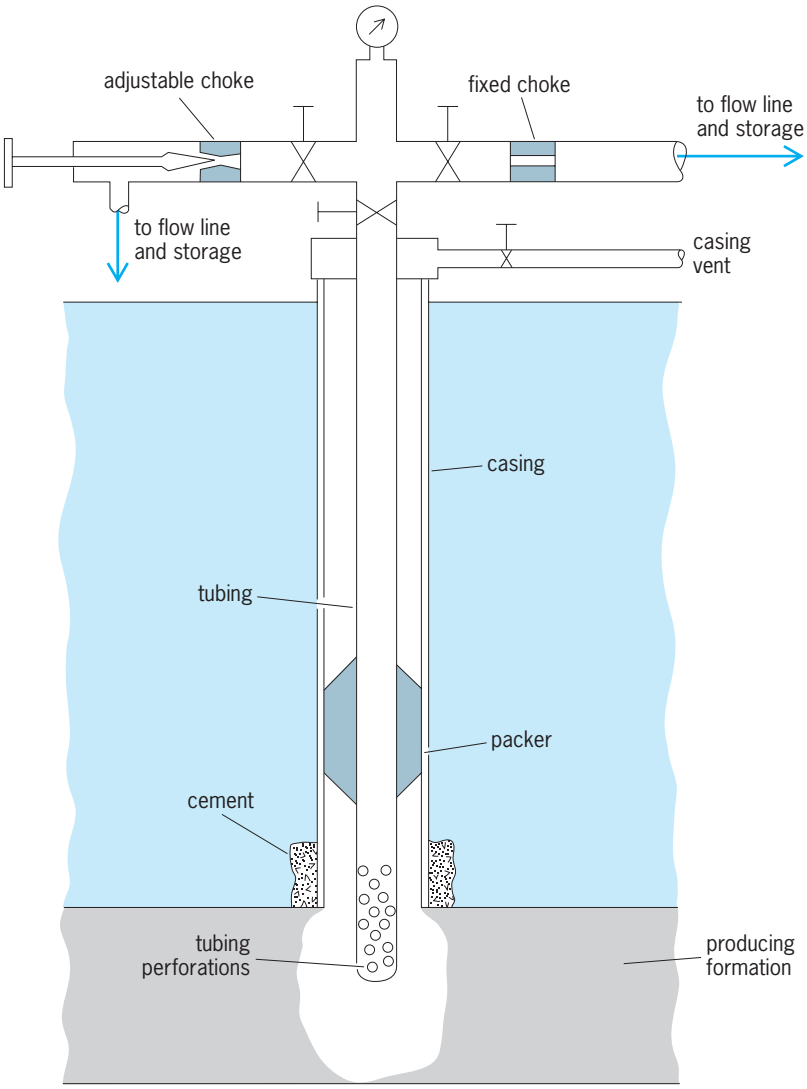

Fig. 1. Schematic view of well equipped for producing by natural flow.

Hydraulic long-stroke pumping. For this the units consist of a hydraulic lifting cylinder mounted directly over the well head and are designed to produce stroke lengths of as much as 30 ft (9 m). Such long-stroke hydraulic units are usually equipped with a pneumatic counterbalance arrangement which equalizes the power requirement on the upstroke and downstroke.

Hydraulic pumping units also are made without any provision for counterbalance. However, these units are generally limited to relatively small wells, and they are relatively inefficient.

Gas lift. Gas lift in its simplest form consists of initiating or stimulating well flow by injecting gas at some point below the fluid level in the well. With large-volume gas-lift operations the well may be produced through either the casing or the tubing. In the former case, gas is conducted through the tubing to the point of injection; in the latter, gas may be conducted to the point of injection through the casing or

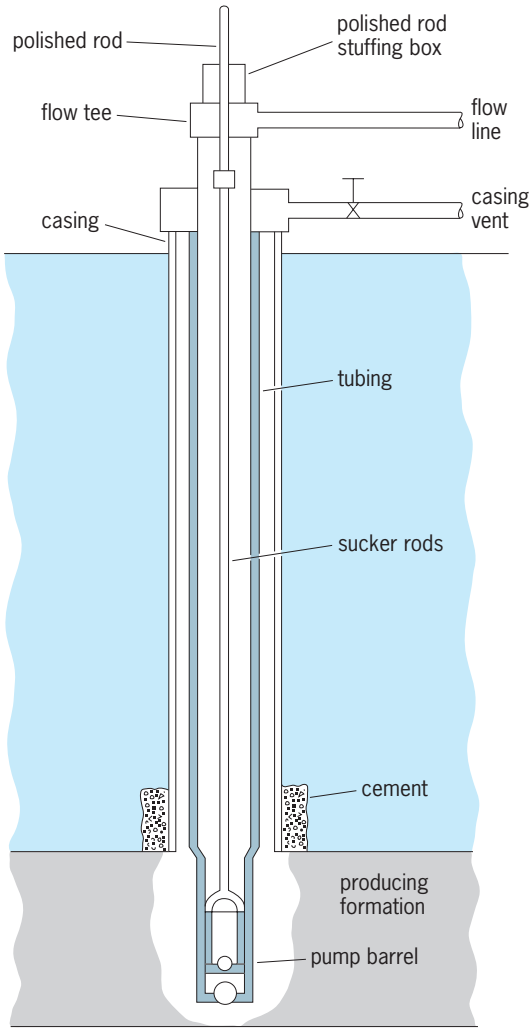

Fig. 2. Schematic view of a well which is equipped for pumping with sucker rods.

through an auxiliary string of tubing. When gas is injected into the oil column, the weight of the column above the point of injection is reduced as a result of the space occupied by the relatively low-density gas. This lightening of the fluid column is sufficient to permit the formation pressure to initiate flow up the tubing to the surface. Gas injection is often utilized to increase the flow from wells that will flow naturally but will not produce the desired amount by natural flow.

There are many factors determining the advisability of adopting gas lift as a means of production. One of the more important factors is the availability of an adequate supply of gas at suitable pressure and reasonable cost. In a majority of cases gas lift cannot be used economically to produce a reservoir to depletion because the well may be relatively productive with a low back pressure maintained on the formation but will produce very little, if anything, with the back pressure required for gas-lift operation. Therefore, it generally is necessary to resort to some mechanical means of pumping before the well is abandoned, and it may be more economical to adopt the mechanical means initially than to install the gas-lift system while conditions are favorable and later replace it.

This discussion of gas lift has dealt primarily with the simple injection of gas, which may be continuous or intermittent. There are numerous modifications of gas-lift installations, including various designs for flow valves which may be installed in the tubing string to open and admit gas to the tubing from the casing at a predetermined pressure differential between the tubing and casing. When the valve opens, gas is injected into the tubing to initiate and maintain flow until the tubing pressure drops to a predetermined value; and the valve closes before the input gas-oil ratio becomes excessive. This represents an intermittent-flow–type valve. Other types are designed to maintain continuous flow, proper pressure differential, and proper gas injection rate for efficient operation. In some cases several such flow valves are spaced up the tubing string to permit flow to be initiated from various levels as required.

Other modifications of gas lift involve the utilization of displacement chambers. These are installed on the lower end of the well tubing where oil may accumulate, and the oil is displaced up the tubing with gas injection controlled by automatic or mechanical valves.

Hydraulic subsurface pumps. The hydraulic subsurface pump has come into fairly prominent use. The subsurface pump is operated by means of a hydraulic reciprocating motor attached to the pump and installed in the well as a single unit. The hydraulic motor is driven by a supply of hydraulic fluid under pressure that is circulated down a string of tubing and through the motor. Generally the hydraulic fluid consists of crude oil which is discharged into the return line and returns to the surface along with the produced crude oil.

Hydraulically operated subsurface pumps are also arranged for separating the hydraulic power fluid from the produced well fluid. This arrangement is

Fig. 3. Pumping unit with adjustable rotary counterbalance. (*Oil Well Supply Division, U.S. Steel Corp.*)

especially desirable where the fluid being produced is corrosive or is contaminated with considerable quantities of sand or other solids that are difficult to separate to condition the fluid for use as satisfactory power oil. This method permits use of water or other nonflammable liquids as hydraulic power fluid to minimize fire hazard in case of a failure of the hydraulic power line at the surface.

Centrifugal well pumps. Electrically driven centrifugal pumps have been used to some extent, especially in large-volume wells of shallow or moderate depths. Both the pump and the motor are restricted in diameter to run down the well casing, leaving sufficient clearance for the flow of fluid around the pump housing. With the restricted diameter of the impellers the discharge head necessary for pumping a relatively deep well can be obtained only by using a large number of stages and operating at a relatively high speed. The usual rotating speed for such units is 3600 rpm, and it is not uncommon for such units to have 50 or more pump stages. The direct-connected electric motor must be provided with a suitable seal to prevent well fluid from entering the motor housing, and electrical leads must be run down the well casing to supply power to the motor.

Swabs. Swabs have been used for lifting oil almost since the beginning of the petroleum industry. They usually consist of a steel tubular body equipped with a check valve which permits oil to flow through the tube as it is lowered down the well with a wire line. The exterior of the steel body is generally fitted with flexible cup-type soft packing that will fall freely but will expand and form a seal with the tubing when pulled upward with a head of fluid above the swab. Swabs are run into the well on a wire line to a point considerably below the fluid level and then lifted back to the surface to deliver the volume of oil above the swab. They are often used for determining the productivity of a well that will not flow naturally and for assisting in cleaning paraffin from well tubing. In some cases swabs are used to stimulate wells to flow by lifting, from the upper portion of the tubing, the relatively dead oil from which most of the gas has separated.

Bailers. Bailers are used to remove fluids from wells and for cleaning out solid material. They are run into the wells on wire lines as in swabbing, but differ from swabs in that they generally are run only in the casing when there is no tubing in the well. The capacity of the bailer itself represents the volume of fluid lifting each time since the bailer does not form a seal with the casing. The bailer is simply a tubular vessel with a check valve in the bottom. This check valve generally is arranged so that it is forced open when the bailer touches bottom in order to assist in picking up solid material for cleaning out a well.

Jet pumps. A jet pump for use in oil wells operates on exactly the same principle as a water-well jet pump. Advantage is taken of the Bernoulli effect to reduce pressure by means of a high-velocity fluid jet. Thus oil is entrained from the well with this high-velocity jet in a venturi tube to accelerate the fluid and assist in lifting it to the surface, along with any assistance from the formation pressure. The application of jet pumps to oil wells has been insignificant.

Fig. 4. Lease tank battery with four tanks and two gas separators. (*Gulf Oil Corp.*)

Sonic pumps. Sonic pumps essentially consist of a string of tubing equipped with a check valve at each joint and mechanical means on the surface to vibrate the tubing string longitudinally. This creates a harmonic condition that will result in several hundred strokes per minute, with the strokes being a small fraction of 1 in. (2.5 cm) in length. Some of these pumps are being used in relatively shallow wells.

Lease tanks and gas separators. A typical lease tank battery consists of four 1000-bbl (159-m^3) tanks and two gas separators (**Fig. 4**). Such equipment is used for handling production from wells produced by natural flow, gas lift, or pumping. In some pumping wells the gas content may be too low to justify the cost of separators for saving the gas.

Natural gasoline production. An important phase of oil and gas production in many areas is the production of natural gasoline from gas taken from the casinghead of oil wells or separated from the oil and conducted to the natural gasoline plant. The plant consists of facilities for compressing and extracting the liquid components from the gas (**Fig. 5**). The natural gasoline generally is collected by cooling and

Fig. 5. Modern natural gasoline plant in western Texas. (*Gulf Oil Corp.*)

condensing the vapors after compression or by absorbing in organic liquids having high boiling points from which the volatile liquids are distilled. Many natural gasoline plants utilize a combination of condensing and absorbing techniques.

Production Problems and Instruments

To maintain production, various problems must be overcome. Numerous instruments have been developed to monitor production and to control production problems.

Corrosion. In many areas the corrosion of production equipment is a major factor in the cost of petroleum production.

For practical consideration, corrosion in oil and gas-well production can be classified into four main types.

1. Sweet corrosion occurs as a result of the presence of carbon dioxide and fatty acids. Oxygen and hydrogen sulfide are not present. This type of corrosion occurs in both gas-condensate and oil wells. It is most frequently encountered in the United States in southern Louisiana and Texas, and other scattered areas. At least 20% of all sweet oil production and 45% of condensate production are considered corrosive.
2. Sour corrosion is designated as corrosion in oil and gas wells producing even trace quantities of hydrogen sulfide. These wells may also contain oxygen, carbon dioxide, or organic acids. Sour corrosion occurs in the United States primarily throughout Arbuckle production in Kansas and in the Permian basin of western Texas and New Mexico. About 12% of all sour production is considered corrosive.
3. Oxygen corrosion occurs wherever equipment is exposed to atmospheric oxygen. It occurs most frequently in offshore installations, brine-handling and injection systems, and in shallow producing wells where air is allowed to enter the casing.
4. Electrochemical corrosion is designated as that which occurs when corrosion currents can be readily measured or when corrosion can be mitigated by the application of current, as in soil corrosion.

Fig. 6. Two pumping wells with tank battery. (*Oil Well Supply Division, U.S. Steel Corp.*)

Corrosion inhibitors are used extensively in both oil and gas wells to reduce corrosion damage to subsurface equipment. Most of the inhibitors used in the oil field are of the so-called polar organic type. All of the major inhibitor suppliers can furnish effective inhibitors for the prevention of sweet corrosion as encountered in most fields. These can be purchased in oil-soluble, water-dispersible, or water-soluble form. *See* CORROSION.

Paraffin deposits. In many crude-oil–producing areas paraffin deposits in tubing and flow lines and on sucker rods are a source of considerable trouble and expense. Such deposits build up until the tubing or flow line is partially or completely plugged. It is necessary to remove these deposits to maintain production rates. A variety of methods are used to remove paraffin from the tubing, including the application of heated oil through tubular sucker rods to mix with and transfer heat to the oil being produced and raise the temperature to a point at which the deposited paraffin will be dissolved or melted. Paraffin solvents may also be applied in this manner without the necessity of applying heat.

Mechanical means often are used in which a scraping tool is run on a wire line and paraffin is scraped from the tubing wall as the tool is pulled back to the surface. Mechanical scrapers that attach to sucker rods also are in use. Various types of automatic scrapers have been used in connection with flowing wells. These consist of a form of piston that will drop freely to the bottom when flow is stopped but will rise back to the surface when flow is resumed. Electrical heating methods have been used rather extensively in some areas. The tubing is insulated from the casing and from the flow line, and electric current is transmitted through the tubing for the time necessary to heat the tubing sufficiently to cause the paraffin deposits to melt or go into solution in the oil in the tubing. Plastic coatings have been utilized inside tubing and flow lines to minimize or prevent paraffin deposits. Paraffin does not deposit readily on certain plastic coatings.

A common method for removing paraffin from flow lines is to disconnect the line at the wellhead and at the tank battery and force live steam through the line to melt the paraffin deposits and flow them out. Various designs of flow-line scrapers have also been used rather extensively and fairly successfully. Paraffin deposits in flow lines are minimized by insulating the lines or by burying the lines to maintain a higher average temperature.

Emulsions. A large percentage of oil wells produce various quantities of salt water along with the oil, and numerous wells are being pumped in which the salt-water production is 90% or more of the total fluid lifted. Turbulence resulting from production methods results in the formation of emulsions of water in oil or oil in water; the commoner type is oil in water. Emulsions are treated with a variety of demulsifying chemicals, with the application of heat, and with a combination of these two treatments. Another method for breaking emulsions is the electrostatic or electrical precipitator type of emulsion treatment.

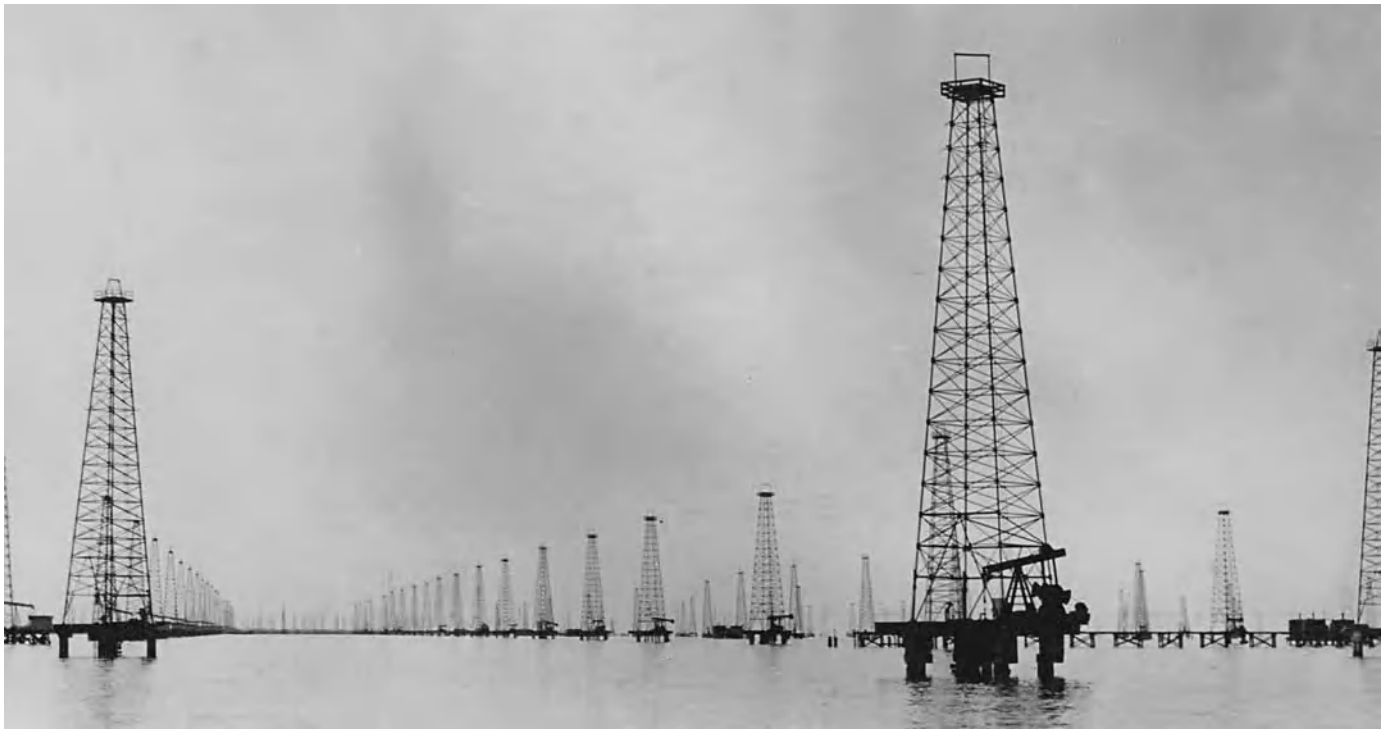

Fig. 7. Numerous offshore wells located in Lake Maracaibo, Venezuela. (*Creole Petroleum Corp.*)

In this method the emulsion to be broken is circulated between electrodes subjected to a high potential difference. The resulting concentrated electric field tends to rupture the oil-water interface and thus breaks the emulsion and permits the water to settle out. A typical tank battery is equipped with a wash tank, or gun barrel, and a gas-fired heater for emulsion treating and water separation before the oil is admitted to the lease tanks (**Fig. 6**). *See* OIL AND GAS STORAGE.

Gas conservation. If the quantity of gas produced with crude oil is appreciably greater than that which can be efficiently utilized or marketed, it is necessary to provide facilities for returning the excess gas to the producing formation. Formerly, large quantities of excess gas were disposed of by burning or simply by venting to the atmosphere. This practice is now unlawful. Returning excess gas to the formation not only conserves the gas for future use but also results in greater ultimate recovery of oil from the formation.

Salt-water disposal. The large volumes of salt water produced with the oil in some areas present serious disposal problems. The salt water is generally pumped back to the formation through wells drilled for this purpose. Such salt-water disposal wells are located in areas where the formation already contains water. Thus this practice helps to maintain the formation pressure as well as the productivity of the producing wells.

Offshore production. Offshore wells present additional production problems since the wells must be serviced from barges or boats. Wells of reasonable depth on land locations are seldom equipped with derricks for servicing because it is more economical to set up a portable mast for pulling and installing rods, tubing, and other equipment. However, the use of portable masts is not practical on offshore locations, and a derrick is generally left standing over such wells throughout their productive life to facilitate servicing. There are a considerable number of offshore wells along the Gulf Coast and the Pacific Coast of the United States, but by far the greatest number of offshore wells in a particular region is in Lake Maracaibo in Venezuela. **Figure 7** shows a considerable number of derricks in Lake Maracaibo with pumping wells in the foreground. These wells are pumped by electric power through cables laid on the lake bottom to conduct electricity from power-generating stations onshore. An overwater tank battery is visible at the extreme right. All offshore installations, such as tank batteries, pump stations, and the derricks and pumping equipment, are supported on pilings in water up to 100 ft (30 m) or more in depth. There are approximately 2300 oil derricks in Lake Maracaibo. Semipermanent platform rigs and even bottom storage facilities are used in Gulf of Mexico waters at depths of more than 100 ft (30 m). *See* OIL AND GAS, OFFSHORE.

Instruments. The commoner and more important instruments required in petroleum production operations are included in the following discussion.

Gas meters, which are generally of the orifice type, are designed to record the differential pressure across the orifice, and the static pressure.

Recording subsurface pressure gages small enough to run down 2-in. (5-cm) ID (inside diameter) tubing are used extensively for measuring pressure gradients down the tubing of flowing wells, recording pressure buildup when the well is

closed in, and measuring equilibrium bottom-hole pressures.

Subsurface samplers designed to sample well fluids at various levels in the tubing are used to determine physical properties, such as viscosity, gas content, free gas, and dissolved gas at various levels. These instruments may also include a recording thermometer or a maximum reading thermometer, depending upon the information required.

Oil meters of various types are utilized to meter crude oil flowing to or from storage.

Dynamometers are used to measure polished-rod loads. These instruments are sometimes known as well weighers since they are used to record the polished-rod load throughout a pumping cycle of a sucker-rod–type pump. They are used to determine maximum load on polished rods as well as load variations, to permit accurate counterbalancing of pumping wells, and to assure that pumping units or sucker-rod strings are not seriously overloaded.

Liquid-level gages and controllers are used. They are similar to those used in other industries, but with special designs for closed lease tanks.

A wide variety of scientific instruments find application in petroleum production problems. The above outline gives an indication of a few specialized instruments used in this branch of the industry, and there are many more. Special instruments developed by service companies are valued for a wide variety of purposes and include calipers to detect and measure corrosion pits inside tubing and casing and magnetic instruments to detect microscopic cracks in sucker rods. *See* PETROLEUM RESERVOIR ENGINEERING.

Roy L. Chenault

Bibliography. R. Baker, *A Primer of Offshore Operations*, 3d ed., 1998; W. D. Berger and K. E. Anderson, *Modern Petroleum: A Basic Primer of the Industry*, 3d ed., 1992; R. A. Dawe and A. G. Lucas (eds.), *Modern Petroleum Technology, vols. 1 and 2*, 6th ed., 2000; M. J. Economides, A. D. Hill, and C. Ehlig-Economides, *Petroleum Production Systems*, 1993; N. J. Hyne, *Nontechnical Guide to Petroleum Geology, Exploration, Drilling and Production*, 2d ed., 2001; M. A. Mian, *Petroleum Engineering Handbook for the Practicing Engineer*, vol. 1, 1992; T. E. W. Nind, *Hydrocarbon Reservoir and Well Performance*, 1989.

Oil and gas storage

Storage, usually in great quantities, of crude oil and natural gas after production from natural reservoirs. Large amounts of refined products are stored as well. Storage is necessary to meet seasonal and other fluctuations in demand; for efficient operation of producing equipment, pipelines, tankers, and refineries; and for emergency use.

Crude oil and refined products. Oil from producing wells is first collected in welded-steel, bolted-steel, or wooden tanks or 100 bbl (16 m^3) or greater capacity located on individuals leases. These tanks, upright cylinders with low-pitched conical roofs, provide temporary storage while the oil is awaiting shipment. Several tanks grouped together are a tank battery. Assemblages of large steel tanks, known as tank farms, are used for more permanent storage at pipeline pump stations, points where tankers load and unload, and refineries.

The trend toward giant tankers, accelerated by the closing of the Suez Canal in 1967, created a need for large storage facilities at both the loading and unloading ends of the tanker runs. Large-capacity excavated reservoirs with concrete linings have been used for many years in California to store both crude and fuel oil. One such reservoir with a fixed roof and elliptical in form is 780 ft (238 m) long, 467 ft (142 m) wide, and 23 ft (7.9 m) deep. It covers $9^1/_4$ acres (3.74 ha) and provides storage for more than 1×10^6 bbl (1.6×10^5 m^3). Another reservoir has a capacity of 4×10^6 bbl (6×10^5 m^3) and covers 16 acres (6.5 ha).

Offshore storage. For offshore producing fields a number of unique storage systems have been designed. In several instances old tankers have been adapted for storage, and barges have been constructed especially for offshore storage use. One underwater installation consists essentially of three giant inverted steel funnels. Each unit is 270 ft (82 m) in diameter and 205 ft (62 m) high, weighs 28×10^6 lb (13×10^6 kg), and has a capacity of 0.5×10^6 bbl (8×10^4 m^3). The bottom is open, and the unit is anchored to the sea floor by 95-ft (29-m) pilings. A reinforced concrete installation features a nine-module storage unit with 1×10^6 bbl capacity surrounded by a perforated wall 302 ft (92 m) in diameter and serves as a breakwater. The outer wall is about 270 ft (82 m) high and extends about 40 ft (12 m) above the water surface. A submerged floating storage tank 96 ft (29 m) in diameter and 305 ft (93 m) high is held in place by six anchor lines and has a capacity of 300,000 bbl (50,000 m^3). One relatively small unit consists of a platform with four vertical legs, each holding 4700 bbl (750 m^3), and four horizontal tanks at the bottom holding 1850 bbl (294 m^3) each, for a total capacity of 26,200 bbl (4170 m^3). A second small unit, utilizing bottom tanks as an anchor, holds 2400 bbl (380 m^3) underwater and 600 bbl (90 m^3) in a spherical tank above the surface. A third small unit consists of a sea-floor base, connected by a universal joint to a large-diameter vertical cylinder, about 350 ft (110 m) high, which extends above the water surface. One design includes an excavated cavern beneath the sea floor; nuclear cavities have also been suggested.

Volatility problems. To minimize vaporization losses, lease tanks are sometimes equipped to hold several ounces pressure. At large-capacity storage sites, special tanks are generally used. Tanks with lifter or floating roofs are used to store crude oil, motor gasoline, and less volatile natural gasoline. Motor and natural gasolines are also stored in spheroid containers. Spherical containers are used for more volatile liquids, such as butane. Horizontal cylindrical containers are used for propane and butane storage.

Refrigerated insulated tank systems enabling propane to be stored at a lower pressure are also in use.

Underground storage. Large quantities of volatile liquid-petroleum products, including propane and butane, are stored in underground caverns dissolved in salt formations and in mined caverns, gas reservoirs, and water sands. Liquid-petroleum products are also stored underground in Belgium, France, Germany, and France. In Pennsylvania an abandoned quarry with a capacity of 2×10^6 bbl (0.3×10^6 m^3) was equipped with a floating roof for storing fuel oil. Refrigerated propane is also being stored in excavations in frozen earth and in underground concrete tanks.

Natural gas. Natural gas is stored in low-pressure surface holders, buried high-pressure pipe batteries and bottles, depleted or partially depleted oil and gas reservoirs, water sands, and several types of containers at extremely low temperature (−258°F or −161°C) after liquefaction.

Low-pressure holders, which store relatively small volumes of gas, basically use either a water or a dry seal, and variations of each type exist. With the displacement of manufactured gas by natural gas in the United States, the need for surface holders greatly diminished and they have disappeared almost entirely.

Underground storage. In the United States gas pipeline and utility companies store large quantities of natural gas in underground reservoirs. In most cases these reservoirs are located near market areas and are used to supplement pipeline supplies during the winter months when the gas demand for residential heating is very high. Since gas can be stored in the summer when the gas demand is low, underground storage permits greatly increased pipeline utilization, resulting in lower transportation costs and reduced gas cost to the consumer. Underground storage is the only economical method of storing large enough quantities of gas to meet the seasonal fluctuations in pipeline loads, and has enabled gas companies to meet market requirements which otherwise could not be satisfied.

Gas was first stored underground in 1915 in a partially depleted gas field in Ontario, Canada. The following year gas was injected into a depleted gas field near Buffalo, New York. Gas reservoirs, which utilize depleted gas and oil fields, water sands, salt caverns, and an abandoned coal mine, are located in 26 states. Gas is also being stored in underground reservoirs in Canada, France, Germany, Austria, Italy, Poland, Romania, Czechoslovakia, and Russia. In the United Kingdom a salt cavern is being used for gas storage.

Gas storage in water sands was first undertaken in 1952, and this method of storage steadily increased, especially in areas lacking gas or oil fields. A cross section of a typical aquifer storage field is shown in the **illustration**. A geologic trap having adequate structural closure and a suitable caprock is needed. The storage sand must be porous and thick enough and under sufficient hydrostatic pressure to hold large quantities of gas. The sand must also be sufficiently permeable and continuous over a wide enough area

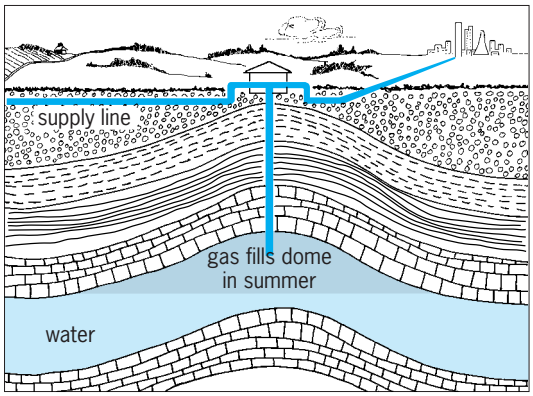

Schematic cross section of typical aquifer gas storage field showing injected gas displacing water. (*Natural Gas Pipeline Company of America*)

so that water can be pushed back readily to make room for the stored gas. In some cases, water removal wells are also utilized.

Reservoir pressure. In operating storage reservoirs only a portion of the stored gas, called working gas, is normally withdrawn. The remaining gas, called cushion gas, stays in the reservoir to provide the necessary pressure to produce the storage wells at desired rates. In aquifer storages some water returns to help maintain the reservoir pressure. The percentage of cushion gas varies considerably among reservoirs. In some instances, the original reservoir pressure of the oil and gas field is exceeded in storage operations. This results in storage volumes greater than the original content and increases well deliverabilities. In aquifer storages the original hydrostatic pressure must be exceeded in order to push the water back.

In Russia, aquifer gas storage has been undertaken in one area where no appreciable structural closure exists. A company in the United States also has experimented with storing without structure. In an inconclusive field test, air was injected into a center well with control of the lateral spread of the air bubble attempted by injecting water into surrounding wells. Storage of gas in cavities created by nuclear explosions has been proposed and seriously considered.

Liquefied gas. Storage of liquefied natural gas throughout the world is in connection with shipment of liquefied natural gas by tanker, and is located at the loading and unloading ends of the tanker runs as well as at peak sharing facilities operated by gas pipeline and local utility companies. In addition to the United States and Canada, England, France, the Netherlands. Germany, Italy, Spain, Algeria, Libya, and Japan also have installations. Storage is in insulated metal tanks, buried concrete tanks, or frozen earth excavations. In two projects using frozen earth excavations, excessive boil-off of the liquefied gas led to replacement with insulated metal tanks. *See* LIQUEFIED NATURAL GAS (LNG); OIL AND GAS FIELD EXPLOITATION; PETROLEUM; PIPELINE. Peter G. Burnett

Bibliography. American Gas Association, *The Underground Storage of Gas in the United States and*

Canada, Annual Report on Statistics; D. C. Bond, *Underground Storage of Natural Gas*, Illinois State Geological Survey, 1975; G. Hobson (ed.), *Modern Petroleum Technology*, pts. 1 and 2, 5th ed., 1984; D. L. Katz et al., *Handbook of Natural Gas Engineering*, 1959; D. L. Katz and P. A. Witherspoon, *Underground and Other Storage of Oil Products and Gas*, Proceedings of the 8th World Petroleum Congress, 1971; E. V. Tiratsoo, *Natural Gas*, 3d ed., 1980; Stone and Webster Engineering, *Gas Storage at the Point of Use*, Amer. Gas Ass. Proj. PL-56, 1965.

Oil and gas well completion

The operations that prepare a well bore for producing oil or gas from the reservoir. The goal of these operations is to optimize the flow of the reservoir fluids into the well bore, up through the producing string, and into the surface collection system. *See* OIL AND GAS FIELD EXPLOITATION; OIL AND GAS WELL DRILLING.

Casing and cement. The well bore is lined (cased) with steel pipe, and the annulus between well bore and casing is filled with cement.

Properly designed and cemented casing prevents collapse of the well bore and protects fresh-water aquifers above the oil and gas reservoirs from becoming contaminated with oil and gas and the oil reservoir brine. Similarly, the oil and gas reservoir is prevented from becoming invaded by extraneous water from aquifers that were penetrated above or below the productive reservoir. *See* AQUIFER.

The casing string is made up of joints of steel pipe which are screwed together to form a continuous string as the casing is extended into the well bore. The common length of an individual joint is 30 ft (9 m). Such factors as the depth of the well, the pressure, temperature, and corrosivity of the fluids to be produced and those in the reservoirs that are to be cased off (behind pipe) are taken into account in specifying the diameter, wall thickness, strength, and chemical composition of the steel pipe for a particular casing job.

In deep wells, one or more intermediate strings of casing are set (**Fig. 1**) in order to cement off either high-pressure intervals which cannot be controlled by the weight of the drilling fluid, or low-pressure intervals into which large volumes of drilling mud may flow and result in lost circulation, preventing further controlled drilling. When drilling into a high-pressure formation, casing is frequently set on top of it in order to facilitate well control operations if a blowout appears to be imminent.

In order to achieve its objectives, the casing must be securely sealed to the well bore itself with cement, although special formulations may be required for specific wells. For example, high-temperature formations or producing formations which will be extensively fractured will require cement that will not set too rapidly at high temperature or will not crack too badly as a result of the pressure shock of hydraulic fracturing, respectively. The cement is pumped down the casing and then on up into the annulus to a predetermined height. Cement returns (to the surface) are not universally required. The cement is mixed, pumped, and metered with highly specialized mobile equipment which is supplied by an appropriate service company. In order not to end up with the casing filled with cement, a specially designed plug is inserted after the required amount of cement has been pumped in and displaced with water until the plug hits the bottom of the casing string. The plug and some minor amount of cement will have to be drilled out after the cement has set.

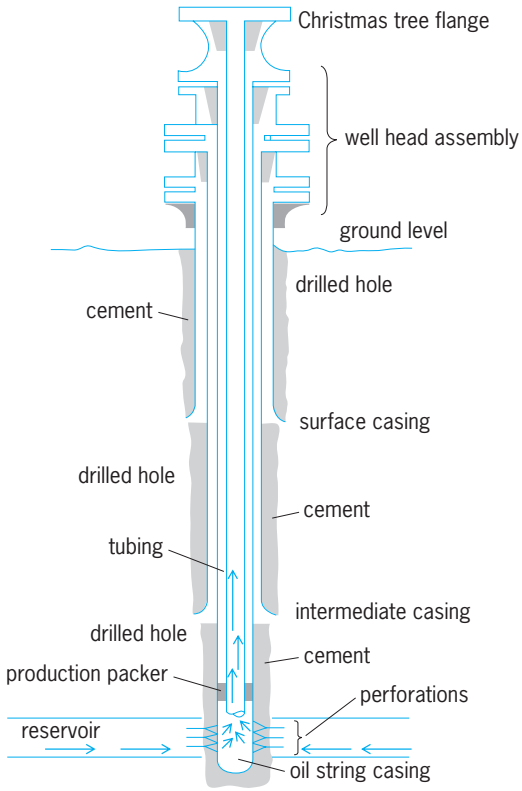

Fig. 1. Casing detail; casing strings in an oil well.

Well bore–reservoir connection. The nature of the reservoir, evaluated from a core analysis, cuttings, or logs, or from experience with like productive formations, determines the type of completion to be used: barefoot, casing set through and then perforated, or a shop perforated or slotted liner.

In a barefoot completion, the casing is set just above the producing formation, and the latter is drilled out and produced with no pipe set across it (**Fig. 2**). Such a completion can be used for hard rock formations which are not friable and will not slough, and when there are no opportunities for producing from another, lower reservoir.

Set-through and perforated completions are also employed for relatively well-consolidated formations from which the potential for sand production is small. However, the perforated completion is used

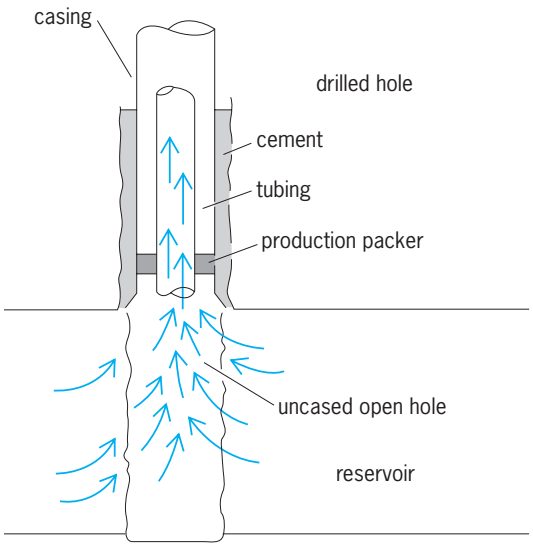

Fig. 2. Diagram of barefoot completion.

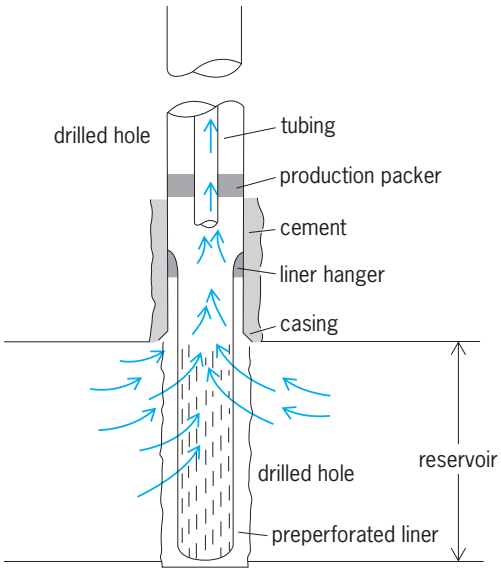

Fig. 3. Liner-type completion; preperforated liner.

when a long producing interval must be prevented from collapse, when multiple intervals are to be completed in the one borehole, or when intervening water sands within the oil-producing interval are to be shut off and the oil-saturated intervals selectively perforated. Perforations are made with bullets or shaped charges (jet perforation). The bullets are fired from a gun with multiple barrels, spaced at desired intervals, which is lowered into the hole on a wire line. An electric impulse detonates the bullets. The holes created by bullets are frequently lined with fused metal and mineral debris and as a result may offer some resistance to fluid influx. *See* WELL LOGGING.

The charges used in jet perforating are similar to the shaped charges used in bazookas. The shaped charges are run into the hole on a glass gun which disintegrates.

A shop-fabricated liner is used for friable formations from which some of the formation sand may flow into the well bore. The passage of such sand into the well bore may cause scoring of the seats and valves in the pump and its consequent failure to be able to lift produced fluid; or it may result in accumulation of a sand plug in the lower joints of casing through which the flow of fluids would be impeded, or in erosion of surface valves and piping. The holes in the liner are designed to screen out any produced sand, and such liners are gravel-packed. A slurry of gravel is circulated (washed-in) down behind the liner, prior to setting the liner hanger (**Fig. 3**). The distribution of particle diameters in the gravel pack is chosen so that the pack is an effective screen for the reservoir sand. A prepacked gravel liner may also be used, but since a gap is left between the well bore and the liner, which may fill up with the fine silt that is carried with the produced fluids, it is generally preferable to use a washed-in gravel pack.

Production. A string of steel tubing is lowered into the casing string and serves as the conduit for the produced fluids. The tubing may be hung from the well-head or supported by a packer set above the producing zone. The packer is used when it is desirable to isolate the casing string from the produced fluids because of the latter's pressure, temperature, or corrosivity, or when such isolation may improve production characteristics.

Artificial lift. The tops of wells from which fluids flow as a result of the indigenous reservoir energy are equipped with a manifold known as the Christmas tree (**Fig. 4**). However, only some reservoirs have sufficient pressure and sufficient gas in solution (which is released at the lower pressure existing in the well bore and therefore lowers the effective density of the

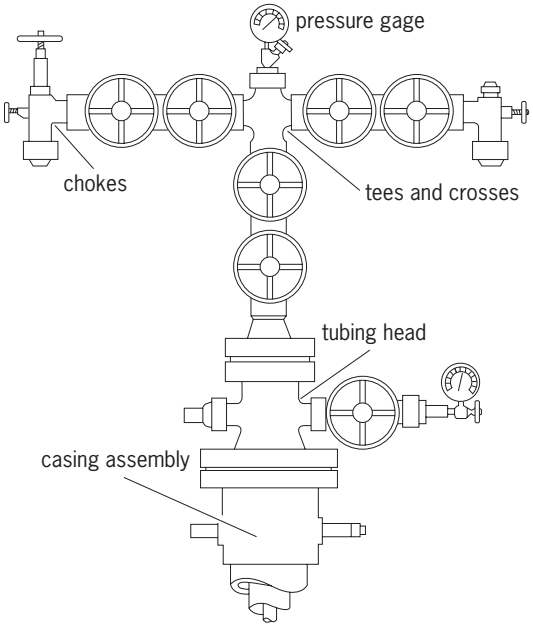

Fig. 4. Typical layout of a Christmas tree manifold.

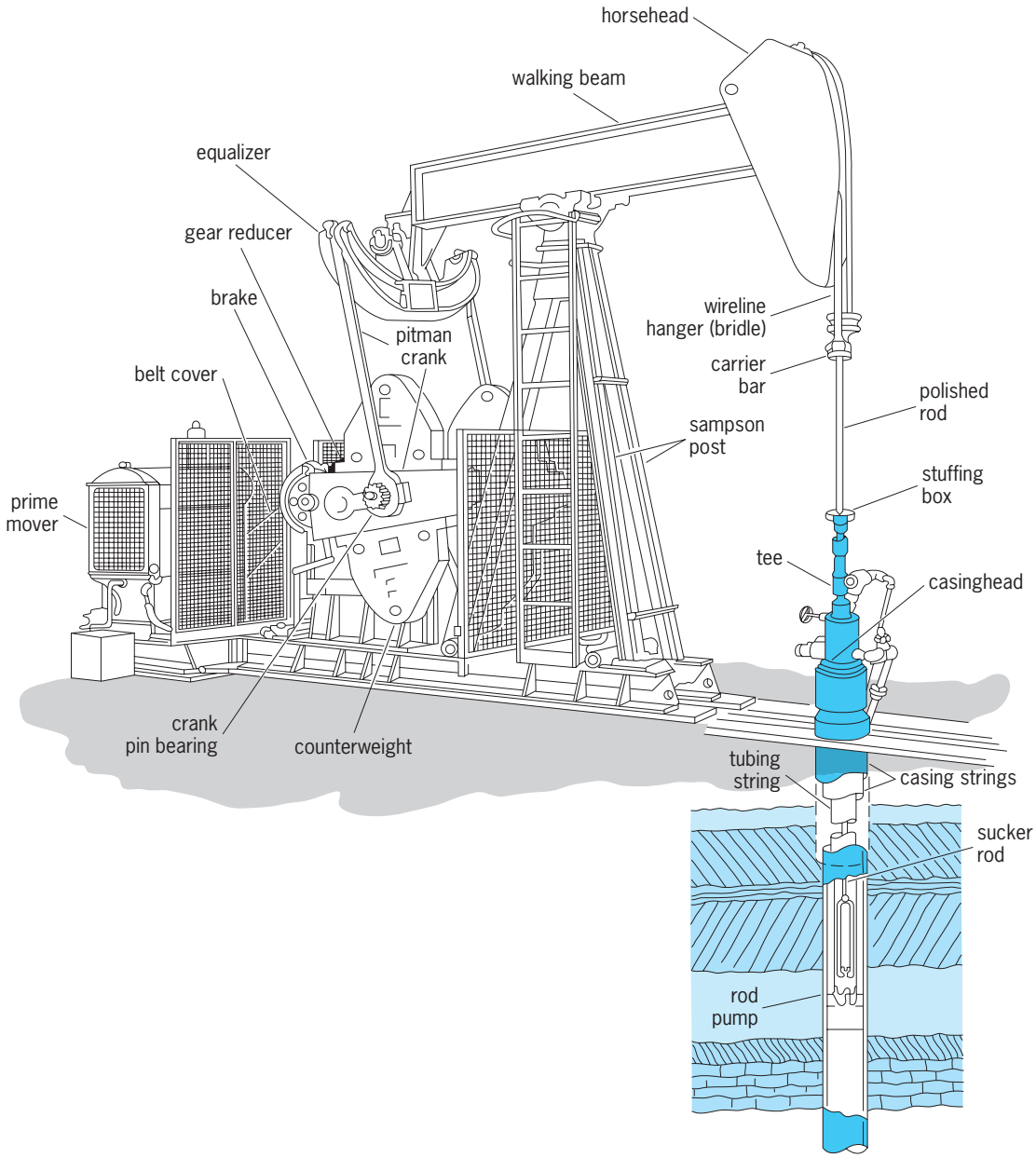

Fig. 5. Schematic diagram of most commonly used downhole pumps.

fluid in the tubing) to permit natural flow to the surface. The reservoir fluids from other reservoirs and, after pressure depletion, even from those which initially flowed must be brought to the surface by one of several methods of artificial lift.

The most common method is the use of a rod pump which is set near the bottom of the hole and operated by reciprocating sucker rods which are in turn attached to the walking beam on the surface (**Fig. 5**). The walking beam is driven by a motor, and by the use of suitable cams and cranks the beam's seesaw movement raises and lowers the sucker rod string. The cycle and stroke length of the sucker rods are adjustable. Tubing pumps attached to the bottom of the tubing string have a relatively high capacity, but the entire string must be pulled to repair a damaged pump. Insert pumps are set within the tubing. Because of their restricted diameter, they have a limited capacity for lifting reservoir fluids, but they have the advantage that they can be pulled and replaced with a wire line without pulling the entire tubing string. For deeper wells, hydraulic motors can be used for which the actuating fluid (crude oil) is pumped down the tubing and returns with the produced fluid up through the annulus. Wells which produce a large amount of fluid (both water and oil) can economically use a submerged centrifugal pump driven by an electric motor for which an electric cable is run down the annulus. *See* CENTRIFUGAL PUMP.

For deep wells which produce a significant amount of gas, gas lift can be employed in which

some of the produced gas is compressed and returned to the casing-tubing annulus. A series of pressure-actuated valves inserted in the tubing string permits the gas to enter the string at various levels to lower the effective density of the fluids in the tubing and propel the fluids to the surface (**Fig. 6**). A plunger lift system to assist with unloading liquids can be easily installed inside tubing without the need to pull the tubing. Such systems are used to produce high-gas-oil-ratio (GOR) wells, water-producing gas wells, or very-low-bottom-hole-pressure oil wells (used with gas lift).

Multiple completion. In some geological provinces, several successive but separated intervals are productive of oil and gas. In some instances, the production from all the intervals may be commingled in a single well bore. However, if the properties of the reservoirs or their fluids are different, then commingling may be unacceptable because of the potential for cross flow between the individual reservoirs. Multiple completions in which the producing zones are separated by the use of packers and individual tubing string are then used (**Fig. 7**).

Water problems. Excessive water production increases the cost of oil production since energy must be expended in lifting the water to the surface. Water production may also jeopardize the production of oil and gas by saturating the oil-productive interval with water. Such damage is more likely to occur in low-pressure formations or formations which contain water-sensitive clays that swell in an excess of water.

Water-exclusion methods. Water exclusion may be effected by the application of cements of various types. If it is determined that water is entering from the lower portion of a producing sand in a relatively shallow, low-pressure well, a cement plug may be placed in the bottom of the hole so that it will cover the oil-water interface of the reservoir. This technique is called laying in a plug and may be accomplished by placing the cement with a dump-bottom bailer on a wire line or by pumping cement down the drill pipe or tubing. For deeper, higher-pressure, or more troublesome wells, a squeeze method is used. Squeeze cementing is the process of applying hydraulic pressure to force a cement into an exposed formation or through openings in the casing or liner. It is also used for repairing casing leaks; isolating producing zones prior to perforating for production; remedial or secondary cementing to correct a defective condition, such as channeling or insufficient cement on a primary cement job; sealing off a low-pressure formation that causes lost circulation of drilling fluids; and abandoning depleted producing zones to prevent migration of formation effluent and to reduce possibilities of contaminating other zones or wells.

The squeeze tool is a packer-type device designed to isolate the point of entry between or below packing elements. The tool is run into the hole on drill pipe or tubing, and the cement is squeezed out between or below these confining elements into the problem area. The well is then recompleted. It may be necessary to drill the cement out of the hole and reperforate, depending upon the outcome of the job performed in the squeeze process.

Water-exclusion plug back. Simple water shutoff jobs in shallow, deep, or high-pressure wells may also be performed in multizone wells in which the lower

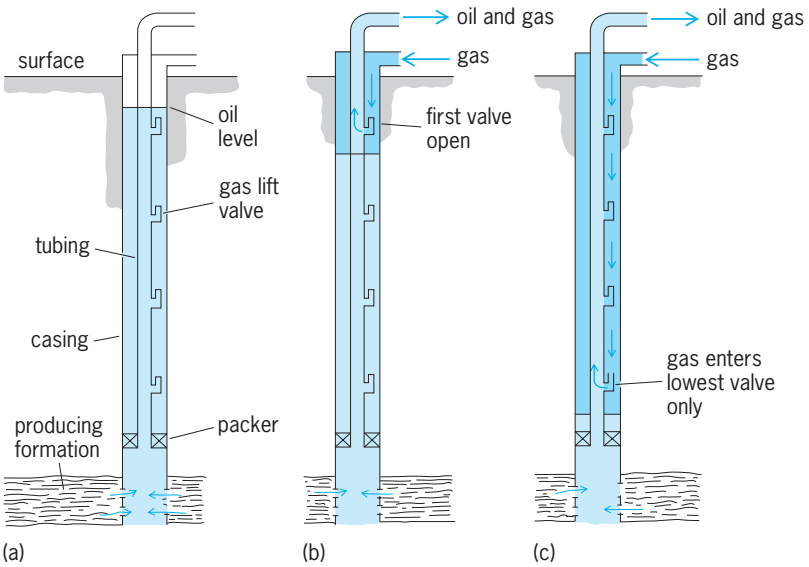

Fig. 6. Schematic representation of operation of a gas lift string. (*a*) Oil level above first valve. (*b*) First valve open and gas entering tubing; oil level in casing/tubing annulus moving downward. (*c*) Oil level has moved downward, and each valve has closed as gas has entered the next lowest valve.

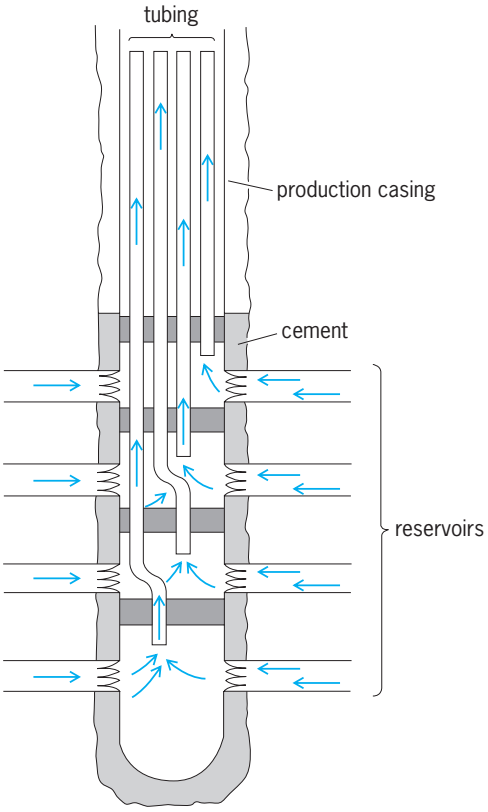

Fig. 7. Schematic diagram of a multiple completion.

producing interval is depleted or the remaining recoverable reserves do not justify recompletion.

Here, water may be excluded by placing a packer-type plug above the interval, then producing formations that are already open or perforating additional intervals that may be present higher up the hole.

Production stimulation. Production may be impaired from a well bore as a result of drilling-mud invasion or of accumulation of clays and fine silts carried by the producing fluids to the borehole, or the lithology of the formation itself may have a naturally low permeability to reservoir fluids. Since the permeability to fluids of the formation within the first few feet of the well bore has an exponential effect on limiting the influx of fluid, the productivity of a well can frequently be increased manyfold by increasing the permeability of this element of the reservoir or removing the skin just at the face of the producing interval. This is accomplished by acidization and fracturing, and in some instances by the use of surfactants, solvents, and explosives. Specialized service companies conduct the work by using their own specially designed equipment.

Acidizing. Inhibited hydrochloric acid contains a chemical additive (an inhibitor) which prevents the acid from attacking steel. In this way the acid can be used for dissolving carbonates, oxides, and other compounds without fear of it attacking the well's steel tubulars. This formulation is used to dissolve limestone and dolomitic matrices and thus enlarge the flow channels in production-impaired reservoirs. Hydrochloric acid is also used to shrink and disperse sheaths of drilling mud on the well bore and to dissolve calcareous cements, which results in larger channels through which fluids can flow to the well bore.

Hydrofluoric acid (released by injecting a mixture of hydrochloric acid and a soluble fluoride salt) is sometimes used in sandstone reservoirs to dissolve and disperse drilling mud that has invaded the reservoir.

Fracturing. Formation fracturing is a hydraulic process aimed at the parting of the formation. Vertical fractures most frequently occur. Horizontal fracturing occurs only in relatively shallow formations, in formations where the major tectonic stress is horizontal, or in relatively plastic formations. The fracturing fluid is injected into the well, and the pressure is raised to maintain a given flow rate until formation breakdown occurs. Injection is continued with a slurry of a selected grade of sand or gravel or particles of other material (such as sintered bauxite or ceramic beads). These particles prop the fracture open after the hydraulic pressure is released. Crude oil, acid, and a variety of gelled liquids are used as fracturing fluids. The propping material guarantees that there will be a high-permeability path into the well bore, and the nature of fluid flow in the vicinity of the well bore is changed to being predominantly linear rather than radial with an associated decrease in pressure drop (or higher flow rate at the same pressure drop).

Other stimulation techniques. Explosives were the first means used to stimulate oil and gas production. However, this technique has largely been supplanted by more effective and safe fracturing and acidizing technology.

Solvents are used when the substances believed to be inhibiting production are asphaltenes, waxes, and emulsions stabilized by such organic materials. Surfactants are frequently used with the solvents to aid in the dispersion of the sediments. Surfactants or alcohols (for example, methanol) are also used alone when the cause of impairment is believed to be a high saturation of water that has accumulated in the reservoir near the well bore.

Sand consolidation. Sand exclusion techniques using liners and gravel packs are not perfect, and therefore technology has been developed that attempts to consolidate friable formations. The consolidating medium must be capable of cementing the grains together without significantly reducing the permeability of the reservoir to fluid flow. Epoxy and phenolic resins have been developed for such purposes; some techniques use thermally deposited nickel metal and precipitated aluminum oxides. However, liners with properly designed gravel packs continue to be the most economical and useful technique for sand control.

Coiled tubing. Many of the well completion or workover techniques can be implemented with a coiled tubing unit that can greatly reduce costs. Instead of moving in a completion rig to lower or pull tubing, a coiled tubing unit may be moved next to the wellhead. With such a unit, instead of having to connect and disconnect stands of tubing, a continuous length of tubing may be uncoiled or coiled into the borehole by using a large spool. In this way, numerous operations may be performed on a well such as acidizing, setting and retrieving bridge plugs or packers, cementing, cleaning out the hole, and even light-duty or slim-hole drilling. A wide variety of remedial operations may be performed. *See* PETROLEUM RESERVOIR ENGINEERING.

Todd M. Doscher; R. E. Wyman

Bibliography. J. Algeroy, Equipment and operation of advanced completions in the M-15 Wytch Farm mulitlateral well, presented at the *2000 Anuual Technical Conference and Exhibition* (Dallas), Pap. SPE 62951, October 1–4, 2000; G. Botto et al., Innovative remote controlled completion for Aquila Deepwater Challenge, *1996 SPE European Petroleum Conference* (Milan), Pap. SPE 36948, October 22–24, 1996; R. A. Dawe and A. G. Lucas (eds.), *Modern Petroleum Technology, vols. 1 and 2*, 6th ed., 2000; M. J. Economides, A. D. Hill, and C. Ehlig-Economides, *Petroleum Production Systems*, 1993; N. J. Hyne, *Nontechnical Guide to Petroleum Geology, Exploration, Drilling and Production*, 2d ed., 2001; V. B. Jackson, Intelligent completion technology improves economics in the Gulf of Mexico, *Amer. Oil Gas Rep.*, June 2000; D. E. Johnson, Reliable and completely interventionless intelligen completion technology: Application and field study, *2002 Offshore Technology Conference* (Houston), Pap. OTC 14252, May 6–9, 2002.

Oil and gas well drilling

The drilling of holes for exploration and extraction of crude oil and natural gas. Deep holes and high pressures are characteristics of petroleum drilling not commonly associated with other types of drilling. In general, it becomes more difficult to control the direction of the drilled hole as the depth increases, and additionally, the cost per foot of hole drilled increases rapidly with the depth of the hole. Drilling-fluid pressure must be sufficiently high to prevent blowouts but not high enough to cause fracturing of the borehole. Formation-fluid pressures are commonly controlled by the use of a high-density clay-water slurry, called drilling mud. The chemicals used in drilling mud can be expensive, but the primary disadvantage in the use of drilling muds is the relatively low drilling rate which normally accompanies high bottom-hole pressure. Drilling rates can often be increased by using water to circulate the cuttings from the hole; when feasible, the use of gas as a drilling fluid can lead to drilling rates as much as 10 times those attained with mud. Drilling research has the objectives of improving the utilization of current drilling technology and the development of improved drilling techniques and tools.

Hole direction. The hole direction must be controlled within permissible limits in order to reach a desired target at depths as great as 25,000 ft (7600 m). Inclined layers of rocks with different hardnesses tend to cause the direction of drilling to deviate; consequently, deep holes are rarely truly straight and vertical. The drilling rate generally increases as additional drill-collar weight is applied to the bit by adjusting the pipe tension at the surface. However, crooked-hole tendency also increases with higher weight-on-bit. A so-called packed-hole technique has been used to reduce the tendency to hole deviation. One version of this technique makes use of square drill collars that nearly fill the hole on the diagonals but permit fluid and cuttings to circulate around the sides. This procedure reduces the rate at which the hole direction can change.

Directional drilling. In mountainous terrain, it is difficult to construct well locations over each subsurface drilling target, and from offshore drilling platforms it is necessary to drill many wells from a single surface location (**Fig. 1**). For these situations, technology has been developed that permits wells to be drilled directionally from the single surface location to the desired subsurface point.

The earliest method used for directional drilling is whipstocking, which involves placing a wedge-shaped piece of steel at the bottom of the hole to force the bit and drill pipe off into the desired direction (**Fig. 2**). The whipstock is set with the aid of instruments that permit the desired angle and its direction to be initiated. The angle of the drill can then be built up by setting successive whipstocks in place.

Another development is the use of turbines and positive displacement motors that are driven by the circulating mud. The motor is located below a bent sub (**Fig. 3**), a component of the drill string, locate djust above the mud motor, which permits the bit to be set in a predetermined direction. This component initiates the new drilling angle. The motor turns only the bit and not the entire drill string as in conventional rotary operations. Again the angle can be built by successive use of bent subs. Twelve or more directional wells can be drilled from one location and through one leg of a drilling platform by using modern directional drilling techniques.

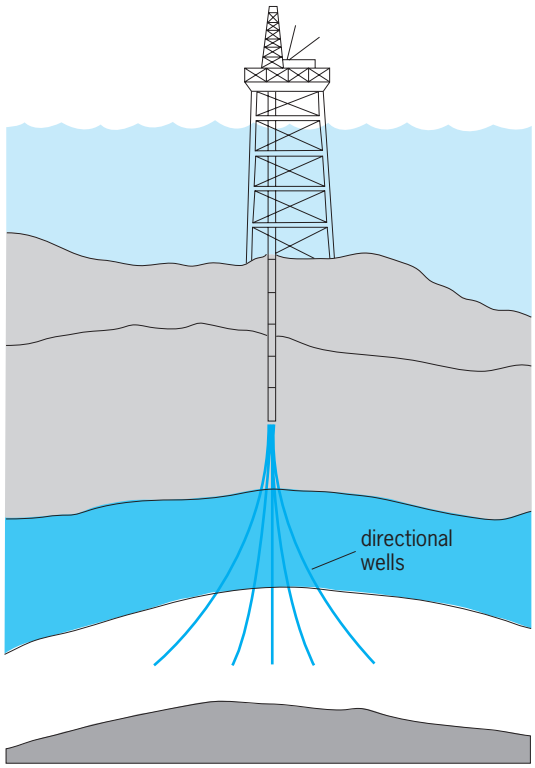

Fig. 1. Directional wells from offshore platform.

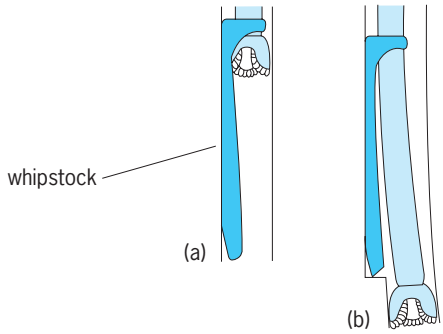

Fig. 2. Whipstocking for deviating a well. (*a*) Whipstock in position. (*b*) Bit and drill pipe deviated by whipstock.

Horizontal drilling. Advances in technology and the need to accomplish special objectives have led to drilling horizontal wells in deep oil or gas reservoirs. The angle of the well is successively built up in order to reach the ultimate horizontal course of the well (**Fig. 4**). The program for drilling the well is developed before hand, based on the technology for actually increasing the angle and then continuing the

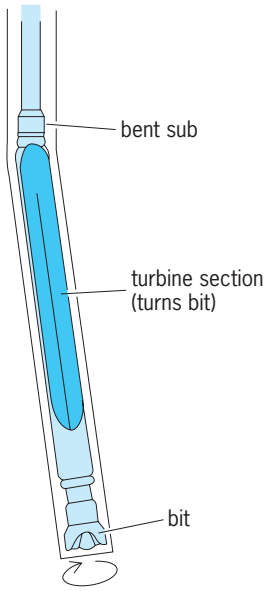

Fig. 3. Turbine for deviating a well.

drilling at the desired angle. In order to be able to drill such a hole without excessive torque and differential sticking, a drilling mud of high lubricity is required. The mud must also have a high yield point and a low fluid loss to prevent hole instability.

One of the shortcomings of horizontal wells is the inability to complete them selectively. Wells have been completed either barefoot (no casing or liner through the producing section) or with noncemented, preslotted lines set through the producing section. The wells are therefore somewhat more vulnerable to being damaged or to collapse by formation transport.

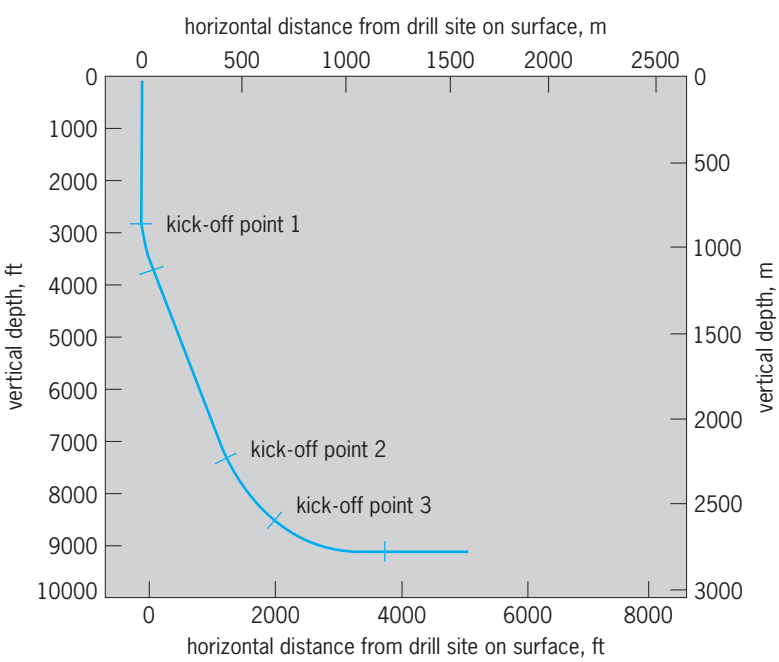

Fig. 4. Trajectory of a horizontal well. At kick-off point 1, the drilling angle is built up 2.5° per 100 ft to a maximum angle of 22.5°. Drilling continues at this angle until kick-off point 2 is reached, where the angle is built at 1.5° per 100 ft to a maximum angle of 44°. At kick-off point 3, the angle is built at 2° per 100 ft to 90°.

The productivity of a horizontal well can be greater than that of a conventional vertical well. Whereas the effective length of a vertical well is limited to that of the producing formation, the length of a horizontal well, theoretically, can be as big as the reservoir itself. The ratio of the productivity of a horizontal well to that of a vertical well increases with the length of the horizontal well, and the effect is more pronounced the greater the value of the parameter L/h, where L is the length of the horizontal well and h is the thickness of the reservoir. The ratio of the horizontal well productivity to that of a conventional well also increases with the value of L/r_e, where r_e is the drainage radius of the reservoir.

The true value of horizontal wells is realized in the following special reservoir situations.

1. The oil saturation and fluid conductivity of the formation may be primarily associated with vertical fractures in echelon. A conventional vertical well may or may not intersect such fractures, depending upon the density and orientation of the fractures. An accurately placed horizontal well can intersect a number of such fractures and effectively drain them.

2. This same advantage will pertain in low-permeability formations, which are often stimulated with hydraulic fractures. These artificially induced fractures will tend to propagate perpendicular to the least principal stress, which is the same direction in which vertical natural fractures are oriented. Hence, horizontal drilling may allow greater production from naturally fractured low-permeability formations by being directed to cross the natural fractures.

3. The oil column is bounded by a gas cap or bottom water, or both. A conventional vertical well under such conditions will be perforated somewhere in the middle of the formation in order to restrict gas coning (the downward flow of gas overlying an oil-saturated interval) and water coning (the entry of bottom water, or mobile water existing below the oil-saturated interval, which rises up and enters the well bore that is open opposite the oil-saturated interval). Such coning actually curtails oil production by reducing the relative permeability of the oil around the wellbore.

4. Another application related to the foregoing is production of the reservoir under a steam drive. Because of the low density of the steam, it overrides the oil column so that oil at the top of the column is heated, mobilized by viscosity reduction, and then displaced by the high-velocity steam to the producing well. As the column of oil is depleted from the top, the cross section available for steam flow increases, and as a result the vapor velocity decreases with decreasing steam vapor velocity, and the ratio of (produced) oil to (injected) steam decreases with a correlative decrease in economic efficiency. A steam-stimulated horizontal well drilled at or near the base of the oil column could promote more economically efficient drainage of oil under the influence of a steam drive.

5. Horizontal drilling may have advantages in certain geological situations by penetrating enhanced porosity or nearby isolated reservoirs that would not be drained by vertical holes. For instance, a number of isolated sand or conglomerate sequences laid down at a low angle as an ocean retreated from a beach could be reached with one horizontal hole. Wells have been drilled that penetrate several miles of formation in a horizontal direction.

Drilling fluids. The increased formation or pore-fluid pressures existing at great depths in the Earth's crust adversely affect drilling. Gushers, blowouts, or other uncontrolled pressure conditions cannot be tolerated; therefore, high-density drilling fluids are used to control well pressures. A normal fluid gradient for salt water is about 0.5 lb/in.^{2} per foot of depth (11 kilopascals/m), and the total stress due to the weight of the overburden increases approximately 1 lb/in.^{2} per foot (23 kPa/m). Under most drilling conditions in permeable formations, the well-bore pressure must be kept between these two limiting values. If the mud pressure is too low, the formation fluid can force the mud from the hole, resulting in a blowout; whereas if the mud pressure becomes too high, the rock adjacent to the well may be fractured, resulting in lost circulation. In this latter case the mud and cuttings are lost into the fractured formation and may also result in a blowout.

High drilling-fluid pressure at the bottom of a borehole impedes the drilling action of the bit. Rock failure strength increases, and the failure becomes more ductile as the pressure acting on the rock is increased. Ideally, cuttings are cleaned from beneath the bit by the drilling-fluid stream; however, relatively low mud pressure tends to hold cuttings in place. In this case, mechanical action of the bit is often necessary to dislodge the chips. Regrinding of fractured rock greatly decreases drilling efficiency by lowering the drilling rate and increasing bit wear.

Drilling efficiency can be increased under circumstances where mud can be replaced by water as the drilling fluid. This might be permissible, for example, in a well in which no high-pressure gas zones are present. Hole cleaning is improved with water drilling fluid because the downhole pressure is lower and no clay filter cake is formed on permeable rock surfaces. So-called fast-drilling fluids provide a time delay for filter cake buildup. This delay permits rapid drilling with no filter cake at the bottom of the hole and, at the same time, prevents excessive loss of fluid into permeable zones above.

In portions of wells where no water zones occur, it is frequently possible to drill using air or natural gas to remove the cuttings. Drilling rates with a gas drilling fluid are often 10 times those obtained with mud under similar conditions. Sometimes a detergent foam is used to remove water in order to permit gas drilling in the presence of limited water inflow. In other instances, porous formations can be plugged with plastic to permit continued gas drilling. However, in many cases it is necessary to revert to either water or mud drilling when a water zone is encountered. Another possibility is the cementing of steel casing through the zone containing water and then proceeding with gas drilling.

Blowout prevention. A drilling well is "kicked" when a high-pressure zone is encountered, and the high-pressure fluids begin to enter the well bore. If the kick is uncontrolled, then the high-pressure fluid continues to enter the well bore and the well may blow out. This is all the more likely to happen if the kick is caused by high-pressure gas, or liquids that contain a significant quantity of dissolved gas, since the gas will continue to expand, displacing more and more drilling mud, as it flows up through the well-bore. The effective density of the drilling mud is thus decreased, and the influx of high-pressure fluid accelerates. The drilling crew will notice the kick by the mud pits' beginning to fill rapidly with the displaced mud.

Drilling operations are immediately stopped, and the blowout preventer (BOP) stacks (**Fig. 5**) are actuated to close in the well. The blowout parameter stack is located just over the well head beneath the drilling floor. The stack is composed of three different seals. The annular preventer is located at the top of the stack and is the first to be actuated. It consists of an expandable, steel-reinforced rubber seal that shuts the annular space between the well bore and drill pipe. If the annular preventer does not succeed in sealing the well, the pipe rams are then closed. These form a steel gate that fits tightly around the drill pipe and seals off the annulus. The blind rams are similar to the pipe rams except that they are used for sealing off the hole if there is no pipe in it. Many blowout preventer stacks are equipped with shear rams, which can cut off the drill pipe and then form the required seal.

As soon as the preventers are closed, dense drilling mud is pumped into the hole down through a choke or kill line. Weighting material is added to the mud until its bottom-hole pressure is sufficient to counterbalance the pressure of the formation that is kicked.

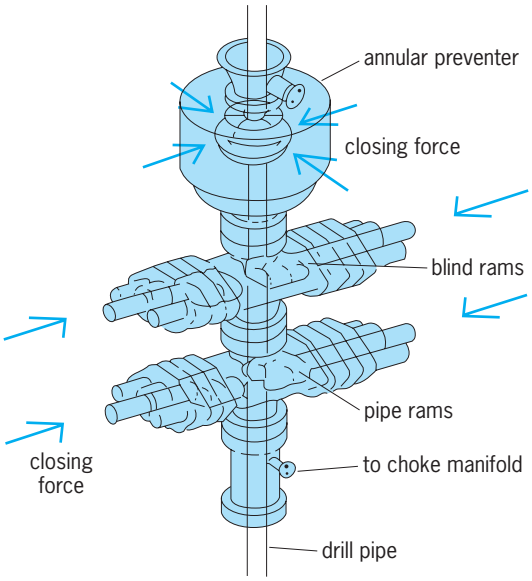

Fig. 5. Components of a BOP stack.

Blowout control. If, despite efforts to prevent a blowout, such an event does occur, it becomes imperative to control the blowout as soon as possible. Usually specialists in blowout control are called to the site to assess the situation and implement control procedures. Various methods may be used, but a typical sequence of events includes cooling the well site with water sprays; removing the surface debris and establishing access to the wellhead; extinguishing the fire if the fluids have ignited; removing damaged wellheads or blowout preventers; and installing a new valve assembly to kill the well.

To extinguish the fire of a blowout, two methods that have been used successfully are to discharge an explosive charge directly above the wellhead, or to direct at the flame a high-pressure spray of a dry chemical such as potassium bicarbonate ($KHCO_3$) in siliconized powder form. After the fire is out, another option to kill a well is to force a hollow tapered tube known as a stinger into the gushing oil. Heavy mud can then be pumped into the well to offset the natural pressures. This method was used extensively to bring oil well fires under control in Kuwait following the Gulf War in 1991.

Cost factors. Drilling costs depend on the costs of such items as the drilling rig, the bits, and the drilling fluid, as well as on the drilling rate, the time required to replace a worn bit, and bit life. The cost per foot increases with depth when encountering geopressures, heavy shale, lost circulation, and well-consolidated, hard formations. Operating costs of a large land rig required for deep holes are much less than comparable costs for an off shore drilling platform or floating drilling vessel. Although weak rock at shallow depth can be drilled at rates exceeding 100 ft/h (30 m/h), drilling rates often average about 5 ft/h (1.5 m/h) in deep holes. Conventional rotary bits have an operating life of 10–20 h. Diamond bits may drill for as long as 50–200 h, but the drilling rate is relatively low and the bits are expensive. For more economical drilling, it is desirable to increase both bit life and drilling rate simultaneously.

Research. Drilling research includes the study of drilling fluids, the evaluation of rock properties, laboratory simulation of field drilling conditions, and the development of new drilling techniques and tools. Fast-drilling fluids have been developed by selecting drilling-fluid additives that plug the pore spaces very slowly, thereby providing the desired time delay for filter cake buildup. Water-shutoff chemicals have been formulated that can be injected into a porous water-bearing formation in liquid form and then, within a few hours, set to become solid plastics. Well-logging techniques can warn of high-pressure permeable zones so that the change from gas or water drilling fluid to mud can be made before the drill enters the high-pressure zone. Downhole instrumentation can lead to improved drilling operations by providing information for improved bit design and also by feeding information to a computer for optimum control of bit weight and rotary speed. Computer programs can also utilize information from nearby wells to determine the best program for optimum safety and economy in new wells.

Since rock is a very hard, strong, abrasive material, there is a challenge to provide drill bits that can penetrate rock more efficiently. A better understanding of rock failure can lead to improved use of present equipment and to the development of better tools. Measurements of physical properties of rocks are correlated with methods for the theoretical analysis of rock failure by a drill bit. A small $1^1/_4$-in.-diameter (32-mm) bit, called a microbit, has been used in scale-model drilling experiments in which independent control is provided for borehole, formation-fluid, and over-burden pressures. These tests permit separation of the effects of the various pressures on drilling rates.

Novel drilling methods that have been explored include studies of rock failure by mechanical, thermal, hydraulic, fusion and vaporization, and chemical means. Jet piercing is widely used for drilling very hard, spallable rocks, such as taconite. Other methods include the use of electric arc, laser, plasma, spark, and ultrasonic drills. *See* DRILLING, GEOTECHNICAL; OIL AND GAS WELL COMPLETION; PETROLEUM GEOLOGY; ROCK MECHANICS; TURBODRILL; WELL LOGGING. Todd M. Doscher

Bibliography. R. Baker, *A Primer of Oilwell Drilling*, 6th ed., 2000; R. A. Dawe and A. G. Lucas (eds.), *Modern Petroleum Technology, vols. 1 and 2*, 6th ed., 2000; M. J. Economides, A. D. Hill, and C. Ehlig-Economides, *Petroleum Production Systems*, 1993; C. T. Franklin, *Handbook of Oil and Gas Operations: Drilling*, vol. 2, 1994; F. A. Giuliano, *Introduction to Oil and Gas Technology*, 1989; N. J. Hyne, *Nontechnical Guide to Petroleum Geology, Exploration, Drilling and Production*, 2d ed., 2001; J. D. Jansen, *Nonlinear Dynamics of Oilwell Drillstrings*, 1993; J. A. Short, *Introduction to Directional and Horizontal Drilling*, 1993; T. Yonezawa et al., Robotic Controlled Drilling: A New Rotary Steerable Drilling System for the Oil and Gas Industry, *Society of Petroleum Engineers*, SPE 74458, 2002.

Oil burner

A device for converting fuel oil from a liquid state into a combustible mixture. A number of different types of oil burners are in use for domestic heating. These include sleeve burners, natural-draft pot burners, forced-draft pot burners, rotary wall flame burners, and air-atomizing and pressure-atomizing gun burners. The most common and modern type that handles 80% of the burners used to heat United States homes is the pressure-atomizing–gun-type burner (**Fig. 1**).

Characteristics. The sleeve burner, commonly known as a range burner because of its use in kitchen ranges, is the simplest form of vaporizing burner. The natural-draft pot burner relies on the draft developed by the chimney to support combustion. The forced-draft pot burner is a modification of the natural-draft pot burner, since the only significant difference between the two types is the means of

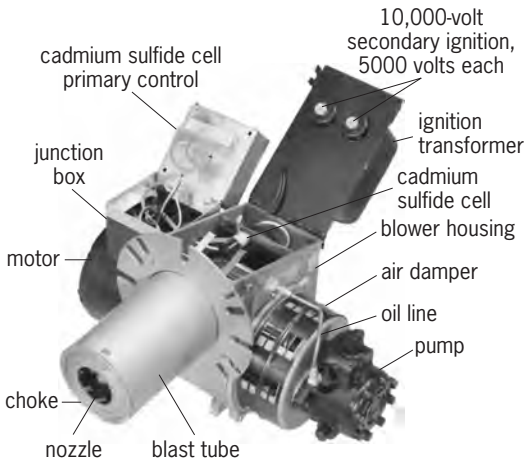

Fig. 1. Oil burner of the pressure-atomizing type. (*Automatic Burner Corp.*)

supplying combustion air. The forced-draft pot burner supplies its own air for combustion and does not rely totally on the chimney. The rotary wall flame burners have mechanically assisted vaporization. The gun-type burner uses a nozzle to atomize the fuel so that it becomes a vapor, and burns easily when mixed with air.

The most important feature of a high-pressure atomizing gun burner is the method of delivering the air. The most efficient burner is the one which completely burns the oil with the smallest quantity of air. The function of the oil burner is to properly proportion and mix the atomized oil and air required for combustion (**Fig. 2**).

Efficiency. If a large quantity of excess air is used to attempt to burn the oil, there is a direct loss of usable heat. This air absorbs heat in the heating unit which is then carried away through the stack with the combustion gases. This preheated air causes high stack temperatures which lower the efficiency of the combustion. The higher the CO_2 (carbon dioxide), the less excess air. Overall efficiencies or stack loss can be estimated by the use of the stack loss chart (**Fig. 3**).

Increased efficiency can be obtained through the use of devices located near the flame end of the burner. CO_2 is not the ultimate factor in efficiency. Burners producing high CO_2 can also produce high smoke readings. Accumulations of $^1/_8$-in. (3-mm) soot layers on the heating unit surface can increase fuel consumption as much as 8%. An oil burner is always adjusting to start smoothly with the highest CO_2 and not more than a number two smoke on a Bacharach smoke scale. For designs in high-efficiency burners commonly known as flame-retention-type burners see Fig. 2.

Nozzle. The nozzle is made up of two essential parts: the inner body, called the distributor, and the outer body, which contains the orifice that the oil sprays through.Under this high pump pressure of

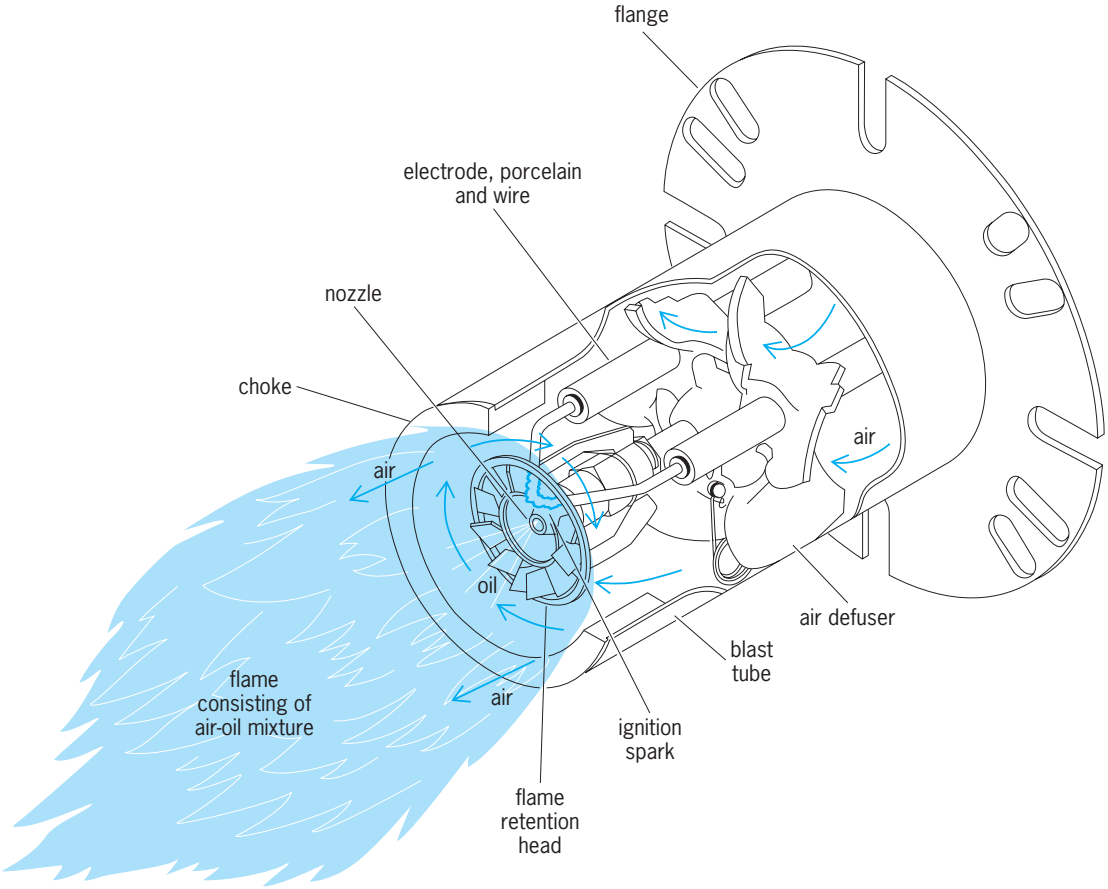

Fig. 2. Flame-retention–type burner, an example of a high-efficiency burner.

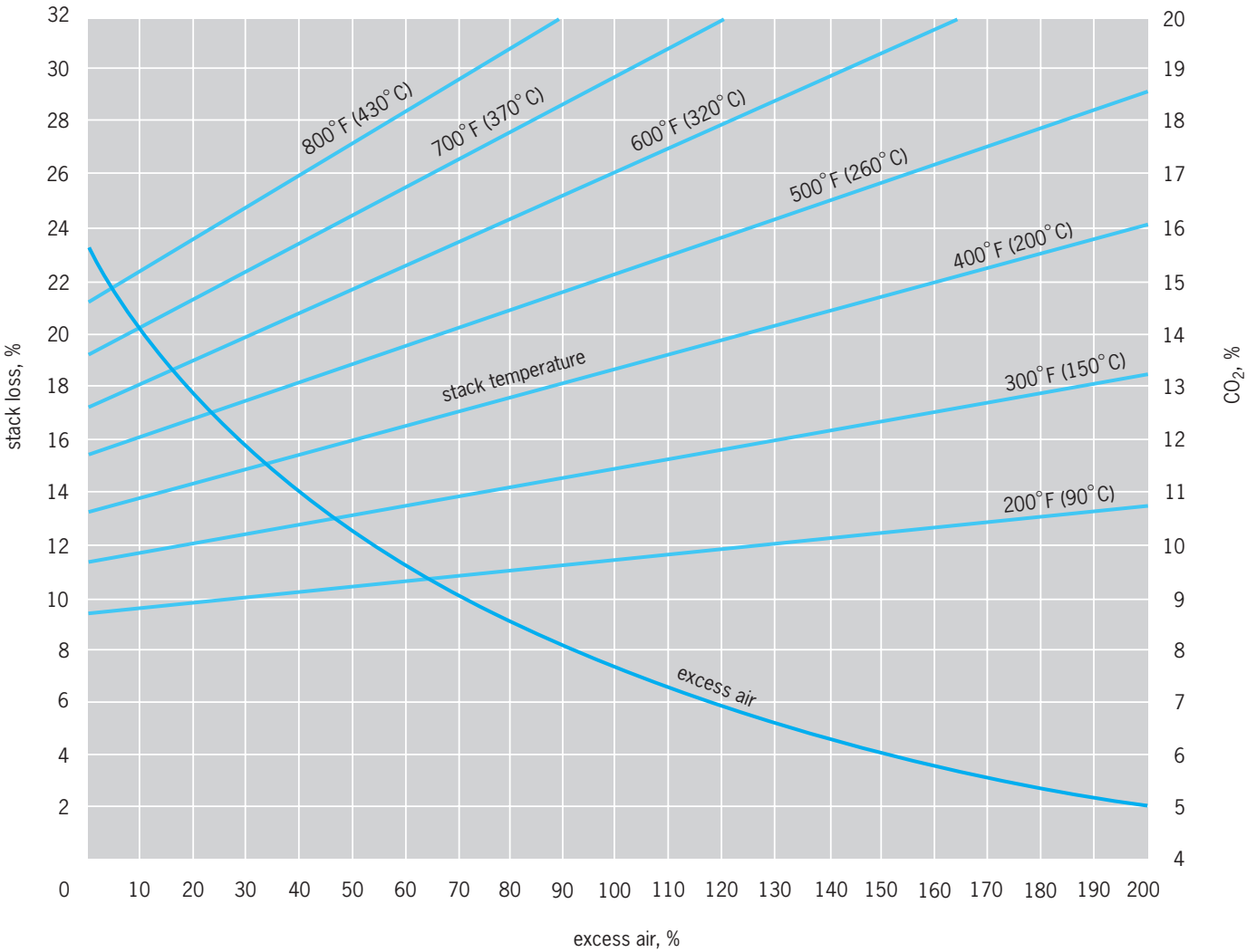

Fig. 3. Typical stack loss chart. To determine stack loss and efficiency, start with correct CO_2 and follow horizontal line to excess air curve, then vertical line to stack temperature, and finally horizontal line to stack loss. Overall efficiency percent = 100 − stack loss.

100–300 psi (700–2100 kilopascals) the oil is swirled through the distributor and discharged from the orifice as a spray (**Fig. 4**). The spray is ignited by the spark and combustion is self-sustaining, provided the proper amount of air is supplied by the squirrel cage blower. Air is delivered through the blast tube and moved in such a manner that it mixes well with the oil spray (Fig. 2). This air is controlled by a damper located on either the intake side or the discharge side of the fan (Fig. 1).

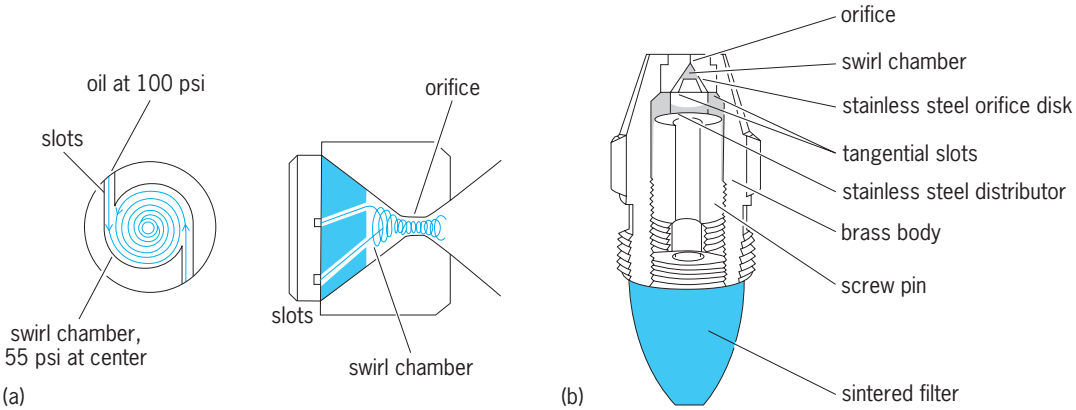

Fig. 4. Oil burner nozzle. (*a*) Operation of a nozzle. 1 psi = 7 kPa. (*b*) Cutaway view. (*Delavan Manufacturing Co.*)

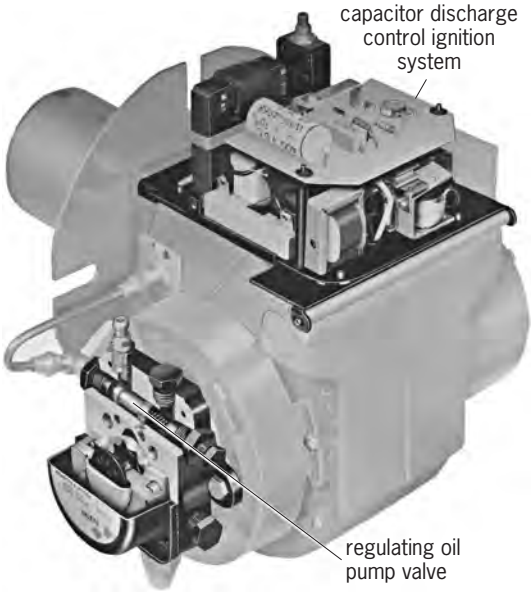

Fig. 5. Oil burner, showing capacitor discharge control ignition system and regulating valve for fuel pump. (*Automatic Burner Corp.*)

The essential parts of the pressure gun burner are the electric motor, squirrel-cage–type blower, housing, fuel pump, electric ignition, atomizing nozzle, and a primary control (Fig. 1 and **Fig. 5**).

The fuel oil is pumped from the tank through the pump gears, and oil pressure is regulated by an internal valve that develops 100–300 psi (700–2100 kPa) at the burner nozzle. The oil is then atomized, ignited, and burned.

The fine droplets of oil that are discharged from the nozzle are electrically ignited by a transformer which raises the voltage from 115 to 10,000 V. The recently developed electric ignition system called the capacitor discharge system incorporates a capacitor that discharges voltage into a booster transformer developing as much as 14,000 V. The electric spark is developed at the electrode gap located near the nozzle spray (Figs. 2 and 5).

The high velocity of air produced by the squirrel cage fan helps develop the ignition spark to a point where it will reach out and ignite the oil without the electrode tips actually being in the oil spray. Figure 2 shows a typical gun-type burner inner assembly.

Installations. Fuel oils for domestic burners are distilled from the crude oil after the lighter products have been taken off. Consequently, the oil is nonexplosive at ordinary storage temperatures. Domestic fuel oils are divided into two grades, number one and number two, according to the ASTM specification. There are several different methods of installing an oil tank system with an oil burner (**Fig. 6**).

One of the most important safety items of an automatic oil burner is the primary control. This device will stop the operation when any part of the burner or the heating equipment does not function properly. This control protects against such occurrences as incorrect primary air adjustments, dirt in the atomizing nozzle, inadequate oil supply, and improper combustion. The most modern type of primary control is called the cadmium cell control. The cadmium sulfide flame detection cell is located in a position where it directly views the flame. If any of the above functional problems should occur, the electrical resistance across the face of the cell would increase, causing the primary control to shut the burner off in 70 s or less.

Draft caused by a chimney or technical means is a very important factor in the operation of domestic oil burners. The majority of burners are designed to fire in a heating unit that has a minus 0.02 in. (0.5 mm) water column draft over the fire. Because the burner develops such low static pressures in the draft tube, over fire drafts are an important factor in satisfactory operation.

If the heating unit is designed to create pressure over the fire, a burner developing high static pressures in the blast tube must be used. These burners develop high static by means of higher rpm on the motor and fan. Burners of this nature are available and are being produced. Flame control is very important. Firing heads similar to Fig. 2 are used on applications of this nature.

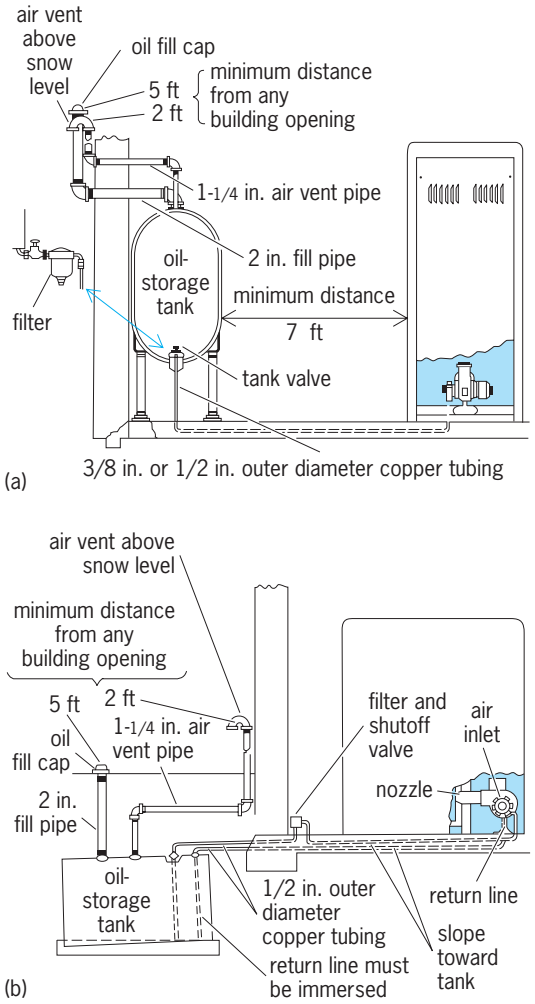

Fig. 6. Two typical installation methods for oil storage tank and oil burner. (*a*) Inside basement. (*b*) Outside underground storage tank. 1 ft = 0.3 m. 1 in. = 2.5 cm.

The oil burner is also used for a wide assortment of heating, air conditioning, and processing applications. Oil burners heat commercial buildings such as hospitals, schools, and factories. Air conditioners using the absorption refrigeration system have been developed and fired with oil burners. Oil burners produce CO_2 in greenhouses to accelerate plant growth. They produce hot water for many commercial and industrial applications. *See* AIR COOLING; COMFORT HEATING; HOT-WATER HEATING SYSTEM; OIL FURNACE.

Robert A. Kaplan

Bibliography. American Society of Heating, Refrigerating, and Air-conditioning Engineers, *ASHRAE Fundamentals Handbook*, 1994, revised 1997; *ASHRAE, Equipment Handbook*, 1992; E. M. Field, *Oil Burners*, 5th ed., 1990.

Oil furnace

A combustion chamber in which oil is the heat-producing fuel. Fuel oils, having from 18,000 to 20,000 Btu/lb (42 to 47 megajoules/kg), which is equivalent to 140,000 to 155,000 Btu/gal (39 to 43 MJ/liter), are supplied commercially. The lower flash-point grades are used primarily in domestic and other furnaces without preheating. Grades having higher flash points are fired in burners equipped with preheaters. The ease with which oil is transported, stored, handled, and fired is advantageous in small installations. The fuel burns almost completely so that, especially in a large furnace, combustible losses are negligible.

Domestic oil furnaces with automatic thermostat control usually operate intermittently, being either off or operating at maximum capacity. The heat-absorbing surfaces, especially the convection surface, should therefore be based more on maximum capacity than on average capacity if furnace efficiency is to be high. The combustion chamber should provide at least 1 ft^3 for each 1.5–2 lb (1 m^3 for each 24–32 kg) of fuel burned per hour. Gas velocity should be below 40 ft/s (12 m/s). The shape of the chamber should follow the outline of the flame. *See* FUEL OIL; OIL BURNER; STEAM-GENERATING FURNACE.

Frank H. Rockett

Oil sand

A loose-to-consolidated sandstone or a porous carbonate rock, impregnated with bitumen, a heavy asphaltic crude oil with a viscosity under reservoir conditions greater than 10,000 millipascal seconds (10,000 centipoise); also known as tar sand or bituminous sand. *See* ASPHALT AND ASPHALTITE.

Heavy-oil reservoirs are distributed throughout the world and were formerly categorized in many different and sometimes conflicting ways. Under the leadership of the United Nations Institute for Training and Research (UNITAR), modern definitions for these heavy substances have emerged: (1) Bitumens (oil sand hydrocarbons) possess viscosities greater than 10,000 mPa · s (10,000 cP) and cannot be produced by conventional methods. (2) Extra-heavy crude oils possess viscosities equal to or less than 10,000 mPa · s (10,000 cP), and an American Petroleum Institute (API) gravity less than 10°. They are heavier than water but have some mobility under reservoir conditions. (3) Heavy crude oils possess API gravities ranging from 10 to 20°.

About 175 experimental projects for the enhanced recovery of heavy oil and bitumen are in operation around the world and, because similar technologies are being used, reserve and recovery estimates for the three categories—heavy crude, extra-heavy crude, and bitumen—frequently are lumped together.

Estimates of the combined world reserves of heavy crude and bitumen exceed 5×10^{12} bbl (8×10^{11} m^3), with bitumen accounting for about half the total. The seven largest of these heavy crude and bitumen deposits contain almost as much oil in place as the world's 300 largest conventional oil fields. The largest proven accumulation of bitumen-containing oil sands occurs in Alberta, Canada, but deposits in Russia rank close behind. Much of the huge accumulation in Venezuela's Orinoco Basin has been reclassified as heavy and extra-heavy crude oil, but this area also, seemingly, contains substantial reserves of bitumen. Smaller reserves occur, in descending order, in the United States, Madagascar, Italy, Albania, Trinidad, and Romania. Estimates of the bitumen that will prove ultimately to be recoverable—generally of the order of 5–25%—are speculative and depend on the development of successful technologies at competitive costs.

Geological factors. The major oil sand and heavy oil deposits are remarkable not only for their size but also for their unconventional geologic settings. The deposits generally exhibit certain characteristics. (1) They tend to occur in a deltaic or nonmarine environment, featuring extensive, fingerlike, fluvial sands with large porosities and permeabilities. Porosities, often of the order of 25–35%, tend to be considerably greater than in most petroleum reservoir sandstones (5–20%). A lack of mineral cement, which occupies the void space in most sandstones, permits these high porosities and gives rise to the term oil sands. (2) The deposits are covered by a widespread regional shale cap which restrained upward escape of the oil migrating into the ancient river channels, the efficient systems that gathered the subsurface fluids. (3) The deposits are emplaced in a gently dipping homocline with updip stratigraphic convergence. (4) Degradation of the original, lighter crude to heavy oil has been effected by water washing and bacterial action. Considerable debate continues about the origin of the heavy oils and whether they are thermally mature or immature. The evidence suggests, however, that the oils were degraded in place.

While most of the world's heavy-oil reservoirs are found in deltaic settings, there are significant exceptions. These exceptions include the Cambrian and Lower Permian carbonate rocks, which host the bulk of Russia deposits, and the Paleozoic Carbonate

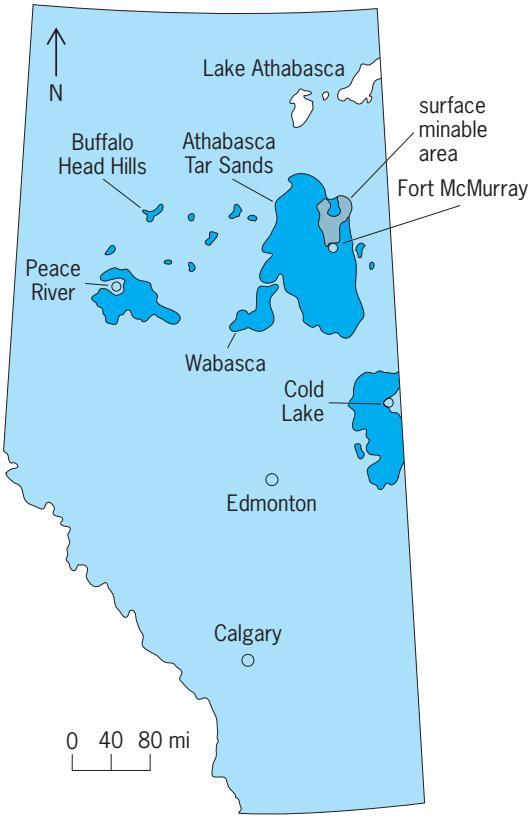

Tar sand deposits in Alberta. 1 mi = 1.6 km.

Triangle in northern Alberta. *See* PETROLEUM GEOLOGY.

Alberta deposits. Canada's Alberta Oil Sands, comprising four enormous and a number of lesser deposits, contain the best known of the world's oil sand reservoirs. Established resources are of the order of 1.25×10^{12} bbl (2×10^{11} m^3). The Athabasca Tar Sands deposit is the largest known petroleum accumulation in the world, with a total area of 13,000 mi^2 (34,000 km^2) and in-place reserves of 9.19×10^{11} bbl (1.46×10^{11} m^3). Cold Lake, rapidly gaining prominence for its in-place projects, contains 2.05×10^{11} bbl (3.26×10^{10} m^3); Peace River, 7.5×10^{10} bbl (1.2×10^{10} m^3); and Wabasca, 4.26×10^{10} bbl (6.8×10^9 m^3). The deposits range across the northern part of the province (see **illus.**).

The Athabasca Oil Sands are located in the McMurray Formation, of Lower Cretaceous age. The formation outcrops along the Athabasca River, north of the city of Fort McMurray. Elsewhere, the deposit is buried under a variable layer of overburden which reaches up to 1720 ft (525 m) in thickness. Fortunately, some of the richest parts of the deposit are covered by the thinnest overburden.

Russian deposits. Hundreds of bitumen occurrences have been charted in Russia; about 70% are found in carbonate rock. Estimates of reserves approximate 1.15×10^{12} bbl (1.8×10^{11} m^3). Of considerable interest is the Olenek anticline in northeastern Siberia where, it is speculated, reserves total 6×10^{11} bbl (9.5×10^{10} m^3). Huge deposits have also been discovered near Baku in Kazakhstan and at Melekess in Volga-Ural. Yarega, in Siberia's Timan-Pechora province, is the only place in the world where underground mining for commercial production of heavy oil is carried out.

Venezuela deposits. Enormous reserves of heavy oil occur in a 36-mi-wide (60-km) band stretching along the northern bank of the Orinoco River for a distance of 400 mi (700 km). The deposits, for many years known collectively as the Orinoco Tar Belt, contain between 7×10^{11} bbl (1.1×10^{11} m^3) and 1×10^{12} bbl (1.6×10^{11} m^3) of heavy hydrocarbons. Some of the crude possesses an API gravity of between 10 and 20°, but most is between 4 and 10°. Most of the heaviest oil, however, shows some mobility because of the below-normal asphaltene content, the temperature of the reservoir, and the presence of dissolved gases. For this reason, the Venezuelan government encourages the use of the term Orinoco Heavy Oil Belt, instead of Tar Belt, to describe these deposits. It is believed, though, that substantial reserves, perhaps as much as one-third of the total hydrocarbons, are too viscous to flow to the well bore and may therefore be categorized as bitumen, or tar oil.

United States deposits. In the United States, 550 tar sand occurrences are known to exist in 22 states, and it is expected the number will increase significantly in future surveys. Seven states (Alabama, California, Kentucky, New Mexico, Texas, Utah, and Wyoming) possess deposits aggregating 1×10^6 bbl (1.6×10^5 m^3) or more. Total United States reserve estimates have remained relatively constant since the mid-1970s at approximately 3×10^{10} bbl (4.8×10^9 m^3), but considerable changes have occurred in the amounts ascribed to the various states.

Utah is the most important tar sand state with an estimated 1.47×10^{10} bbl (2.3×10^9 m^3) of bitumen in place, or about half of the United States total. At one time, though, it was believed that more than 90% of United States tar sand reserves were located there. Over 50 deposits have been identified, but the vast bulk of the oil occurs in six giant deposits, four of which are in the Uinta Basin of northeastern Utah. Three of these (P. R. Spring, Hill Creek, and Sunnyside) occur in the Green River Formation (Eocene), but Asphalt Ridge, probably the best-known tar sand deposit in the country, occurs in older strata.

The Tar Sand Triangle, with reserves of approximately 6.1×10^9 bbl (9.6×10^8 m^3), the largest deposit in the United States, lies in a remote and very rugged area of northeastern Utah, in the White Rim Sandstone Formation. Most of the deposit underlies Federal Recreation and Wilderness areas. Federal agencies also control major sections of the P. R. Spring, Hill Creek, and Sunnyside deposits. The other major ore body, Circle Cliffs, is also located in southeastern Utah.

The thickness of the Utah reservoirs ranges between 0 and 330 ft (0 and 100 m), and the depth ranges from 0 to 2200 ft (0 to 670 m). Some interesting characteristics distinguish the reservoirs: the bitumen occurs in consolidated sandstone; the connate water, at 0.6%, and the sulfur content of the

hydrocarbon, at 0.5%, are about one-tenth of the corresponding values in the Canadian and Venezuelan oil sands. However, the porosity of the formations and the continuity and magnitude of the bitumen saturation are far less. Economic exploitation is therefore in question.

In terms of United States reserves, Alabama has moved into second place, with 4.3×10^9 bbl (6.8×10^8 m^3) attributed to the North Area. Texas (Uvalde area) and California are next, each with 3×10^9 bbl (4.8×10^8 m^3).

In California, the reserves are contained in a number of well-known deposits, Santa Maria, with 2×10^9 bbl (3.2×10^8 m^3), being the largest. The Edna deposit, located midway between Los Angeles and San Francisco, is a rarity among oil sand deposits in that it is considered a marine facies. The Sisquoc deposit is amenable to the hot water extraction process used by the large Canadian projects. The fossiliferous McKittrick deposit is more properly considered a diatomaceous oil shale. *See* OIL SHALE.

Other major world deposits. The Bemolanga deposit in western Madagascar is noteworthy. Covering approximately 154 mi^2 (400 km^2), it contains reserves estimated at 2.1×10^{10} bbl (3.3×10^9 m^3). Another country with multibillion-barrel oil sand deposits is Italy; the Ragusa area in Sicily contains 1.4×10^{10} bbl (2.2×10^9 m^3) of bitumen. G. Ronald Gray

Bibliography. C. J. Borregales, Production characteristics and oil recovery in the Orinoco Oil Belt, *UNITAR 1st International Conference on the Future of Heavy Crude and Tar Sands*, Edmonton, Alberta, 1979; M. Danyluk, B. Galbraith, and R. Omana, Department of Defense, Toward definitions for heavy crude oil and tar sands, *United Nations Institute of Training and Research (UNITAR) 2d International Conference on the Future of Heavy Crude and Tar Sands*, Caracas, Venezuela, 1982; G. J. Demaison, Tar Sands and Supergiant Oil Fields, Can. Inst. Min. Metallurgy Spec. Vol. 17, 1977; V. A. Kuuskraa, S. Chalton, and T. M. Doscher, *The Economic Potential of Domestic (U.S.) Tar Sands*, U.S. Department of Defense, January 1978; R. F. Meyer, P. A. Fulton, and W. D. Dietzman, A preliminary estimate of world heavy crude oil and bitumen resources, *UNITAR 2d International Conference on the Future of Heavy Crude and Tar Sands*, Caracas, Venezuela, 1982; G. D. Mossop, *Geology of the Athabasca Oil Sands*, SCIENCE Reprint Series, January 1980.

Oil shale

A sedimentary rock containing solid, combustible organic matter in a mineral matrix. The organic matter, often called kerogen, is largely insoluble in petroleum solvents, but decomposes to yield oil when heated. Although "oil shale" is used as a lithologic term, it is actually an economic term referring to the rock's ability to yield oil; oil shale appears to be the cheapest source after natural petroleum for large amounts of liquid fuels. No real minimum oil yield or content of organic matter can be established to distinguish oil shale from sedimentary rocks. Additional names given to oil shales include black shale, bituminous shale, carbonaceous shale, coaly shale, cannel shale, cannel coal, lignitic shale, torbanite, tasmanite, gas shale, organic shale, kerosine shale, coorongite, maharahu, kukersite, kerogen shale, algal shale, and "the rock that burns." *See* KEROGEN.

Origin. Oil shale is lithified from lacustrine or marine sediments relatively rich in organic matter. Most sedimentary rocks contain small amounts of organic matter, but oil shales usually contain substantially more. Specific geochemical conditions are required to accumulate and preserve organic matter, and these were present in the lakes and oceans whose sediments became oil shale. R. M. Garrels and C. L. Christ define these conditions in terms of oxidation-reduction potential and acid-base condition (pH) of the water in and around the sediment. Organic matter accumulates under the strongly reducing conditions and neutral or basic pH present in euxinic marine environments and organic-rich saline waters. The organic-rich sediments which became oil shale accumulated slowly in water isolated from the atmosphere. This isolation was achieved through the stratification of the water body. The stratification limited clastic mineral input to the sediment and also protected the sediment from oxidation and from physical disturbance. *See* MARINE SEDIMENTS; PH.

Mineral composition. Quartz, illite, and pyrite (sometimes with marcasite and pyrrhotite) occur in virtually every oil shale. Other clays (particularly montmorillonite) are found in many oil shales, as are feldspar minerals. Most oil shale deposits contain small amounts of carbonate minerals, but some, most notably the Green River Formation in Colorado, Utah, and Wyoming, contain large amounts of dolomite, calcite, and other carbonates. The oil shale minerals were formed probably in the sediment by chemical processes related, at least in part, to the presence of organic matter. *See* CALCITE; CARBONATE MINERALS; DOLOMITE.

Some oil shales, particularly those called black shales because of the coal-like color of their organic matter, have tended to become enriched in trace metals. The reducing conditions necessary to preserve organic matter were conducive to precipitating available trace metals, frequently as sulfides. The Kupferschiefer of Mansfield, Germany, contains an unusually large content of copper, and the Swedish Alum Shale has been exploited for its uranium content. The Devonian Chattanooga Shale of Tennessee and neighboring states contains an average of about 0.006 wt % uranium and has been extensively studied as a potential low-grade source of this element. Vanadium in potentially commercial amounts occurs in the Permian Phosphoria Formation of Wyoming and Idaho. Enrichment of arsenic, antimony, molybdenum, nickel, cadmium, silver, gold, selenium, and zinc has also been noted in black oil shales.

Physical properties. Oil shales are fine-grained rocks generally with low porosity and permeability. Many are thinly laminated and fissile. The colors of

oil shales, which range from black to light tan, are produced or altered by organic matter.

The physical properties of oil shale are influenced strongly by the proportion of organic matter in the rock. The rock's decrease in density with increasing organic content illustrates this most graphically. The mineral components possess densities of about 2.6–2.8 g/cm^3 for silicates and carbonates and 5 g/cm^3 for pyrite, but the density of organic matter is near 1 g/cm^3. Larger fractions of organic matter produce rocks with appreciably lower density. Equation (1)

$$D_T = \frac{D_A D_B}{A(D_B - D_A) + D_A} \qquad (1)$$

quantifies this relationship. Here D_T = density of the rock, A = weight fraction of organic matter, D_A = average density of organic matter, and D_B = average density of mineral matter.

The volume of organic matter in oil shale rock strongly influences its physical properties. The marked increase in volume fraction of organic matter in oil shale with increasing organic weight fraction can be demonstrated by using the above equation. For example, in Green River Formation oil shale the average density of the organic matter (D_A) is about 1.07 g/cm^3, and the average density of the mineral matter (D_B) is about 2.72 g/cm^3. Thus the density of an oil shale containing 4 wt % organic matter (a shale yielding about 6 gal of oil per ton of shale) is calculated to be 2.56 g/cm^3. The organic matter occupies about 10 vol % of this rock. An oil shale containing 15 wt % organic matter (a shale yielding about 26 gal per ton or 110 liters per metric ton) possesses a density of 2.21 g/cm^3, and the organic matter occupies 31 vol % of this rock. An oil shale containing 39 wt % organic matter (a shale yielding about 52 gal per ton or 220 liters per metric ton) possesses a density of 1.86 g/cm^3, and the organic matter makes up about 52 vol % of the rock's volume. In the Green River Formation, in oil shales containing 15 wt % or more organic matter, the organic material is the largest component by volume, and the physical properties of the organic matter predominate in determining the physical properties of the rock. The organic matter makes this rock tough, resilient, and difficult to crush. The richer Green River Formation oil shales tend to deform plastically under load. The volume of organic matter in oil shales can be used to estimate mechanical properties of rocks in oil shale deposits.

Equation (1) can yield a relationship between oil yield and oil shale density by incorporating a factor for conversion of the organic matter to oil. The resulting relationship is useful in calculating resources and reserves in an oil shale deposit. Such relationships have been worked out for Green River Formation oil shales. Similar relationships can be developed for most oil shale deposits.

Organic composition. The organic matter in oil shales and other sedimentary rocks has been studied extensively by organic geochemists, but a specific description of it has not been produced. Although some oil shales contain recognizable organic fragments like spores or algae, most do not, because the reducing conditions associated with oil shale deposition digested and homogenized the organic debris. The resulting organic matter (kerogen) is best described as a high-molecular-weight organic mineraloid of indefinite composition. This composition varies from deposit to deposit and is influenced by the depositional conditions and the nature of the organic debris. Variations in the hydrogen content of this organic matter are significant, because the fraction of organic carbon converted to oil on heating increases as the amount of hydrogen available in the organic matter increases. To illustrate this relationship, the **table** compares the proportion of organic carbon converted and recovered as oil during Fischer assay with the weight ratio of organic carbon to organic hydrogen in several oil shales. For petroleum, the carbon-hydrogen values range from 6.2 to about 7.5; for coal, they range upward from 13. The carbon-hydrogen values for organic matter in the world's oil shales range from near petroleum to near coal.

Relationship between the organic carbon-to-organic hydrogen ratio and the conversion of oil shale organic matter to oil by heating

Deposit sampled	Carbon-hydrogen ratio	Organic carbon converted, wt %
Pictou County, Nova Scotia, Canada	12.8	13
Top Seam, Glen Davis, Australia	11.5	26
New Albany Shale, Kentucky, United States	11.1	33
Ermelo, Transvaal, South Africa	9.8	53
Cannel Seam, Glen Davis, Australia	8.4	60
Garfield County, Colorado, United States	7.8	69

Analytical determination of the elemental composition of the organic matter has been difficult because of the heterogenous nature of oil shales. Carbon, hydrogen, sulfur, oxygen, and nitrogen are the major elements of the organic matter; but they also occur in the mineral matter of oil shales. The organic matter and the mineral matter in oil shales are difficult to separate from each other either physically or chemically. Analytical techniques designed to distinguish between organic and mineral forms of elements, specialized organic matter enrichment techniques, and other specialized evaluation techniques have been developed to aid in the study of oil shales.

Oil production. The Fischer assay is the best known of the specialized analytical procedures used for oil shale. It was developed by the U.S. Bureau of Mines for oil shale resource evaluation and has become a standard ASTM (American Society for Testing and Materials) method. This method, employing a modified Fischer retort, determines quantities of oil products recoverable from an oil shale sample heated under

prescribed conditions. Although the procedure does not measure the total amount of organic matter in the sample, it approximates the oil available by commercial operations. This simple procedure has proved suitable for oil shale evaluation purposes. Resource information for United States oil shales is based on the Fischer assay oil-yield data accumulated by the U.S. Bureau of Mines.

Nuclear magnetic resonance (NMR) on solid oil shale samples has been applied to evaluating the fraction of organic carbon in an oil shale which can be converted to oil by heat. This conversion fraction, correlated above with the laboriously determined organic hydrogen, works with the quantity of organic matter in an oil shale resource to define how much rock must be heated to generate production quantities of oil. The carbon-13 NMR test divides the organic carbon into hydrogen-rich (aliphatic) and hydrogen-poor (aromatic) fractions. The proportion of aliphatic carbon correlates very well with the proportion of organic carbon recovered in the oil produced under the standard heating conditions like Fischer assay. Equation (2) is used for this correla-

$$\mathrm{CONV} = 1.37 \cdot \mathrm{AL} - 0.41 \qquad (2)$$

tion in United States oil shales, where CONV is the fraction of organic carbon converted to oil and AL is the fraction of organic carbon defined by NMR as occurring in aliphatic compounds. **Figure 1** illustrates this easily interpreted NMR characterization of organic carbon for a Green River Formation oil shale from Colorado and a Devonian Black Shale from Kentucky. The *x* axis measures the relative chemical shift generated by the organic matter in the oil shale; the shift is a function of proportion of available hydrogen in the organic matter of the oil shale. The Colorado sample shows much greater aliphatic content and a correspondingly greater conversion of organic carbon to oil. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

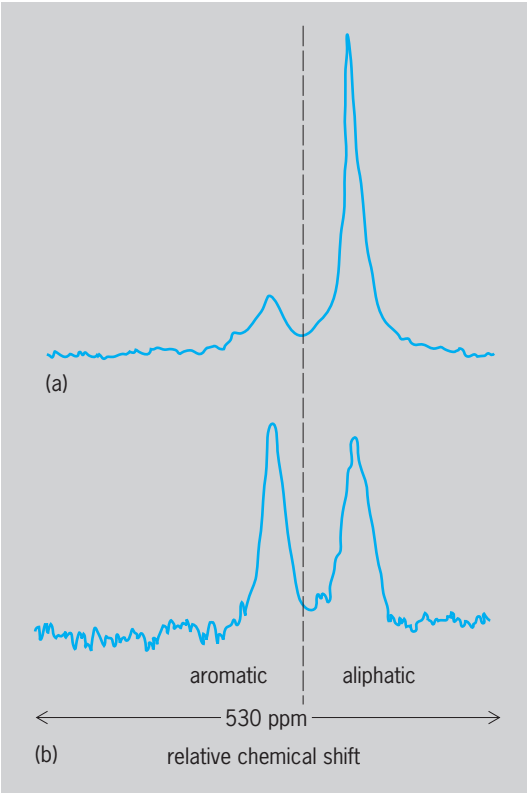

Fig. 1. C-13 nuclear magnetic resonance characterization of organic matter in two United States oil shale samples. (*a*) Colorado oil shale, 80% aliphatic carbon content with 69% conversion of organic carbon to oil. (*b*) Kentucky oil shale, 52% aliphatic carbon content with 30% conversion of organic carbon to oil. The relative area under the two peaks provides the proportion of the aliphatic carbon in the oil shale's organic carbon. The *x* axis is a measurement of the chemical shift of the organic carbon in the oil shale; the shift can be related to the proportion of available hydrogen in the organic matter of the oil shale. (*After F. P. Miknis and J. W. Smith, An NMR survey of United States oil shales, Geochem., 5:93–201, 1984*)

Oil shale units. The United States expresses oil yield determined by Fischer assay in the volume unit of gallons of oil per ton of shale. This unit is converted from weight percent oil yield by multiplying by a conversion factor, 2.4, and dividing the result by absolute oil specific gravity, determined at 15.6°C (60.1°F). The metric volume unit, liters per metric ton, corresponds to gallons per ton. One gallon per ton is equivalent to 4.172 liters/metric ton. Oil shale resource values are expressed in barrels of oil in the United States. A barrel of oil, the 42-gal volume unit used in Western petroleum commerce, has no direct equivalent in metric countries. However, the SI unit equivalent to a barrel of oil represents 0.159 m^3. Specification of oil density, a variable, is necessary to convert the volume unit, barrels, into the metric weight trade unit, metric ton. The 1975 World Energy Conference agreed to define a barrel of oil as 0.145 metric ton, and 1 gal/ton × 0.29 as 1 kg/metric ton, approximations ignoring density variations in oil. Gallons of oil per ton of shale will be the unit of oil yield used in the remainder of this article.

World resources. The world's oil shale deposits represent a tremendous store of fossil energy. It has been estimated that the organic matter in sedimentary rocks contains 1.2×10^{16} tons (1.1×10^{16} metric tons) of organic carbon, nearly 1000 times that found in coals. Although part of that organic carbon has matured to produce oil and gas, most of it is still oil shale. Unfortunately, most of this tremendous resource is not well known. Oil shales occur on every continent in sediments ranging in age from Cambrian to Tertiary. **Figure 2** illustrates this by showing distribution and geologic age of oil shales in the United States. More than 20% of the United States land area is underlain by oil shales, but only the Tertiary Green River Formation and the Devonian black shales, particularly those in Kentucky, have been evaluated in any detail. The balance of the deposits shown are known from associated geologic studies but not as oil shale resources. With the limited amount of information available, one estimate of known oil shale resources of the world was made, and to it was added possible extensions of known resources and geologically based estimates of unappraised and undiscovered

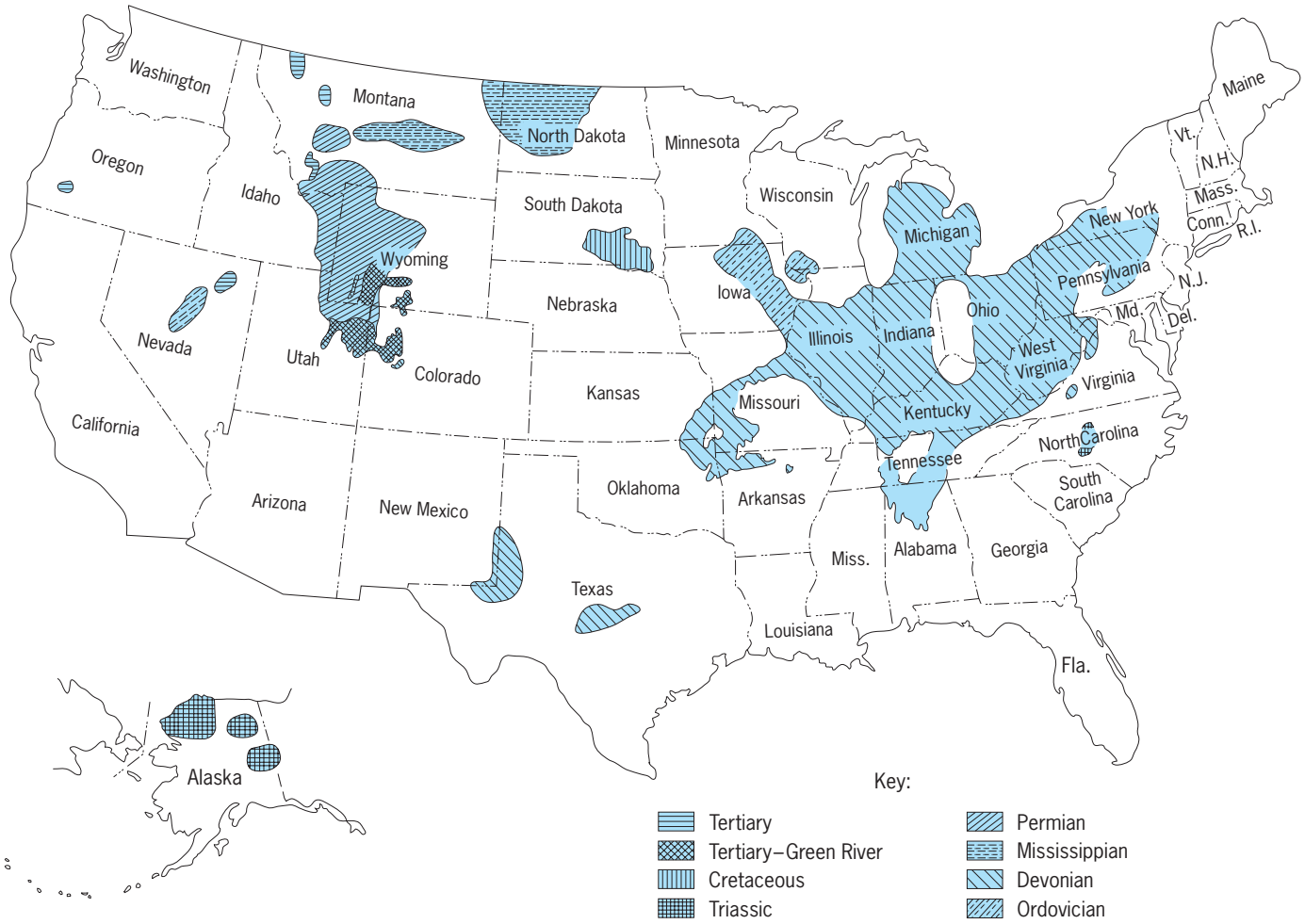

Fig. 2. Oil shale deposits of the United States. (*Courtesy of John Ward Smith*)

resources to obtain order-of-magnitude numbers for the total in-place oil resource in the world's oil shale deposits. These estimates for the total oil resource in shales of all grades reached 1.75×10^{15} barrels. How much of this tremendous resource could actually be produced economically is a variable which changes continuously with the supply of natural oil and developments in technology. However, just 1% of that total shale oil represents more oil than the world is expected to produce as natural petroleum (2×10^{12} bbl). Oil shale represents a tremendous supply of liquid fuels.

World developments. Although the oil potential of the world's oil shales is great, commercial production of this oil has been considered uneconomic generally. Oil shales are lean ores, producing only limited amounts of oil which historically has been low in price. Mining and heating 1 ton of relatively rich oil shale yielding 25 gal/ton produces only 0.6 bbl of oil.

In special situations when other fuels were in short or uncertain supply, or when energy transportation was difficult, energy development from oil shales has been carried out commercially. World War II caused sharp increases in petroleum demand and disrupted both petroleum production and petroleum distribution. Oil shale production operations during and since World War II have been conducted in Germany, France, Spain, Manchuria (China), Estonia and other areas of the former Soviet Union, Sweden, Scotland, South Africa, Australia, and Brazil. *See* ENERGY SOURCES; MINING.

Shale oil. Shale oil is produced from the organic matter in oil shale when the rock is heated in the absence of oxygen (destructive distillation). This heating process is called retorting, and the equipment that is used to do the heating is known as a retort. The rate at which the oil is produced depends upon the temperature at which the shale is retorted. Most references report retorting temperatures as being about 500°C (930°F).

Retorting temperature affects the nature of the shale oil produced. Low retorting temperatures produce oils in which the paraffin content is greater than the olefin contents; intermediate temperatures produce oils that are more olefinic; and high temperatures produce oils that are nearly completely aromatic, with little olefin or saturate content.

In general, shale oils can be refined to marketable products in modern petroleum refineries. There is no really typical shale oil produced from Green River oil shale, but the oils do share many properties in common. They usually show high pour points, 20–32°C

(68–90°F); high nitrogen contents, 1.6–2.2 wt %; and moderate sulfur contents, about 0.5 wt %. High pour points make necessary some processing before the oils are amenable to pipeline transportation. The high nitrogen contents make hydrogenation necessary to reduce the nitrogen contents so that the oils can be processed into fuels. Hydrogenation also reduces the sulfur content. *See* DESTRUCTIVE DISTILLATION; HYDROGENATION.

United States technology. The two general approaches to recovering shale oil from Green River Formation oil shales are (1) mining, crushing, and aboveground retorting, called conventional processing; and (2) in-place processing. The basic problems facing conventional processing are handling and heating huge amounts of low-grade ore and disposing of huge volumes of spent shale. The in-place approach largely avoids the problems of handling and disposal but faces a different basic problem—the impermeability of the oil shale beds.

With the conventional approach, oil shale mining by the room-and-pillar technique developed by the U.S. Bureau of Mines appears capable of producing the huge amounts of ore necessary to operate a large production plant. The procedure has also been tested by industry in Mahogany Zone shales. Outputs on the order of 2500 tons (2250 metric tons) of oil shale per worker shift have been reported from the highly mechanized operation. Crushing technology is well demonstrated.

Retorting must be done continuously in order to reach the throughput necessary for economic production of shale oil. Two general systems for heating a continuous stream of oil shale are outlined in **Fig. 3**. In the internally heated system the oil shale furnishes its own heat because part of its organic matter is burned inside the retort. Examples of the internally fired retort system include: the Bureau of Mines gas combustion retort; one form of the Paraho retort of Development Engineering, Inc.; one form of the Union Oil Company's rock-pump retort; and Superior Oil Company's moving-grate retort. The TOSCO Corporation's TOSCO II retort, in which preheated ceramic balls heat the oil shale; the SGR rock-pump retort of Union Oil; one form of the Paraho retort and the Petrosix pilot retort, in which preheated gas heats the oil shale; and the Lurgi retort, in which preheated sand or spent shale heats the oil shale, are all examples of the externally fired retort system. Several retorting systems, including both internally and externally heated designs, have been tested on pilot or semiworks scales.

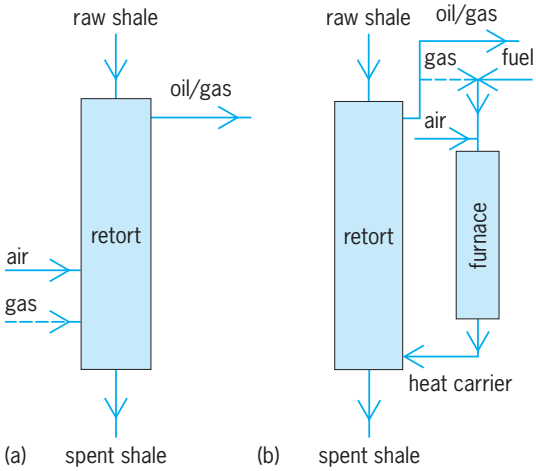

Fig. 3. Oil shale retorting systems: (*a*) internally heated, (*b*) externally heated.

Spent shale disposal has been studied intensively. More than 80% of the mined oil shale remains as residue after oil production. An oil shale plant which produces 50,000 bbl of oil per day from Mahogany Zone shale might mine 75,000 tons (68,000 metric tons) of rock and dispose of 60,000 tons (54,000 metric tons) of spent shale daily. In the vast and largely unpopulated areas of the Green River Formation, dumping these volumes of spent shale is not as large an environmental problem as it appears to be. The spent shale resulting from most surface retorts is largely insoluble, consisting largely of the same native minerals being exposed in huge quantities to natural weathering by the existing extensive erosion of the Green River Formation. Contoured dumping to control water flow will minimize leaching, already low in an arid region. Native vegetation will establish itself on spent shale dumps, and revegetation procedures can accelerate this process. Returning spent shale to the mine to furnish roof support may permit recovery of additional ore, but this approach has not been tested.

Research and development efforts toward in-place processing have concentrated on creating permeability in the impermeable oil shale. In-place processing may be accomplished by two means: (1) a borehole technique in which oil shale is first fractured underground and heat is applied, and (2) a process in which some rock is first removed by mining, then the remaining oil shale is fragmented into the voids created by mining, and finally heat is applied. These two methods are referred to as in-place (no mining) and modified in-place (some mining) processing. Several investigators have field-tested in-place methods. Both of the heating methods outlined for retorts (Fig. 3) have also been applied to in-place processing. Proposals for in-place processing include solvent extraction and radio-frequency electrical heating. *See* SOLVENT EXTRACTION.

Only one-third of the total Green River Formation oil shale resource is considered minable for conventional processing. In-place processing may make much of the remainder of the resource available to production. Conventional processing offers process control advantages, including ready adaptation of many existing industrial procedures, but it is capital-intensive, requiring huge investments before production begins. In-place processing is less easily controlled and evaluated but requires less capital outlay before production begins. *See* ENERGY SOURCES; MINING.

John Ward Smith; Howard B. Jensen

Bibliography. V D. Allred (ed.), *Oil Shale Processing*, 1982; American Society for Testing and Materials, *Standard Test Method for Oil from Oil*

Shale, Procedure D3904–80, 1980; K. P. Chong and J. W. Smith, *Mechanics of Oil Shale*, 1984; Colorado School of Mines, *1st through 17th Oil Shale Symposium Proceedings,* 1963–1984; B. Durand (ed.), *Kerogen: Insoluble Organic Matter from Sedimentary Rocks,* Technip, Paris, 1980; R. M. Garrels and C. L. Christ, *Solutions, Minerals, and Equilibria,* 1965, reprint 1985; R. A. Meyers (ed.), *Handbook of Synfuels Technologies,* 1984; F. P. Miknis and J. F. McKay (eds.), *Geochemistry and Chemistry of Oil Shales,* ACS Symp. Ser. 230, 1983; *Oil Shale: Prospects and Constraints,* Federal Energy Administration Project Independence Blueprint, GPO 4118–00016, 1974; L. C. Ruedisili and M. W. Firebaugh (eds.), *Perspectives on Energy,* 3d ed., 1982; J. W. Smith, *Oil Shale Resources of the United States,* Colorado School of Mines Publications, 1981; J. W. Smith, *Theoretical Relationship Between Density and Oil Yield for Oil Shales,* U. S. Bur. Mines Rep. Inv. 7179, 1969.

Oil field waters

Waters of varying mineral content which are found associated with petroleum and natural gas or have been encountered in the search for oil and gas. They are also called oilfield brines, or brines. They include a variety of underground waters, usually deeply buried, and have a relatively high content of dissolved mineral matter. These waters may be (1) present in the pore space of the reservoir rock with the oil or gas, (2) separated by gravity from the oil or gas and thus lying below it, (3) at the edge of the oil or gas accumulation, or (4) in rock formations which are barren of oil and gas. Brines are commonly defined as water containing high concentrations of dissolved salts. Potable or fresh waters usually are not considered oilfield waters but may be encountered, generally at shallow depths, in areas where oil and gas are produced.

Oilfield waters or brines differ widely in composition and concentration. They may differ from one geologic province to another, from one formation to another within a given geologic province, or from one part of a specific geologic horizon to another. They range from slightly salty water with 1000–3000 parts of dissolved substances in 1 million parts of solution to very nearly saturated brines with dissolved mineral content of more than 270,000 parts per million (ppm).

The most common and abundant mineral found in oilfield waters is sodium chloride, or common table salt. Calcium chloride is next in order of abundance. Carbonates, bicarbonates, sulfates, and the chlorides of magnesium and potassium are present in lesser quantities. In addition to the above mentioned salts, salts of bromine and iodine are also found. Traces of strontium, boron, copper, manganese, silver, tin, vanadium, and iron have been reported. Barium has been reported in many of the Paleozoic brines of the Appalachian region. The commercial value of a brine depends upon the concentration of salts, purity of the products to be recovered, and value and practicability of by-product recovery. Concentrations less than 200,000 ppm are seldom of commercial interest.

Slightly salty waters, while not suitable for human consumption, may be used in some industrial processes or may be amenable to beneficiation for municipal supplies in areas lacking fresh waters.

Classified genetically, oilfield waters are generally considered connate; that is, they are seawaters which (presumably) originally filled the pore spaces of the rock in which they are now confined. However, few analyses of these waters correspond to present-day seawater, thus indicating some mixing and modification since confinement. Dilute solutions suggest that rainwater has percolated into the rocks along bedding planes, fractures, faults, and other permeable zones. Presence of carbonates, bicarbonates, and sulfates in an oilfield water further suggests that at least some of the water had its origin at the surface. Concentrations of dissolved solids greater than that of modern seawater suggest partial evaporation of the water or addition of soluble salts from the adjacent or enclosing rocks.

Waters in most sedimentary rocks increase in mineral concentration with depth. This increase may be due to the fact that, since salt water is heavier than fresh water, the more dense solution will eventually find a position as low as possible in the aquifer. An additional factor would be the longer exposure of the deeper waters to the mineral-bearing rocks. Exceptions have been noted and probably are due to the presence of larger quantities of soluble salts in some geological formations than in others.

Probably the most important geological use of oilfield water analyses is their application to the quantitative interpretation of electrical and neutron well logs, particularly micrologs. In order to compute the connate water saturation of a formation in a quantitative manner from electrical data, it is necessary to know with accuracy the connate water resistivity.

Naturally mineralized waters are frequently the only waters available for water-flooding operations. Water analyses are useful in predicting the effect of the water on minerals in the reservoir rock and on the mechanical equipment employed on the project. Waters which exert a corrosive action on the lines and pumps or which tend to plug up the pay zone are not suitable for water-flooding operations.

Oil field water composition may be an important factor in the determination of the source of water in oil wells which have leaky casings or improper completions with resulting communication between wells, and in identifying and correlating reservoirs in multipay oil pools, particularly in those containing lenticular sand bodies.

Industrial wastes, including mineralized water produced with oil, may be disposed of in underground reservoirs. Between the zone of potable water and the horizon of commercial brines, there commonly are rock formations, the waters of which contain chemicals in amounts sufficient to make the waters unsuitable for domestic, municipal, industrial, and livestock consumption, but not in sufficient

quantity to be considered as a source for recovery of chemicals. Provided there is sufficient porosity and permeability, these rock formations could receive industrial wastes which would otherwise contaminate surface streams and shallow, fresh groundwater horizons into which they might be discharged. *See* GEOPHYSICAL EXPLORATION; PETROLEUM GEOLOGY.

Preston McGrain

Bibliography. A. G. Collins, *Geochemistry of Oilfield Waters (Developments in Petroleum Science)*, 1975; J. Hunt, *Petroleum Geochemistry and Geology*, 2d ed., 1995.

Okra

A warm-season annual, *Hibiscus esculentus*, of Ethiopian origin. Okra, also called gumbo, is grown for its immature pods (see **illus.**), which are

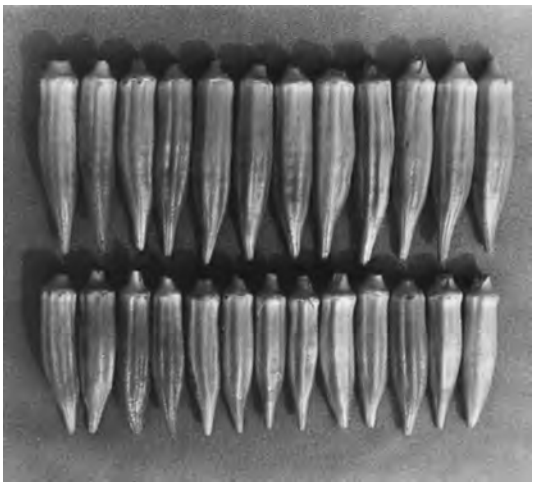

Okra pods. (*Asgrow Seed Co., subsidiary of The Upjohn Co.*)

generally used for preparing soups but are also eaten as a freshly cooked vegetable. It is a member of the order Malvales and is related to cotton. Propagation is by seed. Popular varieties are Clemson Spineless and Green Velvet. Okra is sensitive to low temperatures; commercial production in the United States is primarily in the South. Harvesting begins when the pods are 3–4 in. (7.5–10 cm) long, usually 50–60 days after planting. Georgia, Florida, and Louisiana are important producing states. *See* COTTON; MALVALES.

H. John Carew

Olbers' paradox

The riddle of why is the sky is dark at night. This celebrated riddle originated in the sixteenth century. In 1823, Wilhelm Olbers presented it in the simplest terms: In an infinite universe, populated everywhere with stars, a line of sight in any direction when extended out into space must ultimately intercept the surface of a star. Hence stars should cover the entire sky. And if all stars are sunlike, the sky at every point should blaze as brightly as the disk of the Sun. Olbers has been credited incorrectly with the discovery of the riddle; he did, however, express it in this lucid form, and showed that the riddle still holds even when stars are irregularly distributed in clusters.

Nature of the riddle. A boundless universe of either flat or curved space, in which luminous stars stretch away endlessly, seems a not unreasonable cosmological picture. Yet it leads to a conclusion in total contradiction with experience. The sky at night is dark and not a continuous blaze of bright stars. A calculation, first made in 1744 by Jean-Philippe Loys de Chéseaux, showed that starlight should be 90,000 times brighter than sunlight at the Earth's surface. Chéseaux showed that in an infinite universe, uniformly populated with stars, only a finite number of stars are seen. Visible stars cover the sky out to a distance (the "background distance") where they fuse together and form a continuous background; the stars at greater distances are covered over by the foreground stars and cannot be seen.

A forest analogy helps to understand the riddle. Every (horizontal) line of sight in an endless forest terminates at a tree trunk. Wherever we stand we find ourselves surrounded by trees stretching away into a continuous background. The distance of the background is the average area occupied by a tree and its surroundings halfway to the nearest neighboring trees divided by the typical thickness of a tree trunk at eye level. (For example, the background distance is 50 m if trees are separated from one another by distances of 5 m and have trunk diameters of 0.5 m.)

Solutions. More than a dozen solutions have been proposed since the riddle was first discovered by Thomas Digges in 1576. The riddle has two interpretations, and the proposed solutions are therefore of two kinds.

Missing star light. The first interpretation assumes that the argument is correct and the sky is indeed covered with stars, most of which cannot be seen. The riddle then asks: What happened to the missing starlight? Both Chéseaux and Olbers adopted this interpretation and proposed that starlight is slowly absorbed as it travels large distances in space. The analogy in this case is the foggy forest; only foreground trees are seen and fog obscures background trees. This solution fails because the interstellar absorbing medium quickly heats up and then emits the absorbed radiation. Hermann Bondi in the 1950s also adopted this first interpretation and argued that the light from distant stars is redshifted into invisibility by the expansion of the universe. The redshift solution contributed greatly to the popularity of cosmology, and for several years it was believed that darkness at night proved that the universe is expanding. But Bondi's redshift solution is incorrect in a big bang universe and applies only in the steady-state expanding universe that was disproved in 1965 by the discovery of the radiation from the big bang. *See* BIG BANG THEORY; COSMIC BACKGROUND RADIATION; HEAT RADIATION.

Missing stars. The second interpretation assumes that the sky is not covered with stars, in agreement with observation, and therefore the assertion that every line of sight intercepts the surface of a star is misleading. The riddle, in effect, now asks: What do lines of sight intercept when we look at the dark gaps between stars? Johannes Kepler in the early seventeenth century adopted this second interpretation and argued that we look out between a finite number of stars and see a dark surface enclosing the universe. The analogy is a finite forest enclosed by a high wall. Another solution, consistent with this interpretation and popular with as tronomers in the late nineteenth and early twentieth centuries, was the argument that our Galaxy—the Milky Way—is the only galaxy in the universe, and we look out between the stars to an empty, dark, and infinite space beyond. The corresponding analogy is a finite forest, and we look out between the trees to a treeless plain beyond. Edgar Allen Poe suggested in 1848 that the universe is not old enough for the light from very distant stars to have reached us; hence lines of sight that fail to intercept stars reach back to the darkness that prevailed before the birth of stars. This was investigated by Lord Kelvin in 1901. He assumed that stars are no older than 100 million years, and showed that out to a distance traveled by light in 100 million years, visible stars are insufficient in number to cover the whole sky. The missing stars are actually out there, but we look out in space and back in time to a prenatal era before the stars were born.

Modern calculations confirm and extend Kelvin's conclusions, and the second interpretation of Olbers' paradox is hence correct: Most lines of sight do not intercept stars but extend back to the beginning of the universe. Light travels at approximately 300,000 km/s (186,000 mi/s), and a static universe $10–20 \times 10^9$ years old is not old enough for starlight to reach the Earth from regions sufficiently distant for visible stars to cover the sky. If the sky at night is dark in a static universe, then obviously in an expanding universe of the same age the night sky is even darker because of the redshift of starlight. Calculation shows that the redshift effect is in fact relatively small.

Big bang universe. Olbers' paradox, night has been solved by modren cosmology. In the big bang universe, $10–20 \times 10^9$ years old, we cannot see sufficient stars to cover the sky. Instead, on looking out in space, we look far back in time to the beginning of the universe and see in all directions the big bang covering the sky. The expansion of the universe has reduced the incandescent glare of the big bang and redshifted its radiation into the invisible infrared. *See* COSMOLOGY; UNIVERSE. Edward Harrison

Bibliography. H. Bondi, *Comology*, 2d. ed., Cambridge University Press, 1960; E. R. Harrison, *Cosmology: The Science of the Universe*, 2d. ed., New Cambridge University Press, 2001; E. R. Harrison, *Darkness at Night: A Riddle of the Universe*, Harvard University Press, Cambridge, MA, 1987.

Olfaction

One of the chemical senses, specifically the sense of smell. Olfaction registers chemical information in organisms ranging from insects to humans, including marine organisms. For terrestrial animals, its stimuli comprise airborne molecules. The typical stimulus is an organic chemical with molecular weight below 300 daltons; about a half million such substances exist. A few inorganic chemicals can also stimulate olfaction, notably hydrogen sulfide, ozone, ammonia, and the halogens. For marine organisms, amino acids and proteins, which derive largely from the decomposition of organic matter, form particularly good stimuli.

Anatomy and physiology. The anatomy of olfactory structures and the neurophysiology of olfaction differ significantly among different animal groups.

Insects. Insect olfactory receptors exist within sensory hairs on the antennae. Each hair contains thousands of pores in its cuticle. Stimulating molecules must pass through the pores in order to stimulate the outer segment of a receptor cell. This process of transduction has been studied through the measurement of electric potentials from the receptor neurons. The best known potential, the electroantennogram, reflects the action of many receptor neurons.

Certain insect olfactory receptors are specialists, and respond to a very narrow range of relevant chemicals, such as sex attractants secreted by females but registered only by males. Generalist receptors, on the other hand, respond to a wide range of chemicals. One generalist cell differs somewhat from another in the variety of materials to which it responds, permitting discrimination of different materials in an across-receptor pattern of induced neural activity. *See* CHEMICAL ECOLOGY; INSECT PHYSIOLOGY.

Fishes. The olfactory organ of fishes resides typically in tubular chambers on either side of the mouth (**Fig. 1**). As the fish swims, water passes through the chamber and comes into contact with the olfactory epithelium that lines the folds of a structure called

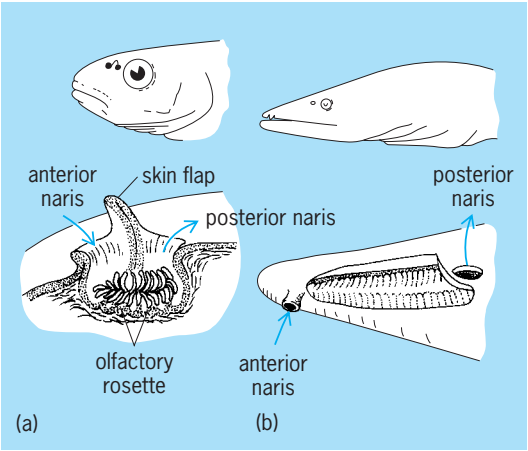

Fig. 1. Position and internal structure of the nose in fishes. (*a*) Minnow (*Phoxinus phoxinus*). (*b*) Eel (*Anguilla anguilla*). (*After H. Teichmann, Was leistet der Gervchssinn bei Fischen?, Umschau Wiss. Tech., 62:588–591, 1962*)

the olfactory rosette. This epithelium has great similarity to that of terrestrial vertebrates. In addition to proteins and amino acids, fishes respond to many of the odorants that form the stimuli for terrestrial olfaction. Hence, fishes respond to stimuli that humans characterize as floral, fruity, putrid, and so on.

Terrestrial vertebrates. In terrestrial vertebrates, the olfactory receptors reside within a sac or cavity more or less similar to the human nasal cavity (**Fig. 2**). The olfactory mucosa patch in the cavity characteristically contains millions of receptor cells, though in some olfactory-dominated (or macrosmatic) mammals, such as the dog and rabbit, it contains tens of millions. The location of the olfactory mucosa relative to air currents in the cavity plays some role in the ongoing olfactory vigilance of the organism. In the human, a microsmatic organism, the mucosa sits out of the main airstream. During quiet breathing eddy currents may carry just enough stimulus to evoke a sensation, whereupon sniffing will occur. Sniffing amplifies the amount of stimulus reaching the receptors by as much astenfold.

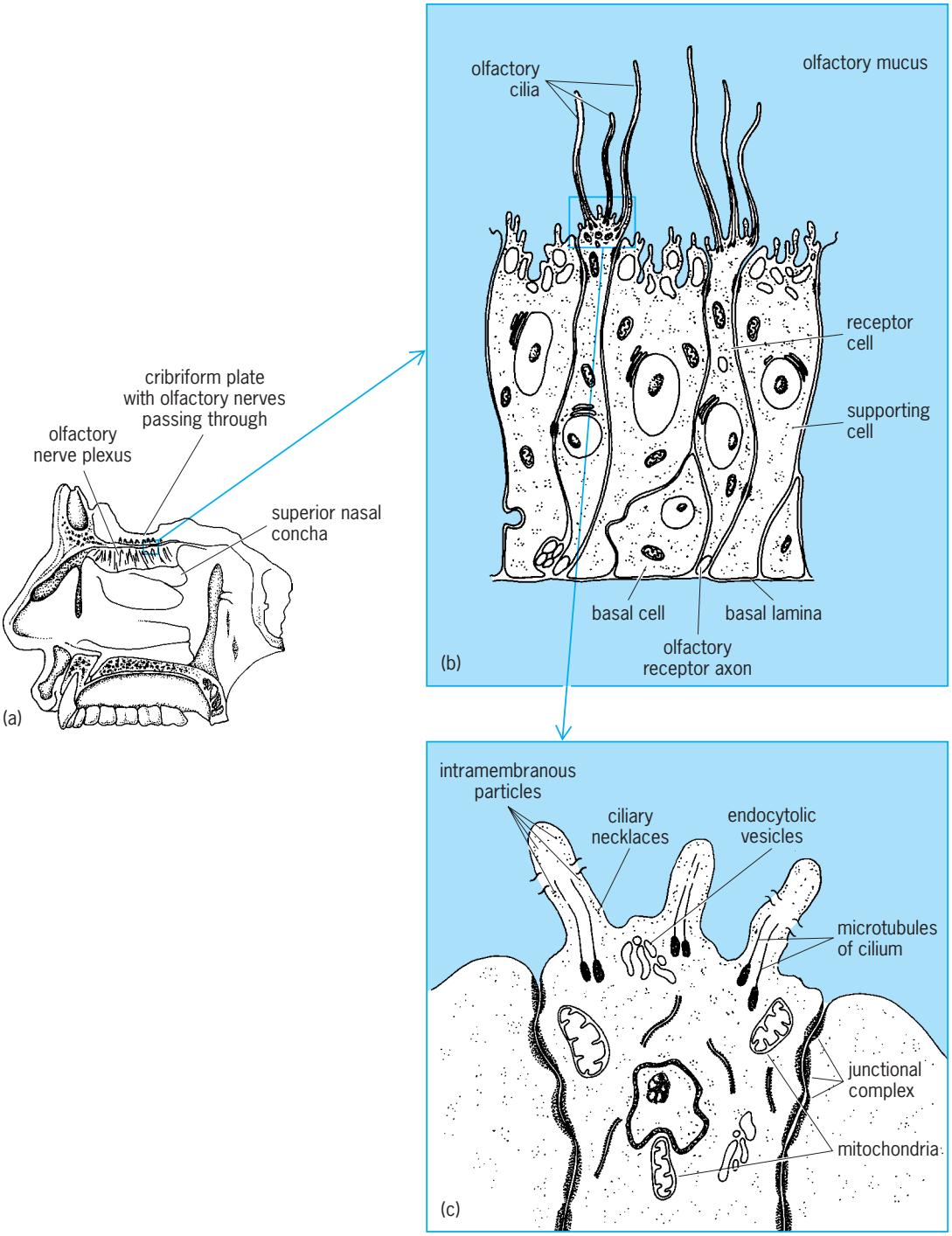

Fig. 2. Human peripheral olfactory system. (*a*) Position of olfactory epithelium. (*b*) Olfactory epithelium. (*c*) Olfactory receptor. (*After H. B. Barlow and J. D. Mollon, eds., The Senses, Cambridge University Press, 1982*)

Reception of the chemical stimulus and transduction into a neural signal apparently occur on the olfactory receptor cilia (Fig. 2). The ciliary membrane contains receptor protein molecules that interact with stimulating molecules through reversible binding. Nonprotein portions of the membrane may also have some chemoreceptive role, through modulation of the conformation or orientation of the receptor proteins.

Vertebrate receptor cells, like the generalist cells of insects, show broad tuning, that is, they respond to many odorants. The breadth of tuning depends in part on cell age—receptor cells have a life-span of just a few weeks, and show broader tuning in their first two to three weeks than thereafter. Since new cells emerge from basal cells beneath the receptor cells, a receptor cell of any age may appear anywhere in the olfactory mucosa. However, there do exist some regional differences in the tuning. In frogs and the tiger salamander, organisms commonly used to study vertebrate olfaction, cells in the anterior portion of the mucosa show a bias toward water-soluble odorants, whereas those in the posterior portion show a bias toward lipid-soluble odorants. An anterior-posterior gradation, seen easily in regional recordings of the receptor potential known as the electroolfactogram, also reveals itself in recordings from the olfactory bulb of the brain.

Adjacent points in the mucosa project generally, though not exclusively, to adjacent points in the olfactory bulb (**Fig. 3**). The synapses between the incoming olfactory nerve fibers and the second-order cells, mitral cells, occur in basketlike structures called glomeruli. On average, a glomerulus receives about 1000 receptor cell fibers for each mitral cell. The tangled arrangement of fibers within a glomerulus invites the conclusion that it must operate as a functional unit. Second-order cells, which are also influenced by collateral neurons, such as the interior granule cell, nevertheless show very similar breadth of tuning to that seen in individual receptor cells. The location of cells within the bulb seems to play a role in encoding odor quality: each odorant stimulates a more or less unique spatial array.

Second-order cells, unlike first-order cells, may change their responsiveness with the state of need of the organism. For instance, mitral cells will respond more vigorously to food odors in a hungry organism. Such modulation presumably results from interplay of incoming neural signals with outgoing activity which originates in more central structures.

The central neural pathways of the olfactory system have a complexity unmatched among the sensory systems. One pathway carries information to the pyriform cortex (paleocortex of the temporal lobe), to a sensory relay in the thalamus (dorsomedial nucleus), and to the frontal cortex (orbitofrontal region) [**Fig. 4**]. This pathway seems rather strictly sensory. Another pathway carries information to the pyriform cortex, the hypothalamus, and other structures of the limbic system. The latter have much to do with the control of emotions, feeding, and sex. The strong affective and motivational consequences of olfactory stimulation seem compatible with projections to the limbic system and with the role of olfaction in certain types of physiological regulation.

In many organisms, odors from other members of the same species have powerful motivational effects. Such stimuli are called pheromones. Sex attractants, trail pheromones, and maternal pheromones are examples. In many vertebrate species, reception of pheromones occurs via an important accessory olfactory organ, known as the vomeronasal organ, which characteristically resides in the hard palate of the mouth or floor of the nasal cavity (Fig. 4). *See* PHEROMONE.

Sensitivity and functional properties. At its best, human absolute sensitivity exceeds that of the most sensitive physical instruments, but sensitivity varies from odorant toodorant over several orders of

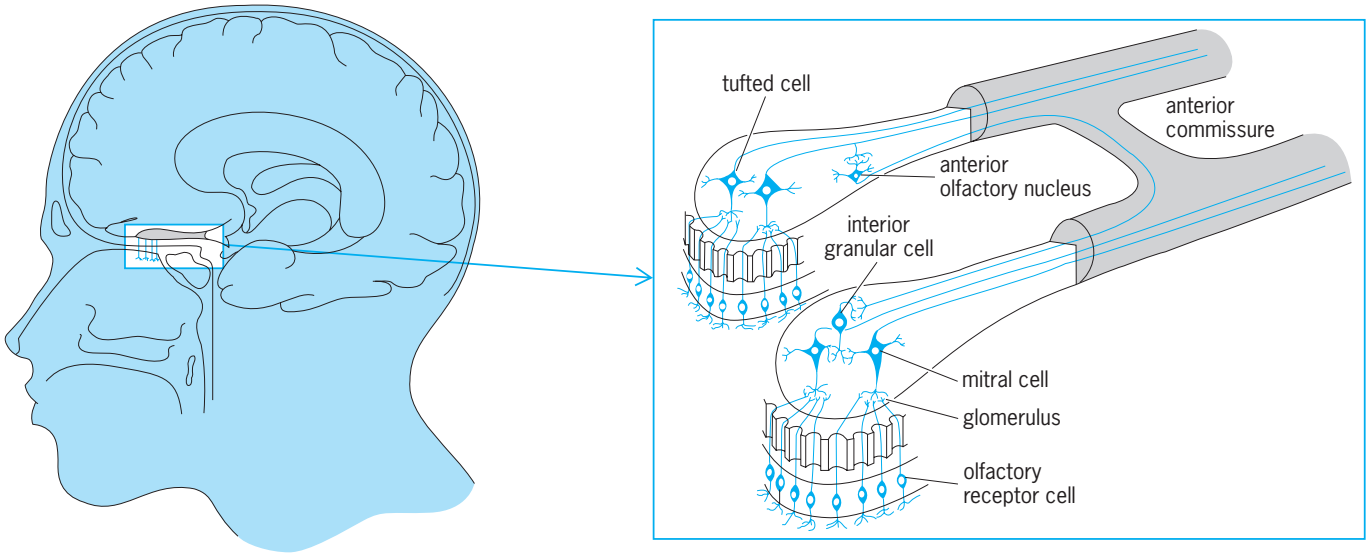

Fig. 3. Location of the olfactory bulbs at the interior surface of the human brain and their connections via the anterior commissure. (*After D. Ottoson, Physiology of the Nervous System, Oxford University Press, 1983*)

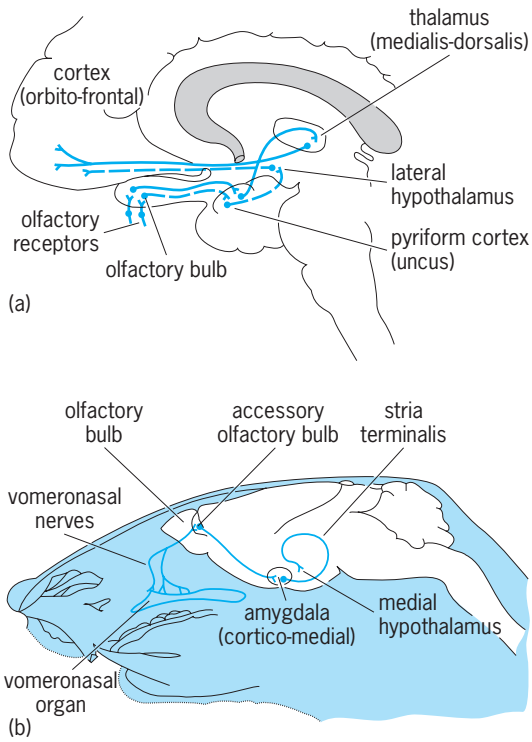

Fig. 4. Olfactory neural pathways in terrestrial vertebrates. (*a*) Central olfactory pathways characteristic of higher primates. Broken and solid lines represent distinct pathways. (*b*) Accessory olfactory pathways characteristic of rodents. (*After H. B. Barlow and J. D. Mollon, eds., The Senses, Cambridge University Press, 1982*)

magnitude. For instance, the threshold for the potent odorant 3-methoxy-3-isobutyl pyrazine (green bell pepper odor) equals about 1 part per 10^{12} parts of air, whereas that for the far less effective odorant methanol (wood alcohol) equals about 100 parts per 10^6. A common range of thresholds for materials used in fragrances and flavors is 1 to 100 parts per 10^9. Physicochemical determinants of molecular deposition and penetration through mucus correlate rather well with this threshold range. In aliphatic series in which physicochemical properties vary progressively with molecular chain length, threshold characteristically changes progressively. Such a correlation holds true for terrestrial vertebrate species other than humans, though absolute values differ. The available data suggest that rats and dogs have best sensitivity, about a thousandfold better than that of humans.

Thresholds gathered from various groups of human subjects permit certain generalities about how the state of the organism affects olfaction. For instance, persons aged 70 and above are about tenfold less sensitive than young adults. Males and females have about equal sensitivity, except perhaps in old age, where females are more sensitive. Females show a small fluctuation in sensitivity during the menstrual cycle. Sensitivity increases prior to ovulation. Smokers have about the same sensitivity as nonsmokers. Persons with certain medical disorders, such as multiple sclerosis, Parkinson's disease, paranasal sinus disease, Kallmann's syndrome, and olfactory tumors, exhibit decreased sensitivity (hyposmia) or complete absence of sensitivity (anosmia).

Above its threshold, the perceived magnitude of an odor changes by relatively small amounts as concentration increases. A tenfold increment in concentration will cause, on average, about a twofold change in perceived magnitude. The perceived magnitude of an odor is often greatly influenced by olfactory adaptation, a process whereby during continuous short-term exposure to a stimulus its perceived magnitude falls to about one-third of its initial value. Concentrations below this conditioning level may then prove imperceptible, but concentrations above the conditioning level will be affected relatively little. Recovery from such adaptation begins as soon as the stimulus is removed, and is generally complete in a few minutes.

Adaptation is largely specific to the conditioning odorant. That is, exposure-induced changes in sensitivity to one odorant will leave sensitivity to most other odorants roughly unchanged. Instances where exposure to one odorant does alter the sensitivity to another exemplify cross-adaptation. Experiments on cross-adaptation serve as a means to explore chemical or structural commonalities among stimuli.

The stimuli for olfaction are commonly complex, that is, they are mixtures. Such products as coffee, wine, cigarettes, and perfumes contain at least hundreds of odor-relevant constituents. Only rarely does the distinctive quality of a natural product, such as a vegetable, arise from only a single constituent. A chemical analysis of most products will not usually allow a simple prediction of odor intensity or quality. One general rule, however, is that the perceived intensity of the mixture falls well below the sum of the intensities of the unmixed components. This suppression of intensity works against the weaker components, which may lose their separate identity in the mixture but may still contribute to overall perceptual quality. A pleasant-smelling mixture of unknown composition may actually be found upon analysis to contain some very unpleasant constituents in small amounts. Their presence is often partially masked by stronger-smelling constituents, yet may make an essential contribution. Perfumery, which plays a wide role in the "deodorization" (often through addition of odor) and reodorization of many manufactured products, entails a careful blending of pleasant- and sometimes unpleasant-smelling materials into a final acceptable odor.

Structure-activity relations. General notions about the properties that endow a molecule with its quality have spawned more than two dozen theories of olfaction, including various chemical and vibrational theories. Most modern theories hold that the key to quality lies in the size and shape of molecules, with some influence of chemical functionality. For molecules below about 100 daltons, functional group has obvious importance: for example, thiols smell skunky, esters fruity, amines fishy-uriny, and carboxylic acids

rancid. For larger molecules, the size and shape of the molecule seem more important. Shape detection is subtle enough to enable easy discrimination of some optical isomers (for example, D-carvone versus L-carvone, which smell like spearmint and caraway, respectively). Progressive changes in molecular architecture along one or another dimension often lead to large changes in odor quality. No current theory makes testable predictions about such changes. *See* CHEMICAL SENSES; CHEMORECEPTION.

William Cain

Bibliography. W. Breipohl (ed.), *Ontogeny of Olfaction*, 1986; T. Engen, *Odor Sensation and Memory*, 1992; A. I. Farbman, *Cell Biology of Olfaction*, 1992; T. L. Payne, M. C. Birch, and C. E. Kennedy (eds.), *Mechanisms in Insect Olfaction*, 1986; M. J. Serby and K. L. Chober (eds.), *The Science of Olfaction*, 1992; E. T. Theimer (ed.), *Fragrance Chemistry: The Science of the Sense of Smell*, 1982.

Oligocene

The third oldest of the seven geological epochs of the Cenozoic Era. It corresponds to an interval of geological time (and rocks deposited during that time) from the close of the Eocene Epoch to the beginning of the Miocene Epoch. The most recent geological time scales assign an age of 34 to 24 million years before present (m.y. B.P.) to the Oligocene Epoch. *See* CENOZOIC; EOCENE; MIOCENE.

In his early subdivision of the Tertiary published in 1833, Charles Lyell established four stratigraphic units, from the oldest to the youngest: Eocene, Miocene, and Older and Younger Pliocene. Stratigraphers in the Netherlands and Germany, however, kept describing strata that they considered to be intermediate in position and characteristics between Eocene and Miocene. They concluded that these sediments represented a major marine transgression in northern Europe at the close of the Eocene Epoch. This eventually led the German stratigrapher E. Beyrich in 1854 to propose the Oligocene as an independent subdivision of the Tertiary based on a sequence of marine, brackish-water, and nonmarine sediments of the Mainz Basin in Germany. He proposed a new Oligocene Epoch and constructed it out of the younger part of Eocene and the older part of Miocene epochs of the Lyellian subdivisions of Tertiary.

An important event that characterizes the Oligocene Epoch was the development of extensive glaciation on the continent of Antarctica. Prior to that time, the world was largely ice-free through much of the Mesozoic and early Tertiary. A significant amount of ice is now known to have existed on the Antarctic continent since at least the beginning of the Oligocene, when the Earth was ushered into its most recent phase of ice-house conditions. This in turn created revolutions in the global climatic and

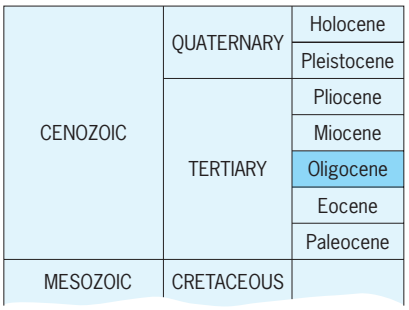

hydrographic systems, with important repercussions for the marine and terrestrial biota. The changes include steepened latitudinal and vertical thermal gradients affecting major fluctuations in global climates, and the shift in the route of global dispersal of marine biota from an ancestral equatorial Tethys seaway, which had become severely restricted by Oligocene time, to the newly initiated circum-Antarctic circulation. *See* GLACIAL EPOCH.

Subdivisions. In modern time scales, this epoch is subdivided into two series, a Lower and an Upper Oligocene. Northern European Rupelian and Chattian stratigraphic stages are designated to be time-equivalent to the Early and Late Oligocene, respectively (**Fig. 1**). Worldwide, the epoch represents an overall regressive sequence when there was a drawdown of global sea level, with relatively deeper, marine facies in the Early Oligocene and shallower-water to nonmarine facies in the Late Oligocene. *See* FACIES (GEOLOGY).

The Rupelian Stage was proposed by A. Dumont in 1849, based on sedimentary strata in Belgium. The Tongrian Stage, also described from Belgium by Dumont in 1839, was originally regarded to be, in part, older than the Rupelian by Beyrich, who considered that the Tongrian and Rupelian together

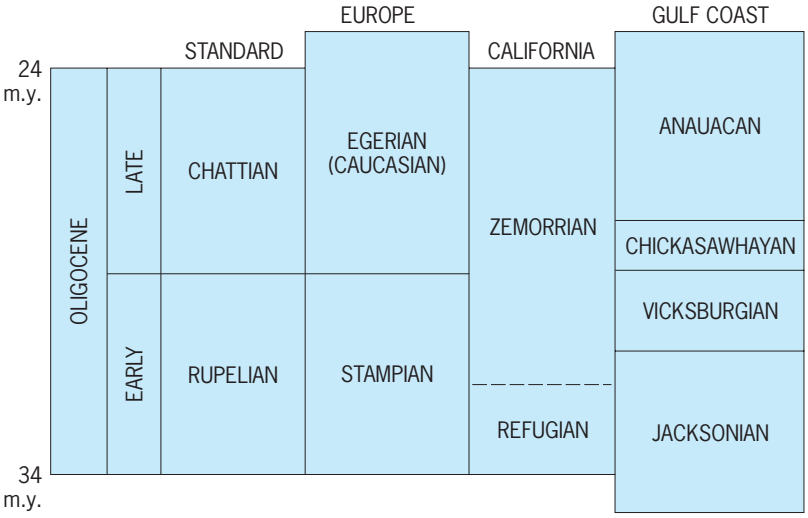

Fig. 1. Oligocene stages and their temporal equivalents.

constituted his new epochal unit, the Oligocene. In 1983 K. Mayer-Eymar erected the Lattorfian Stage with a type locality in Saxony, Germany. Mayer-Eymar considered the Lattorfian to be temporally equivalent to the Early Oligocene. Both the Tongrian and Lattorfian strata were later shown to range downward into the uppermost Eocene, and their use has largely been abandoned in favor of an expanded notion of the Rupelian. Because the Rupelian at its type section spans somewhat less than the time interval included in the Early Oligocene, some French stratigraphers prefer the Stampian Stage to Rupelian. The Stampian was described by A. D'Orbigny in 1852 from marine and lacustrine beds at Etampes in France.

The Chattian, elected by consensus as the standard stage for the Late Oligocene, was based on marine sand beds near Kassel, Germany, and proposed by T. Fuchs in 1894. Fuchs considered the Chattian to be younger than Rupelian and older than the Aquitanian Stage (which was later placed in the Lower Miocene). Later studies have shown that there is a temporal gap between the top of Chattian and the base of Aquitanian as defined at their stratotype sections. However, in stratigraphy it is normal practice to extend the formal concept of standard stages to include such gaps. Thus, to accommodate the gap between the uppermost Eocene stage of Priabonian and the Oligocene Rupelian, and the gap between Chattian and Aquitanian, the formal concepts of Rupelian and Aquitanian were extended downward by later stratigraphers to bridge the intervening lacunae.

Other regional stages of the Oligocene include the Egerian, sometimes used in central and eastern European countries, which is in part equivalent to the Chattian, but may extend into the lower part of Miocene. In Russia the Caucasian Stage is considered to be an equivalent to the Egerian. Marine Oligocene in California is subdivided into the Refugian and Zemorrian stages, the former stage encompassing only the earliest Oligocene, and the latter spanning the bulk of the epoch. In the Gulf coast of the United States the Oligocene extends from the upper part of the Jacksonian Stage to Vicksburgian, Chickasawhayan, and lower part of Anahuacan stages. Temporally, the latter stage extends into the Miocene. Australian stratigraphers include the Willungian, Janjukian, and lower part of Longfordian stages in their Oligocene, and the New Zealanders now consider the epoch to span three local stages: Whaingaroan, Duntroonian, and Waitakian.

Paleontologists who study mammal and other vertebrate fossils often use their own subdivisions to express the ages of terrestrial assemblages. Workers in Europe consider the Oligocene to span three subdivisions, Suevian, Arvernian, and lower part of Agenian. North American vertebrate paleontologists include Orellan, Whiteneyan, and the lower two-thirds of the Arikareean ages in their concept of the Oligocene. *See* PALEONTOLOGY; STRATIGRAPHY.

Tectonics, oceans, and climate. Radical changes occurred in the Oligocene Epoch that revolutionized global oceanographic and climatic conditions. The most prominent change was the shift of circumglobal circulation (and a major means of biotic dispersal) from the equatorial to the southern high-latitude regions. Both the restriction of the equatorial flow between the Indian and the Atlantic oceans through the ancestral Tethys seaway, and the opening of the oceanic gateway at Drake Passage occurred during the Oligocene. These changes had important repercussions that led to the entry of the Earth into predominantly ice-house conditions that continue to the present. The cooling of the polar regions in the Oligocene led to the accentuation of latitudinal and vertical thermal gradients and an increase in seasonality. This resulted in a fundamental shift in the bottom-water regime, from density-driven warm, saline bottom waters in the Cretaceous and Paleocene-Eocene, to cold bottom waters, largely driven by thermal contrast. The accentuated latitudinal thermal contrast also resulted in the expansion of the erosive activity of bottom waters. *See* CRETACEOUS.

The separation of Svalbard and Greenland in the latest Eocene and the breaching of the Rio Grande Rise in the South Atlantic in the earliest Oligocene set the stage for the crossing of the important climatic threshold near the Eocene-Oligocene boundary. It has been suggested that after the development of the connection between the Arctic and the Norwegian Sea that allowed the supply of cold deep water to the Atlantic and the southern high latitudes, more moisture became available around Antarctica. This, combined with the partial isolation of Antarctica, may have led to large-scale freezing at sea level and the initiation of the formation of Antarctic Bottom Water. There is evidence that the temperature of the bottom water dropped by 4–5°C (7–9°F) near this boundary, and the change from warm to cold bottom water may date back to this time. The development of the psychrosphere (deeper, cold layer of the ocean) manifests itself in the presence of widespread erosional hiatuses in the eastern Indian and southwestern Pacific oceans, where scouring by cold bottom waters has stripped away much of the older sediments. The vigorous activity of the deep water is also indicated by widespread drift sediments in the North Atlantic.

Another important tectonic event that had major significance for the oceanic-climatic conditions was the breaching of the straits between South America and Antarctica at the Drake Passage in the mid-Oligocene and complete geographic isolation of Antarctica. This eliminated the last barrier in the path of the circum-Antarctic circulation. The development of the Cirum-Antarctic Current led to further thermal isolation of Antarctica that was conducive to the development of an extensive ice cap on the continent. Overall, there was a dramatic increase in the areas of the desert through the Oligocene and younger epochs that can be ascribed to the cooling higher latitudes and the development of polar ice caps. *See* PALEOCLIMATOLOGY; PALEOGEOGRAPHY; PLATE TECTONICS.

In Oligocene time the global surface circulation patterns had essentially evolved the major features

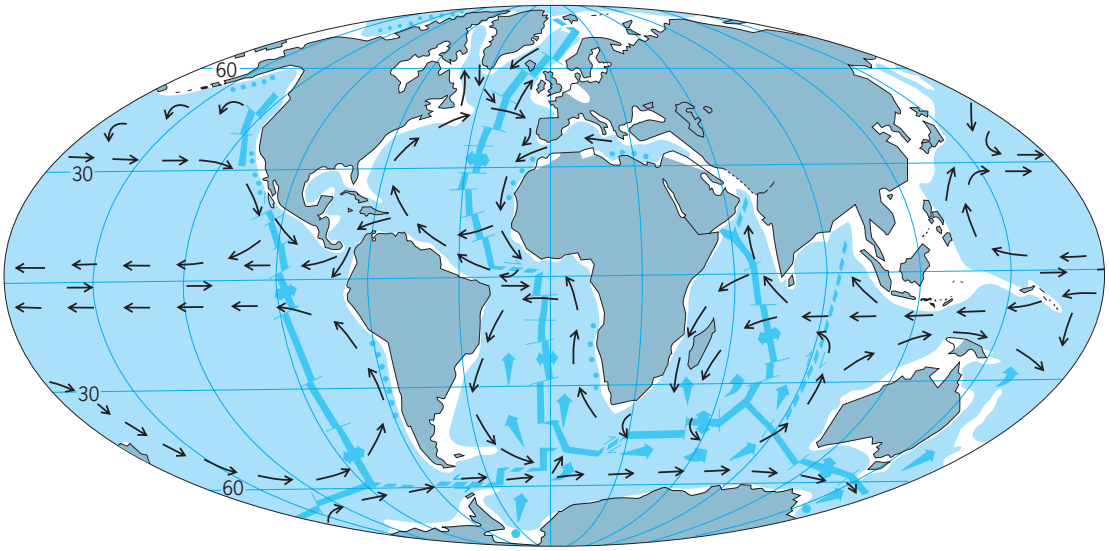

Fig. 2. Paleogeography and oceans of the Oligocene. (*After B. U. Haq and F. W. B. van Eysinga, Geological Time Table, Elsevier, 1998*)

of the modern oceans (**Fig. 2**). One major difference was that in the equatorial region an open Panama isthmus permitted the exchange of waters between the Pacific and the Atlantic oceans. However, by the Early Oligocene the flow of the Tethys Current from east to west had already become sharply reduced, intermittent, and restricted to a narrow passage southwest of the Indian plate, following the collision of the subcontinent with the Asian mainland, initial uplift of the Himalayas in the Eocene, and subsequent raising of the Tibetan Plateau. The convergence between India and Asia and between Africa and Europe closed the Tethys seaway, leaving behind smaller remnants that include the Mediterranean, Black, and Caspian seas.

The worldwide cooling trend that began in the latest Eocene but became accentuated in the Oligocene may have been in part caused by the long-term effect of the uplift of the Tibetan Plateau, which was significant by Oligocene time. The raised plateau is thought to be able to deflect the atmospheric jet stream more vigorously, leading to the strengthening of the summer monsoon, and increased rainfall and weathering in the Himalayas. Since increased weathering and dissolution of carbonates results in greater carbon dioxide drawdown from the atmosphere, the reduced levels of carbon dioxide may have provided a condition that helped the Earth enter into a renewed ice-house phase.

By mid-Oligocene the Circum-Antarctic Current was well established. The ice shelves around Antarctica were the dominant sources of cold bottom waters. Northern sources of cold, deep water had also become operative, and the surface exchange between North Atlantic and the Arctic through Norwegian-Greenland and Labrador passages had become active. There is indirect evidence that suggests the initiation of an ice cover in the formerly ice-free Arctic by Oligocene time. By this time the central Arctic Ocean had become largely landlocked, and the connections to the open sea were restricted enough that the input of fresh water from Eurasian rivers, combined with salinity stratification characteristic of isolated basins, may have allowed winter ice to form. This, however, does not seem to have affected the deeper Arctic waters, as indicated by the presence of benthic faunas well up to the late Miocene.

As a consequence of the cooling higher latitudes and expanding ice sheets, the sea-level history of the Oligocene is one of repeated regressions that were followed by smaller transgressions that successively covered less epicontinental areas than before. The first major sea-level fall occurred near the Eocene-Oligocene boundary, which most likely dates back to the first extensive ice accumulation on Antarctica. A second more pronounced global sea-level fall is indicated in the mid-Oligocene, around 30 m.y. B.P., when the sea level dropped about 150 m (500 ft) and continental margins around the world were laid bare to extensive erosion. This eustatic fall event was also in response to a significant amount of accumulation on the Antarctic ice cap. Later Oligocene eustatic fluctuations were of lesser magnitude, indicating less extensive ice-sheet fluctuations. In the early Miocene a major transgression occurred, and the sea covered much of the coastal areas that were exposed during the Oligocene.

The mid-Oligocene sea-level lowering was accompanied by extensive shelf erosion and stream incision on many continental margins. Prominent canyons were initiated on many of the exposed shelves that remained operative for several million years. The mid-Oligocene eustatic retreat of the seas was characteristically long-lasting. Although this event was followed by minor eustatic recoveries, the sea level never rose back to pre-Oligocene levels, and recovered fully only after the close of Oligocene Epoch. This implies that during the long lowstand of the sea the submarine canyons were kept active continuously, to be filled back only after the full eustatic

recovery in the early Miocene. *See* CONTINENTAL MARGIN; PALEOCEANOGRAPHY; SUBMARINE CANYON.

Life. Due to accentuated thermal gradients and seasonality in the Oligocene (compared to previous Tertiary epochs), marine biotic provinces became more fragmented. Extreme climates, with greater diurnal and seasonal temperature contrasts, are held responsible for reduced diversities in marine plankton. The marine micro and macro fauna and flora of the Oligocene have strong affinities with those of the late Eocene. The typically Paleogene assemblages gradually become extinct during the Oligocene. The boundary between Oligocene and Miocene represents a complete changeover to Neogene faunal elements. Thus, the Oligocene was characterized by transitional faunal features between the Paleogene and the Neogene.

Planktonic foraminifera, which had diversified rapidly in the Eocene, were severely reduced in diversity in the Oligocene. Calcareous nannofossils that had also proliferated earlier, suffered reduction in diversities, though somewhat less severe than the planktonic foraminifera. Benthic fauna fared better and some lineages actually diversified, perhaps due to greater niche fragmentation. However, the last of the *Nummulites* disappeared in the Late Oligocene. The discocyclinids, dominant in the Eocene, were followed by lepidocyclinids in the Oligocene, and later, in the Miocene, gave way to the miogypsinids.

An evolutionary leap was made by the mammals across the Eocene-Oligocene boundary. The Eocene perissodactyls and prosimians saw their last days, giving way to rhinocerids, tapirs, and wild boarlike hog species with strong incisors and large canines. *Hyaenodon* was an Oligocene carnivore with strong canines and sharp molars, much like those of the modern cat species. Grasslands still supported large browsers, and a new group of perissodactyls, typified by the *Titanotherium* (**Fig. 3**), appeared and climaxed, to later die out near the close of Oligocene. The horses that had first appeared in the Eocene continued to increase in size and became three-toed, typified by *Mesohippus*. Other hoofed mammals also diversified during the Oligocene. Elephants made their first appearance near the Eocene-Oligocene boundary and developed a short trunk and two pairs of tusks. An early simian, *Propliopithecus*, made its first appearance in Oligocene, and is considered ancestral to the modern family of gibbons. The general uniformity of mammalian fauna in the Oligocene suggests that the widespread regressions of the sea most likely resulted in land bridges that reconnected some of the Northern Hemispheric land masses, which may have led to transmigrations of some families of mammals between North America, Asia, and Africa. Birds had achieved some of their modern characteristics, and at least 10 modern genera had already made their appearance by the close of Oligocene time. *See* AVES; GEOLOGIC TIME SCALE; MAMMALIA; PALEOECOLOGY; PERISSODACTYLA.

Bilal U. Haq

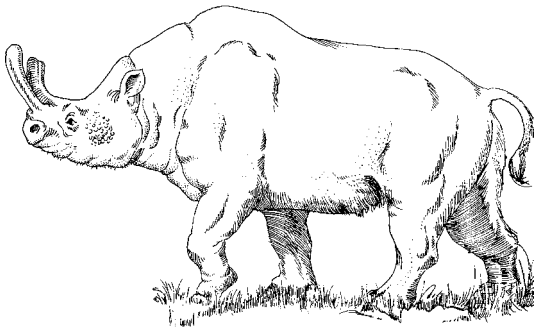

Fig. 3. Oligocene *Titanotherium* (=*Brontotherium*) of North America, a large browsing mammal. The titanotheres disappeared abruptly at the close of the Oligocene Epoch. (*After R. A. Stirton, Time, Life and Man, John Wiley, 1959*)

Bibliography. B. U. Haq, J. Hardenbol, and P. R. Vail, Chronology of fluctuation sea levels since the Triassic, *Science*, 235:1156–1167, 1987; B. U. Haq and F. W. B. van Eysinga, *Geological Time Table*, Elsevier, 1998; C. Pomerol, *The Cenozoic Era: Tertiary and Quaternary*, Ellis Horwood, 1982.

Oligochaeta

A class or subclass (depending on the classification system used) of worms, including the earthworms, within the phylum Annelida. These animals exhibit both external and internal segmentation—that is, furrows in the body wall and transverse partitions (septa) dividing the coelom into chambers. They usually possess chaetae (setae), or bristles, made of chitin, which are not borne on parapodia and are few in number (compared to the Polychaeta). Oligochaetes are hermaphroditic. The gonads are few in number and situated in the anterior part of the body, the male gonads being anterior to the female gonads. The gametes are discharged through the oviducts and sperm ducts. At maturity, a portion of the body wall is swollen into a secretory area called the clitellum (**illus.** *a*). There is no larval stage during the development of oligochaetes. *See* POLYCHAETA.

The oligochaetes are primarily freshwater and burrowing terrestrial animals. About 200 species are marine, mostly occurring in the intertidal zone. They range in size from aquatic species that are 0.02 in. (0.5 mm) long to the giant Australian earthworm (*Megascolides*), which is up to 3 m (10 ft) long. About 25 families and over 3000 species are presently recognized.

Body form and anatomy. Oligochaetes are cylindrical, elongated animals with the anterior mouth usually overhung by a fleshy lobe, the prostomium, and the anus located terminally. The body plan is that of a tube within a tube. The chaetae, usually a pair of ventrolateral bundles, are borne on most segments. Other external features are the pores of the reproductive systems opening on certain segments, the openings of the excretory organs (metanephridia), and in many earthworms dorsal pores which open externally from the coelom (see illus. *b*).

Body wall. The body wall consists of (from interior to exterior) a layer of thin epithelial cells (peritoneum) lining the coelom, muscle tissue arranged

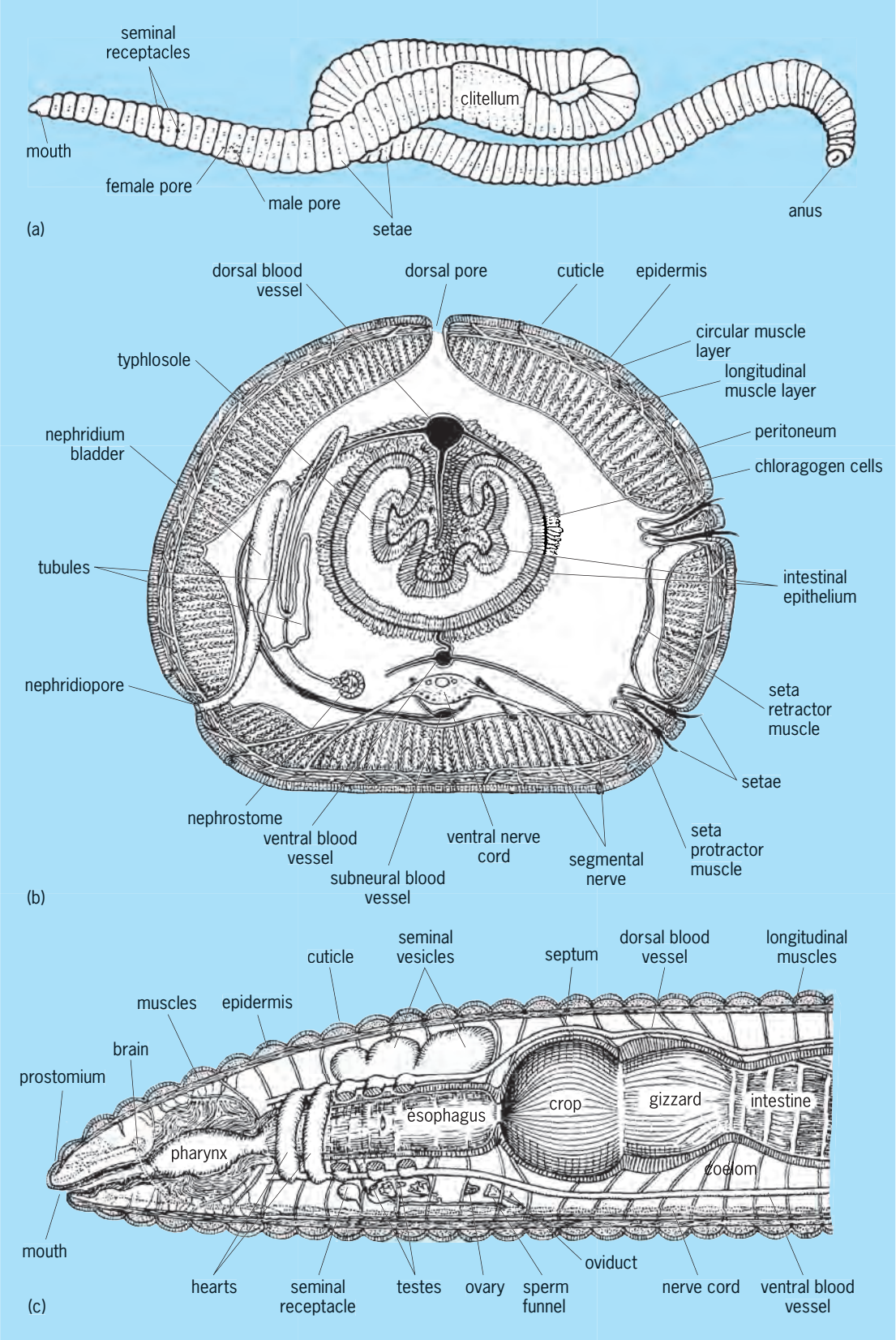

Lumbricidae. (*a*) Earthworm (*Lumbricus terrestris*), external features; (*b*) cross section (left half shows an entire metanephridium and a dorsal pore, and right half includes chaetae but no nephridium); (*c*) internal structures of anterior portion from the left side. (*After T. I. Storer, General Zoology, 3d ed., McGraw-Hill, 1957*)

in inner longitudinal and outer circular layers, and a layer of columnar epithelial tissue (epidermis) covered by a thin secreted cuticle. The epidermis contains interspersed glandular and sensory cells. Movement in terrestrial forms is accomplished by alternate contractions of the muscle layers, which produce a peristaltic motion in which the chaetae aid forward progression by gripping the substrate. Other, smaller muscles protract and retract the chaetae and are associated with the intestine and reproductive organs.

Coelom and gut. The fluid-filled coelom is segmentally partitioned in most cases by the septa. The alimentary canal passes through this series of spaces and is composed of a buccal cavity, a muscular-walled pharynx, a narrow esophagus, and the intestine proper. In the earthworms the posterior portion of the esophagus is often enlarged into a thin-walled chamber (crop) that stores ingested food and a thicker-walled portion (gizzard) that grinds the food (see illus. *c*). The pharynx may contain glands that secrete mucus and digestive enzymes, while the esophagus often demonstrates swellings (calciferous glands) that function in regulation of calcium and carbonate levels. In many of the larger oligochaetes the epithelium lining the gut wall, composed of glandular and absorptive cells, projects downward from the dorsum as a longitudinal fold (typhlosole) into the lumen of the intestine. There are thin muscle layers in the wall of the intestine, and it is invested with a layer of modified peritoneal chloragogen cells; these pigmented cells seem to be important in the synthesis and storage of glycogen and deamination of proteins. The food of most oligochaetes is plant material ingested by a sucking action of the pharynx. Some aquatic forms are carnivorous; a few others may be parasitic.

Nephridia and excretion. Excretion and osmoregulation in the oligochaetes are accomplished by segmentally arranged paired metanephridia. Typically each organ is composed of a funnel (nephrostome), which opens into the coelom, a variously coiled tubule sometimes with an expanded portion (bladder), and a nephridiopore which discharges ventrolaterally to the outside. Nitrogenous wastes and salts enter the metanephridium through the nephrostome and from capillaries surrounding it. Along the tubule, some salts and water may undergo reabsorption into the capillaries, depending on habitat, but products of protein breakdown (such as ammonia and urea) are mostly released at the nephridiopore.

Respiration. Exchange of respiratory gases is in general a function of the body surface. Some aquatic forms (for example, *Dero* and *Branchiura*) have posterior extensions of the body wall—filamentous gills—to increase surface area for diffusion of these gases.

Circulation. The circulatory system consists of a main dorsal vessel, often arising from a perienteric sinus, in which the blood flows anteriorly, and various lateral, ventral, and subneural vessels. The blood contains the respiratory pigment hemoglobin dissolved in it, and phagocytic corpuscles. Contraction of the dorsal vessel or lateral vessels ("hearts") propels the blood. *See* RESPIRATORY PIGMENTS (INVERTEBRATE).

Nervous system. The nervous system is composed of dorsal cerebral ganglia, circumesophageal connectives, and paired ventral nerve cords with segmentally arranged ganglia which give off nerves to the body wall and intestine. Sensory cells of various types are located in the epidermis. Oligochaetes react to a wide range of stimuli.

Reproduction. All oligochaetes are hermaphroditic. The testes are located anterior to the ovaries, and both are derived from the septal peritoneum. The gametes are liberated into the coelom or its pouches and reach the outside through variously differentiated ducts. In copulation, spermatozoa are received in spermatheca (seminal receptacles), and at oviposition, which occurs later, the clitellum secretes a cocoon into which eggs and spermatozoa are discharged. Embryonic development occurs within the cocoon. In the families Aeolosomatidae and Naididae, asexual reproduction by a type of binary fission (paratomy) occurs. *See* INVERTEBRATE EMBRYOLOGY; REPRODUCTION (ANIMAL).

Economic importance. The oligochaetes, especially the common large terrestrial "earthworm" *Lumbricus* (family Lumbricidae) and, more recently, the transparent aquatic "California blackworm" *Lumbriculus* (of the more primitive family Lumbriculidae), have been used in studies of anatomy, physiology, regeneration, and metabolic gradients. Some aquatic forms are important in studies of stream pollution as indicators of organic contamination. Earthworms are important in turning over the soil and reducing vegetable material into humus. It is likely that fertile soil furnishes a suitable habitat for earthworms, rather than being a result of their activity.

Classification. The nature of the reproductive system and the absence of parapodia sharply set the oligochaetes apart from the polychaetes. The Hirudinea (leeches) are much closer to the oligochaetes but constitute a homogeneous group long considered a separate class. The Oligochaeta historically have been treated as a class of the phylum Annelida, coordinate in rank with the Polychaeta and Hirudinea (or Hirudinomorpha). More recently, however, this taxon, regarded by some specialists as polyphyletic, has been merged with the Hirudinea (Hirudinomorpha) into a group (generally given class rank), known at the Clitellata, based on the mutual presence of the clitellum and other common features.

That the oligochaetes are descended from marine polychaete-like ancestors seems certain, but there is no agreement as to the number and relationships of subordinate taxa. Classification into orders and families has largely been based on placement of male gonopores relative to the segment the testes are located in, or similar traits of the reproductive system that are difficult to discern. Two of the three currently recognized orders—Lumbriculida and Moniligastrida—each contain a single family (the aquatic Lumbriculidae and the terrestrial Moniligastridae); both are characterized by four pairs of chaetae per segment. Remaining families are generally included in the order Haplotaxida. Of the aquatic

families, the most important ones ecologically are the larger Tubificidae, often reddish "sludge worms" of over 40 segments; Naididae, smaller worms usually with less than 40 segments (chaetae lacking on the anterior ones), often possessing eyespots but lacking prostomial cilia; and the even smaller (few millimeter) Aeolosomatidae, characterized by a ventrally ciliated prostomium. Tubificid worms generally live in vertical tubes embedded in bottom mud. Some, such as *Tubifex*, can tolerate very low oxygen concentrations, assisting the acquisition of oxygen from the water by waving their tails. Naidids (such as *Dero* and *Stylaria*, the latter bearing a long prostomium) are usually associated with submergent (entirely underwater) vegetation. The Aeolosomatidae (regarded by some as Polychaeta, since representatives lack the clitellum) are usually found gliding around the water-sediment interface. Families with high numbers of large terrestrial species include the widely distributed Lumbricidae (such as *Lumbricus*), the Glossoscolecidae of South America and the Caribbean, and the Megascolecidae of Asia and Australia. *See* ANNELIDA.

Perry C. Holt; Robert Knowlton

Bibliography. R. O. Brinkhurst and B. G. M. Jamieson, *Aquatic Oligochaeta of the World*, 1971; C. A. Edwards and P. Bohlen, *Biology and Ecology of Earthworms*, 3d ed., 1996; M. S. Laverack, *The Physiology of Earthworms*, 1963; E. E. Ruppert et al., *Invertebrate Zoology*, 7th ed., 2004; J. A. Wallwork, *Earthworm Biology*, 1983.

Oligoclase

A plagioclase feldspar with composition in the range $Ab_{90}An_{10}$ to $Ab_{70}An_{30}$, where Ab represents the composition of albite, $NaAlSi_3O_8$, and An represents the composition of anorthite, $CaAl_2Si_2O_8$. The diagnostic properties are hardness on Mohs scale, 6–6.5; density, 2.65 g/cm^3; two good cleavages that intersect at approximately 90°; and color usually white or colorless, transparent to translucent. The presence of minute, mutually parallel inclusions of hematite (Fe_2O_3) causes a golden play of color in the variety of oligoclase called aventurine or sunstone. Repeated twinning is common and results in closely spaced striations visible with a hand-held magnifier. Mean refractive index is 1.545; the mineral may be optically positive or negative, depending on composition. Oligoclase is triclinic. The mineral is common in both plutonic and volcanic silicic igneous rocks, as well as in quartzofeldspathic and pelitic metamorphic rocks. *See* ALBITE; ANORTHITE; CRYSTAL STRUCTURE.

Like all feldspars, oligoclase is a tectosilicate. Each aluminum (Al) and each silicon (Si) atom is surrounded by four oxygen atoms that form a tetrahedron around it; each tetrahedron shares the four oxygen atoms at its corners with four other tetrahedra, forming a three-dimensional atomic framework containing open voids that accommodate the sodium-calcium (Na,Ca) atoms. At the highest temperatures typical of igneous processes, the Al and Si atoms are statistically disordered among the symmetrically distinct tetrahedra. As cooling proceeds, Al and Si order preferentially into symmetrically different tetrahedra. If the composition is more Ab-rich than about $Ab_{83}An_{17}$, then exsolution or spinodal decomposition to an intergrowth of $Ab_{76\pm6}An_{24\pm6}$ and completely ordered $Ab_{99\pm1}An_{1\pm1}$ also occurs. The more calcic lamellae of this so-called peristerite intergrowth are e-plagioclase, consisting of alternating nanometer-scale lamellar domains of albitelike and anorthitelike structure. Oligoclases with bulk compositions more calcic than $Ab_{83}An_{17}$ do not form peristerite intergrowths, but they nonetheless develop into e-plagioclases upon cooling. *See* FELDSPAR; IGNEOUS ROCKS; METAMORPHIC ROCKS.

Dana T. Griffen

Bibliography. W. A. Deer, R. A. Howie, and J. Zussman, *An Introduction to the Rock-Forming Minerals*, 2d ed., 1992; D. T. Griffen, *Silicate Crystal Chemistry*, 1992; J. V. Smith and W. L. Brown, *Feldspar Minerals*, 2d ed., 1988.

Oligonucleotide

A single-stranded short polymer of deoxyribonucleic acid (DNA) or ribonucleic acid (RNA). Although the number of nucleic acid monomers (known as nucleotides) in the polymer varies widely, practical uses for oligonucleotides ranging in size from 4 to over 100 monomers have been reported.

Composition. Each nucleotide in the nucleic acid polymer is composed of three parts: a five-carbon sugar, a nitrogenous base, and a phosphate group (**Fig. 1**).

The five-carbon sugar (ribose or 2′-deoxyribose) is covalently attached to the nitrogenous base through a glycosidic linkage to create a nucleoside. If the sugar is ribose, the resulting oligonucleotide is known as an RNA oligonucleotide, or oligoribonucleotide; if the sugar is 2′-deoxyribose, the resulting polymer is called either a DNA oligonucleotide or an oligodeoxyribonucleotide.

There are two classes of bases: purines (adenine and guanine) and pyrimidines (thymine, cytosine, and uracil). Guanine, cytosine, adenine, and uracil are the four bases found in RNA oligonucleotides; DNA oligonucleotides contain guanine, cytosine, adenine, and thymine. The sequence, or ordering, of these bases determines the specificity of the interaction of one oligonucleotide with a second polymeric strand (either DNA, RNA, or a protein) to form a duplex (**Fig. 2**).

The phosphate molecule covalently links the 5′ hydroxyl group of one nucleoside to the 3′ hydroxyl group of the next nucleoside. This creates a phosphodiester bond between successive sugar molecules, forming the backbone of the nucleic acid strand (Fig. 1). *See* DEOXYRIBONUCLEIC ACID (DNA); NUCLEIC ACID; RIBONUCLEIC ACID (RNA).

Synthesis. Oligonucleotides are synthesized from high-molecular-weight DNA or RNA in a manner that defines the order, or sequence, of purines and

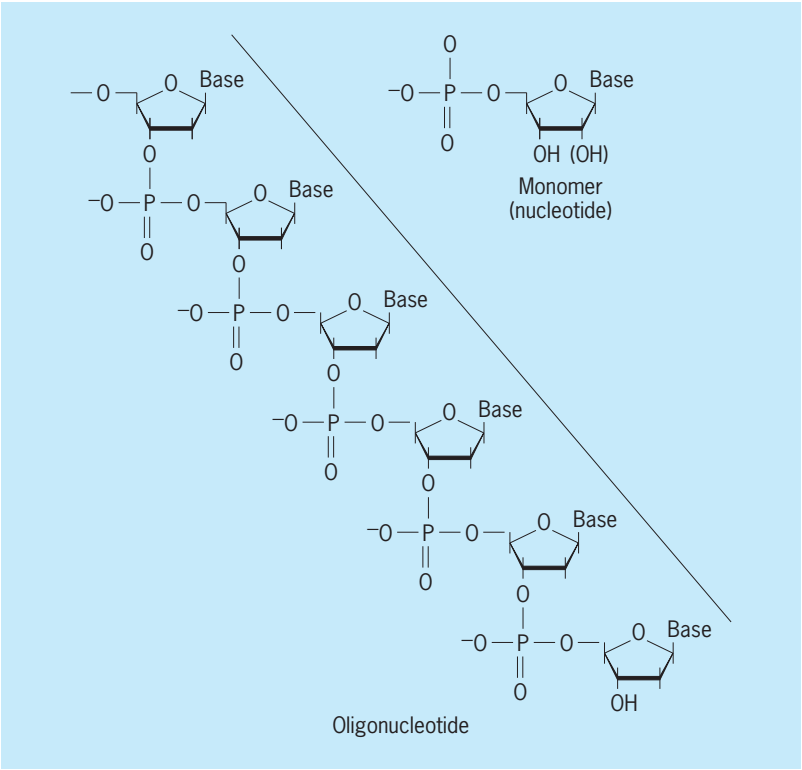

Fig. 1. Oligonucleotides are composed of monomers called nucleotides. The most common chemical form of monomers has either deoxyribose or ribose (an additional OH) attached to a purine (adenine and guanine) or pyrimidine base [thymine (in DNA oligonucleotides), cytosine, uracil (in RNA oligonucleotides)]. Monomers are joined in a consistent and directional manner through the formation of phosphodiester bonds.

pyrimidines. This synthesis is most often carried out chemically using automated instruments, but may also be carried out enzymatically using RNA or DNA polymerase.

Modifications. The basic chemical structure of the oligonucleotide can be modified in a number of ways to enhance desired properties. For instance, purines or pyrimidines can be modified by alkylation (addition or substitution of an alkyl group) or deamination (removal of an amino group) to alter their interaction with other nucleic acids. A methyl group can be added to the sugar at the 2′ carbon to reduce susceptibility to nucleases (enzymes that cleave nucleic acids by attacking the phosphodiester bond), or replaced entirely with other molecules, such as a morpholino. The bridge between monomers can be changed from a phosphodiester linkage to a phosphorothioate (P-S bond) or phosphoramidate (P-N bond). These modifications are made to provide resistance against the numerous cellular nucleases. However, the vast majority of the oligonucleotides currently synthesized have unmodified purine and pyrimidines bases and 2′-deoxyribose as the sugar, and are linked through a normal phosphodiester bond.

Applications. Oligonucleotides are commonly used to bind to other nucleic acids or proteins through base-specific hydrogen bonds. Generally, they are used to form antiparallel duplexes, in which the sugar-phosphate backbones run in opposite directions (one strand runs in the 5′- to 3′-direction and the other runs 3′ to 5′). Oligonucleotides may also bind, in a sequence-specific manner, to a nucleic acid duplex to form three-stranded structures known as a triplexes. Oligonucleotides used to form triplexes with genomic DNA have been used experimentally to inhibit gene expression and to create site-directed mutations.

In vitro: DNA oligonucleotides. Oligonucleotides have been used to assemble genes, or even a small genome, as was demonstrated by the assembly of the complete poliovirus genome in a research laboratory. Perhaps the most common current use of chemically synthesized oligonucleotides is as part of a nucleic acid amplification technique called the polymerase chain reaction (PCR). In a PCR reaction, two different DNA oligonucleotides interact, in a sequence-specific manner, with a long segment of denatured DNA to form two antiparallel duplexes a defined distance apart (a few hundred to a few thousand base-pairs). A unique DNA polymerase known as Taq, which has been isolated from the bacteria *Thermophilus aquaticus*, uses the oligonucleotides as primers to synthesize DNA in vitro in a series of reactions that doubles the amount of newly synthesized DNA with each cycle of denaturation, annealing, and elongation. By repeating this cycle from 35 to 50 times, one copy of a starting DNA template can be amplified to produce useful amounts of a DNA fragment. *See* GENE AMPLIFICATION.

The cost of chemically synthesizing DNA oligonucleotides has decreased dramatically. A result of this important technological breakthrough has been the development of oligonucleotide gene chips, in which tens of thousands of DNA oligonucleotides (representing different potentially expressed genes) are attached to a solid matrix. Messenger RNA (mRNA) molecules from specificcells are first

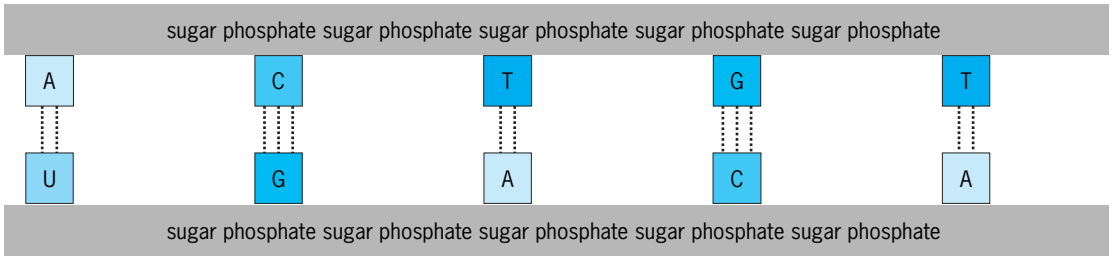

Fig. 2. Oligonucleotides make sequence-specific, noncovalent dimers (or duplexes) with DNA or RNA through Watson: Crick base pairing.

amplified and labeled and then added to the oligonucleotide gene chip. Some of the labeled nucleic acid hybridizes to the complementary DNA oligonucleotides on the gene chip, and the remainder is washed away. A computer determines the location and intensity of the signal generated on the chip, thus determining a pattern of genes expressed in the original cells. This type of transcriptional profiling has revealed groups of genes that are coordinately regulated in some cancers and other disease states.

In vivo: RNA oligonucleotides. In 1998, it was found that long double-stranded RNA oligonucleotide duplexes injected into worms could alter gene expression. This process, known as RNA interference (or RNA silencing), has been described in cells from a number of diverse species. In cells, long RNA oligonucleotide duplexes are first cut into shorter, biologically active 21–25 base-pair fragments. (Alternatively, duplexes can first be made from RNA oligonucleotides 21–25 bases in length, and then introduced into the cell.) After associating with a number of specific proteins, an RNA oligonucleotide: protein complex is formed, in which one strand of the RNA duplex is degraded and the other serves as a guide, leading to the degradation of a complementary mRNA.

In vivo: DNA oligonucleotides. The therapeutic use of DNA oligonucleotides to treat diseases in humans has taken two main forms. The first is used in clinical trials as antiviral or anticancer therapeutic agents. Oligonucleotides act by serving as a specific inhibitor of messenger mRNA translation inside cells. The effect of the oligonucleotide is to reduce the level of proteins (encoded by mRNAs) that allow a virus to propagate or a cancer cell to divide in an unregulated way. The DNA oligonucleotide first binds to a complementary mRNA. This approach is often called "antisense," which refers to the anti-parallel binding rules of guanine binding to cytosine (G-C), and adenine binding to uracil (A-U). The fate of the formed DNA-RNA heteroduplex depends upon the type of DNA oligonucleotide used. If the oligonucleotide has regions with native phosphodiester bonds, the RNA in the heteroduplex will be cleaved by the ubiquitous enzyme ribonuclease H, thus preventing translation of the mRNA. If the oligonucleotide is modified so that the heteroduplex is not recognized by ribonuclease H, then any inhibition of translation must result from preventing the proper assembly of the translational machinery (that is, by blocking ribosome binding).

A second therapeutic use for DNA oligonucleotides is to activate the immune system. The sequence cytosine-guanine (CpG) is usually methylated in vertebrate, but not bacterial, DNA. DNA oligonucleotides with unfamiliar cytosine methylation patterns provide a signal in humans to initiate cellular processes that activate both innate and acquired immunological responses. CpG immune modulators are currently being studied as agents to treat a number of human diseases. *See* GENE; GENETIC CODE; GENETIC ENGINEERING; IMMUNITY.

Daniel Weeks; John Dagle

Bibliography. J. Cello, A. V. Paul, and E. Wimmer, Chemical synthesis of poliovirus cDNA: Generation of infectious virus in the absence of natural template, *Science*, 297:1016–1018, 2002; J. M. Dagle and D. L. Weeks, Oligonucleotide-based strategies to reduce gene expression, *Differentiation*, 69:75–82, 2001; A. Fire et al., Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans, *Nature*, 391:806–811, 1998; A. M. Krieg, CpG motifs in bacterial DNA and their immune effects, *Annu. Rev. Immunol.*, 20:709–760, 2002; R. J. Lipshutz et al., High density synthetic oligonucleotide arrays, *Nat. Genet.*, 21:20–24, 1999; R. J. Lipshutz et al., Using oligonucleotide probe arrays to access genetic diversity, *Biotechniques*, 19:442–447, 1995; A. C. Pease et al., Light-generated oligonucleotide arrays for rapid DNA sequence analysis, *Proc. Natl. Acad. Sci. USA*, 91:5022–5026, 1994; K. H. Rubins et al., The host response to smallpox: analysis of the gene expression program in peripheral blood cells in a nonhuman primate model, *Proc. Natl. Acad. Sci. USA*, 101:15190–15195, 2004; R. K. Saiki et al., Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia, *Science*, 230:1350–1354, 1985; U. Scherf et al., A gene expression database for the molecular pharmacology of cancer, *Nat. Genet.*, 24:236–244, 2000; J. Soutschek et al., Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs, *Nature*, 432:173–178, 2004; C. A. Stein and A. M. Krieg, *Applied Antisense Oligonucleotide Technology*, Wiley, 1998; J. Summerton and D. Weller, Morpholino antisense oligomers: design, preparation, and properties, *Antisense Nucleic Acid Drug Dev.*, 7:187–195, 1997.

Oligopygoida

An order of irregular echinoids in the superorder Neognathostomata resembling clypeasteroids but lacking the accessory ambulacral pores characteristic of that group. Oligopygoids have well-developed petals, and there are characteristic small demiplates present below the petals. These demiplates are wedge-shaped and usually do not reach the inner surface of the test. Each has a simple ambulacral pore. The apical disk is monobasal and the mouth oval and usually deeply sunken. Oligopygoids have a lantern, which closely resembles that of clypeasteroids, and their lantern muscle-attachment structures are a mixture of ambulacral and interambulacral processes.

There are two genera, *Oligopygus* and *Haimea*, containing about 25 species, all from the middle and upper Eocene of the Caribbean and Gulf of Mexico regions. They were probably infaunal deposit feeders like present-day laganiids. *See* ECHINODERMATA; NEOGNATHOSTOMATA.

Andrew B. Smith

Bibliography. A. B. Smith, *Echinoid Palaeobiology*, 1984.

Oligosaccharide

Oligo(few)saccharide(sugar): A carbohydrate molecule composed of 3–20 monosaccharides. Generally, free oligosaccharides do not quantitatively constitute a significant proportion of naturally occurring carbohydrates. Most carbohydrates that occur in nature are in the form of monosaccharides (such as blood sugar, or glucose), disaccharides (such as table sugar, or sucrose, and milk sugar, or lactose), and polysaccharides (such as starch and glycogen, polyglucose molecules, or chitin). *See* GLUCOSE; LACTOSE; MONOSACCHARIDE; POLYSACCHARIDE.

Composition. The monosaccharides of multiple sugar units such as disaccharides, oligosaccharides, and polysaccharides are connected with each other through bonds called glycosidic linkages. Monosaccharides are generally classified as aldose (an aldehyde sugar) or ketose (a ketone sugar). Chemically, these sugar molecules are found mainly as cyclic molecules, which are referred to as hemiacetals. Cyclic hemiacetals exist as five-membered rings (called furanoses) or six-membered rings (called pyranoses). Both classes of sugars are reducing sugars, meaning that the aldehyde or ketone group can become oxidized and in turn will reduce certain compounds. Simple sugars (monosaccharides) are linked primarily to other sugars and to other molecules through these reducing groups. For example, lactose (**1**) has the structure of galactosyl-

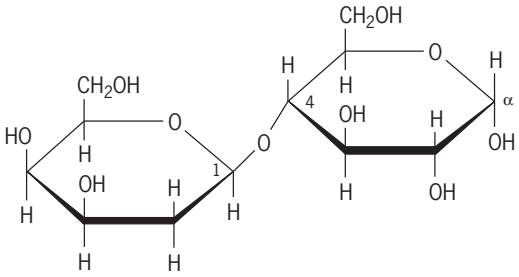

(**1**)

1,4-glucose, meaning that carbon-1 of galactose is linked to carbon-4 of glucose. The proper name for this structure of lactose is β-D-galactopyranosyl-(1-4)-α-D-glucopyranose. In ring structures, the aldehyde or ketone carbon is referred to as the anomeric carbon, and the cyclized aldose or ketose exists in one of two anomeric configurations designated as α or β. The complexity of oligosaccharide chemistry is appreciated when the possible combinations are considered in forming a disaccharide between galactose and glucose. *See* GALACTOSE.

Carbon-1 of galactose can be linked to any of the hydroxyl groups of glucose, giving four possible disaccharides. Since carbon-1 of galactose exists as the anomeric carbon (α or β), the number is multiplied by 2, so that there are eight possible structures based only on these assumptions. The number of possible isomeric oligosaccharides that could exist for a trisaccharide composed of three different monosaccharides with the typical structure given in (**1**) is 1056. If the reducing group of one monosaccharide is linked to the reducing group of another monosaccharide, such as in sucrose (**2**), the resulting disaccharide is nonreducing.

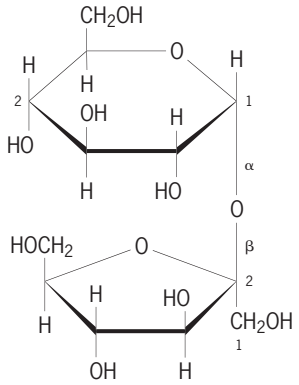

(**2**)

Glycoconjugates. Most naturally occurring oligosaccharides are linked either to proteins (glycoproteins) or to lipids (glycolipids). A number of different sugars can occur in oligosaccharides of glycoproteins and glycolipids. For example, in eukaryotic glycoproteins, seven sugars predominate: these are D-galactose, D-glucose, D-mannose, and L-fucose; the hexosamines *N*-acetyl-D-glucosamine and *N*-acetyl-D-galactosamine; and the nine-carbon sugars called sialic acids (in humans, the sialic acid is primarily *N*-acetylneuraminic acid). The amino sugars have the hydroxyl group on carbon-2 replaced by an amino group (NH_2), which is in turn acetylated. Two of the common linkages of sugars to proteins are O-linked glycoproteins and N-linked glycoproteins. Examples are *N*-acetylgalactosamine, linked to the hydroxyl group of the amino acid serine, and *N*-acetylglucosamine, linked to the amide nitrogen of the amino acid asparagine. *See* GLYCOLIPID; GLYCOPROTEIN.

Glycoconjugates are present in essentially all life forms and particularly in cell membranes and cell secretions. Many hormones are glycoproteins, and an increasing number of enzymes have been shown to have sugars attached. Antigenic properties of the human red blood cell ABO blood group system are determined by glycolipid oligosaccharides. In fact, all the major protein components of blood serum, with the exception of serum albumin, are glycoproteins. *See* BLOOD GROUPS; CELL MEMBRANES.

Many changes in the structures of oligosaccharides of glycoconjugates have been detected in cancer cells. Often these alterations resemble structures present in fetal tissues. Changes or differences in oligosaccharide structures are generally the result of differences in biosynthetic pathways or of degradative pathways.

Biosynthesis. The biosynthesis of essentially all macromolecules such as proteins, nucleic acids, and polysaccharides requires an activated form of the monomeric units. The nature of the first activated sugar was established as a glucose molecule linked to a nucleotide, commonly called uridine diphosphate glucose (UDPglucose). Over 100 different nucleotide

sugars are known wherein both the nucleotide and the sugar can vary.

Specific enzymes called glycosyltransferases transfer one sugar at a time to acceptor molecules. The acceptors can be highly divergent substances such as proteins, lipids, and other sugars. For example, the blood type A glycoprotein of pig salivary secretion (an O-linked glycoprotein) is synthesized by adding the following sugars in sequence: *N*-acetylgalactosamine, galactose, fucose, and *N*-acetylgalactosamine from their respective nucleotide derivatives. Oligosaccharides of N-linked glycoproteins are mechanistically synthesized the same as O-linked glycoproteins; that is, one sugar is added at a time from nucleotide sugars. However, oligosaccharides of N-linked glycoproteins are first formed on a lipid intermediate as a dolichol pyrophosphate-oligosaccharide, and then the oligosaccharide is transferred to an asparagine group of the protein moiety by oligosaccharide transferase. Specific glycosidases (removal of certain sugars) and glycosyltransferases (addition of specific sugars) (together called oligosaccharide processing) are involved in the synthesis of an extremely large number of different oligosaccharide side chains.

Biological functions. There is a great interest in determining the functions of glycoconjugates, and especially the roles of the oligosaccharide side chains. Sugar-mediated glycoprotein clearance from blood serum was first established in studies on Wilson's disease, an inherited disorder of copper metabolism. The normal half-life of most glycoproteins in serum is generally about 56 h; but when the terminal sialic acid was removed from ceruloplasmin (the copper-carrying glycoprotein) by the glycosidase sialidase, liver plasma membranes recognized the exposed galactose moiety and essentially immediately removed the asialoceruloplasmin from circulation. Receptor molecules for hormones, growth factors, and other biological regulators as membrane-bound glycoconjugates have been identified and characterized. Proteins produced by recombinant deoxyribonucleic acid (DNA) technology, or biotechnology, which are potential glycoproteins, require specific glycosylation to elicit a proper biological response. For example, if erythropoietin, a glycoprotein which regulates the production of red blood cells, is not glycosylated properly, it is rapidly removed from the serum. An understanding of glycoconjugates in normal biological systems and in certain disease states is currently of great importance. *See* BIOTECHNOLOGY.

Don M. Carlson

Bibliography. B. Alberts et al., *Molecular Biology of the Cell*, 3d ed., 1994; A. L. Lehninger, D. L. Nelson and M. M. Cox, *Principles of Biochemistry*, 3d ed., 2000.

Olive

Olive fruits, which are produced by a small to medium-sized evergreen tree (*Olea europaea*), can be eaten, after processing, as table olives, or can be extracted for oil that is used on salads, for cooking, for body lotions, or for medicinal purposes. The olive tree (**Fig. 1**) is a historically ancient cultivated plant, having been domesticated by early civilizations in the eastern Mediterranean regions. Olive culture later spread to all the Mediterranean countries and subsequently, during the age of exploration, to South America, California, South Africa, and Australia. The major olive-producing countries are Spain, Italy, and Greece, which provide about 65% of the world's olives. Olives are grown commercially in the United States only in California, and are used primarily as table olives. The olive tree has very precise climatic requirements for good fruit production. First, the trees are killed outright if winter temperatures drop much below 14°F (−10°C). Second, trees of most varieties require a certain amount of winter cold in order to flower in the spring, ruling out the tropics, and third, they need a long, hot, dry summer growing period to properly mature the fruit. *See* EVERGREEN PLANTS; FRUIT, TREE.

Fig. 1. Olive tree with fruits.

Olive trees are very drought-resistant, but they produce better if they are given summer irrigation. The trees also respond markedly to nitrogen fertilizers with increased growth and yields.

Olive trees are unusually free of insect and disease pests. However, one insect, the olive fly (*Dacas olea*), is a serious problem in the Mediterranean countries, causing the fruits to be wormy. Fungus diseases affecting olives are verticillium wilt and peacock spot. A bacterial disease, olive knot, is also a problem in many countries.

Olive trees bloom from early April to late May in the Northern Hemisphere, depending on the season and the region. The wind-pollinated flowers are small and of two types: having both male and female flower parts, or having only male flower parts (**Fig. 2**). The trees tend toward alternate bearing—a heavy crop

Fig. 2. Olive branches showing (left) blooming in the spring and (right) mature fruits ready for harvest for oil in the winter.

one year followed by no crop or only a light crop the next.

Olive fruits are harvested for table olives in the autumn, when the fruits change from green to straw or to a slightly red color. Raw fruits contain a bitter glucoside which makes them inedible, but treatment with an alkali such as sodium hydroxide (lye) neutralizes the bitterness. The lye must subsequently be leached out of the fruits with water. In another method most of the bitterness can be removed by leaching with salt water. A lactic acid fermentation process is widely used in the Mediterranean countries to preserve the olives. In California the lye treatment is mostly used, after which the olives are canned in brine and autoclaved at 240°F (115.5°C) for 1 h. If the fruits are aerated during the lye treatments, they become black, and are called black-ripe olives. If they are kept in the solution and not exposed to air, they remain green, and are called green-ripe olives.

For oil production, the fruits are harvested in midwinter when they have become black and have reached their maximum oil content—15 to 25% of the fresh weight, depending on the variety. In producing olive oil, the freshly harvested fruits including pits are ground, after which this material is placed in burlap or cloth bags and placed under high pressure. Oil and water are extracted and transferred to tanks, where the oil rises to the top and is removed. The oil is washed with warm water several times to remove the bitterness, and again separated by gravity. The oil is passed through centrifuges to remove all water, and through filters to remove suspended matter and give the oil clarity. *See* FAT AND OIL (FOOD).

Some leading table varieties are Sevillano, Manzanillo, Amphissa, Mission, Calamata, and Ascolano. Oil varieties are Picual, Frantojo, Chemlali, and Mission. Many varieties are grown either for table olives or for olive oil.

Hudson T. Hartmann

Olivine

A name given to a group of magnesium-iron silicate minerals crystallizing in the orthorhombic system. Crystals are usually of simple habit, a combination of the dipyramid with prisms and pinacoids. The luster is vitreous and the color olive green, giving rise to the name olivine. Hardness is $6^1/_2$–7 on Mohs scale; specific gravity is 3.27–3.39, increasing with increase in iron content.

Olivine is a nesosilicate with the composition $(Mg,Fe)_2SiO_4$. It comprises a complete solid solution series from the pure iron member fayalite, Fe_2SiO_4, to the pure magnesium member forsterite, Mg_2SiO_4. Minerals of intermediate composition have been given their own names but are usually designated simply as olivine. The magnesium-rich varieties are more common than those rich in iron. The minerals tephroite, Mn_2SiO_4, monticellite, $CaMgSiO_4$, and larsenite, $PbZnSiO_4$, although not in this chemical series, are of the olivine structure type.

Occurrence. Olivine is found in some crystalline limestones but occurs chiefly as a rock-forming mineral in igneous rocks. It varies greatly in amount from an accessory to the main rock-forming constituent. Although it may be present in granites and other light-colored rocks, it is found chiefly in the dark rocks such as gabbro, basalt, and peridotite. The rock dunite is composed almost completely of olivine.

Olivine is one of the first minerals to form upon crystallization of a magma. It is believed that this early-formed olivine accumulated through the process of magmatic differentiation to form the large dunite masses. The type locality is at Dun Mountain, New Zealand; the rock is also found with corundum deposits in North Carolina. *See* MAGMA; PERIDOTITE.

Olivine alters readily to serpentine, a hydrous magnesium silicate. The alteration may take place on a large scale to form great masses of the rock serpentine, or on a small scale to form pseudomorphs of serpentine after single crystals of olivine. *See* SERPENTINE; SERPENTINITE.

At a few localities, notably on St. John's Island in the Red Sea and in Burma, olivine is found in transparent crystals. These are cut into gemstones which go under the name of peridot. Olivine is a major constituent of many stony meteorites. *See* METEORITE.

Cornelius S. Hurlbut, Jr.

Spinel transition. When subjected to very high pressures, minerals with olivine structure are transformed to denser polymorphs with spinel structure. For the composition Fe_2SiO_4, the univariant curve for the olivine-spinel transformation has been determined experimentally. The divariant transition interval for compositions in most of the series Fe_2SiO_4–Mg_2SiO_4 has also been determined at pressures up to about 100 kilobars. Olivine is believed to constitute a high proportion of the upper mantle, estimates ranging between 60 and 90%. The composition of the olivine is generally assumed to be about 90% forsterite (Mg_2SiO_4) and 10% fayalite (Fe_2SiO_4). The experimental results, combined with estimates of the temperature distribution in the

mantle, indicate that olivine of this composition would begin to transform to a more iron-rich spinel phase at a depth of about 220 mi (370 km), with transformation being completed at a depth of about 270 mi (435 km). Gradients of seismic-wave velocities in the transition zone of the upper mantle are abnormally high in two layers 30–60 mi (50–100 km) thick, one beginning at depth of 220 mi (350 km) and the other at about 390 mi (630 km). The olivine-spinel transformation appears to correlate well with the 220-mi (350-km) layer. The spinel present below this layer may be replaced in turn by even more dense phases at the 390-mi (630-km) depth. *See* LITHOSPHERE; SPINEL.

Peter J. Wyllie

Bibliography. W. A. Deer, R. A. Howie, and J. Zussman, *Rock-Forming Minerals, Orthosilicates*, Vol. 2a, 2d ed., 1997.

Omphacite

The pale to bright green monoclinic pyroxene found in eclogites and related rocks. Omphacites are essentially members of the solid solution series between jadeite ($NaAlSi_2O_6$) and diopside ($CaMgSi_2O_6$), but contain smaller amounts of the components acmite ($NaFe^{3+}Si_2O_6$), hedenbergite ($CaFe^{2+}Si_2O_6$), and Tschermak's pyroxene ($CaAlAlSiO_6$). The density ranges from 3.16 to 3.43 g/cm^3 (1.8–2.0 oz/in.^{3}), hardness is 5–6, cleavage is well developed on {110}, and simple or lamellar twinning on {100} is common. With four formula units to the unit cell, omphacite crystallizes in space groups Cs/c, as well as in the lower symmetry group P2/n. The possibility of space groups P2 and P2/c has also been suggested for some omphacites. C2/c structure is characteristic of omphacites crystallized above a P2/n-C2/c inversion temperature (just over 1300°F or 700°C), although there is some evidence for metastable growth at lower temperatures. In this structure (Ca,Na) and (Al,Mg) are more or less randomly distributed over M2 and M1 sites. *See* DIOPSIDE; ECLOGITE; JADEITE; SOLID SOLUTION.

Compositions close to 50% jadeite and 50% diopside ($Jd_{50}Di_{50}$) have a tendency to order their M1 chains into a regular alteration of Al and Mg octahedra. To preserve local charge balance, the Na and Ca ions are partially ordered on the M2 chains. Substitution of the acmite component tends to inhibit this ordering, so that natural P2/n omphacites cluster about the $Jd_{50}Di_{50}$ composition. The schematic phase diagram (see **illus.**) is based on the uncertain assumption that the ordering transition is first-order in character.

Omphacite is stable only at the relatively high pressures of the blueschist and eclogite facies of metamorphism, where it is associated with minerals such as glaucophane and lawsonite or pyropic garnet, respectively. In such environments it also occurs in veins, either on its own or with quartz. *See* GLAUCOPHANE.

High pressures favor minerals in which aluminum is octahedrally coordinated, so that for the reaction

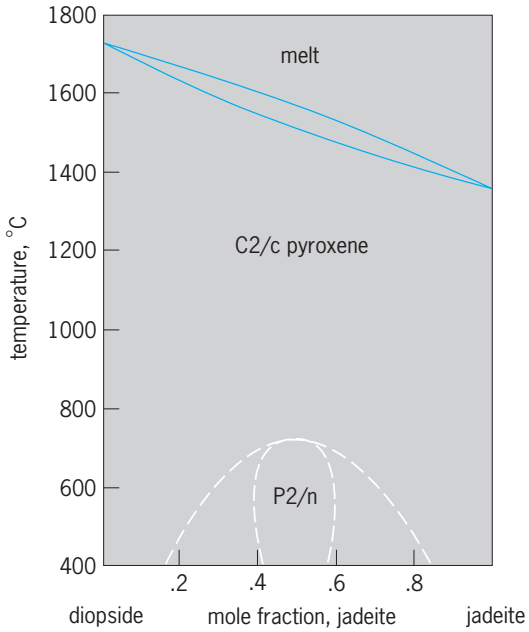

Temperature-composition section at 30 kilobars (3 gigapascals) pressure for the jadeite-diopside system. The low-temperature region is shown schematically. °F = (°C × 1.8) + 32.

$CaMgSi_2O_6 + xNaAlSi_3O_8 = (CaMgSi_2O_6{\cdot}xNaAlSi_2O_6) + xSiO_2$ jadeite solubility in diopside increases with increasing pressure. Experimental work on this reaction allows an estimate of the pressure of equilibration of omphacite + albite + quartz assemblages if a temperature estimate is available. For feldspar-absent assemblages, the pressure estimated for a particular omphacite composition is a minimum value. *See* HIGH-PRESSURE MINERAL SYNTHESIS; PYROXENE.

Timothy J. B. Holland

Bibliography. W. A. Deer, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, Vol. 2B, Double-Chain Silicates, 2d ed., 1997; K. Smulikowski, Chemical variation of clinopyroxenes in eclogites and related rocks, *Bull. Acad. Pol. Sci.*, Serie Sciences de la Terre, 24(1):1–7, 1976.

Oncholaimida

An order of nematodes comprising the single superfamily Oncholaimoidea. These nematodes are principally marine and brackish-water forms with alleged predaceous and carnivorous feeding habits. The external amphidial aperture is an oval or a widened ellipse. Generally, the stoma is armed with one dorsal tooth and two subventral teeth, and its wall may be further fortified with transverse rows of small denticles. The stoma is divisible into the cheilostome (secondary blastocoel invagination) and the esophastome (primary blastocoel invagination). In some species the stoma of the adult male is collapsed or indistinct. The cephalic sensilla are in two whorls: one is circumoral and composed of six papilliform sensilla; the second combinesthe ancestral two whorls

of six and four into a single whorl of ten setiform sensory organs. In some forms these sensilla are papilliform. The cylindrical-to-conoid esophagus may exhibit a series of muscular bulbs posteriorly. The cuticle is generally smooth, and over the length of the body there are scattered sensory setae or papillae. *See* NEMATA. Armand R. Maggenti

Oncofetal antigens

Small proteins or carbohydrates (capable of stimulating an immune response) that are present in normal fetal tissues during early development but also are abnormally expressed by adult tumors. Oncofetal antigens are different from tumor-associated antigens, which are found only in tumors and have no role in development.

One of the goals of modern cancer research is to determine biochemical or immunological differences between normal and tumor cells. Although the malignant state is not necessarily characterized by gross aberrations in metabolic pathways or in the biochemical constitution of cell components, a subtle change does occur in a cell during neoplastic transformation. One area of investigation is the study of oncofetal antigens. These antigens, primarily glycoprotein in nature and produced by cancerous cells, are the products of one or more genes that normally are expressed only during fetal development and then are repressed in adult life. Production of these proteins in association with cancer as a result of activation of control genes, by unknown mechanisms, has been termed retrogenetic expression. Advances in tumor immunology have enabled the use of oncofetal antigens in both the diagnosis and treatment of cancer. In particular, fetal antigens in adult body fluids can serve as tumor markers for the detection of early oncogenic processes and can aid in monitoring the efficiency of, cancer treatment as well as in the development of new treatment modalities.

The two best-known oncofetal antigens are α-fetoprotein and carcinoembryonic antigen (CEA); however, several lesser-known oncofetal antigens, such as pancreatic oncofetal antigen, have been described.

Alpha fetoprotein. Alpha fetoprotein is a substance produced by the liver of a healthy fetus. The exact function of this protein is unknown. However, the physiochemical properties of α-fetoprotein are similar to those of albumin, the major normal serum component in the adult. In fetal serum, an inverse relationship exists between α-fetoprotein level decreases and the albumin level increases with increasing fetal age. After birth, the infant's liver stops producing α-fetoprotein, and an adult liver contains only trace amounts. During pregnancy, the fetus excretes α-fetoprotein in urine, and some of the protein is able to cross the placenta and enter the mother's blood. *See* ALPHA FETOPROTEIN.

Birth defects. Analysis of the amount of α-fetoprotein found in the mother's blood, can aid in the determination of the probability that the fetus is at risk for certain birth defects. Alone, α-fetoprotein screening cannot diagnose a birth defect. It is often part of a "triple check" blood test that analyzes three substances as risk indicators of possible birth defects: α-fetoprotein, estriol, and human chorionic gonadotropin (HCG). When all three substances are measured in the mother's blood, the accuracy of the test results increases. *See* HORMONE; PRENATAL DIAGNOSIS.

Abnormally high α-fetoprotein may indicate that the fetus has an increased risk of a neural tube defect, the most common and severe type of disorder associated with increased α-fetoprotein. These include spinal column defects and anencephaly (a severe and usually fatal brain abnormality). If the screening test indicates abnormally high α-fetoprotein, ultrasound is used for the diagnosis. Levels may also be high if there is too little fluid in the amniotic sac around the fetus, more than one developing fetus, or a pregnancy that is further along than estimated.

For unknown reasons, abnormally low α-fetoprotein may indicate that the fetus has an increased risk of Down syndrome. Down syndrome is a condition that is linked to an abnormality of chromosome 21 (called trisomy 21). If the screening test indicates an abnormally low α-fetoprotein, amniocentesis is used in the diagnosis. Abnormally low levels of α-fetoprotein can also occur when the fetus has died or when the mother is overweight. *See* CONGENITAL ANOMALIES; DOWN SYNDROME.

Cancer. Although α-fetoprotein in human blood gradually disappears after birth, it never disappears entirely. It may reappear in liver disease or in tumors of the liver, ovaries, or testicles. The α-fetoprotein test is used to screen people at high risk for these conditions. After a cancerous tumor is removed, an α-fetoprotein test can monitor the progress of treatment. Continued high α-fetoprotein levels suggest the cancer is growing. *See* CANCER (MEDICINE).

Carcinoembryonic antigen. The carcinoembryonic antigen of human digestive system cancer has been the most studied oncofetal antigen. Originally thought to be a specific antigen of the fetal digestive tract as well as cancer of the colon, carcinoembryonic antigen is now known to occur normally in feces and secretions from the pancreas and bile ducts. It also appears in the plasma due to cigarette smoking and a diverse group of neoplastic and non-neoplastic conditions, including cancers of the colon, pancreas, stomach, lung, and breast; alcoholic cirrhosis; pancreatitis; inflammatory bowel disease; and rectal polyps.

Cancer prognosis. The carcinoembryonic antigen levels in the blood are one of the factors that doctors consider when determining the prognosis, or most likely outcome, of cancer. The carcinoembryonic antigen test is ordered for patients with known cancers, commonly cancer of the gastrointestinal system. These include cancers of the colon, rectum, stomach (gastric cancer), esophagus, liver, and pancreas. It is also used to predict the outcome of cancers of the breast, lung, and prostate gland. In general, a higher carcinoembryonic antigen level predicts a more severe disease, one that is less likely to be curable. However, the carcinoembryonic antigen

test does not provide clear-cut information; therefore, the results are usually considered along with other laboratory or imaging studies to follow the course of the disease.

Cancer treatment response. Once cancer treatment has begun, carcinoembryonic antigen tests have a valuable role in monitoring the patient's progress. The primary use of carcinoembryonic antigen is in monitoring response to treatment of colorectal cancer. A decreasing carcinoembryonic antigen level means therapy is effective in fighting the cancer. A stable or increasing carcinoembryonic antigen level may mean that the treatment is not working, or that the tumor is growing. Serial carcinoembryonic antigen measurements, that is, tests done over a period of time, are the most useful. A single test result is difficult to evaluate, but a number of tests, done weeks apart, show trends in disease progression or regression.

Cancer recurrence. Carcinoembryonic antigen tests are also used to help detect recurrence of a cancer after surgery or other treatment has been completed. A rising carcinoembryonic antigen level may be the first sign of cancer return, and may show up months before other test results or patient symptoms would raise concern. Unfortunately, this does not always mean the recurrent cancer can be cured. For example, only a small percentage of patients with colorectal cancers and rising carcinoembryonic antigen levels will benefit from another surgical exploration. Those with recurrence in the same area as the original cancer, or with a single metastatic tumor in the liver or lung, have a chance that surgery will eliminate the disease. Patients with more widespread return of the cancer are generally not treatable with surgery. However, the carcinoembryonic antigen test does not distinguish between the two groups.

Nonstandard treatments. Patients who are most likely to benefit from nonstandard treatments, such as bone marrow transplants, may be identified on the basis of carcinoembryonic antigen values combined with other test results. Carcinoembryonic antigen levels also may be one of the criteria for determining whether the patient will benefit from more expensive studies, such as computerized tomography (CT) or magnetic resonance imaging (MRI). *See* COMPUTERIZED TOMOGRAPHY; MAGNETIC RESONANCE; MEDICAL IMAGING; TRANSPLANTATION BIOLOGY.

Pancreatic oncofetal antigen. Several other less defined oncofetal antigens, such as the pancreatic oncofetal antigen, have been described. It is associated with pancreatic cancer and is found in extracts of the fetal pancreas, adult human pancreatic tumors, sera from individuals with carcinoma of the pancreas, and—to a lesser degree—in persons with other malignancies. *See* ANTIGEN; ONCOLOGY.

Zoë Cohen; T. Ming Chu

Bibliography. H. Magdelenat,Tumour markers in oncology: Past, present and future, *J. Immunol. Meth.*, 150(1–2):133–43, Jun 24, 1992; R. G. McKinnell et al., *The Biological Basis of Cancer*, 3d ed., Cambridge University Press, 1998, reprint 2000; R. W. Ruddon, *Cancer Biology*, 3d ed., Oxford University Press, 1995.

Oncogenes

Genes that contribute to the conversion of a normal cell into a cancerous cell. Oncogenes can derive from cellular genes that undergo mutations that alter their expression or activity, or from viruses that carry oncogenes within their genome and are transferred into the cell by infection. *See* ANIMAL VIRUS; GENE.

Genetic basis of neoplasia. Most, if not all, neoplasms (cancers) arise when individual cells within the body suffer irreversible genetic damage that leads to unrestrained cell growth, a block in the normal process of differentiation, or programmed cell death (apoptosis). Since genes consist of segments of deoxyribonucleic acid (DNA) that are linked in a specific order along each chromosome, any agent or process that breaks the DNA or alters the individual chemical subunits of the DNA can cause genetic damage.

Genetic lesions in neoplastic cells can affect two classes of genes. The first genes identified were called oncogenes, and the mutations that alter these genes occur on only one of the two similar chromosomes in a cell. The unaffected chromosome retains the normal cellular gene, whereas the mutant form of the gene on the affected chromosome overrides the normal gene function and promotes neoplastic growth. In contrast to oncogenes, mutations in cancer cells can also inactivate or even delete genes whose function is required for normal cell growth. This second class of genes, called tumor suppressors, appears to restrain cell growth or prevent the accumulation of mutations, and a cell remains normal as long as at least one copy of the tumor suppressor gene is functional. However, if both normal copies of the gene are inactivated by mutation or lost from the cell, neoplastic growth can ensue.

Oncogenic mutations in cellular genes are never inherited from parent to child, presumably because the mutations are so detrimental to normal development that an embryo with the mutation would not survive. Instead, these mutations arise during the lifetime of an individual through spontaneous mutation or exposure to environmental carcinogens, both of which damage DNA. In contrast, a small fraction of human tumors are caused by the inheritance of a mutation in a tumor suppressor, giving some families a predisposition to cancer. *See* DEOXYRIBONUCLEIC ACID (DNA); MUTAGENS AND CARCINOGENS; MUTATION; PREDICTIVE GENETICS OF CANCER; TUMOR.

Mechanisms of oncogene formation. Cellular genes have highly specific functions and patterns of expression in normal cells, and the mutations that convert a normal gene into an oncogene within a tumor cell can alter either of these properties. Four major types of oncogene-forming lesions are commonly found in tumor cells: (1) Point mutations can change individual bases within the segment of the gene that encodes a protein, and the resulting mutant protein has enhanced or altered activity that contributes to cancer cell growth (see **illustration**). (2) Chromosomes can break near specific genes and rejoin in such a

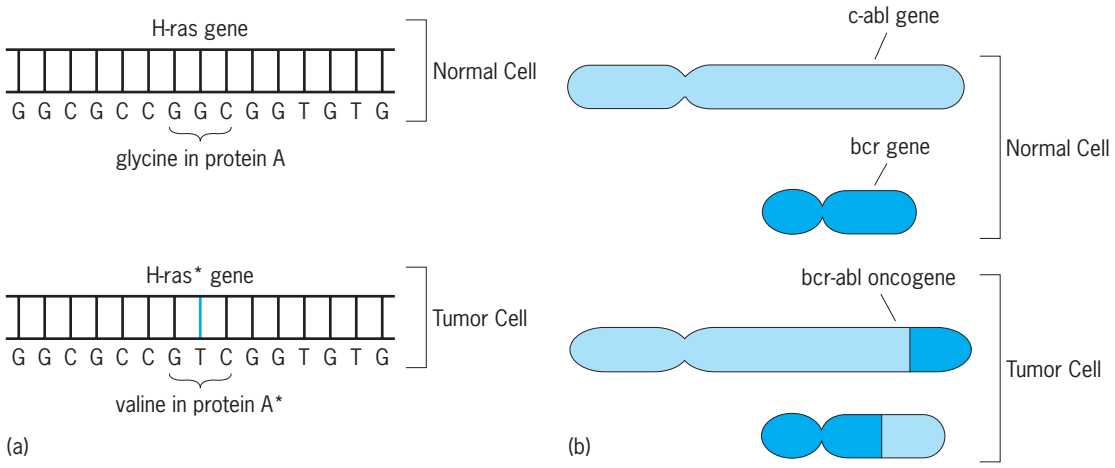

Two common mechanisms of oncogene activation in cancer cells. (*a*) DNA mutation. Mutation of the DNA sequence that encodes a protein involved in growth signals. In this example, the GGC codon directs the amino acid glycine as the 12th amino acid in the H-ras protein. Mutation of a single base changes the codon to GTC and results in a valine as the 12th amino acid of the oncogenic H-ras protein. This mutant H-ras protein was one of the first and most potent oncogenes discovered. (*b*) Chromosomal translocation. Chromosomes can break and rejoin near specific genes. In this example, the *c-abl* tyrosine kinase gene on chromosome 9 breaks and rejoins to the *bcr* gene on chromosome 22. The resulting fusion protein has high tyrosine kinase enzyme activity and contributes to the development of chronic myelogenous leukemia. This chromosomal translocation was the first to be found in human cancer.

way that an oncogene is created through an abnormally high level of gene expression or by the loss of the gene's normal mode of regulation. These lesions are referred to as chromosomal translocations, and they are particularly common in leukemias and lymphomas. Chromosomal translocations can also fuse two genes together so that a new oncogene with altered or enhanced activity is created (see illus.). (3) Oncogenes can be created by the localized amplification of small chromosomal domains, which leads to extra copies of the gene within the amplified segment and elevated levels of oncogene-encoded ribonucleic acid (RNA) and protein. (4) Viruses or mobile genetic elements can be inserted into the chromosome near a specific gene and alter the expression of the gene by imposing new regulatory signals, a process called insertional mutagenesis. *See* CHROMOSOME ABERRATION.

Viral oncogenes. Some viruses contain oncogenes that are part of the viral genetic material and not derived from the cell. Although most human cancer is not associated with infectious agents such as viruses, some cancers have been shown to be caused at least in part by viruses. The most common is cervical carcinoma, where the causative agent is human papilloma virus (HPV). The HPV genome contains two viral oncogenes that contribute to the outgrowth of cancer cells. Rather than providing enzymatic or gene regulatory activities of their own, the HPV oncogenes function by binding to cellular tumor suppressor proteins and inactivating their function. Since the tumor suppressor protein is rendered nonfunctional, the viral infection promotes unrestrained cell growth or genetic instability by the same mechanisms that might otherwise occur through genetic lesions that mutate or delete the tumor suppressor genes themselves. Immunization against HPV has recently become a reality, and eliminating this virus from the population should dramatically reduce the incidence of cervical cancer. *See* ANIMAL VIRUS; TUMOR VIRUSES.

Oncogene functions. Normal cells divide or remain quiescent as the result of diverse extracellular signals (such as growth factors and cell-cell interactions). These signals are interpreted by receptors and cytoplasmic factors, and eventually the signals lead to reprogramming of gene expression in the nucleus. Thus, it is logical that the majority of oncogenes encode proteins that are components of this signal transduction pathway and they inappropriately activate the normal growth signals. Many oncogene-encoded proteins localize to the plasma membrane and are enzymes that phosphorylate other proteins (protein kinases), altering the activity of these substrate proteins. Other oncogenic proteins function as regulatory subunits of enzymes, and the mutations signal the enzymes to be active continuously rather than in a regulated manner. The oncogene-encoded proteins that localize to the nucleus are largely DNA-binding proteins that regulate specific sets of cellular genes involved in growth control. Finally, oncogenes can function not by promoting cell growth but by blocking apoptosis. Since many cells within the body (especially blood cells) are programmed to die after a limited lifespan, overexpression of a protein that blocks this programmed cell death can also contribute to cancerous growth. *See* CELL SENESCENCE AND DEATH; GENE ACTION; GENETIC ENGINEERING.

Oncogene-targeted cancer therapy. The recognition that there are specific genetic mutations that drive the growth of tumor cells provides a new opportunity to develop drugs that target cancer cells. Since these mutations exist only in the tumor, a drug designed to inhibit the mutated or overexpressed oncogenic protein can slow the growth of tumor cells (or even kill them) with reduced side effects on normal

tissues. Several new drugs have been introduced in recent years that selectively target oncogenic mutations, and they are proving to be highly efficacious against certain forms of cancer. For example, antibodies directed against the epidermal growth factor (EGF) receptor selectively target breast cancer cells with amplification and overexpression of the EGF receptor gene. Another successful therapy targets the abnormal enzyme that is created by a chromosomal translocation in chronic myelogenous leukemia. These drugs and many others in the pharmaceutical pipeline will be available to treat the dozens of different cancers, each of which may have a unique collection of causative mutations. Therapies will be chosen based on the underlying molecular pathology of each tumor. Oncogene-targeted therapy promises a new era in the treatment of the cancers that afflict a large fraction of the population. *See* CANCER (MEDICINE); GENE ACTION; ONCOLOGY. Michael Cole

Bibliography. D. W. Kufe et al. (eds.), *Cancer Medicine*, 6th ed., Lewiston, New York [distributor], BC Decker, Hamilton, Ontario, 2003; B. Vogelstein and K. W. Kinzler (eds.), *The Genetic Basis of Human Cancer*, 2d ed., McGraw-Hill, Medical Pub. Div., New York, 2002.

Oncology

The study of cancer. There are five major areas of oncology: etiology, prevention, biology, diagnosis, and treatment. As a clinical discipline, it draws upon a wide variety of medical specialties; as a research discipline, oncology also involves specialists in many areas of biology and in a variety of other scientific areas. The approach to the study of cancer is multidisciplinary because cancer is a fundamental problem in biology. Oncology has led to major progress in the understanding not only of cancer but also of normal biology.

Cancer defies simple definition. It is a disease that develops when the orderly relationship of cell division and cell differentiation becomes disordered. In general, tissues of the body are composed of cells with distinct capabilities. Some cells are specialized to conduct the functions of the particular tissue or organ and such cells do not divide. Division is normally restricted to less well-differentiated stem cells within tissues that proliferate to replace senescent mature, well-differentiated cells. As a proliferating stem cell differentiates, it loses the capacity to divide. In cancer, dividing cells seem to lose the capacity to differentiate, and they acquire the ability to invade through basement membranes and spread (metastasize) to many areas of the body through the bloodstream or lymphatics. Cancer is usually clonal, that is, it develops initially in a single cell. That abnormal cell then produces progeny that may behave rather heterogeneously. Some progeny continue to divide, some develop the capacity to metastasize, and some develop resistance to therapeutic agents. This single cell and its progeny, if unchecked, typically lead to the death of thehost.

Causes of Cancer

What causes cancer? Cancer is generally thought to result from one or more permanent genetic changes in a cell. In some cells a single mutational event can lead to neoplastic transformation, but for most tumors it appears that carcinogenesis is a multistep process. The notion that cancer reflects a permanent genetic change is not absolute, as some cancer cells can develop into normal, functioning, and differentiated progeny if placed under the inductive influence of a fertilized egg. Thus, some cancer cells retain responsiveness to differentiation signals.

Although some rare congenital conditions lead to cancer in infancy, the vast majority of human cancers arise as a result of the complex interplay between genetic and environmental factors. Therefore, estimates of the amount of cancer attributable to each cause seem fruitless. Without question, there are forms of cancer clearly related to particular environmental exposures (**Table 1**); it is equally clear, however, that these factors act on a genetic substrate that may be either susceptible or resistant to the development of cancer. It is well known that lung cancer is much more common in smokers than in nonsmokers. However, the incidence of lung cancer is about 30 times higher among smokers with high lung levels of aryl hydrocarbon hydroxylase, an enzyme that metabolizes benzo(*a*)pyrene in tobacco smoke to highly carcinogenic epoxides, than among other smokers. In this instance, the environmental agent, tobacco smoke, interacts with the genetically determined high level of metabolizing enzyme to facilitate the carcinogenesis.

Genetic factors. The emergence of cancer appears to involve the accumulation of genetic damage in a target tissue. In one study of genetic changes in colon cancer, it appeared that progression from normal colonic epithelium to small adenomatous polyp to large adenomatous polyp to overt carcinoma was accompanied by mutation in the *ras* oncogene followed by deletion of sequences on chromosomes 5 and then 18 and 17. Such complex, sequential genetic changes specific to tissues appear to underlie the progression to cancer. In some tissues, like colon, each genetic lesion produces a clinical abnormality. In other tissues (such as the ovary), stepwise progression has not been documented.

Multistep progression is quite complicated to study in experimental systems. Much work has focused on the identification, isolation, and characterization of oncogenes, which have the ability to transform non-neoplastic cells into cancer cells. More than 50 bona fide or putative oncogenes have been characterized and mapped throughout the human genome. Many were identified because they accounted for the transforming ability of animal retroviruses. However, over time it has become clear that most oncogene sequences are a normal part of the genome, and such endogenous oncogenes, termed proto-oncogenes, serve important roles in cell proliferation and differentiation. Proto-oncogenes encode for growth factors, growth factor receptors, nuclear

TABLE 1. Selected environmental causes of human cancer

Agent	Site or type of cancer
Life-style risk factors	
Tobacco, smoking	Lung, larynx, mouth, pharynx, esophagus, bladder, pancreas, kidney, cervix, breast
Tobacco, smokeless	Mouth, oral cavity
Alcohol	Mouth, pharynx, esophagus, larynx, liver, breast, colon (?)
Sunlight	Skin, lip
Anabolic steroids	Liver
Phenacetin	Renal pelvis
Occupational risk factors	
Asbestos	Lung, pleura, peritoneum
Benzene	Leukemia
Benzo(*a*)pyrene	Lung, skin
Arsenic	Lung, skin, liver angiosarcoma
Aromatic amines	Bladder
Chromium	Lung
Nickel dust	Lung, nasal sinuses
Vinyl chloride	Liver angiosarcoma
Wood dust	Nasal sinuses, Hodgkin's disease
Herbicides	Lymphoma, soft tissue sarcoma
Leather	Nasal sinuses
Infectious risk factors	
Helicobacter pylori	Gastric MALT lymphoma, gastric cancer
Schistosoma haematobium	Squamous carcinoma of the bladder
Clonorchis sinensis	Cholangiocarcinoma of the liver
Epstein-Barr virus	Lymphomas, nasopharyngeal carcinoma
Hepatitis-B virus	Hepatocellular carcinoma
Human T-cell lymphotrophic virus	Adult T-cell leukemia
Papilloma virus	Cervix, skin, vulva
Iatrogenic risk factors	
Chronic alkylating agents	Leukemia, bladder
Radiation	Nearly any site
Estrogens, conjugated	Endometrium
Estrogens, synthetic	Transplacental cervix, vagina
Immunosuppressive agents	Lymphoma, skin, Kaposi's sarcoma
Thorium oxide	Liver angiosarcoma
Chlornaphazine	Bladder

proteins, regulatory enzymes such as tyrosine kinases, and components of intracellular pathways of signal transduction mediated by growth and differentiation factors acting at the cell membrane. When proto-oncogenes are normally regulated, their functions are essential to the cell. However, there are at least six mechanisms by which proto-oncogenes may be altered so that their expression results in neoplastic transformation. These mechanisms include mutation, amplification, translocation, promoter insertion, recombination, and insertion into the cell membrane. They generally result in the escape of a proto-oncogene from its normal controls and lead to aberrant expression or to expression of an altered product with unregulated functions. Some oncogenes require another, complementary oncogene to transform a normal cell. Their aberrant expression commonly results in autonomous cell proliferation and escape of the cell from the normal mechanisms that control proliferation. *See* ONCOGENES.

One or more genetic mechanisms may prohibit oncogenes from causing transformation. Many tumor types are associated with nonrandom chromosomal deletions. The best-studied example is the retinoblastoma gene on chromosome 13, whose gene product is capable of interacting with transforming viral proteins from adenovirus and papilloma virus families. It appears that the transforming genes are more efficient at transforming their target if the retinoblastoma gene product is absent. When either the product or the gene itself is inserted into malignant cells that lack it, the malignant phenotype of the cell is reversed. Thus, these recessive genes appear to function as tumor suppressor genes or antioncogenes, and their absence creates a cell more vulnerable to malignant transformation. *See* HUMAN GENETICS.

There is a close link between oncogenes and growth factors. Unlike normal cells, whose proliferation is often regulated by the availability of a trophic growth factor produced by a distinct cell type, many tumor cells are capable of autocrine growth, that is, they secrete their own growth factors and proliferate autonomously. Another mechanism of uncontrolled proliferation is escape from negative growth control. Many cell types can be inhibited by negative growth factors such as transforming growth factor-beta. Some tumor cells fail to activate this factor (which requires cleavage of an inactive precursor to become active), and others fail to express receptors for the growth-inhibitory polypeptides. The failure of some tumor cells to differentiate normally may also be related to alterations in the expression or function of receptors for differentiation factors. *See* GROWTH FACTOR.

Some cancers are increased in families. About 10% of breast cancers develop in women with a strong family history of breast cancer in first-degree relatives. In about half of these families, germline mutations in *BRCA1*, a gene on chromosome 17 encoding a protein involved in DNA repair, result in a

nonfunctional gene product. Women with *BRCA1* mutations have an 80% lifetime risk of developing breast cancer and a 33% chance of developing ovarian cancer. Women with a strong family history of breast cancer should be encouraged to undergo genetic screening. Those of Ashkenazi Jewish descent may have a specific *BRCA1* mutation, deletion of adenine and guanine at position 185. Currently, women with *BRCA1* mutations are advised to undergo bilateral mastectomy and bilateral salpingo-oophorectomy to prevent the cancers to which they are predisposed. Another breast cancer susceptibility gene has been mapped to *BRCA2* on chromosome 13. *BRCA2* mutations are about half as common as *BRCA1* mutations and predispose both male and female carriers to breast cancer. Familial predispositions to colorectal cancer include familial polyposis coli (mutations in the *APC* gene on chromosome 5) and Lynch syndrome (hereditary non-polyposis colorectal cancer; mutations in mismatch repair genes on chromosomes 2 and 3). These syndromes account for less than 10% of colorectal cancers. Prophylactic total colectomy is often recommended.

Environmental factors. Environmental factors involved in the development of cancers can be chemical, physical, or biological carcinogenic agents (Table 1). At least three stages occur in the natural history of cancer development from environmental factors. The first stage is initiation, which is a specific alteration in the deoxyribonucleic acid (DNA) of a target cell; environmental agents may act by inducing expression of oncogenes. Many chemical carcinogens are capable of inducing mutation, which can result in activation of the transforming potential of an oncogene. The second phase, promotion, involves the reversible stimulation of expansion of the initiated cell or the reversible alteration of gene expression in that cell or its progeny. Agents that serve as promoters are often incapable of inducing cancers on their own, but significantly enhance the development of cancers by initiating agents. Because promotion is thought to be reversible, it is a target for prevention. The final phase of carcinogenesis is progression. It is characterized by the development of aneuploidy and clonal variation in the tumor; these in turn result in invasiveness, metastasis and growth advantage.

Chemical carcinogens. The first clear demonstration of an environmental risk factor for the development of a cancer was made in 1775, when it was noted that chimney sweeps, a group with a considerable occupational exposure to coal tar, had a high incidence of scrotal cancer. Over 140 years later, the precise cause of the cancer in coal tar was found to be benzo(*a*)pyrene.

Epidemiologic studies attempt to identify factors associated with risk, and experimental studies define the precise nature of the factor. Chemical compounds capable of inducing cancer include polycyclic hydrocarbons, aromatic amines, and alkylating agents. Many chemical carcinogens are actually prodrugs activated to become carcinogenic by the body's metabolic machinery, such as the microsomal enzymes that, ironically, evolved to detoxify toxic compounds. The active metabolite of many chemical carcinogens is a free-radical compound, and nearly all chemical carcinogens have been found to interact directly with DNA, forming adducts that can result in errors in base sequence during replication.

Tobacco, the most important environmental carcinogen, accounts for an increasing number of deaths each year from cancer as well as heart and nonmalignant respiratory diseases. Perhaps the most disturbing information about smoking is the evidence that nonsmokers who inhale exhaled smoke are also at risk for smoking-associated cancers.

The role of dietary factors in carcinogenesis is considerably more controversial. Ingestion of alcohol has been associated with increased risk of cancers of the esophagus, pharynx, larynx, and mouth—the same sites associated with cigarette smoking. Some data implicate drinking alcohol with breast and colon cancer, but the carcinogenic potential of other foodstuffs is unclear and the data linking dietary carcinogens to cancer are generally weak. Nitrosamines in smoked meat and salted fish are probably carcinogenic. However, there are foods that appear to have the capacity to detoxify the carcinogens that occur in other foods. Epidemiologic evidence suggests that high-fat and high-calorie diets are associated with an increased risk of colon and breast cancer. However, causal association has not been conclusively shown.

Physical carcinogens. The three major physical carcinogens are ionizing radiation, ultraviolet radiation, and foreign bodies. Ionizing radiation can originate from many sources and can occur in many forms, but the two major categories are electromagnetic radiation, from x-rays and gamma rays, and particle radiation, originating in electrons, protons, neutrons, and alpha particles. The rate at which energy is deposited in the tissue is characteristic of the type of radiation and is called the linear energy transfer. Electromagnetic radiation is weakly ionizing and has a low linear-energy-transfer value, but particle radiation is densely ionizing and has a high linear-energy-transfer value. The biological damage done by radiation is higher with high-linear-energy-transfer radiations than with low-linear-energy-transfer radiations. The major physical features that determine the risk of radiation-induced cancer are linear-energy-transfer value, dose, dose rate, and fractionation. Radiation may induce cancers at any site. More than 80% of radiation exposure is from natural sources such as cosmic rays, terrestrial gamma rays, and radon. Radon may also emanate from the ground and from building materials, disperse in the air, and decay into short-lived aerosolized alpha particles with high-linear-energy-transfer characteristics that are capable of inducing lung cancer when inhaled. Radon daughters, such as polonium-214 and -218, could cause serious radiation injury. Radon exposure varies markedly across the United States, but on average probably accounts for slightly over half the exposure to radiation. Medical uses of radiation account for about 18% of the total radiation exposure. *See* LINEAR ENERGY TRANSFER (BIOLOGY); RADON.

It is known that radiation can induce mutations in DNA and can activate oncogenes, but the precise mechanism of radiation-induced carcinogenesis is unclear. Also, it is not clear why different forms of radiation-induced cancer have such a long latency. For example, cancers of the breast and thyroid can develop more than 20 years after radiation therapy of the chest and neck for Hodgkin's disease. For leukemia, the latent period appears to be much shorter (4–6 years), and the risk for leukemia decreases after 10 years following radiation. Certain genetic conditions, such as familial dysplastic nevus syndrome and ataxia-telangiectasia, may be associated with an increased susceptibility to radiation-induced cancer, but mechanisms are not known. Radiation-induced solid tumors may emerge in radiation-exposed patients at about the same age that tumors of the same tissue develop in nonirradiated persons, but with increased frequency, suggesting that a genetically determined series of processes leads to neoplastic transformation in a particular tissue. Radiation acts to increase the probability that a tissue will undergo transformation, rather than to accelerate the transformation process.

Ultraviolet radiation, mainly from the Sun, is carcinogenic to skin. The incidence of skin cancers other than melanoma is much higher in southern latitudes and on sites exposed to the Sun, and the incidence of melanoma may also be augmented. The UVB portion (280–320 namometers) of the ultraviolet spectrum is most damaging to tissues. Ultraviolet radiation induces the formation of pyrimidine dimers (usually between thymines), which brings about alterations in the normal sequence of bases in DNA. Individuals with red hair and with certain genetic conditions, such as xeroderma pigmentosum, are less able to repair ultraviolet-induced damage and are more susceptible to Sun-induced skin cancers than the general population. Protection against ultraviolet radiation is offered by the ozone layer of the atmosphere and by pigmentation of the skin. Ultraviolet radiation also exerts systemic effects by altering the immune function. *See* RADIATION BIOLOGY; RADIATION INJURY (BIOLOGY).

Foreign bodies are only an occasional cause of tumors, and they rarely induce sarcomas and neoplasms in the site into which they are introduced; and in animal models the composition of the foreign body is less important than its size and shape. The same materials are more carcinogenic if they are fibrous than if they are powdered, porous, or perforated. The malignant transformation is probably related to a derangement during the connective tissue reaction to the foreign body. The foreign body that is most carcinogenic in humans is asbestos, a natural mineral fiber that is inhaled and is associated with lung cancer and mesotheliomas. Other potentially carcinogenic fibers include synthetic vitreous and crystalline fibers. *See* ASBESTOS.

Biological carcinogens. Although viral carcinogenesis is widespread in nature, examples in humans are not numerous, and knowledge of oncogenes has relied on the prevalence of tumors induced by retroviruses in animals; but the precise mechanism of carcinogenesis is not yet clear. There are four documented oncogenic virus types in humans: the hepadnavirus family (specifically hepatitis B virus), associated with hepatocellular carcinoma; the herpesvirus family (specifically Epstein-Barr virus), associated with Burkitt's lymphoma, other lymphomas in the setting of immunodeficiency, nasopharyngeal carcinoma; some papilloma viruses, associated with cervical cancer and skin cancer; and the human retrovirus family (specifically HTLV-I), associated with adult T-cell leukemia/lymphoma. Herpesvirus 6 has been implicated in some lymphomas. *See* ANIMAL VIRUS; EPSTEIN-BARR VIRUS; TUMOR VIRUSES.

Chronic bacterial infections may lead to cancers in certain sites. For example, *Helicobacter pylori* gastritis can lead to gastric MALT (mucosa-associated lymphatic tissue) lymphoma and to gastric adenocarcinoma. Infections with certain parasites can also lead to carcinogenesis, perhaps much like foreign bodies do. An example is bilharzial squamous-cell bladder cancer, which follows chronic bladder infection with *Schistosoma haematobium*.

Hormones may also be considered biological carcinogens. Estrogens can cause endometrial cancer. The synthetic estrogen diethylstilbestrol (DES), administered to women in the 1950s to prevent spontaneous abortion, was found to result in clear-cell carcinoma of the vagina in some of their female offspring between the ages of 15 and 30. Transplacental carcinogenesis in this setting is not understood. *See* ESTROGEN; HORMONE; MUTAGENS AND CARCINOGENS.

Cancer Prevention

An obvious starting point for cancer prevention is avoidance of environmental agents that contribute to carcinogenesis. Eliminating use of tobacco and alcohol would reduce cancer mortality by more than 30%. Skin cancer prevention is a simple matter of using the available tools (clothing and sunscreens) to protect the skin from Sun damage. Maintaining a dark and even tan can result in premature aging of the skin as well as cancer. Thinning of the ozone layer may one day necessitate routine application of sunscreen products before an individual spends any time outdoors.

The role of diet in cancer prevention is controversial. Epidemologic evidence suggests a particularly strong link between a high-fat, high-calorie, low-fiber diet and an increased risk of colon cancer. But a change to a low-fat, low-calorie, high-fiber diet may not alter the risk. Calorie restriction and exercise may reduce the carcinogenicity of carcinogenic agents. The addition to the diet of carotenoids, selenium, vitamins A, D, and E, and some short-chain fatty acids may prevent cancers in high-risk populations, but there is no evidence that any dietary supplement will prevent cancer. *See* NUTRITION.

There are a variety of clinical settings in which surgery may prevent cancer. Repositioning of undescended testes may prevent testicular cancer, and resection of the large intestine may prevent colon

cancer in individuals with familial polyposis, familial colon cancer, or ulcerative colitis. Surgical removal of the thyroid will prevent medullary carcinoma in individuals with certain types of multiple endocrine neoplasia, breast removal can be preventive in familial breast cancer, and removal of the ovaries can prevent cancer in familial ovarian cancer. Preventive medical interventions are being evaluated, including the nonsteroidal anti-inflammatory agent sulindac to block the development of neoplastic polyps in familial polyposis, and tamoxifen and/or aromatase inhibitors in familial breast cancer.

Cancer Biology

The study of cancer biology picks up where cancer etiology leaves off, namely, at the point where the tumor has developed into a clonal cluster of autonomously proliferating cells. The pathological correlate of this stage of tumor development is carcinoma in situ; a condition in which no tissue destruction is evident, but atypical-appearing cancer cells are present at their site of origin. The transition from carcinoma in situ to locally invasive cancer is accompanied by dissolution of the basement membrane, penetration of tumor cells through the membrane and into the supportive tissues, and disruption of the supportive tissues. The tumor cell is thought to bind to the basement membrane and tissue matrix through cell surface receptors for laminin and fibronectin. The tumor cell secretes enzymes that dissolve the membrane or matrix and also secretes an autocrine motility factor that helps the cell move through the membrane and matrix. The tumor may also be affected by chemotactic factors in the host tissues. Expansion of the primary tumor in locally invasive cancer is always accompanied by the development of blood vessels, which are often defective and easily invaded by individual and clumped tumor cells. The tumor cells can also invade regional blood vessels and lymphatics and circulate throughout the body, attaching to endothelium in a distant organ site, inducing retraction of the endothelium, and becoming attached to the endothelial basement membrane. Once attached to the basement membrane, the tumor cells are covered over by the endothelial cells and effectively separated from the flow of blood. Local dissolution of the basement membrane then occurs, allowing the tumor to completely spread into the tissue and reestablish a blood flow in the breached vessel. As it grows, more blood vessel development nourishes the enlarging tumor.

During metastasis, tumor cells must overcome host defenses. They have various mechanisms to do so. For example, they produce new cell surface receptors to facilitate basement membrane and matrix binding; make new enzymes such as collagenases, serine proteases, metalloproteinases, cysteine proteinases, and endoglycosidases to facilitate their invasiveness; and secrete motility factors to enable them to move through the holes and pathways created by their enzymes. They avoid detection by the immune system through a variety of techniques. Unlike animal tumors, most human tumors are poorly immunogenic, and when they do express a tumor-related antigen, its expression is often heterogeneous or it is shed by the tumor cell. Tumor cells may fail to express class I major histocompatibility (MHC) antigens. Since cytotoxic T cells recognize tumors only through MHC antigens, this makes the T cells blind to the presence of the tumor. Tumor cells often produce factors that are immunosuppressive. *See* CELLULAR IMMUNOLOGY; HISTOCOMPATIBILITY.

An unexplained feature of metastasis is the propensity of certain tumor types to spread to specific organs. There must be one or more mechanisms by which some metastatic tumor cells distinguish lung from liver, for example, but the nature of the mechanism is not known.

Cancer Detection, Diagnosis, and Treatment

The curability of a cancer is inversely proportional to its size at diagnosis, and so early diagnosis is critical. However, at the earliest point that a tumor is likely to be detected on physical exam (about 0.4 in. or 1 cm) it already contains 10^8–10^9 tumor cells and has undergone about 30 doublings. A lethal tumor burden occurs at around 40 doublings (about 10^{12} cells); thus, most of a tumor's natural history occurs before detection. Most tumors do not grow exponentially. Exponential growth, which is typical of bacteria, occurs when a cell divides to produce two cells capable of division. Tumors follow Gompertzian growth kinetics, in which not all progeny of a dividing cell are capable of division. Generally the growth fraction of a solid tumor decreases with the size of the tumor (see **illustration**), and contrary to common belief, tumors generally do not proliferate more actively than normal tissues. Not all the factors that

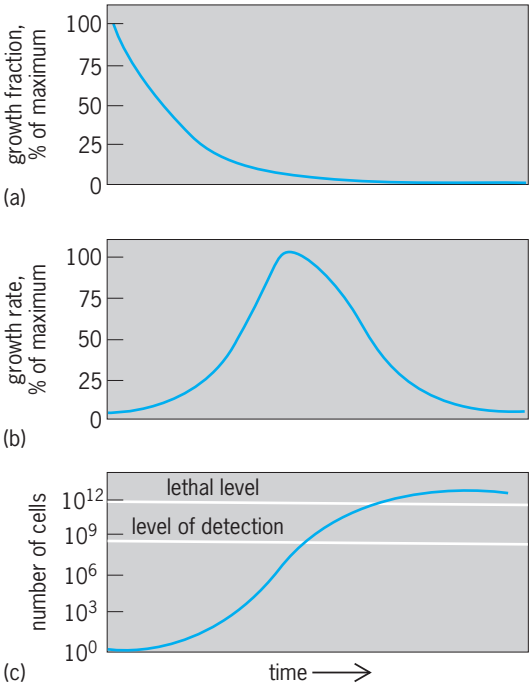

Gompertzian growth kinetics. (*a*) Growth fraction declines exponentially. (*b*) Growth rate of a tumor, a measure of the absolute number of dividing cells plotted as a function of time. (*c*) Tumor size as a function of time.

TABLE 2. Guidelines for cancer screening

Location	Procedures	Baseline age	Screening frequency
Breast	Mammography	35–40	Every 1–2 years between ages 40 and 49; annually thereafter
Cervix	Papanicolaou (Pap) smear + pelvic examination	18 (or earlier if sexually active)	Annually; every 3 years after three consecutive normal results
Colon	Fecal occult blood test + sigmoidoscopy	40	Every 3–5 years following normal baseline examination

regulate tumor cell growth are defined. However, at least some of the tumor cells die or are hypoxic, and others have sustained too much genetic damage to divide. Tumor doubling times range from 3 to 90 weeks in human tumors and, in general, metastases double more quickly than primary lesions, largely because of the selection metastases have undergone to become established.

Tumor detection. There are two major strategies to detect tumors at the earliest possible stage in their history: responding to the seven warning signals of cancer and screening populations at high risk. The seven danger signals of cancer are (1) unusual bleeding or discharge, (2) a lump or thickening in the breast or elsewhere, (3) a sore that does not heal, (4) change in bowel or bladder habits, (5) persistent hoarseness or cough, (6) persistent indigestion or difficulty in swallowing, and (7) change in a wart or mole.

Screening has been enormously successful in reducing mortality from cancer of the cervix, but it has taken many years to develop appropriate guidelines for screening. Large-scale screening can be costly, and the method used must be safe, inexpensive, and accurate. For example, the Papanicolaou (Pap) smear for detecting cervical cancer relies heavily on the interpretive abilities of the technician reading the cytology slide. Screening guidelines for other types of cancer are considerably more controversial (**Table 2**). False positive screening tests increase patient risk and expense associated with follow-up testing. Certain techniques may reduce the size of the population that needs screening, and the screening tests may be able to detect cancers at an earlier stage. The detection of genetic lesions in patient samples promises to further enhance the specificity and sensitivity of screening. For example, screening for colorectal cancer currently depends heavily on finding occult blood in the stool. Detection of mutated genes in stool specimens may improve the detection of colorectal cancer; abnormal cells are continuously shed from tumors, but bleeding in such cancers is only intermittent. *See* MAMMOGRAPHY; RADIOGRAPHY.

Diagnosis. The diagnosis of cancer depends on the careful examination of biopsy material. Light-microscopic analysis of hematoxylin- and eosin-stained material is the primary technique. Most cancers are readily diagnosed in this fashion; the grading of histologic (tissue) atypia can be an important predictor of outcome. However, other techniques can assist the pathologist in defining the diagnosis and in conveying useful prognostic information; special stains can detect intracellular markers of particular tumors, such as melanin and nerve-specific enolase; electron microscopy can reveal pathognomonic cellular organelles; and immunohistochemistry can define cell surface characteristics and cytoplasmic contents of prognostic importance. The detection of genetic lesions characteristic of certain cancers also aids in diagnosis. *See* CLINICAL PATHOLOGY.

Cancers arising in tissues having ectodermal or endodermal origins are generally called carcinomas; those derived from glands are called adenocarcinomas. Cancers arising in tissues derived from mesoderm are called sarcomas; those of lymphohematopoietic origin are lymphomas and leukemias. The cardinal microscopic features of malignancy are anaplasia, invasion, and metastasis.

Once a diagnosis of cancer is made, it is critical to determine the extent to which the disease has spread. This is called staging. It is distinct from grading, which is an assessment of histologic atypia performed with a microscope. Staging entails performing a careful physical examination, various radiographic studies (computed tomography, lymphography, nuclear medicine scanning), and perhaps surgical procedures (biopsies, endoscopies) to examine those sites to which a particular tumor type is most likely to spread. For example, patients with breast cancer often undergo evaluation of the liver, brain, and bones to search for metastatic disease, whereas patients with lymphoma generally require assessment of lymph node groups, bone marrow, and liver. Often the results of such staging tests determine the nature and extent of therapy.

For some tumors, assessment of serum levels of proteins or hormones can act as a measure of total body tumor burden. Such tumor markers can complement but not replace other staging techniques. In some instances they can be used to follow the progress of treatment and detect early relapse (**Table 3**).

Treatment. There are four major approaches to cancer treatment: surgery, radiation therapy, chemotherapy, and biological therapy. These modalities are often used together with additive or synergistic effects. Surgery and radiation therapy are most effective in curing localized tumors and together result in the cure of about 50% of all newly diagnosed cases. Once the cancer has spread to regional nodes or distant sites, it is generally incurable with the use of local therapies alone. Systemic administration of a combination of chemotherapeutic agents may cure another 15–18% of all patients.

Surgery. This treatment modality is used for the following purposes: (1) definitive treatment of primary

TABLE 3. Selected serum tumor markers useful in assessing disease status

Marker	Disease
Human chorionic gonadotrophin (HCG)	Choriocarcinoma
	Testicular cancer
Alpha fetoprotein (AFP)	Testicular cancer
	Hepatocellular carcinoma
Immunoglobulin	Multiple myeloma
CA-125	Ovarian cancer
Carcinoembryonic antigen (CEA)	Colon cancer
	Breast cancer
Prostate-specific antigen	Prostate cancer
Calcitonin	Medullary carcinoma of the thyroid
CA 19–9	Gastrointestinal malignancies
Interleukin-2 receptors	Hairy cell leukemia
	Adult T-cell leukemia/lymphoma
Neuron-specific enolase	Neuroblastoma
Lactate dehydrogenase (LDH)	Lymphoma
	Wilms' tumor
	Ewing's sarcoma
	Testicular cancer

cancer; (2) reduction of tumor bulk to increase the efficacy of subsequent treatment by other modalities; (3) removal of metastases; (4) treatment of oncologic emergencies (spinal cord compression); (5) palliation (relief of biliary or gastrointestinal obstruction); and (6) reconstruction and rehabilitation.

Radiation therapy. This physical method of treating cancer is analogous in some ways to surgery. Radiation will kill only those cells that receive sufficient linear energy transfer. There are two major methods of delivering radiation: teletherapy, where an external beam of radiation is aimed at the tumor site, and brachytherapy, where the radiation source is placed within or near the target. Electromagnetic and particulate forms of radiation can be used to facilitate local control of tumor growth, each of which has advantages for specific clinical situations. Normal tissues vary a great deal in the amount of radiation they can safely tolerate, and it is this tolerance that limits the total acceptable dose of radiation. For radiation to kill a cell, oxygen must be present in the tissue. Many tumors have regions of tissue that are poorly oxygenated, and such tissue is relatively radiation-resistant. Hypoxic cell sensitizers are being tested, and other compounds have been developed in an effort to improve the therapeutic ratio by protecting normal tissues from radiation damage. Radiation therapy is a component critical to the primary treatment of a number of tumor types, including breast cancer, head and neck cancer, cervical cancer, brain tumors, lung cancer, and certain stages of lymphoma and Hodgkin's disease. *See* RADIOLOGY.

Chemotherapy. Chemotherapeutic agents are toxic compounds that exert their maximum antitumor effects when employed at the maximum tolerated dose. Different classes of drugs are used to kill tumor cells by different mechanisms, and are employed mainly for the treatment of metastatic disease. Efficacy is usually directly proportional to dose and inversely proportional to tumor burden. With a drug, as with radiation therapy, toxicity to normal tissues (especially the highly proliferative tissues such as bone marrow and gastrointestinal tract mucosa) limits the amount that can be safely administered. Available agents are capable of curing at least a fraction of patients with 16 different advanced-stage malignancies, and both breast cancer and colorectal cancer appear curable by adjuvant chemotherapy (administered in patients who appear to be free of disease after primary surgery but who have a high likelihood of relapsing if treated with surgery alone). A list of cancers responsive to chemotherapy follows.

Curable in advanced stages
- Choriocarcinoma
- Acute lymphocytic leukemia
- Hodgkin's disease
- Diffuse aggressive lymphoma
- Follicular mixed lymphoma
- Testicular cancer
- Acute myelogenous leukemia
- Ovarian cancer
- Lymphoblastic lymphoma
- Burkitt's lymphoma
- Embryonal rhabdomyosarcoma
- Wilms' tumor
- Peripheral neuroepithelioma
- Neuroblastoma
- Ewing's sarcoma
- Small cell carcinoma of the lung

Curable in the adjuvant setting
- Breast cancer
- Colorectal cancer
- Osteogenic sarcoma
- Soft tissue sarcoma

Responsive but not yet curable in advanced stage
- Head and neck cancer
- Breast cancer
- Colorectal cancer
- Cervical cancer
- Multiple myeloma
- Chronic lymphocytic leukemia
- Chronic myelogenous leukemia
- Bladder cancer
- Adrenal cancer

Prostate cancer
Medulloblastoma
Follicular small-cleaved-cell lymphoma
Melanoma
Endometrial cancer

Two main factors limit the success of chemotherapy: inability to deliver the agents with adequate dose intensity and the development of drug resistance. Theoretically, small increases in the amount of drug delivered could markedly improve the outcome; similarly, small decreases in the amount of drug delivered could drastically reduce the efficacy of the treatment. Analysis of clinical trial results confirms that delivery of agents with less dose intensity results in the cure of fewer patients; however, it has not yet been shown definitively that increasing dose intensity cures more patients. A major emphasis in cancer treatment research is on the development of techniques to deliver chemotherapeutic agents with a higher dose intensity (that is, higher doses of drug per week). Preliminary evidence suggests that the use of high-dose therapy along with bone marrow transplantation to reverse marrow toxicity dramatically improved the outcome in lymphoma and leukemia.

There are diverse mechanisms of drug resistance. The multidrug resistance gene produces a membrane protein that acts as an efflux pump to excrete a variety of structurally unrelated chemotherapeutic agents from the cell. Also, sometimes the tumor cell amplifies the gene whose product is antagonized by the drug; for example, dihydrofolate reductase is amplified to overcome the toxic effects of methotrexate. Drug resistance can result from poor transport into the cell, poor activation of the drug, inactivation of the drug, and altered pools of competing biochemical substrates. Often the use of agents in combinations (combination chemotherapy) overcomes the capacity of the tumor cells to adapt to the treatment. *See* CHEMOTHERAPY.

Biological therapy. This therapy uses an agent that is capable of altering the host-tumor relationship in favor of the host. Biological therapy aims at the rational design of therapeutic interventions to boost host defense or attack the biology of the tumor cell; to inhibit its growth, invasiveness, or metastatic potential; or to promote its differentiation. Biological agents are diverse in constitution and mechanism of action. Cytokines such as interferon, tumor necrosis factor, and interleukin-2 are products of the host and have a variety of mechanisms of action. Monoclonal antibodies are tailored to specifically attack a tumor cell, block critical functional receptors, call forth a cellular response to the tumor, or deliver to it a fatal dose of radiation, toxin, or drug. Adoptive cellular therapy uses expanded host effector cells capable of killing tumor cells in laboratory cultures. Tumor-nonspecific immune stimulants such as BCG (Bacille Calmette-Guérin) are effective in bladder carcinoma in situ. Vaccines against tumors and tumor viruses are showing promising effects in early clinical testing. Differentiating agents such as retinoids are active in acute promyelocytic leukemia. Novel agents targeting oncogene products (such as imatinib mesylate in chronic myeloid leukemia) or signal transduction pathways (gefitinib against epidermal growth factor receptor) are emerging as effective new therapies. Colony-stimulating factors ameliorate the marrow-destroying tendencies of radiation therapy and chemotherapy. Agents aimed at interfering with the metastatic process, such as laminin fragments to prevent tumor cells from adhering to basement membranes, have also been under investigation. A large number of biological therapies are being tested. The use of biological agents in combination and with the other treatment modalities promises to further increase the fraction of patients with disseminated cancer who respond to therapy. *See* MONOCLONAL ANTIBODIES.

Relief of symptoms. Relieving the symptoms of cancer and alleviating the side effects of agents used to treat it is another important aspect of treatment. Many agents and interventions are available for these purposes. Pharmacologic agents can control nausea and vomiting. Various strategies are available to control pain, improve appetite, and combat insomnia and mood changes. Surgical procedures and radiological techniques can palliate many of the complications of cancer that formerly were incapacitating. Even when the hope for a cure has dwindled, the oncologist can relieve much suffering. *See* CANCER (MEDICINE); TUMOR.

Dan L. Longo

Bibliography. M. D. Abeloff et al., (eds.), *Clinical Oncology*, 2004; V. T. DeVita, Jr., S. Hellman, and S. A. Rosenberg (eds.), *Cancer: Principles and Practice of Oncology*, 7th ed., 2004; A. T. Skarin, *Atlas of Diagnostic Oncology*, 2002; I. F. Tannock and R. P. Hill (eds.), *The Basic Science of Oncology*, 3d ed., 1998.

Onion

A cool-season biennial, *Allium cepa*, of Asiatic origin and belonging to the plant order Liliales. The onion is grown for its edible bulbs (**Fig. 1**). *See* LILIALES.

Related species are leek (*A. porrum*), garlic (*A. sativum*), Welsh onion (*A. fistulosum*), shallot (*A. ascalonicum*), and chive (*A. schoenoprasum*). *See* GARLIC.

Propagation. The common onion is grown as an annual and is propagated most frequently by seed sown directly in the field. Onions may also be grown from transplants started in greenhouses or outdoor seedbeds or from small bulbs, called sets, grown the previous year. Field spacing varies; plants are generally grown 1–4 in. (2.5–10 cm) apart in 14–18-in. (36–46-cm) rows. The Egyptian tree or top onion (*A. cepa* var. *vivaparum*) produces little bulbs or topsets in the flower cluster, and the multiplier or potato onion (*A. cepa* var. *aggregatum*) multiplies by branching at the base.

Varieties. Onion varieties (cultivars) are classified mainly according to pungency (mild or pungent) and use (dry bulbs or green bunching). Bulbs may be

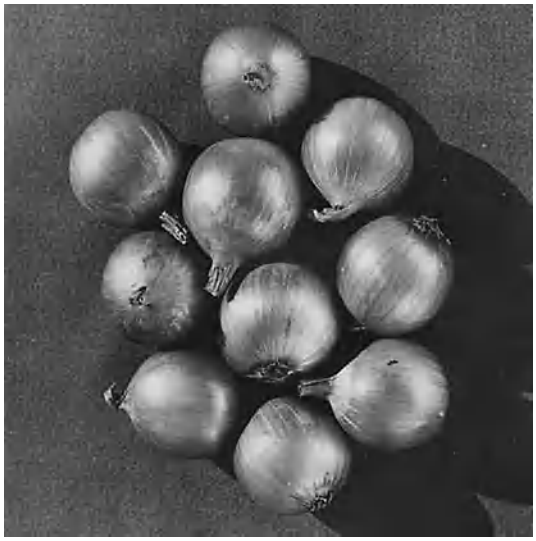

Fig. 1. Onion bulbs. (*Asgrow Seed Company, Subsidiary of The Upjohn Company*)

white, red, or yellow. Varieties differ markedly in their keeping quality and in their response to length of day. Hybrid varieties, produced with male-sterile breeding lines, and with increased disease resistance, longer storage life, and improved quality, are rapidly displacing older varieties.

Harvesting. The harvesting of dry-bulb varieties usually starts after the leaves begin to turn yellow and fall over, generally 3–4 months after planting. Bulbs to be stored are cured by exposure to warm dry air. Bunching onions are ordinarily harvested when the bulbs are $^1/_4$ in. (0.6 cm) or larger in diameter.

Texas, New York, and California are important producing states. H. John Carew

Diseases. The onion and its relatives (chives, garlic, leek, shallot) are subject to many diseases, the most serious of which are caused by fungi and bacteria. Some diseases that occur on the plant in the field affect the yield, quality, or both; others are important in storage and transit, and some diseases are important both in the field and during marketing. Onion diseases cause millions of dollars in damages each year throughout the world. Weather and other environmental conditions, cultivar susceptibility, cultural practices, existence of primary sources of inoculum, and other parameters determine which diseases become limiting economic factors in onion production.

Fungal diseases. White rot, caused by *Sclerotium cepivorum*, is one of the most important onion maladies in the world. White rot usually appears first during the cool, moist weather in the fall and spring. The leaves begin to yellow and wilt and the plants eventually collapse. The soil-borne pathogen invades the roots and the basal parts of the bulb scales, and often develops a white fluffy mass of fungus threads and later black spherical bodies (sclerotia) visible to the naked eye (**Fig. 2**). Tolerant onion cultivars have not yet been tested adequately in the field. Fungicide seed treatments and soil applications often reduce white rot, but are not always reliable. Biological control methods have not yet been applied commercially.

Neck rot and blighting of onion flowers and seed capsules, caused by *Botrytis* spp., are destructive and widespread diseases of onion. Sometimes as much as 50% of the crop is lost. Leaf blight, mainly caused by *B. squamosa*, is a very serious disease during wet growing seasons; it begins in early summer and continues until harvest. Elimination of cull piles and other infested debris will reduce *Botrytis*-induced diseases. Isolation of seed production fields from commercial fields will also reduce these diseases.

Downy mildew, caused by *Peronospora destructor*, varies in destructiveness with locality and season and is enhanced by cool, moist weather. Fungicidal sprays may help if the foliage is thoroughly covered.

At high soil temperatures, a basal and bulb rot, caused by the soil-borne *Fusarium oxysporum*, may cause serious losses to onions, in western valleys and in Northern states after midseason. The roots rot and the bulbs become soft. The disease may then become important in transit.

Black mold, *Aspergillus niger*, is a major cause of bulb losses of yellow onion varieties. The disease is readily identified by observation of black, powdery spores of the mold on the scales at the neck and between the outer scales of the bulb. Drying onion bulbs at or below 3% relative humidity and at 98.5–100°F (37–38°C) may reduce black mold

Colletotrichum circinans, the fungus that causes onion smudge, is confined largely to white cultivars. It appears in the field before harvest as small, dark-green to black dots on the outer scales. Prompt harvest of the crop and avoiding exposure to rain usually reduce this disease.

Smut, caused by *Urocystis cepulae*, is a serious onion disease in upland and muck soils in which the pathogen survives for many years. The disease appears on seedlings as black elongated blisters formed within the scales or leaves. Smut does not affect onion crops starting with sets or transplants. Seed treatments are effective in direct seeded crops.

Pink root, incited by the soil-borne fungus *Pyrenochaeta terrestris*, is occasionally serious in

Fig. 2. White rot (caused by *Sclerotium cepivorum*) on bunching onions. (*P. B. Adams, U.S. Deparment of Agriculture*)

many states in the United States. Infected roots turn yellow, shrivel, and die; they take on a distinct pink color. The disease becomes most apparent at harvest. Crop rotation and use of resistant varieties reduce the disease.

Bacterial disease. Bacterial soft rot, caused by *Erwinia carotovora*, the most destructive post-harvest disease, starts at the neck of the bulb and progresses downward. The organism may enter the bulbs through wounds such as those caused by onion maggots. Removal of diseased bulbs at harvest time, avoiding bruising during harvest and packing, and storing bulbs in dry, well-ventilated areas will reduce soft rot and other storage diseases.

Other diseases. Other diseases, usually of minor importance, are caused by species of *Pythium, Stemphylium botryosum, Alternaria porri*, and *Puccinia porri*, and by virus, nematodes, and *Cuscuta*. Chemical injuries and physiological disorders also occur. *See* PLANT PATHOLOGY. George C. Papavizas

Ontogeny

The developmental history of an organism from its origin to maturity. It starts with fertilization and ends with the attainment of an adult state, usually expressed in terms of both maximal body size and sexual maturity. Fertilization is the joining of haploid gametes (a spermatozoon and an ovum, each bearing half the number of chromosomes typical for the species) to form a diploid zygote (with a full chromosome number), a new unicellular living being. The gametes are the link between one generation and the next: the fusion of male and female gametes is the onset of a new ontogenetic cycle. Many organisms die shortly after sexual reproduction, whereas others live longer and generations are overlapped. Species are usually perceived as consisting mostly of adults, but in most cases the majority of their representation in the environment is as intermediate ontogenetic stages. *See* FERTILIZATION; REPRODUCTION (ANIMAL).

Disciplines. The study of ontogeny covers most aspects of biology and can be split into distinct disciplines. Embryology, often referred to as developmental biology, is the study of embryonic development, usually from fertilization to the beginning of independent life of the new individual. Larval biology is the study of postembryonic and prejuvenile or preadult organisms. Developmental genetics is the study of genetic regulation and specification of development. Life cycle biology involves the study of all the qualitative aspects dealing with organismal growth, from the zygote to the adult stage, identifying developmental stages and their sequences. Life history studies are of ecological interest and consider ontogeny from a quantitative point of view; they examine information such as the number of eggs produced by each individual, the number of viable zygotes produced by each spawning event, and the number of all subsequent stages of the life cycle, having recruitment evaluation [assessment of the numerical value of larval/juvenile recruitment (enrollment in new cohorts); it indicates the number of larvae/juveniles that become adults] as their final result. Evolutionary biology and phylogenyencompass the patterns and processes of evolution, often inferred from ontogenetic patterns and processes.

Development. Zygotes undergo a series of asexual reproductions (via mitoses). In unicellular organisms, mitoses usually lead to the formation of new independent cells (or sometimes to polynucleated cells, or to colonies of identical cells), deriving from a first sexually derived individual, so forming a clone of genetically identical individuals. In multicellular organisms, the products of the asexual reproductions starting with the first division of the zygote are initially clones of one another but remain connected, and eventually differentiate to form an individual. Clonation of individuals occurs even in humans, when the first results of asexual reproduction of the zygote separate from each other, leading to twin formation. *See* MITOSIS.

The ontogeny of a multicellular organism involves segmentation (or cleavage): the zygote divides into two, four, etc., cells which continue to divide. These cells are initially similar to the zygote, although smaller in size. They soon start to differentiate from their ancestors, acquiring special features, and forming specific tissue layers and, eventually, organs. These processes lead to the formation and growth of an embryo. Embryos can develop either freely, or within egg shells, or within the body of one parent; they can grow directly into juveniles (as in humans) or into larvae (with an indirect development, as in insects). Ctenophores, commonly known as comb jellies, are marine organisms that are paradoxical in having sexually mature larvae, which transform into sexually mature adults. *See* ANIMAL GROWTH; EMBRYOLOGY.

Juveniles are similar to adults but are smaller and not sexually mature. Their ontogeny continues until they reach a maximal size and reproductive ability. Usually ontogeny is interrupted at adulthood, but some organisms can grow throughout their life, so that ontogeny ends with death.

Indirect developers undergo rapid but severe morphological and genetic transitions between larval and juvenile/adult bauplans (body plans), referred to as larval metamorphosis. Detection of exogenous stimuli (such as bacterial signals) or internal clocks activate developmental programs. Embryo development and larval metamorphosis usually proceed by a network of signals driving multipotent (stem) cell proliferation and differentiation, apoptosis (programmed cell death), and sometimes cell transdifferentiation.

Stem cells are poorly differentiated or unspecialized cells that retain the potential to form either one or more tissues, or a number of differentiated cell types, by renewing themselves through mitotic division. Stem cells are located in embryos, larvae, and adult organisms. Usually, embryonic stem cells generate multiple cell types of several tissues, whereas adult stem cells are thought to produce cell types of a single tissue or organ. Multipotent stem cells are known in adults of several invertebrate phyla, such as

sponges (archeocytes), cnidarians (interstitial cells), and turbellarians (neoblasts), where they are also involved in cell replacement and regeneration of body parts that are lost by either injury or physiological causes. *See* STEM CELLS.

Inverted cone theory of development. The inverted cone theory states that ontogeny starts with a cell (the zygote) which has the potential to produce a whole adult organism. Further developmental steps—from cleavage to tissue and organ formation—restrict the possibilities of expression of the new cells. A viable mutation in the first ontogenetic steps (at the base of the inverted cone) will thus affect the rest of development, with a sharp modification of the mutated organism, whereas a change in late ontogeny produces a slight modification in adult expression. The higher the mutation in the inverted cone (that is, nearer the end of ontogeny), the smaller the effect; the lower the mutation (nearer the beginning of ontogeny), the greater the effect will be (**Fig. 1**).

When applied to speciation, this theory means that a mutation affecting the base of the inverted cone will lead to a new species similar to the ancestral one (that is, through a process of gradual evolution), whereas a mutation at a lower level of the inverted cone (at the beginning of ontogeny) will affect the rest of development, leading to a new species very different from the ancestral one through a process of saltational (discontinuous) evolution (Fig. 1). Organisms with direct development have a single inverted cone, whereas those with complex life cycles, involving indirect development, have an inverted cone for every developmental stage. *See* ORGANIC EVOLUTION; SPECIATION.

Heterochrony. Heterochrony is a change in the sequence of ontogenetic events within a lineage and is a major evolutionary mechanism, leading to reelaboration of preexisting structures. The evolution of new body plans requires subtraction or reassemblage of old structures (pedomorphosis) as well as the introduction of novel structures added to the ancestral ones (peramorphosis). Ontogenetic changes, through pedomorphosis and peramorphosis, are at the base of evolutionary innovations. Fossil records show that early organisms were simple and that complex organisms are the products of the addition of structures. Since almost no new body plans are known to have evolved after the Cambrian radiation, it is reasonable to presume that most of post-Cambrian evolution occurred via modification of preexisting structures. *See* HETEROCHRONY.

Developmental genetics. The discovery of homeobox, or *Hox*, genes in almost all animals investigated shows that the genes of development are conserved, even though their expression in different lineages leads to different body plans. In the past, embryology had a separate evolution from genetics, but this gap is being bridged by the understanding of ontogenetic processes through an integrated approach. Every cell contains all the necessary information to specify a whole organism (the very basis of clonation) but, during ontogeny, cells differentiate from each other despite their uniform genetic information. It is for this reason that embryology flourished separately from genetics. Apparently, there is a reciprocal information flux through genotype and phenotype, and ontogeny occurs through both genetic and epigenetic processes. *See* DEVELOPMENTAL GENETICS.

Each individual trait has alternative developmental trajectories, that is, potentially realizable phenotypes. These can be defined genetically at birth, but environmental thresholds may drive changes in the ontogenetic history. The shape of trajectories is dynamically modeled by mutations. The interplay of genomic and environmental pressures produces

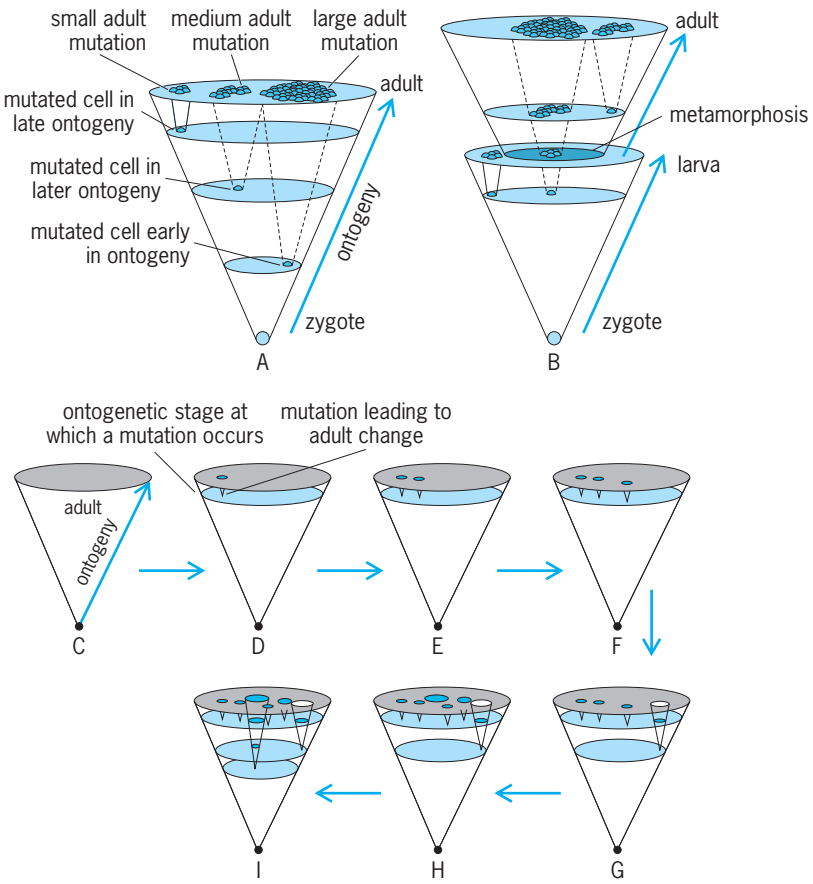

Fig. 1. (A, B) Inverted cone theory of development of an individual organism. (A) The ontogeny of an organism with direct development is seen as an increase in cell number, from the unicellular zygote to the pluricellular adult. Mutations occurring at different ontogenetic stages have a different impact on the final product of ontogeny, that is, the adult. (*After W. Arthur, The Origin of Animal Body Plans, Cambridge University Press, 1998*). (B) The ontogeny of organisms with complex life cycles (indirect development) can be depicted as an inverted cone, with the apex representing the zygote, the area closer to the base representing the (first) larval stage (and more truncated cones, not shown, if there is more than one larval stage), and the base of the distal truncated inverted cone representing the adult. Mutations occurring during ontogeny of larval stages might affect adult structures; however, mutations occurring after metamorphosis affect adult ontogeny only. (C–I) Inverted cone theory of development applied to the evolution of species and higher taxa. (C) Ancestral species, founder of a new lineage. (D) A mutation occurring late in ontogeny (that is, near the end of the inverted cone that represents the adult) leads to a slightly modified new species, sharing most of the features of the ancestral one. (E,F) Additional mutations lead to further anagenesis (or to cladogenesis, not shown), which is the alteration of the genetic properties of a lineage over time. However, the general architecture of the ancestral species is only slightly changed. (G) Speciation due to mutation at an earlier ontogenetic stage modifies the ontogenetic pathway, leading the formation of a new higher taxon (for example, a genus). (H) Gradual evolution continues with a load of mutations affecting late ontogeny. (I) After another load of small changes (not shown), another mutation occurring even earlier in ontogeny causes a further shift from the original body plan, and the founder of a new higher taxon (for example, a family) evolves.

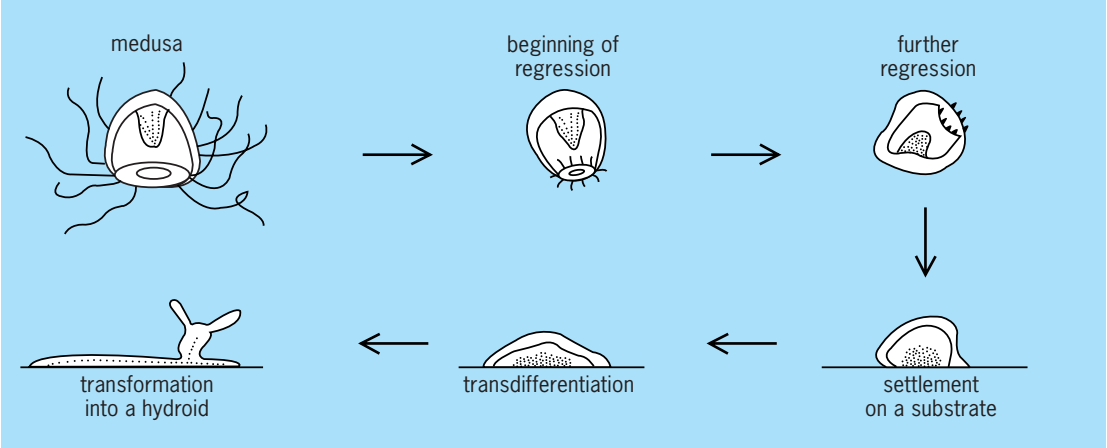

Fig. 2. Ontogeny reversal in the hydromedusa *Turritopsis nutricula*. *(After S. Piraino et al., Reversing the life cycle: Medusae transforming into polyps and cell transdifferentiation in Turritopsis nutricula (Cnidaria, Hydrozoa), Biol. Bull., 190(3):302–312, 1996)*

variations both in phenotypical traits and in their quantitative dimension, which will be screened by selection. Changes in development ultimately lead to evolution. *See* GENE ACTION; MUTATION.

Ontogeny reversal. Ontogeny is considered a one-way chain of events from fertilization to death. The conversion of an adult into a larva (ontogeny reversal) has been documented in species of the hydrozoan genus *Turritopsis*, whose adult medusae can become reduced to a mass of dedifferentiated cells which transdifferentiate into other cell types and form a hydroid colony. This colony is the developmental stage preceding the medusa (**Fig. 2**). Hydroids are normally formed by medusae through sexual reproduction, involving fertilization and planula formation, that is, through embryonic development. Most medusae die after sexual reproduction, but those of *Turritopsis* are able to reverse their ontogeny and, if under stress or after spawning, go back to a hydroid stage, escaping death. These medusae transformed into hydroids are able to produce new medusae, redirecting ontogeny in its normal flow.

Transdifferentiation demonstrates that a differentiated cell can still use the whole potential of its genome. This usually occurs at a cellular level and has almost no bearing on the architecture of the individual organism. In *Turritopsis*, it is the whole adult organism which reassembles its transdifferentiated cells, metamorphosing into an earlier ontogenetic stage.

Ecology of ontogeny. Ontogeny can be put into an ecological framework (and ecology into a development framework). Life cycles allow the persistence of species from one generation to the other. The study of ecology has been centered mainly on both interspecific (food webs) and inter-extraspecific (biogeochemical cycles) fluxes. The importance of intraspecific fluxes (life cycles and histories) has been neglected in the explanation of ecosystem functioning, having been restricted to populational approaches. The widespread occurrence of ontogenetic patterns involving the presence of resting stages leading to the formation of seed banks in terrestrial ecosystems and of resting egg-embryo and cyst banks in aquatic ecosystems is a way of tackling the problem of how matter is maintained in a living state. Ontogeny interruption, usually caused by the onset of adverse conditions, with the formation of resting-stage banks, leads to the formation of a potential biodiversity which is then realized at the onset of a new favorable period when the resting organisms become active again to complete their ontogeny. This process explains sharp discontinuities in the occurrence patterns of many organisms in their adult stage, from terrestrial plants and insects to planktonic protists and metazoans. *See* ECOLOGY.

Ferdinando V. Boero; Jean Bouillon; Stefano Piraino

Bibliography. W. Arthur, *The Origin of Animal Body Plans*, Cambridge University Press, 1998; F. Boero et al., The continuity of living matter and the discontinuities of its constituents: Do plankton and benthos really exist?, *Trends Ecol. Evol.*, 11(4):177–180, 1996; A. Minelli, *The Development of Animal Form. Ontogeny, Morphology, and Evolution*, Cambridge University Press, 2003; S. Piraino et al., Reverse development in Cnidaria, *Can. J. Zool.*, 82 (11):1748–1754, 2004; R. Raff, *The Shape of Life*, The University of Chicago Press, 1996; M. J. West-Eberhard, *Developmental Plasticity and Evolution*, Oxford University Press, 2003.

Onychophora

The only living animal phylum with true lobopods (annulate, saclike legs with internal musculature). There are about 70 known living species in two families, Peripatopsidae and Peripatidae. These terrestrial animals are frequently referred to as Peripatus. Onychophora comprise a single class or order of the same name. They were once considered a missing link between annelid worms and arthropods, but are best considered to be aligned with the arthropods.

They have a cylindrical body, 0.5–6 in. (1.4–15 cm) long, with one antennal pair, an anterior ventral mouth, and 14–43 pairs of stubby, unsegmented legs

ending in walking pads and paired claws. Mandibles are present as modified tips of the first appendage pair. The body surface has a flexible chitinous cuticle; the skin is tuberculate, with numerous transverse folds. The body wall has three layers of smooth muscle, as in annelids, but the coelom is reduced to gonadal and nephridial cavities; the body cavity has an arthropodlike partitioned hemocoel; the heart is tubular with metameric ostia; and the nephridia are segmental with pores on legs or their bases. Gas exchange takes place by means of tracheae; spiracles are minute and numerous, located between skin folds. The brain is bilobed; paired ventral nerve cords are connected by commissures. Slow locomotion is effected by legs and body contractions; the animals can squeeze into very tight spaces. The eyes, located at the antennal base, are the direct type with a chitinous lens and retinal layer. The sexes are separate; the testes and ovaries are paired; and the genital tracts open though the posterior ventral pore. Onychophora are oviparous, ovoviviparous, or viviparous.

The Onychophora are predatory, feeding on small invertebrates. They spurt an adhesive material from modified nephridial glands in papillae at sides of the mouth for food capture and defense. They are largely nocturnal, occurring in humid habitats in forests, being found under logs or in litter.

This is a group of great antiquity. The order Protonychophora was proposed to contain the fossil *Aysheaia pedunculata*, a marine lobopodial onychophorelike animal known from mid-Cambrian shales of British Columbia (over 500 million years old). Stewart B. Peck

Bibliography. A. P. Gupta (ed.), *Arthropod Phylogeny*, 1979; S. M. Manton, *The Arthropoda: Habits, Functional Morphology, and Evolution*, 1977.

Onychopoda

A specialized order of branchiopod crustaceans formerly included in the order Cladocera. The body is up to about 12 mm (0.5 in.) in length, but much of the length of the longest species is made up by a caudal process.

The head and thorax are short, as is the abdomen in some species, but in others it is drawn out into a long caudal process. A carapace is present but is reduced to a dorsal brood pouch, leaving the body naked. A large median compound eye occupies much of the head.

Onychopods swim actively by means of their antennae—the antennules are small and sensory—and seize their food with their four pairs of grasping trunk limbs. Most are predators, but detritus is also eaten by some species. The mandibles are stoutly denticulate.

Reproduction is mostly by parthenogenesis (males are unknown in some species from the Caspian Sea); eggs and young are carried in the brood pouch. Sexual reproduction gives rise to freely shed resistant eggs that overwinter in temperate zone species. Onychopods occur in the sea and fresh water and are worldwide in distribution, but fresh-water species occur only in the Holarctic temperate zone. There is a remarkable group of endemic species in the Ponto-Caspian region. *See* BRANCHIOPODA. Geoffrey Fryer

Onyx

The name onyx is applied correctly to banded chalcedonic quartz, in which the bands are straight and parallel, rather than curved, as in agate. Unfortunately, in the colored-stone trade, gray chalcedony dyed in various solid colors such as black, blue, and green is called onyx, with the color used as a prefix. Because the color is permanent, the fact that it is the result of dyeing is seldom mentioned.

The natural colors of true onyx are usually red or brown with white, although black is occasionally encountered as one of the colors. When the colors are red-brown with white or black, the material is known as sardonyx; this is the only kind commonly used as a gemstone. Its most familiar gem use is in cameos and intaglios, in which the figure is carved from one colored layer and the background in another. *See* CAMEO; CHALCEDONY; GEM; INTAGLIO (GEMOLOGY); QUARTZ. Richard T. Liddicoat, Jr.

Oogenesis

The generation of ova or eggs, the female gametes. Primordial germ cells, once they have populated the gonads, proliferate and differentiate into either sperm in the testis or ova in the ovary. The decision to produce either motile spermatocytes or more sedentary oocytes is based primarily on the genotype of the embryo. In rare cases, this decision can be reversed by the hormonal environment of the embryo, so that the sexual phenotype may differ from the genotype. Formation of the ovum most often involves substantial increases in cell volume as well as the acquisition of organellar structures that adapt the egg for reception of the sperm nucleus, and support of the early embryo. In histological sections, the structure of the oocyte often appears somewhat random, even chaotic, but as the understanding of its chemical and structural organization increases, an elegant but still somewhat cryptic order begins to emerge. *See* OVUM; SPERMATOGENESIS.

Among lower vertebrates and invertebrates, mitotic divisions of the precursor cells, the oogonia, continue throughout the reproductive life of the adult; thus extremely large numbers of ova are produced. In the fetal ovary of mammals, the oogonia undergo mitotic divisions until the birth of the fetus, but a process involving the destruction of the majority of the developing ova by the seventh month of gestation reduces the number of oocytes from millions to a few hundred. Around the time of birth, the mitotic divisions cease altogether, and the infant

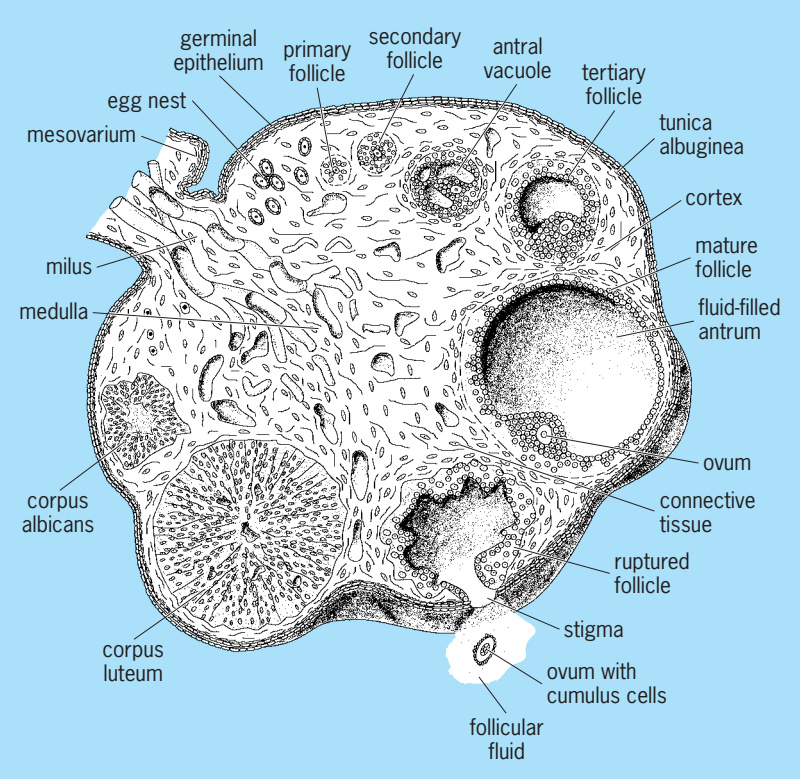

Fig. 1. Three-dimensional view of the cyclic changes in the mammalian ovary.

female ovary contains its full complement of potential ova. At puberty, the pituitary hormones, follicle stimulating hormone (FSH), and luteinizing hormone (LH) stimulate the growth and differentiation of the ova and surrounding cells (**Fig. 1**). *See* MITOSIS.

One important feature of oocyte differentiation is the reduction of the chromosome complement from the diploid state of the somatic cells to the haploid state of gametes. Fusion with the haploid genome of the sperm will restore the normal diploid number of chromosomes to the zygote. The meiotic divisions which reduce the chromosome content of the oocyte occur after the structural differentiation of the oocyte is complete, often only after fertilization. Unlike the formation of sperm, in which the two divisions of meiosis produce four equivalent daughter cells, the cytoplasm of the oocyte is divided unequally, so that three polar bodies with reduced cytoplasm and one oocyte are the final products. Generally, each fertilized oocyte produces a single embryo, but there are exceptions. Identical twins arise from the same fertilized egg; the egg of the armadillo normally gives rise to four embryos; and a single egg of a parasitic wasp may give rise to a thousand embryos. *See* FERTILIZATION; MEIOSIS.

Yolk. The provision of nutrients for the embryo is a major function of the egg, and this is accomplished by the storage of yolk in the cytoplasm. Yolk consists of complex mixtures of proteins (vitellins), lipids, and carbohydrates in platelets, which are membrane-surrounded packets dispersed throughout the egg cytoplasm (ooplasm). The amount of yolk in an egg correlates with the nutritional needs of the embryo. If, as in the case of many invertebrates, the embryo hatches and becomes free-living as a very small larva, the amount of yolk required is small. Sea urchin eggs are a good example of those having a modest amount of evenly distributed yolk, and are classified as oligolecithal. Birds, fishes, and amphibians produce large eggs which sustain considerable growth of the embryo prior to hatching, and are classified as macrolecithal. In addition, the yolk in bird and fish eggs is concentrated at one end of the egg, so they are also described as telolecithal. Although the eggs of mammals are extremely small as compared to the fetus, the bulk of the nutrition is supplied by the placenta; yolk is required only until implantation in the uterine wall.

The contents of the yolk platelets are generally not produced in the oocyte, but are taken up from the blood and surrounding cells. The yolk of vertebrate eggs is thought to originate in the liver and to be carried to the ovary in the blood plasma. Insect yolk proteins are synthesized in the fat body cells and pass between the follicle cells (**Fig. 2**) to the surface of the oocyte, where they are taken in by pinocytosis. Often the transported form of the yolk protein (vitellogenin) is partially degraded upon arrival in the egg cytoplasm to form vitellins. During the formation of the embryo, the vitellins are degraded to amino acids to support the synthesis of new proteins.

Ovary types. Egg cytoplasm also contains large stores of ribonucleic acid (RNA) in the form of ribosomal, messenger, and transfer RNA. These RNAs direct the synthesis of proteins in the early embryo, and may have a decisive influence on the course of development. The mechanism by which the RNA is supplied to the egg is the basis for a major classification of ovary types. Panoistic ovaries, in which the egg nucleus is responsible for the production of all the stored RNA in the ooplasm, are typical of vertebrates,

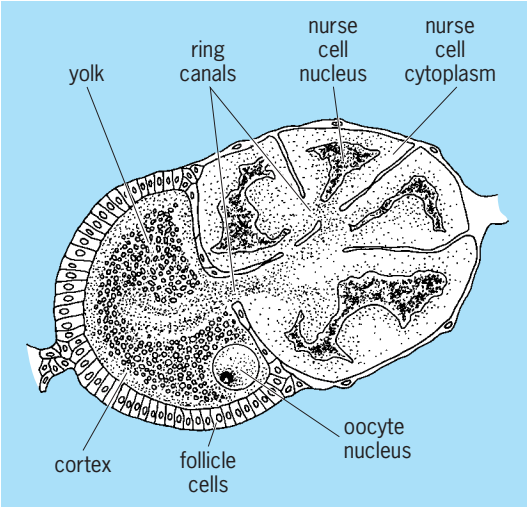

Fig. 2. Developing follicle of the cecropia silkmoth. The nurse cells are sister cells to the oocyte. The ring canals arise from incomplete cytoplasmic divisions and provide direct cytoplasmic connections between oocyte and nurse cells.

primitive insects, and a number of invertebrates. The amounts of RNA produced during the meiotic prophase in such ovaries are much larger than those produced by a somatic cell, and thus special mechanisms seem to be involved in the synthetic process. In the case of ribosomal RNA, extra copies of the ribosomal RNA gene are produced which separate from the chromosomes. The deoxyribonucleic acid (DNA) template for messenger RNAs may not be amplified, but the "lampbrush" configuration often associated with oocyte chromosomes may reflect much more efficient use of the DNA as template for messenger RNA. *See* DEOXYRIBONUCLEIC ACID (DNA); GENE AMPLIFICATION; RIBONUCLEIC ACID (RNA).

More advanced insects and many other invertebrate species have adopted a different strategy for increasing the amount of template DNA required for production of oocyte RNA. Once the oogonial mitoses have been completed, the preoocytes begin a series of divisions in which the nuclei, but not the cytoplasm, divide leaving a cluster of connected cells (cystocytes). Only one cell of each cluster is destined to become an oocyte; the rest will become trophic cells, in some cases called nurse cells. These trophic cells often replicate their chromosomes many times, and thus supply template for oocyte RNA. Ovaries of this type are classified as meroistic, and are further subdivided into polytrophic, where the nurse cells are included in the follicle (Fig. 2), and telotrophic, where the forming oocyte is attached to a mass of trophic cells by a long cord. In panoistic follicles, the RNA is discharged directly from the nucleus into the oocyte cytoplasm. By contrast, the RNA often must be transported for considerable distances from the trophic cells to the ooplasm in meroistic follicles. The transport mechanism is not completely understood, but it is thought that an electrical current flowing from the trophic cells through the oocytes may at least assist in the transport.

Specialized structures of ova. Because eggs are large, sedentary cells, they face difficult physiological stresses, and have developed various specialized structures to cope with these problems. The plasma membrane (oolemma) of developing oocytes is the site of uptake of various compounds from the blood, and to facilitate this activity, numerous fingerlike (microvillar) or platelike projections are found on the surface during oocyte development. These extensions of the surface membrane are often coated with polysaccharides which may assist in the removal of substances from the blood. Numerous coated vesicles, which consist of bubbles of membrane containing concentrated plasma proteins, are found at the base of the microvilli and are the product of this uptake mechanism. The small vesicles near the surface migrate through a yolk-free area, the cortex, and then fuse to form the large yolk platelets.

The cortex is a poorly understood region just under the oolemma, which may be very significant in the formation of the embryo. Many classical experiments in embryology have involved centrifuging fertilized eggs to displace various cytoplasmic elements. One interpretation of the results of these treatments is that unless the cortex is disturbed, the manipulations have little effect on the formation of the embryo, suggesting that the cortex may be of critical importance in determining the structure of the embryo. A special microarchitecture consisting of a dense matrix of fibers has been detected in the cortex of the eggs of many different species. In the silk moth, this structure has been demonstrated to consist of a meshwork of actin fibers. The cortex of eggs is also very rich in messenger RNA (mRNA), which may be held in place by being anchored to this cortical actin meshwork. It is possible that the latent image depends upon maternal mRNA held in a specific spatial configuration by its association with the fibrous cortical meshwork. At least one enzyme crucial to later development, RNA polymerase II, has been found to be stored in the cortex of silk-moth eggs. The ooplasm also contains mitochondria, pigment granules, occasional Golgi apparatus, and of course, the nucleus. *See* CENTRIFUGATION.

In some amphibians and insects a unique area in the ooplasm, germ plasm, has been identified. It is of particular interest since it is the site of formation of the primordial germ cells of the next generation. Germ plasm can often be distinguished by light microscopy of stained sections of oocytes because of its greater affinity for basic dyes. Electron microscopy of these areas reveals large numbers of granules, which are dense concentrations of ribonucleoprotein often intimately associated with mitochondria. The exact nature of the process is not understood, but the blastomeres which come to contain the granules are thereby determined to become the precursors of eggs or sperm. The primordial germ cells can often be identified in later embryos by structure, location, and staining properties, and if they are removed the resulting adult is always sterile.

The activity of the nucleus during oogenesis depends upon the presence or absence of trophic cells. Nuclei of panoistic oocytes, as mentioned earlier, are responsible for the entire RNA contents of the oocyte. Nuclei of meroistic oocytes are relatively "silent" during oogenesis, and often the chromosomes are gathered in a dense body called the karyosome or karyosphere. In the early stages of oogenesis in some insects, a large, temporary, DNA-containing structure, the Giardina body, appears in the nucleus, and then breaks down when yolk begins to appear in the ooplasm (vitellogenesis). The Giardina body contains ribosomal genes, and probably other genes as well.

Oocytes are almost always covered by some type of protective coating which is noncellular and varies in complexity according to the mode of fertilization and the environment into which it is shed. Because they are fertilized internally and maintained in the protection of the uterus, mammalian eggs are simply coated with a noncellular vitelline membrane (area pellucida). A single layer of follicle cells (the corona radiata) remains attached to the outer surface of the area pellucida. Bird eggs, in which the embryos develop outside the body of the female in a dry environment, are coated with a heavy layer of albumen

(egg white) and then a complex shell of mineral and protein which provides support and a barrier to insult and desiccation, but also allows gas exchange with the environment. *See* EGG (FOWL).

Fertilization. Fusion of egg and sperm occur at the top of the reproductive tract in birds, so the albumen and shell are added by oviductal glands after fertilization. Many invertebrates produce eggs with a tough, impervious shell (chorion) secreted by follicle cells before fertilization can occur. A special opening, the micropyle, is incorporated in the structure of the chorion to allow sperm access to the oocyte. Eggs of many lower vertebrates and invertebrates are shed directly into the water and then fertilized by sperm released into the immediate vicinity. Protection in this case is often afforded by jelly coats which expand slowly on contact with water. Although polyspermy, the entry of multiple sperm into an oocyte, is common in some animals, other species have developed mechanisms to prevent multiple sperm entry. The sea urchin egg, for instance, is provided with numerous vesicles just under the oolemma. The first penetration of the oolemma by a sperm causes the eversion of these vesicles, and the release of enzymes that cause the vitelline membrane to pull away from the surface and to harden. This raising and toughening of the vitelline membrane carries away any supernumerary sperm, and affords some protection for the developing embryo. *See* DEVELOPMENTAL BIOLOGY; GAMETOGENESIS. Spencer J. Berry

Bibliography. L. Browder, *Developmental Biology*, 3d ed., 1997; L. Browder (ed.), *Developmental Biology: A Comprehensive Synthesis*, 1985; A. DeLoof, The meroistic insect ovary as a miniature electrophoresis chamber, *Comp. Biochem., Physiol.*, 74A:3–9, 1983; P. D. Nieukoop and L. A. Sutasurya, *Primordial Germ Cells in the Chordates*, 1970.

Oolite

A deposit containing spheroidal grains with a mineral cortex, most commonly calcite or aragonite, accreted around a nucleus formed primarily of shell fragments or quartz grains. The term ooid is applied to grains less than 0.08 in. (2 mm) in diameter, and the term pisoid to those greater than 0.08 in. (2 mm). Accretionary layering (growth banding) is usually developed clearly. A flattened or elongate shape may occur if the nucleus shows that form. Ooids formed on nuclei of shells and shell fragments and composed of fine, radial calcite are shown in **illus.** *a*. These ooids are cemented by coarse, clear calcite. Growth banding is visible in most ooids. The pisoids shown in illus. *b* are composed of many thin layers of very small, tangential (lighter layers) and radial (darker layers) aragonite crystals. These pisoids are cemented with fibrous aragonite.

Ooids are primarily marine, forming in agitated shallow, warm waters such as the Bahama Platform, the Persian Gulf, and off Yucatan. Under those conditions, the ooids are kept intermittently moving, so accretion occurs on all sides. Also, the agitation and

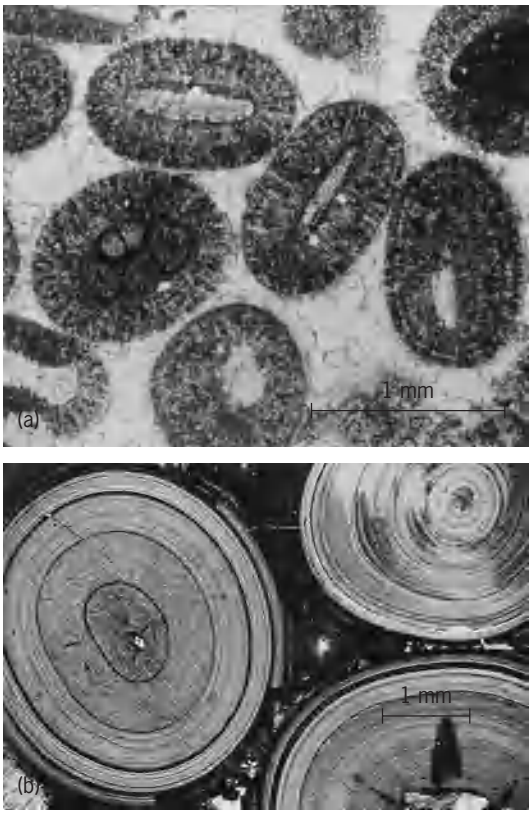

Plain light micrographs of thin sections of ooids and pisoids. (*a*) Oolite of Mississippian age, southwestern Indiana. (*b*) Pisoids from the hot springs of Karlovy Vary (Carlsbad), Czech Republic. The lower right pisoid shows a rock fragment nucleus and dark radial rays, possibly of bacterial origin.

warming leads to loss of CO_2 from the water, thereby enhancing $CaCO_3$ precipitation. Some ooids and most pisoids form in nonmarine environments, such as hypersaline and fresh-water lakes, hot springs, caves, caliche soils, and some rivers. Modern marine ooids are composed commonly of minute needles of aragonite arranged approximately tangentially. Many ancient ooids are composed of radial calcite, and it has long been dogma that those ooids formed by calcite replacement of originally tangential aragonite ooids. However, that widespread misconception is refuted by observations of the calcitization behavior of known aragonites, including ooids in the Pleistocene and Pennsylvanian, which produces irregular, coarser calcite. Hence those radial calcite ooids were never aragonite and may preserve textures close to the original ones.

Primary mineralogy of ooids (aragonite or calcite) appears to have varied through geologic time, probably in response to some factor which is influenced by plate-tectonically mediated global changes in CO_2 abundance.

Although unequivocally secondary replacement of $CaCO_3$ ooids, particularly by such minerals as dolomite, silica, or hematite, is relatively common, some ooids composed of phosphatic (collophane) or ferruginous (chamosite, hematite) minerals appear to be primary, rather than replacement of carbonate

ooids. Even ooids of gypsum and halite have been reported. *See* ARAGONITE; CALCITE.

Philip A. Sandberg

Bibliography. D. P. Bhattacharyya and P. K. Karimoto, Origin of ferriferous ooids: An SEM study of ironstone ooids and bauxite pisoids, *J. Sediment. Petrol.*, 52:849–857, 1982; H. Blatt, *Sedimentary Petrology*, 2d ed., 1991; B. D. Keith and C. W. Zuppann (eds.), *Mississippian Oolites and Modern Analogs*, 1993; W. C. Krumbein and F. J. Pettijohn, *Manual of Sedimentary Petrography*, 1938, reprint 1988; T. Peryt (ed.), *Coated Grains*, 1983; P. A. Sandberg, An oscillating trend in Phanerozoic nonskeletal carbonate mineralogy, *Nature*, 305:19–22, 1983; L. Simone, Ooids: A review, *Earth Sci. Rev.*, 16:319–355, 1981.

Oomycota

A class of fungi in the subdivision Mastigomycotina. They comprise a group of heterotropic, fungallike organisms that are classified with the zoosporic fungi (Mastigomycotina) but in reality are related to the heterokont algae. They are distinguished from other zoosporic fungi by the presence of biflagellate zoospores: the shorter, anteriorly directed flagellum contains two rows of tripartite hairs; the longer, posteriorly directed flagellum is naked and whiplash. Some taxa are nonzoosporic. The thallus may be unicellular, or mycelial and composed of filamentous hyphae that are coenocytic (without cross walls), at least in the early stages of development. The main component of the cell walls is cellulose, or cellulose and chitin. Asexual reproduction involves the release of zoospores from sporangia; in some taxa the sporangium germinates with outgrowth of a germ tube. Sexual reproduction occurs when an oogonial cell is fertilized by contact with an antheridium, resulting in one or more oospores. At least in some, possibly all, species, the thallus is diploid.

There are five orders: Saprolegniales, Rhipidiales, Leptomitales, Lagenidiales, and Peronosporales. The Oomycete systematics is in a state of flux. They have been placed under the Kingdom Protista, Kingdom Chromista, Kingdom Heterokonta, Kingdom Protoctista, and most recently Kingdom Stramenopila.

Oomycetes are cosmopolitan, occurring in fresh and salt water, soil, and as terrestrial parasites of plants. Many species can be grown in pure culture on defined media. The Saprolegniales and Leptomitales are popularly known as water molds; some species are destructive fish parasites, especially in salmon hatcheries, where they may reach epidemic levels. Species of Rhipidales, which once were placed in Leptomitales, grow in stagnant or polluted waters. Many Lagenidiales are parasites of invertebrates and algae. The Peronosporales are primarily plant parasites attacking the root, stem, or leaf, and include some of the more destructive plant pathogens known. One example is *Phytophthora infestans*, the cause of late blight of potato that was responsible for the Irish famine in the midnineteenth century. White rust of crucifers, blue mold of tobacco, and downy mildews are other examples of crop diseases caused by Oomycetes. *See* EUMYCOTA; FUNGI; MASTIGOMYCOTINA.

Donald J. S. Barr

Bibliography. G. C. Ainsworth, K. F. Sparrow, and A. S. Sussman, (eds.), *The Fungi: An Advanced Treatise*, Vol. 4. *A taxonomic review with keys*, Academic Press, New York, 1973; C. J. Alexopoulos, C. W. Mims and M. Blackwell, *Introductory Mycology*, 4th ed., John Wiley, New York, 1996; D. J. S. Barr, Evolution and kingdoms of organisms from the perspective of a mycologist, *Mycologia*, 84:1-11, 1992.; M. W. Dick, Oomycota, in L. Margulis et al. (eds.), *Handbook of Protoctista*, Jones and Bartlett, Boston, 1990.

Opal

A natural hydrated form of silica. Opal is a relatively common mineral in its nongem form, which is known as common opal and lacks the play of color for which gem, or precious, opal is known. All opal is of relatively simple chemical composition, $SiO_2 \cdot nH_2O$, in which SiO_2 represents silicon dioxide, and n represents a variable amount of water (H_2O) that is housed in the internal structural arrangement of the mineral. The water content generally ranges from about 4 to 10 wt %, but it may be as high as 20 wt %. The hardness of opal on the Mohs hardness scale ranges from 5 to 6, the specific gravity from 2.25 to 1.99, and the refractive index from 1.455 to 1.435. Both specific gravity and refractive index decrease as a function of increasing water content. *See* HARDNESS SCALES.

Opal ranges from dense and glassy with a conchoidal fracture to quite porous such as siliceous geyserite (with a specific gravity of 1.8), which occurs at the site of thermal springs and geysers. In addition to being formed at the site of hot springs, opal is deposited by meteoric waters or by low-temperature hydrothermal solutions. It is found lining and filling cavities in rocks, and may replace wood buried in volcanic tuff. It commonly occurs as crusts with botryoidal, globular, and ropy surfaces in stalactitic forms, or as concretionary masses and cavity and vein fillings. Precipitation may result from evaporation or by organic action of organisms such as siliceous tests and diatoms; this form is known as diatomaceous earth (infusorial earth). Most opal is amorphous, that is, apparently lacking an ordered internal, atomic structure. *See* CRYSTAL STRUCTURE; GEYSER; ORE AND MINERAL DEPOSITS; SPRING (HYDROLOGY).

The color of common opal ranges from transparent, glassy, and colorless to white and bluish white. Common pigmenting agents, such as iron, produce yellow, brown, red, and green colors, and frequently several colors in a single specimen. Precious opal has a play of color that is the result of white light being diffracted by the relatively regular internal array of silica spheres. That is, when white light passes through an essentially colorless opal, it strikes the planes of spheres and voids between them, and at

these interfaces the white light is diffracted such that certain wavelengths flash out of the stone as nearly pure spectral colors. Some wavelengths may be completely internally reflected, so the full color spectrum may not escape from a flat surface; however, the curved surface of a cabochon-cut opal allows most or all colors to escape. Because opal is a hydrous mineral, certain opals from specific geologic occurrences may crack because of water loss. Therefore, considerable care is required in the polishing and handling of opal.

Gem classification. Several trade terms are used to describe the appearance of precious opal based on transparency, body color, and the type of play of color. Some of these terms are black opal, which is translucent to almost opaque, with dark gray to black body color, with play of color; fire opal, which is transparent to semitransparent, with yellow, orange, red, or brown body color and with or without play of color; harlequin or mosaic opal, in which the play of color occurs in distinct, broad, angular patches; and matrix opal, which consists of thin seams of high-quality gem opal in a matrix.

Synthetic gem opal has been produced in Switzerland since 1970. This product is similar to natural material in its chemical and physical properties, including a beautiful play of color. A lay person is unable to distinguish this material from natural precious opal. Simulants such as colored synthetic glass with some internal color reflections, and some plastics, imitate some aspects of precious opal. *See* GEM.

Sources. The largest sources of precious opal are in Australia, where it occurs in various districts as vein or void fillings in sedimentary rocks. Important districts are located in Queensland, New South Wales, and South Australia. The opal accumulates in spaces that may be formed by the leaching of fossils, wood, and minerals such as gypsum and glauberite, or in spaces in poorly sorted coarse sediments. In the Lightning Ridge area of New South Wales, the opal occurs as nobbies, which are subrounded shapes up to roughly 40 mm (1.6 in.) in diameter. The origin of the nobbies, not all of which show a play of color, is not understood. *See* SILICATE MINERALS; SILICEOUS SINTER.

Peter J. Darragh; Cornelis Klein

Bibliography. J. E. Arem, *Color Encyclopedia of Gemstones*, 2d ed., 1994; C. S. Hurlbut, Jr., and R. C. Kammerling, *Gemology*, 2d ed., 1993.

Opalescence

The milky iridescent appearance of a dense transparent medium when the medium (or system) is illuminated by polychromatic radiation in the visible range, such as sunlight. Slight changes in the rainbowlike color of the system can occur, depending on the scattering angle, that is, the angle between the directions of incident radiation and of observation. All dense transparent fluids have local density fluctuations due to the thermal motions of molecules, or concentration fluctuations due to the presence of a second component, such as colloidal suspensions or macromolecules in solution. Local fluctuations in density (or concentration) are accompanied by local fluctuations in the refractive index. Since the fluid is optically inhomogeneous, some of the light is scattered to the side. Normally, the amount of light scattered is very small, perhaps of the order of magnitude of 10^{-4} or less of its incident radiation. Whenever the amplitude of fluctuation becomes large, a significant portion of the incident light may be scattered. The transmitted light is then visibly weakened, and the fluid looks turbid. The local optical inhomogeneities may be frozen-in in solids, making the system turbid (opalescent).

Opalescence is a general term which applies to the optical phenomenon of intense scattering in the visible range of the electromagnetic radiation by a system with strong local optical inhomogeneities. The iridescence, or rainbowlike display of interference of colors, arises because the intensity of scattered light is approximately proportional to the reciprocal fourth power of the wavelength of incident light (Rayleigh's law). *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

Critical opalescence. A classical view of the critical point in gas-liquid phase transitions is that it is the state at which the densities of the coexisting gas-liquid phases are equal, and also that it is represented by a characteristic critical temperature above which gas cannot be liquefied, no matter how great the applied pressure is. The corresponding pressure required to liquefy the gas at the critical temperature is the critical pressure. For a one-component fluid the compressibility becomes very large in the neighborhood of the critical point and infinite at the critical point itself. Thus the energy required in the compression of a gas to a given amplitude of fluctuations becomes smaller the closer one approaches the critical point. There the thermal motions of molecules can produce strong density fluctuations, resulting in a very impressive scattering, the so-called critical opalescence. Apart from the critical points of gas-liquid transitions, several other types of second-order phase transitions at which the second derivatives of the free energy are discontinuous, such as critical mixing (consolute) points of binary liquid mixtures and of polymer blends, exhibit critical opalescent behavior. *See* CRITICAL PHENOMENA.

The nature of phase transitions can be studied by observing the size and shape of local fluctuations and their time-dependent changes. By using statistical models, the fluctuations can then be related to thermodynamic and transport properties of the system, such as isothermal compressibility and thermal conductivity. In the critical region, any such properties become very difficult to measure by conventional means. The changes at the critical region in thermodynamic and transport properties are very dramatic; for example, both isothermal compressibility and heat capacity diverge at the critical point. While the scattered intensity is related to the amplitude of local fluctuations in the refractive index, the angular dependence of scattered light reveals the extent of

such fluctuations. As the critical point is approached, the scattered intensity increases because the isothermal compressibility becomes larger; the system consequently looks very turbid. Further analysis shows that the scattered light is concentrated more and more in the forward direction. This indicates that the extensions of fluctuations approach the wavelength of the incident radiation, which is a few hundred nanometers for light in the visible region. These large extensions of local fluctuations result from long-range molecular interactions in the system. A suitable approach is to consider the incident electromagnetic radiation as a measuring scale. Thus, for smaller fluctuations or fluctuations in metal alloys near the consolute point, x-rays or thermal neutrons with wavelengths of the order of 0.5 nanometer or less, instead of visible light with a wavelength of several hundred nanometers, can be used to investigate fluctuation sizes ranging from several to tens (or occasionally hundreds) of nanometers, even when the system is barely opalescent to the naked eye or when it does not transmit visible light.

Time dependency. The density (or concentration) fluctuations that are produced by thermal motions of molecules are time-dependent, so that light is quasi-elastically or inelastically scattered. The spectral distribution of the scattered light characterizes the time dependence of such fluctuations and can be resolved by means of interferometric and optical beat frequency techniques, using a laser as a light source. Changes in the structure of local inhomogeneities during phase transitions, including spinodal decomposition and nucleation processes, can be studied by means of time-resolved (laser) light scattering or time-resolved small-angle x-ray scattering, including the application of intense synchrotron x-rays. However, the divergence in thermodynamic and transport properties results in a slowdown in the relaxation times of thermal fluctuations and an increasing difficulty in reaching thermal equilibrium for the system. Refined experiments to test the very successful applications of the renormalization group theory to critical phenomena become difficult to perform on Earth, and they have begun to be carried out in space under microgravity conditions. Analogy to the critical-point behavior has also been extended to semidilute polymer solutions in general and extremely dilute polymer solutions below the Flory theta temperature, that is, the temperature at which the polymer intermolecular interactions become negligibly small as the second virial coefficient vanishes. Polymer coil collapse could occur below the upper Flory theta temperature or above the lower Flory theta temperature. Care should therefore be exercised when studying the phenomenon of critical opalescence. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; LASER; SYNCHROTRON RADIATION.

Benjamin Chu

Bibliography. S. H. Chen, B. Chu, and R. Nossal (eds.), *Scattering Techniques Applied to Supramolecular and Nonequilibrium Systems*, 1981; B. Chu, *Laser Light Scattering: Basic Principles and Practice*, 2d ed., 1991; B. Chu et al., Critical phenomena and polymer coil-to-globule transition, *J. Appl. Crystal.*, 21:707, 1988.

Open channel

A natural or artificial conveyance through which liquid (typically water) having a free surface moves. The free surface is the interface with a gas (usually the atmosphere), along which the pressure is constant. The liquid is accelerated or decelerated in the flow direction due to an imbalance between the driving gravity force and the viscous boundary resistance force (friction). Such flows occur naturally in rivers, streams, and estuaries as a part of the hydrologic process of surface runoff and artificially in free-surface conduits for the transport of water for irrigation, water supply, drainage, flood control, and other useful purposes. In contrast to full pipe flow, the free surface introduces an additional freedom into the description of open-channel flow, which is the position of the free surface itself as it adjusts to the imposed flow conditions. *See* CANAL; PIPE FLOW; RIVER; VISCOSITY.

Effect of gravity. The effect of gravity as the driving force in open-channel flows is manifested first as a body force (weight) with a component in the direction of motion, parallel to the sloping channel bottom, and second as a gradient (spatial variation) in the pressure with flow depth. Open-channel flows exhibiting changes in depth or velocity at a fixed point over time are classified as unsteady (**illus.** *a*). Flows exhibiting spatial changes in depth or velocity in the flow direction, or the lack thereof, result in a classification as nonuniform or uniform flow, respectively. As a special case, steady uniform flow (illus. *b*) does not change with time and has the property of constant depth and velocity in the flow direction so that the streamwise gravity force is exactly balanced by the boundary resistance force exerted

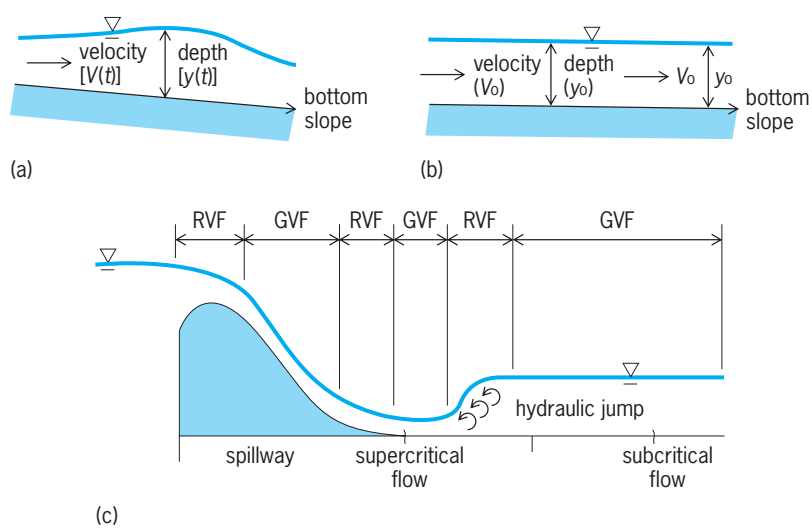

Types of open-channel flow. The inverted triangles denote the free surface of the flow. (*a*) Unsteady flow. (*b*) Steady, uniform flow. (*c*) Steady, gradually varied flow (GVF) and steady, rapidly varied flow (RVF). (*After T. W. Sturm, Open Channel Hydraulics, McGraw-Hill, 2001*)

on the flowing fluid. Artificial channels are often designed for this condition.

Depth of flow. The design of open channels such as storm sewers, roadside drainage ditches, irrigation canals, or water supply aqueducts proceeds from a specification of design flow rate, channel roughness, longitudinal channel slope, and cross-sectional geometry. The depth of flow must be determined from these variables using an empirical formula for steady uniform flow that relates them. An early formula of this type, which determines the channel flow rate Q under such conditions, is Eq. (1), attributed to the

$$Q = (1/n)\,[A]^{2/3}\,S^{1/2} \tag{1}$$

drainage engineer Robert Manning. Here, A is the channel flow area, P is the wetted perimeter of the channel, and S is the bottom slope. The roughness of the channel is characterized by Manning's n factor, which depends partly on the rugosity (wrinkledness) of the roughness elements—from brush and trees in the floodplain to sand, gravel, and cobbles on the bed of an alluvial channel. Some typical values of n are 0.01 for glass, 0.04 for a natural river channel, and 0.10 for a wooded floodplain. The larger the value of n, the greater the uniform-flow depth if all other variables are the same. In the case of sand-bed channels, the bed itself is deformed by the flow into ripples and dunes, which produce variable-form roughness as the river depth rises and falls. Dunes at the beginning of a flood, for example, may be washed out during the peak flow, a process which tends to reduce roughness and buffer depth changes due to floods.

Steady, nonuniform flow. In steady, nonuniform flow, the variation in the streamwise position of the free surface is either gradually varied or rapidly varied (illus. *c*). In the gradually varied case, the streamline curvature is gentle enough that the transverse pressure distribution over the flow depth can be considered to be hydrostatic, and the flow velocity is approximately uniform over the cross section. The shape of the free surface in the streamwise direction is called a water surface profile, and is computed in rivers for the purpose of floodplain mapping and in storm sewers to discover points of surcharging and overflow. Rapidly varied flow, as exemplified by storm surges, dam-break waves, hydraulic jumps, or other abrupt changes in depth, requires analysis of sudden changes in streamwise momentum in the flow region for its prediction. *See* HYDROSTATICS.

Unsteady flow. Unsteady flow is predicted by the Saint-Venant equations, which consist of the general momentum and mass conservation equations for one-dimensional, unsteady, turbulent flow in differential form. In the momentum equation, local and convective acceleration results from the net imbalance in the gravity force component down the slope, the net pressure force arising from a streamwise gradient in depth, and the resisting viscous boundary force. Application of the Saint-Venant equations is made to flood waves in rivers in order to track stage (depth) changes in time and space to provide for warnings and evacuations to prevent loss of life. This general process is referred to as flow or flood routing. *See* FLUID-FLOW PRINCIPLES.

Froude number. Because of the dominance of the gravity force in shaping the free surface of an open-channel flow, the most important dimensionless parameter in open-channel flow is the Froude number, **F**, named after the engineer William Froude, who vastly improved the flow resistance and stability characteristics of ships by building and testing scale models in large towing tanks. He developed the modeling principles required to achieve dynamic similarity between the model and prototype in flows having free liquid surfaces. The primary parameter of this scaling is the Froude number, which represents the ratio of inertial forces to gravity forces and is defined by Eq. (2), in which V is the mean velocity

$$\mathbf{F} = \frac{V}{\sqrt{gD}} \tag{2}$$

of flow, g is gravitational acceleration, and D is the length of the ship. The Froude number is also a primary parameter in open channel flow, for which D is some measure of depth that varies with the cross-sectional shape of the open channel. *See* DYNAMIC SIMILARITY; FROUDE NUMBER; SHIPPOWERING, MANEUVERING, AND SEAKEEPING; TOWING TANK.

Flow regimes. It can be shown from linear wave theory that the speed, or celerity c, of a low-amplitude wave is given by $(gD)^{1/2}$ for the case of shallow water, that is, waves for which the wavelength is large in comparison to the flow depth. Because waves in open-channel flow are generally of this type, it can be observed that the Froude number represents the ratio of flow velocity to wave celerity. For Froude numbers less than one, the velocity is less than the wave celerity; this is called subcritical flow and is analogous to subsonic flow of air with respect to the speed of sound. Similarly, supercritical flow is defined as a flow with velocity greater than the wave celerity so that wave disturbances are swept downstream, analogous to supersonic flow in air. In contrast, waves move both upstream and downstream in subcritical flow. Supercritical flow is more descriptively called rapid flow to reflect the relatively small depths and large velocities at which it occurs, while subcritical flow is referred to as tranquil flow, with relatively large depths and small velocities. *See* WAVE MOTION IN LIQUIDS.

A flow can move from subcritical to supercritical as a result of sudden reductions in channel cross-sectional area or increases in bottom slope. The transition from supercritical to subcritical occurs due to increases in channel cross-sectional area or reduction in bottom slope and often takes the form of a hydraulic jump. In order to illustrate these changes in flow regimes, Boris Bakhmeteff introduced the concept of specific energy, which is the depth or pressure head relative to the channel bottom plus the kinetic energy head. Both these quantities represent energy components per unit weight of fluid; they have the dimensions of length and are referred to as head. The resulting

specific energy diagram, which is a plot of depth versus specific energy, clarifies the existence of critical depth and critical sections in an open channel, where the Froude number is unity and the specific energy is minimum. These sections exert control on the relationship between depth and flow rate, such as might occur in the flow over spillways or in flow- measuring devices such as weirs. *See* FLOW MEASUREMENT; FLUID FLOW; HYDRAULIC JUMP; HYDRAULICS. Terry W. Sturm

Bibliography. E. F. Brater et al., *Handbook of Hydraulics*, 7th ed., McGraw-Hill, 1996; J. A. Cunge, F. M. Holly, Jr., and A. Verwey, *Practical Aspects of Computational River Hydraulics*, reprinted by Iowa Institute of Hydraulic Research, Iowa City, 1994; P. Y. Julien, *River Mechanics*, Cambridge University Press, 2002; I. Nezu and H. Nakagawa, *Turbulence in Open-Channel Flows*, A. A. Balkema, Rotterdam, 1993; T. W. Sturm, *Open Channel Hydraulics*, McGraw-Hill, 2001; B. C. Yen (ed.), *Channel Flow Resistance: Centennial of Manning's Formula*, Water Resources Publications, Littleton, CO, 1991.

Open circuit

A condition in an electric circuit in which there is no path for current between two points; examples are a broken wire and a switch in the open, or off, position. *See* CIRCUIT (ELECTRICITY).

Open-circuit voltage is the potential difference between two points in a circuit when a branch (current path) between the points is open-circuited. Open-circuit voltage is measured by a voltmeter which has a very high resistance (theoretically infinite), such as a vacuum-tube voltmeter. Clarence F. Goodheart

Open-pit mining

The process of extracting beneficial minerals by surface excavations. Open-pit mining is a type of surface excavation which often takes the shape of an inverted cone (**Fig. 1**); the shape of the mine opening varies with the shape of the mineral deposit. Other types of surface mining are specific to the type and shape of the mineral deposit. *See* COAL MINING; PLACER MINING; SURFACE MINING.

The open-pit mine, like any other mining operation, must extract the product minerals at a positive economic benefit. All costs of producing the product, including excavation, beneficiation, processing, reclamation, environmental, and social costs, must be paid for by the sales of the mineral product. A mineral that is in sufficient concentration to meet or exceed these economic constraints is called ore. The terms ore body and ore deposit are used to refer to the natural occurrence of an economic mineral deposit. *See* ORE AND MINERAL DEPOSITS.

Ore bodies occur as the result of natural geologic occurrences. The geologic events that lead to the concentration of a mineral into an ore deposit are generally complex and rare. If those events placed the deposit sufficiently near the surface, open-pit mining may be viable.

Material encountered during the mining process that has little or no economic value is called waste or overburden. One important economic criterion for open-pit mining is the amount of overlying waste which must be removed to extract the ore. The ratio of the amount of waste to the amount of ore is referred to as the strip ratio. In general, the lower the strip ratio, the more likely an ore body is to be mined by open-pit methods.

Mining operations. Modern open-pit mining utilizes large mechanical equipment to remove the ore and waste from the open-pit excavation. The amount of equipment and its type and size depend on the characteristics of the ore and waste and the required production capacity. In general, there are four basic unit operations common to most open-pit mining operations. These are drilling, blasting, loading, and hauling. If the rock types to be excavated are soft enough to permit excavation without blasting, the first two unit operations may not be performed.

The equipment used to accomplish the unit operations works most effectively when applied to a specific set of geometries. These geometric constraints result in the rock being excavated in horizontal slices of equal thickness. A slice is called a bench and may range in height from 10 to 50 ft (3 to 15 m). The specific height of the bench is set, based on the ore body geometry, the distribution of ore and waste, production requirements, and the equipment selected to remove the material. Large base metal mines tend to have benches of 40 or 50 ft (12 or 15 m). Many open-pit gold and silver mines have bench heights of 20 ft (6 m) due to the fact that ore and waste are intermixed, requiring better selectivity (less mixing of the ore and waste), and have lower production rates. Mining bench by bench results in the terraced geometry of the pit wall (**Fig. 2**); this technique aids in the containment of rock falls.

The total depth of the pit is a function of ore geometry, allowable pit slope angle, and the economic constraints of the strip ratio and the ore quality or grade. Large base metal operations can often exceed 1500 ft (460 m) in depth or may be as shallow as 50 ft (15 m). Slope angles depend on the strength of the rock and the natural fractures or joints in the rock. Overall slope angles may vary from 28 to 55° with 35 to 40° being the more common angles.

Drilling. The primary purpose of drilling is to provide an opening in the rock mass for explosives. Explosives work most efficiently when confined within the rock mass to be blasted. The drilling machine generally sits on top of a bench and bores a vertical hole downward into the bench area to be blasted. The drill should be able to penetrate the full bench height plus 2–5 ft (1–2 m) additional depth. The upper limits of the bench height are often set by the effectiveness of the drill. The holes provide a secondary purpose in most metal mining operations in that the cuttings extracted from the hole during the drilling process provide a sample for assaying the grade of the zone near the hole. The results of the assayed

Fig. 1. Aerial view of a mature open-pit mine. (*Kennecott Copper Corp.*)

drill hole cuttings are used to delineate the boundaries between ore and waste and between multiple mineral products, if applicable (for example, copper and molybdenum).

The most common form of drilling is rotary drilling where downward pressure is applied to a rotating tricone rotary bit, a type of bit originally developed by the petroleum industry for oil well drilling. The cuttings are removed from the hole by compressed air forced down the hollow drill pipe, past the rotating bit, and up the annulus between the drill pipe and the blast hole wall. The compressed air serves a second function of cooling the bit and bearings. For rotary drilling of harder formations, the tricone surfaces support buttons made of tungsten carbide, an extremely hard, wear-resistant metal.

When hard and abrasive rocks are encountered, rotary percussion drilling is performed with downhole hammers. This type of blast-hole drilling uses impact by a tungsten carbide bit. After impact, the bit is mechanically retracted and rotated for another impact. This process occurs very rapidly and is generally powered by compressed air. A downhole hammer is a rotary percussion tool that places the reciprocating hammer or impact mechanism immediately behind the bit. In this way, the drill pipe does not incur the stresses of impact. Compressed air forced down the hollow drill pipe powers and lubricates the tool, cools the bit, and removes the cuttings.

The drill equipment and air compressor are generally mounted for transport on a truck-type, rubber-tired carrier or a track-type carrier. The track-type carrier is generally more rugged and able to negotiate steeper grades. Rubber-tired carriers are generally used with smaller drills and where more mobility is required in the pit.

Drill hole diameters vary from roughly 3 to 15 in. (7 to 38 cm). The larger mines generally use holes from $6^3/_4$ in. (17 cm) upward in size. Holes of about $12^1/_4$ in. (31 cm) and larger are generally drilled by track-mounted rotary equipment (Fig. 2). Blast-hole drills can be powered by electric or diesel motors. *See* DRILLING, GEOTECHNICAL.

Blasting. The process of blasting breaks the rock into manageable sizes so that it can be loaded and

Fig. 2. A track-mounted, large-diameter, electrically powered rotary drill. (*Kennecott Copper Corp.*)

moved. While many blasting agents are available, the most common in open-pit mining is ANFO (ammonium nitrate and fuel oil). By itself ammonium nitrate is fairly insensitive and can be handled easily without many of the precautions required for other high explosives. When mixed with roughly 5% diesel fuel, the mixture is a powerful and inexpensive explosive compound. At large mining operations, ammonium nitrate is delivered in bulk in the form of small pellets known as prills, which are stored in bins or silos. A special truck transports the ammonium nitrate to the drill holes, where it is mixed with diesel fuel immediately before being loaded into the drill hole. For small operations, ANFO can be purchased in bags or cartridges. *See* PRILLING.

ANFO cannot be used if the drill holes contain water because the prills are highly soluble and will not detonate if water-soaked. A slurry, which is a mixture of ANFO and one or more additives, can be used to improve water resistance and detonation characteristics. A variety of slurry mixtures can be made to fit specific blasting requirements. Other blasting agents, for example, water gel explosives and gelatin dynamites, are available and are used in specialized applications. Aluminum is a common additive; it increases the energy released by the explosion.

Any high explosive must be detonated to achieve the desired explosive results. Detonation is provided by caps and boosters. Caps are either electric or nonelectric. Nonelectric caps are preferred in open-pit mines as they are not sensitive to electric discharges such as lightning. Propagation of the initial detonation from one drill hole to the next is accomplished with a cord or cable known as det cord or primacord, a hollow plastic tubing filled with high explosive. Det cord is used to connect all drill holes which are filled with explosive in a given blast.

It is not efficient to detonate all holes in a blast at once. A timed pattern of drill hole detonation will result in the most efficient breakage of the rock per given weight of explosive. To accomplish this, delays ranging from milliseconds to seconds are used to arrest the detonation of the primacord momentarily.

After blasting is complete, the muck pile or shot rock is surveyed to set ore and waste boundaries based on the assays from the drill hole cuttings. The material is then ready for loading into trucks and haulage out of the mine. *See* EXPLOSIVE.

Fig. 3. A large (27-yd^3 or 21-m^3) electric track-mounted cable shovel loading a 170-ton (154-metric-ton) haulage truck. (*Kennecott Copper Corp.*)

Loading. The blasted rock is loaded onto some conveyance, such as a truck or a train, for transport. Shovels and rubber-tired front end loaders are the most common pieces of loading equipment in open-pit mines. Track-mounted cable shovels powered by electricity are widely used in large mining operations (**Fig. 3**). These shovels are classified by the volumetric capacity of the dipper (or loading scoop), with common sizes ranging from 22 to 70 yd^3 (17 to 54 m^3). While smaller sizes are available, they are generally being replaced by front end loaders. Electric cable shovels are high-capital-cost items, but they are extremely reliable and can last 15 to 20 years with proper maintenance. Electric shovels are highly productive and consequently have the lowest operating cost of all loading equipment types.

Rubber-tired wheel loaders are commonly used in precious metal mines throughout the world. Wheel loaders for mining applications generally range in size from 5 to 23 yd^3 (4 to 18 m^3) for their loading bucket size (**Fig. 4**). Although both larger and smaller units are available, they are not usually used in open-pit mines. Direct diesel power is common for sizes up to 23 yd^3 (18 m^3), with diesel electric more common for larger units. Wheel loaders do not require trailing power cables and can move from one working area to another rapidly. This mobility is a great benefit where working places are widely separated and the production rate does not require an operating loading unit continuously at each loading area.

Compared to cable shovels, front loaders have lower capital costs but require higher operating costs than shovels of equal capacity. Mining operations with a short mine life or high-mobility requirements are candidates for the application of this type of equipment.

Hydraulic shovels have become more widely applied because of improvements in hydraulic systems

Fig. 4. A 13.5-yd^3 (10-m^3) front end loader putting waste rock onto a 50-ton (45-metric-ton) haul truck. (*Caterpillar Tractor Co.*)

and shovel reliability. Hydraulic front shovels generally range in size from 7 to 45 yd^3 (5 to 35 m^3). The hydraulic machines are generally diesel-powered, but electric-powered models are also available. Although not as mobile as the rubber-tired wheel loader, the hydraulic shovel is usually more productive than the former in terms of tonnage loaded per operating hour.

Two other types of loading equipment used for specific applications in mining are bucketwheel excavators and draglines. Both are used in soft materials that do not require blasting and can be excavated easily, though there are some operators that use draglines to excavate blasted material. Coal stripping is the most common application of both types of equipment; they are used to remove the overburden above the coal. *See* POWER SHOVEL.

Hauling. The broken material is loaded into trucks for transport out of the pit to its destination at the processing plant or the various waste dumps and stockpiles. Truck haulage is the most widely used. Direct conveyor haulage from a loading unit is limited to continuous loading systems such as bucketwheel excavators where the material is small and uniform in size so it may be loaded directly onto a belt. Most open-pit applications of conveyor haulage require the material to be hauled by truck from the loading unit to an in-pit crusher. The crushed material is placed on a belt for transport out of the mine. *See* CRUSHING AND PULVERIZING.

Haul trucks are classified by their payload capacity and vary in size from 20 to 350 tons (18 to 318 metric tons) capacity, with trucks in the 35 to 240 tons (32 to 218 metric tons) capacity range being the most common in open-pits. The trucks with capacity less than 35 tons (32 metric tons) are used in very small mines or quarries. All trucks are powered by diesel engines. The power supplied by the diesel engine is transferred to the wheels by two methods. The first method is by a mechanical transmission, and in the second method electric power is generated and used to drive electric motors located in the truck wheels. Trucks of 85 tons (77 metric tons) capacity and larger often use the diesel-electric method, and smaller trucks typically use mechanical drive. Haul trucks in sizes between 85 and 240 tons (77 and 216 metric tons) can be mechanical or electric drive.

Truck haulage ramps in the pit commonly range between 7 and 10% grade, depending on equipment size, weather conditions, and location in the pit. Haul roads are kept well graded to assure safe and efficient truck operation at speeds up to 35 mi/h (56 km/h). All trucks require braking systems that can stop the vehicle safely when descending these grades.

To augment truck haulage and help reduce fuel costs, a few large mines have installed a trolley assist system to be used with trucks that have electric wheel motors. These are modified trucks that are equipped with pantographs to take electric power from overhead trolley wires instead of from the diesel-alternator system. Such systems are used primarily on relatively permanent haul routes, often on uphill climbs where trucks generally slow down because the diesel engines cannot supply enough power. When not in use, the pantographs are lowered and the trucks operate on power generated by the diesel engine. The system allows the flexibility of truck haulage with the benefit of direct electric power on the more energy-intensive segments of the haul. *See* PANTOGRAPH.

Another modification to material haulage has been the placement of a crusher in or near the pit to shorten the truck haul distance. These crushers are either permanent or movable and feed onto a conveyor system that transports the crushed material to its destination. The system is designed so that the conveyors replace the most costly segment of the truck haul, usually an uphill segment. By doing this, electrical energy replaces diesel energy and reduces the cost of transport. For ore, which must be crushed anyway, crushing in the pit represents no increase in cost. Waste material is not normally crushed, and the cost for this portion of the operation must be offset by the reduction in haul costs.

The implementation of a crush-convey system does have some negative aspects which must be balanced against the reduced haul costs. These include the increase in capital costs early in the mine life derived from installation of the system, and the reduction in the flexibility of the mining operation with semipermanent conveyors and crushers installed in the pit (**Fig. 5**).

In addition to the equipment described above, the open-pit mine has a fleet of auxiliary equipment, including track dozers, rubber-tired dozers, graders, and water trucks. Their function is to maintain the mine working areas, haulage roads, and dump and stockpile areas. *See* BULK-HANDLING MACHINES; CONVEYOR; MATERIALS-HANDLING EQUIPMENT.

Waste disposal. Waste material that is generated during the course of mining at most mines must be

Fig. 5. An in-pit feeder (left), crusher (center), and conveyor system (right and background). (*Cyprus Sierrita Mine*)

discarded as economically as possible without jeopardizing future mining activities but while respecting environmental regulations. Two types of waste material are generated at most mining operations: waste rock and overburden from the mine, and tailings—the waste material from the processing plant after treatment of the ore.

The rock, overburden, and soils removed from the pit and not processed are hauled to dumps or storage piles that can be from tens of feet to several hundred feet high, depending on their location and time of existence. These dumps are located as close to the pit as possible without jeopardizing future pit expansion. Lower-grade material that is not being processed but may become valuable is stockpiled for possible processing at some time in the future. These stockpiles must be located so that the material can be conveniently remined and hauled to the processing plant. Topsoil is often stockpiled for later use in reclaiming the dump and tailings areas. In some operations this is stored in shallow stockpiles to maintain the organic characteristics of the soils.

Considerable attention is being paid to the long-term effects of the disposal of mine waste material, particularly material that contains sulfide minerals or other potentially hazardous materials such as mercury or arsenic. Sulfide minerals are of particular concern because over time the sulfides break down chemically and mix with water to form sulfuric acid. Precautionary planning and monitoring are actively undertaken at all mining operations in the United States. Precautions include lining of tailings and storage ponds, catchment and control of water runoff from dumps and stockpiles due to rain and snowfall, diverting streams and small waterways around mining properties to avoid contamination of their waters, and treatment of contaminated water. Air and water are monitored at most mines to comply with government regulations and to protect the air as well as surface and underground water supplies. These environmental practices are being adopted in mining projects around the world.

Once the reserves have been exhausted and mining activity ceases, reclamation of a mining area is required by many states in the United States. Most large coal surface mines have active land reclamation programs. Many planned mining ventures have budgets that include complete or partial land reclamation. This could involve recontouring and revegetating, covering and planting tailings areas, and filling or partial filling of old pits. These measures add to the cost of mining the ore body and will make the lower-grade mines uneconomic. *See* LAND RECLAMATION.

Computer technology. Computer software is available to assist the mining engineer in ore reserve estimation with the application of geostatistics, mine planning and design, and production and maintenance monitoring and reporting. With the help of high-speed computers the engineering and production staff can evaluate aspects of the mining activities, which allows a more efficient and economical extraction of the mineral commodity. Computerized truck dispatch systems, using global positioning systems (GPS), have been installed at many large mines with fleets greater than 30 trucks. These systems monitor all truck movements, shovel productions and locations, and truck destinations. The computer can instruct the truck drivers to go to specific shovels for loading and to use specific routes to their destinations. Once a particular trip is complete, the computer directs the driver to a specific shovel for the next load. This provides for the best deployment of the truck fleet, minimizing truck-waiting and shovel-waiting time. *See* COMPUTER; DIGITAL COMPUTER; MINING; OPERATIONS RESEARCH; OPTIMIZATION; SATELLITE NAVIGATION SYSTEMS.

Herb Welhener; John M. Marek

Bibliography. S. Bandopadhyay (ed.), *Application of Computers and Operations Research in the Mineral Industry—Proceedings of the 30th International Symposium*, Society for Mining, Mettallurgy, and Exploration, Inc. (SME), 2002; D. W. Gentry and T. J. O'Neil, *Mine Investment Analysis*, Society for Mining, Metallurgy, and Exploration, Inc. (SME), 1984; H. L. Hartman (ed.), *SME Mining Engineering Handbook*, 2d ed., 2 vols., Society for Mining, Metallurgy, and Exploration, Inc. (SME), 1992; W. Hustrulid and M. Kutcha,, *Open-Pit Mine Planning and Design*, 2 vols., A. A. Balkema, 1998; B. A. Kennedy (ed.), *Surface Mining*, 2d ed., Society for Mining, Metallurgy, and Exploration (SME), 1990.

Open-systems thermodynamics (biology)

The application of the physical properties of energy transformations to biological systems.

Thermodynamic reasoning historically evolved along lines which focused on equilibrium states and fictitious "reversible" transitions between them. These transitions have the characteristic that they are frictionless or dissipationless, exemplified by the dissipationless charging and discharging of an ideal electrical capacitor in electronics. This reasoning led to three postulates, commonly called the laws of thermodynamics. These are often augmented by a zeroth law which establishes the notion of thermal equilibrium. Thermodynamics is generally applied to so-called simple systems, namely those which are large enough that certain average properties can be expected to be measured under conditions where fluctuations can be neglected, and which are also homogeneous or at least small collections of homogeneous regions. Thus each system and subsystem becomes defined by a real or imaginary boundary which has well-defined properties with regard to exchange of energy and matter with the environment through the boundary.

The most restrictive condition results from an isolated system, one with a boundary that passes neither energy nor matter. Isolated systems are compelled to achieve equilibrium, and these end points are the focus of the subject matter of classical thermodynamics. Relaxing the constraint against energy flow

through the boundary produces the next level in the hierarchy, the closed system. "Closed" now refers to matter flow since heat does pass into or out of the system. These systems are capable of achieving a class of stationary states away from equilibrium, all involving heat flow through the system due to maintaining a temperature difference across certain boundaries through which the heat flow occurs. If all constraints are removed, an open system results in which boundaries will pass both matter and energy and a broader class of stationary states away from equilibrium can exist. A simple example of this type of open system is an electrical resistor in series with an ideal voltage or current source. The generalization of this idea is the realm of nonequilibrium or stationary-state thermodynamics.

In both open and closed systems, the possibility exists that transients may occur as the system goes from one state to another after an external perturbation at the boundary. The field of network thermodynamics is an extension of thermodynamic reasoning that uses the generalization of classical thermodynamics as the study of all capacitive energy storage, the generalization of nonequilibrium thermodynamics as the study of all dissipative processes without energy storage, and the generalization of inertial energy storage (generalized inductance) to produce a new form of analysis completely isomorphic with modern dynamic systems theory. The network approach to thermodynamics is completed by adding a method for dealing with the connectedness of topology of the system, in the same manner as in electronic networks.

Biological systems. Biological systems obey the laws of physics and chemistry and, in particular, the thermodynamics of open systems. These are also highly organized, hierarchical systems. By using the network approach to their energetics, it is possible to see the flow of matter and energy through them in a particular pattern of connections or topology exactly like the analysis of an electronic network. Since electronic networks have had some powerful methods for their analysis, biological networks are capable of being analyzed by these same, powerful methods. In particular, certain computer simulation techniques developed for the simulation of nonlinear, highly organized, dynamic electrical networks are now being used to simulate some very complex aspects of living networks.

Principles. The principles of the network thermodynamic analysis of biological systems are fairly simple. The system is first seen as a network of discrete elements (simple thermodynamic systems) categorized by the way in which the state variables are related in the elements' description. The state variables are always in conjugate pairs whose product is a power or pseudopower. $P = ef$, where P is the power, and e is an effort or force across the element that arises from a difference in a potentiallike quantity. The flow f is the change in some amount (charge, mass, volume, moles) per unit time and occurs by that amount flowing through the element in a unit of time. Two other state variables are obtained by integration, a displacement [Eq. (1)] and a momentum [Eq. (2)].

$$q = q(0) + \int_0^t f(t)\,dt \tag{1}$$

$$p = p(0) + \int_0^t e(t)\,dt \tag{2}$$

Now four kinds of elements 0 are possible if the elements are to define binary relations among the four state variables: resistance [$R\,(e, f)$], capacitance [$c(q,e)$], inductance [$I(p,f)$], and memristance (membrane resistance) [$M(p,q)$]. Examples of resistances in this general sense are defined by Fick's law for diffusion (permeability^{-1}), chemical reaction kinetics, Poiseuille's law for bulk flow, and Ohm's law. The capacitors are a bit more subtle at first but also are binary relations like resistance. For example, in reactions or mass transfer, the volume v is the capacitance because it relates the displacement, the amount n to the effort's potential, namely concentration c, through a simple binary relation, $c = n/v$ or $n = vc$. The dynamic version of this (easily seen if volume is constant) is expressed by Eq. (3).

$$\frac{dn}{dt} = f = v\frac{dc}{dt} \tag{3}$$

Thus, although resistance yields an algebraic relation between efforts and flows, capacitances (and inductors) yield first-order differential equations. In the steady state, the capacitors are "dormant" since the efforts are no longer changing in time, and they can be omitted from the system's description. Thus arises the earlier statement that the stationary states of nonequilibrium thermodynamics are determined solely by the dissipators (resistances) in the system and also yield purely algebraic relations between the state variables' effort and flow. This is how the physics of the discrete elements is systematically built into the formalism, but the connectedness must also be systematically introduced and combined with the element's description.

Topology of biological networks. Although the method for dealing with a biological network's topology is very general, it is best seen through an example. One such example is shown in **Fig. 1**, which depicts an epithelial membrane such as those lining the gut, kidney tubule, gallbladder, or tongue. For a single, uncharged substance, there are four resistive branches and four compartments with volume for storage, so the network is that shown in Fig. 1 superimposed on the tissue's diagram. The rectangular elements are the resistive diffusion barriers, and the triangles represent the capacitive volume storage elements (essentially capacitors to a "ground" node which is some reference concentration, usually zero). The final step in codifying the system's connectedness is to abstract the network to a linear graph as shown in **Fig. 2**. In this graph, the broken lines are called links or chords and represent the resistive branches, and the solid lines represent the graph's "tree" which is a structure linking all the nodes without any loops being formed. In the

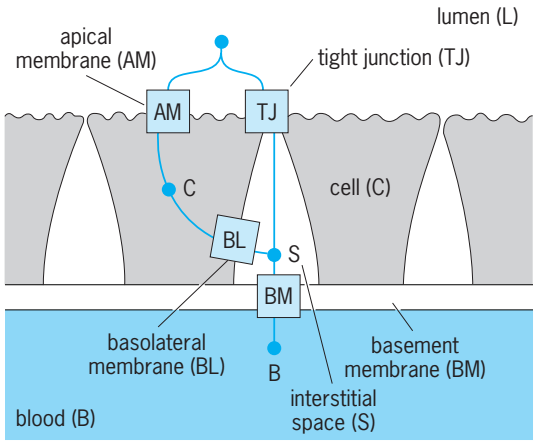

Fig. 1. An epithelial membrane with the network describing the flow through it superimposed on the tissue's morphology. See text for further explanation.

tree, branches are the capacitor-to-ground branches. Now a circle around each node in the graph defines a fundamental cut-set. Each of these circles isolates the node from the rest of the graph by cutting one tree branch and one or more links. This defines a simple code for the graph using the fact that the arrows drawn on each branch define a positive for flow through it and a positive direction for potential drop across it. The code resides in the cut-set–branch incidence matrix Q, having its elements defined as $q_{ij} = +1$ if cut-set i cuts branch j and is oriented in the same direction (cut-sets are directed by the tree branch they cut), $q_{ij} = -1$ if the orientation is opposite, and $q_{ij} = 0$ if they are not incident. In the example above, Notice that Q partitions nicely into

$$Q = [F : I] = \begin{bmatrix} 1 & 0 & 1 & 0 & \cdot & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & \cdot & 0 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & \cdot & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & \cdot & 0 & 0 & 0 & 1 \end{bmatrix}$$

two distinct submatrices F and I, the latter being the identity matrix. This is due to the care taken to arrange the elements so that the links were listed first, then the tree branches in the same order as the cut-sets that cut them. A procedure which is available in any text on network solutions to electronic networks leads to the state-vector equation(4), where

$$\frac{d\bar{x}}{dt} = A\bar{x} + B\bar{u} \qquad (4)$$

$\bar{x}$ is the vector of tree branch efforts, $\bar{u}$ is the vector of sources or inputs energizing the network, and A and B are determined by Q, and the branch resistance and capacitance values. With some modification, the method extends to nonlinear network elements which are characteristic of biochemical reactions. *See* TOPOLOGY.

Kirchhoff's laws and Tellegen's theorem. Another set of relations is easily expressed in this language, namely the conservation of flows at a node, Kirchhoff's flow law, and the closure of the potential drops around any closed path, Kirchhoff's effort law. In the topological code of the incidence matrix, these are $Qf = \bar{O}$, Kirchhoff's flow law, and $X_l = F^T\bar{X}_t$, Kirchhoff's effort law, where l denotes link efforts (resistive), t denotes tree efforts (capacitive), and F^T means the transpose of matrix F. It is easily shown using these relations or variations of them that Tellegen's theorem holds, namely that $\bar{e} \cdot \bar{f} = 0$, or that the vector of efforts is orthogonal to the vector of flows. Since the proof of this depends only on the choice of efforts and flows such that they obey Kirchhoff's laws, the quasipower theorem of Tellegen follows as well: $\bar{e}' \cdot \bar{f} = 0$. In this version of the theorem, $\bar{e}'$ is an effort vector from an entirely different system with the same incidence matrix (topology) as that used to define the flow vector $\bar{f}$. Equally, these could be an effort and flow vector from the same system at different times. This leads to Onsager's reciprocity theorem for coupled systems.

Coupled linear resistive multiport. Onsager's nonequilibrium thermodynamics relates the efforts and flows in a single simple thermodynamic system (membrane, for example) as a set of linear effort flow relations :

$$\begin{aligned} e_1 &= R_{11}f_1 + R_{12}f_2 + \cdots + R_{2n}f_n \\ e_2 &= R_{21}f_1 + R_{22}f_2 + \cdots + R_{2n}f_n \\ &\vdots \\ e_n &= R_{n1}f_1 + R_{n2}f_2 + \cdots + R_{nn}f_n \end{aligned}$$

This set of equations is an extension of the defining relation for a one-port resistor to multiport resistors (a port is a pair of input-output terminals through which a given flow passes and across which a given effort occurs). Onsager's reciprocity law equates all the coupling resistances between pairs of efforts and

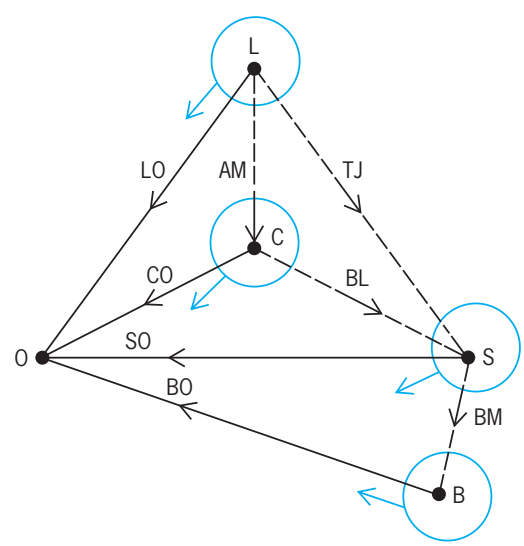

Fig. 2. Graph of the network in Fig. 1, using the same abbreviations. The link branches (AM, TJ, BL, and BM) are the membrane resistances; the tree branches (BO, SO, CO, and LO) are the capacitive storage branches representing accumulation or depletion of material in the compartments. O is a reference or "ground" concentration and is arbitrarily set at zero.

flows $R_{ij} = R_{ji}$ and is a direct result of Tellegen's theorem.

Simulation of living systems. A variety of living systems have now been successfully simulated by using network thermodynamics. Among these are coupled epithelial transport of salt, water, and current; glomerular filtration; microcirculatory absorption and filtration; whole-body pharmacokinetics; cellular pharmacokinetics of anticancer drugs; cellular uptake and metabolism of glucose and its modification by insulin; ion transport systems such as channels and carriers; ecosystems; nervous transmission; synaptic function; and reaction-diffusion systems exhibiting multiple stationary states and hysteresis. *See* THERMODYNAMIC PRINCIPLES; THERMODYNAMIC PROCESSES. D. C. Mikulecky

Bibliography. M. Athans et al., *Systems, Networks, and Computation: Multivariable Methods*, 1974; L. P. Huelsman, *Basic Circuit Theory*, 3d ed., 1991; D. C. Mikulecky, Network thermodynamics: A candidate for a common language for theoretical and experimental biology, *Amer. J. Physiol.*, 245: R1–R9, 1983; G. R. Oster, R. Perelson, and A. Katchalsky, Network thermodynamics: Dynamic modeling of biophysical systems, *Quart. Rev. Biophys.*, 6:1–134, 1973; L. Peusner, *Studies in Network Thermodynamics*, 1986; J. C. White and D. C. Mikulecky, Application of network thermodynamics to the computer modeling of the pharmacology of anticancer agents: A network model for methotrexate action as a comprehensive example, *Pharmacol. Ther.*, 15:251–291, 1982.

Operating system

The software component of a computer system that is responsible for managing and coordinating activities and sharing the resources of the computer. The operating system (OS) acts as a host for application programs that are run on the machine. (In computer jargon, application programs are said to run on top of the operating system.) As a host, one of the functions of an operating system is to handle the details of the hardware operation. This relieves application programs from having to manage these details and makes it easier to write applications. Almost all computers use an operating system, such as handheld computers (including personal data assistants or PDAs), laptop and desktop computers, supercomputers, and even modern video game consoles. *See* COMPUTER SYSTEMS ARCHITECTURE.

Users may interact with the operating system by typing commands or using a graphical user interface (GUI, commonly pronounced "gooey"). The purpose of a GUI is to make it easier to interact with objects in the operating system, such as files, folders, and application programs. The GUI represents these objects as icons and provides easy methods for moving and copying objects, launching and switching between applications, scrolling through windows, and so on. The GUI is generally thought of as part of the operating system; however, most operating systems implement it as a specialized application program or service that runs on top of the operating system. The Microsoft Windows® operating system and Apple's MacOS® each have their own distinctive GUIs that run for the most part outside the operating system. Other operating systems, such as Linux, use GUIs based on the X Window System, which similarly runs on top of the OS. *See* COMPUTER PROGRAMMING; HUMAN-COMPUTER INTERACTION.

Operating systems offer a number of services to application programs and users. Applications access these services through application programming interfaces (APIs) or system calls. By invoking these interfaces (which is usually done by making a special subroutine call in the program), the application can request a service from the operating system, pass parameters, and receive the results of the operation.

Different operating systems exist because no one operating system is capable of being efficient on all types of hardware and for all types of applications. For example, the needs of supercomputer applications are completely different from those of a video game running on a low-cost home game console. Therefore, each operating system provides its own set of APIs that the designers feel are best suited for the intended hardware and application environment. These may be quite different from the APIs of other operating systems, which can make it difficult to take an application written for one operating system and rewrite it to run on another. This is one reason why applications tend to be specific to a certain operating system. Nevertheless, a number of attempts have been made to standardize various subsets of an operating system's APIs so that programs that adhere to the standardized subset can be adapted to run on other operating systems that support the same standard with little effort. One such standard is POSIX, which standardizes a range of APIs for the different variations of the UNIX operating system. Programs that make use of standardized interfaces are said to be "portable" between different operating systems.

Although APIs and services vary from one operating system to another, there are several broad categories of services that most operating systems provide.

Process management. Modern operating systems provide the capability of running multiple application programs simultaneously, which is referred to as multiprogramming or multitasking. Each program running is represented by a process in the operating system. The operating system provides an execution environment for each process by sharing the hardware resources so that each application does not need to be aware of the execution of other processes. For example, since the central processing unit (CPU) of the computer can be used by only one program at a time, the operating system is responsible for sharing the CPU among the processes. This may be done using a technique known as time slicing, whereby each process is given a brief amount of time on the CPU (usually measured in milliseconds) before the CPU is passed to the next process. In this manner,

the processes take turns using the CPU. An operating system may also assign a priority to each process and grant use of the CPU to higher-priority processes ahead of lower-priority processes. This allows the operating system to meet scheduling constraints and deliver deterministic response time for high-priority processes. Operating systems that must meet critical scheduling constraints are referred to as real-time operating systems.

Memory management. The main memory of a computer (known as random access memory, or RAM) is a finite resource. All running applications require the use of memory to store the program's instructions and data while the application is executing. The operating system is responsible for sharing the memory among the currently running processes. When a user initiates an application, the operating system decides where to place it in memory and may allocate additional memory to the application if it requests it. Most operating systems use capabilities in the hardware to prevent one application from overwriting the memory of another. This provides security and prevents applications from interfering with one another. The operating system can also allow portions of memory to be shared among processes. This provides a way for processes on the computer to communicate efficiently. Shared memory is an example of an interprocess communication (IPC) facility.

Many modern operating systems employ a technique called virtual memory (VM). Here, the operating system utilizes special features in the hardware to provide each process with the illusion that it has its own dedicated RAM. This virtual memory may be larger than the actual RAM in the computer. The operating system stores the portions of the virtual memory (referred to as pages) that exceed the available RAM on disk in a paging file or swap area and transparently moves these portions back and forth between memory and disk as needed. This provides a way to share the RAM between multiple processes and to allow processes with very large memory requirements to run on computers that lack large amounts of RAM. *See* COMPUTER STORAGE TECHNOLOGY.

Input/output and device management. Computers can have a variety of input/output (I/O) devices attached to them. Personal computers may include a hard disk, network interface, CD/DVD drive, keyboard, mouse, graphical display, sound, camera, printer, scanner, modem, or other devices. The commands and methods for controlling an I/O device vary from device to device and may even vary for different manufacturers of the same type of device. Because of the great variety of devices available, it would be impossible to imbed the techniques for controlling all possible devices in every application. Therefore, the details of device management are left to the operating system. The operating system provides a set of APIs to the applications for accessing I/O devices in a consistent and relatively simple manner regardless of the specifics of the underlying hardware. The operating system will generally use a software component called a device driver to control an I/O device. A device driver is typically provided by the manufacturer when the I/O device is purchased and is configured into the operating system when the device is first installed. This allows the operating system to be upgraded to support new devices as they become available. *See* COMPUTER PERIPHERAL DEVICES.

File system management. A file system is employed to manage and share the disk space on a computer. A file system allows users and applications to organize their data into files and assign mnemonic names to these files. Modern file systems allow files to be stored in a hierarchical structure for further convenience. File systems also maintain file ownership and permissions, and enforce security by controlling which users can read, write, or execute a given file.

Each file is generally viewed by the application or user as a simple sequence of bytes without regard to the physical location of the data on the disk. Allocating space on the disk is the responsibility of the file system, which may break a given file into pieces and store them in different locations on the disk. Among other benefits, this allows files to be easily extended after they are created. In this manner, the available space on the disk is shared among the different files that are created. This service is completely transparent to applications. However, it can lead to fragmentation of the disk. Some operating systems require that the system administrator periodically run a disk defragmentation utility. Others use a file system designed so that it does not require defragmentation.

So that the file system can be used on many different types of disks, the file system will use the device driver for the disk to manage and control the actual drive that the file system resides on. The device driver used by the file system can also be a virtual driver that controls a set of disks that have been organized to form one large virtual disk. These are generally called redundant arrays of inexpensive disks (RAID) and can allow the file system or even single files to use more space than is available on any single disk drive. RAID can also be used to provide redundancy in various forms so that data are not lost, even if a drive in the array fails.

Network management. A network allows multiple computers to be connected for exchanging data. There is a wide variety of hardware that can be used to create a network. Networks can also be wired or wireless. As with other types of I/O devices, the operating system uses a device driver for the network I/O device. In addition, the operating system includes software known as a network protocol and makes various network utilities available to the user. The protocol provides a standardized method for two computers to communicate and specifies how the data are transferred and how transmission errors are detected. The protocol then provides an API for its services, which allows utility programs and applications to more easily transfer files and data. Applications such as e-mail and Web browsing are implemented outside the operating system by using the APIs provided by the network management portion

of the operating system, while services such as file sharing are generally implemented in the operating system.

Probably the most commonly used protocol suite in use is TCP/IP (transmission control protocol/internet protocol) by which most computers communicate with each other over the Internet. By implementing the network protocol and device driver in the operating system, application programs are relieved of the burden of the communication details. This allows applications to communicate with another computer located around the world as easily as it can with one locally. *See* LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

Security. Operating systems provide security by preventing unauthorized access to the computer's resources. Many operating systems also prevent computer users from accidentally or intentionally interfering with each other. The security policies that an operating system enforces range from none in the case of a video game console, to simple password protection for hand-held and desktop computers, to very elaborate schemes for use in high-security environments. The security policies of an operating system can control who can log into the machine, which users can read or modify which files, which users can install software or reconfigure the operating system, which applications can use the network, which users can terminate applications, and so forth.

With networked computers, a large focus of computer security is in preventing outside users from gaining unauthorized access to a computer's data over the network. This begins in the operating system, which initially determines what outside connections will be accepted. Security continues by having the higher-level networking services and applications (such as e-mail and Web servers) further authenticate outside users before access is granted. Security is most frequently compromised on networked computers due to improper configuration, bugs in the networking or operating system software, or user error that might cause a virus to be installed. Any of these can allow an outside user to gain unauthorized access to the machine. Antivirus software and network firewalls, which further restrict external access to the computer, can be used to help thwart security attacks. *See* COMPUTER SECURITY.

Curt Schimmel

Bibliography. C. Schimmel, *UNIX Systems For Modern Architectures*, Addison-Wesley, 1994; W. Stallings, *Operating Systems: Internals and Design Principles*, 4th ed., Prentice Hall, 2000; A. S. Tanenbaum, *Modern Operating Systems*, 2d ed., Prentice Hall, 2001.

Operational amplifier

A voltage amplifier that amplifies the differential voltage between a pair of input nodes. For an ideal operational amplifier (also called an op amp), the amplification or gain is infinite.

Most existing operational amplifiers are produced on a single semiconductor substrate as an integrated circuit. These integrated circuits are used as building blocks in a wide variety of applications. *See* INTEGRATED CIRCUITS.

Discrete-component and vacuum-tube operational amplifiers preceded the integrated operational amplifiers. The cost of these operational amplifiers was very high, and the performance was modest. The first integrated operational amplifiers appeared in the mid-1960s. By the mid-1970s, inexpensive, general-purpose integrated operational amplifiers with excellent characteristics were widely available. A wide variety of operational amplifiers are now available. Inexpensive general-purpose operational amplifiers are adequate for most applications. Special-purpose operational amplifiers have performance characteristics specifically tailored for niche applications and must be used when the required performance cannot be attained from the general-purpose devices.

Operation. Although an operational amplifier is actually a differential-input voltage amplifier with a very high gain, it is almost never used directly as an open-loop voltage amplifier in linear applications for several reasons. First, the gain variation from one operational amplifier to another is quite high and may vary by ±50% or more from the value specified by the manufacturer. Second, other nonidealities such as the offset voltage make it impractical to stabilize the dc operating point. Finally, performance characteristics such as linearity and bandwidth of the open-loop operational amplifier are poor. In linear applications, the operational amplifier is almost always used in a feedback mode.

A block diagram of a classical feedback circuit is shown in **Fig 1***a*. The transfer characteristic, often termed the feedback gain A_f of this circuit, is given by Eq. (1). In the limiting case, as A becomes very large, the feedback gain is approximated by Eq. (2).

$$\frac{X_o}{X_i} = A_f = \frac{A}{1 + A\beta} \tag{1}$$

$$A_f \simeq \frac{1}{\beta} \tag{2}$$

See FEEDBACK CIRCUIT.

An operational amplifier is often used for the amplifier designated A in this block diagram. Since A_f in the limiting case is independent of A, the exact gain characteristics of the operational amplifier become unimportant provided the gain is large. Although linear applications of the operational amplifier extend

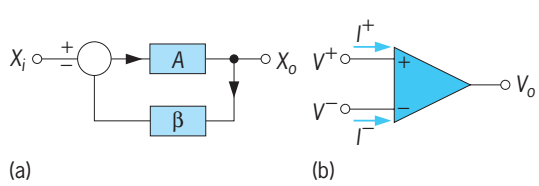

Fig. 1. Basic circuits. (*a*) Classical feedback circuit. (*b*) Operational amplifier symbol typically used in circuit diagrams.

well beyond the simple feedback block diagram of Fig. 1*a*, the applications invariably involve circuit structures with feedback that make the characteristics of the circuit nearly independent of the exact characteristics of the operational amplifier. Such circuits are often termed active circuits.

The commonly used operational amplifier symbol is shown in Fig. 1*b*. In this circuit, the output voltage is related to the gain A of the operational amplifier by Eq. (3), where A is very large and the input currents

$$V_0 = A(V^+ - V^-) \tag{3}$$

I^+ and I^- are nearly zero. As a consequence of the high voltage gain, Eq. (4) follows from Eq. (3). This

$$V^+ \simeq V^- \tag{4}$$

observation is often used in the analysis of active circuits.

Several typical operational-amplifier applications are shown in **Fig. 2** along with the relationship between the input and output. The circuits in Fig. 2*a*–*d* are linear or active circuits using feedback. The circuit of Fig. 2*e* does not use feedback to realize

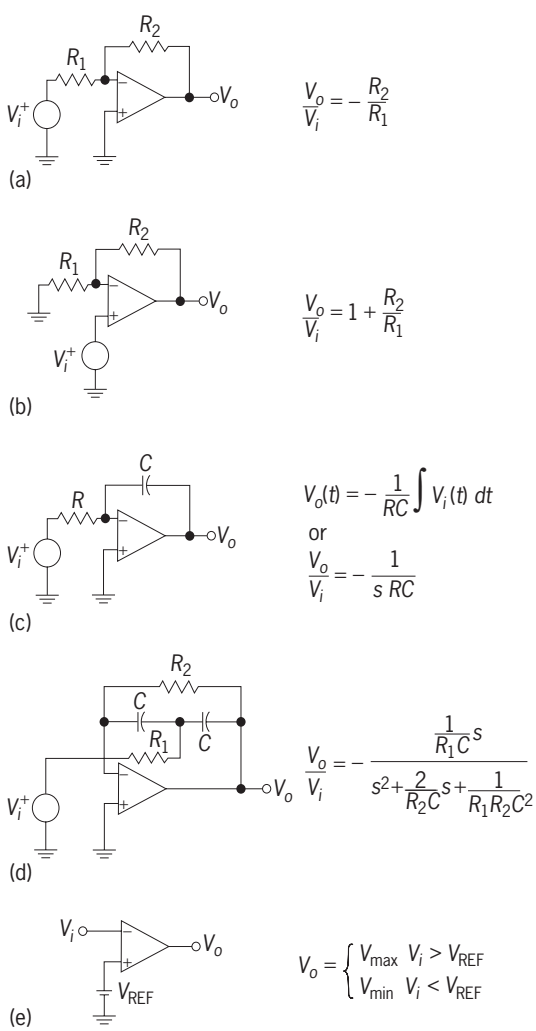

Fig. 2. Typical operational amplifier applications. (*a*) Noninverting amplifier. (*b*) Inverting amplifier. (*c*) Integrator. (*d*) Band-pass filter. (*e*) Voltage comparator.

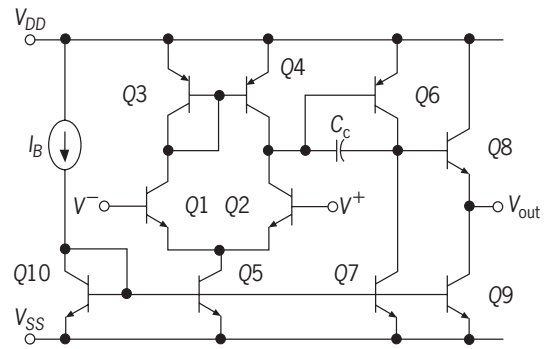

Fig. 3. Simplified circuit diagram of a bipolar operational amplifier.

the nonlinear comparator transfer characteristic. The use of Eq. (4) in the analysis of linear operational-amplifier circuits can be demonstrated easily with the circuit of Fig. 2*a*. Since the voltage at the plus terminal is at 0 V, the potential at the minus terminal must also be at 0 V. Since $I^- \simeq 0$, application of Kirchoff's current law to the node connecting R_1 and R_2 yields Eq. (5), which can be solved to obtain the gain expression, Eq. (6).

$$\frac{V_i}{R_1} + \frac{V_o}{R_2} = 0 \tag{5}$$

$$\frac{V_o}{V_i} = -\frac{R_2}{R_1} \tag{6}$$

Design. Figure 3 is a simplified circuit diagram of a typical bipolar operational amplifier. Every operational amplifier has a differential input stage (here, transistors Q_1 to Q_5), which amplifies the difference between the input voltages, V^+ and V^-, and rejects the sum or common-mode component, followed by one or more secondary amplification stages (here one stage, transistors Q_6 and Q_7). To reduce the output impedance, a low-impedance output stage is added, such as the emitter follower (Q_8 transistors and Q_9) in Fig. 3. A pole-splitting compensation capacitor C_C is added over the secondary gain stages, to push one of the poles in the open-loop gain characteristic to very low frequencies, and at least one of the other poles to higher frequencies, thus improving the stability of the operational amplifier in a feedback configuration. Most commercial operational amplifiers are stabilized for the unity-gain follower configuration, which is the worst case for stability considerations. *See* CONTROL SYSTEMS; EMITTER FOLLOWER.

Specifications. Several specifications are used to characterize the performance of an internally compensated operational amplifier.

The open-loop gain A_{OL} is the gain between the differential input and the output (**Fig. 4*a***) at zero frequency (dc). It is typically high enough to be safely ignored as a source of errors. *See* GAIN.

The gain-bandwidth product (GB) is the product of the open-loop gain A_{OL} and the open-loop 3-decibel bandwidth. For frequencies beyond this 3-dB bandwidth, the gain characteristic falls off at

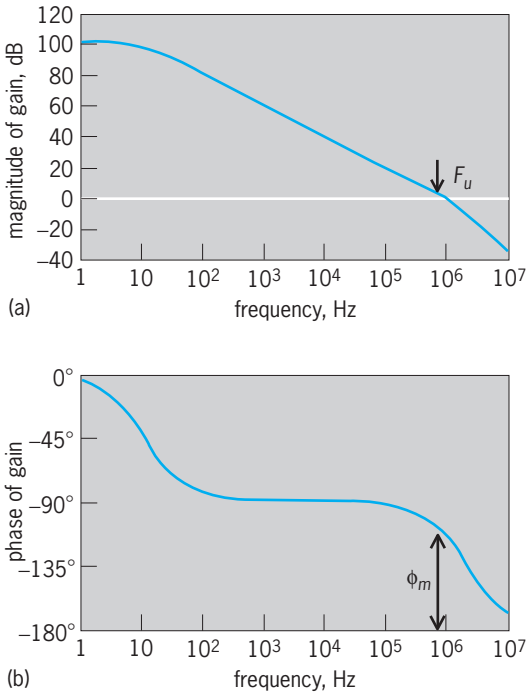

Fig. 4. Open-loop gain characteristics of a typical commercial operational amplifier. (*a*) Magnitude of open-loop gain. (*b*) Open-loop phase characteristic.

20 dB per decade, since a dominant pole is created through pole splitting. The unity-gain frequency F_u, which is the frequency at which the gain of the operational amplifier drops to 0 dB (Fig. 4*a*), is close to the gain-bandwidth product. The gain-bandwidth product is an indication of the bandwidth achievable when feedback is applied around this operational amplifier.

The phase margin ϕ_m is the phase difference between the actual phase of the open-loop gain at the unity-gain frequency F_u and −180° (Fig. 4*b*). Positive values of ϕ_m guarantee stability of a feedback circuit built with that amplifier, with larger values of ϕ_m yielding better stability.

The offset voltage V_{os} is the differential input voltage that must be applied to return the output voltage to 0 V. It is due to device mismatching in the input stage.

The input bias current I_B is the constant (dc) current that is drawn by the input nodes to bias the input devices of the operational amplifier. Related to the input bias current I_B is the small-signal parameter input impedance Z_{in}. The input offset current I_{os} is the difference between the input bias currents drawn by the two input nodes, caused by mismatch in the devices in the input stage. These parameters are primarily important for operational amplifiers with bipolar input transistors, since these have a finite current gain h_{FE}. In operational amplifiers whose input stages are made up of junction field-effect transistors (JFETs), I_B, I_{os}, and Z_{in} can usually be ignored.

The output impedance is specified in the open-loop configuration. Typically, it will exhibit itself as an apparent reduction in the open-loop gain A_{OL} for low-impedance amplifier loads.

The settling time T_{settle} is the time it takes the voltage V_o at the output node of an operational amplifier (in feedback configuration) to settle to within a specific error band. For small signal swings, the settling process is linear. The slew rate (SR) is the maximal variation of the output voltage per unit time (dV_o/dt), which originates from the fact that internal stages are usually biased by constant-current sources. Slewing (operating at maximal output variation) will cause nonlinear distortion in large-swing output signals.

The common-mode rejection ratio (CMRR) is the ratio between the gain of the operational amplifier for difference signals between the input terminals, and the gain for the average or common-mode signal component. The common-mode rejection ratio is important whenever the differential signal is superimposed on a much larger common-mode signal. *See* DIFFERENTIAL AMPLIFIER.

The power-supply rejection ratio (PSRR) is the ratio between the gain of the operational amplifier for difference signals between the input terminals, and the gain for variations of the power-supply voltages. A finite value of the power-supply rejection ratio affects the effective signal-to-noise ratio and the stability.

The input referred noise voltage e_N and the input referred noise current i_N are two noise sources (specified as noise spectral densities) that model the noise performance of acomplete operational

Operational amplifier specifications

Parameter	Ideal value	Typical value*	Optimized value†
Open-loop gain (A_{OL})	∞	100 dB	160 dB
Gain-bandwidth product (GB)	∞	1 MHz	1 GH
Phase margin (ϕ_m)	90°	70°	—
Offset voltage (V_{os})	0	5 mU	20 μV
Input bias current (I_B)	0	200 nA	0.1 pA
Input offset current (I_{os})	0	10 nA	20 fA
Slew rate (SR)	∞	0.7 V/μs	5 V/μs
Common-mode rejection ratio (CMRR)	∞	90 dB	140 dB
Power-supply rejection ratio (PSRR)	∞	80 dB	140 dB
Input referred noise voltage (e_N)	0	20 nU/$\sqrt{\text{Hz}}$	2 nV/$\sqrt{\text{Hz}}$
Input referred noise current (i_N)	0	0.2 pA/$\sqrt{\text{Hz}}$	1 fA/$\sqrt{\text{Hz}}$

*Close to that of a 741 circuit.
†Best value found in commercial operational amplifiers. (These specifications cannot be found in one single operational amplifier.)

amplifier. Two independent sources are required to predict correctly the noise performance of an active circuit. For a low-impedance feedback network around the operational amplifier, e_N dominates the noise performance; for a high-impedance feedback network, i_N dominates. JFET-input operational amplifiers are used primarily for high-impedance signal processing (such as the amplification of capacitive sensor signals), since they have very low values of i_N.

The most important specifications of operational amplifiers are given in the **table**. *See* AMPLIFIER; CIRCUIT (ELECTRONICS). Peter M. VanPeteghem

Bibliography. A. Barna and D. I. Porat, *Operational Amplifiers*, 2d ed., 1989; D. J. Dailey, *Operational Amplifiers and Linear Integrated Circuits: Theory and Applications*, 1989; P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 3d ed., 1993; R. Gregorian and G. C. Temes, *Analog MOS Integrated Circuits for Signal Processing*, 1986; J. Wait, L. Huelson, and G. Korn, *Introduction to Operational Amplifiers: Theory and Applications*, 2d ed., 1992.

Operations research

The application of scientific methods and techniques to decision-making problems. A decision-making problem occurs where there are two or more alternative courses of action, each of which leads to a different and sometimes unknown end result (**Fig. 1**). Operations research is also used to maximize the utility of limited resources. The objective is to select the best alternative, that is, the one leading to the best result. Often, however, it is not simply a matter of searching a table, as there may literally be an infinity of outcomes. More intelligent means are needed to seek out the prime result.

To put these definitions into perspective, the following analogy might be used. In mathematics, when solving a set of simultaneous linear equations, one states that if there are seven unknowns, there must be seven equations. If they are independent and consistent and if it exists, a unique solution to the problem is found. In operations research there may be figuratively "seven unknowns and four equations." There may exist a solution space with many feasible solutions which satisfy the equations. Operations research is concerned with establishing the best solution. To do so, some measure of merit, some objective function, must be prescribed.

		resulting end states				
		θ_1	θ_2	θ_3	θ_4	θ_5
alternatives	a_1	r_{11}	r_{12}	r_{13}	r_{14}	r_{15}
	a_2	r_{21}	r_{22}	r_{23}	r_{24}	r_{25}
	a_3	r_{31}	r_{32}	r_{33}	r_{34}	r_{35}
	a_4	r_{41}	r_{42}	r_{43}	r_{44}	r_{45}

Fig. 1. End states for alternatives in a decision-making process.

There are several terms associated with the subject matter of this program: operations research, management science, systems analysis, operations analysis, and so forth. While there are subtle differences and distinctions, the terms can be considered nearly synonymous. *See* SYSTEMS ENGINEERING.

The field can be divided into two general areas with regard to methods. These are those that can be termed mathematical programming and those associated with stochastic processes. While computers are heavily used to solve problems, the term programming should not be considered in that sense, but rather in the general sense of organizing and planning. Also, the tools of probability and statistics are used to a considerable extent in working with stochastic processes. These areas will be explored in greater detail in a later section. With regard to areas of applications, there are very few fields where the methods of operations research have not been tried and proved successful. Following is a brief history of the field, and then the general approach to solving problems.

History

While almost every art and every science can reach back into antiquity for its roots, operations research can reach back less than a century to find its beginnings. During World War I, F. W. Lancaster developed models of combat superiority and victory based on relative and effective firepower. Thomas Edison studied antisubmarine warfare, and in 1915 F. Harris derived the first economic order quantity (EOQ) equation for inventory. Starting in 1905 and continuing into the 1920s, A. Erlang studied the flow of calls into a switchboard and formed the basis of what is now known as queueing theory.

Empiricists. The formal beginning of operations research was in England during World War II, where the term was and still is operational research. Early work concerned air and coastal defense—the coordination of fighter aircraft, antiaircraft guns, barrage balloons, and radar. Typical of the research groups formed was "Blackett's Circus." This interdisciplinary group consisted of three physiologists, two math physicists, an astrophysicist, an army officer, one surveyor, one general physicist, and two mathematicians. The basic mode of operation was to observe the problem area, and then call on the expertise of the various disciplines to apply methods from other sciences to solve the particular problem. In retrospect, this was an era of "applied common sense," and yet it was novel and highly effective.

Pragmatists. Following World War II, operations research continued to exist mainly in the military area. The operations research groups formed during the war stayed together, and a number of

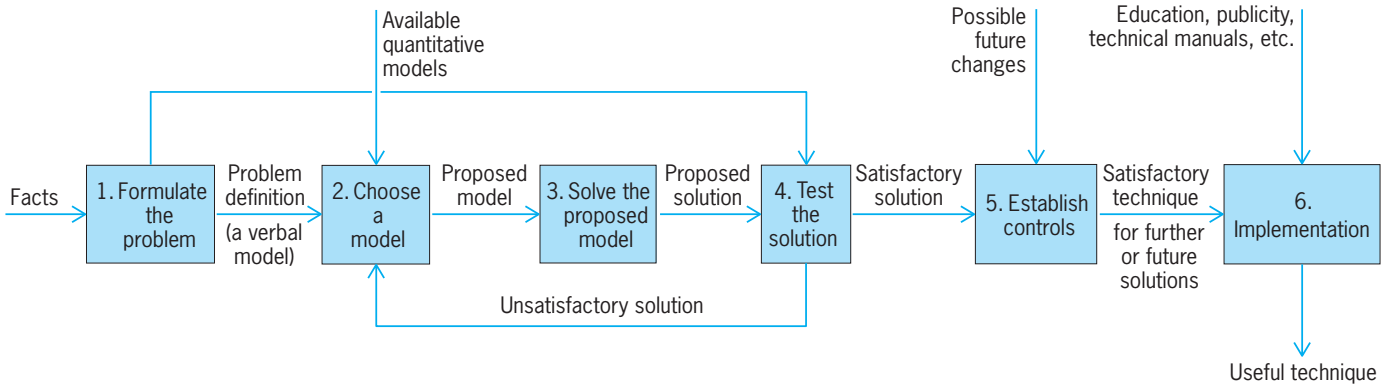

Fig. 2. Operations research approach: the six basic rules of success.

civilian-staffed organizations were established—RAND (1946), ORO (1948), and WSEG (1948). The real impetus to this era and to the whole field was the work done by George Dantzig and colleagues at RAND on Project Scope, undertaken for the U.S. Air Force in 1948. In attacking the problem of assigning limited resources to almost limitless demands, they developed the techniques of linear programming. Perhaps no other method is more closely associated with the field. Its use quickly spread from the military to the industrial area, and a new dimension was added to operations research. No longer was it simply "observe, analyze, and try." For the first time the field became "scientific." It could now "optimize" the solutions to problems. *See* LINEAR PROGRAMMING.

It was during this time that formal courses in operations research were first offered. It was also during this time that operations research suffered its first lapse. Linear programming soon was looked to as the cure for too many of industry's ills. Unfortunately, industry's problems were not all linear, and "straightening them out" to fit resulted in many aborted projects and reports that were simply shelved instead of implemented. The early 1950s saw a growth in the number of industrial operations research groups; by the end of the decade, many had disappeared.

Theorists. Toward the end of the 1950s a number of highly skilled scientists emerged who made some substantial contributions to the field. Operations research came of age and matured. In fact, the movement was so far advanced toward developing a sound theoretical base that a new problem arose—the practicality gap.

Methodology

Operations research today is a maturing science rather than an art—but it has outrun many of the decision makers it purports to assist. The success of operations research, where there has been success, has been the result of the following six simply stated rules:

1 Formulate the problem.
2 Construct a model of the system.
3 Select a solution technique.
4 Obtain a solution to the problem.
5 Establish controls over the system.
6 Implement the solution (**Fig. 2**).

The first statement of the problem is usually vague, inaccurate, and sometimes not a statement of the problem at all. Rather, it may be a cataloging of observable effects. It is necessary to identify the decision maker, the alternatives, goals, and constraints, and the parameters of the system. At times the goals may be many and conflicting; for example: Our goal is to market a high-quality product for the lowest cost yielding the maximum profit while maintaining or increasing our share of the market through diversifications and acquisitions; yielding a high dividend to our stockholders while maintaining high worker morale through extensive benefits, without the Justice Department suing us for constraint of trade, competition, price fixing, or just being too big.

More properly, a statement of the problem contains four basic elements that, if correctly identified and articulated, greatly eases the model formulation. These elements can be combined in the following general form: "Given (the system description), the problem is to optimize (the objective function), by choice of the (decision variable), subject to a set of (constraints and restrictions)."

In modeling the system, one usually relies on mathematics, although graphical and analog models are also useful. It is important, however, that the model suggest the solution technique, and not the other way around. Forcing a model to fit a preferred technique led to some of the bad operations research of the past.

With the first solution obtained, it is often evident that the model and the problem statement must be modified, and the sequence of problem-model-technique-solution-problem may have to be repeated several times. The controls are established by performing sensitivity analysis on the parameters. This also indicates the areas in which the data-collecting effort should be made.

Implementation is perhaps of least interest to the theorists, but in reality it is the most important step. If direct action is not taken to implement the solution, the whole effort may end as a dust-collecting

report on a shelf. Given the natural inclination to resist change, it is necessary to win the support of the people who will use the new system. To do this, several ploys may be used. Make a member of the using group also a member of the research team. (This provides liaison and access to needed data.) Educate the users about what the system does—not to the extent of making them experts, but to alleviate any fear of the unknown. Perhaps the major limitation to successful use of operations research lies in this phase.

Mathematical Programming

Probably the one technique most associated with operations research is linear programming. The basic problem that can be modeled by linear programming is the use of limited resources to meet demands for the output of these resources. This type of problem is found mainly in production systems, but is not limited to this area. Since this method is so basic to the operations research approach, its use will be illustrated.

Linear program model. Consider a company that produces two main products—X and Y. For every unit of X it sells, it gets a $10 contribution to profit (selling price minus direct, variable costs), and for Y it gets a $15 contribution. How many should they sell of each? Obviously there must be some limitations—on demand and on productive capacity. Suppose they must, by contract, sell 50 of X and 10 of Y, while at the other end the maximum sales are 120 of X and 90 of Y. Unfortunately they have only 40 total hours of productive capacity per period, and it takes 0.25 h to make an X an 0.4 h to make a Y. What is their optimal strategy? First, the problem must be formally stated: "Given a production system making two products, the problem is to maximize the contribution to profit, by the choice of how many of each product to make, subject to limits on demand (upper and lower) and available production hours."

If X equals the number of first products made and Y the number of the second products made, the system can be modeled asfollows:

			Constraints line in Fig. 3
Maximize			
Profit contrib. =	$10X + 15Y$		
Subject to	X	≥ 50	1
	X	≤ 120	2
	Y	≥ 10	3
	Y	≤ 90	4
	$.25X + .40Y$	≤ 40	5

The first two constraints are the lower and upper bounds of demand on X, the next two are for Y, and the last constraint refers to the productive capacity.

The problem is illustrated in **Fig. 3**. Any point in the feasible region of solution and the edges will satisfy all the demand and capacity constraints. To pick the best, the objective—"maximize contribution to profit"—is used. Several isoprofit lines, that is, lines where the profit is a constant value, have been added to Fig. 3. Note that the $600 line is below the feasible region, the $1800 line is above it, and the $1200 line runs through it. If other lines were formed between $1200 and $1800, one could graphically find the maximum-profit-level line that still had one point in the feasible region. This would occur at the intersection of the second and fifth constraints. At this point, the solution is as follows:

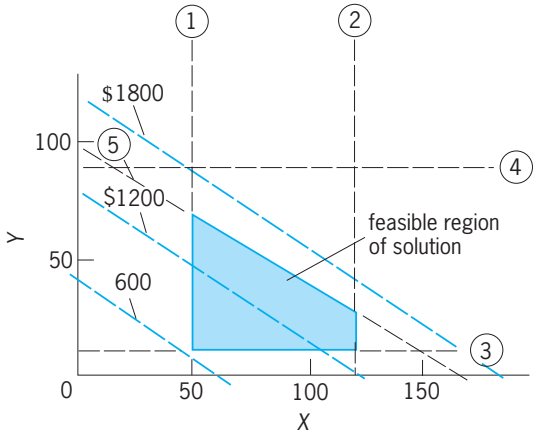

Fig. 3. Graphical presentation of production problem; circled numbers refer to the constraints.

$$\text{Profit contrib.} = \$1575$$
$$X = 120$$
$$Y = 25$$

It is fairly easy to verify that this is the optimal solution, that all the constraints are satisfied, and that no larger profit can be found.

While this simple example did not pose much of a computational problem, actual cases do, in that there may be thousands of variables and thousands of constraints. Problems of scheduling the product mix from a refinery where a blend of different crudes can be used (each with different costs and properties) are in this category. Efficient algorithms and computer codes have been developed to solve these problems.

Other models. There are a number of other linear programming models that relate to specific types of problems. However, advantage is taken of the special structure of the model to develop special and more efficient methods of solution.

Transportation problem. While the transportation problem can be modeled as a linear programming model, it is not efficient to solve it in that manner. Consider the tableau shown in **Fig. 4**. This shows the unit profit of shipping a common item from four plants to six regional distribution points. Supply refers to the number of units that each source can provide. Demand refers to the number of units that each destination requires. The "dummy" column is added to the supply-demand matrix so as to make the supply equal the demand. In Fig. 4, a dummy demand of 300 units is added to make it equal the total 3600 units of supplies available from four sources. Most transportation problems simply minimize the cost of transportation. This example includes variable selling price, variable production cost, and transportation costs that depend on the distance between the plant and destination.

If the problem were to be modeled as a linear programming model, there would be a decision variable for each combination of shipment. For the example problem there are 4 sources and 6 possible destinations, resulting in 24 decision variables: the amount shipped from each source to each destination. There would be 10 constraints—4 indicating the capacity limits on the plants and 6 reflecting the demand constraints. In general, there will be $S \times D$ variables and $S + D$ constraints, where S is the number of sources and D is the number of constraints. For a more realistic problem, there might be a dozen sources and five dozen destinations. Then the linear programming model would have 720 variables and 72 constraints.

Another factor is that the linear programming model matrix will be composed entirely of 0s and 1s. Also these 1s will be regularly placed. Because of this regular structure, more efficient methods have been found to solve the problem. These algorithms typically have two phases, just as in linear programming.

In the first phase, the objective is to find a feasible solution. One such heuristic is called the North-West Corner rule. Simply stated, the solution process starts in the upper left-hand corner of the tableau, and either the minimum of supply or demand is assigned to the first cell. The next step is to move either horizontally or vertically, depending upon where the demand has been satisfied or the capacity has been all used. The result is a set of assignments that generally move down the main diagonal of the tableau.

A more efficient alogrithm is known as Vogel's approximation method (VAM). Again using the transportation tableau, the procedure is to examine each row and column to determine the difference between the highest unit profit and the next highest unit profit. Where only costs are considered, the procedure would be to determine the difference between the lowest cost and the next lowest cost.

These differences are examined to determine the largest difference, and a transportation assignment is made to the highest unit profit that caused this largest difference. The assignment is to ship either the remaining capacity of the plant or the remaining unsatisfied demand of the destination—whichever is smaller.

The basic philosophy of Vogel's approximation method is that these differences represent a measure of "regret"—that is, how much would be lost (or how much more would have to be paid) if the row or column representing highest unit profit was not used. That is, if it is not used, then it might be necessary to settle for "second best." Choosing the largest difference is an attempt to minimize regret.

Vogel's approximation method then proceeds to recalculate these differences for the remaining row and columns (one row or column is eliminated at each iteration) and to repeat the procedure. When the method is completed, the result will not only be a feasible solution but an extremely good one—and possibly the optimal.

INCREMENTAL PROFIT MATRIX								
SOURCES	DESTINATIONS							
	St. Louis	Kansas City	Little Rock	Dallas	Houston	Mont-gomery	dummy	source supply
Atlanta	25.00	27.60	40.50	22.50	32.00	48.50	0.00	1200
Austin	12.00	18.50	30.00	18.00	28.50	32.00	0.00	500
Memphis	32.60	35.20	48.70	30.80	39.80	52.00	0.00	1200
Tulsa	21.40	28.00	37.70	22.50	30.50	38.00	0.00	700
destination demand	750	500	150	750	1000	150	300	3600

Fig. 4. Transportation problem tableau, showing the incremental profit and the supply-demand matrices of sources and destinations.

After either approach, the result is an initial basic feasible solution. The next steps are to determine if the solution is optimal. The specific algorithm is too lengthy to present here.

Assignment problem. The assignment problem is also a linear programming model and a variation of the transportation problem. It can be stated as: Given a set of N tasks to be performed by N people and a table of values showing the cost or time it takes each person to perform each task, the problem is to minimize the total time or cost it takes to perform the N tasks by choice of which task is assigned to each person, with the restrictions that each task must be done and each person is assigned to only one task.

This can be considered a special form of the transportation problem where there are N sources (tasks) and N destinations (people). Also, the capacities are each 1. The linear programming simplex method to solve the problem can be used. There would be $N \times N$ variables and $2N$ constraints. The matrix also would be all 0's and 1's. The transportation algorithm could also be used, but would run into many degenerate solution points.

The special structure of the model is used to form a special algorithm to efficiently solve this model. It is known as the Hungarian method and consists of the following steps:

1. The smallest element in each row of the cost-time matrix is subtracted from each element in that row. (This will create at least one element that is 0.)
2. The same procedure is repeated for each problem.
3. An attempt is made to make one assignment to each row and column using the elements that are 0 as candidates. (The recommended procedure is to start with those rows and columns with only one 0.)
4. If a complete set of assignments has been made, then the optimal solution has been found; if not, the next step in the procedure is followed.
5. Using a minimum number of lines either horizontal or vertical, the 0's in each row and column are crossed out. (This can always be done, as many lines are assignments in the current solution.)

6. The smallest element not crossed out is selected, and it is subtracted from each element not crossed out and added to each element at the intersection of a vertical and horizontal line. (This will result in the creation of at least one new 0 as a candidate for assignment and may result in the loss of one or more 0's if they happened to be at the intersection of two lines.)

7. The entire process is repeated (steps 1–6) until the optimal solution is found.

The solution process is reasonably simple, and even modestly large problems can be solved by hand. Besides the generic assignment problem, this approach has been used for such diverse applications as the machine change-over problem (the sequence of jobs to be run on a machine to minimize setup time; assigning students to projects; allocating parking spaces to employees; and so forth).

Network models. There is another set of linear programming models that fall under the general category of network models:

The shortest-path problem determines the shortest path from one node to another. (The reverse of this is the critical path method used in project management.)

The max flow problem determines the capacity of a network such as a pipeline or highway system.

The min cost problem is a variation of the transportation problem.

Integer linear programming. There is a class of problems that have the linear programming form, but in which the variables are limited to being integer values. These represent the most difficult set to solve, especially where the variables take on values of 0 or 1. (This problem occurs, for instance, in project selection—the project is either funded, 1, or not funded, 0.) While a number of algorithms have been developed, there is no assurance that a given problem will be solved in a reasonable amount of computer time. These algorithms fall into two general categories—the cutting plane method, which adds constraints that cut off noninteger solution points, and the branch and bound method, which examines a tree network of solution points. *See* ALGORITHM.

Nonlinear programming. Another category of problems arises where either the objective function or one or more of the constraints, or both, are nonlinear in form. A series of methods have been developed that have varying degrees of success, depending on the problem structure.

Geometric programming. One of the techniques developed relates to a certain class of nonlinear models that use the arithmetic-geometric mean inequality relationship between sums and products of positive numbers. Such models result from modeling engineering design problems, and at times can be solved almost by inspection. Because of the ease of solution, a considerable effort has been made to identify the various problems that can be structured as a geometric programming model.

Dynamic programming. This technique is not as structured as linear programming, and more properly should be referred to as a solution philosophy rather than a solution technique. Actually, predecessors of this philosophy have been known for some time under the general classification of calculus of variation methods. Dynamic programming is based on the principle of optimality as expressed by Richard Bellman: "An optimal policy has the property that whatever the initial state and initial decision are, the remaining decision must constitute an optimal policy with regard to the state resulting from the first decision." *See* CALCULUS OF VARIATIONS; OPTIMIZATION.

The operative result of this principle is to start at the "end" of the problem—the last stage of the decision-making process—and "chain back" to the beginning of the problem, making decisions at each stage that are optimal from that point to the end. To illustrate this process, the fly-away-kit problem (also known as the knapsack problem) will be briefly described: Given a set of components, each with a unit weight and volume, the problem is to maximize the value of units carried to another location, by choice of the number of each to be taken, subject to limitations on volume and weight that can be carried.

In this problem each component will represent a stage at which a decision is to be made (the number to be included in the kit), and the amount of volume and weight left are defined as state variables. The end of the problem, then, is where there is only one more component to consider (**Fig. 5***a*).

Usually at this stage the solution is almost trivial. As many of the last components (X_1) are selected as can be within the limits of available volume (V_1) and weight (W_1). In practice, all possible values of these two state variables are solved for, since it is not known how much will be left when the last stage is reached. To this solution is now "chained" the problem of how to select the second-last component (Fig. 5*b*).

Here is considered the problem that, for each level of volume and weight, a choice must be made between taking one or more of component 2 or passing it on to the last stage. Combinations of components 1 and 2 must now be looked at to find the best mix, considering the available resources. In like manner, the solution is "chained back" to the beginning of the problem (Fig. 5*c*).

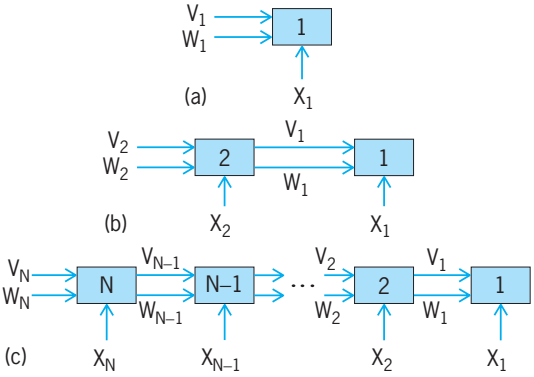

Fig. 5. Dynamic programming problems. (*a*) Final decision stage. (*b*) Two-stage problem. (*c*) "*N*"-stage problem.

The measures of value can be in any terms. If the components are simple cargo, they could have a monetary value. If they are spare parts for some system, a reliability measure could be used.

Stochastic Processes

A large class of operations research methods and applications deals with stochastic processes. These can be defined as processes in which one or more of the variables take on values according to some, perhaps unknown, probability distribution. It takes only one of these random variables to make the process stochastic.

In contrast to the mathematical programming methods and applications, there are not many optimization techniques. The techniques used tend to be more diagnostic than prognostic; that is, they can be used to describe the "health" of a system, but not necessarily how to "cure" it. This capacity is still very valuable. *See* STOCHASTIC PROCESS.

Queueing theory. Probably the most studied stochastic process is queueing. A queue or waiting line develops whenever some customer seeks some service that has limited capacity. This occurs in banks, post offices, doctors' offices, supermarkets, airline check-in counters, and so on. But queues can also exist in computer centers, repair garages, planes waiting to land at a busy airport, or at a traffic light.

Queueing theory is the prime example of what can be said about the state of the system but not how to improve it. Fortunately, the improvements can be made by increasing the resources available. For example, another teller opens up a window in a bank, or another check-out stand is opened in the supermarket. Other possible changes are more subtle. For example, the "eight items or less" express lane in a supermarket minimizes the frustration of a customer in line behind another with two full shopping carts. Some post offices and banks have gone to a single queue where the person at the head of the line goes to the next available window.

While it is not possible to generalize the results of queueing analysis, it is possible to provide some general measures. One is the load or traffic factor. This is the ratio of arrival rate to combined service rate, considering the number of service centers in operation. It is possible to plot queueing statistics against this factor as shown in **Fig. 6**.

When the load factor (ρ) is low, there is excess serving capacity. There may be occasional lines, but not often. As the load factor rises, a fairly linear rise is obtained in any queueing statistics until approximately the point where $\rho = 0.75$. After this there is a sharp and continuous rise in the lengths of line and waits. The system is becoming saturated. If $\rho = 1.00$, the best that can be said is, "the larger the line, the larger the wait."

The following problem is an example of the type of analysis that can be done on a simple queueing system. A check-out counter at a small market can handle customers at a rate of 20 per hour (3 min per customer). Since the number of customers will vary considerably throughout the day and throughout the days of the week, the arrival rate of customers at the check-out counter will also vary considerably. The manager wishes to know when to open up a second check-out counter.

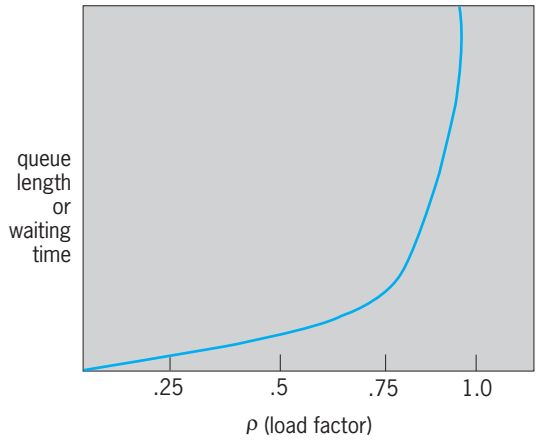

Fig. 6. Plot for general queues relating load factor to serving capacity.

The basic set of queueing formulas are given in Eqs. (1)–(5), where λ = the arrival rate; μ the service

$$\rho = \frac{\lambda}{\mu} \tag{1}$$

$$L = \frac{\lambda}{\mu - \lambda} \tag{2}$$

$$L_q = \frac{\rho^2}{1 - \rho} \tag{3}$$

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} \tag{4}$$

$$W = \frac{W_q + 1}{\mu} \tag{5}$$

rate; ρ = the traffic factor; L = the average number of customers in the system; L_q = the average number of customers in the check-out queue; Wq = the average waiting time for service; and W the average time in the check-out system. If the arrival rate varies from a low value to one approaching the service rate, the effects on the system can be examined. These results are presented in the **table**. From this table, it is evident that the queue builds up very rapidly as the traffic factor increases beyond 0.75. The store manager would need to monitor the actual number of customers in the queue and open another check-out counter when this number exceeds 4 or 5.

The mathematics of the probability distributions of arrival rates and service times often define closed-form solutions, that is, being able to directly solve for an answer. In these cases another technique has proved very useful. *See* QUEUEING THEORY.

Simulation. Simulation is defined as the essence of reality without reality itself. With stochastic processes, values are simulated for the random variables from their known or assumed probability distributions, a simpler model is solved, and the process is

Analysis of a simple queueing system

Arrival (λ)	Traffic factor (ρ)	Average number of customers in the system (L)	Average number of customers in check-out queue (L_q)	Average waiting time for service (W_q)	Average time in check-out system (W)
1	.05	.053	.0026	.0026	.0526
5	.25	.333	.0833	.0166	.0666
10	.50	1.000	.5000	.0500	.1000
15	.75	3.000	2.2500	.1500	.2000
16	.80	4.000	3.2000	.2000	.2500
17	.85	5.667	4.8161	.2833	.3333
18	.90	9.000	8.1000	.4500	.5000
19	.95	19.000	18.0500	.9500	1.0000

repeated a sufficient number of times to obtain statistical confidence in the results.

As a simple example of this, assume that someone makes the statement that the average female student at a university is of a certain height. To verify this assumption the actual heights of all female students could be measured, but this may be a long process. As an alternative, a hopefully random sample may be taken, the heights measured, the average determined, and an inference made as to the whole female population. *See* CHEMOMETRICS.

Many stochastic processes do not lend themselves to such a direct approach. Instead an assumption is made as to what the underlying probability distributions of the random variables are, and these are sampled. This sampling is done with random numbers. True random numbers are difficult to obtain, so pseudorandom numbers are generated by some mathematical relationship. This apparent paradox is justified by the fact that the numbers appear to be random and pass most tests for randomness. As an example, a queueing system could be modeled by having two wheels of fortune—one with numbers in random order from 1 to 15, and the other from 1 to 20. The first could represent the number of arrivals into the system per 10-min period, and the other the number served during the same period. In this manner a system with a load factor of 0.75 could be simulated. Many simulation computer languages have been developed to ease the work of analyzing queueing systems. *See* SIMULATION.

$$\begin{array}{cc} & \begin{array}{cc} \text{brand X} & \text{all others} \end{array} \\ \text{now} \begin{array}{c} \text{brand X} \\ \text{all others} \end{array} & \begin{bmatrix} .60 & .40 \\ .30 & .70 \end{bmatrix} \end{array}$$

Fig. 7. Probability matrix for brand-switching model.

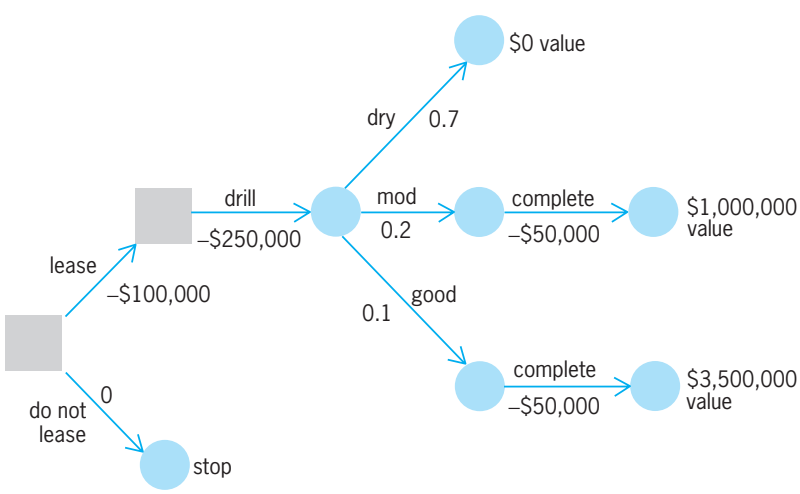

Fig. 8. Decision tree for an oil well problem.

Markov processes. This class of stochastic processes is characterized by a matrix of transition probabilities. These measure the probability of moving from one state to another. Such methods have been used to determine the reliability of a system, movement of a stock in the stock market, aging of doubtful accounts in credit analysis, and brand switching. This last application will be illustrated.

A company is considering marketing a new product. A preliminary market survey indicates that if a customer uses the product, there is a 60% probability that he or she will continue to do so. Likewise, the advertising campaign is such that 70% of those using all other products will be tempted to switch. The question is, what market share will the product be expected to finally obtain? The probability matrix is shown in **Fig. 7**.

Initially the market share of brand X is 0%, but after one period this will rise to 30%; that is, the advertising campaign will induce 30% to try brand X. In the next period, only 60% of these 30% will continue to purchase brand X, but 30% of the other 70% will switch, for a total market share of 39%.

Period by period, there will be transitions from state to state, but soon the variations will dampen out and a steady state will be achieved. This can be found for this problem by matrix value, and the final market share for brand X is 42.8%. *See* PROBABILITY.

Decision trees. While not originated by the field of operations research and not strictly in its domain, decision trees are an important tool in the analysis of some stochastic processes. More properly, they are a part of what may be considered statistical decision making. Their use can be illustrated with an example from the oil industry.

An oil company is developing an oil field and has

the option to lease the mineral rights on an adjacent block of land for $100,000. An exploratory well can be drilled at a cost of $250,000. If the well is considered moderate or good, it will cost an additional $50,000 to complete before production can start.

For simplicity, it can be assumed that the well, if good, has either moderate or good production. Also, the present worth of the net returns on all the future production is $1,000,000 and $3,500,000 for these two states. The problem facing the oil company is what they should do.

To analyze this problem, some subjective probabilities must be estimated. It is assumed that the probability that the well is dry is 70%, that the flow is moderate is 20%, and that it is good is 10%. With these percentages, a decision tree can be constructed (**Fig. 8**).

By a process of "folding back" the values and probabilities, the expected value at each decision point is seen to be represented by a square in Fig. 8. Actually, while there are three sequential decisions—lease, drill, and complete the well—once the decision to lease is made, the sequence is fixed unless the well is discovered to be dry. The expected value of this decision can be calculated as follows:

$$\begin{aligned} E[\text{lease} - \text{drill}] &= -100,000 - 250,000 + 0.70(0) \\ &\quad + 0.20(-50,000 + 1,000,000) \\ &\quad + 0.10(-50,000 + 3,500,000) \\ &= \$185,000 \end{aligned}$$

This expected value is slightly positive. In reality one of three things will happen—the well is dry, and the loss is $350,000; the well is moderate, with a net gain of $600,000; or the well is good, with a net gain of $3,100,000. This expected value can be interpreted as follows: If a large number of wells are drilled, 70% will be dry at a loss of $350,000 each, 20% will be moderate, and 10% will be good. The average gain will be $185,000.

Scope of Application

There are numerous areas where operations research has been applied. The following list is not intended to be all-inclusive, but is mainly to illustrate the scope of applications: optimal depreciation strategies; communication network design; computer network design; simulation of computer time-sharing systems; water resource project selection; demand forecasting; bidding models for offshore oil leases; production planning; assembly line balancing; job shop scheduling; optimal location of offshore drilling platforms; optimal allocation of crude oil using input-output models; classroom size mix to meet student demand; optimizing waste treatment plants; risk analysis in capital budgeting; electric utility fuel management; public utility rate determination; location of ambulances; optimal staffing of medical facilities; feedlot optimization; minimizing waste in the steel industry; optimal design of natural-gas pipelines; economic inventory levels; random jury selection; optimal marketing-price strategies; project management with CPM/PERT/GERT; air-traffic-control simulations; optimal strategies in sports; system availability/reliability/maintainability; optimal testing plans for reliability; optimal space trajectories.

It can be seen from this list that there are few facets of society that do not have an application for operations research. *See* DECISION THEORY; GERT; PERT.

William G. Lesso

Bibliography. R. E. Bellman and S. E. Dreyfus, *Applied Dynamic Programming*, 1962; V. N. Bhat, *Elements of Applied Stochastic Processes*, 2d ed., 1984; F. Budnick, D. McLeavey, and R. Mojena, *Principles of Operations Research for Management*, 1988; B. Cornet and H. Tulkens (eds.), *Contributions to Operations Research and Econometrics*, 1989; F. S. Hillier and G. Lieberman, *Introduction to Operations Research*, 7th ed., 2000; J. W. Sutherland, *Towards a Strategic Management and Decision Technology*, 1989; H. A. Taha, *Operations Research*, 6th ed., 1996.

Operator theory

At one level of abstraction an operator is simply a function whose arguments and values are real- (or complex-) valued functions of one or more real variables; in more naive terms an operator is a rule for converting such real- (or complex-) valued functions into others. The following are simple examples: (i) the operator which takes each differentiable real-valued function of one variable into its derivative; (ii) the operator which takes each twice-differentiable function f of one variable into expression (1); (iii) the operator which takes each

$$\left(\frac{df}{dx}\right)^2 + x^2\,\frac{d^2f}{dx^2} \tag{1}$$

twice-differentiable function f of three variables into expression (2); and (iv) the operator which takes

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \tag{2}$$

the continuous function f of one real variable into the function g where relation (3) holds.

$$g(x) \equiv \int_0^1 \sqrt{x+y}\,f(y)\,dy \tag{3}$$

Since an operator is a function, the usual functional notation is applicable. $L(f)$ may be used to denote the result of operating on f with the operator L. The set of all functions f for which $L(f)$ is defined is called the domain of L, and the set of all functions g such that $L(f) = g$ for some f in the domain of L is called the range of L. It is obvious that solving a differential or integral equation is equivalent (in many ways) to solving an operator equation $L(f) = g$, where g and L are given and it is required to find f. Moreover, the operator concept can be very useful both in theory and practice, producing a great variety of illuminating insights.

In large part the fruitfulness of the operator concept can be traced to two sources. One of these is

the possibility of adding and multiplying operators in such a way that many, though not all, of the laws of ordinary algebra hold. The other is the fact that the ranges and domains of operators behave in many respects like ordinary space and, indeed, may be regarded as contained in infinite dimensional generalizations of the familiar three-dimensional space of solid geometry. This makes it possible to think of an operator as a geometrical transformation and to exploit one's spatial intuition.

Let D be a family of real-valued functions such that $\lambda f + \mu g$ is in D whenever λ and μ are real numbers and f and g are in D. Let L and M be operators with domain D and range included in D. Then the operator which takes each f in D into $L[M(f)]$ is called the product of L and M and is denoted by LM. Moreover, the operator which takes each f in D into $L(f) + M(f)$ is called the sum of L and M and denoted by $L + M$. In particular, one may form powers $L^2, L^3, \ldots$, polynomials $a_0 + a_1L + a_2L^2 + \cdots + a_rL^r$, and in suitably restricted contexts, power series. It is important to note that, while it is always true that $L + M = M + L$, it is not always true that $LM = ML$. On the other hand, it is possible to show that $(LM)N = L(MN)$ and that $(L + M) + N = L + (M + N)$ for all L, M, and N so that parentheses may be omitted just as in ordinary algebra.

In many but not all cases the sets of functions with which one deals derive their spacelike properties from the possibility of assigning a distance $\rho(f,g)$ to each pair f and g of members of the set D under consideration. This is done in such a manner that $\rho(f,g) = \rho(g,f)$, $\rho(f,g) > 0$ if $f \neq g$; $\rho(f,f) = 0$ and $\rho(f,g) \leqq \rho(f,h) + \rho(h,g)$ for all f, g, and h in D. When D is as described in the preceding paragraph, ρ is often chosen so that $\rho(f,g) = \|f - g\|$ where $\|f\| = \rho(f,0)$. If $\|\lambda f\| = |\lambda| \, \|f\|$ for all real numbers λ, then $\|f\|$ is said to be a norm for D. There will usually be more than one way of norming a given D. For example, if D is the set of all real-valued continuous functions defined on the interval $0 \leqq x \leqq 1$, Eq. (4) gives one value for D, and Eq. (5) gives another value for D.

$$\|f\| = \max_{0 \leq x \leq 1} |f(x)| \tag{4}$$

$$\|f\|_1 = \sqrt{\int_0^1 |f(x)|^2 \, dx} \tag{5}$$

The analogy with the familiar space of experience is closest when the second norm is used, but the first is useful also.

The operator L is said to be linear if $L(\lambda f + \mu g)$ is defined and equal to $\lambda L(f) + \mu L(g)$ whenever f and g are in the domain of L, and λ and μ are numbers. Insofar as there is a general theory of operators, it is largely concerned with linear operators, and this article discusses linear operators exclusively. It is useful to develop this theory from axioms.

Axioms. Let F denote either the field of all real numbers or the field of all complex numbers. A vector space over F is a set or collection X whose members are of an unspecified character except that they may be added together and multiplied by the members of F in such a way that the following formal laws are satisfied:

1. $(f + g) + h = f + (g + h)$ and $f + g = g + f$ for all f, g, and h in X.
2. There is a unique zero vector 0 in X such that $f + 0 = f$ for all f in X.
3. $\lambda(\mu f) = (\lambda\mu)f$, $(\lambda + \mu)f = \lambda f + \mu f$, and $\lambda(f + g) = \lambda f + \lambda g$ for all f and g in X and all λ and μ in F.
4. $f = f$ for all f in X.

By generalizing the more special and concrete definition in the obvious fashion, a linear operator is defined to be a function L whose domain is a vector space X, whose range is in a vector space Y, and for which it is true that $L(\lambda f + \mu g) = \lambda L(f) + \mu L(g)$ whenever f and g are in X and λ and μ are in F.

Finite-dimensional case. A vector space X is said to be finite dimensional if it contains a finite subset $v_1, v_2 \ldots, v_n$ spanning the space in the sense that every element in the space may be written in the form $\lambda v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$ where the λ_j are in F. The representation $f = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n$ is unique if, and only if, no v_j is in the span of the rest. In this case $v_1, v_2 \ldots, v_n$ is said to form a basis for X, and the λ_j are said to be the coordinates of f with respect to this basis. It is not hard to show that any two bases for the same space have the same number of elements. This number is called the dimension of the space.

Let L be a linear operator whose domain X is finite dimensional. It follows immediately that the range is also finite dimensional and that the dimension d_R of the range is less than, or equal to, the dimension d_X of the domain. Let Y be the vector space containing the range of L, and let d_Y denote the dimension of Y. This gives $d_R \leqq d_Y$ and $d_R \leqq d_X$. The differences $d_Y - d_R$ and $d_X - d_R$ measure the extent to which the operator equation $L(f) = g$ fails to have a unique solution for all g. In fact if X_1 and X_2 are finite-dimensional subvector spaces of a vector space and $X_1 \supseteq X_2$, then $X_1 = X_2$ if, and only if, X_1 and X_2 have the same dimension. Thus $L(f) = g$ is always solvable if, and only if, $d_Y = d_R$ and $d_Y - d_R$ is a measure of the size of the set of g's for which no solution exists. On the other hand, it is easily seen that if f_0 is a particular solution of $L(f) = g$, then the general solution is $f_0 + h$ where h is any element of X such that $L(h) = 0$. The set of all such h is a vector space N, called the null space of L, whose dimension d_N measures the extent to which the equation $L(f) = g$ has multiple solutions. It is not hard to show that $d_N + d_R = d_X$ so that $d_N = d_X - d_R$. In the special but important case in which $X = Y$, $d_X - d_R = d_Y - d_R$. Thus either $L(f) = g$ has a unique solution for all g (nonsingular case) or else for many values of g, $L(f) = g$ has no solutions, and whenever it has any nonzero solutions, it has many (singular case).

There is a certain sense in which most operators are nonsingular. Let I denote the identity operator which takes every vector into itself. Then it can be shown that $L - \lambda I$ is nonsingular for all but a finite number of values of λ. Indeed $L - \lambda I$ will be singular if, and only if, there exists a nonzero vector f such that

$L(f) - \lambda f = 0$; that is, $L(f) = \lambda f$. Such an f is called a proper vector (or eigenvector) belonging to the proper value (or eigenvalue) λ. It is easy to show that proper vectors belonging to distinct proper values are linearly independent in the sense that no one is in the span of the rest. Thus the number of distinct proper values cannot exceed the dimension of the space. *See* DIFFERENTIAL EQUATION.

Knowledge of the proper values of an operator L yields a great deal of information about the nature of L, especially in the important case in which the domain X admits a basis made up of proper vectors. Let $\upsilon_1, \upsilon_2 \ldots, \upsilon_n$ be a basis for X, and let $L(\upsilon_j) = \lambda_j \upsilon_j$ where the λ_j are (not necessarily distinct) members of F. Very simple computations lead to the following observations.

1. Every proper value of L is equal to some λ_j.
2. L is nonsingular if, and only if, no λ_j is zero.
3. If L is nonsingular, the inverse operator carries $\mu_1\upsilon_1 + \mu_2\upsilon_2 + \cdots + \mu_n\upsilon_n$ into $(\mu_1/\lambda_1)\upsilon_1 + (\mu_2/\lambda_2)\upsilon_2 + \cdots + (\mu_n/\lambda_n)\upsilon_n$.
4. If P is any polynomial with coefficients in F, then $P(L)$ carries $\mu_1\upsilon_1 + \mu_2\upsilon_2 + \cdots + \mu_n\upsilon_n$ into $P(\mu_1)\upsilon_1 + P(\mu_2)\upsilon_2 + \cdots + P(\mu_n)\upsilon_n$.

The structure of $P(L)$ revealed by observation (4) suggests a definition of $F(L)$ where F, instead of being a polynomial, is an arbitrary function with domain and range in F. To wit: $F(L)(\mu_1\upsilon_1 + \mu_2\upsilon_2 + \cdots + \mu_n\upsilon_n) = F(\mu_1)\upsilon_1 + F(\mu_1)\upsilon_2 + \cdots + F(\mu_n)\upsilon_n$. A similar, but of course more subtle, definition in certain infinite-dimensional cases is the source of the modern rigorization of the celebrated operational calculus of O. Heaviside.

Returning to the case in which X need not equal Y, let $\upsilon_1, \upsilon_2 \ldots, \upsilon_n$ and $w_1, w_2 \ldots, w_n$ be bases for X and Y, respectively. Let $L(\upsilon_j) = \alpha_{j1}w_1 + \alpha_{j2}w_2 + \cdots + \alpha_{jn}w_n$ where each α_{ji} is in F. Then $L(x_1\upsilon_1 + \cdots + x_n\upsilon_n) = x_1L(\upsilon_1) + x_2L(\upsilon_2) + \cdots + x_nL(\upsilon_n) = y_1w_1 + y_2w_2 + \cdots + y_nw_n$ where $y_i = a_{1i}x_1 + \alpha_{2i}x_2 \ldots \alpha_{ni}x_n$. The rectangular array in which α_{ji} is in the ith row and jth column is called the matrix of L with respect to the basis in question. It is clear that solving the operator equation $L(f) = g$ in the finite-dimensional case is equivalent to solving m linear algebraic equations in n unknowns. The theory sketched above is the theory of such equations couched in the language of operator theory. When $X = Y$ and $\upsilon_j = w_j$ for all j, then the condition $\alpha_{ji} = \bar{a}_{ij}$ (where the overbar denotes complex conjugate) implies that there exists a basis for X made up of proper vectors of L. However, this condition is by no means a necessary one.

Infinite-dimensional case. When X is not assumed to be finite dimensional, such simple and general theorems as those described above are no longer available. In certain contexts, however, more complicated and less complete analogs of them may be proved. It is with these that the general theory of linear operators is mainly concerned.

Let X be a vector space which is not necessarily finite dimensional but instead is equipped with a norm—defined in the abstract case as suggested by the definition given above for real function spaces. Such a normed vector space is said to be complete if each sequence $f_1, f_2 \ldots$ of members of X which is convergent in the sense defined by Eq. (6) is also convergent in the sense defined by Eq. (7) for some

$$\lim_{\substack{n\to\infty \\ m\to\infty}} \| f_n - f_m \| = 0 \tag{6}$$

$$\lim_{n\to\infty} \| f_n - f \| = 0 \tag{7}$$

f in X. This f is easily seen to be unique and is called the limit of the sequence f_n. A complete normed vector space is called a Banach space. A normed vector space X is said to be separable if there exists a sequence $f_1, f_2 \ldots, f_n$ of elements of X such that every element of X is the limit of some subsequence. The space of continuous functions defined earlier is a separable Banach space. Now let $X = Y$ and let X be a Banach space. Let L be completely continuous in the sense that whenever $f_1, f_2 \ldots$ is a sequence of elements such that $\| f_n \| \leqq 1$ for all n, there is a subsequence $f_{n1}, f_{n2} \ldots$ such that $L(f_{n1}), L(f_{n2}) \ldots$ is convergent in the sense defined in Eq. (6). Then the following theorem can be proved. For each λ in F with $\lambda \neq 0$ there is a pair of vector subspaces M_λ and N_λ of X such that (1) $L(f)$ is in M_λ for all f in M_λ, and $L(f)$ is in N_λ for all f in N_λ; (2) every f in X can be written uniquely in the form $f_1 + f_2$ where $f_1 \epsilon M_\lambda$ and $f_2 \epsilon N_\lambda$; (3) for each $g_1 \epsilon M_\lambda$ there is one and only one element $f_1 \epsilon M_\lambda$ such that $L(f_1) - \lambda f_1 = g_1$; and (4) N_λ is finite dimensional. It follows easily that existence and uniqueness questions for the operator equation $L(f) - \lambda f = g$ reduce to the corresponding questions for the restriction of L to the finite-dimensional space N_λ and, hence, that the simple analysis given earlier applies. It may be proved further that M_λ coincides with X for all values of λ except those in a sequence $\lambda_1, \lambda_2 \ldots$ such that Eq. (8) applies.

$$\lim_{n\to\infty} \lambda_n = 0 \tag{8}$$

Let K be a continuous real-valued function defined on the unit square $0 \leqq x \leqq 1$, $0 \leqq y \leqq 1$. Let X be the Banach space of all continuous real-valued functions defined on the interval $0 \leqq x \leqq 1$ with Eq. (9)

$$\|f\| = \max_{0\leq x\leq 1} |f(x)| \tag{9}$$

applying. Then it can be proved that the operator L_K which takes f into g, where Eq. (10) holds, is

$$g(x) = \int_0^1 K(x, y) f(y)\, dy \tag{10}$$

a completely continuous linear operator. Application of the theorems quoted above yields most of the results of the Fredholm theory of integral equations. Such integral operators occur in inverting the members of a large class of linear differential operators. There are similar results for integral operators in $2n$ variables, it being possible to establish complete continuity whenever the region of integration is bounded.

Infinite-dimensional versions of theorems about bases of proper vectors and functions of operators take their simplest and most complete form when the underlying Banach space is a Hilbert space; that is, when there is defined an F-valued "inner product" $f \cdot g$ for each f and g in X which satisfies the following conditions: (1) $(\lambda f + \mu g) \cdot h = \lambda(f \cdot h) + \mu(g \cdot h)$; (2) $(f \cdot g) = (g \cdot f)$; and (3) $(f \cdot f) = \|f\|^2$ for all f, g, and h in X and all λ and μ in F. The simplest (and original) example of a Hilbert space is the vector space of all sequences $c_1, c_2 \ldots$ of complex numbers such that $|c_1|^2 + |c_2|^2 + \cdots < \infty$, based on the definition $(c_1, c_2 \ldots) \cdot (c'_1, c'_2 \ldots) = c_1\bar{c}'_1 + c_2\bar{c}'_2 + \cdots$.

Now let L be a completely continuous linear operator, the domain of which is a separable Hilbert space H and whose range is contained in H. Let L be self-adjoint in the sense that $L(f) \cdot g = f \cdot L(g)$ for all f and g in H. Then it is a theorem that there exists a sequence υ_1, υ_2, of members of H which has the following properties: (1) Each υ_j is a proper vector for L; (2) if f is any element in H and $c_j = f \cdot \upsilon_j$, then $f = c_1\upsilon_1 + c_2\upsilon_2 + \cdots$ in the sense that the partial sums of this series have f as a limit; and (3) $\upsilon_i \cdot \upsilon_j = 0$ if $i \neq j$ and $\upsilon_i \cdot \upsilon_i = 1$. Such a sequence is said to be an orthonormal basis for H. Just as in the finite-dimensional case, one can form more or less arbitrary functions of the operator L. In rough terms the celebrated spectral theorem is a generalization of the preceding theorem in which the operator L, though self-adjoint, is not required to be completely continuous. Instead of a discrete basis of proper vectors, one finds a sort of continuous basis. More precisely it is possible to map H onto a Hilbert space H' whose elements are complex-valued functions in such a manner that the norms and the vector space operations are preserved and so that L becomes the operator of multiplying the elements of H' by a fixed real-valued function.

In addition to the abstract theory there are many detailed studies of particular operators of importance, such as the Laplace and Fourier transforms, and numerous applications to differential and integral equations. Moreover, in addition to the theory of single operators there are extensive theories of certain kinds of collections of operators. A ring of operators contains the sum, product, and difference of any two of its members, and a group of operators contains the inverse of every operator in it and the product of each two operators in it. The theory of groups and rings of operators is related to the algebraic theory of groups and rings much as the theory of a single operator is related to systems of linear algebraic equations. It has applications to harmonic analysis and to the conceptual foundations of quantum mechanics. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES; GROUP THEORY; INTEGRAL TRANSFORM; REAL VARIABLE; RING THEORY; SET THEORY; TOPOLOGY.

George W. Mackey

Bibliography. N. I. Akheizer and I. M. Glazman, *Theory of Linear Operators in Hilbert Space,* 1961, reprint 1983; J. B. Conway, *A Course in Functional Analysis*, 1994; N. Dunford and J. T. Schwartz, *Linear Operators,* 3 vols., 1958–1971, reprint 1988; L. Hoermander, *The Analysis of Linear Partial Differential Operators,* 2d ed., 1990; A. Taylor and D. Lay, *Introduction to Functional Analysis,* 2d ed., 1980, reprint 1986; T. Yoshino, *Introduction to Operator Theory*, 1994.

Operon

A group of distinct genes that are expressed and regulated as a unit. Each operon is a deoxyribonucleic acid (DNA) sequence that contains at least two regulatory sites, the promoter and the operator, and the structural genes that code for specific proteins (see **illustration**). The promoter (p) site is the location at which ribonucleic acid (RNA) polymerase binds to the operon. RNA polymerase moves down the operon catalyzing the synthesis of a messenger RNA (mRNA) molecule with a sequence that is complementary to DNA. This process is called transcription. The mRNA is used as a template by ribosomes to synthesize the proteins coded for by the structural genes (in the original DNA) in a process called translation. This mRNA is referred to as polycistronic because its sequence directs the synthesis of more than one protein. The operator (o) site is located between the p site and the beginning of the coding region for the first structural gene. It is at this site that molecules called repressors can bind to the DNA and block RNA polymerase from transcribing the DNA, thus shutting off the operon. Some systems can be derepressed by the addition of small molecules called effectors, which bind to the repressor protein and cause a conformational (shape) change that makes it no longer able to bind to the DNA at the operator site. *See* DEOXYRIBONUCLEIC ACID (DNA); RIBONUCLEIC ACID (RNA).

Promoters. Promoters are located just before the start of the first gene in an operon, and have two conserved sequences separated by 16–19 base pairs. The −10 region has the consensus sequence of TATAAT (T = thymine; A = adenine). Although most promoters differ to some degree from the consensus sequence, little variability in the separation of two regions, is tolerated. The reason for this sensitivity is DNA's helical structure, where binding sites spiral around the DNA molecule. The required separation of the −10 and −35 regions determines whether they are on the same side of DNA. *See* MOLECULAR BIOLOGY.

Repressors. There are two types of repressors that block or repress the production of a protein. The

The lactose (*lac*) operon from *Escherichia coli*: *z*, *y*, and *a* are structural genes; *i* is the *lac* repressor gene; *p* is the promoter site; and *o* is the operator site. The arrow indicates length and direction of mRNA synthesis.

better understood are the transcriptional repressors. These proteins bind to the *o* site of the operon. They repress transcription of the gene through blocking binding of RNA polymerase at the promoter or by blocking the movement of RNA polymerase through the gene. In the simplest case, this blockage can be direct by having the operator overlap the promoter. This blockage can also be indirect through the formation of loops of DNA which hide the promoter. Such is the case with the C protein in the arabinose operon from *Escherichia coli*. When the levels of C protein is high, it binds to two sites that are 160 base pairs apart. Portions of the C protein self-associate, forming a loop that inhibits transcription. In this manner, sites that are quite distant from the promoter and the first gene in the operon can influence transcription.

A second type of repression is translational. In order for mRNA to be translated, ribosomes must first bind to the message and then read through the codes adding amino acids to the growing protein chain. In translational repression the repressor binds to mRNA so that the ribosome binding site is blocked.

Activators. At a number of locations within the *E. coli* genome, a regulator called the catabolite gene activator protein acts as a repressor, but there are other sites at which it stimulates the expression of genes within an operon.

Activation is believed to arise from the binding of a protein immediately adjacent to the promoter. The protein provides additional locations with which RNA polymerase can interact; the extra interactions result in an increased amount of polymerase binding to the promoter. Activators are more frequently involved in the regulation of genes in eukaryotes than in prokaryotes.

Antiterminators. Once RNA polymerase begins transcribing a gene, it continues making RNA until a termination site is reached. Antiterminators are proteins that prevent termination at certain sites either by preventing mRNA from forming a hairpin or by preventing the mRNA from falling off if a hairpin is formed. In the presence of these antiterminators, RNA polymerase continues along the genome and transcribes the genes following the termination site until a different class of termination site is encountered.

Attenuation. Attenuation is the premature termination of the mRNA translation. Although the exact mechanism of attenuation has not been unambiguously determined, it is thought that attenuation is due to the formation of a translation termination site in mRNA. As soon as an mRNA strand is synthesized, while it is still associated with DNA, ribosomes begin to bind and produce protein. If the level of tryptophan in the cell is high, the ribosome speeds through this region; the mRNA folds as shown and produces a translation termination site. Since no mRNA for the remaining structural genes is made, the corresponding proteins are not synthesized. If the level of tryptophan is low, the ribosome slows when it comes to that region of mRNA coding for tryptophan. This is believed to give the MRNA time to refold into a structure which has no termination site.

Antirepressor. A protein called an antirepressor does not appear to bind DNA but rather repressors. The production of small amounts of this protein is sufficient to bind a substantial amount of repressors. Transcription is activated as the repressor is removed.

Methylation of DNA. The encoding of a protein that adds a methyl group to the adenine in a particular genetic sequence can occur. When the GATC sequence is found in the promoter of a gene, such methylation might be expected to influence transcription of that gene. In eukaryotic cells, approximately 5% of the cytosines in DNA are methylated. During some viral infections, the expression of the viral genes appears to be repressed by methylation of the viral DNA. It is possible that during development from egg to embryo the expression of some genes may be regulated through methylation of the DNA. The mechanism that causes the level of methylation of the DNA to change is not known. *See* GENE; GENE ACTION; IMMUNOLOGY. Douglas H. Ohlendorf

Bibliography. B. Lewin, *Genes*, 1994; J. Miller and W. Reznikoff (eds.), *The Operon*, 1978.

Ophioglossales

An order of the class Polypodiopsida known as the adder's-tongue ferns. It is a small group with only 3 genera and about 80 species. Two genera, *Ophioglossum* and *Botrychium* (see **illustration**), are widely distributed in tropical and temperate regions and

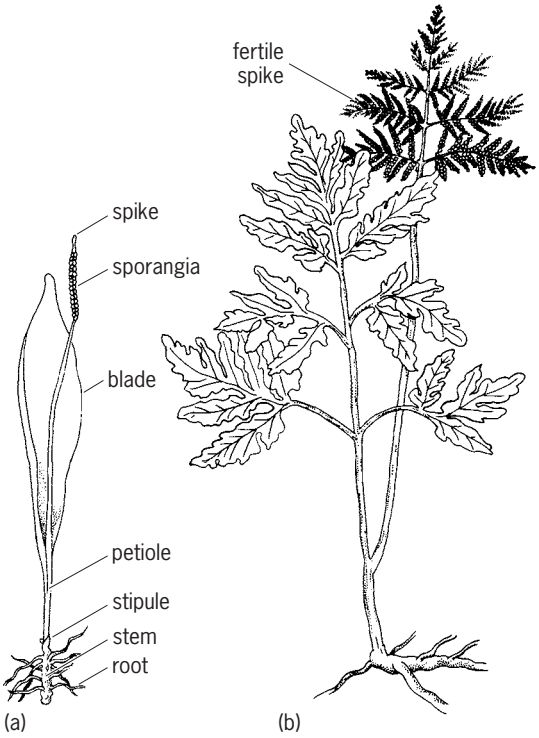

Ophioglossales. (*a*) Adder's-tongue fern (*Ophioglossum*) with a single leaf. (*b*) Grape fern (*Botrychium*), showing the division of the leaf into a sterile foliage part and a fertile spike. (*After H. J. Fuller and O. Tippo, College Botany, Holt, rev. ed., 1954*)

have about the same number of species; the third genus, *Helminthostachys*, is represented by a single species confined to southeastern Asia and Polynesia. These are considered the most primitive of the present-day ferns. No fossils have been reported for this group. The plants are homosporous and eusporangiate; that is, spore sacs develop from groups of epidermal cells. Each sporangium produces a large number of spores, as many as 15,000 in some species. The chromosome number is high in the species of *Ophioglossum*. *Ophioglossum petiolatum*, a tropical species, has a chromosome count of over 1000, the largest number observed in a naturally occurring species of vascular plants. This group is distinguished from other ferns by the arrangement of the sporogenous tissue in the characteristic fertile spike (illus. *a*) of the sporophyte (spore-producing generation). The leaves are erect or merely bent over in bud and not circinate. The gametophyte (gamete-producing generation) is a small, nongreen, fleshy, subterranean saprophyte, associated with an endophytic fungus. The group appears to be an evolutionary dead end. *See* FUNGI; LEAF; POLYPODIALES; POLYPODIOPSIDA; PTEROPSIDA. Paul A. Vestal

Ophiolite

A distinctive assemblage of mafic plus ultramafic rocks generally considered to be fragments of the oceanic lithosphere that have been tectonically emplaced onto continental margins and island arcs. Ophiolite was named by A. Brongniart, a nineteenth-century French naturalist, who considered its scaly appearance and the greenish color of its main constituent rock, serpentinite. An ophiolite is a formation made up of an association of typical rocks in a clearly defined sequence. As shown in **Fig. 1**, a complete idealized ophiolite sequence from bottom to top includes (1) an ultramafic tectonite complex composed mostly of multilayered, deformed harzburgite, dunite, and minor chromitite; (2) a plutonic complex of layered mafic-ultramafic cumulates at the base, grading upward to massive gabbro, diorite, and possibly plagiogranite; (3) a mafic sheeted-dike complex; (4) an extrusive section of massive and pillow lavas, pillow breccias, and intercalated pelagic sediments; and (5) a top layer of abyssal or bathyal sediments, which may include ribbon chert, red pelagic limestone, metalliferous sediments, volcanic breccias, or pyroclastic deposits. Most ophiolites lack complete sections, and are dismembered and fragmented. Their estimated original thickness is variable, ranging from about 2 km (1.2 mi) to more than 8 km (5 mi). *See* EARTH CRUST; LITHOSPHERE.

Occurrence. Ophiolites typically occur in collisional mountain belts or island arcs and define a suture zone marking the boundary where two plates have welded together. The ophiolite complex is interpreted as evidence for a closed marginal ocean or back-arc basin. Typical examples of such sutures are found along the northern flank of the Himalayas (the Indus suture) and in the central Urals; both extend for more than 1000 km (600 mi). The occurrence of a suite of deep-sea sediments, pillow basalts, gabbros, and serpentinized ultramafic rocks within these sutures suggests that they constituted oceanic lithosphere that was subsequently thrust onto the continental margins by a process known as obduction.

According to plate tectonics, the destiny of the ocean floor is subduction into the mantle. In contrast, geologists have proposed the term "obduction" to describe the particular destiny of ophiolites that, during the advance of oceanic lithosphere against a continent, end up stranded on the edge of a continent instead of disappearing into the subduction zone below it. *See* PLATE TECTONICS; SUBDUCTION ZONES.

Throughout the world, ophiolites occur as long narrow belts, up to 10 km (6 mi) wides, that can extend more than 1000 km (600 mi) in length, in two distinct geographic settings. (1) Those in the Alpine-Mediterranean region, Tethyan ophiolite, were formed in small ocean basins that were surrounded by older, attenuated continental crust. Many of the classical ophiolites found in Cyprus and Oman belong to this group, which have a nearly complete sequence and were brought above sea level during the collapse of small ocean basins by the convergence of neighboring plates. (2) Those in western North America and the Circum-Pacific (Cordilleran) region seem to have formed in inter-arc basins. The Cordilleran ophiolites, such as the Trinity ophiolite and the Coast Range ophiolite of California, are generally incomplete, metamorphosed, or dismembered, but they commonly form the basement rocks for many North American continental margin terranes. Tethyan ophiolites are characterized by the occurrence of harzburgitic ultramafic rocks and pronounced thick layers of gabbroic rock, whereas Cordilleran ophiolite, such as the Trinity, have undepleted lherzolites and thin layers of recrystallized and deformed gabbros. Such differences have been attributed to the rate of spreading, with fast spreading accounting for more partial melting of the primary mantle rocks, and hence the thicker mantle residue as harzburgite and greater differentiation of gabbroic magma for the mafic-ultramafic rocks of the Oman ophiolite. *See* BASIN; CONTINENTAL MARGIN; GEODYNAMICS; STRUCTURAL GEOLOGY.

These two types of ophiolites strongly resemble the oceanic lithosphere insofar as the latter is known from dredging, shallow drilling, and geophysical studies. However, the crustal section of most ophiolites is significantly different from the abyssal oceanic crust of the Atlantic and Pacific oceans. Many ophiolitic basalts possess chemical affinities of back-arc basin basalts or island-arc basalts, rather than those of mid-oceanic ridge basalt (MORB). For example, the potassium oxide (K_2O) content in MORB is usually less than 0.2 wt %, whereas it is higher than 0.2 wt % in back-arc basin basalts and the extrusive rocks of the ophiolite sequence. Many ophiolites carry a weak-to-strong subduction zone geochemical imprint with relative depletions in tantalum and niobium in normalized incompatible-element

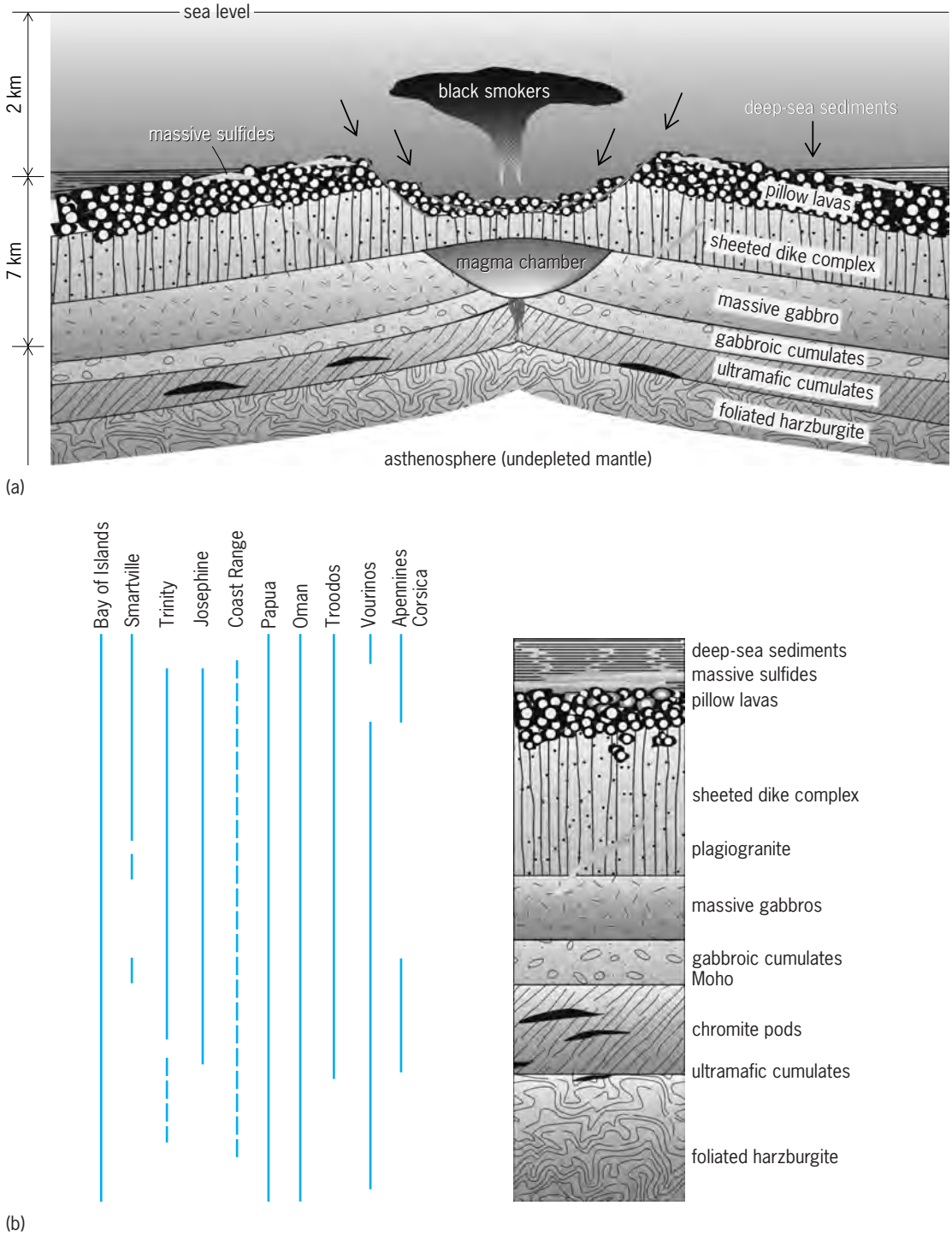

Fig. 1. Idealized cross section of (*a*) an oceanic ridge showing stratigraphy of the intrusive and extrusive sequences, and circulation of heated seawater at the top of the oceanic crust, and (*b*) an idealized ophiolite succession compared with various exposed ophiolites.

distribution diagrams compared to MORB. The sedimentary sequence overlying most ophiolites grades from thin layers of pelagic or cherty sediments upward to turbidite or calc-alkaline volcaniclastics, suggesting that the oceanic crust of most ophiolites formed in back-arc basins. This is also supported by the fact that the crustal sequence of most ophiolites is thin (<5 km; 3 mi) in comparison with average oceanic crust developed in large ocean basins (7–10 km; 4–6 mi). *See* BASALT; MID-OCEANIC RIDGE; OCEANIC ISLANDS.

Origins. Many ophiolites are thought to have formed in submarine extensional tectonic settings with extremely high heat flow, such as the present-day East Pacific Rise and Mid-Atlantic Ridge, where new oceanic crust is being generated. The analogy of

the ophiolite sequence with the oceanic lithosphere (Fig. 1) is supported by the gross similarity in chemistry, metamorphic grades corresponding to temperature gradients existing under spreading centers, the presence of similar ore minerals, and the occurrence of deep-sea sediments. However, in recent years this simple analogy has been challenged. It has been suggested that ophiolites do not represent the typical oceanic lithosphere and do not belong to a unique species; instead they have formed in other extensional regimes, including island-arc or marginal-basin settings, above a subduction zone. Ophiolites are derived from a variety of oceanic sites on which new lithosphere is formed. Despite this ongoing controversy, the general mechanism by which a complete ophiolite succession forms is reasonably well understood and agreed upon.

In extensional oceanic environments, heat flow is high and the mantle asthenosphere is rising. As the pressure decreases in the rising asthenosphere beneath ocean ridges or back-arc basins, undepleted mantle lherzolites partially melt to form basaltic magma according to the general reaction, lherzolite $\longrightarrow$ harzburgite (75%) + basaltic magma (25%). These magmas collect in a chamber (or chambers) at depths about 4–6 km (2.5–4 mi), and 25 km (16 mi) long, and undergo fractional crystallization. Some magma rises as dikes, forming a sheeted-dike complex, and are extruded as pillow lavas in the axial rift on the sea floor. The remaining magma within the chamber fractionates upon cooling to form the layered and massive rocks of the plutonic sequence. Fractional crystallization of the magma gives rise locally to diorite and plagiogranite. The petrologic and chemical data indicate that the lavas, dikes, gabbros, and underlying tectonized harzburgite are all cogenetic; the harzburgite represents a crystalline residue from partial melting in the mantle that produced the overlying igneous rocks. *See* ASTHENOSPHERE; IGNEOUS ROCKS; MAGMA; PLUTON.

Ages of formation and emplacement. At the base of most ophiolites, a thin metamorphic sole occurs and records the travel history of the obduction of the oceanic lithosphere on land. The sole ranges in thickness from 10 to 50 m (33 to 165 ft) and can extend laterally for more than 100 km (60 mi). It is composed of highly deformed amphibolites and metasedimentary greenschists, and shows a sharp decrease in metamorphic grade from top to bottom. When young and hot oceanic lithosphere thrusts upon the oceanic crust near a continent margin, it heats the underlying crustal rocks like a flat iron. These rocks start to recrystallize while being deformed under the moving load of the overlying lithosphere, and are subjected to intense metamorphism. During its oceanic travels at a rate of several centimeters per year, the ophiolite nappe loses heat; the temperature at its base will have dropped and the underlying crust will transform into low-temperature metamorphic rocks, such as greenschist. The metamorphic soles formed, apparently, by successive underplating and welding onto the base of an ophiolite as the hot, young oceanic slab migrated toward the continent. Dating the metamorphic soles' thermally recrystallized minerals provides ages for ophiolite emplacement onto the continental margins. *See* METAMORPHIC ROCKS.

Ages of ophiolite formation can be obtained by direct radioactive lead (U/Pb) dating of zircon from diorite or plagiogranites of the plutonic sequence of an ophiolite. They can also be constrained by determining the ages of radiolarian fossils in pelagic chert overlying these ophiolites. Ophiolites are rather abundant in Phanerozoic orogenic belts: however, several Precambrian ophiolites with ages ranging from 600 million years to about 2 billion years have also been described. Age gaps between the deposition of pelagic sediments and ophiolite emplacement are probably less than 25 million years. Such a short duration is consistent with the life span of a back-arc basin (less than 20 million years), and indicates that obduction of many ophiolites occurred soon after their creation. Ophiolites consequently represent young oceanic lithosphere that was detached while still hot. *See* LEAD ISOTOPES (GEOCHEMISTRY); ROCK AGE DETERMINATION; ZIRCON.

Hydrothermal alteration and formation of massive sulfide deposits. Immediately after new oceanic crust forms, seawater percolates downward through fractures and faults from the flanks of the rifted valley. It easily penetrates the crust down to 2–3 km (1–2 mi) at the base of the dike complex. Heated to 400–450°C (752–842°F), the water circulates toward the ridge and starts to ascend, becoming progressively channeled, and finally discharges along the rift axes at temperatures up to 380°C (716°F). During this high-temperature circulation, the hydrothermal solution alters and corrodes the crustal rocks, dissolving metals. When this discharged hydrothermal solution, consisting of metal sulfides leached from the oceanic crust, meets the cold seawater on the ocean floor, a black cloud of metallic sulfides results. Such hot springs are known as black smokers. Subsequent precipitation at, around, and beneath these chimneys can produce substantial amounts of massive sulfide ores in the upper parts of the extrusive sequence (Fig. 1). Such processes for precipitation of massive sulfide deposits were observed on the present-day ocean floor at locations along Pacific, Atlantic, and Indian ocean ridges. *See* HYDROTHERMAL VENT.

Figure 2 shows a black smoker spewing dark, mineral-rich fluids as observed by scientists during a dive of the deep-sea submersible *Alvin* on the East Pacific Rise (latitude 21°N) in 1979. The dissolved minerals and the heat given off by the black smoker favor the prolific growth of chemoautotrophic bacteria which, in turn, provide nourishment for the surrounding colonies. Around the hot springs, abundant and unusual sea life—giant tubeworms, huge clams, and mussels—has been discovered. *See* MARINE MICROBIOLOGY.

In the upper parts of the extrusive sequences of many ophiolites, thin lenses (up to 50 m or 165 ft thick and ~500 m or 1650 ft across) of sulfide ore occur in some fault-controlled depressions. The

Fig. 2. View of the first high-temperature geothermal vent (380°C or 716°F) ever seen by scientists during a dive of the deep-sea submersible *Alvin* on the East Pacific Rise in 1979. (*Photograph by Dudley Foster, RISE Expedition, courtesy of W. R. Normark, U.S. Geological Survey*)

ore lenses consist of massive sulfide minerals, including pyrite (FeS_2), chalcopyrite ($CuFeS_2$), sphalerite (ZnS), pyrrhotite (FeS), and minor molybdenite (MoS_2), and have been mined as major copper deposits. Such copper-zinc deposits are presently exploited in the basalts of ophiolite sequences. The best-known copper deposit is located on Cyprus and is associated with the genesis of the Troodos ophiolites, and it was known to the ancient Greeks, who named the island after the metal (*cupros*, or copper). These Cyprus deposits contain a total reserve of 1.5 megatons of copper. A similar, contemporaneous mineralization of the Kuroko-type black-ore deposits in Japan was generated along embryonic oceanic ridges, which developed in a back-arc setting. *See* COPPER.

The circulation of heated seawater through cooling, fractured crustal sections also causes significant alteration of the primary minerals. Formation of secondary minerals, including clay minerals and zeolites in basaltic rocks, is selective and incomplete and increases the volatile content of oceanic metabasalts. Hydrothermal metamorphism at depths results in the formation of epidote, chlorite, and amphiboles in the gabbroic sequence, and minor serpentinization of the ultramafic rocks. *See* SERPENTINITE.

Summary. Ophiolites represent new oceanic crust formed in a variety of spreading environments, including oceanic ridge, back-arc basin, and island arcs above a subduction zone, and subsequently emplaced onto the continents. Their occurrence along plate sutures marks the sites of ancient tectonic interaction between oceanic and continental crust. Ophiolites can form in multitude of tectonic settings, and the process of new ocean crust generation in spreading centers can produce different magma types. Depending on the rate of influx of ascending asthenosphere and the spreading rate of the oceanic lithosphere, ophiolites may not consist of plutonic and extrusive sequences resulting from a single magmatic pulse, but several, evidenced by the many intrusive relations. For example, the plagiogranites of the best-preserved Oman ophiolite are the product of fractional crystallization of the second magmatic event. Similarly, many Tethyan ophiolites, including the one in Oman, are characterized by thick harzburgites and well-layered gabbros resulting from fast spreading and large magma chambers. Many Circum-Pacific ophiolites, such as the Trinity ophiolite of California, have relatively thin lherzolitic mantle sequences, which are dismembered and recrystallized as a result of slow spreading at their ridges. In spite of extensive research efforts, it is still difficult to establish what kind of oceanic crust some ophiolites represent, what the actual processes were that formed ophiolites in spreading systems, how particular ophiolites formed in various tectonic settings compared with the oceanic lithosphere, and how they were emplaced onto the continental margins. Nevertheless, ophiolites provide the best opportunity for geologists to walk across the ocean floor on land; they also offer vertical sections in addition to horizontal distributions. Moreover, ophiolite formations record the ages of oceanic fragments that escaped disappearing into subduction zones.

Juhn Liou; Shige Maruyama; Yoshi Ogasawara

Bibliography. American Geological Institute, Penrose Field Conference on Ophiolites, *Geotime*, 17:24–25, 1972; K. C. Condie, *Plate Tectonics and Crustal Evolution*, 1997; R. G. Coleman, The diversity of ophiolites, *Geologic en Mijinbow*, 16:141–150, 1984; R. G. Coleman, *Ophiolites—Ancient Oceanic Lithosphere?*, 1977; I. G. Gass, Ophiolites, *Sci. Amer.*, pp. 122–131, August 1982; R. A. Kerr, Ophiolites: Windows on which ocean crust?, *Science*, 219:1307–1309, 1983; R. Mason, Ophiolites, *Geol. Today*, 1:36–40, 1985; E. M. Moores, Origin and emplacement of ophiolites, *Rev. Geophys. Space Phys.*, 2014:735–760, 1982; A. Nicolas, *The Mid-oceanic Ridges: Mountains below Sea Level*, 1995.

Ophiuroidea

A class of the Asterozoa, known as the brittle stars, in which the arms are usually clearly demarcated from a central disk and perform whiplike locomotor movements, and the tube feet are nonsuctorial sensory tentacles. In all existing ophiuroids the ambulacral plates fuse together in pairs to form articulating joints termed vertebrae, and the ambulacral groove is converted into an internal epineural canal.

As **Fig. 1** shows, typical ophiuroids have a distinctive shape unlike that of the other Asterozoa, and the vigorous lashing movements of the arms distinguish them at once from asteroids. However, fossil ophiuroids are known which approach asteroids in structure, and the two groups are evidently closely related. *See* ASTEROIDEA.

Ophiuroids are usually five-armed, with four- or six-armed individuals occurring as abnormalities; but a few species are regularly six- or seven-armed. In some Euryalae the arms may branch repeatedly; these are the so-called basket stars. In the smallest

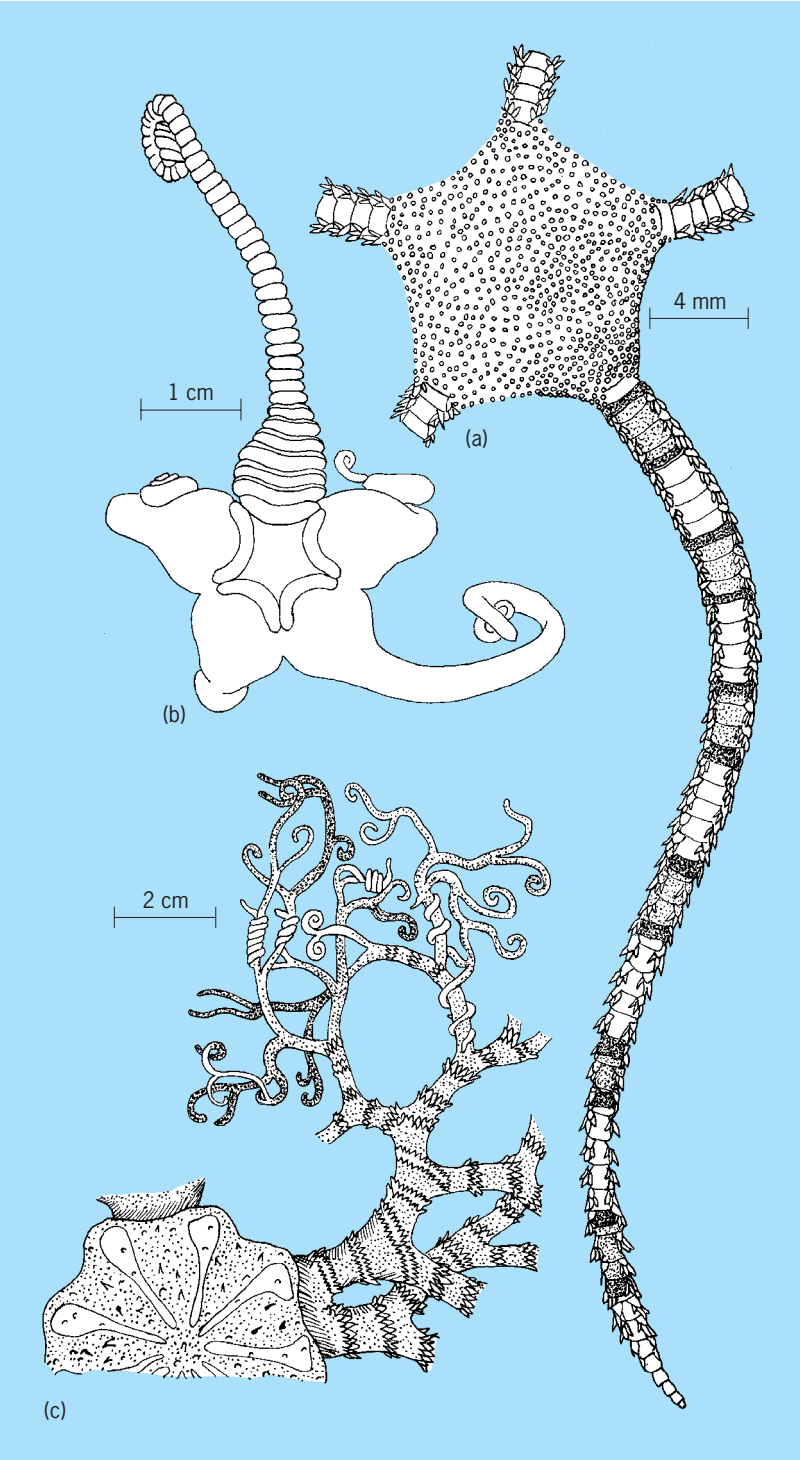

Fig. 1. Representative ophiuroids, which have a quite distinctive shape. (*a*) *Pectinura cylindrica*. (*b*) *Astroporpa wilsoni*. (*c*) *Gorgonocephalus chilensis*.

species the disk and arms together may be no more than a few millimeters across. Large species may have a disk 4 in. (10 cm) across and an arm spread of 20 in. (50 cm). Tropical species are often patterned in contrasting colors, but most ophiuroids tend to match their environment. Their biochromes do not include echinochromes. Some species are luminescent, although not constantly so; the light is emitted by cells at the bases of the arm spines, usually only on stimulation. Such species are known for the genera *Amphiura*, *Amphipholis*, and *Ophiacantha*.

Classification. There are about 1900 extant species referred to 230 genera, arranged to form 3 living orders: Oegophiurida, Phrynophiurida, and Ophiurida. There is also one Paleozoic order, the Stenurida.

Oegophiurida. Oegophiurida is the most primitive order of brittlestars (Ophiuroidea), comprising a single living genus, *Ophiocanops*. In oegophiurids, the gonads, together with outpouches of the digestive system (gastric ceca), extend from the central body cavity into the arms. This is a characteristic that they share with asteroids. The digestive system is confined to the disc in all other ophiuroids (order Ophiuridea), and in only a small number of primitive taxa do the gonads extend into the arms. In addition, the vertebral ossicles that make up the arms in Oegophiurida, though paired and firmly united, are not fused distally, whereas left and right vertebral ossicles are always fused in other extant ophiuroids. Oegophiurida also primitively lack radial and oral shields in their disc, as well as ventral arm plates. Andrew B. Smith

Phrynophiurida. Members of this order of Ophiuroidea are characterized by vertebrae that usually articulate by means of hourglass-shaped surfaces and arms that are able to coil upward or downward in the vertical plane. There is usually a leathery integument, in which calcareous granules or platelets are embedded. Most species are found in deep water, and often the arms are tightly coiled about the branches of black corals, upon which Phrynophiurida feed. Of the families, the Gorgonocephalidae often have branched arms, the Asteronychidae have a large disk and slender arms, and the Asteroschematidae have a small disk and stout arms.

The foregoing families share a number of characteristics and are grouped in one suborder, Euryalina. One remaining family, the Ophiomyxidae, differs in having a soft, unprotected integument, like that of *Ophiocanops*, but lacks the peculiar features of the gut and gonads in oegophiurids. For reasons too specialized to discuss here, it appears best not to associate the Ophiomyxidae with the Euryalina, in the Phrynophiurida, placing the family in a distinct suborder, the Ophiomyxina.

Ophiurida. Members of this order of Ophiuroidea possess vertebrae that articulate by means of ball-and-socket joints, and the arms, which do not branch, move mainly from side to side and do not coil in the vertical plane. The disk and arms are usually sheathed in regularly arranged plates. These are disposed in four series on the arms, namely, one dorsal, one ventral, and two lateral. There is a single madreporite. The order embraces most known genera of brittle stars and includes 13 families.

Howard B. Fell

Stenurida. Individuals in this Paleozoic (Ordovician–Devonian) order have a double row of plates (ambulacra) that abut across the arm axis either directly opposite one another or slightly offset. In contrast, modern ophiuroids have a single series of axial arm

plates termed vertebrae. In stenurids, as in modern ophiuroids, lateral plates are present at the sides of ambulacrals, and prominent lateral spines are typical. Stenurids lack the dorsal and ventral arm shields that are found in most ophiuroids. Proximal ambulacral pairs can be partially separated, forming a buccal slit, an expansion of the mouth frame. The arms of some stenurids are slender and flexible, but those of others are broad and comparatively stiff. The central disk varies from little larger than the juncture of the arms to an expansion that extends most of the length of the arms. The content of the order is poorly established, and fewer than 10 genera are known.

The relationships among ophiuroids, asteroids, and all other echinoderms provide an enduring problem in invertebrate evolution. Developmental and other studies based on modern organisms imply that asteroids and ophiuroids are not closely related within the echinoderms. Stenurid morphology, in contrast, suggests a close common ancestry for the two; the nature of the ambulacral plates is important, but even their general form is transitional.

Daniel B. Blake

Skeleton. The Paleozoic families had open ambulacral grooves on the undersurface of the arms, but in all extant forms the groove has sunk inward, and the ventral midline of the arm usually carries a median row of plates. In five Paleozoic families the ambulacral plates are paired, as in asteroids; in all other ophiuroids these plates fuse in pairs to produce jointed vertebrae, and the corresponding adambulacral plates become the so-called lateral plates on either side of the arm. In most extant forms the arm comprises a series of jointed segments, each one containing a vertebra, and each one covered externally by the ventral plate, right and left lateral plates, and a dorsal plate (**Fig. 2***a*). These features are used in diagnosing the several orders. Spines are often carried by the skeletal plates, especially where the lateral arm plates represent the adambulacral spines of asteroids. In the order Ophiurida they commonly form an erect fringe to the sides of the arms (Fig. 2*b*); in the suborder Euryalina they commonly are transformed into clubs or hooklets, and hang downward.

Muscular system. The arms in extant ophiuroids are provided with well-developed longitudinal muscles linking the successive vertebrae. The two extant suborders are each characterized by a distinctive arm movement, horizontal in the Ophiurida, and vertical in the Euryalina. This distinction is related in either case to the form of the vertebrae, which in turn are defined in the taxonomic diagnoses. Ophiurida move rapidly when disturbed. One of the arms is either trailed or pushed ahead, whereas the other four arms operate as two opposite pairs of levers, thrusting the body forward in a series of rapid jerks. Unlike asteroids, the tube feet play little part in locomotion, or none at all. In very young stages, however, the tube feet may be used as stilts, and seem to be weakly adhesive. Euryalina have bigger vertebrae and smaller muscles, and their movements are less spasmodic;

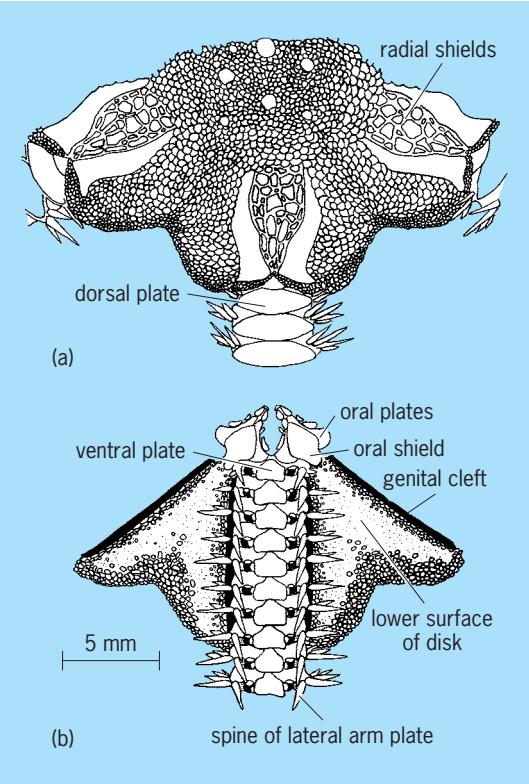

Fig. 2. External features of the ophiuroid *Amphiura*. (*a*) Aboral view. (*b*) Adoral view.

but they can coil the arms very firmly around objects, and keep their hold after death.

Alimentary system. The mouth, in the middle of the underside of the disk, serves for both ingestion and egestion of food, because there is no anus in existing forms. The gut comprises only the saclike stomach, in the walls of which are glandular hepatic cells. The stomach sends blind ceca into the arms in the Ophiocanopidae, a character recalling the asteroids.

Although many ophiuroids are predatory on small organisms when opportunity offers, most of them evidently spend most of their time scavenging in detritus, eating whatever they find. They are selective, because their gut does not permit the gross mud swallowing practiced by asteroids in similar circumstances. Thus, *Pectinura* will feed selectively on beech pollen in season in the New Zealand fiords, where the trees overhang the water. Among the more consistently predatory ophiuroids are certain Euryalina which cling to the branches of black corals and browse upon the polyps. Ophiuroids which live in large, dense populations evidently rely upon a steady flow of suspended matter, and there is evidence that sea-floor currents supply this. Basket stars sometimes live in these conditions rhythmically sweeping the branched arms toward the mouth.

Water-vascular system. This system is typical for the phylum, but the tube feet lack ampullae and

suckers, and can be retracted into so-called tentacle pores. The madreporite normally lies in one interradius of the adoral surface, but some Euryalina have a madreporite in each interradius.

Nervous system. This system is also typical for the phylum. No organs of special sense are known. It may be inferred that photoreceptors and chemoreceptors are distributed on the ectoderm of the ambulacral tube feet. *See* NERVOUS SYSTEM (INVERTEBRATE).

Reproduction. Ophiuroids may take up to 2 years to reach sexual maturity, and full growth may require 3 or 4 years. The life span is unknown but may be estimated at about 5 years. Large Euryalina, such as *Gorgonocephalus*, may well live much longer. The sexes are usually separate, but a few species, such as *Amphipholis squamata*, are hermaphroditic. In a few species the female carries a dwarf male clinging to the disk. The ovaries and testes are confined to the disk, and open indirectly to the exterior by way of interradial pouches in the integument, termed genital bursae. In the Ophiocanopidae, however, the gonads are serially paired in the basal arm joints, and do not open into bursae. Those species producing ophiopluteus larvae are apparently fewer than those with direct development. In viviparous species the young are retained in the bursae. Regeneration of lost organs is widespread. Autotomy (shedding of the arms) is practiced by most species when interfered with; the disk alone regenerates the lost members. The Amphiuridae can also regenerate the gut and gonads, which may be cast off in autotomy. Species of Ophiactidae regularly reproduce by transverse fission, so that the arms (usually six in such forms) occur in two sets, of three large arms and three small arms. In no case have discarded arms been observed to regenerate. *See* ECHINODERMATA; REGENERATION (BIOLOGY).

Relation to humans. No ophiuroids are used as food by humans, and none are venomous. Ophiuroids must have a considerable indirect economic importance in view of their immense numbers and consequent significance in natural food chains involving commercially sought species.

Ecology. Ophiuroids occur in all the oceans from low-tide level downward, often in dense populations which number millions to the hectare. Six families range below a depth of 2 mi (3.2 km); the genera *Ophiura*, *Amphiophiura*, and *Ophiacantha* range below 4 mi (6.4 km). The shallow-water forms hide among algae, under stones, within sponges, or bury the disk in sand or mud, leaving only the arms protruding. Deep-water forms lie in or on the bottom material or adhere to corals or cidarids (urchins).

Among their numerous parasites are Protozoa in the stomach or genital organs; nematodes, trematodes, and Crustacea; Myzostomida (polychaete annelids) sometimes occur. Parasitic algae, such as *Coccomyxa ophiurae*, infest the spines and cause malformations; parasitic mollusks are much rarer than is the case with starfishes and sea urchins.

Howard B. Fell

Bibliography. R. A. Boolootian (ed.), *The Physiology of Echinodermata*, 1966; H. B. Fell, Phylogeny of sea stars, *Phil. Trans. Roy. Soc. London*, ser. B, 246:381–485, 1963; T. Lyman, Report on the Ophiuroidea, in *Report of the Scientific Results of HMS Challenger*, vol. 5: *Zoology*, 1882; T. Mortensen, Studies of Indo-Pacific euryalids, *Videnskabelige Meddelelser fra Dansk Naturhistorisk Forening*, 96: 1-75, 1933; J. Nebelsick and T. Heinzeller, *Echinoderms* Taylor & Francis Group, 2004; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; V. B. Pearse, et al., An accessible population of *Ophiocanops* off NE Sulawesi, Indonesia, pp. 413-418 in R. Mooi and M. Telford (eds.), *Echinoderms: San Francisco*, A. A. Balkema, Rotterdam, 1998; A. B. Smith, A. B., G. L. J. Patterson, and B. Lafay, Ophiuroid phylogeny and higher taxonomy: Morphological, molecular and palaeontological perspectives, *Zool. J. Linn. Soc.*,114: 213-243, 1995; A. B. Smith, et al., From bilateral symmetry to pentaradiality: The phylogeny of hemichordates and echinoderms, pp. 365-383, in J. Cracraft and M. J. Donoghue (eds.), *Assembling the Tree of Life*, Oxford University Press, 2004.

Opiates

Drugs derived from opium, the dried juice of the oriental poppy seed. The pharmacologically active substances, which constitute approximately 25% of the extract, are the alkaloids morphine, codeine, and papaverine. The newer synthetic compounds, which resemble morphine in their action, are called opioids.

The principal effect of opium and opioids is to relieve pain. Even today morphine remains the best analgesic. It also assuages anxiety and causes slight drowsiness, relaxation, and a euphoric state of mind. These psychic effects are so agreeable that many troubled individuals seek solace by ingesting, smoking, or injecting opiates. Other effects are to suppress cough, allay respiratory distress, and relieve the pain and anxiety of a heart attack. The drugs act on the vital centers in the spinal cord and brainstem (medulla, pons, midbrain, and thalamus). *See* MORPHINE ALKALOIDS; PAIN.

Codeine has an action similar to morphine, but its analgesic effects are less. Papaverine has almost no analgesic action, and is used as an antispasmodic to relieve vascular spasm and undesirable contraction of smooth muscle.

The most tragic effect of the analgesic opium compounds is their addictive properties. Prolonged use leads to physical dependence, and discontinuance leads to painful and unpleasant withdrawal symptoms. Once addicted, most individuals become incapable of functioning in school and work and will go to extreme measures to obtain the drug and avoid the abstinence symptoms. Few can break the drug habit unless they enter an addiction center. *See* ADDICTIVE DISORDERS; ANALGESIC; NARCOTIC.

Raymond D. Adams

Opiliones

The harvestmen or daddy longlegs, an order of the class Arachnida; sometimes known as Phalangida. About 4500 species are known. They are common in temperate and tropical climates. Most are red, brown, or black and 5–20 mm (0.2–0.8 in.) long, although a few are only 1–2 mm (0.04–0.08 in.). Some large tropical species have bright iridescent colors and elaborate spines.

The cephalothorax (prosoma) is broadly attached to the segmented abdomen (opisthosoma). The center of the cephalothoracic shield (carapace) bears a tubercle with a simple eye on each side. Many have scent glands opening on the sides which produce repellent fluids with strong odors, possibly phenols or quinones.

The six pairs of appendages include relatively small chelate chelicerae (jaws), leglike (usually) pedipalps, and four pairs of legs. The legs may be very long and slender, the distal ends being able to wrap around plant stalks. There is a genital opening between the fourth coxae.

Harvestmen are predators of small invertebrates or may scavange or eat decaying vegetation. Respiration is by means of tracheae (thin tubes), which open through spiracles on the abdomen; there may be additional spiracles on the legs.

When mating, the male faces the female without courtship and projects a tubular penis between the female's jaws and into the female gonopore. The female deposits eggs into soft soil or leaf litter or under stones through an ovipositor. On hatching, the young resemble adults but are less sclerotized. After a series of molts they become sexually mature. Northern species produce eggs in fall, tropical species toward the end of the rainy season. Opilionids live about a year. Some species may occasionally form aggregations of hundreds of individuals with legs intertwined; the reason for such behavior is not known.

The group is divided into three suborders: the mitelike Cyphophthalmi; the tropical Laniatores with strong pedipalps, and some species adapted for cave life; and the long-legged Palpatores, the most common opilionids in temperate areas. *See* ARACHNIDA.

H. W. Levi

Opisthobranchia

A subclass in the class Gastropoda containing about 4000 living species, arranged in nine orders, including the herbivorous Aplysiomorpha (sea hares) and Sacoglossa and the carnivorous Thecosomata (sea butterflies); Nudibranchia (sea slugs); Cephalaspidea (or Bullomorpha); Gymnosomata; Pleurobranchomorpha (or Notaspidea); Acochlidiacea; and Runciniodea. Primitive members of many of the orders show adaptations for burrowing beneath sand or mud; more advanced members are always active surface-living or pelagic forms. This trend is accompanied by a decrease in the importance of the shell and operculum for passive defense. These are replaced by more dynamic chemical (some species secrete decinormal sulfuric acid through the skin if annoyed), physical (daggerlike calcareous epidermal spicules), or biological (redirected nematocysts derived from coelenterate prey) defensive mechanisms.

The effects of gastropod torsion have been progressively nullified, then abolished in the opisthobranch line of descent from prosobranch ancestors. Many intermediate stages in this evolutionary loss of visceral torsion have survived and may be studied today. The most advanced members of the subclass show external and internal bilateral symmetry of all systems except the reproductive apparatus (situated on the right side to facilitate reciprocal copulation in these hermaphrodite animals).

The adult shells of the primitive opisthobranchs living today are often strongly developed and sometimes colorful, as in *Acteon* (**illus.** *a*) where a horny snug-fitting operculum may be present and functional, but a more typical opisthobranch shell (as in *Micromelo*, illus. *b*, from the Caribbean or *Hydatina* from the Indo-Pacific) is fragile, inflated, and egg-shaped, and has the spire more or less concealed; the operculum is absent after the veliger larval phase. In rather more advanced forms, such as *Umbraculum* (illus. *c*), the external shell has a very wide gape, and in animals like *Berthella* the widely gaping shell is wholly internal, covered by the mantle.

The most varied shells are found in the Sacoglossa, which may have a single capacious coiled shell (*Volvatella*), a flattened open shell situated dorsally (*Lobiger*), or two lateral shells (the bivalved gastropods, *Berthelinia*).

In the highest sacoglossans, some bullomorphs and aplysiomorphs, and in all the nudibranchs, the true shell is completely lost after larval metamorphosis.

Opisthobranch gastropods vary in size from tiny herbivorous sacoglossans, which live upon and consume delicate marine algae, cell by cell, and the sand-dwelling acochlidiaceans, which are sometimes so minute that they can move unharmed between the grains that make up a sandy beach, to the huge nudibranch *Tochiuna* of the Pacific Northwest of America, and the aplysiomorph *Dolabella* of tropical seas, weighing in air as much as 4 lb (2 kg). Some species may be found only by diving, digging, or searching beneath boulders or coral heads; others, like some of the tropical nudibranchs, appear bold and self-advertising.

The only commercial fishery for any opisthobranch is found in coastal regions of China, where dried bodies of *Aplysia* are employed medicinally by physicians. Handling some living opisthobranchs can lead to mild pain and local inflammation. The planktonic nudibranchs *Glaucus* and *Glaucilla*, which feed upon the venomous siphonophore *Physalia* (the Portuguese man-o'-war), utilize for their own defense the nematocysts of this formidable prey. Severe stings by *Glaucus*, resembling those inflicted by *Physalia*, have been reported in eastern Australia by

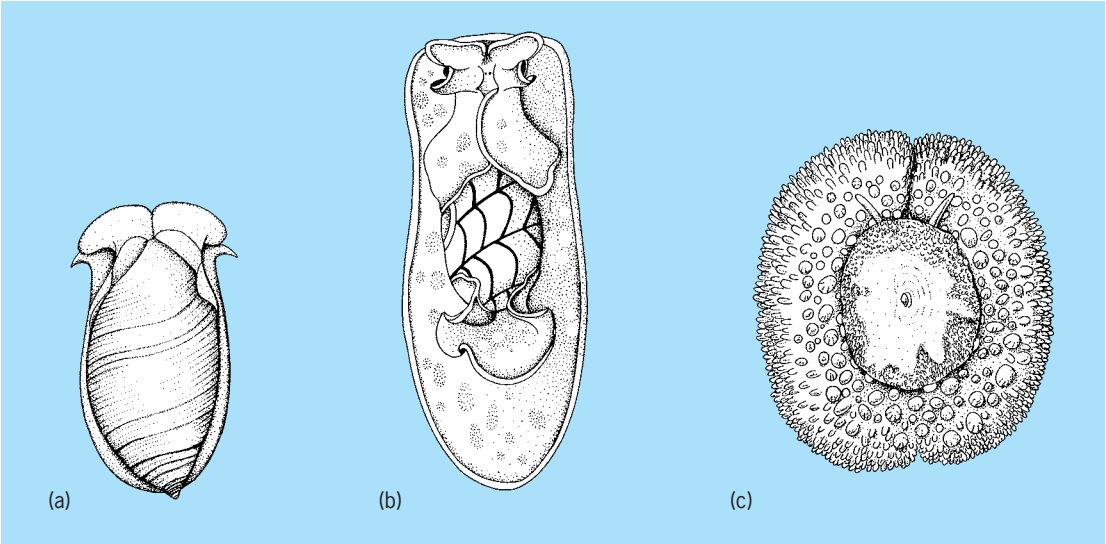

Representative genera of opisthobranchs. (*a*) *Acteon*. (*b*) *Micromelo*. (*c*) *Umbraculum*.

bathers who misguidedly handled these pretty blue mollusks.

Swimming occurs in many opisthobranchs, sometimes brought about by whole-body undulations, sometimes by vigorous movements of expanded lobes of the foot or mantle. The orders Gymnosomata and Thecosomata contain numerous plankton-feeding species which spend their entire lives swimming near the surface of the oceans. *See* GASTROPODA; MOLLUSCA; NUDIBRANCHIA; SACOGLOSSA.

T. E. Thompson

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; C. Lydeard, D. R. Lindberg, and G. J. Vermeij, *Molecular Systematics and Phylogeography of Mollusks*, Smithsonian Books, 2003; T. E. Thompson, *Biology of Opisthobranch Molluscs*, II, 1984; T. E. Thompson, *Nudibranchs*, 1976.

Opisthocomiformes

A small order of birds that contains a single family, the Opisthocomidae. Its one species, the hoatzin, is restricted to South America and has been frequently included in the Galliformes as a suborder; however, little evidence supports this conclusion. More recently, some researchers claimed that it was a typical member of the Cuculidae, the cuckoos, based only on evidence from deoxyribonucleic acid (DNA) hybridization studies. However, important differences in the morphology of the perching foot argue strongly against the inclusion of the hoatzin in Cuculidae. The hoatzin has an anisodactyl perching foot with three anterior toes and a well-developed hallux (first toe) that is completely different from the zygodactyl perching foot of the cuckoos (with the second and third toes anterior, and the fourth toe and hallux posterior). Selective demands for perching could not have resulted in the evolution of the anisodactyl foot of the hoatzin from the zygodactyl foot of the cuckoos. In the absence of definite evidence of its relationship to other avian orders, the Opisthocomiformes are best considered as a distinct order that evolved in South America since its isolation from the rest of Gondwanaland sometime in the late Mesozoic. *See* CUCULIFORMES, GALLIFORMES.

The hoatzin is a unique bird, often considered rather reptilian because of its nearly unfeathered young, which leaves the nest soon after birth to crawl around trees by using its clawed wings as well as its feet (see **illustration**). Hoatzins are adept swimmers, capable of surviving a fall into water. They are medium-sized birds of dark brown plumage streaked with white; a distinct crest of reddish-brown feathers tops the head, which is small and has a short stout

A baby hoatzin climbing a tree using its clawed wings as well as its feet. (*Reprinted with permission from D. W. Linzey, Vertebrate Biology, McGraw-Hill, 2001*)

bill. The wings are long and rounded, the legs are short with strong toes, and the long hallux is on the same level as the anterior toes. Hoatzins are weak fliers, but they are almost strictly arboreal, clambering about in trees to feed on leaves and fruit. A large crop, which displaces the sternum in a posterior direction, is used to store food prior to digestion. The nest is placed in trees overhanging water, and both sexes incubate and care for the young. Hoatzins are found in forests of northern South America, especially along rivers and streams.

The hoatzins are known in the fossil record only by the Miocene *Hoazinoides* from Columbia, a typical form which reveals nothing about the evolutionary history of the order. *See* AVES. Walter J. Bock

Bibliography. B. T. Thomas, Order Opisthocomiformes, pp. 23–32, in J. Del Hoyo et al. (eds.), *Handbook of the Birds of the World*, Lynx Editions, vol. 3, 1996.

Opossum

A New World marsupial belonging to the family Didelphidae. Didelphidae consists of one North American species, the virginia opossum (*Didelphis virginiana*), and 65 Central and South American species, including the mouse opossums (*Marmosa*, *Gracilinanus*, *Marmosops*, *Micoureus*, *Thylamys*, and *Lestodelphys*); short-tailed opossums (*Monodelphis*); gray and black four-eyed opossums (*Philander*); brown four-eyed opossums (*Metachirus*); water opossums (*Chironectes*); woolly opossums (*Caluromys*); black-shouldered opossums (*Caluromysiops*); bushy-tailed opossums (*Glironia*); and thick-tailed opossums (*Lutreolina*). *See* MARSUPIALIA.

Morphology. Opossums are primarily nocturnal and may be either arboreal or terrestrial.

Each front foot has five claw-bearing toes, while each hind foot has four claw-bearing toes and a thumblike, clawless hallux, or big toe, which opposes the other digits. The water opossum, or yapok (*Chironectes*), of Central and South America is semiaquatic with webbed hind feet. The snout is long and pointed, and the tail is usually prehensile (used for grasping). Body size (excluding tail length) ranges from 75 mm (3 in.) in tiny species (*Monodelphis*) to more than 500 mm (20 in.) in the largest species (*Didelphis*). The abdominal pouch (marsupium), well developed in some genera but rudimentary in others, is lined with soft fur and contains the mammae. The skull has a large preorbital region, large zygomatic arches, a prominent median sagittal crest, and a small braincase. Opossums have heterodont dentition consisting of 50 teeth (dental formula: I 5/4, C 1/1, Pm 3/3, M 4/4 × 2). Females have a well-developed marsupium. The usual number of teats is 13 (12 in a circle and one in the center), although all may not be functional. The male reproductive organ is forked, and the female reproductive system consists of paired vaginae and paired uteri. *See* DENTITION; REPRODUCTIVE SYSTEM.

Virginia opossum (*Didelphis virginiana*).

Virginia opossum. The Virginia opossum D. virginiana; (see **illustration**) ranges from southern Canada to Costa Rica. It is a medium-sized mammal with long, rather coarse, grayish-white fur. Individuals have large, thin, leathery, naked ears and a long, scaly, sparsely haired tail. The nose is pink, the eyes are black, and the ears are bluish-black. Adult Virginia opossums are approximately 600–875 mm (23–34 in.) in length, including a 275–350 mm (117–14 in.) tail. Adults usually weigh between 1.3 and 4.5 kg (4 and 13 lb).

Diet and habitat. Opossums are omnivorous, eating a wide variety of foods but preferring animal matter during all seasons. They are solitary mammals and inhabit a variety of habitats from grasslands to forests. Home ranges vary considerably and overlap broadly. There is no evidence of territoriality. Opossums are active throughout the year but may retreat to dens during cold periods. Den sites include hollow trees, fallen logs, brush piles and ground burrows. The nest is composed almost entirely of leaves. Opossums readily climb, using their prehensile tail extensively for balance and support. Although most forms are not aquatic, they can swim and dive readily. These mammals are probably best known for "playing possum," a temporary state of catatonia. This phenomenon, which is a passive defensive tactic that apparently has survival value, seems to be a nonvoluntary (reflex) action and probably is not deliberate or willful.

Mating and development. Mating is promiscuous or polygynous with females being seasonally polyestrous. Two litters are usually produced each year. The gestation of 12 to 13 days is the shortest of any

American mammal. Average litter size ranges from 7 to 9, although as many as 18 young have been recorded. A dozen newborn opossums, each weighing approximately 0.16 g, will fit into a teaspoon. At birth, the front limbs are well developed and are used in a swimming motion to help the young ascend through the mother's fur into the marsupium. The young attach to a teat and remain in the marsupium for approximately 60–80 days. Weaning usually occurs between 80 and 100 days. Opossums reach sexual maturity between 6 and 8 months of age, but females do not breed until the first estrus the following year. Average life expectancy is approximately 1.33 years, although captive animals have lived over 7 years. Major predators include great horned owls, carnivorous mammals including dogs, and humans. Automobiles are responsible for the death of many opossums. Donald W. Linzey

Bibliography. G. A. Feldhamer, B. C. Thompson, and J. A. Chapman (eds.), *Wild Mammals of North America: Biology, Management, and Conservation*, 2d ed., Johns Hopkins University Press, 2003; D. W. Linzey, *The Mammals of Virginia*, McDonald & Woodward Publishing Company, 1998; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford Limited, 2001; R. M. Nowak, *Walker's Mammals of the World*, 2 vols., Johns Hopkins University Press, 1999; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Opportunistic infections

Infections that cause a disease and occur only when the host's ability to fight back (the immune system) is impaired. Some microorganisms can be both opportunistic and nonopportunistic. An example is the bacterium *Mycobacterium tuberculosis*. A healthy individual who is exposed to a sufficient number of *M. tuberculosis* bacteria can develop pulmonary tuberculosis and become quite ill. However, most people with normal immune systems who are exposed to tuberculosis never become clinically ill even though they become infected. In this case, the body's immune defenses are capable of isolating the *M. tuberculosis* and preventing it from spreading throughout the body and causing symptoms or clinical disease. However, if an individual with tuberculosis also has an immune deficiency, such as acquired immunodeficiency syndrome (AIDS), the disease process is much different. Individuals with AIDS who are exposed to an infected person with active tuberculosis are much more likely to become ill, have more widespread disease, infect others, and die from the disease. Fortunately, if the diagnosis is made in time, antituberculous medications work effectively regardless of the underlying condition. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); TUBERCULOSIS.

The classic opportunistic infection actually leads to disease only in individuals with abnormal immunity and never in the normal host. The protozoon *Pneumocystis carinii* infects nearly everyone at some point in life but never causes disease unless the immune system is severely depressed. The most common immunologic defect associated with pneumocystosis is AIDS. Other immunologic deficiencies such as hematologic cancers and profound malnutrition can also predispose the individual to pneumocystosis, but only rarely when compared to AIDS, in which it is the most common life-threatening opportunistic infection.

Compromised host defenses. A compromised host is an individual with an abnormality or defect in any of the host defense mechanisms that predisposes that person to an infection. A variety of specific host defenses can become altered and may lead to a relatively predictable set of opportunistic infections. The altered defense mechanisms or immunity can be either congenital, that is, occurring at birth and genetically determined, or acquired. Congenital immune deficiencies are relatively rare.

Acquired immunodeficiencies are associated with a wide variety of conditions such as (1) the concomitant presence of certain underlying diseases such as cancer, diabetes, cystic fibrosis, sickle cell anemia, chronic obstructive lung disease, severe burns, and cirrhosis of the liver; (2) side effects of certain medical therapies and drugs such as corticosteroids, prolonged antibiotic usage, anticancer agents, alcohol, and nonprescribed recreational drugs; (3) infection with immunity-destroying microorganisms such as the human immunodeficiency virus that leads to AIDS; (4) age, both old (the older one gets, the more the immune system wanes) and young (the younger one is, particularly under 1 month of age, the more immunologically immature); and (5) foreign-body exposure, such as occurs in individuals with prosthetic heart valves, intravenous catheters, and other indwelling prosthetic devices.

Granulocytopenia. Granulocytes are white blood cells that act as a first line of defense against potential pathogens (see **table**). When their numbers are sufficiently depleted, usually as a result of an underlying hematologic cancer, cancer chemotherapy, or alcohol use, infection with aerobic bacteria and fungi that are already present on or in the body becomes more likely. The lack of granulocytes diminishes the overall immune response, making the ability to diagnose an infectious disease even more difficult.

Cellular immune deficiency. Cellular immunity refers to the immune defect that results in the inability of monocytes and macrophages, which are essentially scavenger cells, to kill pathogens that are capable of infecting individual host cells. This type of infection includes such diseases as legionnaires' disease, tuberculosis, pneumocystosis, and cryptosporidiosis. These infections rarely cause disease in the normal host. Conditions associated with impaired cellular immunity include AIDS, Hodgkin's disease, chronic corticosteroid use, irradiation therapy, and collagen vascular diseases such as rheumatoid arthritis and lupus. *See* CELLULAR IMMUNOLOGY.

Humoral immune dysfunction. Humoral immunity refers to the immune defect that results from impaired

Opportunistic infections and immunologic deficiencies

Immunologic deficiency	Associated underlying disease or conditions	Likely opportunistic pathogen	Associated clinical infectious diseases
Granulocytopenia	Acute leukemia Chemotherapy Drugs	*Escherichia coli* *Pseudomonas aeruginosa* *Candida albicans* *Aspergillus fumigatus*	Bacteremia Pneumonia Sinusitis
Cell-mediated deficiency	AIDS Hodgkin's disease Solid tumors Sarcoidosis Aging Malnutrition Lupus	*Pneumocystis carinii* *Toxoplasma gondii* *Cryptococcus neoformans* *Mycobacterium tuberculosis* *Listeria monocytogenes* *Legionella pneumophila* Cytomegalovirus	Pneumonia Encephalitis Meningitis Widely disseminated infection
Humoral immune deficiency	Congenital splenectomy Chronic lymphocytic leukemia Multiple myeloma Sickle cell anemia	*Streptococcus pneumoniae* *Haemophilus influenzae*	Bacteremia Pneumonia Meningitis
Foreign body	Mechanical heart valve Artificial joint Intravenous catheter	*Staphylococcus aureus* *Staphylococcus epidermidis*	Bacteremia Phlebitis Endocarditis
Tumor	Obstructed bronchus	Various anaerobic and aerobic bacteria	Postobstructive pneumonia Enteritis Urinary tract infection

antibody production. Antibodies are important in combining with certain antigens, in this case pathogens, so that the antigen-antibody complex is taken up by host phagocytes and neutralized. The type of infections associated with this deficiency include those by encapsulated bacteria, such as the aerobes *Streptococcus pneumoniae* and *Haemophilus influenzae*. Conditions associated with humoral immune dysfunction include multiple myeloma, congenital deficiencies, chronic lymphocytic leukemia, and splenectomy regardless of the cause. *See* HAEMOPHILUS; STREPTOCOCCUS.

Foreign-body exposure. The presence of any mechanical apparatus, intravascular needle, or prosthetic device such as a joint or heart valve is a risk for infection by organisms that can seed or enter the tissues through the device. The most common organism to do this is a *Staphylococcus* species. It is through this route that many hospital-acquired infections occur. *See* STAPHYLOCOCCUS.

Tumors and other medical conditions. Depending on the nature of the specific malignancy, any number of the host's defense mechanisms may become compromised. In addition, the tumor itself may obstruct or physically invade vital organs, causing yet another form of impaired host defense. An example is the effect of a bronchogenic carcinoma, a type of lung cancer. If the tumor itself obstructs one of the major airways in the lung, a postobstructive pneumonia may result. The usual pathogen is an aerobic or anaerobic bacterium that had formerly colonized the respiratory tract but now is actually causing a pneumonia. Besides tumors, any condition that adversely effects organ function can lead to the development of infection. Examples include asthma, strokes, emphysema, and cystic fibrosis. *See* TUMOR.

Opportunistic pathogens. Virtually any microorganism can become an opportunist. The typical ones fall into a number of categories and may be more likely to be associated with a specific immunologic defect. Examples include (1) gram-positive bacteria: both *Staphylococcus aureus* and the coagulase-negative *S. epidermidis* have a propensity for invading the skin and as well as catheters and other foreign implanted devices; (2) gram-negative bacteria: the most common is *Escherichia coli* and the most lethal is *Pseudomonas aeruginosa*; these pathogens are more likely to occur in cases of granulocytopenia; (3) acid-fast bacteria: *M. tuberculosis* is more likely to reactivate in the elderly and in those individuals with underlying malignancies and AIDS; (4) protozoa: defects in cell-mediated immunity, such as AIDS, are associated with reactivated infection with *Toxoplasma gondii* and *Cryptosporidium*; in individuals with normal immune defenses, these infections are relatively benign; with AIDS, they are sometimes lethal; (5) fungi: *Cryptococcus neoformans* is a fungus that causes meningitis in individuals with impaired cell-mediated immunity such as AIDS, cancer, and diabetes; *Candida albicans* typically causes blood and organ infection in individuals with granulocytopenia. *See* ESCHERICHIA; FUNGI; MEDICAL MYCOLOGY.

Diagnosis and treatment. The first step in treatment of opportunistic infections involves making the correct diagnosis, which is often difficult as many of the pathogens can mistakenly be thought of as benign. The second step involves administration of appropriate antimicrobial agents. As a third step, the underlying immune defect needs to be corrected. In the case of granulocytopenia, physicians can administer cytokines such as granulocyte colony stimulating factor that actually boosts the white blood cell count considerably. In some cases, indwelling catheters and devices must be removed or replaced. Unfortunately, the biggest limitation lies in the fact that the overall

ability to significantly improve immune function is quite limited. *See* IMMUNOLOGICAL DEFICIENCY; INFECTION; MEDICAL BACTERIOLOGY. Robert Murphy

Bibliography. M. A. Sande and P. A. Volberding (eds.), *The Medical Management of AIDS*, 6th ed., 1999; S. T. Shulman, J. P. Phair, and H. M. Sommers (eds.), *The Biologic and Clinical Basis of Infectious Diseases*, 1992.

Opsonin

A term used in serology and immunology to refer to a substance that enhances the phagocytosis of bacteria by leukocytes. As originally proposed by A. E. Wright and E. R. Douglas, the term denoted a thermolabile, relatively nonspecific substance present in normal sera. In modern usages, opsonin is more generally synonymous with the bacteriotropin of F. Neufeld and coworkers, a relatively thermostable antibody, increased in amount during specific immunization, that renders the corresponding bacterium more susceptible to phagocytosis. There is evidence that this action can be promoted to some extent by antibody alone, but that it is substantially increased by the further addition of the thermolabile complement system. Opsonic activity can be displayed by antibodies that also give precipitation, agglutination, lytic, and neutralization reactions.

The opsonic index is a measure of the opsonic activity of sera, found by dividing the average number of bacteria per phagocytic cell, as determined in the presence of an immune serum, by the corresponding value obtained in the presence of normal serum. *See* AGGLUTINATION REACTION; ANTIBODY; IMMUNOLOGY; LYTIC REACTION; NEUTRALIZATION REACTION (IMMUNOLOGY); PHAGOCYTOSIS; PRECIPITIN TEST; SEROLOGY. Henry P. Treffers

Bibliography. J. Kuby, *Immunology*, 4th ed., 2000; J. D. Sleigh and M. C. Timbury, *Medical Bacteriology*, 3d ed., 1990.

Optical activity

The effect of asymmetric compounds on polarized light. To exhibit this effect, a molecule must be nonsuperimposable on its mirror image, that is, must be related to its mirror image as the right hand is to the left hand. An optically active compound and its mirror image are called enantiomers or optical isomers. Enantiomers differ only in their geometric arrangements; they have identical chemical and physical properties. The right-handed and left-handed forms of a molecule can be distinguished only by their optical activity or by their interactions with other asymmetric molecules. Optical activity can be used to probe other aspects of molecular geometry, as well as to identify which enantiomer is present and its purity.

As an example of optical isomers, consider tartaric acid (**Fig. 1**), whichwas one of the first synthetic

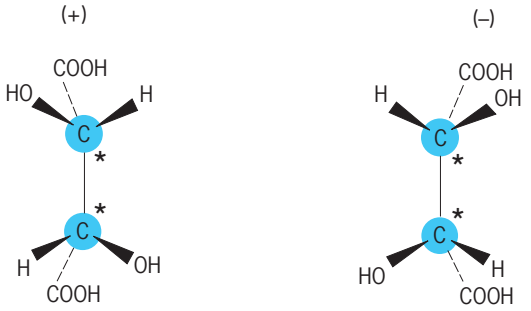

Fig. 1. Enantiomers of tartaric acid.

molecules to be separated into its enantiomers. In this case the asymmetry of each isomer is magnified when trillions of molecules form a crystal; two types of asymmetric crystals are formed.

The physical basis of optical activity is the differential interaction of asymmetric substances with left versus right circularly polarized light. If solids and substances in strong magnetic fields are excluded, optical activity is an intrinsic property of the molecular structure and is one of the best methods of obtaining structural information from a sample in which the molecules are randomly oriented. The relationship between optical activity and molecular structure results from the interaction of polarized light with electrons in the molecule. Thus the molecular groups that contribute most directly to optical activity are those that have mobile electrons which can interact with light. Such groups are called chromophores, since their absorption of light is responsible for the color of objects. For example, the chlorophyll chromophore makes plants green. *See* FARADAY EFFECT; MAGNETOOPTICS; POLARIZED LIGHT.

Methods of measurement. Optical activity is measured by two methods, optical rotation and circular dichroism.

Optical rotation. This method depends on the different velocities of left and right circularly polarized light beams in the sample. The velocities are not measured directly, but both beams are passed through the sample simultaneously. This is equivalent to using plane-polarized light. The differing velocities of the left and right circularly polarized components yield a rotation of the plane of polarization. A polarimeter for observing optical rotation consists of a light source, a fixed polarizer, a sample compartment, and a rotatable polarizer. A cell containing solvent is placed between the polarizers, and one of them is adjusted to be perpendicular to the other, excluding the passage of light. The solvent in the cell is then replaced by a solution of the sample, and the polarizer is rotated to again exclude passage of light. The optical rotation a is the number of degrees the polarizer was rotated. A positive or negative sign indicates the direction of rotation. Enantiomers have rotations of equal magnitude, but opposite signs. The optical rotation depends on the substance, solvent, concentration, cell path length, wavelength of the light, and temperature. Standardized specific

rotations $[\alpha]$ are reported as defined in Eq. (1), where

$$[\alpha]_\lambda^T = \frac{a}{cl} \tag{1}$$

T is the temperature (°C), λ the wavelength (often the orange sodium D line), l the cell path length in decimeters, and c the concentration in grams per milliliter. Alternatively, M_ϕ is defined by normalizing to the rotation for a 1-molar solution, Eq. (2), where

$$M_\phi = \frac{[\alpha]_\lambda^T \mathrm{MW}}{100} \tag{2}$$

M_ϕ is the molar rotation and MW the molecular weight. For polymers, the mean residue rotation, m_ϕ, may be defined by the right side of Eq. (2) by using the mean residue (monomer unit) weight for MW. The variation of optical rotation with wavelength is known as optical rotatory dispersion (ORD).

Circular dichroism. Circular dichroism (CD) is the difference in absorption of left and right circularly polarized light. Since this difference is about a millionth of the absorption of either polarization, special techniques are needed to determine it accurately. Circular dichroism spectrometers consist of a light source, a monochromator to select a single wavelength, a modulator to produce circularly polarized light, a sample compartment, a phototube to detect transmitted light, and associated electronic components. The modulator rapidly switches (typically 50,000 times per second) between left and right circular polarization of the light beam. The absorption of an optically inactive sample is independent of polarization, so that the light intensity at the phototube is constant; thus a constant direct current is generated. The absorption of an optically active sample depends on the polarization, so that the light intensity at the phototube varies at the frequency of the modulator; thus an alternating current is generated. The circular dichroism is proportional to the amplitude of the alternating current. The proportionality constant is determined through calibration by using a compound of known circular dichroism.

Circular dichroism is reported as a difference in absorption, Eq. (3), or as an ellipticity (a measure

$$\Delta\epsilon = \epsilon_L - \epsilon_R = \frac{A_L - A_R}{c'l'} \tag{3}$$

of the elliptical polarization of the emergent beam), Eq. (4), for a 1-molar solution, where ϵ is the ex-

$$M_\theta = 3300\Delta\epsilon \tag{4}$$

tinction coefficient, A is the absorbance [log (I_0/I)], subscripts L and R indicate left or right circular polarization, c' is the concentration in moles per liter, l' is the path length in centimeters, I_0 and I are the light intensities in the absence and presence of the sample, respectively, and M_ϕ is the molar ellipticity. Either $\Delta\epsilon$ or ellipticity, m_ϕ, may be expressed per residue by making c' the concentration of residues (monomer units). As in the case of optical rotation, enantiomers have circular dichroism spectra of equal magnitude but opposite signs.

Variation with wavelength. Optical rotation and circular dichroism are two manifestations of the same interactions between polarized light and molecules. They are related by a mathematical transformation. An important difference between the two measurements is the way in which they vary with wavelength. Optical rotation extends to wavelengths far from any absorption of light. Thus colorless substances still have significant optical rotation at the sodium line. However, all groups which absorb light (chromophores) contribute at all wavelengths, and it can be difficult to extract the contribution of a single group. On the other hand, circular dichroism is confined to the narrow absorption band of each chromophore. Thus it is easier to determine the contribution of individual chromophores, information vital to structural analysis. *See* COTTON EFFECT; DICHROISM; OPTICAL ROTATORY DISPERSION.

Correlation with molecular structure. In synthesizing enantiomers, chemists focus on an asymmetric center, that is, a locus which imparts asymmetry to the whole molecule. A common asymmetric center is a tetrahedral carbon atom with four different groups attached, such as the carbons marked with asterisks in tartaric acid (Fig. 1). However, in correlating optical activity with molecular structure, the focus is on the three-dimensional arrangement of the chromophores which interact most strongly with light.

As examples, consider the nucleoside adenosine and its dimer (**Fig. 2**). The most mobile electrons are in the aromatic ring system, the chromophore (Fig. 2*a*). The electrons in the sugar ribose are more tightly bound and interact less strongly with visible and ultraviolet light. However, all the asymmetric centers are in the ribose part of the molecule. For adenosine, light interacting with the aromatic chromophore is only weakly influenced by the asymmetric centers in ribose, so that small circular dichroism bands are observed (Fig. 2*c*).

In the covalently linked dimer of adenosine, the observed circular dichroism bands are about 10 times larger than those of the monomer. In the 240- to 300-nm region of the spectra (Fig. 2*c*), two bands are observed for the dimer, but only one for the monomer. This indicates strong interaction of the two aromatic chromophores, and hence their close proximity in the dimer. Analysis of the circular dichroism spectra expected for various arrangements of the two chromophores, as well as other types of experimental data, indicates that the aromatic rings are stacked (Fig. 2*b*). The asymmetric centers in ribose cause the formation of the stacked arrangement shown rather than its mirror image. *See* ASYMMETRIC SYNTHESIS.

Stacking of aromatic rings, as exemplified by the adenosine dimer, is a common feature of nucleic acid polymers (deoxyribonucleic acid and ribonucleic acid) isolated from biological sources. Slight differences in the stacking geometry gives each of these polymers a characteristic circular dichroism spectrum. Alterations in the stacking arrangement caused by some pharmacologically active agents can be

detected through alterations in the circular dichroism spectra. These structural changes may in turn be related to the pharmacological action.

A derivative of the amino acid proline (**Fig. 3**) can be used to illustrate another way in which optical activity depends on molecular structure. In this molecule only the OCN group (amide chromophore) which is in the horizontal plane of the drawing and the hydrogen which is marked $H^{\neq}$ need be considered. By forming the N-H bond, $H^{\neq}$ acquires a charge of about $+^1/_3$ electron. It has been predicted that such a positive charge will perturb the motion of the electrons in the amide chromophore in a manner which will produce a negative circular dichroism band when the charge is above the plane of the amide group and to the right of the oxygen. Only for the arrangement shown is the magnitude of the circular dichroism band expected to be as large as observed. Furthermore, it has been shown that there will be no circular dichroism if $H^{\neq}$ is in either of the two planes shown, and that for $H^{\neq}$ in adjacent quadrants the sign of the circular dichroism band alternates (Fig. 3). For this compound, reflection through the horizontal plane will generate the enantiomer. This would place $H^{\neq}$ in the lower right quadrant and generate a positive circular dichroism band with magnitude equal to that of Fig. 3. *See* STEREOCHEMISTRY.

Vincent Madison

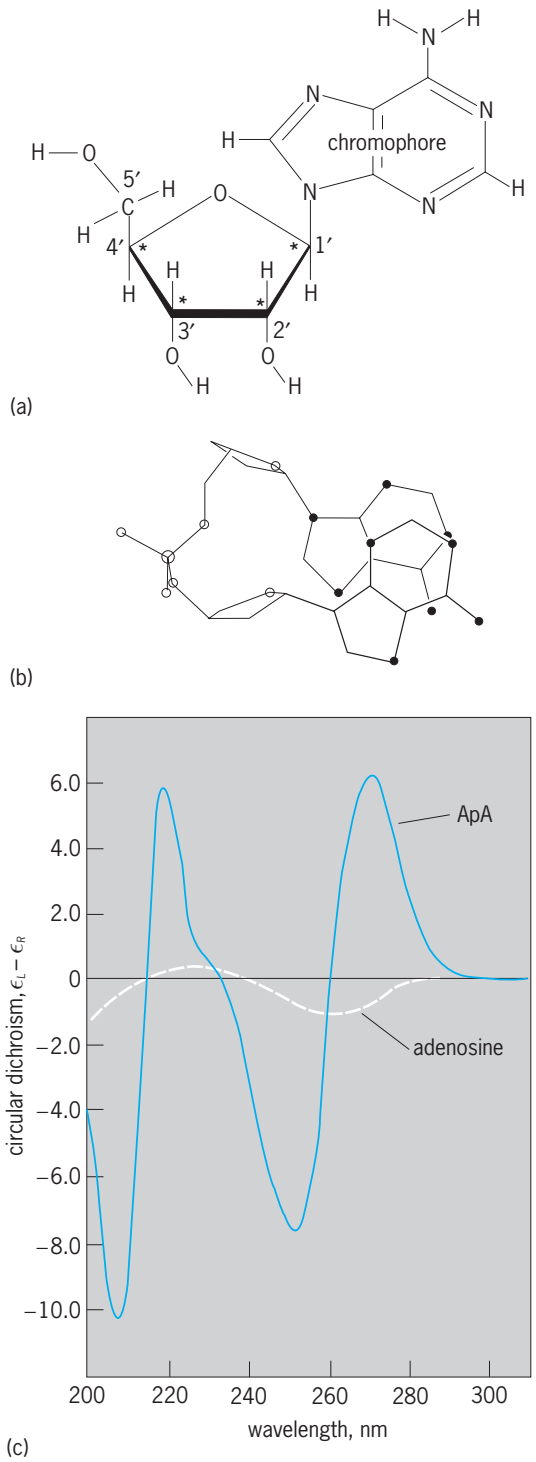

Fig. 2. Adenosine and its dimer. (*a*) Structure of adenosine. Asymmetric centers are marked by an asterisk. (*b*) Stacked arrangement of adenosine dimer (ApA). The 3′ carbon of one adenosine is linked to the 5′ carbon of the other by a phosphate group. (*c*) Circular dichroism spectra of ApA and adenosine at neutral pH in aqueous solution at room temperature.

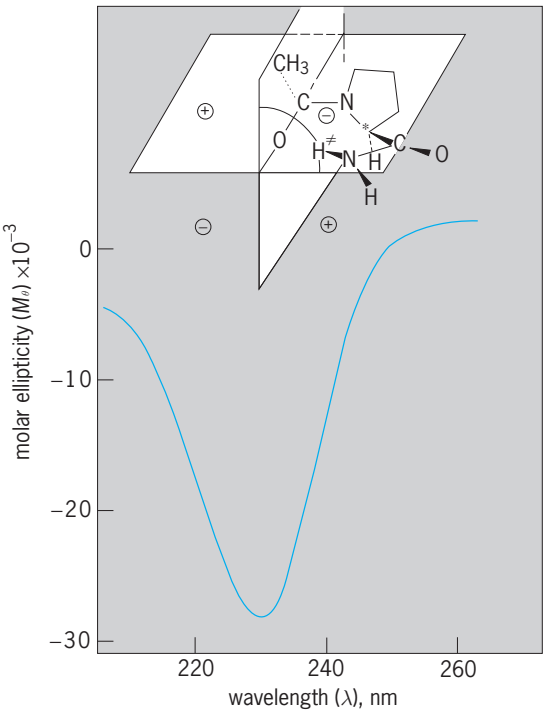

Fig. 3. Folded arrangement of L-proline derivative (*N*-acetyl-L-proline-amide) and its circular dichroism spectrum in *p*-dioxane solution at room temperature.

Bibliography. E. Charney, *The Molecular Basis of Selected Papers on Optical Activity*, 1985; A. Lakhakia (ed.), *Selected Papers on Natural Optical Activity*, 1985; J. A. Schellman, Symmetry rules for optical rotation, *Accounts Chem. Res.*, 1:144–155, 1968; I. Tinoco, Jr., K. H. Sauer, and J. C. Wang, *Physical Chemistry: Principles and Applications in Biological Sciences*, 3d ed., 1995; R. W. Woody, Optical rotatory properties of biopolymers, *J. Polym. Sci. Macromol. Rev.*, 12:181–321, 1977.

Optical bistability

A phenomenon exhibited by certain resonant optical structures whereby it is possible to have two stable steady transmission states for the device, depending upon the history of the input. Such a bistable device may be useful for optical computing elements because of its memory characteristics. The bistability can result from the intrinsic properties of the optical device or from some external feedback such as an electrical voltage supplied by another device. This second type, extrinsic or hybrid optical bistability, is not true optical bistability.

Optical bistability is an inherently steady-state phenomenon, and typically any cycling of the device through its hysteresis cycle must be done adiabatically; that is, changes in the propagating light amplitude, envelope phase, and profile must occur sufficiently slowly that their impact on the evolution of the system may be neglected. This requirement imposes some rather severe frequency-response limitations on the use of intrinsically bistable devices in optical circuits. The two primary types of intrinsic optical bistability, each arising from a distinct physical mechanism, are absorptive bistability and refractive bistability. *See* ADIABATIC PROCESS; HYSTERESIS.

Absorptive bistability. Absorptive optical bistability, discovered in 1969, is based upon coupling the feedback mechanism inherent in an optical cavity with an absorbing nonlinear optical medium in which the absorption coefficient decreases with increasing light intensity (a saturable absorber). The basic theory of operation is as follows: The saturable absorber is placed in the cavity, and the cavity is resonantly pumped. For low light intensities, the transmission coefficient for the cavity is small because of the presence of the highly absorbing medium inside the cavity. As the pump intensity is increased, the absorption of the nonlinear medium decreases. Finally, for some threshold pump intensity, the cavity switches into a high transmission state, because the absorption coefficient is reduced sufficiently that the intrinsic cavity feedback mechanism dominates. The threshold is very sharp because, when the cavity is in a highly transmittive state, the built-up intensity inside the cavity becomes very large compared to the pump intensity (due to the feedback) and effectively bleaches virtually all of the absorption in the nonlinear medium. The intense pump is then largely transmitted, although some energy is stored in the cavity to bleach the absorber. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; LASER; OPTICAL PUMPING.

When the pump wave is reduced in intensity, the built-up intensity inside the cavity does not drop below the point where it effectively saturates the nonlinear absorption until the pump intensity is well below the switch-up threshold. Thus, the cavity stays in its highly transmissive state for intensities well below the threshold. As a result, the transmission characteristic of the cavity exhibits hysteresis and possesses two steady-state outputs for a range of pump intensities.

This device exhibits two characteristics that constrain its usefulness in particular applications. (1) The device is based on an absorption mechanism, so the energy absorbed from the pump light must be dissipated in the bistable element or heat-sinked elsewhere. (2) It is highly frequency-sensitive because its operation is based on the switching characteristics of a resonant cavity.

Refractive bistability. Refractive optical bistability, discovered in 1974, is based on coupling the feedback mechanism inherent in an optical cavity with a nonlinear optical medium that exhibits a change in the refractive index as a function of light intensity. The nonlinear refractive medium is placed inside the optical cavity, and the cavity is pumped slightly off-resonance so that the transmission coefficient is small compared to unity. However, a small amount of light intensity does exist inside the cavity, and changes the effective optical path length inside the cavity by inducing change in the refractive index of the nonlinear medium. As the pump intensity is increased, this change in the effective path length becomes larger, until at some point the cavity switches into, and possibly past, resonance. The transmission coefficient switches abruptly to a value close to unity, and the built-up intensity inside the cavity increases abruptly. If the pump intensity is increased further, it is possible to switch the cavity through a second resonance, with an additional threshold in the transmission coefficient. *See* REFRACTION OF WAVES.

In principle, this process can be repeated for arbitrarily large values of the pump intensity. However, from a practical perspective it is very difficult to obtain more than two or three thresholds because of the intervention of other, higher-order nonlinear processes which arise from the presence of very intense light.

When the pump intensity is reduced, the effective cavity path length returns to its low-intensity value. However, because of the feedback mechanism in the cavity and the presence of the built-up intensity inside the cavity, the switch-down thresholds occur at lower pump intensities than the switch-up thresholds, resulting in hysteresis of the transmission characteristic. In this case, it is possible to observe multiple hysteresis loops, as the cavity is switched through multiple resonances by the built-up intensity within it.

Implementation. The most common implementation scheme for a bistable optical device is the nonlinear Fabry-Perot etalon. The device is typically fabricated from a semiconductor such as indium antimonide (InSb), gallium arsenide (GaAs), or zinc selenide (ZnSe), and consists of a slab of material of approximately 1 micrometer thickness. On each surface of the semiconductor, a highly reflective coating may be deposited to increase the bandwidth of the Fabry-Perot cavity. The choice of a proper nonlinear material is based upon the operating wavelength and the temporal response time desired, and possibly other considerations. Typically for applications in the far-infrared, near-infrared, and visible wavelengths, the proper materials are indium

antimonide, gallium arsenide, and zinc selenide, respectively. Although the response time of all bistable optical devices is limited by the quasi-steady-state nature of the phenomenon, it is also subject to the response time of the nonlinear mechanism that exists in the chosen material. Electronic nonlinear mechanisms such as those found in indium antimonide and gallium arsenide typically have response times in the subnanosecond regime, whereas thermal response times such as those used in zinc selenide interference filter-based bistable devices may be slower. Materials engineering holds the potential to modify the associated response times for various materials through bandgap engineering of devices fabricated by multiple-quantum-well technologies. *See* INTERFEROMETRY; QUANTIZED ELECTRONIC STRUCTURE (QUEST); SEMICONDUCTOR.

Applications. A large number of application schemes for bistable optical devices have been proposed. A major goal of optical bistability technology has been the implementation of optical logic gates. In spite of the fact that most basic logic circuits have been demonstrated experimentally by using optical bistability in some form, the technology has not been competitive with electronic semiconductor technologies. The primary reason is that packing densities for bistable logic circuits are limited by the diffraction of light waves. It is not possible to build circuits by using optical devices significantly smaller than the wavelength of the light being used. In gallium arsenide, for example, theoretically the smallest size of a given device is approximately $0.5\ \mu m \times 0.5\ \mu m$. In contrast, silicon electronic circuits are fabricated with a feature size of $0.1–0.3\ \mu m$. In addition, a certain amount of energy is used by the device in order to cause it to switch from a low-transmission state to a high one, or vice versa. This energy, called the switching energy, is prohibitively large for the large packing densities that are called for in order to compete successfully with electronic technologies. Various computing architecture schemes have been proposed whereby the parallel nature of light propagation may be exploited to perform, in conjunction with bistable optical devices, highly parallel computations that are difficult to do. If these efforts are successful, they will lead the way for the most probable integration of optical computing technology, namely niche computations where some of the inherent properties of light-based circuits and devices may be exploited. *See* INTEGRATED CIRCUITS; INTEGRATED OPTICS; LOGIC CIRCUITS; OPTICAL INFORMATION SYSTEMS.

Active devices. In addition to passive bistable optical devices, it is possible to observe optical bistability in active devices such as laser systems. Although such systems are much more complicated than the passive devices discussed above, several experiments have been performed demonstrating optical bistability, and applications have been proposed. *See* NONLINEAR OPTICAL DEVICES; NONLINEAR OPTICS.

David R. Andersen

Bibliography. H. M. Gibbs, *Optical Bistability; Controlling Light with Light*, 1985; H.-Y. Zhang and K. K. Lee (eds.), *Optical Bistability, Instability, and Optical Computing*, 1988.

Optical coherence tomography

A noninvasive technique for imaging subsurface tissue structure with micrometer-scale resolution. The principles of time gating, optical sectioning, and optical heterodyning are combined to allow cross-sectional imaging. Depths of 1–2 mm (0.04–0.08 in.) can be imaged in turbid tissues such as skin or arteries; greater depths are possible in transparent tissues such as the eye.

Optical coherence tomography complements other imaging modalities commonly used to image subsurface tissue structure, including ultrasound and confocal microscopy. It has a resolution about an order of magnitude better than ultrasound, although the depth of imaging is less. Unlike ultrasound, it does not require a coupling medium between the instrument probe and tissue, facilitating endoscopic applications and imaging of sensitive structures such as the eye. Optical coherence tomography is a type of confocal system, although device parameters are usually set for a lower resolution than is available in commercial confocal microscopes. However, it has a depth of imaging several times that of confocal microscopy. *See* BIOMEDICAL ULTRASONICS; CONFOCAL MICROSCOPY; MEDICAL ULTRASONIC TOMOGRAPHY.

Principles of operation. In a typical optical coherence tomography system (**Fig. 1**), light from a broadband, near-infrared source and a visible aiming beam is combined and coupled into one branch of a fiber-optic Michelson interferometer. Broadband sources include superluminescent diodes, fiber amplifiers, and femtosecond pulse lasers in the wavelength range of 800–1550 nanometers. The light is split into two fibers using a 2×2 coupler, one leading to a reference mirror and the second focused into the tissue. Light reflects off the reference mirror and is recoupled into the fiber leading to the mirror. Concurrently, light is reflected from index-of-refraction mismatches in the tissue and recoupled into the fiber leading to the tissue. Reflections result from changes in the index of refraction within the structure of the tissue, for instance between intercellular fluid and collagen fibers. Light that has been back-reflected from the tissue and light from the reference arm recombine within the 2×2 coupler. *See* INTERFEROMETRY; OPTICAL FIBERS; OPTICAL PULSES.

Because the broadband source has a short coherence length, only light which has traveled very close to the same time (or optical path length) in the reference and tissue arms will interfere constructively and destructively. By changing the length of the reference arm, reflection sites at various depths in the tissue can be sampled. The depth resolution of the optical coherence tomography system is determined by the effectiveness of this time gating and hence is inversely proportional to the bandwidth of the source. An optical detector in the final arm of the Michelson interferometer detects the interference

between the reference and tissue signals. During optical coherence tomography imaging, the reference-arm mirror is scanned at a constant velocity, allowing depth scans (analogous to ultrasound A-scans) to be made. Either the tissue or the interferometer optics is mounted on a stage so that the beam can be scanned laterally across the tissue to build up two- and three-dimensional images, pixel by pixel. *See* COHERENCE.

The optical sectioning capability of optical coherence tomography (its ability to image thin slices of tissue) is determined by the tissue-arm optics. Light from the single-mode optical fiber is essentially a point source, which is focused to a small spot in the tissue. The highly turbid nature of many biological tissues tends to scatter light away from this focus, but the scattered light is not efficiently coupled back into the fiber. Therefore the lateral resolution of the optical coherence tomography system is approximately equal to the focused spot size. Often, the focus-spot diameter and lateral resolution are made approximately 10–20 micrometers to match the depth resolution of the system (using superluminescent diodes) and to provide a relatively long depth of focus. This long working distance is desirable for in vivo systems, since it is difficult to control the exact distance between the instrument optics and the tissue.

Optical coherence tomography has extremely high sensitivity; reflections of less than 10^{-10} of the incident optical power can be detected by using the optical heterodyne technique. Movement of the reference arm mirror induces a modulation of the interferometric signal at the Doppler frequency. In the optical heterodyne detection technique, the envelope of the detector current is recorded by demodulating at this frequency, thus rejecting noise outside the signal bandwidth. The magnitude of the detected signal at the modulation frequency is proportional to the reflectivity of the tissue. *See* DOPPLER EFFECT; HETERODYNE PRINCIPLE.

Color Doppler OCT. Several instruments have been built based on variations of the basic optical coherence tomography system. For instance, polarization-sensitive optical coherence tomography uses polarization-altering optics in the arms of the interferometer to determine the sample birefringence from the magnitude of the back-reflected light. Optical coherence microscopy uses a system of high numerical aperture to achieve resolutions comparable to confocal microscopy but with increased depth of penetration.

Color Doppler optical coherence tomography (CDOCT) is an augmentation capable of simultaneous blood flow mapping and spatially resolved imaging. This technique (also called optical Doppler tomography) makes use of the interferometric phase information ignored in conventional optical coherence tomography. Doppler shifts in light backscattered from moving objects (such as red blood cells) add or subtract from the modulation frequency. If the returned signal is coherently demodulated, a short-time Fourier transform can be performed to find the signal strength at various Doppler-shift frequencies.

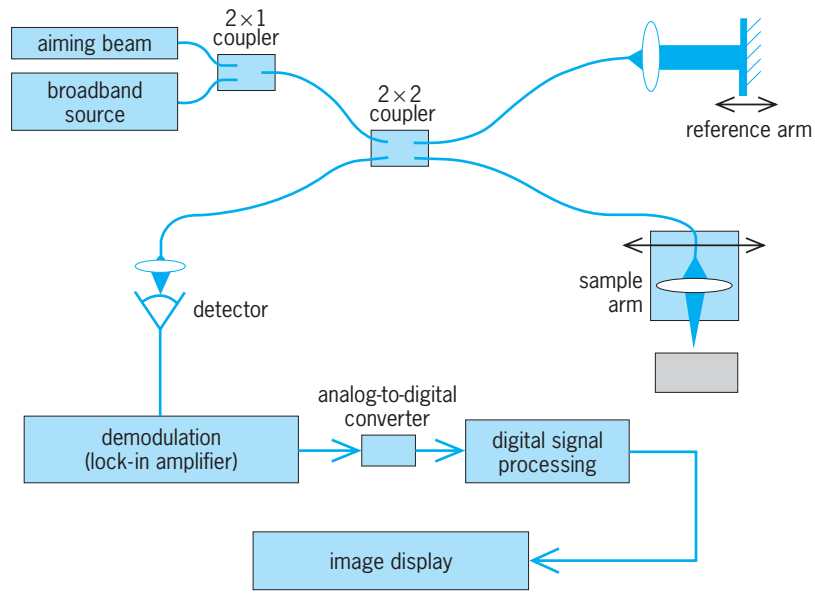

Fig. 1. Typical optical coherence tomography system, based on a broadband source and fiber-optic Michelson interferometer.

The mean velocity of the scatterers, $\overline{V}_{is}$, within each Fourier window can be estimated from the centroid, $\vec{f}_s$, of the localized Doppler frequency spectrum, according to the equation below.

$$\overline{V}_s = \frac{\overline{f}_s c}{2\upsilon_o n_t \cos\theta}$$

Here, c is the speed of light, υ_o is the center frequency of the broadband light source, n_t is the local index of refraction, and $\cos\theta$ is the angle between the instrument probe and the flow direction. The lateral resolution of a Doppler image remains the same as that of the magnitude image obtained in conventional optical coherence tomography; however, the depth resolution is often limited by the Fourier transform window size. A mean velocity can be calculated for each three-dimensional pixel in the image and combined to create a Doppler image separate and complementary to the magnitude image obtained from the amplitude of the signal alone. Alternatively, a threshold can be placed on the signals contributing to the Doppler image in order to eliminate noise, and the Doppler image can be overlaid on the magnitude image.

Applications. Optical coherence tomography can be used to probe the structure of any accessible tissue. Noninvasive studies of the eye and skin are being performed, and a commercial device has been developed for retinal imaging. Using optical coherence tomography, parameters such as eye length can be accurately measured, and the cross-section images of the retina give a clear and quantifiable assessment of retinal separation and macular degeneration, among other pathologies. In skin, the morphology of normal skin layers and components, and disorders such as psoriasis, can be imaged. For example, the expected layers of skin, including epidermis and dermis, fat, muscle, and connective tissue, are seen in a structure

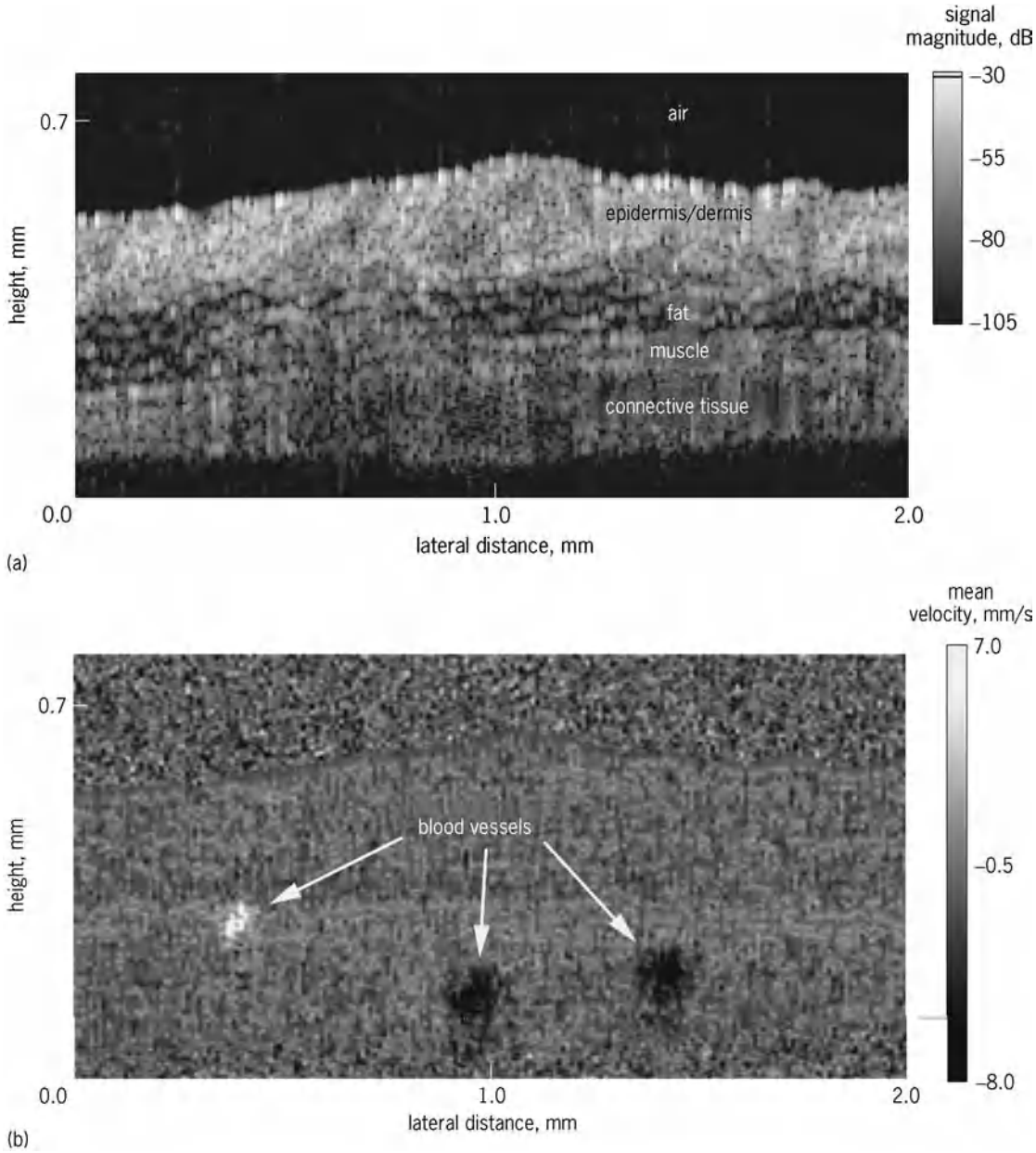

Fig. 2. Optical coherence tomography images of healthy hamster skin in vivo. (*a*) Conventional magnitude image. Various layers of skin can be seen in cross section. (*b*) Result after Doppler processing of the same signal used to create the magnitude image. Three blood vessels are visible because of the Doppler shift caused by their moving red blood cells.

image of healthy hamster skin in vivo (**Fig. 2***a*). Blood vessels located in the connective tissue and fat layers are not seen in this image, because blood and other skin tissues have similar optical properties in the near infrared. When the same signal is processed to distinguish Doppler shifts, however, the moving blood is readily apparent (Fig. 2*b*). *See* EYE (VERTEBRATE); EYE DISORDERS; SKIN; SKIN DISORDERS.

Endoscopes and catheters are used to facilitate imaging of the cardiovascular system and gastrointestinal, respiratory, urinary, and reproductive tracts. Being fiber-based, optical coherence tomography is readily adapted to these applications. Miniature electromechanical and piezoelectric scanning devices have been developed to perform lateral scanning of the interferometer tissue arm optics. In blood vessels, the layered structures of the vessel wall are visible, and distinction can be made between fibrous and calcified plaques. Various mucous membranes of the body have also been successfully imaged. Differences in backscattering properties allow mucosal tissue layers to be diferentiated. Cancerous regions are often marked by a disordering of the layered structure and a more homogeneous appearance of the optical coherence tomography image. Because of this distinction, optical coherence tomography has the potential to guide excisional biopsy and improve the random technique frequently used today. Samples of tissue would be taken from areas which appeared suspect on optical coherence tomography images, for standard histological analysis. *See* FIBER-OPTICS IMAGING.

Potential. Research groups are working on improving the performance of this novel imaging modality. Sources with broader bandwidths and higher powers will increase system resolution and depth of imaging. Signal processing techniques that efficiently extract parameters of interest and display data in an easily comprehended fashion are being developed. Efforts are under way to better understand optical coherence tomography images in terms of the morphology of the tissue under study. Jennifer Kehlet Barton

Bibliography. D. Huang et al., Optical coherence tomography, *Science*, 254:1178–1181, 1991; J. A. Izatt et al., Optical coherence tomography and microscopy in gastrointestinal tissues, *IEEE J. Selected Top. Quantum Electr.*, 2:1017–1028, 1997; X. J. Wang, T. E. Milner, and J. S. Nelson, Fluid flow velocity characterization by optical Doppler tomography, *Opt. Lett.*, 20:1337–1339, 1995.

Optical communications

The transmission of speech, data, video, and other information using light in the frequency range below the infrared limit of the visible spectrum. Optical communications networks now perform this function. An optical communications link consists, at minimum, of a data or information source, a modulated laser transmitter, an optically transparent transmission medium, possibly an amplifier, and a photodetector receiver that recovers the transmitted data (**Fig. 1**). The predominant medium is optically transparent fiber; in this form, the technology is called fiber-optic communications. In another application, optical communications uses free space: free-space optical (FSO) communications.

Light by nature is an electromagnetic wave and it follows the laws of waves, such as refraction, diffraction, interference, and polarization. It is also a particle and exerts pressure; this is particularly noticeable when a satellite enters the Sun's light and starts drifting due to pressure exerted on it. The smallest quantity of light is known as a photon, which, according to atomic physics, is generated when an atom from an excited state makes a transition to either a less excited state or to its nonexcited (ground) state. The frequencies used in optical communications are much higher than those used for radio, television, and microwaves, and lower than those of visible light (**Fig. 2**). Thus, light used in optical communications is at frequencies below the infrared limit of visible light, and therefore it is invisible to the human eye. *See* DIFFRACTION; ELECTROMAGANETIC RADITION; INFRARED RADIATION; INTERFERENCE OF WAVES; LIGHT; PROTON; POLARIZATION OF WAVES; RADIATION PRESSURE; REFRACTION OF WAVES; WAVE MOTION.

The amount of information that a wave can carry depends on its frequency; the higher the frequency, the more the information. Thus, frequencies in the optical spectrum can carry orders-of-magnitude more information than radio and microwave frequencies. For example, the frequency of light is of the order of 3×10^8 MHz as compared with radio and television frequencies that are in the kilohertz-to-megahertz range, and this is what makes optical communications so attractive. Therefore, optical communication is currently the preferred communications technology for very high capacity transmission, and will remain so for many years to come, since it is capable of delivering an unprecedented amount of information per second compared with other communications technologies. *See* BANDWIDTH REQUIREMENTS (COMMUNICATIONS); INFORMATION THEORY.

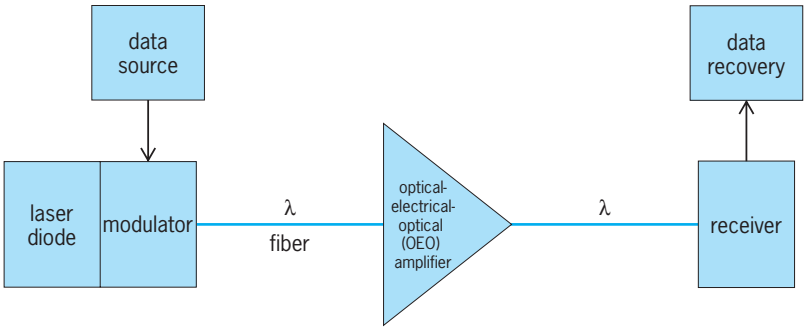

Fig. 1. Block diagram of a simplified optical communications link.

Basic components. For optical communications to be effective, certain basic components are required to construct the communications link. These, in addition to couplers, are a laser transmitter, a modulator, the optical fiber, and a photodetector receiver (Fig. 1).

Laser transmitter. The laser (light amplification by stimulated emission of radiation) used in optical communications is a semiconductor device. When a voltage is applied to it, a strong field is generated within it that excites specific atoms. When excited atoms are stimulated, they emit light at a specific frequency (or wavelength), which is determined by the material, its geometry, and the applied field. Light-emitting diodes (LEDs) are semiconductor devices that emit incoherent light. However, unlike LEDs, a laser radiates stimulated coherent light, that is, light in which all the stimulated emitted wave-photons are in phase. Typical lasers that are used in optical communications generate a continuous optical (yet invisible) beam, and are made with compound materials that consist of the elements indium, gallium, arsenic, and phosphorus. *See* COHERENCE; LASER; LIGHT-EMITTING DIODE.

Modulator. In order to carry information, the continuous optical beam that is generated by the laser must be modulated. That is, the characteristics of the laser light must be altered by impressing on it a stream of digital data or a string of ones and zeros. An optical device that modulates the continuous laser beam is known as a modulator. In optical communications, a modulator may be thought of a fast shutter that acts on the continuous beam. When the shutter opens it allows light to pass through, and when it closes it blocks light; that is, the optical shutter affects the optical power of the beam. This modulation method is known as on-off keying (OOK) and currently is the most used; other modulators affect the frequency or

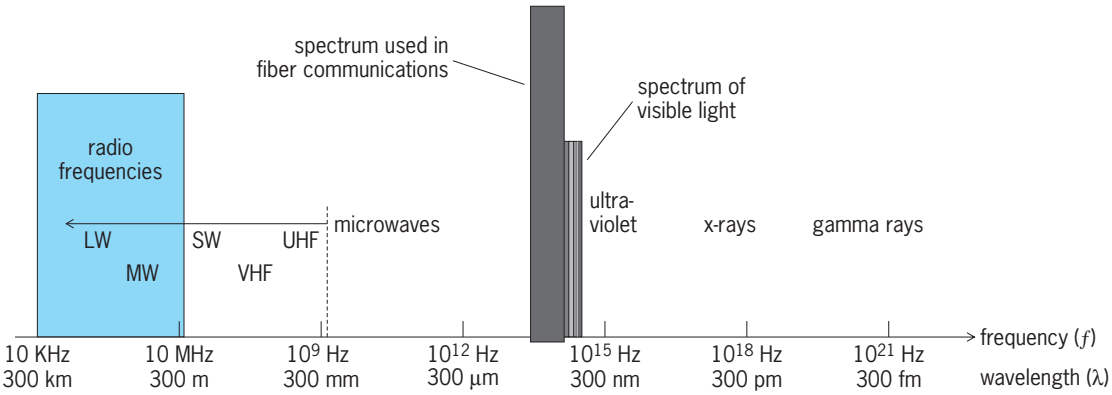

Fig. 2. Optical communications spectrum with respect to the overall electromagnetic spectrum. LW = long-wave, MW = medium-wave, SW = short-wave, VHF = very high frequency, UHF = ultrahigh frequency.

the phase of the laser beam. When laser light passes through the modulator or is blocked for a predetermined period, the period of the optical bit is established. Thus, in 1 second many optical bits may be created by the shutter. Current modulator rates generate up to 4×10^{10} bits per second, or 40 gigabits per second (Gbps). Depending on the technology and modulation rate, the modulator may be external to the laser device or be incorporated internally with it. In low-cost optical communications, it is preferable to incorporate the modulator with the laser device to reduce the number of components and simplify design and manufacturing. *See* MODULATION; MODULATOR.

Optical fiber. The fiber used in optical communications consists of two silica layers, the core and the cladding. What distinguishes the two is the refractive index. The core includes elements, known as dopants, such that its refractive index is higher than the refractive index of the cladding. This refractive index variation helps to keep light within the core in which it is transmitted. The cladding is surrounded by nonsilica layers that are added for strength and for fiber-type identification. Based on geometry, there are two major fiber classifications, the multimode fiber (MMF) and the single-mode fiber (SMF). The multimode fiber has a core diameter of 50 or 62.5 μm surrounded by the cladding layer to a total diameter of 125 μm. The single-mode fiber has a core diameter of only 9 μm surrounded by the cladding layer to a total diameter of 125 μm. In addition to the classification of optical fibers as single-mode and multimode, different types of fiber are produced, such as multimode fiber with a graded refractive index profile in its core, known as GRIN (graded-index), and single-mode fiber with inverted dispersion properties, known as dispersion compensating fiber (DSF). DSF plays a compensating role in the propagation of optical pulses in standard single-mode fiber, since single-mode fiber disperses different frequencies unequally, thus degrading the optical propagation of data at very high rates. *See* OPTICAL FIBERS.

Photodetector receiver. Photodetectors are diode semiconductor devices that generate one or more electron-hole pairs for each photon that enters the biased *p-n* junction. When modulated light impinges on the photodetector device, the generated electron-hole pairs constitute an electric current known as photocurrent. The generated photocurrent emulates the optical digital signal to an electrical digital signal, which is passed through filters and a resistive load to convert it to voltage pulses. The two photodetector types used predominantly in optical communications are, the *p*-intrinsic-*n* (PIN) diode and the avalanche photodiode (APD). *See* OPTICAL DETECTORS; PHOTOVOLTAIC CELL; SEMICONDUCTOR DIODE.

PIN photodiodes consist of a lightly doped (intrinsic) region, which is sandwiched between *p*- and *n*-type semiconductor materials. A reverse-biased PIN photodetector exhibits almost infinite internal impedance with an output current proportional to the input optical power. However, capacitance in the reversed-biased PIN limits its switching speed. In addition, as switching speed increases, parasitic inductance becomes significant and causes shot noise, in addition to the dark noise that is generated due to the temperature of the material. Typically, PIN photodetectors are made with indium gallium arsenide. *See* ELECTRICAL NOISE; MICROWAVE SOLID-STATE DEVICES; PHOTODIODE.

When avalanche photodiodes are reverse-biased, a strong field is developed in the junction region. A photon entering this region generates an electron-hole pair, which, because of the strong field, gains enough energy to generate secondary electron-hole pairs, which generate tertiary pairs, and so on. Thus, this process causes an avalanche, and from a few photons entering the photodetector a substantial photocurrent is generated. However, avalanche photodiodes generate shot noise which is multiplied by the avalanche process as well. Thus, which photodetector type will be used in a particular network application depends on the application requirements that meet the characteristics of the photodetector. Among the key photodetector characteristics, the number of electron-hole pairs produced for each photon determines the efficiency and sensitivity of the photodetector, and this number depends on the material used. The speed of the electron-hole pairs determines how fast a photodetector responds to modulated light. The amount and type of noise

produced determines the expected quality of the signal. Noise is always present and cannot be eliminated or ignored.

The amount of photocurrent due to the optical signal compared with the amount of noise (shot or dark current) generated by the photodetector and associated electronic devices is a measure of the quality of the receiver. This ratio of signal photocurrent to noise, known as the signal-to-noise ratio (SNR), determines whether the original digital signal can be recovered faithfully or not. In practice, the SNR and other factors determine if all the bits that have been received are a faithful reproduction of all the bits transmitted. Calculating or estimating the number of bits in error in 1 billion or trillion bits transmitted for a given data rate provides an important performance parameter in communications, known as the bit error rate (BER). In optical communications, one error bit in 10^{12} bits transmitted is a typical requirement (that is, a BER of 10^{-12}). *See* SIGNAL-TO-NOISE RATIO.

Wavelength-division multiplexing. Initial optical communications networks used a single wavelength at 1310 nm as the carrier frequency, and at modulation rates 2.5 Gbps or less. Soon thereafter, a second wavelength at 1550 nm was added that was suitable for longer fiber lengths. The protocol used was called synchronous optical network (SONET) in the United States, and synchronous digital hierarchy (SDH) in Europe. SONET and SDH, although very similar, differ enough, as a result of specific local needs for services and demographics, to create two different specification standards. As bandwidth requirements increased, another modulation rate was added at 10 Gbps, and more recently another at 40 Gbps.

Despite this incremental and dramatic increase in data rate, bandwidth demand exceeded bandwidth deliverability by a single fiber; this situation is known as capacity exhaust. To meet the ever-increasing bandwidth demand, a need arose for additional fiber, which is very costly. However, at about the same time (in the early 1990s) optical and photonic developments came to fruition that could be applied cost-effectively to optical communications. These developments allowed more than one optical wavelength to be multiplexed in the same fiber, giving birth to a technology called wavelength-division multiplexing (WDM). Providing a multiplicity of wavelengths, or optical channels, in a single fiber, solved the capacity exhaust problem and alleviated the need for additional fibers and networks. *See* MULTIPLEXING AND MULTIPLE ACCESS.

Before WDM technology was deployed, the fiber-optic communications frequency spectrum and the optical channels needed to be defined and standardized for interoperability reasons. The low-loss optical spectrum of silica fiber was subdivided into subbands known as O, E, S, C, and L (see **table; Fig. 3**). Initially, 80 narrow spectral ranges were specified in the C subband and an additional 80 in the L subband, with a separation between center frequencies of adjacent channels, known as channel separation of 50 GHz. The starting reference frequency, from which all other frequencies are calculated, was selected to be a standard frequency at 193.1 terahertz (THz). Since the original recommendation standard, the channel separation has been divided in half to define twice as many optical channels: 160 channels in the C subband and 160 channels in the L. As optical technology evolves and optical communications demand more bandwidth, the number of frequencies may be expanded accordingly.

Subbands of the low-loss optical spectrum of silica fiber

Label	Range, nm	Fiber type*	Applications*
O	1260–1360	SMF	Future DWDM
E	1360–1460	SMF	Future DWDM
S	1460–1530	SMF	Future DWDM
C	1530–1565	SMF	DWDM
L	1565–1625	DSF	DWDM

*SMF = single-mode fiber, DSF = dispersion compensating fiber, DWDM = dense wavelength-division multiplexing.

Figure 3 shows that at wavelengths around 1400 nm the glassy fiber exhibits high attenuation; this is because hydroxyl (O-H) radicals that are trapped in the fiber during the manufacturing process absorb the energy of photons with these wavelengths as they travel in the fiber. However, new manufacturing methods have produced a type of fiber, known as water-free fiber, with low attenuation in this range. If the low-loss spectrum, from below 1300 to over 1600 nm or from O to L, is used, the potential number of channels can be of the order of 1000. If 1000 optical channels are coupled into a single optical fiber, and if each channel is modulated at 10 or 40 Gbps, then the aggregate amount of information transportable in a single fiber is enormous. (For comparison, a digitized uncompressed voice channel reqires 64 kilobits per second, and compressed voice less than 4 kbps.) The technology that makes this possible is known as dense wavelength-division multiplexing (DWDM). *See* HYDROXYL.

Currently, such technology has many manufacturing challenges and is costly; it is used only in high-capacity networks and in long-haul systems. In certain communications applications, a more relaxed and inexpensive optical technology was needed. To address this need, WDM communication standards have defined 18 optical channels with 20-nm

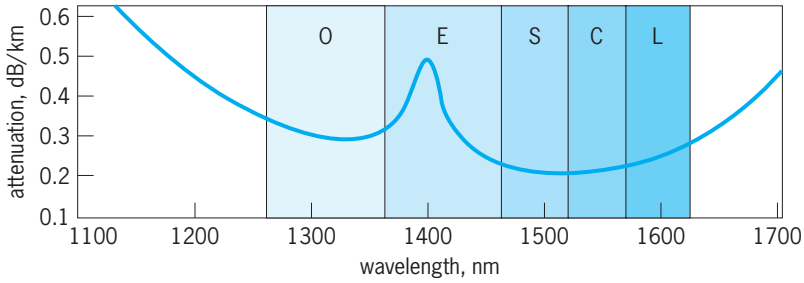

Fig. 3. Low-loss spectrum of silica fiber used in fiber-optic wavelength-division multiplexing communications. The C-band is at a frequency of about 1550 nm (1.55 μm), and the L-band is at about 1600 nm (1.6 μm).

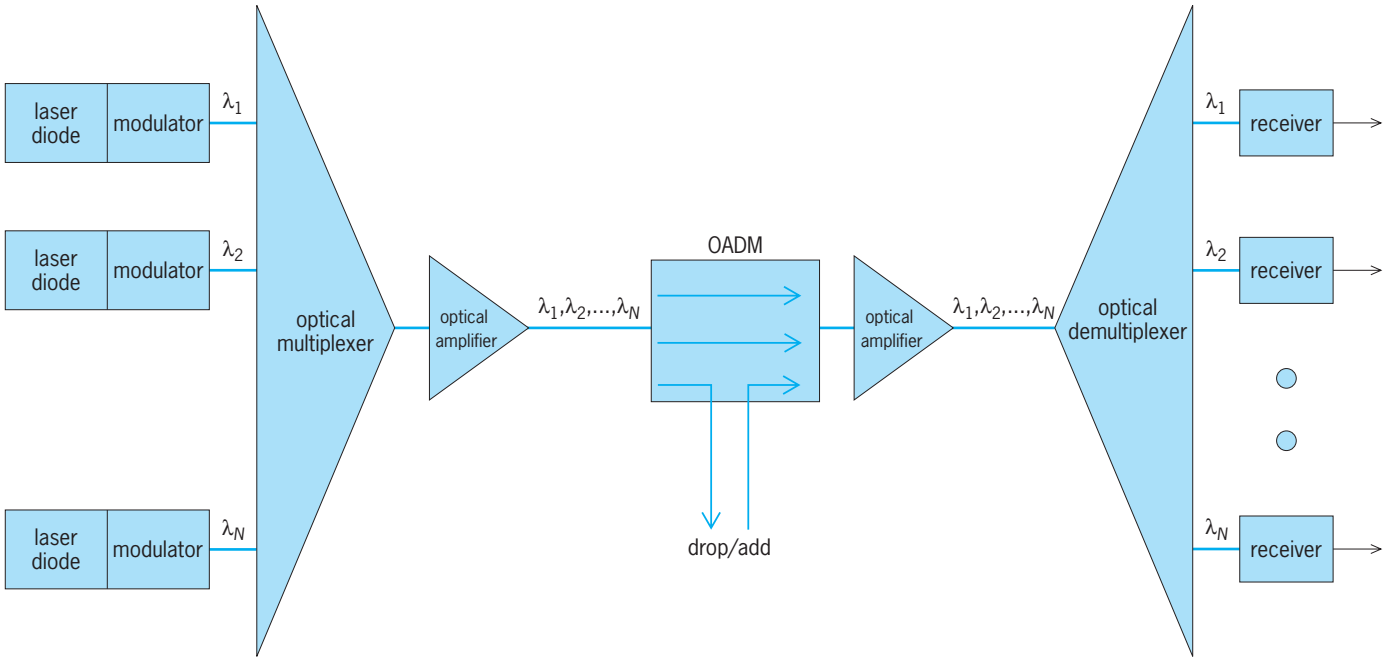

Fig. 4. Point-to-point wavelength-division multiplexing (WDM) link. Shown are laser transmitters, modulators, multiplexers and demultiplexers, and other optical components that make WDM work.

separation (called coarse channels) over the complete low-loss spectrum. This technology, known as coarse wavelength-division multiplexing (CWDM), uses the aforementioned water-free fiber with low loss over the complete spectrum from O to L, including the E subband; CWDM is not shown in the table.

WDM technology. The optical technology that was required in basic optical communications (laser, fiber, and photodiode) was now enhanced with more components to make WDM possible. Key WDM components are optical multiplexers, optical demultiplexers, and optical amplifiers. Other components include couplers, splitters, filters, equalizers, compensators, and optical add-drop multiplexers.

Optical multiplexers and demultiplexers. An optical multiplexer combines many optical channels of different wavelengths into one beam coupled onto the core of a single fiber; an optical demultiplexer separates the WDM beam into its constituent channels; and the optical add-drop multiplexer (OADM) selectively removes one of many channels from one direction and adds the same channel in the same direction but with different data on it (**Fig. 4**). The distance between the optical multiplexer and demultiplexer is of the order of tens to hundreds of kilometers. Therefore, in many applications there is a need to add and drop traffic (that is, one or more channels) in between, and the OADM accomplishes this task. For example, if New York City is connected with Washington, DC, via Philadelphia, with a point-to-point DWDM link, it may be necessary to drop some traffic in Philadelphia (for example, servicing calls in which one of the parties is in the philadelphia area). The one-way communication in Fig. 4 is for illustrative purposes; in reality, the same link exists in the opposite direction. Thus, the OADM must handle adding and dropping traffic in both directions, and it is therefore more complex than that shown in Fig. 4.

Optical amplifiers. The optical amplifier is particularly important in WDM optical communications. It is the optical amplifier that makes WDM optical networks a reality. Optical amplification in a long fiber link is needed to overcome optical signal power loss. Thus, if the signal is amplified every 60–80 km (40–50 mi), it can travel distances that exceed 1000 km (600 mi). In the original optical network with a single channel at 1310 nm, signal amplification was not a major issue because a single regenerator that consisted of a photodetector, an electronic amplifier, electrical filters, and a laser would do; this is known as opaque or optical-electrical-optical (OEO) amplification (Fig. 1). However, in WDM this method is not cost-efficient because many optical channels must be demultiplexed first, then each channel must be amplified by a separate OEO amplifier, and then the channels must be multiplexed and coupled onto the fiber (**Fig. 5**). Thus, the overall cost of amplification is multiplied by the number of channels in the WDM signal. The amplification stage is too complex to be practical; it requires large amounts of space and power because of the hundreds of amplifiers, and maintenance cost is too high. Fortunately, an optical amplifier was developed that simultaneously could amplify all optical channels in a specific subband directly without electronic amplifiers. This amplifier is known as an erbium-doped fiber amplifier (EDFA).

An EDFA is a specialty fiber whose core contains a high concentration of erbium atoms or erbium dopant. The energy of erbium atoms is excited by absorbing optical radiation at a wavelength of 980 or 1480 nm. The excitation frequency in optical amplification is known as the pump. When a photon with a wavelength in the range from 1520 to

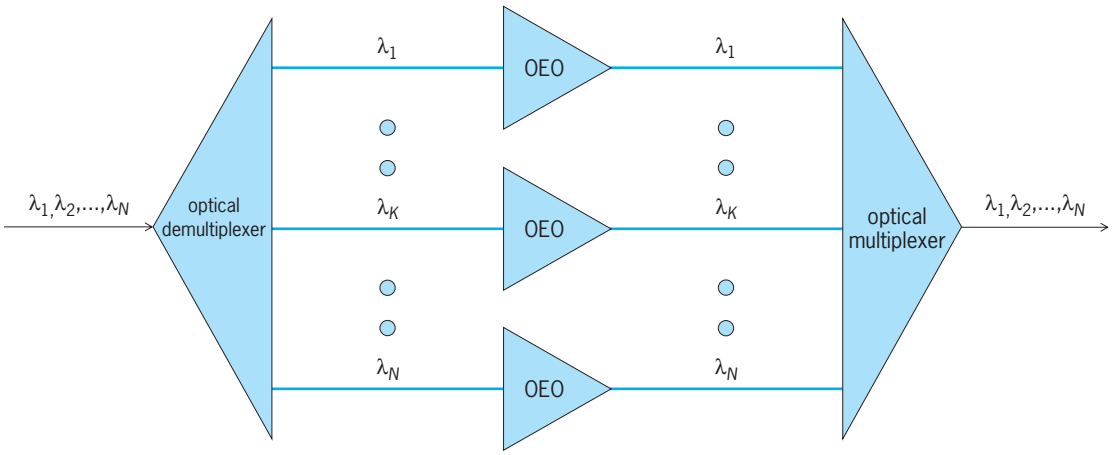

Fig. 5. Use of optical-electrical-optical (OEO) amplifiers in a WDM link. This method is complex and cost-inefficient.

1570 nm (that is, the C-band) passes near an excited erbium atom in the EDFA, it stimulates the atom, and the excited atom emits a photon at the same frequency and phase as that of the stimulating photon; this atom makes a transition to a lower energy level, where it is reexcited by the pump. When the two propagating photons encounter more excited erbium atoms, more photons are emitted. Thus, from one photon many photons are generated at the same frequency and phase, which is manifested as optical amplification.

If two photons with different frequencies A and B, but still in the C-band, enter the excited EDFA, then two excited atoms are stimulated. However, now one atom produces a photon with frequency A and the other atom a photon with frequency B, and this process continues as these photons propagate within the excited EDFA. Thus, two different signals are amplified simultaneously. This example can be extended to more photons of different frequencies, as long as they are within the C-band. When one or more optical signals in the C-band pass through an excited EDFA, they are all amplified. By adding more elements in the core of the fiber, such as aluminum and germanium, the amplification range of EDFAs was extended to include both the C and L subbands. Such fiber-optical amplifiers have been termed extended-range EDFAs.

Thus, an EDFA is a fiber, a few meters long, in which the weak information-carrying WDM signal and the laser pump are coupled and propagate. As they propagate, an evolutionary amplification process takes place (**Fig. 6**). At the end of the EDFA, an optical filter rejects any of the residual pump light, allowing only the amplified WDM signal to continue its travel over the next single-mode fiber link for another 60–80 km (40–50 mi) to the next EDFA. However, the EDFAs produce noise, known as amplifier stimulated emission (ASE), which blends with the WDM signal and cannot be eliminated. This ASE noise is cumulative, and as the signal is amplified by several EDFAs the noise content in the signal increases and degrades the signal-to-noise ratio. This degraded SNR imposes a limit on the number of concatenated EDFAs, and thus a limit on the number of fiber links and on overall fiber length. Thus, in paths that are several thousands of kilometers long, such as transoceanic fiber cables, a regenerator amplifier is placed after every 7 to 10 concatenated EDFAs to filter out noise and reconstitute the signal to its original quality.

Another type of fiber amplifier, known as the Raman amplifier, takes advantage of a phenomenon that is based on the excitation of atoms in the single-mode fiber due to the residual nonlinearity of its dielectric material. When a powerful laser light of a specific wavelength (or frequency) propagates in a single-mode fiber, atoms are excited. Then, if photons at a wavelength that are offset by about 80 nm from the pump wavelength enter the excited fiber, they stimulate the excited atoms. The stimulated atoms emit photons that have the same wavelength as the stimulating photons. Thus, Raman amplification supports signals that are offset about 80 nm from the pump wavelength and within a range of about 40 nm (**Fig. 7**). Unlike EDFAs, which have a fixed usable spectral bandwidth, Raman amplifiers can be used in any spectral subband as long as the Raman pump is offset about 80 nm from the signals to be amplified and the signals are within a 40-nm spectral range. Consequently, several Raman pumps at convenient wavelengths can amplify many signals in a wider spectral band. *See* NONLINEAR OPTICS; RAMAN EFFECT.

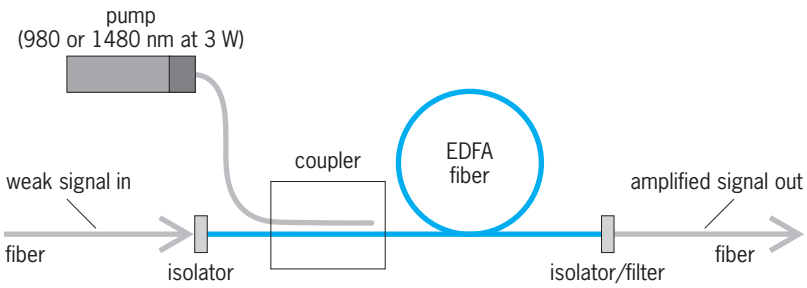

Fig. 6. Fundamental components of an erbium-doped fiber amplifer (EDFA) in a WDM system.

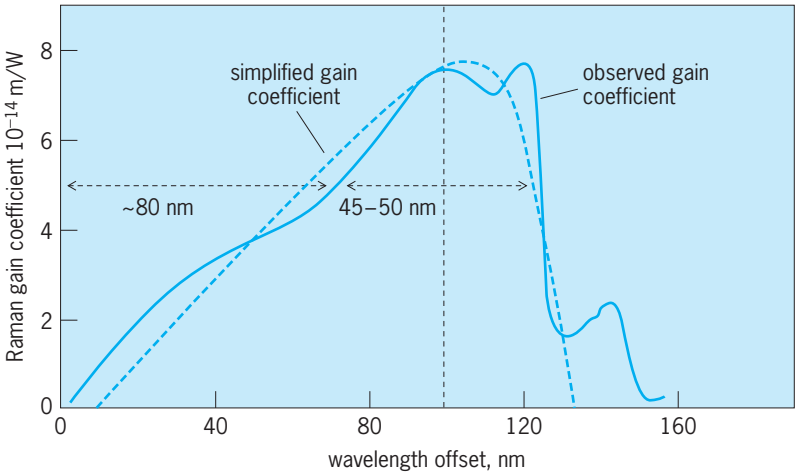

Fig. 7. Typical Raman gain coefficient over the amplification spectrum. In silica fiber at a wavelength of 1550 nm, the coefficient is 6.6 × 10^{-14} m/W.

In addition to the EDFA, there are other types of fiber amplifiers that contain different dopants, such as tellurium (TDFA), praseodymium (PDFA), and yttrium (YDFA). Each has a useful amplification range which is also fixed but is in a different frequency subband than the EDFA. Because DWDM technology currently uses the C and L subbands, these fiber amplifiers have not been widely used yet. As DWDM evolves to include the remaining frequency subbands, the applicability and popularity of these fiber amplifiers will also increase.

Optical switching and routing. Optical networks consist of a mesh of nodes that are interconnected with optical fiber. Nodes receive traffic from one fiber and, based on the predetermined route, they pass it onto another fiber that connects with the next node in the mesh, and so on. This is known as a mesh topology, which is used in the optical networks that connect cities across the United States, also known as the backbone.

To be able to pass traffic from one fiber to another, a node must support an optical switching function. Several technologies provide this function. One such technology uses lithium niobate ($LiNbO_3$) solid-state devices that, by the application of a voltage, couple an optical signal from one waveguide to another, similar to a train switching tracks. This technology is fast but attenuates the signal and is not suitable for very large switches. Another technology uses micromirrors that have been manufactured using a standard photolithography and etching process similar to that used to make integrated circuits. This technology is also known as micro-electro-mechanical systems (MEMS). Tilting the mirrors by applying an electrostatic field will redirect the optical beam in the desired direction, and thus a switching function for many optical channels is accomplished (**Fig. 8**). Multiple micromirrors are manufactured in array or matrix configurations; matrices of up to 1000×1000 micromirrors have been demonstrated. *See* ELECTROOPTICS; MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS); MICRO-OPTO-ELECTRO-MECHANICAL SYSTEMS (MOEMS); OPTICAL MODULATORS.

In current large-mesh optical networks, an end-to-end path is established by routing traffic from node to node using the same optical channel or wavelength. In many networks, this means that the network management system (NMS), that is, the system that overlooks and manages traffic and the operation of the entire network, finds the best route and establishes the end-to-end path. This is accomplished by executing a routing and wavelength assignment (RWA) algorithm and sending commands to nodes in the networks via a protocol; this is also known as path establishment. In practice, the NMS communicates with the element management system (EMS) of each subnetwork, and each EMS then communicates with the network elements (NE) or nodes that are within its responsibility in the subnetwork (**Fig. 9**). Because the number of wavelengths per fiber is finite as well as the number of nodes and links in a mesh network, there is a finite number of paths. Thus, during periods of high demand, when all possible paths have been assigned, any additional path connectivity request will be blocked.

Networks with systems that support wavelength conversion reduce the blocking probability. Wavelength converters are devices that shift an optical wavelength from that of one channel to that of another. Thus, if it happens that the same wavelength is not available over an end-to-end path, but different wavelengths are available in segments and links over the intended path, thenwavelength converters

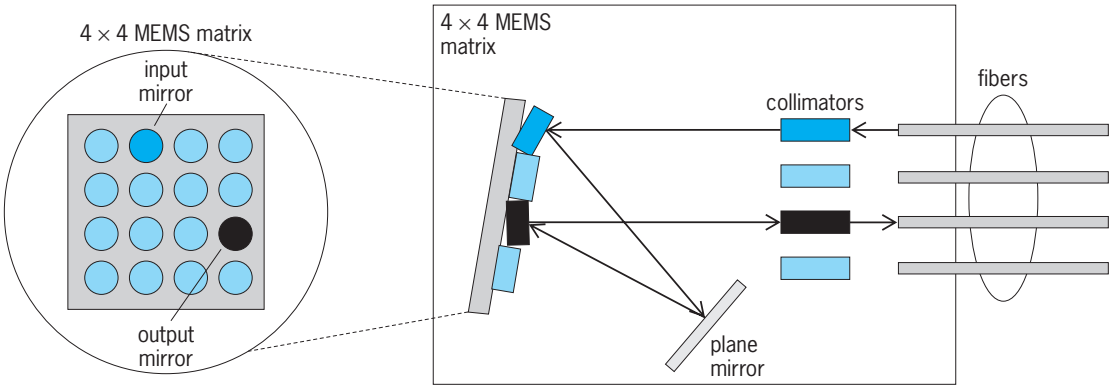

Fig. 8. Optical switch with 4 × 4 MEMS micromirror matrix.

"stitch" different wavelengths together, and the path is established. In this case, the complete path may be envisioned as a colorful array of different wavelengths stitched one after the other; this is known as path establishment with wavelength concatenation (Fig. 9). Although path establishment with wavelength concatenation is more complex and requires protocols and wavelength conversion tables at each node, it greatly decreases the probability for blocking and increases the network bandwidth throughput and bandwidth efficiency of the network.

Protection. As in previous communication networks, the network protection strategy in optical networks is very important. The mesh topology has excellent protection because, if one fiber is cut (fiber cuts happen frequently when digging roads and fields through which fiber is laid), then the switching nodes detect the fiber failure and, by using sophisticated algorithms and control messages, find another route on to which to move the affected traffic. This is similar to moving traffic in a city to another road when a particular road is blocked, but in this case the process is more orderly and takes place with minimal congestion. Another topology of interest is the ring, which is used in many metropolitan and large campus applications, and is thus named the Metro ring. The Metro ring consists of a single fiber, a dual fiber, or a quad fiber. Optical add-drop nodes are placed on the ring to remove and add traffic to be sent to its destination. Clearly, a single ring does not exhibit good protection against fiber cuts unless it is a bidirectional ring; that is, traffic flows in both directions of the ring. The dual-ring Metro has much better protection. In this case, when one fiber ring is cut, traffic from the affected ring is passed onto the healthy one and thus service continues. In the case of a quad fiber, a combination of switching and loopback mechanisms provides superb protection (**Fig. 10**); thus, the quad-ring Metro is preferred in large metropolitan applications that support ultrahigh bandwidth transmission of the order of 400 Gbps or more.

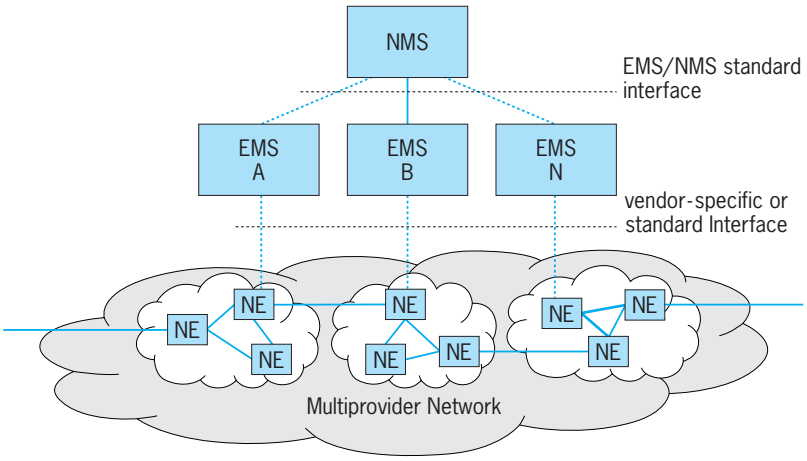

Fig. 9. Large optical network with wavelength converters. Such a network is able to concatenate wavelengths to establish an efficient end-to-end path. NMS = network management system, EMS = element management system, NE = network element.

Free-space optical (FSO) communications. In several applications, a relatively quick point-to-point path must be established in the inner city between two tall buildings. Installing fiber in a busy city is a rather lengthy process because over 1 or 2 km distance many rights of way must be obtained. However, shooting a laser beam from the top of one building to the top of another building that is 1 or 2 km away is much easier, and it does not require rights of way licences. In fact, such a network, using towers and the light of a torch, has been known since antiquity.

Although FSO is quick to deploy, it has a severe drawback. Laser light is severely absorbed by fog. Thus, in cities where fog is frequent, FSO may not be suitable unless an alternative transmission method

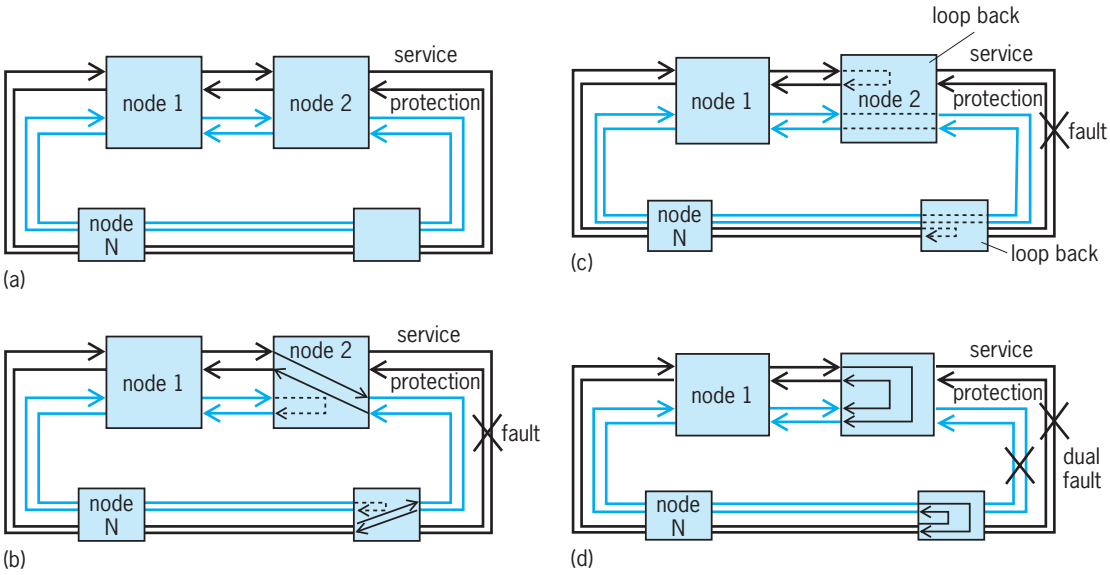

Fig. 10. Large optical four-fiber Metro ring, with protection, loopback, and path switching. (*a*) System operating with no faults. (*b*) System using path switching to operate in the presence of a single fault. (*c*) System using loopback to operate in the presence of a single fault. (*d*) System using loopback to operate in the presence of a dual fault.

can be used, such as microwaves, which are not substantially affected by fog. On the other hand, some microwave frequencies are affected by rain, which does not affect FSO as much. Thus, the two technologies can work side by side, and in case of fog the microwave link becomes operable, but not at the same data rate. [FSO transmits more than 1 Gbps whereas a microwave link transmits several megabits per second (Mbps).] *See* MICROWAVE; RADIO-WAVE PROPAGATION.

FSO technology may also include multiple wavelengths, thus utilizing WDM technology. WDM in this case may reduce the data rate per channel, yet transport more aggregate bandwidth over many channels.

In another application, laser beams may be used as intersatellite links. Satellites form a network of their own, communicating with each other via the laser beams. The neodymium yttrium-aluminum-garnet (Nd:YAG) solid-state laser is suited to this application. Satellites are thousands of kilometers apart, but light can travel such distances in space since the attenuation of light is insignificant there. The key issue is to maintain connectivity as the satellites move, which is a tracking problem and not an optical communications one. The optical issue is the divergence of the laser beam. Although the beam is very narrow at the transmitter, it diverges as it propagates for thousands of kilometers. Because of this divergence, a large-aperture telescope is needed at the receiver, to obtain an adequate signal-to-noise ratio, typically about 10 cm (4 in.) in diameter. While millimeter-wave systems compete with laser beams, laser beams can transfer much higher data rates, exceeding gigabits per second. This is what makes them attractive for satellite communications networks. Such intrasatellite optical communications links are used in the Iridium mobile communications satellite system. *See* COMMUNICATIONS SATELLITE; MOBILE RADIO.

Next-generation optical networks. The initial optical network has grown rapidly since the 1980s, responding chiefly to the rapid growth of voice and data services, as was testified by the initial offering of SONET and SDH protocols. During this evolution, however, electronic devices kept shrinking so that the personal computer became a powerful and inexpensive machine that was ubiquitous, as was mobile telephony. The rapid expansion of the Internet and its services, and also the rapid deployment of mobile and other wireless devices, gave rise to a need for more bandwidth. Thus, on the aggregate, the optical communications network created a bottleneck at access points, which for one reason or another did not follow in step with the evolution of the main optical network. *See* INTERNET; MOBILE RADIO.

However, this situation is changing, and fiber-optic communications are reaching user premises with a passive optical network (PON) technology known as fiber to the premises (FTTP). When this technology is established, the main network is expected to be flooded with all types of traffic, including voice, Ethernet, Internet, video, interactive video, text, games, and music, and all these signals are expected to be integrated over a single optical network. The next-generation network is thus already at hand, and new efficient protocols, known as next-generation SONET/SDH, have been developed that are able to encapsulate a diverse range of traffic onto a single transmission path. *See* DATA COMMUNICATIONS.

For very long haul applications over distances of the order of several thousand of kilometers, the optical transport network (OTN) has been defined with superb bit error detection and correction capability. This capability is necessary to reach such long distances, such as in the case of trans-Atlantic links, where there is no land between the east coast of the United States and the west coast of Europe. On such links, the number of amplifiers is limited as is the number of repairs (to 3 per 25 years) since repairing an undersea cable is an extremely expensive and complex operation. Among the protocols that make this capability possible are the generic framing procedure (GFP), the link access procedure SDH (LAPS), and link capacity adjustment scheme (LCAS). Both the next-generation SONET/SDH and OTN are now defined over WDM technology, and thus the next-generation network will be ready to meet the most aggressive bandwidth demands for many years to come in order to provide reliable, cost-effective, scalable, and secure service. *See* SUBMARINE CABLE.

S. V. Kartalopoulos

Bibliography. A. Borella, G. Cancellieri, and F. Chiaraluce, *Wavelength Division Multiple Access Optical Networks*, Artech House, 1998; E. B. Carne, *Telecommunications Primer*, 2d ed., Prentice Hall, 1995; I. P. Kaminow and T. L. Koch (eds.), *Optical Fiber Communications IIIA and Optical Fiber Communications IIIB*, Academic Press, 1997; S. V. Kartalopoulos, *DWDM: Networks, Devices and Technology*, Wiley-IEEE Press, 2003; S. V. Kartalopoulos, *Fault Detectability in DWDM Systems*, Wiley-IEEE Press, 2001; S. V. Kartalopoulos, *Introduction to DWDM Technology*, Wiley-IEEE Press, 2000; S. V. Kartalopoulos, *Next Generation SONET/SDH*, Wiley-IEEE Press, 2004; S. V. Kartalopoulos, *Optical Bit Error Rate*, Wiley-IEEE Press, 2004; S. V. Kartalopoulos, *Understanding SONET/SDH and ATM*, Wiley-IEEE Press, 1999; J. Nellist, *Understanding Telecommunications and Lightwave Systems*, 2d ed., IEEE Press, 1996; J. C. Palais, *Fiber Optic Communications*, 5th ed., Prentice Hall, 2004; R. Ramaswami and K. N. Sivarajan, *Optical Networks*, 2d ed., Academic Press, 2002.

Optical detectors

Devices that respond to incident ultraviolet, visible, or infrared electromagnetic radiation by giving rise to an output signal, usually electrical. Based upon the manner of their interaction with radiation, they fall into three categories. Photon detectors are those in which incident photons change the number of free carriers (electrons or holes) in a semiconductor (internal photoeffect) or cause the emission of free electrons from the surface of a metal or semiconductor (external photoeffect, photoemission). Thermal

detectors respond to the temperature rise of the detecting material due to the absorption of radiation, by changing some property of the material such as its electrical resistance. Detectors based upon wave-interaction effects exploit the wavelike nature of electromagnetic radiation, for example by mixing the electric-field vectors of two coherent sources of radiation to generate sum and difference optical frequencies. This discussion concentrates on photon and thermal detectors. *See* NONLINEAR OPTICAL DEVICES.

Photon effects. The most widely used photon effects are photoconductivity, the photovoltaic effect, and the photoemissive effect. Photoconductivity, an internal photon effect, is the decrease in electrical resistance of a semiconductor caused by the increased numbers of free carriers produced by the absorbed radiation. The change in resistance is measured by passing a bias current through the photoconductor and measuring the change in resistance between the illuminated and dark state. *See* PHOTOCONDUCTIVE CELL.

The photovoltaic effect, also an internal photoeffect, occurs at a *pn* junction in a semiconductor or at a metal-semiconductor interface (Schottky barrier). Absorbed radiation produces free hole-electron pairs which are separated by the potential barrier at the *pn* junction or Schottky barrier, thereby giving rise to a photovoltage. This is the principle employed in a solar cell. *See* PHOTODIODE; PHOTOVOLTAIC CELL; PHOTOVOLTAIC EFFECT; SEMICONDUCTOR DIODE; SOLAR CELL.

The photoemissive effect, also known as the external photoeffect, is the emission of an electron from the surface of a metal or semiconductor (cathode) into a vacuum or gas due to the absorption of a photon by the cathode. The photocurrent is collected by a positively biased anode. Internal amplification of the photoexcited electron current can be achieved by means of secondary electron emission at internal structures (dynodes). Such a vacuum tube is known as a photomultiplier. Internal amplification by means of an avalanche effect in a gas is employed in a Geiger tube. *See* GEIGER-MÜLLER COUNTER; LIGHT AMPLIFIER; PHOTOELECTRIC DEVICES; PHOTOEMISSION; PHOTOMULTIPLIER; PHOTOTUBE.

Role of materials. Semiconductors are key to the development of most photon detectors. These materials are characterized by a forbidden energy gap which determines the minimum energy that a photon must have to produce a free hole-electron pair in an intrinsic photoeffect. Since the energy of a photon is inversely proportional to its wavelength, the minimum energy requirement establishes a long-wavelength limit of an intrinsic photoeffect. It is also possible to produce free electrons or free holes by photoexcitation at donor or acceptor sites in the semiconductor; this is known as an extrinsic photoeffect. Here the long-wavelength limit of the photoeffect is determined by the minimum energy (ionization energy) required to photoexcite a free electron from a donor site or a free hole from an acceptor site. In a given semiconductor, the long-wavelength limit of an extrinsic photoeffect exceeds that of the intrinsic photoeffect. In either the intrinsic or extrinsic photoeffect, the ability to detect a photon is described by the quantum efficiency, which is the ratio of the photoexcited carriers produced to the number of incident photons. Quantum efficiencies of 0.5 are typical of many photon detectors. *See* SEMICONDUCTOR.

Detector cooling. Free electrons and holes can also be thermally excited by lattice vibrations (phonons) in the semiconductor, whose magnitude is a function of the detector temperature. Detection of a few photoexcited free carriers in the presence of numerous thermally excited carriers is difficult. The problem is eased by cooling the detector in order to reduce the thermally excited carrier background. Long-wavelength infrared photon detectors having spectral responses extending to 10 micrometers have small forbidden energy gaps or small donor or acceptor ionization energies. Accordingly, cooling the detector becomes more important for long-wavelength detectors. Use of liquid nitrogen, which boils at 77 K, or mechanical refrigerators (cryocoolers) is required for photon detectors whose spectral response extends to 10 μm. Even longer-wavelength spectral responses require cooling below 77 K. *See* CRYOGENICS.

Synthetic materials. The usual way to provide a given spectral response for a photon detector is to synthesize a semiconductor whose forbidden energy gap or donor or acceptor ionization energy meets the long-wavelength limit requirement. The semiconductor material most frequently used in long-wavelength military systems is mercury cadmium telluride, wherein the mercury-to-cadmium ratio determines the forbidden energy gap.

An alternative method is to tailor a new material by building up numerous alternating layers of controlled thickness of two semiconductors which have compatible material properties (crystal structure, lattice constant, thermal expansion coefficient). Such structures are known as multiple quantum wells. The most frequently used pair of materials is gallium arsenide and aluminum gallium arsenide. By adjusting the aluminum-to-gallium ratio and the thicknesses of the layers during growth, the spectral response of the photoeffect exhibited by these structures can be tailored within broad limits. When such structures are employed for infrared detection, they are known as quantum-well infrared photodetectors (QWIPs). *See* ARTIFICIALLY LAYERED STRUCTURES; CRYSTAL GROWTH; QUANTIZED ELECTRONIC STRUCTURE (QUEST); SEMICONDUCTOR HETEROSTRUCTURES.

Thermal effects. The choice of materials also plays a role in thermal detectors. The most widely used thermal detector is a bolometer, that is, a temperature-sensitive resistor in the form of a thin metallic or semiconductor film (although superconducting films are also used). Incident electromagnetic radiation absorbed by the film causes its temperature to rise, thereby changing its electrical resistance. The change in resistance is measured by

passing a current through the film and measuring the change in voltage. Materials with a high temperature coefficient of resistance are desired for bolometers, a criterion which usually favors semiconductors over metals. *See* BOLOMETER.

Another widely used thermal detector is based upon the pyroelectric effect. Here a change in temperature causes positive and negative charges to appear on opposite faces of thin samples of certain ferroelectric materials such as triglicine sulfate and strontium barium niobate. This charge can be detected by means of electrodes deposited on the faces. In contrast to a bolometer, a pyroelectric detector requires the incident radiation to be temporally modulated, which can be done by means of a radiation chopper. *See* CHOPPING; PYROELECTRICITY.

A third thermal detector is the radiation thermocouple, formed by two junctions between two dissimilar metals or semiconductors. One junction, exposed to the incident radiation, is warmed by it; the other, shielded junction is not. A voltage generated by the temperature difference between the junctions is detected by an external circuit. *See* THERMOCOUPLE; THERMOELECTRICITY.

In contrast to photon detectors, the spectral response of thermal detectors is determined by the wavelength dependence of their absorption of electromagnetic radiation. If the absorption is invariant with wavelength, the spectral response is invariant with wavelength and is said to be flat.

Elemental and imaging detectors. Optical detectors can also be classified as elemental or imaging. An elemental detector averages any spatial variation in the incident radiation to produce an output signal characteristic of the average incident intensity. An imaging detector produces a time-varying output signal which bears a one-to-one relationship with the spatial variation of intensity falling on it. An example is a television camera tube known as a vidicon, which contains a broad-area thin photoconductive film upon which the radiant image of a scene is projected by a lens. An internal electron beam scanned across the film in raster fashion generates a time-varying electrical signal representing the spatial variation of film resistance due to the projected image. It is also possible to detect an image by mechanically scanning the scene across an elemental detector by using two moving mirrors, one to scan horizontally and the other vertically. Use of a linear array of detectors simplifies the system complexity by requiring scanning in one direction only. *See* TELEVISION CAMERA TUBE.

Applications. The applications of optical detectors can be organized according to the spectral interval within which they respond.

Ultraviolet detectors. These find limited applications in laboratory instrumentation such as spectrometers and in astronomy. Industrial applications include the monitoring of gas-burner flames in large furnaces. Ultraviolet detectors are used also in dual-color (infrared and ultraviolet) missile-guidance schemes. The principal ultraviolet detector is the photomultiplier. *See* ULTRAVIOLET ASTRONOMY.

Visible detectors. Visible radiation detectors have many applications, for example, in television cameras and camcorders. Solid-state arrays based upon silicon charged-coupled-device technology are widely used in simple, inexpensive high-performance imaging systems. Such arrays also find use in astronomy, where their sensitivity at low light levels is of key importance. They are used in industrial automation systems, including those employing machine vision. Electronic still photography employs imaging arrays to replace photographic film in solid-state cameras. *See* CAMERA; CHARGE-COUPLED DEVICES; COMPUTER VISION.

Infrared detectors. Detectors of infrared radiation are widely used in military and, to less extent, commercial night-vision systems. Both image intensifiers, which exploit photoemissive photocathodes responding to the low-level ambient illumination of the night sky at wavelengths out to about 1.1 μm, and thermal imaging systems, employing cryogenic linear and matrix (two-dimensional) arrays of long-wavelength photon detectors, are widely used. Uncooled matrix arrays of bolometers and pyroelectric detectors offer high performance in thermal imaging systems without the complexity and cost of cryogenic systems. *See* INFRARED ASTRONOMY; INFRARED IMAGING DEVICES; INFRARED RADIATION.

Paul W. Kruse

Bibliography. E. L. Dereniak and D. G. Crowe, *Optical Radiation Detectors*, 1984; R. J. Keyes (ed.), *Optical and Infrared Detectors*, 2d ed., 1980; R. H. Kingston, *Optical Sources, Detectors and Systems: Fundamentals and Applications*, 1995; Technology guide detector handbook, *Laser Focus World*, pp. A3–A42, March 1991; R. K. Willardson and A. C. Beer (eds.), *Semiconductors and Semimetals*, vol. 5: *Infrared Detectors*, 1970, vol. 12: *Infrared Detectors*, 1976.

Optical fibers

Transparent strands of glass, plastic, or other flexible material used for light delivery. The light delivery can be for the purpose of signal (data) transmission, light amplification, image transmission, or energy transmission. The light can be confined by total internal reflection, by mirrorlike reflection from an internal coated surface, or by a suitably designed set of interior air holes formed parallel to the axis of the fiber. Optical fibers are, in their most common form, a type of dielectric guide for electromagnetic waves (waveguide). In certain forms, fibers are also termed lightguides. *See* REFLECTION OF ELECTROMAGNETIC RADIATION.

In the most common form, the optical fiber consists of a core of material with a refractive index higher than the surrounding cladding. If a light source such as a laser is directed into the fiber, the light travels (propagates) along the fiber and may be detected at the other end. Information may be encoded on the light by on-off keying (a binary system in which on = 1, off = 0), frequency modulation,

or phase modulation. If the wavelength is chosen correctly, fibers may transmit light over very long distances of 100 km (60 mi) or more. For even longer distances, specialty fibers are fabricated with rare-earth ions, which enable optical amplification. Fiber amplifiers enable both undersea and terrestrial communication over great distances. A fiber amplifier which is equipped with suitable end mirrors may also function as a laser source. *See* FIBER-OPTIC CIRCUIT; FREQUENCY MODULATION; MODULATION; OPTICAL COMMUNICATIONS; PHASE MODULATION; REFRACTION OF WAVES.

Fibers may also be fused together into bundles for image transmission. By assigning one fiber core to each image element (pixel), the fiber provides a flexible device in which the light pattern at the entrance to the bundle is approximately replicated at the output of the bundle. These bundles are often used in medical imaging and inspection. In other cases, optical fibers are a flexible means of laser power delivery. For example, certain medical applications require laser light or similar illumination to a specific point inside the human body. Fibers can be designed to deliver a very high irradiance to places which would otherwise require invasive surgery. Another application for optical fibers is in sensors, where a change in light transmission properties is used to sense or detect a change in some property, such as temperature, pressure, or magnetic field. *See* FIBER-OPTIC SENSOR; FIBER-OPTICS IMAGING.

Fused silica (SiO_2) is the most common fiber material. Plastic fibers are used for low-cost, short-distance optical links. Fluoride-based glasses are required for transmission in the mid-infrared region of the optical spectrum.

Fiber designs. There are five basic types of optical fibers (**Fig. 1**). Propagation in these lightguides is most easily understood by ray optics, although the wave or modal description must be used for exactness. In a multimode, stepped-refractive-index-profile fiber (Fig. 1*a*), the number of rays or modes of light which are guided, and thus the amount of light power coupled into the lightguide, is determined by the core size and the core-cladding refractive index difference. Such fibers are limited to short distances for information transmission due to pulse broadening. An initially sharp pulse made up of many modes broadens as it travels long distances in the fiber, since high-angle modes have a longer distance to travel relative to the low-angle modes. This limits the bit rate and distance because it determines how closely input pulses can be spaced without overlap at the output end. At the detector, the presence or absence of a pulse of light in a given time slot determines whether this bit of information is a zero or one.

A graded-index multimode fiber (Fig. 1*b*), where the core refractive index varies across the core diameter, is used to minimize pulse broadening due to intermodal dispersion. Since light travels more slowly in the high-index region of the fiber relative to the low-index region, significant equalization of the transit time for the various modes can be achieved to reduce pulse broadening. This type of fiber is suitable for intermediate-distance, intermediate-bit-rate transmission systems. For both fiber types, light from a laser or light-emitting diode can be effectively coupled into the fiber. *See* LASER; LIGHT-EMITTING DIODE.

A single-mode fiber (Fig. 1*c*) is designed with a core diameter and refractive index distribution such that only one fundamental mode is guided, thus eliminating intermodal pulse-broadening effects. Material and waveguide dispersion effects cause some pulse broadening, which increases with the spectral width of the light source. These fibers are best suited for use with a laser source in order to efficiently couple light into the small core of the waveguide and to enable information transmission over long distances at very high bit rates. The specific fiber design and the ability to manufacture the fiber with controlled refractive index and dimensions determine its ultimate bandwidth or information-carrying capacity. *See* WAVEGUIDE.

A special class of single-mode fibers comprises polarization-preserving fibers. In an ideal, perfectly circular single-mode fiber core, the polarization state of the propagating light is preserved, but in a real fiber various imperfections can cause birefringence; that is, the two orthogonally polarized modes of the fundamental mode travel at different speeds. For applications such as sensors, where controlling the polarization is important, polarization-maintaining fibers can be designed that deliberately introduce a polarization. This is typically accomplished by using noncircular cores (shape birefringence) or by introducing asymmetric stresses (stress-induced birefringence) on the core. In this manner, the polarization

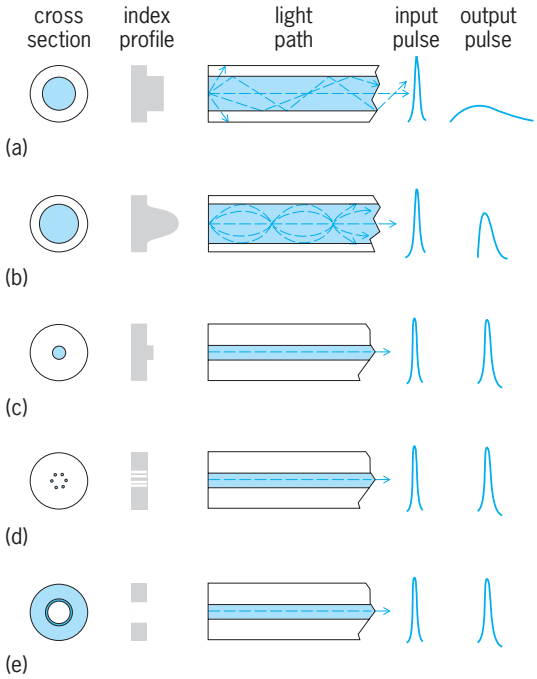

Fig. 1. Types of optical fiber designs. (*a*) Multimode, stepped-refractive-index-profile. (*b*) Multimode, graded-index-profile. (*c*) Single-mode, stepped-index. Graded-index is possible. (*d*) Microstructured optical fiber. (*e*) Hollow-core fiber for high-power laser delivery.

properties of the input light traveling through the fiber can be controlled. *See* BIREFRINGENCE; PHOTOELASTICITY; POLARIZED LIGHT.

There are two fiber structures that confine light without total internal reflection. Microstructured optical fibers (Fig. 1*d*), a category which includes photonic bandgap fibers, confine light in a small core region surrounded by micrometer-size air holes. Such fibers provide useful and interesting applications in nonlinear optics, optical switching, and frequency conversion. The confinement of an ultrashort laser pulse within such a waveguide can result in the generation of a supercontinuum, or a very large range of optical frequencies. A hollow-core optical fiber (Fig. 1*e*) confines light through the use of an interior coating (metal or multilayer dielectric) of high reflectance. Such fibers are used in cases of very high laser power or for wavelengths at which transparent fiber materials are not available.

Attenuation. The attenuation or loss of light intensity is an important property of the lightguide since it limits the achievable transmission distance, and is caused by light absorption and scattering. Every material has some fundamental absorption due to the atoms or molecules composing it. In addition, the presence of other elements as impurities can cause strong absorption of light at specific wavelengths. Fluctuations in a material on a molecular scale cause intrinsic Rayleigh scattering of light. In actual fiber devices, fiber-core-diameter variations or the presence of defects such as bubbles can cause additional scattering light loss. The light loss of a material, after the light has traveled a length L, is related to the initial power coupled into the fiber, P_0, versus the power at the output end, P, by the equation. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; SCATTERING OF ELECTROMAGNETIC RADIATION.

$$\text{Loss (db/km)} = \frac{10}{L/(\text{km})} \log\left(\frac{P_0}{P}\right)$$

Optical fibers based on silica glass have an intrinsic transmission window at near-infrared wavelengths with extremely low losses (**Fig. 2**). Very special glass-making techniques are required to reduce iron and water (OH) to the parts-per-billion level, and have resulted in losses as low as 0.16 dB/km (0.26 dB/mi). Such fibers are used with solid-state lasers and light-emitting diodes for information transmission, especially for long distances (greater than 1 km or 0.6 mi). Plastic fibers exhibit much higher intrinsic as well as total losses, and are more commonly used for image transmission, illumination, or very short distance data links. *See* OPTICAL MATERIALS.

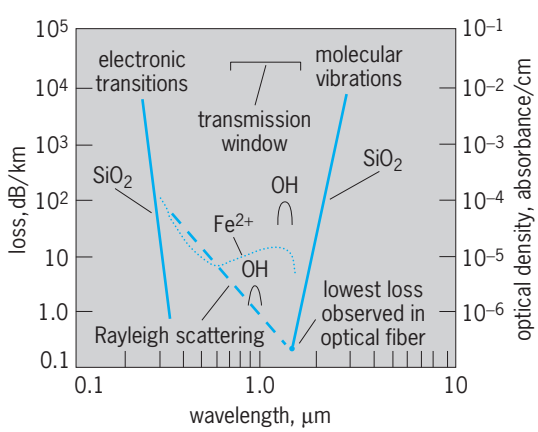

Fig. 2. Loss mechanisms in silica-based fibers.

Many other fiber properties are also important, and their specification and control are dictated by the particular application. Good mechanical properties are essential for handling: plastic fibers are ductile, while glass fibers, intrinsically brittle, are coated with a protective plastic to preserve their strength. Glass fibers have much better chemical durability and can operate at higher temperatures than plastics. Very tight tolerances on core and outer-diameter control are essential for information transmission fibers, especially to allow long lengths to be assembled with low-loss joining or splicing.

Suzanne R. Nagel; Thomas G. Brown

Bibliography. J. A. Buck, *Fundamentals of Optical Fibers*, Wiley, 1995; R. J. Hoss and E. A. Lacy, *Fiber Optics*, 2d ed., Prentice Hall, 1997; S. L. Wymer Meardon, *The Elements of Fiber Optics*, Prentice Hall, 1993; S. E. Miller and I. P. Kaminow (eds.), *Optical Fiber Telecommunications*, vol. 2, Academic Press, 1988; B. P. Pal, *Fundamentals of Fiber Optics in Telecommunication and Sensor Systems*, 1993; J. C. Palais, *Fiber Optic Communications*, 5th ed., Prentice Hall, 2004.

Optical flat

A disk of high-grade quartz glass approximately 0.75 in. (2 cm) thick, having at least one side ground and polished with a deviation in flatness usually not exceeding 0.000002 in. (50 nanometers) all over, and a surface quality of 5 microfinish or less. When two surfaces of this quality are placed lightly together so that the air is not wrung out from between them, they are separated by a film of air and actually touch at only one point. This point is the vertex of a wedge of air separating the two pieces.

If parallel beams of light pass through the flat, part will be reflected against the surface being inspected, while part will be reflected directly back through the flat. Because the distance between the surfaces is constantly increasing along the angle, the beams reflected from the flat and the beams reflected from the workpiece will alternately reinforce and interfere with each other, producing a pattern of alternate light and dark bands (**Fig. 1**). Each succeeding full band from a point of contact means the distance between surfaces is one wavelength thicker. If the light is relatively monochromatic, the wavelength is known. Red with a wavelength of 0.0000116 in. (295 nm) is commonly used. Thus a definite relationship is established between lineal measurement and light waves. Optical flats are used for two general purposes.

Fig. 1. Optical flat being used to determine flatness of seal ring. Interference bands on the seal ring face show lines of constant depth. (*Van Keuren Co.*)

Determination of surface contour. If the surface is flat, the light bands are parallel. Deviation from flatness shows as curvature of the lines. The principle in interpreting a pattern is almost identical with the principle in interpreting a topographical map in that the bands connect points of equal distance from the master surface of the optical flat. Deviation from flatness can be reduced to rational figures.

Comparison of lineal measurement. When an optical flat is placed across a gage block or a buildup of blocks and another object, as an end standard, or a cylinder or sphere, both resting on a precisely flat surface, then the angle between the blocks and the flat can be measured, and the difference in length between the gage blocks and the length or height or diameter of the unknown can be determined (**Fig. 2**). With the high accuracy available in electrical and pneumatic comparators, optical flats are seldom used in this way except for spheres or irregular surfaces where point contact by the comparator is impractical.

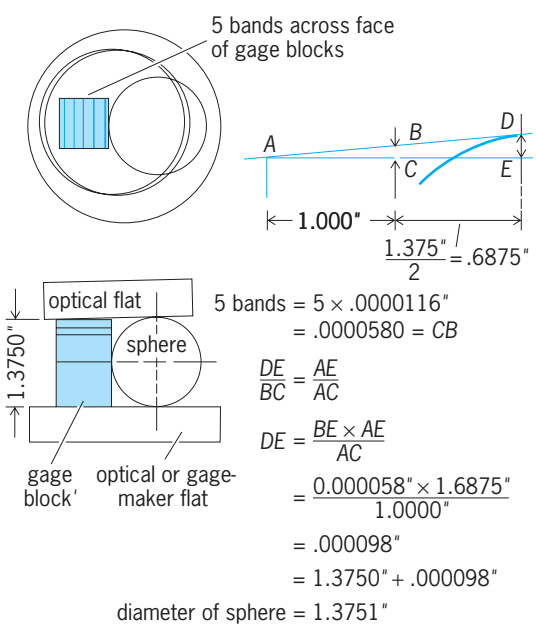

Fig. 2. Measurement of height of sphere determined by means of gage blocks and optical flat. 1″ = 25.4 mm.

The principle of interferometry is the standard method for measurement of gage blocks. However, an interferometer and not the optical flat itself is used. The wavelength of light is the present standard of all lineal measurements. *See* GAGES; INTERFEROMETRY.

Rush A. Bowman

Optical guided waves

Optical-frequency electromagnetic waves confined within an optical waveguide, a structure designed to carry such waves from one place to another somewhat as a pipe carries water. (The terms optical and light are used here in broadest sense to include visible and near-infrared electromagnetic radiation.) The demonstrations of the first semiconductor laser and the first low-loss glass optical fiber initiated a technological revolution. Because of the high data rates that can be achieved, the transmission of information in the form of optical guided waves confined within an optical-fiber waveguide has become the preferred method for the telecommunications industry. That the semiconductor laser, the source that produces the optical signals, also employs an optical waveguide makes clear the pivotal role played by optical guided waves in modern communications technology.

Total internal reflection. Optical waveguides confine light by the method of total internal reflection. At a plane interface between two media (**Fig. 1***a*), the upper and lower media are characterized by indices of refraction n_1 and n_2, respectively. When a beam of light strikes the interface at angle θ_1, part of the beam is reflected at the same angle θ_1, and the remainder of the beam is transmitted into the second medium at angle θ_2. According to Snell's law, Eq. (1),

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{1}$$

the angle θ_2 exceeds θ_1 if the refractive index n_1 exceeds n_2, which means that θ_2 will reach its physical limit of 90° for an incident angle $\theta_1 < 90°$. The incident angle for which $\theta_2 = 90°$, called the critical angle (θ_c), is given by Eq. (2).

$$n_1 \sin\theta_c = n_2 \tag{2}$$

For angles of incidence greater than or equal to the critical ($\theta_1 \geq \theta_c$), there is no transmitted beam and, in the absence of absorption, the energy carried by the incident beam is completely reflected into the first medium at angle θ_1. *See* REFLECTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES.

Rays or waves. It is useful to first think of light as propagating as a ray. Total internal reflection can be used to trap a ray of light in a thin layer (Fig. 1*b*), provided that its refractive index is larger than the

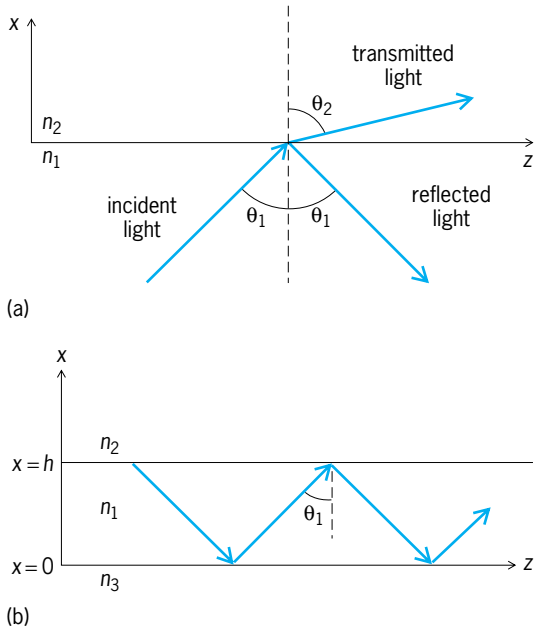

Fig. 1. Total internal reflection. (*a*) Geometry of reflection at the plane interface between two media. (*b*) Light trapped by total internal reflection in a planar optical waveguide of thickness *h* and refractive index n_1. Refractive indices of boundary media, n_2 and n_3, are less than n_1.

refractive indices of the two boundary media. If the internal propagation angle θ_1 exceeds the critical angle for total internal reflection at both the upper and lower interfaces, the light will be unable to escape the thin layer. Each reflection will be perfect, and the light will travel along a direction parallel to the boundaries of the layer without loss in a zigzag pattern until it reaches the end of the structure. A waveguide of this sort is called a planar (or slab) optical waveguide and is used to confine light, for example, in the active region of a semiconductor laser. Optical guided waves can thus be thought of as rays of light trapped within a structure by total internal reflection.

The above argument suggests that light should be able to follow any zigzag path as long as the angle θ_1 exceeds the critical angle for total internal reflection at both interfaces. A more detailed analysis based on Maxwell's equations reveals that the allowed angles in any optical waveguide form a discrete set rather than a continuum. For any waveguide, there exists a mathematical expression called the dispersion relation that shows how the allowed propagation angles are related to the optical wavelength, the size of the waveguide, and the refractive indices of the materials that make up the waveguide structure. The dispersion relation also contains at least one integer, say m, that is used to label a particular solution. The allowed angles are called the mode angles, and each unique trajectory through the waveguide is said to be a mode of the waveguide. The thickness or diameter of an optical waveguide can be made sufficiently small that only a single mode is supported for a given orientation of the electric field (polarization). *See* POLARIZATION OF WAVES.

The ray picture of the propagation of light is necessarily limited. Light is an electromagnetic wave phenomenon that is better described with Maxwell's equations. A rigorous analysis reveals that an optical guided wave, or mode, is best pictured as a unique electric field distribution that propagates along the waveguide axis with a unique velocity. Similar distributions are found for planar waveguides (**Fig. 2**) and optical-fiber waveguides. It is significant to note that not all of the guided-wave's electric field, and hence energy, is confined within the layer. The spillover into the adjacent media is characteristic of a proper treatment of total internal reflection using Maxwell's equations. *See* ELECTROMAGNETIC WAVE TRANSMISSION; MAXWELL'S EQUATIONS.

Optical waveguides. Optical guided waves can propagate in a variety of structures. The optical fiber used in the communications industry consists of two concentric glass cylinders. The inner core region is made of a glass that has a slightly higher index of

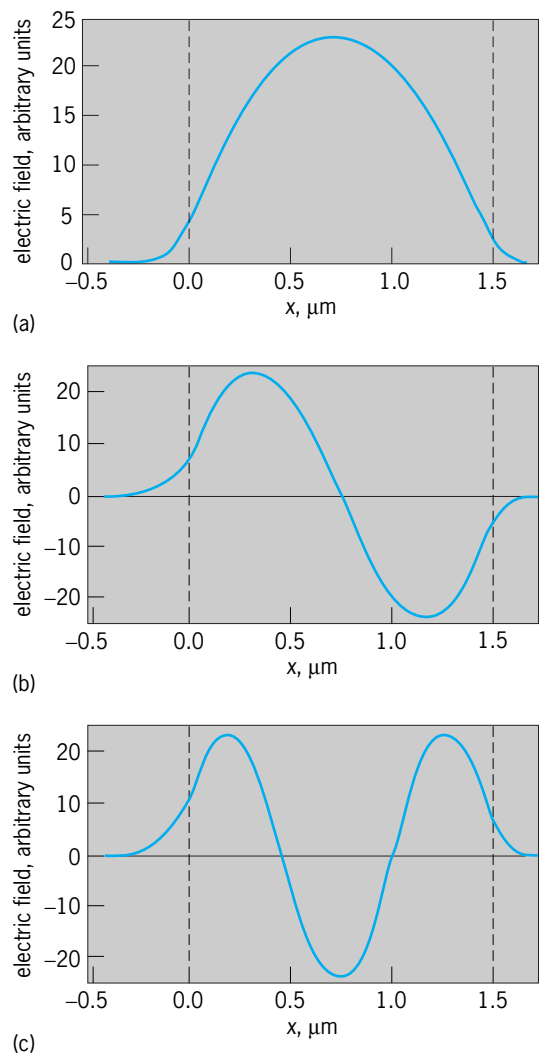

Fig. 2. Electric field distributions for the three lowest-order modes of a planar optical waveguide of thickness $h = 1.5$ μm. Electric field magnitudes are plotted as functions of the *x* coordinate in Fig. 1*b* for an arbitrary, fixed value of *z*. (*a*) Mode $m = 0$. (*b*) $m = 1$. (*c*) $m = 2$.

refraction than the outer cladding region, as required for total internal reflection to occur. In telecommunications applications, core diameters in the range of 4–50 μm are used routinely to carry near-infrared optical radiation. Pulses of light with wavelengths between 0.8 and 1.55 μm are injected into the fiber on the near end. The presence or absence of a pulse is interpreted as a one or a zero by a receiver on the far end, typically many miles (kilometers) away. These bits of information are carried by the optical fiber from the transmitter to the receiver as optical guided waves. It is a feature of optical waveguides that the light pulses broaden as they travel down the optical fiber. If the pulses broaden too much, they begin to overlap, making it difficult to distinguish one from another. This effect provides the fundamental limitation to the information capacity of a fiber-optic communication system. *See* OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

Devices such as semiconductor lasers make use of optical guided waves over distances much shorter than those spanned by optical fibers. In a semiconductor laser, a planar waveguide (Fig. 2) also serves as an electrically driven amplifying medium. Gallium arsenide (GaAs), aluminum gallium arsenide (AlGaAs), or indium gallium arsenide phosphide (InGaAsP) are common examples of semiconductor laser materials. When mirrors are formed on the ends of the semiconductor waveguide, typically a few hundred micrometers apart, light travels back and forth between those mirrors just as with any laser, but in the form of optical guided waves. The semiconductor laser is a waveguide laser. *See* LASER.

Optical waveguides are the basic constituents of an emerging technology known variously as integrated optics, integrated optoelectronics, or photonic integrated circuits. This technology integrates optical and electronic components into or on an optical waveguide. The aim is to process and manipulate light while it is trapped as optical guided waves within the confines of the optical waveguide. *See* INTEGRATED OPTICS; WAVEGUIDE. Dennis G. Hall

Bibliography. M. J. Adams, *An Introduction to Optical Waveguides*, 1981; M. Koshiba, *Optical Waveguide Analysis*, 1992; D. Marcuse, *Theory of Dielectric Optical Waveguides*, 2d ed., 1991; A. R. Mickelson, *Guided Wave Optics*, 1993; T. Tamir et al. (eds.), *Guided-Wave Optoelectronics*, 2d ed., 1990.

Optical image

The image formed by the light rays from a self-luminous or an illuminated object that traverse an optical system. The image is said to be real if the light rays converge to a focus on the image side and virtual if the rays seem to come from a point within the instrument (see **illus.**).

The optical image of an object is given by the light distribution coming from each point of the object at the image plane of an optical system. The ideal image of a point according to geometrical optics is obtained when all rays from an object point unite

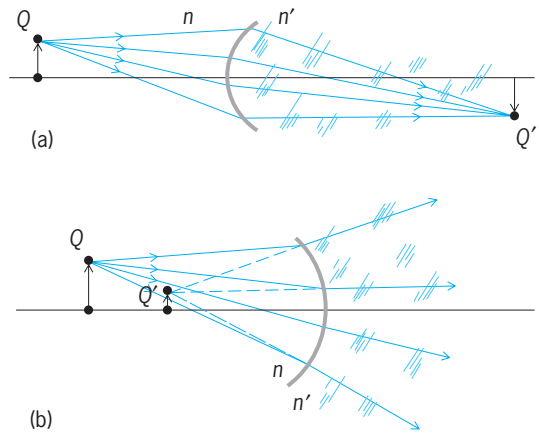

Optical images. (*a*) Real image. Rays leaving object point *Q* and passing through the refracting surface separating media *n* and *n′* are brought to a focus at the image point *Q′*. (*b*) Virtual image. Rays leaving *Q* and refracted by the concave surface separating *n* and *n′* appear to be coming from the virtual image point *Q′*. As the rays are diverging, they cannot be focused at any point. (*After F. A. Jenkins and H. E. White, Fundamentals of Optics, 4th ed., McGraw-Hill, 1976*)

in a single image point. However, diffraction theory teaches that even in this case the image is not a point but a minute disk. The diameter of this disk is about $1.22\lambda/(NA)$, where λ is the wavelength of the light considered and NA is the numerical aperture, the sine of the largest cone angle on the image side multiplied by its refractive index (which is usually equal to unity). *See* FOCAL LENGTH; GEOMETRICAL OPTICS.

Aberrations. From the standpoint of geometrical optics, if this most desirable type of image formation cannot be achieved, the next best objective is to have the image free from all but aperture errors (spherical aberration). In this case the light distribution in the image plane is still circular, resembling the point image; there is a true coordination of object point and image, although the image may be slightly unsharp. If the aperture errors are small, or if the image is viewed from a distance, such an image formation may be very satisfactory.

Asymmetry and deformation errors may be very disturbing if not held in check, because the light distribution of the image of a point in this case has a decidedly undesirable shape.

When the image of an axis point is considered, the rays through a fixed aperture circle converge to an axis point. For this type of imagery, the term half-sharp image will be used. A small object at the object point is then imaged by a circular stop at the focus of the image bundle with a magnification as given by Eq. (1),where u and u' are the angles of the imaging

$$m = \frac{n \sin u}{n' \sin u'} \tag{1}$$

cone in object and image space, respectively, and n and n' are the corresponding refractive indices.

If the axis point is sharply imaged, an object of finite extent is sharply imaged if, and only if, $m = m_0$ (the gaussian magnification) for all values of u (sine condition).

In the case of aperture errors, the most desirable image formation for an axis point is attained when the different images appear under the same angle from the exit pupil. If k' is the distance of the image point from the exit pupil, $\Delta s'$ is the aperture aberration, and if Eq. (2) gives the magnification error

$$\Delta m = \frac{n \sin u}{n' \sin u} - m_o \tag{2}$$

compared with the magnification m_o on the axis, the condition is given by Eq. (3). The fulfillment of this

$$\frac{\Delta s'}{k'} - \frac{\Delta m}{m_o} = \text{constant} \tag{3}$$

condition gives equal quality for an object near the axis of a system with rotation symmetry. Corresponding conditions can be ascertained for the image of an off-axis element if all the asymmetry errors and deformation errors are balanced. *See* ABERRATION (OPTICS).

Resolution. Two points are resolved by an optical system if the two images lie apart. Photometric analysis of an image may indicate the existence of two object points even if their images overlap, but in such an analysis the illumination of the object, as well as the imagery, plays a role.

In interference experiments, it is found that the image of two self-luminous points (that is, two light sources that are sufficiently separated) is incoherent; that is, the intensities of the two beams simply add. If the two object points are illuminated by the same light source, however, the phase relation of the light at the two points has to be taken into consideration. This is of the greatest importance for microscopes and telescopes, which image very small or distant objects. In this case an artificial change of phase by phase plates and apodization may improve resolution. *See* DIFFRACTION; RESOLVING POWER (OPTICS).

Resolving power is not the only consideration in image formation. The eye recognizes only contrast differences, and therefore objects may not be discerned if the contrast difference is too small. Again, for the image of a point, or an object illuminated by a point light source, means can be found to change the apparent contrast, making it possible to discern biological objects, for example, having small differences of refractive index.

Image analysis. Methods have been suggested for obtaining information about optical images by sine-wave analysis. A sinusoidal test object is imaged by an optical system as a sinusoidal image, but altered in phase and amplitude. A large number of sinusoidal test objects with different frequencies (number of maxima per millimeter) are imaged, and the amplitude and phase are measured.

The curve of amplitude versus frequency gives a measure for resolving power and contrast as a function of the frequency of the test object, whereas the adjusted curve of phase versus frequency describes the lack of symmetry in the image. These amplitude-frequency curves can be measured as well as calculated from the spot diagrams, onto which the effects of diffraction can be superimposed if necessary.

Max Herzberger

Bibliography. B. D. Guenther, *Modern Optics*, 1990; M. Herzberger, *Modern Geometrical Optics*, 1958, reprint 1980; F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976; J. Meyer-Arendt, *Introduction to Classical and Modern Optics*, 4th ed., 1995; Optical Society of America, *Handbook of Optics*, 2 vols., 2d ed., 1995.

Optical information systems

Systems that use light to process information. Optical information systems or processors consist of one or several light sources; one- or two-dimensional planes of data such as film transparencies, various lenses, and other optical components; and detectors. These elements can be arranged in various configurations to achieve different data-processing functions. As light passes through various data planes, the light distribution is spatially modulated proportional to the information present in each plane. This modulation occurs in parallel in one or two dimensions, and the processing is performed at the speed of light. Optical processors offer various advantages compared to other technologies: data travels at the speed of light; all data in one-dimensional and two-dimensional arrays are operated on in parallel; multiple planes of data can be processed in parallel by various multiplexing schemes; it is possible to have large numbers of interconnections with no interaction (which is not possible with electrical connections); and power dissipation is less and size and weight can be less for optical processors than for their electronic counterparts. *See* CONCURRENT PROCESSING; MULTIPLEXING AND MULTIPLE ACCESS.

In practice, the processing speed is limited by the rate at which data can be introduced into the system and the rate at which processed data (produced on output detector arrays) can be analyzed. The reusable real-time spatial light modulators used to produce new input data, filters, interconnections, and so forth, are the major components required for these optical information-processing systems to realize their full potential. Spatial light modulators convert electrical input data into a form suitable for spatially modulating input light, or react to an optical input and generate a different optical output. The manipulation of the light passing through the system is controlled by spatial light modulators, lenses, holographic optical elements, computer-generated holograms, or fiber optics. Four major application areas are image processing, signal processing, computing and interconnections, and neural networks. *See* HOLOGRAPHY; OPTICAL FIBERS; OPTICAL MODULATORS.

Components. Inexpensive liquid-crystal television displays (when properly modified) can serve as two-dimensional spatial light modulators. Several commercially available two-dimensional spatial light modulators use magnetooptic techniques, deformable mirrors, and ferroelectric liquid crystals.

However, these devices are binary (except for the deformable mirror device). Optically addressed volume-storage optical media and optically addressed bistable devices and structures that function as optical transistors are not yet competitive with digital technology in size and power dissipation, but research advances may yield such competitive elements. Laser-diode array technology for communications and parallel recording of multiple channels of data on disks, plus optical read/write alterable storage media have greatly advanced the possibilities in this area. Acoustooptic devices represent one of the most mature and robust optical transducer technologies, and architectures that utilize these one-dimensional spatial light modulators to implement various two-dimensional processing functions offer the promise of hardware implementation in the near future. *See* ACOUSTOOPTICS; COMPUTER STORAGE TECHNOLOGY; FERROELECTRICS; LASER; LIQUID CRYSTALS; MAGNETOOPTICS; OPTICAL RECORDING.

Optical image processing. The basic optical image-processing functions will be discussed (**Fig. 1**). An image $g(x,y)$ is placed in a plane, P_1, and illuminated with laser light. The light distribution incident on a second plane, P_2, is the two-dimensional Fourier transform $G(u,v)$ of the input image $g(x,y)$. This Fourier transform distribution is a representation of the spatial frequencies (u,v) present in the input image. Lower spatial frequencies (corresponding to larger input shapes) lie closer to the center of P_2, and higher spatial frequencies (corresponding to smaller input image regions) lie further from the center of P_2. The orientation of each input object is reflected in the angular location of its corresponding spatial frequency distribution in P_2. Detectors placed in P_2 can feed this Fourier-coefficient information to a digital postprocessor for subsequent analysis and classification of the type of input object and its size and orientation. This is one form of optical feature extraction for pattern recognition and precise measurement of object size and orientation. *See* FOURIER SERIES AND TRANSFORMS.

Other architectures can optically compute other features of an input object such as its geometrical moments, its chord histogram distribution (the length and angle of all chords that define the object), or wedge and ring samples of the Fourier transform. A digital postprocessor can analyze these various features to determine the final class and orientation of the input object. Such feature extraction systems are the simplest form of optical pattern recognition image processors. Systems to compute optical wedge-ring Fourier transform samples have been tested and demonstrated on a wide variety of product inspection problems and applications. A Hough transform optical processor, which determines the position, orientation, and length of all lines in an input image has also been developed for product inspection. In this system, a computer-generated hologram consisting of N pairs of orthogonal cylindrical lenses is placed behind an input spatial light modulator on which an image of the product is placed. The output consists of N slices of the Hough transform at N different angles.

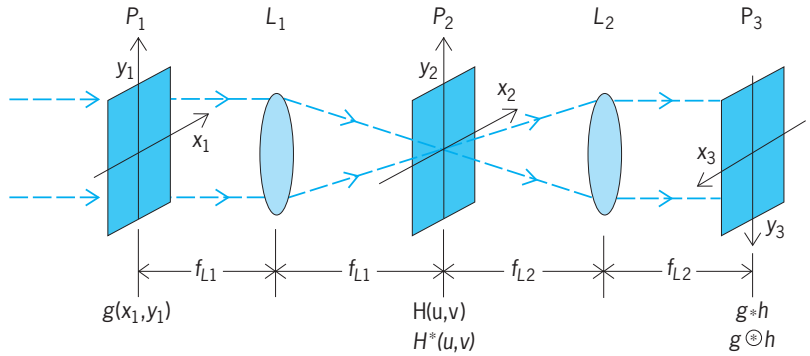

Fig. 1. Optical Fourier transform and correlation image processor. Lenses L_1 and L_2 have focal lengths f_{L1} and f_{L2}.

However, the full optical system of Fig. 1 can also perform the correlation operation. In this case, the transmittance of P_2 is made proportional to $H^*(u,v)$, the conjugate Fourier transform of a reference function $h(x,y)$. The $H^*(u,v)$ filter function at P_2 is formed by holographic techniques. Then, the P_3 pattern is the correlation of the two space functions g and h, written $g \circledast h$. The correlation of two functions is equivalent to translating h spatially over g and, for each translation, forming the product and sum of h and the associated region of g. The regions of an input scene g that most closely match the reference function h being searched for yield the largest output. This template matching operation can thus locate the presence and positions of reference objects in an input image. This object pattern recognition operation is achieved with no moving parts and is quite effective for extracting objects in clutter or high noise and for locating multiple occurrences of an object. Advanced filter synthesis techniques allow such systems to operate independent of scale, rotation, and other geometrical distortions between the input and reference object. Character recognition, robotics, missile guidance, and industrial inspection are the major applications that have been pursued for such architectures. If $H(u,v)$ rather than $H^*(u,v)$ is recorded at P_2, the P_3 output is an image, $g*h$, that is a filtered version of the input. The filter function of the system can be controlled by spatially varying the contents of P_2 to achieve image enhancement and restoration of blurred and degraded images. *See* CHARACTER RECOGNITION; COMPUTER VISION; GUIDANCE SYSTEMS; IMAGE PROCESSING; ROBOTICS.

A major shortcoming of prior optical image processors was the need for different architectures or systems to implement different functions and the lack of attention given to the role for optics in low-level iconic (pixel-based) image-processing functions, which are the more intensive image-processing functions required. Techniques exist to optically implement all basic morphological image-processing functions on the optical correlator in Fig. 1 (using a fixed set of simple structuring-element filters at P_2) and different thresholding operations at P_1 and

P_3. The P_3 output is the convolution of the input image and the structuring element at P_2. With a low or high P_3 threshold, the output is the erosion or dilation of the input image. Intermediate thresholds are more useful and produce rank-order filters. Sequences of these operations yield opening or closing functions. These are useful for filling in regions or gaps in an image, and removing peninsulas, salt-and-pepper noise, and other undesirable features. The size of the structuring element and the number of times it is applied determine the size of the image gaps filled in or the size of image regions removed. Alternatively, the skeleton, median filter, or various edge-enhanced versions of the image can be formed. These operations constitute all of the basic morphological image-processing functions. Optical processors that use them can operate at all levels of computer vision: low-level (iconic processors performing morphological functions), medium-level (optical processors for feature extraction), and high-level (correlators and neural networks).

Other new optical filters produce wavelet and Gabor transforms for object detection. Thus, the repertoire of operations achievable on the same optical processor is quite significant.

Optical signal processing. Signal processing systems are used in radar, sonar, electronic warfare, and communications to determine the frequency and direction of arrival of input signals, and for correlation applications. These systems generally employ acoustooptic transducers to input electrical signals to the optical processor. The electronic input to an acoustooptic cell is converted to a sound wave which travels the length of the cell. If the cell is illuminated with light, the amplitude of the light leaving the cell will be proportional to the strength of the input signal, and the angle at which the light leaves will be proportional to the frequency of the input signal. Since the device and system are linear, if N input signals are simultaneously present, N light waves leave the cell at N angles and strengths. A lens behind the cell focuses each light wave to a different spatial location in an output detector plane. Thus, an optical spectrum analyzer results (**Fig. 2**). Compact systems of this type exist, and bandwidths in excess of 8 GHz can be achieved and up to 1000 signals can be simultaneously processed. If a multichannel acoustooptic cell is used with input signals from different antennas or antenna elements, the two-dimensional Fourier transform of its output provides a two-dimensional display of the simultaneous frequency and direction of arrival of all received signals within the bandwidth of the cell. The angle at which light is deflected can be varied by varying the frequency of the input signal to the acoustooptic cell; a beam deflector or scanner based on this principle is used in some xerographic reproduction systems and laser printers. Other architectures using multiple cells can correlate two or more signals. These systems are used for simultaneous range and Doppler processing, for demodulation and synchronization of signals for communication applications, and for production of adaptive processors that can null interference signals. *See* ELECTRONIC WARFARE; RADAR; SONAR; SPECTRUM ANALYZER.

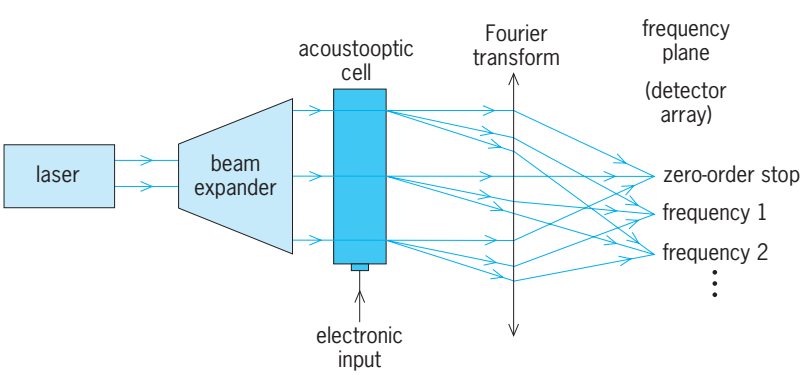

Fig. 2. Acoustooptic signal spectrum analyzer.

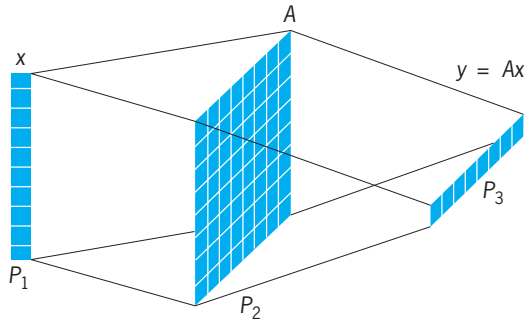

Fig. 3. Optical matrix vector, crossbar, or neural net processor.

Optical computing and interconnections. The term optical computing has had different meanings as optical processing has evolved. It has been used to describe optical array processors that operate on one-dimensional or two-dimensional arrays of data (vectors and matrices) to perform various functions in linear algebra. In the basic version of such a system (**Fig. 3**), the light leaving point modulators (laser diodes or light-emitting diodes) at a vector, P_1, is expanded to uniformly illuminate a two-dimensional data array (matrix), P_2. The light leaving P_2 is then integrated onto a linear detector array, P_3, whose output is the matrix vector product of the P_2 and P_1 data. Implementations of this basic architecture in integrated optics are being studied. *See* INTEGRATED OPTICS; LIGHT-EMITTING DIODE; LINEAR ALGEBRA; MATRIX THEORY.

With N inputs at P_1 and M outputs at P_3, the mask at P_2 can be used to interconnect any of the N inputs to M outputs. One-to-many and many-to-one interconnections are allowed, and a nonblocking optical crossbar switch results. Many optical interconnection systems using fiber optics, holograms, waveguides, and volume materials are rapidly maturing.

Most present optical computing work concerns optical techniques to achieve general-purpose logic and numeric functions. These are generally referred to as optical-digital circuits. There are three basic approaches.

In symbolic substitution, input binary 0 and 1 data are represented by different symbols. Input digit patterns (for example, 00, 01, 10, and 11) are recognized, and different patterns are substituted for each. The pattern substituted determines the operation that the system performs. Substitution rules have been devised to allow all logic and numeric functions to be realized. Different encoding methods allow improved realization with fewer carries required. A variety of optical realizations exist.

A second type of general-purpose optical computer uses acoustooptic devices and optical AND and OR operations. (The AND or OR of multiple optical inputs occurs when several input light beams are incident on a detector or optical device with a high or low threshold.) From a basic parallel optical A-O-I (AND-OR-INVERT) unit, a general optical arithmetic logic unit (ALU) can be fabricated. *See* COMPUTER SYSTEMS ARCHITECTURE; DIGITAL COMPUTER; LOGIC CIRCUITS.

The third major approach uses optical nonlinear or optical bistable devices to develop optical switches (optical transistors). In one use of the device, the wavelength of the input light is offset slightly from the value that gives maximum transmission. However, when the intensity of the input light increases, the transmittance curve for the device shifts to allow transmission of the input wavelength. A hysteresis effect can occur, and the input wavelength can be shifted by different amounts from resonance to produce a family of optical circuits. *See* NONLINEAR OPTICAL DEVICES.

Optical neural networks. The ability of many optical beams to cross with no interference between them, one beam to be broadcast to several locations, or several beams to combine at one location, is the basis for optical interconnections. These same properties are being pursued extensively to produce optical neural networks. In these cases, input points of light (as in Fig. 3) represent neurons, the amounts of output light represent neurons strengths, matrix masks (such as at P_2 of Fig. 3) represent sets of weights or interconnections between neurons, and the P_3 outputs in Fig. 3 represent a second layer of neurons. A cascade of two (or more) such matrix-vector systems thus represents a neural network. With various nonlinear elements used at the different neuron layers, volume materials used for the matrix connections, and feedback between neuron layers, powerful artificial optical neuron networks are possible. *See* NEURAL NETWORK.

David Casasent

Bibliography. N. J. Berg and J. M. Pellegrino, *Acousto-Optic Signal Processing*: *Theory and Implementation*, 2d ed., 1994; D. G. Feitelson, *Optical Computing*: *A Survey for Computer Scientists*, 1988, reprint 1992; H. S. Hinton et al. (eds.), Special issue on optical computing, *Appl. Opt.*, vol. 33, no. 8, March 10, 1994; Special issue on optical interconnections, *Opt. Eng.*, vol. 21, no. 10, October 1986; Special issue on spatial light modulators, *Appl. Opt.*, vol. 28, no. 22, November 15, 1989; H. Szu (ed.), Special issue on adaptive wavelet transforms, *Opt. Eng.*, vol. 21, no. 10, October 1986.

Optical isolator

A device that is interposed between two systems to prevent one of them from having undesired effects on the other, while transmitting desired signals between the systems by optical means. Optical isolators are used for both electrical systems and optical systems such as lasers.

Isolators for Electrical Systems

An optical isolator for electrical systems is a very small four-terminal electronic circuit element that includes in an integral package a light emitter, a light detector, and, in some devices, solid-state electronic circuits. The emitting and detecting devices are so positioned that the majority of the emission from the emitter is optically coupled to the light-sensitive area of the detector. The device is also known as an optoisolator, optical-coupled isolator, and optocoupler. The device is housed in an integral opaque package so that the only optical emission impinging on the detector is that produced by the emitter. This configuration of components can perform as a solid-state electronic transformer or relay, since an electronic input signal causes an electronic output signal without any electrical connection between the input and the output terminals.

Optical isolators are used in electrical systems to protect humans or machines when high-voltage or high-power equipment is being controlled. In addition, optical isolators are used in electronic circuit design in situations where two circuits have large voltage differences between them and yet it is necessary to transfer small electrical signals between them without changing the basic voltage level of either.

Before optical isolators were developed, this circuit function was performed by such devices as isolation transformers or signal relays. The optical isolator provides a number of advantages over such devices in that it is much smaller, much faster, has no contact bounce (unlike a relay), has no inductance (unlike a transformer), and provides a very high voltage isolation between the input and output circuits.

In order to discuss applications of the optical isolator, it is necessary to identify the different types of optical isolators that can be produced through the use of different combinations of light emitters and detectors.

Optical emitters. The optical emitter most commonly selected for use in optical isolators is the gallium arsenide light-emitting diode (LED), which emits in the infrared or near-infrared regions of the spectrum. The typical wavelength of the emission is 850 nanometers. The LED is extremely small (about the size of a transistor chip) and provides extremely fast infrared light pulses. Photodetectors fabricated from silicon are particularly sensitive to the wavelength of light emitted by the gallium arsenide LED, so that the transfer of a signal from the input to the output of the optical isolator is particularly efficient.

The other input devices commonly used are gas-discharge and incandescent lamps. These emitters tend to be somewhat larger and slower than LEDs,

but are used when their particular electronic characteristics are desired, and in conjunction with detectors that have a spectral peak in the visible portion of the spectrum. *See* INCANDESCENT LAMP; LIGHT-EMITTING DIODE; SEMICONDUCTOR DIODE; VAPOR LAMP.

Optical detectors. The light detectors that are used in the construction of optical isolators include light-dependent resistors (such as photocells), light-sensitive devices that generate a voltage without any electrical input (such as photovoltaic devices), light-sensitive devices that switch from one state to another (such as photothyristors), and light-sensitive devices that modify a voltage or current (such as phototransistors, photodiodes, and photodetector-amplifier combinations). *See* OPTICAL DETECTORS; PHOTOCONDUCTIVE CELL; PHOTODIODE; PHOTOELECTRIC DEVICES; PHOTOTRANSISTOR; PHOTOVOLTAIC CELL.

Emitter–detector combinations. Having such a wide variety of emitter and detector combinations available means that circuitry and system designers can use them over a wide range of applications. For example, the combination of a photocell and an incandescent lamp produces a device with entirely different characteristics from those of the combination of an LED and a phototransistor. Each combination is described below and its primary application identified.

Incandescentlamp–photocell optical isolators. In these devices the resistance of the photocell varies from about 10 to 1000 megohms, as the input light is varied from off to maximum brightness. If the input changes instantaneously, the output resistance change occurs in 20 milliseconds (typical). If the input voltage changes gradually, the resistance of the photocell varies somewhat proportionally to the input voltage. As a result of this variable-resistance characteristic, the device whose schematic is shown in **Fig. 1** can be used as a remote-controlled rheostat where there is no electrical connection between the control wire and the circuit being controlled. In applications where signals are required for two separate circuits, without an electrical connection between any of the three circuits, devices are used with two photocells built into the case of the optical isolators.

Neon lamp–photocell combinations. These devices have output characteristics similar to the incandescent lamp–photocell devices. However, the input characteristics vary dramatically, because the neon lamp does not emit light (ignite) until the input voltage reaches about 70 V. Also, the current after ignition is very small. These devices are employed where large input voltage swings are typical and where long life is required. *See* NEON GLOW LAMP.

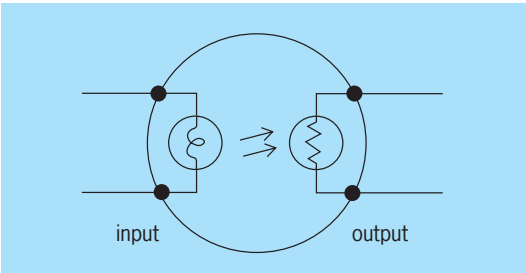

Fig. 1. Schematic diagram of incandescent lamp–photocell optical isolator.

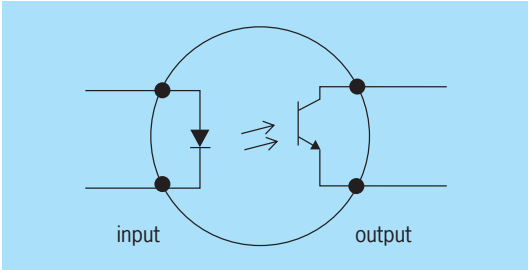

Fig. 2. Schematic diagram of LED–phototransistor optical isolator.

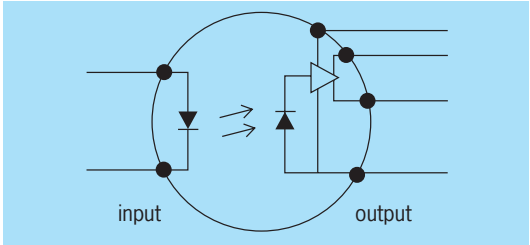

Fig. 3. Schematic diagram of LED–photodiode–integrated-circuit optical isolator.

LED—silicon detector combinations. All of the optical isolators that employ silicon photodetectors have LED emitters. The most common photodetector is the phototransistor. A schematic diagram of a type of LED–phototransistor optical isolator is shown in **Fig. 2**. In such devices the light from the LED causes the phototransistor current to vary as a function of the amount of light impinging on the photosensitive area of the phototransistor. The devices can isolate circuits that differ by as much as 5000 V. The devices are very fast, with the output current changing in some devices in as little as 10–20 nanoseconds after the occurrence of an input pulse. This fast response time means that these devices are particularly useful in computer circuit applications where such fast pulses are common. Not only do these devices provide high voltage isolation between circuits, but because the phototransistor has built-in gain, the detected signal is amplified. These devices are also immune to certain types of noise and are extremely small, the package being just larger than a typical transistor package. The devices can be used in most circuits that require an isolation transformer, and provide the additional advantages of higher speed and higher voltage ratings. Such devices also can perform as a solid-state relay (no mechanical wear and so no contact bounce), as a high-speed chopper, and as a pulse amplifier.

A similar type of optical isolator that produces a higher output amplitude for the same input employs a photo-Darlington sensor and output device. The photo-Darlington is in reality two phototransistors

that are integrally connected. Although it does have increased amplification, its speed is much lower.

LED–photodetector–amplifier. In some optical isolators, integrated circuits are also included in the package to provide specific output characteristics. For example, one optical isolator includes an integrated circuit amplifier so that the optical isolator can perform like a broadband pulse transformer that is compatible with diode-transistor-logic (DTL) computer circuitry as well as provide a frequency response to zero frequency. A schematic diagram of an LED–photodiode–integrated-circuit optical isolator is shown in **Fig. 3**. *See* AMPLIFIER; CIRCUIT (ELECTRONICS); INTEGRATED CIRCUITS; LOGIC CIRCUITS.

Robert D. Compton

Isolators for Optical Systems

The need for optical isolation has broadened considerably since the advent of lasers. It is often necessary to prevent light from reentering the laser, irrespective of any electrical consideration. One example is a small laser followed by high-power laser amplifiers. If the powerful amplified light reenters the small (master oscillator) laser, it can destroy it. Another example is a frequency-stabilized laser, whose oscillation frequency is perturbed by reentering (injected signal) light. The general configuration of an optical isolator for optical systems is shown in **Fig. 4**.

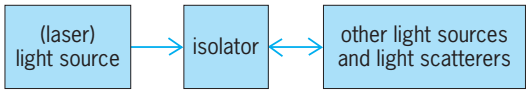

Fig. 4. General configuration of an optical isolator for optical systems.

Quarter-wave polarizer. A polarizer–plus–quarter-wave-plate isolator prevents laser light from reentering the laser when the light is scattered back by specular reflectors. This device cannot ensure isolation if there is diffuse reflection or if polarization-altering (birefringent) optics are encountered. Another limitation of this isolator is that the transmitted light is circularly polarized. The mechanism of operation is illustrated in **Fig. 5**, which shows what happens when light is reflected, double-passing the isolator. (Light reflected back through the system shows as if it were traversing a second system whose components are in reverse order.) *See* BIREFRINGENCE; POLARIZED LIGHT.

Faraday rotator. The Faraday effect is found in many solids, liquids, and gases. When plane polarized light is sent through the material in a direction parallel to a strong magnetic field, the plane of polarization is rotated (**Fig. 6**). The Faraday mechanism differs from that of a wave plate in that the plane of polarization is rotated through an angle θ that is proportional to the path length l. That is, there is no fast or slow axis. This sort of corkscrew effect is called magnetically induced optical activity, in contrast to birefringence in the case of the wave plate. Because the corkscrew rotation has a particular handedness, specularly reflected (double-passed) light receives an additional rotation, which can be stopped by the entrance polarizer. Specifically, if the Faraday rotation

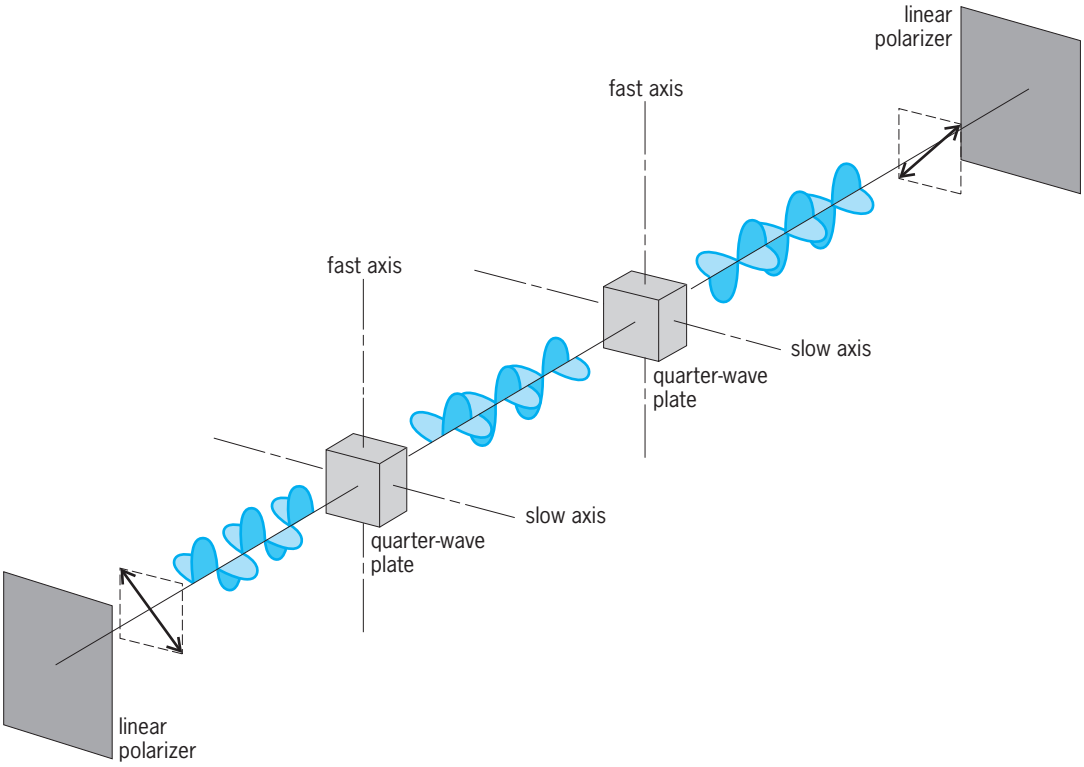

Fig. 5. Device consisting of quarter-wave plate and linear polarizer, showing how it serves as an isolator against specular reflections when they double-pass the system. Light reflected back through the system is shown as if it were traversing a second system whose components are in reverse order.

angle θ equals 45°, then such reflected, double-passed light will be rotated 90° with respect to an entrance polarizer and will not be passed. *See* FARADAY EFFECT.

In contrast to the quarter-wave polarizer isolator, the Faraday isolator can provide truly one-way transmission irrespective of polarization changes from the exit side if an exit polarizer (which passes light that has undergone the Faraday rotation after passing though the entrance polarizer) is used in addition to the entrance polarizer. For example, it isolates against diffuse reflections and any light source on the exit side.

Acoustooptic deflector. The isolation properties of an acoustooptic deflector are based on the fact that light deflected by it is shifted in frequency by an amount equal to the acoustic frequency (**Fig. 7**). The reflected beam, passing through the deflector a second time, is again shifted in frequency by the same amount and in the same sense if the deflector is operated in the Bragg mode. Hence, the reflected light that is returned to the laser is shifted in frequency by an amount $2f$, where f is the frequency of the acoustic wave. Provided the frequency of the light returned to the laser is not close to any resonant frequency of the laser cavity, it will not perturb the laser and will simply be reflected from the output mirror. *See* ACOUSTOOPTICS.

For general use as isolators, acoustooptic deflectors have the limitation that they are really only frequency shifters. On the other hand, acoustooptic deflectors are quite insensitive to the polarization state of the light, which can be a serious problem with quarter-wave-polarizer and Faraday isolators. A combination of acoustooptic and Faraday isolators has been shown to avoid both these limitations.

Stephen F. Jacobs

Bibliography. C. J. Georgopoulos, *Fiber Optics and Optical Isolators*, 1982; A. K. Ghatak and K. Thayagarian, *Optical Electronics*, 1989; H. Lee, Optical isolator using acoustooptic and Faraday effects, *Appl. Opt.,* 26:969–970, 1987; A. Yariv, *Optical Electronics*, 5th ed., 1997.

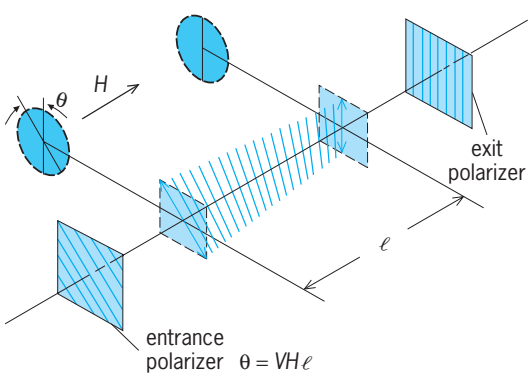

Fig. 6. Faraday effect. θ = angle of rotation, V = Verdet constant, H = magnetic field strength, l = length of medium traversed.

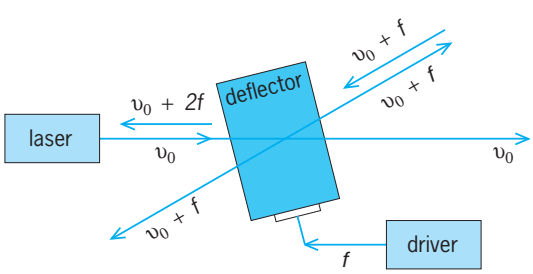

Fig. 7. Acoustooptic isolator, operated in the Bragg mode.

Optical materials

All substances used in the construction of devices or instruments whose function is to alter or control electromagnetic radiation in the ultraviolet, visible, or infrared spectral regions. Optical materials are fabricated into optical elements such as lenses, mirrors, windows, prisms, polarizers, detectors, and modulators. These materials serve to refract, reflect, transmit, disperse, polarize, detect, and transform light. The term "light" refers here not only to visible light but also to radiation in the adjoining ultraviolet and infrared spectral regions. At the microscopic level, atoms and their electronic configurations in the material interact with the electromagnetic radiation (photons) to determine the material's macroscopic optical properties such as transmission and refraction. These optical properties are functions of the wavelength of the incident light, the temperature of the material, the applied pressure on the material, and in certain instances the external electric and magnetic fields applied to the material. *See* ATOMIC STRUCTURE AND SPECTRA; DISPERSION (RADIATION); ELECTROMAGNETIC RADIATION; ELECTROOPTICS; INFRARED RADIATION; LENS (OPTICS); LIGHT; MAGNETOOPTICS; MIRROR OPTICS; OPTICAL DETECTORS; OPTICAL MODULATORS; OPTICAL PRISM; POLARIZED LIGHT; REFLECTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES; ULTRAVIOLET RADIATION.

There is a wide range of substances that are useful as optical materials. Most optical elements are fabricated from glass, crystalline materials, polymers, or plastic materials. In the choice of a material, the most important properties are often the degree of transparency and the refractive index, along with each property's spectral dependency. The uniformity of the material, the strength and hardness, temperature limits, hygroscopicity, chemical resistivity, and availability of suitable coatings may also need to be considered. *See* HARDNESS SCALES; STRENGTH OF MATERIALS.

Transmission. Throughout the visible region of the spectrum, there are many transparent materials such as glasses, crystals, polymers, and plastics with desirable properties; however, in the ultraviolet and infrared regions the transmission range of the individual material becomes of primary importance. The bars in **Fig. 1** represent the wavelength regions in which the optical materials transmit appreciably. More specifically, the bars indicate the wavelength range for which the external transmittance of the

given material is greater than 10% for a sample 1 mm thick. The wavelength scale is logarithmic.

Typical transmission curves (**Fig. 2**) give more detailed information concerning the wavelength dependence of the transmission of selected optical materials. The external transmittance is defined as the ratio of the intensity of light leaving a sample to the intensity of the light entering it. The external transmittance depends on both internal absorption and the reflection losses at the surfaces.

The need for new materials that will withstand high-power laser beams or transmit signals with minimum loss, as in optical fibers, indicates the significance of the transmission (or absorption) characteristics of a material. The equation below, called

$$I = I_o e^{-\alpha x}$$

Bouguer's law, gives the attenuation of a light beam propagating through a solid, where I_o is the beam intensity after entering the solid, I is the intensity at a distance x into the solid, and α is the absorption coefficient of the material. Absorption is the process in which the electromagnetic radiation energy is converted to internal energy (vibrational, rotational, and electronic potential energy) of the electrons, atoms, or ions in the lattice or ensemble of atoms within the material. The absorption loss and thus the transmission vary depending on the individual material, and are characterized by both wide and narrow bands of wavelengths over which the loss of energy occurs (Fig. 2). Reflection at the surface of a material can vary from about 4% (glass) to 36% (germanium). For normally incident light in air, reflection loss is given by $[(n-1)/(n+1)]^2$, where n is the refractive index of the material. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; CRYSTAL ABSORPTION SPECTRA.

The attenuation of the light beam as it traverses a solid can also result from scattering. Scattering varies with wavelength and results from a variety of sources that must be taken into account when choosing the proper material for fabrication into an optical element. Bouguer's law can be adjusted to include a scattering coefficient analogous to the absorption coefficient. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

Refractive index. For many optical elements such as lenses and prisms, the most important optical property is the refractive index, which is defined as the ratio of the speed of light in a vacuum to the speed of light in the material. The refractive index varies with wavelength; this variation is referred to as dispersion. The dispersion curves for selected optical materials are illustrated in **Fig. 3**, where the refractive index is plotted versus wavelength on a logarithmic scale. A wide range of values of refractive index exist for optical materials (see **table**) from as low as 1.3 to greater than 4.0, with optical glasses generally ranging from 1.5 to about 1.7.

The choice of an optical material for a particular application can be facilitated by a reference such as the table, which presents relevant physical properties for a variety of optical materials.

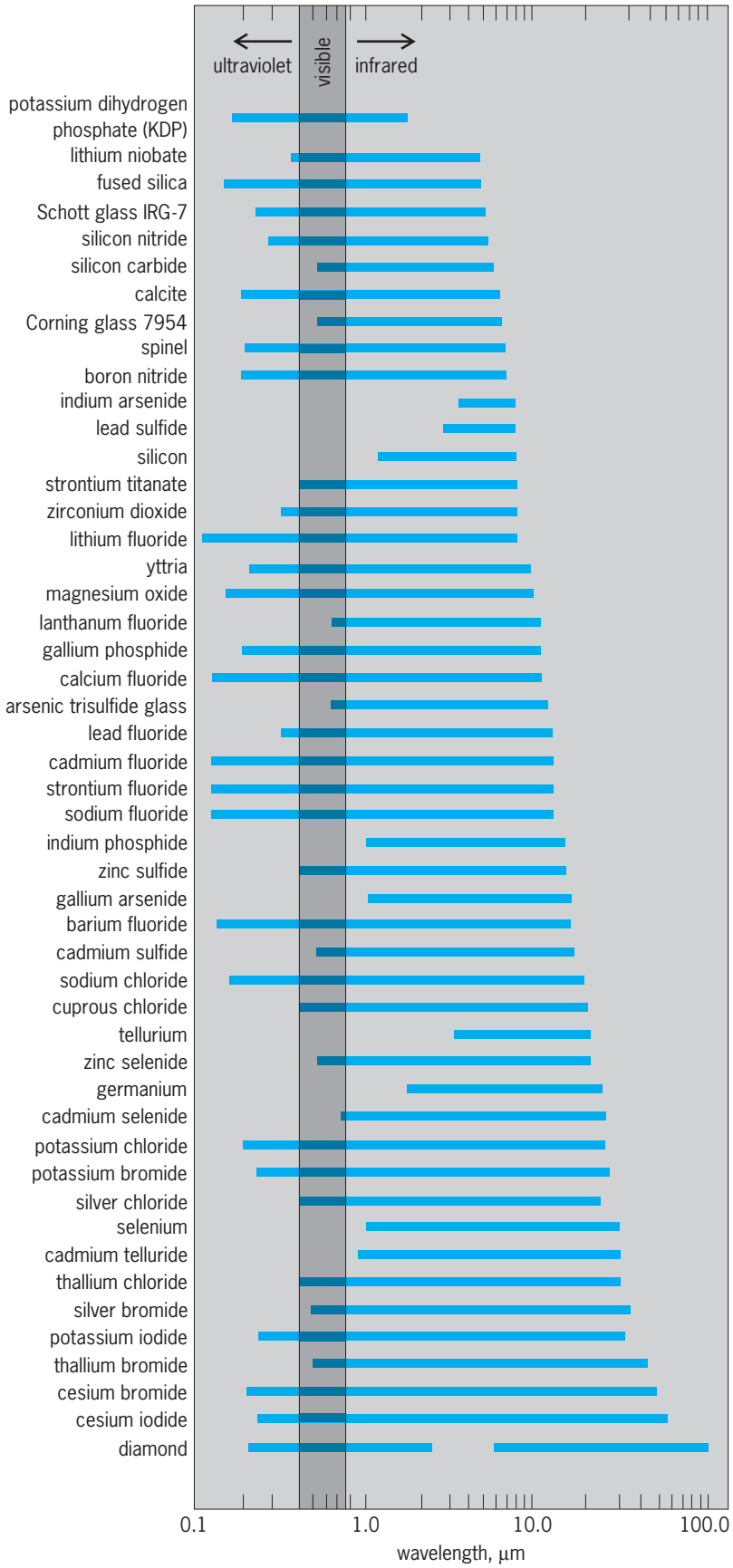

Fig. 1. Transmission range for a variety of optical materials. The bars indicate the wavelength region for which the external transmittance for a 1-mm-thick sample of the material is greater than 10% when measured at room temperature. The wavelength scale is logarithmic.

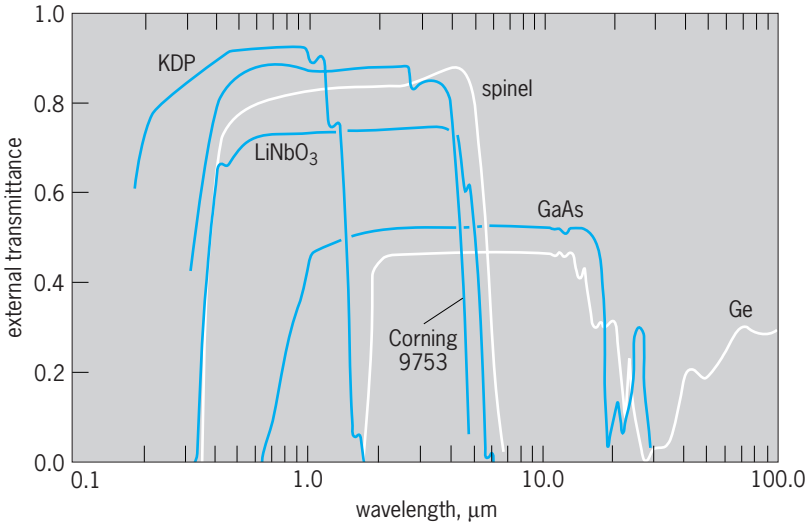

Fig. 2. Transmission versus wavelength for several optical materials. The wavelength scale is logarithmic.

Glass. The development of optical glass was due primarily to E. Abbe, C. Zeiss, and O. Schott. Glasses were the vanguard of optical materials because they transmit in the visible, provide a fairly wide range of refractive indices, can be easily shaped, and are relatively inexpensive. Early optical glasses were classified as either crown or flint. Crown glasses are composed mainly of silica, alkaline earth oxides, and alkali metal oxides. They exhibit low refractive index values and low dispersion. Flint glasses are composed typically of silica, alkali metal oxides, and lead oxides. They exhibit high refractive index values and high dispersion. Advances made by Schott in introducing the borosilicate, fluoride, and barium crown and flint glasses, and later by G. Morey in developing the rare-earth glasses, expanded the range of refractive indices and dispersive powers of optical glasses.

Hundreds of glass compositions are manufactured. Glass suppliers publish catalogs with comprehensive data on each glass. These catalogs contain extensive tables of the refractive index of the glass versus wavelength. The Abbe number indicates the degree of dispersion of the glass. This is the ratio $(n_d - 1)/(n_f - n_c)$, where the subscripts on the refractive index n indicate the wavelength at which it was measured; the subscripts d, c, and f refer to spectral lines at 589.3, 486.1 and 656.3 nanometers respectively. *See* CHROMATIC ABERRATION.

Glass technology provided the foundation for classical optical elements, such as lenses, prisms, and filters. Glasses developed for use in the visible region have internal transmittances of over 99% throughout the wavelength range of 380–780 nm. However, the silicate structure in glasses limits their transmission to about 2.5 μm in the infrared. Chalcogenide glasses, heavy-metal fluoride glasses, and heavy-metal oxide glasses extend this transmission to 8–12 μm. *See* COLOR FILTER.

Beginning with a breakthrough in the development of ultrapure glass in 1970, the process for manufacturing optical fibers changed from melting raw silica to forming the glass from vaporized chemicals. This new process is referred to as outside vapor deposition (OVD) or modified chemical vapor deposition (MOVD). These advances led to the present fiber-optic communication systems that operate in the near-infrared region with windows at wavelengths of 850, 1310, 1550, and 1625 nm. An advanced fiber-optic system, LEAF (Large Effective Area Fiber), was designed to minimize nonlinearities by spreading the optical power over large areas. *See* OPTICAL COMMUNICATIONS; OPTICAL FIBERS; VAPOR DEPOSITION.

The use of photolithography for printing integrated circuits has necessitated the improvement in the transmission of glasses for the ultraviolet region. Fused silica, which transmits to about 180 nm, is well suited for the lithography in the ultraviolet region. However, the crystalline material calcium fluoride, which transmits into the ultraviolet region to about 140 nm, outperforms any glass in printing microchips using fluorine excimer lasers. Corning 7940 and Suprasil are two commercially available fused-silica glasses designed specifically for deep-ultraviolet applications. They have uses as high-energy laser glasses, spacecraft window glasses, and mirror blanks for large astronomical mirrors. These and other fused-silica glasses are available for optical imaging in the ultraviolet region. Ultraviolet glasses combined with advanced ultraviolet laser sources and matched detectors have improved cancer detection by using laser-induced autofluorescence. *See* FLUORESCENCE; INTEGRATED CIRCUITS; TELESCOPE.

Glasses also serve an important function as host materials for lasing ions. Lawrence Livermore National Laboratories has created the largest combined system of glass lasers, known as NOVA. These lasers contain a series of Schott (LG-750 and LG-760) phosphate laser glass amplifiers. *See* GLASS; LASER.

Plastics. The need for an inexpensive, unbreakable lens that could be easily mass-produced precipitated the introduction of plastic optics in the mid-1930s. Although the variety of plastics suitable for

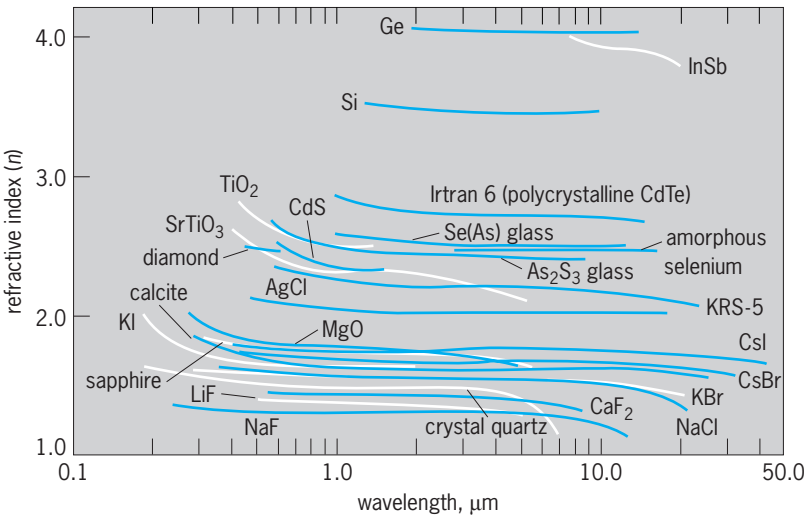

Fig. 3. Refractive index versus wavelength for several optical materials. The wavelength scale is logarithmic.

precision optics is limited compared to glass or crystalline materials, plastics are often preferred when difficult or unusual shapes, lightweight elements, or economical mass-production techniques are required.

The softness, inhomogeneity, and susceptibility to abrasion intrinsic to plastics often restrict their application. Haze (which is the light scattering due to microscopic defects) and birefringence (resulting from stresses) are inherent to plastics. Plastics also exhibit large variations in the refractive index with changes in temperature. Shrinkage resulting during the processing must be considered. *See* BIREFRINGENCE; PHOTOELASTICITY.

The refractive index of most usable plastics is from 1.45 to 1.60. Acrylic (methyl methacrylate) has an index of 1.49 and polystyrene an index of 1.59. An acrylic and polystyrene combination can be used to produce achromatic plastic lenses. Many other polymers, such as polyethylene, can be used in optical elements, along with the optical adhesives and epoxies. *See* POLYACRYLATE RESIN; POLYETHER RESINS; POLYOLEFIN RESINS; POLYSTYRENE RESIN.

Two of the more common plastic optical elements are the Fresnel lenses found in overhead projectors and the dichroic materials used in polarizing sunglasses. The need for mass-produced optical devices such as disposable contact lenses, compact discs, projection television systems, optical fibers, and electrooptical instruments has motivated the search for additional molecularly designed polymers.

Polymeric materials. Organic synthetic polymers are emerging as key materials for information technologies. In many cases, organic polymers compete with and even outperform inorganic materials. Polymers often have an advantage over inorganic materials because they can be designed and synthesized into compositions and architectures not possible with crystals, glasses, or plastics. They are manufactured to be durable, optically efficient, reliable, and inexpensive. Many uses of polymers in photonic and optoelectronic devices have emerged, including light-emitting diodes, liquid crystal-polymer photodetectors, polymer-dispersed liquid-crystal devices (for projection television), optical-fiber amplifiers doped with organic dyes (rhodamine), organic thin-film optics, and electrooptic modulators. Transparent optical fibers have been made from poly(methyl methacrylate). A dye-doped fiber of this polymer has been designed as an optical-fiber laser that is much less complicated and toxic than its liquid-dye laser counterpart. A fiber-optic polymer laser made of poly(methyl methacrylate-co-2-hydroxyethyl methacrylate) containing 0.1% of the rhodamine 6G dye has emitted yellow light pulses of about 0.5 millijoule, achieving high efficiency when pumped by a neodymium YAG laser. *See* LIGHT-EMITTING DIODE; LIQUID CRYSTALS.

Polymers can be produced that exhibit strong nonlinear optical properties that were earlier found only in inorganic crystalline materials. An especially promising class of nonlinear photonic materials is the photorefractive polymers. The photorefractive effect is a nonlinear optical phenomenon in which materials exposed to relatively low-powered laser beams exhibit large refractive index changes. This effect was discovered in 1966 in inorganic crystalline materials. These nonlinear crystals prove more difficult to grow and develop than synthetic organic polymers. In 1991 the photorefractive effect was first observed in an organic material, the epoxy polymer bis-α-NPDA, made photoconductive by doping it with the hole transport agent DEH. The application of the photorefractive effect in polymers is expected to result in a leap forward in high-density information storage and high-speed optical processing. Other promising applications are optical encoding, dynamic real-time holography, image enhancement, phase conjugation, and advanced flat-screen television receivers and computer displays. *See* ELECTRONIC DISPLAY; FLAT-PANEL DISPLAY DEVICE; HOLOGRAPHY; IMAGE PROCESSING; NONLINEAR OPTICS; OPTICAL INFORMATION SYSTEMS; OPTICAL PHASE CONJUGATION; POLYMER.

Crystalline materials. Although most of the early improvements in optical devices were due to advancements in the production of glasses, the crystalline state has taken on increasing importance. Historically, the naturally occurring crystals such as rock salt, quartz, and fluorite plus suitable detectors permitted the first extension of visible optical techniques to harness the invisible ultraviolet and infrared rays.

In the 1930s, synthetic crystal-growing techniques made more readily available single crystals such as lithium fluoride (of special value in the ultraviolet region, since it transmits at wavelengths down to about 120 nm), calcium fluoride, and potassium bromide (useful as a prism at wavelengths up to about 25 μm in the infrared). Many alkali-halide crystals (grown synthetically by the Czochralski technique) are important because they transmit into the far-infrared. The cubic symmetry of the alkali halide crystals results in isotropic optical and physical properties. Mixed crystals made of thallium bromide and thallium iodide called KRS-5 were grown in the late 1940s. *See* CRYSTAL GROWTH; CRYSTAL STRUCTURE; SINGLE CRYSTAL.

In the 1950s, following the invention of the transistor, germanium and silicon ushered in the use of semiconductors as infrared optical elements or detectors. Polycrystalline forms of these semiconductors could be fabricated into windows, prisms, lenses, and domes by casting, grinding, and polishing. Compound semiconductors such as gallium arsenide (GaAs, made of elements from the 13 and 15 columns of the periodic table), ternary compounds such as gallium aluminum arsenide ($Ga_{1-x}Al_xAs$), and quaternary compounds such as indium gallium arsenide phosphide (InGaAsP) now serve as lasers, light-emitting diodes, and photodetectors. *See* SEMICONDUCTOR.

In the 1960s the so-called polycrystalline compacts were developed. Pure powders of several infrared-transmitting materials were heated and compacted into blanks that could be fabricated into

desired infrared optical elements. An improvement over the hot-pressing techniques is the chemical vapor deposition (CVD) process. Polycrystalline zinc sulfide and zinc selenide manufactured by this process provide superior infrared materials available in diameters greater than 1 ft (0.3 m) and thicknesses of 1 in. (2.5 cm) for large windows in optical systems operating at wavelengths from 8 to 12 μm.

Polycrystalline diamond films, or coatings, are produced by a wide variety of chemical vapor deposition techniques. These techniques require high temperatures that preclude substrate materials with melting points below 1000–1400 K (1340–2060°F); thus diamond coatings on plastics are not possible. The most used substrate has been single-crystal silicon. Diamond's extreme properties, high thermal conductivity, low electrical conductivity, low thermal expansion, extreme hardness, and large range of optical transparency make it suitable for diverse applications. Diamond produced by chemical vapor deposition and single-crystal diamond (produced in the laboratory) share these benefits. *See* DIAMOND.

Single crystals are indispensable for transforming, amplifying, and modulating light. Birefringent crystals serve as retarders, or wave plates, which are used to convert the polarization state of the light. In many cases, it is desirable that the crystals not only be birefringent, but also behave nonlinearly when exposed to very large fields such as those generated by intense laser beams. A few examples of such nonlinear crystals are ammonium dihydrogen phosphate (ADP), potassium dihydrogen phosphate (KDP), beta barium borate (BBO), lithium borate (LBO), and potassium titanyl phosphate (KTP) [see table]. *See* CRYSTAL OPTICS.

Nonlinear optics continues at asteady pace with

Physical properties of optical materials*

Material	Symbol	Refractive index (wavelength = 500 nm)	Density, g/cm^3	Hardness: Knoop, kg/mm^2; or Mohs (M)	Solubility, g/100 g H_2O	Specific heat, cal/(g-K)†	Melting or softening temp., K	Thermal conductivity, W/(m · K)‡	Linear expansion coefficent, 10^{-6}/K‡	Young's modulus,‡ GPa
Crystalline materials										
Germanium	Ge	4	5.33	800	Insoluble	0.074	1210	59	6.1	102.66
Lithium fluoride	LiF	1.394	3.5	100	0.27	0.37	1140	11.3	34.4	64.77
Magnesium fluoride	MgF_2	1.39	3.18	415	Insoluble	0.24	1528	21	14(P), 8.9(S)	138.5
Sodium chloride	NaCl	1.53	2.17	15.2	35.7	0.2	1070	6.5	40	39.96
Zinc sulfide	ZnS	2.42	4.08	230	Insoluble	0.112	2100	17	6.6	74.5
Zinc selenide	ZnSe	2.43	5.42	137	0.001	0.0090	1790	19	7	70.97
Barium titanate	$BaTiO_3$	—	5.9	200–580	—	0.103	1870	1.34	19	33.76
Cesium iodide	CsI	1.75	4.51	—	44	0.048	894	1.1	48.3	5.3
Diamond	C	2.4	3.51	5700–10400	Insoluble	0.124	3770	2600	0.8	1050
Lanthanum fluoride	LaF_3	1.6	5.94	4.5(M)	Insoluble	0.121	1766	5.1	11(P), 17(S)	—
Magnesia	MgO	1.74	3.585	910	Insoluble	0.24	3053	40.6	8	—
Potassium chloride	KCl	1.49	1.98	7.2	34.7	0.162	1050	6.7	36.6	29.63
Sapphire	Al_2O_3	1.77	3.98	1370	Insoluble	0.18	2300	33.0	8(P) 5(S)	335
Crystal quartz	SiO_2	1.55	2.65	741	Insoluble	0.17	1740	10.7(P), 6.2(S)	8.0(P), 13.4(S)	97.2(P), 76.5(S)
Barium fluoride	BaF_2	1.47	4.89	0.82	0.17	0.096	1550	11.7	19.9	53.05
Calcium fluoride	CaF_2	1.43	3.18	140	0.0017	0.204	1630	10	18.9	75.79
Cadmium telluride	CdTe	2.69	6.2	56	Insoluble	0.056	1320	6.3	5.9	36.52
Calcite	$CaCO_3$	n_o = 1.665, n_e = 1.490	2.710	3(M)	0.0014	0.203	1612	5.526(P), 4.646(S)	25(P), −5.8(S)	72.35(P), 88.19(S)
Cuprous chloride	CuCl	2.0	4.14	2–2.5(M)	0.0061	—	695	—	10	—
Gallium phosphide	GaP	3.65	4.13	845	Insoluble	0.2	1623	54	4.7	102.6
Indium arsenide	InAs	4.5	—	330	Insoluble	0.06	1215	50	5.3	—
Lead fluoride	PbF_2	1.78	8.24	200	0.064	0.085	1100	—	29	—
Lead sulfide	PbS	4.3 at 3 μm	7.5	—	—	0.050	1387	0.67	18	—
Silicon carbide	SiC	2.68	3.217	2130–2755	Insoluble	0.165	3000	490	2.8	386
Selenium	Se	2.83	4.82	2.6(M)	Insoluble	0.077	490	1.3	48.7	—
Silicon	Si	3.45	2.329	1100	Insoluble	0.18	1690	163	2.6	130.91
Infrared glasses										
Arsenic trisulfide glass	As_2S_3	2.7	3.43	109	Insoluble	0.109	483	0.1674	24.62	15.85
Germanate glass (Schott IRG2)	—	1.9	5.00	481	0.012	0.108	—	0.91	8.8	95.9
Chalcogenide glass (Schott IRG 100)	—	2.73	4.67	150	—	—	624	0.3	15	21
Lead silicate glass (Schott IRG7)	—	1.57	3.06	379	0.171	0.151	—	0.73	9.6	59.7
Ultra-low-expansion (ULE) titanium silicate (Corning 7971)	—	1.484	2.205	459	Insoluble	0.183	1763	1.31	0.015	67.52
Germanate glass (Corning 9754)	—	1.67	3.581	560	—	0.13	1147	1.0	6.2	84.1
Phosphate laser glass (Schott LG750)	—	—	—	—	—	—	—	0.52	13.2	—

Physical properties of optical materials* (*cont.*)

Material	Symbol	Refractive index (wavelength = 500 nm)	Density, g/cm^3	Hardness: Knoop, kg/mm^2; or Mohs (M)	Solubility, g/100 g H_2O	Specific heat, cal/(g-K)†	Melting or softening temp., K	Thermal conductivity, W/(m · K)‡	Linear expansion coefficent, 10^{-6}/K‡	Young's modulus,‡ GPa
Ultraviolet glasses										
Fused silica	SiO_2	1.43	2.203	461	Insoluble	0.22	1448	1.38	0.55	73.1
Fused silica (Corning 7940)	—	1.46	2.202	500	—	0.177	1858	1.38	0.52	73
Nonlinear/photorefractive optical materials										
Ammonium dihydrogen phosphate (ADP)	$NH_4H_2PO_4$	1.51	1.803	—	22.7	—	4.63	.71162(P), 1.2558(S)	39.3	—
Lithium niobate	$LiNbO_3$	$n_o = 2.286$, $n_e = 2.203$ (0.6328 μm)	4.64	5(M)	Insoluble	—	1523	38	2.2(P), 2.0(S)	—
Potassium dihydrogen phosphate (KDP)	KH_2PO_4	1.5	2.338	—	33	0.21	525.6	1.3395	21.6	—
Yttrium vanadate	YVO_4	$n_o = 1.9929$, $n_e = 2.2154$ (0.6328 μm)	4.22	5(M)	Insoluble	—	—	5.23(P), 5.10(S)	11.37(P), 4.43(S)	—
Iron-doped lithium niobate	$Fe:LiNbO_3$	$n_o = 2.286$, $n_e = 2.203$ (0.6328 μm)	4.64	5(M)	Insoluble	—	1523	38	2.2(P), 2.0(S)	—
Silver gallium sulfide	$AgGaS_2$	$n_o = 2.4521$, $n_e = 2.3990$ (1.064 μm)	4.702	—	—	—	1270	0.015	12.5(P), −13.2(S)	—
Silver gallium selenide	$AgGaSe_2$	$n_o = 2.7010$, $n_e = 2.6792$ (1.064 μm)	5.700	—	—	—	1124	—	16.8(P), −7.8(S)	—
Magnesium oxide-doped lithium niobate	$MgO:LiNbO_3$	$n_o = 2.286$, $n_e = 2.203$ (0.6328 μm)	4.64	5(M)	Insoluble	—	1523	38	2.2(P), 2.0(S)	—
Potassium titanyl phosphate (KTP)	$KTiOPO_4$	$n_x = 1.78$, $n_y = 1.79$, $n_z = 1.89$	3.01	5(M)	Insoluble	0.1643	1445	13	—	—
Lithium triborate (LBO)	LiB_3O_5	$n_x = 1.58$, $n_y = 1.61$, $n_z = 1.62$,	2.47	6(M)	—	—	1107	—	—	—
Cadmium sulfide	CdS	2.6	4.82	122	Insoluble	0.0882	1773	15.91	2.1(P), 4(S)	—
Gallium arsenide	GaAs	3.35	5.32	731	Insoluble	0.076	1511	55	5.7	82.68
Beta-barium borate (BBO)	β-BaB_2O_4	$n_o = 1.6749$, $n_e = 1.5555$	3.85	4.5(M)	—	—	1363	1.6(P), 1.2(S)	36(P), 4(S)	—

*All values are given for temperatures near room temperature.
†1 cal = 4.18 J.
‡P denotes values measured parallel to the *c* axis of the crystal, and S denotes values measured perpendicular to the *c* axis.

new types of crystals becoming available on a regular basis. Important applications are second-harmonic generation (frequency doubling), parametric amplification (using two monochromatic waves to amplify a third), parametric oscillation (feedback is added to the parametric amplifier to create an oscillator), optical amplification, and optical phase conjugation.

A promising nonlinear phenomena is the photorefractive effect. This effect was first observed in 1966 in the crystal lithium niobate, and soon afterward a variety of nonlinear optical materials were discovered (see table), including potassium niobate, strontium barium niobate, bismuth silicon oxide, and barium sodium niobate. Many applications of the photorefractive effect were mentioned earlier in regard to polymers.

Other materials. Other optical materials that deserve mention are the liquid crystals used in displays as light valves, materials used in erasable optical discs for computers and in liquid cells (Kerr cells), laser dyes, dielectric multilayer films, filter materials, and the many metals (aluminum, gold, beryllium, and so forth) and alloys that are important as coating materials. *See* COMPUTER STORAGE TECHNOLOGY; KERR EFFECT; OPTICAL RECORDING. James Steve Browder

Bibliography. H. Bach and N. Neuroth (eds.), *The Properties of Optical Glass*, 1995; K. S. Hansen (ed.), *Fiber Optic Reference Guide*, 2d ed., 1999; S. A. Jenekhe and K. J. Wynne (eds.), *Photonic and Optoelectronic Polymers*, 1997; I. M. Khan and J. S. Harrison (eds.), *Field Responsive Polymers*, 1999; P. Klocek (ed.), *Handbook of Infrared Optical Materials*, 1991; Optical Society of America, *Handbook of Optics*, vols. 1, 2, 2d ed., 1995, vols. 3, 4, 2000; R. Wood, *Optical Materials*, 1993.

Optical microscope

An instrument used to obtain an enlarged image of a small object. In general, a compound microscope consists of a light source, a condenser, an objective,

and an ocular or eyepiece, which can be replaced by a recording device such as a photoelectric tube or a photographic plate. The optical microscope is limited by the wavelengths of the light used and by the materials available for manufacturing the lenses.

Lenses

The quality and design of the lens system determines the magnifying power, details of image formation, and color correcting capabilities of a light microscope.

Magnifying power. The magnifying power of a compound microscope is the product of the magnification of the objective and the magnifying power of the eyepiece. The latter is computed like that of any magnifier. The magnification of the objective is equal to the distance from the second focal point to the image formed by the objective, divided by the focal length. An objective of 18-mm (0.7-in.) focal length thus has a power of 10×. It is customary to specify objectives in terms of magnifying power instead of focal length. The distance mentioned is called the optical tube length (generally 180 mm or 7 in.), and is to be distinguished from the mechanical tube length, which is the length of the mechanical tube itself. *See* MAGNIFICATION.

Objectives. A microscope objective consists of a set of achromatic lenses which are partially or wholly corrected for longitudinal color, aperture errors, and asymmetry errors. The numerical aperture (NA) of the system is given by $n \sin u$, where n is the refractive index of the object space and u is the angle made by the ray of largest aperture and the axis.

A microscope objective generally consists of a collection of positive lenses comparatively close together (**Fig. 1***a*). Color magnification and curvature of field in the objective are frequently balanced out at least partly by the ocular. *See* ABERRATION (OPTICS).

For high magnifications, the first lens is planoconvex, with the convex surface either concentric or aplanatic to the aperture rays (Fig. 1*b*). To increase the aperture, the object can be embedded in an immersion liquid (a special oil) having a refractive index equal or nearly equal to the index of the first lens. Such an arrangement is termed a homogeneous immersion system (Fig. 1*c*). A cover glass of fixed but small thickness separates the object from the immersion liquid for biological investigations.

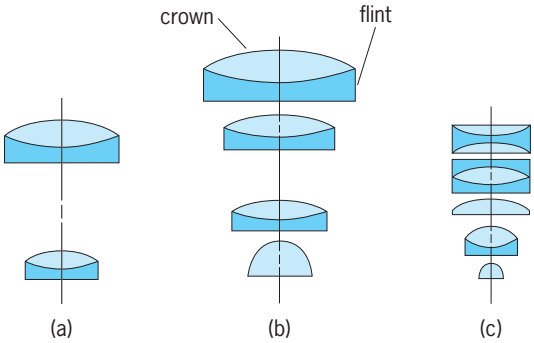

Fig. 1. Microscope objectives: (*a*) 16-mm (0.63-in.) dry achromat; (*b*) typical high-power dry achromat; (*c*) 2-mm (0.08-in.), NA 1.40, oil-immersion apochromat. (*Photographic Service Department, Kodak Research Laboratory*)

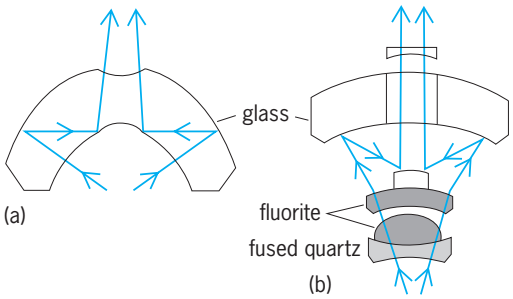

Fig. 2. Two types of catadioptric objective. (*a*) Maksutov type. (*b*) 53X, NA 0.72, ultraviolet objective, designed by Gray. Glass elements in the latter serve purely as reflectors. (*Photographic Service Department, Kodak Research Laboratory*)

Color correction. Achromatic color correction is not good enough for lenses of high aperture. The use of fluorite led to objectives corrected for more than two colors, and most achromatic lenses of high aperture are now made with crystal fluorite. An objective whose chromatic errors are corrected for three colors and whose aperture and asymmetry errors are also corrected is termed apochromatic. The lateral chromatism of such an objective is great and must be removed by a special compensating eyepiece. *See* CHROMATIC ABERRATION; EYEPIECE.

High-aperture microscopes for a field larger than the usual 3° have been designed by adding either a negative lens or a very thick positive lens having a negative Petzval sum. Such objectives are termed plane-achromats.

Because resolving power depends upon wavelength, objectives which are corrected for ultraviolet radiation have been designed. Such objectives are generally corrected for a single wavelength and are termed monochromats. They are made of quartz.

Catadioptric systems. Catadioptric systems have been developed for microscopes. Their great advantage is their comparatively small chromatic aberration. Pure mirror systems have no color aberrations. In catadioptric systems, therefore, it is customary to assign all the power to the mirror or mirrors, keeping the refracting system nearly afocal (**Fig. 2***a*). The chromatic errors of the entire system remain small, and the refracting part can be used to correct the remaining monochromatic errors. However, in catadioptric systems part of the aperture is obscured by the mirror and the ensuing diffraction may damage the fine detail in the image. All microscopic work in the ultraviolet region is done with catadioptric systems (Fig. 2*b*). *See* SCHMIDT CAMERA.

Image formation. Geometrical optics is not sufficient to explain all the details of image formation at high magnifications. According to geometrical optics, a point should be imaged by a perfect objective as a point, but there is diffraction at the aperture, and there may also be diffraction at the object. Diffraction theory applied to the aperture shows that, because

of the finite aperture of the optical system, a spherical concentric wave bounded by the exit pupil produces a light pattern in the image plane in which the light is distributed over a disk of diameter $1.22\lambda/\mathrm{NA}$ with faint rings outside the disk. If the magnifying power of the microscope becomes so large that the rings are visible, the image will contain details which are not in the object. This is undesirable. Thus, the aperture determines the useful magnification of the microscope; it is important that the aperture be enlarged when the magnification has to be increased. *See* DIFFRACTION.

Because the microscope is used to study very small objects at usually high magnifications, diffraction at the aperture is more noticeable in microscopes than in other optical systems. Moreover, most objects viewed with microscopes are so small that there is a significant amount of diffraction at the object, as is seen in the case of dark-field illumination and phase microscopy.

This is of special importance if the object has a periodic structure since even a point light source will then give rise to an imagelike structure. Interference of the rays coming from the light source in phase and being diffracted at the structured object causes this effect and gives rise to sets of images in different planes. This theory (E. Abbe's theory of image formation in the microscope) has been successfully carried further in F. Zernike's theory, which led to the construction of the phase microscope.

In the case of illumination with a large cone of light, these interference patterns fall together at the gaussian focus, so that there is only one image plane. The results in this case, however, can also be derived from Abbe's theory, but only if integration over the aperture is carried out. *See* INTERFERENCE MICROSCOPE; PHASE-CONTRAST MICROSCOPE.

Condensers

An external auxiliary lens is used to condense the light from a light source so that the object is brightly and uniformly illuminated. The usual purpose of a condenser system is to make sure that as much light as possible coming from the object goes through an optical system.

Condensers are used in macroscopic projection, in which an illuminated film or slide is imaged with the help of a projection objective or magnifier. In microscope systems, they are used to direct the light from a light source so that the rays from any object point fill most of the entrance pupil.

A condenser system is usually arranged to image the light source onto the entrance pupil of the optical system (Köhler illumination). The condenser is generally corrected for spherical aberration, color, and sine condition, although the requirements are slightly different than in an image-forming system. Condensers frequently consist of a number of planoconvex lenses with the plane side toward the objective. Sometimes one surface is made aspheric to improve the light concentration. Condensers for projection optics are rarely achromatized, but the effect of color magnification is decreased by vignetting the colored borders. For microscope substage condensers, however, achromatism is a necessary requirement.

The aperture of the condenser must be at least as large as that of the objective with which it is used. Because microscope objectives are generally designed in such a way that they are excellently corrected for color, aperture, and asymmetry errors for only about $\frac{7}{8}$ of their aperture, the condenser need fill only this much of the entrance pupil. A good test of whether the condenser of a microscope is well adjusted is to remove the object and ocular and see whether the exit pupil of the objective is filled uniformly with light up to $\frac{7}{8}$ of its aperture. Max Herzberger

Types of Microscope

The types of microscope discussed in this section are variations of the light- or bright-field microscope.

Light microscope. The mirror, condenser, oculars, and body tube of the light microscope are frequently known as the optical train. The stand, stage, and adjustments comprise the mechanical part of the microscope (**Fig. 3**).

A mirror is usually attached to the substage of the microscope to reflect light along the optic axis of the microscope. When no condenser is used, the concave mirror is used because it concentrates more light on the specimen; a plane mirror is used with a condenser.

Laboratory microscopes are usually supplied with an uncorrected two-lens Abbe condenser with NA 1.25, consisting of a double convex lower and a hyperhemisphere upper lens. A 1.40 NA Abbe is useful for concentrating radiation for fluorescence microscopy. Well-corrected objectives require aplanatic condensers corrected for chromatic aberration for efficient observation.

Objectives vary from a simple doublet lens to complex corrected lens systems. Achromatic objectives are corrected for spherical aberration in one color and for chromatic aberrations in two colors. Apochromatic objectives are corrected to focus three colors together and the spherical aberration is minimized for two colors. Some semiapochromatic and fluorite objectives are of intermediate correction.

With air between the specimen and the objective the maximum NA is about 0.92. Water-immersion objectives have a greater NA, and with immersion oil, an NA of 1.4 is available. Objectives with an NA of 1.6 have been made but require special immersion and mounting media. The resolving power of an objective, the least distance at which two objects can be seen to be separate, is equal to the wavelength of light λ divided by the sum of the numerical apertures of the condenser and objective used. The larger the numerical aperture, the greater is the resolving power. The depth of field seen in focus at one time and the working distance of the objective decrease with an increase in the numerical aperture. The light passed through a microscope is proportional to the square of the numerical aperture and to the inverse of the square of the magnification. Objectives are described also by the equivalent focal length.

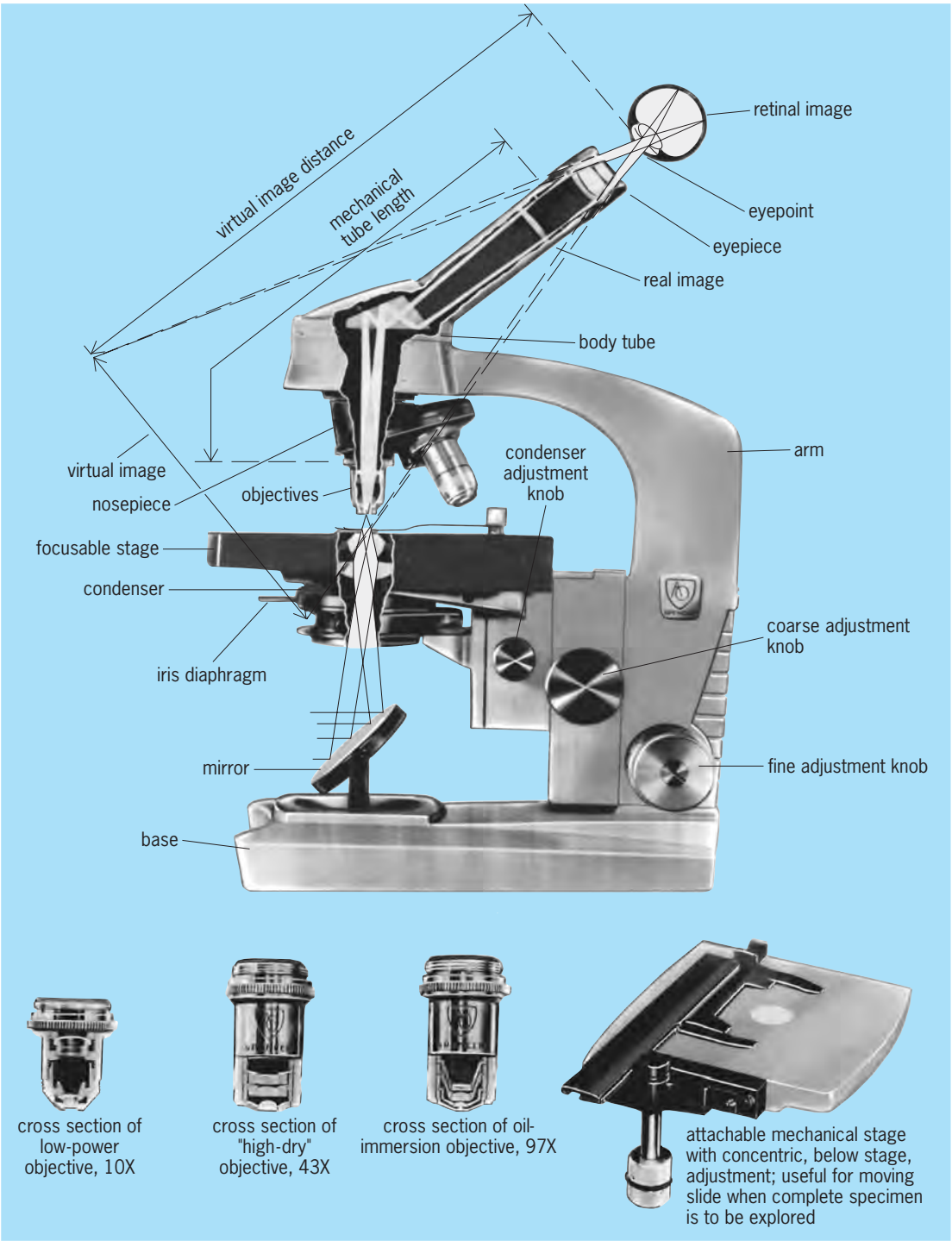

Fig. 3. Diagram of light microscope. (*American Optical Corp.*)

Objectives of shorter focal length have less depth of field, less working distance, and greater magnification.

Photomicrographic objectives are designed to produce a flat image with little distortion. Some objectives obtain a flatter field by means of a concave rather than a flat front lens. Apochromatic objectives are undercorrected and compensating oculars must be used with them to complete the correction for color and for best resolution.

For convenience, two to five objectives can be mounted on a revolving nosepiece to be parfocal and parcentric, so that the specimen remains almost in focus at the center of the field as the objectives are changed. For more critical work, utilizing interference and polarizing microscopes, individually adjustable quick-change nosepieces are employed.

The commonly used Huygenian ocular has a fairly flat field with marked pincushion distortion. Compensating oculars complete the color correction for

apochromatic objectives and have less distortion, but they do have curvature of field. To obtain a flat field with minimum distortion for microprojection or photomicrography special projection oculars are designed, one type with a color-corrected minus lens called a negative amplifier, and the other a positive-projection ocular with a focusable eye lens. Oculars with a high exit pupil are useful for spectacle wearers. Other oculars are designed to give a wide field of view.

The monocular body tube may be of adjustable length. American microscopes are designed for a mechanical tube length of 160 mm (6.3 in.) and a cover-glass thickness of 0.18 mm (0.007 in.). The draw tube is lengthened for thinner and shortened for thicker cover glasses to correct for the spherical aberrations from cover glasses of incorrect thickness.

Binocular bodies are designed for the use of both eyes. Most binocular bodies use prisms to reflect one-half of the light to each eye. Because each eye sees the same field, these binocular bodies do not give stereoscopic vision. The binocular is often longer than the monocular body and the proper tube length is maintained with a compensating lens. Because the tube length is fixed, it is essential for critical microscopy with binocular bodies to use cover glasses 0.18 mm thick. Binocular bodies can be made with stops, or polarizing materials, so that each eye sees with the corresponding half-aperture of the objective and stereoscopic vision is possible, but at one-half the resolving power that would have been obtained with full aperture.

Trinocular bodies are binocular bodies with a third tube for a camera.

The lowest magnification which will resolve the detail required should be used. High magnification gives images that are less bright and therefore difficult to see. Magnification greater than that required for complete visible detail is called empty magnification and is useless for seeing, although sometimes helpful for measurement. For visual microscopy the total magnification (magnification of the objective times that of the ocular) should be about 1000 times the numerical aperture in use. Apochromatic objectives and compensating oculars are desirable for best vision and color photomicrography.

Inverted microscope. The inverted microscope has the body of the microscope, including the objective and the ocular, below the stage and the illumination above the stage for transmitted light. With opaque materials, the vertical illuminator is used under the stage near the objective. The inverted microscope is especially useful for the examination of surfaces (**Fig. 4**). Specimens placed on the stage can be held substantially in focus. Large and awkward specimens can be moved over the stage more readily than with the usual microscope. The inverted microscope is also useful for microdissection and the observation of hanging-drop preparations and is convenient for observing chemical reactions, melting-point determinations, and photomicrography. The camera can be included in the base, as in the metallograph microscope, for stability. Either monocular or binocular bodies can be used with the inverted microscope.

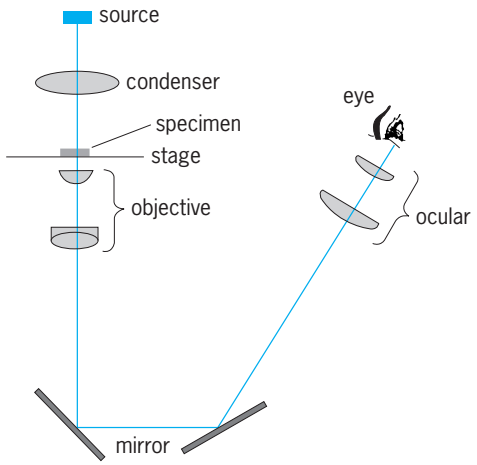

Fig. 4. Optical diagram of inverted microscope.

Comparison microscope. The comparison microscope is an arrangement of two microscopes connected by a special viewing ocular so that the field of one microscope is seen at one side of a vertical dividing line and the field of the other microscope on the opposite side of the dividing line; or it may be a projection type of microscope in which the image is compared with a template or known pattern.

When two microscopes and a comparison ocular are used, the magnifications of each must be matched. The specimens are placed on the microscopes and usually require separate lighting. A common application is the examination of bullets. A test bullet fired from the suspected gun is placed under one microscope and the bullet recovered from the scene of the crime is placed under the other microscope. Holders which allow rotation of the bullets are turned, and if the grooves on the two bullets match, it is evident that they came from the same gun. Comparison microscopes can be used to compare grain size, structure, distribution of elements, color, and various other characteristics of any two specimens. Either transmitted light, vertical illumination, or a combination of lighting types can be used.

A projection microscope with a built-in screen can also be used as a comparison microscope. The image of the part is focused on the screen, and the contours of the image are measured with scales or compared with a template of the desired form, for example, a profile of a gear or a screw thread. The stage is usually modified to have a proper holder for the kind of specimen to be examined.

Dissecting microscope. Dissecting microscopes are of two types. The simplest is a magnifying glass mounted on a support above a glass plate, used for the dissection of materials. A mirror may be present to reflect light to the specimen.

The more usual dissecting microscope, often called a Greenough microscope, is a stereoscopic microscope composed of two separate microscopes

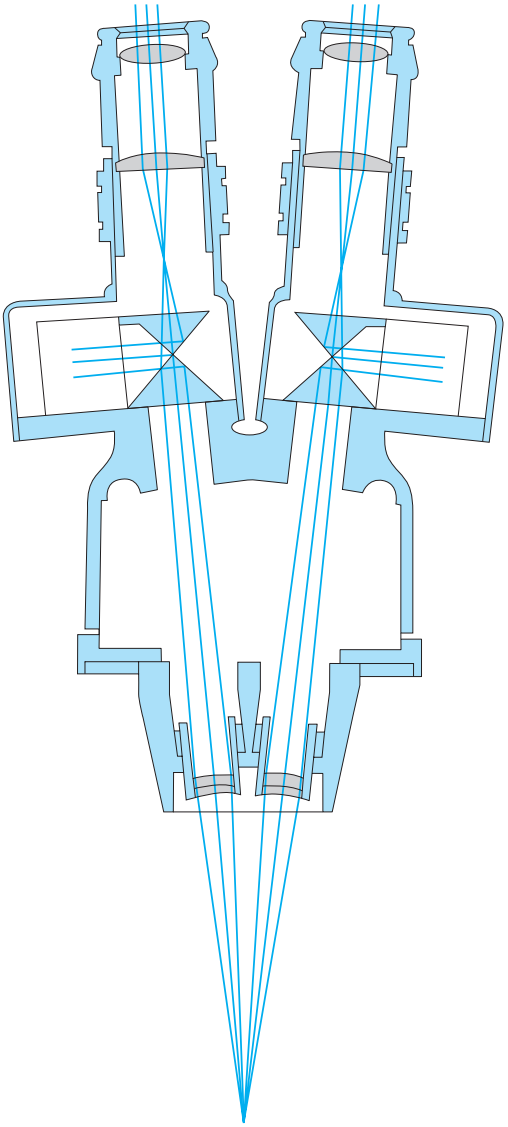

Fig. 5. Diagram of light rays as they pass through a binocular biobjective microscope.

fastened together and used as a single unit on one stand (**Fig. 5**).

This is a truly stereoscopic instrument because the right eye sees the specimen from the right side and the left eye from the left side. Prisms are usually included in the body tube to erect the image; thus movements of the specimen are direct and are not reversed as with the monobjective microscope. Because two objectives are required, the mechanical difficulty of placing the lenses close together limits the numerical aperture to approximately 0.12. There is no advantage in using more than 120 diameters magnification because further magnification would be empty or useless.

The binocular microscope was developed for biological dissecting, but it is used extensively in industry for the examination and assembly of small parts such as transistors. The body and focusing adjustment can be mounted on lathes and other machinery when the work must be controlled visually to close tolerances; they can be mounted on a stand with a long arm for examination of large specimens or may be used to assist the surgeon in the operating room.

When the angle of the objectives and oculars is the same, the specimen is seen in true depth. By changing these angles the perception of depth can be increased or decreased. Hyperstereoscopy is useful for biological dissection.

Examination with the stereoscopic microscope is helpful in orienting the microscopist to the specimen before instruments of greater magnification are used. By proper illumination both opaque and transmitted materials can be examined.

Metallurgical microscope. The metallurgical microscope is a laboratory microscope with a focusing stage and a vertical illuminator, used primarily for the examination of metal surfaces.

The specimens are usually embedded in molded plastic so that the surface is at a definite position, surfaced with a series of increasingly finer abrasives, and polished. To differentiate the constituents of the metal, the surface is etched lightly by chemical treatment before examination with the microscope.

The metallurgical microscope shows the structure of the metal, including grain size, as well as the nature and distribution of the components. The roughness or polishability of the metal can be studied. Because many of the metals used commercially are mixtures, or alloys, rather than pure metals, the metallurgical microscope is important for analysis and for assessing the effects of heat treatment and surface changes.

Microscopy

This discussion of microscopy considers the following with relation to the optical microscope: illumination, calibration and measurement, immersion fluids, mountants, and the hanging-drop method.

Illumination. Illumination of the microscope is obtained from a bright surface or from a luminous source concentrated with the aid of condensing lenses. Sources commonly used are tungsten-filament lamps and carbon, mercury, and xenon arcs.

With an illuminated-surface light source the lamp is positioned and the microscope condenser focused so that the field and aperture of the microscope are lighted as uniformly as possible.

Illuminators with focusing condensers and iris diaphragms, called research lamps, are used for critical, Köhler, and other more efficient methods.

Critical illumination. In critical illumination an image of the light source is focused on the specimen from a uniform source, for example, a ribbon-filament lamp (**Fig. 6**). Parallel light from an illuminator focused to infinity is directed into the microscope condenser which is then focused so that the image of the source is in the plane of the specimen.

Köhler's method. This method is used with coiled filaments or other sources of irregular form or brightness to obtain a uniformly illuminated field. An image of the filament large enough to fill the opening of the iris is focused on the microscope condenser. The microscope condenser then is focused so that the image of the iris diaphragm on the lamp is in

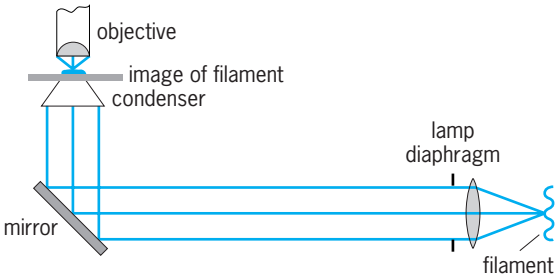

Fig. 6. Light path in critical illumination. (*American Optical Corp.*)

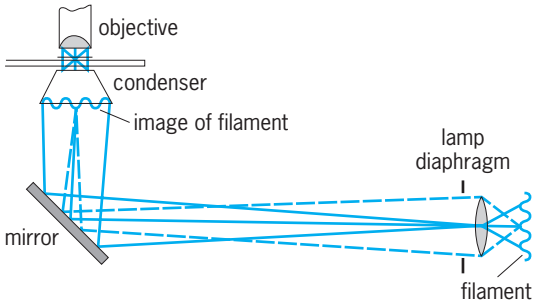

Fig. 7. Light path in Köhler illumination. (*American Optical Corp.*)

focus with the specimen and the lamp iris is opened only enough to fill the field of view. The iris of the microscope condenser is opened only enough to illuminate the back aperture of the objective. No ground glass is used. The condensing lens of the illuminator becomes the effective source and the field is uniformly illuminated, even though the source itself is not uniform. Loss of contrast from misplaced glare light is minimized (**Fig. 7**).

Shillaber's type 3. This method is useful when the lens systems of the condensers are inadequate to illuminate the field and aperture of the microscope.

A ground glass is placed close to the condensing lens of the illuminator between the lens and the lamp bulb, or the surface of the lamp is ground on the side facing the lens. Because the ground surface does not diffuse all the light, the condensing lens can concentrate more light into the microscope than could be obtained from the same area of a surface-type illuminator. An effectively larger source results which is useful with low-power (searcher) objectives for a large field of view and when the lamp filament is not large enough for the condensing system—for example, some sources built into the base of the microscope. This method is preferable when a diffuser must be used, but more stray or glare light is produced than with the Köhler method.

Vertical illumination. Vertical illumination uses a partly silvered mirror, or a prism, to reflect light through the objective onto the specimen. The light reflected back through the objective from the specimen forms the image which is seen.

Epi-illumination is vertical illumination with the illuminating light passing around a special objective to the specimen. Only the light reflected from the specimen passes through the objective. Better definition results from the separation of the paths of the illuminating and viewing light.

Metals, ores, minerals, and opaque materials with adequate reflecting surfaces are examined with vertical illumination.

Dark-field illumination. In dark-field illumination the specimen appears bright, or self-luminous, against a black background. The illuminating beam is a hollow cone of light formed by an opaque stop at the center of the condenser, large enough to prevent any direct light from entering the objective. A specimen placed at the concentration of the light cone is seen with the light scattered or diffracted by it, and the smallest particle revealable depends upon the intensity of the available light, even though the particle may be too small for resolution as to size and shape. Size can sometimes be inferred from the number of particles found in a given volume of the specimen (**Fig. 8**).

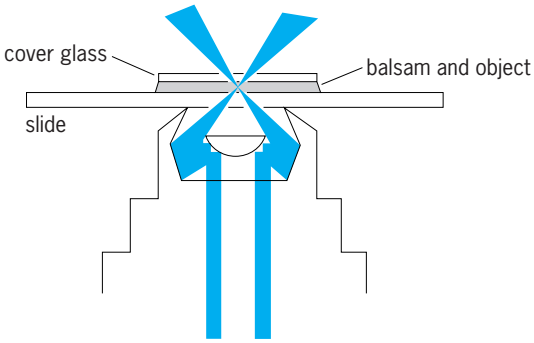

Fig. 8. Dark-field illumination. (*American Optical Corp.*)

Ultramicroscopy, used for the examination of colloids and smokes, is a dark field obtained with an intense narrow beam of light directed through the specimen at right angles to the optical axis of the microscope.

Rheinberg illumination or optical staining is also a modification of the dark-field method. The central disk is transparent and colored, rather than opaque, and an annulus of a complementary color fills the remaining condenser aperture. The specimen is seen in the color of the annulus against a background of the color of the central disk; for example, when the annulus is yellow and the background blue, the specimen appears yellow against a blue background.

Modification by filters. Filters are used between the microscope illuminator and the microscope to control the intensity or quality of illumination. Filters can be liquids in a flat-sided container or solids. Solid filters are made of colored glass, gelatin, or other materials.

Clear filters of water or of heat-absorbing glass are used to remove the excess heat from the lamp beam when delicate specimens must be protected. Blue filters remove excess yellow from tungsten light so that it resembles daylight. Neutral filters of glass or metal (Inconel) deposited on glass are used to reduce the amount of light without altering its color.

Colored filters change the quality of the light by selectively absorbing certain wavelengths and are grouped into broadband and narrow-band types. Broadband filters are used to increase visibility by modifying the color contrast of the specimen; for example, a yellow (minus blue) filter transmits all of the spectrum except the blue. Complementary filters increase contrast, and filters of similar color decrease color contrast. Polychromatic filters, transmitting two or more spectral colors (for example, a minus green filter passing blue and red) are useful with some stained specimens. Smaller regions of the spectrum are isolated with narrow-band-pass filters, interference filters, and by combinations of filters. Monochromatic light usually is obtained from a single spectral line of an arc source rather than with filters. *See* COLOR FILTER; INTERFERENCE FILTER.

Calibration and measurement. The size of the object under examination is frequently a desirable criterion for identification. A reference reticule is placed on the diaphragm in the focal plane of the ocular and seen in focus with the specimen from measurement in microscopy. A cross-hair reticule is satisfactory for position; scales are used for measuring distance and a net reticule for counting and drawing. Oculars with a focusing eye lens permit focusing the scale to the microscopist's eye.

Angular measurements are made by orienting one side of the specimen to the cross hair, reading the position of the graduated stage of the microscope, and rotating the stage until the other side is in line with the cross hair. The difference between the first and second readings of the scale gives the angle. Another method uses a goniometer ocular and measures the angle with the scale of the goniometer.

Linear measurements are made by placing one edge of the specimen at the cross hair, reading the position of the graduated mechanical stage, moving the specimen to the other side, and reading the position of the mechanical stage. The difference of the readings is the distance. This method is useful for larger specimens because the mechanical stages are ordinarily not graduated closer than 0.1 mm (0.004 in.). For measurements of smaller specimens a scale is used in the eyepiece and the size of the specimen obtained from the scale. The actual or true distance of the ocular scale depends upon the magnification of the microscope and is obtained by calibration with the aid of a stage micrometer.

More precise measurements can be made with a screw micrometer eyepiece, a filar micrometer, or a step micrometer eyepiece. The movable scales with graduated controls facilitate measurement, but must also be calibrated.

Immersion fluids. Air or other low-refractive-index material between the specimen and the objective of the microscope limits the angle of light that can enter the objective, the numerical aperture, and the resolving power of the system. To obtain greater values, a medium of higher refractive index must be placed between the condenser of the microscope and the microscope slide, and between the cover slip of the microscope and the front lens of the objective. Immersion fluids may be aqueous substances (such as water or glycerin), oils, or organic materials.

In addition to the correct refractive index and dispersion, a satisfactory immersion oil must possess the following physical properties: chemical inertness, no tendency to spread or creep, no hardening when exposed to air, and little change with age.

Mountants. Mountants for specimens to be observed under the microscope are usually classed as temporary or permanent. Temporary mountants are water, physiological saline, serum, and the like. Such preparations last longer when the cover glass is sealed to the slide with petroleum jelly or paraffin. Semipermanent mounts are made with glycerin, glycerin jelly, or oils, and permanent mountants from gums, resins, or polymers which harden and remain solid at room temperatures. Canada balsam and damar have been largely replaced by synthetic polymers, such as Clarite, and polyvinyl alcohol.

A good mountant should provide enough difference in refractive index for contrast within the specimen, be transparent, not change color or consistency with time, hold the cover glass firmly in place, and not fade stains, or otherwise alter the specimen. *See* MICROTECHNIQUE.

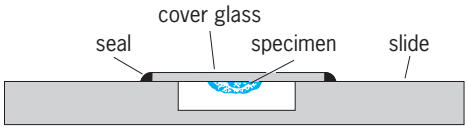

Fig. 9. Hanging-drop preparation.

Hanging-drop preparations. The specimen is placed in a drop of a suitable fluid on a cover slip and the cover slip is inverted over a slide which has a hollow ground in it (**Fig. 9**). Various concavities are used—round, flat, and so on—and the thickness of the slide varies with the preparation. The cover glass is usually sealed to the slide to retard evaporation. More complex preparations have channels so that the environment of the specimen can be changed or kept constant. Some hanging-drop-preparation slides are open at the bottom to make it possible to bring microdissection needles to the specimen for surgical procedures on the specimen while it is observed through the microscope. *See* MICROMANIPULATION.

Oscar W. Richards

Near-Infrared Microscopy

Near-infrared microscopy is an optical method that can be used for studying a variety of materials that are opaque in transmitted visible light (400–700 nm) yet translucent in the near infrared (700–1200 nm). The method utilizes the near-infrared optical microscope, a device for the conversion of the near-infrared image to a visible image.

The heart of the microscope is the image converter tube which is positioned above the microscope objective in the plane of the primary image. The near-infrared to visible conversion is achieved within the

image tube, a cathode tube containing a positive terminal coated with fluorescent material. The impingement of the near-infrared light image onto the negative terminal of the cathode tube releases electrons which are focused by means of an electrostatic lens onto the fluorescent screen.

The image formed on the fluorescent screen either can be viewed directly by means of an eyepiece positioned above the screen or can be photographed with standard cameras using ordinary film. Sharpening of the image is achieved by use of near-infrared filters located in front of the image tube. These filters serve to suppress the visible-light image and limit the wavelength range of the transmitted beam. However, such filters also tend to reduce the percentage of transmitted light so that a need arises for a light source that has a sufficiently intense near-infrared component (such as xenon and tungsten filament lamps).

Lawrence A. Harris

Near-Field Optical Microscopy

A fundamental law of optics, the so-called diffraction limit, states that two objects can be imaged as separate entities only if their distance is larger by about one-half the wavelength of visible light, which ranges from 400 to 700 nm. As a consequence, conventional optical microscopy is restricted to a resolution of about 200 nm, and this is not enough for many important observations. The scanning near-field optical microscope (NSOM or SNOM) circumvents the diffraction limit. In contrast to nonoptical methods of surpassing this limit, such as electron microscopy, it can provide information on such qualities as color, luster, transmissivity, and birefringence, which are sensitive indicators of material composition and status. Furthermore, it operates at ambient conditions, a prerequisite for observations of living organisms.

The ability to image an object is closely related to the ability to confine radiation to a narrow spot. Lenses and mirrors, the common focusing elements, are subject to the diffraction limit and hence can be ruled out. When illuminated, however, holes in an opaque screen or small scattering particles can also act as confined sources of radiation, and they can be made much smaller than half a wavelength.

Light emerging from such a small light source remains confined only over a distance comparable to its size; this is the so-called near-field zone. Beyond this zone, that is, in the far field, it diverges rapidly. The light source (termed an optical probe hereafter) hence must be brought so close to the object to be investigated that its near field is disturbed by the object. This requirement can best be satisfied by mounting the optical probe at the apex of a sharply pointed tip. The most common optical probe is an aperture in an opaque metallic coating which covers the corner of a transparent crystal or the end of an optical fiber. This probe is prepared so that a sharp tip is formed.

Optical detectors such as photomultipliers are fairly large and therefore cannot be mounted in the near field. The disturbance of the near field by the object, however, also influences the far field. This property is illustrated by the case in which the optical probe is a small aperture and the object is a strongly absorbing metal film with small holes. If the probe is positioned directly above a hole, the light passing through the aperture can also pass the object without much attenuation; otherwise, the metal film will act like a lid, reducing the light transmission the closer the aperture is to the film surface (**Fig. 10**).

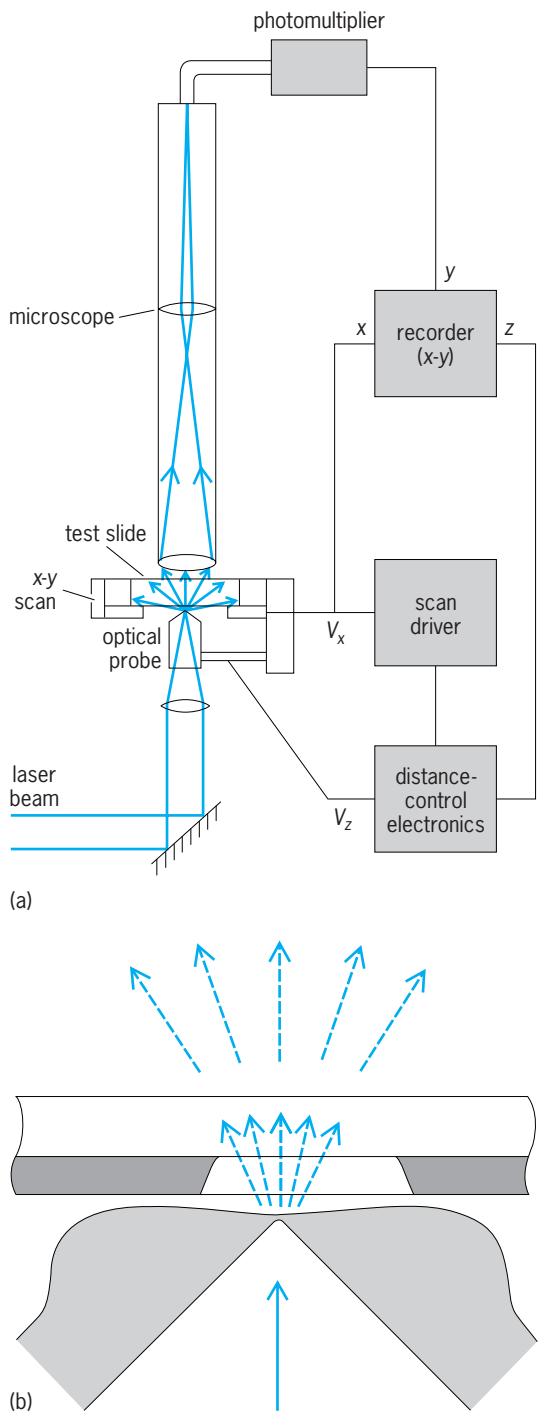

Fig. 10. Scanning near-field optical microscope (SNOM or NSOM). (*a*) Setup. (*b*) Optical probe next to a metallic film with small holes. (*After U. Dürig, D. W. Pohl, and F. Rohner, Near-field optical scanning microscopy, J. Appl. Phys., 59:3318–3327, 1986*)

An image of the object surface can be composed by scanning the probe line by line over the object surface and recording the intensity of radiation as a function of position. An image as it would appear to the eye can be easily created by associating the signal intensity with the gray scale on a computer display; color images can also be created in this way.

The scanning motion is conveniently generated by means of piezoelectric elements which can be made to bend sideways or expand in length when voltage is applied in an appropriate way. Such scanners were originally developed for scanning tunneling microscopy. They can be used simultaneously for distance control. *See* MICROSCOPE; SCANNING TUNNELING MICROSCOPE. D. W. Pohl

Bibliography. R. Barer, A. Bouwers, and D. S. Gray, Reflecting microscope, *Proc. London Conf. Opt. Instrum.*, pp. 43–76, 1951; E. M. Chamot and C. W. Mason, *Handbook of Chemical Microscopy*, vol. 1, 3d ed., reprint 1958; P. J. Duke and A. G. Michette (eds.), *Modern Microscopy: Techniques and Applications*, 1989; P. Gray, *The Microtomist's Formulary and Guide*, 1954, reprint 1975; B. Herman and J. J. Lemasters (eds.), *Optical Microscopy: Emerging Methods and Applications*, 1992; R. P. Loveland, *Photomicrography: A Comprehensive Treatise*, 2 vols., 1970, reprint 1981; M. Pluta, *Advanced Light Microscopy*, 3 vols., 1988, 1989, 1993; D. W. Pohl and D. Courjon (eds.), *Near Field Optics*, NATO ASI Ser. E242, 1993; J. Sanderson, *Biological Microtechnique*, 1994; E. M. Slayter and H. S. Slayter, *Light and Electron Microscopy*, 1992; B. H. Walker, *Optical Engineering Fundamentals*, 1994; V. K. Zworykin and G. A. Morton, Applied electron optics, *J. Opt. Soc. Amer.*, 16:181–189, 1936.

Optical modulators

Devices that serve to vary some property of a light beam. The direction of the beam may be scanned as in an optical deflector, or the phase or frequency of an optical wave may be modulated. Most often, however, the intensity of the light is modulated.

Rotating or oscillating mirrors and mechanical shutters can be used at relatively low frequencies (less than 10^5 hertz). However, these devices have too much inertia to operate at much higher frequencies. At higher frequencies it is necessary to take advantage of the motions of the low-mass electrons and atoms in liquids or solids. These motions are controlled by modulating the applied electric fields, magnetic fields, or acoustic waves in phenomena known as the electrooptic, magnetooptic, or acoustooptic effect, respectively. *See* ACOUSTOOPTICS; ELECTROOPTICS; MAGNETOOPTICS.

For the most part, it will be assumed that the light to be modulated is nearly monochromatic—either a beam from a laser or a narrow-band incoherent source. *See* LASER.

Electrooptic effect. The quadratic or Kerr electrooptic effect is present in all substances and refers to a change in refractive index, Δn, proportional to the square of the applied electric field E. The liquids nitrobenzene and carbon disulfide and the solid strontium titanate exhibit a large Kerr effect. *See* KERR EFFECT.

Much larger index changes can be realized in single crystals that exhibit the linear or Pockels electrooptic effect. In this case, Δn is directly proportional to E. The effect is present only in noncentrosymmetric single crystals, and the induced index change depends upon the orientations of E and the polarization of the light beam with respect to the crystal axes. Well-known linear electrooptic materials include potassium dihydrogen phosphate (KDP) and its deuterated isomorph (DKDP or KD∗P), lithium niobate ($LiNbO_3$) and lithium tantalate ($LiTaO_3$), and semiconductors such as gallium arsenide (GaAs) and cadmium telluride (CdTe). The last two are useful in the infrared (1–10 micrometers), while the others are used in the near-ultraviolet, visible, and infrared regions (0.3–3 μm). *See* CRYSTAL OPTICS; FERROELECTRICS; SEMICONDUCTOR.

The phase increment Φ of an optical wave of wavelength λ that passes through a length L of material with refractive index n is given by Eq. (1). Thus phase

$$\Phi = \frac{2\pi}{\lambda} nL \tag{1}$$

modulation can be achieved by varying n electrooptically. Since the optical frequency of the wave is the time derivative of Φ, the frequency is also shifted by a time-varying Φ, yielding optical frequency modulation. *See* MODULATION.

The refractive index change is given in terms of an electrooptic coefficient r by Eq. (2), where typical

$$n(E) = n(0)\frac{-n^3 rE}{2} \tag{2}$$

values for n and r are 2 and 3×10^{-11} m/V, respectively.

Electrooptic intensity modulation. Intensity modulation can be achieved by interfering two phase-modulated waves as shown in **Fig. 1**. The electrooptically induced index change is different for light polarized parallel (p) and perpendicular (s) to the

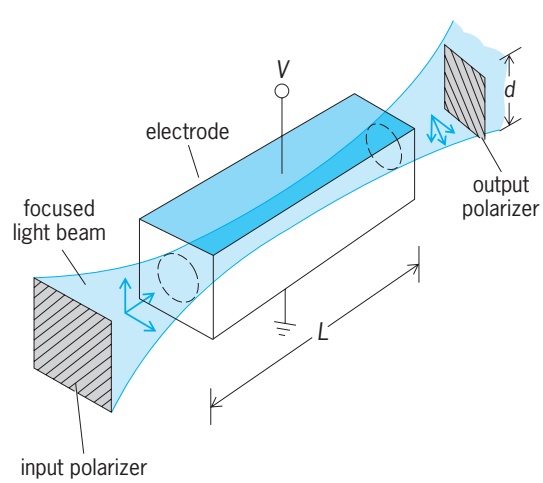

Fig. 1. Electrooptic intensity modulator.

modulating field; that is, $r_p \neq r_s$. The phase difference for the two polarizations is the retardation Γ given by Eq. (3), which is proportional to the ap-

$$\Gamma(V) = \Phi_p - \Phi_s = \frac{2\pi}{\lambda}(n_p - n_s)L \qquad (3)$$

plied voltage V. If it is assumed that $n_p(0) = n_s(0)$ for $V = 0$, then $\Gamma(0) = 0$. Further assume that $\Gamma(V') = \pi$. Then incident light polarized at 45° to the field may be resolved into two equal, in-phase components parallel and perpendicular to the field. For $V = 0$, they will still be in phase at the output end of the electrooptic crystal and will recombine to give a polarization at +45° which will not pass through the output polarizer in Fig. 1. For $V = V'$, however, the two components will be out of phase and will recombine to give a polarization angle of −45° which will pass through the output polarizer. The switching voltage V' is called the half-wave voltage and is given by Eq. (4), with d the thickness of the crystal and r_c

$$V' = \frac{\lambda d}{n^3 r_c L} \qquad (4)$$

an effective electrooptic coefficient. A typical value for $n^3 r_c$ is 2×10^{-10} m/V. *See* POLARIZED LIGHT.

Electrooptic devices can operate at frequencies from dc through hundreds of gigahertz.

Acoustooptic modulation and deflection. All transparent substances exhibit an acoustooptic effect—a change in n proportional to strain. Typical materials with a substantial effect are water, glass, arsenic selenide (As_2Se_3), and crystalline tellurium dioxide (TeO_2).

An acoustic wave of frequency F has a wavelength given by Eq. (5), where C is the acoustic velocity,

$$\Lambda = \frac{C}{F} \qquad (5)$$

which is typically 3×10^3 m/s. Such a wave produces a spatially periodic strain and corresponding optical phase grating with period Λ that diffracts an optical beam as shown in **Fig. 2**. A short grating ($Q < 1$) produces many diffraction orders (Raman-Nath regime), whereas a long grating ($Q > 10$) produces only one diffracted beam (Bragg regime). The quantity Q is defined by Eq. (6). If one detects the

$$Q = \frac{2\pi L\lambda}{n\Lambda^2} \qquad (6)$$

Bragg-diffracted beam (diffraction order + 1), its intensity can be switched on and off by turning the acoustic power on and off.

The Bragg angle θ through which the incident beam is diffracted is given by Eq. (7). Thus the beam

$$\sin\frac{\theta}{2} = \frac{\lambda}{2\Lambda} \qquad (7)$$

angle can be scanned by varying F.

Acoustooptic devices operate satisfactorily at speeds up to several hundred megahertz.

Optical waveguide devices. The efficiency of the modulators described above can be put in terms of the electrical modulating power required per unit bandwidth to achieve, say, 70% intensity modulation or one radian of phase modulation. For very good bulk electrooptic or acoustooptic modulators, approximately 1 to 10 mW is required for a 1-MHz bandwidth; that is, a figure of merit of 1 to 10 mW/MHz. This figure of merit can be improved by employing a dielectric waveguide to confine the light to the same small volume occupied by the electric or acoustic modulating field, without the diffraction that is characteristic of a focused optical beam. This optical waveguide geometry is also compatible with optical fibers and lends itself to integrated optical circuits. *See* INTEGRATED OPTICS; OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

An optical wave can be guided by a region of high refractive index surrounded by regions of lower index. A planar guide confines light to a plane, but permits diffraction within the plane; a strip or channel guide confines light in both transverse dimensions without any diffraction.

Planar guides have been formed in semiconductor materials such as $Al_xGa_{1-x}As$ and $In_xGa_{1-x}As_yP_{1-y}$ by growing epitaxial crystal layers with differing values of x and y. Since the refractive index decreases as x increases, a thin GaAs layer surrounded by layers of $Al_{0.3}Ga_{0.7}As$, for example, will guide light. If one outer layer is p-type and the other two n-type, a reverse bias may be applied to the pn junction to produce a large electric field in the GaAs layer. Very efficient electrooptic modulators requiring about 0.1 mW/MHz at $\lambda = 1$ μm have been realized in such junctions. Further improvement has been realized by loading the planar guide with a rib or ridge structure to provide lateral guiding within the plane. Ridge devices 1 mm long require about 5 V to switch light on or off, with a figure of merit of 10 μW/MHz.

Both planar and strip guides have been formed in $LiNbO_3$ by diffusing titanium metal into the surface in photolithographically defined patterns. Planar guides with surface acoustic wave transducers (**Fig. 3**) have been employed in waveguide

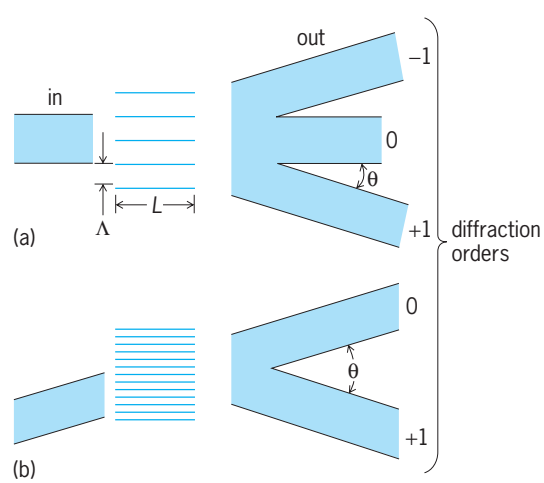

Fig. 2. Acoustooptic modulator-deflector. (*a*) Raman-Nath regime. (*b*) Bragg regime.

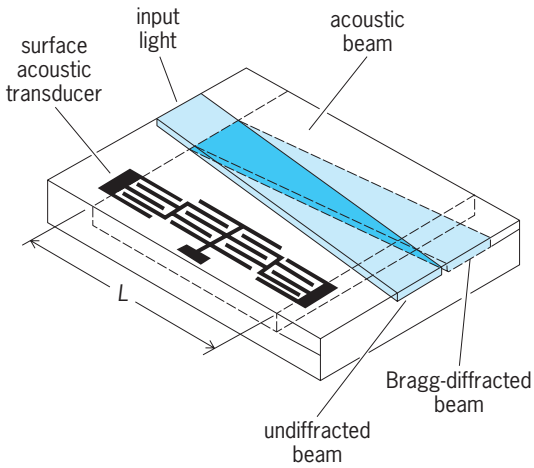

Fig. 3. Waveguide acoustooptic modulator-deflector.

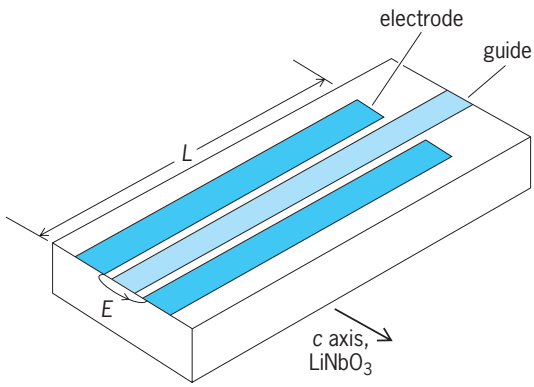

Fig. 4. Titanium-diffused $LiNbO_3$ strip waveguide with coplanar electrodes for electrooptic phase modulation.

acoustooptic devices. Strip guides (**Fig. 4**) have been used to make electrooptic phase modulators 3 cm long requiring 1 V to produce π radians of phase shift at $\lambda = 0.63\ \mu$m; the figure of merit is 2 μW/MHz.

Lithium niobate waveguide modulation. In order to make an intensity modulator or a 2×2 switch, a directional coupler waveguide pattern is formed, and electrodes are applied as shown in **Fig. 5**. With no voltage applied and L the proper length for a given separation between waveguides, light entering guide 1 will be totally coupled to and exit from guide 2; with the proper applied voltage, the phase match between the two guides will be destroyed, and no light will be coupled to guide 2, but will all exit from guide 1. Switches and modulators operating at speeds beyond 10 GHz have been realized in this way.

The directional coupler switch is difficult to manufacture reproducibly because of the small separation between guides, a critical dimension. A more robust modulator structure is based on a waveguide version of the Mach-Zehnder interferometer (**Fig. 6**). In a Mach-Zehnder modulator, the input light is divided equally by a Y-branch splitter and travels along two identical paths before recombining in a second Y-branch. To intensity-modulate the output, equal and opposite phases are induced electrooptically in the two arms by the voltages $+V$ and $-V$ applied to the electrodes. When the phase difference is π, light in the Y-branch combiner interferes destructively and radiates into the substrate with no signal appearing in the output guide. When the phase difference is 0 or 2π, the signals in the two arms add constructively and all the light emerges in the output guide. *See* INTERFEROMETRY.

Lithium niobate Mach-Zehnder modulators are often used in long-distance digital lightwave systems, operating at speeds up to 10 Gb/s. They are well suited to this application because they produce intensity modulation without any associated phase modulation. Any extraneous phase modulation would cause a widening of the laser spectrum, which in turn would cause the optical pulses to spread in time as they traveled along the fiber due to chromatic dispersion in the fiber. Spurious phase modulation is known as chirp. Chirp limits the maximum bit rate and span length of a light-wave system to the point at which the pulse spreading remains less than about one-fourth the width of the data pulses. Chirp is measured as the ratio α of the phase modulation to amplitude modulation. Commercial Mach-Zehnder modulators operate with $\alpha = 0$ and a bit rates in excess of 10 Gb/s.

Semiconductor laser modulation. A semiconductor laser comprises a forward-biased *pn* junction, a waveguide region, and a resonator (**Fig. 7***a*). The dc injection current provides gain and determines the laser output intensity. An additional modulating current may be inserted to produce a directly modulated laser, operating at speeds of a few gigahertz (Fig. 7*b*). This is a low-cost method for providing a modulated output, but its application is limited to lower bit rates and shorter spans by chirp associated with modulation of the waveguide refractive index by the modulating current. Typical values of α in directly modulated lasers are about 4. *See* OPTICAL GUIDED WAVES.

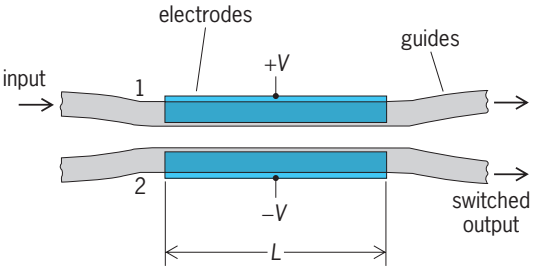

Fig. 5. Directional coupler waveguide modulator switch.

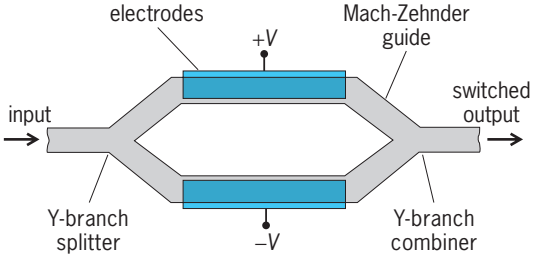

Fig. 6. Mach-Zehnder electrooptic waveguide modulator.

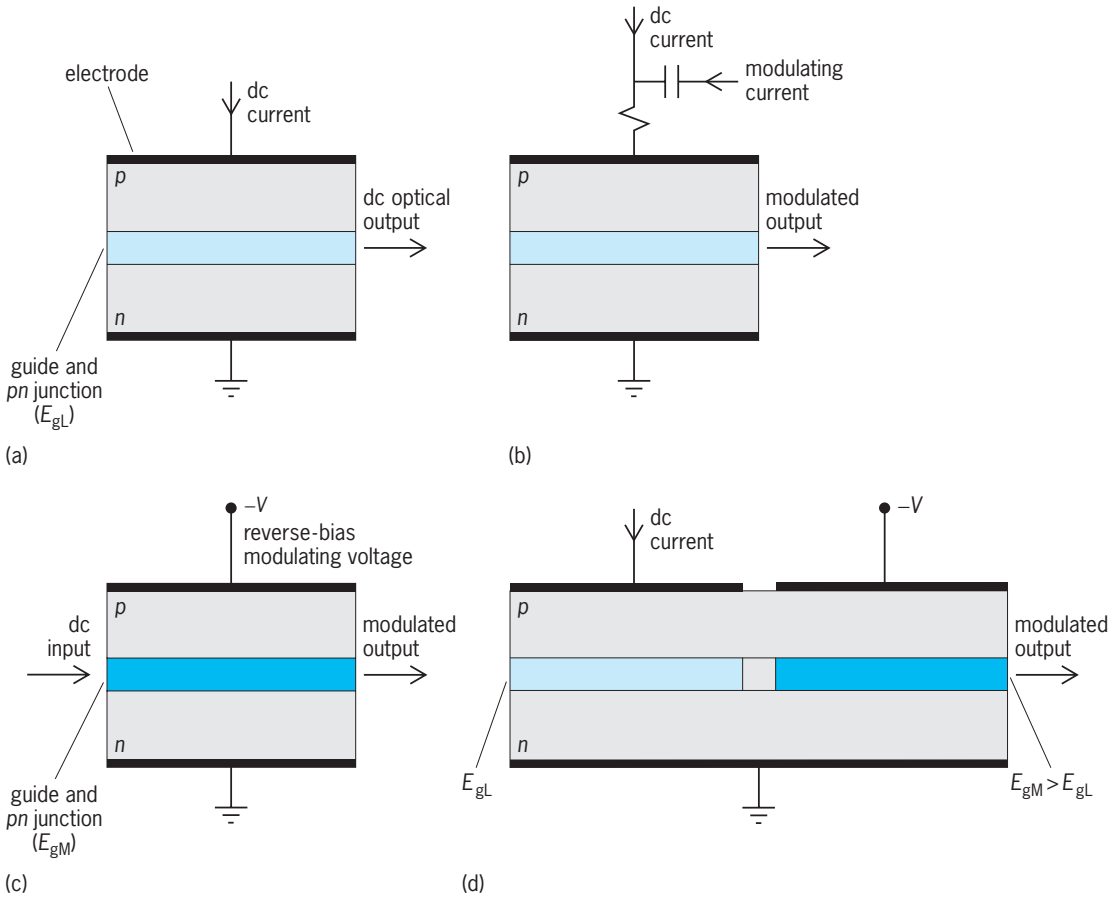

Fig. 7. Semiconductor laser modulation. (*a*) Unmodulated semiconductor laser. (*b*) Directly modulated laser. (*c*) Electroabsorption modulator. (*d*) Integrated externally modulated laser.

When chirp is critical, one must use an external modulator such as a lithium niobate Mach-Zehnder modulator. An alternative is the semiconductor electroabsorption modulator, consisting of a reverse-biased *pn* junction with bias voltage $-V$ and an associated waveguide (Fig. 7*c*). The absorption edge energy E_{gM} of the electroabsorption modulator's waveguide material is designed to be just above the energy of the transmitted laser light $E_{gL}[= hc/\lambda]$. Under these conditions, light entering the waveguide is transmitted to the output when $V = 0$. However, when the reverse-bias V is increased, the absorption edge is shifted to lower energy and reduces the transmission, providing intensity modulation. Electroabsorption modulators can be designed with small but nonzero values of α.

The coupling of an unmodulated semiconductor laser to an external electroabsorption (or Mach-Zehnder) modulator is complicated by requirements for precise alignment of waveguides and minimal reflections from interfaces, which are overcome by integrating on a single chip an unmodulated laser and an electroabsorption modulator (Fig. 7*d*). In a commercial electroabsorption modulator, the substrate is InGaAsP and the laser operates at 1550 nm, the wavelength at which optical fiber has minimum attenuation. The absorption edge in the laser section is near 1550 nm, but the absorption edge in the electroabsorption section is designed to be at a slightly higher energy or, equivalently, shorter wavelength of about 1500 nm to make it transparent to the laser wavelength. Commercial versions of this integrated external modulator laser operate at 2.5 Gb/s and are capable of higher bit rates. *See* INTEGRATED OPTICS; SURFACE-ACOUSTIC-WAVE DEVICES.

Ivan P. Kaminow

Bibliography. I. P. Kaminow and T. L. Koch (eds.), *Optical Fiber Telecommunications IIIB*, Academic Press, 1997; S. E. Miller and I. P. Kaminow, *Optical Fiber Telecommunications II*, Academic Press, 1988; T. Tamir (ed.), *Guided-Wave Optoelectronics*, 2d ed., Springer-Verlag, 1990.

Optical phase conjugation

A process that involves the use of nonlinear optical effects to precisely reverse the direction of propagation of each plane wave in an arbitrary beam of light, thereby causing the return beam to exactly retrace the path of the incident beam. The process is also known as wavefront reversal or time-reversal reflection. The unique features of this phenomenon suggest widespread application to the problems of optical beam transport through distorting or inhomogeneous media. Although closely related, the field

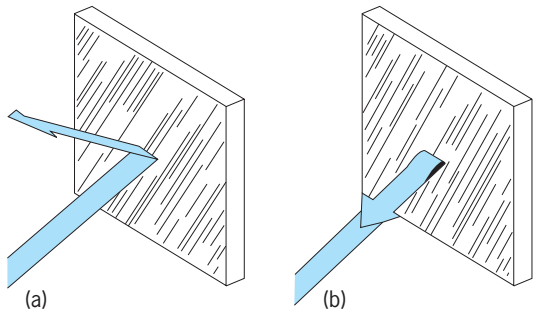

Fig. 1. Comparison of reflections (*a*) from a conventional mirror and (*b*) from an optical phase conjugator. (*After V. J. Corcoran, ed., Proceedings for the International Conference on Laser '78 for Optical and Quantum Electronics, STS Press, McLean, Virginia, 1979*)

of adaptive optics will not be discussed here. *See* ADAPTIVE OPTICS.

Fundamental properties. Optical phase conjugation is a process by which a light beam interacting in a nonlinear material is reflected in such a manner as to retrace its optical path. As **Fig. 1** shows, the image-transformation properties of this reflection are radically different from those of a conventional mirror. The incoming rays and those reflected by a conventional mirror (Fig. 1*a*) are related by reversal of the component of the wave vector $\vec{k}$ which is normal to the mirror surface. Thus a light beam can be arbitrarily redirected by adjusting the orientation of a conventional mirror. In contrast, a phase-conjugate reflector (Fig. 1*b*) inverts the vector quantity $\vec{k}$ so that, regardless of the orientation of the device, the reflected conjugate light beam exactly retraces the path of the incident beam. This retracing occurs even though an aberrator (such as a piece of broken glass) may be in the path of the incident beam. Looking into a conventional mirror, one would see one's own face, whereas looking into a phase-conjugate mirror, one would see only the pupil of the eye. This is because any light emanating from, say, one's chin would be reversed by the phase conjugator and would return to the chin, thereby missing the viewer's eye. A simple extension of the arrangement in Fig. 1*b* indicates that the phase conjugator will reflect a diverging beam as a converging one, and vice versa. These new and remarkable image-transformation properties (even in the presence of a distorting optical element) open the door to many potential applications in areas such as laser fusion, atmospheric propagation, fiber-optic propagation, image restoration, real-time holography, optical data processing, nonlinear microscopy, laser resonator design, and high-resolution nonlinear spectroscopy. *See* MIRROR OPTICS.

Optical phase conjugation techniques. Optical phase conjugation can be obtained in many nonlinear materials (materials whose optical properties are affected by strong applied optical fields). The response of the material may permit beams to combine in such a way as to generate a new beam that is the phase-conjugate of one of the input beams. Processes associated with degenerate four-wave mixing, scattering from saturated resonances, stimulated Brillouin scattering, stimulated Raman scattering, photoreactive phenomena, surface phenomena, and thermal scattering have all been utilized to generate optical phase-conjugate reflections.

Kerr-like degenerate four-wave mixing. As shown in **Fig. 2**, two strong counterpropagating (pump) beams with *k*-vectors $\vec{k}_1$ and $\vec{k}_2$ (at frequency ω) are directed to set up a standing wave in a clear material whose index of refraction varies linearly with intensity. This arrangement provides the conditions in which a third (probe) beam with *k*-vector $\vec{k}_p$, also at frequency ω, incident upon the material from any direction would result in a fourth beam with *k*-vector $\vec{k}_c$ being emitted in the sample precisely retracing the third one. (The term degenerate indicates that all beams have exactly the same frequency.) In this case, phase matching (even in the birefringent materials) is obtained independent of the angle between $\vec{k}_p$ and $\vec{k}_1$. The electric field of the conjugate wave E_c is given by the equation below, where δn is the change in the index of refraction induced by one strong counterpropagating wave, λ_0 is the free-space optical wavelength, and l is the length over which the probe beam overlaps the conjugation region. The conjugate reflectivity is defined as the ratio of reflected and incident intensities, which is the square of the above tangent function. The essential feature of phase conjugation is that E_c is proportional to the complex conjugate of E_p. Although degenerate four-wave mixing is a nonlinear optical effect, it is a linear function of the field one wishes to conjugate. This means that a superposition of E_p's will generate a corresponding superposition of E_c's; therefore faithful image reconstruction is possible. *See* KERR EFFECT.

$$E_c = E_p^* \tan\left(\frac{2\pi}{\lambda_0}\delta n l\right)$$

To visualize the degenerate four-wave mixing effect, consider first the interaction of the weak probe wave with pump wave number two. The amount by which the index of refraction changes is proportional to the quantity $(E_p + E_2)^2$, and the cross term corresponds to a phase grating (periodic phase disturbance) appropriately oriented to scatter the pump wave number one into the $\vec{k}_c$ direction. Similarly, the very same scattering process occurs with the roles of two pump waves reversed. Creating these phase gratings can be thought of as "writing" of a hologram, and the subsequent scattering can be thought of as "reading" the hologram. Thus the four-wave mixing process is equivalent to volume holography in which

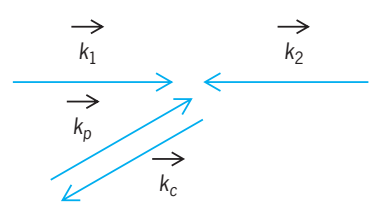

Fig. 2. Geometry of *k*-vectors for optical phase conjugation using degenerate four-wave mixing.

the writing and reading are done simultaneously. *See* DIFFRACTION GRATING; HOLOGRAPHY.

Conjugation using saturated resonances. Instead of using a clear material, as outlined above, the same four-wave mixing beam geometry can be set up in an absorbing or amplifying medium partially (or totally) saturated by the pump waves. When the frequency of the light is equal to the resonance frequency of the transition, the induced disturbance corresponds to amplitude gratings which couple the four waves. When the single frequency of the four waves differs slightly from the resonance frequency of the transition, both amplitude gratings and index gratings become involved. Because of the complex nature of the resonant saturation process, the simple $\tan^2 [(2\pi/\lambda_0)(\delta n 2l)]$ expression for the conjugate reflectivity is no longer valid. Instead, this effect is maximized when the intensities of the pump waves are about equal to the intensity which saturates the transition.

Stimulated Brillouin and Raman scattering. Earliest demonstrations of optical phase conjugation were performed by focusing an intense optical beam into a waveguide containing materials that exhibit backward stimulated Brillouin scattering. More recently, this technique has been extended to include the backward stimulated Raman effect. In both cases, the conjugate is shifted by the frequencies characteristic of the effect. *See* RAMAN EFFECT.

Backward stimulated Brillouin scattering involves an inelastic process where an intense laser beam in a clear material is backscattered through the production of optical phonons. Here, the incoming and outgoing light waves interfere to beat at the frequency and propagation vector of the sound wave, and the incoming light wave and the sound wave interfere to produce a radiating nonlinear polarization at the frequency and propagation vector of the outgoing light wave. The production of the sound wave involves the electrostrictive effect in which the Brillouin scattering medium is forced into regions of high electric field. The scattering of the light by the sound wave can be viewed as Bragg scattering from the sound wave grating with a Doppler shift arising because the sound wave is moving. The coupled process of stimulated backward Brillouin scattering can then be understood by considering that the forward-going beam scatters off of a small amount of sound wave to make more backward-going light, and at the same time the forward-going beam and a small amount of backward-going beam interfere to promote the growth of the sound wave. This doubly coupled effect rapidly depletes the incoming laser beam while producing the backward beam and while leaving a strong sound wave propagating in the medium. *See* DOPPLER EFFECT; ELECTROSTRICTION; LATTICE VIBRATIONS; SCATTERING OF ELECTROMAGNETIC RADIATION.

Photorefractive effects. A degenerate four-wave mixing geometry can be set up in a nonlinear crystal which exhibits the photorefractive effect. In such a material, the presence of light releases charges which migrate until they are trapped. If the light is in the form of an interference pattern, those migrating charges which get trapped in the peaks are more likely to be re-released than would those which become recombined in the valleys. This process leads to the charges being forced into the valleys of the interference pattern. This periodic charge separation sets up large periodic electric fields which, through the Pockels effect, produce a spatially periodic index of refraction. Hence a light-wave grating is converted into an index grating which can then scatter a pump wave into the conjugate direction. This process requires very little power, and therefore can be studied without the use of high-power lasers. *See* ELECTRON-HOLE RECOMBINATION; ELECTROOPTICS; INTERFERENCE OF WAVES; TRAPS IN SOLIDS.

Surface phase conjugation. Here a strong pump wave is directed normal to a surface which exhibits nonlinear behavior. Phenomena which can couple the light waves include electrostriction, heating, damage, phase changes, surface charge production, and liquid crystal effects. In general, the surface phase-conjugation process is equivalent to two-dimensional (thin) holography. *See* LIQUID CRYSTALS; SURFACE PHYSICS.

Thermal scattering effects. Here an optical interference pattern in a slightly absorbing material will be converted to an index grating if the material has a temperature-dependent index of refraction. In the four-wave mixing geometry, these thermal gratings can then scatter pump waves into the conjugate wave direction.

Practical applications. Many practical applications of optical conjugators utilize their unusual image-transformation properties. Because the conjugation effect is not impaired by interposition of an aberrating material in the beam, the effect can be used to repair the damage done to the beam by otherwise unavoidable aberrations. This technique can be applied to improving the output beam quality of laser systems which contain optical phase inhomogeneities or imperfect optical components. In a laser, one of the two mirrors could be replaced by a phase-conjugating mirror, or in laser amplifier systems, a phase-conjugate reflector could be used to reflect the beam back through the amplifier in a double-pass configuration. In both cases, the optical-beam quality would not be degraded by inhomogeneities in the amplifying medium, by deformations or imperfections in optical elements, windows, mirrors, and so forth, or by accidental misalignment of optical elements. *See* IMAGE PROCESSING.

Aiming a laser beam through an imperfect medium to strike a distant target may be another application. The imperfect medium may be turbulent air, an air-water interface, or the focusing mirror in a laser-fusion experiment. Instead of conventional aiming approaches, one could envision a phase-conjugation approach in which the target would be irradiated first with a weak diffuse probe beam. The glint returning from the target would pass through the imperfect medium, and through the laser amplifier system, and would then strike a conjugator. The conjugate beam would essentially contain all the information needed to strike the target after

passing through both the amplifier and the imperfect medium a second time. Just as imperfections between the laser and the target would not impair the results, neither would problems associated with imperfections in the elements within the laser amplifier. *See* NUCLEAR FUSION.

Other applications are based upon the fact that four-wave mixing conjugation is a narrow-band mirror that is tunable by varying the frequency of the pump waves. There are also applications to fiber-optic communications. For example, a spatial image could be reconstructed by the conjugation process after having been "scrambled" during passage through a multimode fiber. Also, a fiber-optic communication network is limited in bandwidth by the available pulse rate; this rate is determined to a large extent by the dispersive temporal spreading of each pulse as it propagates down the fiber. The time-reversal aspect of phase conjugation could undo the spreading associated with linear dispersion and could therefore increase the possible data rate. *See* LASER; NONLINEAR OPTICS; OPTICAL COMMUNICATIONS; OPTICAL FIBERS; OPTICAL INFORMATION SYSTEMS.

Robert A. Fisher; Barry J. Feldman

Bibliography. R. A. Fisher (ed.), *Optical Phase Conjugation*, 1983; M. Gower and D. Proch, *Optical Phase Conjugation*, 1994; J.-I. Sakai, *Phase Conjugate Optics*, 1992.

Optical prism

A simple component, made of a light-refracting and transparent material such as glass and bounded by two or more plane surfaces at an angle (**Fig. 1**), that is used in optical devices, especially to change the direction of light travel, to accomplish image rotation or inversion, and to disperse light into its constituent colors. Once light enters a prism, it can be reflected one or more times before it exits the prism.

A variety of prisms can be classified according to their function. Some prisms, such as the Dove prism, can be used to rotate an image and to change its parity. Image inversion by prisms in traditional binoculars is a typical application (**Fig. 2**). Some prisms take advantage of the phenomenon of total internal reflection to deviate light, such as the right-angle prism and the pentaprism used in single lens reflex cameras (**Fig. 3**). A thin prism is known as an optical wedge; it can be used to change slightly the direction of light travel, and therefore it can be used in pairs as an alignment device (**Fig. 4**). Optical wedges are also use in stereoscopic instruments to allow the viewer to observe the three-dimensional effect without forcing the eyes to point in different directions. A variable wedge can be integrated into a commercial pair of binoculars to stabilize the line of sight in the presence of the user's slight hand

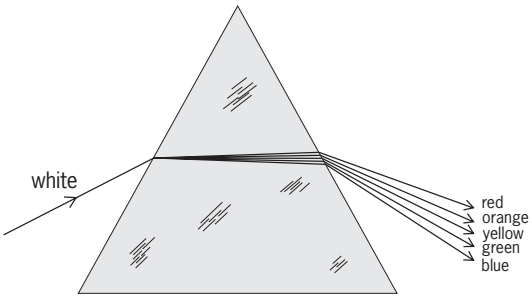

Fig. 1. Optical prism, consisting of a refracting medium bounded by two surfaces. Dispersion of light into its component colors is shown. (*After W. J. Smith, Modern Optical Engineering, 3d ed., SPIE Press–McGraw-Hill, 2000*)

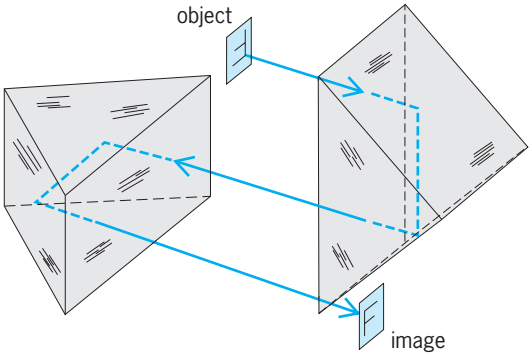

Fig. 2. System of two Porro prisms, used to invert the image in traditional binoculars. (*After R. E. Hopkins and R. Hanau, U.S. Military Handbook for Optical Design, republished by Sinclair Optics, 1987*)

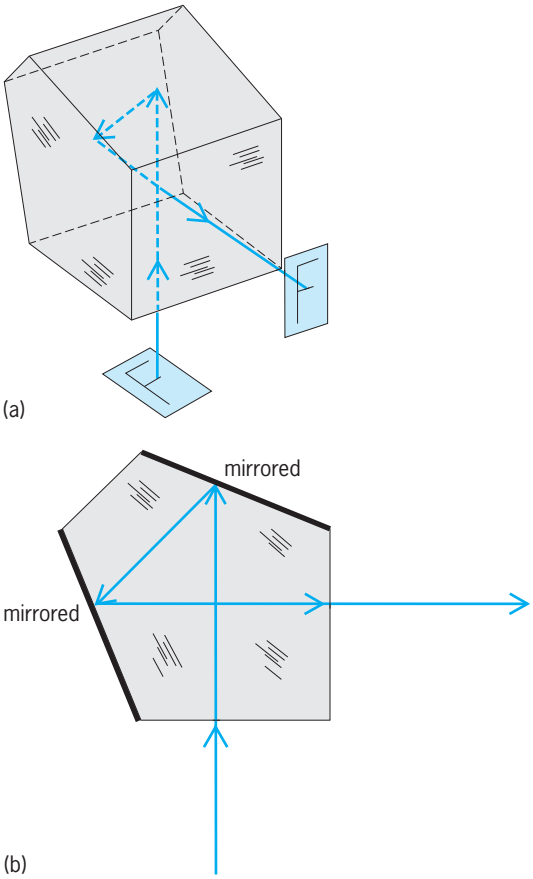

Fig. 3. Pentaprism, used in single lens reflex cameras. (*a*) Perspective view, showing how the prism deviates light through an angle of 90° without inverting or reverting the image. (*b*) Ray diagram. (*After R. E. Hopkins and R. Hanau, U.S. Military Handbook for Optical Design, republished by Sinclair Optics, 1987*)

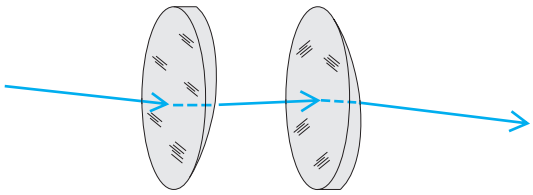

Fig. 4. Set of two thin prisms, known as optical wedges, which can be used as an alignment device. (*After P. R. Yoder, Jr., Non-image forming optical components, in Geometrical Optics, SPIE Proc., 531:206–220, 1985*)

movements. Other prisms such as corner-cubes can be used to reflect light backward, and are fabricated in arrays for car and bicycle retroreflectors. *See* BINOCULARS; MIRROR OPTICS; PERISCOPE; RANGEFINDER (OPTICS); REFLECTION OF ELECTROMAGNETIC RADIATION.

Light dispersion. When light enters at an angle to the face of a prism, according to Snell's law it is refracted. Since the index of refraction depends on the wavelength, the light is refracted at different angles and therefore it is dispersed into a spectrum of colors. The blue color is refracted more than the red. When light reaches the second face of the prism, it is refracted again and the initial dispersion can be added to or canceled, depending on the prism angle. An important application of a prism is to disperse light (Fig. 1). A combination of prisms in tandem can increase the amount of light dispersion. Dispersing prisms have been used in monochromators and spectroscopic instruments. With two prisms of different materials, it is possible to obtain light deviation without dispersion (an achromatic prism) or dispersion without deviation. *See* DISPERSION (RADIATION); REFRACTION OF WAVES; SPECTROSCOPY.

Image manipulation. An anamorphic magnification (magnification whose magnitude is different in each of two perpendicular directions) can be imparted to an image by prisms. For example, the aspect ratio of a beam becomes different as it enters and exits a system of Brewster prisms, and a pair of such prisms is often used to transform the cross section of a diode laser beam from elliptical to circular. (In a Brewster prism, light is incident at Brewster's angle, so that reflection losses for plane-polarized light, that is, light vibrating in the plane of incidence, are minimized.) The angular magnification will also be anamorphic. A prism can also introduce aberrations into a beam of light, depending on how the beam enters and exits the prism. Some prisms can be considered as parallel plates of glass, and therefore they can shift axially the position of an image. *See* ABERRATION (OPTICS); GEOMETRICAL OPTICS; POLARIZED LIGHT.

Other applications. Prisms find many other applications. For example, they have been used to measure the index of refraction, taking advantage of the phenomenon of minimum deviation; they find application in scanning optical systems; and they are used in fingerprint reading devices. *See* FINGERPRINT.

Jose M. Sasian

Bibliography. E. Hecht, A. Zajac, and K. Guardino, *Optics*, 3d ed., Addison-Wesley, 1997; W. J. Smith, *Modern Optical Engineering*, 3d ed., SPIE Press–McGraw-Hill, 2000.

Optical projection systems

Optical projection is the process whereby a real image of a suitably illuminated object is formed by an optical system in such a manner that it can be viewed, photographed, or otherwise observed. Essential equipment in an optical projection system consists of a light source, a condenser, an object holder, a projection lens, and (usually) a screen on which the image is formed (**Fig. 1**).

The luminance of the image in the direction of observation will depend upon (1) the average luminance of the image of the light source as seen through the projection lens from the image point under consideration, (2) the solid angle subtended by the exit pupil of the projection lens at this image point, and (3) the reflective or transmissive characteristics of the screen. Usually it is desirable to have this luminance as high as possible. Therefore, with a given screen, lens, and projection distance, the best arrangement is to have the light source imaged in the projection lens, with its image filling the exit pupil as completely and as uniformly as possible.

The object is placed between the condenser and the projection lens. If transparent, it can be inserted directly in the light beam; however, it should be positioned, and the optical system should be so designed that it does not vignette (cut off) any of the image of the light source in the projection lens. If the object is opaque, an arrangement known as an epidiascope (**Fig. 2**) is used. A difficulty in the design of this system is to illuminate the object so that all portions

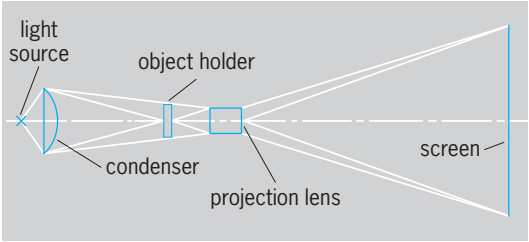

Fig. 1. Simple optical projection system.

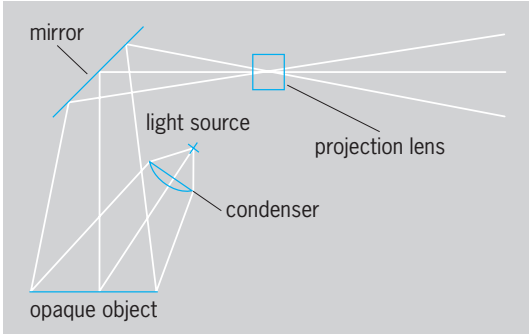

Fig. 2. An epidiascope, or system for projecting an image of an opaque object.

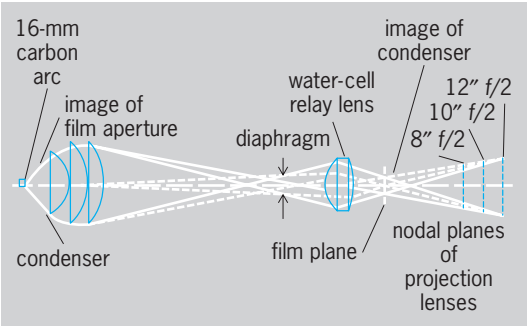

Fig. 3. Relay condenser system having water cell incorporated in second stage. 1″ = 25 mm; 16 mm = 0.63 in.

will show well in the projected image, without excessive highlights or glare.

If a small uniform source which radiates in accordance with Lambert's law is projected through a well-corrected lens to a screen which is perpendicular to the optic axis of the lens, maximum illuminance of the image will occur on this axis, and illuminance away from the axis will decrease in proportion to the fourth power of the cosine of the angle subtended with the axis at the projection lens. In practice, it is possible to design distortion into the condenser so that the illuminance is somewhat lower on the axis, and considerably higher away from the axis than is given by this fourth-power law. Acceptable illumination for most visual purposes can allow a falloff from center to side in the image of as much as 50%, particularly if the illuminance within a circle occupying one-half of the image area does not drop below 80% of the maximum value. *See* PHOTOMETRY.

Light source. Usually, either an incandescent or an arc lamp is used as the light source. To keep luminance high, incandescent projection lamps operate at such high temperatures that their life is comparatively short. Also, they must be well cooled; all except the smallest sizes require cooling fans for this purpose. Filaments are finely coiled and accurately supported, usually being carefully aligned in a prefocus base so that they will be precisely positioned in the optical system. Spacing between coils is such that a small spherical mirror can be used in back of the lamp to image the coils in the spaces between the coils, thus increasing usable light output nearly twofold. *See* LAMP.

When a highly uniform field is required, a lamp consisting of a small disk of ceramic material, heated to incandescence by radio-frequency induction, is available. With this, it is possible to maintain illuminance of a projection field of several square inches with a variation of only 2–3%. *See* INCANDESCENT LAMP.

Arc lamps are used when incandescent lamps cannot provide sufficient flux to illuminate the screen area satisfactorily. Carbon electrodes with special feed control, magnetic arc stabilization, and other devices are usually used to keep the arc as accurately positioned and as flicker-free as possible. The high-pressure xenon arc lamp is also used as a projection light source. It has a color which more accurately duplicates sunlight than the carbon arc. Its intensity is considerably higher than that of any incandescent lamp, and it avoids most of the problems attendant on the burning of carbons. A special power supply, which ordinarily incorporates a striking circuit, is needed. *See* ARC LAMP.

A shield between the arc and the object, called a douser, is used to protect the object while the arc is ignited and to provide a quick shutoff for the light beam. Often water cells and other heat-filtering devices are inserted in the beam to keep the heat on the object as low as possible.

Condenser. The condenser system is used to gather as much of the light from the source as possible and to redirect it through the projection lens. Both reflective and refractive systems are used. Reflectors can be of aluminum, although the better ones are of glass with aluminized coatings. They are usually elliptical in shape, with the light source at one focus and the image position in the projection lens at the other.

Refractive systems may be of heat-resistant glass or fused quartz. With arc lamps particularly, the condenser lens is very close to the source in order to provide the high magnification required if the image is to fill the projection lens. Usually, therefore, the condenser requires special cooling. In larger projectors, several elements are used to give the required magnification; these are often aspherical in shape to give the required light distribution.

A well-designed condenser can pick up light in a cone having a half-angle in excess of 50°. This means that, with a well-designed arc lamp or with an incandescent lamp using an auxiliary spherical mirror, more than one-third of the total luminous flux radiated by the source can be directed through the projector.

To obtain the high magnification required with arc sources and large-aperture lenses, a relay type of condenser (**Fig. 3**) may be used. This images the source

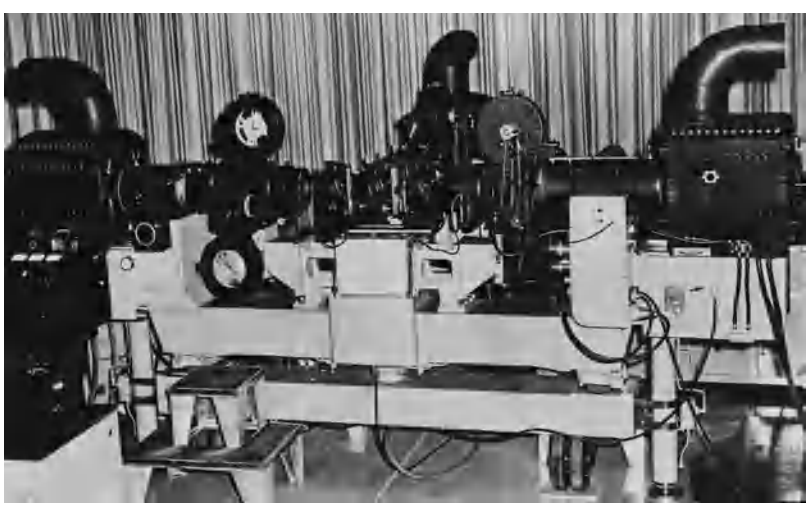

Fig. 4. Triple-head projector with three 250-A arc lamps and 12-in. (30-mm) *f*/2 projection lenses, incorporating iris diaphragms for intensity control. (*Paramount Pictures*)

beyond the first condenser system, and then uses a second lens to relay this image to the projection lens. This arrangement allows for better light-distribution control in the screen image with less waste of light at the object position. Also, an Inconel "cat's-eye" diaphragm at the first image point gives a convenient intensity control for the light beam. For additional information on condensers *see* OPTICAL MICROSCOPE

Object holder. The function of the object holder is to position the object exactly at the focal point of the projection lens. Slight motion or vibration will reduce image sharpness. Therefore, proper mechanical design of this part of the system is most important. When a succession of objects is to be projected, the holder must be able to position these objects precisely and clamp them firmly in a minimum time. Some designs have been suggested and a few tried which allow the object to move while the system projects a fixed image. However, this requires a motion in some portion of the optical system, which invariably reduces the quality of the projected image. Because of this, such systems have not found wide favor.

The object holder must also provide for protection of the object from unwanted portions of the beam, for cooling the object, where this may be required, and for accurately adjusting its position laterally with respect to the beam. In most systems, focusing of the projection lens is provided by moving the lens itself rather than by disturbing the position of the object longitudinally in the system.

Projection lens. The function of the projection lens is to produce the image of the object on the screen. Its design depends upon the use to be made of the projected image. As examples, a profile or contour projector requires a lens which is well corrected for the aberrations known as distortion and field curvature; an optical printer (or enlarger) requires a lens carefully designed to work at comparatively short conjugate focal distances; and picture projectors must preserve the quality which has been captured by the camera lens. *See* ABERRATION (OPTICS).

Since projection lenses usually transmit relatively intense light beams, they must be designed to be heat-resistant. Their surfaces are usually not cemented. Optical surfaces are coated to reduce reflection, but these coatings must be able to withstand heat without deterioration. Because camera lenses are not designed for this type of operation, they should not be employed as projection lenses unless the projector is specifically designed to use them.

An ideal position at which to control intensity is at the projector lens. Large-aperture lenses containing iris diaphragms are used in the large process projectors of motion picture studios (**Fig. 4**).

Projection screen. Usually a projection screen is used to redirect the light in the image for convenient observation. Exceptions are systems which project the image directly into other optical systems for further processing, or which form an aerial image which is to be viewed only from a localized position.

Screens may be either reflective or translucent, the latter type being used when the image is to be observed or photographed from the side away from the projector. An example is the so-called self-contained projector, which has the screen on the housing holding the other elements. Reflective screens may be matte, having characteristics approaching Lambert reflection; directional, producing an image which will appear brighter in the direction of specular reflection; or reflexive, directing most of the light back toward the projector and giving an image of relatively high luminance with low projection intensity, but with a very confined viewing angle. *See* CINEMATOGRAPHY.

Armin J. Hill

Optical pulses

Bursts of electromagnetic radiation of finite duration. Optical pulses are used to transmit information or to record the chronology of physical events. The simplest example is the photographic flash. This was probably first developed by early photographers who used flash powder that, when ignited, produced a short burst of intense light. This was followed by the flash lamp, in which a tube filled with an inert gas such as xenon is excited by a brief electrical pulse. A great advance in the creation of short optical pulses came with the invention of the laser. Lasers are now the most common and effective way of generating a variety of short optical pulses, of different durations, energies, and wavelengths. *See* LASER; STROBOSCOPIC PHOTOGRAPHY.

Q-switching. Pulses of millisecond (10^{-3} s) duration are very simply generated by mechanically modulating a constant light source such as a lamp or a continuous-wave laser. This can be done, for example, by placing a rotating disk with holes in it in front of the light source. Shorter laser pulses, of microsecond (10^{-6} s) or nanosecond (10^{-9} s) duration, are generated by using a technique known as Q-switching. A modulating device is incorporated inside the laser cavity that allows the buildup of the laser radiation inside the cavity and then switches it out in an instant. The modulating device is usually controlled by external electrical pulses. Semiconductor diode lasers, which are used to transmit information (voice or data) over a fiber-optic cable, are pumped by electricity and can be directly pulsed by applying to them a pulsed electrical signal. *See* OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

Mode locking. Ultrashort laser pulses, with durations of the order of picoseconds (1 ps = 10^{-12} s) or femtoseconds (1 fs = 10^{-15} s), are generated by using a general principle known as mode locking. Lasers that directly produce these pulses employ either active or passive mode locking. The Fourier transform relationship implies that an optical pulse must contain a certain number of frequency modes, or equivalently spectral components, which is inversely proportional to its time duration. Mode locking is a technique whereby several frequency modes of the laser structure are made to resonate simultaneously

and with a well-orchestrated relationship so as to form a short-duration pulse at the laser output.

In the case of active mode locking, the laser is stimulated with external pulses, either electrical or optical. For example, the excitation that drives the laser can be another short laser pulse, which has a repetition rate precisely adjusted to the length of the mode-locked laser. Another active technique consists of incorporating into the laser itself an optical element that modifies the laser's properties when the laser is stimulated with electrical pulses, again with a well-adjusted repetition rate.

In the case of passive mode locking, a conceptually more difficult technique to understand, the laser is not actively stimulated with some form of external pulses. Rather, the mode locking originates from the complex interaction of the various elements of the laser cavity. The laser can be pumped by a continuous source of energy, yet still produce ultrashort pulses of light. One of the more standard ways of realizing passive mode locking is by incorporating into the laser resonator a lossy material known as a saturable absorber, in addition to the laser gain medium. The saturable absorber absorbs the front edge of the pulse and then saturates, becoming transparent to the remainder of the pulse. The combination of the gain and the saturable absorber as the pulse makes round trips in the laser cavity leads to a stable short pulse.

There are other, and more exotic, forms of passive mode locking. Typically, lasers mode-locked in this way contain a nonlinear element that acts much as a saturable absorber would. The titanium:sapphire laser (a sapphire crystal doped with the trivalent ion Ti^{3+}) produces extremely short pulses over a wide range of wavelengths (700–1000 nm). The nonlinear interaction that leads to the passive mode locking is believed to originate in the titanium:sapphire crystal itself, which is in fact also the laser gain medium. Lasers fabricated from glass fibers similar to those used in fiber-optic networks can also be passively mode-locked and form very compact sources, as they can be pumped with miniature laser diodes. *See* FOURIER SERIES AND TRANSFORMS; NONLINEAR OPTICS.

Shortest durations. Pulses as short as 11 fs have been produced directly by a passively mode-locked titanium:sapphire laser. The titanium:sapphire laser has also allowed the extension of ultrashort optical pulses to other wavelength ranges, such as the near-infrared (2–10 μm). Dye lasers, based on organic dyes in solution, have achieved durations as short as 27 fs. Ultrashort diode laser pulses have been obtained by active and passive mode locking and produce pulses as short as a few hundred femtoseconds. They are more commonly operated so as to give rise to pulses in the picosecond range, appropriate for optical communication systems.

Optical pulses from many of these lasers can be shortened even further by pulse compression. This is accomplished by passing the pulses through a series of optical elements that include a nonlinear medium such as an optical fiber. Laser pulses as short as 6 fs have been generated in this fashion. The duration of these pulses is equivalent to three cycles of light at the center wavelength of the pulse, 620 nm. These pulses have a correspondingly large spectral bandwidth, about 60 nm, or one-tenth of the entire visible spectrum.

At present, no optical detectors can measure pulse durations of less than 1 ps. The only way, therefore, to accurately measure the duration of a subpicosecond optical pulse is to use the pulse itself. The technique used, autocorrelation, involves splitting the pulse into two equal parts and then recombining them with a varying time delay between the two pulses. The resulting sum signal, which is a function of the relative time delay between the two pulses, can then be analyzed to obtain the temporal duration of the pulse.

Time-resolved spectroscopy. The generation of ultrashort laser pulses has been motivated by the quest for ever better resolution in the study of the temporal evolution and dynamics of physical systems, events, and processes. Such laser pulses are capable of creating snapshots in time of many events that occur on the atomic or molecular scale, a technique known as time-resolved spectroscopy. This stroboscopic aspect of ultrashort laser pulses is their most important scientific application and is used in physics, engineering, chemistry, and biology. For example, ultrashort pulses can excite and take snapshots of molecular vibrations and deformations. They can track the passage of charge carriers through a microscopic semiconductor device. This ability to understand the dynamics of the more elemental building blocks of nature can in turn make it possible to build ever faster devices for use in information processing and information transmission, in addition to providing a better understanding of the physical world. *See* LASER PHOTOCHEMISTRY; LASER SPECTROSCOPY; OPTICAL INFORMATION SYSTEMS; ULTRAFAST MOLECULAR PROCESSES.

High-intensity pulses. Ultrashort pulses also produce extremely high instantaneous intensities of light, since the entire energy of the pulse is compressed into such a short time duration. High-intensity pulses may have important applications in x-ray laser production, high-resolution lithography, and fundamental physics at high intensities. The method used to amplify short pulses to very high intensities is known as chirped pulse amplification. It involves stretching out the pulse in time by spreading out its spectral components, followed by the amplification stage, and finally a compression of the pulse to close to its initial duration. This method avoids the deleterious effects that can accompany the amplification of an ultrashort pulse in a high-energy amplifier. Titanium:sapphire lasers amplifiers contribute significantly to the production of ultrahigh-intensity pulses. Pulses with a duration of 21 fs and an energy per pulse of 0.5 mJ have been demonstrated.

Philippe C. Becker

High-power laser interaction studies. The advent of short-pulse, high-power lasers has opened up new areas of research in atomic physics under high-field

or high-density conditions. In particular, the interaction of high-power lasers with appropriate targets can give rise to incoherent radiation sources for x-ray spectroscopy and x-ray diffraction studies, as well as to coherent soft x-ray sources for x-ray holography. Such studies can enhance understanding of the dynamical properties of material or biological systems.

High-power lasers (in the ultraviolet, visible, or infrared parts of the spectrum) have peak powers of the order of terawatts (1 TW = 10^{12} W). Most of them fall into two groups. Lasers of pulse duration of about 1 ns have been developed primarily for nuclear fusion studies but are also used to generate coherent and incoherent x-ray sources. Lower-energy, smaller lasers of pulse duration of order 1 ps or shorter can give rise to very high fields (by focusing the beam to a very small spot), thereby accessing a new regime of atomic spectroscopy.

High-density physics. In nuclear fusion studies, a multibeam high-power laser is used to spherically compress a target core to densities of the order of 10^{24} electrons per cubic centimeter and temperatures above 1 keV (10^7 K). Under these conditions, similar to those prevailing in stellar interiors, x-ray spectral lines of core elements (heavier than hydrogen) are strongly broadened. This broadening is due to the electric fields produced by neighboring plasma particles (the Stark effect), and it can be used to deduce the plasma density. A stronger effect is associated with the compressed shell surrounding the core, which typically achieves high densities at lower temperatures. Here the plasma ions are in the strongly coupled regime. The electron energy, for the most part, conforms to the Thomas-Fermi distribution. Such experiments can thus be used to study the properties (such as transport and equation of state) of Fermi-degenerate plasmas; a similar state (at much higher densities) exists in white dwarf stars. *See* FERMI-DIRAC STATISTICS; NUCLEAR FUSION; STARK EFFECT; STELLAR EVOLUTION; WHITE DWARF STAR.

Multiphoton ionization. When a high-power, very short-pulse laser is focused onto a low-pressure gas, the high laser flux at focus (above 10^{13} W/cm^2) corresponds to a high-intensity radiation field, and the low gas density isolates the radiation interaction with individual atoms from interparticle interactions. Experiments show a phenomenon termed above-threshold ionization (ATI): the escaped photoelectrons have discrete energies separated by the photon energy of the laser. Each successive peak corresponds to a multiphoton ionization event involving the absorption of one additional photon (generally, a nonresonant process, insensitive to atomic structure). Peaks corresponding to the simultaneous absorption of tens of photons have been observed. At much higher flux levels (above 10^{15} W/cm^2), the charge state of escaping ions indicates the simultaneous absorption of hundreds of photons. When subpicosecond pulses are used, the above-threshold ionization peaks are split into components, related to the fact that high-lying, Rydberg states are shifted (upward) in the presence of a high field. As the intensity changes during the laser pulse, various Rydberg states move into resonance with the total energy of the absorbed photons (corresponding to a given above-threshold ionization peak), thus enhancing the photoionization and giving rise to the component peaks. The short pulse assures that the energy of the escaping photoelectron is not modified by the field. This is therefore a method of measuring atomic level positions in the presence of high radiation fields. *See* ATOMIC STRUCTURE AND SPECTRA; RYDBERG ATOM.

X-ray holography and diffraction. Soft x-ray lasers can be realized by focusing a high-power laser into an elongated focal line on an appropriate target. The high-power laser can form a plasma of the desired ionization state in which plasma electrons or plasma photons can excite population inversion between pairs of ionic energy levels. Spontaneous emission of transitions between such levels is then amplified along the direction of the focal line. Significant coherent output has been achieved at wavelengths from above 20 nm down to below 10 nm (with pulses of energy up to 1 millijoule at the longer end of this range). Studies have aimed at improving the spatial coherence and extending lasing to wavelengths below 5 nm, where the scattering contrast between water and proteins is maximal. These sources can be used for dynamical, in-line holography of hydrated (living) biological samples, with spatial resolution better than 0.1 micrometer. *See* HOLOGRAPHY.

In other examples of laser-based studies, the intense, incoherent x-ray emission from nanosecond-laser-irradiated targets was used for dynamical structure studies of living samples by diffracting off periodic protein molecules; the electron pulse from a picosecond-laser irradiated photocathode was used to dynamically study the annealing of metals by de Broglie diffraction; and a repetitive, picosecond-pulse laser was used to measure the fluorescence decay time of chlorophyll macromolecules. *See* ELECTRON DIFFRACTION; FLUORESCENCE; X-RAY DIFFRACTION.

Plasma-wave spectroscopy. When a high-power laser impinges on a solid target, forming a plasma, a rich spectrum of inelastically scattered radiation is observed. The frequencies in this spectrum are various combinations of the laser frequency and frequencies of natural modes or waves excited in the plasma by the laser itself. These include electron plasma waves (giving rise to Raman scattering, so called in analogy to the Raman scattering from molecular energy levels) and ionic waves (giving rise to Brillouin scattering). The density gradient, characteristic of plasma expanding from an irradiated target, permits the excitation of a wide range of frequencies, enriching the spectrum. The characteristics of such spectra are a source of information on the properties of the laser-heated plasma, in particular the temperature and the density gradient. *See* PLASMA (PHYSICS); RAMAN EFFECT; SCATTERING OF ELECTROMAGNETIC RADIATION.

Relativistic and quantum effects. At moderate flux levels (less than 10^{15} W/cm^2), a free electron oscillates harmonically in a laser field as a mass on a spring. At higher flux levels, the electron's oscillatory

velocity in the field approaches the speed of light and its motion becomes anharmonic. Its accompanying radiation includes the second and higher harmonics of the laser frequency. For counterpropagating electron-laser collisions, this process is multiphoton Compton scattering. Under these conditions, the electric field of the laser is Lorentz-boosted in the rest frame of the electrons. With high-intensity lasers (flux greater than 10^{18} W/cm^2) and relativistic electron beams (energy greater than 50 GeV), the electric field can be greater than the Schwinger critical field, which is the electric field (1.3×10^{16} V/cm) at which an electron is accelerated to an energy equivalent to its rest mass in one Compton wavelength. *See* ANHARMONIC OSCILLATOR; COMPTON EFFECT; LORENTZ TRANSFORMATIONS; QUANTUM ELECTRODYNAMICS; RELATIVITY. Robert L. McCrory

Bibliography. V. R. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Quantum Electrodynamics*, 2d ed., 1982; R. W. Falcone and J. Kirz (eds.), *Short Wavelength Coherent Radiation: Generation and Applications*, Optical Society of America, vol. 2, 1988; J. Hermann and B. Wilhelmi, *Lasers for Ultrashort Pulses*, 1987; H. Hora and G. H. Miley (eds.), *Laser Interactions and Related Plasma Phenomena*, vols. 6–10, 1984–1993; W. Kaiser (ed.), *Ultrashort Laser Pulses: Generation and Applications*, 2d ed., 1993; W. L. Kruer, *The Physics of Laser Plasma Interactions*, 1988.

Optical pumping

The process of causing strong deviations from thermal equilibrium populations of selected quantized states of different energy in atomic or molecular systems by the use of optical radiation (that is, light of wavelengths in or near the visible spectrum), called the pumping radiation.

At thermal equilibrium at temperature T K, the relative numbers of atoms, N_2/N_1, in quantized levels E_2 and E_1, respectively, where E_2 is the higher, are given by $N_2/N_1 = e^{-(E_2 - E_1)/kT}$, where k is Boltzmann's constant. The number of atoms in the higher level is, at equilibrium, always less than that in the lower, and as the energy difference between the two levels increases the number in the higher level becomes very small indeed. By exposing a suitable system to optical radiation, one can, so to speak, pump atoms from a lower state to an upper state so as greatly to increase the number of atoms in the upper state above the equilibrium value. *See* ENERGY LEVEL (QUANTUM MECHANICS).

In an early application of the principle, the levels E_2 and E_1 were not far apart, so that the equilibrium populations of the atoms in the two levels were not greatly different. The system was chosen to possess a third level E_3, accessible from E_1 but not from E_2 by the absorption of monochromatic visible light. (The states involved were the paramagnetic Zeeman components of atomic states, and the necessary selectivity of the transitions excitable by the visible light was secured by appropriate choice of the state of polarization of this light.) The visible light excites atoms from E_1 to E_3, from which they return, with spontaneous emission, with about equal probability, to the lower states E_2 and E_1. After a period of time, provided there has been sufficiently intense excitation by the visible light, most of the atoms are in state E_2 and few are in the lower state E_1—atoms have been pumped from E_1 to E_2 by way of the highly excited state E_3. *See* ATOMIC STRUCTURE AND SPECTRA; ZEEMAN EFFECT.

Optical pumping is vital for light amplification by stimulated emission in an important class of lasers. For example, the action of the ruby laser involves the fluorescent emission of red light by a transition from an excited level E_2 to the ground level E_1. In this case E_2 is relatively high above E_1 and the equilibrium population of E_2 is practically zero. Amplification of the red light by laser action requires that N_2 exceed N_1 (population inversion). The inversion is accomplished by intense green and violet light from an external source which excites the chromium ion in the ruby to a band of levels, E_3, above E_2. From E_3 the ion rapidly drops without radiation to E_2, in which its lifetime is relatively long for an excited state. Sufficiently intense pumping forces more luminescent ions into E_2 by way of the E_3 levels than remain in the ground state E_1, and amplification of the red emission of the ruby by stimulated emission can then occur. *See* LASER. William West

Bibliography. O. Svelto, *Principles of Lasers*, 4th ed., 1998; J. T. Verdeyen, *Laser Electronics*, 3d ed., 1993; A. Yariv, *Optical Electronics*, 4th., 1991.

Optical recording

The process of recording signals on a medium through the use of light, so that the signals may be reproduced at a subsequent time. Photographic film has been widely used as the medium, but in the late 1970s development of another medium, the optical disk, was undertaken. The introduction of the laser as a light source greatly improves the quality of reproduced signals. Digital storage techniques make it possible to obtain extremely high-fidelity reproduction of sound signals in optical disk recording systems. This article first describes optical film recording and then some of the more modern optical recording methods.

Optical Film Recording of Sound

Optical film recording is also termed motion picture recording or photographic recording. A sound motion picture recording system consists basically of a modulator for producing a modulated light beam and a mechanism for moving a light-sensitive photographic film relative to the light beam and thereby recording signals on the film corresponding to the electrical signals. A sound motion picture reproducing system is basically a combination of a light source, an optical system, a photoelectric cell, and a mechanism for moving a film carrying an optical record by means of which the recorded photographic

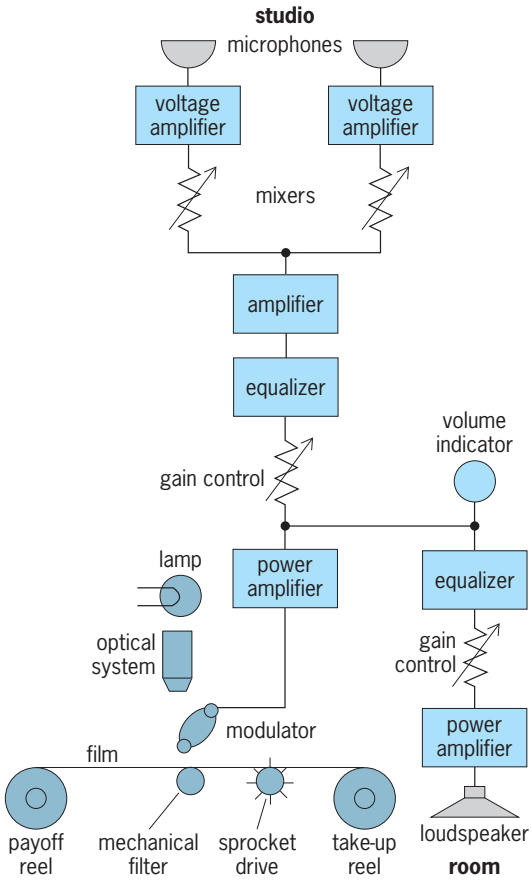

Fig. 1. Schematic arrangement of apparatus in a complete optical sound motion picture recording system.

variations are converted into electrical signals of approximately similar form.

Recording system. In a monophonic sound motion picture recording system (**Fig. 1**), the output of each microphone is amplified and fed to a mixer, a device having two or more inputs and a common output. If more than one microphone is used, such as a microphone for each of two actors, the outputs of the microphones may be adjusted for the proper balance by means of the mixers. An electronic compressor is used to reduce the amplitude range to that suitable for reproduction in the home. An equalizer provides the standard motion picture recording characteristic. A gain control provides means for controlling the overall signal level fed to the power amplifer. A light modulator, actuated by the amplifer, records a photographic image upon the film corresponding to the electrical input. The narrow bands used for the sound record on motion picture film are called sound tracks. A monitoring system consisting of a volume indicator, complementary equalizer, gain control, power amplifier, and loudspeaker or headphone is used to control the recording operation. *See* SOUND RECORDING.

Modulator. In the variable-area recording system (**Fig. 2**), the transmitted light amplitude is a function of the amount of unexposed area in the positive print. This type of sound track is produced by means of a mirror galvanometer which varies the width of the light slit under which the film passes. A triangular aperture is uniformly illuminated by means of a lamp and lens system. The image of this triangular aperture is reflected by the galvanometer mirror focused on a mechanical slit, which in turn is focused on the film. The triangular light image on the mechanical slit moves up and down on the mechanical slit. The result is that the width of the exposed portion of the negative sound track corresponds to the rotational vibrations of the galvanometer. In the positive record, the width of the unexposed portion corresponds to the signal. *See* GALVANOMETER.

A variable-density system has also been used, but it has been largely replaced by the variable-area system. The reproducer can play either one.

Recording film transport. The film-transport mechanism used in recording sound on film consists of a positive drive of the perforated film and a constant-speed drive of the film where the modulated light beam strikes the film. Positive drive of the film is obtained by the sprocket drive, which is interlocked with the camera drive so that synchronism of picture and sound will be obtained. When the film passes over the sprocket drive, variations in the motion of the film at the sprocket-hole frequency are produced. These variations in the film speed must be removed at the recording point to eliminate spurious frequency modulation of the image on the film. Uniform speed

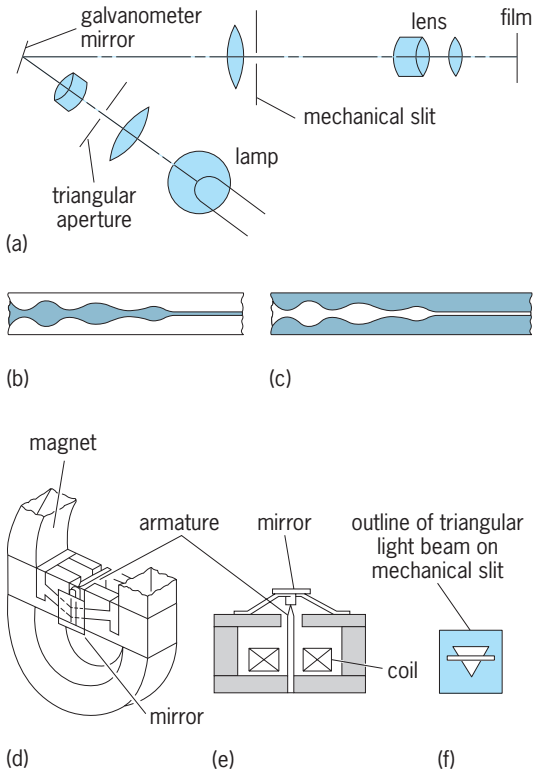

Fig. 2. Elements of a variable-area sound motion picture film recording system. Transmitted light amplitude is a function of the amount of unexposed area in the positive print. (*a*) Schematic of a recording system. (*b*, *c*) Negative and positive sound tracks. (*d*) Perspective view. (*e*) Sectional view. (*f*) Mechanical slit.

at the recording point is provided by a mechanical contrivance called a filter. It is located between the sprocket drive and recording point and consists of the inertia of the recording drum and the compliance of the film between the recording drum and the sprocket drive. The recording drum is driven by a magnetic system from the motor which drives the sprocket and thereby imparts a slight amount of drive to the film. The magnetic drive isolates the variations in the rotational speed of the motor drive from the rotating drum. The combination of the isolating filter and magnetic drive provides a system with very uniform motion of the surface of the drum. The image of the modulator is focused on the film while the film is in contact with the drum.

Film and sound track. In the recording of sound motion pictures, the picture and sound are recorded on separate photographic films. The camera and sound recorder must be synchronized. This is accomplished by the use of an interlock system between the camera and sound recorder and the use of perforated film in the form of sprocket holes along the two edges of the film for both the camera and sound recorder. The sound track on 35-mm film occupies a space about 0.1 in. (2.5 mm) wide just inside the sprocket holes.

Film developing and printing. The mass production of motion picture positive prints proceeds in a series of steps. The negative record is developed, and then the required number of positive prints of both picture and sound is printed from the negative record. These positive records are developed and are used for sound reproduction and picture projection in the theater.

Reproducing system. In a sound motion picture reproducing system (**Fig. 3**), the first element is the optical system, consisting of a lamp and a lens arrangement which produces an illuminated slit of light upon the film. The light beam passes through the film and falls upon a photoelectric cell. When the film is pulled past the slit, the variations in light, which are due to the variable-density or variable-area recording on the film, fall upon the photoelectric cell and are converted into the corresponding electrical variations. The output of the photoelectric cell is fed to an amplifier followed by a filter, which is used to cut the ground noise (residual noise in the absence of the signal) due to the film above the upper limit of reproduction, and by equalizers, which are used to adjust the frequency characteristic to that suitable for the best sound reproduction in the theater. A volume control (gain control) is used for adjusting the level of sound output. The output of a power amplifier feeds stage loudspeakers, located behind the screen, and a monitoring loudspeaker. Except for the stage loudspeakers, the entire equipment including the monitoring loudspeaker is located in the projection booth. *See* CINEMATOGRAPHY; SOUND-REPRODUCING SYSTEMS.

Optical electronic reproducer. In a motion picture film sound-reproducing system, the light source, in the form of an incandescent lamp, is focused upon a mechanical slit by means of a condensing lens. The mechanical slit in turn is focused on the negative film. The height of the image on the film is usually about 0.00075 in. (19 micrometers). Under these conditions the amount of light which impinges upon a photocell behind the flim is proportional to the unexposed portion of the sound track in variable-area recording. When the film is in motion, the resultant light undulations which fall upon the photocell correspond to the voltage variations applied to the recording galvanometer. The voltage output of the photocell is proportional to the amount of light which falls upon the cathode.

Reproducing film transport. The film transport used in reproducing sound on photographic film consists of a positive drive of the perforated film and a constant-speed drive where the light passes through the film to the photoelectric cell. Positive drive of the film is obtained by means of two sprocket drives. The sprocket drives are geared with the positive picture drive so that a constant loop of film is maintained between the sound head and the picture head. The positive drive also ensures that the film speed in reproduction will be the same as that in recording. There is a loose loop of film between the picture head and the sound head, so that variations in the picture drive will not be imparted to the sound head.

After the film enters the sound head, it passes over a drum. The light beam of the reproducing

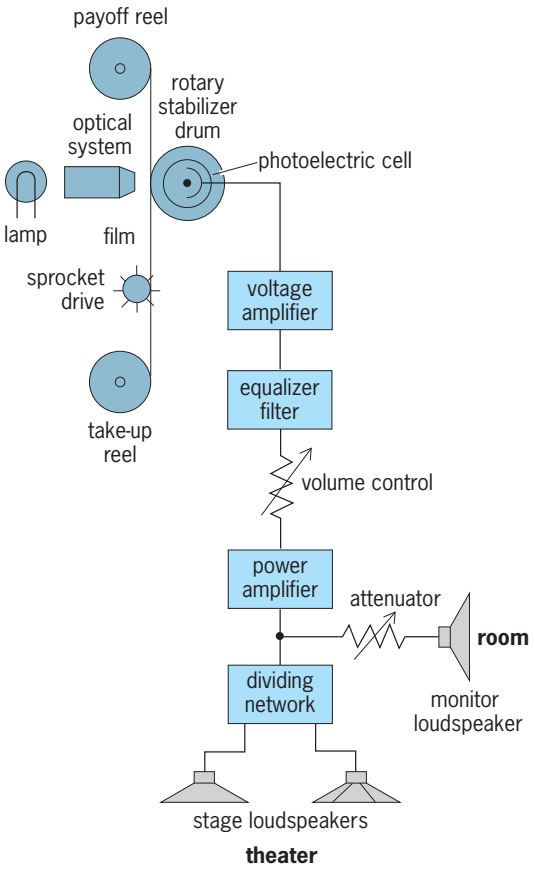

Fig. 3. Schematic arrangement of apparatus in a complete optical sound motion picture reproducing system.

system passes through the film to the photocell located inside the drum. The drum is driven by the first sprocket drive. The compliance of the film between the film and the sprocket provides a mechanical filter system and thereby reduces the sprocket-hole ripple at the drum. Under these conditions, the drum is rotated at a constant speed, and as a consequence the film will move past the light beam at a constant speed. The second sprocket isolates the take-up reel from the reproducing system.

Distortion and noise in reproduction. Most commonly, distortion in an optical reproducing system is due to the inherent nonlinear characteristics of the photographic process. This type of distortion can be reduced to a low value by the use of proper illumination in the recording or duplicating process. The developing processes must also be accurately controlled in order to achieve a low value of nonlinear distortion.

Noise in clean film is due to the inherent grain structure of the photographic medium. Scratches and foreign particles on the film add additional noise.

Another source of distortion is a nonuniform motion of the film in the recording and reproducing process. This is manifested as a frequency modulation of the reproduced signal and is termed flutter and wow.

Laser-beam film recording. In laser-beam film recording, an optical film system utilizes a laser as a light source, a combination of an acoustooptical modulator (AOM) and an acoustooptical deflector (AOD) instead of a galvanometer. A 100-kHz pulse-width modulation (PWM) circuit converts the audio input signal into a PWM signal. The laser beam is made to continuously scan the sound track area at right angles to the direction of the film transport. This is done by means of the acoustooptical deflector, which in turn is driven by a 100-kHz sawtooth signal. Simultaneously, the laser beam is pulse-width-modulated by means of the acoustooptical modulator, which is driven by a 100-kHz PWM signal. The scanning signal and the pulse-width-modulated signal combine and generate the variable-area sound track exposure on the film (**Fig. 4**). The traces of

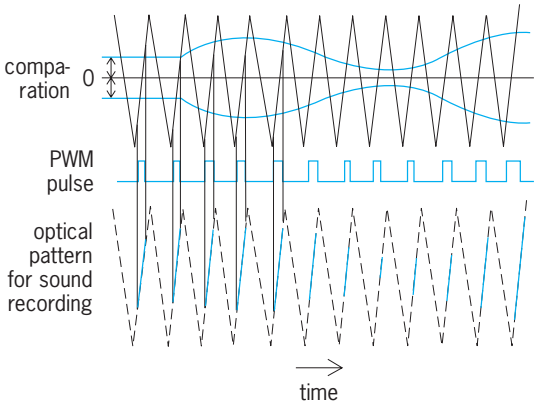

Fig. 4. Schematics of the sound track waveform generation of the laser sound recorder. Actual distance between successive scans is much smaller than shown. (*After T. Taneda et al., A high quality optical sound recording system using a scanned laser beam, SMPTE J., 89:95–97, 1980*)

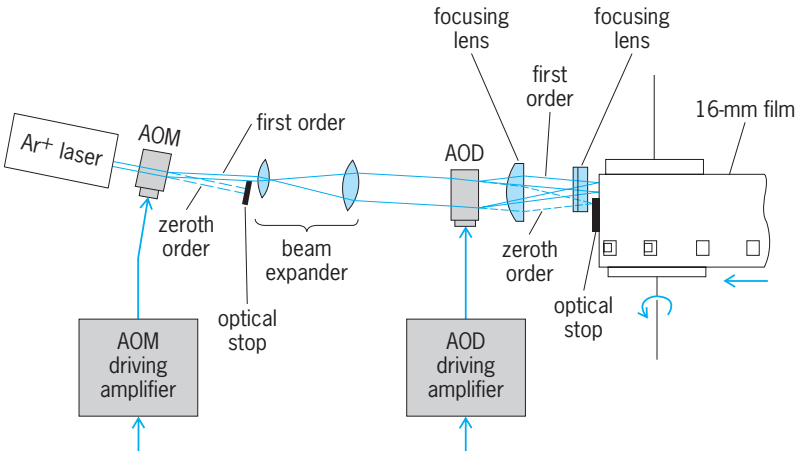

Fig. 5. Basic layout of the main optical components of the laser-beam sound recorder. (*After T. Taneda et al., A high quality optical sound recording system using a scanned laser beam, SMPTE J., 89:95–97, 1980*)

successive scans are fused into a pattern of variable-area recording.

The light from the laser (**Fig. 5**) is at first pulse-width-modulated in the acoustooptical modulator at 10-microsecond (100-kHz) intervals in response to the amplitude of the audio signal. As is normal for acoustooptical modulation, the modulated beam is separated into a zeroth-order (nondeflected) and a first-order (deflected) beam. Only the first-order beam is used, and the zeroth-order beam is blocked by an optical stop. Next, the diameter of the first-order beam is expanded to the size of the aperture. The beam is then deflected by the acoustooptical deflector and is again separated into a first-order and zeroth-order beam. Both beams are now converged to form spot images by means of a pair of cylindrical lenses. The first-order beam is focused on the film, while the zeroth-order beam is blocked by a fixed optical spot just in front of the film. Owing to the nature of the recording system, a wide-band frequency response can be obtained. *See* ACOUSTOOPTICS; LASER; PULSE MODULATION. H. Date

Optical Data Storage

Optical data storage involves placing information in a medium so that, when a light beam scans the medium, the reflected light can be used to recover the information. There are many forms of storage media, and many types of systems are used to scan data.

Storage principles. The storage of data on optical media involves two steps. First, data are recorded on the medium. Recorded data have the property that the reflected light is modulated compared to regions without data. Second, a light beam scans the medium, and modulation in the reflected light is used to detect the data pattern under the scanning spot. *See* MODULATION.

In the recording process (**Fig. 6**), an input stream of digital information is converted with an encoder and modulator into a drive signal for a laser source. The laser source emits an intense light beam that is

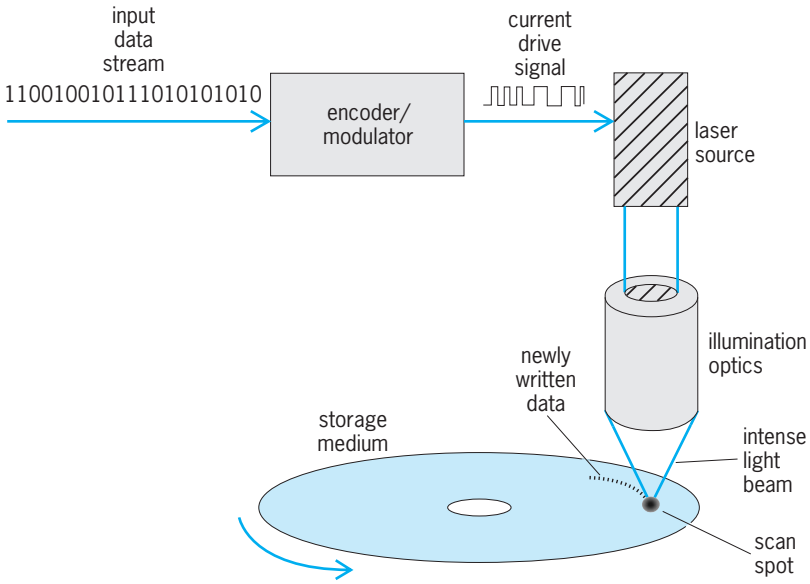

Fig. 6. Recording process for a simple optical medium. Writing data into the recording layers involves modulating an intense laser beam as the layers move under the scan spot.

directed and focused into the storage medium with illumination optics. As the medium moves under the scanning spot, energy from the intense scan spot is absorbed, and a small localized region heats up. The storage medium, under the influence of the heat, changes its reflective properties. Since the light beam is modulated in correspondence to the input data stream, a circular track of data marks is formed as the medium rotates. After every revolution, the path of the scan spot is changed slightly in radius to allow another track to be written.

In readout of the medium (**Fig. 7**), the laser is used at a constant output power level that will not heat the medium beyond its thermal writing threshold. The laser beam is directed through a beam splitter into the illumination optics, where the beam is focused into the medium. As the data to be read pass under the scan spot, the reflected light is modulated. The modulated light is collected by the illumination optics and directed by the beam splitter to the servo and data optics, which converge the light onto detectors. The detectors change the light modulation into current modulation that is amplified and decoded to produce the output data stream.

Configurations for optical media. Optical media can be produced in several different configurations. The most common configuration is the single-layer disk (**Fig. 8***a*), such as the compact disk (CD), where data are recorded in a single storage layer. A substrate provides mechanical support for the storage layer. The substrate is transparent and also provides a measure of contamination protection, because light is focused through the substrate and into the recording layer. Dust particles on the surface of the substrate only partially obscure the focused beam, so enough light can penetrate for adequate signal recovery. *See* COMPACT DISK.

In order to increase data capacity of the disk, several layers can be used (Fig. 8*b*). Each layer is partially transmitting, which allows a portion of the light to penetrate throughout the thickness of the layers. The scan spot isadjusted by refocusing the

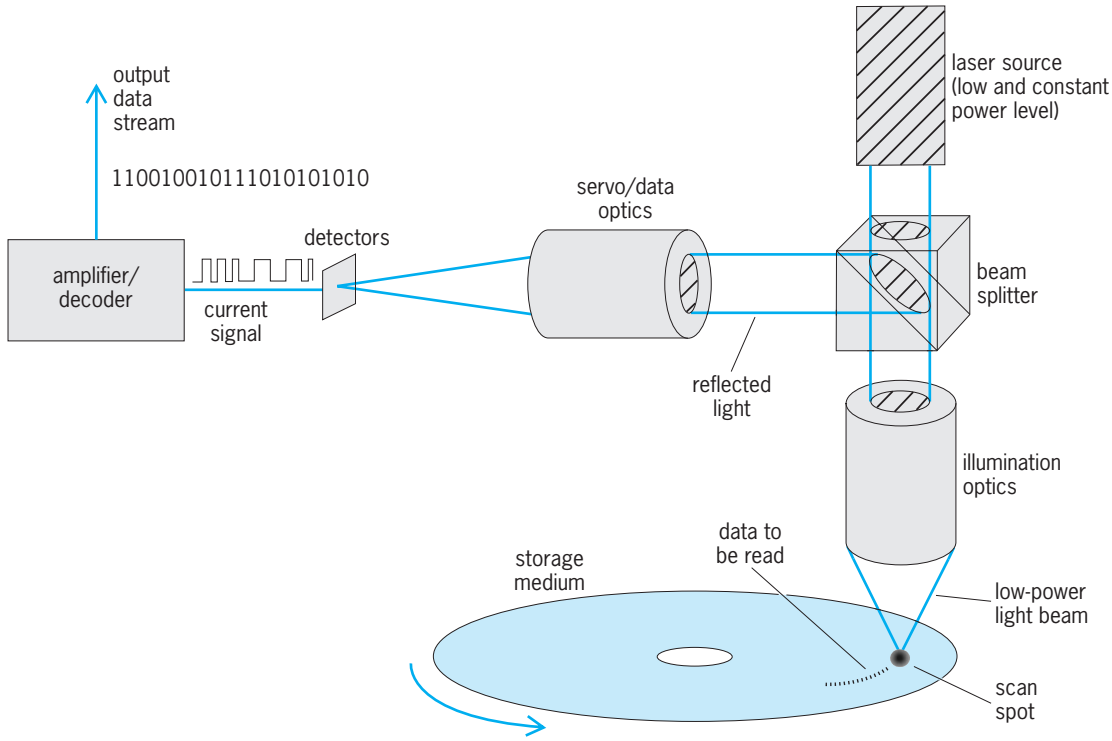

Fig. 7. Readout of an optical medium. Low-power laser beam illuminates the recording layers, and modulation of the reflected light is observed with the detectors. Beam splitter serves to direct a portion of the reflected light to detectors.

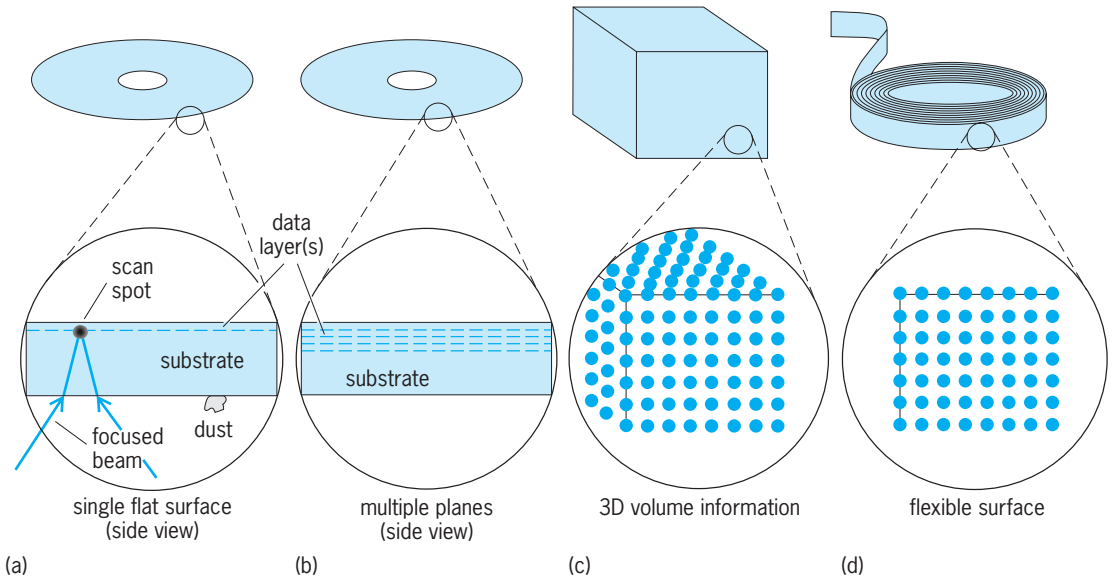

Fig. 8. Configurations for optical media. (*a*) Single-layer disk. (*b*) Multiple-layer disk. (*c*) Volumetric. (*d*) Ribbon.

illumination optics so that only one layer is read out at a time.

Data can also be recorded in volumetric configurations (Fig. 8*c*). As with the multiple-layer disk, the scan spot can be refocused throughout the volume of material to access information. Volumetric configurations offer the highest efficiency for data capacity, but they are not easily paired with simple illumination optics.

The final configuration is to place the information on a flexible surface, such as ribbon or tape (Fig. 8*d*). As with magnetic tape, the ribbon is pulled under the scan spot and data are recorded or retrieved. Flexible media have about the same capacity efficiency as volumetric storage. The advantage of a flexible medium over a volumetric medium is that no refocusing is necessary. The disadvantage is that a moderately complicated mechanical system must be used to move the ribbon.

Commercial media and optical systems. There are several types of optical storage media. The most popular media are based on pit-type, magnetooptic, phase-change, and dye-polymer technologies. CD and digital versatile disc (DVD) products use pit-type technology. Erasable disks using magnetooptic (MO) technology are popular for workstation environments. Compact-disk-rewritable (CD-RW) products [also known as compact-disk-erasable (CD-E)] use phase-change technology, and compact-disk-recordable (CD-R) products use dye-polymer technology. CD and DVD products are read-only memories (ROMs); that is, they are used for software distribution and cannot be used for recording information. CD-R products can be used for recording information, but once the information is recorded, they cannot be erased and reused. Both CD-RW and MO products can be erased and reused.

Pit-type technology for CD and DVD products is based on a simple scattering phenomenon. Data are recorded as small pits in the surface of a disk. The pits are arranged in spiral tracks around the center of the disk. The pit length is of the order of the wavelength of light, or about 0.5–1 micrometer. The widths of the pits along a track are nearly uniform and measure about 0.5–0.8 μm. Information is retrieved by scanning a focused optical spot, about 2 μm wide and 2 μm long, over the pits in a track as the disk spins. As the light spot passes over a pit, the reflected light scatters away from the illumination optics. The remaining light collected is small compared to the amount of light that gets collected when the spot is over a smooth portion of the track, where the disk acts as a mirror to the focused light. The data signal is derived from the detector that senses the amount of collected light.

MO products store information in small magnetic marks, which are about the same size as pits on a CD. Each mark contains magnetic domains oriented in the opposite direction compared to the magnetic domains of the background. The marks have the property that, as a focused light spot passes over it, the polarization of the reflected light is rotated. In order to detect the data signal, a detector is used to sense the change in polarization of the reflected light. To record or erase marks, a higher-power focused spot is used to locally heat the medium, and with the application of an external magnetic field, the magnetic domains of the material can be switched in micrometer-sized regions. A major difference between CD and MO products is that the MO marks are produced in a track with an almost undetectable change in the topology of the track. That is, there is almost no mechanical deformation of the track as the marks are recorded or erased. This property enables MO products to exhibit over 10^6 erase cycles with little if any degradation in performance. *See* MAGNETOOPTICS; POLARIZED LIGHT.

CD-RW products are similar to MO products in that they can be erased with multiple cycles before degradation. However, phase-change technology is based

on the differences of the crystalline and amorphous states of the medium. Marks representing data are stored along tracks. Usually, marks are in the amorphous state and the background is in the crystalline state. Amorphous and crystalline states reflect different amounts of light back into the optical system. As in a CD player, the detector in a CD-RW player senses the amount of collected light. The data signal is derived from the detector current. To record or erase phase-change marks, a higher-power focused spot is used to locally heat the medium in micrometer-sized regions. The thermal cycle of the local regions determines if the region will stabilize in a crystalline or amorphous state. By controlling the energy in the focused spot, the thermal cycle and the state of the material can be controlled. The phase-change process inevitably involves a mechanical deformation of the material. Therefore, the number of erase cycles is limited to several thousand. *See* AMORPHOUS SOLID.

The dye polymers or dye monomers used in CD-R products are organic films that are ablated to form pits along tracks. To form a pit, a high-power focused spot locally heats a micrometer-sized area. The dye polymer absorbs a large percentage of the laser energy. Due to the low thermal conductivity of dye polymers, extremely high temperatures can be reached. In the heated area, the dye material is vaporized or heated to the point that material flows to form a pit. To read data, a low-power laser beam scans the track, and the collected light is sensed with a simple detector. The optical system of a CD-R player is similar to that of a CD-RW player.

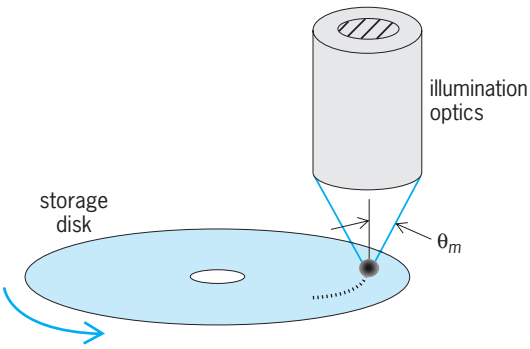

Fig. 9. Geometry that determines the size *s* of the scan spot. The spot size is a function of the laser wavelength λ and the maximum angle, θ_m, of the focused cone of light illuminating the disk: $s = \lambda/\text{NA} = \lambda/\sin\theta_m$, where NA is the numerical aperture.

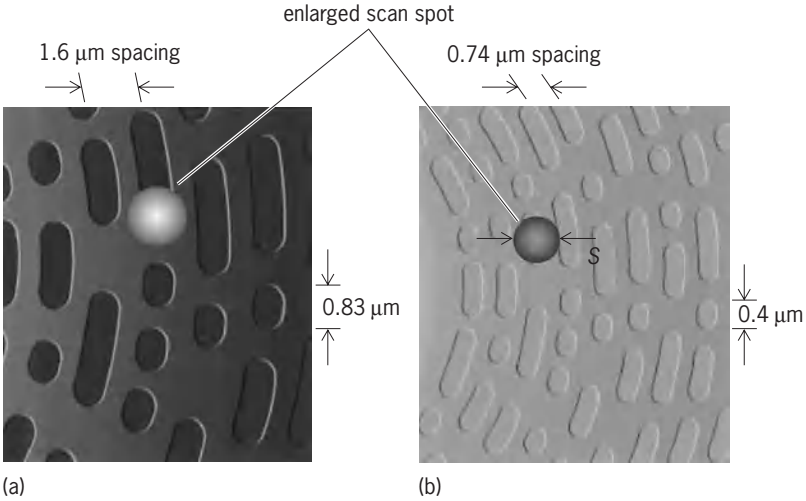

Fig. 10. Comparison of scan spot sizes and sizes of data marks for (*a*) CD and (*b*) DVD media. Spot size *s* is about 1.1 μm for CD and 1.7 μm for DVD.

Performance. Three important performance characteristics of optical data storage devices are the capacity, data rate, and access time.

Capacity is the maximum amount of data that can be stored on a single disk. Capacity is usually specified in terms of gigabytes. (1 GB = 10^9 bytes; 1 byte = 8 bits.) *See* BIT.

Data rate is the number of digital bits per second that are recorded or retrieved from a device during transfer of a large data block. Data rate is usually specified in terms of megabits per second. (1 Mbps = 10^6 bits per second.)

Access time is the latency (time lapse) experienced between when a request is made to access data and when the data start flowing through the communication channel. Access time is usually specified in terms of milliseconds. (1 ms = 10^{-3} second.)

Together, data rate and access time determine the throughput of the device. That is, throughput determines the time required to locate and transmit data to and from the storage device.

The capacity is primarily determined by the size s of the scan spot (the diameter of a circle at which the light intensity drops to $1/e^2$ of its maximum value). As s decreases, smaller transitions between marks and the regions around marks can be sensed. Therefore, marks can be made smaller and the density increases. A useful approximation is $s \sim \lambda/\text{NA}$, where λ is the wavelength of the laser in air and NA is the numerical aperture of the focused beam. Numerical aperture is defined as the sine of the focusing-cone half angle (**Fig. 9**). CD systems exhibit $\lambda = 0.78$ μm and NA = 0.45, which produces $s = 1.7$ μm. DVD systems exhibit $\lambda = 0.65$ μm and NA = 0.60, with $s = 1.1$ μm (**Fig. 10**).

The data rate can be different for writing and reading data on a disk. During writing, the data rate is determined by the highest medium velocity that produces clearly defined marks. During reading, the data rate is determined by the highest medium velocity that produces sufficient signal-to-noise ratio. One straightforward way to increase data rate is to use more than one laser beam at a time. The increase in data rate is nearly proportional to the number of beams.

The access time is determined by the latency due to the time required for disk rotation. The highest latency is the time it takes the disk to make one revolution. Reduction of latency requires spinning the disk faster.

Important considerations for storage are the performance requirements of new and existing applications. For example, as introduced in 1991, the CD-ROM exhibited a capacity of 0.64 GB and a data rate of 1.2 Mbps. By 2000, although the CD-ROM still

Performance data of disk-based products

	CD-1X	CD-40X	DVD-1X	DVD-40X	BD-1X
Capacity, GB	0.64	0.64	4.7	4.7	25
Data rate, Mbps	1.2	48	10	400	36
Capacity-rate product (CRP)	0.77	30.7	47	4700	900
Retrieval time, min	70	1.7	62.7	1.6	92.6

had the same capacity, its data rate had been raised to over 50 Mbps. Its higher speed enabled its acceptance for computer applications over introductory DVD products, which had higher capacity but exhibited a data rate of only 10 Mbps.

A serious limitation exists with disk-based optical data storage. As the data rate increases, the playing time for a fixed capacity decreases. For applications that require long playing times (and correspondingly high capacities), this limitation necessitates compression, which is lossless only for compression factors of 2 to 4. For example, a CD-ROM drive operating at 50 Mbps takes only 102 seconds to read the entire disk. Correspondingly, a hypothetical DVD-ROM drive operating at 400 Mbps (a similar speed multiplier compared to the fast CD drive) takes less than 100 seconds to read a 4.7 GB disk.

A useful figure of merit is the capacity-rate product (CRP), which is the product of the capacity in GB and the data rate in Mbps. The CRP and other performance characteristics of disk-based products are given in the **table**. The data-rate speedup factor is shown as 1X or 40X, where 1X refers to the data rate of products first introduced, such as the CD-ROM in 1991, and 40X refers to a data rate that is 40 times faster than the 1X rate. Also included in the table are preliminary data concerning the Blue-Ray Disc (BD), which is now comercially available. *See* COMPUTER STORAGE TECHNOLOGY.

Tom D. Milster; Glenn T. Sincerbox

Bibliography. T. W. McDaniel and R. Victoria, *Handbook of Magneto-Optical Data Recording: Materials, Subsystems, Techniques*, Noyes Publications, Park Ridge, NJ, 1997; A. B. Marchant, *Optical Recording: A Technical Overview*, Addison-Wesley, Reading, MA, 1990; T. D. Milster, Near-field optics: A new tool for data storage, *Proc. IEEE*, 88(9):1480–1490, 2000; D. Psaltis, D. G. Stinson, and G. S. Kino, Introduction by T. D. Milster, Optical data storage: Three perspectives, *Optics & Photonics News*, pp. 35–39, November 1997.

Optical rotatory dispersion

The change in rotation as a function of wavelength experienced by linearly polarized light as it passes through an optically active substance. *See* OPTICAL ACTIVITY; POLARIZED LIGHT.

Optically active materials. Substances that are optically active can be grouped into two classes. In the first the substances are crystalline and the optical activity depends on the arrangement of nonoptically active molecular units. When these crystals are dissolved or melted, the resulting liquid is not optically active. In the second class the optical activity is a characteristic of the molecular units themselves. Such materials are optically active as liquids or solids. A typical substance in the first category is quartz. This crystal is optically active and rotates the plane of polarization by an amount which depends on the direction in which the light is propagated with respect to the optic axis. Along the axis the rotation is 29.73°/mm (755°/in.) for light of wavelength 508.6 nanometers. At other angles the rotation is less and is obscured by the crystal's linear birefringence. Molten quartz, or fused quartz, is isotropic. Turpentine is a typical material of the second class. It gives rotation of −37° in a 10-cm (3.94-in.) length for the sodium D lines. *See* CRYSTAL OPTICS.

Reasons for variation. In all materials the rotation varies with wavelength. Optical activity is actually circular birefringence. In other words, a substance which is optically active transmits right circularly polarized light with a different velocity from left circularly polarized light.

Any type of polarized light can be broken down into right and left components. Let these components be R and L. The lengths of the rotating light vectors will then be $R/\sqrt{2}$ and $L/\sqrt{2}$. At $t = 0$, the R vector may be at an angle ψ_r with the x axis and the L vector at an angle ψ_l. Since the vectors are rotating at the same velocity, they will coincide at an angle β which bisects the difference as in Eq. (1). If $R =$

$$\beta = \frac{\psi_r + \psi_l}{2} \tag{1}$$

L, the sum of these two waves will be linearly polarized light vibrating at an angle γ to the axes given by Eq. (2).

$$\gamma = \frac{\psi_r - \psi_l}{2} \tag{2}$$

If, in passing through a material, one of the circularly polarized beams is propagated at a different velocity, the relative phase between the beams will change in accordance with Eq. (3), where d is

$$\psi_r' - \psi_l' = \frac{2\pi d}{\lambda}(n_r - n_l) + \psi_r - \psi_l \tag{3}$$

the thickness of the material, λ is the wavelength, and n_r and n_l are the indices of refraction for right and left circularly polarized light. The polarized light

incident at an angle γ has, according to this equation, been rotated an angle α given by Eq. (4).

$$\alpha = \frac{\pi d}{\lambda}(n_r - n_l) \tag{4}$$

This shows that the rotation would depend on wavelength, even in a material in which n_r and n_l were constant and which thus had no dispersion of circular birefingence. In addition to this pseudodispersion, there is a true rotatory dispersion which depends on the variation with wavelength of n_r and n_l.

From Eq. (4) it is possible to compute the circular birefringence for various materials. This quantity is of the order of magnitude of 10^{-8} for many solutions and 10^{-5} for crystals. It is 10^{-1} for linear birefringent crystals. Bruce H. Billings

Specific and molar rotations. For purposes of comparison and compilation, the specific rotation $[\alpha]$, which is the optical rotation that would be given by a sample 1 decimeter (0.1 m) thick and at a concentration of 1 kg/dm^3, is calculated from the observed rotation, as in Eq. (5), where d is the thickness in

$$[\alpha] = \frac{\alpha}{10dw} \tag{5}$$

meters and w is the concentration of optically active solute in kg/dm^3. Molar rotation ($[M]$), the optical rotation of a sample 1 m thick containing 1 mole/dm^3 of the optically active solute, is defined by Eq. (6),

$$[M] = \frac{[\alpha]M}{100} = \frac{\alpha}{dc} \tag{6}$$

where M is the molecular weight of the molecule or, in the case of a polymer, the molecular weight of the average monomer unit, and c is the solute concentration in moles per liter.

Circular dichroism. At wavelengths that are far from any absorption bands, the optical rotation generally shows a monotonic increase in magnitude as the wavelength decreases. The Drude equation (7) describes the optical rotatory dispersion (ORD) in such cases, where A and λ_0 are empirical constants, with λ_0 generally being a wavelength in the far-ultraviolet region. Most colorless optically active substances obey the Drude equation in the visible and near-ultraviolet regions.

$$[M] = \frac{A}{\lambda^2 - {\lambda_0}^2} \tag{7}$$

Wavelengths that are absorbed by the optically active sample, the two circularly polarized components will be absorbed to differing extents. This unequal absorption is known as circular dichroism (CD) and is given by Eq. (8), where ϵ_l and ϵ_r are the molar extinction coefficients for left and right circularly polarized light, respectively. Circular dichroism causes incident linearly polarized light to become elliptically polarized. Optical rotation can still be measured as the angle between the plane of polarization of the incident light and the major axis of the elliptically polarized light. The elliptically polarized light is also characterized by a second angle, the ellipticity θ, which is an alternative measure of circular dichroism. The molar ellipticity, defined by analogy to the molar rotation [Eqs. (6) and (7)], is proportional to $\Delta\epsilon$, as in Eq. (9), where θ is the observed ellipticity,

$$\Delta\epsilon = \epsilon_l - \epsilon_r \tag{8}$$

$$[\theta] = \frac{\theta}{dc} = 3300\Delta\epsilon \tag{9}$$

d is the sample thickness in meters, and c is the molar concentration of the solute. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION.

If a single absorption band is considered, the circular dichroism spectrum has the same shape as the absorption spectrum, whereas the optical rotatory dispersion spectrum shows more complex behavior. For a positive circular dichroism band, the optical rotatory dispersion passes through a maximum on the long-wavelength side of the circular dichroism and absorption maximum (λ_{max}), vanishes at the absorption maximum, then passes through a minimum on the short-wavelength side of the absorption maximum. For a negative circular dichroism band, characteristic of the mirror-image molecule (enantiomer), opposite behavior is observed. The characteristic shapes of optical rotatory dispersion and circular dichroism spectra in the vicinity of absorption bands are referred to as Cotton effects, with a positive optical rotatory dispersion Cotton effect corresponding to a positive circular dichroism band. *See* COTTON EFFECT.

Optical rotatory dispersion and circular dichroism are closely related, just as are ordinary absorption and dispersion. If the entire optical rotatory dispersion spectrum is known, the circular dichroism spectrum can be calculated, and vice versa. Circular dichroism bands are narrow, confined to regions of absorption, in contrast to optical rotatory dispersion bands, which decrease slowly away from the absorption maximum. Circular dichroism is generally preferred because the contributions of individual chromophores are more easily distinguished. Weak circular dichroism bands can be detected and measured without interference from strong bands at shorter or longer wavelengths. With optical rotatory dispersion, weak Cotton effects are difficult to study, because they are often small inflections on a large and strongly curving background. Optical rotatory dispersion still proves useful for molecules such as hydrocarbons and carbohydrates, for which circular dichroism bands are confined to the deep-ultraviolet.

Relationships to molecular structure. In order for a molecule (or crystal) to exhibit circular birefringence and circular dichroism, it must be distinguishable from its mirror image. An object that cannot be superimposed on its mirror image is said to be chiral, and optical rotatory dispersion and circular dichroism are known as chiroptical properties. Helical molecules represent especially clear examples of chiral objects and include the α-helix in polypeptides and proteins, the double helix of deoxyribonucleic acid (DNA), the helix of glucose units in amylose

(starch), and the helices of oxygen-silicon-oxygen (O—Si—O) units in quartz crystals. Molecules that contain a carbon (or other atom) with four different substituents in a tetrahedral arrangement exist in two mirror-image forms and are therefore also chiral, and the two enantiomers have mirror optical rotatory dispersion and circular dichroism spectra. A carbon or other atom that has a chiral arrangement of substitutents is said to be an asymmetric or chiral center. In many complexes of transition-metal ions, the metal ion is a chiral center. Most biological molecules have one or more chiral centers and undergo enzyme-catalyzed transformations that either maintain or reverse the chirality at one or more of these centers. Still other enzymes produce new chiral centers, always with a high specificity. These properties account for the fact that optical rotatory dispersion and circular dichroism are widely used in organic and inorganic chemistry and in biochemistry. *See* ENZYME; STEREOCHEMISTRY.

At one time, measurements of optical rotatory dispersion and circular dichroism were confined to visible and ultraviolet wavelengths, from about 185 to 700 nm. Developments in instrumentation for measuring circular dichroism have opened the very far-ultraviolet (down to about 140 nm) and the infrared (wavelengths up to 15 micrometers). The deep-ultraviolet has proven useful for studies of proteins, nucleic acids, and carbohydrates. Circular dichroism in the infrared derives from transitions between vibrational energy levels, rather than electronic energy levels as in visible-ultraviolet circular dichroism, and it is known as vibrational circular dichroism (VCD). Measurements of vibrational circular dichroism, though technically demanding, promise to provide much information about the stereochemistry and conformation of both large and small molecules. *See* INFRARED RADIATION.

Magnetic effects. In the absence of magnetic fields, only chiral substances exhibit optical rotatory dispersion and circular dichroism. In a magnetic field, even substances that lack chirality rotate the plane of polarized light, as shown by M. Faraday. Magnetic optical rotation is known as the Faraday effect, and its wavelength dependence is known as magnetic optical rotatory dispersion (MORD). In regions of absorption, magnetic circular dichroism (MCD) is observable. In addition to the other advantages of circular dichroism over optical rotatory dispersion, solvent and sample cells do not contribute to the magnetic circular dichroism signal but do contribute to magnetic optical rotatory dispersion, making measurements of the latter on dilute solutions difficult. Magnetic circular dichroism has been used to assign electronic transitions and thus determine the nature of excited states in both transition-metal complexes and aromatic hydrocarbons. In biochemistry, magnetic circular dichroism has provided important information about the geometry of the coordination sites of transition-metal ions such as iron and copper in metalloproteins. *See* COORDINATION CHEMISTRY; COORDINATION COMPLEXES; FARADAY EFFECT; MAGNETOOPTICS.

R. W. Woody

Bibliography. L. D. Barron, *Molecular Light Scattering and Optical Activity*, 1983; S. F. Mason, *Molecular Optical Activity and the Chiral Discriminations*, 1982; K. Nakanishi, N. Berova, and R. Woody, *Circular Dichroism: Principles and Applications*, 1994; N. Purdie (ed.), *Analytical Applications of Circular Dichroism*, 1993.

Optical surfaces

Interfaces between different optical media at which light is refracted or reflected. From a physical point of view, the basic elements of an optical system are such things as lenses and mirrors. However, from a conceptual point of view, the basic elements of an optical system are the refracting or reflecting surfaces of such components, even though they cannot be separated from the components. Surfaces are the basic elements of an optical system because they are the elements that affect the light passing through the system. Every wavefront has its curvature changed on passing through each surface so that the final set of wavefronts in the image space may converge on the appropriate image points. Also, the aberrations of the system depend on each surface, the total aberrations of the system being the sum of the aberrations generated at the individual surfaces. *See* ABERRATION (OPTICS); REFLECTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES.

Optical systems are designed by ray tracing, and refraction at an optical surface separating two media of different refractive index is the fundamental operation in the process. The transfer between two surfaces is along a straight line if, as is usually the case, the optical media are homogeneous. The refraction of the ray at a surface results in a change in the direction of the ray. This change is governed by Snell's law.

Spherical and aspheric surfaces. The vast majority of optical surfaces are spherical in form. This is so primarily because spherical surfaces are much easier to generate than nonspherical, or aspheric, surfaces. Not only is the sphere the only self-generating surface, but a number of lenses can be ground and polished on the same machine at the same time if they are mounted together on a common block so that the same spherical surface is produced simultaneously on all of them.

Although aspheric surfaces can potentially improve the performance of a lens system, they are very rarely used. High-quality aspheric surfaces are expensive to produce, requiring the services of a highly skilled master optician. Moreover, lens systems seldom need aspherics because the aberrations can be controlled by changing the shape of the component lenses without changing their function in the system, apart from modifying the aberrations. Also, many lens components can be included in a lens system in order to control the aberrations. *See* LENS (OPTICS).

On the other hand, mirror systems usually require aspheric surfaces. Unlike lenses, where the shape

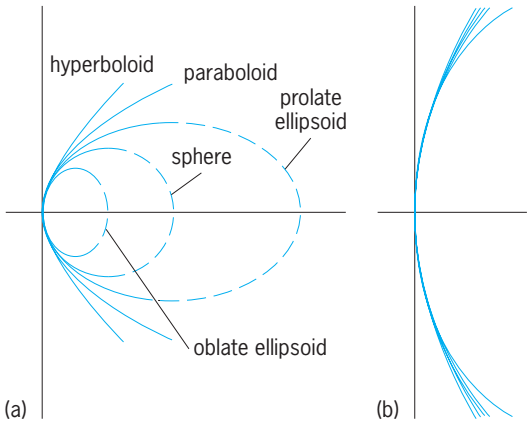

Conics of revolution. (*a*) Cross sections of entire surfaces. (*b*) Cross sections of portions near the optical axis.

can be changed to modify the aberrations, mirrors cannot be changed except by introducing aspheric surfaces. Mirror systems are further constrained by the fact that only a few mirrors, usually two, are used in a system because each successive mirror occludes part of the beam going to the mirror preceding it. *See* MIRROR OPTICS.

Conics of revolution. The most common form of rotationally symmetric surface is the conic of revolution. This is obtained conceptually by rotating a conic curve (ellipse, parabola, or hyperbola) about its axis of symmetry. There are two forms of ellipsoid, depending on whether the generating ellipse is rotated about its major or its minor axis. In the first case it is a prolate ellipsoid, and in the second it is an oblate ellipsoid. There is only one form of paraboloid, and only the major axis is used in generating the hyperboloid. The departure of conic surfaces from spherical form is shown in the **illustration**. *See* CONIC SECTION.

The classical virtue of the conics of revolution for mirrors is the fact that light from a point located at one focus of the conic is perfectly imaged at the other focus. If these conic foci are located on the axis of revolution, the mirror is free of spherical aberration for such conjugate points. For example, the classical Cassegrain design consists of two mirrors, a paraboloidal primary mirror and a hyperboloidal secondary mirror. The paraboloid forms a perfect image of a point at infinity on its axis. When the light converging to this image is intercepted by the convex hyperboloid with its virtual conic focus at the image point, the final image will be formed at the real conic focus of the hyperboloid. Thus, in the classical Cassegrain design the two mirrors are separately corrected for spherical aberration, and so is the system. However, all other monochromatic aberrations remain uncorrected.

Instead of using the two aspheric terms to correct the same aberration for the two mirrors separately, it is more effective to use the two aspherics to correct two aberrations. A Cassegrain system in which both spherical aberration and coma are corrected by the two aspherics is aplanatic, and is identified as a Ritchey-Chrétien system. The aspheric on the primary is a hyperboloid slightly stronger than the paraboloid of the classical primary, and the aspheric on the secondary is also a hyperboloid, slightly different from that for the classical secondary.

General aspherics of revolution. A more general aspheric of revolution is frequently used where terms in even powers of the distance from the axis, usually from the fourth to the twelfth power, are added to the spherical or conic term. This gives greater control over the aberrations, especially if higher-order aberrations are significant. The added terms are necessary also for describing an aspheric departure from a plane surface where a conic cannot be used, as with a Schmidt corrector plate. *See* SCHMIDT CAMERA.

Nonrotationally symmetric surfaces. Occasionally nonrotationally symmetric surfaces are used in optical systems. The most common varieties are cylindrical and toric surfaces, where the cross sections are circular. They are employed, for example, in the optical systems for taking and projecting wide-screen motion pictures, where the image recorded on the film is compressed in the horizontal direction but not in the vertical. The camera lens compresses the image and the projector lens expands it to fill the wide screen. Noncircular cross sections can also be specified by adding terms in even powers of x and y, the coordinates perpendicular to the optical axis, to the description of the surface, but these surfaces are very difficult to make. *See* CINEMATOGRAPHY.

Eccentric mirrors. Another significant class of aspheric surfaces consists of those which are eccentric portions of rotationally symmetric aspheric surfaces. These are used in mirror systems to eliminate the obscuration of the incoming beam by succeeding mirrors. For small mirrors they are cut out of a large rotationally symmetric mirror, but for larger mirrors they can be made by computer-controlled generating and polishing machines. *See* GEOMETRICAL OPTICS.

Roland V. Shack

Bibliography. D. Malacara, *Optical Shop Testing*, 2d ed., 1992; D. Malacara and A. Malacara, *Handbook of Lens Design*, 1994; Optical Society of America, *Handbook of Optics*, 2 vols., 2d ed., 1995; D. C. O'Shea, *Elements of Modern Optical Design*, 1985.

Optical tracking systems

Multipurpose instruments used to make measurements on a remote object, often an airborne vehicle. These systems are used to provide two basic types of data: accurate measurement of the position, velocity, and other motion parameters of the target; and information about the target, such as images or optical spectra.

Many optical tracking systems are used for weapon system qualification. A system tracking a towed aerial target can accurately determine the miss distance of an interceptor missile. One tracking a military aircraft can measure its rate of climb or the release characteristics of its ordnance.

Fig. 1. Little Bright Eyes. (*White Sands Missile Range*)

Another common use of optical tracking systems is for range safety at missile and rocket launch sites. A remotely operated optical tracking system can provide real-time position data and images of the launch vehicle, assuring that all technical parameters are within specification without risk to personnel.

Development. The first optical tracking mount was developed in Germany in 1937 at the Peenemünde Rocket Launch Site by associates of Wernher von Braun. Development continued to support the German rocket development program until the production of the first commercial phototheodolite in 1940. The United States government issued a specification the next year for a similar instrument, but its development was not completed before the end of the war.

With the 1945 arrival of von Braun, his scientific crew, and their V2 rockets in the United States, a longer-range tracking mount was required. James Edson, an antiaircraft artillery officer, married a 35-mm film camera to a gun mount and created the forerunner of modern optical tracking mounts, called Little Bright Eyes (**Fig. 1**). Test ranges incorporated a succession of improvements in each new instrument, but all were based on gun mounts. This changed with the third generation of the Intercept Ground Optical Recorder (IGOR-III), developed and produced in 1960. IGOR-III was designed from the beginning to be a precision optical instrument and incorporated many design features still in use.

Two operators rode on the IGOR system, one controlling the elevation angle and the other the azimuth or bearing direction. This was common practice on optical tracking systems until the development of the Recording Optical Tracking Instrument (ROTI). This instrument was similar in size and capability to the IGOR, but could be directed by a single operator using a joystick. The development of these large systems concluded with production of the Distant Object Attitude Measurement System (DOAMS; **Fig. 2**) in 1976.

IGOR, ROTI, and DOAMS are the high end of optical trackers, with apertures up to 20 in. (0.5 m) and 400-in. (10-m) focal length. Coupled with large-format 70-mm high-speed film cameras, these instruments were perfectly suited for documenting high-altitude and long-range intercepts. These missions constitute only a small portion of range test requirements; most of the rest could be easily satisfied by a smaller instrument.

The cinetheodolite (**Fig. 3**), familiarly referred to as a "cine", was developed in 1952. These instruments quickly became standard equipment at test ranges. Various models were in production for 30 years, with over 400 systems built. The most common of these, the EOTS-F, had a 7.5-in. (190-mm)

Fig. 2. Distant Object Attitude Measurement System (DOAMS). (*L-3 Communications Corp.*)

Fig. 3. Cinetheodolite. (*L-3 Communications Corp.*)

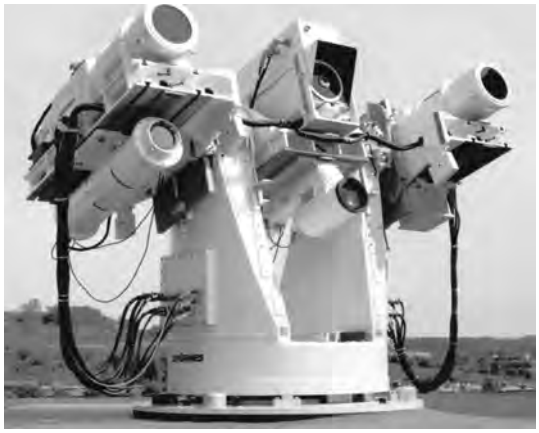

Fig. 4. KINETO tracking mount. (*L-3 Communications Corp.*)

aperture and a 35-mm high-speed film camera. Nearly 100 of these systems are still in use.

The cinetheodolites were extremely accurate, but required specially prepared sites and had a single optical configuration. There was a need for a more versatile instrument, and in 1963 the Cine Sextant, a precision two-axis tracking system mounted on a trailer, was introduced. These systems could carry multiple instruments and be reconfigured for each test. They have been largely supplanted by the Compact Tracking Mount (CTM) and the KINETO tracking mount (**Fig. 4**). These somewhat smaller mounts have all-electric drives instead of the hydraulic/electric combination on the Cine Sextant and as a result are considerably less expensive. Over 150 KINETOs have been produced, making it the most common system in use at test ranges.

Characteristics. To achieve the basic objective, directing an optical instrument at an agile target, an optical tracking system must meet several key performance parameters, as summarized in the **table**.

The aperture of the optical system sets the overall sensitivity; a larger lens collects more light. The focal length determines the size of the image, as in a normal film camera. Optical tracking systems use long-focal-length, telephoto-type lenses. The system resolution is the angular extent of the smallest discernible features. *See* FOCAL LENGTH; GEOMETRICAL OPTICS; LENS (OPTICS); RESOLVING POWER (OPTICS); TELESCOPE.

The angular velocity is a measure of how fast the system rotates. A jet flying at the speed of sound 1 mi (1.6 km) from the tracking system requires a line-of-sight rate of 12°/second. A KINETO tracking mount can move at eight times that rate.

Typical tracking system performance

Parameter	Typical value	Units
Aperture	12 (300)	inches (mm)
Focal length	50 (1250)	inches (mm)
Resolution	1–5	arc-seconds*
Angular velocity	30	degrees/second
Pointing accuracy	5–10	arc-seconds*

*1 arc-second = (1/3600)° = 0.00028°.

Tracking. Optical tracking systems have two basic modes of operation, feedback tracking and ephemeris tracking. With feedback tracking, the position of the target in the field of view of the tracking sensor provides information that corrects the trajectory of the system. Feedback tracking can be as simple as an operator moving a joystick based on what he or she sees on a screen, or it can employ an automatic video tracker (AVT), a specialized image-processing computer that measures the target position in the sensor field of view. The most common type of AVT is called a centroid tracker because it determines the outline of the target and then calculates its center. Some types of objects that have variable size, such as a missile with a plume, are better tracked with an edge tracker. A third type, called a correlation tracker, overlays the current image with the previous one and adjusts the offsets to minimize the differences between them. This is useful in high-clutter environments.

An ephemeris is a mathematical prediction of the path of a celestial body, the usual means of tracking a satellite. Most satellites are visible only when the Sun is shining on them and the angle between the Sun and the observer is favorable. By accurately calibrating the pointing of the tracking system and commanding it from the same equations that determine the motion of the satellite, the line of sight can be directed correctly without feedback. This could be important, for instance, in maintaining a laser communications link to a faint satellite. *See* EPHEMERIS.

Instrumentation. Historically, the instrument of choice for optical tracking systems has been the high-speed film camera. Under favorable conditions, a film could deliver 0.00025-in. (10-μm) resolution in a 70-mm format at 1000 frames per second. To do the same, a digital camera would need 50 megapixels and an 800-gigabit-per-second data rate. No commercial camera has reached this level of performance, but there have been great advances. Delivery of 1000 pictures per second at 2 megapixels has been achieved. *See* CAMERA; PHOTOGRAPHY.

Digital cameras have largely replaced film cameras on the ranges, partly due to the desire for real-time data and partly due to the expense and hazard of film processing. This replacement has been facilitated by the development of standard camera interfaces.

An increasing number of optical tracking systems employ infrared cameras, sometimes called FLIRs after their original application as the forward-looking infrared sensors in aircraft. These sensors do not have the same resolution as visible-band cameras, but have the advantage of being able to track objects that are not sunlit or that have poor contrast. *See* INFRARED IMAGING DEVICES.

It is often necessary to measure the range to a target. This is accomplished by adding a radar ranging system to the optical tracking system or by incorporating a laser rangefinder. Both determine the distance by measuring the time necessary for a pulse to travel to the target and be reflected back to its source.

Most laser rangefinders now meet standards for eye safety and can easily be added to a KINETO or other configurable tracking system. *See* LASER; RADAR.

Special applications. Four applications are described below.

Time space-position information (TSPI). One or more optical tracking systems can be used to accurately fix the position of a target in space. To effectively use this information, the system locations must be accurately leveled and precisely located by survey or the Global Positioning System. One approach is to use a single optical tracking system that is fitted with a laser rangefinder and to convert measurements in this azimuth-elevation-range (AER) coordinate system into Cartesian (XYZ) coordinates so that a target trajectory is established.

By time-tagging the data (accurately determining the time at which the position is recorded), several instruments can make coordinated measurements. By using techniques similar to a surveyor, angle-only measurements from two, three, or four optical tracking systems can determine the Cartesian position of the target. The use of a greater number of tracking systems can reduce target location errors through advanced processing. *See* SURVEYING.

Fire-control systems. The ability of an optical tracking system to accurately determine the position and trajectory of a target leads naturally to its incorporation into air defense systems. The optical line of sight is corrected for the ballistic characteristics of the weapon being used and the calculated motion of the target until intercept.

Space surveillance. The United States government's Groundbased Electro-Optical Deep Space Surveillance (GEODSS) telescopes perform a critical function by mapping the location of artificial and natural objects in high orbits around the Earth. Objects in low-earth orbits are tracked by radar, but this optical tracking system was uniquely configured to detect these small objects at ranges of thousands of kilometers. The tracking system slowly and precisely changes its direction to compensate for the rotation of the Earth, causing the stars to appear stationary in its camera. Satellites, space debris, and minor asteroids appear as short streaks since they move relative to the stars. Once a small part of the sky has been investigated, the system quickly repositions to another location and repeats the process. One GEODSS site has detected over 3 million minor asteroids, of which 200,000 were previously unknown. *See* ASTEROID.

Satellite laser ranging. There is a network of groundbased satellite laser ranging systems run by the National Aeronautics and Space Administration (NASA), foreign governments, and foreign and domestic universities. Each of these systems continuously measures the distance from its location to specially instrumented reflective satellites orbiting the Earth, often day and night. A 4000-mi (6400-km) measurement can be made with an accuracy of better than 0.04 in. (1 mm). The accuracy of these measurements, coupled with the worldwide distribution of the stations, the extreme stability of the satellite orbits, and the long history of observations, allows the motion of the Earth's individual crustal plates, amounting to millimeters per year, to be precisely measured. *See* GEODESY.

James E. Kimbrell

Bibliography. R. Delgado, Photo-optical range instrumentation: An overview, *Opt. Eng.*, 20(5):701–711, September–October 1981.

Optics

Narrowly, the science of light and vision; broadly, the study of the phenomena associated with the generation, transmission, and detection of electromagnetic radiation in the spectral range extending from the long-wave edge of the x-ray region to the short-wave edge of the radio region. This range, often called the optical region or the optical spectrum, extends in wavelength from about 1 nanometer (4×10^{-8} in.) to about 1 mm (0.04 in.). For information on the various branches of optics *see* GEOMETRICAL OPTICS; METEOROLOGICAL OPTICS; PHYSICAL OPTICS; VISION.

In ancient times there was some isolated elementary knowledge of optics, but it was the discoveries of the experimentalists of the early seventeenth century which formed the basis of the science of optics. The statement of the law of refraction by W. Snell, Galileo Galilei's development of the astronomical telescope and his discoveries with it, F. M. Grimaldi's observations of diffraction, and the principles of the propagation of light enunciated by C. Huygens and P. de Fermat all came in this relatively short period. The publication of Isaac Newton's *Opticks* in 1704, with its comprehensive and original studies of refraction, dispersion, interference, diffraction, and polarization, established the science.

So great were the contributions of Newton to optics that a hundred years went by before further outstanding discoveries were made. In the early nineteenth century many productive investigators, foremost among them Thomas Young and A. J. Fresnel, established the transverse-wave nature of light. The relationship between optical and magnetic phenomena, discovered by M. Faraday in the 1840s, led to the crowning achievement of classical optics—the electromagnetic theory of J. C. Maxwell. Maxwell's theory, which holds that light consists of electric and magnetic fields propagated together through space as transverse waves, provided a general basis for the treatment of optical phenomena. In particular, it served as the basis for understanding the interaction of light with matter and, hence, as the basis for treatment of the phenomena of physical optics. In the hands of H. A. Lorentz, this treatment led at the end of the nineteenth century and the beginning of the twentieth to an explanation of many optical phenomena, such as the Zeeman effect, in terms of atomic and molecular structure. The theories of Maxwell and Lorentz are regarded as the culmination of classical optics. *See* ELECTROMAGNETIC RADIATION; LIGHT; MAXWELL'S EQUATIONS.

In the twentieth century, optics has been in the forefront of the revolution in physical thinking caused by the theory of relativity and especially by

the quantum theory. To explain the wavelength dependence of heat radiation, the photoelectric effect, the spectra of monatomic gases, and many other phenomena of physical optics, radical departure from the ideas of Lorentz and Maxwell about the mechanism of the interaction of radiation and matter and about the nature of radiation itself has been found necessary. The chief early quantum theorists were M. Planck, A. Einstein, and N. Bohr; later came L. de Broglie, W. Heisenberg, P. A. M. Dirac, E. Schrödinger, and others.

The science of optics finds itself in a position that is satisfactory for practical purposes but less so from a theoretical standpoint. The theory of Maxwell is sufficiently valid for treating the interaction of high-intensity radiation with systems considerably larger than those of atomic dimensions. The modern quantum theory is adequate for an understanding of the spectra of atoms and molecules and for the interpretation of phenomena involving low-intensity radiation, provided one does not insist on a very detailed description of the process of emission or absorption of radiation. However, a general theory of relativistic quantum electrodynamics valid for all conditions and systems has not been worked out. *See* QUANTUM ELECTRODYNAMICS.

The development of the laser has been an outstanding event in the history of optics. The theory of electromagnetic radiation from its beginnings was able to comprehend and treat the properties of coherent radiation, but the controlled generation of coherent monochromatic radiation of high power was not achieved in the optical region until the work of C. H. Townes and A. L. Schawlow in 1958 pointed the way. Many achievements in optics, such as holography and interferometry over long paths, have resulted from the laser. *See* HOLOGRAPHY; INTERFEROMETRY; LASER.

Richard C. Lord

Bibliography. M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; B. D. Guenther, *Modern Optics,* 1990; F. A. Jenkins and H. E. White, *Fundamentals of Optics,* 4th ed., 1976; J. Meyer-Arendt, *Introduction to Classical and Modern Optics*, 4th ed., 1995; Optical Society of America, *Handbook of Optics*, 2 vols., 2d ed., 1995; O. Svelto, *Principles of Lasers,* 4th ed., 1998.

Optimal control (linear systems)

A branch of modern control theory that deals with designing controls for linear systems by minimizing a performance index that depends on the system variables. Under some mild assumptions, making the performance index small also guarantees that the system variables will be small, thus ensuring closed-loop stability. *See* CONTROL SYSTEM STABILITY.

Classical control theory applies directly in the design of controls for systems that have one input and one output. Complex modern systems, however, have multiple inputs and outputs. Examples include aircraft, satellites, and robot manipulators, which, though nonlinear, can be linearized about a desired operating point. Modern control theory was developed, beginning about 1960, for these multivariable systems. It is characterized by a state-space description of systems, which involves the use of matrices and linear algebra, and by the use of optimization techniques to determine the control policies. These techniques facilitate the use of modern digital computers, which lead to an interactive approach to design of controls for complex systems. *See* DIGITAL COMPUTER; LINEAR ALGEBRA; MATRIX THEORY; MULTIVARIABLE CONTROL; OPTIMIZATION.

The multivariable state-variable description is of the form of Eqs. (1), where $u(t)$ is an m-dimensional

$$x = Ax + Bu \qquad z = Hx \tag{1}$$

control input vector and $z(t)$ is a performance output vector for which there are specified performance requirements. The state $x(t)$ is an internal variable of dimension n that describes the energy storage properties of the system. Matrices A and B describe the system dynamics and are determined by a physical analysis using, for example, Newton's laws of motion. They may in general be time-varying. Matrix H is chosen to select the variables of importance as the performance outputs. In the case of nonlinear systems, the description of Eq. (1) results when the system is linearized about a desired operating point. *See* LINEAR SYSTEM ANALYSIS; NONLINEAR CONTROL THEORY.

State-variable feedback design. Traditionally, modern control-system design assumes that all the states $x(t)$ are available for feedback so that the control input is of the form of Eq. (2), where $K(t)$ is an $m \times n$

$$u = -Kx \tag{2}$$

feedback gain matrix, generally time-varying. Substituting the control of Eq. (2) into the system of Eq. (1) yields the closed-loop system given in Eqs. (3). The

$$\dot{x} = (A - BK)x \qquad z = Hx \tag{3}$$

control design problem is to choose the mn entries of the feedback matrix K to yield a desired closed-loop behavior of the performance output $z(t)$.

Linear quadratic regulator. To obtain satisfactory performance of the system of Eq. (1), a quadratic performance index of the form of Eq. (4) may be chosen,

$$J = {}^1/_2 x^T(T)S(T)x(T) + {}^1/_2 \int_0^T (x^TQx + u^TRu)\, dt \tag{4}$$

where $[0,T]$ is the time interval of interest and the symmetric weighting matrices $S(T)$, Q, and R are design parameters that are selected to obtain the required performance. They must be chosen so that $x^T(T)S(T)x(T) \geq 0$ and $x^TQx \geq 0$ for all $x(t)$, and $u^TRu > 0$ for all $u(t)$. That is, Q and $S(T)$ should be positive semidefinite while R should be positive definite.

The positive semidefiniteness of the weighting matrices ensures that the performance index J is

nonnegative. Then, the performance index can be considered as a generalized energy function, so that, if it is kept small, the states $x(t)$ and controls $u(t)$ that are weighted in Eq. (4) are also small over the time interval $[0,T]$.

The selection of the design parameters $S(T)$, Q, and R is generally not straightforward, and presents one of the challenges in the use of modern design techniques. There has been a large amount of work on the topic, which involves consideration of time-domain responses, frequency-domain responses, and the closed-loop poles.

The optimal control design problem is now to select the feedback gain matrix K in such a way that the performance index is minimized. This is called the linear-quadratic regulator (LQR) problem because the system of Eq. (1) is linear, the performance index of Eq. (4) is a quadratic function, and minimizing J corresponds to keeping the state small or regulating it to zero.

The optimal linear-quadratic feedback gain K may be determined by introducing a Lagrange multiplier $\lambda(t)$ and applying the calculus of variations. By using well-known techniques, it can be shown that the optimal gain K may be computed by solving the system equation (1), in conjunction with the Euler-Lagrange equations (5) and (6). To solve these equations, it is

$$\lambda = -A^T\lambda - Qx \tag{5}$$

$$0 = B^T\lambda + Ru \tag{6}$$

necessary to take into account the boundary equations (7). While the system of Eq. (1) is solved for-

$$x(0) \text{ given} \tag{7a}$$

$$\lambda(T) = S(T)x(T) \tag{7b}$$

ward in time, given an initial condition, the Euler-Lagrange equations must be solved backward in time, given a final condition. *See* CALCULUS OF VARIATIONS.

Solving Eq. (6) for $u(t)$ and substituting into Eq. (1) yields the hamiltonian system in Eqs. (8). Solving this

$$x = Ax - BR^{-1}B^T\lambda \tag{8a}$$

$$\lambda = -A^T\lambda - Qx \tag{8b}$$

system with the boundary conditions of Eqs. (7) is a two-point boundary value problem, since $x(0)$ is specified at the initial time, but Eq. (7*b*) must hold at the final time. By introducing an auxiliary time-varying $n \times n$ matrix variable $S(t)$ and using the so-called sweep method, this can be converted to an easier problem. Thus, the optimal linear-quadratic feedback gain K in the feedback control law given in Eq. (2) may be determined by using Eq. (9), with

$$K = -R^{-1}B^TS \tag{9}$$

$S(t)$ determined by solving the matrix Riccati equation (10), where the final condition $S(T)$ is the design

$$-S = A^TS + SA + Q - SBR^{-1}B^TS \tag{10}$$

matrix in the performance index J.

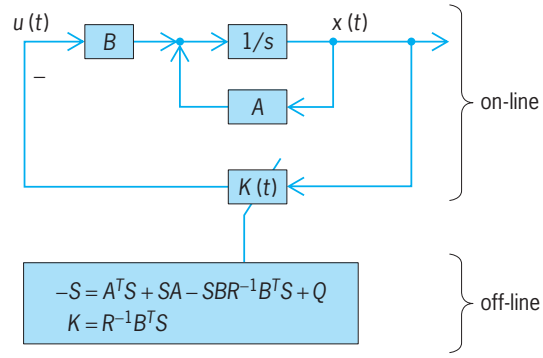

Fig. 1. Feedback formulation of the linear quadratic regulator.

The structure of the linear-quadratic regulator is shown in **Fig. 1**, where s is the Laplace transform variable. It is apparent that it consists of an inner control loop of feedback gains $K(t)$, along with an outer loop required to determine $K(t)$ by solving the quadratic matrix Riccati equation. This hierarchical structure, with a nonlinear outer loop, is typical of modern control systems.

Although the structure of the control system appears complicated, it is not difficult to implement by using modern digital computers. In fact, the Riccati equation may be solved off-line before the control run and the feedback gain $K(t)$ stored in computer memory. Then, during the control run these gains are switched into the system control loop.

It can be shown that, by using the optimal gain $K(t)$ just determined, the performance index takes on the minimum value given by Eq. (11). Thus, $J_{\min}$ may be

$$J_{\min} = x^T(0)S(0)x(0) \tag{11}$$

computed by knowing only the initial state of the system before the control is actually applied to the system. If this value for the performance index is not suitable, new values for the design parameters $S(T)$, Q, and R may be selected and the design repeated. This is an important feature of the linear-quadratic regulator. *See* OPTIMAL CONTROL THEORY.

Constant feedback gains. The optimal linear-quadratic feedback gain $K(t)$ is time-varying. This is in sharp contrast to the situation in classical control theory, where feedback gains are designed by using Laplace transform theory and so are constant. Since constant-feedback gains are easier to implement than time-varying gains, it may be of interest to obtain constant-feedback gains for multivariable systems.

If the final time T approaches infinity, then the time interval of interest becomes $[0,\infty)$, so that the feedback law tries to minimize the performance index and hence regulate the system to zero over all time. This is called the infinite-horizon linear-quadratic control problem. In this case, if the system and performance-index matrices A, B, Q, and R are time-invariant, then under certain conditions the Riccati equation (10) has a limiting solution S_∞ that is constant. If this occurs, then $S_\infty' = 0$, so that

Eq. (10) becomes the algebraic Riccati equation (12).

$$0 = A^T S + SA + Q - SBR^{-1}B^T S \tag{12}$$

Then, S_∞ may be determined directly by solving the algebraic Riccati equation, for which numerical techniques are available using digital computers.

The system of Eq. (1) is called controllable if all of its modes may be independently controlled by using the input $u(t)$. If L is any square root of Q, so that $Q = L^T L$, then (A,L) is called observable if all of the system modes may be independently reconstructed by measuring the linear combination Lx of the state's components. This means basically that all system modes appear in the performance index.

A fundamental result of modern control theory pertains to the case in which the system is time-invariant, Q and R are constant, Eq. (1) is controllable, and (A,L) is observable, where L is any square root of Q. Then, there exists a positive definite limiting solution S_∞ to the Riccati equation that is also the unique positive definite solution to the algebraic Riccati equation. Moreover, the optimal linear-quadratic feedback gain K given by Eq. (9), with S_∞ used there, is constant and guarantees that the closed-loop system of Eq. (3) is stable.

This is an extremely important result, for it shows how to solve the difficult problem of stabilizing a system with multiple inputs and outputs. In fact, under the conditions of controllability and observability, the closed-loop system is stable for any choice of the design parameters Q and R.

Minimum-time control. To obtain different performance objectives, performance indices other than the quadratic performance index in Eq. (4) may be used. If it is desired to drive the system state from a given initial value $x(0)$ to a specified value $x(T)$ in minimum time, the performance index in Eq. (13)

$$J = \int_0^T dt = T \tag{13}$$

is suitable. Choosing the feedback gain K so as to minimize this performance index will result in the minimum time T.

A difficulty with minimum-time control of linear systems is that the solution to the problem of moving from one state to another as quickly as possible is to use infinite control energy. Thus, in order for the problem to have a solution, it is necessary to assume that the control inputs are bounded so that, for example, $|u(t)| \le 1$. The solution to the minimum-time control problem is then found to be the bang-bang control strategy, so called because it involves using a control input that is always at either its maximum or minimum value. The maximum input is used first to accelerate the system quickly, with the minimum control value used to decelerate it so that it comes to rest at the desired final state.

Observer and regulator design. To implement the state feedback control law of Eq. (2), all n components of the state $x(t)$ must be available. However, in practice it is too difficult or expensive to measure all the states. Instead, there is generally available a measured output of the form in Eq. (14) which has only $p < n$

$$y = Cx \tag{14}$$

components. Therefore, if the measured $y(t)$ is used, it is necessary to reconstruct the full state $x(t)$ for use in Eq. (2). This is called the state observer or estimator problem.

One sort of observer is given in Eq. (15). It uses

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}) \tag{15}$$

$u(t)$ and $y(t)$ as inputs, from which the estimate $\hat{x}(t)$ of the state $x(t)$ is reconstructed. The structure of the system and observer is shown in **Fig. 2**. Since $\hat{y} = C\hat{x}$ is the estimate of the output, the observer gain L multiplies the output error $e_y = y - \hat{y}$.

By defining the estimation error $e(t)$ as in Eq. (16), Eq. (15) may be subtracted from Eq. (1) to obtain the error dynamics in Eq. (17). Therefore, to make

$$e = x - \hat{x} \tag{16}$$

$$\dot{e} = (A - LC)e \tag{17}$$

the error go to zero, it is necessary only to select the observer gain L so that $(A - LC)$ is stable, which may be accomplished if (A,C) is observable. This is called the output injection problem and is dual to the state feedback problem, as may be seen by comparing Eq. (17) and Eq. (3).

The state-variable feedback law, Eq. (2), where K has been found by using optimal control theory, may be replaced by the control in Eq. (18), which uses

$$u = -K\hat{x} + v \tag{18}$$

the estimate of $x(t)$ reconstructed from the available measurements $y(t)$. An external input $v(t)$ is included so that the closed-loop system may be manipulated. The structure of the system is shown in Fig. 2, from which it is evident that the observer-plus-feedback arrangement is nothing but a dynamic compensator for the system. It is called a regulator.

It can be shown that the poles of the closed-loop system are just the poles of $(A - BK)$, which depend on the optimal control design, along with the poles of the observer $(A - LC)$. Thus, the feedback

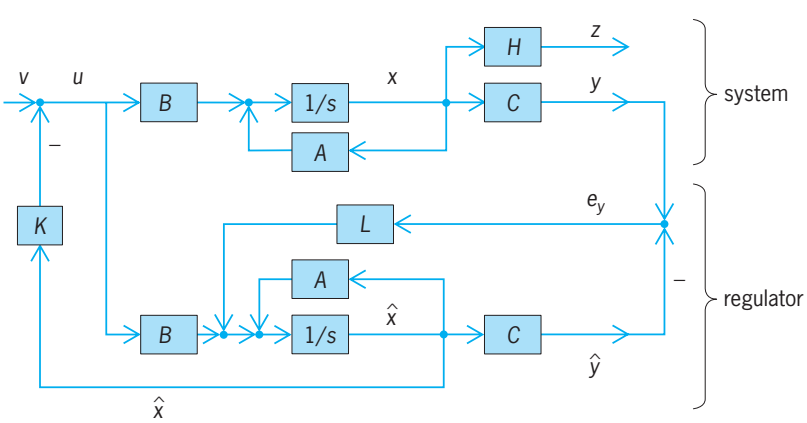

Fig. 2. Structure of state observer and feedback in the regulator configuration.

and the observer may be designed separately. This is the important separation principle of modern control theory.

Other design techniques result in reduced-order observers that have fewer than n states and so are easier to implement than the one described here. If there is noise in the system or measurements, then stochastic notions must be used in the observer design. The resulting estimator is called the Kalman filter. *See* ESTIMATION THEORY; STOCHASTIC CONTROL THEORY.

Output feedback design. An alternative to state-variable feedback design, which generally requires the use of an observer, is to perform an optimal control design directly in terms of the available outputs of Eq. (14). This can result in significant simplifications in the control law, since the additional dynamics of an observer do not need to be built.

If, instead of the state feedback law of Eq. (2), the output feedback control law in Eq. (19) is used

$$u = -Ky + v \tag{19}$$

[where $v(t)$ is an external input], then the closed-loop system is given by Eq. (20). The output feed-

$$\dot{x} = (A - BKC)x + Bv \equiv A_c x + Bv \tag{20}$$

$$z = Hx$$

back linear quadratic regulator problem is to select the output feedback gain K so that the performance index of Eq. (4) is minimized.

In the infinite-horizon case, the gain K is constant. Here it can be shown that the optimal output feedback gain K may be found by solving Eqs. (21), where X, given by Eq. (22), is the expected initial

$$0 = A_c{}^T P + PA_c + C^T K^T RKC + Q \tag{21a}$$

$$0 = A_c S + SA_c{}^T + X \tag{21b}$$

$$K = R^{-1} B^T PSC^T (CSC^T)^{-1} \tag{21c}$$

$$X = E\{x(0)x^T(0)\} \tag{22}$$

mean-square state and S and P are matrix variables. Eqs. (21) are three coupled nonlinear matrix equations. These may be readily solved on a digital computer if the output of the system can be stabilized, that is, if there exists a gain K that makes A_c stable.

In output feedback design it is necessary to have some information about the initial state, namely X. This is not required in state feedback design. This means that the expected value of J, not J itself, must be minimized.

The optimal value of the expected value of the performance index using the optimal linear quadratic gain K is given in Eq. (23) [the trace of a matrix

$$J = {}^1/_2 \operatorname{trace}(PX) \tag{23}$$

is the sum of its diagonal elements]. As in the state feedback case, it may be computed off-line before the control input is ever applied to the system. Then, if it is too high, new values for Q, R, and $S(T)$ may be selected and the design repeated.

Tracking a reference input. The function of the linear quadratic regulator is to hold the performance output near zero. If it is desired for $z(t)$ to follow a nonzero reference input $r(t)$, then the control law must be modified. This is called the tracking or servo design problem.

It is possible to perform an optimal tracker design, where the tracking error $e(t)$ defined in Eq. (24) is

$$e = r - z \tag{24}$$

weighted in the performance index instead of $x(t)$. That is, $x^T Qx$ is replaced by $e^T Qe$ and $x^T(T)S(T)x(T)$ is replaced by $e^T(T)S(T)e(T)$. However, the result is a noncausal system that is difficult to implement.

Alternatively, the tracker problem may be solved by converting it first into a regulator problem. Suppose the system of Eq. (1) with the measured output of Eq. (14) has the applied control of Eq. (19). For perfect tracking, there must exist an ideal plant state x^* and an ideal plant input u^* such that the system of Eq. (1) has a performance output $z(t)$ equal to $r(t)$. If this is not so, then tracking with zero error is impossible. If the state, control, and output deviations are defined as in Eqs. (25), then it may be shown that the deviation has dynamics given by Eqs. (26).

$$\begin{aligned} \tilde{x} &= x - x^* & \tilde{u} &= u - u^* \\ \tilde{y} &= y - y^* & \tilde{z} &= z - z^* \end{aligned} \tag{25}$$

$$\dot{\tilde{x}} = A\tilde{x} + B\tilde{u} \tag{26a}$$

$$\tilde{y} = C\tilde{x} \tag{26b}$$

$$\tilde{z} = H\tilde{x} = -e \tag{26c}$$

Therefore, to keep the tracking error $e(t)$ small it is only necessary to design a regulator to regulate the state of the deviation system of Eqs. (26) to zero. This may be accomplished by finding the output feedback gain K that minimizes a quadratic performance index in terms of $\tilde{x}(t)$ and $\tilde{U}(t)$.

If constant output feedback gains K have been obtained that are optimal with respect to Eqs. (26), then Eq. (27) holds, so that the required control for the plant is given by Eq. (28). That is, the resulting con-

$$\tilde{u} = -K\tilde{y} \tag{27}$$

$$u = u^* + \tilde{u} = -Ky + u^* + Ky^* \tag{28}$$

trol for the servo or tracker problem is the optimal regulator control $-Ky$ plus some feedforward terms that are required to guarantee perfect tracking. These terms may be found by a transfer function analysis of the closed-loop system.

If $r(t)$ is a unit step, then the ideal responses are the steady-state responses, since $x^* = 0$. Then, the feedforward terms $u^* + Ky^*$ involve the inverse of the zero-frequency (direct-current) gain. That is, defining the closed-loop transfer function as in Eq. (29),

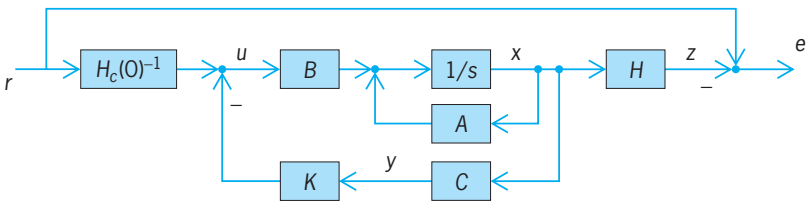

Fig. 3. Structure of tracker (servo) control system.

the servo control of Eq. (28) is given by Eq. (30).

$$H_c(s) = H[sI - (A - BKC)]^{-1}B \tag{29}$$

$$u = -Ky + H_c^{-1}(0)r \tag{30}$$

The structure of the servo control system is shown in **Fig. 3**. *See* SERVOMECHANISM.

Robustness and frequency-domain techniques. In addition to stability of the closed-loop system, an important property is guaranteed stability and performance in the presence of disturbance inputs, sensor noise, and variations in the plant parameters. This property is called robustness and is conveniently studied in the frequency domain.

Many of the classical frequency-domain design techniques may be extended to multivariable systems by using the notions of the singular values and principal phrases of a complex matrix. These quantities extend to nonscalar transfer functions the ideas of the transfer function magnitude and phase, respectively, and may be computed in a straightforward fashion by using modern computer software.

The loop gain of a multivariable closed-loop control system is the matrix $K(s)G(s)$, with $G(s)$ the plant and $K(s)$ the compensator, which may be designed by modern linear quadratic techniques. For good closed-loop disturbance rejection properties, the singular values of the loop gain should be large at low frequencies, for then the sensitivity is small. On the other hand, for good noise rejection, the singular values of the loop gain should be small at high frequencies. A representative plot of the maximum and minimum singular values of the loop gain (and $\bar{\sigma}$ and $\underline{\sigma}$, respectively) is shown in **Fig. 4**.

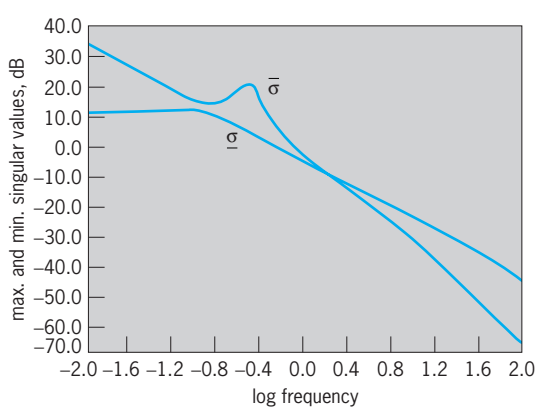

Fig. 4. Plot of maximum and minimum singular values of the loop gain, $\bar{\sigma}$ and $\underline{\sigma}$, as functions of the logarithm of frequency (Bode magnitude plot), for a representative multivariable closed-loop control system.

If the regulator $K(s)$ is a constant full state-variable feedback matrix K designed by using linear quadratic techniques, then the closed-loop system has important guaranteed robustness properties. In fact, it can then be shown that the return difference [that is, $I + K(s)G(s)$] satisfies the constraint in inequality (31).

$$\underline{\sigma}\left[I + KG(j\omega)\right] \geq 1 \tag{31}$$

Thus, the linear quadratic regulator always results in decreased sensitivity.

The linear quadratic constraint of inequality (31) may be used to show that the linear quadratic regulator has an infinite gain margin and a phase margin of at least 60°. The conclusion is that optimality is a stronger condition than stability.

Discrete-time systems. In order to control the system of Eq. (1) by using modern microprocessors, it must first be converted to a discrete-time system by sampling. Optimal control theory for discrete-time systems is well developed, and is in many respects similar to the theory outlined here for continuous-time systems. *See* CONTROL SYSTEMS; MICROPROCESSOR; SAMPLED-DATA CONTROL SYSTEM.

Frank L. Lewis

Bibliography. B. Anderson and J. B. Moore, *Optimal Control: Linear Quadratic Methods*, 1989; M. J. Grimble and M. A. Johnson, *Optimal Control and Stochastic Estimation: Theory and Applications*, 2 vols., 1988; S. Holly and A. H. Hallett, *Optimal Control, Expectations and Uncertainty*, 1989; F. L. Lewis, *Applied Optimal Control and Estimation*, 1992; F. L. Lewis, *Optimal Control*, 2d ed., 1995; E. Mosca, *Optimal, Predictive, and Adaptive Control*, 1995; E. O. Roxin, *Modern Optimal Control*, 1989.

Optimal control theory

An extension of the calculus of variations for dynamic systems with one independent variable, usually time, in which control (input) variables are determined to maximize (or minimize) some measure of the performance (output) of a system while satisfying specified constraints. The theory may be divided into two parts: optimal programming, where the control variables are determined as functions of time for a specified initial state of the system, and optimal feedback control, where the control variables are determined as functions of the current state of the system. *See* CALCULUS OF VARIATIONS.

Examples of optimal control problems are (1) determining paths of vehicles or robots between two points to minimize fuel or time, (2) determining time-varying feedforward/feedback logic to bring a vehicle, robot, or industrial process to a desired terminal state in a finite time using acceptable control magnitudes, and (3) determining constant-gain feedback control logic for vehicles or industrial processes to keep them near a desired operating point in the presence of disturbances with acceptable control magnitudes.

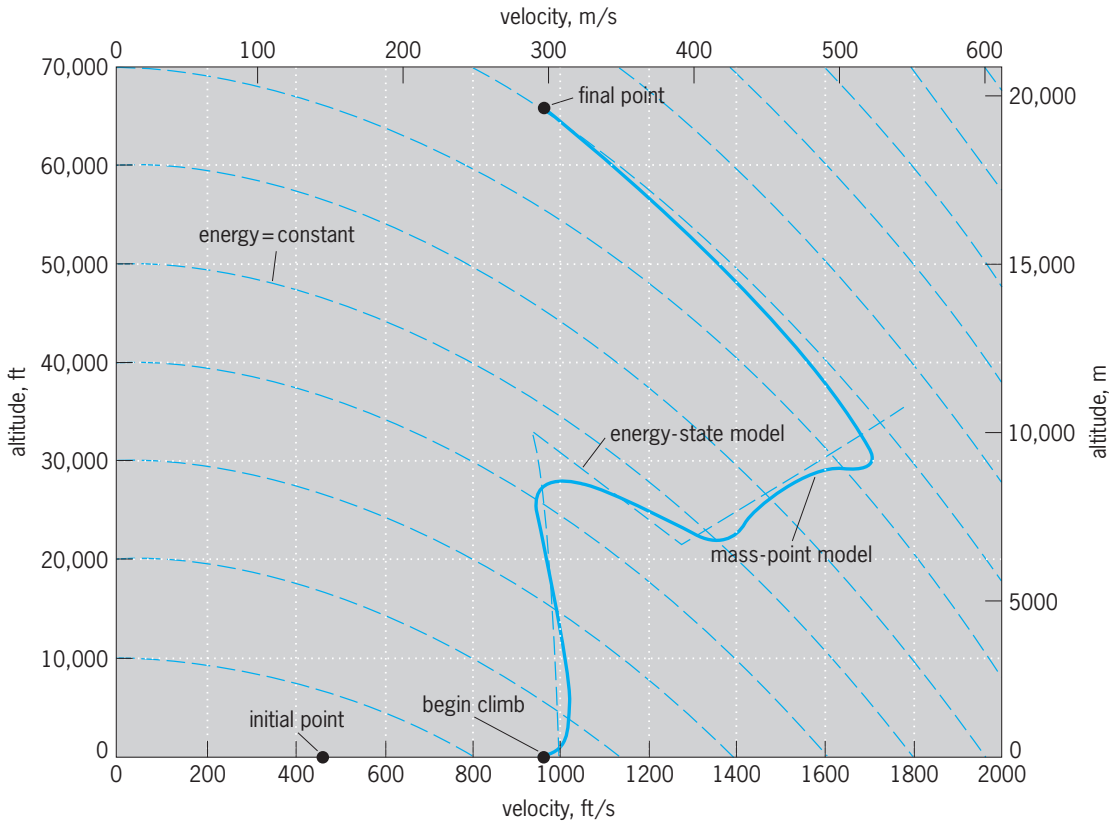

Fig. 1. Minimum-time flight path of a supersonic aircraft from takeoff to an altitude of 20,000 m (65,600 ft), velocity of Mach 1.0, and horizontal flight.

Dynamic systems may be divided into two categories: continuous dynamic systems, where the control and state variables are functions of a continuous independent variable, such as time or distance; and discrete dynamic systems, where the independent variable changes in discrete increments. Many discrete systems are discretized versions of continuous systems; the discretization is often made so that (1) the system can be analyzed or controlled by digital computers (or both), or (2) measurements of continuous outputs are made at discrete intervals of time (sampled-data systems) in order to share data transmission channels.

A large class of interesting optimal control problems can be described as follows: the dynamic system is described by a set of coupled first-order ordinary differential equations of the form of Eq. (1), where x

$$\dot{x} = f(x, u, t) \tag{1}$$

(the state vector) represents the state variables of the system (such as position, velocity, temperature, voltage, and so on); u (the control vector) represents the control variables of the system (such as motor torque, control-surface deflection angle, valve opening, and so on); t is time; and f reprensents a set of functions of x, u, and t. The performance index J, which one desires to minimize, is a scalar function of the final time t_f and the final state $x(t_f)$ plus an integral from the initial time t_0 to the final time t_f of a scalar function of the state and control vectors and time, as given in Eq. (2). Possible constraints include

$$J = \varphi[x(t_f), t_f] + \int_{t_0}^{t_f} L\,[x(t), u(t), t]\ dt \tag{2}$$

(*a*) specified vector functions constraining the initial time t_0 and the initial state $x(t_0)$, as in Eq. 3; (*b*) spec-

$$\alpha[x(t_0), t_0] = 0 \tag{3}$$

ified vector functions constraining the final time and final state, as in Eq. (4); (*c*) specified vector functions

$$\psi[x(t_f), t_f] = 0 \tag{4}$$

constraining the control variables [inequality (5)];

$$C[u(t), t] \leq 0 \tag{5}$$

(*d*) specified vector functions constraining both control and state variables [inequality (6)]; (*e*) specified

$$CS[x(t), u(t), t] \leq 0 \tag{6}$$

vector functions constraining only the state variables [inequality (7)]. Constraints *a* and *b* are equality con-

$$S[x(t), t] \leq 0 \tag{7}$$

straints, whereas constraints *c*, *d*, and *e* are inequality constraints. *See* OPTIMIZATION.

Necessary conditions. The calculus of variations is the basic technique of optimal control theory. To

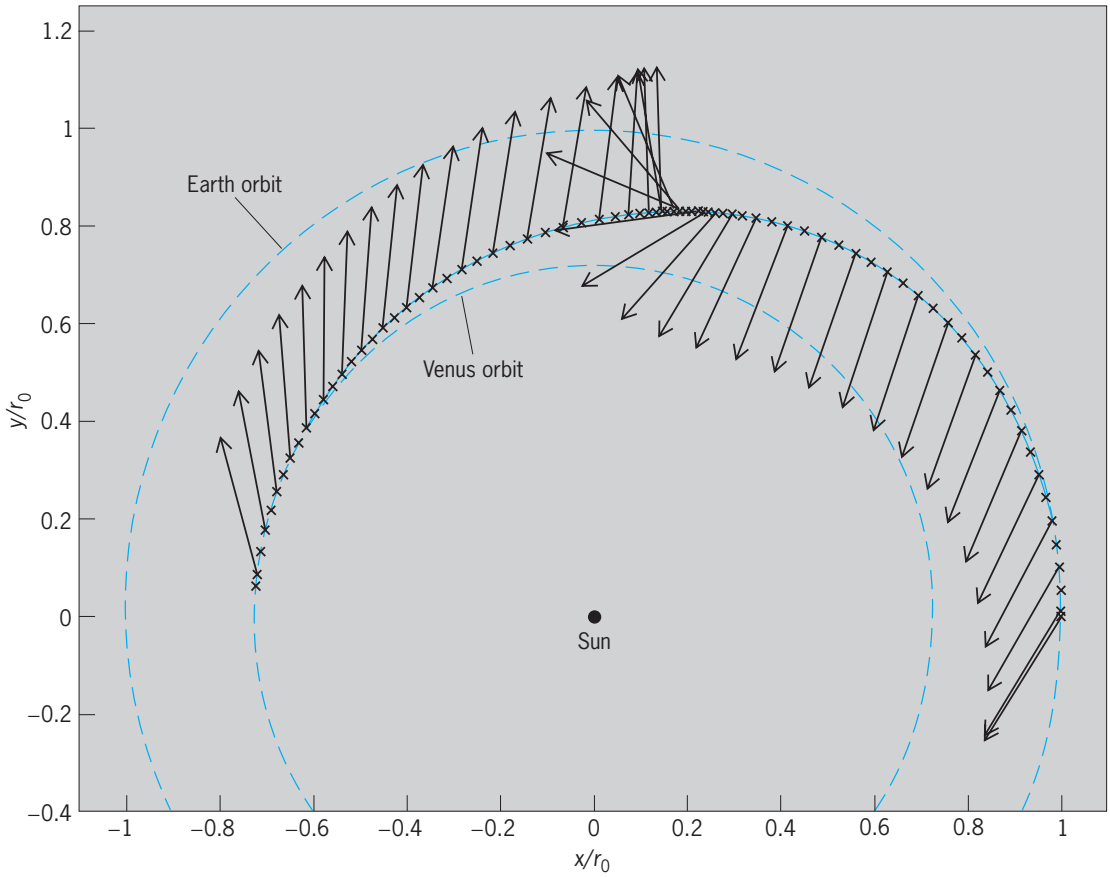

Fig. 2. Minimum-time transfer orbit of a spacecraft from Earth orbit to Venus orbit with low constant thrust. Arrows show direction of thrust.

illustrate the use of necessary and sufficient conditions for a minimum of J, consider the case where (1) the functions of $f(x,u,t)$, $\varphi(x,t)$, and $\psi(x,t)$ are continuous and have continuous first and second partial derivatives; (2) t_0, t_f, and $x(t)$ must satisfy Eqs. (8)–(11), where notation (12) is used ($a \triangleq b$ indicates

$$\dot{x} = f(x, u, t) \qquad x(t_0) \text{ specified} \tag{8}$$

$$0 = \frac{\partial H}{\partial u} \longrightarrow u = u(x, \lambda, t) \tag{9}$$

$$\dot{\lambda}^T = -\frac{\partial H}{\partial x} \qquad \lambda^T(t_f) = \frac{\partial \Phi}{\partial x(t_f)} \tag{10}$$

$$\psi[x(t_f)] = 0 \tag{11}$$

$$\begin{aligned} H &\triangleq L + \lambda^T f \\ \Phi &\triangleq \phi + \nu^T \psi \end{aligned} \tag{12}$$

that a is defined as equal to b); H is called the variational hamiltonian, $\lambda^T(t)$ is a row vector of influence functions (also called Lagrange multiplier functions, adjoint functions, and co-state functions); ν^T is a row vector of influence parameters; and the partial derivatives are evaluated on the optimal path. Equations (8)–(11) constitute a two-point boundary-value problem. Differential equations (8) and (10) are coupled, since x depends on u and, from Eq. (9), u depends on λ, whereas λ depends on x because $\partial f/\partial x$ and $\partial L/\partial x$ are functions of x. The parameters ν must be chosen so that Eq. (11) is satisfied. Equations (9) and (10) are called the Euler-Lagrange equations [although some authors refer to the whole set, Eqs. (8)–(11), by this name]. A stationary solution is one that satisfies Eqs. (8)–(11). Such solutions may not exist for a given problem, and even if a solution is found, it may not be minimizing. A minimizing solution must also satisfy the condition of inequality (13)

$$\frac{\partial^2 H}{\partial u^2} \geq 0 \qquad t_0 \leq t \leq t_f \tag{13}$$

[due to A. M. Legendre and A. Clebsch]. If a solution to Eqs. (8)–(11) satisfies the strong version of inequality (13), that is, inequality (14), then it is a locally min-

$$\frac{\partial^2 H}{\partial u^2} \geq 0 \qquad t_0 \leq t \leq t_f \tag{14}$$

imizing solution. Occasionally, more than one locally minimizing solution exists for a given problem. The locally minimizing solution that gives the smallest J is called the globally minimizing solution. Unfortunately, one cannot always be sure that all of the locally minimizing solutions have been found.

Gradient and shooting algorithms. Realistic optimal control problems must be solved numerically with a digital computer. Gradient algorithms start with a guess of the optimal control history $u(t)$, which determines a nominal state history $x(t)$. Backward

integration of $\lambda(t)$ over this nominal path yields the gradient $\partial H/\partial u$, which determines small changes in $u(t)$ that bring the path closer to feasibility and optimality. After 10–50 iterations, the gradient is close to zero over the whole path and the terminal conditions are met. One common shooting algorithm integrates the coupled Euler-Lagrange equations forward with several initial guesses of $\lambda(t_0)$, determining $u(t)$ from Eq. (9) at each integration step. The correct values of $\lambda(t_0)$ are then interpolated by zeroing the errors in the terminal conditions. Good initial guesses (for examples, from a gradient solution) are required for shooting algorithms to converge since the Euler-Lagrange equations are inherently unstable.

Examples of gradient solutions are shown in **Figs. 1** and **2**. Figure 1 shows the minimum-time flight path of a supersonic aircraft from takeoff to an altitude of 20 km (12 mi), arrriving at Mach number 1.0 and horizontal flight. Full throttle is used all the way; the control is angle-of-attack, which can be changed very rapidly using the elevators. Two solutions are shown, one using a mass-point model with five states and another using a less precise energy-state model with only two states. Figure 2 shows the minimum-time flight path of a spacecraft from Earth orbit to the orbit of Venus. The control is the direction of the low constant thrust force which is shown by arrows in the figure. The spacecraft arrives with the Venus orbital velocity and tangent to the Venus orbit.

Minimum principle. The minimum principle is a generalization of Eq. (14), due to L. S. Pontryagin. It is a statement of sufficient conditions for a local minimum of J for problems with (or without) control variable constraints. In the previous development, Eq. (9) and inequality (13) are replaced by Eq. (15),

$$u = \arg\min_{u} H[x(t), \lambda(t), u(t), t] \tag{15}$$

where H is, again, the variational hamiltonian defined in Eq. (12), and the minimization is carried out over the entire admissible range of u, holding x, λ, and t fixed.

Nonlinear programming algorithms may be used to solve optimal control problems using the minimum principle by parametrizing the control or output variables. They are especially useful for problems with control or state variable inequality constraints since several professional nonlinear programming codes are available. **Figure 3** shows an example, the minimum-time erection of a pendulum from hanging straight down with no angular velocity to straight up with no angular velocity using bounded horizontal force (the control) on the supporting cart. The cart starts and ends at the same position with zero velocity. This is an example of "bang-bang control" since the specific force changes from 1 g (the acceleration of gravity) to the left to 1 g to the right five times during the erection. The figure is a stroboscopic movie of the motion; the square symbols indicate the changing position of the cart while the circle symbols indicate the position of the pendulum bob. *See* NONLINEAR CONTROL THEORY; NONLINEAR PROGRAMMING.

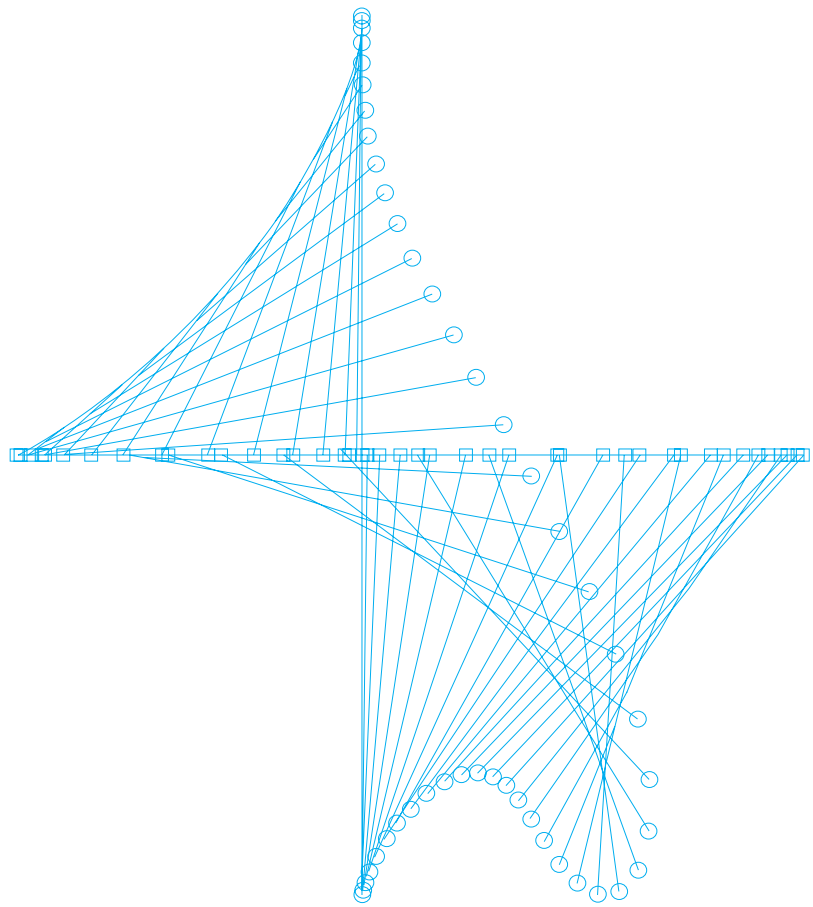

Fig. 3. Minimum-time erection of an inverted pendulum using bounded horizontal force on the supporting cart.

Dynamic programming. This is a technique for determining families of optimum solutions in feedback form, that is, $u = u(x)$. It is based on the principle of optimality, which may be stated as follows: if an optimal path is known from some point in the state-time space $[x(t_0),t_0]$ to a terminal manifold {a set of terminal conditions $\psi[x(t_f,t_f] = 0$}, then this path is also the optimal path to this terminal manifold from any intermediate point on the path. This simple and rather obvious principle can be used to derive a first-order partial differential equation for the optimal return function, $J^{(0)}[x(t),t]$, which is defined as the value of the performance index J in Eq. (2), starting from $x(t)$ at time t and proceeding on an optimum path to the terminal manifold. The partial differential equation is Eq. (16), where Eqs. (17) are

$$\frac{\partial J^{(0)}}{\partial t} = -H^{(0)}\left(x, \frac{\partial J^{(0)}}{\partial x}, t\right) \tag{16}$$

$$\begin{gathered} H^{(0)}\left(x, \frac{\partial J^{(0)}}{\partial x}, t\right) \triangleq \min_{u} H\left[x, \frac{\partial J^{(0)}}{\partial x}, u, t\right] \\ H \triangleq L(x, u, t) + \frac{\partial J^{(0)}}{\partial x} f(x, u, t) \end{gathered} \tag{17}$$

used and the solution must satisfy the terminal condition of Eq. (18). An important by-product of solving

$$J^{(0)}(x, t) = \varphi(x, t) \quad \text{or} \quad \psi(x, t) = 0 \tag{18}$$

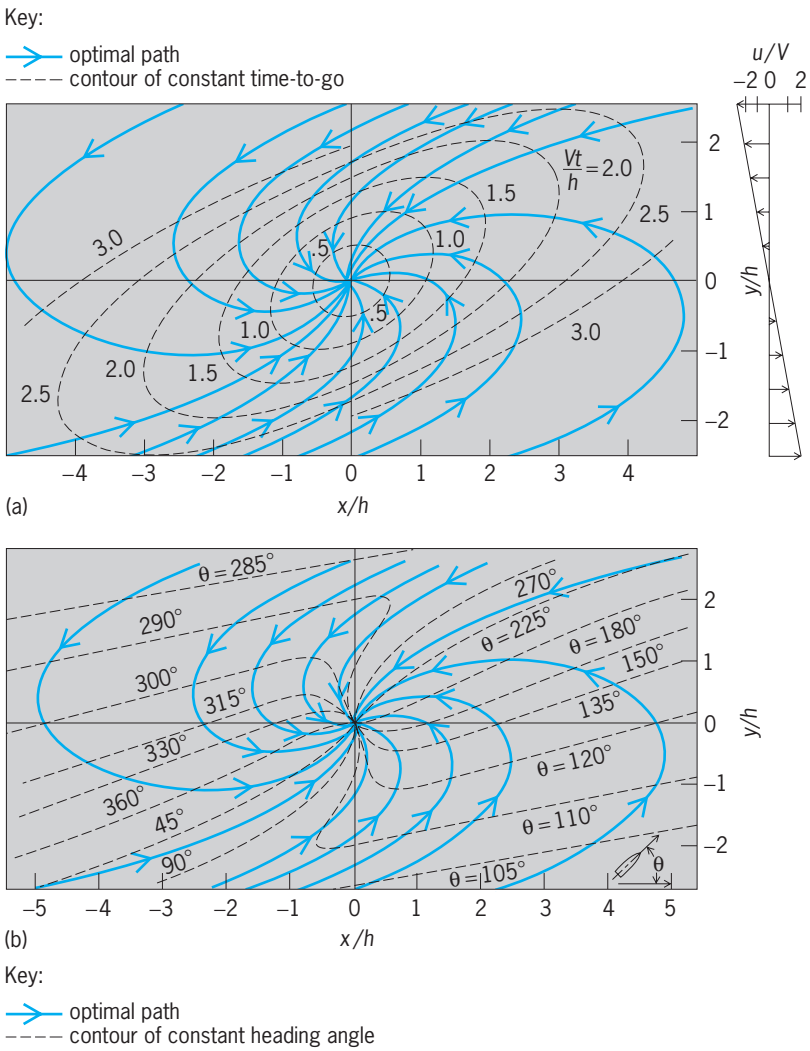

Fig. 4. Minimum-time ship routing with linear variation in current. (*a*) Contours of constant time-to-go and optimal paths. (*b*) Contours of constant heading angle θ and optimal paths.

Eqs. (16) and (17) is the optimal control law (19),

$$u^{(0)}(x) \quad \text{or} \quad u^{(0)}(x, t) \tag{19}$$

which gives the control variables as functions of the state variables (or state variables and time if f or L depend explicitly on time). Comparison on Eqs. (12) and (17) yields Eq. (20), that is, the Lagrange multi-

$$\frac{\partial J^{(0)}}{\partial x} \equiv \lambda^T \tag{20}$$

plier functions introduced in Eqs. (10) are influence functions on the value of the optimal return function for infinitesimal changes in x. Solutions to Eqs. (17) and (19) are constructed backward from the terminal manifold $\psi(x,t) = 0$ and have a wavelike character; that is, contours of constant $J^{(0)}$ are like wavefronts and optimal paths are like rays.

As an example, consider the problem of finding minimum-time paths to $x = y = 0$ for a ship that travels with constant velocity V with respect to the water, but there is a current parallel to the x axis whose magnitude is $-V(y/h)$; hence, at $y = h$, the velocity of the current is in the negative x direction and equals the ship's velocity with respect to the water. The problem may be stated as follows: find $\theta(x,y)$, the heading of the ship as function of position, to go to $x = y = 0$ in minimum time where Eqs. (21) and (22) are given. Optimal paths and

$$\dot{x} = V\cos\theta - \frac{V}{h}y \tag{21}$$

$$\dot{y} = V\sin\theta \tag{22}$$

contours of constant time-to-go (measured in units of h/V) are shown in **Fig. 4*a***. Optimal paths and contours of constant heading angle are shown in Fig. 4*b*; this chart is the feedback control solution since it gives $\theta(x,y)$, that is, optimal heading as a function of position. For continuous dynamic systems with three or more state variables, solutions to Eqs. (17) and (19) require too much computation to be considered practical at present. If feedback solutions are desired for such systems, neighboring optimum feedback control (discussed below) is a more practical approach.

Singular optimal control. In some optimal control problems, extremal arcs ($\partial H/\partial u = 0$) occur on which the matrix $\partial^2 H/\partial u^2$ is only semidefinite; that is, it has one or more zero eigenvalues. Such arcs are called singular arcs. In the classical calculus of variations, an optimal path that contains one or more singular arcs is called a broken extremal since the control variables have discontinuities ("corners") at one or both ends of the singular arcs. Singular arcs may appear in optimal control problems involving cascaded dynamic systems if the direct effect of the control and the effect of a strongly controlled state have opposite signs. Cascaded linear systems of this type are called nonminimum phase (NMP) systems and are characterized by having right-half-plane transmission zeros in the transfer function matrix.

NOFC. Neighboring optimum feedback control (NOFC) about a nominal optimal path is a viable alternative to a "complete" optimal feedback control solution. NOFC constitutes an approximation to the optimal feedback control solution in a region neighboring the nominal optimal path; in the case of the ship routing problem illustrated in Fig. 4, an example of such region is shown in **Fig. 5**. NOFC is a combination of the nonlinear open-loop optimal path plus a linear-quadratic closed-loop solution in the neighborhood of this path. The control logic is of the form in Eq. (23), where $u^{(N)}(t)$ and $x^{(N)}(t)$ are

$$u = u^{(N)}(t) + C(t)\left[x(t) - x^{(N)}(t)\right] \tag{23}$$

the optimal control and state variable programs for the nominal initial and final conditions; $x(t)$ is the current (pertubed) state and $u(t)$ is the current control that should be used in order to proceed along a neighboring optimal path to the specified terminal manifold; and $C(t)$ is the feedback gain matrix obtained by solving a slight generalization of the matrix Riccati equation, using weighting matrices that are second derivatives of the variational hamiltonian

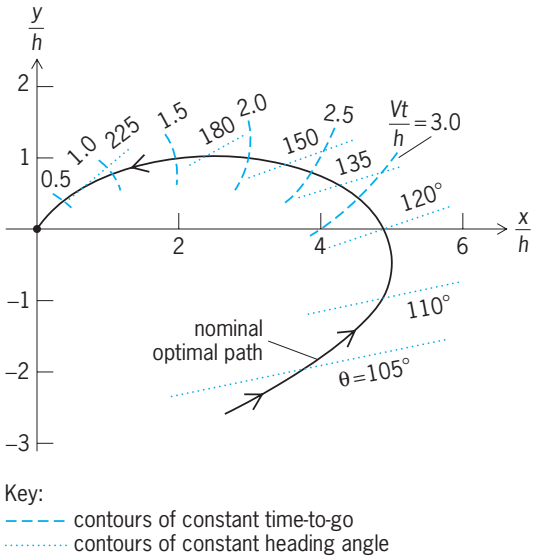

Fig. 5. Neighboring optimum feedback control about a nominal optimal path for the minimum-time ship-routing problem with linear variation in current.

[defined in Eq. (12)], evaluated on the nominally optimal path. *See* OPTIMAL CONTROL (LINEAR SYSTEMS).

If the final time t_f is determined implicitly by the point where the path intersects the terminal manifold [Eq. (4)], the NOFC law, Eq. (23), may be modified to use estimated time-to-go, $t_f^{(e)} - t = \tau$, as the index variable for $u^{(N)}$, $x^{(N)}$, and C; an additional set of feedback gains n must be determined such that Eq. (24) is satisfied. Equation (24) may have to be

$$t_f^{(e)} = t_f + n(\tau)\left[x(t) - x^{(N)}(\tau)\right] \qquad (24)$$

used iteratively to determine a consistent τ, that is, start with $\tau_0 = t_f - t$ to find $\tau_{f0}^{(e)}$, then use $\tau_1 = t_{f0}^{(e)} - t$, then $\tau = t_{f1}^{(e)}$, and so on, until $\tau_{n+1} \cong \tau_n$. *See* CONTROL SYSTEMS. Arthur E. Bryson, Jr.

Bibliography. B. Anderson and J. B. Moore, *Optimal Control: Linear Quadratic Methods*, 1989; A. E. Bryson, *Dynamic Optimization*, 1998; D. N. Kirk, *Optimal Control Theory: An Introduction*, 1992; F. L. Lewis, *Optimal Control*, 2d ed., 1995; E. O. Roxin, *Modern Optimal Control*, 1989; R. F. Stengel, *Stochastic Optimal Control*, 1986.

Optimization

The design and operation of systems or processes to make them as good as possible in some defined sense. The approaches to optimizing systems are varied and depend on the type of system involved, but the goal of all optimization procedures is to obtain the best results possible (again, in some defined sense) subject to restrictions or constraints that are imposed. While a system may be optimized by treating the system itself, by adjusting various parameters of the process in an effort to obtain better results, it generally is more economical to develop a model of the process and to analyze performance changes that result from adjustments in the model. In many applications, the process to be optimized can be formulated as a mathematical model; with the advent of high-speed computers, very large and complex systems can be modeled, and optimization can yield substantially improved benefits. *See* DIGITAL COMPUTER.

Optimization is applied in virtually all areas of human endeavor, including engineering system design, optical system design, economics, power systems, water and land use, transportation systems, scheduling systems, resource allocation, personnel planning, portfolio selection, mining operations, blending of raw materials, structural design, and control systems. Optimizers or decision makers use optimization in the design of systems and processes, in the production of products, and in the operation of systems.

Framework for Optimization

The framework for optimization includes forming a model, establishing and treating constraints, determining feasible solutions, and assigning performance measures.

System models. The first step in modern optimization is to obtain a mathematical description of the process or the system to be optimized. A mathematical model of the process or system is then formed on the basis of this description. Depending on the application, the model complexity can range from very simple to extremely complex. An example of a simple model is one that depends on only a single nonlinear algebraic function of one variable to be selected by the optimizer (the decision maker). Complex models may contain thousands of linear and nonlinear functions of many variables. As part of the procedure, the optimizer may select specific values for some of the variables, assign variables that are functions of time or other independent variables, satisfy constraints that are imposed on the variables, satisfy certain goals, and account for uncertainties or random aspects of the system.

System models used in optimization are classified in various ways, such as linear versus nonlinear, static versus dynamic, deterministic versus stochastic, or time-invariant versus time-varying. In forming a model for use with optimization, all of the important aspects of the problem should be included, so that they will be taken into account in the solution. The model can improve visualization of many interconnected aspects of the problem that cannot be grasped on the basis of the individual parts alone. A given system can have many different models that differ in detail and complexity. Certain models (for example, linear programming models) lend themselves to rapid and well-developed solution algorithms, whereas other models may not. When choosing between equally valid models, therefore, those that are cast in standard optimization forms are to be preferred. *See* MODEL THEORY.

Constraints. The model of a system must account for constraints that are imposed on the system. Constraints restrict the values that can be assumed by variables of a system. Constraints often are classified as being either equality or inequality constraints.

For example, Eq. (1) is a linear equality constraint,

$$x_1 + x_2 + x_3 = 50 \tag{1}$$

whereas inequality (2*a*) is a nonlinear inequality constraint and is equivalent to inequality (2*b*).

$$x_1^2 + x_2^2 + x_3^2 \leq 100 \tag{2a}$$

$$-x_1^2 - x_2^2 - x_3^2 \geq -100 \tag{2b}$$

Constraints (1) and (2) are algebraic constraints and depend on the three real variables x_1, x_2, and x_3, which can be viewed as characterizing three-dimensional euclidean space. In the case of n real variables, the n- dimensional euclidean space R^n is of interest. Other optimization problems involve differential equation constraints, an example of which is Eq. (3).

$$\frac{dx_1}{dt} = -0.5x_1 + 3x_2 \tag{3}$$

See DIFFERENTIAL EQUATION.

The types of constraints involved in any given problem are determined by the physical nature of the problem and by the level of complexity used in forming the mathematical model.

Constraints that must be satisfied are called rigid constraints. Physical variables often are restricted to be nonnegative; for example, the amount of a given material used in a system is required to be greater than or equal to zero. Rigid constraints also may be imposed by government regulations or by customer-mandated requirements. Such constraints may be viewed as absolute goals.

In contrast to rigid constraints, soft constraints are those constraints that are negotiable to some degree. These constraints can be viewed as goals that are associated with target values. In place of Eq. (1), for example, a goal might be to make the left-hand side (a performance index perhaps) as close as possible to the right-hand side, the target value. The amount that the goal deviates from its target value could be considered in evaluating trade-offs between alternative solutions to the given problem.

Feasible solutions. When constraints have been established, it is important to determine if there are any solutions to the problem that simultaneously satisfy all of the constraints. Any such solution is called a feasible solution, or a feasible point in the case of algebraic problems. The set of all feasible points constitutes the feasible region. For example, if Eq. (1) and inequality (2) are the constraints that apply to a given problem, the feasible region is the intersection of two sets of points: the set of points on the plane where the x_i's sum to 50 and the set of points within the sphere of radius 10 that is centered at the origin.

If no feasible solution exists for a given optimization problem, the decision maker may relax some of the soft constraints in an attempt to create one or more feasible solutions; a class of approaches to optimization under the general heading of goal programming may be employed to relax soft constraints in a systematic way to minimize some measure of maximum deviations from goals.

Performance measures. A key step in the formulation of any optimization problem is the assignment of performance measures (also called performance indices, cost functions, return functions, criterion functions, and performance objectives) that are to be optimized. The success of any optimization result is critically dependent on the selection of meaningful performance measures. In many cases, the actual computational solution approach is secondary. Ways in which multiple performance measures can be incorporated in the optimization process are varied. Both single-criterion and multicriteria optimization methods are discussed below.

In some cases, all but one of the performance measures are assigned target values and are converted into soft constraints; the remaining performance measure, say $f(x)$, is then optimized (maximized or minimized, whichever is appropriate) subject to satisfying the constraints. The validity of Eq. (4) allows

$$\min[f(x)] = -\max[-f(x)] \tag{4}$$

an ordinary optimization problem to be cast as either a maximization or minimization problem. The resulting optimization problem is called a single-objective or single-criterion optimization problem, in contrast to multicriteria optimization problems.

The way in which a system is modeled often dictates approaches to obtaining optimal solutions. Unconstrained optimization, constrained optimization with a single performance measure, constrained optimization with multiple performance measures, and optimization of systems that evolve over time (dynamic optimization) each require fundamentally different approaches that will be discussed in turn.

Unconstrained Optimization

A special class of system models are those in which there are no constraints. The independent variable is the column vector $x = [x_1, x_2, \ldots, x_n]^T$ of n variables in n-dimensional euclidean space R^n of real values; the superscript T signifies the transpose of a row vector to form the column vector x. Given a real scalar function $f(x)$ of the n variables, points at which $f(x)$ assumes local maximum and local minimum values are of interest.

Points of zero slope. If $f(x)$ and its first derivatives are continuous, then the maximum and minimum of f (if they exist) must occur at points where the slope of the function is zero, that is, at points where the gradient of f, ∇f, satisfies Eq. (5). **Figure 1** illus-

$$\nabla f(x) \triangleq \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right]^T = 0 \tag{5}$$

trates local maxima and local minima of a function of two variables. It contains a local maximum of 7 in addition to the global maximum of 9, and both points in the x_1,x_2 plane satisfy the necessary condition of Eq. (5), asdoes the minimum point (where the

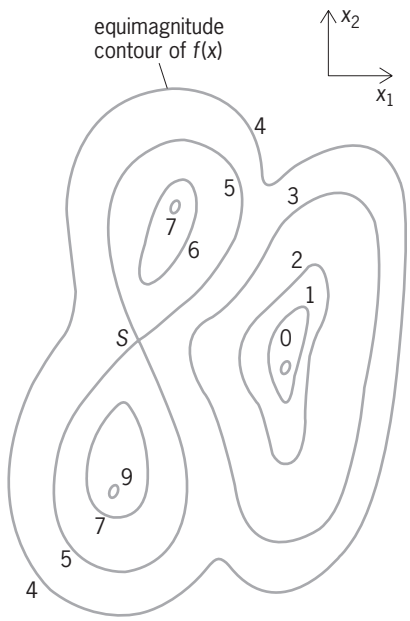

Fig. 1. Equimagnitude contours of a function of two variables.

function equals 0). In addition, the point marked S, which is known as a saddle point, also satisfies Eq. (5). *See* CALCULUS OF VECTORS.

Second-order conditions. To determine if a point x^* that satisfies Eq. (5) is a local maximum point, the determinant of the $n \times n$ matrix of second derivatives of $f(x)$ [the hessian matrix of $f(x)$] may be tested at x^*: if this hessian is negative definite, x^* is a local maximum point; if the hessian is negative semidefinite, x^* may be a local maximum point; otherwise x^* is not a local maximum point. The test for a local minimum point is similar, with "negative" in the preceding statement being replaced by "positive."

Global versus local. To determine if a local maximum (minimum) is the global maximum (minimum), an exhaustive search over all local maxima (minima) may be conducted. If the function $f(x)$ is known to be concave (convex), however, there is only one local maximum (minimum), and therefore also global maximum (minimum), assuming the function is bounded.

Numerical solution methods. A multitude of numerical methods called search techniques have been developed to determine maxima and minima of functions. The most appropriate method to use depends on the nature of $f(x)$. If $f(x)$ is evaluated experimentally, a direct sequential search approach that does not require derivative information can be used. If $f(x)$ is discontinuous, a factorial search or a random search (Monte Carlo method) may be useful. If $f(x)$ and its derivatives are continuous, however, a method from the class of methods that exhibit quadratic convergence may be best. A maximization technique exhibits quadratic convergence if it can locate the exact maximum of a quadratic function (assuming it has a well-defined maximum) in a finite number of calculations. *See* MONTE CARLO METHOD; NUMERICAL ANALYSIS.

Gradient-based methods. The gradient vector $\nabla f(x)$ is central to those search techniques that use derivative information. In Fig. 1, for example, at each point in the figure, the gradient vector is a vector that is perpendicular to the equimagnitude contour and points in the direction of greatest incremental increase in $f(x)$. One simple search strategy is as follows: (1) evaluate a search direction $r = \nabla f(x^\circ)$ at the current best $x = x^\circ$; (2) let $x = x^\circ + \upsilon r$, where υ is a real scalar; (3) search for a maximum of f with respect to values of $\upsilon \geq 0$; and (4) reassign the current best x and return to step 1. It can be shown, however, that this simple gradient search strategy is inefficient when the equimagnitude contours are elongated or irregular in any way, and better gradient-based techniques are available.

Constrained Optimization with Single Performance Measure

Optimization problems may be solved by a variety of techniques that come under the general heading of mathematical programming. The term programming in this regard originated in the sense of scheduling (programming) a solution to the problem and was introduced prior to the widespread availability of computers. A variety of problem types and solution methods are associated with nonlinear programming, linear programming, integer programming, quadratic programming, dynamic programming, goal programming, parametric programming, and so forth.

Nonlinear programming. A common optimization problem is the nonlinear programming problem given by expression (6), where the function $f(x)$ is

$$\text{Maximize } f(x) \qquad x \in S \tag{6}$$

a real scalar-valued function of n real variables contained in the vector x, and where x is constrained to be an element of the set S defined both by equality constraints given by Eq. (7) and by inequality constraints given by inequality (8). The constraint functions $p_i(x) = a_i \qquad i = 1, 2, \ldots, n_a < n$ (7)

$$q_j(x) \leq b_j \qquad j = 1, 2, \ldots, n_b \tag{8}$$

tions p_i's and q_j's are real scalar-valued functions of x; and the a_i's and b_j's are real constants.

Related minimization problems can be cast in the above form by use of Eq. (4). If the constraints of the problem define a convex set of points, and the performance measure is concave (convex), it can be shown that the only local maximum (minimum) is the global maximum (minimum) if it exists.

Karush-Kuhn-Tucker conditions. If the functions in expressions (6), (7), and (8) and the derivatives through second order of the functions are continuous, necessary conditions and sufficient conditions for constrained local maxima are of interest. Such conditions were used by J. L. Lagrange in his formulation of Lagrange multipliers for cases where all of the constraints are equality constraints. W. Karush first presented conditions for testing inequality constraints

with Lagrange multipliers in 1939, an approach that was extended and popularized by H. W. Kuhn and A. W. Tucker in 1951.

The Karush-Kuhn-Tucker conditions depend on the lagrangian function L given by Eq. (9), where

$$L(x, \alpha, \beta) \triangleq f + \sum_{i=1}^{n_a}(a_i - p_i)\alpha_i + \sum_{j=1}^{n_b}(b_j - q_j)\beta_j \quad (9)$$

α and β are column vectors of Lagrange multipliers, the α_i's and β_j's, and where the other terms in Eq. (9) are defined by expressions (6), (7), and (8).

Technical details concerning Karush-Kuhn-Tucker conditions are simplified if the gradients of the active constraints are linearly independent at points of constrained local maxima. A given inequality constraint is said to be active at a point x if it is satisfied in an equality sense at x. Equality constraints necessarily are active constraints at points of constrained local maxima. A point at which the gradient ∇f of f is nonzero and at which the gradients of the active constraints are linearly independent is called a regular point.

First-order necessary conditions for a regular point x^* to be a point of constrained local maximum are as follows: (1) all constraints of Eq. (7) and inequality (8) are satisfied at x^*; (2) $\nabla L(x^*,\alpha^*,\beta^*) = 0$ for some real vectors α^* and β^*; (3) all β_j^*'s are greater than or equal to zero; and (4) if the jth inequality constraint is not active at x^*, then $\beta_j^* = 0$. The ∇L in condition (2) is the gradient of L with respect to x. In effect, these first-order conditions state that the gradient of the performance measure f is a linear combination of the gradients of the active constraints at x^*, with the inequality constraint proportionality factors, the β_j's, necessarily being nonnegative at a constrained local maximum.

Figure 2 depicts the case of one equality constraint and two variables x_1 and x_2. At a typical nonoptimal point x^0 that satisfies the constraint, no multiplier α exists for which $\nabla f = \alpha \nabla p$. At points x^1, x^2, x^3, and x^4, however, ∇f and ∇p are collinear, and real α's exist for which $\nabla f = \alpha \nabla p$. The point x^1 is a constrained local maximum; x^2 is a constrained local minimum; x^3 is the constrained global maximum; and x^4 is a saddle point, which satisfies the first-order necessary conditions but which is neither a constrained maximum nor a constrained minimum.

Figure 3 depicts the case of two inequality constraints and two variables x_1 and x_2. In the interior of the shaded region, both inequality constraints are strictly satisfied (in which case the constraints are not active). On the common boundary point shown, however, both constraints are active, ∇f is a positive linear combination of ∇q_1 and ∇q_2, and the point is a constrained maximum point.

Sensitivity results and higher-order conditions. The Lagrange multipliers associated with a given local maximum can be shown to provide the following sensitivity information. If the right-hand sides of Eq. (7) and inequality (8) can be changed by small amounts Δa_i and Δb_j, respectively, then the constrained local maximum will change by an amount Δf, given by expression (10). Depending on the application, the

$$\Delta f \approx \sum_{i=1}^{n_a} \alpha_i^* \Delta a_i + \sum_{j=1}^{n_b} \beta_j^* \Delta b_j \quad (10)$$

Lagrange multipliers (the dual variables) in expression (10) might represent shadow prices or attribute costs. The costs associated with changes in a's and b's can be weighed against the improvements obtainable in f. The approximation (10) is based on

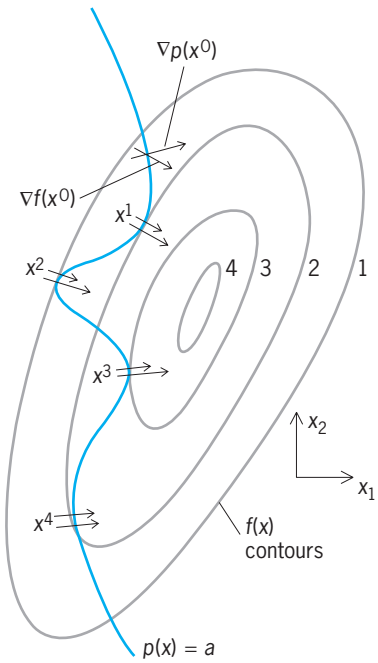

Fig. 2. Conditions for local optima with an equality constraint.

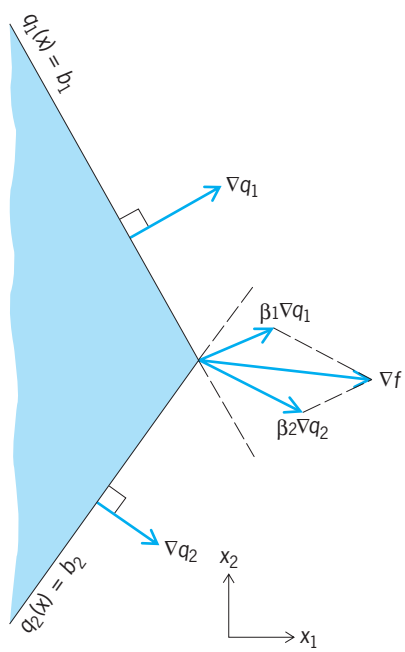

Fig. 3. Conditions for local maximum with inequality constraints.

first-order terms of Taylor series expansions and should not be used if the Δa_i's and Δb_j's are large enough to cause a change in constraint basis, that is, a case in which one or more active (inactive) inequality constraints become inactive (active).

Reasoning based on Taylor series expansions of problem functions can be used to obtain many other sensitivity results and to obtain second-order necessary conditions for local maxima, as well as second-order sufficient conditions.

Solution methods. Many numerical solution methods have been developed for nonlinear programming problems. Approaches to solution include the following: (1) gradient projection methods, in which the gradient of the performance measure is projected onto the constraint boundary as the search progresses; (2) penalty function methods, wherein a penalized performance measure is formed to equal the original performance measure when the constraints are satisfied, but to grow rapidly in magnitude as constraints are violated; and (3) augmented lagrangian methods, which modify the lagrangian of the problem in ways that ensure that the augmented lagrangian has unconstrained local maxima with respect to x at points of constrained local maxima.

When derivatives of problem functions are discontinuous, Monte Carlo or random search methods are popular approaches to solution. When integer or discrete value restrictions are imposed, various combinatorial solution methods may be appropriate. *See* COMBINATORIAL THEORY; NONLINEAR PROGRAMMING.

Linear programming. When all of the functions in expressions (6), (7), and (8) are linear, the problem reduces to a linear programming problem. Linear programming problems have a well-defined mathematical structure: the constraint set S is known to be convex; boundary surfaces of S are known to be planar and to intersect so as to define extreme points (vertices of simplexes); the global maximum of f (if it exists) is known to occur at a vertex of S (but the global maximum may also occur at neighboring boundary points); and the global maximum is the only maximum since no local maxima occur at other vertices. Every linear programming problem has a well-defined dual problem, the solution of which is related to the solution of the original problem and which provides sensitivity information regarding the original problem. *See* LINEAR PROGRAMMING.

Multiobjective Optimization

When more than one performance measure is of interest, approaches to solution inevitably involve trade-offs. In this problem, there is a set $\{f_1(x), f_2(x), \ldots, f_m(x)\}$ of performance measures, each of which is in a form where maximization is desired with respect to the n-vector x of reals. Constraints of the form of Eq. (7) and inequality (8) may apply to the problem and define the feasible set S of allowed values of x.

Efficient solutions. A particular solution is said to dominate another solution if it achieves a higher value for one or more performance measures and a lesser value for none. If a given feasible point x^* is not dominated by any other feasible point, x^* is said to be an efficient solution or a nondominated solution (also called a Pareto optimal solution, after V. Pareto, who developed the concept in 1906). An efficient solution, therefore, is a feasible solution from which an increase in the value of any single performance measure is obtainable only if there is a decrease in value of another performance measure. The set of all nondominated feasible solutions forms the efficient solution set. It is desirable to characterize the efficient solution set in order to apply trade-offs to select one or more members of the set as being most desirable. For linear problems, techniques are available for finding the efficient solution set. For large problems, however, the efficient solutions often consist of a continuum of values that are too numerous to be displayed effectively. Members of the efficient solution set can be obtained by multiparametric programming (discussed below).

Approaches to solution. The approach to solution for a given multiobjective optimization problem depends on the underlying model structure and on the preferences of the problem solver. A given approach may combine elements of the following: (1) treating some of the performance measures as soft equality constraints and using the sensitivity results of expression (10) to determine trade-offs in performance; (2) finding efficient solutions (Pareto optimal solutions); (3) using weighting factors and parametric programming; (4) assigning a goal for each performance measure and striving to minimize deviations from the goals; and (5) assigning rankings or priority levels, a prioritized vector of objective functions, and looking for a lexicographical optimum.

Performance measure weighting. One approach to multiobjective maximization is to assign a positive weighting factor w_i to each performance measure and to maximize the sum of the $w_i f_i$'s. Thus, $f(x,w)$, the single weighted measure to be maximized, is given by Eq. (11). If desired, the w_i's can be normalized

$$f(x, w) = \sum_{i=1}^{m} w_i f_i(x) \qquad w_i\text{'s} \geq 0 \qquad (11)$$

by requiring that they sum to unity. If just one set of w_i's is selected, it is difficult to tell in advance if they will result in a solution deemed as "good." An interactive approach to the problem often is used whereby the solution corresponding to one set of weights guides the selection of alternative weights for alternative solutions.

Multiparametric programming. Solutions corresponding to many different choices of the weighting factors may be of interest. The set of optimal solutions that correspond to all allowed combinations of w_i's may be denoted V. Thus, feasible point x^* is a member of V if and only if there exists some w^* vector of w_i's ≥ 0 for which $f(x^*,w^*) \geq f(x,w^*)$ for all x contained in the feasible set S. The process of obtaining this set V is called multiparametric programming. If the f_i's are concave and the feasible region S is convex, then V and the efficient solution set can be shown to be

identical. Also, the set of w's can be divided into subsets, each of which is associated with a nondominated solution (or with a subset of nondominated solutions).

Sensitivity and weighting factors. If f is the weighted performance measure of Eq. (11), and if $f^* = f(x^*, w^*)$ is a constrained local maximum of f with respect to x, for some fixed w^*, then incremental changes in the weighting factors can be made, giving $w_i = w_i^* + \Delta w_i$, $i = 1, 2, \ldots, m$. An additional term that should then be added to the right-hand side of expression (10) is given by expression (12), where

$$\sum_{i=1}^{m} \left[f_i^* + (\nabla f_i^*)^T \Delta x \right] \Delta w_i \tag{12}$$

$(\nabla f_i^*)^{\mathrm{T}}$ is the transpose of the gradient of f_i with respect to x evaluated at x^*, and Δx is the change in the optimal solution point. As with the previous result (10), expression (12) should not be used if the incremental changes are large enough to cause a change in constraint basis. If the $\Delta x_j \Delta w_i$ terms in expression (12) are sufficiently small in magnitude, the change in the weighted performance measure can be estimated prior to actually evaluating the new optimum point $x^* + \Delta x$.

Goal programming. The point of view taken in goal programming, introduced in 1961 by A. Charnes and W. W. Cooper, is that a decision maker, rather than trying to maximize the $f_i(x)$'s, should assign aspiration levels for the performance measures, thereby creating goals. Similarly, aspiration levels may be assigned for soft constraints. Interest is centered on cases where no solution exists that satisfies all constraints and achieves all goals; the objective then is to find solutions that come as close as possible to satisfying the soft constraints and to achieving the goals, while satisfying all rigid constraints. Different approaches to goal programming use different measures of how close the goals and the soft constraints are satisfied. Different goal programming methods may also include ways to account for prioritized goals.

In forming models for use with goal programming, information regarding the goals must be incorporated. Examples of such information include the priority structure of the goals, establishing which goals are on the same priority level and leading to ranked sets of goals; the minimum performance levels that would be acceptable for the performance measures; the penalties for not achieving goals (as a function of deviations from the goals); and the rates or the values to be assigned to overachievement and to underachievement. Some goals may have desired levels, above which no additional achievement is desired. The minimum acceptable levels for performance measures can be incorporated with rigid inequality constraints.

Lexicographic screening. In some applications, it is appropriate to assign goals and constraints to preemptive priority levels: goals and constraints at higher preemptive priority levels will always take precedence over goals and constraints at lower priority levels, regardless of any scalar weights assigned to the lower-level goals. Rigid constraints can be viewed as being of top priority. Solution candidates for lower-priority levels are restricted to solutions that satisfy rigid constraints and that achieve higher-priority goals within specified tolerances. This approach to optimization is called lexicographical screening or preemptive goal programming.

Dynamic Optimization

Many dynamic systems are characterized by differential equations or by difference equations. A typical example is a dynamic system that is modeled by a set of coupled first-order differential equations: the independent variable in the equations is time; the variables associated with the derivative terms are called state variables; and other variables in the equations may include control variables to be selected in an optimal manner over time. A typical performance measure to be minimized for such a system is the integral with respect to time of weighted values of errors squared and of control actions squared. One approach to the optimization of these systems is to use discrete approximations to produce algebraic models, in which case the methods of optimization described in previous sections apply. However, many useful concepts and solution approaches that are specific to the optimization of dynamic systems have been developed.

Calculus of variations. The classical calculus of variations is concerned with the optimization of performance measures that are functions of functions. For example, a performance measure that is the integral over time of the square of an error maps a function of time (the error squared) into a scalar value. Such performance measures are called functionals. The conditions for optimality associated with functionals involve differential equations. Conditions for optimality and solution methods for these problems have been developed. One of the modern applications of the calculus of variations deals with the optimal control of systems. *See* CALCULUS OF VARIATIONS; OPTIMAL CONTROL (LINEAR SYSTEMS).

In the mid 1950s, work by L. S. Pontryagin and others extended the classical calculus of variations to apply to cases where some of the variables to be selected could be restricted to closed and bounded sets and could exhibit step discontinuities as a function of time. This work has had major applications in the design of optimal nonlinear feedback controllers. *See* NONLINEAR CONTROL THEORY.

Dynamic programming. R. E. Bellman developed the framework for dynamic programming in the 1950s. The cornerstone of dynamic programming is the following principle of optimality: an optimal policy is one having the property that, regardless of the initial state of a system and the initial decision (control) applied, the remaining decisions must constitute an optimal policy with respect to the state resulting from the initial decision. This principle is especially helpful in solving problems that can be separated (separable programming) or decoupled in some fashion. Certain complicated problems can be embedded in a sequence of less complicated but

coupled problems. The solution of the original problem is obtained by solving a sequence of problems. *See* OPTIMAL CONTROL THEORY.

Other Optimization Topics

Optimization includes a variety of other topics, such as stochastic optimization, multiple decision makers, combinatorial optimization, computer codes, and suboptimal solutions.

Stochastic optimization. Many systems contain random events and are modeled with random or stochastic variables. When such systems are optimized, performance measures and constraints may have to be defined in terms of expected values, variances, or other random variable measures. When the systems are modeled by algebraic stochastic relationships, stochastic programming may be useful for obtaining optimal solutions. For dynamic systems, stochastic optimal control and estimation algorithms may be appropriate. *See* ESTIMATION THEORY; STOCHASTIC CONTROL THEORY.

Multiple decision makers. In many applications, system performance is affected by more than one decision maker. The decision makers may have different performance objectives for the system. If goals conflict, or if the direct actions of any one decision maker are not known ahead of time by other decision makers, the resulting optimization problem is classified under the general heading of game theory. There are various classes of game theory problems, ranging from completely antagonistic games to cooperative games. In the latter case, the decision makers have the same goals for the system, but they may not have complete knowledge of the decisions being made by others.

A special case of game theory is the case where there are two decision makers, with one of them being "nature," and it is assumed that nature does everything in its power to make things worse for the ordinary decision maker; in such a case, a maximum strategy of "maximizing the worst case" may be appropriate. This also is called worst-case design. *See* DECISION THEORY; GAME THEORY.

Combinatorial optimization. Many optimization problems are associated with variables that are discrete in nature. For example, in some systems the absence of a certain entity would be represented by 0 and the presence of the entity by 1. Optimization involving such "no/yes" variables is called zero-one programming. Other system models may require certain variables to be integer values (integer programming) or to be selected from a specified set of discrete values. Discrete optimization problems include shortest-path problems that have applications in designing large-scale integrated electronic circuits and in scheduling routes for traveling sales persons. All such problems come under the general classification of combinatorial optimization. In general, solution methods for combinatorial optimization algorithms differ significantly from those discussed above.

Computer codes for solving optimization problems. Computer codes for solving optimization problems can be obtained from many sources, including computer and software vendors, relevant journals and books, and government agencies. Independent evaluations of nonlinear programming codes are available. Relevant available material includes summaries of multiobjective linear programming packages, summaries of goal-programming packages, and codes for unconstrained search methods, the linear programming simplex method, and a combinatorial minimization method.

The continuing development of improved computer systems has a direct bearing on the development of numerical algorithms for optimization. Algorithms that incorporate parallel structures can take advantage of computer systems that have multiple processors executing in parallel. From the user's point of view, optimization codes should be user-friendly, should allow interactive participation, and should employ expert system structures to help the user both in obtaining and in interpreting results. Interactive solution approaches enable decision makers to play "what if" games and to determine sensitivity attributes of candidate solutions. *See* ALGORITHM; CONCURRENT PROCESSING; EXPERT SYSTEMS.

Suboptimal solutions. When dealing with an extremely large or complex system, global optimization may not be practical. In such a case, suboptimal solution may be achieved by reducing the system into subsystems, and by obtaining optimal results for each subsystem with some of the coupling terms ignored. Suboptimal solutions also may be acceptable, or even desirable, in cases where they provide answers with little complexity or where the sensitivity of the solution to uncontrolled parameter changes is reduced.

Donald A. Pierre

Bibliography. E. M. Beale, *Introduction to Optimization,* 1988; J. D. Dennis, Jr., and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, 1983; A. V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, 1983; S. French et al. (eds.), *Multi-Objective Decision Making*, 1983; P. E. Gill, W. Murry, and M. H. Wright, *Practical Optimization*, 1981; J. P. Ignizio, *Linear Programming in Single- and Multiple-Objective Systems*, 1982; M. O'Heigeartaigh, J. D. Lenstra, and A. H. G. Rinnooy Kan (eds.), *Combinatorial Optimization*: *Annotated Bibliographies*, 1985; D. A. Pierre, *Optimization Theory with Applications*, 1986; W. H. Press et al., *Numerical Recipes*: *The Art of Scientific Computing*, 2d ed., 1993.

Oral glands

Glands located in the mouth that secrete fluids that moisten and lubricate the mouth and food and may initiate digestive activity; some perform other specialized functions. Fishes and aquatic amphibians have only solitary mucus (slime) secreting cells, distributed in the epithelium of the mouth cavity. Multicellular glands first appeared in land animals to keep the mouth moist and make food easier to swallow.

These glands occur in definite regions and bear distinctive names. Some glands of terrestrial amphibians have a lubricative secretion; others serve to make the tongue sticky for use in catching insects. Some frogs secrete a watery serous fluid that contains ptyalin, a digestive enzyme. The oral glands of reptiles are much the same, but are more distinctly grouped. In poisonous snakes and the single poisonous lizard, the Gila monster, certain oral glands of the serous type are modified to produce venom. Also many of the lizards have glands that are mixed in character, containing both mucous and serous cells. Oral glands are poorly developed in crocodilians and sea turtles. Birds bolt their food, yet grain-eaters have numerous glands, some of which secrete ptyalin.

Mammals. All mammals, except aquatic forms, are well supplied with oral glands. There are numerous small glands, such as the labial glands of the lips, boccal glands of the cheeks, lingual glands of the tongue, and palatine glands of the palate. Besides these, there are larger paired sets in mammals that are quite constant from species to species and are commonly designated as salivary glands (see **illus.**). The parotid gland, near each ear, discharges in front of the upper jaw. The submandibular (formerly submaxillary) gland lies along the posterior part of the lower jaw; its duct opens well forward under the tongue. The sublingual gland lies in the floor of the mouth, near the midplane. It is really a group of glands, each with its duct. Although not present in humans, the retrolingual gland, situated near the submandibular, is found in many mammals; its duct takes a course similar to that of the submandibular. Other occasional types are the molar gland of the hoofed mammals and the orbital gland of the dog family.

Development. All oral glands develop from the epithelial lining as branching buds. Each gland is organized somewhat after the pattern of a bush that bears berries on the ends of its twigs. The main stem and all branches of the bush correspond to a system of branching glandular ducts of various sizes; the terminal berries correspond to the secretory end pieces. Actually these end pieces are more or less elongate, like the catkins of the willow or birch. The ducts are simple epithelial tubes. The end pieces specialize in different ways. Some elaborate a serous secretion that typically contains an enzyme; others secrete a mucous fluid; still others contain both types of secretory cells. In humans and most other mammals the parotid gland produces a purely serous secretion. The submandibular and sublingual glands of most mammals are mixed (seromucous). The secretion of the sublingual gland tends to be more highly mucous in composition than that of the submandibular.

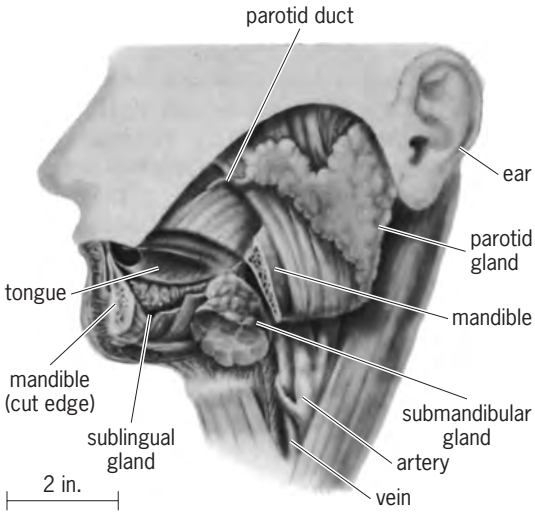

Salivary glands, shown by a partial dissection of the human head. (*After J. C. Brash, ed., Cunningham's Textbook of Anatomy, 9th ed., Oxford, 1951*)

Secretions. Saliva is a viscid fluid containing a mixture of all of the oral secretions. It contains mucus, proteins, salts, and the enzymes ptyalin and maltase. Most of the ptyalin in human saliva is furnished by the parotid gland. The digestive action of saliva is limited to starchy food. Other uses of saliva include the moistening of the mouth and food for easier manipulation by the tongue, the consequent facilitation of swallowing, and a lubrication by mucus that ensures a smoother passage of food down the esophagus to the stomach. The daily amount of saliva produced by an adult is about 1.5 quarts (1.4 liters), by the cow, 65 quarts (61.5 liters). *See* GLAND. Leslie B. Arey

Orange

The sweet orange (*Citrus sinensis*) is the most widely used species of citrus fruit and commercially is the most important. The sour or bitter oranges, of lesser importance, are distinct from sweet oranges and are classified as a separate species, *C. aurantium*. Citrus belonging to other species are also sometimes called oranges. This article deals only with the sweet orange. *See* SAPINDALES.

The sweet orange probably is native to northeastern India and southern China, but has spread to other tropical and subtropical regions of the world. The introduction of this delicious fruit into the Mediterranean region in the fifteenth century aroused much interest in citrus. The orange was first introduced into the Western Hemisphere by Columbus, and has become a major crop in Florida and California and in Brazil. The United States is the largest producer of oranges, followed by Spain, Italy, and Brazil. It is also a major crop in several other countries.

Uses. Sweet orange fruit is consumed fresh or as frozen or canned juice. A large portion of the crop, particularly in the United States, is used as frozen concentrate. After the juice is extracted, the peel and pulp are used for cattle feed. Peel oil is used in perfumes and flavoring, and citrus molasses is used as a livestock feed.

Varieties. The sweet orange tree is a moderately vigorous evergreen with a rounded, densely foliated top. The fruits are round (**Fig. 1**) or somewhat elongate and usually orange-colored when ripe. They can be placed in four groups: the common oranges, acidless oranges, pigmented oranges, and navel oranges. They may also be distinguished on the basis

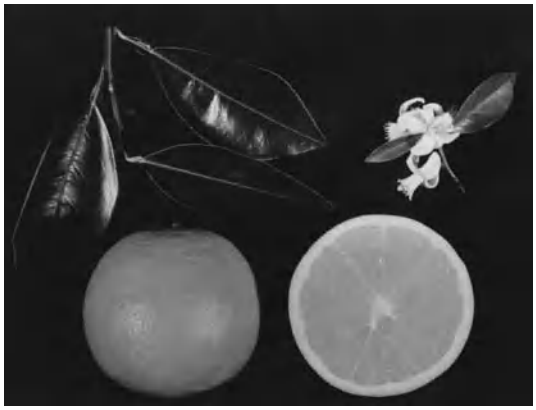

Fig. 1. Leaves, flowers, whole fruit, and cut fruit of *Citrus sinensis*.

of early, midseason, and late maturity. The common oranges are orange-colored and have varying numbers of seeds. Several, such as the Valencia, are important commercial cultivars throughout the world. The acidless oranges are prized in some countries because of their very low acidity, but generally are not major cultivars. The pigmented or "blood" oranges develop varying amounts of pink or red in the flesh and peel and have a distinctive aroma and flavor. Most originated in the Mediterranean area and are popular there, but are little grown in the United States. Navel oranges are so named because of the structure at the blossom end. Because they are usually seedless, some have become important cultivars for the fresh-fruit market. Many, if not all, of the cultivars appear to have been selected as bud sports, with the original type being small, seedy, and nonpigmented. *See* EVERGREEN PLANTS.

Propagation. Sweet oranges are propagated by budding or grafting on rootstocks in order to obtain uniform, high-yielding plantings. Many citrus species and relatives have been used as rootstocks. Diseases have severely restricted the use of some major rootstocks. Tristeza, a virus disease, has caused extensive worldwide losses of trees on sour orange rootstock. Root rot caused by soil-inhabiting fungi is also widespread, attacking many types of rootstocks. Large-scale selection and breeding programs have developed rootstocks that are tolerant to tristeza and root rot. *See* FRUIT; FRUIT, TREE. R. K. Soost

Diseases. Sweet orange trees are affected by several bud-transmitted virus diseases. Psorosis, which causes scaling of the bark, is injurious to the tree irrespective of the rootstock used. Some other virus diseases are injurious only when infected trees are grown on certain rootstocks. For example, exocortis virus stunts trees growing on trifoliate orange rootstock, whereas xyloporosis stunts trees on sweet lime rootstock. The aphid- and bud-transmitted tristeza virus causes a serious decline of trees grown on sour orange rootstock, and alternative rootstocks are essential for sweet orange production in countries where the efficient aphid vector of tristeza, *Toxoptera citricidus*, occurs.

Greening disease causes a devastating tree decline in parts of Asia and South Africa, regardless of the rootstock used. It behaves like a virus disease, but is actually caused by an unnamed bacteriumlike organism. Greening disease is a problem only in those areas where the psyllid insect vectors are abundant. In South Africa, greening disease is controlled by injecting the tree trunks with antibiotics to kill the causal agent and by spraying the orchard with insecticides to control psyllids. Stubborn disease is another viruslike disease that seriously reduces fruit yields. It occurs in California and some Mediterranean countries. The mycoplasmalike organism, *Spiroplasma citri*, which causes stubborn disease is spread by infected budwood and certain leafhopper insects.

Blight is a serious wilt disease that rapidly renders orange trees unproductive. It behaves like an infectious disease; yet all attempts to transmit the disease by budding and grafting have failed. The cause of this disease is presently unknown. In Florida, blight is common in orchards planted on rough lemon rootstock, but is rarely seen on trees growing on sour orange.

Citrus canker, caused by *Xanthomonas citri*, is a bacterial disease that originated in Asia. It reached many other parts of the world through shipment of infected nursery trees, but it was later successfully eradicated in some areas, notably in the United States (Florida), Australia, and South Africa, through the costly destruction of all diseased trees.

Foot rot (trunk gummosis), caused by certain soil-borne *Phytophthora* fungi, occurs almost everywhere that orange trees are grown. Attacks commonly occur at the base of the trunk (**Fig. 2**). The bark and cambium are killed, thereby causing varying amounts of trunk girdling. Prevention depends on the use of resistant rootstocks, high budding to keep the highly susceptible sweet orange bark well above the ground, and good land drainage. These

Fig. 2. Foot rot on sweet orange portion of trunk, with root sprouts emerging from the unaffected disease-resistant rootstock.

Fig. 3. Greasy spot on sweet orange leaf.

same fungi cause brown rot of fruit during periods of prolonged wet weather. Fruit infection is prevented by spraying the tree canopy with copper fungicides.

Other important fungus diseases of sweet orange include greasy spot (caused by *Mycosphaerella citri*; **Fig. 3**), which causes serious defoliation in countries with very hot humid climates; black spot (caused by *Guignardia citricarpa*), which is an important disease in parts of South Africa and Australia; melanose (caused by *Diaporthe citri*), which is prevalent in areas with wet and relatively warm weather during early fruit development; and scab (caused by *Elsinoe australis*), which has been found only in South America. Spray oil or copper fungicides are applied to control greasy spot; benomyl fungicide is used to control black spot; and copper fungicides provide control of melanose and scab. *See* PLANT PATHOLOGY.

Jack O. Whiteside

Bibliography. A. R. Biggs (ed.), *Handbook of Cytology, Histology, and Histochemistry of Fruit Tree Diseases*, 1992; M. L. Flint, *Integrated Pest Management for Citrus*, 2d ed., 1991; J. Janick and J. N. Moore (eds.), *Advances in Fruit Breeding*, 1975.

Orbital motion

In astronomy the motion of a material body through space under the influence of its own inertia, a central force, and other forces. Johannes Kepler found empirically that the orbital motions of the planets about the Sun are ellipses. Isaac Newton, starting from his laws of motion, proved that an inverse-square gravitational field of force requires a body to move in an orbit that is a circle, ellipse, parabola, or hyperbola.

Elliptical orbit. Two bodies revolving under their mutual gravitational attraction, but otherwise undisturbed, describe orbits of the same shape about a common center of mass. The less massive body has the larger orbit. In the solar system, the Sun and Jupiter have a center of mass just outside the visible disk of the Sun. For each of the other planets, the center of mass of Sun and planet lies within the Sun.

For this reason, it is convenient to consider only the relative motion of a planet of mass m about the Sun of mass M as though the planet had no mass and moved about a center of mass $M + m$. The orbit so determined is exactly the same shape as the true orbits of planet and Sun about their common center of mass, but it is enlarged in the ratio $(M + m)/M$. *See* CENTER OF MASS; PLANET.

Parameters of elliptical orbit. The **illustration** shows the elements or parameters of an elliptic orbit. Major axis AP intersects the ellipse AOP at the apsides; the extension of the major axis is the line of apsides. The body is nearest the center of mass at one apside, called perihelion P, and is farthest away at the other, called aphelion A. *See* ELLIPSE.

Shape and size of an orbit are defined by two elements: length of semimajor axis and departure of the orbit from a circle. Semimajor axis a equals CP; this length is expressed in units of the mean distance from the Earth to the Sun. Eccentricity e equals CS/CP where C is the center of the ellipse and S is a focus. For elliptical orbits e is always less than unity.

Position of a body in its orbit at time t can be computed if a, e, and time of perihelion passage p and period of revolution T are known. Let O be the position of a planet at time t and OSP be the area swept out in time $t - p$. From Kepler's area law, area OSP equals $(t - p)/T$ multiplied by the area of the full ellipse.

To describe the orientation of an orbit in space, several other parameters are required. All orbits in the solar system are referred to the plane of the ecliptic, this being the plane of the orbit of Earth about the Sun. The reference point for measurement of celestial longitude in the plane of the ecliptic is the vernal equinox , the first point of Aries. This is the

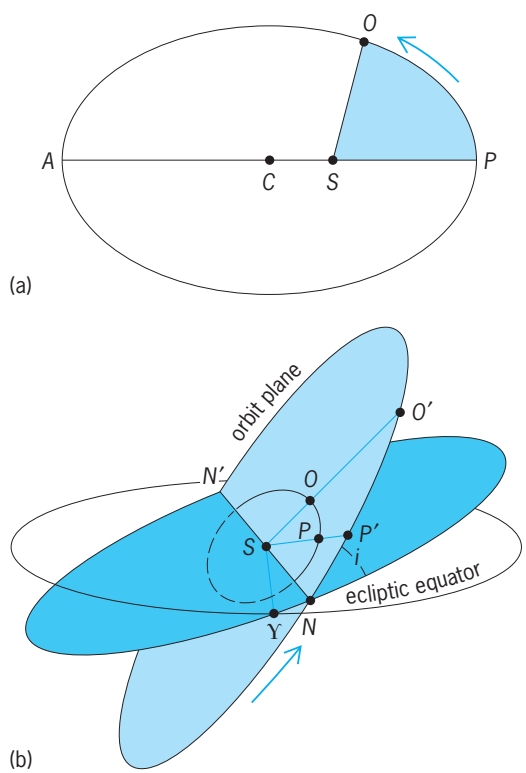

Parameters, or elements, of an elliptical orbit. (*a*) Relative orbit. (*b*) Orbit in space.

point where the apparent path of the Sun crosses the Earth's equator from south to north. The two points of intersection of the orbit plane with the plane of the ecliptic (N and N') are called the nodes, and the line joining them is the line of nodes. Ascending node N is the one where the planet crosses the plane of the ecliptic in a northward direction; N' is the descending node. The angle as seen from the Sun S measured in the plane of the ecliptic from the vernal equinox to the ascending node is ΥSN; it is termed the longitude of the ascending node Ω and fixes the orbit plane with respect to the zero point of longitude. The angle at the ascending node between the plane of the ecliptic and the orbit plane is called the inclination i and defines the orientation of the orbit plane with respect to the fundamental plane. The angle as seen from the Sun, measured in the orbit plane from the ascending node to perihelion, is NSP' and is referred to as the argument of perihelion; it defines the orientation of the ellipse within the orbit plane. The angle $NSP' + \Omega$, measured in two different planes, is called the longitude of perihelion $\tilde{\omega}$. Because dynamically the semimajor axis a and period T of a planet of mass m revolving under the influence of gravitation G about the Sun of mass M are related by Eq. (1),

$$\frac{4\pi^2}{T^2} = \frac{G(M+m)}{a^3} \tag{1}$$

only six elements, $a, e, i, \Omega, \tilde{\omega}$, and p, are required to fix the orbit and instantaneous position of a planet in space. Instead of these elements, however, a position vector x, y, z and the associated velocity vector $\dot{x}, \dot{y}, \dot{z}$ at a given instant of time would serve equally well to define the path of a planet in a rectangular coordinate system with origin at the Sun.

Orbital velocity. Orbital velocity v of a planet moving in a relative orbit about the Sun may be expressed by Eq. (2) where a is the semimajor axis and r is the

$$v^2 = G(M+m)\left(\frac{2}{r} - \frac{1}{a}\right) \tag{2}$$

distance from the planet to the Sun. In the special case of a circular orbit, $r = a$, and the expression becomes Eq. (3). When the eccentricity of an orbit

$$v^2 = \frac{G(M+m)}{a} \tag{3}$$

is exactly unity, the length of the major axis becomes infinite and the ellipse degenerates into a parabola. The expression for the velocity then becomes Eq. (4).

$$v^2 = G(M+m)\left(\frac{2}{r}\right) \tag{4}$$

This parabolic velocity is referred to as the velocity of escape, since it is the minimum velocity required for a particle to escape from the gravitational attraction of its parent body. *See* ESCAPE VELOCITY.

Eccentricities greater than unity occur with hyperbolic orbits. Because in a hyperbola the semimajor axis a is negative, hyperbolic velocities are greater than the escape velocity.

Parabolic and hyperbolic velocities seem to be observed in the motions of some comets and meteors. Aside from the periodic ones, most comets appear to be visitors from cosmic distances, as do about two-thirds of the fainter meteors. For ease of computation, the short arcs of these orbits that are observed near perihelion are represented by parabolas rather than ellipses. Although the observed deviation from parabolic motion is not sufficient to vitiate this computational procedure, it is possible that many of these "parabolic" comets are actually moving in elliptical orbits of extremely long period. The close approach of one of these visitors to a massive planet, such as Jupiter, could change the velocity from parabolic to elliptical if retarded, or from parabolic to hyperbolic if accelerated. It is possible that many of the periodic comets, especially those with periods under 9 years, have been captured in this way. *See* CELESTIAL MECHANICS; COMET; GRAVITATION; PERTURBATION (ASTRONOMY); STELLAR ROTATION. Raynor L. Duncombe

Bibliography. J. E. Prussing and B. A. Conway, *Orbital Mechanics*, Oxford University Press, 1993; A. E. Roy, *Orbital Motion,* 4th ed., Institute of Physics Publishing, 2004; V. G. Szebehely and H. Mark, *Adventures in Celestral Mechanics: A First Course in the Theory of Orbits,* 2d ed., Wiley, 1998; B. D. Tapley, B. E. Shultz, and G. H. Born, *Statistical Orbit Determination*, Elsevier Academic Press, 2004.

Orchid

Any member of the orchid family (Orchidaceae), among the largest families of plants, estimated to contain up to 35,000 species. Orchids are monocots; their flowers have inferior ovaries, three sepals, and three petals. They are distinguished by the differentiation of one petal into a labellum, and the fusion of pistil and stamens into the column (see **illus.**). Pollen is usually contained in pollinia, that is, bundles that are removed intact by pollinators, usually insects or sometimes birds. Self-pollination and asexual reproduction without fertilization also occur. The

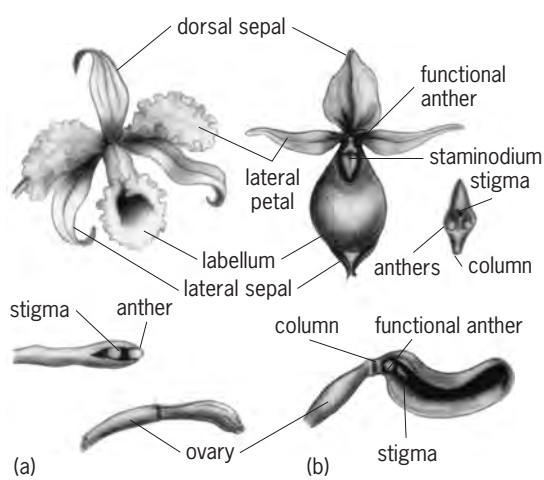

Structure of representative orchid flowers. (*a*) *Cattleya*. (*b*) *Cypripedium*.

combination of lip and column structure, flower color, fragrance, and other factors may limit the range of pollinators. Differing pollination mechanisms often provide barriers to cross-pollination between related species. Each flower can produce large quantities of seeds, with numbers in the millions in some tropical species. Seeds are minute, with undifferentiated embryo and no endosperm. Germination and establishment depends on symbiotic mycorrhizae that provide nutrients and water.

Orchids occur on all continents except Antarctica; they range from arctic tundra and temperate forest and grassland to tropical rainforest, where as epiphytes they reach their greatest abundance and diversity. Vanilla is obtained from seed pods of some species of *Vanilla*, and the beauty and mystique of many orchids make them important horticultural subjects. *See* MYCORRHIZAE; ORCHIDALES; VANILLA.

Charles J. Sheviak

Bibliography. R. L. Dressler, *Phylogeny and Classification of the Orchid Family*, 1993.

Orchidales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Liliidae of the class Liliopsida (monocotyledons). The order consists of four families: the Orchidaceae, with perhaps 15,000–20,000 species; the Burmanniaceae, with about 130 species; the Cordiaceae, with 9 species; and the Geosiridaceae, with 1 species. The Orchidales are mycotrophic, sometimes nongreen Liliidae with very numerous tiny seeds that have an undifferentiated embryo and little or no endosperm. The ovary is always inferior and apparently lacks the septal nectaries found in many Liliopsida although other kinds of nectaries are present. The various species of Orchidaceae (orchids) have numerous, highly diverse adaptations to particular insect pollinators. The flowers are highly irregular (see **illus.**) and usually have only one or two stamens, which are adnate to the thickened style. The pollen grains of most genera cohere in large masses called pollinia, so that a single act of pollination can result in the production of many thousands of seeds. Orchids are most abundant and diversified in moist, tropical regions, where they often grow high above the ground on forest trees. They do not parasitize the trees, but merely roost on them, as epiphytes. The most familiar corsage orchids belong to the genus *Cattleya* or hybrids of it. *See* FLOWER; LILIIDAE; LILIOPSIDA; MAGNOLIOPHYTA; ORCHID; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Eastern American species of moccasin flower (*Cypripedium acaule*). (*U.S. Forest Service photograph by R. Dale Sanders*)

Ordovician

The second-oldest period in the Paleozoic Era. The Ordovician is remarkable because not only did one of the most significant Phanerozoic radiations of marine life take place (early Middle Ordovician), but also one of the two or three most severe extinctions of marine life occurred (Late Ordovician). The early Middle Ordovician radiation of life included the initial colonization of land. These first terrestrial organisms were nonvascular plants. Vascular plants appeared in terrestrial settings shortly afterward. *See* GEOLOGIC TIME SCALE.

The rocks deposited during this time interval (these are termed the Ordovician System) overlie those of the Cambrian and underlie those of the Silurian. The Ordovician Period was about 7×10^7 years in duration, and it lasted from about 5.05×10^8 to about 4.35×10^8 years ago.

The British geologist Charles Lapworth named the Ordovician in 1879, essentially as a resolution to a long-standing argument among British geologists over division of the Lower Paleozoic. Until that time, one school of thought, that of R. I. Murchison and his followers, had maintained that only a Silurian Period encompassed the lower part of the Paleozoic. Adam Sedgwick and his followers advocated that two intervals, the Cambrian and the Silurian, could be recognized in the Lower Paleozoic. By 1879 Lapworth observed that "three distinct faunas" had been recorded from the Lower Paleozoic, and he pointed out that each was as "marked in their characteristic features as any of those typical of the accepted systems of later age."

To the stratigraphically lowest and oldest of the three, Lapworth suggested in 1879 that the appellation Cambrian be restricted. To the highest and youngest, Lapworth stated that the name Silurian should be applied. To the middle or second of three, Lapworth gave the name Ordovician, taking the

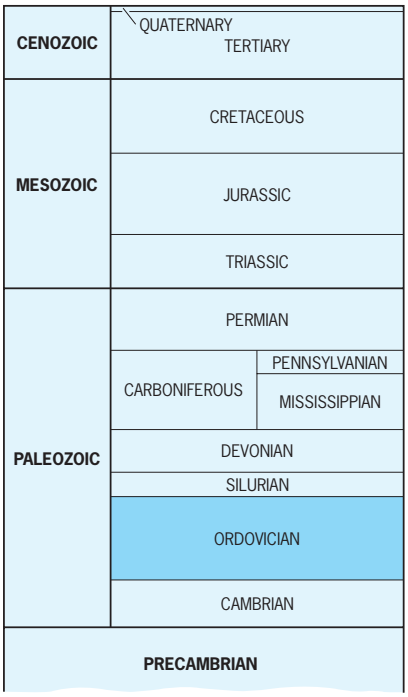

name from an ancient tribe renowned for its resistance to Roman domination.

The type area for the Ordovician System is those parts of Wales and England that include rocks bearing the fossils that composed the second of the three major Lower Paleozoic faunas cited by Lapworth. The Ordovician System in Britain was divided into six major units called series, each distinguished by a unique fossil fauna. The time intervals during which each series formed are epochs. From oldest to youngest, the epochs and series of the British Ordovician are Tremadoc, Arenig, Llanvirn, Llandeilo, Caradoc, and Ashgill. Each of them has a type area in Britain where their characteristic faunas may be collected.

Intervals of shorter duration than those of the epoch are recognized as well in Britain. One set of such intervals is based on the evolutionary development of the fossils called graptolites. These intervals are the graptolite zones recognized by Lapworth and his associates Ethel M. R. Wood and Gertrude Elles. Each graptolite zone was about 3×10^6 to 5×10^6 years in duration. The boundary between the Ordovician and superjacent Silurian System has been designated as the base of the *Parakidograptus acuminatus* graptolite zone by international agreement. The type locality for that boundary is at Dob's Linn, near Moffat, southern Scotland. Black, graptolite-bearing shales are exposed there.

The Ordovician System is recognized in nearly all parts of the world, including the peak of Mount Everest, because the groups of fossils used to characterize the system are so broadly delineated. The British epochs and zones may not be recognized in all areas where Ordovician rocks are found because the fossils used to characterize them are limited to certain geographic areas. Biogeographic provinces limited the distribution of organisms in the past to patterns similar to those of modern biogeographic provinces. Three broadly defined areas of latitude—the tropics, the midlatitudes (approximately 30–60°S), and the Southern Hemisphere high latitudes—constitute the biogeographic regions. Provinces may be distinguished within these three regions based upon organismal associations unique to each province. Epochs and zones are limited to a single region, and consequently each region has a unique set of epochs and zones for the Ordovician.

Dynamic interrelationships. The Earth's crust is essentially a dynamic system that is ceaselessly in motion. Plate positions and plate motions are linked closely with, and potentially exert a primary driving force that underlies, ocean circulation, ocean-atmosphere interactions, climates and climate change, and expansion and reduction of environments. Life responds to these physical aspects of the Earth's crust.

Several lines of evidence, including remanent magnetism, distributions of reefs and other major accumulations of carbonate rocks, positions of shorelines, and sites of glacial deposits, may be used to deduce many aspects of Ordovician paleogeography and paleogeographic changes. The Ordovician configuration of land and sea was markedly different from today. Much of the Northern Hemisphere above the tropics was ocean. The giant plate Gondwana was the predominant feature of the Southern Hemisphere. Modern Africa and South America were joined and occupied most the Southern Hemisphere high latitudes. The South Pole lay approximately in north-central Africa. A massive lobe of the plate extended northward from eastern Africa into the tropics. The modern Middle East, Turkey, India, Antarctica, and Australia constituted much of that huge lobe. A number of small plates lay on the margins of Gondwana. Certain of them may have been joined to Gonwana, and others were close to it. Some of these high-latitude plates included Perunica (modern Czech Republic); Avalonia (possibly in two parts, an eastern and a western), which included parts of Wales, southern Britain, Newfoundland, and Maritime Canada; and a number of plates that today make up southern Europe, including those found in the Alps, those that constitute the Iberian Peninsula, and those that made up Armorica (much of France). Plates that were within about 30–55°S latitude included Baltica (modern Scandinavia and adjacent eastern Europe east to the Urals), the Argentine Precordillera, South China, Tarim (in Asiatic China), the Exploits, and perhaps similar small plates. Parts of the Andes within modern northern Argentina, Peru, and Bolivia were within this midlatitudinal interval. Plates in tropical latitudes included Laurentia (modern North America, Greenland, Scotland, and some of northern Ireland), North China, Siberia, one or more plates that made up the Kazakh plate, and

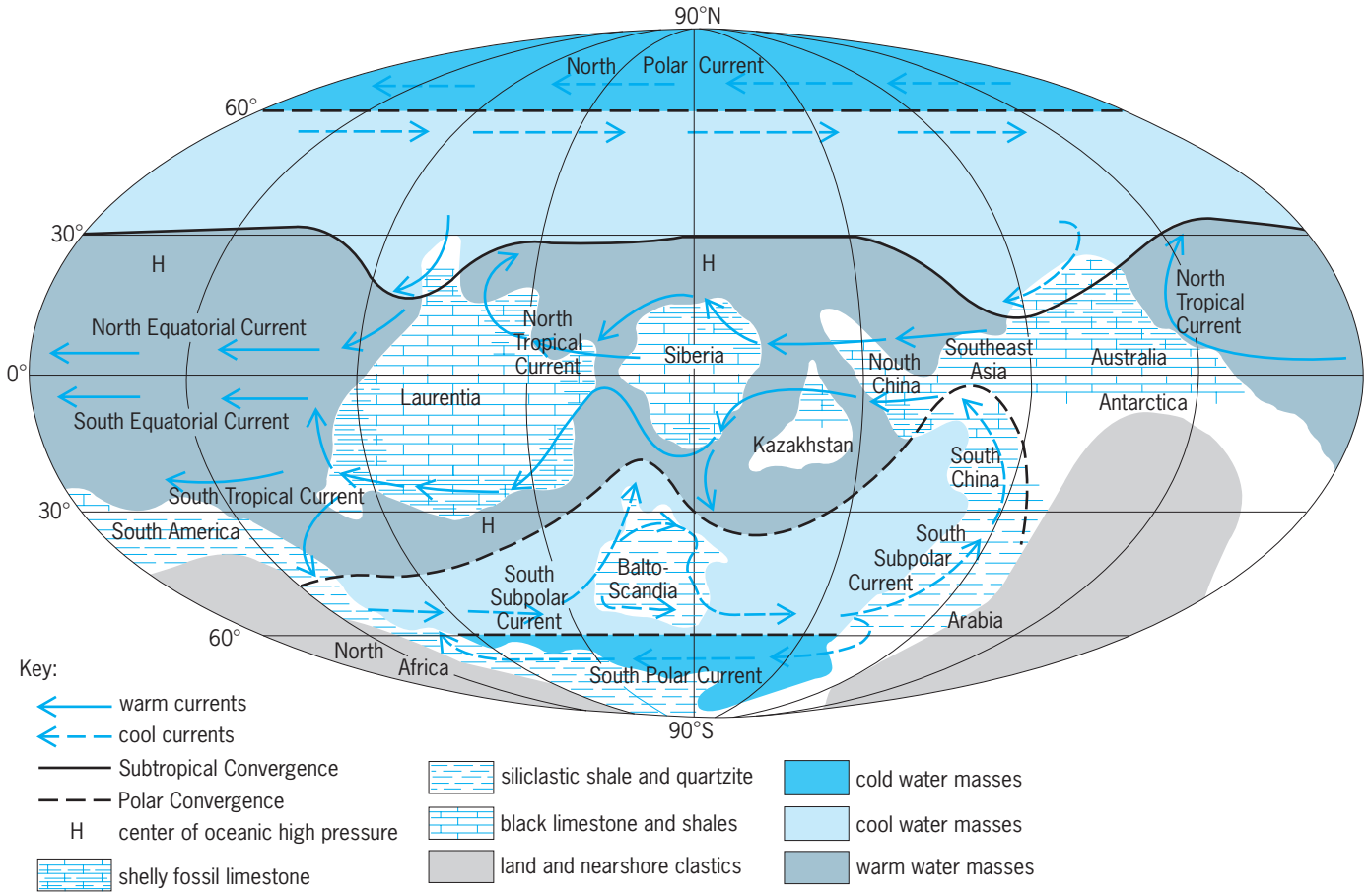

Early Ordovician (Arenig) paleogeography, major rock suites, and potential ocean surface currents; based on remanent magnetism data supplemented by plots of the positions of shelf sea carbonates and reefs.

several plates that were close to or attached to northern or tropical Gondwana (these make up modern southeastern Asia and southeastern China). *See* PALEOGEOGRAPHY; PALEOMAGNETISM; PLATE TECTONICS.

During the early part of the Ordovician (Tremadoc-Arenig), prior to a significant number of plate movements, siliclastic materials (sand, silts, muds) spread northward from a Gondwana landmass into river, delta, and nearshore marine environments on the Gondwana plate and on those plates close to it, especially those in high latitudes. Coeval tropical environments were sites of extensive carbonate accumulations. Most midlatitude plates were sites of siliclastic and cool-water carbonate deposition.

Extensive plate motions and major volcanic activity at the margins of many plates characterize the Arenig-Llanvirn boundary interval of the early Middle Ordovician. Many plates on the Gondwanan margins began a northward movement that continued for much of the remainder of the Paleozoic. In addition, Laurentia bulged upward to such an extent that marine environments, which had covered most of the plate early in the Ordovician, were driven to positions on the plate margins. The Avalonian plates joined and moved relatively quickly northward to collide with the eastern side of Laurentia near the end of the Ordovician. Prior to that collision, the Popelogan or Medial New England plate collided with the Laurentian plate at about the position of modern New England. That collision, which occurred about 455 million years ago, classically has been called the Taconic orogeny. Baltica not only moved northward relatively rapidly, but also rotated about 90° during the latter part of the Ordovician. The Argentine Precordillera plate moved southward across a midlatitudinal interval of ocean to collide with what is today the western side of Argentina in the Middle Ordovician. Africa shifted northward during the Ordovician with the result that northern Africa and the regions adjacent to the Middle East shifted into cool-temperate conditions.

Life and environments. As plate motions took place, environments changed significantly, as well as life in them. Both oceanic and terrestrial settings became the sites of significant radiations.

Early Ordovician (Tremadoc-Arenig) environmental conditions in most areas were similar to those of the Late Cambrian. Accordingly, Early Ordovician life was similar to that of the latter part of the Cambrian. Trilobites were the prominent animal in most shelf sea environments. Long straight-shelled nautiloids, certain snails, a few orthoid brachiopods, sponges, small echinoderms, algae, and bacteria flourished in tropical marine environments. Linguloid brachiopods and certain bivalved mollusks inhabited cool-water, nearshore environments.

Middle Ordovician plate motions were acompanied by significant changes in life. On land, nonvascular, mosslike plants appeared in wetland habitats. Vascular plants appeared slightly later in riverine habitats. The first nonvascular plants occurred in the Middle East on Gondwanan shores. The Middle Ordovician radiation of marine invertebrates is one of the most extensive in the record of Phanerozoic marine life. Corals, bryozoans, several types of brachiopods, a number of crinozoan echiniderms, conodonts, bivalved mollusks, new kinds of ostracodes, new types of trilobites, and new kinds of nautiloids suddenly developed in tropical marine environments. As upwelling conditions formed along the plate margins, oxygen minimum zones—habitats preferred by many graptolites—expanded at numerous new sites. Organic walled microfossils (chitinozoans and acritarchs) radiated in mid- to high-latitude environments. Ostracoderms (jawless, armored fish) radiated in tropical marine shallow-shelf environments. These fish were probably bottom detritus feeders. *See* PALEOECOLOGY.

Glaciation. When the Avalon plate collided with the Laurentian, a major mountain chain developed in a tropical setting. Vast quantities of siliclastic materials were shed from that land to form what is called the Queenston delta in the present-day Appalachians. As the Queenston delta grew, glaciation commenced at or near the South Pole. Continental glaciation spread from its North African center for a 1–2-million-year interval late in the Ordovician. Glacially derived materials (including many drop-stones) occurred in the Late Ordovician strata in Morocco, Algeria, southern France, Germany, Spain, Portugal, and the Czech Republic. Sea level dropped by at least 70 m at the glacial maximum. As a result, most shallow to modest-depth marine environments were drained. Karsts formed across many carbonates that had accumulated in the shallow marine settings. Upwelling along many platform margins ceased or became quite limited. As a consequence, former extensive oxygen minimum zones were markedly diminished. The loss of wide expanses of shallow marine environments and extensive oxygen minimum zones led to massive extinctions of benthic marine organisms, as well as those graptolites living near oxygen minimum zones. These extinctions took place over a 1–2-million-year interval as environments shifted, diminished, or eventually were lost. Oxygen isotope studies on brachiopod shells suggest that tropical sea surface temperatures dropped by as much as 4°C. *See* GEOMORPHOLOGY; PALEOCEANOGRAPHY.

The latest Ordovician stratigraphic record suggests that the ice melted relatively quickly, accompanied by a relatively rapid sea-level rise in many areas. Some organisms—certain conodonts, for example—did not endure significant extinctions until sea levels began to rise and shelf sea environments began to expand. *See* STRATIGRAPHY.

Ocean surface circulation. Surface circulation in Ordovician seas was controlled in the tropics by the several platforms and in the Southern Hemisphere by Gondwanaland. Equatorial surface currents flowed east to west, but they were deflected by the shallow shelf environments. The tropical or warm-water faunal provinces were influenced by these deflections. Homogeneity of the tropical faunas was maintained by the surface water currents. Southern Hemisphere currents were influenced by the relatively long west coast of Gondwanaland and of the Baltoscanian Plate. Upwelling conditions would have been generated along these coasts. Location, size, and relief on Gondwanaland probably led to monsoonal seasonal reversals in surface ocean currents near what is today South China. Absence of lands or shallow shelf seas north of the Northern Hemisphere tropics would permit oceanic surface circulation to be zonal; that is, currents flowed from east to west north of 30° north latitude, and they flowed from west to east between 30 and 60° north latitude.

Economic resources. Ordovician shelf and shelf margin rock sequences in areas where there has been little post-Ordovician volcanic activity or severe deformation have yielded petroleum and natural gas. Quartzites interbedded with carbonates formed in shelf sea environments have been used as a source of silica for glass manufacture. Ordovician carbonates are hosts for lead-zinc-silver ores mined in the western United States, including Missouri and Washington. Significant quantities of gold were recovered from Ordovician graptolite-bearing strata in eastern Australia in the late 1800s. Gold-bearing Ordovician rocks occur in Nevada (western United States) where they are part of one of the most prolific gold-producing areas in the world.

W. B. N. Berry

Bibliography. J. D. Cooper, M. I. Droser, and S. C. Finney (eds.), *Ordovician Odyssey: Short Papers for the 7th International Symposium on the Ordovician System*, Pacific Section, Society for Sedimentary Geology, 1995; P. Kraft and O. Fatka (eds.), Quo Vadis Ordovician?, *Acta Universitatis Carolinae Geologica*, vol. 43, no. 1/2, 1999; C. Lapworth, On the tripartite classification of the Lower Palaeozoic rocks, *Geol. Mag.*, 6:1–15, 1879; T. H. Torsvik, Palaeozoic palaeogeography: A North Atlantic viewpoint, *GFF*, 120:109–118, 1998.

Ore and mineral deposits

Ore deposits are naturally occurring geologic bodies that may be worked for one or more metals. The metals may be present as native elements, or, more commonly, as oxides, sulfides, sulfates, silicates, or other compounds. The term ore is often used loosely to include such nonmetallic minerals as fluorite and gypsum. The broader term, mineral deposits, includes, in addition to metalliferous minerals, any other useful minerals or rocks. Minerals of little or no value which occur with ore minerals are called gangue. Some gangue minerals may not be worthless in that they are used as by-products; for instance, limestone for fertilizer or flux, pyrite for making sulfuric acid, and rock for road material.

TABLE 1. Elemental composition of Earth's crust based on igneous and sedimentary rocks*

Element	Weight %	Atomic %	Volume %
Oxygen	46.71	60.5	94.24
Silicon	27.69	20.5	0.51
Titanium	0.62	0.3	0.03
Aluminum	8.07	6.2	0.44
Iron	5.05	1.9	0.37
Magnesium	2.08	1.8	0.28
Calcium	3.65	1.9	1.04
Sodium	2.75	2.5	1.21
Potassium	2.58	1.4	1.88
Hydrogen	0.14	3.0	

*After T. F. W. Barth, *Theoretical Petrology*, John Wiley and Sons, 1952 (recalculated from F. W. Clarke and H. S. Washington, 1924).

Mineral deposits that are essentially as originally formed are called primary or hypogene. The term hypogene also indicates formation by upward movement of material. Deposits that have been altered by weathering or other superficial processes are secondary or supergene deposits. Mineral deposits that formed at the same time as the enclosing rock are called syngenetic, and those that were introduced into preexisting rocks are called epigenetic.

The distinction between metallic and nonmetallic deposits is at times an arbitrary one since some substances classified as nonmetals, such as lepidolite, spodumene, beryl, and rhodochrosite, are the source of metals. The principal reasons for distinguishing nonmetallic from metallic deposits are practical ones, and include such economic factors as recovery methods and uses.

Concentration. The Earth's crust consists of igneous, sedimentary, and metamorphic rocks. **Table 1** gives the essential composition of the crust and shows that 10 elements make up more than 99% of the total. Of these, aluminum, iron, and magnesium are industrial metals. The other metals are present in small quantities, mostly in igneous rocks (**Table 2**).

Most mineral deposits are natural enrichments and concentrations of original material produced by different geologic processes. To be of commercial grade, the metals must be present in much higher concentrations than the averages shown in Table 2. For example, the following metals must be concentrated in the amounts indicated to be considered ores: aluminum, about 30%; copper, 0.7–10%; lead, 2–4%; zinc, 3–8%; and gold, silver, and uranium, only a small fraction of a percent of metal. Therefore, natural processes of concentration have increased the aluminum content of aluminum ore 3 or 4 times, and even a low-grade gold ore may represent a concentration of 20,000 times. Economic considerations, such as the amount and concentration of metal, the cost of mining and refining, and the market value of the metal, determine whether the ore is of commercial grade.

Forms of deposits. Mineral deposits occur in many forms depending upon their origin, later deformation, and changes caused by weathering. Syngenetic deposits are generally sheetlike, tabular, or lenticular, but may on occasion be irregular or roughly spherical.

Epigenetic deposits exhibit a variety of forms. Veins or lodes are tabular or sheetlike bodies that originate by filling fissures or replacing the country rock along a fissure (**Fig. 1***a*). Replacement bodies in limestone may be very irregular. Veins are usually inclined steeply and may either cut across or conform with the bedding or foliation of the enclosing rocks. The inclination is called the dip, and is the angle between the vein and the horizontal. The horizontal trend of the vein is its strike, and the vertical angle between a horizontal plane and the line of maximum elongation of the vein is the plunge. The veins of a mining district commonly occur as systems which have a general strike, and one or more systems may be present at some angle to the main series. In places the mineralization is a network of small, irregular, discontinuous veins called a stockwork.

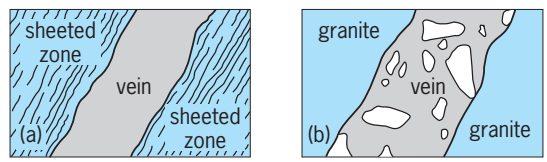

Fig. 1. Typical forms of deposits. (*a*) Vein developed in fissured or sheeted zone. (*b*) Brecciated vein in granite.

Mineral deposits are seldom equally rich throughout. The pay ore may occur in streaks, spots, bunches, or bands separated by low-grade material or by gangue. These concentrations of valuable ore are called ore shoots; if roughly horizontal they are called ore horizons, and if steeply inclined they are called chimneys. After formation, mineral deposits may be deformed by folding, faulting, or brecciation (Fig. 1*b*).

TABLE 2. Abundance of metals in igneous rocks

Element	%
Aluminum	8.13
Iron	5.00
Magnesium	2.09
Titanium	0.44
Manganese	0.10
Chromium	0.02
Vanadium	0.015
Zinc	0.011
Nickel	0.008
Copper	0.005
Tin	0.004
Cobalt	0.0023
Lead	0.0016
Arsenic	0.0005
Uranium	0.0004
Molybdenum	0.00025
Tungsten	0.00015
Antimony	0.0001
Mercury	0.00005
Silver	0.00001
Gold	0.0000005
Platinum	0.0000005

Metasomatism, or replacement. Metasomatism, or replacement, is the process of essentially simultaneous removal and deposition of chemical matter. A striking manifestation of this process in mineral deposits is the replacement of one mineral by another mineral or mineral aggregate of partly or wholly different composition. A large volume of rock may be transformed in this manner, and the resulting deposit is generally of equal volume. Commonly the original structure and texture of the replaced rock is preserved by the replacing material.

Replacement, evidence for which is found in many mineral deposits, operates at all depths under a wide range of temperature. The evidence indicates that the new minerals formed in response to conditions that were unstable for the preexisting ones.

Usually the replacing material moves to the site of metasomatism along relatively large openings such as faults, fractures, bedding planes, and shear zones. It then penetrates the rock along smaller cracks and finally enters individual mineral grains along cleavage planes, minute fractures, and grain boundaries where substitution may take place on an atomic scale until the entire mass has been transformed (**Fig. 2**). After gaining access to individual grains, the replacement may proceed by diffusion of ions through the solid, controlled in large part by imperfections in the crystal structure. In many deposits repeated movement has opened and reopened channelways, which would otherwise have become clogged, to permit continued and widespread replacement. The process may take place through the action of gases or solutions or by reactions in the solid state. *See* METASOMATISM.

Classification. Mineral deposits are generally classified on the basis of the geologic processes responsible for their formation as magmatic, contact metasomatic, pegmatitic, hydrothermal, sedimentary, residual, and regional metamorphic deposits.

Magmatic deposits. Some mineral deposits originated by cooling and crystallization of magma, and the concentrated minerals form part of the body of the igneous rock. If the magma solidified by simple crystallization, the economically valuable mineral is distributed through the resulting rock; diamond deposits found in peridotite are believed by some geologists to be of this type. However, if the magma has differentiated during crystallization, early formed minerals may settle to the bottom of the magma chamber and form segregations such as the chromite deposits of the Bushveld in South Africa. Late-formed minerals may crystallize in the interstices of older minerals and form segregations like the Bushveld platinum deposits. Occasionally, the residual magma becomes enriched in constituents such as iron, and this enriched liquid may form deposits, such as the Taberg titaniferous iron ores of Sweden. It is also possible that during differentiation some of the crystals or liquid may be injected and form sills or dikes. The iron ores of Kiruna, Sweden, have been described as early injections, and certain pegmatites are classed as late magmatic injections. Magmatic deposits are relatively simple in mineral composition and few in number. *See* DIAMOND; MAGMA; PERIDOTITE.

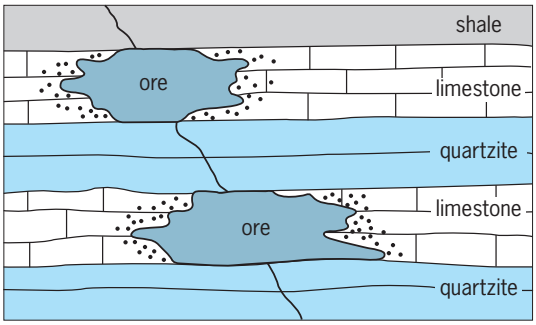

Fig. 2. Replacement of limestone by ore along a fissure. Disseminated ore, indicated by the dots, is forming in advance of the main body.

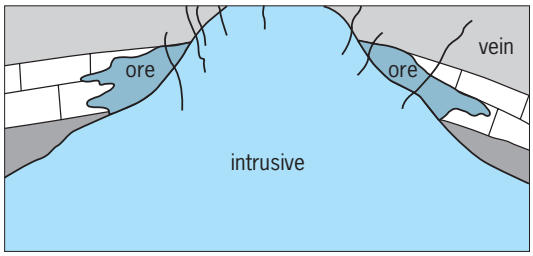

Fig. 3. Association of contact metasomatic and vein deposits with intrusive magmas.

Contact metasomatic deposits. During the crystallization of certain magmas a considerable amount of fluid escapes. This fluid may produce widespread changes near the contacts of magma with the surrounding rocks (**Fig. 3**). Where such changes are caused by heat effects, without addition of material from the magma, the resulting deposits are called contact metamorphic. If appreciable material is contributed by the magma, the deposits are termed contact metasomatic. The term skarn is applied to the lime-bearing silicates formed by the introduction of Si, Al, Fe, and Mg into a carbonate rock; some skarns contain ore bodies. The magmas that produce these effects are largely silicic in composition and resulting mineral deposits are often irregular in form. *See* SKARN.

A complicating case exists where little fluid escaped from the magma but the heat of the intrusion was great enough to cause dissolution and movement of certain metals from the surrounding rocks. It is believed by some investigators that solutions formed in this manner may become concentrated in metal content and subsequently deposit these metals near the contact of the intrusion and the surrounding rocks. In this case, the ore minerals were deposited by replacing preexisting rocks but the source of the ore is the surrounding rocks, not the magma. To further complicate matters, the ore in some deposits appears to consist of material derived from both the intrusion and the surrounding rocks. In such deposits the source of the ore is generally controversial, and the size, amount, and composition of the mineralization would depend upon the relative contributions from the intrusion and the associated rocks.

Under contact metasomatic conditions, the (ore-forming) fluids extensively replace the country rock to produce a variety of complex minerals. Contact metasomatic deposits include a number of important deposits, whereas contact metamorphic deposits are rarely of economic value. Many garnet, emery, and graphite deposits are classed as contact metasomatic, as are such metalliferous deposits as the iron ores of Cornwall, Pennsylvania, Iron Springs, Utah, and Banat, Hungary; many copper ores of Utah, Arizona, New Mexico, and Mexico; the zinc ores of Hanover, New Mexico; and various tungsten ores of California and Nevada.

Pegmatite deposits. Pegmatites are relatively coarse-grained rocks found in igneous and metamorphic regions. The great majority of them consist of feldspar and quartz, often accompanied by mica, but complex pegmatites contain unusual minerals and rare elements. Many pegmatites are regular tabular bodies; others are highly irregular and grade into the surrounding rocks. In size, pegmatites range from a few inches in length to bodies over 1000 ft (3000 m) long and scores of feet across. Some pegmatites are zoned, commonly with a core of quartz surrounded by zones in which one or two minerals predominate. *See* FELDSPAR; QUARTZ.

Pegmatites may originate by various igneous and metamorphic processes. Fractional crystallization of a magma results in residual solutions that are generally rich in alkalies, alumina, water, and other volatiles. The volatiles lower the temperature of this liquid and make it unusually fluid; the low viscosity promotes the formation of coarse-grained minerals. The rare elements that were unable by substitution to enter into the crystal structure of earlier-formed minerals, principally because of differences in size of their atomic radii, are concentrated in the residual pegmatite solutions. Late hydrothermal fluids may alter some of the previously formed pegmatite minerals.

Some pegmatites develop by replacement of the country rock and commonly these are isolated bodies with no feeders or channels in depth. They occur in metamorphic regions usually devoid of igneous rocks and contain essentially the same minerals as those in the country rocks. In some regions small pegmatites have grown by forcing apart the surrounding metamorphic rock, and others have formed by filling a fissure or crack from the walls inward. In both cases growth is believed to have taken place by diffusion and consolidation of material in the solid state. *See* PEGMATITE.

Hydrothermal deposits. Most vein and replacement deposits are believed to be the result of precipitation of mineral matter from dilute, hot ascending fluids. As the temperature and pressure decrease, deposition of dissolved material takes place. It is not altogether certain how important the gaseous state is in the transport of ore material. It may be that at relatively shallow depth and high temperature gaseous solutions transport significant amounts of ore-forming material.

W. Lindgren, who developed the hydrothermal theory, divided these deposits into three groups on the basis of temperature and pressure conditions supposed to exist at the time of formation. Deposits thought to form at temperatures of 50–200°C (120–390°F) at slight depth beneath the surface are called epithermal. Many ores of mercury, antimony, gold, and silver are of this type. Deposits formed at 200–300°C (390–570°F) at moderate depths are known as mesothermal and include ores of gold-quartz, silver-lead, copper, and numerous other types. Hypothermal deposits are those formed at 300–500°C (570–930°F) at high pressures; certain tin, tungsten, and gold-quartz ores belong to this type.

The nature of hydrothermal fluids is inferred by analogy with laboratory experiments, and by investigation of deposits forming around volcanoes and hot springs at the present time. Studies of liquid inclusions in minerals, of mineral textures, and of inversion temperatures of minerals indicate that mineralization takes place at elevated temperatures. Layers of minerals on the walls of open fissures with crystal faces developed toward the openings suggest deposition from solution. In some of these cavities later crystals were deposited on earlier ones in a manner that suggests growth in moving solutions. Certain secondary replacement phenomena, such as weathering and oxidation of mineral deposits, also indicate deposition from liquid solutions. Studies of wall rock alteration where hydrothermal solutions have attacked and replaced rock minerals indicate that these solutions change in character from place to place. Sulfur in such solutions may react with or leach metals from the surrounding rocks or partly solidified magma to form certain kinds of mineral deposits. On the basis of geochemical data it has been estimated that most hydrothermal ore-forming solutions had a temperature in the range 50–600°C (120–1100°F), formed under pressures ranging from atmospheric to several thousand atmospheres, commonly contained high concentrations of NaCl and were saturated with silica but were not highly concentrated in ore metals, were neutral or within about 2 pH units of neutrality; and that the metals probably were transported as complexes.

The principal objections to the hydrothermal theory are the low solubility of sulfides in water and the enormous quantities of water required. W. Lindgren realized this and, for some deposits, favored colloidal solutions as carriers of metals. Laboratory synthesis of sulfide minerals by G. Kullerud shows that some ore-bearing solutions must have been considerably more concentrated than is generally believed.

Two common features of hydrothermal deposits are the zonal arrangement of minerals and alteration of wall rock.

1. Zoning of mineralization. Many ore deposits change in composition with depth, lateral distance, or both, resulting in a zonal arrangement of minerals or elements. This arrangement is generally interpreted as being due to deposition from solution with decreasing temperature and pressure, the solution precipitating minerals in reverse order of their solubilities. Other factors are also involved such

as concentration, relative abundance, decrease in electrode potentials, and reactions within the solutions and with the wall rocks as precipitation progresses.

Zonal distribution of minerals was first noted in mineral deposits associated in space with large igneous bodies, and has since been extended to include zoning related to sedimentary and metamorphic processes in places where no igneous bodies are in evidence. Although many geologists interpret zoning as a result of precipitation from a single ascending solution, others believe deposition is achieved from solutions of different ages and of different compositions.

The distribution of mineral zones is clearly shown at Cornwall, England, and at Butte, Montana. At Cornwall, tin veins in depth pass upward and outward into copper veins, followed by veins of lead-silver, then antimony, and finally iron and manganese carbonates. Such zoning is by no means a universal phenomenon, and, in addition to mines and districts where it is lacking, there are places where reversals of zones occur. Some of these reversals have been explained more or less satisfactorily by telescoping of minerals near the surface, by the effects of structural control or of composition of the host rock in precipitating certain minerals, and by the effects of supergene enrichment on the original zoning, but many discrepancies are not adequately explained.

2. Wall rock alteration. The wall rocks of hydrothermal deposits are generally altered, the most common change being a bleaching and softening. Where alteration has been intense, as in many mesothermal deposits, primary textures may be obliterated by the alteration products. Chemical and mineralogical changes occur as a result of the introduction of some elements and the removal of others; rarely a rearrangement of minerals occurs with no replacement.

Common alteration products of epithermal and mesothermal deposits are quartz, sericite, clay minerals, chlorite, carbonates, and pyrite. Under high-temperature hypogene conditions pyroxene, amphibole, biotite, garnet, topaz, and tourmaline form. In many mines sericite has been developed nearest the vein and gives way outward to clay minerals or chlorite. The nature and intensity of alteration vary with size of the vein, character of the wall rock, and temperature and pressure of hydrothermal fluids. In the large, low-grade porphyry copper and molybdenum deposits associated with stocklike intrusives, alteration is intense and widespread, and two or more stages of alteration may be superimposed.

Under low-intensity conditions, the nature of the wall rock to a large extent determines the alteration product. High-intensity conditions, however, may result in similar alteration products regardless of the nature of the original rock. Exceptions to this are monomineralic rocks such as sandstones and limestones. Wall rock alteration may develop during more than one period by fluids of differing compositions, or it may form during one period of mineralization as the result of the action of hydrothermal fluids that did not change markedly in composition. Alteration zones have been used as guides to ore and tend to be most useful where they are neither too extensive nor too narrow. Mapping of these zones outlines the mineralized area and may indicate favorable places for exploration.

Sedimentary and residual deposits. At the Earth's surface, action of the atmosphere and hydrosphere alters minerals and forms new ones that are more stable under the existing conditions. Sedimentary deposits are bedded deposits derived from preexisting material by weathering, erosion, transportation, deposition, and consolidation. Different source materials and variations in the processes of formation yield different deposits. Changes that take place in a sediment after it has formed and before the succeeding material is laid down are termed diagenetic. They include compaction, solution, recrystallization, and replacement. In general, the sediment is consolidated by compaction and by precipitation of material as a cement between mineral grains. Sedimentation as a process may itself involve the concentration of materials into mineral deposits. *See* DIAGENESIS.

The mineral deposits that form as a result of sedimentary and weathering processes are commonly grouped as follows: (1) sedimentary deposits, not including products of evaporation; (2) sedimentary-exhalative deposits; (3) chemical evaporites; (4) placer deposits; (5) residual deposits; and (6) organic deposits. *See* VOLCANO.

1. Sedimentary deposits. Included in this group are the extensive coal beds of the world, the great petroleum resources, clay deposits, limestone and dolomite beds, sulfur deposits such as those near Kuibyshev, Russia, and the deposits of the Gulf Coast region, and the phosphate of North Africa and Florida. Metalliferous deposits such as the minette iron ores of Lorraine and Luxembourg, and Clinton iron ores of the United States, and the manganese of Tchiaturi, Georgia, and Nikopol in the Ukraine also belong here. There are other deposits of metals in sedimentary rocks whose origin remains an enigma, such as the uranium of the Colorado Plateau, the Witwatersrand in South Africa, and Blind River in Ontario; and the copper deposits of Mansfeld, Germany, and of the Copperbelt of Zambia and the Democratic Republic of the Congo. These deposits have characteristics of both syngenetic and epigenetic types. A controversy centers around the genesis of these and similar deposits of the world. *See* HEAVY MINERALS; SEDIMENTOLOGY.

2. Sedimentary-exhalative deposits. Many large stratiform deposits are found in marine sedimentary rocks associated with volcanic rocks. It is well known that volcanoes and fumaroles carry in their gases a number of metals. On land these gases escape into the atmosphere. Under water the gases, if they carry matter which is insoluble under the existing conditions, will precipitate their metals as oxides, sulfides, or carbonates in the vicinity of the gas emission. If the gases contain matter that is soluble, the metal content of the seawater will increase, and upon reaching saturation level will precipitate an

extensive disseminated ore deposit. Where submarine emissions take place in a large ocean basin, they may be deposited over the floor of the basin as part of the sedimentation process. *See* VOLCANO.

Deposits exemplified by lead-zinc-barite-fluorite mineralization, most commonly found in carbonate rocks, occur in the Mississippi Valley region of North America and also on other continents. These ores are included with the sedimentary-exhalative type, but could also be discussed under several other classes of deposits since they are very difficult to categorize. They have been considered by various geologists to be true sediments, diagenetic deposits, lateral secretion deposits, deposits formed by downward leaching of overlying lean ores, deposits formed by solutions that descended and subsequently ascended, deposits resulting from magmatic-hydrothermal processes, and sea-floor deposits from thermal springs. Most geologists favor either a syngenetic-sedimentary hypothesis or an epigenetic-hypogene one. Some studies hypothesize a source of metal-bearing waters similar to those in deep brines which rise and move through fissures in overlying rocks or are poured out on the sea floor and are added to accumulating sediments. A single generally acceptable hypothesis of origin, if such eventually emerges, must await the accumulation and interpretation of additional geological and geochemical data.

3. Chemical evaporites. These consist of soluble salts formed by evaporation in closed or partly closed shallow basins. Deposits of salt or gypsum that are several hundred feet thick are difficult to explain satisfactorily. Oschsenius suggested that they formed in basins which were separated from the ocean by submerged bars except for a narrow channel (inlet); such barriers are common along coastal areas. Intermittently, seawater flowed over the barrier and was concentrated into saline deposits by evaporation. Modifications of this theory have been proposed to account for the omissions of certain minerals and the interruptions in the succession.

Deposits of gypsum and common salt (halite) are found in many countries, whereas the larger concentrations of potash salts, borates, and nitrates are much more restricted in occurrence. *See* SALINE EVAPORITES.

4. Placer deposits. Placers are the result of mechanical concentration whereby heavy, chemically resistant, tough minerals are separated by gravity from light, friable minerals. Separation and concentration may be accomplished by streams, waves and currents, and air, or by soil and hill creep. The most important economic placer deposits are those formed by stream action (**Fig. 4**).

Stream and beach placers are widespread in occurence and include the famous gold placers of the world, as well as deposits of magnetite, ilmenite, chromite, wolframite, scheelite, cassiterite, rutile, zircon, monazite, and garnet. Placer deposits of diamond, platinum, and gemstones are less common.

5. Residual deposits. Complete weathering results in distribution of the rock as a unit and the segregation of its mineral constituents. This is accomplished by oxidation, hydration, and solution, and may be accelerated by the presence of sulfuric acid. Some iron and manganese deposits form by accumulation without change, but certain clay and bauxite deposits are created during the weathering of aluminous rocks. Residual concentrations form where relief is not great and where the crust is stable; this permits the accumulation of material in place without erosion. *See* WEATHERING PROCESSES.

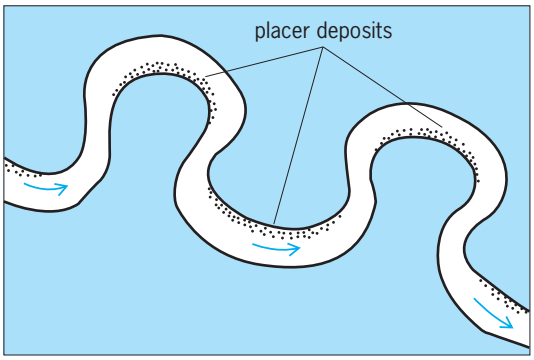

Fig. 4. Deposition of placer by stream action on the inside of meander bends.

Large residual deposits of clay, bauxite, phosphate, iron, and manganese have been worked in many parts of the world, as have smaller deposits of nickel, ocher, and other minerals.

6. Organic deposits. Plants and animals collect and use various inorganic substances in their life processes, and concentration of certain of these substances upon the death of the organisms may result in the formation of a mineral deposit. Coal and peat form from terrestrial plant remains and represent concentration by plants of carbon from the carbon dioxide of the atmosphere. Petroleum originates by the accumulation of plant and animal remains. Many limestone, phosphate, and silica deposits also form by plant and animal activity. Hydrated ferric oxide and manganese dioxide are precipitated by microorganisms; anaerobic bacteria can reduce sulfates to sulfur and hydrogen sulfide. There is considerable controversy, however, as to whether microorganisms are responsible for the formation of certain iron, manganese, and sulfide deposits. Some uranium, vanadium, copper, and other metalliferous deposits are considered to have formed, in part at least, by the activity of organisms.

Deposits formed by regional metamorphism. Regional metamorphism includes the reconstruction that takes place in rocks within orogenic or mountain belts as a result of changes in temperature, pressure, and chemical environment. In these orogenic belts, rocks are intensely folded, faulted, and subjected to increases in temperature. The changes that occur in this environment affect the chemical and physical stability of minerals, and new minerals, textures, and structures are produced, generally accompanied by the introduction of considerable material and the removal of other material.

Some geologists believe that the water and metals released during regional metamorphism can give rise to hydrothermal mineral deposits. Along faults and shear zones movement of fluids could take place by mechanical flow, though elsewhere movement might be by diffusion. The elements released from the minerals would migrate to low-pressure zones such as brecciated or fissured areas and concentrate into mineral deposits. It has been suggested that the subtraction of certain elements during metamorphism also can result in a relative enrichment in the remaining elements; if this process is sufficiently effective, a mineral deposit may result. Certain minerals also may be concentrated during deformation by flow of material to areas of low pressure such as along the crests of folds.

Deposits of magnetite, titaniferous iron, and various sulfides may form in metamorphic rocks, as well as deposits of nonmetallic minerals such as kyanite, corundum, talc, graphite, and garnet.

Opponents of the concept of mineral formation by regional metamorphism believe that a dispersal of minerals, rather than a concentration, would result from the operative processes. However, if movement of material were confined to specific channelways, this objection would not necessarily hold. *See* METAMORPHISM.

Oxidation and supergene enrichment. Many sulfide minerals form at depth under conditions differing markedly from those existing at the surface. When such minerals are exposed by erosion or deformation to surface or near-surface conditions, they become unstable and break down to form new minerals. Essentially all minerals are affected.

The oxidation of mineral deposits is a complex process. Some minerals are dissolved completely or in part, whereas elements of others recombine and form new minerals. The principal chemical processes that take place are oxidation, hydration, and carbonation. The oxidation of pyrite and other sulfides produces sulfuric acid, a strong solvent. Much of the iron in the sulfides is dissolved and reprecipitated as hydroxide to form iron-stained outcrops called gossans. Metal and sulfate ions are leached from sulfides and carried downward to be precipitated by the oxidizing waters as concentrations of oxidized ores above the water table. Oxides and carbonates of copper, lead, and zinc form, as do native copper, silver, and gold. The nature of the ore depends upon the composition of the primary minerals and the extent of oxidation. If the sulfates are carried below the water table, where oxygen is excluded, upon contact with sulfides or other reducing agents they are precipitated as secondary sulfides. The oxidized zone may thus pass downward into the supergene sulfide zone. Where this process has operated extensively, a thick secondary or supergene-enriched sulfide zone is formed. Enrichment may take place by removal of valueless material or by solution of valuable metals which are then transported and reprecipitated. This enrichment process has converted many low-grade ore bodies into workable deposits. Supergene enrichment is characteristic of copper deposits but may also take place in deposits of other metals. Beneath the enriched zone is the primary sulfide ore (**Fig. 5**).

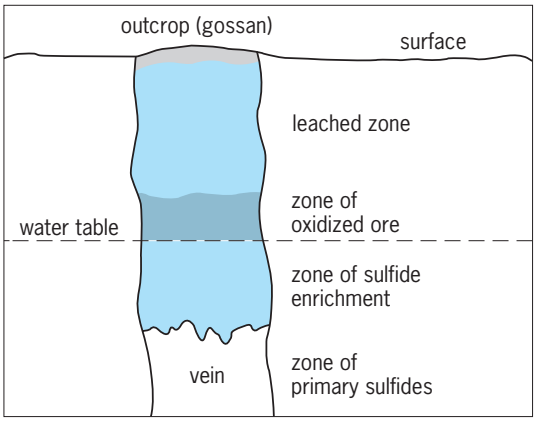

Fig. 5. Vein deposit of sulfide ore, showing changes due to oxidation and supergene enrichment.

The textures of the gossan minerals may give a clue to the identity of the minerals that existed before oxidation and enrichment took place. These have been used as guides in prospecting for ore.

Sequence of deposition. Studies of the relationships of minerals in time and space have shown that a fairly constant sequence of deposition, or paragenesis, is characteristic of many mineral deposits. This sequence has been established largely by microscopic observations of the boundary relationships of the minerals in scores of deposits. Subsequent experimental studies of mineral phases have contributed to the knowledge of paragenesis. In magmatic and contact metasomatic ores, silicates form first, followed by oxides and then sulfides. W. Lindgren presented the paragenesis for hypogene mineral associations, and others have discussed the problems involved. The sequence of common minerals starts with quartz, followed by iron sulfide or arsenide, chalcopyrite, sphalerite, bornite, tetrahedrite, galena, and complex lead and silver sulfo salts. It indicates the existence of some fundamental control but attempts to explain the variations in it have been largely unsuccessful, or are applicable to only part of the series or to specific mineralized areas. Local variations are to be expected since many factors such as replacement, unmixing, superimposed periods of mineralization, structural and stratigraphic factors, and telescoping of minerals may complicate the order of deposition.

Paragenesis is generally thought to be the result of decreasing solubility of minerals with decreasing temperature and pressure. It has also been explained in terms of relative solubilities, pH of the solutions, metal volatilities, decreasing order of potentials of elements, free energies, and changing crystal structure of the minerals as they are deposited. R. L. Stanton has reevaluated paragenetic criteria as applied to certain stratiform sulfide ores in sedimentary and metamorphic rocks. He proposes that the textures of such ores do not represent sequences of deposition but are the result of surface energy requirements

during grain growth, or annealing of deformed minerals. To explain mineral paragenesis more satisfactorily, many additional experiments must be made to determine phase relations at different temperatures and pressures. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; MINERAL.

Mineralogenetic provinces and epochs. Mineral deposits are not uniformly distributed in the Earth's crust nor did they all form at the same time. In certain regions conditions were favorable for the concentration of useful minerals. These regions are termed mineralogenetic provinces and they contain broadly similar types of deposits, or deposits with different mineral assemblages that appear to be genetically related. The time during which these deposits formed constitutes a mineralogenetic epoch; such epochs differ in duration, but in general they cover a long time interval that is not sharply defined. Certain provinces contain mineral deposits of more than one epoch.

During diastrophic periods in the Earth's history mountain formation was accompanied by plutonic and volcanic activity and by mineralization of magmatic, pegmatitic, hydrothermal and metamorphic types. During the quieter periods, and in regions where diastrophism was milder, deposits formed by processes of sedimentation, weathering, evaporation, supergene enrichment, and mechanical action.

During the 1960s numerous studies were made of the regional distribution of mineral deposits associated with long subsiding belts of sediments, or geosynclines, and with platform areas of relatively thin sediments adjoining the thick geosynclinal wedge. Geosynclinal areas commonly suffer folding and later uplift and become the sites of complex mountain ranges. It has been proposed that the outer troughs and bordering deep-seated faults contain ore deposits of subcrustal origin, the inner uplifts contain deposits of crustal origin, and the platforms contain ores derived from subcrustal and nonmagmatic platform mantle rocks. V. I. Smirnov and others have summarized available information on types of mineral deposits characteristic of the processes most active during the evolutionary stages of geosynclinal and platform regions. In the early prefolding stage of subsidence, subcrustal juvenile basaltic sources of ore fluids prevail, and the characteristic metals are Cr, titanomagnetite, Pt metals, skarn Fe and Cu, and deposits of pyritic Cu and Fe and Mn. In the folding episode, rocks of the geosyncline are melted to produce magma from which ore components are extracted or leached by postmagmatic fluids. The most typical ores of this stage are Sn, W, Be, Ni, Ta, and various polymetallic deposits. The late stage is characterized by ore deposits associated with igneous rocks and other deposits with no apparent relationship to igneous rocks. Smirnov believes these ores originated by the combined effect of subcrustal, crustal, and nonmagmatic sources of ore material. Typical metals of this stage include Pb, Zn, Cu, Mo, Sn, Bi, Au, Ag, Sb, and Hg. In the tectonically activated platform areas, deposits of Cu-Ni sulfides, diamonds, various magmatic and pegmatitic deposits, and hydrothermal ores of nonferrous, precious, and rare metals are found. In addition, there are nonmagmatic deposits of Pb and Zn. Some ore material is believed to be both subcrustal and nonmagmatic in origin. Relative proportions of types of mineralization differ from one region to another.

The relationship between mineral deposition and large-scale crustal movements permits a grouping of mineralogenetic provinces by major tectonic features of the continents such as mountain belts, stable regions, and Precambrian shields. The Precambrian shield areas of the world contain the Lake Superior, Kiruna, and Venezuelan iron provinces, the gold provinces of Kirkland Lake and Porcupine in Canada, the gold-uranium ores of South Africa, the gold deposits of western Australia, and the base metals of central Australia. In the more stable regions are the metalliferous lead-zinc province of the Mississippi Valley and provinces of salt and gypsum, iron, coal, and petroleum in different parts of the world. The mountain belts are the location of many diverse kinds of mineral provinces such as the gold-quartz provinces of the Coast Range and the Sierra Nevadas, various silver-lead-zinc provinces of the western United States, the Andes, and elsewhere, and numerous base-metal provinces in the Americas, Africa, Australia, and Europe.

Localization of mineral deposits. The foregoing discussion has shown that mineral deposits are localized by geologic features in various regions and at different times. Major mineralized districts within the shield areas and mountain belts are often localized in the upper parts of elongate plutonic bodies. Specific ores tend to occur in particular kinds of rocks. Thus tin, tungsten, and molybdenum are found in granitic rocks, and nickel, chromite, and platinum occur in basic igneous rocks. In certain regions mineral deposits are concentrated around plateau margins. Tropical climates favor the formation of residual manganese and bauxite deposits, whereas arid and semiarid climates favor the development of thick zones of supergene copper ores. Major mineralized districts are also localized by structural features such as faults, folds, contacts, and intersections of superimposed orogenic belts. The location of individual deposits is commonly controlled by unconformities, structural features, the physical or chemical characteristics of the host rock (**Fig. 6**), topographic features, basins of deposition, ground-water action, or

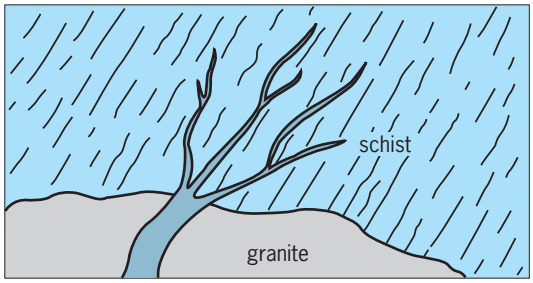

Fig. 6. Strong vein in granite dividing into stringers upon entering schist.

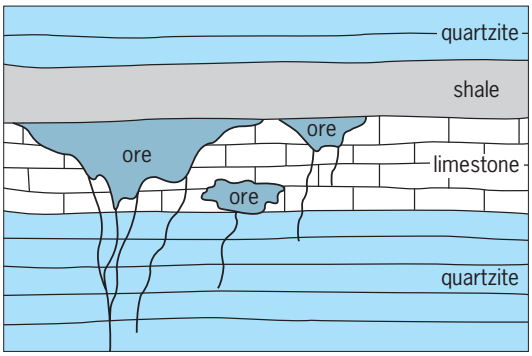

Fig. 7. Ore in limestone beneath impervious shale.

by restriction to certain favorable beds (**Fig. 7**). *See* PLUTON.

Source and nature of ore fluids. Widely divergent views have been expressed as to the original source and mode of transport of mineral deposits. Each view has certain advantages when applied to specific types of deposits. However, the complex nature of some mineralizations and the highly diverse physicochemical environments in which mineral deposits form make it impossible to select one theory to account for the source of all ore-forming materials.

According to one view, the source of the ore material was a juvenile subcrustal basaltic magma from which mineral deposits crystallized by simple crystallization, or in some cases were concentrated by differentiation. Most of the ore deposits associated with such magmas show a close spatial relationship to the enclosing igneous rocks and are similar in composition from one province to another. The exceptions to this are certain pyritic copper and skarn ores that apparently were introduced into sedimentary rocks and are now removed from the postulated source magma.

Another hypothesis holds that many ore deposits associated with granitic rocks were derived from magmas generated by remelting of deep-seated sedimentary rocks, followed by movement of the magma into higher levels of the Earth's crust. As an end product of crystallization and differentiation, an ore fluid was produced containing concentrations of metals originally present in the magma. Commonly, such deposits are confined to the apical portions of granitic plutons that have been altered by postmagmatic fluids. An increasing number of geologists ascribe to the view that the ore material was removed from the solidified magma by these late-stage fluids. Such ore deposits are complex, and their composition, dependent in part on the composition of the remelted rocks, is diverse and variable from one region to another. For certain ores associated with major deep-seated faults or with intersections of extensive fault or fissure systems, an ore source in deeper, subcrustal, regions has been advocated.

Circulation of surface waters may have removed metals from the host rocks and deposited them in available openings; this is the lateral secretion theory. The metals were carried either by cool surface waters or by such waters that moved downward, became heated by contact with hot rocks at depth, and then rose and deposited their dissolved material.

As sediments are compacted and lithified, huge volumes of water may be expelled. It has been suggested that the ore-forming fluid in some sedimentary deposits contains metals that were in solution before the sediment was buried, plus metals acquired during diagenesis of the sediments. For certain ores with colloform textures, it is believed that movement took place as a finely divided suspension, and that the ore minerals initially precipitated as gels. The metals could have been held as adsorbates on clays and other colloids and then released for concentration during later crystallization of the colloids. Crystallization would exclude the metals as finely divided material which could combine with released water and move to places favorable for precipitation.

A considerable amount of experimental work has been done on the geochemistry of ore fluids in an attempt to determine the source, nature, solubility, transport, and deposition of these fluids. Studies of metal-bearing saline waters and of thermal waters and brines in igneous, sedimentary, and metamorphic rocks have also contributed to the knowledge of this complex subject. D. E. White has analyzed and summarized these studies, and he stresses that four mutually interdependent factors must be considered: a source of the ore constituents, the dissolving of these constituents in the hydrous phase, migration of the ore-bearing fluid, and the selective precipitation of the ore constituents in favorable environments. The ore-bearing fluids are Na-Ca-Cl brines that may form by magmatic or connate processes, solution of evaporates by dilute water, or membrane concentration of dilute meteoric water.

During regional metamorphism large quantities of hydrothermal fluids may be released from rocks in deep orogenic zones. These fluids remove metals and other minerals from the country rock and redeposit them at higher levels along favorable structures. Elements may also move by diffusion along chemical, thermal, and pressure gradients.

A number of the famous mineralized districts of the world that have characteristics of both epigenetic and syngenetic deposits have been modified by later metamorphism, thereby further obscuring their origin. In some of these districts the fissure and joint systems in the rocks reflect the pattern in deeper-seated rocks. H. Schneiderhohn has suggested that repeated rejuvenation of these systems by tectonic movements, accompanied by the dissolving action of thermal waters on old ore deposits in depth, would result in upward movement and reprecipitation of metals in higher formations; Schneiderhohn calls these deposits secondary hydrothermal ores. Elsewhere old folded rocks and ore deposits have been greatly deformed, and the ores taken into solution and transported to higher and younger strata; such deposits Schneiderhohn terms regenerated ores. Controversy centers around suitable criteria for epigenetic and syngenetic deposits, the problems of solubility of metals in thermal

waters, their transport over long distances, and whether such rejuvenated and regenerated ores would be dispersed or concentrated by the processes envisaged by Schneiderhohn. A. F. Hagner

Mineral and chemical composition. The common minerals of hydrothermal deposits (**Table 3**) are sulfides, sulfo salts, oxides, carbonates, silicates, and native elements, although sulfates, a fluoride, tungstates, arsenides, tellurides, selenides, and others are by no means rare. Many minor elements which seldom occur in sufficient abundance to form discrete minerals of their own may substitute for the major elements of the minerals and thus be recovered as by-products. For example (as shown in Table 3), the ore mineral of cadmium, indium, and gallium is sphalerite; the major ore mineral of silver and thallium is galena; and pyrite is sometimes an ore of cobalt. *See* ELEMENTS, GEOCHEMICAL DISTRIBUTION OF.

Ore deposits consist, in essence, of exceptional concentrations of given elements over that commonly occurring in rocks. The degree of concentration needed to constitute ore varies widely, as shown in **Table 4**, and is a complex function of many economic and sometimes political variables. The quantity of these elements in the total known or reasonably expected ore bodies in the world is infinitesimal when compared with the total amounts in the crust of the Earth. Thus, each and every cubic mile of ordinary rocks in the crust of the Earth contains enough of each ore element to make large deposits (Table 4). Although there is a large number of geologic situations that are apparently favorable, only a few of them contain significant amounts of ore. Thus, it is evident that the processes leading to concentration must be the exception and not the rule, and obviously any understanding or knowledge of these processes should aid in the discovery of further deposits.

Each step in the process of ore formation must be examined carefully if this sporadic occurrence of ore is to be placed on a rational basis. In order for ores to form, there must bea source for the metal,

TABLE 3. Some common primary minerals of hydrothermal ore deposits

Element	Common minerals	Idealized formulas	Significant minor elements occurring in these minerals, underlined where economically important
Iron	Hematite	Fe_2O_3	
	Magnetite	Fe_3O_4	Mn
	Pyrite	FeS_2	Au,* Co, Ni
	Pyrrhotite	$Fe_{1-x}S$	Ni, Co
	Siderite	$FeCO_3$	Mn, Ca, Mg
	Arsenopyrite	FeAsS	Sb, Co, Ni
Copper	Chalcopyrite	$CuFeS_2$	Ag, Mn, Se
	Bornite	Cu_5FeS_4	
	Chalcocite	Cu_2S	Ag
	Enargite	Cu_3AsS_4	Ag, Sb
	Tetrahedrite	$Cu_{12}Sb_4S_{13}$	Ag, Fe, Zn, Hg, As
Zinc	Sphalerite	ZnS	Fe, Mn, Cd, Cu, Ga, Ge, Sn, In
Lead	Galena	PbS	Ag, Bi, As, Tl, Sn, Se, Sb
Bismuth	Native bismuth	Bi	
	Bismuthinite	Bi_2S_3	
Silver	Native silver	Ag	Au
	Argentite	Ag_2S	
	Various sulfo salts		
Gold	Native gold	Au	Ag, Cu
	Various tellurides of gold and silver		
Mercury	Cinnabar	HgS	
Tin	Cassiterite	SnO_2	
Uranium	Uraninite	UO_2	Ra, Th, Pb
Cobalt	Cobaltite	CoAsS	
	Smaltite	$CoAs_2$	
Nickel	Pentlandite	$(Fe, Ni)_9S_8$	
Tungsten	Scheelite	$CaWO_4$	Mo
	Wolframite	$(Fe,Mn)WO_4$	Mo
Molybdenum	Molybdenite	MoS_2	Re
Manganese	Rhodochrosite	$MnCO_3$	Fe, Mg, Ca
	Rhodonite	$MnSiO_3$	Ca
Others	Calcite	$CaCO_3$	Mn
	Dolomite (and ankerite)	$CaCO_3 \cdot MgCO_3$	Fe, Mn
	Barite	$BaSO_4$	
	Fluorite	CaF_2	
	Quartz	SiO_2	
	Sericite, chlorite, feldspars, clays, and various other silicates		

*Intimately associated as minute particles of metallic gold, but not in the crystal structure of the pyrite.

TABLE 4. Approximate concentration of ore elements in Earth's crust and in ores

Element	Approximate concentration in average igneous rocks, %	Tons/mi^3 rock (metric ton/km^3)	Approximate concentration in ores, %	Concentration factor to make ore
Iron	5.0	560,000,000 (120,000,000)	50	10
Copper	0.007	790,000 (170,000)	0.5–5	70–700
Zinc	0.013	1,500,000 (330,000)	1.3–13	100–1000
Lead	0.0016	180,000 (39,000)	1.6–16	1000–10,000
Tin	0.004	450,000 (99,000)	0.01*–1	2.5–250
Silver	0.00001	1,100 (240)	0.05	5000
Gold	0.0000005	56 (12)	0.0000015*– 0.01	3–2000
Uranium	0.0002	22,000 (4,800)	0.2	1000
Tungsten	0.003	340,000 (74,000)	0.5	170
Molybdenum	0.001	110,000 (24,000)	0.6	600

*Placer deposits.

a medium in which it may be transported, a driving force to move this medium, a "plumbing system" through which it may move, and a cause of precipitation of the pre elements as an ore body. These interrelated requirements are discussed below in terms of the origin of the hydrothermal fluid, its chemical properties, and the mechanisms by which it may carry and deposit ore elements.

Source of metals. It is not easy to determine the source for the metals in hydrothermal ore deposits because, as shown above, they exist everywhere in such quantities that even highly inefficient processes could be adequate to extract enough material to form large deposits.

Fluids associated with igneous intrusion. In many deposits there is evidence that ore formation was related to the intrusion of igneous rocks nearby, but in many other deposits intensive search has failed to reveal any such association. Because the crystal structures of the bulk of the minerals (mostly silicates) crystallizing in igneous rocks are such that the common ore elements, such as copper, lead, and zinc, do not fit readily, these elements are concentrated into the residual liquids, along with H_2O, CO_2, H_2S, and other substances. These hot, water-rich fluids, remaining after the bulk of the magma has crystallized, are the hydrothermal fluids which move outward and upward to areas of lower pressure in the surrounding rocks, where part or all of their contained metals are precipitated as ores. A more detailed discussion of the composition of these fluids is presented below.

Fluids obtained from diagenetic and metamorphic processes. Fluids of composition similar to the above also could be obtained from diagenetic and metamorphic processes. When porous, water-saturated sediments containing the usual amounts of hydrous and carbonate minerals are transformed into essentially nonhydrous, nonporous metamorphic rocks, great quantities of water and carbon dioxide must be driven off. Thus, each cubic mile of average shale must lose about 3×10^9 tons of water (each cubic kilometer, about 6.5×10^8 metric tons) and may lose large amounts of carbon dioxide on metamorphism to gneiss. The great bulk of the water presumably comes off as connate water (entrapped at time of rock deposition) under conditions of fairly low temperature. In many respects this water has the same seawater composition as it had to start with. However, as metamorphism proceeds, accompanied by slow thermal buildup from heat flow from the Earth's interior and from radioactivity, the last fluids are given off at higher temperatures and are richer in CO_2 and other substances. These fluids would have considerably greater solvent power and can be expected to be similar to those coming from cooling igneous rocks.

Role of surface and other circulating waters. It is very likely that the existence of a mass of hot rock under the surface would result in heating and circulation of meteoric water (from rain and snow) and connate water. The possible role of these moving waters in dissolving ore elements from the porous sedimentary country rocks through which they may pass laterally and in later depositing them as ore bodies has been much discussed. The waters may actually contribute ore or gangue minerals in some deposits. The test of this theory of lateral secretion on the basis of preise analyses of the average country rocks around an ore body would involve an exceedingly difficult sampling job. It also would require analytical precision far better than is now feasible for most elements, as each part per million uncertainty in the concentration of an element in a cubic mile of rock represents about 10,000 tons of the element or 1×10^6 tons of 1% ore.

Movement of ore-forming fluids. In addition to the high vapor pressures of volatile-rich fluids acting as a driving force to push them out into the surrounding country rocks and to the surface, there may well be additional pressures from orogenic or mountain-building forces. When a silicate magma has an appreciable percentage of liquid and is subjected to orogenic forces, it moves en masse to areas of lower pressure (it is intruded into other rocks). But if the magma has crystallized 99% or more of its bulk as solid crystals and has only a very small amount of water-rich fluid present as thin films between the grains, and then is squeezed, this fluid may be the only part sufficiently mobile to move toward regions

of lower pressure. (If the residual fluid, containing the ore elements, stays in the rock, it reacts with the early formed, largely anhydrous minerals of the rock to form new hydrated ones, such as sericite, epidote, amphibole, and chlorite, and its ore elements precipitate as minute disseminated specks and films along the silicate grain boundaries.)

The ore-bearing fluid leaves the source through a system comprising joints, faults, porous volcanic plugs, or other avenues. As the fluid leaves the source, it moves some appreciable but generally unknown distance laterally, vertically, or both, and finally reaches the site of deposition. This system of channels is of utmost importance in the process of ore formation.

Localization of mineral deposits. It is stated frequently that ore deposits are geologic accidents; yet there are reasons, however abstruse, for the localization of a mineral deposit in a particular spot. One reason for localization is mere proximity to the source of the ore-forming fluids, as in rocks adjacent to an area of igneous activity or near a major fracture system which may provide plumbing for solutions ascending from unknown depths. Zones of shattering are favored locales for mineralization since these provide plumbing and offer the best possibility for the ore solution to react with wall rock, mix with other waters, and expand and cool, all of which may promote precipitation. Some types of rock, particularly limestone and dolomite, are especially susceptible to replacement and thus often are mineralized preferentially. The chemical or physical properties which cause a rock to be favored by the replacing solutions often are extremely subtle and certainly not understood fully.

Zoning and paragenesis. Mineral deposits frequently show evidence of systematic spatial and temporal changes in metal content and mineralology that are sufficiently consistent from deposit to deposit to warrant special mention under the terms zoning and paragenesis. Zoning may be on any scale, though the range is commonly on the order of a few hundred to a few thousand feet, and may have either lateral or vertical development. In mining districts, such as Butte, Montana, or Cornwall, England, where zoning is unusually well developed, there is a peripheral zone of manganese minerals grading inward through successive, overlapping silver-lead, zinc, and copper zones (and in the case of Cornwall, tungsten, and finally tin). The same sequence of zones appears in many deposits localized about intrusive rocks, suggesting strongly that the tin and tungsten are deposited first from the outward-moving hydrothermal solutions and that the copper, zinc, lead, and silver were deposited successively as the solutions expanded and cooled. In other districts the occurrences of mercury and antimony deposits suggest that their zonal position may be peripheral to that of silver or manganese. The paragenesis, or the sequence of deposition of minerals at a single place, as interpreted from the textural relations of the minerals, follows the same general pattern as the zoning, with the tin and tungsten early and the lead and silver late. With both zoning and paragenesis there are sometimes reversals in the relative position of adjacent zones, and these are usually explained as successive generations of mineralization. Some metals, such as iron, arsenic, and gold, tend to be distributed through all of the zones, whereas others, such as antimony, tend to be restricted to a single position.

The sequence of sulfide minerals observed in zoning and paragenesis matches in detail the relative abilities of the heavy metals to form complex ions in solution. This observation strongly supports the hypothesis developed later that most ore transport occurs through the mechanism of complex ions, since no other geologically feasible property of the ore metals or minerals can explain the zoning relations.

Environment of ore deposition. Important aspects of the environment of ore deposition include the temperature, pressure, nature, and composition of the fluid from which ores were precipitated.

Temperatures. Although there is no geological thermometer that is completely unambiguous as to the temperatures of deposition of ores, there is a surprising number of different methods for estimating the temperatures that prevailed during events long since past that have been applied to ores with reasonably consistent results. Those ore deposits which had long been considered to have formed at high temperatures give evidence of formation in the range of 500–600°C (930–1100°F), or possibly even higher. Those that were thought to be low-temperature deposits show temperatures of formation that are in the vicinity of 100°C (212°F) or even less, and the bulk of the deposits lie between these extremes. *See* GEOLOGIC THERMOMETRY.

Pressures. It would be useful to know the total hydrostatic pressure of the fluids during ore formation. Most of the phenomena used for determination of the temperatures of ore deposition are also pressure-dependent, and so either an estimate of the correction for pressure must be made, or two independent methods must be used to solve for the two variables.

Pressures vary widely from nearly atmospheric in hot springs to several thousand atmospheres in deposits formed at great depth. Maximum reasonable pressures are considered to be on the order of that provided by the overlying rock; conversely, the minimum reasonable pressures are considered to be about equal to that of a column of fluid open to the surface. Pressures therefore range from approximately 500 to 1500 lb/in.^{2} per 1000 ft (10–30 MPa per 1000 m) of depth at the time of mineralization. *See* HIGH-PRESSURE MINERAL SYNTHESIS.

Evidence of composition. Geologists generally concede that most ore-forming fluids are essentially hot water or dense supercritical steam in which are dissolved various substances including the ore elements. There are three lines of evidence bearing on the composition of this fluid. These are fluid inclusions in minerals, thermal springs and fumaroles, and the mineral assemblage of the deposit and its associated alteration halos.

1. Fluid inclusions in minerals. Very small amounts of fluid are trapped in minute fluid-filled inclusions during the growth of many ore and gangue minerals in veins, and these inclusions have been studied intensively for evidence of temperature and composition. Although the relative amounts may vary widely, these fluids will have 5–25 or even more weight percent soluble salts, such as chlorides of Na, K, and Ca, plus highly variable amounts of carbonate, sulfate, and other anions. Some show liquid CO_2 or hydrocarbons as separate phases in addition to the aqueous solution. A few show detectable amounts of H_2S and minor amounts of many other substances. After losing some CO_2 and H_2S through release of pressure and oxidation when the inclusions are opened, the solutions are within 2 or 3 pH units of neutral. There is little evidence of sizable quantities (>1 g/liter or 0.13 oz/gal) of the ore metals in these solutions, and the evidence indicates that the concentrations of the ore elements must generally be very low (<0.1 g/liter or 0.013 oz/gal). Even if the concentrations were in the range of 0.1 g/liter, there should be analytical evidence in the fluid inclusion studies, but this is lacking. In addition, if fluids of such composition were trapped in fluid inclusions in transparent minerals and on cooling precipitated even a fraction of their metal content as opaque sulfides, these should be visible (under the microscope) within the inclusions, but none are seen. If the concentrations of ore elements are much less than 0.001 g/liter (0.00013 oz/gal), the volume of fluids that must be moved through a vein to form an ore body becomes geologically improbable.

2. Thermal springs and fumaroles. These provide the closest approach to a direct look at the processes of ore deposition as some ore and gangue minerals form within the range of direct observation. The solutions from these springs give diluted and possibly contaminated, partly oxidized and partly devolatilized samples of the sort of fluid that presumably forms ore bodies at greater depths. Isotopic studies show that the solutions have been diluted by local meteoric water until less than 5% (if any) of the fluid emitted at the surface is of deep-seated origin. The compositions of these thermal springs, after correction for such dilution, are in good agreement with the data from fluid inclusions.

3. Mineral assemblage. The assemblage of minerals that occurs within a deposit provides a great deal of information about the chemical nature of the fluid from which the ores were precipitated. There are a great number of stable inorganic compounds of the heavy metals known, yet unaltered ore deposits contain only a relatively small number of minerals. For example, lead fluoride, lead chloride, lead carbonate, lead sulfate, lead oxide, lead sulfide, and many others are known stable compounds of lead, yet of these, primary ore deposits contain only the sulfide (galena). Some elements, such as calcium, which occur in combination with several types of anions, for example, the carbonate, fluoride, sulfate, and numerous silicates, are found with the ore minerals. A quantitative approach to the compositional problem may be made by considering such reactions as shown in (1). The equilibrium constant for this reaction is

$$CaCO_3 + 2F^- \rightarrow CaF_2 + CO_3^{2-} \qquad (1)$$

$(CO_3^{2-})/(F^-)^2 = 10^{1.4}$ at 25°C (77°F). Thus when calcite and fluorite are in equilibrium, the requirements for the constant are met, and the $(CO_3^{2-})/(F^-)^2$ ratio is known. A large number of such equations can be evaluated and from comparison with the mineral assemblage known to occur in ores, limits on the possible variation of the composition of the ore-forming fluid may be estimated. Unfortunately, calculations of this sort involving ionic equilibria are limited to fairly low temperatures (less than 100–200°C or 212–390°F) since there are few reliable thermodynamic data on ionic species at high temperature. At any temperature, reactions such as shown in (2) can be

$$2Ag + {}^1/_2S_2 \rightarrow Ag_2S \qquad (2)$$

used to evaluate or place limits on the possible variation of the chemical potential of some components in the ore-forming fluid. *See* IONIC EQUILIBRIUM; SULFIDE PHASE EQUILIBRIA.

The composition of the ore fluid tends to become adjusted chemically by interaction with the rocks with which it comes in contact, and these changes may well contribute to the precipitation of the ore minerals. Thus, the K^+/H^+ ratio may be controlled by such reactions as (3), where the equilibrium constant has the form shown in Eq. (4). Likewise, the

$$\underset{\text{Muscovite}}{4KAl_2(AlSi_3)O_{10}(OH)_2} + 6H_2O + 4H^+ \rightarrow \underset{\text{Kaolin}}{6Al_2(Si_2O_5)(OH)_4} + 4K^+ \qquad (3)$$

$$K = \frac{(K^+)^4}{(H_2O)^6 \cdot (H^+)^4} \qquad (4)$$

quantitatively small but nevertheless important partial pressures of sulfur and oxygen may be governed by such reactions as (5). Such changes in the wall

$$Fe_3O_4 + S_2 \rightarrow Fe_2O_3 + FeS_2 + {}^1/_2O_2 \qquad (5)$$

rock come under the general heading of wall-rock alteration and may be of many types, only a few of the more common of which are mentioned below.

High-temperature alteration of limestones usually results in the formation of water-poor calcium silicates, such as garnet, pyroxenes, idocrase, and tremolite, and the resulting rock is termed skarn. At lower temperatures in the same types of rock, dolomitization and silicification are the predominant forms of alteration, because the partial pressure of CO_2 is too high to permit calcium silicate to form. *See* SILICATE PHASE EQUILIBRIA.

At high temperatures in igneous and metamorphic rocks near granite in composition, the solutions are approximately in equilibrium with the primary rock-forming minerals, and thus there is little alteration except development of sericite and occasionally topaz and tourmaline. At lower temperatures, the

characteristic sequence of alteration from fresh rock toward the vein is first an argillic zone, then a sericitic zone, and finally a silicified zone bordering the vein.

Summary. Summarizing the environment of ore deposition, there are various lines of evidence to show that most hydrothermal ore deposits were formed at temperatures of 100–600°C (212–1100°F) and at pressures ranging from nearly atmospheric to several thousand atmospheres. The solutions were dominantly aqueous and were fairly concentrated in sodium chloride and potassium chloride; however, they were relatively dilute in terms of the ore metals.

Mechanisms of ore transport and deposition. The ore minerals, principally the sulfides, are extremely insoluble in pure water at high temperatures as well as low; the solubility products are so low, in fact, that literally oceans of water would be required to transport the metal for even a small ore body. Thus, it is not easy to explain the mechanism whereby the minerals are solubilized to the extent necessary for ore transport.

Crystals of ore and gangue minerals frequently exhibit evidence of repeated partial re-solution (or leaching) and regrowth. This demonstrates that the process of ore formation may, at least in some instances, be reversible. In such cases studies of artificial systems at equilibrium are applicable.

The re-solution of ore minerals is important in another connection. Some geologists have advocated colloidal solutions, or sols, as an alternative to true solutions for ore transport. This was based on the belief, now known to be generally false, that colloform textures in ore minerals are a result of original deposition as a colloidal gel. Colloidal solutions were attractive also because they permitted ore metal concentrations—even in the presence of sulfide—many orders of magnitude higher than true solutions. The re-solution of ore minerals precludes the process of colloidal ore transport, as colloidal solutions are supersaturated and therefore cannot redissolve a crystal of the dispersed phase. *See* COLLOID.

In addition to the fact that the absolute solubilities, calculated from the solubility products, are extremely low, the relative solubilities of the sulfides are radically different. For example, according to the solubility products, FeS is many, many times more soluble than PbS (about 10^{10} times at 25°C or 77°F), yet the two minerals occur together in ore deposits and behave as if galena were slightly more soluble than pyrrhotite. From this and other lines of evidence, it appears necessary to conclude that the solubilities of the various contemporaneous minerals in a given deposit could not have differed among themselves by more than a few orders of magnitude. *See* SOLUBILITY PRODUCT CONSTANT.

The only geologically and chemically feasible mechanism by which these solubilities may be equalized approximately is the formation of complex ions of the heavy metals. Such complexes can increase the solubilities of heavy metals tremendously. As an example, the activity (thermodynamic concentration) of Hg^{2+} in a solution saturated with HgS (cinnabar) and H_2S at 25°C (77°F), 1 atm (10^5 Pa) pressure, and pH 8, is only about 10^{-47} mole/liter, representing a concentration much less than 1 atom of mercury in a volume of water equal to the entire volume of the oceans of the world. However, in the same solution is formed a very stable sulfide complex of mercury, HgS_2^{2-}, which increases the total concentration of mercury in solution by the impressive factor of about 10^{42}, giving a concentration on the order of 0.001 g/liter (0.00013 oz/gal). Not only does complex formation provide a means to achieve adequate solubility for ore transport, but the relative tendency for metals to form certain types of complexes matches in detail the commonly observed zoning and paragenetic sequences mentioned previously. The metals whose sulfides are the least soluble tend to form the most stable complexes, and metals whose minerals are comparatively soluble form weaker complexes. *See* COORDINATION COMPLEXES.

There are many kinds of complexing ions or molecules (ligands) of possible geologic importance; a few of the more significant are sulfide (S^{2-}), hydrosulfide (HS^-), chloride (Cl^-), polysulfides (S_x^{2-}), thiosulfate ($S_2O_3^{2-}$), sulfate (SO_4^{2-}), and carbonate (CO_3^{2-}), with the first three being most frequently considered. One of the major unsolved problems concerns the behavior of sulfur: What is its oxidation state and concentration relative to metals? If solutions were rich in reduced sulfur species, then the sulfide or hydrosulfide complexes would be dominant. On the other hand, solutions poor in reduced sulfur may transport the metals as chloride complexes.

The precipitation of minerals from complexed solutions takes place either by shifts in equilibrium caused by changing (usually cooling) temperature or by a decrease in the concentration of the ligand, thereby reducing the ability of the solution to carry the metals. This latter alternative can take place in several ways, for example, by reaction with wall rock, by mixing with other solutions, or by formation of a gas phase through the loss of pressure. *See* PRECIPITATION (CHEMISTRY).

Oxidation and secondary enrichment. When ore deposits are exposed at the surface, they are placed in an environment quite different from that in which they were formed, and the character of the deposit is changed through the processes of oxidation and weathering. The sulfides give way to oxides, sulfates, carbonates, and other compounds which are more or less soluble and tend to be leached away, leaving a barren gossan of insoluble siliceous iron and manganese oxides. Some minerals, such as cassiterite and native gold, may leach away at a less rapid rate than does the surrounding material; thus they are concentrated as a surficial residuum.

Where the country rock is relatively inert to the acid solutions generated by the oxidizing sulfides, as in the case of quartzites and some hydrothermally altered rocks, copper and especially zinc are leached away readily; lead and silver may be retained temporarily in the oxidized zone as the carbonate or

sulfate, and the chloride or native metal, respectively; but eventually these too are dissolved away. The various metallic ions are carried downward until they reach unoxidized sulfides in the vicinity of the water table, where the solutions interact with these sulfides to form a new series of supergene sulfide minerals. Copper sulfide is the least soluble sulfide of the base and ferrous metals in the solution, and hence the zone of supergene sulfide enrichment is predominantly a copper sulfide zone with occasional rich concentrations of silver. Zinc nearly always remains in solution and is lost in the ground water.

In reactive wall rocks, such as limestones, reaction with the wall rock prevents the solutions from becoming acid enough for large amounts of metal to be removed in solution; the base metals are retained almost in place as carbonates, sulfates, oxides, and halides, and there is no appreciable sulfide enrichment.

The behavior of some elements is governed by the availability of other materials. Thus, for example, uranium is readily leached from the oxidized zone in many deposits; however, when the oxidizing solutions contain even very small amounts of potassium vanadate, the extremely insoluble mineral carnotite precipitates and uranium is immobilized. Highly soluble materials, such as uranium in the absence of chemicals that precipitate it, may be temporarily fixed in the oxidized zone by adsorption on colloidal materials such as freshly precipitated ferric oxides.

Trends in investigation. There has been a great increase in the degree to which the experimental methods and principles of physical chemistry have been applied to aid in understanding the processes by which ores have formed, and this approach can be expected to be even more fruitful in the future. Several avenues appear promising and are under active investigation in numerous laboratories. Among these are the following:

1. Phase equilibrium studies of both natural and synthetic ore and gangue minerals.
2. Distribution coefficients for trace elements between coexisting phases, and between various forms on the same crystal.
3. Experimental solubility studies in dominantly aqueous solutions.
4. Studies of the composition and origin of thermal spring waters and fluid inclusions in minerals.
5. Thermodynamic properties of minerals.
6. Isotopic fractionation during transportation and deposition processes.
7. Rate studies on crystal growth, habit, diffusion, reaction, and transformation, as well as studies of sluggish homogeneous reactions, such as the reduction of sulfate.
8. Crystal structure determinations and crystal chemical studies of ore and gangue minerals.
9. Distribution of elements in the Earth's crust and in various rock types.
10. Detailed field studies of the relations between minerals in ore deposits. *See* HEAVY MINERALS; PETROLEUM GEOLOGY.

For a discussion of sensitive chemical analytical techniques used in the search for ore deposits *see* GEOCHEMICAL PROSPECTING

For chemical principles involved in ore deposition *see* GEOLOGIC THERMOMETRY; LEAD ISOTOPES (GEOCHEMISTRY); LITHOSPHERE.

Paul B. Barton, Jr.; Edwin Roedder

Bibliography. J. W. Barnes, *Ores and Minerals: Introducing Economic Geology*, 1988; R. L. Bates, *Geology of the Industrial Rocks and Minerals*, 1969; R. L. Bates, *Industrial Minerals: How They Are Found and Used*, 1988; D. Derry, *Concise World Atlas of Geology and Mineral Deposits: Non-metallic Minerals*, *Metallic Minerals and Energy Minerals*, 1980; A. R. Dutton, *Hydrogeology and Hydrochemical Properties of Salt-Dissolution Zones*, 1987; R. Edwards and K. Atkinson, *Ore Deposit Geology and Its Influence on Mineral Exploration*, 1986; R. M. Garrels and C. L. Christ, *Solutions*, *Minerals*, *and Equilibria*, 1965, reprint 1982; M. L. Jensen and A. M. Bateman, *Economic Mineral Deposits*, 3d rev. ed., 1981; K. B. Krauskopf, *Introduction to Geochemistry*, 3d ed., 1994; J. Parnell, H. Kucha, and P. Landais, *Bitumens in Ore Deposits*, 1993; F. Pirajano, *Hydrothermal Mineral Deposits: Principles and Fundamental Concepts for the Exploration Geologist*, 1992.

Ore dressing

Treatment of ores to concentrate their valuable constituents (minerals) into products (concentrate) of smaller bulk, and simultaneously to collect the worthless material (gangue) into discardable waste (tailing). The fundamental operations of ore-dressing processes are the breaking apart of the associated constituents of the ore by mechanical means (severance) and the separation of the severed components (beneficiation) into concentrate and tailing, using mechanical or physical methods which do not effect substantial chemical changes.

Severance. Comminution is a single- or multi-stage process whereby ore is reduced from run-of-mine size to that size needed by the beneficiation process. The process is intended to produce individual particles which are either wholly mineral or wholly gangue, that is, to produce liberation. Since the mechanical forces producing fracture are not susceptible to detailed control, a class of particles containing both mineral and gangue (middling particles) are also produced. The smaller the percentage of middlings the greater the degree of liberation. Comminution is divided into crushing (down to 6- to 14-mesh) and grinding (down to micrometer sizes). Crushing is usually done in three stages: coarse crushing from run-of-mine size to 4–6 in. (10–15 cm) or coarser; intermediate crushing down to about $^1/_2$ in. (1.27 cm); and fine crushing to $^1/_4$ in. (0.64 cm) or less.

Screening is a method of sizing whereby graded products are produced, the individual particles in each grade being of nearly the same size. In beneficiation, screening is practiced for two reasons: as an

integral part of the separation process, for example, in jigging; and to produce a feed of such size and size range as is compatible with the applicability of the separation process. *See* SCREENING.

Beneficiation. This step consists of two fundamental operations: the determination that an individual particle is either a mineral or a gangue particle (selection); and the movement of selected particles via different paths (separation) into the concentrate and tailing products. When middling particles occur, they will either be selected according to their mineral content and then caused to report as concentrate or tailing, or be separated as a third product (middling). In the latter case, the middling is reground to achieve further liberation, and the product is fed back into the stream of material being treated.

Selection is based upon some physical or chemical property in which the mineral and gangue particles differ in kind or degree or both. Thus, in hand picking, the oldest form of beneficiation, color, luster, and shape are used to decide whether a lump of ore is predominantly mineral or gangue. Use is made of differences in other physical or chemical properties, such as specific gravity, magnetic permeability, inductive charging (electrostatic separation), surface chemical properties, bulk chemical properties, weak planes of fracture (separation by screening), and gamma-ray emission (automatic sorting of radioactive materials). *See* MECHANICAL SEPARATION TECHNIQUES.

Separation is achieved by subjecting each particle of the mixture to a set of forces which is usually the same irrespective of the nature of the particles excepting for the force based upon the discriminating property. This force may be present for both mineral and gangue particles but differing in magnitude, or it may be present for one type of particle and absent for the other. As a result, separation is possible, and the particles are collected as concentrate or tailing.

Magnetic separation. Magnetic separation utilizes the force exerted by a magnetic field upon magnetic materials to counteract partially or wholly the effect of gravity. Thus under the action of these two forces, different paths are produced for the magnetic and nonmagnetic particles. **Figure 1** shows a continuously moving endless belt B onto which are introduced the particles to be separated. The magnetic field is produced by a square-shaped bottom pole B and an upper pole A, so curved as to concentrate the lines of force as shown. When a magnetic particle comes within the magnetic field, it is attracted strongly enough to move upward against the force of gravity to the surface of an endless belt A conformed to the surface of the pole. The movement of belt A perpendicularly to belt B carries the particle to a concentrate bin. The unattracted nonmagnetic particle, held by gravity, keeps moving with belt B until it falls into the tailings bin.

All substances placed in a magnetic field acquire magnetic properties. Magnetic permeability is a measure of the ease with which these properties are induced. The ratio of permeabilities may be as low as 5:1 for successful separation. However, for such commonly separated materials as magnetite and quartz it is about 110:1.

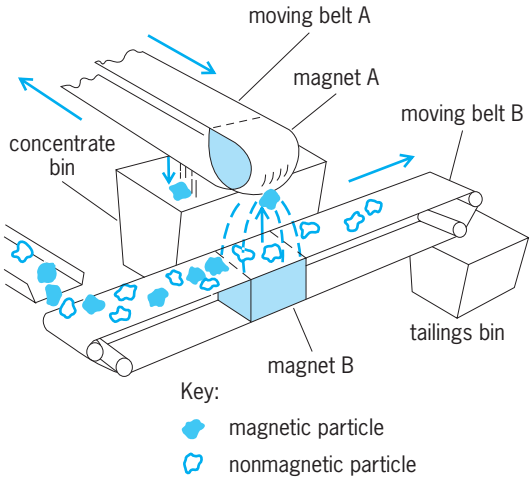

Fig. 1. Magnetic separation.

The most important practical separations are those of the iron ores. In the magnetite ores, magnetite is separated from quartz, feldspars, and the like. In the hematite and limonite ores, the ore is first roasted to convert these iron oxides partly into the magnetic oxide, and this is then separated from gangue. In the preparation of industrial minerals, magnetic separation is used to clean up iron introduced during grinding, as in the preparation of china clay, body slip, or glaze, or to remove trace magnetic minerals, such as biotite, garnet, and tourmaline, from feldspar.

Gravity methods. Gravity concentration is based on a discriminating force the magnitude of which varies with specific gravity. The other force that is usually operating in gravity methods is the resistance to relative motion exerted upon the particles by the fluid or semifluid medium in which separation takes place.

Jigging is a gravity method which separates mineral from gangue particles by utilizing an effective difference in settling rate through a periodically dilated bed. In **Fig. 2**, the mixture of particles (feed)

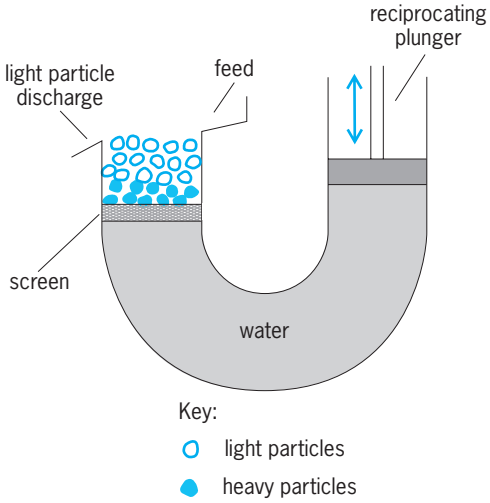

Fig. 2. Diagram of the jigging method.

falls into the jig compartment where it is supported by the screen. Reciprocation of the plunger forces water through the screen and causes periodic dilation and contraction of the bed of particles. During the dilation heavier particles work their way to the bottom while the lighter particles remain on top and are discharged over the lip. Jigging is practiced on materials which are liberated upon being reduced to sizes ranging from $1^1/_2$ in. (3.8 cm) down to several millimeters. It has been used on such diverse ores as coal, iron ores, gold, and lead ores.

Tabling is a gravity method in which the feed, introduced onto an inclined plane and reciprocated deck, moves in the direction of motion while simultaneously being washed by a water film which moves it also at right angles to the motion of the deck. In **Fig. 3**, feed enters at the top of the table, and collects within the valleys formed by the narrow cleats, or riffles, which taper in height from right to left. Under the effect of the reciprocating action, the particles stratify with the heavier particles on the bottom, and they also move from right to left. Owing to decreasing taper of the riffles, the exposed upper surface of the stratified material is acted upon by crosscurrents of water and by the tilt of the deck, as indicated by the arrow, and is moved downhill. The heavier mineral and the lighter gangue are usually collected over the edges of the deck as shown. The boundary between the heavier mineral and lighter gangue particles is roughly a linear diagonal band on the deck of the table. This diagonal band is not stationary; rather it tends to move about a mean position. In practice, therefore, a third product, the middling, is collected between the discharge edges of concentrate and gangue. If the feed to the table has been crushed or ground to produce liberation, then the middling is returned to the feed. If liberation has not been achieved, the middling is returned to the crushing-grinding section of the mill. Tables may be used to treat relatively coarse material (sand tables) or fines (slime tables), with sizes ranging from about 0.08–0.12 in. (2–3 mm) down to 0.003 in. (0.07 mm).

Sink-float separation is the simplest gravity method and is based on existing differences in specific gravity. The feed particles are introduced into a suspension, the specific gravity of which is between that of

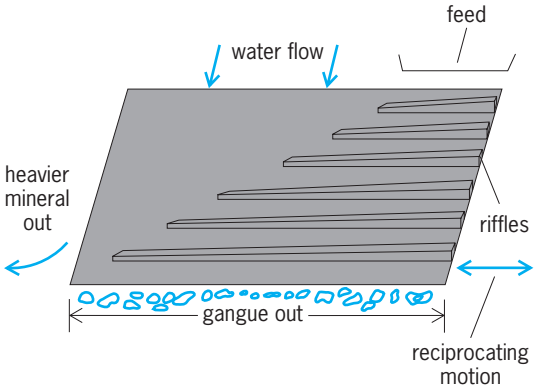

Fig. 3. Diagram of the tabling method.

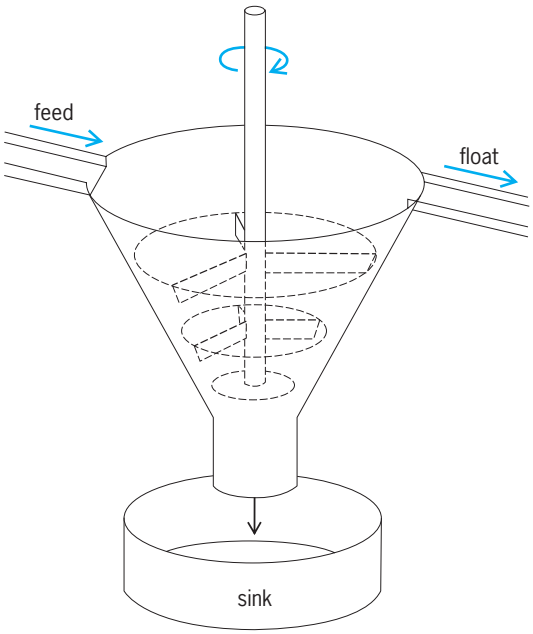

Fig. 4. Sink-float separation.

the mineral and gangue particles, with the result that particles of higher specific gravity sink while those of lower specific gravity float. In **Fig. 4** the separator is a cone equipped with a slowly operated stirrer which serves to impart a slow rotary motion to the suspension and to prevent the suspension from settling out on the walls. Feed is introduced at one point of the circumference and is slowly moved by the rotating motion of the suspension. By the time this material has reached the discharge point on the circumference, those particles whose specific gravity is greater than that of the suspension have moved down through it so that only float particles are discharged at the top. The sink particles are discharged at the bottom.

The suspension may have a specific gravity ranging from 1.3, using quartz, to 2.4, using galena. Magnetite and ferrosilicon are also used at intermediate densities. Although there is no top limit to the size of feed particles, a lower limit of about $^1/_8$ in. (3 mm) exists with the more standard equipment.

Earliest use of sink-float was in the separation of slate from coal using quartz in suspension. A most important use is to produce coarse-size tailing which can be discarded. In this manner it is possible early in the treatment to reduce the quantity of material handled by the concentrating plant, thereby effecting a saving in the capital investment.

Filtration. Filtration is a method of separation based on the differences in size of the things being separated. Water, one of the things being separated, has no lower size, whereas the solid from which it is being separated has a lower size. Consequently if a barrier (filter cloth) having openings which can pass water but not the solids is provided, and if the pulp is placed on one side of the filter, filtration will take place if a pressure is exerted on the pulp. Other methods of concentration are electrostatic

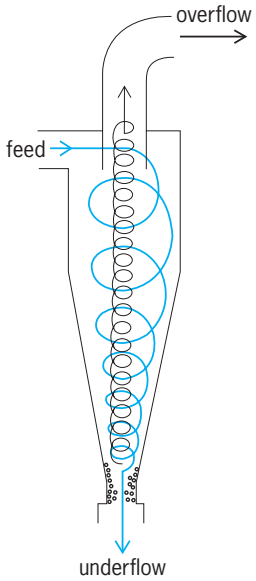

Fig. 5. Separation by cyclone classification.

separation, flotation, and leaching. *See* FILTRATION; FLOTATION; LEACHING.

Cyclone classification. Cyclone classification is a method of separating coarser from finer particles based upon the difference in magnitude of the centrifugal force to which they are subjected. The feed (a mixture of particles and water) is fed tangentially at the top of the cyclone (**Fig. 5**). The angular motion thus imparted to the pulp causes centrifugal forces to act on all the particles. The magnitude of this force depends upon the angular velocity of the pulp and the weight of the particle which, in turn, depends upon its size and specific gravity. This force is greatest on the heaviest particles and least on the lightest. Thus, the coarser particles move outward from the axis of rotation to the inner surface of the cyclone, whereas the finer particles stay closer to the axis of rotation. Because of the geometry of the cyclone, the pulp follows a downward, inward spiraling path with increasing angular velocity which produces increasing centrifugal forces. Thus, the coarser particles collect near the inner surface of the cyclone and move downward to the underflow discharge. The lighter particles with most of the water are carried in an upward spiraling column near the axis of rotation to the overflow discharge. By controlling the diameter of the underflow apex relative to the diameter of the overflow, it is possible to vary the size of the particles which report in overflow and underflow.

The principal use of the cyclone is to classify the discharge from a grinding mill and return the coarser particles to the feed for the grinding mill. Overflow goes to the concentrating plant. This makes it possible to avoid grinding beyond the desired degree of liberation of the valuable minerals from the gangue. Overgrinding can cause problems in later concentration processes, leading to serious losses in recovery of desirable minerals. *See* METALLURGY; PYROMETALLURGY, NONFERROUS. Menelaos D. Hassialis

Bibliography. I. I. Inculet, *Electrostatic Mineral Separation*, 1984; A. Lynch, *Mineral Crushing and Grinding Circuits: Their Simulation, Design, and Control*, 1977; A. L. Mular and R. B. Bhappu (eds.), *Mineral Processing Plant Design*, 2d ed., 1980; B. A. Wills, *Mineral Processing Technology: An Introduction to the Practical Aspects of Ore Treatment and Mineral Recovery*, 6th ed., 1997.

Orectolobiformes

Members of the order Orectolobiformes (carpet sharks), share the following characteristics: two dorsal fins without spines; a short snout that leaves the mouth terminal or almost terminal; and each nostril connected to the mouth by a deep groove, the anterior margin having a well-developed barbel. The classification hierarchy is:

Class Chondrichthyes
Subclass Elasmobranchii
Superorder Euselachii
Order Orectolobiformes

The order consists of about 31 species in 14 genera and 7 families and is represented by at least one family in the seas of the world. Among them is the nurse shark (Family Gingymostomatidae). It occurs near shore in the Atlantic, Indian, and Pacific oceans and is often seen in marine aquariums. This order also includes the whale shark (Family Rhincodontidae), which has a large, terminal mouth, long gill rakers, and exceptionally large gill openings, characteristics indicative of a filter feeder. The whale shark reaches a length of at least 12 m (40 ft), making it the world's largest cold-blooded animal. *See* CHONDRICHTHYES; ELASMOBRANCHII. Herbert Boschung

Bibliography. L. Compagno, M. Dando, and S. Fowler, *Sharks of the World*, Princeton Field Guides, 2005; J. S. Nelson, *Fishes of the World*, 3d ed., 1994.

Oregano

An herb, also known as wild marjoram. The dried leaves of several species of aromatic plants are known as oregano; thus oregano is a common name for a general flavor and aroma rather than the name of a specific plant.

European (*Origanum vulgare*) and Greek (*O. hervacleoticum*) oregano are both in the mint family (lamiaceae). Mexican oregano is obtained primarily from plants of *Lippia graveolens*. These small aromatic shrubs in the verbena family grow wild in Mexico. Origanum oil used in perfumery is steam-distilled primarily from Spanish oregano, *Thymus capitatus*. *See* LAMIALES.

European oregano, the most cultivated oregano species, is often confused with its close relative marjoram because they look and smell similarly. Marjoram has a bittersweet flavor, while European oregano can be distinguished by its strong piquant character

and tall growth with dark, broad leaves; it is a perennial erect herb 2–3 ft tall (0.6–1 m) with pubescent stems, ovate dark green leaves, and white or purple flowers. Native to southern Europe, southwest Asia, and the Mediterranean countries, European oregano is usually found growing in the dry, rocky, calcareous soils of the mountain regions. Greece, Italy, Spain, Turkey, and the United States are the primary sources of European oregano.

Although gathered from wild plants, European oregano is widely cultivated. It is often established with transplants from stem cuttings due to its slow germination and seedling development. European oregano does well on light, slightly alkaline, and well-drained soils that are kept dry. The plants can be harvested several times a year for 4 or 5 years.

Dried oregano leaves are used as a culinary herb in meat and sausage products, salads, soups, Mexican foods, and barbeque sauces. The essential oil of oregano is used in food products, cosmetics, and liqueurs. *See* MARJORAM; SPICE AND FLAVORING.

Seth Kirby

Organic chemistry

The study of the structure, preparation, properties, and reactions of carbon compounds. The term organic was early applied to compounds derived from plant and animal sources. These substances from living systems were usually distillable liquids or low-melting solids and were flammable, in contrast to metals, salts, and oxides from mineral sources. Until about 1830 it was held by some that organic compounds contained some special quality, or vital force. This notion was dispelled, but the term organic remained and became broadened to include carbon compounds in general. *See* CARBON.

Structure

Many organic compounds were known at the beginning of the nineteenth century, but a unifying structural principle was not available until about 1850. A crucial key was the concept that a carbon atom in its compounds always formed four linkages or bonds to other atoms. This was followed by rapid development of a structural theory that included the ideas of linear and cyclic carbon chains, multiple bonds, and isomerism. In 1874, to account for the phenomenon of optical rotation, it was proposed that carbon had a tetrahedral structure. These principles, greatly expanded and elaborated by additional concepts such as the electron-pair bond and later the molecular orbital picture of bonding, provide the underpinning for the hundreds of types of compounds and the millions of individual compounds known today.

The structures of organic compounds are described by a molecular framework of carbon atoms on which substituents may be located at various points. The *Beilstein Handbuch* is a multivolume reference that provides a systematic record of compounds organized into three classes: acyclic (aliphatic), carbocyclic, andheterocyclic (with rings containing atoms other than carbon). Each of these classes is divided into subclasses according to the presence of functional groups.

Structures can be represented in several ways, as illustrated for the three-carbon alcohol 2-propanol (**1**) and the cyclic ketone 2-methyl-3-cyclohexenone (**2**).

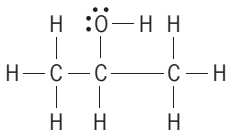

(**1*a***)

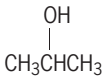

(**1*b***)

(**1*c***)

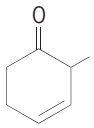

(**2**)

The expanded structure of 2-propanol (**1*a***) shows all bonds and electron pairs, including unshared electrons on oxygen. More compact and convenient is the condensed structure (**1*b***) in which the C—C and C—H bonds are implied. In the bond-line convention (**1*c***), all C—C bonds are indicated by a line, as shown for 2-propanol. Carbon atoms are not shown explicitly, but rather are implied at the ends of each line segment, together with enough hydrogen atoms to complete the tetravalency at each carbon. The bond-line convention is particularly convenient for cyclic structures such as 2-methyl-3-cyclohexenone; each vertex and the end of each line segment represents a carbon and appropriate number of hydrogens. *See* STRUCTURAL CHEMISTRY; VALENCE.

Functional groups. A functional group is an atom other than carbon or a multiple bond, such as the hydroxyl group (OH) of 2-propanol, the double bond (C═C), or the carbonyl group (C═O) of 2-methyl-3-cyclohexenone. The group defines a class of compounds and is the point at which characteristic reactions occur, for example, oxidation, reduction, or addition of an electrophilic or nucleophilic reagent. Some of the principal functional groups are shown in the **table**. *See* ELECTROPHILIC AND NUCLEOPHILIC REAGENTS.

Isomerism. The fact that there can be two or more compounds, known as isomers, with the same molecular composition was one of the key points in development of a structural theory. One type of isomerism, structural or constitutional, is illustrated by the two isomers that have the formula C_4H_{10}, butane (**3*a***) and isobutane (2-methylpropane; **3*b***). The

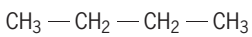

(**3*a***)

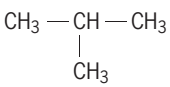

(**3*b***)

number of possible structural isomers becomes enormous in larger molecules. Examples of the molecular formulas of some hydrocarbons and the number of

possible isomers for each are given below.

C_3H_8	1	C_8H_{18}	18
C_4H_{10}	2	$C_{10}H_{22}$	75
C_6H_{14}	5	$C_{20}H_{42}$	366, 319

See MOLECULAR ISOMERISM.

Conformation. Several three-dimensional representations of butane, showing the tetrahedral geometry of the carbon atoms, are given in structures (**4**). As

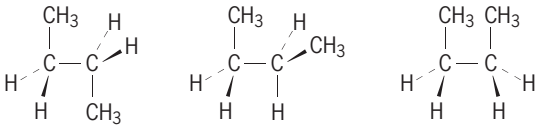

(**4*a***) (**4*b***) (**4*c***)

indicated in these structures, butane can exist in several forms, called conformations, which differ in the relative positions of the carbon atoms, and thus the overall shape of the molecule. However, the barrier to rotation around the central C—C bond is so low that these individual conformational isomers are not separable, and butane is thus a single compound. *See* CONFORMATIONAL ANALYSIS.

In an alkene, rotation around the C═C bond does not occur, and 2-butene, for example, exists as two isomeric compounds, cis (Z) and trans (E) (**5*a*** and **5*b***, respectively).

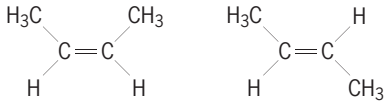

(**5*a***) (**5*b***)

Principal organic functional groups

Composed class	Group	Structure
Alkene	Double bond	>C═C<
Alkyne	Triple bond	—C≡C—
Alcohol	Hydroxyl	—OH
Amine	Amino	$—NH_2(—NR_2)$*
Aldehyde	Carbonyl	—C(═O)H
Ketone	Carbonyl	—C(═O)R
Acid	Carboxyl	—C(═O)OH
Ester	Alkoxycarbonyl	—C(═O)OR
Amide	Carbamoyl	—C(═O)N<
Nitrile	Cyano	—C≡N
Azide	Azido	—N═N═N
Nitro		$—NO_2$
Sulfide		—S—
Sulfoxide		—S(═O)—
Sulfonic acid		$—SO_3H$

*R = any carbon group, for example, CH_3.

Stereoisomerism. Stereoisomers are compounds that have the same bond sequence but differ in the spatial array of the bonds. When a carbon atom is bonded to four unlike atoms or groups, the tetrahedral geometry of carbon causes the atom to be dissymmetric or chiral. A compound, with a chiral atom can exist in two isomeric forms, known as enantiomers. The relative positions of all atoms is identical in the two enantiomers, but they differ in handedness, a characteristic of an asymmetric object and its nonsuperposable mirror image, as in structures (**6**) of 1-chlorobutane.

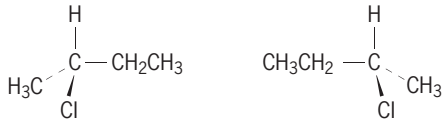

(**6*a***) (**6*b***)

See STEREOCHEMISTRY.

When two chiral centers are present, two stereoisomers can arise from each enantiomer. Thus enantiomer (**7*a***) of chlorobutane can lead to isomeric structures (**7*b***) and (**7*c***), in which the relative

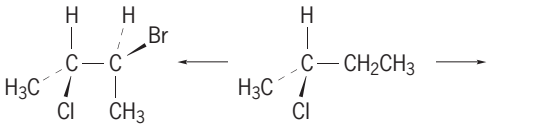

(**7*b***) (**7*a***)

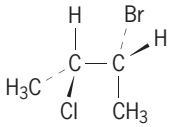

(**7*c***)

positions of the atoms is not identical. In this case, the isomers are known as diastereoisomers. With n chiral centers, there can be 2^n stereoisomers.

Acyclic compounds. The simplest organic compound is methane (CH_4). It is the first member of the homologous series of alkanes, in which successive compounds differ by an additional $—CH_2—$ group (CH_3CH_3, $CH_3CH_2CH_3$, and so forth). Methane is the major constituent of natural gas, and is the most abundant organic compound in the biosphere. It is also released into the atmosphere by the decay of rotting vegetation and fermentation in the rumens of grazing animals. *See* ALKANE; METHANE; NATURAL GAS.

Higher alkanes, $CH_3(CH_2)_nCH_3$ (n = 3–20), and also branched isomers and cyclic hydrocarbons are the principal components of petroleum. These compounds have no reactive functional groups.

However, several reactions that involve major structural changes, such as cracking into smaller molecules, isomerization of carbon chains, and removal of hydrogen are carried out on these compounds by catalytic processes, often at high temperatures, in petroleum refining. *See* CRACKING; HYDROCRACKING; ISOMERIZATION.

Both acyclic and cyclic carbon frameworks can contain multiple bonds; oxygen, nitrogen, and sulfur atoms; and other functional groups listed in the table. Alcohols, alkenes, aldehydes, amines, acids, and so on occur in plant and animal tissues and can be isolated from these sources. If they are required for some use they can be prepared as necessary by laboratory methods, or on large scale by the chemical process industry. As an example, long straight molecules, often containing one or more double bonds and terminating in an ester group, are the characteristic structural features of fats and waxes. Hydrolysis of fats by reaction with aqueous alkali gives the salts of fatty acids [soaps; reaction (1)].

$$\underset{\text{Ester (fat)}}{RC(=O)OR'} \xrightarrow{NaOH} \underset{\text{Soap}}{RC(=O)O^{-}Na} + R'OH \qquad (1)$$

These acids were among the compounds available to organic chemists in early investigations. *See* FAT AND OIL; HYDROLYSIS; SOAP.

Polymers. There is no fixed limit to the length of the carbon chain. Alkenes, with the functional group C═C, can undergo polymerization; successive molecules add to a reactive intermediate (R·). An important example is the formation of polyethylene from ethylene ($H_2C{=}CH_2$), reaction (2), where R· is

$$R\cdot + \underset{\text{Ethylene}}{H_2C{=}CH_2} \longrightarrow (RCH_2CH\cdot) \xrightarrow{H_2C=CH_2} R(CH_2CH_2)\cdot_{n+1} \xrightarrow{Y} \underset{\text{Polyethylene}}{R(CH_2CH_2)_{n+1}Y\cdot} \qquad (2)$$

the initiating radical. The value of n in polyethylene can be as high as 10,000. Other polymers, each with characteristic properties and specific uses, can be obtained from various vinyl monomers [$H_2C{=}CHX$; reaction (3), where X = alkyl, halogen, or some

$$\underset{\text{Vinyl monomer}}{H_2C{=}CHX} \longrightarrow -(H_2C\overset{X}{C}H)_n- \qquad (3)$$

other group]. These macromolecules are manufactured on a scale of millions of tons per year, and are by far the largest-volume product of the chemical process industry. Another major type of polymer is obtained by condensation, in which monomer units are linked by reactions that involve the formation of amide or ester bonds. A typical example is formation of a nylon polyamide [reaction (4)].

$$n\left[{}^{-}OC(=O)(CH_2)_4C(=O)O^{-}\right] + n\left[H_3\overset{+}{N}(CH_2)_6\overset{+}{N}H_3\right] \xrightarrow{-nH_2O} \left[-C(=O)(CH_2)_4C(=O)NH(CH_2)_6NH-\right]_n \qquad (4)$$

See POLYAMIDE RESINS; POLYESTER RESINS; POLYMER; POLYMERIZATION.

Carbocyclic compounds. The two large groups of compounds with rings containing only carbon are alicyclic and aromatic. The parent hydrocarbons in the former series are cycloalkanes and in the latter, benzene.

Alicyclic compounds. Cycloalkanes are named simply by adding the prefix cyclo- to the name of the alkane corresponding to the number of CH_2 groups in the ring. In cyclopropanes and cyclobutanes, with three- and four-membered rings, respectively, the normal tetrahedral bond angles of 109° are compressed, and reactions that lead to breaking a C—C bond are facilitated. Rings with five or more carbons are not significantly strained; the rings are puckered, and the chemistry is much like that in acyclic compounds. A cyclic structure restricts the number of conformations available to a molecule. The lowest-energy conformation of a cyclohexane ring is a chairlike structure (**8**): six bonds (the heavy lines in the structure),

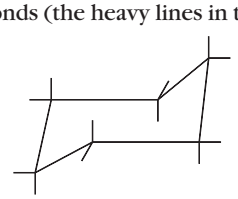

(8)

termed axial, project above and below the ring, and six others, termed equatorial, lie in the mean plane of the ring.

Compounds with several six-membered rings fused together are very common among naturally occurring terpenes and steroids. Highly condensed multiring structures have been prepared by various methods, often taking advantage of intramolecular photocyclizations. An example is the reversible isomerization of norbornadiene to quadricyclane [reaction (5)], which has been suggested as a means of

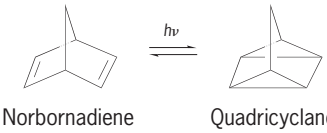

(5)

energy storage. Among the interesting compounds that have been prepared are a tetrahedrane, cubane, and dodecahedrane (**Fig. 1**), with carbon skeletons corresponding to three of the five platonic solids—the tetrahedron, the cube, and the dodecahedron.

Aromatic compounds. Benzene, a hydrocarbon with composition C_6H_6, played a major role in the concept of cyclic structures and later in the development of bonding theory. The structure of benzene (**9**) is a planar six-numbered ring with six electrons in a

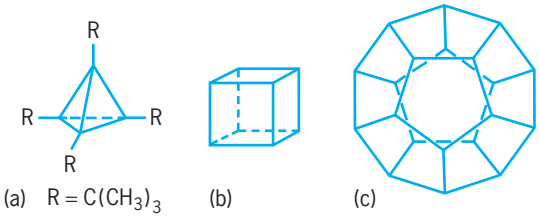

Fig. 1. Highly condensed multiring structures. (*a*) Tetrahedrane. (*b*) Cubane. (*c*) Dodecahedrane.

delocalized array. Since benzene does not show the reactions expected for a highly unsaturated compound, the details of the bonding were obscure for many decades. The Hückel formulation of an aromatic ring as one containing $4n + 2$ π electrons in a cyclic conjugated system provides a unifying basis for the structure of benzene and also nonbenzenoid aromatic compounds. Thus, the five- and seven-membered rings in cyclopentadiene anion (**10**) and cycloheptatriene cation (**11**), respectively, with six π electrons, and the cyclopropenium cation (**12**), with two electrons ($4n + 2$ where $n = 0$), have properties that qualify these species as aromatic. An aromatic system can be symbolized by a circle inscribed within the appropriate ring.

(9)

(10)

(11)

(12)

See AROMATIC HYDROCARBON; BENZENE; RESONANCE (MOLECULAR STRUCTURE).

Benzene, its methyl derivatives toluene and xylenes, and a number of polycyclic aromatic hydrocarbons such as naphthalene and anthracene can be isolated from coal tar distillates, although the major source of benzene and toluene is now cyclization and dehydrogenation of alkanes from petroleum. Benzene was available in quantity in the mid-nineteenth century, and the first extensive studies of the reactions of organic chemistry were those of aromatic compounds. Among the major developments were synthetic dyes based on aniline, which were the foundation of several large German chemical firms. *See* AROMATIZATION; COAL CHEMICALS; DEHYDROGENATION; DESTRUCTIVE DISTILLATION; DYE.

Heterocyclic compounds. A nitrogen, oxygen, or sulfur atom can take the place of carbon in either alicyclic or aromatic rings. The most numerous and important hetrocyclic compounds are those with nitrogen in a five- or six-membered aromatic system. Pyrrole (**13**) and pyridine (**14**) are the simplest compounds, and one or several other heteroatoms can be present at various positions as well, giving rise to thousands of heterocyclic systems with one or two rings. Examples are the pyrimidines (**15**) and purines (**16**), which make up the backbone of the

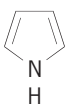

(13)

(14)

(15)

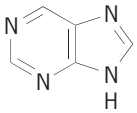

(16)

(17)

nucleic acids. All of the B vitamins, chlorophyll, heme pigments, and alkaloids have heterocyclic structures, for example, 1,2,4-oxadiazole (**17**). *See* CHLOROPHYLL; PURINE; PYRIDINE; PYRIMIDINE; PYRROLE; VITAMIN.

Two major areas of contemporary organic chemistry are the development of new drugs, and also chemical agents for use as selective herbicides, pesticides, and plant growth regulators. The primary thrust in both areas is the preparation and pharmacological or agricultural evaluation of many candidate compounds. These possess, almost without exception, heterocyclic structures. *See* HETEROCYCLIC COMPOUNDS.

Naturally occurring compounds. All living organisms, from prokaryotes to mammals and higher plants, use the same basic groups of compounds in their metabolism. These components are lipids, carbohydrates, nucleic acids, and proteins. The major lipids are fats, which are triesters of glycerol with three long-chain acids, phospholipids with one carboxylic group replaced by a phosphoric acid derivative, and nonsaponifiable compounds such as cholesterol. Carbohydrates consist of sugars, which are five- and six-carbon polyhydroxyaldehydes and their polymeric acetals starch and cellulose. Nucleic acids are very large polymers built up from four different nucleotides by phosphate-ester bonds. Proteins are smaller polymers of α-amino acids linked by peptide (amide) bonds; a total of 20 different amino acids are used in mammalian proteins.

The chemistry of each of these groups of compounds has become a major area of itself, with distinctive experimental methods. Investigations are concerned primarily with the formation, functions, and interactions of these molecules in living cells. Although the original structural studies on these compounds were carried out by organic chemists in the nineteenth century and have added immensely to organic chemistry, contemporary work is usually considered the province of biochemistry. *See* BIOCHEMISTRY; CARBOHYDRATE; LIPID; NUCLEIC ACID; PROTEIN.

Secondary metabolites. An area of natural-products chemistry that was pursued at a very early period was the isolation of active principles of medicinal plants. One group of such compounds comprises the alkaloids. These are cyclic amines with basic properties, and include some compounds with extremely complex structures such as morphine and strychnine. *See* ALKALOID; AMINE; MORPHINE ALKALOIDS; STRYCHNINE ALKALOIDS.

Another large group of plant products comprises the terpenes. These are hydrocarbons, alcohols, and ketones with 10, 15, 20, or 30 carbon atoms. They are present, sometimes in substantial quantities, in many essential oils. Some terpenes from marine plants contain several halogen atoms. Work on the elucidation of the structures of terpenes and their chemistry led to the first clear recognition that the carbon skeleton of a molecule can undergo rearrangement, or scrambling, during a reaction. *See* ESSENTIAL OILS; TERPENE.

Alkaloids, terpenes, and another group of compounds with a phenylpropanoid structure are referred to collectively as secondary metabolites. Their biochemical origin in plant cells has been studied extensively and is understood in some detail, but their functions in the plants and their evolutionary significance are obscure.

Antibiotics. Much attention in contemporary work on naturally occurring compounds has been directed to antibiotic substances produced by various microorganisms. Antibiotics are a very diverse class of compounds, but some of the clinically useful compounds can be grouped into a few general structural types: β-lactam, such as cephalosporin (**18**); tetracycline (**19**), with R = N, Cl, or OH; and macrolide, such as erythromycin (**20**). Each type has several representatives.

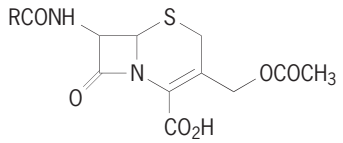

(18)

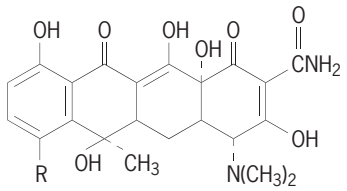

(19)

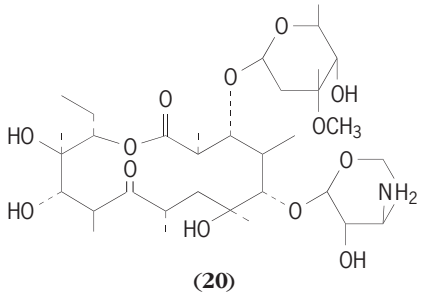

(20)

See ANTIBIOTIC.

Separation and structure determination. Progress in every phase of organic chemistry has depended on the methods available for separation, purification, and structure determination of compounds from sources such as plant extracts, coal tar, petroleum, or complex reactions. The separation of mixtures of compounds with very similar properties has required increasingly refined experimental methods and has provided a constant impetus for the development of new methods.

Early separations depended mainly on fractional crystallization and distillation; both methods require substantial amounts of material. Chromatographic separations based on adsorption and selective elution were first applied to pigments and then to sugars, which were difficult to crystallize and impossible to distill. A major advance was partition chromatography, in which a solution of a mixture is passed over a column of cellulose or strip of paper. The components of the mixture are separated by differential extraction between a stationary phase, for example, water absorbed on the cellulose, and a mobile solvent or vapor phase. This principle is employed in gas chromatography and high-pressure liquid chromatography (HPLC). Modern refinements permit the analysis and separation of extremely complex mixtures on a nanogram scale. A more recent application is the use of a stationary phase with a chiral coating on which enantiomers can be separated by diastereoisomeric complexation, permitting accurate determination of enantiomeric purity. *See* CHEMICAL SEPARATION TECHNIQUES; CHROMATOGRAPHY; CRYSTALLIZATION; DISTILLATION; GAS CHROMATOGRAPHY; LIQUID CHROMATOGRAPHY.

Until about 1940, methods for establishing the structure of a new compound depended entirely on chemical characterization of various features and functional groups, systematic degradation to smaller molecules, and a great deal of insight. This was an arduous task; the final structures of some of the alkaloids that had been isolated in the 1820s were not solved until 1950.

The development of spectroscopic and x-ray diffraction methods revolutionized elucidation of structure in the span of one generation. Ultraviolet, visible, and infrared spectroscopy reveal the presence or absence of certain functional groups and bonds; such information was the key to determination of the structure of penicillin G. A far more powerful tool is nuclear magnetic resonance (NMR). By this method, a signal for nearly every carbon, hydrogen, and nitrogen atom, as well as the environment of the atoms, can be seen in the spectrum. An even more direct view of a molecule is provided by x-ray crystallography, which shows the exact location of atoms in the crystal lattice. With these methods, only microgram-milligram samples are needed for a complete determination of structure, and they have enormously speeded the pace of research. *See* INFRARED SPECTROSCOPY; NUCLEAR MAGNETIC RESONANCE (NMR); X-RAY CRYSTALLOGRAPHY.

Synthesis Reactions

The preparation of compounds occupies much of the effort of organic chemistry, and is the principal business of the chemical industry. The manufacture of drugs, pigments, and polymers entails the preparation of organic compounds on a scale of thousands to billions of kilograms per year, and there is constant research to develop new products and processes. Synthesis of new substances is carried out for many purposes beyond the goal of a commercial

product. A compound of a specified structure may be needed to test a mechanistic proposal or to evaluate a biochemical response such as inhibition of an enzyme. Synthesis may provide a more dependable and less expensive source of a naturally occurring compound; moreover, a synthetic approach permits variations in the structure that may lead to enhanced biological activity.

The term synthesis usually implies a planned sequence of steps leading from simple starting compounds to a desired end product. Each of these steps involves a reaction that may lead to formation of a C—C bond or to the introduction, alteration, or removal of a functional group. Progress in synthesis depends on the availability of a wide range of reactions that bring about these changes in good yield, with a minimum of interfering by-products. An integral part of synthesis is the development of new methods and reagents that are selective for a desired transformation, and, very importantly, proceed with control of the stereochemistry. *See* ASYMMETRIC SYNTHESIS; ORGANIC SYNTHESIS.

Carbon-carbon bond formation. Most syntheses involve the creation of one or more C—C bonds, and there are many ways to accomplish this. Two of the more important general approaches are the reactions of organometallic compounds and the reactions of enolates.

Organometallic reactions. The first generally useful organometallic compounds were organomagnesium halides, discovered by V. Grignard. These compounds are readily prepared from the metal and an organic halide, and react with any carbonyl group. Grignard reagents are effective for the preparation of simple alcohols, for example, but because of their high reactivity they are not selective, and other functional groups interfere. However, conversion of the

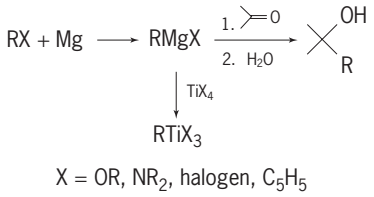

(6)

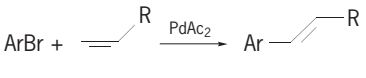

(7)

$$RHC{=}CH_2 + CO + H_2 \xrightarrow{HCo(CO)_4} RCH_2CH_2CH_2OH \quad (8)$$

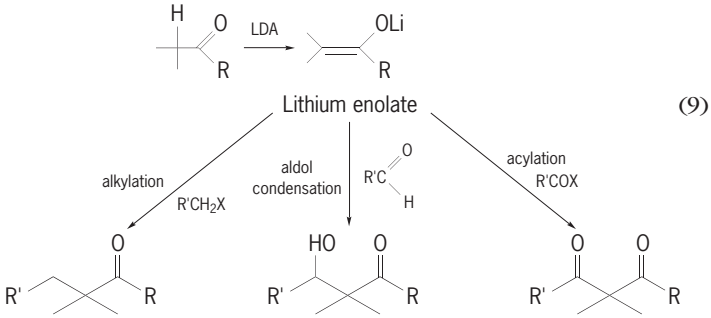

(9)

magnesium compound to another organometallic derivative such as a titanium complex leads to reagents with a range of reactivity and high stereoselectivity [reactions (6)]. *See* GRIGNARD REACTION; ORGANOMETALLIC COMPOUND.

Several other transition metals have important uses in synthesis. In these processes, the reactants are brought together and react in the coordination shell of the metal. An organometallic compound is formed, but in most of these reactions this intermediate is not stable and a catalytic cycle can be maintained. Examples are the coupling of a halide and an alkene [reaction (7)] and hydroformylation [reaction (8)]; the latter is a major industrial process. *See* HYDROFORMYLATION; TRANSITION ELEMENTS.

Enolate reactions. Removal of an α-hydrogen from a carbonyl compound gives an enolate derivative, which can then serve as a nucleophile, that is, an electron pair donor, for alkylation, aldol addition, or acylation [reaction scheme (9); X = any leaving group]. The original carbonyl compound can be an aldehyde, ketone, or ester, and the products can be reduced, dehydrated, or transformed in other ways. Use of the very strong base lithium diethylamide (LDA) permits complete conversion of the carbonyl compound to the lithium enolate at -70°C, greatly facilitating control of the subsequent steps. These reactions provide a very general and flexible approach for the assembly of carbon skeletons with functional groups at specific locations. *See* ACYLATION; ALDEHYDE; ESTER; KETONE.

An illustration of these processes is shown in reaction (10). Alkylation of the diketone by 1,4-addition

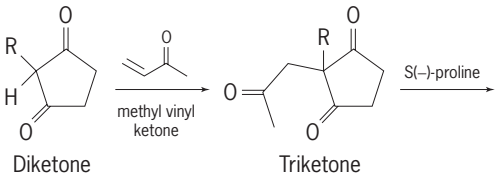

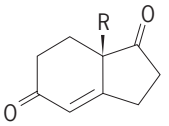

(10)

to methyl vinyl ketone gives the triketone. Aldol condensation of the latter in the presence of S(−)-proline as a chiral catalyst gives a single enantiomer of the bicyclic product, which is an important precursor for the synthesis of steroid hormones and analogs.

Functional group transformations. Many thousands of reactions and reagents have been discovered or invented to introduce, modify, and remove functional groups. These reactions involve all of the general processes of organic chemistry such as substitution, elimination, addition, oxidation, and reduction, and involve multiple bonds, oxygen, nitrogen, sulfur, and numerous other elements. Most frequently, the transformation is the end objective, as in the introduction and reduction of a nitro group, ($-NO_2$), to prepare an amine (RNH_2).

On other occasions a functional group may be modified to obtain a synthon (reagent) to be used

in the formation of a new C—C bond. Examples are conversion of an alcohol to an iodide or trifluoromethysulfonyl ester (CF_3SO_3R) for alkylation, or an aldehyde to the silyl enol ether or dithioacetal [reactions (11)].

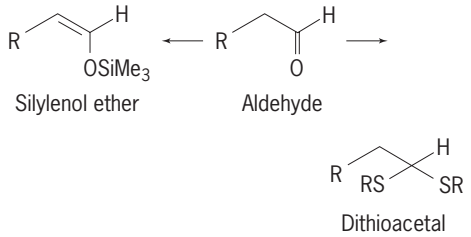

(11)

Similarly, a carboxylic acid can be converted to the more reactive acid chloride (RCOCl) or to a mixed anhydride or activated ester for peptide synthesis. *See* ACETAL; ALCOHOL; CARBOXYLIC ACID; ETHER.

Another objective is the introduction and removal of protective groups. In complex syntheses it is often necessary to block reactions of a functional group until a later stage of the sequence, when it is then regenerated under mild, selective conditions. An example is use of the trichloroethyl ether or ester group to protect an alcohol or acid, respectively, against undesired reactions. The group is removed under nonhydrolytic conditions by reductive elimination with zinc.

Logical design in synthesis. Relatively simple syntheses can be based on an obvious relationship between the desired product and readily available precursors. For example, a cinnamyl alcohol can be obtained by reduction of the corresponding ester, and the latter can be prepared by condensation of the appropriate benzaldehyde [reaction (12)]. The X

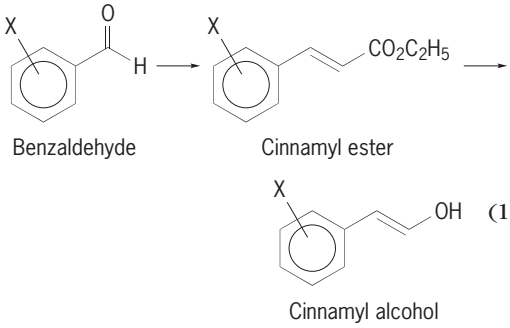

(12)

term shows that this is a general reaction (any group X) at any position.

In the synthesis of more complex structures, such as those of naturally occurring compounds, many different reaction sequences are possible. In this situation, it is desirable to recognize many alternatives and to decide which will be the most practical and efficient. This problem can be facilitated by a method known as retrosynthetic analysis, in which the structure of the target compound is examined in order to discover any bonds whose disconnection will lead to simpler molecules or fragments. An example of this approach is the synthesis of the plant regulator gibberellic acid. Working backward, retrosynthetic steps (indicated by ⇒) suggested that the precursors shown in sequence (13)

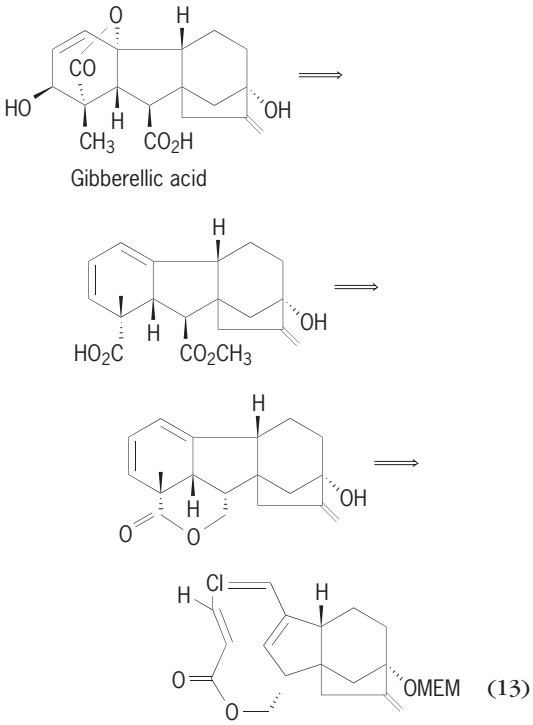

(13)

[OMEM = methoxyethoxymethyl] would provide a suitable route, and the synthesis was accomplished in this way. Computer programs with files of information on methods for bond formation, functional group manipulation, and protective groups have been developed to aid in this type of analysis. *See* COMPUTATIONAL CHEMISTRY.

Theory and Mechanisms

As knowledge of structures and reactions of organic compounds has accumulated, understanding of reaction mechanisms and correlation of structure and reactivity have advanced by application of the methods of physical chemistry and the principles of molecular orbital theory.

Bonding. In 1916, G. N. Lewis suggested that covalent bonds could be described as electron pairs shared by two atoms. A more precise formulation followed from the development of quantum theory, and the concept of atomic orbitals, which are functions of the electron density at various points of space surrounding the atomic nucleus. The orbitals of carbon that are important in bonding are designated 2*s*, $2p_x$, $2p_y$, $2p_z$ or hybrids of these, *sp*, sp^2, and sp^3, in which one, two, or three of the *p* atomic orbitals are "mixed" with the *s* atomic orbital. When the atomic orbitals of two atoms combine, two molecular orbitals, one bonding and one antibonding (MO^*) at a higher energy level, are formed. Overlap along the axis between the nuclei leads to a σ molecular orbital, as shown in **Fig. 2***a* for the combination of two sp^3 methyl radicals to form the σ C—C bond in ethane.

Another type of bond is possible with *p* orbitals, which can overlap sidewise to give a π molecular orbital, as illustrated in the double bond of an alkene

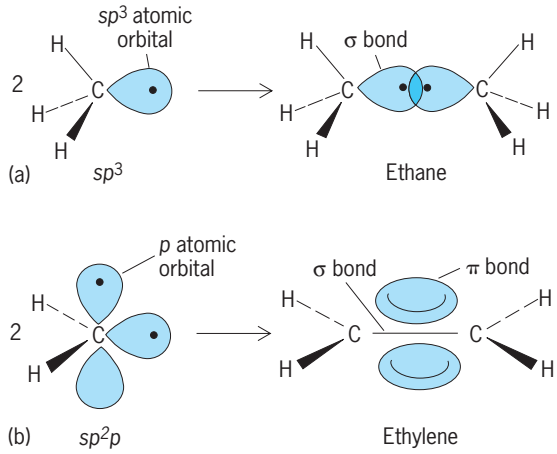

Fig. 2. Bonding orbitals in a (*a*) σ bond and (*b*) π bond.

(Fig. 2*b*). This additional π bond has a nodal plane passing through the atoms, with electron density above and below, and it contributes only 40% of the total bonding energy. A triple bond can be represented as a σ and two orthogonal π bonds.

In Fig. 2 only bonding orbitals are depicted. Treatment of conjugated systems, and the analysis of reactions in molecular orbital terms, requires consideration of the energy levels of both bonding and antibonding molecular orbitals. Diagrams of the π-electron systems of ethylene and 1,3-butadiene are shown in **Fig 3**. The size and shading of the orbital at each position indicates the magnitude and algebraic sign; there is a node, and thus no bonding, between lobes of opposite sign. The highest-energy molecular orbital that is occupied by electrons (HOMO) and the lowest-energy unoccupied orbital (LUMO) are designated the frontier orbitals.

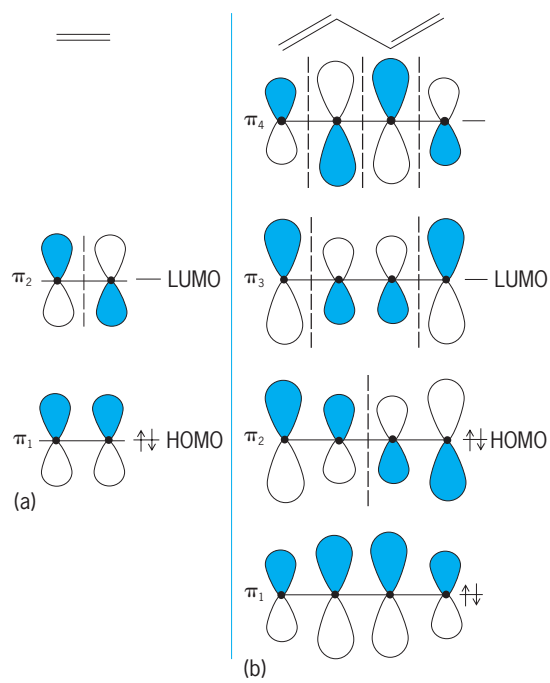

Fig. 3. Molecular orbitals of (*a*) ethylene and (*b*) 1,3-butadiene. (Broken vertical line designates node.)

In the concept of frontier molecular orbital theory, interactions of the HOMO of one reactant with the LUMO of another produces a mutual perturbation of molecular orbitals as bond changes begin. When a reactant has two or more reactive sites, or when two reactants are compared, the product is generally the one that results from the most effective of the possible HOMO-LUMO interactions. Applications of this principle have provided a unifying basis for understanding reactions and observations that are not otherwise accounted for. *See* CHEMICAL BONDING; MOLECULAR ORBITAL THEORY; WOODWARD-HOFFMANN RULE.

Reactions. The course of any reaction is governed by thermodynamics. In the reaction $A + B \rightleftharpoons C + D$, the extent to which the reaction proceeds, K_{eq}, depends on the free-energy change (ΔG), which in turn is determined by the changes in enthalpy (ΔH) and entropy (ΔS) as expressed by Eq. (14).

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ = -RT \ln K_{eq} \qquad (14)$$

In reactions at relatively low temperatures, the entropy term is usually small, and the major factor is ΔH, which represents the difference between the energies of the bonds made in the products and that lost in the reactants. *See* CHEMICAL EQUILIBRIUM; CHEMICAL THERMODYNAMICS; ENTHALPY; ENTROPY; FREE ENERGY; PHYSICAL ORGANIC CHEMISTRY; THERMOCHEMISTRY.

A separate consideration is the reaction rate (k_r) and the kinetic parameters. A widely used concept is the activation energy (E_A) of a transition barrier that must be overcome in the rate-determining step of the reaction. The measured rate and E_A are related by the Arrhenius equation (15), where A is a term known

$$k_r = Ae^{-E_a/RT} \qquad (15)$$

as the preexponential factor, R is the gas constant, and T is the absolute temperature. Measurement of the rate at several temperatures permits evaluation of $\Delta H^\ddagger$ and $\Delta S^\ddagger$ in reaching the transition state. *See* CHEMICAL DYNAMICS; GAS CONSTANT.

Acidity. An important property of organic compounds is proton acidity, as expressed by the equilibrium constant (K_A) for dissociation in water (H_2O), reaction (16), where HA represents the acid and H_3O^+ is the hydronium ion.

$$HA + H_2O \rightleftharpoons H_3O^+ + A^- \qquad (16)$$

$$K_A = \frac{[H_3O^+][A^-]}{[HA]}$$

$$pK_A = -\log K_A$$

See HYDROGEN ION.

For compounds that are weaker acids than water, the equilibrium is determined in a nonaqueous solvent and related to the K_A in water. The acid dissociation constant K_A is usually expressed as the negative logarithm pK_A. In this way, acid strengths ranging

over 10^{50} from alkanes ($pK_A \approx 50$) to strong acids ($pK_A \approx 0$–2) can be compared on a common scale of pK_A. Some representative pK_A values are given below:

CH_3CH_3	50		
$CH_2{=}CH_2$	44	CH_3CH_2OH	15.9
$C_6H_5CH_3$	41	C_6H_5OH	10.0
$CH_3CH_2NH_2$	35	$CH_3\overset{O}{\overset{\|}{C}}OH$	4.7
$HC{\equiv}CH$	25		
$CH_3\overset{O}{\overset{\|}{C}}CH_3$	20	$CH_2Cl\overset{O}{\overset{\|}{C}}OH$	2.8

The pK_A of an acid is a sensitive probe of structural effects. The acidity of a proton bonded to carbon is extremely low, but it increases significantly when the carbon is doubly or triply bonded (sp^2 or sp hybridized) because the electrons of the C—H bond are brought closer to the atom nucleus. An adjacent carbonyl group greatly increases the acid strength of C—H, N—H, or O—H bonds, reflecting the dispersal and stabilization of the charge on the anions, as in the enolate (**21**) or carboxylate (**22**). The enhanced acid-

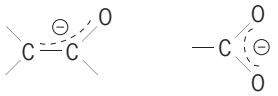

(**21**) (**22**)

ity of an α-chloroacid results from the inductive effect of the electronegative halogen atoms, as shown in notation

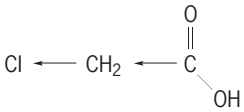

(17)

The influence of electronic factors on pK_A can be seen in the substituted benzoic acids (**23**). An

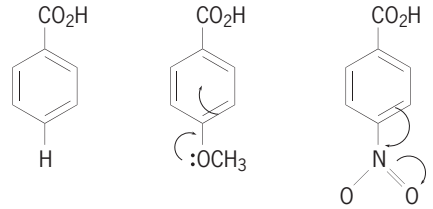

pK_A 4.20 (**23*a***) pK_A 4.47 (**23*b***) pK_A 3.42 (**23*c***)

electron-releasing group, —OCH_3, reduces acid strength, and an electron-withdrawing group, —NO_2, increases it. The effects of substituents on the pK_A of benzoic acids are the basis of a more general correlation termed a linear free-energy relationship. By this correlation, the effect of a given *meta* or *para* substituent on the pK_A of a benzoic acid can be compared with the effect of the same substituent on some other process. This approach provides a powerful tool in studies of mechanism. *See* ACID AND BASE; PK; SOLUTION.

Aliphatic nucleophilic substitution. Substitution of a group X on a sp^3 carbon by a nucleophile that supplies an electron pair is a very general and extensively studied reaction. In a typical case, the leaving group X is a halide or sulfonate anion. Two mechanisms for nucleophilic substitution are designated S_N1 and S_N2, referring to the kinetic order. The S_N2 process depends on both substrate and nucleophile (Nu) concentrations; the transition state involves attack of nucleophile from the side opposite X, with inversion of stereochemical configuration [reaction (18)]. The reaction rate is greatly retarded by the presence of bulky substituents.

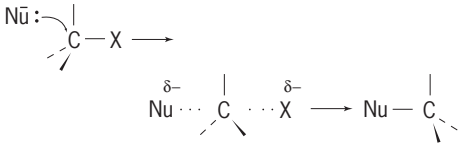

(18)

In the S_N1 mechanism, the group X departs in an initial step with formation of a carbenium ion, R_3C^+, which then reacts with a nucleophile, often a solvent molecule. The rate depends strongly on the stability of the electron-deficient intermediate, which may undergo skeletal arrangement during the process. *See* HALOGENATED HYDROCARBON; STERIC EFFECT (CHEMISTRY); SUBSTITUTION REACTION.

Aromatic substitution. Aromatic compounds undergo several types of substitution. With strongly electrophilic reagents (E^+) such as bromonium (Br^+), nitronium (NO_2^+), acylium (RCO^+), or sulfur trioxide (SO_3), substitution occurs by formation of a cationic intermediate, which then loses a proton [reaction (19)].

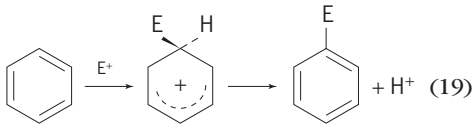

(19)

Nucleophilic substitution of halogen on an aromatic ring with strongly basic reagents can occur by an elimination-addition process, as in the formation of aniline via benzyne intermediate. This mechanism was established by isotopic labeling (the black square represents ^{14}C), which showed that two adjacent positions in the ring became labeled [reaction (20)].

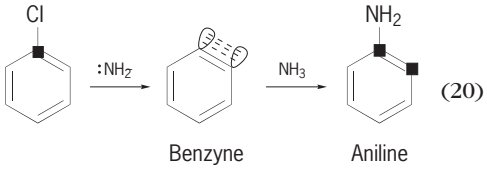

(20)

If an electron-withdrawing group is present, substitution can occur by an addition-elimination sequence [reaction (21)].

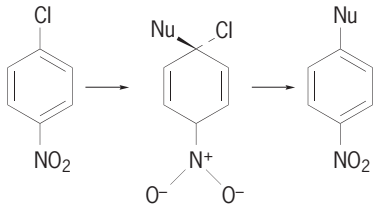

(21)

Addition reactions. A multiple bond can undergo addition, with formation of two new σ bonds and loss of a π bond, by several mechanisms. Addition reactions of alkenes occur by attack of an electrophilic reagent on the electron-rich π bond. With 2-methylpropene, for example, reaction with acids leads to the more stable carbenium ion, which then reacts with a nucleophile such as bromide ion (Br^-) to give the bromide [reaction (22*a*)]. In the presence

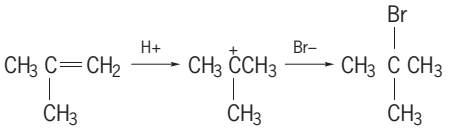

(22*a*)

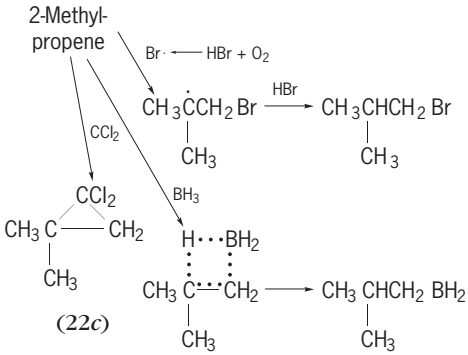

(22*c*)

(22*b*)

(22*d*)

of air, the electrophile is a bromine radical (Br·), and the isometric product is produced [reaction (22*b*)]. Other additions occur by essentially simultaneous formation of both bonds as in the reaction with carbenes to give cyclopropanes [reaction (22*c*)], or by a four-center transition state with borane to give an alkylborane [reaction (22*d*)]. *See* HYDROBORATION; REACTIVE INTERMEDIATES.

Additions to the C═O group of carbonyl compounds are among the most important organic reactions. In this case, addition occurs by attack of a nucleophile at the carbon atom of the polarized C═O bond. A diverse range of products results, depending on the nature of the carbonyl compound and reagent, but the first step in nearly all these reactions is to form the tetrahedral intermediate [reaction (23)].

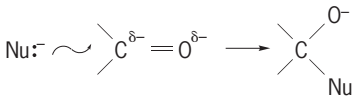

(23)

See ACID HALIDE.

Pericyclic reactions. Several reactions, termed pericyclic, occur by a concerted reorganization of bonds, with no intermediate. In these reactions, the transition state between reactants and products is an aromatic system. Two of these processes, cycloaddition and sigmatropic rearrangement, are very useful synthetic reactions.

Cycloadditions are exemplified by the Diels-Alder reaction, in which a diene combines with a dienophile to form a six-membered ring [reaction (24)]. Another type of cycloaddition involves a dipo-

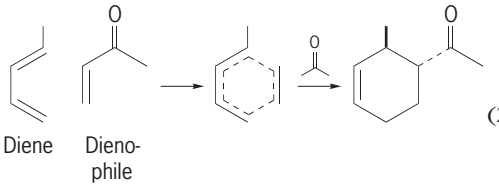

(24)

lar molecule such as a nitrone rather than a diene; the product in this case is a five-membered heterocyclic compound [reaction (25)].

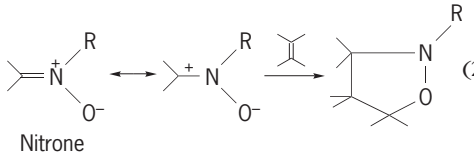

(25)

See DIELS-ALDER REACTION; PERICYCLIC REACTION.

In a sigmatropic rearrangement, a bond in a diene chain is broken and a new bond is formed. The chain may contain heteroatoms and may be part of a cyclic structure [reaction (26)].

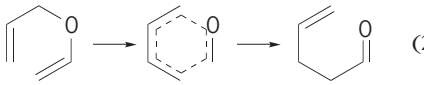

(26)

James A. Moore

Bibliography. F. A. Carey and R. J. Sundberg, *Advanced Organic Chemistry* pts. A and B, 4th ed., 2000; E. J. Corey and X. M. Cheng, *The Logic of Chemical Synthesis*, 1989; T. L. Gilchrist, *Heterocyclic Chemistry*, 3d ed., 1997; H. Hart, *Organic Chemistry: A Short Course*, 9th ed., 1994; J. Lambert et al., *Organic Structural Analysis*, 1976; T. H. Lowry and K. S. Richardson, *Mechanism and Theory in Organic Chemistry*, 3d ed., 1987; J. March, *Advanced Organic Chemistry: Reactions, Mechanisms and Structure*, 5th ed., 2001; A. Streitwieser, Jr., C. H. Heathcock, and E. M. Kosower, *Introduction to Organic Chemistry*, 4th ed., 1992; K. P. C. Vollhardt and N. E. Schore, *Organic Chemistry*, 3d ed., 1998; K. Weissermel and H.-J. Arpe, *Industrial Organic Chemistry*, 3d ed., 1997.

Organic conductor

An organic substance with low electrical resistance. Two major classes of organic conductors are charge-transfer compounds and conducting polymers.

Charge-Transfer Compounds

The charge transfer compounds, a class of well-ordered molecular crystals, display metallic-like electrical conduction and many other unusual phenomena, such as the stabilization of charge-density waves

and spin-density waves, unusual mechanisms for electronic transport, organic superconductivity, and unusual states of matter produced under strong magnetic fields. Most of these phenomena are due to the low dimensionality (one or two dimensions) of the electron gas in these compounds. Superconductivity is observed at temperatures as high as 10 K (−442°F). In addition, experiment and theory suggest a Cooper-pair binding mechanism for electrons in these conducting compounds which, unlike that of normal inorganic superconductors, does not require lattice vibrations but may be mediated by magnetic interactions between electrons. *See* SUPERCONDUCTIVITY.

Molecular structure. Charge-transfer compounds are two-component materials containing anionic and cationic species originating by charge transfer between donor and acceptor entities; these may be two organic molecules or an organic molecule with an inorganic ion. Representative systems are planar molecules with π−orbitals around carbon and heteroatoms with p_z atomic orbitals (**Fig. 1**). Tetrathiafulvalene-tetracyanoquinodimethane (TTF-TCNQ) is the prototype of charge-transfer organic crystals. Its crystal structure exhibits piled-up segregated columns of donor TTF and acceptor TCNQ molecules (**Fig. 2***a*). *See* MOLECULAR ORBITAL THEORY.

Another class of organic conductors is exemplified by radical cation salts such as $(TMTSF)_2X$ where the organic molecule is tetramethyltetraselenafulvalene and X is an inorganic anion such as phosphorus hexafluoride (PF_6^-), perchlorate (ClO_4^-), or nitrate (NO_3^-). In this class of materials (Bechgaard salts) the

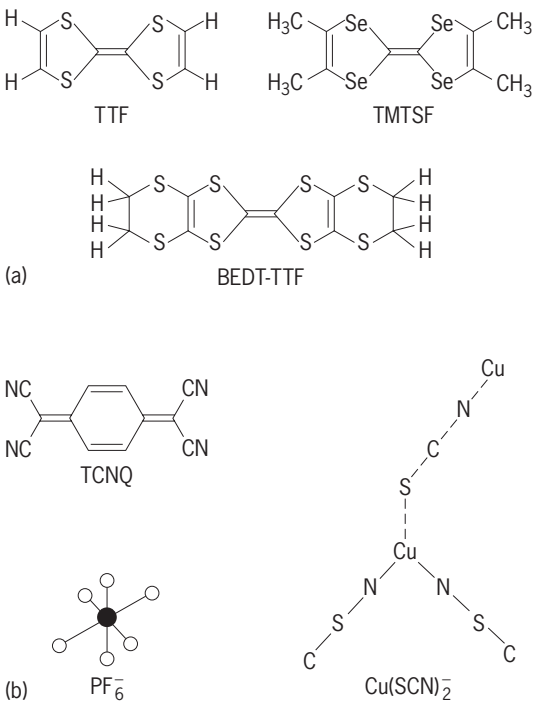

Fig. 1. Structures of organic conductors. (*a*) Electron donors. (*b*) Acceptor molecule (TCNQ) and two monovalent anions. Abbreviations are explained in the text.

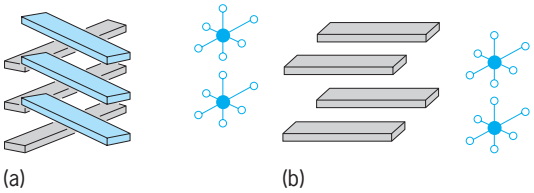

Fig. 2. Stacking of charge-transfer compounds. (*a*) Segregated stacking of TTF-TCNQ-like materials. (*b*) Typical zigzag stacking of the $(TMTSF)_2X$ series.

molecules display a zigzag packing along the stacking axis (Fig. 2*b*), where one positive charge (hole) is shared between two organic molecules.

One-dimensional conduction. The strong overlap between electron clouds of neighboring molecules along the stacks spreads the partially filled molecular electronic states into an energy band 0.5–1 eV wide. This bandwidth is large enough to allow electron delocalization among all molecules on a given stack and to promote electrical conduction similar to that in metal crystals. The resistivity of organic conductors is between 2×10^{-2} and 1×10^{-3} $\Omega \cdot$ cm at ambient temperature and shows a large decrease on cooling, much as with pure metals. A peculiarity of most organic crystals is strong directionality: the conductivity is 100 to 1000 times lower along the directions perpendicular to the packing axis. *See* BAND THEORY OF SOLIDS; DELOCALIZATION; ELECTRICAL CONDUCTIVITY OF METALS.

Charge- and spin-density waves. Unlike ordinary metals, which remain metallic (highly conducting) down to millikelvin temperatures, organic conductors often undergo low-temperature phase transitions toward insulating states. Monochromatic x-ray diffraction of the insulating state of TTF-TCNQ revealed weak satellite diffraction peaks in the vicinity of each Bragg peak related to the crystal structure. These satellites point to the existence of a lattice-distorted Peierls state (charge-density wave) in TTF-TCNQ, following the argument that a uniform conducting one-dimensional electron gas is unstable at low temperature against a lattice distortion. *See* CHARGE-DENSITY WAVE; PHASE TRANSITIONS; X-RAY DIFFRACTION.

In $(TMTSF)_2PF_6$, however, no such satellites of Bragg peaks are found at low temperature. Instead, magnetic data reveal the onset of an internal magnetic modulation (spin-density wave) occurring together with the phase transition at 12 K. In this case, repulsive electrostatic interactions between electrons and low dimensionality are responsible for the existence of the magnetic ground (Overhauser) state. *See* SPIN-DENSITY WAVE.

Superconductivity. The compound $(TMTSF)_2PF_6$ exhibits unusual behavior under pressure. The magnetic and insulating ground state is rapidly suppressed. Above 9 kilobars (0.9 gigapascal), superconductivity accompanied by zero resistance and magnetic flux expulsion (the Meissner effect) is observed below the critical temperature $T_c \approx 1$ K (−458°F). Other investigations of the $(TMTSF)_2X$

series have shown that with the proper anion (in particular ClO_4^-, which is smaller than PF_6^-) superconductivity can be stabilized at ambient pressure below 1.2 K (−457.5°F). Significantly higher critical temperatures have been obtained in the quasi-two-dimensional compound $(BEDT\text{-}TTF)_2Cu(SCN)_2$, where the organic molecules bis(ethylenedithiolo)tetrathiafulvalene (BEDT-TTF, also called the ET molecule) are arranged in two-dimensional conducting sheets sandwiched between insulating copper thiocyanate $[Cu(SCN)_2]$ layers.

High-magnetic-field phenomena. Apart from superconductivity, the response of the electron gas in $(TMTSF)_2ClO_4$ to an applied magnetic field is highly unusual. Under high magnetic fields the conducting state undergoes a transition to a new state above a threshold field that is accompanied by a drastic reduction in the density of carriers from 1 to about 10^{-2} per unit cell. In addition the high-field state exhibits an antiferromagnetic modulation. The interpretation of these phenomena lies in the reinforcement of the one-dimensional character of the quasi-one-dimensional electron gas by the applied magnetic field. Furthermore, the field-induced state reveals a sequence of subphases for which the Hall coefficient remains independent of the magnetic field. This effect clearly resembles the quantum Hall effect in a two-dimensional electron gas. *See* HALL EFFECT.

D. Jérome

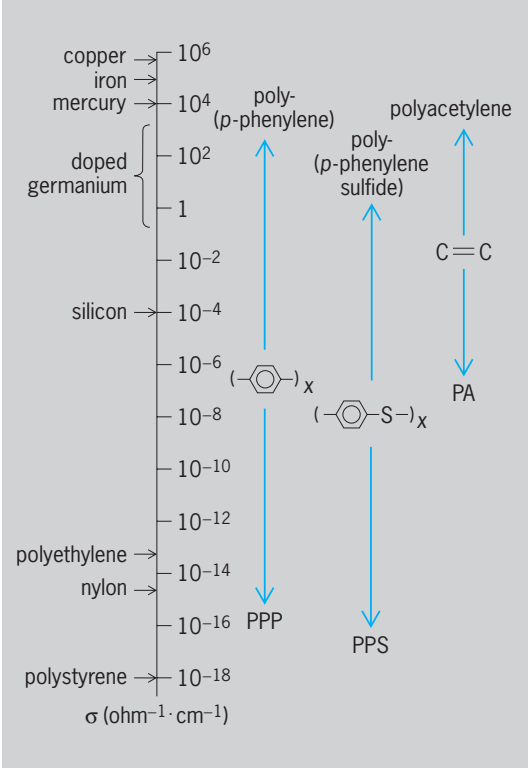

Fig. 3. Comparison of the range of electrical conductivities of some AsF_5-doped conducting polymers with some metals, semiconductors, and nonconducting polymers.

Conducting Polymers

Polymeric materials are typically considered as insulators, and in fact important applications rely on this property. However, research since the late 1970s has led to the discovery of polymeric materials with extremely high conductivity, approaching that of copper. The prospect of materials combining the properties of plastics and metals or semiconductors has led to the pursuit of applications. Improved polymers, particularly polypyrrole and polyaniline (see **table**), no longer suffer from such drawbacks as low stability, processing difficulties, and brittleness. Most conducting polymers can be switched reversibly between conductive and nonconductive states, with the result that their conductivities can span an enormous range (**Fig. 3**). This switching is accomplished through redox chemistry, the conductivity being sensitive to the degree of oxidation of the polymer backbone. This property distinguishes conducting polymers from metals and semiconductors and is the basis of many existing and potential applications. Also, certain polymers become conducting upon oxidation or reduction and thus can exhibit either *p*- or *n*-type conduction. *See* OXIDATION-REDUCTION; POLYMER; SEMICONDUCTOR.

Structure and properties. An oxidant is used to create carriers, and the resulting room-temperature

Examples of dopole polymers, oxidants, dopant ions, and resulting conductivities

Polymer	Formula	Oxidant	Dopant	Conductivity, $(\Omega \cdot cm)^{-1}$
Polyacetylene	$(-CH=CH-)_n$	I_2	Triodide (I_3^-)	10,000
Polypyrrole	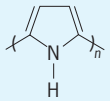	Electrode (anode)	Tetrafluoroborate (BF_4^-)	200
Poly(3-hexyl-thiopene)	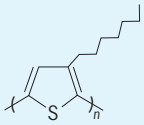	Electrode (anode)	Perchlorate (ClO_4^-)	30
Polyaniline (emeraldine form)	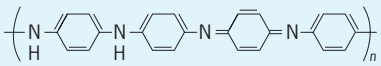	None (protonation leads to internal redox reaction)	Cl^-, CH_3 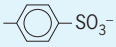 SO_3^-	~200

conductivities, in conducting polymers (see table). The oxidant removes electrons from the π-electron system of the polymer, creating radical cations that, at high concentrations, dimerize to form cation pairs known as bipolarons. Charge-balancing counterions are concomitantly incorporated between polymer chains. The overall process is referred to as doping, and the counteranion (or countercation in the case of reduction) is the dopant. Doping can be carried out by exposing the polymer to vapors or solutions of oxidant or reductant. The polymer can also be oxidized or reduced electrochemically, incorporating dopant ions from the electrolyte. This method is frequently preferred, since dopant ions of precise structure can be introduced and the degree of oxidation or reduction controlled by the electrochemical potential.

Conducting polymers have delocalized π-electrons at least over a few structural units. Such unsaturated polymers favor carrier generation because of the possibility of resonance delocalization of the resulting radical ions; good intramolecular carrier mobility may also result. In addition, the geometry of π-orbitals allows for good orbital overlap and encourages intermolecular carrier transport. The fact that the polymer chains are much shorter than typical sample dimensions indicates that intermolecular transport is dominant, especially in view of the disorder observed in most conducting polymer systems. However, as the number of defects (cross-links or twists that inhibit conjugation) decreases, it might be anticipated that conductivity would be higher because carriers can travel greater distances along a chain before an intermolecular electron transfer of higher activation energy becomes necessary. Indeed, certain stretch-oriented polymers afford significantly greater conductivities than their unoriented counterparts. *See* CONJUGATION AND HYPERCONJUGATION.

Considerable progress has been made on the theory of important parameters such as oxidation potentials, band gaps, and band widths, often with good agreement with experiment. Such work is important for the design of new conductive polymers with specific properties. The elucidation of conduction mechanisms has also received much attention. To a first approximation, carrier transport may be viewed as a combination of intrachain resonance delocalization processes and interchain redox processes involving electron transfer between segments of neutral and oxidized (or reduced) polymer. The interchain processes are reminiscent of intrastack conduction in small-molecule organic radical ion salts (discussed above). *See* ELECTRON-TRANSFER REACTION.

Processibility. Conducting polymers are frequently intractable to processing, although advances in this area have alleviated this problem. For example, alkyl-substituted polythiophenes, unlike the parent polythiophene, are soluble in common solvents and yet have nearly comparable conductivities when doped. Polyacetylene and poly(*p*-phenylene vinylene) can be fabricated into films and fibers from soluble percursor polymers. Doped poly(*p*-phenylene sulfide) can be solubilized if the doping is carried out in solvents such as arsenic fluoride (AsF_3). Many approaches are available for direct synthesis of conducting polymers as films. *See* FILM (CHEMISTRY).

Melt-blending of conducting polymer powders with thermoplastic polymers offers the opportunity to combine conductivity of the conducting polymer filler with desirable mechanical properties of the thermoplastic matrix. Blends of doped polyaniline and poly(vinyl chloride) are available (with the dopants for polyaniline, such as *p* toluenesulfonate, selected to afford maximum thermal stability).

Applications. The utility of most conducting polymers lies in their unique properties or functions, rather than replacing existing materials. Of particular interest is the fact that many of these materials can be oxidized and reduced reversibly, with concomitant changes in conductivity, optical absorption, and wettability.

Conductive polymers appear to be promising materials as battery cathodes or anodes, and may even function as both anode and cathode in a cell (the all-polymer battery). Electrochemical doping of polyacetylene at an anode or cathode is a charging process. Placement of a load across the polymer electrodes results in discharge as the doped polymers neutralize each other, the oxidized polymer (anode) becoming reduced and vice versa. The high surface areas of many conducting polymer films and powders provide for useful power densities. *See* BATTERY.

Other applications include corrosion-inhibiting coatings on steel, biosensors, and matrices for the timed release of chemicals. In the last case, advantage is taken of the fact that the counteranions in *p*-type (oxidized) conducting polymers are released into solution when the polymer is reduced electrochemically. However, it may be desirable for cations rather than anions to be released. This can be done with a polypyrrole–poly(styrene sufonate) composite. The appropriate cation is bound initially to the poly(styrene sulfonate), although upon electrochemical oxidation of polypyrrole the sulfonate moieties of the poly(styrene sulfonate) become the counteranions for the conductive polymer, releasing their previously bound cations into solution. Cations can also be bound with self-doped polyheterocyclic compounds containing pendant sulfonates. The fact that neutral (undoped) polymers are typically hydrophobic whereas their doped forms are polar salts has led to the construction of ion-gate membranes with variable ion permeabilities, depending on the electrochemically controllable oxidation state of the polymer. Doping also changes optical absorption properties, and thus many conducting polymers exhibit electrochromism with possible utility in display technology and so-called smart windows. Microelectrochemical transistors are also fabricated from conducting polymers by using techniques common in the semiconductor industry. *See* COMPOSITE

MATERIAL; ELECTROCHEMISTRY; ELECTROCHROMIC DEVICES; TRANSISTOR.

Poly(phenylene vinylene) and related conjugated polymers (but without doping) exhibit electroluminescence, wherein carriers of opposite sign generated at attached electrodes recombine in the polymer with emission of light. The color of the light can be controlled by tailoring the polymer structure, opening the prospect of flexible, color-tuned displays. *See* ELECTROLUMINESCENCE; SOLID-STATE CHEMISTRY; SOLID-STATE PHYSICS. Gary E. Wnek

Bibliography. P. Chandrasekhar, *Conducting Polymers, Fundamentals and Applications: A Practical Approach*, 1999; J.-P. Farges (ed.), *Organic Conductors, Fundamentals and Applications*, 1994; R. B. Kaner and A. G. MacDiarmid, Plastics that conduct electricity, *Sci. Amer.*, 258(2):106–108, February 1988; V. Z. Kresin and W. A. Little (eds.), *Organic Superconductivity*, 1990; L. Ouahab and E. Yagubskii (eds.), *Organic Conductors, Superconductors and Magnets: From Synthesis to Molecular Electronics*, 2004; T. A. Skotheim, R. L. Elsenbaumer, and J. R. Reynolds (eds.), *Handbook of Conducting Polymers*, 2d ed., 1997.

Organic evolution

Organic, or biological, evolution is the modification of living organisms during their descent, generation by generation, from common ancestors. It is to be distinguished from other phenomena to which the term evolution is often applied, such as chemical evolution, cultural evolution, or the origin of life from nonliving matter. Organic evolution includes two major processes: anagenesis, the alteration of the genetic properties of a single lineage over time; and cladogenesis, or branching, whereby a single lineage splits into two or more distinct lineages that continue to change anagenetically.

Study of evolution. The subject matter of evolutionary biology may be roughly divided into the analysis of the history of evolutionary events and the analysis of the mechanisms, or processes, of evolutionary change. The study of evolutionary history attempts to determine the ancestry of and genealogical relationships among different kinds of organisms, the pathways by which their morphological, biochemical, and other features have become modified, the history by which they arrived at their present geographical distributions, and the changes in the diversity and number of species throughout geological time. The methods by which such inferences are made include analysis of the fossil record and the phylogenetic analysis of living taxa, many having an inadequate fossil record. Phylogenetic analysis, using data on the comparative anatomy, molecular characteristics [for example, protein and deoxyribonucleic acid (DNA) sequences], and geographical distributions of organisms, is part of the province of biological systematics.

The analysis of the mechanisms of evolutionary change addresses primarily the factors that cause changes in the genetic composition of populations and species, and those that influence diversification and extinction of species. The mathematical theory of population genetics is important to this enterprise. Experimental and observational testing of the theory includes molecular, genetic, and developmental analysis of genetic variation and the mechanisms by which it arises; ecological genetics, the study of the impact of ecological factors on genetic change of populations; studies in functional morphology, physiology, behavior, and ecology that address the adaptive value of genetically different traits; and taxonomic and phylogenetic analyses that shed light on processes such as cladogenesis. Thus the study of evolution embraces all of biology.

History. Although some ancient Greek philosophers had vague, often mythological, intimations of evolution, Platonic and Aristotelian philosophy, in which variation represented imperfect reflection of eternal, unchanging essences or "ideas," was antithetical to evolution. The adoption of this framework by Christian theology, and the literal interpretation of the first chapters of Genesis, led to the belief that all living things had been directly created in their current form (special creation) only a few thousand years ago. The first challenges to this view did not arise until the eighteenth century, when speculations on cosmic change, on the antiquity and dynamic nature of the Earth, and on human progress led naturally to the idea that living things might change as well. The French biologist G. de Buffon (1707–1788) was one of the first to hint at evolution, but his countryman J. de Lamarck (1744–1829) was the first to argue forcefully for evolution and to propose a mechanism by which it might occur. Lamarck supposed that new forms arise continually by spontaneous generation, and then progress toward greater complexity and perfection because of "powers conferred by the supreme author of all things" and because their behavioral responses to the environment cause changes in their structure. Lamarckism was soon discredited by the implausibility and vagueness of the mechanisms he postulated, and by forceful arguments against evolution by leading French biologists of the day. Nevertheless, the idea of evolutionary change was "in the air" in the early nineteenth century.

Charles Darwin (1809–1882), son of an English physician, apparently came to think of the possibility of evolution toward the end of his 5-year (1831–1836) voyage as naturalist on the H.M.S. *Beagle*. He conceived of the theory of natural selection in 1838, and spent the next 20 years synthesizing and amassing evidence and refining his ideas until, faced with the possibility that the young naturalist Alfred Russel Wallace, who had independently thought of natural selection, might gain priority for the idea, he published an "abstract" of the massive book which he had been preparing. The "abstract" was a 490-page book, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, published on November 24, 1859. Darwin spent the rest of his life conducting

research and writing books on an extraordinary range of subjects, from the power of movement in plants to the "descent of man," all of which are related directly or indirectly to the themes in his most famous book. But *The Origin of Species*, as it is generally known, is his triumph, one of the most important works in the history of western civilization.

The Origin of Species accomplishes two things. Darwin marshals evidence from every quarter of biology and geology that evolution has in fact occurred: that living things are descended with modification from common ancestors. Second, he presents an explicit, purely mechanistic theory of the causes of evolution. Every species, Darwin points out, has hereditary variation in numerous characteristics. Some variants will be better suited to the exigencies of life than others, and so will survive better or reproduce more prolifically than the inferior variants. Since descendants inherit their superior properties, the proportion of individuals in the population that bear superior characteristics will increase, and the proportion with inferior traits will decrease from generation to generation, until the species has been transformed. The new character itself is subject to further variation and to further alteration by this process of natural selection, so that in the vastness of time the feature comes to differ extremely from the original form—but it is a great change accomplished in small steps. Because different populations experience different environments and adapt to different resources, numerous forms may diverge from an original stock, each adapted to a different environment or way of life. This branching process, continued through the immensity of geological time, gives rise to the great "tree of life."

Darwin's ideas at first met with much opposition, but within 20 years after the *Origin* appeared, most scientists had been convinced of the reality of evolution, and many were stimulated to work in the areas of paleontology, comparative anatomy, and comparative embryology that provide evidence of the historical (phylogenetic) relationships among organisms. But Darwin's theory of the cause of evolution, natural selection, was not widely accepted for lack of sufficient evidence. It fell even deeper into disrepute in the early twentieth century, when the new science of genetics (developing after the rediscovery in 1900 of Mendel's work) seemed to provide alternative mechanisms for evolution, such as mutation.

Modern evolutionary theory began in the late 1920s and early 1930s, when naturalists and systematists amassed evidence on adaptive variation and the nature of species, and when the mathematical geneticists R. A. Fisher and J. B. S. Haldane in England and S. Wright in the United States developed equations showing that natural selection and mendelian genetics (including mutation) are not alternatives, but rather work in combination just as Darwin had proposed. Their theories, which include other causes of evolution in addition to the coaction of natural selection and mutation, are the foundation of mathematical population genetics. *See* POPULATION GENETICS.

During the 1930s through the early 1950s, the theory of population genetics, together with the ideas of experimental workers and taxonomists concerned with evolution, congealed into the "modern synthesis" commonly known as neo-Darwinism. The experimental geneticists S. S. Chetverikov, Th. Dobzhansky, and E. B. Ford, the zoologists E. Mayr, B. Rensch, and J. S. Huxley, the paleontologist G. G. Simpson, and the botanist G. L. Stebbins, were among the chief figures who, together with Wright, Fisher, and Haldane, formulated a coherent, comprehensive theory of evolution and showed that it was consistent with data from diverse fields of study. Since the modern synthesis, new data and theories concerning especially evolution at the molecular level, evolution of behavior and of ecological interactions among species, and long-term patterns of evolution over geological time have been added to the neo-Darwinian framework, and some neo-Darwinian ideas have been called into question. Thus, although most of the neo-Darwinian theory is accepted by most evolutionary biologists, there remain controversies, as in any other field of science, about some aspects of the theory. There is no controversy, however, about the reality of evolution as a historical event. That organisms have descended from common ancestors is accepted by knowledgeable biologists as fact. Molecular and other similarities imply that all living things are related to each other by common ancestry.

Mechanisms of species transformation. Anagenesis consists of change in the genetic basis of the features of the organisms that constitute a single species. Populations in different geographic localities are commonly considered members of the same species if they exchange members at some rate and interbreed with each other (or are thought to be potentially able to interbreed); but unless the level of interchange (gene flow) is very high, some degree of genetic difference among different populations is likely to develop. The changes that transpire in a single population may be spread to other populations of the species by gene flow. *See* SPECIES CONCEPT.

Genetic variation. Almost every population harbors several different alleles (forms of a gene) at each of a great many of the gene loci; hence many characteristics of a species are genetically variable. This is evident by the existence of morphological and physiological variations that are shown to be inheritable, and by recent study of genes themselves (deoxyribonucleic acid, or DNA, sequences) and of their products (proteins). Studies of protein variation on *Drosophila*, humans, and numerous other species suggest that 40% or more of the gene loci may be variable within a population, and analyses of DNA suggest that almost all of a species' genes vary. Thus, to evolve in response to many environmental changes, a population need not wait for new mutations to occur; genetic variations that happen to be suitable for new conditions may already be present. For example, genes for insecticide resistance are present in low frequency in populations of insects that have never been exposed to insecticides.

Sources of genetic variation. All genetic variations ultimately arise by mutation of the genetic material. Broadly defined, mutations include changes in the number or structure of the chromosomes and changes in individual genes, including substitutions of individual nucleotide pairs, insertion and deletion of nucleotides, and duplication of genes. Many such mutations alter the properties of the gene products (RNA and proteins) or the timing or tissue localization of gene action, and consequently affect various aspects of the phenotype (that is, the morphological and physiological characteristics of an organism). Whether and how a mutation is phenotypically expressed often depends on developmental (epigenetic) events, some of which may "canalize" the phenotype to reduce phenotypic variability even if genetic variation exists. The phenotypic effect of mutations ranges from extremely subtle to drastic; many drastic mutations alter development so greatly that the organism's viability is severely reduced. However, some large mutations, as well as more subtle mutations which slightly change body size, the form of an appendage, the activity of an enzyme, or other features, are deleterious, neutral, or beneficial, depending on the organism's environment. Because a gene typically affects several aspects of the phenotype, it may have both harmful and beneficial effects (pleiotropy), the relative magnitude of which determines the net effect of the mutation on survival and reproduction. Because different genes interact in determining phenotypes, the positive or negative value of a mutation may depend on which alleles are prevalent at other gene loci.

The frequency with which a given gene mutates to a recognizably different allele is fairly low (often about 1 mutation per gene per 100,000 gametes), but can be considerably higher when mutation is caused by factors such as transposable segments of DNA that become inserted into a gene and affect its function. Although the mutation rate of any given gene is low, the total flux of mutations is appreciable (at least one new mutation somewhere in the genome of each gamete) because the number of genes is high. Mutation occurs spontaneously because of molecular "noise," but can be enhanced by factors such as heat, radiation, and chemical mutagens. *See* MUTATION.

Mutations are generally believed to be occur at random, not in the sense that all genes mutate at the same rate or that all possible mutations of a gene are equally likely, but rather in the sense that the likelihood of a particular mutation is not influenced by whether it would be advantageous to the organism in its prevailing environment. Genes cannot respond to the environment by mutating in an appropriate direction; nor can the experiences of an individual organism enable it to change in an adaptive manner the genetic basis for the characteristics that the organism uses in its relations to the environment. This is the principal point of difference between the neo-Darwinian and the Lamarckian theories of evolution.

Recombination, arising from sexual reproduction and crossing-over, combines the mutant variants at various loci into a vast number of possible combinations; thus mutation and recombination together generate immense genetic variability among the members of a population. When, as is usually the case, many different loci each contribute to a particular trait such as body size, the result is "continuous" variation, often in the form of a bell-shaped (normal) distribution of the trait in the population. *See* CROSSING-OVER (GENETICS); DISTRIBUTION (PROBABILITY).

Natural selection. The fundamental event in evolution is a change in the frequency of an allele in a population. In its full form, this entails the spread through a population of an allele that, having just come into existence by mutation, is very rare, but which ultimately comprises 100% of the gene copies at that locus (in a population of N diploid organisms, there are $2N$ gene copies at a locus that is not sex-linked). The allele is then said to have been fixed in the population. One of the factors that causes this process is natural selection.

Natural selection is a consistent difference in the average rate at which genetically different entities leave descendants to subsequent generations; such a difference arises from differences in fitness (that is, in the rate of survival, reproduction, or both). In fact, a good approximate measure of the strength of natural selection is the difference between two such entities in their per capita rate of increase r. The entities referred to are usually different alleles at a locus, or phenotypically different classes of individuals in the population that differ in genotype. Thus selection may occur at the level of the gene, as in the phenomenon of meiotic drive, whereby one allele predominates among the gametes produced by a heterozygote. Selection at the level of the individual organism, the more usual case, entails a difference in the survival and reproductive success of phenotypes that may differ (in body size, for instance) at one locus (for example, genotypes *AA* and *aa* differ in size) or at more than one locus (for example, *AABBCC* versus *aabbcc*, where different letters denote different loci and upper- and lowercase letters denote different alleles at a locus). As a consequence of the difference in fitness, the proportion of one or the other allele increases in subsequent generations (**Fig. 1**). Numerous cases of such differences in fitness, often of very considerable magnitude, have been documented in both laboratory-maintained and natural populations of many species. *See* POPULATION ECOLOGY.

The relative fitness of different genotypes usually depends on environmental conditions. Thus, for example, brown shell color in terrestrial snails is advantageous in forests, providing cryptic protection against predators, whereas yellow coloration is advantageous in open fields. Relative fitnesses of genotypes often switch as the environment changes from season to season or year to year, and the frequencies of genotypes often fluctuate in consequence. Thus changes in genetic composition are the passive consequence of changes in environment; a population cannot alter its genetic composition in anticipation of some future environmental change. This implies that populations cannot adapt so as to avoid future

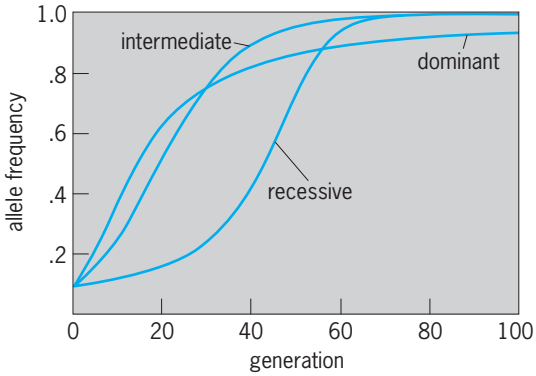

Fig. 1. Increase in the frequency of an advantageous allele within a population, from an initial frequency of 0.1. The fitness, or relative rate of increase, of the fittest genotype is 20% greater than that of the least fit genotype in each of three cases, in which the fittest genotype is dominant, recessive, or partially dominant (intermediate). (*After D. J. Futuyma, Evolutionary Biology, Sinauer Associates, 1979*)

extinction. Natural selection does not act for the benefit of the species as a whole; it is a purely mechanical, mathematical phenomenon: the difference in reproductive success among individual members of the species.

Selection is sometimes directional, meaning that one extreme state of a phenotypic feature is most fit and ultimately becomes fixed if the environmental conditions to which it is adapted prevail long enough (**Fig. 2**). Often, however, selection is stabilizing, meaning that an intermediate phenotype is most fit. If this phenotype is produced by a heterozygote at a single locus, selection maintains both alleles in the population. Several cases of stabilizing selection are known in which the heterozygote is most fit. Genetic variation can also be maintained by diversifying selection, whereby several different phenotypes are favored. Diversifying (disruptive) selection includes the phenomenon of frequency-dependent selection: the rarer a genotype is, the higher its fitness becomes. This phenomenon often arises because of interactions among the members of a population, each of which constitutes part of the environment of every other member. For example, genotypes often differ in their relative ability to use one or another resource, such as different kinds of food. If there is competition for limited food, a rare genotype experiences less competition for its particular food type than the common genotype does for its particular food type; thus the per capita rate of increase of the rarer genotype is relatively greater. When it attains high frequency, the tables are turned, so an equilibrium is attained at which both genotypes persist.

Under directional selection, new mutations that enhance a particular trait may be consistently advantageous over the long run. For example, all other things being equal, greater height is advantageous for a tree because shorter trees are overtopped by their taller neighbors, and so do not receive as much light. But there is a limit to the height that natural selection favors, because countervailing selection pressures come into play: an excessively tall tree will be toppled by wind unless it has a sufficiently strong trunk and root system. Because wind conditions differ from one place to another, and the trunk and root structure differs among species, the relative strength of conflicting selection pressures differs among populations and species, which therefore evolve along divergent lines. In addition to divergent evolution, however, evolution is often convergent: different organisms sometimes experience similar selection pressures and so evolve at least superficially similar characteristics.

Random genetic drift. Different alleles of a gene that provides an important function do not necessarily differ in their effect on survival and reproduction; such alleles are said to be neutral. The proportion of two neutral alleles in a population fluctuates randomly from generation to generation by chance, because not all individuals in the population have the same number of surviving offspring. Random fluctuations of this kind are termed random genetic drift. If the process continues long enough in the absence of countervailing factors, one or another allele will ultimately fluctuate all the way to fixation, and the other alleles will be lost from the population by chance (**Fig. 3**). This process occurs more rapidly, the smaller the population. Many natural populations are quite small; in fact, their effective size, which can be thought of as the number of individuals that actually succeed in reproducing, is considerably smaller than the total number of individuals. Thus all populations are susceptible to genetic drift, and the process

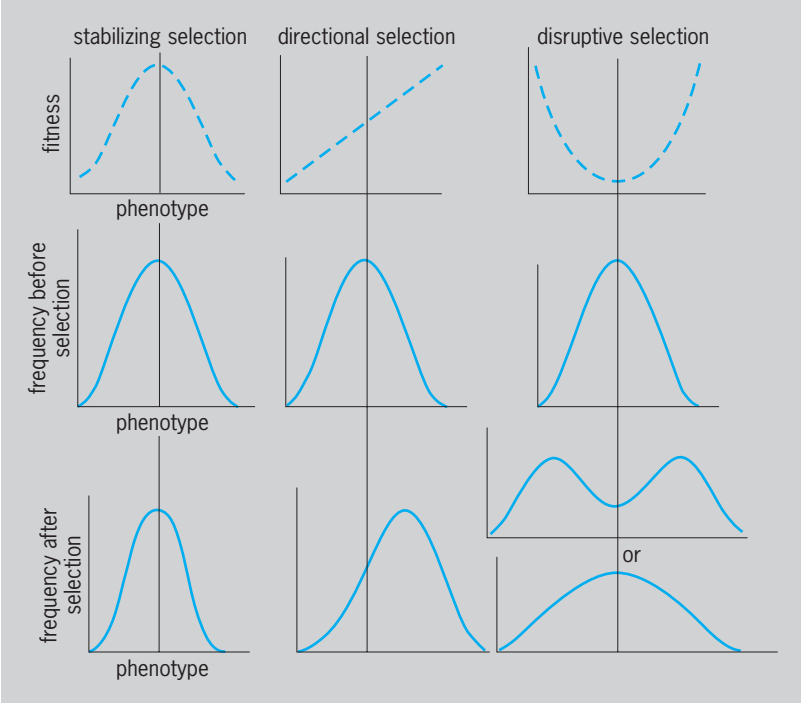

Fig. 2. Three modes of selection on a quantitatively varying phenotypic character. Because of the relation between fitness and phenotype, portrayed in the upper panels, the frequency distribution of the trait in the population changes during one generation of selection from the patterns shown in the middle to those shown in the lower panels. (*After L. L. Cavalli-Sforza and W. F. Bodmer, The Genetics of Human Populations, W. H. Freeman, 1971*)

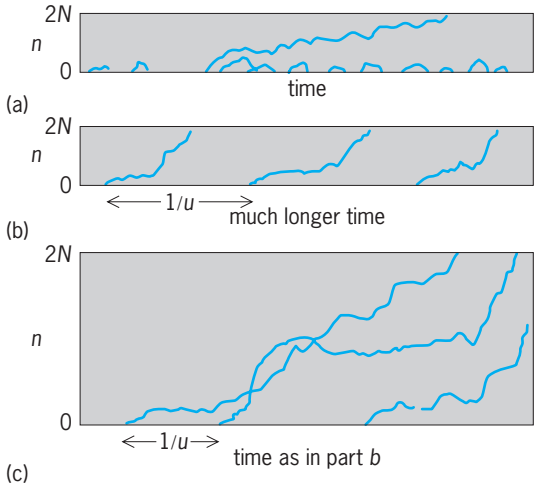

Fig. 3. Gene substitution by genetic drift. A population of *N* diploid individuals has 2*N* genes at any locus. The number of copies *n* of a mutation fluctuates by chance. (*a*) Numerous mutations arise, of which one is fixed while others increase and then decrease in frequency. (*b*, *c*) Only those mutations at various loci which are fixed are shown. The average number of generations between the origin of successive mutations that are ultimately fixed by genetic drift is 1/*u*, where *u* is the mutation rate. In a large population, shown in *c*, several loci are polymorphic at any time, as it takes longer for mutations to reach fixation. (*After J. F. Crow and M. Kimura, An Introduction to Population Genetics Theory, Burgess, 1970*)

is likely to proceed quite rapidly in many species. Because the identity of the fixed allele is only a matter of chance, the allele that becomes fixed will differ from one population to another, so that the genetic composition of different populations or of different species diverges over time, at a rate inversely proportional to their effective population size. The theory of genetic drift has been validated by laboratory experiments and by patterns of genetic difference among natural populations of species.

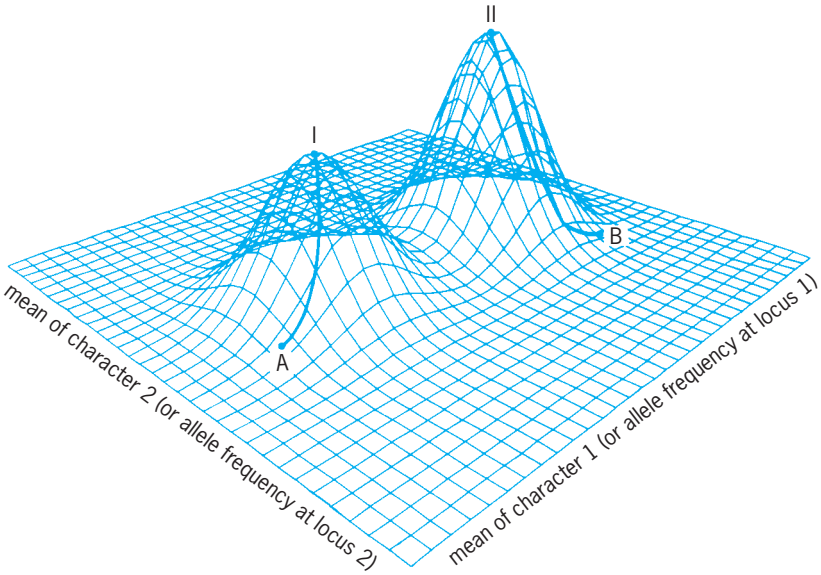

Fig. 4. Natural selection may drive a population to different genetic equilibria depending on initial allele frequencies. The elevation of a point on this landscape represents the mean fitness of a population that has a particular average value for each of two traits (or that has a particular gene frequency at each of two loci). Character combinations I or II represent high fitness, whereas intermediate phenotypes have low fitness. A population beginning with average phenotypes represented by point B will evolve to equilibrium II, whereas one beginning at point A will become stabilized by natural selection at equilibrium I, even though this represents an inferior condition relative to II. (*After D. J. Futuyma and M. Slatkin, Coevolution, Sinauer Associates, 1983*)

Drift versus selection. If different alleles do indeed differ in their effects on fitness, both genetic drift and natural selection operate simultaneously. The deterministic force of natural selection drives allele frequencies toward an equilibrium, while the stochastic (random) force of genetic drift brings them away from that equilibrium. The outcome for any given population depends on the relative strength of natural selection (the magnitude of differences in fitness) and of genetic drift (which depends on population size), just as the trajectory of a dust particle in still air depends on the relative power of gravitation and of brownian motion. Thus the fate of an allele that differs in fitness only slightly from other alleles may be dominated by selection if the population is large, but by genetic drift if the population is small.

The relative importance of genetic drift versus natural selection is the subject of considerable controversy and research, the chief focus of which is on genetic differences at the molecular level, such as different forms of an enzyme or slight differences in nucleotide sequences in DNA. There is reason to believe that many such differences have only a slight impact on fitness and that considerable evolution at the molecular level occurs primarily by genetic drift. The best evidence is provided by synonymous codons of DNA and by the nucleotide sequences of DNA that are not transcribed into RNA and protein. Many such changes at the DNA level are unlikely to affect the organism's fitness, yet comparison among species shows that these sequences have evolved at a high rate—higher than that of DNA sequences that are functional. Fixation of purely neutral alleles, according to mathematical theory, should occur at a constant rate per generation, and there is some evidence that the rate at which species have diverged at the molecular level has been moderately constant. Thus some evolutionary change has certainly transpired by genetic drift, although divergence in clearly adaptive features of morphology, physiology, and behavior has certainly come about primarily by natural selection.

In theory, genetic drift can act together with natural selection to enhance adaptation of a species to its environment. Often the pressure of natural selection can drive the population toward any one of several different genetic constitutions: alternative genetic "solutions" to the same "problem" (**Fig. 4**). Which genetic constitution is achieved depends on which alleles happen to be most prevalent when the environmental challenge arises. The genetic "solution" actually achieved may be inferior to another genetic constitution that might have been achieved had the initial genetic composition of the population been different; yet the equations of gene frequency change show that natural selection alone cannot move the population from one stable genetic equilibrium to a different, superior, genetic equilibrium. If the population is small enough, however,

genetic drift can destabilize the population, carrying the gene frequencies away from the inferior equilibrium enough for natural selection to bring them to a new, superior, genetic configuration. This theory is part of Sewall Wright's shifting-balance theory of genetic change.

Speciation and cladogenesis. The great diversity of organisms has come about because individual lineages (species) branch into separate species, which continue to diverge by the processes described above. This splitting process, speciation, occurs when genetic differences develop between two populations that prevent them from interbreeding and forming a common gene pool. The genetically based characteristics that cause such reproductive isolation are usually termed isolating mechanisms, but there is little reason to believe that they evolve specifically to prevent interbreeding, as the unfortunate term mechanism implies. Rather, reproductive isolation seems to develop usually as a fortuitous by-product of genetic divergence that occurs for other reasons (either by natural selection or by genetic drift).

The most common mode of speciation is undoubtedly genetic divergence among populations that are sufficiently spatially isolated that their gene pools are not homogenized by gene flow. This allopatric mode of speciation may occur when two widespread populations are separated by unsuitable habitat (for example, European and American populations), but is probably more frequent and more rapid when a population in a restricted locality is cut off (for example, by colonization across a habitat barrier) from the main body of the species, and undergoes rapid divergence because of genetic drift and different selection pressures. If sufficient genetic divergence transpires before these populations expand and encounter each other, they will not exchange genes when they meet; if divergence has been insufficient, they interbreed and speciation has not been completed.

This genetic theory of speciation is well supported by evidence from taxonomists' analyses of the relationships among populations, which run the gamut from slight genetic differentiation to complete reproductive isolation. The genetic divergence that provides reproductive isolation sometimes consists of changes in chromosome structure that cause infertility of hybrids, or of genic alterations of development, especially in gametogenesis, that cause hybrids to be inviable or infertile. Hybrid inviability and infertility constitute postmating isolating mechanisms. Many species, however, are capable of forming fertile, viable hybrids, but do not do so in nature because they do not mate with each other; differences in time of reproduction, courtship behavior, or (in plants) flower form are among the premating isolating mechanisms that maintain the distinction between the species. Such differences can arise as pleiotropic by-products of the genetic changes that developed by genetic drift or as responses to different environmental factors. In many groups of animals, courtship signals and responses can diverge rapidly because of sexual selection, in which a characteristic (such as a peacock's train) becomes exaggerated to enable its bearer (often the male) to compete more successfully for mates. Analogous phenomena may occur in some plants, such as orchids. *See* REPRODUCTIVE BEHAVIOR.

At least in plants, speciation can occur sympatrically, without initial geographic isolation, by polyploidy. Under special circumstances, sympatric speciation can theoretically also occur in animals, but the frequency with which this occurs is strongly debated. Ernst Mayr has developed cogent arguments for supposing that it is rare. *See* SPECIATION.

Adaptations. A frequent consequence of natural selection is that a species comes to be dominated by individuals whose features equip them better for the environment or way of life of the species. Such features are termed adaptations. Thus in England in the nineteenth century, dark variants of the peppered moth (*Biston betularia*) replaced the light genotype because they were less evident to predatory birds; their dark coloration may be considered an adaptation to predation. It is sometimes difficult to determine the adaptive significance of a trait, or whether it is an adaptation at all. However, this can often be accomplished by experiment or observation. For example, it has been shown that the color patterns of certain butterflies that resemble distasteful species provide protection against predation, by experimentally altering their color patterns and finding that the altered individuals suffer greater predation. Convergent evolution is often evidence of adaptation: for example, leaflessness has evolved in many groups of desert plants as a mechanism of reducing water loss.

Most adaptive features benefit the individuals that bear them, rather than the population as a whole. Certain features, though, appear at first sight to be altruistic, such as the warning calls that some birds emit when they see a predator. Other members of the flock benefit from being alerted, but the call seems likely to place the warner in jeopardy, so that alleles for such behavior would appear unlikely to increase in the population. However, features like this can evolve because selection can act not only at the level of individual organisms but at other levels. One possibility, the prevalence of which is strongly disputed, is group selection, whereby the genetic composition of the species changes by the differential survival (or proliferation) of whole populations that differ in gene frequency. Thus populations that happen to have a high frequency of genes for warning behavior might survive best, so that these genes increase in the species as a whole, even if they are selected against within populations. A more likely explanation for apparently altruistic traits is a form of selection at the level of the gene known as kin selection. In developing this theory, William Hamilton argued that because relatives share alleles by common descent, an allele may increase in frequency if it causes its bearer to help its relatives survive and reproduce, even if the fitness of the bearer suffers. Kin selection, the cornerstone of the study of sociobiology, is the most likely explanation for the sterility of workers in social insects, and for warning calls and

many other aspects of social behavior. *See* SOCIOBIOLOGY.

Although many features of organisms are adaptive, not all are, and it is a serious error to suppose that species are capable of attaining ideal states of adaptation. Some characteristics are likely to have developed by genetic drift rather than natural selection, and so are not adaptations; others are side effects of adaptive features, which exist because of pleiotropy or developmental correlations. The absence of appropriate genetic variations and the constraints imposed by processes of development limit the variety of adaptive responses of which a species is capable, so that the path of ideal adaptation may be closed to it. The phylogenetic history of a species determines its current state, and thus the kinds of variations that may be available to natural selection. Most species are not capable of adapting ideally to all environmental changes: more than 99% of all the species that have ever lived are extinct.

Origin of higher taxa. Higher taxa are those above the species level, such as genera and families. A taxon such as a genus is typically a group of species, derived from a common ancestor, that share one or more features so distinctive that they merit recognition as a separate taxon. The degree of difference necessary for such recognition, however, is entirely arbitrary: there are often no sharp limits between related genera, families, or other higher taxa, and very often the diagnostic character exists in graded steps among a group of species that may be arbitrarily divided into different higher taxa. Moreover, a character that in some groups is used to distinguish higher taxa (such as the number of cotyledons, one of the features that divides flowering plants into Monocotyledonae and Dicotyledonae) sometimes varies among closely related species or even within species (as does the number of cotyledons in a few flowering plants). In addition, the fossil record of many groups shows that a trait that takes on very different forms in two living taxa has developed by intermediate steps along divergent lines from their common ancestor; thus the inner ear bones of mammals may be traced to jaw elements in reptiles that in turn are homologous to gill arch elements in Paleozoic fishes. Thus evolution of differences of great enough magnitude to define higher taxa appears to proceed usually by gradual incremental changes corresponding to the variations evident within species and among closely related species. However, it is possible that in some instances slight differences at the genetic level have altered developmental patterns so as to yield large discrete changes in a suite of developmentally correlated traits. Neoteny in salamanders, for example, whereby changes in hormone levels cause the retention of a complex of larval characteristics into the reproductive ages, appears to have a simple genetic (or in some cases a purely environmental) basis. The frequency with which such discontinuous, or saltational, changes in phenotype have occurred in evolution is a matter of some controversy. *See* ANIMAL SYSTEMATICS; MACROEVOLUTION; NEOTENY.

Among the primary reasons for changes in the form of a structure is a change in its function. For example, the sting of a wasp is a morphologically and functionally altered ovipositor, the structure used by females of more primitive forms to insert eggs into the plants or animals in which the larvae develop. (Incidentally, this explains why male wasps and bees lack stings.) In many desert plants such as cacti, leaves or branches have taken on a defensive function and are modified into spines.

To a certain degree, simple evolutionary changes in structure or physiology can be reversed during evolution; but complex features, once lost or highly modified, are seldom regained in their original form, for natural selection and genetic drift can act only on the variations of whatever "raw materials" are available. Thus numerous groups of plants, especially those that are wind-pollinated, have lost their petals; in some of these groups, insect pollination has secondarily evolved, but the structures that are colored and otherwise modified to attract insects are leaves, sepals, or other parts, the petals having been lost. *See* PHYLOGENY.

Rates of evolution. The characteristics of a species evolve individually or in concert with certain other traits that are developmentally or functionally correlated. Because of this mosaic pattern of evolution, it is meaningful to speak of the rate of evolution of characters, but not of species or lineages as total entities. Thus in some lineages, such as the so-called living fossils, many aspects of morphology have evolved slowly since the groups first came into existence, but evolution of their DNA and amino acid sequences has proceeded at much the same rate as in other lineages. Every species, including the living fossils, is a mixture of traits that have changed little since the species' remote ancestors, and traits that have undergone some evolutionary change in the recent past.

Rates of evolution are highly variable: whereas many characteristics have changed hardly at all for many millions of years in some lineages, others have responded rapidly to changes in selection pressures. For example, within decades, numerous species of insects have evolved resistance to insecticides, and some plants have become adapted to soils impregnated with toxic metals from mine works. Geographical variation in morphological traits has evolved within a century in house sparrows introduced into North America from Europe. Based on the geological history of their habitats, speciation in some groups of fresh-water fishes is believed to have transpired in less than 10,000 years, whereas in some groups of plants populations isolated for several million years have not become different species. In contrast to the highly variable rate of evolution of morphological and physiological characters, the rates at which nucleotide sequences of DNA and amino acid sequences of proteins have changed appears to be considerably more uniform; Motoo Kimura, Allan Wilson, and some other authors have in fact claimed that the rate of evolution at the molecular level is nearly constant, providing a "molecular clock" that may be used to estimate the time since species diverged from their common ancestors.

The fossil record of certain organisms, especially marine invertebrates, seems to indicate that species often change rather abruptly in morphology after being virtually static for up to a million years; the geological record, however, is usually so coarse that the "abrupt" changes may well have proceeded gradually over a period of many thousands of years. This pattern, of stasis punctuated by brief periods of change, has been termed punctuated equilibrium. Niles Eldredge, Stephen Gould, and Steven Stanley have suggested that the observed changes represent the origin of new species in localized populations, which become evident in the fossil record only after they expand from their site of origin. They also propose that after speciation the species becomes incapable of substantial further change, so that evolutionary change is virtually restricted to divergence during speciation. There is little direct evidence for this notion of genetic paralysis, which is contested by most evolutionary geneticists. *See* MACROEVOLUTION.

There is abundant evidence that rates of evolution are greatest when a lineage adapts to new ecological opportunities—to vacant ecological niches. Rapid divergent evolution is common, for example, when species colonize islands that harbor few competitors; similarly, the rate of evolution is high in lineages that have survived mass extinction events. The usual pattern in such instances is one of adaptive radiation: the origin, by speciation, of numerous descendant lineages that become adapted in different ways to a variety of available resources (**Fig. 5**). A famous example is the radiation of Darwin's finches in the Galápagos Islands, where related species have diverged in beak morphology and have become specialized for feeding on resources that on continents are typically preoccupied by unrelated families of birds. The major adaptive radiation of mammals occurred soon after the demise of the last dinosaurs, leading many authors to suspect that the mammalian radiation was possible only because competition had been alleviated. Thus extinction has played an important role in the history of life, making possible the subsequent diversification of groups that otherwise might not have flourished.

Evolutionary trends. The history of life is not one of progress in any one direction, but of adaptive radiation on a grand scale: the descendants of any one lineage diverge as they adapt to different resources, habitats, or ways of life, acquiring their own specialized features as they do so. There is no evidence that evolution has any goal, nor does the mechanistic theory of evolutionary processes admit of any way in which genetic change can have a goal or be directed toward the future.

Nevertheless, in examining the history of a major group of organisms, it is sometimes possible to discern in retrospect a trend in one or more characters. For example, in one group of horses that became increasingly adapted to grazing and running, body size, tooth size, and the number of the toes changed more or less monotonically (although not at a constant rate) during the Tertiary Period. Once

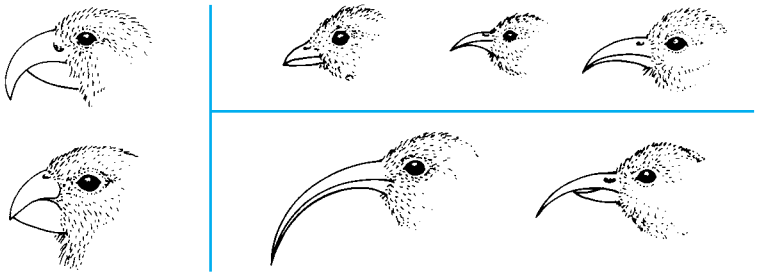

Fig. 5. Some members of an adaptive radiation, the Hawaiian honey-creepers. Although descended from finches, these birds have become adapted to a variety of ecological roles in the Hawaiian archipelago; some feed primarily on seeds (thick beaks), others on insects (short, thin beaks), others primarily on nectar (long, thin beaks). (*After D. J. Futuyma, Evolutionary Biology, Sinauer Associates, 1979*)

embarked on a way of life, a lineage experiences selection for improvement in the features that adapt it to that particular way of life, and the more specialized such features become, the less capable they are of being modified in other directions. A very common trend, for example, is reduction and simplification of parts, as in the reduction of limbs, lungs, and other features of snakes, which are now so morphologically and developmentally committed to their ways of life that they would be incapable of reevolving the lizardlike form of their ancestors. But although such trends can be discerned retrospectively, no one can say that a group such as the snakes was destined from their beginning to take the evolutionary path that they in fact followed: other roads may have been open to them at one time, but are open no longer. All history, including evolutionary history, is contingent on antecedent events.

For particular groups, one can document trends of increasing reduction and simplification; of the converse, increasing structural or behavioral complexity; of increased ecological specialization; of its converse, increased homeostatic freedom from the environment; and of increased mechanical or functional efficiency of morphological and physiological features. Few such trends continue unabated throughout evolutionary time; most are terminated by complete extinction, the fate of most of the higher taxa of organisms that have ever existed.

For life taken as a whole, there is some very tentative evidence that species (at least of marine invertebrates) have on average become more resistant to extinction, so that Cenozoic species have persisted longer than Mesozoic or Paleozoic species; but any such trend has been interrupted repeatedly by mass extinction events (the causes of which are actively disputed) that have eliminated much of the Earth's biota. The total number of species has increased after each such extinction event, and has been amplified in the last 100 million years or so by continental drift, which has accentuated differences in species composition in different parts of the world. More than $1^1/_2$ million living species have been described (of the 5 to 10 million that probably exist); and even though many major taxa have become extinct, the diversity of species now and in the recent past is higher than ever before in Earth's history. For life taken as

a whole, the only clearly discernible trend is toward ever-increasing diversity. Douglas J. Futuyma

Bibliography. J. C. Avise, *Molecular Markers, Natural History, and Evolution*, 1994; D. J. Futuyma, *Evolutionary Biology*, 3d ed., 1997; D. J. Futuyma, *Science on Trial: The Case for Evolution*, 1983; D. L. Hartl and A. G. Clark, *Principles of Population Genetics*, 3d ed., 1997; J. S. Levinton, *Genetics, Paleontology, and Macroevolution*, 2d ed., 2001; W.-H. Li and D. Graur, *Fundamentals of Molecular Evolution*, 2d ed., 2000; J. Maynard Smith, *Evolutionary Genetics*, 2d ed., 1998; E. Mayr, *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*, 1985; E. Mayr, *Populations, Species, and Evolution*, 1970.

Organic geochemistry

The study of the abundance and composition of naturally occurring organic substances, their origins and fate, and the processes that affect their distributions on Earth and in extraterrestrial materials.

Organic geochemistry was born from a curiosity about the organic pigments extractable from petroleum and black shales. It developed with extensive investigations of the chemical characteristics of petroleum and petroleum source rocks as clues to their occurrence and formation, and now encompasses a broad scope of activities within interdisciplinary areas of earth and environmental science. This range of studies recognizes the potential of geological records of organic matter to help characterize sedimentary depositional environments and to provide evidence of ancient life and indications of evolutionary developments through the Earth's history. Organic geochemistry includes determinations of anthropogenic contaminants amid the natural background of organic molecules and the assessment of their environmental impact and fate. Marine organic geochemistry addresses and interprets aquatic processes involving carbon species. It involves investigations of the chemical character of particulate and dissolved organic matter, evaluation of oceanic primary production including the factors (light, temperature, nutrient availability) that influence the uptake of carbon dioxide (CO_2), the composition of marine organisms, and the subsequent processing of organic constituents through the food web. Organic geochemistry extends to broader biogeochemical issues, such as the carbon cycle, and the effects of changing carbon dioxide levels, especially efforts to use geochemical data and proxies to help constrain global climate models. Examination of the organic chemistry of meteorites and lunar materials also falls within its compass, and as a critical part of the quest for remnants of life on Mars, such extraterrestrial studies are now regaining the prominence they held in the 1970s during lunar exploration. *See* COSMOCHEMISTRY; GEOCHEMISTRY.

These activities share the common need for identification, measurement, and assessment of organic matter in its myriad forms. Hence, many advances within this area have paralleled developments in analytical methodology or instrumentation that permit new approaches to the determination of the abundance, structure, and composition of organic matter in geological materials.

Global inventories of carbon. Carbon naturally exists as oxidized and reduced forms in carbonate carbon and organic matter. The major reservoir of both forms of carbon on Earth is the geosphere. It contains carbonate minerals deposited as sediments and organic matter accumulated from the remains of dead organisms. Estimates of the size of the geological reservoir of carbon vary within the range of 5 to 7 $\times 10^{22}$ g, of which 75% is carbonate carbon and 25% is organic carbon. The amounts of carbon contained in living biota (5×10^{17} g), dissolved in the ocean (4×10^{19} g), and present in atmospheric gases (7×10^{17} g) are miniscule compared to the quantity of organic carbon buried in the rock record. Most of it occurs in finely disseminated form within sedimentary rocks, especially shales. The importance of buried organic matter extends beyond its sheer magnitude; it includes the fossil fuels—coal, natural gas, and petroleum—that supply 85% of the world's energy. Global annual consumption of these three fuels represent only approximately 0.4%, 1.5%, and 2.3%, respectively, of their proven economic reserves, which in turn collectively account for only about 0.12% of the total accumulation of organic matter. These figures illustrate the immense quantities of buried organic carbon, which represent a geological legacy inherited from ancient life. They also show how the release of carbon dioxide from burning of these fossil fuels can readily increase the comparatively small atmospheric reservoir of carbon, perturbing the natural balance of the global carbon cycle. *See* BIOGEOCHEMISTRY; CARBON; CARBON DIOXIDE; CARBONATE MINERALS; FOSSIL FUEL; LIMESTONE; SEDIMENTARY ROCKS; SHALE.

Sedimentary organic matter. The vast amounts of organic matter contained in geological materials represent the accumulated vestiges of organisms amassed over the expanse of geological time. Yet, survival of organic cellular constituents of biota into the rock record is the exception rather than the norm. Only a small portion of the carbon fixed by organisms during primary production, especially by photosynthesis, escapes degradation as it settles through the water column and eludes microbial alteration during subsequent incorporation and assimilation into sedimentary detritus. Most biological debris is rapidly broken down or remineralized by surficial processes within the oceans, in lakes, in wetlands, and in soils and other terrigenous environments. Atmospheric and aqueous oxidation reactions contribute to this recycling process, which is also enhanced by consumption associated with food web processes and facilitated by sediment bioturbation (movement of sediment by animals). Protective packaging of organic matter can improve its chance of survival. Rapidly sinking zooplankton fecal pellets contain materials derived from phytoplankton that have survived ingestion and digestion and descend in discrete

capsules, thereby evading degradative influences. Hierarchical communities of aerobic and anaerobic microbes act as the predominant agents for degradation and alteration of organic matter in both the water column and sediments. They use it as their carbon source during a wide range of distinct metabolic processes, including heterotrophy, fermentation, sulfate reduction, methanogenesis, and methylotrophy. Microbial consortia selectively remove target substrates, and augment the surviving organic materials with secondary biomass, by adding distinctive components and isotopic signatures of their activity. The application of various microbiological procedures in examination of sedimentary materials, including incubation tests for specific metabolic processes and characterization of DNA isolates, is rapidly expanding the known distribution of microbial activity in the geosphere. These approaches demonstrate that the highly complex character of sedimentary organic matter often includes diagnostic evidence of its diverse biological sources and of the cumulative effects of alteration processes. *See* BIODEGRADATION; PETROLEUM MICROBIOLOGY.

Two distinct but related phenomena can enhance sequestration of organic matter in sedimentary environments. First, high levels of primary production can generate a supply of organic matter that exceeds the capacity of microbial flora to effectively degrade it. Second, depositional conditions, especially anoxia, can restrict the active populations of microbial communities engaged in the degradative process. Both circumstances can lead to exceptional preservation of organic matter, although their comparative importance in the process continues to be debated. Depositional conditions that favor survival of greater quantities of carbon, in terms of its elemental abundance, also affect the molecular and isotopic compositions of the organic matter by increasing survival of labile components. For example, hypersaline environments tend to facilitate the process of sulfur incorporation, which aids preservation of less stable structural features.

Kerogen and bitumen. Sedimentary organic matter can be divided operationally into solvent-extractable bitumen and insoluble kerogen. Bitumens contain a myriad of structurally distinct molecules, especially hydrocarbons, which can be individually identified (such as by gas chromatography-mass spectrometry) although they may be present in only minute quantities (nanograms or picograms). The range of components includes many biomarkers that retain structural remnants inherited from their source organisms, which attest to their biological origins and subsequent geological fate. *See* BITUMEN; KEROGEN.

For many years, kerogen was thought to form in sediments by a series of condensation reactions that combined the building blocks of life (amino acids, carbohydrates, lipids) into larger insoluble components. Undoubtedly such reactions do occur, aided by microbial alteration processes, but it is now apparent that a significant portion of kerogen represents macromolecular remnants, such as fragments of cell walls, inherited directly from living organisms or partially modified by removal of labile structural elements by degradative processes. The biological sources of organic matter, their survival during sediment deposition, and the extent of alteration effectively establish the elemental composition of kerogen that determines its character. Kerogen is broadly classified into three main types based on the proportions of hydrogen and oxygen relative to carbon (H/C and O/C ratios). Type I kerogen possesses high H/C and low O/C ratios, and is comparatively rare. It arises from contributions of organic matter that are dominated by phytoplankton and bacteria and usually occur in lacustrine (lake) settings. Type III kerogen exhibits low H/C and high O/C ratios, which result from a predominance of terrestrial organic matter that is typical in continental margin settings. It tends to be gas-prone, generating methane (CH_4). Type II kerogen displays intermediate H/C and O/C ratios, which reflects the mixed sources of organic matter that typify marine settings. Most oil source rocks contain this kerogen type, although some enriched in sulfur are labeled type II-S. The process of burial modifies all kerogen types, reducing both H/C and O/C ratios as organic matter experiences thermal alteration. Ultimately, the distinctive characteristics of the separate kerogen types and evidence of their sources are obscured by the burial processes that constitute diagenesis.

Diagenetic alteration of organic matter. Few organic constituents synthesized by living organisms remain stable in the sedimentary domain, although diagnostic features of their structures can survive for eons. To persist in sediments, organic matter must first survive the processes of recycling modification by microbial action during burial and consolidation. Subsequently, the organic matter experiences progressive transformation caused by thermal processes as its host rock undergoes compaction, dewatering, and perhaps remineralization. Modifications of organic matter proceed in a sequential manner and ultimately cause dramatic changes in its character. Hydrogen and oxygen atoms are lost incrementally, and eventually carbon-carbon bonds are broken, liberating mobile hydrocarbons, which is a crucial step in petroleum generation known as catagenesis. Continued alteration during burial at elevated temperature and pressure further depletes the hydrogen in residual organic material and ultimately produces graphite, one of the pure forms of carbon. *See* DIAGENESIS; GRAPHITE.

Petroleum formation. A number of specific geological conditions must be met for generation of recoverable petroleum, which is an inherently inefficient process. First, there is a need for the raw materials that can generate hydrocarbons, which must be provided by a source rock containing appreciable quantities (typically $>0.5\%$) of preserved organic matter of appropriate character. The nature of the organic matter determines the potential yield of petroleum and its composition. Second, the source rock must be buried at sufficient depth that kerogen cracking liberates hydrocarbons. This process is dependent

on a combination of time and temperature; petroleum generation can occur rapidly in regions where the Earth's heat flow is high, or it can proceed slowly over 10^8–10^9 years where thermal gradients are low. Third, the hydrocarbons released must migrate from the source rock to a suitable trap, often in folded or faulted rock strata, where they can accumulate in a reservoir. Primary migration occurs as generated oil moves out of the source rock aided by the thermal expansion and fracturing of kerogen. Subsequently, driven by pressure gradients, it moves through porous conduits displacing formation waters or along planes of weaknesses during the secondary migration to reach the reservoir. Finally, the integrity of the petroleum within its reservoir must be preserved for tens or hundreds of millions of years. Tectonic forces may fracture the reservoir allowing petroleum to escape, or the reservoir may be uplifted and eroded at the surface. Contact with meteoric waters percolating through the petroleum reservoir can lead to water washing, which selectively extracts the soluble petroleum constituents. It may also introduce microbial populations, which can sequentially remove components amenable to biodegradation, leaving residues that are increasingly resistant to alteration. Alternatively, petroleum may seep upward through fractures, to emanate as surface or submarine seeps, where it can evaporate or be washed or degraded by microbes. All of these processes tend to reduce the quality and value of petroleums. *See* PETROLEUM; PETROLEUM GEOLOGY; SEDIMENTOLOGY.

The composition of petroleum reflects its source materials, thermal and migration history, and extent of biodegradation. The combined effects and disparate interactions between these influences produce oils that vary widely in their chemistry and physical properties. Thus, investigation of the compositional characteristics of petroleum enables the cumulative history of these processes, and their effects, to be interpreted. Oil generation is an essential part of a continuum of changes that affect source rocks during their progressive burial, and that lead to successive liberation of petroleum and gases, contingent on the composition of the source material. Thus, the thermal processes that transform kerogen into petroleum continue beyond oil generation, and fragment organic matter into progressively smaller molecules, which leads to production of hydrocarbon gases. The volume and composition of the gases generated is dependent on whether the source rock is oil- or gas-prone, which in turn is a function of its composition, especially its dominant kerogen type. However, as thermal maturity increases, there is a trend in composition from wet gas to dry gas, as the proportion of hydrocarbon gases that condense as liquids (for example butane and propane) decreases while methane becomes predominant. *See* METHANE; NATURAL GAS.

Methane can also be produced in shallow sediments by the microbial activity of methanogenic bacteria. Such bacterial methane forms many recoverable gas accumulations at comparatively shallow depths and represents the source of methane in gas hydrates. Hydrates occur in sediments where the combination of low temperature and high pressure stabilizes the ice cages that can entrap methane. The worldwide frequency of hydrate occurrences supports estimates that they may greatly exceed the combined resources of oil and gas (0.5% versus 0.017% of global carbon). Moreover, massive release of methane from hydrates is a direct consequence of global warming and is now recognized as the climatic perturbation that caused significant changes in Earth's flora and fauna associated with the late Paleocene thermal maximum, 55 million years ago. *See* HYDRATE; METHANOGENESIS (BACTERIA).

Coal. Peat deposits contain an abundance of terrestrial organic matter that is progressively transformed by burial compaction and thermal alteration into lignite, bituminous coal, and ultimately graphite. Coals vary in type and chemical characteristics, largely dependent on the composition of their organic matter (typically type III kerogen), its extent of bacterial degradation, and its thermal history. *See* LIGNITE; PEAT.

Coals contain many plant fragments, which are often altered but can be identified under the microscope and classified according to their origins. These integral parts of coal are called macerals, which are individually described and named according to their source, for example, sporinite (spores), resinite (resin), and vitrinite (woody). Macerals can be recognized in other sediments, where they provide evidence of contributions of organic matter from terrestrial plants and aid assessment of depositional environments. The reflectance of the maceral vitrinite changes systematically with increasing burial and thermal maturity, whether in coals or other rocks. Thus, determination of vitrinite reflectance (R_o) by microscopic examination provides a direct measure of its thermal history, and is widely used in assessing the maturity of petroleum source rocks. *See* COAL.

Biomarkers. Biomarkers are individual compounds whose chemical structures carry evidence of their origins and history. Recognition of the specificity of biomarker structures initially helped confirm that petroleum was derived from organic matter produced by biological processes. For example, the structural and stereochemical characteristics of the hydrocarbon cholestane that occurs in petroleums can be directly and unambiguously linked to an origin from teh sterol cholesterol, which is a widespread constituent of plants and animals. Of the thousands of individual petroleum components, hundreds reflect precise biological sources of organic matter, which distinguish and differentiate their disparate origins. The diagnostic suites of components may derive from individual families of organisms, but contributions at a species level can occasionally be recognized. Biomarker abundances and distributions help to elucidate sedimentary environments, providing evidence of depositional settings and conditions. They also reflect sediment maturity, attesting to the progress of the successive,

sequential transformations that convert biological precursors into geologically occurring products. Thus, specific biomarker characteristics permit assessment of the thermal history of individual rocks or entire sedimentary basins. *See* BASIN.

Carbon isotopes. Carbon naturally occurs as three isotopes: carbon-12 (^{12}C), carbon-13 (^{13}C), and radiocarbon (^{14}C). The proportion of ^{13}C in living organic matter is controlled by the organism's source of carbon, fractionation during carbon fixation, and both biosynthetic and metabolic processes. Measurements of ^{13}C are expressed as a ratio, which compares the amount of ^{13}C in the sample to a standard (carbonate in the Peedee belemnite) and provides a value in parts per thousand (‰), defined as $^{13}C = \{[(^{13}C/^{12}C)_{sample}/(^{13}C/^{12}C)_{standard}] - 1\} \times 1000$. Organic carbon tends to be depleted of ^{13}C (negative ^{13}C values) relative to inorganic carbon (carbon dioxide and carbonate) as organisms preferentially uptake and use ^{12}C because its reactivity is greater than that of ^{13}C. However, the extent of ^{13}C depletion is dependent on biosynthetic pathways. For example, photosynthetic carbon fixation of carbon dioxide by plants that employ the C_3 and C_4 pathways produces organic matter averaging −26‰ and −12‰, respectively. The significant difference is caused by the ability of the latter group, mainly comprising grasses, to store carbon dioxide before fixation. Methylotrophic bacteria use methane, which is highly depleted in ^{13}C, rather than carbon dioxide as their carbon source. Consequently, they produce organic compunds that possess highly negative ^{13}C values (−55‰ to −120‰). By contrast, the products synthesized by bacteria that use the reverse-TCA (tricarboxylic acid, or Krebs) cycle exhibit ^{13}C values typically less than −10‰. Thus, the carbon isotopic composition of many individual components found in sediments attests to the diversity of their origins and demonstrates the variety of microbial processes active in the shallow subsurface. *See* ISOTOPE; PHOTOSYNTHESIS; PLANT RESPIRATION.

Temporal excursions in the ^{13}C values of sediment sequences can reflect perturbations of the global carbon cycle. For example, episodes of enhanced burial of organic matter that occurred during widespread oceanic black shale deposition in the Cretaceous effectively depleted the global ^{12}C reservoir, leading to an enrichment of ^{13}C in sedimentary organic matter. Also, global shifts in carbon isotopic signals provide critical evidence that the late Paleocene thermal maximum was associated with a release of methane from hydrates. Similarly, the remarkable variations in ^{13}C values of organic matter in the Late Proterozoic are consistent with the "snowball Earth" hypothesis that ice sheets extended to the Equator, creating long-term episodes of glaciation. These events ended catastrophically when greenhouse conditions gradually created by a build-up of atmospheric carbon dioxide became sufficient to melt the ice, leading to extraordinarily rapid weathering of continents as their ice burden dissipated, and dramatic deposition of marine carbonates as the ocean again became a sink for carbon dioxide. *See* BLACK SHALE; CRETACEOUS; MARINE SEDIMENTS; PALEOCEANOGRAPHY.

Radiocarbon is widely employed to date archeological artifacts, but the sensitivity of its measurement also permits its use in exploration of the rates of biogeochemical cycling in the oceans. This approach permits assessment of the ages of components in sediments, demonstrating that bacterial organic matter is of greater antiquity than components derived from phytoplankton sources. *See* RADIOCARBON DATING.

Applications. A major impetus for the development of organic geochemistry has been its use as a tool in petroleum exploration. Determination of the quantity, nature, and degree of thermal alteration of organic matter from kerogen analysis provides evidence of the richness and maturity of source rocks and the timing of oil generation. Such information is critical in evaluation of prospects for petroleum formation and accumulation. Similarly, chemical clues provided by the molecular and isotopic composition of petroleums enable the determination of the biological sources of organic matter (marine or terrestrial) and its environment of deposition. Such information can also place constraints on the age of petroleums based on the timing of the appearance of specific biomarkers, based on evolutionary trends in their source organisms. Molecular distributions also provide valuable fingerprints for the correlation of source rocks and oils, and the ^{13}C values of hydrocarbon gases help distinguish their origins from bacterial or thermogenic sources, especially when coupled with the determination of their hydrogen isotope compositions (deuterium values). *See* DEUTERIUM; HYDROGEN.

Molecular and isotopic tools are gaining acceptance as paleoclimate proxies. The proportion of unsaturation (number of carbon-carbon double bonds) in a group of compounds known as alkenones is biosynthetically controlled by the growth temperature of the phytoplankton that synthesize them. The survival of this temperature signal in marine sediment sequences provides a temporal record of sea surface temperatures that reflect past climates, including evidence of the timing and extent of global cooling during ice ages, and the warming associated with El Niño events. *See* CHEMOSTRATIGRAPHY; EL NIÑO.

Also, differences between the carbon isotopic composition of the alkenones and co-occurring carbonate can be interpreted in terms of ancient atmospheric levels of carbon dioxide, when the effect of the growth rates of the alkenone-producing phytoplankton is excluded. *See* PHYTOPLANKTON.

Simon C. Brassell

Bibliography. M. Engel and S. A. Macko, *Organic Geochemistry*, 1993; J. M. Hunt, *Petroleum Geochemistry and Geology*, 2d ed., 1996; S. Killops and V. J. Killops, *An Introduction to Organic Geochemistry*, 1993; L. M. Pratt, J. B. Comer, and S. C. Brassell (eds.), *Geochemistry of Organic Matter in Sediments and Sedimentary Rocks*, 1992.

Organic nomenclature

A system by which a unique and unambiguous name or other designation is assigned to a given organic molecular structure. A set of rules adopted by the International Union of Pure and Applied Chemistry (IUPAC) is the basis for a standardized name for any organic compound. Common or nonsystematic names are used for many compounds that have been known for a long time. The latter names have the advantage of being short and easily recognized, as in the examples below However, in contrast to system-

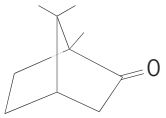

Common name: camphor

Systematic name:
1,7,7-trimethyl-
bicycol[2.2.1]heptan-6-one

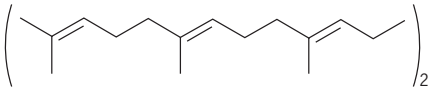

Common name: squalene

Systematic name:
2,6,10,15,19,23-hexamethyl-
2,6,10,14,18,22-tetracosahexaene

atic names, common or trivial names do not convey information from which the structure can be written by reference to prescribed rules.

Aliphatic compounds. In the IUPAC system, a name is formed by combination of a parent alkyl chain or ring with prefixes and suffixes to denote substituents. For aliphatic compounds, the parent name is a stem that denotes the longest straight chain in the structure with the ending -ane. The first four members of the alkane series are methane, ethane, propane, and butane. From five carbons on, the names follow the Greek numerical roots, for example, pentane and hexane. Changing the ending -ane to -yl gives the name of the corresponding radical or group; for example, CH_3CH_2— is the ethyl radical or ethyl group. Branches on an alkyl chain are indicated by the name of a radical or group as a prefix. The location of a branch or other substituent is indicated by number; the chain is numbered from whichever end results in the lowest numbering. Examples are shown below.

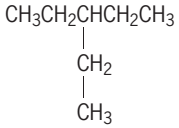

3-Ethylpentane

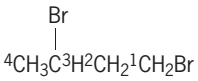

1,3-Dibromobutane

1-Propyl 2-Propyl

See ALKANE.

Double or triple bonds are indicated by the endings -ene or -yne, respectively. The configuration of the chain at a double bond is denoted by *E*- when two similar groups are on opposite sides of the plane bisecting the bond, and *Z*- when they are on the same side. This is illustrated by the structures below.

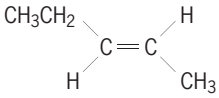

E-2-Pentene

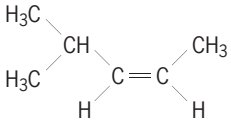

4-Methyl-2*Z* pentene

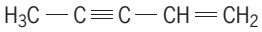

1-Penten-4-yne

See ALKENE; ALKYNE.

A compound containing a functional group is named by adding to the parent name a suffix characteristic of the group. If there are two or more groups, a principal group is designated by suffix and the other(s) by prefix. Examples are given in **Table 1**, in which groups are listed in descending order of priority as the suffix.

Carbocyclic compounds. The same general principles apply to cyclic compounds. For alicyclic rings, the prefix cyclo- is followed by a stem indicating the number of carbon atoms in the ring, as is illustrated in the structure below. In bicyclic compounds, the

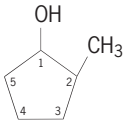

2-Methylcyclopentanol

total number of carbons in the ring system is prefixed by bicyclo- and numbers in brackets which indicate the number of atoms in each connecting chain. The atoms in a bridged bicyclic ring system are numbered by commencing at a bridgeheadatom and

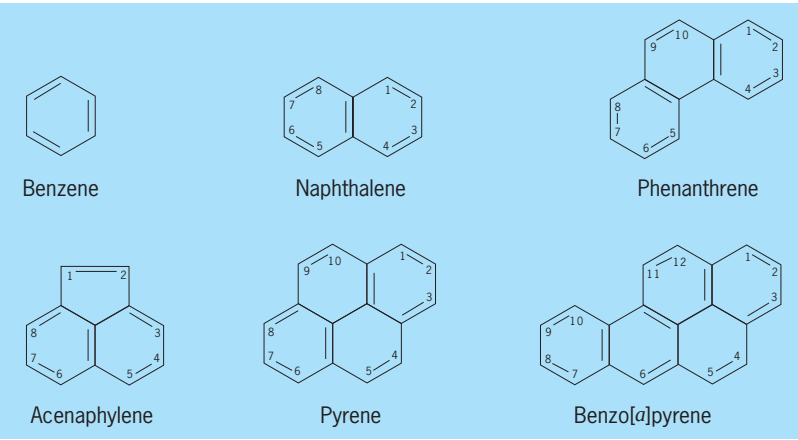

Fig. 1. Structures and position numbering of benzene and some representative polycyclic aromatic structures.

TABLE 1. Names of compounds containing functional groups

Group*	Suffix	Prefix	Structure	Name
$-\overset{O}{\overset{\|\|}{C}}OH$	-oic acid	carboxy-	$CH_3CH_2\overset{O}{\overset{\|\|}{C}}OH$	Propanoic acid
$-\overset{O}{\overset{\|\|}{C}}OR$	alkyl -oate	alkoxycarbonyl-	$CH_3CH_2\overset{O}{\overset{\|\|}{C}}OCH_3$	Methyl propanoate
			$CH_3O\overset{O}{\overset{\|\|}{C}}CH_2CH_2\overset{O}{\overset{\|\|}{C}}OH$	3-Methoxycarbonyl propanoic acid
$-C\equiv N$	-nitrile	cyano-	$CH_3CH_2C\equiv N$	Propanenitrile
			$CH_3CHCN\overset{O}{\overset{\|\|}{C}}OH$	2-Cyanopropanoic acid
$-\overset{O}{\overset{\|\|}{C}}H$	-al	formyl-	$CH_3CH_2\overset{O}{\overset{\|\|}{C}}H$	Propanal
$-\overset{O}{\overset{\|\|}{C}}R$	-one	oxo-	$CH_3CH_2\overset{O}{\overset{\|\|}{C}}CH_3$	Butanone
			$CH_3\overset{O}{\overset{\|\|}{C}}CH_2\overset{O}{\overset{\|\|}{C}}OH$	3-Oxobutanoic acid
$-OH$	-ol	hydroxy-	$CH_3CH_2CH_2OH$	1-Propanol
			$HOCH_2CH_2\overset{O}{\overset{\|\|}{C}}H$	3-Hydroxypropanal
$-NH_2$	-amine	amino-	$CH_3CH_2CH_2NH_2$	1-Propanamine
$-OR$	—	alkoxy-	$NH_2CH_2CH_2CH_2OCH_3$	3-Methoxy-1-propanamine

*R = any alkyl group.

numbering by the longest path to the other bridgehead, and then the shorter bridge, as is illustrated in the structure below.

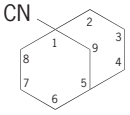

1-Cyanobicyclo[3.3.1] nonane

See ALICYCLIC HYDROCARBON.

The names of aromatic hydrocarbons have the ending -ene, which denotes a ring system with the maximum number of noncumulative double bonds. Each of the simpler polycyclic hydrocarbons has a different parent name; a few of these, with the ring numbering, are shown in **Fig. 1**. Other fused ring systems can be named by adding the prefix benzo-, which denotes attachment of an additional —CH═CH—CH═CH— chain at two adjacent carbons.

To number the positions in a polycyclic aromatic ring system, the structure must first be oriented with the maximum number of rings arranged horizontally and to the right. The system is then numbered in clockwise sequence starting with the atom in the most counterclockwise position of the upper right-hand ring, omitting atoms that are part of a ring fusion. *See* AROMATIC HYDROCARBON.

Heterocyclic compounds. Systematic names for rings containing a heteroatom (O, S, N) are based on a combination of a prefix denoting the heteroatom(s) and a suffix denoting the ring size, as indicated in **Table 2**. Examples are shown in **Fig. 2**; as is often the case, the simplest rings have names that depart from this system. *See* CHEMICAL SYMBOLS AND

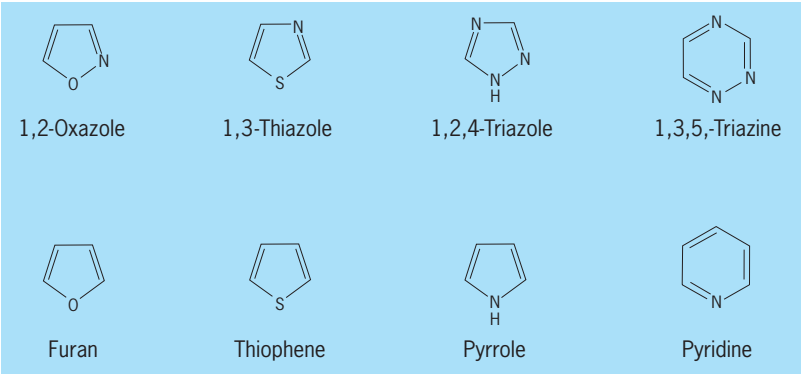

Fig. 2. Examples of structures and names of some representative heterocyclic compounds.

TABLE 2. Basis of nomenclature of heterocyclic compounds

Heteroatom	Prefix	Ring size	Suffix
O	ox(a)-	4	-ete
S	thi(a)-	5	-ole
N	az(a)-	6	-ine
		7	-epine

FORMULAS; ORGANIC CHEMISTRY; STRUCTURAL CHEMISTRY.

James A. Moore

Bibliography. F. A. Carey, *Organic Chemistry*, 2d ed., 1992; International Union of Pure and Applied Chemistry, *Nomenclature of Organic Chemistry*, 1979.

Organic photochemistry

A branch of chemistry that deals with light-induced changes of organic material. Because it studies the interaction of electromagnetic radiation and matter, photochemistry is concerned with both chemistry and physics. In considering photochemical processes, therefore, it is also necessary to consider physical phenomena that do not involve strictly chemical changes, for example, absorption and emission of light, and electronic energy transfer.

Because several natural photochemical processes were known to play important roles (namely, photosynthesis in plants, process of vision, and phototropism), the study of organic photochemistry started very early in the twentieth century. However, the breakthrough occurred only after 1950 with the availability of commercial ultraviolet radiation sources and modern analytical instruments for nuclear magnetic resonance (NMR) spectroscopy and (gas) chromatography. *See* GAS CHROMATOGRAPHY; NUCLEAR MAGNETIC RESONANCE (NMR).

Photo products. In general, most organic compounds consist of rapidly interconvertible, nonseparable conformers, because of the free rotation about single bonds. Each conformer has a certain energy associated with it, and its own electronic absorption spectrum. This is one cause of the broad absorption bands produced by organic compounds in solution. The equilibrium of the conformers may be influenced by the solvent and the temperature. According to the Franck-Condon principle, which states that promotion of an electron by the absorption of a photon is much faster than a single vibration, each conformer will have its own excited-state configuration. *See* CONFORMATIONAL ANALYSIS; FRANCK-CONDON PRINCIPLE; MOLECULAR STRUCTURE AND SPECTRA.

Because of the change in the π-bond order upon excitation of the substrate and the short lifetime of the first excited state, the excited conformers are not in equilibrium and each yields its own specific photoproduct (**Fig. 1**). Though different conformers may lead to the same photoproduct and one excited conformer may lead to several photoproducts, a change in solvent, temperature, or wavelength of excitation influences the photoproduct composition. This is especially true with small molecules; larger molecules with aromatic groups are less sensitive for small wavelength differences.

The influence of wavelength is also of importance when the primary photoproduct also absorbs light and then gives rise to another photoreaction. Excitation with selected wavelengths or monochromatic light by use of light filters or a monochromator, respectively, may then be profitable for the selective production of the primary product. Similarly, irradiation at a low temperature is helpful in detecting a primary photoproduct that is unstable when heated (thermolabile). *See* COLOR FILTER.

In addition to photoreactions and fluorescence from the first excited singlet state (S_1) of an organic compound, the intersystem crossing (isc) to the triplet state (T_1) can also occur. In principle, this state cannot be attained by direct irradiation because of the Pauli exclusion principle. The quantum yield of intersystem crossing depends on the difference in energy between the singlet and triplet state. *See* EXCLUSION PRINCIPLE; FLUORESCENCE; PHOTOCHEMISTRY; TRIPLET STATE.

Compounds containing heteroatoms [other than carbon (C) or hydrogen (H)], especially those with heavy groups such as bromine (Br) or nitro (NO_2), have a high quantum yield of intersystem crossing caused by spin-orbit coupling. Theoretically the conversion of the triplet state to the ground state (S_0) is not allowed, and therefore the lifetime of the triplet state is much longer than that of the first excited singlet state. For this reason, photoreactions are much more probable from the triplet state than from the first excited singlet state. The physical properties of the excited triplet state are different from those of the first excited singlet state, and the unpaired electrons of the triplet state cause the compound to have a biradical character. The photoproducts derived from the two excited states will therefore be different, although a given product may be derived from both states. *See* QUANTUM CHEMISTRY.

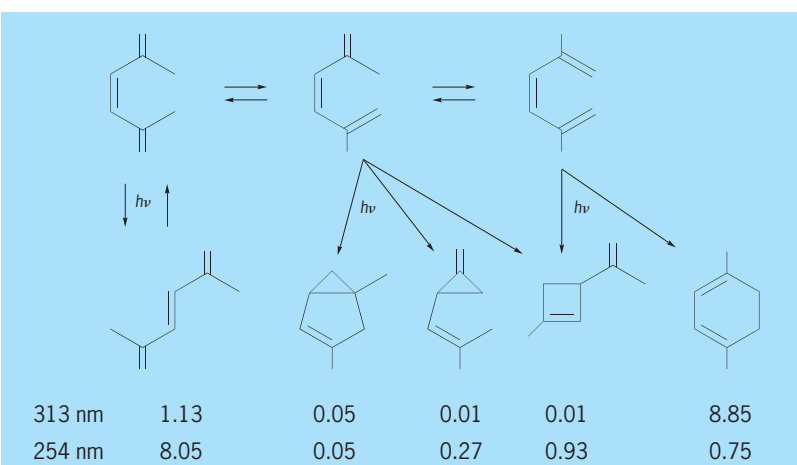

Fig. 1. Reaction scheme of three rapidly interconvertible, nonseparable conformers and their yields of specific photoproducts, showing the composition of the photoproduct mixture (in percent) of the Z isomer of 2,5-dimethylhexa-1,3,5-triene after 10% conversion at two wavelengths.

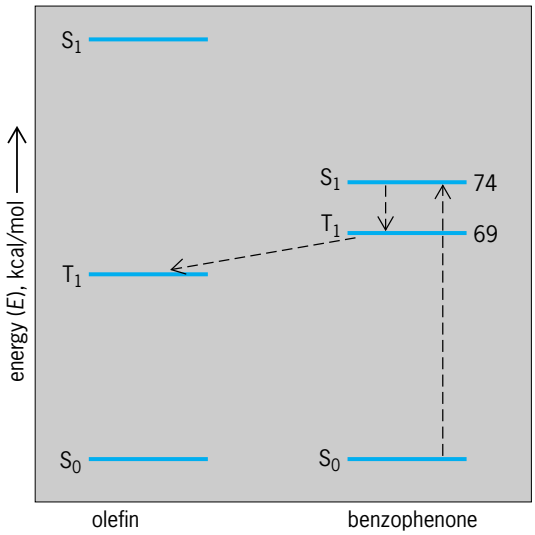

Fig. 2. Benzophenone as a triplet sensitizer for an olefin. Symbols are explained in the text.

Sensitization. To investigate which product is derived from the first excited singlet state, a triplet quencher can to be added to a compound. The triplet energy of the quencher must be lower than that of the compound. However, to identify the products from the triplet state or to obtain only products from this state, the triplet state has to be populated without formation of the first excited singlet state. This can be done by a method known as intermolecular energy transfer. This involves adding a sensitizer to the compound. The sensitizer compound has to possess a high quantum yield of intersystem crossing, and a singlet energy lower and a triplet energy higher than that of the original compound. (**Fig. 2**). Frequently, aromatic ketones and certain dyes that fulfill the electronic requirements are used as sensitizers.

Energy surface description. It is assumed that a potential energy curve controls nuclear motion, except for situations where two surfaces come close together. Each point on such a curve represents a specific geometry (horizontal axis) and a specific energy (vertical axis) [**Fig. 3**]. The energy barrier between the reactant and the product on the ground-state surface may be too high to be overcome by a thermal reaction. Irradiation of the reactant gives rise to the excited species, usually with enough vibrational energy to overcome small barriers in the first excited singlet state surface and arrive at a point that is close to the ground state for a given geometry. By a jump from the first excited singlet state to the ground state at this position, the product and/or the reactant can be attained. This is an example of a nonadiabatic or diabatic photoreaction, which is the most usual type of photoreaction. Thus a small change in the mutual positions of the minimum and maximum of the curves of the first excited singlet state and the ground state due to a substituent, solvent, or another factor will affect the balance of the formation of product and reactant. While organic molecules, consisting of several atoms, possess much more complicated energy hyperfaces, this kind of approximation is found to be useful.

Photoreactions. Organic photochemistry involves all classes of organic compounds. Some examples include the photochemistry of saturated hydrocarbons, unsaturated hydrocarbons, aromatic compounds, and carbonyl compounds.

Saturated hydrocarbons. These compounds have only σ bonds and absorb in the vacuum ultraviolet part of the spectrum. Their photochemistry has not been extensively investigated. Substituted alkanes that absorb above about 200 nanometers are photolyzed on irradiation and form alkyl radicals, which react in the usual way. *See* CHEMICAL BONDING; REACTIVE INTERMEDIATES.

Unsaturated hydrocarbons. These compounds possess one or more double bonds. The second bond of a double bond is known as a π bond; it has a lower energy than the σ bond. With light, the excited $\pi\pi^*$ state is populated.

Because of the large distance between the excited singlet and triplet states of these compounds, the intersystem crossing is inefficient and most photoreactions on direct excitation occur from the $^1\pi\pi^*$ state. Usually the *trans,cis*-isomerization is the most efficient reaction (Fig. 1). Other common intramolecular photoreactions are sigmatropic shifts, electrocyclizations, and cycloadditions.

Not uncommonly in photochemistry, in a transition state or in the primary photoproduct the aromaticity is lost, because of the high energy of the excited state.

Aromatic compounds. Contrary to thermal reactions of aromatic compounds, which occur with retention of aromaticity, photochemical reactions of these compounds give a variety of nonaromatic products. For example, prismane can be derived as a photoproduct of butylbenzene, in more than 70% yield. In general, the yield of the photoproducts depends on the type and position of the substituents in the benzene ring.

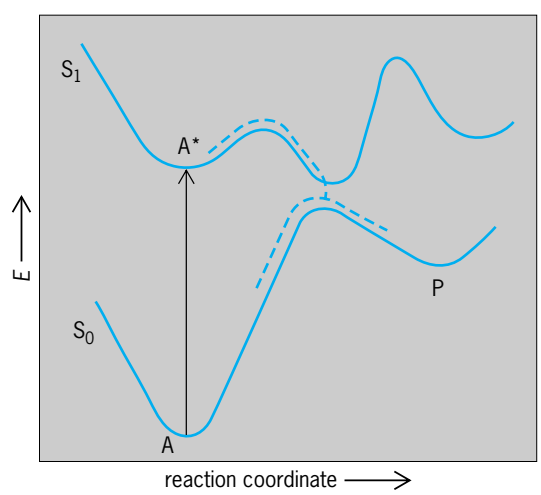

Fig. 3. Energy curves of the ground state (S_0) and the first excited singlet state (S_1) of the photoreaction. By a jump from the first excited singlet state to the ground state, the product (P) or reactant (A) can be attained

Although the energy content of benzene is lower than that of prismane, the latter molecule is stable at room temperature as the thermal ring-opening is not allowed theoretically, because of orbital symmetry considerations. Photochemical isomerizations of benzene derivatives occur via benzvalene or prismane intermediates and not by shifts of the substituents.

Irradiation of benzene and its derivatives in the presence of olefins usually leads to cycloaddition products. The preferred photoreaction with olefins bearing an electron-withdrawing group is a 1,2-(ortho) cycloaddition. The more interesting 1,3-(meta) cycloaddition is generally observed when nonpolar alkenes are used. Not only intermolecular reactions take place, but also intramolecular cycloadditions are found when the double bond is at a suitable position to react. Substitution reactions of aromatic compounds induced by light are usually nucleophilic, again in contrast to those in the ground state, which are electrophilic in nature. *See* SUBSTITUTION REACTION.

Carbonyl compounds. Carbonyl compounds contain a carbonyl group (C═O). In the ground state the C-O bond is polarized, with a partial negative charge on the oxygen atom. Excitation by light causes an n-π^* transition, with the oxygen becoming electron deficient and the carbon more electron rich. The (n,π^*) triplet state of the carbonyl group has a diradical character. These properties determine the reactivities of this class of compounds and make the photochemistry of these compounds rather complicated. In aliphatic carbonyl compounds the intersystem crossing from the first excited singlet state to the triplet state is relatively slow, so that reactions are probable from both states. However, in conjugated aryl ketones and α,β-unsaturated carbonyl compounds, the intersystem crossing is very fast and only the triplet state is the reactive state for reactions. The most important types of photoreaction are α cleavage of the molecule, intramolecular hydrogen abstraction and photoelimination, photoreduction, and cycloaddition to olefins. *See* ORGANIC CHEMISTRY; ORGANIC SYNTHESIS. Wim. H. Laarhoven

Bibliography. A. M. Braun, E. Oliveres, and M. T. Maurette, *Photochemical Technology*, 1991; A. Gilbert and J. Baggott, *Essentials of Molecular Photochemistry*, 1991; W. M. Horspool (ed.), *Handbook of Organic Photochemistry and Photobiology*, 1994; J. Mattay and A. G. Griesbeck (eds.), *Photochemical Key Steps in Organic Syntheses*, 1994; S. A. L. Murov, *Handbook of Photochemistry*, 2d ed., 1993.

Organic reaction mechanism

A complete, step-by-step account of how a reaction of organic compounds takes place. A fully detailed mechanism would correlate the original structure of the reactants with the final structure of the products and would account for changes in structure and energy throughout the progress of the reaction. A complete mechanism would also account for the formation of any intermediates and the rates of interconversions of all of the various species. Because it is not possible to detect directly all of these details of a reaction, evidence for a reaction mechanism is always indirect. Experiments are designed to produce results that provide logical evidence for (but can never unequivocally prove) a mechanism. For most organic reactions, there are mechanisms that are considered to be well established (that is, plausible) based on bodies of experimental evidence. Nevertheless, new data often become available that provide further insight into new details of a mechanism or that occasionally require a complete revision of an accepted mechanism.

Classification of organic reactions. The description of an organic reaction mechanism typically includes designation of the overall reaction (for example, substitution, addition, elimination, oxidation, reduction, or rearrangement), the presence of any reactive intermediates (that is, carbocations, carbanions, free radicals, radical ions, carbenes, or excited states), the nature of the reagent that initiates the reaction (such as electrophilic or nucleophilic), the presence of any catalysis (such as acid or base), and any specific stereochemistry. For example, reaction (1) would

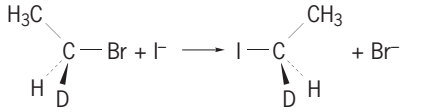

(1)

be described as a concerted nucleophilic substitution of an alkyl halide that proceeds with inversion of stereochemistry. A reaction that proceeds in a single step, without intermediates, is described as concerted or synchronous. Reaction (1) is an example of the S_N2 mechanism (substitution, nucleophilic, bimolecular).

Potential energy diagrams. A common method for illustrating the progress of a reaction is the potential energy diagram, in which the free energy of the system is plotted as a function of the completion of the reaction (see **illus.**).

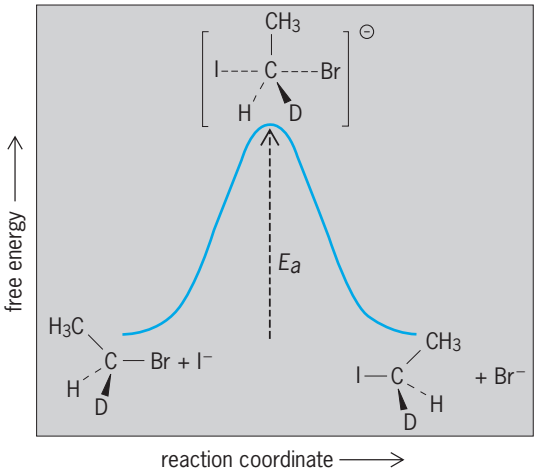

Potential energy–reaction coordinate diagram for a typical nucleophilic substitution reaction that proceeds by the S_N2 mechanism. E_a = activation energy.

The reaction coordinate is intended to represent the progress of the reaction, and it may or may not correlate with an easily observed or measurable feature. In reaction (1), the reaction coordinate could be considered to be the increasing bond length of the carbon-bromine (C-Br) bond as it is broken, or the decreasing separation of C and iodine (I) as they come together to form a bond. In fact, a complete potential energy diagram should illustrate the variation in energy as a function of both of these (and perhaps several other relevant structural features), but this would require a three- dimensional (or higher) plot.

Besides identifying the energy levels of the original reactants and the final products, the potential energy diagram indicates the energy level of the highest point along the reaction pathway, called the transition state. It is important to recognize that the reaction pathway actually taken in a reaction mechanism is chosen because it represents the lowest-energy pathway. The analogy often used is a mountain pass between two valleys, in which the top of the pass represents the transition state; the top of the pass represents the highest point traveled, but not the highest point in the vicinity. Because the transition state represents the highest energy that the molecules must attain as they proceed along the reaction pathway, the energy level of the transition state is a key indication of how easily the reaction can occur. Features that tend to make the transition state more stable (lower in energy) make the reaction more favorable. Such stabilizing features could be intramolecular, such as electron donation or withdrawal by substituents, or intermolecular, such as stabilization by solvent. *See* CHEMICAL BONDING; ENERGY.

Kinetics. Another way to illustrate the various steps involved in a reaction mechanism is as a kinetic scheme that shows all of the individual steps and their rate constants. The S_N2 mechanism is a single step, so the kinetics must represent that step; the rate is observed to depend on the concentrations of both the organic substrate and the nucleophile. However, for multistep mechanisms the kinetics can be a powerful tool for distinguishing the presence of alternative pathways. For example, when more highly substituted alkyl halides undergo nucleophilic substitution, the rate is independent of the concentration of the nucleophile. This evidence suggests a two-step mechanism, called the S_N1 mechanism, as shown in reaction scheme (2), where the k terms represent rate constants.

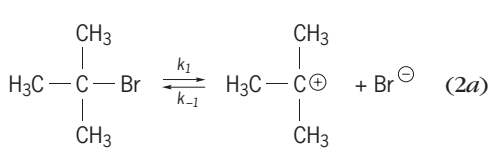

(2*a*)

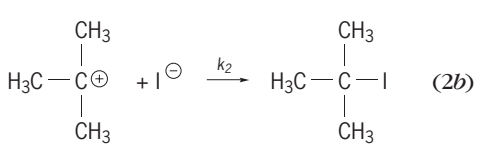

(2*b*)

The S_N1 mechanism accomplishes the same overall nucleophilic substitution of an alkyl halide, but does so by initial dissociation of the leaving group (Br^-) to form a carbocation, step (2*a*). The nucleophile then attaches to the carbocation to form the final product, step (2*b*). Alkyl halides that have bulky groups around the carbon to be substituted are less likely to be substituted by the direct S_N2 mechanism, because the nucleophile encounters difficulty in making the bond to the inaccessible site (called steric hindrance). If those alkyl groups have substituents that can support a carbocation structure, generally by electron donation, then the S_N1 mechanism becomes preferable.

A crucial feature of a multistep reaction mechanism is the identification of the rate-determining step. The overall rate of reaction can be no faster than its slowest step. In the S_N1 mechanism, the bond-breaking reaction (2*a*) is typically much slower than the bond-forming reaction (2*b*). Hence, the observed rate is the rate of the first step only. Thus, kinetics can distinguish the S_N1 and S_N2 mechanisms, as shown in Eqs. (3) and (4), where R is an alkyl group, X is a halo-

$$\text{Rate} = k\,[\text{RX}]\,[\text{Nu}] \quad \text{for an } S_N2 \text{ mechanism} \tag{3}$$

$$\text{Rate} = k\,[\text{RX}] \quad \text{for an } S_N1 \text{ mechanism} \tag{4}$$

gen or other leaving group, Nu is a nucleophile, and the terms in the brackets represent concentrations.

A more complete description of the S_N1 mechanism was recognized when it was observed that the presence of excess leaving group [for example, Br^- in reaction (2)] can affect the rate (called the common ion rate depression). This indicated that the mechanism should include a reverse step [k_{-1} in reaction step (2*a*)] in which the leaving group returns to the cation, regenerating starting material. In this case, the rate depends in a complex manner on the competition of nucleophile and leaving group for reaction with the carbocation. *See* CHEMICAL DYNAMICS; REACTIVE INTERMEDIATES; STERIC EFFECT (CHEMISTRY).

Activation parameters. The temperature dependence of the rate constant provides significant information about the transition state of the rate-determining step. The Arrhenius equation (5)

$$k = Ae^{-E_a/RT} \tag{5}$$

expresses that dependence in term of an exponential function of temperature and an activation energy, E_a; A is called the Arrhenius or preexponential factor, and R is the gas constant. *See* GAS; LOGARITHM.

The activation energy represents the energy difference between the reactants and the transition state, that is, the amount of energy that must be provided in order to proceed along the reaction pathway successfully from reactant to product (see illus.).

A more complete description of transition-state theory expresses the rate constant in terms of contributions from both an enthalpy of activation, $\Delta H^{\ddagger}$, and an entropy of activation, $\Delta S^{\ddagger}$. In this case, enthalpy of activation represents the additional

enthalpy content of the transition state compared to reactants; $\Delta H^{\ddagger}$ can be rather directly correlated with the energy of activation. The entropy of activation, $\Delta S^{\ddagger}$, represents the additional entropy content of the transition state compared to reactants; $\Delta S^{\ddagger}$ is highly informative with respect to the reaction mechanism, because it indicates whether the overall structure becomes more ordered or more disordered as the reaction proceeds from reactants to transition state. For example, the transition state for the S_N2 reaction (see illus.) requires the reactants to be organized in a very specific manner, with the nucleophile approaching the opposite side of the leaving group; the S_N2 reaction typically shows a negative value for $\Delta S^{\ddagger}$. *See* ENTHALPY; ENTROPY.

Stereochemistry. Careful attention to the stereochemistry of a reaction often provides crucial insight into the specific orientation of the molecules as they proceed through the reaction mechanism. The complete inversion of stereochemistry observed in the S_N2 mechanism provides evidence for the backside attack of the nucleophile. Alkyl halides that undergo substitution by the S_N1 mechanism do not show specific stereochemistry, since the loss of the leaving group is completely uncorrelated with the bonding of the necleophile.

In addition reactions, possible stereochemical outcomes are addition of the new bonds to the same or opposite sides of the original pi bond, called syn and anti addition, respectively. The anti addition of bromine to double bonds provides evidence for the intermediacy of a bridged bromonium ion, as shown in reaction (6).

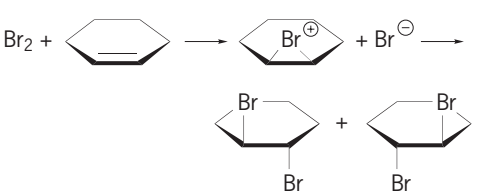

(6)

Experimental probes of mechanisms. Chemists use a variety of techniques to determine the mechanistic course of reactions. The most direct is actual observation of the reactants, intermediates, and products in real time during the reaction. Fast spectroscopic techniques are becoming ever faster and more sensitive, so that direct evidence for short-lived intermediates is often obtainable. A representative fast technique is flash photolysis, in which a compound is rapidly decomposed by an intense pulse of light. Laser light pulses less than a picosecond (10^{-12} s) are available; they, combined with detection methods that can record data on a time scale of picoseconds, allow for direct observation of the spectroscopic or other properties of reaction intermediates. *See* LASER SPECTROSCOPY; ULTRAFAST MOLECULAR PROCESSES.

Another approach is to create an environment that is stabilizing for the reaction intermediates to allow study of their properties by standard techniques. A common approach called matrix isolation creates the intermediate in a frozen matrix, so that it is unable to diffuse and react with other species. In this case, the reactive intermediate must be formed by a unimolecular reaction, since it cannot diffuse to meet another reactant, and typically light-induced photodecompositions are used. An example of the successful use of matrix isolation to generate and observe benzyne is shown in reaction (7).

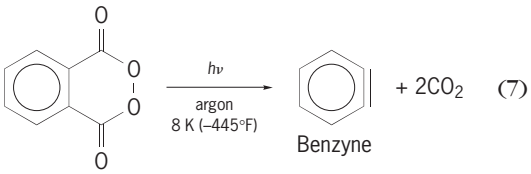

(7)

Isolating benzyne and collecting spectroscopic evidence for its structure under matrix isolation conditions provides evidence that it could be involved as an intermediate in the same reaction under other conditions. At room temperature in solution, benzyne reacts with a variety of substrates to form cycloaddition products, such as in reaction (8). The

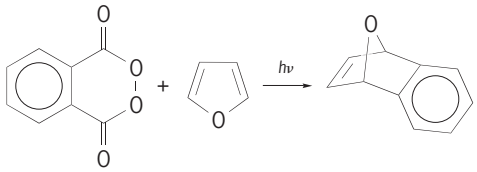

(8)

observation of these same products upon warming a frozen matrix containing benzyne plus the added substrate helps to verify its reactivity patterns and strengthens its implication as a reaction intermediate. *See* MATRIX ISOLATION.

Substituent effects and Hammett equation. By far the major methods for determining mechanisms of reactions as they occur under normal laboratory conditions utilize kinetics and stereochemistry, as already illustrated for the differentiation of S_N1 and S_N2 substitution mechanisms. Variations in the observed kinetics as the structure of the reactants are systematically varied are the basis of one of the most powerful methods of detecting details of a reaction mechanism. Correlation of substituent effects by the Hammett equation is a standard approach that allows the prediction of the extent of electron shifting in a reaction mechanism. Substituents are assigned substituent constants (σ) that indicate the degree to which they are electron-donating (negative values) or electron-withdrawing (positive values) in a standard reaction (the dissociation of benzoic acid). Reactions then display a dependence on the substituent constants, a reaction constant (ρ) that indicates the degree to which the reaction is favored by electron donation (negative values) or favored by electron withdrawal (positive values). The negative ρ value indicates that the reaction is favored by substituents that donate electron density, a point that would not have been obvious from looking at the reactants and products, all of which are uncharged. This evidence suggests that the transition state is polarized. In that transition state, electron donation provides a stabilizing effect, lowers the energy of the transition state, and thereby increases the rate. *See* SUBSTITUTION REACTION.

Isotope effects and isotopic labeling. The use of isotopes is an important tool in mechanistic studies. Kinetic isotope effects are often subtle but useful methods to distinguish details of the transition state. The magnitude of the kinetic isotope effect provides a measure of the amount of bond breaking at the transition state. Isotopes are also used simply as labels to distinguish otherwise identical atoms. For example, the use of ^{13}C isotopic labeling in reaction (9), where

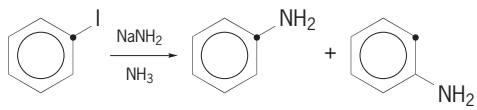

(9)

the heavy dot represents the position of a ^{13}C atom, provided evidence that the mechanism was not a simple nucleophilic substitution. The appearance of ^{13}C equally to two different positions of the product is evidence for a symmetrical intermediate such as benzyne. *See* DEUTERIUM; ISOTOPE.

Theoretical correlations. The theory underlying organic chemistry has developed to the stage that both qualitative and quantitative approaches often provide excellent insight into the workings of a reaction mechanism. The principle of conservation of orbital symmetry, developed by R. B. Woodward and R. Hoffmann and often called the Woodward-Hoffmann rules, provides a simple yet powerful method for predicting the stereochemistry of concerted reactions. The principle states that the formation of new bonds and the breaking of old bonds will be carried out preferentially in a manner that maximizes bonding at all times. *See* PERICYCLIC REACTION; WOODWARD-HOFFMANN RULE.

By knowing the symmetry properties of the molecular orbitals of reactants and products, the preferred (or allowed) pathway can be predicted. The approach can be greatly simplified by recognizing the cyclic delocalized transition state of the reaction and counting the number of electrons involved in the cyclic transition state. If there are $(4n + 2)$ electrons (that is, 2, 6, 10, . . . electrons), the reaction is considered favorable, because all of the bonding can occur on the same sides of the molecules. Thus cycloadditions such as the Diels-Alder reaction (10) are

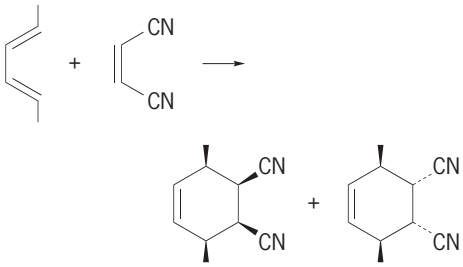

(10)

favorable according to the Woodward-Hoffmann rules, because they involve six electrons in the delocalized transition state.

Reactions involving $4n$ electrons, such as the cycloaddition of two double bonds to form a four-membered ring, are unfavorable by these rules and are not observed except in photochemical reactions, for which the Woodward-Hoffmann rules are exactly reversed. *See* DELOCALIZATION; DIELS-ALDER REACTION; MOLECULAR ORBITAL THEORY; PHOTOCHEMISTRY.

The improvement of computing capabilities has allowed quantitative calculations to become accurate enough to predict energies and structures for simple molecules. Quantum-mechanics calculations may be performed either from first principles, that is, without simplifying assumptions, or semiempirically, that is, with some standardizing parameters. The desired result is a potential energy map that correlates the energy levels of the possible structures of the reactants as they transform to products. Such approaches can help to rule out possible pathways as too high in energy or can suggest alternative pathways that appear feasible based on calculated energies. *See* COMPUTATIONAL CHEMISTRY; MOLECULAR MECHANICS; QUANTUM CHEMISTRY.

Carl C. Wamser

Bibliography. B. K. Carpenter, *Determination of Organic Reaction Mechanisms*, 1984; R. D. Guthrie and W. P. Jencks, IUPAC recommendations for the representation of reaction mechanisms, *Acc. Chem. Res.*, 22:343–349, 1989; T. H. Lowry and K. S. Richardson, *Mechanism and Theory in Organic Chemistry*, 3d ed., 1987; J. March, *Advanced Organic Chemistry*, 5th ed., 2001.

Organic synthesis

The making of an organic compound from simpler starting materials.

Role in chemistry. Organic synthesis plays an important role in chemistry, biochemistry, medicine, agriculture, molecular biology, physics, materials science, electronics, and engineering by allowing for the creation of specific molecules for scientific and technological investigations. In some cases the target molecule has an unusual structure whose characterization may advance understanding of various theoretical aspects of chemistry. Such a molecule may possess particularly unusual patterns of bonding, such as a strained ring system or unique symmetry; examples are cubane (**1**), prismane (**2**), and dodecahedrane (**3**). Once a structurally unusual molecule

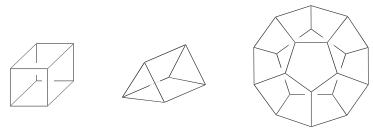

(**1**) (**2**) (**3**)

has been synthesized, its properties can be studied to elucidate concepts of chemical bonding, spectroscopy, and reactivity. *See* CHEMICAL BONDING; ORGANIC CHEMISTRY; STERIC EFFECT (CHEMISTRY).

There are other reasons for synthesizing a particular molecule. For example, a molecule may be isolated as a natural product from some obscure organism, and a synthesis of the molecule in the laboratory can provide larger quantities of it for testing for properties that might indicate its usefulness, such as in medicine or agriculture. Another reason for

synthesizing a molecule is because its structure is different from any known compound and the molecule is predicted to possess some useful property, such as selective ion-binding ability or biological activity such as antibiotic, herbicidal, insecticidal, anticancer, or antiviral activity. In the study of polymers, a particular monomeric subunit might be chosen for synthesis because, upon polymerization, it might yield a polymer having useful physical properties. By hypothesizing that a certain molecule will have a particular property, then obtaining that molecule by synthesis, a chemist can test that hypothesis and thus obtain a better understanding of the relationship between molecular structure and function, and at the same time producing—if the synthesized molecule does possess the desired property—a novel chemical that can be used to better the human condition. *See* POLYMER; POLYMERIC COMPOSITE; POLYMERIZATION.

Finally, a chemist might devise and carry out the synthesis of an already-known compound in order to introduce one or more radioactive isotopes of some of the constituent elements into the product. Such labeled molecules are useful in biology for studying metabolic processes and for isolating enzymes and proteins that bind specifically to that molecule. *See* RADIOISOTOPE (BIOLOGY).

Synthetic strategy. The heart of organic synthesis is designing synthetic routes to a molecule. Organic synthesis can be compared with architecture and construction, where the chemist must devise a synthetic route to a target molecule (a "blueprint"), then utilize a repertoire of organic reactions (the "tools") to complete the "construction project." Along the way, the synthetic chemist must make extensive use of analytical techniques for purifying and characterizing intermediate products as well as the final product. *See* CHROMATOGRAPHY; CRYSTALLIZATION; DISTILLATION; INFRARED SPECTROSCOPY; MASS SPECTROMETRY; NUCLEAR MAGNETIC RESONANCE (NMR).

The simplest synthesis of a molecule is one in which the target molecule can be obtained by submitting a readily available starting material to a single reaction that converts it to the desired target molecule. However, in most cases the synthesis is not that straightforward; in order to convert a chosen starting material to the target molecule, numerous steps that add, change, or remove functional groups, and steps that build up the carbon atom framework of the target molecule may need to be done.

Retrosynthetic analysis. A systematic approach for designing a synthetic route to a molecule is to subject the target molecule to an intellectual exercise called a retrosynthetic analysis. This involves an assessment of each functional group in the target molecule and the overall carbon atom framework in it; a determination of what known reactions form each of those functional groups or that build up the necessary carbon framework as a product; and a determination of what starting materials (synthetic precursors or synthetic intermediates) for each such reaction are required. The resulting starting materials are then subjected to the same retrosynthetic analysis, thus working backward from the target molecule until starting materials that are commercially available (or available by synthesis following an already published procedure) are derived.

The retrosynthetic analysis of a target molecule usually results in more than one possible synthetic route. It is therefore necessary to critically assess each derived route in order to chose the single route that is most feasible (most likely to proceed as written, with as few unwanted side reactions as possible) and most economical (involving the fewest steps and using the least expensive starting materials). The safety of each possible synthetic route (the toxicity and reactivity hazards associated with the reactions involved) is also considered when assessing alternative synthetic routes to a molecule.

Stereoselectivity. Selectivity is an important consideration in the determination of a synthetic route to a target molecule. Stereoselectivity refers to the selectivity of a reaction for forming one stereoisomer of a product in preference to another stereoisomer. Stereoselectivity cannot be achieved for all organic reactions; the nature of the mechanism of some reactions may not allow for the formation of one particular configuration of a chiral (stereogenic) carbon center or one particular geometry (cis versus trans) for a double bond or ring. When stereoselectivity can be achieved in a reaction, it requires that the reaction proceed via a geometrically defined transition state and that one or both of the reactants possess a particular geometrical shape during the reaction. For example, if one or both of the reactants is chiral, the absolute configuration of the newly formed stereogenic carbon center can be selected for in many reactions. Nucleophilic substitution is an example of a reaction that can proceed stereoselectively when the starting material is chiral. Pericyclic reactions also proceed stereoselectively, because they involve transition states that have well-defined geometries. The achievement of stereoselectivity is an important aspect of organic synthesis, because usually a single stereoisomer of a target molecule is the desired goal of a synthesis. Sometimes the target molecule contains a chiral (stereogenic) carbon center; that is, it can exist as either of two possible enantiomers. The possible synthetic routes to the target molecule may not be selective for forming a single enantiomer of the target molecule; each would form a racemic mixture. In some cases, such nonstereoselective synthetic routes to a molecule are acceptable. However, if a synthesis of a single stereoisomer of a target molecule is required, the stereoselectivity of the reactions derived during the retrosynthetic analysis would need to be considered. The development of stereoselective reactions is an active area of research in organic synthesis. *See* ASYMMETRIC SYNTHESIS; ORGANIC REACTION MECHANISM; PERICYCLIC REACTION; RACEMIZATION; STEREOCHEMISTRY.

Chemoselectivity. This term refers to the ability of a reagent to react selectively with one functional group in the presence of another similar functional group. An example of a chemoselective reagent is a reducing agent that can reduce an aldehyde and not a ketone. In cases where chemoselectivity

cannot be achieved, the functional group that should be prevented from participating in the reaction can be protected by converting it to a derivative that is unreactive to the reagent involved. The usual strategy employed to allow for such selective differentiation of the same or similar groups is to convert each group to a masked (protected) form which is not reactive but which can be unmasked (deprotected) to yield the group when necessary. The development and usage of protecting groups is an important aspect of organic synthesis.

Other strategies. Most target molecules for synthesis are complicated, and their syntheses require many steps. An important aspect of the strategy for synthesizing complex molecules is to devise a convergent synthesis, where relatively large subunits of the target molecule are synthesized individually and then attached together to form the complete or nearly complete target. A convergent synthesis strategy is more economical than the alternative linear synthesis strategy, where the target molecule is built up step by step, one group at a time.

Synthetic reactions. A retrosynthetic analysis can be done by using sophisticated computer programs in order to derive as comprehensive a list of possible synthetic routes to a target molecule as possible. One important aspect of synthetic planning is that it depends upon a knowledge (by a chemist, or programmed into a computer program) of known synthetic transformations, reactions that build up carbon skeletons or introduce functional groups or interconvert functional groups. A large variety of organic reactions that can be used in syntheses are known. They can be categorized according to whether they feature a functional group interconversion or a carbon-carbon bond formation.

Functional group interconversions (**Table 1**) are reactions that change one functional group into another functional group. A functional group is a nonhydrogen, non-all-singly-bonded carbon atom or group of atoms. Included in functional group interconversions are nucleophilic substitution reactions, electrophilic additions, oxidations, and reductions. *See* COMPUTATIONAL CHEMISTRY; ELECTROPHILIC AND NUCLEOPHILIC REAGENTS; OXIDATION-REDUCTION; OXIDIZING AGENT; SUBSTITUTION REACTION.

Carbon-carbon bond-forming reactions (**Table 2**) feature the formation of a single bond or double bond between two carbon atoms. This is a particularly important class of reactions, as the basic strategy of synthesis—to assemble the target molecule from simpler, hence usually smaller, starting materials—implies that most complex molecules must be synthesized by a process that builds up the carbon skeleton of the target by using one or more carbon-carbon bond-forming reactions.

Nucleophilic reactions. An important feature of many common carbon-carbon bond-forming reactions is the intermediacy of a carbanionic intermediate, a molecule bearing a carbon-metal bond formed by deprotonation, by a strong base, of a carbon-hydrogen bond that is relatively acidic because of a nearby electron-withdrawing group or because of the insertion of a metal into a carbon-halogen bond. Such carbanionic intermediates are good nucleophiles (electron-rich, partly negatively charged centers) and thus react readily with added aldehydes, alkyl halides, esters, or other electrophilic (electron-poor, partly positively charged) carbon centers to form the carbon-carbon bond. For example, Grignard reagents are formed by the reaction of an organohalide with magnesium metal in ether solvents in the absence of water. Grignard reagents react with aldehydes or ketones by adding the nucleophilic carbon bound to the magnesium to the electrophilic carbonyl carbon of the aldehyde functional group (Table 2). The resulting magnesium alkoxide intermediates will, upon subsequent addition of an acid (usually dilute aqueous hydrochloric acid), undergo protonation to yield neutral alcohol products whose carbon frameworks bear both the carbons of the aldehyde or ketone and the carbons of the organohalide. *See* REACTIVE INTERMEDIATES.

By changing the metal associated with a carbanionic reagent, the reactivity of the reagent can be altered for synthesis. For example, an organocopper (Gilman) reagent will react more readily in nucleophilic substitution reactions with alkyl halides than will organomagnesium (Grignard) reagents (Table 2). Gilman reagents, which are lithium diorganocopper salts, can be formed by the reaction of organohalides, in the absence of water and oxygen, with lithium metal to form the organolithium reagent (RLi), followed by the addition of copper(I) iodide to form the lithium diorganocopper reagent. Subsequent addition of an alkyl halide results in a nucleophilic substitution reaction, where the carbon bound to the copper displaces a halide ion to form a coupled product consisting of the carbon framework of the organohalide precursor to the Gilman reagent connected to the carbon framework of the alkyl halide by a carbon-carbon single bond. *See* ORGANOMETALLIC COMPOUND.

Addition reactions. The Grignard reaction is an example of an addition reaction, where the carbanionic reagent adds itself to the carbon-oxygen double bond of the aldehyde or ketone. A similar addition reaction between a carbanionic reagent and a carbonyl group is the aldol addition reaction (Table 2), where the carbanionic reagent is an enolate formed by the removal, by a strong base, of a hydrogen from a carbon that is bound to a carbonyl group. The electropositive nature of the carbonyl group enhances the acidity of the carbon-hydrogen (C-H) bond next to it, thus allowing such a deprotonation (removal of the hydrogen) to occur, as shown in reaction (1).

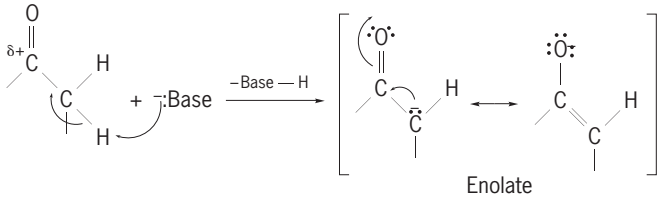

(1)

See GRIGNARD REACTION.

TABLE 1. Examples of some functional-group interconversions

General equation for the reaction*	Net transformation (name)
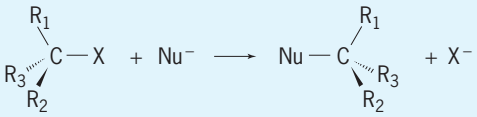 (X = Cl, Br, I, or OSO_2R; Nu = OH, OR, CN, NR_2, others)	Alkyl halide to various functional groups (alcohols, ethers, nitriles, amines, others)
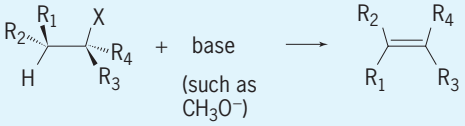	Alkyl halide to alkene (elimination)
$ROH + HX \longrightarrow RX$ (X = Cl, Br, I)	Alcohol to alkyl halide
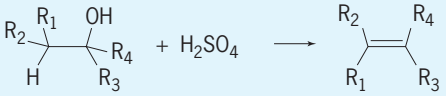	Alcohol to alkene (dehydration)
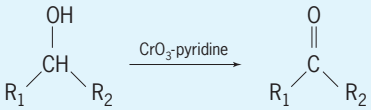	Oxidation of alcohol to ketone or aldehyde
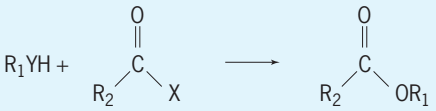 (Y = O or N; X = OH, Cl, others)	Alcohol and carboxylic acid derivative to ester (esterification); amine and carboxylic acid derivative to amide
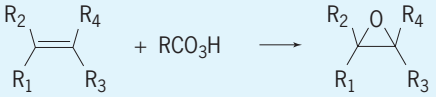	Alkene to epoxide (epoxidation)
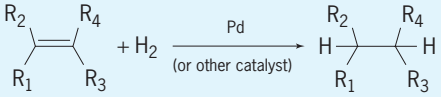	Alkene to alkane (hydrogenation)
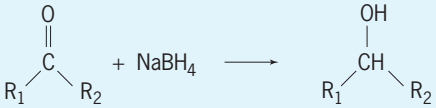	Reduction of ketone or aldehyde to alcohol
$R_1COOR_2 \xrightarrow[\text{2. } H_2O \text{ workup}]{\text{1. LiAlH4}} R_1CH_2OH + R_2OH$	Reduction of ester to two alcohols
$R_1COOR_2 \xrightarrow{H_2O\text{, acid or base}} R_1COOH + R_2OH$	Ester to carboxylic acid and alcohol (ester hydrolysis)
$R-CN \xrightarrow[\text{2. } H_2O \text{ workup}]{\text{1. LiAlH4}} RCH_2NH_2$	Reduction of nitrile to amine
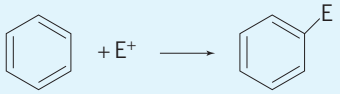 (E = Br, NO_2, R, RCO, others)	Benzene to substituted benzene (electrophilic aromatic substitution)

*R = any organic group (alkyl, aryl, alkenyl) or a hydrogen atom. Nu = nucleophile.

TABLE 2. Examples of some carbon-carbon bond-forming reactions

General equation for the reaction	Name of reaction
$R_1X + Mg \longrightarrow R_1MgX \xrightarrow[2.\ H^+]{1.\ R_2COR_3} R_1-C(OH)(R_2)-R_3$ (X = Cl, Br, I)	Grignard
$R_1X \xrightarrow[2.\ CuI]{1.\ Li} (R_1)_2CuLi \xrightarrow{R_2X} R_1-R_2$ (X = Cl, Br, I)	Gilman
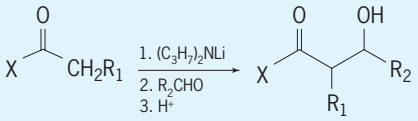 (X = R, RO, NR_2, others)	Aldol addition
$XCH_2CO-Y + RCHO \xrightarrow[2.\ H_2O\ workup]{1.\ Zn} RCH(OH)-CH_2CO-Y$ (X = Cl, Br, or I; Y = R, RO, NR_2, others)	Reformatsky
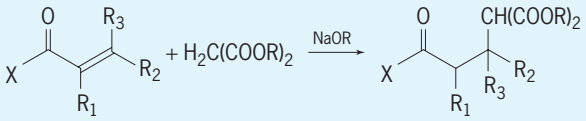 (X = R, RO, NR_2, others)	Michael addition
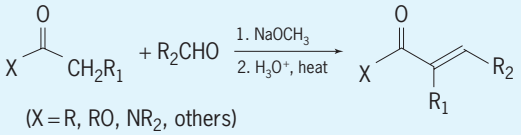 (X = R, RO, NR_2, others)	Aldol condensation
$R_1COOR_2 + R_3CH_2COOR_2 \xrightarrow{NaOR_2}$ 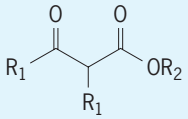	Claisen condensation
$R_1CH_2Br \xrightarrow[2.\ BuLi]{1.\ PPh_3} R_1-CH=PPh_3 \xrightarrow{3.\ R_2CHO} R_1CH=CHR_2$	Wittig
$2RCHO \xrightarrow[2.\ H_2O\ workup]{1.\ Ti(I)} R-CH(OH)-CH(OH)-R$	Pinacol coupling
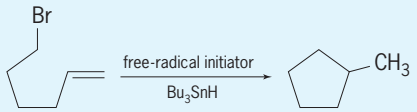	Free-radical cyclization
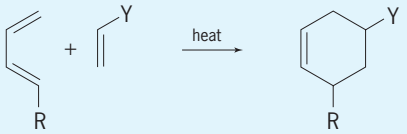 (Y = COOR, COR, CN, others)	Diels-Alder
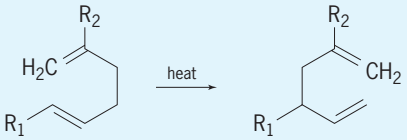	Cope rearrangement

Enolates can be formed by using lithium amide bases (such as lithium diisopropylamide) or, in equilibrium with the carbonyl precursor (most often a ketone, ester, or amide), by using alkoxide bases (such as sodium methoxide). Their structures are resonance hybrids between two resonance structures, one having the negative charge of the enolate centered on the oxygen and the other having the negative charge centered on the carbon. Most of the reactions of enolates used in synthesis involve the anionic carbon atom of the enolate acting as a nucleophile. Enolates can be formed by using weaker bases when the enolate precursor bears two carbonyl groups. *See* RESONANCE (MOLECULAR STRUCTURE).

In the case of the aldol addition reaction, the enolate formed by deprotonation of a carbonyl precursor is allowed to react with an aldehyde; the enolate reacts like a Grignard reagent, adding to the carbonyl carbon atom of the aldehyde to form an alkoxide intermediate. Subsequent addition of water results in the protonation of the alkoxide, thus producing a beta-hydroxycarbonyl product, as shown in reaction (2).

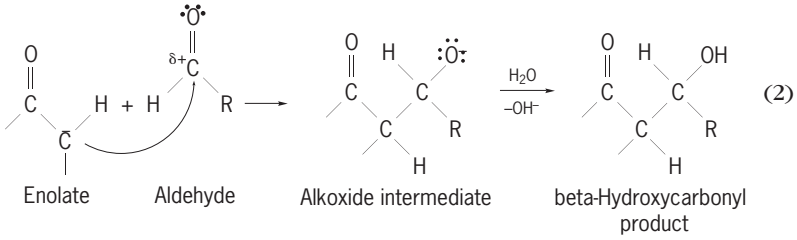

(2)

A similar addition reaction is the Reformatsky reaction (Table 2), where a zinc enolate is formed by the insertion of a zinc metal atom into the carbon-halogen bond of an alpha-halocarbonyl precursor. The zinc enolate is then allowed to react with an aldehyde, adding to it in a manner analogous to the Grignard reaction and aldol addition reaction to yield a zinc alkoxide intermediate which, upon subsequent addition of water, undergoes protonation (addition of a hydrogen) to yield a beta-hydroxycarbonyl product, as shown in reaction (3).

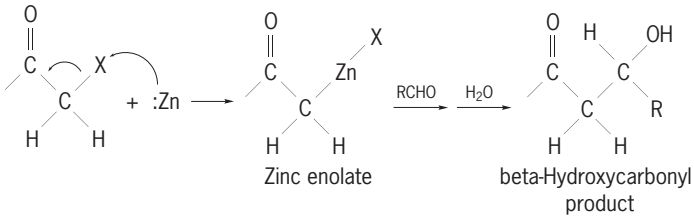

(3)

The Michael addition reaction (Table 2) is similar to the aldol addition in that it involves the formation of an enolate, which then acts as a nucleophile. In the Michael reaction, however, the enolate is formed by deprotonation of a carbon next to two carbonyl groups (usually a diester), and the reaction is typified by the addition of the enolate to a carbon-carbon double bond that is activated by being attached to a carbonyl group (an alpha, beta-unsaturated system). The double bond is polarized by the carbonyl group so that it is electrophilic enough to react with the nucleophilic enolate, thus forming an intermediate enolate which is subsequently protonated to yield an addition product, as shown in reaction scheme (4).

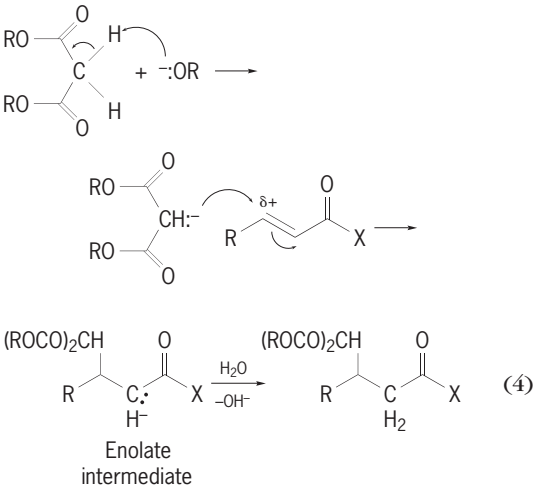

(4)

Condensation reactions. Numerous carbon-carbon bond-forming reactions are classified as condensation reactions, referring to the fact that the reaction combines the two reactants with a net loss of a small fragment, usually water or an alcohol by-product. Condensation reactions typically proceed in a manner similar to that of the Grignard reaction, the aldol reaction, and the Michael addition reaction, involving the initial addition of a carbanionic intermediate to a carbonyl group. The addition step is followed by the loss of the water (as hydroxide) or alcohol (as an alkoxide). For example, the aldol condensation (Table 2) proceeds initially in the same manner as the aldol addition reaction, an enolate adding to an aldehyde to form a beta-hydroxycarbonyl intermediate, but subsequent addition of an acid to the reaction mixture causes the beta-hydroxyl group to eliminate (a dehydration reaction; Table 1) to form an alkene product (specifically, an alpha, beta-unsaturated carbonyl compound). The Claisen condensation (Table 2) involves the addition of an enolate to an ester to form initially an alkoxide intermediate similar to the intermediates formed by enolates adding to an aldehyde. However, this alkoxide intermediate subsequently undergoes an eliminationlike reaction where the alkoxy group of the ester is displaced by the negative charge of the alkoxide to form a ketone carbonyl group; thus the final product of the Claisen condensation is a beta-ketoester, as shown in reaction (5).

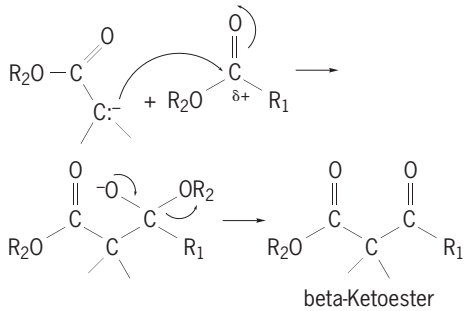

(5)

The Wittig reaction (Table 2) also proceeds by way of a net condensation process. It starts with an alkyl halide, which is first treated with triphenylphosphine to form, by a nucleophilic substitution reaction, an alkyltriphenylphosphonium salt. This salt is then treated, in the absence of water, with a strong base such as *n*-butyllithium, which removes a proton from the carbon attached to the phosphorus atom. As in the case with the formation of enolates, the electropositive nature of the positively charged phosphorus atom enhances the acidity of the C-H bond next to it, allowing the deprotonation to occur. The resulting deprotonated intermediate, which is called a phosphorus ylide, has an anionic carbon center that can add to an electrophile. Thus, when an aldehyde is subsequently added to the reaction mixture, the nucleophilic carbon atom of the ylide adds to the carbonyl carbon to form an alkoxide intermediate. This intermediate then cyclizes to a four-membered oxaphosphetane intermediate, which then undergoes an elimination reaction to form the alkene product along with triphenylphosphine oxide as a by-product, as shown in reaction scheme (6). *See* YLIDE.

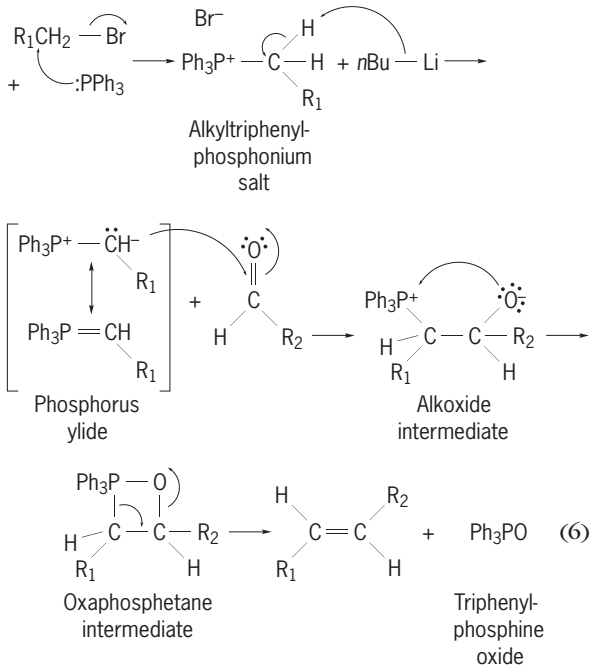

Free-radical intermediates. Some carbon-carbon bond-forming reactions involve free-radical intermediates, where a carbon-centered free radical is an intermediate. The free radicals involved in such reactions are molecules that bear a neutral carbon atom that has three bonds to other atoms and a single, unpaired electron. For example, in the pinacol coupling (Table 2) two aldehyde groups are coupled to each other to form, upon subsequent treatment with water as a proton source, a 1,2-diol product. This reaction requires a reduced metal reagent such as titanium(I), and this metal reagent acts to add an electron to the aldehyde group to form a carbon-centered free-radical metal alkoxide (the metal undergoes oxidation as a result of this electron transfer). The free-radical metal alkoxides then couple to each other (a typical reaction of free radicals) to form a coupled dialkoxide product. Subsequent addition of water results in protonation of the dialkoxide to yield the 1,2-diol product, as shown in reaction scheme (7), where Ph represents the phenyl group (C_6H_5). *See* FREE RADICAL.

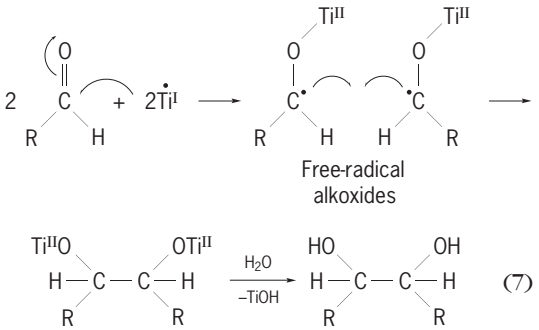

Another common carbon-carbon bond-forming reaction that involves free radical intermediates is the addition of a carbon-centered free radical to an alkene. The form of this reaction that has found greatest utility for organic synthesis is the intramolecular addition of the radical to an alkene group in the presence of tributyltin hydride to form a reduced cyclic product, a cyclization that proceeds most efficiently when it forms a five-membered ring (Table 2). In these cyclizations, the carbon-centered free radical is usually formed by the cleavage of a carbon-bromine bond by a free-radical initiator [usually the isobutyronitrile radical (**5**), which is formed by heating azobisisobutyronitrile, or AIBN (**4**), a commonly used free-radical precursor]. The free-radical initiator abstracts the bromine from the carbon-bromine bond to form the carbon-centered radical, which then attacks the alkene group tethered to it to form a cyclic radical intermediate. This intermediate then abstracts a hydrogen atom from the tin hydride reagent to form the reduced cyclic product and a tin-centered free-radical, which then acts as a radical initiator by abstracting (withdrawing) a bromine atom from another molecule of the starting material. Thus the reaction is a free-radical chain reaction, where free radicals (alternatively, carbon-centered and tin-centered) are continuously produced in a chain, which continues until either the bromine-containing starting material or the tin hydride reagent is consumed, as in reaction scheme (8) where the symbol · identifies a free radical. *See* CHAIN REACTION (CHEMISTRY).

Other reactions. Other carbon-carbon bond-forming reactions proceed in a concerted fashion and do not involve charged or free-radical intermediates. For example, the Diels-Alder reaction (Table 2) is a reaction between a 1,3-diene and an isolated alkene that simultaneously forms two new carbon-carbon bonds and changes the bond order (double bond to single bond, single bond to double bond) of three other carbon-carbon bonds. The Cope rearrangement (Table 2) is a reaction that simultaneously forms and breaks carbon-carbon single bonds while shifting the positions of a carbon-carbon double bond. These

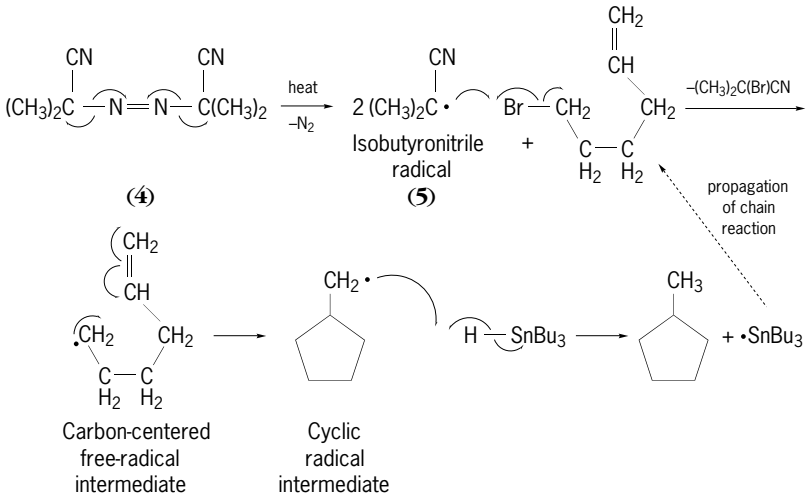

(8)

concerted reactions are examples of pericyclic reactions, a class of reactions that are notable for their stereoselectivity. When a diene is tethered to an alkene group in a starting material, an intramolecular Diels-Alder reaction can proceed, thus forming several rings in a single step. For example, by heating the triene (**6**), the bicyclic product (**7**) can be formed as a single stereoisomer, as shown in reaction (9). Such intramolecular processes are very use-

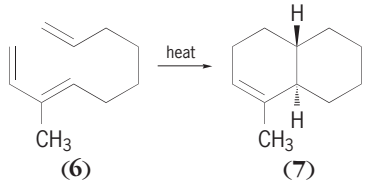

(9)

ful for the synthesis of complex organic molecules. *See* DIELS-ALDER REACTION; WOODWARD-HOFFMANN RULE.

Robert D. Walkup

Protecting groups. A successful chemical synthesis results in a product with a specified structure and set of properties. In designing a chemical synthesis, it is necessary to develop a strategy that will result in a product with these characteristics. In the course of a chemical synthesis, the usual practice is to use compounds with several functional groups. While one of these serves as the reactive center for one reaction, others are saved for later use in the sequence. It is imperative that the conditions required for the desired reaction do not lead to reactions at other groups. Thus, some groups must be protected by first converting them to unreactive derivatives, accomplished with the use of protecting (protective) groups.

A protecting group is a functional group that is attached to a second functional group during a synthesis to prevent interference by or reaction at that site during subsequent chemical transformations. The group should have certain specific properties: (1) It must be possible to install and remove the group in high yield so as not to compromise the efficiency of a synthesis. Every time a protective group is used, it adds two steps to a synthesis. Even if both reactions proceed in 95% yield, the total yield is only 90%. (2) The group should not introduce new stereocenters so that diastereomers are not produced that would complicate analysis by nuclear magnetic resonance and purification. (3) The reagents for its introduction should be low in cost since these ultimately do not add value or carbon to a synthesis. (4) The reagents used for removal of the group should have a high degree of specificity for that group.

Protection of alcohols, phenols, and thiols. Among the protective groups, those for alcohols and amines provide the greatest diversity and will serve to illustrate some more general principles involved in protection and deprotection. Protection of an alcohol usually is accomplished by some modification of a reaction known as the Williamson ether synthesis; that is, the alcohol is treated with a base and a suitable alkylating agent to form an ether. Methods that rely on carbonium ion chemistry are also employed, but they are restricted to substrates that stabilize cations. *See* ALCOHOL.

The greatest variability in protective group chemistry occurs in the methods for deprotection, so that any group in a given molecule can be removed in the presence of the others. This concept is known as orthogonality. Simple ethers or acetals are used as protective groups for alcohols (**Table 3**). The simple ethers are considered the most robust with respect to most synthetic reagents. The methylene acetals (Table 3) offer a large measure of variability in deprotection methods, and consequently greater discrimination is possible when choosing conditions for carbon-carbon bond formation in a synthesis when these are used. *See* ACETAL; ETHER.

A third group consists of silyl ethers ($R_1R_2R_3SiOR$). These have the advantage that they can be introduced and removed with great facility under very mild conditions while maintaining excellent stability to a wide array of synthetic reagents. *See* ORGANOSILICON COMPOUND.

Esters (RCO_2R') are also extensively used to protect alcohols, and they are usually introduced through the acid chloride or anhydride in the presence of a base. Cleavage is normally effected by basic

TABLE 3. Protection for alcohols

Protection group	Deprotection method
Ethers	
$-CH_3$	HI or BBr_3
$-C(CH_3)_3$	Solvolysis in strong acid (CF_3CO_2H)
$-CH_2Ph$	Hydrogenolysis (H_2, Pd−C)
$-CH_2C_6H_4OCH_3$	Oxidation (dichlorodicyanoquinone, DDQ)
$-CH_2C_6H_4-2-NO_2$	Photolysis
$-CH_2CH{=}CH_2$	Isomerization-hydrolysis [$Rh(Ph_3P)_3Cl$; $HgCl_2$, H_2O]
Acetals	
$-CH_2OCH_3$	Strong acid
$-CH_2OCH_2CH_2OCH_3$	Lewis acid ($ZnCl_2$)
$-CH_2OCH_2CH_2TMS$	Fluoride ion (n-Bu)$_4$NF
$-CH_2OCH_2CCl_3$	Reduction with zinc
$-CH_2OCH_2Ph$	Hydrogenolysis (H_2, P−C)
$-CH_2OCH_2SiPh(CH_3)_2$	Oxidation (AcOOH, KBr)
$-CH_2OCH_2C_6H_4OCH_3$	Oxidation (DDQ)
$-CH_2OCH_2CH_2CH_2CH{=}CH_2$	*N*-Bromosuccinimide
$-CH_2OC_6H_4OCH_3$	Oxidation (DDQ)
$-CH_2SCH_3$	$HgCl_2$, $CaCO_3$, H_2O

hydrolysis, and the rate of hydrolysis is dependent upon both electronic and steric effects. By introducing electron-withdrawing groups on the carbon in the alpha position relative to the carbonyl, the rate of base hydrolysis is increased, and this increase can be used to advantage to obtain good cleavage selectivity between several esters. As the crowding or steric demand increases around a particular ester, the rate of hydrolysis decreases because the accessibility to the carbonyl is greatly reduced. For example, the large tertiary butyl ester is much more difficult to hydrolyze with sodium hydroxide than the much smaller methyl ester. As a result, in the case of a substance that contains both these esters only the methyl ester will be hydrolyzed with sodium hydroxide. One advantage of increased steric demand is that it greatly improves the selectivity of protection in polyfunctional substrates. *See* ACID ANHYDRIDE; ESTER.

For alcohol protection, some carbonates have proven to be useful since they are introduced with the ease of an esterification, but can be cleaved by some of the methods used to cleave ethers, thus avoiding the use of basic conditions. Carbonates are generally more stable to base hydrolysis than esters, and thus they can also provide a good level of selectivity in the protection of multiple hydroxyl groups.

Thiol protection is analogous to alcohol protection. The major differences are that with the increased nucleophilicity of a thiol many of the ether-type derivatives are more easily formed, but they are sometimes more difficult to cleave. This is especially true for situations that employ the noble-metal catalysts which are often poisoned with sulfur compounds. As with the methylthio methyl ether, many of the thioether-based protective groups are cleaved with mercury salts. Strong acids such as trifluoroacetic or hydrobromic acid are also frequently used. Protection of a thiol through disulfide formation is an option not available to alcohols. Disulfides are easily prepared by oxidation and cleaved by reduction. *See* CATALYSIS; ORGANOSULFUR COMPOUND.

Protection of amines. Amine protection has its origin in peptide synthesis, where the nucleophilic amine must be rendered temporarily nonnucleophilic in order to activate the carboxylic acid end for coupling with the amine of another amino acid. An amino acid attached to a polymer is coupled with an amino acid protected at the amine and activated at the carboxylate end to form a dipeptide. Deprotection then leads to a new amine, which can then be reacted with a different activated and protected amino acid to form a tripeptide. After the desired number of iterations of the process, the polypeptide is cleaved from the polymer; this step is also a deprotection. *See* PEPTIDE.

The most useful protecting groups for amines are based on carbamates [$R'RNC({=}O)OR''$]. Numerous methods exist for their preparation, with the most common method being through reaction of an amine with a chloroformate or activated carbonate in the presence of a base. The carbamate protecting groups (**Table 4**) exhibit a high degree of orthogonality in that most can be deprotected in the presence of the others. It should be noted that there is a strong correlation between the amine protective groups and the simple ethers used to protect alcohols. Both are often cleaved under the same conditions.

Although it is easy to convert an amine to an amide, most are not useful as protective groups because they are too difficult to hydrolyze. The exception is the trifluoroacetamide, which because of its three strongly electron withdrawing fluorine atoms is relatively easy to hydrolyze. A large number of other methods exist for amine protection, but many of these are of a very specialized nature and are used infrequently. In thinking about amine protection, it is necessary to consider the difference between normal primary and secondary amines and heterocyclic amines such as imidazole and tryptophan. Because of their increased acidity, these are often more easily deprotected. This is especially true of acyl

TABLE 4. Carbamate protective groups

Abbreviation	Structure	Deprotection method
BOC	t-BuOCO—	Strong acid
Fmoc	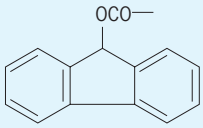	Base
Alloc	$H_2C{=}CHCH_2OCO$—	Pd $(Ph_3P)_4$, nucleophile
Cbz	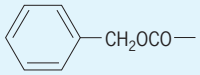	Hydrogenolysis
Troc	CCl_3CH_2OCO—	Reduction with zinc

derivatives which can be cleaved with hydroxide under very mild conditions in contrast to normal amides which are very difficult to hydrolyze. *See* AMIDE; AMINE.

Protection of carbonyls. The most commonly used method for the protection of the carbonyl is by ketal formation. The most common derivatives are the dimethylacetals, 1,3-dioxolanes and 1,3-dioxanes. These are easily prepared by treating a ketone or aldehyde with methanol, ethylene glycol, or 1,3-propane diol respectively and an acid catalyst, while scavenging water either azeotropically with benzene or toluene or chemically with a suitable drying agent. The most frequently employed ketal is the 1,3-dioxolane, for which numerous methods exist for both its introduction and cleavage. Acid-catalyzed hydrolysis is the most commonly employed cleavage method, but many methods that avoid the use of aqueous conditions are also available. For cases where a very robust protective group is required, the carbonyl can be protected as the 1,3-dithiolane or the 1,3-dithiane. These are very resistant to acid cleavage and are usually cleaved with mercury or silver salts, but many other methods have also been developed, the best of which is probably by visible-light photolysis in the presence of the dye methylene green. *See* ALDEHYDE; AZEOTROPIC DISTILLATION; KETONE; PHOTOLYSIS.

Protection of acids. Carboxylic acids are usually protected as an ester (RCO_2R') which serves to tie up the acidic proton, thus avoiding unwanted acid-base reactions. Most commonly the methyl or ethyl ester is used in this capacity. These are easily hydrolyzed with hydroxide. In the event that basic conditions cannot be used to cleave the ester, several other options are available. The allyl and a variety of benzyl esters are often used. These may be cleaved by using conditions similar to the related ethers and carbamates previously discussed. To prevent nucleophilic addition to the carbonyl, sterically hindered esters (that is, esters having bulky substituents near the carbonyl group) such as the *t*-butyl ester are normally used, because they are easily prepared and conveniently cleaved with strong acid. Since hindered esters are not always effective in preventing nucleophilic additions, a second alternative exists: an orthoester can be prepared that completely blocks both the acidity and the carbonyl's susceptibility to nucleophilic attack.

Phosphate protection. This is an exercise in selective deprotection, since phosphoric acid is a trivalent acid. The importance of phosphates has centered on the synthesis of oligonucleotides, that is, gene synthesis. Some of the methods used to protect alcohols and acids such as the use of allyl and various benzyl esters are quite effective. Unlike carboxylic acids, simple phosphate esters are not easily cleaved with mild acid and base. They are often cleaved with trimethylsilylbromide. As the leaving group ability (that is, the ability to stabilize a negative charge) of the ester increases, nucleophilic reagents become much more effective. *See* OLIGONUCLEOTIDE; ORGANOPHOSPHORUS COMPOUND.

Peter G. M. Wuts

Bibliography. F. A. Carey and R. J. Sundberg, *Advanced Organic Chemistry, part B: Reactions and Synthesis*, 4th ed., 2000; W. Carruthers, *Some Modern Methods of Organic Synthesis*, 3d ed., 1986; E. J. Corey and X.-M. Cheng, *The Logic of Chemical Synthesis*, 1989; G. C. Crockett, The chemical synthesis of DNA, *Aldrich Chimica Acta*, 16:47–55, 1983; J. Fuhrhop and G. Penzlin, *Organic Synthesis: Concepts, Methods, Starting Materials*, 2d ed., 1993; T. W. Greene and P. G. M. Wuts, *Protective Groups in Organic Synthesis*, 3d ed., 1999; R. C. Larock, *Comprehensive Organic Transformations: A Guide to Functional Group Preparations*, 2d ed., 1999; R. O. C. Norman and J. M. Coxon, *Principles of Organic Synthesis*, 3d ed., 1993; M. B. Smith, *Organic Synthesis*, 1994; S. Warren, *Organic Synthesis: The Disconnection Approach*, 1982.

Organoactinides

Organometallic compounds of the actinides—elements 90 and beyond in the periodic table. Both the large sizes of actinide ions and the presence of 5*f* valence orbitals are unique features which differ distinctly from most, if not all, other metal ions.

Organometallic compounds have been prepared for all actinides through curium (element 96), although most investigations have been conducted with readily available and more easily handled natural isotopes of thorium (Th) and uranium (U). Organic groups (ligands) which bind to actinide ions include both π- and σ-bonding functionalities. The importance of this type of compound reflects the ubiquity of metal-carbon two-electron sigma bonds in both synthesis and catalysis. *See* CATALYSIS; STRUCTURAL CHEMISTRY.

Synthesis. The most general preparative route to organoactinides involves the displacement of halide by anionic organic reagents in nonprotic solvents. Some examples of the preparation of cyclopentadienyl, C_5H_5 (structure **1**) hydrocarbyl (R = CH_3, C_6H_5, and so forth;), (**2**), pentamethylcyclopentadienyl, $(CH_3)_5C_5$, (**3**), allyl, C_3H_5, (**4**), cyclooctatetraenyl, C_8H_8 (**5**), derivatives are shown in reactions (1)–(5), where M = Th or U. These compounds are

$$MCl_4 + 3Tl(C_5H_5) \longrightarrow M(C_5H_5)_3Cl + 3TlCl \quad (1)$$

(1)

$$M(C_5H_5)_3Cl + RLi \longrightarrow M(C_5H_5)_3R + LiCl \quad (2)$$

(2)

$$MCl_4 + 2(CH_3)_5C_5MgCl \longrightarrow M[(CH_3)_5C_5]_2Cl_2 + 2MgCl_2 \quad (3)$$

(3)

$$MCl_4 + 4C_3H_5MgCl \longrightarrow M(C_3H_5)_4 + 4MgCl_2 \quad (4)$$

(4)

$$MCl_4 + 2K_2C_8H_8 \longrightarrow M(C_8H_8)_2 + 4KCl \quad (5)$$

(5)

all exceedingly reactive and must be rigorously protected from air at all times.

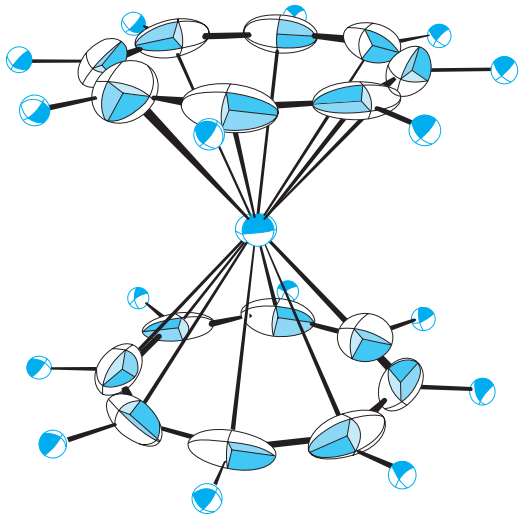

Fig. 1. Molecular structure of $U(C_8H_8)_2$ determined by single-crystal x-ray diffraction. (*After K. O. Hodgson and K. N. Raymond, Inorg. Chem., 12:458, 1973*)

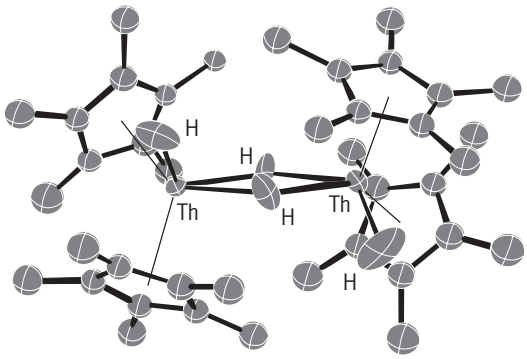

Fig. 2. Molecular structure of dimeric thorium hydride $\{Th[(CH_3)_5C_5]_2H_2\}_2$ as determined by single-crystal neutron diffraction. (*After R. W. Broach et al., Science, 203:172, American Association for the Advancement of Science, 1979*)

Structures. The molecular structures of a number of organoactinides have been determined by single-crystal x-ray and neutron-diffraction techniques. In almost all cases the large size of the metal ion gives rise to unusually high (as compared to a transition-metal compound) coordination numbers. That is, a greater number of ligands or ligands with greater spatial requirements can be accommodated within the actinide coordination sphere. The sandwich complex bis(cyclooctatetraenyl)uranium (uranocene) is an example of this latter type (**Fig. 1**). In most organoactinides, metal-to-ligand atom bond distances closely approximate the sums of the individual ionic radii. This result argues that bonding forces within organoactinides have a large ionic component. Still, a number of spectral (optical, photoelectron, magnetic-resonance, Mössbauer) and magnetic studies reveal that the covalent character of the bonding cannot be ignored, and that there is considerable overlap of metal and ligand orbitals. *See* METALLOCENES; NEUTRON DIFFRACTION; X-RAY DIFFRACTION.

Chemical properties. Studies of organoactinide reactivity have concentrated on understanding the relationship between the identity of the metal, the degree to which the ligands congest the coordination sphere (coordinative saturation), and the types of chemical reactions which the complex undergoes. Employing the methodology of reactions (2) and (3), bis(pentamethylcyclopentadienyl) actinide hydrocarbyl chlorides (**6**) and bishydrocarbyls (**7**) have

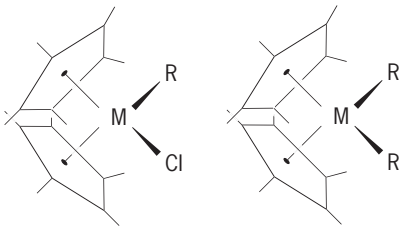

(**6**) (**7**)

been synthesized. These classes of compounds have proved to be some of the most reactive organoactinides, and some of the most informative in terms of revealing new types of reactivity and bonding.

In an atmosphere of hydrogen, thorium-to-carbon and uranium-to-carbon sigma bonds are rapidly cleaved to yield hydrocarbons and the first known organoactinide hydrides, reaction (6). For the dimeric

$$2M[(CH_3)_5C_5]_2R_2 + 2H_2 \longrightarrow \{M[(CH_3)_5C_5]_2H_2\}_2 + 4RH \quad (6)$$

thorium hydride (**Fig. 2**), hydrogen atoms form both terminal (two-center, two-electron) and bridging (three-center, two-electron) bonds to thorium. In solution the organoactinide hydrides are active catalysts for olefin hydrogenation, reaction (7), and for the conversion of C—H bonds in the presence of D_2 to C—D bonds [reaction (8), where M = Th or U].

The high oxygen affinity of actinide ions and the unsaturation of the bis(pentamethylcyclopentadienyl) ligation environment give rise to several unusual new types of carbonylation reactions. The bis(pentamethylcyclopentadienyl) dimethyl derivatives (**8**) react rapidly with CO to produce insertion products in which coupling of four carbon monoxide molecules has occurred to produce two carbon-carbon double bonds and four actinide-oxygen bonds, reaction (9), where M = Th or U. The alkyl chlorides (**9**) undergo an insertion reaction to yield acyl [MC(O)R] complexes (**10**) in which a very strong metal-oxygen interaction takes place, reaction (10). The marked distortion of the bonding in these complexes, away from a classical acyl in which only the carbon atom is bound to the metal, is evident in structural, spectral, and chemical properties. Thus, the metal-oxygen distances are invariably shorter than the metal-carbon distances, the C-O stretching frequencies (about 1460 cm^{-1}) are anomalously low, and the chemistry is decidedly carbenelike. For example, upon heating in solution, the $R = CH_2C(CH_3)_3$ derivative rearranges to a hydrogen atom migration product (**11**), and the $R = CH_2Si(CH_3)_3$ derivative to a trimethylsilyl migration product (**12**). The

$$CH_2{=}CHCH_3 + H_2 \xrightarrow{M-H} CH_3CH_2CH_3 \quad (7)$$

$$C_6H_6 + n/2D_2 \xrightarrow{M-H} C_6D_nH_{6-n} \quad (8)$$

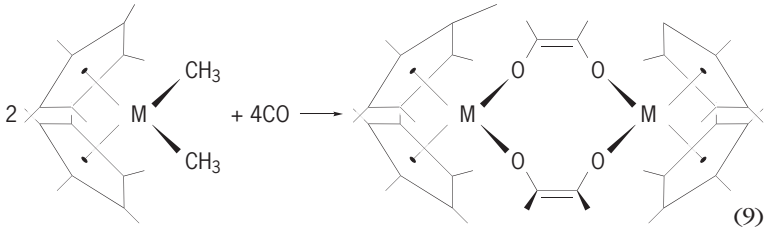

(9)

(8)

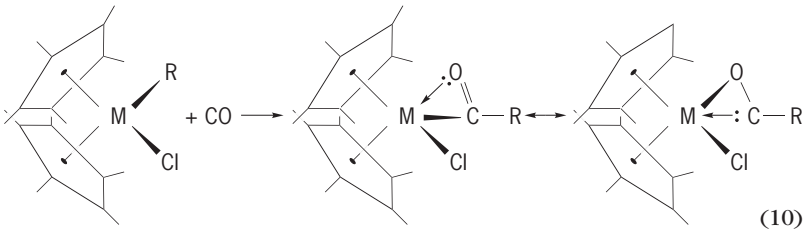

(10)

(9) (10)

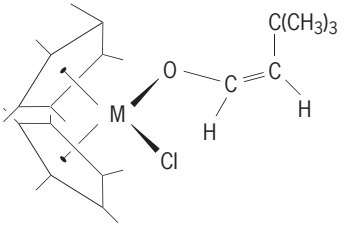

(11)

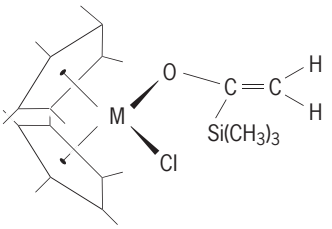

(12)

pronounced oxygen affinity and coordinative unsaturation of the actinide ions in these environments may serve as models for the active surface sites on heterogeneous CO reduction catalysts. *See* ACTINIDE ELEMENTS; COORDINATION CHEMISTRY; HETEROGENEOUS CATALYSIS; HOMOGENEOUS CATALYSIS; ORGANOMETALLIC COMPOUND. Tobin J. Marks

Bibliography. J. Chaiken (ed.), *Laser Chemistry of Organometallics*, 1993; C. Elschenbroich and A. Salzer, *Organometallics: A Concise Introduction*, 2d ed., 1992; T. J. Marks and R. D. Fischer, *Organometallics of the f-Elements*, 1979.

Organometallic compound

A member of a broad class of compounds whose structures contain both carbon (C) and a metal (M). Although not a required characteristic of organometallic compounds, the nature of the formal carbon-metal bond can be of the covalent (**1**; L = ligand), ionic (**2**), or π-bound (**3**) type.

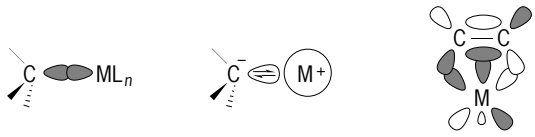

(**1**) (**2**) (**3**)

The term organometallic chemistry is essentially synonymous with organotransition-metal chemistry; it is associated with a specific portion of the periodic table ranging from groups 3 through 11, and also includes the lanthanides. *See* CHEMICAL BONDING; LIGAND; PERIODIC TABLE; TRANSITION ELEMENTS.

This particular area has experienced exceptional growth since the mid 1970s largely because of the continuous discovery of novel structures as elucidated mainly by x-ray crystallographic analyses; the importance of catalytic processes in the chemical industry; and the development of synthetic methods based on transition metals that focus on carbon-carbon bond constructions. From the perspective of inorganic chemistry, organometallics afford seemingly endless opportunities for structural variations due to changes in the metal coordination number, alterations in ligand-metal attachments, mixed-metal cluster formation, and so forth. From the viewpoint of organic chemistry, organometallics allow for manipulations in the functional groups that in unique ways often result in rapid and efficient elaborations of carbon frameworks for which no comparable direct pathway using nontransition organometallic compounds exists.

Early transition metals. In moving across the periodic table, the early transition metals have seen relatively limited use in synthesis, with two exceptions: titanium (Ti) and zirconium (Zr).

Titanium. This transition metal has an important role in a reaction known as the Sharpless asymmetric epoxidation, where an allylic alcohol is converted into a chiral, nonracemic epoxy alcohol with excellent and predictable control of the stereochemistry [reaction (1)]. The significance of the nonracemic

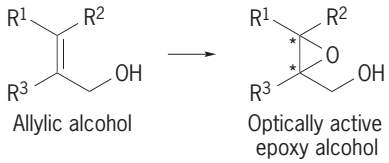

(1)

product is that the reaction yields a single enantiomer of high purity.

An oxophilic element has a strong affinity for oxygen. The ability of oxophilic titanium(IV) to accommodate the necessary chiral ligand [that is, a (+)- or (−)-tartrate ester], a hydroperoxide, and the educt (that is, the starting material) in an octahedral configuration is crucial to chirality transfer that occurs within the dimeric cluster (**4**). Theability to induce

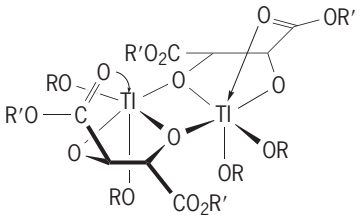

(4)

asymmetry by way of reagents, rather than via existing chirality within a substrate, is extremely powerful, especially if both antipodal forms of the source of chirality are available (as is true with tartaric acid esters). *See* RACEMIZATION.

There are many applications utilizing the Sharpless asymmetric synthesis. Examples of synthetic targets that have relied on this chemistry include riboflavin (vitamin B_2) and a potent inhibitor of cellular signal transduction known as FK-506.

One classic reaction in organic chemistry is the Wittig olefination, where an aldehyde or ketone, upon treatment with a phosphorus ylide, is converted regiospecifically (that is, favoring structural isomer) to an alkene. Such a process, however, is not applicable to a less electrophilic ester carbonyl, the product from which would be an enol ether functional group [reaction (2)]. The oxophilic nature of

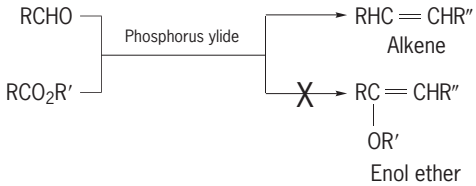

(2)

titanium again has led to a convenient solution to this synthetic problem, in this case coming in the form of the Tebbe reagent (**5**). This organometallic compound can be readily prepared from Me_3Al (Me = CH_3) and Cp_2TiCl_2 (Cp = cyclopentadienyl) in CH_2Cl_2, although in pyridine the starting material can be viewed as the carbene complex (**6**) [reaction (3)]. Either form is an effective methylene

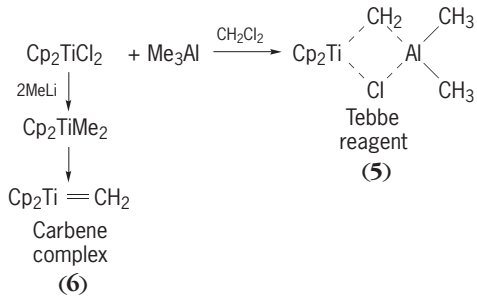

(3)

(5)

(6)

transfer reagent (that is, it donates a CH_2 group) to a variety of carboxylic acid derivatives (RCO-X). *See* ALDEHYDE; ASYMMETRIC SYNTHESIS; TITANIUM; YLIDE.

Zirconium. Below titanium in group 4 in the periodic table lies zirconium, which is evenmore oxophilic than titanium. Most modern organozirconium chemistry concerns zirconium is ready formation of the carbenelike complex [Cp_2Zr:] that because of its mode of preparation, is more accurately thought of as a π complex (**7**). Also important are reactions of the zirconium chloride hydride, $Cp_2Zr(H)Cl$, commonly referred to as Schwartz's reagent, with alkenes and alkynes, and the subsequent chemistry of the intermediate zirconocenes. *See* METALLOCENES.

When the complex Cp_2ZrCl_2 is exposed to two equivalents of ethyl magnesium bromide EtMgBr, the initially formed Cp_2ZrEt_2 loses a molecule of ethane (C_2H_6) to produce the complexed zirconocene [Cp_2Zr: (**7**)], as shown in reaction (4). The term

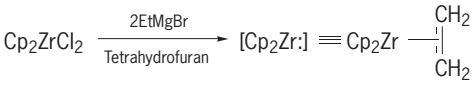

(4)

(7)

β-hydride elimination implies a process whereby a carbon-bound hydrogen with its two electrons (and thus, a hydride) migrates to a metal located on a adjacent carbon, with incipient generation of a π bond between these adjacents carbon atoms. Upon introduction of another alkene or alkyne, a zirconacyclopentane or zirconacyclopentene is formed, respectively [reaction (5)], where the structure above

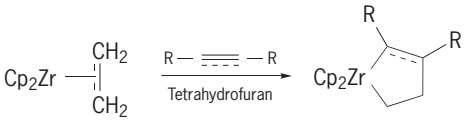

(5)

(7)

the arrow indicates that the chemistry applies to either alkenes (no third bond) or alkynes (with third bond)]. These are reactive species that can be converted to many useful derivatives resulting from reactions such as insertions, halogenations, and transmetalation/quenching [reaction scheme (6)]. When the

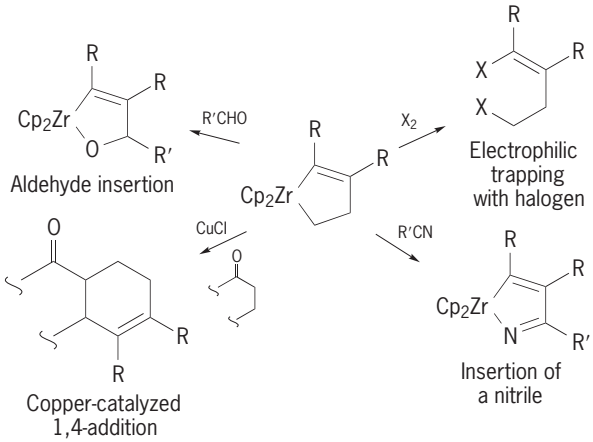

(6)

preformed complex (**7**) is treated with a substrate containing both an alkene and alkyne, a bicyclic zirconacene results that can ultimately yield polycyclic

products as in reaction (7), [for example, pentalenic acid (**8**), a likely intermediate in the biosynthesis of the antibiotic pentalenolactone]. Such an approach has also been used to synthesize dendrobine (**9**), a physiologically active component of the Chinese tonic Chin Shih Hu, and the powerful tumor-promoting agent phorbol (**10**). *See* REACTIVE INTERMEDIATES.

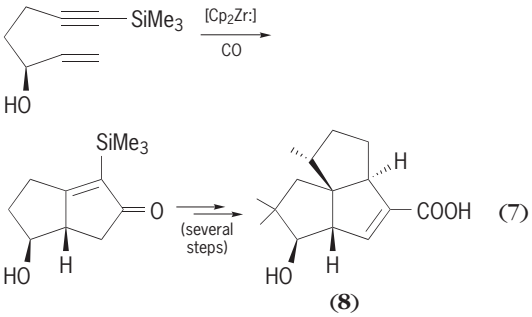

(7)

(**8**)

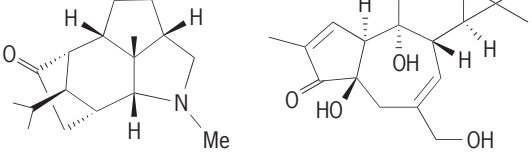

(**9**) (**10**)

Tetrahedral $Cp_2Zr(H)Cl$ readily adds across unsaturated C—C bonds to give alkyl or vinylic zirconocenes, each containing a carbon-zirconium bond. Both are subject to replacement by electrophilic sources of halogen, although perhaps more valuable are their transmetallations to copper (I) and subsequent additions to the β position of an α,β-unsaturated ketone [reaction (8)], or C—C bond-forming reactions, catalyzed by Pd(0), using either vinylic or aryl halides as coupling partners. *See* ZIRCONIUM.

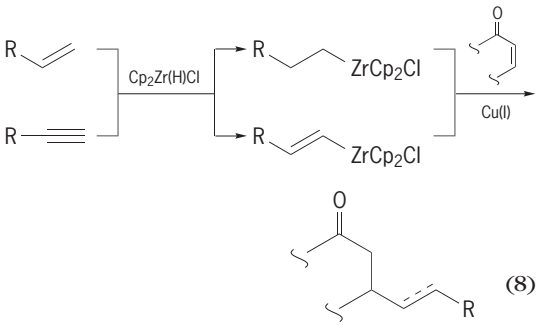

(8)

Mid transition metals. Among the group 6–8 metals, chromium (Cr), molybdenum (Mo), and tungsten (W) have been extensively utilized in the synthesis of complex organic molecules in the form of their electrophilic Fischer carbene complexes, which are species having, formally, a double bond between carbon and a metal. They are normally generated as heteroatom-stabilized species bearing a "wall" of carbon monoxide ligands (**11**).

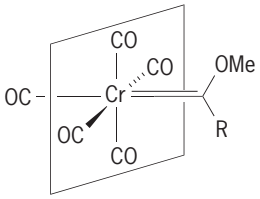

(**11**)

Chromium. Most of the synthetic chemistry has been performed with chromium derivatives, which are highly electrophilic at the carbene center because of the strongly electron-withdrawing carbonyl (CO) ligands on the metal. Many different types of reactions are characteristic of these complexes, such as α alkylation, Diels-Alder cycloadditions of α,β-unsaturated systems, cyclopropanation with electron-deficient olefins, and photochemical extrusions/cycloadditions. The most heavily studied and applied in synthesis, however, is the Dötz reaction. In this reaction, an unsaturated alkoxycarbene is treated with an alkyne thermochemically yielding a hydroquinone following removal of the arene-bound $Cr(CO)_3$ ligand [reaction (9)]. These are remarkable transformations in light of the level of molecular complexity that results from this one-vessel process. Of the many successful applications of this chemistry, those that relate to the aglycones (that is, non-sugar-containing portions) of the antitumor antibiotics in the anthracycline and aureolic acid families have been found to be ideally suited for this chemistry. *See* DIELS-ALDER REACTION; METAL CARBONYL.

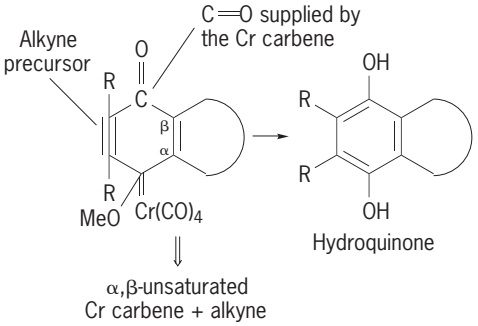

(9)

Another valuable synthetic method involves arenechromium tricarbonyl complexes (**12**), formed from a substituted benzene derivative and $Cr(CO)_6$ or its equivalent. The highly electron-withdrawing tricarbonyl chromium ligand inverts the tendency of the normally electron-rich aromatic ring toward attack by electrophiles, and thus, renders it susceptible to nucleophilic attack. The presence of ring hydrogens results in far greater acidity. As an important element of stereochemical control, the bulk of the $Cr(CO)_3$ residue shields one face of the planar aromatic ring,thereby directing nucleophiles

to approach from the face opposite the metal [reaction (10)].

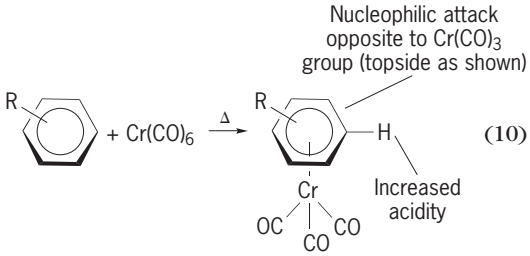

(10)

(12)

See CHROMIUM; ELECTROPHILIC AND NUCLEOPHILIC REAGENTS.

Ruthenium. In its 2+ oxidation state (and hence, a d^6 metal), ruthenium (Ru) forms a highly effective catalyst for asymmetric hydrogenations. The combination of a ruthenium salt such as ruthenium diacetate [$Ru(OAc)_2$] and the bidentate ligand (*R*)- or (*S*)-2,2′-*bis*(diphenylphosphine)-1,1′-binaphthyl (BINAP) leads to octahedral catalysts such as structure (**13**). In the presence of a hydrogen atmosphere and

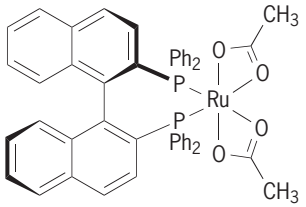

(13)

an alcoholic solvent the combination delivers hydrogen to an olefin with excellent facial selectivity. One case in point is that of an acid [structure (**14**)], where hydrogenation affords the anti-inflammatory agent naproxen (**15**) with 97% optical purity or enantiomeric excess [reaction (11)]. Tolerance to many

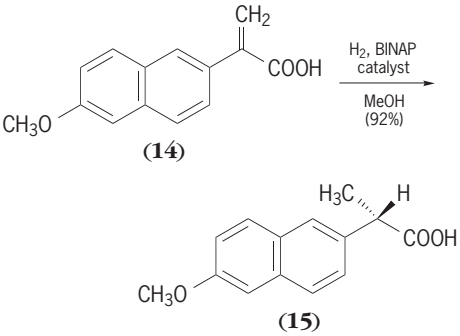

(14)

(15)

(11)

different functional groups exists for example, α,β-unsaturated acids and esters, enamides, allylic alcohols, and unlike most homogeneous catalyst, which are specific for alkene reductions, the octahedral catalyst can likewise reduce a carbonyl moiety as long as a heteroatom is in the α, β, or γ location. *See* CATALYSIS; ELECTRON CONFIGURATION; RUTHENIUM.

Osmium. The OsO_4-catalyzed *cis*-dihydroxylation of an olefin is a well-established means of forming a 1,2-diol. Controlling the facial delivery of the two oxygen atoms to the (prochiral) alkene in an absolute sense, however, has been difficult to achieve. By using both naturally occurring and synthetically modified alkaloid ligands (For example, dihydroquinidine derivatives) as the source of chirality, and potassium ferricyanide [$K_3Fe(CN)_6$] as stoichiometric reoxidant, alkenes of most substitution patterns participate in this Sharpless asymmetric dihydroxylation, shown in reaction scheme (12).

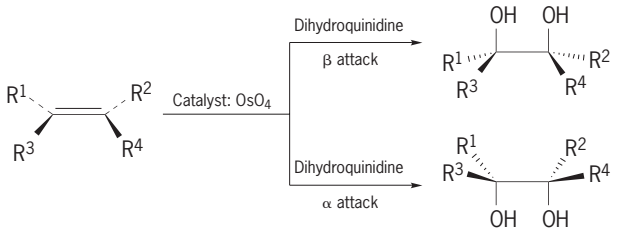

(12)

Many tough mechanistic questions still remain unanswered regarding this remarkably useful process. Nonetheless, as with the Sharpless asymmetric epoxidation, this methodology is so powerful that it was quickly embraced by the synthetic community. *See* PROCHIRALITY; STOICHIOMETRY.

Late transition metals. Groups 9–11 contain transition metals that have been the most widely used not only in terms of their abilities to effect C—C bond formations but also organometallic catalysts for some of the most important industrial processes. These include cobalt (Co), rhodium (Rh), palladium (Pd), and copper (Cu).

Cobalt. When an alkyne is treated with dicobalt octacarbonyl in an ethereal or hydrocarbon medium, two CO ligands are lost and a stable adduct (**16**) is formed [reaction (13)]. The acetylene has become

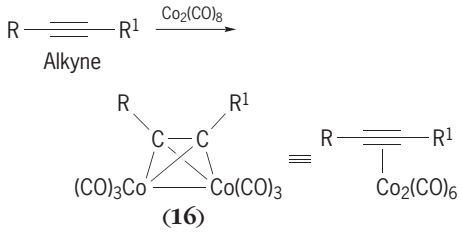

(13)

(16)

a bridging ligand by donating 4π electrons, the hybridization at each carbon now approximating sp^2-like character.

These alkyne derivatives are, in effect, protected from many of the standard reactions of acetylenes, such as hydroboration. Therefore, they posses some unique features that impart synthetic utility. For example, such a complex imparts special stability to a carbocation located on an adjacent carbon (known as the propargylic site), which undergoes coupling with nucleophiles (Nu) at this site. Therefore, in the Nicholas reaction, a propargylic alcohol or ether can be complexed by cobalt to give structure (**17**), then treated with a Lewis acid to generate the corresponding cation (**18**), which is followed by attack of the nucleophile to yield the cobalt-complexed product. Mild oxidation of the metal (for example, with ferric salts) returns the alkyne moiety [reactions (14)]. This type of sequence has proven to be especially valuable en route to the core structure of the enediyne antitumor antibiotics, such as the esperamycins and

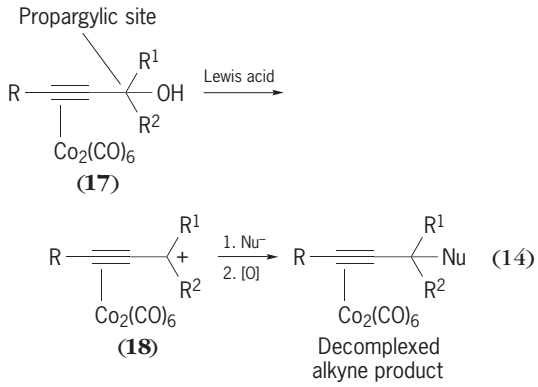

(14)

calicheamycins, regarded as among the most potent naturally occurring antitumor agents known. *See* ACID AND BASE; ANTIBIOTIC.

Another characteristic feature of dicobalt complexes is their participation in ring forming reactions when heated in the presence of an alkene, leading to cyclopentenones (**19**).

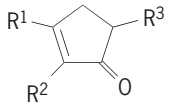

(**19**)

Hydroformylation (the oxo reaction) was discovered in 1938 by O. Roelen. In this type of reaction the components hydrogen (H) and the formyl group (CHO) are formally added across an alkene to arrive at an aldehyde. Indeed, the reaction for the conversion of ethylene to propionaldehyde, as applied commercially to prepare up to C_{15} aldehydes, is the largest-scale industrial process involving a homogeneous catalyst. *See* COBALT; HOMOGENEOUS CATALYSIS.

Rhodium. Just below cobalt in group 9, rhodium is also an extremely effective metal for catalyzing various reactions such as hydrogenations by activation of molecular hydrogen. The species $(Ph_3P)_3RhCl$ (where Ph = C_6H_5), a rhodium(I) derivative referred to as Wilkinson's catalyst, hydrogenates unsaturated C—C bonds, as well as other functional groups, at ambient temperatures and pressures. When $L_2Rh(S)Cl$ (L = a ligand; S = a solvent molecule) is combined with a chiral, nonracemic *bis*-phosphine, the ligands L_2 on Rh are replaced to form a chiral catalyst. Uptake of hydrogen and a reactant such as an alkene into the coordination sphere of the metal leads to an efficient delivery of hydrogen from one preferred face of the olefin, resulting in high levels of stereoinduction; that is, a particular stereochemistry at carbon that was induced or created (in this case) during the hydrogenation of the alkene in the presence of a chiral catalyst. This approach has provided a means of manufacturing the chiral amino acid L-dopa, a clinically useful compound for treatment of Parkinson's disease. *See* PARKINSON'S DISEASE; RHODIUM.

Palladium. There are basically two stable oxidation levels of palladium, Pd(0) and Pd(II). It is the facile transition between these two states that accounts for the C—C bond-forming cross-couplings that have been extensively used to great advantage in organic synthesis. Virtually all reactions that are practical involve catalytic quantities, since palladium metal as well as its numerous salts are relatively expensive. Noteworthy features of palladium chemistry are the remarkable tolerance to functional groups, the breadth of the transformations promoted, and the ready availability and ease of handling of palladium complexes.

The most common source of Pd(0) (a d^{10} metal) is its *tetrakis*(triphenylphosphine) derivative, $Pd(PPh_3)_4$, a yellow crystalline solid with slight sensitivity to air. The species $[L_4Pd(0)]$ at this oxidation level are electron-rich and hence nucleophilic in nature. They readily undergo insertion (that is, oxidative addition) into C_{sp^2}-halogen bonds, the ease of which follows the order iodine > bromine ≫ chlorine. After oxidative addition, which produces a Pd(II) intermediate, another organometallic (R-M) present can transmetallate, that is, exchange its ligand R for the halide on Pd. The initially trans disposition of the two organic groups on the metal can readily isomerize to cis, and then couple (reductively eliminate) to produce the cross-coupled organic product while regenerating the Pd(0) catalyst. There are a number of organometallic nucleophiles that are of general utility in this sequence. An impressive demonstration of the power of this methodology is found in the coupling between C-75 and C-76 present in palytoxin, the toxic principle found in marine soft corals, which is the most poisonous nonpeptidic substance known.

Another common mode of reaction of Pd(0) complexes is their displacement of allylic derivatives such as acetates and carbonates to give rise to an initially σ-bound Pd(II) compound. Isomerization to a π-allyl palladium intermediate occurs that is subject to attack by soft nucleophiles (that is, carbanionic species derived from relatively strong carbon acids), ultimately affording the product of net substitution. Use of this scenario in an intramolecular sense where the π-allyl palladium intermediate is prepared within the same molecule that bears the nucleophile leads to ring-forming reactions. Thus, for allylic carbonate, cyclization gives a key intermediate in the synthesis of (−)-aspochalasin B, a representative of the cytochalasins, which have extraordinary effects on transport across membranes in mammals.

Palladium(II) intermediates R-Pd-X also coordinate alkenes, leading to new σ-Pd-bound intermediates that ultimately undergo β-hydride elimination of H-Pd-X. These Pd(II) species then fall apart to H-X and Pd(0). Thousands of examples are known for this chemistry, usually called the Heck reaction, with many electron-rich or electron-poor olefinic substituents being acceptable.

Olefin insertion into $PdCl_2$ forms the basis of the Wacker process, the industrial route to acetaldehyde from ethylene. Carried out in aqueous media, the alkene is converted to an intermediate that reads to produce the aldehyde and Pd(0). The $CuCl_2$ in the reaction vessel reoxidizes Pd(0) to Pd(II), which reenters the cycle. When applied to substituted alkenes,

ketones are formed in good yields. Alteration of the reaction solvent changes the outcome, thereby leading to many key industrial building blocks. *See* ALKENE; PALLADIUM.

Copper. As a group 11 metal, copper has a valence shell corresponding to $3d^{10}4s^1$. Since virtually all of the organometallic chemistry of copper as it relates to organic synthesis takes place from the 1+ oxidation state [that is, Cu(I)], and hence of d^{10} configuration, there are chemists who would claim that it does not therefore qualify as a transition metal. When copper is considered from the mechanistic perspective, however, where reactions of Cu(I) must be undergoing a transition from Cu(I) to Cu(II) [a 1-electron transfer process involving carbon radical intermedates] or via energetically unfavorable Cu(III) state (implying carbanionlike intermediates), it surely belongs in this classification. From the standpoint of synthetic utility, few would argue with the notion that organocopper complexes are among the most useful transition metal reagents. *See* REAGENT CHEMICALS.

The popularity of organocopper reagents stems from their diversity in reagent makeup and the types of coupling reactions that they facilitate. The combination of two equivalents of an organolithium (RLi) with a copper (I) halide in (ethyloxide) or (tetrahydrofuran) forms cuprates (R_2CuLi), commonly called Gilman reagents. These are Cu(I) monoanionic complexes that are regarded as symmetrical dimers, albeit based on a limited number of solid-state x-ray analyses. Nuclear magnetic resonance spectroscopic studies in solution, however, support this view. The hallmark of their chemistry is the ease with which they undergo 1,4 or conjugate addition reactions with α,β-unsaturated carbonyl compounds rather than competitive 1,2-additions to the carbonyl carbon, as do more basic reagents such as the organolithiums and Grignards reagents. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

Many variations on the nRLi + CuX theme have appeared since the initial report by H. Gilman in 1952. When $n = 1$, the products is RCu (+LiX), which is a relatively unreactive, polymeric material. It can be activated with a Lewis acid to give a reagent also prone toward conjugate addition. With CuCN as the Cu(I) source, RLi + CuCN combine to generate the mixed cyanocuprate RCu(CN)Li, rather than RCu + LiCN. Because of the electron-withdrawing nitrile ligand, which is tightly bound to copper, these mixed cyanocuprates are much less reactive than are Gilman cuprates. The 2RLi + CuCN combination, commonly referred to as higher-order cyanocuprates $R_2Cu(CN)Li_2$, is yet another variation on this theme. These latter species appear to be very reactive toward the usual sorts of electrophilic partners, but they also appear to possess greater thermal stability relative to Gilman reagents. All four reagent types, summarized in reactions (15), have virtues that are routinely called upon in numerous synthetic situations.

Another especially valuable reaction of organocuprates is their cross-couplings with alkyl,allylic,

$$1\ \mathrm{RLi} + \mathrm{CuI} \longrightarrow \mathrm{LiI} + \mathrm{RCu} \xrightarrow{\mathrm{BF_3}} \mathrm{RCu \cdot BF_3} \qquad (15a)$$

$$2\ \mathrm{RLi} + \mathrm{CuI} \longrightarrow \mathrm{R_2CuLi} + \mathrm{LiI} \qquad (15b)$$

$$1\ \mathrm{RLi} + \mathrm{CuCN} \longrightarrow \mathrm{RCu(CN)Li} \qquad (15c)$$

$$2\ \mathrm{RLi} + \mathrm{CuCN} \longrightarrow \mathrm{R_2Cu(CN)Li_2} \qquad (15d)$$

vinylic, and aryl halides, as well as with epoxides. These net substitutions occur via different pathways, and notwithstanding the intense usage of this chemistry, definitive mechanistic schemes remain to be elucidated. Most prevalent are displacements of primary alkyl halides, the order of leaving group ability being $RCH_2I > RCH_2Br > RCH_2Cl$. Other leaving groups based on oxygen are also common, with sulfonates being roughly equal to iodides in reactivity. Oxirane cleavage is particularly valuable, as the alcohol formed on workup remains in the product for further elaboration. Allylic epoxides are highly activated substrates, in this case the R group on copper being transferred to the olefinic carbon with concommitant opening of the oxirane ring. This arrangement of functionality in cyclopentene epoxide allowed for a cyanocuprate to deliver the α chain ultimately needed to complete a synthesis of prostaglandins E_1 and F_{1a}, which are 20-carbon local hormones derived in the organism from the arachidonic acid cascade. *See* EICOSANOIDS; HORMONE.

Much of organocopper chemistry relies on Grignard reagents as an alternative to organolithiums, leading to magnesio cuprates (R_2CuMgX) based on the same stoichiometry: 2RMgX + CuX going to $R_2CuMgX + MgX_2$. These are attractive reagents, given the availability of most Grignard reagents (from the precursor halide R-X + Mg metal). Considering that Grignard reagents readily add to carbonyl groups under similar conditions, implied in these reactions is the in place generation of a highly reactive cuprate (perhaps R_2CuMgX) that undergoes 1,4-addition faster than the RMgX alone can add in a 1,2 fashion to the ketone.

Copper-catalyzed reactions of Grignard reagents are often preferred relative to stoichiometric cuprates as a less costly procedure. The expense factor is even more significant with regard to chemical waste disposal, which became a major issue in the 1990s. Sources of catalytic Cu(I) or Cu(II) for Grignard couplings include not only the usual assortment of salts that have been mentioned for lithio cuprate formation but others such as Li_2CuCl_3 and Li_2CuCl_4, the latter otherwise known as Kochi's catalyst; this is a Cu(II) species that is rapidly reduced to Cu(I) in the presence of RMgX.

As with lithio cuprates, myriad examples exist describing implementation of this coupling approach is synthetic situations. One well-known target is the anti-inflammatory steroid cortisone. An example of a copper catalyzed alkylation of a Grignard reagent can be found in a synthesis of the side chain of vitamin E (α-tocopherol), where isopentyl Grignard displacement of a primary tosylate efficiently produced the desired segment of this

coenzyme. *See* COORDINATION COMPLEXES; COPPER; ORGANIC CHEMISTRY; ORGANIC SYNTHESIS; STEREOCHEMISTRY.

Bruce H. Lipshutz

Bibliography. E. W. Abel, F. G. Stone, and G. Wilkinson (eds.), *Comprehensive Organometallic Chemistry II: A Review of the Literature*, 1982–1994, 2d ed., 1995; J. P. Collman et al., *Principles and Applications of Organotransition Metal Chemistry*, 2d ed., 1987; A. DeMeijere and H. T. Dieck, *Organometallics in Organic Synthesis*, 1988; C. Elschenbroich and A. Salzer, *Organometallics*, 2d ed., 1992; L. S. Hegedus, *Transition Metals in the Synthesis of Complex Organic Molecules*, 2d ed., 1999; L. S. Liebeskind (ed.), *Advances in Metal-Organic Chemistry*, 1991.

Organophosphorus compound

One of a series of derivatives of phosphorus that have at least one organic (alkyl or aryl) group attached to the phosphorus atom linked either directly to a carbon atom or indirectly by means of another element (for example, oxygen). The mono-, di-, and trialkylphosphines (and their aryl counterparts) can be regarded formally as the parent compounds of all organophosphorus compounds; see notation (1). Formal substitution of the hydrogen

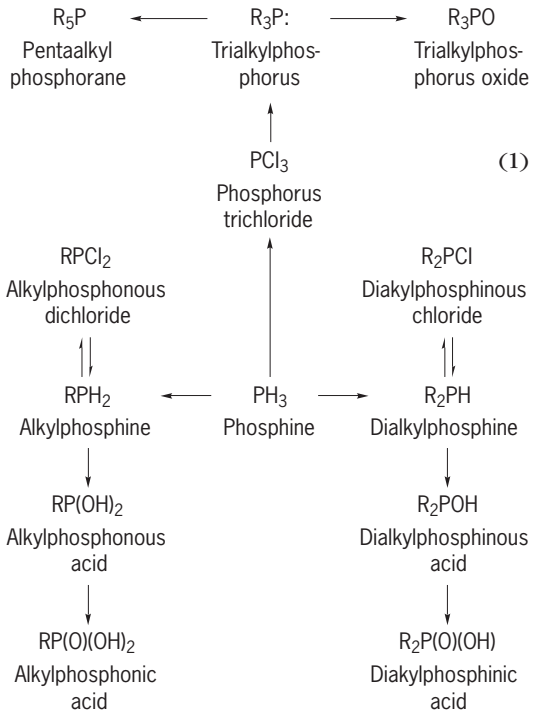

in phosphine (PH_3) by monovalent groups or atoms leads to a number of basic structures; formal addition of bivalent oxygen leads from the mono- and dialkylphosphines to organophosphorus acids and from the trialkylphosphines to their oxides. Many of these organophosphorus molecules have been synthesized in handed (chiral) form, thus serving as stereochemical probes for the study of reaction mechanisms. Formal replacement of the nonbonding electron pair in R_3P: by two substituents gives R_5P. Pentaphenylphosphorane is a stable example in this class. Even hexacoordinate phosphorus (as anionic species) materials are known. There has been considerable research activity devoted to organophosphorus intermediates in atypically bonded structures; see notation (2).

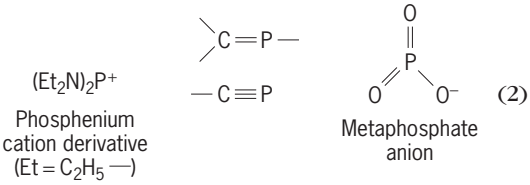

See ORGANIC REACTION MECHANISM; REACTIVE INTERMEDIATES.

Considering the large number of organic groups that may be joined to phosphorus as well as the incorporation of other elements in these materials, the number of combinations is practically unlimited. A vast family in itself is composed of the heterocyclic phosphorus molecules, in which phosphorus is one of a group of atoms in a ring system. *See* HETEROCYCLIC COMPOUNDS.

Preparation. Organophosphorus compounds may be prepared by a variety of methods. Trialkyl esters of phosphorus acid (trialkylphosphites) react with alkyl halides to form dialkyl alkylphosphonates, in the Michaelis-Arbuzov reaction (3). Since

$$P(OR)_3 + R'X \rightarrow R'P(O)(OR)_2 + RX \qquad (3)$$

the preparation involves alkyl halides and depends on the reactivity of the alipathic halide, aromatic organophosphorus compounds can not be obtained in an analogous manner from aryl halides. The preferred method for preparing aromatic compounds is the introduction of a —PCl_2 group into aromatic hydrocarbons by means of PCl_3, with anhydrous aluminum chloride as a catalyst. The resulting arylphosphorus dichlorides ($ArPCl_2$) can react further with the aromatic hydrocarbon to yield the diarylphosphinous chlorides, Ar_2PCl, which are always obtained in addition to the phosphonous dichlorides, but can be made the main product of the reaction by properly choosing reaction conditions.

The reaction of phosphorus halides with organometallic compounds is useful for preparation of trialkyl- or trialphosphines but is less applicable for making compounds with only one P—C bond. Organophosphorus derivatives with one or more P—H bonds can be added across activated alkene or acetylene bonds to form a P—C linkage. The P—H group can also react with carbonyl compounds to give α-hydroxy derivatives. *See* ORGANOMETALLIC COMPOUND.

Reactions. Organophosphorus compounds frequently serve as valuable synthetic reagents. The Wittig reaction (4) is one of the most useful prepar-

$$\underset{\text{Alkylidene phosphorane}}{R_3P{=}CR'_2} + \underset{\text{Carbonyl}}{R''_2C{=}O} \rightarrow \underset{\text{Alkene}}{R''_2C{=}CR'_2} + R_3PO \qquad (4)$$

ative methods known. The alkylidine phosphorane

is usually prepared from a phosphonium salt (R_3P^+-CHR'_2Br^-) and an inorganic or organic base (for example, C_6H_5Li); subsequent treatment with ketones or aldehydes yield alkenes. The thermodynamic driving force for this and many reactions involving organophosphorus is formation of the relatively strong P═O bonds. Polyacetylenes, carotenoids (vitamin A), and steroid derivatives are more synthetically accessible by employment of the Wittig reaction. Both the structure of the phosphorane and reactive substrate have been varied over a wide range of combinations. *See* ORGANIC SYNTHESIS.

Uses. Some organophosphorus compounds have been used as polymerization catalysts, lubricant additives, flameproofing agents, plant growth regulators, surfactants, herbicides, fungicides, and insecticides. Hexamethylphosphoramide, $[(CH_3)_2N]_3PO$, is a remarkable polar solvent used in organic syntheses and capable of forming complexes with organic, inorganic, and organometallic compounds. Cyclophosphamide (CPA) is a phosphorus heterocycle which has been used in the treatment of cancer; it also is an antiinflammatory agent in diseases. Naturally occurring products with a C—P bond have been found (for example, 2-aminoethyphosphinic acid) in protozoa, in marine life, and in certain terrestrial animals. On the other hand, the high mammalian toxicity shown by some methylphosphonic acid derivatives (inhibitors of the enzyme cholinesterase) limits the usefulness of a number of related though much less toxic materials because of potential health hazards. The obvious importance of phosphorus in adenosine triphosphate (ATP) and deoxyribonucleic acid (DNA) in biology is also noteworthy. *See* ADENOSINE TRIPHOSPHATE (ATP); DEOXYRIBONUCLEIC ACID (DNA).

Organophosphorus compounds were made during World War II for use as chemical warfare agents in the form of nerve gases (Sarin, Trilon 46, Soman, and Tabun). *See* INSECTICIDE; PHOSPHORUS.

Sheldon E. Cremer

Bibliography. D. W. Allen and B. J. Walker, *Organophosphorus Chemistry*, vol. 24, 1993; J. I. Cadogan (ed.), *Organophosphorus Reagents in Organic Synthesis*, 1980; R. S. Edmunson (ed.), *Dictionary of Organophosphorus Compounds*, 1988; G. M. Kosolapoff and L. Maier (eds.), *Organic Phosphorus Compounds*, vols. 1–6, 1970–1975; A. D. F. Toy and E. N. Walsh, *Phosphorus Chemistry in Everyday Living*,1987; S. Trippett (ed.), *Organophosphorus Chemistry* vols. 1–19, 1972–1988.

Organoselenium compound

One of a group of compounds that contain both selenium (Se) and carbon (C) and frequently other elements as well, for example, halogen, oxygen (O), sulfur (S), or nitrogen (N). The first examples of organoselenium compounds were reported shortly after the discovery of elemental selenium by J. J. Berzelius in 1818. Their exceptionally unpleasant and persistent odors may have discouraged early investigators. However, less volatile derivatives containing aryl substituents or selenium in higher oxidation states are more easily handled.

Organoselenium compounds have become common in organic chemistry laboratories, where they have numerous applications, particularly in the area of organic synthesis. Organoselenium compounds formally resemble their sulfur analogs and may be classified similarly. For instance, selenols, selenides, and selenoxides are clearly related to thiols, sulfides, and sulfoxides. Despite structural similarities, however, sulfur and selenium compounds are often strikingly different with respect to their stability, properties, and ease of formation. *See* ORGANOSULFUR COMPOUND; SELENIUM.

Selenols and selenides. Selenols have the general formula RSeH, where R represents either an aryl or alkyl group. They are prepared by the alkylation of hydrogen selenide (H_2Se) or selenide salts, as well as by the reaction of Grignard reagents or organolithium compounds with selenium, as in reactions (1). Since they are air-sensitive, they must

$$H_2Se(\text{or } HSe^-) \xrightarrow{RX} \underset{\text{Selenol}}{RSeH} \xrightarrow[2\cdot H_3O^+]{1\cdot Se} \underset{\text{Grignard reagent}}{RMgX} \text{ or } \underset{\text{Organolithium compound}}{RLi} \qquad (1)$$

be stored under an inert atmosphere. Selenols are stronger acids than thiols. They and their conjugate bases are powerful nucleophiles that react with alkyl halides (R′X; R′ = alkyl group) or similar electrophiles to produce selenides (RSeR′); further alkylation yields selenonium salts, as shown in reactions (2), where R″ represents an alkyl group that may be

$$\underset{\text{Selenol}}{RSeH} \xrightarrow{R'X} \underset{\text{Selenide}}{RSeR'} \xrightarrow{R''X} \underset{\text{Selenonium salt}}{R-\overset{\overset{R'}{|}}{Se^+}-R''+X^-} \qquad (2)$$

different from the alkyl group R′. Both acyclic and cyclic selenides are known, but episelenides (three-membered cyclic selenides) are unstable, losing selenium to produce olefins. Hydrogen ions (protons) can be removed from the carbon atom adjacent to the selenium atom of a selenide by treatment with a suitable base, because the selenium atom stabilizes the resulting carbanion. *See* ELECTROPHILIC AND NUCLEOPHILIC REAGENTS; GRIGNARD REACTION; REACTIVE INTERMEDIATES.

Diselenides and selenenyl halides. Diselenides (RSeSeR) are usually produced by the aerial oxidation of selenols; they react in turn with chlorine or bromine, yielding selenenyl halides or selenium trihalides, reactions (3). Selenenyl halides add to olefins [reaction (4) and undergo halide substitution with various nucleophiles. For example, reactions with amines ($R'NH_2$ or R'_2NH), cyanide ion (CN^-), and sulfinates ($R'SO_2^-$) produce selene-namides

($RSeNR'_2$), selenocyanates (RSeCN), and selenosul-

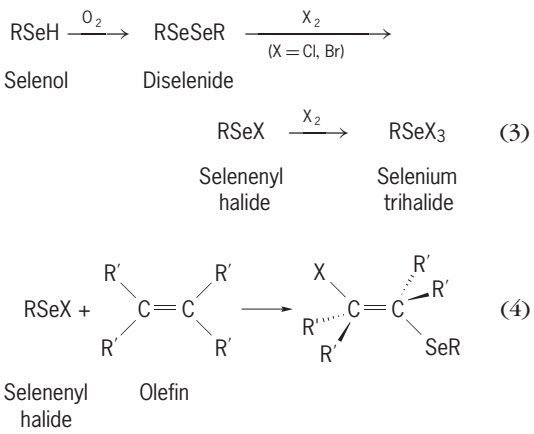

fonates, ($RSeSO_2R'$), respectively. Triselenides and polyselenides are also known.

Selenenic, seleninic, and selenonic acids. Reaction scheme (5) shows the pathways in which diselenides can be oxidized to selenonic acids. Diselenides are

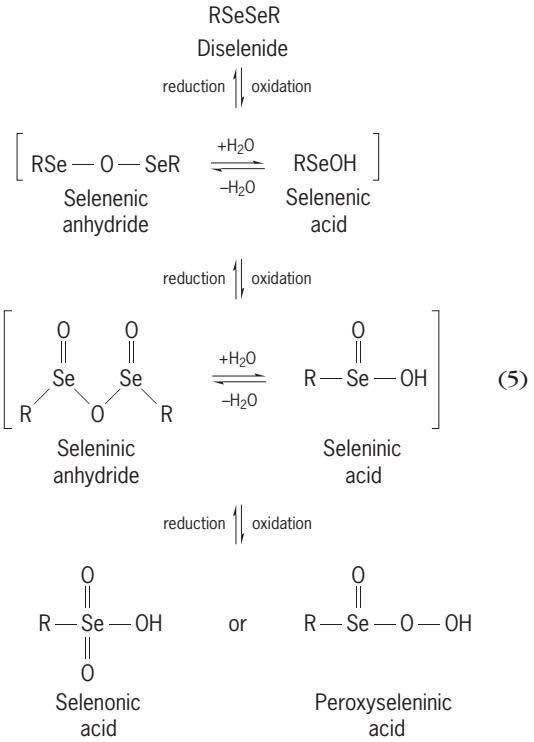

easily oxidized to seleninic acids or anhydrides by reagents such as hydrogen peroxide or nitric acid; further oxidation to selenonic acids is difficult, and the products are relatively unstable. This contrasts with the analogous sulfur compounds where sulfonic acids are stable and easily prepared. The in-place formation of peroxyseleninic acids has also been reported from seleninic acids and hydrogen peroxide. The preparation of selenenic acids (or their anhydrides) from the partial oxidation of diselenides is generally not possible as they disproportionate to diselenides and seleninic acids, as shown in reaction (6). This process is reversible, but the equilibrium favors the products on the right. When generated in the presence of olefins, selenenic acids undergo addition reactions similar to selenenyl chlo-

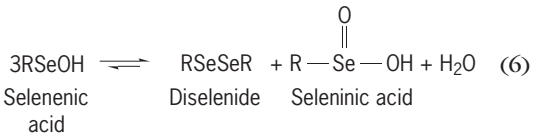

rides [such as reaction (4), except that X = OH].

Selenoxides and selenones. Selenoxides are readily obtained from the oxidation of selenides with hydrogen peroxide (H_2O_2) or similar oxidants [reactions (7)]. Further oxidation to selenones is far more

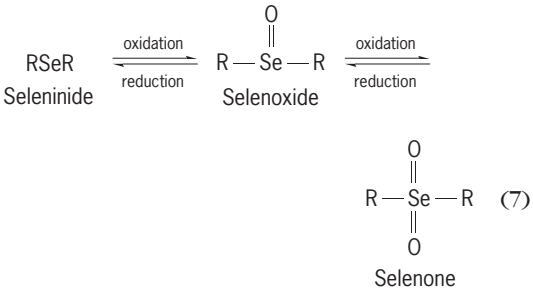

difficult, again in contrast to the corresponding sulfur analogs (sulfones), which are easily prepared from sulfides or sulfoxides. Selenoxides undergo facile elimination to produce olefins and selenenic acids [reaction (8)]. The process is known as syn-

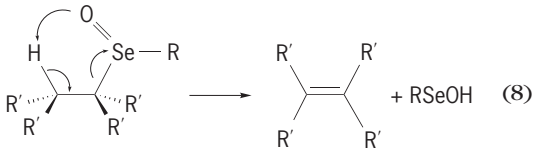

elimination; it requires the selenoxide oxygen atom and the abstracted hydrogen to lie on the same side of the carbon–carbon bond, with all five participating atoms (O, Se, C, C, H) coplanar. The selenium atom of an unsymmetrical selenoxide is a chiral center, and such compounds exist as pairs of enantiomers. They are configurationally stable under anhydrous conditions but racemize rapidly in the presence of water because of the reversible formation of achiral hydrates [reactions (9)].

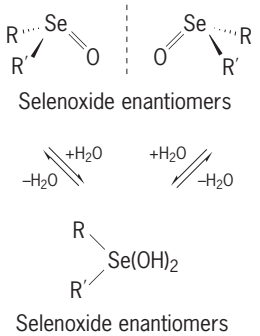

See OXIDATION-REDUCTION; RACEMIZATION; STEREOCHEMISTRY.

Selenocarbonyl compounds. Selenocarbonyl compounds tend to be considerably less stable than their thiocarbonyl or carbonyl counterparts because of the weaker double bond between the carbon and selenium atoms. Selenoamides, selenoesters and related compounds can be isolated, but selenoketones

(selones) and selenoaldehydes are highly unstable unless the selenocarbonyl group is sterically protected by two bulky substituents (for example *t*-butyl). The compounds carbon diselenide (CSe_2) and carbon oxyselenide (COSe) are also known. The general structures of typical selenocarbonyl compounds are shown below.

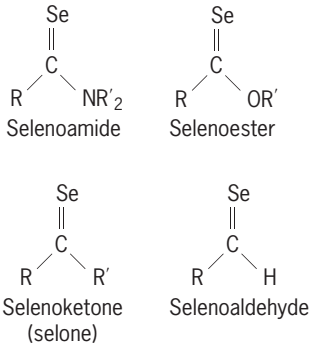

See ALDEHYDE; AMIDE; CHEMICAL BONDING; ESTER; KETONE.

Applications in synthesis. The most widely used procedure employing organoselenium compounds is the selenoxide syn-elimination shown in reaction (8), which is an exceptionally mild olefin-forming reaction. Since it often proceeds readily at room temperature, it obviates the need for the forcing pyrolytic condition associated with other types of syn-eliminations. Scavengers for the by-product selenenic acid are sometimes necessary to prevent its readdition to the product olefin. Seleninic acids and anhydrides, as well as the inorganic reagent selenium dioxide (SeO_2), are useful oxidizing agents. For example, they can be employed in the dehydrogenation of carbonyl compounds, including steroid derivatives [reaction (10); Ph=C_6H_5; phenyl]. Perox-

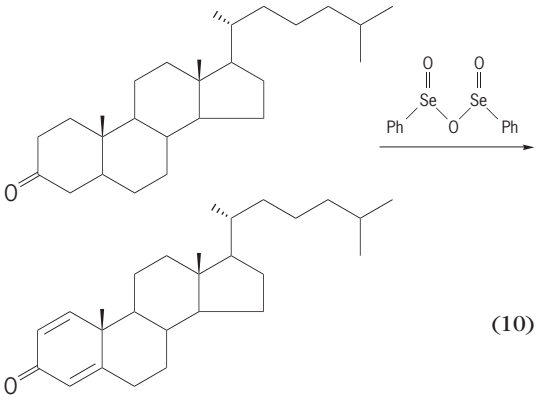

(10)

yseleninic acids are efficient reagents for the epoxidation of olefins or the conversion of ketones to esters or lactones. The additions of selenenyl halides, selenenic acids, and related derivatives to olefins [as reaction (4)] provides a useful means for introducing halide, hydroxyl (OH), and other functional groups into the products. (A functional group is an atom or group of atoms considered to have replaced a hydrogen in a hydrocarbon compound.) Selenium-stabilized carbonions derived from selenides react with alkyl halides and other electrophiles, resulting in the formation of new carbon—carbon bonds [reaction (11)], a vital operation in the synthesis of organic

$$RCH_2SeR' \xrightarrow{\text{base}} \underset{\text{Se-stabilized carbanion}}{R\bar{C}HSeR'} \xrightarrow{R''X} RCHSeR' + X \qquad (11)$$

molecules. *See* ORGANIC REACTION MECHANISM; ORGANIC SYNTHESIS.

Organic conductors. The charge-transfer complexes formed between certain organoselenium donor molecules and appropriate acceptors such as tetracyanoquinodimethane are capable of conducting electric current. Tetraselenafulvalene, its sulfur

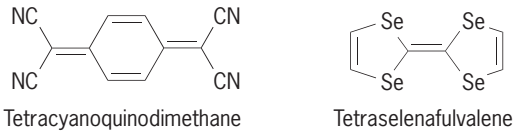

analog, and their various derivatives are particularly effective donors. Selenium-containing polymers are also of interest as organic conductors. *See* COORDINATION COMPLEXES; ORGANIC CONDUCTOR.

Biological importance. Organoselenium compounds are generally quite toxic and must be handled with appropriate care. However, selenium is also an essential trace element, and its complete absence in the diet is severely detrimental to human and animal health. The element is incorporated into selenoproteins such as glutathione peroxidase, which acts as a natural antioxidant. Harmful peroxides formed as by-products during normal oxidative metabolism are rapidly reduced by glutathione in a process that is catalyzed by the selenoprotein. This protects cells from oxidative damage caused by the peroxides. *See* BIOINORGANIC CHEMISTRY; COORDINATION CHEMISTRY; PROTEIN. Thomas G. Back

Bibliography. R. F. Burk, *Selenium in Biology and Human Health*, 1993; D. Liotta (ed.), *Organoselenium Chemistry*, 1987; K. C. Nicolaou and N. A. Petasis, *Selenium in Natural Products Synthesis*, 1984; S. Patai and Z. Rappoport (eds.), *The Chemistry of Organic Selenium and Tellurium Compounds*, vols. 1–2, 1986–1987; C. Paulmier, *Selenium Reagents and Intermediates in Organic Synthesis*, 1986; R. J. Shamberger, *Biochemistry of Selenium*, 1983.

Organosilicon compound

One of a group of compounds in which silicon (Si) is bonded to an organic functional group (R) either directly or indirectly via another atom. Formally, all organosilanes can be viewed as derivatives of silane (SiH_4) by the substitution of hydrogen (H) atoms. The most common substituents are methyl (CH_3; Me) and phenyl (C_6H_5; Ph) groups. However, tremendous diversity results with cyclic structures and the introduction of heteroatoms. Some representative examples and their nomenclature are given in the **table**. *See* SILICON.

Organosilicon compounds are not found in nature and must be prepared in the laboratory. The

Representative organosilicon compounds

Structure		Nomenclature
Single bonds	X = H	Silane
—Si—X	X = Me	Methylsilane
	X = Ph	Phenylsilane
	X = Cl	Chlorosilane
	X = O—CR_3	Alkoxysilane
	X = O—SiR_3	Siloxane
	X = O—H	Silanol
	X = NH_2	Silylamine
Multiple bonds	X = CR_2	Silene (or silaethylene)
Si = X	X = SiR_2	Disilene
	X = O	Silanone
—Si ≡ X	X = CR	Silyne (or silaacetylene)
	X = SiR	Disilyne
Polymers		
$(—SiR_2—)_n$		Polysilane
$(—SiR_2—O—)_n$		Polysiloxane (or silicone)
Cyclic structures		
$[(SiR_2—)_n]$		Cyclosilane
$[(SiR_2—O—)_n]$		Cyclosiloxane

ultimate starting material is sand (silicon dioxide, SiO_2) or other inorganic silicates, which make up over 75% of the Earth's crust. This transformation was first accomplished in 1863 by C. Friedel and J. Crafts, who prepared the first organosilicon compound, tetraethylsilane, by the reaction of silicon tetrachloride ($SiCl_4$) with diethylzinc [$Zn(C_2H_5)_2$]. In the following years, many other derivatives were synthesized, but it was not until the useful properties of silicone polymers were identified in the 1940s that widespread interest in organosilicon chemistry appeared.

Bonding and reactivity. The chemistry of organosilanes can be explained in terms of the fundamental electronic structure of silicon and the polar nature of its bonds to other elements. Silicon appears in the third row of the periodic table, immediately below carbon (C) in group 14 and has many similarities with carbon. However, it is the fundamental differences between carbon and silicon that make silicon so useful in organic synthesis and of such great theoretical interest. *See* CARBON.

In the majority of organosilicon compounds, silicon follows the octet rule and is 4-coordinate. This trend can be explained by the electronic configuration of atomic silicon ($3s^2 3p^2 3d^0$) and the formation of sp^3-hybrid orbitals for bonding. Unlike carbon ($2s^2 2p^2$), however, silicon is capable of expanding its octet and can form 5- and 6-coordinate species with electronegative substituents, such as the 6-coordinate octahedral dianion, $[Me_2SiF_4]^{2-}$. Also, 5-coordinate species with trigonal bipyramidal geometries are proposed intermediates in nucleophilic displacement reactions at silicon. *See* COORDINATION CHEMISTRY; ELECTRON CONFIGURATION; VALENCE.

Silicon is a relatively electropositive element that forms polar covalent bonds ($Si^{\delta+}$—$X^{\delta-}$) with carbon and other elements, including the halogens, nitrogen, and oxygen. The strength and reactivity of silicon bonds depend on the relative electronegativities of the two elements. For example, the strongly electronegative elements fluorine (F) and oxygen (O) form bonds that have tremendous thermodynamic stabilities, 807 and 531 kJ/mol, respectively. For the silicon-carbon bond, where the electronegativity difference is smaller, the bond dissociation energy is 318 kJ/mol and has stability comparable to the carbon-carbon bond, 334 kJ/mol. *See* CHEMICAL THERMODYNAMICS; ELECTRONEGATIVITY.

Before 1981, multiply bonded silicon compounds were identified only as transient species both in solution and in the gas phase. However, when tetramesityldisilene ($Ar_2Si{=}SiAr_2$, where Ar = 2,4,6-trimethylphenyl) was isolated, it was found to be a crystalline, high-melting-point solid having excellent stability in the absence of oxygen and moisture. The key factor for success was the presence of bulky substituents on silicon that impart kinetic stabilization to the double bond by preventing the approach of potential reactants. The bonding is explained in classical terms as a (3*p*-3*p*) pi double bond. Investigations into the reactivity of the silicon-silicon (Si=Si) double bond have revealed a rich chemistry, and many of the reactions of disilenes (compounds with the Si=Si group) are without precedent in ethylene ($H_2C{=}CH_2$) chemistry. For example, disilenes react with molecular oxygen (O_2) to give unusual cyclic siloxanes [reaction (1), where R = organic group].

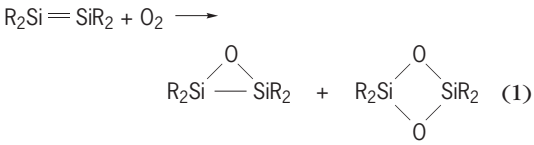

Other multiply bonded silicon species have also been isolated; examples are Si═C and Si═N.

Numerous reactive species have been generated and characterized in organosilicon chemistry. Silylenes are divalent silicon species (SiR_2, where R = alkyl, aryl, or hydrogen), analogous to carbenes in carbon chemistry. The dimerization of two silylene units gives a disilene. Silicon-based anions, cations, and radicals are important intermediates in the reactions of silicon. *See* CHEMICAL BONDING; FREE RADICAL; REACTIVE INTERMEDIATES.

Preparation. The direct process for the large-scale preparation of organosilanes has provided a convenient source of raw materials for the development of the silicone industry. The process produces a mixture of chloromethylsilanes from elemental silicon and methyl chloride in the presence of a copper (Cu) catalyst [reaction (2)]. In a variation on the basic re-

$$\text{Me—Cl} + \text{Si} \xrightarrow{\text{Cu catalyst}} \text{Cl}_2\text{SiMe}_2 + \text{other Cl}_n\text{SiMe}_{4-n} \quad (2)$$

action, methyl chloride can be replaced by other organic halides, including chlorobenzene which yields a mixture of chlorophenylsilanes.

In spite of concentrated research efforts in this area, the synthetic methods available for the controlled formation of silicon-carbon bonds remain limited to only a few general reaction types. These include the reaction of organometallic reagents with silanes catalytic hydrosilylation of multiple bonds, and reductive silylation.

Organometallic reagents with silanes. A versatile approach to the synthesis of organosilanes is the reaction of compounds containing Si-X bonds (where X = halogen, OR, or H), with organometallic reagents (RM; M = metal), such as organolithium and organomagnesium compounds. These reactions can be viewed as nucleophilic displacements at the silicon atom [reaction (3)]. They are facilitated by the presence

$$\text{R—M} + \text{—Si—X} \longrightarrow \text{R—Si—} + \text{MX} \quad (3)$$

of electron-withdrawing substituents on silicon. *See* ORGANOMETALLIC COMPOUND.

Catalytic hydrosilylation. This is used for the preparation of organosilanes from multiply bonded carbon compounds. The reaction effects the addition of a silicon-hydrogen bond across an alkene or an alkyne. It is often catalyzed by peroxides or transition metals, as well as by light. The mechanism is thought to involve silyl radicals (R3Si·) as the reactive species [reaction (4)].

$$\text{R}_3\text{Si—H} + \text{R}_2\text{C}{=}\text{CH}_2 \longrightarrow \text{H—CR}_2\text{—CH}_2\text{—SiR}_3 \quad (4)$$

See ALKENE; ALKYNE; ORGANIC REACTION MECHANISM.

Reductive silylation. A less common reaction for the formation of silicon-carbon bonds is reductive silylation. An example is the silylation of a carboxylic ester with lithium (Li) metal in the presence of chlorotrimethylsilane [reaction (5)].

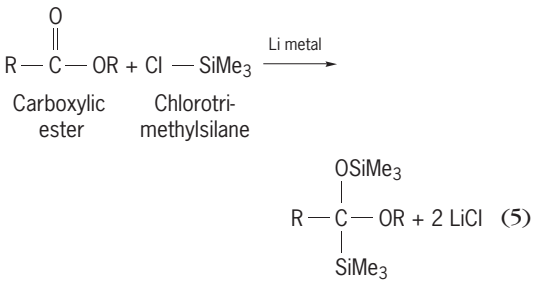

(5)

Reagents in organic synthesis. The role of silicon in organic synthesis is quite extensive, and chemists exploit the unique reactivity of organosilanes to accomplish a wide variety of transformations. Silicon is usually introduced into a molecule to perform a specific function and is then removed under controlled conditions.

A common use of silanes is as a protecting group to mask an active hydrogen functionality (for example, an alcohol, amine, or thiol) so that it is neither altered by, nor does it interfere with, reagents targeted at other functional groups [reaction (6)]. After the

$$\text{R—OH} \underset{\text{hydrolysis or F}^-\text{ ion}}{\overset{\text{silylation}}{\rightleftharpoons}} \text{R—O—SiMe}_3 \quad (6)$$

transformation is completed, the silane is removed by hydrolysis or by treatment with fluoride ion (F^-).

In organic synthesis, the conditions chosen to effect a given chemical transformation at a targeted site often are not compatible with another functionality in molecule. As such, silyl enol ethers are powerful synthetic tools, providing an alternative to the nucleophilic alkylation of carbonyl compounds [reaction (7)]. The two approaches are complementary.

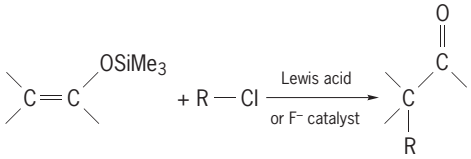

(7)

In certain synthetic reactions, the formation of energetically favorable silicon-fluorine and silicon-oxygen bonds provides a thermodynamic driving force. Two representative examples of this type are shown in reactions (8) and (9). Reaction (8) is a

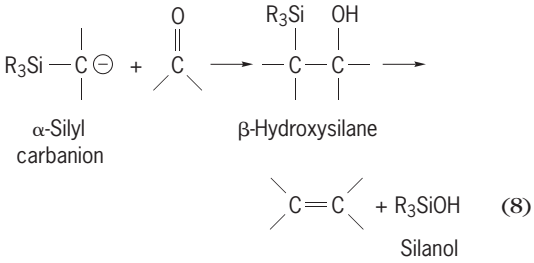

(8)

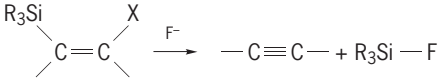

(9)

valuable synthesis for the preparation of carbon-carbon double bonds from carbonyl compounds. An α-silylcarbanion is reacted withan aldehyde or

ketone to form a β-hydroxysilane, and subsequent elimination of a silanol gives the olefin. *See* ALDEHYDE; KETONE.

Reaction (9) is a fluoride ion-assisted elimination of halosilanes to form multiple bonds. The reaction, because it occurs at low temperature under mild conditions, has been used to great advantage in the synthesis of highly strained molecules, such as benzyne.

Polysilanes. These are organosilicon compounds that contain either cyclic arrays or linear chains of silicon atoms. The isolation and characterization of both cyclosilanes (with up to 35 silane units) and high-molecular-weight linear polysilanes (with up to 3000 silane units) have demonstrated that silicon is capable of extended chain formation (catenation). Polysilanes contain only silicon-silicon bonds in their backbone, which differentiates them from polysiloxanes (silicones), which contain alternating silicon and oxygen repeat units. *See* SILICONE RESINS.

The synthetic methods available for the formation of silicon-silicon bonds are extremely limited. The most general procedure for the preparation of polysilanes is via the coupling of halosilanes in the presence of an alkali metal.

The degree of functionalization of the halosilane and the reaction conditions determine the product type. For example, the coupling of monohalosilanes gives disilane products [reaction (10); M = alkali

$$R_3Si{-}Cl \xrightarrow{M} R_3Si{-}SiR_3 \qquad (10)$$

metal]. If dichlorosilanes are used, the result is the formation of cyclosilanes, under mild conditions, or linear polymers, under more extreme conditions [reaction (11); M = alkali metal]. The introduction

$$R_2SiCl_2 \xrightarrow{M} ({-}SiR_2{-})_n \qquad (11)$$

of small amounts of trifunctional silanes leads to branching points, and in the extreme case where only trihalosilane is used, a highly cross-linked polymer network is formed.

The polysilanes have properties that make them attractive materials for many applications. Perhaps the most interesting is that these polymers exhibit extended electron delocalization in the sigma-bonded framework. They are also photosensitive, and the absorption of ultraviolet light results in degradation of the polymers. *See* DELOCALIZATION; PHOTODEGRADATION; POLYMER.

Polysilanes have found applications as photoresists in microlithography, charge carriers in electrophotography, and photoinitiators for vinyl polymerization. Polysilanes also function as preceramic polymers for the manufacture of silicon carbide fibers. *See* ORGANIC SYNTHESIS. Howard Yokelson

Bibliography. E. Colvin, *Silicon in Organic Synthesis*, 1981; S. Patai and Z. Rappoport (eds.), *The Chemistry of Organic Silicon Compounds*, 1989; R. West and T. J. Barton, Organosilicon chemistry, *J. Chem. Educ.*, 57:165–169, 334–340, 1980; G. Wilkinson, F. G. A. Stone, and E. W. Abel (eds.), *Comprehensive Organometallic Chemistry*, vol. 2, 1982.

Organosulfur compound

A member of a class of organic compounds with any of several dozen functional groups containing sulfur (S) and often also oxygen (O), nitrogen (N), hydrogen (H), as well as other elements.

Sulfur is an element of the third row of the periodic table; it is larger and less electronegative than oxygen, which lies above it in the second row. Compounds with an expanded valence shell, that is, compounds bonding to as many as six ligands around sulfur, are therefore possible, and a broad range of compounds can be formed. Moreover, sulfur has a much greater tendency than oxygen to undergo catenation to give chains with several atoms linked together through S—S bonds. *See* CHEMICAL BONDING; PERIODIC TABLE; STRUCTURAL CHEMISTRY; VALENCE.

The structures and names of representative types of organosulfur compounds are shown in the **table**. Some compounds and groups are named by using the prefix thio to denote replacement of oxygen by sulfur. The prefix thia can be used to indicate that one or more —CH_2— groups have been replaced by sulfur, as in 2,7-dithianonane [$CH_3S(CH_2)_4SCH_2CH_3$].

Thiols and sulfides. Thiols and sulfides are sulfur counterparts of alcohols and ethers, respectively, and can be prepared by substitution reactions analogous to those used for the oxygen compounds actions [reactions (1) and (2)]. Sulfonium salts are obtained by further alkylation of sulfides. A better

Some types of organosulfur compounds and groups

Structure	Name
RSH	Thiol (mercaptan)
RSR	Sulfide (thioether)
RSSR	Disulfide
RSSSR	Trisulfide (trisulfane)
$R\overset{+}{S}(R)R\ X^-$	Sulfonium salt
$R_2C{=}S$	Thioketone
$R\bar{N}{=}C{=}S$	Isothiocyanate
$RC(=O)SR$	Thiolate ester (thoic acid S-ester)
$RC(=S)OR$	Thionoate ester
RCS_2R	Dithioate ester
RSOH	Sulfenic acid
RSCl	Sulfenyl chloride
RSOR	Sulfoxide
$RS(=NR)R$	Sulfimide
RSO_2H	Sulfinic acid
RSO_2R	Sulfinate ester
$R_2S{=}R_2$	Sulfonium ylide (sulfurane)
RSO_2R	Sulfone
RSO_3H	Sulfonic acid
RSO_2NH_2	Sulfonamide
RSO_2Cl	Sulfonyl chloride
$ROSO_3R$	Sulfate ester

method for preparing thiols is formation of the isoth-

$$\underset{\text{Bromide}}{RBr} + NaSH \longrightarrow \underset{\text{Thiol}}{RSH} \rightleftharpoons RSNa + H_2S \qquad (1)$$

$$RBr + R'SNa \longrightarrow \underset{\text{Sulfide}}{RSR'} \xrightarrow{R''Br} \underset{\text{Sulfonium salt}}{R\overset{+}{S}(R'')R'\ Br^-} \qquad (2)$$

iouronium salt by reaction with thioureas, followed by hydrolysis [reaction (3)]. In these substitutions,

$$RBr + \underset{\text{Thiourea}}{NH_2C(=S)NH_2} \longrightarrow RSC(=\overset{+}{N}H_2)NH_2 \xrightarrow{H_2O} RSH + \underset{\text{Urea}}{NH_2C(=O)NH_2} \qquad (3)$$

sulfur is a better nucleophile than oxygen. The rate of SN2 reactions with RSNa (Na = sodium) is higher than that with the corresponding oxygen nucleophile (RONa), and the reaction of an ambident nucleophile such as thiosulfate ion gives the *S*-alkyl product exclusively [reaction (4)].

$$RBr + S_2O_3^{2-} \longrightarrow \underset{\text{Alkyl thiosulfate}}{RSSO_3^-} + Br^- \qquad (4)$$

See ALCOHOL; ETHER; ORGANIC REACTION MECHANISM; SUBSTITUTION REACTION.

Although thiols and alcohols are structurally analogous, there are significant differences in the properties of these two groups. Hydrogen bonding (denoted by the broken line) of the type —S–H—S— is very weak compared to —O—H–O—, and thiols are thus more volatile and have lower boiling points than the corresponding alcohols; for example, methanethiol (CH_3SH) has a boiling point of 5.8°C (42.4°F) compared to 65.7°C (150.3°F) for methanol (CH_3OH).

Thiols form insoluble precipitates with heavy-metal ions such as lead or mercury; the older name mercaptan, a synonym for thiol, reflects this property. Both thiols and sulfides are extremely malodorous compounds, recalling the stench of rotten eggs (hydrogen sulfide). However, traces of these sulfur compounds are an essential component of the distinctive flavors and aromas of many vegetables, coffee, and roast meat. *See* MAILLARD REACTION; MERCAPTAN; SPICE AND FLAVORING.

Because of the larger atomic radius and greater polarizability of sulfur, the —SH group is much more acidic than —OH (for example, ethanethiol, pK_a 10.6; ethanol, pK_a 15.9; the lower the value of pK_a, the stronger the acid). For the same reasons, an acyl thiol (thiolester; RS—COR′) is more reactive than an ester (RO—COR′) toward attack of a nucleophile such as hydroxide ion or a carbanion, and an α-hydrogen in CH_3CO—SR is more acidic than that in CH_3CO—OR. *See* PK; REACTIVE INTERMEDIATES.

Thiols readily undergo one-electron oxidation to thiyl radicals, which dimerize to disulfides [reaction (5)]. This reaction and the reverse reduction

$$RSH \underset{[H]}{\overset{[O]}{\rightleftharpoons}} \underset{\text{Thiyl}}{RS'} \rightleftharpoons \underset{\text{Disulfide}}{RS—SR} \qquad (5)$$

can occur under physiological conditions; this redox system has an important role in biochemistry. *See* OXIDATION-REDUCTION.

Thiyl radicals are intermediates in the free-radical addition of thiols to alkenes; this reaction is a useful preparative method for sulfides [reaction (6)].

$$RSH \xrightarrow{[O]} [RS^{\cdot}] \xrightarrow{R'CH=CH_2} R'CHCH_2SR \xrightarrow{RSH} R'CH_2CH_2SR + RS \qquad (6)$$

Thiyl radicals are also involved in the vulcanization of rubber, in which natural or synthetic polymers are converted to an elastomeric substance by cross-linking the chains with sulfur. This process is controlled by use of free-radical accelerators such as mercaptobenzimidazole or tetraethylthiuram disulfide. The latter compound has also been used in the treatment of alcoholism. *See* FREE RADICAL; RUBBER.

Sulfur-stabilized carbanions and ylides. Another important consequence of the high polarizability of sulfur is stabilization of negative charge on carbon by an adjacent sulfur atom. Aryl alkyl sulfides are converted by strong bases to the anion, which can react as a nucleophile with alkylating agents [reaction (7)]. This

$$\underset{\text{Arylalkyl sulfide}}{ArSCH_3} \xrightarrow[\Delta]{\text{butyl lithium}} ArSCH_2^- \xrightarrow{RX} ArSCH_2 \qquad (7)$$

effect is particularly useful with dithioacetals, which are readily deprotonated to the anion; the latter can serve as an acyl anion equivalent, in which the polarity of the original C=O group is reversed, a process known as umpolung. For example, reaction of the dithioanion with a carbonyl compound followed by hydrolysis with a thiophilic metal ion leads to an α-hydroxy carbonyl [reaction (8)].

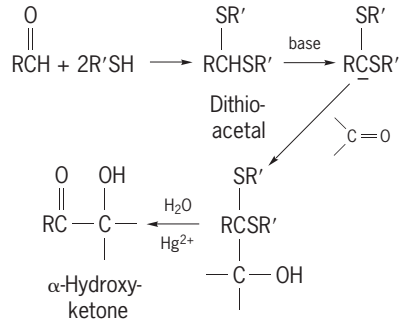

(8)

Removal of a proton (H^+) from a sulfonium salt leads to a sulfur-ylide, analogous to the phosphorus-ylides obtained from a phosphonium salt. The *S*-ylide reacts with a carbonyl compound to give an epoxide [reaction (9)].

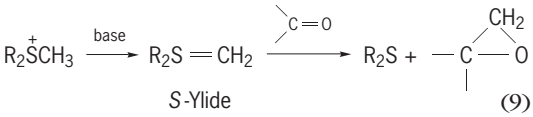

(9)

Thiocarbonyl compounds. These contain a carbon-sulfur double bond (C=S). Thiocarbonyl

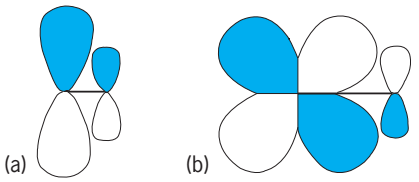

Carbon-sulfur bonding in π-bonds described as orbitals. (*a*) 3*p*-2*p* orbitals. (*b*) 3*d*-2*p* orbitals.

compounds (thiones) are much less common than carbonyl compounds (C═O bond). Simple thioaldehydes or thioketones have a strong tendency to form cyclic trimers, polymers, or other products. Thiocarboxylic acids have the thiol structure rather than the thione, although the latter is the stable form in thioamides, as in the structures below.

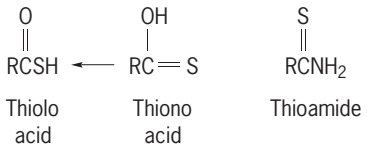

Carbon-sulfur double bonds are weak because the overlap is poor between the small carbon 2*p* orbital and the relatively large and diffuse sulfur 3*p* orbital (**illus.** *a*). Another type of multiple bonding in organosulfur compounds utilizes 3*d* sulfur orbitals or hybrids thereof; it is found in compounds in which the sulfur is in a higher oxidation state than dicovalent sulfur, as in sulfoxides, sulfones, and sulfuranes. The sulfur-oxygen or sulfur-carbon double bonds in these compounds can be described as 3*d*-2*p* π bonds (illus. *b*). *See* MOLECULAR ORBITAL THEORY.

Compounds with higher oxidation states of sulfur. Sulfides can be oxidized sequentially to sulfoxides and sulfones, containing the sulfinyl (—SO—) and sulfonyl (—SO_2—) groups, respectively [reaction (10)].

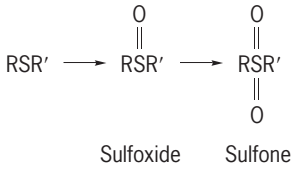

(10)

Dimethyl sulfoxide (DMSO) is available in large quantities as a by-product of the Kraft sulfite paper process. It is useful as a polar solvent with a high boiling point and as a selective oxidant and reagent in organic synthesis. Sulfoxides with unlike groups attached to the sulfur atoms are chiral compounds, and thus the enantiomers can be resolved. *See* DIMETHYL SULFOXIDE; STEREOCHEMISTRY.

Compounds containing the sulfonyl group include sulfones, sulfonyl chlorides, sulfonic acids, and sulfonamides. Sulfones can be obtained by oxidation of sulfoxides or by sulfonation of an aromatic ring with a sulfonyl chloride, as shown in reaction scheme (11). The sulfonyl group resembles a carbonyl in the acidifying effect on an α-hydrogen. The diaryl sulfone unit is the central feature of polysulfone resins, used in some high-performance plastics. Sulfonic acids are obtained by oxidation of thiols or by sulfonation. Sulfonamides, prepared from the chlorides, were the mainstay therapeutic agents in infections until the

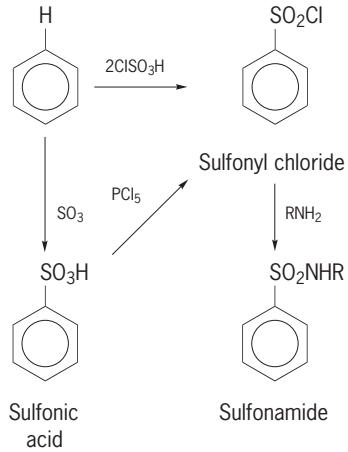

(11)

advent of antibiotics; they are still used for some conditions. *See* POLYSULFONE RESINS; SULFONAMIDE; SULFONIC ACID.

Biochemical compounds. A number of proteins and metabolic pathways in systems of living organisms depend on the amino acid cysteine and other sulfur compounds. In many proteins, for example, in the enzyme insulin, disulfide bonds formed from the —SH groups of cysteine units are an essential part of the structure. Other examples are the keratins (the tough fibrous proteins of hair, wool, and animal horns), which contain large amounts of cysteine. The —SH groups of cysteine, whose structure is shown below, also play a role in the metal-sulfur proteins

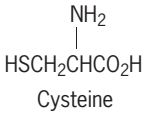

that mediate electron-transport reactions in respiration and photosynthesis. *See* INSULIN; PHOTOSYNTHESIS.

The coenzyme lipoic acid is a cyclic disulfide that functions together with the coenzyme thiamine diphosphate to accept electrons and undergo reduction of the —S—S— bond in the oxidative decarboxylation of pyruvic acid [reaction scheme (12)].

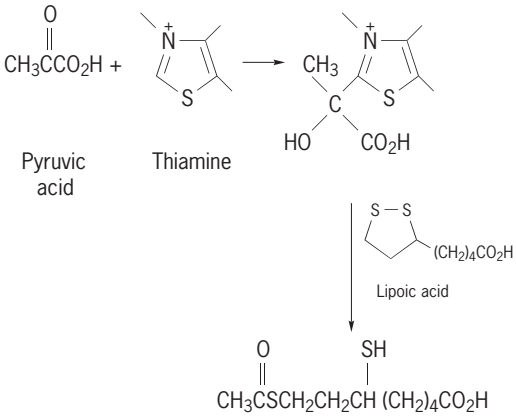

(12)

See BACTERIAL PHYSIOLOGY AND METABOLISM; COENZYME; THIAMINE.

Two other major pathways in metabolism, the transfer of acetyl groups (CH_3CO—) and of methyl (CH_3—) groups, are mediated by organosulfur

compounds. Acetyl transfer, a key step in lipid and carbohydrate metabolism, occurs by way of thioesters. An acetyl group is activated by formation of the thioester with the —SH of coenzyme A. Biological methylations take place by the formation of *S*-adenosylmethionine and attack of a nucleophile at the methyl group of this sulfonium ion, as in reaction (13), where Nuc represents the nucleophile.

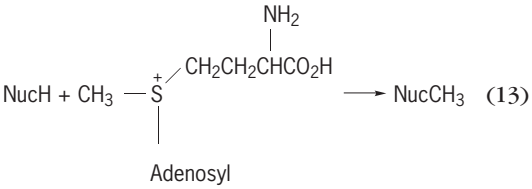

(13)

See CARBOHYDRATE METABOLISM; LIPID METABOLISM.

Naturally occurring compounds. In addition to cysteine, methionine, and several coenzymes, sulfur is present in numerous other compounds found in natural sources. Petroleum contains variable amounts of sulfur, both as simple thiols and sulfides, and also heterocyclic compounds such as benzothiophene. Removal of these is an important step in petroleum refining. *See* PETROLEUM; PETROLEUM PROCESSING AND REFINING.

The presence of sulfur is often signaled by characteristic odors. In onions and garlic, the odors released when the bulb is cut are due to several unsaturated organosulfur compounds. The unique flavor of garlic arises from the disulfide *S*-oxide allicin. The pungent, lacrimatory properties of onions are due to propanethiol *S*-oxide; in horseradish these properties are due mainly to allyl isothiocyanate (see below). Shiitake mushrooms owe their distinctive

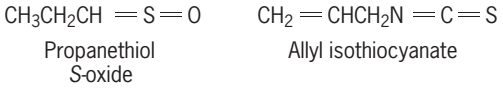

Propanethiol *S*-oxide — Allyl isothiocyanate

aroma to the cyclic polysulfide lenthionine. A mixture of four carbon thiols is responsible for the odor of skunk secretion.

Several sulfur-containing compounds from natural sources have important pharmacological properties. Examples are the β-lactam antibiotics penicillin, cephalosporin, and thienamycin, and the platelet anticoagulating factor ajoene from garlic, produced by a series of complex enzymatic reactions from al-

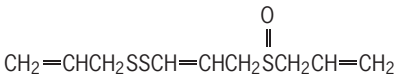

Ajoene

licin. *See* ANTIBIOTIC; HETEROCYCLIC COMPOUNDS; ORGANIC CHEMISTRY; SULFUR. James A. Moore

Bibliography. E. Block (ed.), *Advances in Sulfur Chemistry*, vol. 1, 1988; E. Block, *Reactions of Organosulfur Compounds*, 1978; D. N. Jones (ed.), *Organosulfur Compounds*, vol. 3 of *Barton-Ollis Comprehensive Organic Chemistry*, 1979; A. Senning (ed.), *Sulfur in Organic and Inorganic Chemistry*, vols. 1–4, 1971–1982; T. W. G. Solomons, *Organic Chemistry*, 7th ed., 1999; B. Zwanenburg and A. J. H. Klunder (eds.), *Perspectives in the Organic Chemistry of Sulfur*, 1987.

Oriental vegetables

Oriental vegetables are very important in Asian countries, but are considered as minor crops in the United States and Europe. However, there has been an increased interest in these crops because of their unusual flavors and textures and in some cases their high nutritional values (see **table**). Some of the more common ones are described below.

Chinese cabbage. Chinese cabbage, celery cabbage, napa, or pe-tsai (*Brassica campestris*, pekinensis group; *B. rapa*, pekinensis group; *B. pekinensis*) belongs to the mustard (Cruciferae) family, and is a biennial leafy plant but is grown as an annual. The origin of this crop is obscure; it is thought to have first developed in China, and spread to southeastern Asia and then to Korea and Japan. The harvested part is a head which is composed of broad crinkled leaves with a very wide, indistinct, white midrib. The outer leaves are pale green, and the inner leaves of the head are blanched. Chinese cabbage is usually grown in the cool season.

In the north temperate region, if the summers are relatively cool, seeds are planted in July and August; but if summers are hot, seeding is usually in late August or during September, when the maximum day temperatures have started to decrease. If planted too early during hot weather, plants may bolt (produce seed stalks) before heads can form. The production of Chinese cabbage occurs in the fall, winter, and into early spring. There are many varieties of Chinese cabbage adapted for various climates; the hybrids are very popular because they are more uniform and of better quality than the open-pollinated ones.

The seeds are planted 0.25–0.50 in. (0.6–1.2 mm) in depth on well-prepared beds about 3 ft (90 cm) apart. The seeds may be planted first in a seedling bed for transplanting to the field when plants are at the four- to six-true-leaf stage. Before transplanting to the field, the plants should be hardened by withholding water until they show signs of wilting in the afternoons. Spacing of Chinese cabbage should be 10–12 in. (25–30 cm) between plants. In Asia the thinned plants in direct-seeded fields are often bundled and sold in markets as greens. In 45–75 days from seeding, Chinese cabbage forms a cylindrical compact head. Harvested heads may be stored for several weeks in the refrigerator under high humidity without much loss of quality. It is a good salad vegetable because of the mild flavor and crisp texture.

Pak choy. Most varieties of pak choy, bok choi, Chinese mustard, or celery mustard (*Brassica campestris*, chinensis group; *B. rapa*, chinensis group; *B. chinensis*) are biennials, but some are annuals which flower during the first year without vernalization. The crop is grown as an annual. The origin of pak choy was probably in northern China; the plant subsequently spread to southeastern Asia and to other parts of the world in post-Columbian times. Pak choy does not form a head like most varieties of Chinese cabbage, but forms a celerylike stalk of tall, dark green leaves, with prominent white veins and long white petioles (**Fig. 1**). The base of the

Nutrient composition of some oriental vegetables in amounts per 100-g fresh edible portion[a]

Vegetable	Refuse as purchased, %	Approximate household equivalent[b]	Average food energy, cal[c]	Water	Protein	Fat	Total sugar	Other carbohydrates
Chinese cabbage	10	1½ cups	11	91	1.2	0.2	1.3	0.1
Pak choy	5	1½ cups	13	95	1.5	0.2	1.0	0.2
Chinese winter radish	50[d]	1 cup	13	94	0.6	0.1	2.5	0.2
Edible podded peas	5	¾ cup	35	88	2.8	0.2	4.0	1.8
Yard-long bean	3	¾ cup	30	89	2.8	0.4	3.1	0.7
Bean sprouts								
Mung bean	0	1⅛ cups	25	92	2.7	0.1	2.1	1.4
Jicama[e]	16	1 cup	55	85	1.4	0.2		12.8[f]
Chinese winter melon[e]	25	Slice, 1.5 in.	9	96	0.2	0.1	1.9	0.2
Balsam pear	25	¾ cup	10	94	1.1	0.2	0.8	0.4
Chinese okra	10	2 med. fruits	20	93	1.2	0.2	3.2	0.8
Water spinach	0	1½ cups	25	92	2.6	0.2	0.3	3.1

SOURCE: F. D. Howard, J. H. MacGillivray, and M. Yamaguchi, *Nutrient Composition of Fresh California-Grown Vegetables*. Calif. Agr. Exp. Sta. Bull. 788, 1962.
[a] For vitamins and minerals, amounts are given in milligrams, except for vitamin A, which is given in International Units; elsewhere amounts are in grams, unless otherwise specified.
[b] 1 cup = 237 cm^3. 1 in. = 25 cm.

petiole may be expanded and spoon-shaped; the blade of the leaf is smooth, and not crinkled like that of Chinese cabbage. When grown in cool regions and in moist soils, some varieties may develop enlarged turniplike roots late in the season. Varieties such as Choy Sum are grown especially for the seed stalks, with flower buds and with a few open yellow flowers. The growing of pak choy is very similar to that for Chinese cabbage; however, pak choy is more tolerant of warmer temperatures.

Oriental winter radish. The large, long, white radish (*Raphanus sativus*, longipinnatus group) is often called oriental winter radish, daikon, Chinese winter radish, lobok, and lob paak. Radish is a dicotyledonous herbaceous plant grown for its long, enlarged roots. The origin of radish is thought to be in the eastern Mediterranean region. Radish was an important crop in Egypt about 4500 years ago; it spread to China via India about 2500 years ago, and then to Japan about A.D. 700, where it ranks first in production among the vegetables produced there. Radish roots are generally long and cylindrical, and measure about 2–4 in. (5–10 cm) in diameter and up to 18 in. (45 cm) in length. The turnip-shaped Sakurajima variety can attain diameters as large as 30 cm (12 in.) and weigh 15–22 lb (7–9 kg) each. Rat-tailed radish (*R. sativus*, caudatium group) is grown for the young slender seed pods 12 in. (30 cm) in length, which are eaten raw, cooked, or pickled in India.

Fig. 1. Pak choy or Chinese mustard (*Brassica campestris*). (*University of California Agricultural Experiment Station*)

Winter radish seeds are planted in rows 12–18 in. (30–45 cm) apart in mid-August to mid-November in deep loose soil. Seedlings are thinned to about 8 in. (20 cm). As the hypocotyl and the tap root enlarge, the hypocotyl region pushes out of the ground 3–6 in. (8–15 cm). The roots are ready for harvest 8–10 weeks after planting, but can remain in the ground during the winter months if the temperature remains above freezing; growth continues to take place. The roots should be harvested before the plant bolts (produces a seed stalk) in the spring. Cultivars such as the Tokinashi (all seasons) can be planted in the spring for harvest in the early summer; if the temperature remains cool, it can be planted later for harvest in the summer and into the fall. Spring crops can be harvested 40–50 days after planting.

Edible podded peas. Edible podded peas, China peas, sugar peas, or snow peas (*Pisum sativum*, macrocarpon group) belong to the Leguminosae family. Evidence of peas used for food dates back to about 7000 B.C. in the Near East. The primitive forms of pea are slightly bitter, which prevents their being eaten by animals, and have a tough seed coat, which allows for long dormant periods. In the garden pea the pod swells first as the seeds enlarge, but in the edible podded pea the immature seeds bulge the pod, which demarks the developing seeds. Unlike the regular garden peas which have tough fibery seed pods,

Vitamins					Minerals					
A	Thiamine	Riboflavin	Niacin	C	Ca	Fe	Mg	P	K	Na
1200	0.04	0.05	0.4	27	92	0.5	14	31	230	70
3000	0.04	0.07	0.5	45	105	0.8	27	37	180	100
0	0.02	0.02	0.2	22	27	0.4	22	24	190	30
580	0.15	0.08	0.6	60	43	0.9	22	53	170	6
1400	0.13	0.11	1.0	32	50	1.0	51	59	210	4
25	0.11	0.03	0.6	12	20	0.6	16	35	130	2
Trace	0.04	0.03	0.3	20	15	0.6	—	18	—	—
Trace	0.02	0.03	0.5	14	14	0.4	16	7	200	2
380	0.04	0.04	0.4	84	19	0.5	—	28	—	—
410	0.05	0.06	0.4	12	20	0.4	—	32	—	—
3500	0.03	0.10	0.9	55	95	2.2	49	40	370	6

[c] 1 cal = 4.18 joules. [d] With top as refuse.
[e] Data from W. W. Leung, R. I. Pecot, and B. K. Watt, *Composition of Foods Used in Far Eastern Countries*, USDA Agr. Handb. 34, 1952.
[f] Total carbohydrates.

the edible podded peas were selected for the tender pods and not for the seeds. A variety called Sugar Snap Peas, which has low-fiber-content pods and enlarged seeds, has become popular again. In Taiwan the tender shoots of garden peas are harvested and cooked as greens; the shoots have a very delicate flavor.

Peas are a cool-season crop which can be grown at mean temperatures of 55–65°F (13–18°C). Prior to flowering, pea vines can stand temperatures down to freezing, but the flowers and immature pods are injured by such temperatures. Varieties of edible podded peas have been developed for the tropics and subtropics. Peas are sensitive to acid soils; a pH range of 5.5–7.0 is considered optimal. For early production a light soil is recommended. A good, airy soil is preferred for optimal nitrogen fixation by the symbiotic *Rhizobium* bacteria found in the nodules of the roots. In cool soils some nitrogen fertilizer, together with phosphorus and potassium, is recommended. Late in the fall in mild climates or early in the spring in cold climates, seeds are planted 1–1.5 in. (2.5–4 cm) in depth, 2–3 in. (5–8 cm) apart, and in rows 3 ft (90 cm) apart. For edible podded pea varieties which are usually climbers, poles or trellises should be put up. From seeding to harvest, 60–110 days are required, depending upon the variety and season. For best quality, the pods are harvested just before the seeds begin to swell. The harvested pods can be put in a refrigerator under high humidity for several days without much loss in quality.

Yard-long bean. Yard-long bean or asparagus bean (*Vigna sinensis*, sesquipedelis group) belongs to the Leguminosae family. It is an annual climbing plant which can reach heights of 6–12 ft (2–4 m) and produce pods 1–3 ft (30–90 cm) long. It is a relative of the cow pea (*V. unguiculata*), its primary center of origin was in Ethiopia, and its secondary center of development was in India about 3500 years ago. During the first millennium B.C. the cultivation of yard-long beans spread into southeastern Asia and China.

Yard-long beans grow best in loam soil, but can be grown in heavy clay soils as well. It is a warm-season crop, and cannot stand any freezing temperature. Seeds are planted 1–1.5 in. (2.5–4 cm) in depth in fertile, well-drained soil after all danger of frost has passed, 6–8 in. (15–20 cm) apart in rows of two or three seeds per hill 12–24 in. (30–60 cm) apart; rows or hills should be 3–3.5 ft (90–105 cm) apart to allow for poling or trellising. In some regions of the tropics, yard-long beans are often planted with corn. The beans are planted after the corn is well established and tasseling. The growing bean plant can use the corn stalk for support after the ears have been harvested. Pods can be harvested about 50 days from seeding. They are picked before the pods reach full size and are not fibery or tough, and while the seeds are immature and not starchy. There are two varieties of yard-long beans; the green-podded type, with pods 1.5–3 ft (45–90 cm) long, and the white-podded (pale-green) type, with pods 1–1.5 ft (30–45 cm) long.

Bean sprouts. Mung bean or green gram (*Phaseolus aureus*) sprouts have been used by the Chinese for their remarkable healing qualities for thousands of years, and were described in their earliest medical treatises. Only recently has the Western world recognized the value of sprouted leguminous seeds (mung, soy, and alfalfa sprouts). The unsprouted mung bean seeds contain negligible amounts of vitamin C (ascorbic acid), whereas the sprouted seeds contain 20 mg per 100 g of sprouts, which is as high as tomato juice.

Jicama or yam bean. Jicama (*Pachyrrizus erosus*), also called the yam bean, is indigenous to Mexico and Central America. It belongs to the pea (Leguminosae) family. The name jicama was derived from the Aztec word *xicama* by the Spaniards who first found the crop growing in Mexico. The crop is grown for its enlarged turnip-shaped root, which is eaten raw or cooked, has a crisp texture, and is sweetish in taste. It is an important ingredient in Chinese cookery. Jicama is an herbaceous vine growing to lengths of more than 10 ft (3 m) and under short-day conditions

produces small white to deep violet flowers shaped typically like those of the common bean or pea. It is a tropical plant, and requires a long, warm, frost-free season for the formation of the enlarged root. In the tropics it is grown from sea level to about 5000 ft (1600 m) in elevation.

The seeds are planted 1–2 in. (2.5–5 cm) deep in rows 3–4 ft (1–1.3 m) apart and in rich, sandy or loose soil. The young plants, when about 4 in. (10 cm) in height, are thinned to about 1 ft (30 cm) between plants. The following fertilizers are recommended: nitrogen (35 lb/acre, about 39 kg/ha), phosphorus (30 lb/acre, 34 kg/ha), and potassium (10 lb/acre, 11 kg/ha). Additional nitrogen (about 30 lb/acre, 34 kg/ha) is supplied about the time the plants begin to bloom. For best-quality roots and high yields, the flower buds are pruned off so that seed pods are not allowed to develop. The roots are harvested when they are about 4 in. (10 cm) in diameter; larger roots, and plants left in the ground too long, tend to become fibery and of poor quality. The roots can be stored in a refrigerator or cool place under relatively high humidity for several weeks. The plants, including the enlarged roots, contain rotenone, a natural insecticide, so that no pesticide is needed to grow the crop.

Chinese winter melon. Chinese winter melon, wax gourd, winter gourd, white gourd, ash gourd, Chinese preserving melon, ash pumpkin, or tung kwa (*Benincasa hispida; B. cerifera*) is a monoecious (having male and female flowers separate on the same plant) viny annual cucurbit, probably of Indian or Chinese origin. It has been cultivated for more than 2000 years, and is reported to grow wild in Java. The immature fruits of this species have bristlelike hairs on the surface. The mature fruits are watermelon-shaped (**Fig. 2**); some varieties are somewhat spherical, and are 15–18 in. (38–46 cm) in length, 12–15 in. (30–38 cm) in diameter, and 15–30 lb (7–14 kg) in weight. With increasing maturity, most varieties develop a white waxy bloom on the fruit surface which thickens with time, even after harvest. The flesh of mature fruits is used in making Chinese soups. It can be eaten raw or made into sweet preserves similar to citron or watermelon rind preserves. For special occasions the skin is scraped off the fruit, the pulp and seeds are removed, the fruit cavity is stuffed with previously cooked meat and vegetables, and the entire fruit is steamed for a couple of hours before serving. The fruit is said to have medicinal qualities. A variety of this species is grown especially for the immature fruits, which are 4–6 in. (10–15 cm) long and 2–3 in. (5–7.5 cm) in diameter. These, called mo kwa by the Chinese, are used like summer squash.

Fig. 2. Mature Chinese winter melon (*Benincasa hispida*). (*University of California Agricultural Experiment Station*)

The crop is planted like winter squash or melons, and the vines are allowed to creep on the ground. The seeds are planted in hills about 6–10 ft (2–3 m) apart and in rich, loamy soil 1–2 in. (2.5–5 cm) deep. Fertilizer containing nitrogen, phosphorus, and potassium is applied at planting. Additional nitrogen is applied at first bloom. In southeastern Asia the vines and fruit are sometimes supported on trellises. Chinese winter melon fruits, when mature, store for more than 6 months, and often a year when stored in a cool, dry place such as a cellar. The temperature during storage should not be below 45°F (7.5°C) or above 65°F (18°C). The ideal storage temperature is 50–60°F (10–15°C). If kept at warmer temperatures, the storage time is reduced.

Balsam pear. Balsam pear, alligator pear, bitter gourd, bitter melon, bitter cucumber, or fu kwa (*Momordica charantia*) is a monoecious annual herbaceous vine of southeastern Asia origin; it is now widespread through the world tropics. The fruits are heart-shaped to cylindrical, tapered at the blossom end, and characterized by longitudinal rounded ridges; the entire surface of the fruit is rugose. The mature fruit size varies from 5 to 10 in. (12 to 25 cm) in length and from 2 to 3 in. (5 to 7.5 cm) in diameter. The fruit is extremely bitter; some of the bitterness is removed before cooking by peeling and steeping in salt water. Immature fruits are less bitter than mature ones.

The seeds of balsam pear are sown in rich loam soil about 1–2 in. (2.5–5 cm) deep in hills or beds 3–4 ft (90–120 cm) apart. When the plants are 3–4 in. (7.5–10 cm) in height, they are thinned to 15–18 in. (18–45 cm) between plants and staked or trellised. A complete fertilizer containing nitrogen, phosphorus, and potassium is applied at the planting of seeds, and additional nitrogen is side-dressed shortly after the vines have commenced to climb. As a vegetable, the fruits are harvested just prior to attainment of full size and while the seeds are still immature. The harvested fruits can be stored in a refrigerator for about a week. As with immature cucurbit fruits, prolonged storage in the refrigerator should be avoided, since physiological breakdown, called chilling injury, occurs at temperatures below 50°F (10°C), with subsequent rotting by decay organisms. The mature fruits, with seeds removed, are parboiled in salt water, and some of the bitterness is removed by draining and pressing or squeezing of the pulp, which is then added as an ingredient to other foods being cooked. The red aril of mature fruits is sometimes used as a condiment, and the red pigment is used as food coloring. The tender shoots and immature leaves are used as greens.

Chinese okra. Chinese okra or angled loofa (*Luffa acutangula*) is called zit kwa by the Chinese; it is a

Fig. 3. Chinese okra (*Luffa acutangula*). (*University of California Agricultural Experiment Station*)

close relative of *L. cylindrica*, known as loffa, sponge gourd, or dishcloth gourd. Chinese okra is a monoecious annual climbing vine grown for the immature fruits, which are harvested when the fibers have not yet developed and when the fruits are 4–6 in. (10–15 cm) in length. The origin of *Luffa* is thought to be in India, since wild forms of this genus have been found there. The club-shaped fruits of Chinese okra are oblong and pointed at the blossom end (**Fig. 3**); when mature they are 12–15 in. (30–38 cm) long and 3–3.5 in. (7.5–9 cm) in diameter at the thickest part. The 10 prominent sharp longitudinal ridges that extend from the peduncle to the blossom end characterize the fruit.

The seeds, which are poisonous, are planted in rich loam soil about 1–2 in. (2.5–5 cm) in depth in hills or in rows 6–6.5 ft (2–2.5 m) apart. The plants are spaced at 2–2.5 ft (60–75 cm) in the row. A complete inorganic fertilizer or well-decomposed manure is recommended. The crop is trellised or staked to allow the vines to climb for better quality and increased yields. Though not as large a fruit as *L. cylindrica*, the thoroughly mature fruits of Chinese okra can be made into vegetable sponge by retting until the outer walls are rotted, and then the skin, seeds, and pulp are washed away from the network of fibers. The fibers can be whitened with a dilute solution of bleach, followed by rinsing and drying.

Water spinach. Water spinach, water convolvulous, swamp cabbage, kang kong, or shui ung tsoi (*Ipomoea aquatica*) is a perennial semiaquatic tropical plant grown for its long, tender shoots. It is a relative of the sweet potato (*I. batatas*), but does not produce an enlarged root. Possibly of East Indian origin, it is now widespread throughout the tropics and semitropics, where it has become a very popular green vegetable. There are two types of water spinach: a narrow, pointed leaf type adapted for dryland culture and a broadleaf type adapted for aquatic culture. In the dryland culture the crop is seeded in beds and grown with abundant water. Irrigation is practiced in regions of low rainfall. It is very fast-growing, requiring only about 2 months under ideal conditions from seeding to first harvest. For aquatic culture the crop is propagated by stem cuttings 12–15 in. (30–40 cm) long. The cuttings are planted in puddled soil, similar to rice paddies, to a depth of 6–8 in. (15–20 cm), and the spacing between plants is about 12 in. (30 cm) apart in blocks. For rapid growth and high yields, 40–50 lb/acre (45–56 kg/ha) of ammoniacal nitrogen are recommended after the plants are established.

The first harvest for crops grown in water can be made as soon as the new shoots are more than 15 in. (40 cm) long. About 10–12 in. (25–30 cm) of the shoot is cut, and 8–10 shoots are bundled for marketing. New shoots continue to develop, so that harvest can be made every 7–10 days. For crops grown on dry soil, shorter shoots of 8 in. (20 cm) are harvested, since the stem sections further from the growing point are not as tender. Water spinach is used like common spinach in oriental cooking. It can be stored in a refrigerator under high humidity for a few days without much loss of quality. In southeastern Asia, parts not used for food are harvested and fed to livestock. Tender shoots from the sweet potato vine are often used instead of water spinach for human consumption. M. Yamaguchi

Bibliography. M. Yamaguchi, Production of oriental vegetables in the United States, *Hort. Sci.,* 85:362–370, October 1973; M. Yamaguchi, *Vegetable Crops of the World*, 1978.

Orion

The Hunter, a prominent constellation in the evening winter sky (see **illustration**). Orion is perhaps the easiest constellation to identify in the sky, because of the three bright stars in a line that form the belt of the mythical figure. The cool, red supergiant star

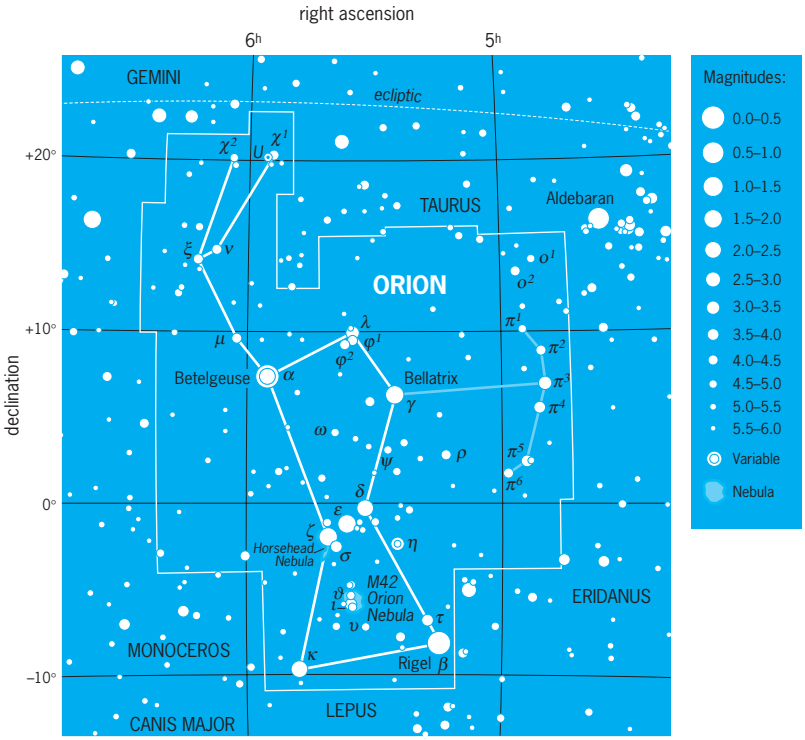

Modern boundaries of the constellation Orion, the Hunter. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

Betelgeuse glows brightly as Orion's shoulder, above which is an upraised club, and the hot, blue star Rigel marks Orion's heel. From the leftmost star of the belt dangles several faint stars and the Horsehead Nebula. The faint stars and the Orion Nebula, M42 (the 42nd object in Messier's eighteenth-century catalogue), one of the most prominent and beautiful emission nebulae and a nursery for star formation, make up Orion's sword. Four bright stars close together (within about 1 light-year) make the Trapezium, which provides the energy to make the Orion Nebula glow. *See* BETELGEUSE; ORION NEBULA; RIGEL.

Orion is pictured holding up his shield to ward off the charging Taurus, the bull. In Greek mythology, Orion met Artemis, goddess of the hunt and of the Moon. To protect her virginity, her brother Apollo sent Scorpius, the Scorpion, to attack Orion. He then tricked Artemis into shooting Orion. When Orion could not be revived, Artemis placed him in the heavens, with the scorpion eternally pursuing him. *See* SCORPIUS; TAURUS.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Orion Nebula

The brightest emission nebula in the sky, designated M42 in Messier's catalog. The Great Nebula in Orion consists of ionized hydrogen and other trace elements (**Fig. 1; Colorplate 1**). The nebula belongs to a category of objects known as H II regions (the Roman numeral II indicates that hydrogen is in the ionized state), which mark sites of recent massive star formation. Located in Orion's Sword at a distance of 460 parsecs or 1500 light-years (8.8×10^{15} mi or 1.4×10^{16} km), the Orion Nebula consists of dense plasma, ionized by the ultraviolet radiation of a group of hot stars less than 100,000 years old known as the Trapezium cluster (**Colorplate 2**). The nebula covers an area slightly smaller than the full moon and is visible with the aid of binoculars or a small telescope. *See* ORION.

In addition to the Trapezium cluster of high-mass stars, the Orion Nebula contains about 700 low-mass stars packed into an unusually small volume of space. Near the cluster center the average separation between stars is only about 5000 astronomical units (4.6×10^{11} mi or 7.5×10^{11} km), about 50 times less thanthe separation between stars in the Sun's

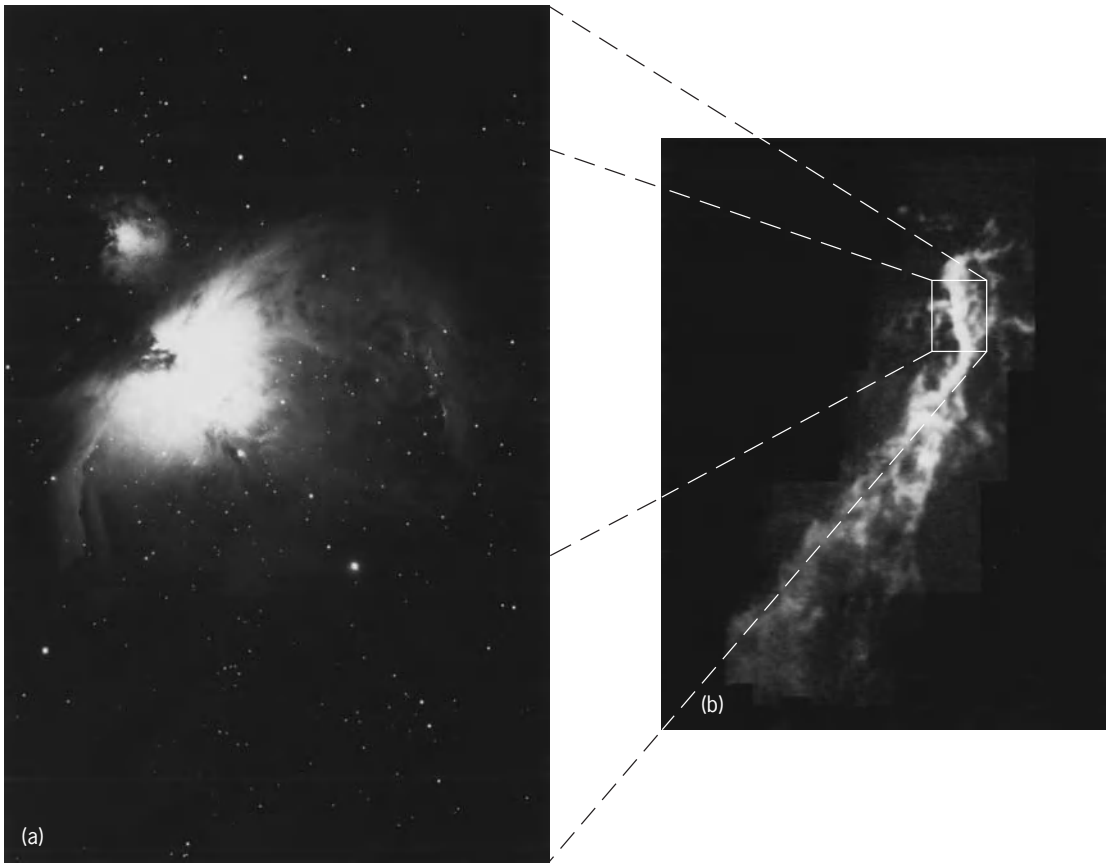

Fig. 1. Orion Nebula and its neighborhood. (*a*) Visual-wavelength view of the Orion Nebula obtained by taking a time exposure on a photographic plate with the 4-m-diameter (158-in.) telescope on Kitt Peak, Arizona (*National Optical Astronomy Observatories*). (*b*) Orion A giant molecular cloud as seen in the 2.6-mm-wavelength emission line of ^{13}CO. Data used to construct this image were obtained with a 7-m-diameter (23-ft) millimeter-wave telescope at Crawford Hill, New Jersey. The image shows a region which extends 5° north-south and 2° east-west. The Orion Nebula, located in front of the northern part of the molecular cloud, is not visible in this image. The approximate dimensions of the optical view are shown by the inset box. (*AT&T Bell Laboratories*)

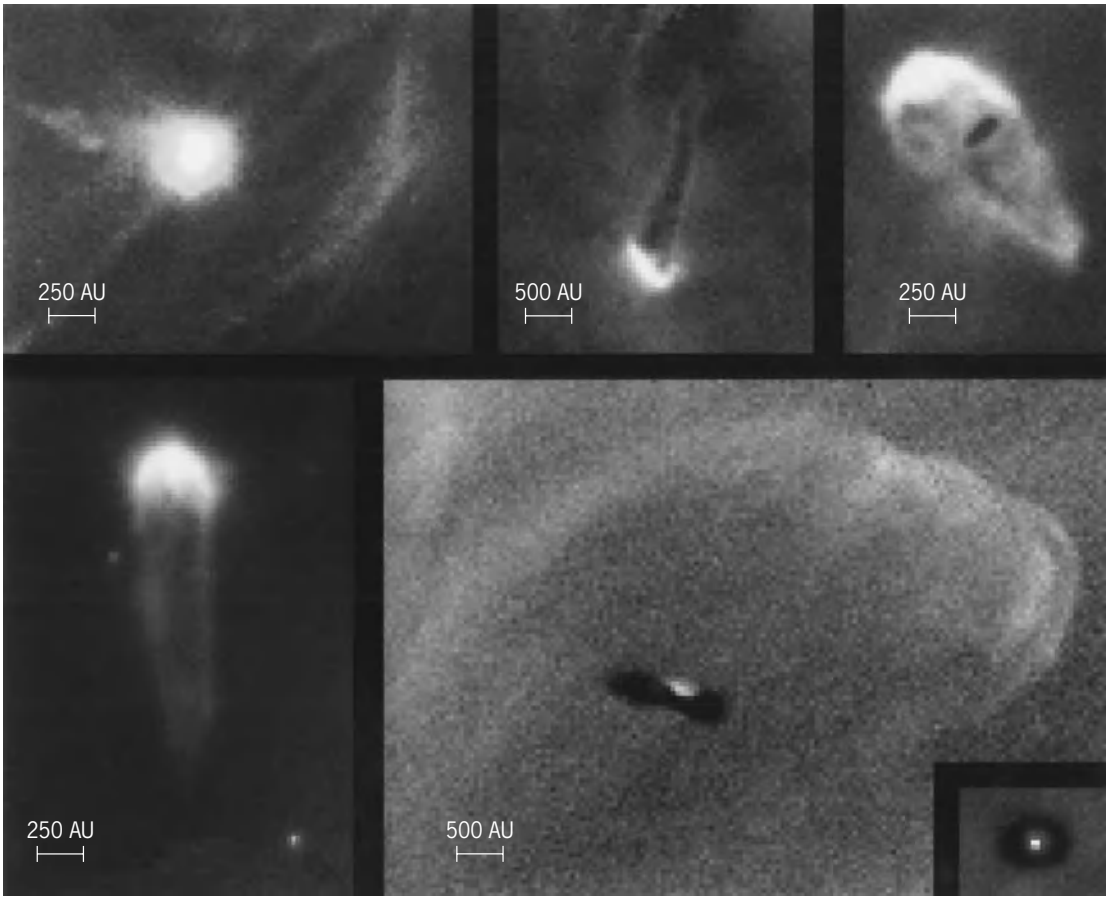

Fig. 2. Hubble Space Telescope image showing closeups of several disks and ionized structures surrounding low-mass stars embedded within the Orion Nebula. The stars located near the Trapezium stars exhibit bright comet-shaped regions of gas expanding away from hidden disks. However, these circumstellar disks can be seen directly in silhouette against background nebular light for stars located sufficiently far from the Trapezium stars. The scale of each image is shown in astronomical units. (*John Bally, Dave Devine, and Ralph Sutherland*)

vicinity. However, as the Orion Nebula evolves, the cluster is expected to expand and possibly dissolve into the smooth background of stars in the Milky Way.

Observations with the Hubble Space Telescope have shown that many low-mass stars in the Orion Nebula are surrounded by disks of dense gas and dust (**Fig. 2**). These so-called proplyds (derived from the term "proto planetary disks") appear to be rapidly losing their mass as they are irradiated by intense ultraviolet radiation produced by high-mass stars in the nebula. They may survive for only 10^4–10^5 years in this harsh environment. However, the proplyds may eventually evolve into planetary systems if most of the solids in these disks have already coagulated into centimeter-sized objects which are resistant to ablation by the radiation field.

Stars form by the gravitational collapse and fragmentation of dense interstellar molecular clouds. A star-forming molecular cloud core (known as OMC1 for Orion Molecular Cloud 1) is hidden behind the Orion Nebula by a shroud of dust. A group of infrared sources, believed to be stars too young to have emerged from their dusty birth sites, is buried within OMC1. Although invisible at optical wavelengths, OMC1 has been investigated in the infrared and radio portions of the spectrum. Observations of the massive young stars in OMC1 have shown that stellar birth is associated with powerful outflows of gas, naturally occurring maser emission, and shock waves that disrupt the remaining cloud core.

The cloud core behind the Orion Nebula is only a small part of a giant molecular cloud (the Orion A cloud), 100,000 times more massive than the Sun (Fig. 1*b*). Although mostly made of molecular hydrogen (H_2), the cloud is best traced in the 2.6-mm-wavelength emission line ^{13}CO, the rarer isotopic variant of carbon monoxide (CO). Over the past 1×10^7 years, the Orion A cloud, together with a second giant molecular cloud in the northern portion of Orion (the Orion B cloud), have given birth to the Orion OB association, a loose, gravitationally unbound grouping of hot massive stars of spectral types A, B, and O. The Orion OB association also contains tens of thousands of low-mass young stars which formed from the Orion molecular clouds. The collective effect of the ionizing radiation and winds produced by these stars, and the supernova explosions occurring at their death, has generated an expanding bubble 70 by 200 pc (1.3 by 3.8×10^{15} mi or 2.2 by 6.2×10^{15} km) in extent in the interstellar medium. The eastern portion of this bubble is

running into dense gas near the plane of the Milky Way Galaxy, where faint optical emission is seen as the 8° radius arc of nebulosity called Barnard's Loop. *See* INTERSTELLAR MATTER; NEBULA; RADIO ASTRONOMY; STAR; STELLAR EVOLUTION; SUPERNOVA.

John Bally

Bibliography. J. Bally, C. R. O'Dell, and M. J. McCaughrean, Disks, microjets, windblown bubbles, and outflows in the Orion Nebula, *Astron. J.,* 119:2919–2959, 2000; A. E. Glassgold, P. J. Huggins, and E. L. Schucking (eds.), Symposium on the Orion Nebula to Honor Henry Draper, *Ann. N. Y. Acad. Sci.*, 395:1–338, 1982; J. M. Pasachoff and A. Filippenko, *The Cosmos: Astronomy in the New Millenium*, 3d ed., Brooks Cole, 2007.

Ornamental plants

The use of ornamental plants developed early in history for it is known that primitive peoples planted them near their dwellings. There was much use of flowers in the Greek and Roman worlds and a small plant nursery was found in the ruins of Pompeii. Gardening in various forms has been a part of the cultures of all known great civilizations. Although increased urbanization has often reduced the areas which can be devoted to ornamentals, nevertheless there is a growing interest in them and the sums spent for plants and horticultural supplies are quite large. New plants from various parts of the world are constantly being introduced by individuals and by botanic gardens and arboreta.

Types and uses. Ornamentals have been selected from virtually the whole kingdom of plants (**Fig. 1**). Ground covers include turfgrasses, herbaceous annuals, biennials or perennials, and some scandent shrubs or vines. The woody structural landscape plants include vines (deciduous and evergreen), shrubs (deciduous and evergreen), deciduous trees, broad-leaved evergreen trees, gymnosperms and cyads, bamboos, and palms.

Plants for special habitats or uses include lower plants such as mosses, selaginellas or horsetails, ferns and tree ferns, sword-leaved monocots, bromeliads and epiphytic plants, aquatic plants, cacti and succulents, and bulbous and cormous plants. Interior decorative plants in a large measure are shade plants from the tropics.

The best choice of an ornamental plant for a given situation depends on the design capabilities of the plant, the ecological requirements, and various horticultural factors. The plant must fulfill the purposes of the designer regarding form, mass, scale, texture, and color. It must also have the ability to adjust well to the ecology of the site in such factors as temperature range, humidity, light or shade, water, wind, and various soil factors, including drainage, aeration, fertility, texture, and salinity or alkalinity. Some important horticultural factors are growth rate, ultimate size and form, predictability of performance, and susceptibility to insects, diseases, deer, gophers, snails, and other pests. The hardiness of the plant is often important, and pruning and litter problems may influence choice.

Large decorative foliage plants are used extensively in homes and offices. In many buildings the architects have included large planter areas where plants may be grown both indoors and outdoors. Where the light level is too low, banks of lamps have been placed where they can assist plants to keep in condition for long periods of time. Because some attention is required to keep plants attractive, there have been some failures in their upkeep. This has encouraged the development of artificial plants and flowers which are increasingly accurate and pleasing representations, especially if they are not inspected closely. Sometimes artificial plants are used in the background, with living plants in the foreground. Artificial flowers, on the other hand, do not seem to have lessened the demand for cut flowers. One way of solving the problem of regular care needed by living plants has been to make available special equipment, professional care, and replacement of plants after their installation. Several automatic watering systems have been developed which aid in maintenance.

Propagation and culture. Scientific advances have been made in plant propagation and culture. The use of small amounts of auxins, such as naphthalene acetic acid or indole butyric acid, promotes root formation on cuttings. Various types of intermittent mist and fog systems prevent cuttings from drying out and often improve rooting. *See* AUXIN.

An improvement in propagation is the introduction of tissue (meristem) culture of orchids into nutrient solutions. This permits the rapid large-scale multiplication of superior cultivars. However, the details of procedure are not yet established for all types of orchids. Meristem culture sometimes permits the elimination of a virus or a fungus from the rapidly expanding growing point of a plant. This system has been used with carnations and chrysanthemums to avoid diseases of several types and to produce clean propagating stock. *See* APICAL MERISTEM; TISSUE CULTURE.

The use of photoperiod to control the blooming time of plants has been important economically. Day length can be augmented by the use of electric lights, or shading with a black cloth may be used to increase the length of the dark periods. Thus it is now possible to buy blooming chrysanthemums and certain other flowers every day of the year. *See* PHOTOPERIODISM.

Automatic watering systems have been devised for plant growing. These can also be connected to tensiometers or several other devices to measure the soil moisture. *See* PLANT-WATER RELATIONS.

Plant growth retardants such as B-9, Phosphon, and Ammo 1618, when applied to certain plants such as chrysanthemum, poinsettias, and azaleas, may produce compact specimens which are superior in appearance. In some instances, the lasting qualities of the plant are greatly improved and flowering may be enhanced.

The composition of soil mixes for the growth of

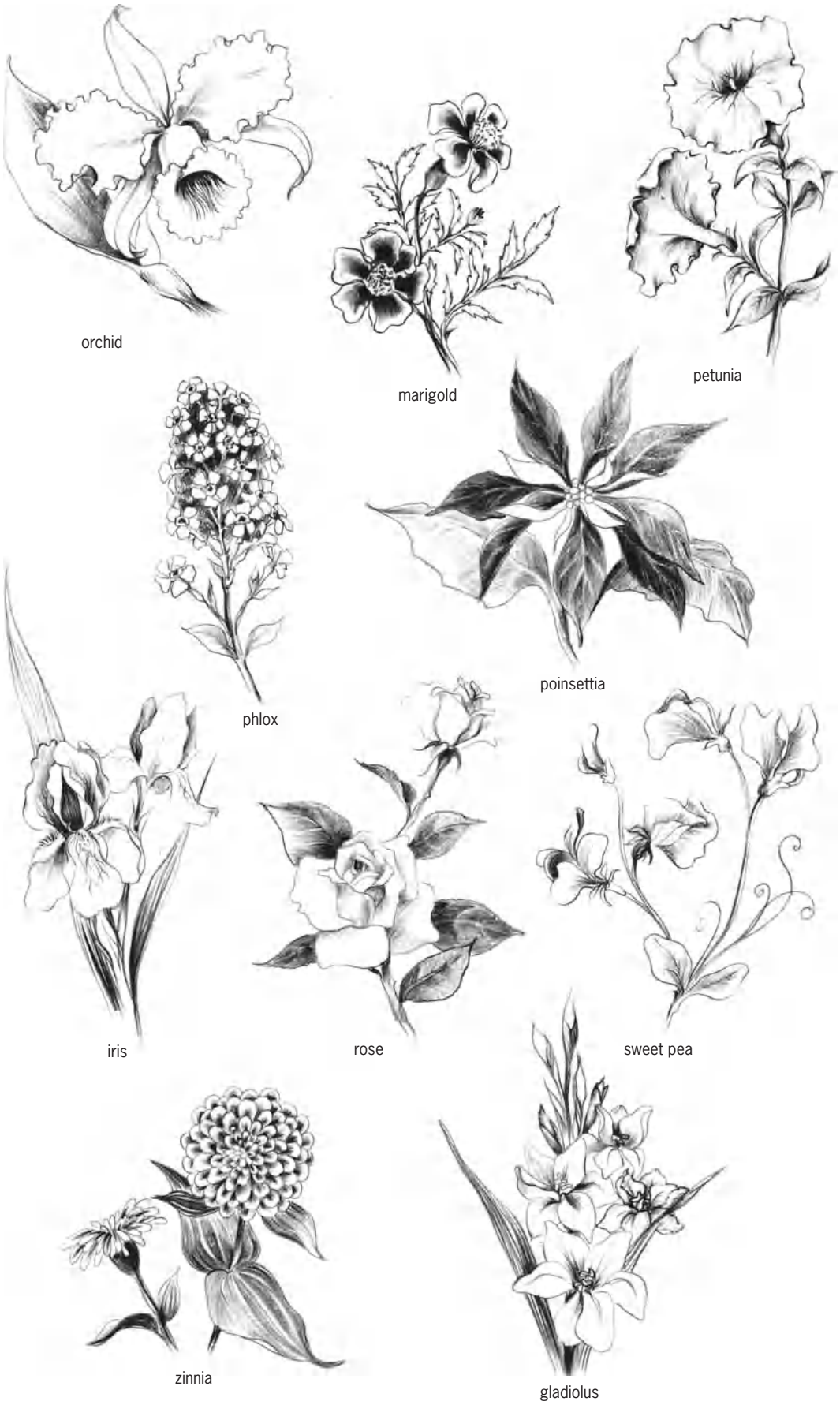

Fig. 1. Representative ornamental flowers used by amateurs and professionals.

container plants either in a commercial nursery or in a garden is important. The outcome of considerable experimentation has been the development of soil mixes for container growing which have a large proportion of relatively inert materials, such as sand, rice hulls, sawdust, forest tree bark, perlite, or vermiculite. Some soil mixes are based mainly on a fixture of fine sand and sphagnum peat moss. These mixtures have an advantage in that they may be used immediately after steam sterilization. Peat and perlite are used in other soil mixes which are quite close in characteristics to those above. Both mixes require a regular application of fertilizers unless slow-release types are incorporated. *See* FERTILIZER.

Many nursery plants are produced in salvage metal containers and several machines are available which assist in shifting small nursery plants. Container growing is more common in warmer areas than in colder parts of the United States.

Horticultural maintenance. Many traditional gardening tools have been improved or mechanized, which reduces labor costs. Electric clippers permit much greater use of clipped hedges; pneumatic machinery is available for tree pruning; and power grass mowers and edgers are widely used.

Herbicides meet the needs of practically any situation and save hand weeding. The new slow-release fertilizers furnish plant nutrients over a long period of time and are particularly helpful in container plant growing. The appropriate chelates furnish trace elements such as iron or zinc to plants and thus largely avoid deficiency symptoms. The use of liquid fertilizers in watering has been simplified by the use of automatic proportioning pumps. Often liquid fertilizer is applied with each watering of plants in the nursery. *See* HERBICIDE.

New and improved fungicides and insecticides are constantly being introduced. The systemic insecticides in which the chemical is applied to the soil and taken up by the plant are especially useful since residue problems do not exist with ornamentals. They are being used by both amateurs and professionals. Both groups are also beginning to use soil fumigation prior to planting. Horticultural maintenance on a contract basis is common now in some parts of the United States and tree surgery is done by specialists. *See* FUNGISTAT AND FUNGICIDE; INSECTICIDE.

Breeding. Outstanding plant breeding has been done by amateurs working with popular garden flowers of minor economic value. Notable examples include such plants as the daylily, iris, and dahlia. These nonprofessionals have frequently formed societies for the exchange of information and breeding stocks. Some of the outstanding developments in breeding ornamental plants are listed in the **table**.

Professional plant breeders work largely with important commercial flowers such as the rose, lily, orchid, gladiolus, and chrysanthemum. The United States plant patent law has given protection to hybridizers and all vegetatively propagated plants are covered. This law has given a great impetus to plant breeding, particularly that of roses. Test gardens for new varieties (cultivars) are located in various parts of the country to assist in evaluating the new varieties. *See* BREEDING (PLANT).

Notable developments in ornamentals

Flower	Development
African violet	Polyploids; many new foliage and flower characters
Aster	Wilt resistance; cut-flower types
Camellia	The Williamsi hybrids (*Camellia japonica* and *C. salenesis*); spectacular new varieties of *C. reticulata* from Yunnan, China; intensive breeding of many types
Carnation	The Sims varieties (U.S.A.), outstanding for commercial cut flowers; spray-type carnations
Fuchsia	Heat-resistant types (California); pure white flowers (California)
Gladiolus	New miniature types; breeding for disease resistance and vigor
Hemerocallis	Purer pinks and reds; recurrent blooming habit; polyploid types
Iris	Superior new varieties of Dutch iris; tall bearded iris, with purer pinks and recurrent blooming habit
Marigold	F_1 hybrids; closer approach to white; scentless foliage types
Orchid	Complex species hybrids; polyploid forms with improved flower quality; miniature flowered cymbidiums; breeding for specific blooming season
Petunia	F_1 hybrids; pure scarlet crimson color; giant ruffled and all-double flower types
Phlox	Tetraploid annual varieties
Poinsettia	New flower and foliage types with longer-lasting qualities
Rose	Large-flowered polyantha types ("floribundas" and "grandifl oras"); lavender flower color; introduction of varieties with flowers of superior keeping quality; new miniature types; intensive breeding for vigor, disease resistance, fragrance, and new colors
Shasta daisy	New single and double types for florist use
Snapdragon	F_1 hybrids; giant tetraploids; reduction of flower shattering; ruffled double-flowered types; rust resistance
Stocks	Mosaic resistance; high double strains
Sweet pea	Heat-resistant strains; many-flowered (multiflora) strains
Zinnia	F_1 hybrids; new twisted-petal types; giant size

Occasionally, ionizing radiation has been applied to ornamentals to produce new forms. The use of colchicine to double chromosomes has been a valuable technique both for producing direct plant changes and for making new crosses possible. Spontaneous mutations often occur when large plant populations are grown. Many new types of annual flowers introduced by seed growers arose in this manner. *See* MUTATION.

The producers of plants grown from seed have no legal protection for their new productions. However, some protection is being obtained by the production of hybrid seed from parents, which are kept by the originator. In this instance seeds from the hybrids may not come true, and thus the seed producer is protected from competition. The flower seed industry does much plant breeding which is concentrated in the cool coastal valleys of central and southern California.

Air transport has revolutionized the growing of floricultural crops and has brought many

competitive advantages to California and to southern Florida. The production is in glasshouses, in cloth houses, and in many instances in open fields. Various types of plastic coverings are used for some crops as a cheaper substitute for glass; however, many of these last only one season. Plastics laminated with fiberglass have a much longer life and are used more. *See* ARBORETUM; BOTANICAL GARDENS; FLORICULTURE. Vernon T. Stoutemyer

Diseases of woody ornamentals. Woody ornamentals are subject to a number of diseases. Some diseases are inconspicuous, causing little or no permanent injury, while others, by severely injuring leaves, twigs, branches, or roots, cause plants to become stunted or to die.

Nonparasitic disease. Environmental stresses commonly affect landscape plants, which are often not genetically adapted to the sites where they are placed. Nutrient stress may be a problem in disturbed or altered soils. In alkaline soils, essential nutrients may not be available to the plants, and chlorosis or yellowing of leaves is a common symptom. Sensitive plants should be grown in a somewhat acid soil.

Drought or flooding are the climatic factors most likely to cause stress damage. Water or air deficiency in the soil causes similar symptoms in the aerial portion of the plant: wilt, leaf necrosis, and, when persisting, death. Both air and water are also important for healthy roots.

Low-temperature injury is more common than high-temperature injury. Rapid drops in temperature to below freezing, following mild weather, often injures plants that are not fully acclimated to low temperatures. Low temperatures dehydrate stem tissues, and vertical slits in the stems (frost cracks) may occur. Below-freezing temperatures coming after shoot emergence in the spring often kill succulent tissues. *See* COLD HARDINESS (PLANT).

Pollutants can be present in the soil or in the air and may result in chronic or acute plant damage. The pollutants that kill plants usually originate in stationary manufacturing plants: sulfur dioxide and fluorides are examples. Ozone and peroxyacetyl nitrate, pollutants associated with automobile emissions, injure leaf surfaces but rarely cause defoliation. Herbicides are considered pollutants when misapplied and may cause leaf and stem distortion, necrosis, and plant death.

Pathogens. Fungi are the infectious agents most frequently reported to cause diseases of landscape plants, but bacteria, viruses, nematodes, and mycoplasmalike organisms also can be serious pathogens.

Pathogens that cause leaf spots or blights do not greatly weaken the host plant unless premature defoliation occurs. Evergreens may be seriously injured after one defoliation, but two or more consecutive years of defoliation usually are required to kill deciduous plants or to cause extensive dieback in them. Mildews, rusts, blights, and leaf spots can be controlled with applications of protective fungicides.

Cankers and dieback result when fungal pathogens kill portions of the vascular cambium of twigs and stems. The extent of injury by these organisms is modified by environmental conditions. Vigorous, well-maintained plants are resistant to many of these pathogens. Pruning out the affected limbs is drastic but is often the only control for susceptible plants.

Root and crown rot pathogens cause plants to have small leaves, thin foliage, dieback, and slow decline. These diseases are most common when plants are in moist or improperly drained soil. Fungicide drenches and dips can reduce the likelihood of their occurrence.

Fungal or bacterial pathogens that become systemic in the vascular tissues of plants bring about the most rapid plant mortality. Those in the wood prevent sap flow: those in the bark prevent carbohydrate translocation. Satisfactory control measures are not now available for vascular wilts. Dan Neely

Diseases of herbaceous ornamentals. Ornamental plants include a large number of herbaceous genera that are susceptible to one or more diseases that affect growth, flowering, or reproduction of the plant. This section of the article describes some of the most common diseases and the kinds of pathogenic organisms that cause them.

Classification. Diseases may be classified as parasitic and nonparasitic. Parasitic diseases are those caused by microorganisms such as fungi, bacteria, viruses, and nematodes. To this group have been added viroids, mycoplasma, and spiroplasma. Nonparasitic diseases are those induced by environmental stress that adversely affect plant growth. Only parasitic diseases will be described below.

Diseases are classified according to the types of symptoms they produce. Microorganisms that cause parasitic diseases may affect roots, stems, leaves, and flowers of fruits. Symptoms of these diseases may be described as leaf spots, blights, cankers, galls, and scorches if they occur on the aboveground portion of the plant where the organism has invaded host tissue. The causal organism can often be isolated from leaves, stems, flowers, and fruits. Fungi and bacteria of different genera can cause one or more of these symptoms.

Fungal and bacterial diseases. Leaf spots are commonly observed on many plants. Fungus leaf spot diseases often spread rapidly by spores produced in the leaf spots. Infected leaves drop prematurely, and susceptible varieties may defoliate rapidly. Many fungi that cause leaf spots of ornamentals overwinter in leaf debris and diseased stems; such stems should be removed and diseased leaves should be destroyed to reduce subsequent infections. Fungicides applied after the diseased leaves are removed will protect the newly formed leaves from infection.

Bacteria also produce leaf spots. Black leaf spot on *Delphinium* is caused by *Pseudomonas delphinii*. Diseased leaves are removed and plants sprayed with streptomycin to control the disease.

Powdery mildew is caused by a group of many closely related fungi that commonly affect the leaves but may also affect the flowers and fruit. A white coating that appears like powder is formed in the spring and summer. The fungus survives the winter as strands or hyphae on shoots that again produce

spores in the spring. The fungus can be controlled by interrupting its life cycle by removing and destroying all diseased leaves to eliminate the overwintering stage. Destruction of infected leaves that drop from the plant will reduce the capability of the fungus to reproduce further and spread during the following spring. This sanitation approach to disease control is an important principle in controlling many diseases caused by foliar pathogens of ornamental plants. During the growing season, chemicals may also be applied to control spread of the disease. Powdery mildew commonly causes diseases on rose, chrysanthemum, lilac, and many other ornamentals.

Blight is a nonspecific term that is usually applied to a disease that affects young growing tissue, including developing buds, flowers, and shoot tips. A common garden blight of ornamentals is caused by the fungus *Botrytis cinerea* and is referred to as gray mold. The disease is common on flowers of rose, lily, geranium, chrysanthemum, carnation, snapdragon, gladiolus, and many other garden flowers. The disease often occurs in the garden during wet weather and may cause rapid rotting of the flowers. The disease may be controlled by removing and destroying affected flowers and by frequent fungicidal applications during periods of wet weather. The fungus forms small black structures (sclerotia) on or in infected tissue, and these structures may survive and overwinter on debris in the soil.

Canker refers to disease of the stem and branches that usually occur on woody ornamentals. The canker appears as a localized eruption that may encircle the stem and cause collapse. Several different fungi produce cankers on rose, and they are controlled by removing and destroying affected canes. Fungicidal sprays are also applied to control these diseases.

Galls are swollen tissues. They may appear as a proliferation of cell growth in flowers, leaves, or roots. Crown gall is a bacterial disease caused by *Agrobacterium tumefaciens*. The bacterium has a wide host range and produces galls on the roots and on stems aboveground. The disease is controlled by cutting off the galls and destroying them and dipping the roots and stem in a solution of the antibiotic streptomycin. Another approach to control is to dip the roots in a suspension of an avirulent strain of the bacterium. *See* CROWN GALL.

Serious root diseases are caused by fungi and nematodes. Disease agents that attack the roots and the bases of the stems are often not recognized until serious damage has been done. Symptoms may appear as wilting and dwarfing. Controls must often be applied as preventive rather than therapeutic measures because the plant may be killed if infection is not prevented. An example of a fungus disease of seedling stems and roots is damping-off caused by *Rhizoctonia solani* and *Pythium* and *Fusarium* species. These fungi often kill germinating seedlings, and are controlled by treating the seed or soil with an appropriate fungicide. Several species of *Fusarium* also cause wilt of plants or individual branches. The conductive tissues in the stem may show a darkening and discoloration. Infected plants show the most prominent symptoms in hot weather, and the plants die. Another root pathogen is *Verticillium*. This fungus thrives in cool weather and will also kill susceptible species. *Fusarium* species are usually specific to one kind of plant, but *Verticillium* attacks many different plants, including annuals and woody species. Both of these fungi survive for long periods in the soil. Infected plants should be removed and burned and resistant varieties planted in soil where the fungus is known to occur.

Rusts are complex fungi that produce as many as five different forms of spores during the life cycle. Rusts are obligate parasites; that is, they must develop in living host tissue. Symptoms appear as reddish-brown spots on the leaves. Some common rusts occur on carnation, hollyhock, snapdragon, aster, and chrysanthemum. Rust diseases may be difficult to control; whereas some rusts complete their life cycle in one host, others must infect two unrelated species to complete the life cycle. Control measures include the application of fungicides and removing and burning diseased leaves to destroy overwintering stages.

Viral diseases. Viral diseases can be very destructive. Viruses move systemically throughout the roots, stems, and leaves. Many viruses are transmitted by rubbing juice from an infected plant onto healthy plants, and many are also transmitted by insects. Diseased plants may be stunted with yellow mottling and ring patterns on the leaves and light- or dark-colored streaks on the flowers. Cucumber mosaic virus infects many species, including gladiolus, lilies, dahlias, and delphiniums. This virus is transmitted by aphids, and the virus spreads rapidly with an increase in aphid populations. Viral disease symptoms may appear only within a particular range of temperature. Disease symptoms may be suppressed in plants grown at unusually high or low temperatures. Infected plants should be removed and destroyed, especially perennial and bulb crops that serve as a source of continuing infection. There are no known chemicals that can be applied to control virus diseases.

Mycoplasmas. Yellows diseases are caused by mycoplasmas that are transmitted by leafhoppers. Aster yellows produces stunting and yellowing of the leaves and growth proliferation producing a "witches'-broom" type of growth. Tetracycline antibiotic sprayed on infected plants has induced a remission of symptoms and resumption of normal growth. However, symptoms reappeared when the antibiotic treatment was terminated.

Viroids. Viroids are a newly described class of disease agents that are much smaller than viruses. Chrysanthemum stunt and chrysanthemum chlorotic mottle are important viroid diseases. Diseased plants should be destroyed.

Control. Parasitic disease control depends on the accurate identification of the pathogen. The kind of pathogen will determine the specific control procedures. Sanitary procedures that reduce the source of infection; healthy seeds, bulbs, and cuttings; disinfection of tools; pasteurization of soil; and chemicals

applied to seed, to soil, or directly to the diseased plant may be used separately or in combination to produce healthy, productive plants. Roger H. Lawson

Diseases of foliage ornamentals. Diseases of foliage ornamental plants, as with any plants in fact, can occur only when the proper combination of the pathogen, host plant, and environmental conditions exists. With the surge of interest in foliage plants for utilization in the home, office, shopping plazas, public agencies, and private institutions, disease problems have become more important. The growth and well-being of such plants, when not grown in the natural environment where they have evolved successfully against adversity, inevitably result in myriads of problems. Of these, diseases often become manifest, often resulting in death of the plant.

Foliage ornamental diseases are usually the result of improper care or lack of knowledge of the requirements for healthy plants, that is, light, moisture, temperature, and nutrition, as well as sanitation. When grown in unnatural situations, they are placed under adverse conditions resulting in debilitation, thus subjecting them to plant-invading organisms.

Plant pathogenic bacteria (*Xanthomonas, Pseudomonas, Erwinia, Agrobacterium, Corynebacterium*) can be soil-borne or occur in infected plants. The most common symptoms of bacterial infections are leaf spots having a water-soaked appearance, wilting of the foliage, and soft-rotting of underground plant parts. Control of bacterial diseases lies in the use of healthy, disease-free plants, "clean" soil, and prevention of contamination of healthy plants by infected plants or contaminated equipment and splashing water. Chemical control is not considered practical for most bacterial diseases of foliage ornamentals.

Pathogenic fungi (*Rhizoctonia, Sclerotium, Sclerotinia, Cercospora, Fusarium, Pythium, Phytophthora, Verticillium, Cylindrocladium,* and so on) are capable of attacking virtually all parts of plants, and are the most common and generally most readily visible of the disease-causing organisms. They cause rotting of the roots, plugging of the vascular system, stem cankers and galls, damping-off (death of the plant just before or soon after emerging from the surface of the soil), and leaf spots (**Fig. 2**) and distortion. Control of fungal diseases can often be achieved with proper chemicals applied at the early stages of disease symptoms, provided that a proper diagnosis is made as to the causal agent.

Viruses or viruslike agents are the most difficult to control. Residing in a plant in almost complete harmony, these organisms are often referred to as the perfect parasites since they rarely, except in a few well-known instances, cause the obvious and readily observed destruction inflicted by bacteria and fungi. Most symptoms of viral diseases are subtle but nevertheless injurious. The range of symptoms include mosaics, flecking and mottling of leaves and sometimes stems, vein necrosis or discoloration, leaf and flower distortion and discoloration, and stunting and general unthriftiness of the plant. Virus-infected plants are best discarded, since there are no practical cures for these diseases. Insects are known to be vectors of some viruses or viruslike agents; hence, they should be controlled with proper insecticides. *See* PLANT PATHOLOGY. S. A. Alfieri, Jr.

Fig. 2. Leaf spot of *Syngonium podophyllum* caused by the fungus *Cephalosporium cinnamomeum*.

Ornithischia

One of two constituent clades of Dinosauria (the other being Saurischia). Ornithischian dinosaurs are characterized by the possession of numerous anatomical features, notably the configuration of the pelvis. In ornithischian hips, the pubis bone (which points downward and forward in the majority of other reptile groups) has rotated posteriorly to lie parallel to the ischium (**Fig. 1**). This configuration is superficially similar to that seen in living birds, but actually represents an example of evolutionary convergence (birds are, in fact, descended from saurischian dinosaurs). Other features that characterize ornithischians include the presence of ossified tendons that form a complex lattice arrangement supporting the vertebral column, a single midline bone called the predentary that lies at the front end of and connects the two lower jaws, and a palpebral bone that traverses the orbit (**Fig. 2**).

Distribution. The earliest known ornithischian, *Pisanosaurus*, was discovered in deposits of Late Triassic age (approximately 225 million years ago) in Argentina. During the Late Triassic–Early Jurassic interval, ornithischians were relatively rare components of dinosaur faunas; however, the group began to radiate spectacularly in terms of species richness, abundance, and morphological diversity during the Middle Jurassic. From the Middle Jurassic onward, ornithischians attained a global distribution and became the dominant terrestrial vertebrates of the Cretaceous. Almost all ornithischians were herbivorous,

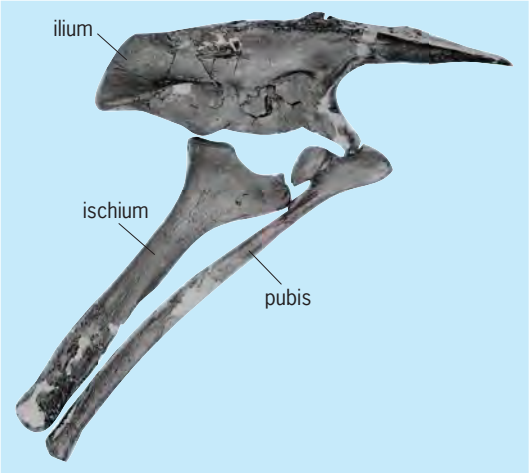

Fig. 1. Pelvis of the Lower Jurassic ornithischian dinosaur *Scelidosaurus harrisonii* from England (Natural History Museum, London specimen BMNH R1111). The front of the animal is to the right of the photograph. Note that the pubis bone has rotated backward to lie alongside the ischium, as in all ornithischian dinosaurs, in contrast to the usual reptilian condition in which the pubis extends forward, away from the ischium.

though a few primitive forms may have been omnivores.

Phylogeny and evolution. Phylogenetic analyses of ornithischian interrelationships indicate that the clade can be subdivided into three major lineages (Thyreophora, Marginocephalia, and Ornithopoda), along with a small number of primitive forms that do not fall within any of these groupings. Primitive ornithischians, such as *Lesothosaurus* from the Early Jurassic (approximately 205 million years ago) of southern Africa, were small [1–2 m (3–6.5 ft) in body length], bipedal animals with leaf-shaped teeth for shearing plant material, long counterbalancing tails, and grasping hands; this basic design was modified and elaborated in later ornithischian groups. Ornithischian evolution was dominated by several major themes: experimentation with stance (quadrupedality vs. bipedality), the appearance of ever more complex feeding mechanisms (such as alterations to teeth and skull morphology), and the development of sociality (gregarious living, parental care, and intra- and interspecific communication).

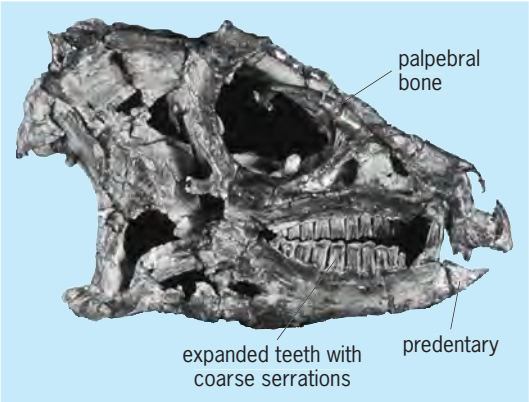

Fig. 2. Skull of the Lower Jurassic ornithischian dinosaur *Heterodontosaurus tucki* from South Africa (Iziko: South African Museum specimen SAM-PK-K1332). Note the predentary bone at the tip of the lower jaw, the palpebral bone crossing the eye socket, and the expanded, coarsely serrated teeth. These features are characteristic of all ornithischian dinosaurs.

Thyreophora. Thyreophora (the "armored" dinosaurs) consists of two clades: Stegosauria and Ankylosauria. Both of these groups were quadrupedal and characterized by the possession of numerous osteoderms (bony plates) that were embedded within the skin.

Stegosauria. Stegosaurs are known from the Middle Jurassic–Early Cretaceous of North America, Africa, Europe, and Asia. The osteoderms were either flat, subtriangular plates (as in *Stegosaurus*: Late Jurassic, North America) or spinelike (as in *Kentrosaurus*: Late Jurassic, Tanzania). The plates and/or spines were arranged in two parallel rows that extended along the back and tail; a large spine was also sometimes present in the shoulder region. The arrangement of plates and spines differs between stegosaur species, perhaps indicating a role in species recognition or intraspecific display. In addition, spines were likely to have had a defensive function; the plates were incised with many grooves for small blood vessels and may have had a thermoregulatory function.

Ankylosauria. Ankylosaurs appeared in the Early Jurassic (*Scelidosaurus* from England; Fig. 1) and persisted until the end of the Cretaceous Period. They were relatively rare during the Jurassic but increased in diversity during the Cretaceous, at which time they had a worldwide distribution. Ankylosaurs possessed an extensive covering of armor, consisting of multiple rows of low armor plates and spines covering almost all surfaces of the body except the underside, giving them an almost impenetrable tanklike exterior. Some genera (for example, *Ankylosaurus*: Late Cretaceous, North America) possessed tail-clubs made of solid bone, which were presumably used as defensive weapons. One genus (*Euoplocephalus*: Late Cretaceous, North America) even possessed a bony eyelid.

Marginocephalia. Marginocephalia are united by the possession of a distinct shelf of bone at the rear of the skull and consist of two groups: Pachycephalosauria ("bone-headed" dinosaurs) and Ceratopsia ("horned" dinosaurs).

Pachycephalosauria. Pachycephalosaurs were bipedal animals whose skulls were thickened, leading to the development of a domelike structure in some cases. This dome [which could reach up to 25 cm (10 in.) thickness in *Pachycephalosaurus*: Late Cretaceous, North America] was probably used for butting rivals during intraspecific combat. Pachycephalosaurs were relatively rare animals that were restricted to the Cretaceous of North America, Asia, and possibly Europe.

Ceratopsia. Early ceratopsians, such as *Psittacosaurus* (Early Cretaceous: China and Mongolia) were small bipeds, lacking horns and possessing only short frills that extended from the back of the skull. However, later forms such as *Triceratops* (Late Cretaceous: North America) were large quadrupeds, which often displayed an array of horns (situated on the tip of the snout and/or above the eyes) and an

impressive frill that could be over 1.5 m (5 ft) in length. All ceratopsians had parrot-like beaks for cropping vegetation and guillotine-like dentitions for processing fibrous plant material. Ceratopsians were limited to the Cretaceous of Asia and North America, though they were locally abundant.

Ornithopoda. A conservative but very successful group that retained the bipedal pose of primitive ornithischians, ornithopods are first known from the Middle Jurassic and persisted until the end of the Cretaceous. All ornithopods possessed sophisticated grinding jaw mechanisms, allowing them to process vegetation as efficiently as living grazers and browsers (for example, horses and cows). One subgroup of ornithopods, the hadrosaurs ("duck-billed" dinosaurs), had teeth arrayed in dense "batteries," consisting of multiple columns and rows of functional teeth that provided a large continuous surface area for chewing. Hadrosaurs were also among the most social of dinosaurs: one genus, *Maiasaura* (Late Cretaceous: North America), is known from dozens of eggs, nests (which are found grouped together in colonies) and babies, as well as numerous associated adult remains. Some evidence (indicating, for example, feeding of babies in the nest) suggests that *Maiasaura* had a high degree of parental care. In addition, colonial nest sites, multiple co-occurring trackways, and accumulations of "bonebeds" indicate gregarious habits for these animals (and for many other ornithopods, such as *Iguanodon* from the Early Cretaceous of Europe, Asia, and North America). Many hadrosaurs (for example, *Parasaurolophus* and *Corythosaurus*: both from the Late Cretaceous of North America) had bizarre cranial crests that were probably used for display and species recognition. *See* ARCHOSAURIA; CRETACEOUS; JURASSIC; DINOSAURIA; SAURISCHIA.

Paul M. Barrett

Bibliography. P. J. Currie and K. Padian (eds.), *The Encyclopedia of Dinosaurs*, Academic Press, San Diego, 1997; J. O. Farlow and M. K. Brett-Surman (eds.) *The Complete Dinosaur*, Indiana University Press, Bloomington, 1997; D. B. Weishampel, P. Dodson, and H. Osmólska (eds.) *The Dinosauria*, 2d ed., University of California Press, Berkeley, 2004.

Orogeny

The process of mountain building. As traditionally used, the term orogeny refers to the development of long, mountainous belts on the continents called orogenic belts or orogens. These include the Appalachian and Cordilleran orogens of North America, the Andean orogen of western South America, the Caledonian orogen of northern Europe and eastern Greenland, and the Alpine-Himalayan orogen that stretches from western Europe to eastern China (**Fig. 1**). It is important to recognize that these systems represent only the most recent orogenic belts that retain the high-relief characteristic of mountainous regions. In fact, the continents can be viewed as a collage of ancient orogenic belts, most so deeply eroded that no trace of their original mountainous topography remains (**Fig. 2**). By comparing characteristic rock assemblages from more recent orogens with their deeply eroded counterparts, geologists surmise that the processes responsible for mountain building today extended back through most (if not all) of geologic time and played a major role in the growth of the continents. *See* CONTINENTS, EVOLUTION OF.

The construction of mountain belts is best understood in the context of plate tectonics theory. Earth's lithosphere is currently fragmented into at least a dozen, more or less rigid plates that are separated by three kinds of boundaries: convergent, divergent, and transform. Plates move away from each other at divergent boundaries. On the continents, these boundaries are marked by rift systems such as those in East Africa; in the ocean basins, they correspond to spreading centers, submarine mountain chains (such as the Mid-Atlantic Ridge) where new oceanic crust is produced to fill the gap left behind by plate divergence. Transform boundaries are zones where plates slide past one another; a familiar example on land is the San Andreas fault system. Orogenic belts form at convergent boundaries, where lithosphere plates collide. *See* FAULT AND FAULT STRUCTURES; MID-OCEANIC RIDGE; PLATE TECTONICS; RIFT VALLEY; TRANSFORM FAULT.

Convergent plate boundaries. There are two basic kinds of convergent plate boundaries, leading to the development of two end-member classes of orogenic belts. Oceanic subduction boundaries are those at which oceanic lithosphere is thrust (subducted) beneath either continental or oceanic lithosphere. The process of subduction leads to partial melting near the plate boundary at depth, which is manifested by volcanic and intrusive igneous activity in the overriding plate. Where the overriding plate consists of oceanic lithosphere, the result is an intraoceanic island arc, such as the Japanese islands. Where the overriding plate is continental, a continental arc is formed. The Andes of western South America is an example. *See* MARINE GEOLOGY; OCEANIC ISLANDS; SUBDUCTION ZONES.

The second kind of convergent plate boundary forms when an ocean basin between two continental masses has been completely consumed at an oceanic subduction boundary and the continents collide. Continent collisional orogeny has resulted in some of the most dramatic mountain ranges on Earth. A good example is the Himalayan orogen, which began forming roughly 50 million years ago when India collided with the Asian continent. Because the destruction of oceanic lithosphere at subduction boundaries is a prerequisite for continental collision, continent collisional orogens contain deformational features and rock associations developed during arc formation as well as those produced by continental collision, and are thus characterized by complex internal structure.

Compressive forces at convergent plate boundaries shorten and thicken the crust, and one important characteristic of orogens is the presence of unusually thick (up to 40–60 mi or 70–80 km) continental crust. Because continental crust is more buoyant than the underlying mantle, regions of thick

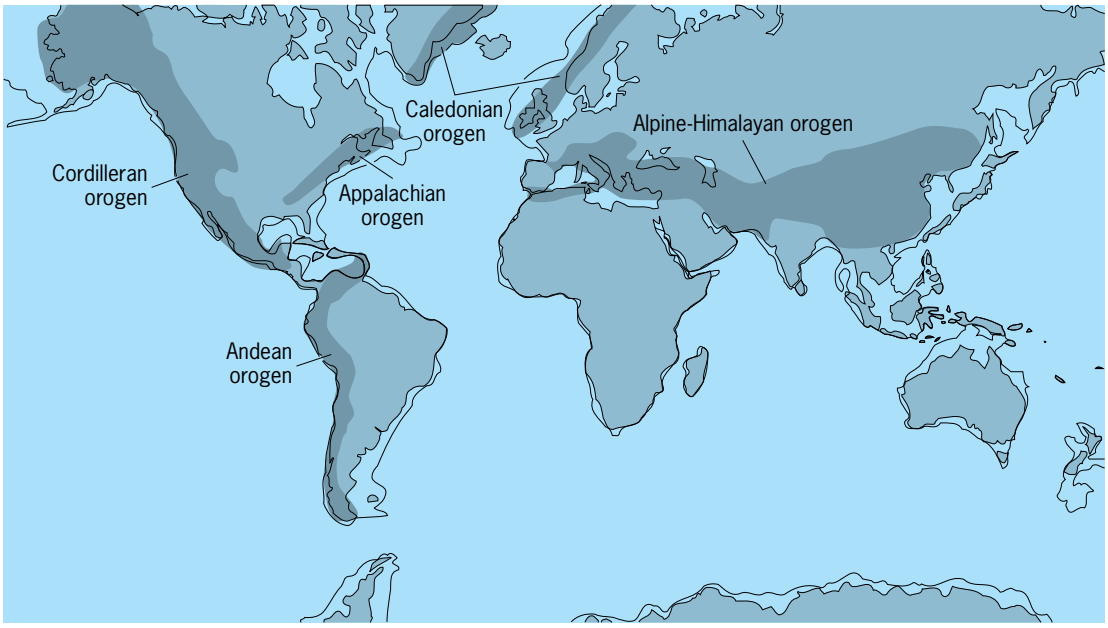

Fig. 1. Distribution of major orogenic belts. (*After E. M. Moores and R. J. Twiss, Tectonics, W. H. Freeman, 1995*)

crust are marked by high elevations of the Earth's surface and, hence, mountainous terrain. Once a mountain range has developed, its high elevation can be maintained by two mechanisms in addition to its own buoyancy. First, compressional forces related to continuing convergence across the plate boundary serve to buttress the orogen against collapse. Second, some mountain ranges are developed on thick, cold, and therefore strong continental lithosphere, and the lithosphere serves as a rigid support. However, such processes work against the gravitational forces that are constantly acting to destroy mountain ranges through erosion and normal faulting. All mountain ranges ultimately succumb to gravity; the resulting denudation exposes deep levels of orogenic belts and provides geologists with important insights regarding the internal architecture. *See* ASTHENOSPHERE; EARTH CRUST; LITHOSPHERE.

Continent collisional orogen. The tectonics of the Himalayas can serve as an example of the basic anatomy of a continent collisional orogen (**Fig. 3**). Prior to collision between India and Asia, the two continental masses were separated by an ocean referred to as Tethys. Most Himalayan geologists are convinced that the southern margin of Asia in late Mesozoic time was an oceanic subduction boundary where

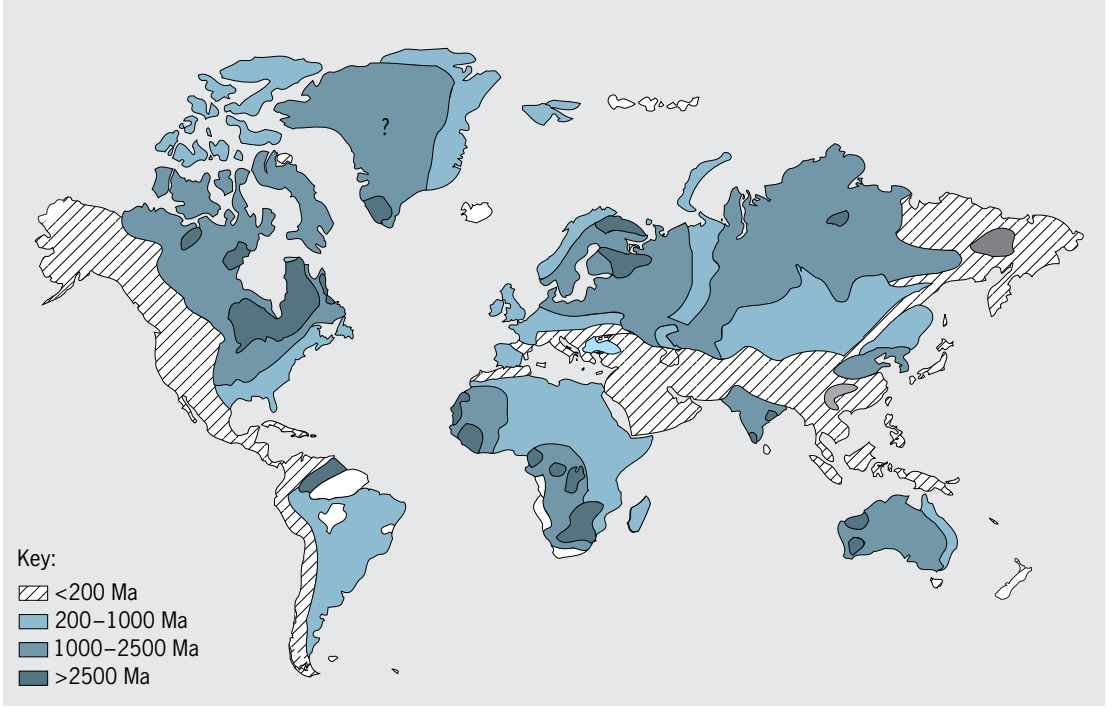

Fig. 2. Distribution of orogenic belts, by time of major orogenic distribution, in the continents. (*After B. C. Burchfiel, The continental crust, in E. M. Moores, ed., Shaping the Earth: Tectonics of Continents and Oceans, W. H. Freeman, 1990*)

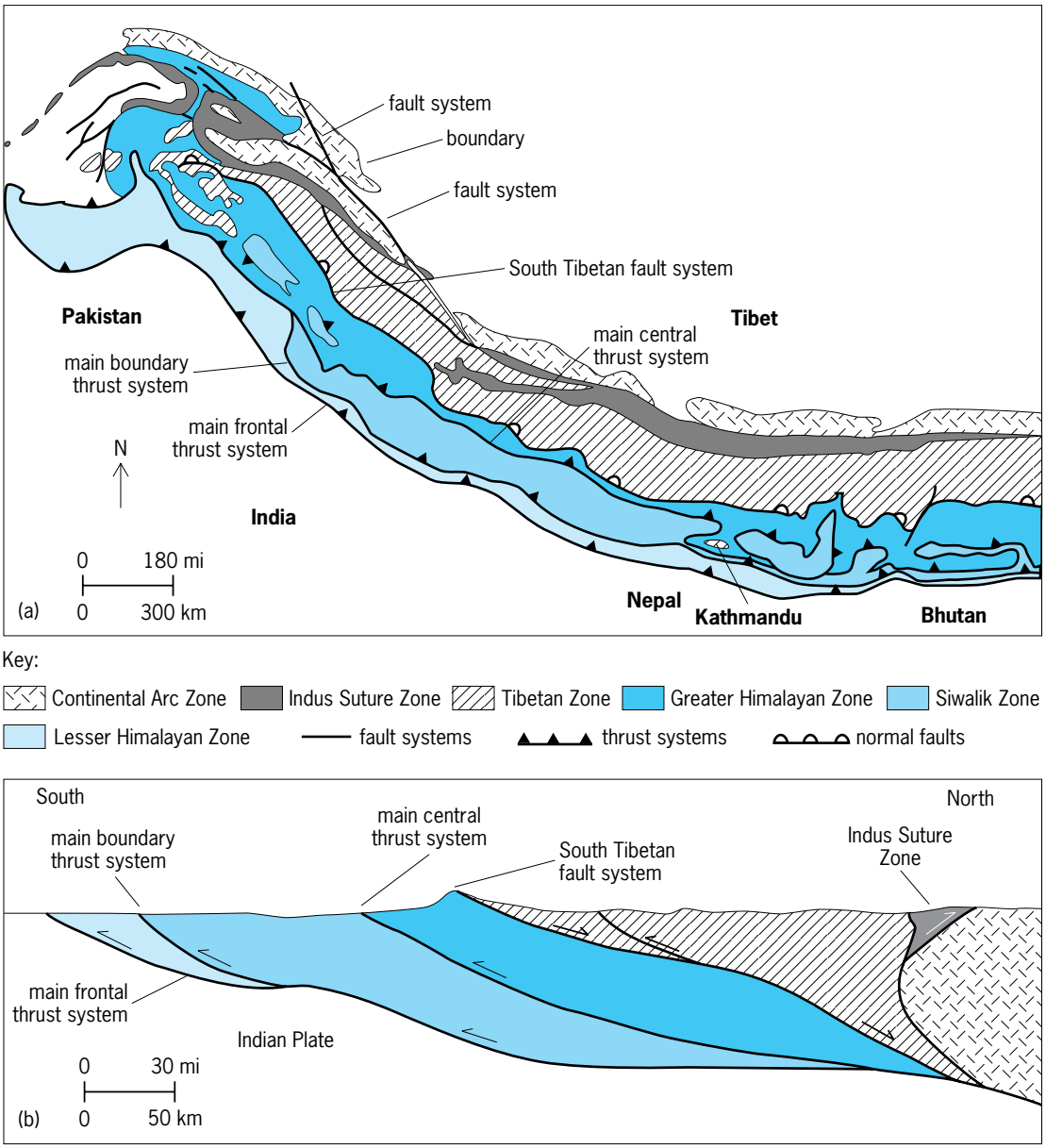

Fig. 3. Himalayan orogen. (*a*) Simplified tectonic map. (*b*) Generalized cross section. Fault systems are labeled at their surface exposure. Half-arrows indicate directions of movement on the fault systems.

Tethyan oceanic lithosphere was thrusting beneath Asia. The principal evidence for this interpretation comes from the existence of a continental arc of appropriate age that was built on Asian lithosphere just north of the Himalayan mountain range (Continental Arc Zone). India and Asia collided when all the intervening oceanic lithosphere had been destroyed at the subduction boundary. The collision zone at the surface of the Earth is marked by remnants of Tethyan oceanic crust (ophiolites) and related rocks exposed in the Indus Suture Zone. All tectonic zones south of the suture are part of the Indian continental lithosphere that was folded and imbricated (overlapped at the margins) during collision. The Tibetan Zone includes Paleozoic to early Tertiary sedimentary rocks originally deposited along the northern margin of the Indian continent. The Greater Himalayan Zone consists of crustal rocks that were deformed, metamorphosed, and partially melted at high temperatures deep within the crust as a consequence of Himalayan orogeny; these rocks form the principal substrate for the high mountains of the Himalayas, including Mount Everest. Farther south, the Lesser Himalayan Zone contains Precambrian to Paleozoic sedimentary and metamorphic rocks that have been deformed at lower temperatures and pressures. As collisional orogeny progressed through Tertiary time, material eroded from the rising Himalayan ranges was transported southward by a variety of ancient river systems and deposited at the southern margin of the orogen to form the Siwalik Zone. These rocks were themselves deformed during the most recent stages of Himalayan orogeny as the locus of surface deformation propagated southward toward the Indian subcontinent. *See* GEOLOGIC TIME SCALE.

A simplified geologic cross section through the orogen reveals that all of these tectonic zones are separated by major fault systems (Fig. 3*b*). Most of

these are dominated by thrust faults that accommodated shortening of the continental crust during collision. Geologists have discovered that the boundary between the Tibetan Zone and the Greater Himalayan Zone is marked by normal faults of the South Tibetan detachment system. Unlike thrust faults, normal faults thin and extend the crust, and identification of them in a continent-continent collisional setting like the Himalayan orogen was surprising at first. Many Himalayan geologists have come to believe that the South Tibetan detachment system developed when the range grew high enough so that the buttressing effects of plate convergence were no longer capable of supporting the weight and the orogen began to collapse, spreading out the load laterally over a larger region.

Other effects. Although mountain belts are the most conspicuous products of orogeny, the effects of plate convergence—especially continent-continent collision—can be found hundreds or even thousands of miles away from the orogenic belt. For example, the Tibetan Plateau, which lies just north of the Himalayan orogen, has an average elevation of about 15,000 ft (5000 m) over an area about the size of France. Development of the plateau was undoubtedly related to India-Asia collision, but the exact age and mechanism of plateau uplift remain controversial. Large-displacement strike-slip faults (where one block of continental crust slides laterally past another) are common in central and southeast Asia, and some geologists hypothesize that they extend from the surface down to at least the base of the crust, segmenting this part of Asia into a series rigid blocks that have extruded eastward as a consequence of the collision. Other geologists, while recognizing the significance of these faults for upper crustal deformation, suggest that the upper and lower crust of Tibet are detached from one another and that the eastward ductile flow of Tibetan lower crust has been an equally important factor in plateau growth.

Topographic expression. Earthquake activity in Bhutan, northern India, Nepal, northern Pakistan, and southern Tibet demonstrates that the Himalayan orogen continues to evolve, and this is principally why the Himalayas include such high mountains. As convergence across an orogen ends, erosion and structural denudation begin to smooth out the topographic gradient produced by orogeny. Rugged and high ranges such as the Himalayas are young and still active; lower ranges with subdued topography, such as the Appalachians, are old and inactive. Some of the oldest orogenic belts on Earth, such as those exposed in northern Canada, have no remaining topographic expression and are identified solely on the basis of geologic mapping of characteristic orogenic zones such as those discussed above for the Himalayas. Mountain ranges such as the Himalayas were formed and eventually eroded away throughout much of Earth's history; it is the preservation of ancient orogenic belts that provides some of the most reliable evidence that plate tectonic processes have been operative on Earth for at least 2 billion years.

Ancient geological processes. One of the most important axioms in the earth sciences is that the present is the key to the past (law of uniformitarianism). Earth is a dynamic planet; all oceanic lithosphere that existed prior to about 200 million years ago has been recycled into the mantle. The ocean basins thus contain no record of the first 96% of Earth history, and geologists must rely on the continents for this information. The careful study of modern orogenic belts provides a guide to understanding orogenic processes in ancient orogens, and permits use of the continents as a record of the interactions between lithospheric plates that have long since been destroyed completely at convergent plate boundaries.

Many of the basic geologic features of modern mountain systems such as the Himalayas can be found in ancient orogens, and geologists have established generic names for the different parts of orogens to facilitate comparisons. The Foredeep contains sediments transported from the rising mountain chain; in the Himalayan orogen, it corresponds to the more southerly portions of the Siwalik Zone (Fig. 3). The Foreland Fold and Thrust Belt (northern Siwalik Zone and Lesser Himalayan Zone) is a complexly deformed sequence of sedimentary and crystalline rocks that were scraped off the upper portion of the down-going plate during collision. The Orogenic Core comprises the central part of an orogen; it may include large tracts of crust that experienced deformation at deep structural levels (Greater Himalayan Zone), rock sequences that were deformed at higher structural levels but are now juxtaposed with metamorphic rocks by movement on normal fault systems such as the South Tibetan detachment (Tibetan Zone), and a discontinuous belt of ophiolites and related materials marking the zone of plate convergence *sensu stricto* (Indus Suture Zone). Ophiolite belts can be viewed as the fossil remnants of oceans that no longer exist. The presence of ophiolite belts of various ages within the continents means that a map of the modern distribution of lithospheric plates provides only a recent snapshot of an ever-changing configuration of lithospheric plates on Earth: mountain ranges such as the Himalayas were formed and eventually eroded away throughout much of Earth's history. Exactly how far back in geologic time plate tectonic processes and their orogenic consequences can be extrapolated is a matter of debate among geologists, but the bulk of evidence suggests that mountain ranges such as the Himalayas provide useful analogs for even the oldest orogens on Earth. The preservation of ophiolite belts as old as 2 billion years in Precambrian orogenic belts provides direct evidence that plate tectonic processes have been operative on Earth for at least the past 50% of Earth history. *See* CORDILLERAN BELT; MOUNTAIN SYSTEMS; OPHIOLITE.

Kip Hodges

Bibliography. K. C. Condie, *Plate Tectonics*, 4th ed., 1997; E. M. Moores (ed.), *Shaping the Earth: Tectonics of Continents and Oceans*, 1990; E. M. Moores and R. J. Twiss, *Tectonics*, 1995; J. T. Wilson (ed.), *Continents Adrift and Continents Aground*, 1996.

Orpiment

A mineral having composition As_2S_3 and crystallizing in the monoclinic system. Crystals are small, tabular, and rarely distinct (see **illus.**); the mineral

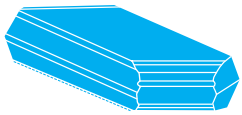

Orpiment crystal habit. (*After C. S. Hurlbut, Jr., Dana's Manual of Mineralogy, 17th ed., John Wiley and Sons, 1959*)

occurs more commonly in foliated or columnar masses. There is one perfect cleavage yielding flexible folia which distinguishes it from other minerals similar in appearance. The hardness is 1.5–2 (Mohs scale) and the specific gravity is 3.49. The luster is resinous and pearly on the cleavage surface; the color is lemon yellow. Orpiment is associated with realgar and stibnite in veins of lead, silver, and gold ores. It is found in Romania, Peru, Japan, and Russia. In the United States it occurs at Mercer, Utah; Manhattan, Nevada; and in deposits from geyser waters in Yellowstone National Park. *See* ARSENIC; REALGAR; STIBNITE. Cornelius S. Hurlbut, Jr.

Orthida

An extinct order of brachiopods that lived in shallow-shelf seas during the Paleozoic. It contains the oldest members of the Rhynchonelliformea and is the stem group for all other brachiopods in this subphylum.

Orthids are characterized by unequally biconvex valves containing distinct ribs and a strophic (straight) hinge line with associated well-developed interareas (**illus.** *a–d*). Openings in the interareas

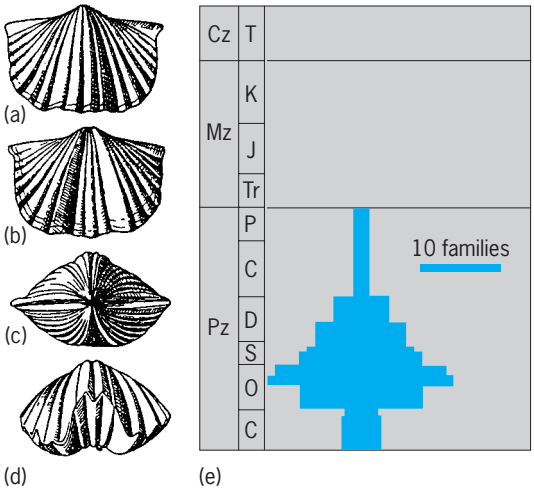

Orthida: (*a*) dorsal, (*b*) ventral, (*c*) posterior, and (*d*) anterior views of *Platystrophia*, Ordovician; (*e*) diversity history of orthid families. (*a–d after R. C. Moore, C. G. Lalicker, and A. G. Fischer, Invertebrate Fossils, McGraw-Hill, New York, 1952; e after J. J. Sepkoski, Jr., A Compendium of Fossil Marine Animal Families, 2d ed., Milwaukee Public Mus. Contrib. Biol. Geol., no. 83, 1992*)

provided space for a fleshy protrusion (pedicle) that was used for attachment to the substrate. One group, the billingsellids, possesses a calcareous structure (pseudodeltidium) that partially covers the pedicle opening; it is a key character suggesting that this orthid group was ancestral to the Strophomenida. All other orthids lack a pseudodeltidium and are the group from which all the rest of the rhynchonelliform brachiopods arose, either directly or indirectly.

Orthids were sessile, attached, epifaunal suspension feeders and were important members of marine epifaunal communities in the Paleozoic. They arose in the Early Cambrian but remained low in diversity and were only minor constituents of Cambrian marine epifaunal communities (illus. *e*). Beginning in the Early Ordovician, orthid brachiopods began to increase dramatically in diversity, taking part in the great Ordovician radiation of marine animal life. During this taxonomic diversification, orthids also became the dominant member of marine epifaunal communities and helped to establish rhynchonelliform brachiopods as one of the main components of the Paleozoic evolutionary fauna. Orthids suffered high extinction in the Late Ordovician mass extinction and never returned to their earlier Ordovician diversity levels. Their importance in marine epifaunal communities also changed as a result of the Late Ordovician mass extinction. Their overall importance decreased mostly at the expense of other brachiopod orders that diversified, such as the strophomenids, pentamerids, and spiriferids. Orthid diversity decreased again as a result of the Late Devonian mass extinction, and orthids were finally extinguished in the late Permian mass extinction. *See* BRACHIOPODA; ARTICULATA (ECHINODERMATA). Mark E. Patzkowsky

Bibliography. D. A. T. Harper and R. Moran, Brachiopod life styles, *Geol. Today*, 13:235–238, 1997; J. R. Richardson, Brachiopods, *Sci. Amer.*, 255:100–106, 1986; M. J. S. Rudwick, *Living and Fossil Brachiopods*, Hutchinson, London, 1970; A. Williams et al., A supra-ordinal classification of the Brachiopoda, *Phil. Trans. Roy. Soc. London B*, 351:1171–1193, 1996.

Orthoclase

Potassium feldspar (Or = $KAlSi_3O_8$) that usually contains up to 30 mole % albite (Ab = $NaAlSi_3O_8$) in solid solution. Its hardness is 6; specific gravity, 2.57–2.5, depending on Ab content; mean refractive index, 1.52; color, white to dull pink or orange-brown; good (001) and (010) cleavages at 90° to one another. Some orthoclases may be intergrown with relatively pure albite which exsolved during cooling from a high temperature in pegmatites, granites, or granodiorites. This usually is ordered low albite, but in rare cases it may show some degree of Al,Si disorder, requiring it to be classified as analbite or high albite. If exsolution is detectable by eye, the Or-Ab composite mineral is called perthite; if microscopic examination is required to distinguish the phases, it is called microperthite; and if exsolution is detectable

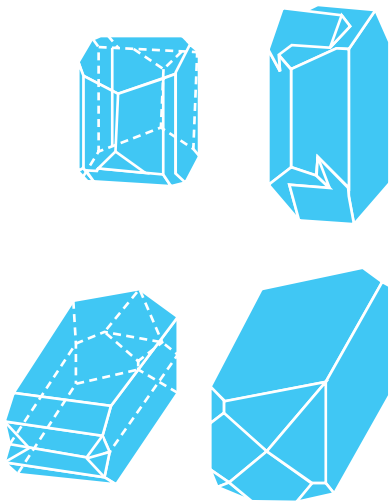

Orthoclase crystal habits. (*After C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*)

only by x-ray diffraction or electron optical methods, it is called cryptoperthite.

Orthoclase is optically monoclinic, ostensibly with C2/m space group and partially ordered Al,Si distribution in the two tetrahedral sites of the feldspar structure. In fact, electron microscopic examinations show most orthoclases to have complex, strained structures that might be considered to represent an intermediate stage in the sluggish Al,Si ordering process between monoclinic sanidine (with its highly disordered Al,Si distribution) and triclinic low microcline (with its nearly completely ordered Al,Si distribution). It has even been described as having a unit-cell-scale, pseudotwinned microstructure. Thus its structure averaged over hundreds of nanometers may be monoclinic, but its true symmetry is triclinic (see **illus.**). *See* ALBITE; CRYSTAL STRUCTURE; FELDSPAR; IGNEOUS ROCKS; MICROCLINE; PERTHITE; SOLID SOLUTION.

Paul H. Ribbe

Orthogonal polynomials

A special case of orthogonal functions that arise in many physical problems (often as the solutions of differential equations), in the study of distribution functions, and in certain other situations where one approximates fairly general functions by polynomials. *See* PROBABILITY.

Each set of orthogonal polynomials is defined with respect to a particular averaging procedure. The average value of a suitable function f is denoted by $E\{f\}$. Examples are shown in Eqs. (1)–(4).

$$E\{f\} = \frac{1}{2}\int_{-1}^{1} f(x)\,dx \tag{1}$$

$$E\{f\} = \frac{\displaystyle\int_{-1}^{1} f(x)\,(1-x)^{\alpha}(1+x)^{\beta}\,dx}{\displaystyle\int_{-1}^{1} (1-x)^{\alpha}(1+x)^{\beta}\,dx} \tag{2}$$

$$E\{f\} = \int_{0}^{\infty} f(x)e^{-x}\,dx \tag{3}$$

$$E\{f\} = (2\pi)^{-1/2}\int_{-\infty}^{\infty} f(x)e^{-x^2}\,dx \tag{4}$$

In general an averaging procedure shown in Eq. (5), a Stieltjes integral, where σ is a distribution

$$E\{f\} = \int_{-\infty}^{\infty} f(x)\,d\sigma(x) \tag{5}$$

function, that is, an increasing function with $\sigma(-\infty) = 0$ and $\sigma(+\infty) = 1$. In the above examples σ has the form

$$\sigma(x) = b^{-1}\int_{-\infty}^{x} \omega(y)\,dy$$

where ω is a nonnegative weight function and

$$b = \int_{-\infty}^{\infty} \omega(y)\,dy$$

Consideration will be given only to averaging procedures for which all the moments

$$\mu_n = E\{x^n\} = \int_{\infty}^{\infty} x^n\,d\sigma(x)$$

exist and for which $E\{|P|\} > 0$ for every polynomial P.

Orthogonal functions. Two functions f and g are said to be orthogonal with respect to a given averaging procedure if $E\{f\bar{g}\} = 0$ where the bar denotes complex conjugation. By the system of orthogonal polynomials associated with the averaging procedure is meant a sequence $P_0, P_1, P_2, \ldots$ of polynomials P_n having exact degree n, which are mutually orthogonal, that is, $E\{P_m\bar{P}_n\} = 0$ for $m \neq n$. This last condition is equivalent to the statement that each P_n is orthogonal to all polynomials of degree less than n. Thus P_n has the form $P_n(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ where $a_n \neq 0$ and is subject to the n conditions $E\{x^kP_n\} = 0$ for $k = 0, 1, \ldots, n-1$. This gives n linear equations in the $n+1$ coefficients of P_n, leaving one more condition, called a normalization, to be imposed. The method of normalization differs in different references. Orthogonal polynomials arising from the average of Eq. (1), Legendre polynomials, satisfy Legendre's differential equation. With the normalization $P_n(1) = 1$ the first few Legendre polynomials are $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = 3/2x^2 - 1/2$, $P_3(x) = 5/2x^3 - 3/2x$. The average in Eq. (1) is the special case of Eq. (2) with $\alpha = \beta = 0$; the orthogonal polynomials corresponding to averages of Eq. (2) are called Jacobi polynomials; those associated with Eq. (3), Laguerre polynomials; with Eq. (4), Hermite polynomials.

The proper setting for the study of expansions in terms of orthogonal polynomials is the Hilbert space H of functions f such that $E\{|f|^2\}$ exists and is finite. The inner product is $(f,\bar{g}) = E\{f\bar{g}\}$. In analogy with the procedure for Fourier series one can write down a formal expansion, relation (6), where the

$$f(x) \sim \sum_{n=0}^{\infty} c_nP_n(x) \tag{6}$$

coefficients are given by Eq. (7). The Nth partial sum

$$c_n = \frac{E\{f\bar{P}_n\}}{E\{|P_n|^2\}} \tag{7}$$

of the series shown in Eq. (8) has the property that

$$s_N(x) = \sum_0^N c_n P_n(x) \tag{8}$$

among all polynomials p of degree not exceeding N, the minimum of the quadratic deviation $E\{|f-p|^2\}$ is achieved uniquely by $p = s_N$. If the only function f in H with the property that $E\{x^k f\} = 0$ for every k is the zero function, one says that the polynomials are "complete" in H. In this case the coefficients in Eq. (7) uniquely determine the function f, and the properties of the series in Eq. (6) are quite analogous to the properties of Fourier series of functions in L^2. The polynomials are always complete when the average is taken over a finite interval, but in general some extra assumption is required. The divergence of the series $\Sigma\mu_{2n}{}^{-1/2n}$ is a sufficient condition for the completeness of the polynomials. (It is fulfilled in each of the examples cited.) *See* FOURIER SERIES AND TRANSFORMS.

The orthogonality property entails certain algebraic properties for the polynomials. For example, the zeros of P_n are all distinct, they lie in the interior of the interval over which the average is taken, and they separate the zeroes of P_{n-1}. Let $X_1{}^{(n)}, \ldots, X_n{}^{(n)}$ be the zeros of P_n. One can find constants $b^1{}_{(n)}, \ldots, b^1{}_{(n)}$ such that $Q_n\{1\} = 1$, $Q_n\{P_k\} = 0$ for $0 < k < n$, where

$$Q_n\{f\} = \sum_{j=1}^{n} b_j^{(n)} f\left(X_j^{(n)}\right)$$

In the case of an average over a finite interval,

$$\lim_{n\to\infty} Q_n\{f\} = E\{f\}$$

for every continuous f. This is of interest in approximate integration, because the integral $E\{f\}$ is approximated by an expression $Q_n\{f\}$ which depends only on the values of f at n points and, what is remarkable, $Q_n\{f\} = E\{f\}$ whenever f is a polynomial of degree $\geqq 2n-1$, whereas one would ordinarily expect an n-point approximation to be exact only for polynomials of degree $\geqq n$.

Ultraspherical polynomials. There exists a theory of orthogonal polynomials in several variables. The most important applications involve averages over spheres in m dimensions. Complete sets of orthogonal polynomials may be chosen among the homogeneous, harmonic polynomials. A polynomial $P(x_1, \ldots, x_m)$ is homogeneous of degree n if $P(\lambda x_1, \ldots, \lambda x_m) = \lambda^n P(x_1, \ldots, x_m)$ for each λ; it is harmonic if it satisfies Laplace's differential equation. Let P be such a polynomial with the property that $P(1, 0, \ldots, 0) \neq 0$. Consider $P(x, y_1, \ldots, y_{m-1})$ as a polynomial in the $m-1$ variables $y_1, \ldots, y_{m-1}$ and take the average over a sphere centered at the origin in $m-1$ dimensions. The result is a polynomial $P_n(x)$ of degree n. The orthogonality over the sphere in m dimensions translates itself into orthogonality on the interval $[-1, 1]$ with the weight function $\omega(x) = (1-x^2)^{(n-3)/2}$. For fixed m, the polynomials obtained this way are the ultraspherical polynomials, special cases of the Jacobi polynomials with $\alpha = \beta = (m-3)/2$. Where $m = 3$ corresponds to three-dimensional space, these are Legendre polynomials. *See* DIFFERENTIAL EQUATION; LAPLACE'S DIFFERENTIAL EQUATION; POLYNOMIAL SYSTEMS OF EQUATIONS; RIEMANNIAN GEOMETRY.

Carl S. Herz

Bibliography. R. Askey, *Orthogonal Polynomials and Special Functions*, 1975; T. S. Chihara, *An Introduction to Orthogonal Polynomials*, 1978; H. F. Davis, *Fourier Series and Orthogonal Functions*, 1963, reprint 1989; P. G. Nevai, *Orthogonal Polynomials*, reprint 1991; J. A. Shohat and J. D. Tamarkin, *The Problem of Moments*, Math. Surv. no. 1, 1943, reprint 1983; G. Szegö, *Orthogonal Polynomials*, 1978, reprint 1991.

Orthonectida

An order of Mesozoa. The orthonectids parasitize various marine invertebrates as multinucleate plasmodia, sometimes causing considerable damage to the host. The plasmodia multiply by fragmentation. Eventually they give rise asexually, by polyembryony, to sexual males and females. Commonly only one sex arises from a given plasmodium.

These sexually mature forms escape as minute ciliated organisms. Structurally they are composed of a single layer of ciliated epithelial cells surrounding an inner mass of sex cells (see **illus.**). The ciliated cells are disposed in rings around the body. Those at the anterior end form the anterior cone, on which the cilia are directed forward, while on the rest of the

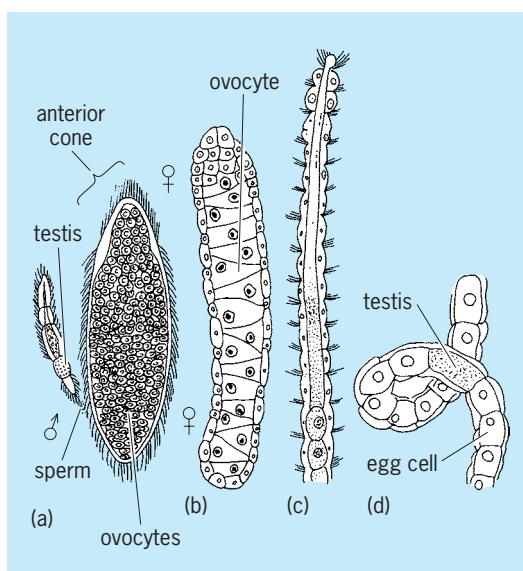

Orthonectids. (*a*) *Rhopalura ophiocomae*, male discharging sperm near genital pore of female. (*b*) *R. metschnikovi* germ cells oriented in a double row. (*c*) *Stoechartrum giardia*, anterior end, and (*d*) part of trunk showing one egg cell to each apparent segment.

body they point backward. Males are smaller than the corresponding females. A few species, however, are hermaphroditic.

After insemination the eggs develop in the female and form ciliated larvae. When they are liberated, these larvae invade new individuals of their host and then disaggregate, liberating germinal cells which give rise to new plasmodia. *See* MESOZOA; REPRODUCTION (ANIMAL).

Bayard H. McConnaughey

Orthoptera

An order that includes over 20,000 species of terrestrial insects, including most of the "singing" insects, some of the world's largest insects, and some well-known pests. Most species of Orthoptera (from "orthos," meaning straight, and "pteron," meaning wing) have enlarged hind legs adapted for jumping. These include grasshoppers and locusts (in the suborder Caelifera, a mainly diurnal group); and the crickets, katydids (bush-crickets), New Zealand weta, and allied families (suborder Ensifera, which comprises mainly nocturnal species). Orthoptera share with other orthopteroid insects, such as mantids and stick insects (now in separate orders Mantodea and Phasmatodea), gradual metamorphosis, chewing mouthparts, and two pairs of wings, the anterior pair of which is usually thickened and leathery and covers the fanwise folded second pair. Wings are reduced or absent in many species. Characters that define the Orthoptera as a natural group (the inclusive set of all species stemming from a common ancestor) are the jumping hind legs, small and well-separated hind coxae (basal leg segments), a pronotum with large lateral lobes, and important molecular (genetic) characters. Food habits range from omnivorous to strictly carnivorous or herbivorous. Habitats are nearly all terrestrial, including arctic-alpine tundra and tropical areas with aquatic floating plants. Female orthopterans usually lay their eggs into soil or plant material. There are no parasitic species, but a few crickets live as cleptoparasitic "guests" in ant nests.

Stridulation and mating. Males of many species are outstandingly noisy or musical. Their songs typically function in obtaining mates. In some species females move to the singing male, and in others a female answering song is involved in pair formation. Males may interact with bouts of singing until one or the other moves away. Song frequencies can range from the audible to the very high ultrasonic. Song patterns typically consist of a complex species-specific series of clicks, buzzes, or musical chirps produced by stridulation, the rubbing of a movable bowlike structure over a precisely arranged series of pegs or ridges. Grasshoppers (Acridoidea) stridulate by rubbing the hind femora against the outer wings, whereas crickets (Gryllidae), katydids (Tettigoniidae), and humped-winged crickets (Haglidae) rapidly rub the forewings together. Some wingless species stridulate with the overlapping edges of abdominal segments or the mandibles. There are at least 20 different modifications of surfaces for acoustical communication, including fanlike snapping of brightly colored hindwings in flight (some grasshoppers).

Hearing organs are similarly diverse in form and location. In the grasshoppers, the first abdominal spiracle is modified as a tympanum, whereas the crickets, katydids, and haglids have small tympanic membranes and acoustic receptor cells set into the front tibiae which are connected via the tracheal system to a specialized thoracic spiracle. There are many species of grasshoppers and crickets, however, that lack specialized sound-producing and -receiving organs. Signal communication in these species may be tactile, visual, or olfactory. *See* ANIMAL COMMUNICATION; PHONORECEPTION.

Sound production and later stages of mating behavior have been the focus of much behavioral research. In particular, male orthopterans are known for their nuptial gifts, which are consumed by females. Examples are specialized expendable wings and leg spines, edible substances attached to sperm packages (spermatophores), and dorsal gland secretions of many katydids and crickets, and male glandular secretions absorbed in the female reproductive, tract in some grasshoppers. Such donations by males can be costly and can mediate sex-role reversals when females compete to mate and acquire important nuptial offerings from males. *See* REPRODUCTIVE BEHAVIOR.

Suborder Ensifera. The first ensiferans appear in the fossil record in the Permian. Three superfamilies are commonly recognized in this natural group: Grylloidea (true crickets and allies), Tettigonioidea (katydids, haglids, and allies), and Gryllacridoidea (camel crickets and allies). Based on analyses of morphological and anatomical characters showing only two natural groups of ensiferans, some authors recognize only the first two superfamilies. Ensiferan antennae are usually longer than the body. Most species are nocturnal, and the night-time stridulation of katydids, crickets, and weta can be very loud, especially when males chorus. The ovipositor is long and sword-shaped, and often bores holes in soft wood, bark, living plant stems, or hard-packed soil to lay eggs singly.

Tettigoniidae (katydids and bush-crickets) and Gryllidae (true crickets) are the most widespread and diverse families. Many species in both families share with acridoids the completion of the life cycle (egg to adult) in one year, and the use of microhabitats in vegetation (many other ensiferan species have life cycles longer than a year and use burrows and crevices as microhabitats). Many katydids are cryptically shaped and colored, some so similar to leaves that decay spots and insect-chewed margins are mimicked. A major western North American pest is the flightless Mormon cricket (see **illustration**), dense groups of which can reach several kilometers in length and cover many miles, damaging crops along the way.

True crickets (Gryllidae) are ground-, tree-, or bush-dwelling, relatively chunky insects with a needle-shaped ovipositor. Members of the genera *Acheta, Teleogryllus*, and *Gryllus* are easily reared,

important laboratory animals, and a standard bait and food item for vertebrates. Although they can occur in large populations, they are not usually serious pests. However, some mole crickets (Gryllotalpidae) are serious pests of tropical crops.

Other ensiferan families include many fossil forms and some bizarre living forms, most of which are flightless. The family Anostostomatidae includes the giant weta of New Zealand (large specimens can be 30–40 grams, or 1.1–1.4 oz, in weight) and related, sexually dimorphic weta and king crickets of New Zealand, Australia, and South Africa, males of which fight using tusks or long mandibles. The North American Jerusalem crickets (Stenopelmatidae) are a diverse group of large-headed desert-inhabiting species. The camel and cave crickets, including cave weta of New Zealand (Rhaphidophoridae), are common in caves and damp forests of most continents. Most of these ensiferans are wingless. In contrast are the hump-winged crickets (Haglidae), which comprise a few species inhabiting the forests and high sage of the North American Rockies and the mountains of temperate Asia. One species, the only nocturnal singing insect at high altitudes, stridulates at the lowest ambient temperature known for any insect, −2°C (28°F).

Suborder Caelifera. This is a natural group that shares a number of characters, including an ovipositor with only four functional valves at the end of the abdomen. In many species, egg laying is accomplished by working the abdomen slowly into the soil and laying eggs in groups. In some grasshoppers the eggs are surrounded by a foamy matrix. Other characteristics that define Caelifera as a natural group are antennae composed of less than 30 segments (being typically less than half the length of the body) and certain genetic characters. There are eight superfamilies: Tridactyloidea (false and pygmy mole crickets and sand gropers); Tetrigoidea (pygmy grasshoppers and grouse locusts); Eumastacoidea (monkey grasshoppers and false stick insects); Pneumoroidea (flying gooseberries, desert longhorned grasshoppers, and razor-back bushhoppers); Pamphagoidea (rugged earth hoppers and true bushhoppers); Acridoidea (grasshoppers and locusts); and Trigonopterygoidea.

Most caeliferans are diurnal. The most familiar and diverse are short-horned grasshoppers (Acridoidea) which tend to be active in bright sunshine. Some species are cryptic, mimicking stones, debris, and grass stems. Although most species of grasshoppers are never abundant enough to become agricultural pests, some have little-understood cycles of abundance that sometimes phase together to become major crop-destroying events, especially in drier regions of North America, Africa, and Australia. In some years, plague locusts such as *Locusta migratoria*, *Schistocerca gregaria*, and others migrate hundreds of miles from an outbreak center to devastate thousands of acres of crops in the tropics and subtropics. Remarkable structural and behavioral changes occur from the solitary generation to the migratory generation. Several North American species

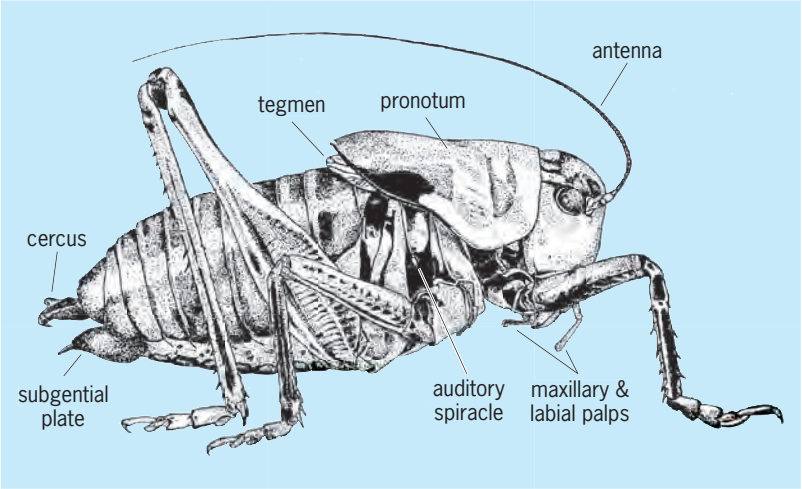

Male Mormon cricket, *Anabrus simplex*, showing some key morphological features of Orthoptera. (*Reprinted from D. T. Gwynne, Katydids and Bush-Crickets: Reproductive Behavior and Evolution of the Tettigoniidae. Copyright © 2001 by Cornell University. Used by permission of the publisher, Cornell University Press.*)

of *Melanoplus* have a latent ability to shift to a migratory phase, though natural outbreaks have occurred only historically, in a species which is now thought to be extinct. The plague locusts are large, hardy insects, extensively used for neurophysiological and general physiological studies.

Pygmy mole crickets and sand gropers (Tridactyloidea) have evolved specialized front legs for digging that resemble the front legs of true mole crickets. Some burrowing sand gropers can be pests of crop fields in Australia. Pneumoroids include the remarkable bladder grasshoppers of Africa, which lack a tympanum in the abdominal ear yet the very loud male stridulation can be detected by females at a distance of almost 2 km (1.2 mi). *See* INSECT PHYSIOLOGY; INSECTA. Darryl T. Gwynne; Robert B. Willey

Bibliography. J. L. Capinera, R. D. Scott, et al., *Field Guide to Grasshoppers, Katydids, and Crickets of the United States*, Cornell University Press, Ithaca, NY, 2004; R. F. Chapman and A. Joern (eds.), *The Biology of Grasshoppers*, Wiley, 1990; L. H. Field (ed.), *The Biology of Wetas, King Crickets and their Allies*, CABI International, Walling ford, 2001; S. K. Gangwere, M. C. Mulalirangen, and M. Mulalirangen (eds.), *The Bionomics of Grasshoppers, Katydids and Their Kin*, CAB International, Oxford, 1997; D. T. Gwynne, *Katydids and Bush-crickets: Reproductive Behavior and Evolution of the Tettigoniidae*, Cornell University Press, Ithaca, NY, 2001; J. A. Marshall and E. C. M. Haes, *Grasshoppers and Allied Insects of Great Britain and Ireland*, Harley Books, Colchester, 1988; D. C. Rentz, *Grasshopper Country: The Abundant Orthopteroid Insects of Australia*, University of New South Wales Press, Sydney, 1996.

Orthorhombic pyroxene

A group of minerals having the general chemical formula $XYSi_2O_6$, in which the Y site contains iron (Fe) or magnesium (Mg) and the X site contains Fe,

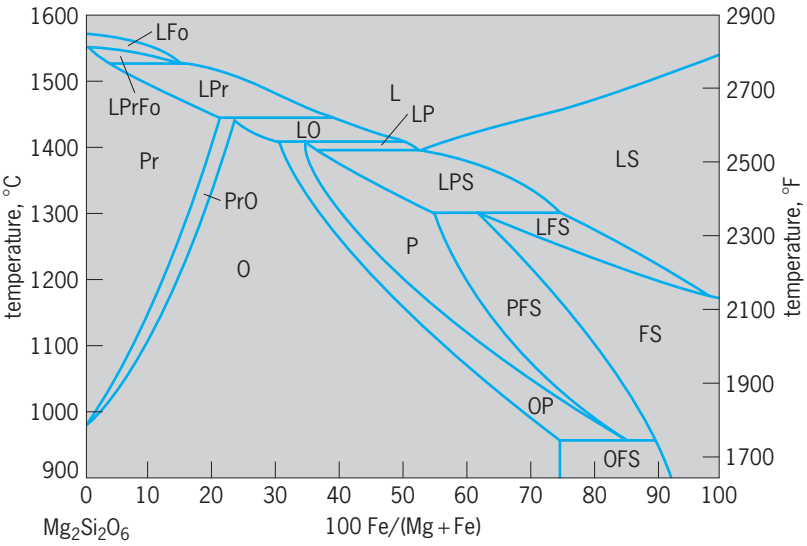

Fig. 1. Temperature-composition relations at low pressure for the system $Mg_2Si_2O_6$–$Fe_2Si_2O_6$. O = orthorhombic phase; Pr = protopyroxene phase; P = monoclinic phase; Fo = magnesian olivine; F = iron-rich olivine; S = silica; L = melt. (*After C. T. Prewitt, ed., Pyroxenes, Reviews in Mineralogy, vol. 7, Mineralogical Society of America, 1980*)

Mg, manganese (Mn), or a small amount of calcium (Ca; up to about 3%). They are characterized by orthorhombic crystal symmetry, which makes it impossible for significant amounts of larger cations such as Ca or sodium (Na) to enter the X position. The most important compositional variation in the orthopyroxene series is variation in the Fe-Mg ratio, leading to a solid solution series bounded by the end members enstatite ($Mg_2Si_2O_6$; abbreviated En) and ferrosilite ($Fe_2Si_2O_6$; Fs). Names that are used for intermediate members of the series are enstatite ($Fs_{0\text{-}50}$) and ferrosilite ($Fs_{50\text{-}100}$). More Fs-rich enstatite may be called ferroan enstatite, and more enstatite-rich ferrosilite may be called magnesian ferrosilite. In addition to the major elements noted above, minor components in orthopyroxene may include aluminum (Al), titanium (Ti), chromium (Cr), nickel (Ni), and ferric iron (Fe^{3+}), although only Al occurs in substantial amounts. Conditions of formation may be important in determining orthopyroxene compositions: for example, Al content is greater at higher temperatures, and the most Fe-rich compositions are stable only at very high pressures. Further, there is a limited solid solution of orthopyroxene toward calcic clinopyroxene, noted above, which increases with increasing temperature. *See* ENSTATITE; SOLID SOLUTION.

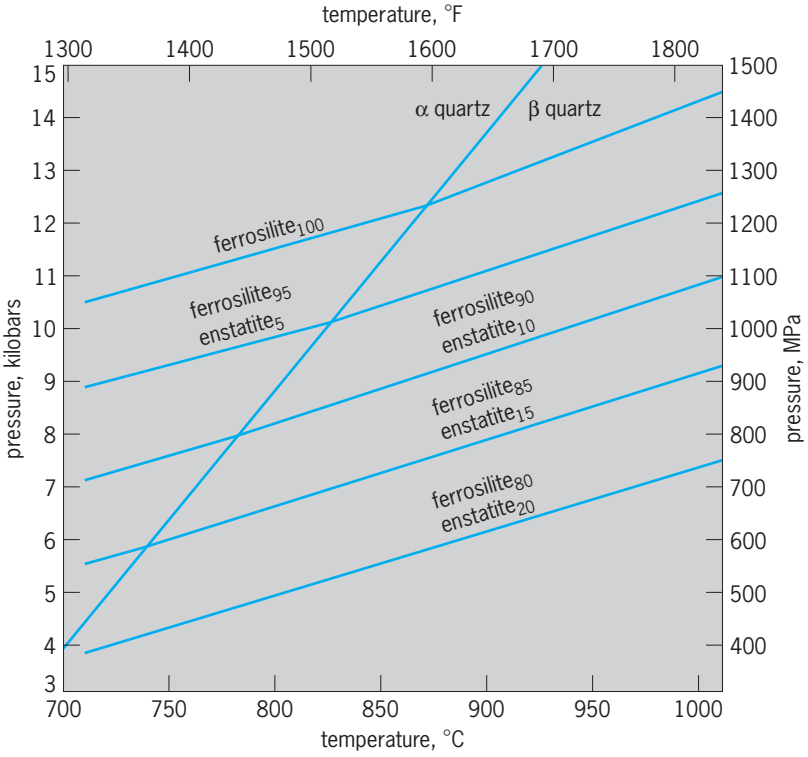

Fig. 2. Pressure-temperature relations of iron-rich orthopyroxenes. Contours show stability of ferrosilite with 0, 5, 10, 15, and 20 mol% of enstatite ($MgSiO_3$). (*After S. R. Bohlen and A. L. Boettcher, Experimental investigations and geological applications of orthopyroxene geobarometry, Amer. Mineralog., 66:951–964, 1981*)

Crystal structure. Orthopyroxenes crystallize in a primitive orthorhombic cell, space group Pbca (a high-temperature Mg-rich form called protoenstatite has slightly different symmetry and is in space group Pbcn). The *b* and *c* unit-cell dimensions are about the same as for monoclinic pyroxene (such as diopside), but the *a* dimension is approximately doubled. This suggests that the absence of large cations in orthopyroxene causes a slight structural shift so that the smallest repeated unit contains twice the number of atoms as in a monoclinic pyroxene unit cell. The crystal chemistry is otherwise similar to that of diopside, with single chains of (SiO_4) tetrahedra alternating in the structure with strips of octahedral sites that accommodate larger cations. As in monoclinic pyroxene, Al is small enough to enter both octahedral and tetrahedral sites. *See* CRYSTAL STRUCTURE; DIOPSIDE.

Properties. Many of the physical and optical properties of orthopyroxene are strongly dependent upon composition, and especially upon the Fe-Mg ratio. Magnesium-rich enstatite ($Fs_{0\text{-}20}$) is commonly pale brown in color; the color becomes much darker for more Fe-rich members of the series. The specific gravity and refractive index both increase markedly with increasing Fe content. Enstatite in the range of $Fs_{10\text{-}30}$ (formerly called bronzite) commonly contains a bronze iridescence that is due either to very fine exsolved lamellae of calcic monoclinic pyroxene that diffract light rays (similar to blue-green iridescence of labradorite plagioclase feldspar) or to very thin, translucent, oriented plates of ilmenite, an iron-titanium oxide. In hand specimens, orthopyroxene can be distinguished from amphibole by its characteristic 88° cleavage angles, and from augite by color—augite is typically green to black, while orthopyroxene is more commonly brown, especially on slightly weathered surfaces. In rock thin sections, orthopyroxene is usually recognized by its parallel extinction and blue-green to pink pleochroism. *See* PETROGRAPHY.

Experimental phase relations. A considerable body of experimental data exists on the pressure-temperature (*P-T*) relations of orthorhombic pyroxenes, both for crystallization from silicate

melts and in the subsolidus regions. The low-pressure thermal behavior of phases in the system $Mg_2Si_2O_6$–$Fe_2Si_2O_6$ is notably pseudobinary, as reflected by the wide variability of the silica content of the liquids with different Fe-Mg ratios, by the incongruent melting of enstatite to Mg-olivine plus siliceous melt, and by the subsolidus breakdown of ferrosilite to Fe-olivine plus silica (**Fig. 1**).

Although orthopyroxene more Fe-rich than Fs_{75} is not stable at low pressure (Fig. 1), higher pressures stabilize these pyroxenes. Pure $Fe_2Si_2O_6$ is stable above about 10 kilobars (100 megapascals), and the stability of orthopyroxene increases markedly with solid solution of small amounts of $Mg_2Si_2O_6$ (**Fig. 2**). In fact, accurate experimental calibration makes Fe-rich orthopyroxene very useful for geobarometry if the composition is known and the temperature can be independently determined.

Aluminum is essentially the only minor element whose effect on pyroxene phase relations has been studied. Investigation of the Al_2O_3 content of enstatite coexisting with forsterite and spinel has demonstrated that pyroxene alumina content increases with increased temperature, and is almost independent of pressure. Aluminous enstatites in spinel peridotites may therefore be very useful as geothermometers. *See* GEOLOGIC THERMOMETRY.

Occurrence. Orthopyroxene is a widespread constituent of both igneous and metamorphic rocks. It is common in peridotites of the Earth's upper mantle (found as nodules in kimberlites or alkali basalts), where it coexists with olivine, augite, and spinel or garnet. Orthopyroxene of intermediate Fe/Mg occurs in many basalts and gabbros and in meteorites, but is notably absent from most alkaline igneous rocks. Ferroan enstatite or magnesian ferrosilite is relatively common in intermediate volcanic and plutonic rocks such as andesite, dacite, and diorite, but rarer in more silicic rocks such as rhyolite and granite. Enstatite or ferroan enstatite is an essential constituent of the type of gabbro called norite, in which augite is subordinate to orthopyroxene in abundance. *See* LITHOSPHERE.

Orthopyroxene occurs in the highest concentration in ultramafic rocks, especially the peridotites and harzburgites of the large layered intrusions such as the Bushveld Complex of southern Africa or the Stillwater Complex of Montana. In these complexes enstatite crystallized from the magma early, along with olivine, and settled through the less dense magma to form cumulate rocks at the bottom of the intrusion. Layered complexes may contain hundreds to thousands of meters of these orthopyroxene-rich cumulate rocks. *See* IGNEOUS ROCKS; MAGMA.

Orthopyroxene of intermediate to high Fs content is a characteristic mineral in granulite-facies metamorphic rocks; granulite-facies terranes are the typical locales for application of geobarometry using compositions of Fs-rich orthopyroxene. In feldspathic rocks such as charnockite, orthopyroxene occurs with garnet, augite, hornblende, alkalic feldspars, and quartz; in mafic gneisses it coexists with augite, hornblende, garnet, and calcic plagioclase. In very high-grade metamorphic rocks, especially in low-pressure contact metamorphism, the dehydration of biotite commonly results in a coexistence of orthopyroxene and potassic feldspar. Magnesian orthopyroxene is a typical constituent of medium- to high-grade metamorphosed ultramafic rocks, in which it coexists with olivine, spinel, chlorite, and amphiboles. *See* FACIES (GEOLOGY); METAMORPHIC ROCKS; MINERALOGY; PYROXENE.

Robert J. Tracy

Bibliography. W. A. Deer, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, Vol. 2B, Double-Chain Silicates, 2d ed., 1997; W. A. Deer, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, vol. 2A: *Single Chain Silicates*, 1978; C. T. Prewitt (ed.), Pyroxenes, *Reviews in Mineralogy*, vol. 7, 1980.

Orthotrichales

An order of the true mosses (subclass Bryidae) consisting of five families and 23 genera. The plants grow in mats or tufts in relatively exposed places, on trunks of trees and on rock. They are rather freely branched and prostrate, or may be sparsely forked and erect-ascending; the habit is more or less pleurocarpous, but sporophytes may be produced at the ends of leading shoots or branches. The leaves are generally oblong and broadly pointed, with a strong midrib. The upper cells are short and often papillose, while the basal cells are longer and often pellucid. The capsules are generally immersed and often ribbed, and the peristome, if present, has a poor development of endostome. The calyptrae are nearly always mitrate and often hairy. The Grimmiales, likewise somewhat pleurocarpous, differ in having single peristomes and a lesser development of neck tissue below the spore-bearing portion of the capsule. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA; GRIMMIALES.

Howard Crum

Osage orange

The genus *Maclura* of the mulberry family, with one species, *M. pomifera*. This tree may attain a height of 60 ft (18 m) and has yellowish bark, milky sap, simple entire leaves, strong axillary thorns (see **illus.**), and

Branch and leaf of *Maclura pomifera*.

aggregate green fruit about the size and shape of an orange. It is planted for hedges and as an ornament, especially in the eastern United States where it is naturalized. The wood is used for fence posts and fuel and as a source of a yellow dye. It has also been used for archery bows, hence one of its common names, bowwood. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Oscillation

Any effect that varies in a back-and-forth or reciprocating manner. Examples of oscillation include the variations of pressure in a sound wave and the fluctuations in a mathematical function whose value repeatedly alternates above and below some mean value.

The term oscillation is for most purposes synonymous with vibration, although the latter sometimes implies primarily a mechanical motion. A device designed to reduce a person's weight by shaking part of the body is likely to be called a vibrator, whereas an electronic device that produces an electric current which reverses its direction periodically is usually called an oscillator. The alternating current and the associated electric and magnetic fields are referred to as electric (or electromagnetic) oscillations.

If a system is set into oscillation by some initial disturbance and then left alone, the effect is called a free oscillation. A forced oscillation is one in which the oscillation is in response to a steadily applied periodic disturbance.

Any oscillation that continually decreases in amplitude, usually because the oscillating system is sending out energy, is spoken of as a damped oscillation. An oscillation that maintains a steady amplitude, usually because of an outside source of energy, is undamped. *See* ANHARMONIC OSCILLATOR; DAMPING; FORCED OSCILLATION; HARMONIC OSCILLATOR; MECHANICAL VIBRATION; OSCILLATOR; VIBRATION.

Joseph M. Keller

Oscillator

An electronic circuit that generates a periodic output, often a sinusoid or a square wave. Oscillators have a wide range of applications in electronic circuits: they are used, for example, to produce the so-called clock signals that synchronize the internal operations of all computers; they produce and decode radio signals; they produce the scanning signals for television tubes; they keep time in electronic wristwatches; and they can be used to convert signals from transducers into a readily transmitted form.

Oscillators may be constructed in many ways, but they always contain certain types of elements. They need a power supply, a frequency-determining element or circuit, a positive-feedback circuit or device (to prevent a zero output), and a nonlinearity (to define the output-signal amplitude). Different choices for these elements give different oscillator circuits with different properties and applications.

Oscillators are broadly divided into relaxation and quasilinear classes. Relaxation oscillators use strong nonlinearities, such as switching elements, and their internal signals tend to have sharp edges and sudden changes in slope; often these signals are square waves, trapezoids, or triangle waves. The quasilinear oscillators, on the other hand, tend to contain smooth sinusoidal signals because they regulate amplitude with weak nonlinearities. The type of signal appearing internally does not always determine the application, since it is possible to convert between sine and square waves. Relaxation oscillators are often simpler to design and more flexible, while the nearly linear types dominate when precise control of frequency is important.

Relaxation Oscillators

Figure 1*a* shows a simple operational-amplifier based relaxation oscillator. This circuit can be understood in a number of ways (for example, as a negative-resistance circuit), but its operation can be followed by studying the signals at its nodes (Fig. 1*b*). The two resistors, labeled r, provide a positive-feedback path that forces the amplifier output to saturate at the largest possible (either positive or negative) output voltage. If v_+, for example, is initially slightly greater than v_-, then the amplifier action increases v_o, which in turn further increases v_+ through the two resistors labelled r. This loop continues to operate, increasing v_o until the operational amplifier saturates at some value V_{max}. [An operational amplifier ideally follows Eq. (1), where A_v is very large,

$$v_o = A_v(v_+ - v_-) \tag{1}$$

but is restricted to output levels $|v_o| \leq V_{max}$.] For the purposes of analyzing the circuit, the waveforms in Fig. 1*b* have been drawn with the assumption that this mechanism has already operated at time 0 and

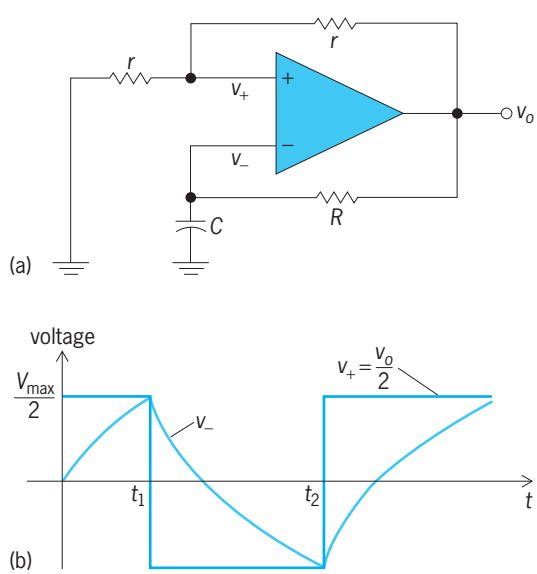

Fig. 1. Simple operational-amplifier relaxation oscillator. (*a*) Circuit diagram. (*b*) Waveforms.

that the initial charge on the capacitor is zero. *See* AMPLIFIER; OPERATIONAL AMPLIFIER.

Capacitor C will now slowly change from v_o through resistor R, toward V_{max}, according to Eq. (2). Up until time t_1, this process continues without any

$$v_- = V_{max}(1 - e^{-t/RC}) \qquad (2)$$

change in the amplifier's output because $v_+ > v_-$, and so $v_o = V_{max}$. At t_1, however, $v_+ = v_-$ and v_o will start to decrease. This causes v_+ to drop, and the positive-feedback action now drives the amplifier output negative until $v_o = -V_{max}$. Capacitor C now discharges exponentially toward the new output voltage until once again, at time t_2, $v_+ = v_-$, and the process starts again. The period of oscillation for this circuit is $2RC \ln 3$.

The basic elements of an oscillator that were mentioned above are all clearly visible in this circuit. Two direct-current power supplies are implicit in the diagram (the operational amplifier will not work without them), the RC circuit sets frequency, there is a resistive positive-feedback path that makes the mathematical possibility $v_o(t) = 0$ unstable, and the saturation behavior of the amplifier sets the amplitude of oscillation at the output to $\pm V_{max}$.

Practical oscillators. While the circuit of Fig. 1 is easy to understand and to experiment with, it does have some limitations because practical operational amplifiers produce a finite and variable slope in the portions of the waveform shown in Fig. 1 as vertical. The finite slope limits the useful frequencies available to the audio range, and the variability of the slope makes it impossible to obtain precise frequency control. **Figure 2***a* and *b* show block diagrams of two widely used commercial oscillators, with generic integrated circuit type numbers 555 and 566, that work on similar principles to the one described above.

The 555 (Fig. 2*a*) allows capacitor C to charge through resistor R until it reaches a threshold that causes the upper comparator to set a latch (memory circuit) that turns on the electronic switch S (actually a transistor), thus discharging C. When C has discharged below the threshold voltage of the lower comparator, the latch is reset (turned off) and the switch opened. The cycle then restarts. This circuit is often used by hobbyists and for applications not requiring high frequencies or great accuracy. *See* ELECTRONIC SWITCH; TRANSISTOR.

The 566 (Fig. 2*b*) uses a transistor current source supplying a current I to charge its capacitor, and then discharges it by switching the sign of the current to $-I$. The Schmitt trigger used to determine the appropriate times for turning the second source on and off uses positive feedback in much the same way as did the operational-amplifier circuit of Fig. 1. The most interesting feature of this circuit is that it can be used to make a voltage-controlled frequency, since the current I can easily be set by an external voltage. In comparison, the circuit of Fig. 1 can be tuned only by varying R or C. Voltage-controlled oscillators are important components of phase-locked loops,

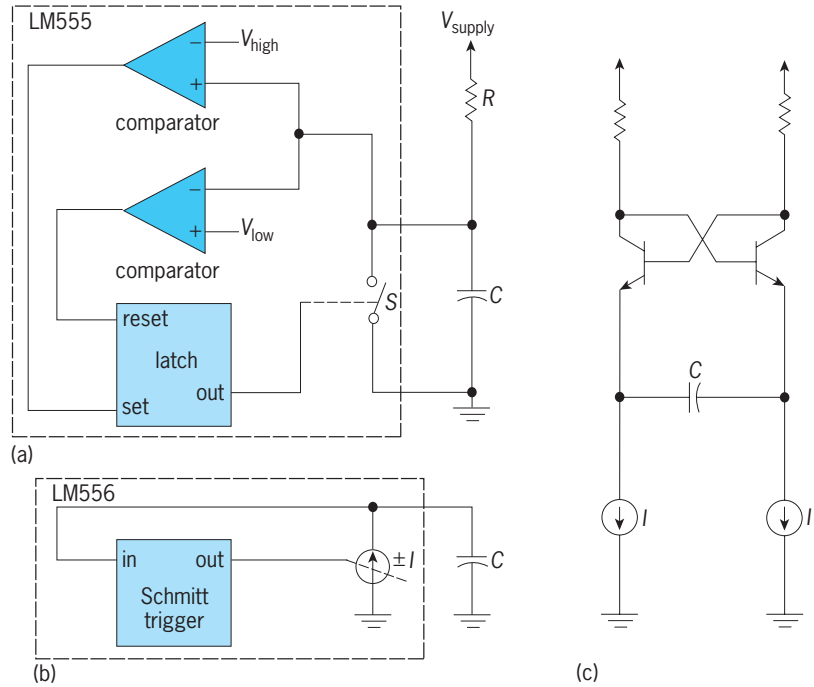

Fig. 2. Commercial relaxation oscillators. (*a*) LM555 timer. (*b*) LM566 voltage-controlled oscillator. (c) Emitter-coupled multivibrator.

which in turn are widely used in radio and data communications. *See* CURRENT SOURCES AND MIRRORS; PHASE-LOCKED LOOPS.

Figure 2*c* shows a simplified circuit for an emitter-coupled multivibrator, which is used as a variable-frequency oscillator in high-speed digital circuits. The cross coupling of each transistor's collector to the other's base is a positive-feedback path, while the charging of C by the adjustable current sources sets the oscillation frequency. *See* MULTIVIBRATOR.

Blocking oscillators. Relaxation oscillators that have a low duty cycle—that is, produce output pulses whose durations are a small fraction of the overall period—are sometimes called blocking oscillators because their operation is characterized by an "on" transient that "blocks" itself, followed by a recovery period. The circuit of **Fig. 3***a*, for example, can be seen to have a positive-feedback loop through the transistor and transformer: if v_o starts to fall, the base voltage v_b will rise (because of the po-

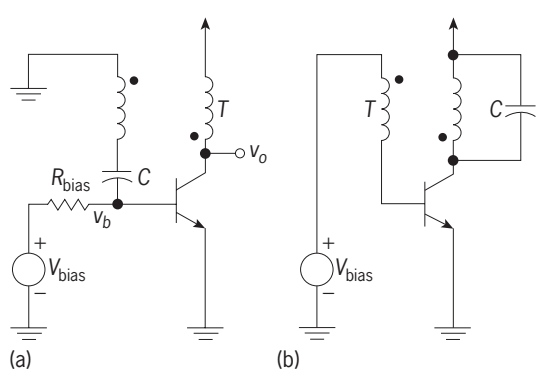

Fig. 3. Transformer-coupled oscillators. (*a*) Blocking oscillator. (*b*) Hartley oscillator.

larity of windings on the transformer), causing the transistor to turn on and so further decreasing υ_o. Thus the transistor will turn on hard, until it saturates and starts to draw a large amount of base current. Transformers cannot provide direct current, because they have imperfect coupling that manifests itself as inductance, and so eventually base current will begin to fall. This starts to turn the transistor off, and the process quickly regenerates through the positive-feedback path. The output can now be said to be blocked.

All of the base current consumed when the transistor was on went through capacitor C, which is therefore charged with a negative voltage on the transistor side. Thus, when the transformer voltage drops to zero, this entire negative capacitor voltage appears at the transistor's base and keeps it turned off. The capacitor must recharge slowly from V_{bias} through R_{bias} before conduction can begin again.

Figure 3*b*, discussed below, uses similar components but a smaller loop gain to make a sinusoidal oscillator.

Oscillators with digital gates. Inverters (digital circuits that invert a logic signal, so that a 0 at the input produces a 1 at the output, and vice versa) are essentially voltage amplifiers and can be used to make relaxation oscillators in a number of ways. A circuit related to that of Fig. 1 uses a loop of two inverters and a capacitor C to provide positive feedback, with a resistor R in parallel with one of the inverters to provide an RC charging time to set frequency. This circuit is commonly given as a simple example, but there are a number of problems with using it, such as that the input voltage to the first gate sometimes exceeds the specified limits for practical gates. *See* LOGIC CIRCUITS.

A more practical digital relaxation oscillator, called a ring oscillator, consists simply of a ring containing an odd number N (greater than 1) of inverters. It has no stable logical state (as can be seen by calling the state on any node x, then following the ring around through the inverters until coming back to the same node, which would have to be the inverse of x, $\overline{x}$, so that $x = \overline{x}$, a contradiction) and in practice (when the number of inverters is three or greater) has no stable state with the outputs at levels intermediate between the voltages representing 0 and 1 either. However, a practical inverter produces an output signal that is a delayed version of the inverse of the input and is given by Eq. (3), where x is the

$$x_o(t) = \overline{x_{\text{in}}(t - \tau)} \tag{3}$$

delay. Because of this delay the circuit solves the problem of the lack of stable states by oscillating. The output of each gate varies periodically, spending time $T/2$ in the "0" state and $T/2$ in the "1" state, where T is the period. A ring oscillator with a ring of N gates oscillates with a period $T = 2N\tau$, where τ is the delay of a single gate. The total time displacement of the output waveform of each gate with respect to that of the previous gate, due to the combination of inversion and delay, is given by Eq. (4). In theory there is another possible mode of

$$\frac{T}{2} + \tau = (N + 1)\tau = \frac{(N + 1)T}{2N} \tag{4}$$

oscillation, in which all states switch back and forth simultaneously between 0 and 1, but this mode is not stable and in practice quickly decays. A ring oscillator is commonly used as a way of measuring gate delays, and also when very fast oscillators with several outputs at finely spaced phases are required.

Sine-Wave Oscillators

Oscillators in the second major class have their oscillation frequency set by a linear circuit, and their amplitudes set by a weak nonlinearity.

Linear circuit. A simple example of a suitable linear circuit is a two-component loop consisting of an ideal inductor [whose voltage is given by Eq. (5), where i

$$v = L\frac{di}{dt} \tag{5}$$

is its current] and a capacitor [whose current is given by Eq. (6)], connected in parallel. These are said to

$$i = C\frac{dv}{dt} \tag{6}$$

be linear elements because, in a sense, output is directly proportional to input, for example, doubling the voltage υ across a capacitor also doubles $d\upsilon/dt$ and therefore doubles i. The overall differential equation for a capacitor-inductor loop can be written as Eq. (7). Mathematically this has solutions of the form of Eq. (8), where $\omega = 1/\sqrt{LC}$ [which means that the

$$i + LC\frac{d^2i}{dt^2} = 0 \tag{7}$$

$$i = A \sin(\omega t + \phi) \tag{8}$$

circuit oscillates at a frequency $1/(2\pi\sqrt{LC})$] and A and ϕ are undefined. They are undefined precisely because the elements in the circuit are linear and do not vary with time: any solution (possible behavior) to the equation can be scaled arbitrarily or time-shifted arbitrarily to give another. Practically, A and ϕ are determined by weak nonlinearities in a circuit. *See* DIFFERENTIAL EQUATION; LINEARITY.

Equation (7) is a good first approximation to the equation describing a pendulum, and so has a long history as an accurate timekeeper. Its value as an oscillator comes from Galileo's original observation that the frequency of oscillation ($\omega/2\pi$) is independent of the amplitude A. This contrasts sharply with the case of the relaxation oscillator, where any drift in the amplitude (resulting from a threshold shift in a comparator, for instance) can translate directly into a change of frequency. Equation (7) also fundamentally describes the operation of the quartz crystal that has replaced the pendulum as a timekeeper; the physical resonance of the crystal occurs at a time constant defined by its spring constant and its mass. *See* HARMONIC MOTION; HARMONIC OSCILLATOR; PENDULUM.

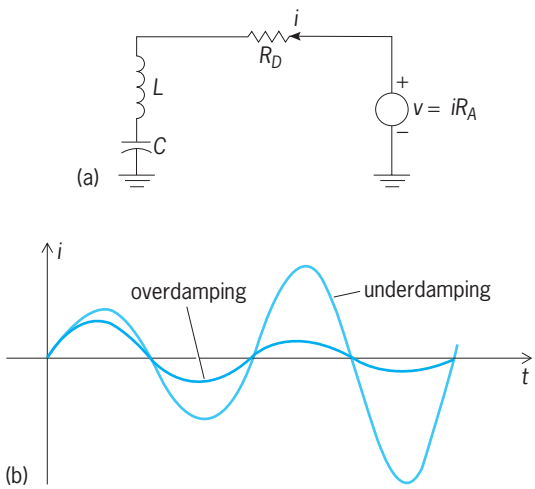

Fig. 4. Negative-resistance oscillator based on a tuned circuit. (*a*) *LC* tuned circuit and controlled-source negative-resistance simulation. (*b*) Output waveforms with overdamping and underdamping.

Amplifier element. Just as the swinging of a practical pendulum gradually dies away, so the oscillations in an *LC* circuit and a crystal die away. Clocks solve this problem with an escapement mechanism that gives the pendulum a small kick of additional energy on each cycle, and electronic oscillators use amplifiers for the same purpose. This mechanism is simultaneously used to set amplitude.

Figure 4*a* shows one implementation of this idea. The heart of the circuit is the resonator formed by the inductor L and capacitor C. This resonator also contains a damping resistance R_D, which might come, for example, from the finite resistance of the wires used to wind the inductor coil. The effect of R_D is to change the differential equation for a practical *LC* loop to Eq. (9). This in turn changes the response

$$i + RC\frac{di}{dt} + LC\frac{d^2i}{dt^2} = 0 \tag{9}$$

of the resonator to an exponentially decaying sinusoid given mathematically by Eq. (10) [where τ is

$$i = Ae^{-t/\tau}\sin(\omega t + \phi) \tag{10}$$

called the decay time and ω is shifted from its undamped value] and graphically by the curve labeled overdamping in Fig. 4*b*. *See* DAMPING.

The transresistance amplifier [current-in, voltage-out, related by Eq. (11), with a gain R_A whose units

$$v_{\text{out}} = R_A i_{\text{in}} \tag{11}$$

are resistance] simulates a negative resistance, which tends to cancel the effect of the physical damping resistor. If $R_A = R_D$ exactly, the output will be given exactly by Eq. (8), but if it is smaller the oscillation will decay (overdamping) and if it is larger the output will grow exponentially (underdamping; Fig. 4*b*). *See* NEGATIVE-RESISTANCE CIRCUITS.

In practice it is not possible to make $R_A = R_D$ exactly, and even if the equality did hold, nothing would define oscillation amplitude. Instead a slightly nonlinear amplifier is used, with higher gain at low voltages than at high voltages. If the initial amplitude is small, the amplifier dominates and the oscillations grow; if large, damping dominates and the oscillations decay. An equilibrium is reached with some amplitude of sine wave that is determined by the nonlinearity.

A rigorous treatment of the behavior of this system is relatively difficult, because it now has a nonlinear differential equation. If the amplifier nonlinearity is cubic, the Van der Pol equation (12) applies. This

$$\frac{d^2x}{dt^2} + \epsilon(x^2 - 1)\frac{dx}{dt} + x = 0 \tag{12}$$

equation has been studied extensively, and it displays most of the interesting properties of real oscillators. In particular, as the "strength" of the nonlinear term increases (that is, as ϵ increases), the sine wave becomes increasingly distorted and eventually its frequency starts to shift. Therefore these oscillators are at their most accurate when the nonlinearity is weakest.

There are practical difficulties with making the nonlinearity vanishingly small. In particular, if the damping resistance is not accurately known or if it may vary, it is necessary to have some safety margin of excess gain and therefore a stronger nonlinearity in order to guarantee oscillation. In some precision circuits a second type of nonlinearity, an automatic gain control, is used to estimate the appropriate gain so that a very weak nonlinearity can be used. *See* AUTOMATIC GAIN CONTROL (AGC).

Barkhausen criterion. The negative-resistance point of view above is one approach to design, but it is often hard to identify the negative resistor in a given circuit. Instead it is common to use a very general amplifier-based analysis and determine stability with the so-called Barkhausen criterion. In this approach the circuit is seen as a loop of an amplifier with gain A and a linear circuit with frequency-dependent gain $\beta(j\omega)$. The loop will oscillate with a perfect sine wave at some frequency ω_0 if at that frequency $A\beta(j\omega_0) = 1$ exactly. This means that at this frequency there is positive feedback with a gain of exactly 1, and any oscillation will sustain itself. The equation is usually expressed as a pair of conditions, one on the magnitude of the gain, $|A\beta(j\omega_0)|$, required for oscillation, and a second that requires the phase, $\phi[A\beta(j\omega_0)]$, to equal 0° or 360°. These conditions define the frequency of oscillation.

Figure 5*a* shows a circuit called a Wien bridge, commonly given as a simple example, that illustrates the use of the Barkhausen criterion. It consists of a voltage amplifier with some gain A, and an *RC* circuit. The gain $\beta(j\omega)$ of the *RC* circuit is given by expression (13). This complex quantity is plotted

$$\frac{j\omega CR}{1 + 3j\omega CR - (\omega CR)^2} \tag{13}$$

in Fig. 5*b*, where the variation of the magnitude and phase of $\beta(j\omega)$ withfrequency are shown. When

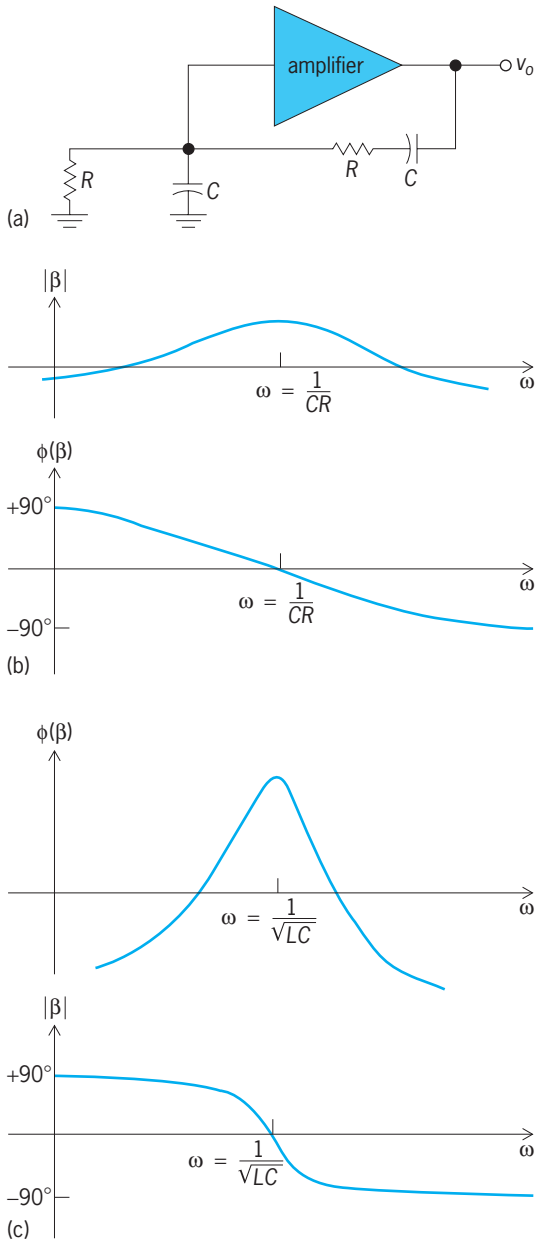

Fig. 5. Design of oscillators based on plotting variation of loop gain with frequency. (*a*) Wien bridge circuit. (*b*) Bode plots for the Wien bridge *RC-CR* network. (*c*) Bode plots for the *LC* resonator of Fig. 4*a*, shown for comparison.

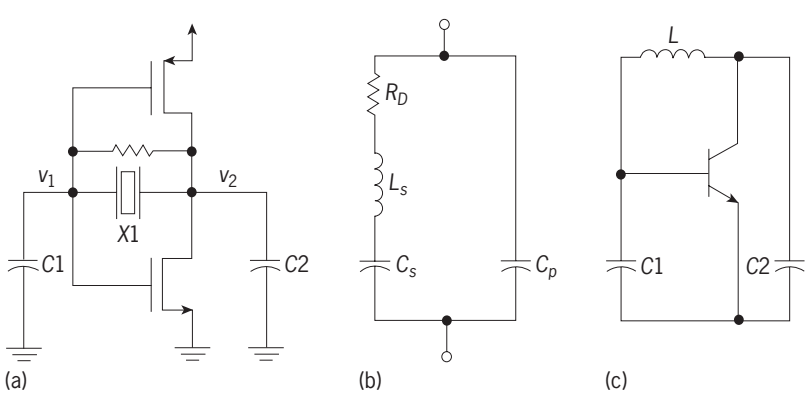

Fig. 6. Clock oscillator, a crystal oscillator. (*a*) Circuit diagram. (*b*) Electrical equivalent of piezoelectric crystal resonator. (*c*) General structure of a Colpitts oscillator.

magnitude and frequency are on logarithmic scales, as here, this is known as a Bode plot. *See* GAIN.

At frequency $\omega = 1/(CR)$, the phase of β is 0°. If the amplifier also has zero phase (ideal frequency-independent amplifiers have a phase of 0° or 180°), then this frequency is a candidate for oscillation if the gain A is appropriately chosen. The phase portion of Barkhausen's criterion therefore gives the oscillator frequency; the magnitude condition then defines the amplifier gain that is needed. At $\omega = 1/(CR)$, the loop gain is $A/3$. Thus the amplifier gain must be at least 3. In practice, A would be chosen somewhat larger than this minimum in order to guarantee oscillation even in the presence of small variations in component values. Oscillation amplitude would be defined by amplifier nonlinearities, such as clipping or by the intentional addition of other nonlinear elements such as diodes to the circuit.

Choice of circuits. There are many linear circuits to choose from, and they have different strengths and weaknesses. The Wien bridge circuit of Fig. 5, for example, is less costly to implement at audio frequencies than the *LC* circuit of Fig. 4, but is less accurate in frequency and produces more distorted waveforms. Bode plots for the Wien bridge (Fig. 5*b*) and the *LC* resonator (Fig. 5*c*) show why.

The phase curve for the *LC* oscillator passes through the critical frequency where phase is 0° with a much steeper slope than does the Wien bridge curve. This makes the *LC* oscillator frequency much less sensitive to the amplifier: suppose in each case that the amplifier has -1° of phase rather than the ideal zero (because a practical amplifier has frequency-dependent gain); then the new oscillation frequency will be that at which $\phi(\beta) = 1^\circ$. This is a much smaller error in frequency for the *LC* circuit than for the Wien bridge.

As to distortion, the *LC* circuit has a magnitude curve with a much sharper peak. Thus distortion, which introduces harmonic frequencies well above the peak into the loop, can be much more effectively filtered in the *LC* case than in the Wien bridge case. The Barkhausen criterion and the use of Bode plots provide a straightforward method of oscillator design as well as insight into important design choices.

Example. The most common oscillator circuit is probably that of **Fig. 6***a*, which generates the clock signals in most digital circuits, and of which a low-voltage variant keeps time in quartz watches. The two transistors are just a digital inverter in complementary metal-oxide-semiconductor (CMOS) technology, and the resistor biases the circuit to operate with $v_1 = v_2$. In these conditions the inverter acts as a transconductance amplifier—the amplifier part of an oscillator. The element $X1$ is a small slice of quartz crystal with electrodes attached to it, and has a mechanical resonance together with a piezoelectric effect that converts between mechanical motion and electric current. This combination gives it the electrical equivalent circuit shown in Fig. 6*b*, in which L_s, C_s, and R_D come from the mechanical resonance as translated by the piezoelectric effect. Capacitor

C_p comes from the ordinary physical capacitance between the electrodes and makes the design somewhat more complicated than for a simple *LC* loop. A crystal is an excellent choice for an oscillator, because it has an exceptionally sharp phase curve and the frequency is defined by very stable physical properties. *See* PIEZOELECTRICITY; QUARTZ CLOCK.

This oscillator is a variant of the classic Colpitts oscillator, of which an idealized equivalent circuit is shown in Fig. 6*c* (without biasing circuitry). Barkhausen analysis of this circuit shows that it oscillates at a frequency given by Eq. (14). The amplifier

$$\omega = 1 \Big/ \sqrt{L\frac{C_1C_2}{C_1+C_2}} \qquad (14)$$

gain (transconductance) required for oscillation depends on the amount of damping in the circuit. The circuit of Fig. 6*a* is obtained from the general form by replacing the inductor with a crystal and adding biasing details.

Practical LC oscillators. The Colpitts structure of Fig. 6*c* is used in many forms to make practical oscillators used at radio frequencies. A second widely used structure is the Hartley oscillator, which generally uses an *LC* tuned circuit and a transformer in the form shown in Fig. 3*b*. Transformer *T* has some inductance *L*, which resonates with *C*, and simultaneously couples the transistor's collector to its base to form a feedback loop. Bias voltage V_{bias} is chosen to keep the transistor in the so-called on condition at an appropriate current. (In practical circuits a more complicated *RC* network is used for bias generation, because operation of the circuit shown is overly sensitive to small changes in V_{bias}. At a frequency $1/(2\pi\sqrt{LC})$, the phase shift around the loop is zero, and the circuit has the potential to oscillate if the magnitude of loop gain exceeds unity. If loop gain is excessive, the nonlinearity of the transistors will become more pronounced, and the circuit behavior will begin to resemble that of the blocking oscillator in Fig. 3*a*.

Frequency Locking

If an external signal is injected into an oscillator, the natural frequency of oscillation may be affected. If the external signal is periodic, oscillation may lock to the external frequency, a multiple of it, or a submultiple of it, or exhibit an irregular behavior known as chaos. For example, in the 555 circuit of Fig. 2*a* a capacitor voltage charges up until it meets a reference voltage, and then rapidly discharges. Because the reference voltage is now varied with an externally generated signal, the two voltages meet earlier than they would have for a constant reference, and so the oscillation frequency is changed. Furthermore, the signals are more likely to meet when the reference voltage is low than when it is high, which tends to synchronize them. *See* CHAOS.

This locking behavior occurs in all oscillators, sometimes corrupting intended behavior (as when an oscillator locks unintentionally to a harmonic of the power-line frequency) and sometimes by design. Whether locking occurs or not depends on the extent of nonlinearity in the oscillator loop, manner of injection, amplitude of the external signal, and the ratio of free-running to external frequencies. The same mechanism can cause collections of oscillators with some coupling mechanism to synchronize to one another. The mathematics of the phenomenon is still being explored.

An important example of an oscillator that exploits this locking principle is the human heart. Small portions of heart muscle act as relaxation oscillators. They contract, incidentally producing an output voltage that is coupled to their neighbors. For a short time the muscle then recovers from the contraction. As it recovers, it begins to become sensitive to externally applied voltages that can trigger it to contract again (although it will eventually contract anyway). Each small section of heart muscle is thus an independent oscillator, electrically coupled to its neighbors, but the whole heart is synchronized by the frequency-locking mechanism. *See* CARDIAC ELECTROPHYSIOLOGY.

Martin Snelgrove

Bibliography. K. K. Clarke and D. T. Hess, *Communication Circuits: Analysis and Design,* 1971, reprint 1994; National Semiconductor Ltd., *Linear Databook 3,* 1988; A. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Oscillatory reaction

A chemical reaction in which some composition variable of a chemical system exhibits regular periodic variations in time or space. It is a basic tenet of chemistry that a closed system moves inexorably toward an unchanging state called chemical equilibrium. That motion can be described by the monotonic increase of entropy if the system is isolated, and by the monotonic decrease of Gibbs free energy

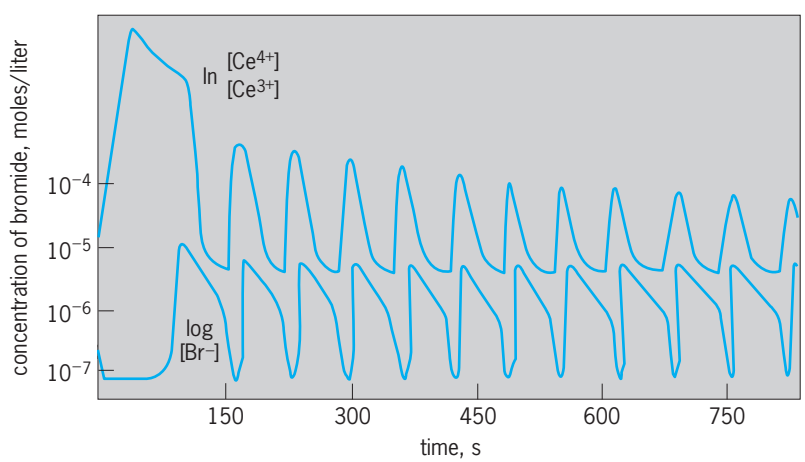

Fig. 1. Oscillatory behavior in a sulfuric acid medium containing potassium bromate, malonic acid, and cerous nitrate. The upper curve is the potential (in unspecified units) of a tungsten electrode and is related to the concentration ratio of cerium(IV)/cerium(III). The lower curve is the potential of an electrode sensitive to bromide ion, and the logarithmic scale at the left relates that potential to the absolute concentration of bromide. (*After R. J. Field, E. Körös, and R. M. Noyes, Oscillations in chemical systems, II. Thorough analysis of temporal oscillations in the bromate-cerium-malonic acid system, J. Amer. Chem. Soc., 94:8649–8664, 1972*)

Fig. 2. Rotating spiral bands of oxidation in a thin layer of solution containing the same reagents as in Fig. 1 except that the redox indicator ferrous phenanthroline has been substituted for cerous ion. (*Courtesy of Prof. A. T. Winfree, University of Arizona; from R. J. Field and R. M. Noyes, Mechanisms of chemical oscillators: Conceptual bases, Acc. Chem. Res., 10:214–221, 1977*).

if the system is constrained to constant temperature and pressure. Because of this universal restriction on what is possible in chemistry, it may appear bizarre when electrodes in a solution generate the oscillating potentials shown in **Fig. 1**.

The species taking part in a chemical reaction can be classified as reactants, products, or intermediates. In a closed system the concentrations of reactants decrease and those of products increase. Intermediates are formed by some steps and are destroyed by others. If there is only one intermediate, and if its concentration is always much less than the initial concentrations of reactants, this intermediate attains a stable steady state in which the rates of formation and destruction are virtually equal. The kind of oscillation reflected in Fig. 1 requires at least two intermediates which interact in such a way that the steady state of the total system is unstable to the minor fluctuations present in any collection of molecules. The concentrations of the intermediates may then oscillate regularly, although the oscillations must disappear before the inevitable monotonic approach to equilibrium.

Periodic chemical behavior may be temporal in a uniform system as illustrated in Fig. 1; it may involve simultaneous temporal and spatial variations as in **Fig. 2**; or it may involve spatial periods in a static system as in **Fig. 3**.

Well-authenticated examples of periodic chemical behavior have been known for almost a century, but until the 1970s most chemists either did not know about them or deliberately ignored them. In 1976 I. Prigogine of Brussels received the Nobel prize for his demonstration a few years before that periodic behavior was indeed consistent with the accepted principles of chemistry. Since that time, interest has developed rapidly, but most examples are still poorly understood. The phenomena are classified here according to types of chemical processes involved. Very different classification schemes may become more appropriate in the future. *See* CHEMICAL EQUILIBRIUM; ELECTRODE POTENTIAL; ENTROPY.

Redox oscillators. The systems whose chemistries are best understood all involve an element that can exist in several different oxidation states. Figure 1 illustrates the so-called Belousov-Zhabotinsky reaction, which was discovered in the Soviet Union in 1951. A strong oxidizing agent (bromate) attacks an organic substrate (such as malonic acid), and the reaction is catalyzed by a metal ion (such as cerium) that can exist in two different oxidation states.

As long as bromide ion (Br^-) is present, it is oxidized by bromate (BrO_3^-), as in reaction (1).

$$BrO_3^- + 2Br^- + 3H^+ \rightarrow 3HOBr \qquad (1)$$

When bromide ion is almost entirely consumed, the cerous ion (Ce^{3+}) is oxidized, as in reaction (2).

$$BrO_3^- + 4Ce^{3+} + 5H^+ \rightarrow HOBr + 4Ce^{4+} + 2H_2O \qquad (2)$$

Reaction (2) is inhibited by Br^-, but when the concentration of bromide has been reduced to a critical level, reaction (2) accelerates autocatalytically until bromate is being reduced by Ce^{3+} many times as rapidly as it was by Br^- when reaction (1) was dominant.

The hypobromous acid (HOBr) and ceric ion (Ce^{4+}) formed in reaction (2) then react with organic matter by a complicated sequence that can be summarized in reaction (3). Reaction (3) creates the

$$\begin{aligned} Ce^{4+} + HOBr + \text{organic matter} \rightarrow \\ Ce^{3+} + Br^- + \text{oxidized organic matter} \\ + \text{brominated organic matter} \end{aligned} \qquad (3)$$

bromide ion necessary to shut off the fast reaction

Fig. 3. Layers of crystallization from an initially uniform magma in the Skaergaard Intrusion in Greenland. (*From A. R. McBirney and R. M. Noyes, Crystallization and layering of the Skaergaard Intrusion, J. Petrol., 20:487–554, 1979*)

(2) and throw the system back to dominance by the slow reaction (1).

As other redox oscillators become understood, they fit the same pattern of a slow reaction destroying a species that inhibits a fast reaction that can be switched on autocatalytically; after a delay the fast reaction generates conditions to produce the inhibitor again. Most of the known examples involve positive oxidation states of chlorine (Cl), bromine (Br), or iodine (I). Elements like nitrogen (N), sulfur (S), chromium (Cr), and manganese (Mn) can also exist in several positive oxidation states, and it may be found that they also drive chemical oscillations.

The traces in Fig. 1 were obtained in a closed system such that no matter crossed its boundaries; it must eventually have decayed to a stable chemical equilibrium. Even more dramatic oscillations can be made to persist indefinitely if fresh chemicals are continuously added while material from the reactor is simultaneously removed. Such a system is called a continuously stirred tank reactor.

If the solution in Fig. 1 were unstirred but had a gradient in composition, oscillations in different regions would get out of phase, and an apparent wave would traverse the medium much the way flashing lights cause a message to move across a theater marquee.

Figure 2 illustrates a still more complex situation. Each light curve is a region of dominance by reaction (2). A combination of reaction and diffusion triggers an advance outward perpendicular to the wavefront into the dark region dominated by reaction (1). These trigger waves annihilate each other when they meet, and Fig. 2 shows two spirals spinning in opposite directions. This kind of behavior has been suggested to explain the fibrillations when a human heart loses its rhythm and degenerates to uncoordinated local contractions that result in death if the condition is not rapidly reversed. *See* OXIDATION-REDUCTION.

Nucleation oscillators. If molecules of a gas are produced at a steady rate in a uniform solution, the evolution of that gas may take place in pulses during which the solution foams up and then subsides. Several examples of this general phenomenon are known. As the chemical reaction proceeds homogeneously, the solution becomes more and more supersaturated, until a threshold is reached at which many microscopic bubbles are nucleated almost simultaneously. These bubbles grow ever faster as they become bigger, until they deplete the supersaturation and escape because of the Earth's gravitational field. The chemical reaction must then raise the supersaturation to the critical level before another burst of nucleation can occur. The result is that the solution foams up repeatedly like a shocked glass of beer and then subsides. Examples of such gas-evolution oscillators are the dehydration of formic acid (HCO_2H) by concentrated sulfuric acid and the decomposition of aqueous ammonium nitrite (NH_4NO_2) to produce elementary nitrogen. *See* SUPERSATURATION.

If a solution is not uniform but has gradients in composition or in temperature, then nucleation and growth of crystals may occur in bands. Figure 3 illustrates a geologic formation on the east coast of Greenland. A large magma chamber cooled very slowly, and the initially uniform material crystallized in a pattern of regular layers.

Oscillations coupled to mass transport. If a reaction in a solution is autocatalytically accelerated by a reactant coming from another phase, the rate may increase until transport of the critical reactant can no longer keep up. For certain situations, the system does not go to the anticipated steady state. Instead, some intermediate is depleted, the concentration of the critical reactant is renewed, and the rate begins to increase again. The only known examples involve reactions with dissolved oxygen in a solution in contact with the air. Detailed mechanisms are not yet understood.

Thermokinetic oscillators. Many chemical reactions give out heat, and rates are strongly dependent on temperature. If heat is not removed too rapidly from the reactor, reaction rate and temperature may couple to generate oscillations.

The known examples involve highly exothermic reactions like combustion or chlorination of organic compounds. No chemical mechanisms are yet understood in detail. In at least some examples, gradients of temperature and of composition are important in addition to changes in local rates of reaction.

Reactions on surfaces. Many important industrial reactions take place on the surfaces of catalysts. The occurrence of such a reaction may temporarily alter that surface or its temperature. Such effects sometimes couple to generate oscillations in the rate of reaction. These oscillations may or may not be of value for the reaction being carried out. Specific examples are being studied actively by chemical engineers.

The surfaces of electrodes may also be influenced by processes taking place on them, and periodic changes in voltage or current are well precedented.

Biological chemistry. Living organisms take in nutrients of high free energy and eliminate waste products of lower free energy. They therefore resemble the continuously stirred flow reactors. The degradation of those nutrients drives the essential vital processes. The pathways involve intricate couplings so that a decrease in free energy of one species contributes in forming species like adenosine triphosphate (ATP) that have increased free energy. Many important intermediates follow repeated cyclic paths while the nutrients are degraded. *See* ADENOSINE TRIPHOSPHATE (ATP).

If all of the processes of metabolism took place at the same rate in the same place, the only net effect would be an increase in entropy, and life would not exist. Processes involving some intermediates must be separated in space or in time from processes involving other intermediates.

Separation in space can be accomplished by membranes permeable to some species but not to others. Such membranes are ubiquitous in biological organisms. *See* CELL MEMBRANES.

Separation in time can be accomplished by oscillatory reactions that turn component processes on and

off much as happens with the Belousov-Zhabotinsky reaction described above. It can hardly be accidental that periodicities are observed in many biological activities.

One of the best chemical examples involves oscillations during oxidation of glucose catalyzed by cell-free yeast extracts. The enzyme phosphofructokinase (PFK) is strongly implicated, and its activity is influenced by the oxidation products.

Aggregation of the slime mold *Dictostelium discoideum* takes place because individual cells move in the direction of a received pulse of cyclic adenylic acid (cyclic AMP) and simultaneously emit a pulse in the opposite direction. Aggregating cells may create spiral patterns resembling those in Fig. 2 except that motion is inward instead of outward.

Circadian rhythms with periods of about 24 h are common in biology. Anybody who has taken a jet halfway around the world becomes well aware of their effects.

Undoubtedly many periodic biological processes of chemical origin are not even recognized. None can yet be described in mechanistic detail. Development of the quantitative explanation characteristic of a mature science is anticipated. Richard M. Noyes

Bibliography. I. R. Epstein et al., Oscillating chemical reactions, *Sci. Amer.*, 248(3):112–123, 1983; R. J. Field and M. Burger (eds.), *Oscillations and Traveling Waves in Chemical Systems*, 1985; H. Haken, *Advanced Synergetics*, 1993; G. Nicolis and I. Prigogine, *Self-Organization in Nonequilibrium Systems*, 1977; S. K. Scott, *Oscillations, Waves, and Chaos in Chemical Kinetics*, 1994; A. T. Winfree, *The Geometry of Biological Time*, 2d ed., 2001.

Oscilloscope

An electronic measuring instrument which produces a display showing the relationship of two or more variables. In most cases it is an orthogonal (x,y) plot with the horizontal axis being a linear function of time. The vertical axis is normally a linear function of voltage at the signal input terminal of the instrument. Because transducers of many types are available to convert almost any physical phenomenon into a corresponding voltage, the oscilloscope is a very versatile tool that is useful for many forms of physical investigation. *See* TRANSDUCER.

The oscillograph is an instrument that performs a similar function but provide a permanent record. The light-beam oscillograph used a beam of light reflected from a mirror galvanometer which was focused onto a moving light-sensitive paper. These instruments are obsolete. The mechanical version, in which the galvanometer drives a pen which writes on a moving paper chart, is still in use, particularly for process control. *See* GALVANOMETER; GRAPHIC RECORDING INSTRUMENTS.

Oscilloscopes are one of the most widely used electronic instruments because they provide easily understood displays of electrical waveforms and are capable of making measurements over an extremely wide range of voltage and time. Although a very large number of analog oscilloscopes are in use, digitizing oscilloscopes (also known as digital oscilloscopes or digital storage oscilloscopes) are preferred, and analog instruments are likely to be superseded. *See* ELECTRONIC DISPLAY.

Analog Oscilloscopes

An analog oscilloscope, in its simplest form, uses a linear vertical amplifier and a time base to display a replica of the input signal waveform on the screen of a cathode-ray tube (CRT) [**Table 1**]. The screen is typically divided into 8 vertical divisions and 10 horizontal divisions. Analog oscilloscopes may be classified into nonstorage oscilloscopes, storage oscilloscopes, and sampling oscilloscopes.

Analog nonstorage oscilloscopes. These oscilloscopes are the oldest and most widely used type. Except for the cathode-ray tube, the circuit descriptions also apply to analog storage oscilloscopes.

A typical oscilloscope (**Fig. 1**) might have a bandwidth of 150 MHz, two main vertical channels plus two auxiliary channels, two time bases (one usable for delay), and a 3.2 × 4 in. (8 × 10 cm) cathode-ray-tube display area; and it might include on-screen readout of some control settings and measurement results. A typical oscilloscope is composed of five basic elements: (1) the cathode-ray tube and associated controls; (2) the vertical or signal amplifier system with input terminal and controls; (3) the time base, which includes sweep generator, triggering circuit, horizontal or x-amplifier, and unblanking circuit; (4) auxiliary facilities such as a calibrator and on-screen readout; and (5) power supplies.

The first four elements will be described in detail; power supplies are used in all electronic equipment and require no special description. *See* ELECTRONIC POWER SUPPLY.

Cathode-ray tube. The central component in a cathode-ray oscilloscope is the cathode-ray tube, which in its simplest form consists of an evacuated glass container with a fluorescent screen at one end and a focused electron gun and deflection system at the other.

Either magnetic or electric fields may be used for focusing and deflection. Electrostatic deflection is almost universally used for oscilloscopes because it is capable of superior high-frequency response.

The cathode-ray-tube designer is faced with four primary technical objectives: (1) high deflection sensitivity up to the desired frequency; (2) a bright image for ease of observation and photography; (3) small spot size relative to the image area; and (4) accurate

TABLE 1. Vertical measurement capabilities of analog oscilloscopes

Maximum bandwidth	Rise time	Deflection sensitivity	Measurement accuracy
1 MHz	350 ns	10 μV/division	5%
40 MHz	8.75 ns	5 mV/division	3%
200 MHz	1.8 ns	2 mV/division	2.5%
350 MHz	1 ns	2 mV/division	2%

deflection geometry. All of these are interdependent so that each tube design is a compromise chosen to best meet the needs of the intended users. *See* CATHODE-RAY TUBE.

Vertical amplifier. The signal to be observed is generally applied to the vertical or y axis of the oscilloscope. The vertical amplifier is required to provide sufficient gain that small signals may be displayed with suitable amplitude on the cathode-ray tube. All modern oscilloscope vertical amplifiers are dc-coupled so as to allow accurate observation of slowly changing signals. The vertical amplifier must amplify the signal so as to minimize distortion in its wave shape. It is thus required to be highly linear and have a suitable frequency, phase, and transient response. In addition, because it is desirable to view the signal that triggers the time base, a signal delay is incorporated in the vertical amplifier to allow the time base to start before the vertical signal reaches the cathode-ray-tube vertical deflection plates. A delay line, either with coaxial cables or of a specially wound type, is generally used for this purpose and is located in the vertical amplifier after the trigger pick-off point. *See* AMPLIFIER; DELAY LINE; DIRECT-COUPLED AMPLIFIER; DISTORTION (ELECTRONIC CIRCUITS).

To obtain a suitable image size on the cathode-ray tube, a convenient means of varying the amplifier gain is needed. The method used must vary the gain without introducing frequency or amplitude distortion. Generally, a high-impedance frequency-compensated attenuator switched in an accurate 1-2-5 sequence is used before the amplifier input, and a continuously variable control covering a range of at least 2.5:1 is used within the amplifier to bridge between the steps of the input attenuator. *See* ATTENUATION (ELECTRICITY).

All oscilloscope vertical amplifiers provide for variable positioning on the cathode-ray-tube screen and a choice of ac or dc coupling for the input signal. The positioning control inserts a variable current controllable by the user at a suitable point in the vertical amplifier, and the input coupling control inserts a capacitor in series with the input terminal when ac coupling is chosen. AC coupling is typically used when a small ac signal is present together with a large dc signal which would otherwise deflect the display off-screen.

The input connector of most oscilloscope signal amplifiers has one terminal connected to the chassis of the instrument. For safety reasons, the chassis is usually connected to ground through the mains lead. Most waveforms also have a ground reference, so the signal and the instrument are compatible. However, there is a danger that the resulting ground loop will introduce errors, especially with low signal levels. Some instruments have isolated inputs to deal with this problem. Add-on input isolators are also available, but these restrict the working bandwidth. Sometimes it is necessary to observe the signal between two points, both of which are at significant potentials relative to ground. Differential amplifiers having two high-impedance input terminals are available for this purpose. Many oscilloscopes allow the difference between two inputs to be displayed, achieving a similar result; however, the two channel gains must be carefully matched at the time of use. *See* DIFFERENTIAL AMPLIFIER.

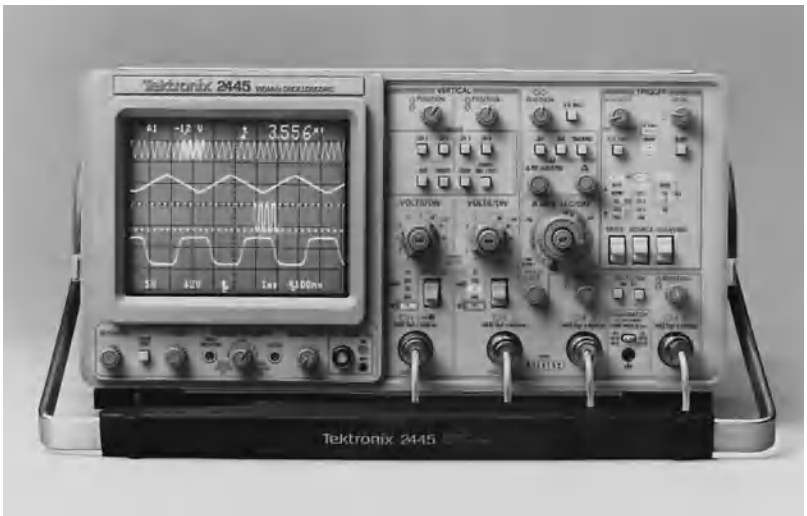

Fig. 1. Four-channel 150-MHz analog oscilloscope. (*Tektronix*)

Most oscilloscopes incorporate a multitrace vertical system which provides for the simultaneous display of two or more different signals for comparison purposes. The most common method used to achieve a multitrace display is to incorporate a vertical input amplifier complete with all controls for each channel and then electronically switch between them at a rate faster than the human persistence of vision. In this case the display will appear flicker-free. To achieve this result, two alternative switching methods are employed with a choice available to the user. One method is to switch during the sweep return time (alternate mode) so that succeeding sweeps display first one channel, then another, and then repeat. This gives a flicker-free display provided the sweep repetition rate is high enough. If this is not the case, then the chopped mode must be used. Here the switching between channels is done continuously at a high rate with cathode-ray-tube beam blanking used to avoid displaying the switching waveform itself.

Some applications require that two simultaneous high-speed signals occurring only once (single shot) or at a very low repetition rate be observed or photographed. Neither of the switching systems above is suited for this measurement, and a dual-beam oscilloscope must be employed. A dual-beam oscilloscope comprises two completely independent vertical amplifiers feeding two separate sets of cathode-ray-tube vertical deflection plates. Either one or two time bases and cathode-ray-tube horizontal deflection systems may be employed. Since the vertical systems are independent, there are no compromises imposed by having to choose between alternate or chopped modes as in a time-shared multitrace oscilloscope. Dual-beam oscilloscopes are more expensive than single-beam types and in general have poorer matching between the transfer response of the vertical channels.

Time bases. In the standard oscilloscope waveform presentation the horizontal axis on the cathode-ray tube represents a linear function of time. This is achieved by the use of a linear sweep generator feeding the cathode-ray-tube horizontal deflection plates via the horizontal amplifier. For a satisfactory display the sweep generator must be correctly triggered and the cathode-ray-tube beam blanked except when a sweep is present.

A cathode-ray-tube presentation that is a linear function of time requires that a voltage changing linearly with time be applied to the cathode-ray-tube plates, generally in push-pull to avoid geometrical distortion in the cathode-ray tube itself. The sweep generator is started by a fast edge related to the vertical input signal. This edge may be applied from an external trigger input or derived within the vertical amplifier. The sweep generator produces a linearly increasing voltage (**Fig. 2***a*), which is amplified by the horizontal amplifier and applied to the horizontal deflection plates. At the end of the sweep the generator rapidly returns to its initial voltage, the time taken being the retrace time. This is followed by the hold-off time, during which the circuits return to a standard initial condition. The cycle is repeated on the receipt of the first trigger signal after the end of the hold-off period. Sometimes the hold-off time may be adjusted by the user, making it possible to choose the overall cycle time and so repetitively display the same part of a complex signal.

Dual time bases are used to obtain a detailed view of part of a complex waveform. One time base (A) is used to establish a delay; the second, faster time base (B) is initiated either by the end of the delay or by the first subsequent trigger pulse. To set up the two time bases, the delaying sweep is first used as a simple time base and adjusted so that the whole complex waveform is presented on the screen. The delayed sweep B appears as a bright region on the trace (Fig. 2*b*). Its position and duration are adjusted in order to cover the region of interest. When the delayed display mode is selected, the complete screen horizontal scan is provided by the delayed sweep (Fig. 2*c*), and the details of a relatively brief part of the waveform are revealed. In better-quality instruments the trace brightness is automatically compensated to suit the change in sweep rate.

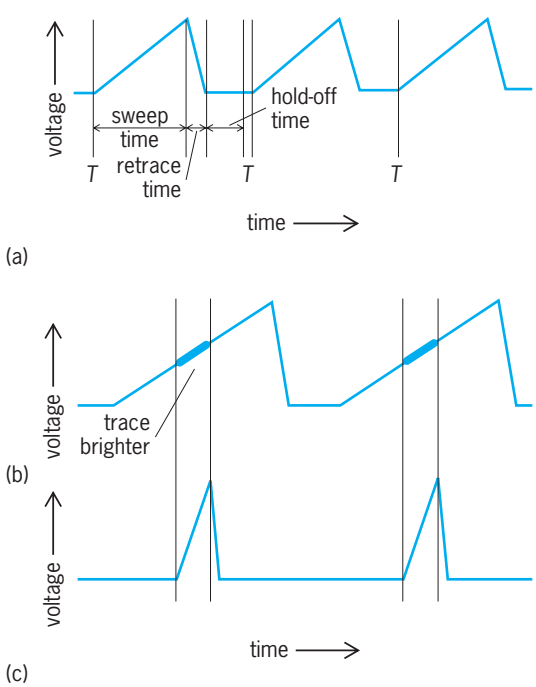

Fig. 2. Sweep generator waveforms. (*a*) Single waveform, showing instant of trigger (T). (*b*) Dual time base in set-up mode: sweep by A, intensified by B. (*c*) Dual time base in delayed display mode: scan delayed by A, sweep by B.

Trigger generator. To ensure a stable display of the incoming signal, each sweep must start at a time corresponding to precisely the same point on the signal waveform. The information needed by the sweep generator to accomplish this may come either from the signal itself or from another signal which has a consistent time relationship to the signal to be observed. The input of the trigger generator is thus capable of being switched between several sources, the most commonly used being a portion of the signal at a suitable point before the delay line in the vertical amplifier. Whatever its source, the signal is applied to a circuit which generates a very fast output pulse every time the incoming signal reaches a certain voltage. This pulse is used to start the sweep generator, providing the hold-off time has ended. Front panel controls allow selection of the slope and voltage of the incoming signal that results in a trigger signal being generated. Most oscilloscopes have an automatic mode of triggering which arranges the trigger circuit to sense the input signal and provide a trigger pulse output at a point on the waveform which will generally result in a satisfactory display. When no signal is present, the time base is made to run, resulting in a bright horizontal line (baseline) being displayed which assists the user in selecting the kind of display that is needed. *See* TRIGGER CIRCUIT.

Calibrators. Most oscilloscopes incorporate an accurate square-wave calibrator which is used to check the vertical-gain calibration and adjust the probe, if one is used, for optimal response. In addition, some calibrators may be sufficiently accurate in time to allow the oscilloscope time-base calibration to be checked.

Sampling oscilloscopes. The limit of speed of vertical response of analog nonstorage oscilloscopes is determined by the design of the vertical amplifier and by the cathode-ray-tube deflection plate bandwidth. The fastest general-purpose oscilloscope available has a bandwidth from 0 (direct current) to 1 GHz; wider bandwidths are possible, but cost and power consumption rapidly increase. Equivalent-time or sampling oscilloscopes provide much higher bandwidths at reasonable cost and power consumption, but the technique requires that the signal to be observed be repetitive at a suitable minimum rate. It makes use of extremely short pulses which are used to sample the incoming signal. Samples are taken at a slightly later time at each recurrence of the trigger pulse. The samples are amplified, lengthened in time, and displayed at a relatively slow sweep speed. Some

important limitations of this type of oscilloscope are that single transients cannot be observed and that it may take a considerable time to build up a complete picture of the incoming signal if its repetition rate is low. A variation on the sampling technique described above is to take samples in a random fashion and then reconstitute the samples in the correct equivalent time and again display at a relatively slow sweep speed. Both sampling methods are used in commercially available oscilloscopes with bandwidths from 0 (dc) to over 10 GHz. *See* VOLTMETER.

Analog storage oscilloscopes. It is often desirable to store a waveform after the originating signal has ceased to exist, particularly in the case of transient signals. The simplest method is to photograph the transient waveform. Photographic recording is, however, inconvenient, time-consuming, and relatively expensive and it requires some skill. Analog storage oscilloscopes, were developed to solve this problem. Digitizing oscilloscopes also store waveforms and, because of their numerous advantages, are rapidly replacing analog storage oscilloscopes.

The analog storage oscilloscope substitutes a storage cathode-ray tube and control circuits for the standard type. There are several types of storage cathode-ray tube, but all have the property of being able to display a waveform for a limited length of time after the originating signal has disappeared. The storage time is at least sufficient for convenient photography and for relatively leisurely visual observation. Most storage oscilloscopes provide several modes of operation such as (1) nonstorage operation similar to a nonstorage oscilloscope, (2) medium-speed storage with long storage time, and (3) high-speed storage with relatively short storage time.

Storage cathode-ray tubes differ from nonstorage types in that they incorporate a target, in some cases the actual phosphor, which changes its electrical state after being scanned by the electron beam and results in a lingering charge storage pattern representing the original waveform. After the original signal has ended, the cathode-ray-tube operating voltages are changed so as to display a replica of the charge storage pattern; simultaneously, all further vertical input signals are stopped from affecting the cathode-ray tube until the user reactivates the oscilloscope. Cathode-ray-tube storage oscilloscopes are available which are capable of storing waveforms at up to 500 MHz.

Photography of waveforms. Many oscilloscopes are provided with fittings that enable matching cameras to be attached. By using these, permanent records of normal or stored waveforms can be obtained. To photograph single fast transients, special cathode-ray-tube phosphors, wide-aperture lenses, and high-speed films can be used. *See* CAMERA; LENS (OPTICS); PHOTOGRAPHY.

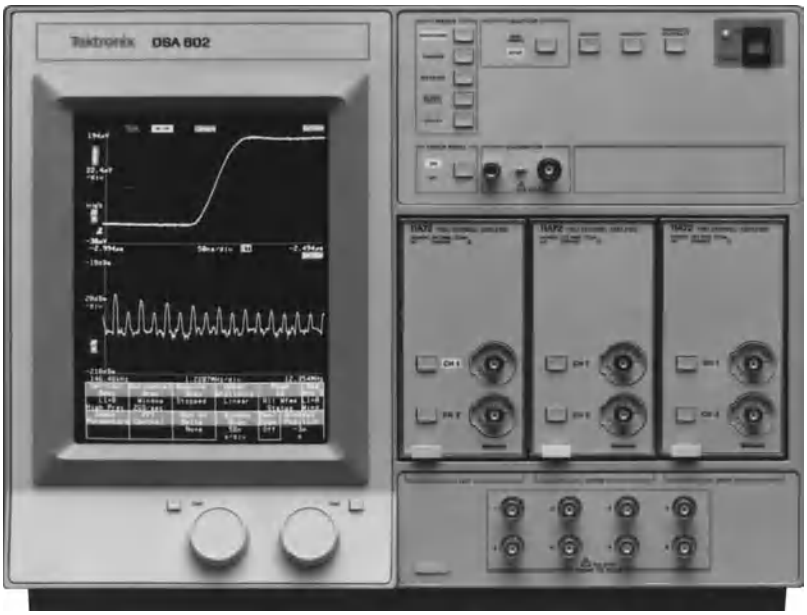

Fig. 3. High-speed general-purpose digitizing oscilloscope. (*Tektronix*)

Digitizing Oscilloscopes

Digital techniques are applied to both timing and voltage measurement in digitizing oscilloscopes (**Fig. 3; Table 2**). A digital clock determines sampling instants at which analog-to-digital converters obtain digital values for the input signals. The resulting data can be stored indefinitely or transferred to other equipment for analysis or plotting. *See* VOLTAGE MEASUREMENT; WAVEFORM DETERMINATION.

Components. In its simplest form a digitizing oscilloscope comprises six basic elements: (1) analog vertical input amplifier; (2) high-speed analog-to-digital converter and digital waveform memory; (3) time base, including triggering and clock drive for the analog-to-digital converter and waveform memory; (4) waveform reconstruction and display circuits; (5) display, generally, but not restricted to, a cathode-ray tube; (6) power supplies and ancillary functions. In addition, most digitizing oscilloscopes provide facilities for further manipulation of waveforms prior to

TABLE 2. Capabilities of digitizing oscilloscopes

Maximum bandwidth	Rise time	Maximum digitizing frequency, samples per second	Record length, samples*	Deflection sensitivity	Measurement accuracy	Resolution
50 MHz	7 ns	2×10^8	1,000	2 mV/division	2%	8 bit
350 MHz	1 ns	1×10^8	60,000	1 mV/division	1.5%	8–12 bit
500 MHz	0.7 ns	2×10^9	50,000	1 mV/division	1.5%	8–12 bit
8 GHz	44 ps	5×10^4	15,000	1 mV/division	0.7%	8 bit

* The maximum record length may require the fitting of a nonstandard option.

display, for direct measurements of waveform parameters, and for connection to external devices such as computers and hard-copy units.

Analog vertical input amplifier. This device performs functions identical to those of a vertical amplifier in a real-time oscilloscope; however, its output feeds an analog-to-digital converter instead of the cathode-ray tube.

High-speed analog-to-digital converter. The analog-to-digital converter samples the incoming vertical signal and provides a digital value for each sample. An 8-bit converter is commonly used, providing a resolution of 1 in 256. Higher resolution can be obtained by using a longer-scale converter or by building up a set of data from many cycles of signal. The latter method can provide the effect of an additional 4 bits of resolution but cannot be applied to single-shot measurements. *See* ANALOG-TO-DIGITAL CONVERTER.

Time base. The analog-to-digital converter is regularly triggered by the time-base clock at a rate selected by the user. During data acquisition the resulting values are stored in adjacent positions of the waveform memory. A trigger circuit similar to that used in an analog oscilloscope identifies a particular point on the waveform. This is used to identify the sample which corresponds to the trigger point. A selected number of further samples are taken after the trigger point. It is therefore possible to arrange that the memory contains a number of samples before, and a different number after, the trigger point. A total record of 1000 samples is typical. When these data are displayed, the effect is as if there was a delay in the signal channel, as information from a period preceding the trigger is included. The effective length of the delay can be adjusted by varying the number of samples which are accepted after the trigger point. No conventional delay line is required in the signal circuit of a digitizing oscilloscope.

Waveform reconstruction and display. The digitized input signal, now in memory, is used along with the knowledge of the original digitizing rate to reconstruct an oscilloscope-type display of amplitude versus time. The display device used is generally a cathode-ray tube, but since the reconstructed display rate can be chosen independently of the original signal speed, other types of display devices may be selected if their characteristics are more desirable. The reconstructed display may display one dot for each sample or may use straight or curved lines to join the dots to make the display easier to interpret.

Real-time and equivalent-time operation. All digitizing oscilloscopes operate in the real-time mode just described, where the incoming signal is digitized continuously and the trigger, in effect, freezes the digitized data for subsequent processing and display. Because the input signal is sampled at discrete points in time, the resulting digitized data can contain no information at a frequency greater than half the digitizing frequency. For example, an oscilloscope with a 100-MHz maximum sampling frequency can provide no information above 50 MHz. Even at this frequency, only two points per cycle are available, and the display is not very useful. By comparison, an analog oscilloscope shows some information above its −3-dB bandwidth frequency. Thus, no direct comparison can be made between the bandwidth of an analog oscilloscope and the maximum digitizing frequency of a digitizing oscilloscope when used in its real-time mode.

Most measured signals are repetitive. For such signals, the effective bandwidth of a digitizing oscilloscope may be considerably extended by using the equivalent-time mode of operation. Here, sampling is done as in an analog sampling oscilloscope except that the sample values are subsequently digitized and stored in digital memory. The limits on effective bandwidth are the same for both. Because a very large number of samples may be taken in this mode, the oscilloscope bandwidth may be specified. Most digitizing oscilloscopes provide both modes of operation.

Advantages and limitations. In almost all respects, digitizing oscilloscopes can match the capabilities of analog types. The existence of the data in digital form gives the technique many advantages, including:

1. The data can be processed within the instrument or sent elsewhere by using standard interfaces.
2. Mathematical processes, from addition and subtraction to Fourier transformations can be carried out. *See* FOURIER SERIES AND TRANSFORMS.
3. Signal parameters such as root-mean-square values and rise times can be calculated and displayed.
4. Signals from repeated sweeps may be averaged, giving the effect of up to 3 bits of additional resolution are reducing random noise.
5. Hard-copy representations of waveforms and their parameters can be created easily.
6. Considerable improvements in accuracy are possible, as the resolution is not limited by optical limitations.
7. The display is not limited to a cathode-ray tube designed to suit the requirements of the signal. Color can be used to enhance legibility. *See* COMPUTER GRAPHICS.
8. Intelligent software can be used to identify sharp transitions and cause a higher sampling rate to be used in such regions. This is called windowing.
9. Completely automatic measuring systems can be constructed.

Digitizing oscilloscopes of modest performance are available in which the techniques are used in conjunction with a liquid-crystal display to provide a versatile hand-held instrument which combines all the functions of an oscilloscope, digital multimeter, and frequency meter. *See* LIQUID CRYSTALS.

Digitizing oscilloscopes suffer from the phenomenon of aliasing, in which a false waveform is recorded which bears little or no similarity to the actual signal. One possible cause is a sampling rate that is too low compared with the highest signal frequency. In perceptual aliasing, the eye of the observer is deceived and tends to link the dots which are nearest on the screen. These may well not be the samples of the waveform which are closest in time. Perceptual aliasing does not occur with displays where the dots are linked by calculated lines.

Plug-in Oscilloscopes

So-called plug-in units are often used in high-performance oscilloscopes. These units are interchangeable hardware modules capable of altering certain major characteristics of an oscilloscope without causing the user to have to buy a completely new instrument. Their most common application is in giving a wider choice of vertical input performance characteristics than would be economically possible to build in. They are also used to alter horizontal characteristics such as triggering or time-base type, the latter only in analog oscilloscopes. Certain performance characteristics or combinations thereof are available only with plug-in oscilloscopes.

Oscilloscope Selection

Higher measurement accuracy is available from digitizing oscilloscopes (Tables 1 and 2). The first decision to be made in choosing an oscilloscope is whether this or any of the other properties exclusive to the digitizing type are essential. If not, the option of an analog design remains. The selected instrument must be appropriate for the signal under examination. It must have enough sensitivity to give an adequate deflection from the applied signal, sufficient bandwidth, adequately short rise time, and time-base facilities capable of providing a steady display of the waveform. An analog oscilloscope needs to be able to produce a visible trace at the sweep speed and repetition rate likely. A digitizing oscilloscope must have an adequate maximum digitizing rate and a sufficiently long waveform memory.

Signal amplitude. The lowest expected signal amplitude should provide over two divisions of deflection at the oscilloscope's minimum deflection factor, and the highest amplitude should provide no more than full screen. Probe attenuation must be taken into account.

Rise time. The fastest rise time of the signal and the degree of acceptable measurement error determine the needed rise time of the oscilloscope. If n is the maximum percentage error acceptable when a signal having a rise time of t_s is viewed, then the oscilloscope's rise time must be no more than that approximately given by expression (1). For instance,

$$\left[\left(1+\frac{n}{100}\right)^2-1\right]^{1/2}\times t_s \qquad (1)$$

if $n = 10\%$ and $t_s = 10$ ns, then the oscilloscope's rise time must be no more than 4.6 nanoseconds.

Conversely, the true rise time of a signal may be approximately estimated from the measured rise time, t_m, and the oscilloscope's known rise time, t_0, using Eq. (2).

$$\text{True rise time} = \left(t_m^2 - t_0^2\right)^{1/2} \qquad (2)$$

Oscilloscope bandwidth. The bandwidth of an oscilloscope is the frequency at which the vertical amplitude response has fallen by 3 dB (about 29%) from its low-frequency value. The way in whichthe amplitude response deteriorates with frequency needs to be carefully controlled if the best possible response to pulses is to be obtained. All modern instruments achieve this when correctly adjusted. Under these conditions the rise time and bandwidth are related approximately by Eq. (3).

$$\text{Bandwidth (in MHz)} = \frac{350}{\text{rise time (in ns)}} \qquad (3)$$

Signal Acquisition

Oscilloscopes of up to 300-MHz bandwidth generally have an input impedance of their signal input equivalent to 1 megohm with 15–30 picofarads capacitance in parallel. Connection to the circuit under test is generally done by using probes. In order to increase the impedance connected to the test circuit, probes with 10 times attenuation are most commonly used. These have a typical input impedance at their tip equivalent to 10 megohms with 10 pF in parallel. The probes are carefully designed to introduce minimum distortion in the signal they transmit. Above 200–300 MHz, the 10-pF-probe input capacitance has an undesirably low reactance (10 pF has a reactance of 53 ohms at 300 MHz), and other methods are needed. Most common is the provision of an accurate 50-ohm input to the oscilloscope. The user can either terminate the signal at the oscilloscope input, or a low-impedance probe can be used which again divides by 10 times, but has an impedance represented by 500 ohms and about 2 pF; at 300 MHz this gives an impedance of 236 ohms and is thus higher than in the previous case. Active probes are also available; they contain a miniaturized input circuit in the probe itself and transmit the signal over a 50-ohm line to the oscilloscope's input. Active probes provide the highest input impedance in the range of 100–1000 MHz but are expensive and relatively fragile. *See* ALTERNATING-CURRENT CIRCUIT THEORY; ELECTRICAL IMPEDANCE.

Calibration. High-impedance voltage divider probes have a capacitor adjustment in the probe that must be correctly set. This is normally achieved by connecting the probe tip to the oscilloscope's square-wave calibrator and adjusting the capacitor until the observed waveform shows a correct step response. In addition, the voltage measurement accuracy of the oscilloscope-probe combination should be optimized by adjusting the oscilloscope's vertical gain for the correct deflection with the probe tip connected to the calibrator. These two steps ensure that the best amplitude and frequency response are obtained. In addition, connections to the signal should be kept as short as possible, since long connections from either the ground or probe tip can lead to the pick-up of unwanted signals and to deterioration in the system's transient and frequency responses.

Current probes. Oscilloscopes indicate and measure voltages. In some circumstances, the current, rather than the voltage, is of interest, and it may not be possible to insert a small resistor in the current path to produce a voltage signal. Systems having probes which clip around a conductor and sense the

magnetic field are available to satisfy this requirement. A variety of products offer the capability of operating down to dc, sensitivities of 1 mA per division, peak current measurements up to kiloamperes, and maximum frequency limits of 15 MHz to 1 GHz. R. B. D. Knight

Bibliography. R. L. Goodman, *Using the Triggered Sweep Oscilloscope*, 1987; I. Hickman, *Oscilloscopes: How to Use Them, How They Work*, 5th ed., 2000; J. D. Lenk, *Handbook of Oscilloscopes*, 1982; R. A. Penfold, *How to Use Oscilloscopes and Other Test Equipment*, 1989; S. Prentiss, *The Complete Book of Oscilloscopes*, 2d ed., 1992.

Osmium

A chemical element, Os, atomic number 76, atomic weight 190.2. The element is a hard white metal of rare natural occurrence, usually found in nature alloyed with other platinum metals. *See* METAL; PERIODIC TABLE; PLATINUM.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Physical properties of the element, which is found as seven naturally occurring isotopes, are given in the **table**. The metal is exceeded in density only by iridium. Osmium is a very hard metal and unworkable, and so it must be used in cast form or fabricated by powder metallurgy. Osmium is a third-row transition element and has the electronic configuration $[Xe](4f)^{14}(5d)^6(6s)^2$; in the periodic table it lies below iron (Fe) and ruthenium (Ru). In powder form the metal may be attacked by the oxygen in air at room temperature, and finely divided osmium has a faint odor of the tetraoxide. In bulk form it does not oxidize in air below 500°C (750°F), but at higher temperatures it yields OsO_4. It is attacked by fluorine or chlorine at 100°C (212°F). It dissolves in alkaline oxidizing fluxes to give osmates (OsO_4^{2-}). *See* ELECTRON CONFIGURATION; IRIDIUM; IRON; RUTHENIUM.

The chemistry of osmium more closely resembles that of ruthenium than that of iron. The high oxidation states VI and VIII (OsO_4^{2-} and OsO_4) are much more accessible than for iron. *See* OXIDATION-REDUCTION.

Osmium forms many complexes. In water, osmium complexes with oxidation states ranging from II to VIII may be obtained. Oxo compounds, which contain Os═O, are very common and occur for oxidation states IV to VIII. Although OsO_4 is tetrahedral in the gas phase and in noncomplexing solvents such as dichloromethane (CH_2Cl_2), it tends to be six-coordinate when appropriate ligands are available; thus, in sodium hydroxide (NaOH) solution, dark purple $OsO_4(OH)_2^{2-}$ is formed from OsO_4. Similarly, $OsO_2(OH)_4^{2-}$ is formed by addition of hydroxide to osmate anion. Analogous reactions with ligands such as halides, cyanide, and amines give osmyl derivatives like the cyanide-deduct $OsO_2(CN)_4^{2-}$, in which the trans dioxo group (O═Os═O) is retained.

Osmium tetraoxide, a commercially available yellow solid (melting point 40°C or 104°F), is used commercially in the important *cis*-hydroxylation of alkenes and as a stain for tissue in microscopy. It is poisonous and attacks the eyes. Osmium metal is catalytically active, but it is not commonly used for this purpose because of its high price. Osmium and its alloys are hard and resistant to corrosion and wear (particularly to rubbing wear). Alloyed with other platinum metals, osmium has been used in needles for record players, fountain-pen tips, and mechanical parts. *See* ALLOY; STAIN (MICROBIOLOGY); TRANSITION ELEMENTS. Carol Creutz

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; J. R. Davis (ed.), *Metals Handbook: Desk Edition*, 2nd ed., ASM International, 1998.

Principal properties of osmium

Property	Value
Density, g/cm^3	22.6
Naturally occurring isotopes (% abundance)	184 (0.018)
	186 (1.59)
	187 (1.64)
	188 (13.3)
	189 (16.1)
	190 (26.4)
	192 (41.0)
Ionization enthalpy, kJ/mol: 1st 2d	840
	1640
Oxidation states	−1 to VIII
Most common	IV, VI, VIII
Ionic radius, Os^{4+}, nm	0.078
Melting point, °C (°F)	3050 (5522)
Boiling point, °C (°F)	5500 (9932)
Specific heat, cal/g °C	0.032
Crystal structure	Hexagonal close-packed
Lattice constant *a* at 25°C, nm *c/a* at 25°C	0.27341
	0.15799
Thermal neutron capture cross section, barns	15.3
Thermal conductivity, 0–100°C, (cal · cm)/(cm^2 · s · °C)	0.21
Linear coefficient of thermal expansion at 20–100°C, (μin./in./°C)	6.1
Electrical resistivity at 0°C, $\mu\Omega$-cm	8.12
Temperature coefficient of electrical resistance, 0–100°C/°C	0.0042
Young's modulus at 20°C, lb/in.^{2}, static	81×10^6

Osmoregulatory mechanisms

Physiological mechanisms for the maintenance of an optimal and constant level of osmotic activity of the fluid within and around the cells, considered to be most favorable for the initiation and maintenance of vital reactions in the cell and for maximal survival and efficient functioning of the entire organism.

Evolution. Practically all living cells function in a fluid environment. This includes isolated unicellular forms, such as paramecia and amebas, as well as cells that make up tissues in air-breathing terrestrial animals. Thus, the ionic composition and osmotic activity of extracellular fluids have biological significance with respect to survival of the living cells. The fluid environment of simple unicellular forms of life consists of the oceans, lakes, or streams in which the forms are found, while that of the complex animal forms consists of the fluid media enclosed by the various compartments of the body.

Presumably, life originated in the primitive ocean, the salinity of which was lower in prehistoric times than in modern times. The first unicellular forms of life remained in a state of osmotic equilibrium with their dilute seawater surroundings; that is, they had no osmoregulatory devices. In the course of evolution, unicellular and multicellular animals migrated from the sea to freshwater streams and eventually to dry land. Survival of such forms was associated with the maintenance of constant osmotic activity of their body fluids through evolution of osmoregulatory devices. *See* PREBIOTIC ORGANIC SYNTHESIS.

Changing osmotic conditions occasioned the evolution of special cells and tissues which permitted retention or exclusion of the proper amount of water and solute for the animal. Examples of these special tissues are the skins of amphibians, the gills of fishes, and the kidney tubules of mammals. In cells of these tissues electrolytes apparently diffuse across one cell membrane and are pumped out across the opposite membrane of the cell. *See* EXCRETION.

Biological mechanisms. The actions of osmoregulatory mechanisms are, first, to impose constraints upon the passage of water and solute between the organisms and its surroundings and, second, to accelerate passage of water and solute between organism and surroundings. The first effect requires a change of architecture of membranes in that they become selectively permeable and achieve their purpose without expenditure of energy. The accelerating effect, apart from requiring a change of architecture of cell membranes, requires expenditure of energy and performance of useful osmotic work. Thus substances may be moved from a region of low to a region of higher chemical activity. Such movement can occur in opposition to the forces of diffusion of an electric field and of a pressure gradient, all of which may act across the cell membrane. It follows that there must be an energy source derived from the chemical reactions of cellular metabolism and that part of the free energy so generated must be stored in molecules driven across the membrane barrier. Active transport is the modern term for such processes. The manner whereby chemical energy can be transferred into osmotic work has not yet been determined. Some examples that illustrate osmoregulation follow.

Fresh-water fish and frogs can pump water out of their bodies via the urine. This process enables these animals to survive in dilute fluids. The salt-water fish pumps salt from its body to the sea across the gills. This removes the salts of ingested ocean water, thus permitting survival with brine as the source of water. Of interest is the absence of glomeruli in the kidneys of some marine forms like the toadfish and the goosefish. Such aglomerular kidneys excrete smaller volumes of urine thando glomerular kidneys and

Summary of osmotic performance in various animals*

Osmotic characteristics	Principal mechanisms	Examples
Osmotic adjustment	No volume regulation	Marine invertebrate eggs; *Phascolosoma*
	Volume regulation	Marine mollusks; *Maja; Nereis pelagica; N. cultrifera*
Limited osmoregulation	Low permeability; salt reabsorption in nephridia (?)	*Nereis diversicolor*
	Water storage	*Gunda*
Fair osmoregulation in hypotonic media	Selective absorption of salts from medium; kidney reabsorption or secretion; low permeability	*Carcinus*
Regulation in hyper- and hypotonic media except at extremes	Unknown	*Uca*
Unlimited regulation in hypotonic media	Hypotonic copious urine; salt reabsorption or water secretion; low surface permeability	Crayfish; fresh-water teleosts, Amphibia
	Water impermeability	Fresh-water embryos
Maintenance of hypertonicity in all media	Urea retention	Elasmobranchs
Regulation in hypertonic media	Extrarenal salt excretion; low water intake	Marine teleosts
	Unknown	*Artemia*
Regulation in moist air	Low skin permeability; salt absorption from medium; salt reabsorption in kidney	Earthworm; frog
Regulation in dry air	Impermeable cuticle; hypertonic urine	Insects
	Hypertonic urine, water reabsorption in kidney	Birds and mammals

*After C. L. Prosser et al. (eds.), *Comparative Animal Physiology,* 2d ed., Saunders, 1956.

possess, in their tubular cells, transport systems for the excretion of salt. The albatross and penguin, by secreting an extremely hypertonic nasal fluid, can survive with seawater as the sole source of water, a unique biological advantage for a terrestrial animal. *See* KIDNEY; SALT GLAND.

The **table** presents a zoological classification of various modes of osmotic defense in several phyla of animals. Efficient osmoregulation is accomplished largely by way of the inherent properties of the cell membranes: (1) The hydrophobic nature of the bilamellar phospholipid-protein array of most membranes constitutes an effective barrier to the free diffusion of polar substances (water and water-soluble solutes); (2) simultaneously, however, most membranes actually facilitate the passage of certain substances into and out of the cell, thus providing selective permeability; and (3) many membranes can actually actively transport, or "pump," certain solutes (for example, Na^+ or glucose) into or out of the cell, which is an important facet of the osmoregulatory process.

In multicellular organisms such mechanisms are integrated into osmoregulatory organ systems under the control of chemical, nervous, and hormonal stimuli. *See* URINARY SYSTEM.

Transport Processes

A stringent definition of active transport of ions requires that a net amount of material be moved unidirectionally across a biological membrane against the forces of diffusion, against the forces of an electrical field, and even against the force of a hydrostatic pressure gradient. For individual ions such a movement is said to be against the electrochemical potential gradient. An equally stringent definition applied to water transport would require movement of water against gradients of hydrostatic pressure and of its chemical potential. In the most general sense active transport of a substance means its movement against its free-energy gradient; therefore, the process requires a source of free energy from cellular metabolism. Exactly how metabolic energy is funneled into a cellular mechanism capable of doing osmotic work is unknown.

The term active transport is generally reserved for the pumping of solutes across membranes. Active transport involves a chemical reaction between the transported solute and some constituent of the membrane. In contrast, the osmotic flow of water through cell membranes is a passive process and does not involve strong chemical interactions. However, during osmoregulation by organs such as the kidney or intestine, water can be moved across the epithelial cell membranes from a lower to a higher chemical concentration of water. This implies that water transfer across osmoregulating epithelial membranes is active—as would be required for an osmoregulating rather than for an osmoaccommodating process. However, the metabolic energy source of osmoregulatory transport of water is derived from the active transport of solute across the membrane. It is generally accepted that the actively transported solute, and not the water, reacts with the membrane constituents. This implies that all transmembrane movement of water is primarily passive and secondary to local osmotic gradients and hydrostatic pressures induced by solute pumping.

Carrier systems. In the majority of cases the actively transported substances involved in osmoregulation are univalent inorganic ions, which are supposed to react with membrane constituents called carriers. Apart from attempts to identify and isolate such carriers, there are good reasons for believing that they exist. To account for active transport in the absence of physical forces, it has been necessary to invoke chemical reactions with carrier substances in the membrane. However, chemical reactions are by their nature scalar, and the rate of reaction depends only upon the concentration of the reactants and products and not upon the spatial location of the reactants and products. In contrast, transport processes are vectorial in nature and result in the translocation of substrate from one place to another. Thus the vectorial property of the transport process must be ascribed to a physical property over and above the chemical process in the membrane. There are two kinds of physical processes which could account for the directionality of transport. These can be classified as moving site models and as mobile carrier models.

Moving site model. This can be represented as an aqueous filled pore, the walls of which are lined with sites having a high affinity for the transported ion. The ion to be transported is first absorbed to the site adjacent to the edge of the membrane. The occupied site then undergoes a decrease in affinity for the transported substance, and the ion is transferred to the next site of high affinity deeper in the membrane. These sequential changes in the affinity of the adsorption sites for the transported ion continue from one side of the cell membrane to the other, in a wavelike manner, thus imparting directionality to the transport process. While the moving site model is adequate, there is no experimental evidence for its existence.

Mobile carrier model. This requires that the transported ions combine with a limited number of mobile carrier molecules and that they then be carried across the cell membrane in the form of carrier-ion complexes. The main tenets of this theory are that (1) the carrier molecule is free to move about in the cell membrane; (2) the range of motion of the carrier molecule is limited to the interior of the cell membrane; (3) the carrier molecule undergoes a change in affinity for the transported ion at one edge of the cell membrane; and (4) the cell membrane is relatively impermeable to the free ion being transported.

The precise mode of motion of the carrier molecule is not known. It could undergo free diffusion; swing across the cell membrane in a pendular motion; revolve about an axis in the plane of the cell membrane; or invert along an axis of symmetry normal to the plane of the membrane.

The experimental evidence for the existence of carrier molecules rests on kinetic studies of the transport of isotopes. Specifically, the kinetic phenomena

of transport are those of saturation kinetics cis inhibition, trans stimulation, and counterflow.

The term cis inhibition describes the following sequence of events. A biological membrane is transporting a mixture of ^{22}Na and ^{23}Na having a given specific activity from left to right at a constant rate. Additional Na^{23} is added to the bathing fluid on the left side of the membrane, reducing the specific activity of the mixture and resulting in a decrease in the absolute rate of transport of ^{22}Na from left to right. Thus the term cis inhibition arises because addition of one isotope of an ion inhibits the transport of another isotope of the same ion when the addition is made to the same side as that from which the ion is being transported.

Trans stimulation is said to occur when addition of one isotope of a transported ion, to the side toward which the transported ion moves, causes an increase in the absolute rate of transport of another isotope of the same ion toward that side.

It has been shown on theoretical grounds that the two kinetic phenomena of cis inhibition and trans stimulation are peculiar to systems where the membrane penetration process is carrier-mediated.

Further important information on carriers has been obtained in certain strains of bacteria which can transport various galactosides, sugars, amino acids, and sulfates. Proteins isolated from the membranes of these bacteria bind specifically to the transportable substrate. Nontransporting mutants of these bacteria do not yield specific binding proteins but can be induced to do so.

In biological systems water flow is usually associated with the transport of ions or sugars. There are four general schemes for the transport of ions: (1) the oxidation-reduction pump; (2) the adenosine triphosphate-driven pump; (3) forced ion exchange; and (4) ion-pair transport. The following is a brief description of these schemes.

Oxidation-reduction systems. The characteristics of such a scheme are that there is an oriented electron-transporting mechanism in the membrane; the reaction results in the movement of a given ion, at a single site from the region of low to a region of high electrochemical activity, with generation of an electromotive force; and the electric field so generated transports another ion of opposite charge in the same direction. Alternatively, the field could move a similarly charged ion, at a separate site in the membrane, in a direction opposite to that of the primary ion movement. In either case the laws of electroneutrality are satisfied for the whole system, while osmotic work is done on the transported particles. The active transport of sodium ion (Na^+) with the passive transport of chloride ion (Cl^-) across the frog skin can be described in a stepwise manner by using the assumptions of an oxidation-reduction mechanism.

Figure 1 illustrates the operation of an oxidation-reduction mechanism designed for the active transport of Na^+ ions. The reaction in the membrane is oriented by means of unspecified constraints at the interface across which the actively transported ion is ejected. Thus Na^+ ion enters the cell membrane by combining electrostatically with a carrier substance X^- to yield undissociated NaX. The undissociated complex diffuses across the membrane, wherein the oriented reaction is between cytochrome-oxidized and cytochrome-reduced. This results in the reaction shown in reaction (1), and the free Na^+ ion formed

$$NaX + Cyt^{3+} \rightarrow Na^+ + X \text{ neutral} + Cyt^{2+} \qquad (1)$$

is ejected from the cell as the electron moves in the opposite direction and reduces the oxidized cytochrome. In order to satisfy electroneutrality, a negative ion (Cl^-) must be transported at a site spatially separated from that of cation transport. (There is evidence in some systems that Cl^- is also actively transported, but active transport of the companion ion is not inconsistent with an electrogenic model.) A separate electron donor substance must be present to convert X neutral to X^- ion, a process requiring metabolic energy. Although never proven rigorously, the assumptions of this scheme may be used to account for many osmoregulatory processes in cells across which a measurable electric field exists. While the scheme depicted in Fig. 1 is adequate for an explanation of sodium transport, it is not a unique oxidation-reduction system. Redox pairs other than cytochrome could be employed to drive the transport. *See* BIOLOGICAL OXIDATION.

ATP systems. An enzyme called sodium-potassium-ATPase (Na-K-ATPase) has been isolated in the microsomal, or membrane fraction, of many tissues that transport sodium, for example, frog skin, toad and turtle bladders, red cells, and brain. The enzyme hydrolyzes the high-energy phosphate bond of adenosine triphosphate (ATP), producing adenosine diphosphate (ADP), inorganic phosphorus, and energy. *See* ADENOSINE DIPHOSPHATE (ADP); ADENOSINE TRIPHOSPHATE (ATP).

The cell stores much of its metabolic energy in the ATP, and the hydrolysis of ATP supplies sufficient free energy to drive an oriented transport mechanism in the membrane as well as many chemical reactions in the cell. The remarkable features of Na-K-ATPase are that it is found only in microsomes; it is the only known ATPase requiring sodium and potassium ions for activation; and it is the only

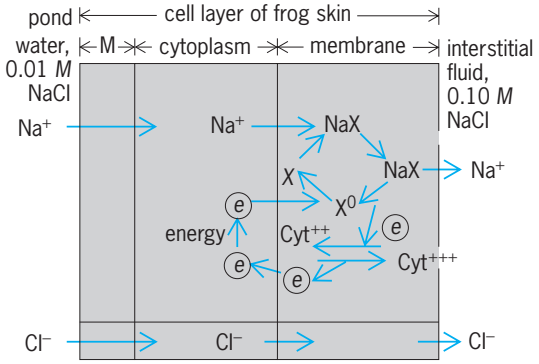

Fig. 1. Transport of Na^+ and Cl^- from pond water across the frog skin to the interstitial fluid. (*After R. F. Pitts, Physiology of the Kidney and Body Fluids, 2d ed., Year Book Medical Publishers, 1968*)

ATPase whose hydrolytic activity is inhibited by the cardiac glycoside, ouabain—which appears to be a specific inhibitor of sodium transport. Whether Na-K-ATPase acts as a carrier and combines directly with sodium during the transport process or whether Na-K-ATPase is involved only in the transfer of energy to the transport machinery is not clear. However, it is obvious that Na-K-ATPase plays some necessary role in sodium transport. In addition, Na-K-ATPase is the first enzyme identified that appears to play a direct role in the transport of an inorganic ion.

Forced ionic exchange. Some investigators believe that electrical potentials across cell membranes are merely diffusion potentials secondary to the ionic gradients produced by a carrier-type active transport. They postulate the existence in the membrane of a charged substance, say X^-, in the case of cation transport. At one interface the carrier of the mechanism ejects Na^+ ion from the cell and simultaneously injects K^+ ion into the cells. This is called a forced ionic exchange process and produces no electrical potential. At another site of the membrane there are water-filled pores, the walls of which are lined with fixed negative charges. Hence, electropositive ions can enter such pores more readily than can electronegative ions. Therefore the potential difference across the membrane interface (at equilibrium) ought to be predicted by the Nernst equation, shown as Eq. (2), where E is the potential difference, R and

$$E = \frac{RT}{F} \ln \frac{a_2}{a_1} \tag{2}$$

T are the gas constant and absolute temperature, F the Faraday and a_1 and a_2 are the chemical activities of the passively transported ion on each side of the membrane. This means that the transmembrane distribution of ions produced by the forced exchange mechanism results in the production of diffusion potentials across the negatively charged pores. According to this scheme most of the observed potential difference is due to the high intracellular concentration of potassium. The absence of a diffusion potential from the high extracellular concentration of sodium has been attributed to a high specific resistance for this ion across the membrane. A Nernst-type relationship has been observed within a limited range of concentrations of K^+ and Na^+ in the external medium of systems like frog skin and squid axon.

The forced ion-exchange model has most frequently been applied to, and investigated in, systems which transport sodium and potassium. This model has certain difficulties. First, derivations of the Nernst equation tacitly assume equilibrium conditions rather than the steady-state condition that prevails in most living systems. Second, the negatively charged pores must select not only between negative and positive ions but also between sodium and potassium. Third, recent isotopic measurements indicate that the ratio of sodium to potassium transport is often not 1 to 1, as would be expected in such a system.

Ion-pair transport. Another mechanism for doing osmotic work is the transport of salts as ion pairs. That is, a carrier could transport neutral sodium chloride molecules. Such a system would do osmotic work, would be nonelectrogenic, and would not necessarily give rise to electrical potentials. Such a transport system has been demonstrated in the proximal renal tubule and in the gallbladder of fish and mammals. *See* ION TRANSPORT.

Membrane Permeability

Implicit in all theories of transport of water or solutes are assumptions concerning the permeability properties of the membrane. For example, in a Na^+ transporting system the carrier concept implies that NaX will diffuse rapidly across the membrane, while free Na^+ ions diffuse slowly or not at all. This is a device for potentiating diffusion of a sluggishly diffusible substance and is economical from a thermodynamic viewpoint. If the membrane were freely permeable to a transported solute, the rate of back diffusion of that solute would increase as its concentration on the transported side increased. Since concentration gradients across cell membranes are high, back diffusion rates would be high. Thus the mechanism would require a large amount of energy to accomplish the net transport of a solute. These and other considerations led to much experimental and theoretical work from which emerged concepts of membrane permeability. *See* CELL MEMBRANES; CELL PERMEABILITY.

Molecular architecture. The cell membrane is viewed as a phospholipid-protein complex. Pairs of phospholipid molecules (such as lecithin) in the form of a bilamellar leaflet are oriented so that the ionizable phosphatidyl head groups of the lipid face toward each aqueous boundary fluid, and the fatty acyl chains of each phospholipid pair face toward each other at the hydrophobic interior of the membrane. Protein molecules, mostly in the alpha-helical and random coil conformation and intercalated with the lipids, penetrate through the entire thickness of the membrane.

Despite the advancing knowledge in the molecular structure of cell membranes (based on patterns of x-ray diffraction, electron microscopy, optical rotary dispersion, and circular dichroism), the exact nature of the path through which solute and solvent molecules pass remains uncertain. Two main models of the intramembrane path have been proposed, the lattice and the aqueous-filled pore. *See* DICHROISM (BIOLOGY); ELECTRON MICROSCOPE; X-RAY DIFFRACTION.

Lattice model. Biologists have proposed that the fatty acyl and protein chains of the cell membrane present a latticelike barrier to the passage of molecules. Thus a penetrant molecule could form a hydrogen bond with an atom of the lattice, then shake loose by thermal agitation, move through an interstitial space to the next bonding site, and so on, until it has traversed the membrane. This picture fits the observed rates of penetration of cell membranes by a large number of nonelectrolyte molecules wherein the rate can be

predicted from the molecular size, the number of hydrogen bonding atoms per molecule, and oil-water solubility ratios. However, the model as constituted falls short in explaining penetration of water (osmosis) and of the electrolytes. *See* OSMOSIS.

Aqueous pore model. Assuming the presence of aqueous-filled pores in membranes, the theory of fixed charges on pore walls was formulated. It is known that artificial membranes, such as collodion, silicates, or protein, when interposed between two electrolyte solutions of different concentration, give rise to an electrical potential difference, depending on the sign and the density of fixed charges lining the pores of the membrane. Such membranes have been called perm-selective, that is, cation- or anion-permeable. When a membrane with positively charged pores separates two salt solutions of different concentration, the dilute solution is negative to the concentrated one; for a membrane with negatively charged pores, the reverse orientation holds. The assumption of fixed charges on pore walls has explained a wide variety of data in artificial perm-selective membranes, and useful analogies between artificial perm-selective and biological membranes have been made. However, many biological membranes (nerve and muscle) can distinguish not only cations from anions but one cation from another, for example. Na^+ from K^+, a finding not accounted for by the fixed charges in the aqueous pore model.

Facilitated diffusion path. Some biologically important materials, such as ions, sugars, and amino acids, move across cell membranes more rapidly than can be explained by the parameters of a lattice or fixed-charge path. Therefore, the transmembrane motion has been explained by assuming that the penetrant combines transiently with a mobile constituent molecule of the membrane. These processes of combination with the carrier, translocation of the loaded complex, and dissociation of penetrant from the carrier are termed facilitated diffusion.

Phenomenological equations. Due to the lack of knowledge of the exact mechanisms whereby materials cross biological membranes, a theoretical approach which circumvented the need for invoking specific mechanisms of motion was applied. The approach of irreversible thermodynamics originated in the work of L. Onsager and was applied to biological systems by C. Kedem and A. Katchalsky. Essentially, a set of simultaneous equations, called phenomenological equations, is used, based upon well-known empirical laws of motion, such as those of A. Fick, J. Fourier, and G. S. Ohm.

It is assumed that all flows can be represented by a system of linear equations in which each flow is set equal to a linear combination of forces multiplied by appropriate coefficients; that is, each flow is taken to be proportional to each of the forces and the system of flows and forces can be represented in matrix form as Eq. (3). Here J_i are the flows, L_{ik} are

$$J_i = \sum_{k=1}^{n} L_{ik} X_k \qquad (i = 1,\ 2,\ 3, \ldots, n) \qquad (3)$$

the coefficients, X_k are the conjugated forces, and the subscript i denotes the row, and k the column position of each term (J_i, L_{ik}, and X_k) in the matrix of simultaneous equations.

To illustrate, by an oversimplified example, consider a nonphysiological problem. Suppose that two sucrose solutions are separated by a rigid membrane and that an osmotic or a hydrostatic pressure can be applied to either one of the two solutions. Since there are but two mobile components in the system, a minimum of two flows—one of sucrose and one of water—all will be considered, and the aforementioned matrix, in algebraic form is expressed as Eqs. (4*a*) and (4*b*), where J_w is the flow of water, J_s

$$J_w = L_{11}\Delta P + L_{12}RT\Delta C \qquad (4a)$$

$$J_s = L_{21}\Delta P + L_{22}RT\Delta C \qquad (4b)$$

is the flow of sucrose, ΔP is the hydrostatic pressure difference across the membrane, ΔC is the difference in the sucrose concentration in the two solutions, L_{11} is the hydraulic permeability of the membrane to water, L_{22} is the permeability of the membrane to sucrose, and L_{12} and L_{21} are cross coefficients representing the influence of the concentration difference on water flow and the influence of hydrostatic pressure difference on sucrose flow. In this case ΔP and ΔC are set at known values experimentally, J_w and J_s are measured, and the four L's which operationally describe the membrane are unknowns.

Onsager has shown on thermodynamic grounds that $L_{21} = L_{12}$, which effectively reduces the unknown to three in number, namely, L_{11}, L_{22}, and L_{12} or L_{21}. By successively forcing ΔP, ΔC, and J_w to be zero in the experiment, it is possible to obtain three equations in three unknowns and solve the system for the L's. In more complicated situations, with more than two flows, more than three experiments are needed to evaluate all of the unknowns, but the method is the same. The advantage of this system of analysis is that it is not necessary to understand the basic mechanisms to evaluate the permeability of the membrane with respect to a given permeant. The disadvantage is that results from one membrane cannot be generalized to another because the method is essentially empirical.

Mechanisms of Water Transport

The forces for water transport in living and in nonliving systems may be listed as follows: hydrostatic forces; osmotic gradients; electroosmosis; and miscellaneous—chemical reactions, thermal gradients, and unknown forces in contractile vacuoles and pinocytosis.

In animals these forces are usually generated by the active transport of solutes and by the circulatory system. When such forces are applied across membranes having perm selective and ion selective properties, such as are possessed by cell membranes, they give rise to the fundamental processes involved in osmoregulation. These processes are filtration, osmosis, diffusion, and maintenance ofGibbs-Donnan

equilibria. In addition to the four simple physical mechanisms, there are three main biological systems which integrate one or more of the simple mechanisms into a device for moving water. They are the three-compartment system, the countercurrent system, and the standing-gradient system. These systems incorporate the processes of filtration, osmosis, diffusion, and the active transport of ions, with multicellular structures of specialized geometric form designed for osmoregulatory functions.

Hydrostatic forces. Operationally, flow of fluid through a tube is usually defined in terms of two main parameters: pressure and resistance to flow. It follows that a bulk flow of fluid across any boundary requires a force and that the direction of such flow is down the gradient of pressure. Filtration refers to the passage of water or solution under the influence of a hydrostatic force through the pores of a membrane. A filtration process separates undissolved constituents from a solution, while ultrafiltration separates dissolved constituents of high molecular weight (protein) from a solution.

In filtration a net movement of a finite mass of solute and solvent occurs, while in diffusion there is no net movement of solvent. By definition, the only force operative in a diffusion cell is that of the gradient of chemical potential of the transported solute material. H. Ussing made use of such differences when he measured permeability coefficients of various biological membranes with and without net movement of solution.

Gibbs-Donnan forces. Ultrafiltration refers to the passage of solute and water across a membrane between two nonidentical solutions in a system like that conceived by J. Gibbs and F. Donnan. The simplest example of such a system is shown below.

Na_1^+ Cl_1^- Prot$^-$		Na_2^+ Cl_2^-
(1)	Membrane	**(2)**

Figure 2 presents an experimental model of the system at equilibrium. Conditions are such that the membrane permits passage of Na^+, Cl^-, and water but not of proteinate. At equilibrium it is possible to derive Eq. (5).

$$[Na_1^+][Cl_1^-] = [Na_2^+][Cl_2^-] \qquad (5)$$

Since $[Na_2^+] = [Cl_2^-]$, it follows that $[Na_1^+] + [Cl_1^-] > [Na_2^+] + [Cl_2^-]$. This means that a hydrostatic pressure must be applied to side I to prevent the osmotic flow of solvent from II to I. The osmotic conditions at equilibrium (no net flow of solvent or solute) are expressed in Eq. (6), where P_I and P_{II} are

$$[Na_1^+] + [Cl_1^-] + [Prot^-] = [Na_2^+] + [Cl_2^-] + \frac{P_{II} - P_I}{RT} \qquad (6)$$

the hydrostatic pressures of compartments I and II, respectively, R is the gas constant, and T is the absolute temperature. Ultrafiltration will occur when the pressure applied to solution I is sufficient to move

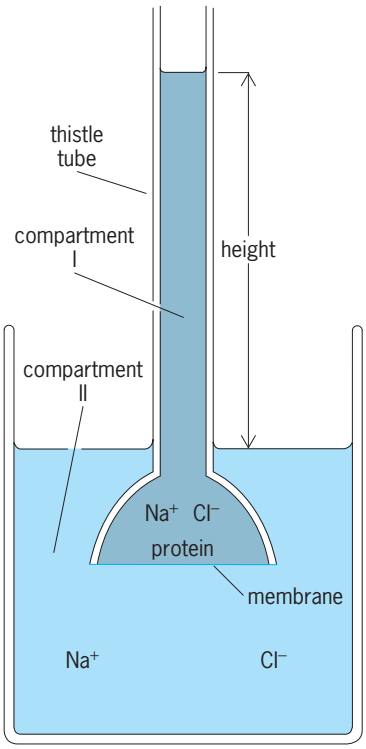

Fig. 2. Simple model illustrating a Gibbs-Donnan equilibrium system. (*After W. D. Stein, Theoretical and Experimental Biology, vol. 6: The Movement of Molecules Across Cell Membranes, Academic Press, 1967*)

the solution from I to II, that is, in a direction opposite to that of the osmotic gradient. *See* DONNAN EQUILIBRIUM.

Filtration across capillaries. Examples of hydrostatic movement of water or solution in biological systems may be found in capillary beds, including that in glomeruli of kidneys. As blood is pumped by the heart into the arterial side of a capillary system, the hydrostatic pressure head pushes an essentially protein-free filtrate of the blood plasma across the capillary wall into the interstitial fluid. When the blood reaches the venous side of the capillaries, the blood pressure has been dissipated and, as predicted from the balance of Donnan forces and of hydrostatic pressure forces, an osmotic flow of fluid occurs from the interstitial fluid, across the capillary wall, and into the venous capillary. In the steady state the amount of fluid filtered equals the amount of fluid reabsorbed from the capillaries. This is a skeletal description of the classic concept of capillary function patterned mainly after the work of E. Starling and E. Landis.

The aforementioned considerations apply not only to all animals possessing a vascular system but to any cellular forms containing protein-rich cytoplasm within a plasma membrane surrounded by a protein-free fluid environment. This criterion can be applied to unicellular animals and to multicellular animals, as well as to the tissues of practically all invertebrates and vertebrates, aquatic and terrestrial. In animal cells hydrostatic pressures are small (about 0.4–2.0 in. Hg or 10–50 mmHg) but sufficient to balance osmotic differences occasioned by the Donnan

forces operative across the plasma membrane. However, in plant cells the hydrostatic pressure can be relatively tremendous (15–30 atm), owing to the tough cellulose wall encasing the plasma membrane.

Osmotic gradients. The presence of osmotic pressure differences across porous membranes is considered to be one of the most important factors causing a net osmotic flow of water in biological systems. The usual assumption made for the system is that the membrane is permeable to solvent but not to solute. Such a requirement for the membrane is approximated in highly selective cation- or anion-permeable membranes. However, most living membranes, such as amphibian skin, nerve, eggs, erythrocytes, stomach, intestine, bladder, capillaries, and even mitochondria, appear to be permeable to solutes of small molecular weight—that is, of molecular weights up to at least 2000—as well as to water. This means that the water activity of cellular fluid tends to equilibrate rapidly with that of extracellular fluid so that appreciable osmotic differences between the phases rarely exist. Despite wide differences in chemical composition between cellular and extracellular fluid, there exist no measurable differences of water activity. Water activity has been evaluated by the usual measures of colligative properties such as freezing-point depression, melting point, or vapor-tension lowering.

The aforementioned remarks are not so general as they might appear, because large osmotic pressure differences are present, or appear to be present, across membranes in many biological systems. All of the osmoregulating forms present such osmotic pressure differences.

Examples of systems with apparent osmotic gradients are [body fluid—gills—pond water] in freshwater fish, crabs, and worms; [body fluid—renal tubular cell—urine] in kidneys producing urine either hypertonic or hypotonic to the body fluids; and [soil water—cell wall—cytoplasm] in plants.

According to van't Hoff's law, the osmotic pressure developed across an ideal membrane is $22.4\Delta C = \pi$, in atmospheres, where ΔC is the concentration difference in the nonpenetrating solute on the two sides of the membrane. However, if the membrane is leaky and allows some of the solute to cross from one side to the other, the stopping pressure will be less than $22.4\Delta C$. The ratio of the measured stopping pressure to that predicted by van't Hoff's law is called the Staverman reflection coefficient, σ. Hence, for real membranes the stopping osmotic pressure will be predicted by $\pi = \sigma \times 22.4\Delta C$, where σ has a value between 1 and 0.

Franck-Mayer hypothesis. The steady-state maintenance of osmotic gradients across cells has been a biological problem for years. The magnitude of some gradients—such as the urine of a hydropenic dog, 1500 milliosmoles per kilogram (mOsm/kg), and plasma, 300 mOsm/kg—is too great to be explained by hydrostatic pressure gradients, except in plant cells. This led to the Franck and Mayer hypothesis of osmotic gradients. Apart from maintenance of gradients, their hypothesis included a mechanism

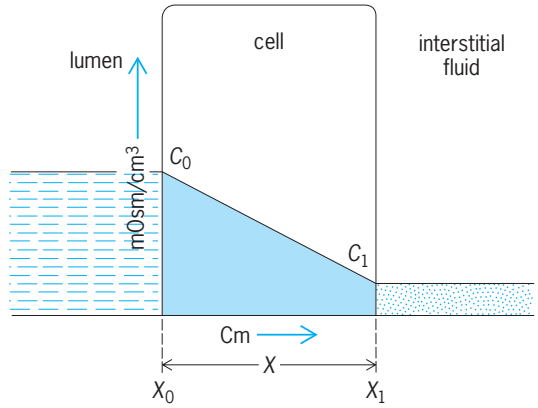

Fig. 3. Scheme of an intracellular osmotic gradient in a kidney tubule during production of osmotically concentrated urine. Osmotic activity *C* is plotted along the ordinate, and cell thickness *X* is plotted along the abscissa. (*After C. A. Giese, Cell Physiology, Saunders, 1959*)

for transporting water (solvent) from a solution of high osmolarity to a solution of low osmolarity; that is, the mechanism could move water up its gradient of activity.

Figure 3 illustrates the essentials of the Franck-Mayer scheme as applied to the process of formation of hypertonic urine. When the intracellular osmotic activity C_0 at the lumen side of a renal tubular cell is slightly greater than that of the luminal fluid (urine) and the intracellular osmotic activity at the interstitial side of the cell C_1 is equal to that of the interstitial fluid, water could be transported from the lumen to the interstitial fluid. The postulated source of solutes at C_0 was a depolymerizing reaction, enzymatically catalyzed, whence a large molecule was split into numerous small particles at the interface. Thermodynamically, the minimum free energy expenditure for maintenance of the gradient by the mechanism, regardless of its origin, must be at least equal to the heat dissipation or to the decrease of free energy of solute diffusing from X_0 to X_1. The number of particles Q diffusing across an area A of cell in unit time for steady-state conditions is expressed as Eq. (7), where

$$Q = \frac{-DA(C_1 - C_0)}{X_1 - X_0} \tag{7}$$

D is the diffusion coefficient and $\Delta X = X_1 - X_0$, the cell thickness of path legend for diffusion. The equation is an integrated form of Fick's equation with the implicit assumption of a flat sheet of cells and zero bulk flow. The rate of decrease of free energy for the diffusion is expressed as Eq. (8), where ΔF is

$$\frac{\partial \Delta F}{\partial t} = QRT \ln \frac{C_0}{C_1} \tag{8}$$

the change of free energy, t is the time, R is the gas constant, and T is the absolute temperature. Since diffusion is an irreversible process, the free-energy loss cannot be funneled back into the transporting or gradient-creating mechanism. An evaluation of the minimum rate of expenditure of free energy, made by A. Brodsky, B. Rehm, C. Dennis, and D. Miller, in the case of mammalian renal cell yielded a value of

87.8 kilojoules/(kg)(h) or 21,000 kcal/(kg)(h). This is about 1000 times the maximal rate of respiration for living cells and imposes a severe limitation on the theoretical use of intracellular osmotic gradients as devices for water transport in biological systems. A major limitation of the Franck-Mayer scheme is the short diffusion path—one cell width, or about 20 micrometers.

Multicellular organisms are often required to maintain substantial osmotic gradients between various body fluids and secretory fluids. It is sometimes necessary for the organism to move water from a concentrated to a dilute solution against the chemical potential gradient for water. Moreover, these tasks must be accomplished without an inordinate expenditure of energy.

Two biological systems which move water against concentration gradients without too much energy expenditure are the three-compartment, two-membrane system and the countercurrent system.

Three-compartment, two-membrane system. This consists of three successive fluid compartments separated by two successive osmotic membranes. One of the two membranes of the system is a tight osmotic membrane, and the other is a leaky osmotic membrane.

Figure 4 is a schematic representation of a three-compartment, two-membrane system. The figure could represent a mucosal cell separating the luminal fluid in the intestine from the interstitial or blood side of the intestinal membrane. In the figure the luminal fluid is represented as being hypertonic to the plasma by 50 mOsm, and the cell fluid as being hypertonic to the luminal fluid by 10 mOsm and hypertonic to the interstitial fluid by 60 mOsm. The net direction of water movement is from hypertonic luminal fluid to the interstitial fluid, against the overall concentration gradient for water.

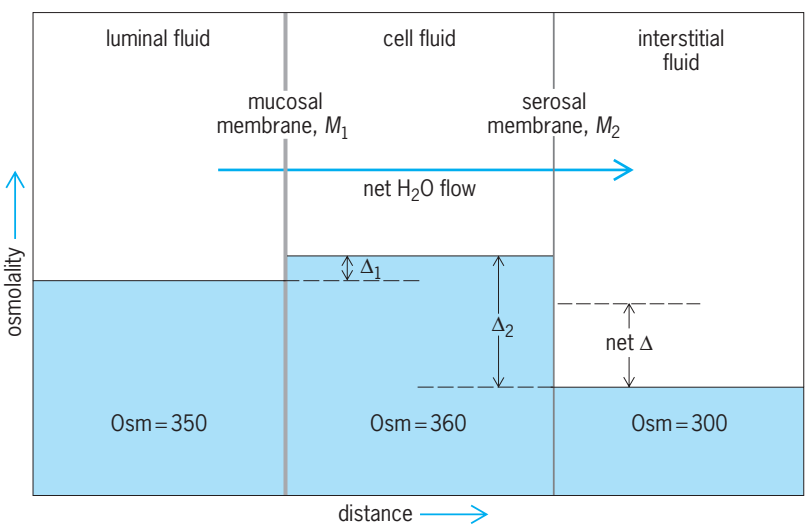

Fig. 4. Osmolality (Osm) versus distance in a three-compartment, two-membrane system. Cell fluid bounded by membranes separates the hypertonic luminal fluid from the interstitial fluid. Cell fluid is hypertonic to both of the other fluids. Magnitudes of osmotic gradients across membranes M_1 and M_2, indicated by Δ_1 and Δ_2, are 10 and 60 mOsm, respectively. The magnitude of the overall gradient between luminal and interstitial fluids, indicated by net Δ, is 50 mOsm. (*After A. S. Crafts, H. B. Currier, and C. R. Stocking, Water in Physiology of Plants, Ronald Press, 1949*)

The system operates as follows: Water (with practically no solute) is pulled across M_1 into the cell by the transmembrane osmotic gradient. This causes an increased hydrostatic pressure within the cell fluid. The intracellular pressure developed is large enough to force water and solute through M_2 in the direction opposite to the osmotic gradient. Thus the water and solute cross the leaky membrane M_2 by a process of pressure-ultrafiltration with some sieving. In other words, water moves across M_1 by an osmotic process and across M_2 by a pressure-filtration process. Nevertheless, the system can be analyzed in terms of two opposing osmometers, one more efficient than the other in separating water from solute.

Now consider the system as two opposing osmometers, one composed of [luminal fluid—membrane (M_1)—cell fluid] and the other composed of [cell fluid—membrane (M_2)—interstitial fluid]. The degree of sieving, or the Staverman reflection coefficient of the tight osmotic membrane M_1, is larger than that for the leaky osmotic membrane M_2; that is, $\sigma_1 > \sigma_2$. The osmotic pressure developed across a membrane is $\Delta\Pi = \sigma 22.4\Delta$ Osm. Thus the net driving force on the water would be expressed as Eq. (9),

$$\Delta\Pi_{net} = 22.4[\sigma_1(\mathrm{Osm}_{cf} - \mathrm{Osm}_{if}) - \sigma_2(\mathrm{Osm}_{cf} - \mathrm{Osm}_{isf})] \quad (9)$$

where σ_1 or σ_2 are the Staverman reflection coefficients of the membranes M_1 and M_2, respectively, and the subscripts on the osmolarities refer to fluid in the lumen fluid (lf), cell fluid (cf), and interstitial fluid (isf).

In the example of Fig. 4, the values for the sigmas must be such that $10\sigma_1 > 60\sigma_2$ in order to cause a net movement of water to the right.

The flow of water through the cell will carry solutes through the leaky membrane M_2 and cause the cell fluid to become diluted. In order to keep the process going there must be a pump which transports the solute into the cell from one side or the other. By choosing the proper location for the solute pump and the proper values for σ_1 and σ_2 in such a model, water or solute, or both, can be made to move in either direction.

Given that the three-compartment system contains two or more solutes (as is the case in biological systems), it can be shown that water will move against the overall osmotic gradient from compartment two to compartment three even in the absence of any intracellular hydrostatic pressure gradient.

Countercurrent system. An ingenious model for moving water economically against osmotic gradients in the kidney is the so-called countercurrent multiplier system. This scheme was originated by analogy with the well-known physical laws of heat flow in concentric pipes containing counterflowing fluids. The elements of the scheme applied to the mammalian kidney are illustrated in **Fig. 5**. The first step in the production of osmotically concentrated urine is glomerular filtration, which produces an ultrafiltrate of 300 mOsm/liter, isoosmotic with plasma. The filtrate in the proximal tubule suffers a reduction of

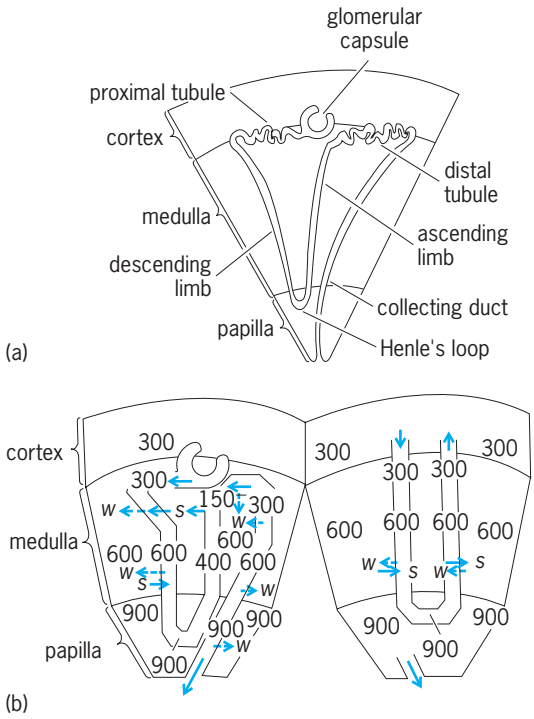

Fig. 5. Osmotic activity in the human kidney. (*a*) An anatomical reference diagram. (*b*) Schematic presentation of a renal countercurrent system for production of an osmotic gradient in the interstitial fluid and consequently for production of urine of osmotic activity greater than that of plasma. Numbers indicate osmotic activity in milliosmoles per liter; *s* = direction of the diffusion of salts; and *w* = direction of diffusion of water. (*After A. Krogh, Osmotic Regulation in Aquatic Animals, Peter Smith Magnolia, 1939*)

volume without change of osmotic activity. Next the fluid enters the descending limb of the loop of Henle, where it suffers further reduction of volume but increases in osmotic activity by equilibrating with hypertonic interstitial fluid. This raises the questions of how hypertonic interstitial fluid is produced and how the osmotic gradient from cortex to papilla is maintained.

It has been shown that the ascending limb of the loop of Henle contains a salt pump which can pump salt from the lumen to the interstitial fluid and that this region of tubule is impermeable to water. Such a pump would tend to produce a region of hypertonicity in the interstitial fluid. The maintenance of the gradient along the long axis of the renal pyramids under steady-state conditions is occasioned by the countercurrent design of the vasa recta (postglomerular branches of arterioles from glomeruli in the subcortical regions of the kidney). As the descending part of the vasa recta plunges into the hypertonic papilla, salts diffuse into the blood, concentrating it toward the hypertonic level of papillar interstitial fluid. As the ascending branch of the vasa recta emerges from the papilla, it moves salt-laden blood toward the cortex, a region isotonic with peripheral blood. Therefore, salts diffuse from the blood toward the interstitial fluid and the osmotic gradient is maintained by what is called a countercurrent exchange system.

The source of energy for this process, referred to as the countercurrent multiplier, consists of the Na^+ pumps along the ascending limb of the loop of Henle. As the hypertonic fluid entering the lumen moves up the water-impermeable ascending limb, it is diluted by the sodium pumps to such an extent that it becomes hypotonic to systemic plasma by the time it reaches the distal tubule. The distal tubular walls are water-permeable; water moves from the lumen to the isotonic interstitial fluid and back of the body via the renal vein. This step is the true water-economy mechanism for the body. A tubular fluid, isotonic with peripheral blood, then enters the cortical portion of the collecting duct. As the collecting duct proceeds toward the papilla, its contents equilibrate osmotically with those of the interstitial fluid. Since osmotic activity of interstitial fluid increases in moving from cortex to papilla, the osmotic equilibration renders the final urine hypertonic to the systemic plasma.

During periods of excess water intake the mammalian kidney elaborates a urine hypotonic to the plasma. Presumably, dilution of body fluids during water ingestion reduces secretion of the antiduretic hormone of the posterior pituitary gland. Absence of this hormone reduces the water permeability of the distal tubules and the collecting ducts. Thus the dilute luminal fluid in these segments does not equilibrate osmotically with the surrounding interstitial fluid.

The biological applications of the countercurrent system have been fruitful for the explanation of diverse phenomena. It has been invoked to account for heat conservation in the extremities or skin of aquatic warm-blooded animals such as seals, whales, or ducks. It is present in gill systems, in swim bladders, or in any system containing or separating two solutions of greatly different osmotic activity. K. Schmidt-Nielson has noted the importance of diffusion length in countercurrent systems, showing that the relative length or renal papillae in desert animals is greater than that in other animals. *See* COUNTERCURRENT EXCHANGE (BIOLOGY).

Standing gradient system. The countercurrent system described above is essentially a standing gradient system. Another type of standing gradient system, less complex than that of the kidney, may be involved in the movement of hypertonic or isotonic solutions across simple epithelial membranes. The mucosal cells of some secretory epithelia at the foot, or the interstitial boundary, of the cell and joined together at the top, or the luminal surface, by short tight junctions. The side walls of such cells are not strongly interdigitated and are not joined together, except at the tight junction at the top of the cell. Thus there is a system of one-ended channels between the cells. The channels, sealed at the luminal end, open into the interstitial fluid at the level of the basement membrane. The lateral walls of the cells are relatively permeable to water and impermeable to ions. Sodium pumps located on the lateral wall of the cell near the tight junction transport sodium from cell fluid into the channel, causing the channel to be filled

with hypertonic fluid at its closed end. As the hypertonic fluid moves down the channel toward the foot of the cell, water moves osmotically from the cell into the channel. As water moves into the channel along the length of its walls, the velocity of fluid flow in the channel increases and the salt concentration decreases toward isotonicity at the open end of the channel. In the steady statee the osmotic profile of the fluid along the axis of the channel changes with distance but not with time. The steady-state gradient is referred to as a standing gradient. Such a system has been proposed by J. Diamond, who has provided a mathematical analysis of the system. The hypertonicity caused by the salt pumps at the head of the channel provides an osmotic force for the movement of the final isotonic fluid against an osmotic or hydrostatic pressure.

Figure 6 is a schematic representation of a standing gradient system. The standing gradient resides in the channel between the two epithelial cells. The lower panel of the figure represents the osmotic profile of luminal fluid—the fluid in the channel and the interstitial fluid. The shaded area at the beginning of the channel represents the zona occludens, or the tight junction between the luminal surfaces of the cells. The thin arrows marked H_2O represent the path of water movement across the epithelial membrane. Sodium chloride diffuses into the cell across the luminal membrane of the cell and is actively transported into the channel near the site marked, P, the NaCl pump.

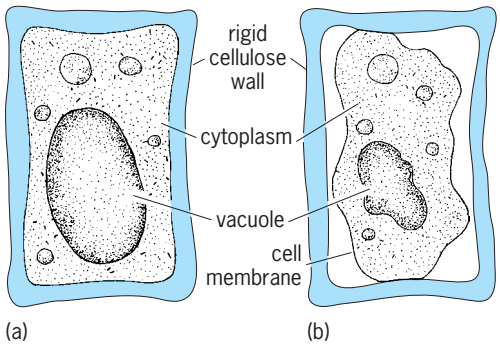

Fig. 7. Osmotic behavior of a plant cell during plasmolysis. (*a*) The plant cell is immersed in an external medium which is hypotonic to the cytoplasm. (*b*) The cell is immersed in an external medium which is hypertonic to the cytoplasm. (*After J. T. Edsall and Wyman, Biophysical Chemistry, vol. 1, Academic Press, 1958*)

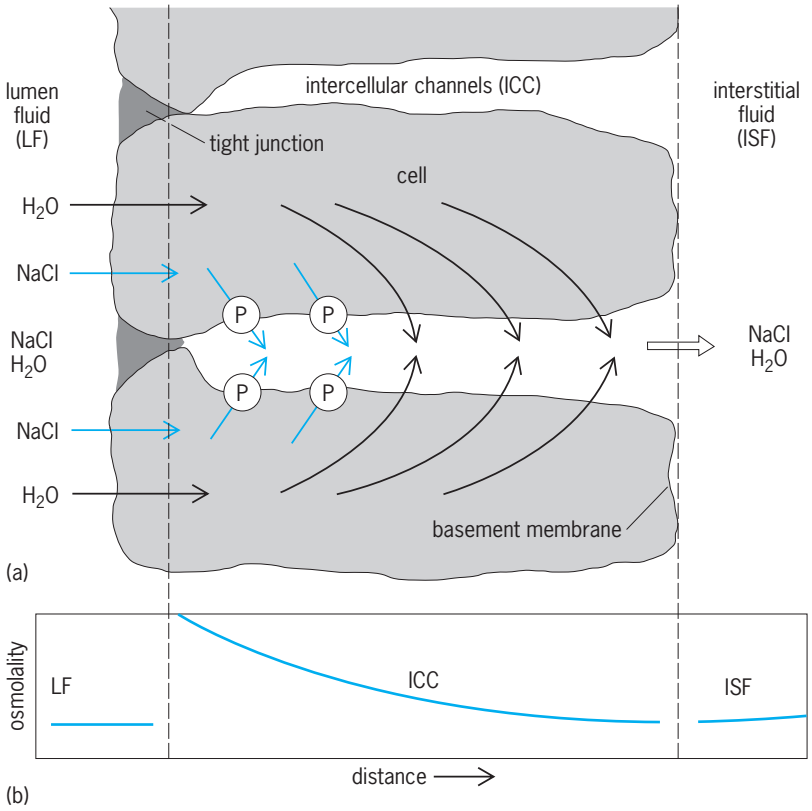

Fig. 6. Movement of an isotonic solution through cells. (*a*) Layer of epithelial cells transporting an isotonic solution of NaCl and H_2O. P = salt pumps. (*b*) Osmotic profile of the standing gradient in the intracellular channels. Vertical broken lines show the spatial relation between the plot and the sketch in *a*. (*After C. L. Prosser et al., eds., Comparative Animal Physiology, 2d ed., Saunders, 1950*)

Osmotic gradients in plant cells. A typical plant in is that found lining the pith, phloem, cortex, or xylem of succulent plant organs. It consists of a cell wall of great tensile strength, cytoplasm, and nuclei. An osmotic gradient between the cell and its surrounding fluid is created by a mechanism transporting salts from the external region of low concentration to the intracellular region of relatively high salt concentration. The mechanism for the transport could be the transmembrane transport of an anion such as nitrate (NO_3^-) at a site separate from that of cation movement. Alternatively, it could be by formation of ion pairs ($NaNO_3$) in the membrane. In either case the osmotic pressure difference across the membrane provides a force oriented to drive water from the exterior to the interior of the cell. Water penetrates into the cell until the internal (turgor) pressure is sufficient to stop the flow. Such turgor pressures are of the order of 10–20 atm (100–200 kilopascals). *See* PLANT CELL; PLANT-WATER RELATIONS.

Plasmolysis. The process (**Fig. 7**) of shrinking the cytoplasm and vacuole by immersing an isolated plant cell in a solution of higher osmotic activity than that of a plant cytoplasm is called plasmolysis. Presumably, no osmotic shrinking or swelling would occur if the cells were immersed in a solution of the same osmotic activity as that of plant cytoplasm. Such data of plasmolysis may not give a true measure of osmotic activity of plant cells because of solute penetration. Nevertheless, a good approximation may be may in cells from which solute leaks very slowly and into which external solute does not penetrate. The data are usually checked by cryoscopic determinations on the juice of pulverized cells. Discrepancies in the results have been found when both techniques have been applied to the same plant tissue.

Electroosmosis. If identical aqueous solutions are separated by a porous membrane containing fixed charges along the walls of the pores and a voltage is applied across the membrane, water will flow across the membrane, against opposing hydrostatic or osmotic forces. If the walls of the pores are lined with negative charges, then some of the positive ions or

H^+ from the dissociated water will be adsorbed to the walls of the pore, leaving the bulk solution in the pore with a net negative charge. Thus, when a voltage is applied across the membrane, water will move toward the positive side of the membrane.

Numerous biological membranes involved in osmoregulatory functions are known to be ion-selective and to have electrical fields across them. Thus electroosmosis may play a role in the movement of water across biological membranes. Electroosmosis and the converse phenomenon of streaming potentials have been observed in the wall of the gallbladder.

Miscellaneous forces. Thermal gradients, if present in living cells, could provide a force for the movement of water. However, the magnitude of temperature gradients is small across most cell membranes and, consequently, the amount of water moved would be small.

Chemical reactions with production or consumption of water would provide a device for water transport. However, the molar quantity of water involved in most known reactions is small compared to the known rates of water transport in animals.

Hydration and dehydration of protein molecules have been invoked as water-transporting processes. It is difficult to tell whether a protein molecule adsorbs pure solvent or bulk solution, and it is even more difficult to see how much a system could drive water in a given direction.

Contractile vacuoles, the structures in the cytoplasm of protozoans such as the paramecia, have been observed extensively while the paramecia were immersed in solutions of varying osmotic activity. The vacuole appears to fill with water, burst, and consequently extrude its contents. The rate of pulsation of the vacuole increases as the external medium is rendered dilute. Vacoular activity depends on metabolic energy, as shown by the suppressing action of cyanide. The mechanism of water transport of contractile vacuoles could be by a secretion of water into the vacuole or by secretion of solute and diffusion of water into the vacuole.

Endocytosis is a term applied to the engulfing of water or solution by pseudopodlike processes of leukocyte membranes. The droplet of water is engulfed by the outer surface of the membrane, wherein it migrates to the inner surface, at which point it is pushed into the cell. *See* ENDOCYTOSIS.

Hormonal control systems. Given the presence of water and solute-transporting mechanisms, the organism requires controlling machinery to govern such processes so that effective osmotic regulation will be maintained. The system in mammals, though complex, has been studied extensively. Therefore, a presentation of the elements of such a system can provide a model that is applicable by analogy to the osmoregulating system of other animal forms. Anatomical elements of the mammalian system are the neurohypophyseal tract of the hypothalamus and the posterior pituitary and the kidney tubule, probably the distal part plus collecting ducts. Osmoreceptor cells, sensitive to small changes of osmolality of the perfusing fluids, are present in the suprahypophyseal tract of the hypothalamus. When these cells respond to osmotic stimuli, impulses are sent through the nerve tract to the posterior portion of the pituitary gland. The pituitary gland responds by increasing or by decreasing the rate of secretion of the antidiuretic hormone (ADH), vasopressin. The hormone, after reaching the blood, is carried to the kidney, where it affects the water-transporting cells of the tubule and collecting ducts.

When the plasma becomes hypertonic (as after dehydration or excessive salt intake), secretion of ADH is increased, causing the walls of the collecting ducts of the kidney to become permeable to water. The fluid in these ducts then equates with the hypertonic papilla of the kidney. Hypertonic urine is produced and free water is returned to the circulation.

When the plasma becomes hypotonic (after excessive water intake), secretion of ADH is reduced and the walls of the collecting duct become impermeable to water; the hypotonic fluid in the collecting ducts fails to equilibrate with the renal papilla. Thus hypotonic urine is produced, whence free water is removed from the circulation.

By virtue of its primary action in the retention or elimination of water per se, vasopressin plays a role in volume control of body fluids. Other hormones, notably the adrenal cortical hormone aldosterone, are involved in volume control. However, their primary action is on the retention or elimination of salts which are followed passively by water. *See* ENDOCRINE MECHANISMS. William A. Brodsky; T. P. Schilb

Bibliography. T. E. Andreoli et al. (eds.), *Membrane Physiology*, 1980; C. J. Duncan and C. R. Hopkins (eds.), *Secretary Mechanisms*, 1980; J. F. Hoffman (ed.), *Membrane Transport Processes*, vol. 1, 1978.

Osmosis

The transport of solvent through a semipermeable membrane separating two solutions of different solute concentration. The solvent diffuses from the solution that is dilute in solute to the solution that is concentrated. The phenomenon may be observed by immersing in water a tube partially filled with an aqueous sugar solution and closed at the end with parchment. An increase in the level of the liquid in the solution results from a flow of water through the parchment into the solution. The process occurs as a result of a thermodynamic tendency to equalize the sugar concentrations on both sides of the barrier. The parchment permits the passage of water, but hinders that of the sugar, and is said to be semipermeable. Specially treated collodion and cellophane membranes also exhibit this behavior. These membranes are not perfect, and a gradual diffusion of solute molecules into the more dilute solution will occur. Of all artificial membranes, a deposit of cupric ferrocyanide in the pores of a fine-grained porcelain most nearly approaches complete semipermeability.

The walls of cells in living organisms permit the passage of water and certain solutes, while preventing the passage of other solutes, usually of relatively high molecular weight. These walls act as selectively permeable membranes, and allow osmosis to occur between the interior of the cell and the surrounding media.

The flow of liquid through such a barrier may be stopped by applying pressure to the liquid on the side of higher solute concentration. The applied pressure required to prevent the flow of solvent across a perfectly semipermeable membrane is called the osmotic pressure and is a characteristic of the solution.

Osmotic pressure. The driving force for osmosis is the decrease in Gibbs free energy accompanying the dilution of the solution by the transfer of solvent to the solution. With G_1^0 representing the molar free energy of the pure solvent, the molar free energy of solvent in the solution is shown in Eq. (1), where R is

$$G_1 = G_1^0 + RT \ln a_1 \tag{1}$$

the molar gas law constant, 8.314 J mole^{-1} K^{-1}, and a_1 is the solvent activity in the solution. This is conventionally taken as unity for the pure solvent, and in ideal solutions, as the mole fraction, x_1. The pressure that must be applied to the solution phase to increase the solvent free energy to that of the pure solvent is the osmotic pressure, π, of the solution. The rate at which the free energy of a system is changed at a fixed temperature by pressure is given by the basic thermodynamic relationship (2). For an incompress-

$$(\delta G/\delta P)_T = V \tag{2}$$

ible liquid solution, the osmotic pressure is, then, shown in Eq. (3), where V_1 is the molar volume of sol-

$$\pi = -(RT/V_1) \ln a_1 \tag{3}$$

vent. Equation (3) represents the fundamental relationship between the osmotic pressure and solution thermodynamics. In an ideal solution containing n_1 moles of solvent and n_2 moles of solute, the solvent activity is $a_1 = x_1 = n_1/(n_1+n_2)$. In terms of the solute, the solvent activity is $(1 - x_2)$, where x_2 is the solute mole fraction, $n_2/(n_1 + n_2)$. For dilute solutions, $\ln a_1 = \ln(1 - x_2) \sim -x_2$ since $n_1 >> n_2$. Then Eq. (3) becomes Eq. (4), where n_1V_1 is replaced by the total

$$\pi = n_2RT/(n_1V_1) = n_2RT/V = cRT \tag{4}$$

volume, V. Equation (4) is recognized as identical in form to the ideal gas law, in which c represents the solute concentration in moles L^{-1}. As applied to osmotic phenomena, Eq. (4) is called the van't Hoff equation. Measurements of osmotic pressure made early in the twentieth century on dilute sucrose solutions using copper ferrocyanide membranes confirm the applicability of (4). For example, a 0.0291 molar aqueous sucrose solution at 15.6°C exhibited an osmotic pressure of 0.684 atm compared with the calculated value of 0.689 atm.

Solution properties reflecting changes in solvent free energy caused by a nonvolatile solute include vapor-pressure lowering, boiling-point elevation, freezing-point depression, as well as osmotic pressure. Of these, the last is the most sensitive to solute concentration. Membrane requirements, however, limit its usefulness in studying solution thermodynamics.

An important application of measuring osmotic pressure is determining a polymer's molar mass. The polymer sample is dissolved in an organic solvent or in a mixture of solvents, and the osmotic pressure is measured using a pretreated cellophane film membrane. An osmometer is used in which the pressure is measured in terms of the height to which liquid rises in a solution cell capillary above that in the surrounding solvent. In accounting for nonideality, Eq. (3) is used in a modified form (5), where c^* is

$$\pi/c^* = RT/M + Bc^* + Cc^{*2} \tag{5}$$

the concentration in g/L, and M, the molar mass in g/mole. B and C are correction terms analogous to gas-phase virial coefficients. RT/M is obtained as a limit of π/c^* as c^* approaches zero. The molar mass obtained is a number average of the polymer molar mass spectrum of the sample.

Aside from the inorganic-type membranes, which are formed by deposition of the salt in a porous matrix, membranes suitable for measurements of osmotic pressures in aqueous systems are not available. A number of synthetic polymeric films, which allow the passage of water and small ionic or molecular species, are sufficiently strong to act as selectively permeable membranes in aqueous solutions. These have been successfully used to study, for example, micellar interactions in salt and buffered media. Although the pressures measured are not true osmotic pressures, thermodynamic solute data may be obtained through a redefined $G^0{}_1$ and use of standard Gibbs-Duhem activity relationships. *See* FREE ENERGY.

Reverse osmosis. Just as the osmotic pressure is that pressure which when applied to the solution phase will prevent the solvent from passing through a semipermeable membrane into the solution, the application of greater pressure will cause solvent to pass from the solution into the pure solvent, or reverse osmosis. Reverse osmosis has long been considered for water purification. Home units for improving the quality of a water supply are effectively pressure filtration devices capable of removing colloidal impurities but are largely ineffective in reducing salt concentrations. On a larger scale, polymeric films or ceramic porous membranes, for example, of aluminum oxide, are described in terms of the fraction of salt rejection as a measure of efficiency. The prospect of using reverse osmosis for desalinating seawater is energetically attractive but technically difficult to realize. The osmotic pressure of seawater is approximately 23 atm. Membrane design and mechanics would have to allow for use at substantially higher pressures with large volumes of water,

over long periods. With advances in membrane engineering, it is possible that reverse osmosis could be used to purify ocean water. *See* EDEMA; OSMOREGULATORY MECHANISMS; PLANT-WATER RELATIONS; SOLUTION. Francis J. Johnston

Bibliography. D. A. Amos et al., Osmotic pressure and interparticle interactions in ionic micellar surfactant solutions, *J. Phys. Chem. B*, pp. 2739–2753, 1998; P. W. Atkins, *Physical Chemistry*, 5th ed., W. H. Freeman, 1994; P. R. Bergethon and E. R. Simons (eds.), *Biophysical Chemistry*, 1984; S. Glasstone, *Textbook of Physical Chemistry*, Van Nostrand, 1946; K. Kotyk, K. Koryta, and K. Janacek, *Biophysical Chemistry of Membrane Functions*, 1988; S. Sourirajan and T. Matsuura, *Reverse Osmosis and Ultrafiltration*, American Chemical Society, 1985.

Ostariophysi

A very large group of fishes containing 27% of all known fishes of the world and accounting for 64% of the freshwater species. Although the perciforms dominate modern seas, the ostariophysans command preeminent faunal importance in freshwaters. At least one-half of the freshwater fishes of each continent, with the exception of Australia, belong to this group. *See* PERCIFORMES.

The classification hierarchy is as follows:

- Class Actinopterygii
  - Subclass Neopterygii
    - Division Teleostei
      - Subdivision Euteleostei
        - Superorder Ostariophysi
          - Series Anotophysi
            - Order Gonorhynchiformes
          - Series Otophysi
            - Order Cypriniformes
            - Order Characiformes
            - Order Siluriformes
            - Order Gymnotiformes

Several characterics alone unite the Ostariophysi: modification of the anterior three to five vertebrae as a protective encasement for a series of bony ossicles that form a chain connecting the swim bladder with the inner ear, the so-called Weberian apparatus; multicellular horn tubercles, in the form of breeding and nuptial tubercles or pearl organs; minute unicellular horny projections called unculi; and an alarm substance (pheromone) that can be released by injured tissue to elicite a fright reaction.

Other identifying characteristics include abdominal pelvic fins, if present, usually with many rays; a pelvic girdle which does not have contact with the cleithra; and a more or less horizontal pectoral fin placed low on the side. Usually there are no fin spines, but one or two spines derived from consolidated soft rays may be found in dorsal, anal, or pectoral fins, and a spinelike structure is present in the adipose fin of certain catfishes. The upper jaw is often bordered by premaxillae and maxillae, the latter commonly toothless and rudimentary. An orbitosphenoid is present, and a mesocoracoid is usually present. There is a swim bladder, usually with a duct (physostomous). Scales are the cycloid type, which may be modified into bony plates, and branchiostegal rays are variable in number and arrangement. *See* EAR (VERTEBRATE); SALMONIFORMES; SWIM BLADDER. Herbert Boschung

Osteichthyes

A group that is recognized herein as consisting of the classes Actinopterygii (rayfin fishes) and Sarcopterygii (lungfishes and lobefins, or coelacanths, but excluding tetrapods). (Euteleostomi is a monophyletic group consisting of Actinopterygii and Sarcopterygii, including tetrapods.) The Sarcopterygii and Actinopterygii are well-defined phyletic lines which had evolved by the Middle Devonian, and many members of the two taxa persist today.

The osteichthyans (bony fishes) include most of the Recent fishes. They differ from the living Agnatha in having jaws, paired nostrils, true teeth, paired pelvic and pectoral fins supported by girdles (unless lost secondarily), three semicircular canals, and bony scales (unless lost or modified). Many fossil agnathans possess bony scales but differ from the higher fishes in the above-mentioned features. Separation of the osteichthyans from the Paleozoic Placodermi and Acanthodii is more difficult because these three groups agree in most basic vertebrate features. The osteichthyans contrast with the Chondrichthyes in having a bony skeleton (although some Recent bony fishes possess a largely cartilaginous skeleton), a swim bladder (at least primitively), a true gill cover, and mesodermal scales (sharks possess dermal denticles, or placoid scales) which are sometimes modified or lost. Fertilization is usually external, but if it is internal, the intromittent organ is not derived from pelvic-fin claspers. Most often a modified anal fin or a fleshy tube or sheath functions in sperm transfer. *See* COPULATORY ORGAN; SCALE (ZOOLOGY); SWIM BLADDER.

Classification. Some authorities rank the Actinopterygii and Sarcopterygii in three subclasses: Actinopterygii (rayfin fishes), Coelacanthiformes or Crossopterygii (lobefin fishes), and Dipnoi (lungfishes). Others may combine the lobefin fishes and lungfishes in one subclass, Sarcopterygii or Choanichthyes. The Paleozoic Acanthodii have been variably assigned as a subclass of the Osteichthyes, of the Placodermi, or (as in this encyclopedia) as a separate class. The Polypteriformes (bichirs) are ranked by some as a subclass of the Osteichthyes, but are here placed as an order of the chondrostean Actinopterygii. The osteichthyans as perceived herein had evolved by the Middle Devonian and all persist today; thus each of these subclasses has a longer history than any of the tetrapods, none of which had then appeared.

The lobefin fishes and lung fishes were destined to

experience moderate success until the early Mesozoic, after which they barely persisted as an element in the world's ichthyofauna. Today's fauna of these fishes includes only four families, four genera, and seven species. Of great interest are the rhipidistians (sarcopterygians in the order Osteolepiformes), the paternity of tetrapods. *Latimeria* is a member of the Coelacanthiformes, the other major radiation of crossopterygians. The Dipnoi are especially interesting scientifically because of their air-breathing habit and the capacity of some to estivate for a protracted period. *See* DIPNOI.

In the late Paleozoic the slowly developing Actinopterygii realized marked success in the palaeonisciform line, a group that was superseded by ascendant development of amiiforms and related groups in the Mesozoic only to be replaced by the enormous and highly successful outburst of salmoniform, cypriniform, siluriform, and perciform descendants in the Upper Cretaceous and early Cenozoic. Most modern fishes belong to these latter groups, which are conveniently termed teleosts (that is, the end of the bony fish line.) The Recent osteichthyan fauna includes 2 classes, 45 orders, 435 families, 4076 genera, and 24,497 species.

A classification of extant and selected fossil osteichthyans follows. Equivalent names are given in parentheses. For more detailed information, see separate articles on each group.

Osteichthyes
- Class Actinopterygii
  - Subclass Cladistia
    - Order Polypteriformes (Brachiopterygii)
  - Subclass Chondrostei
    - Order Acipenseriformes
  - Subclass Neopterygii
    - Order Lepisosteiformes (in part, Semionotiformes)
    - Amiiformes
  - Division Teleostei
    - Subdivision Osteoglossomorpha
      - Order Hiodontiformes
      - Osteoglossiformes
    - Subdivision Elopomorpha
      - Order Elopiformes
      - Albuliformes
      - Anguilliformes (Apodes)
      - Saccopharyngiformes
    - Subdivision Clupeomorpha
      - Order Clupeiformes
      - Subdivision Euteleostei
        - Superorder Ostariophysi
          - Series Anotophysi
            - Order Gonorhynchiformes
          - Series Otophysi
            - Order Cypriniformes
              - Characiformes
              - Siluriformes (Nematognathi)
              - Gymnotiformes
        - Superorder Protacanthopterygii
          - Order Argentiniformes, Osmeriformes
            - Salmoniformes
            - Esociformes (Haplomi, Esocae)
        - Superorder Stenopterygii
          - Order Stomiiformes (Stomiatiformes)
            - Ateleopodiformes
        - Superorder Cyclosquamata
          - Order Aulopiformes
        - Superorder Scopelomorpha
          - Order Myctophiformes
        - Superorder Lampridiformes (Lampriformes)
          - Order Lampriformes
        - Superorder Polymixiomorpha
          - Order Polymixiiformes
        - Superorder Paracanthopterygii
          - Order Percopsiformes
            - Gadiformes
            - Ophidiiformes
            - Batrachoidiformes (Haplodoci)
            - Lophiiformes
        - Superorder Acanthopterygii
          - Series Mugilomorpha
            - Order Mugiliformes
          - Series Atherinomorpha
            - Order Atheriniformes
            - Beloniformes
            - Cyprinodontiformes (Microcyprini)
          - Series Percomorpha
            - Order Stephanoberyciformes (Xenoberyces, in part)
            - Beryciformes
            - Zeiformes
            - Order Gasterosteiformes
        - Suborder Gasterosteoidei (Thoracostei)
        - Suborder Syngathoidei (Solenichthys)
          - Order Synbranchiformes (Symbranchii)
            - Scorpaeniformes
            - Perciformes
        - Suborder Percoidei
          - Elassomatoidei
          - Labroidei
          - Zoarcoidei
          - Notothenioidei
          - Trachinoidei
          - Pholidichthyoidei
          - Blennioidei
          - Icosteoidei
          - Gobiesocoidei (Xenopterygii)
          - Callionymoidei
          - Gobioidei
          - Kurtoidei
          - Acanthuroidei
          - Scombrolabracoidei
          - Scombroidei
          - Stromateoidei

Suborder Anabantoidei (Labyrinthici, in part)
Channoidei (Ophiocephaliformes)
Caproidei
Order Pleuronectiformes (Heterostomta)
Tetraodontiformes (Plectognathi)
Class Sarcopterygii
Subclass Coelacanthimorpha
Order Coelacanthiformes
Subclass Dipnotetrapodomorpha
Infraclass Dipnoi (Dipneusti)
Order Ceratodontiformes

There is a lack of general agreement as to the nomenclature of higher groups of fishes. Broadly, two systems are employed, that of C. T. Regan and the later one of L. S. Berg. In the Berg system, followed here, ordinal names are formed by a standard suffix "-formes" to the step of a type genus.

Reeve M. Bailey; Herbert Boschung

Phylogeny. The Osteichthyes evolved from some group among the primitive, bony gnathostome fishes. Plated agnathans and primitive acanthodians are known from the Silurian, but the oldest osteichthyans do not appear until the Lower Devonian, when lobefins and lungfishes enter the paleontological record. The osteichthyans became well represented in the Middle Devonian, at which time the rayfin fishes, lobefin fishes, and lungfishes were well differentiated. The ancestral stock of the osteichthyans may be sought among the other major groups of gnathostomes, the Placodermi and Acanthodii, but the still-fragmentary character of the early fossil record prevents clear elucidation of the origin and initial history of the group. The Placodermi, abundant fishes of the Devonian, were heavily armored and highly specialized, and they display scant evidence of relationship to early osteichthyans. Advanced acanthodians of the Devonian also were highly modified and are unlikely progenitors of osteichthyans. The occurrence of more generalized types in the Late Silurian, about the time when osteichthyans presumably originated, makes the acanthodians a more plausible ancestral stock. Their overall morphology, bony skeleton, including true bony scales of structure not unlike that of osteichthyans, and the presence of paired pectoral and pelvic fins are all suggestive of kinship with that group. The caudal fin of acanthodians is heterocercal, as in primitive osteichthyans. The other fins had strong spines at their leading edges. Commonly there were paired spines between the pectoral and pelvic fins. Such structures are unknown in primitive osteichthyans. In view of the above, it is a plausible inference that the Osteichthyes and Acanthodii diverged from a common basal stock, presumably in Silurian time. *See* ACANTHODII; PLACODERMI.

Although the beginnings of the osteichthyans are shrouded in uncertainty, advancements in paleontology have laid to rest the belief that they evolved from that other great class of modern fishes, the Chondrichthyes. According to this belief, cartilage preceded bone in phylogeny as well as in ontogeny; cartilaginous fishes were seen as ancestral to bony fishes. Support for this assumption was drawn from the occurrence of largely cartilaginous endoskeletons in modern lungfishes and sturgeons, representatives of ancient lineages of osteichthyans. In contradiction, however, it has been shown that early members of these osteichthyan groups are better ossified than are their secondarily chondrified descendants. The adaptive advantage of cartilage in developmental stages appears to explain its prevalence in young fishes. Bone long preceded the oldest known chondrichthyans, and the Osteichthyes themselves precede the Chondrichthyes in the fossil record. *See* CHONDRICHTHYES.

Fossils. Osteichthyes contains many fossil fishes, including the ancestors of tetrapods. Their earliest satisfactory record is Lower Devonian, but remains of fossils of possible osteichthyan linage have been found in the Upper Silurian.

The upper biting edge of the mouth is formed by dermal premaxillary and maxillary bones to which teeth are fused, rather than by palatoquadrates (as in acanthodians and elasmobranchs). Lateral line canals run through the dermal bones and scales rather than between them. A dermal parasphenoid forms the roof of the mouth, and other dermal bones form an opercular-shoulder girdle complex around the branchial chamber. Primitively the neurocranium is perichondrally ossified in two parts, anterior and posterior, which are partly separated by cranial fissures (lost in higher bony fishes). Also primitively, the fins are supported by dermal lepidotrichia. An air bladder is usually present, sometimes modified into a lung or hydrostatic organ.

John G. Maisey

Bibliography. L. S. Berg, *Classification of Fishes, Both Recent and Fossil*, 1947; P. H. Greenwood et al., Phyletic studies of teleostean fishes, with a provisional classification of living forms, *Bull. Amer. Mus. Nat. Hist.*, 131:339–456, 1966; P. H. Greenwood et al. (eds.), *Interrelationship of Fishes*, 1973; E. S. Herald, *Living Fishes of the World*, 1961; J. A. Moy-Thomas and R. S. Miles, *Palaeozoic Fishes*, 2d ed., 1971; J. S. Nelson, *fishes of the world*, Wiley, 1994; C. Patterson and D. E. Rosen, Review of ichthyodectiform and other Mesozoic teleost fishes and the theory and practice of classifying fossils, *Bull. Amer. Mus. Nat. Hist.*, 158:81–172, 1977; A. S. Romer, *Vertebrate Paleontology*, 3d ed., 1966; D. E. Rosen et al., Lungfishes, tetrapods, paleontology, and pleisiomorphy, *Bull. Amer. Mus. Nat. Hist.*, 167:163–275, 1981; B. Schaeffer and D. E. Rosen, Major adaptive levels in the evolution of the actinopterygian feeding mechanism, *Amer. Zool.*, 1:187–204, 1961.

Osteoglossiformes

An order of teleost fishes consisting of two monophyletic clades, suborders Osteoglossoidei and Notopteroidei. The Osteoglossoidei consists of one

family, the Osteoglossidae (bonytongues and butterfly fishes), which occur in freshwaters of tropical South America, Africa, and Southeast Asia to northern Australia. The Notopteroidei consist of three extant families, Notopteridae (Old World knifefishes and featherfin knifefishes, or featherbacks), Mormyridae (elephantfishes), and Gymnarchidae (gymnarchid eel). The featherbacks are native to central tropical Africa, India, and the Malay Archipelago. The elephantfish family, having a long proboscislike snout in many species, is the most species-rich osteoglossiform taxon but is limited to tropical Africa and the Nile River. The family Gymnarchidae is represented by a single species, *Gymnarchus niloticus*, of tropical Africa including the upper Nile.

These fishes, formerly assigned to the orders Clupeiformes (Isospondyli) and Mormyriformes, are accorded a revised classification as a result of modern phylogenetic research on lower teleosts.

Anatomy and development. Osteoglossiformes have the primary bite between the well-toothed tongue and the strongly toothed parasphenoid and certain pterygoid bones in the roof of the mouth. The mouth is bordered by a small premaxilla, which is fixed to the skull, and the maxilla. The anterior part of the intestine passes posteriorly to the left of the esophagus and stomach, whereas in all but a very few fishes the intestine pass to the right. A unique feature is the presence of paired bony rods at the base of the second gill arch. Additional characters include the absence of epineural intermuscular bones; a single dorsal fin of soft rays (or absent in some featherbacks) [**Figs. 1** and **2**]; lack of adipose fin; caudal fin either well developed, reduced, or absent, its supporting skeleton usually compact and consolidated; pelvic fin usually abdominal, not attached to cleithrum; pelvic rays usually 6 or 7, the fin sometimes rudimentary or absent; supraoccipital separated from frontals by parietals; orbitosphenoid usually present; a mesopterygoid arch; no Weberian apparatus; branchiostegals 3–17; and cycloid scales that are commonly large and often have a complex reticulum formed by the radii. Unlike the Clupeiformes, there is no extension of the cephalic canal system onto the opercle. In contrast to the Elopiformes, development is without a leptocephalous (elongate and flattened side to side) larva. *See* CLUPEIFORMES; ELOPIFORMES; TELEOSTEI.

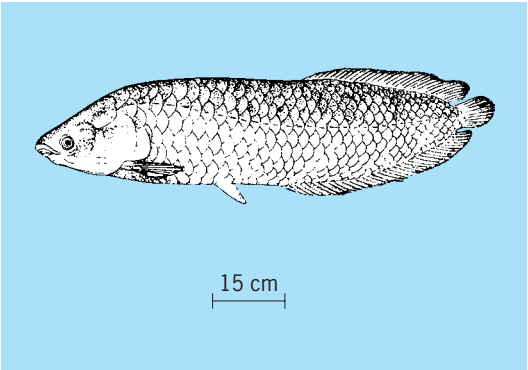

Fig. 1. African bonytongue (*Heterotis niloticus*). (*After G. A. Boulenger, Catalogue of the British Museum, Natural History, vol. 1, 1909*)

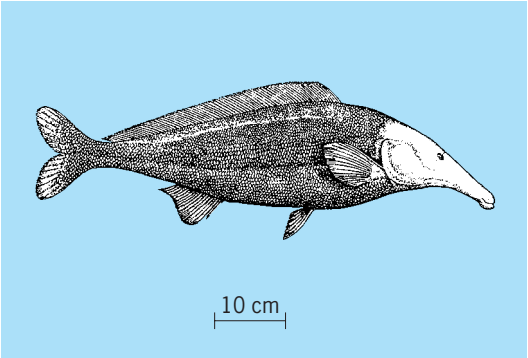

Fig. 2. African elephantnose (*Mormyrus proboscirostris*). (*After G. A. Boulenger, Catalogue of Fresh Water Fishes of Africa in the British Museum, Natural History, vol. 1, 1909*)

Evolution and adaptations. Osteoglossiforms represent one of the basal stocks among the teleosts. The Ichthyodectidae, including such giants as *Xiphactinus* (or *Portheus*), and two related families are well represented in Cretaceous marine deposits, and *Allothrissops* and *Pachythrissops* of the Jurassic and Cretaceous probably belong here.

Recent members may be classified in 2 suborders, 6 families, 29 genera, and about 217 species. All inhabit freshwater, and all are tropical.

Bonytongues are well known because of their economic importance, large size, and frequent exhibition in public aquariums. They are represented by five Recent species: the microphagous, African *Heterotis* (Fig. 1), the giant *Arapaima gigas* (over 10 ft or 3 m long), and the arawana (*Osteoglossum bicirrhosum*) of the Amazon, and two species of *Scleropages* of southeastern Asia, New Guinea, and Australia. The North American Eocene *Phareodus* is a typical osteoglossid. A related group consists of a single small species known as African butterflyfish (*Pantodon buchholzi*). Butterflyfish leap a meter or more out of the water and glide with the help of their expansive pectoral fins. The featherbacks, Notopteridae, with five species from Africa and southeastern Asia, are compressed fishes with tapering bodies and long anal fins that are continuous with the caudal fin; the dorsal fin, if present, is a tiny tuft of a few rays.

Numerically Osteoglossiformes is dominated by the Mormyridae. Of the roughly 198 species, many are valued as food. The elongate snout of some species accounts for the vernacular name elephantfishes (Fig. 2). The caudal peduncle is slim in all species; in most the dorsal and anal fins are approximately equal, but in some the anal is much the longer and in others the dorsal greatly exceeds the anal (Fig. 2). In the related family Gymnarchidae, consisting of a single species that attains a length of 1.5 m, the pelvic, anal, and caudal fins are absent. Mormyrids and gymnarchids are of especial interest

in that they are electrogenic. Modified muscles in the caudal peduncle generate and emit a continuous electric pulse. The discharge frequency varies, being low at rest and high when the fish is under stress. The mechanism operates like a radar device, since the fish is alerted whenever an electrical conductor enters the electromagnetic field surrounding it. The receptor mechanism is not fully understood, but the brain of mormyrids (especially the cerebellum) is the largest of all fishes. Mormyrids have small eyes and often live in turbid water, so this electric mechanism is an effective evolutionary solution to a need for keen awareness of the environment and is comparable, for example, to acute hearing or vision. Electric discharge has been developed in several groups of fishes; in the African mormyrids the discharge is convergent with a similar but wholly unrelated acquisition in the gymnotids or knifefishes of South America. *See* ACTINOPTERYGII; CYPRINIFORMES; ELECTRIC ORGAN (BIOLOGY).

Osteoglossiforms vary greatly in size, shape, and biology. Their trophic biology is greatly diverse, the various species being either piscivores, insectivores, filter feeders, or omnivores. The well-vascularized swim bladder can act as a lung, which is especially helpful in oxygen-deficient habitats. *Gymnarchus niloticus* builds a floating nest; and the osteoglossid genera *Arapaima* and *Heterotis* are substrate nest builders, whereas *Scleropages* and *Osteoglossus* are buccal incubators. *Arapaima gigas* of South America is one of the largest freshwater fishes of the world, at least 3 m, perhaps 4.5 m, in length. Noeopterids reach a meter in length; they are egg layers and the male guards the eggs. Most mormyrids are small, up to 50 cm in length; most are much less, but 1.5-m specimens have been reported. Species vary in having a blunt snout with a terminal or inferior mouth, to a long snout and small terminal mouth, hence variable feeding strategies.

Reeve M. Bailey; Herbert Boschung

Bibliography. T. M. Berra, *An Atlas of Distribution of the Freshwater Fishes of the World*, University of Nebraska Press, 1986; P. H. Greenwood, Interrelationships of osteoglossomorphs, in P. H. Greenwood, R. S. Miles, and C. Patterson (eds.), Interrelationships of fishes, J. Linn. Soc. (Zool.), 53(Suppl. 1):307–332, Academic Press, 1973; V. G. Lauder and K. F. Leim, The evolution and interrelationships of the actinoperygian fishes, *Bull. Mus. Comp. Zool.*, 150:95–197, 1983; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, 1994.

Osteolepiformes

Fossil lobe-finned (sarcopterygian) fishes of Devonian to Permian age (395–270 million years ago), the ancestral group for the land vertebrates, or "tetrapods." *See* SARCOPTERYGII; TETRAPODA.

Phylogeny. The term "osteolepiforms" is applied to a range of fossil fishes, found worldwide in Devonian to Permian strata, that appear to be closely related to tetrapods. After decades of debate about their precise position in the vertebrate family tree, a consensus has emerged that they occupy the bottom part of the tetrapod stem lineage, that is, the evolutionary line that leads toward land vertebrates from the last common ancestor we share with our nearest living fish relatives, the lungfishes (**Fig. 1**). Immediately above the osteolepiforms in the stem lineage, just below the most primitive tetrapods, we find the transitional fossil Panderichthys. *See* PANDERICHTHYS.

Rather than forming a single coherent group, the osteolepiforms spread out along the tetrapod stem lineage (Fig. 1, shaded zone): some are very primitive and fall near the base of the lineage, others are close to the fish-tetrapod transition. Such an arrangement is known as a paraphyletic group. The precise relationships of the different osteolepiforms are still debated, but this article follows the arguments presented by P. Ahlberg and Z. Johanson in 1998.

Morphology. Despite their paraphyletic arrangement along the tetrapod stem lineage, all osteolepiforms look reasonably similar and share many characteristics. This implies that the characteristics in question are ancestral for tetrapods; osteolepiform fossils thus give a clear picture of the specific fish body plan that gave rise to land vertebrates. General characteristics of the osteolepiforms include an intracranial joint running vertically through the skull behind the eyes, which allowed the front part of the head to move slightly up and down relative to the back part; a pair of internal nostrils, or choanae, connecting the nose to the palate; and paired fin

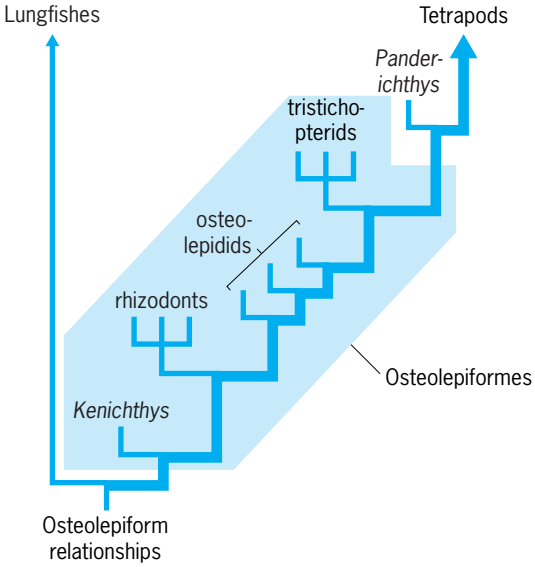

Fig. 1. Schematic phylogeny (family tree) of the osteolepiforms, based on arguments presented by P. Ahlberg and Z. Johanson (1998). The fishes within the grey zone are conventionally designated as Osteolepiformes (the formal group name), but they actually form a paraphyletic segment of the tetrapod stem rather than a well-defined group. The tetrapod stem lineage is indicated in thick line. Arrowheads indicate lineages surviving to the present day.

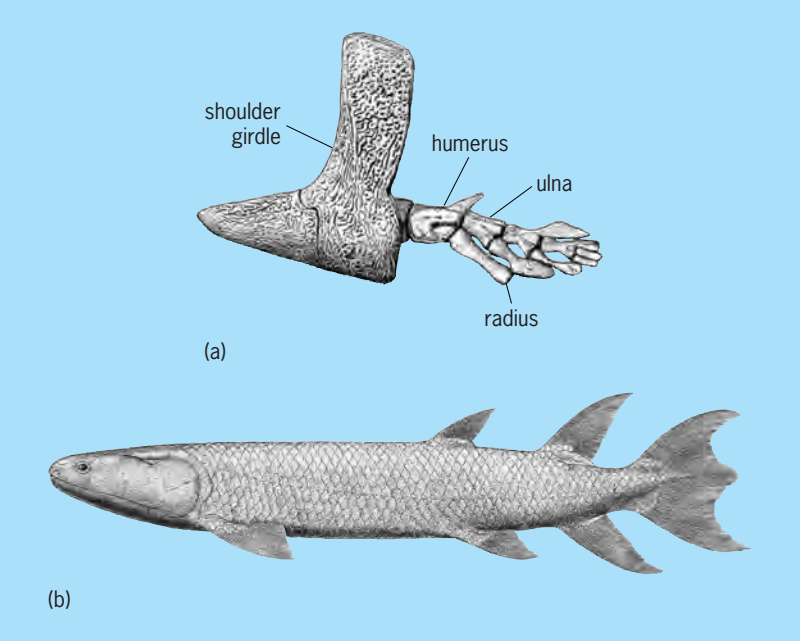

Fig. 2. The tristichopterid *Eusthenopteron foordi*. (*a*) A close-up of its pectoral fin skeleton. (*b*) An illustration of the full specimen. *Eusthenopteron* reached approximately 50–70 cm in length. Its pectoral fin skeleton contains structures equivalent to the tetrapod humerus, radius and ulna, but no equivalent of digits. (*Reprinted from Basic Structure and Evolution of Vertebrates, vol. 1, E. Jarvik, pp. xx, ©1980, with permission from Elsevier*)

skeletons in which one can recognize equivalents of the tetrapod limb bones humerus, radius, ulna (forelimb/pectoral fin) and femur, tibia, fibula (hindlimb/pelvic fin) (**Fig. 2***a*). The intracranial joint was lost during the fish-tetrapod transition—it is immobilized in *Panderichthys* and completely obliterated in early tetrapods—but the internal nostrils and limb skeletons persisted and became basic components of the tetrapod body plan. The similarities between osteolepiform fin skeletons and tetrapod limb skeletons are quite specific, extending to the detailed pattern of muscle attachments on the bones. However, despite their tetrapod-like anatomy, osteolepiforms retained the overall body form of "normal" open-water fishes and do not seem to have been adapted for an amphibious life. Osteolepiforms are known from both marine and nonmarine environments.

Kenichthys. The earliest and most primitive osteolepiform is *Kenichthys* from the Early Devonian of China (about 395 million years ago). It was a small fish, with a head just over 2 cm (0.8 in) long (complete bodies have not been found). The external bones of the skull are covered in cosmine, a smooth shiny coating of dentine and enamel that is also found in many other lobe-finned fishes. The jaws of *Kenichthys* carry broad fields of bulbous, bullet-shaped crushing teeth, suggesting that it may have fed on shelly invertebrates. Its most interesting feature is the snout, which gives a "snapshot" of the evolution of the internal nostril. Most fishes have two external nostrils on each side, both connected to the nasal sac where the olfactory nerve endings are located, whereas tetrapods and osteolepiforms have one external and one internal nostril, or choana, on each side. *Kenichthys*, uniquely, has one external nostril and one right on the edge of the mouth, opening within the upper jaw tooth row. This neatly intermediate state shows that one of the external fish nostrils (the posterior one) migrated onto the palate to become the choana, pushing through the tooth row in the process.

Rhizodonts. Immediately above *Kenichthys* in the tetrapod stem lineage are the rhizodonts, a group of gigantic predatory osteolepiforms. They appear rather late in the fossil record, at the beginning of the Late Devonian (about 370 million years ago), but their position in the tree suggests that they have an unrecorded "prehistory" of some 15 million years or more. It is clear that the rhizodonts originated in Gondwana (the southern supercontinent that includes Africa, South America, India, Antarctica, and Australia), as this is where all the earliest fossils are found. By the end of the Devonian, rhizodonts had spread to the northern Euramerican supercontinent (North America and Europe), and they remained a prominent feature of freshwater faunas through the earlier part of the Carboniferous Period, until their extinction about 300 million years ago.

All known rhizodonts were large fishes. *Gooloogongia* from the Late Devonian of Australia, one of the earliest and most primitive, was a slightly flattened and probably bottom-dwelling predator just over 1 m (3 ft) in length, whereas *Rhizodus hibberti* from the Carboniferous of Scotland reached at least 7 m (23 ft) and is the largest freshwater fish known. Rhizodonts had formidable dentitions of piercing teeth, sometimes equipped with cutting edges, and must have been the top predators of their ecosystems. Their large, rounded pectoral fins, which have very complex fin skeletons, may have been used as paddles to scull through weed-choked waters; they do not appear to have been active swimmers. Later forms such as *Strepsodus* have almost eel-like bodies with very reduced midline fins and were probably lurking ambush predators. This is also suggested by their extremely elaborate lateral line systems, sensory organs that detect vibrations in the water. Rhizodonts lacked cosmine and had round scales.

Osteolepidids and tristichopterids. The remaining "typical" osteolepiforms can be roughly divided into osteolepidids, which had cosmine and rhombic scales, and tristichopterids, which lacked cosmine and had round scales. This similarity between tristichopterids and rhizodonts seems to be convergent (that is, independently evolved in the two groups), as other characteristics indicate that the tristichopterids are much closer to tetrapods than the rhizodonts. The osteolepidids probably form a paraphyletic segment of the tetrapod stem lineage, but the tristichopterids are a coherent group with a single common ancestry (a clade) that forms a well-defined side branch on the tree (Fig. 1).

Osteolepidids. Osteolepidids appear in the fossil record early in the Middle Devonian, about 390 million years ago, and have a worldwide

distribution. Early forms such as *Osteolepis* are usually quite small, about 20–40 cm (8–16 in) total length, and broadly resemble *Kenichthys* except that they had fully developed choanae and a more obviously predatory dentition with narrow rows of piercing teeth. Some later osteolepidid groups, such as the megalichthyids, (which become prominent in the Carboniferous and persisted into the early Permian as the last of all osteolepiforms), reached lengths up to 2 m (7 ft). Megalichthyids often occur alongside rhizodonts but never reached the same gigantic proportions. Like *Kenichthys* and rhizodonts, osteolepidids had broad, rounded heads with short snouts. They usually had heterocercal (that is, asymmetrical, "shark-like") tail fins.

Tristichopterids. The tristichopterids (or eusthenopterids, according to some authors) are the most advanced osteolepiforms. They had narrower, more sharply pointed heads than the other groups, and, in particular, longer snouts. These features are shared with *Panderichthys* and tetrapods, and may reflect a change from suction feeding to a "snapping" mode of prey capture. Tristichopterids had more or less symmetrical tail fins, often with a distinctive three-pronged outline, suggesting they were neutrally buoyant and did not use their tails to generate lift.

The best-known tristichopterid, and arguably the most thoroughly studied of all fossil fishes, is *Eusthenopteron foordi* from the early Late Devonian of Quebec, Canada (Fig. 2*b*). A perfect specimen of this fish was investigated during the 1940s–1960s by Dr. Erik Jarvik of the Swedish Museum of Natural History (Naturhistoriska Riksmuseet), Stockholm, using the painstaking serial grinding technique. This method involved grinding away the head of the fossil a fraction of a millimeter at a time and using the cross sections revealed by the grinding process as templates for building a large wax model; this allowed Jarvik to study not only the external features but also the internal anatomy of the brain cavity and inner ears.

The earliest tristichopterids, such as *Tristichopterus* from the late Middle Devonian of Scotland, were only about 30 cm (12 in) long. Later forms were generally larger, culminating in the 2–3-m (7–10 ft) long *Eusthenodon* and *Hyneria* from the latest Devonian, probably major predators on the early tetrapods that occur in the same faunas. These late, large tristichopterids were in some respects convergent on *Panderichthys* and early tetrapods—they had greatly lengthened snouts and somewhat reduced median fins—but unlike *Panderichthys*, their paired fins do not show any modification toward walking. They died out at the Devonian-Carboniferous boundary and may have been replaced ecologically by the rhizodonts.

Conclusion. Overall, the osteolepiforms are rather conservative fishes. Although they share a suite of tetrapodlike characteristics, there is no sign that they were becoming progressively more amphibious. There is, however, a repeated trend toward greatly increased size, coupled with a reduction of the median fins and the acquisition of enlarged teeth at the front of the jaw, in most groups. Only in the *Panderichthys* lineage was this trend coupled with the emergence of incipient terrestrial adaptations; this became the starting point for the origin of tetrapods. *See* ANIMAL EVOLUTION; CARBONIFEROUS; DEVONIAN; FOSSIL; PERMIAN. Per E. Ahlberg

Bibliography. J. A. Clack, *Gaining Ground*, Indiana University Press, 2002; J. A. Long, *The Rise of Fishes*, The John Hopkins University Press, 1996; C. Zimmer, *At the Water's Edge*, The Free Press, 1998.

Osteoporosis

A metabolic bone disease in which the amount of bone tissue is reduced sufficiently to increase the likelihood of fracture. Fractures of the vertebrae, femur (hip), and wrist are the most common osteoporotic fractures, but other bones such as the ribs, upper arm, and pelvis may also fracture.

Although low bone mass is the major factor in osteoporotic fractures, there may also be qualitative and architectural changes in bone with aging that lead to increased fragility. Osteoporosis can be primary or secondary. Primary osteoporosis occurs independently of other causes. The secondary osteoporoses result from identifiable causes, such as exogenous cortisone administration, Cushing's disease, hyperparathyroidism, hyperthyroidism, hypogonadism, multiple myeloma, prolonged immobilization, alcoholism, anorexia nervosa, and various gastrointestinal disorders. Primary osteoporosis occurring in children is called juvenile osteoporosis; that occurring in premenopausal women and middle-aged or young men is known as idiopathic osteoporosis. Osteoporosis, which is found in older persons, can be classified as postmenopausal (type I) or involutional (type II) osteoporosis (see **table**). *See* ALCOHOLISM; ANOREXIA NERVOSA; GASTROINTESTINAL TRACT DISORDERS; METABOLIC DISORDERS.

Pathogenesis. The pathogenesis of osteoporotic fracture is related not only to the presence of low bone mass but also to the mechanical stress placed on bone. The pathogenesis of low bone mass is multifactorial. Indeed, the changes with aging (see **illus.**) may be viewed as occurring in different stages of the life cycle. The development of maximal peak bone mass prior to involutional bone loss is thought to be related to heredity, nutrition, and physical activity. In young adulthood, men have a higher bone mass than women, and blacks have a higher bone mass than whites. The pattern of bone loss from various skeletal sites is primarily a reflection of the proportion of compact bone near the surface of a bone to less dense bone deeper within a bone. There appears to be substantial bone loss from the femur and spine before menopause. Significant bone loss from the radius bone in the forearm appears to begin after menopause. In older age, renal function

Features of postmenopausal (type I) and involutional (type II) osteoporosis

Factor	Type I	Type II
Age	55–75	Over 70
Sex (female-to-male ratio)	6:1	2:1
Fracture sites	Wrist/vertebrae	Hip/vertebrae, long bones
Hormonal cause	Estrogen deficiency	Calcitriol deficiency
Calcium absorption	Decreased	Decreased
Parathyroid hormone level	Normal	Increased
Importance of dietary calcium	Moderate	High

declines, with diminished ability of the kidney to produce calcitriol—the active form of vitamin D. As a result, calcium malabsorption and secondary hyperparathyroidism occur with further bone loss in the aged. *See* CALCIUM METABOLISM; MENOPAUSE; VITAMIN D.

Prevention. Goals should include prevention of both the underlying disorder (the disease) and its effects (osteoporotic fractures).

Disease prevention. Secondary osteoporoses are managed by eliminating the underlying disorder. To prevent primary osteoporosis, good health-related behavior during childhood and young adulthood has been suggested as the most important factor. Such behavior includes avoiding cigarette smoking and excess alcohol intake, maintaining a normal body weight, and maintaining optimal dietary intake of calcium; preferably, calcium should be obtained by including dairy products with each meal, rather than by taking calcium supplements. It has been suggested that the recommended daily allowance (RDA) for calcium in the postmenopausal period and the recommended intake of vitamin D in the aged should be increased. At menopause, the possibility of estrogen replacement therapy should be considered for women who are at high risk for osteoporosis; this therapy decreases bone loss and reduces the occurrence of osteoporotic fractures. *See* ESTROGEN; NUTRITION.

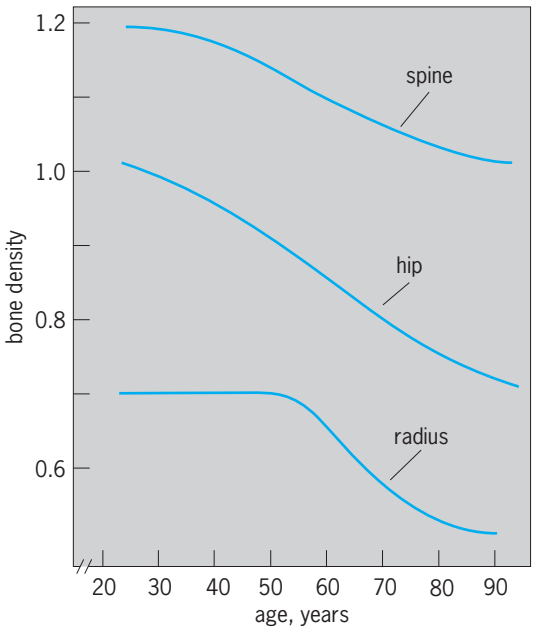

Decrease in bone density with age in women. Before menopause, bone loss affects primarily the spine and hip; after menopause, it is most pronounced in the radius.

Fracture prevention. Fracture prevention should be a lifelong effort. During childhood and the premenopausal years, a maximal peak bone mass should be developed through weight-bearing exercise. Exercise to improve coordination and flexibility and to maintain good posture is also useful. As discussed above, there should be an adequate intake of calcium and vitamin D. The types of exercises noted above should be continued during and after menopause. In the elderly, walking and exercises to promote agility should be encouraged, but the risk of spinal injury must be avoided. It is important to deal with health factors that may lead to falls, such as loss of hearing and vision, the use of certain medications, and postural instability. In addition, efforts should be made to minimize the risk of falling. Installation of handrails, nonskid stair treads, and adequate lighting can help prevent injuries. *See* HORMONE; PARATHYROID HORMONE.

Diagnosis and treatment. The diagnosis of primary osteoporosis is made in the presence of either low bone mass or a characteristic fracture that cannot be attributed to some other cause.

Standard x-rays can show the shape of bones and indicate the presence of fracture. A number of methods that measure bone density may be used to estimate risk for osteoporosis and to assess responses to therapy. These methods include quantitative computed tomography, single-energy photon absorptiometry, and dual-energy x-ray and photon absorptiometry. Neutron activation analysis with whole-body counting is a research tool that can be used to measure total body calcium, which is a measure of skeletal mass. Secondary osteoporoses are excluded by specific laboratory tests. Bone biopsy for histologic evaluation is ordinarily performed only when another form of metabolic bone disease is suspected.

Bone mineral measurement has been recommended for specific patient groups. In estrogen-deficient women, it can be used to identify bone mass low enough to warrant hormone replacement therapy. In patients with vertebral abnormalities or evidence of reduced bone mass on x-rays, it helps establish a diagnosis of spinal osteoporosis. It can assist in monitoring bone mass in patients receiving long-term glucocorticoid therapy and in evaluating

primary asymptomatic hyperparathyroidism. *See* ACTIVATION ANALYSIS; COMPUTERIZED TOMOGRAPHY; MEDICAL IMAGING.

The goals of treatment are rehabilitation and minimization of the risk of future fractures. Goals for rehabilitation include restoring a positive outlook on life, treating depression if it exists, increase of physical activity, restoring independence, relieving pain, restoring muscle mass, and improving posture. Various medications, including estrogen and calcitonin, can maintain bone mass. Basic research on the local control of bone cells and on growth factors holds promise for therapy that can safely increase bone mass. *See* AGING; BONE; SKELETAL SYSTEM DISORDERS. John F. Aloia

Bibliography. J. F. Aloia, *Osteoporosis: A Guide to Prevention and Treatment*, 1989; L. V. Avioli and S. M. Krane, *Metabolic Bone Disease and Clinically Related Disorders*, 3d ed., 1998; H. Broell and M. Dambacher (eds.), *Osteoporosis*, 1994.

Osteostraci

An order of extinct jawless vertebrate fishes, also called Cephalaspida, known from the Middle Silurian to Upper Devonian of Europe, Asia, and North America. They were mostly small, about 2 in. to 2 ft (5 to 60 cm) in length. The head and part of the body were encased in a solid armor of bone, and the posterior part of the body and the tail were covered with thick scales. Some early forms lacked paired fins, though most possessed flaplike pectoral fins. One or two dorsal fins were present. Their depressed shape and the position of the eyes on the top of the head suggest that Osteostraci were bottom dwellers (**Fig. 1**). The underside of the throat region was covered by small plates and there was a small mouth in front.

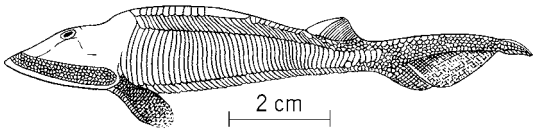

Fig. 1. The ostracoderm *Hemicyclaspis*, a cephalaspid, a Lower Devonian jawless vertebrate. (\hskip.6pt*After E. H. Colbert, Evolution of the Vertebrates, Wiley, 1955*)

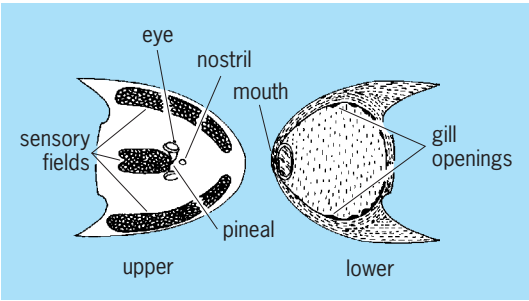

Fig. 2. Upper and lower surfaces of the head in the Devonian ostracoderm *Cephalaspis*. (*After E. H. Colbert, Evolution of the Vertebrates, Wiley, 1955*)

The gill chamber was large and contained as many as 10 pairs of gills, each opening separately to the exterior (**Fig. 2**). It is believed that they fed by sucking in water and straining out food particles in the gills. Because the internal skeleton of the head was often bony, its structure is well known. Details of this structure, particularly the single, dorsal, and median nostril, and the presence of only two pairs of semicircular canals indicate a relationship to extinct Anaspida and to living Petromyzonida. The Devonian *Cephalaspis* is the best-known genus. *See* ANASPIDA; JAWLESS VERTEBRATES; OSTRACODERM. Robert H. Denison

Ostracoda

A major taxon of the Crustacea containing small bivalved animals 0.004–1.4 in. (0.1–33 mm) long, with most between 0.04 and 0.08 in. (1 and 2 mm). They inhabit aquatic environments in nearly all parts of the world. Semiterrestrial species have been described from moss and leaf-litter habitats in Africa, Madagascar, Australia, and New Zealand, and from vegetable debris of marine origin in the Kuril Archipelago. In the oceans, ostracodes live from nearshore to abyssal depths, some swimming and others crawling on the bottom; several are adapted to estuaries of rapidly changing salinity. In fresh water, ostracodes inhabit lakes, ponds, streams, swamps, caves, and ephemeral little standing bodies of water. Of the more than 2000 species extant, none is truly parasitic and most are free-living. However, a few fresh-water and marine forms live commensally on other animals: among the Podocopina, *Entocythere* clings to the gills of crayfish and *Sphaeromicola* and *Paradoxostoma rostratum* to appendages of isopods and amphipods; among the Myodocopina, *Vargula parasitica* and *Sheina orri* are found on the gills of sharks and rays. Ostracodes themselves are parasitized; some marine species are infested with parasitic isopods and copepods, and some fresh-water species serve as intermediate hosts for a cestode (the adult of which parasitizes the black duck) and for an acanthocephalan (the adult of which parasitizes the black bass). Most ostracodes are scavengers, some are herbivorous, and a few are predacious carnivores. Exceptional biological features are known; there are myodocopine ostracodes that produce bioluminescence, and some species of podocopines form a secretion from spinning glands to enable them to climb polished surfaces.

Researchers disagree on the hierarchical classification of the Ostracoda. It is recognized as a distinct class by some and a subclass within the Maxillopoda by others. Consequently, the subdivisions within the Ostracoda may be ranked as subclasses, superorders, or orders. Six major subdivisions are currently recognized: Bradoriida, Phosphatocopida, Leperditicopida, Paleocopa, Myodocopa, and Podocopa. The first four taxa are extinct. The Myodocopa are further subdivided into the

Myodocopida and Halocyprida, and the Podocopa into Platycopida and Podocopida. All fresh-water Ostracoda belong to the Podocopida. Many groups have highly ornamented carapaces, which can be readily classified; a few groups have smooth, bean-shaped carapaces that differ only slightly in shape and present difficulties in taxonomy, particularly in the fossil forms (which have no appendages preserved). *See* MAXILLOPODA; PODOCOPA.

Morphology. Knowledge of the morphology of the Ostracoda is based primarily on extensive study of many species.

Carapace. The two valves, sufficient to enclose the rest of the animal, are joined dorsally along a hinge, that may vary from a simple juncture to a complex series of teeth and sockets. They open by a ligament and close by adductor muscles, which pass through the body. The valves provide an attachment for muscles operating the appendages. Each typically consists of an outer lamella and an inward-facing layer, the inner lamella. The triple-layered outer lamella is commonly composed of a layer of calcite between thin films of chitin, but some large pelagic ostracodes lack the calcareous layer having a nearly transparent, flexible carapace. The rim of the inner lamella may also contain a calcareous central layer to form the duplicature. The chitinous coating is highly colored in some living forms, but it disappears in fossilization.

Body. From the dorsal part of the carapace, the elongate body is suspended as a pliable sac, with lateral flaps of hypodermis extending between the lamellae of the valves. It is unsegmented, but reinforced by chitinous processes for rigidity in the vicinity of the appendages. There is no true abdominal region.

Appendages. Ostracodes of the Podocopida and Myodocopida possess seven pairs of segmented appendages: antennules, antennae, mandibles, maxillules, and three pairs of thoracic legs (**Fig. 1***a*). The body terminates posteroventrally in a pair of furcae or caudal processes. The Platycopida have two pairs of thoracic legs, and the Halocyprida frequently only one pair. Appendages have muscles within them and are connected by musculature to the valves and to a centrally located chitinous structure, the endoskeleton. They are usually specialized for particular functions, such as swimming, walking, food gathering, mastication, and cleaning the interior of the carapace.

Digestive system. The alimentary canal is complete (Fig. 1*d*), consisting of the mouth, esophagus, foregut, midgut, hindgut, and anus. Some have a pair of midgut glands, which discharge into the gut.

Glands. Excretory and secreting glands are scattered in the body, showing no distribution which could indicate original segmentation of the animal.

Respiratory system. A few species of the subfamily Cyclasteropinae have rudimentary gills. All other ostracodes respire through the thin body wall. Water is circulated by waving of branchial plates on certain appendages.

Nervous system. The central system consists of a cerebrum, circumesophageal ganglion, and a ventral chain of partly fused ganglia. Nerves extend to sensory hairs or setae on the outer surface of the carapace and on the appendages. Modified setae are thought to be olfactory or chemoreceptors. Most Myodocopida have lateral eyes with several lenses, in addition to a median frontal organ (presumably light-sensitive); the Podocopina have lateral eyes (one or two lenses each) fused with a median eye. The suborders Cladocopina and Platycopina lack eyes entirely.

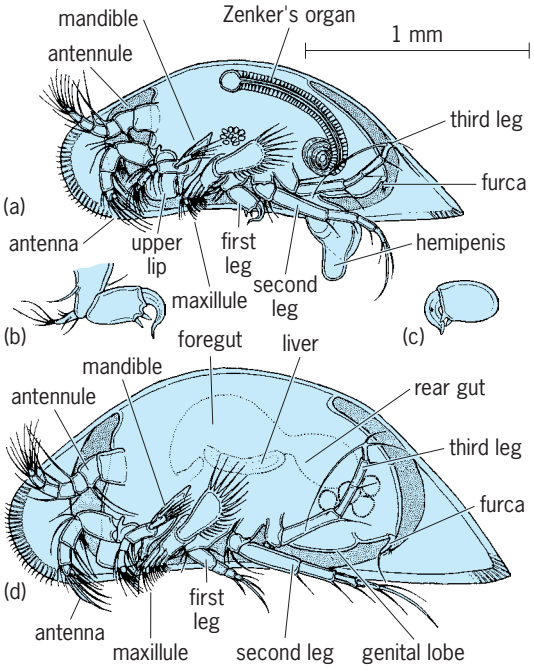

Fig. 1. Ostracode, a typical marine species of the suborder Podocopina, *Macrocypris minna*. (*a*) Male with left valve removed to show appendages; the internal paired Zenker's organs act as pumps to eject the exceptionally large sperm. (*b*) Male left first leg, enlarged; the palp is modified into a pincerlike clasping organ which holds fast to the female during copulation. (*c*) Male palp of right first leg, enlarged. (*d*) Female with left valve removed; a part of the digestive system is indicated by dotted lines. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. Q, 1961, and R. V. Kesling, Contrib. Mus. Paleontol. Univ. Mich., 1969*)

Reproductive system. Ostracode reproductive systems usually are duplicated on left and right sides, the halves not connected. Ostracodes have separate sexes, although many species are parthenogenetic and lack males. Whereas the carapace is a protective armor, its presence creates great difficulties in copulation; males have hemipenes of exceptional size and complexity that can be folded for retraction within the closed valves. In most groups, gonads of male and female lie in the rear part of the body (Fig. 1*a*, *d*), but in some they extend out into the hypodermis alongside the valve wall. Ostracode sperm are among the largest known in the animal kingdom, some exceeding the length of the carapace. To eject these bulky cells, males of the families Cyprididae and Macrocyprididae (suborder Podocopina) have a pair of large seminal pumps, the Zenker's organs (Fig. 1*a*), provided with intricate musculature. Various stages of egg development can be traced in the course of the female ovaries and uteri.

Reproduction and ontogeny. In some species males are unknown and the eggs develop without fertilization (parthenogenesis), but in other species males are produced and fertilize the eggs (syngamy). Most ostracodes lay their eggs on the substrate or on vegetation, but some transfer them to the posterior space within the carapace, where they hatch and the young brood is retained for a time. The eggs of fresh-water species withstand desiccation for many years and hatch only when again placed in water. This adaptation against drought has enabled many fresh-water species to spread over vast regions, the dried eggs being carried by winds or by mud on the feet of migrating birds.

The egg hatches into a bivalved metanauplius. Growth is by ecdysis; the whole of the covering of carapace and appendages is cast off almost simultaneously with secretion of the new enlarged covering. Six to eight instars (growth stages) plus the adult stage are present, according to the group. During the brief, critical interval of ecdysis, the animal adds new appendages and organs and increases the volume of the carapace to twice its former size. Fresh-water ostracodes reach maturity in about a month, depending on temperature of the water, but marine species may require years for their growth.

Dimorphism. In some syngamic species, the valves of male and female are nearly alike. In others, however, the differences are conspicuous. The male may be larger, using the additional volume to house the complicated sex organs, which may constitute one-third of the whole animal; or the female may be larger, using the additional space for egg and brood care. Certain appendages may also be dimorphic, those in the male being modified into pincerlike claspers, chitinous hooks, or suctorial setae for assistance in copulation (Fig. 1*b*, *c*). Patsy A. McLaughlin

Fossils. The fossil record of the Ostracoda is one of the longest and most complete in the animal kingdom. Three aspects contribute to this excellent fossil record. Being microfossils, most with a length of about 0.5–3 mm, the ostracodes are abundant and likely to be buried rapidly, thus enhancing their likelihood of preservation. Most are benthic organisms, and they have dense, heavily calcified carapaces of calcite that are geochemically quite stable. Finally, they have evolved the capacity to live in a very wide variety of environments that range from shallow-marine to abyssal depths, quiet to turbulent water, and even in the fresh-water environments of lakes, ponds, puddles, and streams.

In spite of the excellent fossil record, several problems remain to be solved before the understanding of early Paleozoic Ostracoda can be regarded as satisfactory. The orders Bradoriida and Phosphatocopida, both restricted primarily to Cambrian rocks, are especially problematic. The bradoriids are widely regarded as polyphyletic, and many taxa assigned to the order are certainly not ostracodes at all. Phosphatocopida is marked by carapaces made of calcium phosphate, a feature that, unlike any in other ostracodes, suggests that they too may be unrelated to the Ostracoda. Moreover, two other orders of lower to middle Paleozoic ostracodes present problems as well. Some specialists regard members of the order Eridostracoda as brachiopods, instead,or perhaps

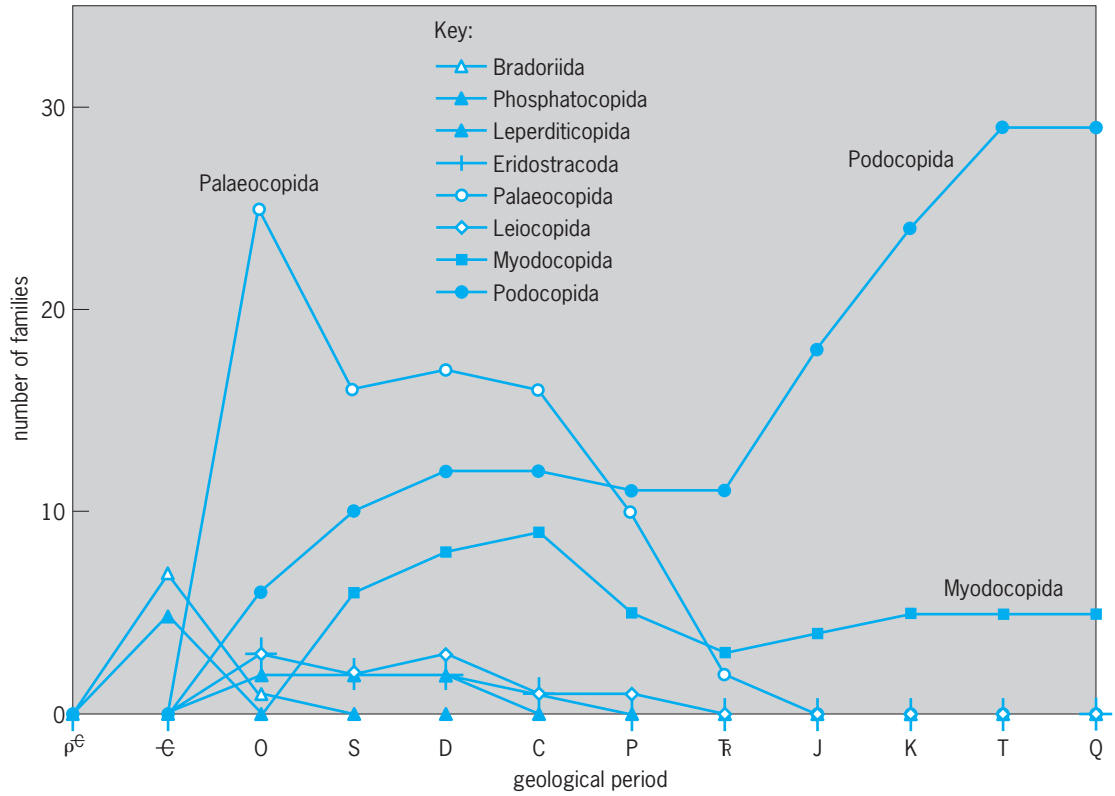

Fig. 2. Biodiversity of the eight orders of Ostracoda showing numbers of families known from the fossil record.

as ostracodes of the order Palaeocopida. Some very large species of the order Leperditicopida are also doubtfully assigned to the class Ostracoda.

Among the most interesting of any ostracodes are the species of the order Palaeocopida, which range from the Ordovician into the Triassic. Many of these ostracodes are marked by pronounced sexual dimorphism. The females develop brood pouches in their carapaces, quite a number of different kinds of which have been described.

The ostracodes that are most common in the modern world, those of the order Podocopida, have been remarkably successful and are found in nearly every marine environment. On at least two separate occasions and perhaps three, the ostracodes have made the difficult transition from marine to freshwater environments. All these invasions of fresh water have been by podocopids. Their occurrence in fresh-water deposits is especially significant because few invertebrate organisms with mineralized skeletons live in fresh water, and the other kinds of invertebrates that do so are likely to have skeletons of aragonite, which tend to recrystallize with much loss of information.

The biodiversities of families known in the fossil record in each of the eight orders of the Ostracoda are shown in **Fig. 2**. Five aspects of the ostracodes' biodiversity are noteworthy. (1) The ostracodes—or some organisms that were confusingly similar to ostracodes—participated in the Cambrian explosion in diversity, but these early orders had become extinct by the middle of the Ordovician Period. (2) Several less diverse orders of Paleozoic ostracodes persisted from the Ordovician to the Permian. (3) The dominant order of Paleozoic ostracodes, the Palaeocopida, reached a peak in familial diversity in the Ordovician and has been declining almost monotonically since, finally becoming extinct by the end of the Triassic Period. (4) The Ostracoda underwent a fundamental change as a result of the end-Permian extinction event, which some paleontologists think may have eradicated as many as 95% of the Earth's marine species. (5) The familial diversity of the Ostracoda in the modern world is due almost entirely to the order Podocopida, which has been steadily increasing in diversity since the Triassic Period. Important to note is the fact that familial diversity does not necessarily provide information about diversity at other taxonomic levels, especially species diversity. *See* CRUSTACEA; PALEOCOPA. Roger L. Kaesler

Bibliography. R. H. Bate, E. Robinson, and L. M. Sheppard (eds.), *Fossil and Recent Ostracods*, 1982; H. G. Bronns, *Klassen und Ordnungen des Tierreichs*, Band 5, *Arthropoda*, 1 Abt.: *Crustacea*, 2 Buch (4), 1966, 1975; R. V. Kesling, Anatomy and dimorphism of adult *Candona suburbana* Hoff, *Four Reports of Ostracod Investigations*, 1965; R. V. Kesling, Terminology of ostracod carapaces, *Contrib. Mus. Paleontol. Univ. Mich.*, 9:93–171, 1951; R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, pt. Q: *Arthropoda 3*, 1961; K. J. Müller, Phosphatocopine ostracods with preserved appendages from the Upper Cambrian of Sweden, *Lethaia*, vol. 1, 1979; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; F. R. Schram (ed.), *Crustacean Issues*, vol. 1: *Crustacean Phylogeny*, 1983; R. N. Smith, Musculature and muscle scars on *Chlamydotheca arcuata* (Sars) and *Cypridopsis vidua* (O. F. Müller), *Four Reports of Ostracod Investigations*, 1965; R. C. Whatley, D. J. Siveter, and I. D. Boomer, Arthropoda (Crustacea: Ostracoda), *in* M. J. Benton (ed.), *The Fossil Record 2*, pp. 343–356 Chapman & Hall, London, 1993.

Ostracoderm

A popular name applied to several groups of extinct jawless vertebrates (fishes). Most of them were covered with an external skeleton or armor of bone, from which is derived their name, meaning "shell-skinned" (see **illus.**). They are known from the Ordovician, Silurian, and Devonian periods, and thus include the earliest known vertebrates. *See* JAWLESS VERTEBRATES. Robert H. Denison

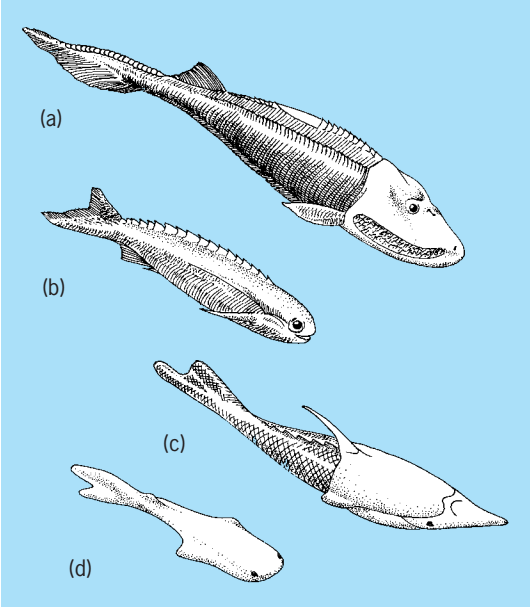

Ostracoderms, drawn to same scale. (*a*) *Hemicyclaspis*, a cephalaspid or Osteostraci. (*b*) *Pterygolepis*, an anaspid. (*c*) *Pteraspis*, a pteraspid or Heterostraci. (*d*) *Logania*, a coelolepid. (*After E. H. Colbert, Evolution of the Vertebrates, Wiley, 1955*)

Otter

A member of the family Mustelidae, which also includes weasels, mink, ferrets, martens, sables, fishers, badgers, and wolverines. Otters are found worldwide except in Australia, Madagascar, and on the islands of the South Pacific. Most inhabit freshwater lakes and rivers. The sea otter (*Enhydra lutris*), however, inhabits the waters and shores of the northern Pacific Ocean from southern California to the Kurile Islands.

General morphology. All otters are much alike no matter where they live. The North American otter (*Lontra canadensis*) has a lithe, muscular body; a broad, flat head; small ears; and a long, powerful, tapering tail which serves as a rudder (see **illustration**). The ears and nostrils are capable of being closed when the otter is underwater. The limbs are short, but the strong hind feet are large, and the five toes are broadly webbed. The short, dense fur is dark brown and is impervious to water. The densely packed underfur consists of approximately 70,000 hairs/cm^2 (450,000 hairs/in.^{2}) and is overlaid by longer guard hairs. Perianal scent glands are present in all species except the sea otter.

Habits and reproduction. Otters are skillful, predators. They usually prey on sluggish varieties of fish, but may also feed on trout, frogs, crayfish, crabs, ducks, muskrats, and young beavers. Otters catch fish with their forepaws, then rip them apart with their teeth. An otter's den is usually a hole in the bank of a stream or lake and has an underwater entrance. A second entrance for ventilation is hidden in bushes on the bank. Except during the breeding season, otters are continually on the move and will travel a 20-mi (32-km) circuit of connecting lakes and rivers in 2 or 3 weeks. Baby otters are born in April or May after a gestation of 60 to 70 days, although gestation in some species with delayed uterine implantation may extend to 10 to 12 months. Average litter size is two. The eyes open at about 4 weeks of age. Normal lifespan in the wild is 8 or 9 years.

Sea otters. Sea otters are among the largest of all otters. Adults may be 4 to 5ft (1.2 to 1.5 m) in length, including a 12-in. (30.5-cm) tail. They may weigh up to 80 pounds. The dental formula is 1 3/2, C 1/1, Pm 3/3, and M 1/2, for a total of 32 teeth. The heavy, thickset body is much less sleek and graceful in appearance than the river otter. The pelt is very full, soft, and deep. It is brownishish-black and more or less finely grizzled. *See* DENTITION.

Sea otters rarely come to land, preferring to live in great beds of floating kelp, a type of brown seaweed. Individuals are born, eat, sleep, grow old, and die in the water. They prefer to float on their back and propel themselves with their tail. For greater speed, they turn right side up and, by undulating their body and using their webbed hind feet, they may travel at a rate of 10 mi (16 km) an hour. Sea otters feed on fish, sea urchins, crustaceans, cuttlefish, mussels, clams, abalone, and other shellfish. They often dive to depths of 30 m (100 ft), with the record dive being 97 m (318 ft). Food is normally spread out on the belly and chest and eaten as the sea otter floats on its back. A flat stone is often carried to the surface, placed on the chest, and used as an anvil to open shell-encased prey.

Sea otters are believed to mate for life. Following a gestation of 9 months, a single pup is born on a thick bed of floating kelp, usually in a sheltered natural harbor. The eyes are open at birth, unlike those of other species of otters. Pups remain with their mother for 6 to 12 months and are often carried on their mother's chest where they can nurse and be cleaned. They are full-grown at 4 years of age. Wild sea otters may live up to 20 years.

North American river Otter (*Lontra canadensis*). (*Alan and Sandy Carey/Getty Images*)

Sea otters have few natural enemies, chief among them being the killer whale. However, this species was almost driven to extinction in the late nineteenth and early twentieth centuries by overhunting for its valuable pelt. It now enjoys complete international protection as the result of a 1911 agreement between the United States, Russia, Japan, and Great Britain. Populations now number approximately 150,000.

Other otter species. Other otters include the giant otter (*Pteronura*) that is native to the Amazon River basin of Brazil and reaches a length of about 5 ft (1.5 m), the giant African otter (*Aonyx*) which may be longer than 5 ft and weigh 60 pounds, and whose claws are rudimentary or absent altogether; the clawless otter (*Paraonynx*) of Africa whose forefeet are small with five naked fingers without claws and whose hindfeet bear minute claws on only the third and fourth fingers; the hairy-nosed otter (*Lutra sumatrana*) of Asia and Sumatra, whose nose pad, naked in most species, is covered with fine hair; and the Eurasian otter (*Lutra lutra*), which ranges from England west to southern China and from the Mediterranean region of Africa north to the Arctic coast. *See* CARNIVORA; MAMMALIA. Donald W. Linzey

Bibliography. H. Kruuk, *Wild Otters: Predation and Populations*, Oxford University Press, 1995; C. F. Mason and S. M. MacDonald, *Otters: Ecology and Conservation*, Cambridge University Press, 1986; R. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford Limited, 2001.

Otto cycle

The basic thermodynamic cycle for the prevalent automotive type of internal combustion engine. The engine uses a volatile liquid fuel (gasoline) or a gaseous fuel to carry out the theoretic cycle illustrated in **Fig. 1**. The cycle consists of two isentropic (reversible adiabatic) phases interspersed between two

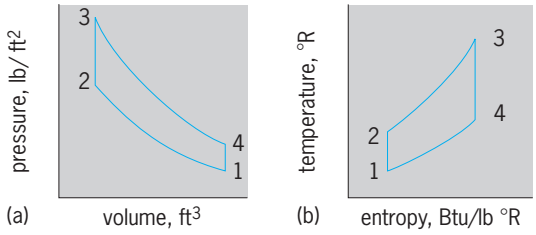

Fig. 1. Diagrams of (*a*) pressure-volume and (*b*) temperature-entropy for Otto cycle; phase 1–2. isentropic compression; phase 2–3, constant-volume heat addition; phase 3–4, isentropic expansion; phase 4–1, constant volume heat rejection.

constant-volume phases. The theoretic cycle should not be confused with the actual engine built for such service as automobiles, motor boats, aircraft, lawn mowers, and other small (generally <300 ± hp or 225 ± kW) self-contained power plants.

The thermodynamic working fluid in the cycle is subjected to isentropic compression, phase 1–2; constant-volume heat addition, phase 2–3; isentropic expansion, phase 3–4; and constant-volume heat rejection (cooling), phase 4–1. The ideal performance of this cycle, predicated on the use of a perfect gas, Eqs. (1), is summarized by Eqs. (2) and (3) for thermal efficiency and power output.

$$\frac{V_3}{V_2} = \frac{V_4}{V_1} \qquad \frac{T_3}{T_2} = \frac{T_4}{T_1}$$
$$\frac{T_2}{T_1} = \frac{T_3}{T_4} = \left(\frac{V_1}{V_2}\right)^{k-1} = \left(\frac{V_4}{V_3}\right)^{k-1} = \left(\frac{P_2}{P_1}\right)^{(k-1)k} \tag{1}$$

$$\text{Thermal eff} = \frac{\text{net work of cycle}}{\text{heat added}}$$
$$= \left[1 - \frac{T_1}{T_2}\right] = \left[1 - \left(\frac{1}{r^{k-1}}\right)\right] \tag{2}$$

$$\begin{aligned}\text{Net work of cycle} &= \text{heat added} - \text{heat rejected}\\ &= \text{heat added} \times \text{thermal eff}\\ &= \text{heat added}[1 - (T_1/T_2)]\\ &= \text{heat added}[1 - (1/r^{k-1})]\end{aligned} \tag{3}$$

In Eqs. (2) and (3) V is the volume in cubic feet; P is the pressure in pounds per square inch; T is the absolute temperature in degrees Rankine; k is the ratio of specific heats at constant pressure and constant volume, C_p/C_v; and r is the ratio of compression, V_1/V_2.

The most convenient application of Eq. (3) to the positive displacement type of reciprocating engine uses the mean effective pressure and the horsepower equation (4), where hp is horsepower; mep is mean

$$\text{hp} = \frac{\text{mep}\ Lan}{33,000} \tag{4}$$

effective pressure in pounds per square foot; L is stroke in feet; a is piston area in square inches; and n is the number of cycles completed per minute. The mep is derived from Eq. (3) by Eq. (5), where 778 is

$$\text{mep} = \frac{\text{net work of cycle} \times 778}{144(V_1 - V_2)} \tag{5}$$

the mechanical equivalent of heat in foot-pounds per Btu; 144 is the number of square inches in 1 ft^2; and $(V_1 - V_2)$ is the volume swept out (displacement) by the piston per stroke in cubic feet. *See* MEAN EFFECTIVE PRESSURE.

Air standard. In the evaluation of theoretical and actual performance of internal combustion engines, it is customary to apply the above equations to the idealized conditions of the air-standard cycle. The working fluid is considered to be a perfect gas with such properties of air as volume at 14.7 lb/in.^{2} absolute and 492°R equal to 12.4 ft^3/lb, and the ratio of specific heats k as equal to 1:4. **Figure 2** shows the thermal efficiency for this air-standard cycle as a function of the ratio of compression r, and the mep for a heat addition of 1000 Btu/lb of working gases. These curves demonstrate the intrinsic worth of high compression in this thermodynamic cycle.

The **table** gives a comparison of the important gas-power cycles on the ideal air-standard base for the case of compression ratio = 10 and 1000 Btu added per pound of working gases. The Otto, Brayton, and Carnot cycles show the same thermal efficiency of 60%. The mean effective pressures, however, show that the physical dimensions of the engines will be a minimum with the Otto cycle but hopelessly large with the Carnot cycle. The Brayton cycle ideal mep is only about one-fifth that of the Otto cycle, and it is accordingly at a distinct disadvantage when applied to a positive displacement mechanism. This disadvantage can be offset by use of a free-expansion, gas-turbine mechanism for the Brayton cycle. The Diesel cycle offers a lower thermal efficiency than the Otto cycle for the same conditions, for example, 42 versus 60%, and some sacrifice of mep, 200 versus 290 lb/in.^{2} As opposed to the Otto engine, the diesel can utilize a much higher compression ratio without preignition troubles and without excessive peak pressures in the cycle and mechanism. Efficiency and

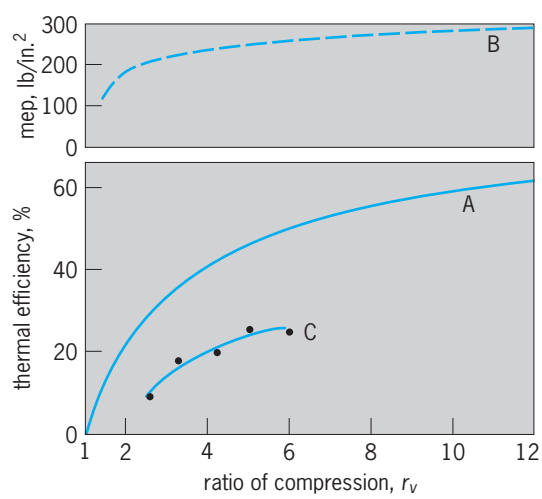

Fig. 2. Effect of compression ratio on thermal efficiency and mean effective pressure of Otto cycle. Curve *A* shows thermal efficiency, air-standard cycle; curve *B* mean effective pressure, air-standard cycle; and curve *C* thermal efficiency of an actual engine.

Thermal eficiency, mean effective pressure, and peak pressure of air-standard gas-power cycles*

Cycle	Efficiency	mep	Peak pressures lb/in.^{2}†
Otto	60	290	2100
Diesel	42	200	370
Brayton	60	61	370
Carnot	60	Impossibly small	Impossibly high

*Ratio of compression = 10; heat added = 1000 Btu per pound working gases (2.3 megajoules per kilogram).
†1 lb/in.^{2} = 6.895 kilojoules.

engine weight are thus nicely compromised in the Otto and diesel cycles. *See* GAS TURBINE.

Actual engine process. This reasoning demonstrates some of the valid conclusions that can be drawn from analyses utilizing the ideal air-standard cycles. Those ideals, however, require the modifications of reality for the best design of internal combustion engines. The actual processes of an internal combustion engine depart widely from the air-standard cycle. The actual Otto cycle uses a mixture of air and a complex chemical fuel which is either a volatile liquid or a gas. The rate of the combustion process and the intermediate steps through which it proceeds must be established. The combustion process shifts the analysis of the working gases from one set of chemicals, constituting the incoming explosive mixture, to a new set representing the burned products of combustion. Determination of temperatures and pressures at each point of the periodic sequence of phases (Fig. 1) requires information on such factors as variable specific heats, dissociation, chemical equilibrium, and heat transfer to and from the engine parts.

N. A. Otto (1832–1891) built a highly successful engine that used the sequence of engine operations proposed by Beau de Rochas in 1862. Today the Otto cycle is represented in many millions of engines utilizing either the four-stroke principle or the two-stroke principle. *See* INTERNAL COMBUSTION ENGINE.

The actual Otto engine performance is substantially poorer than the values determined by the theoretic air-standard cycle. An actual engine performance curve *c* is added in Fig. 2, in which the trends are similar and show improved efficiency with higher compression ratios. There is, however, a case of diminishing return if the compression ratio is carried too far. Evidence indicates that actual Otto engines offer peak efficiencies (25±%) at compression ratios of 15±. Above this ratio, efficiency falls. The most probable explanation is that the extreme pressures associated with high compression cause increasing amounts of dissociation of the combustion products. This dissociation, near the beginning of the expansion stroke, exerts a more deleterious effect on efficiency than the corresponding gain from increasing compression ratio. *See* BRAYTON CYCLE; CARNOT CYCLE; DIESEL CYCLE; THERMODYNAMIC CYCLE.

Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; J. B. Jones and R. E. Dugan, *Engineering Thermodynamics*, 1994; E. F. Obert, *Internal Combustion Engines and Air Pollution*, 1990.

Ovarian disorders

A variety of neoplastic and nonneoplastic disorders that occur in the ovary. Ovarian neoplasms are of greater diversity in histologic appearance and biologic behavior than for any other organ. The nonneoplastic disorders include physiologic cysts, pregnancy luteomas, and polycystic ovarian disease (Stein-Leventhal syndrome). The ovary can also be a site of metastasis from malignant tumors originating in the genital tract, breast, and gastrointestinal tract. *See* CANCER (MEDICINE).

Symptoms. As ovarian enlargement due to tumors occurs, compression of pelvic and abdominal structures produces vague symptoms such as constipation, pelvic discomfort, a feeling of heaviness, and frequent urination. If the size of the ovary exceeds 5 in. (12 cm), the ovary rises out of the pelvis, and the individual may notice abdominal enlargement, sometimes mistakenly attributed to pregnancy or weight gain. Pain can be an initial symptom of both benign and malignant ovarian disorders. The symptoms of malignant ovarian disorders include abdominal pain and swelling, bloating, heartburn, nausea, and anorexia. Because these symptoms are more suggestive of gastrointestinal problems, these individuals are often first evaluated by internists and family practitioners before the ovary is identified as the cause of symptoms.

Physical findings. An examination of the pelvis is the most crucial component in the diagnosis and evaluation of ovarian disorders. Although a mass may be felt during examination of the vagina and abdomen, palpation of the pelvic organs via the rectum is more informative. This often allows the physician to differentiate between disorders of the ovary and other pelvic structures such as the uterus, rectum, and bladder. The location of abnormal ovarian enlargement may offer a clue to the diagnosis as benign cystic teratoma or cysts of endometriosis; most ovarian tumors lie posterior to the uterus. The addition of ultrasound to the pelvic examination may help differentiate benign from malignant disorders. Unilateral, cystic, mobile masses less than 4 in. (10 cm) in size likely indicate benign disease; solid, immobile masses greater than 4 in. (10 cm) likely indicate malignancy.

Benign disorders. Physiologic cysts occur in response to cyclic stimulation of the ovary by hormones. Simple cysts in women of reproductive age which are greater than about 1.2 in. (3 cm) can often be treated with oral contraceptives. Persistence of the cyst for 2 months warrants surgical exploration with removal of the cyst, leaving the ovary intact.

Polycystic ovarian syndrome. This syndrome is caused by abnormal regulation of the hypothalamic-pituitary-ovarian axis. In this condition, the ovaries are bilaterally enlarged with multiple follicular cysts. Clinically, symptoms include amenorrhea, menstrual abnormalities, infertility, and hirsutism (a condition characterized by the growth of hair in unusual places and in unusual amounts). Therapy consists of administering oral contraceptives, progesterone preparations, or clomiphene. Because this is a condition where there is no ovulation for prolonged periods, individuals with this syndrome are at increased risk for hyperplasia and carcinoma of the endometrium.

Endometriosis. This disorder is characterized by ectopic endometrial glandular and stromal tissue that often involves the ovary as well as other pelvic structures. Ovarian endometriosis often produces a cystic mass filled with old blood termed a chocolate cyst. The two most common symptoms of endometriosis are pelvic pain and infertility. This can be treated medically with agents that suppress ovulation, or surgically by removing the ovary.

Neoplasms. Neoplasms of the ovary can originate from epithelial, stromal, or germ cells, and produce a variety of symptoms. Epithelial tumors account for 50–65% of benign neoplasms prior to menopause and 80% after menopause. Mature cystic teratomas account for 50% of benign neoplasms in women between the ages of 20 and 50, and 25–40% of all ovarian neoplasms. Treatment is tailored to the individual's age and desire for future fertility. In a reproductive-age woman who has not completed childbearing or desires conservative therapy, removal of the cyst or one ovary is adequate treatment. In the individual who has completed childbearing or is peri- or postmenopausal, removal of both ovaries is acceptable.

Malignant disorders. As with benign tumors, malignant ovarian neoplasms can arise from epithelial, stromal, or germ cells. Epithelial tumors account for 85% of ovarian cancer, while germ-cell and stromal tumors account for 10 and 5%, respectively. With a high suspicion of malignancy, useful diagnostic tests include serum tumor markers (CA-125, human chronic gonadotropin, alpha-fetoprotein), transvaginal color doppler imaging, chest x-ray, and computerized tomography scan. The primary treatment is surgery; procedures include removal of the ovaries, fallopian tubes, uterus, omentum, and other tumor masses within the abdominal cavity. Survival is directly related to the amount of cancer remaining at the end of surgery. Many individuals receive adjuvant therapy with chemotherapy following surgery. *See* MEDICAL IMAGING.

The ovary is often a site of metastatic cancer from other organs, the most common sources being breast, gastrointestinal tract, and other pelvic organs. Up to 20% of women undergoing removal of an ovary for breast cancer have been found to have metastasis. *See* ONCOLOGY; OVARY; TUMOR. David L. Tait

Bibliography. C. P. Morrow, J. P. Curtin, and D. E. Townsend (eds.), *Synopsis of Gynecologic Onco\-logy*, 5th ed., 1998.

Ovary

A part of the reproductive system of all female vertebrates. Although not vital to individual survival, the ovary is vital to perpetuation of the species. The function of the ovary is to produce the female germ cells or ova, and in some species to elaborate hormones that assist in regulating the reproductive cycle.

The ovaries develop as bilateral structures in all vertebrates, but adult asymmetry is found in certain species of all vertebrates from the elasmobranchs to the mammals. This asymmetry may be morphological as a result of fusion or atrophy, or it may be physiological.

The position of the ovaries in the coelom varies within different vertebrate groups. Those animals which produce large numbers of eggs possess ovaries that almost fill the coelomic cavity during the breeding season, but most vertebrates have relatively small ovaries. The ovaries may be anchored to the dorsal body wall anywhere between the transverse septum (lower vertebrates) and the pelvic cavity (higher mammals).

Histology

The ovary of all vertebrates functions in essentially the same manner. However, ovarian histology of the various groups differs considerably. Even such a fundamental element as the ovum exhibits differences in various groups. The ovum of oviparous forms is large, and it synthesizes and stores large amounts of yolk. The ovum of viviparous forms, especially the mammals, is small and contains little yolk.

Mammals. The mammalian ovary is attached to the dorsal body wall by a mesovarium. The free surface of the ovary is covered by a modified peritoneum called the germinal epithelium (**Fig. 1***a*). This layer may consist of a single or a stratified layer of cells, and the cell shape may vary from squamous to columnar, depending upon the species. The activity of this layer, especially during the embryonic development of the gonad whereby cells are proliferated into the interior, is responsible for its name. The potentialities of the germinal epithelium are still a matter of controversy.

Just beneath the germinal epithelium is a layer of fibrous connective tissue, varying in thickness and density in different species and at different phases of the reproductive cycle in the same species. This is the tunica albuginea (Fig. 1*a*). Most of the rest of the ovary is made up of a more cellular and more loosely arranged connective tissue (stroma) in which are embedded the germinal, endocrine, vascular, and nervous elements (Fig. 1*a* and *b*).

The most obvious ovarian structures are the follicles and the corpora lutea. The smallest, or primary, follicle consists of an oocyte surrounded by a layer of follicle (nurse) cells (Fig. 1*a* and *b*). Follicular growth results from an increase in oocyte size, multiplication of the follicle cells to form several concentric layers called the granulosa, and the differentiation of the perifollicular stroma to form a fibrocellular

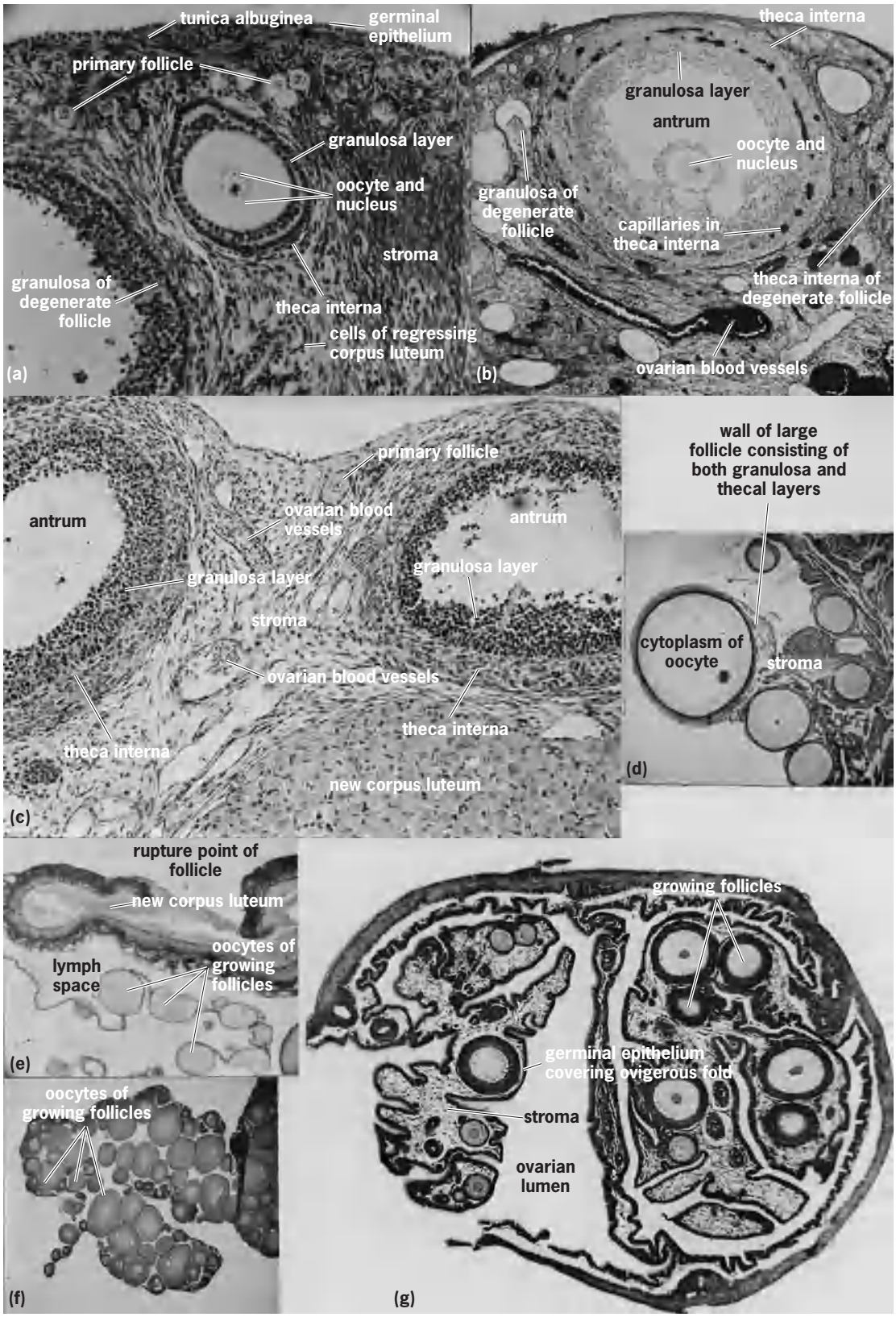

Fig. 1. Ovarian histology. (*a*) Pika (*Ochotona princeps*); (*b*) pocket gopher; (*c*) prairie dog (*Cynomys leucurus*); (*d*) chicken (*Gallus domesticus*) (*from W. Andrew, Textbook of Comparative Histology, Oxford, 1959*). (*e*) Lizard (*Xantusia vigilis*) (*courtesy of M. Miller*). (*f*) Frog (*Rana* sp.) (*from W. Andrew, Textbook of Comparative Histology, Oxford, 1959*). (*g*) Viviparous fish (*Neotoca bilineata*) (*from G. Mendoza, Biol. Bull., vol. 84, 1943*).

envelope called the theca interna (Fig. 1*b* and *c*). Finally, a fluid-filled antrum develops in the granulosa layer, resulting in a vesicular follicle (Fig. 1*b*).

The cells of the theca interna hypertrophy during follicular growth and many capillaries invade the layer (Fig. 1*b*), thus forming the endocrine element that is thought to secrete estrogen. The other known endocrine structure is the corpus luteum (Fig. 1*c*), which is primarily the product of hypertrophy of the granulosa cells remaining after the follicular wall ruptures to release the ovum. Ingrowths of connective tissue from the theca interna deliver capillaries to vascularize the hypertrophied follicle cells of this newly formed corpus luteum; progesterone is secreted here.

Most follicles degenerate before attaining ovulatory size (Fig. 1*a*, *b*, and *c*). In some species the theca interna of such follicles hypertrophies to form the so-called interstitial cells (Fig. 1*b*). The function of this tissue is still uncertain. However, evidence is accumulating that suggests the tissue is involved in steroid secretion.

Birds. The avian ovary has much the same general arrangement as that of the mammal, the chief difference being the size of the mature ova. Other differences are a greatly reduced amount of stroma, the thinness of the theca interna, the reduced number of follicle-cell layers, and the absence of follicular antra and corpora lutea (Fig. 1*d*).

Reptiles and amphibians. The reptilian ovary is quite similar in appearance to that of the bird. The follicles are similar, and some viviparous forms are reported to have corpora lutea, although the physiology of them is uncertain. Centrally placed, epithelial-lined spaces, said to be lymphatic spaces, are present in reptilian ovaries (Fig. 1*e*).

The amphibian ovary is similar to the reptilian type but contains many more smaller follicles. No corpora lutea are formed (Fig. 1*f*).

Fishes. The fishes represent a diverse group as far as ovarian morphology is concerned. Many forms have an arrangement similar to that of the amphibians. Others have the germinal epithelium covering ovigerous folds within a centrally placed ovarian lumen, instead of covering the outer surface (Fig. 1*g*).

Kenneth L. Duke

Physiology

Ovarian tissue is responsible for the production of germ cells (ova) and hormones. Lifetime production of ova is variable and may range from about 400 in the human female to millions in some fishes. Ova typically mature within a follicle, and these germ cells are then periodically released from the ovary (ovulation). In many vertebrates ova pass through accessory ducts, where they receive coverings such as jelly capsules or shells. Fertilization of ova can occur within the ducts or externally. In higher mammals, some reptiles, and elasmobranchs the fertilized ovum develops within an enlarged region of the duct system called the uterus. In mammals this development period (gestation) requires the formation and function of a temporary structure termed a placenta. In some viviparous fish, development of fertilized ova occurs within the ovary.

Steroid synthesis. Several endocrine organs (ovary, testis, adrenal gland, and placenta) have the ability to produce various steroid hormones. The main biosynthetic pathway is similar in all steroidogenic tissues. A simplified scheme is shown below. Each step in the

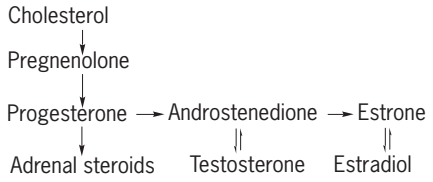

pathway requires the action of one or more enzymes. The presence of the appropriate enzymes and their relative activities can therefore determine which steroid is produced. Ovarian follicular cells possess all the enzymes necessary for estradiol production, whereas the corpus luteum is unable to carry on efficiently the reactions beyond progesterone in the pathway. Tropic hormones from the pituitary, which act on particular steroid-producing tissues, also control the rate and direction of steroid synthesis. *See* ADENOHYPOPHYSIS HORMONE; CHOLESTEROL.

Ovarian hormones. Hormones produced by the ovary maintain reproductive organs and their accessory structures, and they control cyclic events such as estrus, nidation of the fertilized ovum, and parturition. Ovarian hormones are primarily steroid hormones, but in mammals at least one peptide hormone, relaxin, is produced. Relaxin appears to be involved in enhancing the separation of the symphysis pubis, which brings about an enlargement of the birth canal. Relaxin may also stimulate dilation of the uterine cervix.

Steroids elaborated by the ovary are of two principal types; estrogens which produce sexual receptivity (heat) and growth of the uterus and mammary glands; and progesterone which augments estrogen-stimulated growth of the uterus and mammary gland. *See* ESTROGEN; PROGESTERONE; STEROID.

The production of estrogen and progesterone are controlled by follicle-stimulating hormone (FSH) and luteinizing hormone (LH) from the pituitary gland. Estradiol is the most potent natural estrogen and is the major hormone synthesized by granulosa cells of the maturing follicle. Estradiol may be converted to estrone (**Fig. 2**); and estriol, a weak estrogen, is produced as a metabolic end product. Estradiol and estrone have been found in the ovaries of representative species of all vertebrate classes, and estradiol has also been detected in starfish ovaries.

Progesterone (**Fig. 3**) is produced by the ovarian corpus luteum which develops from the ruptured follicle after ovulation. Progesterone production and secretion are stimulated by pituitary LH. The liver converts large amounts of progesterone to the inactive form, pregnanediol, which is excreted in urine. Progesterone, like estradiol, has been found in the ovaries of all vertebrate classes.

Ovarian steroid action. Organs such as the uterus and mammary glands are dependent upon ovarian

estrogens and progesterone for their functional status. These tissues contain receptor molecules (proteins of ~75,000 molecular weight) with high affinities for estradiol and progesterone, a property that allows these tissues (targets) to remove the ovarian steroids from circulating blood. When receptors bind estradiol, the receptor-estradiol complex translocates rapidly to target cell nuclei where specific genes may be activated. As a consequence of its action, estradiol causes a general anabolic response in uterine tissue, which includes increases in a number of biochemical events such as glucose metabolism, lipid synthesis, and ribonucleic acid (RNA) and protein synthesis. Further development of uterine tissue requires the action of progesterone, which is again mediated by a progesterone-receptor complex. This steroid is responsible for the decidual reaction, which is related to the phenomenon of implantation of the fertilized ovum in the uterus. Estrogens and progesterone are also responsible for the development of ducts in mammary glands. *See* LACTATION; MAMMARY GLAND; UTERUS.

Hypothalamic-pituitary-ovarian relationships. In mammals an intricate relationship between hypothalamic gonadotropin-releasing hormone (GnRH), pituitary gonadotropins (FSH and LH), and ovarian steroids is essential for normal reproductive function. GnRH promotes release of FSH and LH, and these pituitary hormones stimulate steroid production by the ovary. The ovarian steroids in turn regulate GnRH and gonadotropin levels and effectiveness by positive- or negative-feedback mechanisms (**Fig. 4**).

Maturation of the follicle requires the action of LH and FSH at precise times. LH binds to receptors on follicular theca cells and stimulates testosterone synthesis. This testosterone is converted to estrogen by follicular granulosa cells under the influence of FSH, which also acts to increase LH receptors on follicular cells. Increasing amounts of LH, including a critically timed LH "surge," bring about ovulation which terminates the follicular phase. The empty follicle becomes transformed into a corpus luteum which responds to FSH and LH by secreting progesterone in addition to estrogens. In rodents, a third gonadotropin, luteotropin (LTH or prolactin) participates in stimulating progesterone secretion.

Circulating estrogen and progesterone act on the pituitary (negative feedback) to inhibit FSH and LH production, and without these hormones the corpus luteum degenerates. As steroid levels fall, FSH and LH secretion resume, and the ovarian cycle may begin again. Although estrogen normally acts to inhibit FSH and LH production, high levels of estrogen at the peak of follicular growth stimulate GnRH secretion and sensitize pituitary cells to GnRH, thus causing the LH surge (positive feedback). In addition to regulation by ovarian steroids, the hypothalamus may also be controlled by other brain centers. Thus, extrinsic environmental stimuli, such as light and temperature, may influence the timing of reproductive cycles by starting a chain of events leading to ovarian activity. *See* ENDOCRINE MECHANISMS.

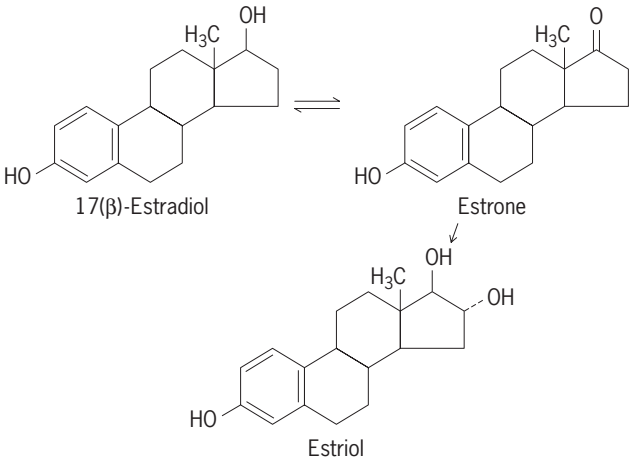

Fig. 2. Structural formulas of the principal natural estrogens.

Sexual cycles. Nonmammalian vertebrates generally exhibit seasonal breeding cycles, and their pattern of ovarian function is quite variable. For example, avian and amphibian ovaries do not form true corpora lutea. In most mammals a period of sexual arousal called estrus typically occurs during the reproductive cycle. The recurrence of estrus (estrus cycle) is variable in mammalian species: for example, rat, 4–5 days; dog, 3–4 months; sheep, 16 days; and pig and cow, 21 days. The estrus period corresponds to the time of ovulation and therefore increases the likelihood that mating and fertilization of ova will take place. In some species ovulation only occurs when stimulated by mating. The sexual cycle in human females (menstrual cycle) lasts for about 28 days and is terminated by the degeneration and loss of the uterine lining (menstruation). This occurs because the corpus luteum of the cycle regresses, resulting in the withdrawal of estrogen and progesterone. Ovulation occurs on about the 13th or 14th day of the menstrual cycle. *See* ESTRUS; MENSTRUATION.

Pregnancy. If the ovum is fertilized, a state of pregnancy will begin. Estrogen and progesterone produced during the cycle promote uterine development, which is conducive to the growth of an embryo. In mammals, a fertilized ovum implants in the uterine wall and begins the formation of a placenta. The placenta helps to maintain pregnancy in some species by its role in prolonging the lifetime of the corpus luteum and also through placental production of progesterone. The corpus luteum

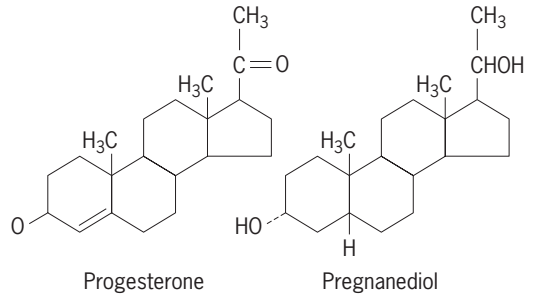

Fig. 3. Active and inactive forms of pregnancy hormone.

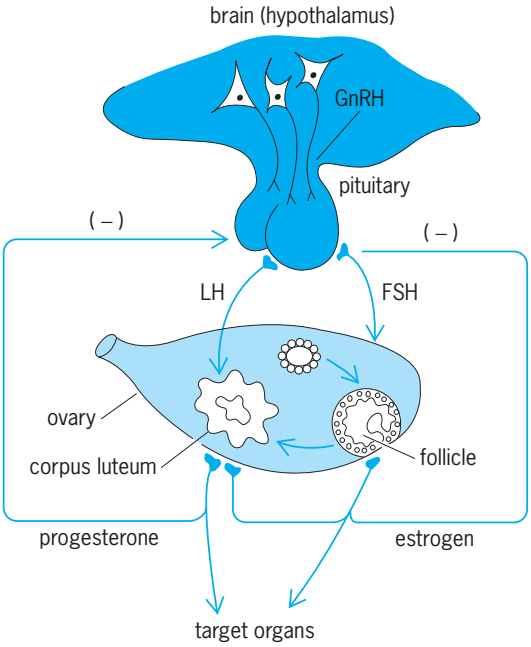

Fig. 4. Hormonal interactions of the hypothalamus, pituitary, and ovary.

persists during pregnancy, and its secretion of progesterone is important in maintaining the early stages of pregnancy. Later stages are dependent upon increased progesterone secretion from the placenta. Progesterone dominance of the uterus during pregnancy keeps the tissue quiescent and promotes its growth. At the end of the gestation period progesterone dominance wanes, uterine muscle begins to contract rhythmically under the influence of estrogen, and the fetus is expelled (parturition). *See* PLACENTATION; PREGNANCY.

Fertility regulation. In cases where human fertility is low because of anovulatory cycles, it has been possible to induce ovulation by administering pituitary gonadotropins. Because of difficulties in controlling gonadotropin doses, however, these treatments frequently result in multiple births. In domestic animal breeding, the administration of GnRH, FSH, and LH at precisely controlled times shows great promise for increasing both numbers of offspring per pregnancy and numbers of pregnancies per lifetime.

Oral contraceptive pills have proved to be extremely effective in decreasing human fertility. These pills contain synthetic progestagens and estrogens which are effective orally because they are not inactivated by the liver. These "combination" pills are taken daily from days 5–24 of the menstrual cycle, and menstruation will usually occur on day 28. An alternative form is the "minipill," which contains only a small quantity of a synthetic form of progesterone. Although less effective than the combination pill, the minipill avoids side effects caused by estrogens. The synthetic steroids in the pills inhibit or alter gonadotropin secretion at the level of the hypothalamus or pituitary. They thus interfere with maturation of follicles and abolish the LH surge which causes ovulation. It is also possible that these compounds affect implantation, fertilization of ova, or the transport of ova. Although rare instances of abnormal blood clotting have been observed in some users of these pills, this method of contraception is considered to be quite safe for most women after a physician's evaluation. *See* BIRTH CONTROL; REPRODUCTIVE SYSTEM.

Donald E. Smith

Bibliography. E. Y. Adashi and P. C. Leung (eds.), *The Ovary*, 1993; A. Altcheck and L. Deligdisch, *Ovarian Disorders: Pathology, Diagnosis, and Management*, 1995; S. I. Fox, *Human Physiology*, 7th ed., 2001; V. Hayssen and A. Van Tienhoven, *Asdell's Patterns of Mammalian Reproduction: A Compendium of Species—Specific Data*, 1993; E. Knobil et al., *The Physiology of Reproduction*, vols. 1–2, 2d ed., 1994; C. G. Scanes and P. K. Pang (eds.), *The Endrocrinology of Growth, Development, and Metabolism in Vertebrates*, 1992; R. L. Stouffer (ed.), *The Primate Ovary*, 1988; J. Tepperman, *Metabolic and Endocrine Physiology*, 1980.

Overvoltage

The difference between the electrical potential of an electrode or cell under the passage of current and the thermodynamic value of the electrode or cell potential under identical experimental conditions in the absence of electrolysis; it is also known as overpotential. Overvoltage is expressed in volts, often in absolute value; it is a measure of the rates of the different processes associated with an electrode reaction.

An understanding of the factors that contribute to the overvoltage is important in the operation of practical electrochemical systems. In batteries, the overvoltage plays a significant role in the available voltage and power. In large-scale industrial electrolysis, overvoltage is a major factor in determining the energy efficiency of a process, and hence, the cost of electricity. *See* DECOMPOSITION POTENTIAL; ELECTROCHEMICAL PROCESS; ELECTRODE POTENTIAL; ELECTROLYSIS.

Basic concepts. Since the overvoltage is governed by kinetic considerations, all of the experimental conditions that can affect the rate of an electrolytic reaction are of importance. These include concentration of electrolyzed substance, temperature, composition of solvent and electrolyte, nature of the electrode surface, mode of mass transfer, and the current density (current per unit area of electrode). A rapid reaction, such as the reduction of mercurous ion on a mercury electrode, occurs with a small overvoltage (a few millivolts). A slow reaction, for example, the reduction of protons or water on a mercury electrode to produce hydrogen gas, requires a large overvoltage (a few volts).

The rates of electrode reactions are frequently determined from current density–potential curves (**Fig. 1**). Since the departure of the electrode or cell potential from the thermodynamic value upon passage of current is sometimes termed polarization, such curves are also known as polarization

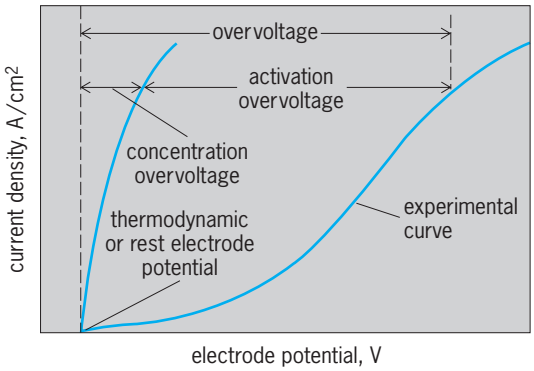

Fig. 1. Current density–potential curve.

curves. These curves are obtained by measurements with three-electrode electrolytic cells. The current density that flows through the electrode of interest (the working electrode) is adjusted with an external direct-current power supply, and the potential of this electrode is measured with respect to a reference electrode whose potential is known and fixed. The measured potential contains a contribution, known as the ohmic drop, that results from the flow of current through the solution resistance between the working electrode and the reference electrode. This drop is minimized by placing these electrodes close together and using various experimental approaches. The thermodynamic potential is obtained from available data for the electrode reaction of interest, corrected for the concentration of the reaction species in the solution under the experimental conditions. It is also sometimes given by the working electrode potential in the electrolysis cell when no current flows (the rest potential). The overvoltage can be read directly from the current–potential curve, as shown in Fig. 1. *See* ELECTRODE; REFERENCE ELECTRODE.

Components. The total overvoltage can be decomposed into different components, which are assigned to different sources of rate limitations, for example, concentration overvoltage, activation overvoltage, reaction overvoltage, and crystallization overvoltage.

Concentration overvoltage. This type of overvoltage occurs when the concentration of the reactants or products at the electrode surface are different from those in the bulk solution. These differences arise because the electroactive reactant is consumed, and products are produced by the passage of current. An example is an aqueous solution containing equal concentrations of iron(III) and iron(II) ions. The rest potential of a platinum electrode immersed in this solution will be near the thermodynamic formal potential for these species. When a cathodic current is passed through the platinum electrode, iron(III) is reduced to iron(II), and the concentrations of these species will be changed near the electrode surface; the concentration of iron(III) increases and that of iron(II) decreases. This causes the potential of the electrode to be more negative than the rest potential. This deviation of potential is known as concentration polarization, and its extent is the concentration overvoltage. The magnitude of this overvoltage depends strongly on the mass-transfer rates existing in the electrolysis cell and is smaller under vigorous stirring conditions. Mathematical treatments are sometimes available to account for transport by diffusion and convection in such cells that allow calculation of the concentration overvoltage. Concentration polarization is the only significant contributor to the overvoltage when all other steps in the electrode reaction are rapid and remain essentially at equilibrium. Such electrode reactions are known as reversible or nernstian reactions.

Activation overvoltage. This component of the overvoltage arises from slowness in the rate of the electron transfer reaction at the electrode surface. To drive an electrode reaction at a given rate (current density), it is necessary to overcome an energy barrier, the energy of activation for the reaction. The additional energy (that is, beyond the thermodynamic requirements) needed to overcome this barrier is provided by the electrical energy supplied to the cell in the form of an increase in the applied potential. The magnitude of this activation overvoltage often depends upon the nature of the electrode material. For example, the reduction of hydrogen ion to elemental hydrogen is sluggish and occurs with a high activation overvoltage at mercury and lead electrodes, but it is quite rapid and occurs with lower overvoltage at platinum. A substance that decreases the activation overvoltage is known as an electrocatalyst.

The theoretical treatment of the relation between potential and current density for an electrode process with a single rate-determining step is based on a model that considers the overall reaction in terms of a forward (reduction) and backward (oxidation) reaction. The forward-reaction contribution to the overall current density is a cathodic one, and the backward-reaction contribution is an anodic current density (**Fig. 2***a*). The net observed current density at any potential is the algebraic sum of these two components. At the rest potential the net current is zero. The magnitudes of the anodic and cathodic current density components under these conditions are equal in absolute magnitude but opposite in sign. This magnitude of a current density component at the rest potential is known as the exchange current density. The exchange current density is generally expressed in units of amperes per square centimeter (A/cm^2) and is directly related to the rate constant for the electrode reaction; a large exchange current density represents a rapid reaction with a large rate constant. The exchange current density cannot be measured directly, but it can be obtained from an analysis of the current-potential behavior, as described below. Both the cathodic and anodic components of the current density vary exponentially with the overvoltage. The cathodic component increases and the anodic one decreases as the overvoltage becomes more negative, while the anodic component increases and the cathodic one decreases as the overvoltage becomes more positive, as shown in Fig. 2*a*. Reactions whose cathodic and anodic components are both appreciable at potentials near the rest

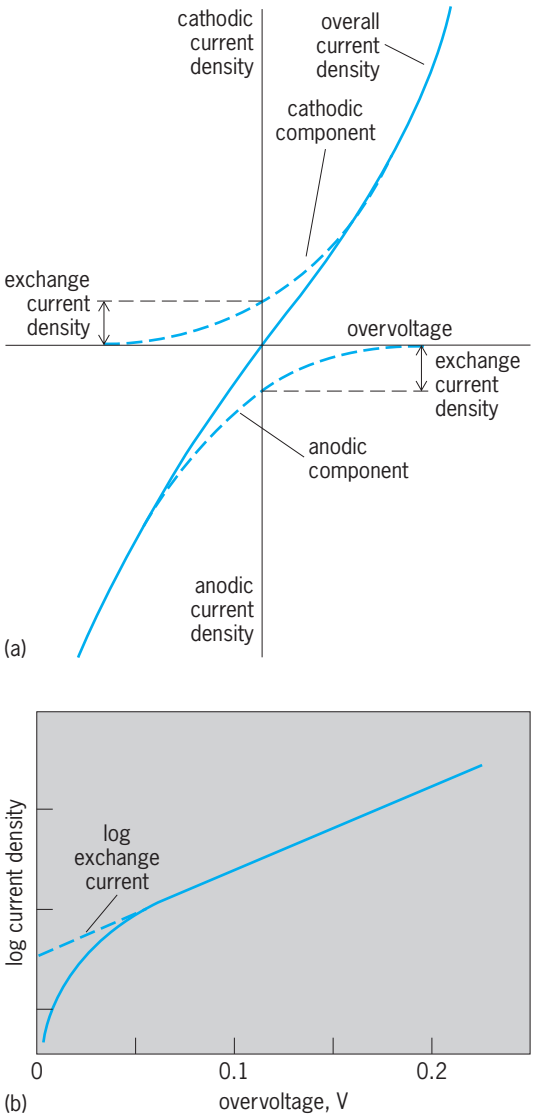

Fig. 2. Current density curves. (*a*) Experimental current–potential curve showing theoretical cathodic and anodic components. (*b*) Tafel plot.

potential, that is, for overvoltages between about +0.1 and −0.1 V, are termed quasireversible reactions. *See* OXIDATION-REDUCTION.

When the overvoltage becomes larger than about 0.1 V in absolute value, one of the current components becomes negligible. Under these conditions a plot of the logarithm of the current density against overvoltage is linear (Tafel law). A Tafel plot based on this relationship (Fig. 2*b*) shows an intercept with the log current density axis (at zero overvoltage) equal to the log exchange current density and a slope that is proportional to a parameter known as the transfer coefficient. An electrode reaction under conditions such that the back reaction makes a negligible contribution to the overall current is known as a totally irreversible reaction.

The exponential dependence of the current density components on overvoltage and the Tafel law were established experimentally. Theoretical treatments of the rates of electrode reactions have led to a justification of these relations and a more detailed understanding of the factors that contribute to the exchange current density and the transfer coefficient. In general, the exchange current density is a function of temperature, concentration of reactants, nature of the solvent and electrolyte, and electrode surface structure.

Other overvoltage contributions. Reaction overvoltage arises when a chemical reaction is associated with the overall electrode reaction. For example, if the electroactive substance is generated from the major reactant by a chemical reaction that precedes the electron transfer, its concentration at the electrode surface will be governed by the rate of this reaction. This preceding reaction will thus affect the potential at which the electrode reaction occurs. Slow steps in the formation of nuclei and the crystal lattice, for example, in the electroplating of a metal, can lead to nucleation and crystallization overvoltages, respectively.

Electrode reaction kinetics measurements. While the overvoltage is a manifestation of slow-rate processes in the electrode reaction, the processes themselves are better characterized by rate constants or exchange currents that are independent of current densities or potentials. If the reaction rates are not too rapid, these rate constants can be determined from the steady-state current–potential curves. However, for more rapid reactions, the observed steady-state behavior is often very similar to that of a reversible reaction, and transient methods at short experimental times (microseconds to milliseconds) must be employed. One approach is to obtain the nonsteady-state current–potential curve at a very rapid scan rate (linear scan or cyclic voltammetry). The potentiostatic method involves application of a potential step to the electrode and observation of the resulting current–time curve. Alternatively, a current step can be applied and the potential–time curve recorded (galvanostatic method). Other methods that have been used involve rapid injection of small quantities of charge (coulostatic methods) or studies of the effect of the superposition of a sinusoidal potential variation on the measured current (ac or sinewave methods). *See* ELECTROCHEMICAL TECHNIQUES; POLAROGRAPHIC ANALYSIS.

Allen J. Bard

Bibliography. A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*, 2d ed., 2000; J. O'M. Bockris and A. Reddy, *Modern Electrochemistry*, 2d ed., 2000; J. S. Newman, *Electrochemical Systems*, 2d ed., 1991; P. H. Rieger, *Electrochemistry*, 2d ed., 1993; K. J. Vetter, *Electrochemical Kinetics*, 1967.

Ovum

The egg or female sex cell. Strictly speaking, the term refers to this cell when it is ready for fertilization, but it is often applied to earlier or later stages. Confusion is avoided by using qualifying adjectives such as immature, ripe, mature, fertilized, or developing ova. The mature ova are generally spheroidal

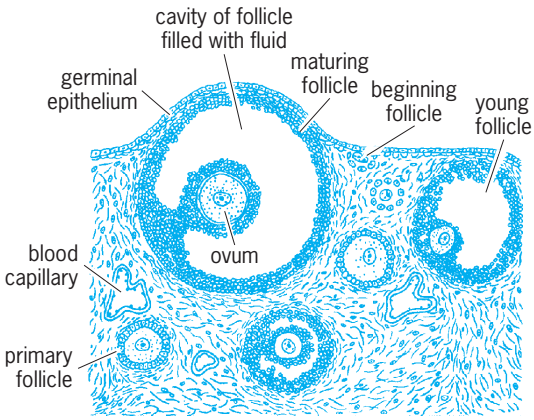

Fig. 1. Section of a mammalian ovary.

and large. In fact, the largest known single cells of a living animal are the mature ova of the ostrich and the shark *Chlamydoselache*, which are about 3 in. (8 cm) in diameter. Among oviparous animals, which spawn eggs at or before the time of fertilization, those that produce larvae capable of feeding at an early stage have small eggs, and their development is generally characterized by radical transitions in appearance, called metamorphosis, before the adult form is attained. The typically viviparous mammals, in which the developing embryo receives nourishment for growth through the uterine tissues of the mother, also characteristically have relatively small eggs which are about 0.1 mm (0.004 in.) in diameter. The number of ova produced at one time varies in different animals, from millions in many marine animals that spawn into the surrounding seawater to about a dozen or less in mammals in which adaptations for internal nourishment of the developing embryo and care of the young are highly developed.

In the ovary the immature ovum is associated with follicle cells through which it receives material for growth. In mammals, as the egg matures, these cells arrange themselves into a structure known as the Graafian, or vesicular, follicle, consisting of a large fluid-filled cavity into which the ovum, surrounded by several layers of cells, projects from the layer of follicle cells that constitutes the inner wall (**Fig. 1**). The fluid contains estrogenic female sex hormone secreted by cells in an intermediate layer of the follicular wall.

Yolk, or deutoplasm, is essentially a food reserve in the form of small spherules, present to a greater or lesser extent in all eggs. It accounts largely for the differences in size of eggs. Eggs are classified according to the distribution of yolk. In the isolecithal type there is a nearly uniform distribution through the cytoplasm (**Fig. 2***a*), as in most small eggs. The yolk in telolecithal eggs is increasingly concentrated toward one pole (Fig. 2*b*), as in the large eggs of fish, amphibians, reptiles, and birds. Centrolecithal, or centrally located, yolk (Fig. 2*c*) occurs in eggs of insects and cephalopod mollusks.

Polarity of organization is manifest in telolecithal eggs by the higher concentration of yolk at the vegetal than at the animal pole, the distribution being radially symmetrical about the polar axis. This distribution is present, though less marked, in isolecithal eggs. The animal pole of the egg is also the place where the polar bodies are formed. It marks the region where the future head of the embryo develops. *See* GAMETOGENESIS; OOGENESIS.

Bilateral symmetry is sometimes evident in the shape of the egg, as in insects. The bisecting plane

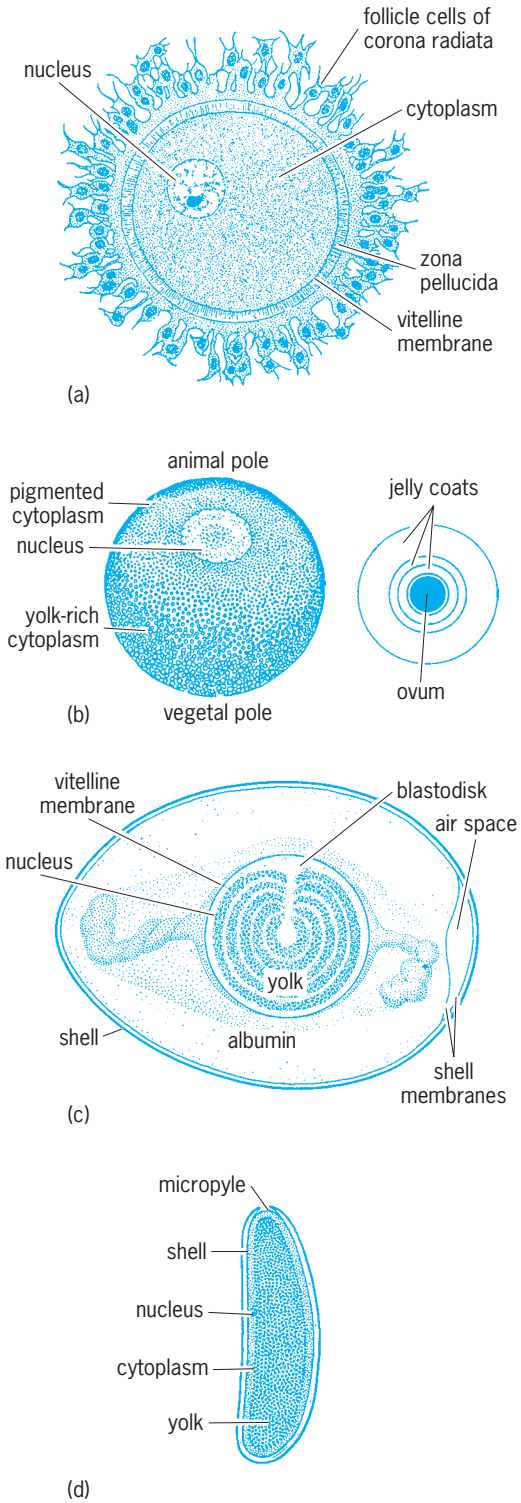

Fig. 2. Types of ova. (*a*) Isolecithal, human. (*b*) Telolecithal, frog. (*c*) Telolecithal, hen. (*d*) Centrolecithal, fly.

represents the future plane of bilateral symmetry of the embryo. In chickens the plane of bilateral symmetry of the developing embryo is generally perpendicular to the long axis of the shell, and its left side is toward the blunt end of the shell. In frogs and many invertebrates it has been discovered that the path which the fertilizing spermatozoon follows in entering the egg is related to the position of the future plane of bilateral symmetry of the embryo, including its dorsoventral axis. The significance of this may be appreciated when it is realized that, together with the polar axis, the position of the dorsoventral axis specifies completely the future location of all the organs of the developing embryo and adult. The establishment of these axes provides the basic pattern of internal organization of the egg. *See* ANIMAL SYMMETRY.

Regulative and mosaic eggs. Despite this patterning at an early stage, it has been possible experimentally in many species of animals to obtain complete individuals from fragments of unfertilized or of developing eggs. This is generally successful when the cuts are made through the polar axis. Eggs in which such fragments develop as a whole are termed regulative and are widely distributed throughout the animal kingdom, including man, as evidenced by the occurrence of identical, multiple births. Eggs in which the fragments develop as structurally defective embryos have been termed mosaic and have been reported principally among the annelids, mollusks, and tunicates. However, the distinction has lost most of its original significance with the demonstration that twins and double monsters were also experimentally obtainable in eggs of the latter group by appropriate procedures. Experiments involving cutting across the polar axis have provided a further clue as to the nature of the internal organization of the egg. When thus cut in half, each fragment develops defectively, but a combination of animal and vegetal quarters develops into a normal individual, as does also the remaining middle fragment. Thus, there are interactions of materials distributed along the polar axis, and a proper balance is essential for normal development. *See* DEVELOPMENTAL BIOLOGY.

Egg membranes. The membranes which surround the egg are designated as primary, secondary, and tertiary according to their derivation from the ovum itself, the surrounding follicular cells of the ovary, and the lining of the oviduct, respectively. The primary membrane in practically all animals is called the vitelline membrane. In many animals this becomes elevated from the surface at fertilization and is then called the fertilization membrane. Secondary membranes are represented by the zona pellucida and the surrounding gelatinous, follicle cell–containing material known as the cumulus oophorus in the mammalian egg. Often, however, it is difficult to decide if a particular coat of the egg is produced by the egg itself, by the surrounding follicle cells, or by both. The distinction also loses some of its significance in view of the evidence that most of the materials of the growing oocyte are supplied to it in practically fully formed state. Tertiary membranes are illustrated by the gelatinous coats of amphibian eggs and by the albumin, shell membranes, and the leathery or calcareous shells of eggs of reptiles and birds. These structures are applied to the egg as it descends the oviduct after ovulation.

In higher plants the egg cell of the embryo sac represents the equivalent of the ovum of animals. In lower plants eggs are formed in an archegonium more nearly analogous to the ovary in animals.

Albert Tyler; Howard L. Hamilton

Oxalidales

An order of flowering plants (angiosperms) in the eurosid I group of the rosid dicots. The order is previously unrecognized in classifications of the angiosperms but is indicated by numerous studies of DNA sequences. Oxalidales consist of five small families: Cephalotaceae (one species), Connaraceae (300 species of tropical trees and vines), Cunoniaceae (250 species of trees and shrubs mostly from the Southern Hemisphere), Elaeocarpaceae (350 species of trees and shrubs from the Southern Hemisphere and Asian tropics), and Oxalidaceae (350 species, mostly in *Oxalis*, mostly herbs that are found throughout the world). Like many of the other newly defined orders based on studies of DNA sequences, Oxalidales are heterogeneous in their morphological traits; some are bizarre, such as the *Cephalotus* (Cephalotaceae), a carnivorous pitcher plant from Australia. Many species of the order are locally economically important, producing timbers and fruits, including zebrawood (*Connarus*, Connaraceae), star fruit (*Averrhoa*, Oxalidaceae), and lightwood (*Ceratopetalum* and *Eucryphia*, Cunoniaceae). *Oxalis* (Oxalidaceae) has some species that are grown as ornamentals and several that are noxious introduced weeds. *See* MAGNOLIOPHYTA; MAGNOLIOPSIDA; ORNAMENTAL PLANTS; PITCHER PLANT; ROSIDAE; WEEDS.

Mark W. Chase

Oxidation process

Processes in which oxygen is caused to combine with other molecules. The oxygen may be used as elemental oxygen, as in air, or in the form of an oxygen-containing molecule which is capable of giving up all or part of its oxygen. Oxidation in its broadest sense, that is, an increase in positive valence or removal of electrons, is not considered here if oxygen itself is not involved. *See* OXIDATION-REDUCTION.

Most oxidations occur with the liberation of large amounts of energy in the form of heat, light, or electricity. The stable ultimate products of oxidation are oxides of the elements involved. These oxidations occur in nature as corrosion, decay, and respiration and in the deliberate burning of matter such as wood, petroleum, sulfur, or phosphorus to oxides of the constituent elements. This article deals only with cases where the object of the oxidation process is

the manufacture of a chemical product rather than the production of energy.

The principal variables to be considered and controlled in any partial oxidation are temperature, pressure, reaction time (or contact time), nature of catalyst, if any, mole ratio of oxidizing agent, and whether the substance to be oxidized is to be kept in the liquid or vapor phase. Only a narrow range of conditions unique to each substance being oxidized and each product desired will give satisfactory yields. It is also essential to maintain conditions outside the range of spontaneous ignition, to avoid explosive mixtures or the accidental accumulation of unstable peroxides, and to choose materials which not only can resist the environmental conditions but also which do not have adverse catalytic effects or otherwise interfere with the desired reaction. *See* COMBUSTION; EXPLOSIVE.

Most oxidations of organic compounds with oxygen appear to proceed through free-radical chain reactions. The specific transient intermediates and sequence of reactions are very complex and are still not completely understood. Many oxidation reactions are autocatalytic. An induction period is therefore often observed during which the concentration of catalyst or intermediate is being built up to the level required for the reaction to proceed at a measurable rate. The reaction rate can continue to increase due to reaction chain branching unless controlled. *See* CHAIN REACTION (CHEMISTRY).

Catalytic effects can be obtained with solid surfaces, generally used in vapor-phase oxidations; with soluble salts, generally used for liquid-phase oxidations; with gases, added in small amounts to the air; or with radiation. Very often catalysts are mixtures in which the action of the major component is modified or maintained by the addition of other agents to the gas, liquid, or solid phase. A limited number of generalizations can be made about oxidation catalysts. The metals, used in the form of their oxides or salts, have a variable valence under the reaction conditions. In vapor-phase reactions the active catalyst is generally deposited on an inert support. Liquid-phase catalysts are usually in the form of a salt, very often containing cobalt or manganese, which is soluble in the organic medium. Inhibitors are also important. Catalyst supports and materials of construction must not have an inhibitory effect and, in liquid-phase oxidations, reactive molecules which terminate free-radical chains, such as phenols, must be avoided. *See* CATALYSIS; INHIBITOR (CHEMISTRY).

In partial oxidations, high selectivity, which may be defined as the equivalents of desired product formed divided by the total moles of feed oxidized, is often made possible by differing degrees of resistance to attack among the various atoms in an organic molecule. Hydrogen atoms attached to aliphatic carbon atoms are more easily oxidized than those attached to carbon atoms in aromatic rings. Among nonaromatics, the ease of oxidation is in the order tertiary > secondary > primary; so that methyl groups are relatively resistant. On the other hand, hydrogen atoms attached to aliphatic carbon atoms are activated by adjacent methyl groups, double bonds, and aromatic rings in roughly that increasing order. These observations are qualitatively illustrated by the examples shown in reactions (1)–(7). Among the compounds already containing oxygen, a rough order of decreasing stability under oxidizing conditions is carboxylic acid anhydrides > esters > carboxylic acids > ketones > secondary alcohols > aldehydes.

In reactions (4) and (5) the anhydride group is relatively stable under the reaction conditions, whereas the corresponding carboxyl group is not. In cases where the reaction conditions required to initiate oxidation of the starting material are severe, it is generally necessary to employ short reaction times, for

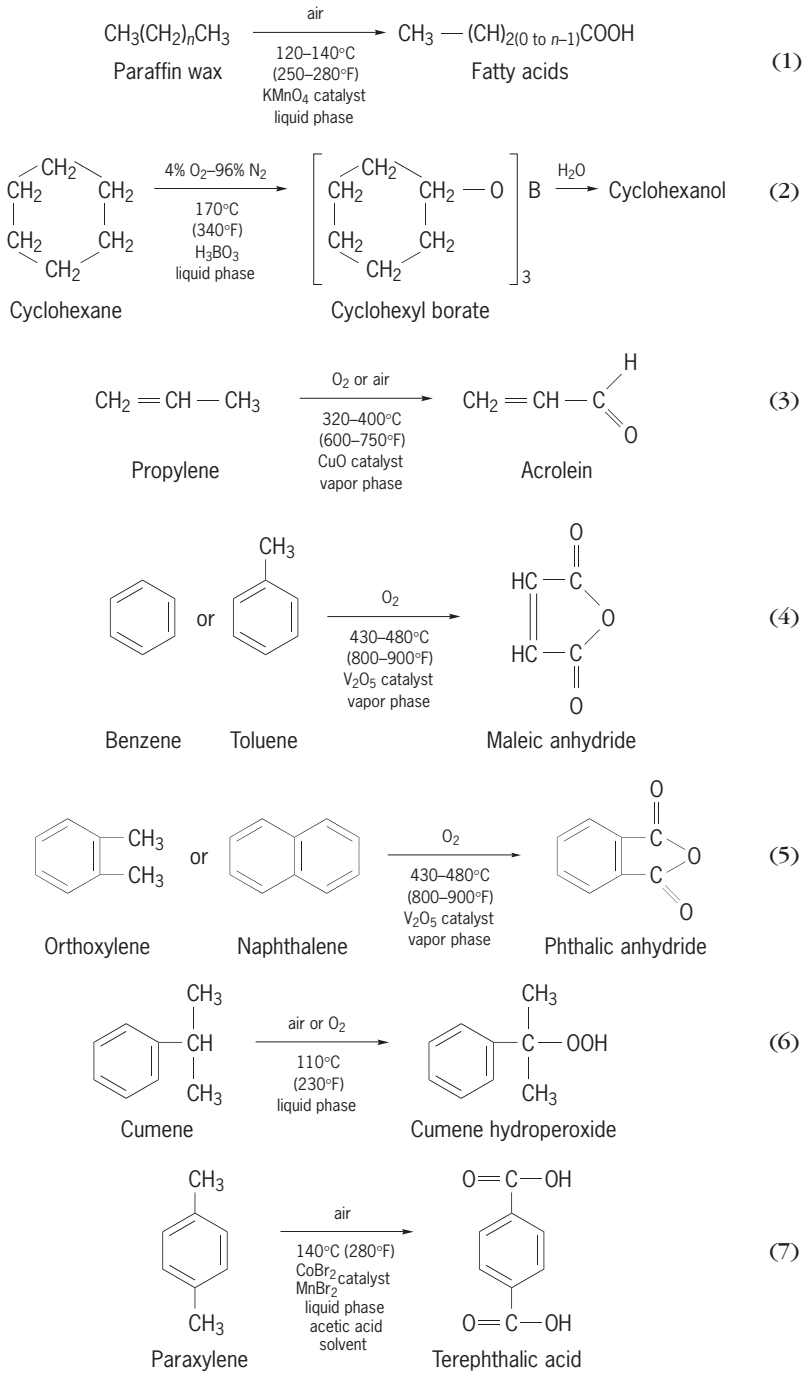

example, 0.001–1 s, by rapidly quenching the reaction mixture to a temperature at which the desired products can survive. In reaction (6) the hydroperoxide is sufficiently stable under the relatively mild conditions required to oxidize cumene for reaction times of an hour or more to be employed. In reaction (2) the unstable secondary alcohol is prevented from decomposing or oxidizing further by formation of the more stable borate ester. Double bonds themselves appear to be relatively resistant to attack by oxygen. Ethylene can be oxidized to ethylene oxide with silver catalysts at high temperatures, but higher olefins require chemical oxidizing agents, such as peroxides, ozone, or hypochlorite, since oxygen attacks other parts of the molecule first.

Plant and equipment design is of utmost importance in achieving efficient control of the oxidation. Vapor-phase reactors with fixed catalyst beds are generally built in the form of tubular heat exchangers if excess air is used. A thermally stable heat exchange medium is circulated outside of the tubes. Molten salts such as sodium nitrate–sodium nitrite mixtures are used for temperatures in the 260–540°C (500–1000°F) region. For temperatures between 100 and 320°C (212 and 600°F), stable organic oils or water under pressure are employed.

If the reaction permits, the catalyst is often used in the form of a fluidized bed; that is, the particle size of the catalyst and support is small enough so that it is suspended in a large vessel by the upward flowing gas. For a given oxidation, reaction time is generally longer than with a fixed bed and the temperature lower. The gas velocity is such that most of the catalyst is not blown out of the reaction zone, and the catalyst which is blown out is removed by cyclone separators or ceramic filters or both and returned. The entire bed is subject to continual mixing, has a high heat capacity, and is easily maintained at a predetermined uniform temperature. Heat is removed by circulating a portion of the hot catalyst through a heat exchanger. Cooling coils can also be put inside the reactor. *See* FLUIDIZATION.

Liquid-phase oxidations are usually carried out in large vessels with internal cooling coils. Inlet air is distributed by spargers or by introducing it at high velocity through a small orifice. Gases leaving the reactor can carry large amounts of the reactants and products, which must be removed by chilling or absorption to minimize losses. In most large oxidation plants, gases are incinerated before release to the atmosphere to avoid pollution.

Commercial processes using oxygen or air. Synthesis gas is manufactured by partial oxidation of hydrocarbons with oxygen. The carbon monoxide–hydrogen mixture obtained is used to make primary alcohols from olefins through the Oxo process and for methanol synthesis. Hydrogen is manufactured from synthesis gas by selective oxidation of the carbon monoxide in the mixture to carbon dioxide and by removal of the latter from the hydrogen by combination with an amine. Carbon dioxide is sometimes recovered from the amine as a separate commercial product.

A number of processes for the oxidation of light aliphatic hydrocarbons are in use for the production of mixtures of acetic acid, methanol, formaldehyde, acetaldehyde, and other products in lesser amounts. The processes are reported to use air or oxygen and to operate with and without catalysts over a wide range of temperatures and pressures in both the liquid and gas phases. Processes also have been developed for oxidizing ethylene to acetaldehyde and vinyl acetate. Aqueous systems with palladium–copper salt catalysts are reported. Ethylene is also oxidized to ethylene oxide in the vapor phase with air in tubular reactors with supported silver catalysts. Temperatures are in the range 200–320°C (400–600°F), and pressures are between 100 and 300 lb/in.^{2} (700–2100 kilopascals).

Propylene oxide is manufactured by oxidizing propylene with the hydroperoxides of either cumene or isobutane, as in reaction (6). Isobutene or styrene are formed as by-products. Cumene hydroperoxide is cleaved in acid solution to phenol and acetone. Phenol is also manufactured by oxidation of benzoic acid with cupric salts, which are regenerated in place with air.

In the manufacture of phthalic anhydride, reaction (5), both fixed- and fluid-bed vanadium oxide catalysts are used. Excess air is employed at substantially atmospheric pressure. Processes for oxidizing benzene to maleic anhydride and of durene to pyromellitic anhydride are similar to those for phthalic anhydride.

Large amounts of terephthalic acid are manufactured from paraxylene by liquid-phase oxidation with air in the presence of manganese and cobalt bromide catalysts. Acetic acid is used as a solvent. A similar process co-oxidizes acetaldehyde along with the xylene in the presence of cobalt acetate, forming acetic acid as a by-product. The acetaldehyde acts to assist oxidation of the xylene by way of a peroxide intermediate, which is probably the actual oxidizing agent. Methyl paratoluate is oxidized to methyl hydrogen terephthalate in the liquid phase with soluble cobalt salts. Processes for isophthalic acid are analogous to those for terephthalic.

Many attempts have been made to manufacture fatty acids commercially by oxidation of paraffin wax. Synthetic fats are reported to have been manufactured from them in Germany during both world wars. It is difficult to separate the oxyacid and lactone by-products from the desired fatty acid, however, and the process does not appear capable of competing in a free economy with fatty acids and fats from natural sources. Fatty alcohols are manufactured by air oxidation of aluminum alkyls and hydrolyzing the resulting aluminum alcoholate.

Cyclohexane is oxidized in the liquid phase with air to form a mixture of cyclohexanol and cyclohexanone. This mixture is further oxidized with nitric acid to adipic acid. Another development is the use of boric acid in the liquid oxidation medium to form the borate ester of cyclohexanol. The ester is hydrolyzed to recover cyclohexanol, and the boric acid is recycled. The use of boric acid is claimed to give

higher total yields and to minimize the formation of troublesome by-products. This process is also used to make secondary alcohols from paraffins for use as detergent intermediates.

Several oxidation processes to produce acetylene have been developed. In these cases, however, the function of the oxidation itself is to produce temperature of around 1650°C (3000°F). Acetylene is produced from natural gas or naphtha introduced into the hot gases.

Two commercial oxidation processes involve oxidation of hydrogen chloride to chlorine over a manganese catalyst in the presence of ethylene to form vinyl chloride and the reaction of propylene and ammonia with oxygen or air to form acrylonitrile. Methanol and ethanol are oxidatively dehydrogenated to the corresponding aldehydes in the vapor phase over silver, copper, or zinc catalysts. Acetaldehyde is oxidized to acetic acid in the liquid phase in the presence of manganese acetate and to acetic anhydride if mixtures of cobalt and copper acetates are used. Polyunsaturated oils are oxidatively polymerized at room temperature in the presence of soluble salts of cobalt and other metals of variable valence. The physical properties of certain asphalts are improved by oxidation with air at elevated temperature, and phenol can be converted to a high polymer by direct oxidation. Microbiological oxidation to produce animal feed and sodium glutamate, as well as to treat domestic and industrial wastes, is becoming increasingly important.

Chemical oxidants. Adipic and terephthalic acids are manufactured from cyclohexanol-cyclohexanone mixtures and paraxylene, respectively, by oxidation with nitric acid of about 30% concentration. Organic and inorganic peroxides are used to manufacture higher olefin oxides, glycols, and hydrogen peroxide. Ozone is used to cleave unsaturated fatty acids to form dibasic acids. Other chemical oxidizing agents for special purposes include sulfur and sulfur compounds, permanganate, perchlorate, hypochlorite, and dichromate. *See* BLEACHING; OZONOLYSIS; PEROXIDE.

Inorganic processes. These are generally carried to the highest stable oxidation state and process control is relatively simple.

Vapor-phase oxidations are used to produce major heavy inorganic chemicals, for example, air oxidation of hydrogen sulfide or sulfur dioxide to sulfur trioxide (sulfuric acid); of ammonia to nitric acid; of phosphorus vapor to phosphorus pentoxide (phosphoric acid); of hydrogen chloride to chlorine; and of vaporized zinc to zinc oxide.

Liquid-phase oxidations of inorganic compounds are rare because so few are liquids. Liquid sulfur is burned to sulfur dioxide. At high temperatures mercuric oxide is made from its elements and litharge from molten lead. Air and oxygen, blown through molten iron, make steel by oxidizing such impurities as carbon, sulfur, and phosphorus. *See* STEEL MANUFACTURE.

Solid-phase oxidations are applied most commonly to obtain oxides from the elements. High-purity carbon dioxide is made from coke in this way. Mixed lead oxides are purified to the monoxide litharge by roasting. Barium peroxide forms from the oxide. Two of the more powerful and costly inorganic oxidizing agents are obtained by processes involving gas-solid phase reactions. Potassium permanganate is produced by roasting a mixture of manganese dioxide and potassium hydroxide with air in a kiln or muffle furnace. In an analogous way, chromite ore and sodium carbonate yield sodium chromate.

I. E. Levine

Bibliography. N. Emanuel and D. Gal, *Modelling of Oxidation Processes*, 1986; N. M. Emanuel, G. E. Zaikov, and Z. K. Maizus, *Oxidation of Organic Compounds: Solvent Effects in Radical Reactions*, 1984; M. Murari, *Mechanism of Oxidation of Organic Compounds by Electron Oxidants*, 1985; R. G. Rice, *Advanced Oxidation Processes*, 1995; R. Sheldon and J. Koch, *Metal-Catalyzed Oxidations of Organic Compounds: Mechanistic Principles and Synthetic Methodology Including the Biochemical Process*, 1981.

Oxidation-reduction

An important concept of chemical reactions which is useful in systematizing the chemistry of many substances. Oxidation can be represented as involving a loss of electrons by one molecule and reduction as involving an absorption of electrons by another. Both oxidation and reduction occur simultaneously and in equivalent amounts during any reaction involving either process.

Some important processes which involve oxidation are the rusting of iron or corrosion of metals in general, combustion of hydrocarbons, and the oxidation of carbohydrates (this takes place in a controlled manner in living cells). In each of the foregoing reactions the agent which is reduced is oxygen. Some important reduction processes are the transformation of carbon dioxide to carbohydrates (this takes place in photosynthesis with water being oxidized), the winning of metals from oxides (carbon is often the reducing agent), electrodeposition of metals (this takes place at the cathode, and an equivalent amount of oxidation occurs at the anode), hydrogenation of fats and of coal, and the introduction of electronegative elements such as oxygen, nitrogen, or halogens into hydrocarbons. *See* BIOLOGICAL OXIDATION.

Oxidation number. The oxidation state is a concept which describes some important aspects of the state of combination of the elements. An element in a given substance is characterized by a number, the oxidation number, which specifies whether the element in question is combined with elements which are more electropositive or more electronegative than it is. It further specifies the combining capacity which the element exhibits in a particular combination. A scale of oxidation numbers is defined by assigning to an oxygen atom in a molecule suchas

SO_4^{2-} the value of 2−. That for sulfur as 6+ then follows from the requirement that the sum of the oxidation numbers of all the atoms add up to the net charge on the species. The value of 2− for oxygen is not chosen arbitrarily. It recognizes that oxygen is more electronegative than sulfur, and that when it reacts with other elements it seeks to acquire two more electrons, by sharing or outright transfer from the electropositive partner, so as to complete a stable valence shell of eight electrons. For compounds of the halogens an analogous rule is followed, but when a halogen atom is in combination with atoms of a more electropositive element, the oxidation number is taken as 1− because only one electron needs to be added to the valence shell to yield a stable octet. The system amounts to a bookkeeping operation on the electrons, so that for this purpose the more electronegative partner is assigned some agreed upon stable electronic configuration, and after taking into account the total charge on the molecule, the net charge left on the electropositive partner is its particular oxidation number. When the combining capacity of an element toward another one is not completely exhausted in a particular combination, as is the case for oxygen in barium peroxide (BaO_2), the electrons shared between atoms of the same kind are evenly divided between them in carrying out the formal decomposition. Thus in the peroxide unit O_2^{2-}, the oxidation number of oxygen is 1−. This is the net charge left on oxygen in the formal decomposition. The ox-

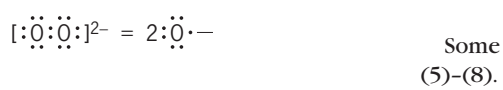

idation number by no means gives a complete description of the state of combination of an atom. Specifically, it is not designed to give the actual charge on an atom in a particular compound. Thus it makes no distinction between fluorine in HF, AlF_3, or NaF, even though the actual charges residing on the fluorine atoms in these three compounds are different.

The utility of the concept is based in part on just this feature because much of the chemistry of these substances can be understood when it is realized that each of them readily yields F^-, as is the case when they dissolve in water. The chemistry of the three substances, in regard to the component fluorine, is concerned with reactions of F^-. Although oxidation number is in some respects similar to valence, the two concepts have distinct meanings. In the substance H_2, the valence of hydrogen is 1 because each H makes a single bond to another H, but the oxidation number is 0, because the hydrogen is not combined with a different element. *See* VALENCE.

The systematization of chemistry based on the concept of oxidation number can be illustrated with reference to the chemistry of iodine. The usual oxidation states exhibited by iodine are 1−, 0, 1+, 5+, and 7+. Examples of substances corresponding to each oxidation state are

7+	IO_4^-, HIO_4, IF_7
5+	I_2O_5, IO_3^-, HIO_3, IF_5
1+	HIP, IO^-, ICl_2^-
0	I_2
1−	IP−, HI, NaI

When the oxidation number of an atom in a species is increased, the process is described as oxidation, no matter what reagent produces it; when a decrease in oxidation number takes place, the process is described as reduction, again without regard to the identity of the reducing agent. The term oxidation has been generalized from its historical meaning, implying combination with oxygen, to combination of an element with an element more electronegative than itself.

When classification by oxidation number is adopted, the reactions fall naturally into two classes. In the first class, no change in oxidation number takes place and, in the second, the class of oxidation-reduction reactions, changes in oxidation number do take place. Some examples of the first class are reactions (1)–(4).

$$I_2O_5 + H_2O \longrightarrow 2HIO_3 \quad (1)$$

$$HIO_3 + OH^- \longrightarrow IO_3^- + H_2O \quad (2)$$

$$HOI + H^+ + 2Cl^- \longrightarrow H_2O + ICl_2^- \quad (3)$$

$$Hg^{2+} + 4I^- \longrightarrow HgI_4^{2-} \quad (4)$$

Some samples of the second class are reactions (5)–(8). [In reaction (8) it is implied that electrons

$$Cl_2 + 2I^- \longrightarrow 2Cl^- + I_2 \quad (5)$$

$$2Fe^{3+} + 2I^- \longrightarrow 2Fe^{2+} + I_2 \quad (6)$$

$$16H^+ + 2MnO_4^- + 10I^- \longrightarrow 8H_2O + 2Mn^{2+} + 5I_2 \quad (7)$$

$$2I^- \longrightarrow I_2 + 2e^- \quad (8)$$

are being extracted from I^- by an anode in an electrolytic process.]

In reactions of the first class, some center regarded as positive undergoes a change in the nature of the groups associated with it, but provided that the group which replaces the electronegative portion is more electronegative than the center, there is no change in oxidation state. Reaction (3) describes the replacement of OH^- on I^+ by Cl^-; both Cl and OH are more electronegative than I. In reactions of the second class, changes in oxidation number occur which may or may not be accompanied by changes in the state of association of the centers in question.

Reactions (5), (6), (7), and (8) illustrate the utility of the concept of oxidation number. A variety of reagents as different in their properties as Cl_2, Fe^{3+} MnO_4^-, and an anode serve to bring about the change, or oxidation, of I^- to I_2. However, their chemical individuality does not affect the state of the product iodine, and nogroup characteristic of the

oxidizing agent is necessarily transferred in the net change. This situation obtains only for reactions in a strongly solvating medium such as water, which provides the groups that associate with the atom being oxidized or reduced. Thus, when the reactions take place in the solid, it is necessary to specify what iodide is being used, whether sodium iodide, NaI, or silver iodide, AgI, for example, and the properties of the reaction would be dependent on the choice.

In representing an element in a particular environment, it is often convenient to specify only the oxidation state, without attempting to identify all of the groups which are attached to the atom in question. Thus, the iron content of a solution made up by dissolving, say, ferric chloride in water will be composed, among others, of the species Fe^{3+}, $FeCl^{2+}$, $FeOH^{2+}$. Collectively, they are correctly, though of course not fully, described by the notation Fe(III). In this kind of usage, the roman numeral represents the oxidation state.

Oxidation-reduction reactions. In an oxidation-reduction reaction, some element decreases in oxidation state and some element increases in oxidation state. The substances containing these elements are defined as the oxidizing agents and reducing agents, and they are said to be reduced and oxidized, respectively. The processes in question can always be represented formally as involving electron absorption by the oxidizing agent and electron donation by the reducing agent. For example, reaction (6) can be regarded as the sum of the two partial processes, or half-reactions, (9) and (10).

$$2I^- \longrightarrow I_2 + 2e^- \tag{9}$$

$$2Fe^{3+} + 2e^- \longrightarrow 2Fe^{2+} \tag{10}$$

Similarly, reaction (7) consists of the two half-reactions (11) and (12), with half-reaction (11) being

$$2I^- \longrightarrow I_2 + 2e^- \tag{11}$$

$$16H^+ + 2MnO_4^- + 10e^- \longrightarrow 2Mn^{2+} + 8H_2O \tag{12}$$

taken five times to balance the electron flow from reducing agent to oxidizing agent.

Each half-reaction consists of an oxidation-reduction couple; thus, in half-reaction (12) the reducing agent and oxidizing agent making up the couple are manganous ion, Mn^{2+}, and permanganate ion, MnO_4^-, respectively; in half-reaction (11) the reducing agent is I^- and the oxidizing agent is I_2. The fact that MnO_4^- reacts with I^- to produce I_2 means that MnO_4^- in acid solution is a stronger oxidizing agent than is I_2. Because of the reciprocal relation between the oxidizing agent and reducing agent comprising a couple, this statement is equivalent to saying that I^- is a stronger reducing agent than Mn^{2+} in acid solution. Reducing agents may be ranked in order of tendency to react, and this ranking immediately implies an opposite order of tendency to react for the oxidizing agents which complete the couples. In the list below some common oxidation-reduction couples are ranked in this fashion:

Strong reducing agent		Weak oxidizing agent
	$Mg = Mg^{2+} + 2e^-$	
	$Zn = Zn^{2+} + 2e^-$	
	$H_2 = 2H^+ + 2e^-$	
	$Cu = Cu^{2+} + 2e^-$	
Increasing reducing power ↑	$I^- = \frac{1}{2}I_2 + e^-$	Increasing oxidizing power ↓
	$Fe^{2+} = Fe^{3+} + e^-$	
	$Br^- = \frac{1}{2}Br_2 + e^-$	
	$Cl^- = \frac{1}{2}Cl_2 + e^-$	
	$4H_2O + Mn^{2+} = MnO_4^- + 8H^+ + 5e^-$	
Weak reducing agent		Strong oxidizing agent

A complete list contains the displacement series of the metals. The most powerful reducing agent shown in the list is magnesium, Mg, although this is not the most powerful known. Magnesium is capable of reacting with any oxidizing agent below it in the list to yield Mg^{2+} and to produce the reduced product resulting from the oxidizing agent. Similarly, permanganate ion, MnO_4^-, in acid, the strongest oxidizing agent shown, is capable of reacting with any reducing agent above it in the list. Conversely, an oxidizing agent at the top of the list will not react appreciably with the reducing agent of a couple below it. The list given, containing nine entries, can be used to predict the results of 72 separate experiments (for example, $Mg + Zn^{2+}$ on the one hand and $Mg^{2+} + Zn$ on the other would be counted as separate experiments in arriving at this figure). *See* ELECTROCHEMICAL SERIES; ELECTRONEGATIVITY.

Since the driving force for a reaction depends on concentrations, the concentrations of all reactants and products must be specified, as well as other conditions, in compiling a list such as that given. The order shown obtains for water solutions at 25°C (77°F), approximately 1 *M* in all soluble reagents and having gases present at approximately 1 atm pressure. A second limitation on the use of this list lies in the fact that it applies only when the expected reaction products are compatible. Although copper is capable in principle of reducing iodine to form I^- and Cu^{2+} at high concentration, these products are not compatible with each other, but they react to form copper (I) iodide, CuI. Allowance for such features can always be made by incorporating the necessary additional half-reactions into the list. Finally, it must be stressed that the list can be used to predict the results of experiments only for systems which reach equilibrium sufficiently rapidly; it does not serve to predict the rate of reaction. To achieve the reduction of Fe^{3+} by H_2 predicted in the list, it would be necessary to use a catalyst in order to realize the reaction in a reasonable time.

The equilibrium information implied by a table of half-reactions can readily be put into quantitative form. Thus, the standard free-energy change for the reaction of 1 equivalent weight of each reducing agent with some common oxidizing agent can be entered opposite each half-reaction. The numerical values of these quantities will be in the same order as

are the half-reactions and can combined algebraically to yield the standard free energy change, and therefore the equilibrium constant, for any reaction which can be written from the table. *See* CHEMICAL EQUILIBRIUM.

A chemist concerned with reactions of the type under discussion will have a ready vocabulary of facts concerning oxidizing agents and reducing agents, such as their oxidizing or reducing powers, the speed with which they react, and the characteristics which may complicate their application. A typical problem in analytical chemistry is to reduce $Cr_2O_7^{2-}$ to Cr^{3+} in acidic (perchloric acid) solution without introducing elements which are not already present. Metallic reducing agents such as zinc and iron or metal ion reducing agents are immediately eliminated from consideration because the products of oxidation may be difficult to remove from the resulting solution. A solution of hydrogen iodide, HI, would be suitable, except that it would be necessary to take special pains to add it in equivalent amount because excess HI would be difficult to remove (the iodine, I_2, produced by oxidation of I^-, however, is easy to remove by extracting it with a suitable solvent such as carbon tetrachloride). Hydrogen would be ideal (the product of its oxidation, H^-, is already present in solution, and excess reducing agent can easily be removed) except that the rate of reaction would be disappointingly slow. A suitable reducing agent would be hydrogen peroxide, H_2O_2; it reacts rapidly, the product of oxidation is oxygen, which escapes from solution, and excess oxidizing agent is easily destroyed by heating the solution. *See* ELECTRODE POTENTIAL.

Mechanisms. The data needed to predict the outcome at equilibrium of the reaction of most common oxidizing and reducing agents are known. A list of the kind shown above can be extended, and when it is elaborated with entries carrying the quantitative information, accurate calculations of the equilibrium state for all the reactions implied by the table can be made. By contrast, though the rates of reaction are also of great importance, they are much less completely understood and less completely described. To understand the rates of reaction, it is necessary to consider how the reactions take place. To illustrate one of the problems of mechanism, a reaction is selected which, though not nearly as complicated as some, suffices for present purposes. When an aqueous solution containing Fe^{2+} is added to one containing Br_2, reaction (13) takes place. For the final stable

$$2Fe^{2+} + Br_2 \longrightarrow 2Fe^{3+} + 2Br^- \qquad (13)$$

products to be produced in a single step would require that two Fe^{2+} and one Br_2 be brought together in an encounter. This course for the reaction is found to be less probable than one in which a single Fe^{2+} encounters one Br_2. Since in forming stable products the oxidation state of iron increases by one unit while that of a bromine molecule decreases by two (one for each atom), the reaction resulting from the encounter must leave either iron or bromine in an unstable state. Reasonable alternatives for the reactive intermediates produced are represented in reactions (14*a*) and (15*b*), together in each case with

$$Fe^{2+} + Br_2 \longrightarrow Fe^{3+} + Br^- + Br \qquad (14a)$$

$$\text{or} \quad Fe^{2+} + Br_2 \longrightarrow Fe(IV) + Br^- \qquad (15a)$$

$$Br + Fe^{2+} \longrightarrow Fe^{3+} + Br^- \qquad (14b)$$

$$\text{or} \quad Fe(IV) + Fe^{2+} \longrightarrow 2Fe^{3+} \qquad (15b)$$

a sequel reaction which leads to the correct overall stoichiometry. *See* REACTIVE INTERMEDIATES.

In the present system, the evidence points to the mechanism represented by the first alternative [reactions (14*a*) and (14*b*)]. But even after a reaction such as (13) is resolved into the different steps, and reactive intermediates are identified, there remain questions about how the changes in oxidation state are brought about in the individual steps. Thus, when Fe^{2+} reacts with Br_2, do the two reactants make direct contact—this would involve replacement of a water molecule on $Fe(H_2O)_6^{2+}$, the form which Fe^{2+} adopts in water, by Br_2—or does reaction occur by electron transfer from intact $Fe(H_2O)_6^{2-}$ to Br_2? If the latter occurs, does electron transfer take place over large distances from $Fe(H_2O)_6^{2+}$ to Br_2?

These questions are important not only for inorganic systems of the kind illustrated by the steps making up reaction (13) but also for reactions at electrodes, and for oxidation-reduction reactions catalyzed by enzymes in biological systems. These kinds of questions are under investigation and have been partly answered for certain systems. *See* ENZYME.

Two different kinds of mechanisms are recognized for oxidation-reduction reactions. So-called outer-sphere reactions are easier to understand in a fundamental way and will be described first. Reaction (16) introduces a typical such reaction. Here the changes in oxidation state, 4+ to 3+ for Ir and 2+ to 3+ for Fe, take place without bonds to either atom being broken, and in this particular system there is not even much change in bond distances attending the changes in oxidation state. Electron transfer is explicit in such a system, and the electron transfer act is subject to the Franck-Condon restriction. This imposes a barrier to electron transfer in that it requires that the environments (coordination sphere and solvent) readjust prior to electron transfer so that after readjustment the energy of the system is the same whether the electron is on one metal or the other. The rates of reactions such as (16) can be estimated

$$IrCl_6^{2-} + Fe(CN)_6^{4-} \longrightarrow IrCl_6^{3-} + Fe(CN)_6^{3-} \qquad (16)$$

at least approximately by calculating the work to bring the partners together, the work of meeting the Franck-Condon restriction, and assuming that electron delocalization when the partners are in contact is adequate. Greater success has been met in attempts at correlating rates and here the equation developed by R. A. Marcus, relating the rate of reaction such as (16) to the standard free energy change

and to the rates of the so-called self-exchange reactions [(17) and (18)], is proving to be useful not only

$$Fe^*(CN)_6^{4-} + Fe(CN)_6^{3-} \longrightarrow {}^*Fe(CN)_6^{3-} + Fe(CN)_6^{4-} \quad (17)$$

$$^*IrCl_6^{3-} + IrCl_6^{2-} \longrightarrow {}^*IrCl_6^{2-} + IrCl_6^{3-} \quad (18)$$

in simple systems but also in understanding electron transfer reactions for large and complex biological molecules.

Much more difficult to systematize and understand is the extensive and very important class of oxidation-reduction reactions in which the changes in oxidation state are linked to bond breaking or bond making. A simple example is provided by reaction (19). Isotopic labeling experiments have shown

$$ClO_3^{*-} + SO_2 + H_2O \rightarrow ClO_2^{*-} + SO_3O^{*2-} + 2H^+ \quad (19)$$

that in the course of the reaction an oxygen atom originating on ClO_3^- is transferred to the reducing agent. Though the process can formally be represented as involving electron loss by SO_2 and electron gain by ClO_3^-, electron transfer as the means by which the changes in oxidation state are brought about is not at all explicit in a reaction of this kind. These so-called inner-sphere mechanisms operate also for reactions involving metal ions. For example, when $[(NH_3)_5CoCl]^{2+}$ reacts with $Cr(H_2O)_6^{2+}$, the Cr(III) product has the formula $Cr(H_2O)_5Cl^{2+}$, and Cl is transferred from Co(III) to Cr(II) when the former oxidizes the latter. This kind of reaction clearly has much in common with that represented by reaction (19). In the latter system atomic oxygen is formally transferred; in the former, atomic chlorine.

A class of inner-sphere reactions of metal-containing molecules which are now recognized as playing an important role in many catalytic processes involves so-called oxidative addition. Reactions of this kind have been known for a long time, but their significance was not appreciated until interest in homogeneous catalytic processes began to develop. A commonplace example of oxidative addition is provided by reaction (20). It will be noted that in reac-

$$SnCl_2 + Cl_2 \longrightarrow SnCl_4 \quad (20)$$

tion (20) both the oxidation number and the coordination number of Sn increase. Oxidative addition for a strong oxidizing agent such as Cl_2 is not surprising, but the reaction type took on new significance when it was discovered that with a suitable metal complex, oxidative addition can be realized also with hydrogen halides, alkyl halides, and even hydrogen. Among the metal complexes which undergo this kind of reaction are Rh(I), Ir(I), or Pt(O) species with 4 or fewer groups attached. In each case, there is the opportunity for an increase in both oxidation and coordination number. A specific example of a molecule which undergoes oxidative addition with H_2 is $[(C_6H_5)_3P]_3RhCl$, which is a useful catalyst for the hydrogenation of alkenes and alkynes. A reaction step in the catalytic sequence is the addition of H_2 to the metal so that the H-H bond is severed; the adduct then reacts with the alkene (or alkyne) transferring two attoms of hydrogen. A substance which will activate H-R bonds (where R is an alkyl radical) in the same way would be desirable. *See* HOMOGENEOUS CATALYSIS.

The fundamental aspects of electron transfer processes in oxidation-reduction reactions have much in common with the electron jump processes in semi-conductors. Recognizing this connection is productive both for those interested in chemical effects accompanying electron transfer (that is, in oxidation-reduction processes) and those interested in electron mobility as a subject in its own right.

Henry Taube

Bibliography. J. P. Collman and L. S. Hegedus, *Principles and Applications of Organotransition Metal Chemistry*, 2d ed., 1987; L. E. Eberson, *Electron Transfer Reactions in Organic Chemistry*, 1987; A. G. Lappin, *Redox Mechanisms in Inorganic Chemistry*, 1993; W. M. Latimer, *Oxidation States of the Elements and Their Potentials in Aqueous Solution*, 2d ed., 1952; H. Taube, *Electron Transfer Reactions of Complex Ions in Solution*, 1970.

Oxide

A binary compound of oxygen with another element. Oxides have been prepared for essentially all the elements, with the exception of the noble gases. Often, several different oxides of a given element can be prepared; a number exist naturally in the Earth's crust and atmosphere: silicon dioxide (SiO_2) in quartz; aluminum oxide (Al_2O_3) in corundum; iron oxide (Fe_2O_3) in hematite; carbon dioxide (CO_2) gas; and water (H_2O).

Most elements will react with oxygen at appropriate temperature and oxygen pressure conditions, and many oxides may thus be directly prepared. Phosphorus burns spontaneously in oxygen to form phosphorus pentoxide, $(P_2O_5)_2$. Sulfur requires ignition and thereafter burns to sulfur dioxide (SO_2) gas if the supply of oxygen is limited. The relative amounts of oxygen and element available often determine which of several oxides will form; in an excess of oxygen, sulfur burns to form some sulfur trioxide (SO_3) gas. Most metals in massive form react with oxygen only slowly at room temperatures because the first thin oxide coat formed protects the metal; magnesium and aluminum remain metallic in appearance for long periods because their oxide coatings are scarcely visible. However, diffusion of the oxygen and metal atoms through the film becomes rapid at high temperatures, and these metals will burn intensely to their oxides if ignited. The oxides of the alkali and alkaline-earth metals, except for beryllium and magnesium, are porous when formed on the metal surface, and they provide only limited protection to the continuation of oxidation, even at room temperatures. Gold is exceptional in its resistance to oxygen, and its oxide (Au_2O_3) must be prepared by indirect means. The other noblemetals, although

ordinarily resistant to oxygen, will react at high temperatures to form gaseous oxides.

Indirect preparation of the oxides may be accomplished by heating hydroxides, nitrates, oxalates, or carbonates, as in the production from the latter of quicklime (CaO) by reaction (1), in which carbon

$$CaCO_3 \rightarrow CaO + CO_2 \quad (1)$$

dioxide is driven off. Gold oxide may be prepared by heating gold hydroxide, as shown by reaction (2). Higher oxides of an element may be reduced

$$2Au(OH)_3 \rightarrow Au_2O_3 + 3H_2O \quad (2)$$

to lower oxides, usually at high temperatures, for example, the reduction of tungsten trioxide by tungsten, as shown by reaction (3). Complete reduction

$$2WO_3 + W \rightarrow 3WO_2 \quad (3)$$

to the element may be performed by other elements whose oxides are more stable, as in the formation of calcium oxide from titanium dioxide by reaction (4).

$$2Ca + TiO_2 \rightarrow Ti + 2CaO \quad (4)$$

Although the solid oxides of a few metals, such as mercury and the noble metals, can be easily decomposed by heating, for example, reaction (5), most

$$2Au_2O_3 \rightarrow 4Au + 3O_2 \quad (5)$$

metal oxides are very stable and many constitute important refractory materials. For example, magnesium oxide, calcium oxide, and zirconium dioxide do not melt or vaporize appreciably at temperatures up to 2500°C (4500°F). A great number of refractories consist of compounds of two or more oxides; silicon dioxide and zirconium dioxide form zirconium silicate by reaction (6).

$$SiO_2 + ZrO_2 \rightarrow ZrSiO_4 \quad (6)$$

Because so many oxides can be easily formed, studies of them have been most important in establishing relative atomic weights of the elements based on the defined atomic weight for oxygen. Furthermore, these studies were fundamental in forming the basis for the laws of definite proportions and multiple proportions for compounds. It is of special significance that, although any gaseous oxide species must necessarily have a definite oxygen-to-element proportion, a number of solid and liquid oxides can be prepared with proportions which may vary continuously over a considerable range. This is particularly true for oxides prepared under equilibrium conditions at high temperatures. Thus, titanium exposed to oxygen until reaction equilibrium is reached at a number of selected conditions of temperature and oxygen pressure will form the solid oxide TiO. It has the same crystal structure as rock salt; that is, every other site along the three coordinate directions of the crystal will be occupied by titanium atoms and the alternate sites by oxygen atoms (each in its ion form Ti^{2+} and O^{2-}) to give the simple Ti/O ratio of 1:1. However, with other selected pressure-temperature conditions, oxides of same structure at every Ti/O ratio from 1:0.7 to 1:1.25 may be prepared. The variable proportions show that variable numbers of oxygen or titanium sites can simply remain vacant in a homogeneous way. The range is referred to as the TiO/O 1:07–1.25 phase or, more loosely, the TiO solid-solution phase.

Most of the nonmetal oxides commonly encountered as gases, such as SO_2 and CO_2, form solids and liquids in which the molecular units of the gas are retained so that the simple definite proportions are clearly maintained. Such oxides melt and boil at low temperatures, because the molecular units are weakly bonded to adjoining molecular units.

Oxides may be classified as acidic or basic according to the character of the solution resulting from their reactions with water. The nonmetal oxides generally form acid solutions, for example, reaction (7),

$$SO_3 + H_2O \rightarrow H_2SO_4 \quad (7)$$

the formation of sulfuric acid. The metal oxides generally form alkaline solutions, for example, reaction (8), for the formation of calcium hydroxide or slaked

$$CaO + H_2O \rightarrow Ca(OH)_2 \quad (8))$$

lime. However, given metals of the groups IV and higher of the periodic table will often have basic, intermediate, and acidic oxides. Here the acid character increases with increasing oxygen-metal ratio. *See* ACID AND BASE; EQUIVALENT WEIGHT; OXYGEN; REFRACTORY. Russell K. Edwards

Bibliography. F. A. Cotton and G. Wilkinson, *Advanced Inorganic Chemistry: A Comprehensive Text*, 6th ed., 1999; G. V. Samsonov, *The Oxide Handbook*, 2d rev. ed., 1982.

Oxide and hydroxide minerals

Mineral phases containing only oxide or hydroxide anions in their structures. By volume, oxide and hydroxide minerals comprise only a small fraction of the Earth's crust. However, their geochemical and petrologic importance cannot be overstated. Oxide and hydroxide minerals are important ores of metals such as iron, aluminum, titanium, uranium, and manganese. Oxide and hydroxide minerals occur in all geological environments. Some form as primary minerals in igneous rocks, while others form as secondary phases during the weathering and alteration of silicate and sulfide minerals. Some oxide and hydroxide minerals are biogenic; for example, iron(III) and manganese(IV) hydroxides and oxides often result from bacterial oxidation of dissolved Fe^{2+} and Mn^{2+} in low-temperature aqueous solutions. *See* HYDROXIDE; MINERAL; ORE AND MINERAL DEPOSITS; RID="622600">SILICATE MINERALS; WEATHERING PROCESSES.

Iron and manganese hydroxide minerals often occur as nanocrystalline or colloidal phases with

high, reactive surface areas. Adsorption of dissolved aqueous ions onto colloidal iron and manganese oxides plays a major role in the fate of micronutrients and heavy metals in soil and ground water and the trace-element chemistry of the oceans. Much current research is focused on measuring the thermodynamics and kinetics of metal adsorption by colloidal Fe-Mn hydroxides and oxides in the laboratory. In anoxic sedimentary environments, bacteria may use iron(III) and manganese(IV) hydroxide minerals as electron acceptors. Consequently, these minerals may facilitate the biodegradation of organic pollutants in soil and ground water. *See* ADSORPTION; COLLOID.

Bonding and crystal chemistry. To a first approximation, the bonding in oxide minerals can be viewed in the ionic (electrostatic) model. According to Pauling's rules, the coordination number of a metal cation (such as Mg^{2+}, Al^{3+}, and Ti^{4+}) is determined by the radius of the cation relative to that of the oxide anion (O^{2-}). This allows one to predict the structures of some of the simpler oxide minerals (**Table 1**). Cations with similar ionic radii (such as Mg^{2+} and Fe^{2+}) are able to substitute for each other and form a solid solution. Sulfide minerals (which are more covalent) show little solid solution. The simple ionic radii arguments will fail when electronic configurations preclude the spherical symmetry of the cations. The ions Cu^{2+} and Mn^{3+}, with nine and four *d* electrons, tend to adopt distorted coordination environments because of the Jahn-Teller effect. Also, *d*-electron configurations can give rise to large octahedral site-preference energies that cause small cations, such as Mn^{4+}, to always adopt octahedral coordination. The magnetic and semiconducting properties of transition-metal oxide minerals, such as magnetite (Fe_3O_4), give useful geophysical signatures for subsurface exploration. *See* IONIC CRYSTALS; JAHNTELLER EFFECT; PROSPECTING; SOLID-STATE CHEMISTRY; STRUCTURAL CHEMISTRY; VALENCE.

TABLE 1. Summary of important simple oxide minerals; classified by structure

Mineral	Formula
X_2O Oxides	
Cuprite	Cu_2O
Argentite	Ag_2O
NaCl Structure	
Periclase	MgO
Manganosite	MnO
Wustite	FeO
Other MO Oxides	
Tenorite	CuO
Zincite	ZnO
Corundum Structure	
Corundum	Al_2O_3
Eskaolite	Cr_2O_3
Hematite	Fe_2O_3
Ilmenite	$Fe^{2+}Ti^{4+}O_3$
Rutile Structure	
Rutile	TiO_2
Cassiterite	SnO_2
Pyrolusite	MnO_2
Spinel Structure	
Spinel	$MgAl_2O_4$
Magnetite	Fe_3O_4
Chromite	$FeCr_2O_4$
Franklinite	$ZnFe_2O_4$
Ulvospinel	$Fe^{2+}(Fe^{2+}Ti^{4+})O_4$
Hausmannite	Mn_3O_4 (distorted)
Fluorite Structure	
Uraninite	UO_2
Thorianite	ThO_2

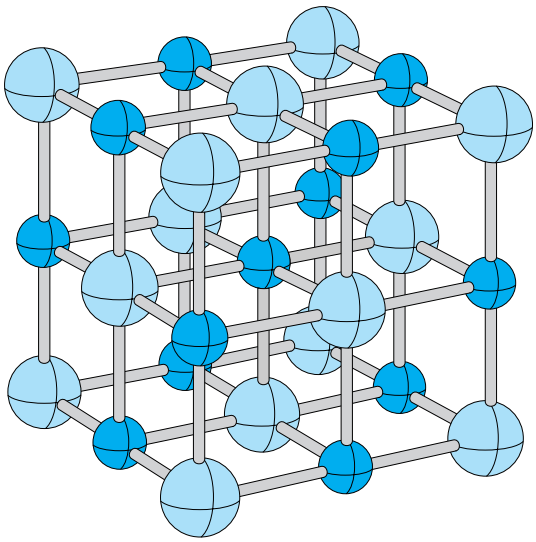

Fig. 1. NaCl structure adopted by XO oxides such as MgO. Small spheres are divalent cations such as Mg^{2+}. Large spheres are O^{2-} anions.

Survey of minerals. The following is a summary of the important oxide and hydroxide minerals by structural classification.

X_2O oxides. The mineral cuprite (Cu_2O) forms during the low-temperature oxidation of primary copper sulfide minerals. It is sometimes an important ore of copper. In the structure of cuprite, Cu^+ ions are in twofold coordination with oxygen. No other monovalent cations form stable X_2O oxide minerals. *See* COPPER.

XO oxides. Oxides of divalent metals, such as Fe^{2+} and Mg^{2+}, with the formula XO will adopt the NaCl structure (**Fig. 1**) in which the metal cation is in sixfold coordination. Fe^{2+} and Mg^{2+} have nearly identical ionic radii, and a phase made up of the solid solution (Mg,Fe)O, called ferropericlase, is probably the second most abundant mineral in the Earth's lower mantle. Experimental evidence indicates that it maintains the NaCl structure over the pressure range of the Earth's interior. The end member MgO (periclase) is quite rare as a crustal mineral; it occurs in some metamorphosed limestones. Most other XO oxide minerals (such as manganosite MnO and bunsunite NiO) are also very rare. Not all XO oxides adopt the NaCl structure. Because the ionic radius of Zn^{2+} is somewhat smaller than that of Mg^{2+}, the mineral zincite ZnO has a structure based on the tetrahedral coordination of Zn. The mineral tenorite (CuO) is a secondary alteration phase of copper sulfides and has a structure based on the square planar coordination of Cu^{2+}. This structure results from the Jahn-Teller distortion of the CuO_6 coordination polyhedron. Larger divalent cations, such as Sr and Ba,

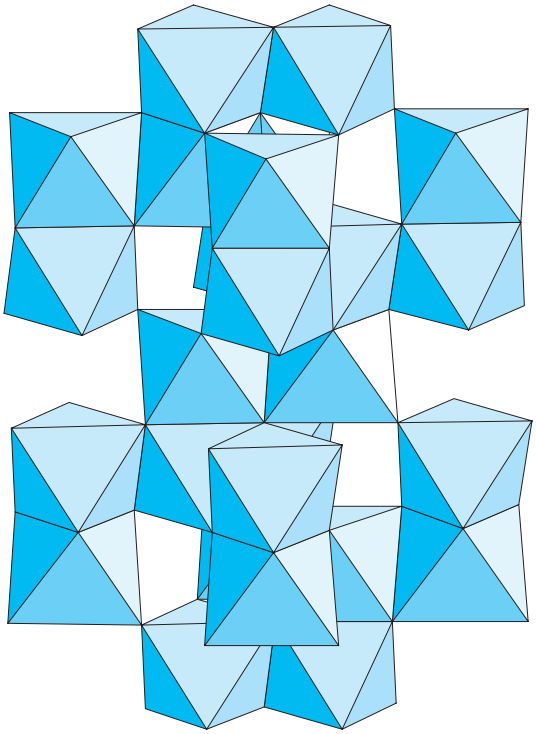

Fig. 2. Polyhedral representation of Al_2O_3 structure adopted by X_2O_3 oxides. Each octahedron represents an M^{3+} cation surrounded by six O^{2-} anions, which define the vertices of the octahedral.

form oxide structures based on eightfold coordination, but these are not stable minerals. *See* CRYSTAL STRUCTURE; SOLID SOLUTION.

X_2O_3 oxides and ilmenite. Trivalent cations, such as Fe^{3+} and Al^{3+}, having radii appropriate for sixfold-coordination with oxygen, will adopt the corundum (α-Al_2O_3) structure (**Fig. 2**). Hematite (α-Fe_2O_3) is a common phase in soils and sediments and forms by the oxidation of Fe^{2+} in primary silicates. (Rust is mostly hematite.) Hematite is the most important ore mineral of iron and is the dominant mineral found in Precambrian banded iron formations. These vast deposits formed because of the oxidation of dissolved Fe^{2+} in the oceans when the Earth's atmosphere accumulated oxygen from photosynthetic bacteria. Corundum is a minor accessory mineral in metamorphic rocks and occurs in peraluminous igneous rocks. Partial solid solution is found between corundum and hematite. The gemstone ruby is Al_2O_3 with minor Cr^{3+}, while sapphire is Al_2O_3 with other chromophores. A modification of the corundum structure is found in the mineral ilmenite ($FeTiO_3$). This is an important accessory mineral in felsic igneous rocks. Above 950°C (1740°F), there is complete solid solution between hematite and ilmenite. Bixbyite (Mn_2O_3) is a distorted corundum structure resulting from the Jahn-Teller effect in Mn^{3+}. *See* BANDED IRON FORMATION; IGNEOUS ROCKS; IRON; METAMORPHIC ROCKS; RUBY; SAPPHIRE.

XO_2 oxides. Tetravalent cations, such as Ti^{4+}, Sn^{4+}, and Mn^{4+}, whose ionic radii favor sixfold coordination, adopt the rutile structure (**Fig. 3**). Rutile (TiO_2) is a common accessory mineral in felsic igneous rocks, gneisses, and schists. It also has two low-temperature polymorphs, anatase and brookite, but these are less common. Cassiterite (SnO_2) is the only significant ore mineral of tin. Cassiterite occurs mostly in granite-hosted hydrothermal deposits such as those in Cornwall, England. Pyrolusite (β-MnO_2) is found in low-temperature hydrothermal deposits. It is less common then previously supposed. MnO_2 also forms another polymorph (ramsdellite, α-MnO_2) based on double chains of MnO_6 polyhedra. This has a structure similar to that of goethite (α-FeOOH). Like pyrolusite, ramsdellite forms in low-temperature hydrothermal deposits. To complicate matters, a phase called nsutite (γ-MnO_2) is a disordered intergrowth of pyrolusite and ramsdellite that forms by the oxidation of manganese carbonate minerals. Synthetic nsutite is used in dry-cell batteries. *See* GNEISS; GRANITE; MANGANESE; SCHIST; TIN.

Large tetravalent cations, such as U^{4+} and Th^{4+}, prefer to be in eightfold coordination with oxygen and form oxides with the fluorite structure (**Fig. 4**). Uraninite (UO_2) is the most important ore of uranium and is a primary mineral in granites. *See* RADIOACTIVE MINERALS; THORIUM; URANIUM.

Spinel oxides. The spinel structure (**Fig. 5**) is adopted by oxides with the formula $X^{2+}Y_2^{3+}O_4$. The spinel structure has one tetrahedral cation site and two octahedral cation sites per four oxygens. In a normal spinel, the tetrahedral site is occupied by a divalent cation such as Mg^{2+}, Fe^{4+}, while the octahedral sites are occupied by trivalent cations such as Fe^{3+}, Cr^{3+}, or Al^{3+}. The inverse spinel structure is a variation where the tetrahedral sites are occupied by trivalent cations and the octahedral sites are occupied by a mixture of divalent and trivalent cations. A variety of solid solutions are possible within the spinel structure oxides.

The most important spinel structure oxide is magnetite (Fe_3O_4). Magnetite is an inverse spinel, so half of the Fe^{3+} cations are in the tetrahedral sites and the remaining Fe^{3+} cations, along with the Fe^{2+}

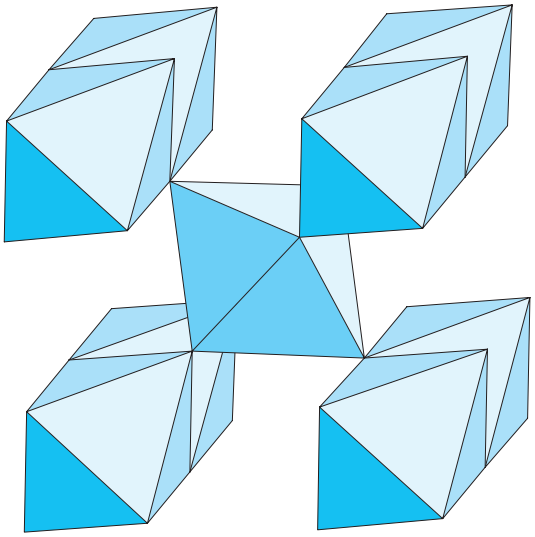

Fig. 3. Rutile structure adopted by XO_2 oxides.

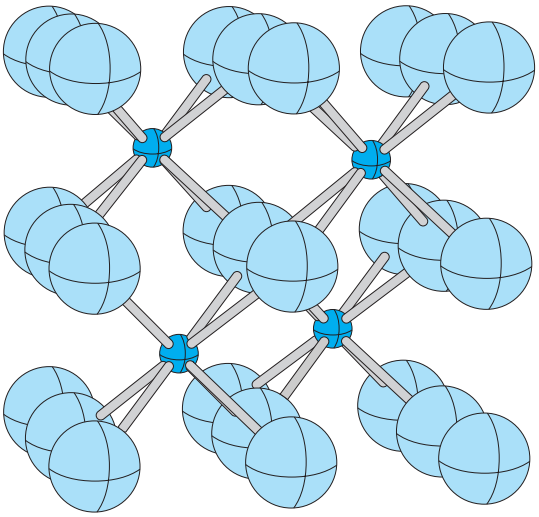

Fig. 4. Fluorite structure adopted by uraninite (UO_2) and thorianite (ThO_2).

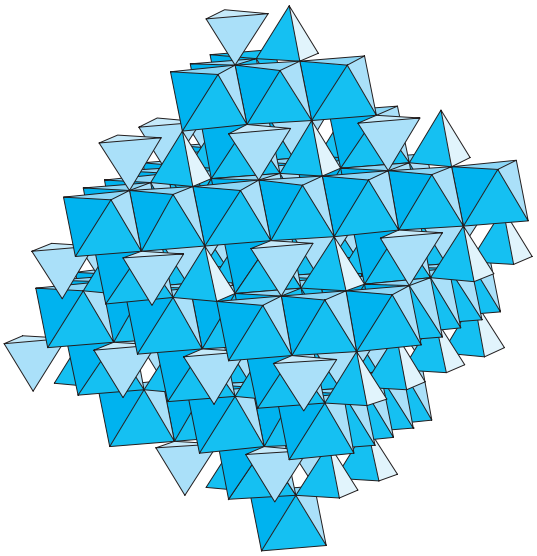

Fig. 5. Spinel structure adopted by XY_2O_4 oxides. In the normal spinel structure (for example, $MgAl_2O_4$), tetrahedral sites are occupied by divalent cations such as Mg^{2+} and octahedral sites are occupied by trivalent cations such as Al^{3+}. The inverse-spinel structure (for example, Fe_3O_4) has trivalent cations in the octahedral sites and a mixture of divalent and trivalent cations in the octoctahedral sites.

cations, are in the octahedral sites. Electron hopping between Fe^{2+} and Fe^{3+} cations in the octahedral sites gives magnetite a high electrical conductivity. The most important geophysical property of magnetite is its ferrimagnetism, with a Néel temperature, the temperature at which an antiferromagnetic material becomes paramagnetic, of 525°C (980°F). As an igneous rock cools, the magnetic moments of individual magnetite domains align with the Earth's magnetic field. This preserves a record of the orientation of the rock relative to the Earth's magnetic field at the time of crystallization. These paleomagnetic signatures in rocks were used to confirm the hypothesis of sea-floor spreading and continental drift. Magnetite often contains significant amounts of other cations such as Cr^{3+} and Ti^{4+}. A complete solid solution between Fe_3O_4 and Fe_2TiO_4 (ulvospinel) is stable above 600°C (1100°F). *See* ANTIFERROMAGNETISM; CHROMIUM; FERRIMAGNETISM; GEOMAGNETISM; MAGNETIC SUSCEPTIBILITY; PALEOMAGNETISM; TITANIUM.

TABLE 2. Manganese(III,IV) oxide minerals

Mineral	Formula
Chain Structures	
Pyrolusite	MnO_2
Ramsdellite	MnO_2
Tunnel Structures	
Hollandite	$BaMn_8O_{16}$
Romanechite	$(Ba,H_2O)_2Mn_5O_{10}$
Cryptomelane	$K_2Mn_8O_{16}$
Coronadite	$PbMn_8O_{16}$
Todorokite	$(Na,Ca)_{0.5}(Mn,Mg)_6O_{12} \cdot nH_2O$
Layer Structures	
Birnessite/Vernadite	$Na_{0.6}Mn_2O_4 \cdot 1.5H_2O$
Chalcophanite	$ZnMn_3O_7 \cdot 3H_2O$
Lithiophorite	$(Li,Al)(Mn^{4+},Mn^{3+})O_2(OH)_2$

The structure of hausmannite (Mn_3O_4), is a distortion of the spinel structure because of the Jahn-Teller effect for octahedrally coordinated Mn^{3+}. Hausmannite is found in high-temperature hydrothermal veins and in metamorphosed sedimentary manganese deposits but is not very common.

The spinel oxide, chromite ($FeCr_2O_4$), is the dominant ore mineral of Cr. Chromite occurs in ultramafic rocks and in the serpentinites that are derived from them; significant ore deposits are found in Iran and Zimbabwe. Because of the high octahedral site preference energy of Cr^{3+}, chromite has a normal spinel structure. *See* COORDINATION CHEMISTRY; SERPENTINITE.

Manganese(III, IV) oxides and oxide hydroxides. Manganese hydroxides and oxides containing Mn^{4+} and Mn^{3+} form a variety of structures based on chains, tunnels, and sheets of MnO_6 polyhedra (**Table 2**). The variations in the Mn oxidation state give variations in the charge of the MnO_6 sheets and tunnel/chain frameworks. The layer and framework charges are compensated by the incorporation of cations (such as K^+, Ba^{2+}, and Pb^{2+}) in the interlayer and tunnel sites. Perhaps the most important example is birnessite (**Fig. 6**) which is a mixed-valence Mn^{4+}-Mn^{3+} layer-structured oxide. This mineral, and the related phase vernadite ("δ-MnO_2"; probably an incoherently stratified birnessite), are major phases in marine ferromanganese nodules and crusts which form on the sea floor. At least two structural modifications are present for birnessite due to the presence of cation vacancies in the sheets; ordering of the vacancies can lower the symmetry of the sheets from hexagonal to triclinic. In the interlayer, cations, such as Li, Al, and Zn^{2+}, will adsorb above the vacancy sites to give structures such as lithiophorite $(Li,Al)(Mn^{4+},Mn^{3+})O_2(OH)_2$ and chalcophanite $(Li, Al)(Mn^{4+}, Mn^{3+})O_2(OH)_2$. The former occurs in manganese nodules formed in acid soils. Chalcophanite is much less common and forms in the oxidized zone of Zn-Mn ore deposits (such as the

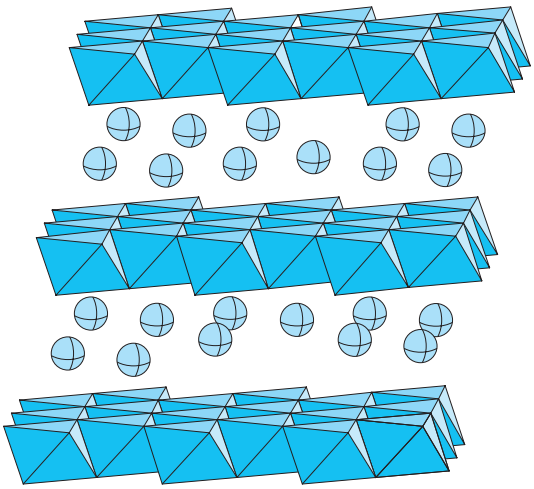

Fig. 6. Structure of birnessite. Between the MnO_2 layers are large exchangeable hydrated cations such as K^+, Ca^{2+}, Na^+.

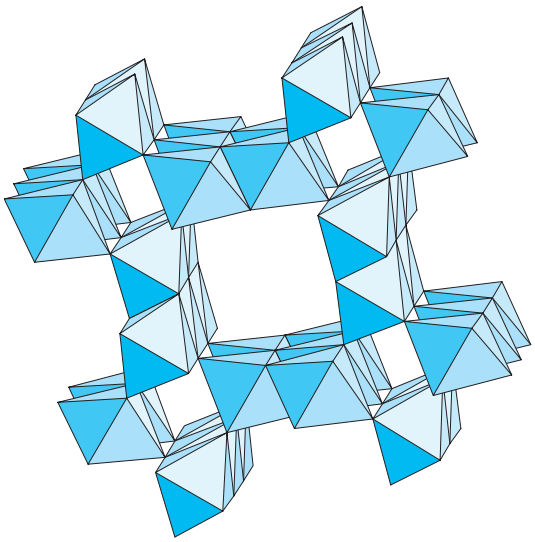

Fig. 7. The 2 × 2 tunnel structure adopted by hollandite and related manganese(IV, III) oxides. The tunnels can accommodate cations such as Ba^{2+}, K^+, Na^+, and Pb^{2+}. The same structure, but with CaI^- anions in the tunnel sites, is adopted by akaganeite (β-FeOOH).

much-studied Franklin and Stirling Hill locality in New Jersey). *See* MANGANESE NODULES.

The simplest tunnel structure (**Fig. 7**) is based on double chains of MnO_6 polyhedra; this is adopted by hollandite and related phases with formula $A_{0\text{-}2}(Mn^{4+}, Mn^{3+})_8(O,OH)_{16}$ (A = Ba, K, Pb, Na) [Table 2]. Todorokite, also found in marine manganese crusts and nodules, is a tunnel structure based on treble chains of MnO_6 polyhedra (**Fig. 8**). Incorporation of ions into manganese oxides must have important controls on the trace-element chemistry of the oceans.

Simple hydroxides. The simplest hydroxide mineral is brucite [$Mg(OH)_2$], the structure of which is based on $Mg(OH)_2$ layers that are held together by hydrogen bonds. Brucite forms during the alteration of magnesium silicates by hydrothermal fluids. It is not very common. The brucite structure also is adopted by hydroxides of other divalent cations such as Ca^{2+}, Fe^{2+}, and Ni^{2+}. The end member $Fe(OH)_2$ is easily partially oxidized, and the charge balance is maintained by the incorporation of various anions (such as CO_3^{2-}, SO_4^{2-}, and Cl^-) between the (Fe^{2+}, Fe^{3+})$(OH)_2$ layers to give "green rusts." These minerals are probably important phases in subsonic and anoxic soils and sediments. The mineral gibbsite [$Al(OH)_3$] has a structure based on layers of $Al(OH)_6$ octahedral with the layers held together by hydrogen bonds. *See* HYDROGEN BOND.

MOOH oxide hydroxides and related minerals. Trivalent cations, such as Fe^{3+} and Al^{3+}, form several oxide hydroxide structures. These minerals usually occur as clay-sized (<2 μm) particles in soils and sediments. Colloidal particles of oxide hydroxide minerals also are suspended in most natural waters. The surfaces of these minerals are quite reactive and strongly adsorb ions from aqueous solutions. In the environment, the aqueous concentrations of many trace micronutrients and toxic heavy metals are probably controlled by adsorption onto iron oxide hydroxide mineral

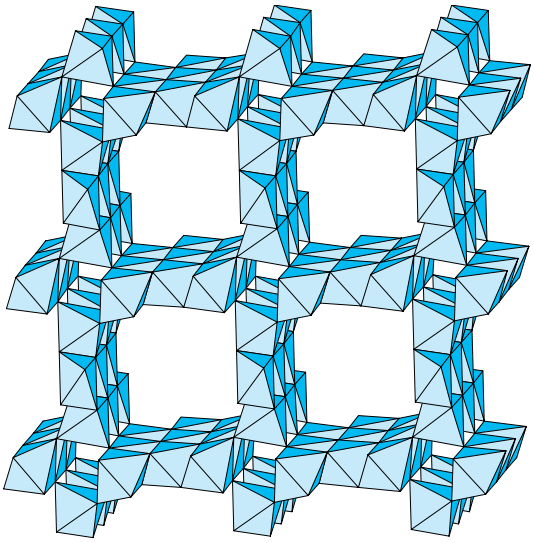

Fig. 8. Structure of todorokite. Within the 3 × 3 tunnels, exchangeable hydrated cations such as K^+, Mg^{2+}, and Ba^{2+} are present.

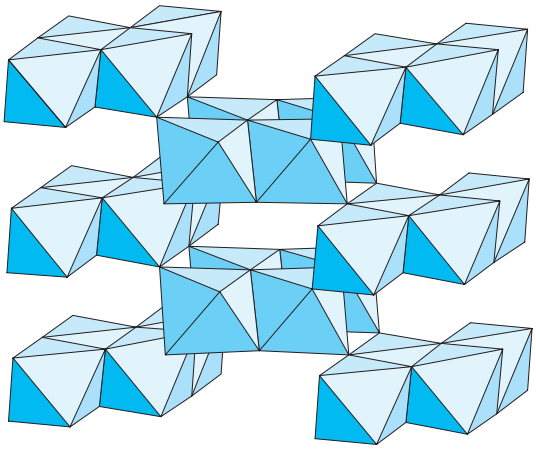

Fig. 9. Structure of goethite (α-FeOOH) and diaspore (α-AlOOH). The MnO_2 polymorph ramsdellite (α-MnO2) also adopts this structure.

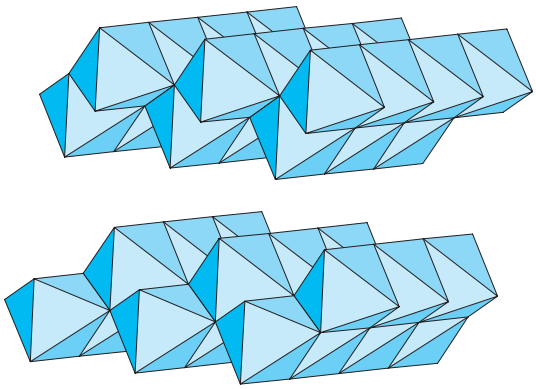

Fig. 10. Structure of lepidocrocite (γ-FeOOH) and boehmite (γ-AlOOH).

TABLE 3. MOOH and related minerals

Mineral	Formula
Goethite	α-FeOOH
Diaspore	α-AlOOH
Groutite	α-MnOOH
Lepidocrocite	γ-FeOOH
Boehmite	γ-AlOOH
Manganite	γ-MnOOH
Akaganeite	"β-FeOOH" = FeOOH ·HCl
Schwertmannite	FeOOH · H_2SO_4
Ferrihydrite	FeOOH—Fe_2O_3

surfaces. The most common FeOOH phases are goethite (α-FeOOH) and lepidocrocite (γ-FeOOH) [**Figs. 9** and **10**]. Goethite tends to form by hydrolysis of dissolved Fe^{3+}, while lepidocrocite forms by oxidation of green rust. The aluminum analogues, diaspore (α-AlOOH) and boehmite (γ-AlOOH), along with gibbsite ($Al(OH)_3$) are the minerals that make up bauxite, the ore of Al formed by weathering of primary silicate minerals such as feldspar in tropical soils. There is only limited solid solution between the FeOOH and AlOOH (or MnOOH) phases. Several other FeOOH minerals are known, but they are less common. Akaganeite ("β-FeOOH") forms by the hydrolysis of Fe^{3+} in chloride-bearing solutions and is a minor phase in marine sediments. Its structure is similar to that of hollandite but the tunnels are occupied by Cl^- anions. The mineral schwertmannite is believed to be similar to akaganeite, but with SO_4^{2-} anions occupying the tunnels. This mineral forms in acid mine drainage, probably by bacterially mediated oxidation of dissolved Fe^{2+}.

Perhaps the most ubiquitous FeOOH type mineral is ferrihydrite. This phase is poorly crystalline and forms by the rapid hydrolysis of dissolved Fe^{3+}. This is facilitated by bacterial oxidation of dissolved Fe^{2+} under less acidic conditions than those favoring schwertmannite. With time, ferrihydrite dissolves and recrystallizes to form the more stable phases goethite and hematite. Nevertheless, the extremely high, reactive surface area of ferrihydrite (up to 600 m^2/gram) means that it can have a strong effect on aqueous geochemistry by sorbing dissolved ions.

David M. Sherman

Bibliography. D. H. Lindsey (ed.), *Oxide Minerals*, Mineralogical Society of America, 1991; J. E. Post, Manganese oxide minerals: Crystal structures and economic and environmental significance, *Proc, Nat. Acad. Sci.*, 96:3447–3454, 1999; U. Schwertmann and R. M. Cornell, *Iron Oxides in the Laboratory: Preparation and Characterization*, 2d ed., Wiley-VCH, Weinheim, 2000.

Oxidizing agent

A participant in a chemical reaction that absorbs electrons from another reactant. In the process a component atom of this substance undergoes a decrease in oxidation number. In this action as an oxidizing agent, the substance undergoes reduction.

A measure of the effectiveness of a reagent as an oxidizing agent is its reduction potential. This is, in electrochemical terms, the equivalent of the free-energy change for the reduction process. The element with the highest reduction potential (and, therefore, the strongest oxidizing agent) is fluorine, F_2. The half-reaction in which fluorine absorbs an electron from another species in aqueous solution is shown by reaction (1). The reduction po-

$$e^- + {}^1/_2F_{2(gas)} \rightarrow F^-_{(aq)} \qquad (1)$$

tential E, although strictly applicable only to thermodynamically reversible systems, is a very useful measure of the tendency of a reaction component to undergo oxidation or reduction. For the F_2–F^- pair, it is given at 25°C (77°F) by Eq. (2),

$$E = 2.87 - 0.1362 \text{ In } \frac{a_{F^-}}{p_{F_2}^{1/2}} \text{ V} \qquad (2)$$

where a_{F^-} is the activity of fluoride ion in the solution, and p is the pressure, in atmospheres of fluorine gas. Under standard conditions, $a_{F^-} = 1$ and $p_{F^-} = 1$, the value for E is 2.87 V. This is E°, the standard reduction potential at 25°C (77°F). The F_2–F^- pair constitute a reduction-oxidation couple.

Thermodynamically, a substance is capable of oxidizing another if the corresponding couple has a lower reduction potential. When dichromate ion in acid solution acts as an oxidizing agent, it undergoes the reduction shown in reaction (3). The standard

$$Cr_2O_7{}^{2-} + 14H^+ + 6e^- \rightarrow 2Cr^3 + 7H_2O \qquad (3)$$

reduction potential at 25°C (77°F) for this reaction is 1.33 V. As this reaction does not proceed reversibly, the standard reduction potential represents only a lower limit. For the reduction of acetaldehyde to ethanol, shown in reaction (4), E° is 0.217 V. Dichro-

$$2e^- + CH_3CHO + 2H^+ \rightarrow CH_3CH_2OH \qquad (4)$$

mate can, therefore, fore, oxidize ethanol according to reaction (5). The corresponding electrochemical

$$Cr_2O_7{}^{2-} + 3CH_3CH_2OH + 8H^+ \rightarrow 3CH_3CHO + 2Cr^{3+} + 7H_2O \qquad (5)$$

potential is 1.11 V. The positive value for the resulting potential indicates a spontaneous process. Actually, in this example, the oxidation by dichromate will go beyond the aldehyde stage to produce the carboxylic acid. Dichromate ion in acid solution is a common oxidizing agent used in organic and analytical chemistry.

In oxidation-reduction (redox) systems composed of couples that are not widely separated, the actual reaction that occurs will depend upon the chemical activities in the system. For example, standard potentials for the $Fe^{3-}-Fe^{2-}$ and $Ag^{-}-Ag$ couples are 0.771 and 0.799 V, respectively. For the redox system shown in reaction (6), the resultant potential at 25°C (77°F) is given by Eq. (7). When the activity quotient

$$Fe^{2+} + Ag^{+} \rightleftharpoons Fe^{3+} + Ag \qquad (6)$$

$$E = 0.028 - 0.1362 \text{ In } \frac{a_{Fe^{3+}}}{a_{Fe^{2+}} a_{Ag^{+}}} \text{ V} \qquad (7)$$

$Q = (a_{Fe}{}^{3}+a_{Fe}{}^{2}+a_{Ag}+) < 1.23$, chemical change will occur from left to right; with $Q > 1.23$, from right to left; and with $Q = 1.23$, the system is at equilibrium and no net chemical change will occur. The above considerations are based only on thermodynamics and allow no prediction concerning the actual rate of the oxidation-reduction process. The practical effectiveness of a given oxidizing (or reducing) agent will depend upon both the thermodynamics and the available kinetic pathway for the reaction process. *See* CHEMICAL THERMODYNAMICS.

Substances that are widely used as oxidizing agents in chemistry include ozone (O_3), permanganate ion (MnO_4^-), nitric acid (HNO_3), as well as oxygen itself. Oxychlorine trifluoride, OF_3Cl, has been suggested and synthesized for possible use as an oxidizer in rocket engines. Organic chemists have empirically developed combinations of reagents to carry out specific oxidation steps in synthetic processes. Many of these utilize transition-metal oxides such as chromium trioxide, CrO_3, vanadium pentoxide, V_2O_5, ruthenium tetroxide, RuO_4, and osmium tetroxide, OsO_4. Oxidation by these species or by specific agents such as sodium perborate, $NaIO_4$, can often be restricted to a single molecular site.

The action of molecular oxygen as an oxidizing agent may be made more specific by photochemical excitation to an excited singlet electronic state. This can be accomplished through energy transfer from a dye, such as fluorescein, which absorbs light and transfers it to oxygen to form a relatively long-lived excited state. *See* PHOTOCHEMISTRY.

In an electrolytic cell, oxidation occurs at the anode, and the designation of a chemical oxidizing agent in the reaction is in general not possible. The electrode acts as an intermediary in the electron transfer. An important example of an anodic oxidation is the Kolbe reaction in which a carboxylic acid anion transfers an electron to the anode, forming a free radical [reaction (8)]. The radicals combine to form a hydrocarbon and CO_2, reaction (9).

$$RCOO^{-} \rightarrow \text{anode} RCOO + e^{-} \qquad (8)$$

$$2\ RCOO \rightarrow R_2 + CO_2 \qquad (9)$$

See ELECTROLYSIS.

Enzymes in living organisms catalyze the exothermic oxidation of carbohydrate, fat, and protein material as a source of energy for life processes. Certain microorganisms are able to oxidize organic compounds at C-H bonds to produce alcohols. These have been suggested as of possible use in cleaning oil spills. *See* BIOLOGICAL OXIDATION; OXIDATION-REDUCTION.

F. J. Johnson

Bibliography. W. R. Adams, *Photosensitized Oxidations*, 1971; D. Benson, *Mechanisms of Oxidation by Metal Ions*, 1976; G. S. Fonken and R. A. Johnson, *Chemical Oxidations with Microorganisms*, 1972; T. E. King, H. S. Mason, and M. Morrison (eds.), *Oxidases and Related Redox Systems*, Advances in the Biosciences Series, vols. 33 and 34, 1982; M. Murari, *The Mechanism of Oxidation of Organic Compounds by Electronic Oxidants*, 1985.

Oxime

One of a group of chemical substances with the general formula RR′C═N—OH, where R and R′ represent any carbon group or hydrogen. Oximes are derived from aldehydes (aldoximes, RHC═NOH) and ketones (ketoximes, RR′C═NOH, where R and R′ are not hydrogen). Oximes and oxime ethers (RR′C═N—OR″) have important pharmaceutical and synthetic applications. The oxime (and oxime ether) functional group is incorporated into many organic medicinal agents, including some antibiotics, for example, gemifloxacin mesylate (see structure); and pralidoxime chloride (see structure) and obidoxime chloride are used in the treatment of poisoning by organophosphate insecticides malathion and diazinon. *See* ALDEHYDE; KETONE.

Hydroxylamine (H_2NOH) reacts readily with aldehydes or ketones to give oximes. The rate of the reaction of hydroxylamine with acetone is greatest at pH 4.5. Oximes are formed by nucleophilic attack of hydroxylamine at the carbonyl carbon (C═O) of an aldehyde or ketone to give an unstable carbinolamine intermediate [reaction (1)]. Since the breakdown of the carbinolamine intermediate to an oxime is acid-catalyzed, the rate of this step is enhanced at low pH. If the pH is too low, however, most of the hydroxylamine will be in the nonnucleophilic protonated form (NH_3OH^+), and the rate of the first step will decrease. Thus, in oxime formation the pH has to be such that there is sufficient free hydroxylamine for the first step and enough acid so that dehydration of the carbinolamine is facile. *See* REACTIVE INTERMEDIATES.

Oximes can also be prepared by acid- or base-catalyzed oximation of active methylene compounds with alkyl nitrites [reaction (2)]. The nitroso compound that is first formed in this reaction rapidly rearranges to the more stable oxime tautomer. *See* TAUTOMERISM.

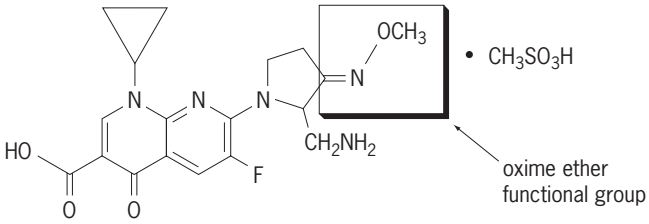

Gemifloxacin mesylate

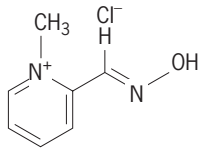

Pralidoxime chloride

Photolysis of alkyl nitrite esters [reaction (3)] is useful in conformationally rigid systems, and leads to oxime formation at a δ-carbon atom. Again, a nitroso tautomer rearanges to the oxime. This reaction has been used to convert cortisone acetate into aldosterone acetate oxime. *See* NITRO AND NITROSO COMPOUNDS.

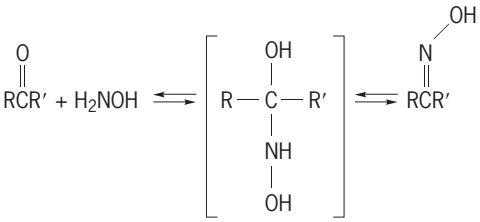

(1)

Carbinolamine intermediate

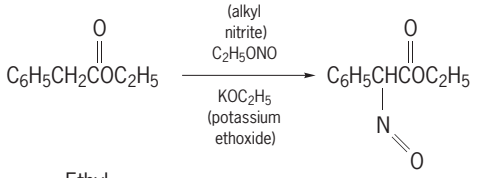

Ethyl phenylacetate

Nitroso tautomer

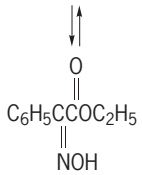

(2)

Oxime

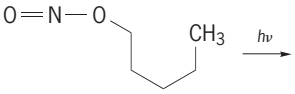

Alkylnitrite ester

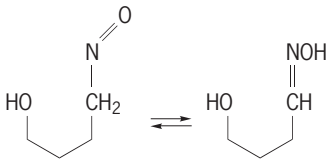

(3)

Nitroso tautomer

Oxime

An oxime can act as both a weak acid and a weak base ($pK_b \cong 12$ and $pK_a \cong 11$). Oxime anions are capable of reacting at two different sites with alkyl halides (R″X). Anions of this type are referred to as ambident (literally, two-sided tooth) nucleophiles. In the case of oxime anions, the reaction with an alkyl halide (an alkylation reaction) can give an oxime ether or a nitrone [reaction (4)].

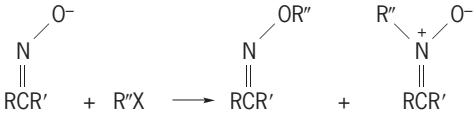

(4)

Alkyl halide

Oxime ether

Nitrone

See ACID AND BASE.

The *Z* and *E* isomers (formerly referred to as syn and anti isomers) of many oximes, such as butyrophenone oxime (see structure), are known. The *E* and *Z* isomers of oxime ethers (RR′C═NOR″) are particularly resistant to thermal isomerization.

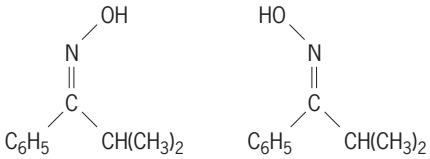

(anti) (*E*)-Butyrophenone oxime

(syn) (*Z*)-Butyrophenone oxime

Many of the original (prior to 1921) assignments of configuration of the *Z* and *E* isomers were based on chemical reactions (the Beckman rearrangement) and were in error because of incorrect assumptions concerning the stereochemistry of these reactions. The assignment of configurations of oximes, when both isomers are available, can be made from ^{1}H and ^{13}C nuclear magnetic resonance spectra. *See* MOLECULAR ISOMERISM; STEREOCHEMISTRY.

The *O*-methyloxime prepared from reaction of acetophenone with hydroxylamine followed by alkylation with methyl iodide has the *E* configuration (OCH_3 and C_6H_5 on opposite sides of the double bond) [reaction (5)]. Ultraviolet irradiation of benzene solution of the *E* isomer gives a mixture of the *E* and *Z* isomers from which the *Z* isomer is obtained by chromatography. Reversion to the more stable *E* isomer can be accomplished by acid-catalyzed isomerization (hydrogen chloride in dioxane) of the *Z* from [reaction (5)].

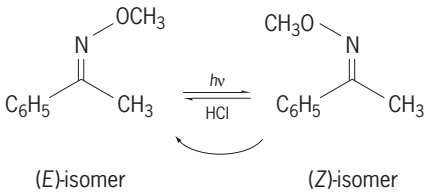

(5)

One of the best-known reactions of oximes is their rearrangement to amides. This reaction, the Beckmann rearrangement, can be carried out with a variety of reagents [such as phosphorus pentachloride (PCl_5), concentrated sulfuric acid (H_2SO_4), and perchloric acid ($HClO_4$)] that induce the rearrangement by converting the oxime hydroxyl group into a group of atoms that easily departs in a displacement reaction (a good leaving group) by either protonation or formation of a derivative. The Beckmann rearrangement has been shown to be a stereospecific process in which the group anti to the leaving group undergoes a 1, 2-shift to give a nitrilium ion. The nitrilium ion reacts with water to form the amide [reaction (6)].

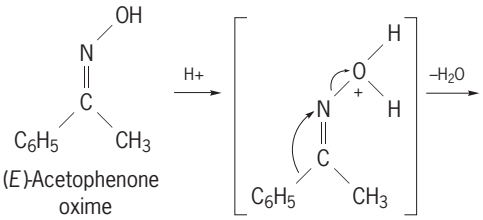

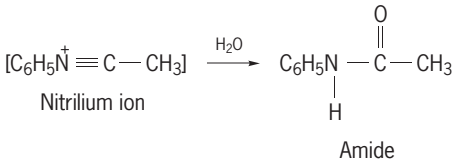

(6)

The Beckmann rearrangement can be carried out with a trialkylalane, which provides a Lewis acid for the rearrangement as well as a nucleophile to react with the intermediate nitrilium ion. For example, the tosylate of cyclohexanone oxime methianesulfonate reacts with tri-*n*-propylalane [$(n\text{-}C_3H_7)_3Al$] to give an imine, which is reduced with diisobutylaluminum hydride to 2-propyl-l-azacycloheptane [reaction (7)].

The industrial synthesis of ϵ-caprolactam is carried out by a Beckmann rearrangement on cyclohexanone oxime. ϵ-Caprolactam is polymerized to the polyamide known as nylon 6, which is used in tire cords [reaction (8)].

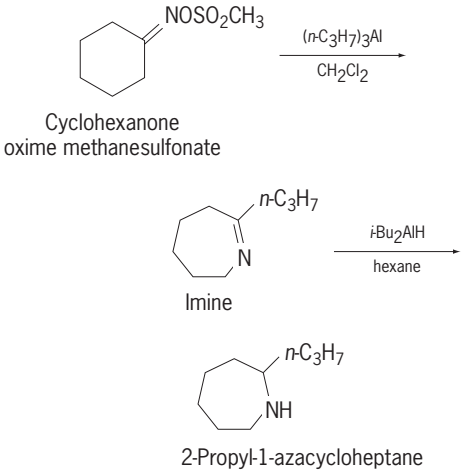

(7)

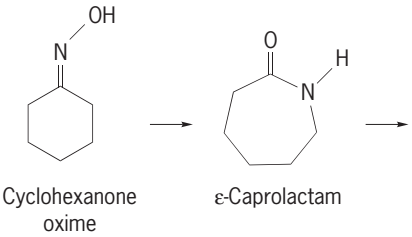

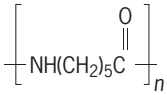

(8)

Nylon 6

The conversion of oxime tosylates into α-amino ketones is known as the Neber rearrangement [reaction (9)]. The intermediate azirine is usually not

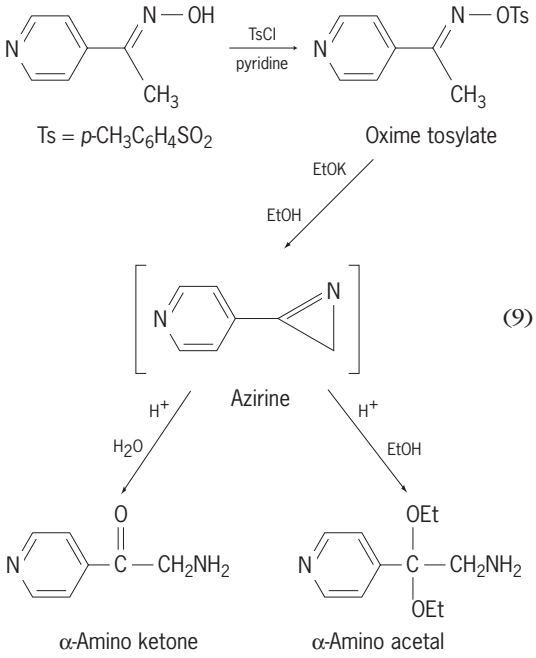

(9)

isolated and is reacted with aqueous acid to give the α-amino ketone. Reaction of the azirine with acid in anhydrous ethanol gives an α-amino acetal.

Aldoximes are dehydrated with acetic anhydride to nitriles (RC≡N). Oxidation of oximes with trifluoroperacetic acid at controlled pH gives nitro compounds ($RR'CHNO_2$). Oximes can be reduced to amines with lithium tetrahydridoaluminate ($LiAlH_4$) or by catalytic hydrogenation. Reductions of oxime ethers with a chiral oxazaborolidine give optically active amines with high enantioselectivity [reaction (10)]. Reduction of oximes with sodium cyanotrihydridoborate ($NaBH_3CN$) gives hydroxylamines. *See* HYDROGENATION.

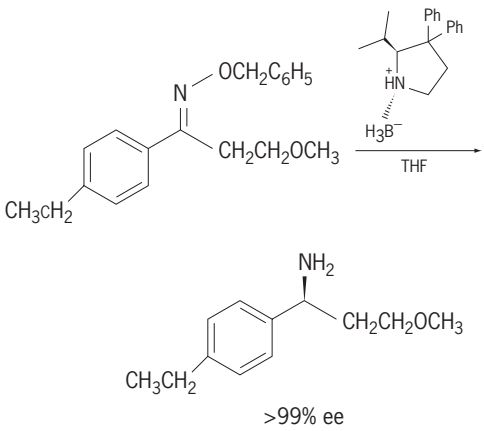

(10)

There are many more complex functional groups that contain the oxime moiety, including:

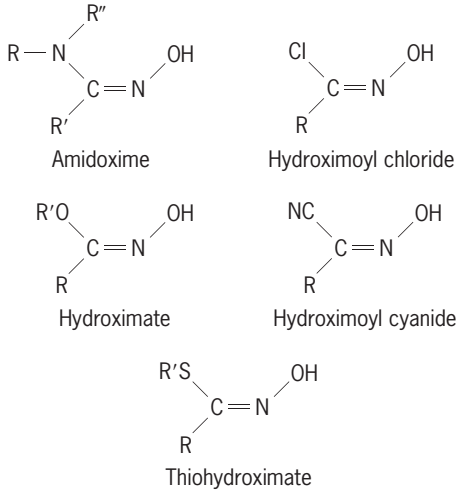

See ORGANIC SYNTHESIS. James E. Johnson

Bibliography. B. R. Brown, *The Organic Chemistry of Aliphatic Nitrogen Compounds*, 1994; F. A. Carey and R. J. Sundberg, *Advanced Organic Chemistry, Part B: Reactions and Synthesis*, 4th ed., 2001; J. P. Freeman (ed.), *Organic Syntheses Collective Volume*, 8, pp. 568–572,1993; J. P. Freeman (ed.), *Organic Syntheses Collective Volume* 7, pp. 149–152,1990; R. E. Gawley, The Beckmann reaction: Rearrangements, elimination-additions, fragmentations, and rearrangement-cyclizations, *Org. React.*, 35:1-420, 1988; L. S. Hegedus (ed.), *Organic Syntheses*, 79:130–138,2002; M. B. Smith and J. March, *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure*, 5th ed.,2001.

Oximetry

Any of various methods for determining the oxygen saturation of blood. Oximeters are instruments that measure the saturation of blood in arteries by transilluminating intact tissue, thus eliminating the need to puncture vessels. Aside from the major advantage of being noninvasive and therefore free of risk, oximeters permit long-term monitoring in a variety of clinical settings, giving a continuous record of oxygen saturation.

Blood color. When dark venous blood passes through the pulmonary circulation or is exposed to air, it becomes bright red. In severe lung disease, the hemoglobin in the blood fails to become adequately oxygenated in its passage through the lungs and the mucous membranes lose their characteristic pink color. The mucous membranes then appear dusky or blue, a physical sign known as cyanosis. Oximeters determine oxygen saturation by making use of this color change; they compare photoelectrically the differences in light absorption between oxygenated and reduced hemoglobin. *See* HEMOGLOBIN.

Light absorption. The relationship between the concentration of a colored substance, length of the light path, and light absorption is expressed by the Lambert-Beer law; Eq. (1), where I is the amount

$$I = I_0 e^{-E' cd} \tag{1}$$

of transmitted light, I_0 is the impingent light, c is concentration, d is the distance traveled by the light through the solution, and E' is a constant at a given wavelength of light for each substance, known as the extinction coefficient. The Lambert-Beer law may also be written in logarithmic form as Eq. (2) where

$$E = \frac{I}{cd} \log \frac{I_0}{I} \tag{2}$$

log (I_0/I) is referred to as optical density (D). Thus, if incident light intensity and depth are held constant, the concentration of a substance in solution is a logarithmic function of the intensity of transmitted light.

As the wavelength of light increases from 300 to 1200 nanometers, the absorption of light by hemoglobin and oxyhemoglobin decreases, each following a slightly different pattern (**Fig. 1**). Between 600 and 800 nm, oxyhemoglobin absorbs less light than reduced hemoglobin, so that more red light is transmitted to produce the bright red color of oxygenated blood. In the infrared portion of the spectrum at 805 nm, there is an isobestic point, a wavelength at which reduced hemoglobin and oxygenated hemoglobin absorb identical amounts of light.

The Lambert-Beer relationship describes the absorption of monochromatic light passing at right angles through glass cuvettescontaining homogeneous

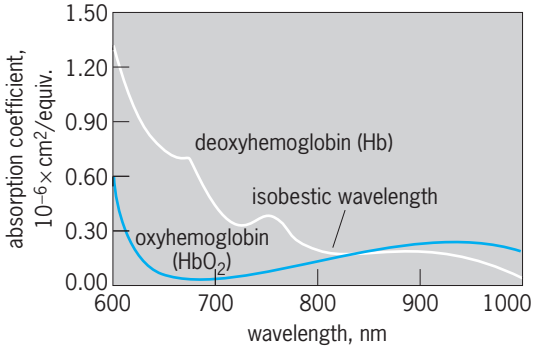

Fig. 1. Absorption coefficients of oxyhemoglobin and deoxyhemoglobin in the red and near-infrared regions. (*After S. Takatani, P. W. Cheung, and E. A. Ernst, Ann. Biomed. Eng., 8:1–15, 1980*)

solutions. While its validity has been demonstrated for hemoglobin solutions and hemolyzed blood in laboratory solutions, it may be inadequate to describe the more complex behavior of light in nonhemolyzed blood. Intact erythrocytes scatter, refract, and reflect incident light. When photometry is applied to hemoglobin in intact blood vessels, usually by transilluminating the helix or lobe of the ear, the situation is further complicated as the ear is an imperfect cuvette. Allowances must thus be made for absorption of light by tissue pigments such as melanin and bilirubin and for the variable depth of the cuvette (ear thickness) from subject to subject.

Ear oximeters. The original ear oximeters used a single wavelength of visible red light to monitor oxyhemoglobin levels. Later models used visible red and green light, allowing a distinction to be made between the effects of changes in hemoglobin concentration and saturation. The Millikan oximeter assessed the ear's blood thickness with an infrared cell, and arterial saturation with green light. The oximeter developed by Earl H. Wood developed the concept of making the ear "bloodless" by inflating a pressure capsule in the earpiece to create a "blank" cuvette, against which the perfused ear could be compared. In the 1970s, eight-wavelength oximeters were developed in which calibration was performed automatically against internal standards. These instruments were unaffected by variations in skin pigmentation and measured changes in arterial oxygen saturation almost immediately, continuously, and with considerable accuracy (**Fig. 2**). Unfortunately, the earpieces, as in earlier instruments, were cumbersome and uncomfortable to wear for prolonged periods of time.

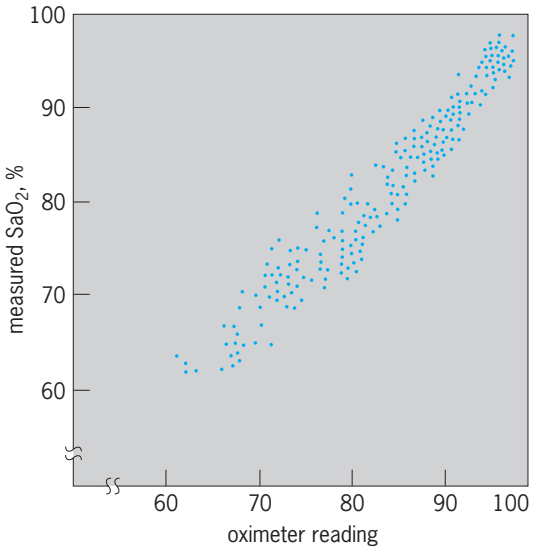

Fig. 2. Comparisons of 223 samples of arterial blood O_2 saturation (SaO_2) using an ear oximeter. (*After N. A. Saunders, A. C. P. Powles, and A. S. Rebuck, Ear-oximetry: Accuracy and practicality in the assessment of arterial oxygenation, Amer. Rev. Resp. Dis., 113:745, 1976*)

Modern two-wavelength oximeters are known as pulse oximeters. These devices estimate oxygen saturation by the simple expedient of ignoring all nonpulsatile light absorbance. During the heart's systolic contractions, the inflow of arterial blood increases light absorbance across perfused tissues. Changes in incident light absorption from the nonflow baseline reflects absorbance by arterial blood. Incident light is emitted by two alternating light-emitting diodes—one in the visible red region (660 nm) and one in the infrared region (940 nm)—to determine the ratio of oxyhemoglobin to deoxyhemoglobin. Small light-emitting diodes are readily incorporated into a variety of probes; sites for transillumination include the earlobe, finger, toe, bridge of the nose and, in neonates, palms of the hand or foot. Pulse oximeters are regarded as accurate in the medically relevant range of 60–100% saturation.

Applications. Oximeters are used widely in modern hospital settings, where their ability to reflect oxygen saturation changes within 1–2 s can alert caregivers to the presence of life-threatening episodes of hypoxemia (inadequate oxygenation). In addition to monitoring the critically ill, oximeters are used routinely in the operating room to safeguard individuals having routine anesthetics. The growing awareness of sleep apnea and other disorders of breathing during sleep has led to widespread use of oximeters in sleep laboratories, where they are invaluable in the diagnostic process. The widespread use of accurate and continuous oxygen-monitoring devices has shown intermittent episodes of hypoxemia to occur much more frequently than was previously suspected. Guidelines for the prescribing of oxygen and other elements of care are being researched and revised in light of this wealth of new diagnostic information.

Kenneth R. Chapman; A. S. Rebuck

Bibliography. K. R. Chapman, A. D'Urzo, and A. S. Rebuck, The accuracy and response characteristics of a simplified ear oximeter, *Chest*, 83(6):860–864, 1983; K. R. Chapman et al., Range of accuracy of two wavelength oximetry, *Chest*, 89:540–542, 1986; M. H. Kryger, T. Roth, and W. C. Dement (eds.), *Principles and Practice of Sleep Medicine*, 2d ed., 1994; J. W. Severinghaus and J. F. Kelleher, Recent developments in pulse oximetry, *Anesthesiology*, 76:1018–1038, 1992.

Oxygen

A gaseous chemical element, O, atomic number 8, and atomic weight 15.9994. Oxygen is of great interest because it is the essential element both in the respiration process in most living cells and in combustion processes. It is the most abundant element in the Earth's crust. About one-fifth (by volume) of the air is oxygen. *See* PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Oxygen is separated from air by liquefaction and fractional distillation. The chief uses of oxygen in order of their importance are (1) smelting, refining, and fabrication of steel and other metals; (2) manufacture of chemical products by controlled oxidation; (3) rocket propulsion; (4) biological life support and medicine; and (5) mining, production, and fabrication of stone and glass products.

Uncombined gaseous oxygen usually exists in the form of diatomic molecules, O_2, but oxygen also exists in a unique triatomic form, O_3, called ozone. *See* OZONE.

Under ordinary conditions oxygen is a colorless, odorless, and tasteless gas. It condenses to a pale blue liquid, in contrast to nitrogen, which is colorless in the liquid state. Oxygen is one of a small group of slightly paramagnetic gases, and it is the most paramagnetic of the group. Liquid oxygen is also slightly paramagnetic. Some data on oxygen and some properties of its ordinary form, O_2, are listed in the **table**. *See* PARAMAGNETISM.

Properties of oxygen

Property	Value
Atomic number	8
Atomic weight	15.9994
Triple point (solid, liquid, and gas in equilibrium)	−218.80°C (−139.33°F)
Boiling point at 1 atm pressure	−182.97°C (−119.4°F)
Gas density at °C and 10^5 Pa pressure, g/liter	1.4290
Liquid density at the boiling point, g/ml	1.142
Solubility in water at 20°C, oxygen (STP) per 1000 g water at 10^5 Pa partial pressure of oxygen	30

Practically all chemical elements except the inert gases form compounds with oxygen. Most elements form oxides when heated in an atmosphere containing oxygen gas. Many elements form more than one oxide; for example, sulfur forms sulfur dioxide (SO_2) and sulfur trioxide (SO_3). Among the most abundant binary oxygen compounds are water, H_2O, and silica, SiO_2, the latter being the chief ingredient of sand. Among compounds containing more than two elements, the most abundant are the silicates, which constitute most of the rocks and soil. Other widely occurring compounds are calcium carbonate (limestone and marble), calcium sulfate (gypsum), aluminum oxide (bauxite), and the various oxides of iron which are mined as a source of iron. Several other metals are also mined in the form of their oxides. Hydrogen peroxide, H_2O_2, is an interesting compound used extensively for bleaching. *See* HYDROGEN PEROXIDE; OXIDATION-REDUCTION; OXIDE; PEROXIDE; WATER.

Arthur W. Francis, Sr.; Lawerence M. Sayre

Bibliography. C. E. Baukal (ed.), *Oxygen-Enhanced Combustion*, Springer, 1998; G. W. Everett and G. W. Everett, Jr., *Preparing and Studying Oxygen and Some of Its Compounds*, 1992; C. Foote et al. (eds.), *Active Oxygen in Chemistry*, Springer, 1996; D. MacKay et al., *Illustrated Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals: Oxygen, Nitrogen, and Sulfur Containing Compounds*, 1995; D. T. Sawyer, *Oxygen Chemistry*, 1991; J. S. Valentine et al. (eds.), *Active Oxygen in Biochemistry*, Springer, 1995.

Oxygen toxicity

A toxic effect in a living organism caused by a species of oxygen. Oxygen has two aspects, one benign and the other malignant. Those organisms that avail themselves of the enormous metabolic advantages provided by dioxygen (O_2) must defend themselves against its toxicity. The complete reduction of one molecule of O_2 to two of water (H_2O) requires four electrons; therefore, intermediates must be encountered during the reduction of O_2 by the univalent pathway. The intermediates of O_2 reduction, in the order of their production are the superoxide radical O_2^-), hydrogen peroxide (H_2O_2), and the hydroxyl radical (HO·). *See* OXYGEN; SUPEROXIDE CHEMISTRY.

The intermediates of oxygen reduction, rather than O_2 itself, are probably the primary cause of oxygen toxicity. It follows that defensive measures must deal with these intermediates. The superoxide radical is eliminated by enzymes that catalyze the reaction below. These enzymes, known as superoxide

$$O_2^- + O_2^- + 2H^+ \rightleftharpoons H_2O_2 + O_2$$

dismutases, have been isolated from a wide variety of living things; they have been found to contain iron, manganese, or both copper and zinc at their active sites. Elimination of superoxide dismutases from

oxygen-tolerant bacteria renders them intolerant of O_2, whereas exposure to elevated levels of O_2 elicits an adaptive increase in the biosynthesis of these enzymes. Both of these responses suggest that O_2^- is a major cause of the toxicity of O_2.

Hydrogen peroxide (H_2O_2) must also be eliminated, and this is achieved by two enzymatic mechanisms. The first of these is the dismutation of H_2O_2 into water and oxygen, a process catalyzed by catalases. The second is the reduction of H_2O_2 into two molecules of water at the expense of a variety of reductants, a process catalyzed by peroxidases. Plant peroxidases are hemoenzymes, that is, enzymes that contain heme as a prosthetic group; they can use ascorbate, phenols, or amines as the reductants of H_2O_2. A selenium-containing peroxidase that uses the tripeptide glutathione as the reductant of H_2O_2 can be found in animal cells. *See* ENZYME.

The multiplicity of superoxide dismutases, catalases, and peroxidases, and the great catalytic efficiency of these enzymes, provides a formidable defense against O_2^- and H_2O_2. If these first two intermediates of O_2 reduction are eliminated, the third (HO·) will not be produced. No defense is perfect, however, and some HO· is produced; therefore its deleterious effects must be minimized. This is achieved to a large extent by antioxidants, which prevent free-radical chain reactions from propagating. For example, the human organism depends upon α-tocopherol (vitamin E) to prevent such chain reactions within the hydrophobic core of membranes and depends upon ascorbate and glutathione to serve the same end in the aqueous compartments of cells. *See* ASCORBIC ACID; ANTIOXIDANT; CHAIN REACTION (CHEMISTRY); FREE RADICAL; PEPTIDE.

Some damage due to oxygen-derived free radicals is sustained on a continuing basis, in spite of the existing multilayered defenses, and must be repaired. Thus, there are enzymes that recognize and hydrolyze oxidatively damaged proteins, and there are other enzymes that repair oxidative damage to deoxyribonucleic acid (DNA). Indeed, analysis of human urine for oxidized pyrimidines, which are removed from DNA during such repair, indicates the occurrence of at least 1000 such events per cell per day.

Thus, the apparent comfort in which aerobic organisms live in the presence of an atmosphere that is 20% O_2 is due to a complex and effective system of defenses against this peculiar gas. Indeed, these defenses are easily overwhelmed, and overt symptoms of oxygen toxicity become apparent when organisms are exposed to 100% O_2. For example, a rat maintained in 100% O_2 will die in 2 to 3 days.

Irwin Fridovich

Bibliography. I. B. Afanas'ev, *Superoxide Ion: Chemistry and Biological Implications*, 1989; I. Fridovich, Superoxide dismutases: An adaptation to a paramagnetic gas, *J. Biol. Chem.*, 264:7761–7764, 1989; H. Jonje, Genetic toxicology of oxygen, *Mutation Res.*, 219:193–208, 1989; A. Touati, Molecular genetics of superoxide dismutases, *Free Radical Biol. Med.*, 5:393–402, 1988.

Oxymonadida

An order of class Zoomastigophorea in the phylum Protozoa. These are colorless flagellate symbionts in the digestive tract of the roach *Cryptocercus* and of certain termites. They are xylophagous; that is, they ingest wood particles taken in by the host. Seven or more genera of medium or large size have been identified, the organisms varying from pyriform to ovoid in shape. At the anterior end a pliable neck-like rostrum attaches the organism to the host intestinal wall, but they are sometimes free (see **illus**.). They can be either uni- or multinucleate. These organisms are termed karyomastigonts and each gives rise to two pairs of flagella in the unattached cells, two flagella to each blepharoplast. In the rostrum there is an axostylar apparatus, fibrils which pass to, and emerge at the posterior part of, the body. The nuclei contain long threadlike persistent chromosomes which appear to pass at random onto an intranuclear spindle. Isogamous (union of similar gametes) sexual processes have been described for three genera. In some cases, at least, these parallel the molting of the host.

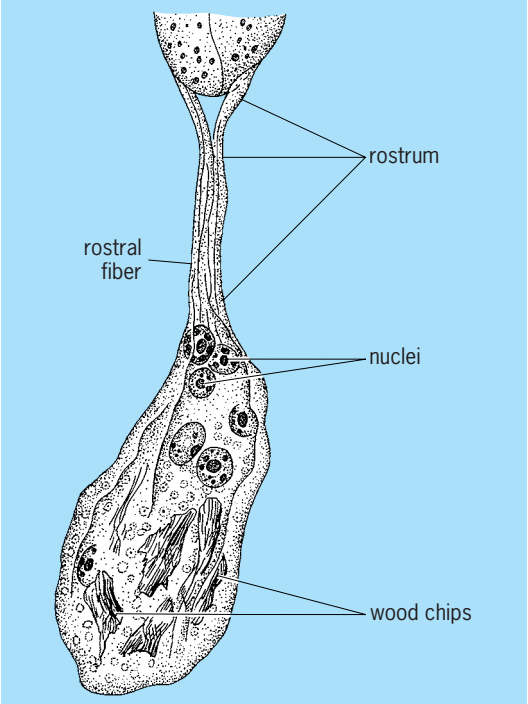

An oxymonad, *Microrhopalodina inflata*.

A peculiar feature is the investment of the cell cuticle with bacteria and spirochetes, often resulting in a thick coating. The genera described have a global distribution. *See* CILIA AND FLAGELLA; PROTOZOA; SARCOMASTIGOPHORA; ZOOMASTIGOPHOREA.

James B. Lackey

Bibliography. L. R. Cleveland et al., The wood feeding roach *Cryptocerus*: Its protozoa and the symbiosis between protozoa and roach, *Mem. Amer. Acad. Arts Sci.*, 17:185–342, 1934; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols.,

1982; J. A. Pechenik, *Biology of the invertebrates*, 5th ed., McGraw-Hill Higher Education: Dubuque, 2005.

Ozone

A powerfully oxidizing allotropic form of the element oxygen. The ozone molecule contains three atoms (O_3), while the more common oxygen molecule has two atoms (O_2).

Ordinary oxygen is a colorless gas and condenses to a very pale blue liquid, whereas ozone gas is decidedly blue, and both liquid and solid ozone are an opaque blue-black color, similar to that of ink. Even at concentrations as low as 4%, the blue color of ozone gas mixed with air or other colorless gas in a tube 1 in. (2.5 cm) or more in diameter and 4 ft (1.2 m) or more long can be seen by looking lengthwise through the tube.

Properties and uses. Some properties of ozone are given in the **table**. Ozone has a characteristic, pungent odor familiar to most persons because ozone is formed when an electrical apparatus produces sparks in air. Ozone is irritating to mucous membranes and toxic to human beings and lower animals. U.S. Occupational Safety and Health Administration standards for industrial workers exposed to ozone on a daily basis limit ozone concentration to 0.1 part per million on the average, with a maximum of 0.3 ppm for short exposures.

High ozone concentrations in liquid- and gas-phase mixtures can decompose explosively when initiated by an electric spark or other high-level energy source. Controlled decomposition to reduce ozone to desirable low concentrations can be accomplished catalytically.

Ozone is a more powerful oxidizing agent than oxygen, and oxidation with ozone takes place with evolution of more heat and usually starts at a lower temperature than when oxygen is used. In the presence of water, ozone is a powerful bleaching agent, acting more rapidly than hydrogen peroxide, chlorine, or sulfur dioxide. *See* OXIDIZING AGENT.

Ozone is utilized in the treatment of drinking-water supplies. Odor- and taste-producing hydrocarbons are effectively eliminated by ozone oxidation.

Some properties of ozone

Property	Value
Density of the gas at 0°C (32°F), 1 atm (10^5 Pa) pressure	2.154 g/liter
Density of the liquid	
−111.9°C	1.354 g/ml
−183°C	1.573 g/ml
Boiling point at 1 atm (10^5 Pa) pressure	−111.9°C (−169.4°F)
Melting point of the solid	−192.5°C (−314.5°F)
Wavelength range of maximum absorption in visible spectrum	560–620 nm
Wavelength range of maximum absorption in the ultraviolet spectrum	240–280 nm

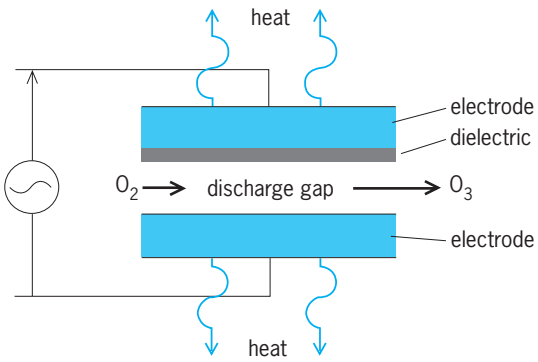

Fig. 1. Diagram of a generic corona cell.

Iron and manganese compounds which discolor water are diminished by ozone treatment. Compared to chlorine, bacterial and viral disinfection with ozone is up to 5000 times more rapid. After treatment, the residual chlorine content leaves a characteristic undesirable taste and odor. In addition, chlorine may yield chloroform and other trihalomethane (THM) compounds which are potentially carcinogenic. *See* WATER TREATMENT.

Plants that use oxygen in aerobic digestion of sewage can add ozone treatment at reduced cost. Ozone can be produced more economically from pure oxygen. By proper integration of the facilities, oxygen not transformed into ozone in its generator passes through the ozonization tank into the aerobic digester with very high efficiency. *See* SEWAGE TREATMENT.

Ozone undergoes a characteristic reaction with unsaturated organic compounds in which the double or triple bond is attacked, even at temperatures as low as −150°F (−100°C), with the formation of ozonides; these ozonides can be hydrolyzed, oxidized, reduced, or thermally decomposed to a variety of compounds, chiefly aldehydes, ketones, or carboxylic acids. Double (C═C) bonds are almost always ruptured in this reaction. Commercially ozonolysis (ozonation followed by decomposition of the ozonide) is employed in the production of azelaic acid and certain chemical intermediates used in the drug industry. *See* OZONOLYSIS.

Natural occurrence. Ozone occurs to a variable extent in the Earth's atmosphere. Near the Earth's surface the concentration is usually 0.02–0.03 ppm in country air, and less in cities except when there is smog; under smog conditions in Los Angeles ozone is thought to be formed by the action of sunlight on oxygen of the air in the presence of impurities, and on bad days the ozone concentration may reach 0.5 ppm or more for short periods of time.

At vertical elevations above 12 mi (20 km), ozone is formed by photochemical action on atmospheric oxygen. Maximum concentration of 5×10^{12} molecules/cm^3 (more than 1000 times the normal peak concentration at Earth's surface) occurs at an elevation of 19 mi (30 km).

Intercontinental air transports cruise at altitudes of 7.5 to 11 mi (12 to 17 km). On flights through northern latitudes, significant concentrations (up to

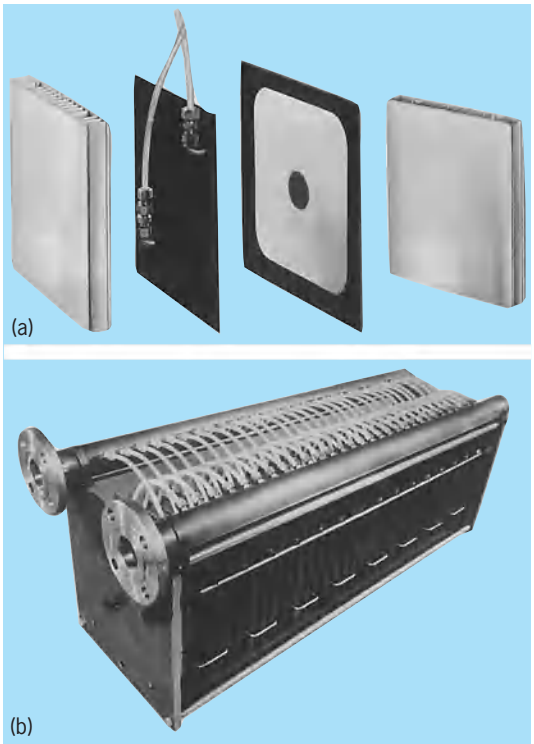

Fig. 2. Lowther cell for ozone generation: (*a*) expanded view of a single cell; (*b*) 30-cell module.

1.2 ppm) of ozone have been encountered. At these levels ozone can cause coughing and chest pains, especially for cabin attendants who are actively working. Carbon filters and catalytic ozone-decomposing equipment have been installed to eliminate the problem.

Absorption of solar ultraviolet radiation by ozone provides enough energy to raise the temperature of the stratosphere (6–30 mi or 10–50 km) significantly above that of the upper troposphere. This increase of temperature with increasing height forms a stable layer resistant to vertical mixing. Gases injected into the stratosphere above 12 mi (20 km) may remain 2 years or longer.

By absorbing most of the short-wavelength light, the ozone layer protects human and other life forms. The layer is thinnest at the Equator, where it permits more ultraviolet radiation to reach ground levels in the torrid zone. This is believed to account for the high incidence of skin cancer in equatorial areas.

The dissociation of ozone to oxygen is catalyzed by several chemicals, especially nitrogen oxides and chlorine. Cosmic rays form nitric oxide in the stratosphere. As solar activity causes Earth's magnetic field to increase, cosmic rays are deflected away from Earth. Consequently, there is less nitric acid and more ozone immediately following the maximum phase of the solar activity cycle.

Volcanic eruptions and cosmic rays result in increased levels of chemicals which dissociate ozone. Above-normal levels of these natural events in previous geologic ages are believed to have reduced the ozone layer to 10% of normal. The resulting increase in ultraviolet radiation reaching the Earth's surface may account for the sudden extinction of some species.

Human activities also influence the ozone layer. Nuclear explosions in the atmosphere, and supersonic aircraft cruising at altitudes around 12 mi (20 km), inject nitric oxide into the stratosphere. A still larger effect may be developing from the release of certain relatively stable fluorocarbons, especially $CFCl_3$ and CF_2Cl_2. This type of compound is believed to remain intact for many years in the atmosphere, permitting the gradual vertical transport from the surface into the stratosphere. Intense photochemical activity decomposes the fluorocarbon molecule, releasing chlorine atoms, each of which may destroy many ozone molecules. *See* FLUOROCARBON; STRATOSPHERIC OZONE.

Preparation. The only method used to make ozone commercially is to pass gaseous oxygen or air through a high-voltage, alternating-current electric discharge called a silent electric discharge. First, oxygen atoms are formed as inreaction (1). Some of

$$O_2 \rightarrow 2O \qquad (1)$$

these oxygen atoms then attach themselves to oxygen molecules as in reaction (2). The excess energy

$$O + O_2 + M \rightarrow O_3 + M \qquad (2)$$

in the newly formed ozone is carried off by any available molecule (M) of gas, thus stabilizing the ozone molecule.

The corona discharge principle employed in all types of commercial ozone generators involves applying a high-voltage alternating current between two electrodes which are separated by a layer of dielectric material and a narrow gap through which the oxygen-bearing gas is passed (**Fig. 1**). The dielectric is necessary to stabilize the discharge over the entire electrode area, so that it does not localize as an intense arc.

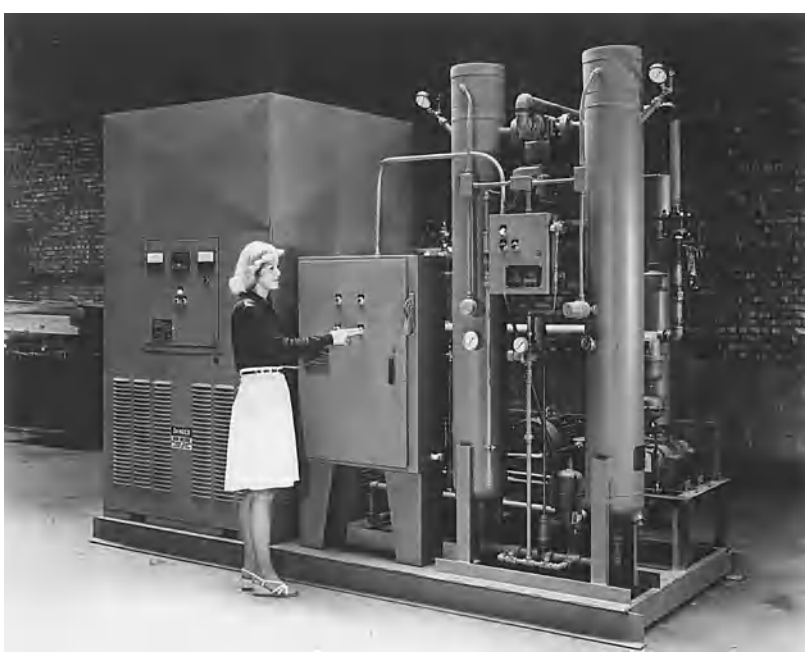

Fig. 3. Packaged ozone generator.

A substantial fraction of the electrical energy is converted to heat. The low volume of gas flowing between the electrodes does not have sufficient capacity to remove this heat. Some external heat sink is necessary, since the decomposition of ozone is accelerated by increasing temperature.

The Lowther cell (**Fig. 2***a*) is an example of a modern, plate-type, air-cooled ozone generator. An individual cell is a gastight sandwich consisting of an aluminimum heat dissipator, a steel electrode coated with a highly stable ceramic dielectric, a spacer to set the width of the discharge gap, a second ceramic-coated steel electrode with an oxygen inlet, and an ozone outlet passing through a second aluminum heat dissipator. Individual cells are stacked into 30-cell modules (Fig. 2*b*), which are grouped with power supplies and controls into packaged ozonators (**Fig. 3**). In the concentric-tube type the oxygen or air to be ozonized passes through the annular space (about 2–3 mm across) between two tubes, one of which must be made of a dielectric material, usually glass, and the other may be either glass or a metal which does not catalyze ozone decomposition, such as aluminum or stainless steel. The internal surface of the inner tube and the external surface of the outer tube, when made of glass, are in contact with an electrical conductor such as metal foil, an electrically conducting paint, or electrically conducting water; these conductors act as electrodes. Between 5000 and 50,000 V at a frequency between 50 and 10,000 Hz is then applied across the electrodes. In some commercial ozone generators the inner and outer tubes are both water-cooled. The latter represents, a simpler type of construction, but does not permit as high an input of electrical power as when both tubes are cooled.

The concentration of ozone in the gas stream leaving commercial ozone generators is usually 1–10% by weight. The yield of ozone is better when oxygen is used instead of air. Other factors which increase the yield of ozone in the silent electric discharge are thorough drying of the oxygen or air before it enters the ozonizer, refrigeration, increasing the pressure to a little above atmospheric, and increasing the frequency of the discharge.

A practical method has been developed for distribution of ozone in small quantities convenient for laboratory use. The ozone is dissolved in a liquefied fluorocarbon. The mixture is maintained at low temperature by a jacket filled with dry ice. Under these conditions, ozone decomposition proceeds very slowly, allowing sufficient time for transport to the user and a modest storage time. Rather high concentrations of ozone may be introduced safely into the laboratory in this manner.

Analytical methods. The analytical determination of ozone is usually carried out in the laboratory by bubbling the gas through a neutral solution of potassium iodide, acidifying the solution, and titrating the iodine thus liberated with standard sodium thiosulfate solution. Ozone in a gas stream may be determined automatically and continuously by passing the gas through a cell with transparent windows and measuring the absorption of either visible light or of ultraviolet radiation beamed through the cell. *See* OXIDATION-REDUCTION; OXYGEN.

Arthur W. Francis

Bibliography. S. B. Majumdar and O. J. Sproul, Technical and economic aspects of water and wastewater ozonization, *Water Res.*, 8:253–260, May 1974; J. B. Murphy and J. R. Orr (eds.), *Ozone Chemical Technology*, 1975; National Academy of Science, *Protection Against Depletion of Stratospheric Ozone by Chlorofluorocarbons*, 1979; National Academy of Science, *Stratospheric Ozone Depletion by Halocarbons*, 1979.

Ozonolysis

A process which uses ozone to cleave unsaturated organic bonds. General olefin cleavage was first extensively studied by C. Harries, beginning in 1903, as a technique to determine the structure of unsaturated compounds by identification of the cleavage products.

Generally, ozonolysis is conducted by bubbling ozone-rich oxygen or air into a solution of the reactant. The reaction is fast at moderate temperatures. Intermediates are usually not isolated but are subjected to further oxidizing conditions to produce acids or to reducing conditions to form alcohols or aldehydes. An unsymmetrical olefin is capable of yielding two different products whose structures are related to the groups substituted on the olefin and the position of the double bond.

The presently accepted mechanism of ozonolysis involves the initial formation of an unstable 1,2,3-trioxacyclopentane "primary ozonide" (**1**) by a 1,3-dipolar cycloaddition of ozone, as shown in one of its resonance structures, with an olefin, in reaction (1). Intermediate (**1**) readily decomposes to a zwit-

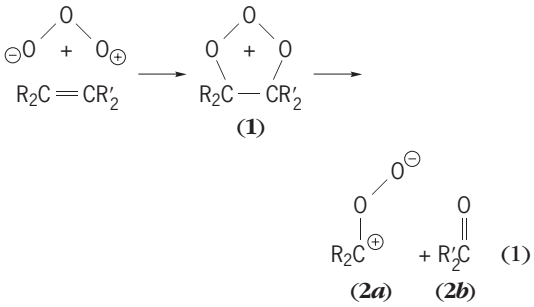

(1)

terion key intermediate, carbonyl oxide (***2a***), and a carbonyl (***2b***). An important reaction of intermediate (***2a***) is with ketones, for example (***2b***), to form a 1,2,4-trioxacyclopentane (**3**), called ozonide, as in reaction (2). The intermediate (***2a***) can also dimer-

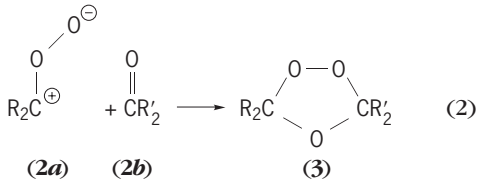

(2)

ize to the diperoxide (**4**) or polymerize to polymeric ozonides and peroxides. The zwitterion produces

oxyperoxides of the general structure (**5**), where the

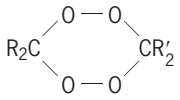

(**4**)

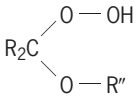

(**5**)

reaction media contains water (R″═H), ethanol (R″ ═OCH_2CH_3), or acetic acid (R″═$OOCCH_3$).

Before World War I, ozonolysis was applied commercially to the preparation of vanillin from isoeugenol. Today the only major application of the technique in the United States is in the manufacture of azelaic and pelargonic acids from oleic acid. *See* ALKENE; OZONE.

Robert K. Barnes

Bibliography. P. S. Bailey, *Ozonation in Organic Chemistry,* vol. 1, 1978; M. Horvath, L. Bilitzk, and J. Huttner, *Ozone*, 1985.

P

Pacific islands

A geographic designation that includes thousands of mainly small coral and volcanic islands scattered across the Pacific Ocean from Palau in the west to Easter Island in the east (see **illus.**). Island archipelagos off the coast of the Asian mainland, such as Japan, Philippines, and Indonesia, are not included even though they are located within the Pacific Basin. The large island constituting the mainland of Papua New Guinea and Irian Jaya is also excluded, along with the continent of Australia and the islands that make up Aotearoa or New Zealand. The latter, together with the Asian Pacific archipelagos, contain much larger landmasses, with a greater diversity of resources and ecosystems, than the oceanic islands, commonly labelled Melanesia, Micronesia, and Polynesia (see illus.). *See* AUSTRALIA; NEW ZEALAND; OCEANIC ISLANDS.

The great majority of these islands are between 4 and 4000 mi^2 (10 and 10,000 km^2) in land surface area. The three largest islands include the main island of New Caledonia (6220 mi^2 or 16,100 km^2), Viti Levu (4053 mi^2 or 10,497 km^2) in Fiji, and Hawaii (4031 mi^2 or 10,440 km^2) the big island in the Hawaiian chain. When the 80-mi (200-km) Exclusive Economic Zones are included in the calculation of surface area, some Pacific island states have very large territories (see **table**). These land and sea domains, far more than the small, fragmented land areas per se, capture the essence of the island world that has meaning for Pacific peoples. In this world of islands, small surface land areas are just one component of a large sea replete with opportunities for exploration. The inhabitants have produced generations of seafarers. *See* EAST INDIES.

Island types. Oceanic islands are often classified on the basis of the nature of their surface lithologies. A distinction is commonly made between the larger continental islands of the western Pacific, the volcanic basalt island chains and clusters of the eastern Pacific, and the scattered coral limestone atolls and reef islands of the central and northern Pacific. One problem with this classification is that islands in the Pacific cannot be grouped neatly by type on the basis of location. There are coral and volcanic islands on the continential shelf in the western Pacific, and

Pacific island developing countries and territories

Countries and territories	Area: Sea Exclusive Economic Zone, 10 km^3	Area: Land, km^2
American Samoa	2,390	197
Cook Islands	1,830	240
Micronesia[a]	3,050	710
Fiji	1,290	18,272
French Polynesia	5,030	3,265
Guam	220	541
Kiribati[b]	3,550	690
Marshall Islands[c]	206	180
Nauru	320	21
New Caledonia	1,740	19,103
Niue	390	259
Northern Marianas	1,870	471
Palau	616	460
Papua New Guinea	3,120	462,840
Pitcairn	800	5
Solomon Islands	1,340	28,369
Tokelau	290	10
Tonga[d]	680	747
Tuvalu	900	26
Vanuatu	680	11,880
Wallis and Futuna	300	255
Western Samoa	120	2,935
Total	30,372	551,476

[a]The Micronesia National Report describes its Exclusive Economic Zone as over 1 million square kilometers.
[b]The Kiribati National Report reports its land area as 822.8 km^2.
[c]The Marshall Islands National Report describes its Exclusive Economic Zone as over 750,000 km^2.
[d]The Tonga National Report describes its total area as 747 km^2 with an inhabited area of 699 km^2; the latter figure has been used in population density calculations. The EEZ is given as 677,021 km^2.
SOURCE: After R. Thustlethwaite and V. Gregory, *Environment and Development: a Pacific Island Perspective*, Asian Development Bank (Manila), 1992.

Pacific islands. (*After R. V. Cole and G. S. Dorrance, Pacific Economic Bulletin, vol. 8, no. 2, December 1993*)

these two types of islands are found in the both the eastern and the northern Pacific.

It has been suggested that a more useful distinction can be drawn between plate boundary islands and intraplate islands. The former are associated with movements along the boundaries of the great tectonic plates that make up the Earth's surface. Islands of the plate boundary type form along the convergent, divergent, or tranverse plate boundaries, and they characterize most of the larger island groups in the western Pacific. These islands are often volcanically and tectonically active and form part of the Pacific so-called ring of fire, which extends from Antarctica in a sweeping arc through New Zealand, Vanuatu, Bougainville, and the Philippines to Japan.

The intraplate islands comprise the linear groups and clusters of islands that are thought to be associated with volcanism, either at a fixed point or along a linear fissure. Volcanic island chians such as the Hawaii, Marquesas, and Tuamotu groups are classic examples. Others, which have their volcanic origins covered by great thickness of coral, include the atoll territories of Kiribati, Tuvalu, and the Marshall Islands. Another type of intraplate island is isolated Easter Island, possibly a detached piece of a mid-ocean ridge. The various types of small islands in the Pacific are all linked geologically to much larger structures that lie below the surface of the sea. These structures contain the answers to some puzzles about island origins and locations, especially when considered in terms of the plate tectonic theory of crustal evolution. *See* MARINE GEOLOGY; MID-OCEANIC RIDGE; PLATE TECTONICS; SEAMOUNT AND GUYOT; VOLCANO.

Island landscapes. Compared with continents, the geology of most oceanic islands is quite simple. They do not display the same degree of variation in rock types as do continental areas that are comparable in size, mainly because most oceanic islands originate as ocean-floor volcanoes. Their growth from the seabed is relatively unimpeded by structures formed in other ways. Coral limestone formations develops on and around these volcanic ridges and cones to form the fringing reefs and atoll lagoons and islets that are characteristic of Pacific islands.

The landscapes of oceanic islands can be grouped into four broad groups: continental, volcanic, high limestone, and low coral reef. The continental or plate boundary islands tend to have steep rugged interiors, often deeply incised by fast-flowing rivers

that have formed coastal plains subject to flooding in the wet season. The more complex geology of these islands has produced numerous mineral deposits, especially nickel, gold, copper, manganese, and bauxite. One of the world's largest nickel deposits forms the backbone of the economy of New Caledonia. Until 1989, the massive Bougainville copper mine at Panguna provided around 40% of Papua New Guinea's export revenues. Fiji's gold mine at Vatukoula has provided export revenues and employment since the 1930s. Mining is often difficult, given the rugged topography and heavy rainfall characteristic of islands in the western Pacific. *See* ORE AND MINERAL DEPOSITS.

The volcanic islands typically have sharp ridges and ampitheater-headed valleys. These valleys are characteristic of islands in the Hawaiian group, in the Society Islands (for example, Tahiti and Moorea), in the Cook Islands (Rarotonga), in Fiji (Viti Levua and Vanua Levu), and in American Samoa (Pago Pago harbor is a drowned ampitheater-headed valley). The principal variables involved in the formation of these spectacular landscapes are river downcutting, chemical weathering, and mass wasting, especially soil avalanching. *See* AVALANCHE; MASS WASTING; WEATHERING PROCESSES.

A common feature of many volcanic islands is a fringing coral reef. This type of reef forms the classical Pacific land-sea interface and is a major source of protein for island residents. A diverse range of fish and crustaceans are obtained from barrier reefs. The fringing coral can also comprise a raised platform (makatea) around the perimeter of volcanic cores, evidence either of sea-level change or tectonic uplift. *See* REEF.

There are also high limestone islands that have no volcanic core above the surface of the sea. Examples include Nayau and Tuvuca in eastern Fiji (Lau), Tongatapu (Tonga), and the isolated islands of Niue, Nauru, and Makatea (French Polynesia). Such islands commonly have a central plateau that dips slightly inland and a series of terraces cut into the plateau flanks. Some of these islands (such as Niue and Makatea) are raised atolls; others are emerged volcanoes draped with limestone. High limestone islands have also been the location of extensive organic phosphate deposits. The isolated oceanic islands of Nauru, Banaba (Ocean Island), and Makatea have been the resting places of migratory birds for thousands of years. Extraction of the resultant bird droppings has been an important mining activity on limestone islands. *See* PHOSPHATE MINERALS.

The low coral islands and reefs constitute a distinctive type of island environment and landscape. The land surfaces of these islands are literally at or near sea level; rarely do atolls exceed 10 ft (3 m) in elevation at their highest points. These islands owe their origin to coral growth, and the main control over coral growth is ocean-water temperature. Coral atolls are confined to a broad band between 20° north and south of the Equator. Atoll landscapes are distinctive because of the low topography, the absence of surface running water, and the amazing diversity in island forms and lagoon shapes. Fresh water comes from the subsurface Ghyben-Herzberg lens, comprising fresh water that floats in hydrostaic balance on the salt water in the substrata of a coral atoll. This lens is the source of most fresh water on coral atolls, where there are no streams and very few lakes. This source of fresh water is becoming very contaminated on coral islands with dense human populations. *See* ATOLL.

Climate. The climate of most islands in the Pacific is dominated by two main forces: ocean circulation and atmospheric circulation. Oceanic island climates are fundamentally distinct from those of continents and islands close to continents, because of the small size of the island relative to the vastness of the ocean surrounding it. Because of oceanic influences, the climates of most small, tropical Pacific islands are characterized by little variation through the year compared with climates in continental areas.

A principal element in climate of Pacific islands near the Equator is the trade winds, which blow steadily for around 9 months of the year—from the northeast in the Northern Hemisphere and from the southeast south of the Equator. These winds moderate temperatures in the subequatorial zone. During the wet season (November to March in the Southern Hemisphere, May to September in the Northern) there is an increased frequency of tropical cyclones in the western Pacific Ocean, especially to the north of Australia and around Hong Kong and the Philippines. Tropical cyclones cause considerable damage to crops and built environments in the western Pacific during most years. They are much less frequent in the central and eastern Pacific. *See* CYCLONE; STORM; TROPICAL METEOROLOGY.

The pattern of tropical cyclones (and the incidence of drought) in the Pacific is affected considerably by variations in air pressure between the Indonesian and South Pacific regions. These variations are termed the Southern Oscillation. Unusually high air pressure in one region is accompanied by unusually low pressure in the other. A measure of the pressure difference between the two is termed the Southern Oscillation Index. When this index is negative, both atmospheric and ocean temperatures in the central and eastern Pacific are higher than normal and an El Niño condition prevails. The warmer sea surface favors the generation of tropical cyclones much farther east than normal, and weather in the eastern Pacific is much wetter and more variable. El Niño conditions also favor drought in the western Pacific. *See* AIR PRESSURE; DROUGHT; EL NIÑO.

When the Southern Oscillation Index is positive, temperatures in the eastern Pacific are cooler, easterly winds prevail, and there is little likelihood of tropical cyclones. This condition is known locally as La Niña. Climate variability under La Niña conditions is much less than that which prevails under El Niño conditions. A major concern for the future is the prospect of pressure difference between the western and eastern Pacific increasing as a result of global warming. Increased variability of climate, associated with this warming, will have profound

impacts on island ecosystems. *See* CLIMATE MODIFICATION.

On the larger high islands, especially those in the western Pacific, local topography can have an important impact on rainfall. Orographic effects are substantial in Fiji, New Caledonia, and Vanuatu and on the larger Solomon Islands. There are distinctive dry and wet zones, and these are reflected in densely forested landscapes on the windward sides and dry savanna landscapes on the leeward sides of islands, with rugged interiors. *See* HYDROMETEOROLOGY; SAVANNA.

Natural hazards. The major natural hazards in the Pacific are associated either with seasonal climatic variability (especially cyclones and droughts) or with volcanic and tectonic activity. The destruction of Rabaul on the island of New Britain in Papua New Guinea in September 1994 was a vivid reminder of the potency of volcanic activity in the western Pacific.

Global climate change has particular significance for islands in the Pacific. One thing that seems clear is the increasing variability in the distribution and incidence of extreme climatic events, especially cyclones and droughts. These, coupled with the degradation of natural ecosystems accompanying development and a scenario of longer-term sea-level rise, make island environments more vulnerable to natural hazards. *See* CLIMATE HISTORY.

Biodiversity. Oceanic islands have long been recognized as having unusual endemic or native biota, largely because of their remoteness from continental landmasses. The small size of many islands has resulted in less species diversity than is common on larger landmasses.

Notwithstanding the smaller range of species and a high level of speciation on oceanic islands in the Pacific, there is considerable interisland variation in biodiversity. The larger islands of western Melanesia show a greater ecosystem diversity than the smaller low-lying atolls of the eastern Pacific, which have perhaps the lowest ecosystem diversity in the world. There also seems to be an attenuation in marine species from west to east, although everywhere marine invertebrates and fish abound on or near reefs and in associated lagoons. Sea life is the chief source of protein for many island peoples. *See* ECOSYSTEM; MARINE FISHERIES; SPECIATION.

A prolific and highly varied bird life is another characteristic of many Pacific islands, especially those used as resting and nesting places for migratory species at different times of the year. Birds and fish feature significantly in island folklore and legend, much more than terrestrial fauna. There are few reptiles, although sea snakes are not uncommon, especially in the atoll territories.

Human impact. There is a complex history of human occupation of the Pacific islands, and settlement of parts of the western Pacific has been dated back to at least 30,000 years ago. Human impact on the island ecosystems has been extensive, although some scientists believe that changes by humans associated with the fashioning of landscape are superficial when compared to the changes in the landscape that occurred earlier in the Quaternary (prehuman). *See* QUATERNARY.

Humans have introduced a number of animals and plants to the islands. The pig and the rat have had a profound impact on indigenous flora and fauna. Plants such as sweet potato and tapioca (cassava or manioc) became staples in the subsistence diet. Some indigenous plants, especially the coconut palm, became the foundation of the monetary economy. On many of the smaller islands, a coconut overlay effectively destroyed quite diverse ecosystems. Cash crops such as coffee, cocoa, and livestock (especially cattle) contributed to significant changes in both island biota and the physical landscape. *See* COCONUT; STARCH; POTATO, SWEET. Richard D. Bedford

Bibliography. H. C. Brookfield, Global change and the Pacific: Problems for the coming half-century, *Contemp. Pacific*, 1(1):1–18, 1989; H. C. Brookfield and D. Hart, *Melanesia: A Geographical Interpretation of an Island World*, 1971; P. D. Nunn, *Oceanic Islands*, 1994; E. Waddell and P. D. Nunn (eds.), *The Margin Fades: Geographical Itineraries in a World of Islands*, Institute of Pacific Studies, University of the South Pacific, 1993; E. Waddell, V. Naidu, and E. Hau'ofa (eds.), *A New Oceania: Rediscovering Our Sea of Islands*, School of Social and Economic Development, University of the South Pacific, 1993; H. J. Wiens, *Atoll Environment and Ecology*, 1962.

Pacific Ocean

The Pacific Ocean has an area of 6.3×10^7 mi^2 (1.65×10^8 km^2) and a mean depth of 14,049 ft (4282 m). It covers 32% of the Earth's surface and 46% of the surface of all oceans and seas, and its area is greater than that of all land areas combined. Its mean depth is the greatest of the three oceans, and its volume is 53% of the total of all oceans. Its greatest depths in the Marianas and Japan trenches are the world's deepest, more than 6 mi (10 km; **Fig. 1**).

Surface currents. The two major wind systems driving the waters of the ocean are the westerlies which lie about 40–50° latitude in both hemispheres (the "roaring forties") and the trade winds from the east which dominate in the region between 20°N and 20°S. These give momentum directly to the west wind drift (flow to the east) in high latitudes and to the equatorial currents which flow to the west. At the continents, water flows from one system to the other, and huge circulatory systems result (**Fig. 2**).

The swiftest flow (greater than 2 knots or 1 m/s) is found in the Kuroshio Current near Japan. It forms the northwestern part of a huge clockwise gyre whose north edge lies in the west wind drift centered at about 40°N, whose eastern part is the south-flowing California Current, and whose southern part is the North Equatorial Current.

Part of the west wind drift turns north into the Gulf of Alaska, thence west again and into the Bering Sea, returning southward off Kamchatka and northern Japan, where it is called the Oyashio Current.

Fig. 1. Principal relief features of Pacific Ocean. 1 fathom = 1. 8 m. (*After F. P. Shepard, The Earth Beneath the Sea, Johns Hopkins, 1959*)

H. U. Sverdrup (1942) reported estimates of the flow in the North Pacific in the upper 4900 ft (1500 m). The Kuroshio carries about 2.3×10^9 ft^3/s (6.5×10^7 m^3/s) at its greatest strength off Japan. The west wind drift in mid-ocean carries about 1.2×10^9 ft^3/s (3.5×10^7 m^3/s), the California Current east of 135°W carries about 5.3×10^8 ft^3/s (1.5×10^7 m^3/s), and the North Equatorial Countercurrent about 1.6×10^9 ft^3/s (4.5×10^7 m^3/s).

A gyre corresponding to the Kuroshio–California–North Equatorial current gyre is found in the Southern Hemisphere. Its rotation is counterclockwise, with the highest speeds (about 2 knots) in the Southeast Australia Current at about 30°S. The current turns eastward and flows around New Zealand to South America, where it turns northward. Along this coast it is called the Humboldt or the Peru Current. It turns westward at the Equator and is known as the South Equatorial Current in its westward flow. It is to be remarked that the northwestern edge of this gyre is severely confused by the chain of islands extending southeastward from New Guinea to New Zealand. This island chain partly isolates the Coral and Tasman seas from the rest of the South Pacific, so that the western equatorial edge of the gyre is not so regular in shape or so clearly defined as its northern counterpart. *See* SOUTHEAST ASIAN WATERS.

In the region of the west wind drift in the South Pacific the ocean is open to both the Indian and the Atlantic, although the eastward passage to the Atlantic through Drake Passage is narrower and shallower than the region between Australia and Antarctica. But part of the water flows around Antarctica with the wind behind it, and it receives more momentum than its northern counterpart. Total transport is several times greater.

G. E. R. Deacon (1937) estimated the transport to the east in the South Pacific part of the west wind drift as more than 3.5×10^9 ft^3/s (1.0×10^8 m^3/s) in the upper 10,000 ft (3000 m). Sverdrup estimates only about 1.0×10^9 ft^3/s (3.5×10^7 m^3/s) for the upper 4900 ft (1500 m) in the North Pacific west wind drift. The gyre which corresponds to the Oyashio–Gulf of Alaska gyre of the North Pacific is thus vaster in transport and area, since much of it passes around Antarctica. *See* ANTARCTIC OCEAN; INDIAN OCEAN.

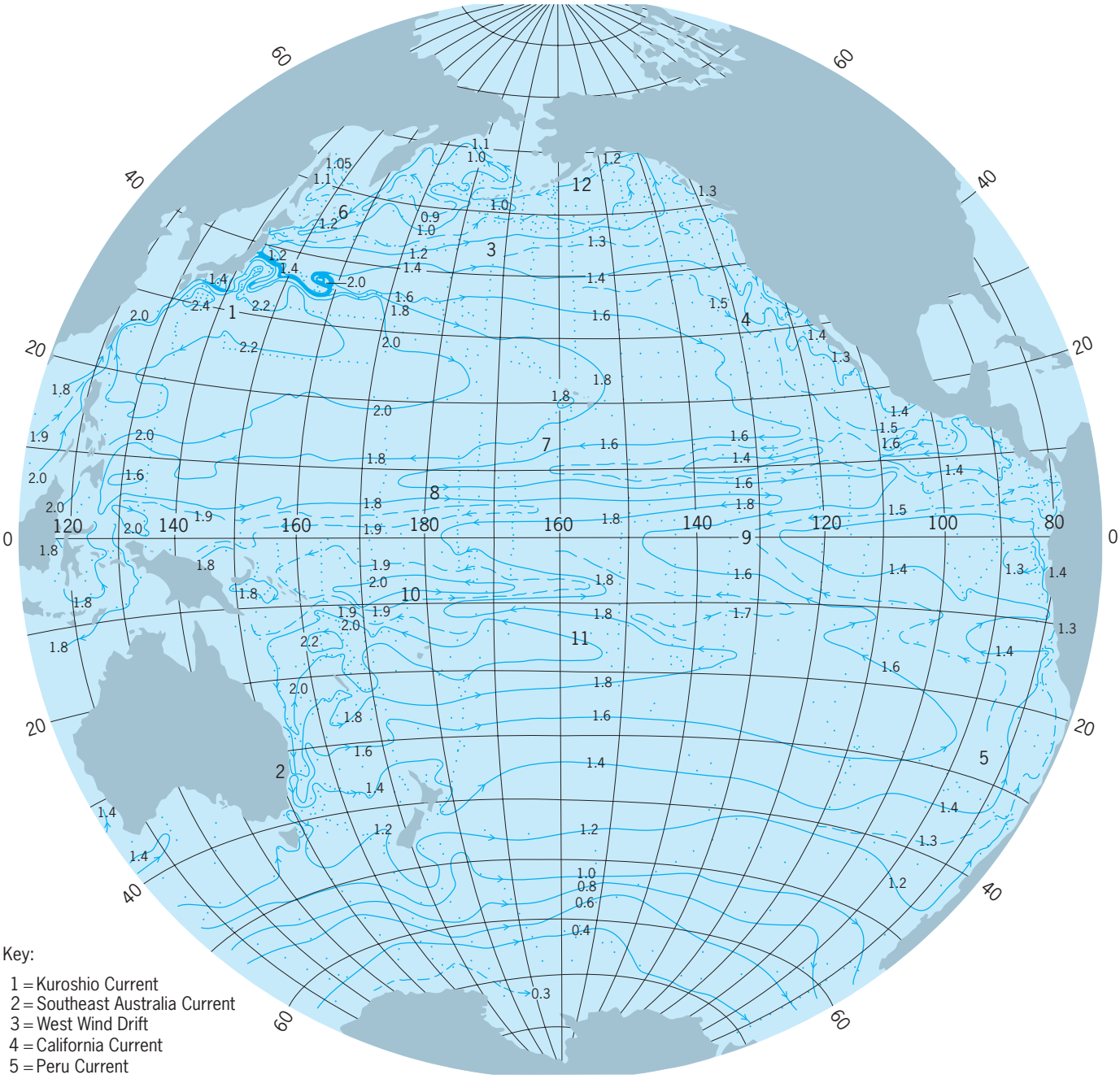

Fig. 2. Principal currents of the Pacific Ocean. Contours give geopotential anomaly in meters at the sea surface with respect to the 1000-decibar (10-MPa) surface, from whose gradient the relative geostrophic flow can be calculated. The large-size numbers refer to the currents which are listed in the key. 1 m = 3.3 ft. (*After J. L. Reid, Jr., On the circulation, phosphate-phosphorus content and zooplankton volumes in the upper part of the Pacific Ocean, Limnol. Oceanogr., 7:287–306, 1962*)

Between the two subtropical anticyclones (Kuroshio–California–North Equatorial Current and the Southeast Australia–Peru–South Equatorial Current) lies an east-flowing current between about 5 and 10°N, called the Equatorial Countercurrent. Sverdrup estimates flow of 8.9×10^8 ft^3/s (2.5×10^7 m^3/s) for this current. There is also an eastward flow along 10°S from 155°E to at least 140°W.

Within the upper 3300 ft (1000 m) the flow of water is generally parallel to the surface flow, but slower. Certain important exceptions occur. Beneath the surface waters of the California and Peru currents, at depths of 660 ft (200 m) and below, countercurrents are found in which some part of the tropical waters is carried poleward. In the California Current this flow reaches to the surface in December,

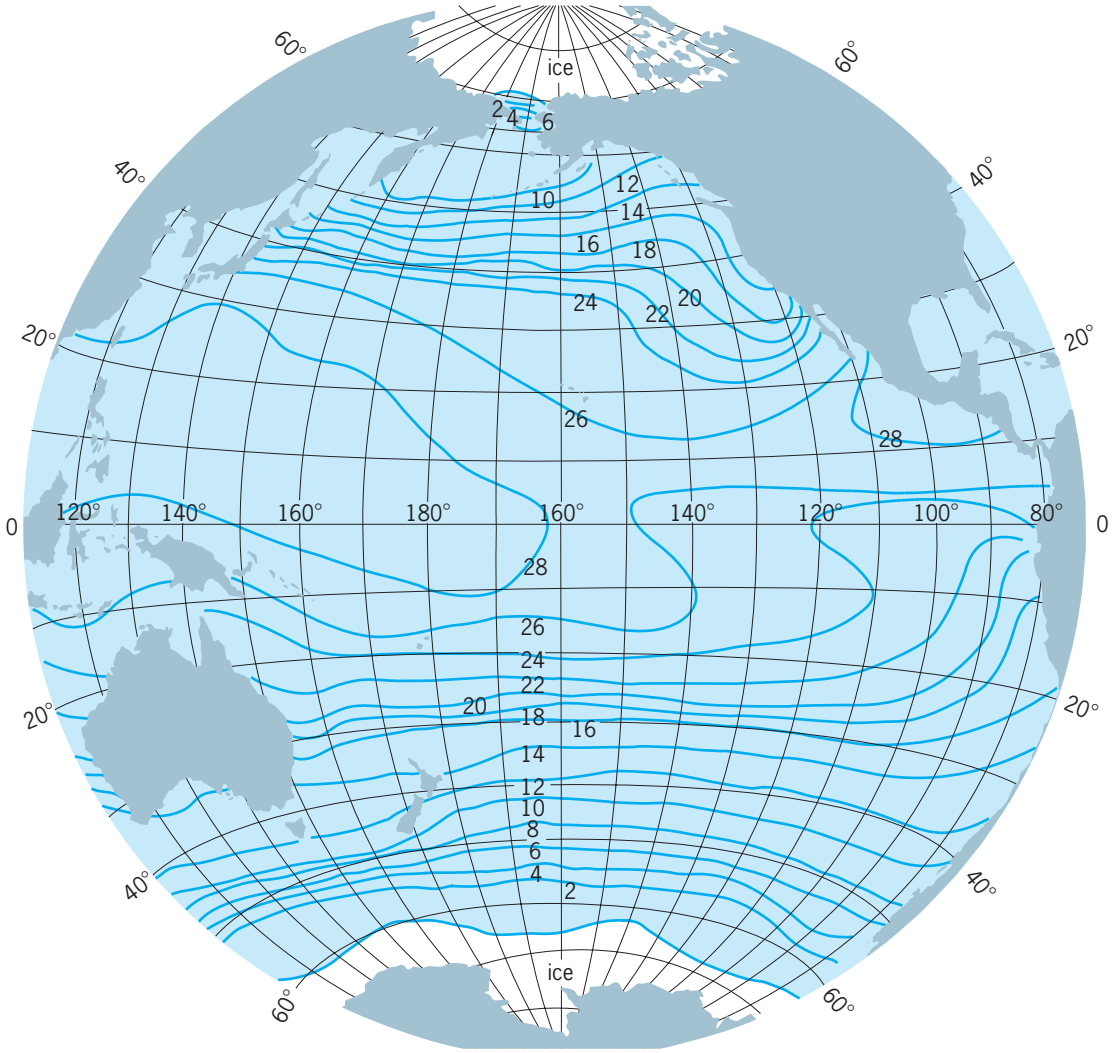

Fig. 3. Temperature at sea surface in August in °C. °F = (°C × 1.8) + 32. (*After U.S. Navy Hydrographic Office, H.O. Publ. 225, 1948*)

January, and February, and it is known as the Davidson Current north of 35°N. It is not known whether a surface poleward flow occurs in southern winter in the Peru Current.

Direct measurements at the Equator and 140°W have revealed that a subsurface flow to the east is found at least from 140°W to 90°W, with highest velocities at depths of about 3300 ft (100 m). Speeds as high as 2–3.5 knots (1–1.8 m/s) to the east were found at this level, while the upper waters (South Equatorial Current) were flowing west at 0.5–1.5 knots (0.25–0.75 m/s). This current has been called the Equatorial Undercurrent, or Cromwell Current after its discoverer. *See* OCEAN CIRCULATION.

Temperature at the sea surface. Equatorward of 30° latitude, heat received from the Sun exceeds that lost by reflection and back radiation, and surface waters flowing into these latitudes from higher latitudes (California and Peru currents) increase in temperature as they flow equatorward and turn west with the Equatorial Current System. They carry heat poleward in the Kuroshio and Southeast Australia currents and transfer part of it to the high-latitude cyclones (Oyashio–Gulf of Alaska Gyral and Antarctic Circumpolar Current) along the west wind drift. The temperature of the equatorward currents along the eastern boundaries of the subtropical anticyclones is thus much lower than that of the currents of their western boundaries at the same latitudes. Heat is accumulated, and the highest temperatures (more than 28°C or 82°F) are found at the western end of the equatorial region (**Fig. 3**). Along the Equator itself somewhat lower temperatures are found. The cold Peru Current contributes to its eastern end, and there is apparent upwelling of deeper, colder water at the Equator, especially at its eastern end [with temperatures running as low as 19°C (66°F) in February at 90°W] as a consequence of the divergence in the wind field.

Upwelling also occurs at the edge of the eastern boundary currents of the subtropical anticyclones. When the winds blow strongly equatorward (in summer) the surface waters are driven offshore, and the deeper colder waters rise to the surface and further reduce the low temperatures of these equatorward-flowing currents. The effect of these seasonal

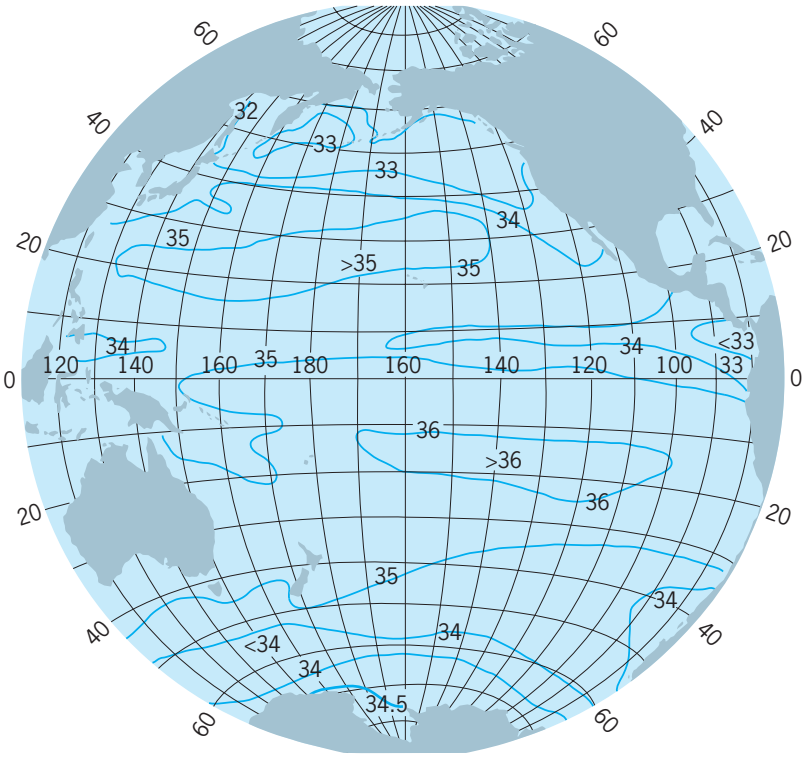

Fig. 4. Salinity at sea surface in northern summer, in parts per thousand by weight.

variations in the winds is thus to reduce the seasonal range of temperature, since the upwelling occurs in spring and summer. The seasonal range of nearshore surface temperature of the California Current at 35°N is less than 4°C or about 7°F (9–13°C or 48–55°F), though the latitudinal mean is about 10°C (50°F). The equatorward winds off Japan occur in winter and the poleward in summer, so that seasonal range is increased at that latitude to more than 16°C or about 29°F (10–26°C or 50–79°F).

The temperatures at the surface of the South Pacific Ocean have not been nearly so well documented, since most of the information comes from measurements made by merchant vessels, and the commercial shipping lines cover but a small part of its great extent. It may be reasoned that most of the temperature characteristics of the North Pacific will have analogies in the Southern Hemisphere. The damped seasonal variation of the California Current seems to occur in the Peru Current as well. But the temperatures of the Southeast Australia Current do not seem to vary so widely through the year as those of the Kuroshio, probably because of the restrictions upon the flow which are imposed by the islands.

The limiting temperature in high latitudes is that of freezing. Ice is formed at the surface at temperatures slightly less than −1°C (30°F) depending upon the salinity; further loss of heat is retarded by its insulating effect. The ice field covers the northern and eastern parts of the Bering Sea in winter, and most of the Sea in Okhotsk, including that part adjacent to Hokkaido (the northern island of Japan). Summer temperatures, however, reach as high as 6°C (43°F) in the northern Bering Sea and as high as 10°C (50°F) in the northern part of the Sea of Okhotsk.

Pack ice reaches to about 62°S from Antarctica in October and to about 70°S in March, with icebergs reaching as far as 50°S. *See* BERING SEA; ICEBERG; SEA ICE.

Salinity at the sea surface. The highest values of salinity observed in the Pacific Ocean are slightly greater than 35.5 and 36.5 parts per thousand (‰), found respectively in the surface water of the centers of the north and south subtropical anticyclones (**Fig. 4**). These anticyclones cover the latitudes in which evaporation exceeds precipitation, and the overlying anticyclonic winds oppose outward flow at the sea surface. Three regions where precipitation most strongly exceeds evaporation are found poleward of 40° latitude and in the eastern tropical Pacific. The result is that the high-latitude cyclones are regions of low salinity (as low as 32.5‰ in the north, 33.8 in the south) which, through mixing with the anticyclones in the region of the west wind drift, contribute water of low salinity to the eastern boundary currents (off California and South America). The greater part of the effect of the eastern tropical precipitation is found at the surface of the North Equatorial Current and Countercurrent. Near Central America, values are less than 33‰ at 10°N, but they rise nearly to 34.5 near the Philippine Islands.

Dissolved oxygen at the sea surface. Above the thermocline the water is in continual overturn and is thus in constant contact with the atmosphere. Oxygen from the atmosphere dissolves in the water until equilibrium is established, and over most of the Pacific the upper layer is very close to saturation in oxygen content with typical values from about 98 to 103% of the saturation value.

The saturated value of dissolved oxygen rises as both the temperature and salinity fall, but the range of surface temperature in the ocean accounts for a wider variation in saturated value than does that of surface salinity, and it is principally variation in space and time of surface temperature which accounts for the oxygen values at the surface. Values greater than 7 ml/liter are found in cold waters of high latitudes and less than 5 ml/liter in warm regions near the Equator.

Nutrients at the sea surface. Nutrients such as inorganic phosphate-phosphorus, silicate-silicon, and nitrate generally increase from the sea surface downward, since photosynthesis and growth in the upper mixed layer tend to use such quantities as are there, and diffusion upward from the higher concentrations is limited by the great stability usually found immediately below the surface layer. At the surface phosphate-phosphorus varies from less than 0.25 microgram-atoms/liter in the centers of the anticyclones to more than 1.5 in the high-latitude cyclones (**Fig. 5**). High values are also found in the California and Peru currents. In a manner similar to the low temperatures found there, their high concentrations are partly the result of transport of mixed water from the cyclones and partly the result of upwelling at the coasts under equatorward winds.

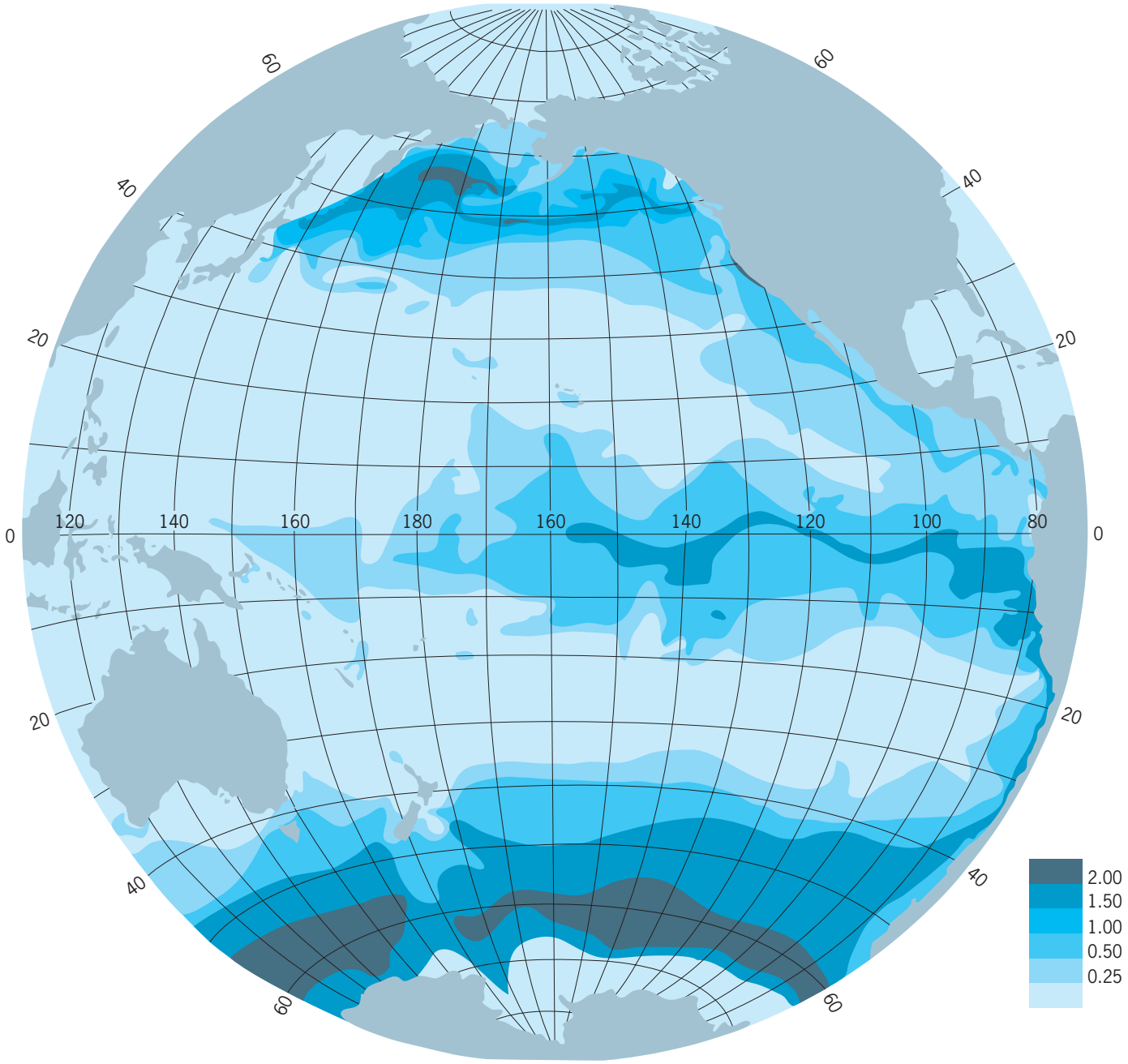

Fig. 5. Inorganic phosphate at sea surface, in microgram-atoms per liter.

Values greater than 1 μg-atom/liter are found in both areas during the summer period of upwelling. At the Equator in the eastern Pacific, upwelling raises values to more than 1 throughout the year.

Silicate-silicon has not been so extensively measured. It also is low in value at the surface and increases with depth. Surface values range from as high as 40 μg-atoms/liter in the high-latitude cyclones and 12 in the upwelling regions of the California Current, to 4 or less in the center of the anticyclones and to values too small to be detected near the Equator in the eastern region. *See* UPWELLING.

Distribution of properties with depth. Surface waters in high latitudes are colder and heavier than those in low latitudes. As a result, some of the high-latitude waters sink below the surface and spread equatorward, mixing mostly with water of their own density as they move, and eventually become the dominant water type in terms of salinity and temperature of that density over vast regions (**Fig. 6**).

The deep and bottom water of all oceans is believed to be formed in the Atlantic high latitudes off Greenland and in the Weddell Sea and arrives in the South Pacific with temperature less than 2°C (36°F), salinity about 34.65–34.7‰, and oxygen about 4.5 ml/liter. The waters filling the Pacific below 10,000 ft (3000 m) retain these values of temperature and salinity everywhere, but in the northern part

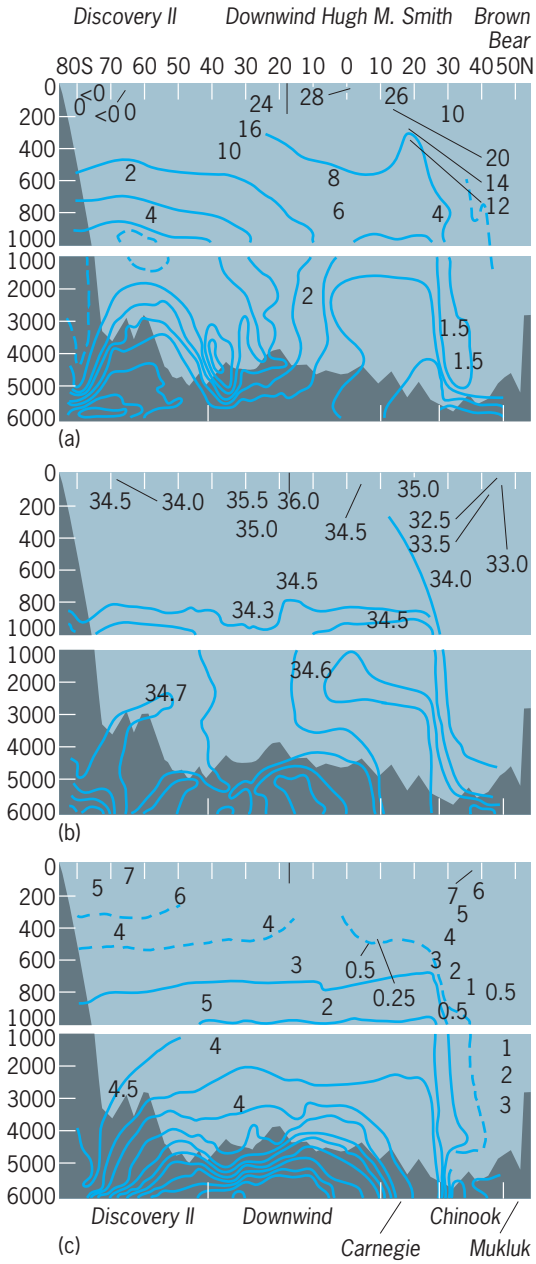

Fig. 6. Vertical sections, central Pacific Ocean, approximately along meridian 160°W, from Alaska (right) to Antarctica (left). (*a*) Temperature in °C. °F = (°C × 1.8) + 32. (*b*) Salinity in parts per thousand. (*c*) Dissolved oxygen in milliliters per liter. Depths are in meters. Depth scale is expanded in upper 1000 m. 1 m = 3.3 ft. (*After Carnegie expedition; NORPAC and EQUAPAC expeditions; Discovery expeditions; and Chinook, Mukluk, and Downwind expeditions*)

the oxygen is reduced to values between 2.5 and 3.5. This would be consistent with depletion by decay and respiration during a slow movement north.

Over most of the North Pacific the temperature does not decrease all the way to the bottom, but beneath a minimum value at about 12,500–13,100 ft (3800–4000 m) it rises again. The increase in temperature is probably in close balance with that in pressure, since no evidence has been found of instability. Two possible explanations for the temperature maximum at the bottom are the flow of heat upward from the ocean floor and an adiabatic rise in temperature as the water flows downward into the deeper basins.

The most conspicuous water masses formed in the Pacific are the Intermediate Waters of the North and of the South Pacific, which on the vertical sections include the two huge tongues of low salinity extending equatorward beneath the surface from about 55°S and from about 45°N. The southern tongue is higher in salinity and density and lies at a greater depth, since the surface waters of the high-latitude cyclone are more saline in the south than in the north.

It has been observed that the higher salinities are found at the surface in the anticyclones. The highest values are in the equatorward halves of the anticyclones. High values penetrate the base of the mixed layer, and tongues of high salinity extend in the thermocline toward the Equator from north and south.

Beneath the mixed layer the waters are not in contact with the atmosphere, and the oxygen that is consumed cannot be replaced directly from the atmosphere. The oxygen quickly falls below the saturated value. Even where high values of oxygen accompany the sinking of water masses, as in the South Pacific Intermediate Water, oxygen is not in saturation below 1000 ft (300 m).

Between saturated waters of the surface and cold bottom waters entering from the south lies a minimum value of oxygen. Beneath the tongue of high oxygen associated with the South Pacific Intermediate Water the minimum is only a little less than 4 ml/liter. Beneath the North Pacific Intermediate Water, however, no water as high as 3.5 is found, and in the minimum itself values less than 0.5 occur over large areas. Values of oxygen beneath the surface are the result of consumption by organisms and replenishment by mixing and renewal of the water. Since at any position the values are nearly constant in time, consumption must everywhere equal replenishment by both flow and diffusion.

Phosphate-phosphorus increases rapidly beneath the sea surface to a maximum which usually lies beneath the oxygen minimum. In regions of upwelling and divergence, higher values of phosphate and other nutrients from depth may from time to time be brought to the surface and thus made available to plants. Such regions (California and Peru currents, South Equatorial Current) are highly productive. The values gradually diminish beneath the maximum to about 2.5 μg-atoms/liter at the bottom in the north and slightly less than 2 in the south. The maximum value is greater than 3.5 in the north and between 2 and 2.5 in the south.

The Pacific Ocean is higher in phosphate concentration than the Atlantic or Indian oceans, exceeding the Atlantic by about 1 μg-atom/liter on the average, though the values at the surface do not differ so much. This excess, like the lower value of oxygen, is undoubtedly related to deep circulation.

Nitrate-nitrogen is present in the ratio of approximately 8:1 by weight of phosphate-phosphorus in those areas where it has been measured, but the details of its distribution are not so well known.

Silicate-silicon has also not been adequately sampled. Its vertical distribution parallels that of phosphate-phosphorus except that beneath the level of the phosphate maximum it remains nearly constant, at about 170 μg-atoms/liter in the center of the northern anticyclone, as high as 220 in the Bering Sea, about 130 in the center of the southern anticyclone. Silicate-silicon, like phosphate-phosphorus, is more highly concentrated in the Pacific than in the Indian and Atlantic oceans. *See* MARINE GEOLOGY; MARINE SEDIMENTS; OCEANOGRAPHY; SEAWATER; SEAWATER FERTILITY. Joseph L. Reid

Bibliography. R. A. Barkley, *Oceanographic Atlas of the Pacific Ocean*, 1969; G. E. R. Deacon, The hydrology of the southern ocean, *Discovery Reports*, vol. 15, pp. 1–124, 1937; J. A. Knauss, Measurements of the Cromwell Current, *Deep-Sea Res.*, 6:265–286, 1960; H. W. Menard, *Marine Geology of the Pacific*, 1964; R. B. Montgomery and E. D. Stroup, *Equatorial Waters and Currents at 150°W in July–August 1952*, Hopkins Oceanographic Studies, vol. 1, 1962; R. A. Muller and T. M. Oberlander, *Physical Geography Today*, 3d ed., 1984; NORPAC Committee, *Oceanic Observations of the Pacific: 1955*, NORPAC Atlas, 1960; G. L. Pickard and W. J. Emery, *Descriptive Physical Oceanography: An Introduction*, 5th ed., 1990; J. L. Reid, Jr., *Intermediate Waters of the Pacific Ocean*, Johns Hopkins Oceanographic Studies, vol. 2, 1965; G. H. Sutton et al. (eds.), *The Geophysics of the Pacific Ocean Basin and Its Margin*, 1976; H. U. Sverdrup, M. W. Johnson, and R. H. Fleming, *The Oceans*, 1942; T. Teramoto (ed.), *Deep Ocean Circulation: Physical and Chemical Aspects*, 1993; W. S. Wooster and G. H. Volkmann, Indications of deep Pacific circulation from the distribution of properties at five kilometers, *J. Geophys. Res.*, 65:1239–1249, 1960.

Packet switching

A software-controlled means of directing digitally encoded information in a communication network from a source to a destination, in which information messages may be divided into smaller entities called packets. Switching and transmission are the two basic functions that effect communication on demand from one point to another in a communication network, an interconnection of nodes by transmission facilities (see **illustration**). Each node functions as a switch in addition to having potentially other nodal functions such as storage or processing.

Switching techniques. Switched (or demand) communication can be classified under two main categories: circuit-switched communication and store-and-forward communication. Store-and-forward communication, in turn, has two principal categories: message-switched communication (message switching) and packet-switched communication (packet switching).

In circuit switching, an end-to-end path of a fixed bandwidth (or speed) is set up for the entire duration of a communication or call. The bandwidth in circuit switching may remain unused if no information is being transmitted during a call. In store-and-forward switching, the message, either as a whole or in parts, transits through the nodes of the network one node at a time. The entire message, or a part of it, is stored at each node and then forwarded to the next.

In message switching, the switched message retains its integrity as a whole message at each node during its passage through the network. For very long messages, this requires large buffers (or storage capacity) at each node. Also, the constraint of receiving the very last bit of the entire message before forwarding its first bit to the next node may result in unacceptable delays. Packet switching breaks a large message into fixed-size, small packets and then switches these packets through the network as if they were individual messages. This approach reduces the need for large nodal buffers and "pipelines" the resources of the network so that a number of nodes can be active at the same time in switching a long message, reducing significantly the transit delay. One important characteristic of packet switching is that network resources are consumed only when data are actually sent.

As a class of store-and-forward switching, packet switching can readily accommodate source and destination end points and transmission links operating at different speeds within the limits imposed by the network. The underlying link-layer protocol (discussed below) can also incorporate error detection and correction capability based on error detection by cyclic redundancy coding (CRC) and correction by retransmission of the packet in error. *See* INFORMATION THEORY.

Packet format. The International Organization for Standardization (ISO) has formulated a conceptual model, known as the open system interconnection (OSI) model, for exchange of information between dissimilar entities. The link-layer and the

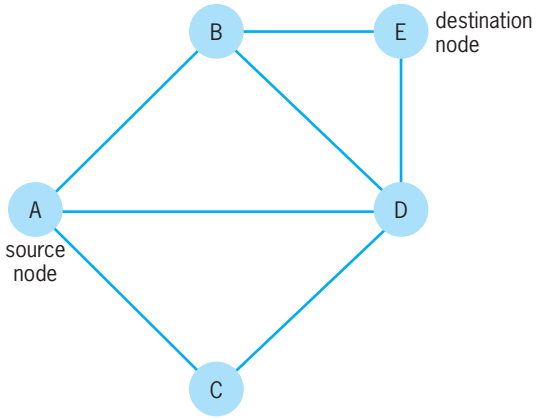

Communication network with five switching nodes interconnected by transmission facilities. Information arriving at the originating or source node can be switched to the destination node through a variety of possible end-to-end paths, for example, ABE, ADE, ABDE, ACDE, ACDBE.

network-layer protocols represent the second and the third layers in this hierarchy. The link-layer protocol is responsible for ensuring the integrity of data across a single link, while the network-layer protocol is responsible for delivering statistically multiplexed data transparently across a network. During their passage through the network, the user data are contained within a packet which is itself contained within (link-layer) frames transmitted from one node to the next.

Packet networks. All public packet networks require that terminals and computers connecting to the network use a standard access protocol. Interconnection of one public packet network to others is carried out by using another standardized protocol.

Services. Packet networks can provide three services: switched virtual circuits, known as virtual calls; permanent virtual circuits; and datagrams. A virtual circuit is a bidirectional, transparent, flow-controlled path between a pair of logical or physical ports. (All bit patterns are transferred across a transparent path or connection without any mutilation.) A switched virtual circuit is a temporary association, while a permanent virtual circuit is a permanent association between two communicating end points of a packet network. Virtual circuits, switched or permanent, are also referred to as providing connection-oriented services. A datagram is a self-contained user data unit containing sufficient information to be routed to the destination without the need for a call to be established. Datagram-based services are referred to as connectionless services.

Satellite and radio networks. Packet-switched networks using satellite or terrestrial radio as the transmission medium are known as packet satellite and packet radio networks, respectively. Such networks are especially suited for covering large areas for mobile stations, or for applications that benefit from the availability of information at several locations simultaneously. The satellite acts as a repeater, amplifying the level of the signal it receives and broadcasting everything back to all the ground stations, including the one that sent the signal. If two or more stations try to transmit to the satellite at the same time, their signals collide, producing useless information. In such a situation, the transmitting stations can easily realize the occurrence of a collision by comparing the signal they received with the one they transmitted, and then can retransmit their information. The other ground stations also detect that the information they received was bad by relying on the check-sum bits that help to detect errors in the packet.

Packet radio networks have lower propagation delays and lower coverage range than satellite-based packet networks. They provide solutions that best fit situations with one or more of the following characteristics: Stations or end points are located in areas with poor telecommunications connectivity among themselves, such as in underdeveloped areas of the world; mobility of the stations or end points is important, such as involving vehicles or a moving data collection point; and stations have relatively infrequent need to communicate.

Asynchronous transfer mode (ATM). This is a type of packet switching that uses short, fixed-size packets (called cells) to transfer information. The term is in contrast to synchronous transfer mode, which is also known as time-division multiplexing, a familiar technique used in digital circuit switching and digital transmission. *See* MULTIPLEXING AND MULTIPLE ACCESS.

The ATM cell is 53 bytes long, containing a 5-byte header for the address of the destination, followed by a fixed 48-byte information field. The rather short packet size of ATM, compared to conventional packet switching, represents a compromise between the needs of data communication and those of voice and video communication, where small delays and low jitter are critical for most applications. Additionally, the ATM protocol takes advantage of high speed and cleaner transmission facilities, such as those provided by optical fibers, by minimizing protocol processing at the intermediate network nodes and relegating functions such as error correction to processing on an end-to-end basis rather than on a node-to-node basis.

Applications. Data communication (or computer communication) has been the primary application for packet networks. Computer communication traffic characteristics are fundamentally different from those of voice traffic. Data communication at speeds from several hundred bits to megabits per second is quite common. Data traffic is usually bursty, lasting from several milliseconds to several minutes or hours. The holding time for data traffic is also widely different from one application to another. These characteristics of data communication make packet switching an ideal choice for most applications. *See* DATA COMMUNICATIONS.

The principal motivation for ATM is to devise a unified transport mechanism for voice, still image, video, and data communication. Multimedia communication services are expected to benefit most from the ATM technology. *See* SWITCHING-SYSTEMS (COMMUNICATIONS).

Internet. The Internet is a global network, the largest packet-switched network in the world. The specific packet-switching protocol used by the Internet is TCP/IP (Transmission Control Protocol/Internet Protocol). Other packet-switching protocols are X.25 (standardized by the International Telecommunications Union) and Frame Relay, a link-layer protocol, commonly used to provide connectivity between enterprise networks (private networks that are designed to serve the communication needs of a particular corporation).

The architecture of the Internet Protocol (IP) is such that it separates applications from the underlying transmission medium. This makes the Internet Protocol suitable for carrying any digital content and, indeed, the Internet is increasingly used to carry any information in digital format. Voice over IP (VoIP) and the transport of visual information are examples of the Internet's growing usage. The precursor to the Internet is the ARPANET (Advanced Research Project Agency Network), funded by the U.S.

Department of Defense, which provided the initial test bed for packet-switching technology in 1969. In addition to the core technology of packet switching, other technological developments which led to the creation of the World Wide Web (WWW) are universal resource identifiers (URI's), the Hypertext Transfer Protocol (HTTP), and Hypertext Markup Language (HTML). *See* INTERNET; VOICE OVER IP; WORLD WIDE WEB. Pramode K. Verma

Bibliography. T. Berners-Lee, *Weaving the Web*, HarperCollins, 1999; Special issue on evolution of Internet technologies, *Proc. IEEE*, vol. 92, no. 9, September 2004; Special issue on packet communication networks, *Proc. IEEE*, vol. 66, no. 11, November 1978; W. Stallings, *Data and Computer Communications*, 7th ed., Prentice Hall, 2003.

Packing

A seal usually used for high pressure as in steam and hydraulic applications. The motion between parts may be infrequent as in valve stems, or continual as in pump or engine piston rods. There is no sharp dividing line between seals and packing; both are dynamic pressure resistors under motion.

Such diverse materials are used for packing as impregnated fiber, rubber, cork, or asbestos compounds. The form of the packing may be square, in ring or spiral form, trapezoidal, or V, U, or O ring in section. In packings, it is necessary that the surface finish of the contacting metal part be smooth for long life of the material. *See* PRESSURE SEAL. Paul H. Black

Pain

An experience of discomfort, primarily associated with tissue damage. Pain is more complex than other sensory systems such as vision or hearing because it not only involves the transfer of sensory information to the nervous system but produces suffering, which then leads to aversive corrective behavior. Unfortunately, sensations other than tissue damage or threatened tissue damage can lead to suffering and yet produce the language and behavioral responses as though tissue damage has occurred.

Pain, especially in its acute form, is usually a reflection of a tissue-damaging or potentially tissue-damaging stimulus. There is a transmission system that conveys this information to the central nervous system. This phenomenon is called nociception. In certain disease states, defects in this transmission system can of themselves generate false information to the nervous system, as though tissue damage were occurring in the periphery. An example of this is phantom limb pain, in which the individual often has a crushing type of pain in a foot that has been amputated. Both of these phenomena lead to a complaint of pain whereby the individual uses the language ("I hurt") and the behavior (rest, inactivity, taking medications, and so on) of pain. Pain causes suffering, but other aversive environmental problems can also cause suffering such as interpersonal relationship problems, loss in its many forms, and such psychological problems as depression, anxiety, or guilt. Many patients with chronic pain complaints use the language and behavior of pain to express this suffering caused by non-tissue-damage problems. Learning is a very potent force, and if this pain behavior is reinforced by environmental contingencies, then the pain complaints can persist purely as a behavioral phenomenon, even though the tissue-damaging or other generators of suffering have ceased.

Acute pain such as occurs with broken bones and other significant injuries is almost inevitably accounted for by the phenomenon of nociception and is probably a purely neurophysiological event. However, the more pain becomes a chronic phenomenon, the more such influences as psychological factors and behavior become part of the expression of pain.

Neurophysiology. Acute pain is a useful warning system. There are specific nerve paths for conducting this sensation (see **illus.**). Pain receptors in the

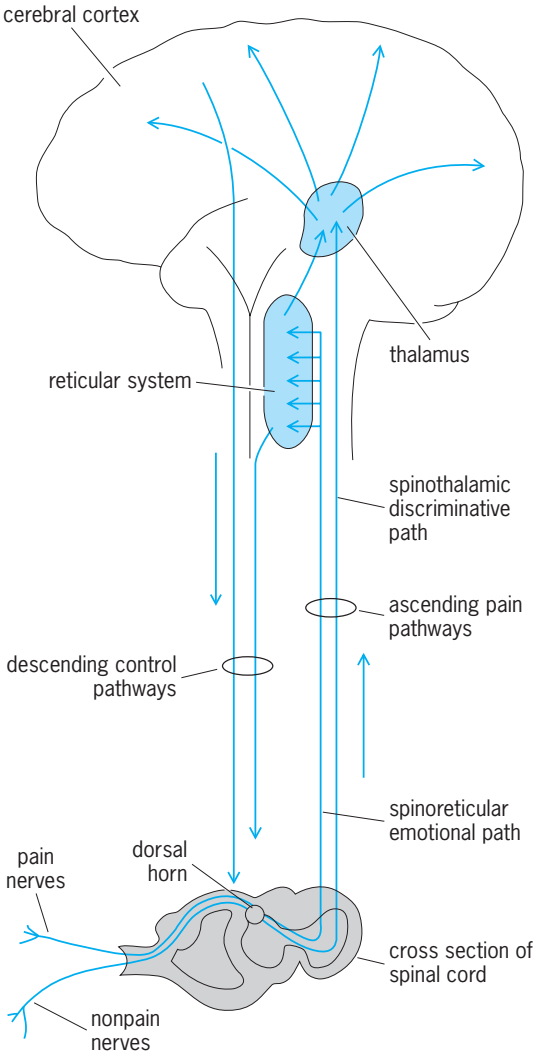

Neurophysiology of incoming pain. Sensation from peripheral receptors travels along specific pain nerves, and is modulated throughout the spinal cord and brain.

skin and other tissues are nerve terminals which lack any special characteristics, and they are probably triggered by a chemical stimulus when potential tissue damage occurs. There appear to be two types of terminals: one responds to many types of painful stimuli, whereas the other specifically responds to either mechanical or thermal energy. When the terminals are stimulated, the pain (that is, nociception message) is carried along specific small sensory fibers called A-delta and C fibers. The A-delta fibers are larger and transmit the "first pain" or "fast pain." The smaller C fibers transmit a secondary dull continuous pain. These nerve fibers were traditionally believed to enter the spinal cord through the dorsal root, but it now seems that some sensory fibers may enter via the ventral route into the spinal cord.

Having entered the spinal cord, these fibers relay in the dorsal horn of the spinal gray matter, an area of considerable regulation and modulation of the incoming pain stimulus which is influenced by other incoming sensory stimuli; that is, touch or pressure sensations can suppress the transmission of signals in the small pain fibers. This helps to explain why when a person is hurting, the pain can be reduced by rubbing the affected part, and this phenomenon forms the basis of some of the treatment strategies of stimulation-produced analgesia. In addition, the incoming pain signal in the spinal cord is also modulated by descending signals from the brain. At times of anxiety, these pain signals may be augmented; however, if the individual is distracted by some interesting task, for example, playing competitive sports, then the pain signal may not be felt at all. From these relay stations in the dorsal horn, the pain signal is carried by two nerve paths up to the brain. The classical pathway is the spinothalamic tract, on the side of the spinal cord opposite to the incoming stimulus, and this leads to the posterior part of the thalamus in the brainstem, and from there nerve paths radiate the pain sensation to many parts of the cerebral cortex, where the pain is appreciated. In addition to this direct path, there is a diffuse ascending path known as the spinoreticular tract which relays to many of the basal ganglia in the brain, and from there to areas of the brain connected with motivational and affective behavior such as the hippocampus and the cingulate gyrus. It is possible that narcotic analgesics exert some of their action on this ascending spinoreticular tract because these drugs tend to reduce the suffering aspects of pain but still preserve many of the discriminative qualities so that individuals can still feel the pain, though it does not bother them so much.

Certain parts of the brainstem around the central canal appear to exert a strong inhibitory effect on incoming pain signals. Stimulation of these areas probably releases endorphins, which are morphinelike substances produced by the body and liberated at various sites on the incoming pain path to suppress these signals. *See* ENDORPHINS.

Treatment. Acute pain often has simple and effective treatment regimens. Chronic pain, however, rarely has simple solutions.

With acute pain, the main management approach is to treat the injured part by combating the infection, immobilizing injured tissues, draining abscesses, splinting broken bones, and so forth, until the natural healing processes resolve the problem. Pain-relieving medications such as anesthetic agents, narcotics, and aspirin can give satisfactory pain relief. *See* ASPIRIN; NARCOTIC.

There have been major improvements in treating acute pain, especially in the postoperative arena. It is now recognized that the patient is probably the best judge of how much analgesic medication is needed postoperatively, and as a result there has been an explosion of interest in and provision of patient-controlled analgesia. The patient is hooked up to an intravenous reservoir of what is usually a powerful narcotic and then pushes a button to deliver a small bolus of pain-relieving drug as needed. These machines, which usually have lock-out safety switches to prevent overdose, have proved most effective, and their use has become widespread.

The novel use of delivering a narcotic such as morphine directly to the spinal cord area has been successful in relieving pain associated with cancer and is becoming common in the treatment of postoperative and posttraumatic acute pain. A small tube catheter placed near the spinal cord can be used to provide analgesia both during surgery and for several days after in a closely monitored postoperative ward. Such systems have provided satisfactory pain relief for as long as 2 or 3 years. However, in most terminal cancer situations the duration of required analgesia is usually much shorter.

Pain that persists for months or even years is much more difficult to control, because such chronic pain is rarely fully explained by the tissue damage model and the neurophysiology of pain explained above. But it is inevitably accompanied by many psychological factors such as depression, anxiety, and personal and other environmental interactions coupled with a learning phenomenon.

Medications, immobilization, and rest, which are very effective for the treatment of acute pain, are usually ineffective and often counterproductive when used for chronic pain.

The potent narcotic analgesics all display the phenomenon of tolerance, and the longer they are used, the less pain-relieving effect they have. The side effects of these drugs, such as constipation and depression, produce secondary, often serious problems, which make them unsuitable for long-term use. The sedative hypnotic drugs such as barbiturates and many tranquilizers have similar disadvantages. If a patient has a fairly finite life span in, say, terminal cancer, then it is often justified to use these potent drugs long-term. But for noncancer chronic pain problems, they are indicated in only a small percentage (5%) of individuals. Aspirin-type drugs do not display tolerance and are useful for chronic noncancer pain problems. *See* ANALGESIC; BARBITURATES; TRANQUILIZER.

For those chronic pains due to a defect in the transmission system whereby nerves have been damaged through either trauma or infection, the antiepileptic

type of drugs are often helpful. They act by reducing the abnormal activity in the defective cells. Surgical section of nerve paths has a useful but limited application and usually in those patients with a finite life span (terminal illness). Sometimes nerves can be blocked by injecting alcohol or phenol into them, especially in those patients debilitated by terminal cancer who are not suitable for surgical operations. These radical forms of pain treatment are rarely (if ever) indicated in noncancer chronic pain because they tend to produce defects in the transmission system which can in themselves produce pain (like phantom pains) at a later stage.

As mentioned above, pain can be relieved by putting in an alternative stimulus such as touch or pressure, and this form of treatment has been used since antiquity with compresses, acupuncture, poultices, and such. There are now small electrical stimulating devices available which can effectively control pain in a certain percentage of patients. It is also feasible to implant electrically stimulating electrodes into the spine or deep areas of the brain such as the thalamus or the central gray matter. However, this approach is still at an experimental stage, and although inserting such systems into spinal areas is becoming increasingly common, the long-term effects are unknown at this time.

For those persons whose pain complaint is an expression of some underlying psychological problem, specific treatments often include the prescribing of antidepressants and antianxiety types of medications rather than directing therapeutic effort at the peripheral pain site. For those whose pain complaint is under control of behavioral environmental factors, it is often necessary to manipulate the environment; most of the pain clinics that are opening up around the world utilize behavioral modification techniques whereby the environmental sequelae of the pain complaints are changed, and individuals can "unlearn" their pain. This usually involves prolonged treatment. The cooperation of the immediate family is vital, and includes major efforts in physical and occupational therapy to improve activity coupled with very strict medication control; as such, these programs are quite effective at restoring these individuals to function and removing drug dependencies and reducing health care utilization. *See* NERVOUS SYSTEM (VERTEBRATE); PSYCHOSOMATIC DISORDERS; SOMESTHESIS. Terence M. Murphy

Bibliography. M. Brody and T. M. Murphy, Pain control and safety monitoring, *Semin. Anest.*, 7:301–306, 1988; R. D. Miller (ed.), *Anesthesia*, 5th ed., 2000; P. D. Wall and R. Melzack, *Textbook of Pain*, 4th ed., 1999; S. W. Weisel, *Pain*, 5 vols., 1989.

Paint and coatings

Substances applied to other materials to change the surface properties, such as color, wear, and chemical or scratch resistance, without changing the bulk properties. The terms paint and coatings often are used interchangeably. However, it has become a common practice to use coatings as the broader term and to restrict paints to the familiar architectural and household coatings and sometimes to maintenance coatings for doors and windows, bridges, and tanks. Another common term for paint and coatings is finish.

Coatings are used for (1) decoration, (2) protection, or (3) some functional purpose. The low-gloss paint used on the ceiling of a room fulfills not only a decorative need but also functions to reflect and diffuse light for even illumination. The exterior coating on an automobile adds beauty, while protecting it from rusting. The public commonly thinks of house paint when talking about coatings; however, all kinds of coatings are important, as they make essential contributions to most high-tech fields.

Coatings may be described by their appearance (for example, clear, pigmented, metallic, or glossy) and their function (such as decorative, corrosion protection, abrasion protection, skid resistance). Coatings may be distinguished as organic or inorganic, although there is overlap. For example, many coatings consist of inorganic pigments dispersed in an organic matrix (binder).

The binder in an inorganic coating is a metal salt. Inorganic coatings also may contain organic additives such as lubricants. Electroplated copper, nickel, and zinc coatings will not be discussed, however; the discussion will be limited to coatings with organic binders that are applied purposefully to a substrate, with organic coatings restricted to materials that can be historically traced back to paints.

Composition. Organic coatings are complex mixtures of chemical substances that can be grouped into four broad categories: (1) binders, (2) volatile components, (3) pigments, and (4) additives. Most coatings contain several substances from each of the four categories, and each substance is usually a chemical mixture. The number of possible combinations is limitless, as are the possible applications.

Binders. These materials form the continuous film that adheres to the substrate, binds together the other substances in the coating, and provides an adequately hard outer surface. The binder governs, largely, the properties of the dried or cured film. The term vehicle usually means the combination of the binder and the volatile components of a coating. Today, most coatings, including waterborne coatings, contain at least some volatile organic solvents. Exceptions are powder coatings and radiation-cured coatings.

Volatile components. Solvents, water, or both are included in a majority of coatings. They play a major role in the process of applying coatings; that is, they make the coating fluid enough for application and they evaporate during and after application.

Pigments. Pigments are finely divided insoluble solids that are dispersed in the vehicle and remain suspended in the binder after film formation. Generally, the purpose of pigments is to provide color and opacity. However, they also have substantial effects on application characteristics and on film properties. While most coatings contain pigments, there are

important types of coatings that contain little or no pigment, commonly called clear coats or clears. Transparent varnishes and clear coats for automobiles are examples. *See* PIGMENT (MATERIAL).

Additives. These are all the materials that are included in small quantities to modify some property of the coating. Almost all coating formulations contain some type of additive. Some additives help to stabilize the coating before use, while others broaden its application properties or improve its durability. Examples are catalysts or initiators for polymerization reactions, stabilizers, and flow modifiers.

Architectural coatings. In general, these coatings are air-drying paints applied by brush, spray, or roller to architectural and structural surfaces for decorative and protective purposes. Materials are classified by formulation type as solvent-based and water-based paints.

Trade sales paints. These coatings are sold over-the-counter and usually applied in the field. Coatings intended for application to industrial structures are not included in this category.

Industrial finishes. Industrial finishes are either those coatings that are applied in the factory or those used for industrial maintenance.

Enamel. Enamel is a type of paint distinguished by its gloss. Enamels differ from flat paints by having a higher percentage of liquid binder, which usually makes them harder, smoother, less porous, more durable, and better able to withstand scrubbing. Commonly, they are either gloss or semigloss. The gloss enamels have a higher percentage of binder and a lower percentage of pigments. Flat enamels contain a flatting agent, a material used to eliminate gloss. Most enamels are solvent-based, but in recent years field-applied latex (water-based) semigloss enamels have been gaining in use.

Lacquer. Lacquer is a clear or colored solution coating that dries by evaporation alone. Drying may take place either at ambient temperature or at elevated temperature by appling heat. *See* LACQUER.

Solvent-based coatings. These form dry films through several mechanisms. Solvent-based coatings may be part of the lacquer or backed enamel families. An air-dry coating is any coating that dries or is chemically cured at ambient temperatures, either in the field or in a controlled factory environment. Shellac is a natural product that is usually dissolved in alcohol and is commonly used as varnish. Coatings based on nitrocellulose, other cellulose derivatives, and acrylic resins are usually called lacquers. *See* POLYACRYLATE RESIN; SHELLAC; SOLVENT; VARNISH.

Solvent-based paints, which dry essentially by evaporation, rely on a fairly hard resin as the vehicle. While the dried coating produces a tack-free and hard film, the coating is sensitive to solvents and may be dissolved or removed by heat or solvent. Such a coating is said to be thermoplastic. Coatings that require higher temperature to accelerate the solvent evaporation also are in this class. Resins for these coatings include shellac, cellulose derivatives, polyurethanes, acrylic, vinyl, and polyester resins. *See* CELLULOSE; POLYESTER RESINS; POLYURETHANE RESINS; POLYVINYL RESINS.

Air-drying finishes once were the conventional factory-applied finishes. These lacquers and varnishes are still used in the finishing of furniture. Until the adoption of baked acrylic and urea finishes, automobile manufacturers used lacquers and air-drying enamels.

Coatings that dry by a combination of air oxidation and solvent evaporation usually are based on unsaturated oils, which are primarily plant-based. An oleoresinous varnish made from a drying oil and phenolic or modified phenolic resin once was commonly used where a hard finish was required for trim or interior surfaces. Such formulations gave way to new resins which cure in shorter times and with improved properties. Modifications of these oils with other chemicals, such as phthalic anhydride and glycerin, produce alkyd resins which can be cured oxidatively at room temperature or at higher temperatures. The oils, such as linseed oil, tung oil, and their alkyd derivatives, usually contain driers to accelerate drying and film formation. The consumption of alkyd- and drying oil-based coatings in United States is declining due to their replacement by environmentally friendly, waterborne coatings. *See* DRIER (PAINT); DRYING OIL; FAT AND OIL; GLYCEROL; OXIDATION-REDUCTION; PHENOLIC RESIN.

Baking finishes. A baking coating is a coating composition, either clear or pigmented, that requires elevated temperature for chemical reactions to occur, that is, curing (crosslinking) the film to a hard, insoluble form. Coatings that require crosslinking are called thermosetting coatings. Solvent-based coatings can also be formulated to form a crosslinked film at room temperature.

Urea and melamine resins polymerize when heated and are used in baking finishes where extreme hardness, chemical resistance, and color retention are required, such as on kitchen and laundry appliances. Acrylic resins crosslinked with melamine–formaldehyde resins and polyisocyanates now dominate the automobile finish market. The reaction of epoxy resins with carboxylic acids and anhydrides requires a minimum temperature of 285°F (140°C) to produce a durable coating. Certain phenolic resins are crosslinked by heat to produce finishes with excellent water and chemical resistance. *See* ACID ANHYDRIDE; FORMALDEHYDE; POLYETHER RESINS; UREA-FORMALDEHYDE RESINS.

Plural component coatings. These coatings begin forming a film by chemical reaction after mixing two or more components. The reactions can proceed at either room or elevated temperature. In theory, neither component reacts by itself; that is, all the components are required to yield a dry film. These coatings generally possess superior durability, adhesion, and corrosion protection properties.

Waterborne coatings. Waterborne coatings are similar to conventional coatings, except that the resin (binder) has been suspended in water through the use of surfactants and the majority of the solvents have been replaced by water. The term waterborne

coating applies to latex, water-reducible (water-thinned), and emulsion paints. Materials formed by emulsion polymerization are described as latex, and the formulated product is called latex paint. The most common latexes used in interior house paint are copolymers of vinyl acetate, butadiene, and acrylic monomers. Exterior house paint usually contains acrylic and styrene polymers. *See* EMULSION; EMULSION POLYMERIZATION; POLYMER; SURFACTANT.

Waterborne coatings can either be air-dried or chemically crosslinked. While latex is said to air-dry, the mechanism of drying is more complicated than evaporation. Ordinary latex coatings used as architectural paints dry by coalescence, forming impermeable films that are no longer soluble in water. Chemical crosslinking, water-reducible, and latex coatings are used as basecoats (color coat) for some automobile finishes. Because they contain lower amounts of volatile organic compounds (VOCs), waterborne coatings are environmentally friendlier than solvent-based paints. Other polymers, such as epoxy, polyurethane, alkyds, and polyesters, can be made into waterborne coatings.

Powder coatings. These coatings are solids. The fine particles of a one-part coating are applied to metals, wood, and some plastics by an electrostatic spray. *See* ELECTROSTATICS.

There are thermoplastic and thermosetting powder coatings. Thermoplastic powder coatings are based on nylon, ethylene, and vinyl polymers. Thermoplastic powders are sprayed onto preheated parts (or dipped into clouds of powder in a fluidized bed) followed by further baking in an oven at 350–450°F (176–232°C). Upon cooling, a tough film much like a conventional coating is obtained. The average film thickness is about 200 micrometers. Typical applications of these powder coatings include office equipment, shopping carts, and refrigerator shelving.

Thermosetting powder coatings include epoxy, polyurethane, polyester resins, and any combination thereof. The fine powder is usually sprayed electrostatically and baked at 275–450°F (135–232°C). The average film thickness is about 75 μm. Since powder coatings are virtually VOC-free, they are more environmentally friendly than solvent-based and waterborne coatings. Unlike liquid coatings, the over-spray is usually reclaimed and reused. Thermosetting powder coatings are applied to a variety of metal, plastic, and wood substrates. Powder coating is used in automotive industry by certain manufacturers as a primer and clear coat.

Thermosetting powder coatings can also be cured by a combination of infrared and ultraviolet radiation. Some heat-sensitive substrates, such as wood and plastics, are ideal for this type of powder coating.

Radiation cured coatings. Coatings for which crosslinking is initiated by radiation, instead of heat, have the potential advantage of being indefinitely stable when stored in the dark. After application, crosslinking occurs rapidly at ambient temperature on exposure to radiation. Most radiation-cured coatings use either acrylic or aliphatic epoxy materials. Generally, the formulations contain no solvent, and emissions are negligible. Rapid curing at ambient temperature is particularly useful for heat-sensitive substrates, such as paper, some plastics, and wood. The two classes of radiation-cured coatings are (1) ultraviolet (UV) where the initial step is the absorption of UV–visible electromagnetic radiation by the photoinitiator to generate a reactive species (ion or radical) which initiates a chemical reaction, and (2) electron beam (EB) in which the initial step is ionization of the coating resin by high-energy electrons. *See* CHARGED PARTICLE BEAMS; FREE RADICAL; PHOTOCHEMISTRY; ULTRAVIOLET RADIATION.

In the case of clear acrylic UV-cured coatings, the curing time required is a fraction of a second. The capital cost for UV curing is low and the energy requirement is minimal. There are limitations and disadvantages. Radiation curing is most applicable to flat sheets or webs for which the distance to the UV source or the window of an EB unit is approximately constant. For UV curing, pigments can limit the thickness of films that can be cured. Radiation cured coatings are essential to the production of computer chips, optical fibers, credit cards, and printed circuit boards, and are used in a variety of other economically important applications.

Electrodeposition. Electrodeposition is a method in which metal objects are virtually plated with resins and color by suspending them in a water bath containing resins, electrolytic stabilizers, and pigments. The objects in the bath become either the anode or cathode, depending on the chemistry of coating, while the anode (cathode) is either the tank or some suitable object suspended in the tank. Resins and pigments for this process were formerly limited to water-soluble epoxies and alkyd-modified phenolic resins and iron oxides, but the selection and color range of the resins have been extended. *See* ELECTROCHEMISTRY.

Corrosion-resistant coatings. In general, these are tough, highly crosslinked coatings that are applied in multiple layers and contain various corrosion-inhibiting pigments and additives. Three basic coating techniques for controlling corrosion are barrier protection (paint), inhibitive primer (coating) protection, and cathodic protection (sacrificial zinc-galvanized steel). *See* CORROSION; INHIBITOR (CHEMISTRY).

When paint is used as a barrier, the substrate is isolated [for example, from salt (conductive electrolyte] by the physical, insulating barrier of the crosslinked film and by the use of flat platy pigments, such as micaceous iron oxide, which provide another barrier. In this case, the adhesion of coating to the substrate is of paramount importance. Coating compositions containing epoxy and phenolic resins generally provide the best corrosion resistance.

Another approach is to use primers that contain special inhibitive pigments that dissolve in water. Unlike the ions of corrosive salt, these materials react with the substrate to form a protective film that further protects steel from corrosion. *See* STEEL.

A somewhat different technique is to alter the surface of the substrate. One way to prevent or reduce corrosion is to use a material that is a sacrificial coating; that is, it dissolves instead of the iron or steel. Galvanized steel has a zinc coating that plays this part. The zinc becomes the anode (a larger one) and slowly dissolves, but the steel (the cathode) does not. A variation of this method exists in which powdered zinc is incorporated into a coating. The sacrificial material used in organic coatings is zinc dust (90–95% of the total weight of the dry coating). Metal contact between the steel and the zinc particles is essential in the early stages of the exposure; but due to the corrosion products, good protection is provided even after the contact has been lost. *See* GALVANIZING; ZINC.

By varying degrees, all organic coatings are permeable to water, salt, and oxygen. To control corrosion, several layers of coatings are used, each having a different function. A base or prime coat is applied to a surface that has been prepared to a specified degree of cleanliness and roughness by abrasive blasting or another method. The prime coat provides adhesion to the substrate for the entire coating system by mechanical anchorage, atomic attractive forces, chemical reaction, or a combination of these. Prime-coat materials once were predominantly a red pigment dispersed in linseed oil and cured by oxidation. Today, the prime coat material commonly specified for application to steel is a dispersion of zinc in a suitable inorganic or organic vehicle. In some circumstances, such as protection of aluminum for aircraft or for substrates such as steel, a one- or two-part epoxy with corrosion-inhibiting pigments is used.

An intermediate coat in a corrosion-resistant system is not always required. When used, intermediate coatings usually are higher solids versions of the same material specified for the topcoat, and are applied only to increase the dry film thickness to increase the barrier properties.

In a corrosion-resistant system, the topcoat may be selected from a variety of formulations, such as a two-part acrylic polyurethane cured with polyisocyanates, chlorinated rubber, epoxy/phenolics, and in some cases epoxy resins cured with polyfunctional amines and polyamides. Topcoat materials are required to have good sealing properties, as well as high resistance to corrosion, erosion, and ultraviolet degradation. *See* METAL COATINGS.

High-temperature coatings. These coatings withstand services temperatures higher than 750°F (400°C). Applications include aircraft engine components, mufflers, heat exchangers, missile technology, chemical reaction vessels, aerospace vehicles, turbines, furnaces, and barbecue grills.

Typical coating formulations, such as alkyds, polyurethanes, and epoxies, have limited service temperatures, well below 500°F (260°C). Compared to other organic polymers, silicone resins have higher thermal stability and greater resistance to UV radiation and oxidation. Coatings based on methyl silicone resins and pigmented with high levels (up to 3 lb/gal or 360g/L) of aluminum are capable of withstanding temperatures greater than 1200°F (650°C). Coatings using some ceramic frits or ferrites give service up to 1500°F (815°C). White and colored silicone-based coatings are suitable for services up to 550–600°F (290–315°C). These decorative coatings are used for incinerators, space heaters, reflectors, clothes driers, and stoves, where requirements for color and gloss retention are combined with the need for high-temperature resistance. *See* ALUMINUM; CERAMICS; CERMET; FERRITE; SILICONE RESINS.

Other high-temperature coatings include zinc dispersed in a suitable vehicle, serviceable to 750°F (400°C); a phosphate bonding system in which ceramic fillers are mixed with an aqueous solution of monoaluminum phosphate, serviceable to 2800°F (1538°C); and a ceramic gold coating used on jet engine shrouds, serviceable to up to 1000°F (538°C). For outer-space coating applications, research continues for higher temperatures and radiation effects.

Smart coatings. A smart coating senses its environment and makes an appropriate response, and is designed so that function and response can be switched on or off. Applications include hygienic coatings and coatings to detect and destroy chemical and biological warfare agents, as well as healable, antifouling, radar-transparent, and conductive coatings. The sensing agent in the coating may be an additive, pigment, or polymer itself. Smart coatings respond to various types of stimuli, such as heat, pressure, pH, impact, vibration, light, viruses, electromagnetic fields, and chemicals.

Jamil Baghdachi

Bibliography. C. H. Hare, *Protective Coatings: Fundamentals of Chemistry and Composition*, Technology Publishing Company, Pittsburgh, 1994; G. P. Turner, *Introduction to Paint Chemistry and Principles of Paint Technology*, 4th ed., 1997; Z. W. Wicks, Jr., F. N. Jones, and P. S. Pappas, *Organic Coatings: Science and Technology*, Wiley-Interscience, New York, 1999.

Palaeacanthocephala

An order of the Acanthocephala, the adults of which are parasitic worms found in fishes, aquatic birds, and mammals. They have the following characteristics. The nuclei of the hypodermis are fragmented and the chief lacunar vessels are lateral. The males have usually two to seven cement glands. The ligament sac in the female breaks down so that the eggs develop in the body cavity. Proboscis hooks occur in long rows and spines are present on the body of some species. Species which commonly occur in vertebrates are *Leptorhynchoides thecatus* and *Corynosoma*.

Leptorhynchoides thecatus. This is one of the most common species of acanthocephalan in North American fresh-water fish. The body is a creamy color, long, slender, and devoid of spines. Both ends of

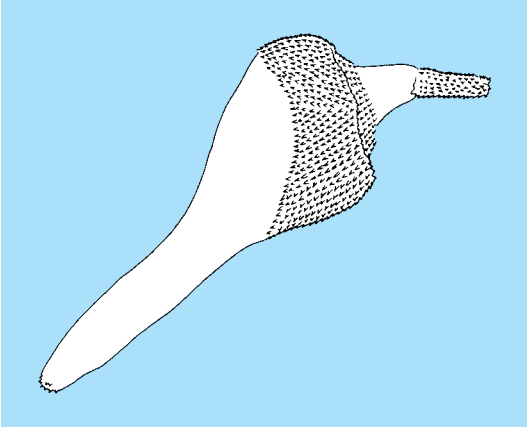

Corynosoma reductum. (*After H. J. Van Cleave, Acanthocephala of North American Mammals, University of Illinois Press, 1953*)

the body are curved ventrally. The females are 0.2–1.0 in. (6–26 mm) and males 0.1–0.5 in. (3–12 mm) in length. Females usually are thicker than the males. The proboscis is long, slightly enlarged in the middle, and bent ventrally at an angle to the body. Proboscis hooks are quincuncial in arrangement with 12 longitudinal rows of 12–16 hooks each. The base of the hook is enveloped by a prominent cuticular sheath. Genital organs of the male occur in the posterior half of the body. Eight cement glands usually are arranged in three levels. The eggs are spindle-shaped with the middle membrane or shell thicker because of polar enlargements. The outer membrane is thick and refractive. The intermediate host is an amphipod, *Hyalella azteca*. Adults have been recorded from the ceca and small intestine of 79 species of fish.

Corynosoma spp. Individuals of the genus *Corynosoma* reach sexual maturity in the small intestine of birds and mammals with aquatic habitats. The body is club-shaped and 0.1–0.4 in. (2–10 mm) long; the anterior end is thickened as an inflated bulb, whereas the hind trunk is narrow and cylindrical (see **illus.**). The body is provided with spines on the anterior extremity of the trunk, which extend farther along the ventral surface than the dorsal. In some species the trunk spines extend the full length of the trunk on the ventral surface. The body spines are of various forms, often sigmoidal, and the tip of each is commonly invested by a cuticular fold. The proboscis is directed ventrally. Proboscis hooks occur in longitudinal rows, and the individual hooks increase in thickness from anterior to posterior. The male organs are restricted to the posterior half of the body with the testes rounded or slightly elongate. Six cement glands, pyriform to clavate in shape, are present. Eggs are spindle-shaped with a short axial prolongation of the middle membrane. The species of this genus are distributed through all continents. Aquatic mammals (seals) and water birds (ducks) are the normal definitive hosts, although fishes serve as the second intermediate or transport host and various crustaceans as the first intermediate host. *See* ACANTHOCEPHALA.

Donald V. Moore

Palaeoisopus

A peculiar arthropod, evidently related to the Pycnogonida and represented by a number of well-preserved fossils from the Devonian Hunsruck shales. It was formerly considered by a number of paleontologists to have an anterior jointed proboscis and a bulbous terminal abdomen, but studies of material under ultraviolet light have demonstrated that the bulbous abdomen is in fact a pair of robust, well-developed chelae (see **illus.**). Furthermore,

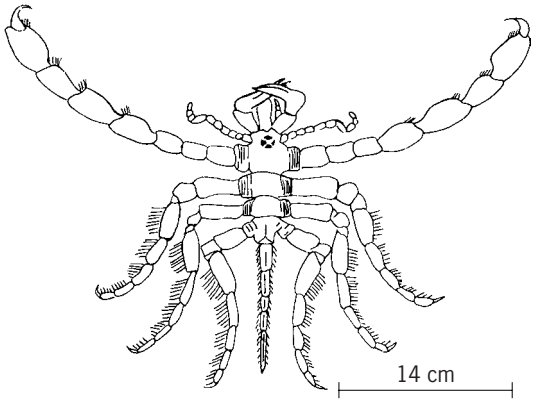

Palaeoisopus problematicus.

additional appendages (palps and ovigers) not apparent in previously studied material have been discerned, bringing its complement of anterior appendages into agreement with that of the Pycnogonida. As reoriented, *Palaeoisopus* has anterior flattened appendages that separate it from extant families of Pycnogonida. *See* PYCNOGONIDA.

Joel W. Hedgpeth

Bibliography. M. W. Lehmann, Neue Entdeckungen an *Palaeoisopus*, *Paleontol. Z.*, 33:96–103, 1959; R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, pt. P, 1955.

Palaeonemertini

A rarer order of the class Anopla in the phylum Rhynchocoela, characterized by an unarmed proboscis, a thin gelatinous dermis, and either a two-layered (outer circular and inner longitudinal strata) or three-layered (outer circular, median longitudinal, and inner circular strata) body musculature. Many members (such as *Tubulanus* = *Carinella*) show primitive features, such as a peripherally located nervous system and the absence of ocelli, ciliated grooves, and intestinal diverticula. Cerebral organs, if present, are generally simple. *See* ANOPLA; ENOPLA; HETERONEMERTINI; RHYNCHOCOELA.

J. B. Jennings

Palaeonisciformes

A large, extinct assemblage of primitive ray-finned (actinopterygian) fishes known from the Silurian to Cretaceous, which were most numerous in

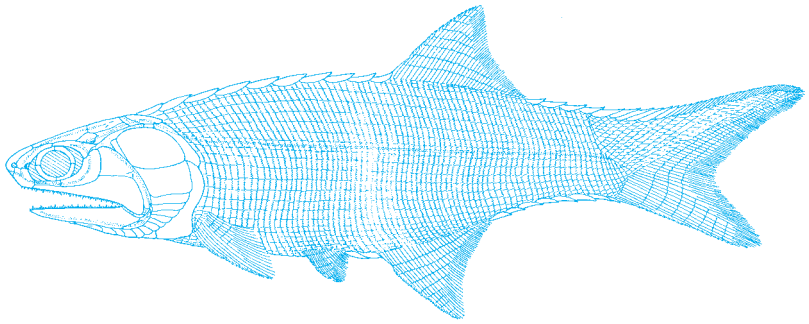

Fig. 1. ***Mimia toombsi*** **(Gardiner & Bartram, 1977), Late Devonian, Gogo, Australia. One of the earliest palaeonisciforms, reaching 15 cm (6 in.) long, and known by complete fossils. (*Reprinted with permission from B. G. Gardiner, The relationships of the palaeoniscid fishes, a review based on new specimens of Mimia and Moythomasia from the Upper Devonian of Western Australia. Bull. Brit. Mus. (Nat. His.), Geology, 37(4):173–428, 1984*)**

Carboniferous and Permian times. They are known as fossils from all continents except Antarctica. Approximately 40 families are recognized but collectively they are not a natural group; some are genealogically more closely related to derived ray-finned fishes (holosteans and teleosts) than to other palaeonisciforms. *See* ACTINOPTERYGII; HOLOSTEI; TELEOSTEI.

Morphology. Primitive palaeonisciforms are mostly small [5–30 cm (2–12 in.) long], with a streamlined body covered with thick, shiny scales and head bones. The scales are rhomboid and covered with a substance called ganoine (in older literature these and a few other fishes are called ganoids). The ganoine consists of a superficial enamellike layer overlying dentine. In turn, the ganoine overlies a thin bone layer—the latter is all that remains in the scales of the vast majority of modern bony fishes. Over most of the body, each scale abuts with its neighbor above and below by means of a peg-and-socket joint. This type of scale, found as isolated remains, is the first evidence of ray-finned fishes in late Silurian rocks, and it gives its name to the group (palaeonisciform literally means "ancient scale").

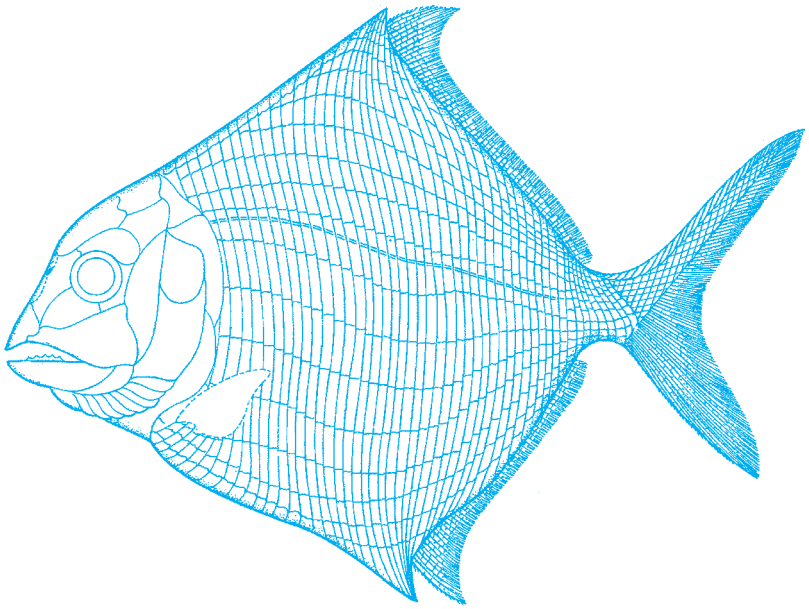

Fig. 2. ***Chirodus granulosus*** **(Young, 1866) was a deep-bodied palaeonisciform that grew to 10 cm (4 in.) and lived in Carboniferous coal swamps of Britain. The teeth formed cutting blades that may have been used to crop vegetation. This reconstruction is taken from Traquair (*Transactions of the Royal Society of Edinburgh, Vol. 29, Plate 5, Fig. 1, 1880*)**

The body tapers to a gently upturned, long, asymmetrical tail, in which all the fin rays are inserted along the ventral edge; there is a single dorsal fin placed midway along the back. The paired fins (pectoral and pelvic) are located close to the ventral midline, and in more primitive palaeonisciforms, the pelvic fin has a broad base of insertion into the body. All fins are supported by finely segmented and branched fin rays, and the leading edges are covered with specialized pointed scales (fulcral scales) that form a cutwater. There are also enlarged scales along the back.

The head tapers to a blunt snout, and the large eyes (placed well forward) suggest that palaeonisciforms relied on sight for capturing prey, such as small crustaceans and other fishes, as well as to avoid predators. The nostrils are tiny and located right at the tip of the snout. The jaws are long and equipped with many tiny pointed teeth that were used to grab and impale rather than cut or grind prey. The tip of each tooth is formed by a glassy cap made of a special type of hard tissue called acrodin; this is characteristic of the teeth of all ray-finned fishes. Thick, heavily ornamented, ganoine-covered bones cover the head and cheek, and most are tightly sutured together to form an inflexible armor. The upper jaw bone (maxilla) is rigidly attached to the other cheek bones so that it is incapable of independent movement. Therefore, the only way to open the mouth is to raise the head at the same time as dropping the lower jaw. The articulation of the lower jaw with the palate is located well behind the level of the eye, such that the jaw suspension from the braincase is markedly oblique. This implies some restriction on the ability of the muscles to close the jaw.

Evolution. Evolutionary trends within palaeonisciforms include thinning of the scales by a reduction of the ganoine layer; development of a more upright jaw suspension and shortening of the jaws that allowed a more efficient jaw closure; a shortening of the base of insertion of the pelvic fin, allowing more flexibility; shortening of the axial lobe of the tail which was then able to generate a more horizontal thrust; and a reduction in the ratio of the dorsal, anal, and caudal fin rays to their respective internal supports, which led to more subtle control over fin movements.

Devonian. Most Devonian palaeonisciforms are small, slender fishes, such as *Mimia*, and lived in shallow seas (**Fig. 1**). *Cheirolepis*, a freshwater representative, is exceptional in that it grew to 70 cm (28 in.) and had tiny studlike scales, very similar to those of acanthodians, but the pattern of skull bones is typically palaeonisciform. *See* ACANTHODII.

Carboniferous. During the Carboniferous, many separate lineages evolved to occupy both marine and freshwater habitats. They were particularly common in deltas and rivers associated with coal forest swamps, and fishes such as *Rhadinichthys* must have swam in large schools. Body forms became more

diverse, and there were several that foreshadowed the range of body forms of modern ray-fins. Some, such as *Chirodus* (**Fig. 2**) and *Platysomus*, are deep-bodied and laterally compressed. In these Palaeonisciforms, the dorsal and anal fins are very elongated and the scales on the flank are greatly enlarged, while the teeth developed as cutting and grinding plates. It is likely that these palaeonisciforms were poor at swimming in open water but highly adept at maneuvering among water vegetation and Palaeozoic coral reefs. The body form, represented by *Tarrasius*, is characterized by an elongated eel-shaped body surrounded by a continuous median fin; the scales are reduced and pelvic fins are absent, as in modern eels.

Late Carboniferous to Late Permian. Typical-shaped palaeonisciforms survived until the late Palaeozoic, during which *Palaeoniscum*—the fish after which the group is named—was common in Late Permian seas. Other palaeonisciforms showed several deviations in body form. In haplolepids, which were common in Late Carboniferous freshwaters, the snout is reduced in deference to the huge eyes. On the body, the scales are thin and the fin rays are unbranched—a very unusual feature in fishes. Most significantly, the jaw suspension is upright and the cheek bones are reduced.

Triassic. Redfieldiids were common inhabitants of Triassic freshwaters. In these fishes, the lobe supporting the tail is shortened, and the dorsal and anal fins are large and placed well back on the body. The snout is prominent and ornamented with tubercles and may have supported a soft lip.

Deep-bodied palaeonisciforms are represented in the Triassic by the bobasatranids. These fishes had small jaws equipped with rounded, crushing teeth. Unlike earlier Carboniferous deep-bodied palaeonisciforms, the pectoral fin of *Bobasatrania* is located on the flank and was probably used more for maneuverability than for generating lift. This type of body form is taken to an extreme in the Late Permian *Dorypterus*, in which the pelvic fin is located well forward and the dorsal fin is produced as a long filamentous leading edge, analogous to that seen in the modern-day John Dory (*Zeus faber*).

Late Triassic to Cretaceous. Most palaeonsicforms had died out by the end of the Triassic; however, one small group, represented by *Coccolepis*, survived into the Cretaceous. In *Coccolepis*, most of the body scales had been reduced to very thin plates covered with thornlike denticles, while the scales on the axis of the tail were elongated. The snout protrudes so that the jaw is decidedly underslung. It is possible that *Coccolepis* is closely related to modern-day acipenseriforms.

Classification. Classification of palaeonisciforms is complicated because most are known only from flattened and fragmentary fossils. Although many separate groups, often called families, are recognized, the relationships of these families to each other and to other fishes are not well understood. Three modern fish groups must have originated from palaeonisciform fishes: the Polypterifomes, or bichirs; the Acipenseriformes, or sturgeons and paddlefishes; as well as fishes classified as holosteans and teleosteans. *See* ACIPENSERIFORMES; HOLOSTEI; OSTEICHTHYES; POLYPTERIFORMES; TELEOSTEI. Peter L. Forey

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, Freeman, New York, 1988; P. Janvier, *Early Vertebrates*, Clarendon Press, Oxford, 1996; J. A. Moy-Thomas and R. S. Miles, *Palaeozoic Fishes*, 2d ed., Chapman & Hall, London, 1971.

Palaeospondylus

A tiny fossil fish found only in rocks of Middle Devonian age within a small geographic region of Caithness, Scotland, and principally at one site, Achanaras Quarry, where it is surprisingly abundant. One of the most enigmatic of all fossil fishes, *Palaeospondylus gunni* was discovered in 1890 by Marcus and John Gunn and described by Ramsay Traquair. Its original environment was a large, deep, freshwater lake system with a broad diversity of fishes. Unusually for a tiny fish, *Palaeospondylus* is typically found in sediments associated with the deepest part of the lake.

Morphology. Specimens of *Palaeospondylus* never exceed 6 cm (2.36 in.) in length. Usually, the head, vertebral column, and a dorsal tail fin are preserved, along with occasional traces of limb girdles (see **illustration**). (*Palaeospondylus* means "ancient vertebrae".) The front of the head includes a strange structure identified as the nasal apparatus or jaws. A characteristic feature of the back of the head is a pair of elongated, backwardly projecting dorsal ribs.

Classification. *Palaeospondylus* has often been taken to be an adult because the head and backbone appear to be well ossified. However, the individual elements have been much transformed in the process of fossilization and often appear fused together. Past attempts to identify *Palaeospondylus* with any known fishes, living or fossil, have always failed, although it has variously been identified as a relative of the lampreys, a placoderm, a lungfish, a shark, a larval amphibian, and even a herring. Debates over its identity became particularly intense in the 1920s and 1930s, as different authors used it to promote rival theories of early vertebrate evolution.

Digital reconstructions of a new series of microscopic thin sections show that *Palaeospondylus* is a larval lungfish, probably the immature stage of

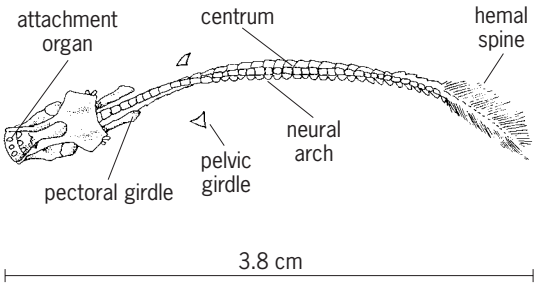

Skeleton of Middle Devonian *Palaeospondylus*, shown dorsally.

the common Caithness lungfish, *Dipterus valenciennesi*. This conclusion is supported by the presence of dorsal ribs, which is a morphological feature unique to living and fossil lungfishes. The structure on the front of the head is a sucker or attachment organ (see illustration) similar to that seen in another type of primitive (but not directly related) fish, the living garfish *Lepisosteus*. The geological environment suggests that the larvae lived in marginal lake environments, where they attached onto rocks or plants and fed on detritus before undergoing metamorphosis into the adult stage. Periodically, swarms of larvae were swept into the middle of the lake where they perished. *Palaeospondylus* is currently the oldest known vertebrate larval form. Keith Thomson

Bibliography. J. A. Moy-Thomas and R. S. Miles, *Palaeozoic Fishes*, Saunders, 1971; K. S. Thomson, A palaeontological puzzle solved? *Ameri. Sci.*, 92: 209–211, 2004; K. S. Thomson, The puzzle of *Palaeospondylus*, *Ameri. Sci.*, 80:216–219, 1992.

Palate

The roof of the mouth in those vertebrates whose mouth cavity and nasal passages are wholly or partially separate. A communication between each nasal sac and the mouth cavity first appears in the specialized lungfishes. The two passages open into the front part of the mouth by apertures called the internal nares, or primitive choanae, and the nose then becomes respiratory as well as olfactory in function. A similar adaptation to air breathing when the mouth is closed occurs in amphibians where the premaxillary, median region of the upper jaw serves to separate the nasal cavities from the mouth in front; elsewhere the mouth is an unpartitioned, common chamber. A distinct advance toward a complete palate distinguishes reptiles and birds, with both groups possessing a pair of bony flanges separated by a median cleft so that the incomplete palate provides something of a conduit for the free passage of air. This inverted trough extends from the primitive choanae back to the pharynx. Crocodiles and mammals create wholly separate channels by interposing a complete palate between the air passages and the mouth cavity. The definitive communication of the air channels with the nasal pharynx is known as the secondary choanae. The premaxillary palate of amphibians and the paired unfused palatal shelves of reptiles and birds correspond to progressive stages occurring during the development of the definitive palate in all mammalian embryos.

Teeth usually occur on the roof of the mouth of bony fishes, amphibians, and reptiles. Many mammals, and especially hoofed and carnivorous forms, have the palate set with transverse ridges of cornified epithelium which aid in the manipulation of food. The toothless whales elaborate these structures into sheets of fringed "whalebone" that serve to strain out minute organisms to be eaten. *See* NOSE; TOOTH.

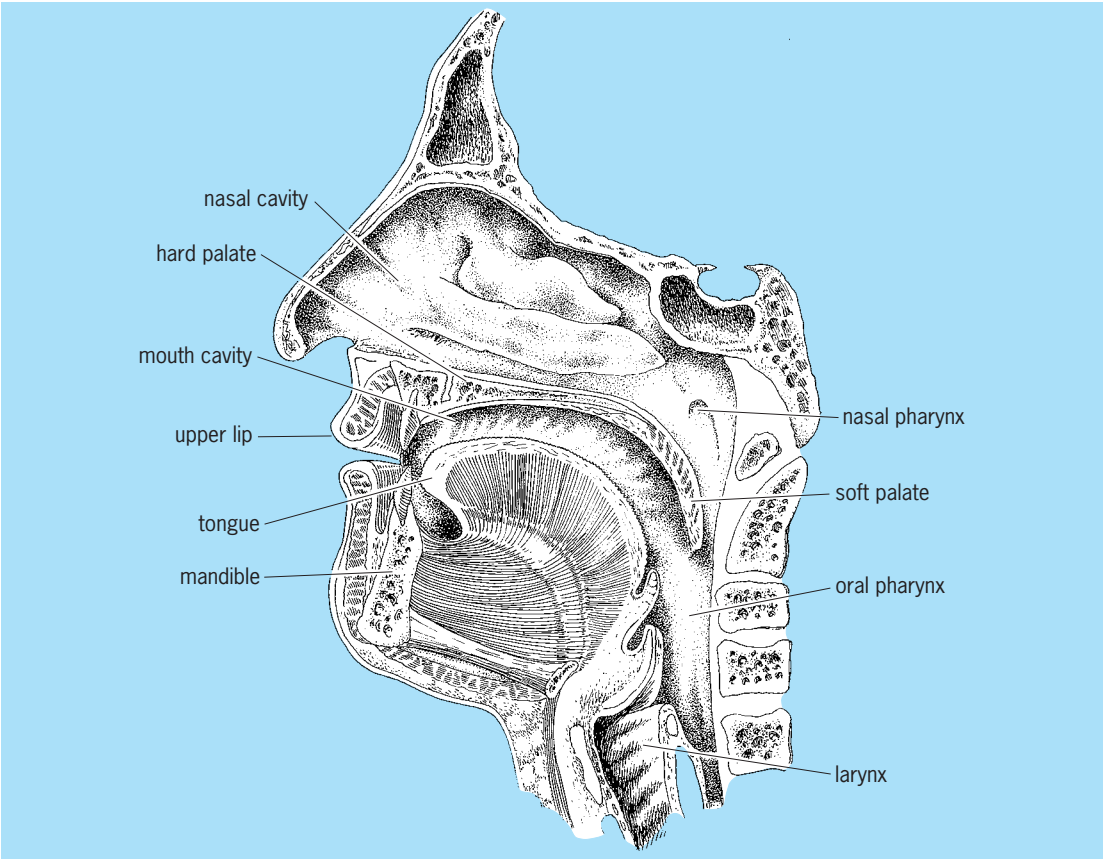

Median section of the human head showing the palate (composed of hard and soft parts) and its relations with nearby structures. (*After M. W. Woerdeman, Atlas of Human Anatomy, vol. 2, McGraw-Hill, 1950*)

Hard palate. The palate of mammals consists of two portions (see **illus.**). The hard palate, more anterior in position, underlies the nasal cavity, whereas the soft palate hangs like a curtain between the mouth and nasal pharynx. The hard palate has an intermediate layer of bone, supplied anteriorly by paired palatine processes of the maxillary bones, and posteriorly by the horizontal part of each palate bone. The oral surface of the hard palate is a mucous membrane, covered with a stratified squamous epithelium. Anteriorly in humans there are four to six transverse palatine ridges; these diminish in prominence between fetal life and old age. A submucosal layer bears pure mucous glands and binds the membrane firmly to the periosteum of the bony component. Above the bone is the mucous membrane that constitutes the floor of the nasal cavity. There is a falsely stratified, ciliated epithelium underlaid by mixed seromucous glands. Nearest to the periosteum of the bone is a layer of elastic fibers.

Soft palate. The soft palate is a backward continuation from the hard palate. Its free margin connects on each side with two folds of mucous membrane, the palatine arches, enclosing a palatine tonsil. In the midline the margin extends into a fingerlike projection named the uvula. Both the hard and soft palate bear a seam, or raphe, along the midline. The oral side of the soft palate continues as the covering of the hard palate, and the submucosa contains pure mucous glands. The intermediate layer is a sheet of voluntary muscle, to which several palatal muscles contribute. The nasal side continues the structures described for the hard palate, but posteriorly, near the free margin, the epithelium becomes stratified because it makes contact at times with the nasopharynx.

Besides separating the nasal passages from the mouth, the hard palate is a firm plate, against which the tongue crushes and manipulates food. The soft palate, at rest, is pendant. In sucking, swallowing, or vomiting, it is raised to separate the oral from the nasal portion of the pharynx. This closure prevents food from passing upward into the nasopharynx and nose. The closing action also occurs in speech, except for certain consonants requiring nasal resonance. The soft palate can also be lowered into contact with the root of the tongue. The palate develops from lateral folds of the primitive upper jaw that meet the fuse in the midline. Bone differentiates in the front half and muscle in the remainder. When the process of fusion fails to any degree along its course, there results a malformation known as cleft palate. *See* CLEFT LIP AND CLEFT PALATE; SPEECH.

Leslie B. Arey

Paleobiochemistry

The study of chemical processes used by organisms that lived in the geological past. Most information on the nature of life in the geological past comes from the study of fossils; a record of biochemical processes that occurred can be found in the organic molecules of sedimentary rocks and fossils. The organic matter in fossil fuel deposits (coal, petroleum, and oil shale) and finely dispersed in shales and limestones represents the debris of cells which have been chemically altered to a more stable form. The relatively reactive organic constituents of cells are subject to bacterial and chemical transformation on the death of the organism. Most organic matter of living organisms is consumed in the metabolism of other organisms and returned to the seas and to the atmosphere. A small amount is deposited in the sediments, where it is preserved or converted to more stable entities. A comparison of the molecular structure of these preserved organic compounds with that of components of living cells enables the researcher to identify similarities and dissimilarities between past and present biochemistry. *See* PETROLEUM.

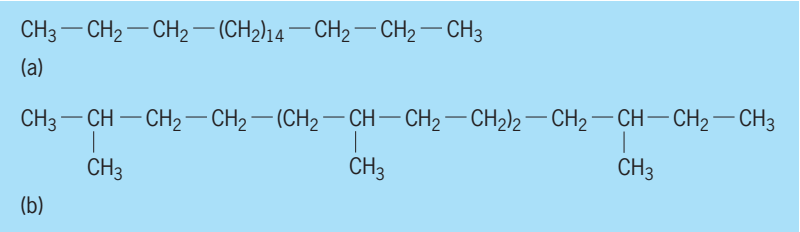

Fig. 1. Alkane isomers in sedimentary rock. (*a*) Normal alkane (*n*-eicosane). (*b*) Isoprenoid alkane (phytane).

Alkanes. Saturated hydrocarbons (alkanes) are among the most stable organic compounds and presumably represent transformation products of the lipid fraction of cells. Two major classes of alkanes in sedimentary rocks are the normal hydrocarbons with linear arrays of carbon-to-carbon bonds, and the isoprenoid hydrocarbons with combinations of five-carbon branched-chain units (**Fig. 1**). *See* ALKANE.

Normal and isoprenoid carbon chains are used preferentially in present-day lipids. There are many thousands of possible spatial arrangements (isomers) of alkanes ($C_{20}H_{42}$); yet those two structural types are the predominant ones both in living cells and in sedimentary rocks. Thus the evidence indicates that this mode of biochemical synthesis has been in use for over 1.5×10^9 years. *See* LIPID METABOLISM.

Tetrapyrrole compounds. Molecules with tetrapyrrole ring structures, such as chlorophyll (**Fig. 2**) and hemin, are used by all living organisms as either photosynthetic or respiratory pigments. In a sedimentary rock, such compounds are transformed into more stable metalloporphyrins and have been found in rocks as old as 1.1×10^9 years. The use of tetrapyrrole-containing molecules in energy-transformation processes obviously evolved very early in the history of terrestrial life. *See* CHLOROPHYLL; HEMOGLOBIN.

Asymmetrical molecules. The ability to synthesize compounds with molecular asymmetry is a characteristic property of living organisms and is due to the use of enzymes as catalysts in biochemical processes. Such asymmetrical compounds display optical activity; that is, their solutions can rotate the plane of polarization of light. Optically active hydrocarbons

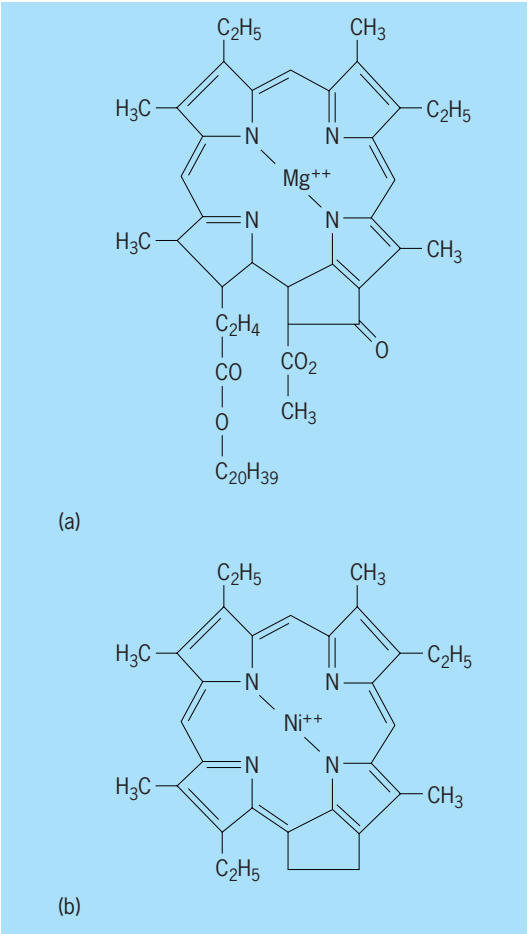

Fig. 2. Structural formulas of two pigment molecules. (*a*) Chlorophyll *a*. (*b*) Metalloporphyrin.

have been detected in petroleums as old as 4×10^8 years. It is believed that the optically active preserved hydrocarbons are the steranes, which have been derived from steroids of cells (**Fig. 3**). If correct, this indicates that the use of enzymes and steroids is an ancient feature of terrestrial life. *See* ENZYME; OPTICAL ACTIVITY; STEROID.

Amino acids. The structural units of proteins are amino acids, of which there are about 25. While the proteins and most of the amino acids are not very stable in the geological environment, seven of the amino acids are stable, can persist for great lengths of time, and have been found in fossils several hundred million years old. These amino acids are alanine, glycine, glutamic acid, leucine, isoleucine, proline, and valine. If any of the other common amino acids, such as serine, threonine, and arginine, which are not very stable, were to be detected in a very ancient fossil, contamination by recent organic matter would be suspected. Such contamination is a serious problem in paleobiochemical studies. *See* AMINO ACIDS; PROTEIN.

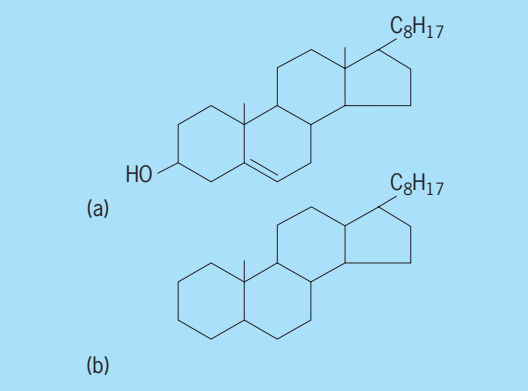

Fig. 3. Structural formulas for a steroid compound and for its possible sterane derivative. (*a*) Cholesterol (steroid). (*b*) Cholestane (sterane).

Paleobiochemical studies have shown that a number of the common chemical processes used by living organisms today have been in use for a very great length of time. Paleobiochemical techniques will be used on materials returned from extraterrestrial sources to determine whether life exists outside of Earth, and if so, to see how it may differ from life on Earth. *See* FOSSIL; PALEOECOLOGY; PALEONTOLOGY.

Thomas C. Hoering

Bibliography. P. H. Abelson (ed.), *Researches in Geochemistry*, vol. 2, 1967; K. A. Kvenvolden (ed.), *Geochemistry of Organic Molecules*, 1980.

Paleobotany

The study of fossil plants of the geologic past. A paleobotanist is a plant historian who carefully pieces together the geologic history of the plant kingdom. Other organisms, including fungi and various types of microscopic plankton, are also studied by paleobotanists. Paleobotany is a branch of paleontology that requires a knowledge of both plant biology (botany) and the geological sciences. *See* BOTANY; FOSSIL; PLANT KINGDOM.

Materials used by paleobotanists to reconstruct plant life through geologic time include fossilized remains preserved in the rock layers of the Earth. Materials such as fossil leaves, seeds, fragments of wood, fruits, and flowers are used to interpret the biology and evolution of ancient plants. In addition, the products and distribution of plants are recorded in the rock record in the form of coal, resins, various chemicals, and other substances produced by plants. Some plant fossils, such as pollen grains and spores, are studied by palynologists; palynology has been especially important in mineral and petroleum exploration and in correlating rock layers that are widely separated geographically. From all of these materials, paleobotanists attempt to reconstruct the habit, structure, and biology of plants that grew on the Earth millions of years ago. *See* PALYNOLOGY.

Paleobotanists are also concerned with the myriad interactions that existed in ancient communities and ecosystems. An examination of fossil plants throughout geologic time points to a succession of organisms that inhabited the Earth, from relatively simple forms that are preserved in rocks about 3.6 billion years old to forms that are progressively more complex and suited to different types of habitats. Paleobotany not

only involves the collection, description, reconstruction, and naming of fossil plants but is also concerned with the evolution of major groups, relationships that exist between fossil and living forms, how ancient plants functioned and reproduced, what type of environment they lived in, how they were fossilized, and many other biological and geological topics.

Plant fossilization. In modern ecosystems, plant litter generally decomposes as a result of fungal and microbial degradation and mechanical disintegration. However, a very small percentage of all plant materials that inhabit the Earth at any one time become buried in accumulating sediments, and are thus preserved as fossils. Plants are preserved in a variety of ways, and various combinations of physical and chemical processes are involved at the time of preservation. For a plant part to become fossilized, several unique circumstances must occur. First, the item must be close to a site where sediments are accumulating; second, it must be rapidly buried. For example, a leaf that falls from a tree has little fossilization potential if it is transported in a stream a long distance, since along the way it will become torn, abraded, or otherwise destroyed. Rapid burial is also necessary to ensure that the biological activities of various microbes do not destroy the tissues of the leaf. *See* BIODEGRADATION.

Plant parts are best preserved in very fine grained silts and shales, which are the lithified muds of ancient deposits. Such sediments generally yield excellent fossils because the small grain size preserves minute details of the leaf; coarser-grained sediments such as sands generally do not reproduce delicate features. Plant parts composed of thick-walled cells or those having fibers within their tissues have a better chance of being preserved (fossilization potential) than more delicate tissues composed of thin-walled cells. The aboveground parts of most plants are covered by a layer of cutin or wax that is generally resistant to decay and aids in the preservation of some tissues.

In addition to the many variables that contribute to the processes of fossilization, the site of burial is critical as to whether fossil remains will ever be discovered. For example, sediments deposited in certain types of lakes, rivers, and swamps may be easily lost as the depositional system changes over time, or as the sediments are eroded away as younger river systems cut into older fossil-bearing rocks. As a result of all of these biological and physical constraints, relatively few organisms become fossils, and an even smaller number are ever found by paleobotanists. In addition, the number of fossils observed is reduced due to the fact that most of the sediment making up the crust of the Earth has never been examined for evidence of ancient plants. It is from this incomplete fossil record that the paleobotanist attempts to reconstruct the flora at a particular point in time and space.

Plants can be fossilized in a variety of ways, each yielding different types of information about the ancient flora. Thus, the paleobotanist must employ different techniques to extract the maximum amount of information from each fossil. There may be more techniques used to study fossil plants than there are ways in which the plants are preserved. There is no single "best" technique. The questions that are asked and the type of preservation will dictate how to examine a particular fossil plant. *See* FOSSIL SEEDS AND FRUITS.

Impressions. Impressions probably represent the most common type of plant fossil. They occur when a plant part is covered by sediment and the water is squeezed from the cells and tissues. Cells that make up the plant are eventually degraded, perhaps as a result of microbial activities, so that only a shallow negative, or imprint, of the plant organ (for example, a fern leaf) remains. The paleobotanist determines the shape, venation, and other structural features of the leaf and notes any characteristics that make it unusual. A leaf forms a fossil impression much like a modern leaf on the surface of wet concrete. The leaf is partially impressed into the surface and results in a negative outline of its shape. In a short time, the leaf is gone, degraded by microorganisms, but the impression remains until it is weathered or abraded away.

Although impressions lack organic material, sometimes when the rock is composed of exceedingly fine grains it is possible to see the outlines of leaf cells and surface features such as hairs. This can be accomplished by using a light microscope or by preparing a latex replica of the surface of the fossil; this replica can be examined in a scanning electron microscope, where greater magnification and resolution are available.

Compressions. The same processes that form impressions are responsible for the formation of compressions. However, compressions result in the surface of the fossil being covered by a thin film of carbonaceous material (**Fig. 1**). This film of carbon represents the original plant, in which the cells have been highly compressed. Most of the best-preserved compressions are found in clay and shale, although some have been discovered in volcanic ash deposits. The most attractive compressions are often preserved in a light-colored matrix that makes the dark carbonaceous film easy to study and photograph. In a few compressions the remains of the plant are exceptionally well preserved and may contain various cell organelles. In some plants the preservation is so good that details can be examined with a transmission electron microscope. In other compressions all that remains is the cuticle that covered the surface of the leaf. Specimens with cuticle can be prepared to show the distribution and types of cells and stomata present on the surface of the leaf. The ultrastructure and chemistry of the cuticle have been observed in some fossil plants. With some compressions the entire fossil can be immersed in acid to dissolve away the rock matrix, leaving the cuticle, which is resistant to the acid, available for detailed examination. Because of the fragile nature of cuticle this technique can be used only with relatively small plants.

Coal is a type of compression fossil. In coals that have not been subjected to extensive heat and

Fig. 1. Compression specimen of the Triassic leaf type *Dicroidium*.

pressure, it is possible to recognize plant materials. Pollen grains (**Fig. 2**), spores, and fragments of plant cuticle can be macerated (disintegrated by cutting into very small fragments) from coals and provide information about the kinds of plants that lived when the coals were being formed.They can also be examined with the aid of certain types of electron microscopes. There are a few examples where the entire coal seam is made up of the cuticle envelopes of plants. After washing and staining, these "paper coals" can be mounted directly on microscope slides for examination. *See* COAL; COAL PALEOBOTANY.

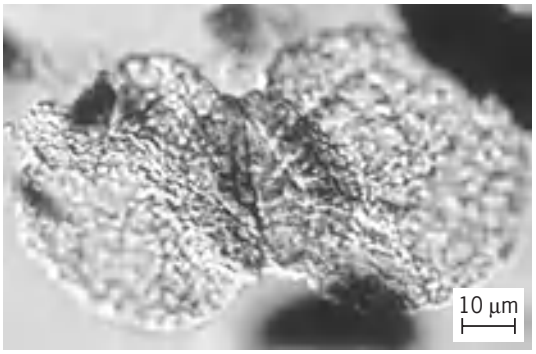

Fig. 2. Permian pollen grain macerated from Antarctic coal.

Molds and casts. Often, three-dimensional plant parts such as stems and seeds become buried in sediment but are not flattened in the process (as in impressions). After a period of time, the plant tissues disintegrate, leaving a negative or mold of the plant organ. Occasionally, the mold becomes filled with new sediment and forms a three-dimensional replica of the original plant part, termed a cast. Molds and casts do not contain any original plant material; however, they show the external form and the original shape of the fossil plant. This type of fossilization process is especially helpful in reconstructing large plants, for example in determining the morphology of underground organs, and understanding the position of large branches.

Permineralizations. Although permineralization is relatively rare, this type of fossilization process provides perhaps the greatest amount of information about the cellular detail and tissue systems in some fossil plants. Permineralizations form when a plant part becomes immersed in solutions containing concentrations of dissolved minerals that permeate the tissues, filling the cell lumens and intercellular spaces. As the minerals crystallize, the plant part becomes entombed by the rock, and thus is not distorted from its original shape as in the case of compression fossils.

The most common entombing minerals in permineralizations are carbonates and silicates, but oxides may also be present. In this fossilization process the cell walls remain organic; however, there apparently is some type of chemical change to the cellulose and lignin. Cell organelles such as nuclei, various membrane systems, and even chromosomes have been described from some permineralizations. Some of the best-known permineralizations are Carboniferous coal balls, and silicified peat blocks collected from Antarctica. To study permineralizations the surface of the rock can be etched with acid so that the plant cell walls are slightly elevated from the surface of the rock. The surface then can be covered with acetone and a thin sheet of cellulose acetate rolled on to the surface. The sticky acetone/acetate mixture flows in and around the elevated cell walls and dries. This thin slice of plant material now embedded in the acetate can be "peeled" from the surface, pulling with it the cells of the permineralized plants. After slightly grinding the surface of the rock again with an abrasive (silicon carbide) powder, a new peel can be prepared. Thus, permineralized plant parts can be sectioned much like that done with extant plants. Peels then can be mounted on microscope slides and photographed (**Fig. 3**). *See* COAL BALLS.

Petrifactions is another type of fossilization which is similar to permineralization. In petrifactions, however, the original cell walls are completely replaced by other minerals. Woods from the Petrified Forest in Arizona, the Cerro Cuadrado Petrified Forest in Patagonia, and Rhynie chert in Scotland are examples of this type of fossilization process. To study petrifactions a thin slice of the petrified plant is cemented to a glass slide and ground with abrasive powder until it transmits light (**Fig. 4**). This thin section is the only

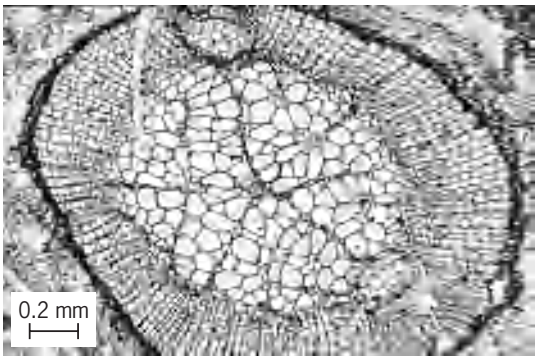

Fig. 3. Cross section of a permineralized stem from a Carboniferous coal ball.

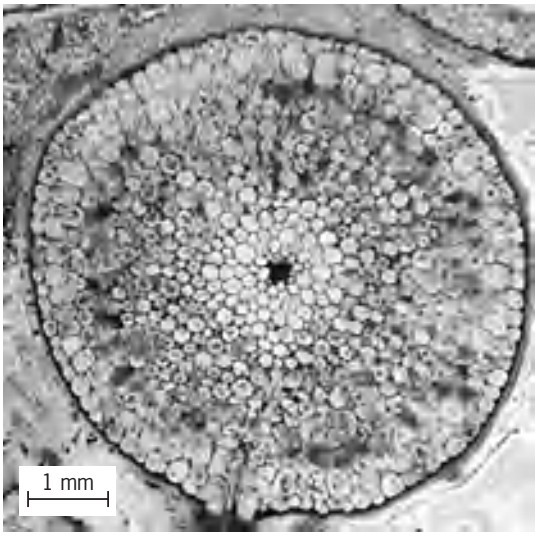

Fig. 4. Thin section of a petrified Rhynie chert stem.

means available to examine the cells that are petrified. *See* PETRIFACTION; PETRIFIED FORESTS.

Unaltered plant material. In special situations, plants may be preserved in an almost unaltered form. Examples are pollen grains and spores, various forms of plant exudate (such as amber), the silica frustules of diatoms, and even certain types of algae that cause calcium carbonate to be formed.

Various chemical signatures that remain in the rock long after the organism is gone represent still another form of unaltered plant material. Chemical compounds such as lignin, sterols, and carboxylic acids are examples of chemical fingerprints that can be detected in rock layers millions of years old. In some instances, it is possible to relate the chemical constituents of the fossils to those found in closely related living plants, and thus to suggest the relationship between living and fossil organisms more clearly. There are several reports of chemical traces in Precambrian rocks, some as old as 3.8 billion years, but the nature of these rocks is being debated. Such "fossils" require critical analysis since contamination from recent organisms is a major problem. A case in point is the deoxyribonucleic acid (DNA) reported from an exceptionally well-preserved fossil leaf of Miocene age. Some believe that this DNA was a contaminant from younger or modern organisms, perhaps even from recent bacteria rather than that of the fossil leaf.

Amber comprises fossilized plant resins of a variety of different types and can be found in the fossil record as early as the Carboniferous. It is an example of unaltered plant material. Phytochemical and x-ray diffraction techniques have been useful in determining the kinds of plants responsible for producing the different forms of amber. Because of its sticky nature, amber also serves as a fossilization matrix for other organisms, including insects. Small flowers, fungi, algae, and various wind-borne plant parts, including pollen grains, spores, and seeds, have also been discovered from samples of amber throughout the geologic column. *See* AMBER.

Other types of preservation are less frequently encountered in the rock record. It is rare that a single fossil is preserved in only one way. Impressions sometimes contain a bit of carbonaceous film, making part of the fossil a compression; and on the outside of permineralized coal balls is often found coal, a type of compression.

Regardless of how plants are fossilized, the most important factor is determining what techniques should be used by the paleobotanist to extract the maximum amount of information. For many fossils (both plant and animals) a particular specimen is the only available example of that organism. Thus, a critical step in paleobotany is first evaluating how the organism is preserved, and then determining how to examine it based on the types of questions to be asked.

Still other paleobotanists use biomechanical principles and morphometric approaches to study fossil plants.

Use of fossil plants. The geological record of past floras can be traced to the present. Fossil plants have provided an enormous body of knowledge about the evolution of plants from the earliest Precambrian unicellular forms to the complex multicellular flowering plants used for food and shelter today. Tracing plants through geologic time has provided the basis for examining how and when the first plants became adapted to a terrestrial environment, when

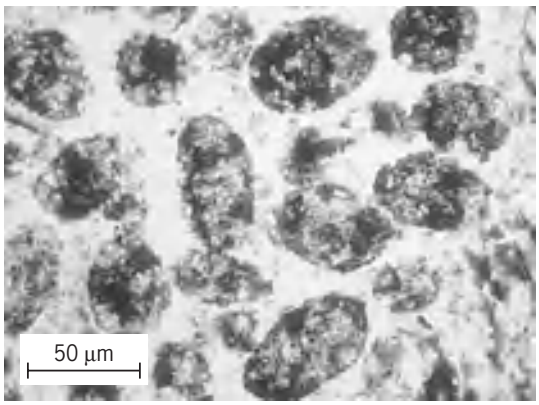

Fig. 5. Several coprolites found within the tissues of a Carboniferous plant.

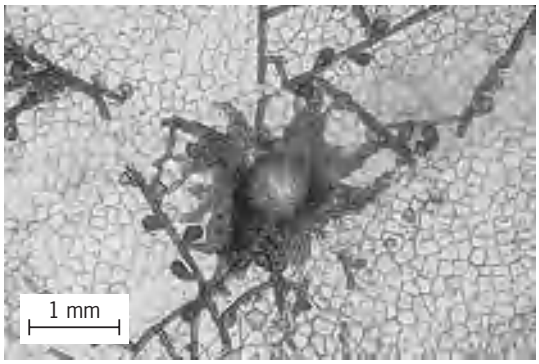

Fig. 6. Cuticle of a fossil (Eocene) leaf showing the outline of the epidermal cells and fungus with specialized hyphae.

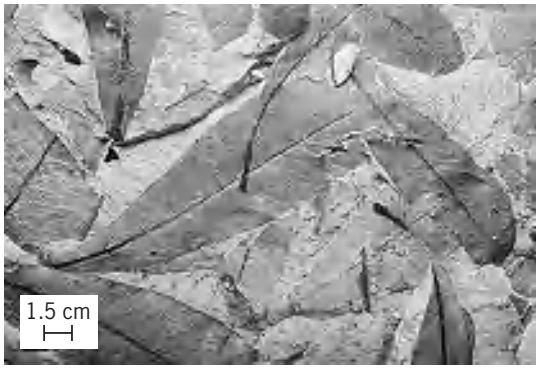

Fig. 7. Several *Glossopteris* leaves from the Permian of Antarctica. Leaves of this type were some of the first evidence that at one time the southern continents were all physically linked into a major continental plate called Gondwana.

the angiosperms first appeared, why seeds evolved, and why some plants acquired the ability to produce wood. Once plants are identified, the composition of the community and ecosystem can be examined, including the types of interactions that existed with other organisms; for example, certain types of herbivores can be identified in permineralized tissues based on coprolites (**Fig. 5**) or wound responses. Fungi (**Fig. 6**) and various types of galls can be identified on fossil leaves that in turn can provide information about the ecosystem. Fossil plants also demonstrate that certain groups that once flourished (for example, the cycads) are now geographically restricted and probably heading, like other plants in the geologic past, toward extinction. *See* PLANT GEOGRAPHY.

The physical environment where the plants lived is an important component of paleobotany, as is the geologic setting. How were the plants preserved, and does the assemblage of fossils that the paleobotanist has collected represent an accurate depiction of the total community? How was preservation affected by seasonality and the types of plants? What types of plants lived together, and how were they distributed in time and space? Another set of questions that can be answered by the paleobotanist relates to biostratigraphy. For example, plant fossils have been used to correlate rock layers that are widely separated geographically. Thus, where the geologic time range of a plant or assemblage of plants has been determined in rocks whose relative age is known, the presence of the same plants in other rocks will indicate an equivalent age (**Fig. 7**). *See* STRATIGRAPHY.

Fossil plants can also be valuable in reconstructing climates of the past. For example, modern woody trees in temperate climates produce different sizes and numbers of cells throughout their growing season (**Fig. 8**). This periodicity results in the production of annual rings in the wood that can be related to climate signals such as moisture, light, temperature, and even the presence of fire. Permineralized fossil wood sometimes shows similar ring features that can be used to interpret the climatic conditions under which the plant was growing many millions of years ago. For example, trees more than 30 m (100 ft) tall grew at approximately 85° South during the Permian and Triassic in what is now Antarctica and produced wide growth rings (**Fig. 9**). These rings are quite different from those produced in temperate climates today. While modern tree rings form in response to changes in temperature and moisture, those in the Antarctic woods formed in response to changing light levels. The overall morphology and various structural features on certain fossil leaves also can be used to interpret climates of the past. For

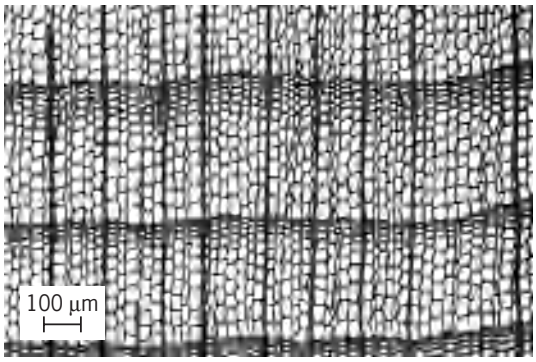

Fig. 8. Section showing tree rings from a temperate plant, showing change in cell size and wall thickness.

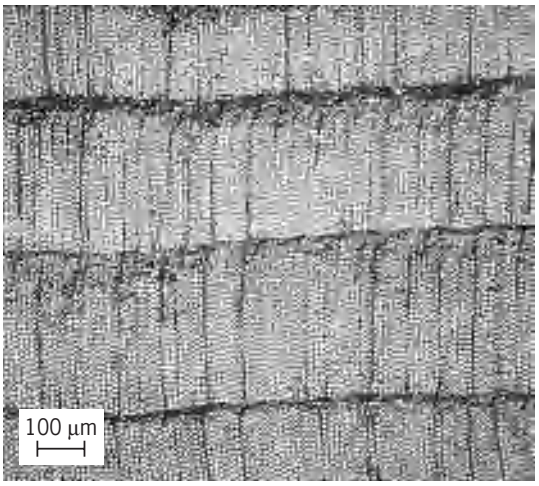

Fig. 9. Section showing rings from a tree with uniform cell size and wall thickness, but growing at approximately 82° South during the Triassic.

example, the number and distribution of stomata on living angiosperm leaves can be measured against known carbon dioxide concentrations in the atmosphere today. This stomatal density can then be calculated for closely related fossil plants. The differences produce a proxy record that is used to infer changes in carbon dioxide concentration through geologic time. *See* PALEOCLIMATOLOGY.

In modern floras, there are certain species of plants that are restricted to specific climatic regimes. Analyzing the margin of the leaves has proven to be a useful method of relating modern and fossil plants, and hypothesizing the climatic conditions under which the plants once grew. For example, the discovery of palms, laurels, magnolias, and cycads in the early Tertiary of southeastern Alaska provides strong evidence of subtropical to warm-temperate, lowland conditions which existed at that time. *See* PALEONTOLOGY; POSTGLACIAL VEGETATION AND CLIMATE.

Thomas N. Taylor

Bibliography. T. P. Jones and N. P. Rowe, *Fossil plants and Spores: Modern Techniques*, Geological Society, 1999; K. J. Niklas, *The Evolutionary Biology of Plants*, University of Chicago Press, 1997; W. N. Stewart and G. W. Rothwell, *Paleobotany and the Evolution of Plants*, 2d Ed., Cambridge, 1993; T. N. Taylor and E. L. Taylor, *The Biology and Evolution of Fossil Plants*, Prentice Hall, 1993; B. A. Thomas, *The Evolution of Plants and Flowers*, Eurobook Ltd., 1981; B. A. Thomas and R. A. Spicer, *The Evolution and Palaeobiology of Land Plants*, Croom Helm, 1987; K. J. Willis and J. C. McElwain, *The Evolution of Plants*, Oxford, 2002.

Paleoceanography

The study of the history of the ocean with regard to circulation, chemistry, biology, and patterns of sedimentation. The source of information is largely the biogenous deep-ocean sediments, so the field may be considered a branch of sedimentology or paleontology. However, there are also strong links to geophysics, marine geochemistry, and mathematical modeling. Geophysical sciences are called upon for reconstruction of geography (position of continents, horizontal motion of the ocean floor), topography (changing depth patterns in the ocean, general subsidence of any given piece of sea floor), and dating (radioisotopes, magnetic patterns on the sea floor and in sediments). Geochemical analyses deliver information on sediment composition (stable and unstable isotopes; major and minor components such as carbonate, opal, and trace elements). Such information is useful in correlating sedimentary sequences and, when combined with geochemical arguments, yields insights about the dynamics of carbon and nutrient cycles. Mathematical modeling introduces a strong quantitative element. It draws upon the knowledge reservoir of modern oceanography and climatology. *See* CLIMATOLOGY; GEOCHEMISTRY; GEOPHYSICS; PALEONTOLOGY; SEDIMENTOLOGY.

Origins and development. The study of marine sediments on land is as old as geology itself. Modern paleoceanography is set apart by the study of sediments recovered from the ocean, especially the deep ocean, and by the use of concepts developed by oceanographers (controls on ocean currents and upwelling, vertical stratification, heat budget, nutrient and carbon cycles, pelagic biogeography and water masses). Important early studies were on cores raised by the German *Meteor* Expedition (1925–1927) to the central and southern Atlantic, by the Südpolar Expedition (*Gauss*, 1901–1903) in the Antarctic, and by the United States cable ship *Lord Kelvin* (1936) in the North Atlantic. The glacial debris zones noted in the *Kelvin* cores are now commonly referred to as Heinrich layers; they are witness to sporadic input of iceberg armadas during the last glacial epoch. *See* GLACIAL EPOCH.

Comparisons of climate-related changes between major ocean basins became possible through the systematic recovery of long cores by the circumglobal Swedish Deep Sea Expedition (1947–1949) on the research vessel *Albatross*. Using new technology developed by Börje Kullenberg in Gothenburg, the expedition retrieved cores up to 15 m (45 ft) long, with records reaching back 500,000–1,000,000 years. Many fundamental paleoceanographic concepts were established by the geologists who analyzed these cores, including the role of trade winds in promoting glacial-age upwelling, the changing supply of North Atlantic Deep Water, the cyclicity of

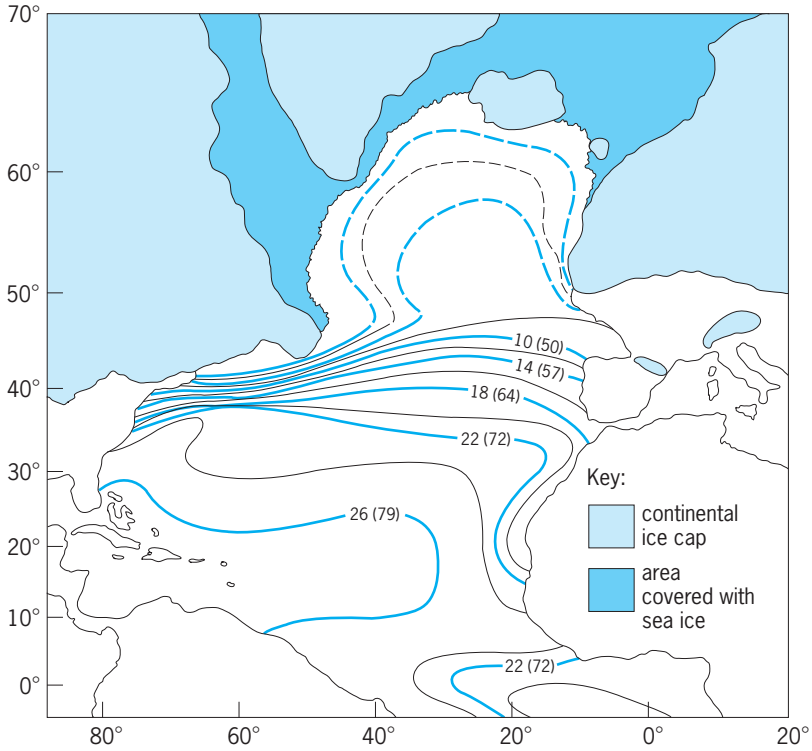

Fig. 1. August temperature of sea surface waters [isotherms in °C (°F) in the North Atlantic during the height of the last glacial 18,000 years before present] broken lines represent extrapolated isotherms (*after A. McIntyre and N. G. Kipp, Geol. Soc. Amer. Mem., 145:61, 1976*). Recent research suggests that the coastal waters off Norway were open during summer.

climatic change in the late Quaternary, and large-scale shifts in biogeographic boundaries.

A quantum jump in paleoceanographic research resulted from the initiation of deep-sea drilling using the research vessel *Glomar Challenger* (1968). Enormous blank regions on the world's map suddenly became accessible for detailed exploration far back into geologic time, that is, into the Early Cretaceous. Highlights of the first decade of drilling results include the documentation of cooling steps as the planet moved into the present ice age; the reconstruction of long-term fluctuations in the carbonate compensation depth; the documentation of large-scale salt deposition in an isolated Mediterranean basin; and the discovery of temporary anoxic conditions in the Cretaceous deep sea.

Other important developments, in the decades following the *Albatross* expedition, were associated with the collection of large numbers of piston cores from many parts of the ocean, the improvement of analytical instrumentation, and the introduction of mathematical statistics to the interpretation of fossil assemblages. In the 1970s, a different emphasis on climatic change arose, stemming from concerns about the impact of human activities on climate, that is, the expectation that the increase in carbon dioxide in the atmosphere would produce global warming. A temperature map of the glacial ocean (the so-called 18,000-year map) resulted from the efforts of the CLIMAP group (**Fig. 1**), which provided an opportunity to test the application of climate models to conditions that are fundamentally different from those of today. *See* CLIMATE MODELING.

Quaternary research. A useful time scale for the ice-age cycles of the Quaternary first emerged from the combination of magnetic reversal stratigraphy and oxygen isotope measurements in the same piston cores. Based on these data, it was possible to show that changes in deep-ocean sedimentation (and hence climate) are influenced (or governed) by orbital forcing, that is, changes in seasonal insolation intensity in high latitudes. However, the physical mechanisms remain unknown by which Milankovitch forcing (changes of up to 8% in the amount of summer sunlight reaching the high latitudes of the Northern Hemisphere due to changes in the geometry of the Earth's orbit with time) is translated into changes of ice mass (and hence of sea level). In particular, it is not clear why there is a pronounced change in the response to forcing 900,000 years ago, when obliquity-dominated cycles gave way to much longer cycles soon displaying a period of about 100,000 years (100 kyr) [the Mid-Pleistocene climate shift; **Fig. 2**]. *See* CHEMOSTRATIGRAPHY; GEOLOGIC THERMOMETRY; INSOLATION; MAGNETIC REVERSALS; PALEOCLIMATOLOGY.

One likely origin for the 100-kyr cycles of the late Quaternary is a general increase of ice mass throughout the previous 2 million years, together with sinking and erosion of the ice-bearing continental crust. An increased amount of ice was thus based below sea level during glacial periods. Upon a minor rise of sea level, at times of exceptionally warm summers

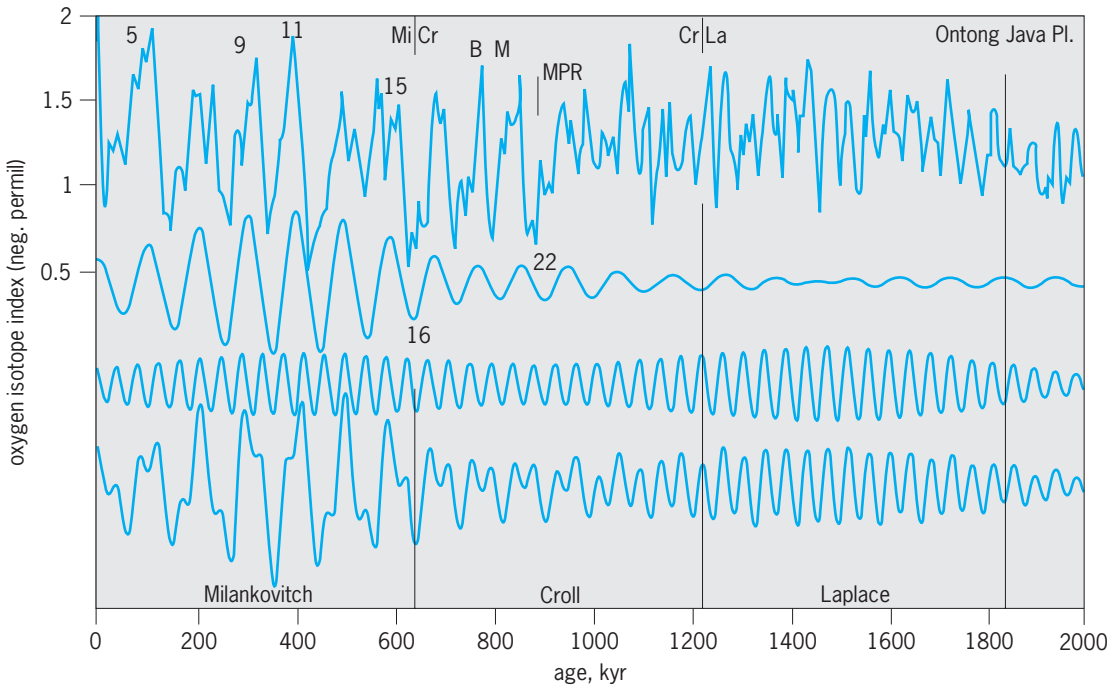

Fig. 2. Oxygen isotope values of the planktonic foraminifera *Globigerinoides sacculifer* from Ontong Java Plateau (Site 806, ODP Leg 130) are shown in the top curve. The two curves in the middle show 100-kyr (eccentricity) and 41-kyr (obliquity) cycles extracted from the isotope curve. The bottom curve is a combination of the two middle curves. The good agreement between top and bottom curves is an indication for orbital forcing of climate variability. The last 2000 kyr can be divided into three periods: Milankovitch period dominated by eccentricity, Laplace period showing obliquity-dominated variations and a transition period (Croll). MPR = mid-Pleistocene revolution. (*Modified from W. H. Berger and G. Wefer, Naturwissenschaften, 79:541–550, 1992*)

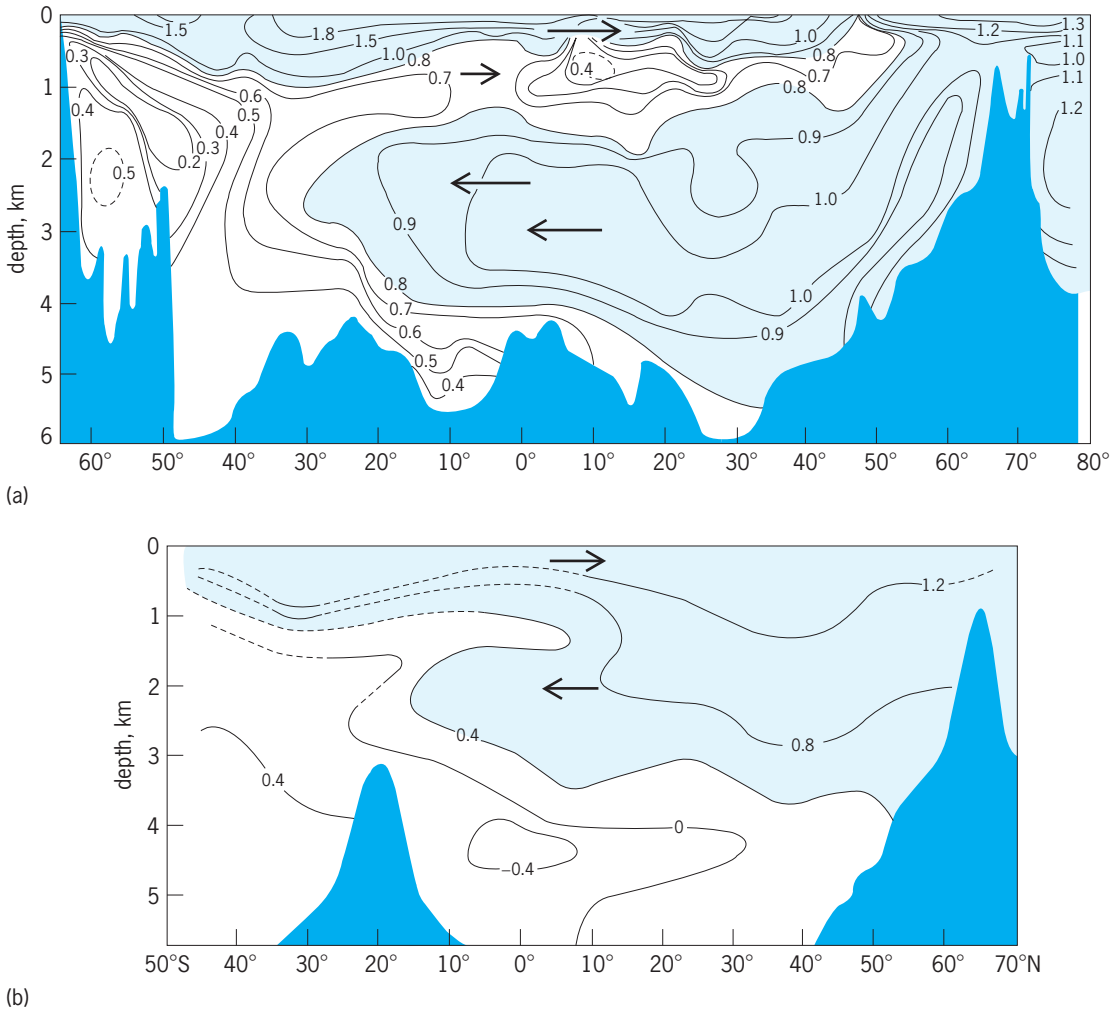

Fig. 3. Deep-water patterns and flow in the South Atlantic (*a*) present and (*b*) last glacial maximum at about 20,000 years ago. Stratification is exhibited by the pattern of $\delta^{13}C$ values of dissolved inorganic carbon (‰ parts per thousand). During the last glacial period, the bottom-near layer of cold Antarctic Water tended to thicken, as the overlying North Atlantic Deep Water flow was reduced. (*Modified from G. Wefer et al., The South Atlantic: Present and Past Circulation, Springer, 1996*)

in northern latitudes, such marine-based ice can become unstable and move out to sea. As a result, support is removed from adjacent land-based ice, which then collapses in surges. Such ice collapse, in combination with changes in ocean circulation providing for changes in heat budget of high northern latitudes, could produce the terminations (that is, sudden ice wasting following maximum glaciations) that dominate the late Quaternary climate history. Thus, the 100-kyr cycles are best understood as an internal oscillation based on ice-shield instability, forced by the Milankovitch mechanism (summer insolation in high northern latitudes). Direct evidence for inherent instability comes from the quasi-cyclic collapse of modestly sized ice masses in Canada (Heinrich events). These iceberg floods occur roughly every 7000 years. *See* QUATERNARY.

Heat conveyor and abrupt climatic change. An important concept concerning the role of the ocean in abrupt climatic change, such as seen during deglaciation, is the Atlantic heat conveyor, which is associated with anti-estuarine circulation in the North Atlantic. Water from the South Atlantic moves north across the Equator at shallow depths to replace North Atlantic Deep Water, which forms in the seas adjacent to Greenland due to cooling of salty surface waters, and moves south across the Equator at depth. At present the Atlantic heat conveyor is very effective in warming western Europe and the northern regions. During glacial times, North Atlantic Deep Water formation was greatly reduced (**Fig. 3**). A slowing down or shutting down of this conveyor (for example, from meltwater input) greatly changes the heat budget of the North Atlantic. Modeling suggests that sudden warming or a large input of fresh water in high latitudes could cause the conveyor to stop, with the oceanic system moving into another equilibrium. The speed of this transformation will depend on the feedback processes involved, which generally take place within a few centuries. According to the circulation models, however, it is possible that strong forcing could change the deep circulation completely or collapse the conveyor in less than 10 years. Once the conveyor was stopped, self-stabilization effects could hinder the resumption of the circulation, even if the induced forcing was no

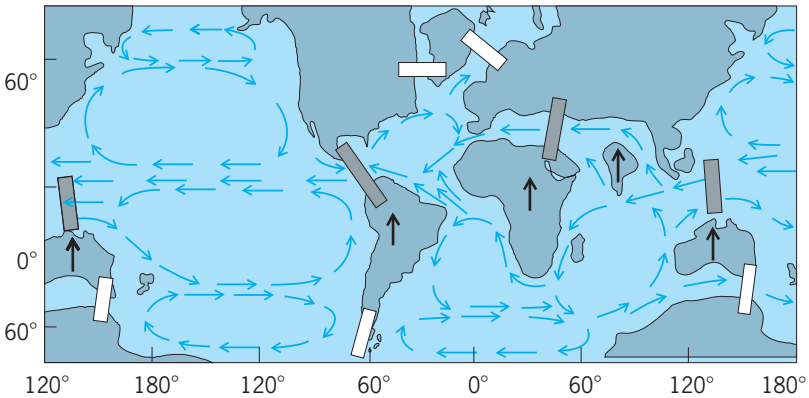

Fig. 4. Geography of the middle Eocene (45 Ma) and major critical valve points for ocean circulation. Tropical valves are closing (shaded rectangles) and high-latitude valves are opening (white rectangles) throughout the Cenozoic. Arrows indicate direction of circulation. (*After B. U. Haq, in E. Seibold and W. H. Berger, The Sea Floor, Springer, 1993*)

longer effective. These modeling results, combined with the paleoceanographic evidence, have important implications for future global warming. It is possible that the Atlantic deep circulation will decrease significantly in the twenty-first century within the framework of the increased greenhouse effect. Although this prognosis is subject to large uncertainties (mainly due to the low resolution of the ocean models), historical experience leaves no doubt that the scenario is a possible one. *See* GREENHOUSE EFFECT.

Cenozoic research. The Cenozoic Era (Age of Mammals) comprises the Paleogene Period (65–25 million years ago, or Ma) and the Neogene Period (25 Ma to present). During the Cenozoic, the Tethys seaway, which allowed free exchange between the major ocean basins in tropical regions, closed, and the main mixing region of the ocean's water masses was transferred to the circumpolar Southern Ocean, ensuring a general cooling trend in deep-water temperatures (**Fig. 4**). Marine life in the early Paleogene (Paleocene) was marked by recovery from the end-of-Cretaceous catastrophe. The deep ocean was as warm as ever, but cooled from about 50–40 Ma on. The end of the Paleocene witnessed a remarkable warming spike, apparently linked to a pulse of methane from the large-scale melting of methane clathrates. The event produced the widespread extinction of benthic foraminifers. In the Eocene, diversity of marine organisms reached a maximum. Around 40 Ma, there was a major change in the ocean. Zones of high productivity developed along the Equator and in coastal regions, extracting silicate and phosphate from the ocean at large, thus impoverishing the pelagic realm. Sea level dropped markedly, and much carbonate was transferred from shelves to the deep ocean. Deep-water sources moved from subtropical regions to higher latitudes. Diversity of both plankton and benthic organisms reached a low point within the Oligocene (the last period of the Paleogene, 36–25 Ma). In the middle to late Oligocene, the first large glaciers grew in Antarctica, as seen in ice-rafted debris around the continent. *See* EOCENE; HYDRATE; PALEOCENE.

In the Neogene, the ocean acquired thoroughly modern characteristics, such as cold deep waters, a strong thermocline, vigorous surface currents, and increased upwelling along the Equator and in eastern boundary coastal oceans. Major geographic events included the opening of the Drake Passage, closing of the Indonesian Passage, isolation of the Mediterranean, and closing of the Panama gateway. Deep-water temperatures decreased further. This cooling occurred in steps, presumably as a result of critical changes in geography (mountain building), as well as climatic threshold events (snow and ice buildup first on and around Antarctica and later in northern latitudes, with attendant positive feedback on cooling from albedo increase). The possibility of positive feedback on cooling from the carbon cycle is of interest, involving increased removal of carbon dioxide from the atmosphere from weathering of fresh rocks (mountain buildup), from increased organic carbon deposition in upwelling areas (thermocline effect), from increased burial in terrigenous marine sediments, and from increased uptake by the ocean (alkalinity increase from transfer of carbonate deposition from shelves to deep ocean, general cooling of the ocean). *See* CENOZOIC.

Cretaceous research. The ocean of the Cretaceous Period (140–65 Ma) was fundamentally different from the present one, because of geography and a generally warmer climate. Large areas of the continents were flooded during much of the period, so marine sediments (especially carbonate) tended to accumulate in shallow seas and the deep ocean was starved. Deep waters were relatively warm (perhaps comparable to today's Mediterranean). The oxygen supply was modest, because of the warm temperature and a more sluggish circulation. Productivity (it may be assumed) was greatly reduced because of a shortage of nitrate (which is used as a source of oxygen by certain bacteria, when oxygen is in short supply) and perhaps a shortage of trace metals (very little windblown material entered the ocean).

Of special interest are organic-rich deposits that occur in thin layers within sequences of clay or chalk, quite commonly in cyclic fashion suggesting orbital forcing. The most widespread black shale deposits of this kind are found during the Aptian and early Albian (110–100 Ma) in the South Atlantic. At the same time, this basin was a long narrow seaway, providing a cul-de-sac opening to the south with opportunities for pulsed estuarine circulation pattern. Black shales also occur elsewhere, however, so some deposits imply a global factor rather than a regional one. A global signal, apparently, is contained in the record of the differences in carbon-13 ratios ($\delta^{13}C$) of pelagic marine limestones, with major excursions in the Aptian and at the Cenomanian-Turonian boundary. It has been suggested that major volcanic activity (related to mantle plumes reaching the Earth's surface and producing flood basalts such as Ontong-Java Plateau) provided for conditions favorable for the accumulation of organic-rich sediments in the deep ocean. If true, volcanism might be ultimately responsible for a large portion of the marine

petroleum deposits that derive from black shales in the middle Cretaceous. *See* BLACK SHALE; CRETACEOUS; MARINE SEDIMENTS.

Cretaceous-Tertiary boundary. A central question in the earth sciences is the reason for the demise of oceanic plankton within a short and well-defined interval at the end of the Cretaceous. Tropical forms were most affected, while high-latitude forms and deep-water forms largely escaped extinction. It is postulated that a planetoid 10 km (6 mi) in diameter hit the Earth and caused major environmental deterioration. A prolonged darkening of the Sun (from particles in the stratosphere), acid rain (from burning nitrogen in superheated air), and drastic changes in temperature, among other effects, are possible. The planktonic $\delta^{13}C$ record shows excursions toward lighter carbon at or immediately after the extinction event, and this evidence has been interpreted as indicating reduction of primary production. However, the cause for the reduction (if real) remains unknown. *See* EXTINCTION (BIOLOGY); OCEANOGRAPHY; PALEOGEOGRAPHY; TERTIARY.

Wolfgang H. Berger; Gerold Wefer

Bibliography. W. W. Hay, Paleoceanography: A review for the GSA Centennial, *Geol. Soc. Amer. Bull.*, 100:1934–1956, 1988; J. P. Kennett (ed.), The Miocene ocean: Paleoceanography and biogeography, *Geol. Soc. Amer. Mem.*, 163:1–337, 1985; E. Seibold and W. H. Berger, *The Sea Floor*, 2d ed., Springer, 1993; G. Wefer et al. (eds.), *The South Atlantic: Present and Past Circulation*, Springer, 1996.

Paleocene

The oldest of the seven geological epochs of the Cenozoic Era, and the oldest of the five epochs that make up the Tertiary Period. The Paleocene Epoch represents an interval of geological time (and rocks deposited during that time) from the end of the Cretaceous Period to the beginning of the Eocene Epoch. Recent revisions of the geological time scales place the Paleocene Epoch between 65 to 55 million years before present (m.y. B.P.). *See* CENOZOIC; EOCENE; GEOLOGIC TIME SCALE; TERTIARY.

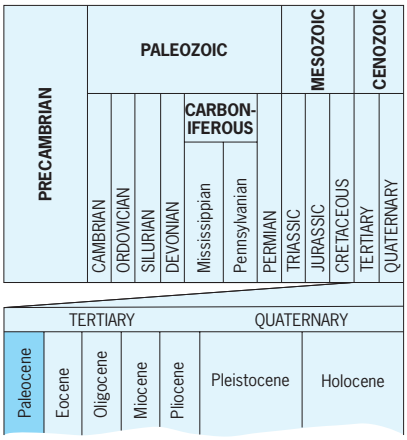

The close of the Cretaceous Period was characterized by the disappearance of many terrestrial and marine animals and plants. The dawn of the Cenozoic in the Paleocene Epoch saw the establishment of new fauna and flora that have evolved into modern biota. The concept of the Paleocene as a separate subdivision of the Tertiary was introduced by the paleobotanist W. P. Schimper in 1874. He observed a distinctive assemblage of plant fossils in the lower Eocene nonmarine sediments of the Paris Basin that he separated out as representing an independent epoch. In the older Lyellian classification, this interval was a part of the Eocene Epoch that constituted the oldest part of the Tertiary Period. The Paleocene deposits had lateral equivalents bearing early Tertiary mammals and were devoid of remains of dinosaurs, which had become extinct at the end of the Cretaceous Period. Schimper also noticed that the Paleocene flora contained numerous components that are now typical of Northern Hemisphere, in contrast to the Cretaceous when Southern Hemispheric floras prevailed.

Subdivisions. Modern schemes of the Paleocene subdivide it into Lower and Upper series, and their formal equivalents, the Danian and Selandian stages. Some authors prefer to use a threefold subdivision of the Paleocene, adding the Thanetian at the top. The older, Danian lithofacies generally tend to be calcium carbonate–rich (pure chalk in the Danian type area), whereas the younger, Selandian and Thanetian facies have greater land-derived components and are more siliciclastic (sand, sandstone, marl). *See* CHALK; FACIES (GEOLOGY); MARL; SAND; SANDSTONE.

The Danian Stage was proposed by the French geologist E. Desor in 1847 with its type locality near Copenhagen, Denmark. Desor, however, regarded the Danian to be the youngest stage of the Cretaceous. He equated the Danian Chalk and Limestone at the Danish type localities of Faxe and Stevns Klint to the uppermost Cretaceous strata of the Paris Basin largely on the basis of similarities in lithology and the contained echinoidal fauna. It was not until 1897 that another French scientist, A. de Grossouvre, suggested that it would make more sense to place the upper limit of the Mesozoic at the base of the Danian and a major extinction level where ammonites, belemnites, rudistids, and dinosaurs disappeared. The Russian geologist P. Bezrukov in 1934 was the first to actually assign the Danian to the Tertiary, based on paleontological studies in the Ural River sections. In the 1960s the Danian Stage was formally placed in the lowermost Paleocene when it was demonstrated that these strata lacked diagnostic ammonites typical of the Late Cretaceous and contained a microfauna with greater affinity to those of the Tertiary. A major faunal break and hiatus below the Danian and equivalent strata in many sections in the world reinforced its definite Tertiary character.

The Selandian Stage was also defined based on sections in Zealand, near Copenhagen, by the Danish stratigrapher A. Rosenkrantz in 1924. Rosenkrantz, however, did not designate a formal type section

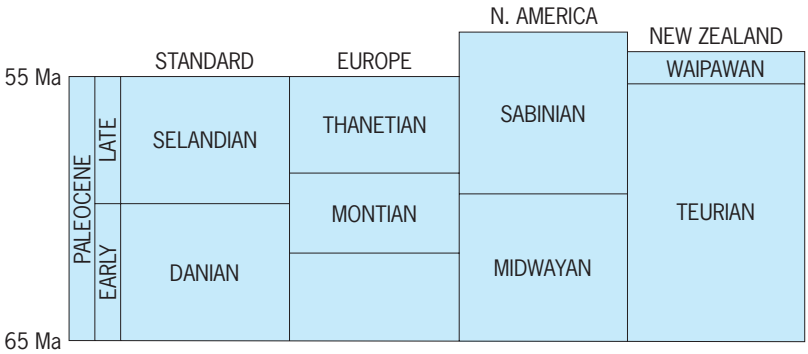

Fig. 1. Standard Paleocene stages and their temporal equivalents in Europe, North America, and New Zealand.

for the stage. Later workers have amended the concept of standard Selandian to include all of the temporal equivalent interval of the Late Paleocene series, although the debate about the lithostratigraphic continuity within these sections continues. This stage is, nevertheless, preferable to the previously used upper Paleocene standard stage of Thanetian, which spans only the very upper part of Late Paleocene.

Regional subdivisions of the Paleocene include the Montian Stage, proposed by G. Dewalque in 1868, with a type locality near Mons, Belgium. Like the Danian, the Montian Stage has also been a subject of much discussion. It is now considered to be temporally equivalent to the youngest part of Danian and older part of Selandian (**Fig. 1**). The Thanetian Stage in Britain was based on the marine Thanet Sands of the Isle of Thanet in Kent. It was first raised to a stage level by E. Renevier in 1873 and is time-equivalent to the upper part of Selandian. The Landenian Stage was proposed by A. Dumont in 1839. It also has its type locality in Belgium and is often used in northwestern Europe as a stage younger than the Montian. Detailed studies have shown the Belgian Landenian also is equivalent to the British Thanetian. In the United States the Gulf coast Paleocene subdivisions include the Midwayan and the older part of the Sabinian stages. The base of the Midwayan shows a major faunal and lithological hiatus. In California, local stratigraphers have often used the Ynezian Stage to represent an informal equivalent of the upper Paleocene. In the former Soviet Union, the Kachinian Stage has been shown to encompass much of the Paleocene, extending slightly into the earliest Eocene. In New Zealand, the Teurian Stage spans most of the Paleocene, with the overlying Waipawan Stage ranging into the youngest Eocene. In Australia, the Late Paleocene strata are sometimes ascribed to the regional Wangerripian Stage. Paleontologists working with vertebrate fossils have often used their own informal age classifications for assemblage associations. In North American, they subdivide the Paleocene into the Puercan, Torrejonian, and Tiffanian ages, whereas in Europe the terms Dano-Montian and Cernayasian have often been used to span the Paleocene in vertebrate paleontological literature. *See* STRATIGRAPHY.

Tectonics, oceans, and climate. Several major tectonic events that began in the Mesozoic continued into the Paleocene. For example, the Laramide Orogeny that influenced deformation and uplift in the North American Rocky Mountains in the Mesozoic continued into the Paleocene as a series of diastrophic movements, which ended abruptly in the early Eocene. *See* OROGENY.

On the ocean floor the most notable tectonic events were the separation of the Seychelles from the rapidly northward-moving Indian plate in the early Paleocene, and the initiation of sea-floor spreading in the Norwegian-Greenland Sea, between North America and Greenland in the late Paleocene. The Indian plate had broken away from eastern Gondwanaland in the Late Cretaceous and started moving relatively rapidly northward some 80 m.y. ago. In the early Paleocene, the spreading ridge between Madagascar and India jumped northeastward toward India. This initiated the spreading between India and the Seychelles Platform and the formation of the Chagos-Laccadive transform ridge.

During the Paleocene Epoch, the Indian plate continued its rapid flight across the eastern Tethys Ocean, the ancestral seaway that occupied the position of the modern northern Indian Ocean, to eventually collide with the Asian mainland in the mid-Eocene around 50 m.y. ago. In the Paleocene, the eastern Tethys was also characterized by another very active transform, the Ninetyeast Ridge. This ridge was formed when the Indian plate moved over a fixed mantle plume hot spot in the Late Cretaceous. *See* PLATE TECTONICS.

The lower-latitude shallow-water seas, as exemplified by the Tethys seaway, received thick deposits of calcium carbonates during the Paleocene. However, the seaway became progressively narrower and shallower, and the nature of carbonate accumulation changed correspondingly.

The sedimentary record of the Paleocene in the northwestern Atlantic indicates that a relatively calm regime of predominantly calcareous sediments typical of the Late Cretaceous was largely replaced by facies that represents more vigorous bottom waters that packed great erosive power due to increased convective overturn and dynamic ocean bottom currents. Although the deep-water connection between the North and South Atlantic was already established by the Early Paleocene, the South Atlantic sedimentary record indicates that in this basin the extensive erosion on the sea floor did not begin until the Late Paleocene. Sedimentary hiatuses representing this dynamic change in bottom-water regime during various times in the Paleocene are also common in other oceans. Thus, the Paleocene deep ocean can be said to have been characterized by extensive erosion and redeposition of sediments on the deep sea floor, reflecting expanded bottom-water activity, compared to the Cretaceous. *See* BASIN.

As a whole, the Cenozoic Era is characterized by a long-term withdrawal of the seas from coastal and inland oceans. In the latest Cretaceous, the global sea level had already begun to fall from the all-time

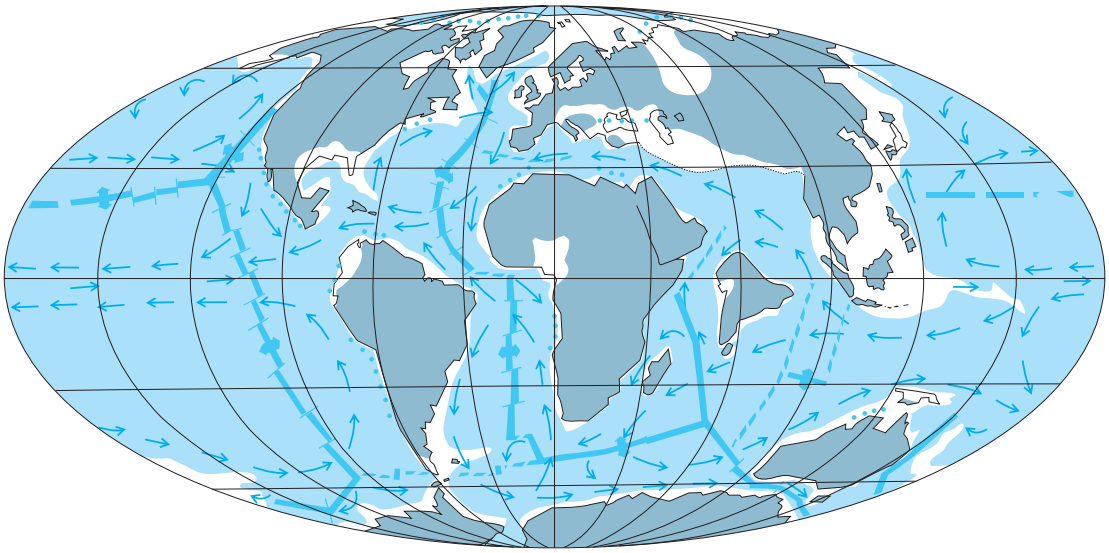

Fig. 2. Paleogeography, oceans, and circulation patterns of the Paleocene. (*After B. U. Haq and F. W. B. van Eysinga, Geological Time Table, Elsevier, Amsterdam, 1998*)

Mesozoic high of the mid-Cretaceous. The Cretaceous-Tertiary boundary event is marked by a general transgression in the Danian, following a sea-level fall in the latest Cretaceous. This trend toward overall regressing seas was further accentuated in the later part of the Paleocene. Whereas the Danian is typified by relatively high sea levels, it is followed by a major fall in the early part of the Selandian that has been recorded in many parts of the world. The restriction of the seas in the Selandian is reflected in the sediment facies of the Paleocene, which changed from more carbonate-dominated during the relatively higher sea levels of the Danian, to more terrigenous and siliciclastic in the Selandian. Latest Paleocene once again saw a sea-level rise that continued into the early Eocene. The Selandian sea-level fall, which is estimated at around 120 m, was large enough that it could have exposed extensive areas of the continental margins. Major coastline retreats are followed by stream incision of the shelf, and accordingly there is evidence that in the Selandian the sea-level fall is associated with major episodes of canyon formation, cut by rivers that migrated to the shelves during the eustatic drop. *See* CONTINENTAL MARGIN.

Paleogeographic-oceanographic considerations of the Paleocene record (**Fig. 2**) suggest that the western Tethyan seaway between Europe and Africa was open and a circumglobal Tethys current flowed through it, dominating the tropical oceanographic scene of the Paleocene. Toward the north a major epicontinental seaway, the Ural Sea, separated Asia from Europe through much of the Paleocene and Eocene. Epicontinental gulfs also existed on the Asian, African, and North and South American continents. These shallow interior seas were sites of extensive carbonate deposition, and also may have been prone to high production and preservation of organic matter that is important for hydrocarbon source-rock accumulation.

The establishment of deeper connections between the North and South Atlantic in the Paleocene facilitated enhanced deep-water flow from the northern to the southern basin. Similar to the Cretaceous, the source of deep water in the Paleocene was most likely in the warm low and middle latitudes, rather than the cooler higher latitudes as in the Neogene. Warm saline bottom water was characteristic of this epoch. In the south, the Drake Passage between South America and Antarctica was still closed, although Australia had already separated from Antarctica by Paleocene time. The lack of circum-Antarctic flow precluded the geographic isolation of Antarctica and the development of cold deep water from a southern source. *See* PALEOCEANOGRAPHY; PALEOGEOGRAPHY.

The isotopic and paleontological climatic proxy indicators all point to an overall rise in global temperatures in the Paleocene that led to a period of peak warming in the latest Paleocene and Early Eocene. A mean sea-surface temperature of around 10°C (50°F) in the higher latitudes is indicated by oxygen isotopic analysis of marine plankton. A prominent, relatively cooler interval that coincides with the major lowering of sea level has been recorded in the Late Paleocene. The marine microplankton (foraminifera and calcareous nannoplankton) exhibit latitudinal migrationary patterns that are consistent with major climatic fluctuations indicated by the Paleocene isotopic record. The Selandian sea-level lowering and concomitant climatic cooling is accompanied by a migration of high-latitude microplankton assemblage toward low latitudes in the Atlantic Ocean, while the latest Paleocene is characterized by a poleward migration of warm, low-latitude assemblages. *See* GEOLOGIC THERMOMETRY; MARINE SEDIMENTS.

Terrestrial floras and faunas corroborate the peak warming in the latest Paleocene and Early Eocene and suggest that the warm tropical-temperate belt may have been twice its modern latitudinal extent.

The temperate floral and faunal elements extended to 60°N, which has been used as an argument to invoke a very low angle of inclination of the Earth's rotational axis in the Paleocene-Eocene. Alternatively, the mild, equable polar climates and well-adapted physiological responses of plants and animals of those times to local conditions may be enough to explain the presence of a rich vertebrate fauna on Ellesmere Island in arctic Canada. *See* CLIMATE HISTORY; PALEOBOTANY; PALEOCLIMATOLOGY.

At the close of the Paleocene Epoch, a prominent carbon-isotopic ($\delta^{13}C$) shift occurred in the global carbonate reservoir that coincides with the peak warming at this time. Recent studies have ascribed this to the dissociation of sediments on continental margins that contained methane hydrates. When hydrates dissociate due to reduced hydrostatic pressure or increased temperature on the sea floor, large quantities of methane can be released into the water and atmosphere. In the latest Paleocene, bottom-water temperature increased rapidly with a coincident negative shift of $\delta^{13}C$ in the global carbon reservoir by 2.5%. This isotopic change was accompanied by important biotic changes in the oceanic microfauna and was synchronous in the oceans and on land. The rapid (<10,000 years) and prominent ^{12}C enrichment of the global carbon reservoir cannot be ascribed to increased volcanic emissions of carbon dioxide, changes in oceanic circulation, and/or terrestrial and marine productivity. The increased flux of methane from gas-hydrates into the ocean-atmosphere system and its subsequent oxidation to carbon dioxide is held responsible for this isotopic excursion in the inorganic carbon reservoir. High-resolution data support the gas-hydrate connection to the latest Paleocene is abrupt climate change. Evidence from two widely separated sites, from the low-latitude and southern high-latitude Atlantic Ocean, indicates multiple injections of methane with global consequences during the relatively short interval at the end of the Paleocene. *See* HYDRATE.

Life. The Paleocene Epoch began after a meteorite struck the Earth, causing massive extinctions at the end of Cretaceous and decimating a large percentage of the terrestrial and marine biota. On land the last of the dinosaurs are the most familiar casualty of this event. In the oceans, all ammonites, genuine belemnites, rudistids, most species of planktonic foraminifera and nannoplankton, and marine reptiles disappeared at the close of the Cretaceous Period. Even though some groups, such as squids, octopus, nautilus, and a few species of marine plankton, survived, the genetic pool was relatively small at the dawn of the Tertiary Period. The recovery of the marine biota was, however, fairly rapid after the mid-Paleocene due to overall transgressing seas and ameliorating climates. By the Late Paleocene, the biota was well on its way to explosive evolutionary proliferation and high diversification of the Eocene. In the Paleocene, endemism in marine and terrestrial biota increased. For example, the larger foraminifera, *Nummulites*, thrived in the shallow seas of the Tethyan margins, but were excluded from the marginal seas of the New World. Planktonic microfauna also show increasing hemispheric and latitudinal provincialism. A new group of warm-water phytoplankton, the discoasters, made their first appearance in the Late Paleocene, and soon thereafter proliferated in the Eocene. The end of the Paleocene Epoch saw marked changes in deep-water circulation of the world ocean that resulted in a massive extinction of the benthic marine species. *See* EXTINCTION (BIOLOGY).

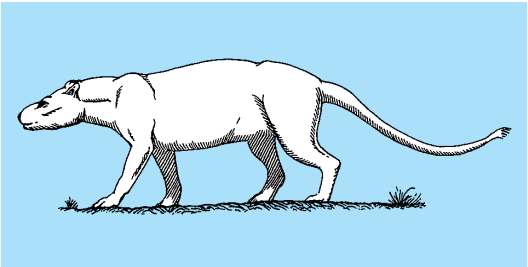

Fig. 3. Paleocene condylarth *Tetraclaenodon*. This five-toed ungulate was near the ancestry of the first horses. (*After R. A. Stirton, Time, Life and Man, John Wiley, 1959*)

On land the large dinosaurs, which had been on the decline for over 20 m.y, died out at the close of the Cretaceous Period. However, smaller reptiles, including alligators and crocodiles, and some of the land flora escaped extinction and continued into the Paleocene. The Paleocene saw the first true radiation of mammals. The mammals of this epoch were characteristically primitive and small in size (50 cm or 20 in. or less). Their principal record is to be found in terrestrial deposits in Asia and North America. A typical example of Paleocene mammals is provided by the condylarths (odd-toed ungulates), which evolved from a Late Cretaceous mammalian lineage. These small animals (**Fig. 3**) were ancestral to many carnivores as well as the important group of perissodactyls, which include horses, rhinos, and tapirs. The appearance of the first true grasses at the close of the Paleocene may have helped the later radiation of these animals. Ancestral insectivores, rodents, and primates also had their beginning in the Paleocene. During the early part of this epoch, North and South America were connected temporarily, allowing free migration of animals, but were separated soon thereafter. This led to evolutionary isolation of South America and survival of an archaic fauna dominated by anteaters, armadillos, and opossums, until the Pliocene when the isolation finally ended following the connection of the two continents at the Panama isthmus. Similarly, as the continent of Australia became more isolated geographically, its mammalian fauna, such as the marsupials, became sequestered and more specialized. *See* DINOSAUR; MAMMALIA; PALEONTOLOGY.

Bilal U. Haq

Bibliography. B. U. Haq and F. W. B. van Eysinga, *Geological Time Table*, Elsevier, Amsterdam, 1998; R. D. Norris and U. Roehl, Carbon cycling and chronology of climate warming during the Palaeocene/Eocene transition, *Nature*, 401:775–778, 1999;

C. Pomerol, *The Cenozoic Era: Tertiary and Quaternary*, 1982.

Paleoclimatology

The study of ancient climates. Climate is the long-term expression of weather; in the modern world, climate is most noticeably expressed in vegetation and soil types and characteristics of the land surface. To study ancient climates, paleoclimatologists must be familiar with various disciplines of geology, such as sedimentology and paleontology, and with climate dynamics, which includes aspects of geography and atmospheric and oceanic physics. A compelling need exists for societies to be able to predict climate change. Understanding the history of the Earth's climate system greatly enhances the ability to predict how it might behave in the future.

Information about ancient climates comes principally from three sources: sedimentary deposits, including ancient soils; the past distribution of plants and animals; and the chemical composition of certain marine fossils. These are all known as proxy indicators of climate (as opposed to direct indicators, such as temperature, which cannot be measured in the past). In addition, paleoclimatologists use computer models of climate that have been modified for application to ancient conditions. *See* GEOLOGY; PALEONTOLOGY.

Climate dynamics. Like modern climatologists, paleoclimatologists are concerned with boundary conditions, forcing, and response. Boundary conditions are the limits within which the climate system operates. The boundary conditions considered by paleoclimatologists depend on the part of Earth history that is being studied. For the recent past, that is, the last few million years, boundary conditions that can change on short time scales are considered, for example, atmospheric chemistry. For the more distant past, paleoclimatologists must also consider boundary conditions that change on long time scales. Geographic features—that is, the positions of the continents, the location and orientation of major mountain ranges, the positions of shorelines, and the presence or absence of epicontinental seaways (large areas of the continents flooded during times of high sea level)—are important for understanding paleoclimatic patterns. In addition, the solar constant and the rotation rate of the Earth have changed through geologic time and also must be considered in studying paleoclimates, particularly before about 4×10^8 years ago (shelly fossils first appeared in the record about 6×10^8 years ago). Forcing is a change in boundary conditions, such as continental drift, and response is how forcing changes the climate system. Forcing and response are cause and effect in paleoclimatic change. Paleoclimatic change during the history of the Earth has taken place on time scales that are very short (thousands of years) and very long (hundreds of millions of years). *See* CONTINENTAL DRIFT; CONTINENTS, EVOLUTION OF; PALEOGEOGRAPHY; PLATE TECTONICS.

Modeling ancient climates. The record of climate in sediments and rocks is remarkable because it illustrates a multitude of abrupt and slow changes in climate throughout Earth history. The record is also incomplete. Many older rocks have been subject to erosion or metamorphism, and farther back in time the record becomes increasingly sparse. In addition, paleoclimatologists do not have true "paleo" barometers or thermometers. Rather, paleoclimate is studied indirectly through its imprint on sediments and organisms. Thus, ancient climates cannot be reconstructed on a global scale. However, the available data are useful if climate is modeled; the data can then be used to test aspects of the model predictions, and thereby the model itself.

Although many aspects of paleoclimatic change are actively studied, some topics are particularly important in that the studies are challenging traditionally accepted ideas. These topics include the role of moutain ranges in determining global climate patterns, the climate systems of supercontinents, warm polar climates, the role of ocean circulation in global climate, and the role of carbon dioxide in long-term climate change.

The study of paleoclimates has benefited greatly from the application of comprehensive computer models. The development of atmospheric and oceanic models based on fundamental physical laws of fluid motion, energy, and momentum was at the leading edge of research in the atmospheric and oceanic sciences during the 1980s. Very rapid development in predictive model capability has occurred because of increased supercomputer power, advances in programming, and a growing inventory of observations about the modern climate system. *See* CLIMATOLOGY; MODEL THEORY; SIMULATION.

Paleoclimatic indicators. Proxy indicators of paleoclimate are abundant in the geologic record. Many were first discovered during the nineteenth century and helped scientists and naturalists appreciate the dynamism of Earth history. Fossil palm leaves and crocodiles in the Canadian Arctic Islands and thick coal seams in Antarctica are among the many seemingly anomalous occurrences that indicate that climate has changed. Many of the changes are explained by continental drift, but other changes represent fundamental differences between modern climatic patterns and ancient ones.

Sedimentary indicators. Important sedimentary indicators forming on land are coal, eolian sandstone (ancient sand dunes), evaporites (salt), tillites (ancient glacial deposits), and various types of paleosols (ancient soils), such as bauxite (aluminum ore) and laterite (some iron ores). Coals may form where conditions are favorable for growth of plants and accumulation and preservation of peat, conditions that are partly controlled by climate, especially seasonality of rainfall. Eolian sandstones and evaporites are indicative of arid climates. Paleosols record water-table levels and intensity of weathering, both of which are dependent on temperature or rainfall. Certain marine sediments also provide important information about ancient atmospheric and oceanic

circulation patterns. These are bedded chert, which is formed from the siliceous shells of microscopic marine plants and animals, and phosphate, both of which form in upwelling zones, that is, zones of high biologic productivity in the oceans. At present, these zones owe their distribution to global wind patterns, and thus geologic evidence of upwelling helps in the interpretation of ancient atmospheric circulation. Extensive deposits of sediments rich in organic substances provide information about stagnation, which may be related to temperature, and productivity in the oceans. In addition to specific types of sedimentary rock, patterns of sedimentation also provide information about paleoclimates. Layering in lake deposits, for example, may be controlled by seasonal changes that are, in turn, controlled by climate. Much of the understanding of more recent climatic change has come from such deposits. *See* BAUXITE; CHERT; COAL; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; LATERITE; MARINE GEOLOGY; MARINE SEDIMENTS; PALEOSOL; SALINE EVAPORITES; UPWELLING.

Fossil indicators. Fossils provide information about climate mostly by their distribution (paleobiogeography), although a few specific types of fossils may be indicative of certain climatic conditions. The latter are usually fossils from the younger part of the geologic record and are closely related to modern species that have narrow environmental tolerances. Paleobiogeographic patterns, on the other hand, commonly follow climatic gradients such as temperature and rainfall; they are useful qualitatively, and sometimes quantitatively, for delineating those gradients. Paleobiogeographic patterns are especially useful in the earlier part of Earth history, because the evidence required for other methods usually was not preserved or is difficult to interpret.

In addition to indicator species and paleobiogeographic patterns, fossils, particularly plants, commonly show morphological features that are related to climate. The morphology of leaves of woody angiosperms (flowering plants) is controlled by mean annual temperature, mean annual range of temperature, and rainfall (**Fig. 1**). In addition, the thickness and structure of the cuticle (waxy outer layer) of angiosperm and other types of leaves are partly dependent on water supply. Leaf-margin analysis of angiosperms has provided a quantitative tool for the estimation of mean annual temperature; the proportion of smooth-margined versus toothed-margined leaves is highly correlated with this climatic parameter. Finally, growth rings in fossil wood are sensitive to frosts, rainfall, and temperature during the growing season. These methods are useful only for the latter part of Earth history, after trees and flowering plants evolved. *See* BIOGEOGRAPHY; DENDROCHRONOLOGY; FOSSIL; PALEOBOTANY.

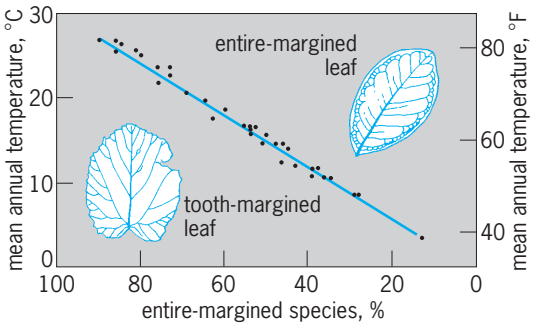

Fig. 1. Correlation of leaf-margin shape with mean annual temperature. (*After J. A. Wolfe, Temperature parameters of humid to mesic forests of eastern Asia and relation to forests of other regions of the Northern Hemisphere and Australasia, U.S. Geol. Surv. Prof. Pap. 1106, 1979*)

Geochemical indicators. The third type of information available for documenting paleoclimatic patterns and change is stable isotope geochemistry of fossils and certain types of sedimentary rock. Many elements that are used by organisms to make shells, teeth, and stems occur naturally in several different forms, known as isotopes. The most climatically useful isotopes are those of oxygen (O). Most oxygen occurs as ^{16}O, an atom with eight protons and eight neutrons in the nucleus; however, some oxygen occurs as ^{18}O, which has two extra neutrons. The rate at which the two isotopes are cycled through the oceans and atmosphere and taken up by organisms differs in response to a number of factors, including temperature; the biology of the organisms also has a strong effect, so that only a few types of organisms can be used for oxygen isotopic studies. In general, organisms living in tropical waters have a higher proportion of ^{16}O in their shells compared with those living in polar waters. In addition, the pro- portions of the two isotopes available in the global system depends on the amount of ice on the Earth's surface. Although the effects of temperature change and ice volume change can be difficult to distinguish, the analysis of oxygen isotopes has provided a powerful quantitative tool for the study of both long-term temperature change and the history of the polar ice caps. Isotopic studies are less reliable for the earlier part of Earth history, because shell material from older rocks tends to be poorly preserved and because the biology of very ancient organisms is unknown. *See* GEOCHEMICAL PROSPECTING; ISOTOPE.

Causes of paleoclimatic change. The rarest of paleoclimatic changes is that which can be attributed confidently to a single cause. A great deal of research in paleoclimatology has been devoted to understanding the causes of climatic change, and the overriding conclusion is that any given shift in the paleoclimatic history of the Earth was brought about by multiple factors operating in concert. Nevertheless, some forcing factors have been identified as being particularly important. The most important forcing factors for paleoclimatic variation are changes in paleogeography and atmospheric chemistry and variations in the Earth's orbital parameters.

Paleogeographic forcing operates on relatively long time scales, millions to tens of millions of years, and occurs in a variety of modes. Changes in continental positions may close off or open up new routes for ocean currents, which can change the distribution of temperature over the Earth's surface. For example, the opening of the Drake

Passage between South America and Antarctica is regarded as a contributing factor to the formation of the Antarctic ice cap. While land connection between the two continents existed, the circumpolar current was deflected equatorward along the western side of the coast while relatively warm water was deflected poleward along the eastern side. With the opening of the Drake Passage, however, the circumpolar current became continuous, isolating Antarctica from the warm, low-latitude waters. Although the paleogeographic change that brought about this isolation—the drifting of Antarctica and South America away from each other—occurred slowly, the climatic change that resulted was relatively rapid. This is an example of a paleoclimatic threshold; the climate system seems to resist change until finally the forcing becomes too great, and then the adjustment takes place quickly.

Sea-level changes also may have brought about paleoclimatic change. Sea level is partly controlled by changes in ice volume at the poles, and thus may sometimes be an effect, rather than a cause, of climatic change. However, sea level also changes in response to mountain-building activity and continental drift, and these changes may be forcing factors for paleoclimatic change. In general, times of high sea level were times of warmth. However, this pattern is not necessarily true in detail, indicating that other factors also must play a role. Among these factors are the location, size, and orientation of mountains and epicontinental seaways. In addition, geochemical models suggest that times of high sea level were also times of high carbon dioxide content in the atmosphere, so the direct paleoclimatic forcing might have been carbon dioxide, not sea level.

Most paleoclimatologic research on atmospheric chemistry has been directed toward modeling the variations in carbon dioxide content of the atmosphere. Carbon dioxide is well known as a so-called greenhouse gas that traps heat at the Earth's surface. Carbon dioxide content is thought to have varied in the past, and strong evidence for such variations during the recent past (the last few thousand years) has been gathered by using detailed records of atmospheric bubbles in ice cores and stable isotopes (**Fig. 2**). In addition, the carbon cycle is relatively well known, and changes in carbon dioxide content in the atmosphere must have taken place, because large fluctuations are observed in the fossil record for other reservoirs for carbon. Finally, carbon dioxide changes appear to be the only mechanism that can account for some global temperature changes. *See* ATMOSPHERIC CHEMISTRY; BIOGEOCHEMISTRY; CARBON DIOXIDE; GREENHOUSE EFFECT.

Characteristics of the Earth's orbit around the Sun vary regularly and lead to cyclical variations in the amount and distribution of solar radiation reaching the Earth. The three characteristics most important to paleoclimatic studies are precession, obliquity, and eccentricity. Precession of the equinoxes determines the time of year when the Earth is closest to the Sun; in the present geological era, this occurs in January. Precession varies with a period of

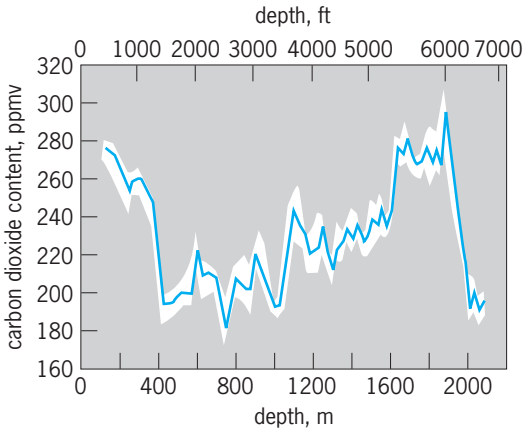

Fig. 2. Carbon dioxide (CO_2) from bubbles in the Vostok ice core from Antarctica. The core spans about 160,000 years. White area represents the data envelope. (*After J. M. Barnola et al., Vostok ice core: A 160,000–year record of atmospheric CO_2, Nature, 329:404–414, 1987*)

about 19,000–23,000 years. Obliquity is the tilt of the Earth's axis relative to the plane of the Earth's orbit around the Sun; at present, obliquity is about 23.5°. Obliquity varies with a period of about 41,000 years. Eccentricity is the shape of the Earth's orbit around the Sun, which varies from circular, allowing solar input to be constant throughout the year, to oval, which creates strong annual variability. Eccentricity varies with a period of about 100,000 years. These are known as Milankovitch cycles, after the Yugoslavian astronomer who developed mathematical models of the cycles. Because the different cycles possess different periods, the paleoclimatic response to Milankovitch cyclicity has been complicated. Nevertheless, Milankovitch cyclicity has been described from sedimentary rocks of many parts of Earth history. *See* EARTH ROTATION AND ORBITAL MOTION; PRECESSION OF EQUINOXES.

Paleoclimatic history. The paleoclimatic history of the Earth can be divided into several stages, based on events and on the increasing resolution in the geologic record toward the present. *See* GEOLOGIC TIME SCALE.

Pre-Phanerozoic climate. Paleoclimatic history before the appearance of the first shelly fossils is poorly known because many of the strongest types of evidence for paleoclimatic studies do not exist in pre-Phanerozoic rocks. It is clear, however, that the interval was punctuated by extensive glaciations, as indicated by the abundance of tillites. *See* CONGLOMERATE.

Pre-Pangaean Paleozoic. During the early part of the Phanerozoic, land consisted of Gondwana, which would later break up into the modern continents of the Southern Hemisphere and pieces of southern Asia, and numerous small continents, which were to become the modern Northern Hemisphere continents. Sea level was high and temperatures generally warm through much of the interval, although sea level and temperatures were low at the very beginning of the interval and large ice sheets formed near the South Pole at the end of the Ordovician Period and again during the early Carboniferous. Reefs

were widespread during the Ordovician through Devonian, and peat swamps were widespread during the Carboniferous. *See* BOG; CARBONIFEROUS; DEVONIAN; ORDOVICIAN; PEAT; REEF.

Pangaean climate. By the end of the Paleozoic, nearly all the continents had come together to form the supercontinent, Pangaea. At the same time, global climate changed dramatically. Abundant evidence exists for aridity or strongly seasonal rainfall. These changes were brought about by the influence that the large landmass had on atmospheric circulation, creating an intensely seasonal climate. Sea level was very low at this time, increasing the size of the exposed continent and augmenting its climatic influence. Temperature is perhaps most difficult to determine for this time because biogeographic patterns, so useful in the previous interval, are extremely weak, and oxygen isotopic determinations of paleotemperatures are still unreliable. *See* PALEOZOIC.

Late Mesozoic to early Tertiary climate. With the breakup of Pangaea, sea level rose and climate again changed, becoming wetter in many parts of the globe and probably cooler. Leaf-margin analysis is possible from this time forward, because flowering plants had evolved and dispersed across the globe; and relatively reliable and abundant oxygen-isotope analyses also begin to be possible for about the same time. The middle Cretaceous through early Tertiary was warm relative to later paleoclimates, but it was punctuated by cooling events. In addition, coal, phosphate, and marine rocks rich in organic substances were abundant at various times during this interval. Particularly impressive was a short-lived but extremely intense episode of deposition of organic-rich rock worldwide. *See* CRETACEOUS; TERTIARY.

Late Tertiary. Climate cooled substantially in several steps through the latter part of the Tertiary Period (**Fig. 3**). Ice caps at both poles were formed, first in the Southern Hemisphere. Rapid and significant changes in oceanic biogeographic patterns took place, and the foundations of the modern vegetation patterns were established.

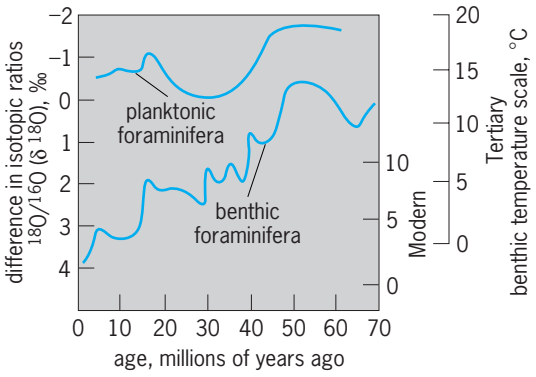

Fig. 3. Temperature curve for the Tertiary Period, based on oxygen isotopes in foraminifera. The modern scale takes into account the effect of ice caps on isotope ratios. °F = (°C × 1. 8) + 32. (*After S. M. Savin, The history of the Earth's temperature during the past 100 million years, Annu. Rev. Earth Planet. Sci., 5:319–355, 1977*)

Quaternary. The Quaternary Period includes the major Northern Hemisphere glaciations of the last 2×10^6 years. Although there is no reason to regard these as different in scale from earlier glacial epochs in Earth history, the proximity of these events to the present day has permitted paleoclimatologists to study them in far greater detail than is possible for the earlier events. Much of the understanding of the dynamics of continental glaciations and paleoclimate has come from studying the global patterns of the Quaternary. *See* CLIMATOLOGY; GLACIAL EPOCH; GLACIOLOGY; QUATERNARY.

Judith Totman Parrish; Eric J. Barron

Bibliography. A. Berger et al. (eds.), *Milankovitch and Climate*, NATO AST Ser., vol. 126, 1984; T. J. Crowley, The geologic record of climatic change, *Rev. Geophys. Space Phys.*, vol. 21, 1983; L. A. Frakes, *Climates Throughout Geologic Time*, 1979; J. E. Kutzbach, Modeling of paleoclimates, *Adv. Geophys.*, vol. 28A, 1985; J. T. Parrish, Climate of the supercontinent Pangae, *J. Geol.*, vol. 101, 1993; P. V. Rich et al., Evidence for low temperatures and biologic diversity in Cretaceous high latitudes of Australia, *Science*, vol. 242, 1988; J. A. Wolfe and G. R. Upchurch, Jr., North American nonmarine climates and vegetation during the Late Cretaceous, *Palaeogeog. Palaeoclim. Palaeoecol.*, vol. 61, 1987.

Paleocopa

Paleocopa, also called Palaeocopida, is an extinct order of the crustacean class Ostracoda. The order is divided into nine superfamilies, one of which, Barychilinacea, is only tentatively assigned to the order. A principal feature of the species in the order is their long, straight hinge that extends along the dorsal margin of the carapace and joins the two valves together (see **illus.**). As is true of most benthic ostracodes, the palaeocopes lack a frontal opening through which to extend their walking legs. Unlike modern ostracodes, however, they have no calcified inner lamella, and their muscle-scar patterns, which are quite useful in the taxonomy of

Paleocopa. (*a*) *Ogmoopsis nodulifera*, left valve; from the Ordovician of Sweden. (*b*) *Eurychilina subradiata*, right valve showing wide complete frill; Ordovician of Minnesota. (*c*) *Sulcicuneus porrectinatium*, male and female carapaces; Devonian of Michigan. (*d*) *Abditoloculina pulchra*, male and female carapaces; Devonian of New York; the bulbous L3 is typical of the family Hollinidae. (*e*) *Dicranella bicornis*, left valve; Ordovician of Minnesota. (*f*) *Sigmoopsis platyceras*, right valves, with histial dimorphism; Ordovician of Estonia. (*g*) *Amphizona asceta*, right valve, a species of the Arcyzonidae, the Devonian forerunners of the Amphissitidae. (*h*) *Amphissites marginiferus*, left valve, Carboniferous of North America. (*i*) *Beyrichia tuberculata*, right valves; Silurian of Europe. (*j*) *Hibbardia lacrimosa*, right valves; Devonian of New York. (*k*) *Hollinella dentata*, juvenile, left valves; Carboniferous of North America. (*l*) *Piretella acmaea*; the carapace of the dimorphic false pouch is formed by incurved, fused, hollow spines; Ordovician of Estonia. (*m*) *Falsipollex laxivelatus*, carapaces; Devonian of Michigan. (*n*) *Tetrasacculus mirabilis*, carapaces; Carboniferous of Illinois. (*o*) *Ctenoloculina cicatricosa*, carapaces; Devonian of Michigan.

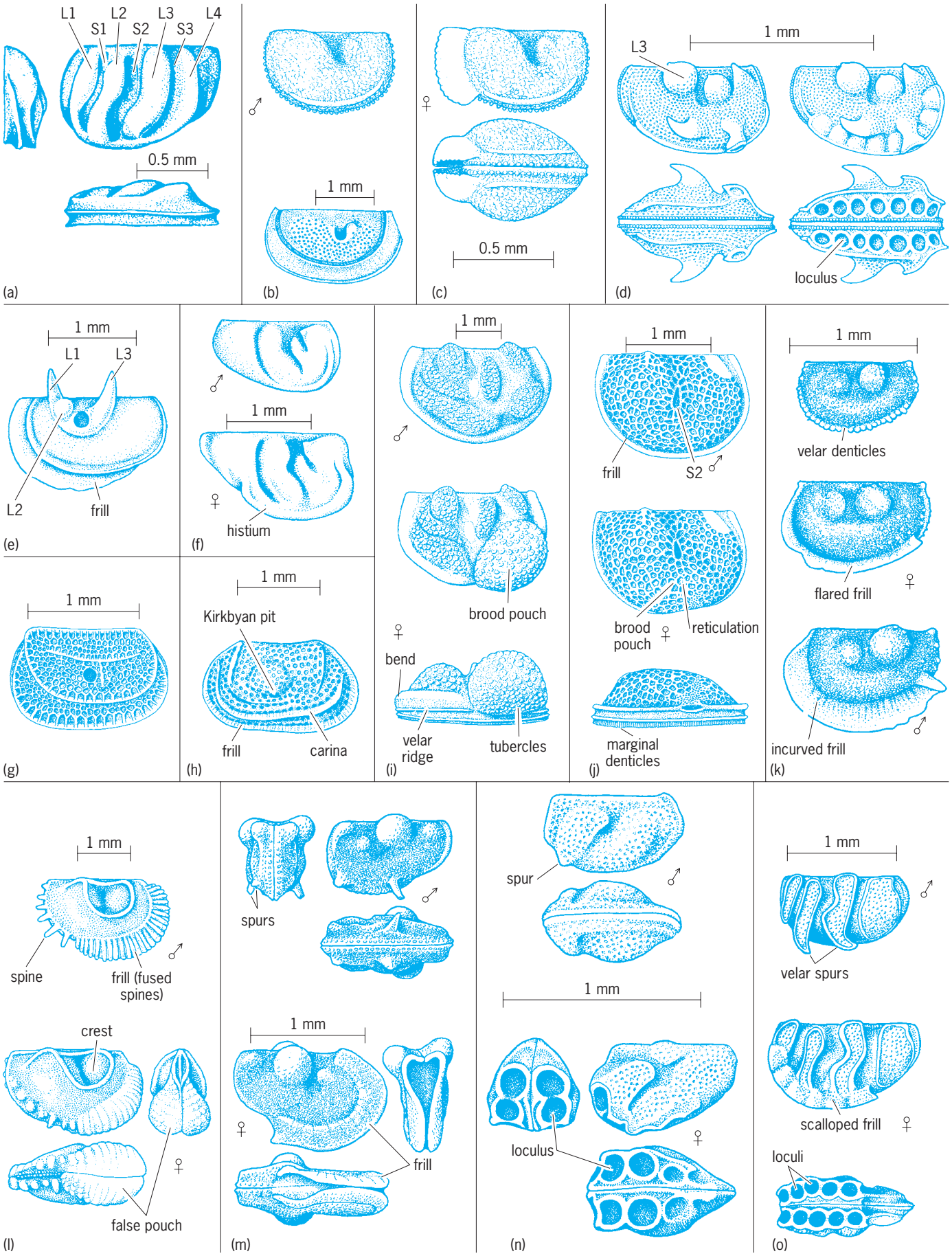
L1
L2
L3
L4
S1
S2
S3
0.5 mm
(a)
♂
1 mm
(b)
♀
0.5 mm
(c)
L3
1 mm
loculus
(d)
1 mm
L1
L3
L2
frill
(e)
♂
1 mm
♀
histium
(f)
1 mm
(g)
1 mm
Kirkbyan pit
frill
carina
(h)
1 mm
♂
brood pouch
♀
bend
velar ridge
tubercles
(i)
1 mm
frill
S2
♂
brood pouch
♀
reticulation
marginal denticles
(j)
1 mm
velar denticles
flared frill
♀
incurred frill
♂
(k)
1 mm
spine
frill (fused spines)
♂
crest
♀
false pouch
(l)
spurs
♂
1 mm
♀
frill
(m)
spur
♂
1 mm
loculus
♀
(n)
1 mm
♂
velar spurs
scalloped frill
♀
loculi
(o)

many other groups of ostracodes, are very poorly known. *See* OSTRACODA.

The carapaces of many palaeocopes are marked by ornamentation in the form of lobes and sulci, which are designated L1, S1, L2, S2, . . . , to L4. Sulcus S2 is the most prominent and is present in all lobate forms. S2 is the external manifestation of the adductor muscle scar, which is located internally in the same spot. Appendages of the palaeocopes are largely unknown except for a few rather poorly preserved specimens.

Sexual dimorphism. In general, Ostracoda are characterized by pronounced sexual dimorphism. The males tend to have more elongated carapaces to accommodate their comparatively gigantic copulatory apparatuses. The carapaces of the females are likely to have a ventral swelling that provides space for carrying the eggs and brooding the young. Not all ostracodes show these differences between the sexes, but most do.

One of the most remarkable morphological features of many species of palaeocope ostracodes is that their sexual dimorphism is quite pronounced and is carried far beyond the rather simple sexual dimorphism that is described above. In typical instances, the males resemble the instars (immature forms), and the females have developed strongly modified morphology that is associated with reproduction—especially the development of pouches for carrying eggs and brooding the young.

The sexual dimorphism of palaeocopes can be of three general kinds: external, internal and distinct, or internal and indistinct. External brood pouches do not open into the interior of the carapace where the animal lives, an area referred to as the domicilium (illus. *b–d*, *k–o*). These external brood pouches are enigmatic because they are positioned on the anteroventral portion of the carapace. In this position it seems that the walking legs would be likely to disrupt the instars in the brood chamber or that the eggs or young might fall out when the valves are opened. Nevertheless, this plan seems to have worked well, for ostracodes with this kind of dimorphism were present for much of the late Paleozoic Era, an interval of several tens of millions of years.

In some instances the brood pouches open directly into the domicilium and are not separated from the domicilium by an internal partition. This indistinct internal dimorphism is shown in illus. *a* and *f*.

By far the most fascinating brood pouches, however, are those that are internal and distinct; that is, they open into the domicilium and are walled off from the domicilium except for a comparatively small opening (illus. *i* and *j*). For a long time the function of the brood pouches was unknown, and which dimorph was the male and which the female was much debated. Careful study, however, has revealed the presence of instars within the brood pouches, and since the time of that discovery the morphs with the pouches have been regarded as females and the others as males.

Given the prominence of sexual dimorphism among some of the palaeocops, it is surprising that some of the superfamilies are not dimorphic at all, apparently not even to the extent described for the ostracodes in general (illus. *g* and *h*). Much more research on the palaeocopes is needed to establish conclusively this surprising evolutionary development and what may have brought it about.

Ontogeny. Arthropods, including the ostracodes, grow by molting, a process that involves shedding the external skeleton, growing quickly, and then secreting a new skeleton. Growth occurs only during molting. Ostracodes are nearly unique among the marine arthropods in having determinate growth. The trilobites and such other crustaceans as the crabs, lobsters, and shrimps continue to grow throughout their lives. The ostracodes, however, grow through only a set, genetically determined number of instars. When an ostracode reaches the adult stage and sexual maturity, it stops growing. The number of growth stages is constant at rather high taxonomic levels—usually at the superfamily level—and varies from seven to nine, or perhaps a few more.

This mode of growth provides the ostracodes and other arthropods with some distinct advantages and disadvantages. During molting, when a palaeocope passes from one instar to another, it is possible for it to undergo a pronounced change in its morphology. Most palaeocopes do not, in fact, alter their morphology radically during ontogeny except for the development of brood pouches by the females, but the potential is there. A disadvantage of growth by molting is that the ostracode is quite vulnerable to predation after it has shed one skeleton and before it has secreted another.

The study of the ontogeny of the palaeocopes presents paleontologists with a unique opportunity to learn about the pathways and mechanisms of evolution. The palaeocopes lend themselves to this sort of study because they molt, have a predetermined number of growth stages, and have pronounced sexual dimorphism that allows one to determine the sex of individuals and the sex ratios of populations. Thus, research on the palaeocopes can reveal the pathways their evolution has followed in a way that is possible for few other kinds of fossil organisms.

When it molts, a palaeocope ostracode roughly doubles its volume before it secretes a new carapace. The cube root of 2 (from doubling) is 1.26. This means that if an ostracode doubles in volume, it should increase in length, height, and breadth by 1.26. This relationship has helped paleontologists determine to which instar an ostracode carapace belongs, although the rule about doubling is only approximate and the exact size can be greatly affected by the environmental conditions.

Paleoecology and biostratigraphy. As is true of other benthic ostracodes, the paleocopes do not have a planktonic larval stage. As a result, they are quite limited in their means of dispersal, and few species are biogeographically widespread. Indeed, most forms are limited primarily to a single basin of deposition. Thus, they are not useful for long-distance, intercontinental correlation, but they are well suited for the study of paleobiogeography. If the sedimentary rocks

in two areas have paleocope faunas that comprise the same species, it is fairly certain that the areas were close enough together, at the time of deposition of the sediments that formed the rocks, to allow the paleocopes to walk from one area to the other or to be transported by fishes or marine currents.

A great deal of work remains to be done to decipher details of the paleoecology of the paleocopes, but they are a promising group for study of organism-sediment interactions because of their small size and intimate association with the sediment substrate. *See* PALEOECOLOGY.

Roger L. Kaesler

Bibliography. R. H. Bate, E. Robinson, and L. M. Sheppard (eds.), *Fossil and Recent Ostracods*, Ellis Horwood Limited, Chichester, 1982; R. L. Kaesler, Superclass Crustacea, in R. S. Boardman, A. H. Cheetham, and A. J. Rowell (eds.), *Fossil Invertebrates*, pp. 241–258, Blackwell Science, Cambridge, MA, 1987; R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, Part Q: *Arthropoda 3, Crustacea, Ostracoda*, Geological Society of America and University of Kansas Press, 1961; R. C. Whatley and C. Maybury, *Ostracoda and Global Events*, Chapman & Hall, New York, 1990; R. C. Whatley, D. J. Siveter, and I. D. Boomer, Arthropoda (Crustacea: Ostracoda), in M. J. Benton (ed.), *The Fossil Record 2*, Chapman & Hall, New York, 1993.

Paleoecology

Ecology of prehistoric times, extending from about 10,000 to about 3.5×10^9 years ago. Although the principles of paleoecology are the same as those underlying modern ecology, the two fields actually differ greatly. Paleoecology is a historical science that must rely on empirical data from fossils and their enclosing sedimentary rocks to make inferences about past conditions. Experimental approaches and direct measurement of environmental parameters, which are critical components of modern ecology, are generally impossible in paleoecology. Furthermore, distortion and loss of information during fossilization means that fossil assemblages and distributions are rarely congruent with living communities. Hence, the resolution of ancient ecosystems must remain relatively imprecise. The lack of precision is compensated for by the fact that paleoecology deals with processes occurring over vast spans of time that are unavailable to modern ecology. Long-term changes in communities (replacement) may be discerned and related to patterns of environmental change. More significantly, overall patterns of ecological change in the global biosphere may be documented; evolutionary paleoecology focuses on recognition and interpretation of long-term ecological trends that have been critical in shaping evolution.

Among the goals of paleoecology are the reconstruction of ancient environments (primarily depositional environments), the inference of modes of life for ancient organisms from fossils, the recognition of recurring groupings of ancient organisms that define relicts of communities (paleocommunities), the reconstruction of the interactions of organisms with their environments and with each other, and the documentation of large-scale and long-term patterns of stasis or change in ecosystems. *See* ECOSYSTEM.

Paleoenvironmental Interpretations

To reconstruct ancient marine environments, many different parameters must be inferred, such as temperature, water salinity, oxygen levels, nutrient concentrations, and water movements and depth (**Fig. 1**). In this regard, paleoecology interfaces directly with the fields of sedimentology and stratigraphy, including study of modern depositional environments. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; STRATIGRAPHY.

Taxonomic uniformitarianism. One of the most useful, but also potentially misused, aspects of paleoecological application is known as taxonomic uniformitarianism. This concept relies on studies of modern organisms to determine limiting environmental factors, such as salinity tolerance, temperature preference, or depth ranges. Fossils of the same or closely related species are then inferred to have had similar environmental preferences, and their occurrence is judged to indicate that particular strata were deposited under a comparable range of environmental conditions. Such an approach is valid for very closely related organisms in relatively recent geologic time. Species and even genera may have relatively uniform environmental ranges through time, but the same cannot always be said of higher taxa such as families. At the level of order or class, only the broadest uniformitarian generalizations apply. For example, it is probably valid to consider fossil nautiloids or echinoderms as indicators of normal marine salinities, as all living representatives of these taxa have very limited abilities for osmotic regulation and therefore are restricted to near-normal salinity. Similarly, the restriction of photosynthetic organisms (such as algae) to the euphotic zone may be useful in determining relative depth. However, the precision and reliability of taxonomic uniformitarianism breaks down in increasingly ancient samples.

Morphologic features. Certain features of the morphology of fossils may be useful in making environmental inferences without reliance on evolutionary relationships. For example, the presence of entire margins and drip tips on leaves of plants is indirectly related to humid, warm climates, and so the proportions of leaf floras with entire margins and drip tips have been used as an index of paleolatitudinal zonation. Growth forms of colonial organisms relate to environmental factors such as turbulence and sedimentation rates. Flexible or articulated skeletons or flat encrusting form in colonial marine animals are associated with highly turbulent shallow-water environments where streamlining becomes important. Delicately branched, inflexible colonies typify quiet areas and, commonly, areas of high turbidity where a branching skeleton may shed sediment more readily than a flat or globose form. Such ecologically related morphology may transcend taxonomic boundaries.

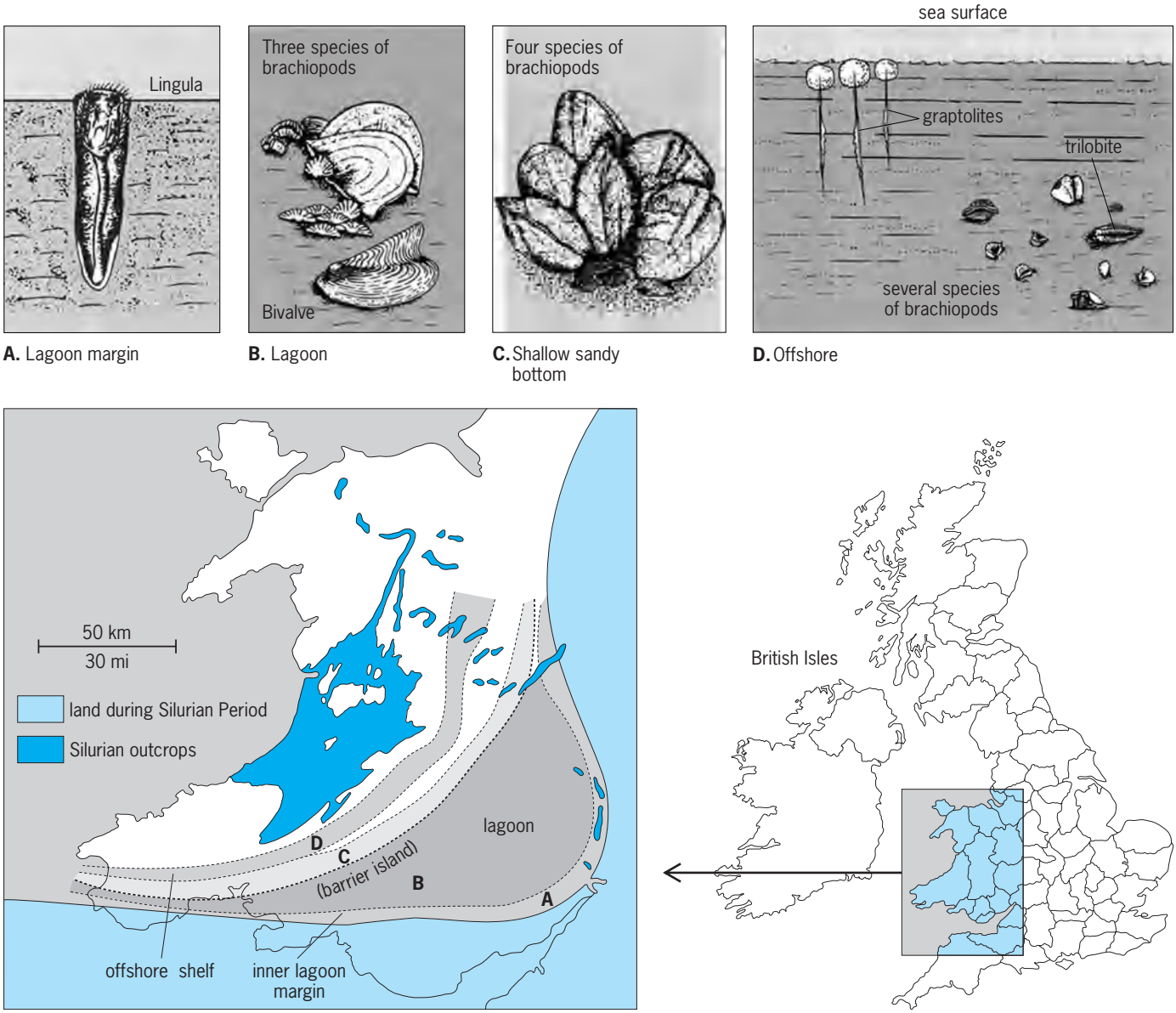

Fig. 1. Zonation of Silurian (425 million year old) marine communities in the Welsh Basin, United Kingdom. Marine animal communities were typically arranged in belts parallel to shoreline and related to water depth. A classic study using fossil communities in basin analysis (Ziegler, 1965) mapped out the distribution of different communities of marine fossils, primarily brachiopods, to show the contouring of belts of ancient environments from the shoreline in the southeast to deep basinal environments. (*Reprinted with permission from S. Stanley, System History, 2d ed., Freeman Co., 2005*)

Skeletal mineralogy and geochemistry. The microstructure and geochemistry of organism skeletons may provide clues about ancient environments. For example, the presence of growth banding in skeletons provides evidence for seasonal variability in climates. The skeletons of fossil organisms, if they are well preserved, also encode valuable environmental information in the form of trace elements and isotopic signatures. For example, the calcium carbonate skeletons of marine invertebrates incorporate trace elements whose proportion is related both to physiology and environmental factors such as temperature and salinity. The isotopic composition of oxygen or carbon within carbonate skeletons is a function of isotopic composition of the seawater in which the skeleton was secreted as well as of water temperature. If temperature can be determined independently, the ratio of ^{18}O to ^{16}O (often expressed as a deviation from a standard and referred to as $\delta^{18}O$) can be used to determine whether a shell was secreted in water of normal (35%) or abnormal salinity. Conversely, if a given shell can be assumed, on independent evidence, to have come from a normal marine environment and is unaltered, then $\delta^{18}O$ may be used to determine paleotemperature. In general, carbonate secreted at lower temperatures is preferentially enriched with respect to ^{18}O, and so, $\delta^{18}O$ is useful for temperature determination. *See* GEOLOGIC THERMOMETRY.

Comparative taphonomy. Taphonomy, which deals with processes and patterns of fossil preservation, has a critical dual role with respect to paleoecology.

On the one hand, preservational processes impose distinct biases on the fossil record that must be considered carefully in any attempt at paleoecological reconstruction. On the other hand, the bodies and skeletons of dead organisms constitute biologically standardized sedimentary particles whose orientations, sorting, and general preservational condition bear the imprint of environmental processes active in the depositional environment. Comparative taphonomy uses the differential preservation of fossils as a source of paleoenvironmental information. The degree of preservation of fossils reflects biostratinomic processes, such as current-wave transport, decay, disarticulation, fragmentation and corrasion of skeletons, and fossil diagenetic factors acting after final entombment of the remains in sediment.

Evidence of mode of death of organisms may also provide critical details. For example, layers of beautifully preserved fish or reptile carcasses signify mass mortalities that involved changes in the water column itself. But such mass mortalities can be recorded only if they were also timed with burial events.

Soft tissues can be preserved only by exceptionally rapid burial in anoxic sediments followed by very early coating or impregnation by minerals. Such deposits not only yield important data on the paleobiology of organisms but also provide detailed insights into depositional environments.

Usually, however, only skeletal remains are preserved. Skeletons composed of bivalved shells (for example, brachiopods and pelecypods) or, particularly, of multiple articulated elements (for example, echinoderms, arthropods, and vertebrates) are sensitive indicators of episodic burial rates. Experimental studies have demonstrated the rapidity of disarticulation under normal marine conditions; most starfish, for example, disintegrate into ossicles in a few days. Hence, intact preservation of these organisms signals episodic burial events. *See* FOSSIL.

Individual skeletons, or parts of skeletons, may become physically fragmented, chipped, or abraded. Such evidence reflects the general degree of turbulence of a particular depositional environment. Similarly, the degree to which skeletal remains are size-or shape-sorted may signify the extent of current and wave processing. Skeletal destruction by bioerosion, physical abrasion, or chemical solution is generally a good indicator of residence time of skeletons on the sea floor prior to burial. The orientation of fossils may yield specific clues as to the extent and types of environmental energy. Pavements of convex-upward valves typically are associated with persistent current reworking, whereas abundant concave-upward shells may signify an episode of stirring of the shells from the sea bottom and resettlement during storms. Furthermore, alignment of elongated shells may provide data on the orientation of unidirectional currents or the propagation direction of oscillatory waves. Vertically embedded specimens of ammonoids are typical of water areas with depths less than 30 ft (10 m). Finally, the early diagenetic features of fossils reflected in solution, compaction, and mineralization may yield information about sediments and bottom water geochemistry, water pH and oxygen content.

Various aspects of biostratinomic and diagenetic fossil preservation can be combined to form predictive models of taphonomic facies or taphofacies. Certain suites of quantifiable preservational conditions, for example, characterize particular environments, and so their recognition by paleoecologists may help to "fingerprint" those environments. *See* SEDIMENTOLOGY; TAPHONOMY; TRACE FOSSILS.

Paleoautecology

Paleoautecology, the interpretation of modes of life (broadly, niches) of ancient organisms, involves a multidisciplinary approach. Although ancient modes of life cannot be determined completely, paleoecologists can often assign fossils to generalized guilds in terms of types of feeding, substrate preference, and degree of activity.

A thorough understanding of the biology of closest modern analogs is particularly important in any attempt to reconstruct paleoautecology. If the species or a closely related species is extant, then its mode of life, general physiology, and even behavior can be inferred with some confidence through the use of taxonomic uniformitarianism, provided that the biology of living relatives is well understood. "Living fossils," or relict extant taxa, such as *Nautilus*, sclerosponges, horseshoe crabs, and modern stalked crinoids provide valuable clues for interpreting the paleobiology of extinct organisms. *See* LIVING FOSSILS.

Functional morphology. For extinct organisms that have no adequate modern analogs, alternative approaches, particularly functional morphology, provide some hints as to life modes. Comparative morphology seeks analogies between the anatomical features of fossil skeletons and those in living forms for which the function can be determined. In some cases, structures in unrelated organisms have evolved convergently, and their function may be interpreted by analogy. When no biological analog exists, a physical or mechanical model, or paradigm, may provide clues to interpreting structures in extinct organisms.

An experimental approach to functional morphology may also provide useful insights. Models of ammonite shells, for example, have been tested in flumes in which artificial currents are produced to determine frictional drag effects of shell shape and sculpture. Certain shell shapes were found to be more hydrodynamically streamlined and probably correspond to more rapidly swimming modes of life. Testing the resistance of different brachiopod shell architectures to crushing, as by predators, has shown that certain features of shell architecture, such as ribbing and deflections of shell margins, can increase shell rigidity.

Fossil data. Certain natural experiments also shed light on the paleobiology of extinct organisms. The fact that oysters encrusted the shells of living ammonites has enabled paleontologists to calculate the buoyancy compensation capabilities of those

ammonoids. Encrustation and boring of cephalopod shells by bryozoans and barnacles that grew preferentially aligned toward currents has demonstrated a predominance of forward swimming motion in these extinct hosts.

Remnants of soft parts, muscle scars, gut contents, and associated trace fossils all provide information useful in the reconstruction of ancient ecological niches. Rare occurrences of rapidly buried fossils in unusual positions can be interpreted as original life positions. Unusual associations with substrates or other organisms also provide insights. Finally, the consistent association of poorly understood fossil species with other fossils whose modes of life are well known or with sediments that indicate particular environments may help to establish the habits and environmental ranges of extinct forms.

Population studies. Certain properties of species, such as mortality patterns, birth rates, and numbers of individuals per age class, can be studied only at the level of populations. Despite the difficulties of studying fossil populations, it is still possible to make some inferences about population parameters. For example, the distribution of individuals of a particular species into different age or size classes may yield some indirect data on the age-frequency distribution of a population that can be used to construct crude mortality curves showing age-at-death relationships. Some species may display a high juvenile mortality, a feature typically associated with stressed environments and rather opportunistic species; others, in stable environments, may display delayed mortality.

Of particular ecological importance is the population strategy of a given species of organisms. Two end-member conditions have been recognized: opportunistic species, sometimes termed r-selected forms, and equilibrium, or k-selected, species. Opportunistic organisms are typically rather generalized in habit and habitat preferences, are commonly stress-adapted, and display exceedingly high rates of reproductive maturation and fecundity. Extremely opportunistic, or "weedy," species of this sort are recognizable in the fossil record by their widespread distribution and occasional presence in extremely dense, monospecific populations on single beds of rock that are otherwise barren of fossils. Equilibrium species, on the other hand, tend to occur in moderate or small numbers in a narrow range of environments commonly associated with diverse assemblages of other species, such as in reef environments. The distinction between equilibrium and opportunistic mode of life may have important implications for understanding the distribution and evolutionary patterns of fossil taxa as well as for interpreting the stability of particular ancient environments. *See* POPULATION ECOLOGY.

Paleosynecology

The study of interrelationships within organism communities that coexist in time and space is known as synecology. At the most basic level of synecology are the interacting pairs of organisms that coexist in a particular environment. Paleosynecology also involves study of ancient community structure and dynamics.

Organism interactions. Interactions range from tolerance to symbiosis, which involves highly dependent and coevolved species pairs. Although interactions are very difficult to determine with fossils, there are cases where strong clues are observed. In some instances, organism interaction may be very indirect. For example, the accumulation of shells on a sea floor may lead to colonization by other organisms that require hard substrates to encrust onto or bore into, a process referred to as positive taphonomic feedback. Conversely, armoring of muddy sea bottoms with shell debris will inhibit burrowing organisms (negative taphonomic feedback).

Shallow-burrowing nuculid bivalves convert muddy sea bottoms into a water-rich pelleted floc. The high turbidity and instability of this fluid substrate may inhibit the settlement of many epifaunal suspension-feeding organisms. Such negative feedback is referred to as trophic group amensalism.

Mutualism, involving symbiotic algae called zooxanthellae, is inferred for many fossil reef-dwelling organisms, including various corals, sponges, and even some bivalves, based on taxonomic uniformitarianism as well as morphology and evidence for prolific skeletal growth. Such mutualism is difficult to substantiate for extinct groups, although distinctive patterns of carbon isotopes within skeletal carbonates may prove a useful fingerprint of secretion aided by zooxanthellae.

Many marine organisms use the skeletons of other living organisms as substrate or to obtain an elevated feeding position without having any effect on the hosts. Evidence of this type of commensal interaction is abundant in the fossil record.

Parasitic interactions are very difficult to observe in fossils because they normally involved the soft tissues of the host. However, rare evidence for paleopathology (fossil diseases) can be documented from Paleozoic times onward in certain organisms such as echinoderms or vertebrates that have an internal skeleton, or endoskeleton. For example, malformations in fossil crinoids may record parasitism.

Fossil evidence for competition is best seen in cases of spatial competition. For example, certain types of bryozoans appear to overgrow other species preferentially. Many aspects of fossil distribution have been attributed to the effects of competition or the evolutionary response for reducing competition by niche partitioning. Examples include the subdivision of many marine communities into distinct feeding groups based on vertical height (tiering) above and below the sediment-water interface. Some researchers have claimed that competition is a primary motor of evolutionary change, often alluding to Darwin's analogy of the wedge, in which more and more species are packed into a particular ecospace by increasingly finely divided specialized niches.

Predation or carnivory is probably one of the most significant ecological interactions in any environment. Direct predator-prey links are difficult to

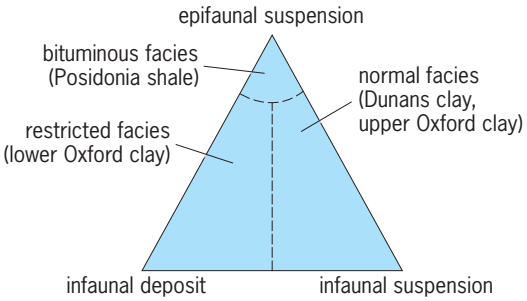

Fig. 2. Classification of shales in terms of percentages of three benthic habitat groups: epifaunal suspension-feeding bivalves (those living at the sediment-water interface and filtering seawater), infaunal (living within and feeding on the sediment), and suspension-feeding groups. (*After K. A. Morris, Comparison of major sequences of organic-rich mud deposition in the British Jurassic, J. Geol. Soc. London, 137:157–170, 1980*)

establish in the fossil record, but there are several lines of evidence that may be used. Bite marks of particular types, such as tooth marks of mosasaurs on ammonoids, provide one line of evidence, so do boreholes of predatory snails on particular prey species and remnants of prey shells preserved in the stomach contents or coprolites (fossilized feces). In turn, numerous morphological trends may signify antipredatory adaptations. The fossil record of predation extends back to the Early Cambrian, as evidenced by bite marks in trilobites, which commonly show a preference for the right side of the prey. The frequency of healed and unhealed predatory fractures in some shells increases significantly in the Paleozoic in concert with the rise of fossil evidence for shell-crushing predators. The earliest shell-drilling snails appear to have been Ordivician in age, but the habit of drilling shells for predation probably evolved independently at least four times in different groups of gastropods.

Paleocommunities. The fossil record contains highly biased remnants of past communities or paleocommunities. Paleocommunities are generally recognized as recurring associations of fossil species. Multivariate statistical techniques such as cluster analysis and ordination analysis are commonly employed to aid in discerning the recurrent groupings of fossil species, or persistent gradients of species composition. Such analyses are based upon field studies in which data on the presence, absence, or

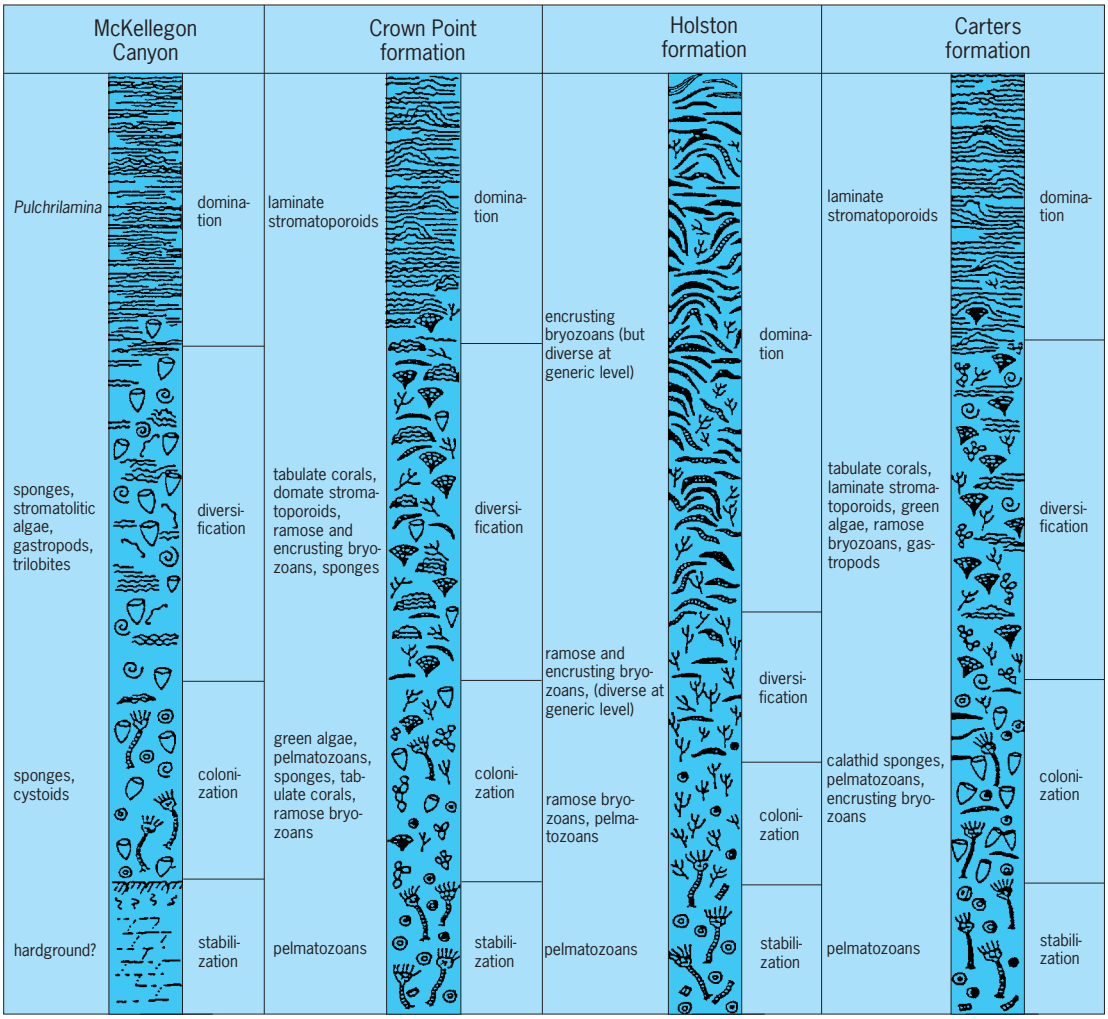

Fig. 3. Comparison of four developmental stages in four ancient reef masses. (*After K. R. Walker and L. P. Alberstadt, Ecological succession as an aspect of structure in fossil communities, Paleobiology, 3:238–257, 1975*)

relative abundance of fossil taxa have been recorded in a large number of samples, typically from many stratigraphic levels.

Taphonomic biases. Statistically defined groupings may or may not represent real ecological entities. For example, in most offshore marine environments, the transport of skeletons between environments is minimal. However, because of differential preservation, the proportions of organisms in the living assemblages (biocoenoses) are not always faithfully reproduced in the death assemblages of skeletal remains (taphocoenoses). Nearly all soft-bodied organisms are lacking in the death assemblage, and those with fragile skeletons tend to be underrepresented. Moreover, because of the accumulation of skeletons over extended periods of time, death assemblages commonly display mixtures of organisms that inhabited slightly differing environments at different times, a phenomenon referred to as time averaging. Fossil assemblages actually may be more diverse than living assemblages of skeletonized organisms at any one time. They record a very biased and averaged-out view of communities that existed over a long period of time.

Relationship to environments. In most studies of paleocommunities, recurrent groups can be related to environments, as inferred from independent evidence such as rock type, sedimentary structures, taphonomy, trace elements, and isotopic studies. Classic studies modeled paleocommunity distribution patterns on relative bathymetry or distance from shoreline, but many later studies emphasized the control of paleocommunity distribution by multiple factors. Depth-related factors such as turbulence, light penetration, and oxygen level are clearly important controls in many cases. However, sedimentation-related factors such as rates of deposition, turbidity, and substrate consistency may be equally important, giving rise to a much more complex array of paleocommunities (Fig. 1).

After recurring associations of fossils are recognized, they are generally analyzed in terms of organism interactions and trophic (feeding) relationships. Primary producers of ancient ecosystems, such as algae, are likely to be poorly preserved or absent from the fossil record, but some links in ancient food chains may be recognizable. One aspect of paleocommunity structure that is commonly analyzed is the proportion of different feeding (trophic) and life-habit guilds. Certain marine paleocommunities are dominated by skeletons of infaunal deposit feeders, others by epifaunal suspension feeders. Unfortunately, live-dead comparisons in modern communities suggest that the original proportions of various life habitat groups are not preserved in the fossil record. But the biased trophic proportions of the taphocoenoses may still relate in a meaningful way to the original environment. Consistent differences in the proportions of infaunal suspension-feeding, infaunal deposit-feeding, and epifaunal suspension-feeding bivalves have been detected in differing ancient oxygen-restricted facies (**Fig. 2**). *See* FOOD WEB.

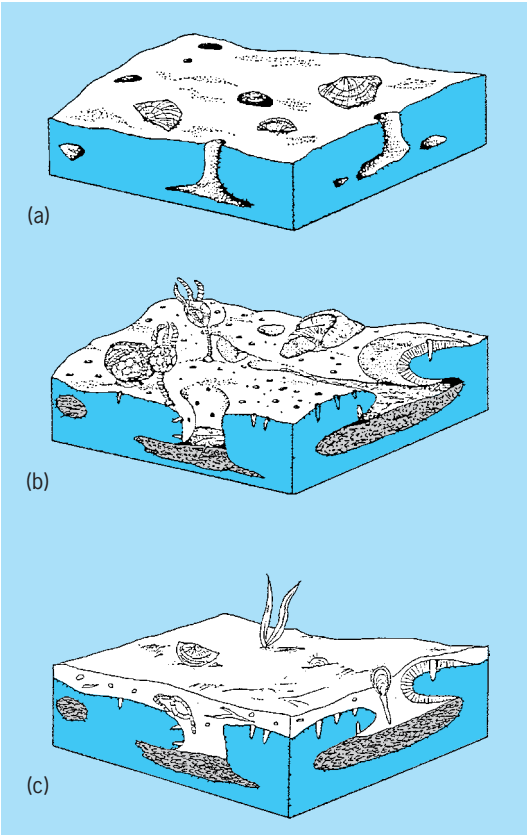

Fig. 4. Formation and burial sequence of Bobcaygeon hardground. (*a*) Soft-bottom community of strophomenid brachiopods and infaunal burrowers inhabiting carbonate mud. (*b*) Hardground community consisting of boring and encrusting organisms. (*c*) Post-hardground community inhabiting muds that blanketed the hardground. (*After C. E. Brett and W. D. Liddell, Preservation and paleoecology of a Middle Ordovician hardground community, Paleobiology, 4:329–348, 1978*)

Temporal changes. Communities and paleocommunities are not static entities in time, but undergo important structural changes on at least three different time scales: succession, replacement, and evolution. Because it operates on a very short time scale, from decades to centuries, ecological succession can be resolved only in a few fossil samples (**Fig. 3**). Some instances of supposed ecological succession, such as encrusting communities upon shells, may in fact reflect taphonomic feedback. Allogenic succession represents changes in communities induced by physical environmental change. Good examples are seen in many hardgrounds, areas of early lithified sea floors (**Fig. 4**). *See* ECOLOGICAL SUCCESSION.

Longer-term changes in community composition, encompassing thousands of years, are not truly succession, but instead record allogenic effects such as sea level or climate variations. These changes are properly termed community replacement, and involve wholesale migration or restructuring of communities at particular locations due to changing environments. In many instances, particular fossil assemblages appear to track shifts in preferred environments and facies within sedimentary cycles. Habitat tracking may provide important clues to deciphering patterns of environmental fluctuations

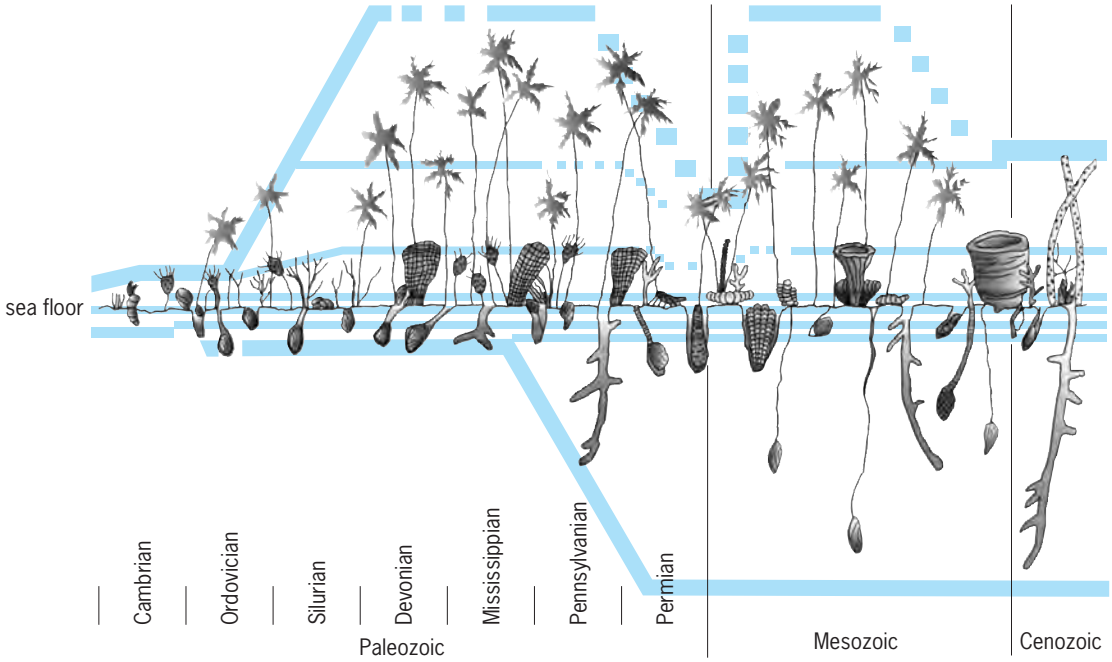

Fig. 5. Tiering or vertical depth stratification in marine communities through time. Note the rapid rise of the highest-tier (level) organisms (crinoids or sea lilies, bryozoans) from a few centimeters to over a meter above the sea floor during the Ordovician Period. Branching bryozoans and shorter-stemmed crinoids took advantage of an intermediate tier 10–20 cm above the sea bottom, while burrowing clams and worms dug down to a tier about the same distance into the sediment. In the middle Paleozoic, still much deeper burrowing forms evolved the ability to mine sediments down to nearly a meter. Note minor readjustment of the tiers associated with mass extinctions at the Paleozoic-Mesozoic and Mesozoic-Cenozoic era boundaries. (*Modified from W. I. Ausich and D. J. Bottjer, J. Geol. Educ., 1991*)

such as transgressive-regressive cycles. On a scale of millions of years, communities show evolutionary changes because their component species have evolved. *See* ECOLOGICAL COMMUNITIES.

Larval ecology and evolution. Ecological patterns such as larval type affect overall patterns in life history. Larval ecology of marine animals controls their geographic distribution. Species with long-lived larvae may be dispersed much more widely than forms with short-lived planktonic phases or direct development from eggs. In turn, geographic distribution, whether localized or cosmopolitan (global), undoubtedly plays an important role in their tendency toward speciation as well as extinction. Thus, it may be possible to develop models to better explain evolutionary patterns in different groups in relation to paleoecology.

Evolutionary Paleoecology

Organisms evolve within the context of other organisms, not in a vacuum. There is substantial fossil evidence to indicate increasing complexity of organism interactions through time. This escalation in the intensity of predatory interactions, for example, may have important implications for evolutionary change. For example, trends of increased spinosity, greater shell thickness, increasingly restricted apertures, and other antipredation adaptations may reflect the intensification of predatory behavior by shell-boring and crushing predators.

Increased vertical stratification or tiering in marine-level bottom communities through time has been recognized (**Fig. 5**). Cambrian communities possessed mainly low-lying suspension-feeding and scavenging organisms that lived mostly just above or below the sediment-water interface. By mid-Paleozoic time, crinoids extended up to heights of several feet or more off the sea floor, and various burrowers extended downward a couple of feet or more into the sediment. The Mesozoic rise of deep-burrowing clams and other infauna increased the infaunal tier to over 3 ft (1 m). The increased vertical structuring of these communities may represent a response to increasing crowding. By feeding at different levels in the water and substrate, organisms were able to further subdivide the resources of a given environment.

Marine animals form a hierarchy of ecological units through the Phanerozoic time interval. These range from blocks of relative stability at time scales of a few million years, to broader intervals of general stability of faunas, to three great evolutionary faunas.

First, at a scale of a few million years, groups of species may show considerable ecological stability punctuated by episodes of abrupt change. Brett and Baird (1995) introduced the concept of "coordinated stasis" to describe a pattern of approximately concurrent long-term stability and abrupt change in many taxa. During a large proportion of geologic time a majority of genera and, in some cases, species show little or no change in morphology. Moreover, general groups of communities or "biofacies" also may be similar throughout blocks of stability referred to as "ecological-evolutionary units and subunits." These

relatively stable intervals, spanning up to several million years, are punctuated by much shorter intervals, perhaps a few tens of thousands of years, of abrupt change across many biofacies, local extinction of many long-standing lineages, immigration and emigration from the local basin, and general faunal turnover. The original example of the Silurian–Devonian (380 to 440 million year old) fossil assemblages of eastern North America—in particular, the Middle Devonian Hamilton Group—features examples of assemblages, separated by up to 5 million years, with nearly identical composition and similar guild structure and even relative abundance. Consideration of a larger number of case studies ranging in age from Cambrian to modern suggests that this original example represents one end member in an array of conditions ranging from similar cases but some with somewhat more species level variability, to examples of nearly continual change in species composition, and biofacies ecological structure. This variability probably depends on local environmental variability. The observed pattern of similarities between samples from cases of coordinated stasis could imply a form of stable, lock-step tracking of certain well-organized "communities." However, this pattern could also be the result of recurrence of a similar assemblage due to persistence of environmental gradients and because species do not drastically change their habitat preferences through time. The retention of habitat preferences by species is perhaps the most important aspect of ecological stasis. It would appear that under appropriate conditions species can simply track shifting preferred environments for millions of years rather than adapt to local change.

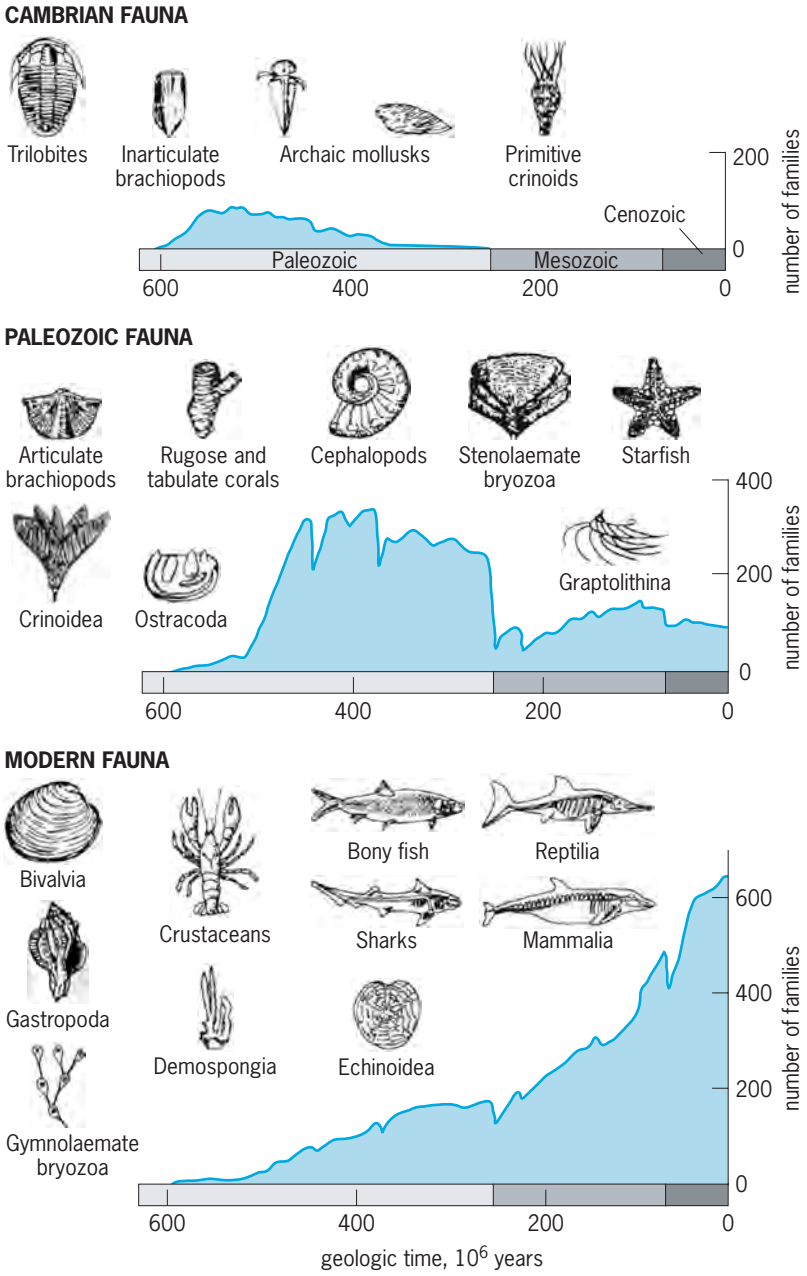

Fig. 6. The three great marine evolutionary faunas of the Phanerozoic. The Cambrian fauna, composed of trilobites, primitive groups of brachiopods, and mollusks arose early during that period to a diversity of about 50 families, then dwindled during the later Paleozoic as more archaic groups migrated offshore and were replaced by the "Paleozoic fauna," typified by rugose and tabulate corals, brachiopods, bryozoans, crinoids, and graptolites. The latter diversified rapidly in the Ordovician Period to over 300 families and then fluctuated around this level until the end of the Paleozoic Era. The great Permian-Triassic extinction reduced the "Paleozoic fauna" and may have favored the rise of the "Modern fauna" during the Mesozoic and Cenozoic eras with diversities as high as 620 families, including especially mollusks, crustaceans, and both sharks and bony fishes. (*Modified from J. J. Sepkoski, Jr., Paleobiology, 1981*)

At the next larger level, marine communities appear to show strong similarities of family- and genus-level composition, as well as general ecological structure (guilds, trophic structure, diversity, and so on) for tens of millions of years. These blocks of relative stability, termed ecological-evolutionary units (EEUs) by Boucot (1990), were terminated by major extinctions. Raup and Sepkoski also recognized five major mass extinctions—the "Big Five" (Late Ordovician, Late Devonian, Permian-Triassic, Late Triassic, and Cretaceous-Tertiary) that stand out from background rates of extinction. These and lesser mass extinctions played critical roles in restructuring the ecology of the biosphere, including changes in guild structure and tiering patterns. Ecological-evolutionary units are bounded by major biotic turnover events, involving widespread extinctions including the "Big Five" mass extinctions. Again, the EEU concept implies that the ecological history of life was not one of continuous, gradual change. Rather, it was characterized by extended periods of near equilibrium that were interrupted by much shorter periods of crisis and major ecological restructuring.

The largest scale of faunal pattern consists of "evolutionary faunas." By analyzing patterns of marine family and genus level diversity using a large database, Sepkoski (1981) recognized three such units through the past 540 million years of the Phanerozoic Eon, each characterized by a different pattern or trajectory of diversification (**Fig. 6**). The "Cambrian fauna"—typified by trilobites, lingulid brachiopods, and certain primitive groups of mollusks and echinoderms—appeared during the earliest Paleozoic, diversified in the Cambrian, and then began to decline as the second or "Paleozoic fauna" diversified. The latter was characterized by

rugose and tabulate corals, articulate brachiopods, bryozoans, and crinoids which formed the major faunas of shallow seas from the Ordovician to the Permian Period and displayed a relatively stable "platform" of family diversity. Finally, the "Modern fauna," characterized by mollusks and crustaceans, arose in nearshore environments during the early Paleozoic, but expanded greatly following the end Permian mass extinctions. Sepkoski and Sheehan (1983) recognized that evolutionary innovations tended to arise first in shallow, nearshore environments. Through time the newly arising groups typical of the "Paleozoic" and then the "Modern" faunas tended to spread offshore, while more archaic forms were displaced to deep ocean "refugia." This is one of the most profound of all paleoecological patterns, and the explanation of this pattern remains imperfectly understood. It may imply that stressed nearshore settings favor evolution of new life strategies and/or that there has been a general intensification of energy utilization through time such that more archaic "Cambrian" or "Paleozoic" faunas were relatively "low energy" and had less competitive ability than physiologically more sophisticated, "high-energy" Modern faunas. *See* ECOLOGY; PALEOCLIMATOLOGY; PALEOGEOGRAPHY; PALEONTOLOGY. Carlton E. Brett

Bibliography. A. J. Boucot, *Principles of Benthic Marine Paleoecology*, 1981; R. J. Dodd and R. J. Stanton, Jr., *Paleoecology, Concepts and Applications*, 2d ed., 1990; J. J. Sepkoski, Jr., A factor analytic description of the Phanerozoic marine fossil record, *Paleobiology*, 7:36–53, 1981; M. J. P. Tevesz and P. L. McCall (eds.), *Biotic Interactions in Recent and Fossil Benthic Communities*, 1983; G. J. Vermeij, *Evolution and Escalation: An Ecological History of Life*, 1987.

Paleogeography

The geography of the ancient past. Paleogeographers study the changing positions of the continents and the ancient extent of land, mountains, and shallow-sea and deep-ocean basins. The Earth's geography changes because its surface is in constant motion due to plate tectonics. The continents move at rates of 2–10 cm/yr (0.75–4 in./yr). Though this may seem slow, over millions of years continents can travel across the globe. As the continents move, new ocean basins form, mountains rise and erode, and sea level rises and falls. The best way to illustrate these changes is through a series of paleogeographic maps. *See* CONTINENTS, EVOLUTION OF; GEOGRAPHY; PLATE TECTONICS.

Paleogeographic maps are necessary in order to understand global climatic change, migration routes, oceanic circulation, mountain building, and the formation of many of the Earth's natural resources, including oil and gas. The maps in this article show the ancient mountains (dark tint), land areas (medium tint), and shallow seas (light tint), as well as the present-day coastline for reference. Past plate boundaries are also shown. The jagged lines in the center of ocean basins are mid-ocean ridges, the curving lines along the edges of continents are deep-sea trenches; the triangular teeth point in the direction of subduction. *See* BASIN; MID-OCEANIC RIDGE; SUBDUCTION ZONES.

Late Precambrian, 650 Ma. In the late Precambrian the continents were colliding to form supercontinents, and the Earth was locked in a major ice age (**Fig. 1**). About 1100 million years ago (Ma), the supercontinent of Rodinia was assembled. Though Rodinia's exact size and configuration are not known, it appears that North America formed the core of this supercontinent. At that time, the east coast of North America was adjacent to western South America, and the west coast of North America lay next to Australia and Antarctica. Rodinia split into halves approximately 750 Ma, opening the Panthalassic Ocean. North America rotated southward toward the South Pole. The other half of Rodinia, composed primarily of Antarctica, Australia, India, Arabia, and China, rotated counterclockwise, northward across the North Pole. Between these two halves lay the Congo continent, made up of much of north-central Africa. The oceans between these three continents were completely subducted by the end of the Precambrian, and the three parts of Rodinia came together to form the supercontinent of Gondwana(land). This major continent-continent collision is known as the Pan-African orogeny. *See* OROGENY; PRECAMBRIAN; PROTEROZOIC; SUPERCONTINENT.

The climate was cold during the late Precambrian. Evidence of glaciation is found on nearly every continent. Why cold conditions were so widespread during the late Precambrian has long puzzled geologists. Some scientists believe the Earth was completely frozen like a snowball. However, as Fig. 1 shows, many continents lay close to the North and South poles at that time. *See* CLIMATE HISTORY; GLACIAL EPOCH; GLACIOLOGY; PALEOCLIMATOLOGY.

Early and middle Paleozoic, 545–360 Ma. The supercontinent that formed at the end of the Precambrian Era, approximately 600 Ma, broke apart at the beginning of the Paleozoic Era (**Fig. 2**). A new ocean, the Iapetus, widened between the ancient continents of Laurentia (North America), Baltica (northern Europe), and Siberia. Gondwana, which was considerably larger than any of the other continents, stretched from the Equator to the South Pole. *See* ORDOVICIAN; PALEOZOIC.

Approximately 400 Ma, the Iapetus Ocean closed, crashing Laurentia and Baltica together. This continental collision, preceded in many places by the collision of marginal island arcs in Maritime Canada and New England, resulted in the formation of the Caledonide Mountains in Scandinavia, northern Great Britain, and Greenland, and the Northern Appalachian Mountains along the eastern seaboard of North America. It is also likely that by middle Paleozoic times North China and South China had rifted away from the Indo-Australian margin of Gondwana, and were headed northward across the Paleo-Tethys Ocean (**Fig. 3**). *See* CONTINENTAL DRIFT.

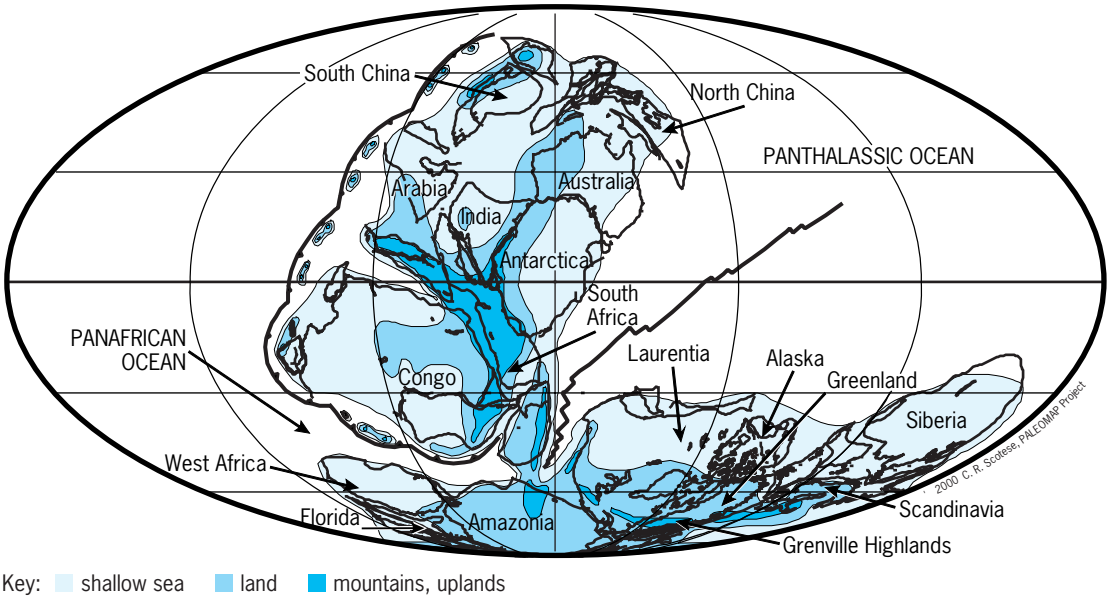

Fig. 1. Vendian (or Ediacaran), 650 Ma. (*Paleogeographic maps by Christoper R. Scotese, PALEOMAP Project, University of Texas at Arlington*)

Late Paleozoic, 360–245 Ma. At the Paleozoic Era, the continents collided to form the supercontinent of Pangea (Fig. 3). Centered on the Equator, Pangea stretched from the South Pole to the North Pole, and separated the Paleo-Tethys Ocean to the east from the Panthalassic Ocean to the west. Though the supercontinent that formed at the end of the Paleozoic Era is called Pangea (literally, "all land"), this supercontinent probably did not include all the landmasses that existed at that time. In the Eastern Hemiphere, on either side of the Paleo-Tethys Ocean, there were continents that were separated from Pangea. These continents were North and South China, and a long, narrow continent known as Cimmeria. Cimmeria consisted of parts of Turkey, Iran, Afghanistan, Tibet, Indochina, and Malaya. It appears to have rifted away from the Indo-Australian margin of Gondwana during the late Carboniferous–early Permian. Together with the Chinese continents, Cimmeria moved northward toward Eurasia, ultimately colliding along the southern margin of Siberia during the Late Triassic. *See* PERMIAN.

Early Mesozoic, 245–144 Ma. The supercontinent of Pangea did not rift apart all at once, but in three main episodes. The first episode of rifting began in the Middle Jurassic, about 180 Ma when North America

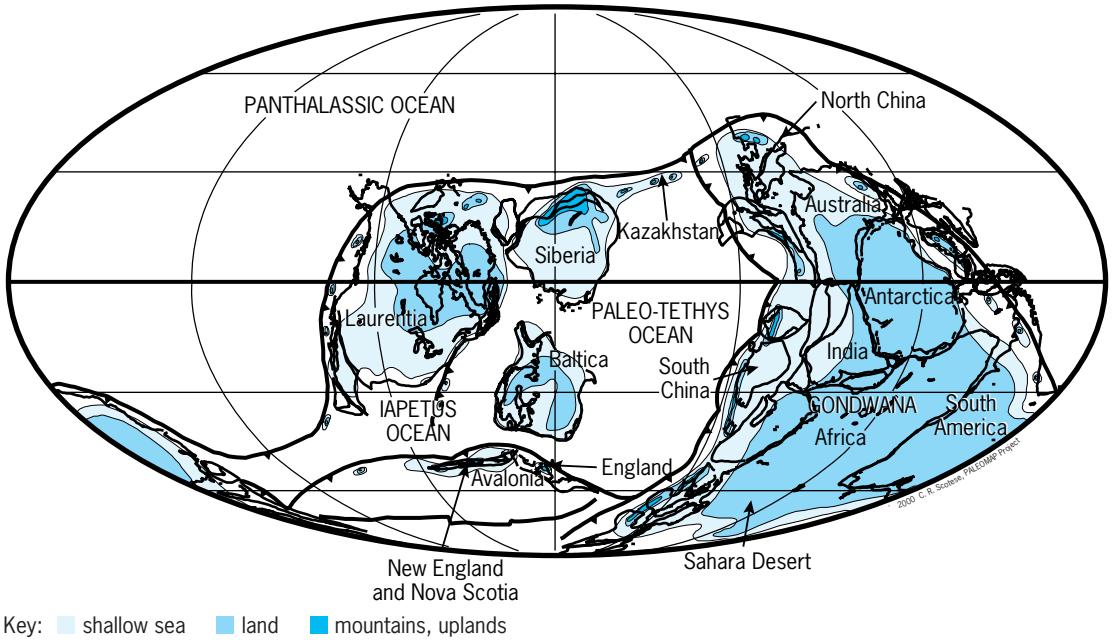

Fig. 2. Late Ordovician, 458 Ma. (*Paleogeographic maps by Christoper R. Scotese, PALEOMAP Project, University of Texas at Arlington*)

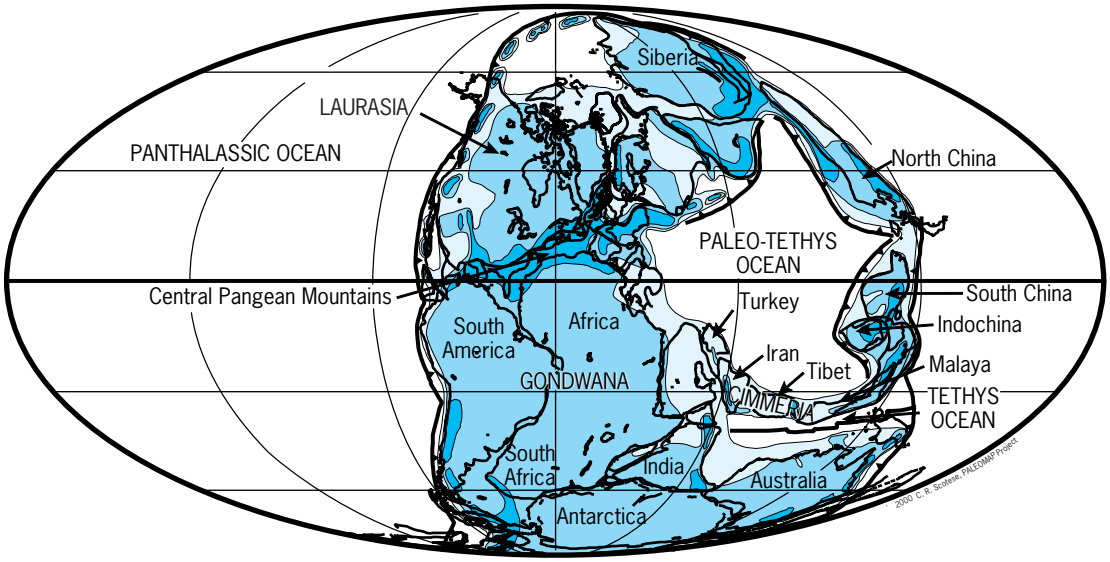

Key: shallow sea land mountains, uplands

Fig. 3. Late Permian, 258 Ma. (*Paleogeographic maps by Christoper R. Scotese, PALEOMAP Project, University of Texas at Arlington*)

rifted away from northwest Africa, opening the Central Atlantic (**Fig. 4**). This movement also gave rise to the Gulf of Mexico. At the same time, on the other side of Africa, extensive volcanic eruptions along the adjacent margins of east Africa, Antarctica, and Madagascar heralded the formation of the Western Indian Ocean. *See* JURASSIC; VOLCANO; VOLCANOLOGY.

During the Mesozoic, North America and Eurasia were one landmass, called Laurasia. As the Central Atlantic Ocean opened, Laurasia rotated clockwise, sending North America northward and Eurasia southward. This clockwise, seesaw motion of Laurasia also led to the closure of the wide V-shaped ocean, Tethys, that separated Laurasia from the fragmenting southern supercontinent, Gondwana. *See* MESOZOIC.

Late Mesozoic, 144–66 Ma. The second phase in the breakup of Pangea began in the Early Cretaceous, about 140 Ma. Gondwana continued to fragment as South America separated from Africa, opening the South Atlantic, and India together with Madagascar rifted away from Antarctica and the western margin of Australia, opening the Eastern Indian Ocean (**Fig. 5**). The South Atlantic did not open all at once, but progressively "unzipped" from south to north. The initiation of rifting between North America and

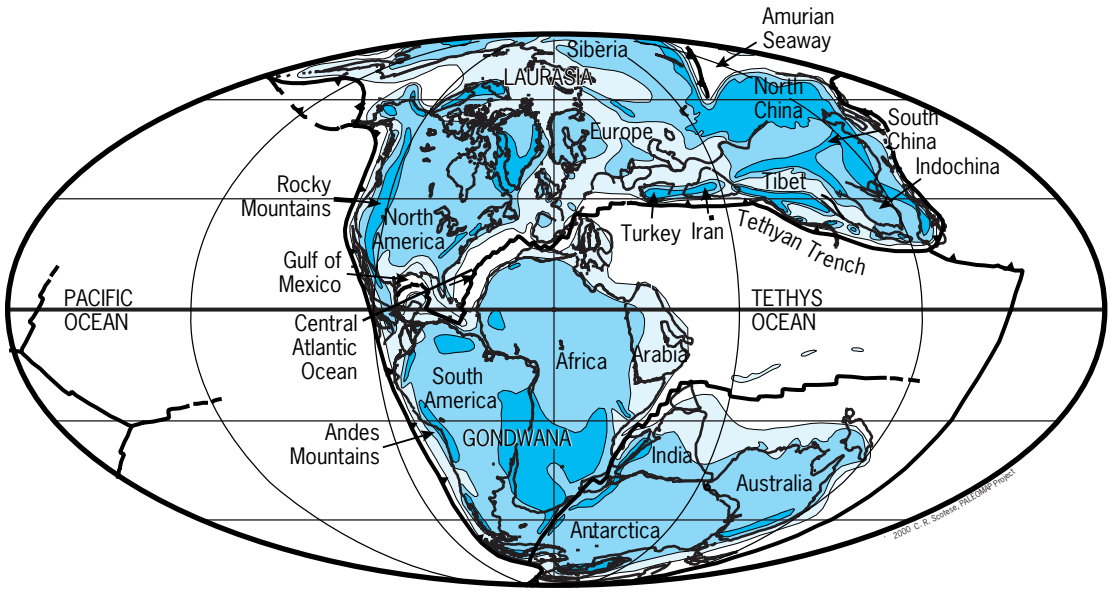

Key: shallow sea land mountains, uplands

Fig. 4. Late Jurassic, 152 Ma. (*Paleogeographic maps by Christoper R. Scotese, PALEOMAP Project, University of Texas at Arlington*)

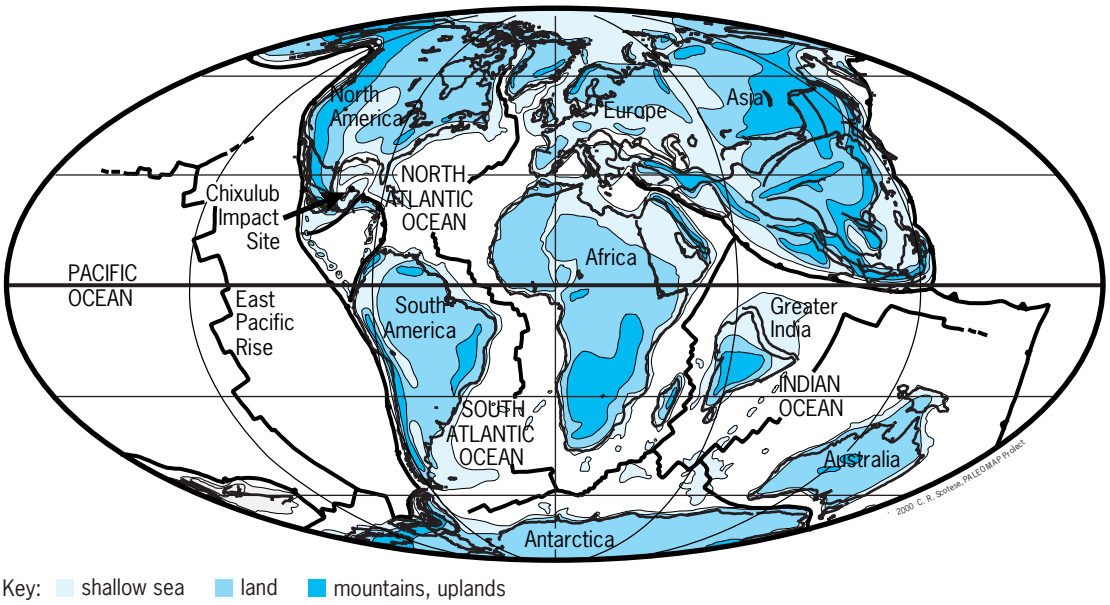

Key: shallow sea land mountains, uplands

Fig. 5. Cretaceous-Tertiary (K/T) boundary, 66 Ma. (*Paleogeographic maps by Christoper R. Scotese, PALEOMAP Project, University of Texas at Arlington*)

Europe, the counterclockwise rotation of Iberia from France, the separation of India from Madagascar, the uplift of the Rocky Mountains, and the arrival of exotic terranes (Wrangellia, Stikinia) along the western margin of North America also took place during the Cretaceous. *See* CRETACEOUS; TERTIARY.

During the late Mesozoic, shallow seaways covered the continents. Higher sea level was due in part to the creation of new rifts in the ocean basins that displaced water onto the continents.

Cenozoic Era, 66–0 Ma. The third and final phase in the breakup of Pangea took place during the early Cenozoic. North America and Greenland split away from Europe, and Antarctica released Australia (**Fig. 6**). Additional, important rifting events have taken place during the last 20 million years of the Cenozoic Era. They include the rifting of Arabia away from Africa, opening the Red Sea; the creation of the East African Rift System; the opening of the Sea of Japan; and the northward motion of California and northern Mexico away from North America. *See* CENOZOIC; EOCENE.

Though several new oceans have opened during the Cenozoic, the last 66 million years of Earth history are better characterized as a time of intense continental collision. The most significant of these collisions was that between India and Eurasia, which began about 50 Ma. During the Late Cretaceous,

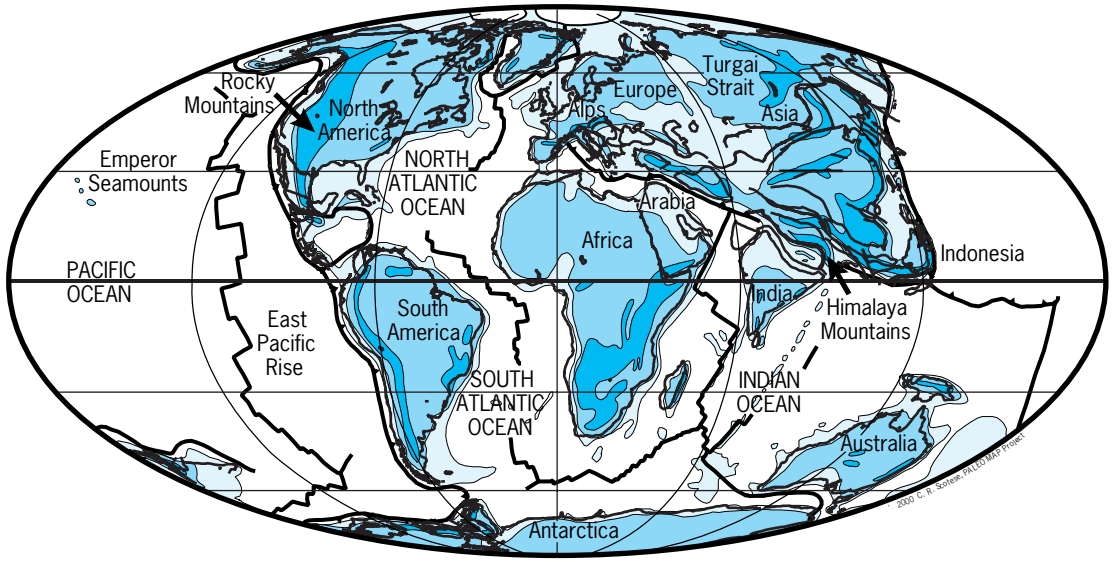

Key: shallow sea land mountains, uplands

Fig. 6. Eocene, 50 Ma. (*Paleogeographic maps by Christoper R. Scotese, PALEOMAP Project, University of Texas at Arlington*)

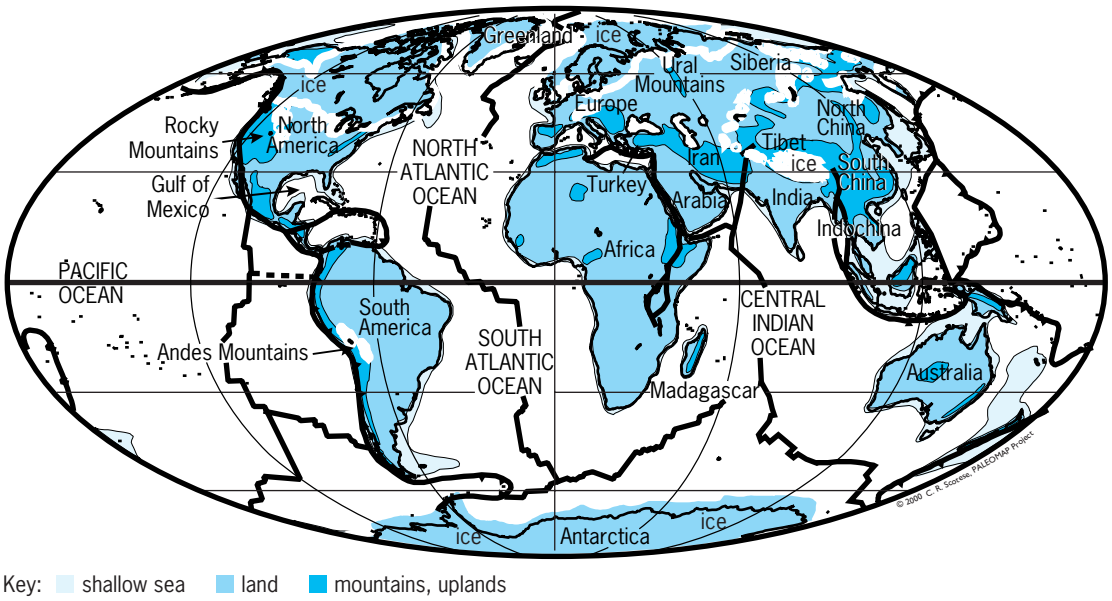

Key: shallow sea land mountains, uplands

Fig. 7. Last glacial maximum, 18,000 years ago. (*Paleogeographic maps by Christoper R. Scotese, PALEOMAP Project, University of Texas at Arlington*)

India approached Eurasia at rates of 15–20 cm/yr (6–8 in./yr)—a plate tectonic speed record. After colliding with marginal island arcs in the Late Cretaceous, the northern part of India, called Greater India, began to be subducted beneath Eurasia, raising the Tibetan Plateau.

The collision of India with Asia is just one of a series of continental collisions that has all but closed the great Tethys Ocean. From east to west these continent-continent collisions involved Spain with France, forming the Pyrenees mountains; Italy with France and Switzerland, forming the Alps; Greece and Turkey with the Balkan States, forming the Hellenide and Dinaride mountains; Arabia with Iran, forming the Zagros mountains; India with Asia; and finally, Australia with Indonesia.

Modern world. About 18,000 years ago, all of Antarctica and much of North America, northern Europe, and the mountainous regions of the world were covered by glaciers and great sheets of ice (**Fig. 7**). These ice sheets melted approximately 10,000 years ago, giving rise to familiar geographic features such as Hudson's Bay, the Great Lakes, the English Channel, and the fiords of Norway.

The continental rifting and collisions that began in the late Cenozoic are continuing today. Notable are the opening of the Red Sea and Gulf of Aden, the rifting of East Africa, the opening of the Gulf of

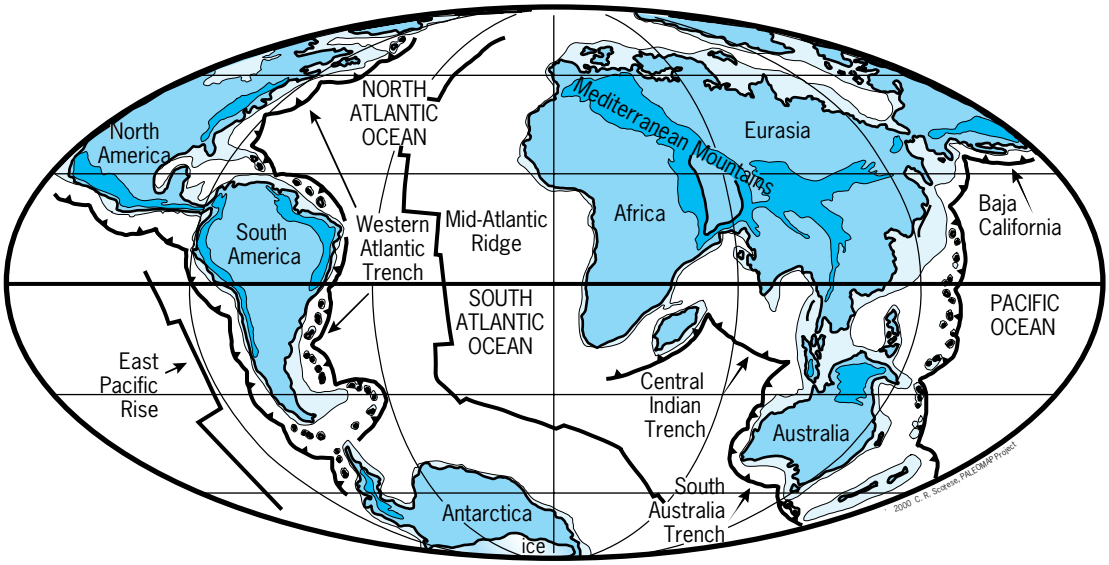

Key: shallow sea land mountains, uplands

Fig. 8. In 50 million years hence. (*Paleogeographic maps by Christoper R. Scotese, PALEOMAP Project, University of Texas at Arlington*)

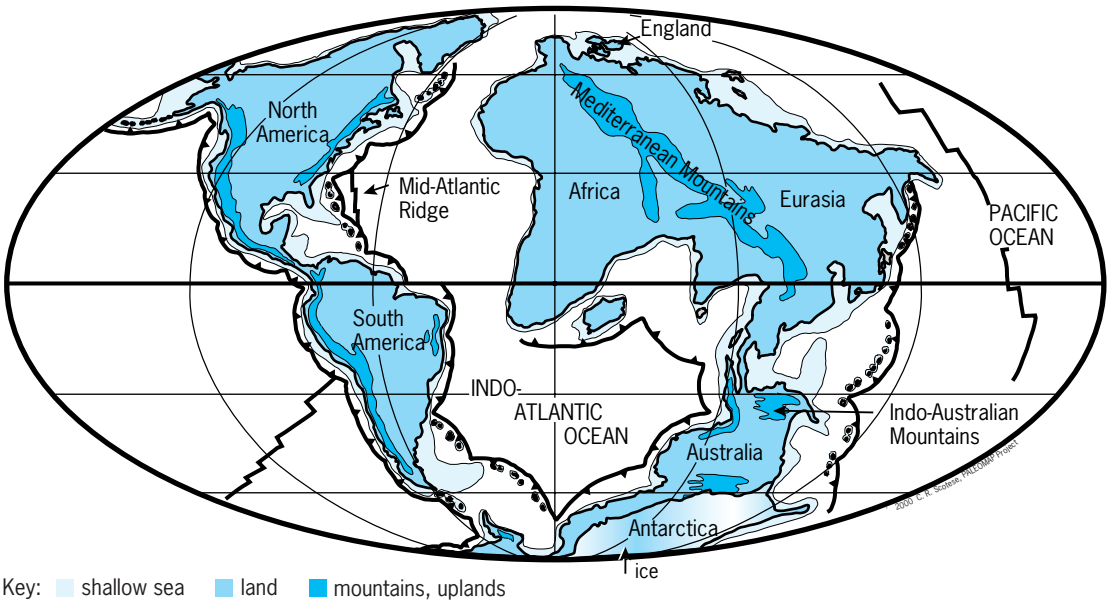

Fig. 9. In 150 million years hence. (*Paleogeographic maps by Christoper R. Scotese, PALEOMAP Project, University of Texas at Arlington*)

California and the northward translation of California west of the San Andreas Fault, and the incipient collision of Australia with Indonesia giving rise to the mountain ranges of New Guinea. *See* CENOZOIC.

Future world. Though there is no way of knowing the future geography of the Earth, it is possible to project current plate motions and make an educated guess. In the final three maps presented here, the Atlantic and Indian oceans continue to widen until new subduction zones recycle the ocean floor in these ocean basins and bring the continents back together in a new Pangean configuration some 250 million years hence.

The reconstruction of the world 50 million years in the future looks slightly askew (**Fig. 8**). North America is rotated counterclockwise, while Eurasia is rotated clockwise bringing England closer to the North Pole and Siberia down to warm subtropical latitudes. Africa has collided with Europe and Arabia, closing the Mediterranean Sea and the Red Sea. Similarly, Australia has beached itself on the doorstep of Southeast Asia, and a new subduction zone encircles Australia and extends westward across the Central Indian Ocean. It is interesting that current plate trajectories suggest that the East African Rift will not grow into a wide ocean.

Though the Atlantic Ocean has widened, extensions of the Puerto Rican Trough and the Scotia Arc

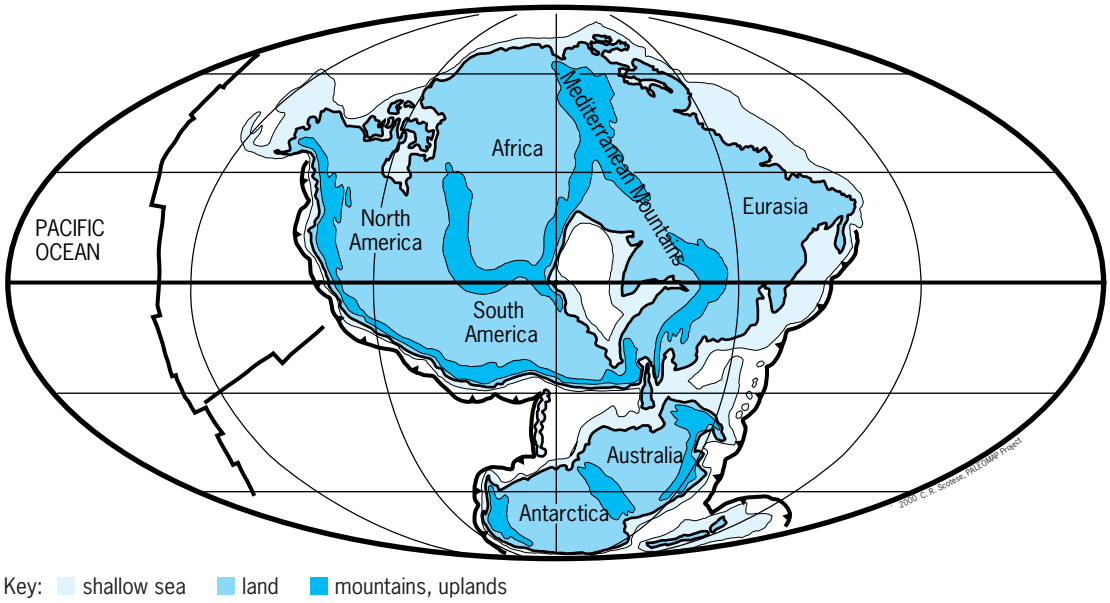

Fig. 10. In 250 million years hence. (*Paleogeographic maps by Christoper R. Scotese, PALEOMAP Project, University of Texas at Arlington*)

have started a new subduction zone along the eastern edge of the Americas. In time, this subduction zone will consume the North and South Atlantic oceans.

The map of the world 150 million years hence shows the contraction of the Atlantic and Indian oceans. Antarctica has collided with Australia, and the Mid-Atlantic Ridge has nearly been subducted beneath the Eastern American subduction zone (**Fig. 9**).

In the final reconstruction of future plate motions (**Fig. 10**), the Indo-Atlantic Ocean is completely closed and a new supercontinent, "Pangea Ultima," was formed. North America has collided back against Africa, but in a more southerly position that juxtaposes Miami with Cape Town. South America wraps around the southern tip of Africa, with Patagonia in contact with Indonesia enclosing a remnant of the Indian Ocean. Antarctica is once again at the South Pole, and the Pacific has grown wider, encircling half the Earth. Christopher R. Scotese

Bibliography. I. W. D. Dalziel, Pacific margins of Laurentia and East Antarctica-Australia as a conjugate rift pair: Evidence and implications for the Eocambrian supercontinent, *Geology*, 19:598–601, 1991; W. S. McKerrow and C. R. Scotese, Palaeozoic Palaeogeography and Biogeography, *Geol. Soc. London Mem.*, no. 12, 1990; C. R. Scotese et al., Paleozoic base maps, *J. Geol.*, 87:217–277, 1979; C. R. Scotese and W. W. Sager (eds.), Mesozoic and Cenozoic Plate Reconstructions, *Tectonophysics*, vol. 155, no. 1–4, Elsevier, Amsterdam, 1989; A. G. Smith, D. G. Smith, and B. M. Funnell, *Atlas of Mesozoic and Cenozoic Coastlines*, Cambridge University Press, 1994; R. Van der Voo, *Paleomagnetism of the Atlantic, Tethys, and Iapetus Oceans*, Cambridge University Press, 1993.

Paleoindian

The oldest archeological cultures of the New World, the ancestors of modern Native Americans, are termed Paleoindian. These colonizing populations of the Americas were *Homo sapiens sapiens* who arrived during the late Pleistocene (Ice Age) from Asia, though precisely when, and whether in a single or multiple pulses of migration, are not yet known.

Arrival in America. The presumed entryway was the Bering Straits, which emerged as dry land during glacial periods, when as much as 5% of the world's water was frozen in massive continental ice sheets (see **illustration**). As glaciers expanded, sea levels dropped 100–150 m below their present levels, exposing shallow parts of the continental shelf worldwide, including that beneath the Bering Sea. The Bering Land Bridge (or Beringia) existed most recently between 25,000 and 11,000 years before present (B.P.), during the last major episode of the Pleistocene. Even at earlier (and later) unglaciated times when Beringia was under water, crossings on foot may have been possible over winter pack ice.

There were no significant geographic or ecological barriers to cross-Beringian traffic. The land bridge was about 1500 km (930 mi) wide, mostly flat, cold, relatively dry, and free of glaciers, except in the mountains on its edges. The absence of barriers explains why the late Pleistocene animal communities of Siberia and Alaska were essentially identical.

The conditions that linked Siberia to Alaska may have simultaneously hindered migration south from Alaska. Groups headed in that direction had two potential routes (broadly defined): down the Pacific coast, or via the continental interior along the eastern flank of the Rocky Mountains. During the late Pleistocene the vast North American glaciers, the Cordilleran and Laurentide, covered much of the Northwest Territories and Canada (at their maximum at 18,000 B.P., extending in western North America from 75° to 48°N), and at times effectively blocked both routes.

Current evidence indicates that a coastal route was largely unglaciated and traversable on foot before 23,000 and after 14/13,000 B.P. (the timing of deglaciation is complex and poorly known). In the millennia in between, 20–50-km-wide tongues of heavily crevassed glacial ice extended down to the coast, obstructing passage. Groups in watercraft may have skirted down the coast, but there is no evidence of such craft, or of a group with the adaptation necessary to survive in a maritime environment (the land being either under ice or resource-poor). By 14,000 B.P., the outer coast was becoming suitable for human colonization: a partly vegetated environment of open tundra with grasses, sedges, and dwarf willows began to develop, and became forested by 12,000 B.P.

For groups headed south via the continental interior, the northern approaches to the eastern slopes of the Rocky Mountains were blocked by the Laurentide glacier, which lapped against the eastern flanks of the Richardson and Mackenzie mountains as early as 30,000 B.P., and remained there until around 11,500 B.P. Groups could have skirted south of those ranges and emerged east of the Rocky Mountains via one of the southern passes (such as along the Liard River), but whether or when such routes would have been traversable is yet unknown. Once the Cordilleran and Laurentide glaciers retreated and opened an ice-free corridor along the Rocky Mountain front, it was 2000–3000 years before the deglaciated landscape was recolonized by plants and animals. Only then, after approximately 11,500 B.P., did the corridor provide sufficient food resources for migrating humans.

There is no evidence yet as to which routes might have been taken. The earliest archeological evidence of a human presence along the Pacific Northwest coast dates to about 9500 B.P., and in the interior corridor to 10,500 B.P.—both well past the presumed entry time. To complicate the picture, the Pleistocene coastline, and any early archeological sites that might be on it, is now mostly underwater following postglacial sea-level rise.

Location of some important North American Paleoindian sites during the late-Pleistocene ice age.

Antiquity. So far, the earliest archeologically confirmed dates put human groups in the Lena Basin and Lake Baikal region of northeast Asia at about 39,000 B.P., in subarctic Siberia by 25,000 B.P., but not in western Beringia (such as Kamchatka) until 14,000 B.P. Humans were in eastern Beringia (Alaska) soon after 12,000 B. P., and present south of the ice sheets in North America by at least 11,500 B.P.—the latter represented by the Clovis culture. Yet, the earliest accepted archeological evidence puts human groups in South America earlier still, by at least 12,500 B.P. at the site of Monte Verde, Chile.

There are no obvious historical or technological affinities between Clovis and the Monte Verde materials, suggesting that the two may represent populations with distinct archeological traditions and separate migratory pulses: a later one (Clovis) that came south through the ice-free corridor soon after it became viable for travel, and an earlier population that perhaps moved along the Pacific coast and reached South America without, so far at least, any traces being found in North America. A hypothesis of multiple migrations, however, must remain tentative, given the small number of early South American sites known (Monte Verde represents one of the only sites of this age, so broader cultural patterns in artifacts or adaptations have yet to be recognized), the liabilities of comparing archeological assemblages so widely separated in time and space, and the evidence from some genetic studies which indicates all Native Americans are descended from a single group. Nonetheless, the possibility of multiple migrations is reasonable, given the absence of barriers to cross-Beringian traffic, and the long period in which potential source populations were present in northeast Asia. There were likely repeated movements from Asia to America and back.

Progenitors. None of the artifact complexes evident among late Pleistocene northeast Asian groups appears similar to those of New World Paleoindians. There are, however, archeological assemblages in Alaska (the Nenana Complex) that slightly predate Clovis and are argued to be similar in form and technology and thus historically linked. Yet, these assemblages lack the diagnostic hallmark of the Clovis technology—fluted projectile points. Fluted points (though not Clovis fluted points) are present (if rare) in Alaska, but none of these specimens has been dated earlier than Clovis, and many appear younger than those farther south.

The absence of an obvious Asian (or Alaskan) predecessor has led some to argue that Clovis is derived from Upper Paleolithic groups in western Europe, known as Solutrean, who supposedly crossed the North Atlantic to reach America. However, current archeological evidence, as well as evidence from genetics, human osteology and teeth, and linguistics, provides no support for such a link (or, for that matter, for a link to any other non-Asian source population). The morphological differences between a few ancient human skeletal remains found in the Americas (such as at Kennewick, WA) and those of modern Native Americans have been seen as supporting the hypothesis that the Americas may have been peopled from continents other than Asia. Yet, most likely all these individuals are descended from northeast Asians, the differences in their form over time the result of long periods of geographic isolation, mutation and genetic drift, and evolutionary change among descendant populations.

North and South America. Clovis and Clovis-like materials are concentrated in North America, and the northern reaches of Central America. Clovis is a widespread entity that first appears on the western Plains and southwest at 11,500 B.P. and in eastern North America at 10,600 B.P. That Clovis and related groups apparently expanded across the continent in what may have been less than 1000 years is all the more remarkable given that they spread at a time of geologically rapid environmental and climatic change marked by continental deglaciation, the extinction of nearly 36 genera of large mammals (megafauna), and the dissolution of long-standing biotic communities. Yet Clovis groups seemingly coped with such adaptive challenges with ease: Their stylistically distinctive projectile points and tool kits—often including bifacial knives, a variety of unifacial scrapers, occasional blades and flake tools, and (more rarely) bone and ivory implements—are surprisingly similar across the continent. These were highly mobile groups who relied on high-quality stone often obtained from geological sources hundreds of kilometers from the sites where the stone was used and discarded. Their rapid radiation, broadly similar tool kits, and long-distance movement bespeak a cultural "founders effect," suggesting that their access to large areas of North America was largely unrestricted—these groups were moving across an essentially empty landscape.

Some argue the similarity in the Clovis tool kit and the rapidity of their movement bespeak a uniform adaptation—big-game hunting. Others take the argument a step further, suggesting that Clovis predation on mammoth and mastodon and other megafauna drove these animals to extinction in the late Pleistocene. Yet, the archeological record for big-game hunting is limited to a dozen sites on the Plains and southwest. In most other areas, there is no evidence that big game was exploited. Instead, Clovis subsistence, it appears, more often involved less risky and smaller prey—and presumably plants, though remains of such are rarely preserved in the archeological record of this period.

Although the timing varies by area, by 10,500 B.P. the Clovis tradition was replaced by regional Paleoindian variants, which generally (though not always) have reduced settlement mobility (relative to Clovis), and include new technologies, prey-specific strategies for hunting and processing (such as the intensive use of bison on the Plains), increasing use of local resources, and distinctive stylistic elements and functional artifact forms. This shift from broad and overall homogeneity in Clovis times to narrower and more regionally restricted complexes in later Paleoindian times likely reflects the setting in of colonizers to specific areas or habitats, and increasingly region-specific adaptations.

The South American Paleoindian record, by contrast, does not evince any artifact forms that dominate the archeological landscape as Clovis does. Instead, this period is marked by more diverse unifacial and bifacial stone tool technologies, often made of stone acquired locally (and not necessarily of superior quality), and includes forms such as bolas—modified spherical stones used in slings or as hand missiles. Projectile points tend to be less common in assemblages here than in North America, and show considerable stylistic variety. While Clovis points per se are absent from South America, some point forms are fluted, which is often cited as evidence of a historical link between the continents. However, there is growing debate about whether that similarity is merely a case of technological convergence.

South American Paleoindians utilized a wide range of animals, and early on even made occasionally heavy use of plants. This is especially evident at Monte Verde, which yielded (in 12,500 B.P. deposits) nearly 70 species of plants, most locally available but some acquired from the distant highlands and coast. Many had food, medicinal, or economic value. An early (pre-10,000 B.P.) use of plants across the continent was followed within just a few millennia by the emergence of domesticated plants (which occur much earlier here than in North America). Generalized foraging, with occasional big-game hunting, is not unexpected, given the considerable ecological variability and richness of South America. In part, this may reflect the fact that glaciation was not nearly so extensive on this continent—its effects were limited to high altitudes and high latitudes. While that prevented early colonization of those areas, in the remainder of the continent early

immigrants could move freely throughout environments that were fairly dynamic in some areas, relatively stable in others, and in places ecologically rich. There was also use of nonterrestrial environments: early use of maritime resources is evident at several sites on the Atlantic and (especially) the Pacific coasts. Such adaptations played a key role in establishing early sedentary lifestyles.

Once the founding population dispersed across South America (over an unknown length of time), subgroups became geographically isolated relatively quickly. From the earliest known site at 12,500 B.P. (Monte Verde) until the end of the Pleistocene (10,000 B.P.), there is a continuing diversification in tool forms and technology, evidently reflecting less mobility, increasing heterogeneity and regional mosaics in culture and adaptations, and less expansive social networks and territories.

All told, it is a different trajectory from the one that unfolded in North America—testimony that the earliest colonization of the two continents, though ultimately derived from the same northeast Asian source, may have taken place at different times under very different circumstances. *See* ARCHEOLOGY; PLEISTOCENE; PREHISTORIC TECHNOLOGY.

David J. Meltzer

Bibliography. J. M. Adovasio and J. Page, *The First Americans: In Pursuit of Archaeology's Greatest Mystery*, Random House, New York, 2002; R. Bonnichsen and K. Turnmire (eds.), *Clovis: Origins and Adaptations*, Center for the Study of the First Americans, Oregon State University, Corvallis, 1991; T. D. Dillehay, *The Settlement of the Americas: A New Prehistory*, Basic Books, New York, 2000; N. Jablonski (ed.), *The First Americans: The Pleistocene Colonization of the New World*, California Academy of Sciences and University of California Press, Berkeley, 2002; D. J. Meltzer, *Search for the First Americans*, Smithsonian Books, Washington, DC, 1993; F. H. West (ed.), *American Beginnings*, University of Chicago Press, 1996.

Paleolithic

The prehistoric period when people made stone tools exclusively by chipping or flaking. John Lubbock proposed and defined the term Paleolithic, or Old Stone Age, in 1865, and also defined a subsequent stage, Neolithic or New Stone Age, during which some stone tools were formed by polishing or grinding. Later archeologists altered these definitions; to many today the Paleolithic is the period during which human beings lived entirely by hunting and gathering, while the Neolithic is the following interval during which plant and animal domestication was introduced. To other archeologists, the Paleolithic is simply a time interval, roughly equivalent to the Pleistocene Epoch, while the Neolithic comprises the early part of the succeeding Holocene (or Recent) Epoch. *See* GEOLOGIC TIME SCALE; NEOLITHIC.

The different definitions sometimes lead to contradictory results, as when hunter-gatherers making chipped stone tools are found to have survived into the Holocene (even to historic times) or when polished stone tools appear to have been made by people who did not know agriculture. To eliminate the contradictions or to refine the definitions, some specialists inject a transitional Mesolithic between the Paleolithic and Neolithic. It is impossible, however, to devise a rigorous, global definition of the Paleolithic or any other cultural stage, because artifact technology and economic practices have changed independently at different times in different parts of the world. Hence, Paleolithic will be used here informally to refer to the time interval between the earliest appearance of stone tools, more than 2.5 million years before present (m.y. B.P.) and the end of the last glacial period, 12,000–10,000 years B.P. (**Fig. 1**).

Artifacts. The oldest artifacts found so far come from sites in Ethiopia, Kenya, and Tanzania where they are dated to between 2.6 and 1.6 m.y. B.P. They comprise crude flakes and the modified pebbles and stone chunks from which the flakes were struck (**Fig. 2**). Collectively they are often assigned to the Oldowan Industrial Complex, named after Olduvai Gorge (Tanzania). Some researchers believe that the Omo industrial tradition, found in a few places in East Africa, preceded the Oldowan. These "tools" are little more than smashed nodules, requiring less technical skill to make than those of the Oldowan. Older artifacts such as the Omo type may be difficult to find even if they exist elsewhere, because their makers might not yet have developed the uniquely human habit of accumulating refuse at repeatedly occupied sites with good archeological visibility.

While it is not certain which type of australopithecine, the earliest of human ancestral forms, may have produced the earliest stone tools, the principal makers of Oldowan artifacts were probably members of the subsequent human species, *Homo habilis*. Damage marks on associated animal bones show that Oldowan tools were sometimes used for butchering, while wear traces observed on the tools themselves show that some were used to cut meat or reeds, and others functioned to scrape or saw wood. Any wooden artifacts that resulted perished long ago in the ground.

Approximately 1.6–1.5 m.y. B.P., at least some people in East Africa, most likely a *Homo erectus* type, began to manufacture the bifacially flaked tools known to archeologists as hand axes. The flaking of the first hand axes was crude, and they differ only subtly from earlier Oldowan flaked pebbles and hunks, which continued to be made. As time passed, however, their flaking tended to become more refined, and the thinness and bilateral symmetry of some later hand axes may reflect esthetic as well as functional considerations.

By 1 m.y. B.P. hand-ax makers had spread through most of Africa and the Near East, and by 900,000–600,000 years B.P. they had reached Europe. All African, Near Eastern, and European artifact assemblages with hand axes are conventionally placed in

the Acheulean Industrial Complex, after the site of St. Acheul, northern France, discovered in the 1850s (**Fig. 3**).

Acheulean artifacts were made by *Homo ergastor/H. erectus* before 500,000–400,000 years B.P., and by early *H. sapiens* after this time. The Acheulean Complex does not appear to have extended to the Far East, where *H. erectus* and perhaps early *H. sapiens* produced artifact assemblages broadly similar to Oldowan ones. It does not follow that the Far East lagged behind Africa and Europe technically, however, since hand ax-like bifacial artifacts have now been recorded at several east Asian early Paleolithic sites, and many important early Paleolithic assemblages in the west also lack hand axes. Early peoples of the Far East simply appear to have employed different means of solving their technological problems than the producers of Acheulean artifacts did. *See* FOSSIL HUMANS.

Lower Paleolithic. In sub-Saharan Africa the Oldowan and Acheulean complexes are frequently placed together in the Early Stone Age, while elsewhere Acheulean and contemporaneous assemblages are commonly referred to the Lower Paleolithic. The Early Stone Age/Lower Paleolithic apparently persisted until sometime between 200,000 and 130,000 years B.P., the exact time perhaps depending on the region.

Middle Paleolithic. In sub-Saharan Africa the Early Stone Age was succeeded by the Middle Stone Age, while in North Africa and Eurasia the Lower Paleolithic was followed by the Middle Paleolithic, also commonly known as the Mousterian, after the cave of Le Moustier, southwestern France, where rich Middle Paleolithic levels were first excavated in the 1860s.

Most Middle Stone Age/Middle Paleolithic assemblages lack hand axes; when hand axes are present they tend to be far smaller and more delicate than Acheulean ones. The principal Middle Stone Age/Middle Paleolithic tools are well-made stone flakes, often modified by edge flaking ("retouch") into types called sidescrapers, knives, denticulates (serrate-edged pieces), and so forth (**Fig. 4**). The typology is based on form alone, and wear traces on some tools show that pieces assigned to a single type may have been used for different purposes. Wear traces also reveal that unmodified flakes were often used as well.

In Europe, Middle Paleolithic people were the Neandertals, *H. (sapiens) neanderthalensis*. Neandertals also lived in the Near East during the early Middle Paleolithic. Neandertal occupation of the region apparently overlapped with that of the some of the first anatomically modern people (*H. sapiens sapiens*) who disappear from the archaeological record for some time, but then return and supplant the Neandertals after ?70,000–60,000 years B.P. Outside of Europe and the Near East, bones representing Middle Paleolithic/Middle Stone Age people are rare, very fragmentary, or both. The few known African fossils suggest that earlier Middle Stone Age people belonged to an archaic variety of *H. sapiens*, while later

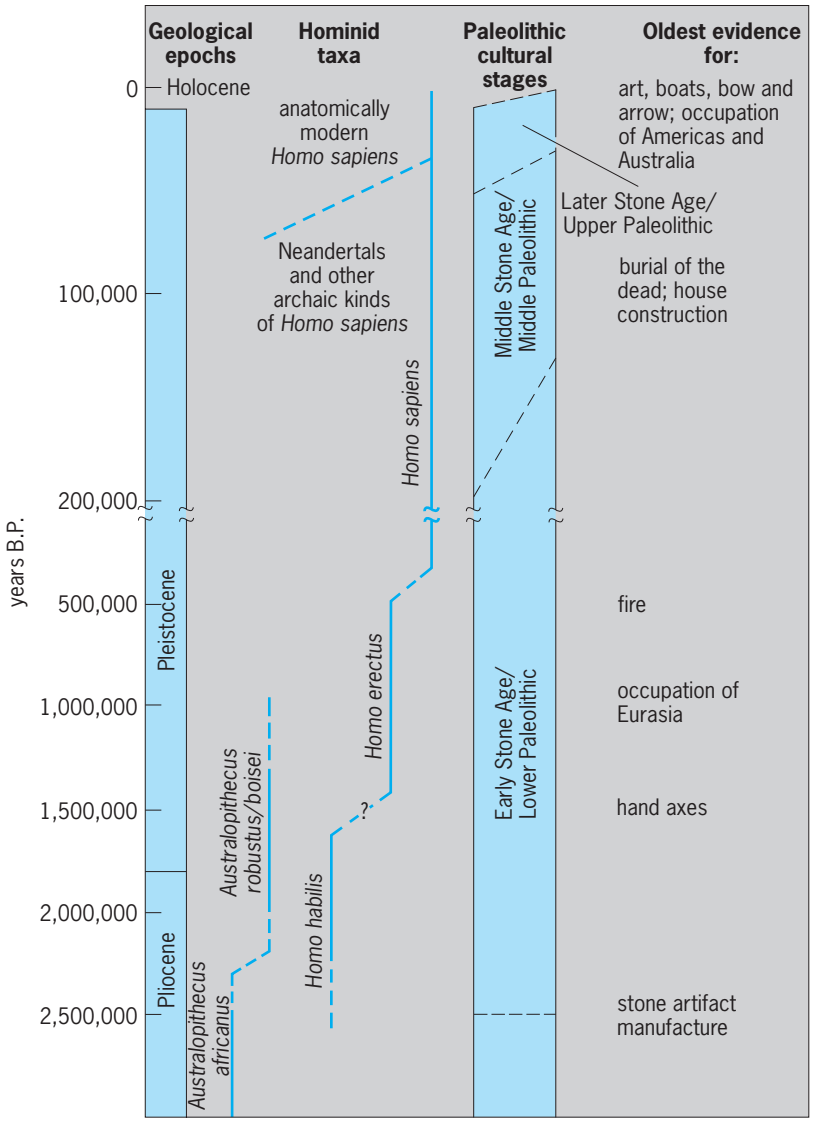

Fig. 1. Relationship between time, hominid taxa, and major cultural developments during the Paleolithic.

Middle Stone Age people (after ?70,000–60,000 years B.P.) may have been anatomically modern. *See* NEANDERTALS.

Upper Paleolithic. In Europe, the Near East, and North Africa, the Middle Paleolithic was followed by the Upper Paleolithic, perhaps 50,000 years B.P. in the Near East, adjacent North Africa, and eastern Europe, and beginning around 40,000 years B.P. in western Europe. The time difference within Europe

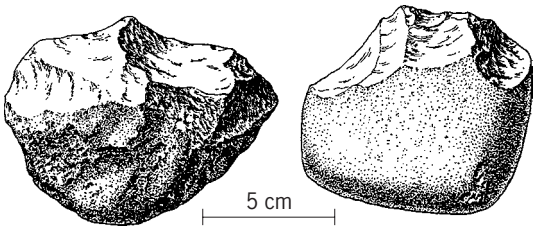

Fig. 2. Oldowan pebble tools from Bed I, Olduvai Gorge. (*After M. D. Leakey, Olduvai Gorge, vol. 3, p. 27, Cambridge University Press, 1971*)

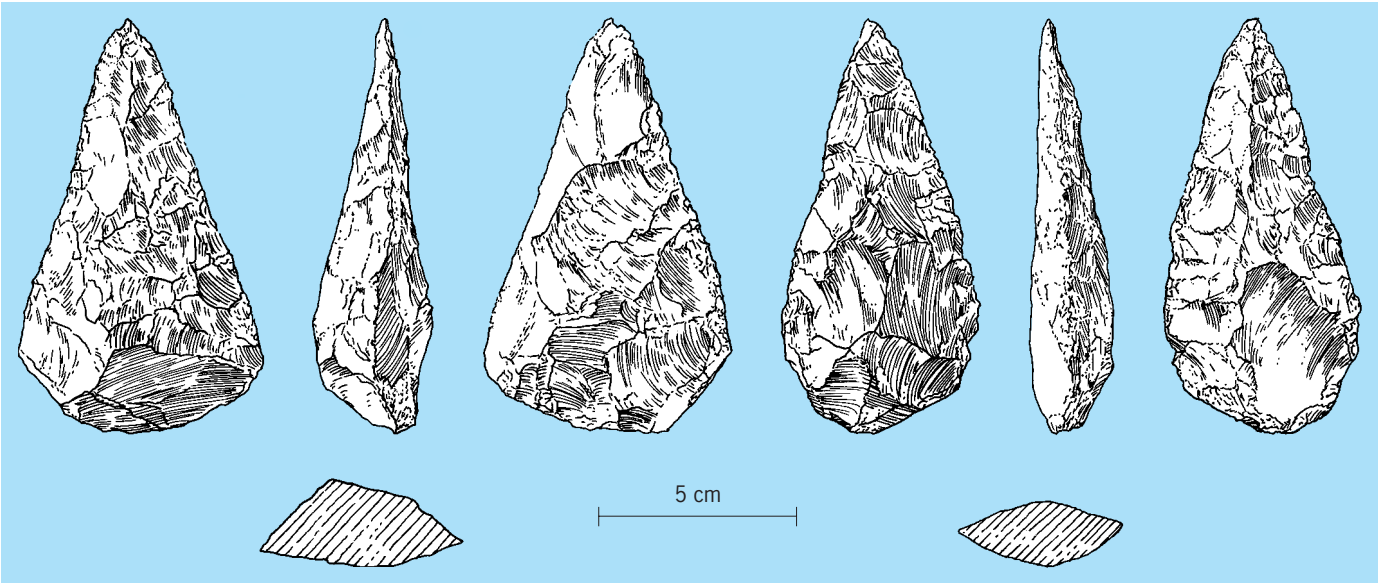

Fig. 3. Late Acheulean hand axes from England. (*After D. A. Roe, The Lower and Middle Palaeolithic Periods in Britain, Routledge Kegan Paul, 1981*)

may reflect the movement of technology, people, or both from east to west. In sub-Saharan Africa, the Middle Stone Age was replaced by the Later Stone Age between 50,000 and 35,000–30,000 years B.P., the precise time again perhaps depending on the region. With the exception of one individual from a site in France, all known Upper Paleolithic/Later Stone Age people were anatomically modern. However, human skeletal remains (archaic or modern) dating to the time of cultural and technological transition itself are extremely rare, and the noted exception is a Neandertal from a level that not all specialists would in fact assign to the Upper (versus the Middle) Paleolithic. The disagreement is part of an ongoing debate about whether the Upper Paleolithic (and modern people) were intrusive into western Europe or evolved there more or less independently.

The Upper Paleolithic and Later Stone Age are difficult to characterize artifactually, because wherever they are well known, they exhibit great variability in time and space. It is often said that Upper Paleolithic "cultures" shared a tendency to manufacture tools on flakes that are at least twice as long as wide (such pieces are called blades) and that the principal Upper Paleolithic tool types were end scrapers and burins (**Fig. 5**). While most Upper Paleolithic assemblages do exhibit these features, there are some that do not, and even among those that do, there is extraordinary variability in the actual frequency of long flakes (blades) and in the kinds of end scrapers and burins that were made. The amount of artifactual change through time and space during the Upper Paleolithic/Later Stone Age far exceeds that in earlier periods and suggests an ability to innovate that earlier people did not exhibit.

Among the most important Upper Paleolithic/Later Stone Age innovations were the spearthrower, the bow and arrow, and tailored clothing. These projectile technologies help to explain why Upper Paleolithic/Later Stone Age people were more effective hunters than their predecessors. The invention of tailored clothing helps explain how Upper Paleolithic people managed to colonize northeastern Europe and northern Asia (Siberia), which were apparently uninhabited earlier. *See* PREHISTORIC TECHNOLOGY.

Upper Paleolithic/Later Stone Age people were also the first to manufacture standardized, formal tool types from bone, antler, and ivory. These include objects that archeologists call points, awls, and needles. In addition, they were the first people to produce what anthropologists all agree is art, including both painted and engraved wall art and numerous portable items—figurines, pendants, beads, and other objects (**Fig. 6**). These were often made of ivory, bone, and antler and their appearance may be connected to Upper Paleolithic/Later Stone Age realization that such organic material could be carved, ground, or otherwise shaped into a wide variety of forms.

Conventionally, the Upper Paleolithic is said to end with the end of the Last Ice Age, 12,000–10,000 years B.P., although Upper Paleolithic/Later Stone Age artifact types and economic practices continued for many millennia throughout much of Eurasia and sub-Saharan Africa.

Economy. Throughout the long Paleolithic time span, all human beings lived by hunting and gathering wild resources. Only from the very end of the Paleolithic, about 12,000–10,000 years B.P., is there evidence that some people domesticated animals, plants, or both. *See* DOMESTICATION.

Like historic hunter-gatherers, most Paleolithic people collected edible wild plants. In lower and middle latitudes, plant foods probably supplied the bulk of Paleolithic diets, but direct evidence is lacking, due to the perishability of plant tissues. In contrast, animal bones, often representing food debris, are preserved at many sites, and have led archeologists to focus on the meat component of the diet out of proportion to its probable importance.

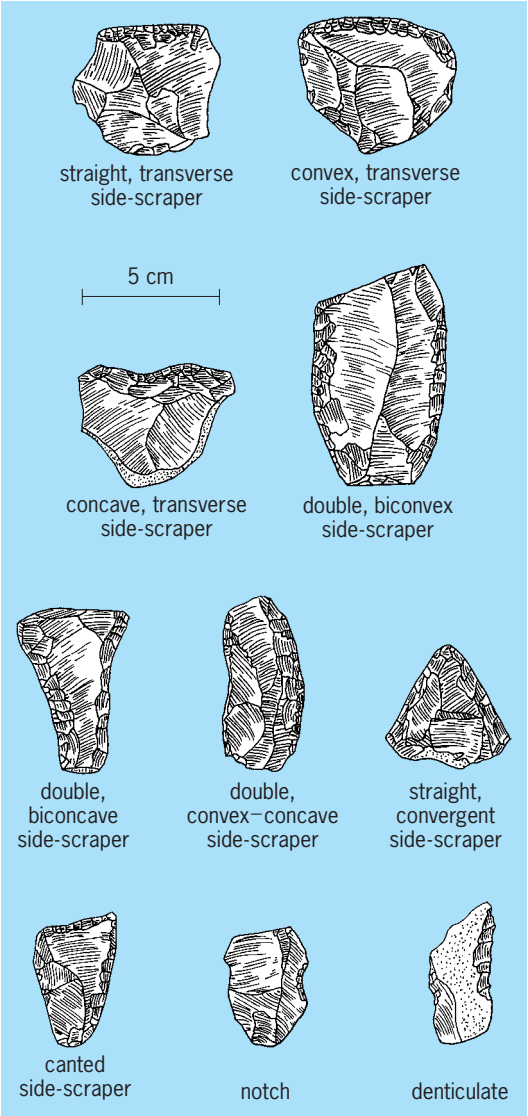

Fig. 4. Middle Paleolithic (Mousterian) flake tools from France. (*After R. G. Klein, Ice-Age Hunters of the Ukraine, University of Chicago Press, 1973*)

Reconstructing the diet of Early Stone Age/Lower Paleolithic people is complicated not only by preservation bias, but also by the geomorphic setting of nearly all the known sites. These are open-air (versus cave) occurrences located on the margins of ancient streams and lakes. At many sites, water action and perhaps kicking by drinking animals have displaced both artifacts and animal bones from their original positions. As a result, spatial associations that might have reflected ancient human butchering activity have been destroyed. There is the further problem that coprolites (fossil feces) and gnawed bones found at many sites show that carnivores were present. At no site can it be simply assumed that people killed all the animals represented or even that they scavenged the carcasses. The problem is to estimate the relative roles of people and other agents. For the moment, the best guess is that as time passed, meat became progressively more important in Lower Paleolithic diets and that human scavenging perhaps declined relative to hunting, but this cannot be demonstrated.

Demonstrating the importance of meat in Middle Stone Age/Middle Paleolithic and Later Stone Age/Upper Paleolithic diets is less problematic, because many of the relevant sites are caves in which people were the principal or sole bone accumulators. This is indicated by the abundance of artifacts, fossil hearths, and other traces of human activity (including cut animal bones) and by the rarity or absence of carnivore coprolites and gnawed bones. From such sites, it is possible to show that the people ate mainly medium-sized ungulates, including antelopes and zebra in Africa, and deer and wild horses in Eurasia.

It is also possible to show that Later Stone Age/Upper Paleolithic people probably ate more meat than Middle Stone Age/Middle Paleolithic people. Bone assemblages from southern African sites indicate that only Later Stone Age people regularly caught fish and flying birds and that they obtained truly dangerous prey such as wild pigs much more frequently than their predecessors did. It is probably not coincidental that only Later Stone Age sites have provided artifacts directly interpretable as fishing and fowling implements along with pieces that were almost certainly parts of arrows. Armed first with the spear-thrower, then later with the bow and arrow, Later Stone Age hunters could have attacked prey from a greater distance, increasing their chances of success and reducing their personal risk.

Other aspects of culture. The oldest reasonably secure evidence for human use of fire comes from the famous Peking man (*H. erectus*) site of Zhoukoudian in north China, tentatively dated to 500,000–240,000 years ago. More equivocal evidence for older or equally old controlled use of fire has been found in Kenya, South Africa, and Europe. Unequivocal fireplaces are found in many sites occupied by European Neandertals and their near-modern African

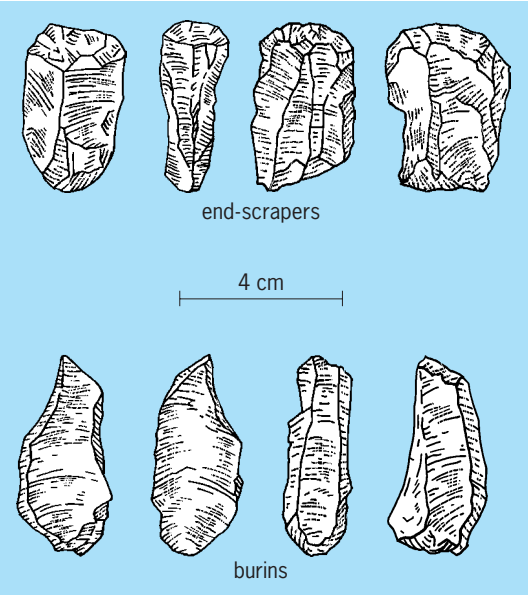

Fig. 5. Upper Paleolithic flake and blade artifacts from the Ukraine. (*After R. G. Klein, Ice-Age Hunters of the Ukraine, University of Chicago Press, 1973*)

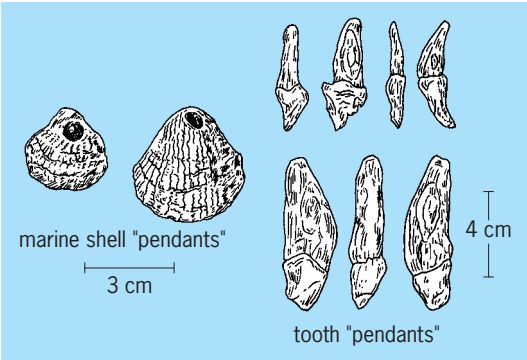

Fig. 6. Upper Paleolithic art objects from the Ukraine. (*After R. G. Klein, Ice-Age Hunters of the Ukraine, University of Chicago Press, 1973*)

contemporaries after 127,000 years B.P. The paucity of evidence for controlled use of fire during the Lower Paleolithic does not necessarily indicate a more limited ability to make fire. Rather, the Lower Paleolithic evidence is mostly from open-air sites near ancient water sources where charcoal fragments and concentrations that mark hearths are much less likely to survive than they are in cave sites, whence comes most of the Middle Paleolithic/Middle Stone Age evidence post 127,000 years ago.

Concentrations of rocks or other debris that may mark the positions of ancient structures such as wind breaks and flimsy tentlike shelters have been uncovered at several Lower Paleolithic sites in Africa and across Eurasia, but the oldest widely accepted "ruins" date from the Middle Paleolithic. However, it is not until 50,000–40,000 years ago that evidence for housing becomes compelling, with such features as fireplaces and substantial wall supports being commonplace.

Upper Paleolithic sites frequently contain spreads of large bones, artificially excavated depressions, or post holes marking the spot of ancient dwellings. Hearths, sometimes elaborate, are always present, and cache or storage pits for ivory, art objects, and other valuable items are common. At some sites, there are multiple ruins, which, if occupied simultaneously, would imply social groups probably far larger than any which existed earlier. Carefully constructed, well-heated dwellings were probably important in permitting Upper Paleolithic people to colonize northeastern Europe and Siberia for the first time.

Finally, the oldest undeniable evidence for burial of the dead comes from the Middle Paleolithic of Europe and the Near East. The deceased were Neandertals, and their skeletons sometimes exhibit pathologies or deformities that must have incapacitated the owners long before death. Perhaps these people could only have survived with care and aid from other members of their group.

Upper Paleolithic people also buried their dead, and the pathological or senile state of some skeletons again suggests a characteristically human care for sick or aged comrades. In addition, unlike Middle Paleolithic graves, Upper Paleolithic ones often contain special objects, such as hundreds of ivory beads, carved pendants, or other body ornaments, perhaps the deceased's personal belongings or items to facilitate the transition to an afterlife. Upper Paleolithic graves, together with Upper Paleolithic art, provide the oldest available evidence for the intangible part of culture called ideology or religion.

Population expansion. Human population growth during the Paleolithic is reflected in the ever-wider area that was occupied. In addition, population probably tended to increase in those areas that had long been inhabited.

Problems in finding and dating Lower Paleolithic/Early Stone Age sites make it impossible to compare their number per unit time to the numbers of later sites, but there are places where the number of Middle Paleolithic sites can be compared to the number of Upper Paleolithic ones. In each instance, Upper Paleolithic sites are much more numerous per unit time, suggesting larger Upper Paleolithic populations. This probably reflects Upper Paleolithic advances in technology, especially related to hunting.

Upper Paleolithic expansion to Siberia was important not only in its own right, but also because it was essential for the colonization of the Americas, which almost certainly occurred across a dry land bridge linking Siberia and Alaska during the Upper Paleolithic. This bridge owed its existence to the drop in sea level caused by the growth of the great continental ice sheets. Conceivably, people reached Alaska only shortly after they occupied Siberia (35,000–30,000 years B.P.), but this remains to be shown. There is a heated debate about when people reached the North American continent and regions farther south. The oldest universally accepted evidence for human presence is only about 15,000 years old. *See* PALEOINDIAN.

Australia was also probably first colonized in Upper Paleolithic times, 40,000–35,000 years B.P., or possibly by 60,000 years ago. In this case, there was no land bridge to the Asian mainland, and some knowledge of boats is implied. In both the Americas and Australia, the arrival of people appears to have been followed shortly by a wave of animal extinctions. If, as some specialists believe, people were largely responsible, then Upper Paleolithic times witnessed not only important innovations but also the first humanly caused ecological catastrophes.

Richard G. Klein; Anne Pike-Tay

Bibliography. O. Bar-Yosef, The Upper Paleolithic Revolution, *Annu. Rev. Anthropol.*, 31:363–393, 2002; R. G. Klein, *The Human Career: Human Biological and Cultural Origins*, 2d ed., University of Chicago Press, 1999; D. Lewis-Williams, *The Mind in the Cave: Consciousness and the Origins of Art*, Thames & Hudson, London, 2002; A. Pike-Tay and R. Cosgrove, From reindeer to wallaby: Recovering patterns of seasonality, mobility, and prey selection in the Paleolithic Old World, *J. Archaeol. Method Theory*, 9(2):101–146, 2002; C. Stringer, R. Barton, and J. Finlayson (eds.), *Neanderthals on the Edge*, Oxbow Books, Oxford, 2000; I. Tattersall and J. Schwartz, *Extinct Humans*, Westview Press, 2001.

Paleomagnetism

The study of the direction and intensity of the Earth's magnetic field through geologic time. Paleomagnetism is an important tool in measuring the past movements of the Earth's tectonic plates. By studying ancient magnetic field directions recorded in rocks, scientists learn how the plates moved relative to the Earth's spin axis and relative to one another. The calibrated history of geomagnetic polarity reversals provides a basis for the temporal correlation of rocks on a local to global geographic scale, called magnetostratigraphy. *See* PLATE TECTONICS; ROCK MAGNETISM; STRATIGRAPHY.

Magnetic field behavior. At all points on the Earth's surface, the geomagnetic field is represented by a vector of specific length and direction. Two features of the geomagnetic field are particularly useful. First, when averaged over 10^4 to 10^5 years, the field can be sufficiently represented by a dipole magnet located at the Earth's center with its north-south axis aligned with the rotation axis (**Fig. 1**). In this model, the time-averaged magnetization vector from a rock sequence magnetized on the Equator would be inclined at 0° with respect to local horizontal and would point toward the north geographic pole. If the same rock sequence were magnetized at the north pole, the time-averaged magnetization vector would be inclined vertically downward. For all points in between, the relationship between latitude and inclination is tan $I = 2 \tan \lambda$, where I is inclination and λ is latitude. Thus, if the magnetization vectors of 10^8-year-old rocks now at 30°N have a mean inclination of 0°, these rocks (and probably the tectonic plate upon which they lie) have moved 30° northward since the magnetization was acquired. Longitudinal motion is undetectable. For example, rocks formed anywhere on the Equator will acquire a magnetization with 0° inclination, and subsequent motion of the rocks along the Equator results in no latitudinal change.

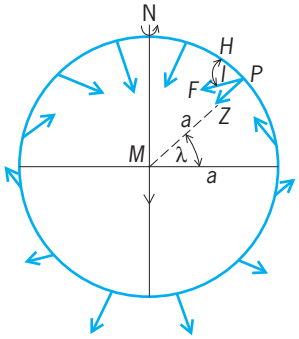

Key:

P = arbitrary point on the Earth's surface	H = horizontal component
a = radius of the Earth	M = magnetic moment of the dipole
λ = latitude	F = total field
Z = vertical component	I = inclination of the total field

Fig. 1. Field of an axial geocentric dipole. (*After M. W. McElhinny, Palaeomagnetism and Plate Tectonics, Cambridge University Press, 1973*)

On time scales shorter than about 10^5 years, the Earth's field at a given point exhibits considerable change in both direction and intensity. These variations, termed secular variations, can be useful for detailed stratigraphic correlation over a limited geographical range and time interval.

A second useful feature of the geomagnetic field is its period polarity reversals. The time-averaged axis of the central dipole remains parallel to the rotational axis before and after the reversal, but the north-seeking pole becomes the south-seeking pole and vice versa. Polarity changes from normal (N) to reverse (R) occur over 3000–5000 years and are effectively an instantaneous geological markers for stratigraphic correlation. Complete N-R-N or R-N-R sequences occur on time scales on the order of 10^5 to 10^7 years, making the geomagnetic polarity time scale a very powerful chronological tool.

Magnetization of rocks. Many rocks acquire remanent magnetizations when they form. These magnetizations are usually parallel to the direction of the ambient magnetic field at the time. Magnetic minerals in igneous rocks, such as basalt or gabbro, acquire a thermoremanent magnetization (TRM) as they cool below the Curie point T_C of the mineral [T_C for magnetite, Fe_3O_4, is 1076°F (580°C) and for hematite, Fe_2O_3, 1238°F (670°C)]. This equilibrium TRM is later "frozen" in the rock at the blocking temperature T_B, whereupon the magnetization vector is no longer in equilibrium with the Earth's magnetic field and can theoretically remain stable for periods longer than 10^{11} years. Blocking temperatures for many magnetite-bearing rocks range from 930 to 1080°F (500 to 580°C). If a preexisting rock is later heated to temperatures near or exceeding T_B, it may acquire a new magnetization upon cooling. *See* CURIE TEMPERATURE; IGNEOUS ROCKS.

When sediment accumulates in a basin, magnetic grains sink through the water column and accumulate at the sediment-water interface together with other sedimentary grains. There, they may become aligned with the ambient magnetic field, and a detrital remanent magnetization (DRM) is the result. As sediment is transformed into sedimentary rock by diagenesis and lithification, a previously acquired DRM can be modified by compaction and dewatering, and an additional chemical remanent magnetization (CRM) may be acquired by the rock. CRM is acquired by the formation of new magnetic minerals as they grow through a critical grain size (about 0.2 micrometer for magnetite, Fe_2O_3). Chemical remanent magnetization may also be acquired at higher temperatures during alteration of plutonic, volcanic, and metamorphic rocks. *See* SEDIMENTARY ROCKS.

An original TRM, DRM, or CRM acquired by a rock when it formed may be partially or completely overprinted by a younger TRM or CRM in a subsequent geologic event such as metamorphism or hydrothermal alteration. Rocks may also be overprinted by viscous remanent magnetization (VRM), the statistical thermal realignment of the magnetization of smaller grains parallel to a younger field direction. At the surface, lightning strikes create large electric

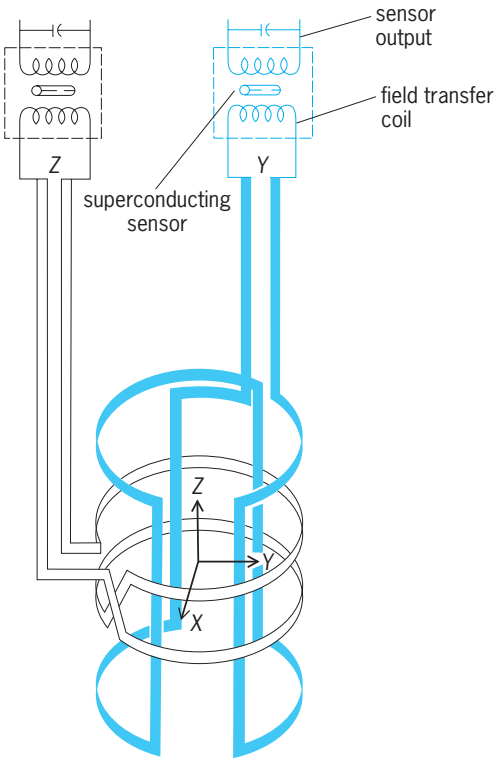

Fig. 2. Measuring head of the SQUID magnetometer (two-axis system).

currents that can impart a spurious isothermal remanent magnetization (IRM). Various demagnetization techniques are employed to selectively remove these undesired magnetizations, isolating the older magnetizations residing in the rock.

Collection and measurement of samples. A sequence of carefully oriented samples is collected, spanning a time interval long enough to average geomagnetic secular variation (for a paleomagnetic study) or spanning a stratigraphic interval (for a magnetostratigraphic study). The samples are cut into small cylinders whose magnetization is measured by using a sensitive magnetometer. Modern instruments use fluxgate or SQUID (Superconducting Quantum Interference Device) detectors with very high sensitivity and low noise (**Fig. 2**). *See* SQUID.

Resolution of the sample's magnetization vectors is accomplished using partial demagnetization techniques. In alternating field demagnetization, the sample is subjected to a sinusoidal alternating magnetic field of smoothly decreasing amplitude within a region of near-zero magnetic field. This procedure progressively randomizes the magnetic grains whose magnetic coercive force is less than or equal to the peak alternating field intensity. These treatments are useful in removing IRM and in resolving one generation of TRM from another. In thermal demagnetization, the sample is heated and then cooled within a region of near-zero magnetic field. By heating the rock to progressively higher temperatures, magnetizations with differing blocking temperatures can be selectively removed. Thermal demagnetization is useful for isolating multiple TRM vector directions and for removing recent VRM. Chemical demagnetization is done by immersing rock samples in concentrated hydrochloric acid for periods of days to weeks. The acid progressively dissolves magnetic grains from the outside in, generally the reverse order in which the CRM was acquired. Chemical demagnetization is useful in separating multiple CRM components, and CRM and DRM in sedimentary rocks. Low-temperature treatments by immersion in a cryofluid such as liquid nitrogen are occasionally applied; this procedure tends to remove the magnetization of large, multidomain grains.

Interpretation. The end product of the laboratory experiments is a suite of magnetization vector directions defined by the inclination I, the angle that the magnetization vector makes with the horizontal, and the declination D, the angle that the projection of the magnetization vector upon a horizontal plane makes with true north. Provided that the samples have recorded the geomagnetic field for 10^4–10^5 years, their collective directions will average geomagnetic secular variation, and representative mean D and I values and an associated uncertainty in direction may be calculated by using statistical techniques. The mean declination and inclination, together with the inclination-latitude relationship, mentioned previously, and some elementary spherical trigonometry, allow the calculation of a representative paleomagnetic pole from the samples. By connecting paleomagnetic poles of different ages in an ordered time sequence, an apparent polar wander path (APWP) may be constructed for a tectonic plate (**Fig. 3**). The APWP specifies the displacement

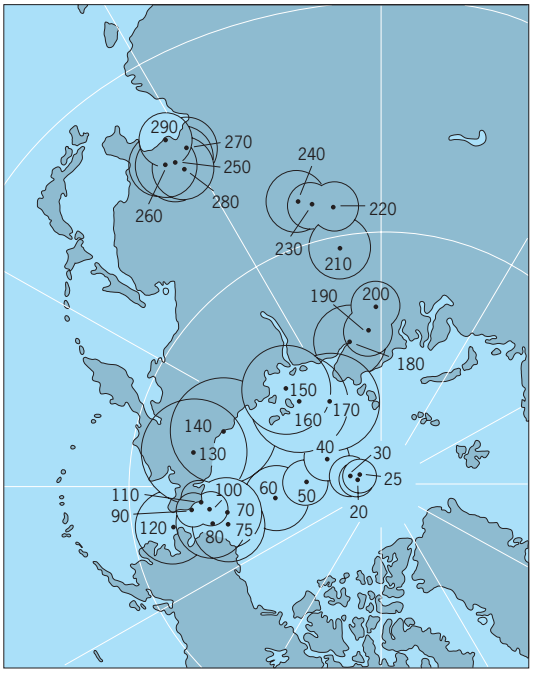

Fig. 3. Apparent polar wander path for North America for Late Carboniferous to recent time. The numbers represent time in millions of years before present. Circles represent one standard error about the mean. (*After E. Irving, Paleopoles and paleolatitudes of North America and speculations about displaced terranes, Can. J. Earth Sci., 16:669–694, 1979*)

history of a plate or continent with respect to the spin axis, and can be directly compared with APWPs from other plates to determine whether relative movements have occurred. For times before about 2.5×10^8 years before present, APWP information is the only direct method of obtaining relative motion information for the continents. Differing APWPs from North America and Europe demonstrated that sea-floor spreading had occurred in the northern Atlantic, hastening the acceptance of continental drift by the earth science community and presaging the development of plate tectonics.

The end product of a magnetostratigraphic study is an ordered set of normal and reversed magnetizations from a rock sequence. By comparing N-to-R and R-to-N transitions in the rock sequence with a reference polarity time scale, and by incorporating other geologic information, magnetostratigraphy makes it possible to correlate rock sequences over intercontinental distances. A reliable polarity time scale has been constructed for about the past 1.5×10^8 years. Together with the apparent polar wander path records from continents, this time scale has proved invaluable in developing plate tectonic theory to its present advanced state. While extension of magnetostratigraphy to earlier times is theoretically possible, the technique is proportionally less accurate because uncertainties in the radiometric calibration of the reversal time scale can exceed the length of at least some of the polarity intervals themselves. *See* GEOMAGNETISM; GEOPHYSICAL EXPLORATION. Michael McWilliams

Bibliography. R. F. Butler, *Paleomagnetism: Magnetic Domains to Geologic Terranes*, 1992; M. W. McElhinny, and P. L. McFadden, *Paleomagnetism: Continents and Oceans*, 2000; R. T. Merrill, M. W. McElhinny, and P. L. McFadden, *The Magnetic Field of the Earth: Paleomagnetism, the Core and the Deep Mantle*, 1996; N. D. Opdyke and J. E. T. Channell, *Magnetic Stratigraphy*, 1996; L. Tauxe, *Paleomagnetic Principles and Practice*, 1998.

Paleontology

The study of animal history as recorded by fossil remains.

The fossil record includes a very diverse class of objects ranging from molds of microscopic bacteria in rocks more than 3×10^9 years old to unaltered bones of fossil humans in ice-age gravel beds formed only a few thousand years ago. Quality of preservation ranges from the occasional occurrence of soft parts (skin and feathers, for example) to barely decipherable impressions made by shells in soft mud that later hardened to rock. *See* FOSSIL; MICROPALEONTOLOGY.

The most common fossils are hard parts of various animal groups. Thus the fossil record is not an accurate account of the complete spectrum of ancient life but is biased in overrepresenting those forms with shells or skeletons. Fossilized worms are extremely rare, but it is not valid to make the supposition that worms were any less common in the geologic past than they are now. *See* EDIACARAN BIOTA.

The data of paleontology consist not only of the parts of organisms but also of records of their activities: tracks, trails, and burrows. Dinosaur footprints, for example, are very common in the Connecticut Valley. Even chemical compounds formed only by organisms can, if extracted from ancient rocks, be considered as part of the fossil record. Artifacts made by people, however, are not termed fossils, for these constitute the data of the related science of archeology, the study of human civilizations. *See* ARCHEOLOGY; PALEOBIOCHEMISTRY.

Paleontology lies on the boundary between two disciplines, biology and geology. Various scientists have tried to place it more firmly in one camp than in the other, but such designations emphasize only a small area of the science's domain and do not acknowledge its entire range. If all types of paleontological research are regarded as equal in importance, then the geological and biological aspects must be granted equal weight. *See* BIOLOGY; GEOLOGY.

Geological Aspects

A major task of any historical science, such as geology, is to arrange events in a time sequence and to describe them as fully as possible. Geology, the study of Earth history, did not become a modern science until the nineteenth century, when a worldwide time scale based on fossils was established; Earth history could not be deciphered until events that occurred in different places were related to one another by their position in a standard time sequence (**Fig. 1**). The data provided by fossils are used to accomplish these tasks in the following ways.

Chronology. Many geologists of the eighteenth century tried to use rock type as a criterion for judging the relative age of rocks, saying that heavy rocks such as granites were older than lighter sandstones and limestones. But granites formed throughout Earth history and Cenozoic granites cannot be distinguished from Cambrian granites by mineral content. A species of organisms, however, is unique; it lives for only a short time before becoming extinct or evolving into something else, and once gone, it never reappears. Thus fossils can be used to correlate rocks; for if species A is found in both a Nevada limestone and a Colorado sandstone, it may be concluded that the rocks are approximately the same age (**Fig. 2**). Fossils of certain species, or index fossils, are ideally suited for correlation because they were abundantly distributed over a large geographic area, yet persisted for a very short time before becoming extinct. *See* INDEX FOSSIL; STRATIGRAPHY.

Of course, fossils only tell that a rock is older or younger than another; they do not give absolute age. The decay of radioactive minerals may provide an age in years, but this method is expensive and time-consuming, and cannot always be applied since most rocks lack suitable radioactive minerals. Correlation by fossils remains the standard method for comparing ages of events in different areas.

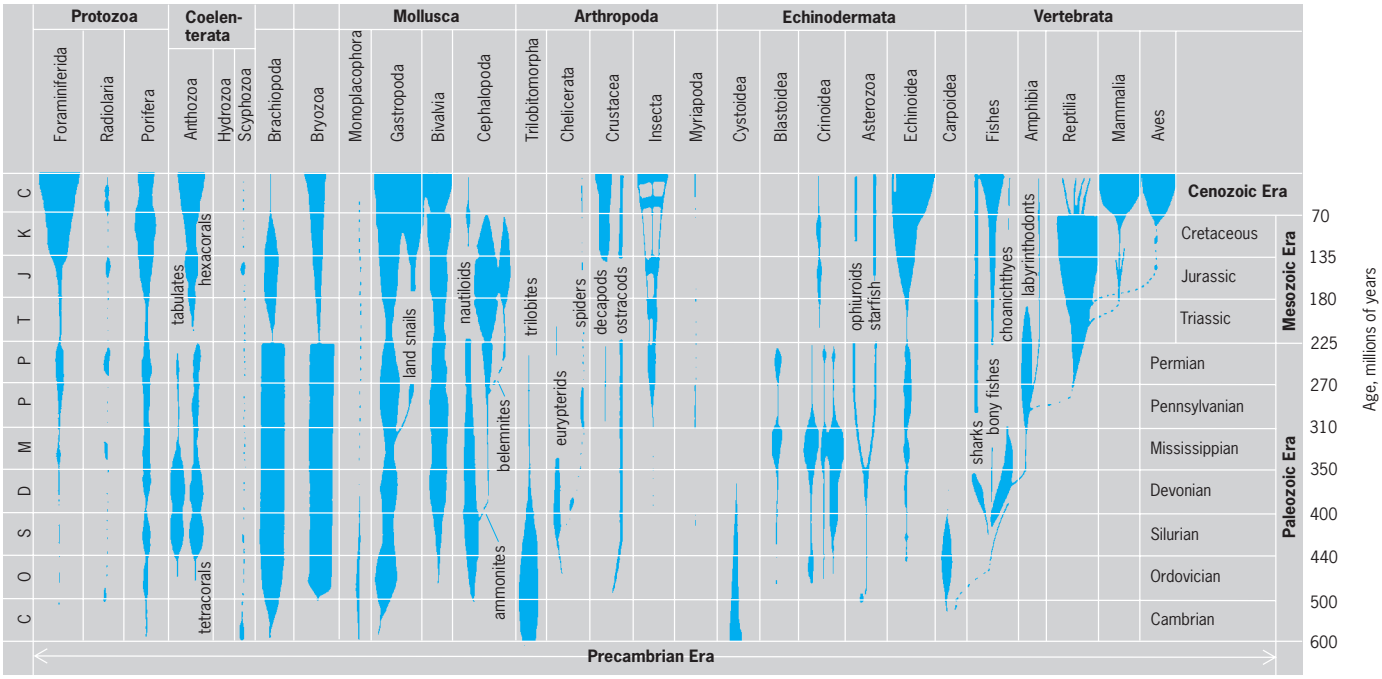

Fig. 1. Geologic distribution of animal life since the beginning of the Cambrian Period.

Determining ancient Earth's appearance. The physical appearance and climate of the Earth during a given period of the geologic past can be described from compilation and analysis of the data which is obtained through studies of the habitats of extant fauna, the geographic distribution of fossils, and the climatic preferences of ancient forms of life.

Local conditions. Life habits of fossils can be inferred by comparing them with modern organisms to which they are closely related. For example, all modern sea urchins live in marine environments. Since the Cretaceous limestones of southern England contain numerous fossil sea urchins, the conclusion is that this area was covered by ocean waters when these rocks were deposited. The study of fossils in relation to their physical and biological environment is called paleoecology. *See* PALEOECOLOGY.

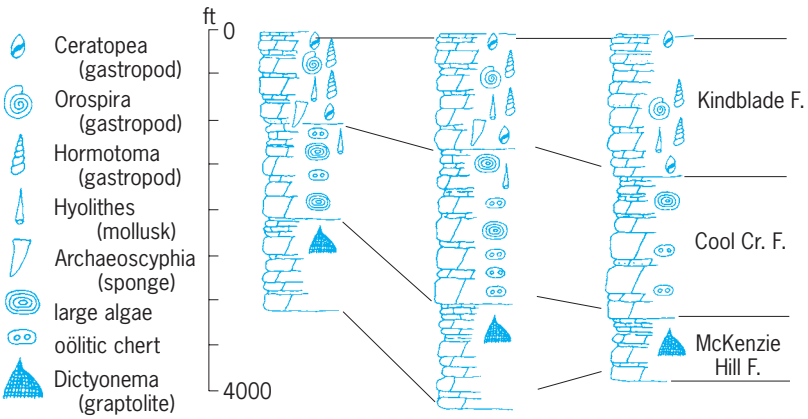

Fig. 2. Correlation of rocks by means of fossils. Although these three rock columns of Ordovician rocks in southern Oklahoma are all very similar in lithology, they can be divided and correlated on the basis of fossils. 1 ft = 0.3 m.

Paleogeography. Fossils are indispensable guides for determining the positions of continents and seas in former times. The formation of the Isthmus of Panama can be dated, for example, by studying the distributions of marine and terrestrial fossils. Before North and South America were connected by this land bridge, Atlantic and Pacific marine faunas were very similar but the mammals of the two continents were completely different. However, in South American rocks deposited after the Isthmus was formed, there are fossils of North American mammals which had migrated over the newly formed land. Likewise, Atlantic and Pacific marine faunas, isolated from each other by the rise of the Isthmus, began to evolve in different directions; this increasing difference can be traced in the fossils of successively younger rocks. Fossils also help the scientist to decide whether continental drift has occurred. If South America and Africa were once united, their faunas should be similar during that time. As they drifted apart, there should be stronger and stronger faunal differences. *See* PALEOGEOGRAPHY.

Paleoclimatology. The natural occurrence of polar bears always implies a cold climate. In the same way, it is possible to learn about ancient temperatures if the climatic preferences of fossilized organisms are known. There is, for example, a one-celled animal which tends to coil its shell one way when the water temperature is cold and the other way when it is warmer.

During the last million years, the oceans have been successively cooled and warmed as giant glaciers of the ice ages grew and melted. The most recent warming of the oceans (melting of the last continental ice sheet) can be dated by finding the age of the sediments in which the change of the fossil's coiling direction is noted. *See* GEOLOGIC THERMOMETRY.

Biological Aspects

Paleontology as a branch of the life sciences may be called the biology of fossils. Any technique applied by biologists to the study of living organisms is theoretically relevant to fossils, although many cannot be applied in practice because of the limitations of paleontological evidence. The rarity of preserved soft parts, for example, severely limits the possibilities for a physiology of fossil animals. *See* FOSSIL.

The most fundamental fact of paleontology is that organisms have changed throughout earth history and that each geological period has had its characteristic forms of life. Before Charles Darwin won acceptance for his theory of evolution, various explanations were offered for the differences between ancient and modern life. Some claimed that fossils were not organic remains, but manifestations of a plastic force in rocks; others attributed fossils to the devastation of Noah's flood; still others envisaged a whole series of catastrophes that destroyed life and led to the creation of an entirely new set of organisms. When Darwin published his theory of evolution in 1859, paleontological evidence for evolution was very meager but it began to accumulate quickly. The "feathered reptile" *Archaeopteryx*, an almost perfect link between reptiles and birds, was discovered early in the 1860s and the first fossil humans were found a few years later. In the light of this evidence and the purview of a more scientific culture, evolution was accepted as the only reasonable explanation for change in the history of life. To a paleontologist, life's outstanding attribute is its evolution.

Evolutionary process and life history. An evolutionist has two major interests: first, to know how the process of evolution works; this is accomplished by studying the genetics and population structure of modern organisms; second, to reconstruct the events produced by this process, that is, to trace the history of life. This is the paleontologist's exclusive domain. Any modern animal group is merely a stage, frozen at one moment in time, of a dynamic, evolving lineage. Fossils give the only direct evidence of previous stages in these lineages. Horses and rhinoceroses, for example, are very different animals today, but the fossil history of both groups is traced to a single ancestral species that lived early in the Cenozoic Era. From such evidence, a tree of life can be constructed whereby the relationships among organisms can be understood.

Life properties. Evolution is responsible for life's outstanding properties: its diversity and its adaptation to environment.

Diversity. A paleontologist studies diversity by classifying organisms into basic units known as species. The number and geographic distribution of species in a more inclusive taxonomic group, such as vertebrates, can then be tabulated for different periods of geologic history. When this is done for all major groups, very definite patterns emerge. At the close of the Permian and Cretaceous periods, for example, extensive extinction of species occurred in many major groups at about the same time. Although the cause for these mass extinctions is unknown (theories range from large-scale shifts of sea level to bursts of cosmic radiation), they are major events in the history of life.

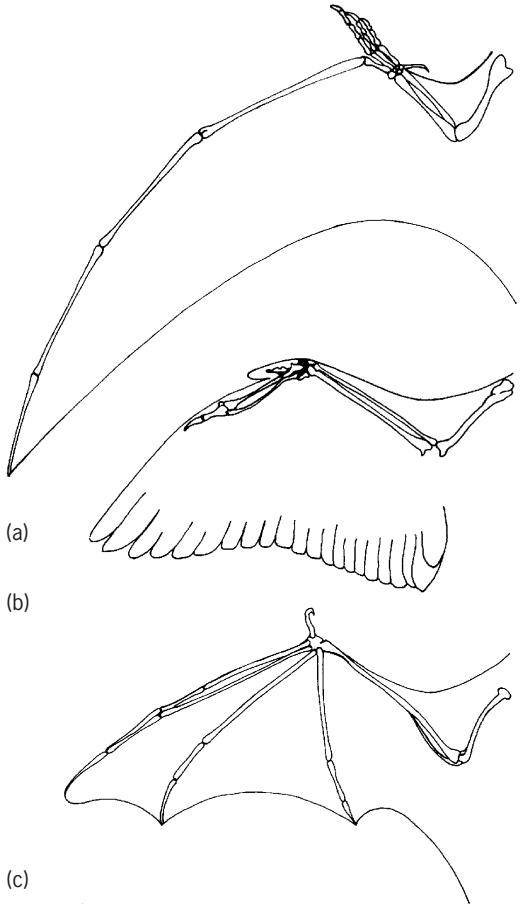

Fig. 3. Adaptation. Three groups of vertebrates independently developed wings for flight. (*a*) Pterodactyl (extinct reptilian pterosaurs). (*b*) Bird. (*c*) Bat.

Adaptation. The term adaptation refers to the way organic form is fitted to function (**Fig. 3**). Given the raw materials of life, an engineer could scarcely design a better flying machine than a bird, yet the bird evolved naturally from reptilian ancestors which did not fly. Paleontologists study the development of new adaptations by a variety of techniques. For example, consider the increase in brain size between apelike ancestors and modern humans. To study this a paleontologist employs the following: (1) studying the rocks in which fossil bones are found to look for signs of any environmental change which might have made increased brain size advantageous; (2) determining the advantages of a larger brain if possible; for example, it may, among other things, have conferred upon ancient humans the ability to learn to control fire for warmth and cooking (charred bones of edible animals in caves inhabited by ancient humans could be used as evidence for such an assertion); and (3) applying mathematical techniques to obtain a precise measurement of the rate of evolution. *See* EXTINCTION (BIOLOGY); SPECIATION.

Capsule History of Life

Although life was not present on Earth when it first formed about 4.5×10^9 years ago, the raw materials necessary for its natural development were available. Electrical discharges and ultraviolet radiation induced the components of the original atmosphere to form complex organic compounds which later combined to form the prototype of a living cell. The oldest recognizable organisms so far discovered are bacterialike cells from Australian cherts dated at 3.5×10^9 years old. By the time the Gunflint cherts of Canada were deposited 2×10^9 years ago, complex algae had evolved. Except for a few forms found in rocks just slightly older than 6×10^8 years, invertebrate animals are not found until the base of the Cambrian Period, 6×10^8 years ago. It is not known why so many groups of invertebrates made their first appearance as fossils in rocks at the same age; perhaps the base of the Cambrian marks a time during which animals first developed hard parts. The first vertebrate fossils, primitive fishes, occur in Lower Ordovician rocks. Vertebrates invaded the land with the evolution of amphibians in the Late Devonian, about 3.5×10^8 years ago. Reptiles evolved soon afterward in Mississippian times; the first true mammals are Jurassic in age. The human is a newcomer, a product of the last few million years. Humans are merely a single species, a natural product of an unplanned evolution; yet from a geological perspective, they have altered the Earth far more than it had changed in any comparable time since life evolved. *See* ANIMAL EVOLUTION; PALEOBOTANY. Stephen J. Gould

Bibliography. E. H. Colbert, *Evolution of the Vertebrates*, 4th ed., 1991; R. Cowan, *History of Life*, 1991; A. Hallam (ed.), *Patterns of Evolution: As Illustrated by the Fossil Record*, 1977; L. B. Halstead, *Search for the Past*, 1983; M. McKinney, *Evolution of Life*, 1993; R. C. Moore, C. G. Lalicker, and A. G. Fischer, *Invertebrate Fossils*, 1952.

Paleopathology

The study of ancient diseases and their origins. Paleopathology is especially important in the understanding of the origins, prevalence, and spread of infectious diseases, including how humans have contributed to the spread of disease and how they can overcome it. *See* EPIDEMIOLOGY; INFECTIOUS DISEASE.

Hypothesis testing of populations has contributed to the field of paleopathology, as has application of macroscopic (visual) examination, routine x-ray, computerized tomography (CAT) scans, magnetic resonance imaging (MRI), electron microscopy, and immunologic, chemical, and mass spectrophotometry techniques to skeletons, soft tissue, and even scat (animal droppings).

Basic Principles and Methods

The scientific method in paleopathology is based upon comparison of archeologic or paleontologic findings with individuals documented to have the disease. To this end the following basic tenets are observed: (1) Tissue must be adequately preserved to allow recognition of disease and distinguish possible pseudopathology or postdeath artifact. (2) The manifestations of a disease must be sufficiently stable across generations to allow comparison of ancient with modern disease. (3) Analysis of entire skeletons is more accurate than analysis of isolated bones. (4) Analysis of afflicted populations (paleoepidemiology) is more accurate than analysis of isolated skeletons.

The range of diagnostic methods used in paleopathology is extensive. Skeletal remains are visually examined to identify occurrence and nature of alterations, mapping their skeletal distribution. Internal structure can then be assessed, preferably by a nondestructive technique. Even fossils are not simply casts of external surfaces, but have a visualizable internal structure.

Microscopes can look at the surface of tissue or at cut sections. Traditional microscopy often requires tissue sections thin enough to allow light to pass through the specimen; another technique bounces the light off the specimen, so destructive thin sections are not required. Preservation of microscopic details has long been recognized in dinosaurs, as well as in the preserved muscles of fish as old as 300 million years. *See* X-RAY MICROSCOPE.

Molecules can be identified in tissue using antibodies or even DNA amplification and analysis, and chemicals can be identified by x-ray diffraction. Molecular preservation extends even to the three-dimensional structure of molecules, such as have been identified in 10,000-year-old collagen from birds and mammals and in blood from *Tyrannosaurus*. *See* ANTIBODY; X-RAY DIFFRACTION.

Rudolf Virchow and Charles Lester Leonard x-rayed mummies as long ago as 1897. As x-ray technique has been refined, computerized tomography and magnetic resonance imaging have also found application. Three-dimensional reformatting has allowed examination of internal surfaces of fossilized skeletons. Computerized tomography scanning with "dissection" of the resulting image and three-dimensional reconstruction can illustrate internal structures, and a plastic model can even be generated. *See* COMPUTERIZED TOMOGRAPHY; MAGNETIC RESONANCE.

Mummies provide an additional source of information. Rehydration of mummy tissue allows standard soft tissue histology, providing information often transcending that available through study of bones. Anthropologic study of artifacts such as daggers that sometimes accompany mummies and skeletons has also contributed to the understanding of ancient lifestyles and the diseases which impacted them. *See* HISTOLOGY.

Diseases of the Past

Following is a review of some of the diseases recognized in living individuals that are believed to have existed as well in ancient life forms.

Dental. Dental caries are an ancient phenomenon, with abscesses present as far back as Devonian lungfish. However, such appear to have been rare occurrences until the introduction of agriculture in human populations. *See* DENTAL CARIES.

Trauma. The oldest example of a broken limb appears to be the radius of the sail-backed Permian lizard, *Dimetrodon*. Fractures are noted only rarely in flying reptiles (for example, *Pteranodon*), suggesting that they were rare or incompatible with survival.

Evidence of injuries among plant-eating dinosaurs is rare. Analysis of 30,000 elements from horned dinosaurs revealed a frequency in the range of 0.025–1.0%. Most were rib fractures, apparently related to side-butting territorial behavior. The most common injuries in duck-billed dinosaurs was to the tall projections off the back of vertebrae, the neural spines, perhaps related to mating activities. Large meateating dinosaurs (for example, *Tyrannosaurus*) seemed to have a much higher frequency of fracture.

Stress fractures related to activity, as have been reported in humans and greyhounds, have also been noted in the toes of horned dinosaurs.

The Neandertal Shanidar I had multiple fractures. The frequency of such fractures has been used to understand ancient "lifestyle." *See* NEANDERTALS.

Vascular phenomena. Death, or necrosis, of bone occurs when its blood supply is compromised. Such has been described in Cretaceous marine lizards and turtles, although its occurrence in marine turtles seems to have diminished in frequency in the early Eocene and nearly disappeared subsequent to the Oligocene. *See* GEOLOGIC TIME SCALE.

Bone necrosis in the mosasaurs, a family of marine lizards, was universal in the deep and repetitive diving genera, yet absent in those whose members remained close to the water surface.

Infections. Suppurative (pus-producing) infections of the postcranial skeleton are recognized by irregular bone destruction and associated draining sinuses. Criteria have also been developed to recognize granulomatous infections such as fungus, tuberculosis, and leprosy, and even for treponemal disease (for example, syphilis). *See* INFECTION.

Tuberculosis. Tuberculosis accompanied by vertebral collapse and fusion, preserving the posterior joints, has been recognized on both sides of the Atlantic dating up to 6000 years before present (BP). Its diagnosis in mummies has been validated by recognition of the causative organism on microscopic examination, as well as by detection of *Mycobacterium tuberculosis* DNA. However, as bone involvement is relatively rare in tuberculosis and as vertebral collapse and fusion represent only a small portion of tubercular bone disease, only a small proportion of individuals afflicted can be recognized in this way. *See* TUBERCULOSIS.

Leprosy. Leprosy, or Hansen's disease, is an infectious disease primarily affecting the skin and peripheral nerves. It is diagnosed on the basis of changes to the tips of fingers and toes, as well as fractures and other defects in the joints. Leprosy has been quite controversial in history, due to imprecise descriptions in biblical and ancient manuscripts. V. Moller-Christensen reported over 300 cases from Medieval hospital cemeteries in Denmark and Sweden. He suggested that 20% of individuals buried in those countries from 800 to 500 years BP had leprosy, but noted it only in isolated cases elsewhere (two dated at 1400 years BP from Egypt, one from France dated 1400 years BP, and 5 from the British Isles dated 1400 years BP). *See* LEPROSY.

Fig. 1. Two views of tibia (lower leg) affected with bejel (nonvenereal syphilis) from nineteenth-century Israel. (*a*) Anterior bowing of tibia with periosteal reaction on surface of both tibia and fibula. (*b*) Anterior and lateral view of tibia with bony enlargement and draining sinuses.

Treponematoses. The treponematoses (caused by the spirochete, or treponeme, type of bacteria) include three clinically distinct disorders known to afflict bone: syphilis, yaws, and bejel (nonvenereal syphilis) [**Fig. 1**]. All three diseases cause skin rashes and bone damage. Syphilis is generally transmitted by sexual contact, while yaws and bejel are contracted by skin contact or sharing food utensils. While syphilis can affect the fetus, yaws and bejel are generally contracted in the first decade of life. Bone changes of syphilis are rare in children (less than 5%), in contrast to yaws and bejel, in which 10–20% of children have bone changes. Syphilis can produce heart, aorta, and nervous system damage, while these systems are generally spared in yaws and bejel.

Syphilis has been of such interest that three theories have been proposed for its occurrence: (1) Old World origin (pre-Columbian hypothesis); (2) New World origin (Columbian hypothesis); and (3) independent New and Old World origin.

The pre-Columbian hypothesis suggests that the epidemic of syphilis in late fifteenth century–early sixteenth century Europe simply represented new diagnostic ability to distinguish syphilis from leprosy. E. H. Hudson suggested an Old World origin of treponemal disease, in the form of yaws, that was subsequently transmitted to the New World.

Patterns reproducible for syphilis have been identified in Michigan and West Virginia (dated at 600 years BP), in Florida (700 years BP), in Ecuador

(800 years BP), in Wisconsin (1000 years BP), and in New Mexico (1500 years BP), but in no pre-Columbian European, African, or Asian sites examined. The worldwide presence of yaws, on the other hand, the geographic separation of syphilis and yaws in the New World, and the chronologic progression to recognizable syphilis (as a regional phenomenon) suggest that syphilis derived from yaws.

The history of treponemal disease in the New World can be definitively traced back almost 8000 years. It has been suggested that yaws migrated to the New World with the first humans to enter North America from Asia.

The first recognized case of treponemal disease appears to have been a case of yaws dated at 1.6 million years BP, suggesting that the onset of treponemal disease coincided with the origins of humans in Africa. *See* SYPHILIS; YAWS.

Parasitic diseases. Parasites, the subjects of the specialized field of paleoparasitology, are well represented in the archeologic record. A fish tapeworm dated at 6000 years BP has been identified in Peru and Chile; dog tapeworm in Medieval England and North Dakota (700 years BP); flukes in coprolytes (fossilized feces) from the southwestern United States (700 years BP) and Medieval Germany; roundworms in South American mummies (4000 years BP) and in the southwestern United States (1500 years BP); hookworms in Peru (1100 years BP), Brazil (5000 years BP), and Tennessee (2200 years BP); whipworm in Arizona (1000 years BP), from pre-Columbian Peru, Ecuador, and Brazil, and Japan (1000 years BP); schistosomiasis in the Sudan (2000 years BP); trichinella in mummies from the Aleutians (500 years BP), Viking-era Greenland, and Medieval Denmark and Switzerland; and Chagas' disease in Chile (4000 years BP). *See* PARASITOLOGY.

Osteoarthritis. Osteoarthritis is a disorder in which bone spurs occur at joints. The underlying bone becomes denser and the space between joints narrows. A type of osteoarthritis characterized as trauma-related degenerative disease has been noted in Neandertals, but osteoarthritis is more frequently a phenomenon of aging. Osteoarthritis is commonly observed in contemporary humans, but only within the past 50 years has it been routinely separated from other forms of arthritis. Thus, clinical records do not allow estimates of its frequency earlier than the most recent past.

As the presence of bone spurs defines the disease, historic occurrences of osteoarthritis are easily discernible by examination of skeletons from cemeteries or mass burials. As joint space–narrowing defines severity, occurrences are harder to measure. If the disease is so severe as to eliminate joint cartilage, bone rubs on bone. This polishes the bone, a condition that can be easily recognized in skeletons. As bone polishing occurs also in other diseases, however, only that associated with spurs indicates the most severe form of osteoarthritis.

Osteoarthritis actually appears to be a relatively new disease. Among dinosaurs it is extremely rare, documented only in *Iguanodon*, affecting the ankles of 2 of 39 individuals examined. This contrasts with the frequent occurrence of osteoarthritis in the flying reptiles, pterosaurs. Osteoarthritis in Pleistocene mammals is quite rare, similar to the frequency observed in contemporary wild caught animals.

The paucity of osteoarthritis in large dinosaurs provides evidence that weight is not the major consideration in its development. Dinosaurs had highly constrained joints. They were not capable of the rotation inherent in the human knee. Thus, they were protected. This finding affirms that it is the stability of the joint, not the weight placed upon it, which determines if osteoarthritis will occur. *See* ARTHRITIS.

Rheumatoid arthritis. Rheumatoid is an inflammatory arthritis that afflicts most of the joints of the body symmetrically but spares the vertebrae (except neck) and sacroiliac (lower back) joints. It originated in the New World. No valid cases earlier than 1785 have been found in the Old World. Rheumatoid arthritis has been found in numerous pre-Columbian New World populations, dating back 6500 years, and the character of this disease has remained unchanged over the time span. Originally found in the Green River region of west-central Kentucky and the west branch of the Tennessee River in northwest Alabama and Tennessee, it remained there for 5000 years, spread into Ohio approximately 1000 years ago, and became more disseminated 200–300 years ago. No valid animal example of rheumatoid arthritis has been identified to date.

The initial geographic localization of the disease, with its subsequent spread in the New World and finally its penetration into the Old World, is highly suggestive of a vector-transmitted disease. Identification of such an agent may allow prevention and possibly specific treatment of the disease in the future.

Spondyloarthropathy. Another form of erosive arthritis commonly produces a more limited peripheral arthritis, a different variety of joint erosion, and joint fusion, and at times affects the vertebrae and sacroiliac joints (**Fig. 2**). There are a number of varieties of disease in this category. Ankylosing spondylitis, psoriatic arthritis, and reactive arthritis are the

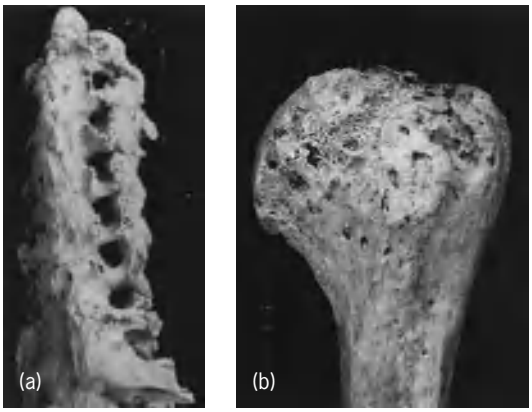

Fig. 2. Spondyloarthropathy. (*a*) Cervical spine (neck vertebra) fusion and (*b*) shoulder (upper arm) erosions affecting both margins and joint surface in individuals from early-nineteenth-century Rochester Poorhouse.

most prevalent. While frequency varies in world populations, the nature and skeletal distribution are quite similar through time and even across species lines. Its frequency in afflicted human populations runs 0.5–8%, seemingly dependent upon sanitary conditions, although some Amerindian groups seem to have a genetic predisposition. The oldest human case reported in the New World so far was dated at 5000–8000 years BP. *See* JOINT DISORDERS.

Spondyloarthropathy is believed to occur in 25% of contemporary bears, apes, and baboons, whose antecedents were largely spared. While the disease was undetected in Miocene apes, the frequency of spondyloarthropathy increased to 20–28% of contemporary gorillas and chimpanzees. While it was undetected in Miocene baboons, its frequency increased from 4% in the early part of the twentieth century to 10% at midcentury and 30% in the 1980s. Spondyloarthropathy truly has become epidemic in both free-ranging and colony-raised animals, and now appears pandemic in mammalian orders including carnivores such as cats and dogs, hoofed animals, and even marsupials such as kangaroos.

Metabolic diseases. Metabolic diseases involve defects of the metabolic pathway that result in an imbalance in nucleic acids, proteins, lipids, or carbohydrates.

Gout. Gout, first reported by Hippocrates, is a disorder in which urate crystals accumulate as space-occupying masses in the blood. Gout has only rarely been reported in ancient humans, but it has been reported in monitor lizards, turtles, alligators and crocodiles, birds, and most recently in *Tyrannosaurus rex*. *See* GOUT.

Osteoporosis. Osteoporosis is a disease in which bones become brittle and likely to break, usually due to a loss of bone density. Unfortunately, modern techniques for measurement of bone density are of limited value for ancient skeletons. Variable effects of diagenesis (the chemical and physical changes undergone by buried bones with increasing time, pressure, and temperature) generally preclude direct application of density-measuring techniques. One indirect technique uses x-rays to measure the cortical thickness of hand bones. This seems an effective technique, although it has yet to be applied in routine population studies. *See* OSTEOPOROSIS.

Osteomalacia and rickets. These diseases, which cause bowing of bones with widened growth plates, are uncommon in the paleontologic record. A case of rickets from the Upper Paleolithic (1200 BCE) of Italy has been reported, as have cases in children from Bahrain 2000 years BP and from sixth-century Negev. An adult case in Austria 1600 years BP has also been reported. *See* SKELETAL SYSTEM DISORDERS.

Anemia. Examination of skeletons may allow recognition of certain forms of congenital anemias (lack of hemoglobin or red blood cells), as well as suggesting those due to the blood loss and other effects of parasitism. A possible case of thalassemia (related to abnormal hemoglobin) in thirteenth-century Greece was reported, as well as cases from late Bronze age Greece and a Khok Phanom Di site 4000 years BP in central Thailand. A DNA abnormality associated with anemia was demonstrated in a child from the early Ottoman period, and possible sickled cells were reported in skeletons dated 150–330 years BP. *See* ANEMIA.

There has been a misconception that localized overgrowth of skull bone marrow, referred to as porotic hyperostosis, can occur as a complication of primary iron-deficiency anemia. Actually, blood loss can produce both iron-deficiency anemia and porotic hyperostosis. The challenge of determining which came first is always problematic when trying to derive relationships from observed associations.

Congenital disorders. Disorders of growth, duplication, or hyper- or hypodevelopment of skeletal elements have been noted as isolated phenomena in both contemporary and ancient animals. Coalitions of foot bones, for example, were noted in the Cretaceous marine lizard *Platecarpus*. Another intriguing isolated phenomenon is the stone baby. A small percentage of fetuses develop outside the uterus. Most die quickly and the body reabsorbs them. Sometimes, the body does not reabsorb the fetus. In those cases, the remains become somewhat mummified. Skeletonized and bound by thickened, calcified membranes, such a fetus—dated at 3100 years BP—was clearly identified in Kerr County, Texas, in 1991.

Spina bifida, failure of the posterior vertebral elements to fully fuse and surround the nerves passing though the sacral canal, has attracted much anthropologic interest. Neural tube defects including spina bifida have been reported from sites in Florida (8000 years BP), Mali (7000 years BP), and northeastern Morocco (10,500–12,070 years BP). *See* CONGENITAL ANOMALIES.

Tumors. One of the simplest forms of neoplasia (excessive tissue production) is an overgrowth of bone called an exostosis. One of the earliest hominid examples is occurrence of an exostosis in the femur of *Homo erectus* from Java. Exostosis has also been found on the scapula of the dinosaur *Allosaurus* and the mandible of *Triceratops*. Multiple hereditary exostoses was recognized in a human 500 years BP from southern Ontario.

A more complicated form of neoplasia actually develops a cartilaginous cap, similar to that seen in joints, but without an opposing bone. Such osteochondromas have been reported from Egypt 3500–2500 years BP and Hawaii 500 years BP. It has also been identified in 64% of Oligocene canids (doglike creatures).

Another form of exostosis affects the external auditory canal. An association with cold water is suggested, possibly related to diving activities. Such cases are of great antiquity, reported in Pleistocene human specimens Shanidar I, la Chapelle aux Saints, and skull X from Zhoudoukien, China, as well as in humans of the Neolithic of southwest Asia, Africa, and North and South America. *See* FOSSIL HUMANS.

Isolated reports include benign bone tumors in humans from England 2000 years BP and in ancient marine lizards, and multiple vertebral hemangioma

(vascular tumors) from humans in Egypt 1000 years BP. Hemangiomas have also been recognized in dinosaurs, including a specimen collected from Emery County, Utah. The fragmentary nature of the 10-pound specimen precluded identification of species or even definite bone element identification, but the general morphology suggested a classic hemangioma.

E. Strouhal reviewed malignant tumors (predominantly osteosarcomas) in the Old World, reporting 10 isolated cases from Egypt, Nubia, and Europe, the oldest being Neolithic. He identified osteolytic metastases in 30 individuals from Egypt, Nubia, Europe, and Asia, dating back to the Early Dynastic of Abu Simbel, Nubia. While the blood malignancy multiple myeloma has been suggested, none of these cases fulfill current criteria for this disease. Multiple myeloma still awaits documentation in the ancient record. *See* TUMOR. Bruce M. Rothschild

Bibliography. J. E. Buikstra and D. H. Ubelaker, Standards for Data Collection from Human Skeletal Remains, *Arkansas Archeol. Surv. Res. Ser.*, no. 44, 1994; M. G. Fiori and M. G. Nunzi, The earliest documented application of X-rays to examination of mummified remains and archaeological materials, *J. Roy. Soc. Med.*, 88:67–69, 1995; C. Greenblatt, *Digging for Pathogens: Ancient Emerging Diseases—Their Evolutionary, Anthropological and Archaeological Context*, Balaban Publishers, Rehovot, Israel, 1998; B. M. Rothschild et al., First European exposure to syphilis: The Dominican Republic at time of Columbian contact, *Clin. Infect. Dis.*, 31:936–941, 2000; B. M. Rothschild et al., Geographic distribution of rheumatoid arthritis in ancient North America: Implications for pathogenesis, *Semin. Arthritis Rheum.*, 22:181–187 1992; B. M. Rothschild and L. Martin, *Paleopathology: Disease in the Fossil Record*, CRC Press, London, 1993; B. M. Rothschild and C. Rothschild, Treponemal disease revisited: Skeletal discriminators for yaws, bejel, and venereal syphilis, *Clin. Infect. Dis.*, 20:1402–1408, 1995; B. M. Rothschild and R. J. Woods, Spondyloarthropathy in the Old World, *Semin. Arthritis. Rheum.*, 21:306–316, 1992; E. Strouhal, Malignant tumors in the Old World, *Paleopath. Newsl.*, 85:1–6, 1994; P. L. Thillaud and P. Charon, *Lesions Osteoarcheologiques: Recueil et Identification*, Kronos, Sceaux, France, 1994; R. A. Tyson, *Human Paleopathology and Related Subjects: An International Bibliography*, San Diego Museum of Man, 1997.

Paleoseismology

The study of geological evidence for past earthquakes. This is a scientific discipline that has contributed greatly to modern understanding of the nature of earthquakes. The patterns of earthquakes, in both space and time, evolve over centuries and millennia and cannot be discovered by modern instruments. Knowledge of these patterns is important for understanding the physics of earthquakes and for forecasting future destructive earthquakes.

In certain natural environments, the features related to ancient earthquakes are preserved in the landforms and superficial layers of the Earth's surface. Geologists use this paleoseismological evidence to extend the short historical and instrumental record of earthquakes into ancient centuries and millennia. Paleoseismological studies have been extended into the ancient past; they have clarified the earthquake record of many parts of the world, including the midcontinent and east coast of the United States, northern Africa, southern Europe, China, Japan, Indonesia, and New Zealand.

Sedimentary record of ancient earthquakes. In certain geological settings, the disruptions produced during large earthquakes become preserved in geological strata. These sediments, therefore, provide a record of the earthquake. In this respect, the sediments are like seismograms. Geologists study the disruptions of these sediments to gain a better understanding of the earthquake. The San Andreas Fault in southern California provides one of the clearest examples.

Most Californians live within 100 km (60 mi) of the San Andreas Fault. Along the fault, coastal California, including Los Angeles, Monterey, and San Diego, has slipped northwestward over 300 km (180 mi), relative to the rest of North America. Despite this large horizontal translation over millions of years, no more than a few meters has occurred along the fault in the two centuries of historical record. These small movements occurred primarily during two great earthquakes—the 1906 San Francisco earthquake in northern California and the 1857 Fort Tejon earthquake in the southern part of the state. *See* FAULT AND FAULT STRUCTURES.

With no more than the short historical and instrumental records, scientists had no clear understanding of the frequency of large, destructive earthquakes produced by the fault. Fortunately, paleoseismological studies in southern California have extended the known record of earthquakes back nearly two millennia. Northeast of Los Angeles, a record of 10 large earthquakes is preserved in peaty and sandy stream and marsh deposits. Faulted sediments record evidence of the great earthquake that occurred around 1480. Overlying layers were deposited after this strand of the fault broke, and thus are undisturbed.

Dating ancient earthquakes. In addition to determining the amount and type of motion along a fault in an ancient earthquake, it is necessary to determine the date. Radiocarbon dating is the most common of many methods. Precise dating of earthquakes along the San Andreas Fault has shown that earthquakes along one part of the fault have struck southern California about every 130 years on average. In the past, 10 large earthquakes occurred along the San Andreas Fault northeast of Los Angeles. The 1857 earthquake is known from historical records. The 1812 earthquake is a historical event, but was not known to have been generated by the San Andreas Fault until the completion of paleoseismologic studies of disturbed trees along the fault. The other quakes have been dated by the radiocarbon method. It is

curious that the intervals between the earthquakes have ranged between 44 and about 330 years. *See* RADIOCARBON DATING.

The reason for the large variation in the dates of the earthquakes is the subject of much debate. Such large natural variations make forecasts of future earthquakes uncertain. Knowledge of the average period between large earthquakes, however, has given Californians a more informed basis for earthquake preparedness. Based upon paleoseismic data from several localities along the fault, the southern 200 km (120 mi) of the fault, near Palm Springs and San Bernardino, are considered the most likely to produce the next great earthquake in California. *See* SEISMIC RISK.

Geomorphic record. The world's greatest earthquakes occur at subduction zones, where great slabs of oceanic lithosphere are sinking diagonally into the Earth's interior. One of the largest earthquakes of the twentieth century, of magnitude 9.2, was generated on the subduction zone of southern Alaska in 1964, when a huge slab of the Pacific Plate lurched suddenly downward about 30 m (100 ft) beneath Alaska. Southern coastal communities were devastated by the long and heavy shaking, by the seismic sea waves (tsunamis) that swept over them, and by permanent submergence of the land. No earthquake of this size had struck this portion of Alaska in the entire two centuries of recorded history. *See* PLATE TECTONICS; SUBDUCTION ZONES; TSUNAMI.

Studies of the effects of this devastating earthquake revealed paleoseismological evidence of previous great earthquakes in this region. There was clear evidence of earlier uplifts on a small island, far out to sea, that had risen 3.4 m (11 ft) in 1964. The flat-topped island exhibits a set of six concentric ancient shorelines, each a few meters higher than its seaward neighbor; these appear as small dark cliffs. Radiocarbon dating of the ancient shorelines revealed that the top of the island, now 50 m (165 ft) above sea level, first rose above the sea about 5000 years ago. In the past 5000 years, uplift occurred episodically, during six ancient earthquakes, 500 to 1500 years apart. The shoreline of early 1964, now also dry and 3.4 m (11 ft) high, forms a ring just inland from the modern shoreline.

Other than a few moderate earthquakes, the Cascadian subduction zone of the Pacific Northwest of the United States has been quiet during the period of historical record. Older building codes in Oregon reflect a historical lack of concern about earthquakes there. Along the coasts of Oregon and Washington, however, there are clear records of sudden and recurring submergence, a paleoseismic record that has been related to great earthquakes on the Cascadian subduction zone. In many coastal estuaries, geologists have found layers of peat intercalated with silt beds. The peats are composed of plants known to grow at certain elevations in the estuaries. The silts, carried to the sea by nearby rivers, are known to settle upon the floor of the estuary in deeper water. Yet, the paleoseismologists have found that the peats now reside well below the level at which they grew and were buried by the silts. Radiocarbon dating of the peats indicates that sudden submergence of the estuaries has occurred several times in the past few thousand years, most recently about 300 and about 1200 years ago. In some localities, geologists have found sand atop the peats, carried into the estuaries from the sea by tsunamis. Since the discovery of paleoseismic evidence for great earthquakes in the Pacific Northwest, much more civic attention has been focused upon seismic building codes and earthquake preparedness there.

The geological preservation of ancient earthquakes also has enabled scientists to compare modern earthquakes with those of the ancient past. In 1983, for example, a sparsely populated region of Idaho was struck by a magnitude-7.3 earthquake. Investigations after the earthquake revealed a fresh, 30-km-long (18-mi) fault scarp running along the western base of the lofty Lost River Range. During the earthquake, Borah Peak, which crowns the range, had jumped 2 m (7 ft) skyward. Inspection of the fresh escarpment produced during the earthquake revealed that it is surmounted by a more subdued, vegetated escarpment of nearly identical length and height. Excavations across this ancient fault scarp showed that it had formed during an event very similar to the earthquake of 1983 but about 5000 years earlier. This is one of several examples of what paleoseismologists call a characteristic earthquake. It appears, from such examples, that some earthquakes are nearly identical repetitions of their predecessors. If nature were always as regimented as this, the prediction of earthquakes would be far simpler. Unfortunately, many examples of irregular behavior of faults also exist. *See* EARTHQUAKE; SEISMOLOGY.

Kerry Sieh

Bibliography. U.S. Geological Survey, *Earthquakes and Volcanos*, bimonthly; R. E. Wallace (ed.), *The San Andreas Fault System, California*, USGS Prof. Pap. 1515, 1990; R. Wesson and R. E. Wallace, Predicting the next great earthquake in California, *Sci. Amer.*, 252:35–43, 1985.

Paleosol

A soil of the past, that is, a fossil soil. Paleosols are most easily recognized when they are buried by sediments. They also include surface profiles that are thought to have formed under very different conditions from those now prevailing, such as the deeply weathered tropical soils of Tertiary geological age that are widely exposed in desert regions of Africa and Australia. Such profiles are generally known as relict paleosols. Those that can be shown to have been buried and then uncovered by erosion are known as exhumed paleosols. The main problem in defining the term paleosol comes not so much from complications such as these arising from its fossil nature, but from defining what is meant by soil, a term that has very different meanings for agronomists, engineers, geologists, and soil scientists. Considering research on soils of Antarctica and Mars and

on paleosols in a variety of rocks ranging back to 3.5 ×nbsp;10^9 years old, soil can be considered distinct from sediment in that it forms in place, but soil need not necessarily include traces of life. At its most general level, soil is material forming the surface of a planet or similar body and altered in place from its parent material by physical, chemical, or biological processes.

Soils buried by till, colluvium, and flood deposits are commonly encountered during roadwork or excavations for foundations. The ways in which such soils can overlap or be eroded are complex (**Fig. 1**). Weakly developed soils may retain features of their parent material, such as relict sedimentary structures. They also may contain fragments or horizons of preexisting soils (pedorelicts). Especially distinctive soil materials, such as laterites, that have been eroded and redeposited are known as pedoliths. Despite these complications, many paleosols are unmistakable with their prominent root traces, gradational soil horizons, and soil structures, such as peds and cutans (**Figs. 2** and **3**). Buried land surfaces of numerous laterally contiguous paleosols are known as geosols and have been widely used as stratigraphic market horizons for the mapping and interpretation of Quaternary deposits. *See* LATERITE; QUATERNARY; STRATIGRAPHY.

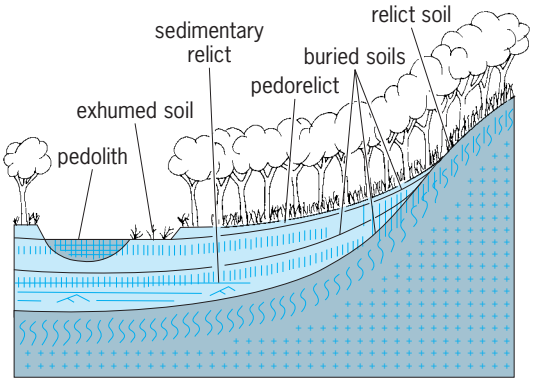

Fig. 1. Diagrammatic landscape profile illustrating key concepts and terms in the study of paleosols.

Fig. 2. Triassic paleosol from the sea cliffs at Long Reef, near Sydney, in southeastern Australia, showing conspicuous drab-haloed root traces (near 1-ft or 0.3-m scale) and the sharply truncated top and gradational lower horizons characteristic of paleosols.

Fig. 3. Middle Miocene paleosols and a similar modern soil in the quarry at Forth Ternan National Monument, southwestern Kenya, showing finely structured surface horizons (crumb peds) that have developed during soil formation at the expense of relict bedding (at level of hammer).

Paleosols also are found at major geological unconformities in the rock record, including the unconformity at the land surface in continental regions that have been exposed to weathering for many millions of years. Some of these relict deep-weathering profiles dating back to the Cretaceous in Australia, Africa, and Brazil are up to 100 m (330 ft) thick. Their bauxitic and lateritic horizons are valued as a source of clay for china, aggregate for construction material, and ores of aluminum and iron. A long fossil record of deeply weathered paleosols at major geological unconformities as old as 3–5 × 10^9 years has been subject to exploration as evidence for prehistoric changes in the atmosphere and life on land. *See* BAUXITE; CRETACEOUS.

Paleosols are especially abundant in volcanic, alluvial, and eolian sedimentary sequences. Along with the fossils, sedimentary structures, and volcanic rocks found in such deposits, paleosols provide an additional line of evidence for ancient environments during times between eruptions and depositional events. *See* PALEOCLIMATOLOGY; SEDIMENTOLOGY.

Paleosols can be characterized by their root traces, horizons, and soil structures visible in the field, supported by laboratory studies of their chemical and

mineralogical composition and their microscopic fabrics as seen in petrographic thin sections. Many ancient paleosols have been altered during burial by physical compaction, by local chemical reduction producing green-gray areas around buried organic matter, by reddening from dehydration of ferric oxyhydrate minerals, and by illitization of clays from the dissolution during burial of microcline and other potash-bearing minerals. Such alteration after burial can limit the reconstruction of a paleosol, its identification in a soil classification, and interpretation of soil-forming factors during its formation. Nevertheless, many features of paleosols, such as horizons rich in clay skins or in calcareous nodules, are surprisingly robust in the face of alteration during deep burial and metamorphism. *See* PETROFABRIC ANALYSIS; SOIL. Gregory J. Retallack

Bibliography. P. W. Birkeland, *Soils and Geomorphology*, 1999; G. J. Retallack, *A Color Guide to Paleosols*, 1997.

Paleozoic

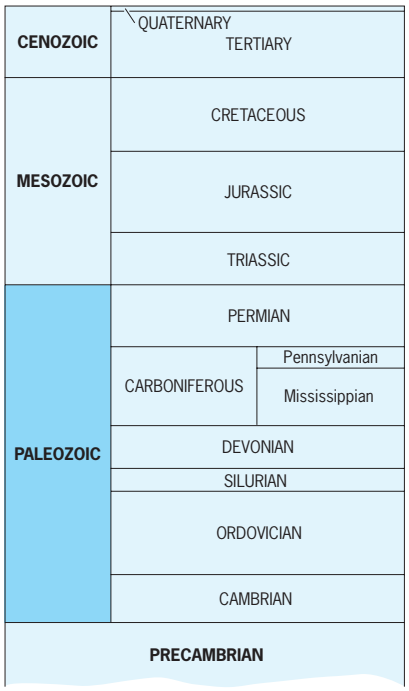

A major division of time in geologic history, extending from about 540 to 250 million years ago (Ma). It is the earliest era in which significant numbers of shelly fossils are found, and Paleozoic strata were among the first to be studied in detail for their biostratigraphic significance. Western Europe, especially the British Isles, was the cradle of historical geology. Early work with rock strata and their fossils was strictly practical; the relative ages of rock units were essential for correlating scattered outcrops to search for natural resources—particularly coal—in the early part of the nineteenth century.

During its first four decades, natural groupings of strata were studied and named for easy reference. Thus the several subdivisions of the Paleozoic, ultimately the six standard systems, were established. The original basis for establishing sequence was superposition. The operational stratigraphic hypothesis is that, in most instances, the strata at the bottom of a sequence are the oldest and the overlying beds are progressively younger. Thus, the basal system of the Paleozoic, in which primitive shelly fossils are found, is the Cambrian. As younger and younger layers were studied, their fossils collected, and the biological affinities suggested, the concept of evolution from simpler to more complex life forms took shape in the minds of the paleontologists and geologists who were studying the rocks. This process did not take place in an orderly way, from oldest to youngest strata, but rather as a consequence of fulfilling a need of the moment, whether to complete a geologic map or to solve a problem of stratigraphic correlation. Consequently, the first Paleozoic system to be named and studied in some detail was the Carboniferous—the great "coal-bearing" sequence—given that name by W. D. Conybeare and W. Phillips in 1822. These strata were to provide the world's major energy resources during the next century and a half. Most of the Northern Hemisphere's coal fields, and much of its oil and gas as well, were produced from Carboniferous rocks. *See* SUPERPOSITION PRINCIPLE.

In the 1830s and 1840s two British geologists, R. Murchison and A. Sedgwick, studied and named the natural groupings of rock strata in the British Isles. Sedgwick named the Cambrian System in 1835, for a sequence of strata that overlies the Primordial (Precambrian) rocks in northwest Wales. Four years later, Murchison gave the name Silurian to the early Paleozoic rocks found in the Welsh borderland. However, there was an almost complete overlap of the Cambrian by Murchison's Silurian. It was not until 1879, when C. Lapworth named the Ordovician System for rocks intermediate between the Cambrian and the "upper" Silurian, that the three early Paleozoic systems were sorted out in the correct order. In the meantime, Murchison and Sedgwick managed to agree on the rocks above the Silurian and, in 1839, they named the Devonian System for rocks exposed in Devonshire, England. The final Paleozoic system, the Permian, was named by Murchison in 1841, after an expedition to Russia, where he recognized the youngest Paleozoic fossil assemblages in the carbonate rocks exposed in the province of Perm. *See* PRECAMBRIAN.

Subdivisions. The Paleozoic Era is divided into six systems; from oldest to youngest they are Cambrian, Ordovician, Silurian, Devonian, Carboniferous, and Permian. The Carboniferous is subdivided into two subsystems, the Mississippian and the Pennsylvanian which, in North America, are considered systems by many geologists. The Silurian and Devonian systems are closer to international standardization than others; all the series and stage names and lower

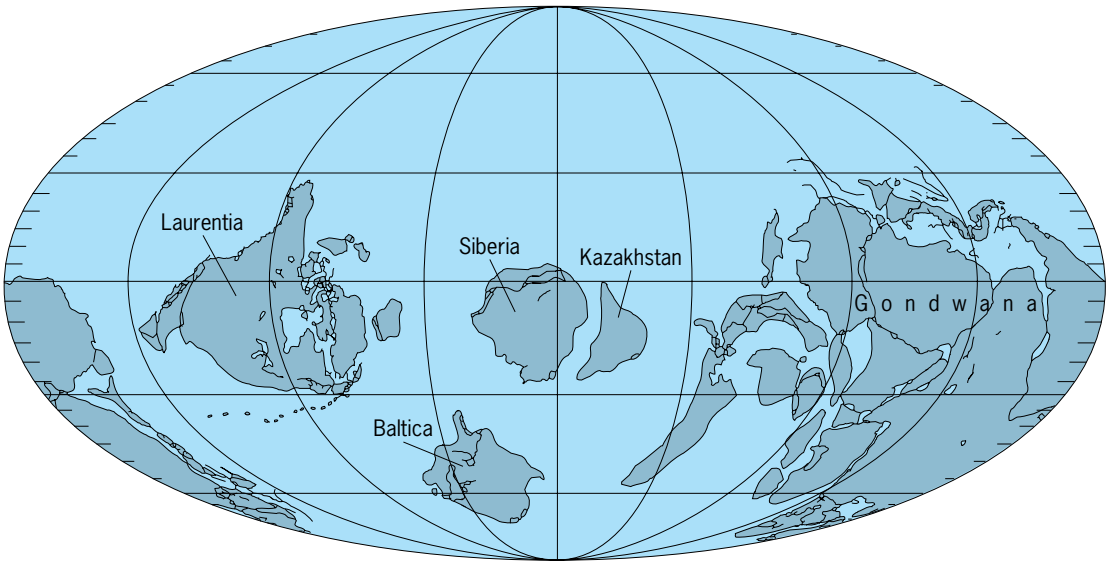

Fig. 1. Paleogeography of the Cambro-Ordovician (Tremadoc) showing most of the northern plates spread east-west in the equatorial regions. (*After W. S. McKerrow and C. R. Scotese, eds., Paleozoic Palaeogeography and Biogeography, Geol. Soc. Mem. 12, Geological Society, London, 1990*)

boundaries have been agreed upon, and most have been accepted. Despite continuing revisions, the major subdivisions of the geologic time scale have been relatively stable for nearly a century. *See* CAMBRIAN; CARBONIFEROUS; DEVONIAN; ORDOVICIAN; PERMIAN; SILURIAN.

Paleotectonics. The Paleozoic oceans, just as those today, surrounded a series of landmasses that formed the cores of ancient plates, always in motion as are their modern counterparts. Sediments were supplied to the seas through a network of river drainage systems and distributed in the oceans, by currents and gravity, very like today. Clastic sediments were supplied by the mountainous regions that were uplifted and eroded in cyclic patterns as the major plates collided and parted; and subduction at the

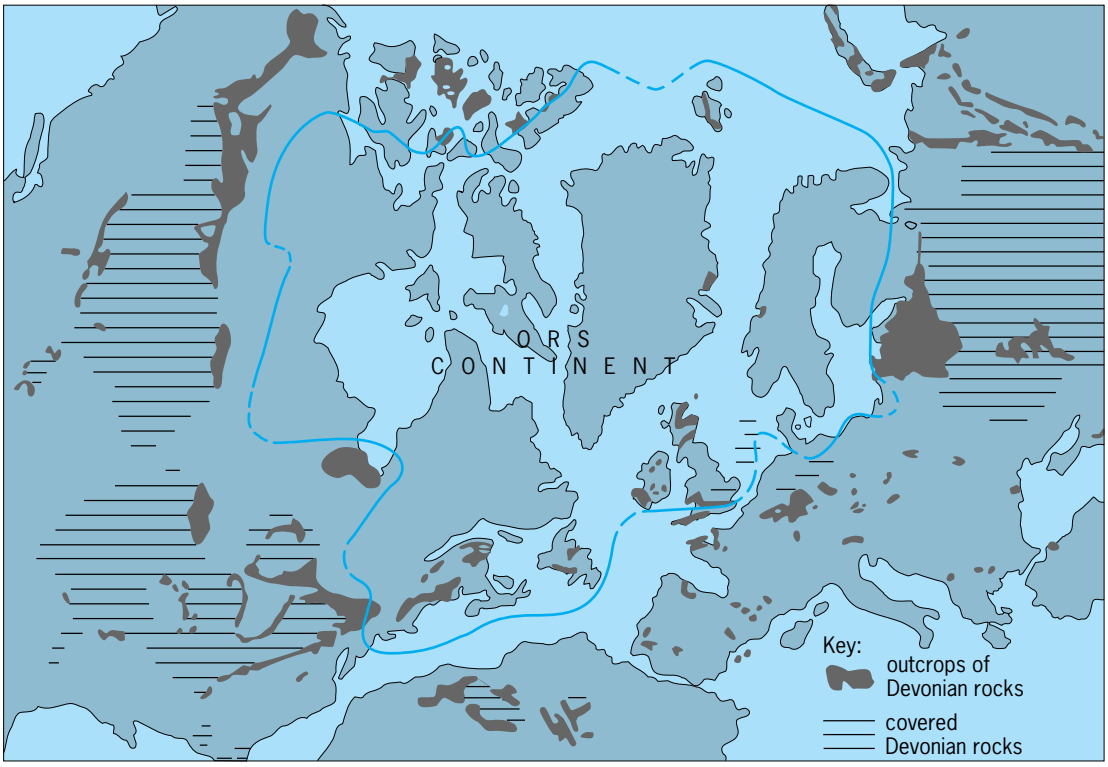

Fig. 2. Paleogeography of Old Red Sandstone continent (outlined in color) in Late Devonian. (*After R. Goldring and F. Langenstrassen, Open shelf and near-shore clastic facies in the Devonian, in M. R. House, C. T. Scrutton, and M. G. Bassett, eds., The Devonian System, Spec. Pap. Palaeont. 23, Palaeontological Association, London, 1979*)

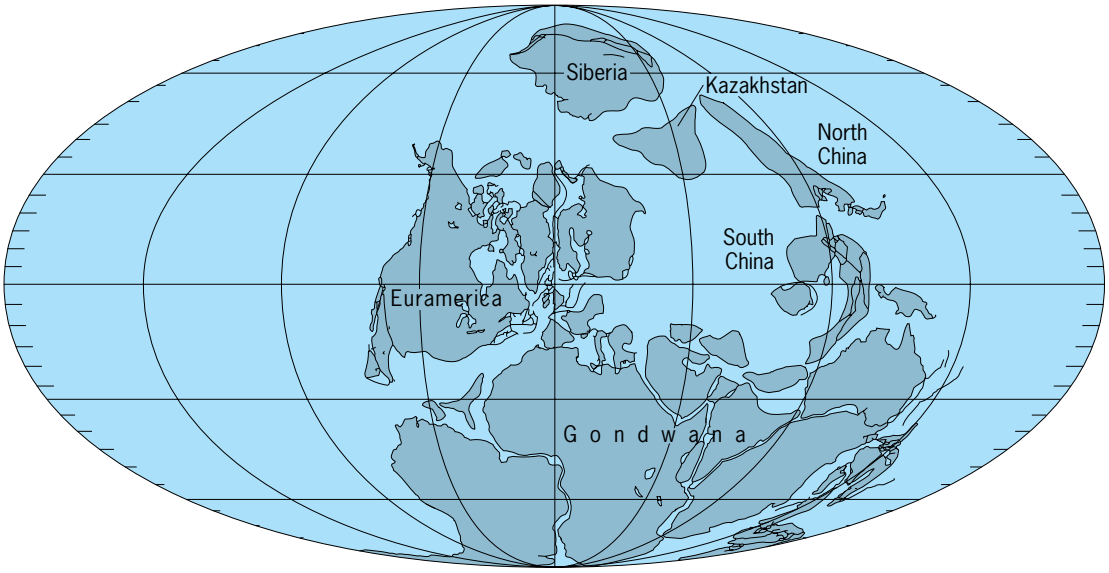

Fig. 3. Paleogeography of the Early Carboniferous (Viséan), showing the Euro-American megaplate centered in equatorial region. (*After W. S. McKerrow and C. R. Scotese, eds., Paleozoic Palaeogeography and Biogeography, Geol. Soc. Mem. 12, Geological Society, London, 1990*)

leading edges of some plates produced volcanic highlands. The plate tectonic theory provides a template for sorting out the periods of mountain building during the Paleozoic. Like the discovery of the stratigraphic systems, periods of orogeny, with their concurrent volcanic and intrusive igneous activities, were revealed by field studies. Tectonic effects (folding and faulting) were analyzed by geologic mapping, as were crosscutting igneous relations and unconformities in the sedimentary sequence. Regional orogenic terranes were named and the general time sequence assigned; these were sharpened as the use of isotopic age analyses of the igneous components became possible in the twentieth century.

Because Alpine and Appalachian mountain chains were among the first studied in detail, orogenies were first named there. In eastern North America, mountain-building effects during the early Paleozoic were ascribed to the Taconic orogeny (Middle and Late Ordovician); middle Paleozoic events were assigned to the Acadian orogeny (Middle and Late Devonian); and late Paleozoic movements were called Appalachian (more accurately Alleghenian) for Permian and, perhaps, Triassic events. Similar, but not precisely correlative, orogenic episodes in western Europe are ascribed to the early Paleozoic Caledonian and the late Paleozoic Variscan (or Hercynian) orogenies. This regionalization, overlapping of

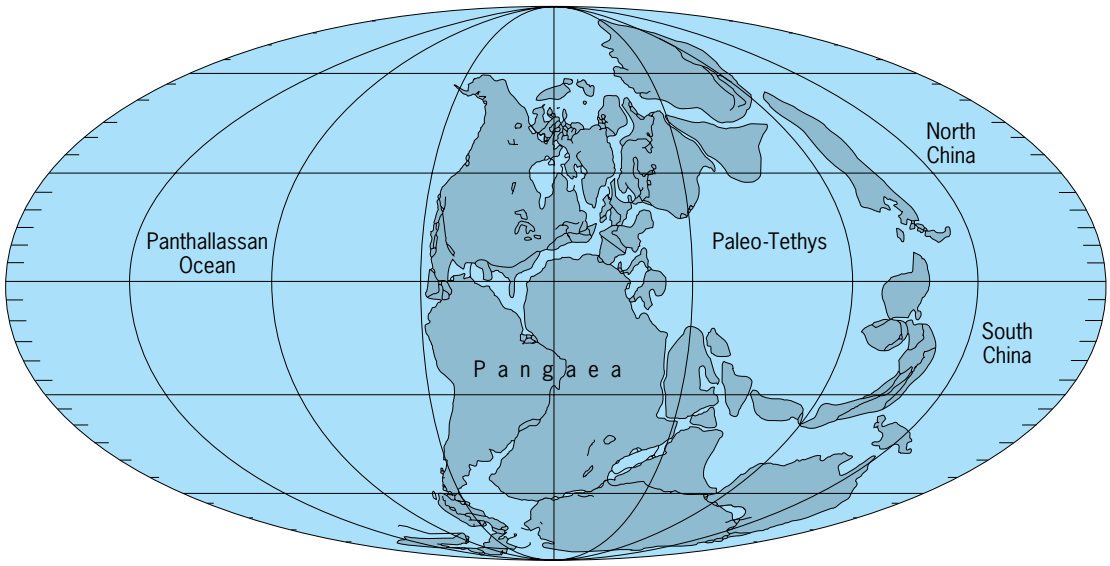

Fig. 4. Paleogeography at the end of the Permian showing the consolidation into Pangaea preparatory to the onset of the Mesozoic tectonic cycle. (*After W. S. McKerrow and C. R. Scotese, eds., Paleozoic Palaeogeography and Biogeography, Geol. Soc. Mem. 12, Geological Society, London, 1990*)

timing of events and lack of correlation of intrusive phases, tectonics, and sedimentation cycles emphasize the universality of the ever-moving plates as the global mechanism responsible for all tectonic events. *See* DATING METHODS; ISOTOPE; OROGENY; PLATE TECTONICS; UNCONFORMITY.

Lithofacies. The major changes in lithofacies during the Paleozoic were also effected by biotic evolution through the era. Limestone facies became more abundant and more diversified in the shallow warm seas as calcium-fixing organisms became more diverse and more widespread. Sediment input from the land was modified as plants moved from the seas to the low coastal plains and, eventually, to the higher ground during the Devonian. Primitive vertebrates evolved during the Cambro-Ordovician, but true fishes and sharks did not flourish until the Devonian. Amphibians invaded the land during the Late Devonian and early Carboniferous at about the same time that major forests began to populate the terrestrial realm. These changes produced an entirely new suite of nonmarine facies related to coal formation, and the Carboniferous was a time of formation of major coal basins on all continental plates.

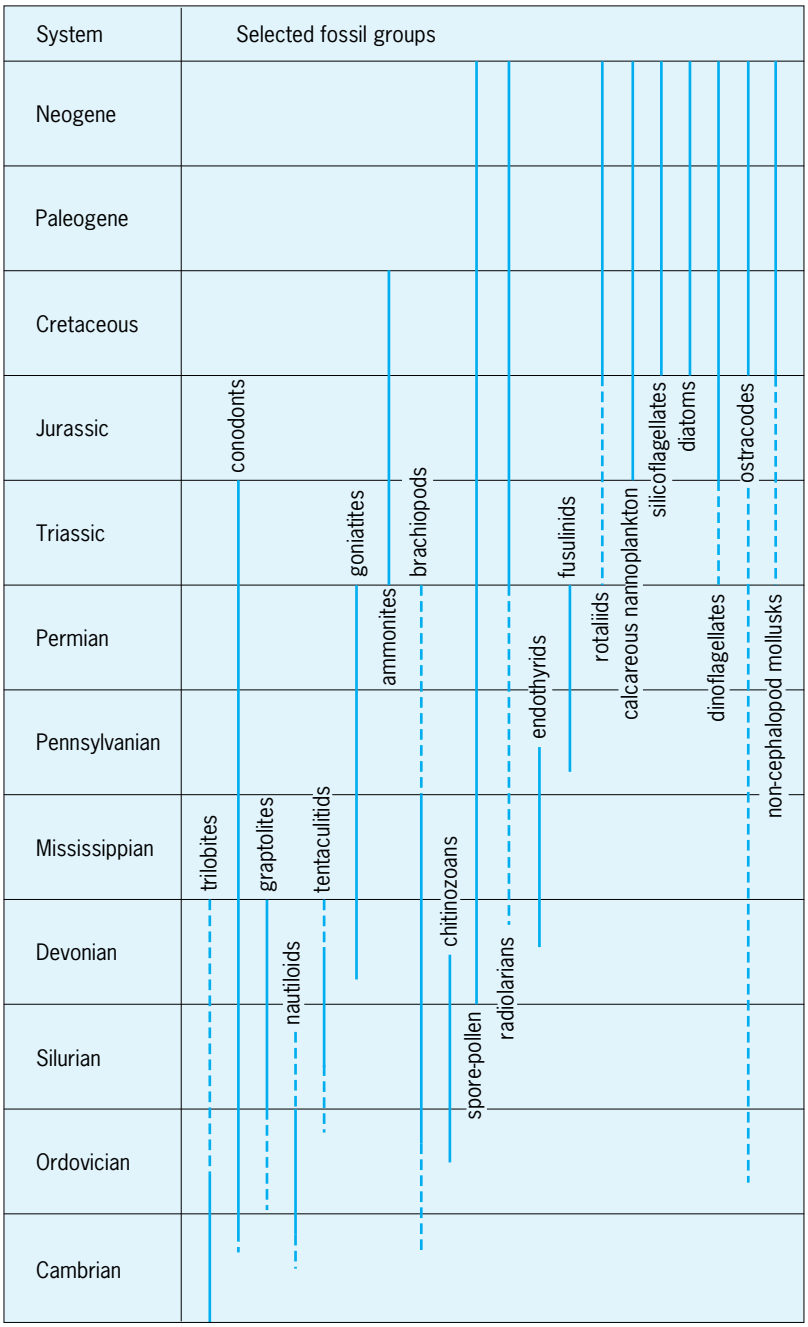

Fig. 5. Major fossil groups used for detailed biostratigraphy of the Paleozoic and younger strata. These are not total ranges of all groups. (*After J. T. Dutro, R. V. Dietrich, and R. M. Foose, eds., AGI Data Sheets, 3d ed., American Geological Institution, 1989*)

Climate continually influenced depositional patterns and lithofacies both on land and in the seas. In the major carbonate basins and platforms, particularly from the Late Ordovician onward, cyclic climatic changes resulted in changes from calcitic to magnesian carbonates and, ultimately, to various saline deposits as the basins dried up. There were great salt deposits in several systems, but spectacular thicknesses developed in many basins during the Silurian and Permian. Major cycles of cold and warm climates were overlaid on depositional and evolutionary patterns, producing periods of continental glaciation when large amounts of the Earth's water were tied up in ice during the Late Ordovician, the Late Devonian, and the late Permian. During the earliest and latest of these periods, icesheets were concentrated in the Southern Hemisphere on a single large Paleozoic continental mass—Gondwana. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; FACIES (GEOLOGY); PALEOCLIMATOLOGY.

Paleogeography. Paleogeographic changes naturally followed the shifting of plates on a megacyclic scale during the Paleozoic. In general, the Paleozoic featured a single southern landmass (Gondwana) for most of the era. This megaplate moved relatively sedately northward during this entire time interval (540–250 Ma) and always contained the magnetic and geographic south poles. Consequently, many of the facies and biologic provinces in the Gondwanan region were influenced by the cooler marine realms and continental and mountain glaciers in nearly every Paleozoic period. Most of the tectonic action that produced major periods of collision, mountain building, carbonate platform building, back-arc fringing troughs with their distinctive faunas and lithofaces, and formation of coal basins and evaporites took place in the Northern Hemisphere. These pulsations produced combinations of Laurentian (North American), Euro-Baltic, Uralian, Siberian, and Chinese plates at various times during the Paleozoic; and these combined units, in turn, moved slowly across the latitudes, producing climatic change; lithofacies changed in response to both the climate and the plate tectonics.

Representative geographies that show the range of change have been deduced (**Figs. 1–4**). The map for the Cambro-Ordovician portrays the general early Paleozoic patterns (Fig. 1); these hold for the entire span of time from the Early Cambrian (about 540 Ma), through the Cambrian, Ordovician, and Silurian, into the early Devonian (about 400 Ma). There were consolidations in the Northern Hemisphere in

the Devonian, leading to a northern landmass—the so-called Old Red Continent (Fig. 2)—the forerunner of the Euro-American megaplate of the Carboniferous (Fig. 3), and culminating at the end of the Permian in the Pangaean continental mass (Fig. 4). This, in turn, set the stage for the breakup of Pangaea during the subsequent Mesozoic tectonic megacyle. *See* PALEOGEOGRAPHY.

Biogeography and biostratigraphy. The complexities of evolution from relatively simple forms at the beginning of the era to more advanced faunas and floras at the beginning of the Mesozoic produced a web of distributions in both time and space during the Paleozoic. In general terms, there were fewer and simpler life forms in the Cambrian—often termed the Age of Trilobites. All groups of invertebrates and plants became more numerous through geologic time. For example, 7 major invertebrate animal groups at the beginning of the Cambrian doubled to 14 by the end of the period, 20 by the end of the Ordovician, 23 at the end of the Devonian, and 25 at the end of the Paleozoic. The pattern for plant diversification, although starting later, is similar. Three simple plant groups became 5 by the end of the Silurian, 7 at the end of the Devonian, and 13 at the end of the Paleozoic. The vertebrates also diversified very slowly. From one or two groups in the Cambro-Ordovician (conodonts are now considered primitive vertebrates), the number of major kinds rose to 6 at the end of the Devonian and 8 at the end of the Paleozoic.

Biostratigraphic usefulness of fossils varies widely. Certain groups have been shown empirically to be more useful than others, and the abundance and diversity within these groups change from system to system during the Paleozoic. Groups with wide dispersal, occurrences in several facies, and rapid rates of evolution have proved most useful (**Fig. 5**). In the Paleozoic, trilobites are most valuable in the Cambrian and Ordovician; conodonts are more widely studied and are providing detailed biochronologic control for many system, stage, and zonal boundaries. Graptolites are indispensible in the deeper-water facies of the Ordovician through Early Devonian; goniatite cephalopods provide standards in the Devonian through the Permian; and fusulinids have long been essential for detailed work in the Carboniferous and Permian. Of course, all groups are useful for other kinds of paleobiologic research. Paleoenvironmental, paleoecological, and paleobiogeographic reconstructions use all appropriate biologic, chemical, and physical data in developing models of ancient Paleozoic worlds. *See* BIOGEOGRAPHY; CEPHALOPODA; CONODONT; FUSULINACEA; GEOLOGIC TIME SCALE; GRAPTOLITHINA; INDEX FOSSIL; PALEOECOLOGY; STRATIGRAPHY; TRILOBITA.

J. Thomas Dutro, Jr.

Bibliography. A. F. Embry, B. Beauchamp, and D. J. Glass (eds.), *Pangea: Global Environments and Resources*, Canadian Society Petroleum Geologists Memoir 17, 1994; F. M. Gradstein, J. G. Ogg, A. G. Smith (eds.), *A Geological Time Scale 2004*, 2005; M. R. House, C. T. Scrutton, and M. G. Bassett (eds.), *The Devonian System*, Spec. Pap. Palaeont. 23, Palaeontological Association, London, 1979; E. G. Kauffman and J. E. Hazel (eds.), *Concepts and Methods of Biostratigraphy*, 1977; W. S. McKerrow and C. R. Scotese (eds.), *Palaeozoic Palaeogeography and Biogeography*, Geol. Soc. Mem. 12, Geological Society, London, 1990; R. C. Moore et al., *Treatise on Invertebrate Paleontology*, Part A, 1979; G. C. Young and J. R. Lauries (eds.), *An Australian Phanerozoic Time Scale*, Oxford University Press, 1996.

Palladium

A chemical element, Pd, atomic number 46, and atomic weight 106.4. A transition metal, palladium occurs in combination with platinum (Pt) and is the second most abundant platinum-group metal, accounting for 38% of the reserves of these metals. *See* PERIODIC TABLE; PLATINUM.

1												13	14	15	16	17	18
1 H	2																2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Palladium is soft and ductile and can be fabricated into wire and sheet. The metal forms ductile alloys with a broad range of elements. Palladium is not tarnished by dry or moist air at ordinary temperatures. At temperatures from 350 to 790°C (660 to 1450°F) a thin protective oxide forms in air, but at temperatures from 790°C (1450°F) this film decomposes by oxygen loss, leaving the bright metal. In the presence of industrial sulfur-containing gases a slight brownish tarnish develops; however, alloying palladium with small amounts of iridium or rhodium prevents this action. Important physical properties of palladium are given in the **table**. *See* ALLOY; METAL.

At room temperature, palladium is resistant to nonoxidizing acids such as sulfuric acid, hydrochloric acid, hydrofluoric acid, and acetic acid. The metal is attacked by nitric acid, and a mixture of nitric acid and hydrochloric acid is a solvent for the metal. Palladium is also attacked by moist chlorine (Cl) and bromine (Br). *See* ELECTROPLATING OF METALS; NONSTOICHIOMETRIC COMPOUNDS.

The major applications of palladium are in the electronics industry, where it is used as an alloy with silver for electrical contacts or in pastes in miniature solid-state devices and in integrated circuits. Palladium is widely used in dentistry as a substitute for

Physical properties of palladium

Property	Value
Atomic weight	106.4
Naturally occurring isotopes (percent abundance)	102 (0.96)
	104 (10.97)
	105 (22.23)
	106 (27.33)
	108 (26.71)
	110 (11.81)
Crystal structure	Face-centered cubic
Thermal neutron capture cross section, barns	8.0
Density at 25°C (77°F), g/cm^3	12.01
Melting point, °C (°F)	1554 (2829)
Boiling point, °C (°F)	2900 (5300)
Specific heat at 0°C (32°F), cal/g	0.0584
Thermal conductivity, (cal·cm)(cm^2·s·°C)	0.18
Linear coefficient of thermal expansion, (μin./in./)/°C	11.6
Electrical resistivity at 0°C (32°F), $\mu\Omega$-cm	9.93
Young's modulus, lb/in.^{2}, static, at 20°C (68°F)	16.7×10^6
Atomic radius in metal, nm	0.1375
Ionization potential, eV	8.33
Binding energy, eV	3.91
Pauling electronegativity	2.2
Oxidation potential, V	−0.92

gold. Other consumer applications are in automobile exhaust catalysts and jewelry. *See* INTEGRATED CIRCUITS.

Palladium supported on carbon or alumina is used as a catalyst for hydrogenation and dehydrogenation in both liquid- and gas-phase reactions. Palladium finds widespread use in catalysis because it is frequently very active under ambient conditions, and it can yield very high selectivities. Palladium catalyzes the reaction of hydrogen with oxygen to give water. Palladium also catalyzes isomerization and fragmentation reactions. *See* CATALYSIS.

Halides of divalent palladium can be used as homogeneous catalysts for the oxidation of olefins (Wacker process). This requires water for the oxygen transfer step, and a copper salt to reoxidize the palladium back to its divalent state to complete the catalytic cycle. *See* HOMOGENEOUS CATALYSIS; TRANSITION ELEMENTS. D. Max Roundhill

Bibliography. G. W. Gribble and J. J. Li, *Palladium in Heterocyclic Chemistry*, Pergamon, 2000; F. R. Hartley, *Chemistry of Platinum and Palladium*, 1973; J. Tsuji (ed.), *Palladium in Organic Synthesis (Topics in Organometallic Chemistry)*, Springer, 2005; J. Tsuji, *Palladium Reagents and Catalysts: New Perspectives for the 21st Century*, Wiley, 2d ed., 2004.

Palpigradi

An order of rare arachnids comprising 21 known species from tropical and warm temperate regions. American species occur in Texas and California. All are minute, whitish, eyeless animals, varying from 0.03 to 0.11 in. (0.7 to 3 mm) in length, that live under stones, in caves, and in other moist, dark places. The elongate body terminates in a slender multisegmented flagellum set with setae. In a curious reversal of function, the pedipalps, the second pair of head appendages, serve as walking legs. The first pair of true legs, longer than the others and set with sensory setae, has been converted to tactile appendages which are vibrated constantly to test the substratum. *See* ARACHNIDA. Willis J. Gertsch

Palynology

The study of pollen grains and spores, both extant and extinct, as well as other organic microfossils. Although the origin of the discipline dates back to the seventeenth century, when modern pollen was first examined microscopically, the term palynology was not coined until 1944.

The term palynology is used by both geologists and biologists. Consequently, the educational background of professional palynologists may be either geologically or biologically based. Considerable overlap exists between some areas of the fields, however, and many palynologists have interdisciplinary training in both the earth and life sciences. Palynologists use a range of sophisticated methodologies and instruments in studying both paleopalynological and neopalynological problems, but the utilization of modern microscopy is fundamental in both subdisciplines.

Palynologists study microscopic bodies generally known as palynomorphs. These include an array of organic structures, each consisting of a highly resistant wall component. Examples include acritarchs and chitinozoans (microfossils with unknown affinities), foraminiferans (protists), scolecodonts (tooth and mouth parts of marine annelid worms), fungal spores, dinoflagellates, algal spores, and spores and pollen grains of land plants. This discussion will focus on the palynomorphs produced by land plants, beginning with a general description of pollen grains and spores and then providing an overview of the primary areas of investigation within neo- and paleopalynological subdisciplines. *See* MICROPALEONTOLOGY.

Pollen Grains and Spores: An Overview

Spores and pollen grains are reproductive structures and play a paramount role in the life history of land plants. The sporophyte generation of nonseed-bearing plants (ferns, for example) produces single-celled spores that ultimately germinate to grow into the haploid gametophyte generation. Homosporous species produce a single type of spore, whereas heterosporous species produce two spore types. Microspores germinate and grow into "male" sperm-producing microgametophytes, and megaspores develop into "female" egg-producing megagametophytes. The gametophytes of most nonseed plants are multicellular and proliferate outside the spore wall during development. All seed-bearing plants

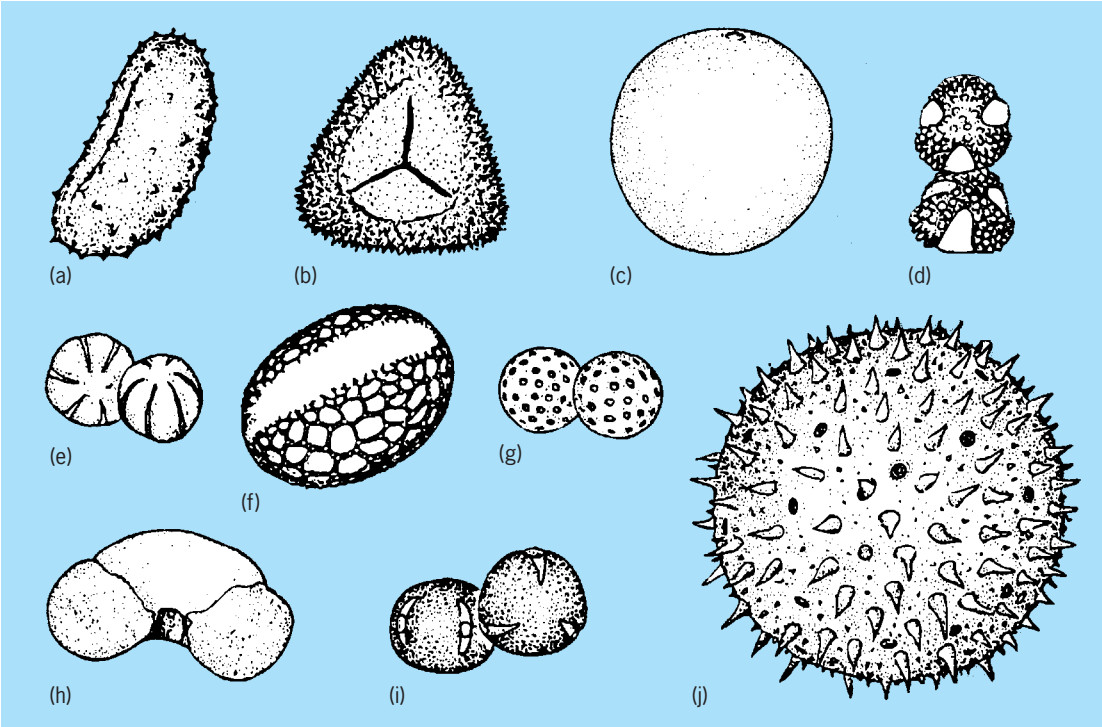

Fig. 1. Pollen and spore morphology, (*a*, *b*) Spores. (*c–j*) Pollen grains. (*After Y. Iwanami, T. Sasakuma, and Y. Yamada, Pollen: Illustrations and Scanning Electron Micrographs, Kodansha and Springer, 1988*)

(gymnosperms and angiosperms) are heterosporous, and their pollen represents the microgametophyte generation. Pollen grains consist of just three to a few cells, and these remain within the microspore wall, where they originally developed. *See* REPRODUCTION (PLANT).

Spores and pollen grains are formed in multiples of fours following meiotic divisions. During development, the four are united into a tetrad that, in most plants, subsequently dissociates into the four individual propagules. In nonseed plants, each spore commonly bears a mark on its proximal surface indicating where it made contact with the others at the center of the tetrad. In most spores this external mark is either straight or Y-shaped (**Fig. 1**), and it is typically characterized by a suture that spans the spore wall and is the site through which germination occurs. In contrast, most pollen grains lack sutures and germinate through thin areas in the wall, or apertures. Apertures are typically located in either a distal or an equatorial position. Common aperture types include elongate furrows, pores, and furrows with a central pore. Aperture type, number, and position are important systematic characters by which fossil and modern taxa can be compared. Other descriptively and systematically relevant characters include size, shape, presence and structure of air bladders, surface ornamentation, and wall ultrastructure. *See* POLLEN.

The wall of spores and pollen grains is known collectively as the sporoderm (or "skin of the spore") and actually consists of two distinct walls (**Fig. 2**). The inner wall, or intine, is primarily composed of cellulose and pectin; as such, it is similar to most other plant cell walls. The outer wall, or exine, is principally composed of sporopollenin, a chemically enigmatic macromolecule that is resistant to biological decay and geological degradation. The exine is further characterized by several ultrastructural layers and an array of sculptural elements. It is the very presence of the exine that allows for the spectacular preservation of pollen and spores in the fossil record.

Neopalynology

This discussion focuses on several subdisciplines of neopalynology, including taxonomy, genetics, and evolution; development, functional morphology, and pollination; aeropalynology; and melissopalynology.

Taxonomy, genetics, and evolution. Taxonomy and systematics are concerned with classifying organisms into hierarchical ranks that reflect evolutionary, or phylogenetic, relationships. Pollen and spore morphology is important systematically, with particular features characteristic of different taxonomic ranks. For example, distinguishing characters may include aperture type for a family, different ornamentation patterns for its subordinate genera, and variation in exine ultrastructure for its congeneric species. Palynological characters are especially useful systematically when evaluated in conjunction with other characters (for example, plant morphological and molecular characters). Cladistics is one technique that has employed such an integrated approach. Cladistic analyses are based on numerical algorithms

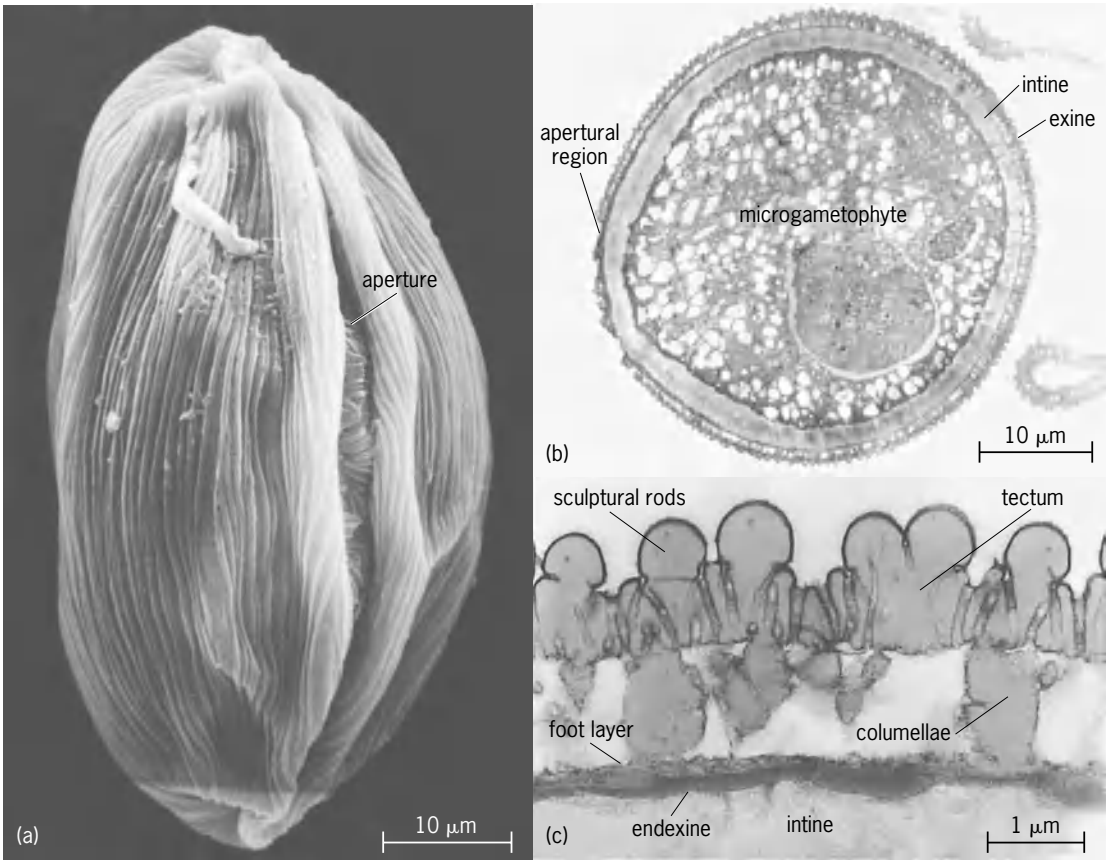

Fig. 2. Pollen morphology and sporoderm ultrastructure of *Cabomba caroliniana* (Cabombaceae), a modern water lily and primitive flowering plant. (*a*) Distal view of grain. (*b*) Cross section through the entire grain. (*c*) Cross section through the sporoderm.

that produce trees, or cladograms, demonstrating phylogenetic lineages among the organisms examined. *See* PLANT EVOLUTION; PLANT TAXONOMY.

Assessing pollen flow is another approach used to study evolutionary questions. Because pollen is the sperm-producing generation, the patterns and rates of pollen transfer are important factors in determining the spread of genes throughout a population. Competition can occur among reproductive organs, and pollen flow may be influenced by the correlation of pollen characters with those of reproductive organs. For example, the ovulate, or female, cones of some gymnosperms, such as pines, are aerodynamically adapted for entraining airborne pollen grains that themselves have particular aerodynamic characters, such as extended air bladders. Furthermore, competition may exist among individual pollen grains. Following pollination in some flowering plants, intraspecific variation can result in the grains of a more reproductively fit plant producing faster-growing pollen tubes than others of the same species. *See* POPULATION GENETICS.

Development, functional morphology, and pollination. In most seed plants, a layer of callose, a carbohydrate, surrounds the entire tetrad and separates each immature pollen grain during development. Formation of the pollen wall and positioning of the aperture begin while each grain is encased within the callose layer. Both the internal ultrastructure and the sculptural surface ornamentation of the outer pollen wall, the exine, are dependent upon the depositional pattern of the chemical that makes up the exine, sporopollenin. Sporopollenin is primarily derived from the developing pollen grain, but can also be released from a specialized layer of cells known as the tapetum, which surrounds the developing pollen grains. The inner pollen wall, or intine, is synthesized last.

When the tapetum breaks down, or undergoes programmed cell death (apoptosis), it also releases several proteins, lipids, and other substances that become isolated within the spaces and on the surface of the developing pollen wall. In flowering plants, many tapetum-derived proteins function as recognition molecules in pollination systems and are important in determining the extent of compatibility of a particular pollen grain on a floral stigma. Other pollen-derived proteins become stored within the intine and may also be involved in compatibility-incompatibility reactions. Several tapetal lipids also play important roles in pollination. Pollenkitt, for example, functions in pollen adhesion and acts as a visual and olfactory attractant. Additionally, pollen morphology and exine architecture may be correlated with pollination systems. For example, the pollen of wind-pollinated plants is typically smooth, has a thin exine, lacks pollenkitt, and may have air bladders, whereas that of animal-pollinated plants is

commonly highly ornamented and bears pollenkitt. *See* FERTILIZATION.

Aeropalynology. Aeropalynology is the study of pollen grains and spores that are dispersed into the atmosphere. Wind-pollinated plants typically produce copious amounts of pollen, thereby enhancing successful pollinations. The abundance of airborne pollen commonly causes allergic reactions in a large proportion of the human population. Pollinosis, allergen rhinitis or hay fever, is elicited when allergen-containing pollen makes contact with the mucous membranes lining the nose, trachea, or bronchi, and the cornea of the eye. Allergens leach out of the pollen and bind to immunoglobulin E antibodies. The antibodies are linked to mast cells that release histamine and other inflammatory chemicals, producing allergy symptoms. Ironically, the allergens that induce pollinosis include many of the same compatibility-incompatibility, recognition proteins involved in pollination.

Knowledge of the temporal, seasonal, and environmental aspects of pollen dispersal is also important in understanding and avoiding hay fever. Flowering time and season vary widely for different plants, and the release of airborne pollen is typically inhibited by high humidity or rain. To monitor risks of pollinosis, the diversity and quantity of various pollen types are assessed by filtering the air throughout the year. *See* ALLERGY; ANTIBODY.

Melissopalynology. Honeybees are the primary pollinators of many flowering plants. Honeybees, and other bees, visit flowers to collect nectar and large quantities of pollen (pollen loads), both of which are used as food sources for developing larvae. Melissopalynologists analyze bee pollen loads and the pollen component within honeys. Although bees primarily produce honey from nectar, 1 mL of honey may contain more than 20,000 pollen grains. The foraging behavior of bees can be determined by microscopically examining their pollen loads and taxonomically identifying the pollen constituents. Honey purity can also be assessed by examining the diversity of pollen grains found within the honey.

Paleopalynology

The main fields of study within paleopalynology are discussed below. The areas addressed involve paleobotany; past vegetation and climate reconstruction; geochronology and biostratigraphy; and petroleum and natural gas exploration.

Paleobotany. Fossil pollen and spores typically consist of only fossilization-resistant exine layers. However, these propagules did at one time contain both gametophytic cells and pectocellulosic intines, and functioned in a similar way to that of their extant counterparts. Fossil pollen and spores can be distinguished into two categories based on the general way in which the palynomorphs are preserved. *Sporae dispersae* grains are those occurring within sediments in a dispersed condition; in most cases, information about the parent plants is unavailable. Investigation of dispersed grains is especially important in the fields addressed below. *In situ* grains occur within intact, megafossil reproductive organs (like flower anthers); as such, morphological data from the parent plant are available and provide for better systematic evaluation. Studies of *in situ* pollen or spores also afford the opportunity to evaluate and interpret fossils in a biological context. For example, developmental information can be inferred by examining pollen-containing organs preserved in various ontogenetic stages, and ancient pollination events, such as pollen germination and pollen tube growth, can be assessed when grains are recovered on receptive structures. These types of data allow paleobotanists to better understand and reconstruct the complete life history of fossil plants.

Past vegetation and climate reconstruction. A significant focus of palynology involves reconstructing the Earth's vegetational history since the last major glaciation event, within the past 10,000 years, or during the Holocene Epoch. This area of postglacial palynology is known as pollen analysis and primarily includes the study of palynomorphs from lake sediments and peat deposits. Sediments are obtained by several methods (mostly core sampling), and palynomorph diversity, distribution, and abundance are plotted on pollen profiles. Pollen analysis can provide historical information regarding both individual taxa and larger plant communities, including data about vegetational succession. Such analyses must consider all possible sources of palynomorphs and take precaution during sample preparation to avoid contamination with extant pollen because many modern taxa may have also existed in the Holocene. However, because of the excellent preservation potential of key palynological characters, such as those described above, fossil pollen grains yield a high degree of taxonomic resolution.

Because many plants inhabit areas exhibiting particular environmental regimes and have limited geographic distributions, palynological analyses contribute to an understanding of paleoclimatic conditions. For example, a palynoflora may be indicative of source vegetation occupying a polar, temperate, subtropical, or tropical habitat. Therefore, palynological information can also be used in conjunction with other megafossil indicators of climate, such as tree ring data. *See* POSTGLACIAL VEGETATION AND CLIMATE.

Geochronology and biostratigraphy. Palynological analyses play a significant role in age determinations of rocks, or geochronology. Dating the geological ages of palynomorph-bearing rocks is dependent upon knowledge of the stratigraphic ranges of extinct plant groups. Because different plant groups are known to have restricted geological time ranges, the pollen and spores produced by their plants are characteristic of particular ages and may serve as index fossils. Comparisons may be made against well-established reference palynomorphs, or index fossils, and palynofloras. Palynological dating techniques are especially useful when correlated with ages of rocks that have been radiometrically dated. *See* FOSSIL; INDEX FOSSIL.

Comparisons of palynomorphs within a given rock section from one site with those of units from other localities are important in documenting stratigraphic similarities among the rock sections, even if the sections exhibit different thicknesses and lithologies. When the occurrence, diversity, and abundance of fossils (palynomorphs, megafossils, or both) are used to correlate geographically separated rock sequences, this is known as biostratigraphy. Historically, biostratigraphic correlation has provided supporting evidence for continental drift theory. For example, some present-day continents, such as Antarctica, Africa, and India, have distinguishing index fossils, of various ages, that are present on these continents and absent from others. These intercontinental correlations are supportive of the previous existence of the single Southern Hemisphere landmass known as Gondwana. *See* PALEOGEOGRAPHY.

Petroleum and natural gas exploration. Economically, palynological biostratigraphy is an important technique used in the exploration for natural gas and petroleum. Biostratigraphic correlations in this context are conducted on a smaller scale, typically within an existing oil field. Besides identifying the locality, it is critical to determine the appropriate level at which to drill. For this endeavor, the palynologist is not necessarily interested in references of depth, but in important palynological indicators of known oil and gas production levels. In addition to the identification of key index fossils, a color evaluation is relevant. Following standard preparations, palynomorphs exhibit a range of colors that indicate their degree of geothermal alteration. Certain palynomorph colors are characteristic of rocks with either oil or gas reservoirs. *See* PALEOBOTANY; STRATIGRAPHY.

Jeffrey M. Osborn

Bibliography. K. Faegri, P. Kaland, and K. Krzywinski, *Textbook of Pollen Analysis*, 4th ed., Wiley, Chichester, 1989; M. M. Harley, C. M. Morton, and S. Blackmore (eds.), *Pollen and Spores: Morphology and Biology*, Royal Botanic Gardens, Kew, 2000; J. Jansonius and D. C. McGregor (eds.), *Palynology: Principles and Applications*, vols. 1–3, American Association of Stratigraphic Palynologists, Salt Lake City, 1996; R. O. Kapp, O. K. Davis, and J. E. King, *Ronald O. Kapp's Pollen and Spores*, 2d ed., American Association of Stratigraphic Palynologists, College Station, 2000; S. Nilsson and J. Praglowski (eds.), *Erdtman's Handbook of Palynology*, 2d ed., Munksgaard, Copenhagan, 1992; A. Traverse, *Paleopalynology*, Unwin Hyman, Boston, 1988.

Pancreas

A composite gland in most vertebrates, containing both exocrine cells—which produce and secrete enzymes involved in digestion—and endocrine cells, arranged in separate islets which elaborate at least two distinct hormones, insulin and glucagon, both of which play a role in the regulation of metabolism, and particularly of carbohydrate metabolism. This article discusses the anatomy, histology, embryology, physiology, and biochemistry of the vertebrate pancreas. *See* CARBOHYDRATE METABOLISM.

Anatomy

The pancreas is a more or less developed gland connected with the duodenum. It can be considered as an organ which is characteristic of vertebrates.

Chordates and lower vertebrates. In *Branchiostoma* (*Amphioxus*) a pancreatic anlage is found in young stages as a thickening of the gut caudal to the liver. The pancreas of cyclostomes, arising from the gut epithelium or from the liver duct, seems to be purely endocrine; it degenerates in later stages of development.

A true pancreas is found in selachians, with an exocrine portion opening into the intestine and an endocrine portion represented by cellular thickenings of the walls of the ducts.

Higher vertebrates. The ganoids (palaeopterygian fishes) show a diffuse pancreas—its principal mass lying between the gut and the liver—in which typical islets of Langerhans are observed. The pancreas of teleosts is either of the massive or dispersed type. Many species, such as the pike, show enormous islets of Langerhans, 10 × 5 mm, from which J. McLeod (1922) extracted insulin. The existence of a pancreas in dipneusts, such as *Protopterus*, is doubtful.

The compact pancreas of the amphibians is located in the gastrohepatic omentum and extends toward the hilus of the liver and along the branches of the portal vein. It develops from three anlagen, one dorsal and two ventral, the evolution of which varies from one species to another. The dorsal anlage would be the only source of endocrine islands. The pancreas of reptiles is very similar to that of amphibians; the number of excretory ducts varies from one to three.

In birds, the massive pancreas always lies in the duodenal loop. It develops from many dorsal and two ventral thickenings of the duodenal epithelium; one (sometimes two) excretory duct persists. The median portion of the dorsal anlage develops into a single mass which subdivides into typical islets of Langerhans. A complete ring of pancreatic tissue surrounds the portal vein.

The pancreas of mammals shows the same variations as in the fishes. The extremes are the unique, massive pancreas of humans, and the richly branched organ of the rabbit. Usually, the main duct, the duct of Wirsung, opens into the duodenum very close to the hepatic duct. Many rodents have this opening of the pancreatic duct as far as 40 cm (15.7 in.) from the hepatic duct. In humans, the pancreas weighs about 70 g (7.5 oz). It can be divided into head, body, and tail. A portion called the uncinate process is more or less completely separated from the head. Accessory pancreases are frequently found anywhere along the small intestine, in the wall of the stomach, and in Meckel's diverticulum. *See* DIGESTIVE SYSTEM.

Histology

The pancreatic parenchyma is formed by two elements; one is the exocrine tissue of which the

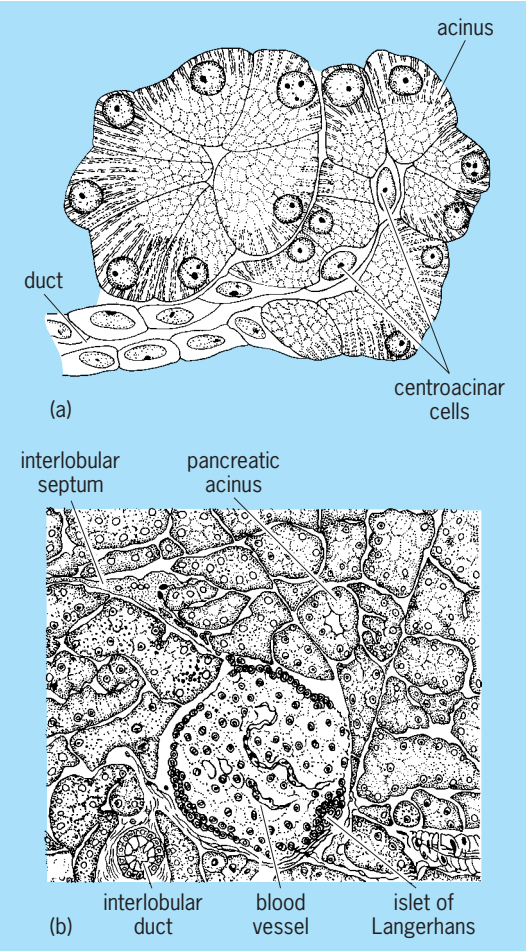

Fig. 1. Histology of the pancreas. (*a*) Centroacinar cells in the pancreas of a guinea pig (*after J. F. Nonidez and W. F. Windle, Textbook of Histology, 2d ed., McGraw-Hill 1953*). (*b*) A section through the pancreas of the rat (*after C. D. Turner, General Endocrinology, 2d ed., Saunders, 1955*).

secretion empties into the pancreatic ducts and ultimately into the duodenum; the latter is the endocrine islands whose secretions enter the blood vessels.

Exocrine pancreas. The acini are the part of the exocrine pancreas that produces the enzymes of the pancreatic juice (**Fig. 1**). The acinar cells are pyramidal in shape and all are of the serous type. The apical pole of an acinar cell is filled with granules whose number varies according to the cell activity: they are particularly numerous in fasting animals, and few after a meal or after injection of pilocarpine. Since the granules contain the enzyme precursors, they are called zymogen granules. The basal pole of an acinar cell is intensely basophilic and contains a material formerly called ergastoplasm, which consists of staggered cisternae of rough endoplasmic reticulum. The well-developed Golgi apparatus is located in the supranuclear region of the cell.

All cells converge toward a central lumen, which is rather narrow when the organ is at rest and becomes dilated during secretion. On their lateral faces, the cells are separated by narrow spaces, the secretion capillaries, which are connected with the central lumen.

The digestive enzymes of the pancreatic juice are peptidases, lipases, esterases, amylase, and nucleases. Their synthesis starts in the basal cytoplasm, inside the endoplasmic reticulum through which they are transported to the Golgi system, where they are segregated in Golgi vesicles and later concentrated into typical granules. The secretory product flows out of the cell through a fissure created by the fusion of the granule-limiting membrane with the plasma membrane. All enzymes, except one lipase and the amylase, are inactive when secreted and become active only in the duodenal lumen. Their production is stimulated by various hormones originating from the gastric and duodenal mucosae.

The lumen of each acinus is continuous with the lumen of a small canal that is limited by so-called centroacinar cells. These cells are pale when stained as compared to the acinar cells because of the scarcity of their organelles. Where the ducts begin, their wall may be made partly with centroacinar cells and partly with acinar cells themselves.

The distal part of the duct system drains proximally into the intralobular or intercalated ducts, which are lined by a low cuboidal epithelium and join together to form the interlobular ducts. These ducts are surrounded by connective tissue. Their wall structure varies with lumen diameter: when the diameter increases, the cuboidal epithelium that lines the ducts becomes cylindrical; goblet cells and isolated hormonal cells are interspersed between the other epithelial cells; some small mucous glands are annexed to the epithelium.

All interlobular ducts join the two main pancreatic ducts. The major pancreatic duct, the duct of Wirsung, runs the entire length of the pancreas, receiving numerous branches, and so gradually increases in size throughout its course. The accessory duct of Santorini lies craniad to the duct of Wirsung and only in the head of the pancreas. In both main ducts there is a layer of dense connective tissue containing elastic fibers that surrounds the stratified columnar epithelium.

Endocrine pancreas. The endocrine portion of the pancreas consists of cellular masses called the islets (or islands) of Langerhans scattered throughout the exocrine portion. In the adult human their number is estimated to range from 200,000 to 2,300,000 and their diameter varies from 30 to 300 micrometers. The islets are demarcated from the surrounding tissue by an irregular, thin layer of connective tissue.

The islet cells are arranged into cords and, in routine preparations, are paler when stained than the ancinous cells. The two main cell types are the alpha or A cells and the beta or B cells, which exist usually in the proportion of 1:4. Another cell type, less abundant (about 5%), is the delta or D cell. The localization of these three cell types varies from species to species. In humans, B cells are usually central and surrounded by A and D cells.

The A cell contains acidophilic granules that are insoluble in alcohol. In electron micrographs these granules, whose diameter ranges from 190 to 310 nanometers, are spherical and very dense. They are

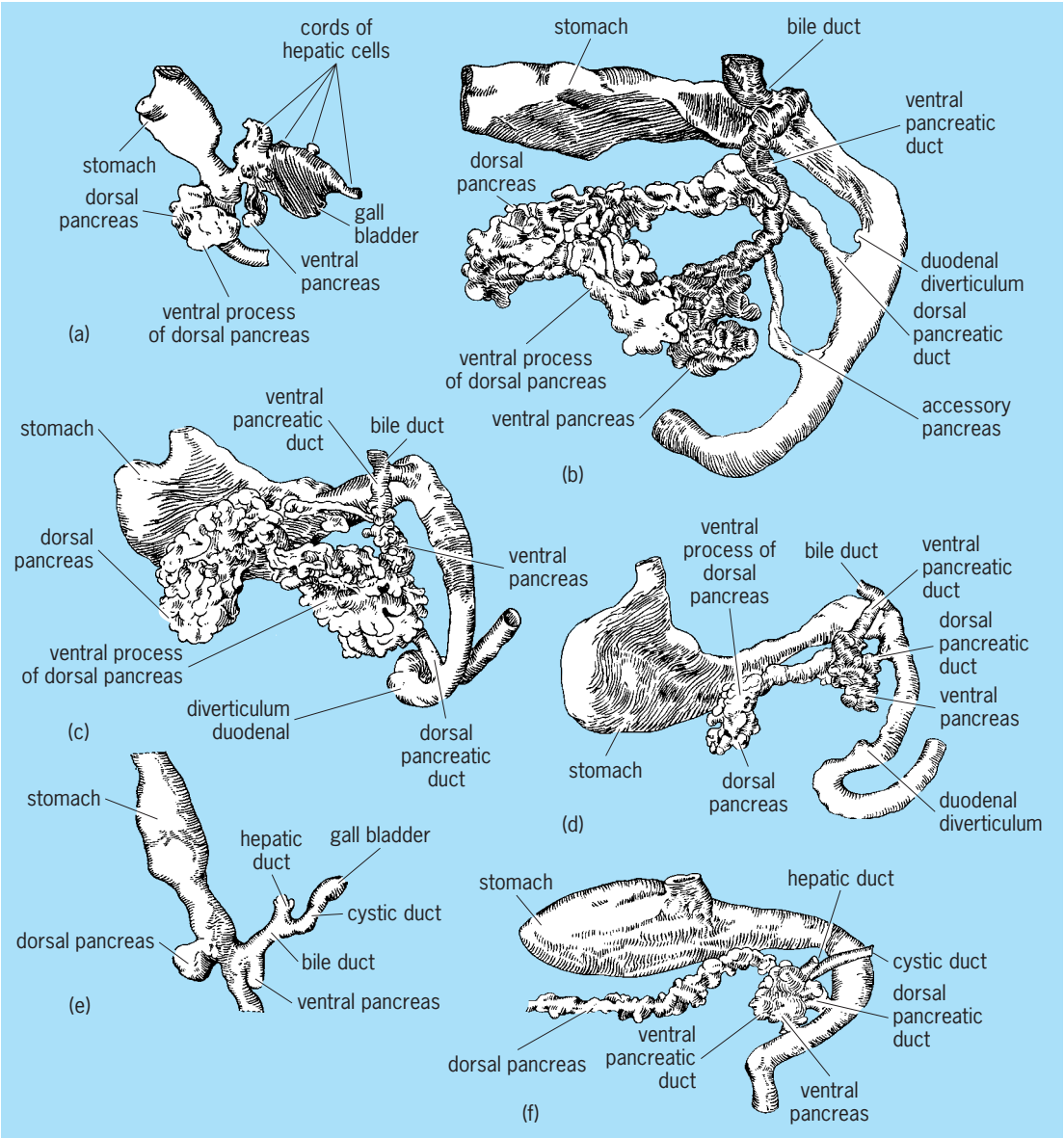

Fig. 2. Pancreas models in vertebrate embryos. (*a*) 5.5-mm pig embryo. (*b*) 20-mm pig embryo. (*c*) 11-mm rabbit embryo. (*d*) 10.7-mm cat embryo. (*e*) 7.5-mm human embryo. (*f*) 13.6-mm human embryo. (*After F. W. Thyng, Models of the pancreas in embryos of the pig, rabbit, cat and man, Amer. J. Anat., 7:488–503, 1907*)

limited by a membrane which is separated from the granule by a narrow clear zone. The core of the granule consists probably of glucagon or of glicentin (GLI 1), the prehormonal form of glucagon. The paler peripheral part of the granule should contain a glicentine-related pancreatic peptide (GRPP). The contents of granules are excreted by exocytosis, and this secretion is stimulated by the lowering of the glucose blood level.

The B cell contains basophilic granules that are soluble in alcohol. In electron micrographs these granules, whose diameter ranges from 225 to 375 nm, are composed from a dense core surrounded by a large very clear halo. It has been shown by immunocytochemical methods that the central part contains insulin and that the peripheral part contains the connection peptide which links the two chains of insulin. The crystals found in the central core are formed by insulin with zinc.

The D cell contains granules considerably less dense than the A or B granules and totally devoid of a clear ring. The D cells were first considered as altered A cells, but it is now clearly demonstrated that they produce somatostatin. It has been speculated that the secretion of somatostatin is necessary for the local regulation of the other pancreatic endocrine secretions. The long slender processes of the D cells which extend between the other cells and are rich in granules could be the morphological support of this paracrine regulation. D cells are most frequent in newborns.

Beside the A, B, and D cells, six other cell types have been identified mostly by immunocytochemical methods. These produce various substances such

as the pancreatic polypeptide, gastrin, secretin, and bombesin; their physiological significance in the islets of Langerhans still remains to be explained.

S. Haumont

Embryology

Ontogenetically, the organ is of entodermal origin. It arises initially from a series of outpocketings of the embryonic digestive tract, which eventually fuse, to various extents in various species, to form a single organ (**Fig. 2**). The adult gland in mammals results from the fusion of two such evaginations, one dorsal, opposite the hepatic diverticulum, and one ventrolateral, at the base of the biliary duct. In forms such as reptiles, anuran amphibians, and birds, two bilateral ventral evaginations are formed in addition to the single dorsal primordium. The elasmobranch pancreas is derived solely from a dorsal evagination, the ventral components never arising.

Primordia. The pancreatic diverticula make their appearance early in development, the dorsal element preceding that of the ventral component or components. In the chick, the pancreatic primordia appear between the third and fourth days of incubation, the dorsal primordia arising about the third day and the ventral on the fourth day. The dorsal component in the pig appears at the 4-mm stage, followed by the formation of the ventral diverticulum at the 5-mm stage. The pancreatic primordia of amphibians appear between the 8- and 10-mm stages, depending on the species.

Pancreatic ducts and tabules. The base of the evaginations eventually develop into the pancreatic ducts, Santorini's duct from the dorsal rudiment, Wirsung's duct from the ventral (**Fig. 3**). Santorini's duct, when it persists, opens directly into the duodenum, while the duct of Wirsung joins the common bile duct. Although both ducts may be retained throughout adult life in some species, such as the chick, horse, and dog, one or the other duct usually disappears, leaving a single pancreatic duct. Therefore, Wirsung's duct is the adult pancreatic duct in humans, sheep, ganoid fish, teleost fish, and the frog; on the other hand, the duct of Santorini serves the adult elasmobranch, pig, and ox.

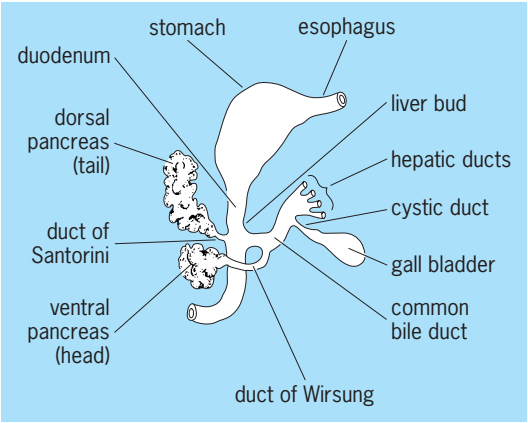

Fig. 3. Embryonic development of liver and gallbladder and of dorsal and ventral pancreatic buds before fusion of the latter. (*After G. C. Kent, Comparative Anatomy of the Vertebrates, McGraw-Hill, 1954*)

The original pancreatic primordia proliferate into masses of undifferentiated cells which then grow into solid cords of cells. These cords eventually are transformed into primitive pancreatic tubules. By a process of budding, the exocrine acini are formed from the primitive tubules. In addition, masses of loosely connected cells, centroacinar cells, are derived from the primitive tubules, and form an anastomosing network between the primary tubule and acinar system.

Endocrine elements. The endocrine cells of the pancreas are grouped in aggregates or islets of varying size which have no connection to the duct system. These aggregates, the islets of Langerhans, are highly vascularized and distributed throughout the organ. The distribution of islets is not uniform, and it has been suggested that this unequal distribution is a reflection of the fact that most of the endocrine cells are derived from the dorsal pancreas. This contention is best supported by the evidence presented from studies of certain urodele species in which fusion of the pancreatic diverticula occurs late and is restricted to the formation of a narrow isthmus of tissue between the pancreatic lobes. The existence of a separate giant islet of endocrine cells in certain teleosts also suggests the predominant role of one diverticulum in the development of the endocrine component.

Islet morphogenesis. The early endocrine elements of the pancreas arise from the cells of the diverticulum itself and from the primitive tubules. These cells form the primary islets of the pancreas, sometimes referred to as the islets of Laguesse. Considerable controversy has been waged over the eventual fate of the primary islets. The alternative views are (1) that the primary islets degenerate and are replaced by the definitive islet tissue, and (2) that the early islets persist into adult life. A second generation of islet cells occurs in the developing pancreas, although many investigators do not make a sharp distinction between primary and secondary islets. These secondary islets are derived largely from centroacinar cells and from the cells of the developing duct system. It has also been suggested that differentiated acinar cells can, by a process of dedifferentiation, give rise to the secondary islets. This contention, however, has been hotly disputed. Transitional stages in the formation of these islets of endocrine cells from ductules have been described in the embryonic organ. In adult organs of certain primitive species, such as the Elasmobranchii, these transitional stages may persist throughout adult life. For example, in some species of elasmobranchs, the islet cells are arranged in sheets next to the duct cells; in others, the islets may be connected to a ductule by a solid cord of cells. Phylogenetically, a progression of pancreatic types exists which resembles the development stages occurring in any one of the higher vertebrates.

Lower chordates. Scattered cells resembling pancreatic cells have been described in the digestive tube of the lancet *Branchiostoma*. In the lamprey

Petromyzon, the pancreas remains embedded in the wall of the intestine, having lost, secondarily, its pancreatic duct. The suggestion has been made that in the Cyclostomata, like the lamprey, the pancreas is primarily endocrine in function. Certain anatomical and histological features of the development of the digestive tract of the ammocoete larvae of the lamprey have led several investigators to speculate that this organism represents a transitional condition in the evolution of the vertebrate pancreas. At the junction of the esophagus and intestine, a mass of darkly staining cells forms from follicles budding off from the intestinal wall. These follicles lose their connection with the intestinal epithelium and, based on the histological features of their component cells, have been identified with the islet tissue of the higher vertebrates. In the anterior portion of the ammocoete intestine, two bilaterally arranged areas containing cells with a distinctly granular cytoplasm have been described, and it has been suggested that they are homologs of the exocrine portion of the vertebrate pancreas. The granules of these cells exhibit staining properties similar to the zymogen granules of the acinar cells of the pancreas of higher vertebrates. The production of a proteolytic enzyme of the tryptic type, thus similar to one of the digestive enzymes of the pancreas, has been localized in the anterior portion of the intestine. The hypothesis is further strengthened by the observation that in the larva of an Australian lamprey, *Geotria australis*, the intestine grows anteriorly as a pair of blind pouches. These pouches are lined with the granular cell type. It is difficult to escape the temptation to speculate that the ammocoete larva represents an evolutionary stage in which the endocrine pancreas has already separated from its original location and that the zymogenlike cells are preparing to follow suit, namely, the condition in the larva of *G. australis*. The next phylogenetic advance is observed in the elasmobranch. In these forms, a single dorsal pancreatic primordium and, consequently, the single duct of Santorini develop. Thus, phylogenetically as well as ontogenetically, the dorsal pancreas arises first. The first orders in which a pancreas similar to the mammalian organ appears are the ganoids and teleosts.

Another observation bespeaking the embryonic origin of the pancreas from the epithelium of the primitive digestive tract is the demonstration that cells with staining properties similar to those of the alpha cells occur in the pyloric portion of the bovine stomach. Extracts of this region give a positive test for the hormone glucagon.

Cell types. Within an islet of Langerhans, at least three distinct granular cell types can be distinguished on the basis of tinctorial differences among the granules demonstrable with polychrome stains. Acidophilic alpha cells are concerned with the synthesis of the hormone glucagon, and the basophilic beta cell is the site of insulin synthesis. The third cell type, the delta cell, may be a transitional form. In addition, a fourth, agranular, cell type may be present in some species. Although, in most species, each islet contains all of the representative cell types, it has been reported that, in the domestic fowl, the individual islets may be predominantly alpha-cell or beta-cell islets. Considerable variation exists as to the time and order of appearance during development of the cell types among the various species. In the rat, the eta cell can first be discerned in the $18^1/_2$-day fetus; alpha cells cannot be demonstrated until 2 days postpartum. The alpha cell, on the other hand, is the first islet cell type recognized in the chick pancreas, appearing on the eighth day of incubation. On the twelfth day, concomitantly with the appearance of large numbers of degenerating cells, beta cells can be distinguished. As yet, no definitive evidence is available which would permit a correlation between the detection of a particular cell type and the production of the hormonal product of that cell type. However, on the basis of two types of observations, it can be safely assumed that pancreatic hormones are elaborated during embryonic life. First, the effects of pancreatic insufficiency or ablation are temporarily alleviated in the pregnant female. Second, insulin and glucagon have been isolated from the pancreata of fetal cattle and swine, the embryonic organ yielding more hormone per unit weight than adult glands.

Irwin R. Konigsberg

Exocrine Physiology

The exocrine function of the pancreas is subserved by acinar and duct cells which form pancreatic juice. Acinar cells make up about 90% and duct cells only about 4% of pancreatic tissue, the rest being endocrine tissue. While secretion of electrolytes and water originates from both acini and excretory ducts, enzyme secretion occurs only in acinar cells. Enzymes are synthesized in the ribosomes attached to the surface of the endoplasmic reticulum, are then transported across the boundary of the endoplasmic reticulum, and are formed into granules within the intracisternal spaces of the reticulum. The granules move toward the Golgi apparatus and become invested with a membrane to become mature zymogen granules which are stored in the apical region of the acinar cell. On stimulation they fuse with the luminal plasma membrane and release their contents into the lumen. The zymogen granule membranes are retrieved and reutilized after degradation into proteins by lysosomes and resynthesis in the endoplasmic reticulum. The juice passes via the duct system to be collected by the main duct which opens into the duodenum, where digestion of proteins, fat, and carbohydrates takes place.

Stimulation of secretion. In humans some 500–800 ml of juice is secreted daily. Secretion is controlled by the vagus nerve and by two main peptide hormones; secretin and cholecystokinin-pancreozymin. While secretin evokes a plasmaisotonic fluid rich in bicarbonate from duct cells, both acetylcholine and cholecystokinin-pancreozymin stimulate enzyme secretion and in some animal species, such as rat, they also stimulate an isotonic NaCl fluid from acinar cells. With increasing flow rate as stimulated by secretin, bicarbonate concentration increases up to ~120 meq/liter and chloride

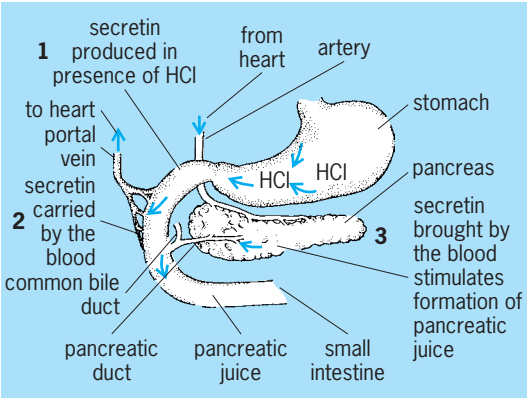

Fig. 4. The path and action of secretin in stimulating production of pancreatic juice. (*After T. I. Storer and R. L. Usinger, General Zoology, 3d ed., McGraw-Hill, 1958*)

concentration decreases, while sodium and potassium concentrations remain constant.

Secretin also evokes enzyme secretion from acinar cells in a variety of animal species. The intracellular mechanism of secretin-stimulated enzyme secretion, however, is different as compared to that of acetylcholine or cholecystokinin-pancreozymin–evoked enzyme secretion.

Intracellular messengers. Stimulation of enzyme secretion by secretagogues, such as acetylcholine or cholecystokinin-pancreozymin from pancreatic acinar cells, is mediated by rise of the cytosolic free Ca^{2+} concentration. This increase in Ca^{2+} is partly due to secretagogue-induced release of Ca^{2+} from intracellular stores and partly due to increased Ca^{2+} influx into the cell.

Enzyme secretion from pancreatic acinar cells is stimulated by secretin and mediated by increase in intracellular adenosine-3′,5′-cyclic monophosphate (cAMP). Fluxes of Ca^{2+} are not changed by secretin; rather, secretin potentiates the response of cholecystokinin-pancreozymin or of acetylcholine to stimulate enzyme secretion. Secretagogues which induce stimulation by mediation of Ca^{2+} also elicit a variety of other intracellular signals, such as increase in guanosine 3′,5′-cyclic monophosphate or of metabolites related to the hydrolysis of phospholipids, such as diacylglycerol, phosphatidic acid, arachidonic acid, and prostaglandins. For some, such as prostaglandins and diacylglycerol, participation in stimulus-secretion coupling seems to be likely, whereas for others, such as cyclic guanine monophosphate, a direct role in enzyme secretion cannot be assumed.

Regulation of secretion. Within a few minutes of taking food, pancreatic juice is secreted. This rapid response occurs reflexively via the vagus nerve and only occurs if the vagi are intact. Conversely, stimulation of the vagus provokes enzyme-rich juice.

The hormones secretin and pancreozymin are produced by the duodenal and jejunal mucosa. The acidity of the duodenal contents as well as fat and products of fat digestion are the most effective stimulants for secretin release, whereas fat, peptones, carbohydrates, and amino acids are most effective in liberating cholecystokinin-pancreozymin. On liberation into the portal venous blood, these hormones gain access to the systemic circulation and reach the pancreatic tissue via the arterial blood (**Fig. 4**).

Irene Schulz

Endocrine Physiology

In 1889 J. von Mehring and O. Minkowsky were the first to remove completely the pancreas of a dog, thereby causing true diabetes. This experiment is easily performed in most animals with the same results. In the rabbit, however, the complete removal of the gland is difficult, because of its ramified constitution. In birds, pancreatectomy is followed only by a very transient hyperglycemia. F. Banting and C. Best (1922) prepared pancreatic extracts which were able to prevent the lethal effects of pancreatectomy. The same effect was obtained with extracts from pancreas in which, after ligature of the duct of Wirsung, the exocrine portion of the gland had disappeared. The insular origin of the active factor, called insulin by De Meyer in 1909, was proved. A considerable amount of histological, embryological, experimental, and pathological material permits the differentiation of the roles of the two main cellular constituents of the islets of Langerhans, namely, the alpha and the beta cells. *See* ENDOCRINOLOGY; GLAND.

Effect of alloxan. The beta cells secrete insulin, that is, the hypoglycemic factor. The injection of alloxan, according to J. Dunn and N. McLetchie (1943), causes a selective destruction of the beta cells and the appearance of the various symptoms of diabetes. The mechanism of the diabetogenic action of alloxan is still a matter of discussion. Repeated injections of the extract of the anterior lobe of the pituitary gland cause alterations of the beta cells and diabetes; the severity of this diabetes is proportional to the degree of degranulation of the cells. The histological study of the pancreases of individuals who have suffered from diabetes frequently shows marked alterations of the beta elements. Certain types of pancreatic adenomas found in the human are essentially constituted by beta cells, and they are associated with hyperinsulinism and hypoglycemia. As further arguments in favor of the beta-cell origin of insulin, the damaging effect of dithizone on the dog's beta cells, associated with diabetes and with the disappearance of zinc from the cells, may be mentioned. Dialuric acid has identical effects, followed by regeneration of beta cells from acinous tissue. Various substances that have been studied by Kadota and Midokawa cause alterations of which the specificity is not evident. *See* DIABETES; INSULIN.

Effect of HGF. Intravenous injections of insulin cause a transitory hyperglycemia, soon followed by the typical hypoglycemia. This was attributed to the existence in the injected insulin or pancreatic extract of another substance, called the hyperglycemic factor or glucagon. This factor has been purified, and its principal effect is to cause hyperglycemia through a process of glycogenolysis. Thus, it is known as the hyperglycemic-glycogenolytic factor or HGF. Besides

the upper two-thirds of the gastric mucosa of dogs and rabbits, the pancreas is the only organ yielding HGF. The insular origin of HGF is shown by the strong correlation between glucagon content and density of insular tissue, the highest content being found in fetal pancreas and the tail part of the adult pancreas. Atrophy of the acinous tissue after ligation of the ducts and degeneration of the beta cells following injection of alloxane do not affect the glucagon content. *See* ENDOCRINE MECHANISMS.

Ernest L. Van Campenhout

Biochemistry

The alpha and beta cells in the islets of Langerhans are the sources of two hormones, insulin from the beta, and glucagon, also known as the hyperglycemic factor, from the alpha cells. The former is a hormone which influences carbohydrate metabolism, enabling the organism to utilize sugar. The latter accelerates the conversion of liver glycogen, the form in which carbohydrate is stored in liver and muscles until needed by the body, into glucose, the principal sugar used by the body to meet its energy requirements. Thus, glucagon elevates the blood sugar level, and its effects are the opposite of those of insulin, so that the two hormones together maintain the sugar metabolism of the body in balance. When the level of sugar in the blood becomes too low, the secretion of glucagon is stimulated.

Insulin. In the normal organism, the blood sugar level remains relatively constant, no matter what the intake or output. When the amount of sugar in the blood is raised, the islets of Langerhans in the pancreas increase the secretion of insulin, which then facilitates the oxidation of sugar as a source of energy. Although the correlation between pancreatic activity and diabetes was made as early as 1889 by J. von Mehring and O. Minkowsky, it was not until 1922 that insulin was discovered, by F. G. Banting, C. H. Best, J. J. R. Macleod, and J. B. Collip, as a life-saving treatment for diabetes mellitus. The hormone was prepared in crystalline form in 1926 by J. J. Abel and coworkers. The effect of insulin in diabetic patients was found to be remarkable; as little as 1 or 2 mg/day administered to an individual unable to manufacture insulin was sufficient to restore his whole sugar metabolism to normal. However, the exact mechanism whereby insulin exerts this action is still a subject of debate. Some investigators believe that the hormone acts on the permeability of the cell membrane, thus permitting the entrance of sugar into cells. Others believe that the enzyme hexokinase, which catalyzes the first reaction in the synthesis of glycogen from glucose, is controlled by insulin. Sugar metabolism is controlled by other factors besides insulin. For example, in 1924 A. B. Houssay demonstrated that the anterior pituitary plays an important role in the etiology of the diabetic state of an organism. *See* DIABETES.

The biological potency of insulin preparations is given in terms of international units (IU), defined as the hormonal activity of 0.125 mg of an international standard preparation. It is generally assayed by injecting the hormone subcutaneously into fasting rabbits and then determining the rate and extent of the lowering of the blood sugar level. Pure insulin possesses an activity equivalent to 24 IU/mg. Insulin loses its physiological activity when treated with alkali to pH 10 or with proteolytic enzymes such as chymotrypsin.

Insulin is a polypeptide with a molecular weight of approximately 6000. It is generally prepared from beef pancreas, and this bovine insulin contains 48 amino acids in two peptide chains (A chain and B chain) joined by sulfur atoms or —S—S—bridges. The sequence of amino acids constituting the primary structure of bovine insulin was described by F. Sanger and coworkers in 1954. In 1972 a three-dimensional (tertiary) structure of insulin was proposed by D. Hodgkin and coworkers; it was derived from the atomic arrangement found by x-ray analysis of rhombohedral crystals of the hormone.

Insulin preparations from other species, such as pig, sheep, horse, and whale, have also been examined by Sanger and coworkers and have been found similar in structure to bovine insulin, except for the amino acid sequence comprising positions 8, 9, and 10 in the A chain of the molecule. The biological potency of insulin has not been found to differ from species to species, nor has it been possible to differentiate the various preparations immunologically. Insulin was isolated from a primitive vertebrate, the Atlantic hagfish (*Myxine glutinosa*) by D. F. Steiner and coworkers in 1975. The complete amino acid sequence of the fish hormone was determined, and it was found that hagfish insulin is in many ways structurally similar to other vertebrate insulins. Apparently, the structure of the insulin molecule has been effectively conserved during the evolution of the vertebrates. The chemical synthesis of insulin was achieved by three groups of investigators in 1963–1965 and confirmed its chemical structure as proposed by Sanger.

In 1974 biologically active, crystallized human insulin was successfully synthesized in a laboratory.

The difference between human and pig insulin is located at the residue position 30 in the B chain: Thr in humans is replaced by Ala in pigs. The conversion of pig insulin to human insulin has been accomplished by chemical means.

While studying the biosynthesis of insulin in 1967, Steiner and P. E. Oyer demonstrated that the hormone itself is formed by the cleavage of the smaller insulin molecule from a larger parent molecule called proinsulin in which the polypeptide chains have been sequentially synthesized. In 1968 Steiner and coworkers and R. E. Chance and coworkers independently isolated and characterized bovine and porcine proinsulins. Chance and his group also determined the complete amino acid sequence of porcine proinsulin (**Fig. 5**).

The precursor forms of rat insulin were identified and characterized by D. F. Stein, W. Gilbert, and their collaborators in 1976–1978. The

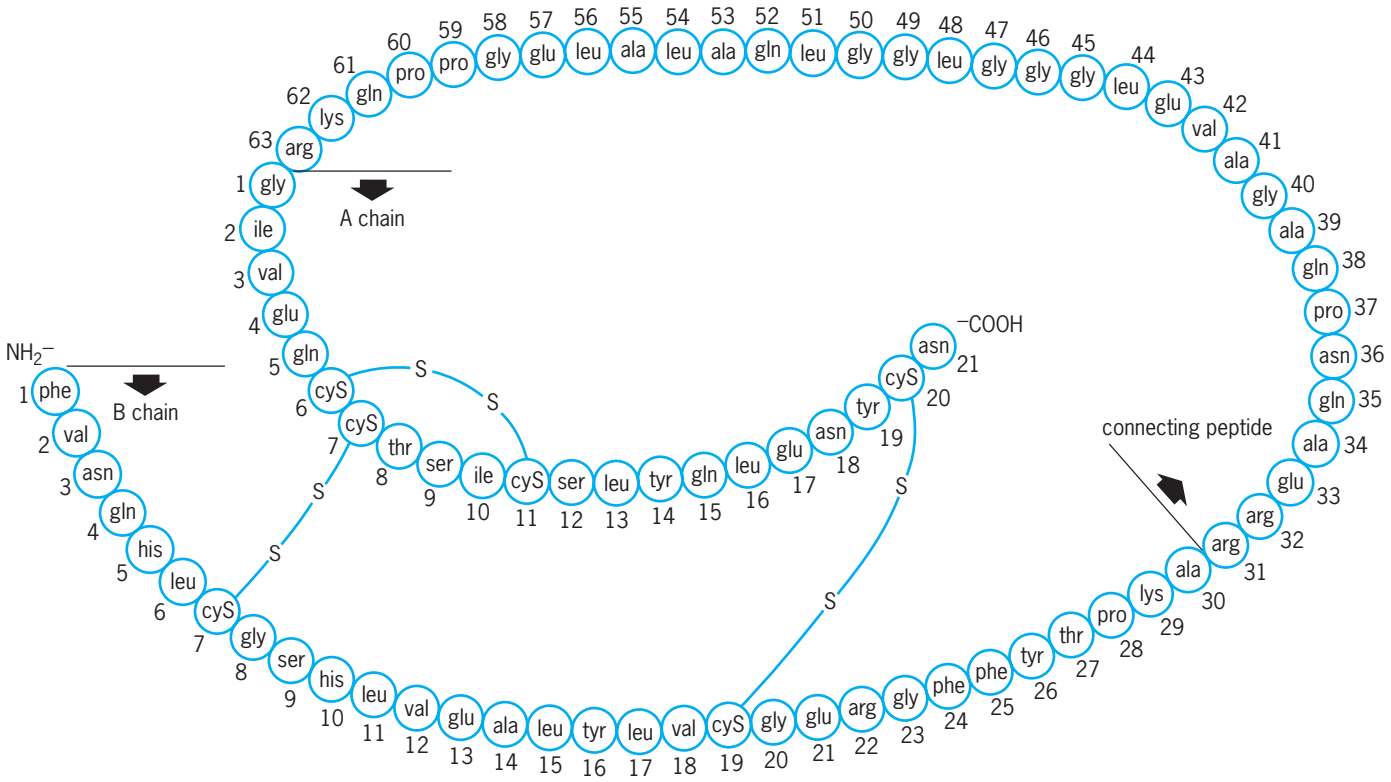

Fig. 5. Amino acid sequence of porcine proinsulin.

recombinant DNA–cloning synthesis of human insulin chains was achieved by D. V. Goeddel and coworkers in 1979.

The first structurally abnormal mutant insulin found in a diabetic patient was suggested in 1980 to be [Leu^{B24}]-insulin on the basis of its antagonistic effect. This was subsequently shown to be [Leu^{B25}]-insulin in 1983. In the same year, the second case of a diabetic patient with an abnormal insulin was demonstrated to be [Ser^{B24}]-insulin.

The initial step of insulin action is the insulin binding to its receptors. The binding kinetics of insulin-receptor interactions have been studied extensively and shown to be similar among different cells, with a dissociation constant of about 1 nM. The insulin receptor has been shown to be a glycoprotein consisting of two α (apparent molecular weight of 125,000) and two β (apparent molecular weight of 94,000) subunits linked by disulfide bonds.

Glucagon. The existence of glucagon in pancreatic extracts was first demonstrated by O. Kimball and J. Murlin in 1923. It is frequently present in commercial preparations of insulin as an impurity; the administration of such insulin preparations causes an initial rise in blood sugar, followed by the lowering of blood sugar (hypoglycemia) characteristic of insulin action. *See* GLUCAGON.

In 1955 A. Staub, O. K. Behrens, and their coworkers obtained glucagon in crystalline form; 2 years later the same group of investigators established the complete structural formula of the hormone as a single polypeptide chain of 29 amino acids with the sequence shown below, where His is histidine; Ser, serine; Gln, glutamine; Gly, glycine; Phe, phenylalanine; Thr, threonine; Asp, aspartic acid; Asn, asparagine; Tyr, tyrosine; Lys, lysine; Leu, leucine; Arg, arginine; Ala, alanine; Val, valine; Try, tryptophan; Met, methionine.

His	Ser	Gln	Gly	Thr	Phe	Thr	Ser
1	2	3	4	5	6	7	8
Asp	Tyr	Ser	Lys	Tyr	Leu	Asp	Ser
9	10	11	12	13	14	15	16
Arg	Arg	Ala	Gln	Asp	Phe	Val	Gln
17	18	19	20	21	22	23	24
		Try	Leu	Met	Asn	Thr	
		25	26	27	28	29	

The three-dimensional structure of crystalline glucagon was determined in 1975; total synthesis of glucagon was achieved in 1967 and again in 1975. The precursor forms of angelfish and rat glucagons have been identified and characterized with molecular weights of 12,000 and 18,000, respectively.

Pancreatectomy. Within 24 h after the surgical removal of the pancreas, hyperglycemia, or elevation of blood sugar, is detectable, followed by glycosuria, or excess sugar secreted in the urine. Without the pancreas, the organism is unable to store or to oxidize sugar. This disturbance of sugar metabolism is also accompanied, when there is pancreas deficiency, by defective fat metabolism. The organism, in an attempt to compensate for the loss of its primary source of energy, sugar, increases the catabolism or

breakdown of fats. When the stores of carbohydrate in the body are depleted, ketone bodies, acid products which are responsible for the metabolism of fats, increase in both the blood and the urine, with serious consequences. The accumulation of these acid products leads to a depletion from the blood of the alkali reserve that balances them in the normal organism. As a consequence, the organism is unable to remove accumulating carbon dioxide in the blood, resulting in acidosis, or air hunger, which leads to coma and death. These symptoms are known to be associated with diabetes mellitus, which in humans is characterized by loss of weight, hyperglycemia, glycosuria, ketonemia, ketonuria, and polyuria.

Another pancreatic disorder, which has been observed in rare instances of tumors of the islet tissue, is known clinically as hyperinsulinism. The disease, associated with low blood sugar and indistinguishable from the consequences of insulin overdosage, is caused by excessive production of insulin by the tumor. *See* CARBOHYDRATE METABOLISM; PANCREAS DISORDERS.

Choh Hao Li

Bibliography. R. E. Chance et al., *Science*, 161:165–167, 1968; V. L. W. Go (ed.), *The Pancreas: Biology, Pathology, and Disease*, 2d ed., 1993; D. C. Hodgkin, *J. Endocrinol.*, 63:1–14, 1974; P. G. Katasoyannis, *Science*, 154:1509–1514, 1966; I. Schulz, Messenger role of calcium in function of pancreatic acinar cells, *Amer. J. Physiol.*, 239:G335–G347, 1980; P. Sieber et al., *Helv. Chim. Acta*, 57:2617–2621, 1974; D. F. Steiner et al., *Ann. N. Y. Acad. Sci.*, 343:1–16, 1980; H. Streb et al., Release of Ca^{2+} from a nonmitochondrial intracellular store in pancreatic acinar cells by inositol-1,4,5-trisphosphate, *Nature*, 306:67–69, 1983; H. S. Tager et al., *Ann. N.Y. Acad. Sci.*, 343:133–147, 1980.

Pancreas disorders

The pancreas is affected by a variety of congenital and acquired diseases. Because of the dual functional role, the diseases of the exocrine portion of the pancreas will be separated from the endocrine lesions in this discussion.

Congenital disorders. The most frequent congenital lesion of the pancreas is more appropriately designated as a developmental abnormality—ectopic or aberrant pancreas. Ectopic pancreas can be found anywhere within the gastrointestinal tract, but is more frequent in the stomach and duodenum. Both acinar and islet tissue may be present in aberrant pancreas. Clinically these tissues have little significance.

Cystic fibrosis. Cystic fibrosis (mucoviscidosis) is a systemic disease in which mucus secretion is altered so that a viscid mucus is produced. The disease is inherited as a mendelian recessive. Cystic fibrosis affects all exocrine glands, including the acinar portion of the pancreas. Production of altered mucus leads to dilation of the exocrine ducts (cystic), destruction of acinar tissue, and replacement of the destroyed tissue by fibrous connective tissue (fibrosis). The islets are not affected by this disease. Elevation in secretion of sodium and chloride in sweat is also common. In addition to affecting the pancreas, the sweat glands, gastrointestinal tract, lungs, liver, and salivary glands are also involved in cystic fibrosis. Destruction of the pancreas can be so extensive that pancreatic insufficiency develops. Careful medical management successfully compensates for pancreatic insufficiency. Involvement of the lungs leads to pneumonia, bronchiectasis, and often death during childhood.

Pancreatitis. Acute hemorrhagic pancreatitis is a serious disease of unknown etiology which causes sudden liberation of activated pancreatic enzymes that digest the pancreatic parenchyma. The digestive process leads to dissolution of fat and production of calcium soaps. In addition, rupture of pancreatic vessels occurs with resultant hemorrhage and shock. This disease is associated with biliary tract disease, especially gallstones (cholelithiasis), alcoholism, hyperlipidemia, and hypercalcemia. *See* ALCOHOLISM; GALLBLADDER DISORDERS.

Chronic pancreatitis, perhaps better designated chronic relapsing pancreatitis, is a condition in which recurrent episodes of pancreatitis occur without the production of symptoms or with the production of mild symptoms. Destruction of the pancreatic tissue, with repair by fibrosis, calcification, and cyst formation, is frequent.

Diabetes. Diabetes mellitus is the principal disease associated with the endocrine portion of the pancreas. Two clinical forms of the disease are recognized—insulin-dependent diabetes mellitus and non-insulin-dependent diabetes mellitus. While many factors are involved in the causation of this disease, basically the disease is a result of the failure of the beta cells of the pancreas to produce appropriate kinds and amounts of insulin to meet metabolic needs. Diabetes mellitus is a disease affecting many organs—a systemic disease. Long-term consequences include the early development of lesions in arteries, especially atherosclerosis, the development of lesions in small blood vessels throughout the body (microangiopathy), and abnormalities of the kidneys and nerves.

The insulin-dependent form frequently occurs in the younger age group, hence it is referred to as juvenile-onset diabetes mellitus. Factors involved in the pathogenesis of this form of diabetes mellitus include injury to the islet cells by viral infection or autoimmune processes. Insulin is necessary in the therapy of these individuals. The majority of individuals with diabetes mellitus develop the disease process as adults—maturity-onset (non-insulin-dependent) diabetes mellitus. Many factors are included in the causation of diabetes mellitus in this group. In maturity-onset diabetes mellitus, the body may have difficulty in utilizing the insulin produced by the pancreatic islets. Increased production of another pancreatic hormone, glucagon, may also play a role. Genetic influences are certainly participating in the development of diabetes mellitus, but the exact role of inheritance has yet to be defined. Environmental factors such as obesity, pregnancy, and infections also

contribute to the development of the disease. *See* DIABETES.

Tumors. Tumors of the pancreas can be either benign or malignant. They affect both the endocrine and exocrine portions of the pancreas.

Benign tumors of the exocrine pancreas are extremely rare. Malignant tumors of the exocrine pancreas arise most frequently from the pancreatic ducts. Acinar carcinomas also exist but are very rare. Exocrine pancreatic carcinomas are very malignant tumors. They are clinically silent until well advanced. Cure by surgery or other therapeutic approaches is rare. Carcinoma of the pancreas affects individuals in the 60- to 80-year-old age group and occurs most frequently in the head and least frequently in the tail of the pancreas. Individuals develop jaundice, as obstruction of the common bile duct is frequent in this disease. Involvement of the nerves in and around the pancreas cause severe back pain. The most frequent histologic type of pancreatic carcinomas is a glandular pattern (adenocarcinoma).

Islet cell lesions are quite rare but may be associated with increased hormone production. The tumors can be single or multiple, benign or malignant, and they can form anywhere in the pancreas. Hyperfunction of the islets of Langerhans can result in three distinct clinical syndromes: hyperinsulinism and hypoglycemia, the Zollinger-Ellison syndrome (gastrinoma), and multiple endocrine neoplasia.

Beta-cell tumors (insulinomas) are the most common tumor, with elaboration of insulin and production of hypoglycemia. Most insulinomas are the result of a solitary benign tumor (adenoma). However, multiple adenomas, islet-cell carcinoma, and islet-cell hyperplasia can also cause the syndrome.

The Zollinger-Ellison syndrome is characterized by a pancreatic islet tumor that produces excessive amounts of gastrin, with a resultant hypersecretion of the mucosa of the stomach and ulceration of the upper gastrointestinal tract (peptic ulceration). The majority of gastrinomas are malignant, with metastasis to lymph nodes and liver. Gastrin is not produced by islet cells under normal conditions.

Individuals with islet-cell tumors may have tumors in other endocrine organs—a condition known as multiple endocrine neoplasia syndrome—including the pituitary, thyroid, parathyroid, and adrenal glands. Several different syndromes have been recognized and are characterized by the organs involved. *See* ONCOLOGY; PANCREAS. H. Thomas Norris

Bibliography. G. P. Burns and S. Bank, *Disorders of the Pancreas*, 1989; R. S. Cotran, V. Kumar, and S. L. Robbins (eds.), *Robbins' Pathologic Basis of Disease*, 6th ed., 1999.

Panda

The family Ailuropodidae contains two species of pandas—the giant panda (*Ailuropoda melanoleuca*) and the lesser or red panda (*Ailurus fulgens*). Until relatively recently, the giant panda had been classified in the family Ursidae with the bears, and the red panda had been included in the family Procyonidae along with raccoons, ringtails, and coatis.

Fig. 1. Giant panda (*Ailuropoda melanoleuca*) (*Copyright © 2001 John White*)

Giant panda. Giant pandas are an endangered species, surviving only in small isolated populations in China. Approximately 1000 giant pandas live in an area of about 29,500 km^2 (10,600 mi^2). They inhabit a narrow zone of bamboo forest from 1200 to 3300 m (3800 to 10,600 ft) in elevation. Below this zone is cultivated land; above this zone is the treeline. There are currently 12 wildlife reserves comprising about 6000 km^2 (2200 mi^2) in which the pandas live.

Giant pandas have massive heads and bodies with an unmistakable black and white fur pattern (**Fig. 1**). The body is whitish except for the ears, eye spots, nose, limbs, and shoulders which are black. The limbs are relatively short, and each forepaw possesses a "pseudothumb," or sixth toe, which is an adaptation for stripping bamboo leaves from the stalk. The dental formula is I 3/3, C 1/1, Pm 4/4, M 2/2 × 2, for a total of 40 teeth. Adults are 150–180 cm (5-6ft) long with a 10–15 cm (4–6 in.) tail. They weigh 75–110 kg (165–242 lb). *See* DENTITION.

Giant pandas are active throughout the year. They are solitary and have broadly overlapping home ranges. Their primary food consists of the sprouts, stems, and leaves of bamboo, although they occasionally consume bulbs, grasses, insects, and rodents. Breeding occurs from mid-March to mid-May. Following a gestation period of 97 to 163 days, one or two cubs are born in a hollow tree or cave, mainly in August and September. Gestation is lengthened because, like bears and some other carnivores, pandas experience delayed implantation of the fertilized egg in the uterine wall. Newborn pandas are blind, pink, and almost hairless. The eyes open at about 45 days, and cubs begin feeding on bamboo at about 5 months of age. They remain with their mother for about 18 months; thus, females give birth only once every 2 or 3 years from age 4 to at least 20. Reproduction in captivity is poor, although artificial insemination and hand raising of young are increasingly successful. A captive specimen lived to an age of approximately 34 years.

Red panda. Red pandas have long, soft, thick rusty to chestnut-brown fur on their dorsal surface (**Fig. 2**). Small dark-colored patches are present

Fig. 2. Red panda (*Ailurus fulgens*). (*Copyright © 2004 John White*)

beneath each eye, and the muzzle, lips, cheeks, and edges of the ears are white. The backs of the ears, the limbs, and the underparts are dark reddish-brown to black. The claws are sharp and semiretractile, and the feet have hairy soles. The soles of the feet have numerous glands opening through minute pores, whose secretions are used for marking runways and trails. Locomotion is plantigrade (individuals walk on the soles of their feet). The bushy nonprehensile tail has 12 inconspicuous dark brown rings on a reddish background. The head is rounded with a short snout, and the ears are large and pointed. The dental formula is I 3/3, C 1/1, Pm 3/4, M 2/2, for a total of 38 teeth. Adults are 79–110 cm (31–44 in.) in total length, including a 28–48 cm (11–19 in.) tail. They weigh 3–6 kg (8–13 lb).

This species inhabits the mountain forests of China and the southeastern mountainsides of the Himalayas ranging in elevation from 2200 to 4800 m (7000 to 15,000 ft). Preferred habitat consists of giant rhododendron, oak, and bamboo forests. They are generally solitary and primarily nocturnal, although researchers have found them to be active in daylight, especially during the summer. Although red pandas are capable climbers, they do most of their feeding on the ground. The principal food is bamboo, but the diet is supplemented by grasses, acorns, fruits, roots, lichens, insects, eggs, young birds, and small rodents. A "pseudothumb" is present to facilitate the handling of bamboo leaves and stems.

Breeding occurs from January to March. Following a gestation of 90–145 days, one to four (usually two) fully haired, blind cubs are born from mid-June to late July in a nest in a hollow tree or a rock crevice. The eyes open at 17–18 days. They attain their adult coloration, leave the nest, and begin feeding on solid food at about 90 days of age. Sexual maturity is attained at 18 months. They may live up to 14 years.

The specialized diet, low reproductive rate, and low population density make this species highly vulnerable. It is classified as endangered by the IUCN (World Conservation Union) which estimates that fewer than 2500 mature individuals exist. *See* CARNIVORA; MAMMALIA. Donald W. Linzey

Bibliography. *Grzimek's Encyclopedia of Mammals*, McGraw-Hill, 1990; D. Lindburg and K. Baragona, *Giant Pandas: Biology and Conservation*, University of California Press, 2004; D. Mcdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford Limited, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; G. B. Schaller et al., *The Giant Pandas of Wolong*, University of Chicago Press, 1985; G. B. Schaller, *The Last Panda*, University of Chicago Press, 1993.

Pandanales

An order of monocotyledons, the composition of which only recently has been revealed by deoxyribonucleic acid (DNA) sequence studies of four genes. Included are four families, two of which have been generally viewed as closely related: Pandanaceae (800 species; the screw pine family) and Cyclanthaceae (230 species; the Panama hat family). These two families have often been considered to be related to Arecales (with a single family, Arecaceae, palms), but actually the relationship is distant. In addition, the order includes Stemonaceae (35 species) and Velloziaceae (200 species), which have been thought to be related to Dioscoreaceae (Dioscoreales) and Bromeliaceae (Commelinales) or Hypoxidaceae (Asparagales), respectively. Pandanaceae are often lianas or large herbs from the Old World; Cyclanthaceae are herbs or lianas from the New World tropics; Stemonaceae are herbs or lianas of the Old World (but with one species in the southeastern United States); and Velloziaceae are herbs or small shrubs of Africa and particularly South America (with one genus in southwest China). Cyclanthaceae, Pandanaceae, and Stemonaceae have flower parts in twos or fours, which is unusual among monocotyledons, in which threes are most common.

Several species in Pandanales are economically important. The leaves of Pandanaceae are fibrous and used for making rope and roofing, and the fruits are eaten in many areas. Cyclanthaceae leaves are fibrous and have similar uses, including the manufacture of Panama hats. *See* ARECIDAE; LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM. Mark W. Chase

Panderichthys

A fossil lobe-finned (sarcopterygian) fish from rock strata of the Baltic region dating to the Middle–Late Devonian Period (about 372–368 million years ago), which provides key evidence about the evolutionary transition from fish to land vertebrates (tetrapods). As the most tetrapod-like fish known, it reveals the earliest steps of the transformation that adapted the fish body plan to terrestrial life. *See* SARCOPTERYGII; TETRAPODA.

Distribution. Fragmentary fossils of *Panderichthys* are widespread in Latvia, Estonia, and western Russia, in sandstones and clays that formed in river deltas along a tropical coast. But only one locality—the clay quarry of Lode in Latvia—yields complete

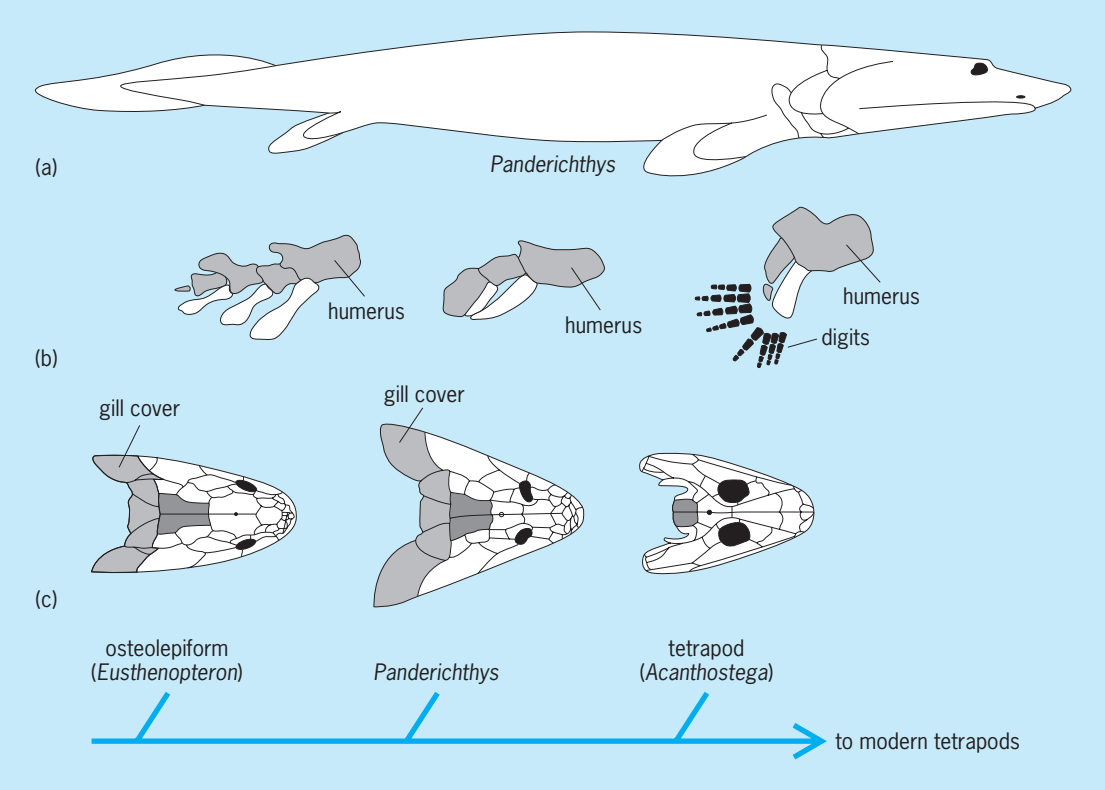

***Panderichthys anatomy*. (*a*) *Panderichthys rhombolepis*, whole-body reconstruction based on specimens from Lode, Latvia. Total length approximately 1.3 m (4.3 ft). Note the crocodile-like head shape, raised "eyebrows," and loss of all midline fins except the tail (*reprinted by permission from Nature, P. E. Ahlberg and A. R. Milner, The origin and early diversification of tetrapods, 368:507–514, © 1994, Macmillan Publishers Ltd.*). (*b*) Pectoral fin/forelimb skeletons of *Panderichthys*, an osteolepiform, and a primitive tetrapod showing proportional changes and the abrupt appearance of digits in the earliest tetrapods. (*c*) Skulls of *Panderichthys* and a primitive tetrapod in dorsal view showing proportional changes (note the position of the eyes and the size of the shaded bones) and the loss of the gill cover during the fish-tetrapod transition.**

bodies that reveal what the animal looked like (**illus.** *a*). Its anatomy proves to be a curious mixture of fish and tetrapod traits.

Anatomy. *Panderichthys* fits into the tetrapod phylogeny (family tree) above the osteolepiform fishes and immediately below the first tetrapods such as *Ichthyostega* and *Acanthostega*. It shares certain tetrapod-like characteristics with the osteolepiforms, such as a pair of internal nostrils (choanae) on the palate, and paired fin skeletons that contain recognizable equivalents of the major tetrapod limb bones, such as the humerus and femur (illus. *b*). However, whereas osteolepiforms still had the body form and fins of "normal" open-water fishes, *Panderichthys* looked quite different: it had a crocodile-like head with a long snout and eyes raised on top under distinct bony "eyebrows" (illus. *c*); a body that is slightly flattened from top to bottom; and a tail carrying only a simple fin fringe (illus. *a*), rather than the separate dorsal, caudal, and anal fins of osteolepiforms. All these features are shared with early tetrapods.

The internal anatomy reveals additional tetrapod features that are not present in osteolepiforms. For example, the skeleton of the pectoral fin, equivalent to the forelimb of tetrapods (illus. *b*), and shoulder girdle resembled those of early tetrapods, not just in construction but in shape, suggesting that the range of movement was becoming limblike. The ribs, which help support the internal organs of the body, were also much larger than in osteolepiforms. However, unlike tetrapods, *Panderichthys* retained well-developed gills, and its paired appendages were still fins with fin rays, not limbs with toes. *See* ICHTHYOSTEGA; OSTEOLEPIFORMES.

Mode of life. These features not only identify *Panderichthys* as a close relative of tetrapods but give hints to its mode of life. Its pectoral fins appear adapted for some degree of "walking," whether on land or in water, while the crocodile-like body form and raised eyes suggest that it operated in very shallow water and habitually looked out above the surface. The retention of gills and a tail fin indicate that *Panderichthys* had not really become terrestrial, while sharp, pointed teeth show that it was a formidable predator. *Panderichthys* can be envisioned as a specialist shallow-water predator of deltaic environments, which used its fins to crawl through water too shallow to support its body, and perhaps also for short excursions over land from one channel to the next. It is possible that it exploited the intertidal zone, emerging to feed on fishes stranded on mudflats and in tidal pools. *See* ANIMAL EVOLUTION; DEVONIAN; FOSSIL.

Per E. Ahlberg

Bibliography. J. A. Clack, *Gaining Ground*, Indiana University Press, 2002; J. A. Long, *The Rise of Fishes*, John Hopkins University Press, 1996; C. Zimmer, *At Water's Edge*, Free Press, New York, 1998.

Panel heating and cooling

A system in which the heat-emitting and heat-absorbing means is the surface of the ceiling, floor, or wall panels of the space which is to be environmentally conditioned. The heating or cooling medium may be air, water, or other fluid circulated in air spaces, conduits, or pipes within or attached to the panels. For heating, electric current may flow through resistors in or on the panels. *See* ELECTRIC HEATING.

Warm or cold water is circulated in pipes embedded in concrete floors or ceilings or plaster ceilings or attached to metal ceiling panels. The coefficient of linear expansion of concrete is 0.000095; for steel it is 0.00081, or 15% less than for concrete. For copper it is 0.000112, or 20% more than for concrete, and for aluminum it is 0.000154, or 60% more than for concrete. Since the warmest or coolest water is carried on the inside of the pipes and the heat is transmitted to the concrete, only steel pipe should be used for panel heating and cooling systems, except when metal panels are used.

Cracks are very likely to develop in the concrete or plaster, breaking the bonds between the pipes and the concrete or plaster. The pipes move freely, causing scraping noises. An insulating layer of air is formed between the concrete or plaster, and this markedly reduces the coefficient of conductivity between the liquid heating medium and the active radiant surfaces.

Heat transfer. Heat energy is transmitted from a warmer to a cooler mass by conduction, convection, and radiation. Radiant heat rays are emitted from all bodies at temperatures above absolute zero. These rays pass through air without appreciably warming it, but are absorbed by liquid or solid masses and increase their sensible temperature and heat content. *See* HEAT TRANSFER.

The output from heating surfaces comprises both radiation and convection components in varying proportions. In panel heating systems, especially the ceiling type, the radiation component predominates. Heat interchange follows the Stefan-Boltzmann laws of radiation; that is, heat transfer by radiation between surfaces visible to each other varies as the difference between the fourth power of the absolute temperatures of the two surfaces, and is transferred from the surface with the higher temperature to the surface with the lower temperature.

The skin surface temperature of the human body under normal conditions varies from 87 to 95°F (31 to 35°C) and is modified by clothing and rate of metabolism. The presence of radiating surfaces above these temperatures heats the body, whereas those below produce a cooling effect. *See* RADIANT HEATING.

Cooling. When a panel system is used for cooling, the dew-point temperature of the ambient air must remain below the surface temperature of the heat-absorbing panels to avoid condensation of moisture on the panels. In regions where the maximum dew-point temperature does not exceed 60°F (16°C), or possibly 65°F (18°C), as in the Pacific Northwest and the semiarid areas between the Cascade and Rocky mountains, ordinary city water provides radiant comfort cooling. Where higher dew points prevail, it is necessary to dehumidify the ambient air. Panel cooling effectively prevents the disagreeable feeling of cold air blown against the body and minimizes the occurrence of summer colds. *See* DEHUMIDIFIER.

Fuel consumption records show that panel heating systems save 30–50% of the fuel costs of ordinary heating systems. Lower ambient air temperatures produce a comfortable environment, and air temperatures within the room are practically uniform and not considerably higher at the ceiling, as in radiator- and convector-heated interiors. *See* COMFORT HEATING; HOT-WATER HEATING SYSTEM.

Erwin L. Weber; Richard Koral

Bibliography. American Society of Heating, Refrigerating, and Air Conditioning Engineers, *ASHRAE Handbook and Product Directory: Systems*, 1992.

Pantodonta

An extinct order of relatively large placental mammals represented by Paleocene-Eocene fossils from western Europe, North America, and eastern Asia. Pantodonts were an early evolutionary experiment in large-bodied herbivory by primitive placental mammals. They first appeared in Asia during the early Paleocene and disappeared during the middle Eocene, leaving no descendants. With the possible exception of the most primitive pantodonts, all were herbivores, and pantodonts were either the largest or among the largest mammals of their time. The adaptive radiation of pantodonts was diverse and encompassed mammals as different as small [1 kg (2 lb) or less in body mass], arboreal herbivores (Asian *Archaeolambda*), and large [650 kg (1430 lb)], ground-sloth-like, terrestrial herbivores (North American *Barylambda*). *See* EOCENE; PALEOCENE.

Pantodonts are unique among placental mammals in having upper third and fourth premolar tooth crowns in which there are V-shaped crests. The lower premolars and molars had broad, transverse crests, and this feature suggests that pantodonts sheared and pulped plant matter as their primary, or only, source of food. All pantodonts were obligate quadrupeds, with four nearly equal-sized limbs. Their wrists and ankles were generally short and massive, and each hand or foot had five toes that bore small hooves, except in a few pantodont genera that had claws. Pantodont skulls were generally small relative to body size and had long, boxlike snouts, small braincases, and small eyes. Most pantodonts had large tusks and long, heavy tails. When compared to living mammals, many pantodonts would have looked somewhat like bears, pigs, or small hippos. *See* DENTITION.

The most primitive pantodonts are assigned to the family Bemalambdidae. All other pantodonts belong to the suborder Eupantodonta, which diverged early into the superfamilies Pantolambdodontoidea

and Pantolambdoidea. The pantolambdodontoids include the Asian family Pantolambdodontidae, the North American family Titanoideidae, and the South American genus *Alcidedorbigniya*. The pantolambdoids include the North American families Pantolambdidae and Barylambdidae and the more cosmopolitan (known from North America and Eurasia) family Coryphodontidae.

The fossil evidence suggests that pantodonts originated in Asia from an insectivorous ancestry during the earliest Paleocene. Thus, the Chinese early Paleocene bemalamdid genera *Bemalambda* and *Hypsilolambda* are the oldest and the most primitive pantodonts. Unlike later pantodonts, their sharp and pointed tooth crowns indicate that they were insectivores or perhaps omnivores.

From their Asian origin, pantodonts rapidly emigrated into North and South America during the early Paleocene. By the late Paleocene, pantolambdodontoids were present in Asia (*Harpyodus*, *Pastoralodon*, *Archaeolambda*, *Altilambda*, and *Pantolambdodon*), North America (*Titanoides*), and South America (*Alcidedorbigniya*). The Asian (and South American) evolution of pantolambdodontoids encompassed small mammals [1 kg (2 lb) or less body mass]. Assuming that the skeleton of Chinese late Paleocene *Archaeolambda tabiensis* is characteristic, Asian pantolambdodontoids were lightly built, arboreal herbivores. By contrast, the North American pantolambdodontoid *Titanoides* was a large [150 kg (330 lb)], terrestrial, clawed quadruped with saber-like upper canines. This animal may have done some digging to obtain its food.

The diversification of the remaining North American pantodonts began during the early Paleocene with *Pantolambda*, a terrestrial quadrupedal browser. Late Paleocene *Caenolambda* was a somewhat unusual (notably, the bladelike first premolar), though functionally similar close relative. These two genera constitute the family Pantolambdidae.

The late Paleocene barylambdid pantodonts were confused early with Asian pantolambdodontids because both families converged on some dental features, especially in their small tusks and incisors. However, the skeletons of the pantodonts in these families are very different. *Barylambda* is the best-known barylambdid and was functionally somewhat analogous to the large ground sloths of the Pleistocene. The graviportal pelvis, massive hindlimbs, and heavy tail suggest that *Barylambda* was capable of bipedal browsing. *Leptolambda* shows some of these modifications as well, but the two smaller barylambdids, *Haplolambda* and *Ignatiolambda*, are too poorly known postcranially to allow firm conclusions on their mode of life.

From the standpoint of abundance of fossils, longevity, diversity, and breadth of geographic distribution, *Coryphodon* was the most successful pantodont. It was the largest land mammal of the latest Paleocene–early Eocene, and its fossils come from the western United States, Ellesmere Island in Arctic Canada, western Europe, Mongolia, and China. *Coryphodon* was a terrestrial, graviportal, subdigitigrade quadruped with an adult body weight of 150–300 kg (330–660 lb). Its teeth somewhat resemble those of a living tapir, suggesting that *Coryphodon* was a browser. There is no compelling evidence that it was amphibious other than the abundance of its fossils in some fluvial and lacustrine deposits. The skeleton of *Coryphodon* most closely resembles that of the living pygmy hippopotamus *Hexaprotodon*.

The extinction of *Coryphodon* at the end of the early Eocene marked the disappearance of pantodonts in North America and Europe. But in Asia, other coryphodontid genera (*Eudinoceras*, *Hypercoryphodon*) persisted through the middle Eocene. Their extinction ultimately was due to changes in climate and vegetation that promoted the rise of early tapirs and rhinoceroses. *See* MAMMALIA.

Spencer G. Lucas

Bibliography. M. J. Benton, *Vertebrate Palaeontology*, 1990; R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988; C. M. Janis, K. M. Scott, and L. L. Jacobs, *Evolution of Tertiary Mammals of North America*, 1998; R. J. G. Savage and M. R. Long, *Mammal Evolution: An Illustrated Guide*, 1986.

Pantograph

A four-bar parallel linkage, with no links fixed, used as a copying device for generating geometrically similar figures, larger or smaller in size, within the limits of the mechanism. In the **illustration** the curve traced by point *T* will be similar to that generated by point *S*. This similarity results because points *T* and *S* will always lie on the straight line $\overline{OTS}$; triangles *OBS* and *TCS* are always similar because lengths $\overline{OB}$, $\overline{BS}$, $\overline{CT}$, and $\overline{CS}$ are constant and $\overline{OB}$ is always parallel to $\overline{CT}$. Distance $\overline{OT}$ always maintains a constant proportion to distance $\overline{OS}$ because of the similarity of the above triangles. Numerous modifications of the pantograph as a copying device have been made. *See* FOUR-BAR LINKAGE.

James Watt applied the pantograph as a reducing motion in his beam engine. The pantograph has also served as a reducing motion for an engine indicator: *S* is attached to the engine crosshead, while the indicator cord is attached to *T*. *See* STRAIGHT-LINE MECHANISM.

A second use of the pantograph geometry is seen in the collapsible parallel linkage used on electric locomotives and rail cars to keep a current-collector bar or wheel in contact with an overhead wire. Two such congruent linkages in planes parallel to the

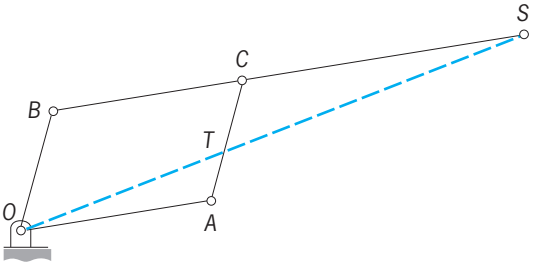

Similar triangles of a pantograph.

train's motion are affixed securely on the top of the locomotive with joining horizontal members perpendicular to each other. The uppermost member collects the current, and powerful springs thrust the configuration upward with sufficient pressure normally to make low-resistance contact from wire to collector. The inevitable variation of distance from wire to track produces, during high speeds, undesirable dynamic behavior of both wire and collector-pantograph, with oscillations of both that can break the connection between them or cause excessive arcing and wear with attendant loss of transmission. Stitching the wire to a stouter supporting wire can minimize the bounce of both collector and wire and assure more reliable current flow. Douglas P. Adams

Bibliography. N. P. Chironis, *Mechanisms and Mechanical Devices Sourcebook*, 3d ed., 2001.

Papaverales

An order of flowering plants, division Magnoliophyta (Angiospermae), subclass Magnoliidae of the class Magnoliposida (dicotyledons). The order consists of only two families: Papaveraceae with some 200 species, and Fumariaceae with about 400 species. Within its subclass, the order is marked by its syncarpous gynoecium, parietal placentation, and only two (seldom three) sepals. Most of the species are herbaceous, and many of them contain isoquinoline alkaloids similar to those in the order Ranunculales. The Papaveraceae, with regular flowers, numerous stamens, and a well-developed latex system, include the poppies (*Papaver* and related genera; see **illus.**), bloodroot (*Sanguinaria*), and celandine (*Chelidonium*). *Papaver somniferum* is the source of opium. The Fumariaceae, with four or six stamens, irregular flowers that usually have some of the petals spurred or saccate, and no latex system, include the bleeding heart (*Dicentra spectabilis*) and some other common ornamentals.

The Papaverales have often been grouped with the Capparales into an expanded order, Rhoeadales, but more recently it has been agreed that the two groups are not closely allied and that the affinities of the Papaverales are with the families of the Ranunculales. *See* MAGNOLIIDAE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM; POPPY; RANUNCULALES. Arthur Cronquist; T. M. Barkley

Oriental poppy (*Papaver orientale*) of the family Papaveraceae and the order Papaverales. (*John H. Gerard, National Audubon Society*)

Paper

A flexible web or mat of fibers isolated from wood or other plants materials by the operation of pulping. Nonwovens are webs or mats made from synthetic polymers, such as high-strength polyethylene fibers, that substitute for paper in large envelopes and tote bags.

Paper is made with additives to control the process and modify the properties of the final product. The fibers may be whitened by bleaching, and the fibers are prepared for papermaking by the process of refining. Stock preparation involves removal of dirt from the fiber slurry and mixing of various additives to the pulp prior to papermaking. Papermaking is accomplished by applying a dilute slurry of fibers in water to a continuous wire or screen; the rest of the machine removes water from the fiber mat. The steps can be demonstrated by laboratory handsheet making, which is used for process control.

Although paper has numerous specialized uses in products as diverse as cigarettes, capacitors, and counter tops (resin-impregnated laminates), it is principally used in packaging (~50%), printing (~40%), and sanitary (~7%) applications. Paper was manufactured entirely by hand before the development of the continuous paper machine in 1804; this development allowed the United States production of paper to increase by a factor of 10 during the nineteenth century and by another factor of 50 during the twentieth century. In 1960, the global production of paper was 70 million tons (50% by the United States); in 1996 it was 240 million tons (30% by the United States). The annual per capita paper use in the United States is 300 kg (660 lb), and about 45% is recovered for reuse.

Material of basis weight greater than 200 g/m^2 is classified as paperboard, while lighter material is called paper. Production by weight is about equal for these two classes. Paperboard is used in corrugated boxes; corrugated material consists of top and bottom layers of paperboard called linerboard, separated by fluted corrugating paper. Paperboard also includes chipboard (a solid material used in many cold-cereal boxes, shoe boxes, and the backs of paper tablets) and food containers.

Mechanical pulp is used in newsprint, catalog, and other short-lived papers; they are only moderately white, and yellow quickly with age because the lignin is not removed. A mild bleaching treatment (called brightening) with hydrogen peroxide or sodium dithionite (or both) masks some of the color of the lignin without lignin removal. Paper made with mechanical pulp and coated with clay to improve brightness and gloss is used in 70% of magazines and catalogs, and in some enamel grades. Bleached chemical pulps are used in higher grades of printing papers used for xerography, typing paper, tablets, and envelopes; these papers are termed uncoated wood-free (meaning free of mechanical

pulp). Coated wood-free papers are of high to very high grade and are used in applications such as high-quality magazines and annual reports; they are coated with calcium carbonate, clay, or titanium dioxide.

Like wood, paper is a hygroscopic material; that is, it absorbs water from, and also releases water into, the air. It has an equilibrium moisture content of about 7–9% at room temperature and 50% relative humidity. In low humidities, paper is brittle; in high humidities, it has poor strength properties.

Wood. Wood, a diverse, variable material, is the source of about 90% of the plant fiber used globally to make paper. Straw, grasses, canes, bast, seed hairs, and reeds are used to make pulp, and in many regards their pulp is similar to wood pulp. Fibers are tubular elements of plants and contain cellulose as the principal constituent. Softwoods (gymnosperms) have fibers that are about 0.12–0.2 in. (3–5 mm) long, while in hardwoods (angiosperms) they are about 0.03–0.06 in. (0.8–1.6 mm) long. In both cases, the length is typically about 100 times the width. Softwoods are used in papers such as linerboard where strength is the principal intent. Hardwoods are used in papers such as tissue and printing to contribute to smoothness. Many papers also include some softwood pulp for strength and some hardwood pulp for smoothness.

Wood consists of three major components: cellulose, hemicellulose, and lignin. The first two are white polysaccharides of high molecular weight that are desirable in paper. Cotton is over 98% cellulose, while wood is about 45%. Hemicellulose, although not water soluble, is similar to starch. The hydroxyl groups of these materials allow fibers to be held together in paper by hydrogen bonding. Adhesives are not required to form paper, but some starch is usually used and has a similar effect to the hemicelluloses in helping the fibers bond together. Lignin makes up about 25–35% of softwoods and 18–25% of hardwoods. It is concentrated between fibers. In mechanical pulping, the lignin is made relatively soft so the fibers can be separated by mechanical action, but the lignin is not removed. In chemical pulping, the lignin is dissolved to separate the fibers. Lignin is a tan material in wood that turns dark brown during chemical pulping. For example, brown paper bags used in grocery stores contain about 10% lignin. Lignin impedes the fiber-to-fiber bonding; therefore, mechanical pulps form much weaker papers than chemical pulps. *See* CELLULOSE; HEMICELLULOSE; LIGNIN.

Extractives are a minor component of wood (1–8%) and comprise relatively small molecules that can be easily removed in the laboratory by using various liquids. Many wood extractives belong to the large class of materials called terpenes. Turpentine (which can be manufactured into pine oil) consists of volatile monoterpenes. This fraction is recovered as the wood chips are heated prior to chemical pulping. Rosin acids and fatty acids are recovered in the kraft recovery process. *See* PINE TERPENE; ROSIN; TALL OIL; TERPENE; WOOD CHEMICALS.

Pulping. The fibers of wood or other materials are separated by mechanical action, chemical action, or a chemical pretreatment followed by mechanical action; these processes are termed mechanical, chemical, and semichemical pulping, respectively. In fact, there are many variations in these processes. However, thermomechanical pulping and kraft chemical pulping account for over 80% of all pulp.

Mechanical pulping. Previously, cotton was the most important source of papermaking fiber because the fibers were already separated from each other, so they did not require pulping, and they have a high cellulose content. Wood bolts were ground against stones in the stone groundwood process, just as flour is made from wheat or corn. This pulp was very weak but was useful as an extender for cotton. As the process improved, the fiber was used in larger and larger amounts. The original grinding stone was sandstone, but it was replaced by synthetic stones embedded with aluminum oxide or silicone carbide.

Stone groundwood pulp has been supplanted by thermomechanical pulp, which has a much higher strength. In this process, wood chips are pulped in refiners with large metal disks that revolve in opposite directions. These disks have bars on them to allow the pulping process to occur. The process is carried out in two steps: a pressurized refiner operating at 110–130°C (230–265°F) softens the lignin and separates the fibers; an atmospheric pressure refining stage refines the pulp. The chips are introduced in the center of the disk, and centrifugal force pushes the pulp outward as it is formed. Energy requirements are 1900–2900 kW-h/ton. The use of a mild chemical pretreatment (using sodium sulfite or hydrogen peroxide under alkaline conditions) prior to thermomechanical pulping is called chemithermomechanical pulping. Nearly 20% of all pulp is made by the thermomechanical pulping process. Hydrogen peroxide is used to brighten mechanical pulp.

Chemical pulping. Kraft pulping (invented in 1879) is the most widely used chemical pulping process because kraft pulp is stronger than any other type. The process is also called the sulfate process because sodium sulfate can be used as a make-up chemical (to replace a small amount of chemical that leaves with the pulp), but sulfate is not involved during the actual pulping process. The white liquor used to cook the wood chips is a highly alkaline aqueous solution of sodium hydroxide and sodium sulfide. The wood is cooked at 160–180°C (320–356°F) in sealed digesters at a pressure of about 800 kilopascals for 0.5–2.0 h. Continuous digesters that may be several hundred feet tall, where wood chips continuously enter the top and pulp continuously exits the bottom, are used to produce most kraft pulp, but batch digesters and other designs are used too. About 65% of all pulp is produced by the kraft process. The pulp yield is only 45–55%.

Kraft pulping must be practiced with a chemical recovery process, which is much more complicated than the pulping process. The recovery process allows the sodium hydroxide and sodium sulfide to be reused, burns the materials dissolved from the wood so they are not a disposal problem, and

creates steam to drive the various processes within the mill. The spent liquor (called black liquor) is washed from the brown pulp (the brownstock). The black liquor is concentrated and burned in a boiler; water-filled tubes inside the boiler capture the heat and form steam. The inorganic chemicals are recovered as smelt (sodium carbonate and sodium sulfide) from the bottom of the boiler; the smelt is dissolved in water to form green liquor; the green liquor is treated with calcium hydroxide to convert the sodium carbonate of the green liquor to sodium hydroxide, forming the white liquor; and the calcium carbonate in washed and burned in a lime kiln to regenerate calcium oxide, which forms calcium hydroxide when it is dissolved in water. Tall oil is recovered from the concentrated black liquor by skimming before it is burned. *See* TALL OIL.

Kraft pulp mills have an odor somewhat like rotten eggs. A few parts per billion of mercaptans in the atmosphere cause this smell. Kraft mills have made much progress in recent years in the reduction of air emissions. Before 1930, the acid sulfite process was the main chemical pulping process. Since chemical recovery was not possible, it was modified in the 1950s as a magnesium or sodium process. However, the acid weakens cellulose and results in a pulp of only moderate strength. Therefore, only smaller, older mills use this method. Sulfite pulp can be used as a source of high-purity cellulose for conversion into cellophane, cellulose acetate, rayon, cellulose nitrate smokeless powder, and other materials. *See* CELLOPHANE; MERCAPTAN.

Bleaching. Chemical pulp for printing paper has 3–6% lignin, which gives the pulp a brown color. This lignin is removed with bleaching chemicals in four to eight stages. Each stage consists of a pump to mix the bleaching agent with the pulp, a retention tower to allow the chemical to react with the pulp for 30 min to several hours, and a washing unit to remove the solubilized lignin and residual chemicals from the pulp. Usually an oxidizing material is followed by a stage of alkali extraction, since lignin becomes more soluble at high pH. A common bleaching sequence involves elemental oxygen, alkali extraction, chlorine dioxide, alkali extraction, and, finally, chlorine dioxide.

Refining, stock preparation, and additives. Chemical pulp fibers are refined prior to papermaking to increase their flexibility and surface area, both of which promote fiber-to-fiber bonding. This is done with a double disk refiner, similar to those used in thermomechanical pulping, except the refining is done at ambient temperature with a finer bar pattern.

Dyes, retention aids, drainage aids, wet-strength adhesives (for paper towels that must hold their strength when wet), biocides, defoamers, and other materials may be added. Fillers are a major additive in printing papers and include clay, precipitated calcium carbonate, and titanium dioxide. These materials add to the brightness of the paper and improve paper opacity, which allows reading one side of the paper without having the print on the other side showing through. Titanium dioxide is especially useful, but its high expense limits it to thin (bible) papers.

Sizing, that is, the ability to resist rapid water absorption, is important to printing paper; otherwise ink spreads rapidly. Internal sizing has been carried out since the early nineteenth century by adding small amounts of rosin and aluminum sulfate to the dilute slurry of fibers at pH 4.5–5. This makes the paper sufficiently acidic to cause the cellulose to lose its strength over a period of decades. Most high-quality printing papers now use calcium carbonate as filler, which requires the slurry to be at neutral or slightly alkaline pH. Sizing agents that work at high pH are also used, and these papers do not deteriorate appreciably with time.

Manufacture. Paper must be made from a dilute slurry of fibers and additives. This slurry travels through screens and cleaners to remove debris prior to papermaking. There is 1 part fiber for every 200 parts water. This low consistency (solids content) is required to make a sheet with good formation, that is, a sheet of even density in all parts of the paper. The purpose of the paper machine is to remove all of this water.

Paper machines are as wide as 33 ft (10 m) and may operate at speeds over 60 mi/h (25 m/s). The dilute slurry of pulp is applied to a continuously moving plastic screen by the headbox. Water drains out by gravity, and then suction devices of increasing strength pull more water out. When the consistency is at 20%, the web is strong enough to be transferred to a felt material. The felt-supported web of fibers travels between three press nips formed where press rolls come together (as in old wringer washing machines) to press additional water from the web into the felt. The consistency at this point is 40–50%. Most of the remaining water is removed by drying the paper against a series of 20–60 steam heated dryer cans that are 6 ft (2 m) in diameter. Surface sizing with a starch or synthetic polymer solution may be accomplished between dryer sections.

The heaviest grades of papers, such as chipboard, are made on multiformer (cylinder) machines that form three to eight layers of fiber mats. These fiber mats are combined prior to pressing and drying. The lightest grades of paper, tissues, cannot withstand numerous felt transfers and are dried on very large Yankee dryers.

The paper may be smoothed against a series of rolls made from metal or rubbery material to impart smoothness or gloss to the paper. The paper may also be coated with a paintlike material to give it high brightness and gloss. In addition, numerous other converting operations may be performed on paper.

Christopher J. Biermann

Bibliography. C. J. Biermann, *Handbook of Pulping and Papermaking,* 1996; J. E. Kline, *Paper and Paperboard: Manufacturing and Converting Fundamentals*, 2d ed., 1991; E. Sjöström, *Wood Chemistry: Fundamentals and Applications*, 2d ed., 1993; D. C. Smith, *History of Papermaking in the United States: 1691–1969*, 1970.

Paprika

A type of pepper, *Capsicum annuum*, with nonpungent flesh, grown for its long red fruit. A member of the plant order Polemoniales and of American origin, it is most popular in Hungary and adjacent countries. Seeds are removed from the mature fruit, and the flesh is dried and ground to prepare the dry condiment commonly referred to as paprika. Production in the United States is limited, with California the only important producing state. *See* PEPPER; SOLANALES.

H. John Carew

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

13 PAR-PLAN

www.MHEST.com

P

Parablastoidea

A small extinct class of relatively advanced blastozoan echinoderms containing five genera (three named), ranging from the middle Early Ordovician to the early Late Ordovician in eastern Canada; northeastern, eastern, south-central, and western United States; south and north Wales; and near St. Petersburg, western Russia. Parablastoids have a bud-shaped theca or body with well-developed pentameral symmetry. Thecal plates include basals, small or large radials, and sometimes other small plates in the lower theca; large distinctive triangular-to-parabolic deltoids between the ambulacra in the upper theca (see **illustration**); and small oral plates and an oral crest surrounding and covering the mouth on the summit. Parablastoids have single or multiple slits through the lower deltoids that are connected by internal folds (cataspires) to pores that open between the single ambulacral plates at the edge of each ambulacrum; presumably they were respiratory organs. Short-to-long biserial brachioles were attached to the edges of the five ambulacra and served as the main food-gathering structures; when not feeding, the brachioles folded in to lie against a high T-shaped ambulacral crest in the center of each ambulacrum in the best-known genus. A stem with one-piece columnals attached the theca to the sea floor, suggesting that parablastoids were attached, medium- to high-level, suspension feeders. Some parablastoids are found in bank or reef deposits, suggesting they may have been adapted for rough-water conditions. Although they converged on blastoids in thecal design and way of life, parablastoids had differences in their plating, ambulacra, and respiratory structures that indicate a separate origin and evolutionary history. This justifies assigning parablastoids and blastoids to different blastozoan classes. *See* BLASTOIDEA; ECHINODERMATA; ORDOVICIAN. James Sprinkle

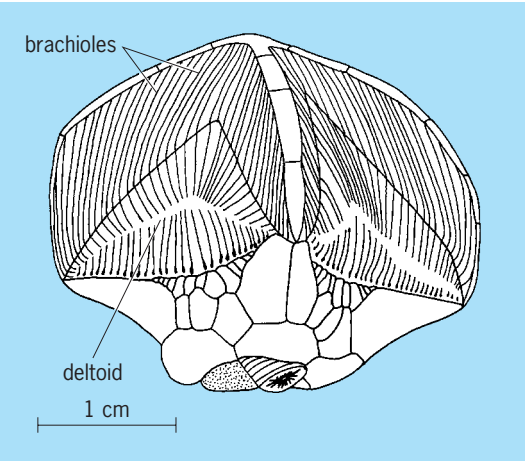

Side view of *Blastoidocrinus carchariaedens* from the Middle Ordovician of New York, showing the thecal plating, large triangular deltoids, and numerous short brachioles attached to the ambulacra. (*After R. O. Fay, in R. C. Moore, ed., Treatise on Invertebrate Paleontology, Pt. S, pp. S293–S296, Geological Society of America and University of Kansas Press, 1968*)

Bibliography. R. O. Fay, Parablastoids, pp. S293–S296 in R. C. Moore (ed.), *Treatise on Invertebrate Paleontology, Part S: Echinodermata 1(1)*, Geological Society of America, Boulder, and University of Kansas, Lawrence, 1968; C. R. C. Paul, The phylogeny of the cystoids, pp. 199–213 in C. R. C. Paul and A. B. Smith (eds.), *Echinoderm Phylogeny and Evolutionary Biology*, Clarendon Press, Oxford, 1988; C. R. C. Paul and J. C. W. Cope, A parablastoid from the Arenig of South Wales, *Palaeontology*, 25(3):499–507, 1982; J. Sprinkle, *Morphology and Evolution of Blastozoan Echinoderms*, Museum of Comparative Zoology, Harvard University, Spec. Publ., 1973.

Parabola

A member of the class of curves that are intersections of a plane with a cone of revolution. It is obtained (see **illus.**) when the cutting plane is parallel

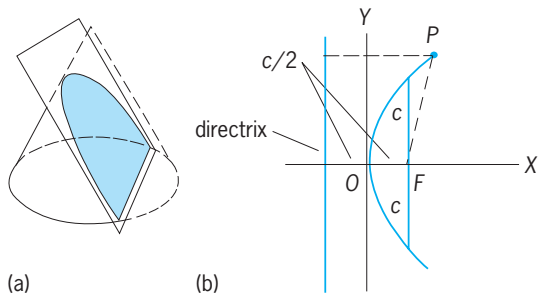

Parabola as (*a*) conic section and (*b*) locus of points.

to an element of the cone. *See* CONIC SECTION.

In analytic geometry the parabola is defined as the locus of points (in a plane) equally distant from a fixed point F (focus) and a fixed line (directrix) not through the point. It is symmetric about the line through F perpendicular to the directrix. To construct a parabola, pin one end of a piece of string to a point F and fasten the other end to one end of a ruler whose length equals that of the string. If the other end of the ruler slides along a line, and the string is kept taut by a pencil, the point of the pencil will trace an arc of a parabola. Hippocrates of Chios (about 430 B.C.) showed that one of the three famous problems of antiquity, duplication of the cube, can be solved by use of parabolas. The problem is to construct the edge of a cube whose volume is twice that of a given cube. If c denotes the edge of the given cube, then the desired edge is obtained by considering the two parabolas whose equations in rectangular cartesian coordinates are $x^2 = cy$, $y^2 = 2cx$. They intersect at the origin and a point $P(x_0,y_0)$, with $x_0{}^3 = 2c^3$.

All parabolas are similar; they differ only in scale. For a discussion of the optical property of parabolas *see* ANALYTIC GEOMETRY.

The curve has numerous other properties of interest in both pure and applied mathematics. (For example, the trajectory of an artillery shell, assumed to be acted upon only by the force of gravity, is a parabola; and the circle that circumscribes the triangle formed by any three tangents of a parabola goes through the focus.) Archimedes found the area bounded by an arc of a parabola and its chord; for example, the area bounded by the parabola $y^2 = 2cx$ and its latus rectum (the chord through F perpendicular to the axis) is $^2/_3c^2$. The length of the parabolic arc that is cut off by the latus rectum is $[\sqrt{2} + \ln(\sqrt{2} + 1)]c$. The volume obtained by revolving this about the axis of the parabola is $\pi c^3/4$, and the surface is $^2/_3(2\sqrt{2} - 1)\pi c^2$. Leonard M. Blumenthal

Parachute

A flexible, lightweight structure, generally intended to retard the passage of an object through the atmosphere by materially increasing the resistive surface. Parachutes have continued to be the simplest and cheapest devices for the deceleration of payloads, people, and vehicles since their first recorded use in 1797. Comprising cloth and suspension lines, their construction is far simpler than that of aircraft. As a result, parachute construction and design is a fairly mature art. However, their very "softness" makes their aerodynamics much more complicated, so complicated in fact that modern supercomputers are not powerful enough to completely simulate both the evolution of the parachute shape and the internal and external airflows, either during inflation or during rapid flight maneuvers. Indeed, unlike aircraft, which are solid structures that deflect the air around them, parachutes not only deflect the surrounding air but also adopt shapes that are dictated by it. During inflation, such feedback is even more dominant as both shape and flows are basically unsteady (that is, do not remain constant in time). Finally, given the lack of streamlining both

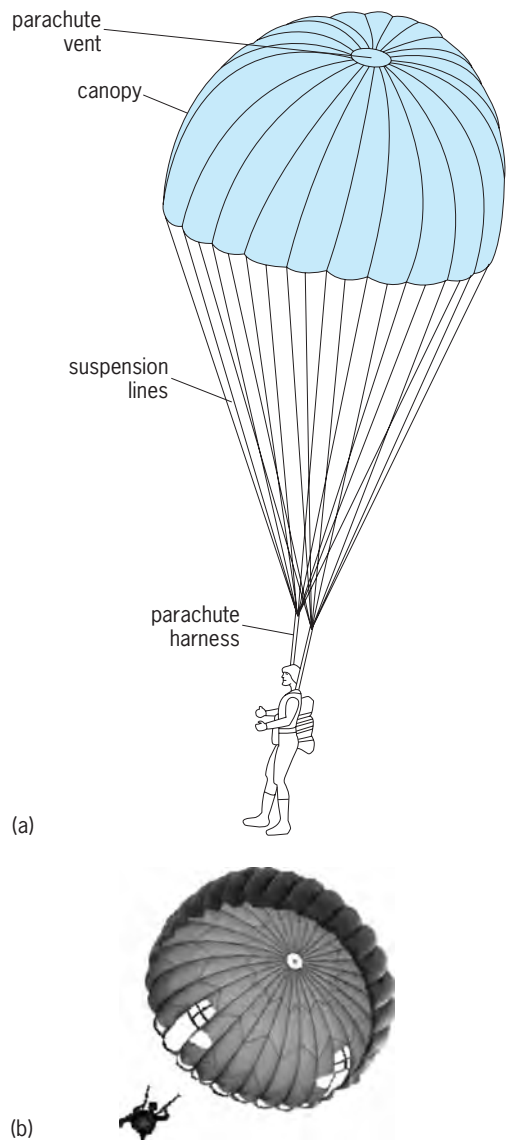

Fig. 1. Hemispherical parachutes. (*a*) T-10 parachute, used for airborne troop insertion. (*b*) Parachute with gaps to enhance stability.

during and after inflation, turbulent flow rather than laminar flow controls the aerodynamics. *See* AERODYNAMICS; LAMINAR FLOW; SUPERCOMPUTER; TURBULENT FLOW.

Types. Parachutes come in two basic shapes: a near-hemispherical cup (**Fig. 1**) or a wing (**Fig. 2**). Most are made of nylon fabric. When fully inflated, some hemispherical parachutes have a slight conical shape, while in others slots, vents, or concentric gaps are cut to enhance stability (Fig. 1*b* and **Fig. 3**). Winglike parachutes, also commonly know as parafoils or ram-air parachutes, are made of bags or cells sewn together to form a wing. They maintain this shape via the wind entering inlets cut on the leading edge of the parachute, thereby pressurizing each cell.

Figure 3 shows several versions of so-called slotted canopies which, by virtue of having concentric rings of small slots, are considered to be the most stable hemispherical parachutes. By stability is meant a reduced propensity to oscillate, that is, to swing the payload sideways, in a direction that is perpendicular to the direction of fall. During their steady descent through the air, stability or lack thereof is caused by the fact that all parachutes accumulate air and build up internal pressure. Parachutes without slots or gaps (such as the parachute in Fig. 1*a*) allow the accumulating internal air to escape by periodically spilling it to one side by the parachute's mouth. Like thrust coming out of a jet engine, such spilling induces motions in directions opposite to the spill, namely sideways. Slotted canopies achieve stability by allowing this accumulating air to flow through a large number of gaps distributed symmetrically over the entire surface of the cloth. But the drawback is that introducing slots and vents yields a lower deceleration capacity—that is, an overall loss of drag-creation performance. Thus the parachute designer must find a compromise between two desired qualities—stability and drag generation—which unfortunately work in opposition. *See* AERODYNAMIC FORCE.

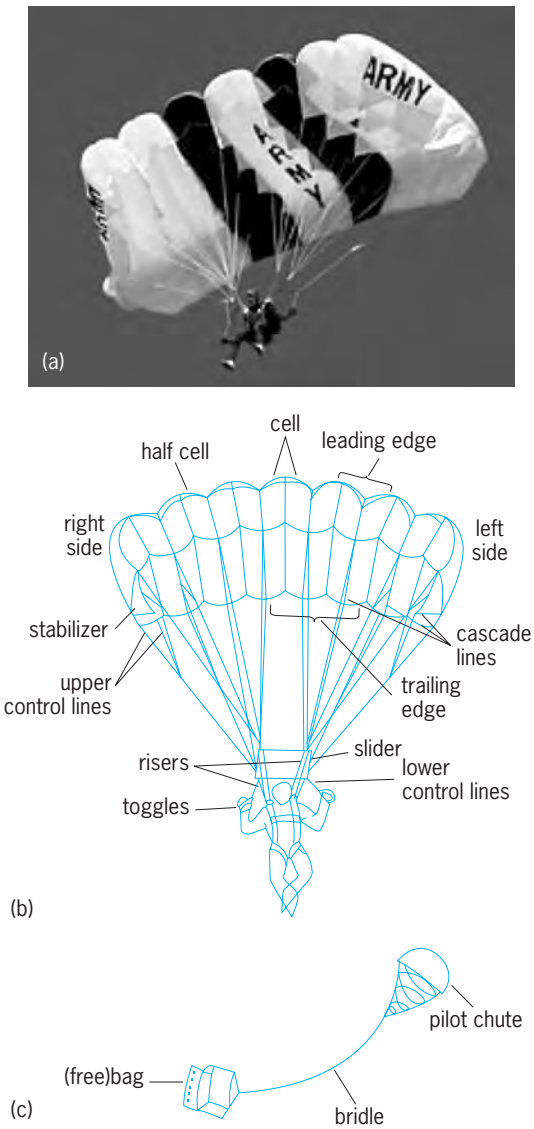

Fig. 2. Wing-shaped parachutes (parafoils). (*a*) Deployed parafoil descending. (*b*) Components of deployed parafoil. (*c*) Deployment mechanism.

The ribbon parachute (Fig. 3*a*) is constructed by using many ribbons, usually 2-in-wide (5-cm) nylon, arranged in concentric circles from the center (vent) of the parachute to the skirt. Between the ribbons are the slots. Similarly, a ring-slot parachute (Fig. 3*b*) is constructed from wider rings of cloth with slots between the rings. For a given size, a ribbon parachute would have perhaps 40 ribbons and slots, and a ring-slot parachute would have 10 rings and slots. A ring-sail parachute (Fig. 3*c*) is similar to a ring-slot parachute, but with the lower rings having added fullness (that is, extra cloth). Such fullness allows the lower edge of each ring of cloth to bulge out and vent air downward, thereby giving the parachute a greater drag-producing capability than a ring-slot parachute.

Deployment and inflation stages. All parachutes are packed into a small container prior to actual use, and therefore require hardware that will extract the parachute out of its container and deploy it into the wind. The initial deployment stages of hemispherical parachutes and parafoils are similar. Both begin with the extraction of the bag containing the folded parachute from the harness-container assembly (or vehicle). The extraction is normally carried out with either a lanyard attached to the aircraft (that is, the so-called static line), or with a "pilot" or "extraction" chute, usually a small hemispherical parachute that is previously deployed into the airstream by a latch and ejection spring mechanism (Fig. 2*c*). As the bag separates from the harness or vehicle, the suspension lines unfold first. Only when these lines are completely stretched is the parachute allowed to unfurl out of the bag and begin inflating. Inflation is typically characterized by several stages whose duration depends on the specific design of the parachute (or parafoil) and on whether the parachute is reefed.

Reefing systems. If not controlled, the opening forces that accompany the inflation process may

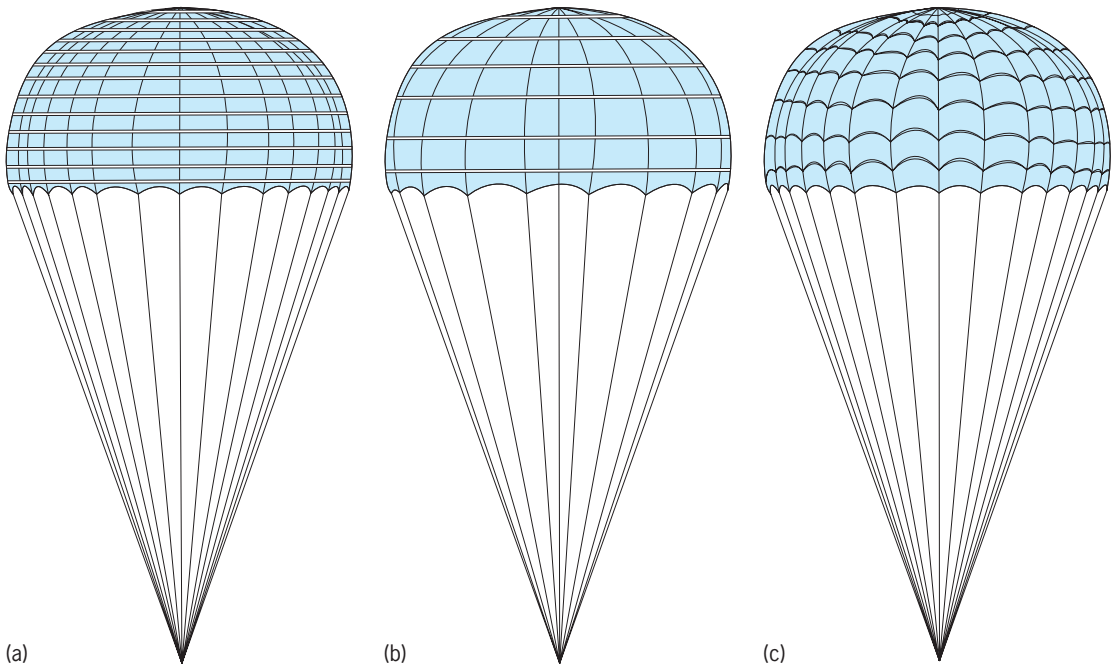

Fig. 3. Slotted canopy parachutes. (*a*) Ribbon parachute. (*b*) Ring-slot parachute. (*c*) Ring-sail parachute. (*After Recovery Systems Design Guide, Irvin Industries Inc., Air Force Flight Dynamics Laboratory, Tech. Rep. AFFDL-TR-78-151, 1978*)

be large enough to destroy the parachute or damage its payload. Typically, the inflation forces are large when a parachute inflates too quickly. In this case the parachute and its payload still travel at high speeds by the time the parachute has spread open, thereby generating a very large amount of drag. Reefing devices are designed to limit the rate of canopy expansion during the early phase of inflation, allowing enough time for the parachute-payload system to reduce its speed with the help of the small but sufficient drag provided by the partially opened parachute. With some hemispherical parachutes, for example, a line routed around the skirt limits the canopy expansion by constraining the parachute's mouth (or skirt) to a small-diameter circle. Usually, a mechanical or pyrotechnic device cuts this line at a preset time or altitude to allow full canopy inflation, albeit at a lower speed. On parafoils, a slider is used to achieve the same result (Fig. 2*b*). Here, a square of nylon fabric is allowed to slide freely and slowly down the suspension lines at a rate that is controlled by the slider's own drag and by the tension of the lines that fan out of the slider.

Pressure dynamics during inflation. The inflation stages of a hemispherical parachute without reefing are shown in **Fig. 4**. After unfurling out of its container, the parachute adopts a rather elongated shape, resembling a vertical tube opened at its lower end (Fig. 4*a*). Because of the system's rapid descent, air rushes in through the tube's opening and accumulates at the apex of the canopy to create a high-pressure "bubble." Continuous inflow builds up internal pressure, allowing the bubble's volume to expand horizontally as well as vertically (Fig. 4*b*). Early in the inflation process, the expansion is more along the vertical; later it is mostly along the horizontal. This expansion continues until the bubble is large enough to occupy the entire "design" volume of the parachute, the size of which is determined by the system's descent rate, by the presence of vents or gaps, and by the balance of pressure outside and inside the parachute (Fig. 4*c*).

The shape of the expanding air bubble trapped in the canopy is dictated by the balance of the aerodynamic forces that act in opposite directions along the boundary defined by the parachute's fabric. The bubble's expansion rate first depends on the pressure differential between the outside and inside of the parachute. As for any object moving through air, wake turbulence generated on the downwind side causes the external pressure to be lower near the apex than the pressure inside the parachute. The faster the parachute, the larger this pressure differential, and the faster the bubble's expansion. However, such rapid expansion generates a large external pressure which squeezes the bubble on its upwind side, thereby slowing down its expansion. This external pressure arises because the bubble deflects outside air sideways, thereby increasing the resistance of the air to the deflection. This balancing act between the bubble's expansion and the resulting external air deflection results in the bubble expanding more along the vertical than the horizontal at first. Only when internal pressure is high enough to overcome the effects of air deflection will the bubble expand along the horizontal. Finally, because the parachute-payload system decelerates during inflation, this pressure and force balance is continuously readjusted as the inflow becomes slower. *See* WAKE FLOW.

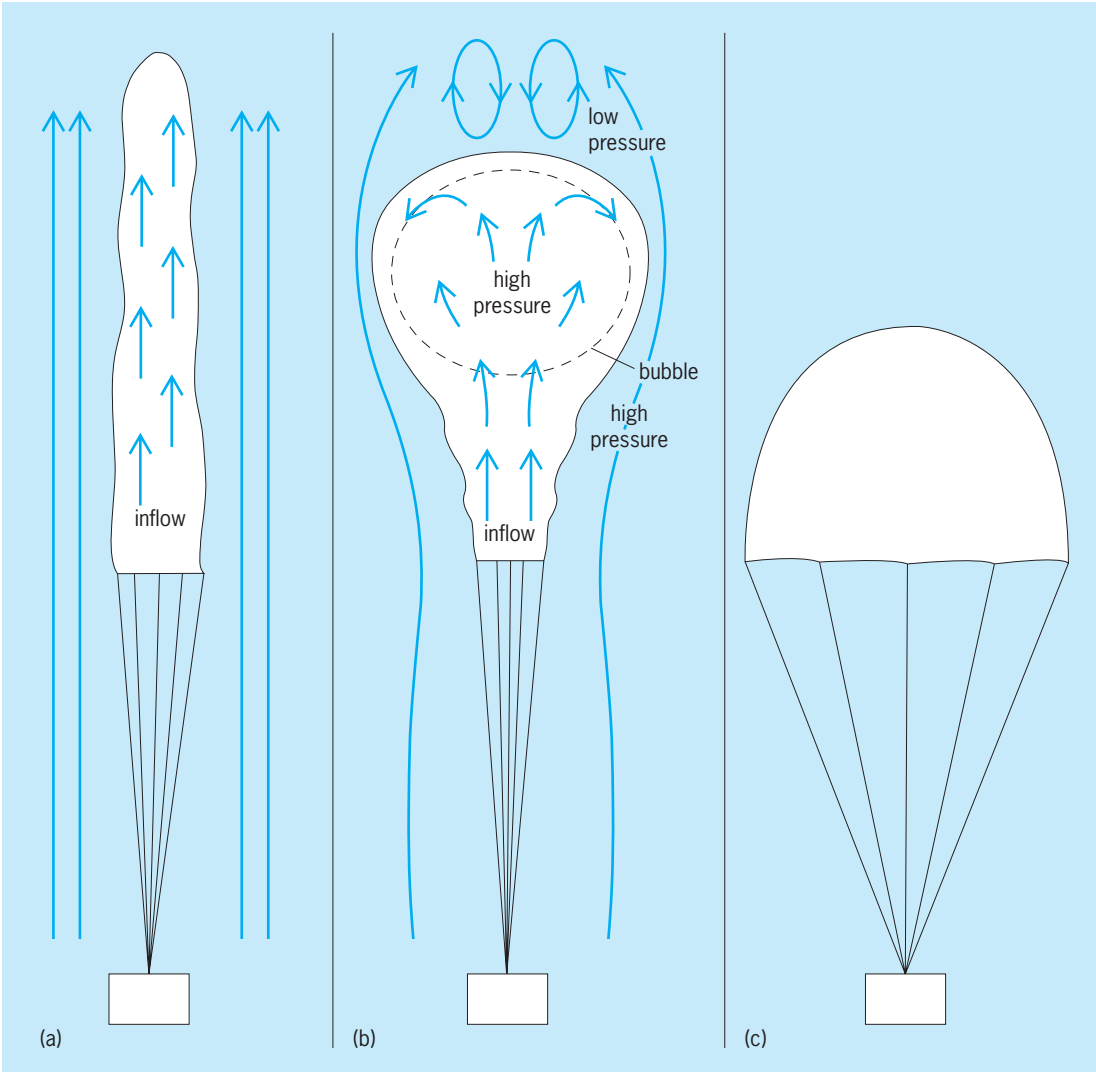

Fig. 4. External and internal airflows and pressures acting on an inflating circular parchute (*a*) after bag unfurling, (*b*) at midpoint, and (*c*) at a late stage.

The inflation of parafoils evolves similarly, at least during the first fraction of a second of the inflation process. Here, inflation takes place in the cells of the parachute, which later define the wing shape of the canopy. Initially, these cells are shaped like collapsed horizontal tubes with the inlets on the wing leading edge scooping in the air coming from below. The walls between the cells have holes that allow the internal pressure to spread more evenly over the entire span of the wing. The inflation process is regulated by the difference of pressure between the outside and inside of the cells in a manner similar to the processes sketched in Fig. 4 except that the air is injected from the side through the inlets.

Technology development. New applications for parachutes constantly require advances in design and packaging. Many current systems have been in use (with minor updates) since the 1950s. Refinements or outright new designs provide more stability while providing the most drag, doing so in harsher environments and at less cost, weight, and pack volume. For example, pack volume reductions have been achieved in an increasing number of military parachute systems by using machinery to pack the parachute under pressure and in a near vacuum.

Computer simulations. The availability of supercomputers has led to the development of computer models that track the interactions of the airflow with the parachute cloth and lines. These models are still approximate in the way that they reproduce fabric crumpling and unsteady turbulent flows. But they are good enough to yield inflation sequences that are very close to reality (**Fig. 5**). In a typical simulation, these models first map the internal and external airflows and pressures on grids that are adapted to the various shapes of the flow near and far from the parachute (**Fig. 6**). Then the fabric and suspension lines are allowed to move according to the pressures acting on them. With this new shape, the airflow is recalculated at the new positions of the parachute parts, which are then moved again according to the

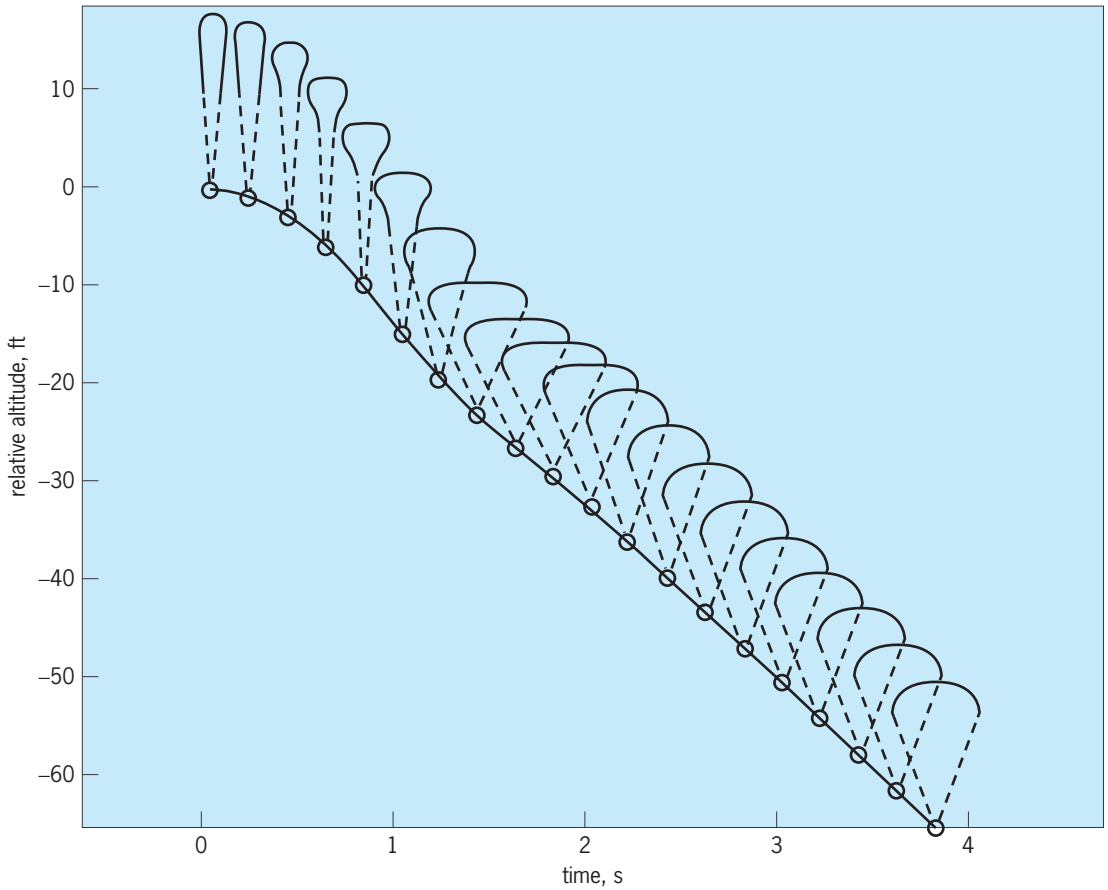

Fig. 5. Computer simulation of an inflating circular parachute released from rest. This process is superimposed on the altitude loss incurred during the process. 1 ft = 0.3 m. (*After R. J. Benney and K. R. Stein, U.S. Army Natick Soldier Center*)

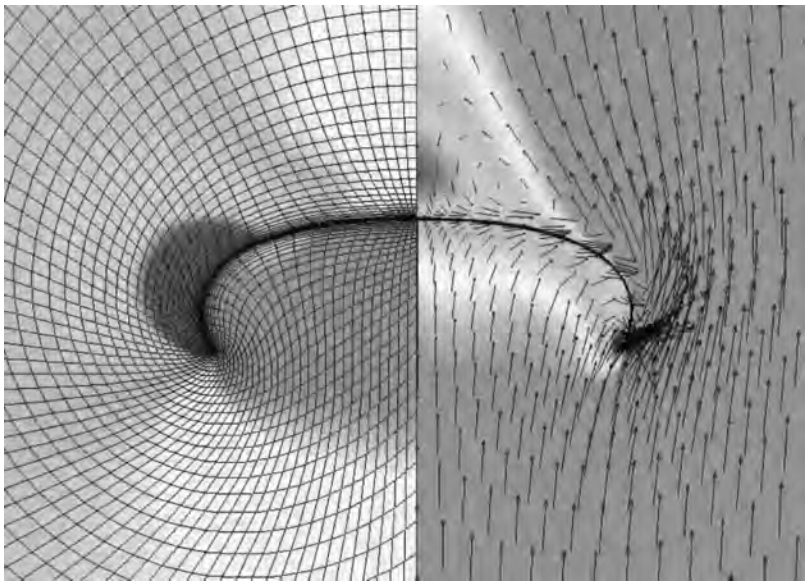

(a) (b)

Fig. 6. Computer simulation of half-scale C-9 hemispherical parachute at 1.6 s into the inflation process. (*a*) Pressure distribution. (*b*) Velocity vectors. (*Simulation by K. R. Stein and R. J. Benney, U.S. Army Natick Soldier Center. Reprinted with permission.*)

new pressures, and so on. Each simulation step is small in time to maintain the accuracy of each calculation. The end results are plots like the one in Fig. 6, which contains the speeds of the airflows and the amount of pressure at various locations. Such maps can be displayed at different times to show the evolution of the inflation process (Fig. 5). *See* COMPUTATIONAL FLUID DYNAMICS. Jean Potvin

Bibliography. T. W. Knacke, *Parachute Recovery Systems Design Manual*, Para Publishing, Santa Barbara, CA, 1992; C. W. Peterson, High Performance Parachutes, *Sci. Amer.*, 262(5):108–115, May 1990; C. W. Peterson, H. Higuchi, and J. H. Strickland, The fluid dynamics of parachute inflation, *Annu. Rev. Fluid Mech.*, 28:361–387, 1996; D. Poynter, *The Parachute Manual*, vols. 1 and 2, 4th ed., Para Publishing, Santa Barbara, CA, 1991; J. H. Strickland and H. Higuchi, Parachute aerodynamics: An assessment of prediction capability, *J. Aircraft*, 33:241–252, 1996.

Paracrinoidea

A medium-sized class of extinct blastozoan echinoderms containing about 17–19 genera, ranging from the late Early Ordovician to the Early

Silurian throughout North America and in northern Europe. Their maximum standing diversity (13 genera) occurred in the early Late Ordovician, when paracrinoids were especially common in eastern and central North America.

Paracrinoids have a lens-shaped to globular body or theca, a short often curved stem, and 2–5 ambulacra or erect arms bearing smaller uniserial appendages with a single row of ossicles. Only genera with 5 ambulacra or arms show any pentameral (five-sided) symmetry. Many other paracrinoids have either two long ambulacra extending down the edges of a lens-shaped theca or very short S-shaped ambulacra on the summit of a globular theca. Three features distinguish paracrinoids from most other blastozoan echinoderms: (1) they have asymmetrical, uniserial, recumbent ambulacra or erect arms typically with smaller, erect, uniserial appendages branching off the right side looking adorally (toward the mouth) and a covered food groove extending up the left side to the mouth (see **illustration**); (2) in many paracrinoids, the summit has shifted so that the anal opening (and not the mouth) lies on the thecal axis opposite the stem facet; and (3) three basal plates, two somewhat larger and one smaller (usually on the left anterior side), form the stem facet at the thecal base.

The theca typically has between 18 and 200 mostly irregularly arranged polygonal plates, either large thin plates without pores, or thicker domed or concave plates with covered pores or slits on the sutures. The erect appendages are usually uniserial "brachioles" that are 15–25 ossicles long, with a small food groove protected by biserial cover plates on one side. The short- to medium-length stem is made up of one-piece round columnals and is often curved just below the theca, and several genera have a one-piece, cemented, distal holdfast attached to a hard substrate or to a skeletal fragment lying on the sea floor. Paracrinoids were low- to medium-level suspension feeders either held some distance up in the water column by the curved stem or sitting recumbent on the sea floor tethered in place with a short buried stem (see illus.).

A few paracrinoid genera show external growth lines on the larger thecal plates, indicating they were blastozoan and not crinozoan echinoderms and must have modified the usual biserial brachioles. They most closely resemble a small group of specialized stemmed blastozoans called rhipidocystids (five genera) that have a flattened theca with a reduced number of elongate marginal and thin central plates. Rhipidocystids are most likely the sister group to paracrinoids, and they should probably be classified

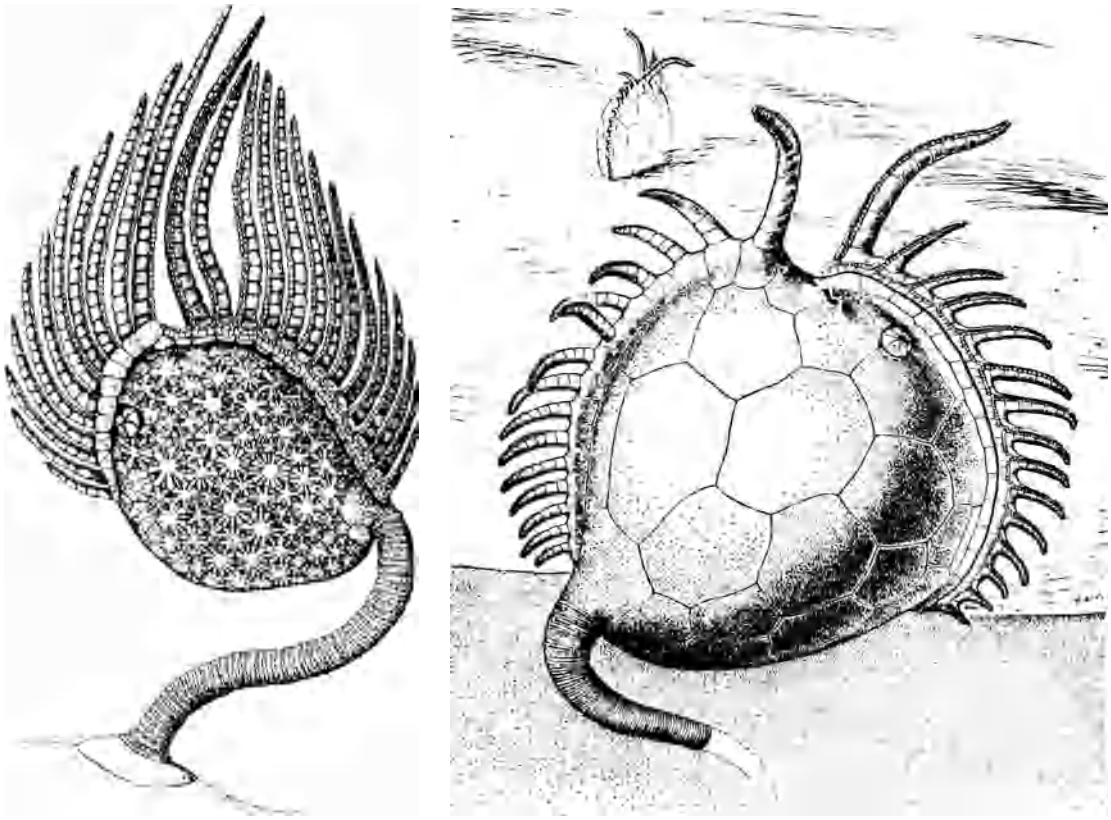

(a) (b)

Reconstructions of two paracrinoids showing different proposed living positions. (*a*) *Amygdalocystites* from the Late Ordovician of Ontario, Canada, held up some distance in the water column by a curved attached stem (*after T. E. Guensburg, J. Paleontol., 65:695, 1991*). (*b*) *Platycystites* from the Late Ordovician of Oklahoma, sitting upright but recumbent on the sea floor with a buried trailing stem (*after R. L. Parsley and L. W. Mintz, Bull. Amer. Paleontol., 68(288):70, 1975*).

together in that class. *See* BLASTOIDEA; CRINOZOA; ECHINODERMATA; EOCRINOIDEA; ORDOVICIAN.

James Sprinkle

Bibliography. T. E. Guensburg, The stem and holdfast of *Amygdalocystites florealis* Billings, 1854 (Paracrinoidea): Lifestyle implications, *J. Paleontol.*, 65(4):693–695, 1991; R. V. Kesling, Paracrinoids, pp. S268–S288 in R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, Part S: *Echinodermata 1(1)*, Geological Society of America and University of Kansas, 1968; R. L. Parsley, Paracrinoids, pp. 212–223 in J. Sprinkle (ed.), *Echinoderm Faunas from the Bromide Formation (Middle Ordovician) of Oklahoma*, Univ. Kans. Paleontol. Contrib. Monogr. 1, 1982; R. L. Parsley and L. W. Mintz, North American Paracrinoidea (Ordovician: Paracrinozoa, new, Echinodermata), *Bull. Amer. Paleontol.*, 68(288):1–115, 1975; S. V. Rozhnov, Comparative morphology of *Rhipidocystis* Jaekel, 1900 and *Cryptocrinites* von Buch, 1840 (Eocrinoidea; Ordovician), pp. 173–178 in B. David et al. (eds.), *Echinoderms Through Time*, A. A. Balkema, Rotterdam, 1994; S. V. Rozhnov, Morphogenesis and evolution of crinoids and other pelmatozoan echinoderms in the Early Paleozoic, *Paleontol. J.*, 36(Suppl. Issue 6): S525–S674 [English translation of Russian monograph in *Paleontologicheskii Zhurnal*], 2002; J. Sprinkle and R. L. Parsley, "Golf-ball" paracrinoid, pp. 224–230 in J. Sprinkle (ed.), *Echinoderm Faunas from the Bromide Formation (Middle Ordovician) of Oklahoma*, Univ. Kans. Paleontol. Contrib. Monogr. 1, 1982.

Paraffin

A term used variously to describe either a waxlike substance or a group of compounds. The former use pertains to the high-boiling residue obtained from certain petroleum crudes. It is recovered by freezing out on a cold drum and is purified by crystallization from methyl ethyl ketone. Paraffin wax is a mixture of 26- to 30-carbon alkane hydrocarbons; it melts at 52–57°C (126–135°F). Microcrystalline wax contains compounds of higher molecular weight and has a melting point as high as 90°C (194°F). The name paraffin is also used to designate a group of hydrocarbons—open-chain compounds of carbon and hydrogen with only single bonds, of the formula C_nH_{2n+2}, where n is any integer. This usage is obsolete. *See* ALKANE; WAX, PETROLEUM.

Allen L. Hanson

Parainfluenza virus

A member of the genus *Paramyxovirus* of the family Paramyxoviridae which is associated with a variety of respiratory illnesses. The virus particles range in size from 90 to 200 nanometers, agglutinate red blood cells, and (like the influenza viruses) contain a receptor-destroying enzyme. They differ from the influenza viruses in their large size, their possession of the larger ribonucleoprotein helix characteristic of the paramyxoviruses, their tendency to lyse as well as agglutinate erythrocytes, and their generally poor growth in eggs. In culture these viruses grow well in primary monkey or human epithelial cells; although they produce little cytopathic effect in the culture cells, they are easily recognized by the hemadsorption method. They are relatively unstable upon storage, even at freezing temperatures. Laboratory diagnosis may be made by hemagglutination-inhibition, complement fixation, and neutralization tests of the patient's blood. *See* COMPLEMENT-FIXATION TEST; EMBRYONATED EGG CULTURE; PARAMYXOVIRUS; TISSUE CULTURE.

Four subgroups are known, designated parainfluenza 1, 2, 3, and 4. Types 1, 2, and 3 are distributed throughout the world, but thus far type 4 has been found only in the United States. Parainfluenza 1 and 3 are ubiquitous endemic agents producing infections all through the year. Types 2 and 4 occur more sporadically. With all of the parainfluenza viruses, most primary infections take place early in life. About half of the first infections with parainfluenza 1, about two-thirds of those with parainfluenza 2, and three-fourths of those with parainfluenza 3 produce febrile illnesses. The target organ of type 3 is the lower respiratory tract, with first infections frequently resulting in bronchial pneumonia, bronchiolitis, or bronchitis. Type 1 is the chief cause of croup, but the other types have also been incriminated, to the extent that one-half of all cases of croup can be shown to be caused by parainfluenza viruses. Antibody surveys of normal populations in the United States indicate that type 3 infection generally occurs very early, and most infants have type 3 neutralizing antibody before they are 2 years old. By the time they enter school, a large proportion of children have acquired antibodies to types 1 and 2 also. Almost all adults possess parainfluenza antibodies, but reinfection is common, particularly with type 3. Although preexisting antibody does not prevent reinfection, it appears to reduce the severity of the illness.

Type 1. This group includes the Sendai virus, also known as the hemagglutinating virus of Japan (HVJ) or influenza D, and the hemadsorption virus type 2 (HA-2). The widespread HA-2 virus appears to be one of the chief agents producing croup in children. In adults it produces symptoms like those of the common cold. Immunization with experimental inactivated parainfluenza type 1 vaccine stimulates the appearance of antibody in the blood but not in nasal secretions. Recipients of such a vaccine remain susceptible to reinfection. In contrast, infection stimulates antibody appearance in the nasal secretions and concomitant resistance to reinfection; however, reinfections may occur with time.

Type 2. This group includes the croup-associated (CA) virus, or acute laryngotracheobronchitis virus of children. Simian viruses 5 and 41 appear to be antigenically related to parainfluenza 2.

Type 3. This virus is also known as hemadsorption virus type 1 (HA-1). Most infants have antibody to type 3 by the time they are 2 years old—usually before they have experienced infection with the other parainfluenza types. Upon subsequent infection with types 1 or 2, a "booster" response is seen in type 3 antibodies, indicating a sharing of antigens.

Type 4. The prototype of this virus is the M-25 strain. It has been associated only with malaise in young children or mild illnesses limited to the upper respiratory tract. It has failed to produce any cytopathic effects in cultures and can be recognized only by the hemadsorption technique. *See* ANIMAL VIRUS; INFLUENZA; VIRUS CLASSIFICATION.

Joseph L. Melnick; M. E. Reichmann

Bibliography. R. M. Chanock, M. A. Mufson, and K. M. Johnson (eds.), Comparative biology and ecology of human virus and mycoplasma respiratory pathogens, *Progr. Med. Virol.*, 7:208–272, 1965.

Parainsecta

A class of hexapod (six-legged) arthropods consisting of the orders Collembola and Protura. Class Parainsecta (also known as Ellipura) is one of several classifications of those hexapods, which, while similar to insects, are sufficiently different that most researchers believe they deserve separate class status.

Collembola and Protura have frequently been lumped with the Diplura (Entotrophi) into one class (Entognatha) on the basis of their mouthparts, which are inside the head capsule rather than exposed as in the Insecta. However, some recent morphological and paleontological studies support a closer link between insects and Diplura. In addition, there are a number of features found in Protura and Collembola that are not found in Diplura. These include the developmental process: Both Collembola and Protura have epimorphic development (that is, development of the young is completed in the egg). In the Parainsecta nervous system the abdominal ganglia are reduced to one ganglion or fused to the last thoracic ganglion, whereas Diplura and primitive insects have at least seven abdominal ganglia. The internal cephalic supportive structures of Parainsecta are different from those seen in Diplura. The postantennal organ present in many Collembola may be homologous with the similar organ of Tömösváry of Protura; no such organ occurs in Diplura or insects. The clypolabium and labium (upper and lower lips) are immobile, unlike those of other adult hexapods. Two unique features of the Parainsecta are the presence of a ventral midline or groove and coxal vesicles (fluid-filled sacs, generally associated with leg bases) that are limited to the first abdominal segment. In the Collembola these vesicles are fused into a single ventral tube. The digestive tract lacks anterior or posterior enlargements, which are often present in Diplura and primitive insects. Other distinguishing features of the Parainsecta are the absence of styli and unpaired pretarsal claws.

Both Collembola and Protura contain very small animals, generally less than 4 mm long. Protura are exclusively, and Collembola primarily, soil and litter inhabitants. Genetic evidence concerning relationships is at present confusing, with some studies showing a closer link between Diplura and Collembola than between either and Protura, and others showing a closer link between Collembola and Protura. In addition, differences of interpretation concerning the phylogenetic importance of various morphological features leave the status of the Parainsecta ambiguous. *See* ARTHROPODA; COLLEMBOLA; DIPLURA; INSECT PHYSIOLOGY; INSECTA; PROTURA.

Kenneth Christiansen

Parallax (astronomy)

The apparent angular displacement of a celestial object due to a change in the position of the observer. With a baseline of known length between two observations, the distance to the object can be determined directly.

Geocentric parallax. The rotation of the Earth or the linear separation of two points on its surface can be used to establish distances within the solar system. The parallax determined is scaled to the equatorial radius of the Earth, which is equal to 6378.1366 km (3963.1903 mi). At the mean distance of the Moon, this baseline subtends an angle of 57′03″, and at the mean distance to the Sun it amounts to 8.794143″. This latter distance is defined as the astronomical unit and serves as a measure of distances within the solar system. One astronomical unit is 149,597,870.691 ± 0.006 km (92,955,807.267 ± 0.004 mi) in length, and its high precision results from tracking interplanetary space probes. *See* ASTRONOMICAL UNIT; EARTH ROTATION AND ORBITAL MOTION.

Heliocentric or trigonometric parallax. The astronomical unit is the baseline for the measure of stellar parallaxes or distances and ultimately every other distance in the universe outside the solar system. Observations made from the Earth in its orbit on opposite sides of the Sun are scaled to the astronomical unit, although as much as twice that distance may be available. The stellar parallax is given in units of arc-seconds and is by definition the reciprocal of the distance in parsecs. One parsec is the distance at which one astronomical unit subtends an angle of one second of arc, and equals 206,264.8 astronomical units or 3.2616 light-years, one of which is 63,240 astronomical units. *See* LIGHT-YEAR; PARSEC.

Attempts to measure stellar parallax were unsuccessful before 1838, when the parallaxes of three different stars were measured independently. One was the triple star Alpha and Proxima Centauri, whose parallax of about 0.76″ is still the largest known. Stellar parallaxes accumulated slowly until about 1900, when long-focus refracting telescopes had become available and the photographic

process was introduced into astrometry. *See* ALPHA CENTAURI; ASTRONOMICAL PHOTOGRAPHY.

The fourth edition of the *Yale Parallax Catalogue*, published in 1995, contains 15,994 parallaxes of 8112 stars. Observations included in this volume were almost entirely obtained by using long-focus refractors with apertures between 0.5 and 1.0 m (20 and 40 in.). However, in 1962 the U.S. Naval Observatory's 1.5-m reflector was put into operation and began to produce high-precision parallaxes for fainter stars. In addition, measurements were increasingly made with automatic measuring engines and their reductions with computers. The practice of measuring only along the direction of right ascension, since most parallactic motion occurs in that direction, was abandoned in favor of utilizing measurements in both directions. Further improvements included increasing the number of comparison stars constituting the reference frame, which allowed more rigorous solutions to be made. Precision was also improved through a more careful balance between observations made before and after midnight, when the Earth is on alternate sides of its orbit.

The external standard error of a single parallax determination made before 1962 averaged about $\pm 0.016''$. By making use of the improved techniques mentioned above, this error has been reduced by about half. The automatic measuring machines on which most work has been done since that date, along with the fine-grain emulsions which were developed a few years later, cut the error in half again. Charge-coupled devices (CCD) replaced the photographic plate in the 1990s and reduced the error still further. At the present time, the only major ground-based parallax programs are located at the U.S. Naval Observatory at Flagstaff, Arizona, and at the Cerro Tololo Interamerican Observatory in Chile. In the latter program a strong effort is being made to find the low-mass red-dwarf stars in the solar neighborhood, and considerable success has been made (2006) in completing the census of the nearby stars. Most ground-based parallaxes now being determined have errors between $\pm 0.0005''$ and $\pm 0.003''$. *See* ASTRONOMICAL IMAGING; CHARGE-COUPLED DEVICES.

Space-based parallaxes. In 1989 the European Space Agency launched the first space satellite entirely devoted to astrometry, named *HIPPARCOS* (from *High Precision Parallax Collecting Satellite*, and also honoring the ancient astronomer Hipparchus). In 1993 the ESA completed its mission by obtaining positions, parallaxes, and proper motions for 118,322 stars listed in its final catalog. The accuracy averages about $0.0015''$ for stars brighter than the completeness limit of about the 8th magnitude, and rises to perhaps twice that amount at the ultimate limiting magnitude of 12. This precision is about equal to that of the best of current ground-based work for the brighter stars and less for faint stars. The advantage of *HIPPARCOS* lies in the much greater numbers of stars, including all the bright ones, with quality parallaxes. Among the common stars, the greatest contribution of the mission is in the calibration of the luminosities of the upper-middle main sequence and the giant and subgiant stars. The contributions to stellar knowledge from this satellite are enormous. While the *Hipparcos Catalogue* is the preferred source for most stars brighter than 9th magnitude, the *Yale Parallax Catalogue* retains its usefulness as the only source for parallaxes of stars fainter than that limit, as well as for stars requiring long time duration such as for the determination of the masses of the stars. *See* BINARY STAR; HERTZSPRUNG-RUSSELL DIAGRAM.

In 1990 the National Aeronautics and Space Administration (NASA) launched the *Hubble Space Telescope* (*HST*) with astrometric capabilities in its imaging cameras and fine guidance sensors. A relatively small number of high-accuracy parallaxes have been produced for stars of special calibration importance, such as Cepheids, planetary nebulae, and novae. The accuracy of the *Hubble Space Telescope* fine guidance sensor parallaxes reaches 200 microarcseconds. *See* HUBBLE SPACE TELESCOPE.

NASA plans to launch the *Space Interferometer Mission* (SIM) to measure parallaxes for a few ten thousand stars at an accuracy of 4 microarcsec, but a cutback in funding in 2006 delayed the launch to the 2015–2016 time frame. The European Space Agency is planning *Gaia* for a 2011 launch to measure parallaxes at an accuracy of about 10 microarcsec for a billion (10^9) stars. These two projects will revolutionize investigations involving galactic structure and setting constraints on the distances of stars, and therefore their masses and luminosities. The hundredfold increase in accuracy over the current best achievable accuracy will make it possible to directly determine parallaxes throughout the Milky Way Galaxy, rather than being limited to a few hundred parsecs, that is, within the solar neighborhood.

Statistical and secular parallax. Distance estimates are possible based on the assumption that the apparent mean space motion of a large number of stars with similar characteristics is a reflection of the peculiar motion of the Sun, derived from previous observations. Two related methods apply; the first equates the mean drift of the stars to the Sun's motion vector. The second requires measurements of radial velocities as well as proper motions, and assumes that the radial velocity distribution at the solar apex and antapex (the directions toward and away from which the Sun is moving, respectively) matches that in the transverse velocities along the direction normal to that of the assumed solar motion. These methods are referred to as secular and statistical parallax, respectively, although the terms are sometimes confused.

Indirect parallaxes. Several methods for the establishment of distance are frequently and incorrectly labeled parallax. These methods, such as spectroscopic or photometric parallax, have in common the reverse of the procedure typical of all true parallax methods. In any true parallax determination, a direct comparison of the same distance in linear and angular measure leads directly to a distance. By

contrast, such methods as spectroscopic and photometric parallax involve a comparison of an object to another of known distance and like luminosity, and scaling the distance as necessary. *See* ASTROMETRY.

William F. van Altena

Bibliography. W. F. van Altena, J. T. Lee, and E. D. Hoffleit, *General Catalogue of Trigonometric Stellar Parallaxes*, 1995; S. Debarbat et al. (eds.), *Mapping the Sky: Past Heritage and Future Directions*, 1988; H. K. Eichhorn and R. J. Leacock (eds.), *Astrometric Techniques*, 1994; European Space Agency, *The Hipparcos and Tycho Catalogues*, SP-1200, 1997; European Space Agency, *Proceedings of the ESA Symposium: Hipparcos Venice '97*, 1997; P. K. Seidelmann and A. K. B. Monet (eds.), *Astrometry in the Age of the Next Generation of Large Telescopes*, ASP Conf. Ser., vol. 338, 2005.

Parallel circuit

An electric circuit in which the elements, branches (elements in series), or components are connected between two points with one of the two ends of each component connected to each point. The **illustration** shows a simple parallel circuit. In more complicated electric networks one or more branches of the network may be made up of combinations of series or series-parallel elements. *See* CIRCUIT (ELECTRICITY).

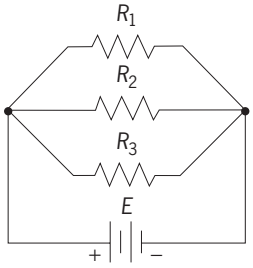

Schematic of a parallel circuit.

In a parallel circuit the potential difference (voltage) across each component is the same. However, the current through each branch of the parallel circuit may be different. For example, the lights and outlets in a house are connected in parallel so that each load will have the same voltage (120 V), but each load may draw a different current (0.50 A in a 60-W lamp and 10 A in a toaster).

For a discussion of parallel circuits *see* ALTERNATING-CURRENT CIRCUIT THEORY; DIRECT-CURRENT CIRCUIT THEORY.

Clarence F. Goodheart

Paramagnetism

A property exhibited by substances which, when placed in a magnetic field, are magnetized parallel to the field to an extent proportional to the field (except at very low temperatures or in extremely large magnetic fields). Paramagnetic materials always have permeabilities greater than 1, but the values are in general not nearly so great as those of ferromagnetic materials. Paramagnetism is of two types, electronic and nuclear.

Paramagnetic substances. The following types of substances are paramagnetic:

1. All atoms and molecules which have an odd number of electrons. According to quantum mechanics, such a system cannot have a total spin equal to zero; therefore, each atom or molecule has a net magnetic moment which arises from the electron spin angular momentum. Examples are organic free radicals and gaseous nitric oxide.

2. All free atoms and ions with unfilled inner electron shells and many of these ions when in solids or in solution. Examples are transition, rare-earth, and actinide elements and many of their salts. This includes ferromagnetic and antiferromagnetic materials above their transition temperatures. For a discussion of these materials

3. Several miscellaneous compounds including molecular oxygen and organic biradicals.

4. Metals. In this case, the paramagnetism arises from the magnetic moments associated with the spins of the conduction electrons and is called Pauli paramagnetism. *See* ANTIFERROMAGNETISM; FERRIMAGNETISM; FERROMAGNETISM.

Relatively few substances are paramagnetic. Aside from the Pauli paramagnetism found in metals, the most important paramagnetic effects are found in the compounds of the transition and rare-earth elements which have partially filled $3d$ and $4f$ electron shells respectively.

Electronic paramagnetism. This arises in a substance if its atoms or molecules possess a net electronic magnetic moment. The magnetization arises because of the tendency of a magnetic field to orient the electronic magnetic moments parallel to itself. The magnitudes of electronic magnetic moments are of the order of a Bohr magneton, which is equal to 9.27×10^{-21} electromagnetic unit, or emu (erg/gauss). *See* ELECTRON SPIN.

Nuclear paramagnetism. This arises when there is a net magnetic moment due to the magnetic moments of the nuclei in a substance. An example is solid sodium, in which each sodium atom has a nuclear magnetic moment of 2.217 nuclear magnetons. One nuclear magneton is equal to 5.05×10^{-24} emu. Nuclear magnetic moments are about 10^3 times smaller than electron magnetic moments. As a result, nuclear paramagnetism produces effects 10^6 times smaller than electron paramagnetic or diamagnetic effects. Therefore, it is usually impossible to detect nuclear paramagnetism by static methods since it will be masked by electronic effects. (An exception is the case of nuclear paramagnetism arising from the protons in solid hydrogen.) However, paramagnetic effects of nuclei are directly observable in resonance experiments. *See* DIAMAGNETISM; MAGNETIC RESONANCE; NUCLEAR MOMENTS.

Langevin theory. The Langevin theory of paramagnetism treats the paramagnetic substance as a classical (non-quantum-mechanical) collection of

permanent magnetic dipoles with no interactions between them. The dipoles are the magnetic moments of the paramagnetic atoms or ions in the substance. The first task of a theory of paramagnetism is to account for the experimentally observed susceptibility (ratio of magnetization to applied field). *See* MAGNETIC SUSCEPTIBILITY.

If an external magnetic field is applied to the paramagnet, each magnetic dipole experiences a torque. Associated with the force which produces this torque is a potential energy given by Eq. (1),

$$V = -\mu H \cos\theta \tag{1}$$

where μ is the magnetic moment of the dipole, H is the applied magnetic field intensity, and θ is the angle between the dipole and the direction of H. Now, in the absence of thermal agitation each permanent magnetic dipole will become oriented in such a way that this potential energy is minimized, that is, oriented parallel to the magnetic field. With all the dipoles lined up, the magnetization (magnetic moment per unit volume), if there are N dipoles per unit volume, would be given by Eq. (2), where the

$$M = N\mu \tag{2}$$

direction of the magnetization would be that of the applied field. Note that in this case an arbitrarily small magnetic field causes all the dipoles to line up so that the susceptibility $\chi = M/H$ would be infinite. In the actual case there is thermal agitation which in part offsets the aligning tendency of the magnetic field. The Langevin theory takes this into account and predicts the paramagnetic susceptibility as a function of temperature.

In the presence of thermal agitation, the magnetic dipoles are not all lined up in the direction of the magnetic field, but there is some distribution of angles made with the field. In this case the magnetization is given by Eq. (3), where $\cos\theta$ is the aver-

$$M = N\mu\,\overline{\cos\theta} \tag{3}$$

age of the cosine of the angle between dipole and field. The average is taken over the distribution of dipoles in thermal equilibrium. According to statistical mechanics, this average is given by Eq. (4), where

$$\overline{\cos\theta} = \frac{\int e^{(-V/kT)}\cos\theta\, d\Omega}{\int e^{(-V/kT)}\, d\Omega} \tag{4}$$

$d\Omega$ is the element of solid angle and $e^{(-V/kT)}$ is the Boltzmann distribution in energy $V = -\mu H \cos\theta$ [Eq. (1)] of a dipole at angle θ with respect to the applied field at absolute temperature T. The integrations may be performed, and the result is $L(a)$, the Langevin function of $a = \mu H/kT$. The result may be combined with Eq. (3) to give Eq. (5). If $a \ll 1$,

$$M = N\mu L(a) \tag{5}$$

then $L(a) \cong b\ a/3$ so that Eq. (6) holds. This is a

$$M \cong \frac{N\mu^2 H}{3kT} \tag{6}$$

good approximation except at low temperatures or extremely high fields. The susceptibility is given by Eq. (7). The $1/T$dependence of the susceptibility is

$$\chi = \frac{M}{H} = \frac{N\mu^2}{3kT} = \frac{C}{T} \tag{7}$$

known as the Curie law. The Curie law was established empirically by P. Curie in 1895 and is obeyed by many gases, liquids, and solids. There are some paramagnetic solids which obey the Curie-Weiss law $\chi = C/(T - \Theta)$ in a certain temperature range. Here Θ is the Curie temperature. The modification often arises because of effective interactions between the dipoles which are neglected in the preceding development. It may also be due to distortion effects. *See* CURIE-WEISS LAW; LANGEVIN FUNCTION.

Experimental data for the paramagnetic susceptibility are often expressed in terms of the effective magnetic moment which must be used for μ in the Curie law [Eq. (7)] in order to give the observed slope of the curve of χ plotted against $1/T$.

Quantum theory. The quantum-mechanical theory of paramagnetism was worked out in detail by J. H. Van Vleck in 1928. This theory is based on the fact that the magnetic moment of the permanent magnetic dipole arises from the total angular momentum of the electrons in the paramagnetic atom, ion, or molecule. Thus an atom with total angular momentum quantum number J has $(2J + 1)$ energy levels in a magnetic field. A collection of such atoms will be distributed among these levels according to a Boltzmann distribution. The magnetization of such a system may be computed by finding the average component of angular momentum parallel to the field. The result is Eq. (8), where g is the spectroscopic

$$M = NgJ\mu_B B_J(a^*) \tag{8}$$

splitting factor (the measure of the energy level splittings of the system), μ_B is the Bohr magneton, $a^* = gJ\mu_B H/kT$, and $B_J(a^*)$ is the Brillouin function of a^* expressed in Eq. (9). The Brillouin function also en-

$$B_J(a^*) = \frac{2J+1}{2J}\coth\frac{(2J+1)a^*}{2J} - \frac{1}{2J}\coth\frac{a^*}{2J} \tag{9}$$

ters the theory of ferromagnetism. If a^* is much less than unity, which is a good approximation except at very low temperatures or in large fields, then Eq. (10)

$$B_J(a^*) \cong \frac{g(J+1)\mu_B H}{3kT} \tag{10}$$

holds. In this case a Curie law again prevails as in Eq. (11). The effective magneton number is defined

$$\chi = \frac{M}{H} = \frac{NJ(J+1)g^2\mu_B^2}{3kT} \tag{11}$$

by $g\sqrt{J(J+1)}$ and is the quantity usually given in

experimental results. *See* NONRELATIVISTIC QUANTUM THEORY.

If only the electron spin contributes to the total angular momentum, $J = {}^1/_2$ and $B_{1/2}(a^*) = \tanh(a^*)$ so that, except at low temperatures or high fields, Eq. (12) holds, which agrees with the classical result.

$$\chi = \frac{N\mu_B^2}{3kT} \tag{12}$$

This case is referred to as the spin-only case.

Rare-earth ions. The paramagnetism of rare-earth ions at room temperature is summarized by some representative examples in **Table 1**.

The calculated effective magneton numbers in Table 1 are the theoretical values for isolated ions. The experimental values are derived from Eq. (11), using experimental values of the paramagnetic susceptibility χ. There is good agreement for all rare-earth ions with the exception of europium and samarium. The experimental results of Table 1 refer to the paramagnetic behavior of rare-earth ions in crystals; different salts of the same ion give the same results.

The experimental result is therefore that at room temperature a crystal containing a number of trivalent rare-earth ions has the paramagnetic susceptibility of that number of free trivalent ions. The reason that there is little influence of the crystalline electric fields on the magnetic behavior is that the electrons responsible for the magnetic moments are in the $4f$ state and therefore occupy an electronic shell lying well inside the ion, a shell that is shielded from outside influence by the $5s$ and $5p$ electrons. This is in contrast to the behavior of iron-group ions discussed later.

At lower temperatures the influence of the crystalline electric fields on the electrons becomes more important and the behavior of the susceptibility can become quite complex. In this case, the susceptibility depends upon the orientation of the magnetic field with respect to the crystal axes.

The behavior of europium and samarium at room temperature is still explainable on the basis of a theory of free ions if the effect of Van Vleck paramagnetism is included.

Van Vleck paramagnetism. This arises when the energy states of an atom or ion divide into two groups, those within an energy kT of the ground (lowest energy) state and those which are separated from the ground state by an energy greater than kT. Here k is Boltzmann's constant and T is the absolute temperature. The situation is shown in the **illustration**. The low-lying states give rise to a susceptibility which follows a Curie law. If these low-lying states arise from a single value of the total angular momentum J, as in the illustration, then the quantum-mechanical derivation applies [Eq. (11)]. The high-lying states give rise to a small temperature-independent susceptibility, an effect which is known as Van Vleck paramagnetism. In intermediate cases, such as in the trivalent europium and samarium ions, the upper states are only a little more than kT away from the ground state so that the temperature dependence is still more complicated.

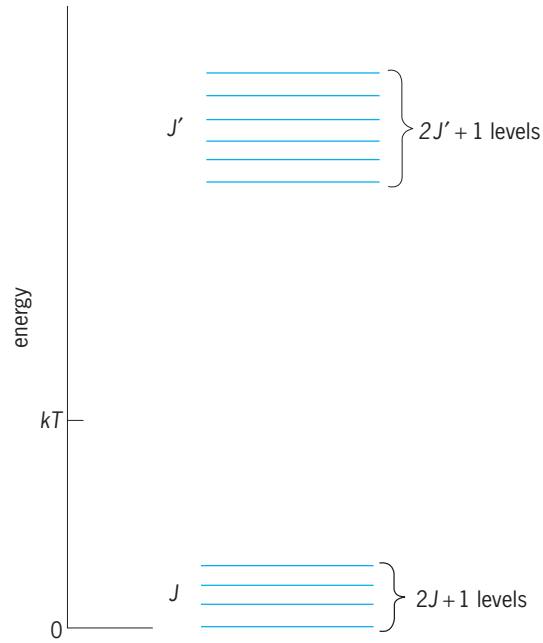

Energy levels in Van Vleck paramagnetism.

Iron-group ions. The paramagnetism of iron-group ions in crystals is summarized in **Table 2**.

Quenching of orbital angular momentum is exhibited in crystal containing ions of the iron group. The last three columns of Table 2 indicate that the orbital angular momentum makes no contribution to the magnetic moment but the iron-group ions behave in crystals as free ions with only the spin S contributing to the magnetic effects. This is evidenced by the fact that the spin-only values of the effective magneton numbers agree well with the experimental results. The orbital angular momentum is quenched because the $3d$ electronic shell, which gives rise to the paramagnetism, is outermost for the iron group; it is therefore exposed to the strong crystalline electric fields arising from neighboring ions. These asymmetric electric fields decouple the orbital angular momentum from the spin angular momentum. This means that the energy levels are no longer specified by the total angular momentum quantrum number J; S alone may determine the levels. More precisely, the $(2L + 1)$ degenerate orbital angular momentum states of orbital angular momentum

TABLE 1. Paramagnetism of some trivalent rare-earth ions

Ion	Electron configuration	Effective magneton number	
		Calculated	Experimental
Ce^{3+}	$4f^1 5s^2 p^6$	2.54	2.4
Nd^{3+}	$4f^3 5s^2 p^6$	3.62	3.5
Sm^{3+}	$4f^5 5s^2 p^6$	0.84	1.5
Eu^{3+}	$4f^6 5s^2 p^6$	0.00	3.4
Gd^{3+}	$4f^7 5s^2 p^6$	7.94	8.0
Yb^{3+}	$4f^{13} 5s^2 p^6$	4.54	4.5

TABLE 2. Paramagnetism of iron-group ions

Ion	Electron configuration	Effective magneton number: Calculated with J	Calculated with S only	Experimental
Ti^{3+}, V^{4+}	$3d^1$	1.55	1.73	1.8
V^{3+}	$3d^2$	1.63	2.83	2.8
Cr^{3+}, V^{2+}	$3d^3$	0.77	3.87	3.8
Mn^{3+}, Cr^{2+}	$3d^4$	0.00	4.90	4.9
Fe^{3+}, Mn^{2+}	$3d^5$	5.92	5.92	5.9
Fe^{2+}	$3d^6$	6.70	4.90	5.4
Co^{2+}	$3d^7$	6.54	3.87	4.8
Ni^{2+}	$3d^8$	5.59	2.83	3.2
Cu^{2+}	$3d^9$	3.55	1.73	1.9

quantum number L may be split by the crystal fields so that the lowest orbital state is nondegenerate (singlet). Then there is no possibility of orienting the orbital angular momentum by a magnetic field so that only the spin contributes to the magnetic moment. It is often said that the crystal field "locks" the orbital angular momentum so that it cannot be oriented by a magnetic field. Partial quenching occurs when the orbital degeneracy is only partially removed by the crystal field. Partial quenching and anisotropic effects can also be caused by spin-orbit coupling. This influence may be observed in spin-resonance and specific-heat experiments. *See* ADIABATIC DEMAGNETIZATION; MAGNETIC RESONANCE.

Pauli paramagnetism. This is the paramagnetism associated with the conduction electrons of a metal. A metal is usually described in terms of a collection of positive ions with closed shells which are arranged on a crystal lattice plus electrons which are essentially free to move about the crystal. Each electron has an intrinsic spin angular momentum, and these momenta give rise to a paramagnetic magnetic moment. At first sight it would seem to be correct to apply the Langevin formula to this "gas" of electrons, but the experimental facts are that the paramagnetic susceptibility of conduction electrons is about one-hundredth of that predicted by the Langevin formula [Eq. (7)]. Furthermore, the susceptibility is temperature independent rather than varying as $1/T$ (Curie law). The explanation was given by W. Pauli in 1927: The electrons obey the quantum statistics of E. Fermi and P. A. M. Dirac rather than the classical statistics which are used in the derivation of the Langevin formula. This means that a given energy state can be occupied at most by two electrons, and their spin angular momenta must be in opposite directions. As a result the net angular momentum is zero, even on application of a magnetic field. Thus most of the electrons in a metal contribute in sum no magnetic moment. That is to say, an electron's spin angular momentum may not orient parallel to an applied magnetic field because there is already an electron in that energy state with its spin parallel to the field. There are, however, a few electrons which are not "paired off," and the spins of these can be oriented by the field. These electrons contribute to the susceptibility according to a Curie law, but the number of them is proportional to the temperature. The combination of the two temperature dependences leads to a temperature-independent susceptibility, smaller than the prediction of the Langevin formula for N electrons per unit volume because only a fraction of these may contribute. The Pauli susceptibility may be written (for a free electron gas) as Eq. (13), where N is the number of electrons per

$$\chi = \frac{3N\mu_B^2}{2kT_F} \tag{13}$$

unit volume and kT_F is the Fermi energy characteristic of the metal. The fraction of electrons contributing to the susceptibility at temperature T is of the order T/T_F. For sodium, for example, $kT_F = 3.12$ eV, $T_F = 37{,}000$ K. The Pauli paramagnetism of metals has been observed in spin-resonance experiments. The total susceptibility arises from Pauli paramagnetism and diamagnetic contributions from conduction electrons and ion cores. *See* EXCLUSION PRINCIPLE; FREE-ELECTRON THEORY OF METALS.

Elihu Abrahams; Frederic Keffer

Bibliography. D. Craik, *Magnetism*, 1994; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; A. H. Morrish, *The Physical Principles of Magnetism*, 1965, reprint 1980; J. H. Van Vleck, *The Theory of Electric and Magnetic Susceptibilities*, 1932.

Parameter

An auxiliary variable, functions of which give the coordinates of a curve or surface. The coordinates of a curve are functions of one parameter. A curve in 3-space has parametric equations (1).

$$x = f(t) \qquad y = g(t) \qquad z = h(t) \tag{1}$$

The coordinates of a surface are functions of two parameters, shown in Eqs. (2).When constant values

$$x = f(u, v) \qquad y = g(u, v) \qquad z = h(u, v) \tag{2}$$

are assigned to the parameters u or v, these equations represent the parametric curves on the surface. General surface curves are obtained by replacing u and v by functions of a new parameter t.

An arbitrary constant in an equation is also called a parameter. Variations in the values of the parameter generate a system of equations which may represent a family of curves or surfaces. Such families are called one-parameter, two-parameter, and so on, according to the number of independent parameters. *See* PARAMETRIC EQUATION.

Louis Brand

Parametric amplifier

A highly sensitive low-noise amplifier for ultrahigh-frequency and microwave radio signals, utilizing as the active element an inductor or capacitor whose reactance is varied periodically at another microwave or ultrahigh frequency. A varactor diode is

most commonly used as the variable reactor. The varactor is a semiconductor *pn* junction diode, and its junction capacitance is varied by the application of a steady signal from a local microwave oscillator, called the pump. Amplification of weak signal waves occurs through a nonlinear modulation or signal-mixing process which produces additional signal waves at other frequencies. This process may provide negative-resistance amplification for the applied signal wave and increased power in one or more of the new frequencies which are generated. *See* VARACTOR.

The parametric amplifier operates through a periodic variation of a circuit parameter, in this case, the capacitance of the varactor. Similar amplification is possible by using a periodically variable inductor which may utilize a saturable magnetic material such as a ferrite. The inductive form of this device has not found application because it requires much larger amounts of pumping power to vary the inductance significantly at microwave frequencies.

Effects. Parametric effects are also found with variable resistive elements such as the varistor or rectifying diode. These elements do not commonly provide amplification of signals but are widely used as frequency converters in radio receivers to translate signals from high frequencies to lower or intermediate frequencies for more convenient signal amplification. *See* VARISTOR.

The varactor diode is a simple *pn* junction semiconductor device designed to maximize the available reactance variation, to minimize signal losses in electrical resistance, and to require low microwave power from the pump. The junction is biased with a low dc voltage in the nonconducting direction and the applied voltage of the microwave pump wave causes the total bias to vary periodically at the microwave pumping frequency. The capacitance of a reverse-biased semiconductor diode varies with voltage, decreasing as the bias voltage is increased. Common diodes with an abrupt junction between the *p* and *n* regions have a capacitance variation according to the equation shown below, where C is the capac-

$$C = \frac{k}{\sqrt{V_i + V_b}}$$

itance in farads, k is a constant depending upon the area of the junction and the doping concentrations of impurities in the two semiconductor regions, V_i is an internal or "built-in" voltage (a few tenths of a volt in most diodes), and V_b is the applied bias voltage (in the nonconducting direction). When used in a parametric amplifier, V_b includes a steady dc component plus a microwave ac value, varying periodically with time at a very high rate.

There are several possible circuit arrangements for obtaining useful parametric amplification. The two most common are the up-converter and the negative-resistance amplifier. In both types, the pump frequency is normally much higher than the input-signal frequency. In the up-converter, a new signal wave is generated at a higher power than the input wave. In the negative-resistance device, negative resistance is obtained for the input-signal frequency, causing an enhancement of signal power at the same frequency. *See* NEGATIVE-RESISTANCE CIRCUITS.

When a varactor is energized with two simultaneous voltage waves of different frequencies, such as a pump wave at a frequency f_p and a signal wave at a frequency f_s, the current induced may have a complex waveform which contains other frequencies. Mathematical analysis shows that the first-order and most significant of these new frequencies are the sum $(f_p + f_s)$ and the difference $(f_p - f_s)$. Depending on the varactor characteristics and the strength of the applied waves, still other frequencies can be produced with frequencies such as $nf_p \pm mf_s$, where m and n are integers of any value, including zero. Normally only the sum or the difference frequency is used.

Analysis also shows that, if the varactor and its surrounding network are so tuned that significant power is generated by the varactor and dissipated at the new frequency $f_p - f_s$ and not at $f_p + f_s$, the varactor will present a negative-resistive input impedance to an input wave at the frequency f_s. Instead of being absorbed by the amplifier network, the input wave will be reflected back down the input transmission line with an increased power. The wave at the frequency $f_p - f_s$ is not utilized outside the amplifier and is called, therefore, the idler frequency. This is the mechanism of the negative-resistance form of the parametric amplifier, the most common type. A functional block diagram of this type of amplifier is shown in the **illustration**.

Alternatively, if the network is so designed that the sum frequency $f_p + f_s$ is generated and transmitted out by a second transmission line, and if significant energy is not generated or dissipated at the

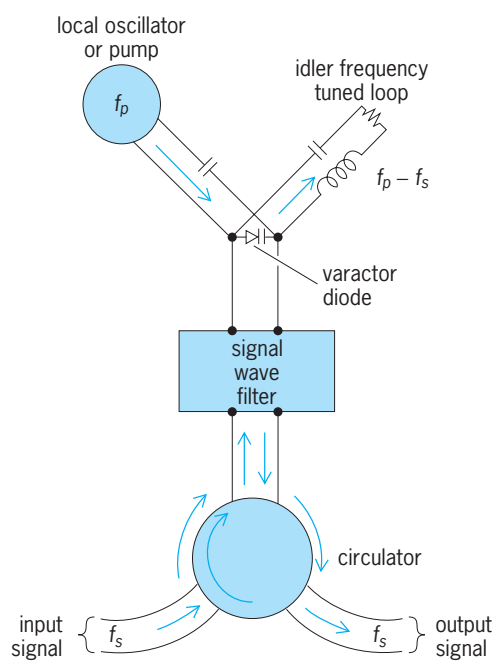

Negative-resistance-type parametric amplifier.

difference frequency $f_p - f_s$, the power output at $f_p + f_s$ can be greater than the power input at f_s by a factor which may approach the ratio of these two frequencies $(f_p + f_s)/f_s$. This is the mechanism of the up-converter form of the parametric amplifier, so named because the signal wave is converted to a new higher frequency band.

Advantages. The most important advantage of the parametric amplifier is its low level of noise generation. In most other amplifiers, current flows through the active device by discretely charged electrons passing from one electrode to another. Random fluctuations in the rate of such flow normally occur, providing a finite background level of electrical noise (called shot noise) in the output which may override very weak signals, destroying their usefulness. Similar weak fluctuation noise is generated by resistive elements at normal temperatures in amplifier networks (called thermal noise). In the parametric amplifier, relatively few electrons pass from one electrode to another in the varactor, reducing the shot-noise effects. Resistance and thermal noise generation effects are also minimized by proper diode and network design.

The parametric amplifier finds its greatest use as the first stage at the input of microwave receivers where the utmost sensitivity is required. Its noise performance has been exceeded only by the maser. Maser amplifiers are normally operated under extreme refrigeration using liquid helium at about 4° above absolute zero. The parametric amplifier does not require such refrigeration but in some cases cooling to very low temperatures has been used to give improved noise performance that is only slightly poorer than the maser. *See* MASER. M. E. Hines

Bibliography. J. C. Decroly et al., *Parametric Amplifiers*, 1973; K. H. Loecherer and C. D. Brandt, *Parametric Electronics: An Introduction*, 1981.

Parametric array

A virtual array of sources of sound in a fluid (more fully termed a parametric acoustic array) that is formed by two collinear acoustic waves propagating in the same direction, which interact because of the nonlinearity (parameter) of the fluid medium to generate new acoustic fields. The array is virtual because it is formed in the fluid medium away from the devices responsible for the primary field components. If the primary interacting field is composed of two essentially monochromatic waves, with frequencies ν_1 and ν_2, then the new fields will form at the sum and difference frequencies, $\nu_1 \pm \nu_2$. If ν_1 and ν_2 are slightly different, then the difference frequency, $\nu_1 - \nu_2$, will be quite small. Since absorption of sound in fluids generally increases rapidly with frequency, this difference-frequency wave will outlast the primary- and sum-frequency waves. In addition to suffering less absorption and propagating further, the difference-frequency wave possesses two remarkable properties: most of its energy is concentrated in a narrow anglar sector in the direction of propagation of the primary waves, and it has no sidelobes. Parametric arrays can be used for both directional transmission and directional reception of sound. *See* DIRECTIVITY; SOUND; SOUND ABSORPTION.

Theory. Liquids and gases are intrinsically nonlinear. Only at infinitesimal amplitudes is the theory of sound in fluids linear. At finite amplitudes, the nonlinear relationship of pressure and density causes the instantaneous speed of propagation of a pressure disturbance to vary with the state of compression: the crest of the wave travels faster than any other part of the wave, and the trough travels slower. A wave that is initially sinusoidal becomes distorted as it propagates: the forward flank of a crest advances toward the perpendicular, and the trailing flank of a trough lags similarly toward the perpendicular. Before these limiting states are attained, other nonlinear effects will begin to act against this steepening of wave fronts. In the extreme case of a shock wave, the wave front does effectively become vertical, characterizing a pressure discontinuity or jump between crest and trough. *See* NONLINEAR ACOUSTICS; SHOCK WAVE.

Several wave types derive from the same underlying set of equations, describing infinitesimal- or small-amplitude, finite-amplitude, and shock waves, but these admit of significant simplification in some cases. The equations describing shock are more complicated than those describing lesser finite-amplitude waves, which in turn are more complicated than those describing infinitesimal- or small-amplitude waves. The complexity of finite-amplitude waves is such that there are gradations in complexity of the simplified equations.

In the case of parametric arrays, the medium is assumed to be weakly nonlinear, but this still admits of rather intense sound sources. Near sea level, pressure amplitudes in water due to primary, source waves can be comparable to the static pressure, hence of order 1 atm or about 10^5 Pa; in the logarithmic domain, this is about 220 dB re 1 μPa at 1 m. *See* DECIBEL; SOUND PRESSURE.

For waves propagating collinearly in the same direction, which is a requirement for the interaction to occur, the simplest wave equation for the pressure field p in the neglect of absorption is Eq. (1).

$$\Box^2 p = -\frac{\beta}{\rho_0 c_0^4}\frac{\partial^2 p^2}{\partial t^2} \tag{1}$$

Here, $\Box^2$ is the d'Alembertian operator,

$$\Box^2 = \nabla^2 - \frac{1}{c_0^2}\frac{\partial^2}{\partial t^2}$$

which includes second-order derivatives with respect to space through the Laplacian ∇^2 and time through $\partial^2/\partial t^2$, c_0 is the small-amplitude sound speed, ρ_0 is the static mass density of the medium, and β is the parameter of nonlinearity, given by Eq. (2),

$$\beta = 1 + \frac{\rho_0}{2c_0^2}\left(\frac{\partial^2 p}{\partial p^2}\right)\rho_0 \tag{2}$$

where the second derivative is evaluated at $\rho = \rho_0$. *See* LAPLACIAN.

Consistent with the level of approximation, it is assumed, as is common in linear acoustics, that propagation of an acoustic wave is an isentropic process: the vibrations are so rapid that heat cannot be exchanged with the surroundings. It is thus possible to distinguish β for gaseous and liquid media. For gases, $\beta = (\gamma + 1)/2$, where γ is the ratio of specific heats at constant pressure and constant volume. Numerical values for γ lie within the range 1–1.67. For monatomic gases such as helium, $\gamma = 1.66$; for diatomic gases including air, $\gamma = 1.4$; for more complex molecules, γ is generally lower. For liquids, $\beta = 1 + B/(2A)$, where A and B are given by Eqs. (3).

$$A = \rho_0 \left(\frac{\partial p}{\partial \rho} \right)_{\rho_0} = \rho_0 c_0^2 \tag{3a}$$

$$B = \rho_0^2 \left(\frac{\partial^2 p}{\partial \rho^2} \right)_{\rho_0} \tag{3b}$$

The quantity B is typically measured. Values of $B/(2A)$ have been tabulated for many liquids. For distilled water it is 5.0; seawater, 5.25; ethanol, 10.5; cod liver oil, 11.0—all at 20°C (68°F). *See* SPECIFIC HEAT.

To solve Eq. (1), the pressure field p is represented as the sum of the primary field p_i and the scattered field p_s, namely that resulting from the nonlinear interaction of the waves composing the primary field. The primary field satisfies the linear wave equation $\Box^2 p_i = 0$. Thus, to the same level of approximation, the scattered field satisfies Eq. (4).

$$\Box^2 p_s = -\frac{\beta}{\rho_0 c_0^4} \frac{\partial^2 p_i^2}{\partial t^2} \tag{4}$$

That is, p_s is due to the nonlinear source term on the right side, being quadratic in p_i.

It is now assumed that the primary field is composed of two sinusoidal waves of equal amplitude p_0 and slightly different frequencies propagating in the same x direction (see **illustration**), and is thus given by Eq. (5),

$$p_i = p_0 \exp(-\alpha x)[\cos(k_1 x - \omega_1 t) + \cos(k_2 x - \omega_2 t)] \tag{5}$$

where α is the absorption coefficient, introduced here in an ad hoc way rather than through Eq. (1), k_1 and k_2 are the respective wavenumbers of the collimated primary waves, and ω_1 and ω_2 are the respective angular frequencies, corresponding to the frequencies ν_1 and ν_2: $\omega_i = 2\pi\nu_i = c_0 k_i$, $i = 1, 2$.

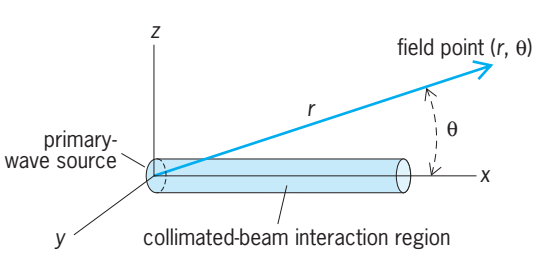

Geometry of a parametric acoustic array.

It is assumed further that the primary waves are collimated, with cross-sectional area S_0. The solution to Eq. (4) for the difference-frequency wave at the farfield range r and angle θ can be derived by integrating the source field along the entire x axis, recognizing that absorption will effectively render this distance finite. The result is Eq. (6),

$$p_s = \frac{p_0^2 \beta S_0 \omega_s^2}{8\pi \rho_0 c_0^4} \frac{\exp[i(k_s r - \omega_s t)]}{r[i\alpha + k_s \sin^2(\theta/2)]} \tag{6}$$

where $\omega_s = \omega_1 - \omega_2 = c_0 k_s$. In this result, absorption of the difference-frequency wave has been neglected. The corresponding intensity is given by Eq. (7).

$$I_s = \frac{p_0^4 \beta^2 S_0^2 \omega_s^4}{128\pi^2 \rho_0^3 c_0^9} \frac{1}{r^2[\alpha^2 + k_s^2 \sin^4(\theta/2)]} \tag{7}$$

See SOUND INTENSITY.

Properties. The solution for the scattered difference-frequency wave is given in Eq. (6). This describes an outgoing spherical wave $\exp[i(k_s r - \omega_s t)]/r$, with constant factor $p_0^2 \beta S_0 \omega_s^2/(8\pi \rho_0 c_0^4)$, times an angular function $[i\alpha + k_s \sin^2(\theta/2)]^{-1}$. This angular dependence is known as the Rutherford scattering function, which also describes the scattering of alpha particles directed at a gold foil, due to the Coulomb interaction. This function has its peak in the forward direction, $\theta = 0$, where its amplitude is $1/\alpha$. It decreases monotonically with increasing scattering angle θ to its minimum at $\theta = \pi$, where its amplitude is $(\alpha^2 + k_s^2)^{-1/2}$. That is, the parametric array has no sidelobes. *See* SCATTERING EXPERIMENTS (NUCLEI).

A measure of beamwidth of the array is given by the angle at which the intensity falls to one-half its peak value. If the fractional absorption of the primary waves over a difference-frequency wavelength is small, hence $\alpha/k_s \ll 1$, this angle is $2(\alpha/k_s)^{1/2}$. Thus, the parametric array is very directional.

Such directionality can also be achieved linearly, but this would require a very large transducer indeed. Admittedly, the efficiency of linear generation of sound is far greater than that of the parametric array, but it is also accompanied by sidelobes, which are absent in the parametric array.

Another special property of the parametric array is its potentially wide band character. Relatively small changes in or modulation of the primary field can produce a relatively broad band in difference frequency.

As earlier noted, the parametric array can serve both to transmit and to receive acoustic waves. If the two primary waves comprising p_i originate in a plane, as by two aligned transducers in proximity, the array acts as a transmitter of a difference-frequency wave. If the parametric array is aligned with the propagation direction of an external wave, and a single primary wave is launched by a transducer, then the two together comprise p_i and interact, enabling the

difference-frequency wave to be received and detected.

A generalization of amplitudes of the primary waves in Eq. (5) is useful. This and other generalizations will be outlined.

Generalizations. The parametric array described thus far follows the seminal development given by its discoverer, P. J. Westervelt. Notwithstanding important developments since then, the development presented above captures the essential elements of the parametric array. At the same time, its limitations are evident in the stated assumptions.

An immediately useful generalization is that of primary waves of unequal amplitude. The consequence of introducing separate primary-wave amplitudes in Eq. (5) is a generalization of Eqs. (6) and (7) in which the factor p_0^2 is replaced by the product of the two pressure amplitudes. This generalization enables the efficiency of the parametric array as a receiver of waves of unknown amplitude to be evaluated, for example, to determine its sensitivity.

Consideration of the source level and absorption rate of the primary waves leads to very useful generalizations of the parametric array, as elaborated by M. B. Moffett and colleagues. Three operating regimes are distinguished. These refer to two quantities: the scaled source level $SL_0 = SL + 20 \log \nu_0$, where SL is the mean source level of the primary waves and ν_0 is the mean primary frequency in kilohertz, and the nearfield absorption $\alpha_0 r_0$, where α_0 is the absorption coefficient at ν_0 and r_0 is the nearfield, or Rayleigh, distance, equal to the area of the source transducer divided by the wavelength at ν_0.

1. If $SL_0 > 280$ dB re 1 μPa at 1 m, then harmonic distortion of the primary beams limits the effective length of the parametric array, hence its gain. This state is called saturation.
2. If $SL_0 < 280$ dB re 1 μPa at 1 m and $\alpha_0 r_0 > 1$ dB, then the small-signal limit assumed by Westervelt applies, and the array is absorption-limited.
3. If $\alpha_0 r_0 \ll 1$ dB, then most of the absorption occurs in the farfield. The array is again absorption-limited.

For $\nu_0 = 18$, 40, and 100 kHz, which are representative of mean primary frequencies for some parametric arrays, $\alpha_0 = 0.0028$, 0.011, and 0.034 dB/m, respectively, assuming seawater of temperature 10°C (50°F), salinity 35 parts per thousand, pH 8, and depth 0 m. The coefficient α in Eq. (5) is derived from α_0 by the equation $\alpha = \alpha_0/8.68$. *See* PH.

Other generalizations of the parametric array have been investigated, but using more general forms of the wave equation than that in Eq. (1). These include dispersive effects in which the sound speed depends on frequency, which can be significant in the presence of air bubbles, as well as certain other inhomogeneities, principally investigated by M. F. Hamilton. Spherically spreading waves, rather than collimated plane waves, have been investigated by J. N. Tjøtta and S. Tjøtta, among others, who have considered more general effects of diffraction. At higher intensities, nonlinear absorption becomes a limiting factor and requires special treatment.

Applications. Many applications have been pursued underwater because of the special properties of the parametric array. The narrowness of the parametric beam is exploited in several applications requiring high angular resolution, for example, measurement of small patches of the sea surface and sea floor, and volumetric inhomogeneities. Objects lying on or buried in the sea bed are detected by means of parametric arrays. These include, for example, hazardous waste, including munitions dumped in the Baltic Sea and mines in many shallow coastal areas. Techniques of imaging and mapping with parametric sonar are used in these applications, as well as in marine archeology, resulting in the discovery of sunken vessels. The acoustic transmission loss has been measured to characterize sediments. Parametric sonar is used for subbottom profiling and seismic analysis, as in oil and gas exploration. *See* GEOPHYSICAL EXPLORATION; SEA-FLOOR IMAGING; SEISMOLOGY; SONAR.

Beamwidth and bandwidth properties of parametric arrays as transmitters are exploited in underwater communications and in fish-swimbladder-resonance absorption spectroscopy. In the second application, effects of excited, breathing-mode resonances of fish swimbladders on transmitted signals are measured, with the aim of determining the numerical density of such fish.

Calibration. Parametric arrays configured as sonar systems have been calibrated to determine the source level and range and angular dependences of the transmitted difference-frequency field. A standard-target method has been proposed for calibrating parametric sonars for backscattering applications. This would add another, quantitative dimension to many of the listed applications.

Kenneth G. Foote

Bibliography. R. T. Beyer, *Nonlinear Acoustics*, Naval Ship Systems Command, Department of the Navy, Washington, DC, 1974; M. J. Crocker (ed.), *Encyclopedia of Acoustics*, vol. 1, Wiley, New York, 1997.

Parametric equation

A type of mathematical equation used, typically, to represent curves in a plane or in space of three dimensions. In principle, however, there is no limitation to any particular number of dimensions. A parameter is actually an independent variable. In elementary analytic geometry a curve in the xy plane is often studied, in the first instance, as the locus of an equation $y = F(x)$ or $G(x,y) = 0$. The form $y = F(x)$ is not adequate for the complete representation of certain curves, whereas the form $G(x,y) = 0$ may be adequate. The circle $x^2 + y^2 - 16 = 0$ affords an example. But the form $G(x,y) = 0$ is not always convenient. The parametric form $x = f(t)$, $y = g(t)$ is often the most convenient; moreover, it is often the naturally occurring form of representation of the curve. For the circle $x^2 + y^2 - 16 = 0$, one possible parametric representation is $x = 4 \cos t$ and $y = 4 \sin t$.

Parametric curves. A pair of equations $x = f(t)$, $y = g(t)$, where f and g are continuous functions defined for some interval of values of t, for example, $a \leqq t \leqq b$, is said to define a parametric curve. The locus of points (x,y) obtained in this way is not always what would be regarded as a curve by a layman. If one thinks of t as time, the equations define the motion of the point (x,y) as t increases from a to b. Clearly the path can cross itself, double back on itself, or the point may even remain motionless. Even more surprising, the point may pass through every position within an entire square unit of area.

In calculus the customary restrictions placed on f and g in the case of a parametric curve are that f and g have derivatives f', g' which are continuous and such that $f'(t)$ and $g'(t)$ are not zero at the same time. Exceptions for isolated values of t are sometimes permitted. The result of these conditions is that a part of the curve near any one point on it can be represented either in the form $y = F(x)$ or in the form $x = F(y)$, where F is a differentiable function with a continuous derivative (a result of implicit-function theory). This implies that the parametric curve really looks like a curve in the intuitive sense and has a continuously turning tangent. The curve may cross itself, but the part generated by small variations of t from any fixed value t_0 is what is called a smooth arc.

The arc length L of a smooth arc from t_0 to t_1 is Eq. (1).

$$L = \int_{t_0}^{t_1} \left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 \right]^{1/2} dt \tag{1}$$

The corresponding differential formula is Eq. (2),

$$ds^2 = dx^2 + dy^2 \tag{2}$$

where s is arc length measured along the curve.

In polar coordinates the formula is Eq. (3).

$$ds^2 = dr^2 + r^2 d\theta^2 \tag{3}$$

The curvature of a curve is defined as $K = d\phi/ds$, where s is arc length and ϕ is the angle from a chosen fixed direction to the tangent drawn in the direction of increasing s. The parametric formula for K is Eq. (4), where the plus or minus sign is to be chosen

$$K = \pm \frac{x'y'' - y'x''}{(x'^2 + y'^2)^{3/2}} \tag{4}$$

according to whether ds/dt is positive or negative. Here Eqs. (5) hold.

$$x' = \frac{dx}{dt} \qquad x'' = \frac{d^2x}{dt^2} \qquad y' = \frac{dy}{dt} \qquad y'' = \frac{d^2y}{dt^2} \tag{5}$$

If $y = F(x)$ is obtained from $x = f(t)$, $y = g(t)$, it follows by the technique of differentials that $dy = F'(x)\,dx$ and $dx = f'(t)\,dt$, $dy = g'(t)\,dt$, whence Eq. (6) holds, and likewise Eq. (7). Higher derivatives

$$F'(x) = \frac{g'(t)}{f'(t)} \tag{6}$$

$$\frac{d^2y}{dx^2} = \frac{dF'(x)}{dx} = \frac{f'(t)\,g''(t) - g'(t)f''(t)}{[f'(t)]^3} \tag{7}$$

of F can also be found in terms of derivatives of f and g.

In some cases it is necessary to include what occurs as $t \to \pm\infty$ to complete a parametric representation in a natural way. For instance, Eqs. (8) rep-

$$x = \frac{2t}{1+t^2} \qquad y = \frac{1-t^2}{1+t^2} \tag{8}$$

resent the circle $x^2 + y^2 = 1$ except for the point $(0, -1)$. This point is obtained in the limit as either $t \to -\infty$ or $t \to +\infty$.

The cycloid $x = a(t - \sin t)$, $y = a(1 - \cos t)$ is an example of a curve that is easily and naturally represented in parametric form but is represented only with great awkwardness by a single equation in x and y.

Curves in space of three dimensions are represented parametrically in the form $x = f(t)$, $y = g(t)$, $z = h(t)$. The conditions for a smooth arc are similar to the conditions in the plane case.

Parametric surface. A parametric surface in space of three dimensions is defined by $x = f(u, v)$, $y = g(u, v)$, $z = h(u, v)$ where f, g, h are continuous functions of the two parameters u, v. In order to have a surface which conforms to the intuitive ideas of a smooth surface, it is sufficient to impose the condition that, for (u, v) in a certain square (or other region) in the $u\,v$ plane, the functions f, g, h possess continuous first partial derivatives such that the three jacobian determinants (9) are never

$$j_1 = \begin{vmatrix} \dfrac{\partial g}{\partial u} & \dfrac{\partial g}{\partial v} \\ \dfrac{\partial h}{\partial u} & \dfrac{\partial h}{\partial v} \end{vmatrix} \qquad j_2 = \begin{vmatrix} \dfrac{\partial h}{\partial u} & \dfrac{\partial h}{\partial v} \\ \dfrac{\partial f}{\partial u} & \dfrac{\partial f}{\partial v} \end{vmatrix} \qquad j_3 = \begin{vmatrix} \dfrac{\partial f}{\partial u} & \dfrac{\partial f}{\partial v} \\ \dfrac{\partial g}{\partial u} & \dfrac{\partial g}{\partial v} \end{vmatrix} \tag{9}$$

zero at the same point (u, v). Under these conditions a sufficiently small square containing a given point in the uv plane is "mapped" onto a smooth piece of surface, and (j_1, j_2, j_3) are components of a vector which is perpendicular to the plane tangent to the surface. The area A of this piece of surface is given by Eq. (10), where the double integral is extended

$$A = \int\int \sqrt{j_1^2 + j_2^2 + j_3^2}\, du\, dv \tag{10}$$

over the square in the uv plane. For a discussion of implicit functions and jacobian determinants *see* PARTIAL DIFFERENTIATION. *See also* ANALYTIC GEOMETRY; CALCULUS.

Angus E. Taylor

Paramo

A biological community, essentially a grassland, covering extensive high areas in equatorial mountains of the Western Hemisphere. Paramos occur in alpine regions above timberline and are controlled by a

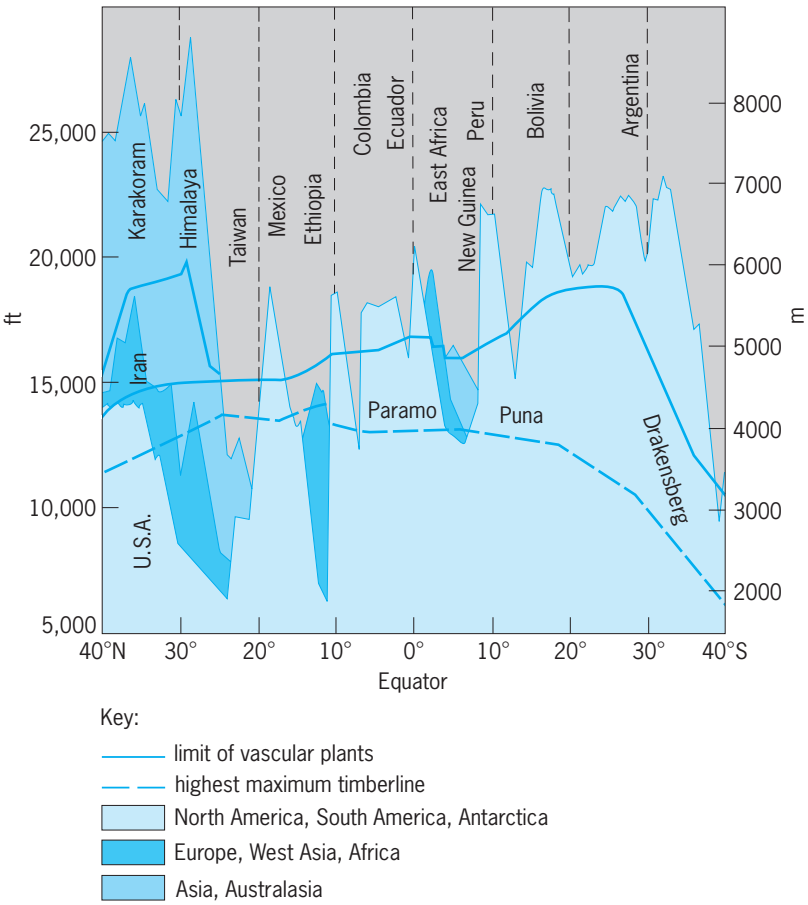

Fig. 1. Alpine and eolian regions of the tropics and subtropics, showing the broad extent of paramo areas. (*After H. E. Wright, Jr., and W. H. Osburn, eds., Arctic and Alpine Environments, Indiana University Press, 1968*)

complex of climatic and soil factors peculiar to mountains near the Equator. The richly diverse flora and the fauna of the paramos are adapted to severely cold, mostly wet conditions. Humans have found some paramos suitable for living and use. Since paramos were defined originally in Spanish-speaking countries, no English equivalent has been given to the term. *See* GRASSLAND ECOSYSTEM.

Geographic and altitudinal zones. Geographically, paramos are limited to the Northern Andes and adjacent mountains, where the climate is generally wetter at high altitudes than in the Central Andes, where related communities are called punas. The only paramo reported outside Colombia, Ecuador, and Venezuela is in Costa Rica. *See* PUNA.

In a broad sense paramos extend from the highest Andean forests to the upper altitudinal limit of vascular plants, the latter corresponding generally to permanent snowline (**Fig. 1**). Areas above this have been termed eolian since only airborne nutrients are available for nonvascular plants. In equatorial alpine regions, the timberline is often lower than the 13,000 ft (4000 m) of Fig. 1, with grassy paramos beginning several hundred feet below this. The lower portions form a transition zone with the forest, termed subparamo or paramillo. Thickets of shrubs and some trees, many in the composite and heath families, become fewer with increasing altitude.

Species composition. Typical paramo is dominated by large bunch grasses with narrow leaves, often appearing like a lush meadow, always green (**Fig. 2**), with the genera *Calamagrostis, Festuca*, and *Agrostis* having the most species. Perennial herbs with brilliant flowers, or growing in compact cushions, are interspersed between the grasses (**Fig. 3**). In Colombia, Venezuela, and northern Ecuador, the paramos are conspicuously dotted with many species of the genus *Espeletia*, family Compositae, reported also in Los Llanganati, a remote eastern area of Ecuador. These plants are known as frailejon (monk's cowl) because of top-heavy rosettes of large leaves which terminate thick stems and bear blooming heads resembling sunflowers, in some species near the ground, in others elevated to 20 ft (6 m) or more (**Fig. 4**). The occurrence in the same habitat of such composites, bunch grasses,

Fig. 2. Paramo de Chisaca, Cordillera Oriental, Colombia (altitude approximately 12,950 ft or 3950 m), showing large bunch grasses in foreground.

Fig. 3. Paramo del Ruiz, Cordillera Central, Colombia (altitude approximately 13,000 ft or 4000 m), showing compact cushions of *Plantago rigida* between grasses.

Fig. 4. Paramo del Cocuy, Cordillera Oriental, Colombia (altitude approximately 14,700 ft or 4500 m), showing dense stand of *Espeletia lopezii* (known as frailejon).

and cushion growths gives many paramos a unique appearance.

Above the typical paramo there exists a zone sometimes termed superparamo, consisting of scattered, low-growing vegetation on rocks and gravel areas. Bunch grass and frailejon are lacking; some perennials not seen readily below include low cushions such as *Draba*, many as dwarf shrubs such as *Hinterhubera*. The species and appearance of the vegetation make the higher levels near the snow seem more like tundra than true paramo. The limit of vascular plants in equatorial regions (paramos) is shown in Fig. 1 at lower level than in the low-latitude alpine regions farther from the Equator (punas). *See* PLANTS, LIFE FORMS OF; TUNDRA.

However, striking differences of paramo and puna, as compared with polar and temperate alpine tundras, include absence of those rhythms and organisms with seasonal adaptation to temporarily favorable growing conditions of summer, which partly offset some of the very severe winter conditions of high latitude and altitude. *See* ALTITUDINAL VEGETATION ZONES.

Climate and soil factors. On tropical mountains, variation in day- and night-length (which serve as signals for plants' seasonal changes) are much smaller than in other cold climates. Depending on altitude and local cloudiness, the incoming solar energy may warm plants and animals above the temperature of the surrounding air and soil. As a result of the outgoing thermal radiation, people and other organisms are chilled every night. It is not surprising therefore that paramos are considered inhospitable areas with penetrating cold, strong winds and heavy fog or mist which saturates air and soil. Early morning sun is usually lost when clouds sweep up from lower altitudes. Nights are always cold, often producing needle ice in the soil and snow in higher parts. Temperatures vary little throughout the year, but considerably diurnally. A range from 32°F (0°C), or slightly above, to around 68°F (20°C) is frequent in a 24-h period. Differences among paramos are more in degree of wetness, depending upon their exposure and position in the Andean ranges, with the eastern slope catching most of the moisture-laden winds. Soils in typical paramo are very black, deep, rich in organic matter, often peaty, frequently saturated even to being swampy, and quite acid.

Plant forms. Many paramo plants present xeromorphic features, as hairy or coriaceous leaves, densely compacted or overlapping in rosettes. The latter growth form occurs in paramo plants other than frailejon. It is also prominent in large *Lobelia* and *Senecio* species of East African mountains, where it is attributed to diurnal temperature changes similar to those mentioned above for the paramos.

Fauna and human habitation. A common paramo animal, evidenced by abundant scat, is the rabbit (*Lepus)*. Dantas or tapirs are hunted in some areas. Wolves, bear, elk, deer, bats, butterflies, bees, many beetles and birds, and one myriapod have been reported in Ecuadorean paramos. Small black and green frogs are abundant in swampy areas.

In Colombia and Venezuela, paramos are rather sparsely inhabited by humans, but they are used as pasture and the grasses are cut for roof thatching. In these countries the potato is the chief crop, but in Ecuador barley and wheat are grown by Quechua Indians up to nearly 13,200 ft (4000 m). The abundant frailejon is used in several ways, and is often injured by humans' repetitious burning of the paramos.

Harriet G. Barclay

Bibliography. H. G. Barclay, Human ecology of the paramos and the punas of the high Andes, *Proc. Okla. Acad. Sci.*, 43:13–34, 1963; J. H. Brown and A. C. Gibson, *Biogeography*, 2d ed., 1998; M. J. Coe, *Ecology of the Alpine Zone of Mount Kenya*, 1967; G. W. Prescott, Ridgepole of the Western Hemisphere, *Ward's Bull.*, 7(48):1–5, 1967; H. E. Wright, Jr., and W. H. Osburn (eds.), *Arctic and Alpine Environments*, 1968.

Paramyxovirus

A group of viruses that belong to the genus *Paramyxovirus* of the family Paramyxoviridae. The family includes two other genera, *Morbillivirus* and *Pneumovirus*. The genus *Parainfluenza* includes viruses such as mumps and Newcastle disease (in fowl and humans). Measles virus is a member of the genus *Morbillivirus*, and respiratory syncytial virus belongs to the genus *Pneumovirus*. Related members exist in nonhuman species; human measles virus is related to canine distemper and bovine rinderpest virus. Simian and bovine parainfluenza viruses also are known. *See* PARAINFLUENZA VIRUS; RESPIRATORY SYNCYTIAL VIRUS.

Like influenza viruses, the paramyxoviruses are ribonucleic acid (RNA)–containing viruses and possess an ether-sensitive lipoprotein envelope. They range in size from 100 to 300 nanometers, and sometimes larger pleomorphic forms are present. The ribonucleoprotein helix of the paramyxoviruses is 18 nm in diameter, twice as thick as that of influenza. *See* ANIMAL VIRUS; RIBONUCLEIC ACID (RNA).

The diseases caused by mumps and measles viruses are well known. The importance of the parainfluenza and respiratory syncytial viruses in causing severe respiratory disease in infants and young children was recognized chiefly after 1960. The most important respiratory tract pathogen of young infants is respiratory syncytial virus, especially in cases of bronchiolitis and pneumonia, and it is the only respiratory virus which has its greatest frequency of human infection in the first 6 months of life. Parainfluenza viruses are next in importance, particularly in cases of croup, but they more commonly infect older children rather than infants. The parainfluenza viruses also frequently cause common colds and other mild upper respiratory illnesses in children and adults. *See* MEASLES; MUMPS; NEWCASTLE DISEASE; PARAINFLUENZA VIRUS; VIRUS CLASSIFICATION. Joseph L. Melnick; M. E. Reichmann

Bibliography. D. W. Kingsbury (ed.), *The Paramyxoviruses*, 1990; D. O. White and F. J. Fenner, *Medical Virology*, 4th ed., 1994.

Paranoia

A mode of thought, feeling, and behavior characterized centrally by false persecutory beliefs, more specifically referred to as paranoidness. The paranoid individual falsely believes he or she is being intentionally harassed, plotted against, threatened, demeaned, interfered with, wronged, exploited, or otherwise persecuted by malevolent others—either a specific person or a group. Commonly associated with these core persecutory beliefs are properties of suspiciousness, fearfulness, hostility, hypersensitivity, rigidity of conviction, and an exaggerated sense of self-reference. These properties are evident with varying degrees of intensity and duration.

Paranoidness represents a specific mode of processing semantic information that has triggering and structuring causes. Triggering causes determine when the paranoid mode is activated; structuring causes are those specific semantic contents that produce the distinctive characteristics of the paranoid class. The historical factors that lead to the formation of the paranoid mode are not well understood.

The paranoid mode can be triggered at either biological or psychological levels. Common precipitating biological causes are brain trauma or tumor, thyroid disorder, cerebral arteriosclerosis, and intoxication with certain drugs, including alcohol, amphetamines, cocaine, other psychostimulants, and hallucinogens such as mescaline or lysergic acid diethylamide (LSD). They can produce disordered activity of central dopaminergic and noradrenergic pathways. At the psychological level, triggering causes include false arrest, birth of a deformed child, social isolation, deafness, and intensely humiliating experiences. *See* NORADRENERGIC SYSTEM.

The paranoid mode is resistant to modification by psychotherapeutic or pharmacological methods. Acute psychotic states of paranoidness accompanied by high levels of anxiety are usually responsive to neuroleptic medication. *See* PSYCHOPHARMACOLOGY.

Paranoidness can occur by itself, or it can be associated with other mental conditions such as depression or schizophrenia, and its course can be quite variable. In some cases it may persist or fluctuate for years, whereas in others it subsides inexplicably. *See* SCHIZOPHRENIA. Kenneth Mark Colby

Bibliography. K. M. Colby, Modeling a paranoid mind, *Behav. Brain Sci.*, 4:515–560, 1981; D. W. Swanson, P. J. Bohnert, and J. A. Smith, *The Paranoid*, 1970.

Parasexual cycle

A series of events, discovered in filamentous fungi, which lead to genetic recombination outside the standard sexual cycle.

Until the 1940s the standard sexual cycle (an alternation of nuclear fusion and meiosis) was the only route to formal genetic analysis and planned breeding. Since that time diverse alternative processes leading to recombination have been discovered in microorganisms. The processes occurring in certain filamentous fungi were termed the parasexual cycle by Guido Pontecorvo. There are three essential steps in the cycle: heterokaryosis; fusion of unlike haploid nuclei in the heterokaryon to yield heterozygous diploid nuclei; and recombination and segregation at mitosis by two independent processes, mitotic crossing-over and haploidization.

The features and operation of the cycle can be outlined in the homothallic ascomycete *Aspergillus nidulans*, the species in which the cycle was first demonstrated. Wild-type *A. nidulans* grows on a minimal medium containing a sugar and inorganic salts, producing green, haploid, uninucleate, vegetative spores (conidia). It is possible to produce mutant strains differing from wild type in various characteristics such as conidial color or growth requirement for an exogenous supply of one or other essential nutrient. Heterokaryons are synthesized by growing together pairs of mutant strains which differ from wild type and each other in nutritional requirements and, preferably, conidial color. Hyphal fusions yield some heterokaryotic cells which carry genetically different nuclei in a common cytoplasm. The heterokaryon, somewhat comparable with a heterozygote, is nutritionally wild type; it can be grown in balanced form on minimal medium which does not permit growth of either parent homokaryon. Most of the conidia formed by a heterokaryon are of one or the other parental haploid type. However, infrequent fusions of unlike nuclei in the hyphae give rare conidia with heterozygous diploid nuclei. Diploid strains are isolated by plating conidia from a heterokaryon on minimal medium which selects only the heterozygotes. The resulting strains are fairly stable at mitosis. They are recognized mainly by their phenotypes, as compared with the parent haploids; their larger conidial diameter; and their ability to show mitotic segregation.

Mitotic crossing-over was first discovered by C. Stern in *Drosophila* and is a rare but regular occurrence in a small percentage of all mitotic divisions at the four-strand stage of mitosis in diploid nuclei. When such crossing-over occurs, there is usually only one exchange, reciprocal between nonsister chromatids, in the whole genome. The subsequent chromatid segregation is mitotic and may yield a nucleus, still diploid, homozygous for all alleles linked in coupling and distal to the exchange. All other loci which are heterozygous remain heterozygous. The modalities of this process permit the ordering of genes on a linkage group arm relative to each other and to their centromere. Mitotic haploidization is independent of mitotic crossing-over and its results are, in effect, like those of meiosis without crossing-over. Haploidization is usually, perhaps always, a stepwise process in which chromosomes are lost successively until a stable haploid state is reached. Haploidization permits the ready assigning of a gene to its linkage group and permits association of linkage group arms mapped by mitotic crossing-over. Detection of mitotic segregants depended initially on visual markers (such as conidial color) revealed in homozygous mutant or haploid condition. Many devices are now available for the automatic selection of mitotic segregants.

Some or all of the features of the parasexual cycle have been demonstrated in many species of fungi. Diploid strains find use in formal genetic analysis and in biochemical research. Mitotic segregation is used for genetic analysis both in species which have a sexual cycle and in imperfect species which have no known sexual cycle. In the former, mitotic analysis is a valuable additional analytical tool; in the latter, it provides the only approach to formal genetic analysis and planned breeding.

Processes similar to the elements of the parasexual cycle have been exploited in the somatic cell genetics of higher plants and animals where, as in fungi, they may overcome limitations of sexual processes. Amphiploids of *Nicotiana* have been produced by selection of the somatic hybrids formed in a mixture of protoplasts from two species; this technique may open the way to the combination of very divergent plant genomes. Techniques are also available for the fusion of cultured animal cells, within and between species. Intraspecific hybrids have a potential in formal genetic analysis and in, for example, tests of allelism of phenotypically indistinguishable conditions determined by recessive mutations. Interspecific hybrids, such as those of mouse and humans, show a sequential loss of whole chromosomes, usually preferentially those of one particular component species; cytological analysis and phenotypic tests of resulting somatic variants provide a means, somewhat parallel to that of haploidization in fungi, for allocation of a gene to its chromosome. *See* BREEDING (PLANT); FUNGI; GENETICS; RECOMBINATION (GENETICS); SOMATIC CELL GENETICS. J. Alan Roper

Bibliography. G. C. Ainsworth and A. S. Sussman (eds.), *The Fungi*, vol. 2, 1966; M. J. Carlile and S. Watkinson, *The Fungi*, 1994; E. Moore-Landecker, *Fundamentals of the Fungi*, 4th ed., 1996; G. Pontecorvo, Alternatives to sex: Genetics by means of somatic cells, in *2d John Innes Symposium*, Oxford, 1975.

Parasitology

The scientific study of parasites and of parasitism. Parasitism is a type of symbiosis and is defined as an intimate association between an organism (parasite) and another, larger species of organism (host) upon which the parasite is metabolically dependent. Implicit in this definition is the concept that the host is harmed, while the parasite benefits from the association. Since well-adapted parasites may be nonpathogenic, parasitism merges into another type of symbiosis in which both partners benefit, that is, mutualism. Parasitism also blends with another ecological relationship, predation. For example, ichneumon fly larvae, which ultimately kill their caterpillar hosts by internal consumption, are indeed predators but are called parasitoids, reflecting the similarity of their life-style to true parasitism. By the definition, above, most species of organisms may properly be called parasites. Although technically parasites, pathogenic bacteria and viruses, fungal, and insect parasites of plants are traditionally outside the field of parasitology.

Parasites often cause important diseases of humans, wildlife, livestock, and pets. For this reason, parasitology is an active field of study; advances in biotechnology have raised expectations for the development of new drugs, vaccines, and other control measures. However, these expectations are dampened by the inherent complexity of parasites and host–parasite relationships, the entrenchment of parasites and vectors in their environments, and the vast socioeconomic problems in the geographical areas where parasites are most prevalent. Therefore, most parasitologists believe that parasite control or eradication will be a slow and expensive process.

Parasite–host relationships. The ecological and physiological relationships between parasites and their hosts are multifaceted and constitute some of the most impressive examples of biological adaptation known. Much of classical parasitology has been devoted to the elucidation of one of the most important aspects of host–parasite ecological relationships, namely, the dispersion and the transmission of parasites to new hosts.

Parasite life cycles range from simple to highly complex. Simple life cycles (transmission from animal to animal) are direct and horizontal with adaptations that include high reproduction rates, and the production of relatively inactive stages (cysts or eggs) that are resistant to environmental factors such as desiccation, ultraviolet radiation, and extreme temperatures. The infective stages are passively consumed when food or water is contaminated with feces that contain cysts. The cysts are then activated in the gut by cues such as acidity to continue their development. Other direct-transmission parasites, such

as hookworms, actively invade new hosts by penetrating the skin. Physiologically more complicated are those life cycles that are direct and vertical, with transmission being from mother to offspring. The main adaptation of the parasite for this type of life cycle is the ability to gain access to the fetus or young animal through the ovaries, placenta, or mammary glands of the mother.

Many parasites have taken advantage of the food chain of free-living animals for transmission to new hosts. During their life cycle, these parasites have intermediate hosts that are the normal prey of their final hosts. Parasites may ascend the food chain by utilizing a succession of progressively larger hosts. This process, called paratenesis, may have been helped in its evolution by the common habit among carnivores of opportunistic scavenging. Another adaptation to food-chain transmission is for parasites to physically impede or alter the behavior of intermediate hosts, thus raising their chances of being caught by the final host. Parasites achieve this by growing in the host's central nervous system or by producing substances that affect host behavior or physiology. This group includes serious pathogens such as the neurotropic larval tapeworms that cause cysticercosis, gid, and hydatid disease in humans as well as animals. *See* FOOD WEB.

Vectors are intermediate hosts that are not eaten by the final host, but rather serve as factories for the production of more parasites and may even carry them to new hosts or to new environments frequented by potential hosts. Blood-sucking arthropods such as mosquitoes and tsetse flies are well-known examples. After acquiring the parasite from an infected host, they move to another host, which they bite and infect. Their efficiency as vectors is enhanced if they permit parasite multiplication internally. In aquatic environments, the role of vector is often taken by snails or blood-sucking leeches. Snails are important vectors for two-host trematodes (flukes), which increase their numbers greatly in the snail by asexual reproduction. The stages that leave the snail may either infect second intermediate hosts that are eaten by carnivorous final hosts, may encyst on vegetation that is eaten by herbivorous hosts, or in the case of the blood flukes (schistosomes) may swim to and directly penetrate the final host.

Parasite–parasite relationships. Besides their ecological relationship with their hosts, parasites interact with other parasites of their own or other species. Many parasites are monoecious (hermaphroditic), allowing self-fertilization, which is an obvious advantage when only one parasite has colonized a host. In dioecious parasites, the opposite sexes must find each other in or on the body of the host. This is facilitated by site specificity in the host and a high density of parasites, but much evidence implicates more refined mechanisms such as pheromones. If crowding may sometimes facilitate sexual contact, it may also be inimical when parasites must share limited resources. Some parasites have evolved mechanisms to limit intraspecific competition. The human pork tapeworm (*Taenia solium*) appears to have a preemptive mechanism against colonization by others of its kind, either through direct chemical action of one worm on another or through the mediation of the host's immune system. Many worm infections exhibit concomitant immunity, whereby the primary infection sensitizes the host to resist subsequent reinfection by the same species while the primary infection worm is spared. *See* PHEROMONE.

Parasites must also compete interspecifically: seldom is a host infected with just one species of parasite. A variety of parasites can be found in certain niches in the host, such as the intestine. Two coexisting species of intestinal worms may adjust to each other's presence by shifting their locations to partially overlapping segments of the bowel, even though in single infections each would have inhabited the identical site. The several species of intestinal coccidia of chickens have become specialized to occupy relatively discrete intestinal regions.

Parasites with free-living transmission stages interact ecologically with other free-living organisms. They are subject to predation, against which they compensate with very high reproductive rates.

Physiologic interactions. The physiologic interactions of parasites with their hosts are as complex and varied as the ecological interactions. Metabolic dependency is the key to parasitism, and parasites employ many ways to feed off their hosts. The simplest is exhibited by the common intestinal roundworm, *Ascaris*, which consumes the host's intestinal contents. Parasites require from their hosts not only energy-yielding molecules but also basic monomers for macromolecular synthesis and essential cofactors for these synthetic processes. Many examples of the specific absence of key parts of energy-yielding or biosynthetic pathways in parasites are known, and these missing enzymes, cofactors, or intermediates are supplied by the host. Theoretically, differences in metabolism between parasite and host are the targets for rational chemotherapy, but these differences are largely useless for this purpose, since the parasite basically parasitizes an already existent host pathway. Other differences must and are being identified for targeted chemotherapy.

Tapeworms are more complex than *Ascaris* in nutritional requirements from the host. They lack a gut, but their surface actively takes up, by facilitated diffusion or active transport, small molecules such as amino acids and simple sugars. They may actually outcompete the host's intestine, as is dramatically illustrated by the broad tapeworm of humans, *Diphyllobothrium latum*, which absorbs vitamin B_{12} so avidly that the host may develop pernicious anemia. The vertebrate intestine is an organ that regulates the concentrations of small molecules at its mucosal surface. But tapeworms do not regulate the uptake of these small organic molecules very well, relying on the host to provide the optimal mix, to which the tapeworm becomes adapted. This has led to the theory that parasites may gain more than just nutrients from the host by also parasitizing host regulatory mechanisms.

Genetics. Because of the obvious absence of organs relating to sensory perception, motility, and digestion, and also because of reductions in biosynthetic and bioregulatory capabilities, parasites have been described as degenerate organisms. In studies comparing genome size of certain parasites with their nonparasitic relatives, the genome size of the parasites was not reduced, suggesting that parasites are not degenerate but instead are highly adapted organisms. The apparent degeneracy of parasites is matched by special requirements for their success. For example, consider a parasite that alternates between a poikilothermic invertebrate host and a homeothermic vertebrate host, which is a common pattern. Different life cycle stages are genetically programmed to respond at the proper times to quite different developmental cues supplied by disparate host physiologies. For example, many parasites are now known to have heat-shock protein genes. The heat-shock response to elevated temperatures could be the signal for a parasite to switch to an alternate metabolism upon infecting a warm-blooded animal. If free-living stages are involved in the parasite's life cycle, then complex host-finding behavior and invasive mechanisms may be required for successful infection, all of which must be encoded in the deoxyribonucleic acid (DNA).

Immunology. The special problems of parasitic existence are well illustrated by the fact that parasites survive and thrive, as foreign invaders, in the bodies of animals with highly evolved immune defense systems. This system's efficiency, which normally protects animals from bacteria and viruses, is underscored by the dire consequences of its failure in congenital or acquired immunodeficiency syndromes. Parasites, by coevolving with their hosts, have the ability to evade the immune response. The best-known evasive tactic is antigenic variation, as found in African trypanosomes, which have a complicated genetic mechanism for producing alternative forms of a glycoprotein that virtually cover the entire parasite. Each type of variant surface glycoprotein is antigenic, stimulating the synthesis of host blood proteins called antibodies that specifically bind to it and no other variant surface glycoprotein. Antibody binding is followed by the trypanosome's death and clearance from the blood. By going through a genetically programmed (but largely unpredictable) sequence of variant surface glycoproteins, the trypanosome population in a host stays one step ahead of immunity and is not eliminated. Other possible immune escape mechanisms in parasites have been discovered and probably cooperate to prolong parasite survival. In the first place, parasites may avoid stimulation of the immune response by poorly expressing antigens, by localizing in sites not well serviced by the immune system, or by synthesizing antigens that resemble those of the host (molecular mimicry), taking advantage of the host's built-in safeguards against autoimmunity. Second, if the immune system is specifically stimulated by parasite antigens, the parasites may avoid its consequences by locally degrading antibodies and other effector molecules. Third, they may redirect the immune response into nonprotective channels by interfering with its normal control processes, which could account for the fact that many of the antibodies and effector cells produced in response to parasitic infection are ineffective and even irrelevant against the parasites.

This is not to say that parasites are altogether exempt from the effects of immunity. Rather than completely eliminating parasites, the immune system more often functions to control their populations in the host. Thus a balance is achieved between hosts and parasites that have lived in long evolutionary association, with both surviving through compromise. Enhancing these particular antiparasite mechanisms and neutralizing the parasite's evasion mechanisms would tip the balance in favor of the host.

The adaptation of parasites to the host immune response has gone even beyond its evasion. Some, perhaps many, parasites may exploit it to their own advantage. Immune exploitation may take several forms: the formation of protective capsules, the movement of parasites or their products through tissues, and the use of host immune molecules as nutrient sources or growth factors. Thus, to label parasites as degenerate organisms is to ignore their intricate adaptations to their way of life. *See* MEDICAL PARASITOLOGY; POPULATION ECOLOGY.

Raymond T. Damian

Bibliography. R. T. Damian, The exploitation of host immune responses by parasites, *J. Parasit.*, 73:3–13, 1987; P. T. Englund and A. Sher (eds.), *The Biology of Parasitism: A Molecular and Cellular Approach*, 1988; E. R. Noble et al., *Parasitology: The Biology of Animal Parasites*, 6th ed., 1989; D. Rollinson and R. M. Anderson (eds.), *Ecology and Genetics of Host-Parasite Interactions*, 1985; G. D. Schmidt and L. S. Roberts, *Foundations of Parasitology*, 5th ed., 1995.

Parasympathetic nervous system

A portion of the autonomic system. It consists of two neuron chains, but differs from the sympathetic nervous system in that the first neuron has a long axon and synapses with the second neuron near or in the organ innervated. In general, its action is in opposition to that of the sympathetic nervous system, which is the other part of the autonomic system. It cannot be said that one system, the sympathetic, always has an excitatory role and the other, the parasympathetic, an inhibitory role; the situation depends on the organ in question. However, it may be said that the sympathetic system, by altering the level at which various organs function, enables the body to rise to emergency demands encountered in flight, combat, pursuit, and pain. The parasympathetic system appears to be in control during such pleasant periods as digestion and rest. The alkaloid pilocarpine excites parasympathetic activity while atropine inhibits it. *See* SYMPATHETIC NERVOUS SYSTEM.

Because the parasympathetic transmission of efferent impulses is restricted to cranial nerves III

(oculomotor), VII (facial), IX (glossopharyngeal), and X (vagus), and to sacral segments 1, 2, 3, and 4, it is also known as the craniosacral system.

Results of stimulation of the parasympathetic components are listed for the following nerves:

1. Oculomotor (III) cranial nerve: eye accommodation in near vision, constriction of the pupil.
2. Facial (VII) cranial nerve: secretion, vasodilation, and constriction of ducts in the submaxillary and sublingual salivary glands and in the lacrimal, nasal, and buccal glands.
3. Glossopharyngeal (IX) cranial nerve: secretion, vasodilation, and constriction of ducts in parotid glands.
4. Vagus (X) cranial nerve: decrease in heart rate, coronary constriction, tracheobronchial constriction, increased smooth-muscle mobility in esophagus, gastrointestinal secretion, internal and external pancreas reduction, increase in gastrointestinal motility, inhibition of ileocolic sphincter, and constriction of pancreatic and biliary ducts.
5. Sacral: evacuation of rectum and bladder, pelvic vasodilation, erection, and secretion of genital tract glands (prostate and Cowper's).

In birds and reptiles the parasympathetic system is very similar to that in mammals. In amphibians and jawed fishes, only the parasympathetic elements of the oculomotor (III) and vagus (X) nerves are distinct. In agnathans (lampreys and hagfish) there is no direct parasympathetic system, but the vagus (X) nerve does innervate the alimentary canal. *See* AUTONOMIC NERVOUS SYSTEM; SPINAL CORD.

Douglas B. Webster

Parathyroid gland

An endocrine organ usually associated with the thyroid gland. Embryologically, the glandular primordia are formed in the endoderm of the third and fourth pharyngeal pouches and are associated with the similarly derived primordia of the thymus. There may be from one to three pairs of the small glands present in individuals of the various vertebrate classes, although two pairs appear most frequently. They are characteristically within, on, or near the thyroid gland. In response to lowered serum calcium concentration, a hormone is produced which promotes bone destruction and inhibits the phosphorus-conserving activity of the kidneys. *See* THYROID GLAND.

Anatomy. Parathyroid glands occur in all vertebrate classes with the exception of the fishes, although cells that appear to be homologs of parathyroid cells are found in cyclostomes at the dorsal and ventral ends of all pharyngeal pouches. Whether they function as parathyroid tissue, however, is unknown. Data relating to the details of the comparative anatomy of these glands are summarized in **Table 1**.

In humans, there are typically four glands situated as shown in **Fig. 1**; however, the number varies between three and six, with four appearing about 80% of the time. Variations in the positioning of the glands (or more properly, glandules) along the craniocaudal axis occur but, excepting parathyroid III which may occasionally be found upon the anterior surface of the trachea, the relation to the posterior surface of the thyroid is rarely lost. The extent of variation in primates generally is greater than that observed in all other mammals.

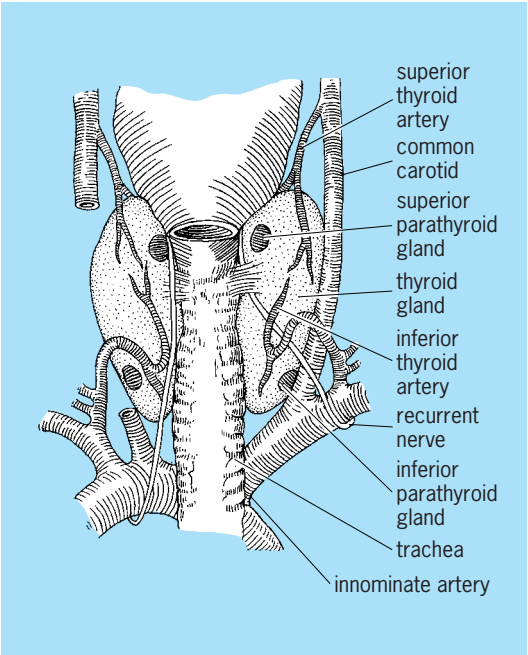

Fig. 1. Common positions of human parathyroid glands on the posterior aspect of the thyroid. (*After W. H. Hollinshead, Anatomy of the endocrine glands, Surg. Clin. N. Amer., 32(4):1115–1140, 1952*)

Macroscopically, the glandules appear as oval, round, or irregularly shaped objects, generally having a smooth, highly vascularized surface, although the latter may be slightly roughened (horse) or finely serrated (cattle, swine). Both size and weight are quite variable and appear to have no apparent correlation with total body weight. The color varies from reddish brown through yellowish pink to grayish red or white, depending upon the proportion of fat cells to secreting cells and the amount of lipid in the secreting cells.

Histology. The histological structure of the parathyroids is based upon nonsecretory supportive or stromal elements and secreting or parenchymal elements.

The glandules are enclosed in a smooth connective tissue capsule from which trabeculae, bearing lymphatics, blood vessels, and nerves, emerge to penetrate the gland and to divide its substance into a roughly lobular pattern. The blood vessels are connected to a rich capillary network, characteristic of the endocrine glands generally, that is entwined in reticular fibers found among the parenchymal cells. In addition, fat cells also occur and increase in number with advancing age.

Parenchymal cells are arranged in interconnecting cords and clumps of cells as well as in occasional follicular masses enclosing a gelatinous protein material

TABLE 1. Comparative anatomy of parathyroid glands

Vertebrate class, order, and examples	Number of glandules, location, and relations*
Amphibia	
Urodeles	
Salamanders	2–3 pairs of glandules separate from thyroid, ventral to thymus, lateral to arterial arches
Anura	
Frogs, toads	2–3 pairs of glandules lie laterally next to external jugular vein; gill remnant is immediately dorsal in frog, just ventral in toad
Reptilia	
Saurians (suborder)	
Lizards	1 pair of glandules, III only; usually attached to thymus
Serpentes (suborder)	
Snakes	2 pairs of glandules which may be II and III or III and IV (probably latter); if so, III is near temporomaxillary joint next to bifurcation of CC; IV lies between thymus III and IV
Chelonia	
Turtles	2 pairs of glandules, III at cranial end, IV at caudal end of thymus
Crocodylia	
Crocodiles	1 pair of glandules, III only; next to thymus, near origin of collateral cervical artery
Aves	
Galliformes	
Fowl (chicken)	2 pairs of glandules, III at inferior pole of thyroid, IV next to III; when fused, connection may be more or less tenuous
Others	In birds other than fowl, there is wide variation in number and position; it is usually associated with thyroid or CC in some way
Mammalia	
Monotremes	
Platypus	3 pairs of glandules, II at bifurcation of CC, III at dorsal end of thymus, IV on lateral surface of thyroid
Marsupialia	
Opossum	2 pairs of glandules, III medial of branching of CC, IV connected with or embedded in thymus IV; also found between the thyroid and CC
Insectivores	
Moles	1 pair of glandules, III near middle CC
Chiroptera	
Bats	2 pairs of glandules, III (larger) on dorsal outer surface of thyroid, IV (smaller) on inner (tracheal) surface of thyroid
Primates	
Monkeys, apes, humans	2 pairs of glandules; in monkeys and apes, extent of variation in number and position exceeds that observed in other mammals; generally, an external pair (III) related to thymus; an internal pair buried in the thyroid
Lagomorpha (rabbits)	2 pairs of glandules, III associated with thyroid, either at superior or inferior pole of lateral lobes, free or sunken into a depression on thyroid; often more closely related to CC; IV at middle of lateral lobe of thyroid; exceptionally at inferior pole
Rodentia	
Muridae (rats, mice)	1 pair of glandules, III only; usually in cavity on lateral surface of thyroid; cavity usually more superficial in mouse than rat
Caviidae (guinea pig)	2 pairs of glandules, III usually removed from thyroid, sometimes vestigial; IV is embedded in thyroid
Carnivores	
Canidae (dogs)	2 pairs of glandules, III on or in front of lateral cranial surface of thyroid; sometimes more or less embedded in surface; IV is small, deeply buried in thyroid; position variable
Felidae (cats)	2 pairs of glandules, III similar to dog; IV near middle of thyroid nearer tracheal than outer surface
Perissodactyla (horse)	2 pairs of glandules, III near (within 1 cm or 0.4 in.), on, or partially embedded in upper medial border of thyroid; occasionally on tracheal surface; IV entirely embedded in thyroid
Artiodactyla	
Suidae (swine)	1 pair of glandules, III only; usually well removed from thyroid, either in thymus or with terminal branches of CC
Bovidae	
Sheep, goats	2 pairs of glandules, III near bifurcation of CC; IV may be either free on surface of thyroid or deeply embedded
Cattle	2 pairs of glandules, III near bifurcation of CC; IV usually lies on inner tracheal surface of thyroid

* Positions given are subject to extensive variation. III or IV is used to designate glandules according to pharyngeal site of origin. In mammals, III is external to the thyroid; IV, to a greater or lesser extent, within its substance. CC is the common carotid artery.

called colloid; the latter resembles thyroid colloid in appearance but does not contain iodine and appears to have no comparable physiological activity.

There are two general types of parenchymal cells (**Fig. 2**): chief, or principal, cells, which occur in the parathyroids of all vertebrates; and oxyphil cells, which are found in humans only after the fourth to seventh year, and in older macaque monkeys and cattle. These observations support the belief that the parathyroid hormone is produced by the chief cells.

Chief cells. These occur in so-called light or dark forms, according to their appearance under the light-microscope; chief cells are generally smaller than oxyphils, are polygonal in outline, and contain a weakly to moderately acidophilic cytoplasm.

Dark chief cells are thought to be active in hormone production and are distinguished by numerous 100–400-nanometer cytoplasmic granules which are distinctively stained by iron and chrome hematoxylin, Bodian silver stain, and aldehyde fuchsin;

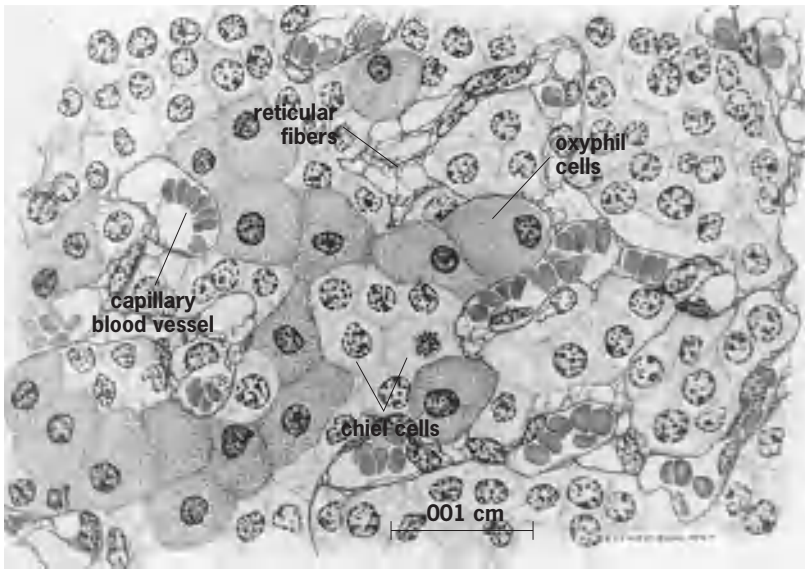

Fig. 2. Section of human parathyroid gland, demonstrating the two main parenchymal cell types. Zenker-formal fixation, Mallory-azan stain. (*After D. Bloom and D. W. Fawcett, A Textbook of Histology, 9th ed., Saunders, 1968*)

nuclei are larger and somewhat more vesicular than their light counterparts. Electron microscopy reveals that the granules, which are membrane-bound aggregates of particles about 10–20 nm, initially appear next to a prominent Golgi apparatus and then pass through the cytoplasmic matrix to be released through the plasma membrane; the particles eventually enter the circulation after transiting the cytoplasm of capillary endothelial cells. The cytoplasm contains sparse amounts of glycogen, and profiles of the cisternae of the granular endoplasmic reticulum occur in parallel arrays.

Light chief cells, which predominate, are thought to be an inactive or nonsecreting phase of a secretory cycle in which the dark cell is the active participant. The cytoplasm of these cells is less chromophilic and less granular than the dark variety, and the nucleus is smaller and denser. Fine structure is characterized by a sharply reduced content of secretory granules, a reduction in the extent of the Golgi apparatus and of the granular reticulum, and a marked increase in glycogen content. Significantly, there are no granules near the Golgi apparatus, and those remaining are at the periphery of the cell.

Oxyphil cells. These are larger than chief cells and have a smooth, rather rounded contour, a strongly acidophilic granular cytoplasm, and a small dark-staining nucleus which may be pycnotic; light and dark forms are said to exist. Oxyphils form only a small part of the cell population and occur singly or in small clusters. Their fine structure is remarkable in that the cytoplasm is literally packed with mitochondria to the virtual exclusion of other organelles; secretory granules occur rarely. The cell has a demonstrable oxidative enzymatic activity, but no definitive function has been discovered; a small body of evidence suggests that some cells may produce hormones. Cells intermediate in structure between chief and oxyphil cells have been found, lending support to the belief that they might be participants in the secretory cycle of the gland.

Embryology. Developmentally, the parathyroids of all vertebrates originate in the endodermal lining of certain pharyngeal pouches. However, differences in details of the site of origin occur among the vertebrate classes (**Table 2**).

TABLE 2. Comparative embryology of parathyroid glands

Vertebrate class, order, and examples	Location of parathyroid primordia: Pharyngeal pouch*	Location of parathyroid primordia: Portion of pouch
Pisces		
Cyclostomes		
Lamprey	?	
Chondrichthyes		
Sharks, skates, and rays	Not known to occur	
Osteichthyes		
Perch, trout	Not known to occur	
Amphibia		
Urodeles		
Salamanders	II(1),III,IV	Ventral
Anura		
Frogs, toads	II(3),III,IV	Ventral
Reptilia		
Sauria (suborder)		
Lizards	II†,III,IV†	Ventral
Serpentes (suborder)		
Snakes	II(?),III,IV,V†	Ventral
Chelonia		
Turtles	III,IV	Ventral
Crocodilia		
Crocodiles	III	Ventral
Aves		
Galliformes		
Fowl (chicken)	III,IV,V(2)	Ventral
Others	III,IV	Ventral
Mammalia		
Monotremes		
Platypus	II,III,IV	Ventral
Marsupialia		
Opossum	III,IV	Dorsal
Insectivores		
Moles	III,IV(1)	Dorsal
Chiroptera		
Bats	III,IV	Dorsal
Primates		
Monkeys, apes, humans	III,IV	Dorsal
Lagomorpha (rabbits)	III,IV	Dorsal
Rodentia		
Muridae (rats, mice)	III,IV†	Dorsal
Caviidae (guinea pig)	III,IV	Dorsal
Carnivores		
Canidae (dogs)	III,IV	Dorsal
Felidae (cats)	III,IV	Dorsal
Perissodactyla (horse)	III,IV	Dorsal
Artiodactyla		
Suidae (swine)	III,IV†	Dorsal
Bovidae		
Sheep, goats	III,IV	Dorsal
Cattle	III,IV	Dorsal

*Numbers in parentheses indicate frequency of occurrence: 1, frequent; 2, occasional; 3, seldom.
†Parathyroid primordium is either not formed or undergoes involution shortly after formation.

The broader aspects of parathyroid development in the various classes are comparable and differ only in detail; the formation in man is given as a representative example. At 35–37 days of development, the endodermal cells that are destined to become the secreting cells of the gland enlarge, become less acidophilic, and aggregate into clumps of cells that constitute the glandular primordia. One primordium is established on the dorsolateral surface of each of the paired third and fourth pharyngeal pouches. Unlike the primordia of the fourth pouches, those of the third are intimately associated with thymic primordia (thymus III). Parathyroid III accompanies thymus III as the latter migrates caudally. When the inferior border of the developing thyroid is reached, the parathyroid primordium is released, while thymus III continues its migration. In consequence of this movement, parathyroid III comes to lie inferior to parathyroid IV. Occasionally fragmentation of parathyroid III occurs during migration, the fragments later developing into isolated masses of functional tissue, accessory parathyroids. For this reason it is often difficult or may even be impossible to render an animal entirely free of parathyroid tissue by surgical means.

During the fourth to fifth months, chief cells appear and a sinusoidal circulatory pattern is established among the interconnecting cords of cells. The onset of functional activity probably coincides with the appearance of this histological pattern. Variations in size, number, and position of matured glands are known to occur, but complete absence is quite rare and is usually associated with other abnormalities of such severity as to ensure fetal or neonatal death. William E. Dossel

Physiology. The parathyroid glands are essential for the regulation of calcium and phosphate concentrations in the extracellular fluids of amphibians and higher vertebrates, but are not present in fishes. Calcium values must be maintained within a small range in order for cells to function normally. Removal or decreased function of the parathyroid glands results in hypocalcemia, increased irritability of nerves, uncontrolled muscle contractions (tetany), and sometimes death. Alternatively, excessive amounts of parathyroid hormone cause increased removal of calcium from bone, resulting in hypercalcemia and abnormal deposition of calcium and phosphate in soft tissues. *See* PARATHYROID HORMONE.

In the normal animal, calcium in the blood and extracellular fluids is regulated at a concentration ranging from 9 to 11 mg/100 ml (mg %). The secretion of parathyroid hormone is controlled directly by the calcium content of the blood. Parathyroid hormone output is stimulated when circulating calcium concentrations decrease below normal; in response to the hormone, blood calcium values will increase. Conversely, parathyroid hormone secretion decreases when calcium values increase above the normal range. This will result in a gradual fall of calcium concentrations back to normal. *See* CALCIUM METABOLISM.

In addition to raising blood calcium values, parathyroid hormone increases the excretion of phosphate in the urine. This urinary loss of phosphate results indirectly in decreased values of blood phosphate. Although phosphate does not appear to be regulated as precisely as calcium, this renal mechanism for elimination of phosphate serves to maintain a balance between phosphate and calcium in the extracellular fluids. Another important effect of parathyroid hormone on the kidney is to increase the renal reabsorption of calcium, thus reducing the loss of calcium in the urine and thereby conserving calcium in the body.

In fishes, some of the calcium-regulating function of parathyroid function appears to be carried out in fishes by calcitonin, a hormone secreted by the ultimobranchial glands.

Mechanism of action. Parathyroid hormone has two major target organs, bone and kidney. It acts on bone in several ways with both short-term and long-term effects. Short-term changes include a rapid uptake of bone fluid calcium into osteoblast cells, which in turn pump the calcium into the extracellular fluids. Long-term effects of parathyroid hormone include stimulating the activity and increasing the number of osteoclasts, bone cells which act to break down bone matrix and release calcium from bone. All of these effects will result in increased blood calcium values. *See* BONE.

Parathyroid hormone also inhibits the renal reabsorption of phosphate, thus increasing the urinary output of phosphate. Phosphate reabsorption across the renal tubule is dependent upon sodium transport, and parathyroid hormone interferes with this sodium-dependent phosphate transport in the proximal tubule. *See* KIDNEY.

There are reports that parathyroid hormone stimulates calcium uptake into the body across the intestine. However, this is not a direct effect. Parathyroid hormone stimulates the production of the most active metabolite of vitamin D, 1,25-dihydroxycholecalciferol, during vitamin D synthesis. This metabolite of vitamin D directly stimulates the intestinal absorption of calcium. *See* ENDOCRINE SYSTEM (VERTEBRATE); PARATHYROID GLAND DISORDERS; VITAMIN D. Nancy B. Clark

Bibliography. S. A. Binkley, *Endocrinology*, 1994; P. Felig, J. D. Baxter, and L. A. Frohman, *Endocrinology and Metabolism*, 3d ed., 1994; A. Gorbman et al., *Comparative Endocrinology*, 1983; M. E. Hadley, *Endocrinology*, 5th ed., 1999; P. K. Pang and A. Epple, *Evolution of Vertebrate Endocrine Systems*, 1980; J. A. Parsons (ed.), *Endocrinology of Calcium Metabolism*, 1982.

Parathyroid gland disorders

Disorders involving excessive or deficient blood levels of parathyroid hormone caused by abnormal functioning of the parathyroid gland. Parathyroid hormone is responsible for keeping the concentration of calcium in blood within a narrow normal

range. If the blood calcium concentration falls, the parathyroid glands respond by secreting hormone which tends to increase the concentration of calcium. Conversely, an increase in blood calcium concentration above the normal range normally suppresses parathyroid hormone secretion. Parathyroid hormone acts directly on bone and kidney, and indirectly on the intestine, to increase the concentration of calcium in blood. It also acts on the kidney to increase excretion of phosphate in the urine, causing a lowering of the concentration of phosphorus in blood.

Hyperparathyroidism. This group of disorders is characterized by excessive parathyroid hormone secretion. Primary hyperparathyroidism is defined as increased parathyroid hormone secretion despite elevated blood calcium. The disorder is readily diagnosed by demonstration of increased blood calcium and parathyroid hormone, and reduction in blood phosphorus. Symptoms and signs range from overt loss of calcium from bone and formation of calcium stones in the kidney to nonspecific manifestations such as weakness and fatigue. With the advent of routine automated measurement of blood calcium, primary hyperparathyroidism is often diagnosed in individuals with no symptoms. The disorder occurs most frequently in women after menopause. There are also inherited forms of the disease. The nonhereditary form of the disease is commonly due to a single benign parathyroid tumor, and very rarely is caused by a parathyroid gland cancer. The hereditary forms are caused by benign enlargement of several or all four of the parathyroid glands. Treatment of primary hyperparathyroidism involves surgical removal of the abnormal gland(s). Secondary hyperparathyroidism is defined as excessive secretion of parathyroid hormone in response to reduction in blood calcium caused by kidney or intestinal disorders or vitamin D deficiency. Treatment is directed at the underlying disorder (for example, kidney transplantation for kidney failure). *See* CALCIUM METABOLISM; VITAMIN D.

Hypoparathyroidism. This group of disorders involves a deficiency in parathyroid hormone secretion or metabolism. Manifestations include involuntary muscle contractions, or generalized seizures in the most extreme cases. Diagnosis is made by demonstrating decreased blood calcium and increased blood phosphorus despite normal kidney function. In most forms of hypoparathyroidism, parathyroid hormone is low or undetectable in blood despite reduced blood calcium because the parathyroid glands are unable to respond appropriately by secreting parathyroid hormone. Hypoparathyroidism can be caused by parathyroid gland destruction, for example, by inadvertent surgical removal. Certain rare forms of hypoparathyroidism are due to resistance to parathyroid hormone action rather than hormone deficiency, and are termed pseudohypoparathyroidism. Blood parathyroid hormone is actually elevated in this form of hypoparathyroidism, because the normal parathyroid glands try to compensate for the inability of the kidney to respond to the hormone. Treatment of all forms of hypoparathyroidism involves administration of calcium and large amounts of vitamin D to restore a normal blood calcium and phosphorus. *See* PARATHYROID GLAND.

Allen M. Speigel

Bibliography. J. P. Bilezikian, R. Marcus, and M. A. Levine (eds.), *The Parathyroids: Basic and Clinical Concepts*, 1994; J. D. Wilson and D. W. Foster (eds.), *Williams Textbook of Endocrinology*, 9th ed., 1998.

Parathyroid hormone

The secretory product of the parathyroid glands. Parathyroid hormone (PTH) is a single-chain polypeptide composed of 84 amino acids. The peptide has no intrachain disulfide linkages and has a molecular weight of 9500. The sequences of human, bovine, and porcine parathyroid hormone are known, and the gene for human parathyroid hormone has been cloned and sequenced (see **illus**.).

Chemistry and biosynthesis. Parathyroid hormone is produced from a larger polypeptide precursor by means of two enzymic cleavages. The gene product is prepro-PTH, a polypeptide of 115 amino acids. The leader sequence is removed in the microsomes, generating pro-PTH, a peptide of 90 amino acids. The 6–amino acid pro-sequence is removed in the Golgi apparatus, generating the final 84–amino acid storage and secretory form of the hormone.

Biological activity resides in the amino terminal portion of the molecule; synthetic parathyroid hormone containing only amino acid residues 1–34 has been shown to contain complete biological activity. The native 84–amino acid secretory form of the peptide is rapidly cleaved into amino-terminal and carboxy-terminal fragments in the body of the organism, but whether amino-terminal fragments constitute an important component of total circulating biological activity is unknown. The multiple forms and fragments of parathyroid hormone in the circulation have complicated the establishment of valid immunoassays for the hormone. Knowledge of the primary structure of the hormone has, however, permitted synthesis of active competitive inhibitors.

Regulation of secretion. The major regulator of parathyroid hormone secretion is the serum concentration of calcium ions, to which the parathyroid cells are exquisitely sensitive. Only a limited amount of parathyroid hormone is stored in secretory granules, so that a hypocalcemic stimulus must ultimately influence biosynthesis as well as secretion of the hormone. Parathyroid secretory protein is a large, acidic glycoprotein which is stored and cosecreted with parathyroid hormone in roughly equimolar amounts; the biological function of parathyroid secretory protein is unknown.

Target tissues and actions. The principal target tissues of parathyroid hormone are bone and the kidney; intestinal effects previously attributed to the hormone are now known to be indirectly mediated by 1,25-dihydroxyvitamin D. *See* BONE; KIDNEY.

The actions of parathyroid hormone at the level of bone are complex. Parathyroid hormone stimulates preformed osteoclasts (the large, multinucleate cells principally responsible for bone resorption) to resorb bone. This process results in the removal of both bone mineral (largely calcium and phosphate ions) and bone matrix (largely collagen). Chronic stimulation with parathyroid hormone results in the recruitment of a larger population of osteoclasts, with an increase in the surface area of bone participating in the resorption process. Chronic stimulation by parathyroid hormone also results in an increase in bone formation (mediated by osteoblasts) and mineralization, the net result therefore being an increase in all aspects of bone metabolism and turnover.

Parathyroid hormone has three important actions in the kidney: (1) it stimulates calcium reabsorption, thereby reducing the quantity of calcium lost into the urine; (2) it inhibits phosphate reabsorption, thereby increasing the quantity of phosphate excreted into the urine; and (3) it stimulates the synthesis of 1,25-dihydroxyvitamin D, the final hormonal product of vitamin D metabolism. The stimulation of 1,25-dihydroxyvitamin D synthesis results in an increase in 1,25-dihydroxyvitamin D–mediated calcium absorption in the small intestine, enhancing the flow of exogenous calcium into the organism.

In all of its target cells, the interaction of parathyroid hormone with its receptors results in a stimulation of adenylate cyclase activity and cyclic adenylic acid (AMP) formation. In many cells that are responsive to parathyroid, the net regulation of cellular physiology appears to result from a complex interaction of the hormone's effects on the cytosolic concentration of cyclic AMP and free calcium ions.

Physiology. Parathyroid hormone is responsible for the fine regulation of the serum calcium concentration on a minute-to-minute basis. This is achieved by the acute effects of the hormone on calcium resorption in bone and calcium reabsorption in the kidney. The phosphate mobilized from bone is excreted into the urine by means of the hormone's influence on renal phosphate handling. Parathyroid hormone also stimulates calcium absorption in the intestine, this being mediated indirectly by 1,25-dihydroxyvitamin D. Thus, a hypocalcemic stimulus of parathyroid hormone secretion results in an increased influx of calcium from three sources (bone, kidney, and intestine), resulting in a normalization of the serum calcium concentration without change in the serum phosphate concentration. *See* CALCIUM METABOLISM.

Disorders. The syndromes of overproduction and underproduction of parathyroid hormone are known as hyperparathyroidism and hypoparathyroidism. Hyperparathyroidism is a common disorder and usually results from a benign tumor (adenoma) involving one of the parathyroid glands. The hallmarks of hyperparathyroidism are an increase in the serum calcium concentration and a decrease in the serum phosphate concentration, and it may be complicated by bone loss and the formation of calcium-containing kidney stones. The hallmarks of hypoparathyroidism are a reduction in the serum calcium concentration and an increase in the serum phosphate concentration, with neuromuscular irritability and occasionally frank seizures being the principal complications. Hypoparathyroidism is uncommon. Pseudohypoparathyroidism is a rare form of the disorder which results from target tissue resistance to the effects of the hormone. *See* PARATHYROID GLAND; PARATHYROID GLAND DISORDERS; THYROCALCITONIN; VITAMIN D.

Arthur E. Broadus

Bibliography. S. A. Binkley, *Endocrinology*, 1994; P. Felig, J. D. Baxter, and L. A. Frohman, *Endocrinology and Metabolism*, 3d ed., 1994.

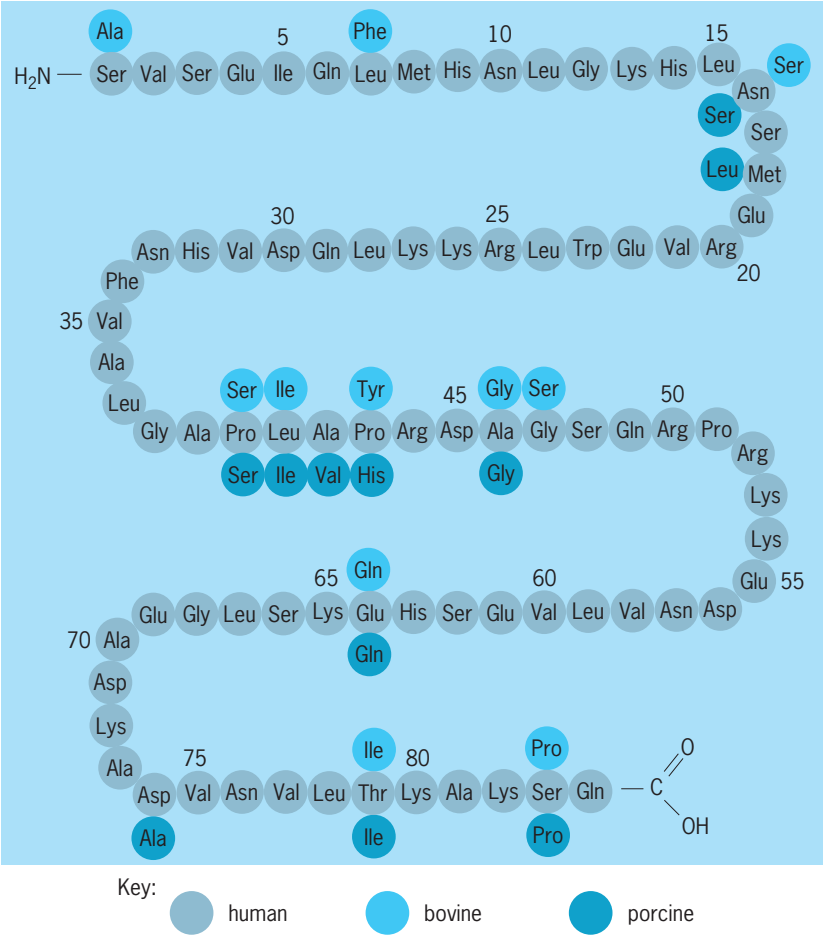
H2N — Ser Val Ser Glu Ile Gln Leu Met His Asn Leu Gly Lys His Leu Ser Asn Ser Ser Leu Met Glu Arg Leu Trp Glu Val Arg Lys Lys Lys Glu Asp Asp Val Leu Val Asn Asp Glu His Ser Glu Val Leu Ser Gln Arg Pro Arg Ser Gly Ala Asp Arg Pro Ala Leu Ala Pro Leu Ser Ile Val His Ser Ile Tyr Gly Ser Gly Ala Ala Leu Val Phe Asn His Val Asp Gln Leu Lys Lys Arg Leu Trp Glu Val Arg Ala Phe Ala Leu Gly Ala Pro Leu Ala Pro Arg Asp Ala Gly Ser Gln Arg Pro Ala Asp Lys Ala Asp Val Asn Val Leu Thr Lys Ala Lys Ser Gln C O OH Ala Ile Pro Ala Ile Pro Gln Glu Gly Leu Ser Lys Glu His Ser Glu Val Leu Val Asn Asp Gln Glu Ser Ala Ala Ser Ser Leu Leu Met Glu Ser Ser Ser

is given to the fact that during the development of sponges with parenchymella larvae, the flagellated external cells of the larva take up an internal position as choanocytes after metamorphosis, whereas the epidermal and mesenchymal cells arise from what was an internal mass of cells in the larva. These facts suggest that either the germ layers of sponges are reversed in comparison with those of other Metazoa or the choanocytes cannot be homologized with the endoderm of other animals. Either interpretation supports the wide separation of sponges from all other Metazoa to form the subkingdom Parazoa or Enantiozoa. *See* CALCAREA; DEMOSPONGIAE; METAZOA; PORIFERA.

Development. On the other hand, there are cogent arguments in favor of the basic similarity of the development of sponges and other Metazoa. Detailed studies of the embryology of Calcarea with amphiblastula larvae suggest an explanation for the reversal of the germ layers seen so strikingly in the development of a parenchymella. The egg cell of *Sycon* (**Fig. 1**), for example, always lies beneath the layer of choanocytes of a flagellated chamber with its long axis parallel to that layer. The sperm enters from a carrier cell at the pole adjacent to the maternal choanocyte layer, and this pole is determined as ectoblastic. Three meridional cleavages are followed by an equatorial cleavage which separates two tiers of eight cells each. The tier in contact with the choanocyte layer gives rise to ectomesenchyme; the other eight cells furnish the choanocytes. Further divisions of the flagellated cells lead to the formation of a hollow blastula with an opening to the outside between the ectomesenchymal cells. The flagella at this stage are located on the internal surface of the embryo. The polarity of the embryo at this stage is identical to that of a hydrozoan planula, where the pole next to the blastostyle and hence near the source of food is ectoblastic. In *Sycon*, however, the blastula is inverted at this stage, so that the flagella of the flagellated cells are on the external surface, and this region of the larva is anterior as the amphiblastula swims. The ectomesenchymal cells are now at the posterior pole. *See* BLASTULATION; CLEAVAGE (DEVELOPMENTAL BIOLOGY); HYDROZOA.

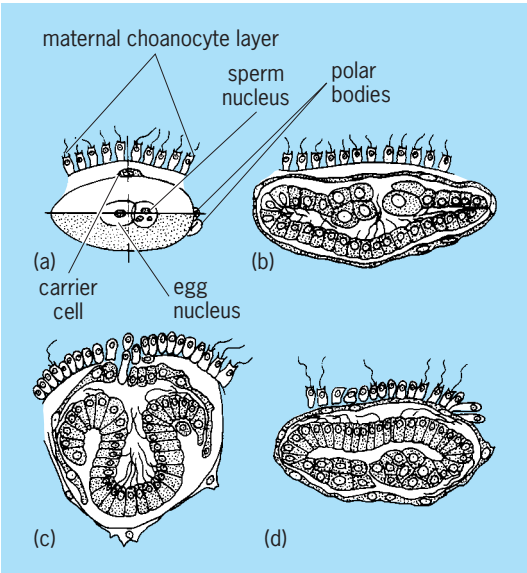

Fig. 1. Development of *Sycon*. (*a*) Fertilized egg beneath the maternal choanocyte layer. (*b*) Stomoblastula. (*c*) Blastula in process of eversion. (*d*) Inverted blastula.

Gastrulation and larvae. The amphiblastula of *Sycon*, upon settling on its anterior end, gastrulates by invagination of the flagellated hemisphere. After the blastopore closes, an osculum breaks through at the opposite, free end. Were it not for the fact that pores also develop in the body wall, such a simple attached olynthus would bear a striking resemblance to a planula developing into a polyp.

The parenchymella larva may be interpreted as an amphiblastula with an accelerated development. No trace of the inversion of the surfaces is apparent, except for the external position of the flagella. The ectomesenchymal cells develop precociously at the posterior pole and eventually fill the interior of the larva, displacing the blastocoele. Gastrulation by invagination is no longer possible; instead the peripheral flagellated cells migrate internally to take up their places in the flagellated chambers. The most accelerated development in sponges is seen in the parenchymellae of spongillids, in which choanocytes begin to differentiate from the internal mass of blastomeres while the larvae are still free-swimming. The external flagellated cells are phagocytized by amebocytes after fixation of the larva to the substrate, and thus take no part in choanocyte formation. *See* GASTRULATION; INVERTEBRATE EMBRYOLOGY; PHAGOCYTOSIS.

Thus, although it cannot be denied that a reversal of the germ layers is a fact in many sponge parenchymellae, this process can be explained as a consequence of the inversion of the surfaces of the larva which brings the flagellated cells into an anterior position with their flagella directed outward. Clear evidence of this process is seen only in the prolonged development of the amphiblastula larvae of certain Calcarea among existing sponges.

Phylogeny. The striking similarity in structural details between the sponge choanocyte and zooflagellates of the group called choanoflagellates (**Fig. 2**), or craspedomonadines, has led most zoologists to look to this group of protozoa as being ancestral to the sponges. Choanoflagellates are holozoic and capture food by means of the collar. *Proterospongia*, found in both fresh and marine waters, has choanocytelike cells embedded in a gelatinous mass in which ameboid cells wander. It shows a striking similarity to free-swimming aggregates of cells sometimes found in cultures of dissociated sponges. Choanocytes not only resemble choanoflagellates but also certain cells among other Metazoa. Some corals have flagellated cells with collars of cytoplasmic filaments, and mollusks and vertebrates have microvilli on the cells making up ciliated epithelia. *See* KINETOPLASTIDA.

The phenomenon of inversion of the surfaces which occurs in the blastulae of certain calcareous sponges bears a striking resemblance to a similar process in the development of daughter colonies as

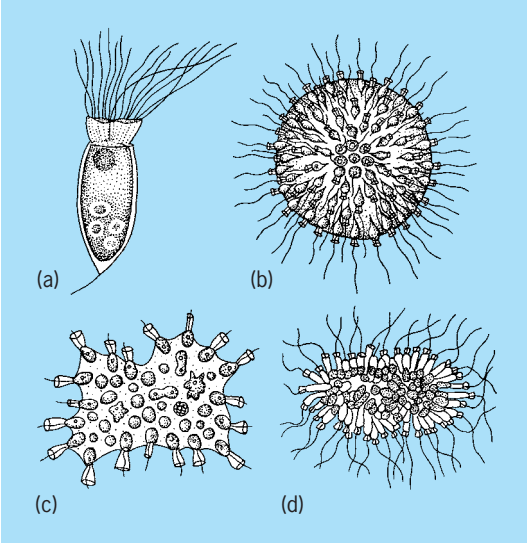

Fig. 2. Choanoflagellate, or craspedomonadine, protozoa. (*a*) *Salpingoeca infusionum*, showing cytoplasmic filaments arising from the collar. (*b*) *Sphaeroeca volvox*, a colonial planktonic form. (*c*) *Proterospongia haeckeli*. (*d*) Free-swimming aggregate of cells of *Grantia compressa* which formed after dissociation of adult cells.

well as sexually produced young in *Volvox*. However, a volvocine ancestry for the sponges is difficult to support on any other grounds. *Volvox* cells lack a collar, have two flagella, are photosynthetic, and have postzygotic reduction divisions. Furthermore, the Volvocales are exclusively fresh-water organisms, whereas the earliest fossil sponges are found in marine strata. *See* ANIMAL KINGDOM. Willard D. Hartman

Parenchyma

A ground tissue chiefly concerned with the manufacture and storage of food. The primary functions of plants, such as photosynthesis, assimilation, respiration, storage, secretion, and excretion—those associated with living protoplasm—proceed mainly in parenchymal cells. Parenchyma is frequently found as a homogeneous tissue in stems, roots, leaves, and flower parts. Other tissues, such as sclerenchyma, xylem, and phloem, seem to be embedded in a matrix of parenchyma; hence the use of the term ground tissue with regard to parenchyma is derived. The parenchymal cell is one of the most frequently occurring cell types in the plant kingdom. *See* PLANT ANATOMY; PLANT PHYSIOLOGY.

Origin. Parenchymal cells are differentiated in the primary plant body (shoot, root) from apical growing zones or meristems. The ground meristem that initates pith, cortex, and leaf mesophyll is the seat of development of parenchyma. In the secondary plant body parenchymal cells are derived from the vascular cambium and cork cambium in the form of ray tissues, xylem and phloem parenchyma, and phelloderm. *See* APICAL MERISTEM; LATERAL MERISTEM.

Variations. Typical parenchyma occurs in pith and cortex of roots and stems as a relatively undifferentiated tissue composed of polyhedral cells that may be more or less compactly arranged and show little variation in size or shape (**Fig. 1**). The mesophyll, that is, the tissue located between the upper and lower epidermis of leaves, is a specially differentiated parenchyma called chlorenchyma because its cells contain chlorophyll in distinct chloroplastids (**Figs. 2** and **3**).

This chlorenchymatous tissue is the major locus of photosynthetic activity and consequently is one of the more important variants of parenchyma. Specialized secretory parenchymal cells are found lining resin ducts and other secretory structures. *See* PHOTOSYNTHESIS; SECRETORY STRUCTURES (PLANT).

Cell contents. Parenchymal cells include the protoplast, common to all living cells, usually with a large vacuole located in the center and the cytoplasm and nucleus next to the cell wall. The kinds of inclusions are partly related to the position of the cell in the plant. Thus cells in leaf mesophyll always have chloroplasts. Cells in root cortex commonly

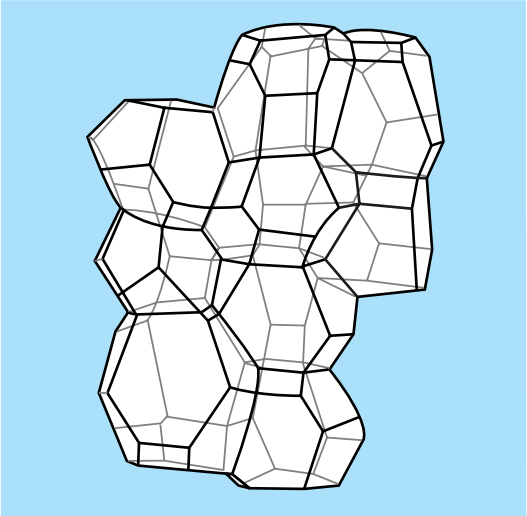

Fig. 1. Schematic representation of a group of cells from parenchyma of *Asparagus* root, illustrating regular arrangement of cells in a compact tissue.

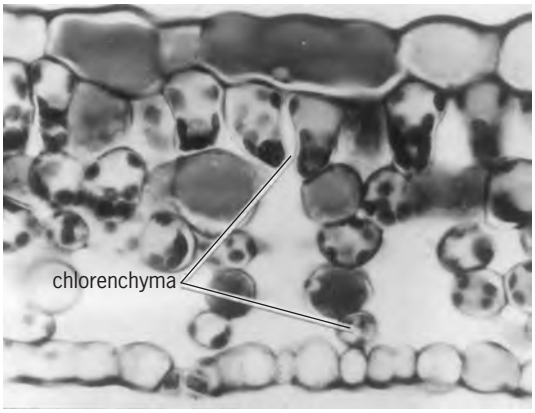

Fig. 2. Transverse section of a leaf of *Tolmiea menziesii*, showing mesophyll or chlorenchyma with prominent intercellular spaces. (*Photograph by R. B. Wylie*)

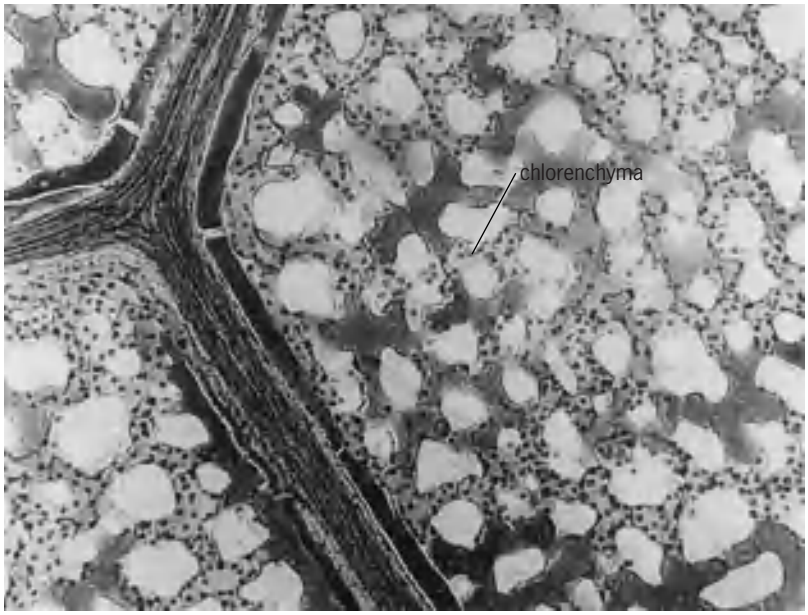

Fig. 3. Section of *Olearia rami* leaf cut parallel to surface, showing mesophyll or chlorenchyma with prominent intercellular spaces. (*Photograph by R. B. Wylie*)

contain grains of storage starch. Cells in pith and cortex of stems contain starch grains and crystals, usually of calcium oxalate. Substances in crystal form are usually interpreted as waste products and are lost with leaf fall, shedding of bark, or decay of pith. Parenchyma in flower parts or fruits may have yellow or red plastids in the cytoplasm or purple or blue anthocyanin pigments in the vacuole.

Cell walls. Parenchymal cells are typically thin and contain cellulose, hemicelluloses, and pectins. They may become thick, hard, and lignified, for example, in wood parenchyma. *See* CELL WALLS (PLANT); CELLULOSE; HEMICELLULOSE; PECTIN.

Cell shape. The three-dimensional cell shape of parenchyma is frequently a function of cell arrangement. In a compact tissue, parenchymal cells are almost isodiametric and polyhedral with an average of approximately 14 sides in contact with 14 other cells. In more loosely arranged tissues with many air spaces, the average number of contacts with adjoining cells is reduced. Parenchymal cells may deviate from the isodiametric shape. Those in the vascular tissues are often conspicuously elongated. Parenchymal cells in the mesophyll and in other tissues with large intercellular spaces may be variously lobed. *See* CORTEX (PLANT); EPIDERMIS (PLANT); FLOWER; FRUIT; LEAF; PHLOEM; PITH; ROOT (BOTANY); SCLERENCHYMA; STEM; XYLEM.

Robert L. Hulbary

Pareto's law

A concept developed by the Italian economist Vilfredo Pareto (sometimes called the 20-80 rule) describing the frequency distribution of an empirical relationship fitting the skewed concentration of a variate-values pattern. The phenomenon wherein a small percentage of a population accounts for a large percentage of a particular characteristic of that population is an example of Pareto's law. When the data are plotted graphically, the result is called a maldistribution curve. To take a specific case, an analysis of a manufacturer's inventory might reveal that less than 15% of the component part items account for over 90% of the total annual usage value. Other analyses may show that a small percentage of suppliers account for most of the late or substandard deliveries; a few engineers get most of the patents; a few employees are responsible for most of the quality rejects, errors, accidents, absences, thefts, and so on. *See* INVENTORY CONTROL.

The key to benefiting from Pareto's law is to recognize its existence in a given situation and then to handle the case accordingly. For example, if there is no Pareto's law effect in a population—for example, the number of assembly errors in a factory is equally spread among all the employees—then one course of action should be taken by the engineer; but if Pareto's law is present and 5–10% of the employees are making 85–90% of the errors, then a very different course of action should be taken. If this important distinction is not made, not only will the engineer's action be the wrong one for the problem, but more importantly the engineer's entire approach to engineering management will be wrong.

Good management requires the devotion of a major portion of the manager's time and energy to the solution of the major problems—with the less important problems delegated to subordinates, handled by an automatic system, dealt with through management by exception, or in some cases simply ignored. A manager who devotes too much time to unimportant things is in peril of failure. All things should not be considered equally when in fact they are not of equal impact on the total problem.

A Pareto's law analysis distinguishes the significant or vital few from the insignificant or trivial many. If it reveals that, for example, 2 out of 20 items (10%) account for 80% of the total problem, then by focusing attention on only that 10% of the population, up to 80% of the problem can be resolved.

The mathematics required to calculate and graph the curve of Pareto's law is simple arithmetic. The **illustration** shows a typical curve, where the cumulative percentage of the total annual usage value (Σ) is plotted against the cumulative percentage of the total number of parts (N) of the different components used in manufacture of a company's products.

It should be noted, however, that the calculations need not be done in all cases. It may suffice to merely make a rough approximation of a situation in order to determine whether or not Pareto's law is present and whether benefits may subsequently accrue. For example, a firm with 100 employees can easily review its personnel records and pick out the 5 employees having the worst chronic absenteeism. Suppose those 5 employees were absent for 120 of the 160 total days of work lost by all 100 employees. Pareto's law would apply, as 5% of the population accounts for 75% of the problem, while the remaining 95% of the employees account for only 25% of the total absences. The value of such a discovery would

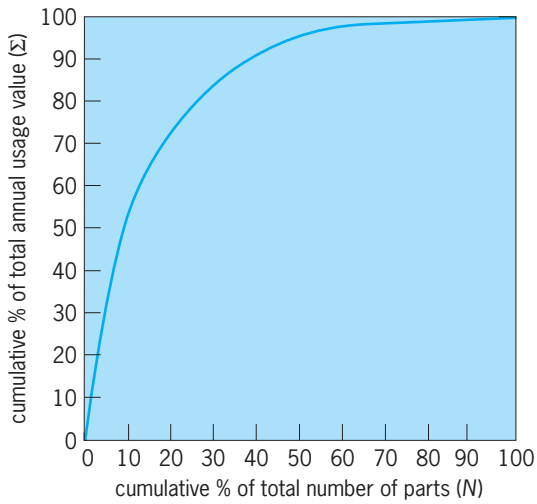

Typical curve representing Pareto's law.

be the realization that absenteeism could be reduced by 75% by dealing with only 5 out of 100 employees. Vincent M. Altamuro

Bibliography. J. M. Juran, *Quality Control Handbook*, 4th ed., 1988; M. G. Kendall, *Dictionary of Statistical Terms*, 4th ed., 1986; M. O. Lorenz, *Methods of Measuring the Concentration of Wealth*, Amer. Stat. Ass. Pub., vol. 9, pp. 209–219, 1904–1905.

Parity (quantum mechanics)

A physical property of a wave function which specifies its behavior under simultaneous reflection of all spatial coordinates through the origin, that is, when x is replaced by $-x$, y by $-y$, and z by $-z$. If the single-particle wave function ψ satisfies Eq. (1), it is

$$\psi(x,y,z) = \psi(-x,-y,-z) \tag{1}$$

said to have even parity. If, on the other hand, Eq. (2)

$$\psi(x,y,z) = -\psi(-x,-y,-z) \tag{2}$$

holds, the wave function is said to have odd parity. These two expressions can be combined in Eq. (3),

$$\psi(x,y,z) = P\psi(-x,-y,-z) \tag{3}$$

where $P = \pm 1$ is a quantum number, parity, having only the two values $+1$ (designated as even parity) and -1 (odd parity). More precisely, parity is defined as the eigenvalue of the operation of space inversion. Parity is a concept that has meaning only for fields or waves and therefore has meaning only in classical field theory or in quantum mechanics. *See* QUANTUM MECHANICS.

Conservation. The conservation of parity follows from the inversion symmetry of space, that is, the invariance of the Schrödinger equation $H\psi = E\psi$ (the wave equation satisfied by the wave function ψ) to the inversion of space coordinates, $\mathbf{r} \to -\mathbf{r}$. As usual, this is shown formally by defining a transformation operator: Let $\mathscr{P}$ be the parity operator which inverts space; that is, $\mathscr{P}$ acting on a wave function yields the wave function at the inverse point of space, $\mathscr{P}\ \psi(\mathbf{r}) = \psi(-\mathbf{r})$. Similarly, for an operator A, $\mathscr{P}A(\mathbf{r})\mathscr{P}^{-1} = A(-\mathbf{r})$. [Here, $\mathbf{r}$ stands for the cartesian coordinates of all the particles of the system.] The statement that the world is symmetrical to inversion means that the hamiltonian H after inversion is the same as before, that is, $\mathscr{P}H\mathscr{P}^{-1} = H$; and thus $\mathscr{P}H - H\mathscr{P} = [\mathscr{P},H] = 0$. Since $\mathscr{P}$ commutes with H, it is a constant of the motion. Further, H and $\mathscr{P}$ can be simultaneously diagonal; that is, the eigenfunctions of H can be simultaneously eigenfunctions of $\mathscr{P}$. In fact, if for an eigenvalue E of H there is only a single eigenfunction (nondegenerate level), this eigenfunction must be an eigenfunction of $\mathscr{P}$. As for the eigenvalues of $\mathscr{P}$, since $\mathscr{P}^2 = 1$, it follows that the possible eigenvalues of $\mathscr{P}$ are $+1$ or -1. That is, an eigenfunction of $\mathscr{P}$ satisfies $\mathscr{P}\ \psi_\pm(\mathbf{r}) = \psi_\pm(-\mathbf{r}) = \pm\psi_\pm(\mathbf{r})$, where the upper (lower) sign indicates an eigenfunction of positive (negative) parity, also known as even (odd) parity, as stated above. *See* EIGENFUNCTION; EIGENVALUE (QUANTUM MECHANICS); NONRELATIVISTIC QUANTUM THEORY.

The parity (or inversion) operator, which changes $\mathbf{r}$ to $-\mathbf{r}$, has the alternative interpretation that the coordinate values remain unchanged but the coordinate axes are inverted; that is, the positive x axis of the new frame points along the old negative x axis, and similarly for y and z. If the original frame was right-handed, then the new frame is left-handed. [A cartesian coordinate system (frame, for short) is called right-handed if it is possible to place the right hand at the origin and point the thumb and first and second fingers along the positive x, y, and z axes, respectively.] Thus, parity would be conserved if the statement of physical laws were independent of the handedness of the coordinate system that was being used. Of course, the fact that most people are right-handed is not a physical law but an accident of evolution; there is nothing in the relevant laws of physics which favors a right-handed over a left-handed human. The same holds for optically active organic compounds, such as the amino acids. However, the statement that the neutrino is left-handed is a physical law. *See* NEUTRINO.

All the strong interactions between hadrons (for example, nuclear forces) and the electromagnetic interactions are symmetrical to inversion, so that parity is conserved by these interactions. As far as is known, only the weak interactions fail to conserve parity. Thus parity is not conserved in the weak decays of elementary particles (including beta decay of nuclei); in all other processes the weak interactions play a small role, and parity is very nearly conserved. Likewise, in energy eigenstates, weak interactions can be neglected to a very good approximation, and parity is very nearly a good quantum number, so that each atomic, nuclear, or hadronic state is characterized by a definite value of parity, and its conservation in reactions is an important principle. *See* FUNDAMENTAL INTERACTIONS; WEAK NUCLEAR INTERACTIONS.

Orbital parity. The parity of a one-particle state of orbital angular momentum l is given by $P = (-)^l$, that is, even $(+1)$ for s, d, . . . , states, and odd (-1) for

$p, f, \ldots,$ states. Thus the deuteron, whose state is a linear combination of $^3S_1$ and $^3D_1$, has even parity; there cannot be any admixture of the odd parity state $^3P_1$. The orbital parity of an n-particle system is the product of the parities of the $n-1$ relative orbital angular momentum states: $P_{\text{orb}} = (-)^{l_1 + \cdots + l_{n-1}}$. Thus the parity of an atom is the product of the parities of the one-electron wave functions; any configurations which mix must have the same parity. The Laporte rule of atomic spectroscopy, which states that an electric dipole transition can occur only between states of opposite parity, depends on the fact that the electric dipole radiation field has odd parity. *See* ANGULAR MOMENTUM; SELECTION RULES (PHYSICS).

Intrinsic parity. The total parity of a system of particles is the product of the (total) orbital parity and the intrinsic parities of the particles. ("Particles" here can mean any subsystems.) But the intrinsic parity of a conserved particle is irrelevant and can be omitted, as was done implicitly in the examples discussed above. For if the particle is conserved in a reaction, so is the contribution of its intrinsic parity to the total parity, so that its intrinsic parity is irrelevant to the balance of parity in the reaction. In fact, if a particle is conserved in all reactions, its intrinsic parity can never be determined. The photon is an unconserved particle; its intrinsic parity is odd. The parity of the π^0 meson (a pseudoscalar) is odd, so that to conserve parity it must be emitted by a nucleon into a P state. By charge independence, the charged π meson must also be emitted in a P state; it is natural to call the parity of the charged π meson odd also, which amounts to defining the parity of the neutron and proton (or the u and d quarks) to be the same. An electron by itself is conserved, but an electron plus a positron can annihilate. Thus the product of the parities of an electron and a positron is well defined. According to the Dirac equation of relativistic quantum theory, the product of their parities is -1. The same result holds for any fermion particle-antiparticle pair. Thus the parity of positronium is -1 times its orbital parity, that is, $-(-)^l$. *See* ELEMENTARY PARTICLE; QUARKS; RELATIVISTIC QUANTUM THEORY.

Spin and momentum correlations. The symmetry of the strong and electromagnetic interactions with respect to inversion implies statements about possible correlations of momenta and spins of the particles emitted as a result of such reactions. The principle is that the probability of a configuration of momenta and spins must be a scalar; that is, it must not change under inversion of the coordinate system. A momentum $\mathbf{p}$ is a vector like $\mathbf{r}$ (it changes sign under inversion: $\mathbf{p} \rightarrow -\mathbf{p}$). On the other hand, orbital angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ is a pseudovector, since under inversion $\mathbf{L} \rightarrow +\mathbf{L}$; the same must hold for spin angular momentum $\mathbf{S}$. Thus $\mathbf{S} \cdot \mathbf{p}$ is a pseudoscalar (it changes sign under inversion), and so such a term cannot occur in the angular distribution of a parity conserving process. This term, $\mathbf{S} \cdot \mathbf{p}$, in an angular distribution, would correlate a particle's spin with its momentum; that is, it would imply a polarization in the momentum direction, or longitudinal polarization, which is accordingly not produced in strong and electromagnetic reactions.

Nonconservation. One of the selection rules which follows from parity conservation is the following: A spin zero boson cannot decay sometimes into two π mesons and sometimes into three π mesons, because these final states have different parities, even and odd respectively. But the positive K meson is observed to have both these decay modes, originally called the θ and the τ mesons, respectively, but later shown by the identity of masses and lifetimes to be decay modes of the same particle. This τ-θ puzzle was the first observation of parity nonconservation. In 1956, T. D. Lee and C. N. Yang made the bold hypothesis that parity also is not conserved in beta decay. They reasoned that the magnitude of the beta-decay coupling is about the same as the coupling which leads to decay of the K meson, and so these decay processes may be manifestations of a single kind of coupling. Also, there is a very natural way to introduce parity nonconservation in beta decay, namely, by assuming a restriction on the possible states of the neutrino (two-component theory). They pointed out that no beta-decay experiment had ever looked for the spin-momentum correlations that would indicate parity nonconservation; they urged that these correlations be sought.

In the first experiment to show parity nonconservation in beta decay, the spins $\mathbf{S}_{\text{Co}}$ of the beta-active nuclei cobalt-60 were polarized with a magnetic field $\mathbf{H}$ at low temperature; the decay electrons were observed to be emitted preferentially in directions opposite to the direction of the ^{60}Co spin (see **illus.**). Thus a $\mathbf{S}_{\text{Co}} \cdot \mathbf{p}_e$ correlation was found or, in terms of macroscopic quantities, an $\mathbf{H} \cdot \mathbf{p}_e$ correlation. The magnitude of this correlation shows that the parity-nonconserving and parity-conserving parts of the beta interaction are of equal size, substantiating the two-component neutrino theory.

It was at first somewhat disconcerting to find parity not conserved, for that seemed to imply a handedness of space which would then not be the empty thing which (since the demise of the ether hypothesis) many physicists think it to be. That is, an ether would be needed to provide a standard of handedness at each point of space, to tell ^{60}Co which direction to decay into. But this is not really the situation; the saving thing is that anti-^{60}Co decays in the opposite direction. Thus, after all, there is nothing

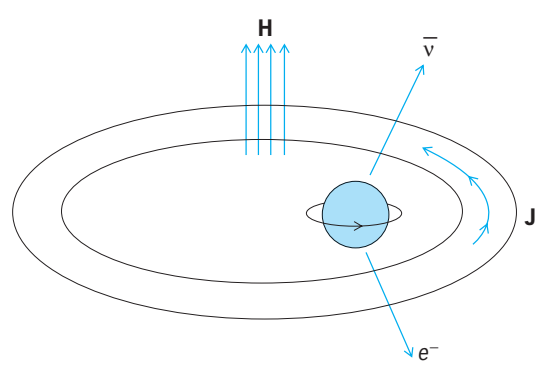

Beta decay from polarized cobalt-60 nuclei. When the spin axes of the cobalt nuclei are not polarized, the preferential emission of the electrons and antineutrinos in the directions shown does not occur.

intrinsically left-handed about the world, just as there is nothing intrinsically positively charged about nuclei. What really exists here is a correlation between handedness and sign of charge.

The present theory of weak interactions is the so-called standard model of the electroweak gauge theory, in which weak and electromagnetic interactions are the consequence of the exchange of the gauge bosons W^+, W^-, Z^0, and γ (the photon). The exchange of $W^\pm$ bosons produces what are called charged-current weak interactions, in which each fundamental fermion that emits or absorbs a $W^\pm$ loses or gains a unit of charge (more precisely, makes a transition between up-type and down-type). This transformation of fundamental fermions is responsible for the weak decays of particles. It is maximally parity-violating in the sense that the $W^\pm$ couples only to the left-hand chiral part of the fermions (chiral coupling); this means that if the mass of the fermion is negligible compared to its energy (which is always true for neutrinos), the fermion is 100% longitudinally polarized. At energies small compared to the mass of the $W^\pm$, this part of the standard model is identical to the older weak interaction theory. *See* ELECTROWEAK INTERACTION; INTERMEDIATE VECTOR BOSON; STANDARD MODEL.

Neutral-current weak interactions. The exchange of Z^0 bosons produces what are called neutral-current weak interactions, in which the fundamental fermions that emit and absorb the Z^0 are unchanged (except in energy and momentum). Since these interactions do not result in decays, they are not as obvious as the charged-current weak interactions, and in fact were not discovered until after their prediction by the standard model. The coupling of the Z^0 to fundamental fermions has two terms: One is a chiral coupling (like the coupling of $W^\pm$) and has opposite sign for up-type and down-type fermions; the other is a vector coupling (like the coupling of the photon γ) and is proportional to the charge of the fermion. The ratio of strengths of these two terms is proportional to $\sin^2 \theta_W$, where the weak angle θ_W is the one new parameter of the standard model. A strong test of the standard model is that when neutral-current weak interactions are described by this form of coupling of the Z^0, the same value of θ_W is found, for all reactions involving any of the fundamental fermions.

At energies not negligibly small compared to the mass of the Z^0, the neutral-current weak interactions are not negligible compared to electromagnetic interactions and can be observed directly in lepton-lepton or lepton-hadron scattering. But at lower energies these interactions can be observed only through the parity nonconservation that they cause. *See* NEUTRAL CURRENTS.

Parity nonconservation in atoms. Parity nonconservation has been observed in several atoms. An example is the work by Bouchiat and colleagues (1986) on cesium. The principle used here is that the M1 radiative decay $7S_{1/2} \rightarrow 6S_{1/2}$ (ground state) is highly forbidden (because the valence electron wave functions in the two states are nearly orthogonal), with a decay rate of only 10^{-6} s^{-1}; compared to this the parity-nonconserving E1 transition amplitude that results from the mixing of $P_{1/2}$ states into the $S_{1/2}$ states by the parity-violating part of the weak interaction (Z^0 exchange) between the valence electron and the nucleus is not small. This parity-nonconserving E1 transition amplitude, proportional to $\mathbf{S}_e$ (the spin of the valence electron), is detected by observing its interference with a parity-conserving transition amplitude, proportional to $\mathbf{E}$ (and also $\mathbf{S}_e \times \mathbf{E}$), where $\mathbf{E}$ is a static electric field applied to the atom. The interference results in circular polarization of light (longitudinal polarization of photons) emitted in the fluorescent transition $7S \rightarrow 6P$ in a direction perpendicular both to the electric field and to the direction of incident circularly polarized light which resonantly excites the $6S \rightarrow 7S$ transition; that is, the observed parity nonconservation signal is $\sim \mathbf{S}_{\text{fluor}} \times \mathbf{S}_{\text{excit}} \cdot \mathbf{E}$, where $\mathbf{S}_{\text{fluor}}$ and $\mathbf{S}_{\text{excit}}$ are the spins of the photons emitted and absorbed in the fluorescent and excitation transitions, respectively. The result shows to high precision that the ratio of coupling of the Z^0 to the u and d quarks is as predicted by the standard model with the value of θ_W as determined by other experiments, and does so at an energy (more precisely, the momentum transferred by the Z^0) much lower than that of any nonatomic experiment. *See* ATOMIC STRUCTURE AND SPECTRA.

Parity nonconservation in nuclei. Parity nonconservation has been observed in many nuclei. Nuclear parity nonconservation effects depend on nuclear and hadronic structure (in contrast to atomic parity nonconservation, where the Z^0 exchanged between the nucleus and an electron has low momentum and so couples to the total weak charge of the nucleus, which is merely the sum of the weak charges of the quarks). Hence, nuclear parity nonconservation is more a test of the understanding of this structure than of the fundamental neutral-current weak interaction. (However, it might detect any additional Z^0-like bosons that are not coupled to leptons and hence unobservable in other experiments.) On the whole, there is a trade-off between the size of the parity violation and size and complexity of the nucleus. At one extreme is a 7% longitudinal polarization of neutrons resonantly scattered from lanthanum-139 (^{139}La) at 0.73 eV, due to the mixing of a great many s-wave states into the narrow p-wave resonance; it seems impossible, in view of the lack of knowledge of the structure of such a nucleus, to quantitatively relate this result to the nucleon-nucleon parity-nonconserving interaction. At the other extreme is a 2×10^{-7}% longitudinal polarization (equivalently, a 2×10^{-7}% dependence of the cross section on longitudinal polarization) observed in low-energy (less than 50 MeV), proton-proton scattering. (In order to gather enough statistics to observe such a small effect, individual events are not observed, but rather detector currents, that is, macroscopic time averages.) The results agree with prediction. *See* NUCLEAR STRUCTURE.

As a result of parity mixing of nuclear energy levels, radiative transitions between them are mixtures of electric and magnetic multipoles, and so nuclear gamma rays have a circular polarization. Unfortunately, the efficiency of detecting their

polarization (by scattering or transmission through magnetized iron) is only a few percent. But offsetting this, some nuclei have parity doublets, pairs of levels with nearly the same energy, of the same spin but opposite parity; the near degeneracy means that the parity-nonconserving nuclear force mixes relatively strongly the two states of the pair (and only these, which simplifies the theoretical calculation). Further enhancement results if the radiative decay rates of the two levels to a common state are very different: Even a small admixture of the rapidly decaying state into the other results in a relatively large admixture of the opposite parity multipole into its decay state, giving a relatively large circular polarization, typically of the order of 0.1%. There are such doublets in four light nuclei, nitrogen-14 (^{14}N), fluorine-18 (^{18}F), fluorine-19 (^{19}F), and neon-21 (^{21}Ne); measurement of the circular polarization of all will yield information on the isobaric-spin dependence of the parity-nonconserving nuclear forces. *See* I-SPIN. Charles J. Goebel

Bibliography. E. G. Adelberger and W. C. Haxton, Parity violation in the nucleon-nucleon interaction, *Annu. Rev. Nucl. Part. Sci.*, 35:501–558, 1985; M. Bouchiat and L. Pottier, An atomic preference between left and right, *Sci. Amer.*, 250(6):100–111, June 1984; M. Gardner, *The New Ambidextrous Universe*, 3d ed., 1991; W. M. Gibson and B. R. Pollard, *Symmetry Principles in Elementary Particle Science*, 1980; G. E. Mitchell, J. D. Bowman, and H. A. Weidenmueller, Parity violation in the compound nucleus, *Rev. Mod. Phys.*, 71:445-457, 1999; R. Novick (ed.), *Thirty Years of Parity Nonconservation: A Symposium Honoring T. D. Lee*, 1987.

Parkinson's disease

A progressive degenerative disorder of the nervous system belonging to a group of conditions called motor system disorders. Parkinson's disease typically begins in middle age with an average age of onset of 59 years, but it can also begin in the 30s or 40s as well as in the 70s or later. Males are affected slightly more commonly than females; all ethnic groups can develop the disease, with Caucasians at somewhat higher risk. Parkinson's disease results in the loss of brain cells that produce dopamine. The cause is unknown but, increasingly, specific environmental and genetic factors are being evaluated for their role in the disease. Typically Parkinson's disease is sporadic, but it can occur in familial clusters; and research on some of these families has resulted in the discovery of a few genetic mutations. To date, however, genetic testing is not usually performed due to the low chance of finding a genetic cause for an affected individual. *See* NERVOUS SYSTEM (INVERTEBRATE).

Characteristics. The cardinal features of Parkinson's disease are tremor, rigidity, akinesia (absence of movement)–bradykinesia (slowness of movement), and altered posture (these features compose the acronym TRAP). This tetrad of features is not present in every patient. The tremor of Parkinson's disease is classically described as pill-rolling due to the appearance of a tremor that resembles the manner in which the fingers move when rolling a pill. It primarily occurs at rest when the limb is in repose (and is therefore called a rest tremor), and lessens when an action is initiated. Classically, the tremor begins on one side of the body, and later affects the other side as well. Rigidity or stiffness can be manifested as reduced arm swing on one side of the body, resulting in some patients mistakenly suspecting that they had a stroke. Another manifestation of rigidity is the accompanying soreness or stiffness that occurs in an arm or leg, often resulting in an orthopedic surgical consultation for a "frozen" shoulder or "bursitis" of a hip. Only with time does the progressive nature of the disease make the diagnosis easier to recognize. Typically, most disabling is the akinesia-bradykinesia. As a result, there is slowed gait with reduced stride length as well as difficulty getting out of low chairs and turning in bed. Often slowness is mistaken for aging. The individual's posture becomes stooped with flexion at the hips, elbows, and wrists. Patients develop balance problems over time, which may result in falls.

Meanwhile other features appear that impact quality of life, including fluctuations in mobility (called motor fluctuations) that occur throughout the day and difficulty initiating walking called freezing of gait. Additionally, patients experience involuntary wriggling movements called dyskinesias which occur with continued disease progression and are elicited by medication. With the development of motor fluctuations, patients lose the ability to function independently due to slowness from the unpredictable response to medication therapy as well as progressive balance difficulties. After 10–20 years, patients may require the use of a walker or a wheelchair. Cognitive changes, which may include a form of dementia that is distinct from Alzheimer's disease, develop as the disease advances. Typically, individuals with Parkinson's disease do not die of the disease per se, but succumb to the complications related to infections or other conditions that come with their progressive immobility. The lifespan of someone with Parkinson's disease is slightly less than is expected in the general population. *See* ALZHEIMER'S DISEASE; MOTOR SYSTEMS.

Mechanisms. The underlying pathology in Parkinson's disease is the progressive loss of dopamine cells that reside in the brainstem, deep within the base of the brain. Within the brainstem, the dopamine-producing nerve cells are clustered in an area called the substantia nigra because of the black pigmentation within the cells. These cells die off prematurely due to the presence of genetic and/or environmental factors. Some believe that the production of additional electrons from cell metabolism or from the environment results in cell death; others suggest that there are impairments in cellular energy production, damage from inflammation, or the generation of glutamate which causes toxic injury. One or more of these processes result in a cascade of events that kills off the dopamine cells. The theory

related to electron-mediated damage arises due to oxidative stress whereby the amount of pro-oxidants exceeds the antioxidants. With oxidative stress there are excessive amounts of chemicals with extra electrons called free radicals which are usually present on oxygen molecules (called reactive oxygen species). These unstable free radicals interact with cell membranes, DNA molecules, and proteins, causing damage and ultimately resulting in cell death. Once the loss of the dopamine nerve cell pool reaches 50–80%, patients begin to experience symptoms. *See* BRAIN; DOPAMINE.

Treatment. Unfortunately, to date there is no treatment that reverses or slows the course of the disease, and treatment is primarily directed at alleviating symptoms by restoring dopamine signaling in the brain. One common strategy is to give the dopamine precursor levodopa (L-dopa). It is combined with another medication called carbidopa so that less of the L-dopa is metabolized by the liver and more L-dopa is available in the brain for dopamine production. L-Dopa/carbidopa remains the most effective treatment for Parkinson's disease. Other medications which simulate the action of dopamine, and are therefore called dopamine agonists, are often used separately or with L-dopa to help manage symptoms. With treatment, symptoms are often reasonably well managed for the first 5–10 years. After this period of stable response to treatment, patients experience an increasing number of hours of impaired movement as the disease progresses. As patients advance into the moderate-to-advanced stages of disease, some choose to undergo deep brain stimulation (DBS) surgery to alleviate the motor fluctuations as well as dyskinesias. DBS surgery is not curative but appears to provide a stabilization of some motor symptoms for several years. Additional research is ongoing to evaluate novel oral medications, gene therapies, and other surgical treatments to help alleviate the suffering of Parkinson's disease and hopefully to alter course. Paul Tuite

Bibliography. M. J. Farrer, Genetics of Parkinson disease: Paradigm shifts and future prospects, *Nat. Rev. Genet.*, 7(4):306–318, 2006; A. E. Lang and A. M. Lozano, Parkinson's disease (2 parts), *N. Engl. J. Med.*, 339(15):1044–1053 and (16):1130–1143, 1998; J. G. Nutt and G. F. Wooten, Clinical practice: Diagnosis and initial management of Parkinson's disease, *N. Engl. J. Med.*, 353(10):1021–1027, 2005.

Parsec

A unit of measure of astronomical distances. One parsec is equivalent to 3.084×10^{13} km or 1.916×10^{13} mi. There are 3.26 light-years in 1 parsec. The parsec is defined as the distance at which the semimajor axis of Earth's orbit around the Sun (1 astronomical unit) subtends 1 second of arc. Thus, because the angle is small, the equation below holds.

$$\frac{1 \text{ astronomical unit}}{1 \text{ parsec}} = 1 \text{ second} = \frac{1}{206{,}265}$$

A parsec is then 206,265 astronomical units; its accuracy depends on the precision with which the distance from Earth to Sun is measured. At a distance of 1 parsec, the parallax is 1 second of arc. The nearest star is approximately 1.3 parsecs distant; the farthest known galaxy is several billion parsecs. *See* PARALLAX (ASTRONOMY). Jesse L. Greenstein

Parsley

A biennial, *Petroselinum crispum*, of European origin belonging to the plant order Umbellales. Parsley is grown for its foliage and is used to garnish and flavor foods. It contains large quantities of vitamins A and C and has been grown for 2000 years or more. Two types, plain-leafed and curled, are grown for their foliage; Hamburg parsley (*P. crispum* var. *tuberosum*), also called turnip-rooted parsley, is grown for its edible parsniplike root. Propagation is by seed. Harvesting begins 70–80 days after planting for foliage varieties, and 90 days after planting for Hamburg parsley. *See* APIALES. H. John Carew

Parsnip

A hardy biennial, *Pastinaca sativa*, of Mediterranean origin belonging to the plant order Umbellales. The parsnip is grown for its thickened taproot and is used primarily as a cooked vegetable. Propagation is by seed; cultural practices are similar to those used for carrot, except that a longer growing season is required. Parsnip seed retains its viability only 1–2 years. Harvesting begins in late fall or early winter, usually 100–125 days after planting. Exposure of mature roots to low temperatures, not necessarily freezing, improves the quality of the root by favoring the conversion of starch to sugar. *See* APIALES; CARROT. H. John Carew

Partial differentiation

A mathematical operation performed on functions of more than one variable. In this article only two or three variables are considered; however, the principles apply to functions of n variables, for any positive integer $n > 1$. If $z = f(x,y)$, the partial derivative $\partial z/\partial x$ is defined as the derivative of $f(x,y)$ with respect to x, with y being regarded as fixed; that is,

$$\frac{\partial z}{\partial x} = \lim_{h \to 0} \frac{f(x+h, y) - f(x, y)}{h}$$

Another notation for $\partial z/\partial x$ is $f_1(x,y)$. The other first partial derivative is $\partial z/\partial y$, also written $f_2(x,y)$. For values at particular points the notation is

$$\left(\frac{\partial z}{\partial x}\right)_{(a,b)} = f_1(a, b)$$

In the case of a function of three variables, $f(x,y,z)$, the expression is

$$\frac{\partial f}{\partial z} = f_3(x, y, z)$$

The second derivatives of $f(x,y)$ are given by Eqs. (1). It can happen that $f_{12}(x,y) \neq f_{21}(x,y)$, but

$$f_{11}(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \quad f_{12}(x,y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$
$$f_{21}(x,y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \quad f_{22}(x,y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \tag{1}$$

this will not happen in common practice, especially with elementary functions. If f_1, f_2, f_{12}, f_{21} are defined in a neighborhood of (a,b), and if f_{12}, f_{21} are continuous at (a,b), then $f_{12}(a,b) = f_{21}(a,b)$. In addition, there are more delicate theorems relating to this matter.

Differentials. The notions of a differential, and of the differentiability of a function, are fundamental in the theory of partial differentiation. The requirement that $f(x,y)$ be differentiable is not the same as the requirement that $f_1(x,y)$ and $f_2(x,y)$ both exist; it is a more inclusive requirement. The geometric meaning of f being differentiable at (a,b) is that the surface defined by $z = f(x,y)$ has a tangent plane not parallel to the z axis when $x = a$, $y = b$. In analytic terms the condition is that if

$$\epsilon = f(a+h, b+k) - f(a, b) - f_1(a,b)h - f_2(a,b)k$$

then
$$\lim_{(h,k)\to(0,0)} \frac{\epsilon}{|h| + |k|} = 0$$

When f is differentiable at (a,b), the expression $f_1(a,b)\,dx + f_2(a,b)\,dy$ is called the differential of f at (a,b) with independent increments dx and dy. It is a linear function of dx and dy, and among all linear functions $A\,dx + B\,dy$, it is the best approximation (in a definite sense of the word) to the expression

$$f(a+dx, b+dy) - f(a,b)$$

In the usual notation $z = f(x,y)$, the differential, evaluated at (x,y), is written as Eq. (2). Here dx and dy

$$dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial}{\partial y}\,dy \tag{2}$$

are independent variables and dz is a dependent variable.

A sufficient condition that f be differentiable at (a,b) is that the partial derivatives f_1, f_2 be defined at all points near (a,b), and continuous at (a,b).

Chain rule. The prime importance of the differentiability concept is that the differentiability property is needed in proving the chain rule for functions of several variables. This rule asserts that a differentiable function of a differentiable function is differentiable, and the rule tells how to compute partial derivatives of the composite function. For example, if $x = f(s,t)$, $y = g(s,t)$, where f and g are differentiable, and if $z = F(x,y)$, where F is differentiable, then the composite function is $G(s,t) = F[f(s,t), g(s,t)]$, and its differential is

$$\frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy$$

where dx and dy, instead of being independent, are given by Eqs. (3), where ds and dt are independent.

$$dx = \frac{\partial f}{\partial s}\,ds + \frac{\partial f}{\partial t}\,dt \qquad dy = \frac{\partial g}{\partial s}\,ds + \frac{\partial g}{\partial t}\,dt \tag{3}$$

Then $z = G(s,t)$ is differentiable as a function of s and t, and

$$\frac{\partial G}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial f}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial g}{\partial s} \qquad \frac{\partial G}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial f}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial g}{\partial t}$$

These equations, expressing the formal part of the chain rule, are often written in the form

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \qquad \frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial}{\partial y}\frac{\partial y}{\partial t}$$

At one occurrence, the status of x and y is that of independent variables, as in $z = F(x,y)$, where they are called variables of the first class. But x, y also occur as dependent variables, depending on the independent variables s, t, which are called variables of the second class.

The chain rule is valid for situations in which there are any number of variables of the first class and, quite independently, any number (the same or different) of the second class.

A typical use of the chain rule occurs when transformations are made on the variables in a problem. For example, one may switch from rectangular to polar coordinates. Then derivatives with respect to x, y, or both, must be converted into expressions involving derivatives with respect to r and θ. Transformations of variables are quite extensively used in studying partial differential equations.

Another interesting instance of the chain rule occurs in the so-called particle differentiation in the flow of fluids. If $\rho = F(x,y,z,t)$ is the density at (x,y,z) in the fluid at time t, and if in a given motion one follows a certain selected particle, denoting the density of the fluid at this particle by $\rho = G(t)$, then

$$G'(t) = \frac{d\rho}{dt} = \frac{\partial \rho}{\partial x}\frac{dx}{dt} + \frac{\partial \rho}{\partial y}\frac{dy}{dt} + \frac{\partial \rho}{\partial z}\frac{dz}{dt} + \frac{\partial \rho}{\partial t}$$

Here ρ has two different meanings: On the left $\rho = G(t)$, and on the right $\rho = F(x,y,z,t)$. In dx/dt, dy/dt, dz/dt, the point (x,y,z) is the position of the particle being followed. Here the variables of the first class are x, y, z, t, and there is just one variable of the second class, namely, t. The role of t also is different on the left and on the right in the equation.

Taylor developments. There is a Taylor's formula with remainder and a Taylor's series of functions of several variables. The easiest way to deal with these things is to think of them as being reduced back to the case of one variable by a device. If one wants to express $f(a+h, b+k)$ by a formula proceeding by terms of various degrees in h and k, consider $g(t) = f(a+th, b+tk)$, develop $g(t)$ in powers of t, and then set $t = 1$. The chain rule is needed to compute the derivatives of g. The general formula is Eq. (4). Here

$$g^n(t) = \left[\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(x,y)\right]_{\substack{x=a+th,\\ y=b+tk}} \tag{4}$$

a symbolic notation with rather evident meaning is used on the right.

Implicit functions. Suppose $F(x,y,z)$ is a function of three variables whose domain of definition is a certain collection of points (x,y,z) in space of three dimensions. As a general rule, it will not be the case that the locus of points for which $F(x,y,z) = 0$ is the graph of an equation $z = f(x,y)$, where f is a single-valued function of two variables. But it may happen that, if (x_0,y_0,z_0) is a point of the locus $F(x,y,z) = 0$, there is a neighborhood of (x_0,y_0,z_0), consisting of all points inside a certain rectangular box centered at (x_0,y_0, z_0), such that the part of the locus $F(x,y,z) = 0$ inside this box is the graph of a function $z = f(x,y)$. There is a standard "implicit function theorem" which covers this situation. It states: Suppose F and its first partial derivatives F_1, F_2, F_3 are continuous throughout some specified neighborhood N of (x_0,y_0,z_0). Suppose also that $F(x_0,y_0, z_0) = 0$ and $F_3(x_0, y_0z_0) \neq 0$. Then there exist certain positive constants a, b, c and a function f of x and y meeting all the following conditions. Let B denote the boxlike region composed of all (x,y,z) such that $|x-x_0| < a$, $|y-y_0| < b$, $|z-z_0| < c$, and let R denote the rectangle in the xy plane composed of all (x,y_0) such that $|x-x_0| < a$, $|y-y_0| < b$. The region B is contained in N; the function f is defined in R, and the graph of $z = f(x,y) = 0$; f is continuous and has continuous first partial derivatives in R, given by

$$f_1(x, y) = -\frac{F_1(x, y, z)}{F_3(x, y, z)} \qquad f_2(x, y) = -\frac{F_2(x, y, z)}{F_3(x, y, z)}$$

where $z = f(x,y)$.

This theorem has two kinds of generalizations: One of the type in which F is a function of n variables and f is a function of $n - 1$ variables, and the other of the type in which the locus $F(x,y,z) = 0$ is replaced by a locus defined by k equations in n variables $(n > k)$, while the equation $z = (x,y)$ is replaced by k equations involving k functions of $n-k$ variables. Sample: $F(x,y,z,u, v) = 0$, $G(x,y,z,u, v) = 0$, $u = f(x,y,z)$, $v = g(x,y,z)$. Implicit function theorems of this second type are proved by mathematical induction with respect to k. The conditions in these theorems involve jacobian determinants.

Jacobians. If $F_1, \ldots, F_k$ are k functions of $z_1, \ldots, z_k$, determinant (5)

$$J = \begin{vmatrix} \dfrac{\partial f_1}{\partial z_1} & \dfrac{\partial f_1}{\partial z_2} & \cdots & \dfrac{\partial f_1}{\partial z_k} \\ \dfrac{\partial f_2}{\partial z_1} & \cdots & & \dfrac{\partial f_2}{\partial z_k} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial f_k}{\partial z_1} & \cdots & & \dfrac{\partial f_k}{\partial z_k} \end{vmatrix} \tag{5}$$

is called the jacobian of $F_1, \ldots, F_k$ with respect to $z_1, \ldots, z_k$, and is denoted by

$$J = \frac{\partial(F_1, \ldots, F_i)}{\partial(z_1, \ldots, z_k)}$$

Notice that the subscripts on the F's are for distinguishing different functions, and do not indicate partial derivatives.

The general implicit function theorem for a system of equations

$$F_1(x_1, \ldots, x_r, z_1, \ldots, z_k) = 0, \ldots, F_k(x_1, \ldots, x_r, z_1, \ldots, z_k) = 0$$

guarantees a local solution of the form

$$z_1 = f_1(x_1, \ldots, x_r) \cdots z_k = f_k(x_1, \ldots, x_r)$$

near a set of values $x_1 = a_i$, $z_j = b_j$ for which $F_1 = \cdots = F_k = 0$ and $J \neq 0$. This is on the assumption that the F's have continuous first partial derivatives. The derivatives of the f's are given by the formulas

$$\frac{\partial f_i}{\partial x_p} = -\frac{J_{ip}}{J}$$

where J_{ip} is what J becomes when its ith column is replaced by $\partial F_1/\partial x_p, \ldots, \partial F_k/\partial x_p$.

If $u = f(x,y)$, $v = g(x,y)$ defines a one-to-one mapping of a region R_1 of the xy plane onto a region R_2 of the uv plane, and if it is known that f and g have continuous partial derivatives and the jacobian $j = [\partial(f,g)/\partial(x, y)]$ is never zero in R_1, then a simple closed curve C_1 and R_1 maps onto a simple closed curve C_2 in R_2 and, as a point P_1 goes counterclockwise around C_1, its image P_2 goes counterclockwise or clockwise around C_2 according to whether $J > 0$ or $J < 0$. Also, if A_1, and A_2 are the areas enclosed by C_1 and C_2 respectively, there is some point inside C_1 such that A_2/A_1 is the value of $|J|$ at that point. If a double integral with respect to u and v, over the region R_2, is converted into a double integral with respect to x and y, over the region R_1, $du\ dv$ is replaced by $|J|dx\ dy$. These results generalize to the case of mappings in space of more than two dimensions. For example, in the passage from rectangular coordinates x, y, z to spherical polar coordinates r, θ, ϕ, by the equations

$$x = r \sin\phi \cos\theta$$
$$y = r \sin\phi \sin\theta$$
$$z = r \cos\phi$$

the jacobian $\partial(x,y,z)/\partial(r,\theta,\phi)$ has the value $-r^2 \sin \phi$, and $dx\ dy\ dz$ is replaced by $r^2 \sin \phi\ dr\ d\theta\ d\phi$ in triple integrals.

If the equations $u = f(x,y)$, $v = g(x,y)$ define a one-to-one mapping from the xy plane to the uv plane (in restricted regions), then

$$\frac{\partial(u, v)}{\partial(x, y)} = \left[\frac{\partial(x, y)}{\partial(u, v)}\right]^{-1}$$

Functional dependence. If $f(x,y)$ and $g(x,y)$ are functionally dependent in a region R of the xy plane, then $[\partial(f,g)/\partial(x,y)] = 0$ in that region. An example of functional dependence would be: $g(x,y) = [f(x,y)]^2 + \sin [f(x,y)]$. In general, f and g are called functionally dependent in R if there is some function F of u and v such that $F[f(x,y), g(x,y)] = 0$ at all points

of R, and yet $F(u,v)$ is not zero throughout any two-dimensional portion of the uv plane. Conversely, if $[\partial(f,g)/\partial(x,y)] = 0$ at all points (x,y) in a neighborhood of (x_0,y_0), then usually f and g are functionally dependent in some (perhaps smaller) neighborhood of the point.

Homogeneous functions. One calls a function $F(x_1, \ldots, x_k)$ positively homogeneous of degree n if $F(tx_1, \ldots, tx_k) = t^n F(x_1, \ldots, x_k)$ for all $t > 0$ and for all $(x_1, \ldots, x_k)$ in the domain of definition of F. The index n need not be an integer. If F is differentiable and positively homogeneous of degree n, the Euler relation

$$x_1 \frac{\partial f}{\partial x_1} + \cdots + x_k \frac{\partial f}{\partial x_k} = nF(x_1, \ldots, x_k)$$

holds. Conversely, if F is differentiable in an open region which contains $(tx_1, \ldots, tx_k)$ for all $t > 0$, provided it contains $(x_1, \ldots, x_k)$, then the validity of Euler's relation in the region implies that F is positively homogeneous in degree n.

Lagrange's method in extremal problems. If $F(x,y,z)$ is a differentiable function of three independent variables in an open region R of (x,y,z) space, and if F reaches a relative maximum or minimum value at a point of R, then necessarily $(\partial F/\partial x) = (\partial F/\partial y) = (\partial F/\partial z) = 0$ there. Sufficient conditions, and tests for discrimination between maximum and minimum values, are sometimes stated in terms of second partial derivatives. Lagrange's method is concerned with the situation in which x, y, z are not independent, but are restricted by a side condition $G(x,y,z) = 0$, where G is a specified function. Example: What is the maximum value of $x^2y^2z^2$ subject to the restriction $(x^2/25) + (y^2/16) + (z^2/9) - 1 = 0$? On the assumption that F and G have continuous first partial derivatives and that one never has $G = G_1 = G_2 = G_3 = 0$ at one point, the Lagrange procedure is to set $u = F + \lambda G$, where λ is a parameter. Then, among all the values of F attained for (x,y,z) such that $G(x,y,z) = 0$, if there is a maximum or minimum value, it will occur for an (x,y,z) point which satisfies the equations $F_i + \lambda G_i = 0$ $(i = 1, 2, 3)$ and $G = 0$, for a certain value of λ. These four equations can in theory be solved for x, y, z, λ, and the extreme value can be located. The Lagrange method extends to other numbers of variables and to more than one side condition. *See* CALCULUS; DETERMINANT; DIFFERENTIATION; OPERATOR THEORY; PARAMETRIC EQUATION. Angus E. Taylor

Bibliography. T. M. Apostol, *Mathematical Analysis,* 2d ed., 1974; R. Courant, *Differential and Integral Calculus,* 2 vols., 1936–1937, reprint 1992; W. Fulks, *Advanced Calculus,* 4th ed., 1988; W. Kaplan, *Advanced Calculus*, 4th ed., 1991; A. E. Taylor and R. W. Mann, *Advanced Calculus,* 3d ed., 1983.

Particle accelerator

An electrical device that accelerates charged atomic or subatomic particles to high energies. The particles may be charged either positively or negatively. If subatomic, the particles are usually electrons or protons and, if atomic, they are charged ions of various elements and their isotopes throughout the entire periodic table of the elements. Before the advent of accelerators, the only sources of energetic particles for research were cosmic rays, or the naturally occurring radioactive atoms that emit subatomic particles such as electrons and alpha particles at various energies ranging from kilovolts to over 8 MeV. *See* RADIOACTIVITY.

The energy of an accelerated atomic or subatomic particle is usually expressed in units of electronvolts (eV). An electronvolt is the amount of energy that a particle with unit charge, such as an electron or proton, receives when accelerated through a potential difference of 1 V. Therefore, when a proton is accelerated with an electrostatic accelerator operating at 10^6 V, its energy will be 10^6 eV; equivalently, if the particle has q units of charge, its resultant energy in this example will be q MeV. Commonly used multiple units are kiloelectronvolts (keV), megaelectronvolts (MeV), gigaelectronvolts (GeV), and teraelectronvolts (TeV). *See* ELECTRONVOLT.

Very high energy atomic and subatomic particles are accelerated in outer space, and constantly bombard the Earth in the form of cosmic rays ranging in energy from hundreds of kiloelectronvolts to hundreds of thousands of megaelectronvolts. These natural sources of particles are uncontrollable and must be used for research in whatever energy, intensity, and kind of natural radiation is available. Particle accelerator devices, on the other hand, allow these same particles to be accelerated to precise energies with complete control of the intensity and energy over wide ranges. The particles may also be directed to collide with specific target materials in whatever way desired so as to greatly expand knowledge of the fundamental interactions of charged particles with each other and with other materials. Moreover, with accelerators it is possible to produce a wide variety of secondary beams of exotic particles with such short half-lives that they would otherwise be unavailable for controlled experimentation. *See* COSMIC RAYS; ELEMENTARY PARTICLE.

Accelerators that produce various subatomic particles at high intensity have many practical applications in industry and medicine as well as in basic research. Electrostatic generators, pulse transformer sets, cyclotrons, and electron linear accelerators are used to produce high levels of various kinds of radiation that in turn can be used to polymerize plastics, provide bacterial sterilization without heating, and manufacture radioisotopes which are utilized in industry and medicine for direct treatment of some illnesses as well as research. They can also be used to provide high-intensity beams of protons, neutrons, heavy ions, pi mesons, or x-rays that are used for cancer therapy and research. The x-rays used in industry are usually produced by arranging for accelerated electrons to strike a solid target. However, with the advent of electron synchrotron storage rings that produce x-rays in the form of synchrotron radiation, many new industrial applications of these x-rays have been realized, especially in the field of solid-state

microchip fabrication and medical diagnostics. *See* ISOTOPIC IRRADIATION; RADIATION BIOLOGY; RADIATION CHEMISTRY; RADIOACTIVITY AND RADIATION APPLICATIONS; RADIOGRAPHY; RADIOISOTOPE; RADIOLOGY; SYNCHROTRON RADIATION.

Particle accelerators fall into two general classes—electrostatic accelerators that provide a steady dc potential, and varieties of accelerators that employ various combinations of time-varying electric and magnetic fields.

Electrostatic accelerators. Electrostatic accelerators in the simplest form accelerate the charged particle either from the source of high voltage to ground potential or from ground potential to the source of high voltage. The high-voltage dc potential may be either positive or negative, and consequently positive or negative particles will be accelerated by being either attracted to or repelled from the high voltage.

All particle accelerations are carried out inside an evacuated tube so that the accelerated particles do not collide with air molecules or atoms and may follow trajectories characterized specifically by the electric fields utilized for the acceleration. Usually the evacuated tube is provided with a series of electrodes arranged to have gradually increasing potentials from ground to the maximum high-voltage potential. In this way, the high voltage is distributed uniformly along the acceleration tube, and the acceleration process is thereby simplified and better control permitted. The maximum energy available from this kind of accelerator is limited by the ability of the evacuated tube to withstand some maximum high voltage. This limitation in energy was a severe problem in the early days of nuclear and atomic research, when only electrostatic accelerators were available, and led to the development of the second kind of accelerator, which uses time-varying electric or magnetic fields, or both, and is not restricted by any particular limit of potential that can be maintained.

Time-varying field accelerators. In contrast to the high-voltage-type accelerator which accelerates particles in a continuous stream through a continuously maintained increasing potential, the time-varying accelerators must necessarily accelerate particles in small discrete groups or bunches. Since the voltage on any given electrode is varying in time, at certain times the voltage will be suitable for acceleration, while at other times it would actually decelerate the particles. For this reason the electrodes must be arranged so that the particle bunches appear in their vicinity only when the voltage is correct for acceleration.

Linear accelerators. An accelerator that varies only in electric field and does not use any magnetic guide or turning field is customarily referred to as a linear accelerator or linac. In the simplest version of this kind of accelerator, the electrodes that are used to attract and accelerate the particles are connected to a radio-frequency (rf) power supply or oscillator so that alternate electrodes are of opposite polarity. In this way, each successive gap between adjacent electrodes is alternately accelerating and decelerating. If these acceleration gaps are appropriately spaced to accommodate the increasing velocity of the accelerated particle, the frequency can be adjusted so that the particle bunches are always experiencing an accelerating electric field as they cross each successive gap. In this way, modest voltages can be used to accelerate bunches of particles indefinitely, limited only by the physical length of the accelerator construction.

All conventional (but not superconducting) research linacs usually are operated in a pulsed mode because of the extremely high rf power necessary for their operation. The pulsed operation can then be adjusted so that the duty cycle or amount of time actually on at full power averages to a value that is reasonable in cost and practical for cooling. This necessarily limited duty cycle in turn limits the kinds of research that are possible with linacs; however, they are extremely useful (and universally used) as pulsed high-current injectors for all electron and proton synchrotron ring accelerators. Superconducting linear accelerators have been constructed that are used to accelerate electrons and also to boost the energy of heavy ions injected from electrostatic machines. These linacs can easily operate in the continuous-wave (cw) rather than pulsed mode, because the rf power losses are only a few watts.

The Continuous Electron Beam Accelerator Facility (CEBAF) uses two 400-MeV superconducting linacs to repeatedly accelerate electrons around a racetracklike arrangement where the two linacs are on the opposite straight sides of the racetrack and the circular ends are a series of recirculation bending magnets, a different set for each of five passes through the two linacs in succession. The continuous electron beam then receives a 400-MeV acceleration on each straight side or 0.8 GeV per turn, and is accelerated to a final energy of 4 GeV in five turns and extracted for use in experiments. The superconducting linacs allow for continuous acceleration and hence a continuous beam rather than a pulsed beam. This makes possible many fundamental nuclear and quark structure measurements that are impossible with the pulsed electron beams from conventional electron linacs. *See* SUPERCONDUCTING DEVICES.

Circular accelerators. As accelerators are carried to higher energy, a linac eventually reaches some practical construction limit because of length. This problem of extreme length can be circumvented conveniently by accelerating the particles in a circular path maintained by either static or time-varying magnetic fields. Accelerators utilizing steady magnetic fields as guide paths are usually referred to as cyclotrons or synchrocyclotrons, and are arranged to provide a steady magnetic field over relatively large areas that allow the particles to travel in an increasing spiral orbit of gradually increasing size as they increase in energy. After many accelerations through various electrode configurations, the particles eventually achieve an orbit as large as the maximum size of the magnetic field and are extracted for use in research and other kinds of applications. *See* MAGNETISM.

A special kind of circular accelerator for electrons using a static magnetic field is called a microtron. A constant frequency in a small acceleration gap or cavity near the edge of a circular-shaped uniform magnetic field is arranged so that it accelerates the electrons tangentially. Electrons are provided at an energy of one or a few megaelectronvolts so that they are moving close to the velocity of light. This means that when the electrons are accelerated their velocity can increase only slightly because the velocity of light cannot be exceeded according to the laws of special relativity. The acceleration and increase in energy are consequently accomplished by increasing their mass instead of their velocity. With constant velocity, the acceleration frequency can also be constant and arranged in such a way that the circular orbits increase in circumference by one or more beam bunches or wavelengths on each revolution of larger diameter. Much larger microtrons in a racetracklike configuration of two opposed halfmoon-shaped magnets have been built with an electron linac between them and aligned with the outermost racetrack orbit.

Practical limitations of magnet construction and cost have kept the size of circular proton accelerators with static magnetic fields to the vicinity of 100 to 1000 MeV. Most circular accelerators can easily operate over a range of 10 to 1 and some as high as 30–50 to 1. However, increasing the size and coupling of smaller to larger machines is necessary for larger ranges in energy. For even higher energies, up to 400 GeV per nucleon in the largest conventional (not superconducting) proton synchrotron in operation, it is necessary to vary the magnetic field as well as the electric field in time. In this way the magnetic field can be of a minimal practical size, which is still quite extensive for a 980-GeV accelerator (6500 ft or 2000 m in diameter). This circular magnetic containment region, or "racetrack," is injected with relatively low-energy particles that can coast around the magnetic ring when it is at minimum field strength. The magnetic field is then gradually increased to stay in step with the higher magnetic rigidity of the particles as they are gradually accelerated with a time-varying electric field. Again, when the particles achieve an energy corresponding to the maximum magnetic field possible in the circular guide field, they are extracted for utilization in research programs, and the magnetic field is cycled back down to its low or near-zero value and the acceleration process is repeated.

Although electron synchrotrons were formerly used in high-energy physics research and as synchrotron radiation sources, they are now used exclusively as injectors for storage rings that provide the research capabililty.

Focusing. One of the chief problems in any accelerator system is that of maintaining spatial control over the beam; this requires some form of focusing. In addition, the presence of focusing elements makes it possible to trade off beam cross-sectional area and angular divergence as the specific utilization may require.

A wide variety of focusing principles and devices are used in different accelerator types. The natural bowing-out of magnetic field lines in a cyclotron magnet results in a net focusing action as will be discussed below; so also does the electrostatic field line configuration between drift tubes of many simple linear accelerators. But these systems are classed generally as weak-focusing approaches. The particle beams under their control can make relatively large excursions from the desired equilibrium orbits with the consequence that large vacuum envelopes are essential if the beams are not to be lost through collision with the envelope walls.

Fortunately, an alternate strong-focusing approach, first described in 1952, makes possible very substantial improvements in this area. It uses electrostatic and magnetic lens systems arranged alternately positive and negative (in the sense of convex and concave optical elements) so that, as in the optical analog, they have a net- and strong-focusing action on charged particle beams. This alternation of positive and negative elements has been implemented in many ways, with many devices and for a variety of geometries. Perhaps its most frequent appearance is in quadrupole (four-pole) or hexapole (six-pole) magnets used as variable-focal-length, variable-astigmatism elements in beam transport both within large accelerators themselves and in extensive beam-transport systems external to the accelerator.

Superconducting magnets. The study of the fundamental structure of nature and all associated basic research require an ever increasing energy in order to allow finer and finer measurements on the basic structure of matter. Since the voltage-varying and magnetic-field-varying accelerators also have limits to their maximum size in terms of cost and practical construction problems, the only way to increase particle energies even further is to provide higher-varying magnetic fields through superconducting magnet technology, which can extend electromagnetic capability by a factor of 4 to 5. Large superconducting cyclotrons and superconducting synchrotrons are in operation. *See* MAGNET.

Storage rings. Beyond the limit just described, the only other possibility is to accelerate particles in opposite directions and arrange for them to collide at certain selected intersection regions around the accelerator. The main technical problem is to provide adequate numbers of particles in the two colliding beams so that the probability of a collision is moderately high. This intensity is usually expressed as luminosity and is in units of number of particles per square centimeter per second. Such storage ring facilities are in operation for both electrons and protons. Besides storing the particles in circular orbits, the rings can operate initially as synchrotrons and accelerate lower-energy injected particles to much higher energies and then store them for interaction studies at the beam interaction points.

Large proton synchrotrons have been used as storage-ring colliders by accelerating and storing protons in one direction around the ring while accelerating and storing antiprotons (negative charge) in the opposite direction. The proton and antiproton beams are carefully programmed to be in different orbits as they circulate in opposite directions and to

collide only when their orbits cross at selected points around the ring where experiments are located. The antiprotons are produced by high-energy proton collisions with a target, collected, stored, cooled, and eventually injected back into the synchrotron as an antiproton beam. It takes 24–36 h to collect and prepare an antiproton beam for injection and use in the collider mode of operation, and the intensity is limited ($^1/_{1000}$ or less than the proton beam) because of the difficulty in producing an antiproton beam.

Electron-positron synchrotron accelerator storage rings have been in operation for many years in the basic study of particle physics, with energies ranging from 2 GeV + 2 GeV to 104 GeV + 104 GeV. The byproduct synchrotron radiation from many of these machines is used in numerous applications. In fact, the production of synchrotron radiation is proportional to the rf power necessary to keep the circulating stored beam on a stable orbit. This synchrotron radiation loss forces the machine design to larger and larger diameters, characterized by the Large Electron Positron Storage Ring (LEP) at CERN, near Geneva, Switzerland (closed down in 2000), which was 17 mi (27 km) in circumference. Conventional rf cavities enable electron-positron acceleration only up to 50–70 GeV (limited by synchrotron radiation loss) while higher energies of 100–150 GeV require superconducting cavities.

Advanced linacs. Although circular machines with varying magnetic fields have been developed because linacs of comparable performance would be too long (many miles), developments in linac design and utilization of powerful laser properties may result in a return to linacs that will outperform present ring machines at much lower cost. As a first example, the 20-GeV electron linac at Stanford University, Palo Alto, California, has been modified to provide simultaneous acceleration of positrons and electrons to energies as high as 50 GeV, while operating in what is called the SLED mode. After acceleration the electrons and positrons are separated by a magnet, and the two beams are magnetically directed around the opposite sides of a circle so that they collide at one intersection point approximately along a diameter extending from the end of the linac across the circle. This collider arrangement is much less expensive than the 17-mi (27-km) ring at CERN and provides electron-positron collisions of comparable energies but at lower intensities.

Performance characteristics. This introduction has shown that there are many kinds of accelerators, varying in size and performance characteristics and capable of accelerating many kinds of particles at varying intensities and energies, in the form of bunched or pulsed beams of particles or as a steady dc current (**Table 1**).

Electrostatic Accelerators

Electrostatic accelerators are used to accelerate atomic or subatomic particles to high energies by means of a high voltage potential that is maintained through some electrical or mechanical means of transporting charge from ground to the high voltage potential. Modern machines of larger size all use mechanical charge transport, but electrical transport systems are widely used in small Cockcroft-Walton or dynamitron-type electrostatic accelerator units. The most commonly used electrostatic generator, invented by R. J. Van de Graaff, utilizes a high-speed rubberized-fabric belt to transport the charge. Many machines utilizing this principle have been homemade or built by various research groups throughout the world.

An improved charging system, the Pelletron® electrostatic accelerator, manufactured by National Electrostatics Corporation, utilizes a chain-charging system instead of the rubberized-fabric belt. The continuous-looped charging chain is fabricated of alternate metal and insulated links and travels at high speed on special pulleys from ground to high potential, an arrangement mechanically similar to the charging belt system. Instead of the charge being sprayed onto the charging belt system from sharp corona points, each metal link is charged and discharged by induction methods so that no electrical sparking or erosion is involved.

Most electrostatic accelerators are housed inside a large high-pressure vessel that is filled to a pressure as high as 15 atm (220 lb/in.^{2} or 1.5×10^6 pascals) with very dry insulating gas. The insulating gas may be pure sulfur hexafluoride (SF_6) or a mixture with carbon dioxide (CO_2) and nitrogen (N_2). This high-pressure insulating gas allows the machine to be housed in a much smaller space than would be possible in air at atmospheric pressure. *See* ELECTRICAL INSULATION.

Basic design. The basic high-voltage electrostatic generator may be considered in the form of a van de Graaff high-voltage generator (**Fig. 1**) in which electric charge is carried on an insulated belt from

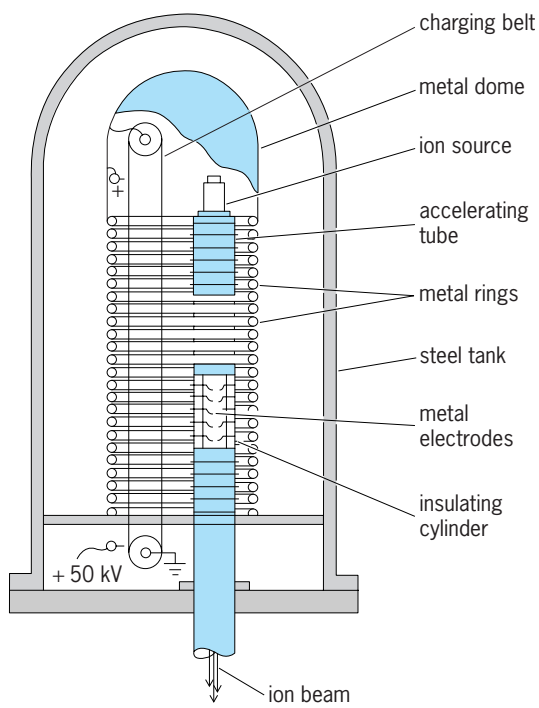

Fig. 1. Van de Graaff generator for operation in high-pressure gas.

TABLE 1. Operating characteristics of particle accelerators

Accelerator type	Particle accelerated	Energy range*	Beam current (average; peak) or beam intensity (luminosity)	Duty cycle	Energy spectrum†	Beam geometry	Development status (2006)
ELECTROSTATIC ACCELERATORS							
Tandetron	p, d, α, heavy ions	To 5 MV	1 μA; 10 μA	Continuous	~0.01%	Small focal spot	1–5 MV operating
Cockcroft-Walton	p, d, α, e, heavy ions	To 4 MV	1 mA; 10 mA	Continuous	~0.01%	Small focal spot	4 MeV operating
Dynamitron	p, d, α, e, heavy ions	To 4 MV	1 μA; 50 mA	Continuous	~0.01%	Small focal spot	4 MV operating
Tandem Van de Graaff	p, d, α, e, heavy ions	To 20 MV	1 μA; 50 μA	Continuous	~0.01%	Small focal spot	20 MV operating
Tandem pelletron	p, d, α, e, heavy ions	To 26 MV	1 μA; 50 μA	Continuous	~0.01%	Small focal spot	26 MV operating
TIME-VARYING FIELD ACCELERATORS							
Circular magnetic types (radio-frequency resonance accelerators)							
Microtron	e^-	To 855 MeV	10 μA; 100 mA	0.1%	0.1%	Small focal spot	30–855 MeV operating
Sector or isochronous cyclotron	p, d, α, heavy ions	To 590 MeV (p); 90 MeV/u	20 μA; 2 mA	Continuous	~0.01%	Internal target or external beam with small focal spot	590 MeV (p), 90 MeV/u (heavy ions) operating
Superconducting cyclotron	Heavy ions	345 MeV/u	1 μA; 100μA	Continuous	~0.01%	Internal or external beam with small focal spot	345 MeV/u operating
Synchrotron (weak focusing)	p, e, heavy ions	1–6 GeV (p); 2 GeV/u	0.1 μA; depends on duty cycle	30%	0.1%	Internal targets; external beam of fair collimation at lower intensity	Mostly closed down
Alternating-gradient synchrotron	p, e^+, e^+	10–980 GeV (p)	1.0 μA (p)	30%	0.1%	(p) internal targets; external beam of fair collimation	980 GeV (p) operating (superconducting); heavy-ion machines operating; ^{197}Au and ^{208}Pb beams operating
	Heavy ions: mass 12–197	11.4 GeV/u	10^8–10^{10} ions/s	30%	0.1%		
	mass 12–208	160 GeV/u	10^8–10^{10} ions/s	30%	0.1%		
Linear accelerators							
Heavy-ion linear accelerator	p, d, α, heavy ions	To 30 MeV/u bbits)	10 μA; 130 mA	~10%	0.5%	Well-collimated and well-focused external beam	30 MeV/u (heavy ion) operating (normal and heavy superconducting)
Linear accelerator	p	50–800 MeV	1 mA; 100 mA	~10%	0.1%	Well-collimated and well-focused external beam	800 MeV (p) operating
TJNAF recirculating superconducting linear accelerator	e^-	0.5–6 GeV	200 μA	Continuous	0.003%	Well-collimated and well-focused external beam	4 GeV operating
Electron linear accelerator	e^+, e^-	6 MeV to 50 GeV	60 μA; 400 μA	~6%	~ 0.2%	Well-collimated and well-focused external beam	50 GeV (SLED mode) closed down
		100 GeV (CM) in e^+, e^- collider mode	10^{29}/cm^{2}s		0.06%		
Colliding-beam storage rings							
Electron-positron storage ring	e^+, e^-	0.3–208 GeV (CM)	~ 10^{31}/cm^2 s	Continuous	0.1%	Small-diameter internal beam	208 GeV (CM) closed down
Proton-proton collider	pp	14 TeV (CM)	~ 10^{34}/cm^2 s	Continuous	0.1%	Small-diameter internal beam	14 TeV (CM) under construction
Proton-antiproton collider	$p\bar{p}$	1.96 TeV (CM)	3×10^{31}/cm^2 s	Continuous	0.1%	Small-diameter internal beam	1.96 TeV (CM) operating
Electron-proton collider	e-p	27 GeV (e), 820 GeV (p)	2×10^{31}/cm^2 s	Continuous	0.1%	Small-diameter internal beam	27 GeV (e), 820 GeV (p) operating
Heavy-ion collider	p, heavy ions	100 GeV/u (^{197}Au), 250 GeV (p)	10^{27}/cm^2 s (^{197}Au), 1.4×10^{31}/cm^2 s	Continuous	0.1%	Small-diameter internal beam	100 GeV/u (^{197}Au) operating

*Voltage range is given for electrostatic accelerators. u = atomic mass unit. CM = center of mass.
†Spread in energy of beam expressed as a percentage of total energy of beam; that is, 1% for the cyclotron means for a 1-MeV beam a spread in energy of 0.01 MeV.

ground to an insulated high-voltage terminal. The high-voltage terminal usually contains an ion source that produces protons, other types of positive ions, or electrons that can then be accelerated by the high voltage of the generator through an evacuated tube—the acceleration tube—back to ground potential. *See* ION SOURCES.

A motor at the base of the machine with a grounded pulley drives either a fabric or rubberized fabric belt over a second pulley that is mounted inside the high-voltage terminal and connected or grounded to it. A comblike set of points is arranged at the base of the machine to spray charge onto the belt, and a similar device removes the charge when it arrives at the high-voltage terminal.

The acceleration tube is fabricated of alternate rings of insulating glass or ceramic and metal electrodes glued or bonded together to provide a hollow tubelike structure, extending from the high-voltage terminal to the base of the machine. Power for the operation of the ion source, which often requires a hot filament and an electric arc power supply, is usually supplied by a generator driven by a belt from the charging pulley. In some designs the generator is inside the charging pulley itself. The acceleration tube is evacuated by vacuum pumps at the base of the machine.

The electrical voltage from the high-voltage terminal to ground is graded in a series of steps corresponding to each of the metal electrodes of the acceleration tube and the basic insulating support column itself. This gradient is provided by resistors strung together in series from the high-voltage terminal to ground, each connected to the metal electrodes and the metal rings. The resistor string then provides a stepped high voltage from the metal dome or terminal of the machine uniformly downward from terminal voltage to ground with each of the metal rings and corresponding electrodes in the tube. In this way, the injected ion from the source is accelerated by a continually accelerating electric field as it moves toward ground potential. After accelerating through the entire gradient of a machine operating at 1 MV, a charged particle has 1 MeV of energy for each unit charge. A similar machine but with a Pelletron charging system instead of a charging belt is called a Pelletron generator.

The main operational difficulty with this design is that the ion source, which is inside the high-voltage terminal, requires regular maintenance, and this in turn requires depressurization of the pressure vessel for maintenance and repressurization for operation. This procedure can take a day or more with the larger machines.

Electrical charging generators. The basic accelerator design with mechanical charging systems (Van de Graaff or Pelletron) is limited to continuous particle beam currents of a few hundred microamperes because of the limited amount of charge that can be transported mechanically to the high-voltage terminal. This limitation prompted the design of electrical charging systems that are capable of transporting a thousand times as much charge or hundreds of milliamperes.

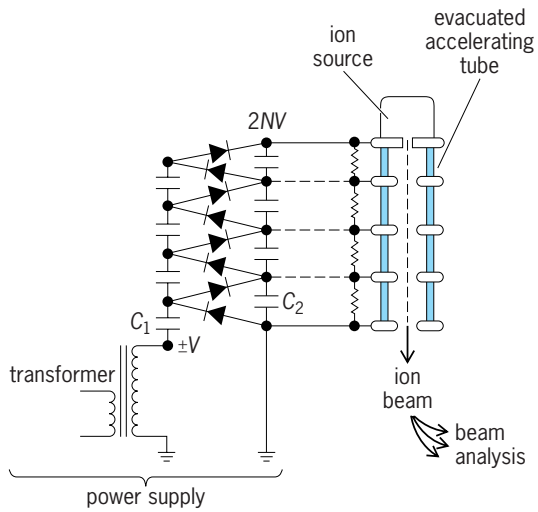

Fig. 2. Cockcroft-Walton power supply and accelerator. The multiplier circuit shown has four stages ($N = 4$).

Cockcroft-Walton accelerator. The original Cockcroft-Walton accelerator transports charge to a high-voltage terminal electrically by using voltage multiplying circuitry. In its most rudimentary form, it contains a high-voltage transformer and a voltage-multiplying circuit (**Fig. 2**). The multiplier circuit consists of two stacks of series-connected capacitors, cross-linked by voltage rectifiers. One set of capacitors is connected to the transformer. During the negative portion of the transformer cycle, the first capacitor in this stack, C_1, becomes charged. During the positive portion, some of this charge is transferred to the first capacitor in the second set, C_2. This process continues, with the transformer effectively pumping charge through the rectifier chain, until all capacitors are charged. Ultimately, the potential difference across each capacitor is $2V$, where V is the peak transformer output voltage, and the voltage across each complete set of N capacitors is $2NV$. Since the second set is referenced to ground, the voltage across it is constant except for ripple caused by power drain. To reduce ripple, alternating voltages at frequencies of several kilohertz are usually supplied to the transformer. Direct-current voltages of either sign can be obtained depending on the orientation of the rectifiers. Modern supplies, which use solid-state rectifiers, have many stages of multiplication with relatively modest transformer voltages. Modern Cockcroft-Walton machines are used mainly as ion injectors for large synchrotron facilities. They are open-air insulated (no pressure vessel) for convenient and rapid ion-source maintenance, and limited to voltages of 0.5–1.0 MV. All of these large dc machines are being replaced by radio-frequency quadrupole accelerators, which are a type of linear accelerator capable of accelerating directly from an ion source to energies suitable for injection into conventional linear accelerators. *See* SEMICONDUCTOR RECTIFIER; VOLTAGE-MULTIPLIER CIRCUIT.

Dynamitron accelerator. Although the Cockcroft-Walton voltage multiplication system provides moderately high voltages (~1 MV) at high currents, it is difficult to go to much higher voltages because the

multiplier arrangement is essentially a series-driven system from ground up to the high voltage. An alternative system, known as the dynamitron accelerator, that operates on a parallel rather than series arrangement can easily drive to much higher voltages (4–5 MV). Most other cascade-type systems operate at low frequency (60–500 Hz) and rely on large energy storage in inductance or capacitance or both. In contrast, the dynamitron uses high-frequency (100 kHz) and low-capacitive coupling with consequently low energy storage, which helps minimize the consequences of component failure.

Two large, semicylindrical rf electrodes operate at approximately 150 kV at 100 kHz, with a resonant inductor and oscillator system (**Fig. 3**). Each split corona shield on the high-voltage column of the accelerator acts as an rf pickup electrode, which in turn drives an individual rectifier. As a pair, the split corona shields also help provide a smooth dc potential distribution, just as do the circular equipotential rings in an electrostatic accelerator.

Modern units use a special solid-state rectifier for rf rectification, which simplifies the design and increases the maximum current capacity. Power is coupled into each rectifier through the shield-to-electrode capacitance.

Low-energy machines. The Cockcroft-Walton series-fed voltage multiplier and the dynamitron parallel-fed voltage multiplier characterize the two main methods of voltage production among a myriad of compact machines. Many machines are small compact accelerators connected to a separate solid-state, high-voltage power supply with a high-voltage cable rated up to 1 MV capacity. Modern production methods require small accelerators for crucial diagnostic capabilities and often in the production process itself.

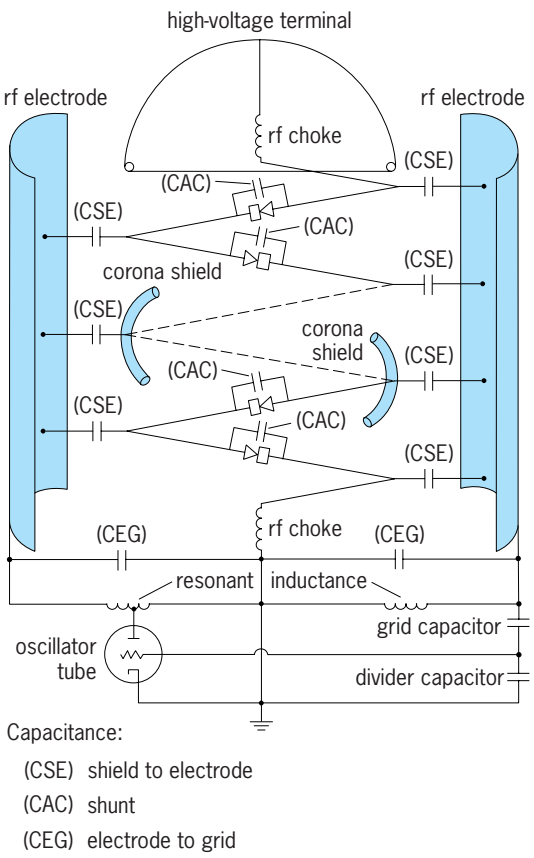

Fig. 3. Radio-frequency coupling and direct-current rectification system utilized in the dynamitron accelerator. (*After R. F. Shea, ed., 1st National Particle Accelerator Conference, IEEE Trans. Nucl. Sci., NS-12(3):227–234, 1965*)

Neutron tube. One interesting example of a tiny accelerator system is the neutron tube accelerator, devised for logging boreholes drilled for oil and mineral prospecting. These accelerators operate at 0.1 MV or less, but this is sufficient for deuterium ions to produce 14-MeV neutrons when they strike a tritium target. The ion source, accelerator, power supplies, and radiation detectors are assembled into a long probe a few inches in diameter which is lowered into the borehole. The accelerator is pulsed on for a few microseconds, during which large numbers of neutrons are produced. These neutrons slow down and are absorbed. The time distribution and types of radiation which result are related to the local nature of the stratum surrounding the borehole.

Tandem electrostatic generator. The tandem electrostatic generator utilizes the basic Van de Graaff generator principle (Fig. 1) in a special form. The pressure vessel is in the form of a long cylinder, and the high voltage terminal is supported in the middle of the cylinder by an insulating column structure extending from both ends of the cylinder. The charging belt operates from the high-voltage terminal to either or both ends of the cylinder, and an acceleration tube passes through from one end of the cylinder to the other so that accelerated particles can pass through the machine. The support column, high-voltage terminal, metal rings, and so forth, are all arranged concentrically about the axis of the cylindrical pressure vessel. This structure is actually two basic electrostatic generators, on the same axis and sharing a common high-voltage terminal.

Negative ions from a negative-ion source outside one end of the cylindrical structure are injected into the acceleration tube at one end of the machine, and then accelerated to the high-voltage terminal. There they pass through a thin foil usually made of carbon approximately 10^{-6} in. (0.02 μm) thick (or a gas stripper), and electrons are stripped away from them, leaving them as positive ions. The particles are then accelerated out of the other end of the acceleration tube just as though they were produced by an ion source in the high-voltage terminal. In this way, a 1-MV machine can provide 2-MeV protons, giving them 1-MeV acceleration as a negative ion and 1 MeV as a positive ion.

Even more advantage is gained when heavy ions of various atomic nuclei are accelerated. The negative ion again achieves 1 MeV acceleration with a 1-MV machine; however, the foil will probably strip the heavy ion to a charge state of two or three units of positive charge, which means it then achieves an additional 2 or 3 MeV of energy, reaching a total energy of 3 or 4 MeV.

The tandem electrostatic generator has been widely used as a research instrument because of the convenience of having the ion source and complicated ion-producing devices of various kinds outside the accelerator at ground potential rather than inside the high-voltage terminal as in the basic electrostatic accelerator.

Folded tandem design. If a shorter tandem electrostatic accelerator is desired, it is possible to combine the two end-to-end machines, making a tandem by folding the two machines together 180° at the high-voltage terminal into a common column structure that now contains two acceleration tubes rather than one. In this way, one tube accelerates the negative ions to the high-voltage terminal, where they are stripped and reversed in direction by a 180° bending magnet and then accelerated back to ground potential by the second acceleration tube.

Large machines. Very large tandem Van de Graaff generators and tandem Pelletron generators have been manufactured commercially, with the largest tandem Van de Graff generator rated at 20 MV, and the largest tandem Pelletron generator rated at 25 MV.

Tandem electrical charging generators. The tandem or two-stage acceleration principle has also been used with the electrical charging generators, specifically the dynamitron and a smaller lower-powered version of the machine called a tandetron. Two 4-MV tandem dynamitrons were built for research use; however, ion accelerators are no longer produced, and only high-current electron accelerators are manufactured for industrial processing. On the other hand, the smaller, lower-powered tandetron is used exclusively for ion acceleration and is widely used in ion implanting, carbon dating, and many different kinds of atomic and solid-state basic research programs. The tandetrons are made in 1-, 2-, and 3-MV sizes and are small enough to fit into most available laboratories. A special 0.7-MV folded version is made for helium backscattering studies and utilizes a specially designed 180° permanent-magnet bending arrangement inside the high-voltage terminal. *See* ION IMPLANTATION; MASS SPECTROSCOPE; RADIOCARBON DATING.

Use as injectors. The ease of energy adjustment, ease of change of ion species, and the high-quality ion beam, which can be easily pulsed to match the input requirements of linacs, cyclotrons, or synchrotrons, make the tandem an ideal injector for other kinds of accelerators, which then may be used to boost the limited energy of the tandem accelerator. Many tandem research laboratories have booster accelerators in the form of either normal or superconducting linacs and cyclotrons.

Applications. Although the original driving need for larger machines in the study of nuclear physics has now largely moved on to energies beyond the practical capability of electrostatic machines, they will continue to support the field as injectors into other machines and be used as powerful tools in the study of atomic, molecular, and solid-state physics. Many application studies and commercial processing operations are carried out with heavy ions from tandem accelerators. The smaller electrostatic accelerators are in constant demand in all fields of research, and industrial applications are expanding on a steady basis. *See* NUCLEAR PHYSICS. Harvey E. Wegner

Circular Accelerators

Circular accelerators utilize a magnetic field to bend charged-particle orbits and confine the extent of particle motion.

Cyclotrons. The cyclotron is a circular accelerator in which the particle orbits start at the center and spiral outward in a guide magnetic field which is constant in time. The basic cyclotron concept, put forward in 1930, is given by Eq. (1), the Lawrence

$$\omega = \frac{qB}{m} \tag{1}$$

equation for the angular frequency of rotation of a particle in a plane perpendicular to a magnetic field. Here, q and m are the charge and mass of the particle and B is the magnetic induction of the perpendicular magnetic field. (SI or mksa units are used throughout this section of the article.) This equation can be derived by noting that (1) $\omega = v/r$, where v is the linear velocity and r is the orbital radius; (2) the magnetic force is perpendicular to the velocity and is given by qvB; (3) this perpendicular force produces a centripetal acceleration and, from Newton's second law, gives $qvB = mv^2/r$. Solving the last equation for v/r leads directly to Eq. (1). The striking statement of Eq. (1) is that if field, charge, and mass are all constant, then angular velocity is likewise constant and, in particular, does not depend on the linear velocity. Faster-moving particles travel on circles of larger radius—the increase in the length of a revolution due to the larger radius just matches the increase in speed, and fast or slow particles thus require the same amount of time to make a 360° revolution. Thus, an accelerating electrode operating at constant frequency can be used. This is the key idea of the Lawrence cyclotron. *See* NEWTON'S LAWS OF MOTION.

In such a cyclotron (**Fig. 4**), a pair of D-shaped accelerating electrodes, called dees, is electrically

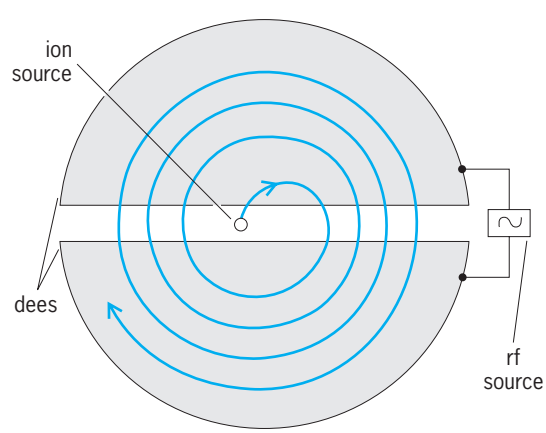

Fig. 4. Principle of the cyclotron. An ion in the accelerator follows the arrowed path. The magnetic field is perpendicular to the page.

attached to an rf voltage source whose frequency is matched to the rotation frequency of the particle in a surrounding magnetic field. Ions are introduced at the center, either from an ion source placed there or from an external ion source, and immediately encounter the high voltage on the facing dee. Those which enter the gap between the two dees at a time when the rf voltage produces an accelerating electric field are accepted for acceleration; the remainder are lost. Beyond the accelerating gap the particle passes inside the dee, which is an electrically shielded region; at the same time it is pulled into a circular path by the perpendicular magnetic field, and after 180° it arrives again at the gap between the dees. Since the frequency of the rf source which drives the dees is selected to match the rotational frequency of the particle, the voltage between the dees will have reversed while the particle is making its 180° turn, and therefore the particle is again accelerated as it crosses from dee to dee. This process can clearly be repeated as often as desired. The particle will thus be repetitively accelerated and, as it speeds up, will gradually move on circles of larger and larger radius. Finally the particle will come to the limit of the magnetic field. At this point an additional electrode—referred to as a deflector—can be inserted to direct all particles out along a single path.

Limitations. This theoretical picture of exactly synchronized orbits and voltage is an oversimplification, and more detailed investigation reveals two potential problems with the cyclotron concept. As any particle speeds up, its total energy E will rise, and therefore, according to the equation $E = mc^2$ of relativistic mechanics, its mass m will increase, the speed of light c being a constant. In a constant magnetic field the gradual increase in mass causes the rotation frequency to steadily decrease as the particle speeds up, and the particle then gradually lags behind a constant rf frequency. A second limitation comes from the magnetic field factor in Eq. (1). To prevent particles from drifting away from the midplane because of their initial velocities there must be a restoring force which pushes particles back toward the plane. In early cyclotrons this was provided by the slight natural decrease in field strength with radius caused by field leakage beyond the poles' outer edges: the field lines bow outward, giving the magnetic force on particles off the midplane a small component back toward it. But the decreasing field also produces a slowing-down according to Eq. (1). This effect adds to that of the mass increase, and the two effects limit the energy of the Lawrence cyclotron. The highest energy achieved by such a cyclotron was 23 MeV for a beam of protons.

Synchrocyclotron. A way to avoid the energy limit on the Lawrence cyclotron is to slow down the frequency of the rf power source at the same rate as the slowing down of the rotational frequency of the particle. But this condition can be maintained only for a reference particle at a particular "synchronous phase" (**Fig. 5**). The usefulness of the process depends on the stability of the motion of neighboring nonsynchronous particles. Will particles which are not exactly at the design values of energy and phase remain close to those values in synchronism with the rf voltage (stability) and accelerate, or diverge from the design values (instability) and not accelerate?

Particles which arrive at the accelerating gap before the synchronous reference particle experience a larger accelerating voltage and energy gain than the reference particle; this causes their mass to increase faster than that of the reference particle, resulting in their revolving more slowly according to Eq. (1). The larger energy gain also causes these early particles to reach a larger radius and thus a lower magnetic field than the reference particle, slowing their revolution even more. They thus gradually lose their early lead and are forced toward the synchronous phase. Similarly, particles which reach the accelerating gap late receive a smaller energy increase, but consequently revolve faster and gradually catch up with the reference particle. This restoring action relative to the phase of the synchronous particle is referred to as phase focusing. It occurs only if the synchronous phase is on the falling side of the rf voltage wave; if it is on the rising side, the phase motion is unstable. The principle of phase stability was discovered toward the end of World War II, and made possible the frequency-modulated cyclotrons or "synchrocyclotrons" that provided the leading advances in ion beam energy in the postwar years. The largest synchrocyclotrons have achieved energies of about 1000 MeV with proton beams; their size is limited by cost and construction difficulty.

In changing from the Lawrence cyclotron to the synchrocyclotron, however, a very valuable beam property is given up, namely, beam intensity. In the Lawrence cyclotron, particles can start their acceleration journey from the ion source on every rf cycle, because every rf cycle is exactly identical. The synchrocyclotron, in contrast, is a batch device. A group of particles leave the source, and the frequency of the accelerating system is then steadily lowered to stay matched with the group while it is being accelerated. During this relatively long period, particles emerging

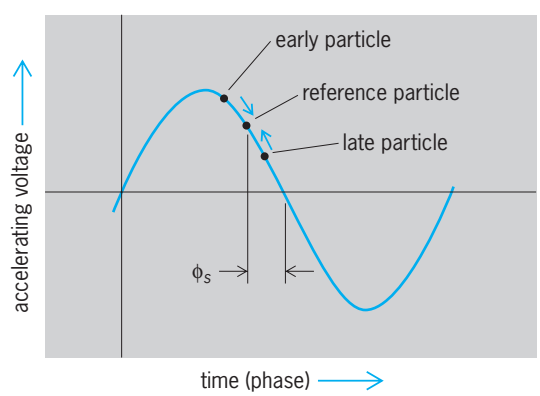

Fig. 5. Principle of phase stability in the synchrocyclotron. The reference particle has resonance energy and synchronous phase, ϕ_s. The early particle accelerates faster to larger radius, where it has a lower rotation frequency which moves it later in time toward the central particle. The late particle is correspondingly moved earlier in time toward the central particle.

from the ion source cannot be accelerated, since the radio frequency of the accelerating system does not match their frequency of revolution in the magnetic field nearby; particles can be "captured" and accelerated again only after the first group has reached full energy and the radio frequency has been returned to its initial value. This results in a large intensity loss. A synchrocyclotron typically accelerates only 1/100 as many particles per second as a Lawrence cyclotron under comparable conditions. Largely because of this low beam intensity, very few synchrocyclotrons remain in use, though their close relatives, FFAG accelerators (discussed below), which are also frequency modulated and pulsed, are experiencing a modern resurgence.

Isochronous cyclotron. Responding to the need for higher beam intensity at energies above 20 MeV, the isochronous (that is, equal-time) cyclotron [sometimes referred to as AVF (azimuthally varying-field) or sector-focused cyclotron] introduced in the 1950s made constant-frequency operation possible at higher energies. Although the concept had been published by L. H. Thomas in 1938, its practical realization was delayed by the complexity of both the magnet design and the associated orbit calculations. The basic idea can be understood by starting with Eq. (1). As the particle speeds up, its mass increases, as discussed earlier, and this will tend to lower its angular frequency ω; however, if the magnetic field B is increased in a compensating way, the frequency can be kept constant, so that pulsed operation is not necessary. Since faster (heavier) particles move in circles of larger radius, a magnetic field is required whose strength increases with the radius—the opposite situation to that found in Lawrence cyclotrons. Such a field can readily be built, but Maxwell's equations then require the field lines to curve toward the axis and hence produce force components pushing particles above or below the mid-plane further away from it. In such an axially defocusing field, very few particles would survive the acceleration process, and so an extra source of axial focusing is required. *See* MAXWELL'S EQUATIONS.

The discussion so far has assumed that the magnetic field is axially symmetrical. If azimuthal variations are allowed, additional terms in Maxwell's equations come into play and can provide additional focusing forces. In practice this is achieved by constructing the magnet from a small number ($N \geq 3$) of wedge-shaped sectors that meet at a point in the center (as in **Fig. 6**, where $N = 6$). In the regions between the magnets, the orbits are almost straight, since the magnetic field is very weak there. They then bend sharply through $360°/N$ within each of the magnets. As a result, orbits cross the edge of each magnet at a nonperpendicular angle, and the ion velocity at the field edge has a component in the radial direction. But the magnetic field lines at the edge of any magnet are bowed outward, so the field there has an azimuthal component, which points in opposite directions above and below the mid-plane. Taking the vector product of the radial component of velocity times the azimuthal component of the field gives an axial force toward the mid-plane, which pushes particles that are out of the plane back toward it. This force is axially focusing whether the particles are above or below the plane, and whether they are entering or leaving the magnet. This "edge" focusing is then an additional force introduced by the sector structure; it can be used to override the defocusing that arises from increasing the average field with radius in order to maintain a constant angular frequency in Eq. (1).

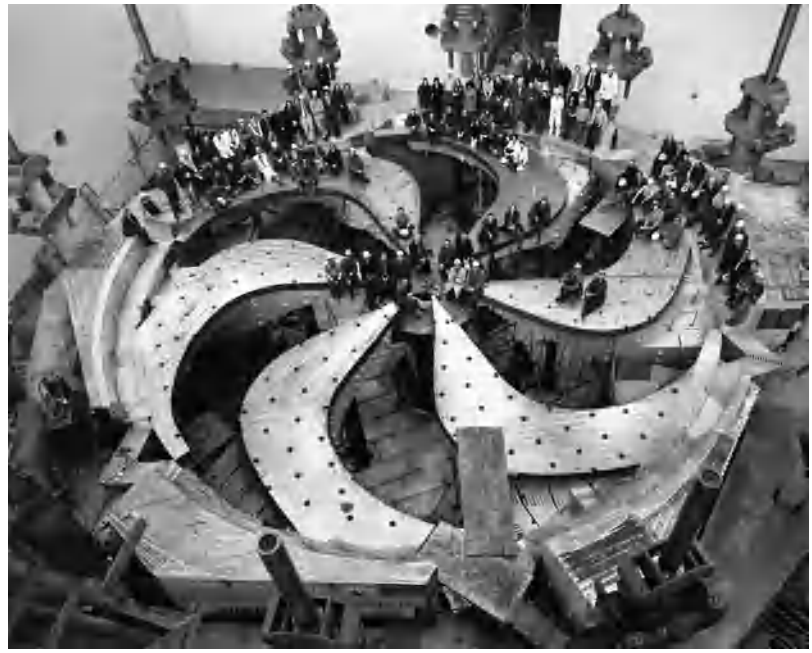

Fig. 6. Partially assembled magnet for the 500-MeV TRIUMF cyclotron, showing how the spiral iron sectors of the lower pole change from radial to spiral to provide stronger focusing as the orbit radius increases. (*TRIUMF, Vancouver, British Columbia*)

Fig. 7. Separated-sector ring cyclotron at PSI in Switzerland, which accelerates protons to 590 MeV. The four rf cavities are visible between the eight spiral sector magnets. (*PSI*)

Isochronous cyclotrons have been built in a broad spectrum of sizes and types. The sectors of many such cyclotrons have a spiral shape at large radii (Fig. 6) to increase the edge axial focusing for higher-energy ions; this introduces a strong-focusing effect (see page 44). In the ring or separated-sector cyclotron (**Fig. 7**), a magnet coil is wound around each sector magnet, leaving open space between sectors for rf accelerating cavities. The beam from a preaccelerator (sometimes a smaller cyclotron) is injected horizontally between the sectors where the magnetic field is very low. The largest spiral-sector and separated-sector cyclotrons (Figs. 6 and 7) produce over 200 microamperes of protons, and are known as meson factories, since large quantities of π mesons are produced when the protons hit a target. The record for proton energy and current is held by the cyclotron at PSI (Paul Scherrer Institut), near Zurich, Switzerland, which delivers 2 milliamperes at 590 MeV to a spallation neutron source (Fig. 7). *See* MESON.

Several laboratories have multistage cyclotron systems. Some of these systems are designed for heavy-ion beams, and electrons are stripped off between cyclotrons to produce higher charge states and higher energies. Other laboratories inject ions into a cyclotron from an electrostatic or linear accelerator, again with the help of electron stripping. Cyclotrons are also used to inject ions into synchrotron and storage rings, which can further accelerate, store, and cool the ions for precision experiments.

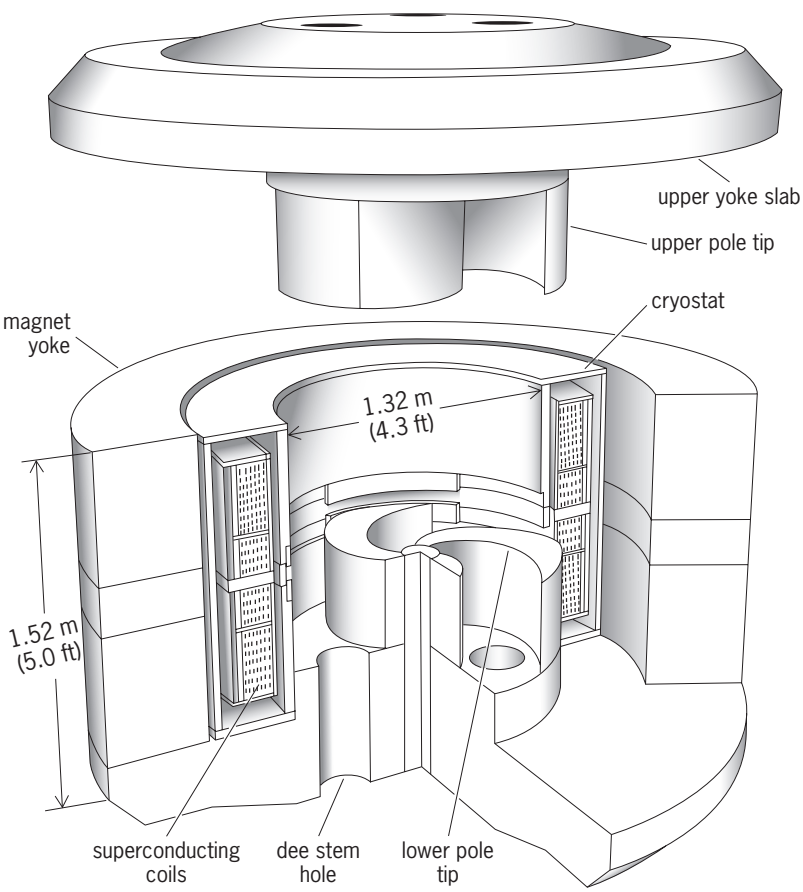

Fig. 8. Conceptual drawing of superconducting cyclotron at Michigan State University. The upper pole cap shown in the raised position is used for maintenance. (*After H. G. Blosser, The Michigan State University superconducting cyclotron program, IEEE Trans. Nucl. Sci., NS-26(2), pt. 1:2040–2047, 1979*)

An important development in nuclear science research is the acceleration of radioactive ion beams to enable the interactions of unstable isotopes to be studied. Cyclotrons are actively used in this area in several ways. One is in the production of radioactive species by bombardment of a target with a cyclotron beam. If the beam consists of heavy ions, a thin target is used, allowing high-energy radioactive ions to emerge for capture and transport directly to the experimental apparatus. If the beam consists of protons, a thick target is used, and low-energy radioactive ions are collected for subsequent acceleration in a cyclotron or linear accelerator. The heavy-ion process is referred to as nuclear fragmentation, the proton process as isotope production on line (ISOL).

Superconducting cyclotron. A powerful extension of cyclotron capabilities has been provided by the use of superconducting magnets—which, in contrast to previous advances, is a technological rather than a conceptual change. In practice, the only superconducting element in the magnet is usually the main coil, which is typically housed in an annular cryostat. Conventional room-temperature components, including pole tips, the rf accelerating system, the vacuum system, and the ion source, are inserted in the warm bore of the cryostat from top and bottom. The superconducting coil allows the strength of the magnetic field to be greatly increased up to a level of approximately 5 teslas (50,000 gauss), or three times higher than was previously typical of cyclotrons. This increase in field reduces the linear size of the cyclotron to approximately one-third, areas to approximately one-ninth, and so on, compared to a normal cyclotron of the same energy. The result is a large reduction in the cost of many cyclotron components. The K1200 superconducting cyclotron, at Michigan State University, is only 9.5 ft (2.9 m) in overall diameter (**Fig. 8**), but can accelerate heavy ions to energies as high as 200 MeV per nucleon. The first cyclotron with separate superconducting sector magnets is now (late 2006) being commissioned at RIKEN, near Tokyo, and has reached 345 en route to 400 MeV per nucleon. Superconducting cyclotrons have also become popular compact sources of 250–300-MeV protons for cancer therapy.

Separated-orbit cyclotrons (SOCs), in which the bending and focusing fields of the sector magnets are specially tailored for each orbit, were first conceived in the early 1960s for the acceleration of high-current (milliampere) proton beams to gigaelectronvolt energies. Their complexity and cost, using normal magnets, deterred potential builders. However, a prototype superconducting SOC has been built in Munich, Germany, and has accelerated a S^{14+} ion beam through six turns to 70 MeV. This is the ultimate in superconducting cyclotron design, since the rf accelerating cavities as well as the magnet are superconducting, and the vacuum is maintained by cryogenic pumping.

FFAG accelerators.. Another early concept that continues to attract interest is that of the fixed-field alternating gradient (FFAG) accelerator—essentially a sector-focused ring synchrocyclotron. FFAGs may be designed with either radial or spiral magnet sectors; in the former case short reverse-bending magnets are placed between the positive bending sectors to enhance the "edge" focusing strength. These ideas were developed by the Mid-western Universities Research Association (MURA) in the late 1950s, and electron models were successfully built and tested, but a full-scale version for protons, which require more powerful magnets and a much wider range of frequency modulation, was too much of a technical challenge at that time. New technology, however, in the form of metallic-alloy rf tuners, has created a resurgence in interest, as it has allowed development of rf cavities producing very high effective accelerating fields (~100 kV/m) at low (megahertz) frequencies and low Q values (~1–15). The low Q allows simple broadband operation, while the high fields permit higher repetition rates (kilohertz) to be used than in synchrotrons, so that the charge per pulse can be much lower, and beam and rf stability more easily maintained. FFAGs are therefore attractive as high-intensity proton machines in the 0.1–10 GeV range. A prototype proton FFAG was successfully commissioned in Tsukuba, Japan, in 2000, and a 150-MeV machine in 2003. A three-stage FFAG near Kyoto (**Fig. 9**) will study accelerator-driven systems (ADS)—that is, subcritical, and therefore safer, nuclear reactors. An FFAG storage ring for low-energy muons is also under construction for the Japan Particle Accelerator Research Complex (J-PARC), in Tokai. Even more ambitious FFAGs for acceleration of protons and muons are being actively considered for neutrino factories and muon colliders. As muons are produced with a wide angular spread, the very large transverse and momentum acceptances of FFAGs are highly advantageous. *See* Q (ELECTRICITY).

All the FFAGs described above have followed the traditional scaling principle, which requires all turns in the spiral orbit to have the same shape and focusing properties, irrespective of particle energy, in order to avoid exciting resonant oscillations that might lead to beam loss. For some applications, however, resonant oscillations pose less of a danger, and several novel and less restrictive nonscaling designs are being explored. One of these uses constant-gradient magnets with the field strength decreasing outward to achieve very high momentum compaction: the orbit spiral becomes extremely tight, allowing the vacuum chamber and magnets to be kept relatively small and inexpensive. One suitable application is the acceleration of muons (lifetime 2.2 μs), as it must take place very quickly. In a 10–20-GeV muon accelerator with a radius of 223 ft (68 m), the width of the orbit spiral would be less than 4 in. (10 cm). A 10–20-MeV electron demonstration model is under construction at Daresbury in the United Kingdom. The same principle is being applied to FFAGs for accelerating high-intensity proton beams to a few gigaelectronvolts and proton and carbon-ion beams for cancer therapy.

Ion sources. An important property of an accelerated ion is its charge state q. The maximum energy of an

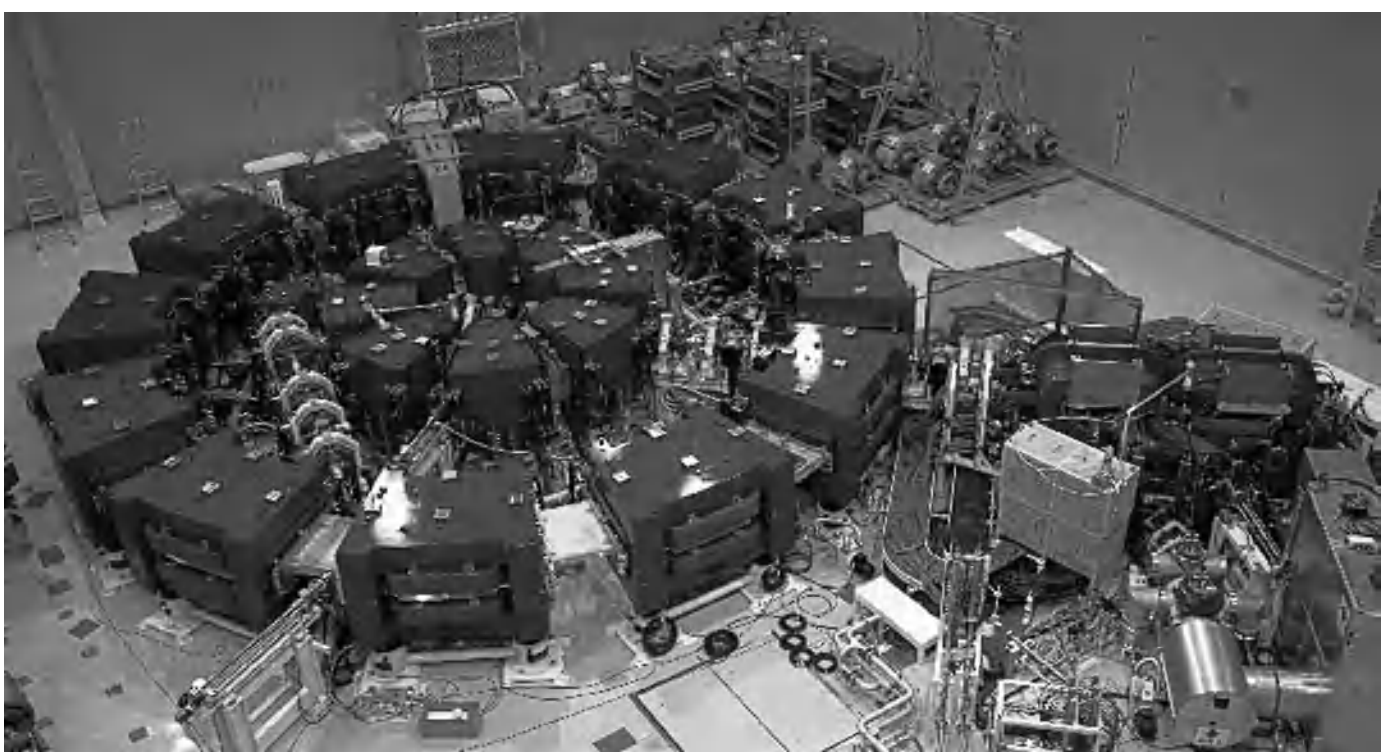

Fig. 9. FFAG complex at Kyoto University Research Reactor Institute: 20- and 150-MeV radial-sector rings (left) and 2.5-MeV spiral-sector injector (right). (*Y. Mori, Kyoto University*)

ion in a cyclotron is proportional to the square of its charge (and inversely proportional to its mass). Thus, for heavier ions, there is a strong incentive to produce ions in high charge states.

Originally the standard ion source was based on the principle of the Penning ion gauge (PIG), in which electrons oscillating in a magnetic field ionize the source gas. The highest-charge-state ions produced by a PIG source in useful intensities were, for example, $q = 5$ for nitrogen (N^{5+}). The electron cyclotron resonance (ECR) source uses microwave-heated electrons in a magnetically confined plasma to produce much higher charge states. For example, the ECR source will produce fully stripped N^{7+} beams. Since maximum ion energy is proportional to the square of ion charge, this doubles the nitrogen energy from that available from the same cyclotron with a PIG source. For heavier ions, fully stripped argon A^{30+} and uranium up to U^{60+} have been accelerated from an ECR source. This source, which is too large to fit in the center of a cyclotron, is placed outside and connected to an injection system which brings the beam axially into the cyclotron center region. Many cyclotrons have upgraded their capabilities by installing ECR sources, and over 30 are in service. These sources are also employed in linacs and synchrotrons.

Polarized protons and deuterons (that is, with their spins aligned) can be produced in special ion sources. Like the ECR source, polarized ion sources are too large to fit into the center of a cyclotron, so ions are injected from an external source. Polarized ions are used for nuclear science experiments which probe the spin-dependent characteristics of nuclear reactions or structure. Polarized radioactive ^{8}Li ions are also used for studies of the surfaces of materials involving β-NMR, a form of nuclear magnetic resonance in which the nuclear spin precession signal is detected through the beta decay of a radioactive nucleus.

Cyclotron applications. There are over 200 cyclotrons in operation. More than 50 are used principally for basic research in nuclear physics and chemistry, aimed at understanding the structure of nuclei, nuclear forces, and nuclear reactions. Another 150 are installed in laboratories and hospitals for the pursuit of practical applications, including isotope production for medical and other uses, cancer therapy using proton and secondary neutron beams, and the study and irradiation of materials. *See* NUCLEAR REACTION; NUCLEAR STRUCTURE. Michael Craddock

Electron synchrotron. An electron synchrotron is a circular accelerator optimized to accelerate electrons or positrons to high energies. The acceleration is achieved as the particles pass through electric fields created in a resonating rf cavity. A string of bending magnets along a circular path forces the particles to travel along a closed loop leading through the accelerating cavity at each revolution. In this way, each particle is accelerated many times during the accelerating cycle by the same cavity, and even a modest accelerating field will eventually lead to a high particle energy. *See* CAVITY RESONATOR.

To make this acceleration mechanism function properly, certain conditions have to be met. The oscillating field in the cavity has to be in the accelerating phase every time a bunch of particles arrives. Thus the time it takes the particles to orbit once around the ring must be an integral multiple of the rf period (the synchronicity condition). In an electron synchrotron the particles are injected at a kinetic energy of some tens of megaelectronvolts and therefore always have a velocity close to the speed of light. As a consequence the revolution frequency of the particles is constant, and it is possible to use a fixed-frequency accelerating cavity. This is the primary difference between electron synchrotrons and proton synchrotrons or cyclotrons, where the frequency of the accelerating rf fields must be adjusted during acceleration by factors of 10 or more because the much heavier protons travel, at typical injection energies, at speeds much slower than the speed of light. A linear accelerator or a microtron is usually used as a preaccelerator to inject the particles into the synchrotron at an energy of tens or hundreds of megaelectronvolts.

In the past, electron synchrotrons have been used extensively for research in high-energy physics and as sources of synchrotron radiation. However, they are now constructed and used exclusively as injectors into storage rings. Electron synchrotrons have been built in sizes from 6.5 ft (2 m) in diameter for an energy of 100 MeV, up to 650 ft (200 m) in diameter (**Table 2**) for a maximum particle energy of 12 GeV.

Acceleration and beam control. The operation of the electron synchrotron is determined by its accelerating cycle, which may range from a few milliseconds

TABLE 2. Some parameters of electron synchrotrons

Parameters	Cornell University	DESY II*	National Synchrotron Light Source (NSLS)	Stanford University
Maximum beam energy, GeV	12	9.2	0.75	3.5
Radius, ft (m)	328 (100)	153 (46.6)	14.8 (4.51)	69.6 (21.2)
Cycling rate, Hz	60	12.5	1	10
Maximum bending magnet field, teslas	0.4	1.13	1.31	1.3
Injection energy, MeV	150	200	70	120
Particles per pulse	3×10^{10}	4×10^{10}	10^{10}	2×10^{10}

*Deutsches Elektronen Synchrotron, located in Hamburg, Germany.

up to seconds. At the beginning of a cycle the preaccelerator is triggered to produce a particle beam that is injected into the synchrotron. The beam from the injector will have certain characteristics as required for optimum acceptance by the synchrotron. The synchronicity condition is fullfilled only for particles which arrive at the right time for acceleration in the synchrotron cavity. The ratio of the radio frequency to the revolution frequency is called the harmonic number and is equal to the maximum number of places (called buckets) around the ring where particles can be placed for stable acceleration. While these buckets rotate around the ring at nearly the speed of light, the injector has to deliver a beam structured and timed so that the particles arrive at the injection point at the same time as the buckets arrive there. Either one or all of the buckets can be filled with particles.

While the synchronicity condition is very important, it has to be fulfilled only approximately. The principle of phase stability provides a restoring force for all particles which do not exactly meet the synchronicity condition, tying together the characteristics of the magnetic guide field and the rf fields in such a way that the particles stay trapped close to and oscillating about the ideal equilibrium position. This is important during the accelerating process. As the magnetic guide field is increased, the principle of phase stability adjusts the particle energy to be proportional to the rising magnetic field. After executing many turns during which the particle energy is increased by a few megaelectronvolts per turn, the particles reach the maximum design energy.

At this point the particle beam is extracted and guided to an experimental station or to a storage ring. The extraction is commonly achieved by the use of a pulsed high-field kicker magnet that deflects the beam out of the synchrotron. After extraction the magnetic guide field is reduced again to the injection condition, and a new accelerating cycle can begin.

The maximum energy that can be achieved in a synchrotron is limited by either the maximum magnetic field necessary to keep the beam on its orbit or by the effect of synchrotron radiation. As the particles are made to follow curved trajectories forming closed orbits, they lose energy by way of this radiation. The energy loss per turn increases as the fourth power of the energy, and the maximum energy is reached when this energy loss is as large as the maximum energy gain per turn in the accelerating cavities.

While a stronger magnet and rf system could increase the maximum energy somewhat, there is another fundamental limit on this energy. During the first phase of an accelerating cycle the beam cross section shrinks from its value at injection (typically a few square millimeters), inversely proportional to its energy. However, the statistical emission of synchrotron radiation photons causes a rapid increase in the beam cross section as the energy is raised. The limit is reached when the beam fills the available aperture in the vacuum pipe.

The maximum intensity in a synchrotron is determined by the performance of the preinjector as well as by beam instabilities. Individual particles can be lost because of collisions with other particles or because of the electromagnetic interaction of all particles in one bunch with the surrounding vacuum chamber or with other bunches. In all cases the limitation on the beam current is more severe at low energies than at high energies, which is why

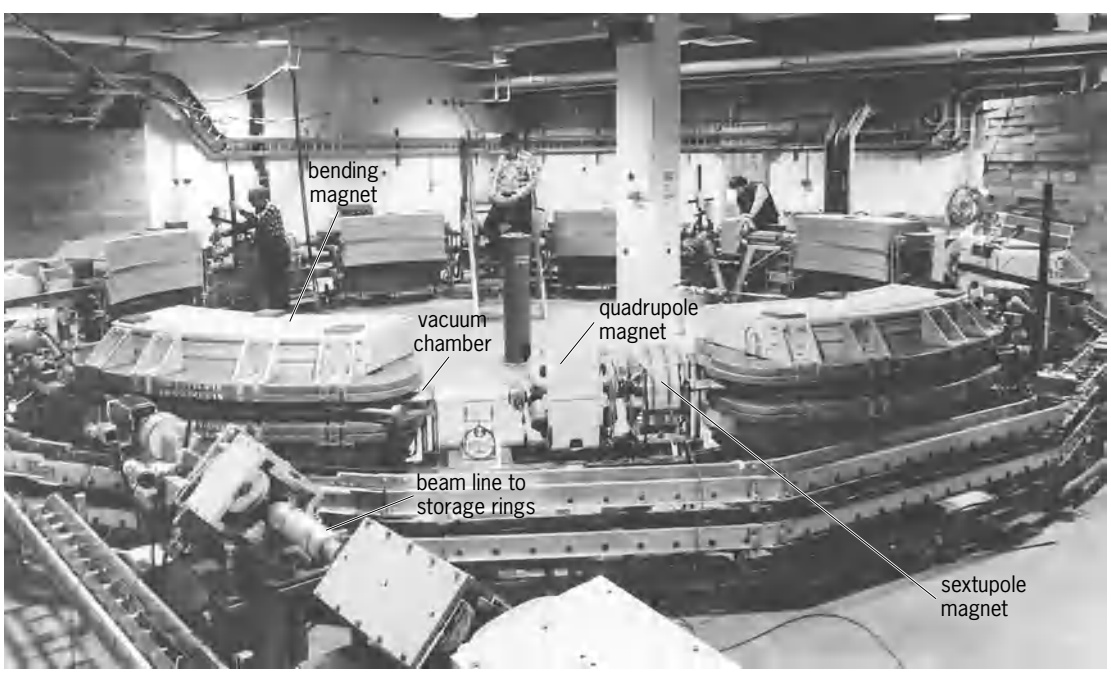

Fig. 10. The 750-MeV electron synchrotron at the national synchrotron Light Source (NSLS) at Brookhaven, National Laboratory.

synchrotrons are used as a step-up injector between the preaccelerator and a storage ring. In a storage ring it would not be possible to accumulate a high beam current at the low energy of the preinjector. In a booster synchrotron a small beam current is accelerated repeatedly to higher energies where instabilities in the storage ring are less severe and high beam currents can be accumulated.

Magnet system. The magnetic guide field in a synchrotron consists of dipole fields to keep the particles on a circular path and a focusing field generated by quadrupoles to keep the beam from diverging. The arrangement of bending and focusing magnets is called the magnet lattice. In the older synchrotrons, both bending and focusing were performed by one type of magnet, and the resulting magnet lattice was called a combined-function lattice. As magnet technology progressed and more flexibility was desired, electron synchrotrons began to employ the so-called separated-function lattice, where bending and focusing is done by different magnets with separate power supplies (**Fig. 10**). To enhance the beam current capability, the newer synchrotrons also use sextupole magnets to correct for effects caused by the energy spread in the beam. In analogy to light optics, these effects are called chromatic aberrations of the particle beam focusing.

Vacuum system. To avoid loss of particles during acceleration, the particles circulate in a vacuum pipe embedded in the lattice magnets. In older synchrotrons this vacuum pipe had to be made from an insulating material to allow the rapid cycling of synchrotrons at 50 or 60 Hz without creating problems due to eddy currents caused by the magnetic fields in the vacuum chamber. In the newer synchrotrons used exclusively as injectors for storage rings, the cycling rate is only about 10–15 Hz or less, and eddy currents are reduced to a negligible level at this lower cycling rate. This makes it possible to use thin stainless steel vacuum pipes, thus reducing cost and improving reliability *See* EDDY CURRENT.

Helmut Wiedemann

Proton synchrotron. The highest-energy particle accelerators are proton synchrotrons (**Table 3**). While many new uses for this type of particle accelerator are being realized, such as medical applications, the primary use of the proton synchrotron remains in the exploration of nuclear and high-energy particle physics. Beams of protons are accelerated to high energy and either are extracted toward fixed targets of material or are directed into collision with other beams of particles (commonly protons, antiprotons, or electrons) to study the most elementary forces of nature. In the case of fixed-target experiments, the proton beams are often used to produce secondary beams of other elementary particles, such as mesons and neutrinos, which are in turn used as probes for physics experiments.

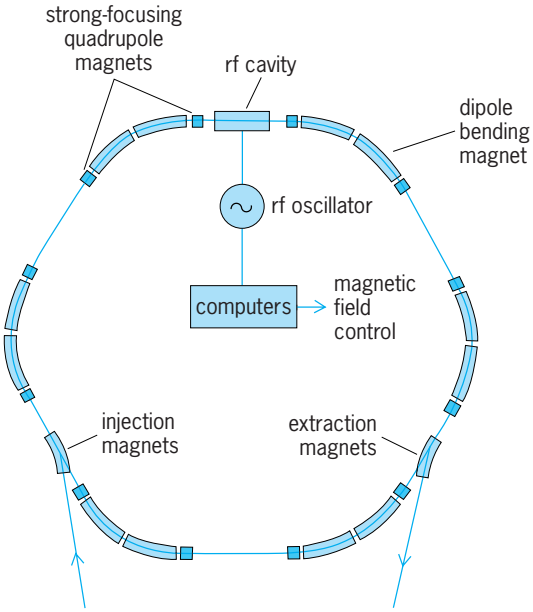

Fig. 11. Idealized proton synchrotron.

Like the cyclotron and the synchrocyclotron, the proton synchrotron makes repetitive use of accelerating fields, in this case by guiding the protons in a nearly circular path by using a time-varying magnetic field which changes in synchronism with the increasing momentum of the particles. The acceleration is performed by radio-frequency (rf) cavities, whose oscillating electric fields point back and forth along the design trajectory of the accelerator. The electric field in the cavities is arranged such that particles are accelerated as they pass through, and then as the field reverses direction, a gap is left in the particle beam. In this way the proton beam is not uniformly continuous about the circumference of the accelerator, but is bunched. The acceleration takes place over a typically short portion of the circumference (**Fig. 11**), while most of the device's circumference is made up of bending magnets which serve to bring

TABLE 3. Large proton synchrotrons currently in operation

Synchrotron	Date of commissioning	Approximate radius, ft (m)	Particle energy, GeV
Proton synchrotron (PS), CERN, Switzerland*	1959	330 (100)	28
Alternating-gradient synchrotron (AGS), Brookhaven, Upton, New York	1961	410 (125)	33
IHEP, Serpukhov, Russia	1967	770 (235)	76
Super proton synchrotron (SPS), CERN	1976	3600 (1100)	400
Tevatron, Fermilab	1983	3300 (1000)	980
Main injector Fermilab	1999	1700 (520)	150
Relativistic Heavy Ion Collider (*pp* mode)	2000	2000 (614)	100

*The CERN PS and the Main Injector at Fermilab are used as injectors for the SPS and the Tevatron, respectively.

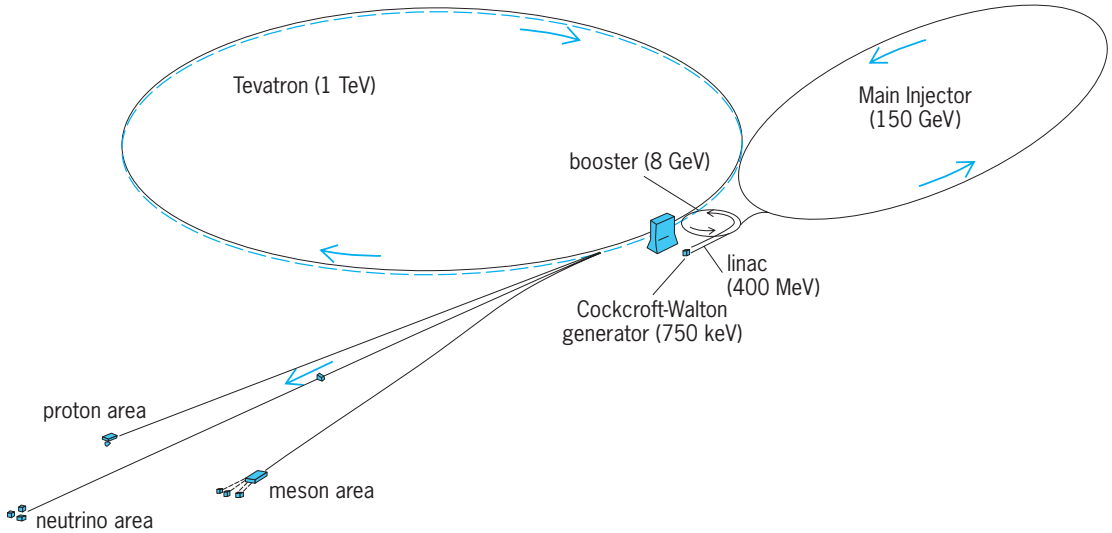

Fig. 12. Fermilab facility, showing sequence of accelerators. At present (2006), protons are routinely extracted from the Main Injector and routed to experimental areas by a "proton switchyard." The extracted beam is used to produce beams of secondary particles. Only the "meson area" beam line shown in the figure is still in use. A new beam line, for neutrino production, is also in operation.

the particles back to the accelerating cavities. As the momentum (and hence, energy) of the particles increases, the fields of the guiding magnets must also increase to maintain the same central trajectory. In addition, as the speed of the particles increases, the frequency of the rf cavities must be adjusted to maintain synchronism. Most modern proton synchrotrons use separate types of magnets to perform the bending (dipole magnets) and the necessary focusing (quadrupole magnets) of the proton beams. A certain number of straight sections are also necessary in the accelerator to perform other functions such as injection and extraction of the particles to and from the accelerator, and various beam diagnostics and control.

Accelerator cascade. As proton beams are accelerated, their transverse dimensions shrink, and hence smaller apertures may be used for the guiding magnets and accelerator hardware. Thus, to generate beams of very high energy protons, it is most economical to use a cascaded series of accelerators, each optimized to accelerate the particles over a particular range of energy. For example, in the Fermilab accelerator complex at Batavia, Illinois (**Fig. 12**), H^- ions (hydrogen atoms each with an extra electron) are generated in a Cockcroft-Walton electrostatic generator and accelerated across a potential of 750 kV. The beam is next transported to a linear accelerator which increases the total kinetic energy of each particle to 400 MeV. The ions are then stripped of their electrons, leaving a pure proton beam, upon injection into a booster synchrotron. This synchrotron accelerates the particles to an energy of 8 GeV and transfers groups or bunches of protons into the Main Injector. At this point, the protons are moving at over 99% the speed of light. The Main Injector accelerates each particle to an energy of 150 GeV, and then the particles are transported, in a single revolution, into the Tevatron accelerator (**Fig. 13**). Two such injections are required to fill the Tevatron. Finally, the Tevatron, a synchrotron made up of superconducting magnets, accelerates the particles to an energy of nearly 1 TeV. In its previous history, protons were then extracted from the Tevatron and transported to fixed-target experiments. The entire acceleration process took roughly 20 s to complete. Today (2006), the Tevatron is used as a proton-antiproton colliding beams accelerator. Fixed-target experiments, however, are still performed using 120-GeV protons directly emanating from the Main Injector.

Particle stability. Implied in this description of a synchrotron is an ideal particle, which at each moment

Fig. 13. Tunnel of the main accelerator at Fermilab. The superconducting magnets of the Tevatron sit near the floor. The magnets above it are remnants of the original Fermilab 400-GeV synchrotron, which has been decommissioned. (*Fermi National Accelerator Laboratory*)

has exactly the right energy and the right time of passage through the accelerating structure so that it receives exactly the right increment of energy to stay in accord with the plan of the accelerator system. But in fact, most (if not all) particles do not meet these very special requirements. At relatively low energies, particles which have a lower momentum than the ideal tend to lag behind because of their lower velocities, and higher-momentum particles move ahead of the ideal particle. At higher energies, as the velocities of the particles approach the velocity of light, increments in energy do not increase their speed significantly. But their momentum continues to increase, and hence the bending magnets do not bend the particles' trajectories as much. The particles with slightly excessive momentum then move in longer orbits around the circumference and hence actually take longer to circulate.

Since the fields in the rf cavities are changing with time, they can be tuned to give different increments of energy to the particles that arrive at the cavities too soon or too late. At lower energies, the particles arriving early receive less than the ideal increment in energy, and those arriving late receive more. A particular particle which has too little energy will keep receiving larger-then-ideal increments until it eventually receives too much. It will then begin arriving at the rf cavities late and hence begin receiving less energy than required. In this way, particles will oscillate about the ideal energy. These energy oscillations are called synchrotron oscillations. If the ideal energy of the particles is high enough, the slightly higher energy particles will arrive at the cavities later than desired, and the phase of the rf oscillating fields must be adjusted so as to give these late particles less energy rather than more. In many proton synchrotrons, there is a transition energy at which the two effects determining the revolution period—the particle's speed and the particle's energy—have offsetting effects. At this energy, the protons all circulate with the same orbital period, and phase stability is momentarily lost. If this energy is crossed quickly enough, and the phase of the rf system switched quickly enough, then no harm is done to the quality of the proton beam.

In addition to having a small spread in energy, the individual protons in the beam enter the accelerator with a variety of trajectories, close to the ideal but slightly off in either displacement or direction. The synchrotron thus contains magnetic elements which can focus the particles, redirecting them toward the ideal trajectory. Early, weak focusing proton synchrotrons used so-called gradient magnets to generate focusing fields that weakened with increasing radius of the particle trajectory. These magnets had the disadvantage that as the design energy, and hence the circumference of the orbit, is increased, aperture also must be increased to allow for particle oscillations. This circumstance led to the invention, in 1952, of alternating-gradient focusing or strong focusing, which allows the size of the beam tube and magnets to be decoupled from the circumference of the accelerator. Most modern proton synchrotrons use quadrupole magnets to perform this focusing. A quadrupole magnet focuses the beam in one plane (horizontal, say) and defocuses it in the other plane (vertical). But a pair of quadrupoles, of opposite polarities, can be arranged to have a net focusing effect in both planes simultaneously.

An individual particle oscillates transverse to the ideal trajectory with a frequency determined by the focusing magnets. This motion is called betatron motion, and the number of transverse oscillations that a particle undergoes during one revolution about the accelerator is termed the "tune" of the accelerator. A particle may undergo many betatron oscillations during one revolution, while it may take hundreds of revolutions to perform one synchrotron oscillation.

Extraction. In a fixed-target synchrotron facility, once the beam is accelerated to its final energy the particles are slowly extracted from the accelerator and sent toward external experimental areas. There, beams of secondary particles, such as mesons and neutrinos, can be created if required, and the final particle beams are directed toward the experimental apparatus to study particle interactions with matter. Frequently, the experimental equipment can handle only a certain rate of events, and so a smooth, continuous delivery of particles to the apparatus is preferred to a single shot of an intense particle beam. (Of course, many other experiments do prefer the latter.) The method of slow resonant extraction is generally used to meet this request. By putting the particles in the synchrotron into resonance with error fields in the accelerator (often introduced purposefully), in a well-controlled fashion, the transverse amplitudes of the particles are made to grow until some of the particles can find themselves on the opposite side of a septum made up of very thin, electrically charged wires. The electric field on one side of the wires deflects the particles into a separate beam pipe and beam-transport channel which guides them toward the experimental areas. The rate of growth of the particle amplitudes is controlled by a feedback system, which ensures that the particles do not all leave at once but are gradually spilled out of the accelerator.

The extraction process can take tens of seconds if desired. Ideally, it may be desirable for experiments to receive the beam continuously from the accelerator, but the actual duty factor—the slow spill time divided by the total cycle time—will be less than one because of the injection, acceleration, and reset time. A fast resonant extraction can also be performed, whereby the entire beam is brought into resonance very quickly, on the order of milliseconds. The extraction rate can also be modulated, changing as the external beam is switched quickly among users with differing requirements.

In some instances, a very fast extraction process is required. Notably in superconducting accelerators, the beam may have to be taken out of the synchrotron in order to avoid damage to the superconducting magnets. In this case, a manually or electronically generated signal will cause kicker magnets with

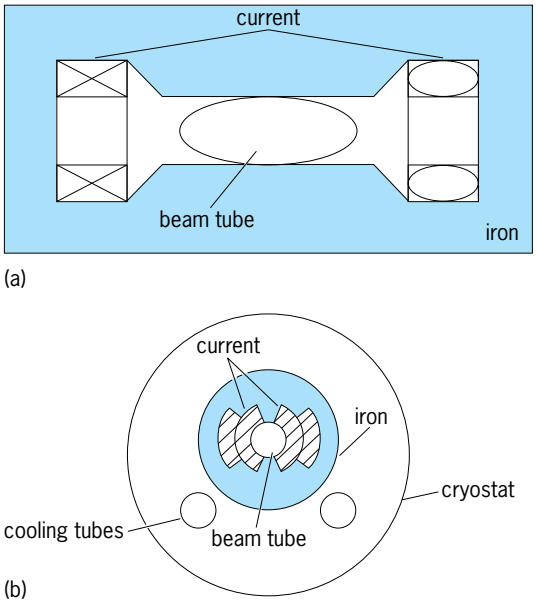

Fig. 14. Cross sections of dipole magnets used in particle accelerators. (*a*) Conventional iron magnet. (*b*) Superconducting magnent.

very fast rise times, on the order of microseconds, to energize and deflect the beam through an extraction channel toward a special beam absorber. The kicker magnets are fast enough to extract the entire beam in a single revolution.

Superconducting synchrotrons. In any synchrotron, the maximum energy of the protons is dictated by attainable field strength of the magnets used to bend the particle trajectory along its course. In conventional (warm) accelerators, the magnetic field is produced by using electromagnets whose field shape and uniformity are governed by the shape of the gap in the iron, through which the particles circulate (**Fig. 14***a*). But in iron, magnetic fields saturate at strengths of approximately 2 T. A significant advance in accelerator technology occurred with the invention of superconducting accelerator magnets. These magnets use superconducting materials for the cable used to carry the current which generates the magnetic field. When cooled to temperatures near absolute zero, these materials lose all electrical resistivity, and very large current densities can be obtained with no heat lost in the cable due to resistance. The fields of these magnets are determined primarily by the placement of the strands of superconducting cable; the iron serves primarily to provide structural support (Fig. 14*b*). Accelerator magnets of this type have been produced in the laboratory which reach fields of roughly 10 T. Niobium-titanium (Nb-Ti) alloys have been the primary superconductor employed. In the presence of the strong magnetic fields involved, these alloys must be cooled to about 4.5 K (−451°F) by using liquid helium as coolant.

In addition to providing higher fields, accelerators with superconducting magnets allow beams to be stored for long periods of time (as in a colliding-beams accelerator, or during slow spill in a fixed-target accelerator) at high field with relatively low electrical power consumption. The most power consumption comes from running the refrigeration system to keep the magnets at their superconducting temperature. (The bus work carrying electricity to the magnets also consumes some energy.) Energy deposition in the magnets, due to radiation, heat leaks, or perhaps most importantly the beam itself, must be carefully controlled in these systems.

Michael J. Syphers

Heavy-ion synchrotron. There are close similarities between heavy-ion synchrotrons and colliders, and their respective proton counterparts. Special considerations of accelerator design for heavy ions arise mostly from their tendency to change their charge states during the course of acceleration and storage, and their high electric charge. In discussing the heavy-ion accelerator, it is customary to express its energy in terms of energy per each nucleon in the ion (GeV/u) instead of the total energy.

Many proton synchrotrons have been augmented by a heavy-ion injector to accelerate heavy ions. For a given peak magnetic (bending) field of a synchrotron, the maximum energy attainable is proportional to the ratio of the charge state to the atomic mass number of the ion species being accelerated. For protons this ratio is unique and equals 1, whereas for heavy ions it varies according to ion species accelerated and the degree of stripping (how many electrons are removed from an atom). The stripping can be accomplished by passing ion beams through thin metal or carbon foils. In this process, a significant attenuation of beam intensity occurs by the ion's interactions, mostly scattering with foil. Light ions can readily be stripped to full charge states. Heavier ions, however, require sequences of acceleration and stripping stages to reach fully stripped states. Fully stripped ions as heavy as gold and lead have been accelerated to energies of 11.4 GeV/u and 160 GeV/u, respectively.

Most of the major heavy-ion synchrotrons (**Table 4**) are utilized for nuclear physics studies.

TABLE 4. Major heavy-ion synchrotrons

Facility and location	Beam energy, GeV/u	Range ion types
Saturne II	1.15	Lithium (Li)
(Saclay, France)†	0.69	Krypton (Kr)
GSI	2.0	Deuterium (D)
(Darmstadt, Germany)	1.0	Uranium (U)
Bevalac	2.1	Carbon (C)
(Berkeley, California)**	0.96	Uranium (U)
Synchrophasotron-Nuclotron	6	Deuterium (D)
(JINR, Dubna, Russia)	4.3	Uranium (U)*
AGS, Brookhaven,	14.5	Oxygen (O)
(Upton, New York)	11.4	Gold (Au)
SPS, CERN	200	Sulfur (S)
(Geneva, Switzerland)	160	Lead (Pb)
HIMAC	0.4	Silicon (Si)
(Chiba, Japan)	0.8	Xenon (Xe)

*Beam intensity less than 10^5 ions per second.
†Closed down in 1997.
**Closed down in 1993.

The high-energy heavy-ion accelerator GSI (Darmstadt, Germany), whose primary application is to nuclear physics, has been used for cancer therapy, and an 800-MeV/u heavy-ion synchrotron, HIMAC, at National Institute of Radiation Science (Chiba, Japan) was constructed with medical applications such as cancer therapy as its primary purpose.

Heavy-ion accelerators require a significantly higher vacuum than their proton counterparts. This is to prevent beam loss from the change of the charge state by electron pick-up or stripping processes through collisions with residual vacuum-chamber gas. The rf power must be boosted for heavy ions to overcome the increased inertia caused by the lower charge-to-mass ratio, and to keep the beam energy synchronized with the rising magnetic fields during acceleration.

Satoshi Ozaki

Linear Accelerator

A linear accelerator accelerates particles in a straight line by means of electric fields developed across a series of accelerating gaps in sequence. Each gap is transversed once and the electric fields must be produced along the entire orbit. Extremely high field strengths and hence high rf power levels are required for room-temperature structures and high beam currents. Superconducting structures requiring low power have been developed for certain applications.

Principles of operation. The synchronism between particles and accelerating field may be achieved by either of two methods: traveling-wave acceleration, whereby a wave with an accelerating field component whose phase velocity is equal to the particle velocity is utilized; or standing-wave acceleration, whereby a standing-wave pattern is produced by the superposition of forward and backward waves in the structure. In the latter case, either the forward-wave phase velocity is made equal to the particle velocity or, more frequently, a set of drift tubes is introduced into the structure to shield the particles from the fields when they are passing through regions where the fields would otherwise be decelerating.

The energy to which a particle can be accelerated is proportional to the particle's charge, the square root of both the length of the accelerating structure and the total power of the rf power sources feeding it, and the $-{}^1/_4$ power of the wavelength. Thus, length and power are equally instrumental in attaining high energies and short wavelengths are also advantageous in terms of power economy. Countering this advantage of short wavelengths are the following disadvantages: (1) higher power levels are generally available from power tubes operating at longer wavelengths and the accelerator guide is capable of high power handling; (2) the absolute tolerances of the mechanical structure are larger at larger wavelengths; (3) longer wavelengths allow for higher beam currents. In practice, electron accelerators use wavelengths of 3–20 cm; proton accelerators, 30–200 cm; and heavy-ion accelerators, 300–1000 cm. In general, the high power levels required for room temperature operation result in pulsed rather than continuously operating machine design, but there is no fundamental reason for pulsed operation and some continuously operating machines are under design.

Successful acceleration of particles requires that acceleration be phase-stable and that the beam of particles remain focused along the orbit. There will be a net defocusing action in a field region where the accelerating field increases in time as the particle crosses the region (**Fig. 15***a*). This is because the defocusing transverse momentum imparted as the particle leaves the region is greater than the focusing momentum imparted as the particle enters it. However, the action is phase-stable, since a particle arriving late will experience a larger acceleration than a faster-traveling particle which arrives early. If the field decreased in time during particle passage,

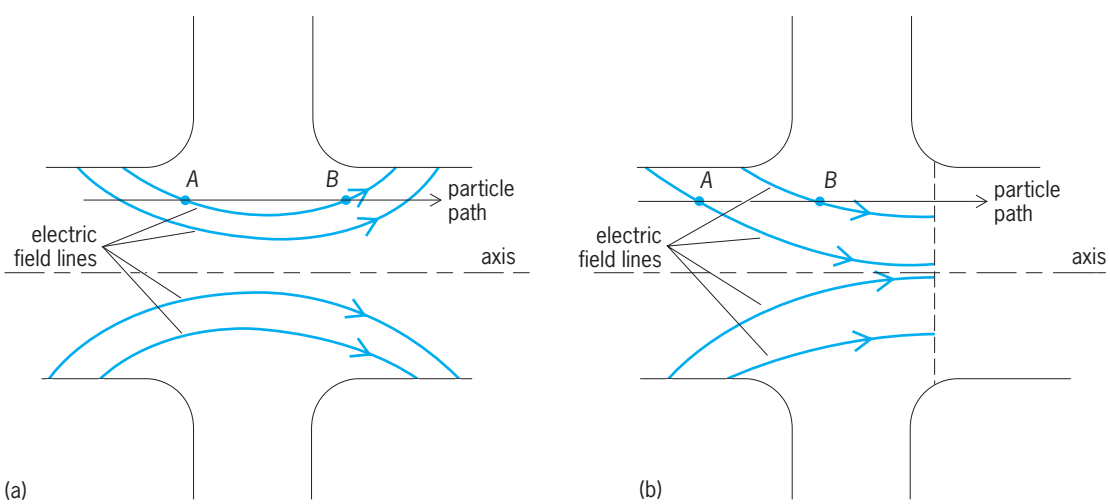

Fig. 15. Phase focusing in a linear accelerator. (*a*) Gap without grids. Focusing transverse momentum imparted at *A* is less than defocusing momentum imparted at *B*. (*b*) Gap with grids. Focusing momentum is imparted at both *A* and *B*.

there would be transverse focusing but longitudinal phase instability. Thus focusing and phase stability are incompatible, and this appeared to be a serious obstacle to the design of long linear accelerators. It has been circumvented in various ways as follows: (1) Charge can be included in the beam to terminate field lines inside the beam and cause a convergent action even at phase-stable transit phases. This charge has typically been induced by grids placed in the field (Fig. 15*b*). In this case, there is focusing momentum in all parts of the region, and in an increasing field the acceleration is also phase-stable. Unfortunately, the grids also intercept some of the beam and therefore prohibit high-current operation. (2) As the particle velocity approaches the speed of light, the incompatibility becomes irrelevant because, for relativistic velocities, the action of the radial electric time-varying field is almost canceled by the accompanying time-varying magnetic field; also, because the velocity is almost invariable, the particles are in neutral equilibrium longitudinally. (3) External magnetic fields (solenoidal or strong-focusing) or electrostatic lenses have been used. (4) By alternating the phase difference, or by suitable field shaping, a region of limited phase stability is attainable.

Electron accelerators. Virtually all electron linear accelerators are of the traveling-wave type. The wave is produced in a loaded waveguide excited in a transverse magnetic (TM) mode which has a longitudinal electric field component. The loading is typically achieved by means of disks which are sized to produce the correct phase velocity to match the particles. The field pattern may be visualized as translating along the axis with a velocity equal to that of the electron beam. Typically solenoid focusing is used over the first few feet of the accelerator, after which the electrons approach the velocity of light whereby all radial forces may be neglected. The momentum component transverse to the axis remains constant, and since the longitudinal component increases continuously, the angle of divergence of the beam decreases continuously. For uniform energy gain per unit length, this corresponds to a beam angle varying as the inverse of distance along the machine and to a beam radius increasing logarithmically. A set of weak magnetic lenses can be used to produce a decrease in beam diameter. *See* MAGNETIC LENS; PHASE VELOCITY; WAVEGUIDE.

Accelerator guides must be manufactured to very close tolerances in order to control the phase velocity to the required accuracy, since there is no phase stability once the energy has grown to several times the rest energy of the electron. However, in practice the accelerator is divided into many separate accelerator guides, each with its own power source. These separate sections may be individually tuned and phased during operation, thus somewhat relieving the tolerance requirements.

Electrons are injected into the accelerator with an energy of typically 80–120 keV, corresponding to a velocity of approximately half the speed of light, and the first accelerator section is often a special bunching device to concentrate the electrons near the crest of the traveling wave in the succeeding sections.

Performance. The total energy capability from a linear accelerator has no fundamental limit since electrons accelerated in a straight line do not lose any appreciable energy because of radiation. The performance of a given accelerator is closely tied to the available power source and the voltage breakdown capability of the accelerating guide. Most machines operate at microwave frequencies and utilize klystron amplifiers driven from a common master oscillator as power sources. *See* KLYSTRON.

The acceleration of electrons at room temperature requires peak radio-frequency power levels of tens of megawatts per source. Costs of power and available power sources limit the beam duty cycle to a few tenths of a percent. Two 400–500-MeV linacs of higher duty factor have been constructed with long pulses and fast repetition rates achieved at the expense of low acceleration rates. A superconducting electron linac which can operate continuously with dissipation of about 1 W/m has been constructed, utilizing standing-wave structures of niobium metal and operated at 1 to 4 K (−458 to −452°F). However, metallurgical and electron loading problems limited the accelerating gradients to about 3 MeV/m. At this level these accelerating guides are useful with recirculating beams as in the microtron (discussed below).

The largest electron linac is 10,000 ft (3050 m) long (**Fig. 16**). Besides producing electrons directly for basic research, it injects two different synchrotron storage rings with electrons and positrons. The Stanford Linear Accelerator Center (SLAC) has increased the output energy of this 20-GeV machine to 50 GeV.

The advent of free-electron lasers as radiation sources and the development of very high power, millimeter-wave radio-frequency sources, brought renewed interest in electron linear accelerators. A key factor in the design of these machines is the brightness of the electron gun, and this has led to the development of rf electron guns. By utilizing a photocathode excited by a pulsed laser, very short electron bunches (∼5 ps) may be produced at output energies of 2–5 MeV. Because of the high energy and short bunch length, very bright beams may be produced with peak currents in excess of 100 A. *See* LASER.

Laser acceleration. Since about 1990 there has been a significant amount of theoretical, and some practical study of the acceleration of electrons at very high acceleration rates of the order of 1 GeV/m with the use of high-power lasers. A number of the experimental studies have produced acceleration. These include inverse Cerenkov and inverse free-electron-laser acceleration.

Electromagnetic radiation is emitted whenever a charged particle passes through any medium in which the phase velocity of light is less than the

Fig. 16. Stanford Linear Accelerator Center. (*a*) Aerial view showing the 2-mi (3.2-km) accelerator and associated buildings and storage rings. Broken lines show path of Stanford Linear Collider (SLC), which operated from 1989 to 1998. (*b*) The accelerator installed in its underground tunnel.

particle velocity. This is known as Cerenkov radiation, and is emitted in a conical surface of half angle ϕ, whose value is given by Eq. (2), where n is the

$$\cos\phi = \frac{1}{\beta n} \tag{2}$$

index of refraction in the medium and β is the particle velocity divided by the velocity of light. Conversely a particle can remain synchronously in phase with an electromagnetic wave in a medium if the particle and wave vector cross at this Cerenkov angle. In this situation the particle may be accelerated, and this is termed inverse Cerenkov acceleration. A practical experiment carried out at the Accelerator Test Facility at Brookhaven National Laboratory in Upton, New York, utilized an axicon lens to focus a radially polarized laser beam onto the same axis as a relativistic electron beam passing through a gas cell. When the electron and light beams were synchronized, acceleration of the electron beam was observed.

In another experiment at the Accelerator Test Facility, inverse free-electron-laser acceleration was achieved. In this scheme the electron beam is made to wiggle back and forth though a light beam from a powerful carbon dioxide (CO_2) laser by applying a series of alternating dipole fields. With proper synchronization of the laser field and the electron beam, accelerating gradients of several hundred megaelectronvolts per meter should be attainable.

An experiment using the inverse free-electron-laser accelerator and the inverse Cerenkov accelerator in tandem has been proposed. Laser wake-field acceleration has also been achieved at Argonne National Laboratory and at JAERI in Tokyo.

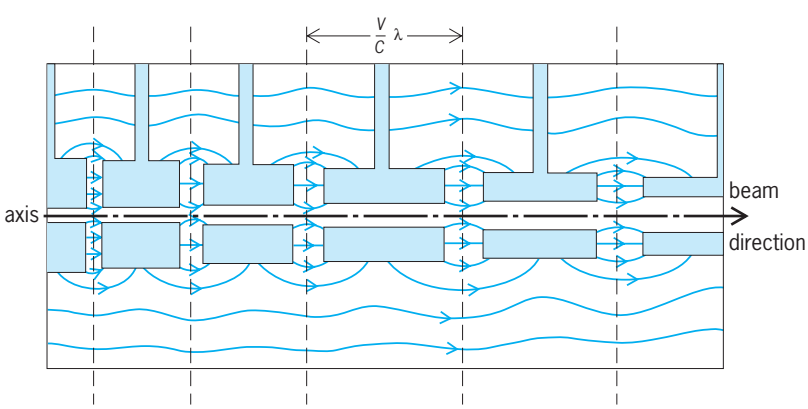

Fig. 17. Electric field configuration in a standing-wave proton accelerator employing a cavity utilizing an Alvarez structure operating in the TM_{010} mode. Drift tubes supported on stems shield the particles when the phase is decelerating. Particles cross gaps at approximately the same phase.

Proton accelerators. Proton linear accelerators are of the standing-wave type since protons do not reach relativistic velocities at typical injection voltages which are in the range 500–2000 keV. The most usual structure is that due to L. W. Alvarez (**Fig. 17**), whereby the particles pass inside copper drift tubes where no electric field is present, and are accelerated in the region between the drift tubes. The distance between the center of each drift tube, or cell length, is equal to $\upsilon\lambda/c$, where λ is the free-space exciting wavelength, υ is the particle velocity at the cell, and c is the speed of light The particles cross each gap at the same phase, usually 20–30° before the time maximum of the electric accelerating field. Focusing

TABLE 5. Heavy-ion linear accelerators

Injector linacs

Name	Location	Resonator type*	Output energy†	Highest-mass ion accelerated	Operating frequency	Injector type*	Injector energy
Super HILAC	Berkeley, California	Alvarez	8.5 MeV/amu	^{40}Ar	200 MHz	RFQ	200 keV/amu
LITL	Tokyo University, Japan	Vane-type RFQ	800 keV/amu	^{84}Kr	100 MHz	—	—
Hyperion II	Saclay, France	Vane-type RFQ	187.5 keV/amu	^{40}Ar	200 MHz	—	—
RILAC	Saitama, Japan	Quarter-wave resonator	0.5–4 MeV/amu	^{197}Au	17–45 MHz	—	—

Posttandem machines

Location	Resonator type	Resonator material	Operating frequency	Velocity range	Total voltage, MV	Number of resonators
Argonne, Illinois	Split-ring	Niobium	145.5 MHz	0.06–0.16 *c*	> 40	42
Stony Brook, New York	Split-ring	Lead	151.7 MHz	0.06–0.21 *c*	20	40
Seattle, Washington	Quarter-wave	Lead	150 MHz	0.06–0.21 *c*	26	32
Bucharest, Romania	Spiral	Copper	—	0.06–0.10 *c*	6	20
Heidelberg, Germany	Spiral and split-ring	Copper	108.5 MHz	0.06–0.10 *c*	12.5	40
Karlsruhe, Germany	Helix	Niobium	108 MHz	0.04–0.08 *c*	~3.5	3
Saclay, France	Helix	Niobium	135 MHz	0.04–0.08 *c*	~25	50

*RFQ = radio-frequency quadrupole.
†amu = atomic mass unit.

is achieved by means of magnetic quadrupoles situated inside the drift tubes. A series of cells is a resonant cavity if each cell resonates at the same frequency. Resonators are sized to match available power sources at the typical operating frequency (around 200 MHz). Special resonant posts or stems are used to improve the energy transfer along the resonator and ease the mechanical and beam loading tolerances. Resonators are placed in tandem to achieve energies up to 200 MeV with the Alvarez structure. Above 200 MeV, the efficiency of the Alvarez structure falls off, and special coupled cavity structures are necessary. The largest proton linac in operation is the 800-MeV accelerator at the Los Alamos Meson Physics Facility (LAMPF).

In some accelerators a structure called the radio-frequency quadrupole (RFQ) is used as an injector instead of the dc injectors previously favored. This device utilizes a special vane-type accelerating structure which also provides quadrupole focusing by electric fields near the axis. Injection energies of 20–50 keV and output energies of 1–2 MeV are typical for such structures.

Heavy-ion accelerators. Modern heavy-ion accelerators are divided into two classes: high-current pulsed machines used mainly for injection into circular accelerators, and low-current continuous-wave machines, which are built as postaccelerators for tandem accelerators (**Table 5**).

In general, the pulsed machines utilize an electron cyclotron resonance ion source or electron-beam ion source, operating at output voltages of 10–100 kV, followed by either a rf quadrupole or a Wideröe structure to increase the output velocity to about 5% of the speed of light.

The posttandem machines employ independently phased resonant accelerating cavities with only a few accelerating gaps and are either copper room-temperature devices or cryogenic cavities made from lead or titanium operating at liquid-helium temperatures. Structures used are either helical, split-ring, or quarter-wave resonant.

Kenneth Batchelor

Microtron

The microtron is an electron accelerator in which the beam energy is raised in steps by recirculating short electron bunches through the same rf accelerating sections. A magnetic guide field provides stable orbits which the electrons follow as they are accelerated. Compared with linear accelerators, only moderate rf power is required because the same accelerating structure is used repeatedly.

Principle of operation. The basic principle of all microtrons containing one or more accelerating sections is embodied in the coherence condition. This condition requires that the geometry of all orbits be adjusted to ensure that the revolution time back to each accelerating section equals an integral multiple of the rf period.

For the circular microtron (**Fig. 18***a*), where the electrons start with zero kinetic energy in front of a short microwave resonator, this condition is met by Eq. (3). Here ΔW is the energy gain per turn;

$$\Delta W = \frac{\nu}{\nu_1 - \nu} \cdot E_0 \qquad E_0 = 0.511 \text{ MeV} \tag{3}$$

ν_1, is the revolution time for the first turn in units of the rf period; and ν is the increase in the revolution time between successive revolutions, also in units of the rf period. The optimum values are $\nu_1 = 2$ and $\nu = 1$ leading to $\Delta W = 0.511$ MeV. Due to thermal heating by rf losses in the cavity walls, the necessary field strength can be achieved only in pulsed operation.

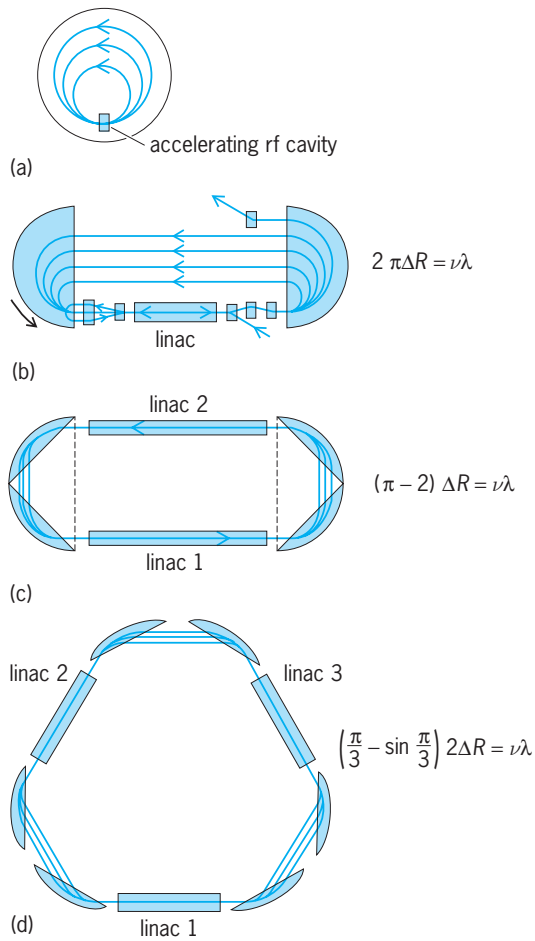

Fig. 18. Various designs for microtrons. (*a*) Circular microtron. (*b*) Race-track microtron. (*c*) Double-sided microtron. (*d*) Hexatron. Also shown for the types in which the particles travel close to the speed of light is the restriction on the change in orbit radius, ΔR, in terms of the rf wavelength, λ, and the number of such wavelengths, ν, by which the orbit length increases between successive passages through the linac.

Much more space for the acceleration section is obtained by splitting the 360° magnet into two separated 180° magnets. However, the velocity of the injected electrons has to be close to the speed of light in this race-track microtron. In modern pulsed race-track microtrons, this is achieved by a hairpin-shaped orbit (Fig. 18*b*) which first guides the beam twice in opposite directions through the accelerating section (linac) before the actual microtron starts. The maximum energy of electrons emerging from the race-track microtron is limited to about 1 GeV because the weight of the magnets increases rapidly with energy. By replacing the 180° magnets with two segment-shaped 90° dipoles, the electron energy can be approximately doubled with the same amount of iron, and a second common axis for another linac is obtained (Fig. 18*c*). To achieve even higher energies, higher-order variants of the microtron geometry have been proposed (Fig. 18*d*).

For particles traveling close to the speed of light, c, successive orbits between two accelerating sections increase in length by an integral number of rf wavelengths (Fig. 18). The resulting restriction on the change in orbit radius, ΔR, is met by adjusting ΔW and the magnetic field in the bending magnets, B, to give the prescribed value of ΔR, according to Eq. (4), where q is the electron charge.

$$\Delta R = \frac{\Delta W}{qBc} \tag{4}$$

Focusing. Due to the fact that the revolution time in all microtrons is proportional to the particle energy, the central phase must be placed on the decaying slope of the rf wave. For instance, electrons starting with higher energy than the central particle will be delayed after one revolution and hence will meet the wave at lower rf field strength so that the energy deviation is reduced. Electrons that have energies lower than the central particle are corrected correspondingly. This effect of phase focusing leads to a very small energy spread of the electrons.

In the circular microtron, transverse focusing takes place only in the horizontal plane due to the geometrical effect that all electrons return to the starting point. Magnetic lenses on the accelerator axes and special field configurations in the bending magnets can be used to focus the beam both in the racetrack microtron and in higher-order microtrons. In the latter, additional effort is required to compensate for the vertical defocusing in the fringe field of the inclined entrance and exit edges.

Applications. The circular microtron finds applications in medicine and industry, furnishing a beam in the energy range of 10–50 MeV. Racetrack microtrons are used in nuclear physics and other research fields, and as injectors for synchrotrons and storage rings.

MAMI microtron cascade. In 1990, MAMI (for Mainz Microtron), a three-stage continuous-wave racetrack microtron cascade with a maximum energy of 855 MeV, was set into operation at the University of Mainz, Germany (**Fig. 19**). Its high-quality beam can be delivered alternatively to three experimental areas for nuclear and particle physics experiments and to one experimental hall for the production and utilization of coherent x-rays. The machine, which runs very reliably for typically 6000 hours per year, has been continuosuly improved, for example, by adding a source of spin-polarized electrons, setting up an energy calibration procedure with an absolute precision of about 100 keV, and installing a system to reduce beam energy fluctuations down to about 1 keV.

In 1999 the construction of a harmonic double-sided microtron (HDSM) as a fourth stage of MAMI was approved to increase the maximum beam energy to 1.5 GeV and thereby extend the experimental possibilities in particle and nuclear physics. It consists essentially of four 270-ton magnets and two continuous-wave linear accelerators with a maximum energy gain of about 9 MeV each (**Fig. 20**). The edge-defocusing in this machine is compensated for all of the 43 revolutions by a decrease in field perpendicular to the magnet edge. The mean bending field thereby drops during acceleration. Therefore, the energy gain per turn must also decrease, in order to fulfill the coherence condition. The energy

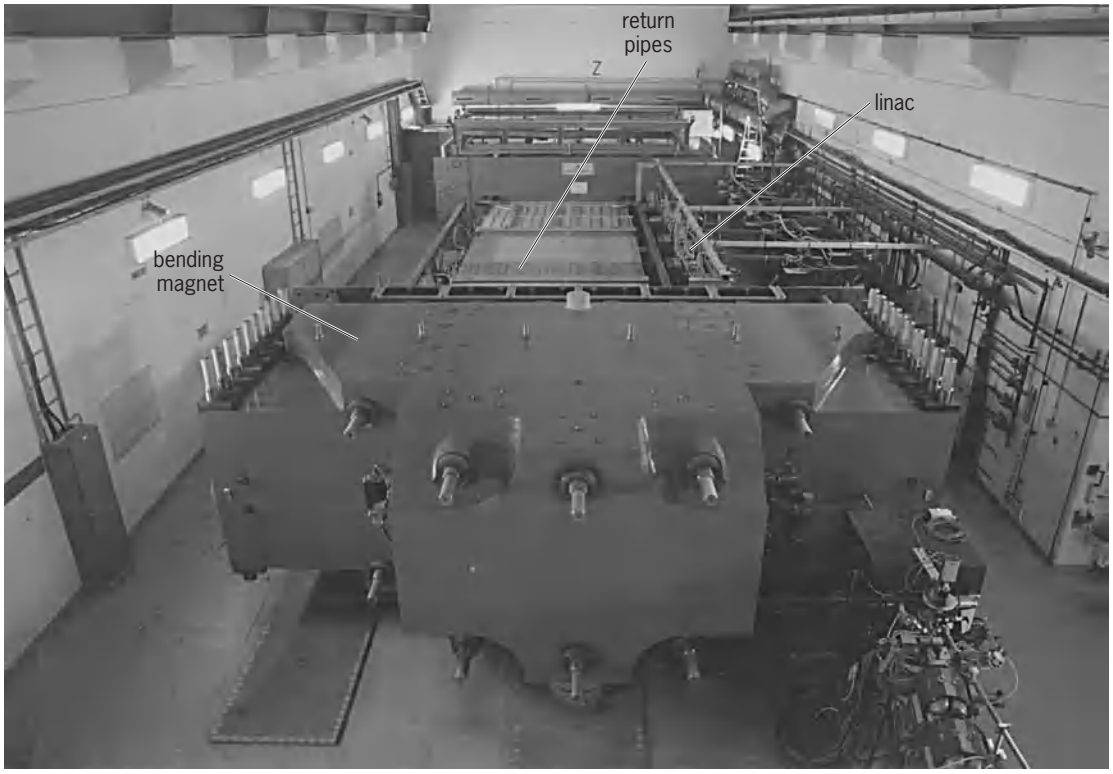

Fig. 19. Third stage of the MAMI continuous-wave microtron cascade. Injection beam energy is 180 MeV, and maximum energy is 855 MeV. One of the two 500-ton (450-metric-ton) bending magnets is in the foreground, and the 7.5-MeV linac and 90 return pipes are visible.

gain reduction is realized by a corresponding phase slide down on the accelerating wave. Due to phase focusing, this happens automatically if the beam is injected with the correct energy and phase. In order to improve the stability against rf misadjustments and drifts, one of the two linacs operates at the first suharmonic frequency. The first beam at full energy was generated in 2006. Karl-Heinz Kaiser

Thomas Jefferson National Accelerator Facility

The Thomas Jefferson National Accelerator Facility, or Jefferson Lab (JLab), in Newport News, Virginia, reached its original 4-GeV design energy in 1995 and its enhanced design energy of 6 GeV in 2000. (It was formerly known as the Continuous Electron Beam Accelerator Facility, or CEBAF.) The scientific purpose of the 6-GeV continuous-wave superconducting recirculating electron accelerator is to study quantum chromodynamics, the fundamental theory of the strong force, in the region where it is incredibly strong—the so-called confinement region. Fundamental to this effort is the precision with which electrons, interacting only through the electroweak force, allow one to probe the structure of matter. With its high current, high duty factor, and unprecedented polarization, Jefferson Lab's CEBAF offers world-leading capabilities. These permit the precise investigation of the motion of quarks within protons and neutrons as well as atomic nuclei, the distribution of protons and neutrons within nuclei, and the correlations between them. *See* FUNDAMENTAL INTERACTIONS; QUANTUM CHROMODYNAMICS; QUARKS.

To attain this combination of characteristics, the accelerator's beam performance objectives are an energy ranging from 0.5 to 6.0 GeV, a maximum current of 200 μA, a duty factor of 100%, a transverse emittance (radius times angular divergence) of $\sim 2 \times 10^{-9}$ meter-radian, and an energy spread of 10^{-5} times the beam energy. Beams are extracted from the accelerator for simultaneous use in three experimental halls.

A continuous-wave device is the approach of choice to produce a high-quality continuous beam. Low peak current for a given average current lowers emittance, and continuously operating rf systems can be controlled more precisely in phase and amplitude, thereby leading to smaller energy spread and smaller variations of average energy. Conceptually, the most straightforward continuous-wave accelerator would be a single linac in which the beam could attain the requisite energy in a single traversal. However, to reach an energy of 6 GeV in one traversal of a linac made up of the superconducting accelerating cavities used in Jefferson Lab would require 2600 ft (800 m) of active accelerating length. To avoid the high capital cost of such a linac, and to minimize total accelerator length, the beam is passed five times through a parallel pair of much shorter linacs connected by recirculation arcs (**Fig. 21**).

This approach is made possible by the fact that electrons move at very nearly the speed of light c at quite modest energies. At only 50 MeV, for instance, the electron velocity is $0.99995c$. Once fully relativistic, an electron's velocity is essentially independent of energy. Thus beams at different energies

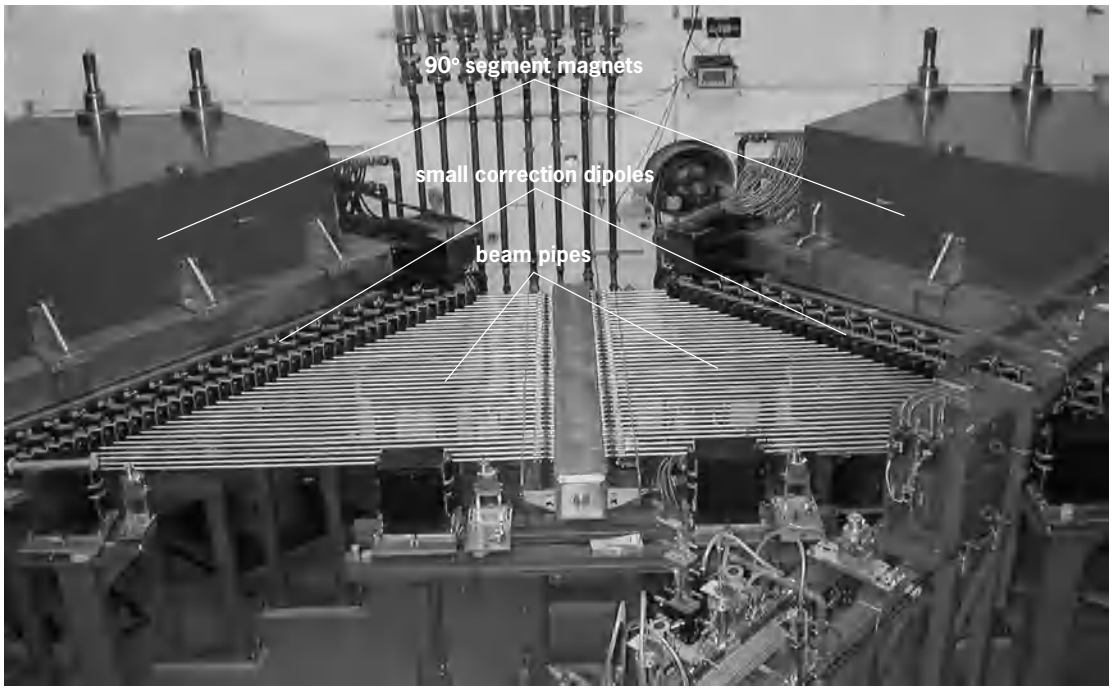

Fig. 20. Return beam pipes between two 7-m-wide (23-ft) segment magnets of the harmonic double-sided microtron (HDSM) of MAMI, immediately prior to completion.

can pass together through the linacs, all maintaining the proper phase relative to the rf field.

In this five-pass system, five electron beams at five different energies are simultaneously present on the same trajectory in the linacs. Each beam requires a separate recirculation path tuned to accommodate its energy. The beam transport lines in the recirculation arcs connecting the two linacs are achromatic and isochronous, provide matching in all phase-space coordinates, and are designed with adequate bend radii and strong focusing to minimize quantum excitation to preserve beam quality.

Accelerating cavities. The accelerator uses 1497-MHz, five-cell (20-in. or 0.5-m) superconducting niobium cavities. A waveguide at one end acts as the fundamental rf power input coupler and as a coupler for extracting some of the higher-order modes that are generated by the beam current; at the other end of the cavity, two waveguides serve as couplers to extract higher-order modes. In longitudinal section, the inner surfaces of the cells comprise elliptical segments; this elliptical shape reduces multipacting, a process that degrades the achievable electric field by leading to excessive heat loss in the cavity wall.

Cavities are paired for operation at 2 K (−456°F) within liquid-helium cryostats called cryounits. Every four such cryounits are integrally linked to form the accelerator's basic operating unit, the cryomodule. Each linac has 20 cryomodules for a total of 338 cavities in the accelerator, including 18 cavities in the injector. Cryomodules are interconnected by warm sections of beam line containing vacuum equipment, beam monitors, and magnets to focus and guide the beam.

Injector and distribution system. The 62.5-MeV injector provides a high-quality electron beam that is sufficiently relativistic to stay in phase with the rf system and the higher-energy, recirculated electron beams present in the first linac. The beam originates in an electron gun, passes through bunching, capture, and initial acceleration (up to 0.5 MeV) regions that operate at room temperature, is further bunched and accelerated to just over 5 MeV in a two-cavity cryounit, and then is accelerated to injection energy in two full-sized cryomodules.

The multiuser beam distribution system has two key elements: the injector and the rf separator (deflecting cavities) in the extraction line. The injector creates three interspersed bunch trains that can have different bunch charges. The rf separator deflects the beam, the optics amplify the initial deflection, and septum magnets extract the beam for simultaneous delivery to all three end stations. Hermann A. Grunder

Storage Rings and Colliders

A storage ring consists of an annular vacuum chamber embedded in a ring of bending and focusing magnets, in which a beam of long-lived, charged particles [e^- (electron), e^+ (positron), p (proton), or $\overline{p}$ (antiproton)], or two counterrotating charged particle–antiparticle beams (e^- and e^+, or p and $\overline{p}$) can be stored for many hours. Storage rings are the major components of circular particle colliders in which two particle beams are made to collide head-on. Storage rings also serve as particle accumulators in which successive particle bunches from an accelerator are added to produce a stored beam of relatively high intensity and with desired space, time, and momentum characteristics suitable for acceleration to higher energies for use in a collider.

Principles of colliding-beam systems. The motivation for using colliding-beam systems is found in the following kinematic considerations. Unlike fixed-target systems (cyclotrons, synchrotrons, linear

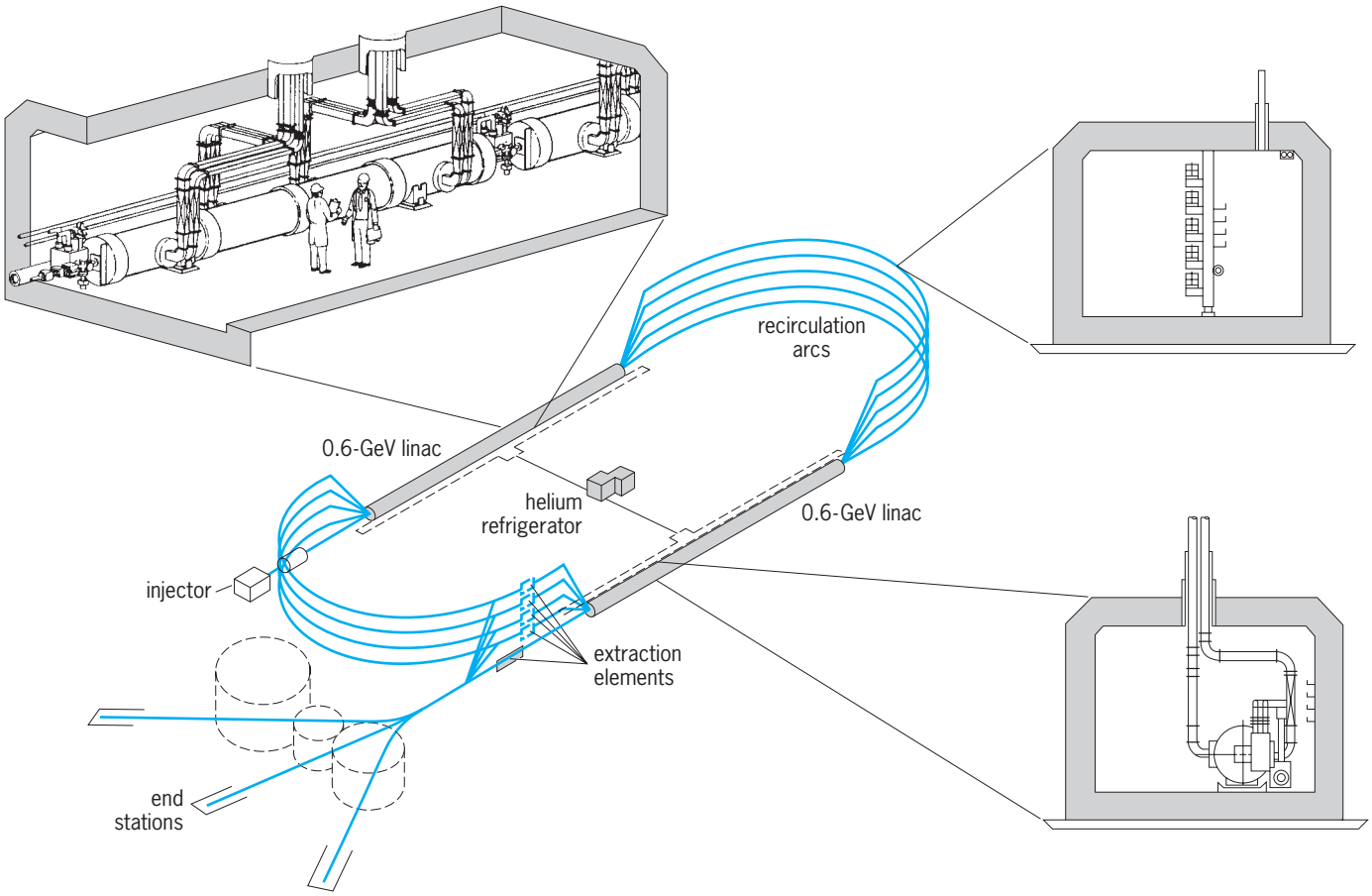

Fig. 21. Configuration of the Thomas Jefferson National Accelerator Facility (Jefferson Lab).

accelerators) in which a beam of accelerated particles traverses a fixed target, colliding-beam systems use the full energy of each particle to produce reactions. When two particles interact, the center-of-mass energy measures the energy available for reactions. For a fixed-target accelerator which produces a beam of high-energy relativistic particles hitting stationary target particles, this useful energy is only a fraction of the available beam energy and increases slowly as its square root. By contrast, when two particles of the same mass and energy collide head-on in a storage ring, all of the available energy is useful. For instance, the CERN storage ring LEP [Large Electron Positron Storage Ring, which was closed down in 2000 to make way for the Large Hadron Collider (LHC)] operated with e^+ and e^- beams of up to 104 GeV yielding a center-of-mass energy of 208 GeV. With a fixed e^- target, an e^+ beam of 42,000,000 GeV would be required to reach the same center-of-mass energy.

Low reaction rate. The high center-of-mass energy of colliding-beam systems is obtained at the expense of reaction rate, which is proportional to the density of target particles and to the beam intensity. It is helpful to think of one stored beam as the target for the other, and to consider the collision of a "beam bunch" with a "target bunch." In a typical e^+e^- storage ring the beam density is about 10^{14} particles/cm^3; a fixed target contains about 10^{24} electrons/cm^3. This huge difference is partially compensated by (1) the fact that at the interaction point each beam bunch collides with each target bunch f times per second (where f is the number of stored particle revolutions per second) and (2) the large intensity of the stored beams compared to accelerator beams which is the result of the accumulation of particles injected over a period of time. Because of the importance of these two factors for obtaining useful reaction rates, beam storage was used in all high-energy colliding beam systems placed in operation before 1987. However, there are limits to the density and intensity of beams that can be stored and made to collide, and thus colliding-beam reaction rates are always much lower than fixed-target reaction rates. Storage ring experiments are designed to take full advantage of the high center-of-mass energy within the limitations imposed by the relatively low reaction rates.

Luminosity. The performance of storage-ring colliding-beam systems is measured by the luminosity. By definition, this parameter is equal to the reaction rate, or number of interactions per second, divided by the interaction cross section.

The LEP storage ring had a circumference of 17 mi (27 km). To reach much higher e^+ and e^- energies than were achieved at LEP, the size of the storage ring would have to become unreasonably large. Instead, it appears possible to obtain useful luminosities in direct head-on collisions between two beams from two very high energy linear accelerators by focusing each beam to a very small cross-sectional area,

further discussed below. The SLAC-Linac-Collider (SLC) project is a prototype which took data at a center-of-mass energy of about 100 GeV from 1989 to 1998.

Circular positron-electron colliders. Positron-electron collisions occur in a storage ring composed of bending and focusing magnets enclosing a doughnut-shaped vacuum chamber in which counterrotating beams of e^+ and e^- are stored for periods of several hours. The two beams are made to collide with each other about 10^6 times per second in straight interaction sections of the vacuum chamber (**Fig. 22**). These are surrounded with detectors for the observation of collision products.

Collisions between e^+ and e^- particles are of particular interest since most of the resulting processes proceed through annihilation of both initial-state particles resulting in a particularly simple final state with the quantum numbers of the photon. This is one of the main reasons why e^+e^- collisions have proved to be particularly fruitful in deepening understanding of the fundamental structure of matter.

Beam instabilities. When particles from beam 1 (e^+ or e^-) pass through a more intense beam 2 (e^- or e^+) in the interaction region, they feel a strong force due to the electromagnetic field set up by beam 2. Thus, beam 2 acts as a very nonlinear lens on beam 1 and tends to cause a diffusive growth in the size of the latter. The force increases rapidly with the density of particles in beam 2 until it becomes so strong that the beam 1 area becomes suddenly very large and the luminosity decreases drastically; this is beam-beam instability. There also exist single-beam instabilities which are caused by electromagnetic forces between particles in the same beam or by fields set up by induced currents in the walls of the vacuum chamber. Maximum luminosity is achieved by operating with two beams of equal intensity just below the beam-beam instability limit, which in a well-designed storage ring is reached before single-beam instabilities become a problem.

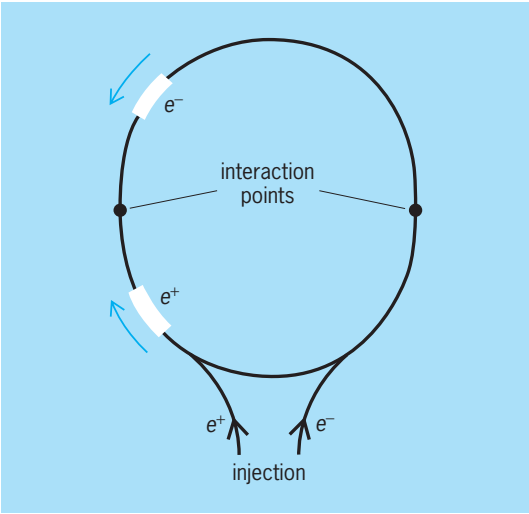

Fig. 22. Schematic diagram of e^+ e^- storage ring. This ring has two interaction regions and one bunch of particles circulating in each direction

Synchrotron radiation. The limitations discussed so far apply equally well in principle to e^+e^- as to p-p or p-$\bar{p}$ colliding-beam systems. There are important differences, however. Because they have a small mass, e^+ and e^- in circular orbits radiate a substantial amount of energy in the form of synchrotron radiation, just as they do in electron synchrotrons. The power radiated is proportional to the total number of orbiting particles (positions and electrons) and to the fourth power of the energy, and it is inversly proportional to the square of the orbit radius. This power, which has to be supplied by the rf system to keep particles on the design orbit, is appreciable and costly; at the peak operating energy it amounts typically to several megawatts in multi-gigaelectronvolt storage rings. In the design of a storage ring for a given energy and luminosity, the most economic solution requires a careful balance between the cost of magnets and buildings (which increases with increasing orbit radius) and the cost of the rf system (which decreases with increasing orbit radius since less power is needed). At the highest operating energies the amount of available rf power limits the luminosity, since for fixed values of radiated power and orbit radius only a steep decrease in the number of orbiting particles can compensate for an increase in energy.

There are also very beneficial consequences from the emission of synchrotron radiation. This emission dampens the amplitude of e^- and e^+ oscillations around the equilibrium orbit, and produces a gradual transverse polarization of the stored beams. As in electron synchrotrons, the radiation itself constitutes a very intense source of x-rays which has proved very useful in various branches of biology, chemistry, and physics. Not only are all operating e^+e^- colliders used as sources of synchrotron radiation, but many e^- storage rings have been built and are used specifically for investigations with this radiation.

Ring. The ring consists of circular sectors of bending and focusing magnets and of straight sections for rf cavities and interaction regions. The bending magnets guide the beam particles around the ring, and the focusing magnets drive them toward the equilibrium orbit around which they execute radial and vertical oscillations. The use of separate magnets for guiding and for focusing permits tight control of beam size to maximize the luminosity as the energy of operation is changed.

Injector. To achieve high luminosity, it is desirable to inject and accumulate e^- and e^+ with energies close to the desired collision energy. Linear accelerators and synchrotrons are used as injectors. The e^- are accelerated directly; e^+ are collected from a target bombarded with high-energy e^- and then accelerated. Accumulation occurs through the addition of particles first to the circulating e^+ beam, then to the circulating e^- beam until desired intensities have been reached. Typical accumulation times are seconds for e^- and minutes for e^+.

Vacuum chamber and system. Very low pressures of 10^{-9} and 10^{-10} torr (10^{-7} and 10^{-8} Pa) are required to prevent the outscattering of beam particles by

collisions with gas molecules, which reduces storage times below useful values. Special techniques are used to achieve such high vacuum in the presence of high-power synchrotron radiation hitting the chamber walls.

The rf power system. The power from the rf system replaces the energy emitted as synchrotron radiation and accelerates the stored beams to the desired energy. In a multigigaelectronvolt storage ring, several megawatts of power at 400–500 MHz are typically delivered to several rf cavities located in the ring.

Beam structure. The stored beams consist of bunches (typically 2 in. or 5 cm long), since rf power is being supplied continuously. For a given number of stored particles, maximum luminosity is achieved with the smallest number of bunches in each beam. This number equals half the number of interaction regions.

Interaction region. Magnet-free straight sections at the center of which e^+e^- collisions occur are provided for the installation of detectors. Special magnets are usually installed on each side of the interaction region to focus the beams to a very small area at the interaction point so as to maximize the luminosity; this is called low beta insertion. Karl Strauch

Proton-proton and p̄p colliders. Proton-proton and proton-antiproton storage rings or circular colliders provide the highest-energy particle beams used to search for and measure the fundamental constituents of matter and their associated forces. Hadron colliders, like circular electron-positron colliders, are synchrotrons and are constructed of a large number of long magnets and other equipment distributed in an approximately circular or racetrack configuration. The magnets have a small aperture or evacuated beam pipe through which the beam can pass repeatedly for many millions of revolutions while being deflected by the magnetic field into the necessary curved trajectory. The so-called bend or dipole magnets cause charged particles, such as protons or antiprotons, traveling in a vacuum chamber to follow the circular layout of the magnets. Additional magnetic elements (quadrupoles), interspersed in a regular way with the bend magnets, focus the beam of particles in a manner analogous to that in which an optical lens focuses light. The beams are injected into the storage rings in special regions or straight sections. Radio-frequency (rf) cavities used to accelerate the beams are also located in these regions, which are free from standard magnetic components.

In proton-proton colliders (**Fig. 23***a*), the two counterrotating beams are of the same charge, but moving in opposite directions. Each beam must be provided with a separate aperture with opposite magnetic field direction in order to follow the same curved trajectory. This can be achieved with two rings of magnets (above or below one another, or side by side), or the magnets can be built to contain the two apertures of opposite magnetic field direction. The beams must then be brought together at specific azimuthal locations around the rings to intersect or pass through one another in order that collisions between individual protons may take place. These regions, called interaction regions, are where high-energy physics detectors are assembled to measure the events from the proton-proton interactions. Proton-proton colliders have the advantage that the two separate rings allow the installation of two separate rf systems and thus the acceleration and storage of different particle species (for example, protons and ions with similar charge-mass ratio).

Proton-antiproton colliders (Fig. 23*b*) have the advantage that only one ring of magnets with one aperature is required. The protons and antiprotons are identical except that they have opposite charge. Because of this, the counterrotating beams can be guided in opposite directions in one ring of magnets with one sign of magnetic field. In this case, collisions are possible at many places around the ring's circumference. As the experiments are located at only a few locations, it is desirable to separate the beams from one another so that they collide only at the experimental areas. Separation into (typically) helical orbits within the single magnetic aperture can be provided by electrostatic deflectors. Though

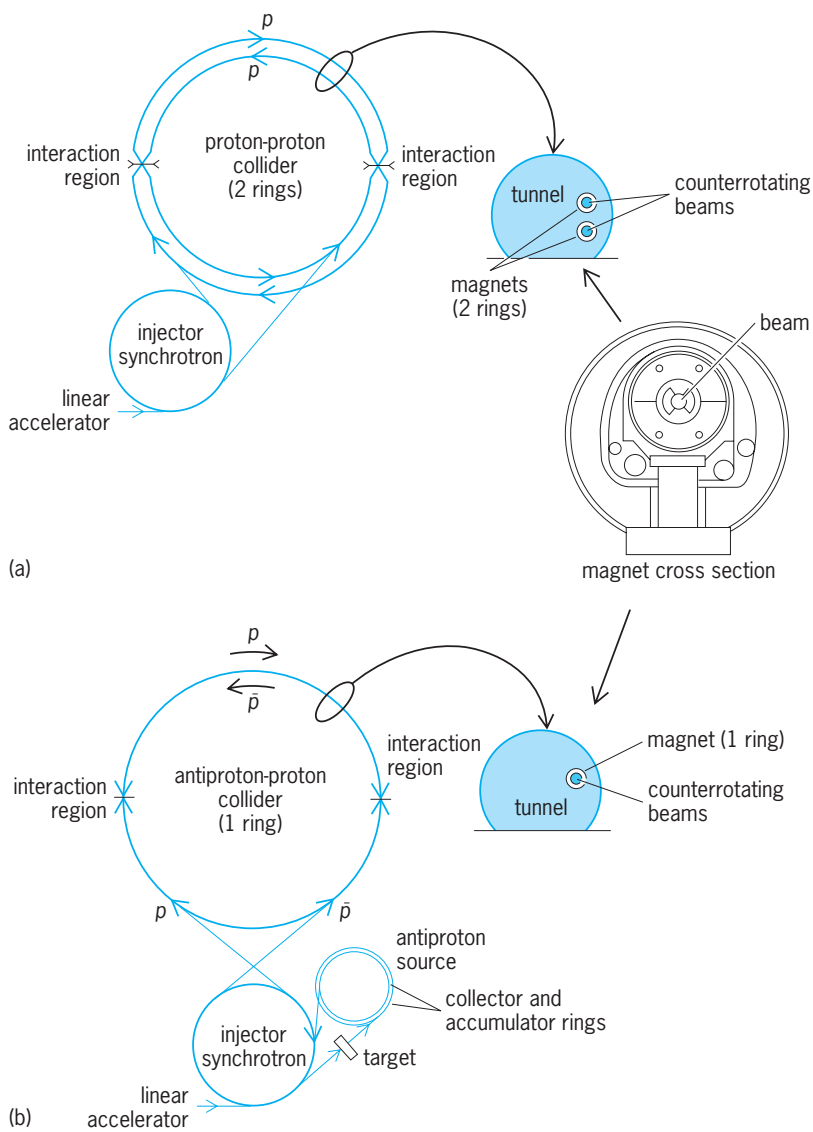

Fig. 23. Layout of colliders. (*a*) Proton-proton collider (two rings). (*b*) Proton-antiproton collider (one ring). The magnet cross section, common to both systems, is shown.

proton-antiproton colliders have a cost advantage in that only one magnet ring is required, they have the disadvantage that the antiprotons must be produced and the antiproton beam intensity is usually limited.

Event rate. At the experimental interaction regions, the beams in a proton-proton or proton-antiproton collider are tightly focused to make their spot size as small as possible. This is to enhance the collision interaction rate needed for high-energy physics experiments. As in positron-electron colliders, luminosity and the event cross section together determine the rate at which events take place. The total cross section for all types of interactions in proton-proton collisions is of the order of 100 millibarns (10^{-25} cm^2). Interesting events on the threshold of understanding have cross sections 10^{-8} of this, in the nanobarn to picobarn region (10^{-34} to 10^{-36} cm^2). Luminosities of 10^{30} to 10^{34} cm^{-2} s^{-1} are achievable in antiproton-proton and proton-proton colliders. The higher range of luminosity is realistic only in proton-proton colliders. Antiproton-proton colliders are limited by the number of antiprotons that can feasibly be produced and collected for use in the collider. The small ratio of the cross section for interesting events over the total cross section implies a large number of "uninteresting" events in proton-proton colliders. This background can be an important limitation for the maximum performance reach of a proton-proton collider.

Maximum energy. The primary limitation to the energy of proton-proton and proton-antiproton colliders is the strength of the magnetic guide field required to bend the protons or antiprotons in the circular orbit. Magnetic fields of the order of 4 to 9 T are produced by superconducting magnets. This field strength then sets the scale of the required circumference of the magnet ring. A continuous array of bending magnets with a $3^1/_3$-T field arranged on a 0.6-mi-radius (1-km) circle would provide the appropriate bending strength for 1-TeV-momentum particles. In practice, space required for magnetic focusing, other necessary beam manipulations such as acceleration from the injection energy (with rf cavities), the injection and extraction processes, and the interaction regions increase circumference of the racetrack to about $1^1/_3$ times that needed just for bending. Thus, the Fermilab Tevatron (**Table 6**), a 0.98-TeV collider with 4.0-T magnetic field, is 3.9 mi (6.3 km) in circumference.

The development of magnets using wire carrying superconducting current has made it possible to operate large rings of magnets at high magnetic excitation for long periods of time with reasonable power usage. If conventional copper magnets were used, power demands would severely limit the maximum beam energy that could be practically realized.

Whereas the energy of proton colliders is limited by the strength of the magnetic field and the circumference that can be realized at reasonable expense, electron storage rings are limited by the energy loss from synchrotron radiation. Both electron and proton rings are forced to larger ring circumferences as the particle energy increases. As discussed above, the Tevatron collider is 3.9 mi (6.3 km) in circumference for 0.9-TeV beam energy, whereas the largest electron collider (LEP at CERN) had a circumference of 17 mi (27 km) and beam energy of 104 GeV. LEP has been decommissioned, and the LEP tunnel is being used for installation of the Large Hadron Collider (LHC), a proton-proton collider which will have an energy of 7 TeV per beam, or 70 times of energy of the electron ring.

Thus, proton-proton and proton-antiproton storage rings would seem to have a large energy advantage over electron rings. This apparent advantage is, however, offset by the fact that the electron is a fundamental particle whereas the proton is made up of quarks and other constituents, only one of

TABLE 6. Proton-proton and proton-antiproton collider storage rings*

Name	Laboratory	Location	Beam energy	$\bar{p}$ stack rate	Luminosity (L), cm^{-2} s^{-1}	Circumference	Type	Status
Intersecting Storage Rings (ISR)	CERN	Geneva, Switzerland	32 GeV		10^{32}	0.6 mi (1 km)	*pp*, 2 rings, conventional magnets	First *pp* ring, decommissioned
Super Proton-Antiproton Synchrotron (S$\bar{p}$pS)	CERN	Geneva, Switzerland	315 GeV	4.8×10^{10}/h	2.8×10^{30}	4.3 mi (6.9 km)	$\bar{p}p$, 1 ring, conventional magnets	First collisions 1981
Tevatron	Fermilab	Batavia, Illinois	0.98 TeV	6×10^{10}/h	3×10^{31}	3.9 mi (6.3 km)	$\bar{p}p$, 1 ring, superconducting magnets	First collisions 1985
Large Hadron Collider (LHC)	CERN	Geneva, Switzerland	7 TeV		10^{34}	17 mi (27 km)	*pp*, 2-beam aperture in one magnet, superconducting magnets	First collisions planned for 2007
Relativistic Heavy Ion Collider (RHIC)	Brookhaven	Upton, New York	100 GeV/u 250-GeV protons		1.4×10^{31} (protons) 1×10^{27} (gold)	2.4 mi (3.8 km)	Heavy ions, 2 rings, superconducting magnets	First collisions 2000
HERA (Hadron Electron Ring Anlage)	DESY	Hamburg, Germany	26–820 GeV		2.0×10^{31}	3.9 mi (6.3 km)	*e-p*, 2 rings, superconducting magnets	First collisions 1991

* Including heavy-ion and electron-proton (*e-p*) colliders.

which is primarily involved in the collision process. Effectively the average useful energy of the protons is reduced by a factor of 3–6. Also, precision interaction energy measurements are more difficult with protons than with electrons. Conversely, proton colliders have a broader exploration potential because very high constituent interaction energies are possible, though at much lower effective luminosities because of the statistics of the energy distribution of the quarks within the proton.

Acceleration and beam control. As in proton synchrotrons, a cascaded chain of accelerators, each providing a factor of 10–20 in acceleration energy gain, is used. The final acceleration is performed in the storage ring itself. In order that colliders can be efficiently used, beams must be stored and collided for the order of a day before a fresh set of particles is injected, accelerated, and stored.

A beam of particles can be characterized by its density distribution in six-dimensional phase space, that is, by the number of particles in a volume element $dxdp_xdydp_ydzdp_z$. Here, p_x is the momentum conjugate to the coordinate x, and so on. In turn, the emittance is a characterization of the volume in phase space occupied by the beam. Typically in accelerators, emittance may be regarded as the product of two loosely coupled components, transverse emittance and longitudinal emittance. Transverse emittance is related to the distribution of position and angle of the particle trajectories and consequently is related to the beam spot size in the focusing lattice of the accelerator. Longitudinal emittance is related to the momentum distribution of the particles and their distribution along the major direction of motion around the ring; it includes the momentum spread of the beam and is important in determining the amount of rf accelerating voltage required and the bunch length of the beam.

Protons in synchrotrons and storage rings differ from their electron counterparts in that great care must be taken that the beam emittance is not enlarged during the acceleration process, during the transfer from one ring to another, or during the storage time when collisions are under way. Such phase space dilution decreases the luminosity and performance of the collider since the luminosity decreases inversely with the beam size at the interaction points. Many sources can cause dilution and even very small nonlinear magnetic fields, variations in power-supply current (ripple), or rf noise from the accelerating cavities can play a major role. An important problem in proton collider design is to calculate accurately how this dilution is expected to take place over hours or days of stored-beam operation. Electron storage rings do not have this problem because the synchrotron radiation which limits the energy also provides a damping mechanism for any enlargement of transverse spot size. *See* RIPPLE VOLTAGE.

Antiproton sources. The antiprotons in proton-antiproton storage rings are not available in nature and must be produced in order to be injected and stored in the collider each time the beams are replenished. The storage rings and associated equipment which provide for the antiproton acceleration is called the antiproton source (Fig. 23*b*). Typically, 20–150 GeV protons extracted from a stage of the proton injector chain are used to strike a target. Secondary particles produced from this target, of as large a momentum and angular acceptance as possible, are then gathered into a collector or debuncher synchrotron. Some of the particles captured in this collector ring will be antiprotons. Their momentum spread can be reduced by debunching the beam or converting very short bunches of particles with large momentum spread into a continuous uniform-density current with small momentum spread. This beam is then passed to the accumulator synchrotron, which stores and accumulates each successive cycle of antiproton beam as it is produced. Once a sufficient quantity of antiprotons has been accumulated and it is desirable to refill the collider, antiprotons are extracted from the accumulator, accelerated, and stored along with protons in the collider ring.

Antiprotons produced in the production target have a large spread in angle and velocity. Even after the manipulations in the collector-debuncher, they are still hot in the sense that they are spread over a large region of phase space (or emittance) that must be reduced to achieve the small beam size necessary for high-luminosity collisions. This cooling process is done in the accumulator. It is accomplished by a stochastic process that relies on monitoring for a very short time a sample of the beam and kicking it back to beam center with high-frequency kickers. This is done repeatedly over long periods of time as the different particles mix with different neighbors, making for different beam sample populations. This combination of damping and mixing in order to create new samples results in cooling or reducing the beam's emittance.

Existing storage rings. The first proton storage ring, the Intersecting Storage Rings (ISR) at CERN, is now decommissioned (Table 6). The next hadron collider to operate, the Super Proton Antiproton Synchroton (SppS), also at CERN, made use of the Super Proton Synchrotron (SPS), which is no longer operated for proton-antiproton collider physics.

The Tevatron is the highest-energy accelerator-synchrotron-storage ring in operation (800–900 GeV, with engineering runs at 980 GeV and plans for physics data taking there). It is also used as a fixed target proton synchrotron; the acceleration of protons in either mode of operation is discussed above. The antiproton source has two rings, the debuncher and the accumulator, operating at 8 GeV. Antiprotons are produced by extracting and targeting a 120-GeV beam from the main ring. The resultant 8-GeV antiprotons are collected first in the debuncher source ring at an approximately 2.5-s-cycle period and then are transferred to the accumulator source ring, where they are allowed to build up for a number of hours and the phase space of the antiproton beam is reduced or cooled. Six equally spaced bunches of 10^{11} protons and 3×10^{10} antiprotons are injected into the Tevatron and accelerated to full field, where they are

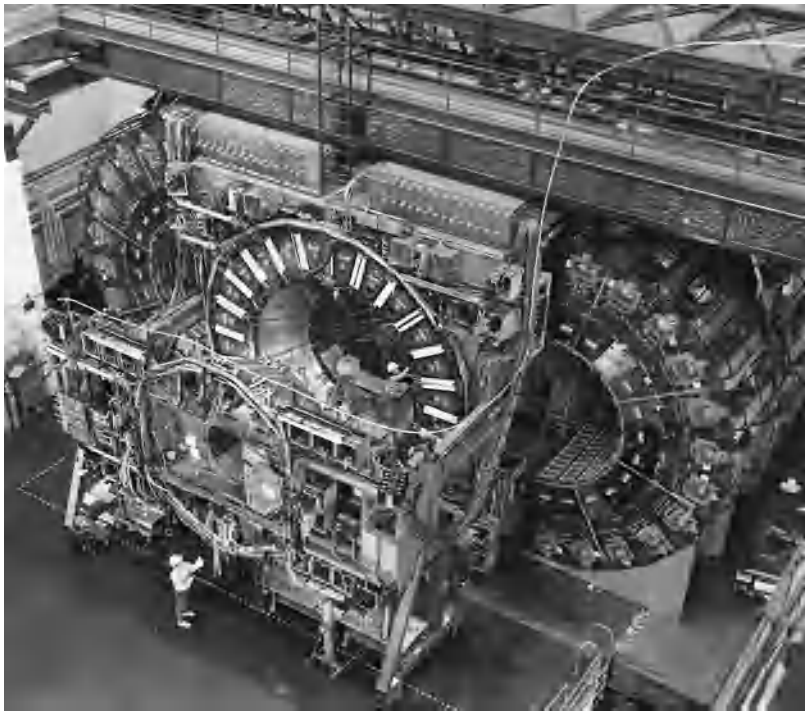

Fig. 24. One of the Fermilab colliding beam detectors used to observe collision events. The detector is made up of a central tracking region surrounding the beam pipe in which the colliding particles counterrotate. Outside the central tracking region is a superconducting magnetic solenoid coil that provides a magnetic field to help analyze the momentum of the produced particles. External to the solenoid are energy calorimeter modules or towers, which absorb, detect, and measure the energy of the events. As shown, the calorimeter modules are pulled away from the central region and coil for maintenance and checkout.

stored and collided for typically 15–20 h. Collisions take place at the two straight sections where the large experimental detectors are located (**Fig. 24**). The beams are separated elsewhere by electrostatic separators.

A second collider using superconducting magnets to circulate the proton beam is the HERA (Hadron Electron Ring Anlage) in Germany. This collider has two very different energy beams of electrons and protons contained in two separate rings, and collides 820-GeV protons against 27-GeV electrons in 170 bunches. It is the only collider of this type.

The third existing superconducting collider facility is the Relativistic Heavy-Ion Collider (RHIC) facility at Brookhaven National Laboratory. RHIC is a two-ring proton-proton and ion collider machine. Proton operation provides collisions of two polarized proton beams with energies of up to 250 GeV.

The Large Hadron Collider, under construction at CERN, is a proton-proton collider which will have a beam energy of 7 TeV and a luminosity of $10^{34}/\text{cm}^{-2}\text{s}^{-1}$. Thus, when it comes into operation, it will be the foremost collider for high-energy physics research. There will be over 1232 dipole magnets, each 49 ft (15 m) long and operating at a nominal magnetic field of 8.33 T, and 392 main quadrupole magnets, each 11 ft (3.25 m) long and operating at a nominal gradient of 223 T/m (resulting in a peak field at the coils of 6.85 T). The magnets are of a unique two-in-one design, with both beam apertures in one cold iron mass and cryostat. The magnets will operate at 1.9 K (−456°F). The cooling of the magnets will be done with superfluid helium.

Helen T. Edwards; Oliver Brüning

Heavy-ion collider. The theory of strong interactions predicts that a phase transition from the so-called normal nuclear structure to a plasma of quarks and gluons will occur in the extreme states of high temperature and high density created by relativistic heavy-ion collisions. Such states resemble the conditions of the early universe, shortly after the big bang. To investigate this new phase of nuclear matter requires a machine capable of providing collisions of heavy ions at energies of the order of 200 GeV/u. The Relativistic Heavy-Ion Collider (RHIC), a colliding-beam accelerator facility at Brookhaven National Laboratory, is dedicated to the investigation of this new phase of nuclear matter, and experiments with its four detectors began in 2000. RHIC can accelerate, store, and collide ions as heavy as gold with a maximum beam energy of 100 GeV/u. Experiments at RHIC have found that ultra-ligh-energy collisions of gold ions create hot and strongly interacting matter that behaves like a perfect fluid. The Large Hadron Collider will be outfitted with a heavy-ion injector and will be used as an ultra-hight-energy heavy-ion collider part of the time with maximum beam energy of 2.77 TeV/u. *See* QUARK-GLUON PLASMA; RELATIVISTIC HEAVY-ION COLLISIONS.

A heavy-ion synchrotron is used as an injector to a heavy-ion collider. Since heavy ions are available only in positive charge states, heavy-ion colliders, like proton-proton colliders, must have two separate storage rings intersecting at several collision points. In addition, for collisions of unequal species of ions, such as proton on gold, the two rings must fulfill either of two conditions in order to maintain the synchronized bunch crossings: (1) the two rings are operable at different magnetic fields to accommodate the difference in the magnetic rigidities of unequal species that have the same rotation frequency (or velocity); or (2) the energies of both beams are relativistic enough that the difference in the rotation frequencies is small and can be accommodated by an adjustment of the average radius of the closed orbits within the dynamic aperture of the rings.

The intrabeam scattering (Coulomb interaction among ions in the beam bunch), which scales as the fourth power of the ion charge divided by the square of the ion mass, and is thus enhanced for heavier ions, results in the rapid growth of transverse and longitudinal dimensions of the beam bunches. The transverse growth of the beam affects the lattice design (favoring short lattice cells and strong focusing) and magnet aperture. The longitudinal growth influences the rf system requirements. Since the bunch growth is particularly rapid at low energies, the time allowed for the injection beam stacking and for acceleration to the top energy must be kept very short (a few minutes). Transverse beam growth beyond the dynamic aperture (the so-called good-field region of the magnetic fields) of the accelerator lattice and longitudinal growth beyond the rf bucket (or capture

phase space) result in beam loss, which reduces the beam lifetime.

In addition to intrabeam scattering, there are a number of unavoidable beam depletion mechanisms, which are dictated by the physics of heavy-ion collisions. The highly Lorentz-contracted Coulomb field of heavy ions such as gold is very intense and produces copious electron pairs as the bunches pass through each other. Ions passing through the cloud of electrons thus produced have a high probability of capturing an electron. This change of charge state causes ions to be lost from the stable orbit, and is a significant beam-loss mechanism. The breakup of the nucleus into several lighter fragments (Coulomb dissociation) by the intense Coulomb field is also significant, but is much smaller than electron capture from the pair production. The beam loss from ordinary nuclear interactions is insignificant. Satoshi Ozaki

Linear electron-positron colliders. While the LHC is pushing the energy frontier for proton-proton collisions, electron-positron colliders with beam energies of 0.25–2.5 TeV are considered useful to complement the experimental program of the LHC. However, power loss due to synchrotron radiation, scaling with the fourth power of beam energy, puts a practical upper limit on achievable beam energies for electron-positron storage rings. A storage ring with energy significantly higher than the 104 GeV achieved with LEP is not considered reasonable.

This limitation can be overcome with a linear collider. In a linear collider an electron linear accelerator and a positron linear accelerator collide their beams heads on. The absence of magnetic bending of the beam orbits eliminates the synchrotron radiation losses and the achievable energy is given by the overall length times the effective accelerating field. Since the accelerated electrons and positrons pass the interaction point only once and are dumped downstream of the interaction point, beam spot size at the interaction point has to be kept very small to reach the required luminosities.

SLC, operated from 1989 to 1998 at SLAC with beam energy of 50 GeV, demonstrated the feasibility of the linear collider scheme. SLC, however, was a kind of hybrid between linear and circular collider. Both beams were accelerated simultaneously in a single linear accelerator, separated with a magnet after acceleration, and collided after appropriate bending in opposite directions by magnetic arcs (Fig. 16*a*).

Several approaches to linear colliders are being studied. They are based on superconducting as well as normal-conducting rf technology for the accelerating structures. Superconducting rf technology makes possible high efficiency in transforming electrical power into beam power, but imposes a fundamental limit on the accelerating field at about 50 MV/m. At this field level the magnetic surface fields in the superconducting cavities reach the critical magnetic field of niobium. In single-cavity rf tests, field levels close to this value have been achieved, while multicell cavities, as required for a linear accelerator, have reached values of about 35 MV/m. Improvements in the field strength and the performance reproducibility in series production for this kind of cavity is a major focus of research. *See* SUPERCONDUCTIVITY.

The superconducting technology for linear colliders was developed by the TESLA (TeV Superconducting Linear Accelerator) collaboration with the TESLA Test Facility (TTF) at DESY (Deutsches Elektronen Synchrotron) as a test facility, and has now been adopted by the International Linear Collider (ILC) collaboration, aiming for a collider in the 0.5–1 TeV center-of-mass energy range.

The normal-conducting designs have the potential to reach higher accelerating fields. Accelerating fields of close to 200 MV/m have been achieved in test structures, but problems related to rf arcs, material erosion, and overall efficiency considerations presently limit the usable accelerating fields to values in the 50–100-MV/m range. Improving these limits is another focus of research.

Major progress on normal-conducting technology for a linear collider with a 0.5–1.5 TeV center-of-mass energy range has been achieved by the NLC/JLC (Next Linear Collider/Japan Linear Collider) collaboration. Key feasibility issues were demonstrated at the NLCTA (Next Linear Collider Test Accelerator) test facility at SLAC. However, the NLC/JLC scheme was abandoned when the collaborating institutes joined in 2004 with TESLA in the ILC collaboration.

The possibility to achieve collision energies in the multi-TeV regime is studied by the Compact Linear Collider (CLIC) collaboration. This study is focused on normal-conducting, high-frequency structures to obtain high accelerating fields. In the CLIC scheme, the rf power to accelerate the main beam is not produced by klystrons but extracted from a second low-energy, high-current beam running parallel to the main linear accelerator. This concept of power transfer from drive beam to main beam has been demonstrated in two test facilities at CERN, the CLIC Test Facilities CTF1 (1992–1995) and CTF2 (1996–2002). Tests to demonstrate the generation scheme for the CLIC drive beam are ongoing at the test facility CTF3.

Other Linear collider research activities are focused on issues related to maximize the luminosity of linear colliders. Test facilities related to these issues are ATF (Accelerator Test Facility) at KEK in Tsukuba, Japan for the production of ultralow emittances, and Final Focus Test Beam (FFTB) at SLAC and ATF2 (at KEK) for beam focusing techniques adapted for the required nanometer-scale spot sizes in the interaction point. Hans Braun; Daniel Schulte

Bibliography. D. A. Bromley (ed.), *Large Electrostatic Accelerators*, 1974; P. J. Bryant and K. Johnsen, *The Principles of Circular Accelerators and Storage Rings*, 1993, paper 2005; A. W. Chao, *Physics of Collective Beam Instabilities in High Energy Accelerators*, 1993; A. W. Chao and M. Tigner (eds.), *Handbook of Accelerator Physics and Engineering*, 1999; D. A. Edwards and M. J. Syphers, *An Introduction to the Physics of High Energy Accelerators*, 1993; H. T. Edwards, The Tevatron energy doubler: A superconducting particle accelerator, *Annu. Rev. Nucl. Part. Sci.*, 35:605–660, 1985;

H. A. Grunder and F. B. Selph, Heavy-ion accelerators, *Annu. Rev. Nucl. Sci.*, 27:353–392, 1977; R. Hellborg (ed.), *Electrostatic Accelerators: Fundamentals and Applications*, Springer-Verlag, 2005; Institute of Electrical and Electronics Engineers, *Proceedings of the 2005 Particle Accelerator Conference*, and similar conferences every 2 years; S. P. Kapitza and V. N. Melekhin, *The Microtron*,1978; J. D. Lawson, *Physics of Charged-Particle Beams*, 2d ed., 1988; S. Y. Lee, *Accelerator Physics*, 2004; E. M. McMillian, A history of the synchrotron, *Phys. Today*, 37(2):31–37, February 1994; *Proceedings of the 2004 Linear Accelerator Conference* and similar conferences every 2 years; *Proceedings of the 2006 European Particle Accelerator Conference*, and other conferences in this series, Edinburgh; *Proceedings of the 17th International Conference on Cyclotrons and Their Applications*, Tokyo, 2005, and other conferences in this series; R. E. Rand, *Recirculating Electron Accelerators*, 1984; M. Reiser, *Theory and Design of Charged Particle Beams*, 1994; H. Schopper (ed.), *Advances in Accelerator Physics and Technology*, 1993; Y. N. Tsipenyuk, *The Microtron: Development and Applications*, 2002; T. Wangler, *RF Linear Accelerators*, 1998; H. Wiedemann, *Particle Accelerator Physics*, 2d ed., vol. 1, 1999, vol. 2, 1998; F. Wilczek, 10^{12} degrees in the shade, *The Sciences*, 34(1):22–30, January/February, 1994; E. J. N. Wilson, *An Introduction to Particle Accelerators*, 2001; R. R. Wilson, The next generation of particle accelerators, *Sci. Amer.*, 242(1):42–57, January 1980.

Particle detector

A device used to detect and measure radiation characteristically emitted in atomic, molecular, or nuclear processes, including photons (light, including visible light, gamma rays, and x-rays), lightweight charged particles (electrons or positrons), nuclear constituents (neutrons, protons, and heavier ions), energetic neutral particles (fast-moving atoms, for example), and subnuclear constituents such as mesons. The device can also be known as a radiation detector. Since human senses generally do not respond to these types of radiation (with the obvious exception of visible light), detectors are essential tools for the discovery of radioactive minerals, for all studies of the structure of matter at the atomic, nuclear, and subnuclear levels, and for protection from the effects of radiation. They have also become important practical tools in the analysis of materials using the techniques of neutron activation and x-ray fluorescence analysis. *See* ACTIVATION ANALYSIS; ELEMENTARY PARTICLE; NUCLEAR REACTION; NUCLEAR SPECTRA; PARTICLE ACCELERATOR; PROSPECTING; RADIOACTIVITY; X-RAY FLUORESCENCE ANALYSIS.

Classification by use. A convenient way to classify radiation detectors is according to their mode of use: (1) For detailed observation of individual photons or particles, a pulse detector is used to convert each such event (that is, photon or particle) into an electrical signal. (2) To measure the average rate of events, a mean-current detector, such as an ion chamber, is often used. Radiation monitoring and neutron flux measurements in reactors generally fall in this category. Sometimes, when the total number of events in a known time is to be determined, an integrating version of this detector is used. (3) Position-sensitive detectors are used to provide information on the location of particles or photons in the plane of the detector. (4) Track-imaging detectors image the whole three-dimensional structure of a particle's track. The output may be recorded by immediate electrical readout or by photographing tracks as in the bubble chamber. (5) The time when a particle passes through a detector or a photon interacts with it is measured by a timing detector. Such information is used to determine the velocity of particles and when observing the time relationship between events in more than one detector. *See* TIME-OF-FLIGHT SPECTROMETERS.

Ionization detectors. Any radiation-induced effect in a solid, liquid, or gas can be used in a detector. To be useful, however, the effect must be directly or indirectly interpretable in terms of either the quantity or quality (that is, type, energy, and so on) of the incident radiation or both. The ionization produced by a charged particle is the effect commonly employed.

Gas ionization detectors. In the basic type of gas ionization detector, an electric field applied between two electrodes separates and collects the electrons and positive ions produced in the gas by the radiation to be measured. Depending on the intensity of the electric field, the charge signal in the external circuit may be equal to the charge produced by the radiation, or it may be much larger. In a proportional counter, the output charge is larger than the initial charge by a factor called the gas amplification. A Geiger-Müller counter provides still larger signals, but each signal is independent of the original amount of ionization. All three types of gas ionization detectors can be used as pulse detectors, mean current detectors, or, with an indicator such as a quartz-fiber electroscope, as integrating detectors. *See* GEIGER-MÜLLER COUNTER; IONIZATION CHAMBER.

Position-sensitive and track-imaging detectors. Position-sensitive detectors and track-imaging detectors are often based on the ionization process. Multiwire proportional chambers and spark chambers are position-sensitive adaptations of gas detectors. The signal division or time delay that occurs between the ends of an electrode made of resistive material is sometimes used to provide position sensitivity in gas and semiconductor detectors. Track-imaging detectors rely on a secondary effect of the ionization along a particle's track to reveal its structure. In Wilson cloud chambers, ionization triggers condensation along particle tracks in a supercooled vapor; bubble formation in liquids is the basis for operation of bubble chambers. A secondary effect of ionization is also employed in photographic emulsions used as radiation detectors where ionization triggers the formation of an image.

Another type of position-sensitive detector uses a microchannel plate or an array of microchannel plates at an amplifier, with an imaging anode. This

technology is used in many practical applications, such as in image intensifiers and night-vision devices. *See* CHANNEL ELECTRON MULTIPLIER; IMAGE TUBE (ASTRONOMY).

Semiconductor detectors. In this type of detector, a solid replaces the gas of the previous example. The "insulating" region (depletion layer) of a reverse-biased *pn* junction in a semiconductor is employed. Choice of materials is very limited; very pure single crystals of silicon or germanium are presently the only fully suitable materials, although other semiconductors can be used in noncritical applications. Collection of the primary ionization is normally used, but an avalanche mode is sometimes employed in which an intense electric field causes charge multiplication. Although semiconductor devices are mostly used as pulse detectors, mean-current and integrating modes are possible when the significant leakage currents of *pn* junctions can be tolerated.

Since solids are approximately 1000 times denser than gases, absorption of radiation can be accomplished in relatively small volumes. A less obvious but fundamental advantage of semiconductor detectors is the fact that much less energy is required (~3 eV) to produce a hole-electron pair than that required (~30 eV) to produce an ion electron pair in gases. This results in better statistical accuracy in determining radiation energies. For this reason, semiconductor detectors have become the main tools for nuclear spectroscopy, and they have also made neutron activation analysis and x-ray fluorescence analysis of materials practical tools of great value. *See* CRYSTAL COUNTER; JUNCTION DETECTOR.

Scintillation detector. In addition to producing free electrons and ions, the passage of a charged particle through matter temporarily raises electrons in the material into excited states. When these electrons fall back into their normal state, light may be emitted and detected as in the scintillation detector. The early scintillation detectors consisted of a layer of powder (zinc sulfide, for example) that, when struck by charged particles, produced light flashes, which were observed by eye and counted. The meaning of the term "scintillation detector" has changed with time to refer to the combination of a scintillator and a photomultiplier tube that converts light scintillations into signals that can be processed electronically. Various organic and inorganic crystals, plastics, liquids, and glasses are used as scintillators, each having particular virtues in regard to radiation absorption, speed, and light output. *See* SCINTILLATION COUNTER.

Neutral particles. Neutral particles, such as neutrons, cannot be detected directly by ionization. Consequently, they must be converted into charged particles by a suitable process and then observed by detecting the ionization caused by these particles. For example, high-energy neutrons produce "knock-on" protons in collisions with light nuclei, and the protons can be detected. Slow neutrons are usually detected by using a nuclear reaction in which the neutron is captured and a charged particle is emitted. For example, in boron trifluoride (BF_3) detectors, neutrons react with boron to produce alpha particles which are detected. Another example is the use of a Channeltron or microchannel-plate detector, in which a fast neutral atom is incident on a surface from which electrons are emitted. The emitted electrons can then be converted to a pulse by an electron multiplication process (discussed below).

Other detector types. Although ionization detectors dominate the field, a number of detector types based on other radiation-induced effects are used. Notable examples are (1) transition radiation detectors, which depend on the x-rays and light emitted when a particle passes through the interface between two media of different refractive indices; (2) track detectors, in which the damage caused by charged particles in plastic films and in minerals is revealed by etching procedures; (3) thermo- and radiophotoluminescent detectors, which rely on the latent effects of radiation in creating traps in a material or in creating trapped charge; and (4) Cerenkov detectors, which depend on measurement of the light produced by passage of a particle whose velocity is greater than the velocity of light in the detector medium.

Another type of detector widely used in atomic and molecular physics is a Channeltron or microchannel-plate detector. High gain (amplification) is achieved by multiplication in secondary electron emission as a single incident particle initiates an electron cascade process. These detectors can be used to detect single particles in a counting mode, to give timing information for time-of-flight measurements, and to measure intensity or radiation flux. Microchannel plate detectors can also be used with an imaging anode to provide position-sensitive (image) detection. *See* CERENKOV RADIATION; PARTICLE TRACK ETCHING; TRANSITION RADIATION DETECTORS.

Large detector systems. The very large detector systems used in relativistic heavy-ion experiments and in the detection of the products of collisions of charged particles at very high energies, typically at the intersection region of storage rings, deserve special consideration. These detectors are frequently composites of several of the basic types of detectors discussed above and are designed to provide a detailed picture of the multiple products of collisions at high energies. The complete detector system may occupy a space tens of feet in extent and involve tens or hundreds of thousands of individual signal processing channels, together with large computer recording and analysis facilities.

The time projection chamber (TPC) typifies this class of detector. This chamber, in its initial configuration, is a cylinder 6.5 ft (2 m) long and 6.5 ft (2 m) in diameter, containing a gas, with a thin-plane central electrode that splits the chamber longitudinally into two drift spaces, each 3 ft (1 m) in length. A high potential (approximately 150 kV) is applied to this electrode to cause the electrons produced by ionization along tracks to drift to the end-cap regions of the cylinder. The end caps contain arrays of a few thousand gas proportional wire chambers designed to locate the position of signals with an accuracy of a small fraction of an inch in both radius and azimuth.

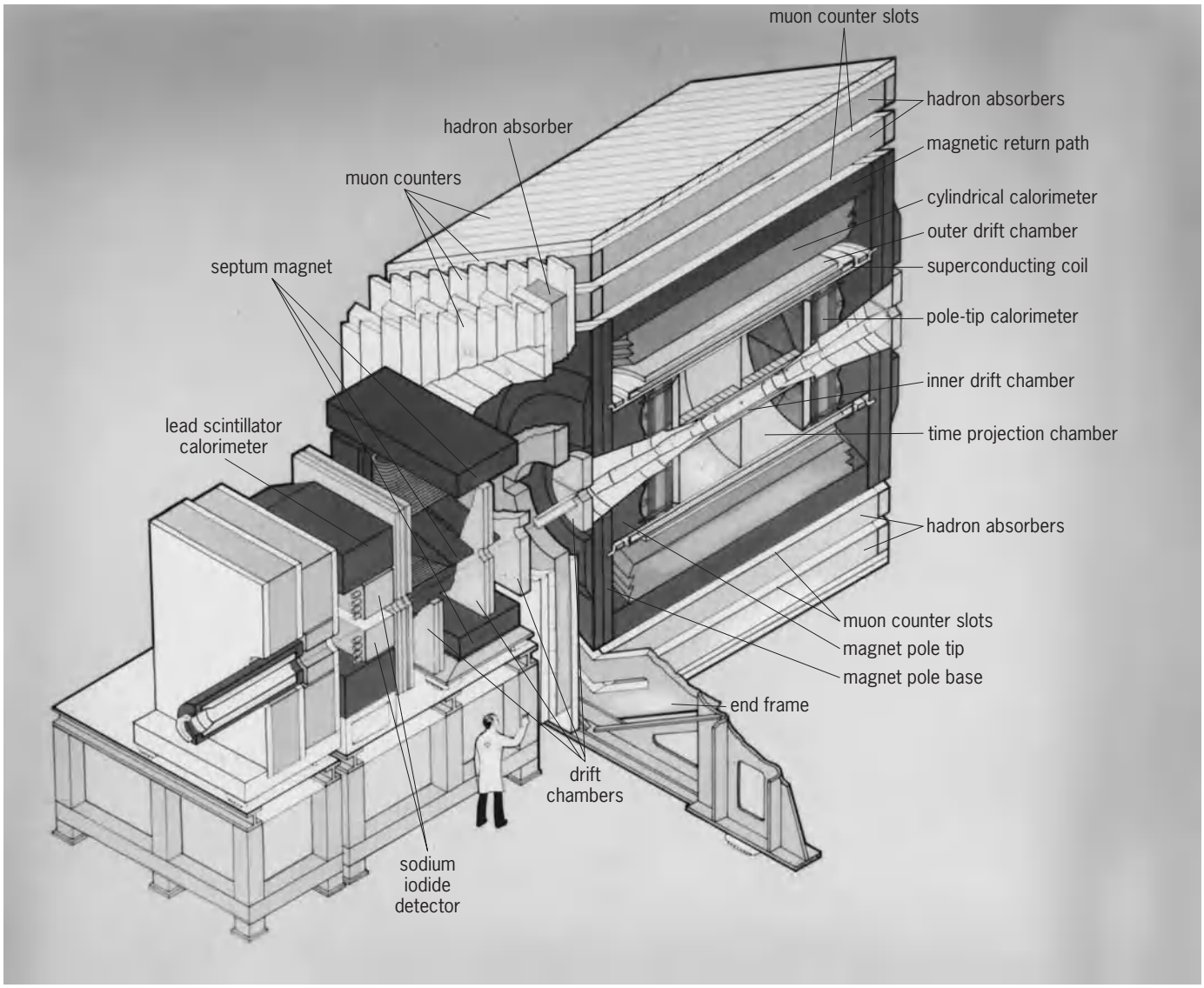

Time projection chamber designed for the PEP (Positron-Electron Projection) storage ring at Stanford University, with surrounding detection equipment and magnets.

The signal amplitude is maintained proportional to the original ionization in the track segment. When this information is combined with the drift time information, the effect is to provide the capability to measure the amount of ionization in individual cells a small fraction of an inch in linear dimensions throughout the whole volume of the chamber. The chamber is immersed in an axial magnetic field so that the bending of the tracks can be used to provide particle momentum information. Finally, the central time projection chamber is surrounded by other detector chambers (see illus.) to provide information on particles leaving the time projection chamber and also to perform ray calorimetry. The time projection chamber may be thought of as a modern version of the cloud or bubble chamber with prompt electrical readout of track information.

Very complex large-area and -volume chambers are also used in heavy-ion physics experiments. Here identification of the type of ion is usually important, and this is achieved partly by measuring the ionization or energy loss at several points along the track and by combining this information with time-of-flight or magnetic bending information. Complex combinations of gas, semiconductor, and scintillation detectors are frequently used in these detector systems. Fred S. Goulding; Alfred S. Schlacter

Bibliography. G. Charpak (ed.), *Research on Particle Imaging Detectors*, 1995; C. F. G. Delaney and E. C. Finch, *Radiation Detectors: Physical Principles and Applications*, 1992; February issues of *IEEE Trans. Nucl. Sci.*, 1970 to present; K. Kleinknecht, *Detectors for Particle Radiation*, 2d ed., 1988; G. F. Knoll, *Radiation Detection and Measurement*, 3d ed., 2000.

Particle flow

Particle flow is important to many industrial processes, including the pneumatic conveying of solids, transport of solids in liquids (slurries), removal of particulates from gas streams for pollution control, combustion of pulverized coal, and drying of

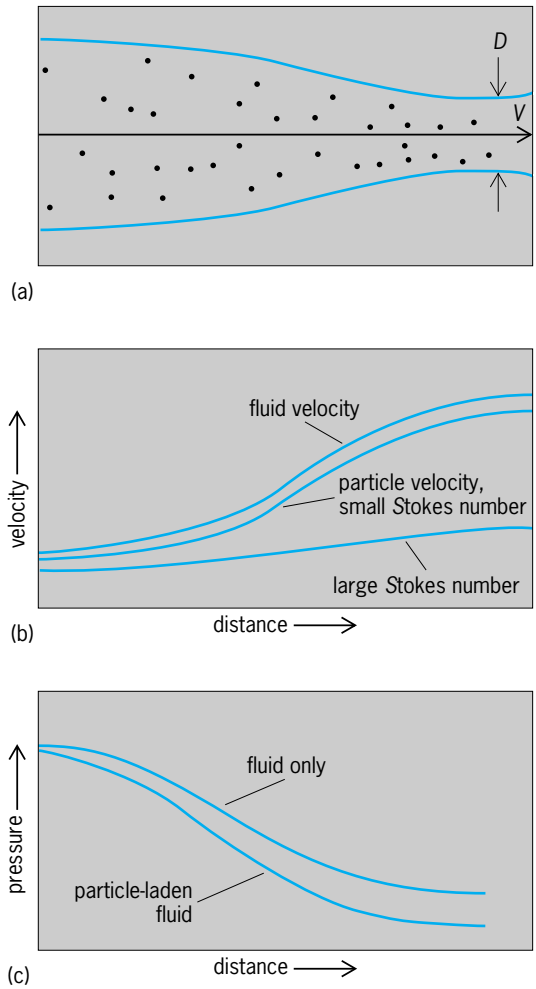

Fluid-particle flow in a converging duct. (*a*) Flow geometry. (*b*) Fluid and particle velocities along the duct in the cases of small and large Stokes numbers. (*c*) Effect of loading on pressure distribution along the duct.

particulates in the food and pharmaceutical industries. The vast majority of flow problems in industrial design involve the flow of gas or liquids with suspended solids. *See* DRYING; FLUIDIZED-BED COMBUSTION; PIPELINE.

Properties. The basic concepts of particle flows can be illustrated by referring to a specific example: the motion of a particle-laden fluid through a converging duct similar to a venturi entrance (**illus.** *a*). A key parameter in fluid-particle flows is the Stokes number, which is the ratio of the response time of a particle to a time characteristic of a flow system. Particle response time is the time a particle takes to respond to a change in carrier flow velocity. For example, if a particle is released from rest into a flow, the particle response time is the time required for the particle to reach 63% of the flow velocity. An example of a time characteristic of a flow system is L/V, where L and V are the characteristic length and velocity scales of the fluid flow.

If the Stokes number is small (say, less than 0.1), the particles have sufficient time to respond to the change in fluid velocity, so the particle velocity approaches the fluid velocity (illus. *b*). However, if the Stokes number is large (say, greater than 10), the particles have little time to respond to the varying fluid velocity and the particle velocity shows little change.

Loading. The relative concentration of the particles in the fluid is referred to as loading. The loading may be defined in several ways, but one definition frequently used is the ratio of particle mass flow to fluid mass flow. Many industrial applications involve highly loaded particle flows.

Coupling. Another term encountered frequently in the particle flow terminology is coupling. If the particle loading is small, the fluid will affect the particle properties (velocity, temperature, and so forth), but the particles will not influence the fluid properties. This is referred to as one-way coupling. If the conditions are such that there is a mutual interaction between the particles and fluid, the flow is two-way coupled.

For the above flow configuration, low particle loading will not affect the pressure distribution, so the flow will be one-way coupled. At high loadings, the pressure drop in the fluid is larger (illus. *c*) because of the momentum needed by the fluid to accelerate the particles. This interaction is defined as two-way coupling. Two-way coupling effects are reduced with increasing Stokes number because the particles undergo less acceleration.

Dilute and dense flows. If the particle motion is controlled by the action of the fluid on the particle, the flow is termed dilute. But, if the particle concentration is sufficiently high, the particles will collide with each other and their motion will be dependent on particle-particle collisions; the flow is then regarded as dense.

Pneumatic transport. An interesting example of a gas-particle flow is pneumatic transport in a horizontal duct. At high velocities the particles are kept in suspension by the fluid motion (homogeneous flow). The flow is essentially dilute. In this case, there is no acceleration of the gas, but the particles are continually accelerating as they collide with the wall and lose their velocity before reentrainment into the gas stream. The force required to accelerate the particles toward the gas flow velocity leads to a higher pressure gradient (larger pressure loss per unit length of tube) or a two-way coupling effect.

As the velocity of the airflow in the horizontal pipe is reduced, the particles tend to settle out on the tube wall. The velocity at which this occurs is called the saltation velocity. The accumulation of particles on the wall leads to a reduced cross-sectional area for the flow, a larger pressure loss for a given flow rate, and the onset of unsteady flow. A series of particle-gas flow patterns are observed as the particles tend to accumulate on the wall. First, dunelike formations occur. A further reduction in velocity leads to the dense-flow condition, where the particles nearly fill the pipe and move through it in a sequence of slugs. For even lower velocities, the particles will plug the pipe, and the air flows through a stationary packed bed or the bed may slide through the pipe. In this case, the particle-particle interaction controls the motion. *See* FLOW OF SOLIDS.

Slurry transport. Slurry transport is similar to pneumatic transport but, because of buoyancy, the liquid

is more capable than a gas of supporting the particles. A slurry is often treated as a single-phase non-newtonian fluid with empirical relationships for shear stress. As with pneumatic transport, there is a critical velocity at which the solid phase begins to settle on the pipe wall, which is called the deposition velocity. It is very important to design a slurry transport system with adequate velocities to avoid the immense problems associated with plugging.

Measurements. Measurements of particle flows are generally difficult to carry out. Instrumentation such as phase-Doppler anemometry, flow visualization with digital image analysis, and other nonintrusive techniques have provided important information. Measurements in dense particle systems are particularly difficult. *See* FLOW MEASUREMENT.

Numerical models. Significant advances have been made in numerical models for particle flows. These models require high-speed computational and extensive storage capability. Also, many of the models are based on empirical input that must be provided by experiment. Still, the models are sufficiently viable to be used to complement the design of particle flow systems. *See* COMPUTATIONAL FLUID DYNAMICS; FLUID FLOW; PARTICULATES. Clayton T. Crowe

Bibliography. G. Hetsroni (ed.), *Handbook of Multiphase Systems*, 1982; G. E. Klinzing, *Gas-Solid Transport*, 1981; G. Matsui et al., *Proceedings of the International Conference on Multiphase Flows—'91 Tsukuba*, vol. 3, 1991; M. C. Roco (ed.), *Particulate Two-Phase Flow*, 1993; C. A. Shook and M. C. Roco, *Principles and Practice of Slurry Flow*, 1993.

Particle track etching

A technique of selective chemical etching to reveal tracks of heavy nuclear particles in a wide variety of solid substances. Developed in order to see fossil particle tracks in extraterrestrial materials, the technique finds application in many fields of science and technology.

Identification of nuclear particles. An etchable track is produced if the charged particle has a sufficiently high radiation-damage rate and if the damaged region in the solid is permanently localized. Thus only highly ionizing particles are detectable; only nonconductors record tracks; and radiation-sensitive plastics can detect lighter particles than can radiation-insensitive minerals and glasses. The conical shape of the etched track depends on the ratio of the rate of etching along the track to the bulk etching rate of the solid. Careful measurements with accelerated ions of known atomic number Z and velocity $v = \beta c$ (where c = velocity of light) have shown that this ratio is an increasing function of Z/β. Measurements of the shapes of etched tracks thus serve to identify particles (**Fig. 1**). The most sensitive detector, a plastic known as CR-39, detects particles with Z as low as 1 (**Fig. 2**) provided $Z/\beta \gtrsim 10$. The clarity and contrast of the images and its high sensitivity make CR-39 a very attractive track-etch detector for nuclear physics, cosmic-ray research, element mapping, personnel neutron dosimetry, and many other applications. Minerals, on the other hand, are insensitive to particles with $Z/\beta \lesssim 150$ and are therefore useful in recording rare, very heavily ionizing, relatively low-energy nuclei with $Z \gtrsim 25$. *See* CHARGED PARTICLE BEAMS.

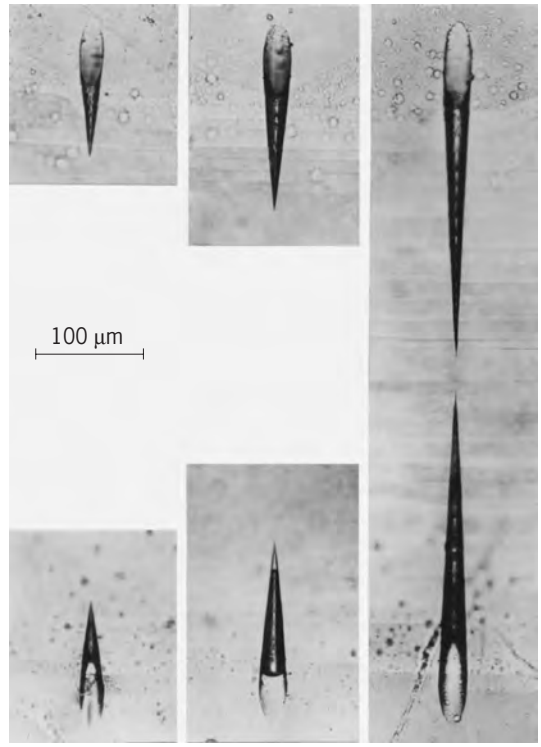

Fig. 1. Upper images show increasing radiation damage produced by a high-energy uranium nucleus in the cosmic radiation as it penetrates a thick stack of Lexan plastic sheets and slows down. The lengths of the etched cones in these three sheets, taken at intervals of several millimeters from within the stack, permit the atomic number and energy of the nucleus to be determined. Lower images show the exit points of the tracks.

Solar and galactic irradiation history. The lunar surface, meteorites, and other objects exposed in space have been irradiated by charged particles from a variety of sources in the Sun and the Galaxy. The

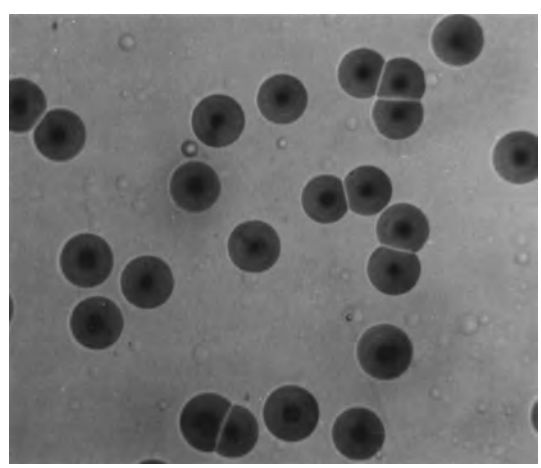

Fig. 2. Etched conical pits in a sheet of allyl diglycol carbonate (CR-39 plastic) irradiated at normal incidence with 60-MeV alpha particles. (*From P. B. Price et al., Do energetic heavy nuclei penetrate deeply into the Earth's atmosphere?, Proc. Nat. Acad. Sci., 77:44–48, 1980*)

particles of lowest energy are produced by the expanding corona of the Sun—the solar wind. Arriving in prodigious numbers, but penetrating only some millionths of a centimeter, the solar wind particles quickly produce an amorphous radiation-damaged layer on crystalline grains. At depths from about 0.00004 to 0.04 in. or 1 micrometer to 1 mm, individual tracks produced in solar flares (sporadic, energetic outbursts on the Sun) can be resolved by electron microscopy or sometimes by optical microscopy. At depths greater than 1 mm, most of the particle tracks are produced by heavy nuclei in the galactic cosmic radiation. *See* SOLAR WIND.

Comparison of fossil particle tracks in lunar rocks and meteorites with spacecraft measurements of present-day radiations has established that solar flares and galactic cosmic rays have not changed over the last 2×10^7 years—the typical time a lunar rock exists before being shattered by impacting interplanetary debris. Observations of grains in stratified lunar cores and lunar and meteorite breccias (**Fig. 3**) enable the particle track record to be extended back more than 4×10^9 years in time. Breccias, which are complex grain assemblages, often contain grains that have high solar flare track densities on their edges, indicating exposure to free space prior to breccia formation. Dating work, using spontaneous-fission tracks, indicates that some of the breccias were assembled soon after the beginning of the solar system some 4.6×10^9 years ago. The study of the time history of energetic radiations in space, using lunar samples and meteorites, elucidate various dynamic processes such as rock survival lifetimes, microerosion of rocks, and the formation and turnover rates of planetary regoliths. *See* METEORITE; MOON.

Nucleosynthesis. In terrestrial crystals, which are well shielded from external radiations, the dominant source of tracks is the spontaneous fission of ^{238}U. Certain meteorites and lunar rocks (Fig. 3*a*) contain additional fission tracks due to the presence of ^{244}Pu (half-life, 8×10^7 years) when the crystal was formed. After corrections are made for chemical fractionation of Pu and U, the data indicate that a large spike of newly synthesized elements was produced at the time of formation of the Galaxy, followed by a period of continuous synthesis and relatively rapid mixing. *See* MILKY WAY GALAXY; NUCLEOSYNTHESIS.

Current solar and galactic irradiation. Studies of tracks in a piece of glass from the *Surveyor 3* spacecraft after a 2.6-year exposure on the lunar surface, and of tracks in plastic detectors exposed briefly above the Earth's atmosphere in rockets, led to the surprising discovery that the Sun preferentially ejects heavy elements in its flares rather than an unbiased sample of its atmosphere. *See* SUN.

The existence of galactic cosmic rays with atomic number greater than 30 was discovered in 1966 when fossil particle tracks were first studied in meteorites. Since then many stacks of various types of plastics and nuclear emulsions up to 220 ft^2 (20 m^2) in area have been exposed in high-altitude balloons and in Skylab in order to map out the composition of the heaviest, rarest cosmic rays. Several particles heavier than uranium have been detected, indicating that cosmic rays originate in sources where synthesis has proceeded explosively beyond uranium. Exposures of giant detectors of about 1100 ft^2 (100 m^2) for a year in space are planned. Hybrid detectors using stacks of track-recording plastics to measure range, and photomultiplier tubes to measure light from Cerenkov radiation and scintillation detectors, have made it possible to determine relative abundances of the isotopes of very heavy elements in the cosmic rays. These data have a bearing both on the history of cosmic rays and on nucleosynthesis in their sources. *See* COSMIC RAYS.

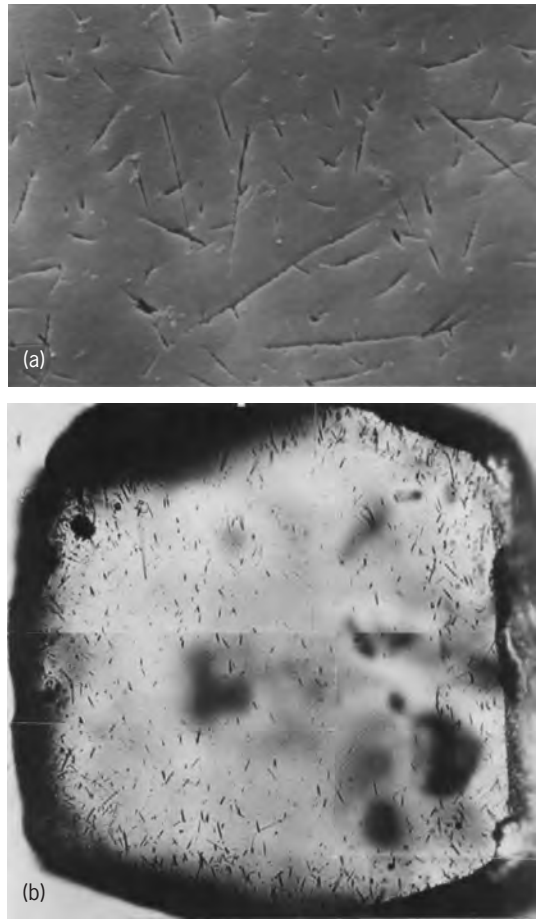

Fig. 3. Etched fossil particle tracks. (*a*) Tracks from spontaneous fission of ^{238}U and ^{244}Pu in a zircon crystal from a lunar breccia, seen by scanning electron microscopy (*from D. Braddy et al., Crystal chemistry of Pu and U and concordant fission track ages of lunar zircons and whitlockites, Proceedings of the 6th Lunar Science conference, pp. 3587–3600, 1975*). (*b*) Tracks of energetic iron nuclei from solar flares, showing a decreasing concentration from edges to center of a 150-μm-diameter olivine crystal from a carbonaceous chondritic meteorite. The irradiation occurred about 4×10^9 years ago, before the individual grains were compacted into a meteorite (*from P. B. Price et al., Track studies bearing on solar-system regoliths, Proceedings of the 6th Lunar Science Conference, pp. 3449–3469, 1975*).

Nuclear and elementary particle physics. Unique advantages of etched-track detectors are their ability to distinguish heavy-particle events in a large background of lightly ionizing radiation and their ability to detect individual rare events by a specialized technique such as electric-spark scanning or

ammonia penetration through etched holes. These advantages have permitted such advances as the measurement of very long fission half-lives; the discovery of ternary fission; the determination of fission barriers; the production of numerous isotopes of several far-transuranic elements; the discovery of several light, neutron-rich nuclides such as ^{20}C at the limit of particle stability; and highly sensitive searches for magnetic monopoles, superheavy elements, and anomalously dense nuclear matter in nature and in accelerators. *See* MAGNETIC MONOPOLES; NUCLEAR FISSION; TRANSURANIUM ELEMENTS.

Geochronology. The spontaneous fission of ^{238}U, present as a trace-element purity, gives tracks that can be used to date terrestrial samples ranging from rocks to human artifacts. Because fission tracks are erased in a particular mineral at a well-defined temperature (for example, 212°F or 100°C in apatite), one can use the apparent fission-track ages as a function of distance from the heat source to measure the thermal (tectonic) history of regions. Examples are the rate of sea-floor spreading (about 1 cm per year) and the surprisingly rapid rate of uplift of the Alps (1.2×10^{-2} to 5.6×10^{-2} in. or 0.3 to 1.4 mm per year) and of the Wasatch range in Utah (4×10^{-3} to 1.6×10^{-2} in. or 0.01 to 0.4 mm per year). Because of its very high uranium concentration (typically a few hundred parts per million), a few tiny zircon crystals less than 4×10^{-3} in. (0.1 mm) in diameter suffice to give a fission track age of sedimentary volcanic ash and thus to determine absolute ages of stratigraphic boundaries. Occasional, anomalously low fission-track ages can sometimes be used to locate valuable ore bodies and even petroleum. *See* FISSION TRACK DATING.

Geophysics and element mapping. In these applications a track detector placed next to the material being studied records the spatial distribution of certain nuclides that either spontaneously decay by charged particle emission or are induced to emit a charged particle by a suitable bombardment. A resolution of a few micrometers is easily attained. Such micromaps make it possible to identify radionuclides in atmospheric aerosols, to measure the distribution and transport of radon, thorium, and uranium, and to measure sedimentation rates at ocean floors. Thermal neutrons, readily available in a reactor, can be used to map ^{235}U (via fission), ^{10}B (via alpha particles), ^{6}Li (via tritons), ^{14}N (via protons), and several other nuclides. Deuterium, a tracer in biological studies, and lead and bismuth can also be mapped by using different irradiations. *See* RADIOACTIVE TRACER.

Practical applications. Filters are produced by irradiating thin plastic sheets with fission fragments and then etching holes to the desired size. Uses include biological research, wine filtration, and virus sizing. In virus sizing a single hole is formed that separates two halves of a conducting solution. When a virus or other tiny object passes through the hole, the resistance increases drastically, and the size, shape, and speed of the object can be determined by analyzing the electric signal.

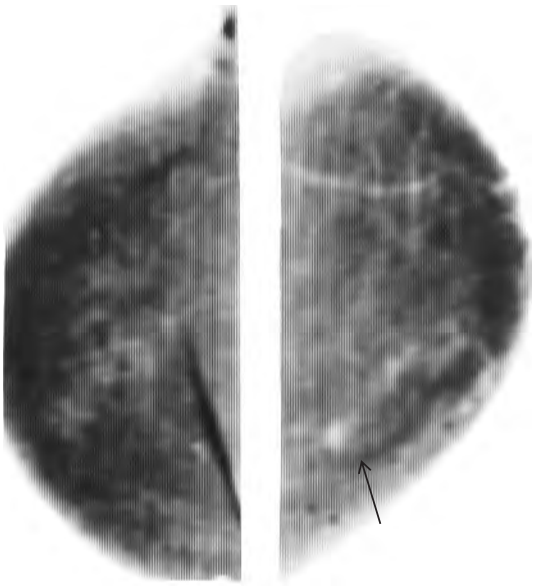

Fig. 4. Reconstructed images of right and left breasts of a patient taken by using a 3600-MeV carbon-ion beam at the Lawrence Berkeley Laboratory Bevalac accelerator. Arrow indicates 1-cm (0.4-in.) carcinoma. (*From C. A. Tobias et al., Lawrence Berkeley Laboratory*)

There are numerous other uses. A uranium exploration method relies on a survey of radon emanation, as measured by alpha-particle tracks in plastic detectors, to locate promising locations in which to drill. CR-39 plastic detectors are used in a Fresnel zone-plate imaging technique to make high-resolution images of the thermonuclear burn region in laser fusion experiments. Plastic detectors are also used in conjunction with a beam of high-energy heavy ions to take radiographs of cancer patients that reveal details not detectable in x-rays. **Figure 4** shows reconstructed images of right and left breasts of a patient taken by using a 3600-MeV carbon-ion beam at the Lawrence Berkeley Laboratory Bevalac accelerator. A single beam pulse is passed through each breast as it is immersed in water, and the stopping points of the ions are recorded in a stack of 30 cellulose nitrate sheets, each 0.01 in. (0.025 cm) thick, which are located behind the water bath. After the sheets are etched, the information on each sheet is converted to digital data by a scanning system, and two images capable of revealing slight differences in density are displayed on a television screen. The patient dose for such images is about 20 to 100 millirads (200 to 1000 micrograys), which is considerably lower than the usual dose from x-ray mammographic imaging. *See* RADIOGRAPHY.

P. Buford Price

Bibliography. B. G. Cartwright, E. K. Shirk, and P. B. Price, A nuclear-track-recording polymer of unique sensitivity and resolution, *Nucl. Instrum. Meth.*, 153:457–460, 1978; R. L. Fleischer, Where do nuclear tracks lead?, *Amer. Sci.*, 67:194–203, 1979; R. L. Fleischer, P. B. Price, and R. M. Walker, *Nuclear Tracks in Solids*, 1975; R. M. Walker, Interaction of energetic nuclear particles in space with the lunar surface, *Annu. Rev. Earth Planet. Sci.*, 3:99–128, 1975.

Particle trap

A device used to confine charged or neutral particles where their interaction with the wall of a container must be avoided. Electrons or protons accelerated to energies as high as 1 teraelectronvolt (10^{12} electronvolts) are trapped in magnetic storage rings in high-energy collision studies. Other forms of magnetic bottles are designed to hold dense hot plasmas of hydrogen isotopes for nuclear fusion. At the other end of the energy spectrum, ion and atom traps can store isolated atomic systems at temperatures below 1 millikelvin. Other applications of particle traps include the storage of antimatter such as antiprotons and positrons (antielectrons) for high-energy collision studies or low-energy experiments. *See* ANTIMATTER; NUCLEAR FUSION; PARTICLE ACCELERATOR; PLASMA (PHYSICS); POSITRON.

Charged-particle traps. Charged particles can be trapped in a variety of ways. An electrostatic (Kingdon) trap is formed from a thin charged wire. The ion is attracted to the wire, but its angular momentum causes it to spiral around the wire in a path with a low probability of hitting the wire.

A magnetostatic trap (magnetic bottle) is based on the fact that a charged particle with velocity perpendicular to the magnetic field lines travels in a circle, whereas a particle moving parallel to the field is unaffected by it. In general, the particle has velocity components both parallel and perpendicular to the field lines and moves in a helical spiral. In high-energy physics, accelerators and storage rings also use magnetic forces to guide and confine charged particles. The earliest form of such a trap is the cyclotron, used to accelerate protons. A tokamak has magnetic field lines configured in the shape of a torus, confining particles in spiral orbits (**Fig. 1***a*). This type of bottle is used to contain hot plasmas in nuclear fusion studies. Another type of bottle uses a magnetic mirror. As the particle approaches the ends of the bottle (Fig. 1*b*), the pinched magnetic field causes it to spiral into ever-tightening helixes. The time-averaged circular motion of the particle is effectively a current loop with an associated magnetic moment that is repelled by the gradient of the magnetic bottle field. Ions are trapped in the Earth's magnetic field (the van Allen belts) by this effect. This type of trap has been used to accumulate positrons for simulations of intergalactic space and to help produce positronium for laser spectroscopy. *See* ELECTRON MOTION IN VACUUM; POSITRONIUM; VAN ALLEN RADIATION.

The radio-frequency Paul trap uses inhomogeneous radio-frequency electric fields to confine particles, forcing them to oscillate rapidly in the alternating field (**Fig. 2**). If the amplitude of oscillation (micromotion) is small compared to the trap dimensions, the trap may be thought of as increasing the (kinetic) energy of the particle in a manner that is a function of the particle position. The particle moves to the position of minimal energy and is therefore attracted to the center of the trap where the oscillating electric fields are weakest. At the center of the trap, the fields are exactly zero, and a single, cold ion or electron trapped there is essentially at rest with almost no micromotion. This trap has stability limits. If the frequency of oscillation is too low, the amplitude of the micromotion approaches the trap dimensions and the oscillations become unstable. On the other hand, if the frequency is too high, the energy associated with the oscillation approaches zero together with the depth of the pseudopotential created by the position dependence of the kinetic energy. The great advantage of this trap is that a single particle is driven to the region of space where the trapping fields are zero. Single ions or a small, countable number of ions have been stored, laser-cooled, and visualized in these traps. When a large number of ions are stored in this type of trap, the mutually repulsive Coulomb forces tend to push the ions into regions where the micromotion is large. This effect is referred to as rf heating.

The Penning trap, with the same electrode configuration as the Paul trap, uses a combination of static electric (**E**) and magnetic (**B**) fields instead of oscillating electric fields. The electric fields create a harmonic potential along the z axis but are antitrapping in the radial direction. The repulsive radial forces are overcome by a static magnetic field directed along the z axis. This additional field causes the particles to move in the x-y plane with a combination of a circular cyclotron motion due to the magnetic field and an $\mathbf{E} \times \mathbf{B}$ drift magnetron motion about the trap axis. In this trap, tens of thousands of ions can be stored and cooled to millikelvin temperatures. *See* MAGNETRON.

Neutral-particle traps. Uncharged particles such as neutrons or atoms are manipulated by higher-order moments of the charge distribution such as the magnetic or electric dipole moments.

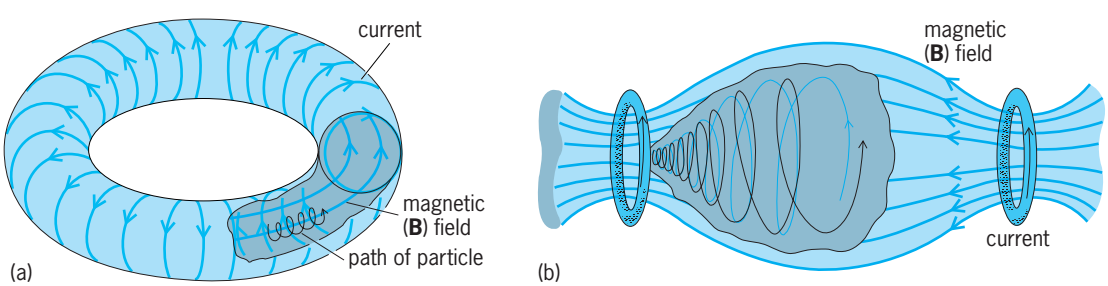

Fig. 1. Magnetic bottles. (*a*) Cutaway view of a tokamak magnetic bottle used to confine hot plasmas. (*b*) Orbit of a charged particle in a magnetic mirror trap as the particle is reflected from the pinched field ends of the trap.

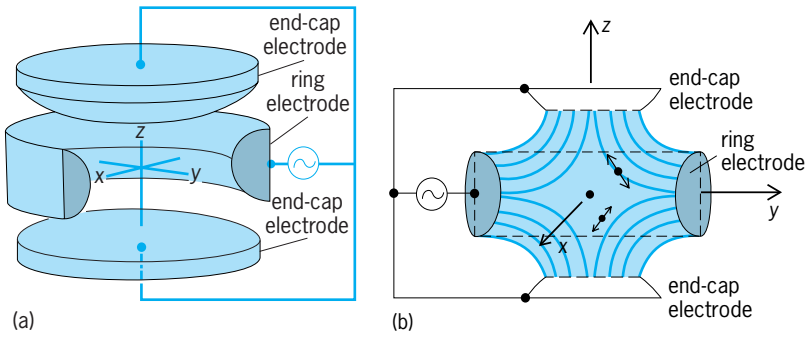

Fig. 2. Radio-frequency Paul trap consisting of two end caps and a ring electrode. (*a*) Cutaway view (*after G. Kamas, ed., Time and Frequency Users's Manual, National Bureau of Standards Technical Note 695, 1977*). (*b*) Cross section, showing the amplitude of the instantaneous oscillations for several locations in the trap.

Magnetic traps of neutral particles use the fact that atoms usually have a magnetic dipole moment on which the gradient of a magnetic field exerts a force. The atom can be in a state whose magnetic energy increases or decreases with the field strength, depending on whether the moment is antiparallel or parallel to the field. A magnetic field cannot be constructed with a local maximum in a current-free region, but a local minimum is possible, allowing particles seeking a weak field to be trapped. In the magnetic field configuration first used to trap an atom (**Fig. 3**), a particle remains in a trapping state as long as its magnetic moment can follow the changing external magnetic field, that is, if the precession frequency of its magnetic dipole moment $\nu = \mu B/h$ (where h is Planck's constant) is high compared to its instantaneous orbital frequency.

Laser traps use the strong electric fields of the laser beam to induce an electric dipole moment on the atom. A laser field tuned below the atomic resonance polarizes the atom in phase with the driving field; the instantaneous dipole moment points in the same direction as the field. Thus the energy of the atom is lowered if it is in a region of high laser intensity. The high-intensity trapping region is formed simply by focusing the beam of a laser. The attraction of the atom to the laser field tuned below resonance resembles the electrostatic attraction of a dust particle to a charged rod. If the laser is tuned above the resonance of the atom, its instantaneous polarization is out of phase with the driving field. In this situation, the atom seeks regions of space with the weakest driving field. For the case in which the atom is surrounded by a laser field with a hole in the middle, this type of trap is analogous to a high-frequency version of the Paul trap. Transparent particles of 0.02 to 10 micrometers suspended in water can also be trapped with a focused laser beam, using the same principles as the neutral atom trap. With these larger particles the trapping forces are helpfully analyzed in terms of the forces on a lens in an inhomogeneous light field. By considering the change in momentum between the incident and refracted-light fields, it can be shown that a particle acts as a converging lens and is drawn to regions of high laser intensity. *See* LASER.

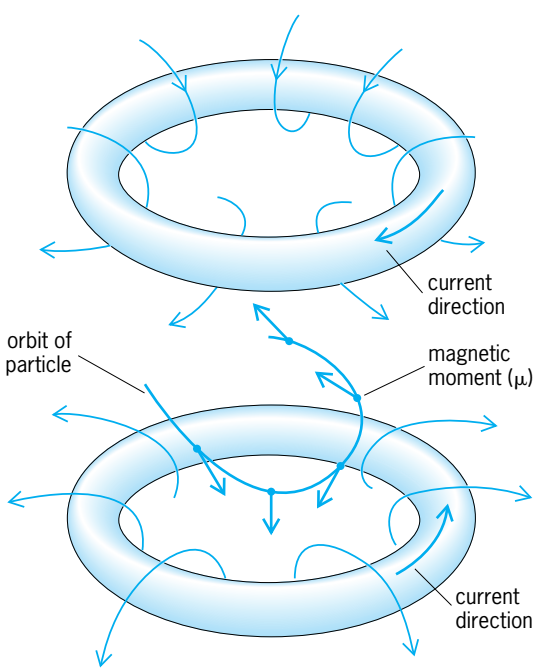

Fig. 3. Spherical quadrupole magnetic trap for slowly moving atoms, showing how the orientation of the atom's magnetic moment remains constant with respect to the local orientation of the magnetic field.

Magnetooptic hybrid traps use, instead of the dipole forces induced by the laser field, the scattering force that arises when an atom absorbs photons with momentum h/λ, where λ is the wavelength. An inhomogeneous magnetic field separates the magnetic substates of an atom in a position-dependent manner. These states interact differently with circularly polarized light. It is possible to arrange a combination of laser beams with proper polarizations to create net scattering forces that drive the atom into the region of zero magnetic field. Such a trap requires much lower laser intensities and weaker magnetic fields. *See* LASER COOLING.

Applications. Particle traps are typically used to hold samples that are too hot or too cold to be permitted to interact with the walls of a containment vessel, or that would react violently with them. Charged-particle traps have been used to confine particles for a very long time. For example, a single positron (named Priscilla by the experimenters) was kept at liquid helium temperatures for 3 months in a radio-frequency trap. Such traps were used to measure the g-factor of the electron and positron to a precision of 4 parts in 10^{15}. This experimental value can be compared with the value calculated by quantum electrodynamics with similar precision. Antiprotons created with high-energy accelerators have been slowed down, trapped, and cooled for over 65 h in a form of Penning trap to be used in precision measurements of the mass of the antiproton and for the creation of antihydrogen. *See* ELECTRON SPIN; QUANTUM ELECTRODYNAMICS.

Long trapping times translate into long measurement times, and potentially more accurate measurement because of the Heisenberg uncertainty principle. Trapped mercury ions form the basis of a

microwave clock with a precision of better than 10^{-14}. Laser-cooled ions in traps may make possible atomic clocks with a frequency uncertainty of less than 10^{-18}.

Laser cooling of ions in these traps to temperatures over a million times colder than that of the vacuum chamber walls has resulted in the Coulomb crystallization of a controlled number of ions. Studies of the quantum interactions of few-body systems under controlled conditions are now possible. *See* ATOMIC CLOCK; UNCERTAINTY PRINCIPLE.

Neutral-particle traps are usually weaker than charged-particle traps. Consequently, only extremely cold atoms can be trapped, and the storage time is limited by collisions between high-temperature background atoms in the vacuum chamber and the trapped atoms. In well-designed traps at liquid-helium temperatures, the vacuum is below 10^{-12} torr (10^{-10} pascal), storage times of several hours have been achieved. Despite the difficulty of trapping neutral particles, there are many instances where their storage and manipulation is feasible. For example, the storage of atoms cooled to a few tens of microkelvins is now possible. Work has been undertaken to cool high-density gases to microkelvin temperatures in order to study their quantum-statistical behavior, with the formation of a Bose condensate or a degenerate Fermi gas as the ultimate goal. Another example of neutral-atom manipulation would be the formation of very slow, intense beams of atoms (including atomic fountains) for use in atomic clocks, atomic-beam interferometers, or sensitive probes in the search for the breakdown of fundamental symmetries such as time reversal invariance. *See* BOSE-EINSTEIN STATISTICS; FERMI-DIRAC STATISTICS; MOLECULAR BEAMS; NEUTRON; SYMMETRY LAWS (PHYSICS); TIME REVERSAL INVARIANCE.

The trapping of particles in the micrometer and submicrometer range has enhanced the possibility of manipulating them. For example, the single-focused-beam laser trap can serve as "laser tweezers" to hold and move microscopic particles such as bacteria and viruses under an optical microscope without apparent damage to the organism. Yeast cells held in this way have reproduced in the trap through 2.5 life cycles. These laser tweezers possess the unique ability to reach inside a living cell and manipulate the contents without puncturing the cell membrane. For example, organelles inside a protozoan cell have been moved about, making it possible to measure the viscosity and elastic properties of the cytoplasm in the region of the trapped organelles. In another application, submicrometer spheres have been attached to the ends of a single deoxyribonucleic acid (DNA) molecule. By using a laser tweezer to grab the sphere, the DNA strand can be stretched out into a straight-line filament for examination. *See* CELL ORGANIZATION; DEOXYRIBONUCLEIC ACID (DNA).

Steven Chu

Bibliography. E. Arimondo, W. D. Phillips, and F. Strumia (eds.), *Laser Manipulation of Atoms and Ions*, 1992; A. Bárány et al. (eds), Proceedings of Workshop and Symposium on the Physics of Low-Energy Stored and Trapped Particles, *Phys. Scripta*, vol. T22, 1988; S. Chu, Laser manipulation of atoms and particles, *Science*, 263:861, 1991; S. Chu, Laser trapping of neutral particles, *Sci. Amer.*, vol. 266, no. 2, February 1992; S. Chu and C. Wieman (eds.), Special issue on laser cooling, *J. Opt. Soc. Amer.*, B6:2058, 1989; W. Itano, J. Bergquist, and D. Wineland, Laser spectroscopy of trapped atomic ions, *Science*, 237:612–617, 1987.

Particulates

Solids or liquids in a subdivided state. Because of this subdivision, particulates exhibit special characteristics which are negligible in the bulk material. Normally, particulates will exist only in the presence of another continuous phase, which may influence the properties of the particulates. A particulate may comprise several phases. **Table 1** categorizes particulate systems and relates them to commonly recognized designations. Fine-particle technology deals with particulate systems in which the particulate phase is subject to change or motion. Particulate dispersions in solids have limited and specialized properties and are conventionally treated in disciplines other than fine-particle technology. *See* ALLOY; GEL.

The universe is made up of particles, ranging in size from the huge masses in outer space—such as galaxies, stars, and planets—to the known minute building blocks of matter—molecules, atoms, protons, neutrons, electrons, neutrinos, and so on. Fine particle technology is concerned with those particles which are tangible to human senses, yet small compared to the human environment—particles that are larger than molecules but smaller than gravel. Fine particles are in abundance in nature (as in rain, soil, sand, minerals, dust, pollen, bacteria, and viruses) and in industry (as in paint pigments, insecticides, powdered milk, soap, powder, cosmetics,

TABLE 1. Types of particulate systems

System			Hydrosol	Aerosol	Powder
Continuous phase		Solid	Liquid	Gas	None (or gas)
Dispersed or particulate phase	Gas	Sponge	Foam	—	—
	Liquid	Gel	Emulsion	Mist Spray Fog Rain	—
	Solid	Alloy	Slurry Suspension	Fume Dust Snow Hail	Single phase Multi-phase (ores, flour)

and inks). Particulates are involved in such undesirable forms as fumes, fly ash, dust, and smog and in military strategy in the form of signal flares, biological and chemical warfare, explosives, and rocket fuels.

Functional attributes. The special properties that particulates possess or accentuate over bulk matter include subdivision, surface area, critical size, curvature effects, interaction with radiation, and mobility.

Bulk materials are subdivided also to increase surface area for purposes of achieving chemical reactions and increased mass or heat transfer, or of altering apparent characteristics (as in coating or encapsulation).

Critical sizes exist or must be met for such phenomena as nucleation, separation, and liberation or beneficiation (including ore dressing and protein shift, a change in the protein content of a food ingredient as the result of mechanical fractionation). Special properties are also exhibited by particles when their size approaches that of a single domain. *See* FOOD ENGINEERING; NUCLEATION; ORE DRESSING.

The curvature of a particle surface affects some of the properties of the material composing the particle, such as surface tension and vapor pressure. Curvature also influences the amount of charge that a particle may acquire in an ionized field. *See* ELECTRIC CHARGE; SURFACE TENSION; VAPOR PRESSURE.

Interactions of particles with electromagnetic radiation (such as light and infrared radiation) include absorption, transmission, reflection, and scattering. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; REFLECTION OF ELECTROMAGNETIC RADIATION; SCATTERING OF ELECTROMAGNETIC RADIATION.

Mobility of particulates, which is important in pneumatic and hydraulic conveying, involves such considerations as particle inertia, diffusion, and drag.

Processing and uses. There are many cases where a particulate is either a necessary or an inadvertent intermediary in an operation. Areas that may be involved in processing particulates may be classified as follows:

- Size reduction
  - Mechanical (starting with bulk material)
    - Grinding
    - Atomization
    - Emulsification
    - Physicochemical (conversion to molecular dispersion)
      - Phase change (spray drying, condensation)
      - Chemical reaction
- Size enlargement (agglomeration, compaction)
  - Pelletizing
  - Briqueting
  - Nodulizing
  - Sintering
- Separation or classification
  - Ore beneficiation
  - Protein shift
- Deposition (collection, removal)
- Coating or encapsulation
- Handling
  - Powders
  - Gas suspensions (pneumatic conveying)
  - Liquid suspensions (non-newtonian fluids)

See ATOMIZATION; CONVECTION (HEAT); DRYING; SINTERING.

End products in which particulate properties themselves are utilized include the following:

- Mass or heat transfer agents (used in fluidized beds)
- Recording (memory) agents
  - Electrostatic printing powders and toners (for optical images)
  - Magnetic recording media (for electronic images)
- Coating agents (paints)
- Nucleating agents
- Control agents (for servomechanisms)
  - Electric fluids
  - Magnetic fluids
- Charge carriers (used in propulsion, magnetohydrodynamics)
- Chemical reagents
  - Pesticides, fertilizers
  - Fuels (coal, oil)
  - Soap powders
  - Drugs
  - Explosives
- Food products

See EXPLOSIVE; FERTILIZER; FLUIDIZED-BED COMBUSTION; FOOD; INK; ION PROPULSION; MAGNETIC RECORDING; MAGNETOHYDRODYNAMIC POWER GENERATOR; PAINT; PESTICIDE; PETROLEUM; SOAP.

Characterization. The processing or use of particulates will usually involve one or more of the following characteristics:

- Physical
  - Size (and size distribution)
  - Shape
  - Density
  - Packing or concentration
- Chemical
  - Composition
  - Surface character
- Physiochemical (including adhesive, cohesive)
- Mechanical or dynamic
  - Inertial
  - Diffusional
  - Fluid drag
    - Dilute suspensions (Stokes' law, Cunningham factor, drag coefficient)
    - Concentrated suspensions (hindered settling, rheology, fluidization)
- Optical (scattering, transmission, absorption)
  - Refractive index (including absorption)
  - Reflectivity
- Electrical
  - Conductivity
  - Charge
- Magnetic

Thermal
  Insulation (conductivity, absorptivity)
  Thermophoresis

See CONDUCTION (ELECTRICITY); CONDUCTION (HEAT); REFRACTION OF WAVES.

Many of the characteristics of particulates are influenced to a major extent by the particle size. For this reason, particle size has been accepted as a primary basis for characterizing particulates. However, with anything but homogeneous spherical particles, the measured "particle size" is not necessarily a unique property of the particulate but may be influenced by the technique used. Consequently, it is important that the techniques used for size analysis be closely allied to the utilization phenomenon for which the analysis is desired.

Particle size. Size is generally expressed in terms of some representative, average, or effective dimension of the particle. The most widely used unit of particle size is the micrometer (μm), equal to 0.001 mm (1/25,400 in.). Another common method is to designate the screen mesh that has an aperture corresponding to the particle size. The screen mesh normally refers to the number of screen openings per unit length or area; several screen standards are in general use, the two most common in the United States being the U.S. Standard and the Tyler Standard Screen Scales.

Particulate systems are often complex (**Fig. 1**). Primary particulates may exist as loosely adhering (as by van der Waals forces) particles called flocs or as strongly adhering (as by chemical bonds) particulates called agglomerates. Primary particles are those whose size can only be reduced by the forceful shearing of crystalline or molecular bonds. *See* CHEMICAL BONDING; INTERMOLECULAR FORCES.

Mechanical dispersoids are formed by comminution, decrepitation, or disintegration of larger masses of material, as by grinding of solids or spraying of liquids, and usually involve a wide distribution of particle sizes. Condensed dispersoids are formed by condensation of the vapor phase (or crystallization of a solution) or as the product of a liquid- or vapor-phase reaction; these are usually very fine and often relatively uniform in size. Condensed dispersoids and very fine mechanical dispersoids generally tend to flocculate or agglomerate to form loose clusters of larger particle size.

Because of the difficulty of distinguishing between the various stages of flocculation, size analysis methods usually aim at measuring discrete particles, because this state is more readily defined and reproduced. Some methods, because of their nature, will measure the size of primary particles as discrete particles, while other methods will measure the agglomerate.

Unfortunately, there is usually some degree of flocculation despite the relatively small adhesive forces between particles. Often this is due to the high concentrations of particles (close proximity and high probability of contact) in practical operations. In those cases, the particles may be flocculated in a dynamic sense. Particles in a floc may be continuously dislodged and replaced by other particles or flocs with the state of dispersion being fixed only in a statistical sense.

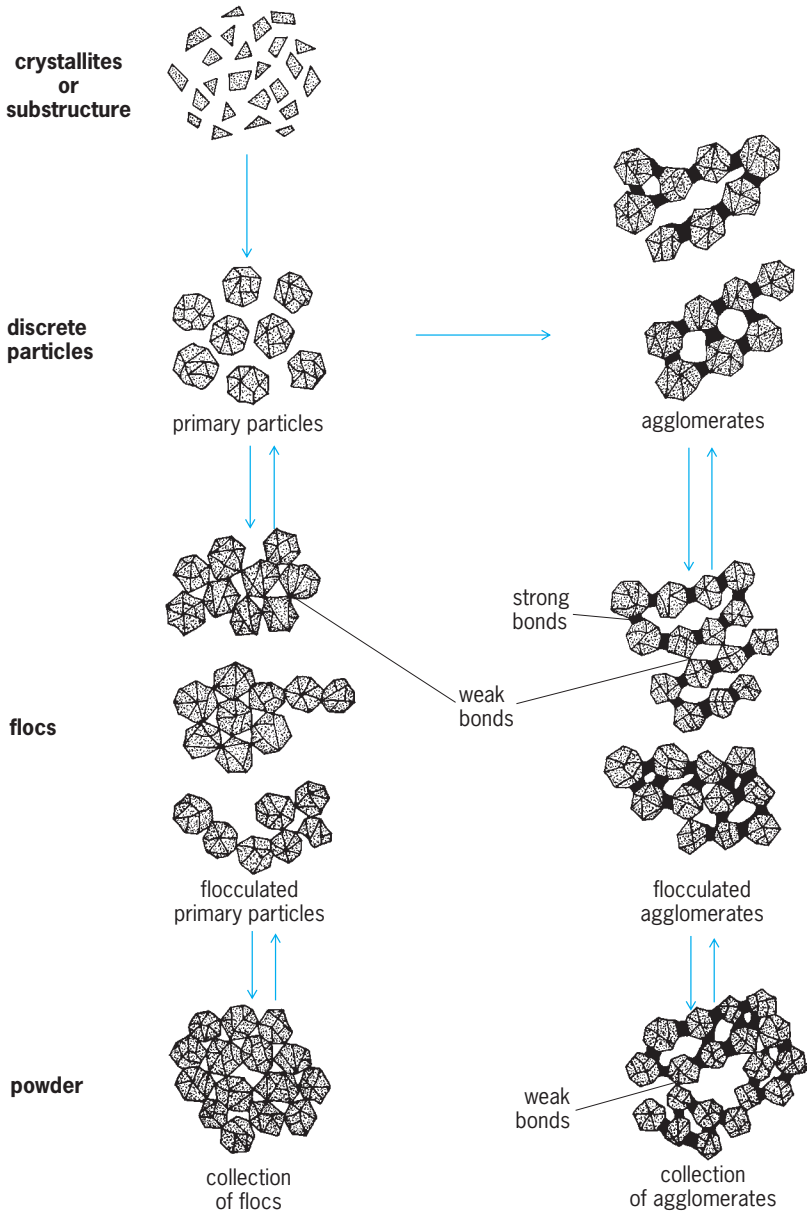

Fig. 1. States of dispersion of particulates. The double arrows imply reversibility with application of light shearing forces.

Differences in state of dispersion of particles are usually of major significance when particle size is less than 10 μm in diameter. With particles larger than 10 μm the magnitude of the surface forces holding particles in flocculated clusters decreases rapidly relative to other forces acting on the particles (such as inertia and gravity), and the particles behave more nearly like discrete particles in a statistical sense.

Methods of representing size. Most real systems are composed of a range of particle sizes. The two common general methods for representing size distribution graphically are shown in **Fig. 2**. The frequency distribution gives the fraction of particles $d\phi$ (on whatever

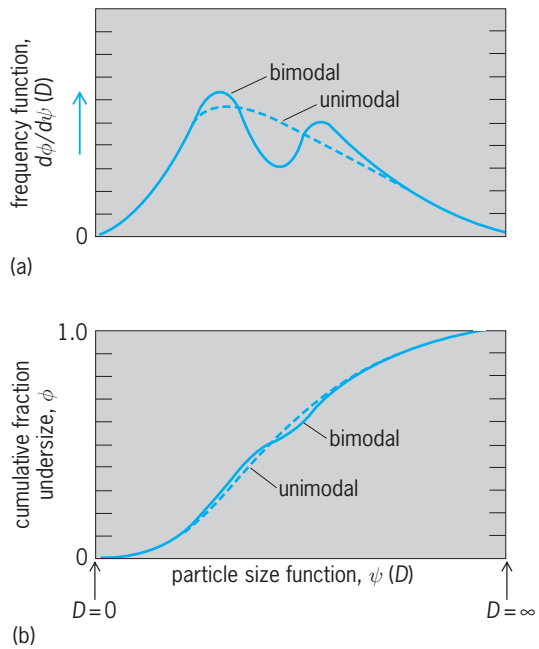

Fig. 2. Methods for representing size distribution. (*a*) Frequency distribution. (*b*) Cumulative distribution.

basis desired) that lie in a given narrow size range dD as a function of the average size of the range (or of some function of the average size). It is essential that the size increment be chosen in a systematic and specified manner (as implied by the denominator of the ordinate of Fig. 2*a*). The shape of the curve will depend on how the increment is chosen, and it will usually take the form of a skewed probability curve. However, it may also have multiple peaks, in which case the distribution is termed multimodal. If an arithmetic increment of size is taken—that is, $d\psi(D) = dD$—the result is commonly referred to as a percent-per-micrometer plot. While the function of diameter in the ordinate need not be the same as that in the abscissa, the area under the curve will be unity if both are the same.

The lower curve of Fig. 2 is a cumulative distribution, the integral of the frequency curve. It gives the fraction ϕ of the particles that are smaller or larger than a given size D. Being an integral, the method of choosing size increments need not be selected or specified. *See* INTEGRATION; STATISTICS.

Bases for representation. As discussed above, size distribution may be represented by the fraction of particles in a specific size range, or over or under a certain size. The fraction, however, must be specified on some basis, which may be number, surface or mass. It is theoretically possible to convert from one basis to another. If the shape and density of the particles are statistically uniform (do not vary with size), Eq. (1)

$$\phi_x = \frac{\sum_0^D D^{x-y}\Delta\phi_y}{\sum_0^\infty D^{x-y}\Delta\phi_y} = \frac{\int_0^D D^{x-y}\,d\phi_y}{\int_0^\infty D^{x-y}\,d\phi_y} \qquad (1)$$

may be used to convert a distribution from one basis (ϕ_y) to another (ϕ_x). Here ϕ_x is the cumulative fraction smaller than size D on a basis corresponding to x; ϕ_y is the cumulative fraction smaller than size D on a basis corresponding to y; D is particle diameter; and x and y are the characteristic subscripts or exponents used to define a specific basis. **Table 2** gives the value of x or y corresponding to various bases for expressing size distribution.

The measurement basis used may markedly influence the precision of the desired answer. A few large particles overlooked in a count analysis will not significantly affect the count distribution, but may result in serious errors in the corresponding calculated mass or volume distribution. Similarly, the count distribution calculated from a mass measurement may be grossly in error at the fines end.

If shape varies with size, conversion between bases requires knowledge of the shape factor and how it varies with size. This knowledge is usually not available. Variation in particle composition introduces additional complications.

Analytical relationships. Various analytical relationships have been proposed for representing size distribution, notably the Rosin-Rammler, the Roller, the Nukiyama-Tanasawa, and the log-probability distributions. Advantages have been claimed for each, usually on the basis that one or the other gives a better fit of data in a particular instance. In addition, claims based on various simplifying assumptions have been made for theoretical justification, and have been made especially in the case of the log-probability distribution.

Actually, it is easier to find exceptions to than agreements with each. There are times, however, when analytical convenience may justify one. The log-probability relationship is particularly useful in this respect. The log-probability frequency distribution is given by Eq. (2), or, in the normalized form by Eq. (3), where ϕ_x is the cumulative fraction smaller

$$\frac{d\phi_x}{d\ \ln D} = \left[\frac{1}{(\sqrt{2\pi})\ \ln\sigma}\right] \times \exp\left\{-\left[\frac{\ln D - \ln D_{mx}}{(\sqrt{2\pi})\ln\sigma}\right]^2\right\} \qquad (2)$$

$$\frac{d\phi_x}{d\ \ln(D/D_{mx})} = \left[\frac{1}{(\sqrt{2\pi})\ln\sigma}\right] \times \exp\left\{-\left[\frac{\ln(D/D_{mx})}{(\sqrt{2\pi})\ln\sigma}\right]^2\right\} \qquad (3)$$

than size D on a basis corresponding to x; D is the

TABLE 2. Bases for expressing size distribution

Type of distribution basis	Value of x or y
Number or count	0
Linear	1
Surface	2
Volume (or mass)	3

particle diameter; σ is the standard geometric deviation; and D_{mx} is median diameter on the basis of a property corresponding to x.

The values of σ and D_{mx} are characteristic constants for a given size distribution. If a material follows a log-probability distribution on one basis (x), it also does on any other basis (y)—with the same value of the standard geometric deviation (σ) but a different value of median size (D_{my}) corresponding to the new basis (y). This is a unique property of the log-probability distribution. Probability graph papers are commercially available (**Fig. 3**) and are convenient for graphical representation of size distribution data even if the distribution does not follow a probability relationship. In addition, the assumption of a log-probability distribution as an approximation permits simple conversion from one basis of representing size distribution, mean size, or median size to another (as described below).

The solid line of Fig. 3 shows a typical size distribution. If the size distribution follows a log-probability relationship, it will be a straight line on this type of plot, as shown by the two broken lines, and can be completely characterized by two numbers, (1) a median diameter, corresponding to the 50% cumulative size, and (2) a standard geometric deviation, a number equal to or greater than unity, given by the ratio of the 84.13% to the 50% size (or of the 50% to the 15.87% size).

The median diameter is a measure of the general size level, whereas the standard geometric deviation is a measure of the degree of uniformity. A completely uniform material (all particles the same size) would show up as a horizontal line in Fig. 3 and have a standard geometric deviation of 1.0. A completely heterogeneous material would be represented by a vertical line, which would have a standard geometric deviation of infinity.

A more realistic but more complex distribution is the upper-limit log-probability distribution proposed by R. A. Mugele and H. D. Evans.

Average particle diameter. For many purposes, it is adequate to represent a particulate material by an average size. The type of average size depends on the nature of the application and can often be established quantitatively by a consideration of the physics involved.

The general size level represented by a particulate material can be expressed in terms of the median size given by Eq. (4), where D_{mx} is the median di-

$$\int_0^{D_{mx}} d\phi_x = \int_{D_{mx}}^1 d\phi_x = {}^1/_2 \qquad (4)$$

ameter based on a property corresponding to x, and ϕ_x is the cumulative fraction of all particles (based on a property corresponding to x) smaller than size D.

Thus, in terms of a specific property x, half the material is finer than the median size D_{mx}, the other half coarser. In terms of mass, a mass median diameter means that half the mass of a powder consists of particles coarser than the mass median diameter, and

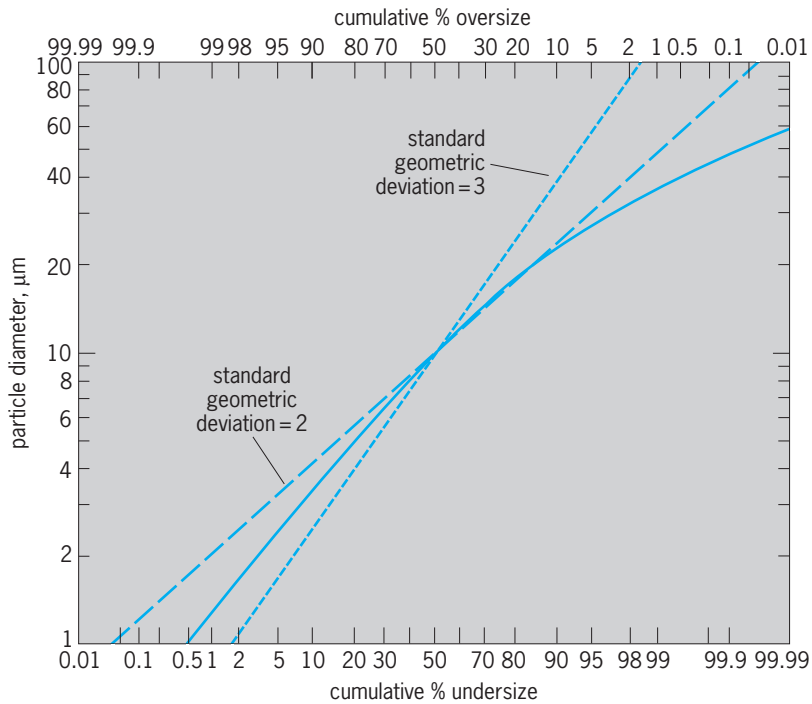

Fig. 3. Probability graph for showing typical size distributions.

half finer. Therefore, the median diameter is given by the 50% point on the cumulative distribution curve obtained in terms of the property desired.

If a material obeys a log-probability relationship, the various median diameters are related to each other by the relationship in Eq. (5), where D_{mx} is

$$D_{mx} = D_{my} \exp\{(x - y)\ \ln^2 \sigma\} \qquad (5)$$

the median diameter on the basis of a property corresponding to x; D_{my} is the median diameter on the basis of property corresponding to y; x and y are the characteristic subscripts or exponents used to define a specific basis as given in Table 2; and σ is the standard geometric deviation. Thus, D_{m0} and D_{m3} may be defined, respectively, as number and volume (or mass) median diameters.

However, while median diameters are readily obtained from size distribution measurements, the effective role of particles in some specific phenomenon can usually be expressed in terms of a mean diameter, defined by Eq. (6), where $\overline{D}_{qp}$ is the mean

$$\overline{D}_{qp} = \left[\frac{\int_0^1 D^q d\phi_n}{\int_0^1 D^p d\phi_n}\right]^{[1/(q-p)]} \qquad (6)$$

diameter; D is any particle diameter; ϕ_n is the fraction of all particles smaller than size D on a number basis; and q and p are the characteristic subscripts and exponents used to define mean diameter.

If the material follows a log-probability distribution, Eq. (6) reduces to Eq. (7), where q, p, and x

$$\overline{D}_{qp} = D_{mx} \exp\{[(q + p - 2x)/2]\ \ln^2 \sigma\} \qquad (7)$$

are the same characteristic subscripts or exponents previously defined.

With a uniform material ($\sigma = 1$), all the ratios would be unity, and all mean and median diameters would be the same. As the material becomes less uniform (as σ increases), the differences between the various means and medians become greater. For a log-probability distribution, the geometric mean and number median diameter are identical. This would not be true for another type of distribution.

Precision and significance of analysis. The precision and significance of size analyses are governed by (1) sampling of particulate source, (2) preparation of sample analysis, (3) type size-discrimination property employed, (4) type of measurement property employed, (5) precision of various analytical procedures, and (6) analyses and interpretation of data.

Sampling involves at least two steps: sampling in the field or plant, and selection of small aliquots for individual analysis. Sampling becomes more difficult as the particle size becomes larger and the size distribution broader because of particle segregation tendencies. Sample preparation can be a determining factor in the significance of the size analysis, especially in the degree of particle dispersion achieved. With liquids, dispersion is improved by the use of dilute suspensions, the addition of dispersing agents, and the use of viscous liquids. *See* CHEMOMETRICS.

Analysis methods. There are two classes of size analysis equipment. In one, the class of counters, each particle is observed individually and its discrimination property measured directly. In the other, particles are measured in bulk (in terms of total number, surface area, or mass) after being subjected to a discriminating action. There are hundreds of size analysis methods, each employing some combination of size discrimination and measurement technique. **Table 3** summarizes the general types of discriminating properties available together with associated variations in measurement techniques. Most techniques do not measure size but some other property that is markedly, but not usually solely, influenced by size. It has become conventional to report the results in terms of a size by some assumed or empirical relationship even though the size does not represent any real dimension.

In some operations (such as chemical reactions, catalysis, comminution, and certain optical effects), a knowledge of the total surface area of the particulate matter may be sufficient. The more common methods for measuring total surface rather than size distribution are the permeability methods (as exemplified by the Blaine and the Fischer Sub-Sieve Sizer instruments) in which the resistance to flow of air through a mass of powder of known porosity is measured. The specific surface of the powder is determined from a semi-empirical equation. In adsorption methods, exemplified by the Brunauer-Emmett-Teller (BET) method, the amount of gas (or liquid, dye, or radioactive material) adsorbed on particles is used as a measure of surface. While

TABLE 3. Types of size analysis methods

Size-discriminating property		
Type	Mechanism	Measurement technique
Geometric	Physical barrier	Sieving (wet or dry) Ultrafiltration
Mechanical or dynamic (in fluids)	Inertia	Impaction on surface Pressure pulse (sonic)
	Terminal settling velocity in force field (gravity, centrifugal, electrical)	Elutriation (liquid or gas; single or series fractionation) Sedimentation (liquid or gas; layer or suspension; differential or integral) Decantation Photosedimentation
	Diffusion	Particle displacement Particle deposition
Optical	Imaging	Light microscopy Ultramicroscopy (mean size only) Electron microscopy
	Spectral transmission	Extinction
	Scattering	Single particle count (right angle, angular, forward, polarization) Macroscopic (mean size only)
	Diffraction	Light, x-ray, laser
Electrical	Resistance	Current flow pulses
	Capacitance	Voltage pulses
	Charge	Triboelectric Induction Corona
Thermal	Particle deposition	Migration in thermal gradient
Physico-chemical	Condensation	Nuclei growth with controlled supersaturation

permeability methods tend to measure the outer exposed surface of primary particles, an adsorption method may include surface associated with internal fissures. Turbidimetric methods and methods involving measurement of energy changes (such as heat of solution, heat of wetting, and coercive force) are also used to measure total surface. *See* ADSORPTION.

Particle concentration. The concentration of particles in a dispersoid is often expressed in terms of grains per cubic foot of gas, where 7000 grains = 1 lb. In air-pollution work, concentrations are sometimes expressed as grams per cubic meter (1 g/m^3 = 0.437 grain/ft^3). Process dust concentrations normally range from 0.01 to 100 grain/ft^3 (0.02 to 200 g/m^3) with 1–10 grains/ft^3 (2–20 g/m^3) being common in process ventilation work. In air conditioning applications concentrations of particulate matter generally range from 0.1 to 10 grains/1000 ft^3 (0.2 to 20 g/m^3). In operations such as pneumatic conveying, the concentrations range from 0.1 to 50 lb solid/lb gas or 0.1 to 50 kg solid/kg gas (about 50–20,000 grains/ft^3 or 100–45,000 g/m^3). In fluidized-solid systems concentrations frequently exceed 1000 lb solid/lb gas or 1000 kg solid/kg gas.

Particulate mechanics. A particle moving in a fluid encounters a frictional resistance given by Eq. (8),

$$F_r = C_D A_p \rho \frac{u_r^2}{2} \tag{8}$$

where F_r is the resisting force in newtons; C_p is the dimensionless drag coefficient; A_p is the area of the particle projected in a plane normal to the direction of motion in m^2; ρ is the fluid density in kg/m^3; and u_r is the velocity of the particle relative to the fluid in m/s. The drag coefficient C_D is a function of the particle shape, the proximity of bounding surfaces, and the Reynolds number N_{Re} ($= D_p u_r \rho/\mu$, where D_p is the particle diameter or other representative dimension in m; and μ is the fluid viscosity in Pa·s). Drag coefficient data are available for a wide variety of particle shapes. *See* REYNOLDS NUMBER.

If a particle suspended in a fluid is acted upon by a force, it will accelerate to a terminal velocity at which the resisting force due to fluid friction just balances the applied force. If a particle falls under the action of gravity, this velocity is known as the terminal gravitational settling velocity and is given by Eq. (9), where u_t is the settling velocity in m/s; g_L

$$u_t = \sqrt{\frac{2 g_L M_p (\rho_p - \rho)}{\rho \rho_p A_p C_D}} \tag{9}$$

is the acceleration due to gravity, 9.807 m/s^2; M_p is the particle mass in kilograms; ρ_p is particle density in kg/m^3; and other terms are as previously defined.

For spherical particles present in dilute concentration, C_D is approximately constant at 0.44 for $N_{\mathrm{Re}} > 500$. For $N_{\mathrm{Re}} < 1$, C_D is equal to $24/N_{\mathrm{Re}}$, and the resisting force is given by Eq. (10), which is known

$$F_r = 3\pi \mu u_r D_p \tag{10}$$

as Stokes' law. Corresponding to this the terminal gravitational settling velocity is given by Eq. (11),

$$u_t = \frac{g_L D_p^2 (\rho_p - \rho)}{18\mu} \tag{11}$$

commonly known as Stokes' law of settling, which is usually applicable for particles smaller than 50 μm in diameter.

When the particle size approaches the magnitude of the mean free path of the molecules of the suspending fluid, the frictional resistance to motion becomes less than (and the settling velocity becomes greater than) that indicated by the above equations. To correct for this molecular slip flow effect, the resistance to motion calculated from the above equations must be divided by (or the settling factor multiplied by) a factor k_m, commonly known as the Stokes-Cunningham correction factor, given by Eq. (12),

$$k_m = 1 + k_{me} \frac{\lambda}{D_p} \tag{12}$$

where k_{me} is a dimensionless constant and λ is the mean free path of the fluid molecules in m. The value k_{me} has been shown experimentally to lie between 1.3 and 2.3 for different gases, particle sizes, and materials. Based on the data of R. A. Millikan, Eq. (13) is

$$k_{me} = 1.644 + 0.552 e^{-(0.656 D_p/\lambda)} \tag{13}$$

obtained, where λ is based on simple kinetic theory ($\lambda = 3\mu p \bar{v}$), and the mean molecular speed $\bar{v}$ is given in ft/s or m/s by Eq. (14), where R is the gas constant,

$$\bar{v} = \sqrt{\frac{8RT}{\pi M}} \tag{14}$$

8.31×10^3 J/(°C) (kg-mole); T is the absolute temperature in °R or K; and M is the molecular weight of the gas in kg/kg-mole. For particles in gases at atmospheric pressure, the Stokes-Cunningham factor becomes significant for particles smaller than 1 μm. At high altitude or low pressure, this factor can become extremely large.

Particles suspended in a fluid partake of the molecular motion of the suspending fluid and hence acquire diffusional characteristics analogous to those of the fluid molecules. This random zigzag motion of the particles, commonly known as brownian motion, is obvious under the microscope for particles smaller than 1 μm. The average displacement of a particle in time t is given by Eq. (15), where Δs is

$$\Delta s = \sqrt{\frac{4 D_v t}{\pi}} \tag{15}$$

the average linear displacement along a given axis, regardless of sign, in m; D_v is the diffusion coefficient for the particle in m^2/s; and t is the time in s. The Einstein equation for the diffusion coefficient D_v for spherical particles is Eq. (16), where N_A is Avogadro's

$$D_v = \frac{k_m RT}{3\pi \mu N_A D_p} \tag{16}$$

number, 6.023×10^{26} molecules/kg-mole, and other terms are as previously defined. This diffusion coefficient may also be used to estimate the migration rate of particles when a concentration gradient exists. *See* AIR FILTER; BROWNIAN MOVEMENT; COLLOID; DIFFUSION; DUST AND MIST COLLECTION; EMULSION; FOAM. Charles L. Lapple

Bibliography. T. Ariman and T. N. Veziroglu (eds.), *Particulate and Multiphase Processes*, 3 vols., 1987; J. K. Beddow (ed.), *Particulate Systems: Technology and Fundamentals*, 1983; R. Clift, *Particulate Technology*, 1991; N. Fuchs, *Mechanics of Aerosols*, 1989; H. E. Hesketh, *Fine Particles in Gaseous Media*, 2d ed., 1986; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997.

Paschen-Back effect

An effect on spectral lines obtained when the light source is placed in a very strong magnetic field, first explained by F. Paschen and E. Back in 1921. In such a field the anomalous Zeeman effect, which is obtained with weaker fields, changes over to what is, in a first approximation, the normal Zeeman effect. The term "very strong field" is a relative one, since the field strength required depends on the particular lines being investigated. It must be strong enough to produce a magnetic splitting that is large compared to the separation of the components of the spin-orbit multiplet. *See* ATOMIC STRUCTURE AND SPECTRA; ZEEMAN EFFECT.

The **illustration** shows, as an example, the Paschen-Back effect of the red line $2s^2S$–$2p^2P$ of lithium. The natural separation of this doublet is very small, only 0.0175 nanometer, so that in the field of 44,200 oersteds (3.52×10^6 amperes/m) for which the diagram is drawn the normal Zeeman splitting of 0.0929 nm greatly exceeds it. The Paschen-Back effect is therefore practically complete, and the resulting pattern is nearly a normal triplet. There is still, however, a residual splitting, which theory gives as two-thirds of the field-free separation. The weak field patterns shown in the figure require a field of only about 1800 oersteds (1.4×10^5 A/m) and correspond to the anomalous patterns obtained in the Zeeman effect.

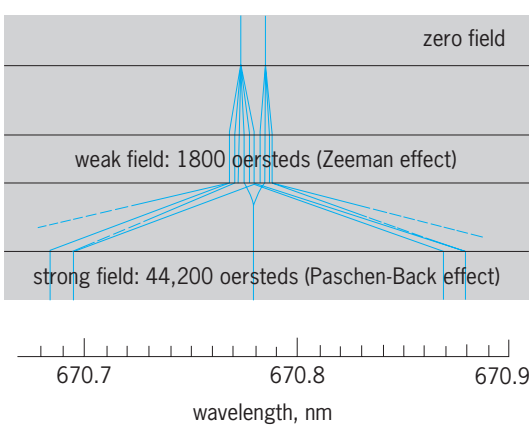

Zeeman and Paschen-Back effects of red lithium doublet, whose natural separation is 0.0175 nm.1 oersted = 79.5 A/m.

Theory explains the transformation from the Zeeman to the Paschen-Back effect as due to the uncoupling of the orbital and spin vectors L and S by the magnetic field. Whereas in a weak field these vectors are coupled magnetically to form a resultant J, a sufficiently strong field causes them to precess independently about the field direction. The correlation of the energy levels between weak and strong fields follows from the principle that the magnetic field is incapable of changing the angular momentum component along the field; that is, the magnetic quantum number M of a level remains constant for all field strengths. A further rule is that two levels having the same value of M (which in the strong field equals $M_L + M_S$) do not cross each other. These rules lead to the correlation of lines shown in the figure. Certain lines fade out during the transition, as is indicated by the change from a solid to a dotted line, and these are the ones for which the direction of polarization would be altered. F. A. Jenkins; W. W. Watson

Passeriformes

The largest and most diverse order of birds, which is found worldwide, including the northernmost tip of Antarctica, most oceanic islands, and all terrestrial habitats. Over 5500 species of passerine birds (of which about 4400 species are Oscines, the true songbirds; see **illustration**) are known, which is just over half of the known species of Recent birds. The most closely related orders may be other land birds, such as the Coraciiformes and the Piciformes. *See* CORACIIFORMES; PICIFORMES.

Classification. The Passeriformes are divided into four suborders as outlined below.

Order Passeriformes
- Suborder Eurylaimi
  - Family: Eurylaimidae (broadbills; 14 species; Old World tropics)
    - Philepittidae (false sunbirds; 4 species; Madagascar)
    - Pittidae (pittas; 24 species; tropical Asia to Australia)
    - Acanthisittidae (New Zealand wrens; 4 species; New Zealand)
- Suborder Furnarii
  - Family: Dendrocolaptidae (woodcreepers; 52 species; Neotropics)
    - Furnariidae (ovenbirds; 218 species; Neotropics)
    - Formicariidae (antbirds; 239 species; Neotropics)
    - Rhinocryptidae (tapaculos; 30 species; Neotropics)

Northern mockingbird (*Mimus polyglottos*). (*Photo by George W. Robinson; © California Academy of Sciences*)

Suborder Tyranni
Family: Tyrannidae (tyrant flycatchers; 377 species; New World, mainly Neotropics)
Pipridae (manakins; 52 species; Neotropics)
Cotingidae (cotingas; 79 species; Neotropics)
Oxyruncidae (sharpbill; 1 species; Neotropics)
Phytotomidae (plant cutters; 3 species; Neotropics)
Suborder Oscines
Family: Menuridae (lyrebirds; 2 species; Australia)
Atrichornithidae (scrubbirds; 2 species; Australia)
Palaeospizidae (fossil; Oligocene, North America)
Alaudidae (larks; 78 species; worldwide, mainly Old World to Australia)
Hirundinidae (swallows; 80 species; worldwide)
Motacillidae (pipits, wagtails; 54 species; worldwide, mainly Old World)
Campephagidae (cuckoo shrikes; 70 species; Old World to Australia)
Palacoscinidae (fossil; Miocene of North America)
Pycnonotidae (bulbuls; 123 species; Old World)
Irenidae (fairy bluebirds, leaf-birds; 14 species; Asia)
Laniidae (shrikes; 65 species; worldwide, mainly Old World)
Prionopidae (helmet shrikes; 9 species; Africa)
Vangidae (vanga shrikes; 13 species; Madagascar)
Bombycillidae (waxwings; 8 species; Paleoarctic, not Australia and Africa)
Dulidae (palm chat; 1 species; Hispaniola)
Prunellidae (hedge sparrows; 12 species; Paleoarctic)
Mimidae (mockingbirds, thrashers; 31 species; New World)
Cinclidae (dippers; 5 species; worldwide, not Australia and Africa)
Turdidae (thrushes; 309 species; worldwide)
Timaliidae (babblers; 276 species; Old World)
Troglodytidae (wrens; 60 species; New World, one species in the Paleoarctic)
Sylviidae (Old World warblers; 349 species; Old World)
Muscicapidae (Old World flycatchers; 113 species; Old World)
Monarchidae (monarch flycatchers; 93 species; Old World to Australia)
Maluridae (Australian wrens; 26 species; Australasian)
Acanthizidae (thornbills; 67 species; Australasian)
Ephthianuridae (Australian chats; 5 species; Australia)
Orthonychidae (rail babblers; 19 species; Australasian)
Rhipiduridae (fantails; 40 species; Australasian)
Petroicidae (Australian robins; 40 species; Australasian)
Pachycephalidae (whistlers; 46 species; Australasian)
Aegithalidae (long-tailed tits; 7 species; Paleoarctic)
Remizidae (bearded tits; 10 species; Paleoarctic)
Paridae (chickadees; 47 species; Holarctic to Africa)
Sittidae (nuthatches; 23 species; Holarctic) Certhiidae (creepers; 6 species; Holarctic) Rhabdornithidae (Philippine creeper; 1 species; Philippines)
Neosittidae (tree runners; 3 species; Australia)
Climacteridae (Australian creepers; 6 species; Australasia)
Dicaeidae (flowerpeckers; 58 species; Old World tropics to Australia)

Family: Nectariniidae (sunbirds; 117 species; Old World tropics to Australia)
Zosteropidae (white-eyes; 83 species; Old World to New Zealand)
Meliphagidae (honeyeaters; 172 species; Australasia to Hawaii)
Vireonidae (vireos; 43 species; New World)
Emberizidae (buntings; 318 species; worldwide except Australasia)
Catamblyrhynchidae (plush-capped finch; 1 species; South America)
Thraupidae (tanagers; 240 species; New World, mainly Neotropics)
Tersinidae (swallow tanager; 1 species; South America)
Parulidae (wood warblers; 125 species; New World)
Icteridae (orioles, blackbirds, troupials; 95 species; New World)
Fringillidae (finches; 122 species; worldwide except Australasia)
Drepanididae (Hawaiian honey creepers; 24 species; Hawaiian Islands)
Estrildidae (waxbills, grass finches; 127 species; Old World to Australia, mainly tropics)
Ploceidae (weaver finches; 117 species; Old World, mainly Africa)
Passeridae (sparrows; 27 species; Old World except Australasia)
Sturnidae (starlings; 111 species; Old World, mainly tropics)
Oriolidae (orioles; 25 species; Old World, mainly tropics)
Dicruridae (drongos; 20 species, Southeast Asia to Australasia)
Callaeidae (wattle-birds; 3 species; New Zealand)
Corcoracidaae (mud-nest builders; 4 species; Australasia)
Artamidae (wood swallows; 13 species; Australasia)
Cracticidae (butcher-birds; 10 species; Australasia)
Pityriasididae (bristle head; 1 species; Borneo)
Ptilonorhynchidae (bower birds; 18 species; Australasia)
Paradisaeidae (birds of paradise; 42 species; Australasia)
Corvidae (crows, jays; 106 species; worldwide)

The composition of most families within those suborders and their relationships to one another are still much disputed. The Eurylaimi are Old World tropical and subtropical birds and may be an artificial group. The other suborders are well delimited and are well-accepted taxa. The Furnarii and the Tyranni appear to be South American in origin and are still Neotropical in distribution (Central and South America), except for about 30 species of tyrant flycatchers and one species of cotingas which occur in North America. The area of origin of the Oscines is not known, but it may be Old World Gondwanaland after the separation of South America.

Several groups within the suborder Oscines are apparent. The largest and most definite are the New World Nine-primaried Oscines, with nine large primary feathers on each wing (the 10th is rudimentary), which include the families from the Vireonidae to the Fringillidae; however, there is some question about the relationships of the Vireonidae to the other members of the Nine-primaried Oscines. This group had its origins in southern North America and spread into South America, with some groups spreading into the Old World. Another is the Australian radiation, which includes the Dicruridae to the Paradisaeidae. Perhaps yet another diverse Australian group would include the Meliphagidae, Zosteropidae, Dicaeidae, Climacteridae. Pachycephalidae, Petroicidae, Rhipiduridae, Orthonychidae, Ephthianuridae, Acanthizidae, Maluridae, and Monarchidae. The Old World Insect-eaters include the families Prunellidae, Mimidae, Cinclidae, Turdidae, Timaliidae, Troglodytidae, Sylviidae, and Muscicapidae, but all of these families may not be natural and/or closely related to one another. Several smaller groups could be listed, but knowledge of the evolutionary history and relationships within the perching birds is still poor.

Fossil record. Because they are small land birds, the passeriforms have a poor fossil record. It is not even certain whether *Palaeospiza bella* is correctly assigned to the Passeriformes; it may be a member of the Coraciiformes. No definite passeriform fossils have been identified in the Northern Hemisphere that predate the beginning of the Miocene, in spite of a number of other small avian fossils found in earlier deposits. The lack of passeriform fossils from the well-worked Northern Hemisphere has led some paleontologists to postulate that the birds originated in the Southern Hemisphere and invaded the northern regions in the mid-Tertiary. Passeriform fossils are found rarely in the Miocene, become somewhat more common in the Pliocene, and are most abundant in the Quaternary. However, the available fossils provide no useful information on the evolutionary history of the passeriforms.

Characteristics. The perching birds are small to medium-sized, ravens (*Corvus corax*) being the largest. The wings are short to medium in length and vary from rounded to pointed. Features that delimit the Passeriformes from other birds are difficult to find; one may be the presence of a second set of air tubes in the lungs, allowing uptake of oxygen during inhaling, as well as the normal avian method of uptake of oxygen during exhaling. A few species, including lyrebirds, scrubbirds, and New Zealand wrens, are almost flightless. The tail varies from nearly absent to long. The bill is widely variable in shape. It may be hooked or straight to strongly decurved, depending largely on food and feeding habits, and it can be small and weak or large and strong. Passeriforms have legs of short-to-medium

length that are usually strong. The four toes show the usual avian anisodactyl arrangement, three in front and a well-developed hallux behind. Most forms can walk or climb well, or both. Plumage varies widely, from all black to mostly white and from bright colors and bold patterns to cryptic coloration.

Feeding habits and food choices show wide variation. Most species eat insects or small animals; however, some groups have become specialized on a diet of nectar, seeds, fruits, small vertebrates, or even leaves and the waxy covering of berries for part of the year. Those species eating wax have enzymes to digest this material. Methods of obtaining food vary as well. Insect eaters include flycatchers and swallows, which catch insects in the air; creepers and nuthatches, which probe crevices in bark; thrashers, which dig into ground litter and soft earth; and dippers feeding under water. Moreover, groups of passerine birds that feed on seeds, on nectar, by flycatching, or by preying on large insects and small vertebrates have evolved independently several times within the order. Sorting out these convergent groups has been a difficult taxonomic task, which is still not completely resolved.

Song is important to most perching birds for species recognition and courtship. The song in young Oscines is learned from adults. The pair bond is usually strong, with both sexes of most species incubating and caring for the young, which remain in the nest until they are able to fly. A number of species have two or even three broods in one breeding season.

Many arctic and cold-temperature passerine species migrate to warmer areas for the cold months. Some of these migratory flights measure several thousand miles. *See* MIGRATORY BEHAVIOR.

Economic significance. The value of the perching birds to humans is difficult to state in simple terms. However, many species are important in literature and music. In addition, the value of birds, and largely perching birds, to human enjoyment from breeding and training canaries to bird watching and bird feeding during winter months is large. Such activities are perhaps of greater total economic value today than hunting and other commercial activities that are associated with wild birds. *See* AVES. Walter J. Bock

Bibliography. P. Alstrom and K. Mild, *Pipits and Wagtails*, Princeton University Press, 2002; D. Brewer, *Wrens, Dippers, and Thrashers*, Yale University Press, 2001; P. Clement, *Thrushes*, Princeton University Press, 2000; J. del Hoyo et al. (eds.), Order Passeriformes: Broadbills to Tapaculos, *Handbook of the Birds of the World*, vol. 8, Lynx Editions, 2003; J. del Hoyo et al. (eds.), Order Passeriformes: Cotingas to Pipits and Wagtails, *Handbook of the Birds of the World*, vol. 9, Lynx Editions, 2004; J. del Hoyo et al. (eds.), Order Passeriformes: Cuckoo-shrikes to Mockingbirds, *Handbook of the Birds of the World*, vol. 10, Lynx Editions, 2005; J. del Hoyo et al. (eds.), Order Passeriformes: Old World Warblers to Monarch Flycatchers, *Handbook of the Birds of the World*, vol. 11, Lynx Editions, 2006; C. B. Frith and D. W. Frith, *The Bowerbirds: Ptilonorhynchidae*, Oxford University Press, 2004; S. Madge and H. Burn, *Crows and Jays*, Princeton University Press, 2001; H. D. Pratt, *The Hawaiian Honeycreepers*, Oxford University Press, 2005.

Pasteurella

A genus of gram-negative, nonmotile, nonsporulating, facultatively anaerobic coccobacillary to rod-shaped bacteria which are parasitic and often pathogens in many species of mammals, birds, and reptiles. It was named to honor Louis Pasteur in 1887. The organisms grow well on enriched complex media, such as blood agar, and preferably at increased atmospheric carbon dioxide and reduced oxygen pressures. They require organic nitrogen sources and various vitamins, and are capable of both aerobic and anaerobic respiration and carbohydrate fermentation. Nitrates are reduced to nitrites. The base composition of the genome deoxyribonucleic acids range from 38 to 45 mol % guanine plus cytosine. The group is closely interrelated with the genera *Haemophilus* and *Actinobacillus*. *See* ACTINOBACILLUS; HAEMOPHILUS.

Genetic studies have shown that the three groups constitute a family, Pasteurellaceae. The genetically redefined genus *Pasteurella* contains organisms that exhibit the following phenotypic features: positive reactions for alkaline phosphatase, indophenol oxidase, catalase, and fermentation of D-glucose, D-galactose, D-fructose, D-mannose, and sucrose; negative reactions for hemolysis, arginine dihydrolase, production of acid from L-sorbose, L-rhamnose, *m*-inositol, adonitol, and salicin, and splitting of esculin.

The genus contains at least 10 species. *Pasteurella multocida* causes hemorrhagic septicemia in various mammals and fowl cholera, and is occasionally transmitted to humans, mainly in rural areas. Human pasteurellosis may include inflammation in bite and scratch lesions, infections of the lower respiratory tract and of the small intestine, and generalized infections with septicemia and meningitis. *Pasteurella dagmatis* (*P.* "gas," *P.* "new species," or *P. pneumotropica*–type Henriksen), *P. canis*, and *P. stomatis* may cause similar, though generally less severe, infections in humans after contact with domestic or wild animals. Other *Pasteurella* species appear to be confined to particular hosts, for example, *P. gallinarum* and *P. avium* (*H. avium*); and several unnamed species, to birds. The control of pasteurellosis in domestic animals is an economically important task, but the problems of immunoprophylaxis are widely unresolved. *See* PASTEURELLOSIS.

The bovine pathogen *P. haemolytica*, the human parasite *P. ureae*, and *P. pneumotropica* (types Jawetz and Heyl), which are common in rodents and some other mammals, do not belong strictly to the genus *Pasteurella* but are closely related to the *Actinobacillus* group. *Pasteurella aerogenes* (frequently occurring in swine) and the *Pasteurella*-like "SP" group (which may affect guinea pigs and humans)

are not true pasteurellas, and their precise taxonomic positions at the genus level are presently unknown. The etiologic agent of plague is no longer classified as a *Pasteurella* species but has been transferred to the genus *Yersinia*. *See* YERSINIA.

Although drug-resistant *Pasteurella* strains have been encountered, human *Pasteurella* infections are as a rule readily sensitive to the penicillins and a variety of other chemotherapeutic agents. *See* ANTIBIOTIC; DRUG RESISTANCE; MEDICAL BACTERIOLOGY; PLAGUE.

Walter Mannheim

Bibliography. C. Adlam and J. M. Rutter (eds.), *Pasteurella and Pasteurellosis*, 1989; M. Kilian, W. Frederiksen, and E. L. Biberstein (eds.), *Haemophilus, Pasteurella and Actinobacillus*, 1981; N. R. Krieg and J. G. Holt (eds.), *Bergey's Manual of Systematic Bacteriology*, vol. 1, 2001.

Pasteurellosis

A variety of infectious diseases caused by the coccobacilli *Pasteurella multocida* and *P. haemolytica*. However, the term pasteurellosis also applies to diseases caused by any *Pasteurella* species. For example, *P. caballi* is considered to have a causal role in upper respiratory infections, pneumonia, peritonitis, and mesenteric abscesses of horses; and *P. granulomatis* has a causal role in a chronic disease of cattle in Brazil characterized by a progressive fibrogranulomatous process. All *Pasteurella* species occur as commensals in the upper respiratory and alimentary tracts of their various hosts. Although varieties of some species cause primary disease, many of the infections are secondary to other infections or result from various environmental stresses. *Pasteurella* species are generally extracellular parasites that elicit mainly a humoral immune response. Several virulence factors have been identified. *See* VIRULENCE.

Pasteurella multocida. This heterogeneous species consists of capsular (polysaccharide) and somatic (lipopolysaccharide) types. The former antigens are designated by capital letters and the latter by numbers. For example, an important serotype causing fowl cholera is designated A:3; E:2 strains cause hemorrhagic septicemia in cattle in Africa (see **table**). Serotyping is useful in studying the epidemiology of *P. multocida* diseases. *Pasteurella multocida* is the most prevalent species of the genus causing a wide variety of infections in many domestic and wild animals, and humans. It is a primary or, more frequently, a secondary pathogen of cattle, swine, sheep, goats, and other animals. As a secondary invader, it is often involved in pneumonic pasteurellosis of cattle (shipping fever) and in enzootic or mycoplasmal pneumonia of swine. Particular serotypes are the primary cause of the important diseases fowl cholera and epidemic hemorrhagic septicemia of cattle and water buffaloes. The latter disease occurs in many tropical and subtropical countries, particularly in southeast Asia and Africa; however, it is rare in the United States and South America. *Pasteurella multocida* is one of the causes of the pleuropneumonia form of snuffles in rabbits. It is responsible for a variety of sporadic infections in many animals, including abortion, encephalitis, and meningitis. It produces severe mastitis in cattle and sheep, and toxin-producing strains are involved in atrophic rhinitis, an economically important disease of swine. Hemorrhagic septicemia, caused by capsular type B strains, has been reported in elk and deer in the United States.

Pasteurella haemolytica. Two different biotypes, A and T, of *P. haemolytica* are recognized based on pathogenicity, antigenic nature, and biochemical activity. Serotypes based upon differences in capsular antigens are designated by numbers. Specific serotypes are associated with certain animal hosts and diseases; for example, serotype 1 of biotype A is the principal cause of bovine pneumonic pasteurellosis (shipping fever). All strains of *P. haemolytica* produce a soluble cytotoxin (leukotoxin) that kills various leukocytes of ruminants, thus lowering the primary pulmonary defense. The importance of the cytotoxin in the pathogenesis of pneumonic pasteurellosis has been clearly demonstrated experimentally. This species has a primary or secondary role in pneumonia of cattle, sheep, and goats. As mentioned above, it is the principal cause of the widespread pneumonic pasteurellosis of cattle. Shipping fever occurs frequently in cattle that have been stressed by transport. Various viruses or mycoplasmas may be the primary cause of a secondary pneumonic pasteurellosis, although

Principal pasteurelloses and their hosts

Species	Serogroup	Biotype	Diseases	Hosts
Pasteurella multocida	A		Fowl cholera	Chickens, turkeys, ducks, other avian species
			Wide variety of infections	Many animal species
			Variety of secondary infections	Many animal species and humans
	B		Epidemic hemorrhagic septicemia	Mainly cattle and water buffaloes
			Hemorrhagic septicemia	Elk, deer, other wild ruminants
	D		Infectious atrophic rhinitis	Swine
			Secondary invader in pneumonia	Swine
			Various infections	Various animal species and humans
	E		Epidemic hemorrhagic septicemia	Cattle and water buffaloes in Africa
	F		Infrequent infections	Avian species, various animals
Pasteurella haemolytica		A	Primary and secondary pneumonia	Cattle, sheep, goats
			Bovine pneumonic pasteurellosis	Cattle
			Septicemia	Nursing lambs
		T	Septicemia	Feeder lambs

P. haemolytica is considered the primary agent in most outbreaks. Stresses other than those related to transport may also predispose cattle to the disease. Other important diseases caused by certain serotypes of *P. haemolytica* are mastitis of ewes and septicemia of lambs.

Diagnosis and treatment. All of the *Pasteurella* species can be isolated by culturing appropriate clinical specimens on blood agar. They are identified generically by their basic characteristics and identified to species by a number of conventional biochemical tests. *Pasteurella* isolates are routinely subjected to antimicrobial susceptibility tests. Multiple drug resistance is frequently encountered, particularly in fowl cholera and bovine pneumonic pasteurellosis. Treatment is effective if initiated early. Among the drugs used are penicillin and streptomycin, tetracyclines, chloramphenicol, sulphonamides, and some cephalosporins.

Prevention. Sound sanitary practices and segregation of affected animals may help limit the spread of the major pasteurelloses. However, the endemic character of pasteurellosis makes control difficult. Live vaccines and bacterins (killed bacteria) are used for the prevention of bovine and ovine pneumonic pasteurellosis, fowl cholera, epidemic hemorrhagic septicemia, and atrophic rhinitis of swine. *See* PASTEURELLA. G. R. Carter

Bibliography. C. Adlam and J. M. Rutter (eds.), *Pasteurella and Pasteurellosis*, 1989; B. E. Patten (ed.), *Pasteurellosis in Production Animals*, 1993.

Pasteurization

The treatment of foods or beverages with mild heat, irradiation, or chemical agents to improve keeping quality or to inactivate disease-causing microorganisms. Originally, Louis Pasteur observed that spoilage of wine and beer could be prevented by heating them a few minutes at 122–140°F (50–60°C). Today pasteurization as a thermal treatment is applied to many foods, including liquid eggs, crab, fruit juices, pickles, sauerkraut, smoked fish, beer, wine, and dairy products. In foods consumed directly, destruction of pathogens to protect consumer health is paramount, while in products without public health hazards, control of spoilage microorganisms is primary. In fermentation processes, the raw material may be pasteurized to eliminate microorganisms that produce abnormal end products, or the final product may be heated to stop the fermentation at the desired level. In many cases, such as with some dairy products, mild heat treatments may be utilized for all of these purposes.

Milk and dairy products probably represent the most widespread use of pasteurization. Several time-temperature combinations have been approved as equivalent: 145°F (63°C) for 30 min; 161°F (72°C) for 15 s; 191°F (89°C) for 1 s; 194°F (90°C) for 0.5 s; 201°F (94°C) for 0.1 s; 204°F (96°C) for 0.05 s; or 212°F (100°C) for 0.01 s. These precise heat treatments are based on the destruction of the rickettsia *Coxiella burnetii*, which is considered the most heat-resistant nonsporeforming pathogen found in milk. Absolute control of the thermal treatment is essential for safety. Pasteurization of milk has successfully eliminated the spread of diseases such as diphtheria, tuberculosis, and brucellosis through contaminated milk. *See* DAIRY MACHINERY; FOOD MANUFACTURING; MALT BEVERAGE; MILK; WINE. Francis F. Busta

Bibliography. J. M. Jay, M. J. Loessner, and D. A. Golden, *Modern Food Microbiology*, 7th ed., 2006; Z. B. Maroulis and G. D. Saravacos, *Food Process Design*, 2003; P. Walstra et al., *Dairy Technology: Principles of Milk Properties and Processes*, 1999.

Patent

Common designation for letters patent, which is a certificate of grant by a government of an exclusive right with respect to an invention for a limited period of time. A United States patent confers the right to exclude others from making, using, or selling the patented subject matter in the United States and its territories. A United States patent covering a process also, under certain conditions, prohibits the unlicensed sale in the United States of articles made by that process anywhere in the world. Portions of those rights deriving naturally from it may be licensed separately, as the rights to sell, to use, to make, to have made, and to lease. Any violation of this right is an infringement.

An essential substantive condition which must be satisfied before a patent will be granted is the presence of patentable invention or discovery. To be patentable, an invention or discovery must relate to a prescribed category of contribution, such as process, machine, manufacture, composition of matter, plant, or design. In the United States there are different classes of patents for different members of these categories.

Utility patents. Utility patents, which include electrical, mechanical, and chemical patents, are the most familiar; they each have a term beginning upon issue and ending 17 years later. This 17-year class of patent is granted (or issued) for a process, manufacture, machine, or composition of matter which meets the statutory criteria. A process may, for example, be a method of inducing or promoting a chemical reaction, of controlling a computer, or of producing a desired physical result (such as differential specific gravity ore separation by flotation). "Manufacture" means any article of manufacture, and includes such diverse items as waveguides, transistors, fishing reels, hammers, buttons, and corks. "Machines" has its broadest conventional meaning, and "composition of matter" includes, for example, drugs, genetically altered living matter, herbicides, and alloys; patentability is not precluded by the fact that the subject matter is alive. This class of patent affords protection for the claimed invention and, under the appropriate circumstances, can also afford protection for equivalents of the subject matter claimed in the patent which do substantially the same thing

as that claimed subject matter in substantially the same way to achieve the same result. The extent of protection for such equivalents is subject to the limitations provided by the prior art and the history of the prosecution of the patent application. Thus mere changes in form, material, inversion, or rearrangement will not avoid infringement of the utility patent. A straightforward substitution of one active device for another, as a transistor for an electronic tube, is not normally such a change as to avoid infringement of a utility patent.

Design patent. This class of patent is granted for any new, original, and ornamental design for an article of manufacture. To the extent that shape is determined by functional, rather than ornamental, considerations, it is not the proper subject matter for a design patent. Unlike the utility patent, the design patent may be avoided by a change in appearance, although the essential function may be retained. The design patent is issued for 14 years.

Plant patents. This class of patent is granted to one who discovers and asexually reproduces any distinct and new variety of plant other than a tuber-propagated plant or a plant found in an uncultivated state. The right of exclusion extends only to the asexual propagation of such a plant. The term of the plant patent is 17 years from the date of issue by the Patent and Trademark Office.

Procedure. The discussion in the balance of this article is limited to utility patents. In the United States, letters patent are granted upon the making of written application to the Patent and Trademark Office. The usual elements of this application are (1) the title; (2) a brief summary of the invention; (3) a brief description of the drawings, which are required where necessary for the understanding of the subject matter to be patented; (4) a detailed description of the invention and the manner and process of making and using it in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use it, and setting forth the best mode contemplated by the inventor for carrying out the invention; (5) the claims, which define the scope of the requested exclusionary right and which particularly point out and distinctly claim that right; (6) a brief abstract of the technical disclosure; (7) an oath or declaration stating, among other things, that the applicant is the original and first inventor, that the invention was not known or used by others in the United States or patented or described in a printed publication in this or a foreign country before the applicant's invention, or printed or described in a printed publication anywhere in the world or in public use or on sale in this country more than 1 year prior to the date of filing of the United States application.

Owing to the complexities of obtaining proper patent protection, it is customary to retain an attorney who has been registered to practice before the Patent and Trademark Office for the purpose of preparing the application and pursuing it through that office.

When the application is received, it is examined at the Patent and Trademark Office by an examiner. The examiner searches the prior art, consisting of prior patents and other publications, to determine whether the application and claims comply with the rigorous standards required by the statute for patentability. If it does not, the examiner points this out in a letter of rejection, specifying the particulars of the basis for rejection. If the application, on the other hand, is allowable, the examiner sends the applicant a notice of allowance. Following the date when the Patent Office mails a rejection and within the time period specified by the Patent Office, ordinarily 3 months but never more than 6 months, the applicant or the attorney must respond. Normally the application would be amended to make it comply with the statutes. If not, the examiner would issue a final rejection, from which the applicant could appeal to the Board of Appeals and Interferences. If the amended application was found to be in accord with the statute, the examiner would issue a notice of allowance. The patent is issued upon the payment of the issue fee. The entire process requires about $1^1/_2$ years on the average.

As early in the pendency of the application as convenient, it is imperative that the patent examiner be informed of the prior art closest to the invention being claimed known to the applicant or to the applicant's attorney. If it is desired to obtain corresponding patent coverage in foreign countries, it is essential that there be no public disclosure of the invention before the United States application is filed.

Invention. Patents are granted for new inventions or discoveries which are within the statutory classes of patentable subject matter, which were not known in the United States before the invention date and not printed anywhere in the world before that date, which are the subject of applications timely filed, and in which the subject matter is not obvious to one skilled in the art to which the invention relates.

It is generally considered that there are two discernible steps in invention. The first is thinking of the invention; the second is constructing it. The first is termed conception. The second is called reduction to practice.

Conception. The formation in the mind of the inventor of a definite and permanent idea of the complete and operative invention as it is thereafter to be applied in practice is conception. It is not merely a perception of what is done or considered desirable to do, but, going beyond this, how it is to be done in terms of a currently realizable instrumentality or group of instrumentalities. Because conception is a mental act, there must be some external, verifying manifestation in the form of impartation to another if the act is to be established in later controversy over priority of invention. Although oral transmission is adequate in theory, the perishability of the human memory favors the unchanging written word, dated, signed by the inventor, and witnessed by the corroborating party. The corroborating party should be a person who fully understands the invention.

It is this which makes the keeping of written records by the inventor in the course of his or her work so important, because a failure to have this verified external manifestation of the conception may cost the inventor the patent if the date becomes important and is challenged. Because the record is not a proof but only a document capable of proof, it is important that the recording be in some permanent form in which undetectable alteration is practically impossible; otherwise the weight of the proof will be diminished. Predating of the document is damaging, although a record, made a few days later, of a previously observed event is valuable, if appearing on its face as such. The parties involved in a record of conception will be tested for verity by cross examination should the dates ever be challenged.

Frequently the moment of conception is clear only in retrospect, when there is a long sequence of experimental effort directed to attaining the desired result. There is no dramatic thunderclap ushering in the birth of most inventions; hence, it is desirable to keep current dated and witnessed records of all work. Such records tend further to corroborate the conception by providing a clue to the entire thought train and revealing the completeness of understanding by the skill with which its principles are later applied.

Conception alone does not give the inventor any vested right in the invention.

Reduction to practice. Two forms of reduction to practice are recognized, actual and constructive. The filing of a patent application which does not become abandoned is a constructive reduction to practice.

Actual reduction to practice requires that the invention be carried out in a physical embodiment demonstrating its practicability. In a process or method, this is sufficiently done when the steps are actually carried out to produce the desired result. In a machine or article of manufacture, it is required that there be a construction showing every essential feature of the claimed invention. Practicality is demonstrated by operating the apparatus under the conditions which it is anticipated will be encountered in actual service. For example, the testing of an automobile hot-water heater by using water from the hot-water tap has been held not to be an actual reduction to practice because some conditions might occur in a motor vehicle which would not be observed in a stationary installation with a heat source of relatively unlimited capacity. Materials are reduced to practice when they are produced, unless utility is not self-evident, for example, in the case of drug compounds, where it is required to establish that the drug is useful for the purpose stated, not merely harmless. To be effective, the reduction to practice must be by the inventor, by one acting as his or her agent, or by one who has acquired rights from the inventor.

A reduction to practice which results from diligent efforts following conception of the invention gives a vested right in the invention, unless followed by an abandonment. If there is a gap in such diligence, the effective date of the right to assert ownership of the invention is only that date which can be connected by a continuous train of diligence with the reduction to practice. Like conception, reduction to practice must be corroborated by a third-party witness, and it is advisable to have a contemporary written record of what was done and what was observed, accompanied by the date of the observations, to refresh the recollection of the witnesses. The witness must have sufficiently acquainted himself or herself with the internal details of any apparatus to be able later to establish the identity between what was demonstrated and what is sought later to be patented. The diligence required is that which is reasonable under all the circumstances. The safest course is to make every effort promptly to reduce to practice consistent with the inventor's physical, intellectual, and financial capacity. For example, if reduction to practice would be clearly within the inventor's financial means, alternative attempts to secure financing could not constitute diligence.

Interferences. When two or more persons are claiming substantially the same invention, a contest to determine priority between the two is instituted by the Patent and Trademark Office. This proceeding is termed an interference. That party who made the invention first will prevail in the interference and be awarded the patent. Making the invention is proved by corroborated acts of reduction to practice and a continuous bridge of diligence, if possible, to the act of conception, provided that those acts have been in the United States. If any or all of the acts were performed outside the United States, the inventor is limited to the date of his or her first efforts in the United States. Exception is made for foreign inventors when they have first filed a corresponding application for patent in a foreign country which is a signatory to the Paris Convention. In such instance the inventor will be credited with a date corresponding to his or her first foreign filing date if the filing in the United States occurs within a year of that date and there has been no sale or publication of the invention in the United States more than 1 year before the actual filing of the application there.

An applicant may provoke an interference with an issued patent which claims an invention that the applicant believes should rightly belong to him or her. This is done within 12 months of issue of the patent by filing in the Patent and Trademark Office, as part of a pending application for patent, all the claims which it is sought to contest, applying them to his or her disclosure, and requesting the declaration of an interference.

Following declaration of an interference, normally the parties are called upon to file preliminary statements under oath setting forth pertinent data surrounding the genesis of the invention and its disclosure to others. When these preliminary statements have been received and approved by the Patent and Trademark Office, the applicant is notified of the setting of a period of time during which motions may be filed. During this time, access to the adversary's application is permitted, marking one of the few times that the secrecy with which the Patent and

Trademark Office surrounds each patent application is penetrated. During the motion period, various motions for modification or termination of the interference may be presented and set for hearing. When an issued patent is involved, there may be no motion for dissolution of the interference upon the ground that the claim in issue is unpatentable. After the motions are disposed of, times for taking the testimony are set, during which the parties may take sworn statements from their witnesses before duly qualified officers, which are then filed with the Patent and Trademark Office, accompanied by any proper exhibits. The burden of proof in the interference is on the party who is entitled to the later filing date and is therefore known as the junior party. If the junior party does not take testimony during the time allotted, the interference is terminated in favor of the senior party without any testimony being taken by him or her.

The conduct of an interference proceeding is frequently an arduous and complex matter, being fraught with many technicalities.

Inventor. In the United States only a natural person may be an inventor, as distinguished from a corporation. Inventors may be either sole or joint.

Assignment. Patents and applications for patents have the general attributes of personal property, and interests in them are assignable by instrument in writing. Such assignments will be recorded by the Patent and Trademark Office upon filing of a request accompanied by a copy of the assignment and payment of the proper fee. To be good against subsequent purchasers without notice, the assignment must be recorded within 3 months from its date, or before the date when such transfer of rights was made.

Witnesses. No joint inventor can serve as a corroborating witness for another joint inventor in connection with the invention which they have jointly made. Beyond this, the rules normally governing witnesses apply. A corroborating witness should fully understand the invention.

Enforcement. Enforcement is available against the manufacturer, the user, and the seller in the United States who perform acts within the scope of the patent. Those who induce infringement by another and those who contribute to the infringement by another by, for example, knowingly providing special parts with no use except in the infringement, are also infringers. Enforcement of patents is through the federal judicial system, action being initiated by the patent owner in the federal judicial district where the defendant resides, or where the defendant has committed the alleged act of infringement and has a regular and established place of business. Action may also be started by one accused of infringement in any place where the patent owner resides or is licensed to do business. Damages, no less than a reasonable royalty, may be awarded, and an injunction granted, prohibiting further infringement by the defendant. Interest on the award prior to the judgment and subsequent to infringement is normally awarded. If damages are awarded, there can be recovery for a period not longer than 6 years preceding the filing of the complaint.

Patents may also be enforced by the International Trade Commission (ITC), which can prevent imports into the United States of goods that infringe a valid patent. No damages are awarded in an ITC proceeding of this type. All appeals from district courts and from the ITC in such cases are heard by the Court of Appeals for the Federal Circuit.

When the infringer is the United States government, or a supplier of the government operating with the authorization and consent of the government, the suit must be filed against the government in the U.S. Court of Claims.

A patent may become unenforceable through improper use, for example, use as a part of an act in violation of the antitrust law.

Under the 6-year statute of limitations, an action may be maintained on a patent up to 6 years after its expiration, the accounting being limited in such instance to damages for infringing activities within the life of the patent.

Licenses. Licenses to operate under a patent may be granted, either nonexclusive or exclusive, and may be in writing or may arise as a necessary implication of other actions of the patentee. Except for an exclusive license, licenses are not ordinarily recorded by the Patent and Trademark Office.

Foreign filing. A United States patent is void if the United States inventor should file an application for the same invention in any foreign country before 6 months from the date when he or she filed in the United States, unless license to do so is obtained from the Commissioner of Patents and Trademarks. Because the United States is a party to the Paris Convention, applications filed in foreign countries that are also members of the Paris Convention within 12 months of the date when the parent case is filed in the United States are accorded an effective filing date in the foreign countries which is the same as the date of filing in the United States. The procedural details of foreign filing vary from country to country and from time to time. The Patent Cooperation Treaty, to which the United States, Japan, most European countries, and other countries are party, has greatly facilitated filing of patent applications in foreign countries. In particular it has established a standardized application and claim format acceptable in the patent offices of all member countries.

Nuclear and atomic energy. Patents may not be issued for inventions or discoveries useful solely in the utilization of special nuclear material or atomic energy in an atomic weapon, nor does any patent confer rights upon the patentee with respect to such uses. As a substitute for the patent incentive for disclosure in this field, there is a mandatory provision in the law that requires anyone making an invention or discovery useful in the production of special nuclear material or atomic energy, in the utilization of such special nuclear materials in the production of an atomic weapon, or in the utilization of atomic energy weapons, to report such invention promptly to the

Department of Energy, unless it is earlier described in an application for patent. Awards may then be requested from the Patent Compensation Board of the Department of Energy.

Aeronautics and astronautics. No patent may be issued for an invention having significant utility in the field of aeronautics or space unless there be filed with the Commissioner of Patents and Trademarks a sworn statement of the facts surrounding the making of the invention and establishing that the invention was made without relation to any contract with the National Aeronautics and Space Administration. Such a contract right is subject to waiver by the administrator, but in the event of waiver the administrator is required by law to retain a license for the United States and foreign governments.

Marking. A patented product may carry a notice of this fact, including the patent number. The affixation of such notice is of advantage in establishing constructive notice and fixing the period for which damages may be collected. In many cases damages may not be collected, in the absence of such marking, without actual notice to the infringer. By statute, false marking is a criminal offense. The marking "Patent pending" or "Patent applied for" gives no substantive rights, but may give rise to sanction under the above-mentioned statute if without foundation in fact.

Foreign patents. The principles guiding most foreign patent systems are essentially the same as those underlying the United States system: the granting of a carefully defined exclusionary right for a limited term of years in return for a laying open of the invention through letters patent. In the socialist countries the grant is sometimes made for innovation as well as invention and establishes a right to compensation. There are some differences in the classes and terms of patent, and in the nature of subject matter which may be patented. The most significant departures are the requirement in other countries that there be no public divulgation of the invention before effective date of the patent application, the measuring of the term, in most instances, from the filing date, and compulsory licensing.

In some examination countries the mounting magnitude of material to be searched so slowed the rate of disposition of patent applications that the delays and the accumulation of pending patent applications created intolerable uncertainties. For example, in Japan the practice of deferred examination has been introduced, according to which the application is given cursory formal examination, and published after 18 months, but examination on the merits is delayed. Examination on the merits is made upon request, but must be made before enforcement of the patent or the end of 7 years, whichever first occurs. The theory was that only a few of the patents are ever brought to litigation, and many are without value after 7 years, so that the load on the examining staff would be reduced to manageable proportions. In fact, however, the system created additional uncertainties for the business community.

Furthermore, in Japan after the patent application has been allowed but before it is issued as a patent, it is published for opposition. Interested parties can then further delay issuance of the patent by asserting prior art not previously considered by the Patent Office alleging it prevents the issuance of a valid patent. Opposition proceedings can last for years.

To eliminate duplicate searching of the same invention by the different patent offices and to standardize the requirements for patent applications, as well as the form of patent claims, the Patent Cooperation Treaty, as noted above, has been entered into by many countries. Under this treaty the application is examined by a competent searching authority, the criteria for which are set by the treaty. The application, together with the results of that search, is then published. Any country to which the applicant subsequently applies has the benefit of that search, which may be utilized by that country in determining whether the invention is in fact patentable in that country.

In somewhat similar fashion the countries of Europe have by treaty established a European Patent Office in Munich to which application can be made. The European Patent Office conducts the search on behalf of the other countries of Europe who are members of that treaty. As in the case of the Patent Cooperation Treaty, further searching may be done by each country, but need not be.

Copyrights and trademarks. Authors of literary, musical, artistic, dramatic, and certain other types of intellectual works such as computer programs protect their creations by copyright, which is provided under the federal law. Copyright protection exists in such original works of authorship when they are fixed in any tangible medium of expression. The exclusive right protected by a copyright is, in general, the right to copy or print the work, to sell or distribute such copies, and to dramatize, perform, or translate the work. The protection provided by copyright extends only to the particular form or expression of the author's idea—as distinguished from the idea itself, which may be freely used by others. The term of copyright protection extends for the author's life plus 50 years. Copyrights are registered in the Copyright Office, which is part of the Library of Congress. Registration is a requirement for enforcement of the copyright. Printed forms are provided by the Copyright Office to facilitate the registration of copyrights; registration of the copyright is effected by completing the forms and sending them to the Copyright Office together with copies of the work copyrighted. Infringers of the copyright are liable for the actual damages of the copyright owner, as well as the infringer's profits—addition to an injunction. The United States has become a signatory to the Berne Convention. As a result, notice of copyright on the copyrighted work is no longer essential; it is, however, highly desirable because monetary recovery for infringement by another is affected by such marking.

A trademark is a word, symbol (that is, a design, a letter, or a number), or any combination of these

elements that is used by a merchant or manufacturer to identify the origin or sponsorship of his or her goods and services from competing goods or services. Technically, a service mark identifies a service, but for purposes of brevity the term trademark should be read to include service mark here. A trade name identifies a company and should not be confused with a trademark, which identifies a product.

A trademark serves three separate functions: it can indicate the ownership and origin of products or services; it can serve as an advertising device; and it provides a guarantee of quality or degree of product or service consistency. Not every word, name, symbol, or design can function as a trademark. The common descriptive name of a product (generic name) cannot be a trademark.

A trademark derives its essential strength in the marketplace by consumers who recognize the differentiating character of such word, symbol, or other device. Trademarks may be registered, after being used on goods described in the registration, in each of the states, and may also be registered federally in the Patent and Trademark Office. Federal registration may be obtained after the mark is used in "commerce" as defined by the Constitution; furthermore, after November 1989 applications to obtain federal registrations could be filed on the basis of good-faith "intent to use" the specified trademark on the goods or services set out in the application. The application does not mature into a registration until a declaration of actual use is filed within the prescribed period of time. In addition, of course, the trademark must be of such nature as to distinguish the goods of the manufacturer or merchant from the goods of others. When registered by the Patent and Trademark Office, the owner of the mark may show such registration by using the symbol®. Rights in trademarks may last indefinitely; certificates of registration must, however, be renewed periodically. Federal registration of a trademark provides nationwide constructive notice of the registrant's claim of ownership. Any person who uses on his or her goods (or services) a mark likely to cause confusion in the marketplace with goods bearing a registered mark is liable in a civil action (with certain exceptions) to an injunction preventing further use of the confusingly similar mark on similar goods and to pay damages sustained by the plaintiff as well as defendant's profits. The district courts of the United States have jurisdiction of actions arising out of infringement of federally registered trademarks, with appeals being taken to the regional circuit courts of appeal. Federally registered trademarks, under certain circumstances, may also be protected from infringement through exclusion by the Custom Service of goods with infringing marks; they may also be enforced by the International Trade Commission, which can order the borders closed to goods of foreign origin carrying marks that infringe the registered mark. Donald W. Banner

Bibliography. H. L. Barber, *Copyrights, Patents and Trademarks: Protect Your Rights Worldwide*, 2d ed., 1994; Matthew Bender and Co., *Patent Law Annual*, annually; Matthew Bender and Co., *Patent Law Developments: Enforceability of Rights*, 1965; Matthew Bender and Co., *Patent Law Developments: Proprietary Intellectual Rights*, 1964; A. W. Deller, *Deller's Walker on Patents*, 1964; P. Goldstein, *Copyright, Patent Trademark and Related State Doctrines: Cases and Materials on the Law of Intellectual Property*, 4th ed., 1999; T. A. Penn and R. D. Foltz, *Understanding Patents and Other Forms of Legal Protection*, 1990; P. D. Rosenberg, *Patent Law Basics*, 1992; P. D. Rosenberg, *Patent Law Fundamentals*, 2d ed., 3 vols., 1980; U. S. Bureau of National Affairs Legal Services Staff, *United States Patents Quarterly*, quarterly; U.S. Department of Commerce, Patent Office, *Manual of Patent Examining Procedures*; H. C. Wegner, *Patent Law in Biotechnology, Chemicals, and Pharmaceuticals*, 1992.

Paterinida

A small extinct order of inarticulate brachiopods that ranges in age from Early Cambrian to Middle Ordovician. The shell is chitinophosphatic in composition, and its outline is circular or elliptical. The ventral (pedicle) valve is more convex than the dorsal (brachial) valve. The pedicle was absent or emerged between the valves. The ventral valve has a flattened posterior area (pseudointerarea) that is divided by a triangular notch which is closed to a variable extent by a plate (homeodeltidium). Comparable, but not as extensive, structures occur at the posterior end of the dorsal (brachial) valve. The muscle scars are unusual compared to other inarticulates, and form narrow triangular tracks radiating from the posterior extremity of each valve (see **illus.**). Except for

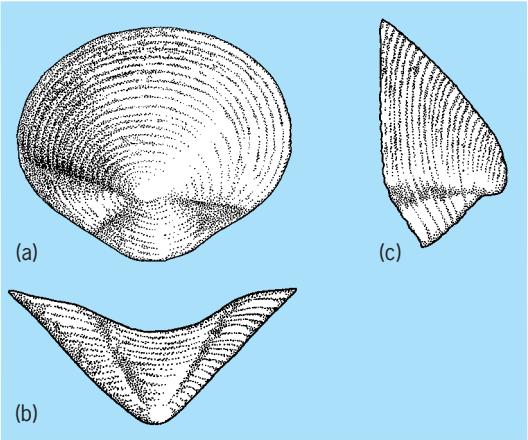

***Paterina*: (*a*) exterior, (*b*) posterior, and (*c*) lateral view of pedicle (ventral) valve. (*After C. D. Walcott, Cambrian Brachiopoda, USGS Monogr., vol. 51, 1912*)**

their shell composition and lack of articulation, the Paterinida resemble articulate brachiopods. Members were presumably epifaunal and sessile. *See* BRACHIOPODA; INARTICULATA. Merrill W. Foster

Pathogen

Any agent capable of causing disease. The term pathogen is usually restricted to living agents, which include viruses, rickettsia, bacteria, fungi, yeasts, protozoa, helminths, and certain insect larval stages.

Pathogenicity is the ability of an organism to enter a host and cause disease. The degree of pathogenicity, that is, the comparative ability to cause disease, is known as virulence. The terms pathogenic and nonpathogenic refer to the relative virulence of the organism or its ability to cause disease under certain conditions. This ability depends not only upon the properties of the organism but also upon the ability of the host to defend itself (its immunity) and prevent injury. The concept of pathogenicity and virulence has no meaning without reference to a specific host. For example, gonococcus is capable of causing gonorrhea in humans but not in lower animals. *See* BACTERIA; FUNGI; IMMUNITY; MEDICAL MYCOLOGY; MEDICAL PARASITOLOGY; MYIASIS; PLANT PATHOLOGY; PLANT VIRUSES AND VIROIDS; PROTOZOA; RICKETTSIOSES; VIRULENCE; VIRUS. Daniel N. Lapedes

Pathology

The study of the etiologies, mechanisms, and manifestations of disease. Techniques and knowledge gained from other disciplines, including anatomy, physiology, microbiology, biochemistry, and histology, are utilized. The information obtained from the study of pathology is necessary prior to developing methods with which to control and prevent disease.

Etiology. The concepts concerning the etiologies of disease have undergone tremendous evolution. For much of history, etiologies were associated with mystical forces. Skeletal remains from prehistoric times had a surprising number of trepanned skulls, that is, in which holes had been cut. It is thought that a belief in evil spirits, trapped within the body and manifesting their presence by producing illness, promoted this surgery. A major shift in thought occurred in the fourth century B.C. when Hippocrates taught that disease is a natural occurrence of the biological world. Unfortunately, during the Middle Ages there was a return to mysticism. The enlightenment of these Dark Ages began in the sixteenth century. Vesalius, an anatomist, broke with tradition and published what he saw rather than the accepted dogma. An understanding of the normal anatomy had to be gained before disease responses could be appreciated. In the seventeenth century modern physiology was born as a result of the work of William Harvey on the circulatory system. The rate of development of scientific knowledge became explosive in the nineteenth century and continues. An entirely new perspective in biology developed and provided a basis for the modern study of disease. A key event was the development of the light microscope. With the aid of this instrument, J. Schleiden (1804–1881) and T. Schwann (1810–1882) understood that all life is composed of cells. R. Virchow (1821–1902) applied the cell theory to the pathogenesis of disease. He recognized that diseases were the product of the alteration of cell structure and function and that these changes were often reflected in gross anatomic alterations.

Thus, with the light microscope it became possible to correlate the observed signs and symptoms in an individual with cellular changes. In its early stages pathology was very descriptive. Diseases were understood and categorized, in part, by how gross and microscopic anatomy was altered. In the last half of the nineteenth century, by using this approach to pathology, coupled with microbiological techniques, it was learned that the major causes of human death were biotic agents: protozoa, bacteria, viruses, and fungi. Infectious diseases took a heavy toll in human lives. Better sanitation and public health measures were instrumental in controlling these diseases, and the production of antibiotics and immunization procedures further reduced their importance.

Mechanisms. Understanding the mechanisms underlying disease responses is a significant task. Whereas the light microscope enabled scientists to look at cells and observe their behavior, the development of the electron microscope by B. von Borries and E. Ruska in 1938 in Germany enabled scientists to look into cells with greater resolution than before. Recognition of subcellular structures became relatively easy. The term molecular biology was introduced in 1938. It is a field of biology which is concerned with the analysis of cell structure and function. Smaller and smaller components of the cells are dissected out and manipulated, in order to understand their normal function and to see what occurs when systems are perturbed in various ways. An offshoot of this science is molecular pathology, which approaches the study of disease mechanism in the same way. *See* MOLECULAR BIOLOGY; MOLECULAR PATHOLOGY.

It is now apparent that all diseases reflect changes at the molecular level. Scientists are beginning to understand what these biochemical alterations are in some diseases. An example of a disease whose underlying mechanism is understood is xeroderma pigmentosum, the result of an inherited disorder. Affected individuals are very sensitive to the ultraviolet rays of sunlight. Their skin ages rapidly, and they develop skin cancer, often as early as their late teens. It is now known that this disease is due to a defect in the enzyme which repairs the damage done to deoxyribonucleic acid (DNA) by ultraviolet light. Unfortunately there is no known cure. Individuals with xeroderma pigmentosum must avoid sunlight. This inhibits the progress of the disease but does not prevent it, since apparently other agents produce the same type of damage as ultraviolet light.

Another rare disease, also attributable to an inborn genetic error, has a more optimistic prognosis now that the biochemical basis of its manifestations has been worked out. This is phenylketonuria,

which is characterized by mental retardation, too little pigment in the skin, and sometimes eczema. These responses have been traced to the absence of an enzyme which converts one of the amino acids, phenylalanine, obtained in the diet, to another product. Thus, phenylalanine accumulates in the blood and spinal fluid. If the disorder is recognized shortly after birth, the effects can be minimized by giving the child a diet which is very low in phenylalanine. Phenylalanine cannot be completely eliminated because it is necessary for the synthesis of proteins. In summary, the goal of molecular pathology is the explanation of the pathogenesis of disease on the basis of active or reactive portions of molecules, energy states, or alterations in molecular constituents. *See* PHENYLKETONURIA.

Manifestations. A diagnosis of a disease is made by observation of manifestations of the process. Alterations within cells are reflected in changes in cellular function and often in structure. When enough cells are involved in a disease response, tissues, organs, and the whole organism may exhibit various lesions. Historically, alterations in gross or microscopic anatomy were very important when identifying a disease process. In human pathology, these observations were used in conjunction with the subjective manifestations, or symptoms, reported by the affected person. This kind of information is still very important. However, added to this is a whole realm of biochemical and physical assays which, in many cases, are very sensitive and give definitive answers to the type of disease process which is occurring.

The interaction between an etiologic agent and an individual may evoke one or more responses. Cells can respond to injury in a limited number of ways. Cellular responses may result in degeneration, death, disturbances of growth and specialization, teratogenesis, inflammation, neoplasia, or aging. The nature of the response is modified by the agent, dose, portal of entry, or duration of exposure, as well as by host factors such as age, sex, nutrition, species, and individual susceptibility. For instance, an important cause of death of humans is a specific type of cancer of the lung, bronchogenic carcinoma. A cell or cells are stimulated to undergo an abnormal proliferation, and a malignant process occurs. It is now known that this disease is produced in response to some agent in cigarette smoke. Since all smokers do not eventually have this cancer, there is obviously individual variation in the cellular response.

Branches. There are many branches of pathology. Divisions are made depending upon focus of interest. Clinical pathology is concerned with diagnosis of disease. As medicine has expanded, subspecialties such as surgical pathology and neuropathology have developed. Experimental pathology attempts to study disease mechanisms under controlled conditions. General pathology covers all areas, but in less detail, and serves in medical education. *See* CLINICAL PATHOLOGY.

Another area of pathology is environmental pathology, which deals with disease processes resulting from physical and chemical agents. At present, the leading causes of death have environmental agents as the known or suspected major etiologic factors; these diseases include heart disease, atherosclerosis, and cancer. It is believed that with understanding, many such diseases, like those produce in response to biotic agents, can be brought under control.

In summary, pathology considers the factors which are known or suspected to be related to alterations of form or function. It attempts to produce or implicate an etiologic agent and seeks to delineate the mechanisms involved in producing the sequential changes in the body. *See* DISEASE.

N. Karle Mottet; Carol Quaife

Pathotoxin

A chemical of biological origin, other than an enzyme, that plays an important causal role in a plant disease. Most pathotoxins are produced by plant pathogenic fungi or bacteria, but some are produced by higher plants, and one has been reported to be the product of an interaction between a plant and a bacterial pathogen. Some pathogen-produced pathotoxins are highly selective in that they cause severe damage and typical disease symptoms only on plants susceptible to the pathogens that produce them. Others are nonselective and are equally toxic to plants susceptible or resistant to the pathogen involved. A few pathotoxins are species-selective, and are damaging to many but not all plant species. In these instances, some plants resistant to the pathogen are sensitive to its toxic product.

Selectivity and disease specificity. Selective pathotoxins, sometimes called host-specific toxins, provided the first mechanisms to account for the remarkable specificity of several important plant diseases. Often, only a single group of plant varieties, which possess a single specific genetic factor, is susceptible to a particular disease. Pathogen-produced selective pathotoxins are equally specific; they are up to 400,000 times more toxic to plants susceptible to the pathogen involved than to those that are resistant. From a scientific viewpoint, selective pathotoxins have been most valuable as model systems in investigations of the physiology of disease and the nature of disease resistance in plants. Substituted for the pathogens that produce them, they have made it possible to follow metabolic changes in diseased plants without complications introduced by metabolic activities of a living pathogen. Moreover, pathological events that occur over a period of days or weeks in infected plants can be compressed to a few hours with toxin treatments.

Selective pathotoxins have had useful practical applications in relation to two destructive epidemics, Victoria blight of oats in the late 1940s and southern corn leaf blight in 1970. Each of these epidemics followed the introduction of a single agronomically desirable genetic character into nearly all commercial cultivars of a crop, and in each case

this desirable character rendered the entire population susceptible to a previously unimportant fungal pathogen.

In Victoria blight of oats, the genetic factor, Pc-2, introduced from the cultivar Victoria, conferred a high degree of resistance to the crown rust pathogen and, at the same time, extreme susceptibility to a new fungal pathogen, *Helminthosporium victoriae*. In artificial culture *H. victoriae* produced a metabolite, victorin, which was highly toxic only to oat cultivars that carried the Pc-2 factor and that hence were susceptible to Victoria blight. After all attempts to combine resistance to rust and Victoria blight by conventional breeding techniques had failed, victorin was used to screen 80,000,000 oat seedlings for resistance to this disease. A few of the survivors (**Fig. 1**) were resistant to both crown rust and Victoria blight. Pathotoxins produced by other plant pathogens have been used in a similar way to screen for resistance to diseases of sorghum, citrus, and sugarcane. Pathotoxins are being used as screening agents for disease resistance in conjunction with regeneration of plants from protoplasts and cell-suspension cultures.

In the southern corn leaf blight epidemic, the introduced character was a cytoplasmic factor for male sterility, designated Tms for Texas male sterile, which made possible the production of hybrid corn seed without the tedious and costly process of mechanical detasseling. Southern corn leaf blight, caused by *H. maydis*, had long been known as a minor disease in the southern regions of the United States. In 1970, however, a new race of the fungus, *H. maydis* race T, highly virulent to plants with Tms cytoplasm, swept north into the Corn Belt and destroyed more than 1,000,000 acres (400,000 hectares) of the crop. Although producers of hybrid corn seed immediately

Fig. 1. Oat seedling resistant to Victoria blight obtained by screening a susceptible population with the pathotoxin victorin, a product of the pathogen *Helminthosporium victoriae*.

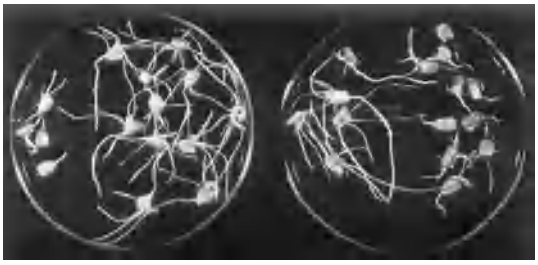

Fig. 2. Separation of corn seedlings resistant (N) and susceptible (T) to southern corn leaf blight by treatment with a pathotoxin produced by *Helminthosporium maydis* race T.

reverted to mechanical detasseling, inadequate supplies of seed with resistant, normal cytoplasm made it necessary to plant much of the corn acreage in 1971 with varying mixtures of seeds with normal and Tms cytoplasm. The selective pathotoxin produced by *H. maydis* race T provided an efficient agent for testing thousands of lots of seeds to verify the proportions of the two cytoplasms in mixtures (**Fig. 2**).

Chemistry and mode of action. Most chemically characterized pathotoxins are nonselective, and some (the gibberellins, for example) function as plant growth regulators. The gibberellins were discovered in Japan during studies of the bakanae or "foolish seedling" disease of rice, caused by the fungus *Gibberella fujikuroi*. In culture this fungus produces gibberellins, and plants infected by the fungus have excessively long internodes, which typify the effects of this group of plant hormones. Fusicoccin (**Fig. 3***a*), produced by *Fusicoccum amygdali*, which causes a wilt disease of peach and almond trees, is another nonselective pathotoxin with growth-regulator properties. This chemical also causes stomates to open. Tentoxin (Fig. 3*b*), a product of *Alternaria alternata*, is species-selective; it causes a striking variegated chlorosis in cucumber, cotton, lettuce, and many other sensitive plants, but has no effect on tobacco, radish, willow, or other insensitive species. Tentoxin binds specifically to coupling factor 1 in chloroplasts of sensitive species, and presumably acts by disrupting energy generation in the photophosphorylative electron transport system. *See* GIBBERELLIN.

The only selective pathotoxins fully characterized chemically are the AM-toxins (Fig. 3*c*), produced by *A. mali*, which attacks apples. AM-toxin I has been reported to disrupt the cellular permeability of susceptible apple tissues at a concentration of $10^{-9}M$, whereas concentrations 10,000 times higher were required to produce similar effects on resistant tissues. HMT-toxin (Fig. 3*d*), produced by *H. maydis* race T, causes rapid swelling and uncoupling of electron transport in mitochondria from Tms corn plants, but has no such effects on mitochondria from disease-resistant plants with normal cytoplasm. These results suggest that HMT-toxin disrupts energy generation in the mitochondrial system, and such a site of action would account for the cytoplasmic inheritance of resistance to southern corn leaf blight.

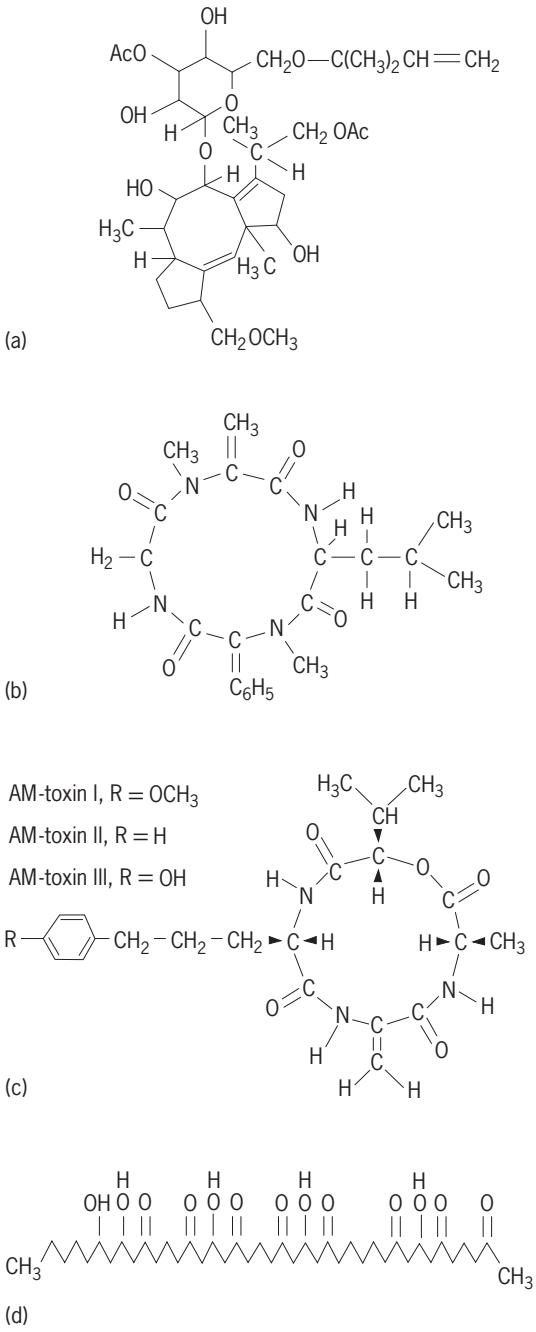

Fig. 3. Chemical structures of four pathotoxins. (*a*) Fusicoccin. (*b*) Tentoxin. (*c*) AM-toxin. (*d*) HMT-toxin.

Extensive attempts to purify other pathotoxins, especially victorin, have been frustrated by instability of the toxin, but these, like the AM-toxins, rapidly disrupt cell permeability. Such disruption would result in the mixing of cell contents, and this in turn would block energy generation through electron transport systems. It thus appears that pathotoxins interfere with the energy generation required for cell maintenance and repair, either by effects on specific energy-generating sites or by disruption of compartmentalizing membranes. *See* PLANT PATHOLOGY; TOXIN.

Harry Wheeler

Bibliography. R. D. Durbin (ed.), *Toxins in Plant Disease*, 1981.

Pattern formation (biology)

The mechanisms that ensure that particular cell types differentiate in the correct location within the embryo and that the layers of cells bend and grow in the correct relative positions. Pattern formation is one of four processes that underlie development, the others being growth, cell diversification, and morphogenesis. Pattern formation can more easily be understood by first briefly clarifying the importance of the other three processes.

Growth refers to an increase in size of a tissue, organ, or body. It may occur through any of a number of events, such as increase in cell number, increase in cell size, or production of extracellular material by a cell. Cell diversification occurs throughout development. At the outset there is only one cell, the fertilized egg, but as development proceeds new cell types with different characteristics emerge. In the human body, for example, there are over 200 identified cell types, each with a precise function and location. The challenge in the study of cell diversification in an embryo is to establish how these cells become different even though they all have the same deoxyribonucleic acid (DNA) complement. *See* ANIMAL GROWTH; CELL DIFFERENTIATION; PLANT GROWTH.

Morphogenesis is the creation of form, that is, the molding of a tissue into a recognizable and precise shape. During development, many cells migrate from one part of the embryo to another, and sheets or layers of cells bend to form pockets, bulges, thickenings, and grooves. Morphogenesis involves cells sticking together or breaking away, changing shape, dividing, and making extracellular material. *See* MORPHOGENESIS; PLANT MORPHOGENESIS.

Cell arrangement. Pattern formation is the creation of a predictable arrangement of cell types in space during embryonic development. The central importance of the study of pattern formation to the understanding of how embryonic development proceeds can be understood by comparing the cell types that are present in a human arm and a human leg. Although the two limbs are made up of identical sets of cell types, such as skin cells, bone cells, muscle cells, nerve cells, and blood cells, anyone can recognize an arm as distinct from a leg. The difference is not what cell types are present but how they are arranged in space relative to one another. Thus, arms and legs have five digits—but each finger is a different shape and length, and fingers are easily distinguished from toes.

Types of patterns. The types of patterns of cell types found in animals and plants can be conveniently described as simple or complex.

Simple patterns. Simple patterns involve the spatial arrangement of identical or equivalent structures such as bristles on the leg of a fly, hairs on a person's head, or leaves on a plant. Such equivalent patterns are thought to be produced by mechanisms that are the same or very similar in the fly and the plant. The glands in the growing skin of a salamander, for example, are found in a regular hexagonal array,

and new glands emerge in the pattern only when the older glands move farther apart than a certain threshold distance in the growing animal. This suggests that the glands themselves produce an inhibitor of new gland development that diffuses into the skin all around the gland. Only when the glands become separated by a prescribed distance in the growing skin does the level of the inhibitor fall low enough to allow new gland formation. The same lateral inhibition type of mechanism could easily explain how leaf patterns develop at the growing apices of plants. The exact nature, timing, and distribution of an inhibitor would determine the exact pattern that leaves have one to another. Such arrangements, called phyllotactic patterns, may be highly varied, from spiral to alternate.

Complex patterns. Complex patterns are those that are made up of parts that are not equivalent to one another. In the vertebrate limb, for example, the structure of the forelimb is different at each level, with one bone (humerus) in the arm, two bones (radius and ulna) in the forearm, and a complex set of bones making up the wrist and the hand. The theoretical framework that allows a basis for understanding how nonequivalent parts are patterned during development is called positional information.

Two stages exist in the positional information framework. First, a cell must become aware of its position within a developing group, or field, of cells. This specification of cellular position requires a mechanism by which each cell within a field can obtain a unique value or address. The second component is the interpretation of the positional address by a cell to manifest a particular cell type by the expression of a particular set of genes. Positional specification mechanisms have been examined in many different systems ranging from vertebrate limbs to insect larvae and coelenterate polyps such as *Hydra*. In each of these diverse animal systems, common principles of positional specification emerge.

Segmentation pattern in Drosophila. A similar principle is seen in the larva of *Drosophila*. The advantage of this system is that genetic knowledge of the fruit fly *Drosophila* is more comprehensive than for any other animal, and mutants have been used in analyzing the positional specification mechanisms. Genetic analysis, both in terms of looking for mutants that show interesting pattern alterations and in terms of molecular genetics and the power of techniques for manipulating deoxyribonucleic acid (DNA), have led to the unraveling of the early stages of pattern formation in this animal. The *Drosophila* larva is segmented and has three regions: head, thorax, and abdomen. The head is derived from a fused set of segments, the thorax has three segments, and the abdomen has eight. Each segment is different, and so it is of interest to determine how this pattern is brought about during development. The explanation is complex but can be summarized in five stages, each involving a particular set of genes. The first four stages involve positional specification and the creation of 14 segments in the larva. The final stage involves the interpretation of positional signals created at earlier stages and is the stage at which the segments differentiate themselves from each other.

The first stage involves the establishment of gradients of morphogens in the newly laid insect egg. In *Drosophila*, there are two gradients of proteins in the egg, one with its high point at the anterior end, and one at the posterior end. Once established, the gradients must be interpreted to establish a segmental pattern. Initially, seven segments, which appear to be equivalent at the outset, form within the embryo. This development involves two sets of genes, called gap genes and pair rule genes. There are a number of gap genes, each one turned on in response to maternally derived gradients. If these genes are mutated, they create large gaps in the larval body. The pair rule genes, of which there are also a number, express in response to the gap gene products. The products of the pair rule genes appear as seven stripes down the length of the embryo from the head to the tip of the abdomen. The exact position of the stripes for each pair rule gene is different, and between them they cover the full length of the larva. The 14 stripes eventually become 14 segments with the expression of the fourth group, the segment polarity genes. The stripes can be visualized by using antibodies to the protein product of the genes (see **illus.**). Finally, with the 14 segments established, the function of the fifth group of genes becomes obvious. These are called homeotic genes, and they are concerned with giving each segment an anatomical character of its own. Mutants in the homeotic genes, therefore, may be very dramatic, turning one region of a larva into another. Because the larva will metamorphose into a fly, these mutants may effect the final anatomy of the fly. Many examples of these homeotic mutations exist. *See* DEVELOPMENTAL GENETICS; EMBRYONIC DIFFERENTIATION; GENE ACTION.

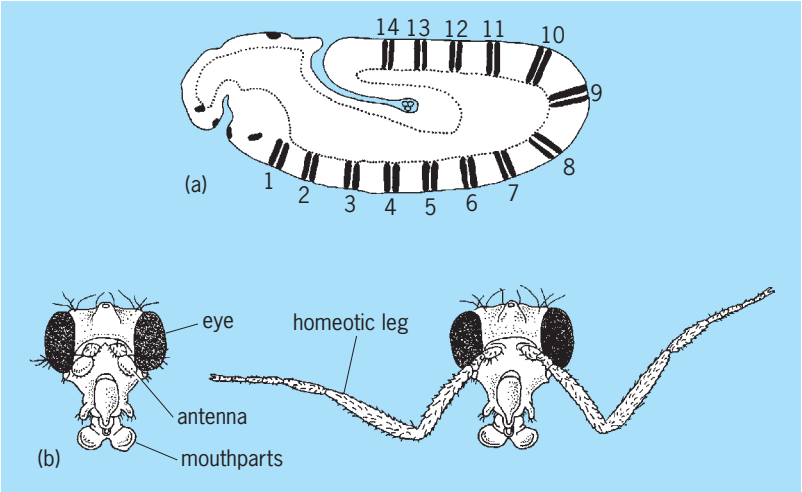

Segmentation and mutagenesis in *Drosophila*. (*a*) Tissue section of the developing *Drosophila* larva showing 14 segments as revealed with antibodies to proteins produced by segment polarity genes. Two such gene products are shown, engrailed (spots) and wingless (stripes). (*b*) Once segment number has been established, homeotic genes act to give each segment a character. Mutations in such genes alter the character of given segments, turning one body part into another. In this example, a mutant (antennapaedia) has converted the antennae into legs.

Pattern formation genes. It seems that the two-step positional information framework is correct in principle, but the interpretation steps following creation of initial spatially regulated patterns of gene expression are very complex and involve a series of different gene classes. Perhaps the most surprising realization is that the patterning mechanisms are largely universal in animals. *See* DEVELOPMENTAL BIOLOGY.

Nigel Holder

Bibliography. M. Fougerean and R. Stora (eds.), *Cellular and Molecular Aspects of Developmental Biology*, 1986; A. F. Hopper and N. H. Hart, *Foundations of Animal Development*, 2d ed., 1985; G. Malacinski, *Pattern Formation*, 1984; V. Walbot and N. Holder, *Developmental Biology*, 1987; C. C. Wylie (ed.), *Determinative Mechanisms in Early Development*, 1986.

Pauropoda

A class, and perhaps the most obscure group, of the Myriapoda. They are pale creatures, no more than 0.04–0.08 in. (1–2 mm) in length, inhabiting damp situations in leaf litter, under bark, stones and debris, and in humus and similar detritus. Their size and retiring habits suggest rarity; this, however, has been refuted by later studies that show them to occur in huge numbers in suitable habitats. Apparently very widely distributed as a class, they have been undiscovered only in deserts and in the arctic and antarctic regions. Like millipedes, they are progoneate (that is, their reproductive tracts open on the anterior end of the body) and have one pair of maxillae, and their trunk segments display a certain degree of amalgamation manifest in the tendency to form diplotergites, but not diplosegments. Their peculiar bifurcate antennae and adult complement of 12 trunk segments with nine pairs of functional legs are distinctive within the myriapod complex (see **illus.**). All pauropods lack eyes, spiracles, tracheae, and a circulatory system.

What they eat and how they live is virtually unknown. They probably subsist largely or entirely upon decaying plant and animal matter; some may consume microscopic animals. Studies based upon a few forms reveal an anamorphic ontogeny similar to that of diplopods. The young pauropod hatches with three pairs of legs, then molts successively until the adult complement of nine is attained.

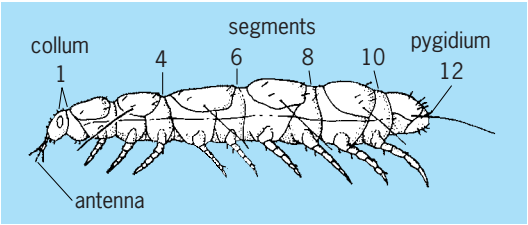

***Pauropus silvaticus*, fully extended adult, enlarged. (*After R. E. Snodgrass, A Textbook of Arthropod Anatomy, Cornell University Press, 1952*)**

The class currently consists of four families; there are over 500 species known. *See* MANDIBULATA.

Ralph E. Crabill

Pavement

An artificial surface laid over the ground to facilitate travel. In this article only road pavements are discussed. The engineering involved is closely similar to that for airport surfacing, another major type of pavement. A pavement's ability to support loads depends primarily upon the magnitude of the load, how often it is applied, the supporting power of the soil underneath, and the type and thickness of the pavement structure. Before the necessary thickness of a pavement can be calculated, the volume, type, and weight of the traffic (the traffic load) and the physical characteristics of the underlying soil must be determined. *See* AIRPORT ENGINEERING.

Traffic load. Traffic data can readily be obtained. Traffic is counted to learn the total volume and the proportion of heavy vehicles. Loadometer scales are placed next to the road and trucks are weighed, front and rear wheels separately. Traffic trends are studied. Such data provide a basis for estimating the total volume of traffic to be carried by a pavement during its service life. They also permit an estimate of the magnitude and frequency of the expected load.

Base and surface courses of pavements are designed to withstand many applications of load over a prolonged period, in some cases up to 30 years. In structural design it is also necessary to give consideration to the direction of traffic. For example, a pavement from a mine to a nearby railroad siding may carry a high percentage of heavily loaded trucks on the inbound lanes and a low percentage on the return lanes.

In general, the larger the volume of heavy vehicles on a highway, the greater the structural capacity required in the pavement. But equally as important as the volume of heavy vehicles are the magnitude of the applied loads and the conditions that will influence the effect of those loads on the pavement. Under the action of vehicular traffic, the surface of a pavement is subjected to a series of highly concentrated forces applied through the wheels of the vehicle. These forces exert an influence throughout the depth of the pavement.

The applied loads vary considerably depending upon the number and spacing of the wheels of each vehicle, the gross weight of the vehicle, and the distribution of that weight among the axles. Two vehicles of the same gross weight may differ widely in the wheel loads applied to the highway. A relatively small truck may cause a load of higher unit stress than a larger vehicle with larger tires and more axles. Consideration of the actual wheel loads has become increasingly important with the large increase in the volume of heavy vehicles on most highways.

All highways, regardless of their design, have some limit to their ability to support the frequent

application of a heavy load. A large number of load applications can be supported by a given pavement if the load does not exceed a particular magnitude, but once this magnitude is exceeded, distress and failure of the pavement becomes increasingly evident. Weather also influences the ability of the pavement to support a load. Definite load limitations are imposed by law on many highways. Further limitations are often imposed on specific highways of lighter design, especially during such conditions as spring freezing and thawing.

In the preceding discussion, wheel loads were treated as static loads. For adequate design analysis, the effects of dynamic loads must be evaluated. The vertical force exerted on pavement by a moving wheel may be considered the sum of the static weight of the wheel and the impact or dynamic force from the wheel's movement over irregularities in the pavement surface. The static load is a constant factor except as the movement of the vehicle along the highway sways or oscillates the load.

An exceedingly variable factor, the dynamic force generated by a moving wheel depends upon (1) the magnitude of the static wheel load; (2) the operating speed of the vehicle; (3) the type, size, and cushioning properties of the tire equipment; and (4) the smoothness of the pavement. An increase in static wheel load or pavement roughness, a decrease in the cushioning qualities of the tires, and, within limits, an increase in vehicle speed all result in increased dynamic forces.

The importance of the foregoing variables and factors has long been recognized by design engineers. The difficulty has been to express the data in terms that could be rationally applied to design formula and then correlated with the performance of foundation soils and materials used to construct the pavement. The problem has not been given a rigorous mathematical solution but rather has depended largely upon field observations under actual operating conditions.

Several methods of load evaluation have been developed and used by various highway agencies and organizations. All of the methods are more or less empirical in approach and thus are subject to revision and adjustment when field observations indicate changing conditions of traffic, climate, or soil performance. No universal method of load evaluation has been developed so far and indeed cannot be until a method is devised that can be readily adjusted when other factors influencing the performance of a pavement structure change.

Methods of evaluation in use include (1) numerical count method, in which the number of vehicles using a particular highway is actually counted and the weight of various vehicles listed as light, medium, heavy, and extra heavy; (2) wheel load method, in which factors based upon the actual weight of the wheel load are used; (3) load frequency method, in which the wheel load weights are combined with the volumetric count of the commercial vehicles; and (4) equivalent wheel load method, in which destructive effects of the actual wheel loads being applied to a pavement are expressed in terms of a standard wheel load. *See* LOADS, DYNAMIC.

Evaluation of subgrade. Factors that must be considered in evaluating the ability of the underlying soil or subgrade to support the pavement include (1) type of soil, such as loam or clay; (2) gradation and variation in particle size; (3) strength or bearing value; (4) modulus of deformation; and (5) swell or volume change characteristics and related properties. Measurement of the supporting power of the subgrade presents numerous difficulties. Tests sometimes used include the plate-bearing test, the direct shear method, the triaxial compression test, and the bearing ratio procedure. *See* SOIL MECHANICS.

Some soil types are unsuitable for supporting pavements because they have low bearing values or undergo changes in volume with variation in moisture content. For example, it is desirable to excavate peat and muck along the road and replace it with soil of higher bearing value.

Rigid and flexible pavements. Once the grading operation has been completed and the subgrade compacted, construction of the pavement can begin. Pavements are either flexible or rigid. Flexible pavements have less resistance to bending than do rigid pavements. Both types can be designed to withstand heavy traffic. Selection of the type of pavement depends among other things, upon (1) estimated construction costs; (2) experience of the highway agency doing the work with each of the two types; (3) availability of contractors experienced in building each type; (4) anticipated yearly maintenance costs; and (5) experience of the owner in maintenance of each type.

Flexible pavements. Flexible paving mixtures are composed of aggregate (sand, gravel, or crushed stone) and bituminous material. The latter consists of asphalt products, which are obtained from natural asphalt products or are produced from petroleum; and tar products, which are secured in the manufacture of gas or coke from bituminous coal or in the manufacture of carbureted water-gas from petroleum

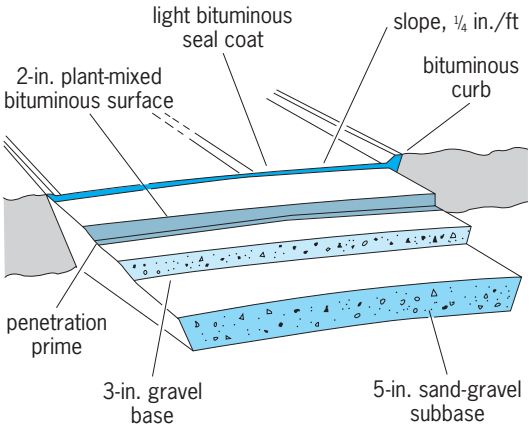

Fig. 1. Flexible pavement design for a city collector street with maximum traffic load of 5 tons (4.5 metric tons) per axle. Right-of-way is 60 ft (18 m) wide and the pavement width is 38 ft (12 m). Berms or boulevards at the sides are sloped in order to drain toward the street. 1 in. = 2.5 cm; 14 in./ft = 2 cm/m.

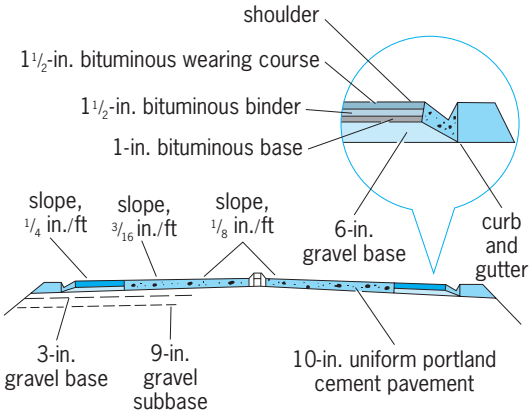

Fig. 2. Rigid main roadway with flexible shoulders. The design is for maximum loads of 9 tons (8 metric tons) per axle. 1 in. = 2.5 cm. 1 in./ft = 8 cm/m.

distillates. Structural strength of a bituminous pavement is almost wholly dependent upon the aggregate, which constitutes a high percentage of the volume of the mixture and forms the structure that carries the wheel load stresses to the base layers. The bituminous material cements the aggregate particles into a compact mass with enough plasticity to absorb shock and jar; it also fills the voids in the aggregate, waterproofing the pavement.

Among the many types of bituminous surface that are used are surface treatments, penetration macadams, road mixes, and plant mixes, as well as variations of these. With surface treatment, a liquid bituminous material is applied over a previously prepared aggregate base.

In building penetration-macadam pavement, a base, usually of crushed stone or gravel, is constructed in layers and firmly compacted by rolling. Often during the rolling, water is applied to make what is termed a water-bound base, or macadam. Then, the keyed and wedged fragments in the upper portion of the base are bonded in place by working alternate applications of bituminous material and choke stone into the surface voids.

With road-mix designs a base is constructed, and a layer of aggregate and bituminous material, mixed on the road with a motor grader, is then uniformly spread and compacted with rollers.

When a plant-mix design is used, construction of a base is also necessary, but the aggregate and bituminous materials are mixed at a central plant, trucked to the job, and then spread or placed with a paving machine and compacted. With the plant-mix procedure more accurate mixing is possible.

For flexible pavements designed to support heavy loads a subbase built of materials similar to those

Fig. 3. A rigid pavement reinforced by welded steel wire mesh. Underneath the mesh are 6 in. (15 cm) of concrete. A 3-in. (8-cm) layer of concrete is being laid on top.

of the base but of poorer quality may also be used. Thickness of the wearing surface, the base, and the subbase depends upon the design load. A typical flexible pavement design for light loads or a 5-ton (4.5 metric ton) axle loading is shown in **Fig. 1**. Where a flexible pavement is designed to carry a large traffic volume, for example, 2000–5000 vehicles per day, including 150–300 heavy commercial vehicles, the gravel subbase would be increased to 10 in. (2.5 cm), the gravel base to 5 in. (13 cm), and the plant-mixed bituminous surfacing to 4 in. (10 cm). It is assumed that the maximum wheel load would be restricted to 9 tons (8 metric tons) per axle. Some flexible pavement designs have called for mixing asphalt with the gravel base material to provide greater pavement strength. Bases treated in this manner are termed bituminous stabilized bases.

Rigid pavements. Coarse aggregate, fine aggregate, and portland cement are mixed with clean acid-free water to produce the concrete used for rigid pavements. The coarse aggregate may consist of coarse gravel or crushed stone and the fine aggregate of sand or crushed-stone screenings. The thickness of the pavement slab may vary from 6 in. (15 cm) for light traffic to 18 in. (45 cm) or more for airport pavements accommodating heavy aircraft. A layer of granular material such as sand, sandy gravel, or slag is generally used as a subbase under the concrete slab to prevent frost heave and to increase the supporting power of the underlying soil. A rigid pavement is shown in cross section in **Fig. 2**.

No steel reinforcement is used with bituminous or flexible pavements. With rigid pavements, especially those designed for heavy loads, reinforcement is often used to strengthen the pavement and to prevent cracking. The reinforcement may consist of welded wire fabric or bar mats assembled by tying transverse and longitudinal steel rods together at their point of intersection (**Fig. 3**). The reinforcement is usually placed about 2 in. (5 cm) below the upper surface of the concrete slab. *See* CONCRETE.

In constructing rigid pavement a longitudinal control joint is used between adjacent lanes. Transverse joints, such as expansion and contraction joints to prevent cracking of the pavement when the temperature changes, may also be included. With flexible pavement no joints are used. *See* HIGHWAY ENGINEERING.

Archie N. Carter

Bibliography. E. R. Brown, *New Pavement Materials*, 1988; S. Brown (ed.), *Pavement Management Systems*, 1993; D. Croney and P. Croney, *Design and Performance of Road Pavements*, 3d ed., 1997; R. Haas, W. R. Hudson, and J. P. Zaniewski, *Modern Pavement Management*, 2d ed., 1994; Y. H. Huang, *Pavement Analysis and Design*, 2d ed., 2003; A. F. Stock (ed.), *Concrete Pavements*, 1988.

Pawl

The driving link or holding link of a ratchet mechanism, also called a click or detent. In **Fig. 1** the driving pawl at *A*, forced upward by lever *B*, engages

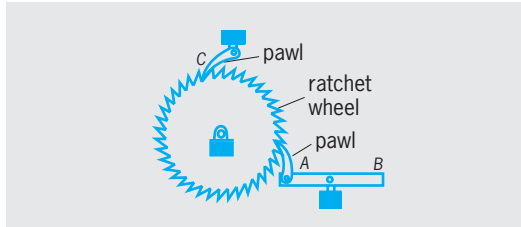

Fig. 1. Holding and driving pawls with a ratchet wheel.

the teeth of the ratchet wheel and rotates it counterclockwise. Holding pawl *C* prevents clockwise rotation of the wheel when the pawl at *A* is making its return stroke. Pawl and ratchet are an open, upper pair.

Driving and holding pawls likewise engage rack teeth on the plunger of a ratchet lifting jack, such as those supplied with automobiles. A ratchet wheel with a holding pawl only, acting as a safety brake, is fastened to the drum of a capstan, winch, or other powered hoisting device.

A double pawl can drive in either direction (**Fig. 2***a*) or be easily reversed in holding (Fig. 2*b*). A cam pawl (Fig. 2*c*) prevents the wheel from turning clockwise by a wedging action while permitting free counterclockwise rotation. This technique is used in the automobile hill-holder to prevent the vehicle's rolling backward.

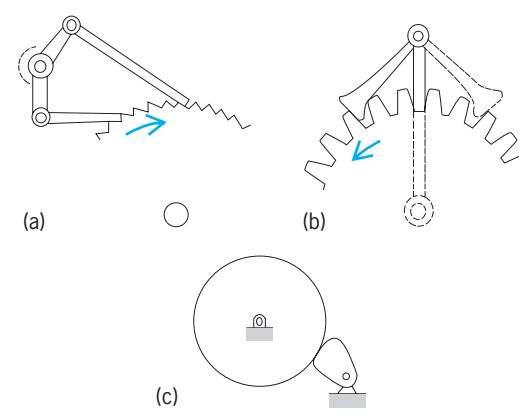

Fig. 2. Some pawl applications. (*a*) Double-acting pawl. (*b*) Reversible pawl. (*c*) Cam pawl.

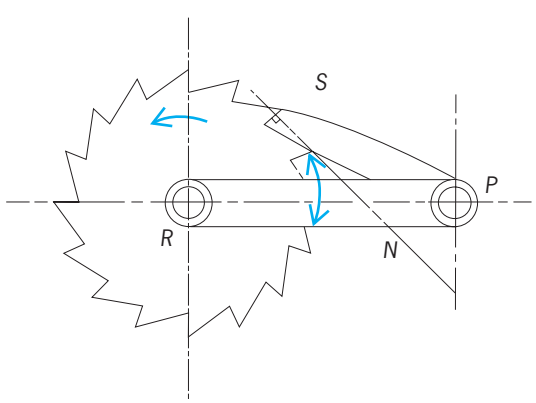

Fig. 3. Essential geometry of pawl design.

The spacing mechanism of a typewriter, although frequently termed an escapement, is properly a ratchet device in which a holding pawl is withdrawn from a spring-loaded rack to allow movement of the carriage, while an arresting pawl is introduced momentarily to permit the holding pawl to engage the next tooth of the rack.

In designing a pawl, care should be taken that the line of contact has a normal *N* passing between centers *R* and *P* (**Fig. 3**); otherwise the pawl will ride out of ratchet step *S*. The pallet of an escapement mechanism is closely related to the pawl. *See* ESCAPEMENT; RATCHET. Douglas P. Adams

Bibliography. H. H. Mabie and F. W. Ocvirk, *Mechanisms and Dynamics of Machinery*, 4th ed., 1987; W. J. Patton, *Kinematics*, 1979; C. E. Wilson et al., *Kinematics and Dynamics of Machines*, 2d ed., 1992.

Paxillosida

An order of sea stars and members of the class Asteroidea. The name is derived from the club-shaped plates, or paxillae, that form the sea star's upper skeletal surface and that have tiny spinelets or granules covering their tips (see **illus.**). Paxillosida encompasses six families, the largest being the Astropectinidae, Luidiidae, and Porcellanasteridae. The Ctenodiscidae, Goniopectinidae, and Radiasteridae are represented by comparatively few members. Astropectinids and luidiids are primarily predators of mollusks and other echinoderms. The former are found over a wide range of depths, whereas the latter live in relatively shallow water. Porcellanasterids are deep-water asteroids that swallow sediment in bulk as they bury themselves to a level just below the surface.

The geologic range of astropectinids extends from Jurassic to Recent; luidiids are known from Miocene to Recent. The Ctenodiscidae are primarily Recent, but a single Cretaceous specimen is known. The other families have not yet been recognized from the fossil record. Asteroids are not easily preserved, and so paxillosidans almost certainly were more common in the past than their fossil record suggests.

In addition to the presence of paxillae, paxillosidans are characterized by a number of unusual features that have been thought to indicate a primitive phylogenetic position among living asteroids. Although the tube feet of most asteroids have suckered disks, those of most paxillosidans are pointed. The digestive system of most asteroids is relatively complex and complete, terminating in an anus, whereas that of the paxillosidans is simple, saclike, and lacking an anus in some members. Most asteroids can extrude their stomach during feeding, but that ability is limited in paxillosidans. A brachiolarian larval stage has been recognized in the development of most asteroids, yet that stage is thought to be absent from paxillosidans.

Research has demonstrated that both suckered tube feet and an anus are present initially but are lost during the ontogeny of certain paxillosidans. Some researchers believe a reduced brachiolarian stage can be recognized as well. Loss of structures during development implies evolutionary reduction rather than primitive organization, and such changes probably reflect an evolutionary adaptation to habit and habitat. Paxillosidans live on unconsolidated sandy or muddy substrates, and they feed by ingesting sediment or shells. Because the large sac-like gut holds shells that are too large to be expelled through an anus and that must be ejected through the mouth, the anus was lost from some. Most paxillosidans bury themselves, a unique ability among the asteroids, so that the paxillae protect the aboral surface from sediment. The brachiolaria is an attached stage that is not suited to the soft, shifting substrates typical of paxillosidans, as is also true of suckered, clinging tube feet. Certain predatory paxillosidans feed at dawn and dusk, then bury themselves in the sediment during digestion. The threat of larger predators, such as fish, may have stimulated evolution of the relatively rapid feeding and hiding behavior. *See* ASTEROIDEA; ECHINODERMATA. Daniel B. Blake

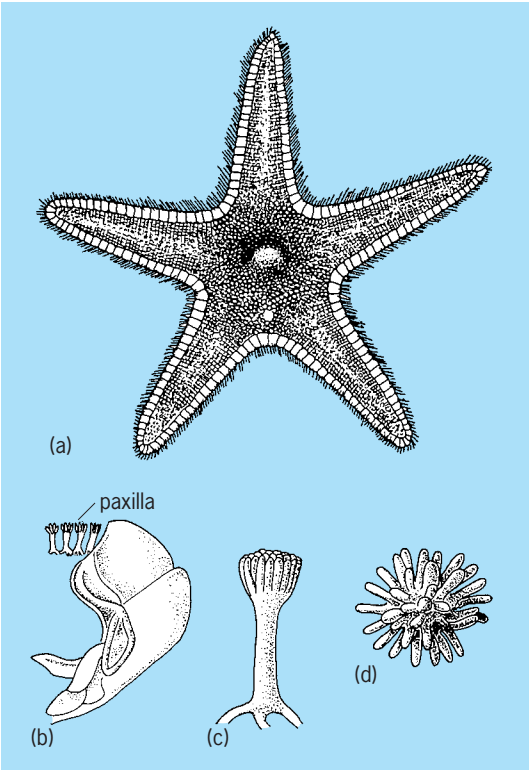

Paxillosida. (*a*) Complete astropecten, a typical paxillosid (*after R. Koehler, Echinodermes: Resultats des Campagnes Scientifiques accomplies sur son yacht par Albert Ier Prince Souverain de Monaco, Imprimerie de Monaco, 1909*). (*b*) Arm arrangement, enlarged to show paxillae; (*c*) paxilla in side view and (*d*) top view (*after W. K. Fisher, Starfishes of the Philippine Seas and Adjacent Waters, Bull. U.S. Nat. Mus. 110, vol. 3, 1919*).

Bibliography. D. B. Blake, A classification and phylogeny of post-Paleozoic sea stars (Asteroidea: Echinodermata), *J. Nat. Hist.*, 21:481–528, 1987; D. B. Blake, Somasteroidea, Asteroidea, and the affinities of *Luidia* (*Platasterias*) *latiradiata*, *Palaeontology*,

25:167–191, 1982; R. D. Burke et al. (eds.), *Echinoderm Biology*, 1988.

Pea

The pea is one of the oldest cultivated crops. It is a native to western Asia from the Mediterranean Sea to the Himalaya Mountains. It appears to have been carried to Europe as early as the time of the lake dwellers of prehistoric times. Peas were introduced into China from Persia about A.D. 400; they were introduced into the United States in very early Colonial days.

Description. Garden peas (*Pisum sativum*; **Fig. 1**) have wrinkled seed coats at maturity when dry; field peas (*P. arvense*) have a smooth seed coat. Both types are annual leafy plants with stems 1–5 ft (0.3–1.5 m) long. Each leaf bears three pairs of leaflets and ends in a slender tendril. The blossoms are reddish-purple, pink, or white, with two or three on each flower stalk. Five to nine round seeds are enclosed in a pod about 3 in. (8 cm) long. Seed color varies from white to cream, green, yellow, or brown. Smooth-seeded varieties may be harvested fresh for freezing or canning, or harvested dry as edible peas. Dry peas may be split or ground and prepared in various ways, such as for split-pea soup. *See* LEGUME.

Culture. Peas require a cool growing season. They should be sown in early spring as soon as a fine, firm seedbed can be prepared. Early-seeded peas develop before the heat of early summer can harm them and usually produce higher yields. Peas are seeded 2–3 in. (5–8 cm) deep. Then the fields are usually rolled immediately to firm the seedbed and aid germination. In the Palouse area of Washington and Idaho, peas are usually planted in early April, while elsewhere irrigated fields are often seeded in March. Garden peas may be seeded as early as February in home gardens in the southern half of the United States.

Fig. 1. Typical wrinkled-seed type of pea, showing fruit (pod) containing seeds (peas). (*Dumas Seed Co., Inc.*)

Applications of lime, sulfur, molybdenum, phosphorus, and potassium are commonly made for pea production.

Harvesting. The peas are mowed and windrowed with a swather. Self-propelled viners, which have replaced stationary viners, pick up the windrowed crop and separate the fresh green peas from the vines and pods. To maintain high quality, this must be done within 2 h after the peas have been mowed. *See* AGRICULTURAL SOIL AND CROP PRACTICES.

A tenderometer is used to determine when green peas for processing should be harvested. Dry peas are harvested with a regular grain combine when the moisture content in the seed drops below 14%. Adjustments are made to minimize damage to the pea seed.

The Grain Division of the Agricultural Marketing Service, U.S. Department of Agriculture, has set standards and issues a grade certificate to define the relative market value of different-quality dry peas.

Kenneth J. Morrison

Production. Wisconsin, Washington, Minnesota, Oregon, Illinois, New York, Pennsylvania, Utah, and Idaho are important in the production of peas harvested green. In dry- and seed-pea production the leading states are Washington, Idaho, Oregon, California, Colorado, Montana, North Dakota, Wyoming, and Minnesota. Seed for all garden varieties is grown primarily in the same areas that grow the dry commercial peas. Washington and Idaho produce most of the total seed and dry commercial pea crop. The two main export markets for dry peas are Europe and Latin America. The European market is somewhat sporadic, depending on the total production of Europe's usually large acreage of peas.

William A. Beachell

Diseases. Diseases are often major limiting factors to acceptable yields in many pea production areas. Causes of major pea diseases include bacteria, fungi, and viruses; nonparasitic problems and nematode-induced diseases are generally of less consequence.

Bacterial diseases. Bacterial blight of pea is caused primarily by *Pseudomonas pisi*, occasionally by *P. syringae*. Symptoms produced by these seed- and rain-borne pathogens are very similar, consisting first of small, often angular, water-soaked spots on leaves (**Fig. 2**) and stipules. When these lesions are numerous, they often coalesce into large, brown, dead areas covering major portions of the foliage, which then dries prematurely to give a blighted appearance. Pods develop water-soaked, slightly sunken lesions that turn brown in 6–10 days. Severely diseased plants have peas which are greatly reduced in yield and quality.

Fungal diseases. The main fungal diseases of peas are root rots, wilts, foliage blights, and mildews. Root rots are the major problem, especially in the United States' most important pea-producing area—the Midwest. They may be caused by *Aphanomyces*

euteiches, *Fusarium solani* f. sp. *pisi*, *Pythium* spp., *Ascochyta pinodella*, and *Rhizoctonia solani*. *Aphanomyces euteiches* causes greater economic loss than all of the others combined. This is because no effective control has been found, even though the Aphanomyces root rot problem has existed since 1925. Warm and wet soils favor the disease. Highly infested fields can be identified by laboratory techniques and can be eliminated from production. These root rot pathogens live naturally in the soil and in the roots of crops other than peas; thus rotation is of limited benefit, and root rots are a recurrent problem. Generally roots show various amounts of brownish discoloration and decay. When severe, the root system may be largely destroyed (**Fig. 3**) and the stunted plants turn yellow and die prematurely.

Two widespread wilt diseases of pea are caused by *Fusarium oxysporum* f. sp. *pisi* race 1 and race 2. Race 5 of the same fungus causes a locally important wilt in Washington. These soil organisms become established in bothersome numbers due to too many pea crops. They attack the roots and grow through the water-conducting vessels up into the stem, interfering with the passage of water to the leaves. Affected plants wilt, sometimes unilaterally (**Fig. 4**), the lower leaves turn yellow, and the plants die in an early stage of development or soon after flowering.

Fig. 2. Bacterial blight lesions on pea leaf.

Fig. 3. Aphanomyces root rot of pea (right); healthy root (left).

Fig. 4. Near-wilt of pea showing unilateral wilt.

Fungus-induced foliage blights include four diseases of variable occurrence and severity, most of them being encouraged by contaminated seed, too many consecutive pea crops, and rainy weather. Ascochyta leaf and pod spot is caused by *Ascochyta pisi*. A similar disease, Mycosphaerella blight, is caused by *Mycosphaerella pinodes*. *Colletotrichum pisi* causes pea anthracnose. All three pathogens cause reddish-brown spots of various sizes on the leaves, stipules, and pods. When leaf lesions are numerous, leaves shrivel and die and the plant appears blighted. Septoria blotch is caused by *Septoria pisi*, and is found primarily on lower, somewhat senescent leaves and stipules. It is characterized by large, yellow to straw-colored areas with indefinite margins.

Powdery mildew is widespread and often quite troublesome on late-maturing plants. It is caused

by *Erysiphe polygoni*, and appears as a superficial whitish "powder" on foliage and pods. Dry weather favors its development. Downy mildew, caused by *Peronospora pisi*, produces tan spots of irregular size and shape on the upper leaf surfaces. A mouse-gray, cottony growth of the fungus develops on the undersides of the foliage immediately beneath the tan spots.

Viral disease. Viral diseases of peas are widespread and common, but are generally of minor consequence. Pea enation mosaic and pea streak are the most important, but pea stunt and seed-borne mosaic also occur. The causal viruses are aphid-borne, often from nearby perennial legumes. *See* PLANT VIRUSES AND VIROIDS.

Disease control. Many of the pea diseases mentioned can be controlled through the use of resistant cultivars. For example, there are cultivars resistant to one or more of the three pea wilt diseases, Ascochyta leaf and pod spot, powdery mildew, downy mildew, and both pea mosaics. Use of disease-free seed, crop rotation, field sanitation, and strict aphid control are helpful disease control measures. Chemical seed protectants are important, and fungicides may sometimes be needed for powdery mildew control. *See* PLANT PATHOLOGY.

D. J. Hagedorn

Bibliography. D. J. Hagedorn, *Compendium of Pea Diseases*, 1984.

Peach

A deciduous fruit tree species (*Prunus persica*) that originated and was first cultivated in western China. It is adapted to relatively moderate climates in the temperate zone. Although most peach cultivars require a substantial amount of winter chilling (temperatures between 32 and 45°F, or 0–7°C) to ensure adequate breaking of winter dormancy and uniform budbreak, peach wood is susceptible to winter injury at temperatures below −15°F (−25°C) and dormant fruit buds are injured by temperatures below 0°F (−18°C). Consequently, commercial cultivation is limited to lower latitudes in the temperate zone or to higher latitudes where large bodies of water have a moderating influence on climate. The principal peach-growing regions in North America, ranked in order of commercial production, are central California, Georgia and the Carolinas, the mid-Atlantic region, the Great Lakes region, and the Pacific northwestern region. Other important peach-growing regions in the world include Italy, southern France, Spain, Japan, China, Argentina, southern Brazil, Chile, South Africa, and southeastern Australia. *See* ROSALES.

Cultivars. Peach cultivars can vary greatly and are usually distinguished by their fruit types. Peach fruits are covered with short epidermal trichomes called fuzz (smooth-skinned peaches are called nectarines) and at maturity are usually yellow or white with a red blush. The internal flesh is also yellow or white. Clingstone cultivars have a relatively firm flesh that adheres to the pit at maturity, and are primarily used for canning. Some principal North American clingstone canning cultivars are Halford, Loadel, Andross, and Carolyn.

Freestones usually have a softer flesh that separates from the pit at fruit maturity, and are primarily used for the fresh market, freezing, and drying. There are more than 150 freestone peach cultivars grown commercially in the United States. The more important western varieties are Springcrest, O'Henry, Fay Elberta, and June Lady. Redhaven, Blake, and Redglobe are important eastern varieties, and June Gold is an important low-chilling variety grown in warm southern locations.

Genetic dwarf peach trees occur naturally; though there have been attempts to develop dwarf peach tree cultivars, none is commercially acceptable.

Propagation and cultivation. Peach cultivars are clonally propagated primarily by budding onto selected rootstocks. The most common rootstock in most of the United States is the Lovell cultivar; however, Nemaguard is commonly used in central California where soil nematodes are a problem, and Halford is often used in the Great Lakes region. Marianna and Myrobolan plum rootstocks are also used commercially. Peach trees grow best on well-drained sandy soils or gravelly loam soils but can also tolerate heavier, clay loams. Traditionally, trees are planted 16–22 ft (5–7 m) apart and trained to an open vase-shaped form with three to five main scaffold branches originating from the trunk at approximately 18–24 in. (0.5–0.8 m) from the ground. There has been a trend to plant trees at higher densities and in hedgerows, but these planting systems have been hampered by the excessive vigor of peach tree growth. Fruit is borne almost exclusively on 1-year-old extension shoots. Trees require extensive pruning to prevent overloading, maintain tree size and shape, and promote regrowth of new fruiting wood in lower areas of the tree. Most peach cultivars are self-pollinating, often set large numbers of fruit, and require fruit thinning by hand or with poles.

Harvesting. Fruit is harvested when its dark-green external coloring changes to a pale yellow or white. In some regions commercial fruit harvest extends from late spring to early fall. Mechanical harvesting is used on some processed fruit, but fruit for fresh market is hand-picked.

Theodore M. De Jong

Diseases. Brown rot, caused by *Monilinia fructicola* and *M. laxa*, is a destructive fungus disease of the peach throughout the world. Blossoms, twigs, and fruit are infected. Major loss is from the decay of the fruit, which is converted into a soggy mass unfit for human consumption. Control is achieved by spraying or dusting the trees at regular intervals with powdered sulfur or with captan, an organic fungicide. For best results it is essential also to control the plum curculio, a common fruit insect, whose punctures facilitate entrance of the fungus into the fruit. *See* COLEOPTERA; FUNGISTAT AND FUNGICIDE.

In contrast to brown rot, which requires strenuous efforts for its control, peach scab (*Cladosporium carpophilum*), although universally present, is

readily controlled by one or two sprays of sulfur shortly after the blossom petals drop. Leaf curl caused by *Taphrina deformans*, a leaf-distorting fungus, can be prevented by one spray of lime sulfur, bordeaux mixture, or ferbam applied before the buds begin to swell.

Bacterial spot, caused by *Xanthomonas pruni*, is a serious disease of peaches in the United States, China, Japan, and New Zealand. The bacteria kill small groups of cells in the leaves, twigs, and fruit. Infected leaves drop prematurely and this devitalizes the trees mainly through reduced photosynthesis, the most serious long-time effect of the disease. Diseased fruit is edible, but its appearance is marred and its market value reduced. The bacteria survive the winter in small cankers formed on the twigs. Control is difficult because of the prolonged infection period. A mixture of zinc sulfate and hydrated lime, applied as a spray at intervals of 10–14 days, reduces the severity of the infection in most years.

Other fungus diseases affecting peaches are rust caused by *Tranzschelia discolor*, peach blight caused by *Coryneum beijerinckii*, and mildew caused by *Sphaerotheca pannosa*. These diseases cause serious losses in various parts of the world. *Rhizopus nigricans*, occurring throughout the world, causes rapid decay of harvested and stored peaches.

Among the more than 20 virus diseases known to affect peaches in the United States, yellows, rosette, phony peach, and mosaic are the most serious. The first two kill the trees in a few years, whereas phony peach and mosaic reduce both quality and quantity of fruit, and the trees eventually become worthless. *See* PLANT VIRUSES AND VIROIDS.

Peaches are also injuriously affected by nutrient deficiencies in the soil, the symptoms of which often resemble those resulting from disease organisms. Fluctuations in temperature during the winter, particularly in the southern states, frequently kill many peach trees. *See* FRUIT, TREE; PLANT PATHOLOGY.

John C. Dunegan

Bibliography. W. H. Chandler, *Deciduous Orchards*, 1942; N. F. Childers (ed.), *The Peach*, 1975; J. Janick and J. N. Moore (eds.), *Advances in Fruit Breeding*, 1975; M. N. Westwood, *Temperate Zone Pomology*, 3d ed., 1993.

Peanut

A self-pollinated, one- to six-seeded legume. It is cultivated throughout the tropical and temperate climates of the world. The oil, expressed from the seed, is of high quality, and a large percentage of the annual world production is used for this purpose. In the United States some 65% goes into the cleaned and shelled trade, the end products of which are roasted or salted peanuts, peanut butter, and confections. *See* LEGUME; ROSALES.

Origin and Description

Peanuts originated in Bolivia and northeastern Argentina where a large number of wild forms are found. The cultivated species, *Arachis hypogaea*, was grown extensively by Indians in pre-Columbian times. Merchant ships carried seed to many continents during the early part of the sixteenth century. Although grown in Mexico before the discovery of America, the peanut was introduced in the United States from Africa.

Botanically, peanuts may be divided into three main types, Virginia, Spanish, and Valencia, based on branching order and pattern and the number of seeds per pod. Marketing standards in the United States include grades for Virginia (large-seeded, grown mostly in North Carolina and Virginia), runner (small-seeded Virginia, grown mostly in Georgia, Florida, and Alabama, with some in Texas and Oklahoma), Spanish (grown mostly in Oklahoma and Texas), and Valencia (grown mostly in New Mexico). *See* SEED.

The peanut's most distinguishing characteristic is the yellow flower, which resembles a butterfly (papilionaceous) and is borne above ground. Following fertilization the flower wilts and, after 5–7 days, a positively geotropic (curving earthward) peg or ovary emerges. Penetrating the soil 0.8–2.8 in. (2–7 cm), the peg assumes a horizontal position and the pod begins to form (**Fig. 1**). *See* FLOWER; PLANT MOVEMENTS.

The pod, a one-loculed legume, splits under pressure along a longitudinal ventral suture. Pod size varies from 0.4 by 0.2 in. (1 by 0.5 cm) to 0.4 by 3.2 in. (2 by 8 cm), and seed weight varies from 0.08 to 0.2 oz (0.2 to 5 g). The number of seeds per pod usually is two in the Virginia and runner type, two or three in the Spanish, and three to six in the Valencia.

The plant may be upright, prostrate, or intermediate between these forms. The main stem is usually upright and may be very short in some cultivars. The leaves are even-pinnate with four obovate to elliptic leaflets. Leaves occur alternately and have a 2:5 phyllotaxy (arrangement on stems). *See* LEAF.

Peanuts are harvested by running a special wing-type plow under the plants. After wilting they are either stacked or allowed to cure in windrows before picking. The main production areas in the United States extend southward from Virginia and

Fig. 1. A typical Virginia-type peanut.

westward to Oklahoma and Texas. Peanuts are under strict acreage controls. *See* AGRICULTURAL MACHINERY; AGRICULTURAL SCIENCE (PLANT); AGRICULTURAL SOIL AND CROP PRACTICES. Astor Perry

Diseases

The peanut plant is susceptible to 46 described diseases that reduce yield of fruit (pods) at least 30%. Causal agents include fungi, nematodes, viruses, bacteria, and insects. Yield of pods is reduced primarily by premature loss of leaves and decay of stems, pods, and pegs. For economical production, disease control measures used by growers involve pesticides, crop rotation, deep burial of surface organic matter, and disease-resistant cultivars.

Fungal diseases. Fungi cause the greatest number of diseases and are of greatest concern to growers. Peanut seed harbors fungi that cause rots, poor germination, and seedling diseases. The most common are species of *Rhizopus* and *Aspergillus*, and consequently remedial procedures must be used. Seed is relatively free of fungi when pods are harvested, but becomes infected because of mechanical injury incurred during the shelling operation and improper handling. It is a common and necessary practice to treat seed prior to planting, and if seed is not coated with a broad-spectrum fungicide serious reduction in stand will result. Leafspot diseases caused by *Cercospora arachidicola* and *Cercosporidium personatum* are the most serious diseases that cause premature leaf loss wherever peanuts are grown (**Fig. 2**). Annual crop losses range from 15 to 50% in yield of pods in many areas. Infection begins on lower leaves near the soil, and movement of fungal units (inoculum) to leaves and subsequent secondary spread of the fungi (conidia) is by wind, rain, insects, or machinery. Foliage is sprayed with fungicides before infection and at 7–14-day intervals through the season for optimum suppression of diseases. The most intensive use of foliar fungicides is in the southeastern United States, where six or seven applications are made each season. Use of crop rotation where peanuts are grown only one out of three years significantly reduces severity of leaf spots. Fungicides used for suppression of leaf spots are also effective against rust caused by *Puccinia arachidis*, a very serious threat to production. The rust fungus is endemic in the West Indies, and has become established throughout Asia and Oceania in Australia and Africa. Web blotch caused by *Phoma arachicola* is widely distributed and may cause significant damage. *See* FUNGISTAT AND FUNGICIDE.

Fig. 2. Blue damage of seed coat and cotyledon (*D. C. Norton*)

Diseases of subterranean plant parts are the most destructive and difficult to control, and growers must depend primarily on cultural practices for disease suppression. Southern blight, caused by *Sclerotium rolfsii*, is the most devastating and occurs in all peanut-growing areas. The fungus attacks succulent tissues at the ground line and either kills plants or causes pod and root decay. Because the fungus depends on organic matter for growth and subsequent invasion, growers remove organic litter from soil surface by deep plowing and plant on flat or raised beds. During cultivation soil is kept away from planting rows so that sclerotia (survival propagules) and organic litter will not be moved to plants and initiate infection. Sometimes a soil fungicide and a nematicide are applied over the row 45–50 days after planting to control southern blight. Other fungi that are widely distributed and cause decay by acting singly or in combination include *Rhizoctonia* spp., *Pythium* spp., and *Fusarium* spp. *Cylindrocladium* spp. and *Sclerotinia* spp. cause sporadic losses in the eastern United States. A pod rot initiated by calcium deficiency is prevented by application of gypsum after flowering.

Nematodes. Plant parasitic nematodes that feed on roots of plants occur in all peanut-growing areas of the world. Most damaging are the northern root-knot nematode (*Meloidogyne hapla*) and the peanut root-knot nematode (*M. arenaria*) which feed on the inside of root, peg, and fruit and cause galls on these parts. Sting nematodes (*Belonolaimus gracilis*) are very destructive and are known to occur only in the lighter soils of the United States. The sting nematodes feed on the outside of the peanut root, peg, and fruit and cause a restricted root system and smaller fruit. The lesion nematode (*Pratylenchus brachyurus*) feeds on the inside of the peanut root, peg, and fruit and causes necrotic lesions on these parts. Nematodes cause damage directly by feeding on roots and other underground plant parts and indirectly by interacting with soil-borne fungi to increase decay. If populations are sufficiently high, nematicides are applied to soil to reduce damage. *See* NEMATA (NEMATODA).

Viral diseases. There are five described virus diseases of peanuts—rosette, mosaic, ringspot, stunt, and mottle. Of these, rosette and stunt are the most significant to growers.

Bacterial diseases. *Pseudomonas solanacearum*, the cause of bacterial wilt, occurs in all peanut-growing areas but causes yield losses only occasionally. Most of the peanut cultivars grown currently in the United States are resistant to bacterial wilt.

Insects. The Southern corn root worm (*Diabroticia undecimpunctata howardii*) interacts with soilborne fungi and bacteria to cause fruit rot. The potato leaf hopper (*Empoasca fabae*) secretes a toxic substance on the leaves and causes a disorder known as hopper burn. The tobacco thrip (*Frankliniella fusca*) causes puckered leaflets and retards the growth of seedlings; however, reduction in yield of pods is minimal. *See* PLANT PATHOLOGY.

R. H. Littrell

Processing

Most peanuts are shelled and processed into peanut butter, salted peanuts, and confections. A few are boiled in 6% salt brine, roasted in shell, salted in shell, exported, or crushed for oil. Outside the United States, peanut oil is the most common end product. It is well suited for deep fat frying and stir frying because it has a high smoke point (446°F or 230°C) and a lack of flavor carry-over. *See* FAT AND OIL (FOOD); SOLVENT EXTRACTION.

Factors affecting flavor and quality. Peanuts must be carefully handled throughout growing, harvesting, curing, storage, shelling, subsequent storage, transportation, processing, packaging, and storage of finished products. Improper handling will produce an unacceptable consumer product with poor flavors. Good roasted peanuts have a slightly sweet taste with a roasted flavor and an appropriate aftertaste. Peanuts are high in oil content (45–50%), and they are easily oxidized to produce off flavors, which cover up good roasted flavor. Some off flavors are bland, stale, fruity, musty, rancid, and beany.

Three major problems face the peanut manufacturers: flavor, foreign material, and aflatoxin. Flavor defects can occur at any stage and can adversely affect product flavor and consumer acceptance. Drought-stressed peanuts have spotted surfaces which are associated with flavor defects. At harvest, not all peanuts are of the same maturity so it is important that they be harvested at optimum maturity to give best flavor, greatest storage stability, and longest product shelf life. Immature peanuts have flavor problems because of high levels of free arginine, threonine, tyrosine, and lysine. Excessive drying also produces off flavors that are associated with acetaldehyde, ethanol, and ethyl acetate. Foreign material includes sticks, stones, small rocks, and glass. A series of screens and blowers will remove much of this material; electronic sorters are also used. Unfortunately, some foreign material may still get into the final product and can cause costly consumer complaints such as broken teeth. Sorting is also used to remove faulty nuts that are discolored and may have off flavors and contain aflatoxins. Aflatoxins are produced by molds such as *Aspergillus flavus* and are associated with peanuts that have become wet or have not been dried properly. United States government regulations minimize the amount of aflatoxins in peanut products. After shelling, cleaning, and aflatoxin testing, peanuts—now ready for processing—are stored in refrigerated warehouses at 32–37°F (0–3°C) with 65% relative humidity to ensure protection against insects, rancidity, and mold growth. *See* AFLATOXIN.

Roasting and oil cooking. Roasting (referred to as dry roasting in the peanut industry) uses hot air (320–446°F or 160–230°C) to develop desirable color, texture, and flavor. An internal temperature of 290°F (143°C) must be reached to develop the roasted peanut flavor and peanutty aftertaste. These hot peanuts must be properly cooled with ambient or refrigerated air to prevent overroasting, which gives burnt flavors. Most of the roasters are of the flat-bed type, but a few gas-fired rotary roasters are used by some smaller processors.

Oil cooking (referred to as oil roasting in the peanut industry) uses hot cooking oils (320–347°F or 160–175°C) to develop the desirable characteristics. Generally, these peanuts have a higher flavor intensity.

Peanut butter. Major peanut butter manufacturers further clean the raw shelled peanuts to remove additional foreign material. Peanuts are roasted, cooled, blanched (to remove the skins which give specks, and to remove the hearts, which impart a grayish color), sorted to remove unblanched and off-colored nuts, ground in a primary mill (usually a plate mill), combined with various ingredients, ground in the secondary mill (usually a cutting mill) to give the desirable particle size, deaerated, cooled, packaged, and then tempered before shipping. Ingredients may include sweeteners, salt, and hydrogenated vegetable oil to improve smoothness, spreadability, and flavor. In the United States, the Food and Drug Administration standards require that peanut butter contain at least 90% peanuts.

There are four basic types of peanut butters. The most common type is molasses flavored (molasses is a sweetener). The second type has roasted peanut flavor with the peanutty aftertaste. The third type has a dark roast and contains some of the burnt flavor with very little roasted and peanutty flavors. The first three types use sweeteners, salt, and stabilizers to prevent oil separation. The fourth is known as the old-fashioned type; it does not contain stabilizers and may not contain sweeteners or salt. The peanut oil separates to the top of the jar and must be remixed before using.

Roasted peanuts. The product known as cocktail peanuts is usually cooked in oil, cooled, and salted. Cocktail peanuts usually have a higher intensity of roasted peanut flavors. Dry roasted peanuts are often coated with gums, sweeteners, and salt and then roasted to make products such as honey roasted peanuts. There are a large number of regional products in this category.

Peanut candies and confections. Wide variations exist in formulations, with peanuts being used as a topping, a component in nut rolls and clusters, coated nuts, peanut brittles, and peanut butter cups (peanut-butter-filled chocolate confections). Peanuts used for candy bars are usually overroasted (a dark roast) so that the chocolate does not cover up the peanut flavor.

In-shell peanuts. These are of two types. There is the so-called roasted-in-shell, which is roasted,

cooled, and packaged. The other type, known as salted-in-shell, is placed in a sodium chloride brine solution under vacuum, dried slowly, roasted, and packaged. *See* FOOD ENGINEERING; FOOD MANUFACTURING.

Clyde T. Young

Bibliography. American Peanut Research and Education Association Inc., *Peanut: Culture and Uses*, 1973; C. R. Jackson and D. K. Bell, *Diseases of Peanut (Groundnut) Caused by Fungi*, Univ. Georgia Coll. Agr. Exp. Sta. Res. Bull. 56, 1969; E. W. Lusas, D. R. Erickson, and W.-K. Nip (eds.), *Food Uses of Whole Oil and Protein Seed*, 1989; H. E. Pattee and C. T. Young, *Peanut Science and Technology*, 1982; D. M. Porter, D. H. Smith, and R. Rodriquez-Kabana (eds.), *Compendium of Peanut Diseases*, 2d ed., 1997; J. G. Woodroof, *Peanuts: Production, Processing, Products*, 3d ed., 1983.

Pear

Any of approximately 20 species of deciduous tree fruits in the genus *Pyrus*. About half of the species are native to Europe, North Africa, and the Middle East around the Mediterrean Sea; the others are native to Asia. Pear culture is documented to have started as early as 1100 B.C. Pears are best adapted to temperate climates with warm, dry summers and cold winters. They require winter cold to break the dormant period but are injured by temperatures below −10 to −15°F (−23 to −26°C). Commercial pear production in the United States is concentrated in the interior valleys of California, Oregon, and Washington.

Commercial types and varieties. Nearly all United States pear production is of the European pear, *P. communis*. The Bartlett variety comprises over 75% of the United States pear crop. Other European pear varieties include d'Anjou, Bosc, Comice, Seckel, and Winter Nelis. The European pear is noted for its soft, juicy flesh. The skin color is medium green to yellow, depending on fruit maturity and the variety. Skin texture can be smooth or rough. Fruit shape ranges from the classic pear shape (round base with narrow neck) to a rounded oblong shape with no clearly defined neck area.

The crisp-fleshed Asian pear, *P. pyrifolia*, is the second most popular type of pear grown worldwide. In the United States, Asian pear production is found primarily in California, with limited commercial production in Oregon and Washington. The Asian pear is characterized by a crisp, juicy flesh that has a gritty texture, and has been referred to as the sand pear. The fruit shape is round, the skin color is generally yellow to amber at maturity, and the skin texture is smooth or corky. Asian pears are sometimes marketed commercially as pear apples because of their flavor and fruit texture; however, they are not a hybrid of apple and pear. The most popular variety of Asian pear grown worldwide is 20th Century. Other commercial varieties include Shinseiki, Hosui, Ya Li, and Kosui. *Pyrus communis* and *P. pyrifolia* hybrids, such as Kieffer and Leconte, have been developed for limited commercial use in the southeastern United States. *Pyrus ussuriensis* has also been selected for Asian pear varieties. The snow pear, *P. nivalis*, is produced in Europe for cider and perry (a fermented liquor).

Propagation and cultivation. The European pear is usually propagated onto seedlings grown from Bartlett or Winter Nelis. Vegetatively propagated *P. communis* rootstocks known as the Old Home × Farmingdale series have been developed for size control and precocity. European pears are also grown on quince (*Cydonia oblonga*) rootstocks, primarily for size control or dwarfing. *Pyrus calleryana* has also been used for European pear rootstock. Soil type and winter hardiness affect the rootstock selection. The Asian pear is not used as a rootstock for European pears because of pear decline, which results from incompatability between the Asian and European pear species. Quince is also incompatible with Bartlett, and interstems of Old Home × Farmingdale are used to alleviate incompatibility symptoms. Pear orchard care and culture are similar to apple. *See* APPLE; QUINCE.

Disease and insect pests. Most commercial pear varieties are susceptible to fireblight, a bacterial disease caused by *Erwinia amylovora*. Fireblight is vectored by insects, such as honeybees, and is spread from the blossoms to the shoot tips and branches. Properly timed sprays of copper or streptomycin provide partial control, but removal of infected branches is always recommended. Resistant varieties are another option, but they generally lack commercial appeal. Fireblight-resistant pear varieties include Harrow Delight, Harvest Queen, and Moonglow. Warm, humid climates favor the development of fireblight, commercial pear production is limited to the western United States. Pear psylla (*Psylla pyricola*) is the major insect pest of pears. To date, chemical control is the preferred method, as there are limited biological control methods. Psylla-resistant varieties have not been developed. *See* PLANT PATHOLOGY.

Harvesting and storage. For optimum eating quality, European pears are picked immature and ripened off the tree. Unlike apples, pears do not continue to ripen in cold storage. Pears are best stored at a temperature of 30 to 32°F (−1 to 0°C) and about 90% humidity. Winter pears, such as d'Anjou, require cold storage for several months to ensure proper ripening after removal from storage. For best dessert quality, pears should be allowed to ripen at 60–70°F (15–21°C) at about 80–90% humidity. Asian pears are picked at the firm-ripe stage, and can be held in cold storage for several months. Asian pears should not soften for optimum eating quality. *See* FRUIT, TREE.

Kathleen M. Williams

Bibliography. M. N. Westwood, *Temperature Zone Pomology*, 1978; K. M. Williams, *Growing Pears in North Carolina*, 1987.

Pearl

The term pearl is applied properly to any mollusk-formed calcareous concretion that displays an orient and is lustrous. There are two major groups of

bivalved mollusks in which gem pearls may form: the salt-water pearl oyster (*Pinctada*), of which there are several species; and a number of genera of fresh-water clams. Usually, jewelers refer to salt-water pearls as Oriental pearls, regardless of their place of discovery, and to those from fresh-water bivalves as fresh-water pearls. *See* GEM.

Formation. Between the body mass and the valves of the mollusk extends a curtainlike tissue called the mantle. Epithelial cells on the side of the mantle toward the shell perform the several stages of the shell-secreting process during the life of the mollusk. One of the stages of shell building is the secretion of nacre, the colorful, lustrous, mother-of-pearl material. In order for a pearl to form, a tiny object such as a parasite or a grain of sand must work through the mantle, carrying with it epithelial cells. When this happens, secretion of nacre around the invading object builds a pearl within the body of the mollusk. Whole pearls form within the body mass of the mollusk, in contrast to blister pearls, which form as protrusions on the inner surface of the shell. Edible oysters produce lusterless concretions, but never pearls.

Pearls occur in a great variety of shapes. The term baroque is used for the common, irregularly shaped forms. The most common and most desirable shape is the spherical or nearly spherical; this is the shape usually chosen for necklaces. Other desirable shapes include those called button, pear, egg, and drop. The colors, more or less in the order of desirability, are rosé, white rosé, light-cream rosé, slightly greenish rosé, white, black, cream, gold, and blue-gray.

Pearls are composed of many tiny overlapping plates of nacreous material. Nacre consists of prismatic pseudohexagonal aragonite crystals (oriented so that the long crystallographic axis is at right angles to the plane of the platelet) held together by conchiolin, a hornlike organic secretion. Chemical analysis of pearls shows calcium carbonate ($CaCO_3$), organic material, and water; the relative quantities vary with the species of the mollusk, the position of the pearl within the shell, and other factors. $CaCO_3$ content is usually 90–92%, but may be somewhat lower. Organic matter usually makes up 4–6% and water 2–4%. *See* ARAGONITE.

Producing areas. The major pearl-producing region in the world is the Persian Gulf. From its pearl fisheries come most of the natural pearls used for gem purposes, but output is now only a fraction of that of a long period during the nineteenth century through the 1920s. Since World War II, the sale of Oriental pearls has never regained earlier peaks. In the Persian Gulf, the important species *Pinctada margaritifera* is found at depths of from about 4 to 8 fathoms (7–14 m) on broad banks that extend for many miles into the gulf from both shores. Since pearl-producing mollusks are only one of many types of mollusk, fishing is a great gamble, and the waste of nonbearing mollusks is enormous. In times of demand, depletion of the mollusk supply is a grave problem. The supply has apparently never fully recovered from the depletion caused by the heavy fishing of the 1920s. Less important pearl sources include the coast of northern Venezuela, to the north of Australia, between Sri Lanka and India, in the Red Sea, and in the South Pacific.

Many of the rivers of the central portion of the United States have pearl-bearing mussels. In the nineteenth century, several rich finds were made that led to rapid exploitation and virtual exhaustion in those regions of the supplies of pearl mollusks; these were principally of the genus *Unio*. Fresh-water pearls vary in color from almost pure white to many that are more strongly tinted than the majority of the salt-water variety.

Cultured pearl. The substitute for natural pearls, to which the name cultured pearl has been given, is usually made by inserting a large bead into a mollusk to be coated with nacre. When the pearl-bearing mollusk (*Pinctada martensii*) reaches maturity at $3^1/_2$ years, the mollusks are gathered and prepared for pearl cultivation using larger nuclei. Smaller beads may be put in younger mollusks.

Workers trained for the task cut a channel into the foot mass or body cavity and insert a large sphere (bead), plus a small section of mantle tissure, with the epithelial cells next to the bead. Ten or more small beads (0.08–0.12 in. or 2–3 mm) may be placed in one mollusk. The number decreases with size. Perhaps two 0.24-in. (6-mm) beads are inserted, but any larger nucleus is placed one to a mollusk. Beads are prepared from large American fresh-water shells. The mollusks are placed in cages suspended from rafts in sheltered bays and usually left for 10 months to $3^1/_2$ years, except for periodic cleaning and inspection. The rate of nacre accretion in Honshu waters is only about 0.0059 in. (0.15 mm) annually, so that the diameter increases about 0.012 in. (0.30) mm annually; thus a 0.28-in. (7.2-mm) cultured pearl usually has a bead center of 0.24 in. (6.0 mm). The rate of accumulation in Kyushu is about double. More rapid nacre accumulation is encountered in South Seas culture stations, which are now producing larger species of *Pinctada*. Baroque fresh-water pearls have been produced without bead nuclei in Japan. *See* BIVALVIA; MOLLUSCA. Richard T. Liddicoat, Jr.

Peat

A dark-brown or black residuum produced by the partial decomposition and disintegration of mosses, sedges, trees, and other plants that grow in marshes and other wet places. Forest-type peat, when buried and subjected to geological influences of pressure and heat, is the natural forerunner of most coal.

Peat may accumulate in depressions such as the coastal and tidal swamps in the Atlantic and Gulf Coast states, in abandoned oxbow lakes where sediments transported from a distance are deposited, and in depressions of glacial origin. Moor peat is formed in relatively elevated, poorly drained moss-covered areas, as in parts of Northern Europe. *See* COAL; HUMUS.

In the United States, where the principal use of peat is for soil improvement, the estimated reserve on an air-dried basis is 13,827,000 short tons (12,544,000 metric tons). In Ireland and Sweden peat is used for domestic and even industrial fuel. In Germany peat is the source of low-grade montan wax. Gilbert H. Cady

Pebble mill

A tumbling mill that grinds or pulverizes materials without contaminating them with iron. Because the pebbles have lower specific gravity than steel balls, the capacity of a given size shell with pebbles is considerably lower than with steel balls. The lower capacity results in lower power consumption. The shell has a nonmetallic lining to further prevent iron contamination, as in pulverizing ceramics or pigments (see **illus.**). Selected hard pieces of the material being ground can be used as pebbles to further prevent contamination. *See* CRUSHING AND PULVERIZING; TUMBLING MILL. Ralph M. Hardgrove

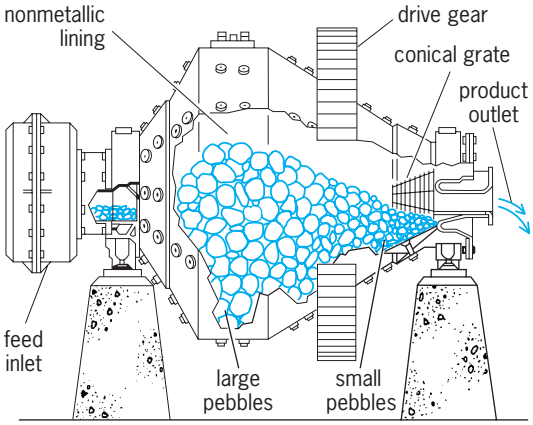

Diagrammatic sketch of a conical pebble mill.

Pecan

A large tree (*Carya illinoensis*) of the family Juglandaceae, and the nut from this tree. Native to valleys of the Mississippi River and tributaries as far north as Iowa, to other streams of Texas, Oklahoma, and northern and central Mexico, this nut tree has become commercially important throughout the southern and southwestern United States and northern Mexico. Limited plantings have been made in other regions, being reported of interest in such diverse areas as Arizona and California, South Africa, Israel, Brazil, and Peru. The tree's major importance is in areas with a long growing season of over 200 days and midsummer average temperatures of 79°F (26°C) or higher, though a few early-ripening cultivars are grown in slightly cooler regions. There are different cultivars for arid, irrigated regions and humid summer regions. The United States cultivars have an indistinct winter chilling requirement but fruit best where the coldest winter month averages less than 61°F (16°C). *See* JUGLANDALES.

Nuts and kernels of the Cherokee pecan cultivar developed by the U.S. Department of Agriculture. (*From Pecan Quarterly*)

Botany. The tree is deciduous with compound, pinnate leaves. Flowers are unisexual with staminate flowers borne on year-old wood and pistillate flowers formed terminally on new spring growth. The period of pollen shedding of a cultivar usually does not overlap the time of stigma receptivity, so a choice of cultivars is necessary in large plantings to ensure cross pollination. Pollen is windblown for distances up to 0.6 mi (1 km), and usually there is enough diversity of bloom habits to ensure fruit set in smaller orchards and home plantings. Because of the bloom habits, pecans are heterozygous and seedlings are highly variable.

Production. Nuts from native seedlings and from named grafted cultivars are marketed in roughly equal amounts. The former are more important to Texas, Oklahoma, and Louisiana while the latter are more important in Georgia, Alabama, other southeastern states and in new producing areas of western Texas, New Mexico, and Arizona. The tree requires good soil fertility and culture, and usually some spraying to give satisfactory production. Total annual production varies as much as 20% from year to year because of strong tendencies toward biennial bearing. Heavy production in one season, combined with premature leaf drop from disease, insects, or poor nutrition will result in light cropping the next year.

Breeding. As in all agricultural endeavors, the cultivar is the base for success. Old selections like Stuart and Western, dating from the late nineteenth century, are still grown extensively. Intensive breeding work produced cultivars (see **illus.**) that come into production early and have 0.25–0.28 oz (7–8 g) nuts with over 60% kernel content. Combined with tree spacings as close as 33 ft (10 m), these can make orchards profitable sooner and double or triple yields. *See* NUT CROP CULTURE. Ralph H. Sharpe

Processing. Harvesting of pecans is done either by mechanical means or by collecting the nuts that have fallen to the ground. Contact of pecan nuts with the

ground, particularly with wet ground, may contaminate the meat with pathogenic microorganisms, such as *Salmonella*, which can enter intact pecan shells. Processing, therefore, involves not only stabilizing the nut meat against spoilage microorganisms such as mold but also separation of the nut meat from the shell, removal of immature or damaged meats, and separation of shell fragments and pith from the nut meat.

Pecans are usually purchased from individual growers by wholesalers, who may process the pecans or sell them to pecan shellers. Mechanically harvested pecans must be separated from sections of tree limbs and other debris. Prior to storage in large lots in collection centers, inshell pecans are dried to 4.5–6% moisture by using low-temperature dry flowing air in trailers or bins. Moisture content below 6% is needed to prevent mold growth. Inshell pecans may be processed immediately or may be stored at 34°F (1°C) and 65% relative humidity prior to processing.

Approximately 94% of the total pecan crop is marketed as shelled pecans and 6% as inshell pecans. Inshell pecans are processed by first passing the nuts on a perforated conveyor belt under a vacuum hood to remove light nuts. This step is followed by a light spray of coloring agent, drying, application of a thin coat of wax, and packaging.

The processing of shelled pecans is more involved. Since pecan nut meat is a high-value item, shellers must recover most of the nut meats. The highest-valued product is the pecan halves, and the lowest-valued is the pecan meal. Thus, pecan shelling requires the utmost care in maintaining pecan pieces as large as possible. The pecan nut consists of a hard outer shell, and soft pithy material between the two halves of the meat and between the shell and the meat. The pith has a bitter flavor, and the hard shell can cause tooth damage to consumers, so the nut meats must be free of these materials. In addition, larvae of the pecan weevil in the nut mix with the meat fragments and must be removed. Pecan shelling operations involve several types of separation processes.

The first step in pecan shelling is grading the nuts into several size categories. Cracking machines are set to handle specific pecan sizes so that the nuts receive just the right amount of impact to crack the shell without damaging the nut meat. The nuts, separated into the proper size categories, are conditioned to add moisture to the meat, preventing them from shattering in the cracking machine. Conditioning is done by soaking the nuts in chlorinated warm water, followed by draining, and then equilibrating at room temperature. Use of chlorinated water in all operations where water contacts the inshell pecans and the nut meats minimizes microbial counts in the final product.

Cracking is accomplished by positioning individual nuts on a carrier that advances them at high speed between an anvil and a spring-operated impact tool. The cracked nuts are conveyed to a rotating cylinder that loosens the meat from the shell. The mixture of nut meat, pith, and shell fragments enters a primary water flotation tank, where the large shell fragments and most of the pith settles while the floating nut meats are removed by a conveyor. This mixture of nut meats of various sizes, including a few shell fragments and pith, is dried with hot air and separated by size by means of vibrating screens. The halves retained on the top screen are virtually free of shell and pith. These are conveyed to color sorters, where dark-colored ones are ejected electronically.

The sound nut meats are either packaged raw or fried in oil and packaged as roasted pecans. The smaller fragments may contain shell fragments and pith. They go through electronic sorting machines where off-colored ones and some shell fragments are ejected. These small nut meats contain shell fragments and pith which are separated by pressure flotation. They are placed in a tank full of water and pressurized to force water into the shell fragments and pith, which settle. The floating nut meats are skimmed from the top of the tank and air dried. Some nut meat fragments are of the same size as the pecan weevil larvae. These fragments are first electronically sorted to remove those with off-color. The larvae may be separated by flotation in dilute ethanol solution, where the larvae float while the nut meats settle. Following alcohol flotation, the nut meats are floated under pressure and air dried. As an alternative to alcohol flotation, larvae are removed by inspectors, who pick out the fluorescent larvae as the nut meats are passed on a conveyor under ultraviolet light. *See* FLOTATION.

Shelled pecans rapidly become rancid, so they are stored at 14°F (−10°C) or below in odor-free rooms. For applications such as ice cream, which requires very low microbial counts in the product, the shelled pecans may be decontaminated by exposure to ethylene oxide gas. *See* ICE CREAM. Romeo T. Toledo

Bibliography. J. Ogawa and H. English, *Diseases of Temperate Zone Tree Fruit and Nut Crops*, 1991; C. R. Santerre (ed.), *Pecan Technology*, 1994.

Pectin

A group of polysaccharides occurring in the cell walls and intercellular layers of all land plants. They are extractable with hot water, dilute acid, or ammonium oxalate solutions. Pectins are precipitated from aqueous solution by alcohol and are commercially used for their excellent gel-forming ability.

Source and isolation. Commercially, the primary source of pectin is the peel of citrus fruits such as lemon and lime, although orange and grapefruit may be used. A secondary source is apple pomace and sunflower heads.

Plant tissue is treated with sulfurous acid to inactivate pectic enzymes and washed with water to remove bitter glycosides and free sugars. The tissue is then suspended in boiling water at pH 1.5–3.0, obtained by addition of hydrochloric, sulfuric, or nitric acid. Acid extraction is continued for 30–60 min, and the mixture is filtered with filter aid and precipitated

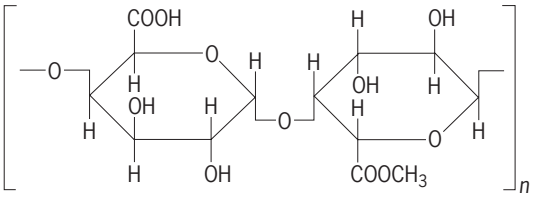

Partial methyl ester of pectic acid.

with isopropanol. The precipitate is pressed, dried, and milled to approximately 60-mesh.

Structure and properties. Native pectin is a mixture of polysaccharides, with the major component a polymer of α-D-galacturonic acid, mainly as the methyl ester (see **illus.**) and often with some acetyl groups on the hydroxyls at C-2, although some may be on the C-3 hydroxyls. The range of acetyl content may vary from pectin to pectin over the range 0.2–4%, with apple pectin being 0.45%. A minor component is a polymer of α-L-arabinofuranosyl units (an arabinan), joined by α-L-(1→5) linkages. Sugar units are substituted occasionally by α-L-arabinofuranosyl groups linked α-D-(1→3). Another polysaccharide component is a linear chain of β-D-galactopyranosyl units (a galactan) linked (1→4). Arabinan and galactan are largely broken down and removed during the extraction and purification steps. Commercial pectin can be regarded as mainly an α-D-galacturonan with varying degrees of methyl esterification.

Gel-forming properties of pectins are largely dependent on their content of methyl ester groups. Pectins are classified according to their degree of methylation. Those with methyl content 70% or greater are called high-methoxyl (or rapid-set) pectins, while those of methyl content less than 50% are called low-methoxyl pectins. Intermediate values of 60–65% are characteristic of slow-set pectins. Amidated pectins are those that have been treated with ammonium hydroxide to partially saponify the methyl ester groups and to replace some methyl ester groups with amide groups. Such pectins have reduced sensitivity to a calcium ions in gel formation.

Immature plant tissues contain an insoluble pectin precursor, protopectin, which, due to the action of various pectinases, becomes more soluble as tissue maturity and ripening proceed.

The molecular weight range is 30,000–300,000, and the linear nature of pectin is illustrated by its film-forming ability, by x-ray diffraction, viscosity, and other physical measurements.

Methyl groups are removed slowly by acid hydrolysis to leave at intermediate stages a partially methylated polysaccharide, pectinic acids, wherein the ester groups are randomly distributed. With the enzyme pectinesterase, obtained from such sources as roots, leaves, and fruits of all higher plants and also from a number of microorganisms, the ester groups are quickly removed. However, at intermediate stages the molecule contains both segments that have the full complement of ester groups and segments that are completely devoid of ester groups. Pectic acid has a lower water solubility than pectinic acid and is more affected by multivalent ions, such as calcium, that cause gelatin by forming structural cross-links.

Aqueous solutions of commercial pectin of 2–3% concentration may be easily prepared in warm or hot water. These solutions are viscous, with the viscosity depending on the polymer molecular weight range, the pH, the degree of methylation, and the solution composition, such as the presence of polyvalent ions and sugar. The free acid groups of pectin can be titrated in the normal manner to give titration curves that resemble those of simple, monobasic acids. Pectins are usually characterized by their anhydrouronic acid content, their solution viscosity (non-newtonian), and their degree of methylation.

Uses. Pectin is widely used in the food industry, principally in the preparation of gels. Food gels based on pectin have been known for over 500 years. Pectins are standardized on the basis of their jelly grade, that is, the number of weight units of sucrose that one weight unit of pectin will support in a standard 65%-sugar jelly containing a specified amount of acid. Good commercial-grade pectins have jelly grades in the 150–300 range. Pectin has been used in the preparation of jams, jellies, preserves, canned fruits, fruit juices, and confectionery products. It is also used as a stabilizer in some dairy products and frozen desserts, such as sherbet. Low-methoxyl pectins are useful as stabilizers for pudding mixes and also as edible protective coatings for sausages, almonds, candied dried fruit, and soft dates.

Pectin has also been used for the treatment of diarrhea, and several commercial products have a pectin-based formula. Intravenously administered pectin shortens the blood coagulation time and is thus useful in controlling hemorrhage. Work has shown that orally administered pectin lowers blood cholesterol and is an aid in improving carbohydrate tolerance in diabetics.

Pectins, in combination with plant hemicelluloses and lignin, may be useful dietary constituents in preventing coronary heart disease, diverticular disease, ulcerative colitis, and a variety of other Western diseases. D-Galacturonic acid prepared from pectin can be used to synthesize vitamin C. Pectin has numerous other medical and pharmaceutical uses. *See* COLLOID; GEL; GUM; HEMICELLULOSE; POLYSACCHARIDE.

Roy L. Whistler; Jams R. Daniel

Bibliography. M. L. Fishman and J. Jen (eds.), *Chemistry and Function of Pectins*, 1986; R. H. Walter (ed.), *The Chemistry and Technology of Pectin*, 1997.

Pectolite

A mineral inosilicate with composition $Ca_2NaSi_3O_8$ (OH), crystallizing in the triclinic system. Crystals are usually acicular in radiating aggregates (see **illus.**). There is perfect cleavage parallel to the front and basal pinacoids yielding splintery fragments elongated on the *b* crystal axis. The hardness is 5 on Mohs

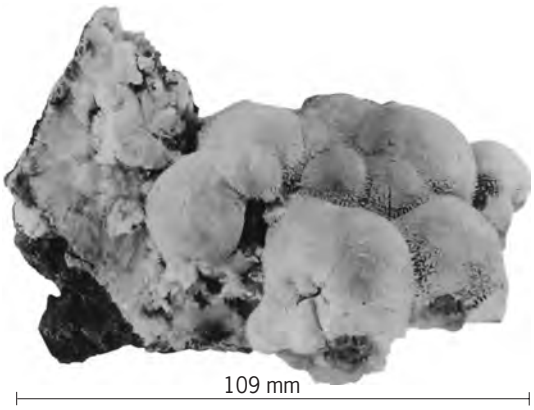

Globular masses of needlelike pectolite crystals with calcite at left, found in Paterson, New Jersey. (*American Museum of Natural History specimen*)

scale, and the specific gravity is 2.75. The mineral is colorless, white, or gray with a vitreous to silky luster. Pectolite, a secondary mineral occurring in cavities in basalt and associated with zeolites, prehnite, apophyllite, and calcite, is found in the United States at Paterson, Bergen Hill, and Great Notch, New Jersey. *See* SILICATE MINERALS. Cornelius S. Hurlbut, Jr.

Pediculosis

Human infestation with lice. There are two biological varieties of the human louse, *Pediculus humanus*, var. *capitis* and var. *corporis*, each showing a strong preference for a specific location on the human body. *Pediculus humanus capitis* colonizes the head and *P. h. corporis* lives in the body-trunk region. *See* ANOPLURA.

These lice are wingless insects which are ectoparasites. Their mouthparts are modified for piercing skin and sucking blood. The terminal segments of their legs are modified into clawlike structures which are utilized to grasp hairs and clothing fibers.

After a blood meal, the male and female lice copulate, and the female deposits (cements) eggs to hair shafts (*P. h. capitis*) or clothing fibers (*P. h. corporis*); these eggs are called nits and are visible to the unaided eye. After emerging from the egg, a sexually immature louse feeds on blood and goes through three stages (instars) of development before becoming a sexually mature adult louse. This life cycle requires approximately 3 weeks.

Lice are important vectors of human diseases. Their habit of sucking blood and their ability to crawl rapidly from one human to another transmit relapsing fever (spirochetal). The body fluids and feces of infected lice transmit these diseases. Epidemic typhus (*Rickettsia prowazeki*), European epidemic relapsing fever (*Borrelia recurrentis*), and trench fever (*Rochalimaea quintana*) are transmitted by the body louse. The role of head lice and pubic lice (*Pthirus pubis*) in disease transmission has not been demonstrated, and their involvement in the transmission of hepatitis and acquired immune deficiency syndrome (AIDS) remains to be determined. However, mild-to-severe dermatitis, which appears to be an allergic response to the secretions and excretions of the lice, may occur during louse infestation. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); HEPATITIS.

Historically, evidence of head lice infestations of humans has been indicated by the presence of *Pediculus h.c.* eggs (nits) on hair combs dating back to the first century B.C. in the Middle East. Egyptian mummies have also been found with evidence of head lice. With the advent of regular bathing, the human louse populations were reduced, and the number of cases of louse-borne diseases diminished. Disruption of human populations during disasters caused by nature or humans, such as crowding into refugee camps, tends to increase louse populations owing to unsanitary conditions. Infestation of school children by head lice continues to be a public health problem; these school-centered infestations are cosmopolitan.

Pediculosis can be controlled by public education, good personal hygiene, and the use of insecticides, such as lindane or pyrethrin with piperonyl butoxide. Richard Sudds

Bibliography. L. R. Barker, J. R. Burton, and P. D. Zieve, *Principles of Ambulatory Medicine*, 5th ed., 1998; W. K. Joklik et al., *Zinsser Microbiology*, 20th ed., 1995; T. M. Peters, *Insects and Human Society*, 1988; S. A. Schroeder, M. A. Krupp, and L. M. Tierney, *Current Medical Diagnosis and Treatment*, 1988; H. Zinsser, *Rats, Lice, and History*, 1935.

Pedinoida

An order of Diadematacea, making up those genera which possess solid spines and a rigid test. The ambulacra show typical diadematoid structure, and the tubercles are noncrenulate. The single known family, Pedinidae, includes 15 genera, ranging from the Late Triassic onward, though only one genus, *Caenopedina*, survives today; the latter inhabits the Indo-Pacific and Caribbean seas, ranging from the continental shelf down to 660 ft (2000 m). *See* ECHINODERMATA; ECHINOIDEA. Howard B. Fell

Bibliography. R. E. Moore (ed.), *Treatise on Invertebrate Paleontology*, U-3(1):340–366, 1966.

Pedology

Defined narrowly, a science that is concerned with the nature and arrangement of horizons in soil profiles; the physical constitution and chemical composition of soils; the occurrence of soils in relation to one another and to other elements of the environment such as climate, natural vegetation, topography, and rocks; and the modes of origin of soils. Pedology so defined does not include soil technology, which is concerned with uses of soils.

Broadly, pedology is the science of the nature, properties, formation, distribution and function of

soils, and of their response to use, management, and manipulation. The first definition is widely used in the United States and less so in other countries. The second definition is worldwide. *See* SOIL; SOIL CONSERVATION; SOIL MECHANICS. Roy W. Simonson

Pegasidae

The sea moths or sea dragons, a small family of peculiar actinopterygian fishes also known as the Hypostomides. The precise relationships are unknown. The body is encased in a broad, bony framework anteriorly and has bony rings posteriorly, simulating the seahorses and pipefishes (see **illus.**). Unlike those fishes, which have the mouth at the tip of the produced snout, sea moths have enlarged nasal bones which form a rostrum that projects well forward of the small, toothless mouth. The preopercle, pterosphenoid, intercalar, entopterygoid, and metapterygoid bones are absent, and the gill cover has a single bone. The greatly expanded horizontal pectoral fin belies its appearance and does not function in aerial gliding. The pelvic fin is abdominal and consists of a slender spine and one or two long rays. The short, opposed dorsal and anal fins have no spines. There is no swim bladder. *See* GASTEROSTEIFORMES.

There are two genera, *Pegasus* and *Eurypegasus*, and five species. They live amidst vegetation on Indo-Pacific shores from East Africa to Japan, Australia, and Hawaii. They rarely exceed 4 in. (10 cm) in length. There is no fossil record. *See* ACTINOPTERYGII. Reeve M. Bailey

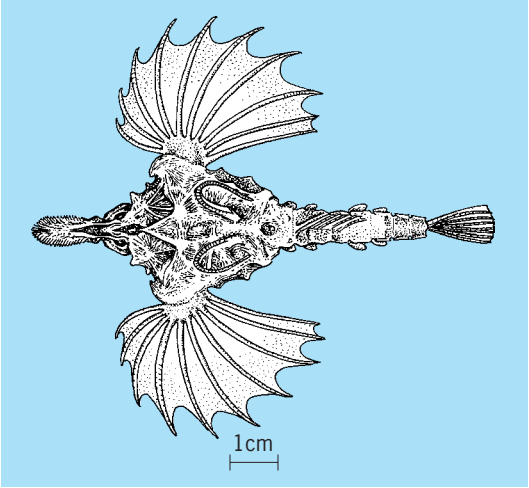

Sea moth (*Pegasus draconis*). (*After D. S. Jordan and J. O. Snyder, vol. 24, Leland Stanford University Contributions to Biology, 1901*)

Pegasus

The Flying Horse, a giant northern constellation (see **illustration**). It is noticeable for a huge square of stars (of which one is now considered to be in

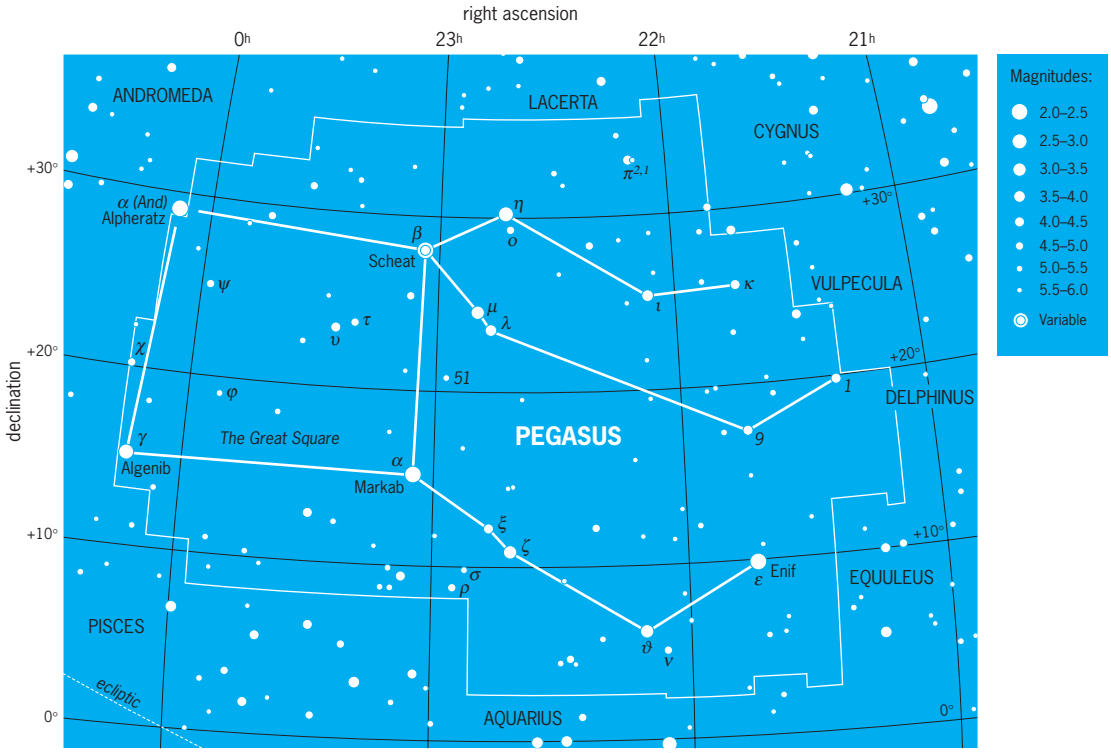

Modern boundaries of the constellation Pegasus, the Flying Horse. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

neighboring Andromeda), the Great Square of Pegasus. Each side is half again larger than a fist held at the end of the observer's outstretched arm. The Great Square is marked by Alpheratz (now alpha Andromedae), Scheat (beta Pegasi), Markab (alpha Pegasi), and Algenib (gamma Pegasi). Pegasus is prominent in the autumn evening sky. *See* ANDROMEDA.

In Greek mythology, Pegasus was born out of the blood of the gorgon Medusa after she was killed by Perseus. Pegasus was then tamed by Bellerophon. When Bellerophon tried to ride Pegasus to Olympus, Zeus killed him, but Pegasus continued to climb, winding up in the heavens. *See* PERSEUS.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Pegmatite

Exceptionally coarse-grained and relatively light-colored crystalline rock composed chiefly of minerals found in ordinary igneous rocks. Extreme variations in grain size also are characteristic (**Fig. 1**), and close associations with dominantly fine-grained aplites are common. Pegmatites are widespread and very abundant where they occur, especially in host rocks of Precambrian age, but their aggregate volume in the Earth's crust is small. Individual bodies, representing wide ranges in shape, size, and bulk composition, typically occur in much larger bodies of intrusive igneous rocks or in terranes of metamorphic rocks, from most of which they are readily distinguished by their unusual textures (**Fig. 2**) and often by concentrations of relatively rare minerals. Many pegmatities have been economically valuable as sources of clays, feldspars, gem materials, industrial crystals, micas, silica, and special fluxes, as well as beryllium, bismuth, lithium, molybdenum, rare-earth, tantalum-niobium, thorium, tin, tungsten, and uranium minerals. *See* APLITE; CLAY MINERALS; FELDSPAR; GEM; MICA.

Compositional types. Most abundant by far are pegmatite bodies of granitic composition, corresponding in their major minerals to granite, quartz monzonite, granodiorite, or quartz diorite. In some regions or districts they contain notable quantities of the less common or rare elements such as As, B, Be, Bi, Ce, Cs, La, Li, Mo, Nb, Rb, Sb, Sn, Ta, Th, U, W, Y, and Zr. Diorite and gabbro pegmatites occur in many bodies of basic igneous rocks, but their total volume is very small. Syenite and nepheline syenite pegmatites are abundant in a few regions characterized by alkaline igneous rocks. Some of them are noted for their contents of As, Ce, La, Nb, Sb, Th, U, Zr, and other rare elements.

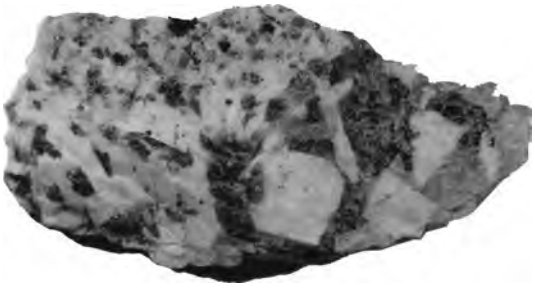

Fig. 1. Contrasting textures in pegmatite of quartz monzonite composition, New Hampshire. Upper part is relatively fine-grained and granitoid, and remainder comprises large, roughly faced crystals of alkali feldspar (light) set in slightly smoky quartz (dark).

Shape and size. Individual bodies of pegmatite range in maximum dimension from a few inches to more than a mile, and in shape from simple sheets, lenses, pods, and pipes to highly complex masses with many bulges, branches, and other irregularities. Tabular to podlike masses of relatively small size commonly occur as segregations or local injections within bodies of cogenetic igneous rocks, where they represent almost the full spectrum of known pegmatite compositions. Granitic and syenitic pegmatites also occur in metamorphic rocks, some of them distributed satellitically about igneous plutons. Their contacts with the enclosing rocks may be sharp or gradational. Many bodies formed in strongly foliated or layered rocks are conformable with the host-rock structure (Fig. 2); others are markedly discordant, as if their emplacement were controlled by cross-cutting fractures or faults.

Mineral composition. Essential minerals (1) in granitic pegmatities are quartz, potash feldspar, and sodic plagioclase; (2) in syenitic pegmatites, alkali feldspars with or without feldspathoids; and (3) in diorite and gabbro pegmatites, soda-lime or lime-soda plagioclase. Varietal minerals such as micas, amphiboles, pyroxenes, black tourmaline, fluorite, and calcite further characterize the pegmatites of specific districts. Accessory minerals include allanite, apatite, beryl, garnet, magnetite, monazite, tantalite-columbite, lithium tourmaline, zircon, and a host of rarer species.

Most pegmatites are mineralogically simple, with only one or two varietal constituents and a few sparse accessories. More complex pegmatites typically contain much albite (commonly the variety cleavelandite) and groups of minerals that reflect unusual concentrations of rare elements such as beryllium (beryl, chrysoberyl, gadolinite, phenakite), boron (axinite, tourmaline), and lithium (amblygonite, lepidolite, petalite, spodumene, triphylite-lithiophilite, zinnwaldite). Cavities and pockets in some pegmatite bodies contain beautifully formed crystals, and such occurrences have yielded transparent gemstones of beryl, garnet, quartz, spodumene, topaz, tourmaline, and other minerals.

Texture. The most distinctive features of pegmatites are coarseness of grain and great variations in grain size over short distances (Fig. 1 and **Fig. 3**). The average grain size for all occurrences is on the order of 4 in. (10 cm), but mica and quartz crystals 10 ft (3 m) across, beryl and tourmaline crystals 10 to 20 ft (3 to 6 m) long, feldspar crystals 30 ft (9 m) long, and spodumene crystals nearly 50 ft (15 m) long have been found, as have individual crystals of allanite,

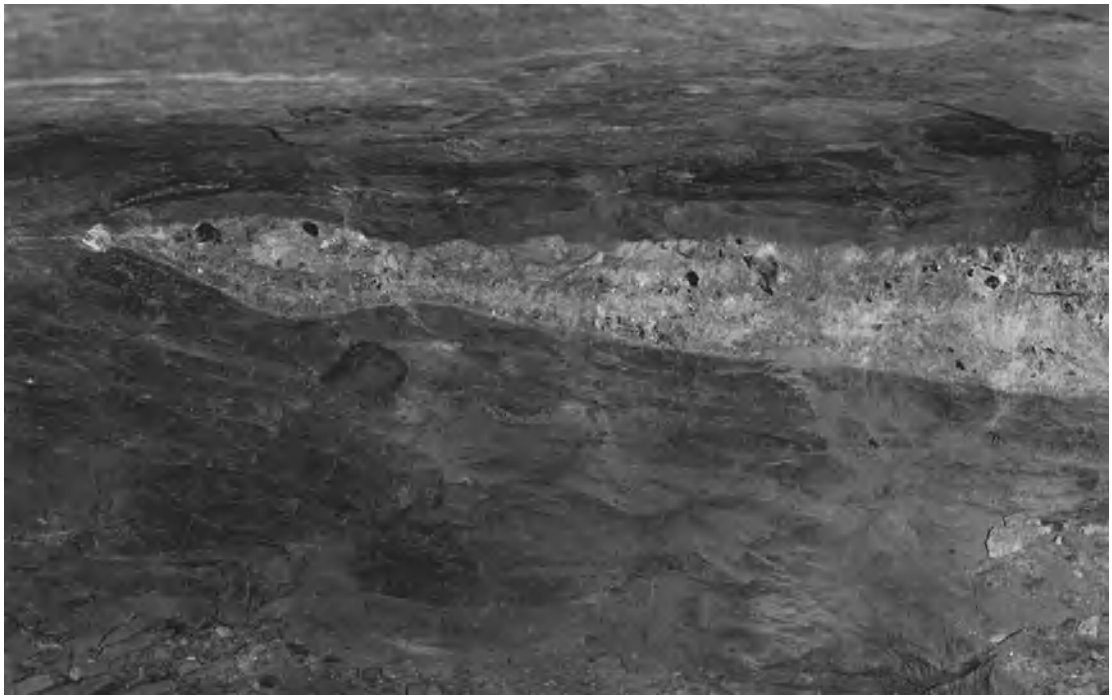

Fig. 2. Sill of faintly layered quartz diorite pegmatite, 6 to 12 in. (15 to 30 cm) thick, in fine-grained granitic gneiss, North Carolina. Sodic feldspar (light), quartz (darker), and muscovite (very dark) are the principal pegmatite minerals. (*From R. H. Jahns, The study of pegmatites, Econ. Geol., 50th Anniv. Vol., pp. 1025–1030, 1955*)

columbite, monazite, and topaz weighing hundreds of pounds.

On much smaller scales, sodic plagioclase is intergrown with host potash feldspar to form perthite, and quartz with feldspars in cuneiform fashion to form graphic granite. More or less regular intergrowths of hematite and magnetite in muscovite, and of muscovite, garnet, or tourmaline in quartz also are common. Evidence of corrosion and replacement of some minerals by others typically is present over a wide range of scales. *See* IGNEOUS ROCKS.

Fig. 3. Four contrasting zones in upper part of thick, gently dipping dike of very coarse-grained granite pegmatite overlain by dark-colored metamorphic rocks, New Mexico. From top down, the zones consist of: quartz with alkali feldspars, apatite, beryl, and muscovite (light band); quartz (dark band); quartz with giant lathlike crystals of spodumene (thick band with comblike appearance); and quartz, alkali feldspars, spodumene, lepidolite, and microlite (mined-out recesses). (*From E. N. Cameron et al., Internal structure of granitic pegmatites, Econ. Geol., Monogr. 2, 1949*)

Internal structure. Many pegmatite bodies are grossly homogeneous, whereas others, including nearly all of those with economic mineral concentrations, comprise internal units of contrasting composition or texture. Such internal zoning is systematically expressed as successive layers in nearly flat-lying bodies with tabular form (Figs. 2 and 3), and elsewhere as layers or shells that also reflect the general form of the respective bodies (**Fig. 4**). In horizontal section the zones tend to be arranged concentrically about a single or segmented core, but in vertical section they are typically asymmetric and less continuous (Fig. 4). Some pegmatite bodies are distinguished by relatively fine-grained outer zones that are more regular and continuous in three dimensions. Associations of major minerals within successive zones consistently follow a definite sequence of worldwide application.

Fracture-filling masses of tabular form transect parts of some pegmatite bodies. They range in thickness from less than an inch to nearly 10 ft (3 m), and most are composed of quartz or of quartz, feldspars, muscovite, and other common minerals in various combinations. Less regular masses of lenticular, branching, or network form, and a few inches to several hundred feet in maximum dimension, appear to have developed wholly or in part at the expense of earlier pegmatite. Typical examples are quartz, potash feldspar, and muscovite replaced by aggregates of cleavelandite or sugary albite, alkali feldspar and quartz by muscovite, and potash feldspar and spodumene by lepidolite. Some replacement masses also contain notable concentrations of rare minerals.

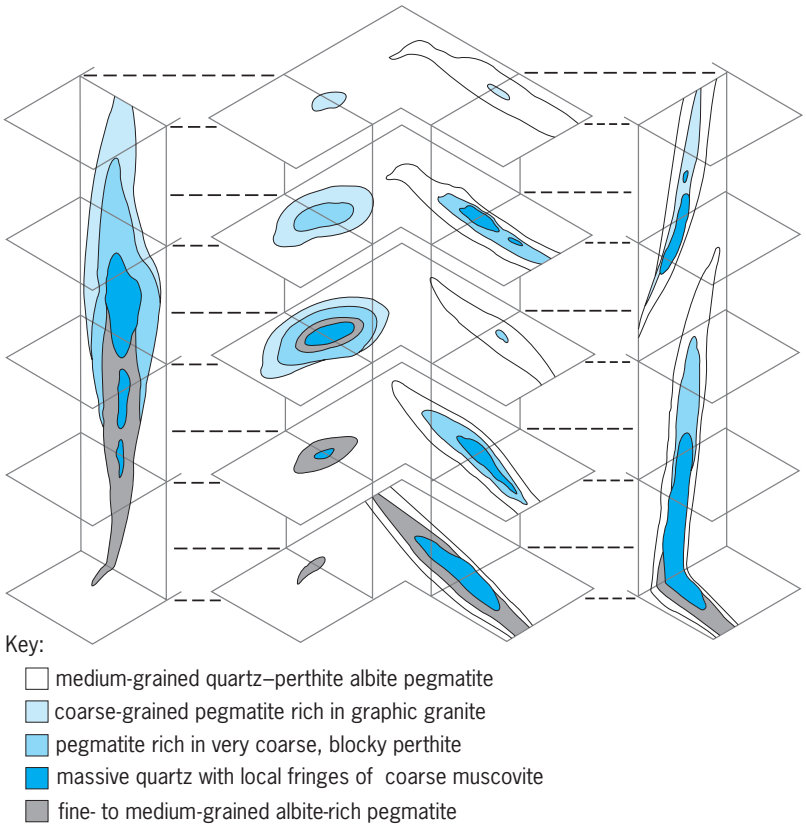

Fig. 4. Isometric diagram comprising vertical sections of a cucumber-shaped pegmatite body (left) and parts of two pegmatite dikes (right), together with successive horizontal sections of all three bodies (center). Internal zoning is shown in idealized form without a specific scale, and its configuration within relatively horizontal bodies can be visualized through a figurative vertical compression of the bodies shown here.

Origin. Many pegmatites must have formed from the residual fractions of crystallizing rock melts (magmas) or from the initial fractions of deeply buried crustal materials undergoing partial melting. Such fractions would be silicate liquids relatively rich in water, halogens, and other volatile constituents, and also in many of the less common elements. They could remain within the crystallized or partly melted parent rocks to form bodies of segregation pegmatite with irregular and gradational boundaries, or they could be forced away to form homogeneous or zoned bodies of pegmatite elsewhere in the parent rocks or in rocks beyond. *See* MAGMA.

Other pegmatites, generally regarded as metamorphic in origin, consist of materials probably derived from the enclosing rocks via aqueous pore fluids and concentrated at favorable sites in the form of secretionary or replacement bodies. Such transfer of materials could account for numerous veins, pods, and irregular bodies of pegmatite with relatively simple composition and internal structure, and for the development of pegmatite on a grand scale in Precambrian terranes of metamorphism and granitization in regions such as Fennoscandia, India, central Africa, and eastern Canada. *See* METAMORPHIC ROCKS; METAMORPHISM.

Long-time studies of pegmatites and experimental investigations of pertinent silicate systems indicate that the unifying factor in pegmatite genesis is the presence of a high-temperature aqueous fluid. When such a fluid exsolves from a crystallizing water-rich magma, it serves as a low-viscosity medium into which other constituents can be preferentially transferred from the magma, and through which they can diffuse rapidly but selectively in response to temperature-induced concentration gradients. Thus it is an effective vehicle for the nourishment of large crystals and development of coarse textures, and for the gross segregation of constituents to form zonal structure in many pegmatite bodies. Upward gravitational rise of low-density aqueous fluid also can account for observed vertical asymmetry of zonal structure (Fig. 4).

The aqueous fluid promotes exsolution of alkali feldspars, mobilization of the released albitic constituents, and extensive reactions among earlier-formed crystals both before and after all magma has been used up. Thus pegmatites of igneous origin reflect metasomatic activities as well. Moreover, the transfer of materials and the mineral replacements implied in the origin of metamorphic pegmatites can be expected to occur if an interstitial aqueous fluid is available in the hot host rocks and a temperature or pressure gradient is present, whether or not a magma is at hand. *See* METASOMATISM. Richard H. Jahns

Bibliography. D. Hyndman, *Petrology of Igneous and Metamorphic Rocks*, 2d ed., 1985; A. Miyashiro, *Metamorphic Petrology*, 1994; B. B. Rao, *Metamorphic Petrology*, 1986.

Pelecaniformes

A small order of diverse aquatic, mainly marine, fish-eating birds that includes the pelicans (see **illustration**), boobies, and cormorants. The members of the order, found worldwide, are very different, and some researchers believe that this "order" is an artificial group; however, all members are characterized by several unique features (for example, a bare throat pouch).

Although a few scientists place the whale-billed stork (Balaenicipitidae) in the Pelecaniformes, this is a quite controversial conclusion. The five suborders and nine families of the order are listed below. *See* CICONIIFORMES.

Order Pelecaniformes
- Suborder Phaethontes
  - Family: Prophaethontidae (fossil; Eocene of England)
    - Phaethontidae (tropic birds; 3 species; oceanic pantropical)
- Suborder Odontopterygia
  - Family Pelagornithidae (fossil; lower Eocene to Miocene; worldwide marine)
- Suborder Pelecani
  - Family Pelecanidae (pelicans; 8 species; worldwide)
- Suborder Sulae
  - Family: Sulidae (boobies, gannets; 9 species; oceanic worldwide)
  - Plotopteridae (fossil; Oligocene to Miocene; North Pacific)
  - Phalacrocoracidae (cormorants; 33 species; worldwide)
  - Anhingidae (anhingas; 4 species; freshwater pantropical)
- Suborder Fregatae
  - Family Fregatidae (frigate birds; 5 species; oceanic worldwide)

Some taxonomists have enlarged the Ciconiiformes to include the pelecaniform pelicans and allies, as in the Sibley-Ahlquist taxonomy (a radical new approach to taxonomy, based on DNA hybridization studies).

***Pelecanus erythrorhynchos*, the American White Pelican. (*Courtesy of Dr. Lloyd Glenn Ingles, © California Academy of Sciences*)**

Fossil record. The pelecaniforms have an extensive, and indeed exciting, fossil record. *Prophaethon* is an Eocene primitive tropical bird, and *Lintnofregata* is an early Eocene primitive frigate bird. Fossil pelicans are known from the lower Miocene, boobies from the Oligocene, and anhingas from the late Miocene. The Eocene *Proloplotus* may possibly belong to a distinct family, but on the other hand may be related to the Plotopteridae and to cormorants from the Eo-Oligocene phosphorite beds of France. Most families have a relatively complete record continuing to the present. Of greater interest are the two purely fossil families, the Pelagornithidae and the Plotopteridae.

The Pelagornithidae, or pseudodontorns, are a remarkable family of extinct gigantic marine birds with long bills that bore numerous bony projections up to 2 in. (5 cm) long, the pseudoteeth. They are found in deposits dating from the early Eocene to the late Miocene or early Pliocene from many locations throughout the world. They were flying birds with wingspans as great as 18–20 ft (5.5–6 m), longer than that of any known albatross and rivaling that of the largest fossil Teratornithidae (fossil New World vultures). The pseudodontorns appear to have had behavior patterns similar to those of albatrosses. *See* PROCELLARIIFORMES.

A second remarkable group is the Plotopteridae, flightless, wing-propelled diving relatives of the cormorants that have been found in a number of sites around the North Pacific from California to Japan. They are known from the late Oligocene to their disappearance in the middle Miocene. The plotopterids range from medium-sized cormorants to birds larger than emperor penguins. A nondescript femur suggests a form that would be the largest known diving bird. The plotopterids are closely convergent to the penguins, but evolved from a different ancestral stock and in a different region of the world. The reasons for the extinction of the plotopterids is as mysterious as for the Pelagornithidae. *See* SPHENISCIFORMES.

Characteristics. The pelecaniforms are medium-sized to large marine birds that are characterized by a foot with four toes united in a common web (totipalmate). The legs are short and stout or weak, and they are used for swimming and perching. The birds are poor walkers, but all living forms—with the exception of the flightless Galápagos cormorant, *Namiopterum barrisi*—are excellent fliers, particularly the frigate birds. The bill varies widely in size and shape, from the short stout bill of tropical birds to the long hooked bill of frigate birds and the long, flat, flexible bill of pelicans. All species have a bare throat pouch; the largest such pouch is found in the pelicans, which use it as a fishing net. The brilliant red pouch of the male frigate bird is blown up in courtship like a balloon. Pelecaniform plumage is black, gray, or white. All pelecaniforms feed on fish, crustaceans, and squid that are caught by diving from the air or the water surface, are taken from the water surface, or are stolen from other birds. Aerial diving of the sulids and the brown pelican are spectacular.

The clavicle is fused to the sternum for additional protection against injury from the impact with water during dives.

Pelecaniforms are usually gregarious, breeding in colonies. The pairs are monogamous, and both sexes incubate and care for the young, which remain in the nest until they are able to fly.

Economic significance. The Chinese and Japanese once trained cormorants for fishing, but that method is now uneconomical. Several species of cormorants are valuable as guano (phosphate- and nitrogen-rich, partially decomposed excrement of seabirds) producers; thus, in some areas, large platforms are built to encourage nesting of cormorants. The guanay (*Phalacrocorcax bougainvilli*) is the chief contributor to the guano deposits of Peruvian islands, and is considered to have been the most valuable wild bird because of accumulation of deep beds of guano, now largely mined. Several species, especially the brown pelican of North America (*Pelecamus occidentalis*), have been severely reduced in numbers because of concentrations of the insecticide DDT in their food. Some species of cormorants have been shot by fishermen because of the misconception that they deplete stocks of fish. *See* AVES. Walter J. Bock

Bibliography. P. A. Johnsgard, *Cormorants, Darters and Pelicans of the World*, 1993; J. B. Nelson, *Pelicans, Cormorants and Their Allies*, Oxford University Press, 2005.

Pelmatozoa

A division of the Echinodermata made up of those forms which are anchored to the substrate during at least a part of the life history. Formerly treated as a formal unit of classification with the rank of subphylum, pelmatozoans are now realized to be a heterogeneous assemblage of forms with similar habits but dissimilar ancestry, their common features having arisen by convergent evolution. Most pelmatozoan echinoderms are members of the subphylum Crinozoa, but some echinozoans also exhibit a sedentary, anchored life, with modifications for such existence. *See* ECHINODERMATA; ECHINOZOA; ELEUTHEROZOA. Howard B. Fell

Peltier effect

A phenomenon discovered in 1834 by J. C. A. Peltier, who found that at the junction of two dissimilar metals carrying a small current the temperature rises or falls, depending upon the direction of the current. Many different pairs of metals were investigated; bismuth and copper were among the first. The temperature rises at a junction where the flow of positive charge is from Cu to Bi and falls where the flow is from Bi to Cu. A reversible output of heat occurs at the first-named junction and a reversible intake at the second. In view of the experiments of Quintus Icilius (1853), which established that the rate of intake or output of heat is proportional to the magnitude of the current, it can be shown that an electromotive force resides at a Cu-Bi junction, directed from Bi to Cu. Electromotive forces of this type are called Peltier emf's. *See* SEEBECK EFFECT; THERMOELECTRICITY; THOMSON EFFECT. John W. Stewart

Pelycosauria

A paraphyletic grade of fossil (extinct) synapsids that is ancestral to later forms, such as the therapsids and mammals. (Paraphyletic means not containing all of the descendants of the most recent common ancestor.) Modern taxonomic practice does not formally recognize ancestral groups—hence the use of pelycosaur-grade synapsid or Permo-Carboniferous synapsid for this assemblage.

The earliest pelycosaur-grade synapsids are known from fragmentary material recovered from the Middle Pennsylvanian of North America, approximately 310 million years in age (Ma). By Late Pennsylvanian

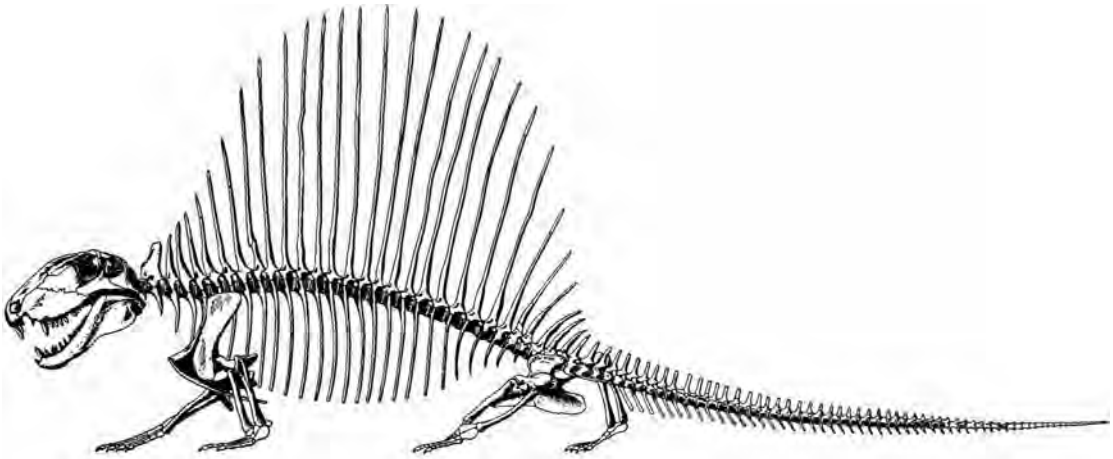

Skeletal reconstruction of the pelycosaur-grade synapsid *Dimetrodon*. Remains of this animal are relatively common in the Lower Permian rocks of north-central Texas, where it was the top carnivore of its ancient ecosystem. *Dimetrodon* grew up to 10 ft (3 m) in total body length. (*Reconstruction modified from A. S. Romer and L. I. Price, Review of the Pelycosauria, Geol. Soc. Amer. Spec. Pap., 28:1–538, 1940*)

and Early Permian times (305–270 Ma), however, synapsids were the dominant terrestrial vertebrates and had diversified into herbivorous, insectivorous, and carnivorous groups. Except for a handful of late surviving forms, pelycosaur fossils are typically restricted to Texas, Oklahoma, and the surrounding states, as well as several localities in Europe. Interestingly, these areas all lie close to the paleoequator (the Earth's equator as positioned in the geologic past), which suggests that the cold climates reconstructed for the Permo-Carboniferous acted as a barrier to early synapsid dispersal.

Paleontologists currently recognize six family-level pelycosaur clades (Caseidae, Edaphosauridae, Eothyrididae, Ophiacodontidae, Sphenacodontidae, and Varanopidae) that encompass approximately 30 genera. Ecologically, large-bodied sphenacodontids, such as the sailback *Dimetrodon* (see **illustration**), were the earliest land-living vertebrates adapted to prey upon similarly sized animals. Caseids and edaphosaurids, by contrast, were among the first terrestrial herbivores. For example, the giant early Permian caseid *Cotylorhynchus* is estimated to have reached 15 ft (4.6 m) in length and weighed upward of 800 lb (360 kg), which represents the largest vertebrate to have evolved by that point in Earth history. Pelycosaur-grade synapsids are not known after the Middle Permian (260 Ma). *See* MAMMALIA; SYNAPSIDA; THERAPSIDA. Christian A. Sidor

Bibliography. M. J. Benton, *Vertebrate Palaeontology*, 2005; R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988; T. S. Temp, *Origin and Evolution of Mammals*, 2005.

Pendulum

A rigid body mounted on a fixed horizontal axis, about which it is free to rotate under the influence of gravity. The period of the motion of a pendulum is virtually independent of its amplitude and depends primarily on the geometry of the pendulum and on the local value of g, the acceleration of gravity. Pendulums have therefore been used as the control elements in clocks, or inversely as instruments to measure g.

Pendulum motion. In the schematic representation of a pendulum shown in the **illustration**, the line OC makes an instantaneous angle θ with the vertical. In rotary motion of any rigid body about a fixed axis, the angular acceleration is equal to the torque about the axis divided by the moment of inertia I about the axis. If m represents the mass of the pendulum, the force of gravity can be considered as the weight mg acting at the center of mass C. Therefore, the angular acceleration α is determined by relation (1), where

$$-mgh \sin\theta = I\alpha = I\frac{d^2\theta}{dt^2} \tag{1}$$

h is the distance OC, and t represents time.

If the amplitude of motion is small, $\sin\theta \approx \theta$ in radian measure. In this approximation the motion is simple harmonic. Equation (2) has for its solution

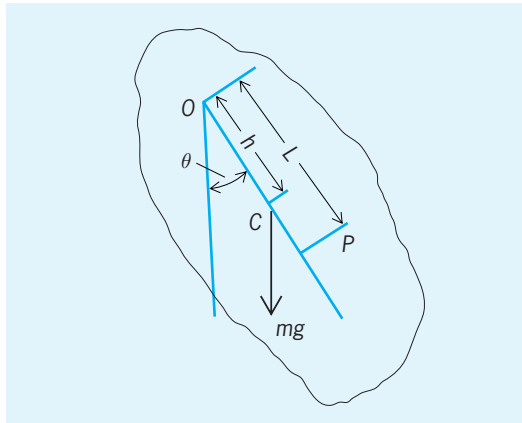

Schematic diagram of a pendulum. *O* represents the axis, *C* is the center of mass, and *P* the center of oscillation.

Eq. (3), where the amplitude A and the phase δ are

$$-mgh\theta = I\frac{d^2\theta}{dt^2} \tag{2}$$

$$\theta = A \sin(\omega t - \delta) \tag{3}$$

arbitrary constants. The angular frequency ω is given by Eq. (4). The period T, time for a complete vibra-

$$\omega^2 = \frac{mgh}{I} \tag{4}$$

tion (for example, from the extreme displacement right to the next extreme displacement right), and frequency f, number of vibrations per unit time, are given by Eq. (5).

$$T = \frac{1}{f} = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{I}{mgh}} \tag{5}$$

See HARMONIC MOTION.

The actual form of a pendulum often consists of a long, light bar or a cord that serves as a support for a small, massive bob. The idealization of this form into a point mass on the end of a weightless rod of length L is known as a simple pendulum. An actual pendulum is sometimes called a physical or compound pendulum. In a simple pendulum, the lengths h and L become identical, and the moment of inertia I equals mL^2. Equation (5) for the period becomes Eq. (6).

$$T = 2\pi\sqrt{\frac{L}{g}} \tag{6}$$

Because the value of g in metric units (about 9. 8 m/s^2) is very nearly equal to π^2, a simple pendulum 1 m (3.28 ft) in length has a period very close to 2 s; the time for a single swing from right to left or left to right is approximately 1 s.

Center of oscillation. Equation (6) can be used to define the equivalent length of a physical pendulum. Comparison with Eq. (5) shows that Eq. (7) holds.

$$L = \frac{I}{mh} \tag{7}$$

The point P on line OC of the figure, whose distance

from the axis O equals L, is called the center of oscillation. Points O and P are reciprocally related to each other in the sense that if the pendulum were suspended at P, O would be the center of oscillation.

The proof of this relation follows from the parallel axis property of moments of inertia. If the moment of inertia of the pendulum about its center of mass is defined by Eq. (8), then Eq. (9) holds; and if defined by Eq. (7), then Eq. (10) holds.

$$I = mb^2 \tag{8}$$

$$I = m(h^2 + b^2) \tag{9}$$

$$hL = h^2 + b^2 \tag{10}$$

For a given value of b (the radius of gyration about the center of mass) and L, h is determined by Eq. (10) to be either of the quantities given in Eq. (11). The

$$h = \frac{L}{2} \pm \sqrt{\frac{L^2}{4} - b^2} \tag{11}$$

sum of these two values is equal to L: therefore, if one value of h is the distance OC, then CP must represent the other value of h that will give the same equivalent length.

If some particular body with a definite value for b is to be mounted about an arbitrary axis to make a pendulum, Eq. (10) shows that L can never be less than $2b$, and that L will have this minimum value (and the period T will be a maximum) if h is made equal to b.

Center of percussion. The points O and P share another reciprocal property. If the body is free to move in the plane of the figure, instead of fixed on an axis, and an impulsive force is applied to the body at O, the initial motion of the body will be a rotation about P. For this reason P is sometimes called the center of percussion about O.

If the motion of a pendulum is not limited to small amplitudes, Eq. (2) is not an adequate substitute for the correct Eq. (1). The angular velocity, $d\theta/dt$, can be derived as a function of displacement by multiplying both sides of Eq. (1) by $d\theta/dt$ and then integrating. The result, which can also be obtained directly from the principle of conservation of energy, is given in Eq. (12), where θ_0 is the maximum displacement

$$\left(\frac{d\theta}{dt}\right)^2 = 2\omega^2(\cos\theta - \cos\theta_0) \tag{12}$$

or amplitude of the motion. Here ω is still the characteristic constant of the pendulum defined by $\omega^2 = mgh/I = g/L$, although the relation between ω and frequency is no longer as simple as in the approximate Eq. (5).

To obtain θ as a function of time, introduce an angle ψ by relation (13), where $k \equiv \sin(\theta_0/2)$.

$$\sin\psi \equiv \frac{1}{k}\sin\frac{\theta}{2} \tag{13}$$

Equation (12) becomes Eq. (14). If time is chosen

$$\begin{aligned}\omega\, dt &= \pm\frac{d\theta}{\sqrt{2(\cos\theta - \cos\theta_0)}} \\ &= \pm\frac{d\theta}{2\sqrt{\sin^2(\theta_0/2) - \sin^2(\theta/2)}} \\ &= \pm\frac{d\psi}{\sqrt{1 - k^2\sin^2\psi}}\end{aligned} \tag{14}$$

zero when θ is zero, then Eq. (15) holds, where

$$\omega t = F(k, \psi) \tag{15}$$

$F(k,\psi)$ is the standard elliptic integral of the first kind, as defined by Eq. (16). Conversely, the angle

$$F(k, \psi) = \int_0^\psi \frac{dz}{\sqrt{1 - k^2\sin^2 z}} \tag{16}$$

θ can be expressed as an elliptic function of time, as in Eq. (17).

$$\sin\frac{\theta}{2} = k\,\sin(\omega t) \tag{17}$$

The accurate expression for the period T, obtained from Eq. (15), can be written in terms of the complete elliptic integral of the first kind, $K(k)$, as Eq. (18). Numerical values for the ratio of the period for amplitude θ_0 to the period for infinitesimal amplitude are listed in the **table**.

$$\omega T = 4F\left(k, \frac{\pi}{2}\right) = 4K(k) \tag{18}$$

Pendulum types. The following paragraphs describe important types of gravity pendulums.

Kater's reversible pendulum. This type is designed to measure g, the acceleration of gravity. It consists of a body with two knife-edge supports on opposite sides of the center of mass as at O and P (and with at least one adjustable knife-edge). If the pendulum has the same period when suspended from either knife-edge, then each is located at the center of oscillation of the other, and the distance between them must be L, the length of the equivalent simple pendulum. The value for g follows from Eq. (6) or Eq. (18).

Ballistic pendulum. This is a device to measure the momentum of a bullet. The pendulum bob is a block of wood into which the bullet is fired. The bullet is

Ratio of the period for amplitude θ_0 to the period for infinitesimal amplitude

θ_0	$T(\theta_0)/T(O)$
0°	1.0000
20°	1.0077
40°	1.0313
60°	1.0732
80°	1.1375
100°	1.2322
120°	1.3729
140°	1.5944
160°	2.0075
180°	∞

stopped within the block and its momentum transferred to the pendulum. This momentum is determined from the amplitude of the pendulum swing. *See* BALLISTICS.

Spherical pendulum. This is a simple pendulum mounted on a pivot so that its motion is not confined to a plane. The bob then moves over a spherical surface. A Foucault pendulum is a spherical pendulum suspended so that its plane of oscillation is free to rotate. Its purpose is to demonstrate the rotation of the Earth. If such a pendulum were mounted at the North Pole, the rotation of the Earth under the pendulum would make it appear to a terrestrial observer that the plane of the pendulum's motion rotated 360° once every day. The plane of motion of a Foucault pendulum set up at a lower latitude rotates at a reduced rate, proportional to the sine of the latitude. *See* FOUCAULT PENDULUM.

Torsional pendulum. Despite its name, a torsional pendulum is not a pendulum. It is an example of a torsional harmonic oscillator, consisting of a disk or other body of large moment of inertia mounted on one end of a torsionally flexible rod. The other end of the rod is held fixed. If the disk is twisted and released, the torsional pendulum oscillates harmonically. Gravitation plays no part in its motion. *See* CLOCK (MECHANICAL); DIMENSIONAL ANALYSIS; SCHULER PENDULUM. Joseph M. Keller

Bibliography. A. P. Arya, *Introduction to Classical Mechanics*, 2d ed., 1997; V. D. Barger and M. G. Olsson, *Classical Mechanics: A Modern Perspective*, 2d ed., 1995; J. L. Meriam and L. G. Kraige, *Engineering Mechanics*, 4th ed., 1998; L. Shepley and R. Matzner *Classical Mechanics*, 1991.

Penicillin

One of a series of beta-lactam antibiotics, all of which possess a four-ring beta-lactam structure fused with a five-membered thiazolidine ring. These antibiotics are nontoxic and kill sensitive bacteria during their growth stage by the inhibition of biosynthesis of their cell wall mucopeptide. *See* PLANT CELL.

The antibiotic properties of penicillin were first recognized by A. Fleming in 1928 from the serendipitous observation of a mold, *Penicillium notatum*, growing on a petri dish agar plate of a staphylococcal culture. The mold produced a diffuse zone which lysed the bacterial cells. Commercial production of penicillin came from the pioneer work of E. Chain and H. W. Florey in 1938, first in England and then in the United States, where it was developed from an academic project into a collaborative war effort between industry and government research. Penicillin (as penicillin G) was made available to the allied troops in Europe in the latter part of the World War II.

Penicillin is produced from the fungal culture *P. chrysogenum* that was isolated from a moldy cantaloupe. The pharmaceutical industry uses highly mutated strains cultured in large, highly aerated, stirred tank fermentors controlled to optimize antibiotic production and to efficiently use raw materials such as corn syrup, corn steep liquor, and cottonseed flour. Temperature, pH, dissolved oxygen, soluble nitrogen and ammonia levels, and sugar feed rates are important control factors. The biosynthesis of penicillin is known in detail, and all the enzymes involved in the formation of this secondary metabolite have been isolated and purified.

Penicillin G, the most commonly fermented penicillin, is produced by the addition of a precursor phenylacetic acid to the growing culture. Use of phenoxyacetic acid as a precursor produces penicillin V. Both penicillins are recovered by extraction into organic solvents at acid pH and precipitation as their potassium or sodium salt.

Penicillin G is generally given by injection against penicillin-sensitive streptococci such as pneumococci (meningitis), and in treatment of endocarditis and gonorrhea. Penicillin G procaine salt is used for intramuscular injection to provide quick distribution of the antibiotic. Penicillin G benzathine salt is used for slower release. Penicillin G can be given in combination with probenecid, a compound which delays urinary excretion. Penicillin V is acid stable and is usually given orally. It is effective in the treatment of upper respiratory infections and periodontal work. *See* GONORRHEA; MENINGITIS.

More effective semisynthetic penicillins are produced by coupling different side chains to the active penicillin nucleus 6-aminopenicillanic acid. This nucleus is produced from either penicillin G or penicillin V by using specific immobilized enzymes. The various side chains confer different antibiotic properties on the penicillin.

The fermented penicillin G and penicillin V are susceptible to destruction by an enzyme (beta-lactamase) produced by certain bacteria which makes them resistant. The penicillins methicillin, oxacillin, nafcillin, cloxacillin and dicloxacillin all are resistant to hydrolysis by beta-lactamases and are used to treat staphylococcal infections. Cloxacillin and dicloxacillin are used orally. Ampicillin and amoxicillin are penicillins with extended spectra as they are effective against many gram-negative bacteria. They are used mainly orally against streptococci and other respiratory-tract pathogens, including *Haemophilus influenzae*, in the treatment of sinusitis, bronchitis, and pneumonia. They are used extensively in pediatrics and against *Listeria monocytogenes* and *Salmonella* spp. *See* HAEMOPHILUS; SALMONELLOSES; STREPTOCOCCUS.

Hetacillin, pivampicillin, and bacampicillin are effective prodrug forms of ampicillin. Amoxicillin is formulated with a beta-lactamase inhibitor clavulanic acid as Augmentin for the treatment of ampicillin-resistant infections. Carbenicillin, ticarcillin, azlocillin, mezlocillin, and piperacillin were developed to combat pseudomonad, enterobacter, and serratia infections resistant to ampicillin. They are used mainly to treat urinary *Pseudomonas aeruginosa* infections, sepsis from burns, and chronic infections of the respiratory tract. *See* ANTIBIOTIC; DRUG RESISTANCE; PSEUDOMONAS. D. A. Lowe

Bibliography. I. Kawamoto, *Antibiotics I: Lactams and Other Microbial Agents*, 1992; S. F. Queener, J. A. Webber, and S. W. Queener (eds.), *Beta-Lactam Antibiotics for Clinical Use*, 1986.

Penis

The male organ of copulation, or phallus. In mammals the penis consists basically of three elongated masses of erectile tissue. The central corpus spongiosum (corpus urethrae) lies ventral to the paired corpora cavernosa. The urethra runs along the underside of the spongiosum and then normally rises to open at the expanded, cone-shaped tip, the glans penis, which fits like a cap over the end of the penis. Loose skin encloses the penis and also forms the retractable foreskin, or prepuce. The organ is held firmly in place by fibrous tissue and ligaments that bind it to the posterior surface of the pubic arch. In some mammals, although not in man, there may be spines on the glans penis, or there may be a bone (the baculum) within the penis; both of these characters are useful in classifying certain types of mammals.

Erection of the penis is caused by nervous stimulation resulting in engorgement of the spiral helicine arteries and the plentiful venous sinuses of the organ. In most mammals other than Primates the penis is retracted into a sheath when not in use.

In submammalian forms the penis is not as well developed. Crocodilians, turtles, and some birds have a penis basically like that of mammals, although not as well developed, lying in the floor of the cloaca. When erected, it protrudes from the cloaca and functions in copulation. In most birds the penis is rudimentary. Snakes and lizards have paired hemipenes, which are quite different organs serving the same function. They are sacs in the posterior wall of the cloaca which can be everted and used as intromittent organs. They are retracted by muscles and lack erectile tissue. The surface of the hemipenis is frequently covered with spines and ornamentation important in classifying these animals.

Other vertebrates lack a penis, although various functionally comparable organs may be developed such as the claspers on the pelvic fins of sharks and the gonopodia on the anal fins of certain teleost fishes. *See* COPULATORY ORGAN. Thomas S. Parsons

Pennatulacea

An order of the cnidarian subclass Alcyonaria (Octocorallia), commonly called the sea pens. These animals lack stolons and live with their bases embedded in the soft substratum of the sea. The colony consists of a distal rachis bearing many polyps and a polypless proximal peduncle, whose terminal end sometimes expands to form a bladder. The colony of *Pennatula* looks like a feather (**Fig. 1***a*), being formed of numerous secondary polyps which arise from leaf-shaped lateral expansions of the very elongated primary axial or terminal polyp. In the other form of colony the polyps arise directly from the primary one, as in *Renilla* and *Veretillum* (Fig. 1*b*

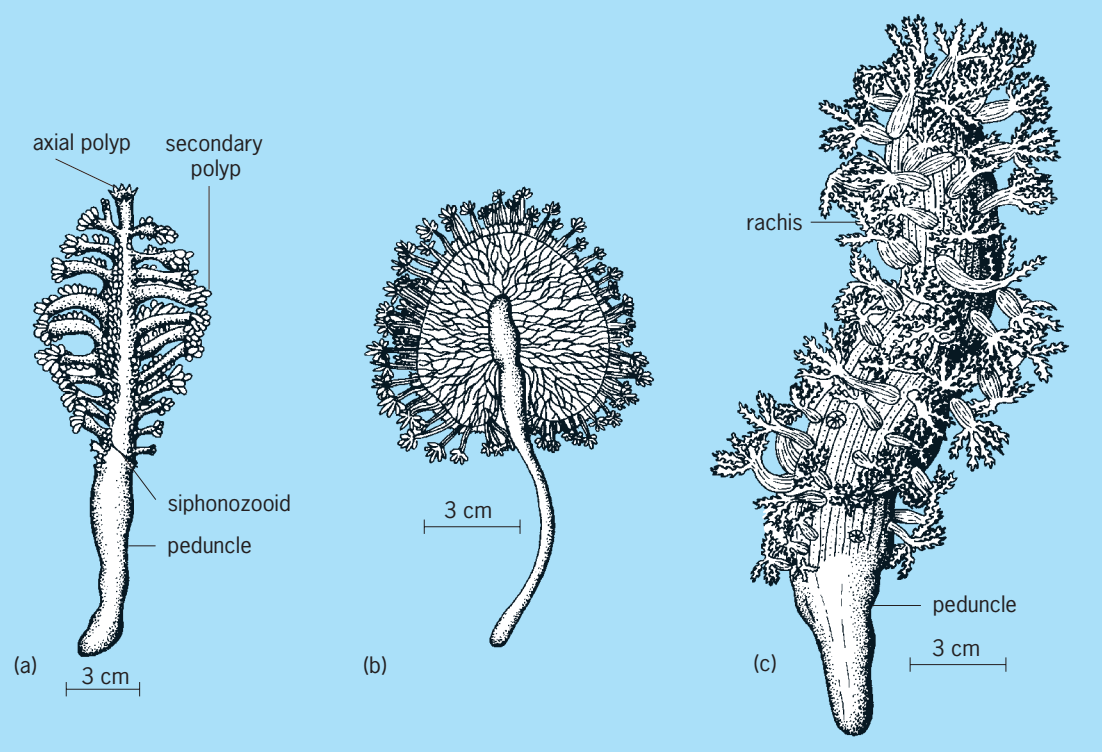

Fig. 1. Pennatulacea colonies. (*a*) Young colony of *Pennatula phosphorea*. (*b*) *Renilla amethystina*. (*c*) *Veretillium cynomorium*.

and *c*) and *Cavernularia* (**Fig. 2**). The colony has a horny unbranched axial skeleton composed of pennatulin and some calcium carbonate and phosphate, but *Cavernularia*, including the luminous species, has a rudimentary one and *Renilla* lacks a skeleton. Smooth, calcareous spicules in oval, rod, spindle, and needle shapes are found in the coenenchyme.

The gastrovascular cavity of the primary polyp is divided into four principal canals which run longitudinally around the axial skeleton through a spongy tissue containing solenial networks (**Fig. 3**). The cavity of each secondary polyp is interconnected by such canal systems. The musculatures are strongly developed, especially in the peduncle. Longitudinal muscles contract the body, and circular ones serve to force the water out of the main canals. In some littoral species the polyp contracts or expands according to the tidal or diurnal and nocturnal rhythms.

All Pennatulacea are dimorphic. The two kinds of polyps resemble those of the other orders. In addition to the genera indicated above, the order also includes *Virgularia*, *Umbellula* (**Fig. 4**), *Pteroeides*, and others. *See* OCTOCORALLIA (ALCYONARIA).

Kenji Atoda

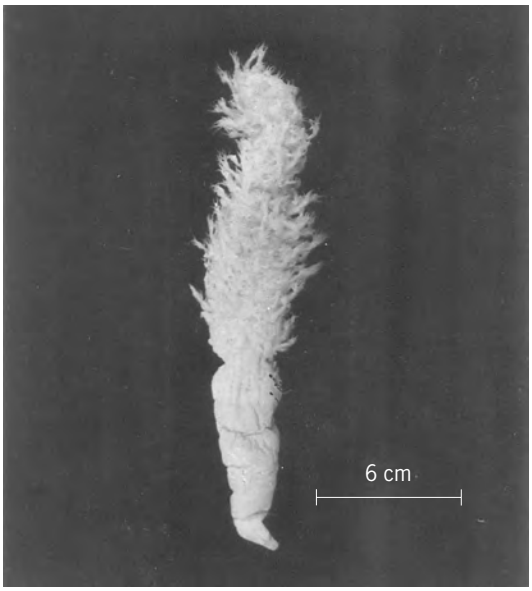

Fig. 2. ***Cavernularia habereri*** **(preserved state).**

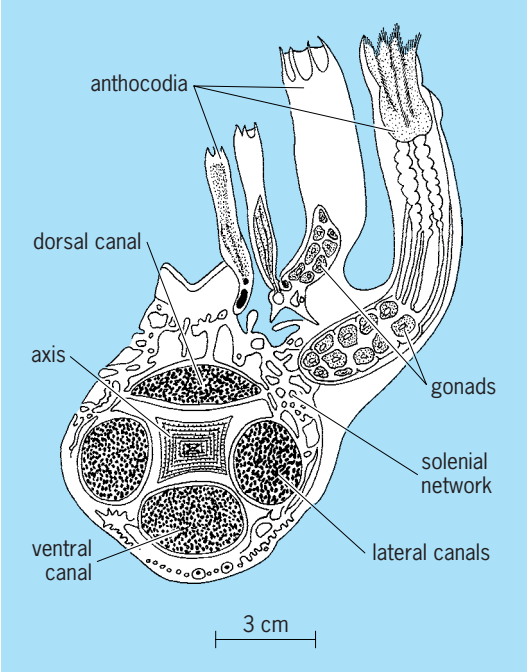

Fig. 3. Cross section of ***Funiculina quadrangularis*****.**

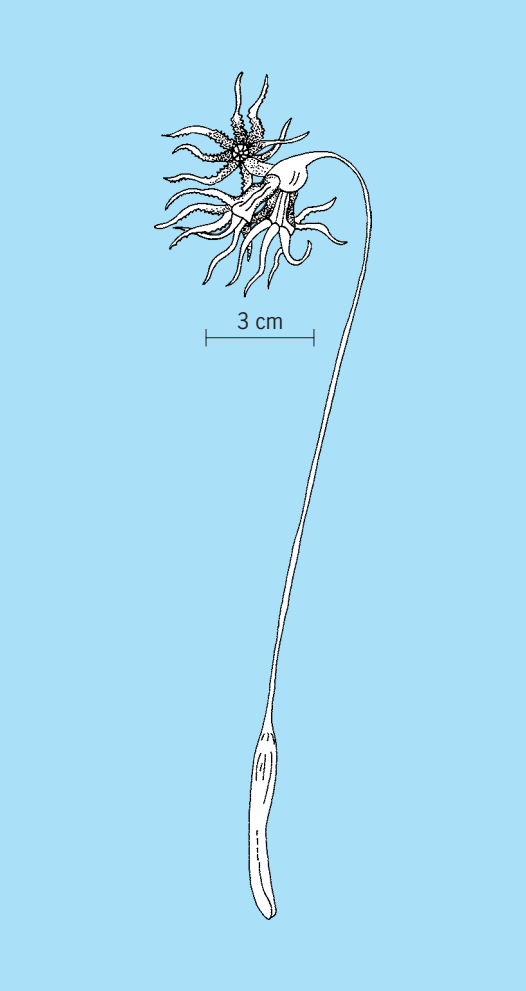

Fig. 4. ***Umbellula encrinus*****.**

Pennsylvanian

A major division of late Paleozoic time, considered either as an independent period or as the younger subperiod of the Carboniferous. In North America, the Pennsylvanian has been widely recognized as a geologic period and derives its name from a thick succession of mostly nonmarine, coal-bearing strata in Pennsylvania. Radiometric ages place the beginning of the period at approximately 320 million years ago and its end at about 290 million years ago. In northwestern Europe, strata of nearly equivalent age are commonly designated as Upper Carboniferous and in

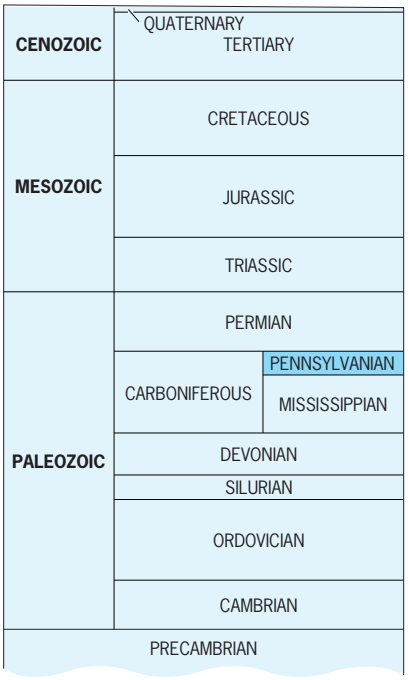

eastern Europe as Middle and Upper Carboniferous. *See* CARBONIFEROUS.

In North America, the Pennsylvanian Period was characterized by the progressive growth and enlargement of the Alleghenian-Ouachita-Marathon orogenic belt, which formed as the northwestern parts of the large continent Gondwana (mainly northwestern Africa, the area that is now Florida, and northern South America) collided against and deformed the eastern and southern parts of the North American continent.

During the Pennsylvanian and continuing into the early part of the Permian, the orogeny generally progressed with a large jawlike rotational movement, closing from northeast to southwest. However, the outlines of the two cratons on opposite sides of this jaw collided in such a way that the continents were irregularly forced against each other. This resulted in the reactivation of many older structures and, more importantly, in forming numerous new northwest trending shear folds and associated tear faults. These new intracratonic structures comprise important large graben basins, half-graben basins, and horst uplifts and half-horst uplifts that were particularly well developed as intermontane troughs and block-faulted mountains in Oklahoma, Texas, New Mexico, Colorado, Arizona, Utah, Nevada, Wyoming, and Montana. *See* CRATON; FAULT AND FAULT STRUCTURES; GRABEN; HORST; OROGENY.

The uplifted orogenic belt contributed great amounts of clastic sediments onto the edge of North America either as terrestrial clastic deposits where the adjacent colliding blocks were uplifted, or as dark marine turbidites and other marine clastics where one side of the colliding blocks was tectonically depressed beneath the other side. This contrast in the type of deposits resulted because of the irregularies along the colliding margins and the wrinkling and tearing of the North America craton. This is well shown by the mainly nonmarine fluvial clastics that form the Pennsylvanian deposits in the Appalachians sector as compared to the deep-water marine turbidites in the Ouachita and Marathon sectors which formed in fore-deep basins just ahead of the advancing orogenic belt. The intracratonic horsts and half-horsts were in many places sources of large amounts of sediments for adjacent graben basins and half-graben basins. Some of these basins had marine deposits, others had evaporite deposits, and still others were essentially nonmarine. Yet others, such as the Permian Basin of western Texas and southern New Mexico, were more complicated as two rapidly subsiding subbasins, the Midland and Delaware basins, were separated by a much less rapidly subsiding horstlike structure, the Central Basin Platform, and surrounded on the east, north, and west by relatively stable cratonic blocks. *See* BASIN; TURBIDITE.

Much of North America remained a stable, low-lying cratonic platform during the Pennsylvanian and was covered by a relatively thin veneer of shallow-water marine carbonates and marine and nonmarine clastic sediments. These were deposited as sea level repeatedly rose and fell as the polar glaciers of southern Gondwana contracted and expanded. *See* CARBONATE MINERALS; MARINE SEDIMENTS.

In Pennsylvania and along the western part of the present Allegheny Plateau, Pennsylvanian strata are predominantly nonmarine deposits made up of channel sandstones, floodplain shales, siltstones, sandstones, and coals. Farther to the west in the Illinois-Kentucky basin, the nonmarine clastics intertongue through a complex series of distributary deltas with shallow marine beds. In western Missouri and eastern Kansas, marine beds form the greater proportion of the sediments.

During the Pennsylvanian, the western margin of cratonic North America was much farther east of the present coast and extended north-northeastward from southeastern California across eastern Nevada, eastern Idaho, and western Montana into easternmost British Columbia and Alberta. This margin was tectonically unstable and was being pushed westward into an oceanic basin during the Pennsylvanian Period. Later Permian, Mesozoic, and Cenozoic orogenic movements and their structures obscure much of this Pennsylvanian history; however, it is possible to identify several basins and uplifts along the old cratonic margin. Locally, basins on this margin include complete successions of marine sediments across the Mississippian and Pennsylvanian boundary, as at Arrow Canyon near Las Vegas, Nevada, where a stratigraphic reference section for the lower boundary of the Pennsylvanian was established in 1996.

During Pennsylvanian time, the paleoequator extended across North America from southern California to Newfoundland, through northwestern Europe into the Ukrainian region of eastern Europe, and across parts of northern China. This was a time of extensive coal deposition in a tropical belt that appears to have included areas from 15 to 20° north and south of the paleoequator. Coal of this age is

abundant and relatively widespread and has great economic importance. Its geographic distribution strongly influenced the location and development of the Industrial Revolution in the mid-1800s. These coals supplied the energy for the rapid increase in the standard of living in western Europe and North America during the late 1800s, were one of the major prizes of the Franco-Prussian war, and influenced many military objectives of World Wars I and II. As a source of readily available combustible fuel, coals of Pennsylvanian age remain of major importance. *See* COAL.

Petroleum is commonly trapped in nearshore marine deposits of Pennsylvanian age, particularly in carbonate banks near the edge of shelves, in longshore bars and beaches, in reefs and mounds, and at unconformities associated with transgressive-regressive shore lines. Many of these traps contribute significantly to petroleum production and include the Horseshoe atoll of the Texas Panhandle and carbonate banks around the Midland basin in Texas. *See* PETROLEUM.

Pennsylvanian paleogeography changed significantly during the period as the supercontinent Pangaea gradually was formed by the joining together of Gondwana and Laurasia. North America and northern Europe, which had been combined into the continent Laurasia since the late Silurian, and South America and northwestern Africa, which formed the northern part of the continent of Gondwana, came together along the Ouachita–Southern Appalachian–Hercynian geosyncline. The result was an extensive orogeny, or mountain-building episode, which supplied the vast amounts of sediments that make up most of the Pennsylvanian strata in the eastern and midwestern parts of the United States.

Evidence in the form of well-developed tree rings, less diverse fossil floras and faunas, and glacial deposits indicates that temperate and glacial conditions were common in nonequatorial climatic belts during Pennsylvanian time. Climatic fluctuations during the period caused significant increases and decreases in the amount of water that was temporarily stored in the glaciers in Gondwana and contributed to eustatic changes of sea level. The Early Pennsylvanian (Morrowan Stage) was a time of generally low sea levels, and marine sediments are preseved only along lower parts of the cratonic margins and marginal basins. Sea levels rose high enough by the middle of the Atokan Stage (early part of the Middle Pennsylvanian) to extend onto, and across, many of the cratonic platforms. At least 60 of these major cycles of sea-level fluctuations repeatedly exposed much of the cratonic areas during Pennsylvanian time so that on the cratons these strata have many unconformities. *See* CYCLOTHEM; PLATE TECTONICS; UNCONFORMITY.

Charles A. Ross; June R. P. Ross

Bibliography. E. D. McKee et al., Paleotectonic Investigations of the Pennsylvanian System in the United States, *USGS Prof. Pap.*, no. 853 (3 pts.), 1975; C. A. Ross and J. R. P. Ross, Biostratigraphic zonation of late Paleozoic depositional sequences, in C. A. Ross and D. Haman (eds.), Timing and Depositional History of Eustatic Sequence: Constraints on Seismic Stratigraphy, *Cushman Found. Foram. Res. Spec. Publ.*, no. 24, pp. 151–168, 1987; C. A. Ross and J. R. P. Ross, Late Paleozoic sea levels and depositional sequences, in C. A. Ross and D. Haman (eds.), Timing and Depositional History of Eustatic Sequence: Constraints on Seismic Stratigraphy, *Cushman Found. Foram. Res. Spec. Publ.*, no. 24, pp. 137–149, 1987; J. W. Skehan et al., The Mississippian and Pennsylvanian (Carboniferous) Systems in the United States, *USGS Prof. Pap.*, no. 1110-A–DD, pp. A1–DD16, 1979.

Pentamerida

An extinct order of Brachiopoda that lived from the Lower Cambrian to Upper Devonian (540 to 375 million years ago), with a peak during the Silurian Period. Pentamerides (**Figs. 1** and **2**) are the largest brachiopods in the Lower Paleozoic, with adults ranging from 1 to 18 cm (0.4 to 7 in.) in length, and the order forms part of the subphylum Rhynchonelliformea (previously termed the Articulata). *See* BRACHIOPODA; RHYNCHONELLIFORMEA.

Characteristics. The two valves, always composed of calcite ($CaCO_3$), are convex, with a larger ventral valve and a smaller dorsal valve, articulated at the posterior and which opened anteriorly during feeding. Externally, most pentamerides are smooth, apart from concentric growth line increments, but some have radial ribs (costae), such as the abundant Silurian *Kirkidium* (Fig. 2), which may be smooth or sinuous in cross section. From comparison with living brachiopods and by the analysis of growth lines, it can be deduced that the average lifespan was 8

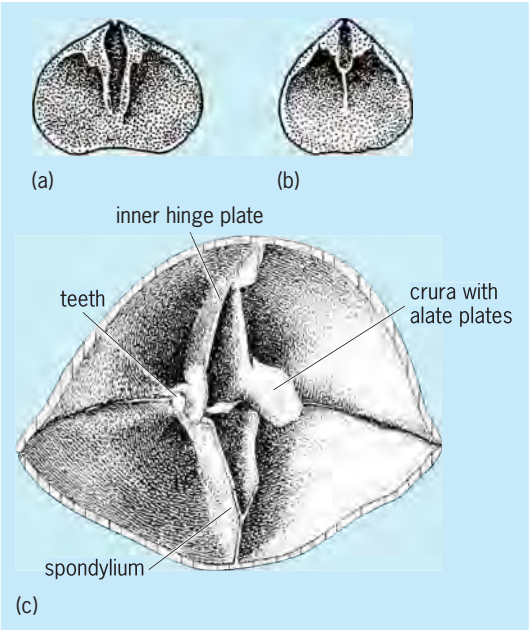

Fig. 1. Interior views of pentamerides. (*a*) *Porambonites*, dorsal interior. (*b*) *Camerella*, ventral interior. (*c*) Interior of the posterior end of *Parastrophina* with dorsal valve above and ventral valve below.

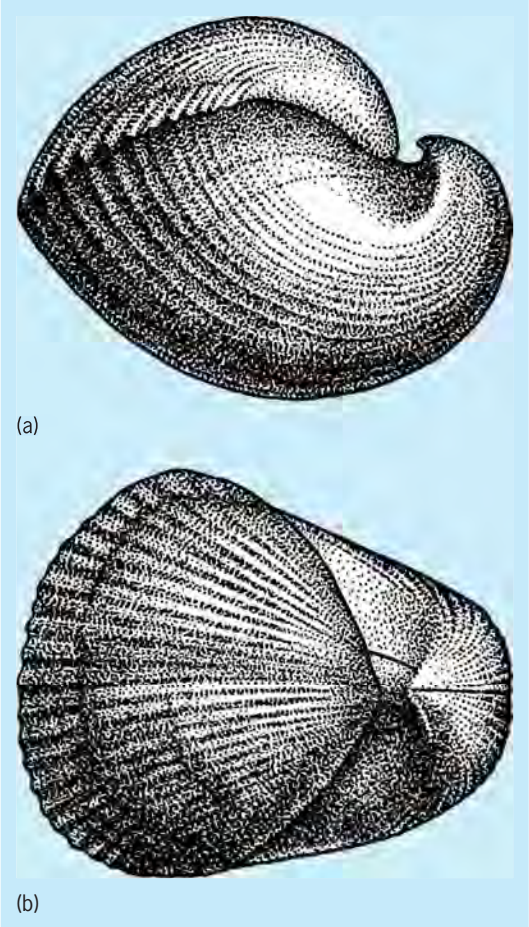

Fig. 2. ***Kirkidium*, of the Silurian, (*a*) lateral and (*b*) dorsal views.**

to 9 years. Internally, the space between the two pentameride valves is divided by variably developed walls (septa), with at least one septum in the ventral valve and two in the dorsal valve (although a few syntrophiidine genera have many more, up to a maximum of 12 septa), thus dividing the internal space into several divisions. In some genera, the septa unite anteriorly; and in the posterior part of the shell, the ventral valve septum divides to form a platform termed the spondylium (Fig. 1*b* and *c*), which was used as an attachment area for the muscles which opened and closed the valves.

The form and disposition of the soft parts are not well known. However, by comparison with other brachiopods living today, it is probable that the fleshy, food-gathering lophophore would have occupied much of the internal space between the valves, leaving the main body organs (mouth, stomach, reproductive gonads, etc.) enclosed within an epithelium near the posterior of the shell interior. During the earliest growth stages, all pentamerides had a functional fleshy pedicle which emerged through an opening in the pedicle valve to enable the animal to be attached to the substrate. However, in the vast majority of pentameride families and genera, this functional pedicle did not persist into adult life, and the pedicle atrophied while the animal was relatively young, with the pedicle opening becoming closed and filled with secondary calcite. Most pentamerides lived vertically as adults, with the posterior ends of both valves settled downward into the substrate. These posterior parts of the valves were much thickened by secondary calcite deposition on the valve interiors, and the increased posterior weight enabled adults to maintain stability with the posterior downward. Many pentamerides, including *Pentamerus* itself, which lived during the Early Silurian, occurred in enormous quantities, dominating the midshelf communities, which are often named after them. In some cases, the pentameride shells were packed together as tightly as modern oyster beds and thus occur in rock-forming abundance in many localities and geological horizons. *See* BIOSTROME; LOPHOPHORE.

Classification. Within the order Pentamerida, there are currently 214 genera, which have been described from Paleozoic rocks of all continents. The order is subdivided into the suborder Syntrophiidina, with 85 genera ranging from the Early Cambrian (probably Fortunean and certainly Branchian) to the Early Devonian (Pragian), and the suborder Pentameridina, with 129 genera ranging from the Late Ordovician (Ashgill) to the Late Devonian (Frasnian). The acme of the syntrophiidines was in the Early Ordovician; for example, in the paleocontinents of Siberia (central Siberia, Mongolia, and beyond), Laurentia (most of North America today), and Baltica (northern Europe). In Baltica, the endemic Early Ordovician syntrophiidine *Lycophoria* occurred in great abundance, not only dominating its shallow to midshelf community but also providing a useful indicator to the margins of the paleocontinent. The acme of the pentameridines was in the Middle to Late Silurian; however, since the paleocontinents were closer together than in the Ordovician [and thus the spat (brachiopod spawn) could cross the intervening deep oceans], there were fewer endemic genera. *See* ORDOVICIAN; PALEOGEOGRAPHY.

L. Robin M. Cocks

Bibliography. S. J. Carlson et al., Pentamerida, in R. L. Kaesler (ed.), *Treatise on Invertebrate Paleontology: Part H, Brachiopoda Revised*, Geological Society of America and Kansas University Press, 2002.

Pentastomida

A subclass of bloodsucking crustacean arthropods, parasitic in the respiratory organs of vertebrates, that frequently are referred to as the Linguatulida or tongue worms. The adults are vermiform, with a short cephalothorax and an elongate, annulate abdomen that may be cylindrical or flattened. On the ventral surface the cephalothorax bears two pairs of hooklike, retractile claws. Respiratory, circulatory, and excretory organs are absent. The digestive tract is a straight tube with cuticle-lined anterior and posterior ends. The nervous system consists of subesophageal ganglia and a circumesophageal ring.

Paired nerves extend from the ganglia to cephalic structures, and a midventral pair extends posteriad the length of the body. The sexes are separate; the females are two to five times the size of the males. The male genital pore is medial and ventral, at the anterior border of the abdomen. The female genital pore is anterior in the Cephalobaenida but located at or near the posterior end of the body in the Porocephalida.

The subclass is divided into two orders: the Cephalobaenida, a more primitive group and the Porocephalida, a more specialized one. The first has six-legged larvae and the other four-legged larvae. The mitelike form of the larvae, with short stumpy legs, demonstrates relationship to the arthropods. Characteristic arthropod features include the presence of (1) jointed appendages in the larvae; (2) stigmata or breathing pores in the body wall; (3) specialized reproductive organs, especially those of the male; and (4) ecdysis or molting of larvae and nymphs. More than 50 species have been described.

These animals are parasitic, both as larvae and as adults, in a wide variety of vertebrates, chiefly in tropical regions. With the exception of two genera—*Reighardia*, whose only species occurs in birds, and *Linguatula*, with species only in mammals—all other pentastomes are parasites of reptiles. Adults occur in snakes, lizards, turtles, and crocodiles. Larval stages are encysted in fishes, amphibians, reptiles, birds, and mammals. Life cycles are not well known; usually an intermediate or transfer host harbors the larval stage, but development may be completed in a single host. Late larval stages are sometimes described as nymphs. Fertilization is internal, and the eggs contain fully formed larvae when passed from the uterus. The eggs are ingested by a vertebrate, and the larva emerges in its digestive tract. The larva bores through the tissues and soon encapsulates. In the capsule it undergoes metamorphosis and attains the adult form, except that it has a circle of spines around each annulus and is not sexually mature. Some authors believe that the nymph leaves the cyst and migrates to another location, where it encysts again. When liberated in the intestine of the final host, the larvae migrate to the lungs or air passages, where they become sexually mature. Human infection occurs frequently in Africa and Asia, where humans are accidental intermediate hosts of the nymphal form. The liver is a common site of infection, and large numbers of larvae may produce serious and even fatal effects. *See* ARTHROPODA; CEPHALOBAENIDA; CRUSTACEA; POROCEPHALIDA.

Horace W. Stunkard

Bibliography. H. R. Hill, Annotated bibliography of the Linguatulidae, *Bull. S. Calif. Acad. Sci.*, 47:56–73, 1948; R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; L. W. Sambon, A synopsis of the family Linguatulidae, *J. Trop. Med. Hyg.*, 25:188–208, 391–428, 1922; C. W. Stiles and A. Hassal, Key-catalogue of the Crustacea and arachnoids of importance in public health, *Hyg. Lab. Bull.*, no. 148, pp. 197–289, 1927.

Pepper

The garden pepper, *Capsicum annuum* (family Solanaceae), is a warm-season crop originally domesticated in Mexico. It is usually grown as an annual, although in warm climates it may be perennial. This species includes all peppers grown in the United States except for the "Tabasco" pepper (*C. frutescens*), grown in Louisiana. Other cultivated species, *C. chinense*, *C. baccatum*, and *C. pubescens*, are grown primarily in South America. Some 10–12 strictly wild species also occur in South America. Peppers are grown worldwide, especially in the more tropical areas, where the pepper is an important condiment. *Piper nigrum*, the black pepper, a tropical climbing vine, is botanically unrelated to the common pepper.

Pepper. (*a*) Sweet pepper. (*b*) Long Red cayenne pepper plant. (*W. Atlee Burpee Co.*)

Cultivation and harvesting. Propagation is by planting seed directly in the field or by plants started in greenhouses or hotbeds and transplanted to the field after 6–10 weeks. Field spacing varies: plants 12–24 in. (30–60 cm) apart in 30–36 rows are common. Long warm, but not hot, seasons favor yield and quality.

Green peppers are harvested when fully developed but before appearance of red or yellow color, about 60–80 days after transplanting. Hot peppers, picked red ripe, require about 70–90 days. Florida, California, North Carolina, and New Jersey, in that order, are the most important growing areas.

Uses. Sweet (nonpungent) peppers (**illus.** *a*), harvested fully developed but still green, are widely used in salads or cooked with other foods. Popular varieties (cultivars) are California Wonder and Yolo Wonder. Perfection pimento, harvested red ripe, is used for canning. Paprika is made from ripe red pods of several distinct varieties; the pods are dried and ground. *See* PAPRIKA; PIMENTO.

Hot (pungent) peppers (illus. *b*) may be picked before ripening for fresh use, pickles, or canning. "Wax" peppers are yellow instead of green before ripening. Ripe pods dried and ground are used to make red pepper powder or mixed with spices to make chili powder. Cayenne and Anaheim Chili, respectively, are the major varieties for grinding. Hot sauce is made from the fresh fruit of the Tabasco pepper ground with vinegar. Pungency is due to the compound capsaicin, located in fragile glands on the internal partitions or "ribs" of the fruit. Seed and wall tissue are not pungent.

The ripe color of most varieties is red, a few varieties are orange-yellow, and in Latin America brown-fruited varieties are common. Nutritionally, the mature pepper fruit has three to four times the vitamin C content of an orange, and is an excellent source of vitamin A. *See* ASCORBIC ACID; VITAMIN A.

Paul G. Smith

Diseases. Virus diseases are the most common field problem of peppers throughout the world. Four potyviruses—pepper mottle, pepper veinal mottle, potato Y, and tobacco etch—are particularly troublesome. Potato Y virus is more widely distributed and known to be present on pepper in Africa, Asia, Australia, the Americas, and Europe, while pepper veinal mottle is restricted to Africa, and pepper mottle to Central and North America. Tobacco etch is known to occur on pepper in North America, but its presence elsewhere has not been established. Early infection with each of these viruses causes stunted plants with vein-banding mottle leaf symptoms and fruits reduced in size and severely disfigured in shape, with streaked ovulary walls.

Tobacco mosaic virus causes only a transient, chlorotic mottle and some reduction in yield, but early infection with cucumber mosaic may result in multiple symptoms that may include severe stunting, systemic granular mottling to shoestring leaf malformation, and extreme reduction in size and number of fruits set. Oak-leaf, concentric line patterns in leaves are caused by alfalfa mosaic infection. Tomato spotted wilt virus causes conspicuous chlorotic and necrotic ring and line patterns on leaves and fruits, with progressive necrosis of one or more growing tips.

Delay in field incidence of potato Y virus has been reported from repeated application of oil sprays. Cultivars with resistance to three of the potyviruses and tobacco mosaic, namely, Yolo Y, Florida VR-2, and Delray Bell, have been developed at the Florida Agricultural Experiment Station.

Bacterial leaf spot (caused by *Xanthomonas vesicatoria*) results in dark brown, translucent lesions that coalesce and cause defoliation under warm, rainy, and windy conditions. Pathotypes of the bacterium have been distinguished. Copper fungicides provide adequate control, except under conditions favorable for the disease.

Leaf spot caused by *Cercospora capsici* [mostly circular lesions to 0.4 in. (1 cm) in diameter with gray-white centers and dark margins] may cause severe defoliation if lesions are abundant. Carbamate fungicides have been recommended for control.

Phytophthora capsici causes a rapid, wet, systemic rot of plants of all ages. The disease spreads quickly in wet weather, when control is practically impossible, but the disease usually disappears when dry weather returns.

Bacterial canker (caused by *Corynebacterium michiganense*) has been found to infect pepper naturally in Australia, California, and Israel. *See* PLANT PATHOLOGY.

A. A. Cook

Peppermint

The mint species *Mentha piperita* (family Lamiaceae), a sterile interspecific hybrid believed to have occurred in nature from the hybridization of fertile *M. aquatica* with fertile *M. spicata*. Peppermint (see **illus.**) grows in wet, marshy soil along streams in Europe at latitudes of 40–55°. Overwintering stolons, the leafless underground stems, produce new plants, giving the species a perennial habit.

Peppermint (*Mentha piperita*). (*Photo by Michael Thompson/U.S. Department of Agriculture*)

Strains. Numerous clonal strains brought by pioneers for medicinal use in colds and stomach disorders or as flavoring herbs have persisted in North America in wet areas for more than 100 years. These and a few strains cultivated prior to 1890 are frequently called American peppermint to distinguish them from the single clonal strain called Black Mitcham peppermint, obtained from Mitcham, England, near London. Due to its excellent flavor and aroma, Black Mitcham, with a red anthocyanin stem color, became the exclusive cultivar source of United States peppermint oil from 1910 to 1972. A similar, green-stemmed strain called White Mitcham proved less winterhardy and is no longer cultivated. Irradiation breeding during 1955–1972 produced two *Verticillium*-resistant varieties. These were vegetatively propagated and distributed to farmers by certified seed-producing agencies at the state universities of the mint-producing states. Mitcham peppermint is also grown in Bulgaria, Italy, and Russia and to a limited extent in other countries but is restricted by its long-day–short-night photoperiod to an area north of 40° latitude that is not subject to severe mid-June freezes.

Cultivation and harvesting. In mint farming the dormant stolons are scattered by hand in 10–12-in.-deep (25–30-cm) furrows and covered with soil in rows 36–38 in. (90–95 cm) apart that can be cultivated for weed control with a maize cultivator. In the United States machine planting and preemergence application of terbicil herbicides (such as Sinbar) for weed control are standard practice to minimize labor costs. Peppermint has rather high water requirements; hence supplementary overhead or sprinkler irrigation may be used on the organic soils of the Midwest area (southern Michigan, northern Indiana, Wisconsin) and must be used on alluvial soils of the Willamette Valley of western Oregon and the Madras area of central Oregon. In all these areas the first-year crop is row mint, and all subsequent crops are meadow mint in a solid stand, whereas rill or ditch irrigation necessitating row culture has generally been used on the volcanic soil of eastern Washington and in part on alluvial soils of the Snake River Valley of Idaho.

The crop is mown and windrowed, as hay, about the time that peak vegetative growth is attained or flower buds are beginning to form (July 15–September 10). The windrowed, partially dried hay is picked up by a loader that chops the hay and blows it into a mobile metal distillation tank mounted on four wheels or on a dump truck chassis. *See* AGRICULTURAL SOIL AND CROP PRACTICES.

Peppermint oil. The oil of commerce is obtained by steam distillation from the partially dried hay. Steam under pressure traverses the chopped hay in the mobile tank, removing the oil, and is then cooled in a condenser. The oil floats on the water in a receiving receptacle and is removed to metal drums for storage and sale. Stills are expensive since automatic valves and stainless steel construction are needed. The oil is complex, having 3 major constituents and over 80 minor ones that are necessary to the flavor and aroma. Oil yields in pounds per acre are generally 40–65 (45–73 kg/hectare) in the Midwest, 65–85 (73–96 kg/hectare) in the Willamette Valley and Idaho, and 85–110 (96–125 kg/hectare) in eastern Washington. The United States produces 50,000–80,000 acres (202–324 km^2) of peppermint. The main uses of peppermint oil are to flavor chewing gum, confectionary products, toothpaste, mouthwashes, medicines, and as a carminative in certain medical preparations for the alleviation of digestive disturbances. *See* ESSENTIAL OILS; MENTHOL; SPICE AND FLAVORING.

Diseases. Principal diseases are Verticillium wilt, which occurs everywhere, and Puccinia rust in the western United States. In the Willamette Valley, where rust is severe, fields are flamed with a propane burner to interrupt the rust cycle in early spring as plants are emerging. Genetic control of Verticillium wilt is through use of the new varieties. *See* PLANT PATHOLOGY.

Merritt J. Murray

Bibliography. J. E. Robbers and V. E. Tyler, *Tyler's Herbs of Choice: The Therapeutic Use of Phytomedicinals*, 2d ed., Haworth Press, 1999.

Pepsin

A proteolytic enzyme found in the gastric juice of mammals, birds, reptiles, and fish. It is formed from a precursor, pepsinogen, which is found in the stomach mucosa. Pepsinogen is converted to pepsin either by hydrochloric acid, naturally present in the stomach, or by pepsin itself. A pepsin inhibitor, peptide, remains attached to the pepsin molecule at pH values above 5.0, preventing activation of the enzyme. The approximate molecular weight of pepsin is 34,500 and that of pepsinogen is 42,600. Maximum activity occurs at pH 1.8, and broad specificity for peptide bonds is demonstrated. *See* ENZYME; PH.

Pepsin is prepared commercially from the glandular layer of fresh hog stomachs. Official National Formulary (NF) pepsin has an activity of 1:3000 and digests 3000–3500 times its weight of egg albumin in 2.5 h at 126°F (52°C). Activities up to 1:30,000 are available commercially.

Pepsin is a part of the crude preparation known as rennet, used to curdle milk in preparation for cheese manufacture. Pepsin is also used to modify soy protein and gelatin (providing whipping qualities), to modify vegetable proteins for use in nondairy snack items, to make precooked cereals into instant hot cereals, to prepare animal and vegetable protein hydrolysates for use in flavoring foods and beverages, and to prevent the formation of cloudiness in beer during refrigeration. *See* CHEESE; FOOD ENGINEERING; MILK.

Myron Solberg

Bibliography. H. Holzer and J. Tschesche (eds.), *Biological Functions of Proteinases*, 1980; T. Nagodawithana and G. Reed (eds.), *Enzymes in Food Processing*, 3d ed., 1997; L. Stryer, *Biochemistry*, 4th ed., 1995.

Peptide

A compound that is made up of two or more amino acids joined by covalent bonds which are formed by the elimination of a molecule of H_2O from the amino group of one amino acid and the carboxyl group of the next amino acid. Peptides larger than about 50 amino acid residues are usually classified as proteins.

In the reaction shown, the α linkage between the α-amino group of alanine and the α-carboxyl group

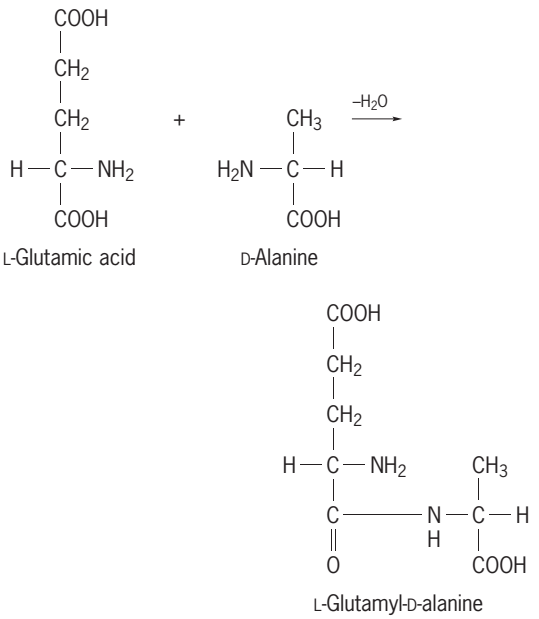

of glutamic acid is by far the most common type of linkage found in peptides. However, there are a few cases of γ linkage in which the side-chain carboxyl group of one amino acid is linked to an α-amino group of the next amino acid, as in glutathione and in part of the structure of the bacterial cell wall. Linear peptides are named by starting with the amino acid residue that bears the free α-amino group and proceeding toward the other end of the molecule. Sometimes the configuration of the asymmetric carbon atom of each residue is included in the name, as shown in the reaction. A three-letter notation is convenient to use, for example, L-Glu-D-Ala.

Occurrence. Peptides of varying composition and length are abundant in nature. Glutathione, whose structure is γ-L-glutamyl-L-cysteinyl-glycine, is the most abundant peptide in mammalian tissue. Hormones such as oxytocin (8), vasopressin (8), glucagon (29), and adrenocorticotropic hormone (39) are peptides whose structures have been deduced; in parentheses are the numbers of amino acid residues for each peptide. Some of the hormone regulatory factors which are secreted by the hypothalamus are peptides that govern the release of hormones by other endocrine glands. In the peptides indigenous to mammalian tissues all of the amino acid residues are of the L-configuration.

Many of the antimicrobial agents produced by microorganisms are peptides that can contain both D- and L-amino acid residues. Penicillin contains a cyclic peptide as part of its structure. Other peptide antibiotics include the gramicidins, the tyrocidines, the polymyxins, the subtilins, and the bacitracins. *See* ANTIBIOTIC.

Synthesis. For each step in the biological synthesis of a peptide or protein there is a specific enzyme or enzyme complex that catalyzes each reaction in an ordered fashion along the biosynthetic route. However, although the biological synthesis of proteins is directed by messenger RNA on cellular structures called ribosomes, the biological synthesis of peptides does not require either messenger RNA or ribosomes. *See* RIBONUCLEIC ACID (RNA); RIBOSOMES.

In the methods most commonly used in the laboratory for the chemical synthesis of peptides, the α-carboxyl group of the amino acid that is to be added to the free α-amino group of another amino acid or peptide is usually activated as an anhydride, an azide, an acyl halide, or an ester, or with a carbodiimide. To prevent addition of the activated amino acids to one another, it is essential that the α-amino group of the carboxyl-activated amino acid be blocked by some chemical group (benzyl oxycarbonyl-, *t*-butyloxycarbonyl-, trifluoroacetyl-) that is stable to the conditions of the coupling reaction; such blocking groups can be removed easily under other conditions to regenerate the free α-amino group of the newly added residues for the next coupling step. In order to ensure that all the new peptide bonds possess the α linkage, the reactive side chains of amino acid residues are usually blocked during the entire synthesis by fairly stable chemical groups that can be removed after the synthesis. For a successful synthesis of a peptide or protein, all of the coupling steps should be complete, and none of the treatments during the progress of the synthesis should lead either to racemization or to alteration of any of the side chains of the amino acids. In the earlier techniques for the chemical synthesis of peptides the reactions were carried out in the appropriate solvents, and the products at each step were purified if necessary by crystallization ("solution method"). An innovation devised by R. B. Merrifield uses a "solid phase" of polystyrene beads for the synthesis. The peptide, which is attached to the resin beads, grows by the sequential addition of each amino acid. The product can be washed by simple filtration at each step of the synthesis. *See* AMINO ACIDS; PROTEIN.

James M. Manning

Peracarida

A large crustacean superorder generally classified within the phylum Arthropoda, subphylum Crustacea, class Malacostraca, subclass Eumalacostraca. The Peracarida are defined as having the following characteristics: a telson (posteriormost body segment) without appendages; one pair (rarely two or three pairs) of maxillipeds (thoracic appendages modified as mouthparts); mandibles (primary jaws)

with articulated accessory processes in adults, between the molar and incisor teeth, called the lacinia mobilis; a carapace, when present, that is not fused with posterior thoracic segments and that is usually reduced in size; basalmost leg segments of the thorax with unique, thin, flattened ventral plates (oöstegites) that enclose a ventral brood pouch (marsupium) in all orders except the Thermosbaenacea (which use the carapace to brood embryos); and having the young hatch as a prejuvenile stage called a manca, lacking the last pair of legs (that is, no free-living larvae occur in Peracarida). The roughly 21,500 species of peracarids are divided among nine orders. The orders Mysida and Lophogastrida were formerly combined (as the Mysidacea), but most authorities now treat them separately. *See* CRUSTACEA; MALACOSTRACA.

Natural history. The peracarids are an extremely successful group of crustaceans and are known from many habitats. Although most are marine, many also occur on land and in freshwater, and several species live in hot springs at temperatures of 30–50°C (86–122°F). Aquatic forms include planktonic as well as bottom-dwelling species at all depths. The group includes the most successful terrestrial crustaceans—the pillbugs and sowbugs of the order Isopoda (suborder Oniscidea)—and a few amphipods that have invaded land and live in damp leaf litter and gardens. Peracarids range in size from tiny interstitial (living among sand grains) forms only a few millimeters long to planktonic amphipods over 12 cm (4.7 in.) long (*Cystisoma*) and deep-water bottom-dwelling isopods that grow to 50 cm (20 in.) in length (*Bathynomus giganteus*). Peracarids exhibit all sorts of feeding strategies, including carnivory, herbivory, and detritivory, and a number of them, especially isopods and amphipods, are symbionts (parasites or commensals). *See* DEEP-SEA FAUNA.

Classification. The Peracarida include nine orders: Mysida, Lophogastrida, Cumacea, Tanaidacea, Mictacea, Spelaeogriphacea, Thermosbaenacea, Isopoda, and Amphipoda (see **illustration**).

The order Mysida comprises nearly 1000 known species of marine shrimplike forms ranging in length from about 2 to 80 mm (0.08 to 3.2 in.). *See* MYSIDA.

The order Lophogastrida (~40 pelagic marine species) is similar to the Mysida; most are 1–8 cm (0.4–3.2 in.) long, although the giant *Gnathophausia ingens* reaches 35 cm (13.8 in.).

The order Cumacea comprises about 1000 species of small, odd-looking, bottom-dwelling marine and freshwater crustaceans with a large, bulbous anterior and a long slender posterior (0.5–2 cm or 0.2–0.8 in. in length). *See* CUMACEA.

The order Tanaidacea comprises about 1500 benthic marine and brackish-water species (0.5–2 cm or 0.2–0.8 in. in length). Tanaids often live in burrows or tubes and are known from all ocean depths. *See* TANAIDACEA.

The four known species that compose the order Mictacea are small (2–3.5 mm or 0.08–0.14 in. in length) creatures living in refugial habitats (that is, areas that have escaped the great changes that occurred in the regions as a whole, providing conditions in which relic colonies can survive). *See* MICTACEA.

The order Spelaeogriphacea is known from only three species. These rare, small (<1 cm or 0.4 in.) peracarids were long known only from a single species living in a freshwater stream in Bat Cave on Table Mountain in Cape Town, South Africa. A second species was recently reported from a freshwater cave in Brazil, and a third from an aquifer in Australia. Little is known about the biology of these animals. *See* SPELAEOGRIPHACEA.

The order Thermosbaenacea contains 11 known species. *Thermosbaena mirabilis* is known from freshwater hot springs in North Africa, where it lives at temperatures in excess of 40°C (104°F). Other thermosbaenacean species occur in much cooler freshwaters, typically in ground water or in caves, and still others are marine or inhabit underground marine pools. *See* THERMOSBAENACEA.

The largest order of Peracarida is the Isopoda, characterized by a complete lack of a carapace, a heart located primarily in the abdomen, and biphasic molting (posterior region molts before anterior region). The isopods comprise about 10,000 described marine, freshwater, and terrestrial species ranging in length from 0.5 to 500 mm (0.02 to 20 in.), the largest being species of the benthic genus *Bathynomus* (family Cirolanidae). Isopods are common inhabitants of nearly all environments, and some groups are exclusively (suborder Epicaridea) or partly (suborder Flabellifera) parasitic. The suborder Oniscidea includes about 5000 species that have invaded land (pillbugs and sowbugs); they are the most successful terrestrial crustaceans. Their larvaless development, flattened shape, good osmoregulatory capabilities, thickened exoskeleton, and aerial gas exchange organs (pseudotracheae) allow most oniscideans to live divorced from aquatic environments. Most isopods are herbivorous or omnivorous scavengers, but direct plant feeders, detritivores, and predators are also common. Some are parasites (for example, on fishes or on other crustaceans). One isopod fish parasite, *Cymothoa exigua* (family Cymothoidae), consumes the tongue of its host fish (snappers) but remains attached to the tongue stub such that the fish continues to thrive using the isopod as a "replacement tongue"—this is the only known case of a parasite functionally replacing a host's organ that it has destroyed. *See* ISOPODA.

Like the isopods, the order Amphipoda is characterized by the lack of a carapace, but the heart is thoracic and molting is not biphasic. The roughly 8000 species of amphipods range in length from 1-mm (0.04-in.) forms to deep-sea bottom-dwelling species reaching 25 cm (10 in.) and some planktonic forms exceeding 10 cm (4 in.). They have invaded most marine and many freshwater habitats and constitute a large portion of the biomass in many marine regions. The principal amphipod suborder is Gammaridea. A few gammarideans are semiterrestrial in moist forest leaf litter or on sandy beaches (for

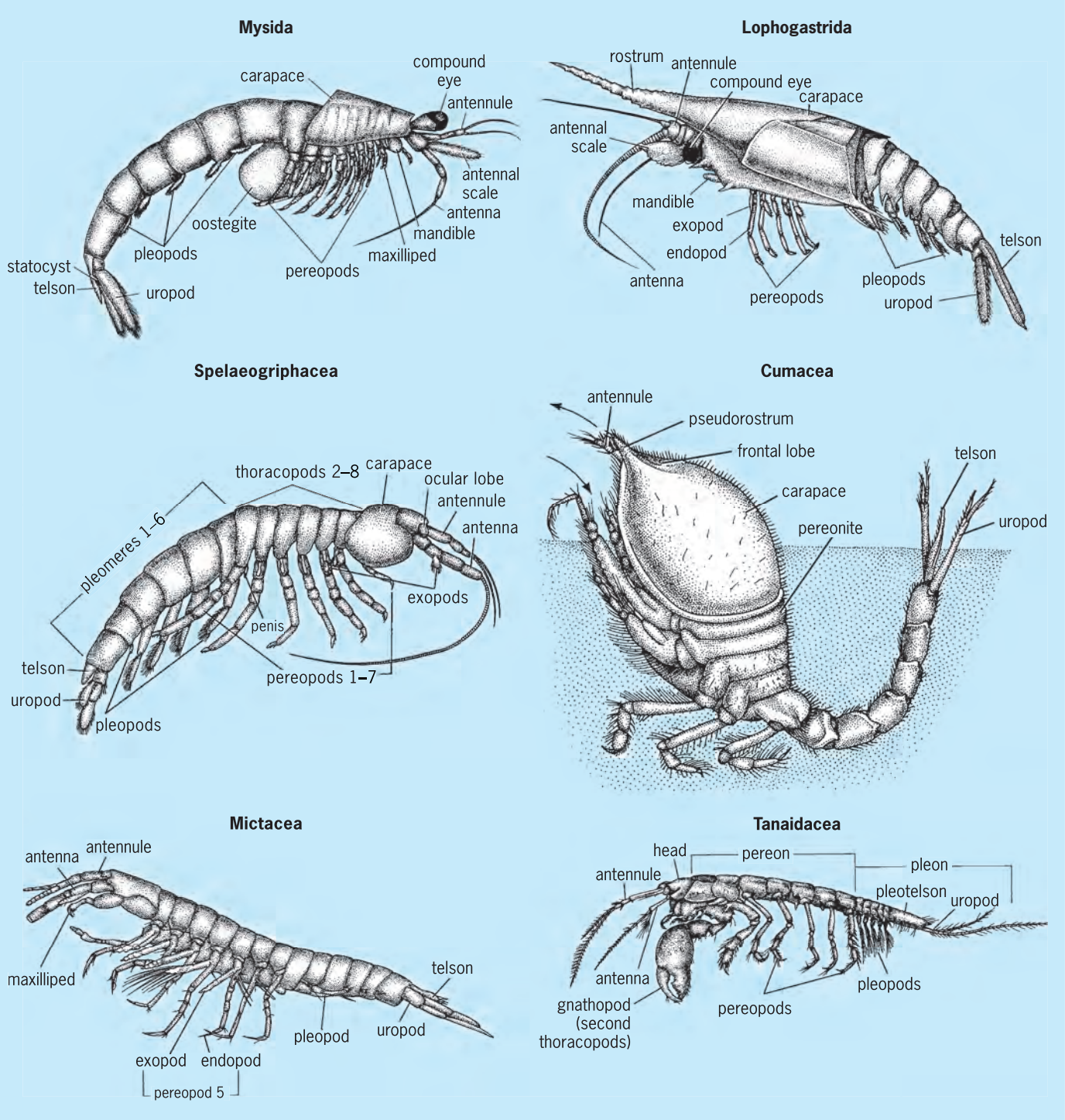

Representatives of the nine orders of Peracarida. (*From R. Brusca and G. Brusca, Invertebrates, 2d ed., Sinauer Associates, 2003*)

example, beach hoppers); a few others live in moist gardens and greenhouses (for example, *Talitrus sylvaticus* and *T. pacificus*). They are common in subterranean ground-water ecosystems of caves, the majority being obligatory ground-water species characterized by reduction or loss of eyes, pigmentation, and occasionally appendages. Most gammarideans, however, are marine species. The three other amphipod suborders are the open-ocean Hyperiidae (most, if not all, associate with gelatinous zooplankton such as jellyfishes and salps), the Ingolfiellidea (about 30 species of primarily subterranean fresh- and brackish-water species, and some marine species), and the Caprellidea. Caprellids include about 300 species of marine "skeleton shrimp," highly modified for clinging to other organisms, including

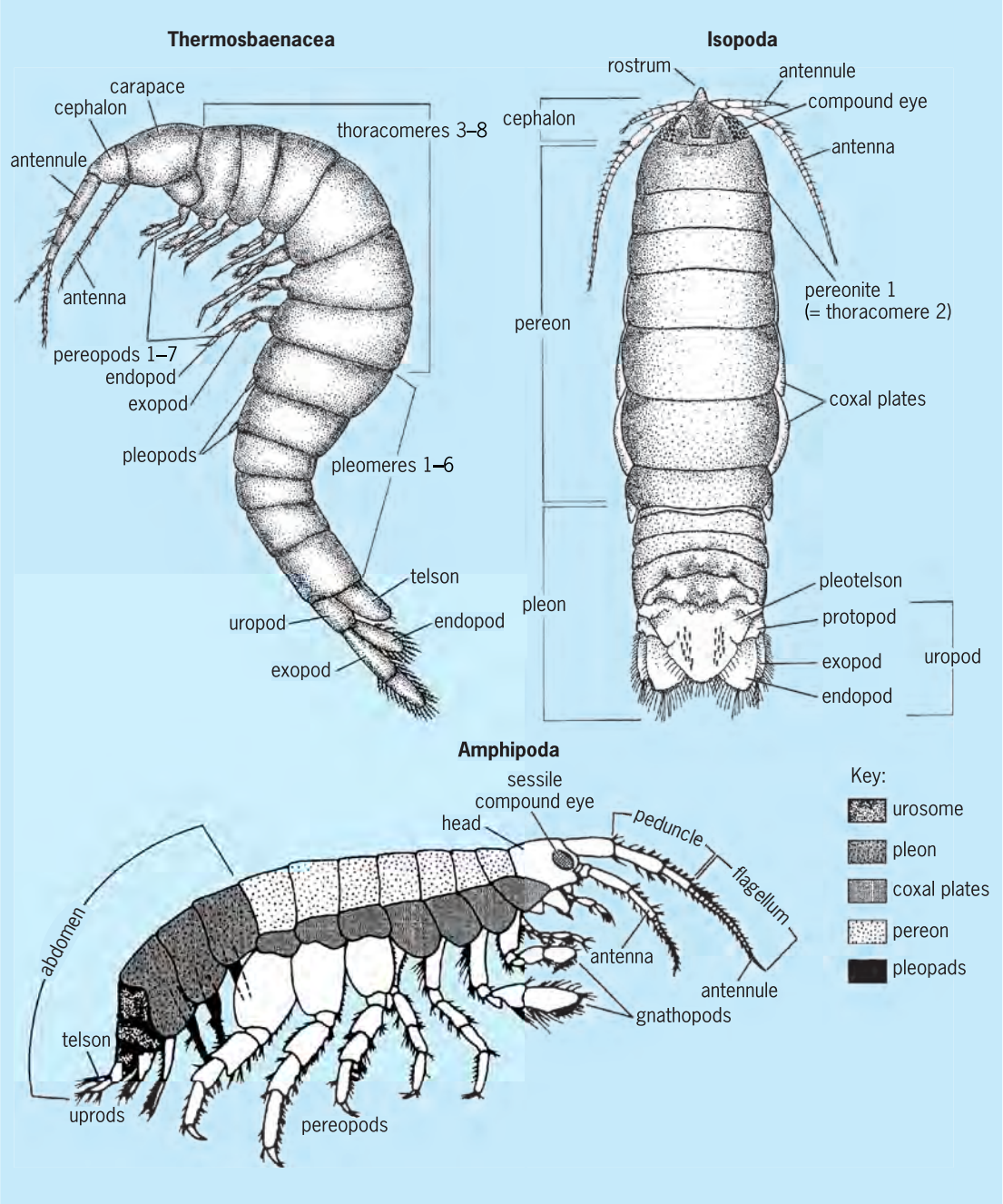

filamentous algae and hydroids. In one family of caprellids, the Cyamidae (28 species), individuals are obligate symbionts on cetaceans (whales, dolphins, and porpoises) and have flattened bodies and prehensile (adapted for seizing or grasping) legs. *See* AMPHIPODA; CAPRELLIDA. Richard C. Brusca

Bibliography. J. L. Barnard, The families and genera of marine gammaridean Amphipoda (except marine gammaroids), *Rec. Austral. Mus.*, Suppl. 13, 1991; T. E. Bowman et al., Mictacea, a new order of Crustacea Peracarida, *J. Crust. Biol.*, 5:74–78, 1985; R. C. Brusca and G. J. Brusca, *Invertebrates*, 2d ed., Sinauer Associates, Sunderland, MA, 2003; R. C. Brusca and M. Gilligan, Tongue replacement in a marine fish (*Lutjanus guttatus*) by a parasitic isopod (Crustacea: Isopoda), *Copeia*, 3:813–816, 1983; J. R. Grindley and R. R. Hessler, The respiratory mechanism of *Spelaeogriphus* and its phylogenetic significance, *Crustaceana*, 20:141–144, 1970; B. Kensley and R. C. Brusca (eds.), *Isopod Systematics and Evolution*, Balkema, Rotterdam, 2001; N. S. Jones, *British Cumaceans*, Academic Press, New York, 1976; P. Laval, Hyperiid amphipods as crustacean parasitoids associated with gelatinous zooplankton, *Oceanog. Mar. Biol. Annu. Rev.*, 18:11–56, 1980; S. L. Sutton and D. M. Holdich (eds.), *The*

Biology of Terrestrial Isopods, Clarendon Press, Oxford, 1984; H. P. Wagner, A monographic review of the Thermosbaenacea, *Zool. Verh.*, 291:1–338, 1994.

Percent

A ratio comparison of two quantities expressed by using 100 equal parts, or hundredths; symbolized %. There are three major uses of percent: part of a whole, rate, and comparison of any two quantities.

Part of a whole. The basic idea of percent is as a ratio that shows a part of a whole. The technical name for the whole is base.

Ratios as percents. If 89 out of 100 problems are correct, the part/whole comparison shows 89/100 or 89% correct. If the whole is not already divided into equal parts, an equivalent ratio to 100 is found.

Certain ratios are easy to express as hundredths. When 3 baskets are made in basketball out of 10 attempts, an equivalent ratio using 100 is found from which the percent is obvious: 3/10 = 30/100 = 30%.

For 35 hits out of 126 times at bat, the ratio 35/126 is not easily expressed as a ratio using 100. Hundredths will be more obvious by dividing 35 by 126, and then reading the number of hundredths to find the percent: 35 ÷ 126 ≈ 0.278 or 27.8%. The percent is obtained by moving the decimal point two places to the right.

Another method is to set up a part-to-whole proportion with n representing the part out of 100, as in Eqs. (1). The ratio is 27.8%.

$$\frac{\text{part}}{\text{whole}} \quad \frac{35}{126} = \frac{n}{100} \quad \frac{\text{part}}{\text{whole}}$$

$$126n = 3500 \text{ (multiplying each side by } 126 \times 100) \qquad (1)$$

$$n = 27.8$$

The usage for the word percentage has changed over time from its original meaning of the "part." In sports statistics, pct. for percentage really means average, not percent. If a batting average of .278 were a percent, it would be reported as 27.8%.

Discounts and savings from bargain sales are examples of a part of a whole. A discount of \$35 (the part of the whole) on a coat with an original price of \$105 (the whole or base is 105) is a savings of 35/105 or 33.3% off the original price, as in Eq. (2).

$$\frac{35}{105} = \frac{1}{3} = \frac{33\frac{1}{3}}{100} = 33\tfrac{1}{3}\% \text{ or } 33.3\% \qquad (2)$$

Percent decrease shows a relative amount of decrease. The amount of decrease, the part, is compared to the original amount, the whole or base. If the value of a house drops \$9000 from an original price of \$80,000, the percent decrease is 9000/80,000, 0.1125, or 11.25%.

Percent of a number. To find a part when the percent and the whole (base) are known, a percent of a number is found.

There are three ways to find the part saved in buying a chair marked at \$300 (the whole or base) with a 25% discount: using a fraction and division, as in notation (3); using multiplication by a fraction or decimal, as in notation (4); or using a proportion, as in Eqs. (5).

$$\begin{aligned} 25\% \text{ of } \$300 &= \tfrac{1}{4} \text{ of } \$ \\ \tfrac{1}{4} \text{ of } \$300 &= \$300 \div 4 \text{ or } \$75 \end{aligned} \qquad (3)$$

$$25\% \text{ of } \$300 \text{ is } 0.25 \times \$300 \text{ or } \tfrac{1}{4} \times 300 = \$75 \qquad (4)$$

$$\begin{aligned} \frac{25}{100} &= \frac{n}{300} \\ 25 \times 300 &= 100n \\ n &= 75 \end{aligned} \qquad (5)$$

The advantage of the first method is that it can be done by easy computation, often mentally. The second method is easier on a calculator. The third method illustrates the proportional connection for percent. In this example, the sale price is \$300 − \$75 or \$225. The sale price can also be found by taking 75%, that is, 100% − 25%, of \$300.

Income tax is a part (the tax) of the whole or base (income). For example, a tax of \$12,500 on income of \$65,000 is 12,500/65,000 or 19.2%.

Calculator use. On calculators, percents can be expressed as decimals and calculations can be done with the decimals. The amount saved in a 35% sale on a \$467 desk can be keystroked as .35 × 467 =. If a percent key, %, is used, conversions are not made to decimals and the % key is pressed last: 467 × 35%. (On some calculators, = is pressed after the %.) After seeing \$163.45 as the amount saved, pressing the minus sign (−) and = keys shows the balance after the discount, \$467 − \$163.45 = \$303.55.

If \$83 is saved on a printer originally priced at \$425, keystroking 83 ÷ 425% shows 19.5, a 19.5% saving.

Rate. While percent always means a comparison to 100, percent can show a rate of so many per 100, not so many out of 100. A sales tax of 6% means a rate of 6% for each dollar, and this 6% is in addition to the dollar. The tax amount can often be calculated mentally by multiplying the 6% rate by the number of dollars. The tax on a \$30 book is 6 × 30% or 180%, \$1.80, and the total cost is \$31.80. The total is 106% of the original price.

Methods of calculation are the same. On a chair priced at \$275 with a sales tax of 8.5%, the tax is 8.5% of \$275. The tax can be found by solving either 0.085 × 275 = n or Eq. (6). Note in the proportion

$$\begin{aligned} \frac{8.5}{100} &= \frac{n}{275} \\ n &= \$23.38 \end{aligned} \qquad (6)$$

that the bases of 100 and 275 appear as the second number for each ratio. The keystroking on a calculator is 275 × 8.5%, and perhaps = also. After finding the tax, $23.38, it must be added to the price to find the total cost, which is $298.38.

Mental computation is needed, for example, to calculate the amount of tip on a restaurant bill of $62.50, as in notation (7).

10% tip:	15% tip:
$^1/_{10}$ of $62.50 = $6.25	10% is $6.25
Move the decimal point	5% more is $3.13
one place to the left	15% is $9.38

20% tip: (7)
10% is $6.25
20% is $12.50
Or divide by 5,
since 20% = $^1/_5$

Interest paid or interest received is done by using rates. Simple interest at a rate of 6% means $6 per $100 for a full year. If interest is calculated monthly on the unpaid balance and the yearly rate is 18%, the monthly rate is approximately 18% ÷ 12 or about 1.5%. At this rate, a balance of $650 for one month results in interest of 1.5% of $650, or $9.75.

With compound interest, the amount of interest is added each compounding period, and this total amount is subject to compounding for the next period. For example, $1.00 invested at 8% will be worth 108% of $1.00 at the end of one year, or $1.08. The worth at the end of the second year is 108% of $1.08, or $(1.08)^2$. At this same rate, after 10 years the compounded value is $(1.08)^{10} = 2.1589247$, or $2.15, using 1.08 as a factor for 10 times.

Similarly, if the birth rate, above deaths, of a country is 4% per year, evaluating $(1.04)^n$ when $n = 18$ shows that the population will double in about 18 years.

It is through compound interest that an annuity builds up value over time. Income from an annuity depends on the rate on investment as well as the number of years that payment is to be made before the annuity matures.

Comparing any two quantities. Percent is used to compare any two quantities, but special care must be given to the base for the comparison. Comparison of city A with a population of 42,000 people and city B with 67,000 people will depend on the base. A compared to B is 42,000/67,000, 62.7%; A is 62.7% of B. B compared to A is 67,000/42,000, 1.595, or 159.5%; B is 159.5% of A. Total sales are often compared year to year. For example, if sales this year are $240,000 and last year they were $175,000, then this year is 240,000/175,000 or 137.1% of last year, an increase of 37.1%. *See* ARITHMETIC. Joseph N. Payne

Bibliography. G. D. Allinger and J. N. Payne, *Percent Level G*, 1991; E. F. Krause, *Mathematics for Elementary Teachers*, 1991; H. L. Schoen (ed.), *Estimation and Mental Computation*, 1986.

Perception

Those subjective experiences of objects or events that ordinarily result from stimulation of the receptor organs of the body. This stimulation is transformed or encoded into neural activity (by specialized receptor mechanisms) and is relayed to more central regions of the nervous system, where further neural processing occurs. Most likely, it is the final neural processing in the brain that underlies or causes perceptual experience, and so perceptionlike experiences can sometimes occur without external stimulation of the receptor organs, as in dreams.

In contemporary psychology, interest generally focuses on perception or the apprehension of objects or events, rather than simply on sensation or sensory processes (such as vision or hearing). While no sharp line of demarcation between these topics exists, it is fair to say that sensory qualities are generally explicable on the basis of mechanisms within the receptor organ, whereas object and event perception entails higher-level activity of the brain. For example, color vision appears to be based upon the activity of cone cells in the retina of the eye. Accordingly, an adequate explanation of why a particular hue is seen is provided by an analysis of the activation of those cells that are primarily sensitive to a particular range of wavelengths of light in the visible spectrum. However, the perception of an object's size, shape, orientation, or the like cannot be accounted for solely by events within the retina of the eye or by the direct neural transmission of these events into the brain. *See* COLOR VISION; HEARING (HUMAN); SENSATION; VISION.

Since objects or events are not experienced only through vision, the term perception obviously applies to other sense modalities as well. Certainly things and their movement may be experienced through the sense of touch. Such experiences derive from receptors in the skin (tactile perception), but more importantly, from the positioning of the fingers with respect to one another when an object is grasped, the latter information arising from receptors in the muscles and joints (haptic or tactual perception). The position of the parts of the body are also perceived with respect to one another whether they are stationary (proprioception) or in motion (kinesthesis), and the position of the body is experienced with respect to the environment through receptors sensitive to gravity such as those in the vestibular apparatus in the inner ear. Auditory perception yields recognition of the location of sound sources and of structures such as melodies and speech. Other sense modalities such as taste (gustation), smell (olfaction), pain, and temperature provide sensory qualities but not perceptual structures as do vision, audition, and touch, and thus are usually dealt with as sensory processes. In this article the emphasis will be almost exclusively on visual perception. *See* OLFACTION; PAIN; PROPRIOCEPTION.

In many respects, the eye is analogous to a camera. Thus, both eye and camera admit light through

a small opening and focus that light by means of a lens upon a light-sensitive surface such that any given point in the environment is represented by a single point on the retina or film. The sum of all these points can be thought of as an image of the external scene. Moreover, the principles of optics apply equally to both eye and camera. For example, distant objects yield smaller images than nearer objects of the same size, and all objects produce images that are inverted and left-right reversed. But perception cannot be explained simply on the basis of the "picture" on the retina, as if nothing more were required but the neural transmission of this picture back into the visual area of the brain. First of all, the eye is in a constant state of tremor, which has important consequences for vision, and eye movements are essential to bring various regions in the scene into the center of the retina so that they can be seen clearly. But more important for the present discussion, if perception were nothing more than the registration in the brain of the retinal image, the perceived size of objects would vary as a function of their distance, the perceived shape of objects would vary as a function of their slant, the perceived orientation of objects would change as the head was tilted, the perceived direction of objects would vary as a function of the position of the eyes within the head, objects would be seen to move when the observer moved, and so forth.

Constancy. By and large, these perceptual properties of objects remain remarkably constant despite variations in distance, slant, and retinal locus caused by movements of the observer. This fact, referred to as perceptual constancy, is perhaps the hallmark of perception and more than any other, serves to characterize the field of perception.

Size. Consider first the problem of size perception. Except at very great distances, objects appear to be about the same size whether seen nearby or farther away. For example, an object at one end of a room will not appear to be any smaller than one of the same size next to the observer despite the fact that the size of the retinal image (or visual angle) subtended by the more distant object will be much smaller than that subtended by the near object (**Fig. 1**). Size perception is constant and thereby correct (or veridical). This would not be true if perception depended exclusively on the size of the object in the retinal "picture." Also, size constancy is perceptual and not simply an act of judgment or knowledge. In other words, objects appear to be the same size at varying distances. It is not the case that they appear to be small at a distance, but that it is realized or reasoned that they are not. This is supported by the fact that animals and young children behave as if their perception of object size, despite variation of object distance, is constant. It is unlikely that they would do so if constancy were a matter of reasoning rather than of perception.

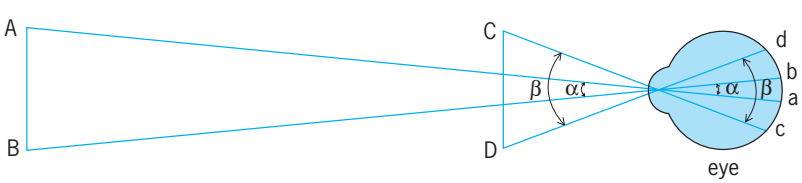

Fig. 1. Objects of equal size A-B and C-D at different distances from the observer produce retinal images of unequal size a-b and c-d. The size of these retinal images can also be described in terms of the visual angles subtended by the objects at the eye for a-b and c-d.

Shape. A similar problem arises in the perception of shape. Object shape also remains perceptually constant. A circle seen from the side projects an elliptical rather than a circular image on the retina. Therefore, if perceived shape were simply a function of retinal-image shape, the circle would appear to be elliptical from most vantage points. The fact that it is seen as circular and not elliptical is referred to as shape constancy.

Orientation and position. The orientation of an object's image on the retina changes when the head is tilted, and the position and stability of an image changes when the observer moves the entire body, head, or eyes. Nonetheless, objects maintain their perceived orientation, location, and stability with such tilts or motions. These properties are referred to as orientation constancy and position constancy, respectively.

Lightness. Whether a surface reflects most or little of the light it receives is a function of its physical characteristics. There is little variation in the perceived lightness of such surfaces—whether they appear black, gray, or white—despite enormous difference in the amount of light falling upon them and in the amount of light reaching the eye from them (luminance). This property is referred to as lightness constancy. *See* LUMINANCE.

Experiments. A typical experiment on constancy, for example, size constancy, may involve a subject viewing an object such as a triangle at a particular distance. The subject is instructed to compare the size of that triangle (called the standard) with a series of other triangles of graduated size shown one at a time (called the comparison series) at a much closer distance; the subject then selects the triangle that appears to be the same size as the standard. The comparison series can be presented alternately in ascending and descending order for several trials and the average match computed. The same procedure is then followed with the standard at an entirely different distance. By using many subjects and averaging their average judgments for each distance, the degree of constancy achieved can be determined for the population sampled for a range of distances. There are alternatives to this manner of presenting the comparison objects, referred to as the psychophysical methods. Experiments of this kind on constancy usually yield average matches that reflect a high degree of constancy except when very great distances are utilized, when the adequacy of distance perception is much reduced, or when young children are tested. *See* PSYCHOPHYSICAL METHODS.

Basis for constancy. While there is as yet no accepted explanation of perceptual constancy, there is general

agreement that it is based on taking account of certain additional information available to the perceptual system. For instance, in the above example, the far object is seen to be far away; however, the perceptual system could compute by a process analogous to multiplication that the small retinal image yielded by a distant object represents an object equal in size to that represented by a larger retinal image projected by a nearer object. In other words, the computation would entail multiplying image size by detected distance, and since the size of the image is inversely proportional to distance, perceived size should remain constant as long as the object's distance is correctly perceived. Where distance is underestimated or not adequately registered, the object's size should be computed as small. Therefore at very great distances, size perception no longer remains correct or constant—things then do look diminutive. There is evidence that constancy requires focused attention, and that without it, perception regresses to conform to the retinal image alone.

Veridicality and illusion. Constancy implies veridicality, that is, agreement with the real properties of the object. Veridicality is vital for survival. An animal must be able to perceive that an approaching but still distant object is an adult lion and not a cub, and it must be able to distinguish whether a moving retinal image signifies a moving external thing or is caused by its own motion while viewing a stationary thing. But the process that has evolved—or possibly has been learned—and that leads to veridicality and constancy can often lead to nonconstancy or illusion. That is why investigators take seriously and investigate illusions of all kinds. It is presupposed that such illusions, once understood, will lead to an improved understanding of perceptual processing in general.

Motion perception. Perception is not only of objects but of events as well. The kind of event that has been investigated in great detail is the perception of motion. As already noted, perceived movement cannot simply be explained by the motion of an object's retinal image since image motion caused by observer or eye movement does not lead to perceived object movement. Moreover, an object tracked by smooth-pursuit eye movements will appear to move, although in that case there is essentially no motion of the object's image over the retina. Similarly, an afterimage will appear to move during eye movement even in a completely darkened room. Where ordinarily the movement of the retinal image caused by the moving eye is computed to signify "no object motion," thus yielding position constancy (since the image motion and eye motion are equal in magnitude), the same computational rule must signify "object motion" in the case of the afterimage.

Mechanisms. In any event, motion will be perceived when an object or afterimage changes its angular direction with respect to the observer at a rate above some threshold value, and it does not matter whether the information concerning such change of angular direction is based upon eye movement, with retinal image stationary, or retinal-image movement, with eyes stationary. Clearly it is the combined information from retina and eye that matters, and thus perception is not explicable simply in terms of events in the receptor mechanism. There are cells in the visual nervous system that are responsive to motion of contours over the retina. However, the discharging of these so-called motion-detector cells should not be thought of as the explanation or cause of perceived motion, at least in animals high on the evolutionary scale where the eyes move. Rather these mechanisms should be thought of as providing the information that an image is displacing over the retina. What that information then leads to as far as motion perception is concerned depends upon other information, such as what the eyes are doing at that moment.

Induced movement. Other phenomena also attest to the operation of more complex processes underlying motion perception. For example, when clouds pass in front of the Moon, the Moon appears to move in the opposite direction. This kind of illusion is referred to as induced movement. It is studied in a dark room by moving a large object such as a luminous rectangular perimeter back and forth while the observer views a stationary luminous spot inside the rectangle. Typically the spot appears to move in a direction opposite to that of the rectangle. While induced movement is not yet fully understood, it seems evident that it is based upon the change in position of spot and rectangle relative to one another. The rectangle is taken as the frame of reference, and the relative motion between spot and rectangle is thus interpreted by the perceptual system as motion of the spot. If change of position of one object relative to others is information that can produce an illusion, then it is probable that it is important in daily life. Therefore when an object moves in the environment, there are two independent determinants that account for the perception of it as moving: it changes its angular direction with respect to the observer at an above-threshold rate, and it changes its position relative to neighboring objects. Of the two, the latter may be the stronger factor.

Since one's body is itself an object in the field, it would seem to follow that if the body is surrounded by a structure that could serve as a frame of reference for it, that is, for one's own self, motion of that structure would induce movement of the self in the opposite direction. In fact, this effect occurs often in daily life, as when a nearby automobile moves and the observer in another automobile is stationary. This effect is studied in the perception laboratory by surrounding a stationary observer with a moving drum or artificial room. When observers experience themselves as moving, the moving, surrounding room appears to be stationary.

Stroboscopic motion. Still another illusion of motion occurs when a stationary object is first seen briefly in one location, and, following a short interval, is seen in another location. This phenomenon, known

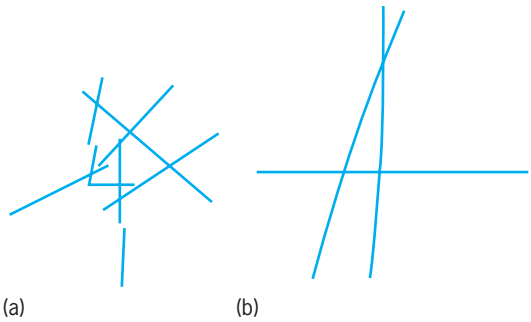

Fig. 2. Perceptual organization. (*a*) The figure of a four is immediately and spontaneously perceived despite the presence of other overlapping and adjacent lines. (*b*) The four, although physically present, is not spontaneously perceived and is even difficult to see when one knows it is there.

as stroboscopic motion (or alternatively, as apparent motion, beta motion, or the phi phenomenon), underlies the perception of motion in movies or television. A moving picture is based on the successive projection of stationary frames in the film strip. Thus there is no motion itself on the screen, but motion is nonetheless perceived. Many erroneously believe that this effect is based on the neural persistence of the image of a frame during the interval between frames, but clearly this only explains the absence of a flickering effect. The motion perceived must be based upon the tendency of the perceptual system to interpret the successive appearance of an object in different locations as signifying motion of the object. The effect occurs only if the temporal and spatial separations are within a particular range of values. It is worth emphasizing that in both induced and stroboscopic motion, there is no motion over the retina of the object which is seen to move.

Causality. Another type of event besides motion that has been investigated concerns the factors that lead to the impression that one object has caused a certain behavior in another object. For example, when one billiard ball hits another, the movement of the second ball is experienced as having been caused by the movement of the first. This kind of effect is studied in the laboratory by a sequence of visual transformations (as in an animated cartoon) in which the behavior of the visual objects A and B can be experimentally manipulated. It has been found that the impression of causality occurs if and only if certain stimulus conditions prevail: the movement of B must occur immediately upon impact, the direction of B must be the same as that of A, and so forth. For these reasons it is now claimed that causality can be directly perceived, whereas in the past it would have been held that it is simply a judgment based on knowledge and that all that is directly perceived is the sequence of motions.

Form perception. One question requiring study is why a triangular object appears triangular when a triangular image of the object is present on the retina. Before attempting to answer this question, it is important first to make certain distinctions. Form perception means the experience of a shaped region in the field. Recognition means the experience that the shape is familiar. Identification means that the function or meaning or category of the shape is known. For those who have never seen the shape before, it will be perceived but not recognized or identified. For those who have, it will be perceived as a certain familiar shape and also identified. Recognition and identification obviously must be based on past experience, which means that through certain unknown processes, memory contributes to the immediate experience that one has of the triangle, giving the qualities of familiarity and meaning.

Perceptual organization. But even if the discussion is restricted to form or shape perception per se, there are still many aspects of such perception that require explanation. For example, there is perceptual organization. The figure of a 4 in **Fig. 2***a* is seen as one unit, separate from other units in the field, even if these units overlap. This means that the parts of the figure are grouped together by the perceptual system into a whole, and these parts are not grouped with the parts of other objects. Moreover, a region will usually be seen as a figure in front of a background (ground) rather than as a hole in a surface. In general, the contour tends to give shape to the region on only one side at any one time, the one organized as a figure. A change in such organization can thus lead to a very different perceived shape (**Fig. 3**). Perceptual organization is an achievement of perceptual systems without which the perception of the world would presumably be an incoherent mosaic of unrelated points and patches of different lightnesses and hues. It is held that this organization is based on certain principles, many of which are known. For example, in Fig. 2*b* the figure of a 4 that is physically present is not spontaneously perceived, as it is in Fig. 2*a*, because each of its sides is organized as part of the longer lines by virtue of the principle known as good continuation. The necessity for such organization is one reason why form perception cannot be regarded as given directly by the presence of a retinal image of a particular shape.

Geometrical relationships. There are other problems about form perception that remain to be unraveled. For example, the size of a figure can vary, as can its locus on the retina or even its color or type of

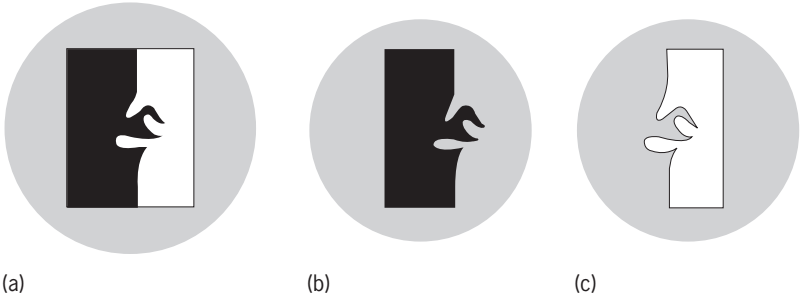

Fig. 3. Figure-ground organization. The pattern in *a* tends to be organized as a black figure on a white ground or a white figure on a black ground. The shapes of these two figures are very different despite the fact that the central contour is the same in the two cases, as is clearly indicated in *b* and *c*.

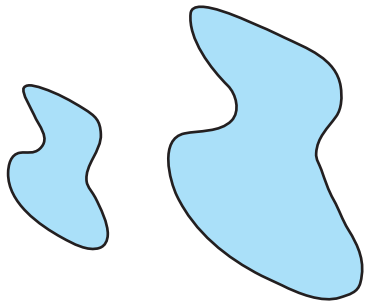

Fig. 4. Transposition of form; the two shapes look the same despite the difference in size.

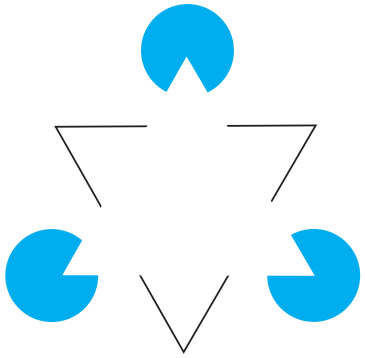

Fig. 5. Illusory contours. Although most of the contours of the upright triangle are not physically present, the triangle is perceived as a solid white figure with contours.

contour, without affecting its perceived shape (**Fig. 4**). This means that the perception of a particular shape does not depend upon the stimulation of particular cells of the retina or on the action of particular cells in the visual area of the brain, the occipital cortex, to which the retinal pattern is projected. What matters for form perception is the patterning or geometrical relationships between the parts of a figure. In other words, just as a melody can be transposed in key or octave, so a form can be transposed in various ways. Common to both cases is the fact that while the parts are no longer the same, the experience of the whole remains unchanged. This implies that the perceptual system is focusing on what does not vary with transposition, namely, the relational information about the melody or form. In the former it is the rhythm and pitch relationships, and in the latter the geometrical relationship of contour components to one another.

Physical contours. Ordinarily, form perception depends upon the detection of the contours which delineate a figure, and the contours are represented on the retina by an abrupt change of light intensity (or luminance) across each edge. Since sensory physiologists have discovered cells in the brain that "detect" the presence of contours or edges in particular orientations on the retina, some might be inclined to believe that form perception simply reduces to the detection of all the contours of a figure. Besides the problems of organization and transposition already mentioned, there are further difficulties with such a reductionistic explanation. Physical contours need not be present on the retina for "subjective" or illusory contours to be perceived (**Fig. 5**). Moreover, a figure can be perceived when it is moved behind an opaque surface containing a narrow slit so that only a thin slice of it is visible at any moment, a phenomenon referred to as anorthoscopic perception. These facts suggest that physical contours are ordinarily the markers of the locations of the boundaries of a figure but are not essential for form perception.

Orientation. A further fact about form perception is that it is dependent upon orientation (**Fig. 6**). It is a commonplace observation that printed or written words are difficult to read when inverted, and faces look very odd or become unrecognizable when upside down. Simple figures also look different when their orientation is changed: a square looks like a diamond when tilted by 45°. It is not easy to explain these effects on the basis of the altered orientation of the figure's contours on the retina since its perceived shape will not be changed simply by tilting the head (and therefore the eyes). In fact, the square tilted in the environment will look like a diamond regardless of head position. What seems to matter here are changes in the regions of the figure that are regarded as "top," "bottom," and "sides." The assignment of these locations to a figure by the perceptual system plays an important role in form perception.

Attention. Then there is the question of the role of attention in perception. Is form perception the automatic result of the neural transmission of an image from retina to brain, or is attention necessary? Suppose, as in a well-known experiment, the observer is shown an array of many letters of the alphabet in a very brief presentation, less than a second. Are they all perceived? There is disagreement among psychologists on this question, but the fact is that unless attention is focused on certain letters during

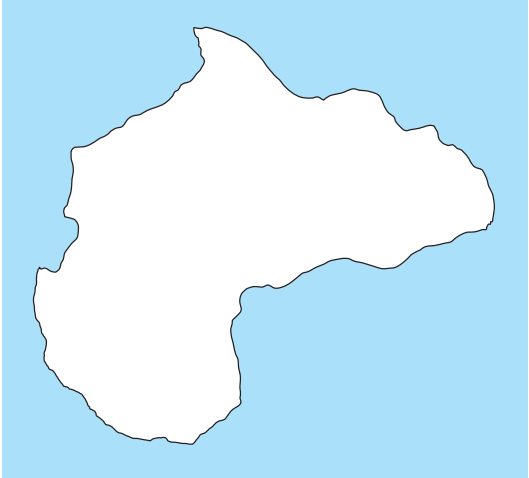

Fig. 6. This figure is not recognizable. If, however, it is tilted 90° clockwise, it will be recognized as an outline map of Africa.

or shortly after their presentation, they will not be remembered, even directly afterward. Therefore these unattended letters are either not perceived or, if they are, they do not lead to the establishment of even fleeting memories. In another famous experiment on hearing, observers are presented with two voices simultaneously reading different messages, and they are required to attend only to one of them by repeating each word in it. Afterward, memory tests indicate that very little if any of the unattended message is recalled. This experiment brings into the laboratory the well-known situation in everyday life where a person tries to listen only to one other person when several people are speaking at once, the so-called cocktail party effect. The impression is of perceiving only that person's conversation which is being attended to, and the rest is experienced as mere background "noise." Similar effects of inattention have been found for visual perception. What is not yet clear from these experiments, however, is whether the unattended figures or messages are not identified or are not even perceived as specific visual forms or organized auditory units.

One theory states that attention is not necessary to process certain elementary features in the visual array, such as colors, edges, or motions. Indeed, the presence of many such elements seems to be detected "in parallel." On the other hand, when combining such features in perception, they must be integrated—for example, perceiving a red triangle—and such feature integration does require attention. Support for this theory derives from experiments suggesting that integrated features can be processed only serially. Further support derives from discoveries that different object features are processed in entirely different regions of the brain.

Geometrical illusions. Related to the topic of form perception is the misperception of the size or direction of parts of figures that constitutes many of the geometric illusions. In an illusion figure, one particular part is perceived to be either longer or shorter than another part, although they are objectively equal (**Fig. 7***a*); or the direction of a contour is perceived to be different from that of another contour although they are the same (Fig. 7*b*). For reasons still not understood despite a century of investigation of these illusions, the background or context of the rest of the figure affects these parts. One theory is that the illusion pattern is interpreted by the perceptual system as three-dimensional—whether the observer is consciously aware of this or not—and thus constancy operations are inappropriately applied. For example, in the Ponzo illusion illustrated in Fig. 7*a*, the converging lines function as the perspective representation of parallel lines receding into the distance. Consequently the horizontal lines are taken to be at different distances from the observer and, if so, would have to appear different in size. However, there are difficulties with this theory, and several other theories have been advanced and investigated.

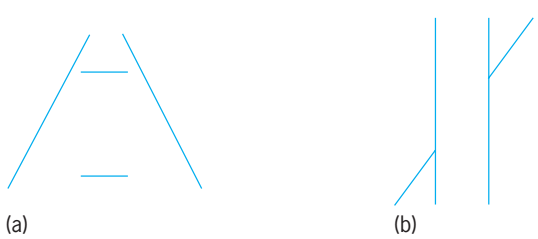

Fig. 7. Geometrical illusions. (*a*) The Ponzo illusion in which the two horizontal lines of equal length appear unequal. (*b*) The Poggendorff illusion in which the two oblique line segments are aligned with one another (that is, are colliner) but appear to be misaligned.

Innate or learned? A central problem is whether the perception of properties such as form and depth or the achievement of veridical perception as in the constancies is innately determined or is based on past experience. By "innate" it is meant that the perception is the result of evolutionary adaptation and thus is present at birth or when the necessary neural maturation has occurred. By "past experience" it is meant that the perception in question is the end result of prior exposure to certain relevant patterns or conditions, a kind of learning process. Despite centuries of discussion of this problem by philosophers, and considerable experimental work on it during the century since psychology became an empirical science, there is still no final answer to the question. However, one method of studying vision in infants which has made it possible to provide some answers makes use of the fact that infants tend to look at presented objects but within a matter of seconds will cease doing so (habituation). The direction of their eyes can be observed and recorded. If another object is then presented, they will look at it provided that it appears sufficiently different from the first object (dishabituation). In this way it is possible to ascertain what objects appear similar to the infant, which in turn presupposes adequate vision of such objects. Using this and other methods it has been determined that infants, even shortly after birth, perceive shape and achieve figure-ground organization and perceptual constancy such as the size and shape. It seems clear that certain kinds of perception are innate, but past experience also is a determining factor.

Depth perception. The perception of the distance of objects is present in various species of animals early enough to rule out past experience as a causal factor. One experiment demonstrating this consists of placing a very young animal in the center of a large horizontal sheet of glass. To one side an opaque cloth is directly under the glass but, to the other side, the floor is several feet below. When placed in the center of this "visual cliff," the young of virtually all species tested will avoid moving to the side of the glass that appears to entail a steep drop. The experiment makes use of an innate fear of height, but of course only if that height can be perceived should the animals be expected to avoid one side of the glass. What is not yet known with certainty, however, is which of the many "cues" to distance is relevant in this experiment. Since a similar effect

occurs when vision is permitted in only one eye, some investigators believe the relevant innate cue here is one governed by the displacement of the retinal image resulting from movement of the observer, known as movement parallax. In fact, a similar, vivid depth effect can be achieved in adult observers by motion of an array of points or contours with the observer stationary, currently referred to as depth (or structure) through motion, the stereokinetic effect, or the kinetic depth effect. By using the method of habituation with infants, it has been found that very young infants perceive distance and depth within objects, particularly when motion of the infant or object is allowed; binocular depth perception based on steroposis seems to take several months to develop.

Of course, one of the important cues to depth derives from the slightly different images the two eyes receive by virtue of their slightly different position in the head, an effect known as retinal disparity. Such depth perception will occur even if nothing recognizable (or even no visible contours) is presented to each eye, but instead a random array of small elements or black and white dots is presented to each eye (the so-called random-dot stereogram). These stereograms are constructed by selecting some region in one picture and displacing a corresponding region in the other picture with respect to the background. In viewing monocularly, one perceives only a random array of dots. In viewing binocularly, however, the observer perceives the region as floating in front of or behind the remainder of the pattern.

By way of contrast, cues to depth can be based on the content of the scene itself, such as perspective, shadow, or the interposition of one object in front of another. Since these cues are effective in creating in impression of depth even in pictures, they are called pictorial cues. There is some experimental evidence that past experience plays a role in the utilization of this kind of stimulus information in the perception of depth, both in real-life scenes and in pictures. Moreover, it seems plausible that this kind of experience could and would be acquired. Patterns of this kind are encountered virtually all the time in daily experience under conditions where depth perception is given by other information that has an innate basis. Thus the pictorial cues could be acquired by a process of associative learning.

Picture perception. The topic of picture perception has been receiving increasing attention. It is a complex problem because observers simultaneously perceive what the picture represents, for example, depth, and that it is something that is two-dimensional on a surface. Of particular interest is the question of whether or not past experience with pictures is necessary to perceive them appropriately or whether exposure to certain kinds of environments, such as the carpentered world of technologically advanced societies, plays a crucial role in picture perception. These questions have led to the study of picture perception in animals, in young children, and in "primitive" societies. If object perception is examined separately from depth perception in such pictures, it seems clear that the objects themselves are universally perceived and recognized in all societies and by very young children and some animals. This is particularly true if one removes the sometimes obfuscating effect of the background of a picture.

Basis of constancy. As to the question of whether or not perceptual constancy is learned or innate, the results of the habituation method with infants indicate the very early presence of constancy. In the case of lightness constancy, there is wide agreement that it is based on stimulus relationships, specifically on the ratios of luminance of neighboring regions. Indeed, there is evidence that the crucial information about such luminance ratios is given only at the edges between differing regions. If ratios govern perceived lightness, it would explain why constancy prevails, that is, why changes in overall illumination do not affect perceived lightness. It is unlikely that such a mechanism for constancy would have to be learned. However, a further problem here is that the perceptual system must be able to distinguish edges caused by lightness differences from those caused by illumination differences, such as shadows or three-dimensional corners. The perceptual system seems to be able to do so, but the mechanism is unknown.

Optical distortions. There has, however, been a good deal of research concerned with the capacity of animals and humans to adapt perceptually to optical distortions created by viewing the world through prisms and lenses of various kinds. For example, if an observer views the world through lenses that invert the retinal image—that is, turn it around 180°F from its customary orientation and cause the world to appear upside down—will the observer, after wearing the lenses for days or weeks, ultimately perceive the world as upright again? Within certain limits, there is evidence that adaptation does occur to prismatically created optical displacements, tilts, changes of curvature, and the like, although the precise nature, basis, and mechanism of such perceptual learning are not yet fully understood. However, the evidence is not convincing that adaptation to a reinverted retinal image occurs.

Some of this research bears on the question of the learned basis of constancy. For example, experiments have been conducted in which certain prisms or lenses worn by the subject create a new rate or direction of motion of the retinal image of a stationary scene resulting from the observer's movement. At the outset, position constancy is abolished, and the scene appears to move whenever the observer moves. Following continual viewing through these devices combined with movement of the observer, motion of the scene begins to lessen and eventually it ceases. Therefore position constancy is known to be subject to relearning even if, as seems to be the case, it did not have to be learned in the first place by the infant.

Experience and form perception. It has never been entirely clear why past experience has been considered essential for the emergence of various

kinds of perception, particularly in the case of form. One might ask what is perceived prior to the occurrence of the necessary past experience. Nor has it been clear how experience would play a determining role such that, following it, the appearance of objects and events would be different than before it. One difficulty is the fact that if past experience is "carried" in the brain in the form of specific memories, these memories (rather than other irrelevant ones) must be contacted or evoked by the stimulus now impinging on the sense organ if they are to play a role in shaping the percept. Ordinarily, some psychologists believe, memories are evoked on the basis of similarity to what is now perceived, yielding recognition and identification. For past experience to determine form perception per se, however, the memories would have to be evoked before the form is perceived, and that seems unlikely, or at least it poses difficulties. This analysis suggests that the role of experience in form perception may be limited. It cannot entirely determine the form perceived, but might be able to enrich it in certain respects once it is perceived, and a bridge to the relevant memories is thus established. The impression of depth in the case of a picture of a three-dimensional object might well be an example of this kind. *See* MEMORY.

Characteristics. One very important principle is that, with few exceptions, perception is not determined by knowledge about the nature of the objects or events in a given situation. Thus, illusions are not abolished simply by knowing they are illusions. Perception follows its own laws, but this does not necessarily deny that past experience can automatically and unconsicously influence perception in certain cases.

Ambiguity. Another importance fact is that the stimulus impinging on a sense organ, such as the retinal image of an object, is ambiguous as concerns certain properties of the object it represents. An image subtending a given visual angle logically can represent an object in the world of any size, depending upon its distance. Conversely, different visual angles can represent an object of the same size. Therefore, the same stimulus can represent varying objects, and different stimulus can represent the same object. Ambiguity also refers to the fact that the same stimulus can, over time, give rise to alternating perceptions. Such stimulus patterns are known as reversible figures (**Fig. 8**). Reversal has been much studied in the perception laboratory. However, research has shown that if the observer is not familiar with the figure and is not told it is reversible, then such reversal often will not occur. In the case of other patterns that can represent two (or more) objects or events, they often tend to be perceived in only one of these ways, suggesting that the perceptual system has some basis for a preference for one over the other. For example, when no other depth information is available, an expanding or contracting image of an object will generally be seen as an object approaching or receding rather than as one changing its size.

The Gestalt psychologists suggested that there is a preference for the simplest of all possible perceptions, which they referred to as the tendency toward *Prägnanz*, and they based it on their corresponding notion that brain processes tend toward a minimum of energy distribution. Thus Fig. 8*a* was held to be seen as three-dimensional rather than two-dimensional because as such it is regular and symmetrical. There is now renewed interest in this concept, since it is possible to formulate it in terms of information theory. *See* INFORMATION THEORY (BIOLOGY).

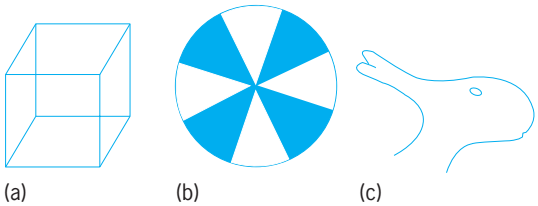

Fig. 8. Reversible figures. (*a*) The Necker cube, an example of reversible perspective. (*b*) A variation of the Maltese cross in which either the colored or white cross can be seen as the figure. (*c*) The duck or rabbit figure.

Relationships. Still another general characteristic of perception is that what is perceived is often determined by relationships or context as mentioned above. Thus, percepts are often relationally determined. There are various examples of this kind, such as the fact that a vertical rod seen in a dark room within a tilted rectangle will appear tilted from the upright in the opposite direction, the so-called rod-and-frame effect. Clearly it is the angular relationship of rod to rectangular frame that matters here. One explanation of the effect is that the rectangle is taken as the frame of reference, as if it were a surrogate of the vertical and horizontal axes of space. That being so, a rod not aligned with either axis of the rectangle must appear tilted. Induced motion and lightness perception are other examples.

Visual capture. One other general fact about perception concerns the dominance of vision over other sense modalities. That is, when a conflict is created such that vision indicates one value and touch or hearing another value, not only is vision dominant in determining what is perceived, but what is felt or heard conforms to what is seen. This latter fact is referred to as visual capture. Some examples have already been mentioned. Induced movement of the self implies that, despite sensory information that one is stationary, visual motion of the surroundings leads to the experience that one is in motion in the opposite direction. Another example is a straight rod that appears curved when viewed through a wedge-shaped prism. If one then runs one's hand along the rod while viewing it (which of course, with eyes closed, would lead to the impression of a straight rod), it will feel as if it is curved. Thus vision captures touch. Sound localization is also dominated by vision. While a speaker inside one's automobile in a drive-in theater will produce sounds of voices localized inside the car before the movie begins, after it does commence, the voices are perceived as coming from the actors seen on the

distant screen. Here vision captures hearing.

Theories. Three different theories of perception have dominated this field over the last century.

Helmholtz theory. The first, associated with H. von Helmholtz, is that perception is the result of a process of unconscious inference concerning what the stimulus reaching the sense organ most probably represents in the world. Helmholtz believed that the inference process is based on past experience and is unconscious because the individual obviously is not aware of making it. For example, a white surface will appear white whether seen in sunlight or in dim indoor illumination because allowance is made for the differing illumination on the basis of experience. The achievement of veridical perception despite the ambiguity of the stimulus is thus accounted for in terms of a process of inference.

Gestalt theory. The second theory, represented by Gestalt psychology and founded by M. Wertheimer, W. K. Köhler, and K. Koffka, arose in part as a reaction against the empiricist view as it developed in the hands of psychologists such as Wundt and Titchener. These latter theorists had argued that individuals do in fact perceive on the basis of the stimulus impinging on the sense organ (which the empiricists called sensations) but that they interpret them in accord with what they have learned. To avoid confusion interpretations with true sensations, Titchener advocated training observers to become expert at attending to their sensory experiences (analytical introspection). The Gestaltists mounted a powerful critique of this approach, noted many logical and empirical difficulties with it, and emphasized the danger of distorting the description of the way the world appears in order to fit a preconceived theory. They also objected to the assumption for which there was no solid evidence, that perception is a function of prior experience. Instead they postulated innate organizing or interactive processes in the brain as the basis of perception. For example, stroboscopic motion perception was held to result from a neural interaction of the regions in the brain where the stimulus from each flashing object was transmitted. In general, the achievement of veridical perception despite the ambiguity of the stimulus was held to be explicable on the basis of relational determination and perceptual organization. Stimuli were said to interact and that is why relationships matter.

Psychophysical theory. The third theory represented and attempt to get around the difficulties concerning the ambiguity of the stimulus and those that arise because perception does not seem to correspond in any direct way with the stimulus. It was those same difficulties, for example, that size perception is not simply a direct function of visual angle, that led to the first two theories. According to this third theory, represented chiefly by James J. Gibson and his associates, there are features of the stimulus that do correspond to, or are correleated with, the perception, but they depend upon a more abstract or higher-order analysis of the stimulus. This approach has been characterized as a psychophysical or direct theory since it maintains that perception can be completely explained on the basis of stimulation, and therefore it is not necessary to posit such nonobservable hypothethical events as unconscious inference, interpretation of sensation, or neural interaction. As example of a higher-order feature of the stimulus is the gradient of density of texture that is present when an image of a uniform or random distribution of elements on a plane is projectd onto the retina. For example, the blades of grass and space between them from the nearby region of a field will yield relatively large retinal images and separations among them, but the farther away the region of grass, the smaller these images and separations will be. Thus a gradient of texture density could well be the stimulus for the apparent depth of a plane such as the ground. With this redefinition, proponents of this theory argue that the stimulus information is in fact not ambiguous.

Difficulties. There are contemporary psychologists who espouse one or another of these three theories, with or without certain modifications. However, each kind of theory encounters certain difficulties in accounting for all the known facts. Therefore, most investigators in perception today extract those aspects of these theories that seem to have been verified or that seem most plausible. Thus there is general agreement with the Gestalt view that a process of perceptual organization must be assumed and that relationships among stimulus components are indeed the basis of many perceptions. There is substantial agreement with the helmholtzian view that constancy often entails a process very much like unconscious inference, of computing object properties on an basis of available information, but it is not yet clear that such a process is necessarily based on past experience. There are many other examples in perception which suggest that what is perceived is the result of an intelligent, thoughtlike mental operation or inference. If so, such an operation must be very rapid (since perception occurs within a fraction of a second), not verbal (since animals and young children seem to perceive much as adults do), and unconscious. Finally, many researchers agree with the direct theory that there are certain higher-order attributes of the stimulus that can be abstracted (which were previously undetected or overlooked) and that can account for certain perceptions.

Trends. Independent of these theories are certain trends in the field. One approach that has flourished is to view perception as a stage in the temporal sequence of processing of information (called information processing). Perception follows earlier stages such as the neural encoding of sensory stimulation, but it precedes later stages such as the establishment of memories. Processing at each state requires a certain amount of time and, in principle, can be isolated by certain experimental techniques. *See* INFORMATION PROCESSING (PSYCHOLOGY).

Another development arises out of the interest in simulating the perceptual and other intelligent achievements of animals and humans by computers. Thus investigators in the field of artificial intelligence have begun to develop computer

programs that would make it possible to arrive at a correct representation of the outer scene, given the same kind of information the eye receives in the form of an image (referred to as machine vision). Connectionism, or parallel distributed processing, challenges the assumptions underlying the "symbolic paradigm" of the artificial intelligence approach to perception, namely, that mental processes can be modeled as programs running on a digital computer. Instead, connectionism proposes that mental processes can be thought of as large-scale dynamic networks of simple neuronlike processing units. Perception (as well as other cognitive processes) can then be understood as the flux of global patterns of activation over the entire network. Since the pattern of activation tends toward a state of minimum energy, the theory is in some respects similar to Gestalt theory except that it is compatible with contemporary knowledge of neutral networks. *See* COGNITION; NEURAL NETWORK.

There is also renewed interest in the relationship between perception and imagery. Research has suggested that imagery has characteristics similar to those of perception and that perception itself can sometimes be affected by imagining. *See* ARTIFICIAL INTELLIGENCE; CONSCIOUSNESS; INTELLIGENCE.

Irvin Rock

Bibliography. K. R. Boff, L. Kaufman, and J. P. Thomas, *Handbook of Perception and Human Performance*, 2 vols., 1986; I. Rock, *Perception*, 1984; R. Sekuler and R. Blake, *Perception*, 3d ed., 1994.

Perciformes

The largest order of fishes, comprising 20 suborders, 160 families, about 1540 genera, and over 10,000 species, which is about 36% of all fish species. Eighty percent of the perciforms are marine species; in contrast, the Ostariophysi, which comprise the next three largest orders (Cypriniformes, ~3300 species; Characiformes, ~1700 species; and Siluriformes, ~2900 species), are virtually all freshwater inhabitants. The ostariophysans are dominant in the freshwaters of the world; the perciforms are dominant in the oceans. *See* OSTARIOPHYSI; OSTEICHTHYES.

The order Perciformes is not a monophyletic group, and it evades clear definition, cannot be defined cladistically, and has no unique specialized character or combination of specialized characters; however, without a doubt, it is the most diverse order of fishes. The following characters (or lack of characters) can suffice to characterize the perciforms: spines in the dorsal, anal, and pelvic fins (**Fig. 1**); dorsal and anal fins free from the caudal fin; pelvic girdle attached directly or by ligaments to cleithra or coracoids of the pectoral girdle (**Fig. 2**); pelvic fin with 1 spine and 5 soft rays, thoracic or jugular in position, and its base more or less vertical and placed well up on the side; no adipose fin; upper jaw bordered largely or entirely by premaxillae; no orbitosphenoid, mesocoracoid, epipleural, and epicentral bones; physoclistic (closed) swim bladder; double nostrils on each side of the snout; ctenoid scales; and caudal fin with 17 principal rays (15 branched). *See* SCALE (ZOOLOGY); SWIM BLADDER.

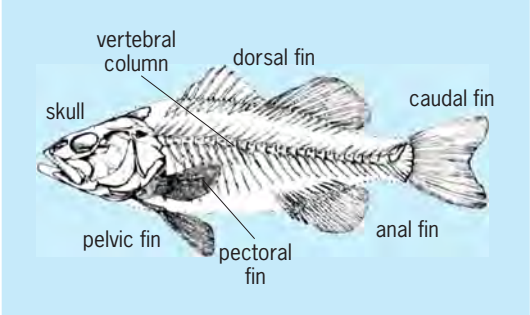

Fig. 1. Skeleton of a generalized perciform fish. The axial skeleton (head, vertebral column and ribs, and dorsal, anal, and caudal fins) is unshaded; the appendicular skeleton (pectoral and pelvic girdles and associated fins) is darkly shaded. (*From H. T. Boschung and R. L. Mayden, 2004*)

Hardly a single character listed above is without exception, as is seen in the suborder and family accounts that follow.

Adaptive radiation. The Perciformes first appeared in the Upper Cretaceous, experienced tremendous diversification, and throughout the Cenozoic became the dominant fish group in the marine environment, as well as having several large groups in freshwater environs. Many of the basic structural types, as well as most major perciform derivatives, such as the Pleuronectiformes and Tetraodontiformes, were present in the Eocene. The few families of perciforms that have been notably successful in freshwater are the Cichlidae in Africa and South America, the Centrarchidae in North America, the Percidae in North America and Eurasia, and the Anabantidae in Africa and southeastern Asia. Other families (for example, Chiasmodontidae) have effectively adapted to life in the deep seas, and still others (for example, Scombridae, Stromateidae, and Coryphaenidae) have become specialized for pelagic existence. However, it is in the shore areas, the offshore banks, the coral reefs, the coastal beaches and lagoons, and the intertidal zone that the perciforms have attained their greatest diversity. *See* PISCES (ZOOLOGY).

Classification. Ichthyologists are rapidly extending knowledge of phyletic relationships in this great order of fishes, so it is inevitable that classification is changing and controversial. For example, groups ranked as separate orders by some are included as suborders of the Perciformes by others; further, as a consequence of new discoveries, a family may be transferred from one order to another, or a new order may be named to contain it. The taxonomy of many of the 20 nominal suborders featured by J. S. Nelson (2006) is so tentative that a characterization of some of them is not attempted here.

Some 52 families of perciforms have a single genus, 23 have a single species (that is, are monotypic), and 21 have 100 or more species. Three suborders—Percoidei, Labroidei, and Gobioidei—account for over three-quarters of the species. The eight largest families—Gobiidae, Cichlidae, Serranidae, Labridae, Blenniidae, Pomacentridae,

Apogonidae, and Sciaenidae—together account for 5479 species, or about 55% of the perciform fishes.

A brief account is given for selected families, especially North American families, with emphases on the largest or the most interesting.

Suborder Percoidei. Percoidei is the typical perciform group; it lacks highly modified features such as are found in most other suborders. With certain exceptions, which are evident in some of the family accounts, the characters given for the order serve to define this suborder equally well. It is the largest suborder of perciforms, with 79 families, 549 genera, and about 3200 species. Twenty-six families have a single genus, and 10 are monotypic. The 10 largest families of Percoidei—that is, Serranidae (sea basses and groupers), Apogonidae (cardinalfishes), Sciaenidae (drums and croakers), Percidae (perches), Haemulidae (grunts), Carangidae (jacks, scads, and pompanos), Chaetodontidae (butterflyfishes), Pseudochromidae (dottybacks), Sparidae (porgies), and Lutjanidae (snappers)—collectively with about 2000 species, account for about 62% of the suborder.

Moronidae (temperate basses). The moronids are characterized by the following features: two-part dorsal fin, first with 8 to 10 spines, second with 1 spine and 10 to 13 soft rays; anal fin with 3 spines and 9 to 12 rays; lateral line extending onto the caudal fin; two opercular spines; and 7 branchiostegal rays.

Four species and several hybrids, in a single genus (*Morone* = *Roccus*), are native in freshwater, brackish, and coastal marine habitats of the Atlantic and Gulf of Mexico drainages. Two species in the genus *Dicentrarchus* occur in the eastern Atlantic from Norway to Morocco, the Canary Islands, and Senegal, and in the Mediterranean Sea and Black Sea. The striped bass (*Morone saxatilis*) [**Fig. 3**], the largest species of the family, is popular among surf fishers on the Atlantic coast from Maine to Florida. *See* BASS.

Serranidae (sea basses). The largest family of the suborder Percoidei, with 475 species classified in three subfamilies and 64 genera, is identified by the following characters: dorsal fin usually continuous or deeply notched, less commonly in two distinct parts; 7 to 13 dorsal spines, usually 10 or fewer; continuous lateral line, not extending onto the caudal fin; scales usually ctenoid, but cycloid in some species; 3 opercular spines; and usually 7 branchiostegal rays. They are hermaphroditic, and usually the two sexes do not develop at the same time. In addition to sea bass, species of serranids are known by a number of other common names, including bass, grouper, soapfish, hind, basslet, hamlet, and anthias.

Serranids are generally colorful fishes, dressed in all hues known in fishes, and color patterns may be vertical bars, longitudinal stripes, blotches, mottling, spotting, and combinations of the patterns (**Fig. 4**). Often juvenile coloration is quite different from that of adults. Serranids vary greatly in size: the smallest is a reef bass (*Pseudogramma guineensis*) in the eastern central Atlantic that is only about 2.7 cm (1.1 in.) in total length, and the largest, the goliath grouper (*Epinephelus itajara*), has a recorded weight of 455 kg (1000 lb). However, most serranids are less than 101 cm (40 in.) in length, and many are less than 10 cm (4 in.). Serranids are represented in all tropical and temperate seas of the world and in a variety of habitats. The majority inhabit coral reefs and rock outcrops, as well as benthic artifacts. Some species inhabit sandy bottoms, some are benthopelagic, and very few are freshwater inhabitants. Serranids feed

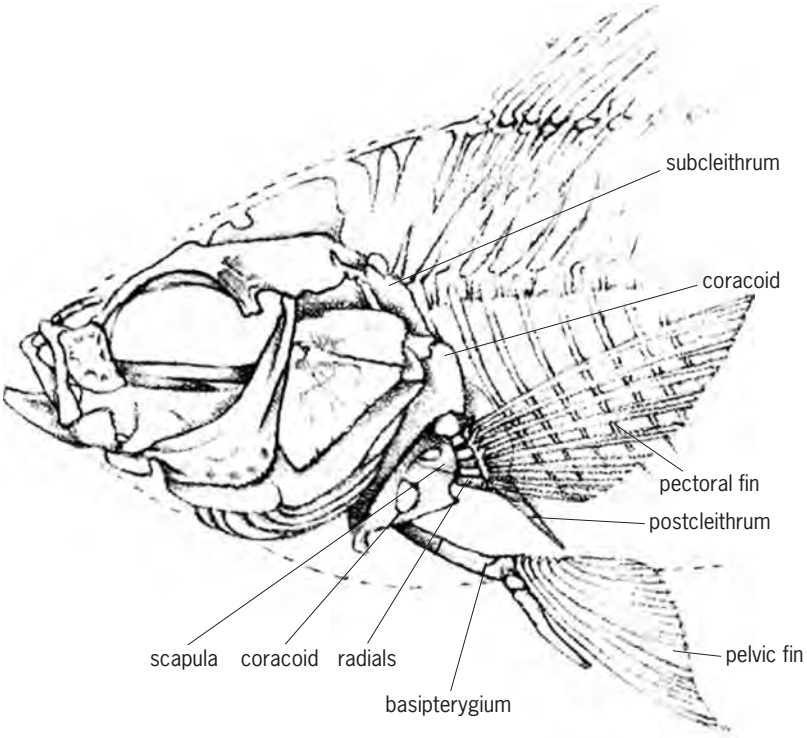

Fig. 2. Pectoral and pelvic girdles of a generalized perciform fish. (*From H. T. Boschung and R. L. Mayden, 2004*)

Fig. 3. Striped bass (*Morone saxatilis*). (*Photo from U.S. Environmental Protection Agency*)

Fig. 4. Yellowspotted grouper (*Epinephelus timorensis*). (*Photo by Richard C. Wass; Checklist of the Fishes from the National Park of American Samoa/NPS*)

on crustaceans and fishes. The larger species tend to be solitary ambush predators. These fishes are economically important as food and game, as well as supporting the aquarium trade. However; some of the species are known to have ciguatera poison, and certain soapfishes (tribe Grammistini) have a bitter-tasting skin toxin, grammistin, which can kill other animals in an aquarium. Unfortunately, several of the large groupers known in North American waters—for example, the speckled hind (*Epinephelus drummondhayi*), the Warsaw grouper (*E. nigritus*), and the goliath grouper (*E. itajara*)—are critically endangered [as per the International Union for the Conservation of Nature and Natural Resources (IUCN)]. *See* ENDANGERED SPECIES; MARINE FISHERIES; POPULATION VIABILITY.

Pseudochromidae (dottybacks). This is a relatively large family (about 29 genera and 120 species) of small fishes of the tropical Indo-Pacific. Pseudochromids typically have 1 to 3 spines in each dorsal and anal fin; pelvic fins with 1 spine and 3 to 5 soft rays, inserted below or in front of the pectoral fin base; 6 branchiostegal rays; variable lateral line (interrupted or continuous); lower jaw with ligamentous attachment to hyoid arch (a very unusual trait among percid fishes); and eggs with chorionic filaments.

Certain species differ markedly from the typical dottybacks. For example, the ellblennies (subfamily Congrogadinae) exhibit the following features: long and eellike body with small cycloid scales; only 1 spine in the dorsal fin (absent in one species); no anal spine; no pelvic fins on the majority of species, but in the jugular position when present; and dorsal and anal fins confluent with caudal fin in some species, slightly separate in most.

Dottybacks, or snakelet and eelblenny as a few species are called, are among the world's most colorful fishes and as a group among the world's smallest marine percids, traits that fit them well for the aquarium trade. The majority are less than 11 cm (4.3 in.) in length: the smallest, *Pseudoplesiops rosae*, is only about 2.6 cm (1 in.) in total length. In contrast, *Congrogadus subducens* attains a total length of 45 cm (18 in.) and at this size is about twice as long as the next largest member of the family.

In nature, the great majority of these fishes occur in or near holes and crevices in coral rock and rubble, where they feed on small invertebrates and fishes; however, some species (for example, eelblennies) inhabit coarse sand bottoms where weeds and sponges abound.

Opistognathidae (jawfishes). The jawfishes, comprising three genera and more than 78 species, are represented in the western and central Atlantic, Indian, and western and eastern Pacific oceans, including 16 species in North American waters, including Mexico. They are distinguished by a large mouth; cycloid scales on body, but with naked head; pelvic fins with 1 spine and 5 soft rays (the inner 3 weak and branched; the outer 2 stout and unbranched) that are underneath or usually anterior to the pectoral fins; continuous dorsal fin, with 9 to 12 spines and 12 to 22 soft rays; and lateral line following dorsal contour of the body and terminating at middle of the dorsal fin. Species of the genus *Stalix* are unique among fishes in having the tips of some of the dorsal fin spine transversely forked. Jawfishes are for the most part small colorful fishes ranging from about 3.5 to 55 cm (1.4 to 22 in.), with the majority less than 25 cm (10 in.). They live in sandy burrows with their heads exposed and feed on benthic and planktonic invertebrates. The males are mouthbrooders.

Centrarchidae (sunfishes). The family is identified by 3 or more anal fin spines; 6 to 13 dorsal fin spines confluent with soft dorsal fin, but usually separated by notch; rounded, emarginate or slightly forked caudal fin; 5 to 7 branchiostegal rays; small concealed pseudobranchiae; separate branchiostegal membranes, not connected to isthmus; villiform teeth in both jaws, vomer, and palatine bones, and tongue of some species; ctenoid scales (cycloid in *Acantharchus*) present over the entire body, as well as cheeks and opercles; and complete lateral line (except in *Lepomis symmetricus*). Bright colors on the cheeks and belly, especially of breeding males, are generally characteristic of sunfishes and are the character from which they probably take their name.

All centrarchids follow the same general reproductive behavioral pattern. The male constructs a nest by fanning a depression in the substrate. During spawning, a nuptial pair slowly swims side by side in a tight circle over the nest. Several females may spawn in a single nest. Fertilization is external, and the eggs sink to the bottom, where they adhere to the substrate. After spawning, the female either leaves the nest voluntarily or is chased away by the male. The male then becomes the sole guardian of the developing eggs and young, and defends them with great vigor, as attested by his willingness to attack intruders much larger than himself. He fans the eggs and larvae in order to oxygenate them and keep them free of silt.

Centrarchids are endemic to North America east of the Continental Divide, with the exception of the Sacramento perch (*Archoplites interruptus*), which is native to the Sacramento and San Joaquin valleys in California. Thirty-one species are native to the United States, 11 are native to southernmost Canada, and five are native to northernmost Mexico. Centrarchidae is not the largest North American perciform family, but it is undoubtedly the best known, made famous by the black basses (genus *Micropterus*), the most popular freshwater game fish in North America, and the ubiquitous sunfishes (colloquially brim or bream), popular pan fishes in the genus *Lepomis*.

Two subfamilies are recognized: Centrarchinae, with five genera, *Ambloplites* (4 species), *Archoplites* (1), *Centrarchus* (1), *Enneacanthus* (3), and *Pomoxis* (2); and Lepominae with one genus, *Lepomis* (12 species). *Acantharchus pomotis* (mud sunfish) and the genus *Micropterus* (7 species) are awaiting subfamily affiliation.

Genus *Ambloplites* has a moderately elevated and compressed oblong body; projecting lower jaw; well-developed supramaxilla; opercle with 2 flat

projections; angle of preopercle is serrated; teeth on the tongue, vomer, and palatine bones; continuous dorsal fin, with shallow gap between the spines and rays; 10 or 11 dorsal fin spines; 5 to 7 anal fin spines; and no bright red, orange, blue, or green colors. The species are *A. ariommus* (shadow bass), *A. cavifrons* (Roanoke bass), *A. constellatus* (Ozark bass), and *A. rupestris* (rock bass) [**Fig. 5**]. Maximum lengths of the smallest and largest species range from about 28 to 43 cm (11 to 17 in.).

Archoplites, a monotypic genus (*A. interruptus*, Sacramento perch), has a deep, moderately compressed body; large mouth with maxilla extending to under the middle of the eye; dorsal fin base about twice as long as anal fin base; 12 or 13 dorsal fin spines; and 6 or 7 anal spines. Maximum published length is 73 cm (29 in.) and weight is 1.44 kg (3.2 lb).

Centrarchus, also a monotypic genus (*C. macropterus*, flier), has a short and compressed body, with its depth about equal to its length; anterior profile concave over eye; relatively small and terminal mouth, with a projecting lower jaw; maxilla not reaching past the middle of the eye; large eye, with its diameter being greater than the length of the snout; well-developed supramaxilla; opercle without flat projections; teeth on the tongue, vomer, and palatine bones; 30 to 35 long slender gill rakers; 36 to 42 lateral scale rows; relatively tall and continuous dorsal and anal fins, about equal in size and shape; 11 to 13 dorsal fin spines; and 7 or 8 anal spines. The maximum length is 19 cm (7.5 in.).

Enneacanthus is identified by a short, deep, compressed body; oblique terminal mouth; maxilla not reaching posteriorly to the middle of eye; well-developed supramaxilla; opercle with 2 flat projections; preopercle not serrated; a rounded caudal fin; short and round pectoral fins; teeth on vomer and palatine bones, none on the tongue; continuous dorsal fin; 8 or 9 dorsal fin spines; and 3 anal fin spines. There are three species, *E. chaetodon* (blackbanded sunfish), *E. gloriosus* (bluespotted sunfish), and *E. obesus* (banded sunfish). These are colorful sunfishes, with a maximum length of less than 10 cm (4 in.), that inhabit vegetated lakes, ponds, quiet-waters sand- and mud-bottomed pools and backwaters of creeks, and small-to-medium rivers of the Atlantic and Gulf coasts below the Fall Line (the interface of upper coastal plains with terranes of high relief).

Fig. 5. Rock bass (*Ambloplites rupestris*). (*Photo by Eric Engbretson/U.S. Fish and Wildlife Service*)

Pomoxis (crappies) is a very distinctive centrarchid genus characterized by a strongly compressed oblong body with a long sloping anterior dorsal profile, which is concave above the eye; a projecting lower jaw; a well-developed supramaxilla; maxilla reaching beyond the pupil of the eye; an opercle with 2 flat projections; teeth on the tongue, vomer, and palatine bones; a continuous dorsal fin; 5 to 8 dorsal fin spines; 5 to 7 anal fin spines; and no bright red, orange, blue, or green colors. There are two species.

The presumed natural range of *P. annularis* (white crappie) is east of the Continental Divide and High Plains to the Appalachians, and from Minnesota and the lower Great Lakes south through the Mississippi Basin and the Gulf states from Alabama to Texas. The white crappies inhabit oxbow lakes, sloughs, impoundments, and large sluggish pools of low-gradient streams, often in rather turbid water and over bottoms of mud, clay, sand, or gravel. They usually are associated with submersed aquatic vegetation or submerged brush, logs, stumps, and tree roots. The largest white crappie listed by the International Game Fish Association (IGFA) is 2.35 kg (5.2 lb).

Pomoxis nigromaculatus (black crappie) is native to the upper Missouri River and the Lake of the Woods in southeastern Manitoba, Canada, to lower St. Lawrence River drainage, south on the Atlantic seaboard to peninsular Florida, and west to eastern Texas, including the Mississippi Basin east of the Great Plains. The black crappie typically inhabits large streams, backwaters, lakes, oxbows, ponds, and impoundments. It seems to be less adaptable than the white crappie to conditions typical of artificial ponds and impounded rivers. It fares best in clear waters amid submersed aquatic vegetation. The National Freshwater Fishing Hall of Fame (FHF) record is 2.72 kg (6 lb) from the Westwego Canal, Louisiana.

The two species can be distinguished by coloration, by the relative length of the dorsal fin base, and by the number of dorsal fin spines (5 or 6 in *P. annularis*; 7 or 8 in *P. nigromaculatus*).

Lepomis, with 12 species, is the largest centrarchid genus. It is characterized by a deep, rather strongly compressed body; jaws of about equal size; reduced or absent supramaxilla; opercle without a flat spinous projection, usually with a flap (ear); smooth-edged opercle, not serrate; teeth on the vomer and palatine bones, none on the tongue; fewer than 55 lateral scale rows; a continuous dorsal fin, with a shallow gap between the spines and rays; 9 to 12, typically 10, dorsal fin spines; 10 to 12, typically 11, rays; 3 anal fin spines; and bright red, orange, blue, or green colors.

The species are *L. auritus* (redbreast sunfish), *L. cyanellus* (green sunfish), *L. gibbosus* (pumpkinseed), *L. gulosus* (warmouth; previously in *Chaenobryttus*), *L. humilis* (orangespotted sunfish),

L. macrochirus (bluegill), *L. marginatus* (dollar sunfish), *L. megalotis* (longear sunfish), *L. microlophus* (redear sunfish), *L. miniatus* (spotted sunfish), *L. punctatus* (redspotted sunfish), and *L. symmetricus* (bantam sunfish).

Sunfishes are notorious for hybridizing, especially those in the genus *Lepomis*. Hybridization in nature may occur in areas where there has been ecological disturbance (thus negatively affecting visual acuity and species recognition), in ponds with an overabundance of vegetation and consequent reduction of open areas suitable for nest building, and in situations in which one species is greatly outnumbered by another. Also, sunfishes of different species build nests close together, and milt (sperm-containing fluid of a male fish) of one species can drift from one nest to another. Fortunately, fertilization by drifting sperm is minimized somewhat by the fact that the functional life of a sperm is only about 1 minute. Sunfishes, including the black basses (*Micropterus*), are hybridized artificially to produce hardier, faster-growing, and easier-to-catch progeny. *See* FISHERIES ECOLOGY.

Some sunfishes, especially bluegills, practice a reproductive strategy called female mimicry. Female mimics are small, sexually mature males that mimic the details of female behavior and gain access to functional females that were attracted to the nests of large, aggressive territorial males. Mimics are found in predominantly female aggregations, but when alone they do not attempt to construct or defend nest sites. They develop the barred color pattern of breeding females, approach the aggressive territorial male with impunity, and swim around in the nest with resident males. Ordinarily, female mimics become sexually mature at an earlier age than nesting males, and their testes are twice or more the size of potential nesting males of the same body size.

Acantharchus is another monotypic genus (*A. pomotis*, mud sunfish). It is unique among centrarchids in having cycloid scales. In addition, it has a rounded caudal fin; 3 or 4 parallel black stripes across the face and along the body; large mouth, with upper jaw extending to below the middle of the eye; 10 to 12 dorsal spines; and usually 5 anal spines. The maximum length is 21 cm (8.3 in.). The mud sunfish is an uncommon species that inhabits vegetated lakes, pools, and backwaters, usually over mud and detritus of creeks, and small to medium rivers of the Atlantic Coastal Plain and lower Piedmont drainages from the Hudson River to St. Johns River, and Gulf Coastal Plain drainages of Florida and Georgia from the Suwannee River to St. Marks River.

The genus *Micropterus*, comprising seven species, is commonly called black basses because of the black color of its larvae. It is identified by a moderately compressed, oblong body; large mouth, with projecting lower jaw; well-developed supramaxilla; 2 flat projections on the opercle; teeth on the tongue, vomer, and palatine bones; a distinct notch between the spinous and soft portions of the dorsal fins; typically 10 dorsal fin spines and 3 anal fin spines; an emarginate to slightly forked caudal fin; 59 or more lateral scale rows (other centrarchids have 54 or less); and no bright orange or blue colors.

Micropterus cataractae (shoal bass) is endemic to the Apalachicola drainage in Alabama, Georgia, and Florida, where it inhabits main-channel rapids and rock shoals. Unlike some of the other black basses, it is intolerant of reservoir conditions. For many years, the shoal bass was not distinguished from the redeye bass; in fact, the IGFA record shoal bass with a weight of 3.99 kg (8.8 lb) appears under the name redeye bass.

Micropterus coosae (redeye bass) is native to the Mobile Basin above the Fall Line (that is, the Coosa, Tallapoosa, Cahaba, and Black Warrior river systems), as well as upper Savannah and Chattahoochee river drainages in Alabama and Georgia. Its typical habitat is small upland streams that offer vegetative cover, and undercut banks and logs as well as rock ledges and boulders in flowing pools, but the species cannot sustain permanent populations in reservoirs and navigational pools. The redeye bass is the smallest species of the genus, with a maximum weight of about 1.2 kg (2.6 lb), but it has a well-deserved reputation for its "scrappy" resistance to capture. Where streams are too warm to sustain recreational trout fishing, the redeye bass serves as a good substitute.

Micropterus dolomieu (smallmouth bass) [**Fig. 6**] is endemic to the Great Lakes and St. Lawrence River drainages and the upland Mississippi Basin, but is absent from the Coastal Plain stream of the Atlantic seaboard and the Gulf of Mexico. It inhabits cool, clear streams having rocky and gravel substrates and moderate current, and pools between riffle areas (shallow areas of a stream in which the water flows rapidly over a rocky or gravelly bed). It also inhabits lakes, reservoirs, and navigational pools on large rivers, and is especially attracted to areas strewn with large rocks, as well as rock ledges in lakes and impoundments. The record weight (IGFA) is 4.93 kg (10.8 lb).

Micropterus notius (Suwannee bass) is a small bass, attaining a weight of about 1.75 kg (3.9 lb). It is endemic to the Suwannee River drainage in Florida and the Ochlockonee River drainage in Florida and Georgia, where it inhabits rocky riffles, runs, and pools, as well as large springs and spring runs.

Fig. 6. Smallmouth bass (*Micropterus dolomieu*). (*Photo by Eric Engbretson/U.S. Fish and Wildlife Service*)

Micropterus punctatus (spotted bass) ranges from the lower and central Mississippi Basin northward to southern Illinois and east in the Ohio River system to West Virginia, and also Gulf coastal streams from the Apalachicola drainage west to the Guadalupe River drainage in Texas. It occurs in several Atlantic coast rivers, but its native range does not extend east of the Appalachians. Spotted bass inhabit medium-sized and large streams, as well as reservoirs and navigational pools on large rivers, and it is quite tolerant of turbid waters and silty or muddy substrates. The record spotted bass (IGFA) of 4.44 kg (9.8 lb) is from an introduced population in California.

Micropterus salmoides (largemouth bass): As is true of most centrarchids, the original range of the largemouth bass has been so obliterated by early undocumented introductions that lines drawn now are conjectural. The presumed native range is Great Lakes drainages (exclusive of Lake Superior), the lower Mississippi Basin, drainages of the Atlantic Slope from the Savannah River southward throughout Florida, and Gulf drainages west to northern Mexico. Among the black basses, the largemouth bass is best adapted to life in lakes, ponds, oxbows, reservoirs, and still pools of medium-sized to large rivers. It can tolerate turbidity, at least for short periods, but prefers clear water as it depends primarily on sight to detect food. Also, it fares best in waters containing ample cover such as fallen trees, tree stumps and roots, and vegetation. In coastal areas, it readily enters brackish water to salinities as high as 24 parts per thosand (ppt). Largemouth bass found in brackish waters attain a smaller adult size. The largest specimen reported by the IGFA weighed 10.09 kg (22.2 lb).

Micropterus treculii (Guadalupe bass) is native to the Brazos, Colorado, Guadalupe, and San Antonio river drainages in Texas, where it inhabits gravel riffles, runs, and flowing pools of creeks and small to medium rivers. It is similar to the spotted bass but much smaller, attaining a weight of 1.67 kg (3.7 lb). *See* SUNFISH.

Percidae (perches). The Percidae, comprising 10 genera and 201 species, is by far the largest family of freshwater perciforms in the world. The family is identified by the following combination of characters: 2 dorsal fins, entirely separate or narrowly joined (broadly joined in the European genus *Zingel*); usually 2 (occasionally 1) anal spines; thoracic pelvic fins, with 1 spine and 5 rays; a serrate preopercle, except in many darters (Etheostomatinae); 5 to 8 branchiostegal rays; branchiostegal membranes free from the isthmus but sometimes united across it; well-developed to rudimentary pseudobranchiae; protractile or nonprotractile premaxilla; ctenoid scales; lateral line varying from complete to only a few pored scales in several species; present or absent breeding tubercles; villiform or cardiform teeth in the upper and lower jaws; teeth on vomer and palatine bones of most species; and swim bladder absent or reduced in darters, and well developed in *Perca* and *Sander*. The males of most darters exhibit brilliant colors, especially during the breeding season.

North American percids are confined entirely to the region east of the Rocky Mountains. All other North American fish families, with the exception of the cavefishes (family Ambylopsidae), either occur in western North America today or are known to have occurred there in the past. This distribution strongly indicates that Percidae immigrated from Eurasia (where it is generally conceded to have arisen) into North America relatively recently in a geological time. Relatively recent immigration makes the explosive evolution of the family in North America quite remarkable when one considers the small number of living species (about 15) inhabiting Eurasia.

Three subfamilies are recognized: Percinae, Luciopercinae, and Etheostomatinae.

Percinae is identified by a strongly serrate preopercle; greatly enlarged first interhaemal bone (bone between the haemal spines of the vertebrae that are associated with the spines and rays of the anal fin); well-developed anal fin spines; and a well-developed swim bladder. There are 3 genera and 8 species. One, *Perca flavescens* (yellow perch) [**Fig. 7**], is native to the United States and Canada. These are relatively large percids; the IGFA record for a yellow perch is 1.91 kg (4.2 lb).

Luciopercinae differs from Percinae in that none of the interhaemal spines are greatly enlarged, and the anal fin spines are weak. It comprises three genera and nine species, two of which are native to the United States and Canada. The sauger, *Sander* (long known as *Stizostedion*) *canadensis*, is native in large rivers from central Alberta to Quebec, the St. Lawrence River, southward west of the Appalachians to the Tennessee River drainage in Alabama, and north to Montana. The IGFA record size for a sauger is 3.96 kg (8.7 lb). *Sander vitreus* (walleye) is native east of the Continental Divide from northwesternmost Canada to the Gulf Coast. The Atlantic Slope and lower Mississippi Basin populations, considered introduced by most authorities, may represent natural populations. Populations west of the Continental Divide unequivocally are introduced. Walleyes inhabit lakes, large rivers, and tributary reservoirs but avoid mainstream reservoirs and navigational pools. The IGFA record size for a walleye is 11.34 kg (25 lb).

The 185+ species in the subfamily Etheostomatinae are identifiable by the following features: small size, usually less than 10 cm (4 in.) in length;

Fig. 7. Yellow perch (*Perca flavescens*). (*Photo by Eric Engbretson/U.S. Fish and Wildlife Service*)

preopercular margin smooth or only partly serrated; first interhaemal bone enlarged; usually 5 or 6 branchiostegal rays; and swim bladder reduced or absent. There are four genera, and all except eight species are in *Etheostoma* (136+) and *Percina* (41+).

Ammocrypta (sand darters) is characterized by an elongate, very slender, terete body; truncate or emarginate caudal fin; frenum absent; usually 6 branchiostegal rays; opercular spine absent; preopercular margin smooth or partially serrated; lateral line complete, not arched anteriorly; no scutes on breast and belly; 1 anal spine; swim bladder absent; breeding tubercles present in males, absent in females; live flesh translucent; and adult size small, less than 70 mm (2.7 in.). The genus comprises six species, which inhabit medium-sized to large sandy-bottomed streams and are noted for burying themselves in the sand. This behavior may serve several purposes: conservation of energy that would otherwise be used to maintain a position in the water current; cover from predators; or cover from which prey can be ambushed. The genus is represented in central Gulf Slope river systems below the Fall Line.

Crystallaria asprella (crystal darter) is the only species of the genus. It has an elongate, terete body, tapering posteriorly to a long, slender caudal peduncle; opercular spine; long and pointed snout; complete lateral line, not arched anteriorly; naked breast and belly; and maximum total length of about 160 mm (6.3 in.). It occurs in the upper Mississippi and Ohio river basins south to Gulf Slope streams from west Florida to Louisiana, typical in raceways and swift deep riffles of medium to large streams with an abundance of clean sand, gravel, or both.

Etheostoma (darters), with 136+ species, is the largest genus of fishes in North America. They are identified by the following combination of characters: body shape variable, elongate, terete to moderately compressed; flesh in life opaque; caudal fin slightly emarginate, truncate, or rounded; usually 5 or 6 branchiostegal rays, rarely 7; branchiostegal membranes narrowly to broadly connected across the isthmus; preopercular margin usually smooth; lateral line complete or incomplete, straight or arched upward anteriorly; no scutes on breast and belly; 1 or 2 anal fin spines; swim bladder absent or poorly developed; and adult size usually less than 110 mm (4.3 in.) in total length. The smallest species, golden darter (*E. denoncourti*), is about 34 mm (1.3 in.) in length, and the largest, greenside darter (*E. blennoides*), has a length of 170 mm (6.7 in.).

Percina (also called darters), with 41+ members, is similar to *Etheostoma*, but differs in having scutes on the breast and belly; opercular spine may be present; lateral line complete and straight; all with 2 anal spines; 6 branchiostegal rays, rarely 7; and larger size, generally over 110 mm (4.3 in.) total length. The smallest species, coal darter (*P. brevicauda*), has a length of about 58 mm (2.3 in.), and the largest, freckled darter (*P. lenticula*), is 200 mm (7.9 in.) in length.

Darters occur in a wide range of aquatic habitats varying from the dark waters and fetid substrates of swamps and backwaters to clear and clean waters trickling over stones and cobbles of mountain brooks. They inhabit the smallest rills and the largest rivers, still clear waters of springs, and torrential waters of large creeks and rivers, with substrates varying from soft mud to large boulders. Some species are very widely distributed, for example, the johnny darter (*Etheostoma nigrum*) occurs from the Hudson Bay region south to the Gulf Slope, whereas others are restricted to a mile or so of a single small stream or to several nearby springs.

Unlike the large percids (yellow perch, sauger, and walleye), which exhibit a more or less random spawning behavior and provide no parental care of eggs and fry, most of the darters have evolved a completely different reproductive strategy whereby they select a precise spawning site, and many species provide parental care of the eggs and fry. Also, the large percids release all their eggs at one time, whereas darters are partial (that is, fractional) spawners, releasing only part of their seasonal egg production at any one time. The hypothesized ancestral spawning behavior, egg burying, presumably is characteristic of all species in the genera *Ammocrypta*, *Crystallaria*, *Percina*, and some species of *Etheostoma*. Egg attaching involves sticking eggs to objects such as plant stems, leaves, roots, logs, and rocks. In most cases, the female selects objects onto which the eggs are to be attached. The majority of the species of *Etheostoma* attach their eggs in this way. A third reproductive strategy is egg clustering, which involves depositing eggs in clusters, most often by sticking them to the roof of a rock shelter prepared by the male. He clears the site of silt and debris by vigorously fanning his tail. As eggs are deposited, they are fertilized, and the cluster and ensuing fry are guarded, kept clean, and aerated by the male. The food of darters consists primarily of benthic arthropods, especially immature stages of aquatic insects.

Conservation. Percidae is the most imperiled large family of North American fishes. For example, in the southern United States (including subspecies and undescribed species but excluding hybrids), 31 are vulnerable, 23 are threatened, and 17 are in danger of extinction. Thus, about one-third of the darters are in some degree of jeopardy. Specialized reproductive behavior and high endemism are factors contributing to the large numbers of jeopardized darters.

Systematics. The monophyly of the genus *Etheostoma* has never been clearly demonstrated in any phylogenetic study. It is possible that some species of *Etheostoma* are more closely related to species of *Percina* or *Ammocrypta* than to other species of *Etheostoma*.

Apogonidae (cardinalfishes). Apogonidae is a family of primarily marine fishes comprising 28 genera and about 275 species of the Atlantic, Indian, and Pacific oceans. These are small fishes, varying from about 22 to 250 mm (0.86 to 9.8 in.), but the majority are less than 100 mm (4 in.) in total length. Cardinalfishes

have 2 dorsal fins, usually well separated (united in one genus), the first with 6 to 8 spines, the second with 1 spine and 8 to 14 soft rays; anal fin with 2 spines and 8 to 18 soft rays; scales usually ctenoid, cycloid in several groups, and completely absent in another; and many species are mouthbreeders (oral incubation). Many cardinal fishes do well in aquariums. Three genera, *Apogon* (19 species), *Astrapogon* (3), and *Phaeoptyx* (3), occur in North American waters, mostly on the Atlantic side.

Pomatomidae (bluefish). Pomatomidae is a monotypic family of predatory and voracious fish ranging throughout the tropical and temperate seas of the world, except for the eastern and central Pacific areas. A single species, *Pomatomus saltatrix*, makes up the family. *See* BLUEFISH.

Carangidae (jacks). Carangidae, a very diverse family of about 32 genera and 140 species, is characterized by the following characteristics: great variation in body shape, from very deep and extremely compressed to fusiform and moderately compressed; most species with small cycloid scales, ctenoid in a few; scales along the lateral line often modified into spiny scutes; 2 dorsal fins, first with 3 to 9 spines and second with 1 spine and usually 18 to 37 soft rays; first dorsal fin often very low or hidden by flesh, thus making the dorsal appear as a single fin; anal fin usually with 3 spines (first 2 separate from the rest of fin) and usually 15 to 31 soft rays; several genera with detached finlets behind the dorsal and anal fins; caudal fin deeply forked; and caudal peduncle slender and strengthened in many species with a row of strong scutes on each side. Generally the young differ markedly from adults in shape and color pattern.

The smallest carangid is the Indo-West Pacific razorbelly scad (*Aleles kleini*), measuring 16 cm (6.3 in.) in fork length, and the largest is the subtropical circumglobal yellowtail amberjack (*Seriola lalandi*), which is recorded to weigh 96.8 kg (213 lb).

Carangids occur in tropical and warm-temperate waters of the Atlantic, Indian, and Pacific oceans, where they inhabit coastal bays and lagoons, waters over reefs, or the open ocean. They are, for the most part, swift-swimming predators that feed on a wide variety of invertebrates and fishes. Worldwide, carangids are commercially important as food and game fishes; however, in the United States they are not exploited commercially, except the Florida pompano (*Trachinotus carolinus*) and recently the greater amberjack (*Seriola dumerili*).

Fifteen genera and 55 species of carangids occur in the Atlantic and Pacific waters of North America (including Mexico). The genera are listed below.

The following 11 genera have scales in the lateral line modified as scutes, quite small in some species, large in others:

- *Alectis ciliaris*, African pompano—body deep and compressed; steep anterior profile; bony scutes supporting caudal peduncle; lateral line strongly arched anteriorly; juveniles with very long trailing dorsal and anal fin filaments. Atlantic and Pacific.
- *Caranx* (14 species), *Hemicaranx* (3), and *Uraspis* (2), the jacks and trevallies (= crevalle)—body ovate to oblong and compressed but less so than in *Selene*; strong scutes on posterior lateral line and caudal peduncle; pectoral fin falcate; color mostly silvery, dusky dorsally; fins often yellow. Atlantic; mostly Pacific.
- *Decapturus* (5 species), *Selar* (1), and *Trachurus* (2), the scads—body elongate, moderately compressed; first dorsal fin high and separate from second by a gap; bony scutes supporting caudal peduncle; 1 dorsal and 1 anal finlet in *Decapterus*; lateral scutes extensive, extending entire length of body in *Decapterus*. Atlantic and Pacific.
- *Gnathanodon speciosus* (golden trevally)—similar to *Caranx*; color of juveniles and young adults bright yellow to silvery, with alternating broad and narrow black bars on body and yellow fins; lips thick and fleshy; breast completely scaled; bony scutes supporting caudal peduncle; pectoral fins falcate; and anal fin with 2 detached spines. Pacific.
- *Selene* (6 species), the lookdowns and moonfishes—very deep body and extremely compressed; anterior profile very steep; no large scutes supporting caudal peduncle; anterior lobe of second dorsal short in moonfishes, very long in lookdowns. Atlantic and Pacific.
- *Chloroscombrus* (2 species), the bumpers—body ovate, dorsal and ventral contours about equal; no large scutes supporting caudal peduncle; black spot on opercle and caudal peduncle. One species each in the Atlantic and Pacific.

The following 5 genera lack bony scutes on the caudal peduncle.

- *Naucrates ductor* (pilotfish)—relatively elongate; first dorsal low and hardly seen; caudal peduncle with fleshy keel; 5 to 7 broad black bands on body and extending onto dorsal and anal fins. Atlantic and Pacific.
- *Oligoplites* (3 species), the leatherjacks—body relatively long and rather strongly compressed; dorsal and ventral profiles equally curved; series of finlets following dorsal and anal fins; pectoral fin short; scales tiny and embedded; body silvery, bluish dorsally; fins yellow. This genus has special interest in that its dorsal and anal fin spines are venomous. A local species, *O. saurus*, arches its back and swims on its side, mimicking a floating leaf, undoubtedly a behavior designed to ambush an unsuspecting prey. Atlantic and Pacific.
- *Seriola* (6 species), the amberjacks, jacks, and rudderfish—body elongate, moderately deep; dorsal fin spines low but obvious; pectoral fin small; flesh known to cause ciguatera poisoning. This genus includes the greater amberjack (*S. dumerili*), which attains a weight of 80 kg (178 lb). Atlantic and Pacific.
- *Trachinotus* (7 species), the pompanos, permit, and palometa—body deep, not extremely compressed; second dorsal and anal fin falcate. Atlantic and Pacific.
- *Elagatis bipinnulata*, rainbow runner—body shape and fin configuration similar to *Seriola*, but more slender and certainly more colorful; greenish

blue dorsally, with more or less alternating longitudinal stripes of blue and yellow of varying widths on the side; and finlets behind the dorsal and anal fins. Atlantic; more common in the Pacific, where it tends to form large schools.

Lutjanidae (snappers). Lutjanidae, a family of 17 genera and 105 species, is characterized by the following features: dorsal fin continuous or with a shallow notch, with 9 to 12 spines and 9 to 18 soft rays; anal fin with 3 spines and 7 to 11 soft rays; pelvic fins inserted just behind the pectoral fin base; mouth terminal, moderate to large; usually 1 or 2 large canine teeth near the front of each side of the jaws, and small teeth on palatine bone and usually vomer bone; maxilla slips beneath preorbital shelf when mouth closed; supramaxilla absent; seven branchiostegal rays; caudal fin truncate to deeply forked; and 24 vertebrae (10 trunk + 14 tail). Snappers from deep rocky slopes are usually red, whereas reef types tend to be tan and yellow, often with some red colors. The smallest known adult, the spotstripe snapper (*Lutjanus ophuysenii*) from the Northwest Pacific, is 15.5 cm (6 in.) in length, and the largest is the cubera snapper (*L. cyanopterus*) from the western Atlantic, which weighs 57 kg (125 lb).

Snappers are represented in warm temperate to tropical waters of the Atlantic, Indian, and Pacific oceans, generally near the bottom, from the littoral zone to about 550 m (1800 ft). Juveniles of shallow-water species are common in seagrass beds, around mangroves, and on shallow reefs. Only three species are restricted to freshwater. Most species prey on crustaceans and fishes, and several are planktivores. Adults are primarily nocturnal feeders.

Worldwide, snappers are important food and game fishes; however, the flesh of certain species, if eaten, may cause ciguatera poisoning (for example, dog snapper and cubera snapper).

Seven genera and 28 species of the family Lutjanidae occur in marine waters of North America. Only a few bear a common name that does not include the word snapper. They include *Lutjanus*, with 20 species, including the schoolmaster; *Apsilus*, *Etelis*, *Oxycurus*, *Rhombloplites*, and *Hoplopagrus* (barred pargo), each with a single species; and *Pristipomoides*, with three species, two of which are called wenchman. Field identifications of snappers rely heavily on the colors of the body, fins, and iris; on fin shapes, especially the dorsal, anal, and caudal; and to a lesser extent on scalation, head profile, and shape of the vomerine tooth patch.

Haemulidae (grunts). Haemulidae, also known as Pomadasyidae, comprises 17 genera and 145 species. The family is identified by the following combination of characters: dorsal fin continuous, with 9 to 14 spines and 11 to 26 soft rays; anal fin with 3 spines and 6 to 18 soft rays; mouth small, with thick lips; teeth in jaws usually cardiform; vomerine teeth generally absent; usually enlarged chin pores; 7 branchiostegal rays; 26 or 27 vertebrae (10 or 11 trunk + 16 tail); and inside of mouth orange-red in *Haemulon*. These fishes produce sounds by grinding their pharyngeal teeth—hence the name grunt. For reasons unknown, some species (for example, bluestriped, white, and French grunts) pair in an unusual kissing performance. A maximum length of 105 cm (41 in.) is reportedly achieved by *Plectorhinchus harrawayi*, a western Indian Ocean species. The largest grunt in North American waters probably is the margate, *H. album*, reported to be 79 cm (31 in.) in total length.

Haemulids are chiefly marine, with many inhabiting brackish water (at least temporarily) and rarely freshwater. They occur mostly in tropical and subtropical coastal waters of the Atlantic, Indian, and Pacific oceans; relatively few species occur in temperate waters. Adults, typically inactive during the day, disperse to feed at night on benthic invertebrates over open sandy, muddy, or grassy areas. Juveniles are found mostly in seagrass beds in bays, lagoons, and coastal waters. A few species are planktivores. Worldwide, these are very important commercial and recreational fishes.

Nine genera and 44 species occur in marine waters of North America. The principal genus, *Haemulon*, whose species are named grunt, margate, tomtate, cottonwich, and sailors choice, differ from the other genera in having dense scales covering the basal area of the soft dorsal and anal fins. The other genera are *Anisotremus* (8 species), grunt, sargo, and porkfish; *Conodon* (2); *Haemulopsis* (4); *Microlepidotus* (2); *Orthopristis* (4, one called pigfish); *Pomadysys* (5), grunts; and *Xenichthys* (1) and *Xenistius* (1), called salema.

Sparidae (porgies). Sparidae, comprising 33 genera and 115 species, has a deep and compressed head and body. It is further identified by large eyes, small mouth, continuous dorsal fin that usually has 10 to 13 spines and 10 to 15 soft rays, anal fin with 3 spines and 8 to 14 soft rays, deeply forked caudal fin, maxilla covered by suborbital sheath when mouth closed, incisor-like or canine-like front teeth and molariform rear teeth, a pharyngeal "mill" of molariform teeth used to crush and grind shellfish, 24 vertebrae (10 trunk + 14 tail), and 6 branchiostegal rays. Most species are silvery, often with iridescent hues; others may have black or dusky vertical bars or dark blotches on the side, and uncommonly lateral stripes and body colors. Many species are hermaphroditic, some have male and female gonads simultaneously, and others change sex as they get larger. The maximum length is about 1.2 m (4 ft), but most species fall in the 30 to 60 cm (12 to 24 in.) category; the maximum weight is said to be more than 70 kg (150 lb).

Sparids inhabit seagrass beds in bays and coral reefs in shallow coastal waters and near banks of the continental shelf in the temperate to tropical zones of the Atlantic, Indian, and Pacific oceans. The greatest diversity of sparids is in the coastal waters around Africa. The only North American sparid known to enter freshwater is the Atlantic sheepshead, *Archosargus probatocephalus*. Sparids are important food and sport fishes; however, a few species have been implicated in cases of ciguatera poisoning.

The family is represented in North American waters by six genera and 18 species. The genera are *Archosargus* (2 species), sheepshead; *Calamus* (20), porgy; *Doplodus* (2), porgy and pinfish; *Lagodon* (1), pinfish; *Pargus* (1), porgy; and *Stenotomus* (2), scup and porgy. The largest species are *A. probatocephalus* (sheepshead) and *Pargus pargus* (red porgy), each at 91 cm (3 ft) and 9 kg (20 lb).

Sciaenidae (drums and croakers). Sciaenidae is a large family comprising 70 genera and 270 species. The combination of two simple characters—1 or 2 anal fin spines and a lateral line that extends to the end of the caudal fin—distinguish sciaenids from other percoid fishes. Additional family characters include the following: long dorsal fin, with a deep notch separating spinous from soft portion (rarely completely separate), first with 6 to 13 spines and second with 1 spine and usually 20 to 35 soft rays; second anal spine often larger than first; 6 to 13 anal rays; caudal fin slightly emarginate to rounded; upper bony edge of opercle forked, with bony flap present above the gill opening; barbels, when present, single or in a patch of small barbels on the chin; conspicuous pores on snout and lower jaw and frequently large cavernous canals (part of the acoustico-lateralis system) on the head; no teeth on roof of mouth; some species with molarlike teeth on pharyngeal bones; otoliths exceptionally large; swim bladder usually complex, with many branches and used as a resonating chamber; and 24 to 30 vertebrae.

Sciaenids are typically continental shore fishes of the Atlantic, Indian, and Pacific oceans. They are uncommon on island shores, and the farther from a continent the fewer the number. For example, there are only two genera and four species in the Bahamas and none in Hawaii. Continental distribution of sciaenids may be explained by their dependence on estuaries as nursery grounds for the young. A few species are restricted to freshwater, for example, *Aplodinotus grunniens* (freshwater drum), a North American species that ranges east of the Continental Divide from the Hudson Bay region to Guatemala (absent from Atlantic Slope streams and Gulf coastal streams east of Mobile Bay, Alabama), plus several genera in rivers of South America.

Twenty-six genera and 79 species of sciaenids are known from the coasts of North America (including Mexico), 51 of which are on the Pacific coast. Of these, 45 species occur in Mexico and eight northward. *Totoaba macdonaldi* (totoaba or totuava), one of the world's largest sciaenids, weighing 100 kg (220 lb), is endemic to the Gulf of California. Unfortunately this giant is in critical danger of extinction. Its size mate, *Argyrosoma regius* (meagre) of the eastern Atlantic is reported to weigh 103 kg (227 lb). The overall smallest sciaenids are in the Indo-West Pacific genus *Johnius*, one of which measures only 7.7 cm (3 in.) in total length.

Of the 12 genera and 28 species on the east coast of North American, only three genera, represented by four species, are not shared with the west coast; these are *Leiostomus xanthurus* (spot), *Pogonias cromis* (black drum), *Equetus lanceolatus* (jackknife-fish), and *E. punctatus* (spotted drum).

The following genera require mention: *Bairdialla*, *Odontoscion*, *Stellifer*, and *Larimus* are small, usually under 25 cm (10 in.) in length, with relatively large terminal mouth, no chin barbels, and well-developed second anal spine. *Micropogonias* (croaker) has chin pores and chin barbels, and *Pogonias* (drum) has numerous chin barbels and a much enlarged second anal spine, attains a weight of 50 kg (110 lb). *Leiostomus* (spot) lacks chin barbels and is the only local sciaenid with a distinctly forked caudal fin. *Sciaenops* (red drum) lacks chin barbels and has a conspicuous black spot (or spots) on the caudal peduncle. *Menticirrhus* (kingfish) and *Umbra* (sand drum) have a single stout chin barbel and inferior mouth. *Equetus* is reef-inhabiting and has the most conspicuous color pattern of western Atlantic sciaenids. Its extremely high first dorsal fin has a broad stripe from its tip to the body, where it curves posteriorly and onto the caudal fin; a similar stripe from the nape, across the opercle, and onto the pelvic fin; and another through the eye and across the cheek. *Cynoscion*, whose members are commonly called weakfish, seatrout, or corvina (in Mexico), differs somewhat from the typical sciaenids, with an elongate and moderately compressed body, large mouth with projecting lower jaw, and two canine teeth anteriorly in the upper jaw.

Sciaenids are, for the most part, bottom-dwelling fishes, primarily in muddy or sandy bays, estuaries, and shallow banks. They are carnivorus, feeding on benthic invertebrates and small fishes. Some are important food and game fishes, for example, species of weakfishes and seatrout, kingfish, and the red drum. The Atlantic croaker, *Micropogonias undulatus*, and some other sciaenids are principal components of pet food.

Chaetodontidae (butterflyfishes). Chaetodontidae is a very distinctive family comprising 11 genera and 122 species. It is distinguished by the following characters: disk-shaped strongly compressed body, with steep anterior profile; continuous dorsal fin, with 6 to 16 spines and 15 to 30 soft rays; anal fin with 3 to 5 spines (usually 3) and 14 to 23 soft rays; caudal fin, with 15 branched rays, rounded to emarginate; well-developed pelvic axillary process (a scute-like scale); no spine at angle of preopercle; some species with small serrations on preopercle; dorsal and anal fins densely scaled; mouth small, terminal, and protractile, with band or rows of small brushlike teeth; coiled gut; and two anteriorly directed processes in the swim bladder. Most species are brightly colored, and many feature dark vertical bands on the body and a dark band from the nape, through the eye, and across the cheek, as well an eyespot (ocellus) on the dorsal or posterior part of the body. Such a color pattern may be designed to confuse predators. Additionally, butterflyfishes are quick and agile, and their short, deep, compressed body allows quick turns so they can dart quickly into coral crevices to escape predators. If a defense posture is necessary, the fish lowers its head and erects its dorsal fin spines.

Butterflyfishes are mostly small fishes of coastal waters in warm-temperate and tropical seas. They are especially common around coral reefs and on rock bottoms, or near pilings and seawalls. By far, the greatest diversity is in the Indo-West Pacific, especially around Australia. A few species occur in brackish water. In coastal waters of North America, 7 species occur in the Atlantic and four in the eastern Pacific.

Butterflyfishes feed diurnally on coral polyps, small invertebrates, and filamentous algae. Some species are adapted for planktivory. The long-nosed species can reach into coral crevices and holes to obtain prey that most other fishes cannot reach.

Butterflyfishes are pelagic spawners. Postlarvae of most species have a tholichthys stage; that is, the head is encased in an armor of bony plates that extend onto the body. The tholichthys plates are lost during the juvenile stage. The larval stage was once thought to be an undescribed genus and was given the name *Tholichthy*—hence the term tholichthys. Juveniles are usually solitary, whereas adults of most species occur as heterosexual pairs.

Pomocentridae (angelfishes). Pomocentridae, a family of 8 genera and 82 species, is very similar to the Chaetodontidae but differs in having a sharp strong spine at the angle of the preopercle; no well-developed pelvic axillary process; swim bladder lacking anteriorly directed extensions; and larvae lacking a tholichthys stage. The family is further characterized by the following features: dorsal fin with 9 to 15 spines and 15 to 37 soft rays and anal fin with 3 spines and 14 to 25 soft rays; dorsal and anal fin rays with long extensions in some species; and larvae with spiny scales. Angelfishes possess spectacular coloration, which is markedly different between juveniles and adults of many species.

Angelfishes occur in tropical zones of the Atlantic and Indo-Pacific oceans, with the greatest diversity in the western Pacific. They generally inhabit areas near coral reefs, usually at depths less than 20 m (66 ft), where a variety of foods are consumed. Some species feed primarily on filamentous algae, and some on zooplankton; yet others are planktivorous; however, most species consume sponges, sessile invertebrates, and fish eggs. Most species do well in the aquarium, but some food specialists are difficult to maintain.

As far as known, all angelfishes are protogynous (that is, changing sex from female to male) hermaphrodites, maintain a haremic social system, and are pelagic spawners.

There are only three genera in North American waters: *Centropyge* (1 Atlantic species), *Holocanthus* (3 Atlantic and 2 Pacific), and *Pomacanthus* (2 Atlantic and 1 Pacific).

Suborder Elassomatoidei (pygmy sunfishes). Elassomatoidei consists of one family, Elassomatidae; one genus, *Elassoma*; and six species of small fishes. The species range in length from 3 cm (1.2 in.) to 4.6 cm (1.8 in.); however, specimens exceeding 4 cm (1.6 in.) are rare. The family is further characterized by the following: continuous dorsal fin, with 3 to 5 spines and 8 to 13 soft rays; anal fin with 3 spines and 4 to 8 soft rays; rounded caudal fin; only lacrimal bones of suborbital series remain; cycloid scales; no lateral line on body; branchiostegal membranes united across the isthmus; and bright orange and red colors lacking.

The family, with one exception, inhabits well-vegetated sluggish streams, backwaters, swamps, sloughs, and roadside borrow pits, usually over a mud substrate, in the Coastal Plain of the southeastern United States, on both the Atlantic and Gulf slopes. The exception, *Elassoma alabamae* (spring pygmy sunfish), is limited to spring-fed vegetated pools above the Fall Line in the Tennessee River drainage in Alabama.

Pygmy sunfish are carnivorous throughout life, feeding primarily on small crustaceans, mollusks, and larval insects, especially midges. Spawning takes place amid thick vegetation, and the fertilized eggs are adhesive and stick to plants, where they develop. After spawning, the female leaves the area willfully or is chased away by the male, which protects the developing eggs and fry. A life span of usually only 1 year makes these little fishes especially vulnerable to environmental insults that would prevent them from reproducing.

For many years, elassomatids were considered part of the sunfish family, Centrarchidae. The consensus is that elassomatids merit recognition as a distinct family, but their ordinal placement remains controversial. *See* SUNFISH.

Suborder Labroidei. Labroidei is a huge group of 6 families, 235 genera, and about 2275 species. Beside the esoteric characters pertaining to the pharyngeal region, especially the food-processing pharyngeal jaws, which distinguish the Labroidei, the group is noted for its numerous species that are protogynous and that display a wide diversity of color patterns associated with size and sex. Various populations may be monandric (adjective for monandry, comprising protogynous hermaphrodites among which all males are sex-reversed females) or diandrous (referring to monandry plus males incapable of sex change).

Cichlidae (cichlids). Cichlidae, comprising 112 genera and about 1350 species, is the most species-rich non-ostariophysan family of freshwater fishes in the world. For the most part, cichlids have a rather deep and compressed body similar to the familiar sunfishes in the family Centrarchidae; however, some are disk-shaped (body depth about equal to length—for example, the South American genus *Symphysodon*)—and others elongate and slender—for example, julies, genus *Julidochromis*, endemic to Lake Tanganyika, Africa. The family is further identified by the following characterstics: continuous dorsal fin with 7 to 25 spines and 5 to 30 soft rays; anal fin with 3 to 9 spines, 3 in the majority and a few species with up to 15; single nostril on each side; interrupted lateral line; usually 20 to 50 scales in lateral series, 100 or more in some species; and no subocular shelf.

Species of cichlids vary greatly in size. The smallest may be the South American genus *Apistogramma*,

many of which vary in length from 2.1 to 6.5 cm (0.8 to 2.6 in.). The largest reported cichlid (IGFA) is the speckled pavon, *Cichla temensis*, in the Amazon and Orinoco basins, which weighed 12.2 kg (27 lb). Another large cichlid, the butterfly peacock bass (*C. ocellaris*), which is one of about 20 cichlids introduced into the United States, attains a weight of 6.8 kg (15 lb). Cichlids exhibit a huge variety of colors and color patterns, and many are popular aquarium fishes. Especially popular in home aquariums are the atypical freshwater angelfishes (genus *Petrophyllum*).

Cichlids have highly organized breeding strategies involving parental care. One group, mouthbrooders (oral incubation), are usually polygamous, and it is usually the female that carries the fertilized eggs and fry in the mouth. Another group is called substrate brooders, that is, the eggs are deposited on the substrate, where both sexes of monogamous parents provide care. In yet another smaller group, the two strategies are combined; for example, eggs are laid on the substrate and are guarded while developing, but the newly hatched fry are taken into the parent's mouth, where they continue to develop. Mouthbrooding is common among African cichlids but is known from only a relatively few species in Central and South America. An interesting parental phenomenon is seen in the South American discus fishes, where the females secrete a whitish milklike substance from the skin that provides nourishment for their young.

Cichlids occur in Africa, Madagascar, Syria, Israel, Iran, Sri Lanka, coastal southern India, Central and South America, and the West Indies. The Rio Grande cichlid (*Cichlasoma cyanoguttatum*) in Texas is the only cichlid native to the United States. There are about 900 described species in Africa and several hundred more yet to be described. Lakes Malawi, Victoria, and Tanganyika each have several hundred species of cichlids. Such diversity is called a species flock, which is of special interest to students of evolutionary biology.

Embiotocidae (surfperches). Embiotocidae is a coastal marine family of 13 genera and 23 species, 19 of which are limited to the Pacific coast of North America and the remaining four occur on the coasts of Japan and Korea. One species, the Tule perch (*Hysterocarpus traskii*), is a freshwater species in north-central California. Surfperches differ from cichlids in being viviparous (giving birth to live young) and having an uninterrupted lateral line, and all species have 3 anal fin spines and a forked caudal fin. Maximum sizes vary from about 19 cm (7.5 in.) for a species from Japan to 47 cm (18.5 in.) for the rubberlip surfperch (*Rhacochilus toxotes*) that ranges from northern California to Baja California.

Pomocentridae (damselfishes). Pomocentridae is an all-marine family of 28 genera and 348 species occuring in tropical seas of the world, with the greatest diversity in the Indo-Pacific, especially from the Philippines to Australia. The family is characterized by the following: usually deep and compressed body (terete body in a group called plankton-pickers); continuous dorsal fin, with base of spinous part longer than base of soft part; 2 anal spines, rarely 3; small mouth, usually single nostril on each side; subocular shelf present; lateral line incomplete or interrupted; maximum length about 35 cm (14 in.); and parental care of developing demersal (bottom-deposited) eggs provided by males. Coloration of a species may vary from one locality to another.

Damselfishes exhibit a variety of life styles. One group of Indo-West Pacific species, called anemonefishes, have a commensal relationship with sea anemones (noted for their dangerous nematocysts), living near or within them with impunity. Some species have a commensal relationship with corals. Many species are highly territorial. Trophic types are herbivores, omnivores, or planktivores.

Six genera and 29 species occur in North American, with 14 in Atlantic waters and 15 in the Pacific. The genera *Chromis* (9 species) and *Azurina* (1) are in the subfamily Chrominae, which have the upper and lower edges of the caudal peduncle with 2 or 3 short spiny procurrent caudal rays. Conversely, the genera *Abudefduf* (4 species), *Hypypops* (1), *Microspathodon* (3), and *Stegastes* (11), in the subgenus Pomacentrinae, lack spiny procurrent caudal rays.

Labridae (wrasses). Labridae is the second largest family of marine fishes and the third largest perciform family, with about 68 genera and about 453 species. The family is identifiable by the following characteristics: protractile mouth; jaw teeth usually projecting outward; dorsal fin with 8 to 21 spines (usually fewer than 15) and 6 to 21 soft rays; anal fin with 2 to 6 spines (usually 3) and 7 to 18 soft rays; scales cycloid, generally large, with 25 to 80 in a lateral series, but some species have small scales numbering 100 or more; and lateral line continuous or interrupted.

Labridae is one of the most diversified of all fish families in shape, color, and size. The variety in body and fin form is unmatched in any other family of perciform fishes. Many species are highly colorful (**Fig. 8**), and several color patterns may exist within a species. Regardless of the huge number of wrasses, there is hardly a species among them that does not rival in absolute beauty the most exotic birds. Most species change color and sex with growth; from an initial phase of both sexes, the females can change

Fig. 8. Bluespotted wrasse (*Anampses caeruleopunctatus*). (*Photo by Richard C. Wass; Checklist of the Fishes from the National Park of American Samoa/NPS*)

into an often brilliantly colored male, the terminal phase. The largest wrasse is the humphead wrasse (*Cheilinus undulatus*) from the Indo-West Pacific and is about about 2.3 m (90 in.) in length, with a weight of 191 kg (420 lb). Unfortunately this giant wrasse is in danger of extinction. Many wrasses are less than 15 cm (6 in.) in length, with the shortest, *Minilabrus striatus*, from the Red Sea, measuring only 4.5 cm (1.8 in.).

Wrasses are for the most part nonschooling reef dwellers. They are active during the day and rest at night by burrowing into the sand or by secreting a mucous cocoon in which they rest. Trophic types range from planktivores to carnivores, the latter feeding on a variety of small benthic invertebrates, mollusks, as well as small fishes. Some small species are parasite pickers, that is, larger fishes "invite" the wrasses to enter their mouth and gill cavity where, in addition to the exterior, they feast on ectoparasites, thus ridding the large fish of an otherwise inescapable nuisance. Actually the wrasses establish "service stations" where the infester fishes visit on a regular basis.

Wrasses occur mostly in tropical zones of the Atlantic, Indian, and Pacific oceans, with the greatest diversity in the Indo-West Pacific. Three of the relatively few north temperate species occur on the Atlantic coast: the tautog (*Tautoga onitis*) ranges from Nova Scotia to South Carolina and is the one of the largest wrasses in the western North Atlantic, weighing 10 kg (22 lb), and is equaled in size by the hogfish (*Lachnolaimus maximus*), which ranges from Nova Scotia to northern South America; and the cunner (*Tautogolabrus adspersus*) ranges from the Gulf of St. Lawrence to Chesapeake Bay. Twelve additional genera and 39 species occur in subtropical and tropical waters of North America (including Mexico), with the numbers of Atlantic and Pacific species being equal.

Scaridae (parrotfishes). Scaridae comprises 10 genera and 88 species of perciform fishes. Parrotfishes have several physiological and behavioral characteristics in common with wrasses, such as sex changes, sexual dichromatism, being active by day and resting at night (some in a mucous cocoon), pelagic spawning, and mode of casual swimming, which is by sculling with the pectoral fins. However, the two families have some trenchant morphological differences. Parrotfishes are characterized by the following combination of features: remarkably consistent meristic (referring to number of body parts) characters; large thick cycloid scales, with 22 to 24 in lateral series; dorsal fin continuous, without a notch, and 9 spines and 10 soft rays; anal fin with 3 spines and 9 soft rays; jaw teeth fused (sometimes only at the base) and parrotlike; and mouth not protrusible. Coloration usually is brilliant, especially in terminal males where reds and greens dominate. Usually males and females are similarly colored at maturity.

Probably the smallest species is the western Atlantic bluelip parrotfish (*Cryptotomus roseus*), which attains a total length of 13 cm (5 in.). The largest reported is the green humphead parrotfish (*Bolbometopon muricatum*) from the Indo-West Pacific, which is 130 cm (51 in.) in length and weighs 40 kg (88 lb).

In the process of scraping algae from dead coral, bits of substrate are consumed and ground by the heavy pharyngeal teeth, a pharyngeal mill as it is called. Much of the fine sand found on and around coral reefs is thus produced. Parrotfishes are mostly tropical in the Atlantic, Indian, and Pacific oceans. The range of 14 species in 4 genera includes the Atlantic coast of the United States. The yellowtail parrotfish (*Sparisoma rubipinne*), occurs northward to Massachusetts, and the emerald parrotfish (*Nicholsina usta*), to New Jersey. The loosetooth parrotfish (*N. denticulate*) is the lone species on the west coast of the United States, and the range of 5 species includes the west coast of Mexico.

Herbert Boschung

Suborder Zoarcoidei. This all-marine group is one of the most atypical suborders. In some of its species, the dorsal and anal fins are confluent with the caudal fin; others lack pelvic fins, scales, and a lateral line.

Suborder Trachinoidei. The family Ammodytoidei (sand lances) is included in this suborder. Sand lances are marine, cool- or cold-water shore fishes with slender eel-shaped bodies. The family Chiasmodontidae has some species that have huge jaws equipped with long sharp teeth and that are capable of swallowing prey larger than themselves.

Suborder Blennioidei. One of the larger suborders of the Perciformes, Blennioidei contains chiefly small fishes that live in coral and rock reefs or in the tidal zone. Such groups as the large family Blenniidae (combtooth blennies) [**Fig. 9**] and Dactyloscopidae (sand stargazers) are mostly tropical or subtropical. [Most of the fishes of the permanently frigid waters surrounding Antarctica, the Nototheniidae and their relatives, have been placed in a separate suborder, the Notothenioidei.] In blennioids, pelvic fins either are lacking or are located on the throat, and there is an exact numerical correspondence between dorsal and anal soft rays and the vertebrae between them. Many blennioids are slender or eellike in form.

Suborder Callionymoidei. This suborder consists of two closely related families known as dragonets—colorful, little bottom fishes of marine waters of the

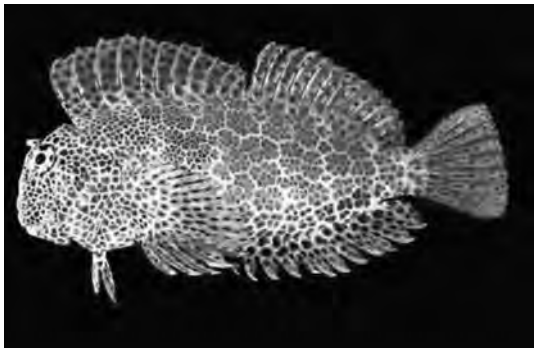

Fig. 9. Leopard (or shortbodied) blenny (*Exallias brevis*). (*Photo by Richard C. Wass; Checklist of the Fishes from the National Park of American Samoa/NPS*)

Fig. 10. Halfspotted goby (*Asterropteryx semipunctatus*). (*Photo by Richard C. Wass; Checklist of the Fishes from the National Park of American Samoa/NPS*)

Fig. 11. Whitespotted surgeonfish (*Acanthurus guttatus*). (*Photo by Richard C. Wass; Checklist of the Fishes from the National Park of American Samoa/NPS*)

tropics and subtropics. The body is elongate, the head is flattened, and one or more bones in the opercular area are spinous. The gill aperture is commonly a small pore located high on the head.

Suborder Gobioidei. The gobies make up one of the predominant groups of fishes in shallow shore waters, reefs, and estuaries of the tropics and subtropics; a few live in temperate waters. Some enter freshwaters near the coasts, and many skip nimbly over mudflats or about tide pools many feet from water. They vary from the smallest of fishes to nearly 2 ft (60 cm) in length, from somber hues to the brilliant and gaudy, and from short and stocky to a sinuous eel-shaped build. All lack a lateral line. Most belong to the family Gobiidae (**Fig. 10**), in which the pelvic fins are more or less united to form a sucking disk on the breast. The number of known kinds has been increasing rapidly, in part because the development of scuba gear has made the habitat of secretive reef species available to scientific collectors. The eight families usually recognized include over 200 genera and perhaps nearly 2000 species, but these figures are tentative because of the incomplete state of knowledge of the group.

Suborder Acanthuroidei. The Acanthuridae (surgeonfishes) [**Fig. 11**], Zanclidae (moorish idols), and Siganidae (rabbitfishes) are mostly inhabitants of the tropical and subtropical Indo-Pacific shores, though a few enter streams and some surgeonfishes live in the Atlantic. Many are highly colorful and showy in aquariums or on their native reefs. They are chiefly herbivorous.

Suborder Scombroidei. Scombroids are moderate-sized to large shore and oceanic fishes, all of which have a fixed premaxillae; that is, the jaws are not protrusible. There are six families, 46 genera, and about 147 species.

Sphyraenidae (barracudas): Elongate body, moderately compressed; jaws with large strong teeth; first dorsal fin, with 5 spines, well separated from second, with 1 spine and 9 soft rays. One genus (*Sphyraena*) [**Fig. 12**] and about 21 species in tropical to warm temperate Atlantic, Indian, and Pacific oceans. Once aligned with the atherinids, mugilids, and polynemids, *Sphyraena* is now considered the primitive sister group to the remaining scombroids. *See* BARRACUDA.

Trichiuridae (cutlassfishes): Elongate, strongly compressed bodies; long-based dorsal fin, continuous from the nape to the caudal fin; pelvic fins rudimentary or absent; and usually a small forked caudal fin. The common Atlantic cutlassfish, *Trichiurus lepturus*, known in the eastern Atlantic and western Atlantic from Cape Cod to the Gulf of Mexico to Argentina, lacks a pelvic and caudal fin, and the anal fin is hardly noticeable.

Gempylidae (snake mackerels): Similar to trichiurids but differing in having isolated finlets behind the dorsal and anal fins. Snake mackerels inhabit tropical and subtropical seas, often in very deep water.

Scombridae (mackerels and tunas): The Scombridae (mackerels and tunas) [**Fig. 13**] are streamlined, fast-swimming denizens of the high seas. Found worldwide, they are some of the most valuable food

Fig. 12. Bigeye barracuda (*Sphyraena forsteri*). (*Photo by Richard C. Wass; Checklist of the Fishes from the National Park of American Samoa/NPS*)

Fig. 13. Yellowfin tuna (*Thunnus albacares*). (*Photo by Richard C. Wass; Checklist of the Fishes from the National Park of American Samoa/NPS*)

and game fishes, being extremely popular for commercial fishing. *See* MACKEREL; TUNA.

The Istiophoridae (billfishes, including marlins) and Xiphiidae (swordfishes) have tremendously produced bills. They are highly prized sport fishes.

Suborder Stromateoidei. This suborder consists of six families. The Stromateidae (butterfishes) are mostly high-seas fishes, some of which live in association with or feed on coelenterates. All have teeth located in pharyngeal pockets in the throat. Some are excellent food fishes.

Suborder Anabantoidei. The labyrinth fishes include three families of freshwater fishes from southeastern Asia and Africa: Anabantidae, Helostomatidae, and Osphronemidae (climbing perches, gouramies, and relatives). Anabantoids have a specialized labyrinthine mechanism in the gill region that permits air breathing. Gouramies and their relatives are important food fishes, and many anabantids, such as the paradisefish and Siamese fighting fish, are aquarium favorites.

Suborder Channoidei. The Channidae (snakeheads) are swamp-living fishes that are valued as food in southeastern Asia. Like the anabantoids, they possess a specialized labyrinthine mechanism in the gill region that permits air breathing.

Suborder Kurtoidei. This suborder consists of a single genus, *Kurtus*, with two species of small Indo-Pacific fishes. They have a unique ossification that encloses the upper part of the swim bladder, and a conspicuous occipital hook in the male that is used to hold the eggs during brooding.

Other suborders. The remaining perciform suborders are Notothenioidei (Antarctic blennies), Pholidichthyoidei (convict blennies), Icosteoidei (ragfishes), Gobiesocoidei (clingfishes), Scombrolabracoidei (longfin escolar), and Caproidei (boarfishes).

Reeve M. Bailey

Bibliography. H. T. Boschung and R. L. Mayden, *Fishes of Alabama*, Smithsonian Books, Washington, DC, 2004; P. C. Heemstra and J. E. Randall, *FAO Species Catalogue*, vol. 16: *Groupers of the World (Family Serranidae, Subfamily Epinephelinae): An Annotated and Illustrated Catalogue of the Grouper, Rockcod, Hind, Coral Grouper and Lyretail Species Known to Date*, FAO Fish. Synop., 125(16):1–382, 1993; IGFA (International Game Fish Association), *World Record Game Fishes*, Pompano Beach, FL, 1999; *National Freshwater Fishing Hall of Fame*, FHF, Hayward, Wisconsin, 1999; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; L. M. Page and B. M. Burr, *A Field Guide to Freshwater Fishes*, Houghton Mifflin, Boston, 1991; C. R. Robins and G. C. Ray, *A Field Guide to Atlantic Coast Fishes of North America*, Houghton Mifflin, Boston, 1986; K. J. Roe, P. M. Harris, and R. L. Mayden, Phylogenetic relationships of the genera of North American sunfishes and basses (Percoidei: Centrarchidae) as evidenced by the mitochondrial cytochrome *b* gene, *Copeia*, 2002(4):897–905, 2002; W. F. Smith-Vaniz, Opistognathidae, jawfishes, pp. 2588–2589 in K. E. Carpenter and V. H. Niem (eds.), *FAO Species Identification Guide for Fishery Purposes: The Living Marine Resources of the Western Central Pacific*, vol. 4: *Bony Fishes, Part 2* (*Mugilidae to Carangidae*), FAO, Rome, 1999.

Percopsiformes

An order of the Paracanthopterygii, a fish superorder of questionable monophyly. Percopsiforms are identifiable by the following combination of characters: nonprotractile premaxilla; tooth-bearing ectopterygoid and palatine bones; absence of a suborbital shelf, orbitosphenoid, and basisphenoid bones; weak spines usually present in the dorsal fin; and pelvic fins, if present, behind the pectoral fins. There are three families, seven genera, and nine species, all confined to freshwaters of North America.

The order as presented herein is considered by some authors to be an unnatural group in that strong evidence linking the sister groups Aphredoderidae and Amblyopsidae with Percopsidae is lacking. *See* ACTINOPTERYGII; OSTEICHTHYES; TELEOSTEI.

Percopsidae (trout-perches). The family Percopsidae is distinguished from the other families of Percopsiformes by the following characteristics: alveolar premaxilla; toothless vomer; ctenoid and cycloid scales; scaleless head; dorsal and anal fins with 1 or 2 spines, and 9 to 12 and 6 or 7 soft rays, respectively; subthoracic pelvic fins, with 8 rays; adipose fin present; lateral line complete or nearly so; and anus in front of the anal fin. The family was well represented in Eocene times and presently represented by one genus and two species: *Percopsis omiscomaycus* (trout-perch; **Fig. 1**) and *P. transmontana* (sand roller).

The trout-perch occurs in lakes, deep flowing pools of creeks, and rivers from Alaska eastward to Quebec and southward to the Great Lakes and Mississippi River basins to West Virginia, Kentucky, Illinois, Missouri, North Dakota, and Montana. It is apparently extirpated in the Potomac and Susquehanna rivers. It attains a maximum length of 20 cm (8 in.).

Sand rollers are usually found near vegetation over sand in quiet backwaters and pool margins of small to large streams of the Columbia River drainage in Idaho, Washington, and Oregon. They attain a maximum length of 9.6 cm (3.8 in.).

Aphredoderidae (pirate perch). This monotypic family is represented by *Aphredoderus sayanus* (**Fig. 2**),

Fig. 1. Trout-perch (*Percopsis omiscomaycus*). (*Photo by Konrad P. Schmidt*)

Fig. 2. Pirate perch (*Aphredoderus sayanus*). (*Photo by Konrad P. Schmidt*)

which ranges in lowlands of Atlantic Slope streams from Long Island to Florida, the Gulf Slope west to the Brazos River drainage in Texas, the Mississippi Basin to Michigan, and disjunctly to Lakes Erie and Ontario, avoiding Appalachia. It is identified by the following combination of characters: robust, moderately deep body; anus in the normal position, just ahead of the anal fin, in juveniles and migrating forward in adults to between gill membranes (jugular position); ctenoid scales; scaly sides on side of head; normal eyes; dorsal fin with 3 or 4 spines and 9 to 11 soft rays; anal fin with 2 or 3 spines and 6 or 7 soft rays; subthoracic pelvic fin, with 6 or 7 rays; 18 branched caudal rays; adipose fin absent; lateral line incomplete; and numerous cutaneous sensory organs on the head and nape.

The pirate perch inhabits still backwaters and pools of creeks and rivers, swamps, oxbow lakes (crescent-shaped lakes formed when a meander of a river or stream is cut off from the main channel), and borrow-pits (areas, or pits, where earth has been dug up for use elsewhere), where dense vegetation abounds on a mud substrate. It feeds aggressively at night, primarily on isopods and amphipods, and attains a maximum total length of about 13 cm (5.2 in.).

Amblyopsidae (cavefishes). This family of five genera and six species is known from unglaciated southern and eastern regions of the United States. It is identified by the following characteristics: cycloid scales; scaleless head; small to vestigial eyes; dorsal and anal fin, each with 0 to 2 spines and 7 to 12 soft rays; pelvic fins usually absent; sensory papillae in rows on the head, body, and tail; myodome lost; anus between gill membranes (jugular position) in adults; adipose fin absent; and lateral line absent or incomplete. An annotated list of species is listed below.

Fig. 3. Ozark cavefish (*Amblyopsis rosae*). (*Photo from U.S. Fish and Wildlife Service*)

Chologaster cornuta (swampfish). The swampfish occurs in the Atlantic Coastal Plane from Roanoke River drainage, Virginia, to Altamaha River drainage, Georgia. It is found in quiet pools and backwaters of streams, and in swamps amid vegetation and debris. Characteristics include small functional eyes; no pelvic fins; strongly bicolored; 9 to 11 branched caudal rays; and maximum length of 6.8 cm (2.7 in.).

Forbesichthys agassizi (spring cavefish). The spring cavefish is found in central Tennessee, western Kentucky, southern Illinois, and southeastern Missouri, in springs and caves. Characteristics include very slender body; small functional eyes; pelvic fins absent; dark brown above, with three thin stripes on the side; 11 to 16 branched caudal fin rays; and maximum length of 8.4 cm (3.3 in.).

The remaining species, described below, are totally blind, usually without pigment, appearing white or pinkish, and are restricted to limestone caves.

Amblyopsis spelaea (northern cavefish). The species occurs in south-central Indiana to Mammoth Cave area in Kentucky. Characteristics include large broad head; very small abdominal pelvic fins present, with up to 6 rays, or rarely absent; 4 to 6 rows of papillae on caudal fin; 9 to 11 dorsal fin rays; usually 9 or 10 anal fin rays; 11 to 13 branched caudal fin rays; and maximum length of 11 cm (4.4 in.).

Amblyopsis rosae (Ozark cavefish). Occurring in southwestern Missouri, northeastern Oklahoma, and northwestern Arkansas, the Ozark cavefish is similar to the northern cavefish. Characteristics include lack of pelvic fins; 7 or 8 dorsal fin rays; usually 8 anal fin rays; 9 to 11 branched caudal fin rays; and maximum length of 6.2 cm (2.5 in.). This fish is protected as a threatened species (**Fig. 3**).

Typhlichthys subterraneus (southern cavefish). This species occurs in southernmost Indiana, central Kentucky, central Tennessee, northwest corner of Georgia, and northern Alabama and disjunctly west of the Mississippi River in central and southeastern Missouri, and northeastern Arkansas. Characteristics include long, rather flat head; prominent nares; vestigial eye tissue under the skin; terminal mouth; 11 to 15 branched caudal fin rays; numerous neuromast (sensory organs) ridges on the head and body; none, one, or two rows of papillae on the caudal fin; and maximum length of 8.6 cm (3.4 in.).

Speoplatyrhinus poulsoni (Alabama cavefish). This fish is known only from Key Cave, Tennessee River drainage in the northwest corner of Alabama. Characteristics include long flat head, constricted behind the snout; pelvic fin absent; unbranched rays in all fins (other amblyopsids have branched fin rays); incised fin membranes; 21 or 22 caudal fin rays; and maximum length of 7.2 cm (2.8 in.). It is protected as an endangered species. Herbert Boschung

Bibliography. H. T. Boschung and R. L. Mayden, Percopsiformes, pp. 361–367 in *Fishes of Alabama*, Smithsonian Books, Washington, DC, 2004; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; L. M. Page and B. M. Burr, *A Field*

Guide to Freshwater Fishes, Houghton Mifflin, Boston, 1991; W. J. Poly, Family Percopsidae Agassiz 1850—trout perches and sand rollers, *Calif. Acad. Sci. Annotated Checklists of Fishes*, no. 23, 2004; W. J. Poly, Family Aphredoderidae Bonaparte 1846—pirate perches, *Calif. Acad. Sci. Annotated Checklists of Fishes*, no. 24, 2004.

Performance rating

A procedure for determining the value for a factor which will adjust the measured time for an observed task performance to a task time that one would expect of a trained operator performing the task, utilizing the approved method and performing at normal pace under specified workplace conditions. Since the introduction of stopwatch time study more than a hundred years ago, there has been a need to adjust an observed operator time to a "normal time." Normal time is the time that a trained worker requires to perform the specified task under defined workplace conditions, employing the assumed philosophy of "a fair day's work for a fair day's pay."

Normal performance must ultimately be subjectively based on the observer's opinion of what represents a fair or equitable pace of activity for a worker with a specified task under specified conditions. It is assumed under the "fair day's work for a fair day's pay" philosophy that skilled operators of average physique for the tasks they are performing, performing widely differing tasks under widely varying conditions, will be equally fatigued at the end of an average day's work. During a normal workday, a specified amount of time is set aside as nonwork time and is called allowances. The performance rating process is concerned with determining normal pace during the work portion of an average day and must, therefore, consider the fatigue recovery aspects of allowance times occurring during the day.

The following two equations relate factors in determining how much time a worker will be allowed per unit of output:

$$\text{Standard time} = \text{normal time} + \text{allowances}$$
$$\text{Normal time} = \text{observed time} \times \text{rating factors}$$

If the observed time for a task is adjusted by the performance rating factor to determine normal time, and allowance time is added for nonwork time, the standard time will represent the allowed time per unit of production. For example, if the observed time is 60 min, and the rating factor is estimated to be 80%, and the allowances are 15% of normal time, then the normal time is $60 \times 0.8 = 48$ min, and the standard time is $48 + 0.15(48) = 55.2$ min. In a typical 8-h shift, the expected production from this worker would be $8(60)/55.2 = 8.7$ units of production.

Two frequently used benchmarks of normal pace are walking and dealing playing cards. These benchmarks are walking at a pace of 3 mi/h (1.33 m/s), and dealing a deck of 52 cards into four piles 1 ft (30 cm) apart in 0.5 min. Films of these activities being performed at different speeds are frequently viewed by performance raters so as to have their ratings be more consistent and more standardized.

Leveling factors for performance rating

Effort			Skill		
Category	Code	Value	Category	Code	Value
Excessive	A1	+0.13	Superskill	A1	+0.15
	A2	+0.12		A2	+0.13
Excellent	B1	+0.10	Excellent	B1	+0.11
	B2	+0.08		B2	+0.08
Good	C1	+0.05	Good	C1	+0.05
	C2	+0.02		C2	+0.03
Average	D	+0.00	Average	D	0.00
Fair	E1	−0.04	Fair	E1	−0.05
	E2	−0.08		E2	−0.10
Poor	F1	−0.12	Poor	F1	−0.16
	F2	−0.17		F2	−0.22

What has been observed in rating has been variously called speed, effort, tempo, pace, or some other word connoting rate of activity or exhibited effort. In 1927 a technique referred to as leveling, which required the evaluation of four factors—skill, effort, conditions, and consistency—was developed by the Westinghouse Corporation. Evaluation scales for the first two factors, effort and skill (see **table**) were employed in determining values for these factors. The factors of skill, conditions, and consistency were employed to adjust what was observed to what was intended as standard for the specified task. However, if an average trained operator is observed performing the specified task under specified conditions, there is no underlying theoretical justification for considering these factors. Conditions and consistency, as adjustment factors, are not in common use today in employing this technique.

The word "effort" has been defined in the past in such unfortunate terms as "demonstration of the will to work effectively," which is possibly unobservable and likely to be unmeasurable and irrelevant. Regardless of an employee's will, it is the pace of a worker that the employer has a right to expect (that is, employers do not pay for trying, they pay for accomplishment).

In 1949 the original Westinghouse method was revised extensively, and the new plan, called the performance rating plan, included three major classifications of factors—dexterity, effectiveness, and physical demand. Each classification contains from two to four factors. As indicated above, there is no theoretical justification for any factor other than pace, if the appropriate employee utilizes the specified method under specified conditions.

If two tasks of widely varying job difficulty are rated without consideration of the inherent relative job difficulty, and the two employees are paid on an incentive basis, the employee with the more difficult task is penalized in terms of potential incentive earnings, because although it may be an equally demanding task to perform at day work pace, the more difficult task is more difficult to increase in pace in

comparison to the simpler task. The objective rating technique, first published in 1946, considers job difficulty and, in so doing, adjusts for this problem. Such subfactors as amount of body movement, use of foot pedals, bimanualness, eye-hand coordination, handling requirements, and weight are considered in evaluating the primary factor—job difficulty. Pace is the other primary factor considered.

The most commonly employed rating technique throughout the history of stopwatch time study, including the present, is referred to today as pace rating. A properly trained employee of average skill is time-studied while performing the approved task method under specified work conditions. Rating consists only of determining the relative pace (speed) of the operator in relation to the observer's concept of what normal pace should be for the observed task, including consideration of expected allowances to be applied to the standard. *See* HUMAN-FACTORS ENGINEERING; METHODS ENGINEERING; WORK MEASUREMENT. Philip E. Hicks

Bibliography. R. M. Barnes, *Motion and Time Study*, 7th ed., 1980; M. E. Mundel, *Measuring Total Productivity in Manufacturing Organization: Algorithms and PC Programs*, 1987; M. E. Mundel and D. L. Danner, *Motion and Time Study: Improving Productivity*, 7th ed., 1994; B. W. Niebel, *Motion and Time Study*, 9th ed., 1992.

Performing arts medicine

A subspecialty in medicine that deals with problems specific to the activities of dancers, musicians, and vocalists. It is an outgrowth of the fields of sports medicine and occupational medicine.

Performing artists have medical problems similar to those of many athletes in that their activities involve certain parts of the body repetitively performing a specific motion. Athletes and performance artists often practice at least 3–5 h each day on their own and then may have team practice or rehearsal and finally actual competition or a performance. Unlike athletics, however, the performing arts are typically not seasonal, and so time for rest and rehabilitation is limited.

The vast majority of injuries in performers can be attributed to overuse. Unlike athletes, who may lift weights, run, or swim as a part of their training regimen, performing artists train by executing the same motions as when they perform. This adds to the overuse incurred during their work schedule. Many performing artists, particularly dancers, avoid lifting weights because they fear development of defined musculature and as a result may not develop proper muscle tone. Poor technique often plays an important role in overuse injuries: poor body positioning and inefficient use of abdominal musculature adversely affect dancers, singers, and musicians. Inappropriate choice of dance style, musical instrument, or vocal repertoire also may lead to overuse injuries. A musician should select an instrument that can be held and played comfortably, a singer should select pieces that fit his or her register, and a dancer should choose a dance form that complements his or her lower extremity anatomical structure. Another problem stems from the fact that professional performance artists do not have trainers, and they may be unaware of proper training techniques including warm-up, cool-down, stretching and strengthening exercises, and even proper nutrition.

Types of injuries. Each field of the performing arts tends to have its own characteristic injuries, which afflict the parts of the musculoskeletal system subjected to the greatest stress. Tendinitis in dancers occurs most commonly in the rectus femoris tendon in the thigh and the Achilles tendon in the leg. Stress fractures most often affect the metatarsal bones in the foot, the fibula in the leg, and the spine. Many dancers also suffer from pain related to mechanical stress on the kneecap, a consequence of jumping and deeply bending the knees. Kneecap pain is particularly prevalent among dancers whose lower extremity alignment is not ideally suited to their dance form and dancers who use poor technique.

Eating disorders are also common in dancers, particularly female ballet dancers, who strive to remain thin. For dancers, weight control is a major issue, and malnutrition is common; some dancers even have anorexia nervosa. In some cases, dietary disorders, low body weight, and intense training routines combine to produce amenorrhea. This lack of normal menstrual cycling can have disturbing secondary effects such as osteoporosis, which increases the likelihood of stress fracture. Most musicians, singers, and dancers perform on empty stomachs, and eat a large meal shortly before bedtime; doing so tends to result in the escape of stomach acids into the esophagus and even laryngitis and tends to add body fat. *See* ANOREXIA NERVOSA; MALNUTRITION; OSTEOPOROSIS.

Tendinitis in musicians most commonly occurs in the extensor tendons of the wrist and fingers, the flexor tendons of the fingers, and the rotator cuff muscles in the shoulder, producing pain with use of the upper extremities. Neck pain is also common, as is compression of a nerve (nerve entrapment) in the upper extremities. Nerve entrapments often manifest themselves with pain, weakness, and numbness and tingling in the fingers and arm.

Musicians playing wind instruments suffer from problems related to use of the embouchure, which is the opening from the upper airway into a wind instrument. Coordinated action of the lips and tongue in relation to the palate, teeth, and mouthpiece controls the flow of air into the instrument to create tone, pitch, and volume. Overuse produces neck muscle fatigue, pain at the articulation between the upper and lower jaws, and lip sores.

Singers are prone to overuse injuries of the vocal cords. These injuries take the form of swelling (laryngitis), nodules, and even hemorrhage. Excessive muscle tension in the neck, inadequate abdominal support, and singing beyond one's range may contribute to vocal cord injury. Vocal cord abuse may be

aggravated by allergy, upper respiratory tract infection, and breathing smoke or dry air, such as during airplane travel.

Evaluation and treatment. Determining the cause of a performing artist's medical problem is most important for treatment. Data on the performing artist's schedule, equipment, environment, and technique—as well as general stamina, specific muscular strength, and flexibility—are important in evaluating an injured performer. Identifying and changing specific factors that contributed to the injury not only help treat the problem but also prevent recurrence. Watching the artist perform is very helpful in detecting generalized tension and, particularly, harmful nuances of body positioning and muscle use.

Most performing artists want to remain active despite injury, and so the treatment should be designed to permit such activity whenever possible. Some injuries heal with relative rest, which may involve continuing to perform and rehearse but decreasing the practice schedule and eliminating certain parts of the repertoire that are particularly painful or that stress use of a certain body part. For the vocalist, relative rest may involve staying in midregister and at moderate volume intensity. Physical therapy modalities such as ultrasound, ice, and whirlpool are frequently used for spine and extremity problems, and hydration and steam inhalation are recommended for inflamed vocal cords. In addition, nonsteroidal anti-inflammatory medications are often prescribed. If poor technique, inadequate warm-up, poor nutrition, or psychological stress is contributing to the problem, it should be addressed. Relaxation techniques are especially helpful. During the rehabilitative phase of treatment, exercises to gain or maintain endurance, flexibility, and strength are critical. *See* SPORTS MEDICINE.

Carol C. Teitz

Bibliography. E. S. Gorman and C. A. Warfield, Pain syndromes in musicians, *Hosp. Practice*, 22:68–71, 1987; W. G. Hamilton, Foot and ankle injuries in dancers, *Clin. Sports Med.*, 7:143–173, 1988; J. A. Howard and A. T. Lovrovich, Wind instruments: Their interplay with orofacial structures, *Med. Prob. Perform. Art.*, 4:59–79, 1989; A. J. Ryan and R. E. Stephens, *Dance Medicine*, 1987; R. T. Sataloff, Common diagnoses and treatments in professional voice users, *Med. Prob. Perform. Art.*, 2:15–20, 1987; L. M. Vincent, *Competing with the Sylph: Dancers and the Pursuit of the Ideal Body Form*, 1979.

Pericycle

As commonly defined, the outer boundary of the stele of plants. Originally it was interpreted as a band of cells between the phloem and the innermost layer (endodermis) of the cortex. Such pericycle is commonly found in roots and, in lower vascular plants, also in stems. In higher vascular plants, however, a distinct layer of cells may not be present between the phloem and the cortex. The homogeneous groups of fibers located on the outer boundary of the primary phloem in many stems are not pericyclic in origin but develop in the earliest part of the phloem, whose remaining cells are obliterated. The pericycle, if present, may be composed of parenchyma or sclerenchyma cells with relatively thin or heavily thickened walls. It may be one to several layers in radial dimensions (**Fig. 1**).

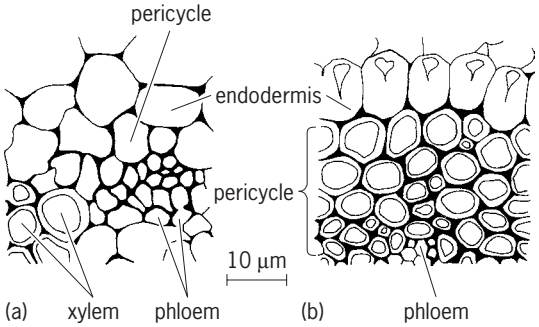

Fig. 1. Pericycles. (*a*) Part of transection of root *Actaea alba*, including xylem and phloem. Pericycle thin-walled and one cell in radial dimension. (*b*) Part of transection of root of *Smilax herbacea*, including phloem. Pericycle thick-walled and four or five cells in radial dimension (brace).

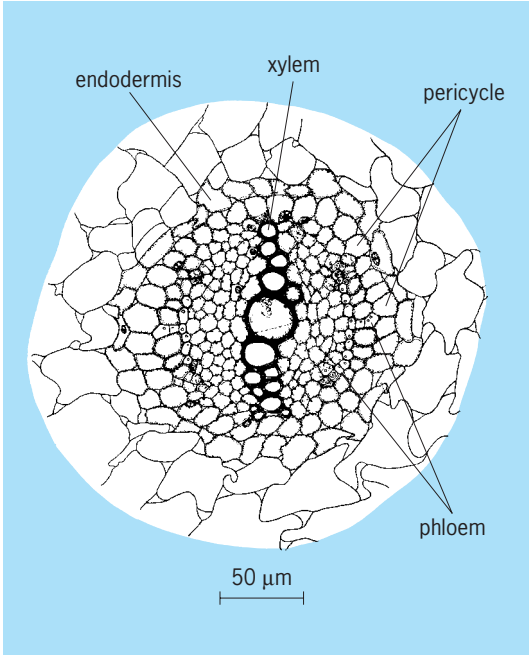

Fig. 2. Transection of central part of sugarbeet. (*After K. Esau, Hilgardia, 9(8), 1935*)

Primordia of branch roots commonly arise in the pericycle in seed plants, most frequently outside the xylem ridges (**Fig. 2**). The first cork cambium may also arise in the pericycle of those roots that have secondary vascular tissues. In roots, a part of the vascular cambium itself (that outside the primary xylem ridges) originates from pericycle cells. *See* CORTEX (PLANT); ENDODERMIS; LATERAL MERISTEM; PARENCHYMA; PHLOEM; PLANT TISSUE SYSTEMS; ROOT (BOTANY); SCLERENCHYMA; STEM; XYLEM.

Vernon I. Cheadle

Pericyclic reaction

Concerted (single-step) processes in which bond making and bond breaking occur simultaneously (but not necessarily synchronously) via a cyclic (closed-curve) transition state. Although a given reaction may appear formally to be pericyclic, it cannot be assumed to be a concerted process. In each case, the detailed mechanism of the reaction must be established experimentally. Pericyclic reactions can be promoted either by heat or by light; the stereochemistry of the reaction is determined by the mode of activation employed and the number of electrons that are delocalized in the transition state. *See* CHEMICAL BONDING; STEREOCHEMISTRY.

Classification. Four types of pericyclic reactions that are frequently encountered in organic chemistry are electrocyclic processes, cycloadditions, sigmatropic shifts, and cheletropic reactions.

Electrocyclic processes. These are reactions that involve either cyclization across the termini of a conjugated π-system with concomitant formation of a new σ-bond or the microscopic reverse. The sequence of steps involved in the forward reaction must be the same, in the reverse order, as that in the reverse direction when the forward and reverse reactions are carried out under identical conditions. This statement is known as the principle of microscopic reversibility.

The effect of the mode of activation upon the stereochemistry of an electrocyclic process is shown in reaction (1), where Me = methyl, for the hexatriene-

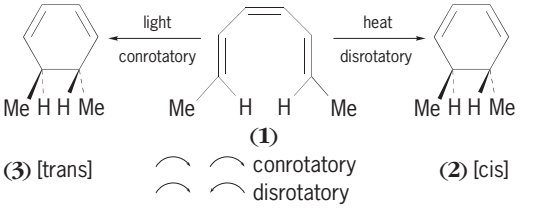

(1)

cyclohexadiene interconversion (a six-electron electrocyclic process). Thus, when *trans,cis,trans*-2,4,6-octatriene [structure (**1**)] is heated, disrotatory motion of the two terminal 2*p* orbitals occurs; that is, they rotate in opposite directions, thereby resulting in exclusive formation of *cis*-5,6-dimethylcyclohexa-1,3-diene (**2**). The corresponding photochemical process results in conrotatory motion of the termini in structure (**1**); that is, the two terminal 2*p* orbitals rotate in the same direction, thereby yielding *trans*-5,6-dimethylcyclohexa-1,3-diene (**3**) exclusively.

The stereochemistries of the corresponding thermal and photochemical four-electron processes (for example, butadiene-cyclobutene interconversion) are just the opposite of the results cited above for the hexatriene-cyclohexadiene interconversion. The thermal four-electron process occurs in conrotatory fashion, whereas the corresponding photochemical reaction is disrotatory.

Cycloadditions. These occur when two (or more) π-electron systems react under the influence of heat or light to form a cyclic compound with concomitant formation of two new σ-bonds that join the termini of the original π-systems. The stereochemistry of this reaction is classified with respect to the two molecular planes of the reactants. Thus, if σ-bond formation occurs from the same face of the molecular plane across the termini of one of the component π-systems, the reaction is said to be suprafacial on that component. If instead σ-bond formation occurs from opposite faces of the molecular plane, the reaction is said to be antarafacial on that component. This distinction is illustrated in reaction (2) for two thermal processes, where symbol $\neq$ indicates the structure of the transition state. Reaction (2*a*) shows the Diels-Alder [4+2] cycloaddition of butadiene (**4**) to ethylene (**5**), a six-electron pericyclic reaction in which additions across the termini of the diene (four-electron component) and dienophile (two-electron component) both occur suprafacially. Reaction (2*b*) shows a [14+2] cycloaddition in which σ-bond formation occurs suprafacially on the two-electron component tetracyanoethylene (**6**)] and antarafacially on the fourteen-electron component [heptafulvalene (**7**)]. [Note the trans relationship between the two circled hydrogen atoms in the product (**8**)]. In the

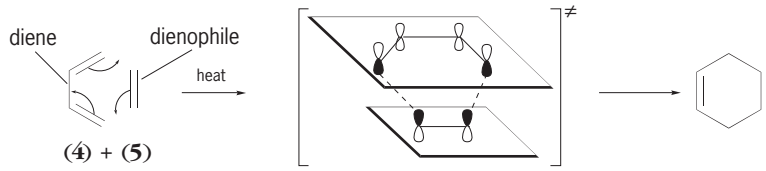

(2*a*)

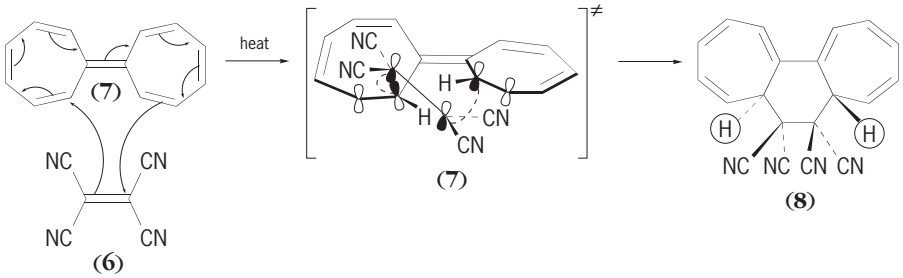

(2*b*)

orbital representations of the transition states for these reactions, the π-lobes which interact constructively to form new σ-bonds are shaded and are connected by dotted lines. *See* DIELS-ALDER REACTION.

A different mode of activation, that is, light rather than heat, is required to promote cycloaddition between *trans*-stilbene (**9**) and tetramethylethylene (**10**) [reaction (3), where Me = methyl, Ph = phenyl].

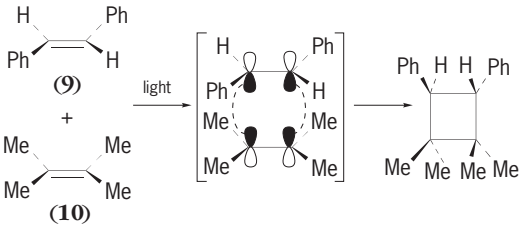

Ph = phenyl (3)

This photochemical [2+2] cycloaddition proceeds through the lowest-energy electronically excited state (rather than through the ground state). Two new σ-bonds are formed; the overall process is suprafacial on both components. *See* PHOTOCHEMISTRY.

Sigmatropic shifts. These reactions involve migration of a σ-bond that is flanked at either (or both) ends by conjugated π-systems. Either one or both ends of the σ-bond may migrate to a new location within the one or more flanking π-systems. The shift is said to be of order [*i*,*j*] when the σ-bond migrates from the [1,1] position to some distant position, [*i*,*j*], along the one or more conjugated chains. Thermal rearrangement of 3-methyl-1,5-hexadiene (**11**) to 1,5-heptadiene (**12**) is an example of a sigmatropic shift [reaction (4), where Me = methyl.]. This reaction, the Cope

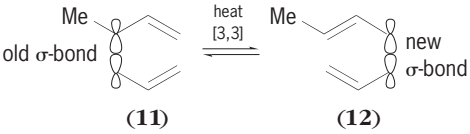

(4)

rearrangement, is a sigmatropic shift of order [3,3].

The stereochemistry of a [*i*,*j*] sigmatropic shift is classified with respect to the movement of the migrating group along the π-system with respect to the plane that contains the π-electron system. If the migrating group moves across one face of the π-system, the reaction is suprafacial. If instead the migrating group crosses the molecular plane during this process, the reaction is antarafacial. In both instances, migration can occur with either retention or inversion of configuration of the migrating group. Reaction (5) shows the thermal rearrangement of structure (**13**) to structure (**14**); this involves a suprafacial sigmatropic shift of order [1,3] that proceeds with inversion of the configuration of the migrating group. [Note that the substituents OAc and D are mutually trans in structure (**13**), but in the product, structure (**14**), they are cis].

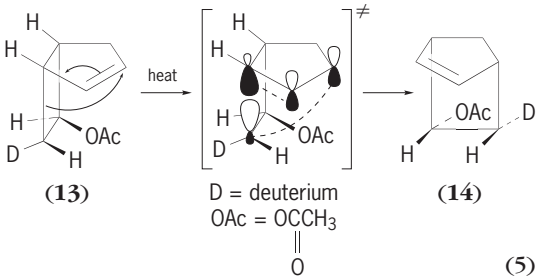

(5)

Cheletropic reactions. These are reactions that involve extrusion of a fragment via concerted cleavage of two σ-bonds that terminate at a single atom or the reverse process. Cheletropic fragmentations may be either linear or nonlinear [reaction (6*a*)]. An example is reaction (6*b*), in which 7-norbornadienone (**15**)

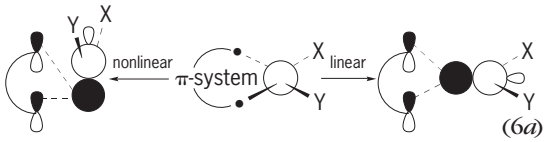

(6*a*)

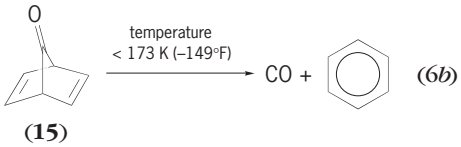

(6*b*)

decomposes at very low temperature via cheletropic extrusion of carbon monoxide (CO) with resulting aromatization. (In this example, it is not possible to demonstrate experimentally whether carbon monoxide is extruded in linear or nonlinear fashion.)

Theory. R. B. Woodward and R. Hoffmann introduced an application of molecular orbital theory that permits prediction of rates and products of pericyclic reactions. They utilized symmetry properties of molecular orbitals to estimate relative energies of diastereoisomeric transition states for structurally similar pericyclic reactions. In their approach, the energy changes that accompany orbital interactions as reactants are converted to products along the reaction coordinate are followed via construction and analysis of orbital correlation diagrams. An element of symmetry is chosen that is preserved throughout the reaction and that bisects bonds that are either being broken or formed during the pericyclic reaction. Central to their argument is the principle of conservation of orbital symmetry; that is, the symmetry of individual molecular orbitals with respect to the chosen symmetry element will be preserved as the reaction coordinate is traversed. If the variation in total energy of the system with reaction geometry along the reaction coordinate is small, little resistance to reaction will occur, and the reaction is predicted to occur readily; that is, the reaction is said to be symmetry-allowed. However, if any of the occupied molecular orbitals experience considerable destabilization while traversing the reaction coordinate, hindrance to reaction will result; that is, the reaction is symmetry-forbidden. It should be noted that these arguments apply to the rates of pericyclic reactions and not to the position of thermodynamic equilibria (which is determined by the relative stability of reactants versus products).

In an alternative theoretical approach to understanding pericyclic reactions, the transition state

is examined directly, and attempts to estimate the degree of electronic stabilization (allowedness) or destabilization (forbiddenness) inherent in that transition state are made. One such approach emphasizes the importance of frontier orbitals (highest-occupied–lowest-unoccupied molecular orbitals) in determining the course of a pericyclic reaction. Thus, if donation of electrons from occupied to unoccupied frontier orbitals results in a stabilizing interaction, the reaction should be allowed; if instead this results in a destabilizing interaction, the reaction will be forbidden.

A second approach directly estimates the energies of relative transition states by noting whether delocalization of electrons results in an aromatic (symmetry-allowed) or an antiaromatic (symmetry-forbidden) transition state. As an example, ground state (thermal) [4+2] cycloadditions proceed via a six-electron cyclic transition state that is Hückel-aromatic and thus is symmetry-allowed. Thermal [2 + 2] cycloadditions, however, proceed via a four-electron cyclic transition state that is isoconjugate with cyclobutadiene and thus is Hückel-antiaromatic and symmetry-forbidden.

Another example is provided by the butadiene-cyclobutene interconversion. The disrotatory electrocyclic process is isoconjugate with cyclobutadiene and hence antiaromatic (a symmetry-forbidden ground-state reaction). In contrast, the corresponding conrotatory process results in an array of π-orbitals that resembles a Möbius strip (that is, the top of one π-terminus overlaps with the bottom of the other). This latter array is isoconjugate with so-called Möbius cyclobutadiene and is aromatic (a symmetry-allowed ground-state reaction). *See* CONJUGATION AND HYPERCONJUGATION; DELOCALIZATION; MOLECULAR ORBITAL THEORY; ORGANIC REACTION MECHANISM; PHYSICAL ORGANIC CHEMISTRY; WOODWARD-HOFFMANN RULE. Alan P. Marchand

Bibliography. G. Desimoni et al. (eds.), *Natural Products Synthesis through Pericyclic Reactions*, 1983 S. N. Ege, *Organic Chemistry*, 5th ed., 1998; R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; I. Fleming, *Pericyclic Reactions*, 1998; I. Fleming, *Frontier Orbitals and Organic Chemical Reactions*, 2d ed., 2006; E. A. Halevi, *Orbital Symmetry and Reaction Mechanisms: The OCAMS View*, 1992; G. M. Loudon and J. E. Stowell, *Organic Chemistry*, 4th ed., 2003; T. H. Lowry and K. S. Richardson, *Mechanism and Theory in Organic Chemistry*, 3d ed., 1987; S. Sankaraman, *Pericyclic Reaction—A Textbook: Reactions, Applications and Theory*, 2005; R. B. Woodward and R. Hoffmann, *The Conservation of Orbital Symmetry*, 1970.

Periderm

A group of tissues which replaces the epidermis in the plant body. Its main function is to protect the underlying tissues from desiccation, freezing, heat injury, mechanical destruction, and disease. Although periderm may develop in leaves and fruits, its main function is to protect stems and roots. The fundamental tissues which compose the periderm are the phellogen, phelloderm, and phellem (**Fig. 1**). *See* EPIDERMIS (PLANT).

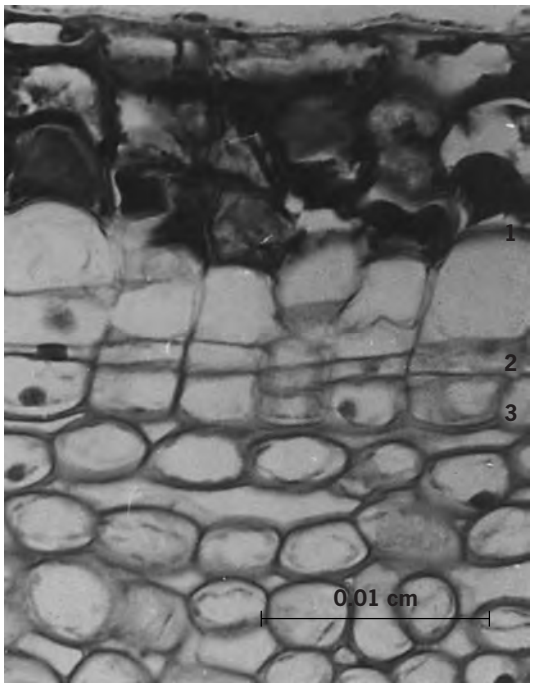

Fig. 1. Cross section of the stem periderm of *Sambucus nigra*: (1) phellem; (2) phellogen; (3) phelloderm.

Phellogen. The phellogen is the meristematic portion of the periderm and consists of one layer of initials. These exhibit little variation in form, appearing rectangular and somewhat flat in cross and radial sections, and polygonal in tangential sections (**Fig. 2**).

It is commonly accepted that the first phellogen of the shoot is derived from mature parenchymatous or collenchymatous cells of the cortex or from epidermal cells. Subsequent phellogens may originate from parenchymatous cells of the cortex or phloem. Phellogen initation occurs a short time before or concurrently with the rupture of the epidermis. Such a rupture may result from an internal increase in diameter or from external damage. The cells that give rise to the phellogen usually cannot be distinguished from other cells of the cortex by their shape. However, shortly before their first division, those cells

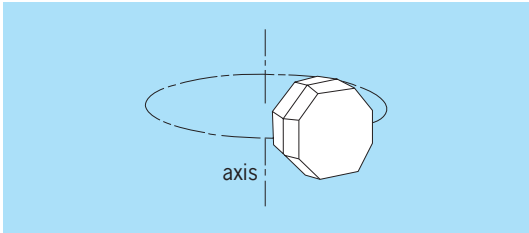

Fig. 2. Schematic representation of a phellogen cell, shown in its natural orientation.

lose their central vacuoles, their cytoplasmic volumes increase, and prominent quantities of tannins, leucoanthyocyanins, and chlorogenic acid are accumulated. The time at which the first phellogen is initiated varies with species, organs, and environmental conditions. In most species, phellogen is initiated within a month or two after germination, or at least within the first year. However, in a few temperate and subtropical tree species (for example, *Ceratonia siliqua*) the initiation of the phellogen and the subsequent formation of the periderm take place only after several years.

The cell divisions that follow phellogen initiation are mostly periclinal (that is, walls are formed parallel with the circumference). However, occasional anticlinal divisions occur which increase the circumference of the phellogen initial layer. *See* CORTEX (PLANT); PARENCHYMA; STEM.

Periods of phellogen activity alternate with periods of inactivity, even during the same year. The activity of the phellogen within a growth period shows variations in rate and can be modified by environmental conditions. Short-day conditions stimulate phellogen activity, while long-day conditions hinder it, as with *Robinia pseudacacia*. Extreme temperatures—both high and low—accelerate activity, as with *Acacia* sp. Hyperhydric conditions seem also to cause both early initiation and intensive activity of the phellogen, as with *Eucalyptus* sp. In many cases, conditions that inhibit activity of the vascular cambium stimulate the activity of the phellogen, and vice versa. The first phellogen may enter dormancy and resume activity several times, but in most cases it finally loses its meristematic activity and undergoes a complete differentiation. In such a case, a new layer of phellogen, or subsequent phellogen, is initiated in the tissues beneath. *See* PHOTOPERIODISM; PLANT-WATER RELATIONS.

The phellem produced by subsequent phellogens will usually enclose layers of dead cells of the cortex and phloem. Such a dead outer periderm is named rhytidome, or shell bark. *See* PHLOEM.

The manner in which the subsequent periderms are formed determines to a large extent the form of rhytidome exfoliation. Usually when the first periderm is shallow beneath the epidermis, the subsequent periderms are formed in overlapping patches, and the rhytidome is exfoliated in the form of scales (scaly bark). When the first periderm is initiated deep in the cortex, subsequent periderms will also encircle the axis and the rhytidome will be exfoliated in the form of a hollow cylinder (ring bark).

The phellogen of the roots is usually initiated in the pericycle. *See* PERICYCLE; ROOT (BOTANY).

Phelloderm. The phelloderm cells are phellogen derivatives formed inward. The number of phelloderm layers varies with species, season, and age of the periderm. In some species, the periderm lacks the phelloderm altogether. The phelloderm consists of living cells with photosynthesizing chloroplasts and cellulosic walls. These cells may show a less organized tier arrangement than the phellogen or phellem cells and are hard to distinguish from the parenchymatous cells of the cortex and phloem. Sclerified phelloderm cells are found in various species of *Eucalyptus* and *Acacia*.

Phellem. The phellem, or cork, cells are phellogen derivatives formed outward. These cells are arranged in tiers with almost no intercellular spaces except in the lenticel regions. After completion of their differentiation, the phellem cells die and their protoplasts disintegrate. The cell lumens remain empty, excluding a few species in which various crystals can be found. Distinguishable rings of thin-walled and wide-lumened cells and thick-walled and narrow-lumened cells develop seasonally or intraseasonally. The walls of the cork cells consist of a thin cellulosic primary wall and a thick suberized secondary wall (suberin lamella). In some species an inner cellulosic lamella can also be found. The remarkable impermeability of the suberized cell walls is largely due to their impregnation with waxes, tannins, cerin, friedelin, and phellonic and phellogenic acids. In various species, nonsuberized cells named phelloids may also be found in the phellem.

Except for a very few cork-producing species, the production of phellem cell layers may be as low as one or two cell layers per month. This rate can be accelerated under various environmental conditions. *See* CELL WALLS (PLANT).

Lenticels. Lenticels are loose-structured openings that develop usually beneath the stomata and that facilitate gas transport through the otherwise impermeable layers of phellem (**Fig. 3**).

The lenticels have a round or oblong shape in surface view. In their center, a mass of loose, thin-walled cells (complementary tissue) is formed. In a few cases these cells are released outward while in other cases they are held in place by bands of normal dense phellem cells (closing layers).

The lenticels develop prior to or concurrently with the initiation of the first phellogen and may

Fig. 3. A lenticel as seen in a cross section of the stem periderm of *Sambucus nigra*.

Fig. 4. Surface view of a bottle cork. A few annual rings and a perennial lenticel can be distinguished.

persist for many years. The lenticels of subsequent periderms are formed continuous with the lenticels of the first periderm. More lenticels may be formed in subsequent periderms, usually inside cracked periderm layers.

Cork. Cork is the most important commercial product of the periderm. Commercial cork is produced mainly by two species, *Quercus suber* and *Q. occidentalis*, both natives of the west Mediterranean region. This cork consists mostly of layers of phellem cells with perennial lenticels and is a product of a subsequent phellogen which is formed only after the first phellogen has been exfoliated (**Fig. 4**). The phellogen of the cork oaks may function for scores of years.

Wound periderm. A special type of periderm which is worth notice is the wound periderm. This is also a protective tissue that develops within injured plant organs beneath wound surfaces. Its formation is rapid and is accelerated under high temperature and high humidity conditions. The development of a similar periderm is also associated with any kind of normal plant abscission. Prior to abscission, a separation layer is formed between the abscising and the persisting tissues, and below this layer a phellogen is initiated. *See* ABSCISSION; BARK; SCLERENCHYMA.

Y. Waisel; H. Wilcox

Bibliography. G. B. Cooke, *Cork and the Cork Trees*, 1961; K. Esau, *Anatomy of Seed Plants*, 4th ed., 1990; A. Fahn, *Plant Anatomy*, 3d ed., 1982; N. Harris and K. J. Oparka (eds.), *Plant Cell Biology: A Practical Approach*, 1994.

Peridotite

A rock consisting of more than 90% of millimeter-to-centimeter-sized crystals of olivine, pyroxene, and hornblende, with more than 40% olivine. Other minerals are mainly plagioclase, chromite, and garnet. Much of the volume of the Earth's mantle probably is peridotite.

Mineralogy. Peridotites have various names depending upon the subordinate minerals (**Fig. 1**). Generally the aluminous mineral is mentioned as a modifier, as in feldspathic lherzolite. Other minerals, if significant, are similarly used as modifiers—hornblende peridotite, for example.

Structure and texture. Some peridotites are layered or are themselves layers; others are massive. Many layered peridotites occur near the base of bodies of stratified gabbroic complexes. Other layered peridotites occur isolated, but possibly once composed part of major gabbroic complexes. *See* GABBRO.

Both layered and massive peridotites can have any of three principal textures: (1) rather well-formed crystals of olivine separated by other minerals (**Fig. 2**); (2) equigranular crystals with straight grain boundaries intersecting at about 120°; and (3) long crystals with ragged curvilinear boundaries. The first texture probably reflects the original deposition of an olivine sediment (or cumulate) from magma. The second texture may result from slow cooling whereby recrystallization leads to a minimization of surface energy. The third texture probably results from internal deformation.

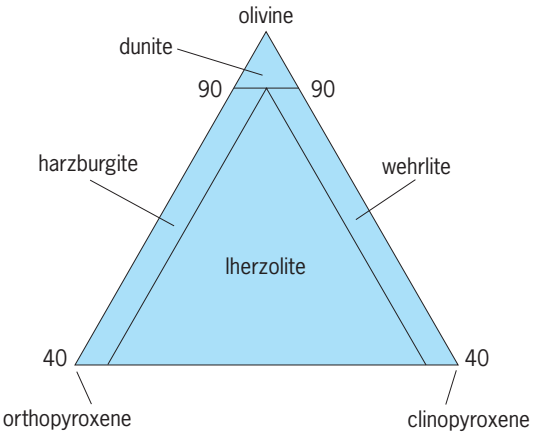

Fig. 1. Diagram of various peridotites which depend on principal mineral content. The numbers indicate 90 vol % olivine; and 40 vol % olivine and 60 vol % pyroxene (ortho or clino or both).

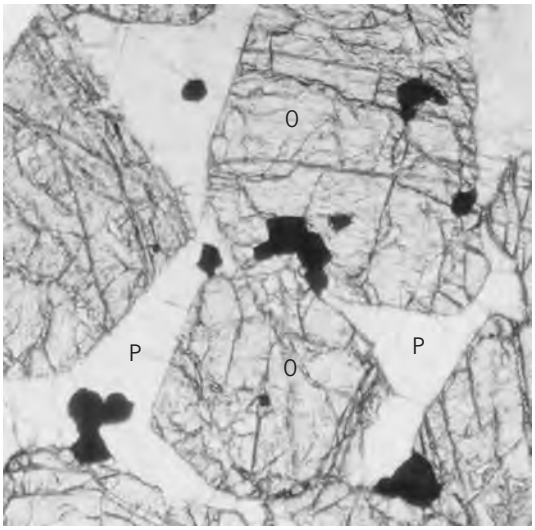

Fig. 2. Millimeter-sized crystals of olivine (O) with good crystal form are outlined by plagioclase (P) which fills the space between the accumulated crystals of olivine. Small crystals of black chromite occur in the olivine and plagioclase.

Physical properties. Peridotites are comparatively dense rocks (3.3–3.5 g/cm^3 or 1.9–2.0 oz/in.^{3}) with rather high velocities ($V_p \approx 5$ mi/s or 8 km/s) for body-wave propagation. The electrical conductivity is sensitive to oxidation state, lattice defects, and temperature. Remanent magnetization is variable in peridotites, being particularly high in altered varieties rich in small grains of magnetite.

Occurrence. Peridotites have three principal modes of occurrence corresponding approximately to their textures: (1) Peridotites with well-formed olivine crystals occur mainly as layers in gabbroic complexes; so-called Alpine peridotites generally have irregular crystals and occur as more or less serpentinized lenses bounded by faults in belts of folded mountains such as the Alps, the Pacific coast ranges, and the Appalachian piedmont. (2) Peridotite nodules in alkaline basalts and diamond pipes generally have equigranular textures, but some have irregular grains. (3) Peridotite also occurs on the walls of rifts in the deep sea floor and as hills on the sea floor, some of which reach the surface. Many Alpine peridotites occur in the ophiolite association: peridotite, gabbro, diabase sill-and-dike complex, pillow basalt, and red chert. Some peridotites rich in amphibole have a concentric layered structure and form parts of plutons called Alaskan-type zoned ultramafic complexes. Certain volcanic rocks, sometimes called komatiites, are sufficiently rich in olivine and pyroxene to be termed peridotite. Small pieces of peridotite have been found in lunar breccias.

Composition. Peridotites are rich in magnesium, reflecting the high proportions of magnesium-rich olivine. The compositions of peridotites from layered igneous complexes vary widely, reflecting the relative proportions of pyroxenes, chromite, plagioclase, and amphibole.

Special interest focuses on the compositions of nodules of peridotites found in certain basalts and diamond pipes (kimberlites) because they may provide clues to the composition of the Earth's mantle. There are serious difficulties in evaluating the analytical data, particularly for minor and trace elements because of the complex history which such rocks have undergone. At least two principal varieties of peridotites exist in the Earth's mantle: (1) lherzolites, which have relatively high proportions of basaltic ingredients (in the minerals garnet and clinopyroxene); and (2) harzburgites, which have relatively low proportions of basaltic ingredients (because garnet and clinopyroxene are lacking or minor; see **table**). *See* AMPHIBOLE; BASALT; GARNET; PYROXENE.

Origin. Layered peridotites are igneous sediments and form by mechanical accumulation of dense olivine crystals (Fig. 2). Peridotite nodules are pieces of mantle rock more or less modified by partial melting. Mantle lherzolites may be the principal source rock for basaltic magmas, whereas mantle harzburgites probably form both from the crystalline residue left after basaltic magma migrates out of lherzolite and from a crystalline accumulation of early solidification products of some basaltic magmas within the mantle. Alpine peridotites of the ophiolite association probably formed in the oceanic crust and uppermost mantle by transfer of partial melt from the mantle to the crystal, followed by tectonic emplacement of both crust and mantle along thrust faults in mountain belts. Peridotites associated with Alaskan-type ultramafic complexes probably formed in the root zones of volcanoes. *See* PLATE TECTONICS.

Composition of two principal varieties of peridotite*

Ingredients	A	B
SiO_2	47.52	43.70
TiO_2	<0.02	0.25
Al_2O_3	0.63	2.75
Cr_2O_3	0.27	0.28
Fe_2O_3	2.37	1.38
FeO	3.50	8.81
MnO	0.10	0.13
MgO	43.21	37.22
CaO	0.49	3.26
Na_2O	0.06	0.33
K_2O	0.02	0.14
H_2O^{+}†	1.67	1.94
H_2O^{-}‡	0.07	0.05
Total	99.93	100.24

*A is garnet harzburgite, sample number PHN 1569, Thaba Putsoa kimberlite, Lesotho, South Africa. B is garnet lherzolite, sample number PHN 1611, Thaba Putsoa kimberlite, Lesotho, South Africa.
†Water released above 110° C (230° F).
‡Water released at or below 110° C (230° F).

Use. Peridotite is an important rock economically. Where granites have intruded peridotite, asbestos and talc are common. Pure olivine rock (dunite) is quarried for use as refractory foundry sand and refractory bricks used in steelmaking. Serpentinized peridotite is locally quarried for ornamental stone. Tropical soils developed on peridotite are locally ores of nickel. The sulfides associated with peridotites are common ores of nickel and platinoid metals. The chromite bands commonly associated with peridotites are the world's major ores of chromium. *See* IGNEOUS ROCKS.

Alfred T. Anderson, Jr.

Bibliography. F. R. Boyd and H. O. A. Meyer (eds.), *The Mantle Sample: Inclusions in Kimberlites and Other Volcanics*, American Geophysical Union, 1979; A. R. McBirney, *Igneous Petrology*, 2d ed., 1992; L. R. Wager and A. Hall, *Igneous Petrology*, 1987; P. J. Wyllie (ed.), *Ultramafic and Related Rocks*, 1976, reprint 1979.

Period doubling

A scenario for the transition of a natural process from regular motion to chaos. Various natural processes develop in time in a way that depends upon prevailing environmental details. A quantity that specifies the particular state of the environment of a process is called a parameter, and is taken as a fixed value over the course of development of the process. It is a frequent natural occurrence for a process to have a regular and easily describable motion for some range of parameters, but to have complex, irregular, and difficult-to-describe motions for other ranges of

parameters. In the context of fluid flow, the latter circumstance is termed turbulence. In a more general context it is called chaos (which includes fluid turbulence but presages an underlying generality). *See* TURBULENT FLOW.

Sometimes, as the environmental parameters are varied, a process may systematically exhibit more irregular motions, turning over into chaotic motion beyond some parameter value. In analogy to the phenomenology of phase transitions, this circumstance is termed a transition to chaos. There are a variety of qualitatively different transitions to chaos, each termed a scenario. Some of them are now well understood.

Period doubling is one frequently encountered scenario leading to chaos for which a full theoretical account exists. Since it occurs in a wide variety of processes of significantly divergent physical characters (for example, fluid flow, chemical reactions, and electronic devices), it is sensible to consider it as a phenomenon in its own right. Indeed, this panphysical character is a motivation to pose the study of the nature of complex motions as a subject in itself, and hence justifies the embracing generalization from turbulence to chaos. The affirmation of this vantage point is afforded by the fact that period doubling, wherever it appears, appears in a way that is quantitatively universal. This means that accurate numerical relations that govern this scenario hold for every such physical process, be it a fluid or a Josephson junction device.

Qualitative behavior. In order to observe this scenario, it is sufficient that all but one parameter is held fixed. Over some range of this varied parameter (it shall be defined to increase over the range of investigation) the motion is observed to be periodic. Above a certain value of the parameter the motion grows more complicated (a bifurcation has occurred): after the amount of time T for which the motion exactly repeated itself just prior to the bifurcation, the motion now slightly fails to do so, exactly repeating, however, after another T seconds. That is, the period has doubled from T to $2T$. As the parameter is further increased, the error to repeat after the first half of the new period systematically increases. A still further increase of parameter produces another bifurcation resulting in a new doubling of the period: the motion slightly fails to repeat after two roughly periodic cycles, exactly doing so after four. As the parameter is further increased, there are successive period-doubling bifurcations, more and more closely spaced in parameter value until at a critical value the doubling has occurred an infinite number of times, so that the motion is now no longer periodic and hence of a more complex character than had yet been encountered. Unpredictably complex motions occur for values of the parameter above its critical value, although ranges of parameter still exist for which the system exhibits new periodic motions. Indeed, any period-doubling system exhibits the same sequence of truly chaotic motion and interspersed periodicities as its parameter increases. Thus there is a strong degree of qualitatively universal behavior for all systems experiencing this scenario.

Quantitative behavior. However, there is also a precise quantitative universality. That is, without knowing the system (or its equations) essentially all measurable quantities can be predicted: By looking at the data alone, it would not be possible to guess the physical system responsible for that data. In order to describe these facts, some formal definitions are required. Denoting the parameter by p, let p_n be the value of p at the nth bifurcation, so that, for large n, T_n, the new periodicity is asymptotically given by Eq. (1) for an appropriate value of T_0. The theory of this scenario then predicts that Eq. (2) is satisfied, where δ is a universal mathematical constant of nature with value given by Eq. (3). The content

$$T_n \sim 2^n T_0 \tag{1}$$

$$\frac{p_{n+1} - p_n}{p_{n+2} - p_{n+1}} \sim \delta \tag{2}$$

$$\delta = 4.6692016\ldots \tag{3}$$

of Eq. (2) is that the p_n's accumulate to the critical value p_∞ geometrically for large n at rate δ, as given by Eq. (4).

$$p_\infty - p_n \sim \delta^{-n} \tag{4}$$

Furthermore, the errors to reproduce after T_n seconds when the period is $2T_n$ at each value of time over the entire motion are related to the errors at the same time when the periodicity was T_n in a precise way. To specify this, define $x_n(t)$ as the value of any quantity in the system when the periodicity is T_n. The error to reproduce after half the period is then given by Eq. (5). The theory of period doubling then determines that Eq. (6) is satisfied, where σ is a universal function which in first approximation has two distinct values (and their negatives) given by α^{-1} and α^{-2} where α is given by Eq. (7).

$$d_n(t) = x_n(t) - x_n\left(t + \frac{T_n}{2}\right) \tag{5}$$

$$d_{n+1}(t) \sim \sigma\left(\frac{t}{T_{n+1}}\right) d_n(t) \tag{6}$$

$$\alpha = 2.50290787\ldots \tag{7}$$

The meaning of σ is that the quantitatively universal aspects of this scenario are scaling relationships for the motion: in first approximation, after the overall variation of $x(t)$ is measured to set a "yardstick" for the given system, all the details of $x(t)$ can be constructed from Eq. (6) by successive scalings. At p_∞, the motion $x_\infty(t)$ is self-similar at all scales. That is, the "noise" in x_∞ is not statistical in nature, but rather the complex outcome of repeated elaboration of details at all scales. It is a technical exercise to compute the power spectrum of $x_n(t)$ once Eq. (6) is known. *See* FRACTALS.

Thus the manner in which the parameter axis is laid out, given by Eq. (4), and the manner in which the system modifies its response at each parameter value, given by Eq. (6), are precisely determined apart from some overall scales in a way completely independent of the nature of the given system. That is, δ and $\sigma(t)$ are measurable quantities (and have been measured in a variety of real physical systems) and basically exhaust all the information that could be measured, or could be computed from the specific equations governing each specific system investigated. Thus, reminiscent of thermodynamics, questions can be posed and answered in a general manner that bypasses the specific mechanisms governing any particular system.

Universality of the scenario. The results discussed above hold over some parameter range for almost all nonlinear dissipative systems with enough "parts" (degrees of freedom), where enough means three or more (or even infinitely more). For some such systems the relevant range of parameters is sufficiently large to make period doubling a prominent part of the system's repertoire. For others it may be unmeasurably small, and so of no physical consequence. Indeed there do exist alternative scenarios, some of which satisfy theories highly reminiscent of that of period doubling, that can prominently mark a system's transition to chaos.

All of the features described are the consequences of a rigorous renormalization group theory, so that this scenario is similar in spirit to second-order phase transitions. In particular, δ is a critical exponent, α is a rescaling parameter, and the function σ is constructed from the fixed point of an appropriate transformation. The other scenarios akin to period doubling differ in that the renormalization group transformations differ. *See* PHASE TRANSITIONS; RENORMALIZATION. Mitchell J. Feigenbaum

Bibliography. J. Argyris, G. Faust, and M. Haase, *An Exploration of Chaos*, 1994; R. L. Devaney, *A First Course in Dynamical Systems: Theory and Experiment*, 1992; J. Gleick, *Chaos: Making a New Science*, 1992; E. Lorenz, *The Essence of Chaos*, 1993.

Periodic motion

Any motion that repeats itself identically at regular intervals. If $x(t)$ represents the displacement of any coordinate of the system at time t, a periodic motion has the property defined by Eq. (1) for every value of

$$x(t+T) = x(t) \qquad (1)$$

the variable time t. The fixed time interval T between repetitions, or the duration of a cycle, is known as the period of the motion. Frequency is the number of repeating cycles per unit time, and is numerically equal to the reciprocal of the period T.

The motion of the escapement mechanism of a watch, the motion of the Earth about the Sun, and the more complicated motion of the crankshaft, piston rods, and pistons in an engine running at uniform speed are all examples of periodic motion.

The vibration of a piano string after it is struck is a damped periodic motion, not strictly periodic according to the definition. Although the motion very nearly repeats itself, and with a fixed repetition time, each successive cycle has a slightly smaller amplitude. *See* DAMPING.

Any periodic motion can be expressed as a Fourier series—a sum of sine and cosine terms whose frequencies of integral multiples of the frequency f of the periodic motion. Thus Eq. (2) holds, where the

$$x(t) = A_0 + \Sigma_n A_n \cos(2\pi nft) + \Sigma_n B_n \sin(2\pi nft) \qquad (2)$$

A's and B's are constant coefficients, and the sums may be taken over all positive integer values of n. For the special case in which the coefficients all vanish for $n > 1$ *see* HARMONIC MOTION; FOURIER SERIES AND TRANSFORMS.

Many systems with more than one degree of freedom, whose motion is not simply periodic, are multiply periodic. The motion may be resolved into parts (for example, horizontal and vertical components, radial and tangential components), each of which is periodic, but with periods that are not commensurate. One example is the vibration of a bell, whose overtone frequencies are not simply related to the fundamental frequency. The motion of the solar system is multiply periodic because it never exactly repeats itself, even though each planet moves periodically. *See* VIBRATION; WAVE MOTION. Joseph M. Keller

Periodic table

A list of chemical elements arranged along horizontal rows in increasing atomic number. It is organized such that the vertical columns consist of elements with remarkably similar properties. The first column, known as the alkali metals (albeit with hydrogen, a nonmetal on top), contains elements with just one outer (valence) electron. The last column has completely filled valence orbitals leading to chemically inert elements called the noble gases. The position of elements in the periodic table provides a powerful method of classifying not only the physical properties of elements but also their expected properties in molecules and solids. *See* ALKALI METALS; ATOMIC NUMBER; ELECTRON CONFIGURATION; INERT GASES; VALENCE.

The periodic table dates back to around 1870 when the Russian chemist D. Mendeleev used the similarities in chemical reactivity attributed to different elements to group them according to increasing atomic mass. However, this left blank spaces that Mendeleev boldly predicted would be filled by elements that were then undiscovered. Not only did other scientists discover missing elements including scandium, gallium, and germanium, but remarkably these elements possessed properties similar to

Mendeleev's predictions. Although the layout of the periodic table has changed over time with the addition of new elements, the essential information remains comparable to the original periodic table. *See* ATOMIC MASS.

Groups. The modern periodic table is divided into 18 columns called groups or families (see **illustration**). Elements in each family tend to have similar properties. In column 1, each alkali metal is soft, relatively low-melting, and highly reactive toward air and water. Column 2 contains the alkaline earth metals, which have higher melting points and are less reactive. Columns 3–12 are filled by the transition metals, which are shiny and good conductors of both heat and electricity. Columns 13–18 are often discussed along with columns 1 and 2, and collectively they are known as the main group or representative elements. Column 15, headed by nitrogen, is known as the pnicogens; column 16, beginning with oxygen, as the chalcogens; column 17, starting with fluorine, as the halogens; and column 18, starting with helium, as the noble gases. *See* ALKALINE-EARTH METALS; TRANSITION ELEMENTS.

Periods. The horizontal rows of the periodic table are called periods. Atomic mass generally increases from left to right across a period, while atomic size generally decreases. The decrease in size is due to incomplete screening of the positive nuclear charge by the valence electrons, which causes the outer electron shells to contract. Other properties follow periodic trends, including the ionization potential (the energy needed to remove an electron), electron affinity (the energy released on accepting an electron), and electronegativity (the ability of an atom in a compound to attract electron density). *See* ATOMIC STRUCTURE AND SPECTRA; ELECTRON AFFINITY; ELECTRONEGATIVITY; IONIZATION POTENTIAL.

After element 57 (lanthanum) comes a series of 14 metallic elements numbered 58–71 with very closely related properties. These originally were named the rare earths since they are all nearly the same size, have similar chemical reactivity, and are difficult to separate. However, they are not rare and are now more appropriately called the lanthanides. Technically the lanthanides should be placed between elements 57 (lanthanum) and 72 (hafnium). Since this would nearly double the width of the periodic table, they are usually placed below all the other elements. Keeping in mind that size decreases left to right across a period, hafnium (72) is much smaller than its neighbor lanthanum (57). In fact, hafnium is essentially the same size as the element above it, zirconium (40). With comparable chemical reactivity and size, zirconium and hafnium are difficult to separate. This also suggests that the second and third rows of the transition metals will possess many common chemical features, as they do. The decrease in size due to the 14 elements between

KEY

Atomic number → 79
Symbol → Au
Atomic weight → 196.97

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 H 1.0079																	2 He 4.0026
3 Li 6.941	4 Be 9.012											5 B 10.811	6 C 12.011	7 N 14.007	8 O 15.999	9 F 18.998	10 Ne 20.180
11 Na 22.990	12 Mg 24.305											13 Al 26.982	14 Si 28.086	15 P 30.974	16 S 32.065	17 Cl 35.453	18 Ar 39.948
19 K 39.098	20 Ca 40.078	21 Sc 44.956	22 Ti 47.867	23 V 50.942	24 Cr 51.996	25 Mn 54.938	26 Fe 55.845	27 Co 58.933	28 Ni 58.693	29 Cu 63.546	30 Zn 65.39	31 Ga 69.723	32 Ge 72.64	33 As 74.922	34 Se 78.96	35 Br 79.904	36 Kr 83.80
37 Rb 85.468	38 Sr 87.62	39 Y 88.906	40 Zr 91.224	41 Nb 92.906	42 Mo 95.94	43 Tc (98)	44 Ru 101.07	45 Rh 102.91	46 Pd 106.42	47 Ag 107.87	48 Cd 112.41	49 In 114.82	50 Sn 118.71	51 Sb 121.76	52 Te 127.60	53 I 126.90	54 Xe 131.29
55 Cs 132.91	56 Ba 137.33	57–71	72 Hf 178.49	73 Ta 180.95	74 W 183.84	75 Re 186.21	76 Os 190.23	77 Ir 192.22	78 Pt 195.08	79 Au 196.97	80 Hg 200.59	81 Tl 204.38	82 Pb 207.2	83 Bi 208.98	84 Po (209)	85 At (210)	86 Rn (222)
87 Fr (223)	88 Ra (226)	89–103	104 Rf (261)	105 Db (262)	106 Sg (266)	107 Bh (264)	108 Hs (277)	109 Mt (268)	110 Ds (271)	111 Rg (272)	112 Uub (285)	113	114	115	116	117	118

Lanthanides	57 La 138.91	58 Ce 140.12	59 Pr 140.91	60 Nd 144.24	61 Pm (145)	62 Sm 150.36	63 Eu 151.96	64 Gd 157.25	65 Tb 158.92	66 Dy 162.50	67 Ho 164.93	68 Er 167.26	69 Tm 168.93	70 Yb 173.04	71 Lu 174.97
Actinides	89 Ac (227)	90 Th 232.04	91 Pa 231.04	92 U 238.03	93 Np (237)	94 Pu (244)	95 Am (243)	96 Cm (247)	97 Bk (247)	98 Cf (251)	99 Es (252)	100 Fm (257)	101 Md (258)	102 No (259)	103 Lr (262)

Periodic table of the elements. Atomic weights are those of the most commonly available long-lived isotopes on the 1999 IUPAC Atomic Weights of the Elements. A value in parentheses denotes the mass number of the longest-lived isotope.

lanthanum and hafnium is called the lanthanide contraction. Below the lanthanides are 14 more metallic elements (90–103) called the actinides. *See* ACTINIDE ELEMENTS; LANTHANIDE CONTRACTION; RARE-EARTH ELEMENTS.

Elements. Each box in the periodic table contains a one- or two-letter symbol representing a different element such as C for carbon (6) or Sg for Seaborgium (106) [see illustration]. The number in the upper left corner is the atomic number indicating how many protons are in the atom's nucleus. The atomic mass generally appears below the symbol indicating the average mass observed for that element. For example, most carbon (99%) contains 6 protons and 6 neutrons, leading to a mass of 12. However, since about 1% of carbon has an extra neutron, the average mass of carbon as given in the periodic table is 12.011. If an element has no stable isotopes, then the mass of the longest-lived isotope is given in parentheses. More complex periodic tables often include information on density, melting points, and boiling points. Separate tables are available, indicating crystal structures, magnetic properties, radioactive decay patterns, and other properties.

Of the current elements in the periodic table, 111 have been officially ratified by the International Union of Pure and Applied Chemistry (IUPAC). After a newly discovered element has been independently verified, the original discoverer (often a team) earns the right to propose a name to IUPAC. The elements beyond 92 (uranium) do not occur naturally and are produced using nuclear reactions. Elements beyond 100 are not particularly useful, since they generally undergo rapid nuclear decay by emitting radiation. *See* TRANSURANIUM ELEMENTS.

Other properties. Many periodic tables include a stair-step line separating metals from the metalloids and nonmetals. Most elements are metals and generally have physical properties that include luster (high reflectivity), good conductivity for both electricity and heat, high density, usually high melting points, ductility (the ability to be drawn into a wire), and malleability (the ability to be hammered into thin sheets). The chemical properties of most metals include corrosivity such as iron rusting and silver tarnishing, as well as the ability to give up electrons. Nonmetals are found to the right of the metals and their characteristics are the inverse. This means most have no luster (appearing dull), are poor conductors of electricity and heat, have low density, low melting points, and are brittle. Chemically, nonmetals like to gain electrons and often react with metals to produce salts. For example, combining an alkali metal with one valence electron with a halogen that needs one electron to complete its valence shell produces an alkali halide salt such as sodium chloride or common table salt. Metalloids straddle the stair-step line and often have properties in between metals and nonmetals. With intermediate conductivities, elements such as silicon form important semiconductors used in computer chips and solar cells. *See* METAL; METALLOID; NONMETAL; SEMICONDUCTOR. Richard B. Kaner

Bibliography. J. Emsley, *The Elements*, Clarendon Press, 3d ed., 1998; N. N. Greenwood and A. Earnshaw, *Chemistry of the Elements*, Butterworth-Heinemann, 2d ed., 1997; J. E. Huheey, E. A. Keiter, and R. L. Keiter, *Inorganic Chemistry*, 5th ed., 2002; G. Rayner-Canham and T. Overton, *Descriptive Inorganic Chemistry*, W. H. Freeman, 4th ed., 2006; E. Scerri, *The Periodic Table: Its Story and Significance*, Oxford University Press, 2006.

Periodontal disease

An inflammatory lesion caused by bacteria affecting the tissues housing the roots of the teeth. The disease, sometimes called pyorrhea, increases in prevalence and severity with increasing age, and is a principal cause of tooth loss in adult humans throughout the world. When only the gum tissue or gingiva is affected the disease is called gingivitis, but when the process extends into the deeper structures it is known as periodontitis. The diseased tissues appear abnormally red and slightly swollen, and they tend to bleed, sometimes profusely, when the teeth are brushed. In some cases the gums may become thickened and scarred, and they may recede, exposing the root surfaces. As the disease advances, the attachment of the gum to the tooth is lost, creating a periodontal pocket; a large portion of the gum tissue is destroyed; and the bone surrounding the roots is resorbed (see **illus.**). The teeth become loose, abscesses form, and extraction is required.

Etiology. Both gingivitis and periodontitis are caused by bacteria that form plaques on the surfaces of the teeth at the gingival sulcus or pocket. These plaques may contain 250 or more separate microbial species. Plaques of any microbial composition can cause gingivitis, but specific bacteria appear to be necessary for induction of periodontitis. Among the bacteria involved in periodontitis are various species of *Porphyromonas*, *Bacteroides*, *Actinobacillus*, *Eikenella*, *Fusobacterium*, *Wolinella*, and other less well-characterized species. Spirochetes are present in active lesions, but their role remains unclear. The bacteria extend apically along the interface between the tooth root and the gingival tissue and cause periodontal pockets to form.

Pathogenesis. The principal features of the pathogenesis of periodontitis have been described. The lesions begin as an acute inflammatory response, followed by a dense accumulation of lymphoid cells. There is a net loss of collagen in the areas nearest the junctional epithelium and periodontal pocket, with scarring and fibrosis of the connective tissues at more distant sites. The junctional epithelium is converted into an ulcerated pocket epithelium, the alveolar bone housing the tooth roots resorbed, and the periodontal ligament is destroyed. Products released by infiltrating leukocytes, including prostaglandins, interleukins and collagenase, and other hydrolytic enzymes, are involved in tissue destruction.

Bacterial colonization and extension activate several host defense mechanisms. The most effective of these is the accumulation of functional neutrophilic granulocytes between the surface of the plaque and the gingival tissue. These cells tend to counter and limit microbial extension. The bacteria appear to invade the periodontal connective tissues, where they induce immunopathologic and other destructive inflammatory reactions in the host, and these lead, in major part, to the observed tissue destruction. Periodontal destruction is episodic, with periods of exacerbation characterized by highly acute inflammation, followed by periods of quiescence.

Predisposing factors. Although bacteria are essential for induction of the disease, predisposing factors are also important, though their elucidation is not complete. Individuals who manifest functionally abnormal neutrophilic granulocytes or monocytes are unusually susceptible to the severe early-onset forms of periodontitis. The leukocyte abnormality appears to be genetically transmitted. The early-onset forms have been designated as prepubertal, juvenile, and rapidly progressive periodontitis; adult periodontitis has a later onset and does not seem to be related to leukocyte abnormalities. Some persons with acquired immune deficiency syndrome (AIDS) manifest a highly destructive, unique form of periodontitis. Other predisposing conditions include unusually stressful situations, and periods of hormone imbalance occurring at puberty, during pregnancy, and in some women taking birth control drugs.

Prevention and treatment. Good daily oral hygiene practices, including vigorous brushing of the exposed surfaces of the teeth and use of dental floss, interproximal brushes, and other devices to clean between the teeth, constitute the most effective measures to prevent periodontal disease. Basic ingredients of treatment of existing disease include bringing the infection under control and establishing conditions which preclude reinfection. All of the microbial deposits must be removed from the crown and root surfaces. In individuals with severe forms of periodontitis, these procedures may be supplemented by use of antibiotics either systematically or directly into the pocket. These procedures usually lead to reduction of the inflammation and to some shrinkage in the gums, but the periodontal pockets remain. Based on the traditional view that treated pockets may become reinfected and the disease may continue to spread, surgical treatment may be performed with the aim of reducing pocket depth and restoring normal tissue contours. Alternatively, regenerative procedures use various grafting materials, including freeze-dried decalcified bone or bone substitutes and guided tissue regeneration. To perform guided tissue regeneration flaps are opened in the gingival tissue and the root surfaces are thoroughly cleaned; a porous membrane is placed around the tooth covering the bone defect with or without placing grafts, and flaps covering the membranes are sutured into place. The membrane permits the wound site to become populated with cells having the capacity to generate new bone, cementum, and periodontal attachment. *See* TOOTH DISORDERS.

Roy C. Page

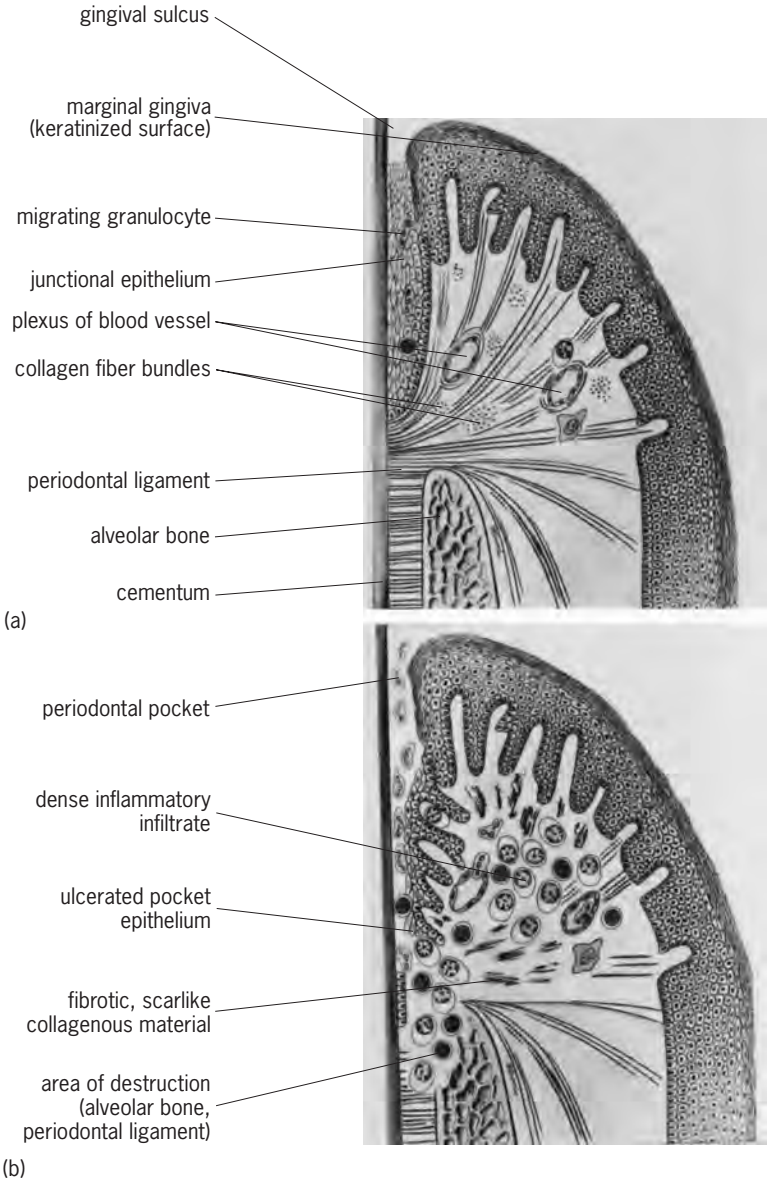

Structural features of (*a*) the normal periodontium and (*b*) the advanced stage of periodontal disease. (*After S. Schluger, R. C. Page, and R. Yuodalis, Periodontal Disease: Basic Phenomena, Clinical Management and Restorative Interrelationships, Lea and Febiger, 1990*)

Bibliography. R. C. Page, Periodontal therapy: Prospects for the future, *J. Periodontal*, 64:744–753, 1993; S. Schluger et al., *Periodontal Disease*, 2d ed., 1990.

Perischoechinoidea

A subclass of Echinoidea lacking stability in the number of columns of plates that make up the ambulacra and interambulacra. The ambulacral columns vary from 2 to 20, the interambulacral from 1 to 14. There

are two orders, Bothriocidaroida and Echinocystitoida. *See* ECHINODERMATA; ECHINOIDEA.

Howard B. Fell

Periscope

An optical instrument that permits viewing along a displaced or deflected axis, providing an observer with the view from a position which may be inaccessible or dangerous. Periscopes range in complexity from the simple unit-power tank periscope to the complex multielement submarine periscope.

Tank periscope. This device, intended to protect the user from bullets, employs a pair of plane, parallel, reflecting surfaces (either mirrors or prisms), so arranged in a mount that the path of light through the instrument forms a crude letter Z (**Fig. 1***a*). If powers greater than unity are desired or if the periscope is to be used for sighting, a terrestrial telescope can be added to the periscope, either as a simple, internally contained system (Fig. 1*b*) or entirely in front of or behind the periscope itself, as desired. The reflecting elements of the system which are responsible for deflecting the optical axis are independent of the refracting (telescope) elements which provide the optical power. *See* TELESCOPE.

It is also possible to arrange the two mirrors of a periscope at right angles to each other, in which case the observer views an inverted image while facing away from the direction from which light enters the instrument. By adding an inverting (astronomical type) telescope to this system, the image is reinverted for the observer.

Periscopes of this type cannot be used for scanning the horizon by rotating the upper mirror because of the image rotation which accompanies such movement. In the panoramic sight (**Fig. 2**), this difficulty is overcome by providing the system with a dove prism which rotates at half the angular speed of the right-angle prism through the action of a differential gear linkage. The combined inversions of the dove prism and the amici prism at the bottom

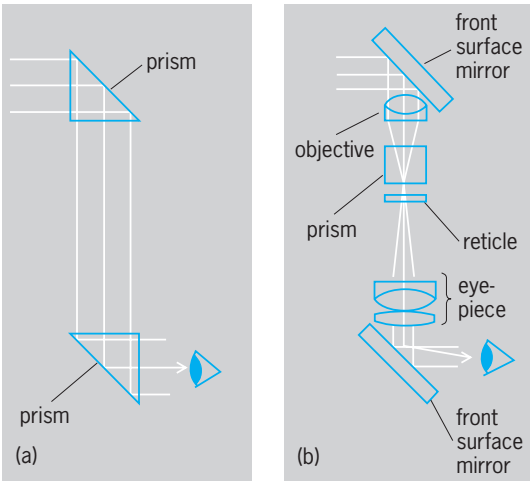

Fig. 1. Tank periscope. (*a*) Simple, with parallel reflecting surfaces. (*b*) With terrestrial telescope.

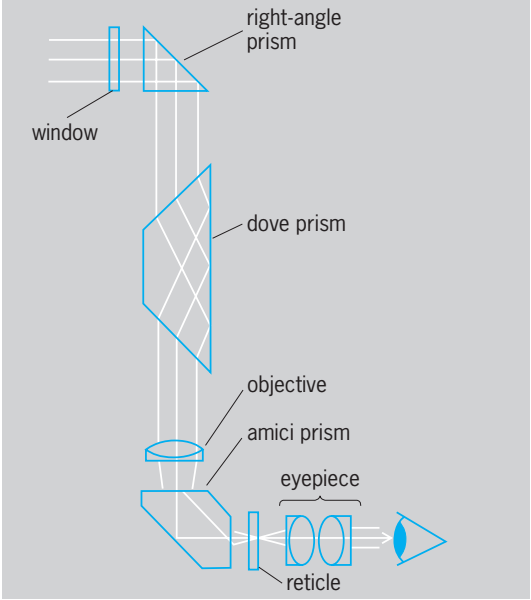

Fig. 2. Panoramic sight, with erect image.

completely compensate for the inversions of the telescope system, while the relative motions of the right-angle prism and the dove prism maintain the image erect during scanning. *See* MIRROR OPTICS; OPTICAL PRISM.

Submarine periscope. In this device, it is necessary to employ a telescope system having a wide field of view and uniform illumination across a field which can be fitted into a long, narrow tube whose length-to-diameter ratio may be 50 or greater. This is achieved by utilizing a plurality of lenses so spaced along the length of the tube as to cause the incoming principal rays from the edge of the field to be deviated from side to side within the tube. In general, the greater the number of lenses, the wider the field of view. One example of the periscopic relay train is shown in **Fig. 3***a* and employs six lenses with three inversions. The typical submarine periscope (Fig. 3*b*) may be considered to be a pair of telescopes facing each other, with such a relay train between them. The usual magnification of submarine periscopes is 6, although some U.S. Navy periscopes have dual magnifications of 6 and 1.5, the latter being achieved by inserting an inverted galilean telescope into the optical path before the top objective. *See* LENS (OPTICS); MAGNIFICATION.

The submarine periscope can be provided with a built-in rangefinder for fire-control purposes. A conventional coincidence or split-field type of rangefinder may be attached either vertically or horizontally to the upper end of the periscope, the objective of which receives an image from each of the entrance windows of the rangefinder. *See* RANGEFINDER (OPTICS); SUBMARINE.

Other types. Various modifications of the basic optical systems described here are employed as viewing periscopes in military aircraft and as viewing devices in particle accelerators and nuclear reactors. The cystoscope and endoscope are slender, sometimes

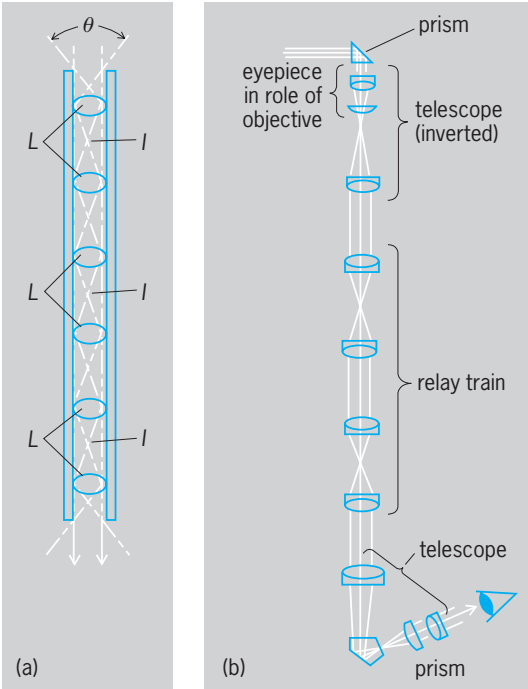

Fig. 3. Periscope relay train. (*a*) Showing lenses *L*, inversions *I*, and angle of view θ. (*b*) Between a pair of facing telescopes in a submarine periscope.

mechanically flexible periscopes used for visual examination and photography of body cavities inaccessible to direct observation; an entirely different basis for the design of such instruments is in the use of bundles of optical fibers. *See* OPTICAL FIBERS.

Edward K. Kaprelian

Bibliography. Optical Society of America, *Handbook of Optics*, 2d ed., 1995; A. G. Smith, *Easy to Make Periscope*, 1990; B. H. Walker, *Optical Engineering Fundamentals*, 1995.

Perissodactyla

An order of herbivorous, odd-toed, hoofed mammals, including the living horses, zebras, asses, tapirs, rhinoceroses, and their extinct relatives. They are defined by a number of unique specializations, but the most diagnostic feature is their feet. Most perissodactyls have either one or three toes on each foot, and the axis of symmetry of the foot runs through the middle digit. The woodchucklike hyraxes, or conies, are apparently closely related to perissodactyls, although there is still some question about the relationships.

The perissodactyls (other than hyraxes) are divided into three groups: the Hippomorpha (horses and their extinct relatives); the Titanotheriomorpha (the extinct brontotheres); and the Moropomorpha (tapirs, rhinoceroses, and their extinct relatives). *See* HORSE PRODUCTION; RHINOCEROS; TAPIR.

Origins. Perissodactyls were once thought to have evolved in Central America from the phenacodontids, an extinct group of archaic hoofed mammals placed in the invalid taxon Condylarthra. However, in 1989 a specimen recovered from deposits in China about 57 million years old was described and named *Radinskya*. This specimen shows that perissodactyls originated in Asia some time before 57 million years ago (Ma), and were unrelated to North American phenacodonts. *Radinskya* is very similar to the earliest relatives of the tethytheres (elephants, manatees, and their kin). This agrees with other evidence that perissodactyls are more closely related to tethytheres than they are to any other group of mammals.

By 55 Ma, the major groups of perissodactyls had differentiated, and migrated to Europe and North America. Before 34 Ma, the brontotheres and the archaic tapirs were the largest and most abundant hoofed mammals in Eurasia and North America. After these groups became extinct, horses and rhinoceroses were the most common perissodactyls, with a great diversity of species and body forms. Both groups were decimated during another mass extinction about 5 Ma, and today only five species of rhinoceros, four species of tapir, and a few species of horses, zebras, and asses cling to survival in the wild. The niches of large hoofed herbivores have been taken over by the ruminant artiodactyls, such as cattle, antelopes, deer, and their relatives.

Horses. From their Asian origin, the hippomorphs spread all over the northern continents. North America became the center of evolution of true horses, which occasionally migrated to other continents. In Europe, the horselike palaeotheres substituted for true horses. From *Protorohippus* (once called *Hyracotherium* or *Eohippus*), a dog-sized ancestor with four toes on the front feet that lived 55–50 Ma, horses evolved into many different lineages living side by side. The three-toed horses *Mesohippus* and *Miohippus* (from beds dated about 30–37 Ma) were once believed to be sequential segments on the unbranched trunk of the horse evolutionary tree. However, they coexisted for millions of years, with five different species of the two genera living at the same time and place. From *Miohippus*-like ancestors, horses diversified into many different ecological niches, with some species adapted for eating soft, leafy vegetation, and others with high-crowned teeth for eating tough, gritty grasses. About 15 Ma, there were as many as 12 different lineages of three-toed horses in North America, each with slightly different ecological specializations; and they were analogous to the diversity of modern antelopes in East Africa. On two different occasions (*Pliohippus* and *Dinohippus*), three-toed horses evolved into lineages with a single toe on each foot. About 5 Ma, most of these three-toed and one-toed horse lineages became extinct, leaving only *Dinohippus* to evolve into modern *Equus*.

On a few occasions in the last 20 million years, true horses migrated back from North America to Eurasia, and eventually to Africa (evolving into zebras and wild asses) and South America. At the end of the last ice age (about 10,000 years ago), horses became extinct in the New World. However, they were reintroduced to their ancestral homeland by Columbus in 1493. Wild horses that have escaped

from domesticated stock are known as mustangs.

Most extinct horses were browsers and ate soft, leafy vegetation, but all living horses are grazers, using their sharp incisors and mobile lips to crop low-growing grasses. The only common wild horse, the plains zebra, lives in large herds (up to 100 individuals) and migrates over large areas of grasslands in search of food. However, desert-dwelling asses and Grevy's zebra live in small herds, with a stallion guarding a small harem of mares. Most species of wild horses, including the Grevy's and mountain zebras, all species of onagers and asses, and Przewalski's horse (an ancestor of domesticated horses), are nearly extinct in the wild.

Brontotheres. These beasts began as pig-sized, hornless animals about 53 Ma, and quickly evolved into multiple lineages of cow-sized animals with long skulls and no horns. Between 37 and 34 Ma, their evolution culminated with huge, elephant-sized beasts bearing paired blunt horns on their noses. The extinction of brontotheres about 34 Ma was due to a global climatic change (triggered by the first Antarctic glaciers) that eliminated most of the soft, leafy vegetation on which they fed. *See* EXTINCTION (BIOLOGY).

Moropomorphs. The earliest moropomorphs, such as *Homogalax*, from strata about 55 million years old, are virtually indistiguishable from the earliest horses. From this unspecialized ancestry, a variety of archaic tapirlike animals diverged. Most retained the simple leaf-cutting teeth characteristic of tapirs and, like brontotheres, died out about Ma when their forest habitats shrank. Only the modern tapirs, with their distinctive long proboscis, survive in the jungles of Central and South America (three species), and southeast Asia (one species). All are stocky, piglike beasts with short stout legs, oval hooves, and a short tail. They have no natural defenses against large predators (such as jaguars or tigers), so they are expert at fleeing through dense brush and swimming to make their escape.

The horselike clawed chalicotheres are closely related to some of these archaic tapirs. They apparently used their claws to haul down limbs and branches to eat leaves (much as ground sloths did), rather than for digging. *Chalicotherium* had such long forelimbs and short hindlimbs that it apparently knuckle-walked like a gorilla, with its claws curled inward.

Rhinoceroses have been highly diverse and successful throughout the past 50 million years. They have occupied nearly every niche available to a large herbivore, from dog-sized running animals, to several hippolike forms, to the largest land mammal that ever lived—the 18-ft-tall (6-m), 44,000-lb (20,000-kg) *Paraceratherium* (once called *Baluchitherium* or *Indricotherium*). Most rhinoceroses were hornless. Rhinos with horns first appeared about 28 Ma; two different lineages independently evolved paired horns on the tip of the nose. Other groups evolved a single nasal horn, or huge horn on the forehead, or a pair of horns in tandem on nose and forehead. Unlike the horns of cattle, sheep, and goats, rhino horns are made of cemented hair fibers, and have no bony core.

Between 20 and 5 Ma, rhinos diversified into several browsing (leaf-eating) lineages, and hippolike grazing lineages, and browser-grazer pairs of rhinos were found all over the grasslands of Eurasia, Africa, and North America. The mass extinction event that occurred about 5 Ma wiped out North American rhinos and decimated most of the archaic rhino lineages in the Old World. During the ice ages, woolly rhinos and their relatives were common all over Eurasia. Their only surviving descendant is the endangered Sumatran rhinoceros. Only a few hundred individuals still live in the mountainous jungles of Sumatra.

Four other species of rhino survive in Asia and Africa, but all are on the brink of extinction because of heavy poaching for their horns. The Javan rhinoceros is rarely seen in the dense jungles of Java and Vietnam; fewer than 50 individuals may still be living in the wild. The Indian rhinoceros and the African rhinos (the browsing black rhinoceros and the grazing white rhinoceros) inhabit grasslands, although they prefer to hide in dense vegetation. Rhinos have very poor eyesight but excellent senses of smell and hearing, so they can detect danger long before others can see them. Most live in small family groups of females with their calves; however, bull rhinos tend to be solitary. *See* MAMMALIA.

Donald R. Prothero

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988; R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999; D. R. Prothero, The rise and fall of the American rhino, *Nat. Hist.*, 96:26–33, 1987; D. R. Prothero and R. M. Schoch (eds.), *The Evolution of Perissodactyls*, 1989; R. J. G. Savage and M. R. Long, *Mammal Evolution: An Illustrated Guide*, 1986.

Peritoneum

The membranous lining of the coelomic, especially the abdominal, cavity, composed mainly of flattened epithelial cells that produce a small amount of watery, or serous, fluid. In the embryo the coelomic wall is lined by this membrane, which continues over the developing viscera so that they are suspended and supported by the reflected peritoneum, principally from the dorsal body wall, but also from the ventral body wall in the region of the liver.

As the organs develop, enlarge, and assume their adult form and arrangement, the supporting peritoneum becomes modified, some parts being lost

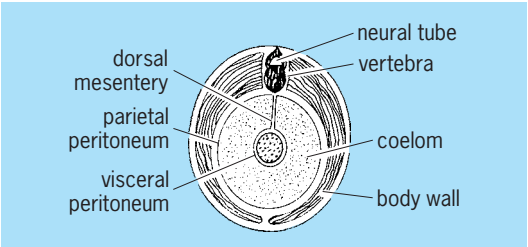

Diagrammatic cross section of a vertebrate, showing the relations of the layers of peritoneum and the coelom. (*After H. M. Smith, Evolution of Chordate Structure, Holt, Rinehart and Winston, 1960*)

and others becoming thickened or twisted, or otherwise adapting to normal growth (see **illus.**). After development is completed, those portions which line the interior of the body wall are called parietal peritoneum, and the supporting sheets are known collectively as mesenteries, many areas of which have received specific names. The remaining peritoneum, which covers most of the organs, is called visceral peritoneum and forms the outer layer, or serosa, of the walls of most of the gastrointestinal tract. The fluid-filled space between the serosa and the parietal peritoneum is the remnant of the coelom, or body cavity. *See* EPITHELIUM; FETAL MEMBRANE. Thomas S. Parsons

Peritonitis

Inflammation of the peritoneum, the serous membrane which lines the abdominal cavity and surrounds most of the abdominal organs. The condition may be caused by infectious organisms or foreign substances introduced into the abdominal cavity. The small amount of serous fluid normally present as a lubricant acts as an excellent culture medium for bacterial growth and also as a means of spreading invading materials. The source of such substances or organisms is commonly a gastrointestinal inflammation, especially if perforation has occurred. Appendicitis, peptic ulcer, cancer of the bowel, gallbladder disease, and dysentery are common sources of infection that may produce peritonitis, as well as blood-borne forms of tuberculosis and pneumonia. *See* PERITONEUM.

Infection may also stem from spread of bacterial organisms from the female organs, the kidneys, and the pancreas. Each form of peritonitis may show both common and specific features.

Contamination of the abdominal cavity may occur from penetration of the abdominal wall during an accident or following a wound. Bile, blood, or fluid from a ruptured abdominal cyst or ectopic pregnancy may also induce a peritoneal inflammation.

Where peritonitis occurs, the normally glistening peritoneum becomes dull, the blood vessels engorge, and a fibrin-containing exudate is produced which may later lead to adhesions. A tendency for localization is apparent, with loops of intestine or other organs forming pockets of inflammation.

The clinical course is quite variable, depending on the agent involved, the type and severity of reaction, and concomitant disease. *See* APPENDICITIS; BACILLARY DYSENTERY; GALLBLADDER DISORDERS; PNEUMONIA; PREGNANCY DISORDERS; TUBERCULOSIS. Edward G. Stuart; N. Karle Mottet

Peritrichia

A specialized subclass of the class Ciliatea composed of a large group of unusual-looking ciliate protozoa. It has excited the curiosity of microscopists for nearly 300 years. Many are sessile and stalked, while some form colonies which may reach a large size. A number are attached as ectocommensals to a variety of animals and plants. A free-swimming stage in the life cycle, indispensable for distribution, is known as the telotroch. It is a small, mouthless form equipped with a single girdle of posteriorly located locomotor cilia. This is quite unlike the morphology of the mature, sedentary form, which is an inverted bell form atop a long stalk. The body is naked of ciliature except for conspicuous wreaths of buccal ciliary organelles at the oral end. Much is made by many protozoologists of the fact that the adoral zone of membranelles, in this instance, winds counterclockwise toward the mouth. Actually, this may be of little real importance, although convenient in taxonomic keys. *Vorticella* (**illus.** *a*) and *Epistylis* are probably the best-known

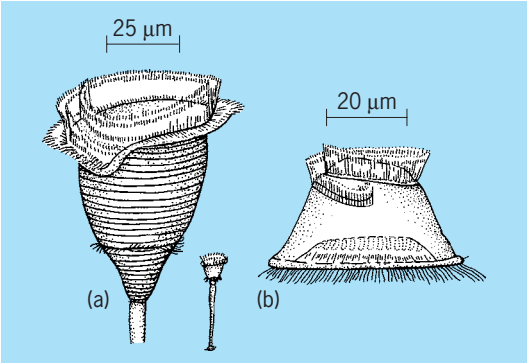

Peritrichida. (*a*) *Vorticella*, a stalked peritrich. (*b*) *Trichodina*, a mobile peritrich.

stalked forms. The former is a solitary ciliate, the latter a colony builder. *Trichodina* (illus. *b*) belongs to the group of mobile peritrichs. Its species are associated with a wide variety of invertebrate and vertebrate hosts on which its actions range from those of a harmless commensal to a pathogenic parasite. The subclass Peritrichia contains a single order, the Peritrichida, whose species may have evolved from such holotrichian orders as the Hymenostomatida or the Thigmotrichida. *See* CILIATEA; CILIOPHORA; HYMENOSTOMATIDA; PROTOZOA. John O. Corliss

Permafrost

Perennially frozen ground. Two classes of frozen ground are generally distinguished: seasonally frozen ground, which freezes and thaws on an annual basis; and permafrost, which is formed when subsurface earth materials remain below 0°C (32°F) for more than a year, without regard to composition, phase, or cementation by ice. Permafrost occurs at most locations where the mean annual temperature at the ground surface is below freezing. Because substantial differences often exist between the temperature at the surface and that measured in the air, climate statistics do not provide a reliable guide to the existence or distribution of permafrost. Although permafrost is ultimately a climatically determined phenomenon, its presence or absence is strongly influenced by local factors, including microclimatic

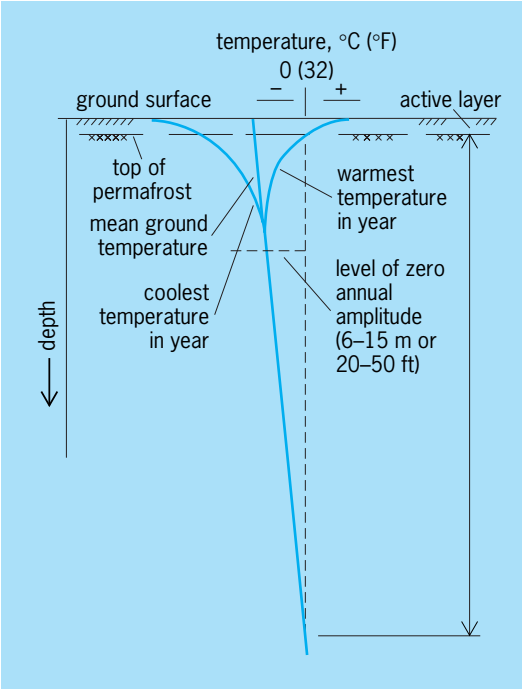

Fig. 1. Typical ground thermal regime in permafrost (area to right of broken line). (*After P. J. Williams, Pipelines and Permafrost: Science in a Cold Climate, Carleton University Press, 1986*)

variations, circulation of ground water, type of vegetation cover, and thermal properties of subsurface materials.

The range of temperatures experienced annually at the ground surface in permafrost terrain decreases progressively with depth, down to a level at which only minute annual temperature variation occurs, termed the level of zero annual amplitude (**Fig. 1**). A zone of near-surface, seasonally frozen ground above permafrost called the active layer often experiences complex heat-transfer processes. Below the permafrost table, the upper limit of material that experiences a maximum annual temperature of 0°C (32°F), heat transfer occurs largely by conduction. In situations where the bottom of the active layer is not in direct contact with the top of the permafrost, the permafrost is a relic of a past colder interval and is not in equilibrium with the present-day climate.

Distribution. Approximately 25% of the Earth's land surface is underlain by permafrost. Substantial areas of subsea permafrost also occur around the land margins of the Arctic Ocean. Permafrost distribution is often classified on the basis of its lateral continuity (**Fig. 2**). In the Northern Hemisphere the various classes form a series of concentric zones that roughly follow the parallels of latitude. Southward deviations in the extent of permafrost in Siberia and central Canada are a response to lower mean annual temperatures in the continental interiors. In the continuous permafrost zone, perennially frozen ground underlies most locations, the principal exception being beneath large bodies of water that do not freeze to the bottom annually. In the discontinuous zone, permafrost may be widespread, but a substantial proportion of the land surface can be underlain by seasonally frozen ground because of variations in local factors. Some scientists also refer to a zone in which permafrost is sporadic, occurring largely as isolated patches that reflect combinations of local factors favorable to its formation or maintenance. Because the thermal properties of peat are conducive to its formation and maintenance, subarctic bogs are the primary locations of permafrost in southerly parts of the subarctic lowlands. In the Southern Hemisphere, permafrost occurs in ice-free areas of the Antarctic continent and in some of the subantarctic islands. Permafrost also occurs extensively in such midlatitude mountain ranges as the Rockies, Andes, and Himalayas, as a response to progressively colder mean annual temperatures at higher elevations. Given sufficient altitude and favorable local conditions, patches of permafrost can exist even in the subtropics, such as in the crater of Mauna Kea at an elevation of 4140 m (13,580 ft) on the island of Hawaii. *See* ANTARCTICA; ARCTIC OCEAN; BOG; PEAT.

The thickness of permafrost at a specific location is determined by the mean annual temperature at the ground surface, the thermal properties of the substrate, and the amount of heat flowing from the Earth's interior. Permafrost thickness ranges from very thin layers to about 1500 m (4920 ft) in some unglaciated parts of Siberia. In general, permafrost thickness increases toward the poles. *See* EARTH, HEAT FLOW IN.

Ice content and surface features. Although the presence of ice is not a criterion in the definition of permafrost, ground ice is responsible for many of the distinctive features and problems in permafrost regions. Ice can occur within permafrost as small individual crystals in soil pores, as lenses or layers of nearly pure ice parallel to the ground surface, and as variously shaped intrusive masses formed when water is injected into soil or rock and subsequently frozen. Many remarkable landforms exist in permafrost regions, although only some are unambiguous indicators of the presence of permafrost. Surface features formed under cold, nonglacial conditions are known collectively as periglacial landforms. Periglacial features that serve as good indicators of the presence of permafrost include ice wedges, vertically oriented ice inclusions created when water produced from melting snow seeps into cracks formed in fine-grained sediments during severe cold-weather events earlier in winter. Repeated many times over centuries or millennia, this process produces tapered wedges of foliated ice more than a meter wide near the surface and extending several meters or more into the ground. Viewed from the air, networks of ice wedges form striking polygonal patterns over extensive areas of the Arctic. Other landforms diagnostic of the presence of permafrost are pingos, ice-cored hills frequently over 10 m (33 ft) in height that can form when freezing fronts encroach from several directions on saturated sediments remaining after drainage of a deep lake. Palsas are smaller mound-shaped forms often found in

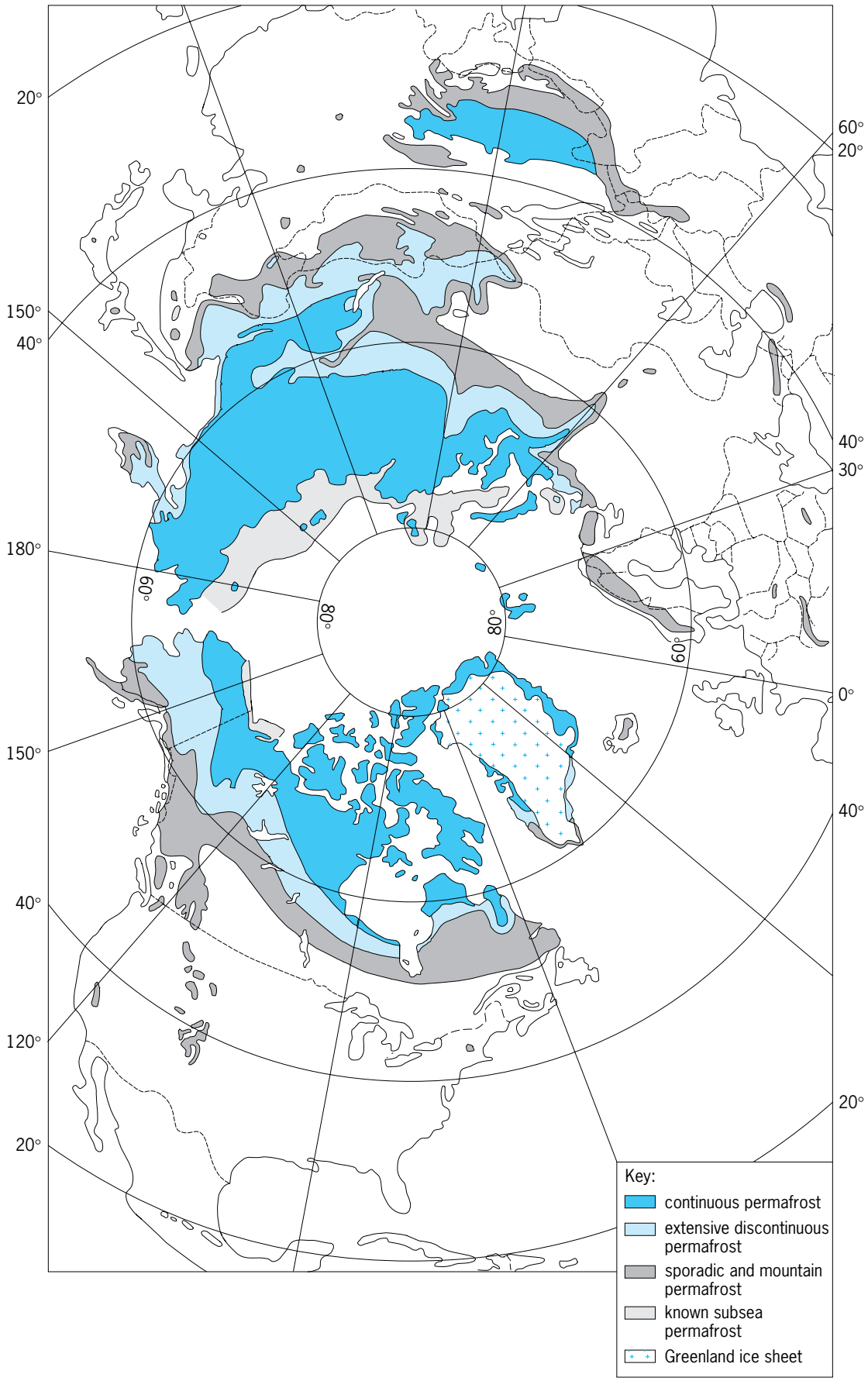

Fig. 2. Map of permafrost distribution in the Northern Hemisphere. (*After J. A. Heginbottom, Geological Survey of Canada*)

subarctic peatlands, and can form in a variety of ways. Thawing of ice-rich permafrost, whether through climatic changes or local disturbance, can result in irregular subsidence of the ground surface and can form a series of pits, mounds, and water-filled depressions known collectively as thermokarst terrain.

Other periglacial landforms occur frequently in association with permafrost, but are not necessarily diagnostic of its presence. When ice-rich permafrost or another impermeable layer prevents infiltration of water, the soil on a hillside may become vulnerable to a slow, flowlike process known as solifluction, giving rise to a network of lobes and terraces that impart a crenulated or festooned appearance to the slope. Other small landforms frequently encountered in subpolar and polar regions, collectively referred to as patterned ground, include small hummocks and networks of striking geometric forms arranged into circles, polygons, and stripes of alternating coarse and fine-grained sediment. The bold appearance of these forms is often accentuated by vegetation patterns.

Engineering considerations. Construction in permafrost regions requires special techniques at locations where the terrain contains ice in excess of that contained within soil pores. Prior to about 1970, many projects in northern Alaska and elsewhere disturbed the surface significantly, triggering thermokarst processes and resulting in severe subsidence of the ground surface, disruption of local drainage patterns, and in some cases destruction of the engineered works themselves. The route of a winter road constructed in 1968–1969 by bulldozing the tundra vegetation and a thin layer of soil can be seen as a scar on the ground surface (**Fig. 3**). This disturbance altered the energy regime at the ground surface, leading to thawing of the underlying ice-rich permafrost and subsidence of up to 2 m (6.6 ft) along the road, which became unusable several years after construction.

Environmental restrictions in North America, based on scientific knowledge about permafrost, have been established to regulate construction activities and minimize their impact on terrain containing excess ice. The Trans-Alaska Pipeline, which traverses 1300 km (812 mi) from Prudhoe Bay on the Arctic coastal plain to Valdez on Prince William Sound near the Gulf of Alaska, carries oil at temperatures above 60°C (140°F). To prevent development of thermokarst and severe damage to the pipe, the line is elevated in sections of the route where surveys indicated the presence of excess ice. To counteract conduction of heat into the ground, many of the pipeline's vertical supports are equipped with heat pipes that cool the permafrost in winter, lowering the mean annual ground temperature and preventing thawing during summer. In several short sections of ice-rich terrain where local aboveground conditions required burial of the line, the pipe is enclosed in thick insulation and refrigerated. *See* PIPELINE.

Fig. 3. Linear thermokarst feature in ice-rich permafrost terrain near Prudhoe Bay, Alaska. (*Frederick E. Nelson*)

Other unusual engineering techniques devised for use in ice-rich permafrost include constructing heated buildings on piles, which allows circulation of air beneath the structures and prevents conduction of heat to the subsurface. Roads, airfields, and building complexes are frequently situated atop thick gravel pads or other insulating materials. In relatively large settlements, water and sewage are transported in insulated, elevated pipes known as utilidors.

Effects of climatic change. Despite implications contained in the term permafrost, perennially frozen ground is not permanent. Permafrost extended farther into the midlatitudes during past episodes of continental glaciation. Evidence for the former existence of permafrost, including ice-wedge casts and pingo scars, has been found in many European countries now far removed from the permafrost regions. In North America, evidence that permafrost existed during colder intervals is scattered along the glacial border from New Jersey to Washington state.

Much of the world's permafrost is within only a few degrees of the melting point of ice and is potentially unstable. Many scenarios of global climatic change suggest that warming may be most severe in the polar regions. Pronounced warming in those portions of the Arctic containing large areas of thaw-sensitive permafrost could lead to development of thermokarst terrain at regional scales, damage many existing engineered works, and cause the boundaries of the permafrost zones (Fig. 2) to be displaced poleward. Moreover, a large amount of carbon is stored in the permafrost regions. An increase in the active-layer thickness or widespread disappearance of permafrost could lead to the release of much of this carbon, intensifying the effects of global warming. The relationship between permafrost and future changes in climate involves much more than temperature fluctuations, and constitutes an important area for research. *See* CLIMATE HISTORY; CLIMATE MODIFICATION.

Frederick E. Nelson

Bibliography. G. H. Johnston, *Permafrost: Engineering Design and Construction*, 1981; F. E. Nelson and S. I. Outcalt, A computational method for prediction and regionalization of permafrost, *Arctic Alpine Res.*, 19(3):279–288, 1987; A. L. Washburn, *Geocryology: A Survey of Periglacial Processes and Environments*, 1980; P. J. Williams, *Pipelines and*

Permafrost: Science in a Cold Climate, 1986; P. J. Williams and M. W. Smith, *The Frozen Earth: Fundamentals of Geocryology*, 1989.

Permeance

The reciprocal of reluctance in a magnetic circuit. It is the analog of conductance (the reciprocal of resistance) in an electric circuit, and is given by Eq. (1),

$$P_m = \frac{\text{magnetic flux}}{\text{magnetomotive force}} = \frac{\int\int \mathbf{B} \cdot d\mathbf{S}}{\oint \mathbf{H} \cdot d\mathbf{l}} \qquad (1)$$

where **B** is the magnetic flux density, **H** is the magnetic field strength, and the integrals are respectively over a cross section of the circuit and around a path within it. *See* CONDUCTANCE.

From Eq. (1), it can be shown that Eq. (2) is valid,

$$P_m = \frac{\mu A}{l} \qquad (2)$$

where A is the cross-sectional area of the magnetic circuit, l its length, and μ the permeability. If the material is ferromagnetic, as is often the case, then μ is not constant but varies with the flux density, and the complete magnetization curve of B against H may have to be used to determine the permeance.

If there are parallel magnetic circuits with the total flux divided between them, the permeances simply add, so that the combined permeance is given by Eq. (3), where $P_1, P_2, \ldots$, are the permeances of the individual branches.

$$P_m = P_1 + P_2 + \cdots \qquad (3)$$

For permeances in series, on the other hand, as for an iron circuit with an air gap, the reciprocals add, as given by Eq. (4).

$$\frac{1}{P_m} = \frac{1}{P_1} + \frac{1}{P_2} + \cdots \qquad (4)$$

See RELUCTANCE.

A. Earle Bailey

Permian

The name applied to the last period of geologic time in the Paleozoic Era and to the corresponding system of rock formations that originated during that period. The Permian Period commenced approximately 290 million years ago and ceased about 250 million years ago. The system of rocks that originated during this interval of time is widely distributed on all the continents of the world.

The Permian System is presently divided into three series: the lower Cisuralian Series with type sections in the western Ural region of Russia; the middle Guadalupian Series with type sections in western Texas and southern New Mexico; and the upper Lopingian Series with type sections in southern China.

Permian rocks contain evidence for a paleogeography that was greatly different from present geography. The Permian Period was a time of variable and changing climates, and during much of this time latitudinal climatic belts were well developed. During the later half of Permian time, many long-established lineages of marine invertebrates became extinct and were not immediately replaced by new fossil-forming lineages. Rocks of Permian age contain many resources, including petroleum, coal, salts, and metallic ores. *See* LIVING FOSSILS.

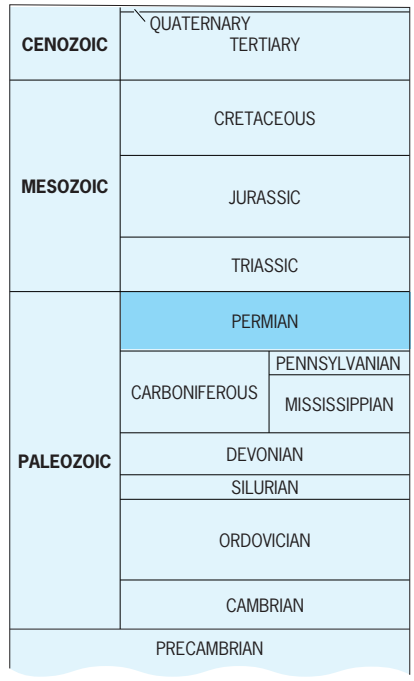

Eastern European Permian. The Permian System was proposed in 1841 by R. Murchison for a succession of marine, brackish, evaporitic, and nonmarine deposits exposed in the former province of Perm on the western flank of the Ural Mountains in Russia (**Fig. 1**). During the Permian Period, this region was near the boundary between the stable Russian Platform on the west and the tectonically active Ural Geosyncline to the east (**Fig. 2**). The ancestral Ural Mountains were gradually formed by the uplift and compressive collapse of the Ural Geosyncline during this time. The geosynclinal deposits are mostly clastic sediments that reach a thickness of several thousand meters. During the early part of the Permian, the edge of the Russian Platform was marked by the development of extensive carbonate banks and reefs. These carbonate deposits contrast markedly with the thick clastic beds of the geosynclinal facies and the thin shales and limestones of the platform facies.

Continental deposition. By the middle of Early Permian time, the marine connection from the Russian Platform southward into the Tethyan ocean was closed, apparently by uplifts along the southern and western margins of the platform. During middle Permian time, the Russian Platform was covered by a broad, shallow, brackish gulf which had limited water circulation with an ocean to the north. Brackish conditions fluctuated with evaporites, redbeds, and

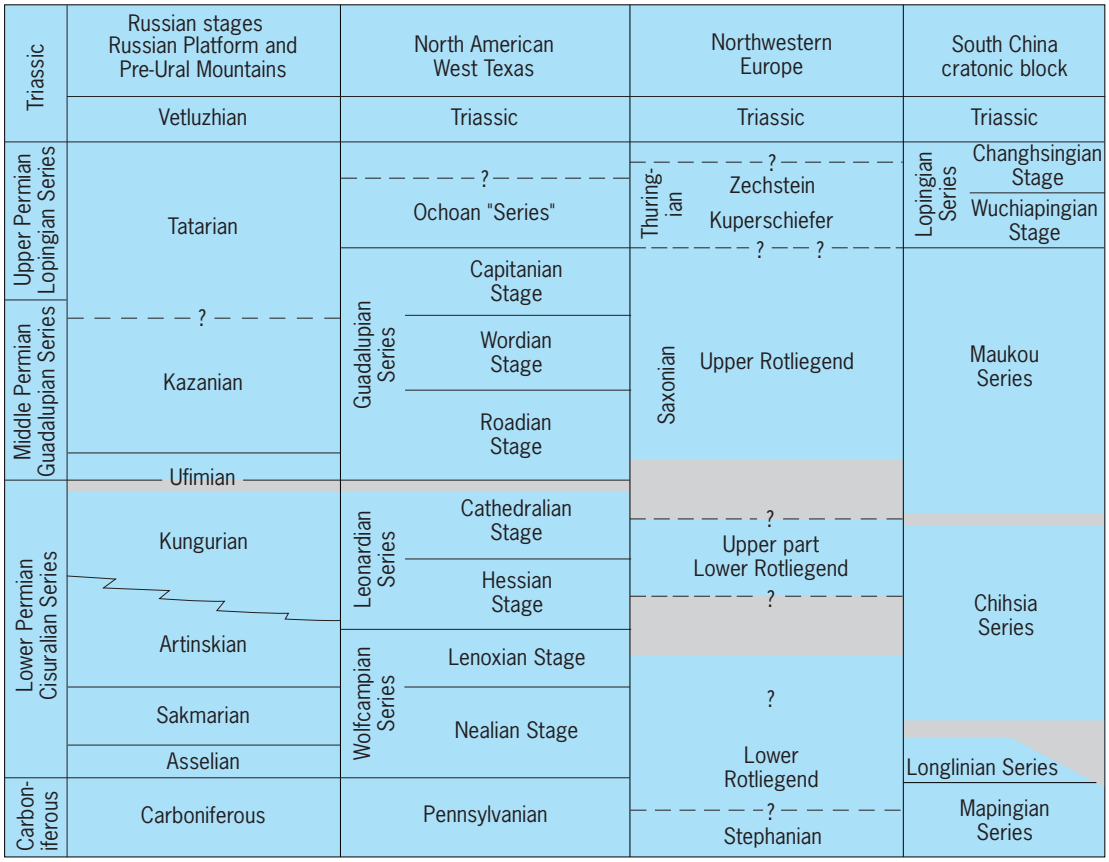

Fig. 1. Correlation of major Permian rock units.

other continental deposits. During later Permian time, detrital material eroded from the uplifted ancestral Ural Mountains, filled the remnants of the deformed geosynclinal depression, and were progressively transported westward onto the Russian Platform. Continental deposition continued into Triassic time.

Boundaries. Both the lower and upper boundary of the Permian System have been subject to considerable discussion. As Murchison originally defined the system, the lower boundary was at the base of the evaporitic facies of the Kungurian beds and the upper boundary at the top of the continental facies of the Tatarian beds. Based on similarities in lithology, Murchison correlated the thick Artinskian clastics that lie beneath the Kungurian within the Ural Geosyncline with the lower Carboniferous of western Europe. He believed the limestones and shales that lie beneath the Kungurian beds on the Russian Platform were part of the upper Carboniferous. Later studies showed that these original correlations were not correct. A. P. Karpinsky in 1889 described Artinskian ammonoids that were closely related to faunas considered Permian in age in other

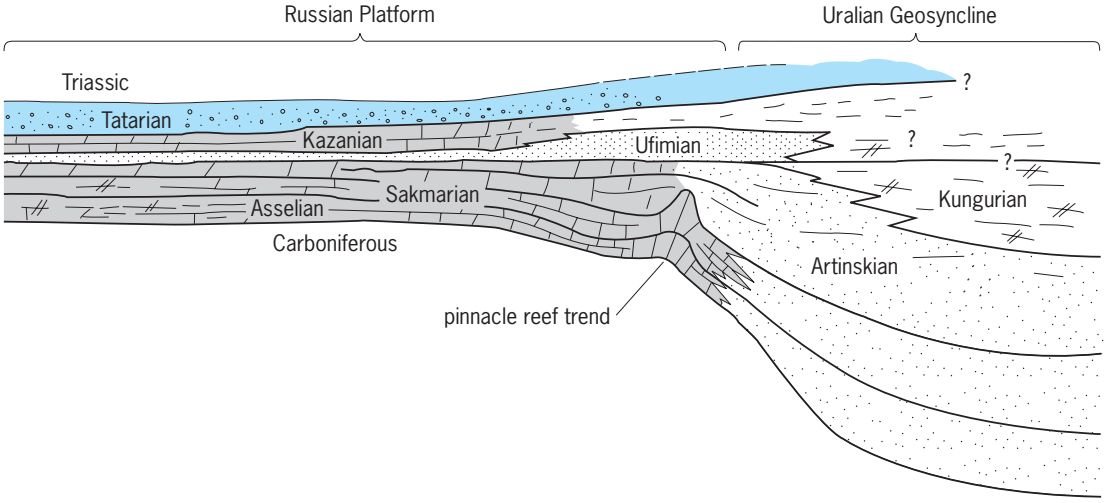

Fig. 2. East-west section of the Permian System in its type area.

parts of the world. The suggestion to lower the Permian boundary in its type region to include these Artinskian faunas gained general acceptance.

Cisuralian Series. The limestones of the Ufa Plateau continued to be classified as upper Carboniferous until after 1930. During the 1930s V. E. Ruzhentsev recognized prolific ammonoid faunal assemblages that were present in both the carbonates and shales of the Ufa Plateau and the Artinskian clastics. Based on these studies, Ruzhentsev subdivided the original Artinskian into a lower Asselian Stage, a middle Sakmarian Stage, and an upper (revised) Artinskian Stage (Fig. 1). In the 1990s, these stages were combined with the Kungurian Stage to form the lower Permian Cisuralian Series (Fig. 1).

The Asselian, Sakmarian, and Artinskian stages include normal marine faunas with abundant fusulines, conodonts, and ammonites, which are extremely useful in interregional correlation. Large, massive reef mounds and pinnacle reefs, 1800 ft (600 m) or more in thickness, are common on the edge of the Russian Platform (Fig. 2), and these pass eastward into thick shales and sandstones in the Ural Geosyncline and westward into thin dolomitic limestone, dolostone, and evaporites.

On the Russian Platform and on the margin of the Ural Geosyncline, the upper part of the Artinskian beds intertongues with Kungurian red shales and evaporites. The Kungurian extends stratigraphically higher and locally is the source of major salt, anhydrite, and potash salts, as near Solikamsk. Fossils are rare and not diverse, and show affinities at some localities to Artinskian species.

Middle Permian deposits. On the Russian Platform, the Ufimian beds lie at the base of the middle Permian and are less continuously distributed and are red shales and sandstones, variegated clays, and marly shales less than 450 ft (150 m) thick. They primarily contain a meager fresh-water fauna of bivalves and ostracodes. Eastward near the Ural Mountains, the Ufimian reaches 4500 ft (1500 m) and is coarser grained.

The Kazanian Stage unconformably overlies the Ufimian and consists of about 300 ft (100 m) of greenish-gray impure limestone that has brackish-water or restricted marine faunas that are mostly bryozoans and brachiopods. The Kazanian deposits were laid down in an elongate shallow basin that paralleled the western flank of the Ural Mountains. This basin was closed to the south, where it locally includes evaporites, and open to the north, where a few additional marine invertebrates, including rare ammonoids, are locally present.

Upper Permian deposits. The youngest stage of the Russian Permian is the Tatarian, which is formed of brightly colored, variegated sandstone, conglomerate, and other clastic beds that were deposited by rivers, streams, and lakes, and includes some dune sands. Plants and terrestrial animals, including insects and vertebrates, are well known from these sediments. The original Tatarian of Murchison had five main faunal zones; however, only the lower two are now considered Permian, and the name Tatarian is restricted to these two zones. The upper three zones are included in the Vetlushian Stage and are Triassic. *See* REDBEDS.

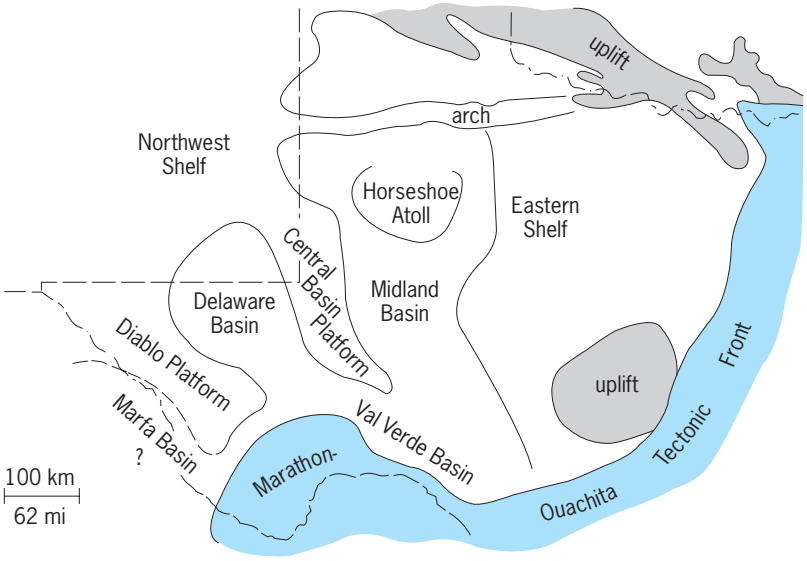

Fig. 3. Depositional and structural relationships in western Texas during Permian time.

Western Texas and southern New Mexico. One of the finest Permian sections known is in the Delaware Basin of western Texas and southeastern New Mexico (**Fig. 3**), where the Permian sequence reaches a thickness of more than 12,000 ft (4000 m). The rocks of the lower 6600 ft (2200 m) are of marine origin and are highly fossiliferous. Although a few fossils from this region were described in 1858, this succession of strata remained virtually unknown until 1909. Intensive study followed the discovery of oil in the region about 1920. In 1939 a committee of Permian specialists proposed to subdivide the American Permian section into four series, and this usage was widely accepted. It has become the standard section for North America and, to a considerable degree, for the world. It is divided into four stages in ascending order: Wolfcampian Stage, 1500 ft (500 m); Leonardian Stage, about 3000 ft (1000 m); Guadalupian Stage, about 3000 ft (1000 m); and Ochoan Stage, about 4500 ft (1500 m).

The Wolfcampian correlates closely with the Russian Asselian, Sakmarian, and part of the Artinskian stages (Fig. 1). The Leonardian correlates with the upper part of the Artinskian and the Kungurian, and has larger and more varied faunas. The Guadalupian cannot be correlated in detail with the middle part of the Russian section, because the latter has restricted faunas. The Ochoan is virtually unfossiliferous, except for a thin zone near the top (in the Rustler Dolostone) which contains productid brachiopods and a few other types of Paleozoic invertebrates.

Wolfcampian and Leonardian. During the Wolfcampian Epoch, most of the region was a broad, shallow marine intracratonic basin in which a rich and varied fauna thrived. A fore-deep basin bordered the northern flank of the active Marathon orogenic belt. By Leonardian time, three distinct basins (Fig. 3) were subsiding more rapidly than the surrounding area, which became a broad shelf occupied by wide, shallow lagoons. Light-colored, fossiliferous

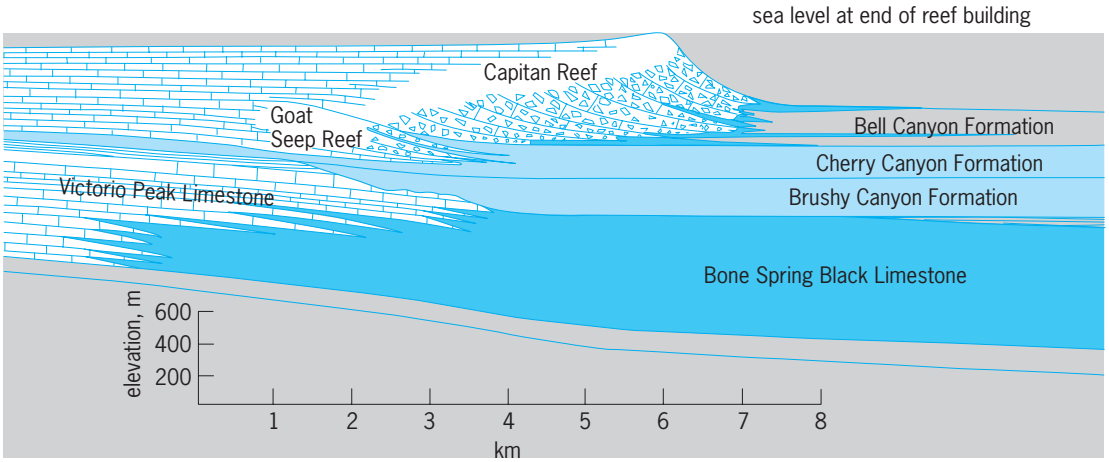

Fig. 4. Section across the Capitan Reef Complex. 1 m = 3.3 ft; 1 km = 0.6 mi.

limestones (Victorio Peak Limestone) accumulated on the shelves, while black limestone and black shale accumulated in the basins. Evidently the threshold to the basins, which was somewhere in Mexico, was so shallow that water in the basins was density-stratified and the bottom was stagnant and foul, so that almost no benthic organisms could survive. The black Bone Spring Limestone is generally barren of fossils.

Guadalupian. During Guadalupian time (middle Permian), the basins continued to deepen, and as the climate became markedly arid, surface water flowed radially out of the basins onto the adjacent platforms to replace the water lost by evaporation in the shallow lagoons. Narrow limy banks grew along the margins of the basin to form the great Capitan Reef. **Figure 4** shows the complex relations within the Leonardian and Guadalupian deposits along the face of the Guadalupe Mountains. Probably no other great reef complex is so well exposed or has been so intensively studied as that of the Capitan Reef. Massive deposits of reef talus were derived from the growing front of the reef. These dip steeply into the basin, become finer down dip, and grade out into thin tongues of calcarenite. The back-reef deposits are calcareous for distances of 1.5–6 mi (2–10 km) and then grade rapidly into gypsiferous shales and anhydrite. Gray, beach, and offshore-bar sands intertongue from the landward margins of the lagoon. Farther back the sands pass into redbeds.

Late Permian. During Ochoan time, the basins became evaporitic under intensely arid conditions. Enormous deposits of anhydrite and, later, of halite were precipitated. Interbedded in the salt in the center of the Delaware Basin are several lenses of potash salts—sylvite, carnallite, and polyhalite.

Other areas. In central Texas, Oklahoma, and Kansas, the Wolfcampian equivalents are largely of marine origin. Marine conditions persisted well into Leonardian time in central Texas, but in Kansas a very large salt deposit (in the Wellington Shale) is followed by a thick redbed sequence. By Guadalupian time, most of the deposits in central Texas were nonmarine redbeds in which several thin marine dolostones intertongued from the west. In Oklahoma, all but the lower part of the Wolfcampian is in redbed facies because of the local influence of the Oklahoma mountains.

In other areas of the central and southern Rocky Mountain states, block faulting during Wolfcampian time resulted in a series of horsts, which were local sources of sediments, and adjacent grabens, which became local basins of deposition. Although most of these basins had initial deposits of marine origin, they were mainly filled by the end of Wolfcampian time. Subsequently, Leonardian and early Guadalupian sediments formed thin, blanketlike continental deposits over much of the region. Dunes, redbeds, and local evaporites are common.

The western margin of the North American craton during the Permian was located far to the east of the present west coast. The Permian margin extended generally northward from southern California through southern and eastern Nevada, southeastern Idaho, eastern British Columbia, and just into the eastern edge of Alaska. The southern part of this margin had a thick succession of Wolfcampian, Leonardian, and early Guadalupian limestones deposited on it. In northern Utah and Idaho, sandstone becomes more prevalent, and in the Guadalupian, phosphorite and chert become important deposits. Farther north, sandstone and chert continue to increase in abundance and form most of the Permian sediments on the old cratonic margin.

West of the Permian cratonic margin, Permian strata are found in four or five tectonically disturbed belts. Although these belts include rocks of the same age, each structural belt contains different fossil faunal assemblages that are quite distinct, and in each the lithologic succession and depositional history appears to be independent. These different assemblages of faunas and sediments were apparently deposited at considerable distances from one another, and their present-day geographic proximity indicates that each belt has been added by tectonic accretion to the western margin of the North American craton in post-Permian time—that is, during the Mesozoic or Cenozoic. Of all these structural belts, the best known is the Cache Creek (British Columbia) belt, which extends from southern British Columbia north-northwestward into southern Yukon. The

Cache Creek belt is composed of oceanic ribbon cherts, dark shales and sandstones, basic igneous flows, and massive limestone reefs having a tropical fossil fauna and flora which show affinities to the Tethyan faunal realm rather than to the Midcontinent-Andean realm of North American and South American cratons.

Northwestern Europe. Permian beds are widespread in northwestern Europe and the North Sea and produce large amounts of gas and oil, which have had an important influence on the economies of northwestern European countries since the late 1960s. These Permian beds are divisible into two parts. The lower part (Rotliegend beds) is mainly red sandstone. The upper part (Thuringian Series) consists of conglomerate, chalcopyritic shale, dolomitic limestone, evaporites, and shale. The Rotliegend beds are subdivided into the Autunian Stage (or Lower Rotliegend) in the lower part and Saxonian Stage (or Upper Rotliegend) in the upper part. Saxonian strata are younger than the last major movements of the Hercynian orogeny, and an unconformity separates them from the overlying Thuringian.

The Thuringian contains a basal conglomerate, a thin copper-bearing shale (Kupferschiefer), dolomitic limestones (Zechstein), and seven important evaporitic units that near Stassfurt include potassium salts. The Thuringian sea extended east across Poland, and possibly connected with the Kazanian sea of the Russian Platform. The Magnesian Limestone of England is a thin western tongue of the Thuringian Series.

The Atlantic Ocean did not start to open until middle Mesozoic time, so that the Permian deposits of northwestern Europe are very similar to, and are thought to have been formerly continuous with, Permian beds found along the central east coast of Greenland, on Spitsbergen, and on the Canadian Arctic islands. They also connected, by way of the Barents shelf, to the northern entrance to the Ural Geosyncline and Russian Platform.

Tethyan regions. South of the Hercynian orogenic belt, the sediments and faunas and floras of the Permian change markedly in Mediterranean Europe, northern Africa, and southern and central Asia (**Fig. 5**). Limestone, dark shales and sandstones, cherts, and basaltic volcanics are common. The species and generic diversity is great in contrast to that of northwestern Europe and the Russian Platform.

The marine Tethyan invertebrate fauna evolved rapidly during the middle part of the Permian and formed a distinctive biogeographic realm, called the Tethyan realm. This diverse fauna is characterized by the verbeekinid fusulinaceans (foraminiferal Protozoa) and colonial and solitary waagenophyllid corals (Rugosa). Later orogenic movements during the Mesozoic and Cenozoic have greatly complicated the interpretation of these Permian deposits; however, several linear belts of Permian rocks appear to be present. Each may have included shallow reef deposits and shallow- and deeper-water clastic and carbonate deposits, and some belts may have fringed a number of small fragments of cratonic crustal blocks. In post-Permian time, sea-floor spreading and crustal subduction apparently displaced these several depositional belts and cratonic fragments and lodged them as accreted terranes in orogenic belts against larger cratons, such as Europe, central and eastern Asia, and western North America.

The Permian Tethyan faunas were apparently tropical, and mostly shallow-water because of the abundance of calcareous algae, bryozoans, brachiopods,

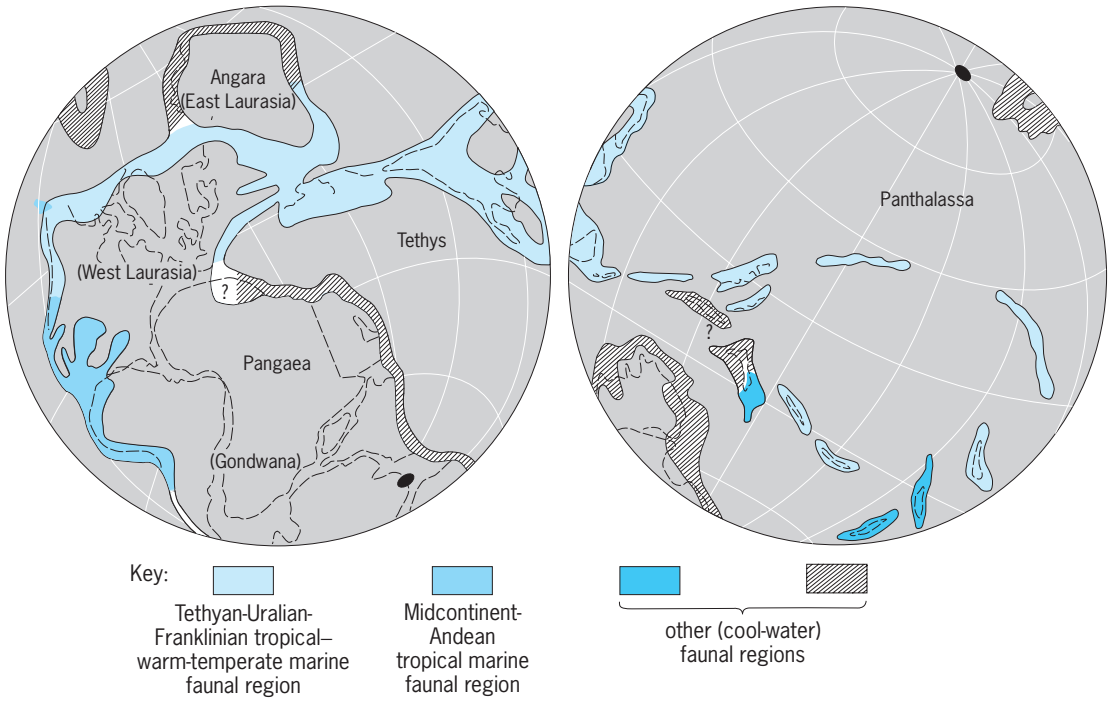

Fig. 5. Paleogeography during the earliest part of the Permian Period.

and echinoderms, and the reeflike nature of most of the limestones. The close association of these limestones with basic igneous rocks (including pillow lavas), ribbon cherts, and dark sandstones and shales suggests that much of the Tethyan region was made up of island arcs and oceanic carbonate plateaus similar to those of the present tropical Pacific Ocean.

Southern China. The lower part of the Permian in southern China includes the upper part of the Maping Stage and the Longlinian (Changshan) Stage which are widespread, but locally deeply eroded, beneath a middle lower Permian unconformity. Overlapping this unconformity are extensive limestones of the Chihsia Series which comprise the remainder of the lower Permian in the area. An extensive unconformity at the top of the Chihsia represents the mid-Permian sea-level event. Higher, widely distributed limestones of the Maukou Series are present. They show major lateral variations in lithologic facies.

Lopingian Series. The type area for the upper series of the Permian, the Lopingian Series, overlies the Maukou in continuous succession in much of southern China. This series is divided into the the Wuchiapingian Stage (below) and the Changhsingian Stage (above) and the top of the succession passes conformably into the Lower Triassic.

Gondwana continents. The Permian successions are remarkably similar in all of the parts of southern Africa, Australia, South America, Pakistan and India, and Antarctica that formed the large supercontinent of Gondwana during the late Paleozoic. In southern Africa these deposits form the lower part of the Karoo System and commence with a tillite (Dwyka Tillite), which is followed by dark shales (Ecca Series) overlain by sandstones and red shales (Beaufort Series). The Ecca includes the late Paleozoic coals of southern Africa, and the Beaufort has a distinctive and extensive mammallike reptilian fauna.

The Permian deposits of much of South America have close similarities to those of southern Africa. In the Paraná Basin the Guata Group with glacial sediments, coal, and marginal marine beds, and, in its upper part, *Eurydesma* (Bivalvia) and *Glossopteris* (plant) is of probable Early Permian age. Above, the Irati Formation has the same distinctive reptiles found in the Beaufort of southern Africa and also has an extensive plant fossil succession.

Antarctica also has a generally similar late Paleozoic succession to those found in southern Africa and southern South America, including coal and plant and reptile fossils.

Australia, Pakistan, and India have late Paleozoic successions preserved in a number of fault-bounded basins. The adjacent parts of the Gondwanan craton were apparently emergent. Many of these basins include tillites and glacial marine deposits in their Carboniferous and earliest Permian deposits. These are commonly followed by marine deposits having some early and middle Permian limestones. Several important New South Wales coalfields, such as the Newcastle Coal Measures, are of late Permian age and have nontropical Tatarian insect faunas. Coal formation continued into the Mesozoic in many of these basins.

Paleogeography. During the Permian Period, several important changes took place in the paleogeography of the world. The joining of Gondwana to western Laurasia (Fig. 5), which had started during the Carboniferous, was completed during Wolfcampian time (earliest Permian). The addition of eastern Laurasia (Angara) to the eastern edge of western Laurasia finished during Artinskian time (middle to latest early Permian) and completed the assembly of the supercontinent Pangaea (**Fig. 6**). The climatic effects of these changes were dramatic. Instead of having a circumequatorial tropical ocean, such as during the middle Paleozoic, a large landmass with several high chains of mountains extended from the South Pole across the southern temperate, the tropical, and into the north temperate climatic belts. One very large world ocean, Panthalassa and its western tropical branch, the Tethys, occupied the remaining 75% of the Earth's surface, with a few much smaller cratonic blocks, island arcs, and atolls. *See* CONTINENTAL DRIFT; CONTINENTS, EVOLUTION OF.

Early Permian glacial deposits of the Gondwanan continents lie within an area about 40° from the south paleopole. Recently reported Permian glacial beds in eastern Angara would have been within 30 to 40° of the north paleopole. Middle and late Permian sediments, such as sand dunes, marine beds having a few invertebrate fossils with warmer-water affinities, and coals, suggest warmer conditions but not tropical or subtropical. Considerable evidence also suggests the world climate became progressively milder through a series of fluctuating warming and cooling steps during the later part of the early Permian and late Permian. Desert conditions became widespread in many parts of tropical and subtropical Pangaea with dune sands, evaporites, redbeds, and calcic soil zones.

Vast deposits of salt and anhydrite accumulated in Kansas, New Mexico, and the Permian Basin in western Texas. On the eastern part of the Colorado Plateau, extensive dune sands were deposited, such as in the Canyon DeChelly area. Elsewhere similar conditions are shown by the evaporites of the Kungurian on the Russian Platform and the dune sands and salt deposits of the Thuringian of northwestern Europe. *See* PALEOGEOGRAPHY; SALINE EVAPORITES.

Life. Most marine invertebrates of the Early Permian were continuations of well-established phylogenetic lines of middle and late Carboniferous ancestry. During early Permian time, these faunas, dominated by brachiopods, bryozoans, conodonts, corals, fusulinaceans, and ammonoids, gradually evolved into a number of specialized lineages. The tropical shallow-water faunas of southwestern North America evolved almost in isolation from those of the Tethyan region, because faunal exchanges had to cross either the cooler waters of a temperate shelf or deep waters of Panthalassa. Fluctuating climates permitted rare dispersals of some faunas between these two tropical faunal realms during the early part of early Permian time. The closure of the Uralian seaway during Artinskian time, however, extended that dispersal path north around Angara and into cold boreal waters through which the tropical species could not disperse. With the extension of this dispersal

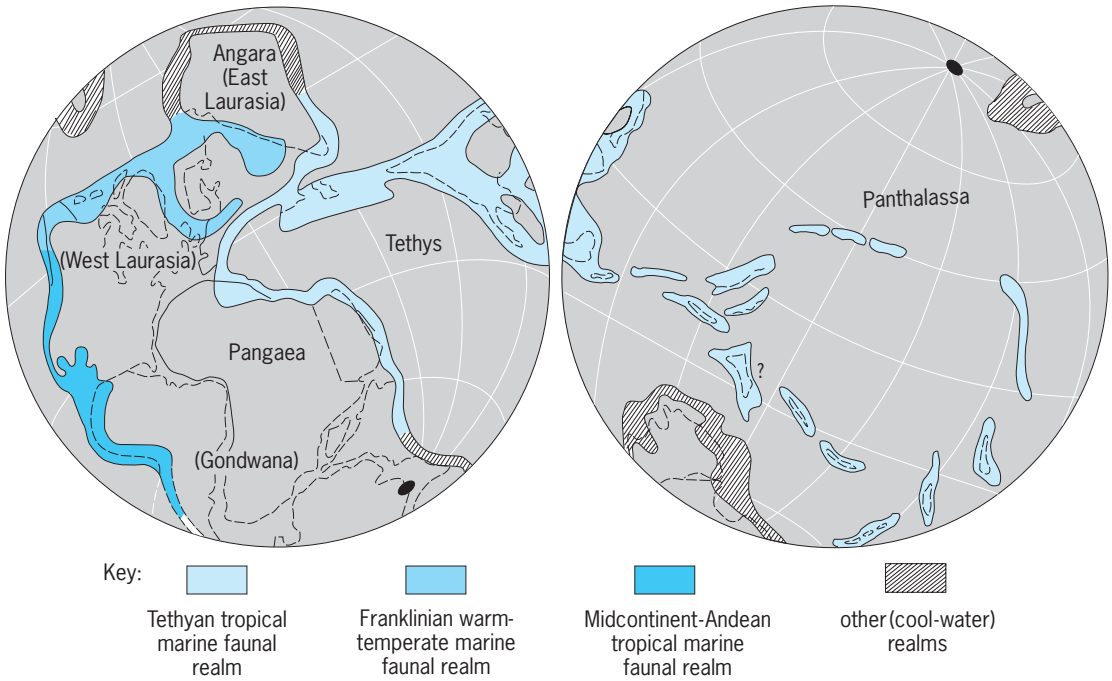

Fig. 6. Paleogeography near the middle of the Permian Period.

path, the faunas of the two tropical realms evolved independently, with only extremely rare dispersals between them.

The Siberian traps, an extensive outflow of very late Permian basalts and other basic igneous rocks (dated at about 250 million years ago), are considered by many geologists as contributing to climatic stress that resulted in major extinctions of many animal groups, particularly the shallow-water marine invertebtares. The end of the Permian is also associated with unusually sharp excursions in values of the carbon-12 isotope (^{12}C) in organic material trapped in marine sediments, suggesting major disruption of the ocean chemistry system.

Foraminiferans. The warm to tropical shallow-water foraminiferal faunas during the Permian were dominated by the fusulinaceans. One group, the verbeekinids, which evolved in the Tethyan realm, formed the distinctive foraminiferal fauna of reefs and atolls of the Tethyan realm. Schwagerinids were less abundant and occurred in sandy sediments adjacent to the reefs. Many of the schubertellids appeared to be adapted to lagoonal environments. In the tropical Midcontinent-Andean realm along the western coast of Pangaea, the schwagerinids filled more of the shallow water niches than in the Tethyan realm, but they were not as diverse in number of new genera and species. Away from the paleotropics in both realms, species diversity decreases rapidly, nearly in proportion to the amount of limestone that was deposited in the succession. Near the end of Guadalupian time, fusulinaceans declined markedly. In the Lopingian specialized genera persisted in the Tethyan region until the end of the period. *See* FORAMINIFERIDA.

Brachiopods. Among the brachiopods, the strophomenids became highly specialized and important framework builders in many bioherms and small reefs. Rhynchonellids and spiriferids are of interest because of their relict pattern of distribution after their evolutionary diversification during the Devonian. Locally they are abundant and important. Near the end of the Guadalupian, brachiopods also were greatly reduced, and about half of those that survived the Lopingian became extinct before the Triassic. Only a few genera and species in one-fifth of the families that were present in the early Permian survived into the Triassic. *See* BRACHIOPODA.

Bryozoans. These have much the same history as other marine invertebrates during the Permian. Several families of cryptostomes, such as the polyporids, hyphasmoporids, and nikiforovellids, and several families of trepostomes, such as the eridotrypellids and araxoporids, became increasingly diverse during the middle part of the Permian. Their greatest geographical diversity was in the Tethys. During Guadalupian time, many genera within the bryozoan families became extinct and, by the end of that epoch, more than 10 families were extinct. Six more families became extinct during the Lopingian (latest Permian). At least five families range into the Triassic. *See* BRYOZOA.

Ammonoids. This important Permian fossil group also had a rapid middle Permian diversification, followed by a rapid reduction in genera late in Guadalupian time. Of about 55 Guadalupian genera, 15 survived into the earliest Lopingian (earliest late Permian), where they evolved into about 50 genera. Only 6 of these survived into the late Lopingian (late late Permian), however; they evolved into nearly 30 genera. Only one ammonoid genus survived the end of the Permian and ranged into the Early Triassic.

Insects. Terrestrial faunas included insects which showed great advances over those of the Carboniferous Coal Measures. Several modern orders emerged, among them the Mecoptera, Odonata, Hemiptera,

Trichoptera, Hymenoptera, and Coleoptera. Extensive insect faunas are known from the lower Permian rocks of Kansas and Oklahoma, the Permian of Russia, and the upper Permian of Australia. *See* INSECTA.

Land plants. During the Permian, plants, including lepidodendrons and cordaites, were well adapted to the moist conditions of low-lying coal swamps. Several lineages also adapted to the drier, well-drained conditions of mountains and alluvial plains, particularly conifers. In glaciated areas of Gondwana, the tongue ferns *Glossopteris* and *Gangamopteris* are common and were apparently adapted to cold climates.

Vertebrates. Of the vertebrates, labyrinthodont amphibians were common and varied; however, reptiles showed the greatest evolutionary radiation and the most significant advances. Reptiles are found in abundance in the lower half of the system in Texas and throughout most of the upper part of the system in Russia and also are common in Gondwana sediments. Of the several Permian reptilian orders, the most significant was the Theriodonta, or mammallike reptiles, that evolved in the Triassic into mammals. These reptiles carried their bodies off the ground and walked or ran like mammals. Unlike most reptiles, their teeth were varied—incisors, canines, and jaw teeth as in the mammals—and all the elements of the lower jaw except the mandibles showed progressive reduction. Most of the known theriodonts are from South Africa and Russia. *See* PALEOZOIC; REPTILIA; THERAPSIDA. Charles A. Ross; June R. P. Ross

Bibliography. C. O. Dunbar and K. M. Waage, *Historical Geology*, 3d ed., 1969; A. L. DuToit, *Geology of South Africa*, 3d ed., 1954; H. Falke (ed.), *The Continental Permian in West Central and South Europe*, 1976; C. R. Handford et al., *Regional Cross Sections of the Texas Panhandle: Precambrian to Mid-Permian*, 1982; R. T. Magginett, C. E. Stevens, and P. Stone (eds.), *Early Permian Fusulinids from the Owens Valley Group, East-Central California*, 1988; N. D. Newell et al., *The Permian Reef Complex of the Guadalupe Mountains Region, Texas and New Mexico*, 1953; C. A. Ross and J. R. P. Ross, Permian, in R. A. Robinson and C. Teichert (eds.), *Treatise on Invertebrate Paleontology*, pt. A, pp. 291–350, 1979; S. M. Stanley, *Earth and Life Through Time*, 1985; D. H. Tarling, *Paleomagnetism: Principles and Applications in Geology, Geophysics, and Archaeology*, 1983.

Permineralization

A fossilization, or taphonomic, process characterized by the deposition of a mineral or mineraloid (an amorphous "gel mineral") in the pore spaces of an organism's body structures. Pores in bone, shell, or plant material may be impregnated by minerals or mineraloids precipitated from ground water, lake or stream water, or ocean water. Minerals commonly involved in permineralization are quartz, calcite, pyrite, siderite (iron carbonate), and apatite (calcium phosphate). Opal, a mineraloid, is also a common agent of permineralization. Normally, the structural parts of organisms undergo little, if any, shape change as the pores are filled, leading in many examples to exquisite preservation of the fine structural details of bodily remains. *See* FOSSIL; TAPHONOMY.

Permineralization can occur in two ways. One way is by the simple infilling of pore spaces. This often occurs in the permineralization of bone (notably dinosaur bone) by silica, or the formation of a coal ball (a subspherical mass containing mineral matter embedded with plant material) by carbonate mineralization of peat. The fossil retains its original shape because its pore spaces were filled before sediment compaction occurred. The second type of permineralization, called petrifaction, involves the filling of pore spaces together with replacement of organic matter or skeletal parts by mineral or mineraloid matter. Large pores and smaller spaces within and between cells of the tissue are filled first. Simultaneously or shortly afterward, the structural framework of the organism's remains, usually organic tissues, decays or dissolves. Finally, to ensure preservation of delicate structures in physically undistorted condition, replacement of the lost material must occur before sediment is compacted. A common form of petrifaction is that of fossil wood: as infilling of pore spaces occurs, cellulose of the woody material decays or dissolves and is quickly replaced by silica or another substance. *See* COAL BALLS; PETRIFACTION.

Normally, the source of silica for precipitation of quartz or opal is water saturated in silica percolating through sediment. Silicification, or permineralization by means of silica, is often associated with hydrothermal sources, volcanic sources, or other igneous settings. Hot water is usually required to dissolve silica and carry its ions. Why silica precipitates preferentially on organic remains and not in adjacent sediment has not been fully determined. It is likely that when silica-saturated water comes in contact with decaying organic remains, microbial biofilms associated with decay mediate the initial precipitation of silica. (Biofilms are aggregations of microorganisms, enveloped by the extracellular biopolymers they produce, adhering to the interface of a liquid and a surface.) *See* OPAL; QUARTZ; SILICA MINERALS.

Petrified wood, in addition to preserving fine structural details such as cells and growth rings, is often colorful and prized as a gemstone. The colors in wood silicified by chalcedony (cryptocrystalline quartz, including jasper and agate) are imparted by small quantities of impurities in the silica. Nearly pure chalcedony is clear, white, or gray. Carbon is a source of black pigmentation. Greens and blues are the result of copper, chromium, or cobalt as impurities. Iron oxides yield red, pink, brown, and yellow hues; manganese yields blue, violet, or black color; and manganese oxides yield pink to orange hues. *See* PALEOBOTANY; PETRIFIED FORESTS.

Permineralization of remains by pyrite (called pyritization), calcite (calcification), siderite (siderite mineralization), phosphate (phosphatization), or other minerals takes place in the presence of water, either in fully aqueous environments or in ground-water settings. For pyrite to precipitate, water must be

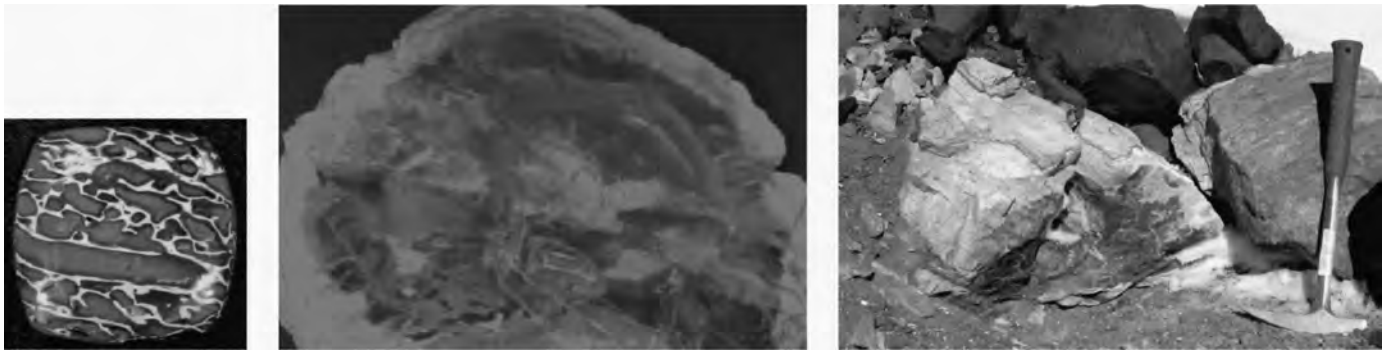

(a) (b) (c)

Fossils preserved by permineralization. (*a*) Polished specimen (a cabochon) of dinosaur bone from the Jurassic of Utah showing original phosphatic bone (white) with pore spaces filled in by jasper, a form of chalcedony (cryptocrystalline quartz, or silica). Length of specimen is 16 mm (0.6 in.). (*b*) Polished slab of petrified wood from the Neogene of Washington showing infilling of pores and replacement of woody tissue by chalcedony. Length across specimen is about 10 cm (4 in.). (*c*) Petrified logs (reddish brown and gray colors in original) weathering free of sedimentary matrix in association with dolerite (or diabase, a basaltic rock, dark brownish red to black in color) from the Jurassic of the Central Transantarctic Mountains, Antarctica. Igneous magma is the source of the silica that petrified the logs.

enriched in iron and sulfur ions. Anoxia (lack of oxygen) seems to be a prerequisite condition. Mounting evidence suggests that anoxic conditions can be quite limited in scope, in some cases limited to a chemical microenvironment created by a microbial biofilm, or decay halo, around decaying organic tissues. Microbes associated with the biofilm mediate the initial precipitation of iron sulfide within weeks of the start of the decay process. For carbonates to precipitate, water must be rich in bicarbonate ions. Siderite precipitation also requires the presence of iron ions. Phosphatization requires ions of calcium and phosphate in water. Similar to pyritization, carbonate and phosphate mineralization seem to occur preferentially within a decay halo in which microbes mediate the initial precipitation of minerals. Loren E. Babcock

Bibliography. R. S. Boardman, A. H. Cheetham, and A. J. Rowell (eds.), *Fossil Invertebrates*, 1987; P. S. Borkow and L. E. Babcock, Turning pyrite concretions outside-in: Role of biofilms in pyritization of fossils, *Sed. Rec.*, 1(3):4–7, 2003; R. M. Feldmann and M. Hackathorn (eds.), Fossils of Ohio, *Ohio Div. Geol. Surv. Bull.*, no. 70, 1996; C. L. Fenton and M. A. Fenton, *The Fossil Book*, 1958; R. C. Moore, C. G. Lalicker, and A. G. Fischer, *Invertebrate Fossils*, 1952.

Permittivity

A property of a dielectric medium that determines the forces that electric charges placed in the medium exert on each other. If two charges of q_1 and q_2 coulombs in free space are separated by a distance r meters, the electrostatic force F newtons acting upon each of them is proportional to the product of the charges and inversely proportional to the square of the distance between them. Thus, F is given by Eq. (1), where $1/(4\pi\epsilon_0)$ is the constant of proportionality, having the magnitude and dimensions necessary to satisfy Eq. (1). This condition leads to a value for ϵ_0, termed the permittivity of free space, given by Eq. (2), where c is the velocity of light in vacuum.

$$F = \frac{q_1 q_2}{4\pi\epsilon_0 r^2} \qquad \text{newtons} \tag{1}$$

$$\epsilon_0 = \frac{1}{4\pi 10^{-7} c^2} \simeq 8.8542 \times 10^{-12} \qquad \text{farads/meter} \tag{2}$$

In all that follows, dielectric media are taken to be homogeneous and isotropic. If now the charges are placed in such a medium, the force on each of them is reduced by a factor ϵ_r, where ϵ_r is greater than 1. This dimensionless scalar quantity is termed the relative permittivity of the medium, and the product $\epsilon_0\epsilon_r$ is termed the absolute permittivity ϵ of the medium.

A consequence is that if two equal charges of opposite sign are placed on two separate conductors, then the potential difference between the conductors will be reduced by a factor ϵ_r when the conductors are immersed in a dielectric medium compared to the potential difference when they are in vacuum. Hence a capacitor filled with a dielectric material has a capacitance ϵ_r times greater than a capacitor with the same electrodes in vacuum would have. Except for exceedingly high applied fields, unlikely normally to be reached, ϵ_r is independent of the magnitude of the applied electric field for all dielectric materials used in practice, excluding ferroelectrics. *See* CAPACITANCE; FERROELECTRICS.

Polarization. Since, for a given fixed distribution of charges, the introduction of a dielectric medium reduces the electric field everywhere by a factor ϵ_r, the medium must produce a flux P coulombs/meter^2, termed the polarization, which opposes the flux D produced by the charges. The reduction in electric field E can be expressed in terms of ϵ_r [Eq. (3)] or in terms of P [Eq. (4)]. The dimensionless ratio of P to $\epsilon_0 E$ is known as the electric susceptibility χ_e of

$$E = \frac{D}{\epsilon_0\epsilon_r} \qquad \text{volts/meter} \tag{3}$$

$$E = \frac{D - P}{\epsilon_0} \qquad \text{volts/meter} \tag{4}$$

the dielectric. Eliminating D between Eqs. (3) and (4) yields Eq. (5). Since the value of ϵ_r for vacuum

$$\chi_e = \frac{P}{\epsilon_0 E} = \epsilon_r - 1 \tag{5}$$

is unity, χ_e represents the component of the relative permittivity which is due to the electrical response of the medium itself to the applied field. *See* ELECTRIC SUSCEPTIBILITY; POLARIZATION OF DIELECTRICS.

Dipole moment. When two charges of equal magnitude q and opposite signs are separated by a fixed distance d, they constitute a dipole. If the line joining the charges were to make a right angle with an electric field E, the dipole would experience a torque qEd. The torque in unit field is termed the dipole moment μ. Thus Eq. (6) holds. In a dielectric material in

$$\mu = qd \qquad \text{coulombs} \times \text{meter} \tag{6}$$

which a polarization P exists, an elementary cuboid may be imagined of volume δv with two faces perpendicular to P of area δA, separated by a distance δx. Its dipole moment is $P\delta A\delta x$, or $P\delta v$, and so the polarization P is the dipole moment per unit volume of dielectric. This dipole moment is determined by the polarizability of the submicroscopic entities of the material. *See* DIPOLE; DIPOLE MOMENT.

Mechanisms of polarization. Polarization may arise due to both nonpolar and polar molecules.

Nonpolar molecules. Even if the distribution of charges within a molecule is such that it has no dipole moment, the application of an electric field E can cause a change in this distribution and so cause each molecule to have an induced dipole moment. The dipole moment μ_e caused by the displacement of electrons within the atom, and the dipole moment μ_a caused by the displacement of atoms within the molecule, give rise to electronic and atomic polarization respectively. Normally electronic polarization is larger than atomic polarization by a factor of, say, 5–10, but for some materials, for example, for alkali halide crystals, atomic polarization is considerable.

The electric field E_l that acts at the molecular site and causes charge displacements is termed the local field. Although E_l is proportional to the applied field E, E_l is in general not equal to E, since E_l contains an additional component arising from neighboring induced dipoles. The induced dipole moments and E_l are continuously fluctuating, and a rigorous treatment demands the use of sophisticated statistical mechanics. Time-averaged values are implied here throughout. *See* STATISTICAL MECHANICS.

Since μ_e and μ_a are proportional to E_l, electronic polarizability α_e and atomic polarizability α_a may be defined by Eqs. (7) and (8). The electronic and

$$\alpha_e = \frac{\mu_e}{E_l} \qquad \text{farads} \times \text{meter}^2 \tag{7}$$

$$\alpha_a = \frac{\mu_a}{E_l} \qquad \text{farads} \times \text{meter}^2 \tag{8}$$

atomic polarizations P_e and P_a are obtained from the induced dipole moments μ_e and μ_a by multiplying by the number density of molecules N. Thus the induced polarization P_i is given by Eq. (9).

$$P_i = N(\alpha_e + \alpha_a)E_l \qquad \text{coulombs/meter}^2 \tag{9}$$

Polar molecules. Some molecules have a charge distribution such that, even when no electric field is applied, the molecule possesses a dipole moment. Such molecules are termed polar molecules, and their dipole moment μ_p is termed permanent, to distinguish it from the induced dipole moment which, in common with nonpolar molecules, they also possess. In general, depending on the structure of the material, the molecules will have some freedom of rotational movement. In the absence of an applied electric field, the permanent dipole moments are randomly directed, and so the sum of their components in any given direction is zero. However, when a steady field is applied, the summation of the dipole moments in the field direction yields a net positive value. Thus the ordering effect of the field is in competition with the thermal disordering effect of the environment, and the resulting orientational polarization P_0 is inversely proportional to absolute temperature T and, except for exceedingly high fields unlikely normally to be reached, directly proportional to E_l, where E_l now includes a component arising from neighboring permanent dipoles. It is easily shown, following a calculation carried out by P. Langevin in 1905 relating to magnetic dipoles, that for all realistic fields P_0 is given by Eq. (10), where

$$P_0 = \frac{N\mu_p^2 E_l}{3kT} \tag{10}$$

k is Boltzmann's constant. Thus the effective orientational polarizability α_p is given by Eq. (11). The total polarizability α of a polar molecule is thus given by Eq. (12), and the total polarization P is given by Eq. (13). Similarly, the total electric susceptibility χ_e

$$\alpha_p = \frac{\mu_p^2}{3kT} \qquad \text{farads} \times \text{meters}^2 \tag{11}$$

$$\alpha = \alpha_e + \alpha_a + \alpha_p \qquad \text{farads} \times \text{meters}^2 \tag{12}$$

$$P = N\alpha E_l \qquad \text{coulombs/meters}^2 \tag{13}$$

is the sum of three components, χ_{el}, χ_a, and χ_p. *See* POLAR MOLECULE.

Local field. Substituting the expression for P of Eq. (13) into Eq. (5) gives Eq. (14). In order to evaluate this equation, it is necessary to know E_l/E. It can be shown that E_l is given by Eq. (15), where the

$$\chi_e = \epsilon_r - 1 = \frac{N\alpha}{\epsilon_0} \times \frac{E_l}{E} \tag{14}$$

$$E_l = E + \frac{\gamma P}{\epsilon_0} \qquad \text{volts/meter} \tag{15}$$

dimensionless constant γ is dependent on the internal arrangement of molecules in the material. For the special cases of either solids with cubic atomic structure, or media in which the molecules are in complete disorder, $\gamma = {}^1/_3$; then E_l is given by

Eq. (16). A classical calculation carried out by

$$E_l = E + \frac{P}{3\epsilon_0} \qquad \text{volts/meter} \tag{16}$$

H. A. Lorentz gives this expression for E_l, which is therefore known as the Lorentz field.

Substituting the expression for P of Eq. (5) into Eq. (16) gives Eq. (17). Substituting this value for E_l/E into Eq. (14) gives Eq. (18). This is the Clausius-

$$\frac{E_l}{E} = \frac{\epsilon_r + 2}{3} \tag{17}$$

$$\frac{\epsilon_r - 1}{\epsilon_r + 2} = \frac{N\alpha}{3\epsilon_0} \tag{18}$$

Mossotti equation, relating the macroscopic quantity ϵ_r with the submicroscopic quantity α, for materials for which $\gamma = {}^1/_3$. It takes account of dipole-dipole interaction, whereas failure to do so would result in the expression for ϵ_r given in Eq. (14) with the term E_l/E put equal to unity.

Complex permittivity. When an ac electric field is applied to a dielectric, each polarization mechanism can only operate effectively up to a limiting frequency. Then, further increase in frequency will cause the mechanism to become gradually less effective until it virtually ceases to operate. The frequency range over which the polarization is diminishing is known as the dispersion region. Because of the springlike nature of the forces involved in electronic and atomic polarization, their disappearance is accompanied by resonant behavior of ϵ_r. The dropping out of the orientational polarization is, however, a relaxation process, possibly involving a broad distribution of relaxation times, and therefore it occurs over a much wider frequency range. Typically, the midfrequencies f_0 involved would be about 10^{15} Hz for electronic polarization and 10^{12} Hz for atomic polarization, while for orientational polarization f_0 might be anywhere in a wide range, say 10^4 to 10^{10} Hz, depending on the material and its temperature. *See* RESONANCE (ALTERNATING-CURRENT CIRCUITS).

Each dispersion region is associated with an energy loss, or absorption, and the energy loss per period per unit volume, J, exhibits a peak at f_0. The existence of energy loss means that, on a phasor diagram, D, instead of being in phase with E, lags E by a loss angle δ. Thus it is evident from Eq. (3) that ϵ_r is a complex quantity, given by Eq. (19), where Eq. (20) holds. Similarly χ_e is also complex, given by

$$\epsilon_r = \epsilon_r' - j\epsilon_r'' \tag{19}$$

$$\frac{\epsilon_r''}{\epsilon_r'} = \tan\delta \tag{20}$$

Eq. (21), where the real and imaginary parts of χ_e are related to those of ϵ_r by Eqs. (22) and (23). From this

$$\chi_e = \chi_e' - j\chi_e'' \tag{21}$$

$$\chi_e' = \epsilon_r' - 1 \tag{22}$$

$$\chi_e'' = \epsilon_r'' \tag{23}$$

it follows that J is given by Eq. (24) or (25), where E

$$J = 2\pi\epsilon_0\epsilon_r'' E^2 \qquad \text{joules/meter}^3 \tag{24}$$

$$J = 2\pi\epsilon_0\epsilon_r' E^2 \tan\delta \qquad \text{joules/meter}^3 \tag{25}$$

is the root-mean-square value of the applied field (in V/m), and ϵ_r'' and $\tan\delta$ are known as the loss factor and loss tangent respectively.

At optical frequencies it is usual to characterize the material by its complex refractive index $\hat{n}$, given by Eq. (26), where n is the real part of the refractive

$$\hat{n} = n - jk \tag{26}$$

index and k is the absorption coefficient. Refractive index $\hat{n}$ is related to ϵ_r through the relevant expressions for propagation velocity, so that $\epsilon_r = \hat{n}^2$, with Eqs. (27) and (28) resulting. The use of these rela-

$$\epsilon_r' = n^2 - k^2 \tag{27}$$

$$\epsilon_r'' = 2nk \tag{28}$$

tionships allows values of ϵ_r at high frequencies to be derived from optical measurements. For low k, $\epsilon_r' \simeq n^2$. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES.

The frequency variations of ϵ_r' and ϵ_r'' are not independent (**Fig. 1**). The drop in ϵ_r' between the effective beginning and end of a dispersion region, namely $\Delta\epsilon_r'$, is called the strength of the dispersion, and is given by the Kramers-Kronig relationship, Eq. (29).

$$\Delta\epsilon_r' = \frac{2}{\pi}\int_{-\infty}^{+\infty} \epsilon_r''\, d(\ln f) \tag{29}$$

If ϵ_r' and ϵ_r'' are plotted against $\log_{10} f$, the reduction in ϵ_r' for the dropping out of a given polarization mechanism is therefore about 1.466 times the area beneath the ϵ_r'' curve for that mechanism. In practice, the integral will be truncated to avoid the inclusion of area attributable to other mechanisms. *See* DISPERSION RELATIONS.

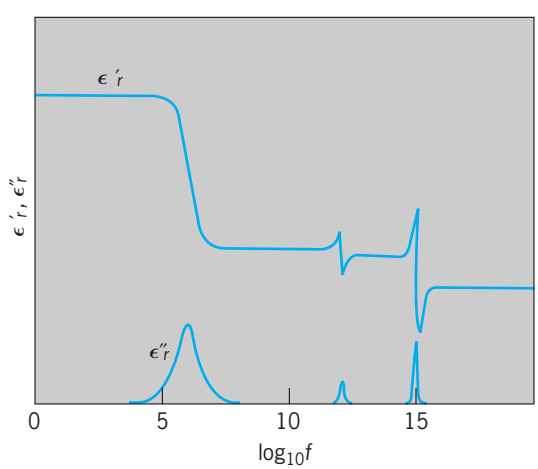

Fig. 1. Variation of ϵ_r' and ϵ_r'' with frequency f. Abscissa is $\log_{10} f$, where f is expressed in hertz.

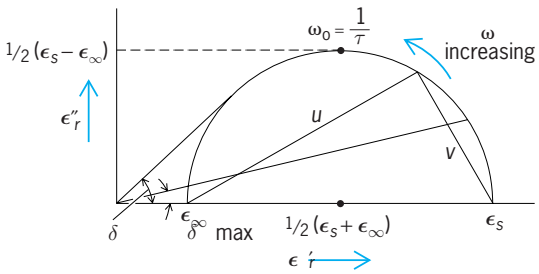

Fig. 2. Plot of ϵ''_r versus ϵ'_r for Debye behavior.

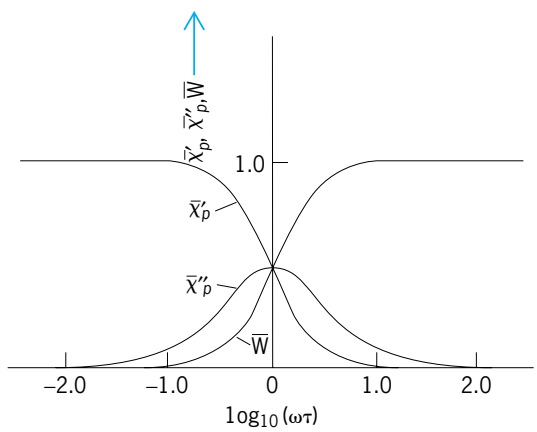

Fig. 3. Plot of $\overline{\chi}'_p = \chi'/\chi_{ps}$, $\overline{\chi}''_p = \chi''/\chi_{ps}$, and $\overline{W} = W\tau/\epsilon_0\chi_{ps}E^2$ for Debye behavior (χ_{ps} is the static orientational susceptibility).

In addition to the submicroscopic polarization mechanisms, the existence of macroscopic discontinuities at internal structure defects or at blocking electrodes, or the inclusion of impurities causing conducting regions, result in behavior classified as Maxwell-Wagner effects. These give rise to an effective polarization and associated loss, the frequency behavior (though not the temperature behavior) of which is similar to that of orientational polarization. The dispersion may lie in the region of 1 Hz or lower. *See* CRYSTAL DEFECTS.

Some dielectrics may exhibit an ohmic conductivity σ because they contain free charge carriers. This causes an energy loss $\sigma E^2/f$ per period. Comparison with Eq. (24) indicates that the associated apparent ϵ''_r is given by Eq. (30). Thus away from dispersion

$$\epsilon''_r = \frac{\sigma}{2\pi\epsilon_0 f} \tag{30}$$

regions, a curve of log ϵ''_r versus log f is a straight line of negative unit slope, with an intercept yielding σ. With this information, a correction can be made to measurements in the dispersion regions to give the value of ϵ''_r attributable to polarization processes alone.

Orientational polarization. When orientational polarization is operative, it is usually the dominant polarization present, and it normally gives rise to a stronger dispersion than do other mechanisms.

Debye theory. The classical theory of this mechanism is due to P. Debye, and for a single relaxation time τ, the variation of ϵ_r with angular frequency ω is given by Eq. (31) where ϵ_∞ and ϵ_s are the relative permittivities at frequencies much higher and much lower respectively than those of the dispersion region arising from the relaxation of permanent dipoles, but not so far removed that other dispersion mechanisms are operative between these frequency limits. Equating real and imaginary parts gives Eqs. (32) and (33).

$$\frac{\epsilon_r - \epsilon_\infty}{\epsilon_s - \epsilon_\infty} = \frac{1}{1 + j\omega\tau} \tag{31}$$

$$\frac{\epsilon'_r - \epsilon_\infty}{\epsilon_s - \epsilon_\infty} = \frac{1}{1 + \omega^2\tau^2} \tag{32}$$

$$\frac{\epsilon''_r}{\epsilon_s - \epsilon_\infty} = \frac{\omega\tau}{1 + \omega^2\tau^2} \tag{33}$$

See COMPLEX NUMBERS AND COMPLEX VARIABLES.

This shows that if ϵ''_r is plotted against ϵ'_r (**Fig. 2**), a semicircle of radius $^1/_2(\epsilon_s - \epsilon_\infty)$ results, with its center on the ϵ'_r axis at $^1/_2(\epsilon_s + \epsilon_\infty)$ making intercepts of ϵ_s and ϵ_∞ on that axis. The loss factor ϵ''_r has its maximum value when $\omega = \omega_0 = 1/\tau$, and according to Eq. (24) this corresponds to a maximum value of J given by Eq. (34). The maximum value of $\tan\delta$ occurs when $\omega\tau = (\epsilon_s/\epsilon_\infty)^{1/2}$ and is given by Eq. (35). For any

$$J_{\max} = \pi\epsilon_0(\epsilon_s - \epsilon_\infty)E^2 \qquad \text{joules/meter}^3 \tag{34}$$

$$\tan\delta_{\max} = {}^1/_2(\epsilon_s - \epsilon_\infty)(\epsilon_s\epsilon_\infty)^{-1/2} \tag{35}$$

point on the semicircle the frequency is given by Eq. (36), where u and v are the distances of the

$$\frac{\omega}{\omega_0} = \frac{v}{u} \tag{36}$$

point from the intercepts of the semicircle with the ϵ'_r axis, as shown in Fig. 2.

The power loss W associated with orientational polarization rises monotonically with increase in ω, and if plotted to a suitable scale against log $\omega\tau$ has the same shape as the χ'_p curve (where χ_p, as introduced above, is the component of the electric susceptibility that is due to orientational polarization), but reversed (**Fig. 3**). As $\omega \to \infty$, W rises

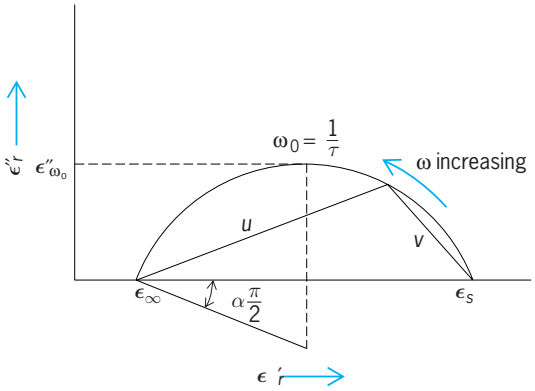

Fig. 4. Cole-Cole plot. Value of ϵ''_{ω_0} is given in the table.

Values of permittivity in the dispersion region for orientational polarization

Plot	$\frac{\epsilon_r - \epsilon_\infty}{\epsilon_s - \epsilon_\infty}$	Frequency ω	ϵ'_r	ϵ''_r
Debye	$\frac{1}{1 + j\omega\tau}$	$\omega_0 = \frac{1}{\tau}$	$\epsilon'_{\omega_0} = \frac{1}{2}(\epsilon_s + \epsilon_\infty)$	$\epsilon''_{\omega_0} = \frac{1}{2}(\epsilon_s - \epsilon_\infty)$
Cole-Cole	$\frac{1}{1 + (j\omega\tau)^{1-\alpha}}$	$\omega_0 = \frac{1}{\tau}$	$\epsilon'_{\omega_0} = \frac{1}{2}(\epsilon_s + \epsilon_\infty)$	$\epsilon''_{\omega_0} = \frac{1}{2}(\epsilon_s - \epsilon_\infty) \tan\frac{(1-\alpha)\pi}{4}$
Cole-Davidson	$\frac{1}{1 + (j\omega\tau)^2}$	$\omega_0 = \frac{1}{\tau}$	$\epsilon'_{\omega_0} = \epsilon_\infty + (\epsilon_s - \epsilon_\infty)2^{-\beta/2} \cos\frac{\pi\beta}{4}$	$\epsilon''_{\omega_0} = (\epsilon_s - \epsilon_\infty)2^{-\beta/2} \sin\frac{\pi\beta}{4}$
		$\omega_m = \frac{1}{\tau}\tan\frac{\pi}{2(1+\beta)}$	$\epsilon'_{\omega_m} = \epsilon_\infty + (\epsilon_s - \epsilon_\infty)\left[\cos\frac{\pi}{2(1+\beta)}\right]^\beta \cos\frac{\pi\beta}{2(1+\beta)}$	$\epsilon''_{\omega_0} = (\epsilon_s - \epsilon_\infty)\left[\cos\frac{\pi}{2(1+\beta)}\right]^\beta \sin\frac{\pi\beta}{2(1+\beta)}$

asymptotically to a theoretical value given by Eq. (37), which is twice the power loss at $\omega\tau = 1$.

$$W = \frac{\epsilon_0(\epsilon_s - \epsilon_\infty)E^2}{\tau} \qquad \text{watts/m}^3 \tag{37}$$

However, at very high frequencies, the Debye expression for ϵ''_r, and so Eq. (37) for W, may not hold. Various modifications have been suggested to make the calculation more satisfactory at the higher frequencies.

Cole-Cole plot. Frequently, experimental results yield a circular arc, rather than a semicircle, with its center below the abscissa. This modified plot (**Fig. 4**) is given by the Cole-Cole equation (38), where $0 < \alpha < 1$. The maximum value of ϵ''_r still occurs at $\omega = \omega_0 = 1/\tau$, but is reduced to ϵ''_{ω_0} (see **table**). For any point on the arc the frequency is given by Eq. (39).

$$\frac{\epsilon_r - \epsilon_\infty}{\epsilon_s - \epsilon_\infty} = \frac{1}{1 + (j\omega\tau)^{1-\alpha}} \tag{38}$$

$$\frac{\omega}{\omega_0} = \left(\frac{v}{u}\right)^{1/(1-\alpha)} \tag{39}$$

Such behavior can be predicted if a suitable distribution of relaxation times is postulated, as is also the case for the skewed-arc plot.

Cole-Davidson plot. A variety of other shapes are observed, such as the skewed arc which approximates to a straight line intersecting the ϵ'_r axis at an acute angle ξ at the high frequency end, curving round into a circular arc intersecting the ϵ'_r axis at right angles at the low frequency end (**Fig. 5**). Such a curve is represented by the Cole-Davidson equation (40),

$$\frac{\epsilon_r - \epsilon_\infty}{\epsilon_s - \epsilon_\infty} = \frac{1}{(1 + j\omega\tau)^\beta} \tag{40}$$

where $0 < \beta < 1$. The maximum value of ϵ'_r occurs not at ω_0 but at angular frequency ω_m given in the table, which also gives the values of ϵ'_r and ϵ''_r at ω_0 and ω_m.

Time domain. It follows from the foregoing that if a steady dc voltage were to be applied to a specimen exhibiting Debye behavior, or if the specimen were to be short-circuited after the polarization was fully established, then the circuit current would decay exponentially with a time constant τ. However, for many materials, particularly polymers, over a range of time the current decays as t^{-n}, a relationship known as the Curie–von Schweidler law. In the frequency domain, this corresponds to the type of linear relationship between ϵ'_r and ϵ''_r that is approached at the high-frequency end of the Cole-Davidson plot. The modifications to the Debye equation, devised to yield plots which fit such experimental observations, are sometimes criticized on the ground that they are not based on physical models. It has been suggested particularly with regard to materials that obey the Curie–von Schweidler law, that explanations should be sought not in terms of modifications to the Debye equation based on postulated distributions of relaxation times but in terms of quite different concepts arising directly from the linear relationships between ϵ'_r and ϵ''_r observed over quite wide ranges of frequency. However, a physical model based on protonic conduction has been put forward which yields a distribution of relaxation times from which such behavior can be predicted. James H. Calderwood

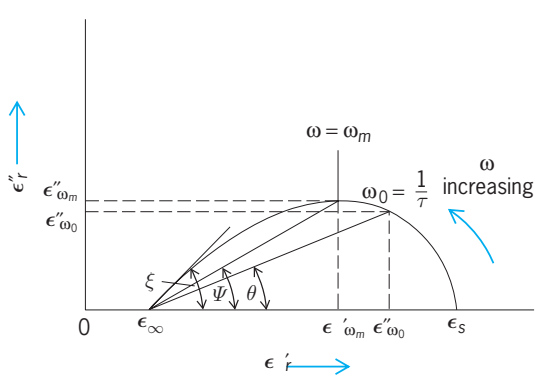

Fig. 5. Cole-Davidson plot. $\xi = (\pi/2)\beta$; $\psi = (\pi/2) \times [\beta/(1 + \beta)]$; $\theta = (\pi/4)\beta$. Values of ω_m, ϵ'_{ω_0}, ϵ''_{ω_0}, ϵ'_{ω_m}, and ϵ''_{ω_m} are given in the table.

Bibliography. C. J. F. Böttcher et al., *Theory of Electric Polarization*, 2d ed., vol. 1: *Dielectrics and Static Fields*, 1973, vol. 2: *Dielectrics in Time Dependent Fields*, 1978; W. T. Coffey et al., *Molecular Diffusion and Spectra*, 1984; P. Debye, *Polar Molecules*, 1929; V. I. Gaiduk and J. R. McConnell, *Dielectric Relaxation and Dynamics of Polar Molecules*, 1995; A. K. Jonscher, *Dielectric Relaxation in Solids*, 1983; H. Kliem, Dielectric

small signal response by protons in amorphous insulators, *IEE Trans. Elec. Insul.*, 24:185–197, 1989; B. K. P. Scaife, *Principles of Dielectrics*, 1989; A. von Hippel, *Dielectrics and Waves*, 2 vols., 1954.

Perovskite

A minor accessory mineral, formula $CaTiO_3$, occurring in basic rocks. Perovskite has given its name to a large family of materials, synthetic and natural, crystallizing in similar structures. The crystal structure is ideally cubic, with a framework of corner-sharing octahedra, containing titanium (Ti) or other relatively small cations surrounded by six oxygen (O) or fluorine (F) anions (see **illus.**). Within this framework are placed calcium (Ca) or other large cations, surrounded by twelve anions. Tilting of the octahedra and other distortions often lower the symmetry from cubic, giving the materials important ferroelectric properties and decreasing the coordination of the central cation. This flexibility gives the structure the ability to incorporate ions of different sizes and charges. Substitution of niobium (Nb), cerium (Ce), and other rare-earth elements in natural calcium titanate ($CaTiO_3$) is common and can make perovskite an ore for these elements. *See* COORDINATION CHEMISTRY; CRYSTAL STRUCTURE; RARE-EARTH ELEMENTS.

At pressures over 200,000 atm (20 gigapascals), reached in the Earth's lower mantle at depths below 700 km (420 mi), silicate minerals transform to a perovskite structure with a composition close to magnesium silicate ($MgSiO_3$). Synthesized in tiny quantities at great expense in the laboratory, this silicate perovskite may in fact be the most abundant single mineral on Earth. Its properties may relate to plate tectonics, seismic discontinuities, and the origin of deep earthquakes. *See* EARTHQUAKE; PLATE TECTONICS.

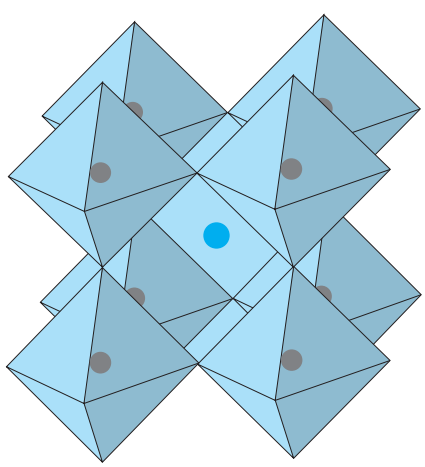

Ideal cubic perovskite structure: anions are at the corners of each octahedron, small cations (gray circles) are at the center of each octahedron surrounded by six anions, while the large cation (colored circle) is in the central cavity surrounded by twelve anions. (*After A. Navrotsky and D. J. Weidner, eds., Perovskite: A Structure of Great Interest to Geophysics and Materials Science, American Geophysical Union, 1989*)

A number of synthetic perovskites are of major technological importance. Barium titanate ($BaTiO_3$) and lead zirconate-titanate (PZT) ceramics form the basis of a sizable industry in ferroelectric and piezoelectric materials crucial to transducers, capacitors, and electronics. Lanthanum chromate ($LaCrO_3$) and related materials find applications in fuel cells and high-temperature electric heaters. *See* FUEL CELL.

It is anticipated that the high-temperature oxide superconductors, with critical temperatures above the boiling point of liquid nitrogen, will open up a variety of possible applications; they are largely based on the perovskite structure. In these, another fundamental feature of this structure plays a crucial role, namely its ability to incorporate large deviations from the ideal composition by cation or anion vacancies and by intergrowth, on a molecular scale, of other structural elements. One important family of oxide superconductors, the strontium–lanthanum–copper oxide ($Sr_xLa_{2-x}CuO_{4-y}$) type, is based on the potassium-nickel fluoride (K_2NiF_4) structure, which can be described as an intergrowth of perovskite and rock salt layers. Another group, the so-called 1:2:3 compounds like yttrium-barium-copper oxide ($YBa_2Cu_3O_{7-y}$), has complex structures derived from perovskite by the ordered removal of oxygen atoms, leaving both planes and chains of copper atoms. Superconductivity in both these types depends strongly on the oxygen content, which is controlled by preparation conditions and determines the formal oxidation state of copper. *See* CERAMICS; OXIDE; SOLID-STATE CHEMISTRY; SUPERCONDUCTIVITY.

Alexandra Navrotsky

Peroxide

A chemical compound that contains the peroxy (—O—O—) group. The simplest member of the series is hydrogen peroxide (HOOH), and the higher members of the series result from substituting a group for one or both of the hydrogen atoms in HOOH. Peroxidic compounds occur in nature, and they have commercial applications

Peroxidic compounds in nature. Both plants and animal cells reduce oxygen by one electron to form superoxide as a by-product of the reduction of oxygen to water, the process in which the cell burns foodstuff to make energy. Two superoxide radical ions react to give a dismutation reaction, that is, simultaneous oxidation and reduction, in which one ion is reduced to hydrogen peroxide and one is oxidized back to oxygen [O_2; reaction (1)]. In 1968

$$2O_2^{\bullet -} + 2H^+ \rightarrow H_2O_2 + O_2 \qquad (1)$$

J. McCord and I. Fridovich discovered the enzyme that catalyzes this reaction in cells, superoxide dismutase. Virtually all aerobically living cells contain some form of superoxide dismutase. Cells also contain enzymes called lipoxygenases, which catalyze the substitution of oxygen into the position next to a

double bond in unsaturated fatty acids. This position is called the allylic position, and the allylic hydrogen atom is particularly easily removed in free-radical reactions [reaction (2)].

$$\underset{\text{Lipid}}{LH} + O_2 \rightarrow \underset{\text{Lipid hydroperoxide}}{LOOH} \qquad (2)$$

See FREE RADICAL; EICOSANOIDS; ENZYME; SUPEROXIDE CHEMISTRY.

Lipid hydroperoxides can also be produced by the nonenzymatic oxidation of lipids, a process called autoxidation or lipid peroxidation. This is the process that causes butter and other fats to turn rancid when they are exposed to air. Lipid peroxidation involves an initiation step in which radicals are formed, followed by two propagation steps [reactions (3) and (4)] in which the number of radicals is conserved.

$$L^{\bullet} + O_2 \rightarrow LOO^{\bullet} \qquad (3)$$

$$LH + LOO^{\bullet} \rightarrow L^{\bullet} + LOOH \qquad (4)$$

See AUTOXIDATION.

A termination reaction then occurs in which two radicals combine and nonradical products are formed.

The reactions that cause fats such as butter to become rancid also occur in cells, causing the gradual deterioration of the fats. However, lipoxygenase enzymes also catalyze these processes, producing complex lipid hydroperoxides that play a role in many biological processes. For example, all of the prostaglandin and leukotriene hormones are produced from oxidation of the 20-carbon fatty acid that contains four double bonds, arachidonic acid, to the compound prostaglandin G, which contains both an endoperoxide bridge and a hydroperoxide bond [structure (**1**)].

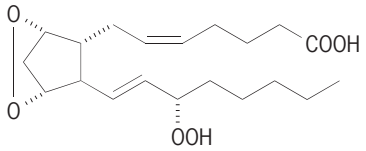

(**1**)

Nitric oxide (NO) is a hormone produced by a number of types of cells, including those in the cells of the artery lining (called endothelial cells), where its function is to reduce muscle tone and thus blood pressure. Compounds such as nitroglycerin release nitric oxide. Since nitric oxide is a free radical, it combines with superoxide [reaction (5)], which is

$$\bullet NO + O_2^{\bullet -} \rightarrow {}^{-}O{-}O{-}N{=}O \qquad (5)$$

(**2**)

produced by many of the same types of cells that release nitric oxide. This gives the peroxidic compound peroxynitrite (**2**). The related acid, peroxynitrous acid (**3**), a potent oxidant in the cell, is formed by the reaction of peroxynitrite with the hydrogen ion [reaction (6)].

$${}^{-}O{-}O{-}N{=}O + H^{+} \rightarrow HO{-}O{-}N{=}O \qquad (6)$$

(**3**)

Peroxides in commerce. In commerce, peroxides are used in oxidations, purifications, syntheses, polymerizations, and oxygen generation. Inorganic peroxides include persulfates, sodium peroxide, and bivalent metal peroxides. Organic peroxides include peroxyacetic acid, benzoyl peroxide, cumene peroxide, and a number of addition products of hydrogen peroxide and aldehydic compounds.

Inorganic peroxides. Peroxydisulfates, familiarly called persulfates, are produced by electrolytic oxidation of aqueous sulfuric acid or ammonium bisulfate. The potassium salt of persulfuric acid (KHS_2O_8) is used in the manufacture of styrene-butadiene synthetic rubber, and smaller quantities are used in hair bleaches. Peroxymonosulfates, also called monopersulfates, are salts of peroxymonosulfuric acid (Caro's acid, $HO{-}SO_2{-}OOH$); they are powerful oxidants. *See* OXIDIZING AGENT.

Hydrogen peroxide. The most widely used peroxy compound is hydrogen peroxide, a waterlike liquid manufactured as aqueous solutions. Hydrogen peroxide is not combustible; water is a safe diluent and coolant. With organic compounds, hydrogen peroxide can form detonable mixtures. Directions for safe handling and storage of hydrogen peroxide are available.

Hydrogen peroxide is manufactured by electrolytic and organic oxidation processes. The former involves electrolytic production of the peroxydisulfate intermediate, followed by steam hydrolysis to hydrogen peroxide. One organic process uses an anthraquinone that is dissolved in organic solvents. The quinone, similar to that shown in (**4**), is catalytically hydrogenated to the hydroquinone (**5**), which undergoes autoxidation. Isopropyl alcohol also can be oxidized at moderate temperatures and pressures to give hydrogen peroxide and acetone coproducts. *See* ELECTROCHEMICAL PROCESS.

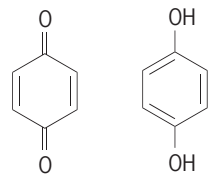

(**4**) (**5**)

Hydrogen peroxide applications include commercial bleaching, which consumes more than half of that produced. Applications include bleaching of practically all wool and cellulosic fibers, as well as of major quantities of synthetics, and paper and pulp mill bleaching of ground wood and chemical pulps.

Hydrogen peroxide also is used to synthesize epoxides and glycols from unsaturated petroleum hydrocarbons, terpenes, and natural fatty oils. The resultant products are plasticizers, stabilizers, diluents, and solvents for vinyl plastics and protective

coating formulations. Power generation applications include use in specialized propulsion units for aircraft, missiles, torpedoes, and submarines. The hot oxygen-system mixture from catalytically decomposed hydrogen peroxide powers the feed pumps of many large liquid-propellant rockets. *See* BLEACHING; HYDROGEN PEROXIDE; TURBINE ENGINE SUBSYSTEMS; ROCKET PROPULSION.

Organic peroxides. As a group, organic peroxides are more hazardous than the inorganic peroxides. Many of the organic compounds are flammable or detonable, thus restricting their availability. Manufacture is chiefly through reaction of the organic substrate with hydrogen peroxide. However, air oxidation also can sometimes be used.

Peracetic acid [$CH_3(C{=}O)OOH$] is prepared from acetic acid and hydrogen peroxide. Peracetic acid, as well as performic or perpropionic, may also be generated from hydrogen peroxide, organic acid, and catalyst. Major applications of peracetic acid are in the synthesis of epoxidized and hydroxylated compounds and as a bactericide, fungicide, and sterilizing agent for processing equipment.

Cumene hydroperoxide [$C_6H_5{-}C(CH_3)_2{-}OOH$] is another peroxide used in commerce. It is a colorless to pale-yellow liquid, produced by air oxidation of isopropyl benzene. It cleaves in acid solution to give phenol and acetone, and this reaction is one of the commercial syntheses of phenol.

Benzoyl peroxide. Benzoyl (or dibenzoyl) peroxide ($C_6H_5{-}CO{-}O{-}O{-}CO{-}C_6H_5$) is the most important aromatic acyl peroxide. It is a colorless, crystalline solid, melting point 106–108°C (223–226°F), that is virtually insoluble in water but very soluble in most organic solvents. It is stable at room temperature but explosible with heating or scratching when in the pure state. For this reason, it is shipped in paper barrels and plastic bags, rather than glass bottles, since crystals caught in the threads could explode when the cap is unscrewed. It is manufactured by reaction of benzoyl chloride and alkaline hydrogen peroxide. Major uses include polymer manufacture and flour bleaching. *See* POLYMERIZATION.

Benzoyl peroxide is synthesized from the reaction of benzoyl chloride with hydrogen peroxide in alkaline aqueous solution or with sodium peroxide in aqueous acetone. Recrystallization from hot solvents is hazardous; peroxides should be heated only in dilute solution. Benzoyl peroxide is an example of a peroxidic initiator, a material that decomposes at a controlled rate at moderate temperatures to give free radicals. Benzoyl peroxide has a half-life (10 h at 73°C or 163°F in benzene) for decomposition that is convenient for many laboratory and commercial processes. This, plus its relative stability among peroxides, makes it one of the most frequently used initiators, for example in vinyl polymerization. It also is the preferred bleaching agent for flour, and is used to bleach many commercial waxes and oils. Benzoyl peroxide is the active ingredient in commercial acne preparations.

Hazards. Benzoyl peroxide itself is neither a carcinogen nor a mutagen. However, when coapplied with carcinogens to mouse skin it is a potent promoter of tumor development. Many peroxidic compounds also are tumor promoters. All peroxides should be treated as potentially explosive, but benzoyl peroxide is one of the least prone to detonate. *See* MUTAGENS AND CARCINOGENS.

Reactions. Like all peroxides, benzoyl peroxide decomposes by rupture of its oxygen-oxygen peroxide bond to give two benzoyloxyl radicals. Thus, benzoyl peroxide provides a source of both phenyl and benzoyloxyl radicals that can perform chemical reactions. For example, the phenyl radicals can add to the double bond of vinyl chloride ($H_2C{=}CHCl$) to form a new radical and thus initiate the polymerization of the vinyl monomer.

Accelerated and induced decomposition. Many materials cause benzoyl peroxide to decompose at a rate accelerated over that observed in pure benzene. For example, solvents with high reactivity toward free radicals (such as dimethyl ether) cause a greatly accelerated decomposition. Compounds that can act as nucleophiles, reducing agents, or electron donors also can greatly accelerate the decomposition; examples are amines and sulfides. Transition-metal ions such as iron or copper also cause an increased rate of decomposition; this can be utilized in commercial practice to achieve a lower temperature or more rapid decomposition. *See* OXIDATION-REDUCTION. W. A. Pryor

Bibliography. F. A. Cotton and G. Wilkinson, *Advanced Inorganic Chemistry: A Comprehensive Text*, 6th ed., 1999; D. C. Nonhebel and J. C. Walton, *The Chemistry of Peroxides*, 1984; W. A. Pryor, *Free Radicals*, 1966; W. A. Pryor (ed.) *Free Radicals in Biology*, vols. 1–6, 1976–1983; D. Swern (ed.), *Organic Peroxides*, vol. 2, 1971; A. Wataru (ed.), *Organic Peroxides*, 1992.

Peroxisome

Peroxisomes are intracellular organelles that are found in all eukaryotes except the archezoa (original lifeforms). They are bounded by a single membrane and contain a granular matrix and sometimes a paracrystalline core. In electron micrographs, peroxisomes appear round with a diameter of 0.1–1.0 micrometer (see **illus.**), although there is evidence that in some mammalian tissues peroxisomes form an extensive reticulum (network). The name peroxisome derives from the fact that many enzymes, particularly the oxidases, that generate hydrogen peroxide, and catalase, the enzyme that catalyzes the dismutation of hydrogen peroxide to water and molecular oxygen, reside in the organelle. However, peroxisomes contain more than 50 characterized enzymes and perform many biochemical functions, including detoxification. *See* CELL ORGANIZATION.

Function. Peroxisomes are important for lipid metabolism. In humans, the β-oxidation of fatty acids greater than 18 carbons in length occurs in peroxisomes, while fatty acids of shorter length undergo β-oxidation primarily in mitochondria. In yeast, all fatty acid β-oxidation occurs in peroxisomes.

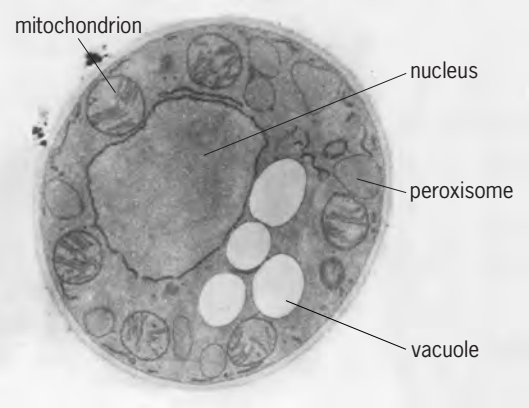

Electron micrograph of a cell of the yeast *Yarrowia lipolytica* grown in fatty acid–containing medium.

Peroxisomes contain the first two enzymes—alkyl dihydroxyacetonephosphate (DHAP) synthase and DHAP acyl transferase—required for the synthesis of plasmalogens (a group of phospholipids in which a fatty acid group is replaced by a fatty aldehyde). Peroxisomes also play important roles in cholesterol and bile acid synthesis, purine and polyamine catabolism, and prostaglandin metabolism. In plants, peroxisomes are required for photorespiration. *See* LIPID METABOLISM; PHOTORESPIRATION.

Peroxisomes are related to two other organelles, the glyoxysomes of plants and the glycosomes of certain parasites, particularly the trypanosomatids. Together these organelles constitute the microbody family of organelles. All three organelles contain enzymes involved in the β-oxidation of fatty acids. In addition, glyoxysomes harbor the enzymes of the glyoxylate pathway, while glycosomes contain a number of glycolytic enzymes. *See* ENZYME.

Peroxisomes show dramatic changes in number, volume, and protein composition in response to changes in environmental or physiological conditions. Various yeasts show one or few peroxisomes when grown on glucose, but peroxisomes readily proliferate and there is an increased overall peroxisomal protein content when yeasts are grown on carbon sources such as oleic acid and methanol (the metabolism of these sources requires peroxisomal activity). Similar peroxisomal multiplication effects have been observed in some mammals, particularly rodents, treated with peroxisome proliferative agents such as the fibrate family of hypolipidemic drugs (substances that produce a decrease in the level of lipids in the blood).

Human disorders. A number of recessively inherited peroxisomal disorders have been described and grouped into three categories. Group I is the most severe and is characterized by a general loss of peroxisomal function. Many of the enzymes normally localized to the peroxisome are instead found in the cytosol. Mature peroxisomes are absent in the cells of these patients, and instead peroxisomal ghosts are found. These ghosts do not contain matrix enzymes, but most do have a normal complement of peroxisomal membrane proteins. Among the diseases found in group I are Zellweger syndrome, neonatal adrenoleukodystrophy, and infantile Refsum disease. Patients with these disorders usually die within the first years after birth and exhibit neurological and hepatic (liver) dysfunction, along with craniofacial dysmorphism (malformation of the cranium and the face).

Group II peroxisomal disorders are characterized by a loss of peroxisomal function less severe than in group I. Examples are rhizomelic chondrodysplasia punctata, in which DHAP acyl transferase, alkyl DHAP synthase, phytanic acid oxidase, and thiolase are absent from peroxisomes, and Zellweger-like syndrome, in which DHAP acyl transferase and the enzymes of peroxisomal β-oxidation are absent.

Group III comprises those genetic diseases in which one peroxisomal enzyme is disfunctional. Examples are acatalasemia, X-linked adrenoleukodystrophy, and deficiencies of the individual enzymes of peroxisomal β-oxidation.

Peroxisome biogenesis. Peroxisomes do not contain deoxyribonucleic acid (DNA). They do not have independent protein synthesis machinery, as do mitochondria and chloroplasts. Accordingly, all peroxisomal proteins are encoded in the nucleus. Most peroxisomal proteins, both soluble and membrane-associated, are posttranslationally imported into peroxisomes. Import of proteins into peroxisomes requires energy, but it is uncertain whether there is a requirement for an electrochemical gradient across the peroxisomal membrane. Interestingly, peroxisomes are capable of oligomeric protein import. Peroxisomes can arise by growth and fission of preexisting peroxisomes; however, there is accumulating evidence for the de novo synthesis of peroxisomes from vesicular precursors.

Most peroxisomal proteins are synthesized without any precursor extensions or posttranslational modifications. Many matrix proteins are targeted to peroxisomes by a conserved tripeptide peroxisomal targeting signal, called PTS1, located at their extreme carboxyl termini. A small subset of matrix proteins, the best known being the β-oxidation enzyme 3-ketoacyl-CoA thiolase, are targeted by an aminoterminal PTS2. PTS2 motifs may or may not be cleaved from the newly synthesized protein within the peroxisomal matrix. Little is known about the signals that target proteins to the peroxisomal membrane.

Mutants of peroxisome assembly (*pex* mutants) have been made in a number of yeast species and in Chinese hamster ovary cells. These *pex* mutants mimic human peroxisomal biogenesis disorders, particularly Zellweger syndrome. Complementation of *pex* mutants has led to the identification of more than 20 genes whose proteins are required for peroxisome assembly. Pex proteins are localized both to the cytosol and to peroxisomes. Pex proteins that act as PTS receptors, as docking proteins for PTS receptors, and as regulators of peroxisome division have been identified. Nevertheless, the functions of many Pex proteins remain unknown. *PEX* genes isolated from yeast and rodents have been used to identify the corresponding human genes through sequence

homology. Many of these human *PEX* genes have been shown to be mutated in the various complementation groups of Zellweger syndrome. The gene encoding Pex7p, the PTS2 receptor, has been shown to be mutated in patients of rhizomelic chondrodysplasia punctata. The identification of human *PEX* genes by homology to *PEX* genes of yeast has been one of the triumphs of the use of simple model systems to identify human genes. *See* GENE; PROTEIN.

Richard A. Rachubinski

Bibliography. B. Alberta et al., *Molecular Biology of the Cell*, 3d ed., Garland Publishing, 1994; G. M. Cooper, *The Cell: A Molecular Approach*, ASM Press, 1997.

Peroxisome proliferator-activated receptor

A member of the nuclear hormone receptor superfamily of transcription factors, similar in both structure and function to the steroid, thyroid, and retinoid receptors. Peroxisome proliferator-activated receptors (PPARs) respond to specific factors by altering gene expression in a cell-, developmental-, sex-specific manner. PPARs were initially characterized as being activated by peroxisome proliferators (PPs), a large group of hypolipidemic agents and pollutants. Subsequently, it was shown that PPARs respond to endogenous fatty acids and control a variety of target genes involved in lipid homeostasis. Further, PPARs were shown to play a key role in the response to antidiabetic drugs. The multifaceted responses of PPARs are actually carried out by three subtypes of PPARs expressed in different tissues. Currently, the subfamily has been defined as PPARα, PPARβ (also called PPARδ and NUC1), and PPARγ, each with a possibility of different ligands, target genes, and biological roles see **illustration**. *See* GENE; LIPID; LIPID METABOLISM.

Ligands. Many peroxisome proliferators and fatty acids are PPAR activators. The xenobiotic peroxisome proliferators such as the fibrate hypolipidemic drugs, are structurally similar to endogenous and dietary fatty acids and their metabolites, containing a carboxylic acid functional group and a hydrophobic tail. Certain mono- and polyunsaturated fatty acids bind directly to PPARα and PPARγ at physiological concentrations. The eicosanoid 8(S)-hydroxyeicosatetraenoic acid is a PPARα ligand, while 15-deoxy-delta12,14-prostaglandin J2 (PGJ_2) specifically associates with PPARγ. The chemotactic agent leukotriene B4 and the hypolipidemic drug Wy 14,643 are reasonably specific for the PPARα subtype. A great deal of interest in PPARs has been initiated by the discovery of highly specific PPARγ ligands among the class of antidiabetic agents termed thiazolidinediones (TZDs). Many TZDs, including ciglitazone, pioglitazone, and BRL49653, are highly specific and potent activators of PPARγ.

Few high-affinity ligands for PPARβ are known, either xenobiotic or endogenous. However, fatty acids are weak activators of this receptor, with roughly the same preference as PPARα. In addition, certain prostaglandins (PGA1, PGD2, and PGD1) can activate PPARβ in reporter assays.

Target genes. Similar to other steroid hormone receptors, PPARs are ligand-activated transcription factors that control gene expression by interacting with specific DNA response elements (PPREs) located upstream of responsive genes. Interestingly, the PPRE site is virtually identical for the three PPAR subtypes; this signifies that much of the differences in function of the PPARs is derived from tissue-specific expression as well as subtype-specific activators. The motif recognized by PPAR contains two copies of a core consensus sequence AGGTCA with a one nucleotide spacer, termed a DR-1 site. A second nuclear receptor, retinoid-X-receptor (RXR), is required for PPAR to bind to DNA and regulate gene expression. There are several genes that contain this PPRE motif, and they are often referred to as target genes (see **table**). *See* DEOXYRIBONUCLEIC ACID (DNA).

The ultimate response of a cell to PPAR activators is the sum total of the genes being regulated in that cell. Both PPARα and PPARγ play key roles in regulating fatty acid metabolism, albeit in seemingly opposite directions (see table). The result of PPARα activation in hepatocytes and certain other tissues is an increase in the peroxisomal and mitochondrial oxidation of fatty acids. The array of genes regulated by PPARγ in adipocytes is indicative of fatty acid accumulation. This regulation of gene expression is concomitant with increased differentiation of immature adipocytes into mature fat-storing cells. The key genes regulated by PPARγ that are associated with diabetes have not been conclusively demonstrated, but may involve adipocyte-secreted cytokines and hormones such as TNF-α and leptin.

The genes regulated by PPARγ in adipocytes are similar to those in the macrophage and include lipoprotein lipase and CD36. Treatment of macrophages with PPARγ synthetic agonists inhibits the production of several cytokines such as interleukin 1-β and TNF-α and may result in an anti-inflammatory response. Another link between PPARγ and inflammation is the fact that 15-deoxy PGJ2, a product of the cyclooxygenase pathway, and nonsteroidal anti-inflammatory drugs (NSAIDS) are potent activators of PPARγ.

Biological role. The potency of various chemicals to activate PPARs is subtype-specific, and the expression of PPARα, β, and γ varies widely from tissue to tissue. In numerous cell types from ectodermal, mesodermal, or endodermal origin, PPARs are coexpressed, although their concentration relative to each other varies widely. PPARα is highly expressed in hepatocytes (liver cells), cardiomyocytes (heart cells), enterocytes (intestinal cells), and the proximal tubule cells of kidney. PPARβ is expressed ubiquitously and often at higher levels than PPARα and γ. PPARγ is expressed predominantly in adipose tissue, colon epithelium, and macrophages and exists as two distinct forms, $\gamma 1$ and $\gamma 2$, which arise by differential transcription start sites and alternative splicing. The distinct tissue distribution suggests that the PPAR subtypes play different biological roles. In fact, of

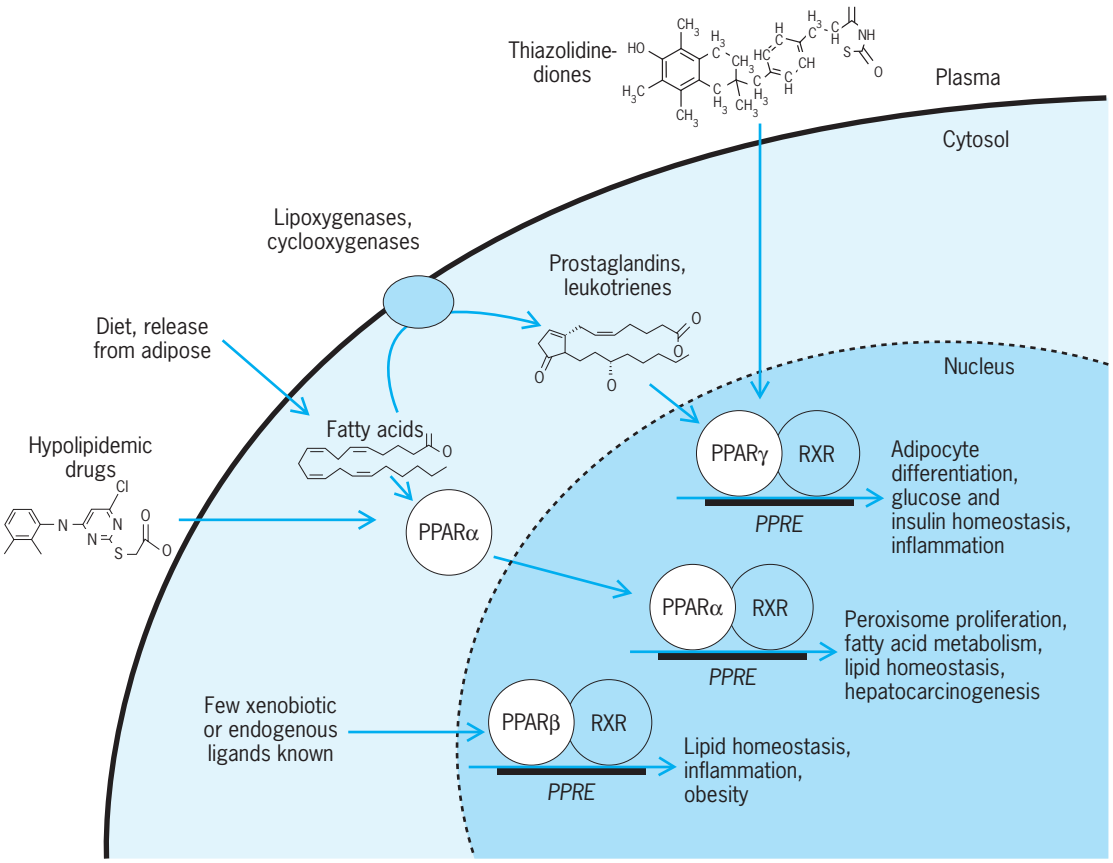

Basic mechanism of action of PPARα, β, and γ and effects on gene expression. Similar to other nuclear hormone receptors, PPAR acts as a ligand-activated transcription factor. Upon binding fatty acids or hypolipidemic drugs, PPARα interacts with RXR and regulates the expression of target genes. These genes are involved in the catabolism of fatty acids. Conversely, PPARγ is activated by prostaglandins, leukotrienes, and antidiabetic thiazolidinediones and affects the expression of genes involved in the storage of fatty acids. Although few xenobiotic or endogenous activators of PPARβ are known, this subtype appears to regulate genes important for lipid metabolism and for inflammation, and may be a target for antiobesity drugs.

the different subtypes, PPARα predominates in hepatic lipid metabolism and PPARγ plays a pivotal role in adipogenesis and immune responses. There is also evidence to support the role of PPARs in cell growth and differentiation.

PPARs, in particular PPARγ, have been clearly established to be involved in differentiation of several cell types. The wide variety of cells that can be induced to differentiate with peroxisome proliferators, fatty acids, thiazolidinedione, and other PPAR ligands suggest that this subfamily of proteins are "master regulators" of differentiation. This classification is reserved for genes that specify the fate of a particular cell, such as transcription factors capable of activating the program of differentiation. PPARs have been shown to induce differentiation of adipocytes, oligodendrocytes, myoblasts, keratinocytes, and monocyte/macrophages. Ectopic expression of PPARγ in fibroblasts regulated development of the adipose lineage in response to endogenous lipid activators. That is, expression and activation of PPARγ is sufficient

Summary of genes regulated by PPARα and PPARγ

PPARα-regulated genes in liver	Function	PPARγ-regulated genes in adipose	Function
Acyl-CoA oxidase	Fatty acid oxidation	aP2	Fatty acid uptake
Cytochrome P450 4A	Fatty acid oxidation	Lipoprotein lipase	Fatty acid uptake
Enoyl-CoA hydratase	Fatty acid oxidation	Fatty acid transporter protein	Fatty acid uptake
Peroxisomal thiolase	Fatty acid oxidation	Acyl-CoA synthetase	Fatty acid esterification
Acyl-CoA binding protein	Fatty acid transport	Phosphoenolpyruvate carboxy kinase	Glyceroneogenesis
Carnitine acyltransferase	Fatty acid transport	Malic enzyme	Lipogenesis
Fatty acid binding protein	Fatty acid transport		
Acyl-CoA synthetase	Fatty acid esterification		
Lipoprotein lipase	Fatty acid uptake		
Apolipoprotein AI	Lipid export		
c-*myc*	Growth regulation		
ZFP-37	Growth regulation		

to result in a phenotypic change in fibroblasts. Also, ligand activation of PPARγ is sufficient to induce growth arrest in fibroblasts, a key step in committing a cell to differentiate. *See* CELL DIFFERENTIATION.

In contrast to PPARγ, PPARα may play a larger role in cellular proliferation than in differentiation. Activators of PPARα have long been known to be rodent liver tumor promoters, and it is generally held that altered gene expression is the key. Support for this theory includes the induction of growth regulatory genes such as c-*myc*, c-Ha-*ras*, *fos*, *jun*, *egr-1*, and *ZFP-37* by peroxisome proliferators. The most convincing data regarding a PPAR–cell proliferation–tumor promotion connection come from the PPARα-null mouse model system. Remarkably, the mice that lack PPARα do not display the typical pleiotropic response when challenged with the peroxisome proliferators, such as hepatomegaly, peroxisome proliferation, and transcriptional activation of target genes. These mice display abnormal lipid homeostasis, including fatty acid metabolism. Importantly, in PPARα-null mice fed a peroxisome proliferator in the diet, there was no increase in hepatic cell proliferation or hepatocellular neoplasms, whereas all of the wild-type mice developed tumors.

The biological role of PPARβ has not been clearly established, although the development of high-affinity ligands as well as the use of the PPARβ-null mouse model has shed some light on the subject. Addition of PPARβ ligands to obese mice results in an improvement of metabolic parameters and reduction in adiposity. Also, the absence of PPARβ (null mouse) results in altered response to xenobiotics, in particular in the skin and immune system.

John P. Vanden Heuvel

Bibliography. P. Escher and W. Wahli, Peroxisome proliferator-activated receptors: Insight into multiple cellular functions, *Mutat. Res.*, 448:121–138, 2000; S. Kersten, B. Desvergne, and W. Wahli, Roles of PPARs in health and disease, *Nature*, 405:421–424, 2000; J. M. Peters and J. P. Vanden Heuvel, Peroxisomes, peroxisome proliferators and peroxisome proliferator-activated receptors (PPARs), in J. P. Vanden Heuvel et al. (eds.), *Cellular and Molecular Toxicology*, vol. 14, pp. 133–158, Elsevier Science, Amsterdam, 2002; B. M. Spiegelman, PPAR-gamma: Adipogenic regulator and thiazolidinedione receptor, *Diabetes*, 47:507–514, 1998; J. P. Vanden Heuvel, Diet, fatty acids, and regulation of genes important for heart disease, *Curr. Atheroscler. Rep.*, 6:432–440, 2004; T. M. Willson et al., The PPARs: From orphan receptors to drug discovery, *J. Med. Chem.*, 43:527–550, 2000.

Peroxynitrite

A nitrogen oxyanion containing an O—O peroxo bond that is a structural isomer of the nitrate ion. These species are generally distinguished as $ONOO^-$ and NO_3^-, respectively. Other names for peroxynitrite include pernitrite and peroxonitrite; the systematic name recommended by the International Union of Pure and Applied Chemistry (IUPAC) is oxoperoxonitrate (1−). Energy calculations indicate that there are two stable conformations of $ONOO^-$, for which all of the atoms lie in a plane with the peroxo O—O and N═O bonds forming dihedral angles of approximately 0° (cis isomer) or 180° (trans isomer):

See CHEMICAL BONDING.

Characteristics. Although they are nearly isoenergetic, the cis isomer appears to be slightly more stable than the trans isomer. Peroxynitrite is a strong base and is protonated in weakly acidic solutions, forming peroxynitrous acid (ONOOH; IUPAC name: hydrogen oxoperoxynitrate). Calculations indicate that the gas-phase molecule also possesses stable cis and trans conformations; in this case, the energy of the cis isomer is further stabilized relative to the trans by an internal hydrogen bond between the bound hydrogen atom and the N═O terminal oxygen atom. In the trans isomer, the peroxo O—H bond is almost perpendicular to the plane of the other atoms. As expected from the high stability of the nitrogen dioxide ($^{\bullet}NO_2$) radical relative to the hydroxyl radical ($^{\bullet}OH$), the peroxo O—O bond in ONOOH is very weak. Its bond dissociation energy is estimated at ~22 kcal/mol compared to ~51 kcal/mol for HOOH. The acid dissociation constant of ONOOH, defined by the equilibrium $ONOOH \rightleftharpoons ONOO^- + H^+$, is $\sim 2 \times 10^{-7}$ *M*. This value is unusually high for a peroxo acid; the relatively weak O—H bond is attributed to the strongly electron-withdrawing character of the N═O substituent group. One consequence of the relatively high acidity of ONOOH is that, unlike hydrogen peroxide and organic peroxides, the anion is a major form in neutral solutions. *See* ACID AND BASE; HYDROGEN BOND; HYDROXYL.

Formation. Peroxynitrite is formed in nitrate salts or nitrate-containing solutions when exposed to ionizing radiation or ultraviolet light. Solutions can also be prepared by a variety of chemical reactions, including the reaction of hydrogen peroxide with nitrous acid (1); reaction of the hydroperoxide anion

$$HOOH + HNO_2 \rightarrow ONOOH + H_2O \quad (1)$$

with organic and inorganic nitrosating agents (2); reaction of ozone with the azide ion (3); or, ap-

$$HOO^- + RONO \rightarrow ROH + ONOO^- \quad (2)$$

$$2O_3 + N_3^- \rightarrow ONOO^- + N_2O + O_2 \quad (3)$$

parently, reaction of O_2 with compounds capable of generating the nitroxyl anion (NO^-) [4]. These

$$2O_2 + NH_2OH + 2OH^- \rightarrow ONOO^- + HO_2^- + 2H_2O \quad (4)$$

preparations invariably contain unreacted materials or decomposition products, particularly nitrite ion, which can significantly modulate the peroxynitrite chemical reactivity. Peroxynitrite is also formed in

radical-radical coupling reactions, notably superoxide ($^{\bullet}O_2^-$) with nitric oxide ($^{\bullet}NO$) [reaction (5)], and hydroxyl radical with nitrogen dioxide [reaction (6)].

$$^{\bullet}NO + {}^{\bullet}O_2^- \rightarrow ONOO^- \quad (5)$$

$$^{\bullet}OH + {}^{\bullet}NO_2 \rightarrow ONOOH \quad (6)$$

See SUPEROXIDE CHEMISTRY.

Peroxynitrite has been isolated as the tetramethylammonium salt by carrying out reaction (5) in liquid ammonia. Formation of peroxynitrite in both solids and solutions is indicated by the appearance of yellow coloration, which is due to tailing of intense near-ultraviolet absorption bands into the visible region.

Reactions. The peroxynitrite anion is relatively stable in both solid matrices and aqueous solution; it has, for example, been found in soil samples recovered from the Martian landscape. Peroxynitrous acid, however, undergoes spontaneous decomposition to give nitric acid and/or nitrous acid plus O_2. Both the decomposition rate and product distributions depend upon the medium conditions. Below pH 5, decomposition is first-order and only nitric acid is formed; the ONOOH lifetime is ~1 s at room temperature. Although, in principle, the reaction could occur by a concerted intramolecular rearrangement, theoretical calculations suggest that the lowest-energy pathways involves homolytic cleavage of the O—O bond to form a hydrogen-bonded {$^{\bullet}NO_2$, $^{\bullet}OH$} radical pair that subsequently recombines with rearrangement to nitric acid. This conclusion is supported by kinetic studies made in alkaline solutions, where the decomposition rate slows progressively with increasing solution alkalinity and is accompanied by formation of increasing amounts of O_2 and NO_2^- in 1:2 proportions. In this reaction domain, the O_2 yield exhibits an exceptionally complex dependence upon added radical scavengers that is quantitatively predicted by a dynamical model based upon known radical reactions. Crucial steps include dissociation of ONOOH to $^{\bullet}NO_2$ and $^{\bullet}OH$ radicals, dissociation of $ONOO^-$ to $^{\bullet}NO$ and $^{\bullet}O_2^-$ radicals, oxidation of $ONOO^-$ by $^{\bullet}OH$, and a radical chain cycle involving dinitrogen trioxide (N_2O_3). *See* CHEMICAL DYNAMICS.

Decomposition. Peroxynitrite decomposition is catalyzed by other Lewis acids, including metal ions and carbon dioxide. Higher-valent oxo ions have been detected as intermediates in decompositions catalyzed by metalloporphyrins, indicating that these reactions proceed by oxidation-reduction cycling of the metal ion. The reactant species in the carbon dioxide–catalyzed decomposition are $ONOO^-$ and CO_2, which apparently combine to form an adduct called nitrosoperoxycarbonate ($ONOOCO_2^-$; IUPAC name: 1-carboxylato-2-nitrosodioxidane). Because association is the rate-limiting step in this reaction, the intermediate does not accumulate, and its properties must be inferred by indirect means. Kinetic studies have established that the intermediate lifetime is shorter than ~1 ms and that CO_2 acts as a true catalyst; that is, it is the immediate product of the decomposition. However, when NO_2^- or other reductants are present, alternate decomposition pathways are expressed that lead to conversion of carbon dioxide to bicarbonate ion (HCO_3^-). It is therefore possible under some reaction conditions for dehydration of HCO_3^- to CO_2 to become rate-limiting in the catalytic decomposition cycle. In the lowest-energy configuration calculated for nitrosoperoxycarbonate, the —NO and —CO_2 substituents are cis about the peroxy O—O bond and oriented perpendicular to each other. The peroxy bond is even weaker than in ONOOH, having a calculated bond dissociation energy of ~9 kcal/mol. This difference can be attributed to the greater stability of the carbonate radical ($^{\bullet}CO_3^-$) formed upon O—O bond homolysis of $ONOOCO_2^-$ than the hydroxyl radical formed upon O—O bond homolysis of ONOOH. Calculations suggest that, as with ONOOH, the decomposition of $ONOOCO_2^-$ to NO_3^- and CO_2 involves homolytic O—O bond cleavage to form an intermediary {$^{\bullet}NO_2$, $^{\bullet}CO_3^-$} radical pair. *See* COMPUTATIONAL CHEMISTRY.

Oxidation. Peroxynitrite is a powerful oxidant that has been shown to react with a wide variety of inorganic and organic reductants, as well as hydroxylate and nitrate aromatic compounds, including benzene. Its pH-dependent one-electron standard reduction potential is estimated to be ~2.1 V (pH 0). Nitrosoperoxycarbonate is also a strong oxidant that is capable of oxidizing compounds whose reduction potentials are as high as ~1.3 V. It nitrates phenolic compounds more effectively than ONOOH, but appears unreactive toward benzene and other simple aromatics. Kinetic studies have demonstrated that, unlike most nitration reactions, conversion of tyrosine to 3-nitrotyrosine does not involve electrophilic attack of the aromatic ring by an intermediary nitronium (NO_2^+) ion, but probably proceeds via one-electron oxidation of the phenol to the corresponding phenoxy radical, followed by radical coupling to the $^{\bullet}NO_2$ generated in the initial redox step. Metal-catalyzed phenol nitrations by ONOOH are thought to occur by similar redox mechanisms.

Peroxynitrite and nitrosoperoxycarbonate can engage in both one-electron and two-electron oxidations of reacting partners. Two-electron oxidations are generally limited to good nucleophiles that have highly polarizable atoms or functional units; for ONOOH, these reactions occur by direct bimolecular interaction of the reactants. In contrast, most (if not all) one-electron oxidations occur by rate-limiting unimolecular activation of ONOOH with rate parameters that are identical to those for its decomposition. Furthermore, the maximal product yields obtained for reactions of both ONOOH and $ONOOCO_2^-$ are always substantially less than limits based upon stoichiometric consumption of the oxidant. This behavior indicates the existence of at least two intermediates along the reaction pathway, only one of which is oxidizing. Two fundamentally different hypotheses concerning the nature of these

intermediates have been advanced. One is that the unreactive and reactive species are different configurational isomers similar to the cis-trans isomers of ONOOH; the other is that the unreactive species are the {$^{\bullet}NO_2$,$^{\bullet}OH$} and {$^{\bullet}NO_2$,$^{\bullet}CO_3^-$} radical pairs that either rapidly recombine to form NO_3^- or escape the cage to give as reactive species discrete $^{\bullet}NO_2$ and $^{\bullet}OH$ (or $^{\bullet}CO_3^-$) radicals. Although this issue is unresolved, the dynamical behavior of the reactive species indicates that they have reactivities that are comparable to free $^{\bullet}OH$ and $^{\bullet}CO_3^-$ radicals. *See* OXIDATION-REDUCTION; REACTIVE INTERMEDIATES.

Physiological aspects. Interest in these reactions has been greatly stimulated by recognition that $^{\bullet}NO$ and $^{\bullet}O_2^-$ radicals are generated in the bloodstream, neuronal tissues, and phagocytic cells of animals in sufficient quantities to form peroxynitrite [reaction (5)]. Correspondingly, major roles for this powerful oxidant have been proposed both in diseases and tissue damage associated with oxidative stress and in natural cellular defense mechanisms against microbial infection. Indirect evidence that endogenously generated peroxynitrite might be involved in cellular damage includes the observations that tyrosyl nitration levels in proteins from affected tissues of patients suffering from oxidative diseases are exceptionally high, that tissues can be protected from oxidative damage by treatments that inhibit $^{\bullet}NO$ formation, and that ONOOH is highly toxic to cells and can cause substantial damage to proteins, carbohydrates, nucleic acids, membrane lipids, and cellular antioxidant systems. Physiological buffer systems contain high concentrations of CO_2; consequently, a major fraction of the peroxynitrite formed by living systems is expected to be converted to nitrosoperoxycarbonate before reacting with biological components. Because $ONOOCO_2^-$ has a considerably shorter lifetime than ONOOH, it is unclear whether its formation will act to protect cells from oxidative damage by promoting peroxide decomposition or will potentiate the damage by facilitating reactions with biological components.

Several additional factors that make it difficult to assess the contributions of these peroxy compounds to cellular oxidative processes are that ONOOH reacts rapidly with oxyhemoglobin and peroxidases (which could thereby constitute major sinks for its oxidizing equivalents), that ONOOH is freely diffusible across biological membranes (and therefore not confined to a particular intracellular environment), and that peroxidases also catalyze formation of other reactive nitrogen species (that might account, at least in part, for the elevated levels of tyrosyl group nitration detected at sites of oxidative damage). *See* BIOINORGANIC CHEMISTRY.

James K. Hurst

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; J. Loscalzo and J. A. Vita (eds.), *Nitric Oxide and the Cardiovascular System*, 2000; Z. Rappoport (ed.), *The Chemistry of Peroxides*, 2006.

Perpetual motion

The expression perpetual motion, or perpetuum mobile, arose historically in connection with the quest for a mechanism which, once set in motion, would continue to do useful work without an external source of energy or which would produce more energy than it absorbed in a cycle of operation. This type of motion, now called perpetual motion of the first kind, involves only one of the three distinct concepts presently associated with the idea of perpetual motion.

First kind. Perpetual motion of this type refers to a mechanism whose efficiency exceeds 100%. Clearly such a mechanism violates the now firmly established principle of conservation of energy, in particular that statement of the principle of conservation of energy embodied in the first law of thermodynamics. (Indeed, the first law of thermodynamics is sometimes stated as "A perpetuum mobile of the first kind cannot exist.") Thus, with the establishment of the energy conservation principle in the middle of the nineteenth century, the possibility of obtaining perpetual motion of the first kind could be denied. However, prior to that time, some ingenious machines, generally dependent on gravity for their "operation," had been devised. One of the most famous of these is the wheel of the Second Marquis of Worcester (**Fig. 1**). *See* CONSERVATION OF ENERGY.

The original model is described as having an outer rim 14 ft (4.3 m) in diameter, a concentric inner rim 12 ft (3.7 m) in diameter, and 40 spokes. Forty 50-lb (23-kg) weights are fastened at the center of ropes which are 2 ft (0.6 m) long. The weights are then attached to the wheel in the following fashion. One end of a given rope is secured to spoke *a*, say at the outer rim, and the other end of this rope is fastened to the adjacent spoke in a clockwise direction, spoke *b*, at the inner rim. All weights are attached in this way. The actual orientation of the weights

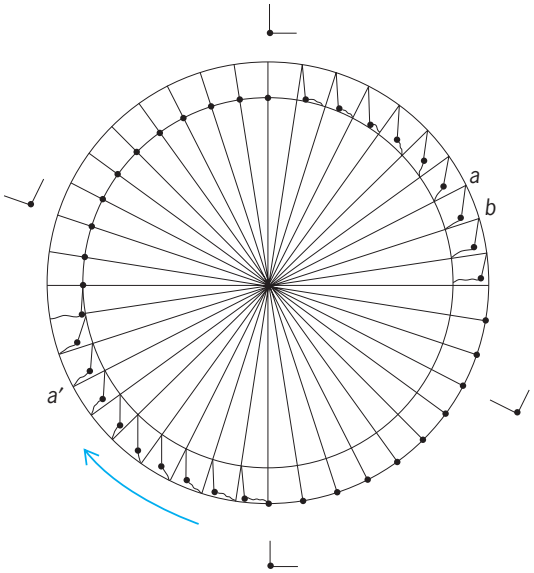

Fig. 1. Wheel of the Marquis of Worcester.

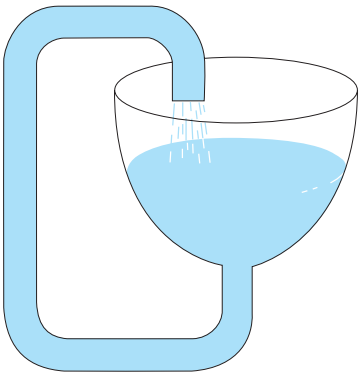

Fig. 2. Perpetual-motion device of the first kind based on the hydrostatic paradox.

is indicated both on the wheel in Fig. 1 and in the small sketches appearing outside the periphery of the wheel. To assess the motional tendency of the wheel, consider a pair of spokes making up a wheel diameter, say the pair *a*-*a*′. The weight supported by spoke *a* is supported at the outer rim, whereas the weight supported by spoke *a*′ is supported at the inner rim. Thus there is a clockwise torque associated with the spoke pair *a*-*a*′. Since, from Fig. 1, this appears to be the case for all spoke pairs that make up wheel diameters, the wheel presumably rotates clockwise perpetually. After describing the wheel and commenting upon the torque imbalance, the Marquis enigmatically remarks, "Be pleased to judge the consequence."

Also in the category of perpetual motion of the first kind were the proposals for hydrodynamical devices, essentially waterwheels which could resupply their own millstreams. These devices, many of extreme complexity, usually involved the hydrostatic paradox. In the example in **Fig. 2**, the water would supposedly flow continuously because its weight in the large vessel exceeds that in the tube.

Second kind. Perpetual motion of the second kind refers to a device that extracts heat from a source and then converts this heat completely into other forms of energy, a process which satisfies the principle of conservation of energy. A dramatic scheme of this type would be an ocean liner, which extracts heat from the nearly limitless oceanic source and then uses this heat for propulsion. This type of perpetual motion is, however, precluded by the second law of thermodynamics, also established in the nineteenth century, which is sometimes stated as "A perpetuum mobile of the second kind cannot exist."

Third kind. The third type of perpetual motion is, in contrast to the two types described above wherein useful output was the goal, merely a device which can continue moving forever. It could result in actual systems if all mechanisms by which energy is dissipated could be eliminated. For example, if all bearing friction could be removed, a wheel spinning in a vacuum would continue to spin indefinitely. Since experience indicates that dissipative effects in mechanical systems can be reduced, by lubrication in the case of friction, for example, but not eliminated, mechanical perpetual motion of the third kind can be approximated but never achieved.

An example of a genuine case of this kind occurs in a superconductor. A superconductor is a metal which, if cooled to a very low temperature, loses all of its resistance to the passage of a direct current; that is, a dissipative effect has been truly eliminated rather than reduced. If a direct current is caused to flow in a superconducting ring (**Fig. 3**), this current will continue to flow undiminished in time without application of any external force. *See* SUPERCONDUCTIVITY.

This last example leads perilously close to the microscopic realm where, of course, perpetual motion is the rule. Electrons in atoms and atoms themselves are in constant motion, this latter motion manifesting itself, for example, in brownian motion. The existence of this microscopic perpetual motion does not, in any event, force a reevaluation of the conclusions above, for the traditional ideas of perpetual motion deal with the macroscopic world of machines and devices. Since the laws of thermodynamics are themselves applicable to matter in macroscopic quantities and, as such, contain implicitly the consequences of the microscopic behavior, the impossibility of the useful types of perpetual motion, that is, of the first or second kinds, is at present deemed an unequivocal conclusion. *See* BROWNIAN MOVEMENT; THERMODYNAMIC PRINCIPLES.

Modern examples. Designs and working models of devices promising perpetual motion, in particular of the first kind, continue to be described frequently. Since the impossibility of perpetual motion of the first or second kind is now commonly accepted, the inventors of these devices are usually quick to dissociate their inventions from perpetual motion and to attribute device operation to "new (and not understood) principles." Though mechanical systems continue to be popular in this realm, the new breed of device is more likely to feature an operating system based on electromagnetic concepts. Particularly popular are "motors" incorporating complex and often ingenious configurations of permanent magnets, which are claimed to operate indefinitely with no power supplied. That this type of "motor" can operate, even for extended periods, need not be a contradiction of any of the known laws of physics. However, the operation is limited by the energy available,

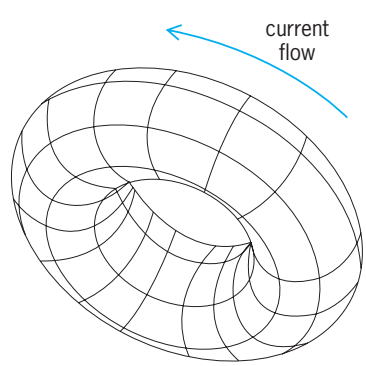

Fig. 3. Perpetual motion of the third kind.

namely, the energy contained in the magnetic field established by the configuration of magnets. Thus, the dream of perpetual motion continues unrealized and, in the view of essentially all scientists, unrealizable. K. L. Kliewer

Bibliography. A. W. Ord-Hume, *Perpetual Motion: The History of an Obsession*, 1977; K. S. Pitzer and L. Brewer, *Thermodynamics*, 3d ed., 1995; M. Tinkham, *Introduction to Superconductivity*, 1975, reprint 1980.

Perseus

A northern constellation named for a Greek hero (see **illustration**). Perseus and adjacent Cassiopeia are prominently located in the northern Milky Way. In Johan Bayer's (1603) star atlas, the bright star Mirfak marks Perseus' shoulder. *See* CASSIOPEIA.

In Greek mythology, this son of Zeus and Danae killed the gorgon Medusa, who was dangerous to look at, by looking only at her reflection in his shield and using winged sandals he got from Hermes. Perseus went on to use Medusa's terrifying head to kill Cetus, the sea monster, who was besieging Andromeda. In Johannes Bode's star atlas (1801), Perseus is carrying Medusa's head. *See* ANDROMEDA.

The double cluster, a pair of star clusters known as *h* and χ Persei, are fine objects for small telescopes. Algol, an eclipsing binary star known as the Demon Star, shows a cycle of regular dimming by a factor of 6 with a period of 2 days 21 hours on a predictable schedule. *See* BINARY STAR; ECLIPSING BINARY STAR; STAR CLUSTERS.

Each August 12, the Perseid meteor shower appears to come from this part of the sky. *See* METEOR.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION. Jay M. Pasachoff

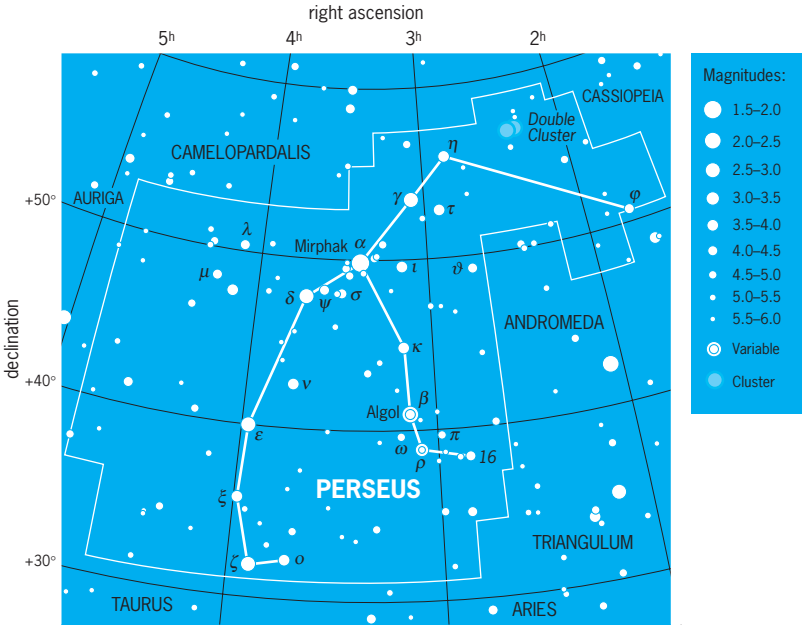

Modern boundaries of the constellation Perseus, a Greek hero. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

Persian melon

A long-season cultivar of muskmelon, *Cucumis melo*, of the gourd family, Cucurbitaceae. The fruit is 7–10 in. (18–25 cm) in diameter, weighs 6–8 lb (2.7–3.6 kg), and is round and without sutures; it has dark-green skin and thin, abundant netting. The flesh is deep orange, very thick and firm, and distinctly sweet in flavor. The fruit separates from the stem at full maturity and soon becomes overripe. Plants are large and vigorous with large leaves. The vines usually bear andromonoecious flowers, and pollination is performed by bees. Persian melons require a long frost-free growing season of about 115 days. Average growing temperatures of 70–75°F (21–24°C) are preferred. Fruits are about 90% water and contain 7.5 g of carbohydrate per 100 g of tissue. The flesh is very rich in potassium, vitamin A, and vitamin C. *See* VIOLALES.

In the United States the Persian melon is a very minor crop and is grown chiefly in the Central Valley of California, where low humidity and absence of summer rain tend to prevent the fungus diseases that often defoliate plants in humid areas. It is a luxury product that is very susceptible to damage during transit and storage.

Important foliar diseases are powdery and downy mildews, which are intensified by high humidity. Sulfur dusts are effective for powdery mildew control. Fusarium and Verticillium wilts inhabit the soil and cause wilting of the foliage and early death. Watermelon mosaic virus, cucumber mosaic virus, and squash mosaic virus are transmitted by insects and cause severe losses. Aphids, cucumber beetles, leaf miners, and red spider mites are the major insect pests. Nematodes attack the roots of the plants grown on infested soils. *See* MUSKMELON; PLANT PATHOLOGY. Oscar A. Lorenz

Bibliography. S. P. Doolittle et al., *Muskmelon Culture*, USDA Agr. Handb. 216, 1961.

Persimmon

A deciduous fruit tree species, *Diospyros kaki*. Persimmons originated in the subtropical regions of China but were cultivated more extensively in Japan where, until 1900, it was the most important fruit. Now, although it is only the fifth most important crop there, Japan leads in world production. Persimmons have been introduced to a number of temperate zone countries; however, they have not attained substantial popularity, and only small commercial plantings exist in the United States (primarily in California), Italy, Brazil, and Israel.

Persimmon cultivars adapt to a wide climatic

range. Although they have a low chilling requirement, they can tolerate temperatures as low as 5°F (−15°C) if they are dormant. However, areas with late spring and early fall temperatures below 27°F (−3°C) should be avoided, as young growth and maturing fruit will be damaged. Mean annual temperatures averaging 57–59°F (14–15°C) are required for good growth and quality.

Cultivars. For horticultural purposes, persimmons are classified as astringent or nonastringent. Astringent persimmons have water-soluble tannins in the flesh that decrease as the fruit softens to ripeness. They are conical in shape. Nonastringent persimmons, are firm when ripe, as their soluble tannins decrease with pollination. They have an oblate shape. Astringent cultivars produce best in cool climates, whereas nonastringent types require hot, dry climates for good quality. Both astringent and nonastringent cultivars are subdivided into pollination-constant and pollination-variant types. Pollination-variant persimmon flesh darkens in response to the presence of seeds, produced by pollination, but pollination-constant persimmon flesh does not. Among major persimmon cultivars grown in Japan, California, and New Zealand, wide variation exists among the astringent and nonastringent, pollination-constant and variant cultivars, suggesting centuries of chance and deliberate cross breeding. Nonastringent cultivar production has increased at the expense of astringent cultivars.

Propagation and cultivation. Persimmons are clonally propagated by budding onto seedling rootstocks. Seed extracted from mature fruits of *D. lotus*, *D. kaki*, and *D. virgiana* are germinated in the greenhouse in the fall. Seedlings are transplanted in outdoor nursery rows when temperatures are over 55°F (13°C). After a growing season, seedlings are of sufficient size to chip-bud in September. The dormant, budded whip is maintained in the nursery row until the danger of frost subsides. During the following spring, it is removed from the nursery rows and held in moist sawdust as a bare root whip until planting. Trees are planted at 12–15 ft (3.5–4.5 m) in the row, with 16–18 ft (5–5.5 m) between rows. Whips are topped at 20 in. (50 cm) when planted and are trained to a modified central leader tree through five dormant seasons. Persimmons bear laterally on current season's growth. As persimmons are vigorous and tend to alternate crops, mature-tree dormant pruning is used to maintain an upright tree and thin crop load. Persimmons can be monoecious, dioecious, and hermaphroditic. The major cultivars are dioecious, with female flowers capable of producing fruit without fertilization. Male pollinating trees may be planted at a ratio of one male to eight females to reduce fruit drop and increase fruit quality. This is a common practice among nonastringent cultivars. However, pollination produces seeds, an undesirable quality. *See* PLANT REPRODUCTION.

Diseases. Persimmons are relatively disease free in California. The most troublesome disease is crown gall, caused by *Bacterium tumifaciens*. The major symptom is large galls at the juncture of large roots and the trunk on young trees. These eventually die and slough off, inviting secondary infections that girdle the roots and crown. The disease is easily avoided by soaking rootstock seeds in sodium hypochlorite prior to germination and stratifying the seeds in nursery beds rather than refrigerating them. If the nursery stock is not clean, the young trees can become susceptible to disease any time the trunk or crown roots are wounded during digging and planting. Another common disease in California is mushroom root rot, caused by *Armillaria mellea*. This fungal disease can exist in the soil for years on the roots of previously infected hosts. It attacks new root growth in spring and early summer, when soil temperatures range between 62 and 75°F (17 and 24°C), and results in a slow decline of the tree, manifesting as poor shoot growth, leaf drop, and eventual death. During the host summer months, trees can suddenly collapse. The best method of preventing disease is by developing a resistant rootstock. The most damaging postharvest disease is caused by the omnipresent fungus *Penicillium*. It is controlled by avoiding injury during picking and transport. *See* PLANT PATHOLOGY.

Harvesting. Multiple, selective harvests are done in late fall to ensure that the fruits achieve the desirable deep orange-red color. Fruits are harvested by closely clipping the stem above the calyx with clippers.

Astringent cultivars are marketed fresh or dry, made into jams and jellies, or frozen as a pulp for cooking. Nonastringent persimmons, which can be maintained in cold storage for as long as 2 months, are primarily marketed as fresh fruit. *See* FRUIT, TREE.

Louise Ferguson

Bibliography. J. Janik, *Advances in New Crops*, 1990.

Personality theory

A branch of psychology concerned with developing a scientifically defensible model or view of human nature—the modern parlance, a general theory of behavior.

Personality theory is an outgrowth of nineteenth-century French and German psychiatry. Nineteenth-century psychiatry was a practical rather than an academic discipline, concerned with the origins and treatment of neurotic and even psychotic disorders. It dealt with abnormal rather than normal thought processes, focusing on the unconscious rather than the conscious mind. Perhaps most distinctively, it considered a single person as the unit of study, rather than a single perception or behavioral response as in academic psychology. These themes continue today. Thus, with only a few exceptions, the prominent personality theorists have been psychiatrists or clinical psychologists, and personality psychology has developed largely outside the mainstream of academic psychology. The major exceptions are those personality theorists with a behaviorist orientation who stress the role of learning in personality

development—John Dollard, Neal Miller, and Albert Bandura.

Personality theory as a portion of academic psychology came into existence in the late 1930s with the publications of Gordon Allport and Ross Stagner. The field has grown steadily, until today, in conjunction with social psychology, personality is one of the most active areas of research in all psychology.

There are two primary usages for the word "personality" in ordinary language. In the first sense, personality refers to the distinctive impression a person makes on others. In thinking of someone (a parent, a friend, a favorite or hated teacher), that person's face is rarely evoked; rather, the thought concerns the unique way that person acted on a particular occasion, the kind of atmosphere that person created by his or her presence. This is known as the surface definition of personality. In the second sense, which is usually preferred by philosophers, theologians, playwrights, and novelists, personality refers to the inner core of a person, his or her soul or "true" nature. This inner nature may be expressed in terms of the unique impression a person makes on others, but personality is defined in terms of its core rather than surface aspects.

Basis for Theories

A personality theory is a set of assumptions and expectations about oneself and other people that guides one's perceptions of, and actions with regard to, oneself and others. Personality theories answer three kinds of questions: How or in what ways are people all alike—for example, are humans naturally aggressive? How or in what ways are people different—for example, what were the sources of Beethoven's creativity? What is the meaning of a particular, inexplicable action by someone directly or indirectly known—for example, why did Lee Harvey Oswald assassinate John F. Kennedy? However they may differ, all personality theories eventually respond to these three questions, but of course they provide unique answers to them.

Differences. Personality theories differ with regard to certain features which are often used as a means of evaluating them.

Scope. Theories differ in terms of their scope. Cognitive dissonance theory and attribution theory, for example, each offer a single psychological principle to explain the manner in which people think about social behavior. On the other hand, psychoanalysis has great scope and complexity, and attempts to account for all important aspects of social and psychological experience. Academic psychologists tend to prefer theories with limited scope; clinical psychologists and psychiatrists tend to prefer theories with larger scope.

Explicitness. Theories also differ in terms of their explicitness—the degree to which it is clear what a theory would predict in a particular case. Those personality theories inspired by behaviorism and learning theory (Dollard, Miller, Bandura, B. F. Skinner) tend to be admirably explicit; those inspired by psychoanalysis (Sigmund Freud, Carl Jung, Alfred Adler, Erik Erikson) tend to be inexplicit—one never quite knows what these theories will predict in a given case.

Theory versus fact. Theories differ in terms of the degree to which they conform to empirical facts. This, however, is a subtle judgment because what counts as a fact depends on the theory, and all theories point to certain facts that support them, and ignore facts that contradict them. That is, all theorists ignore the facts that embarrass their theories, but this tendency is a deeply human characteristic.

Supporters. Finally, theories differ in the degree to which they attract advocates and supporters. Far and away the two most influential perspectives on personality are the behaviorist models of Skinner, Bandura, and other social learning theorists, and the psychoanalytic model of Freud and his later followers. Although behaviorism and psychoanalysis represent widely divergent perspectives on human nature, the work of Dollard and Miller was largely aimed at synthesizing these apparently antithetical traditions. Curiously, the popularity of various theories often seems as closely related to the politics of science and the charisma of the theorist as it does to the intellectual merit of the ideas involved.

Classification. Most personality theories can be classified in terms of two broad categories, depending on their underlying assumptions about human nature. On the one hand, there are a group of theories that see human nature as fixed, unchanging, deeply perverse, and self-defeating. These theories emphasize self-understanding and resignation; in the cases of Freudian psychoanalysis and existentialism, they also reflect a distinctly tragic view of life—the sources of human misery are so various that the best that can be hoped for is to control some of the causes of suffering. On the other hand, there are a group of theories that see human nature as plastic, flexible, and always capable of growth, change, and development. Human nature is basically benevolent; therefore bad societies are the source of personal misery. Social reform will produce human happiness if not actual perfection. These theories emphasize self-expression and self-actualization—the cases of Carl Rogers and Abraham Maslow, they reflect a distinctly optimistic and romantic view of life.

Core Ideas

There are six core ideas with which any competent theory of personality must deal. Theories differ in terms of how well they treat these ideas, which provide another means for evaluating them.

Motivational theory. Motivational theory is concerned with what makes people go, what provides the energy or reasons for their actions. Motives are the major, if not primary, explanatory variables in personality theory; people's actions are normally explained by appealing to motivational concepts. Motivational terms vary widely across personality theories; nonetheless, each theory tends to specialize in one of two broad types of motivational concepts.

Biological theories. The first type of motivational concept is oriented toward biology or physiology. This

category includes terms such as drive, need, instinct, passion, and urge, so that when a theorist uses these terms, a biological form of motivational theory is being endorsed. Biological theories (such as psychoanalysis) describe human motivation in terms of a small number of fixed and unchanging motives, and everyone is assumed to share this same set of motives. These motives are usually described as unconscious, as beyond rational control, and as becoming harmful or dangerous if frustrated or ignored for a long time. *See* INSTINCTIVE BEHAVIOR; PSYCHOANALYSIS.

Psychological theories. The second category of motivational concepts is psychological or mental in origin, and includes terms such as values, expectations, attitudes, preferences, and goals; and so, when a theorist uses these terms, a psychological form of motivation is being considered. Theorists of this type (Allport, G. A. Kelly, D. C. McClelland) tend to think of human motivation in terms of a relatively large number of personalized or idiosyncratic motives, so that each person's complement of motives is unique. These motives are usually regarded as conscious, as rationally chosen, and as capable of being modified through education or self-reflection. In contrast with biological motives, which are seen as driving behavior, psychological motives are seen as guiding and directing it. Consequently, psychological motives—the desire for fame, power, or revenge—are more subtle in their operation than the biological motives. *See* MOTIVATION.

Personality development. Again, personality theories differ considerably in the degree to which they deal with personality development. For some (Erikson), their theory of development *is* their theory of personality. Others, especially the behaviorists, rarely consider the subject.

Classic view. Among those theories concerned with personality development, two views prevail. In the first, classic developmental theory, personality development is seen as proceeding through a series of levels or stages which are usually correlated with a person's age. Personality is thought to change markedly as a person moves from one stage to the next; that is, the movement from one stage to the next is accompanied by a qualitative transformation in the structure of personality. Classic developmental theory also assumes that earlier experiences have a lasting impact on later development; those experiences that occur earliest are the most significant. Thus the events of the first year of life are often crucial for, and are reflected in, adult personality. The sources of adult neurosis, therefore, are located in infancy and early childhood according to this classic view. Europeans such as Freud, Jung, and Erikson by and large adopted this view of personality development.

Learning. In contrast to the classic view, many British and American theorists adopt what might be called the "development as learning" view. For them, most notably the factor analysts (R. B. Cattell, H. J. Eysenck) and behaviorists (Skinner, Bandura), development is linear rather than stagelike. No structural transformations take place during development; an adult differs from a child primarily in being more experienced, in having learned more ways of dealing with the world. Children are inexperienced adults, with essentially the same needs, capacities, and interests. In this case, development is not tied to or programmed by age; development is a function of sheer experience—the more experience, the more development. These theorists also tend to deny that the timing of experiences is important; early experiences are no more crucial than later experiences. The effects of early childhood trauma are readily overcome in later development. That which makes an experience important is its frequency, not its timing. The sources of adult neurosis, therefore, are not located in a person's childhood but in the current circumstances of that person's life. Since development is not stagelike, the trend of development can be altered at any time by changing a person's living circumstances.

Moral development. One aspect of personality development is so important that it has become a separate field of study in itself. The topic used to be called socialization, but now many psychologists focus on the problem of values and their transmission and on the moral implications of conformity, rebellion, and delinquency. Despite the seeming trendiness of this interest in moral development, it is actually a very old preoccupation among personality theorists—for example, it is the central problem in classic psychoanalytic theory. There are four problems to be distinguished with regard to moral development.

Effects of moralization. The first contrast concerns the effects of moralization on personality development. Psychoanalytic theory, for example, describes moral development as absolutely necessary for the survival of civilization but as devastating for the individual who undergoes it. The moralization process transforms an individual from an antisocial child to a civilized member of society, but at the cost of that child's spontaneous, natural tendencies, and ultimately at the cost of that person's happiness. To live in society, the natural instincts of a child must be ruthlessly suppressed. According to psychoanalysis, most pleasure comes from the gratification of instincts. By definition, then, living in society is unpleasurable and contrary to human instinctual nature. In contrast to psychoanalysis, classic sociological theory maintains that children become truly human only when they internalize the morality of their culture. Far from crushing or stultifying one's personality, the process of internalizing or incorporating the morals of one's family and culture, according to classic sociology, deepens and enriches one's personality and gives meaning and purpose to one's life. From this viewpoint, children of nature, happy savages, and persons who have somehow escaped from the "bondage" of civilization are shallow, empty, and less than human. It is ties to others, concern for their welfare, and a feeling of respect for the rules and traditions of one's culture that makes someone a whole person and a well-rounded personality.

Causes. The second problem concerns the causes of moral development—what instigates, triggers, or

sets off this process? Some feel that moral development is somehow preprogrammed in the psyche, that it is an unfolding process that occurs naturally and spontaneously in every human infant. According to J. Piaget, for example, the child-training practices of well-meaning adults in reality only interfere with the moralization process. To the degree that adults intervene in children's spontaneous growth, they impede the normal process of moral development as it unfolds in each child. The alternative viewpoint is that without clear, direct, and even forceful adult intervention, children will remain forever childlike, antisocial, narcissistic, and impulsive—unfit as companions for themselves or other people. On this second view, best exemplified by Freud, there is no natural unfolding, no spontaneous evolution of personality. Parents and society are responsible for training children, and adult personality reflects the quality of this training.

Socialization. The third question concerns what it means to be socialized—what distinguishes a person whose moral development is normal from someone who has not developed normally? Traditionally there have been two answers to this. On the one hand, psychoanalysts and behaviorists think of moral development in terms of the number of rules and values that a person has internalized, and in terms of that person's capacity to feel guilt. A person with a well-developed conscience has internalized many rules and tends to feel guilty if breaking one of the rules. Conversely, criminals have internalized few rules and feel little guilt about breaking them. An alternative way of stating this first definition of moral development is in terms of a person's attitudes toward authority—to be well socialized means having positive attitudes toward authority and therefore being willing to comply with the requests and orders of authority figures. Conversely, persons who are poorly socialized are hostile to authority, and much of their behavior can be seen as defiance or rebellion against authority.

A second way of thinking about socialization has to do with social sensitivity, especially sensitivity to other persons' needs and expectations. Here, the more responsive one is to the expectations of one's family and other members of society, the more socialized one is and the more closely will one's actions conform to the rules and values of society. A person may feel guilty when violating these expectations. From this perspective, criminals are insensitive to the needs and expectations of others. Consequently, the ability to respond to others' expectations is seen as a fundamental part of what it means to be human.

Autonomy. The final question in theories of moral development concerns the topic of autonomy. For many theorists the defining feature of moral development is a person's willingness to conform to the norms and laws of family and culture. Conformity is essentially equated with moral development; psychoanalytic and behaviorist theories are good examples of this approach. The alternative, exemplified in the works of W. McDougall and the existentialists, emphasizes autonomy as the essential feature of moral behavior. These writers argue that when one conforms to the rules of society and to others' expectations, one is not acting in a truly moral way. That which is necessary for true moral behavior is personal, autonomous choice. Morally mature people may act contrary to the laws of their society and expectations of their families; in these cases they act autonomously.

Self concept. Personality theories differ greatly in terms of how they deal with the self concept. Behaviorists, psychoanalysts, and factor theorists tend to give the topic a skimpy treatment at best. Among those who deal with the self concept, it is usually seen as a central aspect of personality. Some, notably Jung, regard the self as a state of being—attaining selfhood is the major goal of life; others see the self as equivalent to personality (Allport, Carl Rogers); and a third group see the self concept as a view or image that one holds of oneself, which image serves to filter and screen action. But in this third case the self concept is just one aspect of personality. Thus, there are four ways of dealing with the self concept: minimize its importance; postulate it as a state of being; equate it with personality; and regard it as one component of personality.

Assuming that people can spell out their self-images, how do these images compare with how the individual would like to be viewed? The correspondence between real and ideal self-images is often regarded as an index of mental health. Phenomenologically oriented theorists (such as Rogers) use this correspondence as a way of evaluating the effects of psychotherapy—as one's adjustment improves, one's real self-image will become more like one's ideal self-image.

Unconscious processes. The idea that an unconscious but dynamically active mind operates outside the control of the individual and causes a person to do things he or she would not consciously do, is one of the oldest and most interesting notions in psychiatric theory and popular psychology. Although traditional psychiatric theory accepts the notion of a dynamic unconscious as an article of faith, other psychologists are suspicious of the idea. Not surprisingly, then, personality theories differ considerably in terms of the importance they assign to unconscious processes. Generally speaking the older the theory and the more it is concerned with explaining the origins of neurotic disorders, the more heavily will it appeal to unconscious processes. Conversely, the more recent the theory and the more it is concerned with problems other than neurotic disorders—creativity, self-actualization, personal soundness, status—the less will that theory be concerned with unconscious processes.

Among those theories that are concerned with unconscious mental processes, many issues are still disputed. Perhaps the most important issue concerns how thoughts, wishes, and desires become unconscious in the first place. There are probably four answers to this problem. Jung argues that the nature of the mind is such that a portion of it is and always will be unconscious, that the most that can be

done is be alert for signs of unconscious thoughts and wishes leaking through into everyday conduct. Freud maintains that thoughts become unconscious through a direct action of the mind; one portion of the mind, an unconscious censor, decides that certain thoughts are unacceptable and blocks them out of consciousness by directing mental energy against them.

A third viewpoint, strongly rejected by Freud, is called dissociationism. This view, widely popular in the nineteenth century, gained prominence through the work of Ernest Hilgard. Here the notion is that the mind operates along separate channels, and the contents of consciousness depend on the channel one has tuned in; the contents of any channel will be unconscious so long as one is tuned in elsewhere. The final viewpoint comes from existentialism. Existentialists argue that people often lie to themselves about certain unpleasant facts (about being too fat, too lazy, and such). These lies include the lie that certain memories have been "forgotten," or that the person did not realize what he or she was doing or saying on particular occasions. The unconscious is composed of these lies.

Emotional dynamics. The topic of emotional dynamics concerns the various psychological factors that regulate one's life, how they are interrelated, and most importantly, how these factors can be put out of kilter. Ultimately, then, emotional dynamics concerns how and why people break down psychologically, how they become neurotic or even psychotic, and what may be done to help them. The study of emotional dynamics overlaps all the earlier central ideas—personality development, motivation, the self concept, and unconscious processes are involved to some degree in every explanation of why and how people break down.

Five approaches to the origin of neurosis can be found in the various theories of personality.

Classic psychodynamics. The first of these is the classic psychodynamic view that certain powerful instinctual impulses or potent emotions and memories somehow get sealed off in the unconscious mind where they then begin to fester like a psychic boil. Eventually these buried unconscious impulses and memories begin to affect everyday behavior, but always in peculiar ways that are self-defeating for the person involved. Such problems are relieved or cured by a careful, elaborate, and often time-consuming exploration of the person's unconscious mind and personality development. The goal is to bring the forgotten memories to consciousness where they may be analyzed and overcome.

Phenomenology. The second approach is the phenomenological perspective of such theorists as Jung, Rogers, and the existentialists. Here the assumption is that people are motivated by a need for personal growth and development, for self-actualization and transcendence. Unfortunately, according to these theories, the path of life is strewn with difficulties, and certain traumatic events in one's life can block the process of personal growth. Over time these frustrated strivings for wholeness and completeness turn into neurotic symptoms. The cure for neurosis, in these cases, is to remove the psychological forces that block the process of self-actualization.

Learning theory. The third approach comes from learning theory. Here the primary assumption is that the unconscious is irrelevant for understanding neurotic symptoms, because the origins of neurotic behavior are available to public inspection; they are learned patterns of behavior that turn out to be inappropriate and self-defeating. But because they are learned, neurotic symptoms can, in principle, be unlearned. Two kinds of neurotic behavior patterns can be learned. On the one hand, a person may learn to fear certain objects, people, or situations, although fear is inappropriate and may become very disruptive in later life. On the other hand, people can learn certain habits, certain ways of acting that may work in one special situation but be terribly inappropriate in other situations. So, for example, a child may learn that it can get its way with the parents by throwing tantrums. In later life, when that person's wishes are opposed by others, the person may react with anger, screams, and hostile emotional displays. In this second case, the maladaptive behavior pattern is like a bad habit that needs to be unlearned. The essential feature of any learning theory approach to the origins of neurotic behavior is that the source of the problem is the behavior itself—the problem is not buried in the unconscious mind. If the behavior can be changed, the problem will be eliminated. Learning approaches to the study of neurosis tend to challenge the validity of psychodynamic and phenomenological approaches.

Sociocultural theory. A fourth way of thinking about neurosis may be called the sociocultural theory, a viewpoint that comes from anthropology and sociology. Here mental illness is seen as a function of changes in society and the social environment. Wars, economic depressions, famines, earthquakes—any anthropogenic or natural disaster that uproots people, disturbs their standard of expected life style—will produce stress. When this stress reaches a certain level, disordered behavior breaks out in the population—alcoholism, drug addiction, violent crime, and mental illness are all symptoms of a society under pressure. Environmentally produced stress may be confined to a remote mountain village—a mudslide or flood may affect only a localized and very specific population. On the other hand, a major economic depression may affect the entire population of a country.

In sociocultural theory, emotional disorders, because they are caused by environmental conditions, are treated by modifying environmental factors. Abused children are placed in foster homes where presumably they will receive better care; underprivileged children are provided federally financed preschool education in order to stimulate their intellectual development. Within clinical psychology and psychiatry, group therapy and family therapy are based on sociocultural theory.

Biophysical model. A final way of thinking about emotional dynamics is the biophysical model. This model

assumes that biological or neurological defects, not environmental stress or frustrated desires, produce mental disorders. This viewpoint is consistent with medical theory and practice, but not very popular with psychologists. It received its strongest support with the discovery around the turn of the century that a distinctive group of mental patients, classified dementia paralytica, were in fact suffering from an advanced state of syphilis. This finding led to the hope that some kind of neurological defect underlies all other forms of mental disorder, from anxiety attacks to schizophrenia.

Biophysical theory assumes biological causes for emotional disorders. Consequently, this theory recommends biophysical treatment for such disorders. Whatever the emotional problem, it will be treated with biochemical or physical means.

These six topic areas—motivation, personality development, moral development, the self concept, the unconscious, and emotional dynamics—form the substance of personality theory. All theories treat these six topics and differ largely in terms of how they handle them. Taken together, the various personality theories reflect the best modern judgment regarding the nature of human nature. *See* PSYCHOSIS.

Robert Hogan

Bibliography. R. B. Ewen, *Introduction to Theories of Personality*, 4th ed., 1992; P. Fonagy and A. Higgitt, *Personality Theory and Clinical Practice*, 1985.

PERT

An acronym for program evaluation and review technique; a planning, scheduling, and control procedure based upon the use of time-oriented networks which reflect the interrelationships and dependencies among the project tasks (activities). The major objectives of PERT are to give management improved ability to develop a project plan and to properly allocate resources within overall program time and cost limitations, to control the time and cost performance of the project, and to replan when significant departures from budget occur.

The basic requirements of PERT, in its time or schedule form of application, are:

1. All individual tasks required to complete a given program must be visualized in a clear enough manner to be put down in a network composed of events and activities. An event denotes a specified program accomplishment at a particular instant in time. An activity represents the time and resources that are necessary to progress from one event to the next.

2. Events and activities must be sequenced on the network under a logical set of ground rules.

3. Time estimates can be made for each activity of the network on a three-way basis. Optimistic (minimum), most likely (modal), and pessimistic (maximum) performance time figures are estimated by the person or persons most familiar with the activity involved. The three-time estimates are used as a measure of uncertainty of the eventual activity duration.

4. Finally, critical path and slack times are computed. The critical path is that sequence of activities and events on the network that will require the greatest expected time to accomplish. Slack time is the difference between the earliest time that an activity may start (or finish) and its latest allowable start (or finish) time, as required to complete the project on schedule.

5. The difference between the pessimistic (b) and optimistic (a) activity performance times is used to compute the standard deviation ($\hat{\sigma}$) of the hypothetical distribution of activity performance times [$\hat{\sigma} = (b - a)/6$]. The PERT procedure employs these expected times and standard deviations (σ^2 is called variance) to compute the probability that an event will be on schedule, that is, will occur on or before its scheduled occurrence time.

In the actual utilization of PERT, review and action by responsible managers is required, generally on a biweekly basis, concentrating on important critical path activities. A major advantage of PERT is the kind of planning required to create an initial network. Network development and critical path analysis reveal interdependencies and problem areas before the program begins that are often not obvious or well defined by conventional planning methods.

Joseph J. Moder

Bibliography. M. O. Ball et al., *Handbooks in Operations Research and Management Science*, vol. 7: *Network Models*, 1995; J. J. Moder and C. R. Phillips, *Project Management with CPM and PERT*, 3d ed., 1995.

Perthite

Any of the oriented intergrowths of potassium- and sodium-rich feldspars, $(K,Na)AlSi_3O_8$, whose proportions are determined in part by the initial composition of the alkali feldspar from which they exsolved and whose physical properties are thus somewhat variable. The early stages of perthite formation from homogeneous, usually monoclinic $(K,Na)AlSi_3O_8$ may be observed experimentally by high-magnification electron microscopy.

Spinodal decomposition and homogeneous nucleation are the dominant mechanisms of exsolution, producing compositional modulations that initially are lamellar in nature and coarsen with increasing time and temperature of annealing from 10 to 50 nanometers or more. The coherent K- and Na-rich lamellae are subparallel and have interfaces oriented within a few degrees of the $(\bar{8}01)$ or $(\bar{6}01)$ crystallographic planes, a consequence of minimizing the elastic strain energy between them. Details of the exsolution process vary from one bulk composition and rock annealing history to another, and the anorthite ($CaAl_2Si_2O_8$) component is more important in some perthites than others. Water has a critical effect on cation diffusion rates and the kinetics affecting the final textures. One suggested path for the coarsening of a cryptoperthite from the earliest stages of

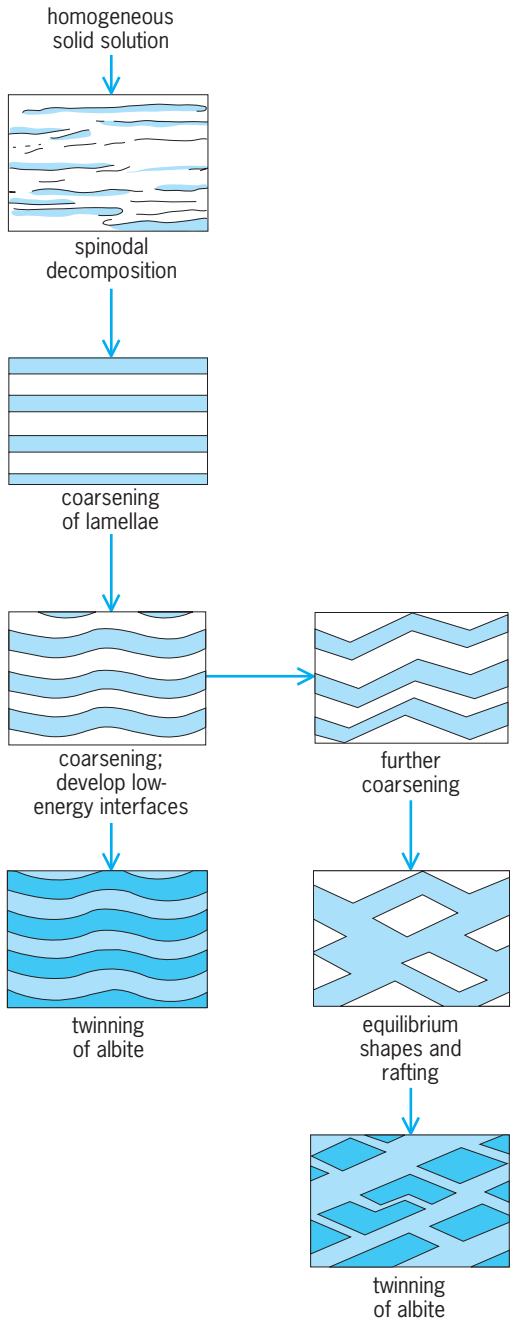

Schematic showing the development of two perthitic intergrowth textures. (*After G. W. Lorimer and P. E. Champness, The origin of the phase distribution in two perthitic alkali feldspars, Phil. Mag., 28:1401, 1973*)

exsolution is shown in the **illustration**. If the final K- and Na-rich lamellae or particles are submicroscopic, the composite feldspar is called cryptoperthite. If the particles are small enough and the feldspar relatively clear, Rayleigh-type scattering of light may occur, giving rise to the beautiful blue-to-whitish luster of the semiprecious gem called moonstone. If coarsening has progressed to the micrometer scale and can be seen on a polarizing microscope, the composite is called microperthite; and if the two feldspars are visible to the eye in hand specimen, it is called perthite or macroperthite. Often the albite phase will appear as white veins or blotchy patches against a colored K-rich phase, which may be green to blue microcline (amazonite) or dull pink to orange-brown orthoclase. *See* ALBITE; ANORTHITE; ANORTHOCLASE; CRYSTAL STRUCTURE; FELDSPAR; IGNEOUS ROCKS; MICROCLINE; ORTHOCLASE.

Paul H. Ribbe

Perturbation (astronomy)

Departure of a celestial body from the trajectory it would follow if moving only under the action of a single central force. Perturbations may be caused by either gravitational or nongravitational forces.

Corrections to elliptic orbits. In the solar system, orbits of planets may be adequately represented by mean elliptical elements to which are added small corrections due to the mutual planetary attractions. Although such motion is referred to as disturbed, it is as much a consequence of the law of gravitation as is undisturbed elliptic motion. *See* ORBITAL MOTION; PLANET.

Another method of representing perturbed motion is to augment the position derived from the mean ellipse by the actual displacements in the coordinates due to the disturbing forces. These perturbations of the elements, and perturbations of the coordinates, are represented by infinite series; usually, many terms are required to represent the disturbed motion accurately. These analytical expressions are referred to as general perturbations and, with their associated mean elements, form a general theory of the motion. In some instances, such as the orbit of an outer satellite of Jupiter or the motion of a comet moving with nearly parabolic velocity, the analytical expressions representing the perturbing function become so involved (mainly because of lack of convergence of the Fourier series) that general perturbations are not attempted. Instead, the perturbed positions are computed from a step-by-step numerical integration of the equations of motion; this is known as the method of special perturbations. In practice, many long-term studies of perturbed motion are now carried out using numerical integrations on digital computers. *See* FOURIER SERIES AND TRANSFORMS; NUMERICAL ANALYSIS; SERIES.

Long- and short-term disturbances. Planetary orbits are subject to two classes of disturbances: secular, or long-term, perturbations; and periodic, or relatively short-term, perturbations. Secular perturbations, so called because they are either progressive or have excessively long periods, arise because of the relative orientation of the orbits in space. They cause slow oscillatory changes of eccentricities and inclinations about their mean values with accompanying changes in the motions of the nodes and perihelia. The periods of time involved in these oscillations may extend from 50,000 to 2,000,000 years. Periods and major axes of orbits are not affected by secular change. For the orbit of the Earth, the present inclination to the invariable plane is $1^\circ 35'$. This will diminish to a minimum of $47'$ in approximately

20,000 years. The eccentricity, presently 0.017, is diminishing also and will reach a minimum of 0.003 in about 24,000 years.

Periodic perturbations arise from the relative positions of the planets in their orbits. When the disturbed and disturbing planets are aligned on the same side of the Sun, the perturbation reaches a maximum, and reduces to minimum when alignment is reached on opposite sides of the Sun. The size of a periodic perturbation is a function of the mass of the disturbing body and of the length of time the two planets remain near the point of closest approach. Periodic perturbations continually shift a planet away from the position it would occupy in undisturbed motion, moving it above or below the orbital plane, nearer to or farther from the Sun, and forward or backward in the orbit.

Commensurable motions. If the mean motion of the disturbed planet were exactly a submultiple, say 1/2, of the mean motion of the disturbing planet, the maximum perturbation produced by their close approach would always occur in the same part of the disturbed orbit. The displacement in position of the disturbed planet would increase with each coincidence until the character of the orbit became modified to the point where exact commensurability of the mean motions would cease to exist.

Because the solar system is middle-aged, cosmically speaking, few examples of commensurability of mean motions exist today. None is found in the motions of the major planets, although Neptune is locked in a 2:3 commensurable motion with the dwarf planet Pluto. Cases of near commensurability exist which give rise to long-period periodic terms of large amplitude. As an example, the periods of Jupiter and Saturn are nearly in the ratio of 2:5. Thus, after nearly five revolutions of Jupiter, the two planets return to approximately the same juxtaposition. Their line of coincidence, however, sweeps slowly around Jupiter's orbit, completing a circuit in about 850 years and thus producing a perturbation of this period.

In the early history of the solar system, the newly formed planets were still orbiting in the midst of a remnant disk of planetesimals. This could have resulted in significant orbital evolution as the planets gravitationally interacted with the disk. One possible outcome is that Jupiter and Saturn could have encountered the strong 1:2 commensurability. Numerical simulations suggest that this could account for many of the dynamical characteristics of the solar system. *See* SOLAR SYSTEM.

Among the four inner planets, the periodic perturbations are small, amounting in orbital longitude at most of 0.′25 for Mercury, 0.′5 for Venus, 1′ for Earth, and 2′ for Mars. Periodic perturbations of the outer planets are larger, reaching in the case of the long-period terms to 30′ for Jupiter, 70′ for Saturn, 60′ for Uranus, and 35′ for Neptune.

Because the amplitude of a periodic perturbation depends on the mass of the disturbing planet, observational measurement of this amplitude affords a method of determining the disturbing mass. For the planets Mercury and Venus, which do not have satellites, this is the only method of determining the mass without using direct measurements obtained by tracking spacecraft as they pass near the planet.

The orbits of the minor planets are affected in varying degree by the attractions of the major planets. Those orbits passing close to Jupiter suffer large perturbations which, if the mean motions were commensurable with that of Jupiter, would be augmented at each close approach until the trajectories were sufficiently altered to reduce the commensurability. In the overall distribution of mean motions of the minor planets, there are noticeable gaps near the points where the period would be an exact submultiple (1/2, 1/3, 2/5, . . .) of the period of Jupiter. Furthermore, the distribution has also been influenced by secular perturbations whereby the natural rate of the motion of the apsides of a minor planet's orbit is close to one of the rates of the major planets. The combination can lead to chaotic evolution of some asteroid orbits, even resulting in their becoming Earth-crossing. A small group of minor planets, called the Trojan asteroids, has been so completely captured by Jupiter that they oscillate about the 60° points which form equilateral triangles with Jupiter and the Sun. *See* ASTEROID; TROJAN ASTEROIDS.

Effect on comets. Planetary perturbations also affect the orbits of comets. Studies of the motion of Halley's Comet indicate that the time from one perihelion passage to the next has varied by almost 5 years because of perturbations. Most comets approach the Sun at nearly parabolic speeds in randomly oriented orbits, but if a comet approaches close to one of the more massive major planets, the planet may so alter the trajectory that the comet pursues an elliptical orbit thereafter. A number of short-period comets whose orbits agree only in that all pass close to Jupiter illustrate the perturbing effect of this planet on cometary orbits. *See* COMET; HALLEY'S COMET.

Nongravitational causes. Objects may also be subject to several nongravitational types of perturbation. For example, the rarefied matter that is given off by the head of the comet is forced into a trajectory away from the Sun by the pressure of solar radiation. This is an instantaneous effect that changes the orbit of the emanating dust particles, but there are other, more subtle effects. The nonuniform reradiation of absorbed direct sunlight or even reflected sunlight gives rise to Poynting-Robertson light drag. This is most effective for small particles with a size comparable to the wavelength of the incident light. It results in the gradual shrinking of a particle's orbit. The Yarkovsky effect is experienced by rotating objects large enough to have a night-day temperature difference across their surfaces. The net result can lead to orbit expansion or contraction over millions of years at a rate depending on the rotational properties of the object. The Yarkovsky effect is thought to play a role in the orbital evolution of some asteroids. In the presence of a magnetic field, a charged body is also subjected to a Lorentz force, another form of

nogravitational force that is important for micrometer-sized ring particles orbiting in the magnetic fields of the giant planets. *See* MAGNETISM; RADIATION PRESSURE.

Associated with many of the periodic comets are swarms of smaller particles that appear as meteors upon collision with the upper atmosphere of the Earth. The density of these swarms is so tenuous that they cannot hold themselves together by their own gravitation, and planetary perturbations of speed and direction soon spread the components completely around the orbit. The annual meteor showers, such as the Perseids, reflect this dispersal of particles along the orbit. The effect of the Earth's attraction on a meteor trajectory depends on the relative velocity, that is, whether the Earth is overtaking the meteor or meeting it head on. Once the meteor enters the upper reaches of the Earth's atmosphere, its motion is subject to a nongravitational perturbation caused by atmospheric drag. This resistance to the passage of the particle is evidenced by the trail of incandescent gas and vapor which forms until the particle is consumed or continues in its trajectory greatly decelerated. *See* METEOR.

Many asteroids can be classified as belonging to groups or families, whose members share similar orbits and even spectral properties. Asteroid families are believed to originate in the collisional breakup of parent bodies, the resulting debris being subjected to the perturbing effects of the planets as well as nongravitational forces, leading to a gradual dispersal.

Perturbations of rings. Planetary rings are often in locations where they are subjected to the gravitational perturbations of nearby satellites. In some cases, observations of the resulting effect on the rings can be used to determine the mass of the satellite. For example, the Saturnian satellites Pan and Daphnis orbit in Saturn's Encke and Keeler gaps, respectively. Each satellite produces wavelike structures on the adjacent ring edges. The amplitude of these features is directly proportional to the mass of the perturbing satellite, and this gives an estimate of the mass. *See* SATURN.

Perturbations of satellite orbits. The motions of planetary satellites, natural and artificial, reflect both gravitational and nongravitational perturbations. The centrifugal force arising from the rotation of a planet causes a deformation or oblateness of figure. In such a case, the central mass does not attract as if it were concentrated at its center. For a close satellite, the principal perturbation arises from the attraction of this equatorial bulge. The effect of this attraction on an otherwise undisturbed satellite orbit is a gradual regression of the line of nodes on the equatorial plane and a rotation of the line of apsides. Both rotations vary with the inclination of the satellite orbit. Nearer to the primary, the tidal forces may become so great that a satellite would be literally torn to pieces. For a fluid satellite of the same density as the planet, the limit within which this disruptive perturbation occurs is about $2^1/_2$ times the radius of the planet. *See* ROCHE LIMIT.

The gravitational effect of a natural satellite acts to distort the nonrigid body of the planet it orbits. This tidal distortion affects the long-term motion of the satellite, causing expansion or contraction of its orbit depending on whether it lies inside or outside the synchronous location where the orbital period matches the rotational period of the planet. The resulting orbital evolution gives rise to the large number of commensurable pairs of satellites in the Jupiter and Saturn systems.

Satellite motions are also disturbed by the direct attraction of other satellites, the Sun, and, to a lesser amount, by other planets. Observation of the orbital displacements caused by the mutual perturbations of satellites in the systems of Jupiter and Saturn makes possible the determination of the masses of these satellites. Perturbations between locked pairs of satellites (the phenomenon of resonance) can excite the orbits of the satellites, in some cases giving rise to significant tidal heating. The solar attraction is significant in the orbits of the outer satellites of Jupiter and Saturn. *See* JUPITER.

The orbit of the Moon is disturbed mainly by the Sun, with some changes in motion due to the oblateness of the Earth, the figure of the Moon, and smaller perturbations caused by the planets. The attraction of the Sun on the Moon is more than twice the Earth's attraction, but because both the Earth and Moon are free to move, it is only their relative acceleration with respect to the Sun which determines the motion. This relative acceleration toward the Sun is always less than 1/80 of the acceleration of the Moon toward the Earth. The eccentricity and inclination of the Moon's orbit oscillate slowly about their mean values, while the line of apsides advances with an average period of almost 9 years and the nodes regress through one revolution in 18.6 years. *See* MOON.

The observed motion of the lunar node and perigee affords one means of measuring the oblateness of the Earth. The present lunar theory incorporates the value 1/294. The *International Astronomical Union System of Astronomical Constants* (1976) contains a reference ellipsoid of revolution for the Earth having a flattening of 1/298.257. This value has been derived mainly from measures of the motions of the nodes and apsides of artificial Earth satellites. Lunar and solar perturbations of artificial Earth satellite orbits are minor for orbits 500 mi (800 km) above the surface, but grow with increasing distance from the Earth. Atmospheric drag perturbations are significant at this altitude, but decrease with increasing altitude. *See* CELESTIAL MECHANICS; EARTH; GRAVITATIONAL FIELD OF.

Raynor L. Duncombe; Carl D. Murray

Bibliography. J. M. A. Danby, *Fundamentals of Celestial Mechanics*, 2d ed., 1988; A. Milani, A. M. Nobili, and P. Farinella, *Non-Gravitational Perturbations and Satellite Geodesy*, 1987; A. Morbidelli, *Modern Celestial Mechanics: Aspects of Solar System Dynamics*, 2002; C. D. Murray and S. F. Dermott, *Solar System Dynamics*, 1999; A. E. Roy, *Orbital Motion*, 4th ed., 2005; V. G. Szebehely and H. Mark, *Adventures in Celestial Mechanics*, 2d ed., 1998.

Perturbation (mathematics)

A modification in the mathematical structure of a problem changing the problem from one that can be solved exactly, the unperturbed problem, to one, the perturbed problem, for which it is usually possible to obtain only an approximate solution. The methods employed for this purpose form perturbation theory. These methods attempt to express the solution of the perturbed problem in terms of the properties of the solutions of the unperturbed problem.

Examples. Examples of perturbation problems can be found in nearly every branch of mathematics and physics, and in astronomy. The simplest case occurs in ordinary algebra. Suppose that the roots of the equation $f(x) = 0$ are known (the unperturbed problem), and that the roots of the equation $f(x) + \epsilon g(x) = 0$ are to be found (the perturbed problem). The parameter ϵ measures the size of the perturbation. Another set of examples occurs in linear differential equations and in particle dynamics. Possible perturbations include changes in the forces considered to be acting on the particle as well as changes in initial conditions. *See* PERTURBATION (ASTRONOMY).

Several examples occur in partial differential equations. One physical realization occurs in the theory of wave propagation where the perturbations can be changes in the index of refraction, changes in initial conditions, or changes in the nature or shape of the surfaces encountered by the waves. All of these changes can occur separately or concurrently. The first of these changes is called a volume perturbation, the second a perturbation of initial conditions, and the third a perturbation of boundary conditions. Similar examples can be taken from quantum mechanics, where the volume perturbation corresponds to a change in the hamiltonian, and perturbation of initial conditions to quantum-mechanical time-dependent perturbation theory. Other partial differential equations of physics, such as the Laplace equation, the diffusion equation, and the equations of hydrodynamics, furnish further examples. *See* PERTURBATION (QUANTUM MECHANICS).

As a final illustration of these various types of perturbation, consider possible modifications in an equation (as well as boundary and initial conditions) describing the motion of particles such as neutrons or electrons moving through a medium which can scatter and absorb them. The equation is known as the Lorentz-Boltzmann equation and changes in it occur as a consequence of modifications of the laws of scattering and absorption, that is, because of changes in the medium.

All of these problems are linear and can therefore be cast into an equation of the form $A\psi = \lambda\psi$, where ψ is the unknown quantity, λ is a constant, and A is an operator involving among other possibilities differentiation and integration. The quantity ψ may be a scalar, a vector, or more generally a matrix quantity. When solutions can be obtained for only special values of λ, the eigenvalues, the equation is called the eigenvalue equation, and the associated problem is called the eigenvalue problem. The operator A contains the perturbation; that is, A equals $A_0 + \epsilon A_1$, where A_0 is the unperturbed operator and ϵA_1, the perturbing term.

Iteration method. The method generally employed to obtain an approximate solution is called the iteration method. Rewrite the equation $A\psi = \lambda\psi$ as $(A_0 - \lambda)\psi = -\epsilon A_1\psi$. Let the unperturbed solution be ϕ_0, where $(A_0 - \lambda_0)\phi_0 = 0$. Then ψ_1, a first approximation to ψ, is obtained as a solution of $(A_0 - \lambda)\psi_1 = -\epsilon A_1\phi_0$. A second approximation is the solution of $(A_0 - \lambda)\psi_2 = -\epsilon A_1\psi_1$. The nth approximation is obtained in terms of the $(n - 1)$ approximation, ψ_{n-1}, from the equation $(A_0 - \lambda)\psi_n = -\epsilon A_1\psi_{n-1}$. It is assumed that the properties of the unperturbed operator, A_0, are completely known so that the solution of these equations can be obtained. If the sequence $\phi_0, \psi_1, \ldots, \psi_n \ldots$, converges, it will converge to a solution of the problem. For an eigenvalue problem, the procedure must be modified. The first approximation to λ is λ_0; the nth approximation is λ_n. Then the equation determining ψ_n in terms of ψ_{n-1} is $(A_0 - \lambda_{n-1})\psi_n = \epsilon A_1\psi_n$. It is important for the practicality of this procedure that the approximation λ_n can be expressed in terms of the approximation ψ_{n-1} and the operators A_0 and ϵA_1.

In a related and more familiar formulation both ψ and λ are expanded in a power series in ϵ; that is, $\psi = \phi_0 + \epsilon\phi_1 + \epsilon_2\phi_2 + \cdots$ and $\lambda = \lambda_0 + \epsilon\lambda_1 + \epsilon_2\lambda_2 + \cdots$. Then the equation $A\psi = \lambda\psi$ reduces to a set of equations for ϕ_n. For example, the equation for ϕ_1 is $(A_0 - \lambda_0)\phi_1 = -(A_1 - \lambda_1)\phi_0$ and the equation for ϕ_2 is $(A_0 - \lambda_0)\phi_2 = -(A_1 - \lambda_1)\phi_1 + \lambda_2\phi_0$, and so on. This formulation is more complex, and often yields slower rates of convergence than the method just outlined.

The iteration method can be generalized in two respects. First, it is not necessary to use ϕ_0 as the zeroth approximation to ϕ. If by reason of other information a better approximation, say ψ_0, is known, the iteration sequence starts with the equation $(A_0 - \lambda)\psi_1 = -\epsilon A_1\phi_0$. Second, the iteration method can be employed in the treatment of nonlinear as well as the linear problems discussed in detail here.

For the iteration method to be at all possible, it is necessary for the sequence $\psi_0, \psi_1, \ldots$ to exist and to converge. The usefulness of the method increases with increasing rate of convergence. The sequence exists only if the singularities of the perturbation are not too strong, or if the initial zero approximation is properly chosen, or both. When the sequence exists, it will converge for a range in values of the parameter ϵ. The largest value of ϵ is the radius of convergence. This is found to be that value of ϵ for which the equation $A\psi = \lambda\psi$ has at least two degenerate solutions, that is, solutions with identical values of λ. There are various methods of increasing the radius of convergence. For example, the general techniques of analytic continuation, such as the Euler transformation, can often be employed. A clever choice of ψ_0, the zeroth approximation, will often produce the desired effect. The variational method can generate the appropriate choice for ψ_0. A more general method was

developed by I. Fredholm in which the solution of $A\psi = \lambda\psi$ is given as the ratio of two functions, each of which can be expressed as a series of ϵ. For a wide class of operators A, each of these series will have an infinite radius of convergence.

The eigenvalue problem can be reduced to the problem of the solution of a set of homogeneous linear simultaneous equations which is generally but not always infinite. A nontrivial solution of these equations is possible only if the determinant of the coefficients is zero. Because the coefficients involve the eigenvalue λ, this condition yields an equation, the secular equation, which determines the possible values of λ. The determinant is known as the secular determinant. An example follows. Let the solutions of the unperturbed problem be $\phi^{(p)}$ with eigenvalues $\lambda^{(p)}$; that is, $A_0\phi^{(p)} = \lambda^{(p)}\phi^{(p)}$. Moreover, suppose that the set $\phi^{(p)}$ is complete, which roughly means that an arbitrary function can be represented as a linear combination of $\phi^{(p)}$. Therefore let ψ, the solution of the perturbed problem, be $C_0\phi^{(0)} + C_1\phi^{(1)} + C_2\phi^{(2)} + \cdots$, where C_p are constants. By substituting this expression for ψ in the equation $A\psi = \lambda\psi$ and employing the properties of the set $\phi^{(p)}$ which follow from the nature of the operator A_0, it is possible to obtain a set of equations for C_p. In a typical case these equations have the form:

$$C_0(\lambda - \lambda^{(0)}) + C_1(\epsilon A_1)_{10} + C_2(\epsilon A_2)_{20} + \cdots = 0$$

$$C_0(\epsilon A_1)_{01} + C_1(\lambda - \lambda^{(1)}) + C_2(\epsilon A_1)_{21} + \cdots = 0$$

$$C_0(\epsilon A_1)_{02} + C_1(\epsilon A_1)_{12} + C_2(\lambda - \lambda^{(2)}) = 0$$

and so on. The elements $(\epsilon A_1)_{pq}$ are numbers which depend upon $\phi^{(p)}$, $\phi^{(q)}$, and the operator ϵA_1. The consequent secular equation is obtained by setting the determinant of the coefficients of C_p in this sequence of equations equal to zero. The solution of these simultaneous equations for the coefficients C_p can be obtained by the iteration method, which yields a particular representation of each of the approximations ψ_n. If there are only a finite number of C_p, the secular determinant is of finite order and reduces to a finite polynomial in λ so that in the finite case solutions of the secular equation can always be obtained without using perturbation methods. Once allowed values of λ are known, corresponding values of C_p can be determined. *See* EIGENVALUE (QUANTUM MECHANICS).

Degenerate perturbation theory. A special technique is required when the unperturbed problem is degenerate, that is, when there are several solutions of the equation $A_0\phi = \lambda\phi$ for a single value of the eigenvalue λ. The number of such independent solutions is the order of the degeneracy. The corresponding method is designated degenerate perturbation theory. The objective of the special method adopted for this case is the determination of the appropriate linear combinations of these degenerate solutions for use as the initial approximation, ψ_0, in the iterative method. To this end, all terms of the equations for C_p are dropped when p refers to an unperturbed solution which is not one of the degenerate solutions under consideration, and only those C_p which do refer to these degenerate solutions are retained. The resulting secular equation for λ has a number of roots equal to the order of the degeneracy. For each root there is a corresponding set of values for C_p which determine a particular linear combination of the degenerate unperturbed solutions. Each of these combinations can be employed as the initial approximation, ψ_0, in the iterative method. It is often the case that the determination of the possible ψ_0 is sufficient for the evaluation of the major effects of the perturbation. Herman Feshbach

Bibliography. H. Baumgartel, *Analytic Perturbation Theory for Matrices and Operators*, 1985; E. J. Hinch, *Perturbation Methods*, 1991; J. Kevorkian and J. D. Cole, *Perturbation Methods in Applied Mathematics*, 1991; P. M. Morse and H. Feshbach, *Methods of Theoretical Physics*, 1953; A. H. Nayfeh, *Introduction to Perturbation Techniques*, 1981, reprint 1993.

Perturbation (quantum mechanics)

An expansion technique useful for solving complicated quantum-mechanical problems in terms of solutions for simple problems. Perturbation theory in quantum mechanics provides an approximation scheme whereby the physical properties of a system, modeled mathematically by a quantum-mechanical description, can be estimated to a required degree of accuracy. Such a scheme is useful because very few problems occurring in quantum mechanics can be solved analytically. Consequently an approximation technique must be employed in order to give an approximate analytic solution or to provide suitable algorithms for a numerical solution. Even for problems which admit an exact analytic solution, the exact solution may be of such mathematical complexity that its physical interpretation is not apparent. For these situations, perturbation techniques are also desirable. Here the discussion of the application of perturbation techniques to quantum mechanics will be limited to the domain of nonrelativistic quantum theory. Applications of a similar but mathematically more intricate nature have also been made in quantum electrodynamics and quantum field theory. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; QUANTUM MECHANICS.

Perturbation theory is applied to the Schrödinger equation, $H\Psi = (H_0 + \lambda V)\Psi = i\hbar(\partial/\partial t)\Psi$ [where $\hbar$ is Planck's constant h divided by 2π, and $(\partial/\partial t)$ represents partial differentiation with respect to the time variable t], for which the exact hamiltonian H is split into two parts: the approximate (unperturbed) time-independent hamiltonian H_0 whose solutions of the corresponding Schrödinger equation are known analytically, and the perturbing potential λV. The basic idea is to expand the exact solution Ψ in terms of the solution set of the unperturbed hamiltonian H_0 by means of a power series in the coupling constant λ. Such a procedure is expected to be successful

if the system characterized by the unperturbed hamiltonian closely resembles that characterized by the exact hamiltonian. Supposedly the differences are not singular in character, but change as a continuous function of the parameter λ.

Perturbation theory is used in two contexts to provide information about the state of the system, which in quantum mechanics is determined by the wave function Ψ. If λV is time-independent, an objective may be to find the stationary states of the system Ψ_n whose time dependence is given by $\exp(-iE_n t/\hbar)$, where $i = \sqrt{-1}$ and E_n represents the energy of the stationary state labeled by n. If λV is either time-independent or time-dependent, an objective may be to find the time evolution of a state which at some specified time was a stationary state of the unperturbed hamiltonian. The perturbing potential is then considered as causing transitions from the original state to other states of the unperturbed hamiltonian, and application of time-dependent perturbation theory provides the probability of such transitions.

Time-independent perturbation theory. In time-independent perturbation theory, the stationary state $\Psi_n = \psi_n \exp(-iE_n t/\hbar)$, which satisfies the time-independent Schrödinger equation (1), is solved by

$$(H_0 + \lambda V)\psi_n = E_n \psi_n \tag{1}$$

a perturbation technique employing the normalized set of eigenfunctions of H_0, Eq. (2). The approach

$$H_0 \phi_m = \epsilon_m \phi_m \tag{2}$$

used in Rayleigh-Schrödinger perturbation theory, which was first utilized by Lord Rayleigh in his study of acoustic modes of vibration, is to expand both the eigenfunctions ψ_n and the eigenvalues E_n in a power series in λ, as in Eqs. (3) and (4), where s is a label that

$$\psi_n = \sum_{m=0}^{\infty} \lambda^m \psi_n^{(m)} = \sum_{m=0}^{\infty} \lambda^m \sum_s a_{ns}^{(m)} \phi_s \tag{3}$$

$$E_n = \sum_{m=0}^{\infty} \lambda^m E_n^{(m)} \tag{4}$$

ranges over the complete set of the unperturbed set ϕ_s [the symbolic sum over s representing a sum over discrete (bound-state) unperturbed eigenfunctions, as well as an integral over continuum unperturbed eigenfunctions], and $a_{ns}^{(m)}$ are numerical coefficients to be determined. The labels s are arranged so that in the limit $\lambda \to 0$, $E_n \to \epsilon_n$, and $\psi_n \to \phi_n$. If only one eigenfunction ϕ_n has the value ϵ_n as its associated eigenvalue (that is, the eigenvalue is nondegenerate), then equating the coefficients of like powers of λ in the time-independent Schrödinger equation leads to Eqs. (5)–(7), where generically $V_{mn} = \int \phi^*_m V \phi_n$

$$a_{ns}^{(0)} = \delta_{ns} \qquad E_n^{(0)} = \epsilon_n \tag{5}$$

$$a_{nn}^{(1)} = 0 \; a_{ns}^{(1)} = \frac{-V_{sn}}{\epsilon_s - \epsilon_n} \text{ for } n \neq s \tag{6}$$

$$E_n^{(1)} = V_{nn} \qquad E_n^{(2)} = -\sum_m{}' \frac{V_{nm} V_{mn}}{\epsilon_m - \epsilon_n} \tag{7}$$

(integration taking place over all configuration space), and the prime indicates that the $m = n$ term is excluded. The expansion also provides recursion relations from which higher-order contributions may be obtained. The criteria for validity of the expansion is usually taken to be $|a_{ns}^{(1)}| \ll 1$ for $n \neq s$, although formal proofs of convergence exist only in a limited number of cases.

Degenerate perturbation theory. If $N > 1$ zero-order (unperturbed) eigenfunctions have the same value of the eigenvalue ϵ_n, the zeroes in the denominators invalidate the preceding expressions, and one must employ degenerate perturbation theory. Essentially this consists of taking appropriate linear combinations of the zero-order degenerate wave functions for the unperturbed basis. The degenerate set $\{\phi_s\}$ is replaced by the j linear combinations of Eq. (8), and matching like powers of λ then leads to the condition of Eq. (9).

$${}^j\phi_s = \sum_{s'=1}^{N} {}^j b_{ss'} \phi_{s'} \qquad j = 1, 2, \ldots, N \tag{8}$$

$$\sum_{s'=1}^{N} {}^j b_{ss'} \left[V_{ks'} - \delta_{ks'} {}^j E_s^{(1)} \right] = 0 \tag{9}$$

The first-order energy corrections for the jth combination, ${}^jE^{(1)}_s$, are found by setting the determinant of the coefficients of ${}^jb_{ss'}$ equal to zero. This technique, which may be considered as a first-order diagonalization process, together with the normalization condition on the coefficients, gives the coefficients ${}^jb_{ss'}$ for the jth combination of unperturbed wave functions. In general, the first-order shifts ${}^jE_s^{(1)}$ split the original unperturbed energy levels ϵ_s, and higher-order correction formulas contain denominators involving differences of the first-order shifts. If the shifts are the same for two or more different values of j, a linear combination of the corresponding ${}^j\phi_s$ must then be taken for the unperturbed starting basis for the second-order treatment. The appropriate combination is determined by a second-order diagonalization process. Higher-order degeneracies are removed in a similar fashion.

Brillouin-Wigner expansion. An alternative approach to time-independent perturbation theory is provided by the Brillouin-Wigner expansion. In this approach the exact time-independent wave function ψ_n is expanded in terms of unperturbed eigenfunctions

$$\psi_n = \sum_s a_{ns} \phi_s$$

which leads to relation (10). The basic idea here is

$$a_{nm}(E_n - \epsilon_m) = \lambda \sum_s a_{ns} V_{ms} \tag{10}$$

to use this expression to generate an iterative expansion for the off-diagonal coefficients a_{ns} (that is, for $n \neq s$). It follows that Eq. (11) holds, where the

$$E_m - \epsilon_m = \lambda V_{mm} + \lambda (a_{mm})^{-1} \sum_s{}' a_{ms} V_{ms} \tag{11}$$

prime indicates only a sum over off-diagonal a's, and

the expansion is provided by repeated iteration of Eq. (12). The convergence of the Brillouin-Wigner

$$\frac{a_{ms}}{a_{mm}} = \frac{\lambda V_{sm}}{E_m - \epsilon_s} + \sum_k{}' \frac{\lambda a_{mk} V_{sk}}{E_m - \epsilon_s} \qquad (12)$$

series is usually much faster than the convergence of the corresponding Rayleigh-Schrödinger series, but since the energy levels E_m are given only implicitly (to a given order in terms of a solution of a polynomial), the Brillouin-Wigner series expansion may be less convenient, although formally the expansion is much simpler.

Applications. Among the problems for which time-independent perturbation theory is employed are estimation of level splitting in the Zeeman effect, calculation of electric and magnetic susceptibilities, and a host of other problems concerning energy-level determination in atomic and molecular physics. Special techniques are employed in many-body perturbation theory in which a product basis of single-particle unperturbed eigenfunctions is utilized.

WKB method. An approximation technique for solving the time-independent Schrödinger equation by exploiting the classical limit of quantum mechanics is the WKB (Wentzel-Kramers-Brillouin) method. Here one substitutes $\psi = A \exp [iW/\hbar]$ into Schrödinger's equation, and expands the amplitude A and phase W in powers of h. In the zeroth approximation as $h \to 0$, the phase W satisfies the classical Hamilton-Jacobi equation; higher-order contributions give the peculiarly quantum-mechanical effects. The applicability of the WKB approximation is usually confined to one-dimensional potential well or barrier problems, or three-dimensional problems for which the well or barrier depends only on the radial coordinate. The WKB approximation is expected to give reliable results when $d|\lambda(x)|/dx \ll 1$, $\lambda(x)$ being the local de Broglie wavelength $h/p = h\,[2m(E - \lambda V)]^{-1/2}$. The WKB method has been used to give an approximate treatment of alpha-particle emission from a nucleus, as well as of a number of other problems where penetration of a barrier is involved. *See* HAMILTON-JACOBI THEORY; WENTZEL-KRAMERS-BRILLOUIN METHOD.

Time-dependent perturbation theory. In time-dependent perturbation theory, where the perturbing potential may or may not be time-dependent, the time-dependent wave function Ψ is expanded in terms of the set of time-dependent unperturbed eigenfunctions by means of expansion coefficients which are themselves time-dependent. This method is sometimes known as the method of variation of constants. Specifically one expresses a particular state, labeled by n, by means of Eq. (13). The coefficients c then obey the rate relation, Eq. (14), where

$$\Psi_n = \sum_s c_{ns}(t)\phi_s \exp \frac{-i\epsilon_s t}{\hbar} \times \exp\left[\frac{-i\lambda}{\hbar}\int_{t_0}^{t} V_{ss'}\,dt'\right] \qquad (13)$$

$$i\hbar\frac{\partial c_{nm}}{\partial t} = \lambda \sum_{s\neq m} c_{ns} V_{ms} \exp\left[\frac{i}{\hbar}\int_{t_0}^{t} \gamma_{ms}\,dt'\right] \qquad (14)$$

$\gamma_{ms} = (\epsilon_m + \lambda V_{mm}) - (\epsilon_s + \lambda V_{ss})$ is the difference of the perturbed energies for the states m and s. The coefficient c_{ns} is defined in Eq. (13) so as not to include the phase involving the integral over V_{ss}, in order to prevent so-called secular terms (those which have a nonoscillatory behavior at large times) from appearing in an iterative expansion of the coefficients. The appearance of such terms would inhibit convergence of the resulting expansion. Usually, in time-dependent problems, one assumes that at time $t = t_0$ the system is in state n so that $c_{ns}(t_0) = \delta_{ns}$. The problem then is to determine the probability of transition to other states at a later time.

Transient perturbations. For transient perturbations, that is, those for which $V \to 0$ for $t \to \pm\infty$, the initial time t_0 may be taken as $-\infty$, and it is clear that as $t \to \infty$, the expansion coefficients of Eq. (15) be-

$$a_{ns}(t) = c_{ns}(t) \exp\left[\frac{-i\lambda}{\hbar}\int_{t_0}^{t} V_{ss}\,dt'\right] \qquad (15)$$

come constant since the ϕ's are eigenfunctions of the final hamiltonian operator. Calculations of $a_{ns}(\infty)$ to successive orders in λ, leading to what is called successive Born approximations, give the long-term excitation probability $P_{nm}(\infty) = |a_{nm}(\infty)|^2$ for excitation from initial state n to final state m. Problems such as scattering of particles in a continuum state from atomic systems, and inelastic collisions of atomic systems, are amenable to this type of treatment. *See* SCATTERING EXPERIMENTS (ATOMS AND MOLECULES).

Persistent perturbations. When a persistent perturbation is present, the perturbing potential which was originally zero ultimately reaches a nonvanishing time-independent limit. A first-order solution for the coefficients c gives Eqs. (16), where the notation of Eqs. (17) and (18) is used. The coefficients $\beta_{nm}(t)$

$$c_{nn}(t) = 1 \qquad c_{nm}(t) = \alpha_{nm}(t) + \beta_{nm}(t) \qquad (16)$$
$$n \neq m$$

$$\alpha_{nm}(t) = \int_{t_0}^{t}\left[\exp\frac{i}{\hbar}\int_{t_0}^{t'}\gamma_{mn}\,dt''\right] \times \frac{d}{dt'}\left[\frac{V_{mn}(t')}{\gamma_{mn}(t')}\right]dt' \qquad (17)$$

$$\beta_{nm}(t) = \frac{-\lambda V_{mn}(t)}{\gamma_{mn}(t)} \exp\left[\frac{i}{\hbar}\int_{t_0}^{t}\gamma_{mn}\,dt'\right] \qquad (18)$$

merely provide for the first-order propagation of the initial state in time, whereas the coefficients $\alpha_{nm}(t)$ give the excitation probability $P_{nm}(t) = |\alpha_{nm}(t)|^2$. If the perturbation increases from zero to its asymptotic value in such a short time that the exponential in the definition of α_{nm} is essentially constant, that is, in times shorter than $\hbar/(\epsilon_m - \epsilon_n)$, then $|\alpha_{nm}(\infty)|$ may be approximated by $|\lambda V_{mn}/\gamma_{mn}|$ where the relevant quantities are evaluated at their asymptotic values. This approximation, known as the sudden approximation, leads to useful results for transition probabilities arising from short-time perturbations, such

as occur in the radioactive decay of ^{3}H atom into a doubly ionized ^{3}He ion.

Adiabatic approximation. Yet another approach, the adiabatic approximation, is useful when the perturbing potential changes very slowly. This technique employs a set of basis wave functions χ_s (with a time-independent phase) and associated eigenvalues $E_s(t)$ satisfying Eq. (19), which would follow for the exact

$$[H_0 + \lambda V(t)]\chi_s = E_s(t)\chi_s \tag{19}$$

solution of the time-dependent Schrödinger equation if the time variation of the potential were neglected. The basic idea is to expand the exact time-dependent wavefunction Ψ_n in terms of this basis, as in Eq. (20), obtain the rate equation giving the

$$\Psi_n = \sum_s c_{ns}(t)\ \chi_s \exp\left[\frac{-i}{\hbar}\int_{t_0}^{t} E_s(t')dt'\right] \tag{20}$$

time-development of the coefficients, and solve for these coefficients to the desired order in λ by an iterative technique, assuming that the coefficients c_{ns} differ little from their initial values. The use of the basis χ rather than an unperturbed basis hopefully provides for a more satisfactory lower-order treatment. This approach has been successfully used in atomic-collision problems for which the interaction distance changes slowly with time. *See* PERTURBATION (ASTRONOMY); PERTURBATION (MATHEMATICS).

David M. Fradkin

Pesticide

A material useful for the mitigation, control, or elimination of plants or animals detrimental to human health or economy. Algicides, defoliants, desiccants, herbicides, plant growth regulators, and fungicides are used to regulate populations of undesirable plants which compete with or parasitize crop or ornamental plants. Attractants, insecticides, miticides, acaricides, molluscicides, nematocides, repellants, and rodenticides are used principally to reduce parasitism and disease transmission in domestic animals, the loss of crop plants, the destruction of processed food, textile, and wood products, and parasitism and disease transmission in humans. These ravages frequently stem from the feeding activities of the pests. Birds, mice, rabbits, rats, insects, mites, ticks, eel worms, slugs, and snails are recognized as pests.

Materials and use. Materials used to control or alleviate disease conditions produced in humans and animals by plants or by animal pests are usually designated as drugs. For example, herbicides are used to control the ragweed plant, while drugs are used to alleviate the symptoms of hay fever produced in humans by ragweed pollen. Similarly, insecticides are used to control malaria mosquitoes, while drugs are used to control the malaria parasites—single-celled animals of the genus *Plasmodium*—transmitted to humans by the mosquito.

Sources. Some pesticides are obtained from plants and minerals. Examples include the insecticides cryolite, a mineral, and nicotine, rotenone, and the pyrethrins which are extracted from plants. A few pesticides are obtained by the mass culture of microorganisms. Two examples are the toxin produced by *Bacillus thuringiensis*, which is active against moth and butterfly larvae, and the so-called milky disease of the Japanese beetle produced by the spores of *B. popilliae*. Most pesticides, however, are products which are chemically manufactured. Two outstanding examples are the insecticide DDT and the herbicide 2,4-D.

Evaluation. The development of new pesticides is time-consuming. The period between initial discovery and introduction is frequently cited as being about 5 years. Numerous scientific skills and disciplines are required to obtain the data necessary to establish the utility of a new pesticide. Effectiveness under a wide variety of climatic and other environmental conditions must be determined, and minimum rates of application established.

Insight must be gained as to the possible side effects on other animals and plants in the environment. Toxicity to laboratory animals must be measured and be related to the hazard which might possibly exist for users and to consumers. Persistence of residues in the environment must be determined. Legal tolerances in processed commodities must be set and directions for use clearly stated. Methods for analysis and detection must be devised. Economical methods of manufacture must be developed. Manufacturing facilities must be built. Sales and education programs must be prepared.

Regulation and restriction. By the mid-1960s, it became apparent that a number of the new pesticides, particularly the insecticide DDT, could be two-edged swords. The benefits stemming from the unmatched ability of DDT to control insect pests could be counterbalanced by adverse effects on other elements of the environment. Detailed reviews of the properties, stability, persistence, and impact upon all facets of the environment were carried out not only with DDT, but with other chlorinated, organic insecticides as well. Concern over the undesirable effects of pesticides culminated in the amendment of the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) by Public Law 92-516, the Federal Environmental Pesticide Control Act (FEPCA) of 1972. The purpose behind this strengthening of earlier laws was to prevent exposure of either humans or the environment to unreasonable hazard from pesticides through rigorous registration procedures, to classify pesticides for general or restricted use as a function of acute toxicity, to certify the qualifications of users of restricted pesticides, to identify accurately and label pesticide products, and to ensure proper and safe use of pesticides through enforcement of FIFRA.

Selection and use. Pesticides must be selected and applied with care. Recommendations as to the product and method of choice for control of any pest problem—weed, insect, or varmint—are best obtained from county or state agricultural extension specialists. Recommendations for pest control and pesticide use can be obtained from each state

agricultural experiment station office. In addition, it is necessary to follow explicitly the directions, restrictions, and cautions for use on the label of the product container. Insecticides are a boon to the production of food, feed, and fiber, and their use must not be abused in the home, garden, farm, field, forest, or stream. George F. Ludvik

Persistence. Considerable scientific evidence has been accumulated that documents the deleterious effects on the environment of several of the organic insecticides. Over 800 compounds are used as pesticides and, with increasing technological sophistication, the list can be expected to grow. Until World War II pesticides consisted of inorganic materials containing sulfur, lead, arsenic, or copper, all of which have biocidal properties, or of organic materials extracted from plants such as pyrethrum, nicotine, and rotenone. During World War II a technological revolution began with the introduction of the synthetic organic biocides, particularly DDT.

There is an increasing number of organic biocides. They vary widely in toxicity, in specificity, in persistence, and in the production of undesirable derivatives. The factors that determine how, when, and where a particular pesticide should be used, or whether it should be used at all, are therefore necessarily different for each pesticide. The use of chemicals to achieve a measure of control over the environment has become an integral part of technology, and there are no valid arguments to support the proposition that the use of all of these chemicals should cease.

Some take the view that pests include insects, fungi, weeds, and rodents. Many insects, however, particularly the predatory species that prey upon the insects that damage crops, are clearly not pests. Fungi are essential in the processes that convert dead plant and animal material to the primary substances that can be recycled through living systems. In many ecosystems rodents are also an important link in the recycling process. On the other hand, an increasing number of species traditionally considered desirable are becoming "pests" in local contexts. Thus chemicals are used to destroy otherwise useful vegetation along roadsides or in pine plantations, or vegetation that might provide food and shelter to species undesirable to human populations. The control of any component of the global environment through the use of biocidal chemicals is therefore a problem much more complex than is initially suggested by the connotation of the word pest. The highly toxic chemicals that kill many different species are clearly undesirable in situations in which control over only one or two is sought.

Other pest-control measures. Sophisticated methods of pest control are continually being developed. One technique involves the raising of a large number of insect pests in captivity, sterilizing them with radioactivity, and then releasing them into the environment, where they mate with the wild forms. Very few offspring are subsequently produced. Highly specific synthetic insect hormones are being developed. In an increasing number of pest situations, a natural predator of an insect has been introduced, or conditions are maintained that favor the propagation of the predator. The numbers of the potential pest species are thereby maintained below a critical threshold. An insect control program in which use of insecticides is only one aspect of a strategy based on ecologically sound measures is known as integrated pest management. *See* AGRICULTURAL CHEMISTRY; CHEMICAL ECOLOGY; FUNGISTAT AND FUNGICIDE; HERBICIDE; INSECT CONTROL, BIOLOGICAL; INSECTICIDE. Robert W. Risebrough

Bibliography. A. J. Burn, T. H. Coaker, and P. C. Jepson (eds.), *Integrated Pest Management*, 1988; R. Carson, *Silent Spring*, 25th anniversary ed., 1987; G. P. Georghiou and T. Saito (eds.), *Pest Resistance to Pesticides*, 1983; R. Honeycutt, *Advances in Understanding the Mechanics of Pesticide Movement*, 1994; P. Hurst, N. Dudley, and A. Hay, *The Pesticide Handbook*, 1989; R. L. Metcalf and W. Luckmann (eds.), *Introduction to Insect Pest Management*, 3d ed., 1994; H. F. Van Emden, *Pest Control: New Studies in Biology*, 2d ed., 1989; G. W. Ware, *The Pesticide Book*, 5th ed., 1999.

Petalite

A rare pegmatitic mineral with composition $LiAlSi_4O_{10}$. Its economic significance is markedly disproportionate to the number of its occurrences. It is the only basic raw material suitable for production of a group of materials known as crystallized glass ceramics (melt-formed ceramics). These extremely fine-grained substances, formed by controlled nucleation and crystal growth, are based on a keatite-type structure (stuffed silica derivative). Among the desirable properties of such submicroscopic aggregates are their exceedingly low thermal expansion and high strength, making them suitable for use in cooking utensils and telescopic mirror blanks.

Petalite is mined from a large lithium-rich pegmatite at Bikita, Rhodesia, the only major world source. Other significant occurrences are at Varuträsk, Sweden (largely exhausted); Karibib district, South-West Africa; and Londonderry, Western Australia. About 30 other localities are reported, chiefly for specimens only.

Petalite is characterized as follows: white, pink, pale green, or gray to black; monoclinic (Pa); hardness $6^1/_2$; optically (+); $\alpha = 1.504$–1.507, $\beta = 1.510$–1.513, $\gamma = 1.516$–1.523, $\gamma - \alpha = 0.011$–0.017, $2V = 82$–84° optic plane $\{001\}$; $\beta \wedge Z = 24$–30°; cleavages, $\{001\}$ perfect, $\{201\}$ good, lamellar twins $\{001\}$ common.

In pegmatites it is associated with other lithium minerals such as spodumene, lepidolite, eucryptite, and amblygonite, and common also with the rare cesium silicate, pollucite. It is an early-formed pegmatitic species and may be converted isochemically to a spodumene-quartz intergrowth via the reaction

$$\underset{\text{Petalite}}{LiAl_4O_{10}} \rightarrow \underset{\text{Spodumene}}{LiAlSi_2O_8} + \underset{\text{Quartz}}{2SiO_2}$$

See LITHIUM; PEGMATITE. E. William Heinrich

Petrifaction

One of the most remarkable mechanisms by which the remains of extinct organisms are preserved in the fossil record. In petrifactions (though chiefly in the case of plants rather than animals) the original shape and topography of the tissues, and occasionally even minute cytological details, are retained relatively undeformed.

Formation. The term petrifaction was adopted as a scientific term before knowledge existed of the geochemical mechanism or processes involved. Initially, it was believed that in the formation of a petrifaction the organic matter of the organism or tissue was replaced molecule by molecule with mineral material entering in solution in percolating ground water. It is now evident that what actually happens is that the mineral fills cell lumena and the intermicellar interstices of cell walls with insoluble salts depositing from solution. Petrifaction is hence a form of mineral emplacement or embedding, by which the organic residues are filled with solid substance which infiltrates in solution. The most common substances involved in petrifactions are silica, SiO_2, and calcium carbonate, $CaCO_3$ (calcite). Occasionally phosphate minerals, pyrite, hematite, and other less common minerals make up all or part of the petrifaction matrix. The most perfect preservation of original structure is found in siliceous petrifactions. The clear, transparent, microcrystalline silica renders excellent optical properties to thin sections of such fossils and makes possible the use of transmitted light in microscopic study, in much the same manner as with recent biological material.

Calcified types. An unusual type of calcareous petrifaction known as coal balls occurs in Carboniferous coal seams of parts of Europe and North America. They comprise nodular, usually spheroidal or ovoidal masses of relatively uncompressed plant tissues completely permeated with calcite or dolomite. They represent irregularly spaced and localized areas of mineral precipitation with resulting petrifaction of the coal-forming plant debris. Mineralization occurred while the coal was still in the peat stage. After mineralization the infiltrated plant parts failed to compress so that their structure and cellular organization are preserved. Much of what is known of the internal organization and anatomy of ancient plants and the evolution of their organs, both vegetative and reproductive, is derived from petrifactions, which occur throughout the geologic record from Precambrian to Recent. Coal balls, though limited to the Carboniferous, have provided an unusually comprehensive body of knowledge on the morphology and anatomy of the rich and varied vegetation of that geologic period. *See* COAL BALLS.

Silicified types. Despite the widespread abundance of petrifactions (chiefly of plants but also of the hard parts of animals) throughout the geologic column, there is very little known of the geochemical mechanisms which induce their formation. The problem is particularly baffling in the case of silicification; in many examples entire trunks of trees are completely mineralized with no visible evidence of the sequence of changes following sedimentation. The fact that silicified plant parts often show virtually no physical distortion or compression indicates that the process occurs relatively rapidly. On the other hand, the amount of silica in groundwater ($\pm$70–100 parts per million) is so small that the process must proceed with extreme efficiency. It is probable that decay processes proceeded to at least the early stages, since the histology of silicified (also calcified) plant tissues exhibits varying stages of degradative alteration, to nearly total loss of structure. The percentage by weight of organic matter retained in silicified wood may range from a few hundredths of 1% to more than 15%. Ordinarily it is only a few percent. *See* FOSSIL; PALEOBOTANY; PETRIFIED FORESTS.

Elso S. Barghoorn

Petrified forests

Concentrations of appreciable numbers of petrified tree trunks exposed at the earth surface, standing upright in the position of growth, lying prostrate, or both. They typically consist of the mineralized remains of woody plants and sometimes are called fossil forests (**Figs. 1** and **2**). *See* FOSSIL; PETRIFACTION.

Although petrified wood is relatively common, petrified forests are quite rare because special circumstances, such as rapid burial of living or dead trees, are necessary for them to form. Recent studies indicate that only about two dozen are presently known to science. However, as time passes, more

Fig. 1. Tall petrified tree trunks on Specimen Ridge in Yellowstone National Park, Wyoming. (*Photo by William W. Dunmire/NPS*)

may be discovered in remote parts of the world. For example, there are unconfirmed reports of forests in parts of Russia, and there are areas in Africa and South America that have not been thoroughly explored by geologists. The known forests are scattered from the Antarctic to the Arctic but are most common in the midlatitudes of the Northern Hemisphere (**Fig. 3**). They range in age from the Devonian (about 390 million years old) to the Neogene (about 3 million years old), but they do not become common until the Mesozoic about 248 million years ago (**Fig. 4**). [Note that the bold numbers in parentheses that follow in the text refer to the designated locations found in Figs. 3 and 4.] Many of the forests such as the Late Jurassic forest in Xinjiang, China (**10**), and the Middle Triassic Fremouw forest in Antarctica (**15**) were preserved in sedimentary rocks deposited by rivers, but a few, such as the Chemnitz forest in Germany (**17**) and the forest in Florissant Fossil Beds National Monument in Colorado (**4**), were preserved in stratified volcanic rocks that originated from adjacent volcanoes. Most of the forests consist of petrified trunks of woody plants, such as those of conifers, as in the Late Jurassic Cerro Cuadrado, Argentina (**11**), and Early Cretaceous Alexander Island forest in Antarctica (**8**); but angiosperms dominate in most Cretaceous and younger forests, including the Late Neogene Calistoga, California (**1**), and Late Paleogene Cairo, Egypt (**3**), forests. Notable exceptions are the Middle Devonian Gilboa, New York (**19**), and Late Mississippian Victoria Park, Glasgow, Scotland (**18**), forests that contain stump casts of nonwoody treelike lycopods, and the Permian Tocantins, Brazil (**16**), and Chemnitz, Germany (**17**), forests that contain numerous petrified trunks of tree ferns and other nonwoody plants. There also are small concentrations of petrified wood of various types scattered about the world that range from Pennsylvanian to Neogene in age and are informally called petrified forests, but they are not considered here because they do not meet the definition of petrified forests used here. *See* PALEOBOTANY.

Fig. 2. Petrified wood log tumble in Petrified Forest National Park, Arizona. (*Photo by T. Scott Williams/NPS*)

Fig. 3. Map showing the location of the principal petrified forests listed in the stratigraphic chart of Fig. 4 and mentioned in the text. Note that the location of the Fremouw forest in east Antarctica is not visible on this map projection.

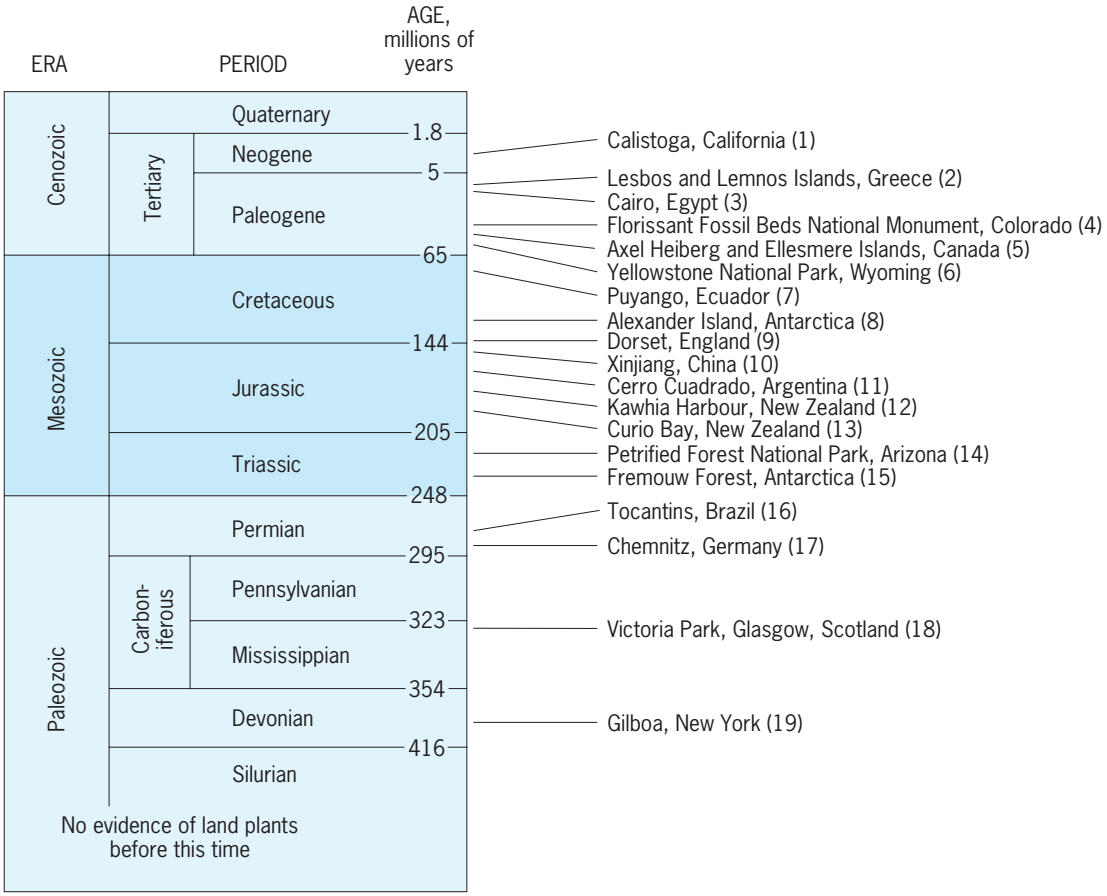

Fig. 4. Chart showing the stratigraphic position of the known principal petrified forests. The numbers in parentheses following the forest names correlate with the numbers showing the forest locations on the map of Fig. 3.

Petrified Forest National Park. Petrified forests are significant because they contain the remains of past life—particularly trees—usually preserved in or close to the terrestrial setting that they once inhabited and sometimes even in the position of growth. Also they often are associated with the remains of coexisting invertebrates and vertebrates, which permits a more complete reconstruction of ancient environments. Furthermore, they help establish the paleoclimates of the time and demonstrate the profound climatic changes that have taken place in the past. *See* CLIMATE MODIFICATION; PALEOCLIMATOLOGY.

A petrified forest that illustrates this concept particularly well is in the Petrified Forest National Park (**14**), which covers about 218,000 acres (88,200 hectares) in east-central Arizona. It is about 225 million years old and occurs at several horizons in the Chinle Formation of Late Triassic age, which was deposited by sluggish rivers in a broad low-lying swampy basin and in lakes. At that time, this part of Arizona was much closer to both the Equator and the Pacific Ocean, and the paleoclimate was subtropical with abundant rainfall, which may have been somewhat seasonal although annual rings have not been observed in the petrified wood. Many of the floral and faunal elements of the ancient terrestrial ecosystem that coexisted with the woody trees in the forest are also preserved in the Chinle Formation in the park. Much of the wood has been replaced by quartz and is brightly colored in tones of red, pink, brown, yellow, and even purple and blue and is called Rainbow Wood. Some of the wood, however, is more somber in color and even black. The forest itself includes thousands of prostrate silicified trunks lying exposed on the surface of the ground, but only about 100 stumps still stand in the position of growth. Also, the ground surface between the logs and stumps is generally covered with smaller fragments of petrified wood. Furthermore, as time passes, additional logs are exposed by erosion, and weathering continues to slowly break down previously exposed wood into smaller fragments. The logs range up to about 120 ft (36 m) in length and 10 ft (3 m) in diameter at the base, but most logs are shorter, usually somewhat less than 100 ft (30 m) in length and only about 4 ft (1.2 m) in diameter at the base. *See* QUARTZ.

The plant fossils found in the Chinle Formation appear to represent mainly local streamside environments, although there are also more limited deposits that probably originated in small swamps. Most of the prostrate logs (Fig. 2) in the Chinle Formation apparently are the remains of trees that grew along the banks of the ancient streams not far from where their remains are now found. It is possible that a few of the logs were carried a somewhat longer distance before burial, but probably they do not represent a

distant upland environment. Shortly after they fell into the streams, some of the logs became tangled with each other, as commonly happens in modern streams, and they formed log jams; but more often they were washed out onto the adjacent floodplains, where they were buried by the next flood and fossilized. The majority of logs in the Chinle Formation are assigned to *Araucarioxylon arizonicum*, a conifer of uncertain affinities. However, the petrified remains of other types of trees occur here, including other conifers, cycads, ginkgoes, and a tree fern. In addition, abundant leaf and stem compressions of horsetails, ferns, seed ferns, cycads, cycadeoids, and conifers as well as their reproductive structures such as cones, seeds, pollen, and spores occur in beds of mudstone adjacent to the deposits of wood. Small particles of amber also occur in the park, with at least some being derived from the dominant tree preserved in the park, *A. arizonicum*. The remains of the other elements of the ancient streamside environment that coexisted with the trees and are found here include clams, snails, beetles, other insects, crayfish, horseshoe crabs, fish, and a variety of amphibians and reptiles, including some of the earliest known dinosaurs in the world. It is apparent from the fossils found here that this area was characterized by an understory of horsetails, ferns, cycads, and scattered trees of several types and was inhabited by a large variety of both invertebrates and primitive vertebrates, as well as dinosaurs. To an extent, the area was probably similar to that of the Florida Everglades. *See* TREE; TREE FERNS.

Axel Heiberg and Ellesmere Islands. Petrified forests that contrast strongly with the one in Arizona were recently described from Axel Heiberg and Ellesmere Islands (**5**) in Canada's Arctic Archipelago about 700 mi (1100 km) from the North Pole. They are much younger than the Arizona forest, and the wood generally has not been replaced by quartz. Instead, in one forest the wood has been replaced by the mineral calcite, and in many others the wood is brown to black in color and has been naturally mummified so that it is flexible and still burns easily. Also, the forests occur over a much broader area and at many more horizons, perhaps as many as 30, and the forest floor in which they were rooted is still preserved. They are found in the Buchanan Lake Formation of Early Paleogene age and are about 45 million years old and consist primarily of the remains of woody trees that lived in shallow wetlands on the floors of a broad basin. The standing stumps are the remains of tall trees as some are as much as 10 ft (3 m) in diameter, which implies that the living trees were as tall as 100 ft (30 m) or more. It appears that most of the stumps represent the dawn redwood *Metasequoia* and the Chinese swamp cypress *Glyptostrobus*, although other types of trees also may have been present. A layer of slightly decomposed leaf litter covers the soil in which the stumps are rooted, and consists of the beautifully preserved foliage and cones of these trees together with the needles of several other types of conifers such as pine and spruce, as well as the broad leaves of deciduous angiosperms such as alder and birch. In addition, the delicate remains of fern leaves and mosses, as well as pollen and spores of these and many other plants, are found in the leaf litter. Adjacent beds of rock contain the remains of invertebrate and vertebrates that existed at the same time the plants were living and show that the swamps were inhabited by alligators, crocodiles, turtles, snakes, and primitive mammals.

It is evident from all of these fossils that the paleoclimate was warm and subtropical, similar to that of the Florida Everglades, during the Early Paleogene when the forests of Axel Heiberg and Ellesmere Island were living. However, at that time the area of the two islands was near its present geographic position, and the plants and animals were adapted to living with long periods of darkness. For example, the dominant trees of the area, the dawn redwood, swamp cypress, and so on, were deciduous and lost their leaves as the darkness set in, and thus preserved their stored food supplies until spring when they began to grow rapidly once again and develop new foliage in the long sunny days of the Arctic summer. Since that time, general cooling of the Earth caused the extinction of what was once an extensive lush Arctic forest. The presence of these subtropical forests in the high Arctic 45 million years ago clearly illustrates the dramatic climatic changes that the Earth has undergone in the past.

Southern Hemisphere petrified forests. The petrified forests of the Southern Hemisphere generally are not too well known because most are so remote from population centers that casual visitors have difficulty in reaching them. Two such are the Alexander Island forest (**8**) and the Fremouw forest (**15**) in east Antarctica which are accessible only during part of the year and even then require special transportation. These forests include standing stumps and prone trunks of coniferous trees that were adapted to living in a paleoclimate characterized by several months of total darkness each year. Probably the most famous and best-known petrified forest in this part of the world is the Cerro Cuadrado forest (**11**) near the southern tip of South America, in Patagonia, Argentina, which was not discovered by the scientific community until the second decade of the twentieth century. This forest has attracted the attention of both scientists and tourists alike because it contains not only long prone trunks and standing stumps of tall trees but all other parts of the trees, including seed cones, seeds, foliage, and leafy twigs. The petrified material is so well preserved that it is possible to study the internal cellular structure of the fossils. The forest is Late Jurassic in age, and most of the plant fossils preserved in it are the remains of the distinctive tree *Araucaria*, which is common now in the Southern Hemisphere.

Sidney R. Ash

Bibliography. S. R. Ash, *Petrified Forest: A Story in Stone*, Petrified Forest Museum Association, Petrified Forest, AZ, 2005; J. F. Basinger, The fossil forests of the Buchanan Formation (early Tertiary), Axel Heiberg Island, Canadian Arctic Archipelago: Preliminary floristics and paleoclimate, *Geol. Surv. Can. Bull.*, no. 403, pp. 39–65, Ottawa, 1991;

M. G. Calder, A coniferous petrified forest in Patagonia, *Brit. Mus.* (*Nat. Hist.*) *Bull.*, *Geol.*, 2:99–138, London, 1953; U. Dernbach, *Petrified Forests: The World's 31 Most Beautiful Petrified Forests*, D'Oro-Verlag, Heppenheim, Germany, 1996; H. W. Meyer, *The Fossils of Florissant*, Smithsonian Books, Washington, DC, 2003; W. G. Parker, S. R. Ash, and R. B. Irmis (eds.), A century of research at Petrified Forest National Park, 1905–2006: Geology and paleontology, *Mus. North. Ariz. Bull.*, no. 62, Flagstaff, 2006.

Petrochemical

Any of the chemicals derived from petroleum or natural gas. The definition of petrochemicals has been broadened to include the whole range of aliphatic, aromatic, and naphthenic organic chemicals, as well as carbon black and such inorganic materials as sulfur and ammonia.

Petrochemicals are made or recovered from the entire range of petroleum fractions, but the bulk of petrochemical products are formed from the lighter (C_1–C_4) hydrocarbon gases as raw materials. These materials generally occur in natural gas, but they are also recovered from the gas streams produced during refinery operations, especially cracking. Refinery gases are particularly valuable because they contain substantial amounts of olefins that, because of their double bonds, are much more reactive then the saturated (paraffin) hydrocarbons. Also important as raw materials are the aromatic hydrocarbons (benzene, toluene, and xylene) that are obtained from various refinery product streams. For example, catalytic reforming processes convert nonaromatic hydrocarbons to aromatic hydrocarbons by dehydrogenation and cyclization. *See* PETROLEUM; PETROLEUM PRODUCTS.

Thermal cracking processes (such as coking) are focused primarily on increasing the quantity and quality of gasoline and other liquid fuels, but also produce gases, including lower-molecular-weight olefins such as ethylene ($CH_2{=}CH_2$), propylene ($CH_3CH{=}CH_2$), and butylenes (butenes, $CH_3CH{=}CHCH_3$ and $CH_3CH_2CH{=}CH_2$). Catalytic cracking is a valuable source of propylene and butylene, but it is not a major source of ethylene, the most important of the petrochemical building blocks. *See* CRACKING; ETHYLENE.

The starting materials for the petrochemical industry (**Table 1**) are obtained from crude petroleum in one of two ways. They may be present in the raw crude oil and are isolated by physical methods, such as distillation or solvent extraction; or they are synthesized during the refining operations. Unsaturated (olefin) hydrocarbons, which are not usually present in natural petroleum, are nearly always manufactured as intermediates during the refining sequences (**Table 2**). *See* DISTILLATION; PETROLEUM PROCESSING AND REFINING; SOLVENT EXTRACTION.

The main objective in producing chemicals from petroleum is the formation of a variety of well-defined chemical compounds, including (1) chemicals from aliphatic compounds; (2) chemicals from olefins; (3) chemicals from aromatic compounds; (4) chemicals from natural gas; (5) chemicals from synthesis gas (carbon monoxide and hydrogen); and (6) inorganic petrochemicals.

TABLE 1. Hydrocarbon intermediates used as sources of petrochemical products

Carbon number	Hydrocarbon type: Saturated	Unsaturated	Aromatic
1	Methane		
2	Ethane	Ethylene Acetylene	
3	Propane	Propylene	
4	Butanes	*n*-Butenes Isobutene Butadiene	
5	Pentanes	Isopentenes (isoamylenes) Isoprene	
6	Hexanes Cyclohexane	Methylpentenes	Benzene
7		Mixed heptenes	Toluene
8		Di-isobutylene	Xylenes Ethylbenzene Styrene
9			Cumene
12		Propylene tetramer Tri-isobutylene	
18			Dodecylbenzene
6–18		*n*-Olefins	
11–18	*n*-Paraffins		

TABLE 2. Major sources of petrochemical intermediates

Hydrocarbon	Source
Methane	Natural gas
Ethane	Natural gas
Ethylene	Cracking processes
Propane	Natural gas, catalytic reforming, cracking processes
Propylene	Cracking processes
Butane	Natural gas, reforming and cracking processes
Butene(s)	Cracking processes
Cyclohexane	Distillation
Benzene Toluene Xylene(s) Ethylbenzene	Catalytic reforming
Alkylbenzenes	Alkylation
$>C_9$	Polymerization

An aliphatic petrochemical is an organic compound that has a chain of carbon atoms, be it normal (straight), such as *n*-pentane ($CH_3CH_2CH_2CH_2CH_3$); or branched, such as, *iso*-pentane (2-methylbutane, $CH_3CHCH_3CH_2CH_3$); or unsaturated. Unsaturated compounds, called olefins, include important starting materials such as ethylene ($CH_2{=}CH_2$), propylene ($CH_3CH{=}CH_2$), 1-butene ($CH_3CH_2CH_2{=}CH_2$), *iso*-butylene [$(CH_3)_2C{=}CH_2$], and butadiene ($CH_2{=}CHCH{=}CH_2$).

An aromatic petrochemical is an organic compound that contains, or is derived from, a benzene ring.

A significant proportion of the basic petrochemicals are converted into plastics, synthetic rubbers,

TABLE 3. Chemicals produced from methane (natural gas)

Basic derivatives and sources	Uses
Carbon black	Rubber compounding Printing ink, paint
Methanol	Formaldehyde (mainly for resins) Methyl esters (polyester fibers), amines, and other chemicals Solvents
Chloromethanes	Chlorofluorocarbons for refrigerants, aerosols, solvents, cleaners, grain fumigant

TABLE 4. Chemicals from ethylene

Basic derivatives and sources	Uses
Ethylene oxide	Ethylene glycol (polyester fiber and resins, antifreeze) Di- and triethylene glycols Ethanolamines Nonionic detergents Glycol esters
Ethyl alcohol	Acetaldehyde Solvents Ethyl acetate
Polyethylene	
Low-density	Film, injection molding
High-density	Blow molding, injection molding
Styrene (from ethylene and benzene)	Polystyrene and copolymer resins Styrene-butadiene rubber and latex Polyesters
Ethylene dichloride	Vinyl chloride Scavenger in antiknock fluid Ethyleneamines
Ethyl chloride	Tetraethyllead Minor amounts for ethylations (such as ethyl cellulose)
Ethylene dibromide	Scavenger in antiknock fiuid
Acetyls	Plastics and chemical intermediates
Linear alcohols and α-olefins	Detergents, plasticizers

and synthetic fibers. These materials, known as polymers, are high-molecular-weight compounds made up of repeated structural units. The major polymer products are polyethylene, poly(vinyl chloride), and polystyrene, all derived from ethylene, and polypropylene, derived from propylene. Major raw material sources for synthetic rubbers include butadiene, ethylene, benzene, and propylene. Among synthetic fibers the polyesters, which are a combination of ethylene glycol and terephthalic acid (made from xylene), are the most widely used. They account for about one-half of all synthetic fibers. The second major synthetic fiber is nylon, its most important raw material being benzene. Acrylic fibers, in which the major raw material is the propylene derivative acrylonitrile, make up most of the remainder of the synthetic fibers. *See* MANUFACTURED FIBER; POLYACRYLATE RESIN; POLYAMIDE RESINS; POLYESTER RESINS; POLYMER; POLYMERIZATION; POLYOLEFIN RESINS; POLYURETHANE RESINS; POLYVINYL RESINS; RUBBER.

Aliphatics. Methane, obtainable from petroleum and natural gas or as a product of various conversion (cracking processes), is an important raw material source for aliphatic petrochemicals (**Table 3**; **Fig. 1**). Ethane, propane, and butane are also important sources of petrochemicals because of their relative ease of conversion to ethylene, which provides more valuable routes to petrochemicals (**Table 4**; **Fig. 2**). The higher-boiling members of the paraffin

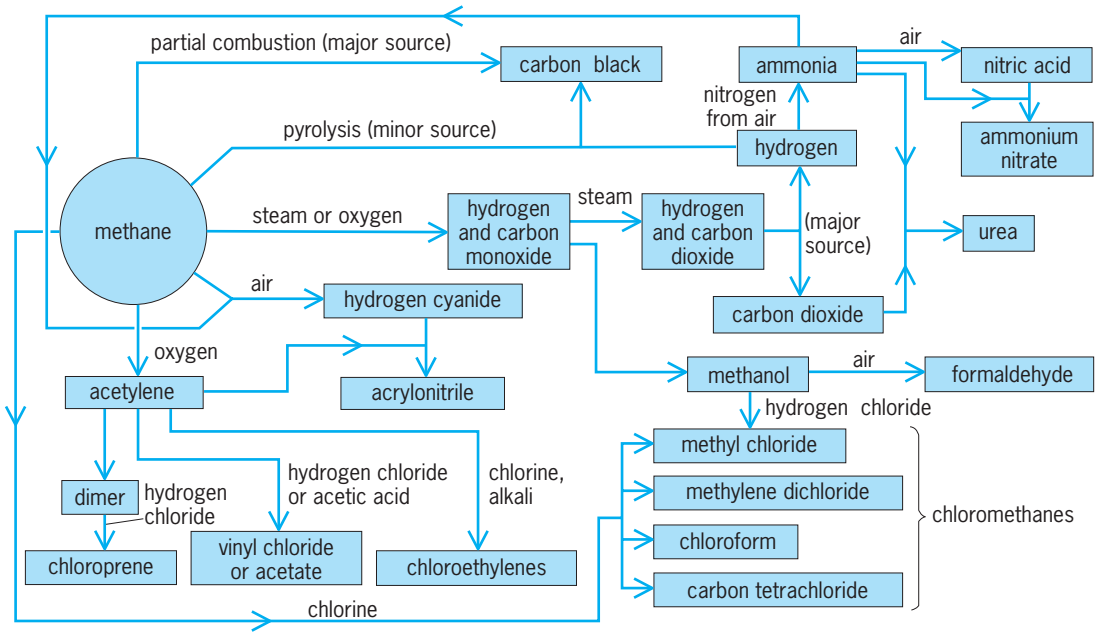

Fig. 1. Chemicals from methane.

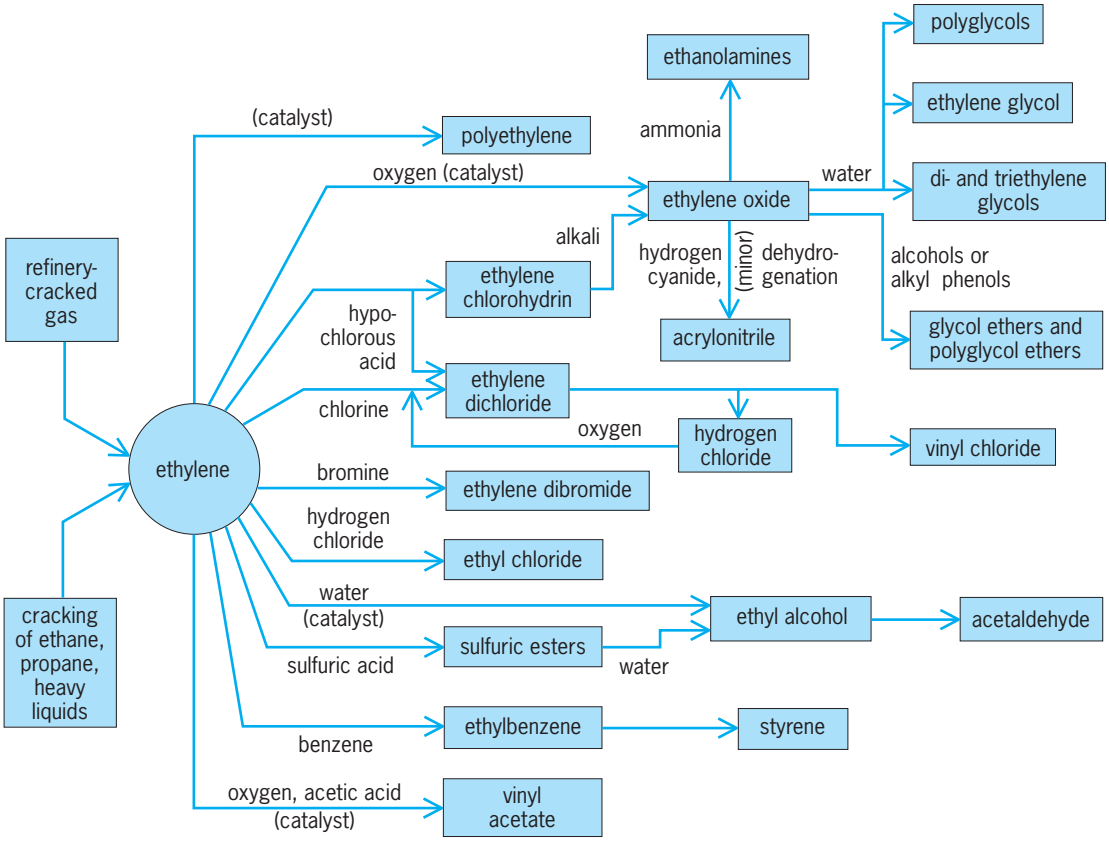

Fig. 2. Chemicals from ethylene.

series are less easy to fractionate from petroleum in pure form; moreover, the number of compounds formed in each particular chemical treatment makes the separation of individual products quite difficult. *See* PARAFFIN.

Halogenation. The ease with which chlorine and bromine can be introduced into the molecules of all the hydrocarbon types present in petroleum has resulted in the commercial production of a number of widely used compounds. With saturated hydrocarbons the reactions are predominantly substitution of hydrogen by chlorine or by bromine, and are exothermic, difficult to control, and inclined to be explosive [reaction (1)].

$$RH + Cl_2 \longrightarrow RCl + HCl \qquad (1)$$

Moderately high temperatures are used, about 250–300°C (480–570°F), for the thermal chlorination of methane; but as the molecular weight of the paraffin increases, the temperature may generally be lowered. A mixture of chlorinated derivatives is always obtained; and many variables, such as choice of catalyst, dilution of inert gases, and presence of other chlorinating agents (antimony pentachloride, sulfuryl chloride, and phosgene), have been tried in an effort to direct the path of the reaction.

An example of the chlorination reaction is the formation of ethyl chloride by the chlorination of ethane [reaction (2)]. Ethyl chloride (CH_3CH_2Cl)

$$CH_3CH_3 + Cl_2 \longrightarrow CH_3CH_2Cl + HCl \qquad (2)$$

is also prepared by the direct addition of hydrogen chloride (HCl) to ethylene ($CH_2{=}CH_2$) or by reacting ethyl ether ($CH_3CH_2OCH_2CH_3$) or ethyl alcohol (CH_3CH_2OH) with hydrogen chloride. The chlorination of *n*-pentane and *iso*-pentane does not take place in the liquid or vapor phase below 100°C (212°F) in the absence of light or a catalyst, but above 200°C (390°F) it proceeds smoothly by thermal action alone. The hydrolysis of the mixed chlorides obtained yields all the isomeric amyl (C_5) alcohols except *iso*-amyl alcohol. Reaction with acetic acid produces the corresponding amyl acetates, which find wide use as solvents.

The alkyl chloride obtained on substituting an equivalent of one hydrogen atom by a chloride atom in kerosine is used to alkylate benzene or naphthalene for the preparation of a sulfonation stock for use in the manufacture of detergents and antirust agents. Similarly, paraffin wax can be converted to a hydrocarbon monochloride mixture, which can be used to alkylate benzene, naphthalene, or anthracene. The product finds use as a pour-point depressor for retarding wax crystal growth and deposition in cold lubricating oils. *See* HALOGENATED HYDROCARBON; HALOGENATION.

Nitration. Hydrocarbons that are usually gaseous (including normal and *iso*-pentane) react smoothly in the vapor phase with nitric acid to give a mixture of nitro compounds, but there are side reactions, mainly oxidation. Only mono nitro derivatives are obtained with the lower paraffins at high temperatures, and they correspond to those expected if

scission of a C-C and C-H bond occurs. For example, ethane yields nitromethane and nitroethane [reaction (3)]. Propane yields nitromethane, nitroethane, 1-nitropropane, and 2-nitropropane [reaction (4)].

$$2CH_3CH_3 + 2HNO_3 \longrightarrow CH_3NO_2 + CH_3CH_2NO_2 + 2H_2O \quad (3)$$

$$4CH_3CH_2CH_3 + 4HNO_3 \longrightarrow CH_3NO_2 + CH_3CH_2NO_2 + CH_3CH_2CH_2NO_2 + CH_3CH(NO_2)CH_3 + 4H_2O \quad (4)$$

The nitro derivatives of the lower paraffins are colorless and noncorrosive and are used as solvents or as starting materials in a variety of syntheses. For example, treatment with inorganic acids and water yields fatty acid (RCO_2H) and hydroxylamine (NH_2OH) salts, and condensation with aldehydes ($RCH{=}O$) yields nitroalcohols [$RCH(NO_2)OH$]. *See* NITRATION; NITRO AND NITROSO COMPOUNDS; NITROPARAFFIN.

Oxidation. The oxidation of hydrocarbons and hydrocarbon mixtures is generally uncontrollable, and the complex mixtures of products have made resolution of the reaction sequences extremely difficult. Therefore, except for the preparation of mixed products having specific properties, such as fatty acids, hydrocarbons higher than pentanes are not used for oxidation because of the difficulty of isolating individual compounds. For example, when propane and butane are oxidized in the vapor phase, without a catalyst, at 270–350°C (520–660°F) and at 50–3000 psi (345–20,685 kPa), a wide variety of products are obtained, including C_1–C_4 acids, C_2–C_7 ketones, ethylene oxide, esters, formals, and acetals. *See* OXIDATION PROCESS; OXIDATION-REDUCTION.

In contrast to the oxidation of paraffins, cyclohexane is somewhat selective in its reaction with air at 150–250°C (300–480°F) in the liquid phase in the presence of a catalyst, such as cobalt acetate. Cyclohexanol is the initial product, but prolonged oxidation produces adipic acid. Oxidation of cyclohexane and methylcyclohexane over vanadium pentoxide at 450–500°C (840–930°F) affords maleic and glutaric acids.

The preparation of carboxylic acids from petroleum, particularly from paraffin wax, for esterification to fats or neutralization to form soaps is comparatively slow with very little reaction taking place at 110°C (230°F), with a wax melting at 55°C (130°F) after 280 hours. At higher temperatures the oxidation proceeds more readily; maximum yields of mixed alcohol and high-molecular-weight acids are formed at 110–140°C (230–285°F) at 60–150 psi (414–1034 kPa); higher temperatures (140–160°C; 285–320°F) result in more acid formation [reaction (5)]. Acids from formic (HCO_2H) to that

$$\text{Paraffin wax} \longrightarrow \underset{\text{Alcohol}}{ROH} + \underset{\text{Acid}}{RCO_2H} \quad (5)$$

with a 10-carbon atom chain [$CH_3(CH_2)_9CO_2H$] have been identified as products of the oxidation of paraffin wax. *See* CARBOXYLIC ACID; ESTER; SOAP.

Alkylation. Alkylation chemistry contributes to the efficient use of C_4 olefins generated in the cracking operations. *Iso*-butane has been added to butenes (and other low-boiling olefins) to give a mixture of highly branched octanes (such as heptanes) by a process called alkylation. Typically, sulfuric acid (85–100%), anhydrous hydrogen fluoride, or a solid sulfonic acid is employed as the catalyst in these processes. The first step in the process is the formation of a carbocation by combination of an olefin with an acid proton [reaction (6)], followed by the addition

$$(CH_3)_2C{=}CH_2 + H^+ \longrightarrow (CH_3)_3C^+ \quad (6)$$

of the carbocation to a second molecule of olefin to form a dimer carbocation. The extensive branching of the saturated hydrocarbon results in high octane. In practice, mixed butenes are employed (*iso*-butylene, 1-butene, and 2-butene), and the product is a mixture of isomeric octanes that has an octane number of 92–94. With the phase-out of leaded additives in gasoline, octane improvement is a major challenge for the refining industry, and alkylation provides a route to high-octane products. *See* ALKYLATION (PETROLEUM); GASOLINE; OCTANE NUMBER.

Thermolysis. Thermal cracking is the major process for generating ethylene and the other olefins that are the reactive building blocks of the petrochemical industry. In addition to thermal cracking, other important processes that generate sources of hydrocarbon raw materials for the petrochemical industry include catalytic reforming, alkylation, dealkylation, isomerization, and polymerization. *See* CATALYSIS; DEHYDROGENATION; HETEROGENEOUS CATALYSIS; ISOMERIZATION; REFORMING PROCESSES.

At the high temperatures of refinery processes (usually approximately 500°C, or 950°F, and higher), cracking is favored. Unfortunately, in the cracking process certain products interact with one another to produce products of increased molecular weight relative to the molecular weight of the original feedstock. Thus some products are taken from the cracking units as useful products (such as olefin gases, gasoline, and other distillates), but other products include cracked residuum and coke.

Olefins. Olefins (C_nH_{2n}) are the basic building blocks for a host of chemical syntheses (**Tables 4**, **5**, and **6**). These unsaturated materials enter into polymers and rubbers, and with other reagents react to form a wide variety of useful compounds, including alcohols, epoxides, amines, and halides. In addition, ethane (also available from natural gas and cracking processes) is an important source of ethylene that, in turn, provides more valuable routes to petrochemical products (Fig. 2). *See* ALKENE.

Ethylene is usually made by cracking gases—ethane, propane, butane, or a mixture of these as might exist in a refinery's off-gases. Ethylene and coproducts are produced from steam cracking of hydrocarbons such as ethane, propane, butane, naphtha, and gas oil in tubular reactor coils installed in externally fired heaters. Alternative processes include naphtha cracking, whole crude oil cracking, and heavy oil cracking processes; in such cases

TABLE 5. Chemicals from propane and propylene and their uses

Basic derivatives and sources	Uses
Isopropyl alcohol	Acetone
	Solvents, drugs, and chemicals
Cumene	Phenol and acetone
Acrylonitrile	Acrylic fibers
	Nitrile elastomers and acrylonitrile-butadiene-styrene resins
Polypropylene	Molding, fiber, and film
Propylene oxide	Propylene glycol
	Dipropylene glycol and poly(propylene glycol)
	Polyurethane
Oxochemicals	
Isooctyl alcohol	Phthalate esters
Butyraldehydes (propylene only)	Intermediate for butanols, 2-ethylhexanol, *n*-butyric acid
Dodecene (tetramer)	Dodecyl benzene
	Dodecyl phenol
Nonene (trimer)	Decyl alcohol and nonylphenol
Epichlorohydrin	Glycerol and epoxy resins
Polyisoprene	Elastomers

TABLE 6. Chemicals from butane and butylene and their uses

Basic derivatives and sources	Uses
Butadiene	Styrene-butadiene rubber and resins
	Polybutadiene
	Adiponitrile
	Nitrile rubber
	Acrylonitrile-butadiene-styrene plastics
sec-Butyl alcohol	Methyl ethyl ketone
Butyl rubber (from isobutylenes)	Tire products
Polybutenes (from isobutylene)	Lube oil additives
	Caulking and sealing compounds
	Adhesives
	Rubber compounding

significant quantities of higher-molecular-weight olefins and aromatics are also produced. Other processes include ethanol dehydrogenation, syngas-based processes, dehydrogenation of paraffins, oxidative coupling of methane, and olefin disproportionation.

Ethylene manufacture via the steam cracking process is widely practiced throughout the world. The operating facilities are similar to gas oil cracking units, operating at temperatures of 840°C (1550°F) and at low pressures (24 psi; 165 kPa). Steam is added to the vaporized feed to achieve a 50-50 mixture, and furnace residence times are only 0.2–0.5 second. Ethane extracted from natural gas is the predominant feedstock for ethylene cracking units. Propylene and butylene are largely derived from catalytic cracking units and from cracking a naphtha or light gas oil fraction to produce a full range of olefin products.

The olefinic materials ($RCH{=}CHR'$) present in gaseous products of cracking processes offer promising source materials. Cracking paraffin hydrocarbons and heavy oils also produces olefins. For example, cracking ethane (CH_3CH_3), propane ($CH_3CH_2CH_3$), butane ($CH_3CH_2CH_2CH_3$), and other feedstocks such as naphtha and gas oil produces ethylene. Propylene ($CH_3CH{=}CH_2$) is produced from thermal and catalytic cracking of naphtha and gas oils, as well as propane and butane.

Hydroxylation. The earliest method for converting olefins into alcohols involved their absorption in sulfuric acid to form esters, followed by dilution and hydrolysis, generally with the aid of steam. In the case of ethyl alcohol, the direct catalytic hydration of ethylene can be used. Ethylene is readily absorbed in 98–100% sulfuric acid at 75–80°C (165–175°F), and both the mono- and diethyl sulfates are formed; hydrolysis takes place readily on dilution with water and heating.

The direct hydration of ethylene to ethyl alcohol is practiced over phosphoric acid on diatomaceous earth or promoted tungstic oxide under 100 psi (690 kPa) pressure and at 300°C (570°F). Purer ethylene is required in direct hydration than in the acid absorption process and the conversion per pass is low, but high yields are possible by recycling. Propylene and the normal butenes can also be hydrated directly.

One of the first alcohol syntheses practiced commercially was that of isopropyl alcohol from propylene. Sulfuric acid absorbs propylene more readily than it does ethylene, but care must be taken to avoid polymer formation by keeping the mixture relatively cool and using acid of about 85% strength at 300–400 psi (2068–2758 kPa) pressure; dilution with an inert oil may also be necessary. Acetone is readily made from isopropyl alcohol, either by catalytic oxidation or by dehydrogenation over metal (usually copper) catalysts. Secondary butyl alcohol is formed on absorption of 1- or 2-butene by 78–80% sulfuric acid, followed by dilution and hydrolysis. Secondary butyl alcohol is converted into methyl ethyl ketone by catalytic oxidation or dehydrogenation.

There are several methods for preparing higher alcohols. One method, the Oxo reaction, involves the direct addition of carbon monoxide (CO) and a hydrogen (H) atom across the double bond of an olefin to form an aldehyde ($RCH{=}O$) that, in turn, is reduced to the alcohol (RCH_2OH). Hydroformylation (the Oxo reaction) is brought about by contacting the olefin with synthesis gas (1:1 carbon monoxide–hydrogen) at 75–200°C (165–390°F) and 1500–4500 psi (10,342–31,027 kPa) over a metal catalyst, usually cobalt. The active catalyst is held to be cobalt hydrocarbonyl $HCO(CO)_4$, formed by the action of the hydrogen on dicobalt ocatcarbonyl.

A wide variety of olefins participate in this reaction, and the olefins that contain terminal unsaturation ($—CH{=}CH_2$) are the most active. The hydroformylation reaction is not specific, and the hydrogen and carbon monoxide add across each side of the double bond. Propylene gives a mixture of 60% *n*-butyraldehyde and 40% *iso*-butyraldehyde. Terminal ($RCH{=}CH_2$) and nonterminal ($RCH{=}CHR'$) olefins, such as 1- and 2-pentene ($CH_3CH_2CH_2CH{=}CH_2$ and $CH_3CH_2CH{=}CHCH_3$), give the same distribution of straight- and branched-chain C_6 aldehydes because

of isomerization. Simple branched structures add mainly at the terminal carbon; *iso*-butylene forms 95% *iso*-valeraldehyde and 5% trimethylacetaldehyde. Commercial application of the synthesis has been most successful in the manufacture of *iso*-octyl alcohol from a refinery C_3C_4 dimer, decyl alcohol from propylene trimer, and tridecyl alcohol from propylene tetramer. *See* ALCOHOL.

The hydrolysis of ethylene chlorohydrin ($HOCH_2CH_2Cl$) or the cyclic ethylene oxide produces ethylene glycol ($HOCH_2CH_2OH$). The main use for this chemical is for antifreeze mixtures in automobile radiators and for cooling aviation engines; considerable amounts are used as ethylene glycol dinitrate in low-freezing dynamites. Propylene glycol is also made by the hydrolysis of its chlorohydrin or oxide.

Glycerin [$HOCH_2CH(OH)CH_2OH$] can be derived from propylene by high-temperature chlorination to produce alkyl chloride, followed by hydrolysis to allyl alcohol and then conversion with aqueous chloride to glycerol chlorohydrin, a product that can be easily hydrolyzed to glycerol (glycerin).

Halogenation. At ordinary temperatures, chlorine and bromine react with olefins by addition. For example, ethylene dichloride ($ClCH_2CH_2Cl$) is made in this way from ethylene. At slightly higher temperatures, substitution occurs. In the chlorination of propylene, a rise of 50°C (90°F) changes the product from propylene dichloride [$CH_3CH(Cl)CH_2Cl$] to allyl chloride ($CH_2{=}CHCH_2Cl$).

Polymerization. The polymerization of ethylene under pressure (1500–3000 psi; 10,342–20,684 kPa) at 110–120°C (230–250°F) in the presence of a catalyst or initiator, such as a 1% solution of benzoyl peroxide in methanol, produces a polymer in the 2000–3000 molecular-weight range. Polymerization at 15,000–30,000 psi (103,422–206,843 kPa) and 180–200°C (355–390°F) produces a wax melting at 100°C (212°F) and 15,000–20,000 molecular weight, but the reaction is not straightforward since there are branches in the chain. Considerably lower pressures can be used over catalysts composed of aluminum alkyls (R_3Al) in the presence of titanium tetrachloride ($TiCl_4$), supported chromic oxide (CrO_3), nickel (NiO), or cobalt (CoO) on charcoal, and promoted molybdena-alumina (MoO_2-Al_2O_3), which at the same time give products more linear in structure. Polypropylenes can be made in similar ways; and mixed monomers, such as ethylene-propylene and ethylene-butene mixtures, can be treated to give high-molecular-weight copolymers of good elasticity. Lower-molecular-weight polymers, such as the dimers, trimers, and tetramers, are used as constituents of gasoline. The materials are normally prepared over an acid catalyst. Propylene trimer (dimethylheptenes) and tetramer (trimethylnonenes) are applied in the alkylation of aromatic hydrocarbons for the production of alkylaryl sulfonate detergents, and as olefinic feedstocks in the manufacture of C_{10} and C_{13} oxo-alcohols. Phenol is alkylated by the trimer to make nonylphenol, a chemical intermediate for the manufacture of lubricating oil detergents and other products. *See* PHENOL.

Isobutylene also forms several series of valuable products; the di- and triisobutylenes make excellent motor and aviation fuel additives; they can also be used as alkylating agents for aromatic hydrocarbons and phenols and as reactants in the oxo-alcohol synthesis. Polyisobutylenes in the viscosity range of 10,000 m^2/s or 55,000 SUS (Saybolt Universal Seconds) [at 38°C; 100°F] have been used as viscosity index improvers in lubricating oils. 1-Butene ($CH_3CH_2CH{=}CH_2$) and 2-butene ($CH_3CH{=}CHCH_3$) participate in polymerization reactions by the way of butadiene ($CH_2{=}CHCH{=}CH_2$), the dehydrogenation product, which is copolymerized with styrene (23.5%) to form GR-S (styrene-butadiene) rubber, and with acrylonitrile (25%) to form GR-N (nitrite) rubber. Derivatives of acrylic acid (butyl acrylate, ethyl acrylate, 2-ethylhexyl acrylate, and methyl acrylate) can be homopolymerized using peroxide initiators, or copolymerized with other monomers to generate acrylic resins.

Oxidation. The most striking industrial olefin oxidation process involves the oxidation of ethylene by air using a silver catalyst at 225–325°C (435–615°F) to give pure ethylene oxide in yields ranging from 55 to 70%. Analogous higher-olefin oxides can be prepared from propylene, butadiene, octene, dodecene, and styrene via the chlorohydrin route or by reaction with peracetic acid. Acrolein ($CH_2{=}CHCHO$) is formed by air oxidation or propylene over a supported cuprous oxide catalyst or by condensing acetaldehyde and formaldehyde. When acrolein and air are passed over a catalyst, such as cobalt molybdate, acrylic acid is produced; or if acrolein is reacted with ammonia and oxygen over molybdic oxide, the product is acrylonitrile. Similarly, propylene may be converted to acrylonitrile.

Oxidation of the higher olefins by air is difficult to control, but at temperatures between 350 and 500°C (660 and 930°F) maleic acid is obtained from amylene (2-methyl-2-butene) and a vanadium pentoxide catalyst; higher yields of the acid are obtained from hexene, heptene, and octene. *See* OXIDATION PROCESS; OXIDATION-REDUCTION.

Aromatics. Aromatic compounds are valuable starting materials for a variety of chemical products. Reforming processes have made benzene, toluene, xylene, and ethylbenzene economically available from petroleum sources. A further source of supply is the aromatic-rich liquid fraction produced in the cracking of naphtha or light gas oils during the manufacture of ethylene and other olefins. The aromatics are generally recovered by extractive or azeotropic distillation, by solvent extraction (with water-glycol mixtures or liquid sulfur dioxide), or by adsorption. Naphthalene and methylnaphthalenes are present in catalytically cracked distillates. *See* AROMATIC HYDROCARBON; BENZENE; POLYNUCLEAR HYDROCARBON.

The major aromatic feedstocks are benzene (C_6H_6), toluene ($C_6H_5CH_3$), xylene ($CH_3C_6H_4CH_3$), and the two-ring condensed aromatic compound

naphthalene ($C_{10}H_8$) [**Table 7**; **Fig. 3**]. Benzene is used to make styrene ($C_6H_5CH{=}CH_2$), the basic ingredient of polystyrene plastics. It is also used to make paints, epoxy resins, glues, and other adhesives. Toluene is used primarily to make solvents, gasoline additives, and explosives. Xylene is used in the manufacture of plastics and synthetic fibers and in the refining of gasoline. Naphthalene is used to manufacture insecticides. Toluene is used as a source of trinitrotoluene (TNT); and alkylation with propylene, followed by dehydrogenation, yields methylstyrene, which can be used for polymerization. Alkylation of toluene with propylene tetramer yields a product suitable for sulfonation to a detergent-grade surfactant. *See* NITROAROMATIC COMPOUND; STYRENE.

Ortho-xylene is oxidized by nitric acid to phthalic anhydride, *m*-xylene to *iso*-phthalic acid, and *p*-xylene with nitric acid to terephthalic acid. These acid products are used in the manufacture of fibers, plastics, and plasticizers. Phthalic anhydride is also produced by the aerial oxidation of naphthalene at 400–450°C (750–840°F) in the vapor phase at about 25 psi (172 kPa) pressure over a fixed-bed vanadium

TABLE 7. Chemicals from cycloaliphatics and aromatics and their uses

Chemical	Uses
Benzene	Styrene
	Cyclohexane
	Phenol
	Detergent alkylate
	Maleic anhydride
	Aniline
	DDT
Toluene	Dealkylation to benzene
	Solvents
	Toluene diisocyanate
	Motor and aviation gasoline
	TNT
Xylenes	*p*-Xylene
	o-Xylene
	m-Xylene
	Solvents and chemicals
	Gasolines
Ethyl benzene	Styrene
Cyclohexane	Nylon intermediates
	Nonnylon uses (cyclohexanone and adipic acid)
Naphthalene	Phthalic anhydride
	Insecticides
	β-Naphthol and mothballs

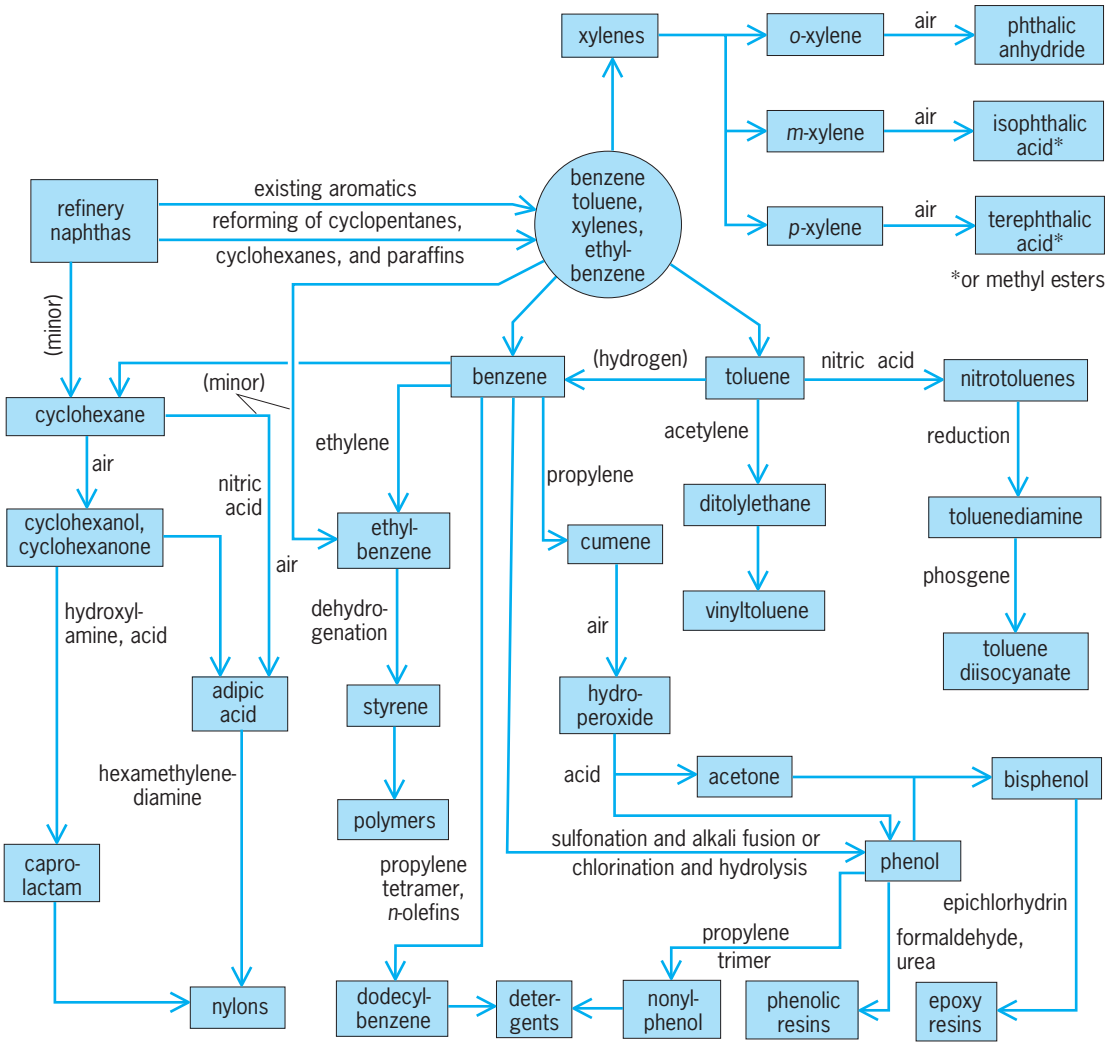

Fig. 3. Chemicals from cycloaliphatic compounds and from aromatic compounds.

pentoxide (V_2O_5) catalyst. Terephthalic acid is produced in a similar manner from *p*-xylene.

Chemicals from natural gas. Natural gas can be used as a source of hydrocarbons (such as ethane and propane) that are higher-molecular-weight than methane and that are important chemical intermediates. The preparation of chemicals and chemical intermediates from natural gas should not be restricted to those described here, but should be regarded as some of the building blocks of the petrochemical industry. In summary (Table 3; Fig. 1), natural gas can be a very important source of petrochemical intermediates and solvents. *See* NATURAL GAS.

The availability of hydrogen from catalytic reforming operations has made its application economically feasible in a number of petroleum-refining operations. Previously, the chief sources of large-scale hydrogen (used mainly for ammonia manufacture) were the cracking of methane (or natural gas) and the reaction between methane and steam. In the latter, at 900–1000°C (1650–1830°F) conversion into carbon monoxide and hydrogen results [reaction (7)].

$$CH_4 + H_2O \longrightarrow CO + 3H_2 \qquad (7)$$

If this mixture is treated further with steam at 500°C (930°F) over a catalyst, the carbon monoxide present is converted into carbon dioxide and more hydrogen is produced [reaction (8)]. The reduction of

$$CO + H_2O \longrightarrow H_2 + CO_2 \qquad (8)$$

carbon monoxide by hydrogen is the basis of several syntheses, including the manufacture of methanol and higher-boiling alcohols. In fact, the synthesis of hydrocarbons by the Fischer-Tropsch reaction (9)

$$nCO + 2nH_2 \longrightarrow (CH_2)_n + nH_2O \qquad (9)$$

occurs in the temperature range 200–350°C (390–660°F), which is sufficiently high for the water-gas shift to take place in presence of the catalyst [reaction (10)]. The major products are olefins and

$$CO + H_2O \longrightarrow CO_2 + H_2 \qquad (10)$$

paraffins, together with some oxygen-containing organic compounds that may be varied by changing the catalyst or the temperature, pressure, and carbon monoxide–hydrogen ratio. *See* FISCHER-TROPSCH PROCESS.

The hydrocarbons formed are mainly aliphatic, and on a molar basis methane is the most abundant; the amount of higher hydrocarbons usually decreases gradually with increased molecular weight. *Iso*-paraffin formation is more extensive over zinc oxide (ZnO) or thoria (ThO_2) at 400–500°C (750–930°F) and at higher pressure. Paraffin waxes are formed over ruthenium catalysts at relatively low temperatures (170–200°C; 340–390°F), high pressures (1500 psi; 10,342 kPa), and with a carbon monoxide–hydrogen ratio. The more highly branched product made over the iron catalyst is an important factor in a choice for the manufacture of automotive fuels. On the other hand, a high-quality diesel fuel (paraffinic character) can be prepared over cobalt.

The small amount of aromatic hydrocarbons found in the product covers a wide range of possibilities. In the C_6–C_9 range, benzene, toluene, ethylbenzene, xylene, *n*-propyl- and *iso*-propylbenzene, methylethylbenzenes, and trimethylbenzenes have been identified; naphthalene derivatives and anthracene derivatives are also present.

Synthesis gas. Synthesis gas is a mixture of carbon monoxide (CO) and hydrogen (H_2) that is the beginning of a wide range of chemicals (**Fig. 4**). Many refining catalysts promote the steam-reforming reaction (11) that leads to a product gas containing more

$$CH_4 + H_2O \longrightarrow CO + 3H_2 \qquad (11)$$

hydrogen and carbon monoxide and fewer unsaturated hydrocarbon products than the gas product from a noncatalytic process. The catalyst also influences the reactions rates in the thermal cracking reactions, which can lead to higher gas yields and lower tar and carbon yields.

Synthesis gas can be produced from heavy oil by a partial oxidizing via the oxidation process that leads to the formation of carbon monoxide and hydrogen [reaction (12)]. The initial partial oxidation step con-

$$[2CH]_{petroleum} + O_2 \longrightarrow 2CO + H_2 \qquad (12)$$

sists of the reaction of the feedstock with a quantity of oxygen insufficient to burn it completely, making a mixture consisting of carbon monoxide, carbon dioxide, hydrogen, and steam. *See* OXYGEN.

Inorganic petrochemicals. An inorganic petrochemical is one that does not contain carbon atoms; typical examples are sulfur (S), ammonium sulfate [$(NH_4)_2SO_4$)], ammonium nitrate (NH_4NO_3), and nitric acid (HNO_3). Of the inorganic petrochemicals, ammonia is by far the most common. Ammonia is produced by the direct reaction of hydrogen with nitrogen, with air being the source of nitrogen [reaction (13)].

$$N_2 + 3H_2 \longrightarrow 2NH_3 \qquad (13)$$

Ammonia production requires hydrogen from a hydrocarbon source. Traditionally, the hydrogen was produced from a coke and steam reaction, but refinery gases, steam reforming of natural gas (methane) and naphtha streams, and partial oxidation of hydrocarbons or higher-molecular-weight refinery residual materials (residua, asphalt) are the sources of hydrogen. The ammonia is used predominantly for the production of ammonium nitrate (NH_4NO_3) as well as other ammonium salts and urea (H_2HCONH_2) that are major constituents of fertilizers. *See* AMMONIA; AMMONIUM SALT; FERTILIZER; UREA.

Carbon black (also classed as an inorganic petrochemical) is made predominantly by the partial combustion of carbonaceous (organic) material in a limited supply of air. The carbonaceous sources vary from methane, to aromatic petroleum oils, to coal tar by-products. The carbon black is used primarily

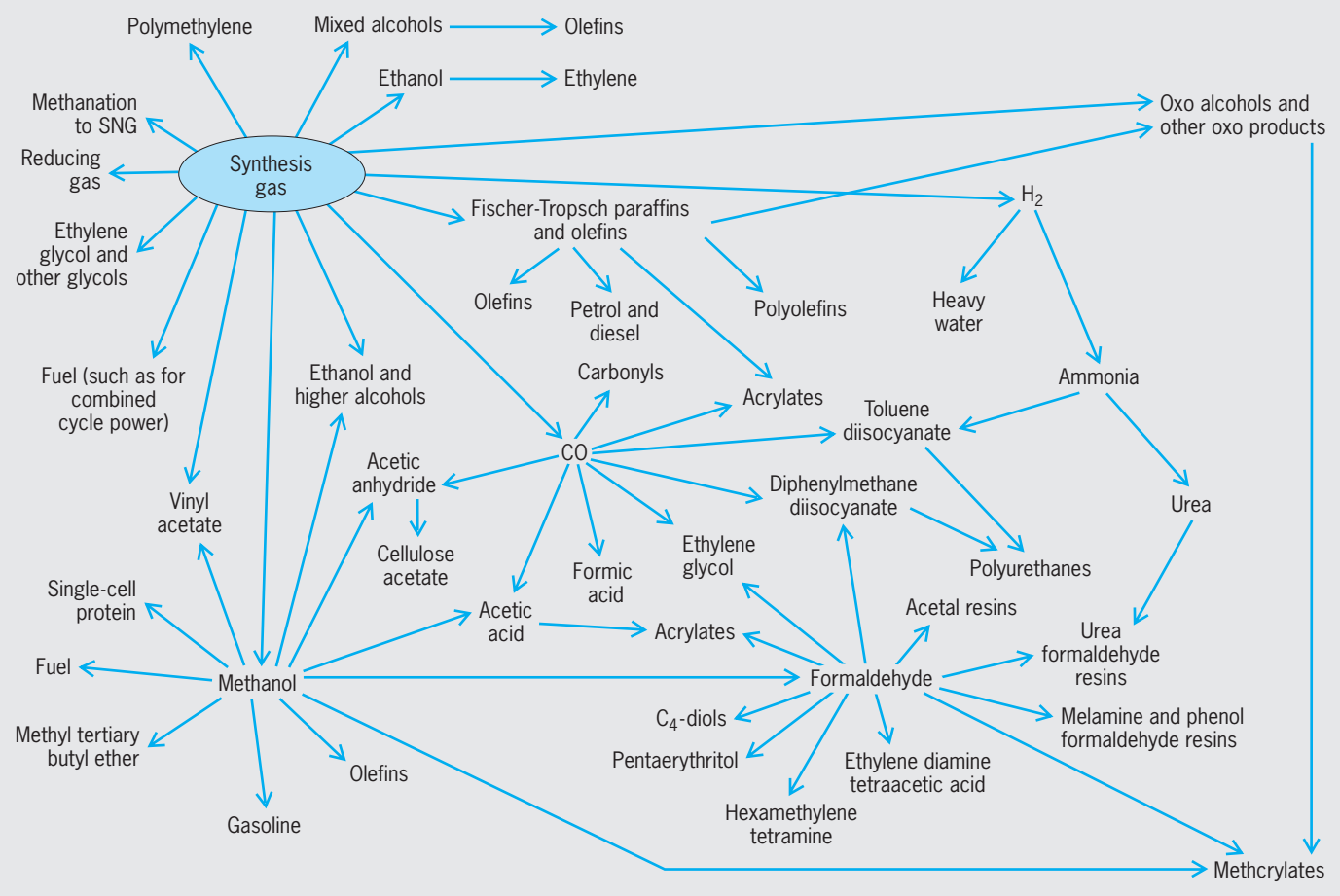

Fig. 4. Chemicals from synthesis gas. (*After J. G. Speight, The Chemistry and Technology of Petroleum, 3d ed., p. 852, Marcel Dekker, New York, 1999*)

in the production of synthetic rubber. *See* CARBON BLACK.

Sulfur, another inorganic petrochemical, is obtained by the oxidation of hydrogen sulfide [reaction (14)]. Hydrogen sulfide is a constituent of

$$H_2S + O_2 \longrightarrow H_2O + S \qquad (14)$$

natural gas and also of the majority of refinery gas streams, especially those off-gases from hydrodesulfurization processes. Most of the sulfur is converted to sulfuric acid for the manufacturer of fertilizers and other chemicals. Other uses for sulfur include the production of carbon disulfide, refined sulfur, and pulp and paper industry chemicals. *See* PAPER; SULFUR; SULFURIC ACID.

James G. Speight

Bibliography. R. A. Meyers (ed.), *Handbook of Petroleum Refining Process*, McGraw-Hill, New York, 1997; J. G. Speight, *The Chemistry and Technology of Coal*, 2d ed., Marcel Dekker, New York, 1994; J. G. Speight, *The Chemistry and Technology of Petroleum*, 3d ed., Marcel Dekker, New York, 1999; J. G. Speight, *The Desulfurization of Heavy Oils and Residua*, 2d ed., Marcel Dekker, New York, 2000; J. G. Speight (ed.), *Fuel Science and Technology Handbook*, Marcel Dekker, New York, 1990; J. G. Speight (ed.), *Petroleum Chemistry and Refining*, Taylor & Francis, New York, 1998.

Petrofabric analysis

The systematic study of the fabrics of rocks, generally involving statistical study of the orientations and distributions of large numbers of fabric elements. The term fabric denotes collectively all the structural or spatial characteristics of a rock mass. The fabric elements are classified into two groups: (1) megascopic features, including bedding, schistosity, foliation, cleavage, faults, joints, folds, and mineral lineations; and (2) microscopic features, including the shapes, orientations, and mutual arrangement of the constituent mineral crystals (texture) and of internal structures (twin lamellae, deformation bands, and so on) inside the crystals. The aim of fabric analysis is to obtain as complete and accurate a description as possible of the structural makeup of the rock mass with a view to elucidating its kinematic history. The fabric of a sedimentary rock, for example, may retain evidence of the mode of transport, deposition, and compaction of the sediment in the size, shape, and disposition of the particles; similarly, that of an igneous rock may reflect the nature of the flow or of gravitational segregation of crystals and melt during crystallization. The fabrics of deformed metamorphic rocks (tectonites) have been most extensively studied by petrofabric techniques with the objective

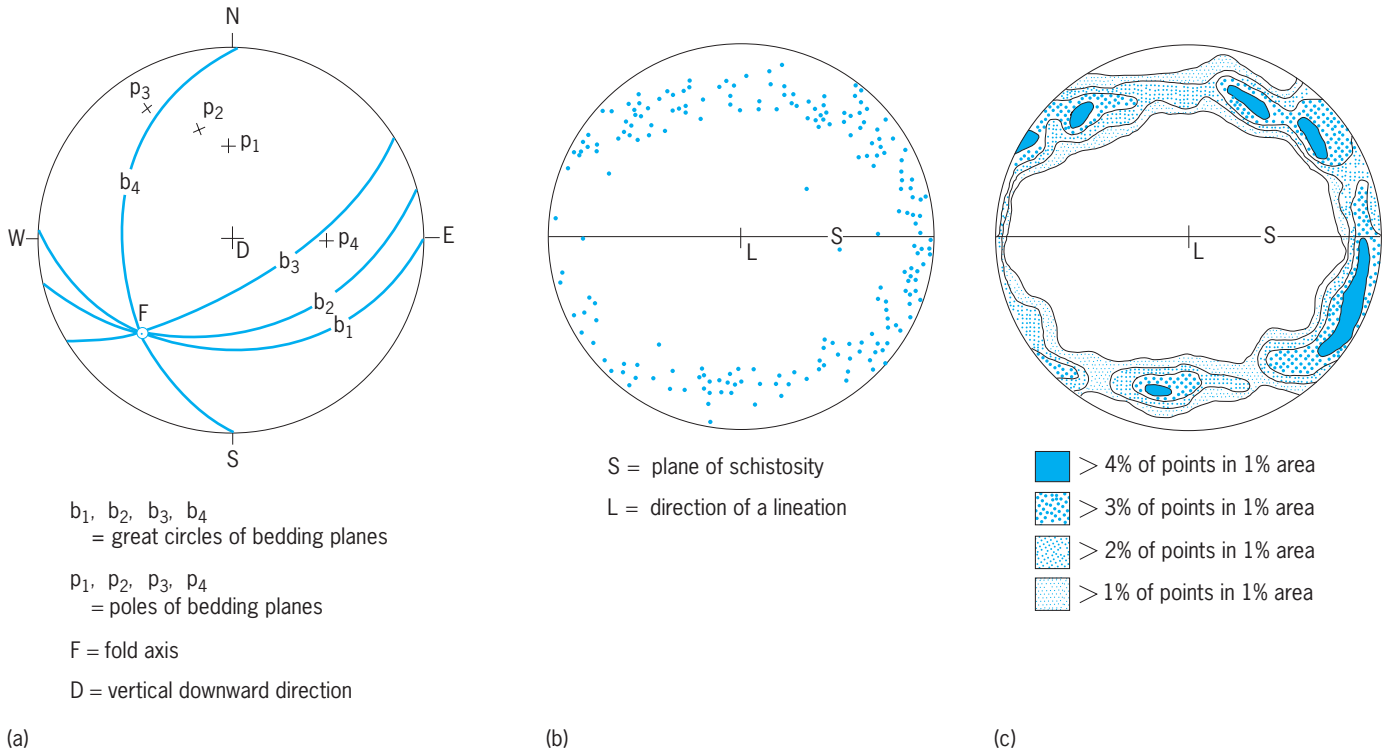

Fig. 1. Representation of the orientations of structures in spherical projections. (*a*) Stereographic projection of the orientations of four bedding planes in a cylindrical fold. The poles of the planes are distributed on a great circle perpendicular to the fold axis. (*b*) Equal-area projection of 200 *c* axes of quartz crystals in a metamorphic quartzite. The crystal axes show a strong preferred orientation. (*c*) Same data as *b*, contoured to show the distribution of the crystal axes.

of determining the details of the history of deformation and recrystallization. *See* FAULT AND FAULT STRUCTURES; FOLD AND FOLD SYSTEMS; STRUCTURAL GEOLOGY; STRUCTURAL PETROLOGY.

Techniques. The methods of investigation of fabrics include mapping the distribution and recording the orientations of megascopic structures in the field, and collecting selected specimens of recorded orientation for laboratory study. Measurements of the orientations of microscopic structures are made on cut, polished, or etched surfaces of oriented specimens with a binocular microscope or in thin sections of known orientation with a polarizing microscope and universal stage. The orientations of planar and linear structures are analyzed by means of spherical (stereographic and equal-area) projections; each plane or line is considered to pass through the center of the sphere, so that the orientations of planes or lines are represented by a great circle or a point, respectively, in the projections (**Fig. 1***a*). If the equal-area projection is employed, the distribution of large numbers of points on the projection (representing directions such as poles of bedding, mineral lineations, or crystallographic axes of mineral grains) may be contoured to give a statistical representation of the preferred orientation of the fabric elements (Fig. 1*b* and *c*). The spatial distribution of groups of elements may be analyzed on maps or thin sections by statistical techniques. The fields of homogeneity of subfabrics of different elements must be defined. *See* PETROGRAPHY.

Applications. Significant information may be derived directly from a synthesis of the fabric data. The geometry of a rock mass is described in quantitative terms. In areas that have suffered repeated deformation it is generally possible by association to identify groups of structures assignable to separate phases of deformation. It is also possible in some instances to establish the sequence of the deformations from simple geometric considerations.

Although an exhaustive study of tectonic fabrics on all scales is seldom made, analysis of some elements of the microfabric is commonly done to supplement the information obtained from megascopic structures. Tectonites usually display textures that clearly originated by recrystallization, and many common minerals (such as quartz, micas, calcite, and dolomite) show preferred orientations of their crystal directions (Fig. 1*b* and *c*) analogous to those in deformed metals. Microscopic textures are especially informative in determining the sequence of recrystallization, and deformation in a rock, as recrystallization obliterates preexisting textures and internal structures in the mineral grains. Thus microstructures induced after metamorphism are preserved, whereas those produced before or during metamorphism are removed in the process of recrystallization and growth of new mineral phases. Fortunately, a record of the deformation during metamorphism is also retained in the textures and preferred orientations of crystals, though the exact interpretation of these records is still not fully understood. One

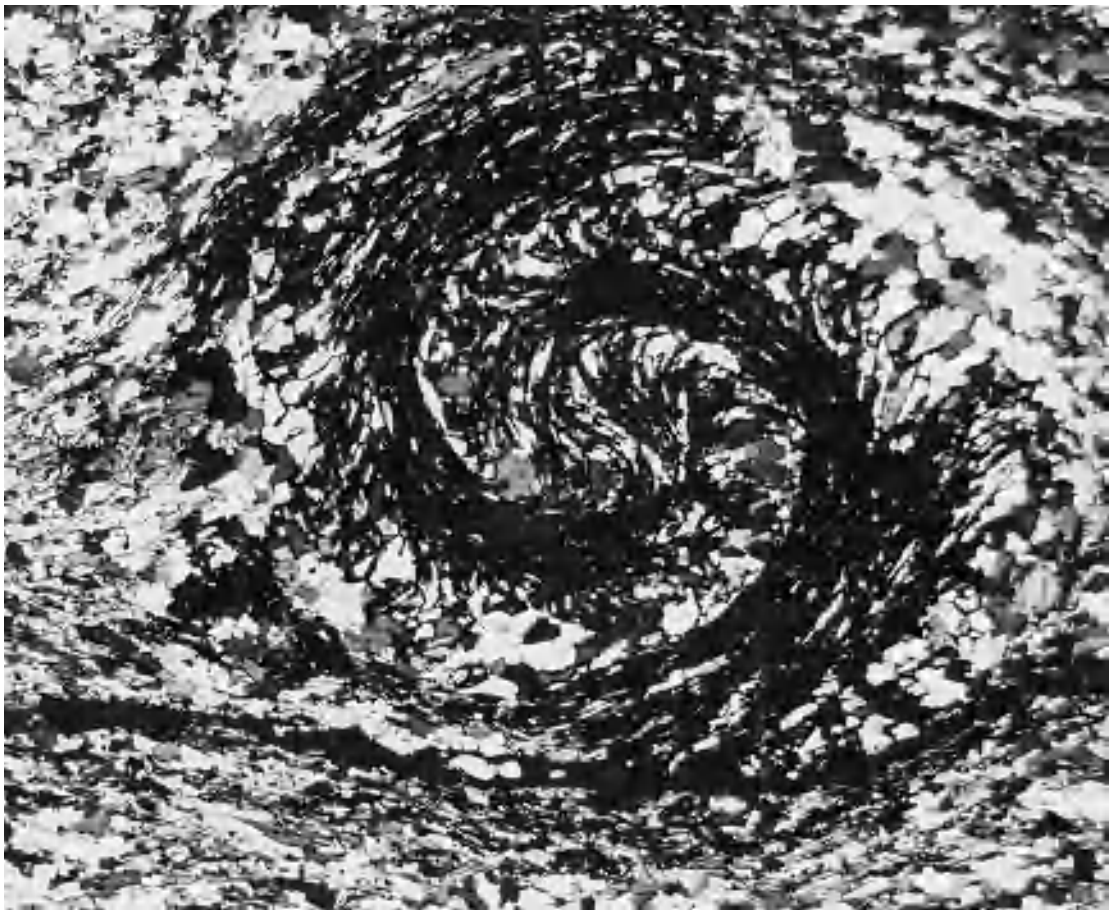

Fig. 2. Photomicrograph of a thin section of schist from Proctorsville, Vermont, showing a "snowball" garnet (crossed polarizers). The garnet (circular black area) contains many inclusions of quartz, feldspar, and mica arranged in a helical distribution as a result of rotation of the garnet during its growth. Garnet is 6 mm (0.2 in.) in diameter. (*Courtesy of J. L. Rosenfeld*)

spectacular example of such a record is seen in the trains of mineral inclusions in "rolled," or "snowball," garnets, which were rotated because of strain in the enclosing rock during their growth (**Fig. 2**).

Traditionally the strain history has been inferred from fabrics, largely by the use of symmetry arguments of the type introduced by B. Sander in his classic studies of rock fabrics: The overall symmetry of a fabric is considered to be the same as that of the deformation that produced the fabric. The general validity of such arguments is borne out by experiments, if the influence of the predeformational fabric and the sequence of incremental contributions to the strain (that is, the history of strain) are taken into account.

This approach does not yield a unique picture of the strain history, however, and more precise interpretation of fabrics is based on results of experimental deformation of rocks and mineral crystals at high temperatures and pressures. Certain types of folding, microscopic textures, and preferred orientations of minerals have been reproduced in the laboratory under known conditions closely analogous to those existing at depth in the Earth's crust. The results of these experiments provide a reliable basis for specifying the mode of origin of many features of rock fabrics. Theoretical analysis and experiments with dimensionally matched models are also contributing appreciably to the understanding of rock structures and the development of rock fabrics. *See* HIGH-PRESSURE MINERAL SYNTHESIS. John M. Christie

Bibliography. K. Bucher and M. Frey, *Petrogenesis of Metamorphic Rocks*, 6th ed., 1994; R. Mason, *Petrology of the Metamorphic Rocks*, 2d ed., 1990; A. Miyashiro, *Metamorphic Petrology*, 1994; B. B. Rao, *Metamorphic Petrology*, 1986; F. J. Turner, *Metamorphic Petrology*, 2d ed., 1981.

Petrography

The description of rocks with goals of classification and interpretation of origin. Most schemes for the classification of rocks are based on the size of grains and the proportions of various minerals. Interpretations of origin rely on field relations, structure, texture, and chemical composition as well as sizes and proportions of different kinds of grains. The names of rocks are based on the sizes and relative proportions of different minerals; boundaries between the names are arbitrary. The conditions of formation of a rock can be estimated from the types and textures

of its constituent minerals. The **illustration** shows examples of petrographic features.

Rock description. The description of rocks begins in the field with observation of the shape and structure of bodies of rock at the scale of centimeters to kilometers. The geometrical relations between and structures within mappable rock units are generally the domain of field geology, but are simply rock descriptions at a reduced scale.

A petrographer can correctly name most rocks in which most crystals are larger than about 0.04 in. (1 mm) simply by examining the rock with a 10-power magnifying lens. Rocks with smaller grains require either microscopical examination or chemical analysis for proper classification.

Sizes, shapes, and orientations of grains and voids are the most important features of a rock relevant to its origin. The same features also affect density, porosity, permeability, strength, and magnetic behavior. It is also essential to know the identity, abundance, and compositions of minerals constituting the grains in order to name a rock and infer its conditions of formation.

Rock sections. Sections of rock are used for measurement of grain size and other purposes. There are two kinds of rock sections: flat surfaces (commonly polished) which are looked at, and thin slices which are looked through. To make rock sections, the rock is first impregnated with a plastic to hold it together. The rock is then sawed into a wafer a few millimeters thick using a cutting edge embedded with diamonds. The wafer is smoothed with abrasives, glued to a glass slide, and ground to a thinness of about 0.001 in. (0.03 mm). The wafer may be polished but not thinned, or it may be both thinned and polished.

The polarizing microscope is used by petrographers to routinely examine sections of rock. Polarized light is useful in identifying minerals because most minerals are optically anisotropic and this property affects the transmission and reflection of light in characteristic ways. Rock slices are mounted in standard glue, and the glue helps reveal the relative refractive index of ground mineral grains when they are viewed under the microscope. Opaque materials are polished and then examined under the petrographic microscope using reflected light (sometimes called a metallographic microscope). Identification of minerals under reflected light depends on variable reflectivity, color, hardness, and reflection anisotropy (see illustration).

Separation of minerals. Separation of grains of different minerals is necessary for some chemical and physical studies. Magnetic and electrostatic devices and liquids with densities up to about 4.2 g/cm^3 (2.4 oz/in.^{3}) are commonly used. Similar methods are used industrially to concentrate ore minerals.

Grains. In sedimentary and some igneous rocks, grains commonly consist of more than one crystal. To determine grain size in sedimentary rocks, the petrographer must be able to discern individual particles of deposition and determine whether they are obvious pebbles in a conglomerate or fecal pellets in a mud. In igneous rocks, aggregates or intergrowths which formed individual entities during crystallization must be recognized. Discernment of grains and determination of grain size are partially subjective.

Grain size is measured by sorting of loose grains and by examination of two-dimensional sections. Grains can be loosened from certain sediments and sedimentary rocks by agitation or by the dissolution of cementing material in an appropriate solvent.

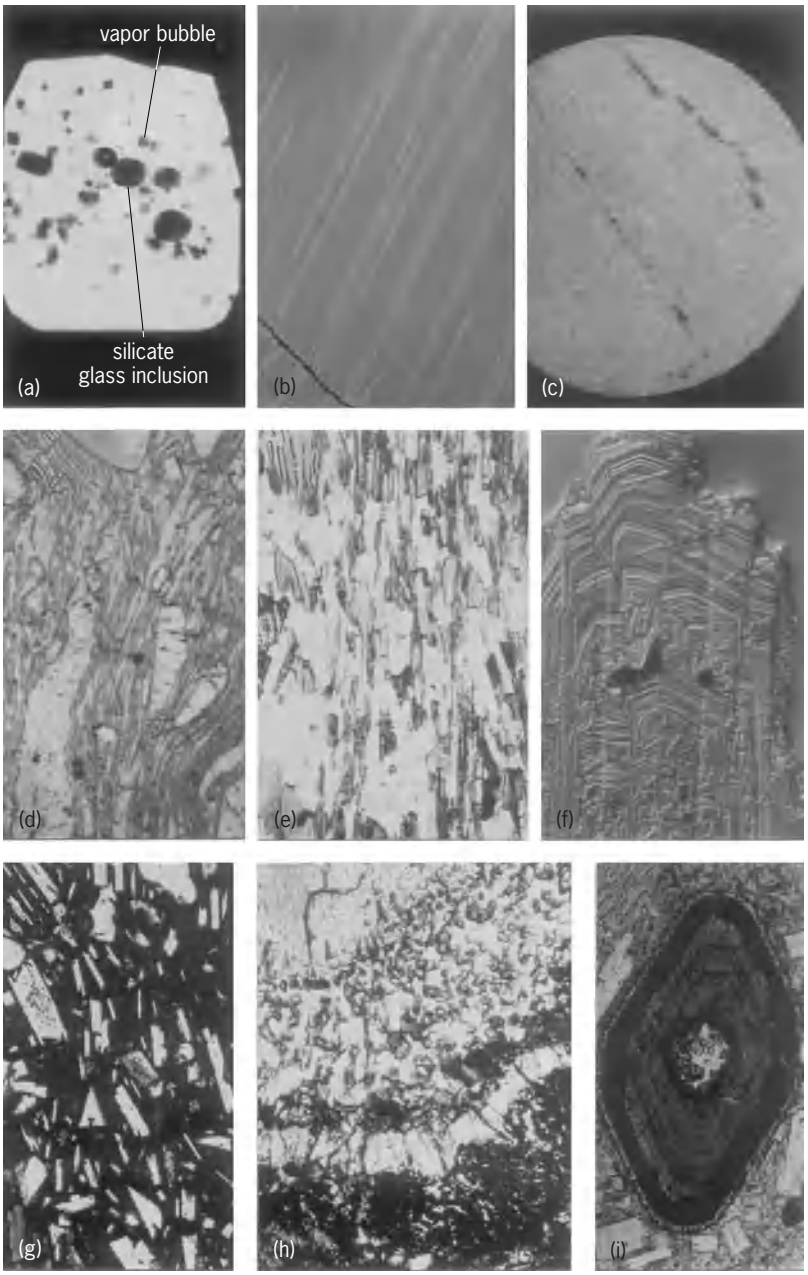

Examples of petrographic features. Images *b, c*, and *f* are reflected light images; the crystal in *f* has been etched and coated with a reflective coating to reveal contrasts. Images *d, e, g, h*, and *i* are transmitted light images, all in plane polarized light. Image *a* shows an anisotropic crystal of olivine under crossed polarized light. (*a*) Faceted crystal of olivine (Mg_2SiO_4) erupted from Kilauea volcano in 1959. (*b*) Lamellar intergrowth of hematite (Fe_2O_3, in light gray) and ilmenite ($FeTiO_3$). (*c*) Globule of sulfide minerals surrounded by basaltic glass from a submarine lava near Kilauea volcano. (*d*) Welded ash flow tuff from the Jemez Mountains, New Mexico. (*e*) Schist with aligned crystals of mica and sillimanite from southeastern Quebec, Canada. (*f*) Crystal of plagioclase ($CaAl_2Si_2O_8$–$NaAlSi_3O_8$) erupted from Fuego volcano, Guatemala, 1974. (*g*) Crystals in an andesitic lava flow from Mount Shasta, California. (*h*) Zones of minerals in a metamorphic rock. (*i*) Crystal of amphibole in an andesite from Mount Shasta.

Loose grains may be freed from igneous rocks such as some tuffs and pumices by dissolving the glass in an appropriate acid. The grains may then be sorted mechanically either by using sieves or by measuring the dimensions of individual grains. Alternatively the disaggregated grains may be sorted dynamically according to their rate of fall through air or water. For certain purposes the dynamical size is more important than the geometrical size.

The measurement of grain size using sections is complex because grains have three dimensions. The grain size evident in a section is smaller than that revealed by measurement of loose grains. Shape and orientation effects render interpretation of measurements in sections difficult. Crystals and grains may also be imaged using x-rays.

Grains in a rock vary in size. The size distribution characterizes the variation and has important implications for origin: for example, well-sorted sands form in environments where currents deliver and deposit grains of similar settling velocity. Even-grained, igneous rocks may suggest a particular range of cooling rates. Size grading of grains (large grains concentrated near the base of a particular unit) suggests that the grains settled through a liquid or gas. Examples exist in both sedimentary (water-laid) and igneous (melt-laid) rocks.

Grain shapes are measured by two parameters: sphericity (equidimensionality) and roundness (smoothness). Most crystals develop flat faces and sharp, angular edges and corners as they grow from liquids and gases, but are rounded by abrasion and solution. Needles and wafers may have the same sphericity and roundness, but they differ in shape. Some crystals in igneous rocks may have the same shape, but they differ in habit if their bounding faces differ crystallographically. The surface texture of particles may be revealed at high magnification with the aid of a scanning electron microscope. The surface features may indicate impacts between grains, etching of defects, overgrowths, and so on. Shape is a complex, varied, and revealing feature of grains.

Voids. These have size and shape also. Round voids (vesicles) surrounded by many grains indicate shaping due to surface tension and therefore the presence or former presence of a liquid, because most solids are too strong to be shaped by the weak forces of surface tension. Cavities with angular, faceted crystals projecting into them suggest growth of crystals from a gas or liquid.

Minerals in grains and rocks. Minerals, and sometimes their composition, are identified microscopically according to their optical properties. Compositional zoning is common and can reveal the progressive crystallization of igneous rocks. More complete and quantitative chemical analyses are possible at the scale of 0.0002 in. (0.005 mm) with the electron probe x-ray microanalyzer. Ion-beam microanalyzers can determine ages of crystallization of some grains or parts of grains in a rock. The compositions of certain associated minerals in rocks can help establish the temperature and pressure of crystallization or recrystallization, if the minerals were in equilibrium (zoned minerals cannot be in equilibrium).

The occurrence of certain garnets and diamonds in some rocks indicates that the rocks are derived from depths of 60 mi (100 km) or more. The identity and composition of certain minerals in sedimentary rocks can suggest certain kinds of source rocks, the weathering, and tectonic environments. For example, because uraninite (U_3O_8) is readily oxidized and dissolved, round grains of it in certain ancient conglomerates suggest transport in an environment with much less atmospheric oxygen than at present.

The proportions of various minerals in a rock are measured under the microscope. Commonly a mechanical stage moves a thin section by fixed increments, and the minerals encountered at the grid points are added up. The result is called the mode. It must be corrected for a thickness effect if some of the minerals occur mostly as grains as small as, or smaller than, the thickness of the section. Other methods are used to find the mode of loose grains or of polished surfaces. The mode together with the grain size generally suffices to name the rock. *See* CHEMICAL MICROSCOPY; REFLECTING MICROSCOPE; SECONDARY ION MASS SPECTROMETRY (SIMS); X-RAY FLUORESCENCE ANALYSIS; X-RAY SPECTROMETRY.

Structural modification within minerals. Structural characterization of certain minerals can help elucidate the cooling history of igneous rocks. The degree of order in certain mineral structures can be revealed optically or by x-ray crystallography or Mössbauer spectral analyses. Such properties may be inherited by the sand grains in sedimentary rocks; consequently, it may be possible to recognize sandstones derived from eroded volcanoes. *See* MÖSSBAUER EFFECT; X-RAY CRYSTALLOGRAPHY.

Intergrowths of crystals within grains. Mineral intergrowths provide a vast array of textures useful in interpreting how rocks form, but are difficult to generalize. Intergrowths within grains reveal twinning, alteration, and unmixing. Twinning may reveal certain conditions of growth or recrystallization in the presence of nonuniform stress. Particles of clay dispersed throughout minerals indicate partial alteration at temperatures below a few hundred degrees Celsius. Regular intergrowths of feldspars, pyroxenes, amphiboles, and oxide and sulfide minerals commonly are interpreted in terms of unmixing of an initially uniform (homogeneous) solid during cooling. The widths and compositions of the separate portions help evaluate the cooling history. For example, the lamellae of nickel-rich alloy in the metal of meteorites have been used to infer the cooling rate and size of the parent body. Certain regular intergrowths between unlike minerals such as quartz and feldspar are interpreted to result from competing, simultaneous growth. Sieve textures where many small grains are included as islands in a large crystal of another mineral may indicate successive growth. Reaction and solution of the smaller grains are suggested if they are widely separated and if it can be

inferred that initially the small grains formed an interlocking or touching aggregate. The various textures are named in an elaborate terminology that is falling into disuse in favor of simple words, pictures, and drawings. *See* CRYSTALLOGRAPHY.

Orientations of grains. Fabric and grain distribution in rocks suggest various processes of rock formation. The alignment of large crystals suggests laminar flow or deformation. If aligned parallel to the boundary of a rock body, such fabric may be the principal evidence of igneous origin. Parallel arrangements of platy or leaflike crystals yield a foliation (which means something else to foresters). Foliation may reveal the deformational history of the rock. For example, slates are metasedimentary rocks which recrystallized in a stress field such that micas grew parallel to each other. The shiny flat surfaces of slates are formed by the cleavage surfaces of aligned micas.

The crystallographical orientations are determined by studying thin sections with a device called a universal stage. It consists of two hemispheres of glass between which the rock section is sandwiched. Three or more mutually perpendicular axes of rotation allow the section to be rotated in three dimensions. The rotations are necessary to determine the orientations of the grains. The orientations of grains in rocks help reveal how the grains were deposited or deformed. The petrographer can determine from the orientations whether a particular rock has rotated loosely relative to bedrock, and can thereby guide engineers in their search for firm foundations for bridges and dams. *See* PETROFABRIC ANALYSIS.

Liquids and gases in rocks. Many crystals in ores contain inclusions of gas and liquid. The freezing temperatures of the inclusions help reveal the salinities of ore-depositing solutions. The temperatures of homogenization of liquid and vapor can yield information about the temperature and pressure at the time of ore formation.

Applications. Petrographers study organic as well as inorganic objects, and petrographic analyses are useful to both paleontologists and petroleum geologists. The quality of coal is revealed with polarizing and reflecting light microscopes. Fossil wood was one of the first objects studied with a polarizing microscope, in 1827. The daily and seasonal tree-ring-like increments of growth of certain corals and mollusks are revealed microscopically, and have permitted paleontologists to estimate the greater number of days in the year hundreds of millions of years ago. Tiny fossils in cherty rocks about 2 billion years old reveal the existence of algae at a time when the atmosphere probably began to contain free oxygen. The families of shell fragments in limestones are evident under the microscope, and are routinely used for purposes of correlation and characterization of the environment of deposition. Inclusions of petroleum and brine in crystals of silicates and salt in rocks help scientists infer how petroleum formation is connected with cementation and other modifications of buried sediments. *See* PALEONTOLOGY; PETROLEUM GEOLOGY.

Petrographers also study synthetic objects. The textures of metals and alloys are scrutinized by petrographers in order to understand what makes these materials strong and resistant to corrosion. Flaws in glasses and ceramics are revealed by microscopical and polarizing techniques. Rims of hydrated glass on obsidian artifacts increase in thickness with time after burial and help anthropologists date them. The microscopic roughness of the cutting edges of stone tools can reveal their principal use. Fragments of minerals and rocks in some pottery can help point to its source and help trace prehistoric routes of trade. The industrial, agricultural, and natural sources of particles in the air and water may be established from petrographic study. *See* MINERALOGY; PETROLOGY.

Alfred T. Anderson, Jr

Bibliography. H. Blatt, R. J. Tracy, and B. Owens, *Petrology: Igneous, Sedimentary, and Metamorphic*, 3d ed., 2005; R. V. Dietrich and B. J. Skinner, *Gems, Granites, and Gravels: Knowing and Using Rocks and Minerals*, 1990; H. B. Milner, *Sedimentary Petrography: Geology and Mineralogy*, 2003; A. Philpotts, *Petrography of Igneous and Metamorphic Rocks*, 1988; L. A. Raymond, *Petrology: The Study of Igreous, Sedimentary and Metamorphic Rocks*, 2d ed., 2001.

Petrolatum

A smooth, semisolid blend of mineral oil with waxes crystallized from the residual type of petroleum lubricating oil. The wax molecules contain 30–70 carbon atoms and are straight chains with a few branches or naphthene rings. They are microneedles and hold a large amount of oil in a gel. Petrolatums are useful because they cling, lubricate, and resist both moisture and oxidation. They serve as lubricants in baking and candymaking; as carriers in polishes, cosmetics, and ointments; as rust preventives; as waterproofing agents for paper; and in other uses calling for an inert greaselike material. *See* LUBRICANT; PETROLEUM PRODUCTS; WAX, PETROLEUM.

J. K. Roberts

Petroleum

Unrefined, or crude, oil is found underground and under the sea floor, in the interstices between grains of sandstone and limestone or dolomite (not in caves). Petroleum is a mixture of liquids varying in color from nearly colorless to jet black, in viscosity from thinner than water to thicker than molasses, and in density from light gases to asphalts heavier than water. It can be separated by distillation into fractions that range from light color, low density, and low viscosity to the opposite extreme. In places where it has oozed from the ground, its volatile fractions have vaporized, leaving the dense, black parts of the oil as a pool of tar or asphalt (such as the Brea Tar Pits in California). Egyptians used such tar for embalming the dead, and Mesopotamians used

it for adhering bricks together. Much of the world's crude oil is today produced from drilled wells. *See* PETROLEUM ENGINEERING.

Composition. Petroleum consists mostly of hydrocarbon molecules. The four main classes of hydrocarbons are paraffins (also called alkanes), olefins (alkenes), cycloparaffins (cycloalkanes), and aromatics. Olefins are absent in crude oil but can be formed in certain refining processes.

Alkanes consist of straight or branched chains of carbon atoms surrounded by hydrogen atoms. Each carbon atom is capable of forming four chemical bonds with other atoms, while each hydrogen atom can from only one bond. The simplest hydrocarbon is one carbon atom bonded to four hydrogen atoms (chemical formula CH_4), and is called methane (structure **1**). The proportions of hydrogen

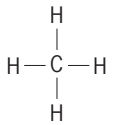

(1)

and carbon atoms in the alkane series is expressed as C_nH_{2n+2}, where n is the number of carbon atoms in each molecule. *See* ALKANE; PARAFFIN.

Branched-chain alkanes are called isoalkanes. An example is isooctane (2,2,4-trimethylpentane, C_8H_{18}). Sometimes the carbon atoms are joined as a closed ring, called a cycloalkane. This can occur only for a ring of three to seven carbon atoms, although only five- and six-membered-ring compounds occur naturally. *See* ALICYCLIC HYDROCARBON.

When carbon atoms are joined by a double bond in a straight or branched chain, they are called olefins (alkenes). If double bonds occur in a ring, only every other carbon-carbon bond can be double. The simplest six-membered (aromatic) ring is called benzene (C_6H_6; structure **2**). Napthalene ($C_{10}H_8$; structure **3**)

(2)

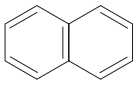

(3)

is formed when a four-carbon chain containing two double bonds is attached at each end to adjacent carbons in a benzene ring. It is also possible to have aromatics with multiple fused rings. *See* AROMATIC HYDROCARBON.

Petroleum usually contains all of the possible hydrocarbon structures except alkenes, with the number of carbon atoms per molecule going up to a hundred or more (see **table**). These fractions include compounds that contain sulfur, nitrogen, oxygen, and metal atoms. Sulfur and oxygen may be covalently bonded to nonring hydrocarbon compounds or may occur as one atom of a carbon ring. Nitrogen is found only as one atom of a ring. Metal atoms occur in the lubricating oil range of some crude oils as salts of organic acids (similar to acetic acid but much larger). The proportion of compounds containing these atoms increases with increasing size of the molecule.

Range of carbon numbers per molecule for the fractions of petroleum commonly produced by distillation

Carbon numbers	Fraction
1–2	Natural gas
3–4	Liquefied petroleum gas (LPG)
5–11	Gasoline
12–14	Kerosene or jet fuel
15–19	Fuel oil or diesel fuel
20–40	Heavy gas oil, light to medium lubricating oils
40+	Asphalt (residue left by distillation)

Asphaltic molecules contain many cyclic compounds in which the rings contain sulfur, nitrogen, or oxygen atoms; these are called heterocyclic compounds. An example is pyridine (structure **4**).

(4)

There can also be porphyrin structures having four nitrogen-containing rings with the nitrogens facing each other and linked weakly to a metal atom between the four. Some refineries use processes to remove the metals from heavy gas oil fractions before catalytic cracking (a process used to reduce the molecular weight of hydrocarbons by breaking molecular bonds). This is because the metals may deposit on the cracking catalyst and make it less active. *See* CRACKING; HETEROCYCLIC COMPOUNDS; PORPHYRIN.

Asphalt contains a high proportion of heavy, viscous oil in the lubricating oil range, plus a number of high-molecular-weight aromatic-heterocyclic solids (asphaltenes). The asphaltenes are dispersed in the oil in a very fine colloidal form by similar molecules of about half the molecular weight, called resins. Typical road asphalt contains about 25% asphaltenes, 5–15% resins, and the rest viscous oil, as does the asphalt used in roofing tar and shingles. *See* ASPHALT AND ASPHALTITE; PAVEMENT; ROOF CONSTRUCTION.

Composition changes in refining. In refining petroleum at the present time, it is normally necessary to change the hydrocarbon structures to improve the amount and performance of the desired products. Crude oil usually has too many heavy fractions. This is remedied by catalytically cracking the heavy gas oil or even thermally cracking the asphalt and making gas, gasoline, and fuel oil from it. The gasoline obtained directly from the crude oil has too low an octane number for use in modern automobiles, so it is catalytically reformed or isomerized to raise the octane number. Catalytic refining processes use catalysts which speed the reactions and preferentially make the desired products. Sometimes the small olefins resulting from catalytic and thermal cracking can be linked together by another catalyst to make a gasoline-range isoalkane. This was how isooctane was first made. Alkylate (a gasoline component)

made this way is currently used in premium gasoline. *See* GASOLINE; OCTANE NUMBER; PETROCHEMICAL; PETROLEUM PROCESSING AND REFINING; PETROLEUM PRODUCTS.

Lubricating oils, which have had wax and aromatics removed, are often treated catalytically with hydrogen to remove undesired trace components. The asphalt is air-blown to make it more suitable for manufacturing roof shingles or for paving roads. In air-blowing, part of the resins are converted into asphaltenes and the product becomes stiffer, better able to stand hot weather.

Some of the intermediate refining products go to petrochemical plants where they are converted catalytically into rubber, plastics, and textile fibers.

The yield of valuable products from modern refineries is usually slightly greater than the amount of crude oil from which they are made. Chemical additives are used in most refinery products to improve their end-use performance.

Origin. It is generally agreed that petroleum formed by processes similar to those which yielded coal, but was derived from small animals rather than from plants. Geophysical knowledge of plate tectonics has revealed that moving plates of the Earth's surface layers have piled up mountain ranges where the plates collided. The erosion of these mountains by water and ice, wind, and gravity led to the grinding up of enormous amounts of rock into sand, silt, and mud which settled to the bottom of lakes, rivers, and seas. The lakes and seas teemed with mostly small organisms. Some of these organisms extracted calcium from the salts in the water and reacted the calcium with carbon dioxide to form calcium carbonate shells that protected them from predators.

Dead organisms have been buried in mud over millions of years. Further layers deposited over these mud layers have in some cases reached a thickness of thousands of feet, and compacted the layers beneath them, until the mud has become shale rock. When a layer was particularly rich in broken sea shells, it was compacted into limestone. Sandy layers became sandstone. The deep interior of the Earth is as hot as the Sun's surface, and heat moving up toward the Earth's surface makes the temperature in a given layer hotter the deeper it is buried. Thus, the mud layers were heated and compressed by the layers above. The bodies of the organisms in the mud were decomposed and converted into fatty liquids and solids. Heating these fatty materials over a very long time caused their molecules to break into smaller fragments and combine into larger ones, so the original range of molecular size was spread greatly into the range found in crude oil. Bacteria were usually present, and helped remove oxygen from the molecules (which began mostly as carbohydrates, comprising carbon, oxygen, and hydrogen) and turned them into hydrocarbon compounds. The great pressure of the overlying rock layers helped to force the oil out of the compacted mud (shale) layers into less compacted limestone, dolomite, or sandstone layers next to the shale layers. *See* DOLOMITE; LIMESTONE; ORGANIC GEOCHEMISTRY; PETROLEUM GEOLOGY; SANDSTONE; SEDIMENTOLOGY; SHALE.

Water with dissolved salts is present everywhere underground. Oil is less dense than water, so it moves up through the permeable rock wherever it can go. If it can reach the surface, it does. But if the rock layer which it enters has been bent by tectonic processes into a ridge or a dome structure, the oil is trapped there. A liquid's pressure in underground porous rock is equal to the pressure of an equal depth of water, and can even be higher. Trapped oil is under enough pressure that it can keep in solution a considerable amount of the hydrocarbon gas made by the oil generation processes.

At depths greater than about 25,000 ft (7620 m), the temperature is so high that the oil conversion processes go all the way to natural gas and soot. Crude oil is found generally at depths of a few hundred feet to about 20,000 ft (6100 m). Layers have often had overlying layers eroded away, so strongly converted oil can be found at shallower depths. Natural gas formed by the conversion processes is now also found over a variety of depths which do not indicate the depth and temperature of their origin. *See* NATURAL GAS.

Nature of oil reservoirs. The oil formed by the natural thermal and bacterial processes was squeezed out of the compacting mud layers into sandstone or limestone layers and migrated upward in tilted layers. Tectonic processes caused such uptilting and bulging of layers to form ridges and domes. When the ridges and domes were covered by shale already formed, the pores of the shale were too tiny to let the oil through, so the shale acted as a sealing cap. In some cases, the migrating oil came to a place where the tilted sandy or limy layer through which they were oozing had been eroded away and a new layer of shale formed as a horizontal layer above, cutting off the tilted layer. When the oil could not rise farther, it was trapped. Often, a number of layers were bent or folded at the same time, and the oil migrated simultaneously into a number of stacked layers in a given ridge or domal structure, underlain and surrounded by water. Porous rock in such a structure that contains oil or gas is called an oil or gas reservoir.

Over geologic time, sandy deserts have been formed on land. The sand was later submerged and covered with mud, which became shale. Oil has in some cases migrated into such buried desert sands and been trapped.

Shallow seas and vast lakes have existed in past times and were later buried, then uplifted. Rivers with sand bars and muddy bottoms and deltas have also contributed source rocks and reservoir rocks.

Certain mixtures of sand, silt, and mud retained their animal content while the conversion processes generated hydrocarbons, and in some cases tarry oils, which probably remained where they were generated. The Athabasca Tar Sands and the Orinoco Tar Belt may have formed this way. The heavy oil from these deposits requires preprocessing, that is, thermal cracking or hydrocracking, prior to refinery

processing. It is therefore less economical to process, relative to lighter crude oils. *See* OIL SAND.

In other cases, the organic fatty acids, which could have been thermally converted to hydrocarbon liquids, reacted with alkaline minerals present in their source rock, as did the oil shale deposits in Colorado and elsewhere. Organic acids are characterized by the formula RCOOH, where R is the large hydrocarbon end of the molecule, and H is a positively charged hydrogen ion which can react with an alkali (in this case, sodium bicarbonate, HOCOONa) to form water, HOH, and RCOOCOONa. Upon heating in a pyrolysis reactor, part of the compound is converted to CO_2 (carbon dioxide), the alkali is regenerated, and the rest becomes a hydrocarbon vapor, which can be condensed to produce shale oil. This oil is usually high in sulfur and nitrogen and metal compounds, and before entering a normal refining processes, it must be hydrotreated to reduce the amount of these deleterious compounds. The hydrogen necessary for the hydrotreating can be made from part of the hydrocarbons produced, but this consumes a significant part of the shale oil. Disposal of the spent alkaline rock is also a potential environmental problem. Thus, the recovery of oil suitable for further refining from shale oil deposits is generally not economical, compared to normal petroleum. *See* OIL SHALE.

Even with light crude oils, the recovery from typical reservoirs is not as high as might be thought. Multiple-layer reservoirs will typically contain oil-bearing layers with a wide range of permeability. When recovery from the highest-permeability layers is as complete as it can be, the low-permeability layers will usually have been only slightly depleted, despite all efforts to improve the recovery. The approximate success of various oil recovery processes is as follows (varying from case to case): primary depletion (by solution gas liberation and pumping), ~10–20%; secondary water flooding, ~15–35%; tertiary processes (propane slug, carbon dioxide slug), ~10–15%; steam flooding (including steam soak and backflow), ~5–40%. The recovery from primary, secondary, and tertiary processes is additive, from steam flooding is not. Despite recovery efforts, half or more of the oil originally present in United States oil reservoirs is still in them. *See* PETROLEUM ENHANCED RECOVERY; PETROLEUM RESERVES; PETROLEUM RESERVOIR ENGINEERING.

Oil is still being formed by natural processes—perhaps as much as hundreds or a few thousands of barrels per year. Of course, formation cannot keep up with consumption, so known oil and gas deposits are being depleted at such a rate that most of these will be consumed by some time in the twenty-first century. Elmond L. Claridge

Heavy oil. Heavy oil and tar sand oil (bitumen) are petroleum hydrocarbons found in sedimentary rocks. They are formed by the oxidation and biodegradation of crude oil, and occur in the liquid or semiliquid state in limestones, sandstones, or sands. *See* BITUMEN.

These oils are characterized by their viscosity; however, density (or API gravity) is also used when viscosity measurements are not available. Heavy oils have gas-free viscosities between 100 and 10,000 millipascal seconds (centipoise) at reservoir conditions, and their densities range from 1000 kg/m^3 (10°API) to 943 kg/m^3 (20°API) measured at 15.6°C and atmospheric pressure. Extraheavy oils or tar sand oil (bitumen) has gas-free viscosities greater than 10,000 millipascal seconds (centipoise) and densities greater than 1000 kg/m^3 (less than 10°API). Heavy oils contain 3 wt % or more sulfur and as much as 200 ppm vanadium. Titanium, zinc, zirconium, magnesium, manganese, copper, iron, and aluminum are other trace elements that can be found in these deposits. Their high naphthenic acid content makes refinery processing equipment vulnerable to corrosion. *See* OIL AND GAS FIELD EXPLOITATION.

Worldwide sources of heavy oils are estimated at over 950 billion cubic meters (6000 billion barrels) compared to conventional oil reserves of 258 billion cubic meters (1620 billion barrels). However, the true resource base is believed to be much higher since heavy oil discoveries are not well defined or documented. Presently, major heavy oil deposits are found in Canada, Venezuela, and Russia. Canada has the world's largest heavy oil reserves at the Athabasca Tar Sands, Cold Lake, and Peace River. Deposits in Venezuela are located in the Orinoco Belt. Russia has reserves in the Lena-Anabar basin (eastern Siberia) and Tataristan. There are other countries where such deposits are known to exist, including the United States (Utah, Texas, and California), Turkey (Bati Raman), China, and Indonesia. These resources are generally found within 1500 m (4500 ft) of the Earth's surface. Production, transportation, and refining these hydrocarbons require technologies that are expensive and environmentally sensitive.

Heavy oils that are in the liquid state in the reservoir as well as at surface conditions can be recovered using techniques known as steam injection, in situ combustion, or chemical flooding with conventionally drilled oil wells. These processes, known as in situ recovery methods, all aim at decreasing the viscosity of the oil so that it will flow easily through the porous reservoir rock to the production wells. *See* PETROLEUM ENHANCED RECOVERY.

In steam injection, steam generated at the surface is injected down through the wells into the reservoir. Steam injection temperatures up to 250°C have been used. The viscosity of the oil decreases as the hot fluids contact the oil in the reservoir rock, increasing the flow toward the production wells.

For in situ combustion, air is injected down the well, and when it contacts the oil reservoir combustion starts with some initial downhole heating. The continuous injection of air or air-water mixtures helps drive the combustion front through the reservoir, displacing more oil toward the production wells. Temperatures are as high as 500°C in this application. Both of these heat injection methods require specialized production equipment to withstand the high temperatures.

There are several different types of chemical flooding; however, the method known as carbon dioxide flooding has found application in heavy oil fields. Carbon dioxide is injected into the reservoir through the wells and dissolves in the oil, thus reducing its viscosity as it moves through the porous rock. Oil is recovered by the continuous injection of carbon dioxide or injection and production from the same wellbore, known as the huff and puff method.

The application of these methods may increase heavy oil recovery from 1–3% to 10–30% of the oil in place.

Extraheavy oils or tar sand bitumens—semisolid or solid hydrocarbons found close to the Earth's surface—are mined. Following the mining operation, the tar sands are crushed into smaller-sized particles, and then the bitumen is extracted. There are three main processes for extracting bitumen from tar sands. The hot water process uses sprayed hot water, often with added steam and caustic to separate the bitumen from the sand. The resulting mixture is sent to a flotation cell to separate the sand from the bitumen. The bitumen is treated to reclaim the solvents, and clean bitumen is sent for various end uses. In the solvent extraction process, crushed sand is mixed with an organic solvent, and then the dissolved bitumen and sand are separated. Following processing, the solvent is treated for recycling. The thermal processes involve the thermal decomposition of the bitumen. Crushed sand is sent to a reactor where it is heated. The product is a thermally cracked liquid, and the by-products are coke and sand. Another type of thermal process, known as the Taciuk process, has been developed and tested in Canada. The oil produced by the above methods is called synthetic crude oil or syncrude. With today's technology it is possible to produce 0.159 m^3 (1 barrel) of syncrude from 2 tons of tar sand. *See* SYNTHETIC FUEL. Ender Okandan

Bibliography. G. D. Hobson (ed.), *Modern Petroleum Technology*, 5th ed., 1984; G. D. Hobson and E. N. Tiratsoo, *Introduction to Petroleum Geology*, 2d ed., 1985; R. R. Kinghorn, *An Introduction to the Physics and Chemistry of Petroleum*, 1983; E. Okandan (ed.), *Heavy Crude Oil Recovery*, Martinus Nijhof, 1984; J. H. Tatsch, *Petroleum Deposits; Origin, Evolution, and Present Characteristics*, 1974; B. P. Tissot, *Petroleum Formation and Occurence*, Springer-Verlag, 1984; *Unitar 7th International Conference on Heavy Crude and Tar Sands*, Beijing, China, October 27–30, 1998; A. J. Walker, *The Oil and Gas Book*, 1985; R. R. Wheeler and M. Whited, *Oil from Prospect to Pipeline*, 5th ed., 1985.

Petroleum engineering

An eclectic discipline comprising the technologies used for the exploitation of crude oil and natural gas reservoirs. It is usually subdivided into the branches of petrophysical, geological, reservoir drilling, production, and construction engineering. After an oil or gas accumulation is discovered, technical supervision of the reservoir is transferred to the petroleum engineering group, although in the exploration phase the drilling and petrophysical engineers have played a role in the completion and evaluation of the discovery.

Petrophysical engineering. The petrophysical engineer is perhaps the first of the petroleum engineering group to become involved in the exploitation of the new discovery. By the use of down-hole logging tools, hydrocarbon analysis from the circulating mud and drill cuttings, and visual or laboratory analysis of cores or drill cuttings, the petrophysicist estimates the porosity, permeability, and fluid content of the reservoir rock. In addition, the petrophysical engineer is called upon to assist the geologist, geophysicist, and drilling engineer with subsurface information. *See* WELL LOGGING.

Geological engineering. The geological engineer, using petrophysical and petrographic data, seismic surveys conducted during the exploration operations, and an analysis of the regional and environmental geology, develops inferences concerning the lateral continuity and extent of the reservoir. However, this assessment usually cannot be verified until additional wells are drilled and the geological and petrophysical analyses are combined to produce a firm diagnostic concept of the size of the reservoir, the distribution of fluids therein, and the nature of the natural producing mechanism. Additional geophysical data, particularly three-dimensional seismic, may be useful to help delineate a complex field. As the understanding of the reservoir develops with continued drilling and production, the geological engineer, working with the reservoir engineer, selects additional drill sites to further develop and optimize the economic production of oil and gas. *See* PETROLEUM GEOLOGY; SEISMIC EXPLORATION FOR OIL AND GAS.

Reservoir engineering. The reservoir engineer, using the initial studies of the petrophysicist and geological engineers together with the early performance of the wells drilled into the reservoir, attempts to assess the producing rates (barrels of oil or millions of cubic feet of gas per day) that individual wells and the entire reservoir are capable of sustaining. One of the major assignments of the reservoir engineer is to estimate the ultimate production that can be anticipated from both primary and enhanced recovery from the reservoir. The ultimate production is the total amount of oil and gas that can be secured from the reservoir until the economic limit is reached. The economic limit represents that production rate which is just capable of generating sufficient revenue to offset the cost of operating the reservoir. The proved reserves of a reservoir are calculated by subtracting from the ultimate recovery of the reservoir (which can be anticipated using available technology and current economics) the amount of oil or gas that has already been produced. *See* PETROLEUM RESERVES.

Primary recovery operations are those which produce oil and gas without the use of external energy except for that required to drill and complete the

wells and lift the fluids to the surface (pumping). Enhanced oil recovery, or supplemental recovery, is the amount of oil that can be recovered over and above that producible by primary operation by the implementation of schemes that require the input of significant quantities of energy. In modern times waterflooding has been almost exclusively the supplementary method used to recover additional quantities of crude oil. However, with the realization that the discovery of new petroleum resources will become an increasingly difficult achievement in the future, the reservoir engineer has been concerned with other enhanced oil recovery processes that promise to increase the recovery efficiency above the average 33% experienced by the United States (which is somewhat above that achieved in the rest of the world). The restrictive factor on such processes is the economic cost of their implementation. *See* PETROLEUM ENHANCED RECOVERY.

In the past the reservoir engineer was confined to making predictions of ultimate recovery by using analytical equations for fluid flow in the framework of an overall definition of the geological and lithological description of the reservoir. Extensive references to reservoirs that are considered to have analogous features, and history matching, such as curve fitting and extrapolation of the declining production, are important tools of the reservoir engineer. The analytical techniques are limited in that reservoir heterogeneity and the competition between various producing mechanisms (solution gas drive, water influx, gravity drainage) cannot be accounted for, nor can one reservoir be matched exactly by another. The reservoir engineer has also been able to use mathematical modeling to simulate the performance of the reservoir. The reservoir is divided mathematically into segments (grid blocks), and the appropriate flow equations and material balances are repeatedly applied to contiguous blocks. Extensive history matching is still required to achieve a reliable predictive model, but this can usually be achieved more quickly and reliably than with the use of analytical solutions and analogy. Numerical simulation is almost routine with the power of desktop computers; complex simulation of enhanced recovery processes such as thermal recovery or miscible flooding may be most effectively run on a powerful mainframe computer. *See* COMPUTER; MICROCOMPUTER; PETROLEUM RESERVOIR ENGINEERING; PETROLEUM RESERVOIR MODELS.

Drilling engineering. The drilling engineer has the responsibility for the efficient penetration of the earth by a well bore, and for cementing of the steel casing from the surface to a depth usually just above the target reservoir. The drilling engineer or another specialist, the mud engineer, is in charge of the fluid that is continuously circulated through the drill pipe and back up to surface in the annulus between the drill pipe and the borehole. This mud must be formulated so that it can do the following: cool the bit; carry the drill cuttings to the surface, where they are separated on vibrating screens; gel and hold cuttings in suspension if circulation stops; form a filter cake over porous low-pressure intervals of the earth, thus preventing undue fluid loss; and exert sufficient pressure on any gas- or oil-bearing formation so that the fluids do not flow into the well bore prematurely, blowing out at the surface. In special situations, air may be used instead of mud. As drilling has gone deeper and deeper into the earth in the search of additional supplies of oil and gas, higher and higher pressure formations have been encountered. This has required the use of positive-acting blow-out preventers that can firmly and quickly shut off uncontrolled flow due to inadvertent unbalances in the mud system. Important new technology being used routinely by the drilling engineer includes downhole motors to rotate the bit, novel bit designs, measurement of borehole and formation characteristics while drilling, and horizontal drilling. *See* OIL AND GAS WELL DRILLING.

Production engineering. The production engineer, upon consultation with the geological, petrophysical, reservoir, or completion engineers, plans the completion procedure for the well. This involves a choice of setting a liner across the formation or perforating a casing that has been extended and cemented across the reservoir, selecting appropriate pumping techniques, and choosing the surface collection, dehydration, and storage facilities. The production engineer also compares the productivity index of the well (barrels per day per pounds per square inch of drawdown around the well bore) with that anticipated from the measured and inferred values of permeability, porosity, and reservoir pressure to determine whether the well has been damaged by the completion procedure. Such comparisons can be supplemented by a knowledge of the rate at which the pressure builds up at the well bore when the well is abruptly shut-in. Using the principles of unsteady-state flow, the reservoir engineer can evaluate such a buildup to assess quantitatively the nature and extent of well bore damage. Damaged wells, or wells completed in low-permeability formations, can be stimulated by acidization, hydraulic fracturing, additional perforations, or washing with selective solvents and aqueous fluids. *See* OIL AND GAS WELL COMPLETION.

Construction engineering. Major construction projects, such as the design and erection of offshore platforms, require the addition of civil engineers to the staff of petroleum engineering departments, and the design and implementation of natural gasoline and gas processing plants require the addition of chemical engineers. *See* CIVIL ENGINEERING; OIL AND GAS, OFFSHORE; PETROLEUM; PETROLEUM PROCESSING AND REFINING.

The technology has become increasingly sophisticated and demanding with the implementation of new recovery techniques, deeper drilling, and the expanding frontiers of the industry into the hostile territories of arctic regions and deep oceans.

Todd M. Doscher; R. E. Wyman

Data analysis and mapping. In petroleum exploration, relational databases and advanced computer graphics have radically altered the work environment of computers from processing applications to interpretive usages. There is a heavy emphasis on facile

gathering of data and extraction of selected items to provide effective displays and interpretations.

The analysis of subsurface petroleum information poses a dual problem: there is either too much or insufficient data for any given application. In exploration, the volumes of data from seismic processing and interpretation can strain the limits of most mass-storage devices, and sophisticated processing routines can require supercomputers. On the other hand, data from wells can be minimal to nonexistent and may require extreme extrapolation and model-based solutions. Production applications can also be computer-intensive. As an example, reservoir oil simulation is feasible only in advanced modules on supercomputers. In most cases, there is no happy medium. Because of this dichotomy, computing environments in the petroleum industry run the full range from laptop field applications to main-frame installations. The computing requirements have become visually oriented, requiring a large dependency on interactive, graphic-display, and visualization routines.

In general, petroleum computing can be viewed on three levels: geological computing, geophysical computing, and engineering applications. Geological computing trends have focused on database and spatial system configurations, with specialty applications such as cross-section balancing or geochemical modeling. Geophysical computing tends to be computer-intensive; interpretive installations are, like all interactive workstation environments, driven by graphics. Engineering applications are also computer-intensive; they are generally classified as either simulation or process types.

Geological computing. Proliferation of high-quality workstation environments has given an individual geoscientist the tools to create maps and perform analyses on a large scale and to incorporate data to a much larger extent than was previously possible. The demand for data has led to the development of major products for exploration applications. Databases, language structures, and hardware that facilitate access and management of data have become a necessary part of computing environments in companies of all sizes.

In addition to database software, geographic information systems (GIS), initially developed for resource planning and cultural analysis, have been adapted to geological computing. The internal structure of geographic information systems differs from conventional computer-aided design (CAD) technology in the referencing of digital elements in a geographic (spatial) relation to all other elements in the database, not to layers (two-dimensional drawings) as symbols independent of their locations. *See* DATABASE MANAGEMENT SYSTEM.

Geophysical computing. Geophysical computing is based on an interpretive workstation dedicated to tasks of seismic interpretation. The display, manipulation, and interpretation of both two- and three-dimensional seismic data is done interactively using both stand-alone workstations and integrated mainframe systems. Seismic lines and traces are displayed in color, and the system permits interactive identification of horizons from line to line. Most of these programs have automated facilities for enhancement of attributes as well as for correction of errors in interpretation between seismic lines (mis-ties). The computer-aided interpretation frees the explorationist or production geophysicist from the extremely large volume of paper-record sections; at the same time it allows for consistent and realistic mapping data.

Engineering computing. The demands of engineering computing in development of petroleum resources are intense, and hardly any area remains untouched. The giant reservoir simulation routines known as black-oil simulators have given way to even more sophisticated segmented-reservoir simulators. These modeling packages generally require supercomputing speeds to adequately produce reasonable representations of reservoirs. The addition of geostatistics (kriging) and fractal geometries in the simulation procedure permits analysis of reservoir heterogeneities. Brian Robert Shaw

Bibliography. J. Gluyas and R. Swarbrick, *Petroleum Geoscience*, 2003; R. C. Laudon, *Principles of Petroleum Development Geology*, 1995; A. I. Levorsen, *Geology of Petroleum*, 2d ed., 1967, reprinted 2001; S. M. Luthi, *Geological Well Logs: Their Use in Reservoir Modeling*, 2001; M. Nikravesh, F. Aminzadeh, and L. A. Zadeh (eds.), *Soft Computing and Intelligent Data Analysis in Oil Exploration*, 2003; *Petroleum Engineer International*, monthly; R. C. Selley, *Elements of Petroleum Geology*, 2d ed., 1997; B. P. Tissot and D. H. Welte, *Petroleum Formation and Occurrence*, rev. ed., 1992.

Petroleum enhanced recovery

Technology to increase oil recovery from a porous formation beyond that obtained by conventional means. Conventional oil recovery technologies produce an average of about one-third of the original oil in place in a formation. Conventional technologies are primary or secondary. Primary technologies rely on native energy, in the form of fluid and rock compressibility and natural aquifers, to produce oil from the formation to wells. Secondary technologies supplement the native energy to drive oil to producing wells by injecting water or low-pressure gas at injection wells. The target of enhanced recovery technologies is that large portion of oil that is not recovered by primary and secondary means. *See* PETROLEUM ENGINEERING.

Process fundamentals. The boundary between secondary and enhanced recovery technologies is blurry. Many of the challenges encountered by secondary technologies are identical to those encountered by enhanced recovery technologies. Those challenges include reducing residual oil saturation, improving sweep efficiency, fitting the technology to the reservoir heterogeneities, and minimizing upfront and operating costs.

Residual oil remains trapped in a porous rock after the rock has been swept with water, gas, or any other

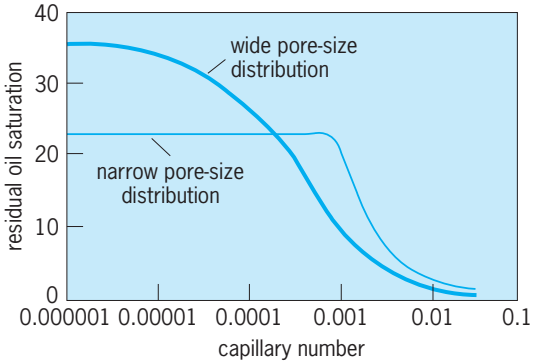

Fig. 1. The capillary number is the ratio of viscous forces to capillary forces in the rock, or simply the viscosity of the recovery agent times its velocity through the rock divided by the interfacial tension between the recovery agent and the oil. Capillary numbers for secondary technologies fall to the left of the graph; thus residual oil saturations for these processes lie between 20 and 40%.

recovery fluid. The residual oil saturation is the percentage of the pore space occupied by the residual oil. The residual oil saturation depends on the pore size distribution and connectivity, the interfacial tension between a recovery agent and the oil, the relative wettability of the rock surfaces with respect to the recovery agent and the oil, the viscosity of the fluids, and the rate at which the fluids are moving through the rock. The qualitative relationship between residual oil saturation and these various factors is shown in **Fig. 1**.

The sweep efficiency specifies that portion of a reservoir that is contacted by a recovery fluid. Sweep efficiency increases with volume of injected fluid. It also depends on the pattern of injection and production wells in a formation, on the mobility of the oil and the recovery fluid, and on heterogeneities in the formation (**Fig. 2**).

Oil reservoirs, which consist of porous rock formations, have complex structures created by the whims of nature in deposition, diagenetic, and tectonic processes. The combined effect of these processes yields formations that are heterogeneous at microscopic and macroscopic scales. The size distribution and connectivity of pores, the presence of clay particles in the pores, and the mineralogy of the surfaces of the pores all contribute to heterogeneity at the pore level. Layering or bedding of the formation produces heterogeneity at a slightly larger scale. At a larger scale, winding of river channels, anomalies in reef structures, and fractures caused by tectonic process contribute to additional heterogeneity. These various levels of heterogeneity affect the ease with which oil can be recovered. At the microscopic scale, heterogeneities affect residual oil saturation, as shown in Fig. 1. At larger scales, heterogeneities decrease sweep efficiency. Enhanced oil recovery technology must be selected or adapted to cope with the particular heterogeneities of each reservoir. *See* PETROLEUM GEOLOGY.

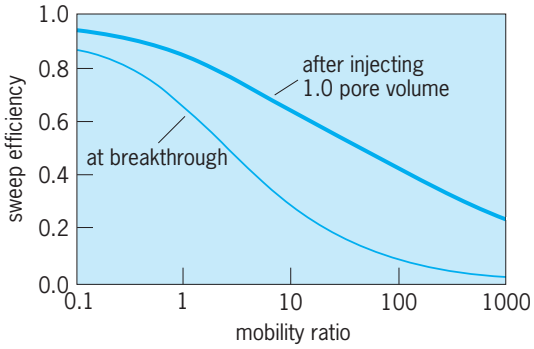

Fig. 2. The sweep efficiency of a five-spot pattern (a producer at each corner of a square area with an injector in the center) is shown as a function of the ratio of the mobility of the recovery agent to the mobility of the oil. Breakthrough sweep efficiency is that fractional volume swept when the recovery agent first arrives at the producing wells.

For economic reasons, access to reservoirs is extremely limited. The distance between wells that penetrate a formation is often 1000 ft or more. Limited access to the formation compounds the challenge of adapting recovery technology to cope with heterogeneity. Another consequence of limited access is that months of operation may be required to see the benefit of enhanced recovery technology because the rate at which fluids move through formation is a few feet per day. So, after starting injection of a recovery agent, one might wait as much as one year before any increased oil productivity results. The delayed benefit of a technology translates to increased economic risk. An enhanced recovery technology with high up-front and operating costs will likely fail to yield acceptable oil recovery performance. *See* OIL AND GAS WELL DRILLING.

Process technologies. A wide variety of processes have been considered for enhancing oil recovery: thermal processes, high-pressure gas processes, and chemical processes. The fundamental motivations for these processes are found in the relationships of Figs. 1 and 2. Specifically, low residual oil saturation can be obtained by selecting a recovery fluid that provides a very low interfacial tension between the oil and the fluid. With very low interfacial tension, the capillary number is large. And high sweep efficiency can be obtained by selecting a recovery agent with low mobility or by increasing the mobility of the oil.

In thermal processes, the mobility of the oil is increased by heating the oil-bearing formation, thus reducing the oil's viscosity. Thermal processes consist of cyclic steam injection, continuous steam drive, and in situ combustion. Typically, thermal processes have been applied to formations with high oil viscosity, usually greater than 100 millipascal seconds (centripoise). Such formations are rarely candidates for secondary recovery technologies, so the oil saturations are relatively high. Cyclic and continuous steam injection processes are widely used for recovery of high-viscosity oils in southern California, Venezuela, and western Canada. In the cyclic steam injection process, the same well serves as an injector

and producer. First, steam is injected for a period of one month. After a short soak period of one week, the same well is put on production for several months to a year. In continuous steam injection, the injected steam propagates through the reservoir as it moves from the injector toward the producer. Steam injection processes are limited to shallow reservoirs, less than 610 m (2000 ft) below the surface. For deeper reservoirs, or for reservoirs in formations less than 15 m (50 ft) thick, heat losses in the injection well and to surrounding formations can be too large for profitable operation with steam injection. For such reservoirs, in situ combustion processes may be more suitable. In these processes, air is injected to produce combustion of a portion of the oil in the formation. The remainder of the oil is driven ahead of the combustion front by steam and gases produced by combustion. *See* PETROLEUM.

Carbon dioxide and natural gas (mostly methane) dominate applications of high-pressure gas processes, although nitrogen is injected into some reservoirs. Gas injection processes are particularly suited to reservoirs with low permeability. With sufficiently high operating pressure, the interfacial tension between the injected gas and the displaced oil reduces to zero; that is, the gas becomes miscible with the oil. The pressure required to develop miscibility depends on the compositions of the gas and the oil, and on the reservoir temperature. Most miscible carbon dioxide applications operate between 10,342 and 17,237 kPa (1500 and 2500 psi). For miscible natural gas and nitrogen processes, the pressures must be greater than 27,580 kPa (4000 psi). If a high-pressure gas process is to be applied to a reservoir, the reservoir must be deep enough to contain the pressure without fracturing and leaking to other formations or the surface. Usually, a reservoir must be at least 610 m (2000 ft) deep for every 6895 kPa (1000 psi) of operating pressure. High-pressure carbon dioxide processes have been applied to many oil reservoirs in western Texas and southeastern New Mexico. The carbon dioxide for these reservoirs is transported in pipelines from northern New Mexico and southern Colorado, where nearly pure carbon dioxide is found in deep formations.

Chemical processes include polymer, micellar-polymer, caustic or alkaline, and microbial, as well as combinations of these processes. The mobility of water injected in a secondary process can be reduced by adding a suitable polymer, usually polyacrylamide or a xanthan gum. With decreased mobility, the sweep efficiency of a waterflood improves. Polymer injection, however, does not increase the ultimate recovery (does not reduce the residual oil saturation) but increases the rate of recovery. With surfactants, very low interfacial tension between oil and water can be produced. A wide variety of surfactants have been used, including surfactants produced by sulfonation of crude oils. In caustic processes, surfactants are produced in situ by injecting alkaline agents, dissolved in water, into a reservoir. The alkaline agents react with portions of the oil in the reservoir to produce surfactants. A major obstacle to success of polymer and surfactant processes is loss of the chemicals by precipitation and adsorption onto mineral surfaces. Precipitation is caused by reaction of the chemicals with divalent ions, such as magnesium and calcium. High-surface-area clays are the main actor in adsorption. Chemical processes are also limited to formations with moderate to high permeabilities. While many chemical processes were tested in the United States during the early 1980s, there are few in operation today. *See* CLAY MINERALS; PETROLEUM MICROBIOLOGY; POLYMER; SURFACE AND INTERFACIAL CHEMISTRY; SURFACTANT.

Future developments. With the decline in oil prices in 1985 and 1986, interest and research efforts in enhanced oil recovery technologies waned. Many enhanced recovery operations were terminated as prices remained low. The surviving operations in the United States are dominated by steam injection (in California) followed by carbon dioxide injection in western Texas.

While continued instability in oil prices in the 1990s suppressed technological development, there was limited effort on technologies that are less costly to implement. For example, combinations of alkaline, surfactant, and polymer processes have been implemented in a few reservoirs in recent years. These processes rely on the capacity of some oils to produce useful surfactants when reacted with alkaline agents. The cost of alkaline agents is low. Microbial processes provide another example of continuing field testing. Microbial processes may potentially generate useful low-cost chemicals (surfactants and polymers) and gases (such as carbon dioxide) from bacterial colonies growing in a reservoir formation.

Recent developments in horizontal and multilateral drilling technologies have provided greater access to reservoirs, and they may stimulate some renewed interest in enhanced recovery technology. Increased access to reservoirs by large-scale tunneling into the formations has been proposed and was tested sporadically over the last 30 years. Very high well densities are possible by short-reach drilling from tunnels. Whether such technology will ever gain wider application is hard to project. An extreme example of reservoir access is found in the Athabasca Tar Sands near Fort McMurray in northern Alberta, Canada. There, the near-surface tar sands are transported by large shovels and trucks to a central processing facility which separates the tar from the sand with a thermal process and converts the tar to coke, sulfur, and a high-quality crude oil. Large increases in oil production from these tar sands are planned. *See* OIL SAND.

Richard L. Christiansen

Bibliography. *Enhanced Oil Recovery*, National Petroleum Council, U.S. Department of Energy, 1984; D. W. Green and G. P. Willhite, *Enhanced Oil Recovery*, Textbook Series, Society of Petroleum Engineers, 1998; *Improved Oil Recovery*, Interstate Oil Compact Commission, Oklahoma City, 1983; L. W. Lake, *Enhanced Oil Recovery*, Prentice Hall, Englewood Cliffs, NJ, 1989.

Petroleum geology

The practice of utilizing geological principles and applying geological concepts to the discovery and recovery of petroleum. Related fields in petroleum discovery include geochemistry and geophysics. The related areas in petroleum recovery are petroleum and chemical engineering. *See* CHEMICAL ENGINEERING; GEOCHEMISTRY; GEOPHYSICS.

Occurrence of petroleum. Petroleum occurs in a liquid phase as crude oil and condensate, and in a gaseous phase as natural gas. The phase is dependent on the kind of source rock from which the petroleum was formed and the physical and thermal environment in which it exists. As a liquid, oil may be readily transported from producing fields to points of consumption, thus moving in a global market. Natural gas is moved chiefly by pipelines and is therefore tied to land-locked markets. *See* NATURAL GAS; PETROLEUM.

Most petroleum occurs at varying depths below the ground surface, but generally petroleum existing as a liquid (crude oil) is found at depths of less than 20,000 ft (6100 m) while natural gas is found both at shallow depths and at depths exceeding 30,000 ft (9200 m). In some cases, oil may seep to the surface, forming massive deposits of oil or tar sands, such as the Athabasca oil sands of Alberta, Canada, and the tar sands of the Faja de Orinoco in Venezuela. Natural gas also seeps to the surface but escapes into the atmosphere, leaving little or no surface trace. *See* OIL SAND.

Most petroleum is found in sedimentary basins in sedimentary rocks, although many of the 700 or so sedimentary basins of the world contain no known significant accumulations. Although oil or gas accumulations are commonly known as pools, oil and gas occur in relatively small voids in rocks, such as pores and fractures, and not as underground pools or streams. *See* BASIN; SEDIMENTARY ROCKS.

Several conditions must exist for the accumulation of petroleum: (1) There must be a source rock, usually high in organic matter, from which petroleum can be generated. (2) There must be a mechanism for the petroleum to move, or migrate. (3) A reservoir rock with voids to hold petroleum fluids must exist. (4) The reservoir must be in a configuration to constitute a trap and be covered by a seal—any kind of low-permeability or dense rock formation that prevents further migration. If any of these conditions do not exist, petroleum either will not form or will not accumulate in commercially extractable form.

Petroleum source. Source rocks for most petroleum generation are those containing high concentrations of animal and plant organic matter. To yield high concentrations of total organic carbon, accumulation of the animal or plant remains should be in a chemical reducing environment so that the carbon is not oxidized. Accumulated organic matter eventually must be buried at depths sufficient to have temperatures of at least 140°F (60°C). On average, temperature increases with depth at a rate of about 18°F/1000 ft (10°C/300 m), but such a geothermal gradient exhibits wide variations. Organic matter exposed to sufficiently high temperatures, over time, changes or matures through stages to a liquid or gaseous hydrocarbon. The process of converting organic matter to petroleum is essentially a cooking process. *See* ORGANIC GEOCHEMISTRY.

Petroleum migration. Rocks that form source beds are commonly fine-grained and low in permeability, and thus make poor reservoirs. In some cases, however, the generated petroleum hydrocarbons do not leave the source rock or place of formation. Excellent examples are the so-called oil shales of the western United States, where vast quantities of petroleum hydrocarbons occur but extraction and recovery are expensive. Generally, migration from the source rock to a porous and permeable reservoir is necessary for commercial extraction. Such migration is initiated as buried source rocks are compacted and fluids are expelled and moved laterally and vertically to reservoirs, but the precise mechanisms for primary migration from source rocks to reservoirs are subjects of continuing debate. *See* OIL SHALE.

Petroleum trapping and reservoir. Migration ends once petroleum liquids and gases are trapped in a reservoir, or if they are not trapped, when they escape to the surface. The reservoir must have sufficient porosity and permeability, either with pores or fractures, to accumulate fluids. Those voids are generally filled initially with water, so that emplacement of petroleum must involve the displacement of some or all of the water. Most reservoir rocks are either sandstone or carbonate; these rocks show wide ranges of porosity and permeability, and they vary significantly in reservoir quality, as demonstrated by the efficiency in extraction. In fact, actual recovery of oil from a reservoir varies from as little as 5% of total original volume to as much as 95%, depending largely on the quality of the reservoir.

A trap is any arrangement of strata that allows the accumulation of oil and gas and precludes further migration. A variety of geologic mechanisms exist. Strata may be arranged by folding to give an anticline or a convex-upward trap. Areal changes in permeability of a reservoir from porous to impermeable will stop migration and cause a trapping of oil and gas. A reservoir may terminate in any direction due to subsequent erosion or to the original process by which it formed, such as a coral reef; such terminations are pinch-out traps. A fault or a vertical displacement of strata may result in permeable strata abutting impermeable strata, creating a fault trap. Or traps may be formed by vertical movement of salt or shale to form diapir traps.

Reservoirs must be covered by a seal of impervious strata, such as salt or shale, to prevent further migration and to effect petroleum accumulation. With changing geologic conditions over time, the reservoir may be breached or modified, causing the trapped oil and gas to escape. This movement, known as secondary migration, may be to the ground surface or to another reservoir and trap. *See* DIAPIR; LIMESTONE; SALT DOME; SANDSTONE; SHALE.

Prospecting. The aim of petroleum geologists is to find traps or accumulations of petroleum. The trap

not only must be defined but must exist where other conditions such as source and reservoir rocks occur. Most structures or rock configurations that could contain and hold petroleum in fact do not. Thus, even after applying the best scientific information, a well that is drilled to an accurately defined trap may not encounter hydrocarbons.

A variety of techniques are used by the petroleum geologist to reconstruct geologic events and define an area or a prospect to be drilled. Geologic structures shown at the surface, such as folds or unconformities, can be projected into the subsurface. Accumulation of hydrocarbons at the ground surface may occur as seeps. In the early days of oil and gas exploration, reliance was placed on such surface manifestations; and in remote, unexplored areas of the world, surface conditions are yet a clue to the existence of petroleum at depth.

Most of the world's oil and gas accumulations show no surface manifestation. To locate these traps, the geologist must rely on subsurface information and data gathered by drilling exploratory wells and data obtained by geophysical surveying. These data, once interpreted, are used to construct maps, cross sections, and models that are used to infer or to actually depict subsurface configurations that might contain petroleum. Such depictions are prospects for drilling. If, on drilling, the trap is found to be as reconstructed by the petroleum geologist and if all other conditions for oil and gas accumulation exist, a discovery is made. If the prospect does not exist as envisaged, the well encounters no oil and is recorded as a dry hole. On average, less than one exploratory well in ten contains commercial quantities of oil or gas. Exploration efficiency has improved remarkably in recent years with advances in the technology of seismic reflection acquisition and processing. Running seismic surveys in a closely spaced grid allows three-dimensional imaging of the Earth's subsurface. In certain areas, especially offshore where seismic is shot through a uniform medium of seawater, character of reflections commonly allows direct detection of hydrocarbon accumulations. Accordingly, in areas amenable to advanced seismic technology, success rates for oil and gas discovery are 35% or higher. *See* GEOPHYSICAL EXPLORATION; OIL AND GAS WELL DRILLING.

Economic deposits. Oil and gas must be trapped in an individual reservoir in sufficient quantities to be commercially producible. That quantity is determined by the price or value of oil or gas and the cost to find it. In the United States, extensive drilling since around 1890 has led to the discovery of nearly 30,000 oil and gas fields. These fields range in size from some with less than 100,000 barrels (16,000 m^3) of oil to the largest field so far discovered in the United States, Prudhoe Bay in Alaska, which contained more than 10^{10} bbl (1.6×10^8 m^3) of recoverable oil. Worldwide, 25% of all oil discovered so far is contained in only ten fields, seven of which are in the Middle East. Fifty percent of all oil discovered to date is found in only 50 fields.

Worldwide, about 4×10^6 oil and gas wells have been drilled. About 85% percent of the wells have been drilled in the United States. As a result, most of the large and fairly obvious fields in the United States have been discovered, except those possibly existing in frontier or lightly explored areas such as Alaska and the deep waters offshore. Few areas of the world remain entirely untested, but many areas outside the United States are only partly explored, and advanced techniques have yet to be deployed in the recovery of oil and gas found so far. *See* PETROLEUM RESERVES.

Although in the United States the most obvious and most readily detectable prospective traps have been tested, many traps remain that are subtle and not easily definable. It is known that petroleum is accessible in the subtle traps because many have been found, commonly by accident. This occurs as a particular prospect is being tested and drilling encounters an unexpected accumulation. The challenge to the petroleum geologist is to develop models and to use ever-advancing technology to improve the ability to detect these subtle structures.

In addition to exploration for difficult-to-detect traps, greater efforts in petroleum geology along with petroleum engineering are being made to increase recovery from existing fields. Of all oil discovered so far, it is estimated that there will be recovery of only 35% on the average. In the United States alone, the amount of oil known in existing reservoirs and classed as unrecoverable is more than 3.25×10^{11} bbl (5.2×10^{10} m^3), twice the volume produced to date and 50 times the amount of oil the United States uses every year. Recovering some part of this huge oil resource will require geological reconstruction of reservoirs, a kind of very detailed and small-scale exploration. These reconstructions and models have allowed additional recovery of oil that is naturally movable in the reservoir. If the remaining oil is immobile because it is too viscous or because it is locked in very small pores or is held by capillary forces, techniques must be used by the petroleum geologist and the petroleum engineer to render the oil movable. If oil is too viscous or heavy to flow, steam can be injected into the reservoir to raise the temperature and thus lower the oil viscosity. If oil is locked in small pores, gas can be injected to expand the fluid and cause it to escape and move.

William L. Fisher

Bibliography. J. Gluyas and R. Swarbrick, *Petroleum Geoscience*, 2003; R. C. Laudon, *Principles of Petroleum Development Geology*, 1995; A. I. Levorsen, *Geology of Petroleum*, 2d ed., 1967, reprint 2001; L. B. Magoon and W. G. Dow (eds.), *The Petroleum System—from Source to Trap*, 1994; R. C. Selley, *Elements of Petroleum Geology*, 2d ed., 1997; B. P. Tissot and D. H. Welte, *Petroleum Formation and Occurrence*, rev. ed., 1992.

Petroleum microbiology

Those aspects of microbiology in which crude oil, refined petroleum products, or pure hydrocarbons serve as nutrients for the growth of microorganisms or are altered as a result of their activities.

Applications of petroleum microbiology include oil pollution control, enhanced oil recovery, microbial contamination of petroleum fuels and oil emulsions, and conversion of petroleum hydrocarbons into microbial products.

Biodegradation of petroleum hydrocarbons. Many species of bacteria, fungi, and algae have the enzymatic capability to use petroleum hydrocarbons as food. Biodegradation of petroleum requires an appropriate mixture of microorganisms, contact with oxygen gas, and large quantities of utilizable nitrogen and phosphorus compounds and smaller amounts of other elements essential for the growth of all microorganisms. Part of the hydrocarbons are converted into carbon dioxide and water and part into cellular materials, such as proteins and nucleic acids. The requirement for a mixture of different microorganisms arises from the fact that petroleum is composed of a wide variety of different groups of hydrocarbons, whereas any specific microorganism is highly specialized with regard to the type of hydrocarbon it can digest. For example, different bacteria are involved in the degradation of aliphatic cyclic and aromatic hydrocarbons. The bacterial genera that contain the most frequently isolated hydrocarbon degraders are *Pseudomonas*, *Acinetobacter*, *Flavobacterium*, *Brevibacterium*, *Corynebacterium*, *Arthrobacter*, *Mycobacterium*, and *Nocardia*. The fungal genera that contain oil utilizers include *Candida*, *Cladosporium*, *Rhodotorula*, *Torulopsis*, and *Trichosporium*.

By the use of genetic engineering techniques, a *Pseudomonas* species has been constructed which has the enzymatic capability of degrading several different groups of hydrocarbons. Although practical applications have not yet been found for this particular strain, it has provided a precedent for patenting genetically constructed microorganisms. *See* BIODEGRADATION; GENETIC ENGINEERING.

Clean-up of oil spills. Oil pollution results from natural hydrocarbon seeps, accidental spills, and intentional discharge of oily materials into the environment. Once the oil is released and comes into contact with water, air, and the necessary salts, microorganisms present in the environment begin the natural process of petroleum biodegradation. If this process did not occur, the world's oceans would soon become completely covered with a layer of oil. The reason that oil spills become a pollution problem is that the natural microbial systems for degrading the oil become temporarily overwhelmed.

Various methods have been tested for increasing the rates of petroleum biodegradation and thereby stimulating microbial clean-up of oil spills, the most successful of which involve modification of environmental parameters, such as addition of nitrogen and phosphorus fertilizers. Since most environments have an indigenous population of oil-degrading microorganisms that can multiply rapidly under appropriate conditions, additions of "seed" bacteria have little value in stimulating clean-up of oil spills. *See* WATER POLLUTION.

Enhanced oil recovery. The largest potential application of petroleum microbiology is in the field of enhanced oil recovery. Microbial products, as well as viable microorganisms, have been used as stimulation agents to enhance oil recovery from petroleum reservoirs. Xanthan, a polysaccharide produced by *Xanthomonas campestris*, is used as a waterflood thickening agent in oil recovery. Emulsan, a lipopolysaccharide produced by a strain of *Acinetobacter calcoaceticus*, stabilizes oil-in-water emulsions. A number of other microbial products are being tested for potential application in enhanced oil recovery processes, such as polymer flooding, micellular flooding, wetting of reservoir rock, reduction of oil-water interfacial tensions, and release of hydrocarbons from tar sands and oil shale. Field tests have indicated that injection of viable microorganisms with their nutrients into petroleum reservoirs can lead to enhanced oil recovery, presumably due to production of carbon dioxide gas, acids, and surfactants. *See* PETROLEUM ENHANCED RECOVERY.

Contamination of hydrocarbons. The ability of microorganisms to utilize petroleum also has its detrimental aspects, particularly with respect to the deterioration of petroleum fuels, asphalt coatings, and oil emulsions used with cutting machinery. All hydrocarbons become contaminated if they come into contact with water during storage. The problem is most serious with kerosine-based jet aircraft fuels which are consumed at very high rates by jet engines; thus, even a small amount of microbial sludge in the fuel becomes a hazard.

A significant economic problem is microbial degradation of asphalt pipeline coatings leading to metal corrosion. Petroleum-in-water emulsions are widely used as coolants and lubricants for high-speed metal-cutting machines. These emulsions rapidly become contaminated with microorganisms, resulting in breakage of the emulsion and potentially causing dermatitis to the workers. Antimicrobials that concentrate at oil-water interfaces are most effective at reducing the rate of microbial contamination of hydrocarbons. *See* CORROSION.

Conversion of petroleum hydrocarbons. A variety of valuable materials, such as amino acids, carbohydrates, nucleotides, vitamins, enzymes, antibiotics, citric acid, long-chain dicarboxylic acids, and biomass can be produced by microbial processes using petroleum hydrocarbons as substrates. Most of these materials also can be produced from more traditional microbial feedstocks, such as molasses. The main advantage of using hydrocarbons as substrates is their lower cost. Also, certain products, such as tetradecane-1,14-dicarboxylic acid, a raw material for preparing perfumes, are synthesized in higher yields on hydrocarbon than on carbohydrate substrates.

The most active area of research and development in petroleum microbiology since the mid-1960s has been in the large-scale production and concentration of microorganisms for animal feed and human food. Dried microbial cells are collectively referred to as single-cell protein. The advantages of single-cell protein over the more traditional sources of

protein include the following: (1) Microorganisms grow rapidly. For example, yeast growing on hydrocarbons will double their mass every 4 h, leading to more than a hundredfold increase in yeast protein in 30 h. (2) Microorganisms have a high protein content. Depending on the particular type of microorganism and the growth conditions, the protein content of microbes varies from 30 to 70%. By comparison, legumes, wheat, and rice contain 25, 12, and 8% protein, respectively. (3) Microorganisms can be easily modified genetically to produce favorable traits. It takes many years to alter animals and plants genetically. Most strains that are used today for food have undergone continuous strain improvement for centuries. It should be possible in a relatively short time to improve microbial food strains by mutation and selection as well as by genetic engineering. (4) Microorganisms can be cultivated on a large scale anywhere in the world, independent of soil or climatic conditions. A medium-sized single-cell protein plant, occupying an area of 3 acres (1.2 hectares), produces the same amount of protein as 300,000 acres (120,000 hectares) devoted to soybeans.

In spite of the above advantages, single-cell protein has not yet played a significant role in providing protein for animal feed or human consumption. Manufacturing costs are still quite high when compared with the major competing protein source, soy, which is produced with little fertilizer and requires only a limited amount of processing to produce the soy meal. There is no way to predict when all of the conditions necessary to make single-cell protein a major force in the food and feed markets will occur, but many scientists are optimistic about its potential. *See* ANIMAL FEEDS; BACTERIAL PHYSIOLOGY AND METABOLISM; INDUSTRIAL MICROBIOLOGY.

Eugene Rosenberg

Bibliography. A. L. Demain, J.E. Davis, and R. M. Atlas (eds.), *Manual of Industrial Microbiology and Biotechnology*, 2d ed., 1999; I. Goldberg, *Single Cell Protein*, 1986; D. L. Gutnick and E. Rosenberg, Oil tankers and pollution: A microbiological approach, *Annu. Rev. Microbiol.*, 31:379–396, 1977; N. Kosaric (ed.), *Biosurfactants*, 1993; J. G. Leahy and R. R. Colwell, Microbial degradation of hydrocarbons in the environment, *Microbiol. Rev.*, 54:305–315, 1990; J. H. Litchfield, Single-cell protein, *Science*, 219:740–746, 1983; E. Rosenberg, Exploiting microbial growth on hydrocarbons: New markets, *Trends Biotech.*, 11:419–423, 1993.

Petroleum processing and refining

The separation of petroleum into fractions and the treating of these fractions to yield marketable products. Petroleum is a mixture of gaseous, liquid, and solid hydrocarbon compounds that occurs in sedimentary rock deposits throughout the world. In the crude state, petroleum has little value but, when refined, it provides liquid fuels (gasoline, diesel fuel, aviation fuel), solvents, heating oil, lubricants, and the distillation residuum asphalt, which is used for highway surfaces and roofing materials. *See* PETROLEUM; PETROLEUM PRODUCTS.

Crude petroleum (oil) is a mixture of compounds with different boiling temperatures that can be separated into a variety of fractions (**Table 1**). Since there is a wide variation in the composition of crude petroleum, the proportions in which the different fractions occur vary with origin. Some crude oils have higher proportions of lower-boiling components, while others have higher proportions of residuum (asphaltic components).

Petroleum processing and refining involves a series of steps by which the original crude oil is converted into products with desired qualities in the amounts dictated by the market. In fact, a refinery is essentially a group of manufacturing plants that vary in number with the variety of products in the mix. Refinery processes must be selected and products manufactured to give a balanced operation; that is, crude oil must be converted into products according

TABLE 1. Petroleum fractions and their uses*

Fraction	Boiling range °C	Boiling range °F	Uses
Fuel gas	−160 to −40	−260 to −40	Refinery fuel
Propane	−40	−40	Liquefied petroleum gas (LPG)
Butane(s)	−12 to −1	11–30	Increases volatility of gasoline, advantageous in cold climates
Light naphtha	−1 to 150	30–300	Gasoline components, may be (with heavy naphtha) reformer feedstock
Heavy naphtha	150–205	300–400	Reformer feedstock; with light gas oil, jet fuels
Gasoline	−1 to 180	30–355	Motor fuel
Kerosine	205–260	400–500	Fuel oil
Stove oil	205–290	400–550	Fuel oil
Light gas oil	260–315	500–600	Furnace and diesel fuel components
Heavy gas oil	315–425	600–800	Feedstock for catalytic cracker
Lubricating oil	>400	>750	Lubrication
Vacuum gas oil	425–600	800–1100	Feedstock for catalytic cracker
Residuum	>600	>1100	Heavy fuel oil, asphalts

* From J. G. Speight (ed.), *The Chemistry and Technology of Petroleum*, 3d ed., Marcel Dekker, New York, 1999.

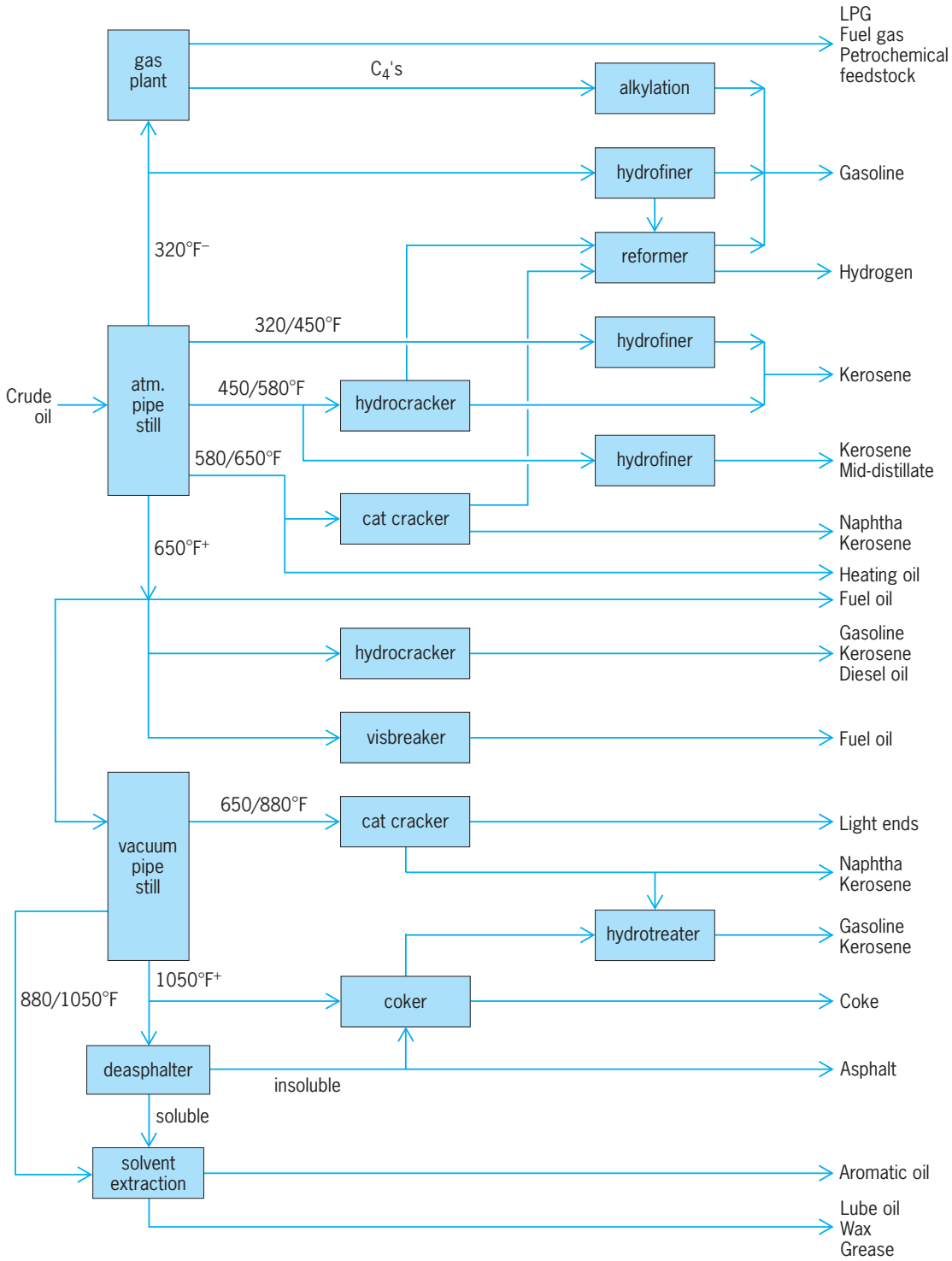

Fig. 1. Schematic of a petroleum refinery showing the various units. (*After J. G. Speight, ed., The Chemistry and Technology of Petroleum, 3d ed., Marcel Dekker, New York, 1999*)

to the demand for each. For example, the manufacture of products from the lower-boiling portion of petroleum automatically produces a certain amount of higher-boiling components. If the latter cannot be sold as, say, heavy fuel oil, these products will accumulate until refinery storage facilities are full. To prevent such a situation, the refinery must be flexible and able to change operations as needed. This usually means more processes, such as thermal processes to change excess heavy fuel oil into gasoline with coke as the residual product, or vacuum distillation processes to separate heavy oil into lubricating oil stocks and asphalt.

Early refineries (70 years ago) were predominantly distillation units with, perhaps, ancillary units to remove objectionable odors from the various product

TABLE 2. Desalting and dewatering operations

Separation method	Temperature, °F (°C)	Type of treatment
Chemical	140–210 (60–99)	0.05–4% solution of soap in water 0.5–5% solution of soda ash in water
Electrical	150–200 (66–93)	10,000–20,000 V
Gravity	180–200 (82–93)	Up to 40% water added
Centrifugal	180–200 (82–93)	Up to 20% water added (sometimes no water added)

streams. The refinery of today (**Fig. 1**), the result of major evolutionary trends, is a highly complex operation. Most of the evolutionary adjustments to refineries have occurred since 1940. In the petroleum industry, as in many other industries, supply and demand are key factors in efficient and economic operation. Innovation is also key.

Distillation. Distillation, the first method by which petroleum was refined, remains a major process. The original technique involved a batch operation in which the still, a cast-iron vessel, was mounted on brickwork over a fire and the volatile materials passed through a pipe or gooseneck at the top of the still to a condenser, a coil of pipe (worm) immersed in a tank of running water.

In a modern refinery, after dewatering and desalting operations (**Table 2**), distillation remains the prime means by which petroleum is refined. The distillation section of a modern refinery (**Fig. 2**) is the most flexible unit in the refinery since conditions can be adjusted to process a wide range of feedstocks from the lighter crude oils to the heavier, more viscous crude oils. The maximum preferential temperature (in the vaporizing furnace or heater) to which the feedstock can be subjected is 350°C (660°F), unless the goal is to produce a thermally decomposed residuum. Thermal decomposition (cracking) occurs above this temperature, which can lead to coke deposition in the heater pipes or in the tower itself, resulting in failure of the distillation unit. Higher temperatures (up to 393°C; 740°F) are used in distillation units, but the petroleum residence time must be adjusted to prevent cracking.

Atmospheric distillation. In a present-day petroleum distillation unit (actually a collection of distillation units), a tower is used for fractionation (separation). The feedstock to the tower (pipe still) flows through one or more pipes arranged within a large furnace (pipe-still heater or pipe-still furnace) where the feed is heated to a temperature at which a predetermined portion of the feed changes into vapor. The heated feed is introduced into a fractional distillation tower where the nonvolatiles or liquid portions pass downward to the bottom of the tower and are pumped away, while the vapors pass upward through the tower and are fractionated into gas oils, kerosine, and naphthas.

Pipe-still furnaces vary greatly and can accommodate 50,000 (US) barrels (bbl; 2,100,000 US gallons) or more of crude petroleum per day. The walls and ceiling are insulated with firebrick, and the interior of the furnace is divided into two sections: a smaller convection section where the oil first enters the furnace, and a larger section (fitted with heaters) where the oil reaches its highest temperature.

Heat exchangers, used to preheat the feed, are made of bundles of tubes arranged within a shell so that the unheated feedstock passes through a series of tubes opposite the direction of the heated feedstock also passing through the shell. In this way, hot products from the distillation tower are cooled, passing the heat to the incoming cold crude oil. *See* HEAT EXCHANGER.

The primary fractions from a distillation unit contain some proportion of the lighter constituents characteristic of a lower-boiling fraction. These lower-boiling materials are removed by stripping (stabilization) before storage or further processing.

Vacuum distillation. Vacuum distillation is used in petroleum refining to separate the less volatile products, such as lubricating oils, from petroleum without subjecting the high-boiling products to cracking conditions. When the feedstock (reduced crude, atmospheric residuum) is required for the manufacture of labricating oils, further fractionation without cracking is desirable, which can be achieved by distillation under vacuum conditions.

Operating pressure for vacuum distillation is usually 50–100 mm of mercury (6.7–13.3 kilopascals) [atmospheric pressure = 760 mm of mercury] and, in order to minimize large fluctuations in pressure

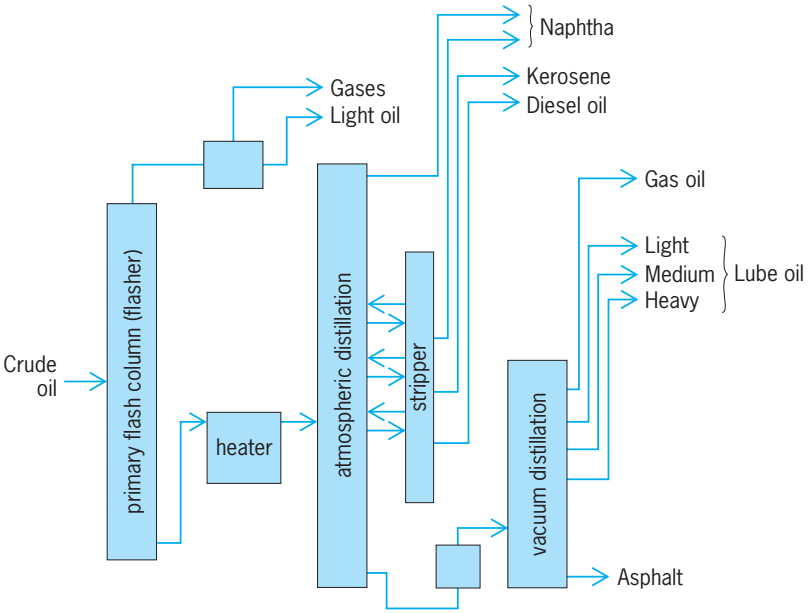

Fig. 2. Distillation section of a refinery. (*After J. G. Speight, ed., The Chemistry and Technology of Petroleum, 3d ed., Marcel Dekker, New York, 1999*)

in the vacuum tower, the units are necessarily larger in diameter than the atmospheric units. Some vacuum distillation units have diameters on the order of 45 ft (14 m). By this means, a heavy gas oil that has a boiling range in excess of 315°C (600°F) at atmospheric pressure may be obtained at temperatures of around 150°C (300°F), and lubricating oil, having a boiling range in excess of 370°C (700°F) at atmospheric pressure may be obtained at temperatures of 250–350°C (480–660°F). The feedstock and product temperatures are kept below the temperature and residence time where cracking will occur. The partial pressure of the feedstock constituents is reduced further by steam injection. The steam added to the column, principally for the stripping of asphalt in the base of the column, is superheated in the convection section of the heater.

The fractions obtained by vacuum distillation of the reduced crude (atmospheric residuum) from an atmospheric distillation unit depend on whether or not the unit is designed to produce lubricating or vacuum gas oils. In the former case, the fractions include (1) heavy gas oil, an overhead (volatile) product used as catalytic cracking stock or a light lubricating oil; (2) lubricating oil (light, intermediate, and heavy); and (3) asphalt (or residuum), the bottom product, which is used directly as, or to produce, asphalt or is blended with gas oil to produce a heavy fuel oil. Atmospheric and vacuum distillation are major parts of refinery operations, and no doubt will continue to be used as the primary refining operation.

Other distillation processes. As distillation techniques in refineries became more sophisticated, they were able to process a wider variety of crude oils to produce marketable products or feedstocks for other refinery units. However, it became apparent that the distillation units in the refineries were incapable of producing specific product fractions. In order to accommodate this type of product demand, refineries, beginning about midcentury, incorporated azeotropic distillation and extractive distillation in their operations.

All compounds have definite boiling temperatures, but a mixture of chemically dissimilar compounds will sometimes cause one or both of the components to boil at a temperature other than that expected. A mixture that boils at a temperature lower than the boiling point of any of the components is called an azeotropic mixture. When it is desired to separate close-boiling components, the addition of a nonindigenous component will form an azeotropic mixture with one of the components of the mixture, thereby lowering the boiling point and facilitating separation by distillation.

The separation of components of similar volatility may become economic if an azeotropic entrainer can be found that effectively changes their relative volatility. It is also desirable that the entrainer be reasonably cheap, stable, nontoxic, and readily recoverable from the components. In practice it is probably recoverability that limits the application of extractive and azeotropic distillation. The majority of successful processes are those in which the entrainer and one of the components separate into two liquid phases on cooling if direct recovery by distillation is not feasible. A further restriction in the selection of an azeotropic entrainer is that the boiling point of the entrainer be in the range of 10–40°C (50–100°F) below that of the components. *See* AZEOTROPIC DISTILLATION; AZEOTROPIC MIXTURE; DISTILLATION.

Thermal processes. Cracking is used to thermally decompose higher-boiling petroleum constituents and products to lighter oils. The first units were installed in refineries in 1913 as a means of producing more gasoline. In the early (pre-1940) processes used for gasoline manufacture, the major variables involved to achieve the cracking of the feedstock to lighter products with minimal coke formation were feedstock type, time, temperature, and pressure.

As refining technology evolved, the residuum or heavy distillate from a distillation unit became the feedstock for the cracking process. The residua produced as the end products of distillation processes, and even some of the heavier virgin oils, often contain substantial amounts of asphaltic materials, which preclude their use as fuel oils or lubricating stocks. However, subjecting these residua directly to thermal processes has become economically advantageous, since the end result is the production of lower-boiling materials. The asphaltic materials in the residua are regarded as the unwanted coke-forming constituents.

As the thermal processes evolved and catalysts were used with more frequency, poisoning of the catalyst (and this reduction in the lifetime of the catalyst) become a major issue for refiners. To avoid catalyst poisoning, it became necessary to remove from the feedstock as much of the nitrogen and metals (such as vanadium and nickel) as possible. Most of the nitrogen, oxygen, and sulfur as well as the metals are contained in, or associated with, the asphaltic fraction (residuum).

A number of thermal processes, such as residuum separation (flash distillation), vacuum flashing, visbreaking, and coking, became widely used by refiners to upgrade feedstocks by removing the asphaltic fraction.

Thermal cracking. One of the earliest conversion processes used in the petroleum industry was the thermal decomposition of higher-boiling materials into lower-boiling products. This process is known as thermal cracking.

The majority of the thermal cracking processes use temperatures of 455–540°C (850–1005°F) and pressures of 100–1000 psi (690–6895 kPa). For example, the feedstock (reduced crude) is preheated by direct exchange with the cracking products in the fractionating columns. Cracked gasoline and heating oil are removed from the upper section of the column. Light and heavy distillate fractions are removed from the lower section and are pumped to separate heaters. Higher temperatures are used to crack the more stable light distillate fraction. The streams from the heaters are combined and sent to a soaking chamber where additional time is provided to complete the cracking reactions. The cracked products

are then separated in a low-pressure flash chamber where a heavy fuel oil is removed as bottoms. The remaining cracked products are sent to fractionating columns. *See* DISTILLATION COLUMN.

The thermal cracking of higher-boiling petroleum fractions to produce gasoline is now virtually obsolete. The antiknock requirements of modern automobile engines together with the different nature of crude oils (compared to those of 50 years ago) has reduced the ability of the thermal cracking process to produce gasoline on an economic basis. Very few new units have been installed since the 1960s, and some refineries may still operate the older thermal cracking units.

Visbreaking. Visbreaking (viscosity breaking) is a mild thermal cracking operation that can be used to reduce the viscosity of residua to allow the products to meet fuel oil specifications. Alternatively, the visbroken residua can be blended with lighter product oils to produce fuel oils of acceptable viscosity. By reducing the viscosity of the residuum, visbreaking reduces the amount of light heating oil that is required for blending to meet fuel oil specifications. In addition to the major product, fuel oil, material in the gas oil and gasoline boiling range is produced. The gas oil may be used as additional feed for catalytic cracking units or as heating oil.

In a typical visbreaking operation (**Fig. 3**), a crude oil residuum is passed through a furnace where it is heated to a temperature of 480°C (895°F) under an outlet pressure of about 100 psi (689 kPa). The heating coils in the furnace are arranged to provide a soaking section of low heat density, where the crude oil residuum remains until the visbreaking reactions are completed and the cracked products are then passed into a flash-distillation chamber. The overhead material from this chamber is then fractionated to produce a low-quality gasoline as an overhead product and light gas oil as bottom. The liquid products from the flash chamber are cooled with a gas oil flux and then sent to a vacuum fractionator. This yields a heavy gas oil distillate and a residual tar of reduced viscosity. Visbreakers come as soaker type (Fig. 3) or coil type. In the coil type, the heating and thermal changes occur in the heating coil.

Delayed coking. Delayed coking is a thermal process for continuously converting (atmospheric or vacuum) residua into lower-boiling products, such as gases, naphtha, fuel oil, gas oil, and coke. Gas oil, often the major product of a coking operation, serves primarily as a feedstock for catalytic cracking units. The coke obtained is typically used as fuel; but specialty uses, such as electrode manufacture, and production of chemicals and metallurgical coke are also possible, increasing the value of the coke. For these uses, the coke may require treatment to remove sulfur and metal impurities. *See* COKE; FUEL OIL; NAPHTHA.

Delayed coking is a semicontinuous process (**Fig. 4**) in which the heated charge is transferred to large soaking (or coking) drums, which provide the long residence time needed to allow the cracking reactions to proceed to completion. The feed to

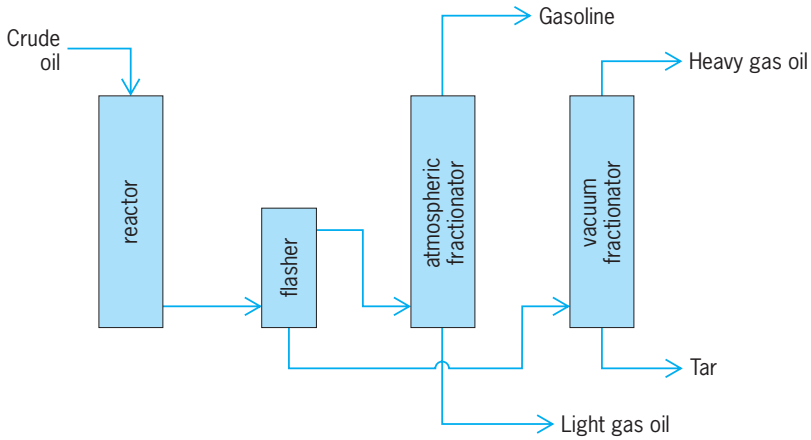

Fig. 3. Visbreaking process. (*After J. G. Speight, ed., The Chemistry and Technology of Petroleum, 3d ed., Marcel Dekker, New York, 1999*)

these units is normally an atmospheric residuum, although cracked residua are also used. The feedstock is introduced into a product fractionator where it is heated and the lighter fractions are removed as a side streams. The fractionator residuum, including a recycle stream of heavy product, is then heated in a furnace whose outlet temperature varies from 480 to 515°C (895 to 960°F), and enters one of a pair of coking drums where the cracking reactions continue. The cracked products leave as distillate, and coke deposits form on the inner surface of the drum. The temperature in the coke drum ranges from 415 to 450°C (780 to 840°F), and pressures are from 15 to 90 psi (101 to 620 kPa). Overhead products go to the fractionator, where the naphtha and heating oil fractions are recovered. The nonvolatile material is combined with preheated fresh feed and returned to the coking drum. *See* CRACKING.

For continuous operation, two drums are used; while one is on stream, the other is being cleaned. A coke drum is usually on stream for about 24 hours before becoming filled with porous coke, which is

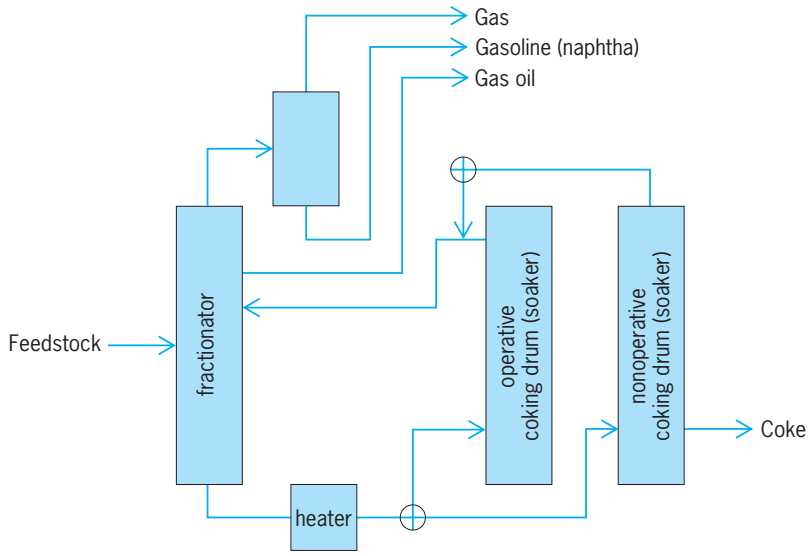

Fig. 4. Delayed coking process. (*After J. G. Speight, ed., The Chemistry and Technology of Petroleum, 3d ed., Marcel Dekker, New York, 1999*)

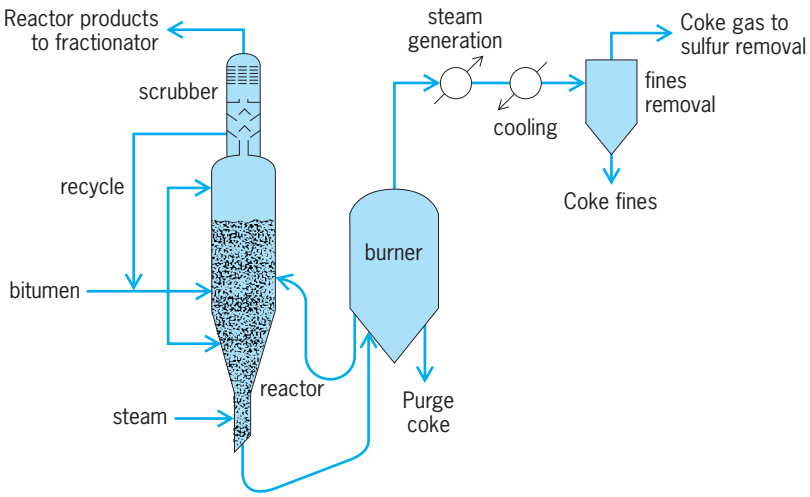

Fig. 5. Fluid coking pocess. (*After J. G. Speight, ed., The Chemistry and Technology of Petroleum, 3d ed., Marcel Dekker, New York, 1999*)

removed hydraulically. Normally, 24 hours are required to complete the cleaning operation and prepare the coke drum for use.

Fluid coking. Fluid coking (**Fig. 5**) is a continuous process that uses a fluidized bed to convert atmospheric and vacuum residua into more valuable products. The residuum is sprayed into a fluidized bed of hot, fine coke particles, which permits the coking reactions to be conducted at higher temperatures with shorter contact times than are possible in delayed coking. These conditions result in decreased coke yields and the recovery of greater quantities of more valuable liquid product.

Fluid coking uses two vessels, a reactor and a burner. Coke particles are circulated between them to transfer heat (generated by burning a portion of the coke) to the reactor. Steam is introduced at the bottom of the reactor to fluidize the coke particles. *See* FLUIDIZATION.

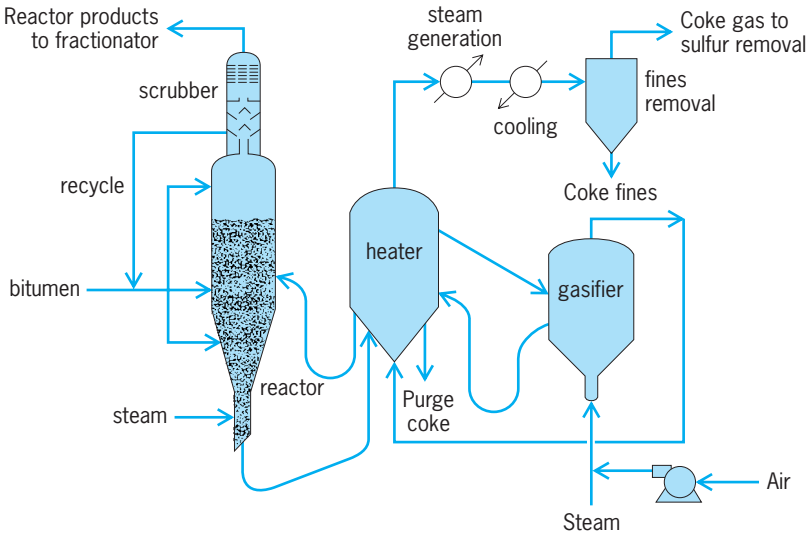

Fig. 6. Flexicoking process. (*After J. G. Speight, ed., The Chemistry and Technology of Petroleum, 3d ed., Marcel Dekker, New York, 1999*)

Flexicoking. Flexicoking is also a continuous process (**Fig. 6**) that is related to fluid coking. The unit uses the same configuration as the fluid coker but has a gasification section in which excess coke can be gasified to produce refinery fuel gas. *See* COKING (PETROLEUM).

Catalytic cracking. There are many processes in a refinery that use a catalyst to improve process efficiency (**Table 3**). The original incentive arose from the need to increase gasoline supplies in the 1930s and 1940s. Since cracking could virtually double the volume of gasoline from a barrel of crude oil, cracking was justifiable on this basis alone.

Catalytic cracking (**Fig. 7**) was introduced to refineries in 1936 as a method for producing higher-octane gasoline. This process is basically the same as thermal cracking, but differs by the use of a catalyst, which directs the course of the cracking reactions to produce more of the desired higher-octane hydrocarbon products. Catalytic cracking has a number of advantages over thermal cracking.

1. The gasoline produced has a higher octane number because of the presence of isoparaffins and aromatics, which also have greater chemical stability than monoolefins and diolefins.
2. Olefinic gases are suitable for polymer gasoline manufacture.
3. Smaller quantities of methane, ethane, and ethylene are produced by catalytic cracking.
4. Sulfur compounds are changed in such a way that the sulfur content in catalytically cracked gasoline is lower than in thermally cracked gasoline.

Catalytic cracking also produces less coke and more of the useful gas oils than does thermal cracking. *See* ALKENE; GASOLINE; OCTANE NUMBER.

Catalytic cracking is regarded as the modern method for converting high-boiling petroleum fractions, such as gas oil, into gasoline and other low-boiling fractions. The usual commercial process involves contacting a gas oil faction with an active catalyst at a suitable temperature, pressure, and residence time so that a substantial part (>50%) of the gas oil is converted into gasoline and lower-boiling products, usually in a single-pass operation. However, during the cracking reaction, carbonaceous material is deposited on the catalyst, which markedly reduces its activity. Removal of the deposit is usually accomplished by burning the catalyst in air.

The several processes currently used in catalytic cracking differ mainly in the method of catalyst handling, although there is overlap with regard to catalyst type and the nature of the products. The catalyst, which may be an activated natural or synthetic material, is employed in bead, pellet, or microspherical form and can be used as a fixed bed, moving bed, or fluid bed. Fixed-bed was the first process to be used commercially and involves a static bed of catalyst in several reactors, which allows a continuous flow of feedstock to be maintained. Thus, the cycle of operations consists of (1) flowing feedstock through the catalyst bed, (2) discontinuing the feedstock flow and removing the coke from the catalyst by

TABLE 3. Catalytic processes

Process	Materials charged	Products recovered	Temperature of reaction	Type of reaction
Cracking	Gas oil, fuel oil, heavy feedstocks	Gasoline, gas, and fuel oil	875–975°F 470–525°C	Dissociation or splitting of molecules
Hydrogenation	Gasoline to heavy feedstocks	Low-boiling products	400–850°F 205–455°C	Mild hydrogenation; cracking; removal of sulfur, nitrogen, oxygen, and metallic compounds
Reforming	Gasolines, naphthas	High-octane gasolines, aromatics	850–1000°F 455–535°C	Dehydrogenation, dehydro-isomerization, isomerization, hydrocracking, dehydrocyclization
Isomerization	Butane, C_4H_{10}	Isobutane, C_4H_{10}		Rearrangement
Alkylation	Butylene and isobutane, C_4H_8 and C_4H_{10}	Alkylate, C_8H_{18}	32–50°F 0–10°C	Combination
Polymerization	Butylene, C_4H_8	Octene, C_8H_{16}	300–350°F 150–175°C	Combination

* From J. G. Speight (ed.), *The Chemistry and Technology of Petroleum*, 3d ed., Marcel Dekker, New York, 1999.

burning, and (3) Putting the reactor back on stream. *See* CHEMICAL REACTOR.

The moving-bed process uses a reaction vessel (in which cracking takes place) and a kiln (in which the spent catalyst is regenerated), and catalyst movement between the two is provided by various means.

The fluid-bed process differs from the fixed-bed and moving-bed processes insofar as a powdered catalyst is circulated essentially as a fluid with the feedstock. The several fluid catalytic cracking processes in use differ primarily in mechanical design. Side-by-side reactor-regenerator construction and unitary vessel construction (the reactor either above or below the regenerator) are the two main mechanical variations.

Cracking of crude oil fractions occurs over many types of catalytic materials, but high yields of desirable products are obtained with hydrated aluminum silicates. These may be either activated (acid-treated) natural clays of the bentonite type of synthesized silica-alumina (SiO_2-Al_2O_3), or silica-magnesia (SiO_2-MgO) preparations. Their activity to yield essentially the same products may be enhanced to some extent by the incorporation of small amounts of other materials, such as the oxides of zirconium, boron (which has a tendency to volatilize away on use), and thorium. Natural and synthetic catalysts can be used as pellets, beads, or powder; all of the forms require replacement because of attrition and gradual efficiency loss. It is essential that they are able to withstand the physical impact of loading and thermal shocks, and the chemical action of carbon dioxide, air, sulfur and nitrogen compounds, and steam. Synthetic catalysts, or certain selected clays, appear to be better in this regard than average untreated natural catalysts.

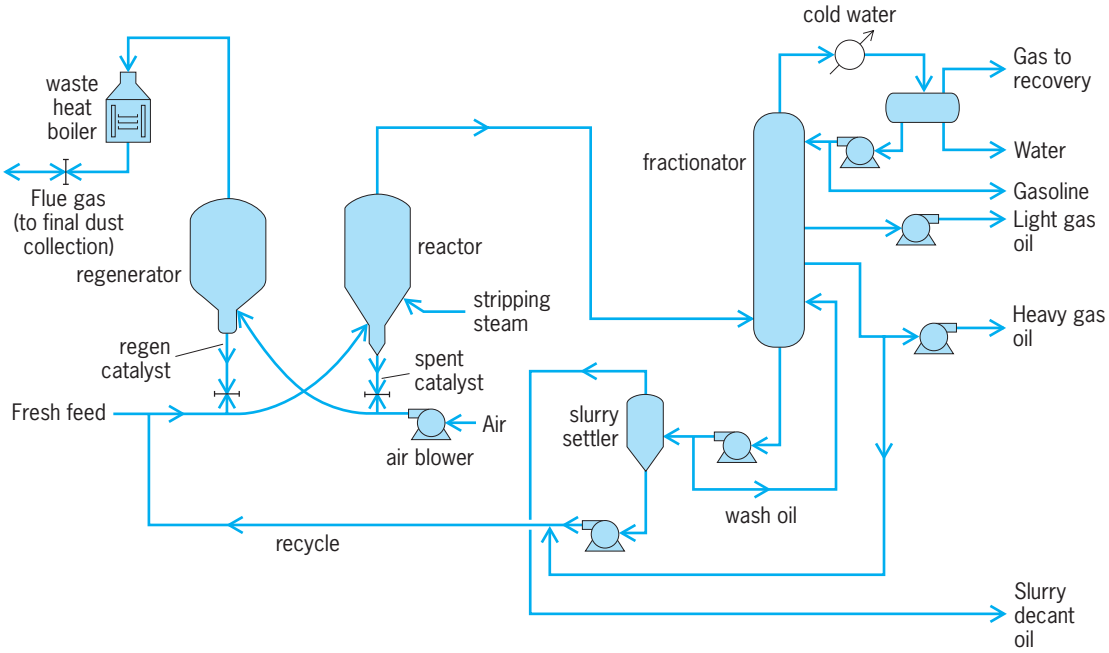

Fig. 7. Catalytic cracking process. (*After J. G. Speight, ed., The Chemistry and Technology of Petroleum, 3d ed., Marcel Dekker, New York, 1999*)

Hydroprocesses. The use of hydrogen in thermal processes was perhaps the single most significant advance in refining technology during the twentieth century. The process uses the principle that the presence of hydrogen during a thermal reaction of a petroleum feedstock will terminate many of the coke-forming reactions and enhance the yields of the lower-boiling components, such as gasoline, kerosine, and jet fuel. Hydrogenation processes for the conversion of petroleum fractions and petroleum products may be classified as destructive and nondestructive. *See* HYDROGENATION.

Destructive hydrogenation (hydrogenolysis or hydrocracking) is characterized by the conversion of the higher-molecular-weight constituents in a feedstock to lower-boiling products. Such treatment requires severe processing conditions and the use of high hydrogen pressures to minimize the polymerization and condensation reactions that lead to coke formation. *See* HYDROCRACKING.

Nondestructive hydrogenation is used for improving product quality without appreciable alteration of the boiling range. Nitrogen, sulfur, and oxygen compounds undergo reaction with the hydrogen, forming ammonia, hydrogen sulfide, and water, respectively. Unstable compounds that might lead to the formation of gums or insoluble materials are converted to more stable compounds.

Hydrotreating. Hydrotreating (nondestructive hydrogenation) is carried out by charging the feed to the reactor, together with hydrogen in the presence of catalysts, such as tungsten-nickel sulfide (W-NiS), cobalt-molybdenum-alumina (Co-Mo-Al_2O_3), nickel oxide-silica-alumina (NiO-SiO_2-Al_2O_3), and platinum-alumina (Pt-Al_2O_3). Most processes use cobalt-molybdena catalysts that generally contain about 10% molybdenum oxide and less than 1% cobalt oxide supported on alumina (Al_2O_3). The temperatures used are in the range of 300–345°C (570–655°F), while the hydrogen pressures are about 500–1000 psi (3447–6895 kPa).

The reaction generally takes place in the vapor phase but, depending on the application, may be a mixed-phase reaction. Generally it is more economical to hydrotreat high-sulfur feedstocks prior to catalytic cracking than to hydrotreat the cracking products. The advantages of hydrotreating are that (1) sulfur is removed from the catalytic cracking feedstock, and corrosion is reduced in the cracking unit; (2) carbon formation during cracking is reduced, resulting in higher conversions; and (3) the cracking quality of the gas oil fraction is improved.

Hydrocracking. Hydrocracking (destructive hydrogenation) is similar to catalytic cracking with hydrogenation superimposed and with the reactions taking place either simultaneously or sequentially. Hydrocracking was initially used to upgrade low-value distillate feedstocks, such as cycle oils (high aromatic products from a catalytic cracker which usually are not recycled to extinction for economic reasons), thermal and coker gas oils, and heavy-cracked and straight-run naphthas. Catalyst improvements and modifications have made it possible to yield products from gases and naphtha to furnace oils and catalytic cracking feedstocks. The hydrotreating catalysts are usually cobalt plus molybdenum or nickel plus molybdenum (in the sulfide) form impregnated on an alumina (Al_2O_3) base. The hydrotreated operating conditions are such that appreciable hydrogenation of aromatics will not occur [1000–2000 psi (6895–13,790 kPa) hydrogen and about 370°C (700°F)]. The desulfurization reactions are usually accompanied by small amounts of hydrogenation and hydrocracking.

In a hydrotreating process, the feedstock is heated and passed with hydrogen gas through a tower or reactor filled with catalyst pellets. The reactor is maintained at a temperature of 260–425°C (500–800°F) at pressures from 100 to 1000 psi (690 to 6895 kPa), depending on the particular process, the nature of the feedstock, and the degree of hydrogenation required. After leaving the reactor, excess hydrogen is separated from the treated product and recycled through the reactor after the removal of hydrogen sulfide. The liquid product is passed into a stripping tower where steam removes the dissolved hydrogen and hydrogen sulfide; after cooling, the product is stored or, in the case of feedstock preparation, pumped to the next processing unit.

The most common hydrocracking process is a two-stage operation (**Fig. 8**) that maximizes the yield of transportation fuels and has the flexibility to produce gasoline, naphtha, jet fuel, or diesel fuel to meet seasonal demand. The processes also operate at higher temperatures than the hydrotreating processes and use a more temperature-resistant catalyst.

Reforming processes. When the demand for higher-octane gasoline developed during the early 1930s, attention was directed to improving the octane number of fractions within the boiling range of gasoline. Straight-run (distilled) gasoline frequently had very low octane numbers, and any process that would improve the octane numbers would aid in meeting the demand for higher-octane-number gasoline. Such a process (thermal reforming) was developed and used widely, but to a much lesser extent

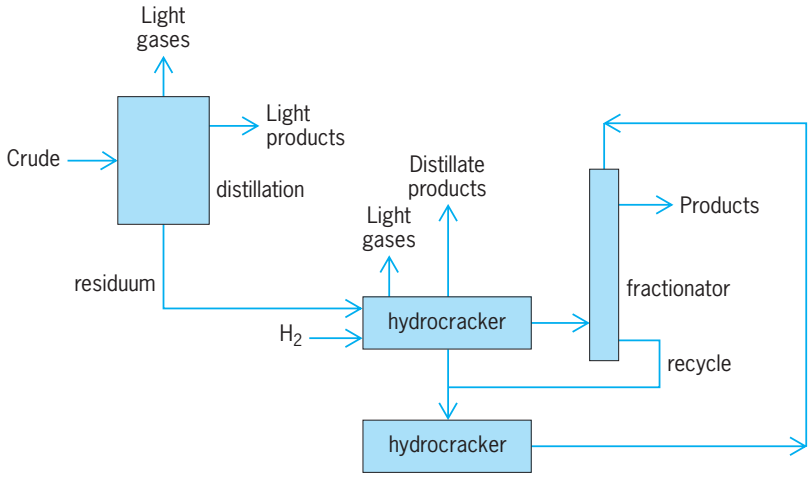

Fig. 8. Two-stage hydrocracking process. (*After J. G. Speight, ed., The Chemistry and Technology of Petroleum, 3d ed., Marcel Dekker, New York, 1999*)

than thermal cracking. Thermal reforming converts (reforms) gasoline into higher-octane gasoline. The equipment for thermal reforming is essentially the same as for thermal cracking, but uses higher temperatures.

Thermal reforming. In a thermal reforming process, the feedstock, such as 205°C (400°F) end-point naphtha or a straight-run gasoline, is heated to 510–595°C (950–1100°F) in a furnace, much the same as a cracking furnace, with pressures from 400 to 1000 psi (2758 to 6895 kPa). As the heated naphtha leaves the furnace, it is cooled or quenched by the addition of cold naphtha and then enters a fractional distillation tower where the heavy products are separated. The remainder of the reformed material leaves the top of the tower and is separated into gases and reformate. The reformate has a high octane number due to cracking of the longer-chain paraffins into higher-octane olefins.

The products of thermal reforming are gases, gasoline, and residual oil or tar, the latter being formed in very small amounts (about 1%). The amount and quality of gasoline, known as reformate, is very dependent on the temperature. As a rule, the higher the reforming temperature, the higher the octane number, but the lower the yield of reformate.

Catalytic reforming. Like thermal reforming, catalytic reforming converts low-octane gasoline into high-octane gasoline (reformate). While thermal reforming could produce reformate with octane numbers of 65–80 depending on the yield, catalytic reforming produces reformate with octane numbers on the order of 90–95.

Catalytic reforming usually is carried out by feeding a naphtha (after pretreating with hydrogen if necessary) and hydrogen mixture to a furnace for heating to the desired temperature, 450–520°C (840–965°F), and then passed through fixed-bed catalytic reactors at hydrogen pressures of 100–1000 psi (689–6895 kPa). Normally pairs of reactors are used in series, with heaters located between adjoining reactors to compensate for the endothermic reactions taking place. Sometimes as many as five reactors are kept on stream in series while one or more is being regenerated. The on-stream cycle of any one reactor may vary from several hours to many days, depending on the feedstock and reaction conditions.

The commercial processes available for use are broadly classified as the fixed-bed, moving-bed, and fluid-bed types. Fixed-bed processes use predominantly platinum-containing catalysts in units equipped for cycle, occasional, or no regeneration, whereas the fluid- and moving-bed processes use mixed nonprecious metal oxide catalysts in units equipped with separate regeneration facilities.

Dehydrogenation is a major chemical reaction in catalytic reforming, producing large quantities of hydrogen gas. The hydrogen is recycled though the reforming reactors, providing the atmosphere necessary for the chemical reactions and preventing carbon deposition on the catalyst, thus extending its operating life. An excess of hydrogen, above whatever is consumed in the process, is produced. As a result, catalytic reforming processes are unique in that they are the only petroleum refinery processes to produce hydrogen as a by-product. *See* DEHYDROGENATION.

The reforming catalyst composition is dictated by the composition of the feedstock and the desired reformate. The catalysts used are principally molybdena-alumina (Mo_2O_3-Al_2O_3), chromia-alumina (Cr_2O_3-Al_2O_3), or platinum on a silica-alumina (Pt-SiO_2-Al_2O_3) or alumina (Pt-Al_2O_3) base. Nonplatinum catalysts are widely used in regenerative process for feeds containing sulfur, which poisons platinum catalysts, although pretreatment processes (hydrodesulfurization) may permit platinum catalysts to be used.

The purpose of platinum on the catalyst is to promote dehydrogenation and hydrogenation reactions, that is, the production of aromatics, participation in hydrocracking, and rapid hydrogenation of carbon-forming precursors. For the catalyst to have an activity for isomerization of both paraffins and naphthenes—the initial cracking step of hydrocracking—and to participate in paraffin dehydrocyclization, it must have an acid activity. The balance between these two activities is most important in a reforming catalyst. In fact, in the production of aromatics from cyclic saturated materials (naphthenes), it is important that hydrocracking be minimized to avoid loss of the desired product; and, thus, the catalytic activity must be moderated relative to the case of gasoline production from a paraffinic feed, where dehydrocyclization and hydrocracking play an important role. *See* REFORMING PROCESSES.

Isomerization processes. During World War II, isomerization processes found commercial applications for making high-octane aviation gasoline components and additional feed for alkylation units. The lowered alkylate demands in the postwar period led to the majority of the butane isomerization units being shut down. In recent years the greater demand for high-octane motor fuel has resulted in new butane isomerization units being installed.

The earliest process of note was the production of isobutane, which is required as an alkylation feed. The isomerization may take place in the vapor phase, with the activated catalyst supported on a solid phase, or in the liquid phase with a dissolved catalyst. In the process, pure butane is mixed with hydrogen (to inhibit olefin formation) and passed to the reactor, at 110–170°C (230–340°F) and 200–300 psi (1379–2068 kPa). The product is cooled, the hydrogen is separated, and the cracked gases are then removed in a stabilizer column. The stabilizer bottom product is passed to a superfractionator where the normal butane is separated from the isobutane.

During World War II, aluminum chloride was the catalyst used to isomerize butane, pentane, and hexane. Since then, supported metal catalysts have been developed for use in high-temperature processes, which operate in the range 370–480°C (700–900°F) and 300–750 psi (2068–5171 kPa), while aluminum chloride plus hydrogen chloride is universally used for the low-temperature processes. *See* ALKANE.

A nonregenerable aluminum chloride catalyst is

used with various carriers in a fixed-bed or liquid contactor. Platinum or other metal catalyst processes are used in fixed-bed operation, and can be regenerable or nonregenerable. The reaction conditions vary widely, depending on the particular process and feedstock [40–480°C (100–900°F) and 150–1000 psi (1034–6895 kPa)]. *See* HETEROGENEOUS CATALYSIS; HOMOGENEOUS CATALYSIS; ISOMERIZATION.

Alkylation processes. The combination of olefins with paraffins to form higher isoparaffins is called alkylation. Since olefins are reactive (unstable) and are responsible for exhaust pollutants, their conversion to high-octane isoparaffins is desirable when possible. In refinery practice, only isobutane is alkylated, by reaction with isobutene or normal butene, and isooctane is the product. Although alkylation is possible without catalysts, commercial processes use aluminum chloride, sulfuric acid, or hydrogen fluoride as catalysts, allowing the reactions to take place at low temperatures and minimizing undesirable side reactions, such as polymerization of olefins.

Alkylate is composed of a mixture of isoparaffins having octane numbers that vary with the olefins from which they were made. Butylenes produce the highest octane numbers, propylene the lowest, and pentylenes the intermediate numbers. All alkylates, however, have high octane numbers (>87), making them particularly valuable.

The alkylation reaction as now practiced in petroleum refining is the reaction through catalysis, of an olefin (ethylene, propylene, butylene, and amylene) with isobutane to yield high-octane branched-chain hydrocarbons in the gasoline boiling range. Olefin feedstock is derived from the gas produced in a catalytic cracker, while isobutane is recovered from refinery gases or produced by catalytic butane isomerization. To accomplish this, either ethylene or propylene is combined with isobutane at 50–280°C (125–450°F) and 300–1000 psi (2068–6895 kPa) in the presence of a metal halide catalyst, such as aluminum chloride. Conditions are less stringent in catalytic alkylation; olefins (propylene, butylenes, or pentylenes) are combined with isobutane in the presence of an acid catalyst (sulfuric acid or hydrofluoric acid) at low temperatures and pressures (1–40°C, 34–104°F; 14.8–150 psi, 6.9–1034 kPa).

Sulfuric acid, hydrogen fluoride, and aluminum chloride are the general catalysts used commercially. Sulfuric acid is used with propylene and higher-boiling feeds, but not with ethylene, because it reacts to form ethyl hydrogen sulfate. The acid is pumped through the reactor and forms an air emulsion with the reactants, and the emulsion is maintained at 50% acid. The rate of catalyst deactivation varies with the feed and isobutane charge rate. Butene feeds cause less acid consumption than the propylene feeds.

Aluminum chloride is not widely used as an alkylation catalyst, but when it is, hydrogen chloride is used as a promoter and water is injected to activate the catalyst as an aluminum chloride/hydrocarbon complex. Hydrogen fluoride is used for alkylation of higher-boiling olefins bacause it is more readily separated and recovered from the products. *See* ALKYLATION (PETROLEUM).

Polymerization processes. In the petroleum industry, polymerization is the process by which olefin gases are converted to liquid products that may be suitable for gasoline (polymer gasoline) or other liquid fuels. The feedstock usually consists of propylene and butylenes from cracking processes or may even be selective olefins for dimer, timer, or tetramer production.

Polymerization may be accomplished thermally or in the presence of a catalyst at lower temperatures. Thermal polymerization is regarded as not being as effective as catalytic polymerization, but it can be used to polymerize saturated materials that cannot be induced to react by catalysts. The process consists of vapor-phase cracking of, for example, propane and butane, followed by a prolonged period at high temperature (510–595°C; 950–1100°F) for the reactions to proceed to near completion.

Olefins can also be conveniently polymerized by means of an acid catalyst in which an olefin-rich feed stream is contacted with a catalyst (sulfuric acid, copper pyrophosphate, phosphoric acid) at 150–220°C (300–425°F) and 150–1200 psi (1034–8274 kPa), depending on feedstock and product requirement.

Phosphates are the principal catalysts used in polymerization units; the commercially used catalysts are liquid phosphoric acid, phosphoric acid on kieselguhr (diatomite), copper pyrophosphate pellets, and phosphoric acid film on quartz. The last is the least active, but the most often used and the easiest to regenerate by washing and recoating, the serious disadvantage is that tar must occasionally be burned off the support. The process using liquid phosphoric acid catalyst is the one most often associated with attempts to raise production by increasing temperature. *See* POLYMERIZATION.

Solvent processes. Solvent processes, such as deasphalting, occur after the distillation units and are used chiefly to produce asphalt, deasphalted oil, and aromatic oils (Fig. 1). *See* SOLVENT EXTRACTION.

James G. Speight

Bibliography. R. A. Myers (ed.), *Handbook of Petroleum Refining Process*, McGraw-Hill, New York, 1997; J. G. Speight, *The Chemistry and Technology of Petroleum*, 3d ed., Marcel Dekker, New York, 1999; J. G. Speight, *The Desulfurization of Heavy Oils and Residua*, 2d ed., Marcel Dekker, New York, 2000; J. G. Speight (ed.), *Fuel Science and Technology Handbook*, Marcel Dekker, New York, 1990; J. G. Speight (ed.), *Petroleum Chemistry and Refining*, Taylor & Francis, New York, 1998.

Petroleum products

Petroleum products are those fractions derived from petroleum that have commercial value as a bulk product (**Table 1**). Petrochemicals, in contrast, are individual chemicals, derived from bulk fractions, that are used as the basic building blocks of the chemical industry. Gases and liquid fuels are currently the main

products of the petroleum industry (**Table 2**). However, other products, such as lubricating oils, waxes, and asphalt, have also added to the value of petroleum resources. *See* PETROCHEMICAL; PETROLEUM.

Petroleum products are hydrocarbon compounds, containing combinations of hydrogen and carbon with various molecular forms. Many compounds occur naturally. Other compounds are created by commercial processes for altering one combination to form another. Each combination has its unique set of chemical and physical properties. *See* PETROLEUM PROCESSING AND REFINING.

Many products replace similar products historically derived from nonpetroleum sources—for example: illuminating oil to replace sperm oil from whales; synthetic rubber to replace natural rubber from trees; and manufactured fibers to replace textiles from animals and vegetation. Each new use for a product imposes additional specifications regarding application and manufacturing methods. Specifications

TABLE 1. Petroleum products

- Hydrocarbon gases
  - *Liquefied gases* — Metal cutting gas, illumination gas
  - *Petroleum ether*
  - *Polymers* — Antiknock fuels, lubricating oils
  - *Alcohols, esters, ketones* — Solvents
  - *Acetylene*
    - Aldehydes — Resins
    - Acetic acid — Esters
    - Synthetic rubber
    - Acetylene black — Batteries
  - *Gas black* — Rubber tires, inks, paints
  - *Fuel gas*
  - *Light naphthas*
    - Light naphtha
      - Gas machine gasoline
      - Pentane, hexane
    - Intermediate naphthas
      - Aviation gasoline
      - Motor gasoline
      - Commercial solvents
        - Rubber solvent
        - Fatty oil solvent (extraction)
        - Lacquer diluents
      - Blending naphtha
    - Heavy naphthas
      - Varnish-makers and painters naphtha
      - Dyers and cleaners naphtha
      - Turpentine substitutes
- Light distillates
  - *Naphthas*
  - *Refined oils*
    - Refined kerosine — Stove fuel, lamp fuel, tractor fuel
    - Signal oil — Railroad signal oil, lighthouse oil
    - Mineral seal oil — Coach and ship illuminants, gas absorption oils
- Intermediate distillates
  - *Gas oil*
    - Water gas carburetion oils
    - Metallurgical fuels
    - Cracking stock for gasoline manufacture
    - Household heating fuels
    - Light industrial fuels
    - Diesel fuels oils
  - *Absorber oil* — Gasoline recovery oil, benzol recovery oil
- Heavy distillates
  - *Technical oils*
    - White oils
      - Techincal
        - Tree spray oils
        - Bakers machinery oil, fruit packers oil
        - Candymakers oil
        - Egg packers oil
        - Slab oil
      - Medicinal — Internal lubricant, salves, creams, ointments
    - Saturating oils — Wood oils, leather oils, twine oils
    - Emulsifying oils — Cutting oils, textile oils, paper oils, leather oils
    - Electrical oils — Switch oils, transformer oils, metal recovery oils
    - Flotation oils
  - *Paraffin wax*
    - Candymakers and chewing gum wax
    - Candle wax, laundry wax, sealing wax, etchers wax
    - Saturating wax, insulation wax — Match wax, cardboard wax, paper wax
    - Medicinal wax
    - Canning wax
    - Paraflow
    - Fatty acids — Grease, soap, lubricant
    - Fatty alcohols and sulfates — Rubber compounding, detergents, wetting agents
  - *Lubricating oils*
    - Light spindle oils
    - Transformer oils
    - Household lubricating oils
    - Compressor oils
    - Ice machine oils
    - Meter oils
    - Journal oils
    - Motor oils
    - Diesel oils
    - Engine oils
    - Steam cylinder oils
    - Valve oils
    - Turbine oils
    - Dust-laying oils
    - Tempering oils
    - Transmission oils
    - Railroad oils
    - Printing ink oils
    - Black oils
    - Lubricating greases
  - *Petroleum grease* — Petrolatum
    - Medicinal
      - Salves, creams, and ointments
      - Petroleum jelly
    - Technical
      - Rust-preventing compounds
      - Lubricants
      - Cable-coating compound

(*continued*)

TABLE 1. Petroleum products (*cont.*)

- Residues
  - *Residual fuel oil*
    - Wood preservative oils
    - Boiler fuel
    - Gas manufacture oils
    - Metallurgical oils
  - *Still wax*
    - Roofing material
  - *Asphalts*
    - Liquid asphalts
    - Binders
      - Roofing saturants, road oils, emulsion bases
    - Fluxes
    - Steam-related asphalts
      - Briqueting asphalts
      - Paving asphalts
      - Shingle saturants
      - Paint bases
      - Flooring saturants
    - Oxidized asphalts
      - Roof coatings
      - Waterproofing asphalts
      - Rubber substitutes
      - Insulating asphalts
  - *Coke*
    - Carbon electrode coke, carbon brush coke, fuel coke
- Refinery sludges
  - *Acid coke*
    - Fuel
  - *Sulfonic acid*
    - Saponification agents
    - Demulsifying agnets
    - Emulsifiers
  - *Heavy fuel oils*
    - Refinery fuel
  - *Sulfuric acid*
    - Fertilizers

SOURCE: After P. Albery Washer, Texas A&M, College Extension Division (First Session).

for petroleum products are based on properties such as density and boiling range to assure that a petroleum product can perform its intended task.

Gaseous products. Fuels with four or fewer carbons in the hydrogen-carbon combination have boiling points lower than room temperature. Therefore, these products are normally gases. Common classifications for these products are as follows.

Natural gas. This is predominantly methane (CH_4), which has the lowest boiling point and least complex structure of all hydrocarbons. Natural gas from an underground reservoir, when brought to the surface, may contain other, higher-boiling-point hydrocarbons, and is often referred to as wet gas. Wet gas is processed to remove the entrained hydrocarbons that are higher-boiling than methane. The high-boiling hydrocarbons that are isolated and liquefied are called natural gas condensates. *See* METHANE; NATURAL GAS.

Still gas. This is a broad term applied to low-boiling hydrocarbon mixtures, and is the lowest-boiling fraction isolated from a distillation (still) unit in the refinery. If the distillation unit is separating only low-boiling hydrocarbon fractions, the still gas will be almost entirely methane with only trace amounts of ethane (CH_3CH_3) and ethylene ($CH_2{=}CH_2$). If the distillation unit is handling higher-boiling fractions, the still gas may also contain propane ($CH_3CH_2CH_3$), butane ($CH_3CH_2CH_2CH_3$), and their respective isomers. Fuel gas and still gas are terms that are often used interchangeably. However, fuel gas is intended to denote the product's future use as a fuel for boilers, furnaces, or heaters. *See* ETHYLENE.

Liquefied petroleum gas (LPG). This is composed of propane and butane and is stored under pressure in order to keep these hydrocarbons liquefied at normal atmospheric temperatures. Before liquefied petroleum gas is burned, it passes through a pressure

TABLE 2. Commercial names and uses for major petroleum products

Crude oil cuts	Refinery blends	Consumer products
Gases	Still gases	Fuel gas
	Propane/butane	Liquefied petroleum gas (LPG)
Light/heavy naphtha	Motor fuel	Gasoline
	Aviation turbine, Jet-B	Jet fuel (naphtha type)
Kerosine	Aviation turbine, Jet-A	Jet fuel (kerosine type)
	No. 1 fuel oil	Kerosine (range oil)
Light gas oil	Diesel	Auto and tractor diesel
	No. 2 fuel oil	Home heating oil
Heavy gas oil	No. 4 fuel oil	Commercial heating oil
	No. 5 fuel oil	Industrial heating oil
	Bright stock	Lubricants
Residuals	No. 6 fuel oil	Bunker C oil
	Heavy residual	Asphalt
	Coke	Coke

relief valve that causes a reduction in pressure and the liquid vaporizes (gasifies). Winter-grade liquefied petroleum gas is mostly propane, the lower-boiling of the two gases, that is easier to vaporize at lower temperatures. Summer-grade liquefied petroleum gas is mostly butane. *See* LIQUEFIED PETROLEUM GAS (LPG).

Gasoline. Gasoline (motor fuel) is a complex mixture of hydrocarbons that boils below 200°C (390°F) and is intended for most spark-ignition engines (such as those used in passenger cars, light-duty trucks, motorcycles, and motorboats). The properties of gasoline are intended to satisfy the requirements of smooth and clean burning, easy ignition in cold weather, minimal evaporation in hot weather, and stability during long storage periods. *See* GASOLINE; INTERNAL COMBUSTION ENGINE.

The hydrocarbon constituents in this boiling range are those that have 4 to 12 carbon atoms in their molecular structure. Gasolines can vary widely in composition; even those with the same octane number may be quite different. For example, low-boiling distillates with high aromatic content (above 20%) can be obtained from some crude oils. The variation in the content of aromatics, as well as normal paraffins, branched paraffins, cyclopentanes, and cyclohexanes, involve the characteristics of any individual crude oil and influence the octane number of the gasoline. *See* AROMATIC HYDROCARBON.

During the first decade of the twentieth century, the gasoline produced was originally present in crude oil or condensed from natural gas. However, it was soon discovered that if the heavier portions of petroleum (for example, the fraction which boiled higher than kerosine, such as gas oil) were heated to more severe temperatures, thermal degradation (or cracking) occurred to produce smaller molecules within the range suitable for gasoline. Therefore, gasoline not originally present in the crude petroleum could be manufactured. *See* CRACKING.

At first, cracked gasoline was regarded as an inferior product because of its comparative instability on storage; but as more gasoline was required, the petroleum industry revolved around processes by which this material could be produced (such as, catalytic cracking, thermal and catalytic reforming, hydrocracking, alkylation, and polymerization), and the problem of storage instability was addressed and resolved. *See* ALKYLATION (PETROLEUM); HYDROCRACKING; POLYMERIZATION; REFORMING PROCESSES.

Because of the differences in composition of gasoline fractions, blending the components is necessary to meet gasoline specifications. The physical process of blending is simple, but determining how much of each component to include is difficult. The operation is carried out by simultaneously pumping all the components of a gasoline blend into a pipeline that leads to the gasoline storage, but the pumps must be set to deliver automatically the proper proportion of each component. Baffles in the pipeline are often used to mix the components as they travel to the storage tank.

Regular and premium grades. These are the relative classifications for the octane numbers of gasoline. An octane number is a measure of a gasoline's ability to resist spontaneous detonation. It is critical that detonation be at a precise time for a gasoline-air mixture in a spark-ignition engine. That time is determined by the electrical spark system. After ignition, the course of the detonation should progress smoothly, with a flame front moving across the combustion chamber.

If the fuel has a low octane number, the temperature and pressure wave generated by the spark-timed flame front can allow the remaining fuel-air mixture to spontaneously ignite, causing an extra pressure pulse (knock), more fuel energy to be lost as heat, and the engine to deliver less power. *See* SPARK KNOCK.

The knocking tendencies of two pure hydrocarbons are used as references to relate final engine conditions to an octane number. One of these hydrocarbons, isooctane (2,2,4-trimethylpentane), is assigned an octane number of 100, and the other hydrocarbon (*n*-heptane) is assigned an octane number of 0. Mixtures of these reference hydrocarbons are assigned octane numbers equal to the volume percent of isooctane in the mixture. To determine the octane rating of a sample fuel, its knock intensity in a test engine is compared to those from these reference mixtures. *See* OCTANE NUMBER.

Research (RON or R) and motor octane (MON or M) grades. These grades identify other test engine variables (such as engine speed and intake air temperature), with the motor octane number being representative of the performance of the engine at higher revolutions per minute and higher air intake temperature. Both ratings are important to a multicylinder automobile engine because it usually operates over a wide range of conditions. The arithmetic average of the research and motor octane numbers [(RON + MON)/2] is used for marketing purposes.

Leaded and unleaded grades. These grades of gasoline denote whether the gasoline mix includes lead additives. Lead compounds, such as tetraethyllead and tetramethyllead, are additives to improve the octane rating of a gasoline mix. However, the introduction of the catalytic muffler to automobile engines (to reduce exhaust emissions) led to the phase-out of leaded gasoline (**Table 3**). *See* LEAD.

TABLE 3. Trends in gasoline specifications

Time frame	Specifications
Pre-1994	Phase-out of lead in gasoline
	Lower Reid vapor pressure
1995–2000	Reformulated gasoline (Phase I)
	Lower aromatics content: 25 vol % maximum
	Benzene: 1 vol % maximum
	Oxygen: 2 wt % minimum
	Reid vapor pressure: 8.1 psi maximum
Post-2000	Reformulated gasoline (Phase II)
	Sulfur: 30 ppm maximum
	Aromatics content: 22 vol % maximum
	Benzene: 0.8 vol % maximum
	Oxygen: >2 wt %
	Olefins: 4 vol %
	Reid vapor pressure: 7.0 psi maximum

Volatility. This is also an important property of gasoline, since it is related to ease of startup, length of warmup, extent of carburetor icing, speed of acceleration, occurrence of vapor lock, quality of manifold distribution, and extent of crankcase dilution. Volatility comprises two extreme properties: (1) enough low-boiling hydrocarbons to vaporize easily in cold weather; and (2) enough high-boiling hydrocarbons to remain a liquid in an engine's fuel supply system during hot periods.

The reduction of the lead content of gasoline and the introduction of reformulated gasoline have been very successful in reducing automobile emissions due to changes in gasoline composition (Table 3). Further improvements in fuel quality have been proposed for the early decades of the twenty-first century. These projections are accompanied by a noticeable and measurable decrease in crude oil quality. The reformulated gasoline will help meet environmental regulations for liquid fuel emissions, but will be subject to continuous review because of the potential for environmental impact. *See* AUTOMOBILE.

Aviation fuel. This comes in two types: gasoline and jet fuel. Aviation gasoline, now usually found in use in light aircraft and older civil aircraft, has a narrower boiling range (38–170°C or 100–340°F) than conventional (automobile) gasoline (−1 to 200°C or 30–390°F). The narrower boiling range ensures better distribution of the vaporized fuel through the more complicated induction systems of aircraft engines. Since aircraft operate at high altitudes (5.3 km or 3.3 mi) where the prevailing pressure (0.5 atm or 7.4 psi) is less than the pressure at the surface of the Earth (1.0 atm or 14.7 psi), the vapor pressure of aviation gasoline must be limited to reduce boiling in the tanks, fuel lines, and carburetors. Jet fuel is classified as aviation turbine fuel, and the specifications and ratings relative to octane number are replaced with properties connected with the ability of the fuel to burn cleanly. *See* AIRCRAFT FUEL.

Solvents. Petroleum solvents (naphthas) have been available since the early days of the petroleum industry. They are valuable as solvents because of their nonpoisonous character and good dissolving power. The wide range of naphthas available and the varying degree of volatility possible offer products suitable for many uses. *See* NAPHTHA.

Petroleum naphtha is a generic term applied to refined, partly refined, or unrefined petroleum products. Naphthas are prepared by several methods, including (1) fractionation of distillates or crude petroleum, (2) solvent extraction, (3) hydrogenation of distillates, (4) polymerization of unsaturated (olefinic) compounds, and (5) alkylation processes. The naphtha may also be a combination of product streams from more than one process.

The main uses of petroleum naphthas fall into the general areas of (1) solvents (diluents) for paints, (2) dry-cleaning solvents, (3) solvents for cutback asphalts, (4) solvents in rubber industry, and (5) solvents for industrial extraction processes. Turpentine, the traditional solvent for paints, has been almost completely replaced by the cheaper and more abundant petroleum naphthas. *See* SOLVENT.

Kerosine. Kerosine was the major refinery product before the onset of the automobile age and the ensuing emphasis on gasoline. Kerosine originated as a straight-run (distilled) petroleum fraction that boiled between approximately 205 and 260°C (400 and 500°F). In the early days of petroleum refining, some crude oils contained kerosine fractions of very high quality; but other crude oils, such as those having a high proportion of asphaltic materials, must be thoroughly refined to remove aromatics and sulfur compounds before a satisfactory kerosine fraction can be obtained.

The kerosine fraction is essentially a distillation fraction of petroleum. The quantity and quality of the kerosine vary with the type of crude oil; some crude oils yield excellent kerosine, while others produce kerosine that requires substantial refining. Kerosine is a very stable product, and additives are not required to improve the quality. Apart from the removal of excessive quantities of aromatics, kerosine fractions may need only a lye (alkali) wash if hydrogen sulfide is present. Kerosine is used as a fuel for heating and cooking, jet engines, and lamps, for weed burning, and as a base for insecticides. *See* KEROSINE.

Diesel fuel. Diesel fuel is a distillate product that has a higher boiling point than gasoline (or naphtha) but that also must self-ignite easily. This is determined through the cetane rating, derived from the reference fuel *n*-cetane. Cetane number is a measure of the tendency of a diesel fuel to knock in a diesel engine. The scale is based upon the ignition characteristics of two hydrocarbons, *n*-hexadecane (cetane) and 2,3,4,5,6,7,8-heptamethylnonane. *See* CETANE NUMBER; DIESEL ENGINE; DIESEL FUEL.

Cetane has a short delay period during ignition and is assigned a cetane number of 100; heptamethylnonane has a long delay period and is assigned a cetane number of 15. Just as the octane number is meaningful for automobile fuels, the cetane number is a means of determining the ignition quality of diesel fuels, and is equivalent to the percentage by volume of cetane in the blend with heptamethylnonane, which matches the ignition quality of the test fuel.

Sulfur content is an important specification for diesel fuels, as a predictor of sulfur emissions into the atmosphere. While most sulfur compounds are removed from gasoline prior to use, the sulfur content of diesel fuel (and fuel oil) requires consideration. *See* SULFUR.

Diesel fuel oil is essentially the same as furnace fuel oil, but the proportion of cracked gas oil is usually less since the high aromatic content of the cracked gas oil reduces the cetane value of the diesel fuel.

Fuel oil. Fuel oil is classified as two main types: distillate and residual. It is also divided into subtypes based on use (**Table 4**). *See* FUEL OIL.

Distillate fuel oil is vaporized and condensed during a distillation process, has a definite boiling range, and does not contain high-boiling constituents such as asphaltic components. A fuel oil that contains any

TABLE 4. Properties of types of fuel oils

Types	Properties
No. 1 fuel oil	Similar to kerosine or range oil (fuel used in stoves for cooking) Defined as a distillate intended for vaporizing in pot-type burners and other burners where a clean flame is required
No. 2 fuel oil	Often called domestic heating oil Similar to diesel and higher-boiling jet fuels Defined as a distillate for general-purpose heating in which the burners do not require the fuel to be completely vaporized before burning
No. 4 fuel oil	A light industrial heating oil that is intended where preheating is not required for handling or burning Two grades differing primarily in safety (flash) and flow (viscosity) properties
No. 5 fuel oil	A heavy industrial oil that often requires preheating for burning and, in cold climates, for handling
No. 6 fuel oil	A heavy residuum oil Commonly referred to as Bunker C oil when it is used to fuel oceangoing vessels Preheating required for both handling and burning

amount of the residue from crude distillation or thermal cracking is a residual fuel oil. Diesel fuel oil is generally a distillate fuel oil, but residual oil has been successfully used to power marine diesel engines, and mixtures of distillate and residua have been used on locomotive diesel engines. *See* LOCOMOTIVE.

The terms "distillate fuel oil" and "residual fuel oil" are losing their significance, since fuel oil is now made for specific uses and may be distillate, residuum, or a mixture of the two. The terms "domestic fuel oil" and "heavy fuel oil" are more indicative of the uses.

Domestic fuel oil. This term is applied to fuel oil used primarily in the home, and includes kerosine, stove oil, and furnace fuel oil. Stove oil is a straight-run (distilled) fraction from crude oil, whereas other fuel oils are usually blends of two or more fractions. The straight-run fractions available for blending into fuel oils are heavy naphtha, light and heavy gas oil, and residua. Cracked fractions such as light and heavy gas oil from catalytic cracking, cracking coal tar, and fractionator bottoms from catalytic cracking may also be used as blends to meet the specifications of the different fuel oils.

Heavy fuel oil. This term includes a variety of oils, ranging from distillates to residual oils that must be heated to 260°C (500°F) or higher before they can be used. In general, heavy fuel oil consists of residual oil blended with distillate oil to suit specific needs and to meet designed specifications. Included among the category of heavy fuel oil are various industrial oils; when used to fuel ships, heavy fuel oil is called bunker oil.

Heavy fuel oil usually contains residuum that is mixed (cut back) to a specified viscosity with gas oils and fractionator bottoms. For some industrial purposes where flames or flue gases contact the product (ceramics, glass, heat treating, open hearth furnaces), the fuel oil must be blended to have a minimum specified sulfur content.

Lubricating oil. After kerosine, the early petroleum refiners wanted paraffin wax for the manufacture of candles. Demand for the lubricating oils that were the by-products of paraffin wax manufacture did not grow until heavy industry increased, after the 1890s; at this time, petroleum largely replaced animal oils (such as lard oil, sperm whale oil, and tallow) and vegetable oils as the source of lubricants. *See* LUBRICANT.

Lubricating oils are distinguished from other fractions of crude oil by a high boiling point (>400°C or 750°F), as well as high viscosity. Materials suitable for the production of lubricating oils are composed principally of hydrocarbons containing 25–35 carbon atoms per molecule, whereas residual stocks may contain hydrocarbons with 50–80 carbon atoms per molecule. *See* VISCOSITY.

Lubricating oil manufacture was well established by 1880, the method used depending on whether the crude petroleum was processed primarily for kerosine or for lubricating oils. Usually the crude oil was processed for kerosine, and primary distillation separated the crude into three fractions—naphtha, kerosine, and a residuum. To increase the production of kerosine, the cracking distillation technique was used, and this converted a large part of the gas oils and lubricating oils into kerosine.

The development of vacuum distillation provided the means of separating more suitable lubricating oil fractions with predetermined viscosity ranges and removed the limit on the maximum viscosity that might be obtained in a distillate oil. Vacuum distillation prevented residual asphaltic material from contaminating lubricating oils, but did not remove other undesirable materials such as acidic components or components that caused the oil to thicken excessively when cold and become very thin when hot. *See* DISTILLATION.

Lubricating oils may be divided into many categories according to the types of service they are intended to perform. However, there are two main groups: oil used in intermittent service, such as motor and aviation oil; and oil designed for continuous service, such as turbine oil.

Oils used in intermittent service must show the least possible change in viscosity with temperature, and these oils must be changed at frequent intervals to remove the foreign matter that collects during service. The stability of such oils is, therefore, of less importance than the stability of oils used in continuous service for prolonged periods without renewal. Oils used in continuous service must be extremely stable because such engines operate at fairly constant temperature without frequent shutdown.

White oil, insulating oil, and insecticides. The term "white distillate" is applied to all the refinery streams with a distillation range between approximately 80 and 360°C (175 and 680°F) at atmospheric pressure and with properties similar to the corresponding straight-run distillate from atmospheric crude distillation. Light distillate products, such as naphthas,

kerosine, jet fuels, diesel fuels, and heating oils, are manufactured by appropriate blending of white distillate streams.

For many years, most white distillate products originated from naphthenic stocks, but now they are prepared from paraffinic, mixed-base, or naphthenic fractions, depending on the final use of the oil. Naphthenic crude oils give products of high specific gravity and viscosity, desirable in pharmaceutical use; whereas paraffinic stocks produce oils of lower specific gravity and lower viscosity suitable for lubrication purposes.

White oils. This category of petroleum products falls into two classes: those often referred to as technical white oils, which are employed for cosmetics, textile lubrication, insecticide vehicles, paper impregnation, and so on; and pharmaceutical white oils, which may be employed as laxatives or for the lubrication of food-handling machinery. The colorless character of these oils is important in some cases, as it may indicate the chemically inert nature of the hydrocarbon constituents. Textile lubricants should be colorless to prevent the staining of light-colored threads and fabrics. Insecticide oils should be free of reactive (easily oxidized) constituents so as not to injure plant tissues when applied as sprays. Laxative oils should be free of odor, taste, and also hydrocarbons, which may react during storage and produce unwanted by-products. These properties are attained by the removal of oxygen-, nitrogen-, and sulfur-containing compounds, as well as reactive hydrocarbons by, say, sulfuric acid.

The medicinal oils require a test showing minimal color change, but depending on its intended use, a technical oil showing rather marked color change may be satisfactory. The only further distinction between pharmaceutical and technical oils is that the high-quality medicinal oils are made as viscous as possible, whereas the technical oils are likely to be made of the less viscous fractions.

Insulating oils. These fall into two classes: those used in transformers, circuit breakers, and oil-filled cables; and those employed for impregnating the paper covering of wrapped cables. The first are highly refined fractions of low viscosity and a comparatively high boiling temperature range that resemble heavy burning oils, such as mineral seal oil, or the very light lubricating fractions known as nonviscous neutral oils. The second are usually highly viscous products, often naphthenic distillates, and are not usually highly refined.

The insulating value of fresh transformer oils seems to vary little with chemical constitution, but physical purity, including freedom from water, is highly significant. A water content of 0.1% lowers an original dry insulating value by a factor of about 10; higher water content causes little additional change.

Insecticides. These are petroleum oils that are usually applied in water-emulsion form and have marked killing power for certain species of insects. For many applications in which their own effectiveness is too slight, the oils serve as carriers for active poisons, as in household and livestock sprays.

The physical properties of petroleum oils, such as their solvent power for waxy coatings on leaf surfaces and insect bodies, make them suitable as carriers for more active fungicides and insecticides. The additive substance may vary from fatty acids and soaps, the latter intended chiefly to affect favorably the spreading properties of the oil, to physiologically active compounds such as pyrethrum, nicotine, rotenone, thiocyanates, methoxychlor, and lindane. *See* INSECTICIDE.

Grease. Grease is a lubricating oil in which a thickening agent has been added for the purpose of holding the oil to surfaces that must be lubricated. The most widely used thickening agents are soaps of various kinds, and grease manufacture is essentially the mixing of soaps with lubricating oils.

Soap is made by chemically combining a metal hydroxide with a fat or fatty acid [reaction (1)]. The

$$\underset{\text{Fatty acid}}{RCO_2H} + \underset{\text{Sodium hydroxide}}{NaOH} \longrightarrow \underset{\text{Soap}}{RCO_2^-Na^+} + \underset{\text{Water}}{H_2O} \qquad (1)$$

most common metal hydroxides used for this purpose are calcium hydroxide, lye, lithium hydroxide, and barium hydroxide. Fats are chemical combinations of fatty acids and glycerin. If a metal hydroxide is reacted with a fat, a soap containing glycerin is formed. Frequently a fat is separated into its fatty acid and glycerin components, and only the fatty acid portion is used to make soap. Commonly used fats for grease-making soaps are cottonseed oil, tallow, and lard. Among the fatty acids used are oleic acid (from cottonseed oil), stearic acid (from tallow), and animal fatty acids (from lard). *See* FAT AND OIL; SOAP.

Grease may contain 30–50% soap, and although the fatty acid influences the properties of a grease, the metal in the soap has the most important effect. For example, calcium soaps form smooth buttery greases that are resistant to water but are limited in use to temperatures under about 95°C (200°F). Soda (sodium) salts form fibrous greases that disperse in water and can be used at temperatures well over 95°C (200°F). Barium and lithium soaps form greases similar to those from calcium soaps, and can be used at both high temperatures and very low temperatures; hence barium and lithium soap greases are known as multipurpose greases. In addition to soap and oil, grease may contain additives that are used to improve the ability of the grease to stand up under extreme bearing pressures, to act as a rust preventive, and to reduce the tendency of oil to seep or bleed from a grease. Graphite, mica, talc, or fibrous material may be added to grease that is used to lubricate rough machinery, and other chemicals can make grease more resistant to oxidation.

Wax. Petroleum wax can be categorized as two types: paraffin wax in petroleum distillates, and microcrystalline wax in petroleum residua. The melting point of wax is not directly related to its boiling point, because waxes contain hydrocarbons of different chemical structure. Nevertheless, waxes are graded according to their melting point and oil content. *See* WAX, PETROLEUM.

Paraffin wax is a solid crystalline mixture of straight-chain (normal) hydrocarbons ranging from C20 to C30 and higher. Wax constituents are solid at ordinary temperatures (25°C or 77°F), whereas petrolatum (petroleum jelly) does contain both solid and liquid hydrocarbons. *See* PARAFFIN; PETROLATUM.

Wax production by a process called wax sweating was originally used in Scotland to separate wax fractions with various melting points from the wax obtained from shale oils. Wax sweating is still used to some extent but is being replaced by the more convenient wax recrystallization process. In wax sweating, a cake of slack wax is slowly warmed to a temperature at which the oil in the wax and the lower-melting waxes become fluid and drip (or sweat) from the bottom of the cake, leaving a residue of higher-melting wax. *See* DEWAXING OF PETROLEUM.

The amount of oil separated by sweating is now much smaller than it used to be, due to the development of highly efficient solvent dewaxing techniques. Wax sweating is now more concerned with the separation of slack wax into fractions with different melting points. The three main methods used in modern refinery technology are: (1) solvent dewaxing, in which the feedstock is mixed with one or more solvents; then the mixture is cooled down to allow the formation of wax crystals, and the solid phase is separated from the liquid phase by filtration; (2) urea dewaxing, in which urea forms adducts with straight-chain paraffins that are separated by filtration from the dewaxed oil; (3) catalytic dewaxing, in which straight-chain paraffin hydrocarbons are selectively cracked on zeolite-type catalysts and the lower-boiling reaction products are separated from the dewaxed lubricating oil by fractionation.

Wax recrystallization, like wax sweating, separates wax into fractions; but instead of relying upon differences in melting points, the process makes use of the different solubilities of the wax fractions in a solvent such as a ketone. When a mixture of ketone and wax is heated, the wax usually dissolves completely; and if the solution is cooled slowly, a temperature is reached at which a crop of wax crystals is formed. These crystals will be of the same melting point, and if they are removed by filtration, a wax fraction with a specific melting point is obtained. If the clear filtrate is cooled further, a second batch of wax crystals with a lower melting point is obtained. Thus, alternate cooling and filtration can subdivide the wax into a large number of wax fractions, each with different melting points.

Recrystallization can also be applied to the microcrystalline waxes obtained from intermediate and heavy paraffin distillates, which cannot be sweated. Indeed, the microcrystalline waxes have higher melting points and differ in their properties from the paraffin waxes obtained from light paraffin distillates; and thus, wax recrystallization has made new kinds of waxes available.

Asphalt. Asphalt is a residuum that cannot be distilled even under the highest vacuum since the temperatures required to volatilize the residuum promote the formation of coke. Asphalts have complex chemical and physical compositions that usually vary with the source of the crude oil. *See* ASPHALT AND ASPHALTITE.

Asphalt manufacture commences with distilling everything possible from crude petroleum until a residuum with the desired properties is obtained. This is usually done in two stages: (1) Distillation at atmospheric pressure removes the lower-boiling fractions and yields a reduced crude that may contain higher-boiling (lubricating) oils, asphalt, and even wax. (2) Distillation of the reduced crude under vacuum removes the oils (and wax) as volatile overhead products, and the asphalt remains as a bottom (or residual) product. If the residuum obtained by distillation does not meet the desired specification, blowing the residuum with air can then be applied. Propane deasphalting also produces asphalt.

There are wide variations in refinery operations and in the types of crude oils and different asphalts that will be produced. Blending with higher- and lower-softening-point asphalts may make asphalts of intermediate softening points. If lubricating oils are not required, the reduced crude may be distilled in a flash drum that is similar to a distillation tower but has few, if any, trays. Asphalt descends to the base of the flasher as the volatile components pass out of the top. Asphalt can be made softer by blending the hard asphalt with the extract obtained in the solvent treatment of lubricating oils. Soft asphalts can be converted into harder asphalts by oxidation (air blowing).

Road oil is liquid asphalt that is intended for easy application to earth roads to provide a strong base or a hard surface and will maintain a satisfactory passage for light traffic. Liquid road oils, cutback asphalt, and emulsion asphalt are of recent date, but use of asphaltic solids for paving goes back to the European practices of the early 1800s. Cutback asphalt is a mixture in which hard asphalt has been diluted with a lighter oil to permit application as a liquid without drastic heating. Cutback asphalt may be described as rapid-, medium-, and slow-curing, depending on the volatility of the diluent that governs the rate of evaporation and consequent hardening.

An asphaltic material may be emulsified with water to permit application without heating. Such emulsions are normally of the oil-in-water type. They reverse or break on application to a stone or earth surface, so that the oil clings to the stone and the water disappears. In addition to their usefulness in road and soil stabilization, they are useful for paper impregnation and waterproofing. The emulsions are chiefly the soap or alkaline type and the neutral or clay type. The former break readily on contact, but the latter are more stable and probably lose water mainly by evaporation. Good emulsions must be stable during storage or freezing, suitably fluid, and amenable to control for speed of breaking.

In the post-1980 period, a shortage of good-quality asphalts developed in the United States and in other parts of the world. Due to the tendency of refineries to produce as much liquid fuel (such as gasoline) as

possible to meet market demand, residua that would have once been used for asphalt manufacture was being used to produce liquid fuels (and coke). Now asphalt, once the garbage product of a refinery, has again become a valuable refinery product.

Coke. Petroleum coke is the residue left by the noncatalytic destructive distillation (thermal decomposition with simultaneous removal of distillate) of petroleum residua. The coke formed in catalytic cracking operations is usually nonrecoverable because it adheres to the catalyst employed as fuel for the process. *See* COKE.

The composition of the coke varies with the source of the crude oil, but in general, large amounts of high-molecular-weight complex hydrocarbons (rich in carbon but correspondingly poor in hydrogen) make up a high proportion.

Petroleum coke is employed for a number of purposes, but the major use is in the manufacture of carbon electrodes for aluminum refining, which requires a high-purity carbon (that is, low in ash and sulfur-free). In addition, petroleum coke is employed in the manufacture of carbon brushes, silicon carbide abrasives, and structural carbon (such as pipes and Rashig rings), as well as in the manufacture of calcium carbide (CaC_2) from which acetylene is produced [reaction (2)].

$$\underset{\text{Calcium carbide}}{CaC_2} + \underset{\text{Water}}{H_2O} \longrightarrow \underset{\text{Acetylene}}{HC{\equiv}CH} \qquad (2)$$

In the refining before the 1920s, coke was an unwanted by-product and was usually discarded. As more coking operations became integral parts of refinery operations in the post-1920 era, coke was produced in significant amounts, and the demand for coke as an industrial fuel and for graphite/carbon electrode manufacture increased after World War II.

The use of coke as a fuel must proceed with some caution. With the acceptance by refiners of the heavier crude oils as refinery feedstocks, the higher content of sulfur and nitrogen in these oils yields a product coke containing substantial amounts of both elements. Both will produce unacceptable pollutants—sulfur oxides (SO_x) and nitrogen oxides (NO_x)—during combustion. They must also be regarded with caution in any coke that is scheduled for electrode manufacture, and removal procedures for these elements are continually being developed.

Sulfonic acids and sulfuric acid sludge. Petroleum distillates are generally treated with sulfuric acid to dissolve unstable or colored substances and sulfur compounds, as well as to precipitate asphaltic materials. When drastic conditions are employed, as in the treatment of lubricating fractions with large amounts of concentrated acid or when fuming acid is used in the manufacture of white oils, considerable quantities of petroleum sulfonic acids are formed [reaction (3)]. Two general methods are applied for the

$$\underset{\text{Paraffin}}{RH} + \underset{\text{Sulfuric acid}}{H_2SO_4} \longrightarrow \underset{\text{Sulfonic acid}}{RSO_3H} + \underset{\text{Water}}{H_2O} \qquad (3)$$

recovery of sulfonic acids from sulfonated oils and their sludge. In one case the acids are selectively removed by adsorbents or by solvents (generally low-molecular-weight alcohols); and in the other case the acids are obtained by salting-out with organic salts or bases.

Petroleum sulfonic acids may be roughly divided into those soluble in hydrocarbons and those soluble in water. Because of their color, hydrocarbon-soluble acids are referred to as mahogany acids, and the water-soluble acids are referred to as green acids. The composition of each type varies with the nature of the oil sulfonated and the concentration of the acids produced. In general, those formed during light acid treatment are water-soluble; oil-soluble acids result from more drastic sulfonation.

Sulfonic acids are used as detergents made by the sulfonation of alkylated benzene. The number, size, and structure of the alkyl side chains are important in determining the performance of the finished detergent. The salts of mixed petroleum sulfonic acids have many other commercial applications, as anticorrosion agents, leather softeners, and flotation agents; they have been used in place of red oil (sulfonated castor oil) in the textile industry. Lead salts of the acids have been employed in greases as extreme-pressure agents, and alkyl esters have been used as alkylating agents. The alkaline earth metal (magnesium, calcium, and barium) salts are used in detergent compositions for motor oils, and the alkali metal (potassium and sodium) salts are used as detergents in aqueous systems.

The sulfuric acid sludge from sulfuric acid treatment is frequently used as a source (through thermal decomposition) to produce sulfur dioxide (SO_2) that is returned to the sulfuric acid plant, and sludge acid coke. The coke, in the form of small pellets, is used as a substitute for charcoal in the manufacture of carbon disulfide. Sulfuric acid coke is different from other petroleum coke in that it is pyrophoric in air and also reacts directly with sulfur vapor to form carbon disulfide. *See* SULFURIC ACID. James G. Speight

Bibliography. American Society for Testing and Materials, *Annual Book of Standards*, 2000; G. W. Mushrush and J. G. Speight, *Petroleum Products: Instability and Incompatibility*, Taylor & Francis, New York, 1995; J. G. Speight, *The Chemistry and Technology of Petroleum*, 3d ed., Marcel Dekker, New York, 1999; J. G. Speight, *Fuel Science and Technology Handbook*, Marcel Dekker, New York, 1990; J. G. Speight (ed.), *Petroleum Chemistry and Refining*, Taylor & Francis, New York, 1998.

Petroleum reserves

Proved reserves are the estimated quantities of crude oil liquids which with reasonable certainty can be recovered in future years from delineated reservoirs under existing economic and operating conditions. Thus, estimates of crude oil reserves do not include synthetic liquids which at some time in the future may be produced by converting coal or oil shale, nor

do reserves include fluids which may be recovered following the future implementation of a supplementary or enhanced recovery scheme.

Indicated reserves are those quantities of petroleum which are believed to be recoverable by already implemented but unproved enhanced oil recovery processes or by the application of enhanced recovery processes to reservoirs similar to those in which such recovery processes have been proved to increase recovery.

Thus, crude oil reserves can be called upon in the future with a high degree of certainty, subject of course to the limitations placed on production rate by fluid flow within the reservoir and the capacity of the individual producing wells and surface facilities to handle the produced fluids. It is important to bear in the mind the distinction between resources and reserves. The former term refers to the total amount of oil that has been discovered in the subsurface, whereas the latter refers to the amount of oil that can be economically recovered in the future. The ratio of the ultimate recovery (the sum of currently proved reserves and past production) to the resource or original oil in place is the anticipated recovery efficiency. *See* PETROLEUM ENHANCED RECOVERY.

Levels. In earlier years crude oil reserves were estimated by first defining the volume of the resources from drilling data, the nature of the natural producing mechanism from the performance of the reservoir, particularly the rate of decline in productivity, and then applying a recovery factor based on analogy with similar reservoirs. Although more sophisticated technology is in use today, earlier rule-of-thumb estimates have proved to be surprisingly valid.

Reserves are increased by the discovery of new reservoirs, by additions to already discovered reservoirs by continued drilling, and by revisions due to a better-than-established anticipated performance or implementation of an enhanced recovery project. Reserves are decreased by production, and by negative revisions due to poorer-than-anticipated performance or less-than-projected reservoir volumes.

In the 15 years between 1954 and 1969, reserves of crude oil in the United States remained relatively stable. In 1969 the reserves jumped to a value of 39×10^9 bbl (6.2×10^9 m^3) as a result of the discovery of the gigantic Prudhoe Bay oil field on the North Slope of the Brooks Range in Alaska. From 1969 to the early 1990s, the crude oil reserves declined to less than 25×10^9 bbl (40×10^8 m^3) in the United States, which is the lowest level since 1950.

Discovery and production. In retrospect, it can be seen that reserves in the United States were sustained during the 1950s and early 1960s as a result of extensions and revisions, the latter due primarily to the installation of waterfloods, rather than as a result of significant new discoveries. The discovery at Prudhoe Bay indicated the necessity of exploration in new frontiers if new reserves were to be added to the United States total. Subsequent exploration failures in other frontier areas, such as the eastern Gulf of Mexico, the Gulf of Alaska, and the Baltimore Canyon, however, again emphasized the limited occurrence of crude oil in the Earth's crust.

The existence of an accumulation of petroleum fluids in the Earth's crust represents the fortuitous sequence of several natural events: the concentration of a large amount of organic material in ancient sediments, sufficient burial and thermal history to cause conversion to mobile fluids, and the migration and eventual confinement of the fluids in a subsurface geological trap. Although some oil and gas have been produced in 33 of the 50 states, 4 states account for 85% of the proved reserves: Texas and Alaska account for over 60%, with California and Louisiana ranking third and fourth. Further evidence for the spotty concentration of oil in the crust is gleaned from the fact that although there are some 10,000 producing oil fields in the United States, some 60 of them account for over 40% of the productive capacity.

In this context, it was to be anticipated that oil discovery and production would reach a peak followed by a monotonic decline in reserves and productive capacity. Only the date of peaking might be in doubt, but even this was predicted with a high degree of accuracy at least a decade before it occurred in the late 1960s in the United States. Additional oil discoveries will be made in some of America's remaining new frontiers, such as the deep ocean and in the vicinity of the Arctic Circle; however, most estimates of the amount of oil remaining to be found have been steadily decreasing, and it is highly unlikely that the decline will be substantially affected by future events.

Natural gas. Natural gas reserves are estimated in a manner similar to that used for the estimation of crude oil reserves. Some natural gas is found in association with crude oil in the same reservoir, and some is found in reservoirs which do not have a gas-oil contact, that is, nonassociated gas. Gas represents a higher maturation level than that of crude oil, and therefore some geographical areas are more gas-prone than others. The history of reserves of natural gas in the United States parallels that of crude oil, with a peak having been reached in the late 1960s. The ratio of gas reserves to oil reserves in the United States is approximately twice the ratio for the world. *See* NATURAL GAS; PETROLEUM. Todd M. Doscher

Bibliography. Worldwide report, *Oil Gas J.*, published in last issue of each year.

Petroleum reservoir engineering

The technology concerned with the prediction of the optimum economic recovery of oil or gas from hydrocarbon-bearing reservoirs. It is an eclectic technology requiring coordinated application of many disciplines: physics, chemistry, mathematics, geology, and chemical engineering. Originally, the role of reservoir engineering was exclusively that of counting oil and natural gas reserves. The reserves—the amount of oil or gas that can be economically recovered from the reservoir—are a measure of the wealth available to the owner and operator. It is also

necessary to know the reserves in order to make proper decisions concerning the viability of downstream pipeline, refining, and marketing facilities that will rely on the production as feedstocks.

The scope of reservoir engineering has broadened to include the analysis of optimum ways for recovering oil and natural gas, and the study and implementation of enhanced recovery techniques for increasing the recovery above that which can be expected from the use of conventional technology.

Original oil in place. The amount of oil in a reservoir can be estimated volumetrically or by material balance techniques. A reservoir is sampled only at the points at which wells penetrate it. By using logging techniques and core analysis, the porosity and net feet of pay (oil-saturated interval) and the average oil saturation for the interval can be estimated in the immediate vicinity of the well. The oil-saturated interval observed at one location is not identical to that at another because of the inherent heterogeneity of a sedimentary layer. It is therefore necessary to use statistical averaging techniques in order to define the average oil content of the reservoir (usually expressed in barrels per net acre-foot) and the average net pay. The areal extent of the reservoir is inferred from the extrapolation of geology and fluid content as well as the drilling of dry holes beyond the productive limits of the reservoir. The definition of reservoir boundaries can be heightened by study of seismic surveys, particularly 3-D surveys, and analysis of pressure buildups in wells after they have been brought on production. *See* PETROLEUM GEOLOGY; WELL LOGGING.

If the only mobile fluid in the reservoir is crude oil, and the pressure during the production of the crude remains above the bubble point value (at which point the gas begins to come out of solution), the production of crude oil will be due merely to fluid expansion, and the appropriate equation (1). Here c is

$$\text{Production} = N_P = (c \cdot V \cdot \Delta P)B_0 \qquad (1)$$

the compressibility of the fluid, V is the volume of crude oil in the reservoir, and ΔP is the decrease in average reservoir pressure associated with the production of N_P barrels of crude oil measured at the surface. The term in parentheses is the expansion measured in reservoir barrels, and B_0 is the formation volume factor (inverse of shrinkage) which corrects subsurface measurements to surface units. Thus, if N_P is plotted as a function of $(c \cdot \Delta P)B_0$, the slope of the resulting straight line will be V, the volume of the original oil in place in the reservoir.

For most reservoirs, the material balance equation is considerably more complicated. Few reservoirs would produce more than 1 or 2% of the oil in place if fluid expansion alone were relied upon. The reasons for the increased complexity of a material balance equation when the pressure falls below the bubble point are as follows: (1) The pore volume occupied by fluids in the reservoir changes with reservoir pressure as does the volume of the immobile (connate) water in the reservoir. (2) As the pressure declines below the bubble point, gas is released from the oil. (3) The volume of gas liberated is greater than the corresponding shrinkage in the oil, and therefore the liberated gas becomes mobile upon establishing a critical gas saturation, and the gas will flow and be produced in parallel with oil production. (4) There may be an initial gas cap above the oil column (above a gas-oil contact), and some of this gas, depending upon the nature of the well completion, may flow immediately upon opening the well to production. (5) Below the oil column, there may be an accumulation of water (below a water-oil contact), and if sufficiently large, the water in this aquifer will expand under the influence of the pressure drop and replace the oil that is produced from the oil column.

Thus the simple material balance equation (1) for fluid expansion must be expanded to take into account all the changes in fluid and spatial volumes and the production of oil and gas. By the suitable manipulation of the various terms in the resulting equation, the material balance can be reduced to the equation of a straight line with the slope and intercept of the line providing information about the magnitude of the original oil in place and the significance of the water encroachment and the magnitude of the gas cap. However, the accuracy of the material balance equation depends strongly on having obtained accurate knowledge about the *PVT* (pressure-volume-temperature) behavior of the reservoir crude, and knowing the average reservoir pressure corresponding to successive levels of reservoir depletion. There are significant limits to acquiring such information.

The material balance equation is not time-dependent; it relates only average reservoir pressure to production. Although it can be extrapolated to provide information on the amount of oil and gas that may be produced in the future when the pressure falls to a given level, it does not provide any information as to when the pressure will fall to such a level. It can be used to estimate the ultimate production from the reservoir if an economically limiting pressure can be specified. This usually can be done since pressure and rate of production are of course explicitly interrelated by the equations for fluid flow, and the economic limit of production is that rate at which the economic income is insufficient to pay for the cost of operation, royalties, taxes, overheads, and other facets of maintaining production.

Fluid flow in crude oil reservoirs. In order to develop some understanding of the future performance of a reservoir on a real-time basis, the reservoir engineer has two options. One is to predict the flow rate history of the reservoir by using the equations for fluid flow, and the second is to extrapolate the already known production history of the reservoir into the future by using empiricism and know-how generated by experience with analogous reservoirs.

The flow of fluids in the reservoir obeys the differential form of the Darcy equation, which for radial

flow is Eq. (2), where q is the volumetric flow rate

$$q = \frac{k \cdot h \cdot 2\pi r \left(\frac{\Delta P}{\Delta r}\right)}{\mu} \tag{2}$$

measured in subsurface volumes, r is the radial distance from the well bore to any point in the reservoir, $(\Delta P/\Delta r)$ is the pressure gradient at the corresponding value of the radius, μ is the viscosity of the mobile fluid, and k is the permeability of the reservoir to the mobile fluid.

This equation could be integrated directly if $\Delta P/\Delta r$ and the pressure at the outer limits of the reservoir were constant. However, this condition is true only where there exists a strong natural water drive resulting from a contiguous aquifer. Steady-state conditions then prevail and the flow equation is simply Eq. (3), where q is in barrels per day, k

$$q = \frac{7.08kh(P_e - P_w)}{\mu \ \ln\left(\frac{r_e}{r_w}\right)} \tag{3}$$

in darcies, P_e (pressure at the outer boundary) and P_w (pressure at the well bore) in pounds per square inch, μ in centipoises, and r in feet.

When steady-state conditions do not exist, the Darcy equation must be combined with the general radial diffusivity equation (4), where ϕ, the porosity,

$$\frac{\partial^2 P}{\partial r^2} + \frac{1}{r}\frac{\partial P}{\partial r} = \frac{\phi \mu c}{k}\frac{\partial P}{\partial t} \tag{4}$$

is the only term not elsewhere defined. Appropriate boundary conditions must be specified in order to obtain an integrated analytical equation for fluid flow in the reservoir. Solutions have been given for the important two sets of boundary conditions: constant production rate at the well bore, and constant production (well bore) pressure. Solutions are available for both the bounded reservoir in which the pressure declines from its initial value at the physical reservoir boundary, or in the case of multiple wells in a reservoir at the equivalent no-flow boundary between wells, and the infinite reservoir (usually only a transient condition in real reservoirs) in which the pressure has not declined below initial value at the boundary.

However, the use of the solutions to these equations still does not permit prediction of reservoir performance. The solutions are usually presented in terms of dimensionless time and dimensionless pressure or dimensionless cumulative production. The permeability to the mobile fluid and its compressibility enter into these dimensionless parameters, and neither of these functions is constant, nor can they be explicitly calculated. The reasons are readily traced.

It has already been noted that as the pressure in the reservoir is reduced, the gas saturation builds up. Accompanying this increase, there is an increase in the permeability to the mobile gas phase, and a decrease in the permeability to oil. The changes are

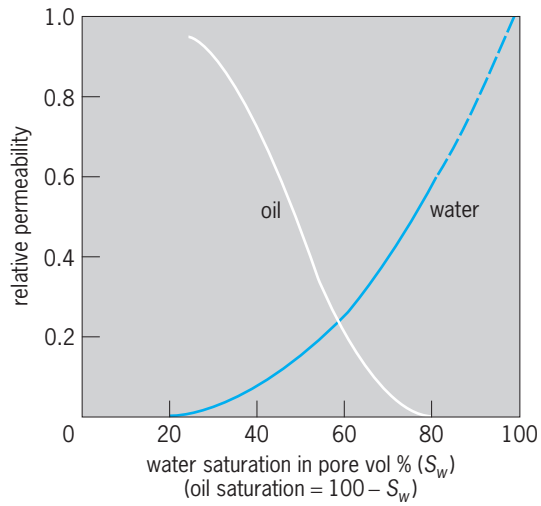

Fig. 1. Typical relative permeability relationship for oil and water in a porous medium.

not linear (see **Fig. 1**). Since the pressure is not uniform throughout the reservoir, neither will the gas saturation be uniform, nor will the corresponding permeabilities to the gas and oil. The compressibility of the reservoir fluids will obviously vary directly as the saturation of the compressible gas. Thus, the integrated forms of the radial diffusivity equation cannot be used with any high degree of accuracy over the entire reservoir. The modern high-speed digital computer has made it possible, however, to represent the entire reservoir by a three-dimensional grid system in which the reservoir is mathematically subdivided, and the fluid flow equations and material balance across each block are iteratively and compatibly solved for adjacent blocks. The use of mathematical simulation techniques requires a detailed specification of the reservoir geometry and lithology, and the relative permeabilities to the mobile fluids under conditions of both diffuse and stratified flow. Again, such knowledge is limited, and as a result it is necessary to calibrate the mathematical simulator by history-matching available production and pressure history for the reservoir. Unique predictions are difficult to achieve and as a result such simulations must be constantly updated. Pressure-transient analyses of individual wells are analyzed by superimposing the flow regimes for the flowing period and the shut-in period to yield data on effective permeability to the flowing fluids and the average reservoir pressure at the time of shutting in the well. Such data are important in calibrating the reservoir simulator for a given reservoir.

Before the high-speed digital computer was available to the reservoir engineer, and today when the data available for numerical simulation are not available or when the cost of such simulation studies cannot be justified, the reserves may be predicted by decline curve analysis. Depending on the production mechanism, the production rate will decline in keeping with exponential, hyperbolic, or

constant-percentage decline rates. The appropriate coefficients and exponents for a particular decline rate are estimated by history matching, and then the decline curves are extrapolated into the future. A variety of nomographs are available for facilitating the history matching and extrapolation.

Reservoir recovery efficiency. The overall recovery of crude oil from a reservoir is a function of the production mechanism, the reservoir and fluid parameters, and the implementation of supplementary recovery techniques. In the United States, recovery shows a geographical pattern because of the differences in subsurface geology in the various producing basins. Many reservoirs in southern Louisiana and eastern Texas, where strong natural water drives, low-viscosity crudes, and highly permeable, uniform reservoirs occur, show a recovery efficiency well over 50% of the original oil in place. Pressure maintenance, the injection of water to maintain the reservoir pressure, is widely used where the natural water influx is not sufficient to maintain the original pressure. In California, on the other hand, strong water drives are virtually absent, the crude oil is generally more viscous than in Louisiana, and the reservoirs, many of which are turbidites, show significant heterogeneity. Both the viscous crude oil and the heterogeneity permit the bypassing of the oil by less-viscous water and gas. Once a volume of oil is bypassed by water in water-wet rocks (the usual wettability), the oil becomes trapped by capillary forces, which then require the imposition of unfeasible pressure gradients or the attainment of very low interfacial tensions for the oil to be released and rendered mobile. As a result, the overall displacement in California reservoirs is relatively poor, and even after the implementation of water flooding, the recovery efficiency in California will probably be only 25–30%.

In general, recovery efficiency is not dependent upon the rate of production except for those reservoirs where gravity segregation is sufficient to permit segregation of the gas, oil, and water. Where gravity drainage is the producing mechanism, which occurs when the oil column in the reservoir is quite thick and the vertical permeability is high and a gas cap is initially present or is developed on producing, the reservoir will also show a significant effect of rate on the production efficiency.

The overall recovery efficiency in the United States is anticipated to be only 33% unless new technology can be developed to effectively increase this recovery. Reservoir engineering expertise, together with geological and petrophysical engineering expertise, is being used to make very detailed studies of the production performance of crude oil reservoirs in an effort to delineate the distribution of residual oil and gas in the reservoir, and to develop the necessary technology to enhance the recovery. Significant strides have been made in California, where the introduction of thermal recovery techniques has increased the recovery from many viscous oil reservoirs from values as low as a few percent to well over 50%. *See* PETROLEUM ENHANCED RECOVERY.

Todd M. Doscher

Well Testing

Well testing broadly refers to the diagnostic tests run on wells in petroleum reservoirs to determine well and reservoir properties. The most important well tests are called pressure transient tests and are conducted by changing the rate of a well in a prescribed way and recording the resulting change in pressure with time.

The information obtained from pressure transient tests includes estimates of (1) unaltered formation permeability to the fluid(s) produced in the well; (2) altered (usually reduced) permeability near the well caused by drilling and completion practices; (3) altered (increased) permeability near the well created by deliberately stimulating the well by injecting either an acid that dissolves some of the formation or a high-pressure fluid that creates fractures in the formation; (4) distances to flow barriers located in the area drained by the well; and (5) average pressure in the area drained by the well. In addition, some testing programs may confirm hypothesized models of the reservoir, including important variations of formation properties with distance or location of gas/oil, oil/water, or other fluid/fluid contacts.

Tests. Pressure transient tests may be broadly categorized as either single-well tests or multiwell tests. Common single-well tests include pressure buildup tests, pressure drawdown tests, injection tests, and injection falloff tests. In pressure buildup tests, a well is produced at a fixed or stabilized rate for a given period of time, following which the well is shut-in (well production is stopped), and the change in the downhole pressure is measured with time.

In drawdown or flow tests, a well is produced at specified rates, while the pressure downhole is measured as it declines with time. For injection wells, the preferred tests are injection tests and injection falloff tests. Injection tests are run by injecting into a well using a specified rate schedule (usually constant), and measuring the bottomhole pressure as it increases with time. Injection falloff tests are run by injecting a fluid into a well at fixed rate for a given period of time, then shutting-in the well and measuring the bottomhole pressure as it declines with time.

The major multiwell tests are interferences tests and pulse tests. An interference test involves producing (or injecting into) one well (called the active well) and measuring the pressure response in one or more offset wells, called observation wells. The size of the pressure change and the time at which it is large enough to measure are related to the permeability and porosity/compressibility product in the area between the active and observation wells.

Interference tests may indicate pressure changes at an observation well that may have been caused by reservoir activities at wells other than the active well. For that reason, pulse tests may be a more practical alternative for sending a pressure signal from an active well to observation wells. In a pulse test, the active well is produced (or injected into) for a few hours, shut-in, and returned to production at the same rate

for several cycles. In this way, a signal is sent from active wells to observation wells. This signal is usually clearly distinguishable from pressure responses to other activities in the reservoir.

Interpretation. Pressure transient tests are usually interpreted by comparing the observed pressure-time response to the predicted response by a mathematical model of the well/reservoir system. The method is an example of an inverse problem, in which an observed response from a system that is modeled mathematically, but with a model having parameters that are unknown—in this case, well and formation properties. The solution to this inverse problem is the unknown properties themselves. Often, however, (1) the mathematical model is usually highly idealized and (2) the solution is not unique—more than one set of formation properties may appear to be accurate.

Graphical techniques are used to calculate permeability. More sophisticated graphical techniques involve matching changes in pressure to preplotted analytical solutions (type-curve matching).

Regression analysis is used to match observed pressure-time data to mathematical models. Although analytical solutions are being found for more and more complex reservoir models each year, many reservoirs are still so complex that their behavior cannot be described accurately by analytical solutions. In such cases, finite-difference approximations to the governing flow equations can be used in commercial reservoir simulators, the reservoir properties treated as unknowns, and properties found that fit the observed data well.

W. John Lee

Special Core Analysis Using Computerized Tomography

Special core analysis is used to determine the spatial distribution of solids, voids, and fluids in a rock sample and to quantify the sample rock's ability to transport oil, gas, and water. The rock solid is classified by its lithology, composition, and density. The void space is characterized by its porosity, spatial distribution, and connectivity. The fluid saturation is described by the fraction of the void volume that it occupies. Multiphase fluid transport properties of a rock are described by absolute and relative permeabilities. Special core analysis results are used in reservoir prediction models to design and optimize the recovery processes.

X-rays, a form of short-wavelength radiation, are able to penetrate many materials. The attenuation of an x-ray beam as it passes through matter is controlled by a combination of Compton scattering and photoelectric effects, and is sensitive to the density and the apparent atomic number of the material. The ability of x-rays to pass through matter and yet be affected by it led to the development of computerized tomographic (CT) mathematical algorithms and instruments known today as CT scanners or imagers. CT technology has had a significant impact on special core analysis procedures. It provides a nondestructive technique for obtaining the spatial distribution of solids and fluids in a rock sample under static or dynamic conditions. *See* ATOMIC NUMBER; COMPTON EFFECT; COMPUTERIZED TOMOGRAPHY; X-RAYS.

Special core analysis uses CT to determine the distributions of densities, atomic numbers, lithologies, porosities, and fluids in reservoir rock samples. Density and apparent atomic number maps are obtained through dual-energy scanning of dry extracted cores. The same maps can be generated for saturated cores if some assumptions are made about porosity and fluid saturations. The lithology of the rock is obtained by density–atomic number cross-plots, similar to wellbore logging analysis. The structure of the rock is mapped by a nondestructive method using x-ray CT. A layered sample with a diameter of 5 cm is identified by a gray-scale display (**Fig. 2***a*). In this display, light shades denote high density and low porosity; dark shades denote low density and high porosity. The layers are parallel to each other, and the two dark layers at the left are about 3 mm thick. Fractures are also detected by x-ray CT (Fig. 2*b*). Open fractures without solid deposits appear as dark, low-density elongated features. The figure shows an oil-filled shale with a diameter of 10 cm. The fracture width is less than the resolution of the x-ray scanner, and hence the fracture appears wider than it actually is.

There are two primary methods for determining porosity. In the first method, dual-energy scanning is used to determine the overall density of the sample. The porosity distribution is determined using estimates of the densities of the rock matrix and the

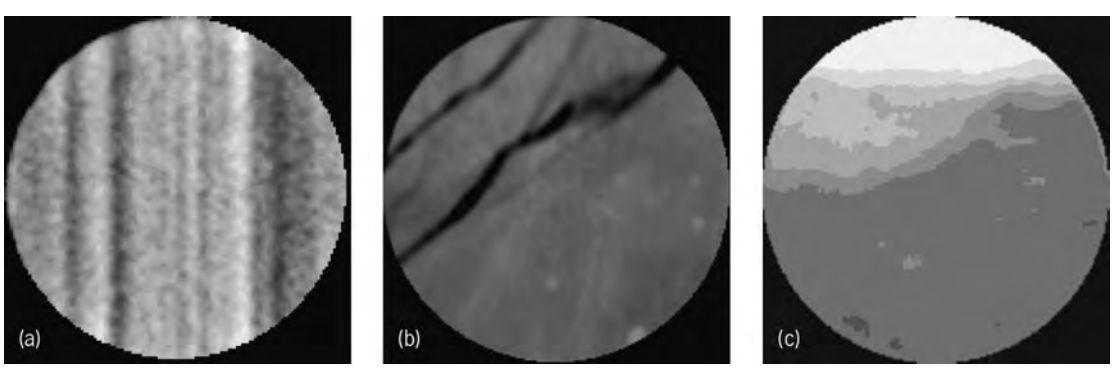

Fig. 2. Core analysis examples using x-ray computerized tomography for rock: (*a*) layers, (*b*) fractures, and (*c*) fluid phases.

fluids. In the second method, two high-energy scanning sequences, dry and saturated, are acquired. Subtracting the dry images from the saturated ones yields a map of the net fluid distribution in the sample. Quantitative porosity distribution is obtained either by combining the CT responses of the rock and the saturating fluid, or by calibrating the net fluid maps to external average porosity values.

X-ray CT provides the distribution of residual saturations in a rock sample and helps explain fluid displacement processes. It is also used to track fluid distributions in a sample during flow tests. In two-phase flow cases, one energy level, typically greater than 120 kV, is used. The fluids must have a strong x-ray attenuation contrast to be detected. The contrast between the fluids must be increased as the porosity of the sample is decreased. Elements such as iodine are used to artificially tag one of the fluids. Fluid saturations in a rock are tracked by careful calibrations and by computing CT attenuation changes. The independent determination of three-phase saturations using CT requires fluid tagging and dual-energy scanning at every saturation stage. It is a difficult process to perform without altering the intensive properties of the fluids. A net map of three fluid phases in a sample with a diameter of 5 cm is shown in Fig. 2*c*. The light-shaded region at the top denotes a gravity-driven low-density phase. This phase is tagged with iodine and has a high x-ray attenuation. The dark-shaded region in the lower half of the figure denotes an oil phase bordered at the top with a water-saturated region. This image was obtained during a dynamic displacement experiment with three nonmiscible phases. Abraham S. Grader

Petroleum Reservoir Models

Reservoir behavior can be simulated using models that have been constructed to have properties similar either to an ideal geometric shape of constant properties or to the shape and varying properties of a real (nonideal) oil or gas reservoir. *See* MODEL THEORY; SIMULATION.

In early reservoir modeling, analog electrical physical models were tested against the mathematical solutions that could be obtained for a single fluid by the method of conformal transformation for such ideal geometries. *See* CONFORMAL MAPPING.

For application to petroleum reservoirs, it is necessary to predict the simultaneous flow behavior of more than one fluid phase having different properties (water, gas, and crude oil). The permeability, the relative permeability, and the density and viscosity of each phase constitute its transport properties for calculating its flow. The relative permeability is a factor for each phase (oil, water, gas) which, when multiplied by the permeability for a single phase such as water, will give the permeability for the given phase. It varies with the volume fraction of the pore space occupied by the phase, called the saturation of the given phase. Generally, the relative permeability of the water phase depends only on its own saturation, and likewise for the gas phase. The relative permeability of the oil phase is a function of the saturations of both gas and water phases. *See* FLUID-FLOW PRINCIPLES; FLUID MECHANICS.

For mathematical predictions, the flow is equal to a coefficient times a driving force divided by a resistance. For flow through porous, permeable rocks, the relationship is called Darcy's law [Eq. (2)], in which the coefficient is called the permeability, the driving force is the cross-sectional area times a combination of the gravity force and the pressure gradient per unit area, and the resistance is the fluid viscosity. The equation is written for each phase present. For use in three-dimensional reservoir models, the equations are converted into partial differential equations for each phase (with the three space dimensions and time as independent variables), and then into numerical finite difference equations to allow for finite steps in each dimension and in time. The reservoir is subdivided into rectilinear blocks for the spatial steps, and the porosity, permeability, permeability for each phase, saturation of each phase, temperature, and pressure are specified initially for each grid block of the reservoir model. It is then possible to compute flows under the driving forces between blocks over the entire volume (including blocks specified to contain either injection wells or production wells, where fluids are injected or withdrawn), over a time step small enough not to incur mass-balance errors larger than a specified maximum. The computations must include changes of given components, such as natural gas between liquid phase and gas phase, and the changes in transport properties of the phases with temperature, pressure, and saturation. By repeating such calculations using computers, the process can be carried from an initial state to a desired end point, usually dictated by an economic minimum oil rate or maximum water-oil or gas-oil ratio at producing wells. *See* FINITE ELEMENT METHOD.

Using three-dimensional grids, the number of grid blocks may easily become quite large, approaching a million in some cases, and thus computer reservoir simulation can tax the abilities of the highest capability of computer available. For mega-cell simulation studies, parallel processing is used rather than conventional sequential processing. If the reservoir architecture is modeled using geostatistical variances, it will be necessary to scale the multi-million-cell model to a size with, at most, a few hundred thousand cells. If the crude-oil, gas- and water-phase chemical compositions are initially given and allowed to vary by specified processes as the computation continues, further demands are placed on computer capabilities, and certain compromises are usually required.

For short-term predictions of oil production rates where the rate is declining with time, decline curve analysis issued to give a suitable exponent between rate and time for the available data, so that further prediction to a given end point can be made.

An experimental prediction can also be obtained by constructing a laboratory model in which the geometric shape is similar to that of a given oil or gas reservoir. The injection rate and fluid properties

are modified from those of the actual reservoir so that given combinations of the fluid properties, pressure differences, and flow rates can be made by expressing ratios of forces or flow resistances, and the dimensions of each combination cancel out (giving dimensionless numbers). The design method of such models is called dimensionless scaling. Dimensionless numbers are generally found by examining the groups which occur when the partial differential equations for the system involved are cast into dimensionless form. One of the earliest users of dimensionless modeling was Osborne Reynolds, whose experimental work on fluid flow in pipes led to the use of the Reynolds number [Re; Eq. (5)]. He observed

$$\mathrm{Re} = \frac{\rho V L}{\mu} \tag{5}$$

that streamline or turbulent regions of fluid flow in pipes could be determined simply by the value of the dimensionless number formed by the pipe diameter (L) times the flow velocity (V) times the fluid density (ρ), divided by the fluid viscosity (μ). It has been found that laboratory models will exhibit appropriate flow behavior when given properties are changed from those in an oil reservoir but certain important dimensionless groups are made the same for both model and reservoir. The important set of dimensionless groups vary for different reservoir processes. *See* DIMENSIONAL ANALYSIS; DIMENSIONLESS GROUPS; REYNOLDS NUMBER.

The compilation of the reservoir data for either computer modeling or experimental models is a major task. Expert geophysicists, geologists, petrophysicists, logging experts, and reservoir engineers are required. In order to ameliorate the need for experts, computer-based expert systems have been developed for processing data. One method, called neural networks, involves computing methods in which weighting factors are applied to given parts of the data used to compute a given result. A set of data, which should give a certain result, is fed to the system. The weighting factors are varied until the computation gives the known answer. This is done for several different learning sets, and adjusted as needed to give the best set of answers. Then sets are tried for which the answers are withheld until the computations are made. If the neural network still gives the correct answers, it is concluded that the network has learned the proper weighting and may be used when the correct answer is unknown. *See* NEURAL NETWORK. Elmond L. Claridge

Bibliography. J. K. Ali, *Neural Networks: A New Tool for the Petroleum Industry?*, Soc. Petrol. Eng. Pap., no. 27561, presented at the European Petroleum Computer Conference, Aberdeen, March 1994; J. J. Arps, Analysis of decline curves, *Trans. AIME*, 160:219–227, 1945; L. E. Baker, *Three-Phase Relative Permeability Correlations*, Soc. Petrol. Eng. Pap., no. 17639, presented at the SPE/DOE Enhanced Oil Recovery Symposium, Tulsa, April 1988; H. B. Crichlow, *Modern Reservoir Engineering: A Simulation Approach*, Prentice Hall, Englewood Cliffs, NJ, 1977; R. C. Earlougher, Jr., *Advances in Well Test Analysis*, SPE Monogr. Ser., no. 5, 1977; A. S. Emanuel et al., *Reservoir Performance Prediction Methods Based on Fractal Statistics*, Soc. Petrol. Eng. Res. Eng., p. 311, August 1989; R. N. Horne, *Modern Well Test Analysis*, 2d ed., Petroway, Palo Alto, CA, 1995; M. K. Hubbert, *The Physical Basis of Darcy's Law and Its Importance in the Exploration and Production of Petroleum*, Soc. Petrol. Eng. Pap., no. 639, 1963; P. K. Hunt, P. Engler, and C. Bajsarowicz, Computed tomography as a core analysis tool: Applications, instrument evaluation, and image improvement techniques, *SPE Formation Eval.*, pp. 1203–1210, September 1988; R. E. Johnstone and M. W. Thring, *Pilot Plants, Models, and Scale-Up Methods in Chemical Engineering*, McGraw-Hill, New York, 1957; W. J. Lee, *Well Testing*, SPE Textbook Ser., no. 1, 1981; C. S. Matthews and D. G. Russell, *Pressure Buildup and Flow Tests in Wells*, SPE Monogr. Ser., no. 1, 1967; M. Muskat, *Physical Principles of Oil Production*, McGraw-Hill, New York, 1949; S. Siddiqui, P. J. Hicks, Jr., and A. S. Grader, Verification of Buckley-Leverett three-phase theory using computerized tomography, *J. Petrol. Sci. Eng.*, 15:1–21, 1996; H. J. Vinegar and S. L. Wellington, Tomographic imaging of three-phase flow experiments, *Rev. Sci. Instrum.*, pp. 96–107, January 1987; S. Y. Wang, S. Ayral, and C. C. Gryte, Computer-assisted tomography for the observation of oil displacement in porous media, *Soc. Petrol. Eng. J.*, pp. 53–55, February 1984.

Petrology

The study of rocks, their occurrence, composition, and origin. Petrography is concerned primarily with the detailed description and classification of rocks, whereas petrology deals primarily with rock formation, or petrogenesis. Experimental petrology reproduces in the laboratory the conditions of high pressure and temperature which existed at various depths in the Earth where minerals and rocks were formed. A petrological description includes definition of the unit in which the rock occurs, its attitude and structure, its mineralogy and chemical composition, and conclusions regarding its origin. For a discussion of mineral identification, petrographic analysis, and the classification of rocks *see* MINERALOGY; PETROGRAPHY; ROCK.

Igneous rocks. Volcanic rocks are igneous rocks that reach the Earth's surface before solidification. They occur as lavas or as pyroclastic (fragmental) rocks.

Volcanic activity varies greatly in intensity, duration, periods between eruptions, and quantities of gases, liquid rock, and solidified fragments expelled. The important factors influencing these differences are chemical composition of the magma; amount of gas dissolved in it; extent of crystallization or cooling before eruption; and configuration of the conduit and depth to the magma chamber. *See* MAGMA.

Volcanic structures are of a variety of types

TABLE 1. Types of volcanic structure

Name	Characteristics
Shield	Low height, broad area; formed by successive fluid flows accumulating around a single, central vent
Cinder cone	Cone of moderate size with apex truncated; circular in plan, gently sloping slides; composed of pyroclastic particles, usually poorly consolidated
Spatter cone	Small steep-sided cone with well-defined crater composed of pyroclastic particles, well consolidated (agglomerate)
Composite cone	Composed of interlayered flows and pyroclastics; flows from sides (flank flows) common, as are radial dike swarms; slightly concave in profile, with central crater
Caldera	Basins of great size but relatively shallow; formed by explosive decapitation of stratocones, collapse into underlying magma chamber, or both
Plug dome	Domal piles of viscous (usually rhyolitic) lava, growing by subsurface accretion and accompanied by outer fragmentation
Cryptovolcanic structures	Circular areas of highly fractured rocks in regions generally free of other structural disturbances; believed to have formed either by subsurface explosions or sinking of cylindrical rock masses over magma chambers

(**Table 1**), influenced by the style of activity of the volcano. Volcanic rocks are presently forming mostly at plate boundaries: at the midocean ridges where voluminous basalts are produced and at convergent plate boundaries (island arcs and continental margins) where most explosive and more silicic volcanism occurs. *See* VOLCANO.

Intrusive igneous rocks occur in many different types of units or intrusive masses, which are classified chiefly by their shape and structural relations to their wall rocks (**Table 2**). Bodies that crystallize at great depths (such as batholiths) are referred to as plutonic; those consolidated under shallow cover are designated as hypabyssal. *See* PLUTON.

The crystallization of the larger intrusives may result in profound alterations in the adjacent wall rocks (exomorphism). Where stocks and batholiths have invaded sedimentary rocks, an aureole of contact metamorphism is developed. This results from recrystallization under increased temperature and may be accompanied by chemical transformations (pyrometasomatism) produced by hydrothermal solutions generated during the latter stages of magmatic differentiation. Where batholiths have been intruded into rocks which are already regionally metamorphosed, the contact rocks formed are injection gneisses or migmatites. *See* CONTACT AUREOLE.

Igneous rocks make room for themselves by forceful injection (dilatance), by engulfing wall rock blocks (magmatic stoping), or by subsidence of overlying rocks. The hypothesis of granitization maintains that granites result from the wholesale transformation of sedimentary or metamorphic rock layers by solutions operating through mineral replacement or by ionic emanations acting through solid diffusion.

Blocks of wall rock included in an intrusive mass are xenoliths; their partial destruction by reaction may produce irregular clumps of mafic minerals called schlieren. In some instances such endomorphic effects are sufficiently intensive to result in modification of the composition of the magma (syntexis). *See* XENOLITH.

Crystallizing under equilibrium conditions, early magmatic minerals react with remaining fluid to yield new species (**Fig. 1**). Interruption of the sequence will yield liquid fractions richer in silicon dioxide, alkalis, iron, and water than the original magma, and crystalline fractions richer in calcium and magnesium than the parent magma (magmatic differentiation).

Igneous rocks occur in distinct associations (**Table 3**) which can be put in a specific tectonic

TABLE 2. Characteristics of intrusive igneous rock masses

Name	Shape	Structural relations to wall rocks	Size and other features
Dikes	Tabular, lensoid	Discordant	Few feet to hundreds of miles long
Sills	Tabular, lensoid	Concordant	Up to several hundred feet thick
Laccoliths	Plano-convex or doubly convex lenses	Generally concordant	1–4 mi (1.5–6.5 km) in diameter; several thousand feet thick
Volcanic necks	Pipelike	Discordant	Few hundred feet to a mile in diameter; cores of eroded volcanoes
Stocks	Irregular, with steep walls	Crosscutting	A small batholith or its upward projection; outcrop area less than 40 mi^2 (104 km^2)
Batholiths	Irregular, contacts dip steeply or outward; no bottoms known	(1) Discordant (2) Concordant in general, may be crosscutting in detail	Some cover 16,000 mi^2 (41,000 km^2); some composite intrusives of varied petrology
Plutons	Irregular	Usually crosscutting	Usually large; used as general name for intrusive masses that do not fit other definitions

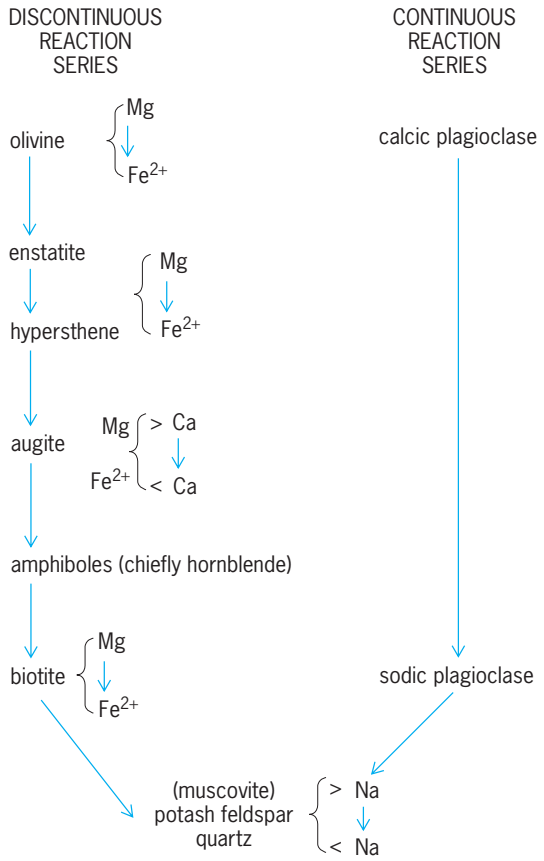

Fig. 1. Reaction series of Bowen (modified).

context. The sources of magma are the mantle and lower crust of the Earth. The diversity of magmas is caused by variations in the source rock, variations in the conditions and depth of origin, differentiation at various depths, assimilation of other rocks by the magma, mixing of two magmas which originated separately, and unmixing (immiscibility) of the melt. A scheme which shows the possible complexity of the origin of common igneous rock associations is given in **Table 4**. *See* IGNEOUS ROCKS.

Sedimentary rocks. Sedimentary rocks are broadly divided into two classes: clastic sediments, such as sandstones and conglomerates, which are composed largely of fragments of preexisting rocks and minerals and chemical sediments, such as evaporites and many limestones, which form as chemical precipitates from oceans or lakes. With the exception of material deposited by glaciers (till or the consolidated form tillite), sedimentary rocks show bedding or stratification. This separation into generally parallel layers (beds, strata) results from sorting according to grain size during deposition, from differences in composition or texture, or from variations in the rate of deposition. The development of most clastic sediments proceeds in the following stages: There is a source rock, any older rock or, for organic sediments, a supply of organically originated material. By weathering, the older rock is mechanically comminuted, chemically altered, or both, to form unconsolidated surficial rock debris called mantle. Particles are transported by streams, ocean and lake currents, wind, glaciers, or by the direct action of gravity which causes particles to slide and roll down slopes. Material moved by rolling, suspension, or solution is deposited. Deposits usually are consolidated by the processes of cementation (sandstones), compaction (shales), and recrystallization (limestones).

Chemical changes accompanying consolidation are termed diagenetic. Weathered material not transported may become a residual sedimentary rock (bauxite). Sedimentary rocks are deposited either on land areas (continental) or in ocean waters (marine). Most marine sedimentation takes place on the submarine extensions of the continents called continental shelves. Examples of types of sedimentary deposits are listed in **Table 5**. Features characteristically found in sedimentary rocks, in addition to stratification, are cross-bedding, concretions, ripple marks, mud cracks, and fossils. *See* DIAGENESIS; SEDIMENTOLOGY; WEATHERING PROCESSES.

A formation, which is the basic unit of stratigraphy, is a series of rocks deposited during a specific unit

TABLE 3. Igneous rock associations

Name	Main rock types	Environment
Oceanic basalts	Tholeiitic basalt	Mid-ocean ridges, much of the ocean floor
Spilites	Spilite, keratophyre, pillow basalt	Edges of island arcs; obducted ocean floor
Alkali basalts	Olivine basalt, trachyte, phonolite	Oceanic islands, atectonic continental sites
K-rich basalts	Leucite basalt, trachyte	Isolated atectonic continental sites
Tholeiitic flood basalts	Basalt (generally olivine-free), diabase	Continental flood basalt regions, associated with rifting
Layered mafic intrusives	Gabbro, norite, anorthosite, peridotite	Continental rifting sites
Calc-alkalic volcanic rocks	Basalt, andesite, rhyolite	Convergent plate margins, island arc and continental margin
Alpine peridotites	Peridotite, serpentinite	Convergent plate boundaries
Precambrian anorthosite	Andesine or labradorite anorthosite, norite, syenite, monzonite	Domed pluton of massifs in Precambrian terrains
Granite batholith	(a) Simple: granite, granodiorite (b) Complex: gabbro, tonalite, granodiorite, minor granite	Precambrian shields; cores of mountain ranges
Minor granitic intrusive	Granite (some alkalic), quartz syenite, syenite, diorite	Hypabyssal, in mountain ranges and as their outliers
Nepheline syenite	Feldspathoidal rocks, carbonatites	(a) Simple plutons (b) Ring complexes
Lamprophyre	Minette, kersantite, camptonite	Dike swarms

TABLE 4. Synopsis of magmatic evolution*

Primary magmas	Mode of origin	Common types of igneous rocks		
Alkaline olivine basalt magmas	By differentiation and plutonic crystallization	Teschenites Theralites Essexites Shonkinites Pyroxenites Picrites Peridotites		Syenites Nepheline syenites Ijolites → Carbonatites
	By differentiation and volcanic crystallization	Mugearites Trachybasalts Olivine basalts Picrite basalts		Alkali rhyolites ↑ Trachytes ↓ Phonolites
		Nepheline basalts Soda lamprophyres Melilite basalts		Phonolites
	By contamination involving reaction with "granite" of continental basement	Leucite basalts Biotite lamprophyres		Leucite phonolites Latites Trachytes
	By differentiation and modification in geosynclines	Keratophyres Spilites		
Granodiorite granite magmas	By differentiation and plutonic crystallization	Granites Granodiorites Tonalites Diorites Syenites		Pegmatites Aplites
	By differentiation and volcanic crystallization	Rhyolites Dacites Andesites Latites		
	By mixing with basaltic magma	Andesites		Dacites Rhyolites
	By assimilative reaction with slates, limestones amphibolites, etc.	Syenites Nepheline syenites		
		Diorites Tonalites		
Tholeiitic magmas	By differentiation and plutonic crystallization	Granophyres Diorites Anorthosites (bytownite) Gabbros Pyroxenites Peridotites		
	By differentiation in thick sills	Granophyres Quartz diabases Diabases Olivine diabases		
	By differentiation and volcanic crystallization	Tholeiitic basalts Picrite basalts	?	Rhyolites Andesites

*After F. J. Turner and J. Verhoogen, *Igneous and Metamorphic Petrology,* 2d ed., McGraw-Hill, 1969.

of geologic time and consisting either of a particular rock type or of several types deposited in a sedimentary cycle. Such a cycle is the changing sequence of deposits reflecting, for example, advance or retreat of marine waters in a particular area.

However, while sandstone may be deposited at one time in one place in the sedimentary basin, limestone may be formed simultaneously elsewhere. Such lateral variation in a formation is referred to as facies. *See* CYCLOTHEM; FACIES (GEOLOGY); STRATIGRAPHY.

By means of detailed studies of the fossils of a formation and its lithology, composition, structure, and distribution, the paleoecology of the area may be reconstructed. Correlation of formations is attempted chiefly on the basis of fossils, with supplementary data from the lithology, stratigraphic position, insoluble residues (in acid-soluble rocks), heavy detrital minerals (in clastic rocks) and in drill holes by electrical conductivity, radioactivity, and seismic-wave velocities. *See* SEDIMENTARY ROCKS.

TABLE 5. Selected examples of sedimentary deposits under various environments

Agent	Deposit	Resulting rock
Continental		
Streams	Valley fill	Sandstone
	Alluvial fan	Conglomerate
	Delta	Siltstone
Lakes	Varved clay	Shale
Springs		Travertine
		Siliceous sinter
Swamps	Peat	Coal
Wind	Dune	Sandstone
	Dust	Loess
	Volcanic ash	Tuff
Glaciers	Moraine	Tillite
Ground water	Stalactite	Dripstone
Gravity	Talus	Breccia
	Avalanche	Conglomerate
	Landslide	
Marine		
Breakers and alongshore currents	Beach	Sandstone
		Conglomerate
Longshore currents		Sandstone
		Shale
Marine organisms	Reefs and other shell deposits	Shell limestone
		Coquina
		Diatomite
Marine water	Evaporites	Rock salt
		Rock anhydrite
Marine water	Colloidal precipitates	Phosphorite
		Manganese oxide concretions
		Chert

Metamorphic rocks. Metamorphism transforms rocks through combinations of the factors of heat, hydrostatic pressure (load), stress (directed pressure), and solutions. Most of the changes are in texture or mineral composition; major changes in chemical composition are called metasomatism. The major types of metamorphism are presented in **Table 6**. Rocks that can serve as parent material for metamorphic derivatives include igneous, sedimentary, and older metamorphic rocks as well. The complexity of the possible metamorphic mineral assemblages stems not only from the variety of possible parent rocks and from the imposition of the several kinds of metamorphism but also from variation in the intensity of particular types of metamorphism (grade), and from the difficulty of readily achieving chemical equilibrium through solid-state reactions. Various features characteristic of metamorphic rocks include foliation (slaty cleavage, schistosity, and gneissic structure), lineation, banding, and relict structures. *See* METAMORPHISM; METASOMATISM.

TABLE 6. Types of metamorphism and their factors

Type	Factors	Changes in rock
Cataclastic	Stress, low hydrostatic pressure	Fragmentation, granulation
Contact (thermal)	Heat, low to moderate hydrostatic pressure	Recrystallization to new minerals or coarser grains; rarely melting
Pyrometa-somatism	Heat, additive hydrothermal solutions, low to moderate hydrostatic pressure	Reconstitution to new minerals; change in rock composition
Regional (dynamic)	Heat, weak to strong stress, moderate to high hydrostatic pressure, ±nonadditive solutions	Recrystallization to new minerals or coarser grains; parallel orientation of minerals to produce foliation

The facies principle is employed in attempting to reconstruct the environment under which a metamorphic rock was developed. A metamorphic facies consists of all rocks, without respect to chemical composition, that have been recrystallized under equilibrium, within a particular environment of stress, temperature, load, and solutions. The first two factors are considered critical. The facies are named after metamorphic rocks deemed diagnostic of such restricted conditions. In practice, a group of related rocks of different compositions is assigned to a particular facies upon presence of such a key assemblage. Facies and their type descriptions are as follows:

I. *Facies of contact metamorphism*. Load pressure low, generally 100–3000 bars (10–300 megapascals). Water pressure highly variable, in some cases possibly exceeding load pressure, in a few cases very low. Facies listed in order of increasing temperature for given range of pressure conditions.

1. Albite-epidote hornfels.
2. Hornblende hornfels.
3. Pyroxene hornfels.
4. Sanidinite—corresponds to minimum pressures (load, P_{H_2O}, P_{CO_2}) and maximum temperatures—pyrometamorphism.

II. *Facies of regional metamorphism*. Load and water pressures generally equal and high (3000–12,000 bars or 300–1200 MPa). Facies listed in order of increasing temperature and pressure.

1. Zeolite (very low grade).
2. Prehnite-pumpellyite (very low grade).
3. Greenschist (low grade).
   a. Quartz-albite-muscovite-chlorite.
   b. Quartz-albite-epidote-biotite.
   c. Quartz-albite-epidote-almandine.
4. Glaucophane schist (represents a divergent line of metamorphism conditioned by development of unusually high pressures at low temperatures).
5. Almandine amphibolite (medium to high grade).
   a. Staurolite-quartz.
   b. Kyanite-muscovite-quartz.
   c. Sillimanite-almandine.
6. Granulite (high grade).
   a. Hornblende granulite.
   b. Pyroxene granulite.
7. Eclogite (very high pressure).

Because experimental petrology has allowed the development of a "petrogenetic grid" of mineral

reactions, most facies may be subdivided if mineral assemblages are carefully detailed. This subdivision can allow inference of pressure-temperature conditions of metamorphism.

Regional variations in grade may be mapped by means of isograds, lines formed by the intersection of planes of isometamorphic intensity with the Earth's surface. These are defined on the appearance of a specific mineral known to reflect a major increase in the intensity of metamorphism.

The primary cause of stresses acting during regional metamorphism is diastrophism of the mountain-building type. These stresses may precede, accompany, or follow the heating, and the order of thermal and structural events can be unraveled by petrographic study. Most regional metamorphic rocks record several heating and deformation episodes. The higher temperatures may result from deep burial, owing to the geothermal gradient of the Earth, in part to concentrations of radiogenic heat, or in part to heat supplied by cooling masses of magma. In contact metamorphism this last is the sole heat source. *See* GEOLOGIC THERMOMETRY.

William Ingersoll Rose, Jr.

Experimental mineralogy and petrology. One aim of mineralogy and petrology is to decipher the history of igneous and metamorphic rocks. Detailed study of the field geology, the structures, the petrography, the mineralogy, and the geochemistry of the rocks is used as a basis for hypotheses of origin. The conditions at depth within the Earth's crust and mantle, the processes occurring at depth, and the whole history of rocks once deeply buried are deduced from the study of rocks now exposed at the Earth's surface. One approach used to test hypotheses so developed is experimental petrology; the term experimental minerals refers to similar studies involving minerals rather than rocks (mineral aggregates).

The experimental petrologist reproduces in the laboratory the conditions of high pressure and high temperature encountered at various depths within the Earth's crust and mantle where the minerals and rocks were formed. By suitable selection of materials the petrologist studies the chemical reactions that actually occur under these conditions and attempts to relate these to the processes involved in petrogenesis.

The experiments may deal with the stability range of minerals and rocks in terms of pressure and temperature; with the conditions of melting of minerals and rocks; or with the physical properties or physical chemistry of minerals, rocks, and rock melts (magmas), or of the vapors, gases, and solutions coexisting with the solid or molten materials. These experiments may thus be related to major geological processes involved in the evolution of the Earth: the conditions of formation of magmas in the mantle and crust and their subsequent crystallization either as intrusions or lava flows; the evolvement of gases by the magmas during their crystallization and the precipitation of ore deposits from some of them; the processes leading to the development of volcanic arcs and mountain ranges; and the metamorphism and deformation of the rocks in the mountain chains, and thus the origin and development of the continents. Representative experimental results are shown in **Fig. 2**.

Experimental probes into the Earth. The pioneer work of James Hall of Edinburgh on rocks and minerals sealed within gun barrels at high pressures and temperatures earned him the title "father of experimental petrology." His experiments, published early in the nineteenth century, heralded an era of laboratory synthesis in mineralogy and petrology. Experimental petrology gained tremendous impetus in 1904, when the Geophysical Laboratory was founded in Washington, D.C., and new techniques were developed for the study of silicate melts at carefully controlled temperatures. The results obtained by N. L. Bowen, J. F. Schairer, and others have persuaded most petrologists that physicochemical principles can be successfully applied to processes as complex as those occurring within the Earth, and they have formed the basis for much of the work performed at high temperatures and high pressures. The design of equipment capable of maintaining high temperatures simultaneously with high pressures has been successfully achieved only since about 1950, with a few notable exceptions. The extent of the experimental probe into the Earth varies with the type of apparatus used. The Tuttle cold-seal pressure vessel, or high-pressure test tube, reproduces conditions at the base of the average continental crust. Internally heated, hydrostatic pressure vessels reproduce conditions similar to those just within the upper mantle. Opposed-anvil devices, or simple squeezers, readily provide pressures of 200 kilobars or more, but temperatures attainable are considerably lower than those existing at depths in the Earth corresponding to these pressures.

The deepest experimental probe is provided by a variety of internally heated opposed-anvil devices which can produce pressures of 200 kbar (20 GPa) with simultaneous temperatures of 2000°C (3630°F) or more, although the pressure and temperature measurements become less accurate with the more extreme conditions. This deepest probe produces conditions equivalent to depths of 180–240 mi (300–400 km) in the upper mantle, which are not very deep compared with the 1740-mi (2900-km) depth at the core, but are well within the Earth's outer 300 mi (500 km) where many petrological processes have their origins. *See* HIGH-PRESSURE MINERAL SYNTHESIS.

Studies related to Earth structure. Geophysical studies and inferences about the compositions of deep-seated rocks based on petrological studies of rocks now exposed at the surface, together with laboratory measurements of physical properties of selected materials, provide the basis for theories concerning the composition of the Earth and those parts of the Earth which may be treated as units. *See* EARTH INTERIOR.

At high pressures, most silicate minerals undergo phase transitions to denser polymorphs, and it is believed that polymorphic transitions within the mantle contribute to the high gradients of seismic-wave

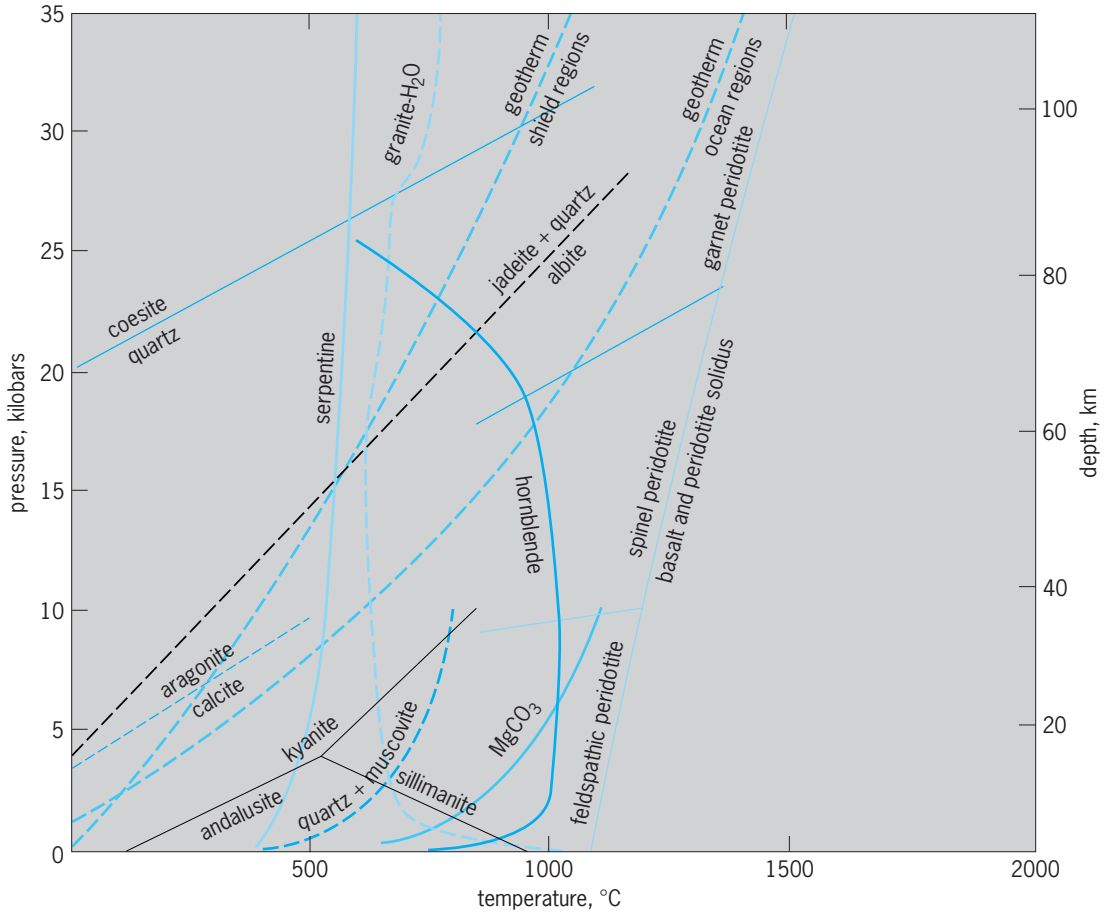

Fig. 2. Selected results in experimental mineralogy and petrology, averaged from many sources. The two geotherms show the temperature distribution with depth in the Earth. The solidus for dry basalt and peridotite provides the upper temperature limit for magma generation to occur in the mantle. The mineral facies of a peridotite upper mantle are shown below the solidus. The lower limit for magma generation is given by the curve granite-H_2O. Solid-solid mineral transitions and reactions plotted are quartz-coesite, the breakdown of albite, calcite-aragonite, and the polymorphic transitions among kyanite-andalusite-sillimanite. A decarbonation reaction is shown by the curve for $MgCO_3$ (magnesite), and dehydration reactions are shown for serpentine, for the assemblage muscovite + quartz, and for hornblende in mafic or ultramafic rocks. Note the effect of pressure on the hornblende reaction in the garnet-peridotite facies. 1 kilobar = 10^2 megapascals; °F = (°C × 1.8) + 32.

velocities occurring at certain depths. Experimental studies of the olivine-spinel transition have been successfully correlated with the beginning of the transition zone of the upper mantle at a depth of about 210 mi (350 km). It has been widely held that the Mohorovičić (M) discontinuity marks a change in chemical composition from basalt of the lower crust to peridotite of the upper mantle, but recent experimental confirmation that basalt undergoes a phase transition to eclogite, its dense chemical equivalent, suggests that the M discontinuity could be a phase transition. This now appears to be unlikely for several reasons, but more detailed experimental studies are required because there remains a strong possibility that this phase transition is involved in tectonically active regions of high heat flow where the M discontinuity is poorly defined. *See* MOHO (MOHOROVIČIĆ DISCONTINUITY).

Igneous petrology. The Earth's mantle transmits shear waves and is therefore considered to be crystalline rather than liquid; but the eruption of volcanoes at the Earth's surface confirms that melting temperatures are reached at depth in the Earth from time to time and from place to place. One group of experiments is therefore concerned with the determination of fusion curves for minerals believed to exist in the Earth's mantle and of melting intervals for mineral aggregates, or rocks. The position of the solidus curve provides an upper temperature limit for the normal geotherm, and the experiments dealing with melting intervals at various pressures have direct bearing on the generation of basaltic magmas at various depths within the mantle. They also provide insight into the processes of crystallization and fractionation of these magmas during their ascent to the surface. This is one of the major problems of igneous petrology. *See* IGNEOUS ROCKS; MAGMA.

Many reactions among silicate minerals are very sluggish, and the presence of water vapor under pressure facilitates the experiments. Water is the most abundant volatile component of the Earth, with carbon dioxide a poor second, and the water in the experiments often plays a role as a component as well as a catalyst. In the presence of water vapor under pressure, melting temperatures of silicate minerals and rocks are markedly depressed, and curves for the

beginning of the melting of rocks in the presence of water provide lower limits for the generation of magmas. It is widely believed that many of the granitic rocks constituting batholiths, the cores of mountain ranges, were derived by processes of partial fusion of crustal rocks in the presence of water. The hypothesis appears to be consistent with experimental results on feldspars and quartz in the presence of water at pressures up to 10 kbar (1 GPa) and with similar experiments on plutonic igneous rock series ranging in composition from gabbro to granite. However, evidence suggests that some magmas forming batholiths, and andesite lavas of equivalent composition in tectonically active regions, originated directly from mantle material. Experimental petrologists are therefore extending their studies of the minerals and rocks to granitic and andesitic compositions to pressures greater than 10 kbar (1 GPa), corresponding to upper mantle conditions, with and without water present. *See* SILICATE PHASE EQUILIBRIA.

The path of crystallization of a magma and the mineral reaction series thus produced are apparently quite sensitive to many variables, including pressure, pressure or fugacity of water, and oxygen fugacity. Techniques have been developed for studying the effect of oxygen fugacity on crystallization paths and also for controlling oxygen fugacity at very low values while the water pressure is simultaneously maintained at very high values.

Metamorphic petrology. When rocks are metamorphosed, they recrystallize in response to changes in pressure, temperature, stress, and the passage of solutions through the rock. The facies classification is an attempt to group together rocks that have been subjected to similar pressures and temperatures on the basis of their mineral parageneses. The metamorphic facies are arranged in relative positions with respect to pressure (depth) and temperature scales, and the experimental petrologist attempts to calibrate these scales by delineating the stability fields of minerals and mineral assemblages under known conditions in the laboratory. Potentially the most useful reactions for this purpose are solid-solid reactions involving no gaseous phase, such as the reactions among the polymorphs of Al_2SiO_5, kyanite, sillimanite, and andalusite.

Progressive metamorphism of sedimentary rocks produces a series of dehydration and decarbonation reactions with progressive elimination of water and carbon dioxide from the original clay minerals and carbonates. Experimental determination of these reactions, along with the solid-solid reactions, provides a petrogenetic pressure-temperature grid, in which mineral assemblages occupy pigeonholes bounded by specific mineral reactions. If mineral assemblages in metamorphic rocks are matched with the grid which has been calibrated in the laboratory, this provides estimates of the depth and temperature of the rock during metamorphism. Unfortunately, the temperatures of dissociation reactions are very sensitive to partial pressures of the volatile component involved in the reaction, and therefore there are complications introduced in application of the grid to metamorphic conditions. However, continued experimental studies of the stability of minerals and mineral stabilities under a wide range of laboratory conditions (in the presence of H_2O and CO_2 gas mixtures, for example, and with oxygen fugacity varied as well) may eventually provide a guide to the composition of the pore fluid that was present during metamorphism. If this composition is known, then the experimental data can be applied with greater confidence. *See* METAMORPHIC ROCKS. Peter J. Wyllie

Bibliography. T. F. W. Barth, *Theoretical Petrology*, 2d ed., 1962; M. G. Best, *Igneous and Metamorphic Petrology*, 2d ed., 2007; K. Bucher and M. Frey, *Petrogenesis of Metamorphic Rocks*, 7th ed., 2007; G. V. Chilinger, H. J. Bissel, and R. W. Fairbridge (eds.), *Carbonate Rocks*, 2 vols., 1967; Y. Guegen and V. Palciauskas, *Introduction to Rock Physics*, 1994; H. H. Hess and A. Poldervaart (eds.), *Basalts*, 2 vols., 1967; A. Miyashiro, *Metamorphic Petrology*, 1994; A. Philpotts, *Principles of Igneous and Metamorphic Petrology*, 1990; P. C. Ragland, *Basic Analytical Petrology*, 1989; M. Wilson, *Igneous Petrogenesis*, 1988; J. D. Winter, *Introduction to Igneous and Metamorphic Petrology*, 2001.

Petromyzontiformes

An order of jawless fishes also known as the Hyperoartii or commonly as the lampreys. It is the only order of the class Petromyzontida. Characteristics of the class include the following combination of characters that describe the order: eellike body; no bones; no paired fins; one or two dorsal fins; diphycercal (isocercal) tail (that is, with symmetrical upper and lower lobes) in adults, hypocercal (that is, with a larger lower lobe) in ammocoetes (larval lampreys); no barbels; teeth on oral disc and tongue (**Fig. 1**); two semicircular canals; seven pairs of external lateral gill openings; well-developed eyes in adults, and rudimentary eyes in ammocoetes, usually laterally placed; single median nostril (a nasohypophyseal sac with an external opening only); spiral valve and cilia in intestinal tract; small cerebellum; separate sexes; and oviparous, with small and numerous eggs. The ammocoetes undergo radical metamorphosis, and all species spawn in freshwater. *See* JAWLESS VERTEBRATES.

Biology and range. Lampreys are either parasitic or nonparasitic, and it is believed that nonparasitic species are derived from a parasitic species. The derived nonparasitic species are called satellite species, and the parasitic species are called stem species. For example, the Southern brook lamprey (*Ichthyomyzon gagei*) is thought to be a satellite of the chestnut lamprey (*I. castaneus*). The nonparasitic lampreys are usually called brook lampreys. Brook lampreys are confined to freshwater and spend most of their lifetime (usually about 3 to 5 years) as ammocoetes. The ammocoetes lack teeth, oral disc, and functional eyes (Fig. 1). As filter

feeders, they passively extract microorganisms from the surroundings, and they use little energy in the process. Also, by burrowing into the substrate, they are protected from predators. The size of organisms ingested is determined by a meshwork of branched cirri (the velum) at the posterior end of the oral hood. The food may include algae, bacteria, and partially decayed phytoplankton and detritus. It is no accident that the highest concentrations of ammocoetes are usually found in beds of extremely soft, detritus-laden substrates, which are rich sources for this type of food. Shortly following metamorphosis, the digestive tract atrophies, and the energy reserves of the body are used for rapid gonadal development. In preparation for spawning, both parasitic and nonparasitic lampreys usually migrate to small streams where the substrate is sand and small pebbles. Here, they excavate a shallow nest by carrying the pebbles in their mouth and by fanning the sand substrate. After spawning, the lampreys die.

Parasitic lampreys, by contrast, may be freshwater or anadromous (ascending into freshwater streams from the sea to spawn), but most species are limited to freshwater; however, no parasitic freshwater lampreys are known from the Southern Hemisphere. Parasitic lampreys have a prolonged adult life, and gonadal development is slower. Rasplike teeth rake the flesh of a host fish, allowing the lamprey to feed on its blood and tissue fluids, which is kept liquid by enzymes (lamphedrins) secreted by the lamprey's buccal glands. Parasitic adults continue to grow and at maturity are somewhat longer than related nonparasitic species. Before spawning, parasitic lampreys cease feeding, the intestine atrophies, the eyes degenerate, and the gonads rapidly increase in size. Some species are serious predators on commercially important fishes, with the notorious sea lamprey (*Petromyzon marinus*) being a good example. It is an anadromous species that occurs in Atlantic drainages of the United States, Canada, Iceland, and Europe, including the Mediterranean Sea, and is landlocked in the Great Lakes and other large lakes of that region. It is by far the largest lamprey, attaining a total length of 120 cm (4 ft), whereas all other lamprey species vary from about 12 to 76 cm (4.8 to 30 in.). *See* PARASITOLOGY; PREDATOR-PREY INTERACTIONS.

Lampreys are nearly antitropical (found in both hemispheres but not in equatorial regions) and are absent entirely from the broad belt around the world between the Tropic of Cancer and the Tropic of Capricorn. The earliest fossil agnathans (jawless fishes) date to the late Cambrian, about 500 million years before the present (ybp). Most members were extinct by the close of the Devonian, almost 200 million years later. Except for the fossil family Mayomyzontidae, a single family, Petromyzontidae, comprising eight genera and 34 species, occurs in the Northern Hemisphere, essentially north of 30°N. Only two genera and four species in two families, Geotriidae and Mordaciidae, occur in the Southern Hemisphere.

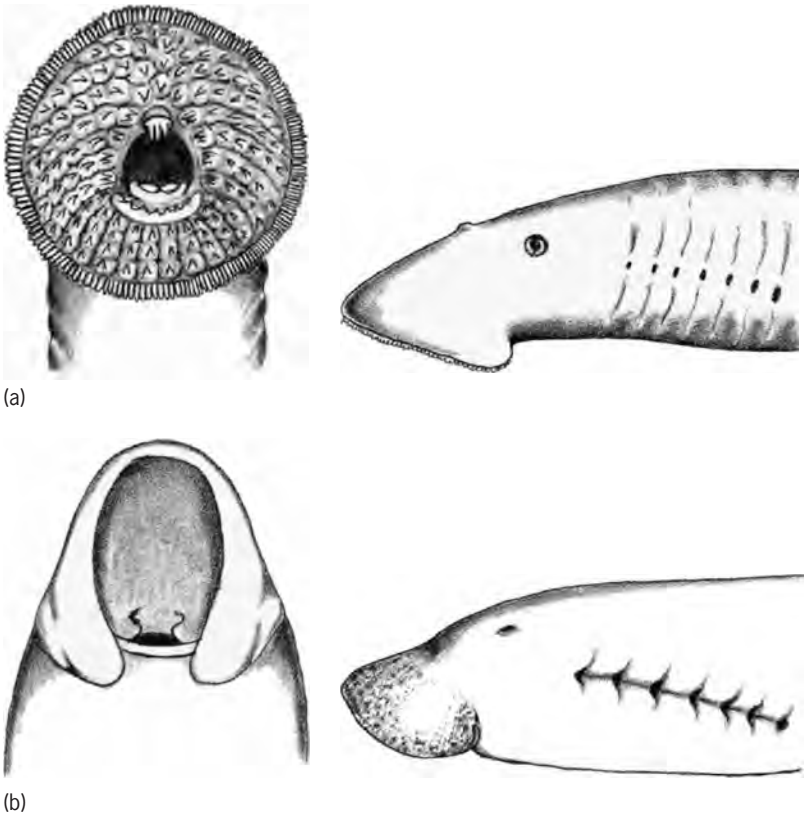

Fig. 1. Anteriors of (*a*) adult lamprey showing oral disc and teeth and (*b*) ammocoete showing oral hood and absence of teeth. (*From H. T. Boschung and R. I. Mayden, 2004*)

Petromyzontidae (northern lampreys). In addition to the characters given for the order, the family is further identified by dorsal fins continuous or at least slightly contiguous in adults; teeth in the anterior field of the oral disc in alate (winglike) rows; pointed teeth; infraoral lamina; and diphycercal tail. Three genera and 20 species occur in North America (including Mexico); the remainder are Eurasian. *Ichthyomyzon* (**Fig. 2**) is limited to streams and rivers of eastern North America that drain into the Gulf of Mexico, St. Lawrence River, and Hudson Bay. *Lampetra* is holoarctic in freshwater; and *Petromyzon* is anadromous.

Geotriidae (southern lampreys). This family is represented by a single species which differs in a number of characters, including well-separated dorsal fins and spatulate-shaped teeth on the oral disc. It is an

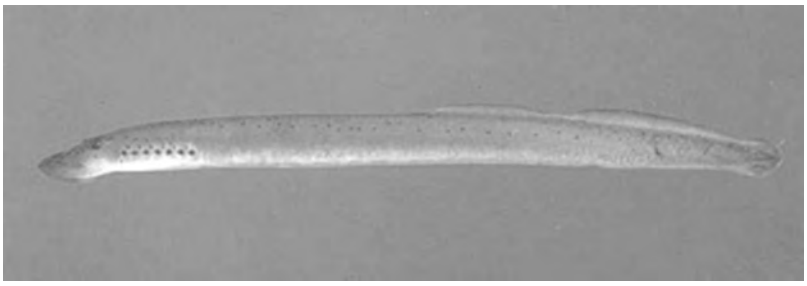

Fig. 2. Ohio lamprey (*Ichthyomyzon bdellium*). (*Bob Jenkins and Noel Burkhead; EFISH <http://www.cnr.vt.edu/efish/families/oh_lamp.html>*)

anadromous parasitic species that occurs in southern Australia, Tasmania, New Zealand, Chile, Argentina, and the Falkland and South Georgia islands.

Mordaciidae (southern topeyed lampreys). The one genus and three species of three parasitic and nonparasitic lampreys can be distinguished simply by combining the following three characters: separate dorsal fins; pointed or rounded disc teeth; and dorsally placed eyes in mature individuals and dorsolaterally placed eyes in the immature.

Mayomyzontidae. The single species assigned to this family is known from the Pennsylvanian Period (about 300 million ybp) in Illinois. It is a relatively small species differing from extant species in dentition.

Classification. Because of recent major changes in the classification of craniates and in order to put the Petromyzontiformes in perspective, the old and new classifications are compared. The first classification of the craniates, as used by many authors in the past, is as follows (the dagger marks extinct forms):

Phylum Chordata
- Subphylum Vertebrata (=Craniata)
  - Superclass Agnatha
    - Class Myxini
    - Class Pteraspidomorphi
      - Order: †Arandaspidiformes
        - †Pteraspidiformes
        - †Thelodontiformes
    - Class Cephalaspidomorphi
      - Order: Petromyzontiformes
        - †Anaspidiformes
        - †Galeaspidiformes
        - †Cephalaspidiformes (=Osteostraci)
  - Superclass Gnathostomata

The revised classification, reflecting recent studies of agnathans, summarized and documented by J. S. Nelson (2004), is as follows:

Phylum Chordata
- Subphylum Craniata
  - Superclass Myxinomorphi
    - Class Myxini
      - Order Myxiniformes (Hyperotreti—hagfishes)

All remaining craniates may be placed in the clade of vertebrates:

- Superclass Petromyzontomorphi
  - Class Petromyzontida
    - Order Petromyzontiformes (Hyperoartii—lampreys)
      - Family: Petromyzontidae
        - Geotriidae
        - Mordaciidae
        - †Mayomyzontidae
- Superclass Conodonta
  - †Class Conodonta
- Superclass Pteraspidomorphi
  - Class Pteraspidomorphi (Diplorina)
    - Order: †Astraspidiformes
      - †Arandaspidiformes
      - †Cyathaspidiformes
      - †Pteraspidiformes
- Superclass Anaspida
  - †Class Anapsida
    - †Order Anaspidiformes
- Superclass Thelodonti
  - Class Thelodonti
    - Order: †Loganelliformes
      - †Shieliiformes
      - †Phlebolepidiformes
      - †Thelodontiformes (+Coelolepidiformes)
      - †Furcacaudiformes
- †Superclass Osteostracomorphi
  - Class Cephalaspidomorphi (Monorhina)
    - Order: †Cephalaspidiformes (Osteostraci)
      - †Galeaspidiformes
      - †Pituriaspidiformes
- Superclass Gnathostomata

In the recent classification, the terms Craniata and Vertebrata are not considered as synonyms, but are at different levels of the hierarchy. Members of the subphylum Craniata (less the Myxinomorphi) may be placed in the clade of vertebrates to be recognized as the infraphylum Vertebrata. Agnatha and Cyclostomata are no longer used as taxon names; however, the common name "agnathans" is used for all jawless fishes (a paraphyletic group), and "cyclostomes" is a term used for the extant jawless fishes (it too is a paraphyletic group).

The presence of a round "mouth," horny teeth, large notochord, and pouch-shaped gills in hagfishes and lampreys led to their placement together in the Cyclostomata. Both modern groups differ from Paleozoic agnathans in the total lack of bones, scales, and paired fins. The interrelationships of cyclostomes are often the topic of debate. It is now generally believed that living agnathans (cyclostomes) are not closely related to each other and that the closest living relatives of lampreys are the gnathostomes. This notion is supported by a number of morphological and physiological characters; however, the deoxyribonucleic acid (DNA) sequence is inconclusive. *See* CHORDATA; CYCLOSTOMATA (CHORDATA); GNATHOSTOMATA; MYXINIFORMES; VERTEBRATA. Herbert Boschung

Bibliography. H. T. Boschung and R. L. Mayden, *Fishes of Alabama*, Smithsonian Books, Washington, DC, 2004; R. L. Carroll, *Vertebrate Paleontology and Evolution*, W. H. Freeman, New York, 1988; R. E. Foreman et al. (eds.), *Evolutionary Biology of Primitive Fishes*, Plenum Press, New York, 1985; P. L. Forey and P. Janvier, Evolution of the early vertebrates, *Amer. Sci.*, 82:554–565, 1994; H. S. Gill et al., Phylogeny of living parasitic lampreys (Petromyzontiformes) based on morphological data, *Copeia*, 2003(4):687–703, 2003; M. W. Hardisty and I. C. Potter (eds.), *The Biology of Lampreys*, vol. 1, Academic Press, London, 1971; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Pewter

A tarnish-resistant alloy of lead and tin always containing appreciably more than 63% tin. Other metals are sometimes used with or in place of the lead; among them are copper, antimony, and zinc. Pewter is commonly worked by spinning and it polishes to a characteristic luster. Because pewter work-hardens only slightly, pewter products can be finished without intermediate annealing.

Early pewter, with high lead content, darkened with age. With less than 35% lead, pewter was used for decanters, mugs, tankards, bowls, dishes, candlesticks, and canisters. The lead remained in solid solution with the tin so that the alloy was resistant to the weak acids in foods.

Addition of copper increases ductility; addition of antimony increases hardness. Pewter high in tin (91% tin and 9% antimony or antimony and copper, for example) has been used for ceremonial objects, such as religious communion plates and chalices, and for cruets, civic symbolic cups, and flagons. *See* ALLOY STRUCTURES; LEAD ALLOYS; TIN ALLOYS.

Frank H. Rockett

pH

An expression for the effective concentration of hydrogen ions in solution. The activity of hydrogen ions or, more correctly, hydronium ions, that is, hydrated hydrogen ions $H(H_2O)_n^+$, affects the equilibria and kinetics of a wide variety of chemical and biochemical reactions. Because these effects are activity dependent, it is extremely important to distinguish between the hydrogen-ion concentration and activity. The concentration, or total acidity, is obtained by titration and corresponds to the total concentration of hydrogen ions available in a solution, that is, free, unbound hydrogen ions as well as hydrogen ions associated with weak acids. The hydrogen-ion activity refers to the effective concentration of unassociated hydrogen ions, the form that directly affects physicochemical reaction rates and equilibria. This activity is therefore of fundamental importance in many areas of science and technology. The relationship between hydrogen-ion activity (a_{H^+}) and concentration (C) is givenby Eq. (1), where the activity coefficient γ is a

$$a_{H^+} = \gamma C \qquad (1)$$

function of the total ionic strength (concentration) of the solution and approaches unity as the ionic strength approaches zero; that is, the difference between the activity and the concentration of hydrogen ion diminishes as the solution becomes more dilute. *See* ACTIVITY (THERMODYNAMICS); CHEMICAL EQUILIBRIUM; HYDROGEN ION.

The effective concentration of hydrogen ions in solution is expressed in terms of pH, which is the negative logarithm of the hydrogen-ion activity [Eq. (2)].

$$\text{pH} = -\log_{10} a_{H^+} \qquad (2)$$

Approximate pH of some common solutions

Solution	pH	a_{H^+}
Battery acid	0	1
Cola soft drink	2	10^{-2}
Vinegar	3	10^{-3}
Beer	5	10^{-5}
Pure water	7	10^{-7}
Milk of magnesia	10	10^{-10}
Household ammonia	12	10^{-12}
Caustic drain cleaner	14	10^{-14}

Because of the negative logarithmic (exponential) relationship, the more acidic a solution, the smaller the pH value (see **table**). The pH of a solution may have little relationship to the titratable acidity of a solution that contains weak acids or buffering substances; the pH of a solution indicates only the free hydrogen-ion activity. If total acid concentration is to be determined, an acid-base titration must be performed. *See* ACID AND BASE; BUFFERS (CHEMISTRY); TITRATION.

Two methods, electrometric and chemical indicator (optical), are used for measuring pH. The more commonly used electrometric method is based on the glass (or polymeric) pH electrode used along with a reference electrode and pH meter (an electronic voltmeter that provides a direct conversion of potential differences to values of pH). This procedure is based on measurement of the difference between the pH of a test solution and that of a standard solution. The pH scale is defined by a series of reference buffer solutions that are used to calibrate the pH measurement system. The instrument measures the potential difference developed between the pH electrode and a reference electrode of constant potential. The difference in potential obtained when the electrode pair is removed from the standard solution and placed in the test solution is converted to the pH value. In the indicator method, the pH value is obtained by simple visual comparison of the color of pH-sensitive dyes to standards (for example, color charts) or by use of calibrated optical readout devices (photometers), often in combination with fiber-optic sensors. *See* ACID-BASE INDICATOR; ELECTRODE; REFERENCE ELECTRODE.

The determination of pH is one of the most widely performed analytical measurements, ranging in application from basic research and clinical analyses to industrial process control and environmental monitoring. Although electrometric pH measurements are among the most specific with respect to possible interferences, a number of potential sources of error can affect the measurements and care must be exercised to ensure the accuracy of the measurements.

Richard A. Durst

Bibliography. R. G. Bates, *Determination of pH: Theory and Practice*, 2d ed., 1973; G. Eisenman (ed.), *Glass Electrodes for Hydrogen and Other Cations*, 1967; Y. C. Wu, W. F. Koch, and R. A. Durst, *Standardization of pH Measurements*, NBS Spec. Publ. 260-53, 1988.

pH regulation (biology)

The processes operating in living organisms to preserve a viable acid-base state. In higher animals, much of the body substance (60–70%) consists of complex solutions of inorganic and organic solutes. For convenience, these body fluids can be subdivided into the cellular fluid (some two-thirds of the total) and the extracellular fluid. The latter includes blood plasma and interstitial fluid, the film of fluid that bathes all the cells of the body. For normal function, the distinctive compositions of these various fluids are maintained within narrow limits by a process called homeostasis. A crucial characteristic of these solutions is pH, an expression representing the concentration (or preferably the activity) of hydrogen ions, $[H^+]$, in solution. The pH is defined as $-\log [H^+]$, so that in the usual physiological pH range of 7 to 8, $[H^+]$ is exceedingly low, between 10^{-7} and $10^{-8}M$. Organisms use a variety of means to keep pH under careful control, because even small deviations from normal pH can disrupt living processes. *See* HOMEOSTASIS; PH.

Normal values. The most accessible and commonly studied body fluid is blood, and it, therefore, provides the most information on pH regulation. Blood pH in humans, and in mammals generally, is about 7.4. This value indicates that blood is slightly alkaline, because neutrality, the condition in which the concentration of hydrogen ions ($[H^+]$) equals the concentration of hydroxyl ions ($[OH^-]$), is pH 6.8 at mammalian body temperature of 98°F (37°C). The pH within cells, including the red blood cells, is typically lower by 0.2–0.6 unit, and is thus close to neutrality. The cell fluid is more complex and far less accessible than blood, but is of obvious importance and has received considerable research attention with techniques for measuring cell pH such as micro pH electrodes, the distribution characteristics of weak acids and bases, nuclear magnetic resonance, and fluorescent dyes.

In most animals other than warm-blooded mammals, blood pH deviates from the familiar value of 7.4. The major reason is that body temperature has an important influence on pH regulation, and as a rule, the blood pH of invertebrates and lower vertebrates varies inversely with body temperature (see **illus.**). For example, the blood pH of a reptile at 86°F (30°C) may be 7.5, but when the same reptile is tested at 50°F (10°C) its blood pH may be 7.8. Values of blood pH from a great many animals fall on a broad continuum at body temperatures between 32 and 104°F (0 and 40°C; see illus.), and humans and other mammals occupy a stationary position near the high-temperature end. The significance of this regulatory pattern is controversial, but one hypothesis is that the net electric charge state of proteins remains constant when pH varies in this manner. Preservation of charge state will stabilize the three-dimensional structure of proteins that may be essential for their normal roles as enzymes or membrane receptor molecules. Consequently, animals that experience significant changes in body temperature have no single normal pH at which they regulate, but rather a series of values depending on body temperature.

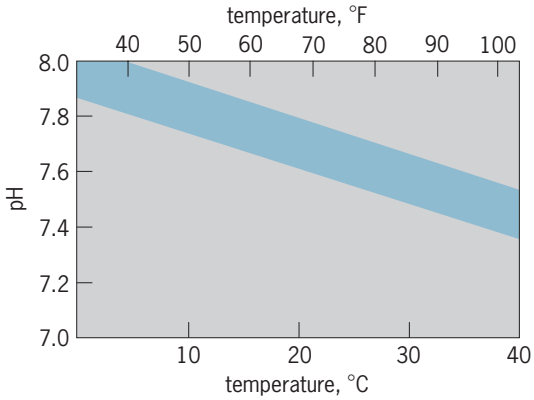

The pH of blood samples from animals at various body temperatures varies inversely with temperature. (*After D. C. Jackson, Blood acid-base control in ectotherms, in C. R. Taylor, K. Johansen, and L. Bolis, eds., A Companion to Animal Physiology, Cambridge University Press, 1982*)

Blood. Blood pH regulation is necessary because metabolic and ingestive processes add acidic or basic substances to the body and can displace pH from its proper value. In addition, organisms excrete waste products and exchange substances with their surroundings, and these processes can affect acid-base balance. The immediate impact on pH of any of these events, however, is moderated by the chemical buffers of the blood and other body fluids. Buffers are weak acids and bases that are partially dissociated at physiological pH. Their dissociation properties help set the fluid pH at its normal value and resist displacements of pH away from that value by combining with most H^+ or OH^- added to the system. Buffering is a passive, physicochemical process, though, that cannot precisely regulate the pH of a living organism. For true regulation, active physiological mechanisms are required that can alter the acid-base composition of the blood in a controlled fashion. *See* BUFFERS (CHEMISTRY).

It is conventional to identify these control mechanisms with their effects on the principal buffer of the extracellular fluid, carbon dioxide (CO_2). Carbon dioxide is produced by cellular metabolism and distributes readily throughout the body because of its high solubility and rapid diffusion. In solution, CO_2 is hydrated to carbonic acid (H_2CO_3) which dissociates almost completely to H^+ and bicarbonate ions (HCO_3^-). Each of these reactions is reversible, as shown below. Further dissociation of HCO_3^- to

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$

carbonate and H^+ is negligible at normal blood pH. Dissolved CO_2 can equilibrate and exchange with CO_2 gas, so that the dissolved CO_2 corresponds to a particular partial pressure of CO_2 (P_{CO_2}). The ready loss of CO_2 from the body, by diffusion into either the surrounding water or air, is a key property of this molecule and accounts for its importance in pH regulation. To regulate blood pH, organisms have mechanisms to independently control P_{CO_2} and $[HCO_3^-]$.

Pco_2 control. To control Pco_2, animals utilize gas exchangers, such as lungs or gills. By appropriately adjusting the rate at which air ventilates the lungs or water flows through the gills, the respiratory loss of CO_2 is matched to the metabolic production of CO_2, and body fluid Pco_2 is controlled. The value of blood Pco_2 can be quite different depending on the habitat of an organism. Fishes and other aquatic animals have low blood Pco_2, on the order of 2–3 mmHg (0.3 kilopascal), because of the large volumes of water that must be breathed to obtain enough O_2 from it. Due to the close link between fish respiration and O_2 supply, the gills are not commonly employed for pH control. In contrast, air breathers generally have higher values of Pco_2 [human arterial blood Pco_2 is close to 40 mmHg (5.3 kPa)], and their breathing is less tied to O_2 uptake in the more O_2-rich aerial environment. Consequently, air breathers can use their lungs to rapidly and accurately control Pco_2 and pH. Animals such as lungless salamanders that rely on passive diffusion of CO_2 and O_2 through the body surface and have no active breathing mechanism may lack the accurate and rapid control of Pco_2. *See* RESPIRATION.

Bicarbonate control. To control [HCO_3^-], animals use various ion exchange or excretion mechanisms that differentially transfer strong cations (such as Na^+) and strong anions (such as Cl^-) and cause a shift in the equilibrium state of CO_2 and other weak acids. Assuming that Pco_2 is set by the respiratory system, these ion exchanges will exert their effects on pH and [HCO_3^-]. In most terrestrial vertebrates, the kidneys serve this function and can produce and excrete urine having a highly acid or alkaline composition, depending on the needs of the organism. In fishes, the gills are the primary site for ionic pH regulation. Unlike the kidneys, which are strictly excretory organs, the gills exchange ions with the water passing through; for example, in a fresh-water fish, ammonium ion (NH_4^+) or H^+ ions may be pumped into the water in exchange for Na^+, or HCO_3^- may be exchanged for ambient Cl^-; these processes respectively raise or lower the blood [HCO_3]. Other tissues, such as the amphibian skin, may also excrete or exchange ions for acid-base regulation, but each of these mechanisms, whether kidneys, gills, or skin, shares the common effect of controlling the ionic balance of the organism and thereby controlling [HCO_3^-] and pH. Ionic mechanisms are slower than respiratory Pco_2 changes, but are probably the principal pH regulators in fishes and in animals with passive gas exchanges. *See* KIDNEY; OSMOREGULATORY MECHANISMS; RESPIRATORY SYSTEM.

Cells. The respiratory and ionic mechanisms already discussed serve the interests of all the body fluids, although their immediate effects may be on the blood. The cells, while benefitting from the stability afforded by these whole body mechanisms, also have local means for their own pH regulation. An acute acid load on a cell, whether from its intrinsic metabolism or from an external source, is dealt with first by the cell's chemical buffering capacity, a capacity that exceeds that of the blood by severalfold. Other cellular mechanisms include the conversion of organic acids to neutral compounds through metabolic transformations, and the transfer of acid equivalents from the primary cell fluid, the cytoplasm, into cellular organelles. Ultimately, however, to restore its normal acid-base state, the cell must expel the excess acid to the exterior; cell membrane ion exchangers have been described in several cell types that can do this. For example, in mammalian cells, a coupled Na^+/H^+ exchange exists. In more easily studied giant cells of various invertebrates, other coupled ion transfers, such as $H^+ + Cl^-$ out in exchange for $Na^+ + HCO_3^-$ in, have been described.

Presumably, all cells possess similar transmembrane transport systems to independently regulate the pH of their contained fluid and to act as a second line of defense behind the overall body pH control mechanisms. Cell pH regulation is complicated by the heterogeneous nature of cell contents. The cytoplasm is the bulk of the cell fluid, but is itself the fluid bathing the cell organelles, and the fluid within these organelles differs in pH from the cytoplasm. For example, mitochondrial fluid is more alkaline than the cytoplasm, so that a H^+ gradient exists between these two compartments that is created and maintained by H^+ pumps on the inner mitochondrial membrane. Oxidative phosphorylation, the transfer of metabolic energy to adenosine triphosphate (ATP), is thought to depend on this H^+ gradient. *See* CHEMIOSMOSIS.

Significance. The integrity of cellular function requires a normal pH in the extracellular fluid bathing the cells and, above all, in the fluids within the cells. Examples of vital functions dependent on normal pH include membrane permeability to charged molecules and catalyzed metabolic reactions. The permeability function underlies the excitable properties of nerve and muscle that give rise to action potentials and nerve transmission. A seemingly small reduction in blood pH, from 7.4 to 7.0, can severely depress nervous function in a human. The basis for this effect is probably a change in the distribution of fixed charges within membrane ion channels. For metabolic reactions, enzymes have pH optima where their catalysis of substrate reactions is maximal. Disturbances in pH can alter surface charges on these protein molecules, leading to changes in their three-dimensional structure, and loss of their catalyzing properties. *See* BIOPOTENTIALS AND IONIC CURRENTS; ENZYME.

Donald C. Jackson

Bibliography. N. L. Jones, *Blood Gases and Acid-Base Physiology*, 2d ed., 1987; H. A. Neidig and J. N. Spencer, *pH: Acids and Bases*, 1992.

Phaeophyceae

A class of plants, commonly called brown algae, in the chlorophyll *a-c* phyletic line (Chromophycota). Brown algae occur almost exclusively in marine or brackish water, where they are attached to rocks, wood, sea grasses, or other algae. Approximately 265 genera and 1500 species are recognized, arranged in about 15 orders. *See* ALGAE.

Characteristics. Phaeophyceae are characterized primarily by biochemical and ultrastructural features. The cells are typically uninucleate and contain one or more chloroplasts with or without pyrenoids. Photosynthetic pigments include chlorophyll *a* and *c*, β-carotene, and several xanthophylls, principally fucoxanthin. Varying proportions of these pigments cause the thalli to range in color from yellow-brown to olive green. Food reserves are chiefly laminaran (a polysaccharide in which the glucose units have β-1,3-linkages) and mannitol (a soluble sugar alcohol). The cell walls are composed of an inner layer of cellulose and an outer layer of alginate (a linear polymer of mannuronic and glucuronic acids) and fucans (highly branched sulfated polysaccharides). The mucilage that is often secreted copiously by the larger brown algae (kelps and rockweeds) is composed of various polysaccharides, especially fucans. Motile reproductive cells are usually pear-shaped and have two laterally inserted flagella, an anteriorly directed tinsel flagellum, and a posteriorly directed smooth flagellum. *See* ALGINATE.

Structure. The simplest thallus exhibited by Phaeophyceae is generally considered to be an erect, unbranched or branched, uniseriate filament arising from a prostrate filamentous base. Some phycologists, however, believe that the unicellular algae assigned to the order Sarcinochrysidales of the class Chrysophyceae may be considered primitive brown algae. Many Phaeophyceae are crustose or bladelike. Complex thalli differentiated into macroscopic organs are produced by kelps and rockweeds. Growth is diffuse or, more often, localized at the apex of the thallus (apical growth), at the base of uniseriate filaments (trichothallic growth; see **illus.**), or in special intercalary tissues (transition meristem of kelps, meristoderm of kelps and rockweeds). Colorless uniseriate hairs with a basal meristem (phaeophycean hairs) are common.

Embryonic sporophyte of *Desmarestia anceps*, showing trichothallic growth resulting in an embryonic stipe and several pairs of embryonic branches.

Life histories and reproduction. Most Phaeophyceae have a life history involving an alternation of both somatic and cytological phases. The diploid thallus (sporophyte) bears unilocular sporangia in which the initial nucleus undergoes meiosis. The resulting haploid nuclei undergo several mitotic divisions so that 32 or 64 zoospores are commonly produced within a sporangium. A zoospore develops directly into a haploid thallus (gametophyte), which produces isogametes, anisogametes, or eggs and spermatozoids. The two somatic phases may be identical in size and form (isomorphic) or as different (heteromorphic) as the microscopic gametophyte and 100-ft-long (30-m) sporophyte of the giant kelp, *Macrocystis*. The more primitive brown algae are markedly plastic in their life histories, with one or both somatic phases, and sometimes an interpolated juvenile phase, recycling by means of zoospores. Rockweeds, such as *Fucus*, have a life history with only one somatic phase, a diploid thallus that produces eggs and sperm. The gametes, however, do not incorporate the nuclei that result directly from meiosis (as in animals), but instead the derivatives of those nuclei. The intervening mitotic divisions are interpreted by some phycologists as a manifestation of a vestigial or incipient haploid somatic phase.

Classification. The order Ectocarpales comprises simple filamentous forms with diffuse growth and isomorphic somatic phases. In the Sphacelariales, the filaments have longitudinal cell divisions (pluriseriate) and terminate in conspicuous apical cells that initiate growth. The somatic phases are isomorphic, and sexual reproduction is anisogamous or oogamous. In the Chordariales, the filaments are often compacted into a firm or mucilaginous branched thallus, and the two somatic phases are heteromorphic. The thallus of Dictyotales is a blade in which growth is effected by apical cells, either singly at the tip of each branch or in marginal rows. Sexual reproduction is oogamous, with the sperm unique in lacking a posteriorly directed flagellum. The meiospores formed on the isomorphic soporophyte are also distinctive within the Phaeophyceae in that they are nonmotile. Thalli of Dictyosiphonales are parenchymatous and in the form of a solid or hollow cylinder or a blade. The orders Laminariales (kelps) and Fucales (rockweeds), which exhibit the most specialized anatomy in the algae (including food-conducting tissues), are very interesting biologically and economically important. Large kelplike thalli may be formed in the Desmarestiales, but growth is pseudoparenchymatous, being effected primarily by divisions of a row of cells at the base of filaments that

terminate the branches, and secondarily by the production of external corticating filaments as well as internal thickening filaments (see illus.). *See* FUCALES; LAMINARIALES.

Distribution. The geographic distribution of Phaeophyceae is bimodal. Kelps are most abundant and diverse on surf-swept rocky shores of the North Pacific, but they form an ecologically important vegetation belt in the lower intertidal and upper subtidal zones on all cold-water shores except Antarctica, where they are replaced by members of the Desmarestiales. Rockweeds are similarly abundant on cold-water shores, forming conspicuous belts both in the upper intertidal zone, but they reach their peak of diversity in Australia and New Zealand. Rockweeds also form extensive stands in salt marshes in the Northern Hemisphere. Tropical waters, on the other hand, support a diverse array of Dictyotales and members of the fucalean family Sargassaceae. Five genera of brown algae have been recorded from fresh-water habitats, especially rapidly flowing streams, but so infrequently that it is impossible to deduce a meaningful pattern of geographic distribution. Paul C. Silva; Richard L. Moe

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae*: *Structure and Reproduction*, 2d ed., 1985; M. N. Clayton, Evolution of the Phaeophyta with particular reference to the Fucales, *Prog. Phycol. Res.*, 3:11–46, 1984; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Phagocytosis

A mechanism by which single cells of the animal kingdom, such as smaller protozoa, engulf and carry particles into the cytoplasm. It differs from endocytosis primarily in the size of the particle rather than in the mechanism; as particles approach the dimensions and solubility of macromolecules, cells take them up by the process of endocytosis.

Feeding mechanism. Cells such as the free-living amebas or the wandering cells of the metazoa move about their environment by ameboid movement. They often can "sense" the direction of a potential food source and move toward it (chemotaxis). If, when the cell contacts the particle, the particle has the appropriate chemical composition, or surface charge, it adheres to the cell. The cell responds by forming a hollow, conelike cytoplasmic process around the particle, eventually surrounding it completely. Although the particle is internalized by this sequence of events, it is still enclosed in a portion of the cell's surface membrane and thus isolated from the cell's cytoplasm. The combined particle and membrane package is referred to as a food or phagocytic vacuole, one of the many normal components of cytoplasm in cells capable of phagocytosis. *See* FEEDING MECHANISMS (INVERTEBRATE); VACUOLE.

Not all cell contacts with particles result in phagocytosis, primarily because the particles fail to stick to the cell surface. Most amebas respond to moving objects such as fine glass needles or swimming paramecia or other food organisms which repeatedly contact them by attempting phagocytosis. This seems to be a nonspecific response to mechanical stimulation rather than a specific response to the chemical properties of the object. There is another type of nonspecific phagocytosis (surface phagocytosis) in which the cell engulfs any small object which it traps in a surface depression during normal locomotion.

Defense mechanism. At the end of the nineteenth century E. Metchnikoff first proposed that the ameboid cells of the metazoa can selectively remove foreign particles, bacteria, and other pathogens by phagocytosis. He demonstrated his proposition by introducing a thorn into a transparent starfish larva and noting the large numbers of wandering cells that accumulated around the thorn. Subsequent studies by Metchnikoff and others established that the macrophages and some white blood cells of humans can remove bacteria and other foreign particles, such as splinters, which invade the body. Of particular significance in this defense mechanism is the highly specific interaction between the surface of the bacteria and the surface of the macrophage. The same mechanism governs the transplant of organs from one individual to another: If the normal activities of the macrophages are not suppressed by drugs, most transplants fail. *See* TRANSPLANTATION BIOLOGY.

Immunity factors. The surface of the macrophage may adhere spontaneously to the surface of the bacteria or the foreign cell with no intervening factors. However, substances in the blood and tissue fluids usually act to facilitate sticking, thereby accelerating phagocytosis. These substances (opsonins) are usually antibodies and blood factors which give the body immunity to foreign cells and viruses. The concentration of specific opsonins in the host, the rate at which the macrophages move by chemotaxis toward the invading bacteria, the ability of the macrophage to stick to the bacteria, and the completeness with which the macrophage can digest or destroy the microbes, balanced against the rate of microbe entry and growth, determine the course of the disease caused by the microorganism. *See* DISEASE; IMMUNITY; OPSONIN.

Digestion and destruction. After the foreign particle or microorganism is trapped in a vacuole inside the macrophage, it is usually digested. To accomplish this, small packets (lysosomes) of lytic proenzymes are introduced into the phagocytic vacuole, where the enzymes are then dissolved and activated. However, not all ingested microbes are killed and digested; some (for example, the tubercle bacillus) survive and multiply. In such instances the body must rely on other methods to combat the infection just as it must when the microbe initially fails to stick to the surface of the macrophage. The destructive abilities of the enzymes secreted into the phagocytic vacuole are supplemented by the ionic conditions found in the cytoplasm. The vacuole fluid is rapidly equilibrated with the cytoplasm and becomes therefore rich in potassium ions and deficient

in calcium ions. This condition, which is necessary and normal within cells, is intolerable to most forms of cellular life when it exists both inside and outside the cell membrane. Thus the microbe loses the life-supporting juxtaposition of very differently composed salines if it becomes trapped in the cytoplasm of a wandering cell by the process of phagocytosis. *See* ENDOCYTOSIS; LYSOSOME. Philip W. Brandt

Phanerozoic

The uppermost eonothem (a time-stratigraphic unit composed of rocks formed during an eon of geologic time) of the Standard Global Chronostratigraphic Scale. The Phanerozoic Eonothem is the physical stratigraphic representation of the Phanerozoic Eon, a geologic time unit, embracing the Paleozoic, Mesozoic, and Cenozoic eras (see **illustration**). Phanerozoic time, which extends from 542 million years ago to the present, is estimated to be slightly less than 12% of the Earth's history. *See* GEOLOGIC TIME SCALE; STRATIGRAPHIC NOMENCLATURE; STRATIGRAPHY.

The Phanerozoic essentially represents the most recent cycle of supercontinent breakup, assembly, and subsequent breakup. Following the dispersal of Pannotia, beginning in the latest Proterozoic, rifted crustal pieces (Gondwana, Laurentia, Baltica, Siberia, Avalonia, and others) reassembled as the supercontinent Pangea in the late Paleozoic. Pangea split into the modern continents beginning in the mid-Mesozoic, and has continued to undergo rifting and dispersal through the Cenozoic. *See* PALEOGEOGRAPHY; SUPERCONTINENT.

The Phanerozoic Eonothem is characterized by a good visible fossil record in sedimentary strata. Through much of the Phanerozoic, global sea level was at positions higher than its present position, and continental shelves were flooded with widespread shallow seas that have left a rich marine fossil record. Sea level was near or below its present position in parts of the late Paleozoic, Mesozoic, and late Cenozoic, and these are intervals for which a good terrestrial fossil record is also available. *See* FOSSIL.

Paleozoic. The beginning of the Paleozoic Era (the Cambrian Period) represents one of the most significant points in evolutionary history. The boundary between the Proterozoic Eonothem and the overlying Phanerozoic Eonothem is a zone of transition from a fossil-poor stratigraphic record to a fossil-rich stratigraphic record. The base of the Phanerozoic Eonothem (which is the same as the bases of the Paleozoic Erathem and the Cambrian System) is marked by the appearance of the earliest assemblage of intricate marine trace fossils. These traces are an indication of the evolution of complex animal behavior. Animal body fossils are present in the Proterozoic, but they are neither diverse nor common in many places. Succeeding the basal Paleozoic trace fossil assemblage are distinctive assemblages of body fossils representing the great diversification of marine animals during the Cambrian Period, an event characterized as the Cambrian explosion. Assemblages of small shelly fossils (tiny biomineralized skeletal parts) are replaced about midway through the Cambrian by more diverse organisms such as trilobites, brachiopods, mollusks, echinoderms, sponges, and hyoliths. Trilobites dominate most Cambrian fossil deposits, except for deposits of exceptional preservation, in which biomineralized and nonbiomineralized organisms are preserved. The most spectacular deposits are the Burgess Shale of British Columbia and the Chengjiang deposit of Yunnan, China, both of which contain remains of many types of arthropods, worms, chordates, cnidarians, algae, macroscopic bacteria, and many other organisms. *See* BURGESS SHALE; CAMBRIAN; TRACE FOSSILS; TRILOBITA.

The early Paleozoic was a time of dramatic changes not just in marine life forms but also in the consistency of marine sedimentary surfaces. At the end of the Proterozoic, sediment–water interfaces were stabilized by microbial biomats (aggregates of microorganisms composed mainly of bacteria and algae). With the introduction of diverse benthic (bottom-dwelling) and nektobenthic (dwelling just above the ocean bottom) animals in the Cambrian, sediment surfaces were increasingly burrowed and fluidized. Burrowing worms and burrowing predatory trilobites were major contributors to the fluidization of marine sediments through the Cambrian and Ordovician periods.

From the Ordovician through the end of the Paleozoic (Permian Period), marine biotas were dominated by brachiopods, mollusks, corals, echinoderms, stromatoporoid and other sponges, and trilobites. Ammonoid cephalopods, which first appeared in the Devonian, soon diversified to become one of the dominant marine predators of the mid-Paleozoic to the end of the Mesozoic. Crinoids, a group of stalked echinoderms, formed vast "meadows" in shallow tropical seas of the Carboniferous.

In the Paleozoic, the land was invaded by vertebrates, arthropods, plants, and other organisms. Jawless fishes evolved in the early Paleozoic and gave rise to a wide variety of jawed fishes in the mid-Paleozoic (Silurian and Devonian periods). One group of jawed fishes, the lobe-finned fishes, gave rise in the Devonian Period to amphibians, the earliest land-dwelling vertebrates. Amphibians, in turn, spawned the terrestrial reptiles, which underwent initial diversification in the Carboniferous-Permian. Invasion of the land by plants is indicated first by spores in the Ordovician and then vascular tissues in the Silurian. The first seed plants (gymnosperms) and forested wetland areas had developed by the Devonian; and widespread wetlands dominated by lycopods, early conifers, scouring rushes, ferns, seed ferns, and progymnosperms formed extensive "coal swamps" in coastal, low-latitude areas in the Carboniferous. Arachnid arthropods invaded the land by the Silurian, and winged insects evolved from wingless terrestrial arthropods in the Devonian. After undergoing adaptive radiation in the Carboniferous, the insects have maintained a position as the most evolutionarily successful group of animals. Today,

Eons	Eras	Periods	Epochs	Geologic ages, Ma	Major Evolutionary Events
Phanerozoic	Cenozoic	Neogene	Holocene	0.01	Mass extinction of terrestrial mammals; *Homo sapiens* appears
			Pleistocene	1.8	
			Pliocene	5.3	First hominids; grasses radiate
			Miocene	23.0	
		Paleogene	Oligocene	33.9	
			Eocene	55.8	Mammals, angiosperms, insects diversify
			Paleocene	65.5	Mass extinction; ammonite cephalopods, marine reptiles, dinosaurs other than birds become extinct
	Mesozoic	Cretaceous		145.5	First angiosperms First birds
		Jurassic		199.6	Mass extinction First dinosaurs and mammals First scleractinian coral reefs
		Triassic		251.0	Mass extinction; trilobites become extinct
	Paleozoic	Permian		299.0	Reptiles diversify First reptiles
		Carboniferous		359.2	Insects and land plants diversify; vast wetlands form Crinoids diversify Mass extinction
		Devonian		416.0	First amphibians (land-dwelling vertebrates) First ammonoid cephalopods, ray-finned fishes, insects, gymnosperms, and forests First jawed fishes
		Silurian		443.7	First vascular land plants and terrestrial arthropods Mass extinction Terrestrial plant spores appear
		Ordovician		488.3	Ordovician radiation of invertebrates Extinctions of trilobites and other animals
		Cambrian		542.0	Trilobites appear; animals diversify Small shelly fossils, brachiopods, mollusks appear Complex trace fossils
Proterozoic				2500	First animals, trace fossils, and biomineralized skeletons Earliest eukaryotic life
Archean				~4560	Earliest life (prokaryotes)

Geologic time scale showing the eons, eras, and periods of the Phanerozoic; epochs of the Cenozoic Era; and major events in the history of life. Geologic ages are given in millions of years ago (Ma). The beginning of the Archean Eon is undefined at present. Geologic time units (depicted here) and their equivalent chronostratigraphic (time-rock) equivalents are as follows: eras/erathems, eons/eonothems, eras/erathems, periods/systems, and epochs/series.

insects account for at least 80% of the animal species on Earth.

A series of mass extinctions, each followed by a short recovery interval, altered the composition of marine communities in stepwise fashion through the Paleozoic. Notable extinctions occurred during the Cambrian, and at the end of the Ordovician, Devonian, and Permian periods. The end-Permian extinction is estimated to have decimated some 90% of marine life-forms. *See* EXTINCTION (BIOLOGY); PALEOZOIC.

Mesozoic. In the Mesozoic Era, terrestrial environments were dominated by vertebrates. Dinosaurs and mammals evolved in the Triassic Period, and dinosaurs underwent a series of radiations through the Mesozoic. One group of theropod dinosaurs evolved feathers and other features to become birds in the Jurassic. Land plants of the Mesozoic were dominated by gymnosperms until the Cretaceous, when flowering plants (angiosperms) underwent their initial diversification.

Mesozoic marine biotas were rich in shelly fossils. Swimming animals included ray-finned fishes, marine reptiles (especially ichthyosaurs, mosasaurs, and plesiosaurs), and ammonite and belemnoid cephalopods. Benthic communities saw great increases in the diversity of mollusks (especially oysters), echinoderms (especially sea urchins), and crustacean arthropods. In the Triassic, scleractinian corals formed reefs; in the Cretaceous,

reefs were formed by rudistid clams.

Major extinction events occurred near the end of the Triassic and at the end of the Cretaceous. The end-Cretaceous extinction had a devastating effect on life on Earth. Dinosaurs (other than birds), ammonite and belemnoid cephalopods, rudist clams, and most marine reptiles became completely extinct. Communities of marine brachiopods, mollusks, sea urchins, diatoms, and dinoflagellates suffered huge declines in species. *See* MESOZOIC.

Cenozoic. In the Cenozoic, land biotas were dominated by mammals, flowering plants, and insects, all of which diversified in the Paleogene Period, and marine biotas were dominated by mollusks, ray-finned fishes, sharks, mobile echinoderms, and microscopic foraminiferans. In the mid-Cenozoic, grasses diversified, providing an important food source for mammalian herbivores and omnivores. Primate diversification intensified in the Neogene Period, with hominids appearing first in the Miocene Epoch, and the genus *Homo* evolving in the Pliocene Epoch. *Homo sapiens* first appeared in the Pleistocene Epoch.

Major advances of continental glaciers occurred during various intervals of the Phanerozoic when large landmasses encroached on polar regions. In the Ordovician, Devonian, and Carboniferous-Permian, glaciers advanced from the south polar region of Gondwana. In the Neogene, glacial ice sheets expanded during two intervals in both the Northern and Southern hemispheres. The culmination of the most recent glacial age essentially marks the end of the Pleistocene Epoch and the beginning of the Holocene Epoch (or Recent). Overhunting by early humans may have partly contributed to the extinction of some large terrestrial mammals at the end of the Pleistocene. *See* CENOZOIC; GLACIAL EPOCH; GLACIAL GEOLOGY. Loren E. Babcock

Bibliography. R. H. Dott, Jr., and D. R. Prothero, *Evolution of the Earth*, 5th ed., 1994; S. J. Gould (ed.), *The Book of Life*, 1993; F. Gradstein, J. Ogg, and A. Smith, *A Geologic Time Scale 2004*, 2005; H. L. Levin, *The Earth through Time*, 7th ed., 2003; S. M. Stanley, *Earth System History*, 1999; R. Wicander and J. S. Monroe, *Historical Geology*, 1993.

Pharetronidia

A subclass of sponges of the class Calcarea in which the skeleton is formed of quadriradite spicules cemented together into a compact network, or includes an aspiculous massive basal skeleton formed of irregular calcitic spherulites (see **illus.**). Spicules shaped like tuning forks are characteristically present, and free triradiates, quadriradiates, and diactinal spicules may occur. *Murrayona* has a surface layer of scalelike spicules.

The characteristics of the pharetronidians fail to support their inclusion in either the calcinean or calcaronean line of the Calcarea, and they are best regarded as a separate subclass. They have left a rich fossil record that begins in the Permian and is best developed in the Mesozoic Era. Recent forms are typical inhabitants of cryptic habitats of Indo-Pacific coral reefs. *Petrobiona* occurs in caverns in the Mediterranean. *See* CALCAREA. Willard D. Hartman

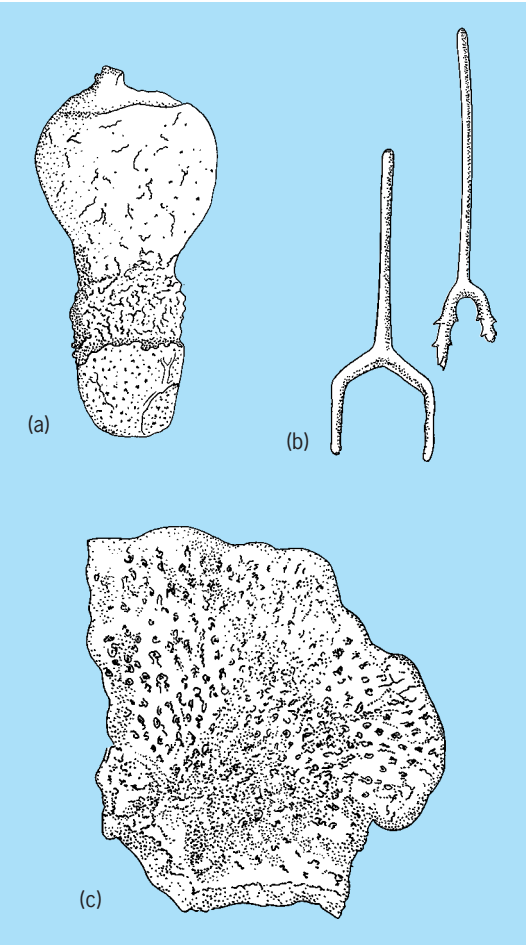

Representative pharetronidians. (*a*) *Petrobiona massiliana*. (*b*) Tuning-fork spicules of *Minchinella*. (c) *M. lamellosa*.

Pharmaceutical chemistry

The chemistry of drugs and of medicinal and pharmaceutical products. The important aspects of pharmaceutical chemistry are as follows:

1. Isolation, purification, and characterization of medicinally active agents and materials from natural sources (mineral, vegetable, microbiological, or animal) used in treatment of disease and in compounding prescriptions.

2. Synthesis of medicinal agents not known from natural sources, or the synthetic duplication, for reasons of economy, purity, or adequate supply, of substances first known from natural sources.

3. Semisynthesis of drugs, whereby natural substances are transformed by means of comparatively simple steps into products which possess more favorable therapeutic or pharmaceutical properties.

4. Determination of the derivative or form of a medicinal agent which exhibits optimum medicinal activity and at the same time lends itself to stable formulation and elegant dispensing.

5. Determination of incompatibilities, chemical and biological, between the various ingredients of a prescription.

6. Establishment of safe and practical standards, with respect to both dosage and quality, to assure uniform and therapeutically reliable forms for all medication.

7. Improvement and promotion of the use of chemical agents for prevention of illness, alleviation of pain, cure of disease, and search for new therapeutic agents, particularly where no satisfactory remedy now exists.

Chemistry in its various facets, along with cognate disciplines from the physical and biological sciences, is essential for pharmaceutical chemistry. It is difficult to characterize any individual procedure or reaction as solely or preponderantly pharmaceutical. For example, the procedure for converting alcohol to ether, an excellent general anesthetic, may with slight modification be adjusted for the production of ethylene, an important industrial chemical; and the step by which adiponitrile is hydrogenated to diaminohexane, a synthetic fiber intermediate, is useful for the conversion of streptomycin to dihydrostreptomycin. Drug standardization uses many techniques and methods of analytical chemistry.

In the nineteenth century, beginning with the isolation of morphine from opium by F. W. A. Sertüurner (1805), an apothecary from Einbeck, Germany, effort was primarily directed toward the isolation of the active constituents of plants; this resulted in the discovery and description of many alkaloids, glycosides, carbohydrates, volatile oils, and fixed oils. The terms plant chemistry and pharmaceutical chemistry were used interchangeably, and this field of investigation remains active today and promises to continue so for years come, because it is estimated that less than 3% of the known flora have been examined.

The food and drug legislation of 1906 led to the broadening of research activities to include the development of analytical procedures for the maintenance of the quality and potency of drugs. These endeavors are continuing and expanding, as can be seen in the quinquennial revisions of the *National Formulary* and the *U.S. Pharmacopeia*, both official standards, as supplemented by state and federal regulations and tentative methods proposed by agencies charged with the enforcement of these official standards. *See* PHARMACOGNOSY; PHARMACOLOGY; PHARMACY.

Walter H. Hartung

Bibliography. J. N. Delgado (ed.), *Wilson and Gisvold's Textbook of Organic Medicinal and Pharmaceutical Chemistry*, 10th ed., 1998; W. O. Foye, *Principles of Medicinal Chemistry*, 4th ed., 1995; United States Pharmacopeial Convention, Inc., *U.S. Pharmacopeia and the National Formulary*, USP XXII–NF XVII, 1990.

Pharmaceuticals testing

Techniques used to determine that pharmaceuticals conform to specified standards of identity, strength, quality, and purity. Conformance to these standards assures pharmaceuticals which are safe and efficacious, of uniform potency and purity, and of acceptable color, flavor, and physical appearance. Pharmaceuticals are medicinal products which are prescribed by medical doctors and dispensed through pharmacies and hospitals. They are usually taken orally or by parenteral injection.

Standards or specifications and their attendant procedures are designed to provide desired characteristics and acceptable tolerances for all raw materials, intermediates, and finished products. These standards thus provide an objective determination of whether pharmaceuticals are properly constituted. Two components are vital to the makeup of such standards: appropriate analytical procedures to permit a comprehensive examination, and a list of specifications to define the acceptable limits for each property tested. Standards are established by the pharmaceutical manufacturer, official compendiums (*U.S. Pharmacopeia* and the *National Formulary*), and regulations promulgated by the U.S. Food and Drug Administration (FDA).

The steps in the production cycle for pharmaceuticals must be rigidly and uniformly controlled so that each phase is completely accurate. The four control phases are generally designated as raw materials, manufacturing procedures, finished product testing, and control of identity.

Raw materials. These are usually referred to as components and are purchased on specifications. If the raw materials are officially recognized in the *Pharmacopeia* or the *National Formulary*, the specifications are provided in monographs in these compendiums. For raw materials which are not covered by these compendiums, specifications are prepared by the pharmaceutical manufacturer on the basis of requirements for the finished product. In the case of antibiotics there are federal regulations which specify acceptable characteristics of identity, strength, quality, and purity. These criteria are correlated with physical, chemical, biological, and microbiological examinations consisting of identity tests, safety tests, tests for limiting contaminants, and assay procedures. *See* ANTIBIOTIC.

Physical specifications include such characteristics as bulk density, mesh size, color, odor, extraneous contamination such as fibers, and homogeneity. Chemical specifications usually include such characteristics as chemical or physiological potency, melting point, boiling range, optical rotation, moisture, heavy-metals content, chemical identity, solubility, and presence of chemical contaminants. In addition, complex drugs cannot be adequately evaluated and controlled without such special instrumentation and skills as ultraviolet spectrophotometry, infrared spectrophotometry, nonaqueous titrimetry, column chromatography, gas chromatography, paper chromatography,

thin-layer chromatography, polarography, x-ray diffraction, x-ray fluorescence, spectrophotofluorimetry, and radioactive tracer techniques. *See* CHROMATOGRAPHY; POLAROGRAPHIC ANALYSIS; RADIOACTIVE TRACER; TITRATION; X-RAY DIFFRACTION.

Samples are taken upon receipt of a specific batch of raw material. These may be composite samples, composed of small portions from each container in the batch, or one or more random small portions of the entire batch. In some cases the sampling may be based on a statistical, or square-root, plan (except in shipments of five or less containers, when all containers must be sampled). The samples are tested to ensure conformance to each specification. Only after the raw material has been checked against each of the specifications can it be approved for use in pharmaceuticals.

Manufacturing procedures. To ensure products of the highest quality, pharmaceuticals must be manufactured under strictly regulated procedures and with adequate checks during each operation. Each dosage unit, though produced by different people from various raw materials, must furnish the patient with the exact kind and amount of drug specified on the label, in a safe and suitable formulation.

Pharmaceutical manufacturers must make products in conformance with the Good Manufacturing Practices as prescribed by the FDA. These regulations provide criteria on the following: buildings, equipment, personnel, components, master-formula and batch-production records, production and control procedures, product containers, packaging and labeling, laboratory controls, distribution records, stability, and complaint files. Proper production of drug products requires highly qualified and trained personnel, adequate manufacturing facilities and equipment, and competent testing laboratories. The quality-control system must provide regular and continuous use of all reasonable procedures, methods, and operations that are necessary to ensure uniform safety and effectiveness of the pharmaceutical.

Batch-production records must describe each manufacturing step in detail. The completion of each step or operation is initialed by the operator. Exact processing temperatures, specific mixing times, designated equipment, and precise details of operations, such as mixing sequence, filtration, or compression, are carefully specified on the batch-production record. All raw materials are double-checked for identity and quality before being incorporated in the process.

In-process assays are used to ensure homogeneity of mixing or completeness of a reaction in the manufacturing process. Such assays can range from simple pH measurements to complex chromatographic determinations. In accord with the FDA's Good Manufacturing Practices regulations, there must be adequate in-process controls, such as checking the weights and disintegration time of tablets, the fill of liquids, the adequacy of mixing, the homogeneity of suspensions, and the clarity of solutions. Upon completion of the manufacturing operation, batch-production records are checked by competent and responsible personnel for actual yield against theoretical yield of a batch of drug and to ensure that each step has been performed and signed for.

Representative samples of the pharmaceutical are taken by inspectors and submitted to the chemical or biological testing laboratory for final assays. Although the Good Manufacturing Practices regulations provide for testing adequately representative samples, the antibiotic regulations provide for a sampling ratio. Samples of the finished antibiotic dosage form must be collected by taking single tablets at intervals throughout the entire processing of the batch so that the quantities tableted during the sampling intervals are approximately equal. In no case may more than 5000 tablets be compressed during each interval of sampling. The nature of the production of a given batch of a pharmaceutical determines the method used to obtain an adequately representative sample. The regulations also provide for a reserve sample, of at least twice the quantity of drug required to conduct all the tests performed on the batch of drug, to be retained at least 2 years after distribution has been completed.

Finished-product testing. Finished pharmaceuticals must conform to appropriate standards of identity, strength, quality, and purity. Accordingly each batch of a pharmaceutical must satisfy five requirements: conformance with (1) the label claim for potency, (2) homogeneity standards, (3) standards of pharmaceutical elegance, (4) identity specifications, and (5) regulatory standards if they are applicable to the specific pharmaceutical.

Potency standards require that the drug meet label claims within specified limits. Monographs of the *Pharmacopeia* or the *National Formulary* usually specify maximum and minimum acceptance limits for official products. Limits for unofficial products are established by the manufacturer and usually parallel those for official products. New drugs and antibiotics are the subject of formal applications which must be approved by the FDA. A part of such application provides for specific acceptance limits on the potency of the product. In the case of antibiotics these standards are specified in regulations which are issued for every antibiotic preparation.

Potency assays vary in complexity from simple tests on pharmaceuticals containing only one active ingredient (ascorbic acid tablets) to very complex chemical and biological tests on pharmaceuticals containing two or more active ingredients (antibiotic and steroid combination tablets). Similarly, biological products must meet analogous standards for potency.

Some special types of pharmaceuticals require additional complex tests. All parenteral products, intended for injection, must meet sterility requirements. Tests are frequently required on parenteral products for absence of pyrogens and for safety (toxicity). In some cases ophthalmic solutions require sterility tests. Products such as parenterals and

ophthalmic preparations should be examined for conformance to appropriate standards on foreign particles. These additional tests are mandatory to obviate undesirable physiological reactions.

Statistical quality-control trend charts on certain characteristics of a pharmaceutical, such as tablet weights, ampul-fill volume, or random assay values, assure conformance to the appropriate standards. Stability tests, which are usually more complex and specific than the production potency tests, ensure the effectiveness and safety of the pharmaceutical during normal shelf life. At the same time such tests confirm the absence of any harmful degradation substances. As a result of these tests, many pharmaceuticals have a full-effectiveness expiration date on the label.

The most perfect pharmaceutical is no better than the container in which it is held. In accordance with the Good Manufacturing Practices regulations, suitable specifications, test methods, and cleaning and sterilization procedures must be used to assure that containers, closures, and other component parts of drug packages are suitable for their intended use. They must not be reactive, additive, or absorptive to any extent that significantly affects the identity, strength, quality, or purity of the drug and must furnish adequate protection against its deterioration or contamination.

Pharmaceutical elegance refers to the physical appearance of the dosage units. These standards include inspection to ensure that solutions are sparkling clear, tablets are not capped or chipped, parenteral ampuls are free of floaters, and colored products are of the right hue or shade. These standards govern physical quality.

Identity. Identity is the final requirement in pharmaceutical testing. Identity techniques guarantee proper labeling, that is, that the right product is in the right bottle with the right label. Serious consequences would result from a bottle of strychnine tablets carelessly labeled as saccharin tablets. To maintain the identity of the product, extensive checks are made throughout the manufacturing operation, including the use of duplicate label tags on all bulk goods, and very rigid controls are applied to printing, storage, and application of labels on finished pharmaceuticals.

Only when all operations in the production of pharmaceuticals, from securing raw materials to labeling the final container, are rigidly controlled through testing and checking procedures can it be assured that the pharmaceutical is pure, safe, and efficacious. *See* BIOASSAY; QUALITY CONTROL.

W. Brooks Fortune

Bibliography. American Pharmaceutical Association, *Evaluations of Drug Interactions*, 4th ed., 1988; H. C. Ansel, *Introduction to Pharmaceutical Dosage Forms and Drug Delivery Systems*, 5th ed., 1989; K. A. Connors et al., *Chemical Stability of Pharmaceuticals: A Handbook for Pharmacists,* 2d ed., 1986; *U.S. Pharmacopeia*, 22d ed., 1990; *U.S. Pharmacopeia*, *National Formulary*, XVII, 1990.

Pharmacognosy

The general biology, biochemistry, and economics of nonfood natural products of value in medicine, pharmacy, and other health professions. The products studied are of biologic origin, either plant or animal. They may consist of entire organs, mixtures obtained by exudation or extraction, or chemicals obtained by extraction and subsequent purification.

Pharmacognosy literally means knowledge of drugs, as do pharmacology and pharmacy. The center of interest in pharmacology, however, is on the mode of action of all drugs on the animal body, particularly on humans. In pharmacy major attention is directed toward provision of suitable dosage forms, their production and distribution. Pharmacognosy is restricted to natural products with attention centered on sources of drugs, plant and animal, and on the biosynthesis and identity of their pharmacodynamic constituents.

Sources of materials. Organs, or occasionally entire plants or animals, are dried or frozen for preservation and are termed crude drugs. They may be used medicinally in essentially this form, as in the case of the cardiac drug, digitalis, or the endocrine drug, thyroid, or as sources of mixtures or of chemicals obtained by processes of extraction.

Mixtures obtained by exudation from living plants include such drugs as opium, turpentine, and acacia. Processes of extraction are required to obtain such mixtures as peppermint oil (steam distillation), podophyllum resin (percolation), and parathyroid extract (solution). For a discussion of classes of natural products with medically significant members of this type *see* ESSENTIAL OILS; FAT AND OIL (FOOD); GUM; TERPENE; WAX, ANIMAL AND VEGETABLE.

Pure chemicals may be extracted from a crude drug (for example, the glycoside digitoxin from digitalis or the hormone insulin from pancreas), from a mixture obtained by exudation (for example, the alkaloid morphine from opium), or from an extracted mixture (for example, the terpene menthol from peppermint oil). For a discussion of natural products of this type *see* ALKALOID; GLYCOSIDE; HORMONE.

Vitamins as a class of natural products are within the scope of pharmacognosy, although many are obtained commercially by laboratory synthesis. Included also are antibiotics and biologicals (serums, vaccines, and diagnostic biological products). *See* ANTIBIOTIC; VITAMIN.

The general biology of pharmacognosy is largely descriptive. It includes the taxonomic position of the natural source of the product, the part of the plant or animal yielding the drug, the scientific and common names of the biologic source, the gross and histologic anatomic characterization of the part used, and the principal uses of the product in the health professions.

The biochemistry is both descriptive and experimental. It includes the chemical nature and percentage of the medically significant constituent, the mechanisms of biosynthesis of the constituent, and

the role of the constituent in the economy of the plant. Attention is also given to mechanisms of biosynthesis by the use of radioactive precursors of medically active constituents to follow biosynthesis step by step. The isolation and chemical identification of new, potentially useful plant and animal constituents are an important aspect of biochemical research in pharmacognosy.

The economic aspects include the discovery and study of natural sources of crude drugs and their derivatives, development of cultivated sources where feasible, improvement of the yield of useful constituents, and protection of medically useful crop plants from their natural enemies. Methods of harvesting, drying, curing or other processing treatment, storing, packaging, and shipping enter into the commerce of drugs.

For a single drug, for example, menthol from peppermint oil, these several aspects are frequently inseparable. The commercial grower must know the species of *Mentha* yielding commercially profitable quantities of the essential oil and of menthol, percentage yields from various species, conditions suitable for growth, natural enemies of the growing plant such as specific viruses, methods of extraction of the oil and of isolation of the menthol, and proper conditions for packaging, storing, and shipping the purified drug.

Uses of materials. Medical uses are chiefly as therapeutic, prophylactic, or diagnostic agents. Prior to the twentieth century the materia medica of all countries was preeminently of natural products; it still is in some underdeveloped countries. Research in the twentieth century has contributed many synthetic and semisynthetic drugs to the modern materia medica, but a significant number of the crude drugs are still the drugs of choice in therapy or serve as the source of widely used purified mixtures or chemicals. Digitalis and its glycosides, the alkaloids of opium and belladonna, penicillin, thyroid, insulin, and poliomyelitis vaccine are examples.

Pharmaceutical uses are chiefly in the production of palatable and stable dosage forms: gums and mucilages in emulsions and suspensions, starch and lactose in tablets, sugar and essential oils in elixirs, oils and waxes in ointments. Many natural products of insignificant or questionable therapeutic value continue to be used in home remedies.

Uses in other health professions include antiseptics, protectives, and local anesthetics used by dentists; rodenticides, insecticides, and other pesticides used in the protection of the public health; and a variety of prophylactic and therapeutic agents used by veterinarians.

A large number of natural products of value in the health professions are used also in cosmetology (essential oils, gums, fats, and waxes), in the culinary arts (spices, essential oils, and condiments), and in industry (naval stores, mucilages, fats, and waxes).

The role of a medically active chemical produced and used by the animal organism is usually well understood. Physiologic function within the organism and therapeutic use by humans are usually closely related, as in the case of pepsin, thyroid, or the sex hormones.

Corresponding knowledge of medically active plant constituents is almost nonexistent. The role of menthol in the economy of *Mentha piperita*, of digitoxin in *Digitalis purpurea*, of morphine in *Papaver somniferum*, or of reserpine in *Rauwolfia serpentina* is unknown. Through the centuries, however, it has been discovered that certain plants relieve the symptoms of or cure certain diseases. With the discovery of alkaloids in the early nineteenth century and of the other major classes of medically active plant constituents, the chemicals responsible for therapeutic actions have been identified one by one, but not the function of these chemicals in their respective plant sources.

Types of materials. Classes of therapeutic agents have frequently been discovered by study of biosynthesized medicinal chemicals. Most such classes, in fact, have been developed from chemicals orginally known from crude drugs or from their exudates or extractives. The first uses of opium as a narcotic and analgesic drug are lost in antiquity, but its position in the medical practice of the day has been primary for over 2000 years. Morphine was among the first alkaloids to be isolated and has been widely used for more than 150 years; from its study has developed a class of analgesic drugs of wide application in medical practice. *See* ANALGESIC; MORPHINE ALKALOIDS.

An analogous pattern has given the modern classes of hypotensive drugs and tranquilizers. The Indian drug rauwolfia after centuries of use in folk medicine eventually found its way into scientific medical practice in the Orient. Study of its chemical derivatives during the 1930s and 1940s revealed the presence of many alkaloids, one of which, reserpine, was shown to be an effective antihypertensive agent. Subsequent therapeutic use of reserpine and other rauwolfia products demonstrated the tranquilizing action. A large class of drugs having hypotensive action, tranquilizing effect, or both developed rapidly. Various species of the genus have been characterized morphologically, and there has been intensive study of practical methods of culture. *See* TRANQUILIZER.

Another major class of modern drugs of natural origin, the antibiotics, has been developed largely since the beginning of World War II. The prototype, penicillin, was discovered in part as a result of fortuitous accident, but the many other commercially available antibiotics have been developed as a result of carefully planned systematic search.

Not infrequently the clue that has led to collection and scientific investigation of a crude drug as a possible source of medically significant constitutents has been use of the drug by natives for a nonmedical, but to them desirable, purpose such as narcosis or as a poison against wild animals or humans. The use of opium as a narcotic by the laity undoubtedly preceded its medical use. Coca was chewed or sucked by the Indians of South America to increase endurance; condemned by the Spanish conquerors, it nevertheless was introduced into medical practice in Europe. Discovery of the local anesthetic action

of the alkaloid cocaine led to the development of a new and important class of therapeutic agents, the local anesthetics. *See* COCA; NARCOTIC; OPIATES.

A second native drug from South America, curare, was first used by the Indians as an arrow poison to kill wild animals for food. The neuromuscular paralysis caused in game by the drug suggested therapeutic use as a muscular relaxant. Studies of the plants yielding crude curares revealed several species of two principal genera. *Strychnos* and *Chondodendron*, as the main sources. A number of crystalline alkaloids were isolated from crude curares; eventually the alkaloid tubocurarine was identified as having the therapeutic potentialities suggested by the paralyzing action of the native drug. The botanical source was established as *C. tomentosum*.

Comparable studies of the African arrow poisons inee and kombe added ouabain and strophanthin to the class of cardioactive glycosides, of which digitoxin is the most widely used. The arrow poisons are prepared from African species of *Strophanthus* and are used by natives of both the eastern and western coasts.

The alkaloid physostigmine, also from an African poison and useful in the treatment of glaucoma, was discovered as a result of use of its plant source as a human poison in the trial by ordeal of those accused of offenses. The alkaloid is the toxic constituent of the seeds of *Physostigma venenosum*, which were fed to the accused. Toxic symptoms were taken as evidence of guilt; those who vomited the material were considered guiltless.

Synthetic materials. Development of synthetic drugs related chemically to the active constituent of a natural product has frequently followed investigation of native use of the natural product as drug or poison. The objective of such development is usually to produce a drug having fewer undesirable side effects while retaining the useful therapeutic action. Substitutes for morphine, reserpine, cocaine, tubocurarine, and physostigmine are among a host of synthetic drugs which accomplish the objective to a greater or lesser degree and whose discovery depended on study of natural products.

Intermediates useful in the laboratory synthesis of drugs often exist as therapeutically inactive chemicals in natural products. Plants and animals biosynthesize many such compounds with chemical structures similar to, but not identical with, medicinally useful substances. A slight change in molecular configuration may yield a potent therapeutic agent. A simple example is pinene, a chemical abundant in turpentine oil and convertible by laboratory procedures into camphor. The resulting "synthetic" camphor is actually semisynthetic and possesses the therapeutic and most other properties of natural camphor.

An important class of natural intermediates are the steroids, widely distributed in both plants and animals. Some chemical variations are active physiologically and as drugs, for example, sex and adrenal cortical hormones. Natural sources, glands of domesticated animals used as food by humans, are not available in quantities adequate to fulfill the drug needs for these products. Many plants contain steroids suitable as intermediates for sex hormones. Natural intermediates readily converted into adrenal cortical hormones are uncommon, and extensive search for such steroids has been made. Field studies involve collection and identification of plants judged to be potentially good sources of steroids, preliminary extraction and determination of the presence or absence of these intermediates, and further collection, drying, and preserving of larger quantities of promising species. *See* STEROID.

Systematic screening of plants of reputed therapeutic value and indigenous to a country or other restricted geographic area is a costly and time-consuming procedure—a major reason it has been done for relatively few regions. Notwithstanding, such surveys give promise of uncovering new sources for known classes of drugs, adding to knowledge of such little-known classes of hallucinogens and anticarcinogenic drugs, developing entirely new classes of therapeutic agents and providing profitable natural sources of intermediates useful in drug synthesis. *See* ANTIMICROBIAL AGENTS; BIOCHEMISTRY; BIOSYNTHESIS; PATHOLOGY; PHARMACOLOGY; PHARMACEUTICAL CHEMISTRY; PHARMACY; PLANT PHYSIOLOGY.

Richard A. Deno

Bibliography. W. A. Creasy, *Drug Disposition in Humans: The Basis of Clinical Pharmacology*, 1979; G. E. Trease and W. C. Evans, *Pharmacognosy*, 12th ed., 1983; V. E. Tyler et al., *Pharmacognosy*, 9th ed., 1988.

Pharmacology

The science of detection and measurement of the effects of drugs or other chemicals on biological systems. The effect of chemicals may be beneficial (therapeutic) or harmful (toxic) when administered to humans, other mammals, or other living systems. The pure chemicals or mixtures may be of natural origin (plant, animal, or mineral) or may be synthetic compounds.

The broad area covered may be conveniently divided into a number of categories: chemotherapy, the use of chemicals to destroy invading organisms such as bacteria and molds in or on the host; pharmacotherapy, the use of drugs to restore or replace normal function in various tissue cells, organs, or integrated units; pharmacodynamics, studies on the mechanism of action of drugs which may utilize physiological, biochemical, or electrical techniques; toxicology, the study of the poisonous effects of chemicals; psychopharmacology, the study of the effects of chemicals on the behavior of humans or animals; biochemical pharmacology, the effects of chemicals on biochemical reactions in living systems, and the effects of these systems on the chemicals, that is, their metabolism; structure-activity relationship, relationship of biological activity to chemical structure and molecular properties; and clinical pharmacology, the study and

evaluation of the effects of drugs in humans. *See* CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; PATHOLOGY; TOXICOLOGY.

The chemicals which have a beneficial effect in humans (such as restoring function or behavior toward the normal state, relieving pain, destroying harmful invading organisms, or aiding in the diagnosis of disease) are called drugs; those chemicals which produce only harmful effects are called poisons. All drugs may be poisonous, however, if administered in large enough amounts. *See* POISON.

It is a fundamental property of drugs that their effects increase in intensity as the dose is increased. Determination of this relationship delineates the dose-response curve. This relationship can be usefully summarized by the dose required to produce a standardized effect and also by the slope of the curve. In many situations the numerical expression of the dose required to produce 50% of the maximal response permits comparison of drugs having the same type of action. This ED_{50} (dose effective at the 50% level) is usually the standard-effect dose used to compare the potency of drugs having the same action, since it can be determined with greater precision than other doses producing a smaller or larger percentage of the maximal effect. Measurement of toxicity as manifested by lethal effects is expressed similarly as the LD_{50} (lethal dose 50). The relative safety of drugs is estimated by means of the therapeutic index, which is the ratio of the LD_{50} to the ED_{50}. The higher the ratio, the safer the drug. *See* EFFECTIVE DOSE 50; LETHAL DOSE 50.

The necessity to use such terms as ED_{50} and LD_{50}, which are averages, implies that individuals have been found to differ in the amount of drug that each will require to produce the desired effect. It is common knowledge that various individuals tolerate widely varying quantities of ethyl alcohol. Individual differences occur in the response to all drugs, and what may be an effective dose in one person may not be effective in another. Individuals may also differ qualitatively in their response to a drug. In some cases these differences have been shown to be due to genetic differences. This has led to expanded research of genetic influences on drug actions, which is generally referred to as pharmacogenetic research.

Depending upon the effect under study, placebos (pills or other preparations containing starch, sugar, and coloring matter but no pharmacologically active agent) may also produce a favorable response in some individuals. The administration of a placebo to a person with hypertension may result in a temporary fall in blood pressure, for example; and in several studies of agents used to relieve pain, as many as 35% of the individuals tested obtained relief of pain from the administration of the placebo. Observations such as these have pointed out the necessity of using a placebo control or a positive control (a drug of established activity) in experiments designed to evaluate a new drug. Because the attitude of the physician who administers a drug may also influence the way patients respond, this variable is controlled by use of the double-blind technique; that is, neither the individual nor the doctor knows which sample is drug and which is placebo.

The use of animals has been and is absolutely essential to progress in pharmacology and other medical sciences; however, different species may differ markedly in their response to drugs both quantitatively and qualitatively. Drugs which produce sedation in one species may produce excitement in another. Fortunately, most chemicals produce similar effects in many species, including humans. Unfortunately, not all drugs found effective in animals are effective in humans, presumably because of differences in metabolism or excretion or unrecognized differences in the nature of the disease process or the physiological mechanism of maintenance of the same function in different species. Study of these differences has yielded much basic information on mechanism of actions of drugs and on physiological and biochemical differences among species. *See* BIOASSAY; PHARMACEUTICAL CHEMISTRY; PHARMACY.

Charles J. Kensler

Bibliography. A. Gilman et al. (eds.), *Goodman and Gilman's The Pharmacological Basis of Therapeutics*, 9th ed., 1996; B. G. Katzung, *Basic and Clinical Pharmacology*, 8th ed., 2000; D. H. Robertson and C. R. Smith, *Manual of Clinical Pharmacology*, 1981.

Pharmacy

The health profession concerned with the discovery, development, production, and distribution of drugs. Drugs are substances (other than devices) used to diagnose, prevent, cure, or relieve the symptoms of disease. For relations to closely allied fields *see* MEDICINE; PHARMACEUTICAL CHEMISTRY; PHARMACOLOGY.

General practice. This part of the profession is carried on in exclusive prescription pharmacies, semiprofessional pharmacies, and drug stores. It consists of compounding and dispensing drugs on order of the physician, dentist, or veterinarian; serving as consultant on drugs to the health professions and to the public; and selling other health supplies such as antiseptics, bandages, and home remedies. Combination of nonprofessional with professional activities is customary in the United States, however, and is commonly termed retail pharmacy.

Hospitals. In addition to qualities characteristic of general pharmacy practice, a hospital pharmacy includes special administrative features, provision of drugs for nursing stations, manufacturing of pharmaceutical preparations, teaching of nurses and medical and pharmacy interns, service to the hospital committee on pharmacy and therapeutics, preparation and revision of a hospital formulary, and monitoring the drug regimen of the individual patient (clinical pharmacy). The pharmacist may have charge of investigational drugs, radioactive pharmaceuticals, medical and surgical sterile supplies, and gaseous drugs for inhalation therapy.

Research. One type of research is in pharmaceutical chemistry, synthetic if the objective is to produce new and improved drugs by laboratory procedures, and analytic if the objective is to provide improved methods of assay for quality control of pharmaceutical production. Research may be in product development, aimed at provision of more palatable, stable, economical dosage forms. It may be on drugs from natural sources (pharmacognosy), or on the mode of action of drugs from any source (pharmacology). It draws heavily on investigations in organic chemistry, biochemistry, and microbiology.

Manufacturing. The pharmaceutical industry is the mass-production medium of drugs and suitable dosage forms. It produces some of its own raw materials, but depends on the closely allied chemical industry for most of them. It sponsors and conducts extensive pharmaceutical research.

Pharmaceutical distribution and promotion are specialized activities within the pharmaceutical industry which are now commonly called pharmacy administration; this term also includes the commercial phases of retail pharmacy.

Jurisprudence. This is a highly specialized area in pharmacy. General pharmacy practitioners and hospital pharmacists are subject to federal and state laws on drugs and pharmacy practice. A board of pharmacy serves as the law enforcement agency and the examining and licensing body in each state; various federal agencies administer the provisions of the several federal laws governing pharmacy.

Other special fields of pharmacy include journalism, administration of national and state associations, and provisions of modern drug standards.

Education. From 1932 to 1960 the education of pharmacists consisted of collegiate studies for at least 4 years in basic science, mathematics, general education, and professional areas. In 1960 the minimal educational program was extended to a total of 5 years of collegiate study. A flood of new and complex drugs, released at the rate of hundreds per year since the end of World War II, has made necessary for the pharmacist a strong background in the physical and biological sciences, highly specialized training in drugs, and the sociologic and humanistic understanding desirable in all professional people. Richard A. Deno

Bibliography. *Remington's Pharmaceutical Sciences*, 17th ed., 1985; A. Wertheimer, *Pharmacy Practice*: *Social and Behavioral Aspects*, 3d ed., 1989.

Pharynx

A chamber at the oral end of the vertebrate alimentary canal, leading to the esophagus. Because of divergent specializations in the various classes of vertebrates, it cannot be described in general terms except at embryonic stages. In adult humans it is divided anteriorly by the soft palate into a nasopharynx and an oropharynx, lying behind the tongue but anterior to the epiglottis; there is also a retropharyngeal compartment, posterior to both epiglottis and soft palate. The nasopharynx receives the nasal passages and communicates with the two middle ears through auditory tubes. The retropharynx leads to the esophagus and to the larynx, and the paths of breathing and swallowing cross within it. In adult fishes, the pharynx is not segregated from the mouth cavity but is pierced by a number of paired gill slits. *See* ESOPHAGUS; LARYNX; PALATE.

Embryology. Shortly after the germ layers of the embryo are in place, the pharyngeal cavity appears as a simple enlargement of the anterior end of the endoderm tube. Its lateral or lateroventral walls promptly become thickened as a series of paired pillars, the pharyngeal segments (branchial arches or visceral arches), separated from their fellow by paired pouches which push outward from the endodermal lining of the cavity. These pharyngeal pouches are approached on each side by corresponding branchial grooves which push inward from the head ectoderm. The matching grooves and pouches may meet. Their touching surfaces may then become thinned as closing plates, and they may actually break through as pharyngeal clefts. The mouth perforates into the pharyngeal cavity in a similar way anteroventrally. This embryonic appearance of the pharynx as an enlarged anterior gut chamber whose walls are marked by pouches, grooves, clefts, and intervening pillarlike segments, is one of the few characteristics found in all members of the phylum.

The more primitive vertebrate embryos, such as those of lampreys and sharks, have up to seven or even more pairs of large open pharyngeal clefts of fairly uniform size (**Fig. 1**). Embryos of reptiles, birds, and mammals show fewer segments and sharp reduction of the more posterior ones (**Fig. 2**). Also, several of the last pharyngeal pouches fail to break through as clefts. Whereas in most fishes several posterior clefts become enlarged and the intervening segments become equipped with gills and valves during development, few vestiges of either pouches or

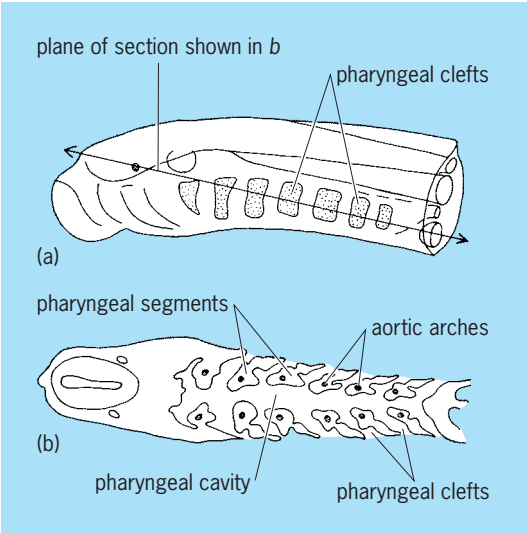

Fig. 1. Lamprey embryo. (*a*) Side view and (*b*) horizontal section of pharynx region.

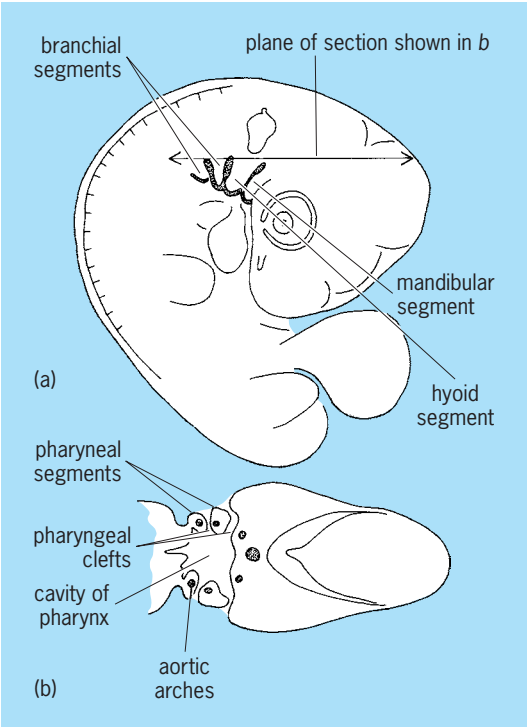

Fig. 2. Chick embryo. (*a*) Side view and (*b*) horizontal section of pharynx region.

clefts remain at adult stages of land animals. Nevertheless, in all vertebrates the solid tissues of the pharyngeal segments give rise to numerous structures of the head and neck.

Each pharyngeal segment of the embryo is lined by ectoderm (the future epidermis) on the outside, and by endoderm (the future mucous membrane) on the inside. The mesoderm enclosed between these layers is of hypomeric or lateral-plate origin, and differentiates chiefly into muscles and arteries. In addition to these constituents, tongues of neural crest cells migrate down into the pharyngeal segments and differentiate there into skeletal elements, except in cyclostome fishes. *See* NEURAL CREST; RESPIRATORY SYSTEM.

Pharyngeal derivatives. The first pair of pharyngeal segments, called the mandibular segments, enclose the mouth between them, and their skeleton-forming cells produce the mandibular cartilages, which are the earliest rudiments of the lower jaw. Their muscle-forming cells differentiate into a dorsal group of muscles whose principal function in all gnathostome vertebrates is to snap the jaws shut, and a ventral group which supports the tongue by forming a sheet between the halves of the lower jaw. All these muscles are supplied by the motor division of the trigeminal nerve.

The second pair is called the hyoid segments. They arise just posterior to the mandibular segments, usually separated from them temporarily by hyomandibular clefts. They lie ventral to the ear vesicles on the sides of the embryonic head. Their skeleton-forming cells form important parts of the tongue support, and their muscles are used in some vertebrates for opening the mouth, for tongue manipulation, for facial expression, and other functions. The motor division of the facial nerve innervates the hyoid group of muscles. In higher vertebrates the hyomandibular pouch is involved in the formation of the auditory tube and the middle ear space. *See* EAR (VERTEBRATE).

The third and all more posterior pairs are called branchial segments since in fishes they usually form gill-bearing arches; they tend to be repressed or dispersed in the development of the higher vertebrates. Their skeleton-forming cells either form a succession of jointed rodlike bones or cartilages for the support of gills, as in fishes, or join the hyoid skeleton in tongue support and concentrate ventrally in the larynx cartilages, as in terrestrial vertebrates. In fishes the muscles derived from the first branchial segment (the third in the pharyngeal series) are strictly innervated by the glossopharyngeal nerve, and those of all the rest of the segments by individual branches of the vagus nerve. They function for manipulation of the gills, and for grinding and swallowing food. In higher vertebrates, the branchial segments become less and less distinct during development and the nerve supply, while still derived from the glossopharyngeal and vagus nerves, is not so clearly segmental, since the striated muscles themselves become arranged in a more or less continuous pharynx-constrictor sheet or congregate in the laryngeal cartilages, serving nerve mechanisms of swallowing and sound production. *See* SPEECH.

Histology. The pharynx is in general lined by simple mucous membrane and backed by fibrous connective tissue and a double layer of striated muscle. In bony fishes its skeletal arches may be thickly studded with simple teeth, or may even be developed into grinding or crushing plates. Terrestrial vertebrates show simple tubular glands emptying, usually in great numbers, through the mucous membrane and tonsillar collections of lymphoid tissue in the submucosa. *See* TONSIL.

Gross derivatives. Elaborate gill pouches are developed in all aquatic groups and subject to many special adaptations. They differ sharply in design in lampreys, hagfishes, cartilaginous fishes, and the bony fishes. A median ventral evagination from the posterior end of the pharynx gives rise to the entire respiratory system of the land vertebrates, including lungs, larynx, and trachea. A similar but often dorsal evagination, usually from the pharynx-esophagus boundary, gives rise to the air bladder in bony fishes. *See* SWIM BLADDER.

The epithelium of the pharyngeal pouches and of the pharynx floor produces a constellation of endocrine glands and other structures. *See* PARATHYROID GLAND; THYMUS GLAND; THYROID GLAND; ULTIMOBRANCHIAL BODIES. William W. Ballard

Defects and disorders. Congenital defects of the human pharynx that are commonly seen are malformed or split uvulae, or soft palates, and extension of a cleft palate backward to the pharyngeal region. *See* CLEFT LIP AND CLEFT PALATE.

Inflammations may be local or part of a systemic involvement. Acute pharyngitis may be caused by almost any irritant and typically does not involve an infection by a microorganism. Acute follicular pharyngitis is caused by infectious bacteria, usually streptococci; it is also called septic sore throat. Acute tonsillitis involves the masses of lymphoid tissue found in the back of the pharynx. The tonsils may also be the seat of peritonsillar abscesses. All of these inflammations may persist as subacute or chronic diseases, but usually the more prolonged forms occur as a result of repeated attacks or low-grade, persistent infections. *See* STREPTOCOCCUS.

The most common benign tumor of this region is the papilloma; hemangiomas, fibromas, and other less common nonmalignant growths are also found here. The malignant tumors of the pharynx include several varieties of carcinoma and sarcoma, particularly lymphosarcomas in children. *See* ONCOLOGY.

Nervous disorders seen not infrequently are paresthesias (abnormal sensation) and hyperesthesia (increased sensation). Neuralgia of the glossopharyngeal nerve is marked by severe pain in the neck-ear-jaw region. Motor disorders which affect swallowing may originate from local irritations or from central nervous system disease, as in the case of rabies. *See* SOMESTHESIS. Edward G. Stuart; N. Karle Mottet

Bibliography. W. Becker et al., *Ear, Nose, and Throat Diseases: A Pocket Reference*, 1994; K. Haskins (ed.), *Nose and Paranasal Sinuses: Larynology and Bronchoesophagology*, 1994; E. Jarvik, *Basic Structure and Evolution of Vertebrates*, 2 vols., 1981.

Phase (astronomy)

The changing fraction of the disk of an astronomical object that is illuminated, as seen from some particular location. The monthly phases of the Moon are a familiar example (see **illus.**). When the Sun is approximately on the far side of the Moon as seen from Earth (conjunction), the dark side of the Moon faces the Earth and there is a new moon. The phase waxes, beginning with crescent phases, as an increasing fraction of the illuminated face of the Moon is seen. At quadrature, when half the visible face of the Moon is illuminated, the phase is called the first-quarter moon, since the Moon is now one-quarter of the way through its cycle of phases. The waxing moon continues through its gibbous phases until it is in opposition; the entire visible face of the Moon is illuminated, the full moon. During the full moon, the Moon and the Sun are on opposite sides of the Earth, a configuration known as a syzygy. Then the Moon wanes, going through waning gibbous, third-quarter, and waning crescent phases until it is new again. The cycle of moon phases takes approximately 29.53 days and explains the origin of the word month.

The Earth, Moon, and Sun are not directly in line at the times of new moon and full moon, because the Moon's orbit around the Earth is inclined 5° to the plane of the Earth's orbit around the Sun. When they are directly in line, a lunar eclipse occurs when the Earth's shadow falls on the Moon, and a solar eclipse when the Moon's shadow falls on the Earth. *See* ECLIPSE; MOON.

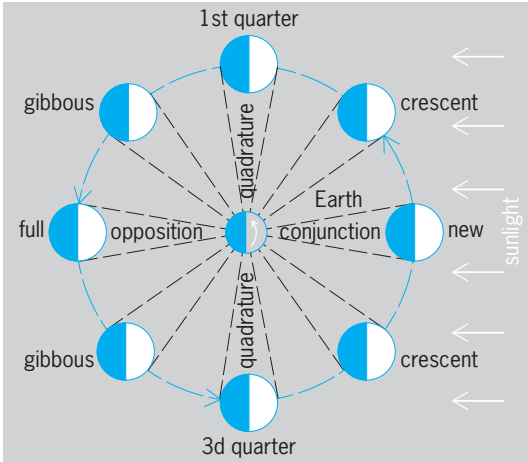

From Earth, different fractions of the illuminated half of the Moon are seen at different times as the Moon goes through a 29.53-day cycle of phases.

Galileo discovered the phases of the planet Venus when he observed the sky with his telescope in 1610. Giovanni Zupus discovered the phases of the planet Mercury in 1639. Because of the angle at which the outer planets are seen from Earth, and because of their great distance, they do not appear to go through phases as seen from Earth. Mars, Jupiter, Saturn, Uranus, and Neptune have now all been seen by spacecraft as crescents. Jay M. Pasachoff

Phase (periodic phenomena)

The fractional part (portion) of a period through which the time variable of a periodic quantity (alternating electric current, vibration) has moved, as measured at any time instant from an arbitrary time origin. In the case of a sinusoidally varying quantity, the time origin is usually assumed to be the last point at which the value of the quantity was zero, passing from a negative to a positive value. It is customary

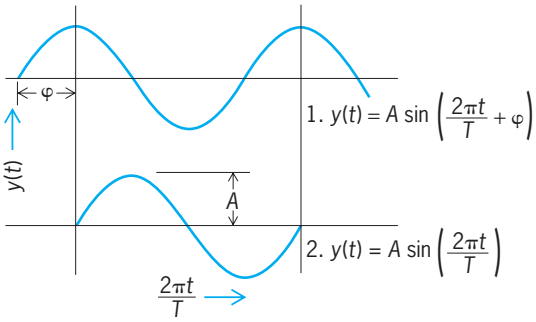

Illustration of the meaning of phase for a sinusoidal wave. The difference in phase between waves 1 and 2 is φ and is called the phase angle. For each wave, *A* is the amplitude and *T* is the period.

to choose the origin so that the fractional part of the period is less than unity.

In comparing the phase relationships at a given instant between two time-varying quantities, the phase of one is usually assumed to be zero, and the phase of the other is described, with respect to the first, as the fractional part of a period through which the second quantity must vary to achieve a zero of its own (see **illus.**). In this case, the fractional part of the period is usually expressed in terms of angular measure, with one period being equal to 360° or 2π radians. Thus two sine waves of a given frequency are said to be 90°, or $\pi/2$, out of phase when the second must be displaced in time, with respect to the first, by one-fourth period in order for it to achieve a zero value. *See* PHASE-ANGLE MEASUREMENT; SINE WAVE.

William J. Galloway

Phase-angle measurement

Measurement of the time delay between two periodic signals. The phase difference between two sinusoidal waveforms that have the same frequency and are free of a dc component can be conveniently described as shown in **Fig. 1**. It can be seen that the phase angle can be considered as a measure of the time delay between two periodic signals expressed as a fraction of the wave period. This fraction is normally expressed in units of angle, with a full cycle corresponding to 360°. For example, in Fig. 1, where the voltage v_1 passes through zero $^1/_8$ cycle before a second voltage v_2, it leads by 360°/8 or 45°. Phase angle is usually defined from the fundamental component of each waveform; therefore distortion of either or both signals can give rise to errors, the extent of which depends on the nature of the distortion and the method of measurement. *See* DISTORTION (ELECTRONIC CIRCUITS).

Many techniques have been developed to measure phase differences or to generate signals with known phase differences.

Three-voltmeter method. This method, also known as the law-of-cosines method, can be used when the voltages involve a common point. **Figure 2***a* shows

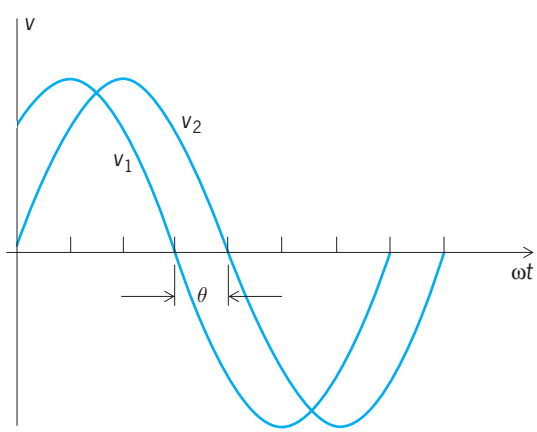

Fig. 1. Phase angle θ between voltages v_1 and v_2.

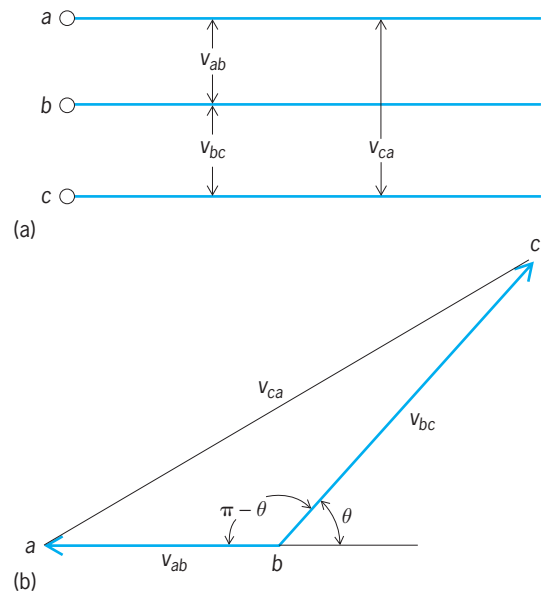

Fig. 2. Voltages employed in three-voltmeter method. (*a*) Circuit diagram. (*b*) Vector diagram.

three terminals *a*, *b*, and *c*. If a high-impedance voltmeter is used to measure the amplitudes of the signal voltages v_{ab} and v_{bc} and the differential voltage between them v_{ca}, they can be plotted to give a triangle, as in Fig. 2*b*. The phase angle θ can then be determined from the law of cosines in trigonometry, as shown by Eq. (1).

$$\theta = \cos^{-1}\left(\frac{v_{ab}^2 + v_{bc}^2 - v_{ca}^2}{2v_{ab}v_{bc}}\right) \qquad (1)$$

See TRIGONOMETRY.

Zero-crossing-detector phase meter. Many modern phase-measuring devices are based on the use of zero-crossing detectors. The time at which each signal crosses the zero-voltage axis is determined, usually by means of a squaring-up circuit (for example, an overdriven amplifier) followed by a high-speed comparator. This produces, in each channel, a trigger pulse that is used to drive a bistable flip-flop. The output of the flip flop is a rectangular wave, the duty cycle of which is proportional to the phase difference between the input signals. If this signal is integrated by means of a suitable filter, a dc voltage is produced that is an analog representation of the phase angle. This voltage is then displayed on a panel meter (analog or digital) suitably scaled in degrees or radians. Instrumentation using this principle is capable of measuring phase differences to approximately ±0.05° over a wide range of amplitudes and frequencies. **Figure 3** is a timing diagram illustrating the principles involved. *See* AMPLIFIER; COMPARATOR; ELECTRIC FILTER; MULTIVIBRATOR; SWITCHING CIRCUIT; WAVE-SHAPING CIRCUITS.

FFT phase meter. If the waveforms to be measured are sampled at a number of equally spaced points within each cycle and these samples are digitized, a fast Fourier transform (FFT) algorithm can be used to compute, among other parameters, the phase difference between the signals. The advantage of this

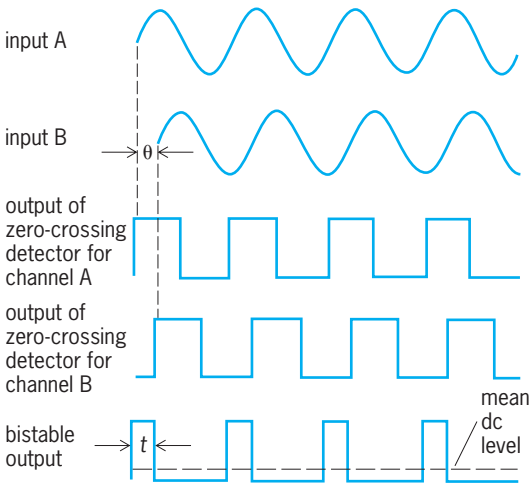

Fig. 3. Zero-crossing-detector timing diagram.

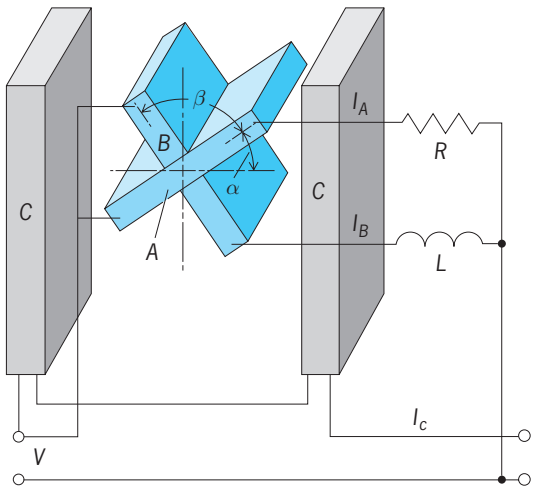

Fig. 4. Crossed-coil phase meter.

technique is that only the phase of the fundamental component needs to be evaluated, so that possible effects of distortion are virtually eliminated.

Crossed-coil phase meter. Also known as the Tuma phase meter, this is the basic element of power-factor meters and of synchroscopes. When used for power-factor indication, it contains two movable coils A and B on a common shaft (**Fig. 4**). The two coils move as a unit, the angle β between them remaining fixed. There is no restraining spring acting between them (as there is in a conventional electrodynamic meter movement), and the system will revolve as long as the currents produce an average nonzero torque. *See* POWER-FACTOR METER.

The meter also contains a fixed coil C, which carries the load current. The fixed coil is made in two sections so that its magnetic field will be nearly uniform in the vicinity of the movable-coil assembly.

The instantaneous force induced on coil A is proportional to the product of the instantaneous currents in coils A and C and therefore to the sine of the angle between the planes of A and C. If the currents in A and C are sinusoidal, with phase displacement θ, the average torque T is proportional to expression (2), and this can also be expressed as relation (3).

$$(I_A \sin \omega t) I_C \sin (\omega t + \theta) \cos \propto \qquad (2)$$

$$T \propto I_A I_C \cos \theta \cos \propto \qquad (3)$$

The movable coil A is placed in a resistive circuit so that its current I_A is in phase with the applied voltage V. The other movable coil B is placed in an inductive circuit so that its current lags the voltage by approximately 90°.

In practice, the circuit of coil A cannot be made completely noninductive and, equally, the circuit of coil B cannot be nonresistive. The phase angle between the currents in coils A and B is therefore less than 90°. The scaling of the phase meter is constructed to take this into account.

If the current in coil B leads the current in fixed coil C by the angle θ, and the current in coil B lags the current in coil A by the angle ϕ, then the current in coil A will lead the current in the fixed coil by the angle $\theta + \phi$, and the average torque on A is proportional to $\cos (\theta + \phi) \cos \alpha$, whereas the average torque on B is proportional to $\cos \theta \cos (\alpha + \beta)$ and in the opposite direction. If these torques do not cancel, the shaft will rotate ($\propto$ will vary) until the two become equal, or when Eq. (4) is satisfied.

$$\cos (\theta + \phi) \cos \propto = \cos \theta \cos (\alpha + \beta) \qquad (4)$$

From this equation it follows that if β equals ϕ, then $\propto$ equals θ, and the phase meter indicates directly the phase angle between the sources supplying coil C and the crossed coils. A change of 5° in phase angle produces a deflection of 5° of the phase-meter pointer.

Oscilloscope methods. There are many methods available that use an oscilloscope to measure phase difference. The easiest and most convenient is to use a dual-channel oscilloscope, one signal being applied to each channel. The phase difference may readily be calculated from the relative positions of the two waveforms on the oscilloscope screen. *See* OSCILLOSCOPE.

Phase-order indicators. These devices are used to indicate which phase voltage of a polyphase circuit leads or lags another. If the voltage vectors of a polyphase circuit are as indicated in **Fig. 5*a***, the voltage from neutral to line 2 reaches a maximum $^1/_3$ of the wave period after the voltage of line 1 and $^1/_3$ of the wave period before the voltage of line 3. The

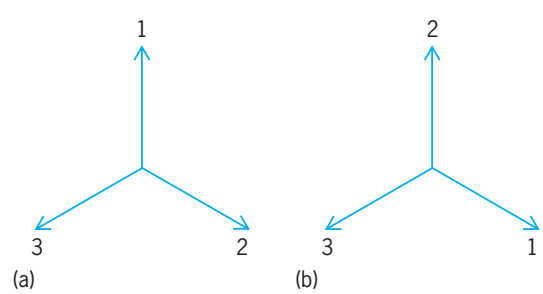

Fig. 5. Phase sequence. (*a*) Sequence 1-2-3. (*b*) Sequence 1-3-2.

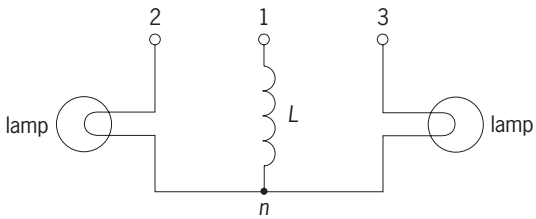

Fig. 6. Circuit diagram of phase-sequence indicator.

phase order (or sequence) is then said to be 1-2-3.

Relative motion between the armature conductors and the magnetic field induces the voltages in an alternator. Therefore, if the alternator were to be rotated in the opposite direction, the order in which the phase voltages reached maximum would be reversed. The phase sequence would then be 1-3-2, as shown in Fig. 5*b*.

A miniature three-phase motor designed to rotate clockwise when connected to a three-phase system possessing a phase sequence of 1-3-2 is used as a phase-sequence indicator. Counterclockwise rotation indicates a 1-2-3 phase sequence.

A common type of phase-order indicator consists of an inductance and two lamps connected in a Y configuration to the three-phase lines as in **Fig. 6**. In the absence of the inductor, equal currents would flow through the lamps; however, the inductor produces a 90° phase shift in the current injected at point *n*. This produces an imbalance in the current through the lamps, the overall effect being to increase the current in lamp 2 when the phase sequence is 1-2-3, and to increase the current in lamp 3 when the phase sequence is 1-3-2.

Phase-relation indicators. Frequently employed in the electrical utility industry, these devices are used to indicate when two generators or sources of alternating voltage are in phase with one another. If two voltages reach maximum at the same time, they are in phase. If two generators are to be connected in parallel, they should have the same voltage, frequency, and phase. A voltmeter and a tachometer (or frequency meter) can be used to indicate equality of the voltage and frequency. The phase relation between the two sources is shown by means of phasing lamps or by means of a synchroscope or synchronizer.

Phasing lamps. Phasing lamps placed across the open switch used to parallel two generators will often suffice to indicate an in-phase condition. Depending on the relative phase of the two generators, the lamp voltage varies from the sum to the difference of the generator voltages. As the frequencies of the two generators approach one another, the lamps flicker, changing from full brightness to dim at a decreasing rate. If the two frequencies are equal, the lamp will maintain a fixed brilliance. Usually the oncoming generator is set with a slightly higher frequency so that it will provide additional power rather than be an additional load on the system. As the lamps slowly go through the dim phase, the switch is closed, connecting the incoming generator to the system.

Synchronizer or synchroscope. This is based on the Tuma phase meter. The current in the fixed coil is supplied by one generator; the current in the movable crossed-coils is supplied by the other generator. If the two generators are in synchronism, their frequencies are equal and the movable coils will take up a position depending on their relative phase angle. If the frequencies are different, the phase will continue to vary and the movable coils will rotate, the direction of rotation indicating which generator is running faster. As in the case of phasing lamps, the incoming generator is generally running at a slightly higher speed; the two are connected together when the synchroscope pointer drifts past the zero mark. *See* SYNCHROSCOPE.

High-frequency measurement. Conventional phase meters, such as the zero-crossing-detector type, have an upper frequency limit of a few hundred kilohertz. This limit is imposed mainly by the ability of the arrangement consisting of a comparator and a flip-flop to maintain a clean and precise rectangular waveform under conditions of high-speed operation. In order to measure phase angle at frequencies between about 100 kHz and several gigahertz, it is necessary to down-convert the radio-frequency signals to a frequency that can be handled correctly by the phase meter. A superheterodyne system that can fulfill this function is shown in **Fig. 7**. *See* HETERODYNE PRINCIPLE.

Sampling systems are sometimes used to achieve the same result. Any phase difference caused by the mixers can be measured by applying zero phase angle (for example, via a power splitter) and adjusting a compensating trimmer in one channel for a zero reading. In order to ensure that phase errors due to radio-frequency mismatch are minimized, these instruments normally terminate the radio-frequency signals in a 50-ohm load. At radio frequencies, a major source of error can arise due to differences in cable length, and users of such equipment must ensure that such effects are minimized.

At microwave frequencies, instruments such as slotted lines, air lines, and vector network analyzers are also used for phase-angle measurements. *See* MICROWAVE MEASUREMENTS.

Phase-angle generation techniques. Many techniques have been developed for generating signals of known phase difference. A classic method involves establishing two signals in quadrature and

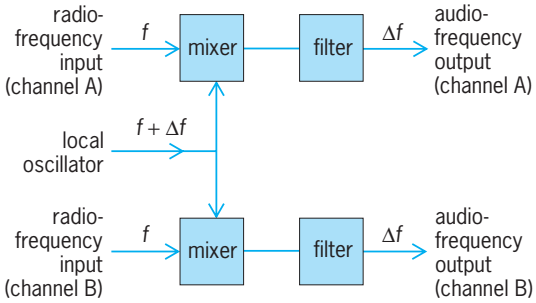

Fig. 7. High-frequency phase measurement using down-conversion technique.

mixing them together via calibrated voltage dividers to produce the required phase angle. Such methods have been superseded by modern digital techniques, which offer greater accuracy, versatility, and ease of use.

Although the various digital techniques differ in detail, the principles of operation are all similar. A clock and address counter are used to sequentially address the memory locations, thus producing bit patterns that, when applied to a digital-to-analog converter, will produce a stepped approximation to the stored waveform. Each step is maintained for one clock period, at the end of which the next set of digital data is applied to the digital-to-analog converter. A low-pass filter is used to remove the steps. If two memories are used, then a dual-channel generator is available. A phase displacement between the two signals can then be obtained by either deliberately offsetting the relative memory locations or recalculating the data contained in each memory location in accordance with the required phase angle. *See* DIGITAL-TO-ANALOG CONVERTER.

Such techniques are used in national metrology laboratories to provide phase-angle generating capabilities with uncertainties as good as $\pm 0.01^\circ$. Commercial instrumentation using similar techniques is also available. *See* ELECTRICAL MEASUREMENTS.

J. Hurll

Bibliography. H. M. Berlin and F. C. Getz, Jr., *Principles of Electronic Instrumentation and Measurement*, 1988; Dranetz Engineering Laboratories Inc., *Application Handbook for Precision Phase Measurement*, 1975; C. G. Coombs, *Electronic Instrument Handbook*, 3d ed., 1999; A. D. Helfrick and W. D. Cooper, *Modern Electronic Instrumentation and Measurement Techniques*, 1990; L. Schnell (ed.), *Technology of Electrical Measurements*, 1993; L. M. Thompson, *Electrical Measurements and Calibration: Fundamentals and Applications*, 2d ed., 1994.

Phase-contrast microscope

A microscope used for making visible differences in phase or optical path in transparent or reflecting specimens. It is an important instrument for studying living cells and is used in biological and medical research.

Microscopy of transparent objects. When a light wave passes through an absorbing object, it is reduced in amplitude and intensity. Since the human eye and the photographic plate are sensitive to variations in intensity, such an object will give a visible image when viewed through an ordinary microscope. A perfectly transparent object does not absorb light, so that the intensity remains unaltered, and such an object is essentially invisible in an ordinary microscope. However, the light that has passed through the transparent object is slowed down by it and arrives at the eye a minute fraction of a second later than it would otherwise have done. Such delays, or in technical language, phase changes or differences in optical path, are not detected by the eye, and the fundamental problem in the microscopy of transparent objects, of which living cells are important examples, is to convert them into visible intensity changes. For further discussions of basic principles *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; DIFFRACTION; INTERFERENCE OF WAVES; LIGHT.

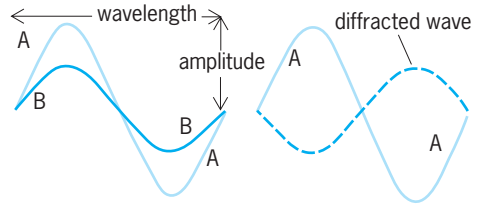

Fig. 1. Effect of a partially absorbing object on a light wave. The diffracted wave is half a wavelength out of phase with the incident wave. A = incident wave; B = transmitted wave.

Simple theory. A consideration follows of what happens when a light wave passes through a partially absorbing object, such as a stained biological specimen. In **Fig. 1**, let A be the incident wave. When the wave passes through the object, energy is absorbed so that the transmitted wave B is reduced in height or amplitude. The intensity is proportional to the square of the amplitude. Wave B can also be represented as the sum of the incident wave A and another wave shown by a dotted line. This wave has a real physical existence and represents the light scattered or diffracted by the object. The German physicist E. Abbe was the first to stress the importance of diffraction in image formation and showed that the final image was formed by interference or addition of the incident and diffracted light. If the incident wave and the diffracted wave are added together algebraically, that is, upward displacements being treated as positive and downward ones as negative, the resultant is wave B. In the case of a partially absorbing object, therefore, the trough of the diffracted wave coincides with the crest of the incident wave, that is, the two waves differ in phase by half a wavelength.

Consideration of a perfectly transparent refractile object follows. Since no light is absorbed, the height of the wave is unchanged, but if the object has a higher refractive index than its surroundings, the wave will be delayed and can be represented by wave B in **Fig. 2**. Once again, wave B can be represented as the sum of the incident wave A and a diffracted wave (Fig. 2*b*). This diffracted wave is no longer half a wavelength out of phase with wave A. If wave B is only slightly delayed, the diffracted wave is approximately one-quarter of a wavelength out of phase with A. Suppose now that the incident wave A and the diffracted wave could somehow be separated and their phase relationships altered; in particular, suppose the phase difference between them could be increased from about one-quarter of a wavelength to half a wavelength, as in Fig. 2*c*. The crest of one wave would now coincide with the trough of the other, and the new resultant wave would be indistinguishable from wave B in Fig. 1. In other words, the

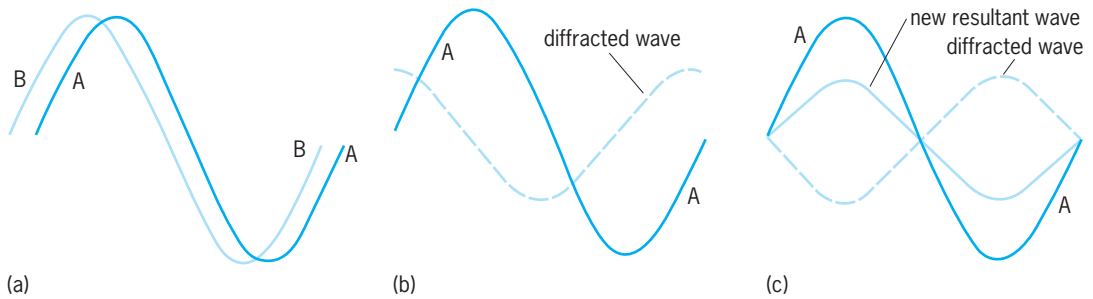

Fig. 2. Effect of a transparent object on a light wave. (*a*) Incident wave A is unaltered in amplitude but delayed in phase and is transmitted as wave B. (*b*) B can again be represented as the sum of the incident wave A and a diffracted wave. The diffracted wave is no longer half a wavelength out of phase with A, but if the phase difference can be changed to half a wavelength (*c*) the new resultant wave will resemble that produced by a partially absorbing object.

invisible transparent object would be indistinguishable from a partially absorbing object. This was first realized by Frits Zernike, who was awarded the 1953 Nobel Prize in physics for his invention of phase contrast.

Practical realization. The essential features of a phase-contrast microscope are shown in **Fig. 3**. The practical problem is to find some way of separating the incident or direct light from that diffracted by the object. This is done by placing a diaphragm D of easily recognizable shape, such as an annulus, at the front focal plane of the substage condenser C. Light from each point of the focal plane passes as a parallel pencil of rays through the specimen S and is brought to a focus at the rear focal plane P of the objective O. Thus, on removing the eyepiece, an image of the annulus will be seen at the back of the objective lens. This image corresponds to the incident light. In addition, when a specimen is present, some light is diffracted by it and spreads out to fill the whole of the back lens of the objective. Thus, apart from the small area of overlap over the image of the annulus, the direct and diffracted waves are essentially separated at the plane P. A phase plate is now inserted at this level. This can be a transparent disk with an annular groove of such dimensions that it coincides exactly with the image of the diaphragm D. All the direct light now passes through the groove in the phase plate, whereas the diffracted light passes mainly outside the groove. Since the diffracted light has to pass through a greater thickness of transparent material than the direct light, a phase difference, depending on the refractive index of the phase-plate material and on the thickness of the groove, is introduced between them. If this phase difference is about one-quarter of a wavelength, the basic conditions for phase contrast will have been achieved (Fig. 2). If the phase plate is made to retard the incident wave by a quarter of a wavelength, the crests and troughs of the two waves will coincide, giving a resultant of greater amplitude. Refractile details will appear bright (negative contrast) instead of dark (positive contrast). The commercial Anoptral system is a type of negative phase contrast.

In practice, a partially absorbing layer of metal or dye is usually deposited over the phase-plate annulus. Figure 2*c* shows that the amplitude of the direct wave is usually greater than that of the diffracted wave so that, although the resultant wave is reduced in amplitude, it is still far from zero, and a transparent object will not appear perfectly black, but gray. If the amplitude of the direct wave is reduced after it has passed through the object, it can be made equal to that of the diffracted wave so that the resultant is zero, and the object appears perfectly black. Unfortunately, in order to match the amplitude of all possible diffracted waves, it would be necessary to make a variable-absorption phase plate. Such a device is

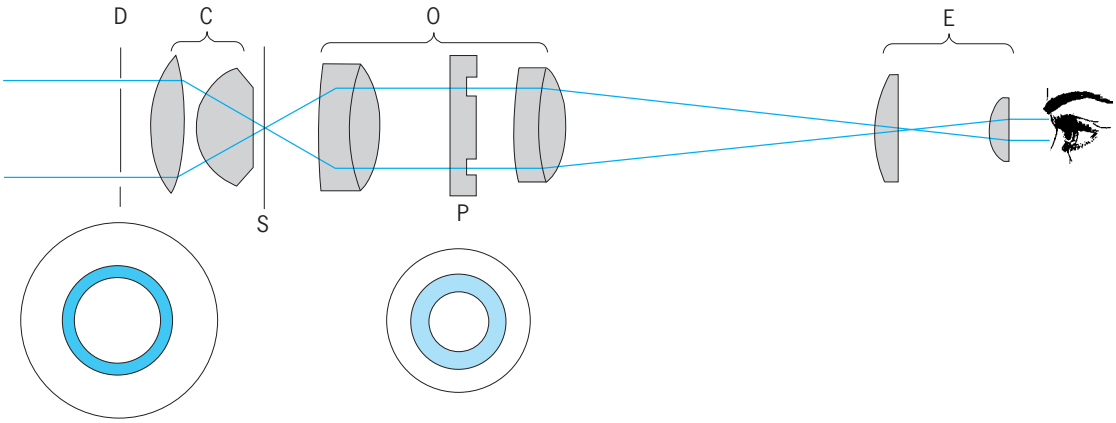

Fig. 3. Diagram of a phase-contrast microscope. D = annular diaphragm; C = substage condenser; P = phase plate at rear focal plane of objective; S = specimen; E = eyepiece.

expensive, and in practice most manufacturers use fixed phase plates with absorptions of between 50 and 80%. Each objective must have a corresponding substage diaphragm. The commonest arrangement is a rotatable wheel in the substage condenser containing three or four diaphragms of different sizes and also a clear circular aperture which enables the instrument to be used with conventional illumination.

Image interpretation. The phase change or optical path difference ϕ introduced by an object is defined by the relationship $\phi = (n_o - n_m)t$, where n_o is the refractive index of the object, n_m that of the surrounding medium, and t the object thickness; ϕ is usually expressed either in terms of the wavelength, in which case t must also be expressed in the same unit, or in angular measure. In the latter notation, one wavelength is equal to 360° or 2π radians. Although the phase-contrast microscope converts variations in ϕ into variations in intensity, the relationship between these quantities is not a simple one. In general, as ϕ increases, the image becomes darker; but beyond a certain value of ϕ, it becomes lighter again, and finally, for very large values of ϕ, the contrast actually becomes reversed, so that the image becomes brighter than the background. Thus, it cannot be assumed that because one part of the image appears darker than another, it will necessarily correspond to a region of greater optical path. Even these basic relationships are disturbed because the theoretical requirement of complete separation between the direct and diffracted light can never be achieved in practice. This results in the appearance of a bright halo around every dark detail and an accentuation of edges and sharp discontinuities in the object. Although these effects do not allow the instrument to be used for quantitative measurements of optical path, they are not really disadvantageous for purely observational work.

Biological applications. The phase-contrast microscope is a routine instrument for the examination of living cells. It made it possible to study the structure of living cells under excellent optical conditions and with no loss in resolving power. Accurate observations thus became much easier to make, and in particular, the use of phase-contrast cinemicrography made it possible to study changes in cell structure during the movement and division of cells with great clarity. The method is also useful for the study of unstained tissue sections and for the comparison of material in the electron and optical microscopes.

A quantitative application is microrefractometry. It follows from the basic definition of phase change, namely $\phi = (n_o - n_m)t$, that if the refractive index n_m of the mounting medium is made equal to that of the object n_o, ϕ becomes zero irrespective of the object thickness t, and since the intensity is a function of ϕ, the object will become invisible. The phase-contrast microscope can thus be used as a very sensitive null indicator for measuring refractive indices. The principle is to immerse the object in a series of media of graded refractive index until one is found that makes the object invisible. This is one of the most sensitive methods of microrefractometry available, and it can

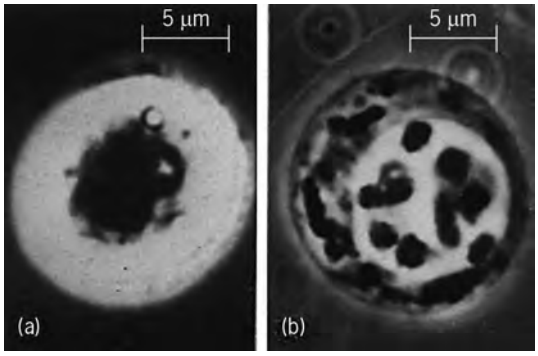

Fig. 4. Phase-contrast photograph of living locust spermatocyte. (*a*) In a saline medium. The refractive index of the cell is so much higher than the index of the medium that all internal detail is obscured. (*b*) Similar cell immersed in 9% bovine plasma albumin solution. The refractive index difference is much less, and the internal nucleoplasm and chromosomes can be clearly seen.

be extended to the quantitative refractometry of living cells **(Fig. 4)**. In this case, a suitable nontoxic medium must be used. The most suitable medium is a concentrated protein solution such as Armour's bovine plasma albumin fraction V, with added salts. The importance of cell refractometry rests on the fact that there is a linear relationship between the refractive index n of a solution and the concentration C of the dissolved substance. Thus, $n = n_o + \alpha C$ where n_o is the refractive index of the solvent, for example, water, and α is a characteristic constant known as the specific refraction increment. The mean value of α for most cellular constituents may be taken as 0.0018, and the formula can be used to calculate C, the total solid concentration in the cell. The water concentration can also be deduced, and if the cell volume is known, the dry and wet mass of the cell is also known. *See* INTERFERENCE MICROSCOPE.

Nonbiological applications. The study of transparent specimens such as crystals and fibers calls for little comment except to point out that, if the specimen is birefringent, a single polarizer must be used in order to avoid a confused image.

Industrial applications include the examination of chemicals, oils, waxes, soaps, paints, foods, plastics, rubber, resins, emulsions, and textiles.

Surface structure can often be studied by phase contrast with vertical illumination. Either the specimen must be naturally reflecting, or a thin film of reflecting material must be deposited on it or on a surface replica. A typical arrangement is shown in **Fig. 5**. A condenser lens forms an image of a lamp on an annular diaphragm. A semireflecting cover slip is placed just above the objective. A projector lens and the objective combine to form an image of the annulus at a plane just above the semireflector, and a phase plate is inserted here (compare Fig. 3). Normally, it is desirable to place the phase plate at the rear focal plane of the objective; but if this were done here, the incident light would have to pass through the metalized part of the phase plate and would lose intensity. Stray light and glare would also occur because of reflections at the surface of the phase plate. Apart from these modifications, the system is similar to that for

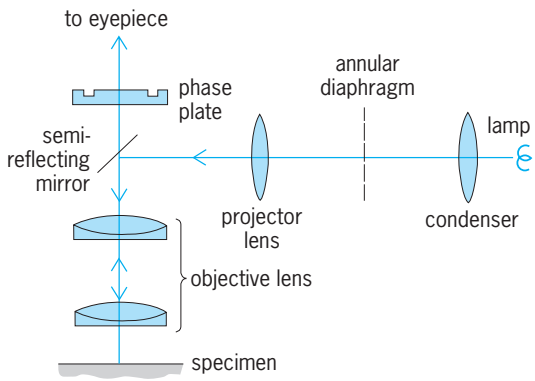

Fig. 5. Phase-contrast microscope with vertical illumination for reflecting specimens.

phase contrast with transmitted light. Differences in height in the reflecting specimen will produce differences in optical path length, and these will be converted into intensity differences. Robert Barer

Bibliography. A. Briggs, *Acoustic Microscopy*, 1992; T. G. Rochow and P. Tucker, *Introduction to Microscopy by Means of Light, Electrons, X-Rays, or Acoustics*, 2d ed., 1994; G. Wade (ed.), *Acoustic Imaging: Cameras, Microscopes, Phased Arrays, and Holographic Systems*, 1976.

Phase equilibrium

A general field of physical chemistry dealing with the various situations in which two or more phases (or states of aggregation) can coexist in thermodynamic equilibrium with each other, with the nature of the transitions between phases, and with the effects of temperature and pressure upon these equilibria. Many superficial aspects of the subject are largely qualitative, for example, the empirical classification of types of phase diagrams; but the basic problems always are susceptible to quantitative thermodynamic treatment, and in many cases, statistical thermodynamic methods can be applied to simple molecular models.

Thermodynamics requires that when two phases, α and β, are free to exchange heat, mechanical work, and matter (chemical species), the temperature T, the pressure P, and the chemical potential (partial molar free energy) μ_i of each particular component i must be equal in both phases at equilibrium. Algebraically, equilibrium exists when $T_\alpha = T_\beta$, $P_\alpha = P_\beta$, $\mu_{i,\alpha} = \mu_{i,\beta}$.

These conditions of thermal, mechanical, and material equilibrium need not all be present if the equilibrium between phases is subject to inhibiting restrictions. Thus, for a solution of a nonvolatile solute in equilibrium with the solvent vapor, the condition of equality of solute chemical potentials $\mu_{2,\alpha} = \mu_{2,\beta}$ need not apply, since there can be no solute molecules in the vapor phase. Similarly, in osmotic equilibria, in which solvent molecules can pass through a semipermeable membrane, whereas solute molecules cannot, $\mu_{1,\alpha} = \mu_{1,\beta}$ and $T_{1,\alpha} = T_{2,\beta}$, but the solute chemical potentials μ_2 are unequal, as are the pressures on opposite sides of the membrane. *See* OSMOSIS; SOLUTION.

If a system consists of P phases and C distinguishable components, there are $C + 2$ thermodynamic variables (C chemical potentials μ_i, plus the temperature and pressure) which are interrelated by an equation for each phase. Since there are P independent equations relating the $C + 2$ variables, one needs to fix only $F = C + 2 - P$ variables to define completely the state of the system at equilibrium; the other variables are then beyond control. This relation for the number of degrees of freedom F, or variance, is called the phase rule and was first derived by Willard Gibbs in 1873. It has proved to be a powerful tool in interpreting and classifying types of phase equilibria.

When chemical changes may occur in the system, the number of components C is the number of independent components, that is, the number of components whose amounts can be varied by the experimenter; this is equal to the total number of chemical species present less the number of independent chemical equilibria between them.

An invariant system has no degrees of freedom ($F = 0$), for which the number of phases $P = C + 2$. For a one-component system, such an invariant point is a triple point at which three phases coexist at a single temperature and pressure only; for a two-component system, a quadruple point (four phases) would be invariant. *See* TRIPLE POINT.

In a univariant system ($F = 1$), $P = C + 1$. With a one-component system, one can fix the temperature at which two phases (liquid and gas, for instance) can coexist in equilibrium; then the pressure (here the vapor pressure) is determined and not subject to external control. A univariant system is described by a line in a phase diagram, for example, a plot of vapor pressure versus temperature. The differential equation (1) for such a univariant line in a one-component

$$\frac{dP}{dT} = \frac{\Delta H}{T\,\Delta V} \qquad (1)$$

system was first deduced by B. P. E. Clapeyron in 1834. In the equation ΔH and ΔV are the enthalpy change (heat absorbed) and volume change, respectively, for the transition from one phase to another. *See* VAPOR PRESSURE.

In systems of two or more components, more complicated equations for the univariant line replace the Clapeyron equation, but in special cases they may reduce to the simpler form. For example, in the chemical decomposition of a solid calcium carbonate to solid calcium oxide and gaseous carbon dioxide, reaction (2), the three-phase system (two solids and

$$CaCO_3(s) \rightleftharpoons CaO(s) + CO_2(g) \qquad (2)$$

one gas) is univariant (since the number of independent components C is two), the equilibrium pressure is the decomposition pressure of calcium carbonate ($CaCO_3$), and the ΔH and ΔV of the Clapeyron equation are the enthalpy and volume changes associated with the chemical reaction.

Binary Systems

Phase diagrams and binary systems containing two components are easily classified. Typical examples of the important classes (liquid-gas, liquid-liquid, solid-liquid, solid-solid) have been selected for description.

Liquid-gas equilibrium. In a one-component system, liquid and vapor are in equilibrium at the boiling point. For two-component systems, the two-phase situation is bivariant and more complex. A complete temperature-pressure-composition diagram would be three-dimensional, so most phase diagrams are made for either constant pressure or constant temperature. **Figure 1** shows the simplest type of binary liquid-vapor temperature-composition diagram, exemplified by the system carbon tetrachloride–stannic chloride, which forms essentially ideal solutions. The regions labeled G and L are one-phase regions, gas (vapor) and liquid, respectively; the region labeled $L + G$ is a two-phase region in which liquid and vapor coexist. If the temperature of a liquid mixture of 40 mole percent $SnCl_4$ (mole fraction = 0.40) is increased at a constant pressure of 1 atm, the change in the system can be traced along the straight line $ab'cd''e$. At low temperatures, only one phase, the liquid, is present, but at 87.5°C (190°F), point b', a vapor phase appears. The composition of this vapor phase is given by point b'' (mole fraction $SnCl_4 = 0.18$), and the two conjugate phases are connected on the diagram by the tie line $b''b'$. As the temperature is increased further, more vapor is formed; since the vapor is rich in CCl_4, this component becomes relatively depleted in the liquid phase, and the liquid composition moves along the line $b'c'd'$, while the vapor composition moves along the line $b''c''d''$.

At 90°C (194°F) the overall composition of the two-phase system is represented by point c, but the compositions of vapor and liquid separately are given by

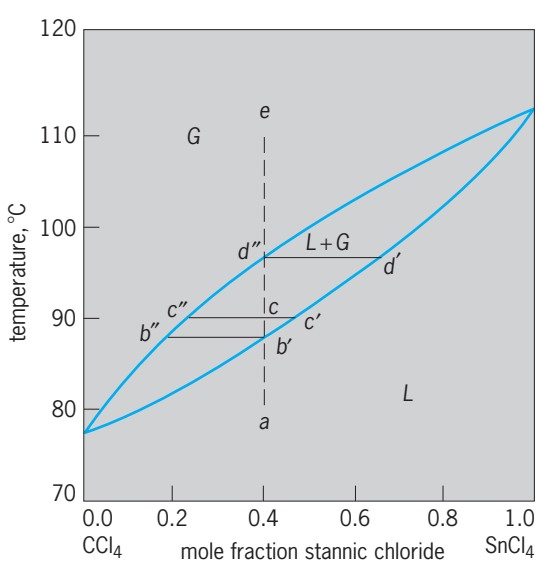

Fig. 1. Binary liquid-vapor temperature-composition diagram for the system carbon tetrachloride + stannic chloride. G = gas (vapor) phase; L = liquid phase; $L + G$ = two coexisting phases. Pressure = 1 atm (10^2 kPa).

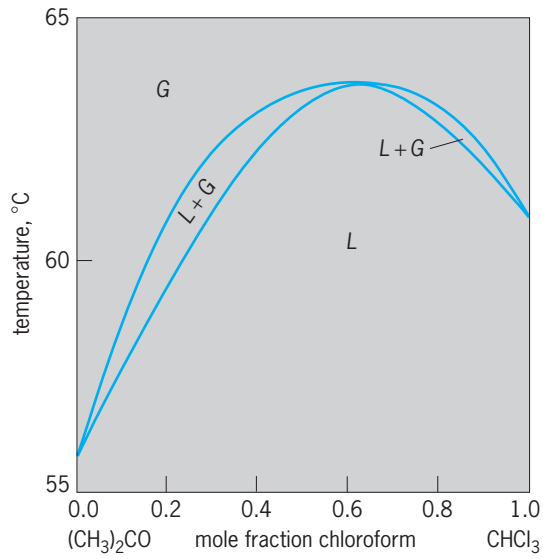

Fig. 2. Temperature-composition diagram for acetone + chloroform, showing maximum boiling point. Pressure = 1 atm (10^2 kPa).

the two ends of the tie line, points c'' and c', respectively (mole fractions of 0.22 and 0.47, respectively). The relative amounts of the two phases are given by the lever arm principle of physics. The ratio of the moles of vapor to moles of liquid is given by the ratio of the length cc' to the length $c''c$, here 0.07:0.18, or 28%, in the vapor phase. Further increase in temperature produces more and more vapor until, at 97°C (207°F), the liquid phase (point d', mole fraction 0.62) has become vanishingly small; at higher temperatures, it disappears, and only the vapor phase (point d'', mole fraction = 0. 40) remains. Further increase in temperature (along the line $d''e$) is uneventful.

In this simple system, there are no maxima or minima in the liquid and vapor curves; consequently, such liquid mixtures can be separated completely into the two pure components by fractional distillation. Systems with maximum boiling points (acetone + chloroform; **Fig. 2**) or minimum boiling points (ethanol + benzene; **Fig. 3**) cannot be so separated into the pure substances. At the maximum or minimum, the composition of the liquid is identical with that of the vapor with which it is in equilibrium; continued boiling will not alter these compositions, so that these solutions are called constant-boiling mixtures or azeotropes. It should be noted that the solid and liquid lines are smooth curves tangent to each other at such a point; any phase diagram which shows a sharp corner is thermodynamically incorrect. *See* AZEOTROPIC MIXTURE; DISTILLATION.

Maximum boiling mixtures are associated with negative deviations from ideal behavior; this type of deviation usually arises from strong attractions between molecules of the different species (sometimes called compound formation). Minimum boiling mixtures are associated with positive deviations from ideal behavior; this usually arises when the attraction between two unlike molecules (1-2) is weaker

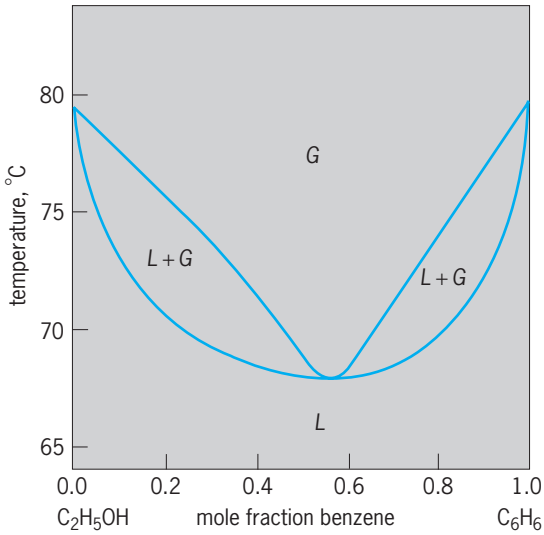

Fig. 3. Temperature-composition diagram for ethanol + benzene, showing minimum boiling point.

than the average of two like pairs (1-1 and 2-2); extreme examples of this arise when one component may be described as associated. The simpler type of phase diagram (Fig. 1) occurs when the two components mix nearly ideally or when the boiling points are very different.

Liquid-liquid equilibrium. When two liquids are sufficiently different in their intermolecular forces, they may not mix in all proportions but instead be only partially miscible, part of the phase diagram being occupied by a two-phase region of two immiscible liquid phases. Most liquids (but not all) become more miscible as the temperature increases and are completely miscible at the critical solution temperature (also called the consolute temperature). The system aniline + *n*-hexane (**Fig. 4**) is a typical example of such liquid-liquid immiscibility and critical phenomena. The liquid-liquid phase boundary and the critical solution temperature are only slightly dependent upon pressure. The size of the two-phase region increases with pressure if (as is usual) the two liquids expand on mixing at constant pressure.

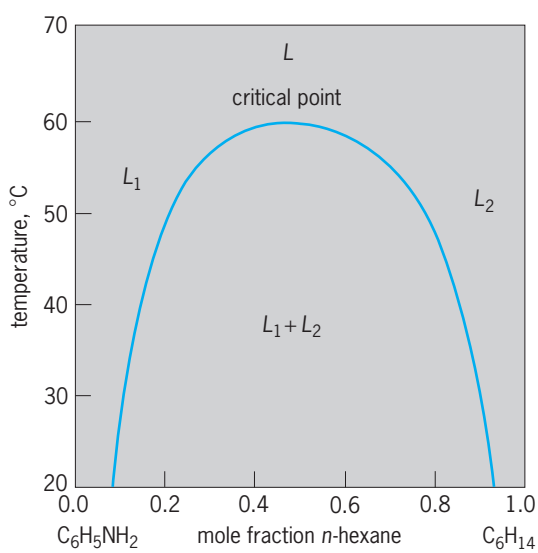

Fig. 4. Liquid-liquid equilibrium for aniline + *n*-hexane.

Solid-liquid equilibrium. When two substances are completely miscible with each other and form a complete series of solid solutions, the solid-liquid phase diagrams are entirely analogous to the liquid-vapor diagrams illustrated above. Those with no maximum or minimum, usually associated with nearly ideal liquid and solid solutions, such as methane + krypton, are called type I, according to the Bakhuis Roozeboom classification; those with a maximum melting point, an exceedingly rare type exemplified by *d*-carvoxime + *l*-carvoxime, are type II; and those with a minimum melting point (bromobenzene + iodobenzene), type III. *See* SOLID SOLUTION.

In most binary systems, however, extensive solid solutions are impossible because of the incompatibility of the size, shape, and crystal lattices of the two components. In the absence of solid solution formation, the addition of a solute to a liquid solvent invariably depresses the freezing point, that is, the temperature at which, upon cooling, the first trace of solid solvent appears. This depression of the freezing point of a dilute solution is a convenient method (the so-called cryoscopic method) for determining the molecular weight of a solid solute. *See* MOLECULAR WEIGHT.

As the freezing-point curve of a liquid continues to lower temperatures with higher concentrations of the second component, the conventional roles of solute and solvent become reversed, and one speaks of the solubility of a solute rather than of the depression of the freezing point of the solvent; no point of demarcation exists, and the two situations are in fact two aspects of the same phenomenon.

The second component also has a freezing-point–solubility curve marking the temperature at which the solution is in equilibrium with pure solid 2. The point of intersection of the two phase boundaries is the eutectic point which defines a eutectic temperature and a eutectic composition. Below this temperature, the system consists of two solid phases (S_1 + S_2 in **Fig. 5***a*).

Solid-liquid phase diagrams are conveniently determined by thermal analysis of cooling curves. An initially homogeneous liquid is cooled gradually, and the temperature plotted against time. Figure 5*a* shows a simple eutectic-type phase diagram, and Fig. 5*b* sketches typical cooling curves obtained for various compositions (those marked *A* to *F* in Fig. 5*a*). Curves *A* and *F* are for the pure components; each shows a single temperature arrest, a horizontal section of the curve for the melting point (between *f* and *m*), at which the temperature remains constant from the time the first bit of solid is formed until the last bit of liquid disappears. *See* THERMAL ANALYSIS.

Curves *B* and *E* are for solutions rich in components 1 and 2, respectively. At a temperature below the melting point of the pure substance, the first bit of solvent begins to freeze out; this is indicated by a change in slope (point *f*). The freezing section of the

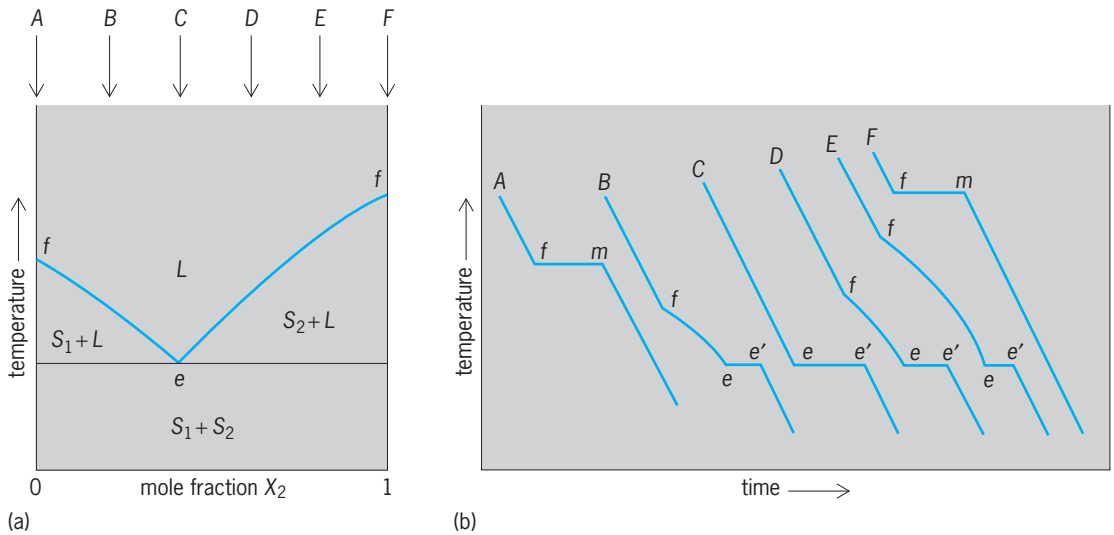

Fig. 5. Solid-liquid equilibrium. (*a*) Schematic diagram when there are no compounds of solid solutions. Points *f* are freezing points of pure substance; *e* is the eutectic point. *A*, *B*, *C*, *D*, *E*, and *F* are the compositions corresponding to the cooling curves of part *b*. (*b*) Schematic cooling curves for system shown in part *a*; *f* = freezing point, *m* = melting point, *e* = beginning of eutectic freezing, *e′* = end of eutectic freezing.

cooling curve (between points *f* and *e*) is not horizontal; as solvent freezes out, the liquid phase becomes richer in solute, and the freezing point is further depressed. When the composition of the liquid phase reaches the eutectic composition, the second component begins to freeze out as well as the first. No further change in liquid composition occurs, so the temperature remains constant until no liquid remains (between points *e* and *e′*). The solid which freezes out at the eutectic (the eutectic mixture) appears superficially very different from either pure solid; it is a mixture of very small crystals of each of the two components which have crystallized together. This microcrystalline two-phase mixture is in no sense a compound.

Curve *D* is a cooling curve for still another composition, and curve *C* is for the eutectic composition. Alone, curve *C* is indistinguishable from that for the freezing of a pure substance; only by combining the information from a series of cooling curves can one construct the whole diagram.

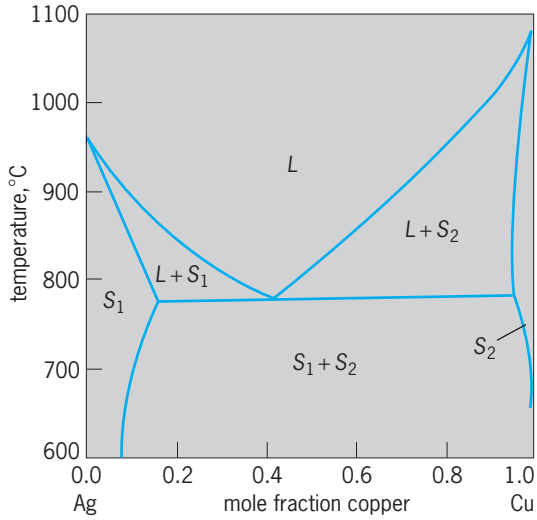

Fig. 6. Solid-liquid equilibrium in the system silver + copper (type-V solid solutions).

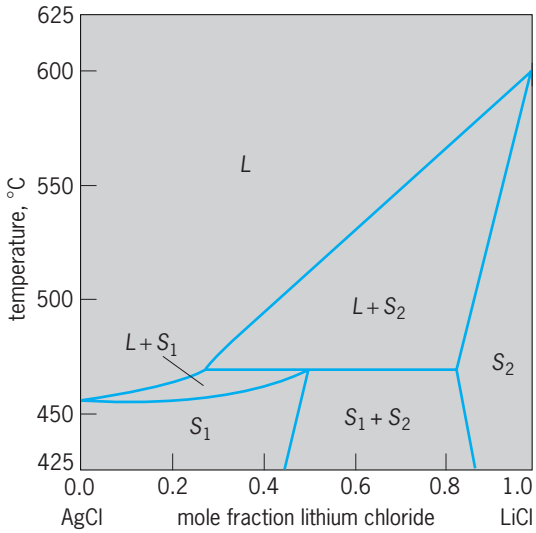

Fig. 7. Solid-liquid equilibrium in the system silver chloride + lithium chloride (type-IV solid solutions).

When the components form an incomplete series of solid solutions (partial miscibility), the phase diagram can be of the eutectic type illustrated by the system silver + copper shown in **Fig. 6**. Indeed, since the mutual solubilities are never exactly zero, eutectic diagrams, such as that in Fig. 5*a*, are in fact merely extreme examples of this more general case.

When a solid phase upon melting transforms into a liquid phase and a solid phase of different composition, one speaks of incongruent melting and a peritectic-type phase diagram. A simple example is the type-IV solid-solution diagram illustrated by the system silver chloride + lithium chloride (**Fig. 7**); similar phenomena occur in systems without any appreciable solid-solution formation.

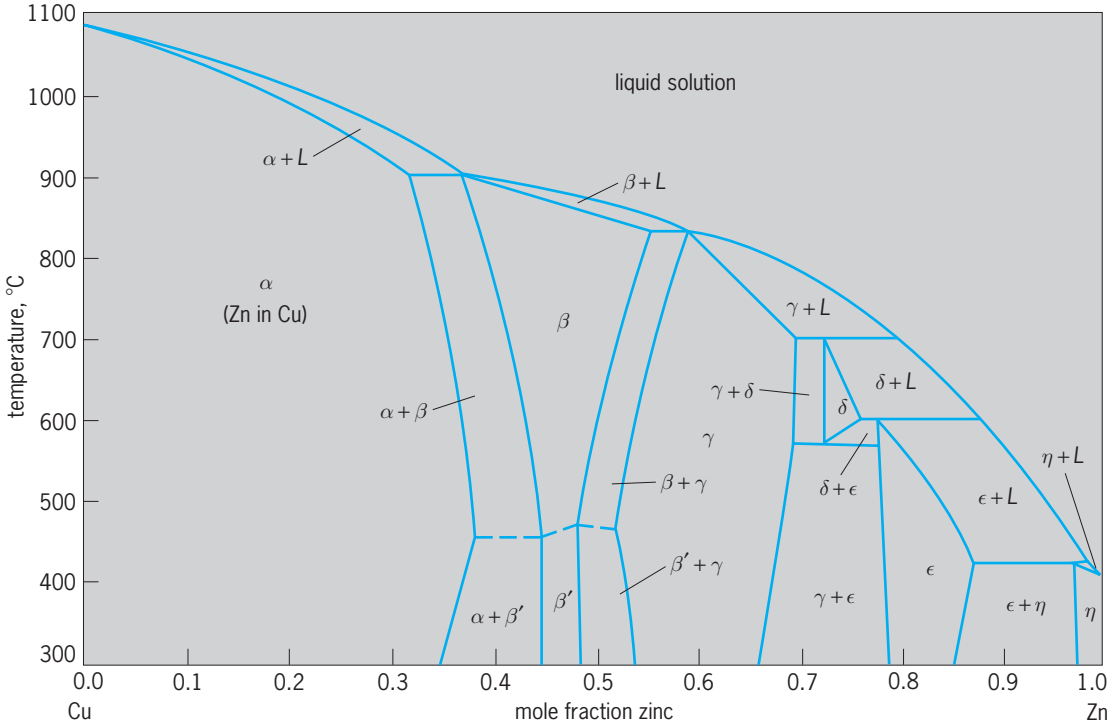

Fig. 8. Temperature-composition diagram for the system copper + zinc (brass). Note the six different solid phases. The broken line separating β from β' denotes a second-order, order-disorder transition.

Many solid-liquid phase diagrams are complicated by the existence of intermediate crystalline phases of different crystal structure. Usually, these intermediate phases are at compositions close to simple mole ratios of the components and have the two kinds of molecules distributed in a regular arrangement; consequently, it is convenient to call these compounds, even if no specific chemical interactions can be demonstrated unequivocally. There is a rich variety of such complex phase diagrams, especially among binary systems of ionic salts ($CaCl_2$-KCl) and among binary systems of metals (alloys). **Figure 8** shows the Cu + Zn system (brass), in which a whole series of crystal structures appears: α-brass has the crystal structure of pure copper (face-centered cubic) with an occasional Zn atom in the lattice; β-brass has a body-centered cubic structure with a Cu-Zn ratio of approximately 1:1; γ-brass has a very complex structure related to the formula Cu_5Zn_9; η-brass has the crystal structure of pure Zn (hexagonal, close-packed); the δ and ϵ phases have still different structures. Note the five peritectic transitions [for example, at 600°C (1112°F) where the ϵ-solid melts incongruently to the δ-solid and a liquid solution rich in zinc].

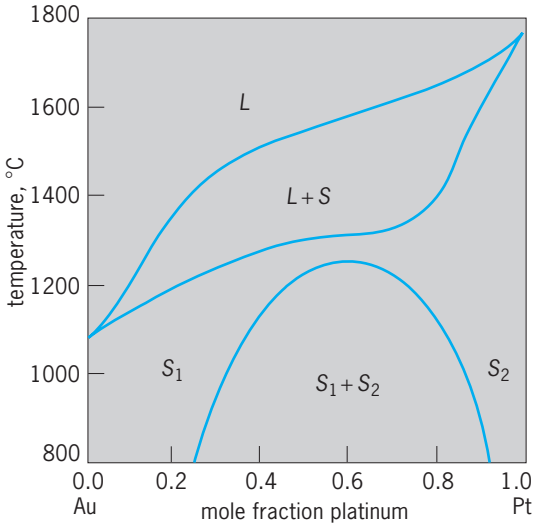

Fig. 9. Solid-liquid and solid-solid equilibria in the system gold + platinum. Critical region of the solid-solid phase boundary nearly touches the solidus curve of the solid-liquid phase boundary. If these touched and coalesced, a type-IV diagram (Fig. 7) would result.

Solid-state equilibrium. These are of several types. One can have solid critical solution temperatures (with solid solutions of the same crystal structure) which are analogous to the more familiar liquid-liquid case (the system gold + platinum shown in **Fig. 9**). More common are the transitions between one crystalline form and another, which can occur even in pure substances. One such occurs in the system nitrogen + carbon monoxide (**Fig. 10** shows solid-solid, solid-liquid, and liquid-vapor transitions for this system). *See* TRANSITION POINT.

Multicomponent Systems

As one proceeds from binary systems to systems with three or more components, the phase diagrams become more complex. Each component adds another dimension to the representation of the phase

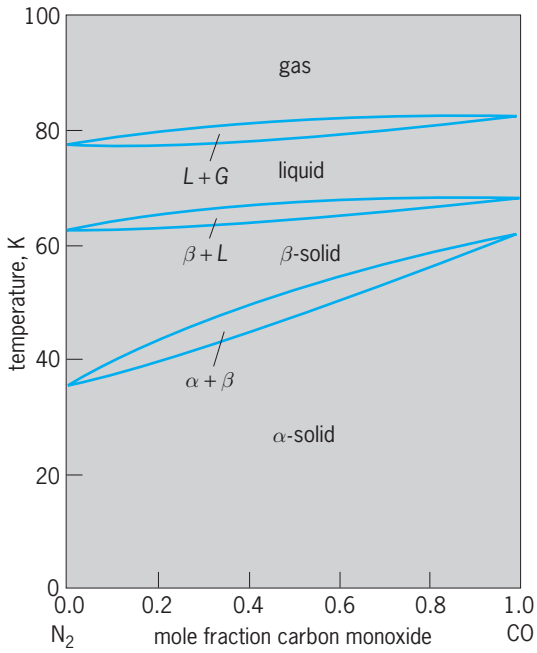

Fig. 10. Phase equilibria for nitrogen + carbon monoxide, being solid-solid, solid-liquid, and liquid-vapor.

equilibria. Thus, for three components, two dimensions are required to represent the phase diagrams for a single temperature and pressure; these are conveniently depicted by a triangular diagram in which each vertex represents a pure component. **Figure 11** shows a diagram of a ternary system in which three liquid phases can coexist. Even when there are three phases, the system is still bivariant. (An example of such a system is water + succinonitrile + diethyl ether.)

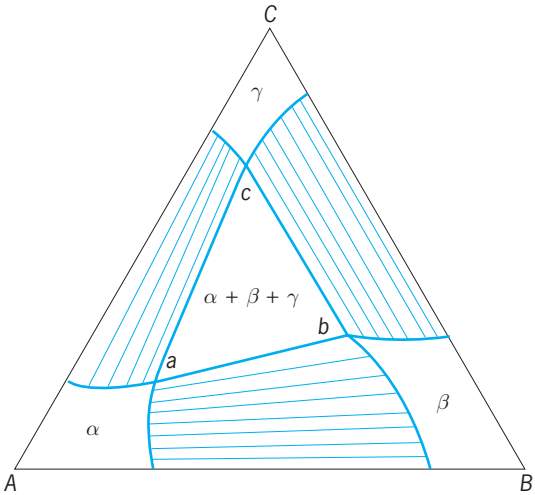

Fig. 11. Schematic diagram of a three-component system at a fixed temperature and pressure. Points *A*, *B*, and *C* represent the pure liquids. The composition corresponding to a point in the diagram is determined by the positions along a line from each vertex to the opposite side; thus point *b* is 20% *A*, 50% *B*, 30% *C*. Regions α, β, and γ correspond to single phases rich in *A*, *B*, and *C*, respectively; $\alpha + \beta + \gamma$ is a three-phase region; the three saturated solutions have the compositions given by the points *a*, *b*, and *c*. Three two-phase regions, $\alpha + \beta$, $\alpha + \gamma$, and $\beta + \gamma$, are indicated by drawing in the tie lines.

A special case of a three-component system is that in which there are two immiscible solvents and a third component, soluble in both, distributed between the two phases. The ratio of the concentrations of the solute in the two solvents is the distribution coefficient; in dilute solutions, this is independent of the concentration, but at higher concentrations, nonideal behavior of the solute can produce systematic variations of the distribution coefficient which ends in the most concentrated solutions with the ratio of the solubilities in the saturated solutions.

Distribution effects are important in separating similar materials. A small difference in distribution coefficients is amplified by multistage equilibria, such as those used in countercurrent extraction and partition chromatography.

The examples used to illustrate the various types of phase equilibria are not supposed to suggest that these types are restricted to the particular kind of chemical substances shown. In general, examples of each could have been selected from many kinds of substance such as metals, nonmetallic elements, inorganic salts, and organic nonelectrolytes. *See* ALLOY STRUCTURES; CHEMICAL EQUILIBRIUM; CHEMICAL THERMODYNAMICS; COUNTERCURRENT TRANSFER OPERATIONS; CRYSTAL STRUCTURE; EXTRACTION; INTERFACE OF PHASES; SILICATE PHASE EQUILIBRIA; SULFIDE PHASE EQUILIBRIA. Robert L. Scott

Bibliography. G. M. Barrow, *Physical Chemistry*, 6th ed., 1996; R. Ginell, *Association Theory: The Phases of Matter and Their Transformations*, 1979; I. N. Levine, *Physical Chemistry*, 4th ed., 1995; Research and Education Association Staff, *The Essentials of Physical Chemistry*, vols. 1 and 2, 1994; F. E. Wetmore and D. J. LeRoy, *Principles of Phase Equilibria*, 1951.

Phase inverter

A circuit having the primary function of changing the phase of a signal by 180°. The phase inverter is most commonly employed as the input stage for a push-pull amplifier. Therefore, the phase inverter must supply two voltages of equal magnitude and 180° phase difference. A variety of circuits are available for the phase inversion. The circuit used in any given case depends upon such factors as the overall gain of the phase inverter and push-pull amplifier, the possibility that the input to the push-pull amplifier may require power, space requirements, and cost. *See* PUSH-PULL AMPLIFIER.

Overall fidelity of a phase inverter and push-pull amplifier can be adversely affected by improper design of the phase inverter. The principal design requirement is that frequency response of one input channel to the push-pull amplifier be identical to the frequency response of the other channel. In this respect, popular phase-inversion circuits are capable of providing precisely 180° phase difference only after careful selection of components. Some phase-inverter circuits can perform inversion at only one

frequency; at other frequencies, distortion is introduced because of unequal frequency-response characteristics.

Operational-amplifier inverter. A model phase inverter using two operational amplifiers is shown in **Fig. 1**. Operational amplifer 1 is a voltage follower; the output voltage is equal to the input voltage. Operational amplifier 2 is a unity-gain inverter, and so the output is the negative of the input. The two operational amplifiers can be fabricated in one integrated circuit chip. The circuit is shown in simplified form in that each operational amplifier needs a dual voltage supply, perhaps plus and minus 15 V. Furthermore, the circuit needs some extra resistors to eliminate the effects of the dc offset voltages and currents. The frequency responses of operational amplifiers 1 and 2 must be identical over the frequency range of operation so that the outputs will truly be V and $-V$ as shown. As noted above, this is a requirement for all phase-inverter circuits. The output resistance of each operational-amplifier path in Fig. 1 is very low because of the feedback. This is a desirable property, not found in the circuits discussed below. *See* AMPLIFIER; OPERATIONAL AMPLIFIER.

Paraphase amplifiers. An amplifier that provides two equal output signals 180° out of phase is called a paraphase amplifier. If coupling capacitors can be omitted, the simplest paraphase amplifier is illustrated in **Fig. 2**. Approximately the same current flows through R_L and R_E, and therefore if R_L and R_E are equal, the ac output voltages from the collector and from the emitter are equal in magnitude and 180° out of phase. The gain of the circuit is less than unity, which is one factor that limits its applicability. A second important factor is that the addition of coupling capacitors and biasing resistors, necessary when the circuit is coupled to the push-pull stage, causes the phase inversion to be other than 180° over the frequency range of expected operation. One of the most stable and important paraphase amplifiers is the emitter-coupled phase inverter (**Fig. 3**). If the emitter resistance R_E is large compared to the impedance seen looking into the emitter of each transistor, current i_1 will equal i_2. Under this condition the voltage at one collector is exactly the negative of that at the other collector, and push-pull operation is achieved. The collector-to-collector gain is equal to the gain which would be provided by a single-transistor grounded-emitter amplifier with collector load R_L. *See* TRANSISTOR.

If the phase-inverter circuit is to produce two voltages 180° out of phase, the equivalent circuits governing the behavior of the two output voltages must be identical. The midfrequency gain of each must be identical, and the phase-shift functions must be identical. The phase-shift requirements are often compromised in the interests of simplicity of the final circuit and freedom from critical adjustments of key circuit parameters.

Transformer inverter. The simplest form of phase-inverter circuit is a transformer with a center-tapped secondary (**Fig. 4**). Careful design of the transformer assures that the secondary voltages are equal. The

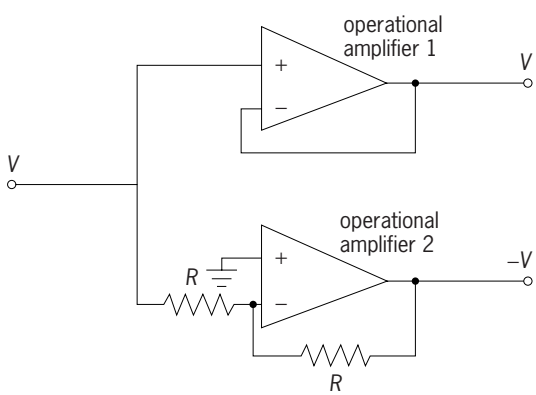

Fig. 1. Phase inverter using operational amplifiers.

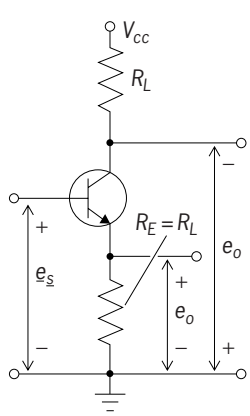

Fig. 2. Single-transistor inverter.

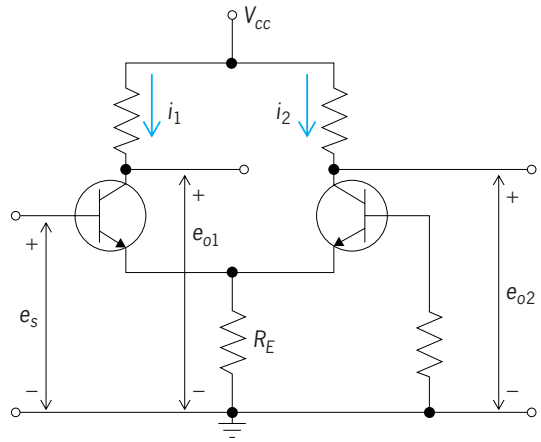

Fig. 3. Emitter-coupled phase inverter.

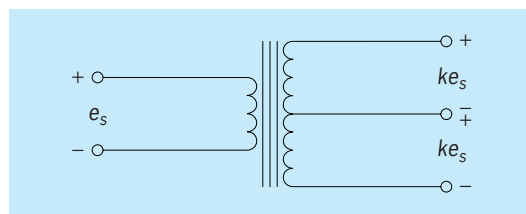

Fig. 4. Transformer, simplest form of phase-inverter circuit. The turns ratio is represented by *k*.

transformer forms a good inverter when the inverter must supply power to the input of the push-pull amplifier. The turns ratio can be adjusted for maximum power transfer. *See* TRANSFORMER.

The transformer inverter has several disadvantages. It usually costs more, occupies more space, and weighs more than a transistor circuit. Furthermore, some means must be found to compensate for the frequency response of the transformer, which may not be as uniform as that which can be obtained from solid-state circuits. *See* PHASE (PERIODIC PHENOMENA); PHASE-ANGLE MEASUREMENT.

Harold F. Klock

Bibliography. P. Chirlian, *Analysis and Design of Integrated Electronic Circuits*, 2d ed., 1986; M. Ghausi, *Electronic Devices and Circuits: Discrete and Integrated*, 1995; S. Gibilisco (ed.), *Encyclopedia of Electronics*, 2d ed., 1990.

Phase-locked loops

Electronic circuits for locking an oscillator in phase with an arbitrary input signal. A phase-locked loop (PLL) is used in three fundamentally different ways: (1) as a demodulator, where it is employed to follow (and demodulate) frequency or phase modulation; (2) as a tracker of a carrier or synchronizing signal which may vary in frequency with time; and (3) as a frequency synthesizer, where an oscillator is locked to a multiple of an accurate reference frequency. When operating as a demodulator, the phase-locked loop may be thought of as a matched filter operating as a coherent detector. When used to track a carrier, it may be thought of as a narrow-band filter for removing noise from the signal and regenerating a clean replica of the signal. When used as a frequency synthesizer, a voltage-controlled oscillator (VCO) is divided down to a reference frequency that is locked to a frequency derived from an accurate source such as a crystal oscillator. *See* DEMODULATOR; ELECTRIC FILTER; FREQUENCY-MODULATION DETECTOR; FREQUENCY MULTIPLIER; OSCILLATOR; PHASE-MODULATION DETECTOR.

Basic operation. The basic components of a phase-locked loop are shown in **Fig. 1**. The input signal is a sine or square wave of arbitrary frequency. The voltage-controlled-oscillator output signal is a sine or square wave of the same frequency as the input, but the phase angle between the two is arbitrary. The output of the phase detector consists of a direct-current (dc) term, and components of the input frequency and its harmonics. The low-pass loop filter removes all alternating-current (ac) components, leaving the dc component, the magnitude of which is a function of the phase angle between the voltage-controlled-oscillator signal and the input signal. If the frequency of the input signal changes, a change in phase angle between these signals will produce a change in the dc control voltage in such a manner as to vary the frequency of the voltage-controlled oscillator to track the frequency of the input signal.

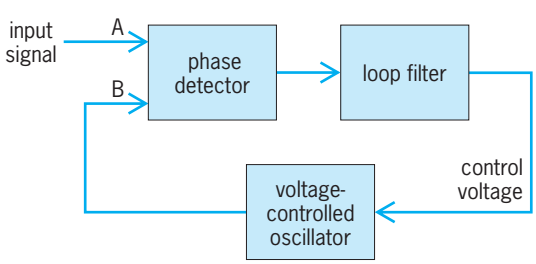

Fig. 1. Basic components of a phase-locked loop.

Two qualities of the loop specify its performance: the lock range and the capture range. The lock range is the maximum change in input frequency for which the loop will remain locked. It is governed by the dc gain of the loop. As the input frequency is changed, the change in phase of the two signals to the phase detector will produce a dc control voltage that will change the frequency of the voltage-controlled oscillator. As the input frequency is further changed, the phase angle will continue to increase until it reaches 0 or 180°, when the loop will unlock. If an amplifier is added to the loop, a greater control voltage will be generated which will decrease the phase error in the phase detector, and hence further detuning can occur before unlocking takes place.

The capture range is that range of frequencies that the loop will lock to if it is initially unlocked. Suppose the loop is unlocked and the voltage-controlled oscillator is running at frequency f_1. If the input signal f_2 is applied but is out of the pull-in range, a beat note $f_1 - f_2$ will appear at the output of the phase detector. The filter components will govern the amplitude of this beat note at the input to the voltage-controlled oscillator. If the frequency difference is reduced, the frequency of the beat note will decrease and the amplitude at the voltage-controlled oscillator input will increase. At some point, the amplitude will drive the voltage-controlled oscillator far enough over in frequency to match the input frequency and locking will occur. The lower the roll-off frequency of the loop filter, above which its attenuation begins to increase, the less will be the capture range.

Phase-locked loops are generally designed to have narrower capture range than lock range. This is the advantage of a phase-locked loop over more conventional types of filters. The capture range is analogous to filter bandwidth and may be made as narrow as desired by suitable choice of the resistance-capacitance (RC) filter, while the center frequency may be any desired value.

Voltage-controlled oscillators. Voltage-controlled oscillators may take many forms, two common types being the reactance modulator and the varactor modulator. These circuits are commonly used in narrow-band phase-locked loops where narrow lock and capture ranges are desired. *See* VARACTOR.

Another common form of voltage-controlled oscillator utilizes a multivibrator type of circuit where the timing capacitor charging current may be varied by a dc control voltage. This circuit relies on the fact that the time required to charge the timing capacitance is inversely proportioned to the frequency of

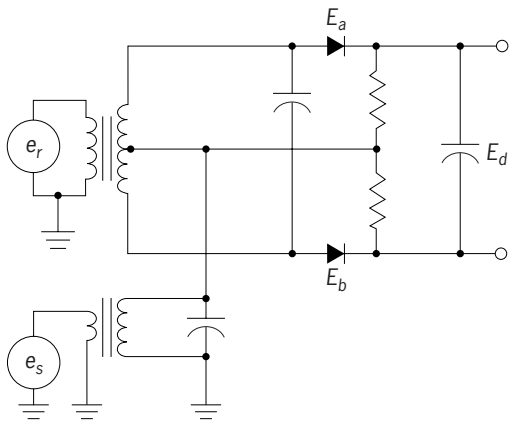

Fig. 2. Diode phase detector.

oscillation, hence the control voltage will vary the frequency of the oscillator. *See* MULTIVIBRATOR.

Reactance- and varactor-type modulators are used with crystal-controlled oscillators where tuning ranges of 0.25–0.5% are desired, or with inductance-capacitance (*LC*) oscillators where tuning ranges of 5% or more can be achieved. Multivibrator types of circuits are used where wide tuning ranges of up to 1000 to 1 have been achieved. Conversely, phase noise and stability are best with a crystal-controlled oscillator and least with multivibratortype circuits.

Phase detectors. Two commonly used phase detectors are the diode phase detector and the double-balanced phase detector. The diode phase detector has been used historically because of lower cost and better performance at high frequencies (above 50 MHz). With the advent of monolithic integrated circuits, the double-balanced phase detector has become useful at lower frequencies. *See* INTEGRATED CIRCUITS.

Diode detector. In the diode detector (**Fig. 2**) the voltages applied to the two peak rectifiers are the sum and difference of the two input signals, e_r and e_s. Rectified output voltages are equal to the amplitudes of each of the sums E_a and E_b. The phase detector output voltage E_d is equal to the difference of the two rectified voltages, $E_a - E_b$. Considering that e_r and e_s are sinusoidal waves of the same frequency but varying phase angle, the dc output E_d will be zero when the two input signals are in quadrature (90°), a maximum positive value when in phase, and a maximum negative value when 180° out of phase.

Double-balanced detector. The double-balanced phase detector (**Fig. 3***a*) works well at low frequencies. The upper four transistors operate as a double-pole double-throw switch driven by the reference signal e_r. The lower transistors operate as an amplifier, their collector currents being proportional to the input signal e_s. Figure 3*b* shows the output waveform for different phase angles between the two input signals. The output signal shown is that which would be obtained if the filter capacitor *C* were removed.

Digital phase/frequency detector. A third type of phase detector, used in frequency synthesizers, is shown in **Fig. 4**. This digital three-state phase/frequency comparator with a charge-pump (switched current source) output circuit has the advantage of producing a maximum dc component at its output when not in lock, so that the loop is driven to lock even when a large error exists between the two frequencies at the detector input. The output of this detector has a three-state capability; depending on which input signal is higher or lower in frequency, its output is a positive or negative current source or an open circuit. When the two input frequencies are the same, the detector's output will drive high or low depending on the phase error between the input signals. When a very small phase error exists between the triggering edges of the two input signals A and B, either the pull-up or the pull-down signal will generate a small output error signal to maintain lock. Unlike the linear phase detectors described above, this circuit operates at zero phase error between its inputs,

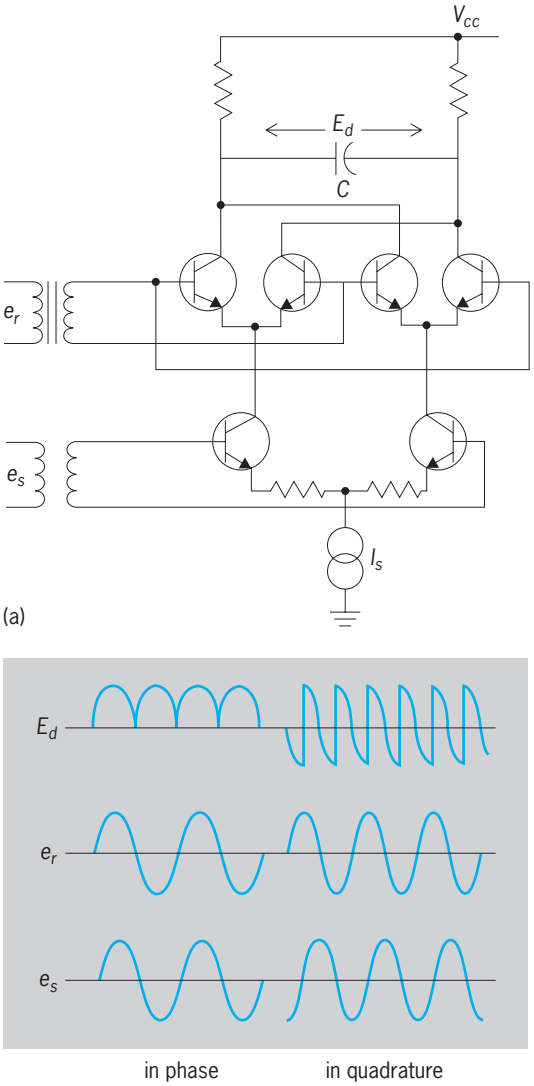

Fig. 3. Double-balanced phase detector. (*a*) Circuit. (*b*) Waveforms.

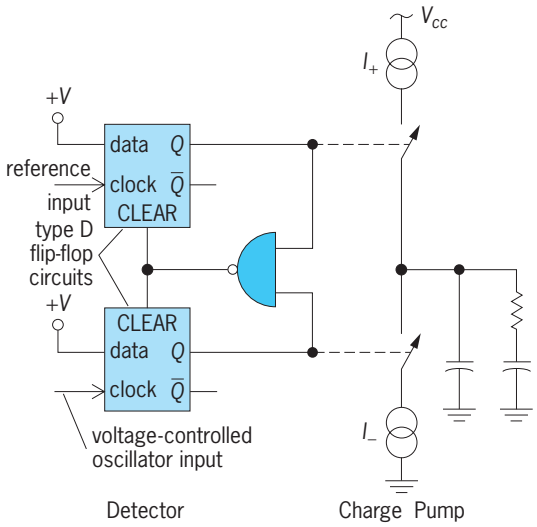

Fig. 4. Digital phase/frequency detector. Q = output; $\overline{Q}$ = inverting output.

and under these conditions the output is a high impedance, with essentially no output ripple voltage. This is an ideal type of phase detector for loops operating with very quiet input signals; however, any noise will drive the output high or low until the next cycle edge occurs, causing a major disturbance to the loop. This property of the digital-logic-gate type of phase detector makes it a poor choice for demodulator functions operating with noisy input signals.

Loop filter. The loop filter determines the dynamic characteristics of the loop: its bandwidth and response time. This filter is generally composed of a lag-lead network (**Fig. 5**). The transfer function of the voltage-controlled oscillator has the characteristic of an integrator: a step change in the dc control voltage will produce a ramp change in output phase, increasing without bound. The low-pass filter characteristic begins attenuating frequencies above its break point determined by R_1C_1; at a higher frequency determined by R_2C_1, no further attenuation occurs. This is important because the phase shift associated with the voltage-controlled oscillator is 90°, and that associated with the roll off due to R_1C_1 is also approaching 90°. With a phase shift of 180°, the loop will oscillate and become unstable. The leading phase shift associated with R_2C_1 reduces the phase shift around the loop and enhances stability.

Uses of phase-locked loops. The most widespread use of phase-locked loops is undoubtedly in television receivers. Synchronization of the horizontal oscillator to the transmitted sync pulses is universally

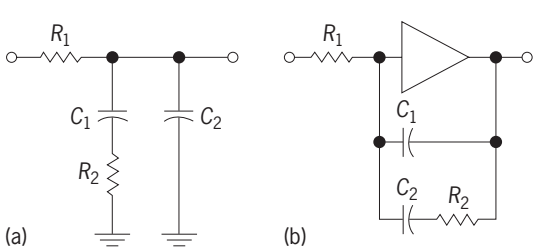

Fig. 5. Loop filter configurations. (*a*) Passive. (*b*) Active.

accomplished with a phase-locked loop; here it is desirable to have a stable, noise-free reference source to generate the scanning line since the eye is intolerant of any jitter (or phase noise). The color reference oscillator is often synchronized with a phase-locked loop; here it is necessary to maintain less than 5° of phase error between the transmitted color reference signal and the locally generated reference in the receiver. In both of these applications, the voltage-controlled oscillator signal is the useful output, and the design of the loop filter is such as to produce a narrow-band characteristic so that noise is reduced and the filter "remembers" the proper phase during short periods of noise and loss of signal. *See* TELEVISION RECEIVER.

Phase-locked loops are also used as frequency demodulators. Since the voltage-controlled oscillator is locked to the incoming carrier frequency, and the control voltage is proportional to the voltagecontrolled oscillator frequency, as long as the loop remains locked, the control voltage will be a replica of the modulating signal.

Stereo decoding. Phase-locked loops have been applied to stereo decoders made on silicon monolithic integrated circuits. This design technique eliminates the coils used in previous decoder designs. The circuit consists of a 76-kHz voltage-controlled oscillator, which, after being divided by 4, is locked to the transmitted 19-kHz pilot carrier. The phase-locked loop therefore acts as a carrier regenerator (or narrowband filter). The 76 kHz is also divided by 2 to obtain 38 kHz that is applied to a double-balanced demodulator, which decodes the stereo signal into left and right channels. *See* STEREOPHONIC RADIO TRANSMISSION.

Amplitude demodulation. High-performance amplitude demodulators may be built by using phase-lock techniques. If, instead of a diode detector, a double-balanced demodulator is used, the lower port of which is driven with the signal to be demodulated and the upper port is driven with a noise-free replica of the carrier, a very linear (low-distortion) demodulator results. The phase-locked loop locks to the carrier frequency; the loop filter should be narrow enough that modulation is not present in the phase-locked loop. The phase detector is driven in quadrature with the input signal so that very little amplitude information is demodulated in the loop. The voltage-controlled oscillator signal is also phase-shifted 90° and applied to a second phase detector, which is also driven with the signal to be demodulated. Since the carrier and reference are in phase in this detector, the output will be proportional to the amplitude of the incoming carrier. Demodulators using this technique have been built with linearity better than 0.5% distortion for 90% modulated signals. *See* AMPLITUDE-MODULATION DETECTOR.

Frequency synthesis. A frequency-synthesizer application of a phase-locked loop is used in the majority of AM-FM radio and television receivers and most modern two-way communications equipment. A crystal oscillator generates a reference signal F_x, which is divided down by a number M to a reference frequency

F_r. A voltage-controlled oscillator running at the frequency to be synthesized is also divided down by a number N to the reference frequency to which it is phase-locked. In this way, the voltage-controlled oscillator can be tuned to exact mutiples of the reference frequency. In an example of a local oscillator for an FM broadcast receiver, the crystal oscillator may operate at 4 MHz, and be divided by 160 in the M counter to produce a reference frequency of 25 kHz. It is desired to tune the local oscillator from 98.6 to 118.6 MHz in 25-kHz steps. In this case, the divide ratio in the N counter would vary between 3944 and 4744. In radio and communications applications, fast switching speed is not a primary requirement, and loops are designed to have a relatively narrow closed-loop bandwidth in order to reduce noise associated with the reference frequency. A microprocessor is used to load N in the programmable divider and provide for frequency display, keyboard reading, and so forth in a radio application. A variety of single- and multiple-chip synthesizers are available from manufacturers of integrated circuits. Most of these integrated circuits contain the crystal oscillator, both M and N counters, and the phase detector. Varactor-tuned inductance-capacitance (LC) oscillators are usually used in these applications. *See* RADIO RECEIVER.

Data-disk synchronization. With the requirements for large amounts of mass storage, the computer industry has turned to magnetic recording techniques using both hard and floppy magnetic disks. Optical storage techniques have become available for extremely high-density data storage. In all of these systems, it is necessary to lock a clock signal to the data being read from the disk or an internal clock signal. This is accomplished with a phase-locked loop using a digital-logic-gate type of phase detector described above. The voltage-controlled oscillator is locked to either the system clock or the encoded data signal from the disk depending on the state of the read gate. In this way, the voltage-controlled oscillator signal is available to the coding and decoding logic circuits in the drive and is coherent in phase with the data being processed at all times. Generally, response speed is important in phase-locked loops for disk applications, and loop bandwidth may approach 15% of the data rate or hundreds of kilohertz. Many loops are designed for both high- and low-speed operation by switching the gain of the phase detector. Single-chip integrated circuits are available from many manufacturers to provide the phase-lock functions for both fixed and floppy disk drives. Resistance-capacitance (RC)–coupled multivibrator voltage-controlled oscillators are often employed in these applications. *See* COMPUTER STORAGE TECHNOLOGY; OPTICAL RECORDING. Thomas B. Mills

Bibliography. R. E. Best, *Phase-Locked Loops: Design, Simulation, and Applications*, 5th ed., 2003; W. F. Egan, *Phase-Lock Basics*, 1998; F. M. Gardner, *Phaselock Techniques*, 3d ed., 2005; W. C. Lindsey and C. M. Chie (eds.), *Phase-Locked Loops*, 1985; D. H. Wolaver, *Phase-Locked Loop Circuit Design*, 1991.

Phase modulation

A technique used in telecommunications transmission systems whereby the phase of a periodic carrier signal is changed in accordance with the characteristics of an information signal, called the modulating signal. Phase modulation (PM) is a form of angle modulation. For systems in which the modulating signal is digital, the term "phase-shift keying" (PSK) is usually employed. *See* ANGLE MODULATION.

In typical applications of phase modulation or phase-shift keying, the carrier signal is a pure sine wave of constant amplitude, represented mathematically as Eq. (1), where the constant A is its amplitude,

$$c(t) = A \sin \theta(t) \tag{1}$$

$\theta(t) = \omega t$ is its phase, which increases linearly with time, and $\omega = 2\pi f$ and f are constants that represent the carrier signal's radian and linear frequency, respectively.

Phase modulation varies the phase of the carrier signal in direct relation to the modulating signal $m(t)$, resulting in Eq. (2), where k is a constant of pro-

$$\theta(t) = \omega t + km(t) \tag{2}$$

portionality. The resulting transmitted signal $s(t)$ is therefore given by Eq. (3). At the receiver, $m(t)$ is re-

$$s(t) = A \sin \{\omega t + km(t)\} \tag{3}$$

constructed by measuring the variations in the phase of the received modulated carrier.

Phase modulation is intimately related to frequency modulation (FM) in that changing the phase of $c(t)$ in accordance with $m(t)$ is equivalent to changing the instantaneous frequency of $c(t)$ in accordance with the time derivative of $m(t)$. *See* FREQUENCY MODULATION.

Properties. Since only the phase of the carrier signal $c(t)$ is affected by the modulation process, the amplitude of the modulated signal $s(t)$ is constant over time. Consequently the peak power and average power in $s(t)$ are constant as well, independent of the form of the modulating signal $m(t)$ and the value of k, and completely determined by the amplitude of $c(t)$. Because of this feature, phase modulation has several advantages over other forms of modulation, such as amplitude modulation (AM). Among the advantages are superior noise and interference rejection, enhanced immunity to signal fading, and reduced susceptibility to nonlinearities in the transmission and receiving systems. Phase modulation also exhibits the capture effect, whereby the stronger of two signals with nearly equal amplitudes suppresses the other. *See* DISTORTION (ELECTRONIC CIRCUITS); ELECTRICAL INTERFERENCE; ELECTRICAL NOISE.

Although theoretically the bandwidth of $s(t)$ is infinite and independent of $m(t)$, in practical terms it is strongly affected by the bandwidth of $m(t)$. Thus, under certain conditions, if the power of $m(t)$ is distributed over the range of frequencies from f_1 hertz at the lower end of the spectrum to f_2 hertz at the

upper end, the amount of power in $s(t)$ contained outside the frequency range from $f - f_1$ hertz to $f - f_2$ hertz and from $f + f_1$ hertz to $f + f_2$ hertz is of minor consequence. The bandwidth of $s(t)$ is then taken to be $2f_2$ hertz, centered at f hertz, the linear frequency of $c(t)$.

Phase modulators and demodulators. For modulating signals that produce only small variations in phase, a simple form of phase modulation can be achieved by generating a carrier signal in a local oscillator, shifting its phase by 90°, and multiplying it by the modulating signal. Adding this product to the original unmodulated carrier signal yields a good approximation of the desired result. For larger phase variations, switching-circuit modulators are frequently employed. *See* PHASE MODULATOR.

Some demodulators, known as zero-crossing detectors, rely on the fact that the points in time when the modulated carrier signal assumes a value of 0 are representative of the original modulating signal. Other demodulators employ phase-locked loops, which track the phase of the modulated carrier signal by adjusting the output of the demodulator through a feedback mechanism. Phase-locked loops find application in low signal-to-noise ratio environments. *See* PHASE-LOCKED LOOPS; PHASE-MODULATION DETECTOR.

Phase-shift keying. When the modulating signal $m(t)$ is digital, so that its amplitude assumes a discrete set of values, the phase of the carrier signal is "shifted" by $m(t)$ at the points in time where $m(t)$ changes its amplitude. The amount of the shift in phase is usually determined by the number of different possible amplitudes of $m(t)$. In binary phase-shift keying (BPSK), where $m(t)$ assumes only two amplitudes, the phase of the carrier differs by 180°. An example of a higher-order system is quadrature phase-shift keying (QPSK), in which four amplitudes of $m(t)$ are represented by four different phases of the carrier signal, usually at 90° intervals.

The bandwidth of a PSK signal, defined as the range of frequencies over which the power in the modulated carrier signal is distributed, is determined by the rate R at which the phase of the carrier signal is shifted, and is independent of the number of amplitudes of $m(t)$. To a good approximation, it can be taken to equal $2R$ hertz, extending over the frequency range from $f - R$ to $f + R$ hertz. Thus the higher-order systems such as QPSK require less bandwidth per unit information than BPSK. *See* MODULATION.

Hermann J. Helgert

Bibliography. B. P. Lathi, *Modern Digital and Analog Communication Systems*, 3d ed., Oxford University Press, 1998; M. Schwartz, *Information Transmission, Modulation, and Noise*, 4th ed., McGraw-Hill, 1990.

Phase-modulation detector

A device that recovers or detects a signal from a phase-modulated carrier. A phase-modulating (PM) signal operating on a sinusoidal carrier can be recovered in a variety of ways by phase-modulation detection (or demodulation). *See* CARRIER; DEMODULATOR; MODULATION; PHASE MODULATION.

Phase-locked loop. A phase-locked loop (PLL) is a closed-loop servo that comprises a phase detector for comparing the waveforms of the input sinusoidal signal with the output of a frequency-controlled oscillator whose frequency is responsive to an input-control voltage (see **illustration**). The oscillator's input (frequency)-control voltage is obtained by integrating and loop-filtering (in order to provide closed-loop stability) the output of the phase detector. The basic phase detector is a simple multiplier of the in-phase input signal and the quadrature component of the oscillator output. If we have both in-phase and quadrature components of both the input and oscillator signals, the products of the in-phase component of the input and the quadrature component of the oscillator output can be combined with those of the quadrature component of the input and the in-phase component of the oscillator output to mechanize a superior phase detector. If the loop filter is designed to provide a slow servo-loop response, the oscillator will run at the precise carrier frequency of the input signal, and the phase detector will provide a control signal that is proportional to the sine of the phase difference between the input signal and the oscillator output signal, plus a small and easily removed signal at twice the frequency of the input signal. If the inputs to the phase detector are hard-limited (prevented from exceeding a set output), the phase detector output will provide a signal that is linearly proportional to the phase difference between the input signal and the oscillator output plus a small double-frequency noise signal. The simple PLL is a very good phase-modulation detector of analog signals. *See* OSCILLATOR; PHASE-LOCKED LOOPS; QUADRATURE; SERVOMECHANISM.

Costas loop. The PLL detector is a coherent demodulator (or detector) because, in the absence of a provided demodulation reference, it reconstructs the carrier signal within the PLL. There exists a superior means for recovery of a data signal, such as a binary phase-shift keyed signal (BPSK). Phase-shift demodulation recovery always requires a reference waveform for measuring phase shift. The superior alternative approach to obtaining the reference waveform is the two-phase PLL, called the Costas loop, which takes advantage of the fact that the data bits are ± 1, so the squared data values are always one. The oscillator outputs are both sine and cosine at the input carrier frequency with a slight phase error, θ. The input signal is demodulated with respect to both sine and cosine outputs of the oscillator, and

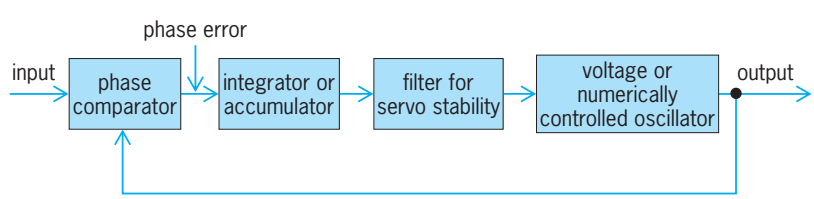

Basic phase-locked loop (PLL).

low-pass-filtered, and then the two resulting data-modulated signals are multiplied together to obtain a data-independent oscillator frequency-control input signal that is proportional to the sine of double the phase difference between the input signal and the oscillator output. The Costas-loop phase detector's performance can be improved, as with the simple PLL, by using complex input and oscillator-signal components.

Signal recovery. Recovery of differential binary phase-shift keyed (DBPSK) signals is accomplished by delaying the signal by one symbol time, then correlating the input and output of the delay element. A correlator simply multiplies two signals together and integrates over their product waveform; just multiply and average. When the signals are the same (no transition) the correlator output is high; when the signals are dissimilar the correlator output is zero or even negative (low). This is the function required by differential recovery.

Recovery of multiple phase-shift keyed (MPSK) signals requires coherent detection, that is, correlation with all the M possible symbols (M phase-shifted sinusoids), and selecting the largest value during each symbol time. Here the symbol duration, τ_s, is related to the bit duration, τ_b, and M by $\tau_s = \tau_b \log_2 M$. We will multiply the input signal by the M candidate symbol waveforms in M parallel processing channels and integrate each over τ_s seconds. The symbol corresponding to the largest output is the proper symbol for that symbol time.

Recovery of quadrature amplitude-modulation (QAM) signals is a different situation because we are attempting to decode simultaneously both amplitude and phase modulation. The first step is to correlate the input waveform with the sine and cosine of the carrier over one symbol time. We are not looking for the largest output from the two channels; rather, the outputs from these two channels will be applied to digital-to-analog converters (DACs) to recover amplitude information. The data sets are interleaved to produce the desired digital output. *See* AMPLITUDE MODULATOR.

Stanley A. White

Bibliography. G. A. Breed (ed.), *Wireless Communications Handbook*, Cardiff Publishing, 1992; S. Gibilisco, *Handbook of Radio and Wireless Technology*, McGraw-Hill, 1999; J. G. Proakis, *Digital Communications*, 4th ed., McGraw-Hill, 2001; T. S. Rappaport, *Wireless Communications*, 2d ed., Prentice Hall, 2002; M. S. Roden, *Analog and Digital Communications Systems*, 4th ed., Prentice Hall, 1995; M. E. Van Valkenburg and W. M. Middleton (eds.), *Reference Data for Engineers: Radio, Electronics, Computer and Communications*, 9th ed., Butterworth-Heinemann, 2002.

Phase modulator

A device that shifts the phase of a sinusoidal carrier signal, $c(t)$, in proportion to the instantaneous amplitude of the information signal, $s(t)$. Since the definition of frequency is the rate of change of phase, phase modulator and phase modulation are very closely related to frequency modulator and frequency modulation. (Integrate the frequency signal with respect to time in order to obtain the phase signal; differentiate the phase signal with respect to time to obtain the frequency signal.) The form of the unmodulated sinusoidal carrier signal is $c(t) = A \cos (2\pi f_c t)$, where the amplitude and frequency of the sinusoid are A and f_c, respectively.

Either analog or digital information may be modulated onto the sinusoidal carrier, but the methods may differ. Sets of bits that comprise a digital information unit, called a symbol, can have only a finite number of well-defined possible values (usually an integer power of 2), so the digital phase-modulating signal is constrained to a specific set of values. This modulating process is known as phase-shift keying, or PSK, referring to an early practice of closing one of multiple (M) switches at prescribed instances, thereby "keying" the output of an M-tap phase-shifting network. By this means, the phase of the carrier may be shifted to one of M prescribed angles for transmission. The transmitted waveform is a symbol-modulated carrier, and the period of time for its transmission is called a symbol time or symbol period. Because there are M signal states, this is called M-ary PSK in which data are compressed by a factor of $\log_2 M$ because that many data bits are combined to form the symbol. For example, simple two-state (on-off) binary phase-shift keying (BPSK) requires but a single bit to signify its state. On the other hand, M = 8 requires 3 bits/symbol. *See* ANGLE MODULATION; FREQUENCY MODULATION; MODULATION; PHASE MODULATION.

Continuous phase modulation. An analog phase-modulating signal can take on an infinite number of values, and is an excellent choice for transmitting any continuous signal, such as audio, by continuous phase modulation (CPM). The phase-modulated carrier is of the form $A \cos [2\pi f_c t + k_p s(t)]$, where k_p is the phase-modulation index in radians, and $s(t)$ is the amplitude-normalized information signal, where $|s(t)| \leq 1$. In the case of digital data, $s(t) = S_{k(t)} \cdot 2\pi/M$, where the PSK data values, S, are 0, 1, 2, . . . , M−1. The objective is to reduce the required transmission bandwidth, so data bits are collected in sets to represent one of M points in the signal space (a complex plane on which are indicated the relative strengths of in-phase and quadrature components of possible signals of some fixed time duration) around the unit circle, or, alternatively, an angle, θ, where the angular resolution is $\delta\theta = 2\pi/M$. *See* PHASE-ANGLE MEASUREMENT; QUADRATURE.

Consider the most general CPM case. Suppose that we wish to shift that unmodulated sinusoidal carrier through a dynamically changing phase angle $\theta(t)$, where $\theta(t) = k_p s(t)$. Then $A \cos [2\pi f_c t + \theta(t)] = A \cos (2\pi f_c t) \cos [\theta(t)] - A \sin (2\pi f_c t) \sin [\theta(t)]$. We will see later why, for digital transmission, we want $\theta(t)$ to be large, often $\pi/2$. In the continuous case, if $\theta(t)$ is constrained to be small,

$\cos[\theta(t)] \approx 1$ and $\sin[\theta(t)] \approx \theta(t)$, and therefore the phase-shifted sinusoid is very nearly equal to $A \cos(2\pi f_c t) - \theta(t) A \sin(2\pi f_c t)$. This narrow-band phase modulation can therefore be achieved with a simple 90° phase shift of the cosine to generate the sine, multiplication of that sine by $\theta(t)$, and subtraction of the product from the initial cosine.

Wideband phase modulation of a continuous signal offers no such shortcut. We have to form both $\sin(\theta)$ and $\cos(\theta)$ as well as the two products, or rely on frequency multiplication. In frequency multiplication, the signals are produced at a frequency $1/N$th of the target frequencies, and then scaled up in frequency by a factor of N by using a nonlinear device, called a frequency multiplier, to generate a rich set of harmonics, followed by a narrow bandpass filter to preserve only the desired frequency components. *See* FREQUENCY MULTIPLIER.

Binary phase-shift keying. For transmitting digital data, the simplest form of PSK is binary, or BPSK. We can generalize to M-bit PSK, or MPSK. The two-state BPSK data are transmitted by a plain carrier signal for a logical "1" and inverting of its phase for a logical zero. The inverse convention may also be used, but both carry a hidden cost at the receiver or detector because in order to demodulate phase information one must have a reliable reference phase available from a phase-"unshifted" reference signal. *See* PHASE MODULATION DETECTOR.

The BPSK signal $A \cos[2\pi f_c t + k_p s(t)]$ is constrained by $s(t) = \pm 1$, so we may rewrite it as $A \cos(2\pi f_c t) \cos[k_p s(t)] - A \sin(2\pi f_c t) \sin[k_p s(t)] = A \cos(2\pi f_c t) \cos(k_p) - A \cdot s(t) \sin(2\pi f_c t) \sin(k_p)$. The data-independent power of the carrier at frequency f_c is $A^2 \cos^2(k_p)/2$, while the power of the information component is $A^2 \sin^2(k_p)/2$. In order to maximize the useful signal power and minimize the wasted carrier power, it should be obvious that we want to set $k_p = \pi/2$. The carrier vanishes (suppressed carrier), and all the power is profitably invested in the signal. The time duration of each bit is τ seconds, so the energy of each data bit is the product of power and duration, $A^2\tau/2$. The square root of the energy per bit gives the distance between the origin and the point in signal space of the bit. The resulting signal-space locations of our BPSK states are therefore $\pm A(\tau/2)^{1/2}$. The distance between these signal-space locations shows the noise resistance. (How much noise must we add to one point in signal space to move it sufficiently close to another point to confuse the interpretation?) The square of the distance from the origin is proportional to the power required to transmit that information.

Differential binary phase-shift keying (DBPSK) at the transmitter avoids the data recovery problem at the receiver (because an absolute phase reference was not available) by encoding not the phase of the absolute data, but the change in phase from one clock cycle to the next. For example, a "0" may be represented by a change in state (for example, "0" to "1" or "1" to "0"), and a "1" by no change in the following clock cycle. Now, we no longer require an absolute phase reference, but are more susceptible to a new class of errors—a missed transition can invert the interpretation of the data unless error-detection coding is built into the data stream.

We can double the data rate over BPSK by transmitting bit pairs in one of two ways. Suppose we interpret each pair of data bits as an amplitude (for example, $+3, +1, -1, -3$). In a single symbol time, we can transmit this amplitude information. As we noted above, the signal-space locations of these states are the amplitudes scaled by $A(\tau/2)^{1/2}$. The noise immunity looks good because the closest pair of states is the same as in the BPSK case. The cost is increased average power required, $[3^2 + 1^2 + (-1)^2 + (-3)^2]/4$, is five times greater than the BPSK case. On the other hand, let us interpret each pair of bits as $1, j\pi/2, -1, -j\pi/2$. The points in signal space are on a circle of radius $A(\tau/2)^{1/2}$ rather than unity, that is, unit amplitude but varying phase. The signal-space distance between the closest pair is $\sqrt{2}$ (not a bad tradeoff). The noise immunity can be kept at the worst-case BPSK level by bumping up the transmitted power by a factor of $\sqrt{2}$ while the data rate is doubled! Mechanization is simply a 90° phase-shifting network and an inverter. An alternative mechanization is by summing of two BPSK signals on orthogonal carriers, sine and cosine. There are more advantages to phase-encoding bit pairs than there are in amplitude coding.

MPSK. We can combine both amplitude coding and phase coding to improve bandwidth efficiency and again double the data rate. Quadrature PSK (QPSK) encodes data in pairs of bit pairs to signal a phase shift of $0, \pm\pi/2$, or π and amplitudes of ± 1 and ± 3. Mechanization of a phase shift of π can be accomplished either by a phase-shift network of by modulating the signal onto a quadrature carrier, that is, a second carrier that is orthogonal to (or in quadrature with or 90° out of phase with) the first carrier. *See* AMPLITUDE MODULATOR.

Other MPSK schemes are implemented by quadrature amplitude modulation (QAM). ($M = 16$ is frequently used in cable television converters.) As we try to increase bandwidth efficiency by combining more bits into a single symbol in signal space, the separation of those symbols in signal space decreases, making the transmissions more susceptible to noise-induced errors.

Stanley A. White

Bibliography. G. A. Breed (ed.), *Wireless Communications Handbook*, Cardiff Publishing, 1992; S. Gibilisco, *Handbook of Radio and Wireless Technology*, McGraw-Hill, 1999; J. G. Proakis, *Digital Communications*, 4th ed., McGraw-Hill, 2001; T. S. Rappaport, *Wireless Communications*, 2d ed., Prentice Hall, 2002; M. S. Roden, *Analog and Digital Communications Systems*, 4th ed., Prentice Hall, 1995; M. E. Van Valkenburg and W. M. Middleton (eds.), *Reference Data for Engineers: Radio, Electronics, Computer and Communications*, 9th ed., Butterworth-Heinemann, 2002.

Phase rule

A relationship used to determine the number of state variables F, usually chosen from among temperature, pressure, and species compositions in each phase, which must be specified to fix the thermodynamic state of a system in equilibrium. It was derived by J. Willard Gibbs between 1875 and 1878. The phase rule (in the absence of electric, magnetic, and gravitational phenomena) is given by Eq. (1), where C is the

$$F = C - P - M + 2 \tag{1}$$

number of chemical species present at equilibrium, P is the number of phases, and M is the number of independent chemical reactions. Here phase is used to indicate a homogeneous, mechanically separable portion of the system, and the term independent reactions refers to the smallest number of chemical reactions which, upon forming various linear combinations, includes all reactions which occur among the species present. The number of independent state variables F is referred to as the degrees of freedom or variance of the system.

Examples. The system constituted by liquid water and water vapor contains two phases ($P = 2$) and one component ($C = 1$), and there are no chemical reactions ($M = 0$). Therefore this system has $F = 1 - 2 - 0 + 2 = 1$ degree of freedom. This is in accord with the observation that the vapor and liquid forms of water exist in equilibrium only for values of temperature and pressure along the coexistence curve, so that specifying either of these variables fixes the other.

At high temperatures the three reactions shown in (2) occur between sulfur and oxygen in the gas

$$S + O_2 \longrightarrow SO_3 \tag{2a}$$

$$SO_2 + {}^1/_2O_2 \longrightarrow SO_3 \tag{2b}$$

$$S + {}^3/_2O_2 \longrightarrow SO_3 \tag{2c}$$

phase. There are only two independent chemical reactions in this single-phase, four-component system, since the last reaction is the sum of the first two. Therefore this system has $F = 4 - 1 - 2 + 2 = 3$ degrees of freedom.

Derivation. The derivation of the phase rule starts with the experimental observation that a single phase of C nonreacting components has $C + 1$ degrees of freedom (for example, once temperature, pressure, and $C - 1$ mole fractions are specified, no other intensive variables of the system can vary). Now consider chemical and phase equilibrium in a more general system in which there are C components and P phases, and in which M independent chemical reactions occur. It might appear, based on the experimental observation above, that such a system should have $P(C + 1)$ degrees of freedom. However, since the system is in equilibrium, the degrees of freedom are reduced as follows:

1. At equilibrium the temperature of each phase must be the same. Thus it is not possible to set the temperature of each of the P phases separately; once the temperature of one phase is specified, the temperature of all phases is fixed. Consequently the requirement that the temperature must be the same in all phases at equilibrium eliminates $P - 1$ degrees of freedom.
2. At equilibrium the pressure must be the same in each phase. This eliminates another $P - 1$ degrees of freedom.
3. At equilibrium the chemical potential (partial molar Gibbs free energy) of each species must be the same in each phase. This restriction, which holds for each of the C components, eliminates an additional $C(P - 1)$ degrees of freedom.
4. For each chemical reaction (3) to be in equilib-

$$aA + bB + \cdots \longrightarrow rR + sS + \cdots \tag{3}$$

rium, Eq. (4) must be satisfied, where μ_i is the chem-

$$a\mu_A + b\mu_B + \cdots \longrightarrow r\mu_R + s\mu_S + \cdots \tag{4}$$

ical potential of species i. (If this equation is satisfied in one phase, it is, by the equality of the chemical potentials of a given species in its various phases, satisfied in all phases.) This requirement eliminates another M degrees of freedom.

Therefore, the actual number of degrees of freedom in a multicomponent, multiphase, chemically reacting system is given by Eq. (5), which is the

$$\begin{aligned} F &= P(C + 1) - 2(P - 1) - C(P - 1) - M \\ &= C - P - M + 2 \end{aligned} \tag{5}$$

same as Eq. (1). *See* CHEMICAL EQUILIBRIUM; CHEMICAL THERMODYNAMICS; PHASE EQUILIBRIUM; THERMODYNAMIC PROCESSES.

Stanley I. Sandler

Phase-transfer catalysis

A process in which the rate of an organic reaction in a heterogeneous two-phase system is enhanced by the addition of a compound that transfers one of the reactants across the interface between the two phases.

An important factor, which contributes to the slowness of many organic reactions, is the lack of homogeneity of the reaction mixture. This is particularly the case with nucleophilic substitution reactions (1), where RX is an organic reagent and Nu^- is

$$RX + Nu^- \rightarrow RNu + X^- \tag{1}$$

the nucleophilic reagant. The nucleophilic reagent is frequently an inorganic anion, which is soluble in water in which the organic substrate is insoluble, but is insoluble in the organic phase. The encounter rate between Nu^- and RX is consequently low as they can only meet at the interface of the heterogeneous system. The water-soluble anion is also frequently highly solvated by water molecules, which stabilize the anion and thus reduce its nucleophilic reactivity. These problems have been overcome in the past by the use of polar aprotic solvents, which will dissolve both the organic and inorganic reagents, or by the use of homogeneous mixed-solvent systems, such as

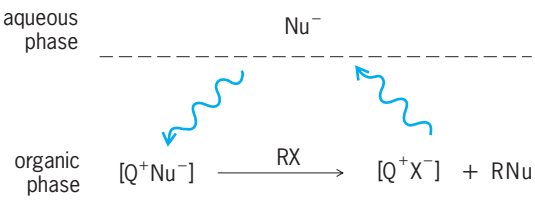

Fig. 1. Formation and transportation of the reactive ion pair in phase-transfer catalysis.

water:ethanol or water:dioxan. Homogeneous reaction systems can also be established by the use of surfactants, and the interfacial area can be increased by rapid agitation or mechanical emulsification. Although these procedures may increase the rate of reaction, they have disadvantages, such as difficulties in the isolation and purification of the products and, on an industrial scale, they can be costly in terms and solvents and energy.

Phase-transfer catalysis involves the transportation of the inorganic anion, Nu^-, from the aqueous phase into the organic phase by the formation of a nonsolvated ion pair with a cationic phase-transfer catalyst, Q^+. It was originally believed that this process involved the formation of the ion pairs in the aqueous phase, followed by their partition between the aqueous and organic phases. A more current explanation is that with highly lipophilic catalyst, the reactive ion pair $[Q^+Nu^-]$ is formed at the interface between the aqueous and organic phases, followed by rapid transportation into the bulk of the organic phase (**Fig. 1**). The rate of the reaction is enhanced, because the encounter rate of the nucleophile, Nu^-, with the organic reagant, RX, in the single phase will be significantly higher than at the interface. Moreover, as the anion is transferred virtually without any water of solvation, its nucleophilic reactivity can be considerably higher in the organic phase than in the aqueous phase. Rate enhancements of greater than 10^7 have thus been observed.

It has been noted, however, that several reactions in which the nucleophile is generated by abstraction of a proton from an organic substrate under basic two-phase conditions are ineffective when the reaction is conducted using the aqueous-organic two-phase conditions, described above, and are better accomplished using solid powdered potassium hydroxide [KOH; or what is even better, an intimate mixture of powdered potassium hydroxide and potassium carbonate (K_2CO_3)] instead of aqueous sodium hydroxide. It has been proposed that under these conditions the pH of the medium can be greater than 30. It has been further observed that the efficiency of many phase-transfer catalyzed liquid:liquid two-phase reactions can be improved by the omission of the aqueous solvent and the use of finely divided solid inorganic salts in solid:liquid two-phase conditions.

Although the process does not yet appear to have widespread application, phase-transfer catalyzed gas:liquid and gas:solid two-phase reaction systems have also been reported.

The efficiency of the catalyst in both liquid:liquid and solid:liquid two-phase systems depends upon the ability of the cation, Q^+, to transfer the anion across the interface of the heterogeneous system; that is, in the case of the aqueous:organic two-phase systems, the rate should be proportional to the partition coefficient of the ion pair $[Q^+Nu^-]$ and the water solubility of $[Q^+X^-]$ should be higher than that of $[Q^+Nu^-]$. It has been found that bulky quaternary ammonium cations, R_4N^+, such as tetra-*n*-butylammonium and methyltrioctylammonium ions, and phosphonium cations, R_4P^+, such as tetra-*n*-butylphosphonium cations, have a high propensity to form "strong" ion pairs, which have a high solubility in organic solvents, and are excellent phase-transfer catalysts.

As the partition coefficient of the ion pair is a significant factor in the catalytic effect, the choice of the organic phase is important. The most commonly used organic phases are the semipolar solvents, dichloromethane and 1,2-dichlorobenzene, and the less polar solvent, toluene.

An alternative phase-transfer catalytic procedure utilizes polyethers, which are capable of complexing alkali metal cations. Ion pairs of the complexed cation and the nucleophilic anion are formed and are transported across the water:organic phase interface in the same manner as the quaternary ammonium and phosphonium salts. Polyether catalysts as shown in **Fig. 2***a* may be acyclic [glymes; (**1**)], or monocyclic [crown ethers; (**2**)], or polycyclic (cryptands; (**3**)]. The most commonly used polyether catalysts are the crowns, such as dicyclohexanono-18-crown-6, which complex readily with potassium and sodium ions to form complexes of type (**4**) [Fig. 2*b*].

Advances in phase-transfer catalyst design has led to increased use of tridents [tris(polyalkoxyalkyl)-amines], of which TDA-1 [tris(3,6-dioxaheptyl)-amine; (**8**) Fig. 2*d*] is an example. These compounds are acyclic forms of the cryptands and have the capability to form complexes with a wide range of cationic species. The trident catalyst therefore has a potential use in oxidation:reduction reactions through the transport of transition-metal cations, as well as in the standard nucleophilic substitution reactions. This type of compound has particularly good application in the catalysis of solid:liquid two-phase reactions.

There are also descriptions of the use of bis-quaternary ammonium salts as phase-transfer catalysts. Compounds such as BBDE [bis(2-(benzyldiethylammonio)ethyl)ether dichloride; (**9**); Fig. 2*e*] not only have the capability of transporting two equivalents of a mononucleophilic species across the interface of two phases, but also efficiently transport dinucleophilic species, for example, the phthalate dianion. *See* COORDINATION COMPLEXES; MACROCYCLIC COMPOUND.

Applications. Phase-transfer catalysts can be used to enhance reaction rates in nucleophilic substitution reactions, oxidation-reduction reactions, generation of reactive intermediates, asymmetric induction, and carbonylation reactions.

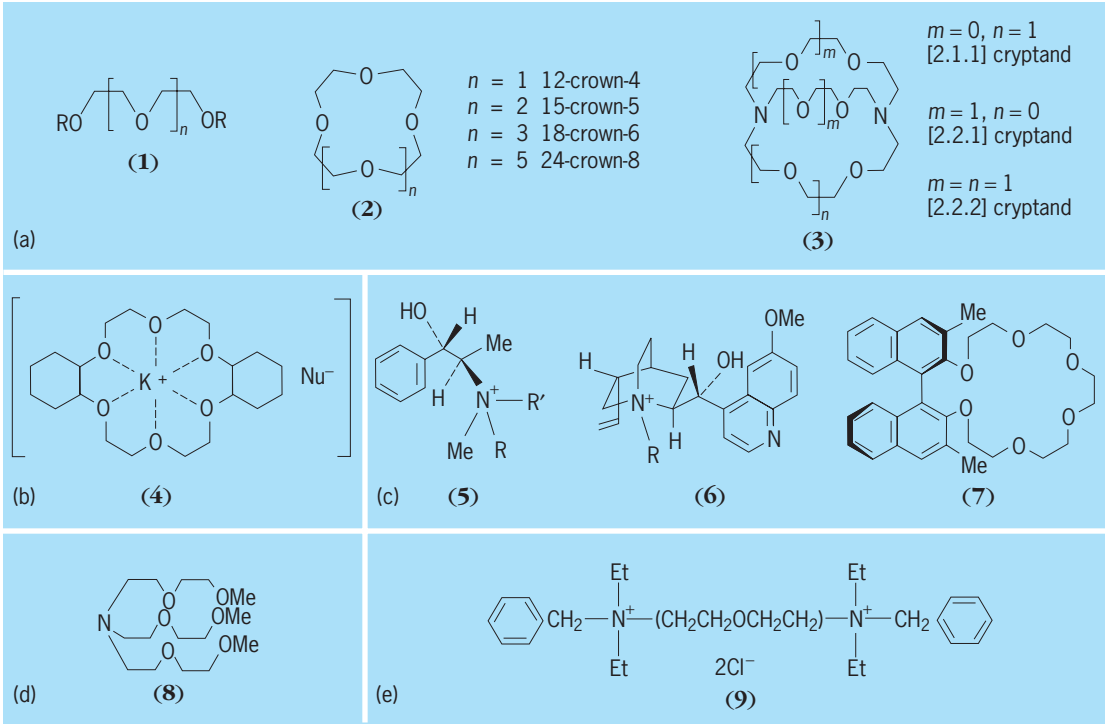

Fig. 2. Structures of some involved in phase-transfer catalysis. Me = CH_3 Et = C_2H_5. (*a*) Typical polyether catalysts used in phase-transfer catalyst procedure. (*b*) Typical complex formed with a polyether catalyst. (*c*) Chiral catalysts used in asymmetric induction. (*d*) Typical tris(polyoxaalkyl)amine (trident) catalyst. (*e*) Example of bis-quaternary ammonium salt catalyst.

Nucleophilic substitution reactions. By using phase-transfer catalysis, it is possible to enhance the rate of most nucleophilic substitution reactions, as in reaction (1), with a wide range of nucleophilic anions.

Kinetic studies have shown that the nucleophilicity of the nonhydrated halide anions, produced upon dissolution of the quaternary-onium halides in organic solvents, follow the order fluorine > chlorine > bromine > iodine. Under these conditions the fluoride ion is a strong base and can be used in preference to the hydroxide ion in base-catalyzed reactions. *See* STERIC EFFECT (CHEMISTRY).

Oxidation and reduction reactions. Purple benzene (or purple dichloromethane) and yellow benzene (or yellow dichloromethane) are readily obtained by the dissolution of permanganate or dichromate ions, respectively, in the otherwise colorless benzene (or dichloromethane) by using phase-transfer catalysts, and the solutions have been used for a wide range of oxidation reactions in high yields with minimal side reactions. Solid tetraalkylammonium permanganates and dichromates are potentially explosive.

Hydride reductions have been conducted in organic solvents by using tetraalkylammonium borohydrides.

Diborane, B_2H_6, produced from the reaction of tetraalkylammonium borohydride in dichloromethane with simple alkyl halides [reaction (2)] has an

$$2[Q^+BH_4^-] + 2CH_3I \rightarrow 2CH_4 + 2[Q^+I^-] + B_2H_6 \qquad (2)$$

activity similar to that of diborane in ether and can be used for all the usual reduction and hydroboration reactions. *See* HYDROBORATION; OXIDATION-REDUCTION.

Generation of reactive intermediates. Reactive carbanions, which normally require strictly anhydrous reaction conditions for their preparation, can be generated by using the phase-transfer catalytic technique [reaction (3)], where Z = acyl, alkoxycarbonyl, cyano, ano, or nitro, and R = H, alkyl, or aryl.

$$\underset{\text{(org)}}{RCHZ_2} + \underset{\text{(aq)}}{HO^-} \xrightarrow{[R'_4N^+X^-]} \underset{\text{(org)}}{[RCZ_2^-R'_4N^+]} \xrightarrow{R''X} \underset{\text{(org)}}{RR''CZ_2} \qquad (3)$$

Similarly, other systems, which are labile in the presence of water such as RCOCN, can be synthesized by the phase-transfer catalytic procedure.

Tetraalkylammonium hydroxides react with chloroform to yield, initially, the trichloromethyl carbanion, which loses a chloride ion to generate the highly reactive dichlorocarbene [reaction (4)]. It has

$$\underset{\text{(org)}}{CHCl_3} \xrightarrow{HO^-(aq)} \underset{\text{(interface)}}{CCl_3^- + H_2O} \xrightarrow{[Q^+Cl^-]} \underset{\text{(org)}}{[Q^+CCl_3^-]} \rightarrow \underset{\text{(org)}}{[Q^+Cl^-]} + \underset{\text{(org)}}{C\ddot{C}l_2} \qquad (4)$$

been proposed that the abstraction of the proton from the chloroform occurs at the interface and that the role of the catalyst is the transfer of the trichloromethyl carbanion into the bulk of the organic solvent. The dichlorocarbene reacts rapidly and in high product yield with a wide range of organic substrates.

Other dihalogenocarbenes, and vinylidene and vinyl carbenes have been obtained by phase-transfer catalyzed procedures. *See* REACTIVE INTERMEDIATES.

Carbonylation reactions. Reactions of organic substrates with metal carbonyl compounds, which frequently require high pressures of carbon monoxide (CO), are feasible under relatively mild conditions when conducted under phase-transfer catalytic conditions. For example, the reaction of iodoalkanes with dicobalt octacarbonyl to give acylcobalt tetracarbonyls [reaction (5)] is catalyzed by quaternary ammonium salts. Subsequent reaction of the acylcobalt derivative with alkenes or alkyl halides leads to formation of ketones, while further reaction with a base leads to carboxylic acids. Reduction and oxidation reactions, which are initiated by metal carbonyl derivatives, are also catalyzed by the addition of quaternary ammonium salts. *See* METAL CARBONYL.

$$\underset{\text{Iodoalkane}}{RCH_2I} + \underset{\text{Dicobalt octacarbonyl}}{Co_2(CO)_8} \xrightarrow[Q^+\ Cl^-]{\text{aqueous NaOH}} \underset{\text{Acylcobalt tetracarbonyl}}{RCH_2COCo(CO)_4} \quad (5)$$

Asymmetric induction. Chiral ammonium catalysts having both a chiral quaternary nitrogen atom and chirality within the carbon skeleton, such as structures (**5**) and (**6**) [Fig. 2*c*], have been used successfully in the stereochemical control of reactions involving carbonyl-containing compounds. A critical feature of the chiral catalyst is the presence of a β-hydroxyethyl substituent on the chiral nitrogen atom, which can hydrogen-bond with the carbonyl group of the organic substrate, thereby preferentially presenting one "face" of the substrate for reaction. Reactions in which an enantiometric excess is >60% have been observed.

Stereochemical control of organic reactions has also been accomplished by using chiral crown ethers, such as structure (**7**). *See* ASYMMETRIC SYNTHESIS.

Solid-supported phase-transfer catalysis. As an elaboration of liquid:liquid catalysis, the quaternary ammonium group or crown ether is attached to an organic polymeric support, such as polystyrene. Quaternary ammonium salts have also been immobilized on zeolites. Although the efficiency of such catalysts differs little from the conventional soluble phase-transfer catalysts, there are advantages in the use of triphase (that is liquid:liquid:catalyst) systems. The catalyst is readily removed by filtration from the reaction system after use and, on an industrial scale, there is a potential for continuous phase-transfer catalyzed processes. *See* CATALYSIS; HETEROGENEOUS CATALYSIS; HOMOGENEOUS CATALYSIS; QUATERNARY AMMONIUM SALTS; STEREOCHEMISTRY; ZEOLITE.

R. Alan Jones

Bibliography. H. des Abbayes, Hydroxide ion initiated reactions under phase-transfer catalysis conditions, *Israel J. Chem.*, 26:249–262, 1985; E. V. Dehmlow and S. S. Dehmlow, *Phase-Transfer Catalysis*, 3d ed., 1993; W. E. Keller (ed.), *Compendium of Phase-Transfer Reactions and Related Synthetic Methods*, 2 vols., 1979, 1987; C. Liotta, C. M. Starks, and M. Halpern, *Phase-Transfer Catalysis: Fundamentals, Applications, and Industrial Perspectives*, 1994; M. Rabinovitz, Y. Cohen, and M. Halpern, Metal carbonyls in phase-transfer catalysis, *Angew. Chem. Int. Ed. Engl.*, 25:960–970, 1986; G. Soula, Tris(polyoxaalkyl) amines (Trident), a new class of solid-liquid phase-transfer catalysts, *J. Org. Chem.*, 50:3717–3721, 1985; C. M. Starks (ed.), *Phase-Transfer Catalysis: New Chemistry, Catalysts, and Applications*, American Chemical Society Symposium Series, no. 326, 1987; W. P. Weber and G. W. Gokel, *Phase-Transfer Catalysis in Organic Synthesis*, 1977.

Phase transitions

Changes of state brought about by a change in an intensive variable (for example, temperature or pressure) of a system. Some familiar examples of phase transitions are the gas-liquid transition (condensation), the liquid-solid transition (freezing), the normal-to-superconducting transition in electrical conductors, the paramagnet-to-ferromagnet transition in magnetic materials, and the superfluid transition in liquid helium.

Characteristics

Typically the phase transition is brought about by a change in the temperature of the system. The temperature at which the change of phase occurs is called the transition temperature (usually denoted T_t). As examples: the solid-liquid transition occurs at the melting point (denoted T_m); the liquid-gas transition occurs at the boiling point (denoted T_b). The condition of temperature and pressure, unique for each substance, at which the distinction between liquid and gaseous fluids disappears is called the critical point (denoted T_c). T_c is also used to denote the critical temperature below which ferromagnetism begins to grow in some materials.

Order and entropy. The two phases above and below the phase transition can be distinguished from each other in terms of some ordering that takes place in the phase below the transition temperature. For example, in the liquid-solid transition the molecules of the liquid get "ordered" in space when they form the solid phase. In a paramagnet the magnet moments on the individual atoms can point in any direction (in the absence of an external magnetic field), but in the ferromagnetic phase the moments are

lined up along a particular direction, which is then the direction of ordering. Thus in the phase above the transition the degree of ordering is smaller than in the phase below the transition. One measure of the amount of disorder in a system is its entropy, which is the negative of the first derivative of the thermodynamic free energy with respect to temperature. When a system possesses more order, the entropy is lower. Thus at the transition temperature the entropy of the system changes from a higher value above the transition to some lower value below the transition. *See* ENTROPY; FERROMAGNETISM; PARAMAGNETISM.

Continuous and discontinuous transitions. This change in entropy can be continuous or discontinuous at the transition temperature. In other words, the development of order in the system at the transition temperature can be gradual or abrupt. This leads to a convenient classification of phase transitions into two types, namely, discontinuous and continuous.

Discontinuous transitions involve a discontinuous change in the entropy at the transition temperature. A familiar example of this type of transition is the freezing of water into ice. As water reaches the freezing point, order develops without any change in temperature. Thus there is a discontinuous decrease in the entropy at the freezing point. This is characterized by the amount of latent heat that must be extracted from the water for it to be "ordered" into the solid phase (ice). Discontinuous transitions are also called first-order transitions.

In a continuous transition, entropy changes continuously, and hence the growth of order below T_c is also continuous. There is no latent heat involved in a continuous transition. Continuous transitions are also called second-order transitions. The paramagnet-to-ferromagnet transition in magnetic materials is an example of such a transition.

Order parameter. The degree of ordering in a system undergoing a phase transition can be made quantitative in terms of an order parameter. At temperatures above the transition temperature the order parameter has a zero value, and below the transition it acquires some nonzero value. For example, in a ferromagnet the order parameter is the magnetic moment per unit volume (in the absence of an externally applied magnetic field). It is zero in the paramagnetic state since the individual magnetic moments in the solid may point in any random direction. Below the critical transition temperature T_c, there exists a preferred direction of magnetic ordering. As the temperature is decreased below T_c, more and more individual magnetic moments start to align along the preferred direction of ordering, leading to a continuous growth of the magnetization or the macroscopic magnetic moment per unit volume in the ferromagnetic state. Thus the order parameter changes continuously from zero above to some nonzero value below the transition temperature. In a first-order transition the order parameter would change discontinuously at the transition temperature.

The behavior of systems very near the onset of phase transitions is an important field of study in itself. For a detailed description of this behavior *see* CRITICAL PHENOMENA.

Examples

Several modern examples of phase transitions or quasi-phase transitions will be discussed. By quasi-phase transition is meant an apparent, diffuse transition which may have an underlying thermodynamic basis but for which kinetic factors also play a role in determining the apparent transition temperature. A number of well-known transitions are mentioned only briefly or excluded from the discussion here; for these transitions *see* FERROMAGNETISM; SUPERCONDUCTIVITY.

Glass transitions. When certain substances are taken very rapidly from the liquid state to the solid state, or condensed from the vapor phase to the solid phase, they are found to have an amorphous or glassy structure whose atomic arrangement differs from those of both crystals and gases. Similar features are observed in substances composed of chain-like molecules in random arrangement. In the crystal, the atoms are placed on a regular, periodic lattice, and only discrete interatomic spacings between the atoms are possible (**Fig. 1***a*); this describes a state of long-range order. In the gas, the atoms are constantly

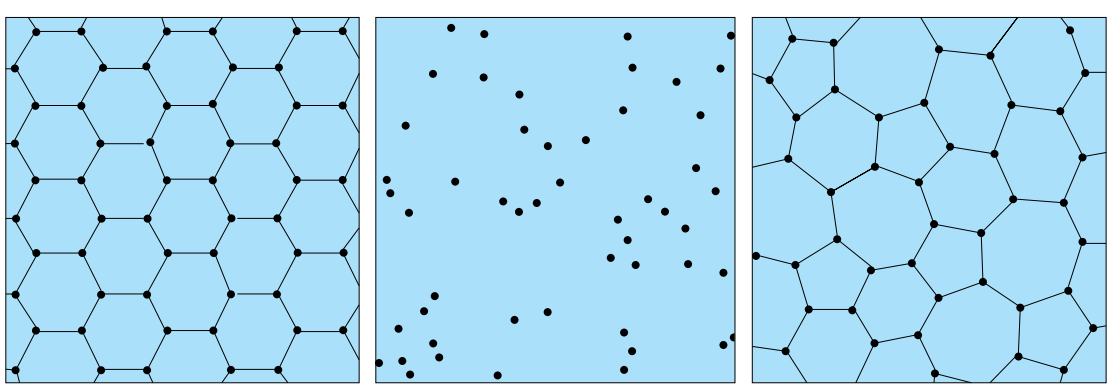

(a) (b) (c)

Fig. 1. Two-dimensional representation of atomic arrangements in three phases: (*a*) Crystal. (*b*) Gas. (*c*) Glass. Each dot represents position of an atom at a given instant of time. (*After E. W. Montroll and J. L. Leibowitz, Fluctuation Phenomena, North-Holland Publishing, 1979*)

colliding with each other and the walls of the container (Fig. 1*b*)—there is no long-range order. The glass phase has a structure intermediate between that of the crystal and gas (Fig. 1*c*). The local environment or short-range order of each atom in the glass shown is similar to that of the crystal, but there is no long-range order. Some glasses have atomic arrangements, at any instant of time, very similar to those of a liquid of the same composition. The structure of certain amorphous metallic alloys is similar to that which would be obtained by packing hard spheres into a container to obtain the largest density.

The glass transition temperature T_g is defined to be the temperature below which the atoms are frozen into relatively stable positions, even though the glassy phase is a metastable one whose energy is higher than the crystalline phase. When a glass is heated above T_g, the material can transform into the crystalline state by the formation of nucleation centers in the solid state. In metallic glasses, this crystallization process can take place with only a little heating above T_g. In other glasses, such as fused silica, crystallization does not take place above T_g, no matter how long one waits; the next transition in such a material as the temperature is increased is to the liquid state, that is, melting. The amorphous solid-liquid transition is an example of a quasi-phase transition as defined above. The glass transition appears to possess an underlying equilibrium thermodynamic character but is modified by kinetic effects. Some examples of insulating or semiconducting glasses are silicon dioxide (SiO_2), arsenic selenide (As_2Se_3), silicon, (Si), and germanium (Ge). Examples of metallic glasses are $Pd_{80}Si_{20}$, $Fe_{80}B_{20}$, $Gd_{65}Co_{35}$, $Zr_{40}Cu_{60}$, and $Mg_{70}Zn_{30}$. Both classes of materials have been intensively studied with a view toward significant technological applications. *See* AMORPHOUS SOLID; GLASS; METALLIC GLASSES.

Ferromagnetism and antiferromagnetism. As a preliminary to discussing several more complicated magnetic phase transitions, the fundamentals of simple ferromagnetic and antiferromagnetic transitions will be outlined. The fundamental interaction energy E_{ij} between two spins S_i and S_j can be represented by Eq. (1), where $\mathscr{J}_{ij}$ is called the exchange interaction

$$E_{ij} = -\mathscr{J}_{ij}\vec{S}_i \cdot \vec{S}_j \tag{1}$$

because it originates in the quantum-mechanical requirement that the wave function of a system of electrons must be antisymmetric (change sign) under exchange of the coordinates of any two electrons. Although Eq. (1) is a quantum-mechanical effect, it will be discussed from a classical viewpoint since much insight can be gained in this way. Classically, Eq. (1) can be written as Eq. (2), where θ_{ij} is the angle be-

$$E_{ij} = -\mathscr{J}_{ij}S_iS_j\cos\theta_{ij} \tag{2}$$

tween the two spins. When $\mathscr{J}_{ij}$ is positive, E_{ij} is minimized for $\vec{S}_i$ and $\vec{S}_j$ parallel; when $\mathscr{J}_{ij}$ is negative, E_{ij} is minimized for $\vec{S}_i$ and $\vec{S}_j$ antiparallel. These two cases are shown schematically in **Fig. 2***a* and *b* as ferromagnetic and antiferromagnetic order, respectively. The magnetic moment vector associated with each spin is proportional to the magnitude of the spin and points along the same line. In the absence of thermal fluctuations, that is, at absolute temperature $T = 0$, the spins of a ferromagnet are all parallel in a single domain so that a large magnetization results (Fig. 2*a*). For an antiferromagnet, again at $T = 0$, there are two sublattices whose spins are oppositely directed so that there is no net magnetization (Fig. 2*b*). As the temperature is raised from near zero, the spins in either case begin to deviate from their original positions until, at the magnetic ordering temperature, a phase transition to the paramagnetic phase takes place (Fig. 2*c*). In the paramagnetic phase, there is a complete lack of long-range order in the spin directions. *See* ANTIFERROMAGNETISM; EXCHANGE INTERACTION.

Magnetism in crystalline alloys. A disordered crystalline alloy is a mixture of two elements in which the atoms of the mixture are found at more or less random positions on a crystal lattice. If the concentration of one element is x (and the other $1 - x$), then some typical examples of disordered binary alloys are $Cu_{1-x}Mn_x$, $Au_{1-x}Fe_x$, and $Ag_{1-x}Pd_x$. If one of

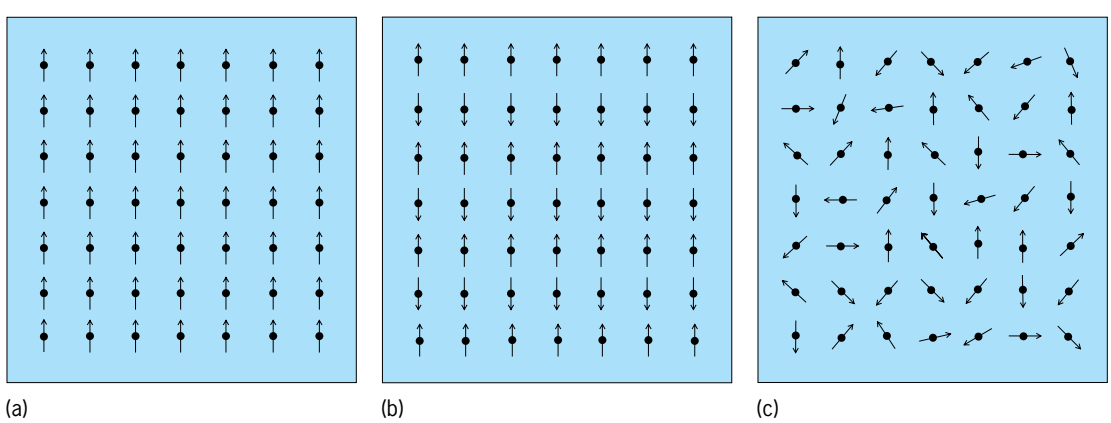

Fig. 2. Ordering of spins in a crystal lattice. Arrows show the direction in which the spin (or the magnetic moment) at a particular lattice site is pointing. (*a*) Ferromagnet. (*b*) Antiferromagnet. (*c*) Paramagnet.

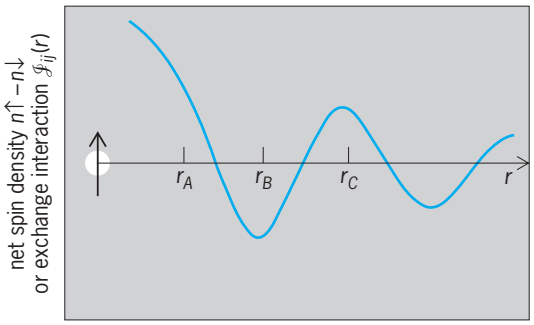

Fig. 3. Net spin density of conduction electrons (number of electrons per unit volume with spin up $n\bar{u}$ minus number with spin down $n\bar{d}$) as function of distance r from magnetic moment. Graph also shows exchange interaction $\mathscr{J}_{ij}(r)$ between magnetic moments i and j as function of distance r between them. Magnetic moment at the origin will couple ferromagnetically to magnetic moments at positions r_A and r_C, but antiferromagnetically to a magnetic moment at r_B.

the constituents of the alloy has a magnetic moment and therefore spin associated with it, there will be disorder in the atomic positions of the spins, yet it is possible for certain types of magnetic order to exist. In the three examples mentioned above, the manganese (Mn) and iron (Fe) atoms do have localized magnetic moments associated with them.

Spin-glass state. Consider a relatively dilute alloy such as $Cu_{0.95}Mn_{0.05}$. In such a case the conduction electrons will scatter from the manganese impurity atoms and create a disturbance in the net spin density in the vicinity of each manganese atom. This disturbance (net spin up minus spin down) is shown in **Fig. 3**. If another manganese atom is found in the vicinity of the atom at the origin, it will have a spin that is coupled to the latter according to Eqs. (1) and (2). However, in this case $\mathscr{J}_{ij}$ is a function of the distance r between the two spins and has the behavior shown by Fig. 3. Thus $\mathscr{J}_{ij}(r)$ oscillates in sign and decreases in magnitude as r increases. A magnetic moment at the origin of Fig. 3 will couple ferromagnetically to magnetic moments at positions r_A and r_C, but antiferromagnetically to a magnetic moment at r_B. Since there will be a considerable number of different Mn-Mn interatomic spacings, there will be a mixture of ferromagnetic and antiferromagnetic exchange interactions in such an alloy. This mixture leads to a new magnetic phase called a spin glass. In analogy with the ordinary glass shown in Fig. 1*c*, a spin glass has short-range magnetic order but no long-range magnetic order, even below the spin-glass ordering temperature T_{SG} (**Fig. 4***a*). Above T_{SG}, thermal fluctuations will destroy the ordered spin-glass state, and the magnetic system will become paramagnetic, with no short- or long-range magnetic order (Fig. 4*b*). *See* SPIN GLASS.

Ferromagnetic state. In certain disordered alloys such as $Au_{1-x}Fe_x$, for x values larger than about 0.16, the spin-glass phase is replaced by the ferromagnetic state (Fig. 4*c*). At this iron concentration, called the percolation limit, it is possible for nearest-neighbor ferromagnetic couplings of iron spins to propagate completely throughout the crystal, thus leading to the state in which all iron spins are parallel. This situation depends on $\mathscr{J}_{ij}(r)$ having positive values for nearest neighbors, as depicted schematically in Fig. 3 for $r = r_A$. If the nearest-neighbor magnetic couplings are negative (antiferromagnetic), as is thought to be the case in $Cu_{1-x}Mn_x$, then the ordered state at higher manganese concentrations becomes a random antiferromagnetic state in which, to the extent possible, nearest-neighbor spins are antiparallel.

The question of whether the spin-glass and spin-glass–to–ferromagnet transitions are true or quasi-phase transitions is still a subject of much controversy.

Speromagnetic and asperomagnetic states. Speromagnetic and asperomagnetic transitions, which occur in substances having amorphous atomic structures, were characterized in the 1970s. In **Fig. 5***a* a schematic diagram is shown of a glass in which each atom possesses a magnetic moment which is assumed to be coupled to its neighbors by the Heisenberg interaction of Eq. (1). Even though the interatomic spacings are smeared out compared to the

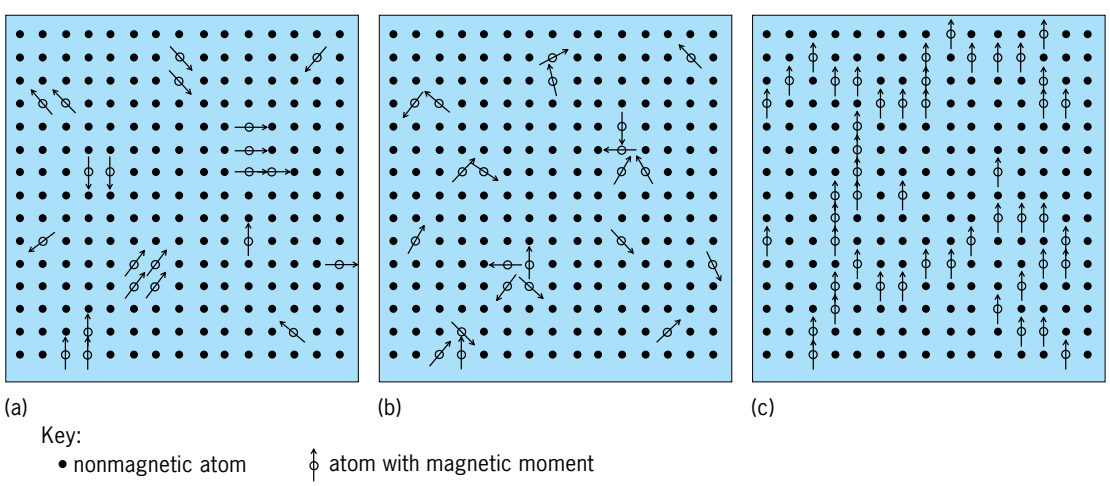

Fig. 4. Arrangement of magnetic moments in crystal lattice of a disordered crystalline alloy. Arrows show the direction in which the magnetic moment at a particular site is pointing. (*a*) Spin-glass state. (*b*) Paramagnetic state. (*c*) Ferromagnetic state.

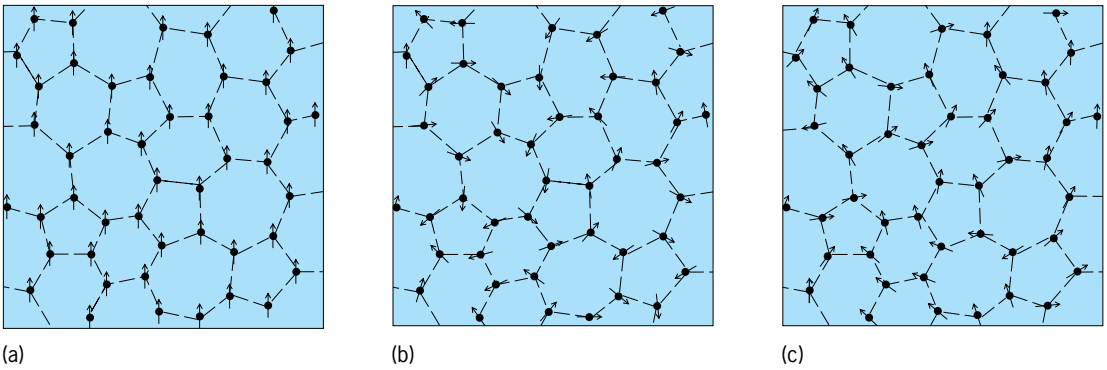

Fig. 5. Arrangement of magnetic moments in (*a*) ferromagnetic glass, (*b*) speromagnetic glass, and (*c*) asperomagnetic glass. Arrows show the direction in which the magnetic moment at a particular site is pointing.

case of a crystal, if all the exchange interactions $\mathscr{J}_{ij}$ are positive the state of magnetic order will be as shown in Fig. 5*a*. An example of a ferromagnetic metallic glass is $Gd_{65}Cu_{35}$, in which all the gadolinium (Gd) moments are aligned while the copper (Cu) atoms possess no moments.

A novel state of order can exist in rare-earth glasses such as $Dy_{75}Au_{25}$, $Tb_{75}Ga_{25}$, and $Er_{65}Co_{35}$. In these cases the rare-earth atoms have magnetic moments which are due to both electron spin and orbital angular momentum. The electric fields which exist in such a glass in combination with the spin-orbit interaction make it difficult for the total angular momentum vector $\vec{J}$ and hence the magnetic moment of each rare-earth atom to be aligned in an arbitrary direction. There can be an "easy" axis $\hat{k}_i$ for each $\vec{J}_i$ such that $\vec{J}_i$ tries to align itself parallel or antiparallel to $\hat{k}_i$. A model that describes this situation of the energy of the ith ion is given by Eq. (3), where $\hat{k}_i$ is a

$$E_i = -D(\hat{k}_i \cdot \vec{J}_i)^2 \tag{3}$$

unit vector and D the strength of the local anisotropy. The random anistropy model of magnetism in glasses assumes that the amorphous structure leads to completely random anistropy axes $\hat{k}_i$ for each magnetic moment. Then the total energy is given by Eq. (4).

$$E = -\mathscr{J}\sum_{i,j}\vec{J}_i \cdot \vec{J}_j - D\sum_i(\hat{k}_i \cdot \vec{J}_i)^2 \tag{4}$$

If $\mathscr{J}$ is positive, the first term tends to cause parallel alignment of the $\vec{J}_i$ vectors while the second term causes a scattering of the $\vec{J}_i$ vectors. This leads in the ordered state to either a speromagnetic structure (Fig. 5*b*) or an asymmetric speromagnetic or asperomagnetic structure (Fig. 5*c*). The origin of the term speromagnetism is the Greek word for scattered, *spero*. The speromagnetic state has a more or less random distribution of moment directions which are frozen in below the magnetic ordering temperature (**Fig. 6***a*). This state of magnetic order has some common features with the spin-glass state, but the origin of the scatter in the frozen moment directions is different. In the asperomagnetic state, the positive exchange interaction is assumed to lead to mostly parallel or nearly parallel near-neighbor moment directions so that the moment distribution would be in the upper hemisphere (Fig. 6*b*). The true ground state of the system, even with all exchange interactions positive, has been the subject of considerable controversy, but research suggests that the ground state is not the asperomagnetic state but rather has some fraction of the moments in the lower hemisphere.

Metallic glasses of the form $Fe_{80}G_{20}$, where G represents a glass-forming element such as boron, phosphorus, or silicon, may have possible uses as soft magnetic materials in transformer cores. Metallic glasses containing rare-earth elements and transition metals may find applications as magnetooptic storage and magnetic bubble computer memory devices. *See* MAGNETOOPTICS.

Order-disorder transition. In an ordered binary alloy AB of the β'-brass type, A occupies the corner position and B the body-centered position of the

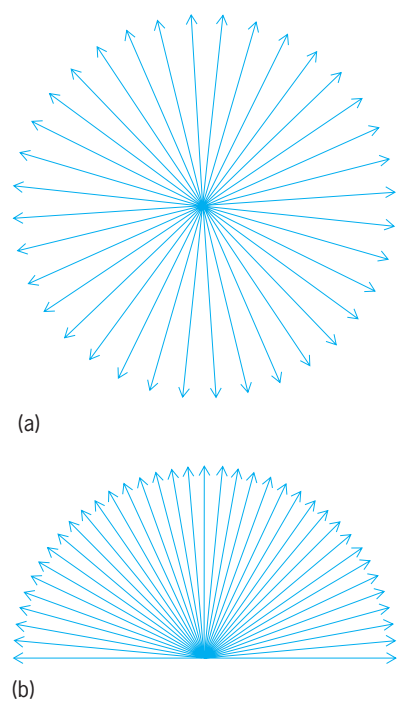

Fig. 6. Distribution of magnetic moments in (*a*) speromagnetic glass and (*b*) asperomagnetic glass.

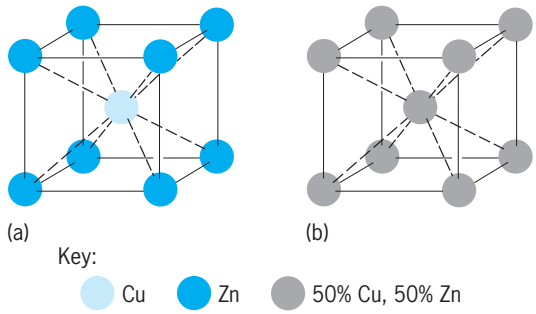

Fig. 7. Structures of (*a*) ordered, or β', brass and (*b*) disordered, or β, brass.

body-centered cubic structure at absolute zero. The prototype of this structure is β'-brass (β'CuZn; **Fig. 7***a*). The ordered structure is possible if the AB bond is energetically favored over the AA and BB bonds. As the temperature is raised, the thermal vibrations introduce disorder in the structure, the degree of disorder depending on the relative strengths of the various bonds. Above the order-disorder transition temperature, T_{od}, the A and B atoms occupy the lattice sites more or less at random. For example, in the β'-brass structure, above T_{od} the structure is the disordered β-brass structure (Fig. 7*b*). This is an entropy-driven transition where the system minimizes its free energy at a given temperature and pressure. If the order-disorder energy difference is much larger than kT_m, where T_m is the melting temperature and k is Boltzmann's constant, the system will remain ordered until it melts. If this energy difference is small compared to kT_m, the order-disorder transition to a disordered state takes place primarily through diffusion via vacancies. Diffraction techniques such as x-rays can be used to study this transition. *See* DIFFUSION; X-RAY DIFFRACTION.

Structural phase transition. The smallest entity of a perfect crystalline solid that repeats over and over again is known as its primitive unit cell. The structure of a unit cell at a given temperature and pressure is determined by the atomic interactions and the entropy through the minimization of the free energy. A crystal undergoes a commensurate structural phase transition at a certain temperature if its unit cell volume multiplies by an integral number as the temperature is lowered. In a commensurate phase transition, a crystal goes from one periodic structure to another at the transition temperature. The layered compound titanium diseleride ($TiSe_2$) undergoes such a phase transition at -100°F (200 K), with the volume of the unit cell in the low-temperature phase being eight times that in the high-temperature phase. This transition is thought to be driven by charge-density waves (discussed below). A second type of structural phase transition is an incommensurate phase transition, in which a superlattice of periodicity $\{l'\}$ is superimposed on the primary lattice of periodicity $\{l\}$ such that $\{l'\}$ is not contained in $\{l\}$. The superlattice manifests itself as secondary intensity spots around the primary spots in various scattering experiments (x-ray, neutron, electron, and so forth). The incommensurate phase lacks translational periodicity. Paraelectric potassium selenate (K_2SeO_4) undergoes an incommensurate phase transition at 127 K (-231°F) that changes to a ferroelectric phase at 97 K (-285°F). The ferroelectric phase is commensurate with the paraelectric phase, with the unit cell of the former being seven times that of the latter. *See* ELECTRON DIFFRACTION; FERROELECTRICS; NEUTRON DIFFRACTION.

Charge-density waves. These transitions have been observed in low-dimensional conductors, that is, materials that conduct electricity in only one or two dimensions. For example, if a material consists of linear chains of molecules such that the interaction between neighboring molecules on different chains is much smaller than that between neighboring molecules on the same chain, the electrons can move relatively easily along the chain, but motion of electrons from one chain to another is highly unlikely. In this case the material behaves as a quasi one-dimensional conductor. Some examples are $NbSe_3$ and $TaSe_3$ (**Fig. 8***a*). There are also materials whose molecules are arranged in sheets, with the interaction between neighboring molecules on different sheets being much smaller than that between neighboring molecules in the same sheet. In this case the electrons can move relatively easily within the sheet, but motion of electrons between sheets is highly restricted. Such materials behave as quasi-two-dimensional conductors. Some examples of two-dimensional conductors are $NbSe_2$ and $TaSe_2$ (Fig. 8*b*).

Below the transition temperature the electronic charge-density is modulated, the periodicity depending upon the topology of the Fermi surface of the electrons. The electronic charge density as a function of position $\vec{r}$ can be represented by Eq. (5), where

$$P(\vec{r}) = P_0(\vec{r})[1 + \alpha \cos(\vec{q}_0 \cdot \vec{r} + \phi)] \qquad (5)$$

$P_0(\vec{r})$ is the electronic charge density above the transition (that is, in the normal state), α is the amplitude of the charge-density wave (which is zero above the

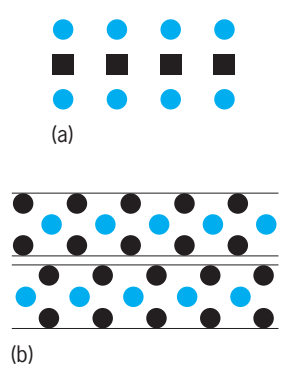

Fig. 8. Schematic representations of low-dimensional conductors. (*a*) One-dimensional conductor. Conducting molecules (circles) are separated by nonconducting molecules (squares). (*b*) Two-dimensional conductor. Conducting atoms (colored circles), for example, niobium, form a sheet. Nonconducting atoms (black circles), for example, sulfur, form sheets above and below the sheet of conducting atoms and thus help to isolate it from the next layer or sheet of conducting atoms.

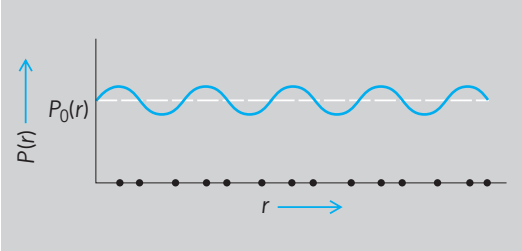

Fig. 9. Charge-density wave. Horizontal straight line shows electronic charge density in absence of wave. Curve shows sinusoidal variation of charge density below transition temperature. Dots at bottom show distorted positions of atoms in lattice.

transition and grows as the temperature is lowered below the transition temperature), $\vec{q}_0$ is the wave vector of the charge-density wave which determines the periodicity, and ϕ is the phase. This modulation of the negative electronic charge alone is energetically unfavorable because of the Coulomb interaction. But the modulation of the electronic charge density produces a distortion in the lattice of the ions, such that the attractive Coulomb interaction between electronic charge and ionic charge helps to stabilize the charge-density wave (**Fig. 9**). It is this accompanying lattice distortion that can be observed by x-ray, neutron, and electron diffraction techniques and thus provides evidence for the charge-density-wave transition.

The lattice distortion causes a gap in the electronic energy spectrum, which leads to a reduction in the number of electrons available for conduction and hence causes an increase in the resistivity of the material. The size of the gap grows from zero at the transition temperature to a maximum at the absolute zero of temperature. In some cases the gap is so large that the material becomes an insulator at the onset of the lattice distortion, leading to a metal-insulator transition. *See* BAND THEORY OF SOLIDS; CHARGE-DENSITY WAVE.

Spin-density waves. In a charge-density-wave state below the transition, the modulation in the electronic density has the same phase for electrons of both spins. If the modulation in the density of spin-up electrons is 180° out of phase with the modulation in density of spin-down electrons, then the charge density would remain unmodulated, but there would be a modulation of the spin density. Such a state is called a spin-density-wave state. The electronic density for spin-up (+) and spin-down (−) electrons can then be written in a fashion similar to that for the charge-density-wave case as Eq. (6) [**Fig. 10**]. Spin-density-

$$P_{\pm}(\vec{r}) = {}^1/_2 P_0(\vec{r})[1 \pm \alpha \cos(\vec{q}_0 \cdot \vec{r} + \phi)] \qquad (6)$$

wave transitions are much less common than charge-density waves, but are known to exist in metallic chromium below 312 K (102°F). *See* SPIN-DENSITY WAVE.

Metal-insulator transition. The atomic energy levels of the constituents of a solid are broadened into bands due to the overlap of their electron clouds. For semiconductors and insulators, the highest band that is completely full is the valence band. The next band higher in energy than the valence band is called the conduction band. The conduction band is empty in a nonmetal and partly filled in a metal. In response to an applied electric field, the electrons in the conduction band of a metal are accelerated by being excited to a neighboring empty state in the band. Therefore a metal is a good electrical conductor. An insulator with a reasonable band gap (5 eV) between the valence and the conduction bands cannot conduct electricity because the electrons from the valence band cannot be excited to the conduction band with the usual electric fields. In principle, the lattice parameter can be varied to change the band gap and hence induce a metal-insulator transition in a system. This is indeed possible for some systems. For example, iodine becomes metallic under a pressure of 160 kilobars (16 gigapascals). Liquid mercury becomes a semiconductor at high temperature such that its density is below 5 g/cm^3, down from 13.5 g/cm^3 at 68°F (20°C). Many, but not all, metal-insulator transitions can be explained by one-electron band-structure calculations. For the others, it is necessary to go beyond the one-electron model and consider many-body mechanisms. *See* ELECTRICAL CONDUCTIVITY OF METALS; ELECTRICAL RESISTIVITY.

Superfluid transition. Helium gas condenses at 4.2 K (−452°F) and is not known to solidify under atmospheric pressure down to the lowest temperature observed. A pressure of 25 atmospheres (2.5 megapascals) is required to solidify helium. This is due to weak interatomic interactions and the low mass of the helium atom. Thus liquid helium is a quantum liquid in which the energy of zero-point motion is comparable to that of interatomic interactions so that the helium atoms cannot be localized to form a solid. Helium-4 (^{4}He) undergoes a very interesting phase transition at 2.18 K (−455.75°F) from helium I to helium II such that a part of helium II behaves like a normal fluid and a part like superfluid. The superfluid component has zero entropy and zero viscosity, allowing it to flow without resistance through openings of 10^{-2} cm or even smaller. This property is due to the fact that ^{4}He has all of its spins paired, giving it a net spin of zero. Thus ^{4}He is a quantum liquid obeying Bose-Einstein statistics. In Bose-Einstein statistics, there is no restriction on the number of particles that can occupy a given energy level. Thus

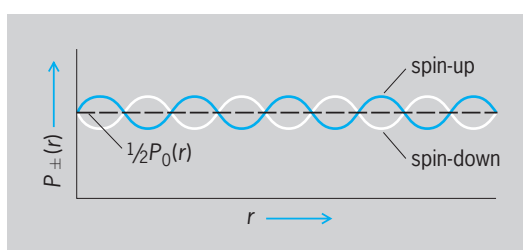

Fig. 10. Spin-density wave. The two curves show the variations of the densities of spin-up and spin-down electrons.

at 0 K (−459.67°F) all the particles are in the ground state and belong to the superfluid state with zero entropy and zero viscosity. As the temperature is raised, some of the particles occupy the excited states and have finite entropy and viscosity. The number of superfluid atoms N_0 is related to the total number of atoms N by Eq. (7), where T_0 is the transition temperature above which the fluid is normal. If ^{4}He were a noninteracting Bose-Einstein gas, T_0 would be 3.14 K (−454.02°F), which is in qualitative agreement with the experimental value of 2.14 K (−455.82°F).

$$\frac{N_0}{N} = 1 - \left(\frac{T}{T_0}\right)^{3/2} \tag{7}$$

The superfluid component of helium II is a very good thermal conductor. This can be used to find T_0 experimentally. Under normal pressure, helium boils at 4.2 K (−452°F) like a typical liquid with many gas bubbles due to poor heat conduction. As the boiling point is lowered by lowering the pressure, the bubbles disappear as soon as the boiling point is the same as the transition temperature. The heat is then conducted away before the gas bubbles can be formed.

The less abundant isotope ^{3}He has a spin of $^1/_2$, and hence ^{3}He liquid obeys Fermi-Dirac statistics. It becomes superfluid below 3 millikelvins by forming Cooper-type pairs of ^{3}He atoms that behave as bosons. This pairing is similar to that occurring in superconductivity, and will not be discussed here. *See* LIQUID HELIUM; SUPERFLUIDITY.

Liquid-crystal transition. An isotropic liquid made up of long organic molecules can undergo a phase change to an anisotropic liquid known as a liquid crystal. The liquid crystal has viscosity similar to that of the liquid but with a long-range order in the orientation of its molecules. Thus a liquid crystal is an intermediate phase between a liquid and a solid. Most of the known liquid crystals have large optical anisotropies so that molecular reorientations produce dramatic changes in optical birefringence, absorption, reflection, scattering, and color. The phase change is determined by a delicate balance of intermolecular forces and changes dramatically because of small outside stimuli such as weak electric fields. These properties are the basis of the display devices used in pocket calculators, wristwatches, clocks, and so forth. Since liquid-crystal displays modify the ambient light instead of emitting light, they use much less power than light-emitting displays. *See* LIQUID CRYSTALS.

D. J. Sellmyer; S. S. Jaswal

Bibliography. B. Bergersen and M. Plischke, *Equilibrium Statistical Physics*, 2d ed., 1994; G. Burns, *Solid State Physics*, 1985; D. Chowdhury, *Spin Glasses and Other Frustrated Systems*, 1986; P. A. Cox, *The Electronic Structure and Chemistry of Solids*, 1987; N. E. Cusack, *The Physics of Structurally Disordered Matter*, 1987; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; K. M. Moorjani and J. M. D. Coey, *Magnetic Glasses*, 1984; N. Mott, *Metal-Insulator Transitions*, 2d ed., 1990; J. M. Yeomans, *Statistical Mechanics of Phase Transitions*, 1992.

Phase velocity

The velocity of propagation of a pure sine wave of infinite extent. In one dimension, for example, the form of the disturbance for such a wave is $y(x,t) = A \sin [2\pi(x/\lambda - t/T)]$. Here x is the position at which the disturbance $y(x,t)$ exists at time t, λ is the wavelength, T is the period which is related to the wave frequency by $T = 1/f$, and A is the disturbance amplitude. The argument of the sine function is called the phase. The phase velocity is the speed with which a point of constant phase can be said to move. Thus $x/\lambda - ft =$ constant, so the phase velocity v_p is given by $dx/dt = v_p = \lambda f$. This is the basic relationship connecting phase velocity, wavelength, and frequency. *See* PHASE (PERIODIC PHENOMENA); SINE WAVE; WAVE MOTION.

For a simple sine wave on a string, one can perceive this phase velocity (provided the wave does not move too quickly) by merely focusing attention on any particular wave crest and observing its apparent motion along the string. A fairly slack, heavy cord has a slow, easily observable phase velocity.

The phase velocity for waves in a medium is determined in part by intrinsic properties of the medium. For all mechanical waves in elastic media, the square of the phase velocity is proportional to the ratio of the appropriate elastic property of the medium to the appropriate inertia property. For example, the square of v_p for transverse waves of small amplitude in a stretched string is $v_p^2 = T/\mu$, where T is the tension in the string and μ is the mass per unit length. The tension is proportional to Young's modulus of elasticity Y for the material from which the string is made, so $v_p^2 \propto Y/\mu$. The phase velocity of electromagnetic waves depends upon the medium as well. In vacuum or (usually to good enough approximation) air, the phase velocity c is given by $c^2 = 1/\epsilon_0\mu_0 \approx 9 \times 10^{16}$ m^2/s^2, where ϵ_0 and μ_0 are respectively the permittivity and permeability of the vacuum. *See* ELECTROMAGNETIC RADIATION; YOUNG'S MODULUS.

Phase velocity may also depend upon the mode of wave propagation. For example, transverse waves in a bar travel with a phase velocity which is different from that for longitudinal waves of the same frequency traveling in the same bar.

Phase velocity will also depend, in general, upon the frequency of the wave. Waves of different frequencies will travel at different speeds, resulting in a phenomenon called dispersion. A beautiful example of this is the dispersion of white light into the colors of the visible spectrum by a prism. There the phase velocity is given by c/n, where n is the index of refraction of the glass from which the prism is made, which depends quite markedly upon the frequency (hence wavelength, hence color) of the light. The equation for determining the disturbance at any place and time, given suitable starting values, is called the wave equation. If this wave equation involves only second-order rates of change of the disturbance with respect to both space coordinates and time, then the phase velocity is frequency-independent. Otherwise, the phase velocity will be frequency-dependent.

See GROUP VELOCITY; LIGHT; REFRACTION OF WAVES; WAVE EQUATION. S. A. Williams

Bibliography. D. Giancoli, *Physics for Scientists and Engineers*, 3d ed., 2000; D. Halliday, *Fundamentals of Physics*, 6th ed., 2000; I. G. Main, *Vibrations and Waves in Physics*, 3d ed., 1993.

Phasmatodea

The order of insects comprising the leaf insects and the stick insects, also known as walking sticks. There are about 2800 species of phasmids (Phasmatodea is alternatively known as Phasmida) cataloged into six families, the majority of which live in the tropics. All are herbivores, and most have a strong resemblance to twigs and sticks (see **illustration**) with long slender bodies and legs. Phasmatodea species have cryptic coloration: either brown like a stick or green like a leaf. Many species have knobs, spines, and lobes to camouflage themselves, thus avoiding predation from visual predators such as birds and spiders. Stick insects also behave to make themselves appear more sticklike; often they hold their front pair of legs out straight to increase their length and rock side-to-side as if they were swaying in the wind. Members of the family Timemidae (=Phyllidae) are leaf mimics; their bodies and legs are flattened and marked to resemble leaves. *See* INSECTA; PROTECTIVE COLORATION.

Characteristics. Phasmids range in size from tiny *Timema* males from North America that are only 1.2 cm (0.5 in.) long to giants of Malaysia in the genus *Phobeticus* that can be 0.5 m (1.65 ft) long, including body and legs. Some species are winged and some wingless. In some instances the wings of the male are fully developed while those of the fatter female are reduced or absent. In general, male phasmids are smaller and thinner compared to females.

Reproduction and development. In the majority of species, females drop eggs from where they are feeding high among the plant leaves. The eggs with hard cases fall to the ground like the seeds they resemble, and remain hidden among the leaf litter. Phasmids hatch out as perfect miniatures of their parents. Phasmid eggs vary a great deal in size and shape, and egg capsule morphology is important in the classification of this group. A number of species have evolved a strategy to increase egg survival by employing other insects to carry off and hide their eggs. The stick insect provides a meal for the ant in the form of the egg shell capitulum (a knoblike proturberance), which is removed and eaten by the ant. The stick insect egg is left by the ants in their colony rubbish dump, where it is safe from predators and fires. It is not known how many phasmid species have formed a mutualism of this type with ants, but a similar strategy is used by some plants to have their seeds dispersed, stored safely, and provided with a suitable germination site. Other stick insect species are known to bury their eggs, insert them under the bark of trees, or attach them to branches with a gluelike substance. *See* MUTUALISM.

Although sexual reproduction is the standard method within Phasmatodea, parthenogenesis (a special type of reproduction in which an egg develops without entrance of a sperm) is also common. About 10% of species can produce viable eggs without fertilization, and this has made them a popular and convenient laboratory animal and pet. Studies of phasmid asexual reproduction have shown that different genetic mechanisms are employed by different species [automictic (where offspring result from meiosis and chromosome doubling so are homozygous), apomictic (where offspring are formed with meiosis, and are identical to mother), and hemiclonal (where half the genome passes intact to the next generation while the other half is replaced)] to produce eggs without sex. Some entirely parthenogenetic species have evolved via interspecies hybridization when two bisexual species have mated. Mary Morgan-Richards

Bibliography. P. D. Brock, *The Amazing World of Stick and Leaf Insects*, Amateur Entomologists' Society, 1999; D. Otte and P. D. Brock, *Phasmida Species File: A Catalog of the Stick and Leaf Insects of the World*, Orthopterists' Society, 2003; S. Trewick and M. Morgan-Richards, *New Zealand Wild: Stick Insects*, Reed Publishing, 2005.

Walking stick insect. (*Photo from National Park Service*)

Phenacetin

One of a general class of medicinals known variously as analgesics, antifebrins, or antipyretics, of which acetanilide is the best known. Phenacetin is an acetyl derivative of *p*-phenetidine. It is made by the reaction of

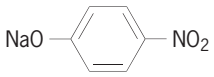

with C_2H_5Cl to form

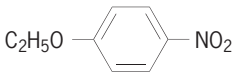

which is reduced to the corresponding amine and acetylated to form phenacetin,

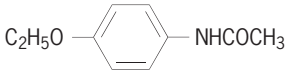

The reaction is continued as a cyclic process in which phenol, acetic acid, and ethyl chloride are continuously supplied. Phenacetin is less toxic than acetanilide, but it lowers the ability of blood to combine with oxygen. *See* ASPIRIN; HYPOTHERMIA.

Allen L. Hanson

Phenocryst

A relatively large crystal embedded in a finer-grained or glassy igneous rock. The presence of phenocrysts gives the rock a porphyritic texture (see **illus.**).

Granite (quartz monzonite) from the Sierra Nevada of California showing numerous phenocrysts of microcline feldspar in parallel orientation with banded structure of the rock. Hammer is 10 in. (25 cm) long. (*USGS photograph by W. B. Hamilton*)

Phenocrysts are represented most commonly by feldspar, quartz, biotite, hornblende, pyroxene, and olivine. Strictly speaking, phenocrysts crystallize from molten rock material (lava or magma). They commonly represent an earlier and slower stage of crystallization than does the matrix in which they are embedded. Phenocrysts are to be distinguished from certain relatively large crystals (porphyroblasts) which develop late in solid rock as the result of metamorphism or metasomatism. If the origin of a large crystal is in question, the nongenetic term megacryst should be used. *See* CONTACT AUREOLE; IGNEOUS ROCKS; PORPHYROBLAST; PORPHYRY.

Carleton A. Chapman

Phenol

The simplest member of a class of organic compounds possessing a hydroxyl group attached to a benzene ring or to a more complex aromatic ring system. Phenol itself, C_6H_5OH, may also be called hydroxybenzene or carbolic acid. Pure phenol is a colorless solid melting at 42°C (108°F), moderately soluble in water, and weakly acidic (p*K* 9.9).

Phenol has broad biocidal properties, and dilute aqueous solutions have long been used as an antiseptic. At higher concentrations it causes severe skin burns; it is a violent systemic poison. *See* ANTISEPTIC.

Structure and nomenclature. Phenol has the structure shown . Simple substituted phenols, such as the three isomeric chlorophenols, are named as indicated, using the ortho, meta, and para prefixes. In more highly substituted phenols the positions of substitution are indicated by numbers, for example, 2,4-dichlorophenol. Compounds with more than one

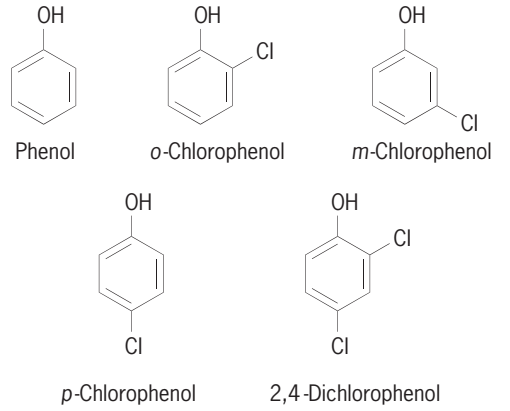

hydroxyl group per aromatic ring are known as polyhydric phenols, and include catechol, resorcinol, hydroquinone, phloroglucinol, and pyrogallol.

Reactions. The chemistry of phenol can be discussed under two headings.

Reactions of hydroxyl group. Phenols react readily with alkali hydroxides to form salts, reaction (1),

$$C_6H_5OH + NaOH \rightarrow \underset{\text{Sodium phenoxide}}{C_6H_5ONa} + H_2O \qquad (1)$$

and those phenols containing electronegative substituents (for example, trichlorophenol) are even acidic enough to react with bicarbonate.

Like their aliphatic analogs, the alcohols, phenols form esters and ethers. Treatment of phenol with acetic anhydride yields phenyl acetate. The methyl ether of phenol (anisole) can be prepared by treatment of sodium phenoxide with methyl sulfate. Etherification of 2,4-dichlorophenol with sodium chloroacetate, reaction (2), yields

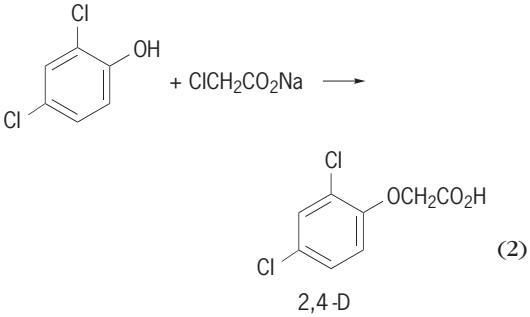

(2)

2,4-dichlorophenoxyacetic acid (2,4-D), which is a powerful herbicide.

Reactions of aromatic ring. Under the conditions for electrophilic substitution reactions (for example, nitration, halogenation, sulfonation, alkylation, and diazo coupling), phenol is one of the most reactive aromatic compounds. Such reactions invariably lead to introduction of the substituent into an ortho or

para position, as illustrated by the bromination of phenol, reactions (3). At low temperature in carbon

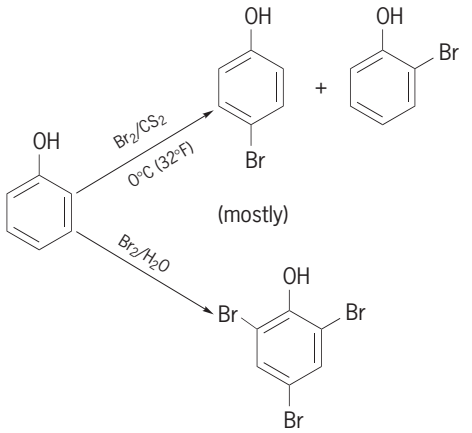

(3*a*)

(3*b*)

disulfide solution, careful addition of one equivalent of bromine leads to the monobromo products of which the para isomer predominates, reaction (3*a*). Treatment of phenol with excess bromine in water yields 2,4,6-tribromophenol, in which both of the ortho and the para hydrogens have been replaced by bromine, reaction (3*b*). Reaction (3*b*) has been used as a fairly sensitive test for phenol, based upon the rapid precipitation of the insoluble tribromo product.

Two further examples of electrophilic substitution are the reaction with acetone, reaction (4), to form

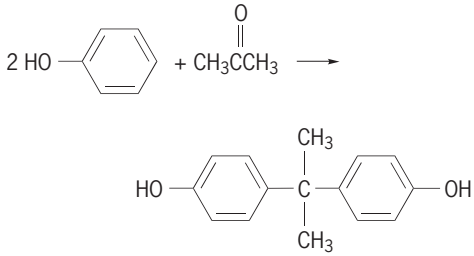

(4)

Bis-phenol A

bis-phenol A, an important plastics intermediate, and the carboxylation of sodium phenoxide, reaction (5), to yield salicylic acid and thence aspirin.

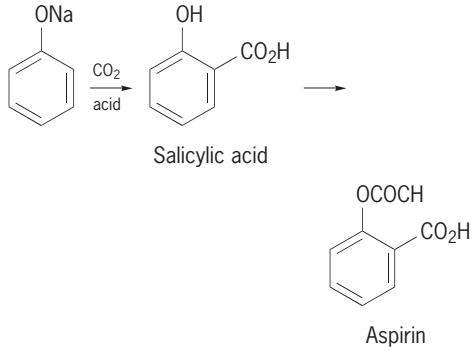

(5)

Phenol is particularly sensitive to oxidation. Quinone is formed by chromic acid oxidation of phenol, but most oxidizing agents convert phenol to a complex mixture of products derived from coupling of intermediate phenoxy radicals (oxidative coupling products). Certain 2,4,6-trisubstituted phenols are oxidized to relatively stable free radicals, and this property is used to advantage in the design of phenolic antioxidants. Reduction of phenol to cyclohexanol can be effected by hydrogen in the presence of catalysts such as nickel.

Production. Until World War I phenol was essentially a natural coal tar product. However, synthetic methods replaced extraction from natural sources. There are many possible syntheses of phenol; six commercially significant ones are listed below.

1. The oldest process, that of sulfonation of benzene followed by neutralization of the sulfonic acid thus produced and fusion with caustic soda, was introduced into the United States about 1915. This is shown by reactions (6) and (7).

$$C_6H_6 + H_2SO_4 \rightarrow C_6H_5SO_2OH \xrightarrow{NaOH} C_6H_5ONa \quad (6)$$

$$C_6H_5ONa \xrightarrow{acid} C_6H_5OH \quad (7)$$

2. The cumene hydroperoxide process yields the largest production in the United States, reactions (8) and (9). This process grew very rapidly after its intro-

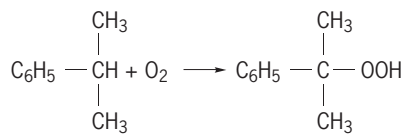

(8)

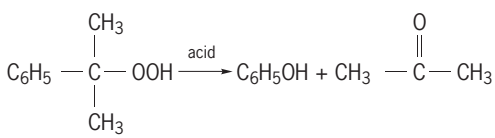

(9)

duction in 1955. The coproduct acetone is of considerable value, which undoubtedly contributed to the widespread acceptance of the process.

3. The Raschig process was introduced in 1940 and involves a first-stage chlorination of benzene using an air–hydrochloric acid mixture, as shown by reactions (10) and (11). The by-product HCl is re-

$$C_6H_6 + HCl + {}^1/_2O_2 \rightarrow C_6H_5Cl + H_2O \quad (10)$$

$$C_6H_5Cl + H_2O \rightarrow C_6H_5OH + HCl \quad (11)$$

cycled to the first stage. High temperatures are employed, and high alloy reaction vessels are used to minimize corrosion.

4. The chlorination, or Dow, process was introduced in 1924 and is reasonably straightforward. Starting with chlorobenzene, the process proceeds as in reaction (12). Copper catalysts and high temperatures and pressures are employed.

$$C_6H_5Cl + NaOH \rightarrow$$

$$NaCl + C_6H_5ONa \xrightarrow{acid} C_6H_5OH \quad (12)$$

5. A process developed in 1962 starts with toluene and proceeds through oxidation to benzoic

acid, as seen in reaction (13). This is followed by fur-

$$C_6H_5 \longrightarrow CH_3 + {}^{3}/_{2}O_2 \rightarrow C_6H_5COOH + H_2O \qquad (13)$$

ther oxidation or decarboxylation in a second step using cupric benzoate as a catalyst, as shown by reaction (14).

$$C_6H_5COOH + {}^{1}/_{2}O_2 \rightarrow C_6H_5OH + CO_2 \qquad (14)$$

6. The last process starts with cyclohexane, which is oxidized to a mixture of cyclohexanol and cyclohexanone (mixed oil). The next step is hydrogenation to convert all of the mixture to cyclohexanol. This intermediate is then, by classical methods, ring-dehydrogenated to yield phenol and hydrogen, which is recycled, as shown in reactions (15) and (16). Boron compounds are employed as catalysts or mediators.

$$C_6H_{12} + O_2 \rightarrow \text{(mixed oil)} \xrightarrow{\text{hydro-genation}} C_6H_{11}OH \qquad (15)$$

$$C_6H_{11}OH \xrightarrow{\text{catalyst}} C_6H_5OH + 3H_2 \qquad (16)$$

Uses and derivatives. Phenol is one of the most versatile and important industrial organic chemicals. It is the starting point for many diverse products used in the home and industry. A partial list includes nylon, epoxy resins, surface active agents, synthetic detergents, plasticizers, antioxidants, lube oil additives, phenolic resins (with formaldehyde, furfural, and so on), cyclohexanol, adipic acid, polyurethanes, aspirin, dyes, wood preservatives, herbicides, drugs, fungicides, gasoline additives, inhibitors, explosives, and pesticides.

Naturally occurring phenols. Phenol and the methyl phenols (cresols and xylenols) are obtained from coal tar. Phenol has been found as a minor constituent of various plant materials, including tobacco leaves and pine needles. In the structures shown below, carvacrol, thymol, and vanillin are examples of natural phenols of plant origin. Phenolic rings are also present in such biologically important compounds as the amino acid tyrosine and in estrone, a female sex hormone.

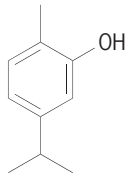

Carvacrol

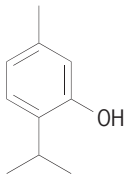

Thymol

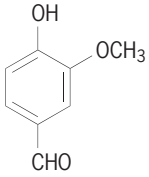

Vanillin

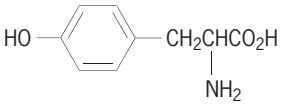

Tyrosine

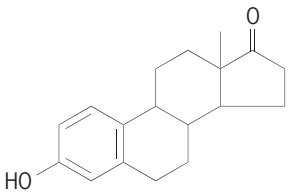

Estrone

Robert I. Stirton; Martin Stiles

Bibliography. D. H. R. Barton and D. Ollis, *Comprehensive Organic Chemistry*, vol. 4, 1979; S. N. Ege, *Organic Chemistry*, 4th ed., 1998; R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998.

Phenolic resin

One of the condensation products of phenols or phenolic derivatives with aldehydes such as formaldehyde and furfural. The term phenoplasts is sometimes used to refer to the whole group of products. The phenol-formaldehyde resins, developed commercially between 1905 and 1910, were the first truly synthetic polymers and have found wide usage for electrical insulation, molded objects, shell molds for metals, laminates, adhesives, and many other applications. They are characterized by low cost, dimensional stability, high strength, and resistance to aging. The combination of low cost and good properties is reflected in the fact that phenolic resins are produced in greater volume than any other thermosetting resin.

Phenol is prepared by the hydrolysis of chlorobenzene, by the alkali fusion of sodium benzene sulfonate, by the oxidation of toluene, by the dehydrogenation of cyclohexanol/cyclohexane, or by the decomposition of cumene hydroperoxide, as shown in reaction (1).

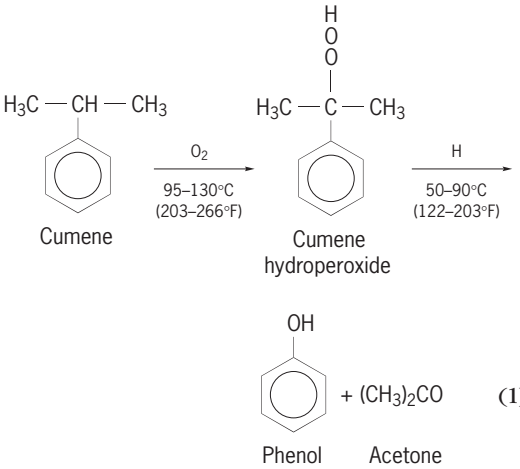

(1)

Resorcinol, obtained by the alkaline fusion of *m*-benzene disulfonic acid, and resorcinol from coal tar are also used.

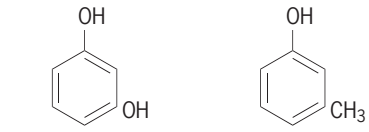

m-Dihydroxybenzene or resorcinol

m-Cresol

Formaldehyde is produced by the oxidation of

methane or methyl alcohol, as in reaction (2), and

$$\underset{\text{Methyl alcohol}}{CH_3OH} + O_2 \xrightarrow[\substack{550\text{–}600^\circ C \\ (1020\text{–}1110^\circ F)}]{\text{Ag, Fe, or Mo catalyst}} \underset{\text{Formaldehyde}}{HCHO} + H_2O \qquad (2)$$

furfural is obtained by the hydrolysis of oat hulls.

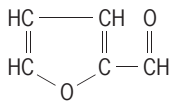

Furfural

Polymerization. In the presence of an acid or base, phenol and aqueous formaldehyde react to form a solution of phenolic alcohols or methylol derivatives with the methylol groups in the ortho and para positions, reaction (3). This reaction takes place quickly

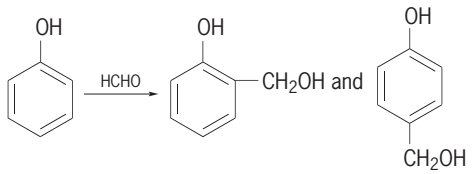

(3)

in a basic medium and slowly in an acidic medium.

The methylol phenols formed initially in a basic medium with formaldehyde in excess condense with each other and with additional formaldehyde to yield an "A-stage" resin or "resole," a brittle resin which is soluble and fusible. The resole resin consists of a mixture of isomers containing free methylol groups, which are available for subsequent cross-linking reactions to form a less-soluble "B-stage" resin. Many structural variations are possible. The structure for a typical resole component is

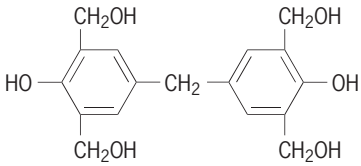

In the presence of acid and less than 0.86 mole of formaldehyde per mole of phenol, the primary alcohols react to yield diphenylmethane polymers called novolacs, which are soluble and fusible and contain an average of 5 or 6 phenol units per molecule. These resins may also be referred to as A-stage resins. Novolacs may also be reacted with epichlorohydrin to yield epoxy polymers. The structure for a typical novolac resin is

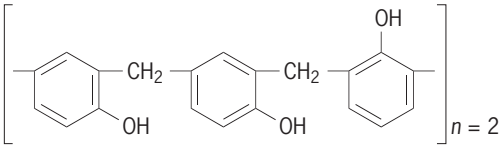

Hardening of all of these is effected by further cross-linking. A resole-type resin is inherently capable of cross-linking itself on heating and is sometimes referred to as a one-stage resin. On the other hand, a novolac has no free methylol groups and must be mixed with an aldehyde to undergo further reaction; hence a novolac is sometimes called a two-stage resin.

In the production of phenolic-resin molding compositions, it is common practice to neutralize, concentrate, and dry the B-stage resin; to mix it with fillers and in, in the case of a novolac, a curing agent or hardener; and finally to compact it into the form of pellets or briquets. Other ingredients may be present also, such as curing accelerators, pigments, lubricants, and plasticizers. The curing agent for a novolac is hexamethylenetetramine, which at the temperature of molding reacts with water to form formaldehyde and ammonia. In the presence of ammonia and the additional formaldehyde, the B-stage resin cures in the mold to yield a highly cross-linked, insoluble, and infusible C-stage product. The structure for a C-stage phenolformaldehyde resin is

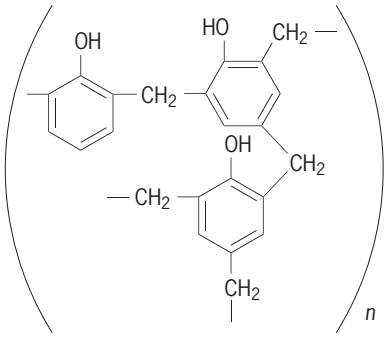

By use of *m*-phenol derivatives, such as resorcinol or *m*-cresol, resins are obtained which cure rapidly at low temperature because the meta substituents activate the ortho and para positions. An ortho or para alkyl phenol which has only two active sites available for reaction (difunctional) can be used, such as *p*-tert-butylphenol

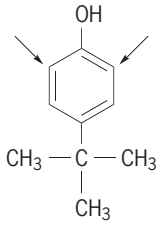

Then oil-soluble, thermoplastic resins are formed instead of the cross-linked materials obtained from the trifunctional phenols just discussed. These products, somewhat more expensive than ordinary phenolic resins, are used in special paint, varnish, and adhesive formulations.

With the use of furfural instead of formaldehyde, the B-stage resin has the unique property of remaining thermosplastic for a relatively long time. The phenol-furfural compositions are useful for molding large complex forms in which extra time is needed for the resin to fill the mold completely.

Fabrication and use. Phenolic resins can be cast from syrupy intermediates or molded from B-stage solid resins. Laminated products can be produced by impregnating fiber, cloth, wood, and other materials with the resin. After heating, laminated sheets can be pressed into any shapes desired. Most of the phenolic plastics can be machined if necessary.

Another important type of phenolic resin product is rigid foam. One type is prepared by incorporating a blowing agent in a curing mixture so that the heat of reaction decomposes the blowing agent. A second type, "syntactic" foam, consists of microscopic hollow spheres of phenolic resin mixed with a curable binder such as an epoxy resin or polyester. Foams of this type have very high strength after curing.

Cured phenolic plastics are rigid, hard, and resistant to chemicals (except strong alkali) and to heat.

Some of the uses for phenolic resins are for making precisely molded articles, such as telephone parts, for manufacturing strong and durable laminated boards, or for impregnating fabrics, wood, or paper. Phenolic resins are also widely used as adhesives, as the binder for grinding wheels, as thermal insulation panels, as ion-exchange resins, and in paints and varnishes. *See* ADHESIVE; ION EXCHANGE; PHENOL; PLASTICS PROCESSING; POLYMERIZATION; TEXTILE CHEMISTRY.

John A. Manson

Bibliography. J. A. Brydson, *Plastic Materials*, 7th ed., 1999; F. A. Carey and R. J. Sundberg, *Advanced Organic Chemistry*, pt. A: *Structures and Mechanisms*, 4th ed., 2000; A. Gardziella, L. A. Pilato, and A. Knop, *Phenolic Resins: Chemistry, Applications, Standardization, Safety and Ecology*, 2d ed., 1999; C. A. Harper and E. M. Petrie, *Plastics Materials and Processes: A Concise Encyclopedia*, 2003.

Phenylketonuria

An inborn error of metabolism in which affected individuals lack the liver enzyme phenylalanine hydroxylase (PAH), which is needed to metabolize phenylalanine, an amino acid essential for normal growth and development. If untreated, affected individuals may become severely mentally retarded, become microcephalic, have behavioral problems, develop epilepsy, or show other signs of neurological impairment. Phenylketonuria (PKU) is inherited as an autosomal recessive trait and is found in all ethnic groups but most frequently in individuals of northern European descent. Its incidence is about 1 per 14,000 births in the United States. Classically, persons with phenylketonuria exhibit blood phenylalanine concentrations of 20 mg/deciliter or more (normal concentrations are about 1–2 mg/dl), normal blood tyrosine levels, and excessive phenylalanine metabolites in the urine while on a normal diet. Phenylketonuria variants have blood phenylalanine concentrations of 10–20 mg/dl, but may not have phenylalanine metabolites in their urine unless they have ingested excessive amounts of protein.

Benign hyperphenylalaninemia is the term applied when blood phenylalanine concentrations are 4–10 mg/dl. It causes no clinical symptoms, does not affect intelligence, and requires no treatment. About 1–3% of individuals with phenylketonuria have an atypical form caused by a defect in the synthesis of the enzyme dihydropteridine reductase or the coenzyme tetrahydrobiopterin. This coenzyme is a cofactor that is needed for the conversion of phenylalanine to tyrosine. Infants with deficiencies of tetrahydrobiopterin have high blood levels of phenylalanine and exhibit progressive neurological deterioration when treated with only a low-phenylalanine diet.

The phenylalanine hydroxylase locus has been mapped on the distal tip of the long arm of chromosome 12. The gene has been cloned, and, by using molecular genetic techniques, prenatal identification of phenylketonuria homozygotes and carriers has been achieved in about 90% of affected families. Since phenylalanine hydroxylase is not expressed in amniotic cells, amniocentesis, a common screening procedure, is not useful for the prenatal diagnosis of phenylketonuria. *See* PRENATAL DIAGNOSIS.

Newborn screening programs for phenylketonuria have been successful in identifying most cases within a few weeks of birth. The most widely used screening method is the Guthrie test, which measures the blood phenylalanine level by bacterial inhibition assay using a dried blood spot on filter paper obtained by heel puncture as close as possible to the time of discharge from the hospital nursery. A low phenylalanine diet with restriction of proteins, if initiated early in infancy, can prevent the development of severe intellectual and neurological handicaps that would otherwise occur in virtually all untreated cases. *See* MENTAL RETARDATION.

The success of nationwide screening programs for phenylketonuria created an unexpected problem as women with the disorder reached childbearing age. Retrospective surveys have revealed a high rate of mental retardation, microcephaly, congenital heart defects, and intrauterine growth retardation among offspring of women with untreated maternal phenylketonuria. Therefore, in order to reduce the morbidity associated with the offspring of women with maternal phenylketonuria, a phenylalanine-restricted diet should be followed throughout pregnancy and should probably be initiated prior to pregnancy in order to have maximum effectiveness. *See* HUMAN GENETICS; PROTEIN METABOLISM.

Felix de la Cruz

Bibliography. G. M. Addison et al. (eds.), *Studies in Inherited Diseases: Prenatal and Perinatal Diagnosis*, 1989; F. G. Cunningham, P. C. MacDonald, and N. F. Gant, *Williams Obstetrics*, 21st ed., 2001; N. A. Holtzman et al., Effect of age at loss of dietary control on intellectual performance and behavior of children with phenylketonuria, *N. Engl. J. Med.*, 314:593–598, 1986; R. R. Lenke and H. L. Levy, Maternal phenylketonuria and hyperphenylalaninemia: An international survey of the outcome of untreated and treated pregnancies, *N. Engl. J. Med.*, 303:1202–1208, 1980; A. A. Lidsky, F. Guttler, and S. L. C. Woo, Prenatal diagnosis of classical phenylketonuria by DNA analysis,

Lancet, 1:549–551, 1985; A. C. Whittle, *Prenatal Diagnosis in Obstetric Practice*, 2d ed., 1995.

Pheromone

A substance that acts as a molecular messenger, transmitting information from one member of a species to another member of the same species. The first pheromone to be characterized chemically was bombykol, an unsaturated, straight-chain alcohol that is secreted in microgram amounts by females of the silkworm moth (*Bombyx mori*) and is capable of attracting male silkworm moths at large distances. The electroantennogram (EAG) technique was developed in connection with this work; the electrical signals that are generated by the pheromone stimulation of an insect's chemoreceptor cells are recorded and analyzed.

A useful distinction has been made between releaser pheromones, which elicit a rapid, behavioral response, and primer pheromones, which elicit a slower, developmental response and may pave the way for a future behavior.

Fungi. Communication via pheromones is common throughout nature, including some eukaryotic microorganisms that exchange vital chemical signals. For example, female gametes of *Allomyces*, an aquatic fungus that reproduces sexually in one part of its life cycle, secrete the sesquiterpenoid alcohol (–)-sirenin, which serves to attract male gametes. Another example is the water mold *Achyla*; the female secretes a steroidal pheromone, antheridiol, which induces the development of male structures. The resultant antheridial branches then secrete a second steroid, dehydro-oogoniol, which induces further development of the female reproductive structures. The later stages in this reproductive process remain to be elucidated chemically. *See* FUNGI.

Cellular slime molds. The cellular slime molds (*Acrasiomycetes*) are soil microorganisms with a particularly interesting life cycle. A mobile, unicellular ameboid form, when the local food supply is exhausted, gathers into large aggregations of up to 10^5 amebas. These unite to form a sorocarp made up of a long, slender stalk that supports a spore-containing fruiting body. A pheromone, originally called acrasin, is responsible for the aggregation. It is cyclic adenosine monophosphate (cAMP), the same compound that serves as a second messenger in many hormonal systems. However, different species of slime mold may utilize different aggregation pheromones.

Algae. In several species of algae, relatively simple hydrocarbons act as sperm attractants. The first natural algal pheromones to be characterized were 1,3,5-octatriene and ectocarpine, from the aquatic species *Fucus serratus* and *Ectocarpus siliculosus*, respectively; these are extremely hydrophobic messenger molecules.

Insects. By far the largest number of characterized pheromones come from insect species. In part this is a consequence of the large number of insect species on Earth (well over a million), and in part it reflects the great interest in this remarkably successful group of animals. In social insects, such as termites and ants, there may be as many as a dozen different types of messages that are used to coordinate the complex activities which must be carried out to maintain a healthy colony. These activities might require specialized pheromones such as trail pheromones (to lead to a food source), alarm pheromones (recruiting soldiers to the site of an enemy attack), or pheromones connected with reproductive behavior. In the honeybee, isoamyl acetate serves as a natural releaser of aggressive behavior. A large number of pheromones have been characterized from important pest species; there has been some elucidation of the biosynthetic pathways whereby these pheromones are produced from fatty acid precursors. The gypsy moth sex attractant, (+)-disparlure, has been characterized; and one of the most complex sex pheromones is that of the American cockroach (*Periplaneta americana*), a pheromone that has been synthesized.

Many female Lepidoptera make use of sex attractants that are most often straight-chain molecules with relatively simple functionalities. An interesting group of nitrogen-containing aphrodisiac pheromones is produced by certain male Lepidoptera. Unlike the females, males produce these compounds in relatively large quantities (around 1–100 micrograms per insect). They are usually applied directly to the female antennae by special organs, present only in the male, during courtship. These pheromones are derived from pyrrolizidine alkaloids, which the insects obtain from plants. Since the alkaloid itself provides protection against predators, such as birds and spiders, the pheromones derived from them may serve to inform a female about the defensive status of a potential mate, thereby serving as a chemical criterion for sexual selection. *See* INSECT PHYSIOLOGY; SOCIAL INSECTS.

Mammals. Much less is known about mammalian pheromones because mammalian behavior is more difficult to study. There are, however, a small number of well-characterized mammalian pheromones from pigs, dogs, hamsters, mice, and marmosets. There is a steroid that is present in the saliva of male pigs which, when sprayed on the face of a sow, induces a rigid stance in the female that facilitates mating. (It has been shown that pigs are effective truffle hunters because truffles excrete this same steroid.) In spite of a great deal of interest in the characterization of possible human pheromones, there have been no definitive chemical and behavioral results.

Research techniques. Pheromone isolation has become much easier since the 1960s as a result of rapidly improving techniques, especially gas chromatography and high-performance liquid chromatography, for the separation of complex mixtures of natural products on a micro scale. The characterization of a pheromone also requires a reliable bioassay. In many cases, several compounds must be present together, and in the correct proportions, in order to convey a message. Once compounds of

established biological activity are isolated, their structures may be determined by organic-chemical and spectroscopic techniques. Syntheses, sometimes stereospecific, can provide material for additional biological studies, as well as for confirmation of the proposed structures.

Applications. There is great potential for controlling the behavior of a given species by manipulating its natural chemical signals. For example, pheromones have been used to disrupt the reproduction of certain insect pests. One simple application involves the use of pheromones in baiting traps that can indicate when a particular insect population is growing. This approach can lead to reduced use of pesticides as well as advances in the control of both agricultural pests and disease vectors. *See* CHEMICAL ECOLOGY; CHEMORECEPTION; INSECT CONTROL, BIOLOGICAL.

J. Meinwald

Bibliography. W. C. Agosta, *Chemical Communication: The Language of Pheromones*, 1992; E. S. Albone, *Mammalian Semiochemistry: The Investigation of Chemical Signals Between Mammals*, 1984; B. A. Leonhardt and M. Beroza, *Insect Pheromone Technology: Chemistry and Applications*, 1982; D. W. MacDonald, D. Muller-Schwartz, and S. Natynczok (eds.), *Chemical Signals in Vertebrates*, 1991; G. D. Prestwich and G. J. Blomquist (eds.), *Pheromone Biochemistry*, 1987.

Philosophy of science

The subfield of philosophy that treats fundamental questions pertaining to science. What is science, as contrasted with technological innovation? What are its objectives and methods? Are there aims or goals that are proprietary to science—aims and goals that belong in the first instance to the scientific enterprise, even if they do not belong to any particular individual engaging in the enterprise? Would the methods of the science with which we are familiar be equally successful if the world were quite different? Does science make genuine progress, and if so how? What kind of a world does science as we know it testify to? Is it a world of order or one of chaos? If it is a world of order, how can we understand that order? Should we understand it in terms of laws, akin to those we devise for the sake of regulating our various community interactions and undertakings? And if not in terms of laws like these, then what kinds of laws should we think of science as revealing? Science, in our present era, is divided into specialties. How do specialties originate and evolve? How are the subject matters of diverse scientific specialties related? These questions are as urgent today as they were in the days of Aristotle, who originated the first theory of science, as well as engaged in some of the very earliest scientific inquiries and defined the very first disciplinary boundaries. *See* PHYSICAL SCIENCE; SCIENCE; SCIENTIFIC METHODS; TECHNOLOGY.

Aristotle's theory of science. Aristotle's theory of knowledge and his theory of science were the same. He held that intellectual inquiry—inquiry for its own sake—falls into natural categories that today we might refer to as disciplines, and that these disciplines are as different from each other as they are from common sense. Aristotle insisted that standards of inquiry differ by discipline. For example, the standards of proof and evidence in logic and mathematics should not be applied to the study of nature, or to politics. Aristotle said that common sense is not knowledge because common sense has no standards of inquiry. It is not a discipline. It is a practical activity, as contrasted with scientific or disciplinary life. Knowledge, as contrasted with common sense, is relatively rare because it is the product (as the target or goal) of scientific inquiry—a relative newcomer on the stage of natural history. Science is common only conditionally—common only to the same extent that hard-won fruits of scientific inquiry are transmitted or disseminated without barriers. If science were never invented, there could be no knowledge, however piecemeal—just abject common sense.

Aristotle's theory of science deserves special attention, if only for its contrariety with a certain current in our present culture. Science, as we know it, has come under attack in numerous quarters, some of them academic. The grounds for criticism most heavily subscribed to is the charge that science is not—and could never be—either as objective or as disinterested as it purports to be. The charge is that science—or, more precisely, scientists—do not aim at truth. Instead, science is just one more social institution entangled in the usual power struggles that characterize any scheme of social arrangements. It is both a player in the familiar political drama and a convenient host or node for other such players, and so it is not deserving of the authority it routinely enjoys. This contemporary conception of scientific endeavors contrasts with Aristotle's conception in that it does not acknowledge science as an enterprise with a life and reach of its own—a pull and momentum largely independent of the forces exerted by the scientists who practice it. Thus this more contemporary perspective does not acknowledge (whereas by contrast Aristotle did) that science, as an enterprise, comprises goals and standards far broader and richer than the lives and reaches, goals, and standards of the scientists who enter into it. *See* LOGIC.

Pragmatism. The dominant conception of science among academics today—the conception we just contrasted with Aristotle's—draws heavily on the philosophical doctrine of pragmatism. Classical philosophical pragmatism was an American phenomenon. It originated in the scholarship of Charles Sanders Peirce, William James, and John Dewey, and lives on today in the intellectual lineages of Donald Davidson and Daniel Dennett. Many of the themes originated and advanced today by these Anglo-American philosophers now run like golden threads through the doctrines of numerous thinkers of the European continent—"deconstructionists" or "constructivists" as they are known. (Widely read, among those writing in that tradition today, are Jacques Derrida, Michel Foucault, and later writings of Richard Rorty.) Fundamental to classical pragmatism

is its commitment, ironically, to Darwinism in the form of a doctrine that the favorable designation of "knowledge" should fall to beliefs in human history that happen to have been recruited for overcoming obstacles to the satisfaction of human needs or wants. The name of "science" therefore refers to those beliefs or attitudes that serve to confer upon certain organisms in the lineage of *Homo sapiens* an evolutionary advantage. It emphatically does not refer to beliefs purporting to portray the world "as it really is," for how could a species have acquired the ability to represent the universe—especially the universe as it really is, as opposed to how it is usefully described, relative to the particular needs of members of that species? Pragmatists today are legion and diverse, and are found in all areas of the academy, as well as in all areas of philosophical inquiry. Many of them reject the original doctrine of classical pragmatism, according to which science does not seek to portray the world as it is. Still, uniting pragmatists is the idea that science is "natural" in the sense that it is continuous with common sense, and hence that science serves each human organism—to the extent it serves at all—in the way that a bat's wing serves the bat, or (perhaps more aptly) in the way that a political institution might serve those in a position to benefit from it.

Origins of science. These fundamentally opposing ideas—Aristotle's original idea on one side and the pragmatists' uniting idea on the other—lie at the foundations of the raging controversy today known as the "science wars." While not all pragmatists oppose the authoritative—or at least favorable—status that science has enjoyed in recent history, they agree with those who do so oppose it on the matter of the natural history of science. Those who side with Aristotle—who view science as exerting a special force over its own history, a drive to the truth—need to present a contrary natural history. We await that history, according to which science is a development in the natural order, but in which science falls into a different category from common sense. Such a history will present science quite differently. Unlike the bat's wing, which serves the bat whose wing it is, science (according to this still-awaited account) does not serve directly the interests of those who further it. This long-awaited account must present science as an entity susceptible also of a natural history, distinct from the natural history of the species *Homo sapiens*, and distinct as well from the natural history of technology (which, as has recently been shown by Jared Diamond, has a natural history that is firmly entangled with the natural history of our species and the local conditions under which each of its societies has had to face).

While we await the outcome of these debates on the evolutionary origins of science, the field of philosophy of science continues to debate the questions with which the field, as a subspecialty in philosophy, first came to life: What is a scientific explanation? What is a law of nature? What is causality that even the newborn is capable of recognizing it on the ground? And how are causal regularities related to the laws of nature? How do we attain knowledge of these laws (if they exist) by which nature is governed? And how are the laws of, say, biology related to the laws of physics? Are there natural laws that govern the human social sphere? Each of these questions opens out onto an area of academic inquiry in which philosophical methods interact today with the products of science themselves, and the outcome is very often a wealth of new and exciting philosophical and empirical questions. One illustration of this phenomenon pertains to research on the topic of reason—which is at such a premium in very science itself. In Renaissance times, the laws governing reasoning—the laws of right thinking, as René Descartes referred to them, by which we entertain hypotheses regarding the nature of the natural world—were not thought to be of an empirical nature: no amount of laboratory research, it was thought, could bear upon how to think correctly. But as debate on the rules of right thinking has advanced, this position is increasingly under pressure to give way to the idea that economy is a desideratum, even in thought; and since what counts as economical depends upon what resources are available to the thinking entity, the facts pertaining to the conditions of evolution will have a great deal of bearing upon "right reason." And so the findings of cognitive psychologists, and even more profoundly evolutionary psychologists, will surely bear upon the question of what counts as right reason. And so a new tangle of questions is revealed, pertaining to how specifically the findings of empirical science ought to bear upon that original question of right reason, formerly thought to belong exclusively to the philosopher's private preserve. *See* EMPIRICAL METHOD; EXPERIMENT; PHYSICAL LAW.

Scientific progress. Science is a relative newcomer on the evolutionary scene, while technological innovation is considerably older. What is perhaps indisputable is that science grew up alongside philosophy. It responded favorably to the ministrations of those cultures that promoted and protected arenas (however exclusive of women, foreigners, and the underclasses) for disputation of ideas. The concept of an open society plays a large role in the philosophies of the most important twentieth-century figures in the philosophy of science: Karl Popper and Thomas Kuhn. Science requires an integrated and highly interactive community. No individual thinker can attain the vision for scientific innovation unless he or she will submit to standing upon the shoulders of those giants who came before—and this is increasingly true as history marches forward. And so we are left with a tension between innovation and deference: a culture of disputation is the most fertile ground for science; but no long-term scientific enterprise can flourish without accumulation of insight. In the face of these facts, the questions of how exactly science makes progress—still a central open question in the philosophy of science—becomes more urgent than ever.

The traditional cumulative conception of scientific

knowledge (that scientific knowledge grows simply through accumulation or accretion), accepted without debate up until the nineteenth century, was effectively challenged in the twentieth century by Karl Popper and Thomas Kuhn, in quite similar ways. They argued convincingly that the foundations of scientific thinking (sometimes discussed as the logic of science, or the rationality of science) is a logic of disputation—as Popper put it, a logic of conjectures and refutations, a logic of challenge and reply. No scientific hypothesis is immune to challenge. And so, far from being a predominantly antiquarian exercise in maintaining old knowledge and increasing it when possible, the primary scientific exercise is better characterized as disputation of old orthodoxy and proposals for its replacement with a challenger, when the challenger can stand up to further challenges. *See* HYPOTHESIS.

The cumulative view of scientific progress, still quite common today, was an important ingredient in the optimism of the eighteenth-century enlightenment thinkers—Frederick Hegel and Jean Jacques Rousseau. But if (as philosophers currently believe) science does not make progress in this straightforward way, is there any reason to think that human society itself can be progressive? In what sense can science—as distinct from mere technological innovation—be part of the solution to the problems faced by human societies today? What can science bring to the negotiation table when the high stakes associated with public policies crafted in the twenty-first century and beyond are involved? Can science help us understand ourselves and our impact upon our fragile planet? And perhaps most importantly, can science help us craft policies and polities that will effectively safeguard the future of the Earth and the diversity of its life. These are now, as they have always been, open questions for philosophers of science.

Mariam Thalos

Bibliography. M. Curd and J. Cover (eds.), *Philosophy of Science: The Central Issues*, W. W. Norton, 1998; C. G. Hempel, *Philosophy of Natural Science*, Prentice Hall, 1966; P. Kitcher, *Science, Truth and Democracy*, Oxford University Press, 2003; H. Longino, *The Fate of Knowledge*, Princeton University Press, 2002; P. Machamer and M. Silberstein (eds.), *Blackwell Guide to the Philosophy of Science*, Blackwell, 2002.

Phlebitis

An inflammation of a vein. Individuals with phlebitis typically experience tenderness, redness, and hardness along the course of the vein. The cause of the inflammation may be related to injury of the vein or infection. The presence of varicose veins and the long-term use of indwelling intravenous catheters or irritating intravenous solutions place individuals at risk of developing phlebitis. In addition, those with certain diseases, including systemic lupus erythematosus, vasculitis, or malignancy, are at increased risk.

The condition is usually diagnosed on the basis of the clinical symptoms described above. If there is uncertainty regarding the diagnosis, a noninvasive ultrasound study can be used to differentiate among inflammation of the lymph system, tissue inflammation, and phlebitis. Two varieties of phlebitis are recognized: phlebothrombosis and thrombophlebitis.

Phlebothrombosis is a condition in which a blood clot develops within an inflamed vein. As the clot enlarges, it may detach and travel to the lung, becoming a pulmonary embolism. Thrombophlebitis begins with an inflammatory reaction in the vein wall. When the lining of the vein is damaged, three reactions influence the development of thrombosis. Initially, damage to the lining results in adherence of white blood cells, coagulation, and a loss of the lining's nonthrombogenic characteristics. Subsequently, the deep lining of the vein is exposed, bringing it into contact with blood and allowing platelets to adhere and aggregate. Finally, the exposed lining and activated platelets result in changes in coagulation, causing more platelets to interact with deep-lining structures. These factors are influenced by the velocity of blood flow in the affected area. *See* EMBOLISM; THROMBOSIS.

Symptomatic thrombophlebitis usually results in a clot which is firmly adherent to the vein wall with a decreased risk of embolizing. Some individuals may develop symptoms suggestive of deep venous thrombosis such as pain and swelling, and should undergo noninvasive ultrasound examination of the deep veins of the leg. This is particularly important in individuals with cancer. There are other factors that can predispose people to phlebitis, including stasis, immobility, tissue destruction, cardiac failure, obesity, varicose veins, and infection.

In the absence of deep venous thrombosis, the goal of treatment of superficial phlebitis is symptomatic relief. Analgesics, warm compresses and elevation of the affected limb may be beneficial. Anticoagulants are not routinely prescribed unless the thrombus is enlarging despite treatment or is in a long leg vein extending to the groin. Late effects of phlebitis include damage to the vein wall and destruction of the venous valves or obliteration of the vein. When the deep veins of the lower extremity are involved, many individuals develop chronic venous insufficiency and its associated morbidity. *See* CIRCULATION; INFLAMMATION.

Lazar J. Greenfield; Mary C. Proctor

Phloem

The principal food-conducting tissue in vascular plants. Its conducting cells are known as sieve elements, but phloem may also include companion cells, parenchyma cells, fibers, sclereids, rays, and certain other cells. As a vascular tissue, phloem is spatially associated with xylem, and the two together form the vascular system. Less is known of phloem than of xylem, partly because of its lesser direct economic importance and partly because the sieve elements function for a short time (usually one season)

and then undergo marked structural and functional changes, such as crushing or sloughing off as a result of periderm formation in the case of woody plants. *See* XYLEM.

Sieve elements. Sieve elements differ from phloem parenchyma cells in the structure of their walls and to some extent in the character of their protoplasts. Sieve areas, distinctive structures in sieve element walls, are specialized primary pit fields in which there may be numerous modified plasmodesmata. Plasmodesmata are strands of cytoplasm connecting the protoplasts of two contiguous cells. These strands are often surrounded by callose, a carbohydrate material that appears to form rapidly in plants placed under stress. How the very specialized end walls and associated sieve areas in some sieve elements (sieve-tube members) form is not fully understood, but both callose and cisternae of endoplasmic reticulum are often observed at future pore sites and these might be involved in the selective deposition or removal of cellulosic wall material in these regions. The walls of sieve elements often increase in thickness by the deposition of the so-called nacreous thickening. The protoplast of an immature sieve element is usually indistinguishable from that of a typical parenchyma cell of the same age. In the course of development into a functioning sieve element, however, several distinctive changes appear to occur. Commonly, there is a disintegration or change in appearance of the nucleus. The tonoplast that normally delimits the vacuole from the cytoplasm probably becomes disrupted during preparation of phloem samples for microscopy; thus, mature sieve elements often appear enucleate and without vacuoles. Ribosomes and dictyosomes normally disappear. A plasmalemma (cell membrane), endoplasmic reticulum, mitochondria, and plastids remain, and are generally located next to the wall.

Difficulty in providing an accurate account of the contents of mature, functioning sieve elements lies in the fact that these cells readily show signs of disturbance. This is due to their contents being under a relatively negative water potential (high turgor), because of the high solute (sucrose) concentration. Also, different groups of plants (such as gymnosperms, dicotyledons, and monocotyledons) exhibit different types of sieve elements with contents that are somewhat varied. In some sieve elements the inner lumen appears filled with the contents of the "former" vacuole and some cytoplasmic material. P-protein (formerly called slime) arises as distinct bodies in the cytoplasm of young sieve elements of dicotyledons, but may later disperse throughout the cell. P-protein is the source of the "slime" plugs, accumulations of protein on sieve areas of injured sieve elements. In many species the plastids produce a special kind of starch or proteinaceous material (**Fig. 1***b*). In monocotyledon species, where P-protein is lacking, these plastids or their contents (Fig. 1*a*) might serve to plug the sieve plate pores. Again, it should be emphasized that because of the delicate nature of sieve element contents, many of the structures (nucleus, tonoplast) reported lacking in mature sieve elements are, in fact, destroyed or displaced during tissue processing. Therefore, if one were to reconstruct the pretreatment condition, sieve elements for angiosperms would more than likely contain structures such as those diagrammed in **Fig. 2***a–f*. When a comparison is made between a sieve element that was purposely traumatized (Fig. 1*a*) and one that was not (Fig. 1*b*), it can be seen that sieve element contents look much like those of highly vacuolate parenchyma cells, except that the nucleus might be absent or appear necrotic.

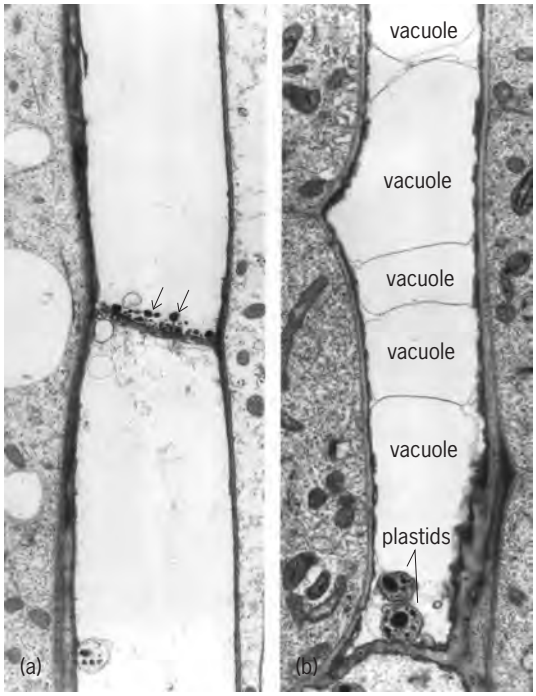

Fig. 1. Electron micrographs showing longisections of sieve elements. (*a*) Sieve elements were traumatized (cut) during pretreatment and their contents (arrows) have surged toward the sieve plate. (*b*) Sieve element with plastids and vacuoles more or less intact as a result of "gentle" pretreatment (no cutting). (*From J. E. Melaragno and M. A. Walsh, Ultrastructural features of developing sieve elements in Lemna minor L.: The protoplast, Amer. J. Bot., 63:1145–1157, 1976*)

Figure 2 shows some features that are somewhat controversial, including the nucleus, tonoplast, P-protein, and sieve plate pores. The nucleus, if retained, appears degenerate or necrotic. The tonoplast persists and may delimit one or more vacuoles in shorter sieve elements (as in Fig. 2*b* and *e*) that are more or less in direct contact with cells providing the source of photoassimilates; or it may become fused with the tonoplasts of other sieve-tube members so as to constitute one continuous vacuole throughout a series of sieve-tube members (as shown in Fig. 2*d* and *f*). The P-protein is shown associated only with sieve-tube members of dicotyledons (Fig. 2*c*, *e*, and *f*) and it persists in the form of discrete bodies that are peripheral in distribution. It should be noted that the P-protein comes in a variety of forms (fibrillar, tubular, and crystalline) and is not restricted to the sieve-tube members but is sometimes found in other phloem parenchyma cells. The

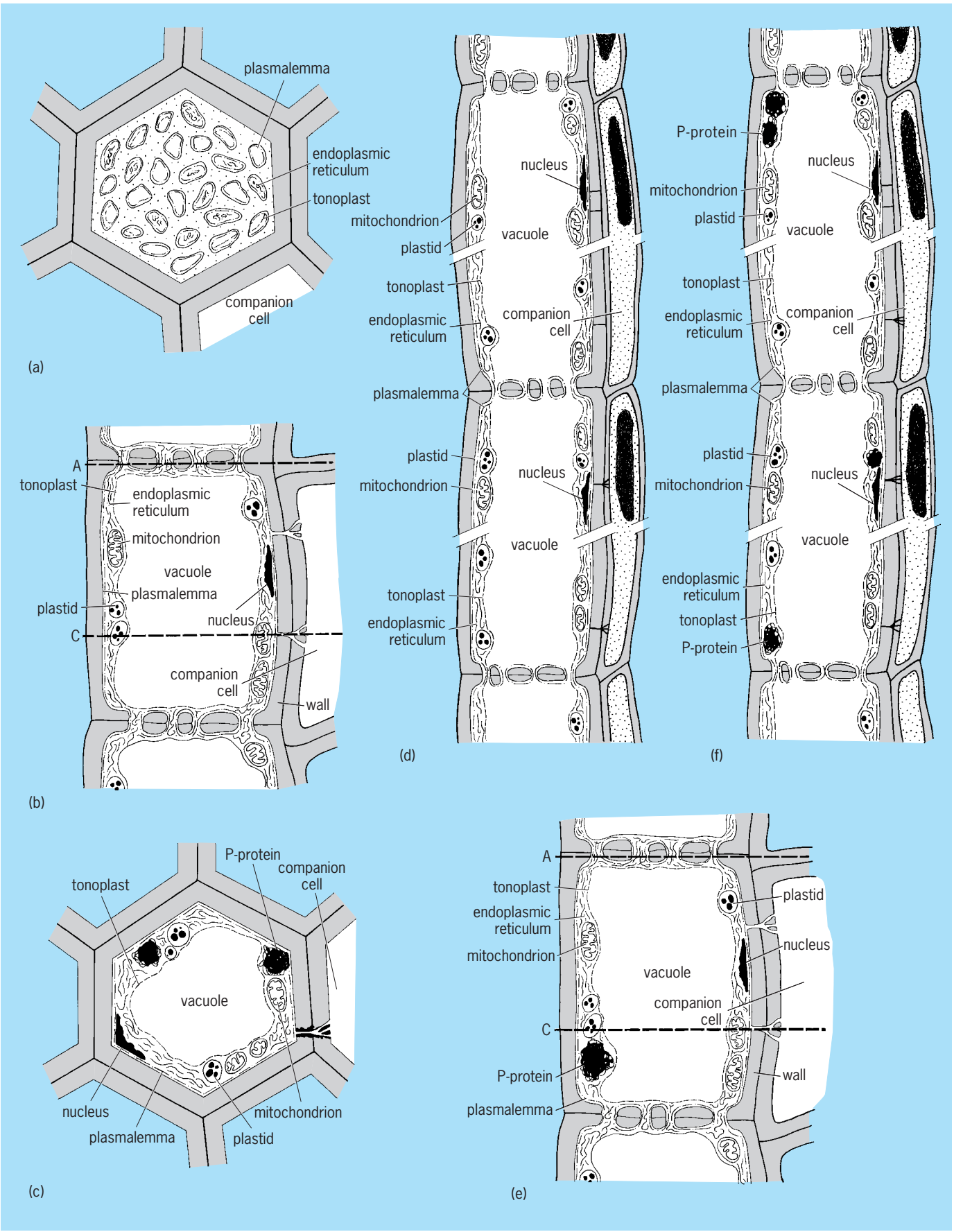

Fig. 2. Mature sieve elements in angiosperms: (*a*,*b*,*d*) portions of monocotyledon sieve elements; (*c*,*e*,*f*) dicotyledon sieve elements. In *b* and *e* the cell is bisected by lines A and C, referring to views shown in *a* and *c*.

pores are relatively unoccluded and are lined by the plasmalemma, and some may be traversed by endoplasmic reticulum and portions of a tonoplast. Sieve-tube members are accompanied by companion cells that have branched cytoplasmic connections in their walls. The connections are shown unbranched in the wall of the sieve-tube members. Plastids, mitochondria, and smooth endoplasmic reticulum are present in sieve elements of both dicotyledons and monocotyledons. Though not distinguished in Fig. 2, the plastids of monocotyledons possess only the unusual crystalline, proteinaceous material, while those of dicotyledons may possess either proteinaceous inclusions or starch.

As a sieve element becomes nonfunctioning, its contents disintegrate. The callose first increases in amount and then disappears along with the connecting strands, leaving empty pores in the wall. The element becomes filled with air and may be crushed later by enlargement of surrounding parenchyma cells.

Sieve cells and sieve-tube members. Typical sieve cells are long elements in which all the sieve areas are of equal specialization, though sieve areas may be more numerous in some walls than in others. In contrast, a sieve-tube member has some sieve areas more specialized than others; that is, the pores, or modified plasmodesmata, are larger in some sieve areas. Parts of the walls containing such sieve areas are called sieve plates. Simple sieve plates have one specialized sieve area that generally occurs on a transverse end wall; compound sieve plates have two or more on an oblique end wall (**Fig. 3**). Sieve tubes are composed of an indeterminate number of sieve-tube members arranged end to end. Sieve-tube members are shorter than sieve cells and have become progressively more so with evolutionary change.

Companion cells. Companion cells are specialized parenchyma cells that occur in close ontogenetic and physiologic association with sieve-tube members. They arise from the same meristematic cell that produces the sieve-tube member and vary in size, position, and number, but always retain their nucleus. Some sieve-tube members lack companion cells. The precise functional relationship between these two kinds of cells is unknown, but they become nonfunctioning simultaneously.

Parenchyma cells. Parenchyma cells in the phloem have thin or somewhat thick walls and occur singly or in strands of two or more cells. They store starch, frequently contain tannins or crystals, commonly enlarge as the sieve elements become obliterated, or may be transformed into sclereids or cork cambium cells. Parenchyma cells in secondary phloem may arise from a meristematic cell (phloem initial) that produces only such cells or from one that also eventually produces one or more sieve-tube members and companion cells. Parenchyma cells seem to intergrade with companion cells in the angiosperms.

Fibers. Phloem fibers vary greatly in length (from a fraction of a millimeter in some plants to 20 in. or 50 cm in the ramie plant). The secondary walls are commonly thick and typically have simple pits, but

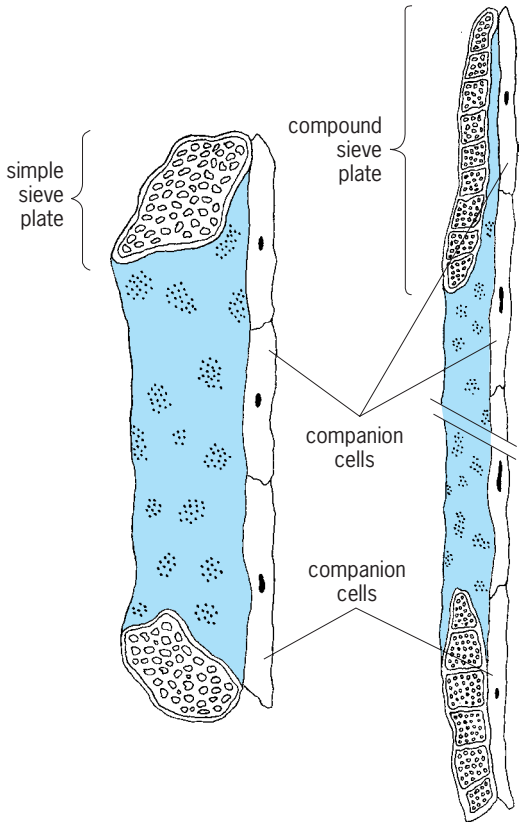

Fig. 3. Sieve-tube members and their sieve plates. Simple sieve plates are shown in the sieve element on the left, and compound sieve plates in the sieve element on the right. Lateral sieve areas are reduced in size and more or less evenly distributed along the lateral walls.

may or may not be lignified. In secondary phloem some fibers do not increase in length beyond the size of their primordia, but others may elongate extensively by apical intrusive growth. In primary phloem immature fibers elongate, sometimes hundreds of times over their original length. The fibers may become septate, are frequently multinucleate, and may intergrade with sclereids. *See* SCLERENCHYMA.

Primary phloem. Primary phloem differentiates from derivatives of the apical meristem. The earliest primary phloem (protophloem) contains sieve elements, with or without companion cells, and parenchyma cells. The sieve elements function for a brief time and then are usually obliterated. The remaining cells may become collenchymatous, as in many leaves, or be transformed into long protophloem fibers, often erroneously called pericyclic fibers. Metaphloem is formed after growth in length of surrounding cells is completed. Sieve elements, companion cells (in angiosperms), and parenchyma cells occur in such phloem, but typical fibers are generally lacking. If secondary phloem is absent, the metaphloem functions throughout the life of the plant. *See* APICAL MERISTEM.

Secondary phloem. Secondary phloem is produced by the same vascular cambium that forms secondary xylem. Such phloem consists of two interpenetrating systems, the vertical or axial and the horizontal or ray (**Fig. 4**). The phloem rays are basically similar

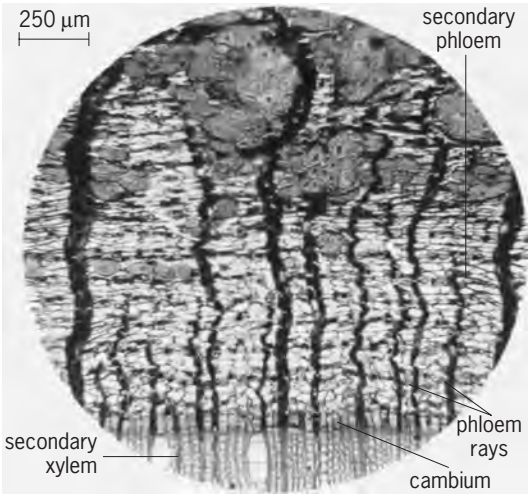

Fig. 4. Photomicrograph of cross section of the secondary phloem of paper birch (*Betula papyrifera*). (*Forest Products Laboratory*, *USDA*)

to xylem rays, but their component cells differ in typically lacking secondary walls. Moreover, as the girth of the stem or root increases, the older phloem ray cells increase in width and may divide radially. This dilatation does not occur in all phloem rays, but it is a common feature of secondary phloem and stops only at the time of periderm formation within the ray. The vertical system contains sieve elements, parenchyma, often fibers or sclereids, and infrequently other elements such as laticifers. The fibers may occur singly, in dispersed groups, or in tangential bands.

Conifers and lower vascular plants. The phloem of conifers contains long sieve cells and parenchyma cells and frequently fibers. In secondary phloem of conifers (**Fig. 5**), these cells may be arranged in regularly alternating bands that give an orderly appearance to the phloem as seen in transection. In the ferns and fern allies (lower vascular plants) the phloem is mostly of a primary nature and also

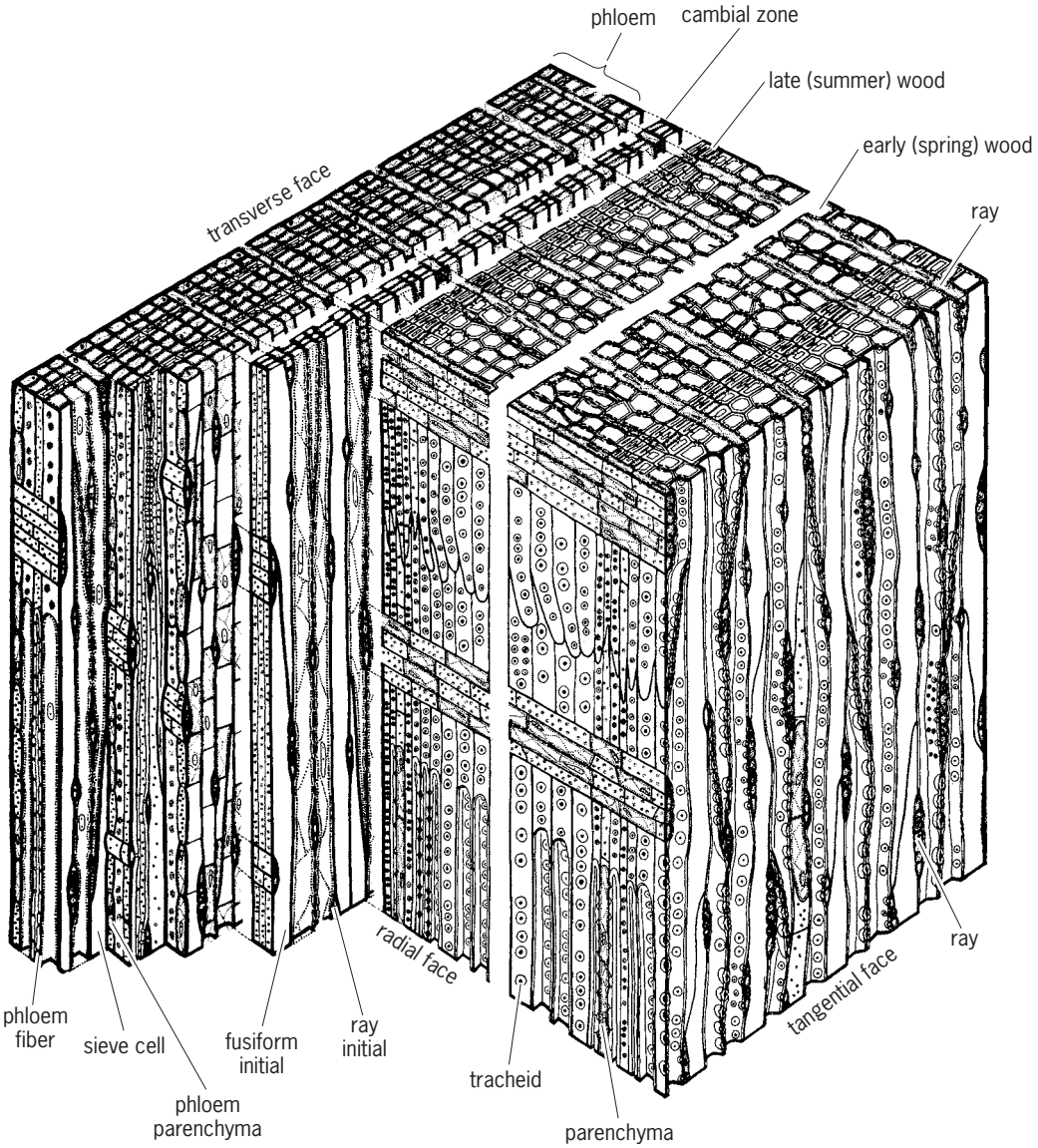

Fig. 5. Block diagram showing the three primary tissues: the secondary xylem, the cambial zone, and the secondary phloem of the conifer (gymnosperm) *Thuja*, or white cedar.

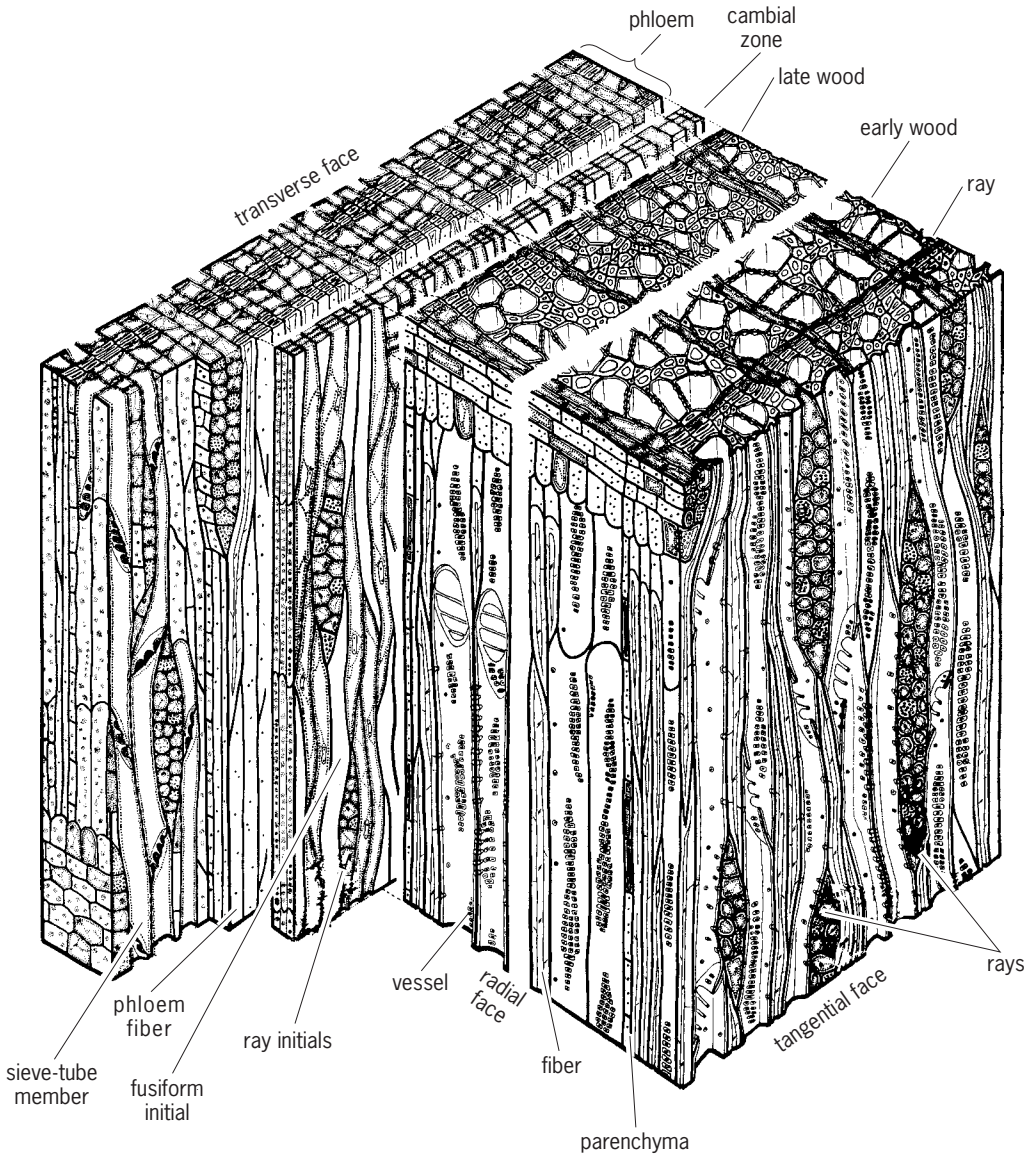

Fig. 6. Block diagram showing the three primary tissues: the secondary xylem, the cambial zone, and the secondary phloem of the dicotyledon (angiosperm) *Liriodendron* (tulip tree).

consists of sieve cells, parenchyma cells, and fibers. Special structures called refractive spherules occur in the sieve cells of lower vascular plants. Their origin and function are not fully understood, though they are believed to form from the endoplasmic reticulum and Golgi apparatus. Cytochemical studies reveal these to consist of proteinaceous material. *See* PINALES; POLYPODIOPHYTA.

Dicotyledons. The phloem in dicotyledons (**Fig. 6**) has greater diversity of cell structure and of arrangement than that in the conifers. It contains in varying proportions and groupings sieve-tube members, companion cells, parenchyma cells, and often fibers, sclereids, and various other kinds of cells or cell groups, such as secretory cells. The various cells may be arranged in alternating bands or have no regular spatial disposition. The functioning phloem is generally more orderly in appearance than the nonfunctioning. This difference results from partial or total collapse of the older sieve elements and associated companion cells, and frequently from the concurrent enlargement of neighboring parenchyma cells. *See* MAGNOLIOPSIDA.

Moncotyledons. Phloem in monocotyledons is much like that in dicotyledons. However, because no secondary growth is derived from a vascular cambium, there is no secondary phloem in monocotyledons. For this reason much of the phloem, at least in perennial monocotyledons, remains functional for 50–100 years or more, unlike that of some species of woody dicotyledons which might maintain functional phloem for only 5–10 years. In addition, sieve-tube members in monocotyledons appear to lack the P-protein material associated with sieve-tube members of dicotyledons. On the other hand, the sieve-tube members of all monocotyledons thus far investigated possess plastids that contain quasicrystalline, proteinaceous materials. *See* LILIOPSIDA.

Nonvascular plants. Although it may seem a contradiction, there are some nonvascular plants that

contain conducting tissues, and cells which may be reminiscent of sieve elements. Examples are the larger mosses (bryophytes) and members of the brown algae (Phaeophyta). Here the special conducting cells are called leptoids and trumpet hyphae, respectively. *See* BRYOPHYTA; PHAEOPHYCEAE; PLANT TRANSPORT OF SOLUTES. Michael A. Walsh

Bibliography. H. D. Behnke and R. D. Sjolund (eds.), *Sieve Elements*, 1990; J. Cronshaw, W. J. Lucas, and R. T. Giaquinta (eds.), *Phloem Transport*, vol. 1, 1987; A. Fahn, *Plant Anatomy*, 1990.

Phlogopite

A mineral with an ideal composition of $KMg_3(AlSi_3)O_{10}(OH)_2$. Phlogopite belongs to the mica mineral group. It has been occasionally called bronze mica.

Phlogopite is a trioctahedral mica, where all three possible octahedral cation sites are occupied by magnesium (Mg). The magnesium octahedra, $Mg(O,OH)_6$, form a sheet by sharing edges (see **illus.**). As in all micas, tetrahedra are located on either side of the octahedral sheet, which may be occupied by aluminum (Al) or silicon (Si). Adjacent tetrahedra share corners to form a two-dimensional network of sixfold rings, thus producing a tetrahedral sheet. Two opposing tetrahedral sheets and the included octahedral sheet form a 2:1 layer. Potassium (K) ions are located between adjacent tetrahedral sheets in the interlayer region.

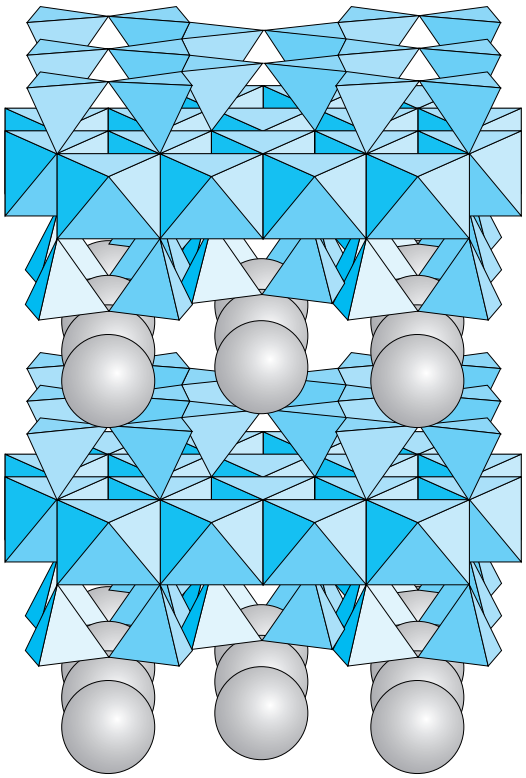

Polyhedral model illustrating the one-layer monoclinic phlogopite atomic structure. The structural unit consists of only one layer; note that both layers shown are identical. The corners of the tetrahedra represent oxygen atoms, whereas octahedral corners represent oxygen or hydroxyl. Potassium ions are depicted as large spheres between and connecting the layers.

Solid solutions between phlogopite and siderophyllite or between phlogopite and lepidomelane are complete, balanced by substitutions of aluminum and silicon, or vacancies in the sites usually occupied by magnesium. Proper usage of the term phlogopite refers to end members, for example, with less than 10% substitutions for magnesium. In the past, the term phlogopite was often used for samples greater than 70% magnesium. Biotite is used to describe the series along or near the hypothetical annite-phlogopite join. Solid solutions between phlogopite and the dioctahedral micas are quite limited. *See* BIOTITE; SOLID SOLUTION.

A perfect basal cleavage occurs because of the layer structure. Specific gravity is 2.86, hardness on the Mohs scale is 2.5–3.0, and luster is vitreous to pearly. Thin sheets are flexible. Color is yellow brown, reddish brown, or green, and thin sheets are transparent. Thermal stability varies greatly with composition, with iron or fluorine substitutions reducing or increasing stability, respectively. Weathering of phlogopite may produce vermiculite. *See* HARDNESS SCALES; VERMICULITE; WEATHERING PROCESSES.

Phlogopite occurs in marbles produced by the metamorphism of siliceous magnesium-rich limestones or dolomites and in ultrabasic rocks, such as peridotites and kimberlites. *See* DOLOMITE; LIMESTONE; PERIDOTITE.

Phlogopite is used chiefly as an insulating material and for fireproofing. It has high dielectric properties and high thermal stability. *See* MICA; SILICATE MINERALS. Stephen Guggenheim

Bibliography. S. W. Bailey (ed.), *Micas*, Reviews in Mineralogy, vol. 13, 1984.

Phobia

An intense irrational fear that often leads to avoidance of an object or situation. Phobias (or phobic disorders) are common (for example, fear of spiders, or arachnophobia; fear of heights, or acrophobia) and usually begin in childhood or adolescence. Psychiatric nomenclature refers to phobias of specific places, objects, or situations as specific phobias. About one in two individuals report that they are afraid of public speaking; in very severe cases, this is considered a form of social phobia. Social phobias also include other kinds of performance fears (such as playing a musical instrument in front of others; signing a check while observed) and social interactional fears (for example, talking to people in authority; asking someone out for a date; returning items to a store). Individuals who suffer from social phobia often fear a number of social situations.

Types. Although loosely regarded as a fear of open spaces, agoraphobia is actually a phobia that results when people experience panic attacks. Panic attacks are unexpected, paroxysmal episodes of anxiety and accompanying physical sensations (for example, racing heart; shortness of breath) that can occur at

any time in susceptible individuals. When individuals have recurrent panic attacks, they often begin to avoid situations where they believe that a panic attack might occur. Agoraphobia develops in approximately two out of three individuals who experience recurrent, spontaneous panic attacks. School phobia is a misnomer, because affected children do not actually fear school. Some fear being away from their parents; that is, they suffer from separation anxiety. Others fear having to perform in front of or interact with their peers or teachers; that is, they suffer from social phobia. Individuals with phobias about dirt and germs may suffer from an anxiety disorder called obsessive-compulsive disorder. Their dirt and germ phobia is a manifestation of an obsession with contamination fears.

Origins. The origin of phobias is varied and incompletely understood. Most individuals with specific phobias have never had anything bad happen to them in the past in relation to the phobia. In a minority of cases, however, some traumatic event occurred that likely led to the phobia. It is probable that some common phobias, such as a fear of snakes or a fear of heights, may actually be instinctual, or inborn. Both social phobia and agoraphobia run in families, suggesting that heredity plays a role. However, it is also possible that some phobias are passed on through learning and modeling, such as when a child sees a parent shy away from encounters with strangers.

Occurrence. Phobias occur in over 10% of the general population. Social phobia may be the most common kind, affecting approximately 7% of individuals. Agoraphobia occurs in approximately 1% of the general population. Most phobias are distressing to the affected person, and in some cases they are even disabling. Some people with social phobia quit school because of the intense distress they experience when they give an oral presentation to the class. Others avoid so many situations where they fear a panic attack that they become housebound.

Reaction. When persons encounter the phobic situation or phobic object, they typically experience a phobic reaction consisting of extreme fearfulness, physical symptoms (such as racing heart, shaking, hot or cold flashes, or nausea), and cognitive symptoms (particularly thoughts such as "I'm going to die" or "I'm going to make a fool of myself"). These usually subside quickly when the individual is removed from the situation. The tremendous relief that escape from the phobic situation provides is believed to reinforce the phobia and to fortify the individual's tendency to avoid the situation in the future.

Treatment. Many phobias can be treated by repeated and prolonged exposure of the individual to whatever is feared. Often, the individual is gradually encouraged to approach the feared object and to successively spend longer periods of time in proximity to it. In some cases, the individual may just have to imagine that he or she is in the situation, but usually it is more effective to have the individual actually encounter it. Cognitive therapy is also used (often in conjunction with exposure therapy) to treat phobias. It involves helping individuals to recognize that their beliefs and thoughts can have a profound effect on their anxiety, that the outcome they fear will not necessarily occur (for example, the likelihood of their plane crashing is very low), and that they have more control over the situation than they realize.

Medications are sometimes used to augment cognitive and exposure therapies. For example, musicians with performance phobias (stage fright) sometimes use beta-adrenergic blocking agents, such as propranolol, prior to the performance. This lowers their heart rate and reduces tremulousness, and leads to reduced anxiety. For agoraphobia and social phobia, certain kinds of antidepressants and anxiolytic medications are often helpful. It is not entirely clear how these medications exert their antiphobic effects, although it is believed that they affect levels of neurotransmitters in regions of the brain that are thought to be important in mediating emotions such as fear.

Murray B. Stein

Bibliography. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed., 2000; R. J. McNally, *Panic Disorder: A Critical Analysis*, 1994; M. B. Stein, *Social Phobia: Clinical and Research Perspectives*, 1995.

Phoenicopteriformes

The flamingos, a small monotypic order of wading birds that includes the family Phoenicopteridae, which has five species in one genus *Phoenicopterus* found worldwide in tropical to temperate marine and fresh waters; four species are found in South America, of which three live in the Andes. The flamingos were formally included in the Ciconiiformes and are still placed there by many researchers; others have advocated a close relationship with either the Anseriformes or the Charadriiformes. None of these classifications, however, are strongly supported by available evidence, and so it is best to place these bizarre birds in a unique order. *See* ANSERIFORMES; CHARADRIIFORMES; CICONIIFORMES.

Fossil record. Several Cretaceous fossils, including an Early Cretaceous form, had been labeled flamingos; however, such attributions are most dubious. The earliest definite flamingo fossil is *Juncitarsus*, from the middle Eocene of Wyoming; it possesses several primitive traits for the family. Beginning in the late Oligocene, modern flamingos are found in the fossil records from all areas of the world.

Characteristics. Flamingos (see **illus.**) are long-legged, long-necked wading birds with a thick bill which is sharply bent downward at the midpoint. The bill is deep, with many transverse lamellae on both the mandible and maxilla, and the tongue is thick and fleshy. During feeding, the head is held upside down with the anterior portion of the bill parallel to and just below the water surface. Water is pumped through the bill by the tongue so that small food particles are trapped between the lamellae; however the structure and mechanics of the filtering method of flamingos differs from that in the waterfowl. The three anterior toes are webbed, possibly for walking on soft substrates. The long and

Greater flamingo (*Phoenicopterus ruber*). (*Photo by Dr. Lloyd Glenn Ingles; California Academy of Sciences*)

broad wings enable the bird to fly well. The adult plumage is pink to light red, with black flight feathers. Flamingos are gregarious, often congregating in flocks in excess of 1 million in the rift valley lakes of Africa.. They breed monogamously in large colonies, and both parents incubate the one or two eggs in a nest formed of a stout pillar of mud 1 ft (0.3 m) high on a mud flat. The downy, gray young leave the nest after hatching to be cared for by both parents, while the older young are kept together in large groups. Most flamingos are worldwide, found in tropical waters, either in shallow lakes or along the coasts, although three species live on shallow lakes in the high Andes of South America. *See* AVES. Walter J. Bock

Bibliography. J. Kear and N. Duplaix-Hall, *Flamingos*, Poyser, 1975; J. del Hoyo, Order Phoenicopteriformes, in J. del Hoyo et al. (eds.), *Handbook of the Birds of the World*, vol. 1, pp. 508–526, Lynx Edicions, 1992.

Pholidota

An order of mammals comprising the living pangolins, or scaly anteaters, and their poorly known fossil predecessors. All living pangolins are assigned to the genus *Manis*. They are found in Africa south of the Sahara and in southeastern Asia, including certain islands of the East Indies.

Pangolins feed principally on termites and ants. The elongate tubular skull without teeth, long protrusive tongue, small eyes with heavy eyelids, thick skin, strong legs, five-toed feet with large claws, and large tail enable these unique animals to rip open ant nests and termite dens and devour the animals therein. The greatest peculiarity of animals in the genus *Manis* is a covering of all but the undersides of the body by an armor of large imbricating dermal horny scales. Living pangolins are frequently characterized as being animated pine cones. The position and number of hairs in relation to the scales are peculiar to each modern species.

Scattered and rare fossil bones from the Oligocene and Miocene rocks of western Europe have been named *Necromanis*, *Leptomanis*, *Teutomanis*, and *Galliaetatus*, but the validity and complete significance of these names remain doubtful. *Manis* or a closely related genus is found in the Pleistocene of Hungary and southeastern Asia.

Although the phyletic origin of the Pholidota is conjectural, it is believed that the general anatomical similarities of these scaly anteaters to the true anteaters (order Edentata) and to the aardvarks (order Tubulidentata) are the product of convergent evolution between three groups of distantly related eutherian mammals. A common ancestor for anteaters, exclusive of the marsupial anteaters, may be represented by a Late Cretaceous eutherian that is close to or in the order Insectivora. *See* EDENTATA; EUTHERIA; INSECTIVORA; MAMMALIA; TUBULIDENTATA. Donald E. Savage

Bibliography. M. J. Benton, *Vertebrate Paleontology*, 1991; R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988.

Phonetics

The science that deals with the production, transmission, and perception of spoken language. At each level, phonetics overlaps with some other sciences, such as anatomy, physiology, acoustics, psychology, and linguistics. In each case, phonetics focuses on phenomena relevant to the study of spoken language. The outline presented below follows a speech signal from its production by the speaker through the medium of transmission to its perception by the hearer.

Articulation. Speech is normally produced by exhaling air from the lungs through the vocal tract. (Other airstream mechanisms are used in some languages.) The vocal tract extends from the larynx through the pharynx and the oral cavity to the lips. If the velum (soft palate) is not raised, the air also passes through the nasal cavities. The shape and size of the oral cavity can be varied by the movement of active articulators: tongue, lips, and velum. *See* PALATE.

Phoneticians usually describe speech sounds with reference to their point (or place) of articulation (**Fig. 1**) and their manner of articulation. The point of articulation of a sound is the place of maximum constriction within the vocal tract. The great majority of sounds are produced by moving some part of the tongue toward some region on the roof of the mouth.

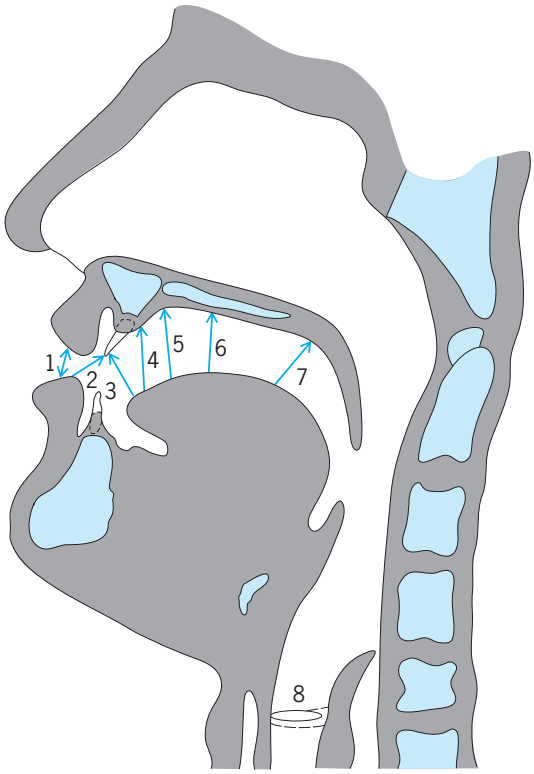

Fig. 1. Sagittal section of the head, showing the vocal tract and identifying points of articulation used in English. 1 = bilabial; 2 = labiodental; 3 = dental; 4 = alveolar; 5 = palatoalveolar; 6 = palatal; 7 = velar; 8 = glottal.

Exceptions are articulations involving lips and those sounds in which the vocal folds serve as articulators (see **table**).

Manners. At most of these points of articulation, sounds can be produced with several manners of articulation. One way to classify manners of articulation refers to the degree of stricture employed in producing the sound. Sounds produced with complete constriction of the vocal tract are stops, or plosives. The term "stop" refers to the fact that the airstream is completely stopped at the point of articulation. The term "plosive" focuses on another aspect of the production of stops: the fact that if a complete closure is formed in the vocal tract, while the lungs continue to produce an egressive airstream, there is a pressure buildup behind the closure until the hold is suddenly released, resulting in a weak explosion.

If the closure is incomplete, but the articulators are brought close enough so that the air passing between them is set into turbulent motion, the resultant sounds are fricatives or spirants. If the articulators are approximated but the constriction remains large enough that air can pass through without friction, the sounds are called approximants—vowellike sounds functioning as consonants. The distinction between vowels and consonants is based primarily on their function: vowels may constitute a syllable nucleus and may be produced in isolation, without any accompanying consonants, while consonants ordinarily constitute syllable margins and normally are not produced in isolation. There are numerous counterexamples, though; in some languages, [r] and [l] sounds can be syllable nuclei, for instance.

Other manners of articulation include nasals, which are produced with a lowered velum, so that the airstream passes through the nasal cavity while an articulation is maintained in the oral cavity; laterals, which are produced in such a way that the airstream passes over one side of the tongue while a constriction is formed by the other side; trills, which are produced by adjusting the tension of the muscles of the articulators and the rate of airflow in such a manner that an articulator is set into vibration; taps and flaps, in the production of which an articulator executes a single rapid closure. Most of these sounds can be produced at several points of articulation. Nasals are produced at the same points of articulation as stops (nasals are sometimes called nasal stops); laterals may be dental, alveolar, palatal, or even velar; trills can be produced with the lips, the tip of the tongue, or the uvula; and taps and flaps are produced with the tip of the tongue or the uvula. Symbols for consonant sounds are shown in **Fig. 2**.

Voiced and voiceless sounds. Most of these consonant sounds can be voiced or voiceless; vowels are normally voiced. The terms "voiced" and "voiceless" refer to the presence and the absence of vocal fold vibration. According to the generally accepted myoelastic-aerodynamic theory of phonation, the process consists of the following steps. The vocal folds are approximated and tensed through the action of the laryngeal musculature. The airstream passing through the narrowed opening between the vocal folds (the glottis) increases its velocity, and a negative pressure is generated (the Bernoulli effect) that pulls the folds together. Air pressure from the lungs builds up behind the vocal folds and blows

The points of articulation used in English and sample sounds produced at them

Point of articulation	Articulators and movement	Example
Bilabial	Lips	Initial sound of *peel*
Labiodental	Lower lip against upper teeth	Initial sound of *feel*
Dental	Tip of tongue against upper teeth	Initial sound of *thin*
Alveolar	Tip of tongue against the alveolar ridge	Initial sound of *tea*
Palatoalveolar the alveolar ridge and the hard palate	Blade of the tongue toward a region between	Initial sound of *she*
Palatal	Central part of the tongue toward the hard palate	Initial sound of *hue*
Velar	Back of the tongue against the soft palate	Initial sound of *cool*
Glottal	Vocal folds serving as articulators	Occurs in English at the beginning of emphatically produced vowel-initial words

consonants	bilabial	labiodental	dental and alveolar	retroflex	palato-alveolar	alveolo-palatal	palatal	velar	uvular	pharyngeal	glottal
plosive	p b		t d	ʈ ɖ			c ɟ	k g	q ɢ		ʔ
nasal	m	ɱ	n	ɳ			ɲ	ŋ	ɴ		
lateral			l	ɭ			ʎ				
lateral fricative			ɬ ɮ								
rolled			r						ʀ		
flapped			ɾ	ɽ					ɽ		
rolled fricative			ɾ								
fricative	ɸ β	f v	θ ð l s z ɹ	ʂ ʐ	ʃ ʒ	ɕ ʑ	ç j	x ɣ	χ ʁ	ħ ʕ	h ɦ
frictionless continuants and semivowels	w ɥ	ʋ	ɹ				j (ɥ)	(w) ɣ	ʁ		

(a)

	bilabial	labiodental	dental	alveolar	postalveolar	retroflex	palatal	velar	uvular	pharyngeal	glottal
plosive	p b		t d			ʈ ɖ	c ɟ	k g	q ɢ		ʔ
nasal	m	ɱ	n			ɳ	ɲ	ŋ	ɴ		
trill			r						ʀ		
tap or flap			ɾ			ɽ					
fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç j	x ɣ	χ ʁ	ħ ʕ	h ɦ
lateral ficative			ɬ ɮ								
approximant			ɹ			ɻ	j	ɰ			
lateral approximant			l			ɭ	ʎ	ʟ			
implosive	ɓ		ɗ			ᶑ	ʄ	ɠ	ʛ		
ejective stop	p' b'		t' d'			ʈ' ɖ'	c' ɟ'	k' g'	q' ɢ'		

(b)

Fig. 2. Consonant chart of the International Phonetic Association. (*a*) Chart used in phonetic transcription before 1989 (*after The Principles of the International Phonetic Association, International Phonetic Association, London, 1949, reprint 1961*). (*b*) Revised chart adopted in 1989; where symbols appear in pairs, the one to the right represents a voiced consonant (*after Report on the 1989 Kiel Convention, J. Int. Phonetic Ass., 19(2):67–80, 1989*).

them open. The elasticity of the vocal fold muscles, their inward momentum, and the Bernoulli force cause the vocal folds to return to their original position and close, and thus the cycle is repeated over and over again—the vocal folds are set in vibration, which continues as long as the aerodynamic forces and muscular tension are maintained at appropriate levels. (In certain types of phonation the vocal folds do not close completely as part of the cycle.) Research, primarily utilizing the microbeam technique, has greatly clarified the functions of the individual laryngeal muscles in producing various types of phonation and in controlling the rate of vocal-fold vibration. *See* BERNOULLI'S THEOREM.

Classifying vowels. There are three generally recognized methods for classifying vowels. The traditional method of classification employs the notion "highest point of the tongue" and locates that highest point in a three-dimensional space whose axes are the high-low (vertical) dimension, the front-back (horizontal) dimension, and lip-rounding. While this system has some pedagogical utility, x-ray studies of vowel articulation have shown that the notion "highest point of the tongue" cannot be supported by observation. A modification of this system was developed by Daniel Jones, whose Cardinal Vowel System (**Fig. 3***a*) is based on the following principles. There are two "hinge" vowels that are articulatorily defined: a high front vowel (Cardinal Vowel One) produced with a tongue position as far front as possible and as high as possible without generating friction, and the lowest possible back vowel (Cardinal Vowel Five). Based on these two "hinge" vowels, there are six additional cardinal vowels, for a total of four front vowels and four back vowels, produced so that the "highest

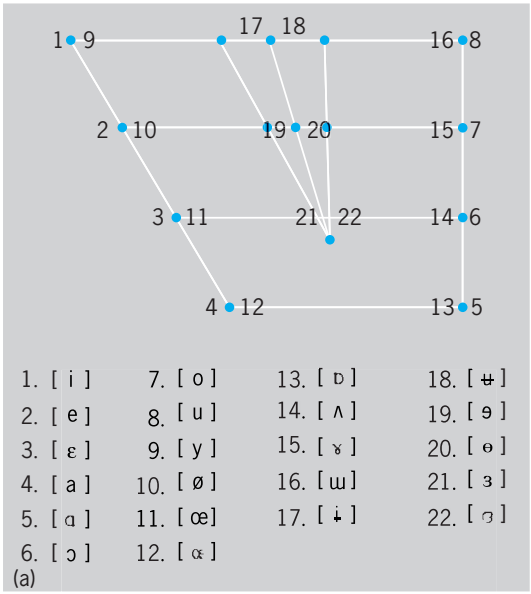

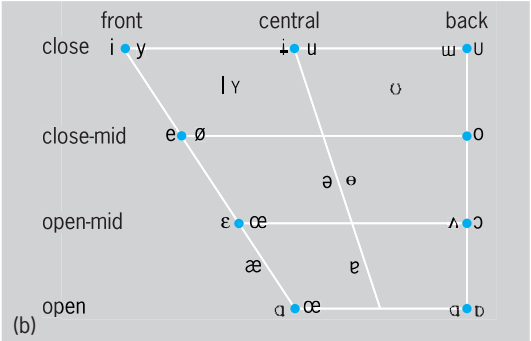

Fig. 3. Charts of vowel sounds. (*a*) Diagram showing the positions of Daniel Jones's Cardinal Vowels. The symbols for the Cardinal Vowels are identified by 1–22. (*b*) Revised chart adopted by the International Phonetic Association in 1989. Where symbols appear in pairs, the one to the right represents a rounded vowel. (*After Report on the 1989 Kiel Convention, J. Int. Phonetic Ass., 9(2):67–80, 1989*)

point of the tongue" for each of them lies on the extreme outside limits of the articulatory vowel space, and so that they are at equal distances from each other. The cardinal vowels are arbitrarily selected reference points (they do not necessarily coincide with any vowels in a given language), and thus can be used to describe all languages of the world. Of the original eight cardinal vowels, five were unrounded and three rounded; later additions to the system include the missing rounded/unrounded counterparts of the original (primary) cardinal vowels as well as a series of central vowels.

Attempts have been made to bring the International Phonetic Alphabet more in line with contemporary linguistic thought. One of these attempts tries to incorporate the insights of distinctive features theory into the classification of speech sounds. In phonological theory, the fundamental units of speech are not phonemes (contrastive speech sounds) but distinctive features, which can be used to represent the sounds of all languages of the world. Groups of features tend to be implemented simultaneously to form segments, which have been called bundles of distinctive features. Efforts have been made to develop articulatory as well as acoustic-perceptual correlates of the features. One inventory of features organizes them into two groups depending on whether the vocal tract is relatively open or constricted, corresponding to vocalic and nonvocalic features. The vocalic features are high, low, back, round, nasal, spread glottis, and constricted glottis. The nonvocalic features are sonorant, continuant, coronal, strident, consonantal, anterior, lateral, distributed, and voice. The International Phonetic Association charts in Figs. 2*a* and 3*a* remain a basic tool for interpreting phonetic transcriptions made up to 1989. However, in 1989 the Council of the International Phonetic Association approved a revised version of these charts. The revised consonant chart (Fig. 2*b*) contains columns for the following places of articulation: bilabial, labiodental, dental, alveolar, postalveolar, retroflex, palatal, velar, uvular, pharyngeal, and glottal. The manners of articulation included in the revised consonant chart are plosive, nasal, trill, tap or flap, fricative, lateral fricative, approximant, lateral approximant, ejective stop, and implosive. Modifications of the vowel chart (Fig. 3*b*) are less extensive. The Council also approved various changes in the use of diacritics.

Jones had originally assumed that the equal distance between the vowels meant articulatory distance, but since x-rays showed this not to be true, the definition was changed to refer to auditory equidistance. This definition retains a degree of subjectivity. A relatively more objective way to describe vowels is provided by acoustic analysis.

Acoustic phonetics. This branch of phonetics deals with the manner in which the spoken message is encoded in the sound waves. According to the generally accepted source-filter theory of speech acoustics, sound is generated at a source (which for phonated speech is constituted by the vibrating vocal folds) and passed through the vocal tract. The opening and closing of the vocal folds creates a succession of condensations and rarefactions of air molecules—variations in air pressure—and transforms kinetic energy into acoustic energy. The sound wave that is generated at the glottis can be considered, for practical purposes, a complex periodic wave, and as such it contains energy at frequencies that are multiples of the fundamental frequency (harmonics). *See* HARMONIC (PERIODIC PHENOMENA); SOUND.

The vocal tract acts as a filter, transmitting more energy at those frequencies that correspond to the resonances of the vocal tract than at other frequencies. This modified sound wave is then radiated through the mouth-opening at the lips and propagated through the outside air. Energy concentrations at the resonance frequencies of the vocal tract are referred to as formants (the formants may also be defined as the resonances of the tract).

In principle, the source and filter are independent of each other; consider the fact that the same vowel can be sung at different fundamental frequencies (pitches), and different vowels can be produced at the same pitch. The sound wave can be described by specifying its fundamental frequency, amplitude, and spectrum. The fundamental frequency corresponds to the rate of vibration of the vocal folds. Amplitude is dependent upon the rate of airflow and the efficiency with which muscular energy is converted to acoustic energy in the production of the sound. The spectrum—the distribution of energy in the frequency domain—depends on the size and shape of the vocal tract. In perception, fundamental frequency is correlated with perceived pitch, amplitude with perceived loudness, and spectrum with the phonetic quality of the generated speech sounds. The sound source can also be a nonperiodic one, which is the case with fricatives and the release-bursts of stops. *See* LOUDNESS; PITCH.

Sound spectrograph. While many instruments have been developed for the study of the acoustic structure of spoken language, the sound spectrograph remains a basic tool. The spectrograph is an electronic device that analyzes the acoustic signal into its constituent parts and converts it into a visual display, with time represented on the horizontal axis, frequency on the vertical axis, and intensity conveyed through the darkness of the patterns traced on the spectrogram by a stylus (**Fig. 4**). The selection of analyzing filters makes it possible to concentrate either on the harmonics or on the formant structure of the speech signal. A study of spectrograms of vowels reveals that as a first approximation, the vowels produced with a high tongue position have a first formant with a relatively low frequency, while vowels produced with a low tongue position have a high first formant. Front vowels have relatively high second formants, while back vowels have low second formants. A standard acoustical vowel diagram displays the position of vowels in a two-dimensional acoustic space whose coordinates are the first and second formants (**Fig. 5**). Since pitch perception is nonlinear, it is customary to employ logarithmic scales.

Speech perception. Acoustic phonetics makes it possible to describe what is present in the signal, but such analysis does not explain how the discovered

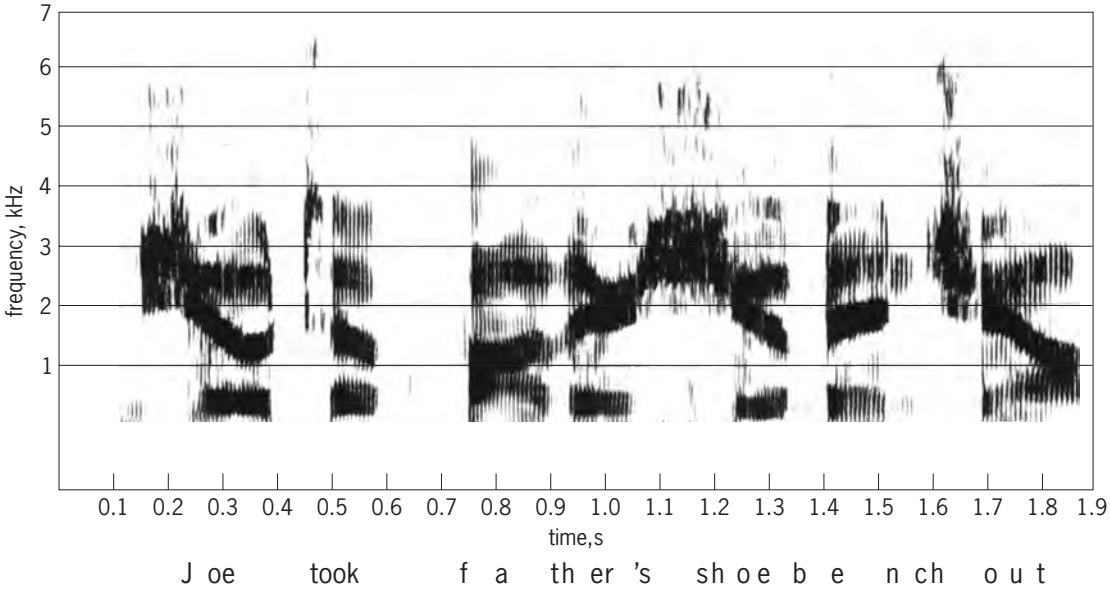

Fig. 4. Broadband spectrogram of the utterance "Joe took father's shoebench out," produced by a male speaker.

patterns are relevant in speech perception. Various speech synthesizers enable the investigator to produce modifications of the signal and study their perceptual significance. *See* VOICE RESPONSE.

Cues for the perception of the manner of articulation are generally present during the most constricted or central segment of the articulated sound. For the vowels, these cues consist of formant positions and the relative intensities of the formants (for most vowels, the first three formants provide sufficient information for perceptual identification). For consonants, the class of plosives is recognized on the basis of the presence of a gap (corresponding to the closure of the stop) followed by a sudden short burst of energy (corresponding to the explosive release). The class of fricatives is recognized by the presence of noise (aperiodic energy). Nasals are characterized by the presence of additional nasal formants and by the wide bandwidth and relatively weak intensity of the first (oral) formant. Laterals have vowellike formants in the lower part of the spectrum and an antiresonance in the region where the third formant would be expected.

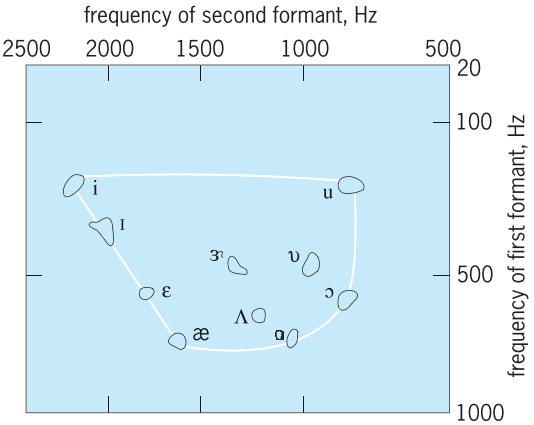

Fig. 5. Acoustical vowel diagram of vowels occurring in the English words *heed, hid, head, had, hod, hawed, hood, who'd, hud, and heard*.

The point of articulation of a consonant is reflected in the influence of the consonant on an adjacent vowel. Particularly significant perceptually is the frequency position of the second formant at the moment of transition between the consonant and the vowel. Sounds produced at the same point of articulation are characterized by the same frequency region at which the second formant starts its movement toward the target position of the vowel, or toward which it moves in the transition from the vowel toward a postvocalic consonant. This spectral region is called the hub or locus of the consonant. The point of articulation may also be deduced from spectral cues, especially in the case of fricatives and plosive bursts.

Prosodic aspects of language. The subject matter of phonetics is not limited to the production and perception of vowels and consonants; of equal importance are such prosodic and suprasegmental aspects of spoken language as duration, fundamental frequency, and intensity. Suprasegmental features relate to segmental features by constituting an overlaid function of inherent features (features that can be defined with reference to the segment itself). The fundamental frequency of an inherently voiced segment, besides characterizing the segment as voiced, may also serve to signal a tonal or intonational pattern. To be recognizable as a segment, every segment has a certain duration in the time domain; at the same time, that duration may be contrastive (for example, may characterize the segment as being distinctively short rather than long). Every segment also has a certain amount of intensity; whatever the acoustic and physiological correlates of stress, they consist of intensifying phonetic factors already present in a lesser degree. Furthermore, suprasegmental features differ from segmental features by the fact that suprasegmental features are established by comparison of items in sequence, whereas segmental features are identifiable by inspection of the segment itself.

Suprasegmental features are also involved in signaling various boundaries, including the boundaries of morphemes, words, phrases, clauses, and sentences, as well as still-higher-level units such as paragraphs in a longer stretch of speech. For example, modification of the time dimension—use of different degrees of preboundary lengthening—may serve to distinguish between sentence boundaries within a paragraph and those sentence boundaries that simultaneously terminate the paragraph.

As is the case with segmental features, phonetic science is concerned with the production of suprasegmental features, with the way in which the suprasegmental features are manifested in the acoustic representation of the signal, and with the ways in which the encoded information is extracted from the signal through the process of speech perception. This brief sketch has omitted a number of other relevant aspects of phonetics, such as the neurophysiological control of the process of speech production and perception, and the specifically linguistic problems relating to the ways in which the phonological systems of different languages are realized in the phonetic domain. Phonetics also has a number of practical applications in areas such as foreign language teaching, speech therapy, and communications engineering. *See* SPEECH. Ilse Lehiste

Bibliography. J. Clark and C. Yallop, *Introduction to Phonetics and Phonology*, 2d ed., 1995; A. Cutler and D. R. Ladd, *Prosody: Models and Measurements,* 1983; P. Ladefoged, *A Course in Phonetics*, 4th ed., 2000; J. Laver, *Principles of Phonetics*, 1994.

Phonolite

A light-colored, aphanitic (not visibly crystalline) rock of volcanic origin, composed largely of alkali feldspar, feldspathoids (nepheline, leucite, sodalite), and smaller amounts of dark-colored (mafic) minerals (biotite, soda amphibole, and soda pyroxene). Phonolite is chemically the effusive equivalent of nepheline syenite and similar rocks. Rocks in which plagioclase (oligoclase or andesine) exceeds alkali feldspar are rare and may be called feldspathoidal latite. *See* FELDSPATHOID.

Rapid cooling at the surface causes lavas to solidify with very fine-grained textures. Most phonolitic lavas, however, carry abundant large crystals (phenocrysts) when they are erupted, and these are soon frozen into the dense matrix to give a porphyritic texture. Generally very little material congeals as glass. The phenocrysts, many visible to the naked eye, include alkali feldspar, feldspathoids, and mafics. These may be well formed (euhedral) or moderately well formed (subhedral). *See* PHENOCRYST.

Most other features of phonolites can be seen only microscopically. The alkali feldspar is principally soda-rich sanidine and orthoclase. It generally occurs in the rock matrix, but if abundant it may also form as phenocrysts. Plagioclase is not abundant except in nepheline latites where it may form abundant phenocrysts.

Nepheline may occur as euhedral crystals (square or hexagonal), some of which may be phenocrysts. Otherwise it is irregular (anhedral) and interstitial. Nosean, hauyne, and sodalite, as euhedral or partly corroded crystals, may occur as phenocrysts and matrix grains. These twelve-sided (dodecahedral) crystals generally show hexagonal outlines in thin sections of the rock. Eight-sided euhedral crystals of pseudoleucite may occur as phenocrysts in potash-rich rocks. More rounded grains of leucite may form part of the matrix. Leucite is commonly altered to pseudoleucite, but the euhedral outline is retained. Analcite occurs principally as matrix material but in some rocks it is abundant and as large euhedral phenocrysts.

Biotite is not common but may form large strongly resorbed phenocrysts. Amphiboles are usually soda-rich (riebeckite, hastingsite, and arfvedsonite). They may occur as phenocrysts or as interstitial clusters. They may show resorption or may be replaced by pyroxene. The most important mafic is soda pyroxene. As phenocrysts it is commonly zoned with cores of diopside surrounded by progressively more sodic shells of aegirine-augite and aegirite. Aegirite is the common pyroxene of the rock matrix.

Accessory minerals are varied and include sphene, magnetite, zircon, and apatite.

The structures and textures of phonolite are similar to those of the more common rock trachyte. Fluidal structure, formed by flowage of solidifying lava and expressed by lines or trails of phenocrysts, may be seen without magnification. Under the microscope, flowage is shown by subparallel arrangement of elongate feldspar crystals. *See* TRACHYTE.

Phonolites are rare and highly variable rocks. They occur as volcanic flows and tuffs and as small intrusive bodies (dikes and sills). They are associated with trachytes and a wide variety of feldspathoidal rocks.

The origin of phonolites and related rocks constitutes an interesting problem. There is still a considerable difference of opinion as to how the phonolitic magma (molten material) originates. One theory assumes an origin from basaltic magma by differentiation. Certain early-formed crystals are removed (perhaps by settling), causing the residual magma to approach the composition of phonolite. Another theory supposes these peculiar magmas to form when a more normal rock melt assimilates large quantities of limestone fragments. Volatiles, notably carbon dioxide, are considered by many to play an important role in transferring and concentrating certain constituents (like potassium) in the magma. The great variety of rock types and modes of association strongly suggests that several different mechanisms may operate to form these feldspathoidal rocks. *See* IGNEOUS ROCKS; MAGMA. Carleton A. Chapman

Phonon

A quantum of vibrational energy in a solid or other elastic medium. This vibrational energy can be transported by elastic waves. The energy content of each wave is quantized. For a wave of frequency f, the energy is $(N + {}^1/_2)hf$, where N is an integer and h is

Planck's constant. Apart from the zero-point energy, $^1/_2hf$, there are N quanta of energy hf. In elastic or lattice waves, these quanta are called phonons. Quantization of energy is not related to the discreteness of the lattice, and also applies to waves in a continuum. *See* QUANTUM MECHANICS; WAVE MOTION.

The concept of phonons closely parallels that of photons, quanta of electromagnetic wave energy. The indirect consequences of quantization were established for phonons just as for photons in the early days of quantum mechanics—for example, the decrease of the specific heat of solids at low temperatures. Direct evidence that the energy of vibrational modes is changed one phonon at a time came much later than that for photons—for example, the photoelectric effect—because phonons exist only within a solid, are subject to strong attenuation and scattering, and have much lower quantum energy than optical or x-ray photons. *See* PHOTOEMISSION; QUANTUM ACOUSTICS; SPECIFIC HEAT OF SOLIDS.

Like photons, phonons can be regarded as particles, each of energy hf and momentum proportional to the wave vector of the elastic or lattice wave. Such a particle can be said to transport energy, thus moving with a velocity equal to the group velocity of the underlying wave. *See* LATTICE VIBRATIONS; PHOTON.

Phonon-electron interactions. Scattering of conduction electrons by lattice vibrations is described by single phonons. An electron absorbs or emits a phonon, changing its own energy by an amount equal to the phonon energy (or its negative, in the case of phonon emission), and changing its own wave vector (which is proportional to its momentum) by an amount equal to the sum of the phonon wave vector (or its negative) and a reciprocal lattice vector, which may equal zero. The change in the electron's energy results from energy conservation, and the change in its wave vector results from constructive interference between the electron wave and the elastic wave. In a discrete lattice, the phonon-electron interaction can be combined with Bragg reflection (involving a nonzero lattice vector), resulting in an Umklapp (flip-over) process. *See* ELECTRON DIFFRACTION; UMKLAPP PROCESS.

At low temperatures, the phonon wave vector is small compared to the nonzero reciprocal lattice vectors, and most processes are said to be normal; that is, the reciprocal lattice vector involved is zero. Thus, the change in the electron wave vector expresses conservation of the combined electron and phonon momentum. The electron-phonon interaction in metals at low temperatures involves scattering of an electron inelastically through a small angle, approximately equal to the ratio of the phonon wave vector to the electron wave vector. Each process is a step in a random walk which gradually changes the direction of the electron. The resulting electrical resistivity varies with temperature in proportion to the product of the square of the temperature and the lattice specific heat (which is itself proportional to the cube of the temperature). The thermal resistivity, however, varies as the lattice specific heat divided by the temperature, because each step changes the electron energy. This makes the thermal resistivity larger than expected from the Wiedemann-Franz law. *See* ELECTRICAL RESISTIVITY; THERMAL CONDUCTION IN SOLIDS; WIEDEMANN-FRANZ LAW.

At low temperatures, when the phonon wave vector is much smaller than the nonzero reciprocal lattice vectors for the important phonons, electron-phonon interactions conserve the combined total momentum of the electron and phonon gases. The net momentum of electrons in an electric current causes net momentum of the phonon gas. This phonon drag sets up phonon heat flow and an addition to the Peltier coefficient. In metals this is of the order of the product of the temperature and the lattice specific heat divided by the electronic charge density. In semiconductors, the specific heat is only that of those phonons which can interact with electrons. At higher temperatures, phonon drag is reduced by other phonon scattering processes and by electron-phonon processes with Bragg scattering (that is, scattering involving a nonzero reciprocal lattice vector). *See* THERMOELECTRICITY.

Other phonon interactions. Magnetic ions and other spin sites change their spin state with the emission or absorption of phonons. At low temperatures, just one phonon is involved, which matches the energy difference of the spin states (a direct process). If all spins have the same level spacing, the phonons are confined to a narrow frequency band, which can sometimes be overpopulated, slowing down the spin-lattice interaction (a phonon bottleneck). *See* PARAMAGNETISM.

Inelastic neutron diffraction also involves single phonons, allowing the phonon frequency to be measured as a function of the phonon wave vector. Phonons form a significant part of thermal excitations in liquid helium, which has also been studied by neutron diffraction. *See* LIQUID HELIUM; NEUTRON DIFFRACTION; SLOW NEUTRON SPECTROSCOPY.

Paul G. Klemens

Bibliography. F. J. Blatt, *Physics of Electronic Conduction in Solids*, 1968; G. Grimvall, *Thermophysical Properties of Materials*, 1986; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996.

Phonoreception

The perception of sound by animals through specialized sense organs. A sense of hearing is possessed by animals belonging to two divisions of the animal kingdom: the vertebrates, which form the main subphylum of the phylum Chordata, and the insects, which make up the most important class of the phylum Arthropoda. The sense is mediated by the ear, a specialized organ for the reception of vibratory stimuli. Such an organ is found in all except the most primitive vertebrates, but only in some of the many species of insects. The vertebrate and insect types of ear differ in evolutionary origin and in their modes of operation, but both have attained high levels of performance in the reception and discrimination of sounds. *See* SOUND.

Vertebrates

The vertebrate ear is a part of the labyrinth, located deep in the bone or cartilage of the head, one ear on either side of the brain. A complex assembly of tubes and chambers contains a membranous structure which bears within it a number of sensory endings of different kinds. *See* EAR (VERTEBRATE).

The membranous labyrinth is shown in a generalized schematic form in **Fig. 1**. It is convenient to recognize two divisions, a superior division, which includes the three semicircular canals and the utricle, and an inferior division, which includes the saccule and its appendages, the lagena and the cochlea. The superior division is remarkably uniform in character from the higher fishes upward, but the inferior division shows many variations. The saccule is always present. The lagena is present in all classes except the mammals, although it is missing in occasional species. The cochlea is found in reptiles, birds, and mammals.

The sensory endings within these parts of the labyrinth also vary in the vertebrate series. Again there is uniformity for the superior division. There is a crista in each ampulla of the three semicircular canals and a utricular macula. In all but the mammals (with a few individual exceptions), there is a macula neglecta, usually located on the floor of the utricle or close to the junction of utricle and saccule. All forms have a saccular macula. All those with a lagena (in general, all except the mammals) have a lagenar macula. All the amphibians have a papilla amphibiorum, but it is found in no other forms. A basilar papilla appears in certain amphibians, is continued in the reptiles, and then is developed in a more elaborate form as the cochlea of higher reptiles, birds, and mammals.

These endings contain ciliated cells (hair cells) which are supplied by fibers of the eighth cranial (auditory) nerve. In the cristae the cilia of the hair cells are particularly long and are embedded in a gelatinous substance that forms a cap or cupola. In the maculae the cilia are surmounted by a flat plate of gelatinous material in which numerous granules of calcium carbonate (otoliths) are usually embedded. The ciliated cells in the papillae lie on a movable membrane (basilar membrane) and have a membranous covering, the tectorial membrane.

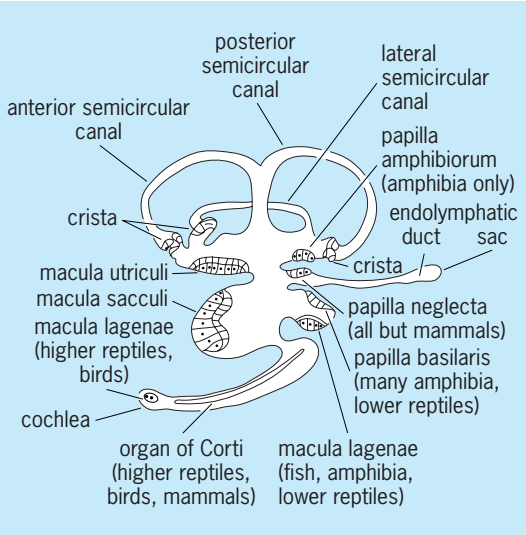

Fig. 1. Generalized sketch of the vertebrate labyrinth. The three cristae, macula ultriculi, and macula sacculi are always present in vertebrates, and the other endings appear as indicated, with a few exceptions.

The superior part of the labyrinth generally serves for bodily posture and equilibrium, whereas the saccule and its appendages (lagena and cochlea) serve for hearing. However, there are exceptions to this rule, the most important of which is that in the higher vertebrates, including mammals and probably birds and reptiles, the saccule serves only for equilibrium.

Beginning with the amphibians, which are the earliest vertebrates to spend a considerable portion of their lives on land, there appears a special mechanism, the middle ear, whose function is the transmission of aerial vibrations to the endings of the inner ear. All the vertebrates above the fishes, and certain of the fishes as well, have some type of sound-facilitative mechanism.

Fishes. The maculae of the utricle, saccule, and lagena in the bony fishes have a peculiar form. Instead of numerous calcareous particles there is a single otolith, a large body of distinctive form. The macula neglecta is sometimes lacking.

Few questions have been more actively debated than the ability of fish to hear. Experiments on this question began with G. Parker in 1903, who observed the natural reactions of fish when exposed to a sudden sound, and were carried forward by F. Westerfield and others in 1922 by the introduction of conditioned-response methods. This work culminated in the series of studies by K. von Frisch and his associates, who trained fish to make feeding responses at the sounding of a tone. These experiments proved that fish may be divided into two groups according to hearing ability, those that hear only crudely and those that hear well. The first group includes the great majority of fish species, with what may be called the basic type of labyrinth and lacking any accessory mechanism. The fish that hear well have one of two general types of sound-facilitating structure: either an air vesicle adjacent to some part of the labyrinth or a connection with the swim bladder.

The second of these types of accessory structure is the more common, and is found in a large group of fresh-water fishes known as the Ostariophysi. Between the labyrinth and the anterior part of the swim bladder is a chain of three or four small bones, known from their discoverer, E. H. Weber, as the Weberian ossicles (**Fig. 2**). Weber correctly supposed, when he described this apparatus in 1820, that it serves for the facilitation of sound reception, for it has been demonstrated that the hearing is impaired after removal of the swim bladder and after an interruption of the ossicular chain.

The most extensive study of hearing in fish was carried out by von Frisch and his students on the minnow (*Phoxinus laevis*), one of the Ostariophysi

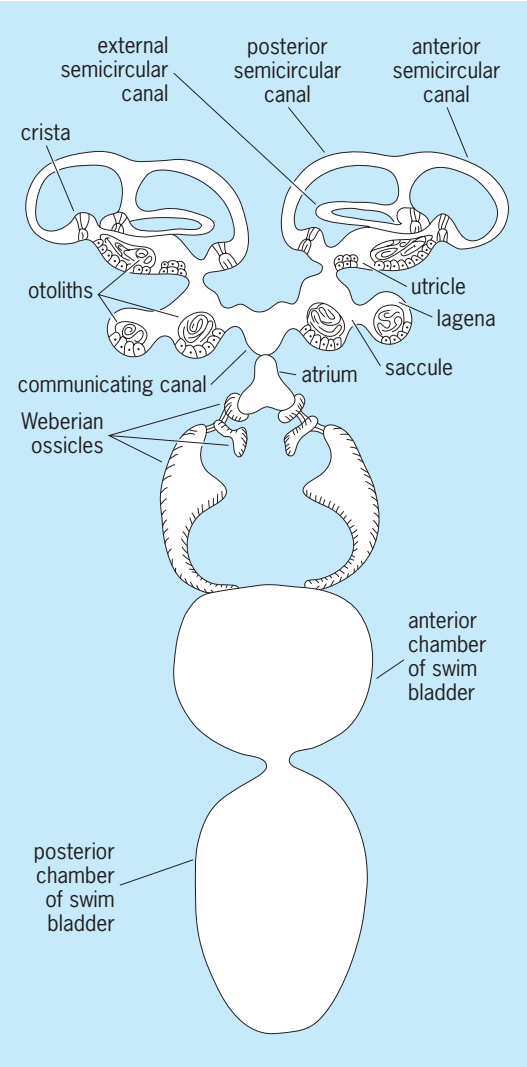

Fig. 2. Diagram (from above) of the two labyrinths of an ostariophysan fish and their connections with the swim bladder through the Weberian ossicles.

that responds readily to training procedures. This fish was reported to respond to tones over a range from about 32 to 5000–6000 Hz, and in the lower part of its range to be able to discriminate a change of frequency of about 3%. The dwarf catfish [*Ictalurus* (*Ameiurus*) *nebulosus*], another ostariophysan, was said to respond to tones as high as 13,000 Hz. Questions have been raised concerning some of the procedures of these experiments, and the tests need to be repeated. More recent studies have not usually indicated upper limits as high as these. Responses only up to 3000 Hz have been obtained in a carefully controlled study of the goldfish (*Carassius auratus*).

In nonostariophysans the observed limits are considerably lower. It was found, in studies of nine species of marine fishes, that the upper limits were around 1000–1200 Hz in eight of them, and there was a limit of 2800 Hz in one, *Holocentrus ascensionis*. Other species have shown even more limited ranges, such as 600 Hz for *Anguilla* and 800 Hz for *Gobius paganellus*.

The mormyrids and labyrinthine fishes, which lack the Weberian ossicles but have an air vesicle adjacent to the labyrinth, have usually shown relatively good hearing, comparable to that of the Ostariophysi.

A number of experiments, mostly on minnows, have dealt with the problem of the particular parts of the fish labyrinth that are concerned with hearing. Removal of the superior portion, which includes the utricle and semicircular canals, does not impair the responses to sound, but seriously affects the posture and swimming ability. After this operation the fish may assume an inverted position, and swims erratically. These parts must therefore contain organs of equilibrium. Removal or impairment of either the saccule alone or the lagena alone leaves the fish able to hear, but the removal of both saccule and lagena abolishes the responses to all tones except those of very low frequency which are perceived through skin or lateral-line receptors. Hence, the endings of the saccule and lagena are auditory in function. Exceptional in this respect are the herring and sardine (clupeids), in which an air vesicle is applied to the wall of the utricle; in these the sense of hearing is probably mediated by the utricular macula.

Amphibians. The three orders of amphibians—the Apoda (legless), including wormlike forms such as caecilians; the Urodela (tailed), including mud puppies, newts, and salamanders; and the Anura (tailless), including frogs and toads—all have some type of middle-ear mechanism.

The first two orders include animals whose ears show a great variety of accessory structures, some of which look as though they might function well in sound reception, whereas others seem crude. Of these only the salamander has been studied experimentally. In 1938, S. Ferhat-Akat trained larvae to come for food at the sounding of a tone, and got results for tones up to 244 Hz in one specimen and up to 218 Hz in three others.

Higher amphibians, such as the frog, possess a well-developed middle-ear mechanism, consisting of a disk of cartilage flush with the lateral surface of the head and covered with skin, and a rod of cartilage and bone, called the columella, leading inward from the disk and expanding to form the stapes, which is embedded in an opening (oval window) in the wall of the otic capsule (**Fig. 3**).

The active and often loud croaking of frogs in the breeding season has focused attention upon the problem of their hearing. R. Yerkes in 1905 first succeeded in obtaining experimental evidence of their auditory sensitivity by showing that sounds may

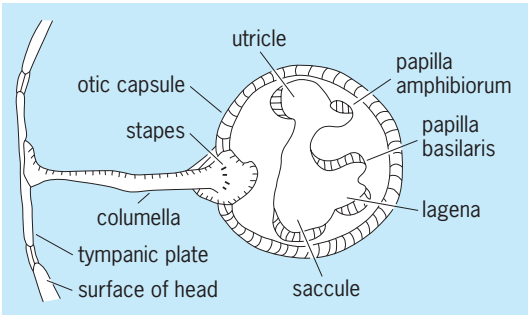

Fig. 3. Simplified diagram of the frog ear.

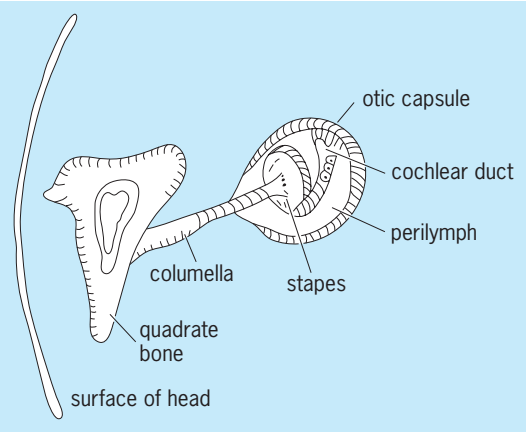

Fig. 4. Diagram of the ear of a snake. The nonauditory parts and endings are not shown.

enhance or inhibit their response to a strong tactual stimulus. Several studies in which the impulses from the eighth cranial nerve were recorded on stimulation with tones showed results only for low frequencies, up to 500–600 Hz, or at most to 1024 Hz. The most extensive study of the electrical responses of the ear by W. Strother in 1959 showed responses in *Rana catesbeiana* over a range from below 100 Hz to about 3500 Hz. The sensitivity was best in two regions, around 500 Hz and around 1500 Hz, and there is reason to believe that these two regions correspond to the actions of the two papillae of the frog, the papilla basilaris and the papilla amphibiorum.

Reptiles. The living reptiles belong to four important groups, represented by snakes, turtles, chameleons and lizards, and crocodiles and alligators.

Many authorities have asserted that snakes are completely deaf, or that their ears are sensitive only to vibrations conducted to the head through the ground. This impression has arisen partly from the fact that snakes do not have any external ear and do not show obvious reactions to sounds. There is no tympanic membrane to receive aerial sound pressures, but its purpose is served by one of the bones of the skull, the quadrate bone, which is loosely attached to the main part of the skull. Although it lies beneath the skin and other tissues of the side of the head, the quadrate bone presents a flat surface for the action of sounds, and communicates them to a thin bony rod (columella) running inward to expand as the stapes in the oval window (**Fig. 4**).

Experiments have shown that electrical potentials are produced in the inner ears of snakes in response to low-frequency sounds, to both the sounds conducted through the substratum and those conducted through the air in the usual way. Hence, it is safe to conclude that snakes have hearing, although only for the lower range of sounds and not as highly sensitive as that of most other animals.

Doubt has often also been expressed about the ability of turtles to hear, but here again the evidence is that they do. They have a well-developed middle ear, including a cartilaginous disk on the side of the head beneath the skin, and a columella leading to a stapes in the oval window of the otic capsule. Two investigators have succeeded in training turtles to make positive reactions to an acoustic signal, although others have failed in this attempt. The electrophysiological method yields positive results. Indeed, the observations show that for low tones, those of 100–700 Hz, the turtles have excellent sensitivity. The wood turtle (*Clemmys insculpta*) exceeds other species studied (**Fig. 5**).

Structurally the ear of the lizard is superior to that of the turtle. With a few exceptions, there is a membranous drum, an extracolumella, and a columella whose expanded end constitutes a stapes in the oval window. C. Berger was able to train two species of lizards (*Lacerta agilis* and *L. vivipara*) to make feeding movements in response to various sounds, including tones over a range of 69–8200 Hz. Electrical potential studies carried out in a great many species indicate considerable species variations in sensitivity. In general, the geckos, most of which are nocturnal, have the best hearing. The frequency range as represented by the cochlear potentials ranges at least 100–10,000 Hz, though these ears seem especially susceptible to damage by the uppermost frequencies. The structural development of the cochlea varies widely in the different species; thus the iguanids in general have a short and undifferentiated basilar membrane with about 60 hair cells,

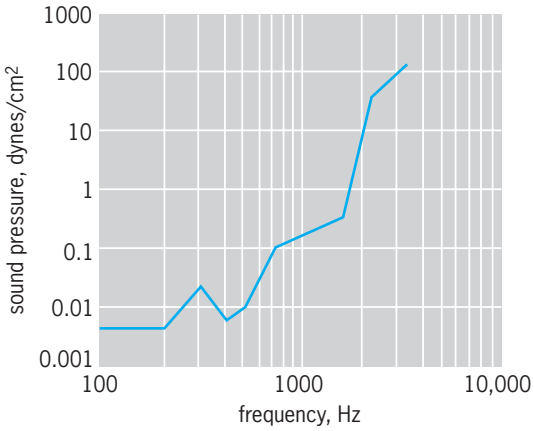

Fig. 5. Auditory sensitivity of a wood turtle (*Clemmys insculpta*), as shown by the potentials produced in its ear by sounds. The curve shows the sound pressure necessary to produce a potential of 0.3 microvolt.

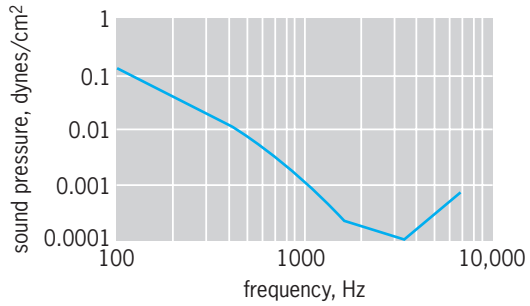

Fig. 6. Graph showing threshold sensitivity in the bullfinch (mean of four birds). (*After J. Schwartzkopff, Uber Sitz und Leistung von Gehöor und Vibrationssinn bei Vögeln, Z. vergl. Physiol., 31:527–608, 1949*)

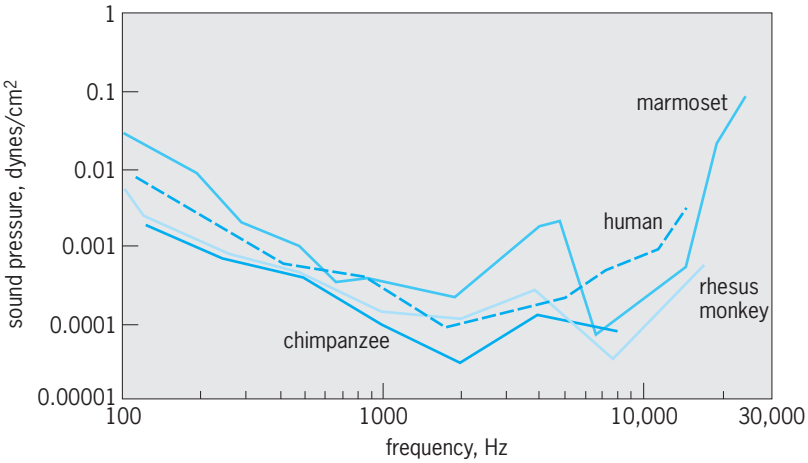

Fig. 7. Auditory thresholds in primates. The curves show the sound pressures that are barely audible for humans, chimpanzees, rhesus monkeys, and marmosets.

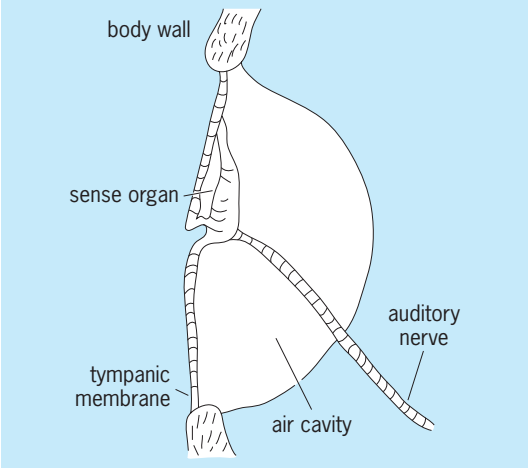

Fig. 8. Diagram of the ear of a grasshopper.

whereas the gekkonids have a long and tapered basilar membrane with several hundred hair cells (up to 1600 in *Gekko gecko*).

Certain chamelons (*Chamaeleo senegalensis* and *C. quilensis*) have a peculiar auditory structure somewhat similar to that of the snake in that there is no external ear opening or eardrum; but a plate of bone (a part of the pterygoid bone) lying beneath the skin and muscle layers at the side of the head serves as a receptive surface for sounds and transmits the vibrations through the columella to the inner ear.

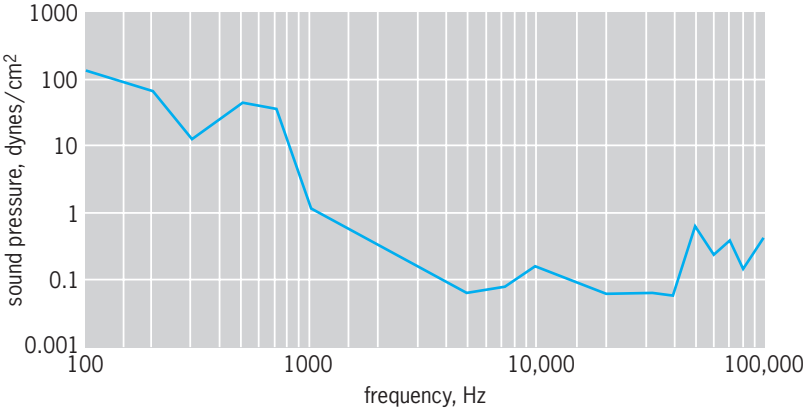

Fig. 9. Auditory thresholds of the katydid (*Conocephalus strictus*). Sensitivity is greatest in the 7000–60,000-Hz region. (*After E. G. Wever and J. A. Vernon, The auditory sensitivity of Orthoptera, Proc. Nat. Acad. Sci., 45:413–419, 1959*)

The crocodiles have an ear representing a distinct advance over other reptiles, and there is a curved cochlea similar to that of birds. Probably they have excellent hearing, although the experimental evidence is scanty as yet. Many general observations indicate that they use sounds in mating activities, and the males are capable of producing loud roars. F. Beach was able to provoke captive animals into roaring and making movements by stimulation with low tones, especially sounds at 57 Hz, but also others at about 300 Hz. E. Wever and J. Vernon found that young caimans gave cochlear potentials in response to sounds over a range of 20–6000 Hz.

Birds. The labyrinth in birds is generally similar to that of the higher reptiles. There is a membranous eardrum, a columella leading inward to the stapes, and a curved cochlea. There are only minor variations in the forms of these structures among the various species.

Most birds have a range of hearing of about 50–20,000 Hz, and an absolute sensitivity that probably approaches that of humans in the medium-high-tone range. The threshold sensitivity of a finch (*Pyrrhula p. minor*) is shown in **Fig. 6**, as determined by J. Schwartzkopff by a training method. In general, the small songbirds are more sensitive than the larger birds, such as chickens and pigeons. The owl is exceptional in having an ear with a particularly large drum membrane and other special features that make for high sensitivity. Thus, an owl in a tree at dusk can hear and locate a mouse rustling in the grass below.

Pitch discrimination in songbirds and parrots is about as keen as that of humans (0.3–0.7%), but in pigeons it is relatively poor (6%).

Mammals. The auditory apparatus attains its highest development in the mammals. The columella of lower forms has been replaced by a chain of three ossicles, which connect the tympanic membrane with the inner ear. In the egg-laying mammals, such as the platypus, the cochlea is a curved tube as in crocodiles and birds, but in all other mammals it is a spiral of one to four turns. The great extension of the cochlea and the corresponding multiplication of sensory cells have enhanced the capacities of the mammals to deal with the varieties and complexities of sounds.

Despite intense interest in mammalian hearing, precise information is available on only a few mammals apart from humans. Experimental studies have been carried out on some of the subhuman primates and on a few of the common laboratory animals. Only fragmentary information is available on the many other species of mammals, although it can be assumed from their general behavior, because they seem to respond to much the same range and intensities of sounds that humans do, that their hearing is similar to humans'. *See* HEARING (HUMAN).

Training experiments are easily carried out on the subhuman primates; **Fig. 7** presents threshold

curves for the chimpanzee, the rhesus monkey, and a species of marmoset, with the human curve for comparison. It will be noted that these animals have auditory sensitivity similar to man's in the low-tone range, but are superior to man in the high-tone range.

The most extensive studies have been made on the cat. Its sensitivity also is similar to man's over the lower range, but extends far above the human limit to 60,000 Hz or more. Electrical potentials have been recorded from the cat's cochlea to tones as high as 100,000 Hz.

Other animals whose hearing has been studied experimentally are the dog, rat, guinea pig, rabbit, bat, and certain species of mice.

The bats are of special interest, more particularly the small insectivorous species, because they repeatedly produce vocal sounds of high frequency, up to 40,000 Hz or more, as they fly about in search of insect prey. They locate the prey by hearing the echoes of their cries, and they guide themselves in the total darkness of caves by echoes from the walls. The hearing of bats extends far into the high frequencies, up to 100,000 Hz at least. *See* AUDIOMETRY.

Invertebrates

The group of invertebrates which has received the most attention has been the insects. Other arthropods, such as certain crustaceans and spiders, have also been found to be sensitive to sound waves.

The insect ear consists of a superficial membrane of thin chitin with an associated group of sensilla called scolophores. Such an apparatus is shown in simplified form in **Fig. 8**. These ears are found in most species of katydids, crickets, grasshoppers, cicadas, waterboatmen, mosquitoes, and nocturnal and spinner moths. They occur in different places in the body: on the antennae of mosquitoes, on the forelegs of katydids and crickets, on the metathorax of cicadas and waterboatmen, and on the abdomen of grasshoppers. Probably these differently situated organs represent separate evolutionary developments, through the association of a thinned-out region of the body wall with sensilla that are found extensively in the bodies of insects and that by themselves seem to serve for movement perception.

The insects mentioned above are noted for their production of stridulatory sounds made by rubbing the edges of the wings together, or a leg against a wing, or by other means. These sounds are produced by the males and serve for enticing the females in mating.

The sensitivity of insect ears is keenest in the high frequencies. **Figure 9** shows threshold curves obtained on a katydid by observing the potentials produced in the auditory nerve during stimulation with sounds. As will be seen, the sensitivity in this species is greatest in the region of 7000–60,000 Hz. It extends to even higher frequencies, usually as high as 120,000 Hz and sometimes beyond. Other species have distinctly different sensitivity curves, and there is reason to believe that there is a relation to the range of the stridulatory sounds. There is evidence that these sounds are discriminated, for the females of one species respond to the males of their own kind. J. Regen placed a cage of chirping male crickets in a field, and found that females of their species sought them out. This seeking activity ceased, however, when the males were deprived of their stridulating organs or the females were made deaf by removal of their ears. A most striking adaptation is that shown by mosquitoes: The ear of the male mosquito is sensitive only to a narrow range of frequencies around 380 Hz, and this frequency is the one which is produced by the wings of the female in flight. M. Tischner found that when the ear of the male mosquito was made nonfunctional, the mosquito failed to find a mate.

Ernest G. Wever

Bibliography. C. R. Pfaltz (ed.), *New Aspects of Cochlear Mechanics and Inner Ear Pathophysiology*, 1990; A. N. Popper and R. R. Fay (eds.), *Comparative Hearing: Mammals*, 1994; A. N. Popper and R. R. Fay (eds.), *Comparative Studies of Hearing in Invertebrates*, 1980.

Phoresy

A relationship between two different species of organisms in which the larger, or host, organism transports a smaller organism, the guest. It is regarded as a type of commensalism in which the relationship is limited to transportation of the guest. The term is credited to P. Lesne following his observations on the biology of a small fly, *Limosina sacra*, which is transported by a scarabeid, one of the dung beetles, into its burrow. These burrows are suitable breeding sites for both animals. *See* POPULATION ECOLOGY.

Charles B. Curtin

Phoronida

A small phylum of marine animals; in the recent past, they have been grouped with Brachiopoda and Bryozoa into the Lophophorata. The Phoronida phylum or class has no intermediate hierarchical level until the generic level. Two genera, *Phoronis* and *Phoronopsis*, are recognized with respectively seven and three well-defined species.

Habitat and distribution. Phoronids may occur in vertical tubes embedded in soft sediments (sand, mud, or fine gravel) or form tangled masses of many individuals, buried in or encrusting limestone rocks and shells of dead mollusks. *Phoronis australis* is embedded into the tube of cerianthid anemones. Phoronids secrete characteristic rigid tubes consisting of layers of chitin to which particles and debris adhere. Phoronids are found in all oceans (except the polar seas) at depths ranging from the intertidal zone to about 400 m (1300 ft).

Morphology. The body is elongate, ranging in length from about 0.1 to 18 in. (2.5 mm to more than 45 cm), and bears a terminal, bilaterally symmetrical crown of tentacles (the lophophore) that surrounds

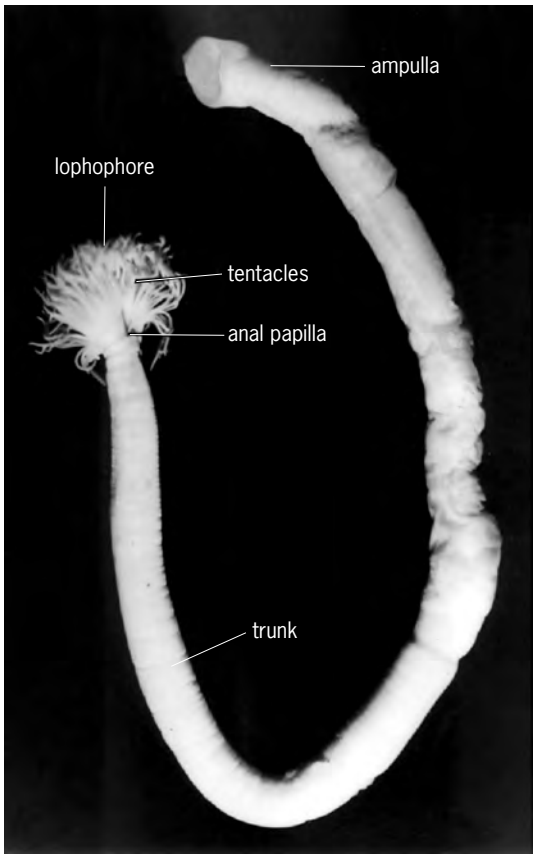

Fig. 1. ***Phoronopsis harmeri*** **removed from its tube. Length is about 8 in. (20 cm).**

the mouth (**Fig. 1**). The anus occurs at the level of the mouth and is borne on a papilla immediately outside the lophophore. The digestive tract is therefore U-shaped, with the mouth and anus opening close together at one end.

The lophophore may be oval to horseshoe shaped with two spiral to helical coils; the complexity of the lophophore is generally proportional to the increase in body size. The main functions are feeding, respiration, and protection. The tentacles vary in number from 11 to over 1500 and have complex arrays of cilia that create a current which carries food particles to the mouth and absorbs oxygen for respiration. Associated with the mouth is a ciliated flap of tissue known as the epistome that occurs along the inner row of tentacles. *See* LOPHOPHORE.

The digestive tract consists of a short esophagus, a long prestomach, and a stomach surrounded by a blood plexus. A muscular pylorus separates the stomach and intestine, the latter being a long slender ascending tube, which ends at the anus. The junction of the stomach and intestine occurs in the bulblike posterior or aboral extremity of the animal called the ampulla that anchors the body in the rear end of the tube. The food seems to consist chiefly of microscopic plankton such as diatoms, flagellates, peridinians, small larvae, and detritus. *See* FEEDING MECHANISMS (INVERTEBRATE).

There is a blood vascular system in which nucleated red blood corpuscles containing hemoglobin circulate. The closed vascular system consists basically of two longitudinal vessels, known as the lateral (afferent) and median (efferent) vessels, which are continuous with one another at the proximal end of the body and flow through the stomach blood plexus, a blood sinus that surrounds the gut and develops a large number of blind blood ceca. Distally, the vessels connect with two semicircular lophophoral vessels, immediately below the tentacles. Within each tentacle is a single blind vessel which branches into two at its base and so connects with both semicircular vessels. In the living animal, corpuscles can be seen pulsating up and down in the tentacular and longitudinal vessels.

The body is divided into two parts in larval and adult forms, with each part containing its own coelomic cavity: the mesosome, represented by the lophophore, and its cavity, the mesocoelom; and the trunk or metasome and its cavity, the metacoelom. Both body parts are separated by a horizontal mesentery named a diaphragm. The metacoelom is subdivided by a series of longitudinal mesenteries extending from the digestive tract to the body wall. In most phoronids, there are four such mesenteries occupying oral, anal, right lateral, and left lateral positions, thus dividing the metacoelomic cavity into four chambers (**Fig. 2**).

The paired excretory organs are metanephridia, located in the trunk on either side of the intestine at the level of attachment of the lateral mesenteries. They open into the metacoelom by one or two funnels and discharge their contents to the exterior via a nephridiopore, located on the anal papilla (or sometimes on a nephridial ridge) on each side of the anus. The morphological characteristics of the nephridia are of prime taxonomic importance to identify a phoronid species. The nephridia also act as gonoducts.

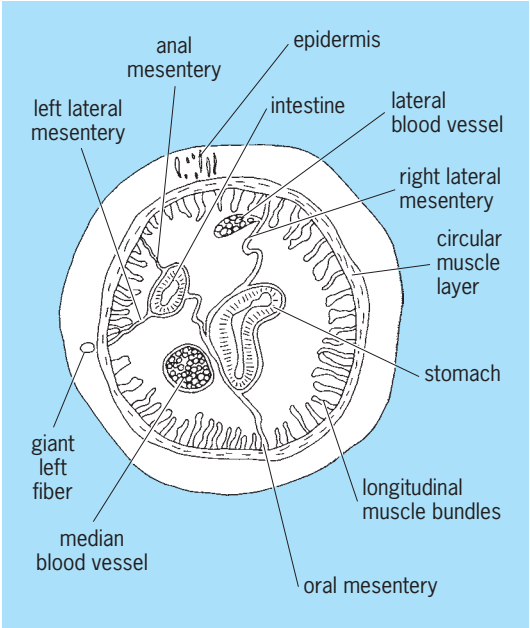

Fig. 2. Cross section through ***Phoronopsis harmeri*****.**

The body wall consists of an outer layer of epithelial cells, many of them secretory and concerned with the building of the tube, and two layers of muscle. The outermost layer of muscle consists of circular fibers, and inside this is a series of bundles of longitudinal fibers (Fig. 2). A nerve plexus underlies the epidermis and continues as a more dense concentration in the form of a nerve ring at the level of the diaphragm. Extending proximally from nerve cells in this ring are either one on the left or two giant nerve fibers which taper and disappear at the proximal end of the ampulla. The giant nerve fibers are known in all species, but often are lacking in *Phoronis ovalis*, and are probably concerned with the rapid retraction of the body into the tube.

Reproduction. Phoronids include both dioecious animals and hermaphrodites. Phoronids reproduce asexually by transverse fission or sexually by internal fertilization. Reproductive tissue is formed from cells that multiply first on the thin walls of the blood ceca housed mainly in the ampulla. When ripe, the gametes are shed into the body cavity and find their way to the nephridia that serve as gonoducts. After internal fertilization, egg cleavage is total, equal, and typically radial. Three types of developmental patterns occur: (1) brooding in the parental tube, (2) brooding on nidamental glands in the lophophore until the larval stage is reached, (3) shedding fertilized embryos into the sea water that grow into a characteristic ciliated, free-swimming pelagic larva, the *Actinotrocha* or actinotroch. The actinotroch consists of an anterior preoral lobe with a nervous ganglion on the apical area, a tentacular ridge, a pair of protonephridia, and posteriorly a ciliated ring around the anus (**Fig. 3**). The larvae undergo planktotrophic development during 2–3 weeks before settling. The metamorphosis from the actinotroch larva to a slender young phoronid is "catastrophic," occurring in less than 30 minutes.

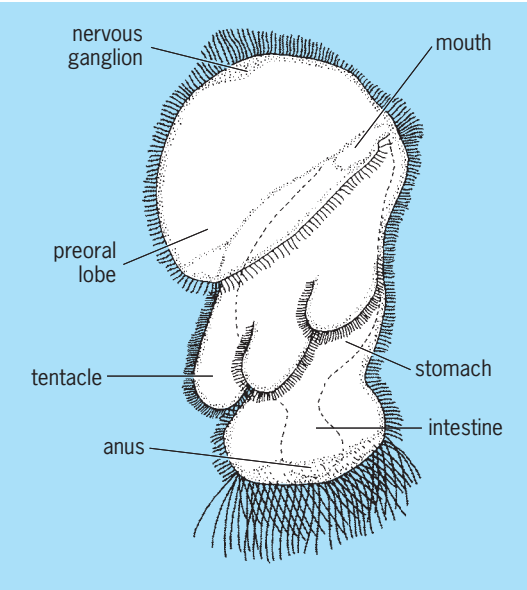

Fig. 3. ***Actinotrocha vancouverensis*, the larva of *Phoronis ijimai*.**

Christian C. Emig; Joan R. Marsden

Bibliography. C. C. Emig, The biology of Phoronida, *Adv. Mar. Biol.*, 19:1–89, 1982; C. C. Emig, British and other Phoronids, in D. M. Kermack and R. S. K. Barnes (eds.), *Synopses of the British Fauna*, no. 13, Academic Press, London, 1979; C. C. Emig, Phoronida, in S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, vol. 2, McGraw-Hill, New York, 1982; C. C. Emig, Phylogenetic systematics in Phoronida (Lophophorata), *Z. Zool. System. Evolut.-forsch.*, 23(3):184–193, 1985; L. H. Hyman, Phoronidea, in *The Invertebrates*, vol. 5: *Smaller Coelomate Groups*, McGraw-Hill, New York, 1959; J. R. Marsden, Phoronidea from the Pacific coast of North America, *Can. J. Zool.*, 37(2):87–111, 1959.

Phosphatase

One of a large group of enzymes, having various chemical structures, that catalyze the hydrolysis of phosphoric acid esters or anhydrides. The phosphatases are a complex group, some of them having highly specific functions. In general, phosphatases promote the degradation of organic phosphorus compounds into inorganic phosphate and an organic moiety (see **illustration**). In addition, phosphatases readily bind phosphates and transport them to cells. *See* ENZYME; PHOSPHORUS.

Phosphatases may be intracellular or extracellular. Intracellular phosphatases are found in blood plasma, cartilage, intestinal mucosa, spleen, liver, kidney, and plant tissues; they are found in the cytoplasm or periplasm, or attached to some cells. Intracellular phosphatase production of inorganic phosphates from organic phosphates is crucial to cellular metabolism and the growth and development of organisms. Soil and water contain a large amount of extracellular phosphatases, in either free or immobilized forms, which are important in the biogeochemical cycling of phosphorus. *See* BIOGEOCHEMISTRY; CELL (BIOLOGY); PHOSPHORUS CYCLE.

Reaction factors. Phosphatases are adaptive enzymes, so their production is stimulated by a deficiency in inorganic phosphorus. In contrast, the presence of inorganic phosphates inhibits both phosphatase activity and synthesis in various tissues. In soils and aquatic systems, production of extracellular phosphatases is stimulated by the presence of organic matter containing phosphorus.

Phosphatases require certain metal ions, called activators, to catalyze reactions. During the reaction process of a phosphatase and a substrate (for example, phosphate ester), the metal ion forms a metal–substrate complex. Ions of calcium, cobalt, copper, magnesium, manganese, molybdenum, sodium, potassium, and zinc are known to participate in phosphatase reactions.

The ionic state of phosphatases and therefore their reactivity are also influenced by the effect of pH directly on functional groups of the active site and

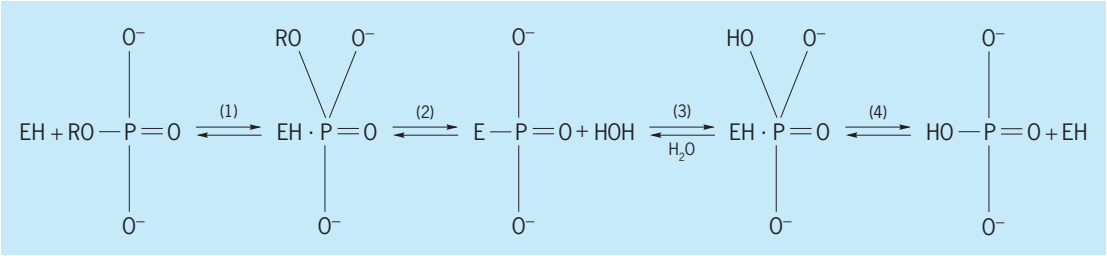

Reaction mechanism for the phosphatase-catalyzed hydrolysis of a phosphoric acid ester. (1) Noncovalent binding of the phosphatase (EH) to the organic substrate (RPO_4^{2-}). (2) Release of an alcohol from the complex and covalent binding of inorganic phosphorus (orthophosphate) to phosphatase, forming a phosphoryl–enzyme compound. (3) Conversion of the phosphoryl—enzyme compound to a noncovalent complex through the uptake of a water molecule. (4) Production of a phosphate ion and regeneration of phosphatase for further catalysis of other organic molecules. R = organic substituent.

by the physical and chemical state of the enzyme–substrate and enzyme–product complexes. However, immobilized phosphatases are more resistant to changes in pH, temperature extremes, inhibitors, and proteolytic attack (degradation inside cells) than their free counterparts are. Clay particles and humic colloids can immobilize free phosphatases in soils and sediments. *See* CLAY; HUMUS.

Types. Different phosphatases typically have maximum hydrolyzing capacities at different pH values. However, all phosphatases are able to tolerate a range of pH values. *See* PH.

Alkaline phosphatase. Alkaline phosphatases have higher catalytic activity above pH 7, with optimum activity well above pH 7. They often require metal ions for their activities and are inhibited by metal-complexing agents such as ethylenediaminetetraacetic acid (EDTA). In humans, large increases in alkaline phosphatase levels are indicators of pathological conditions such as obstructive jaundice, rickets, and bone disease. *See* MALNUTRITION; SKELETAL SYSTEM DISORDERS.

Acid phosphatase. Acid phosphatases have higher catalytic activity below pH 7, with optimum activity around pH 4–6. For hydrolysis of phosphoric acid esters, it is essential for acid phosphatases to have an oxygen linkage. Metal-complexing agents do not inhibit acid phosphatases. However, nonspecific anions (such as fluoride) inhibit their activity.

Extracellular phosphatase. Some plants may take up some forms of organic phosphorus directly from soil. However, most organic phosphorus must first undergo enzymatic hydrolysis to inorganic phosphate in the form of $H_2PO_4^-$ and HPO_4^{2-} ions. The activity of extracellular phosphatases is maximum in plant root zones. In most studies, positive correlation has been found between phosphatase activity and organic phosphorus content in soil and aquatic systems. Bacteria, fungi, yeast, soil invertebrates, mycorrhizae, and all plants produce extracellular phosphatases. *See* BACTERIA; FUNGI; MYCORRHIZAE; RHIZOSPHERE; ROOT (BOTANY); SOIL. H. K. Pant

Bibliography. T. E. Barman, *Enzyme Handbook II*, Springer-Verlag, New York, 1985; R. R. Burns, Interactions of enzymes with soil mineral and organic colloids, in P. C. Hang and M. Schnitzel (ed.), Interactions of soil minerals with natural organic and microbes, *Spec. Pub. Soil Sci. Soc. Amer.*, no. 17, 1986; V. P. Hollander, Acid phosphatases, in P. C. Boyer (ed.), *The Enzymes I*, Academic Press, New York, 1971; R. B. McComb, G. N. Bowers, and S. Posen, *Alkaline Phosphatases*, Plenum Press, New York, 1979; G. Nowak and J. Nowak, Prolonged humification of ^{14}C-labelled oat residues in extensively fertilized soils, *Soil Biol. Biochem.*, 22:1089–1109, 1990; H. K. Pant and P. R. Warman, Enzymatic hydrolysis of soil organic phosphorus by immobilized phosphatases, *Biol. Fertil. Soils*, 30:306–311, 2000; J. C. Tarfdar and N. Claassen, Organic phosphorus compounds as a phosphorus source for higher plants through the activity of phosphatase produced by plant roots and microorganisms, *Biol. Fertil. Soils*, 5:308–312, 1988.

Phosphate metabolism

Organic phosphate compounds are present in the structural units of every animal cell, and inorganic phosphate is associated with calcium in bone and teeth. The total phosphorus in the adult human body is about 12 g/kg, with only 1.4 g/kg present in the soft tissues and the remainder in mineralized tissue in the form of apatite crystals. Blood phosphate plays an important role in regulating neutrality, and it is in equilibrium with both bone and cellular organic phosphates. In plasma it is primarily orthophosphate with HPO_4^{2-} and $H_2PO_4^-$ present in a ratio of about 4:1. The blood level is held relatively constant by regulating phosphate excretion by the kidney. This control is primarily mediated by action of parathyroid hormone. Vitamin D enhances the entry of phosphate into bone. Phosphate plays an important role in absorption of sugars from the intestine and reabsorption of glucose from the kidney. *See* PARATHYROID HORMONE; VITAMIN D.

The central role of phosphates in life processes is indicated by their occurrence in ribonucleic acid (RNA) and deoxyribonucleic acid (DNA), which are important in protein synthesis and in the function of chromosomes in the processes of growth and heredity. Through the formation of lecithins, phosphates are involved in fat metabolism. Phosphates play a major role in the conservation and transfer of energy, particularly of the energy produced in the tricarboxylic acid cycle (Krebs cycle), in glycolysis,

and in the pentose shunt. They do so by participating in many phosphorylation and transphosphorylation reactions involving sugars and other organic compounds. *See* CARBOHYDRATE METABOLISM; CHROMOSOME; CITRIC ACID CYCLE; LIPID METABOLISM; NUCLEIC ACID.

In phosphorylation reactions, compounds such as phosphocreatine (PC) and adenosine triphosphate (ATP) are formed which are capable of yielding relatively large amounts of free energy (5000–11,000 cal or 21–46 kilojoules per mole) when the phosphate bonds are broken by hydrolysis. ATP and PC have a central role in energy storage and transfer in all tissues. The transfer of phosphate from one organic molecule to another is mediated by enzymes without liberation of inorganic phosphate. *See* ADENOSINE TRIPHOSPHATE (ATP).

Phosphorus-containing coenzyme systems include the pyridine (nicotinamide) and the riboflavin nucleotide systems concerned with oxidation-reduction reactions; coenzyme A, the functional form of pantothenic acid, concerned with transacetylation, acylation, and condensation reactions; the diphosphothiamine system concerned with decarboxylation; and pyridoxal phosphate concerned with transamination. *See* BIOCHEMISTRY; BIOLOGICAL OXIDATION; COENZYME; ENERGY METABOLISM.

Morton K. Schwartz

Bibliography. C. L. Comar and F. Bronner (eds.), *Mineral Metabolism*, vol. 2, pt. A, 1964; L. Stryer, *Biochemistry*, 4th ed., 1995.

Phosphate minerals

Any naturally occurring inorganic salts of phosphoric acid, $H_3[PO_4]$. All known phosphate minerals are orthophosphates since their anionic group is the insular tetrahedral unit $[PO_4]^{3-}$. Mineral salts of arsenic acid, $H_3[AsO_4]$, have a similar crystal chemistry. There are over 150 species of phosphate minerals, and their crystal chemistry is often very complicated. Phosphate mineral paragenesis can be divided into three categories: primary phosphates (crystallized directly from a melt or fluid), secondary phosphates (derived from the primary phosphates by hydrothermal activity), and rock phosphates (derived from the action of water upon buried bone material, skeletons of small organisms, and so forth).

Primary phosphates. These phosphate minerals have crystallized from fluids, usually aqueous, during the late stage of fluid and minor element segregation and concentration. Certain rocks, particularly granite pegmatites, frequently contain phosphate minerals which sometimes occur as enormous crystals. Giant crystals up to 10 ft (3 m) across are known for apatite, $Ca_5(F,Cl,OH)[PO_4]_3$; triphylite-lithiophilite, $Li(Fe,Mn)[PO_4]$; amblygonite, $(Li,Na)Al(F,OH)[PO_4]$; and graftonite, $(Fe,Mn,Ca)_3[PO_4]_2$. Rare-earth phosphates, such as monazite, $(La,Ce)[PO_4]$, and xenotime, $Y[PO_4]$, are known from some pegmatites and are mined for rare-earth oxides. The $[PO_4]^{3-}$ anionic group is often considered a "mineralizer"; that is, it contributes to decreased viscosity of the fluid and promotes the growth of large crystals. Primary phosphates are segregated since they usually do not combine with silicate-rich phases to form phosphosilicates. *See* AMBLYGONITE; APATITE; PEGMATITE.

Primary phosphates also occur with other rock types, notably ultrabasic rocks such as nepheline syenites, jacupirangites, and carbonatites. Enormous quantities of apatite occur with nepheline syenite in the Kola Peninsula, Russia. Carbonatites often contain rare-earth phosphates such as britholite, $(Na,Ce,Ca)_5(OH)[(P,Si)O_4]_3$. Primary phosphates also occur in metamorphosed limestones, usually as apatite. Apatite also occurs with magnetite, $FeFe_2O_4$, segregated from basic rocks such as norites and anorthosites. Rarely, phosphates occur in meteorites, including apatite, whitlockite, β-$Ca_3[PO_4]_2$, and sarcopside, $(Mn,Fe,Ca)_3[PO_4]_2$. *See* CARBONATITE.

Secondary phosphates. A large spectrum of secondary phosphates is known, particularly because they have formed at low temperatures and over a range of pH and pO_2 conditions. Over 50 species are known to have been derived from the action of water on triphylite-lithiophilite. Their crystal chemistry is very complicated; most contain Fe^{2+}, Fe^{3+}, Mn^{2+}, and Mn^{3+}, octahedrally coordinated by $(OH)^-$, (H_2O), and $[PO_4]^{3-}$ ligand groups, and can be considered coordination complexes. The most common species are strengite, $Fe[PO_4](H_2O)_2$; ludlamite, $Fe_3[PO_4]_2(H_2O)_4$; and vivianite, $Fe_3[PO_4]_2(H_2O)_8$. Because they contain transition-metal cations, they are often beautifully colored and are highly prized by mineral fanciers. *See* VIVIANITE.

Members of another group, known as the fibrous ferric phosphates, arise by reaction of phosphatic water with goethite, α-FeO(OH), and occur in certain limonite beds, sometimes in generous quantities. These include rockbridgeite, $Fe^{2+}Fe_4^{3+}(OH)_5[PO_4]_3$, and dufrenite, $Fe^{2+}Fe_5^{3+}(OH)_5[PO_4]_4(H_2O)_2$. There are several other species, but this group of interesting minerals is poorly understood. *See* GOETHITE; LIMONITE.

Rock phosphates. These are phosphates that are derived at very low temperatures by the action of water on buried organic material rich in phosphorus, such as bones, shells, and diatoms. Also included are phosphates obtained by the reaction of phosphatic waters with carbonate materials, such as corals. These waters were enriched in the phosphate anion by previous action upon organic material.

Large beds of phosphatic oolites are known. The mineralogy is poorly understood because of the small grain size. The major mineral is carbonate apatite, but minor amounts of monetite, $CaH[PO_4]$, and brushite, $CaH[PO_4](H_2O)_2$, are also encountered, as well as many gel-like substances which are not strict mineral entities.

Crystals of several hydrated magnesium phosphates have been recovered from bat guanos which were suitably preserved. Common minerals include newberyite, $MgH[PO_4](H_2O)_3$, and struvite,

$NH_4Mg\text{-}[PO_4] \cdot 6H_2O$. Guanos are often marketed for fertilizer because of their high phosphorus content. *See* FERTILIZER.

Phosphate minerals of this category are also known to occur as urinary calculi. Most frequent is apatite. Other species are struvite, newberyite, and whitlockite. Struvite crystals have also been found in canned sardines. *See* MINERAL. Paul B. Moore

Phospholipid

A lipid that contains one or more phosphate groups. Like fatty acids, phospholipids are amphipathic in nature; that is, each molecule consists of a hydrophilic (having strong affinity for water) portion and a hydrophobic (lacking affinity for water) portion. Due to the amphipathic nature and insolubility in water, phospholipids are ideal compounds for forming biological membranes. Phospholipids are present in plasma lipoproteins, egg yolk, soya bean, and animal tissues such as brain. Commercially, phospholipids are used to make chocolates, caramels, and many other prepared foods. They are important components in cosmetics and are also used as solubilizing and texturizing agents in pharmaceutical preparations. Medically, phospholipids are used to form therapeutic liposomes—microscopic spheres enclosed by phospholipid membranes and containing specific drugs for delivery into the body. The phospholipid sphere fuses with the plasma membrane of the target cell, and the drug inside the liposome is emptied into the cell. Some liposomes are preferentially recognized by certain organs in the body, thus facilitating the targeted delivery of drugs. *See* LIPID; LIPOSOMES.

Classes. There are two classes of phospholipids: those that have a glycerol backbone and those that contain sphingosine (synthesized from serine and palmitic acid). Both classes are present in the biological membrane.

Phosphoglycerides. Phospholipids that contain a glycerol backbone are called phosphoglycerides (or glycerophospholipids), the most abundant class of phospholipids found in nature. The simplest form of phosphoglyceride is phosphatidic acid (structure **1**).

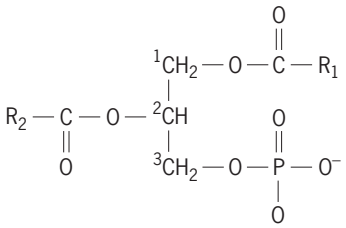

(**1**)

It contains a glycerol backbone with fatty acid esters attached to the *sn*-1 and *sn*-2 positions of the glycerol molecule, and a phosphate group attached at the *sn*-3 position. Other phosphoglycerides share a similar structure, but with appropriate head groups esterified to the phosphate moiety in phosphatidic acid. For example, choline is esterified to the phosphate group in phosphatidic acid to form phosphatidylcholine (structure **2**)

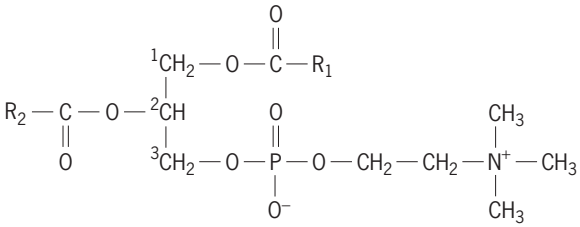

(**2**)

and ethanolamine is esterified to the phosphate group to form phosphatidylethanolamine. Phosphatidylcholine (lecithin) is the most abundant type of phosphoglyceride in eukaryotes, whereas phosphatidylethanolamine is the major phosphoglyceride found in bacteria. Other naturally occuring phosphoglycerides include phosphatidylserine, phosphatidylinositol, phosphatidylglycerol, and cardiolipin. The structural diversity within each type of phosphoglyceride is due to variability in chain length and degree of saturation of the fatty acid ester groups. In general, a saturated (containing no double bonds between carbon atoms) fatty acid ester group is attached to the *sn*-1 position, whereas an unsaturated (containing one or more double bonds between carbon atoms) fatty acid ester group is attached to the *sn*-2 position.

Sphingomyelin. Sphingomyelin is the major sphingosine-containing phospholipid in eukaryotic cells. Its general structure consists of a fatty acid attached to sphingosine by an amide linkage, producing a ceramide. The alcohol group in the ceramide is esterified with phosphorylcholine, producing sphingomyelin (structure **3**).

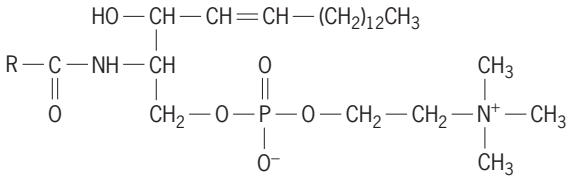

(**3**)

Biosynthesis. All organisms have the ability to synthesize phospholipids. In the eukaryotic cell, phosphatidylcholine is the major phosphoglyceride; in the prokaryotic cell, phosphatidylethanolamine predominates. Phosphoglycerides are synthesized either by the condensation of cytidine diphosphate (CDP)–diacylglycerol directly to the head group or by the condensation of 1,2-diacylglycerol to a CDP-bound head group. Sphingomyelin is synthesized from the condensation of ceramide and CDP-choline.

Since dipalmitoyl-phosphatidylcholine is the major constituent of lung surfactant, failure to produce an adequate amount of this phospholipid by the type II pneumocytes (lung cells) causes the collapse of lung alveoli in newborn babies,

a disorder termed respiratory distress syndrome (RDS). An exogenous surfactant, composed mainly of dipalmitoyl-phosphatidylcholine, has been used successfully since the 1990s as part of the routine care for preterm neonates with RDS. *See* INFANT RESPIRATORY DISTRESS SYNDROME.

Biological membranes. The biological membrane is a barrier that separates the inside of a cell from its outside environment. It also serves as a barrier to separate cellular contents into different compartments. A bilayer membrane is formed spontaneously when phospholipids are dispersed in an aqueous solution. In this bilayer structure, phospholipids are arranged in two leaflets with the hydrophobic tails facing each other, and the hydrophilic ends exposed to the aqueous medium. One distinct feature of the biological membrane is the asymmetrical distribution of phospholipids between the inner and outer leaflets. In a typical plasma membrane, the majority of phosphatidylcholine is located on the outer leaflet, whereas the majority of phosphatidylethanolamine is located in the inner leaflet. The shape of the bilayer is determined by the asymmetrical distribution of phospholipid types, the chain length of each phospholipid, as well as the degree of saturation of the acyl groups in the phospholipid. Individual phospholipid molecules are able to move freely in the lateral plane of the bilayer but not in the transverse plane (flip-flop). Small uncharged molecules are able to diffuse through the bilayer structure, but the permeability of larger or charged molecules is restricted. The arrangement of phospholipid molecules into a bilayer in an aqueous medium follows the laws of thermodynamics and represents the structural basis for the formation of all biological membranes. Selected protein molecules, which are responsible for other membrane functions such as providing channels for large or polar molecules to enter or exit the cell, are embedded in the phospholipid bilayer. *See* CELL MEMBRANES.

Phospholipases and signal transduction. For a long time, phospholipids were regarded as merely building blocks for the biological membrane. It was discovered in the mid-1970s, however, that phospholipids participate extensively in the transduction of biological signals across the membrane. For example, a special form of phosphoglyceride, 1-alkyl-2-acetyl-glycero-3-phosphocholine, acts as a very powerful signaling molecule. It causes the aggregation and degranulation of blood platelets, and is known as platelet-activating factor (PAF). Other effects of this compound include increasing pulmonary edema, hypersensitivity, acute inflammatory reactions, and anaphylactic shock. *See* ANAPHYLAXIS; EDEMA; INFLAMMATION.

The production of signal molecules is largely through the action of phospholipases. Phospholipases are enzymes responsible for the degradation (via hydrolysis) of phosphoglycerides.

Phospholipase A_2. The most common type is phospholipase A_2 (PLA_2), which is found in all tissues, in pancreatic juice, as well as in snake and bee venoms. The reaction below shows the hydrolysis reaction

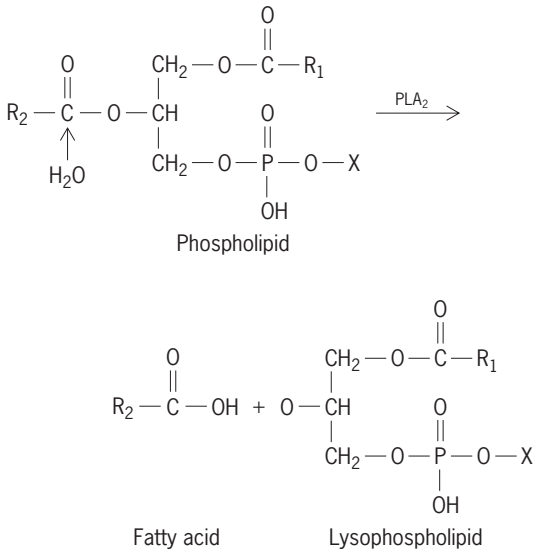

catalyzed by PLA_2. Over 100 molecular forms of PLA_2 have been identified, and the enzyme can be divided into two general classes: the histidine- and the serine-containing enzymes.

The histidine-containing enzymes are normally secreted in venoms or digestive juice. The major biological role of these enzymes is the digestion of food. They have small molecular weights (<20 kD) but consist of a large number of disulfide bonds. This class of PLA_2 has an absolute requirement of calcium for activity and consists of a histidine/aspartic acid pair at the catalytic site.

The serine-containing forms of PLA_2 are normally intracellular enzymes. They have larger molecular weights and consist of a serine molecule at the catalytic site. This class of enzymes does not exhibit an absolute requirement for calcium, although enzyme activity can be modulated by intracellular calcium concentrations. The serine-containing enzymes are responsible for the intracellular release of arachidonic acid for prostaglandin biosynthesis. When the cell is stimulated, a change in the intracellular calcium level causes the activation of PLA_2 and the release of arachidonic acid from phospholipids. The arachidonic acid released is subsequently converted to prostaglandins by the action of cyclooxygenase (COX) enzymes. *See* EICOSANOIDS.

Phospholipase C. The hydrolysis of a phospholipid by phospholipase C results in the production of a diacylglycerol and a phosphorylated head group. The product(s) from this reaction may also act as signal molecules. For example, when the hormone vasopressin is bound to its receptor on the plasma membrane of a liver cell, the binding sets off a cascade of reactions that result in the activation of phospholipase C. The activated phospholipase C then catalyzes the breakdown of a special form of phosphatidylinositol called phosphatidylinositol bisphosphate (PIP_2), producing diacylglycerol and inositol triphosphate. These two hydrolytic products of PIP_2 are known as second messengers. Second

messengers are molecules generated inside the cell in response to the binding of the primary messenger (a hormone) to the receptor. The diacylglycerol generated stimulates the activation of protein kinase C (a Ca^{2+}-dependent enzyme that transfers phosphate groups to certain cellular proteins, thereby regulating important physiological functions, such as cell growth), whereas the inositol triphosphate causes an increase in the intracellular calcium concentration. The activation of protein kinase C and the increase in cellular calcium, initiated by vasopressin binding, lead to the enhanced breakdown of glycogen in the liver cell, thus producing more glucose.

Another signal transduction pathway also involves PIP_2 and its conversion into phosphatidylinositol 3,4,5-trisphosphate (PIP_3). Upon stimulation, the insulin receptor causes the activation of a kinase, which converts the PIP_2 molecule into PIP_3. The PIP_3 formed is a lipid signaling molecule that activates several effectors, including the proto-oncogene (an oncogene precursor) product protein kinase B (also known as Akt). This proto-oncogene product could cause cellular cancer if it is not deactivated. When signaling by PIP_3 is no longer required, the molecule is converted back to PIP_2 by the action of a protein called PTEN, which acts as a phosphatase (an enzyme that removes phosphoryl groups from proteins). PTEN is also known as a tumor suppressor because it attenuates the stimulation of protein kinase B/Akt by PIP_3. A mutation of the PTEN gene may affect its phosphatase activity, which in turn hampers its ability to deactivate PIP_3. The continued activation of protein kinase B/Akt, in combination with other factors, could lead to an increase in cell growth and possible tumor development. *See* INSULIN; ONCOGENES; PROTEIN KINASE; TUMOR.

Phospholipase D. Phospholipase D is the enzyme that catalyzes the hydrolysis of phosphatidylcholine for the release of phosphatidic acid and the choline headgroup. The phosphatidic acid released may itself act as a signal molecule or be further hydrolyzed to form diacylglycerol (see above). Phospholipase D is ubiquitously present in plants, and the enzyme is located in the plasma membrane of the mammalian cell. *See* SECOND MESSENGERS.

Genetic disorders. Sphingomyelinase, a lysosomal enzyme, hydrolytically degrades sphingomyelin into phosphorylcholine and ceramide. The ceramide, whether produced from the hydrolysis of sphingomyelin or synthesized from serine and palmitate in a de novo pathway, is regarded as an important cellular signal for inducing apoptosis (programmed cell death). Niemann-Pick disease is a genetic disorder caused by a defect in the production of sphingomyelinase, which leaves the cell with no or limited ability to degrade sphingomyelin. In a severe form (type A) of this disease, the liver and spleen are sites of lipid deposits and are therefore tremendously enlarged. The lipid deposits consist primarily of the sphingomyelin that cannot be degraded. Infants with this disease suffer severe mental retardation and death in early childhood. The type A form of this disease is detected predominantly in Ashkenazi Jews, with a carrier frequency of 1 in 100. A number of less severe variants of this disease are known in which minor damage to neural tissues has been detected. In addition, the kidney, spleen, liver, and bone marrow are affected, and the life expectancy of these individuals is reduced to early adulthood. *See* HUMAN GENETICS; LIPID METABOLISM. Patrick Choy

Bibliography. D. E. Vance and J. Vance, *Biochemistry of Lipids, Lipoproteins and Membranes*, Elsevier, 2002.

Phosphorescence

A delayed luminescence, that is, a luminescence that persists after removal of the exciting source. It is sometimes called afterglow.

This original definition is rather imprecise, because the properties of the detector used will determine whether or not there is an observable persistence. There is no generally accepted rigorous definition or uniform usage of the term phosphorescence. In the literature of inorganic luminescent systems, some authors define phosphorescence as delayed luminescence whose persistence time decreases with increasing temperature. According to this usage, luminescence whose persistence time is independent of temperature is called fluorescence regardless of the length of the afterglow; a temperature-independent afterglow of long duration is called simply a slow fluorescence, which implies that the atomic or molecular transition involved is forbidden to a greater or lesser degree by the spectroscopic selection rules. In nonphotoconductive inorganic systems, phosphorescence arises when some excitation process has placed an atom (or ion or molecule) in a metastable energy state M (from which transitions to the state of lowest energy, or ground state G, are highly improbable or forbidden), and energy from the thermal vibrations of the system subsequently raises the atom to a higher energy state E from which luminescent transitions are highly probable or allowed (see **illus**.). The most common mechanism of phosphorescence in photoconductive inorganic systems, however, occurs when electrons or holes, set

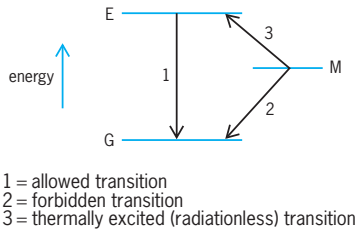

usage

process	1	2	3 followed by 1
inorganic	fluorescence	fluorescence	phosphorescence
organic	fluorescence	phosphorescence	delayed fluorescence

Atomic and molecular transitions involved in luminescence, and corresponding terminology.

free by the excitation process and trapped at lattice defects, are expelled from their traps by the thermal energy in the system and recombine with oppositely charged carriers with the emission of light. In these cases, the level M represents the state of the system with the electron or hole trapped. *See* HOLE STATES IN SOLIDS; SELECTION RULES (PHYSICS).

In the study of organic systems much attention has been given to the G→E→M process, which is the excitation of a molecule from its ground state to an excited state, both usually spectroscopic singlet states, followed by a radiationless transition (called an intersystem crossing) from the singlet excited state to the metastable triplet state M. In the organic literature the term phosphorescence is reserved for the forbidden luminescent transition M → G, while the afterglow corresponding to the M→E→G process is called delayed fluorescence. The spectrum (color) of organic "phosphorescence," defined in this way, is necessarily different from the spectrum of the ordinary fluorescence, because the emitting states (E and M) in the two cases are different and the final (ground) state G is the same. Conversely, the spectra of ordinary fluorescence and delayed fluorescence in organic systems are the same, because the luminescent transition takes place between the same emitting state E and the ground state G in both cases.

The temperature-dependent (M→E→G process) luminescence of a given system can exhibit a wide range of persistence times. At very low temperature where there is insufficient energy available to raise atoms from metastable to emitting states, or to expel electrons from traps, little or no afterglow is observed. At some higher temperature a low-intensity, long-lived afterglow will be observed; at a still higher temperature the afterglow will be brighter but of shorter duration. Finally, at some high temperature where rate of expulsion of atoms from metastable states or rate of expulsion of electrons from traps is very rapid, afterglow can become immeasurably short.

The time dependence of the luminescence intensity (the decay law) can be extremely complex, depending on the number and energies of the metastable states or electron traps involved. Phosphors which give a phosphorescent emission visible to the eye for about half a day at normal temperatures have been synthesized. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; FLUORESCENCE; LIGHT; LUMINESCENCE.

Clifford G. Klick; James H. Schulman

Phosphorus

A chemical element, P, atomic number 15, atomic weight 30.9738. Phosphorus forms the basis of a very large number of compounds, the most important class of which are the phosphates. For every form of life, phosphates play an essential role in all energy-transfer processes such as metabolism, photosynthesis, nerve function, and muscle action. The nucleic acids which among other things make up

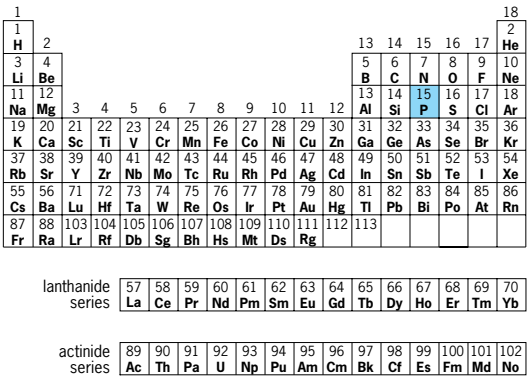

the hereditary material (the chromosomes) are phosphates, as are a number of coenzymes. Animal skeletons consist of a calcium phosphate. *See* PERIODIC TABLE; PHOSPHATE.

About three-quarters of the total phosphorus (in all of its chemical forms) used in the United States goes into fertilizers. Other important uses are as builders for detergents, nutrient supplements for animal feeds, water softeners, additives for foods and Pharmaceuticals, coating agents for metal-surface treatment, additives in metallurgy, plasticizers, insecticides, and additives for petroleum products. *See* DETERGENT; WATER SOFTENING.

Of the nearly 200 different phosphate minerals, only one, fluorapatite, $Ca_5F(PO_4)_3$, is mined chiefly from large secondary deposits originating from the bones of dead creatures deposited on the bottom of prehistoric seas and from bird droppings on ancient rookeries. *See* PHOSPHATE MINERALS.

Research in phosphorus chemistry indicates that there may be as many compounds based on phosphorus as on carbon. In organic chemistry it has been customary to group the various chemical compounds based on carbon into families which are called homologous series. This can also be done in the chemistry of phosphorus compounds, even though many phosphorus-based families are incomplete. The best known of the families of compounds based on phosphorus is the group of chain phosphates. Phosphate salts consist of cations, such as sodium, along with chain anions, such as $(P_nO_{3n+1})^{(n+2)-}$, which may have 1–1,000,000 phosphorus atoms per anion.

The phosphates are based on phosphorus atoms tetrahedrally surrounded by oxygen atoms, with the lowest member of the series being the simple PO_4^{3-} anion (the orthophosphate ion). The family of chain phosphates is based on a row of alternating phosphorus and oxygen atoms in which each phosphorus atom remains in the center of a tetrahedron of four oxygen atoms. There is also a closely related family of ring phosphates, a member of which, the trimetaphosphate, is shown in **Fig. 1**.

An interesting structural characteristic of many known phosphorus compounds is the formation of cagelike structures. Such cagelike molecules are exemplified by white phosphorus, P_4, and one of the phosphorus pentoxides, P_4O_{10} (**Fig. 2**). Network

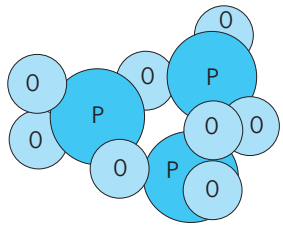

Fig. 1. Ring phosphate anion, $(P_3O_9)^{3-}$.

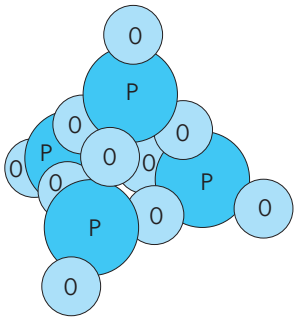

Fig. 2. Phosphorus pentoxide, P_4O_{10}, in vapor state.

structures are also common; for example, black phosphorus crystals in which the atoms are bonded together in the form of vast, corrugated planes (**Fig. 3**).

In the majority of its compounds, phosphorus is chemically bonded to four neighboring atoms. There is a large number of compounds in which one of the four neighboring atoms is absent, and in which its place is taken by an unshared pair of electrons. There are also a few compounds in which there are five or six neighboring atoms bonded to the phosphorus. These compounds are very reactive and tend to be unstable.

A large number of organic-phosphorus compounds have been prepared. Most of these chemical structures involve three or four neighboring atoms bonded to the phosphorus, but stable structures having two, five, or six neighboring atoms per phosphorus are also known. *See* ORGANOPHOSPHORUS COMPOUND; PHOSPHATE.

Essentially all of the phosphorus used in commerce is in the form of phosphates. The majority of phosphatic fertilizers consist of highly impure monocalcium or dicalcium orthophosphate, $Ca(H_2PO_4)_2$ and $CaHPO_4$. These phosphates are salts of orthophosphoric acid. *See* FERTILIZER.

The phosphorus compound of major biological importance is adenosine triphosphate (ATP), which is an ester of sodium tripolyphosphate, widely employed in detergents and water-softening compounds. Practically every reaction in metabolism and photosynthesis involves the hydrolysis of this tripolyphosphate to its pyrophosphate derivative, called adenosine diphosphate (ADP).

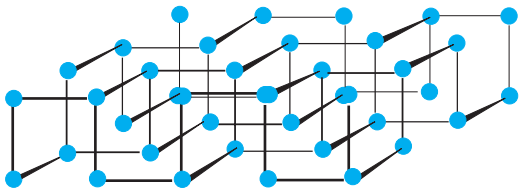

Fig. 3. Black phosphorus, P_n.

John R. Van Wazer

Bibliography. D. E. C. Corbridge, *Phosphorus 2000: Chemistry, Biochemistry & Technology*, 2000; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; E. N. Walsh et al. (eds.), *Phosphorus Chemistry: Developments in American Science*, 1992.

Photoacoustic spectroscopy

A technique for measuring small absorption coefficients in gaseous and condensed media, involving the sensing of optical absorption by detection of sound. It is frequently called optoacoustic spectroscopy. Although the technique dates back to 1880 when A. G. Bell used chopped sunlight as the source of radiation, it remained dormant for many years, primarily because of the lack of suitable powerful sources of tunable radiation. However, the usefulness of optoacoustic detection for spectroscopic applications was recognized early in its development, and pollution monitoring instruments (called spectrophones) dedicated for detection of specific gaseous constituents have been used intermittently since Bell's work.

Methods of measuring absorption. During the transmission of optical radiation through a sample (gas, liquid, or solid), the absorption of radiation by the sample can be measured by at least three techniques. The first one is the straightforward detection technique which requires a measurement of the optical radiation level with and without the sample in the optical path. The transmitted power P_{out} and the incident power P_{in} are related through Eq. (1), where

$$P_{out} = P_{in}e^{-\alpha l} \tag{1}$$

α is the absorption coefficient and l is the length of the absorber. With this technique, the minimum measurable αl is of the order of 10^{-4} unless special precautions have been taken to stabilize the source of radiation.

The second of the techniques is the derivative absorption technique where the frequency of the input radiation is modulated at a low radio frequency or audio frequency, ω_m. The transmitted radiation then contains a time-varying component at ω_m, if the optical path contains absorption which has a frequency-dependent structure. (For structureless absorption, modulated absorption spectroscopy does not provide a signal that can characterize the amount of absorption.) For situations where the absorption has well-defined structure, the modulation absorption spectroscopy can be used to measure αl as small as about 10^{-8} for sufficiently high input powers. The ability to measure the small absorption effects is independent of the input and output power levels for the straightforward measurement technique as long as the noise contributed by the detector is not a factor in determining the signal-to-noise ratio. For the

derivative absorption technique, the smallest αl that can be measured varies as $(P_{in})^{-1}$ until the shot noise of the detector begins to be appreciable.

The third technique, optoacoustic detection, is a calorimetric method where no direct detection of optical radiation is carried out but, instead, a measurement is made of the energy, with power P_{abs}, absorbed by the medium from the incident radiation, Eq. (2). Thus the optoacoustic signal, V_{oa}, is given by Eq. (3), where K is the constant describing the con-

$$P_{abs} = P_{in}(1 - e^{-\alpha l}) \tag{2}$$

$$V_{oa} = K[P_{in}(1 - e^{-\alpha l})] \approx KP_{in}\alpha l \qquad (\text{for } \alpha l \ll 1) \tag{3}$$

version factor for transforming the absorbed energy into an electrical signal using an appropriate transducer. It has been tacitly assumed that the absorbed energy is lost by nonradiative means rather than by reradiation. The optoacoustic detection scheme implies that the absorbed energy will be converted into acoustic energy for eventual detection.

From Eq. (3), the optoacoustic signal is proportional to the incident power and the absorption-length product αl. Thus, for given sources of noise from the detection transducers, the signal-to-noise ratio improves as the incident energy is increased. Put differently, the smallest amount of absorption that can be measured using the optoacoustic technique varies as $(P_{in})^{-1}$, with no limitation on the level to which P_{in} can be increased for detecting small absorptions. Values of αl as small as 10^{-10} can be measured in the gas phase. The techniques generally used for gases and those used for condensed-phase optoacoustic spectroscopy differ in detail somewhat.

Gases. If optical radiation is amplitude-modulated at an audio frequency, the absorption of such radiation by a gaseous medium that has been confined in a cell with appropriate optical windows for the entrance and exit of the radiation and the subsequent nonradiative relaxation of the medium, will cause a periodic variation in the temperature of the column of the irradiated gas (**Fig. 1**). Such a periodic rise and fall in temperature gives rise to a corresponding periodic variation in the gas pressure at the audio frequency. The audio-frequency pressure fluctuations (that is, sound) are efficiently detected using a sensitive gas-phase microphone. The intrinsic noise limitation to the optoacoustic detection scheme arises from the Brownian motion of gas atoms/molecules, and Kreuzer showed that the minimum detectable absorbed power is $P_{min} \approx 3.6 \times 10^{-11}$ W for a 11.9-cm-long (4.7-in) cell. Substituting P_{min} for P_{abs} in Eq. (2), and noting, as in Eq. (3), that $(1 - e^{-\alpha l}) \cong \alpha l$ for $\alpha l \ll 1$, it follows that α_{min} varies as (P_{min}/P_{in}) as indicated above. The usefulness of the optoacoustic detection for measurement of small absorption coefficients became evident with the development of a variety of tunable high-power laser sources which could take advantage of the $(P_{in})^{-1}$ dependence. Using a spin-flip Raman laser tunable in the 5.0- to 5.8- micrometer range, with a power output of approximately 0.1 W,

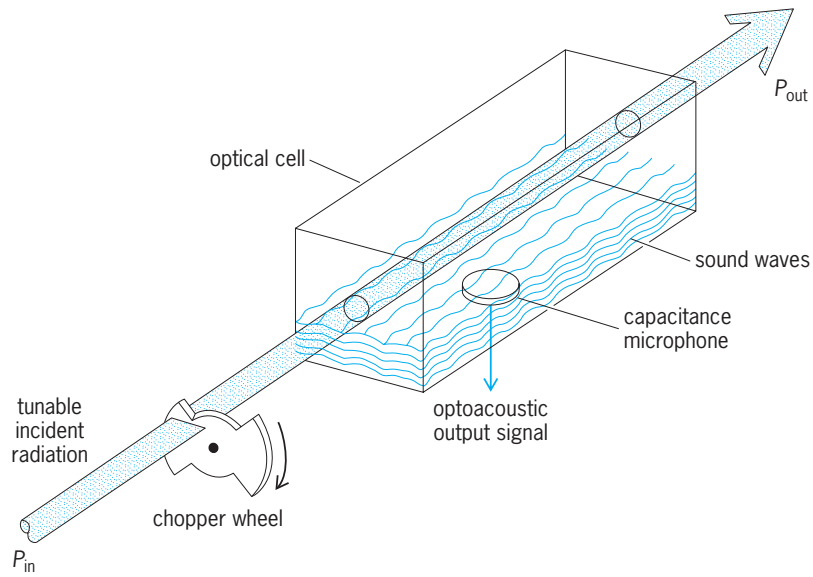

Fig. 1. Optoacoustic cell for gaseous spectroscopy.

C. K. N. Patel and R. J. Kerl were able to detect α_{min} of approximately 10^{-10} cm^{-1} for a cell length of 10 cm (4 in.). These studies used a miniature optoacoustic cell (**Fig. 2**) with a total gas volume of approximately 3 cm^3 (0.18 in.^{3}). The absorber used was nitric oxide diluted in nitrogen. It is estimated that for a signal-to-noise ratio of approximately 1, and a time constant of 1 s, it is possible to detect a nitric oxide (NO) concentration of approximately 10^7 molecules cm^{-3}, corresponding to a volumetric mixing ratio of approximately 1:10^{12} at atmospheric pressure. *See* LASER.

The capability of measuring extremely small absorption coefficients and correspondingly small concentrations of the absorption gases has many applications, including high-resolution spectroscopy of isotopically substituted gases, excited states of

Fig. 2. Sensitive optoacoustic gaseous spectroscopy cell.

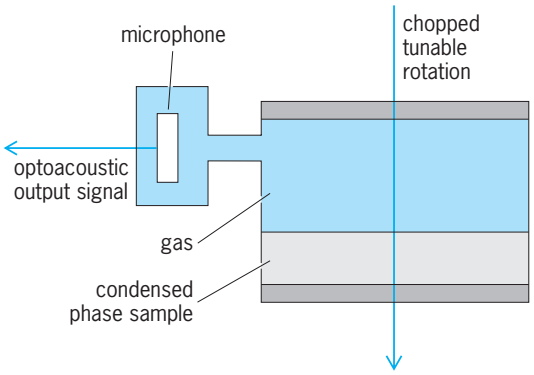

Fig. 3. Arrangement for condensed-phase photoacoustic spectroscopy.

molecules and forbidden transitions, and pollution detection. In the last application, both continuously tunable lasers, such as the spin-flip Raman laser and dye lasers, and step tunable infrared lasers, such as the carbon dioxide (CO_2) and carbon monoxide (CO) lasers, have been used as sources of high power radiation. The pollution measurements have demonstrated that the optoacoustic spectroscopy technique in conjunction with tunable lasers can be routinely used for on-line real-time in-place detection of undesirable gaseous constituents at subparts-per-billion levels. Specific examples include the measurement of nitric oxide on the ground and in the stratosphere (where nitric oxide plays an important role as a catalytic agent in the stratospheric ozone balance) and measurements of hydrocyanic acid (HCN) in the catalytic reduction of $CO + N_2 + H_2 + \cdots$ over platinum catalysts. These studies point toward expanding use in the future of optoacoustic spectroscopy in pollution detection. *See* LASER; LASER SPECTROSCOPY.

Condensed-phase spectroscopy. A straightforward application of the gas-phase optoacoustic spectroscopy technique to the study of condensed phase (liquid or solid) spectra involves enclosing the condensed phase material within the gas-phase optoacoustic cell (**Fig. 3**). The "photoacoustic" signals generated in the sample due to the absorption of optical radiation are communicated to the gas-phase microphone via coupling through the gas filling the chamber. The inefficiency of such a system is high because of the very poor acoustical match (coupling efficiency approximately 10^{-5}) between the condensed-phase sample and the gas. In reality, because of the large acoustical mismatch, the detection scheme is really "photothermal" rather than "photoacoustic," and this scheme provides a capability of measuring fractional absorption at a level of approximately 10^{-4} when a continuous-wave laser power of approximately 10 W is used. A more severe drawback of the scheme, however, lies in the difficulty of interpretation of the data because of the intimate dependence of the observed optoacoustic signal from the microphone on the chopping frequency, absorption depth, and heat diffusion depth. However, in spite of its shortcomings, the gas-phase microphone technique for condensed-phase optoacoustic spectroscopy has found applications.

A very sensitive calorimetric spectroscopic technique has been developed for the study of weak absorption in liquids and solids. This technique uses a pulsed tunable laser for excitation and a submerged piezoelectric transducer, in the case of a liquid, or a contacted piezoelectric transducer, in the case of a solid, for the detection of the ultrasonic signal generated due to the absorption of the radiation and its subsequent conversion into a transient ultrasonic signal (**Fig. 4**). The major distinction between the above condensed-phase "photoacoustic" spectroscopy technique and the pulsed-source, submerged or contacted piezoelectric transducer technique, is the high coupling efficiency of approximately 0.2 for the ultrasonic signal in the liquid to the submerged transducer, or an efficiency of approximately 0.9 for coupling the ultrasonic wave in a solid to a bonded transducer. Because of this high efficiency, the pulsed-laser, submerged or bonded

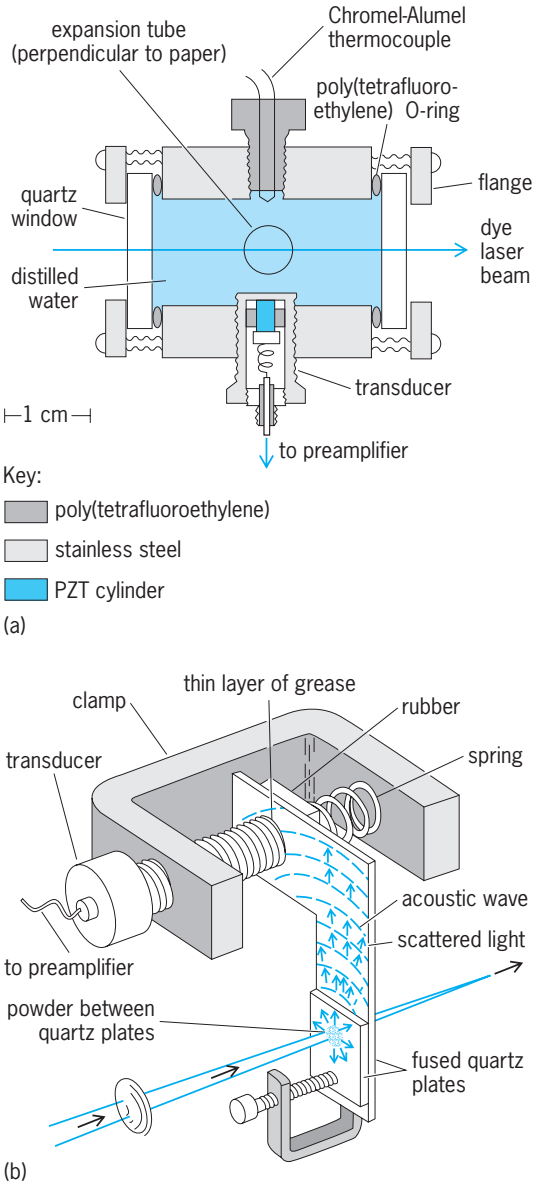

Fig. 4. Arrangement for pulsed-laser (*a*) immersed and (*b*) contacted piezoelectric transducer optoacoustic spectroscopy.

optoacoustic spectroscopy technique has been shown to be useful for measuring fractional absorptions (that is, values of αl) as small as 10^{-7} when using a laser source with pulse energy of approximately 1 millijoule, pulse duration of approximately 1 microsecond, and a pulse repetition frequency of 10 Hz. There is room for improvement by increasing the laser pulse energy. There is a possibility of the electrostriction effect giving rise to an unwanted background signal, but this signal is not dependent on the light wavelength, and can be minimized by proper choice of experimental parameters.

The pulsed-laser, submerged or bonded piezoelectric transducer technique has a further advantage that time-gating of the ultrasonic signal output can be utilized for the rejection of spurious signals since the sound velocity in condensed media is known and hence the exact arrival time of the real optoacoustic pulse can be calculated. This technique has been used for measurement of very weak overtone spectra of a variety of organic liquids, optical absorption coefficients of water and heavy water in the visible, two-photon absorption spectra of liquids, Raman gain spectra in liquids, absorption of thin liquid films, spectra of solids and powders, and weak overtone spectra of condensed gases at low temperatures. Because of the capability of measuring very small fractional absorptions, the technique is clearly applicable to the area of monitoring water pollution, impurity detection in thin semiconductor wafers, transmission studies of ultrapure glasses (used in optical fibers for optical communications), and so forth. Further, even though in all of the present studies use is made of only the optical radiation, there is no reason to restrict "optoacoustic" spectroscopy to the optical region. By using pulsed x-ray sources, such as the synchrotron light source or pulsed electron beams, the principle described above for a pulsed-light-source, submerged or bonded piezoelectric transducer, gated-detection technique can be extended to x-ray acoustic spectroscopy and electron-loss acoustic spectroscopy. These extensions are likely to have major impact on materials and semiconductor fabrication technology. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; SYNCHROTRON RADIATION.

C. K. N. Patel

Bibliography. P. Hess and J. Pelzl (eds.), *Photoacoustic and Photothermal Phenomena*, 1988; V. S. Letokhov and V. P. Zharov, *Laser Optoacoustic Spectroscopy*, 1986; A. Mandelis (ed.), *Principles and Perspectives of Photoacoustic Spectroscopy*, 1991; M. Rosencwaig, *Photoacoustics and Photoacoustic Spectroscopy*, 1980, reprint 1990.

Photoaffinity labeling

Photoaffinity labeling is a process by which a macromolecule can be labeled at or near its binding or active site. The method can be applied to proteins, nucleic acids, and lipids—all macromolecules—by chemists who need to identify the parts of these molecules that are significant for particular biological functions. Otherwise, x-ray crystallography can frequently determine the complete structure of proteins and sometimes of nucleic acids, although the method does not necessarily highlight the reactive groups in a biomolecule. Nuclear magnetic resonance spectroscopy is increasingly useful in illuminating such structures. *See* NUCLEAR QUADRUPOLE RESONANCE; X-RAY CRYSTALLOGRAPHY.

Many important biological processes involve the formation of a complex between a biopolymer and a chemical reagent; for example, enzymes form complexes with their substrates on the way to catalysis, and antibodies form tight complexes with their antigens. Many biochemical receptors (such as hormone receptors) are complexed with lipid membranes. These complexes cannot usually be crystallized, but information about them can often be obtained by photoaffinity labeling.

In ordinary affinity labeling, a modified substrate that contains a reactive group is chosen. When such a substrate forms a complex with a macromolecule, its reactive substituent can often react with, and so mark, nearby nucleophiles. Although affinity labeling has proved an important method of marking groups essential to enzymatic catalysis, it is grossly inadequate because most of the chemical groups in active sites or binding sites are insufficiently reactive. In many instances, this problem can be solved by photoaffinity labeling.

The idea behind photoaffinity labeling is to place in the binding site a compound that is essentially inert but can be photoactivated at will to yield a highly reactive intermediate. Preferably, this intermediate will react with almost any chemical structure. In the first published example of photoaffinity labeling, a diazoacetyl group was placed on the reactive serine residue of the enzyme chymotrypsin (Chy) [reaction (1); the asterisk indicates an atom of carbon-14]. When the diazo group was photoacti-

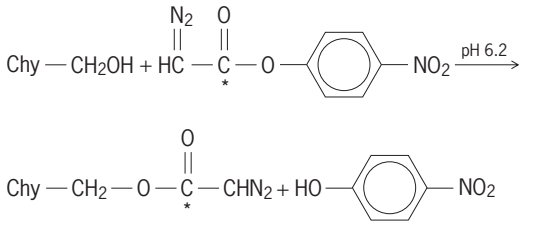

(1)

vated, it decomposed to yield a carbene [reaction (2)], a derivative of divalent carbon, that will react

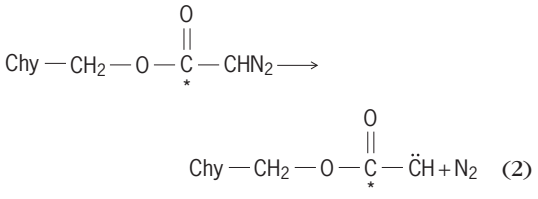

(2)

with almost anything; it will even react with unactivated carbon-hydrogen bonds.

This carbene, and another similar one, marked

Examples of photoaffinity labeling reagents

Reagent	Type	Photoproduct	Type
$N_2CH-CO_2-C_6H_4-NO_2$	Diazo	$:CH-CO_2-R$	Carbene
$CF_3-CN_2-CO_2-C_6H_5$	Diazo	$CF_3-\ddot{C}-CO_2-R$	Carbene
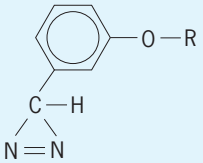	Diazirine	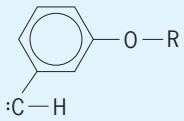	Carbene
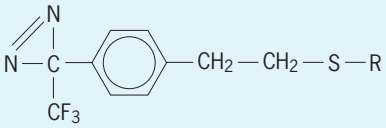	Diazirine	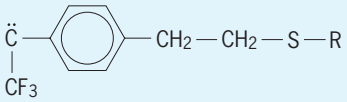	Carbene
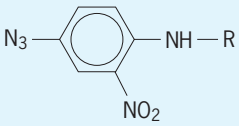	Azide	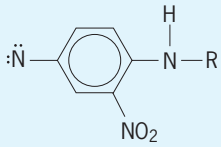	Nitrene
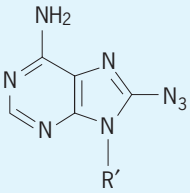	Azide		Nitrene

a number of the amino acids near the active sites of chymotrypsin and trypsin, including a tyrosine residue and the methyl group of an alanine residue. But the process with diazoacetyl chymotrypsin also revealed several of the potential problems with the method. The major reaction product of the carbene proved to be that of internal (Wolff) rearrangement [reaction (3)], rather than an attack on nearby structures.

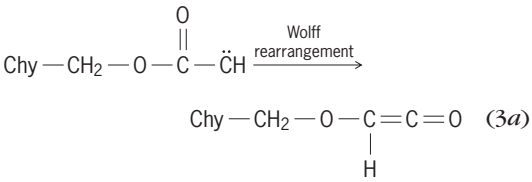

(3*a*)

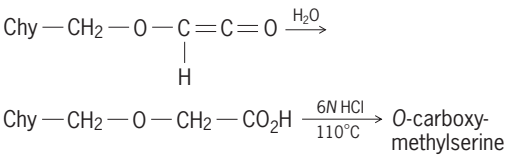

(3*b*)

Another difficulty is nonspecific labeling, that is, labeling of surrounding nucleophiles that occurs when the photoproduct is insufficiently reactive and survives long enough to diffuse out of the pocket where it was formed, only to label amino acids or other structures that are distant from the binding site. Still another problem concerns the occasional transformation of singlet carbenes and nitrenes into the corresponding triplets, which are free-radical structures less reactive than the singlets. Photoaffinity labeling is successful only when the primary photoproducts (the singlet structures) react with their surroundings more rapidly than they undergo intersystem crossing to the relatively unreactive triplets.

Although such problems have also plagued other photoaffinity labeling reagents, they have largely been overcome. Many useful reagents have been prepared, and many successful labeling reactions have been carried out. Computer search of the literature suggests that the method has already been used as many as 5000 times. All 20 of the natural amino acids in proteins have been marked by one reagent or another (see **table**).

The successful reagents include diazirines that (like diazo compounds) yield carbenes on photolysis, arylazides that are photochemically decomposed to nitrenes (highly reactive compounds that contain univalent nitrogen), and ketones that yield free radicals on irradiation. Special structural features (such as trifluromethyl substituents) can enhance the activity of the reagents. The more indiscriminantly reactive the intermediate that results from photolysis, the better; ideally, such intermediates will react with and so identify all the surrounding proteins, lipids, nucleic acids, or other macromolecules. *See* AZIDE.

In order to identify the products of reactions, many of which occur only in low yield, the reagent must be made radioactive. This is usually accomplished with carbon-14 (^{14}C) or tritium (^{3}H), radioactive isotopes of relatively low energy that nevertheless can easily be traced.

Frequently, photoaffinity labeling can be utilized to cross-link two parts of a major assembly of molecules and to reveal its quaternary structure. A good example of the utilization of the method is determination of the packing of the protein helices in purple membrane. The relative positions of helices and of the ordering of proteins in membranes can probably be best determined by photoaffinity techniques. F. H. Westheimer

Bibliography. H. Bayley, *Photogenerated Reagents in Biochemistry and Molecular Biology*, Elsevier, 1983; V. Chowdhry and F. H. Westheimer, *Annu. Rev. Biochem.*, 48:293, 1979; F. Kotzyba-Hilbert, I. Kapfer, and M. Goeldner, *Angew. Chem. Int. Ed. Engl.*, 34:1206, 1995; M. S. Platz,, *Photochem. Photobiol.*, 65:193, 1997; J. F. Resek, D. Farrens, and H. G. Khorana, *Proc. Nat. Acad. Sci. USA*, 91:7643, 1994; A. Singh, E. Thornton, and F. H. Westheimer, *J. Biol. Chem.*, 237:3006, 1962.

Photochemistry

The study of chemical reactions of molecules in electronically excited states produced by the absorption of infrared (700–1000 nanometers), visible (400–700 nm), ultraviolet (200–400 nm), or vacuum ultraviolet (100–200 nm) light. Bond making and bond breaking as well as electron transfer and ionization are often observed in both organic and inorganic compounds as a consequence of such excitation. *See* ELECTROMAGNETIC RADIATION.

Electronic absorption. An important generalization sometimes called the first law of photochemistry is that only light that is absorbed can induce chemical change. The absorption of a photon induces an electronic transition in which an electron originally present in a molecular orbital, usually a bonding or nonbonding molecular orbital of the ground state of the absorbing molecule, is promoted to a higher-lying orbital. The excited state produced by absorption of light has a different electronic structure than its ground-state precursor and can reasonably be regarded as an isomeric species with distinct and characteristic chemical and physical properties. *See* PHOTON.

The Born-Oppenheimer approximation allows electronic transitions to be considered as distinct from spins of the electrons and the relative positions of nuclei. The Franck-Condon principle states that because an electronic transition is rapid compared with nuclear motion, the positions of all nuclei remain frozen at the ground-state equilibrium values during an electronic transition. In a Heitler-London plot of the dependence of energy on nuclear configuration (**Fig. 1**), light absorption is thus represented as a straight vertical line: the Franck-Condon principle asserts that a vertical transition (one that takes place without a change in ground-state geometry) is most probable. *See* FRANCK-CONDON PRINCIPLE; RENNER-TELLER EFFECT.

Excited state. The probability of producing an excited state upon exposing a ground-state molecule to light is directly proportional to the square of the transition dipole moment, which in turn depends on the electron redistribution attained by the transition. Selection rules are used to define those factors that permit or forbid a specific electron redistribution. The most favorable transitions are those that take place without changes in state symmetry, without spin interconversion, and where excess vibronic energy is minimized. *See* GROUND STATE; SELECTION RULES (PHYSICS).

Most organic molecules exist as ground-state singlets in which all electrons are paired. Because photoexcitation causes the promotion of only a single electron, two singly occupied orbitals are produced upon excitation. If the electronic transition takes place without a spin inversion, these two electrons have opposite spins, and a singlet excited state is produced (**Fig. 2**).

The number of unpaired electrons in a molecule determines its multiplicity: a molecule with no unpaired spins is a singlet; one with one unpaired spin is a doublet; one with two unpaired spin is a triplet; one with three unpaired spin is a quartet, and so forth. If an electronic transition were to take place

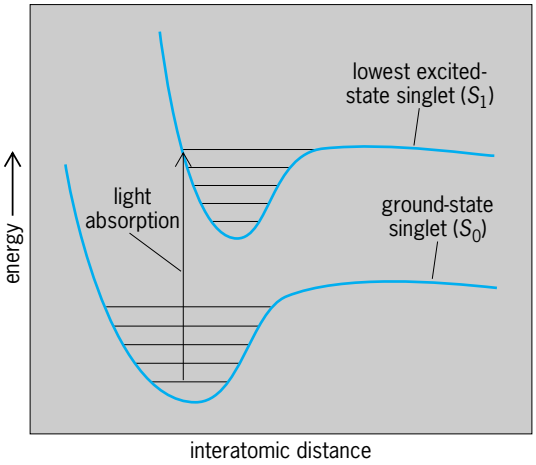

Fig. 1. Representation of a vertical Franck-Condon transition.

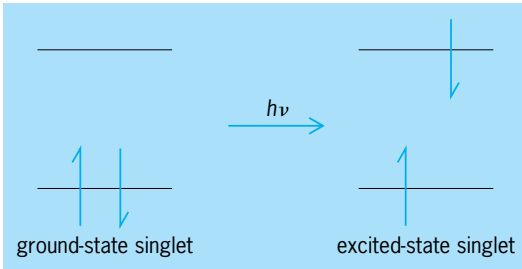

Fig. 2. Spin selection rules dictate that no spin inversion can take place during photoexcitation.

with a spin inversion, the two singly occupied orbitals would be populated by electrons with parallel spins, producing a triplet excited state. Spin restrictions forbid spin inversion during excitation, and only singlet-singlet electronic transitions are easily observed spectroscopically. After excitation, however, a change in state multiplicity can take place by a process called intersystem crossing. The facility of intersystem crossing is influenced by the magnitude of spin-orbital coupling, which can be enhanced by the presence of a heavy atom (an atom in the third row or below of the periodic table), either bound to the absorbing molecule or present externally as solvent. *See* PERIODIC TABLE; TRIPLET STATE.

For a molecule to be converted to an excited state by photoexcitation, it must have an available state with an energy that exactly matches that of the incident photon, for the energy of each photon is quantized. The energy of a photon is specified by its frequency (ν), from the Bohr frequency condition, as in Eq. (1), where E is the transition energy, h is Planck's

$$\Delta E = h\nu \tag{1}$$

constant, and ν is the frequency of the radiation. The frequency (ν), often expressed in cm^{-1}, is related inversely to the wavelength (λ) by the expression $\nu = c/\lambda$, where c is the velocity of light. Because energy is directly proportional to frequency, which is inversely proportional to wavelength, higher-energy transitions take place at shorter wavelengths: for example, ultraviolet transitions are of higher energy than those in the visible region of the electromagnetic spectrum. The energy of a given transition can be calculated directly from its wavelength (in nm) by substituting and inserting Avogadro's number (N_A) into the Bohr equation (2). From such a calculation,

$$\begin{aligned}\Delta E &= N_A hc/\lambda \\ &= 1.1963 \times 10^5 \text{ kJ mol}^{-1}/\lambda \text{ (in nm)}\end{aligned} \tag{2}$$

it can be shown that the energy emitted by a low-pressure mercury arc at 254 nm has an energy corresponding to 471 kJ mol^{-1} (113 kcal mol^{-1}). The lowest energy transition in many molecules is that in which an electron is promoted from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO). *See* EXCITED STATE; MOLECULAR ORBITAL THEORY.

Transitions. A chromophore is that part of the molecule that accounts for its absorption of light and its photochemical activity. The absorption corresponding to a particular chromophore depends on the type of transition involved in that particular excitation. The promotion of an electron from a π-bonding molecular orbital to a π-antibonding orbital is referred to as a π,π^* (read pi to pi star) transition. Such transitions are frequently encountered in alkenes, alkynes, aromatic molecules, and other unsaturated compounds. Because the spatial overlap of π and π^* orbitals is substantial, such a transition typically has high oscillator strength and a large extinction coefficient (absorbtivity). Promotion of an electron from a nonbonding molecular orbital to a π-antibonding orbital, referred to as an n,π^* transition, involves orbitals that are nearly orthogonal; and it takes place only inefficiently, that is, it has a low oscillator strength and a small extinction coefficient. Such transitions are often encountered in compounds containing carbon-heteroatom or heteroatom-heteroatom double bonds. Because nonbonding molecular orbitals lie at higher energy than bonding ones, n,π^* transitions are of lower energy than the corresponding π,π^* transitions. Both n,π^* and π,π^* transitions are usually found in the ultraviolet region of the electromagnetic spectrum. Transitions involving sigma (σ) bonds (for example $n\sigma^*$ transitions in amines, alcohols, ethers, and alkyl halides and σ,σ^* transitions in alkanes) are usually encountered at the high-energy end of the ultraviolet spectrum or in the vacuum ultraviolet region. *See* CHEMICAL BONDING; ULTRAVIOLET RADIATION.

Electronic transitions in inorganic compounds often involve a metal-based orbital. Promotion of an electron from one metal-based orbital to another is called a d, d transition; such transitions are commonly encountered in simple metal complexes. Promotion of an electron from a metal-based orbital to an orbital strongly associated with an attached ligand is referred to as a metal-to-ligand charge transfer (MLCT) transition. Conversely, promotion of an electron from a ligand-based orbital to a metal-based one is called a ligand-to-metal charge transfer (LMCT) transition. The d, d, MLCT, and LMCT transitions are usually found in the visible region of the spectrum. *See* COORDINATION COMPLEXES; LIGAND.

Each allowed transition of a compound registers as a band in the absorption spectrum, with the intensity of the transition (measured by its extinction coefficient) being governed by the operative selection rules. The transition intensity of a given absorption is measured by integrating over the whole absorption band. The resulting integrated absorption coefficient is directly proportional to the oscillator strength of the transition. The oscillator strength, a measure of the allowedness of an electric dipole transition compared to that of a free electron oscillating in the three dimensions, is directly related to an experimentally measured value, the extinction coefficient (ϵ). Beer's

law is given by Eq. (3), where A is the observed absorbance, ε is the extinction coefficient, b is the path length (in centimeters) of the cell used for the measurement, and c is the molar concentration of the absorbing species. This law is used to correlate the observed absorbance with the extinction coefficient and concentration of the absorbing species. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION.

$$A = \varepsilon bc \tag{3}$$

Photophysics. The excited state produced by absorption of a photon is not generally a stable species. After a characteristic lifetime that can vary from femtoseconds (10^{-15} s) to hours, the excited molecule will either relax to its ground-state precursor or undergo a chemical transformation. The term photophysics is used to describe nonreactive relaxation processes, which include radiative (taking place with the emission of light) and nonradiative (taking place without the emission of light) pathways.

When absorption of light produces a Franck-Condon excited state that is vibrationally excited, the excess vibrational energy is rapidly dissipated in a process called a vibrational cascade during a period of several femtoseconds to hundreds of picoseconds (10^{12} s), producing the lowest vibrational level of the excited state. When a more highly energetic excited state than the lowest-lying state is produced by the absorption event, relaxation from the upper excited state to a lower one of the same multiplicity also takes place rapidly (typically within a few picoseconds) by a process called internal conversion. Internal conversion is often accomplished isoenergetically, producing a highly vibrationally excited state of the lower-lying excited state, which relaxes very rapidly by a fast vibrational cascade to the lowest vibrational level of the lowest-energy excited state. Internal conversion is so fast, compared with radiative relaxation, that it is generally assumed that radiative processes always occur from the lowest electronic state of a given multiplicity (Kasha's rule) and that the efficiency of the emission is therefore independent of the energy absorbed (Vavilov's rule).

Emission of light associated with a relaxation that does not involve a spin inversion, for example, from an excited singlet state to a ground-state singlet, is called fluorescence. Emission associated with a relaxation in which spin inversion occurs, as from an excited triplet state to a ground-state singlet, is called phosphorescence. Because both fluorescence and phosphorescence result from states produced by fast vibrational relaxation and internal conversion, the emission is observed at a lower energy (longer wavelength) than the absorption that produced the excited state. A transition at a lower energy than a reference wavelength is said to be redshifted; one at higher energy than the reference is said to be blueshifted. In rare cases, thermally activated reverse intersystem crossing from a vibrationally excited triplet state to the excited singlet state can take place; because of the time delay associated with the double intersystem crossing, this emission is called E-type delayed fluorescence. *See* FLUORESCENCE; PHOSPHORESCENCE.

The energies of the lowest singlet and triplet excited states (relative to the ground state) can be obtained from the longest wavelength band of the fluorescence and phosphorescence spectra, respectively. This band is called a 0,0 band to indicate a transition between the lowest vibrational levels of the lowest-lying states. Singlet and triplet energies can also be determined indirectly by measuring quenching efficiencies. The shift between the 0,0 bands for absorption and emission in a single molecule is called its Stokes shift. A small stokes shift is usually observed when the excited state has a geometry similar to the ground state. A Jablonski diagram (**Fig. 3**) is often used to graphically depict the relationship between competing photophysical processes.

The term fluorescence lifetime is used for the time required for the emission of an excited singlet state to decay to half of its initial intensity [typically about 10 nanoseconds (10^{-10} s) for a fluorescent organic molecule]; phosphorescence lifetime, for that of an emissive triplet excited state [typically several microseconds (10^{-6} s) in the absence of a reactive quencher]. The longer triplet lifetime results from the forbiddenness of the phosphorescent emission, which must take place with a concomitant spin flip. The lifetime (τ) is defined as the reciprocal of the rate constant k_r for radiative decay (Eq. 4).

$$\tau = 1/k_r \tag{4}$$

Emission lifetimes can be measured directly or indirectly. In a direct measurement, the excited state is produced by a laser or electron pulse of very short duration (femtoseconds to nanoseconds), producing an excited state whose emission decay is monitored as a function of time. In an indirect determination, the emission intensity is monitored for solutions containing increasing concentrations of a

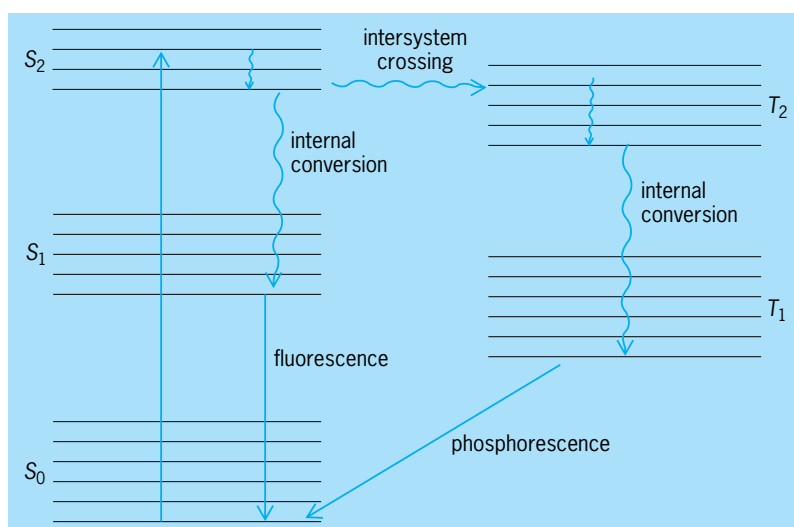

Fig. 3. Jablonski diagram. Solid arrows represent radiative processes; and wavy arrows nonradiative processes. *S* terms = singlet states; *T* terms = triplet states.

second chemical species that interacts efficiently (often at a diffusion-controlled rate) with the excited state. A relationship known as the Stern-Volmer equation is used to relate the relative emission intensity to the concentration of the quencher.

Quantum yield, or quantum efficiency, is defined as the number of molecules participating in a given photophysical process or reaction divided by the number of photons absorbed. The quantum yield ranges between zero and one for photoreactions induced by a single photon; values larger than one are indicative of a chain process in which product is formed in a repeating, dark cycle initiated by the photoexcitation. For a photochemical reaction, the number of molecules participating in the reaction is determined spectroscopically or chromatographically as a chemical yield per volume unit per time. The number of photons absorbed is obtained by measuring with a radiometer the light flux per volume unit per time or by employing a chemical actinometer, a known chemical reaction for which the quantum yield is known and accepted as a standard. One of the most frequently used chemical actinometers is the photoreduction of potassium ferrioxalate to ferrous ion (Fe^{2+}), which is determined colorimetrically as its red phenanthroline complex. The quantum yield can also be obtained as the ratio of the rate constant for the desired process to the sum of the rate constants for all competing processes emanating from the same excited state. *See* QUANTUM CHEMISTRY.

Bimolecular excited-state interactions. When an excited state molecule, M^*, is produced in the presence of high concentrations of the ground state, M, a complex called an excimer, $(MM)^*$, is produced, namely a dimer that is stable only in an electronically excited state [reaction (5)].

$$M^* + M \rightarrow (MM)^* \tag{5}$$

An excimer is stabilized relative to an isolated pair because the occupied HOMO interacts strongly with the isoenergetic HOMO of the ground state, inducing an orbital splitting which is energetically favorable. The formation of an excimer is often indicated by a strongly quenched monomer fluorescence coupled with a broad and structureless emission at a wavelength substantially redshifted from that of the monomer. Both singlet and triplet excimers are known, although the existence of the latter are usually inferred from kinetic measurements rather than from direct spectroscopic observation. The binding energy of an excimer is usually less than 85 kJ mol^{-1} (20 kcal mol^{-1}) and typically has a sandwich structure to permit optimal π,π^* interaction.

The bimolecular encounter of two excited triplet states, $^3M^*$, can sometimes lead to a singlet excimer, $^1(MM)^*$, in a process called triplet annihilation [reaction (6)]. If the resulting excimer emits, a time delay

$$^3M^* + {}^3M^* \rightarrow {}^1(MM)^* \tag{6}$$

associated with the triplet lifetime is observed and the resulting fluorescence is called P-type delayed fluorescence.

In parallel to the association of an excimer, the encounter between an excited molecule, M^*, and a quencher, Q, can also result in an excited-state complex called an exciplex [reaction (7)]. Exciplex formation is most favorable when the energy difference between the LUMOs of M^* and Q or the HOMOs of M^* and Q are small, namely when the excited state has substantial charge-transfer interaction with the quencher. Like that in excimers, the fluorescence from an exciplex is broad, structureless, and redshifted, and its binding energy can be estimated from a relationship known as the Rehm-Weller equation.

$$M^* + Q \rightarrow (MQ)^* \tag{7}$$

Reaction partners that exhibit strong charge-transfer interactions in the ground state are said to form electron donor-acceptor (EDA) complexes, excitation of which can also yield an emissive exciplex. The presence of a ground-state EDA complex is inferred from the appearance of new redshifted absorption bands or nonadherence to Beer's law upon mixing the electron-rich and electron-poor reagent.

Energy transfer. The process by which an excited state molecule, M^*, in an excited singlet or triplet state transfers all or part of its excitation energy to a reaction partner or quencher, Q, is called energy transfer or quenching when the molecule of interest is M [reaction (8)].

$$M^* + Q \rightarrow M + Q^* \tag{8}$$

This same process is called sensitization when the molecule of interest is Q. In the latter case, M is called the sensitizer. Energy transfer permits an exception to the first law of photochemistry in that Q^* is produced without having absorbed the incident light.

For energy transfer to take place, an incident wavelength must be chosen so that M is primarily excited, producing an excited state M^* whose energy lies above that of Q^*. Symmetry selection rules require that all energy transfer events preserve spin multiplicity. Thus, if M^* is an excited singlet and Q is a ground-state singlet, M will be produced as a ground-state singlet and Q^* as an excited singlet. If M^* is an excited triplet and Q is a ground-state singlet, M will be produced as a ground-state singlet and Q^* as an excited triplet.

There are several mechanisms by which energy transfer can occur. These include (1) trivial or radiative energy transfer, in which the light emitted in the radiative decay of M^* is immediately reabsorbed by Q; (2) coulombic or Förster energy transfer, in which dipole-dipole interactions act over long intermolecular distances (**Fig. 4**); and (3) electron-exchange or Dexter energy transfer, in which two electrons are simultaneously exchanged between M and Q during a collision (**Fig. 5**). The first mechanism requires overlap of the absorption spectrum of Q with the emission spectrum of M^* but does not require the interacting molecules to be near each other. Its

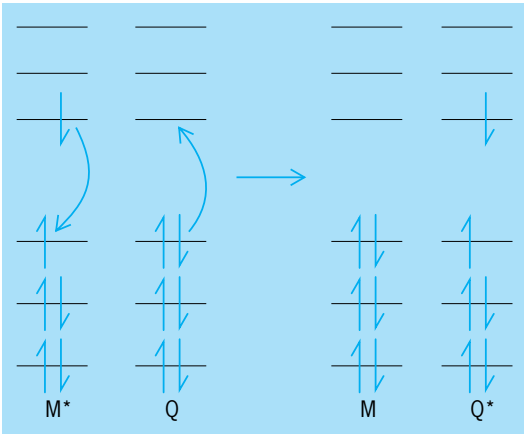

Fig. 4. Dipole-dipole (Förster) energy transfer.

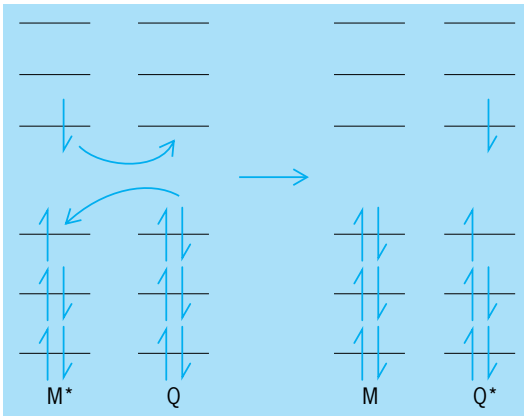

Fig. 5. Electron-exchange (Dexter) energy transfer.

efficiency is governed by those factors and those selection rules that control the oscillator strengths of these two transitions at the common wavelengths. The second mechanism couples M* and Q through a long-range electrostatic interaction. The effect of the oscillating dipole of M* leads to a redistribution of electrons in the HOMOs and LUMOs of the interacting pair, but the electrons of both M* and Q are specifically retained on each molecule (Fig. 4). The magnitude of the dipole-dipole interaction falls off with the sixth power of the distance separating M* and Q. The electron exchange requires a collision in which an electron from the LUMO of M* transfers to the LUMO of Q while an electron from the HOMO of Q simultaneously moves into the vacancy in the HOMO of M* (Fig. 5).

The excited state Q* produced by energy transfer exhibits the same chemical and physical properties as when it is produced by direct light absorption. Sensitization, however, makes it possible to avoid unwanted competing photophysical processes that precede formation of the excited state. For example, higher yields of triplet-derived products can be obtained by a sensitization pathway than are accessible from direct irradiation of a molecule in which the rate of intersystem crossing is low compared with that of radiative or nonradiative relaxation of the singlet state. Therefore, an effective triplet sensitizer will have a small singlet-triplet splitting and a high intersystem crossing yield. Benzophenon [triplet energy = 320 kJ mol^{-1} (76 kcal mol^{-1})] is a commonly used triplet sensitizer.

The triplet energy of a molecule of interest can be determined indirectly by measuring the efficiency of quenching by a series of quenchers with varying, known triplet energies: diffusion-controlled quenching is observed when the triplet energy of the sensitizer is greater than that of the acceptor. The energy transfer can be followed by monitoring the efficiency of reactions derived from the quencher triplet. Some commonly employed triplet quenchers and their sensitized reactions include room temperature solution phase phosphorescence of biacetyl, photosensitized dimerization of 1,3-butadiene, and the geometric isomerization of *cis*-piperylene (*cis*-1,3-pentadiene).

Photoinduced electron transfer. Because photoexcitation promotes an electron to an energetically higher-lying orbital than is occupied in the ground state, both oxidation (removal of the higher-energy electron) and reduction (addition of an electron to the hole created by photoexcitation) are easier in the excited state than in the ground state (**Fig. 6**). That is, the ionization potential (IP), defined as the energy required to remove an electron from the energetically highest-lying orbital to an unbound state, is greater in the ground state than in the excited state; and the electron affinity (ϵ_A), defined at the energy released when an electron is inserted from an unbound state into the energetically lowest-lying molecular orbital, is greater in the excited state than in the ground state. Transfer of an electron from an excited state M* to a quencher Q is called oxidative quenching [reaction (9)], and that from the quencher Q to the

$$M^* + Q \rightarrow M^+ \cdot + Q^- \cdot \qquad (9)$$

excited state M* is called reductive quenching [reaction (10)].

$$M^* + Q \rightarrow M^- \cdot + Q^+ \cdot \qquad (10)$$

The free-energy change associated with these electron transfers can be estimated from free energies of

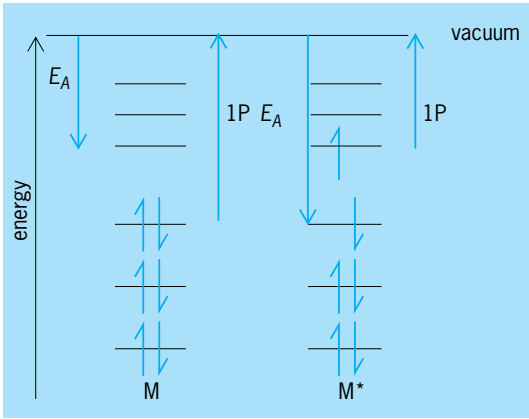

Fig. 6. Diagram showing ionization potentials (IP) and electron affinities (E_A) of ground (M) and excited states, (M*)

the redox half reactions ($M^*/M^{+\cdot/-\cdot}$) and $Q/Q^{-\cdot/+}$). Here the superscripts indicate one-electron oxidation to produce a cation or one-electron reduction to produce an anion. The free energy of the excited-state redox reaction ($M^*M^{+\cdot/-\cdot}$) differs from its ground-state redox potential by the excited-state energy $\Delta E_{0,0}$, as modified by inclusion of a work term that describes the coulombic energy associated with formation of the two ions ($M^{+\cdot/-\cdot}$ and $Q/Q^{-\cdot/+\cdot}$) held at some fixed distance.

Marcus theory describes the relationship between the energy of activation for a bimolecular electron transfer and the thermodynamic driving force expressed as this free energy. In this theory, the nuclear configuration required for electron transfer is assumed to be sensitive to variations in relative vibrational motions of the interacting molecules (inner-sphere reorganization energy) and in the positions, orientations, and induced polarizations of the surrounding solvent molecules (outer-sphere reorganization energy). Marcus theory predicts that the reaction rate will increase with thermodynamic driving force until the binding energy becomes large and negative, causing the reaction rate to decrease with further increases in exothermicity. This phenomenon is described as the Marcus inverted region. *See* CHEMICAL THERMODYNAMICS.

The radical cation/radical anion pair produced by photoinduced electron transfer exhibits enhanced reactivity compared with that of their neutral precursors. A wide variety of chemical reactions results from these activated species, making photoinduced electron transfer a principal means by which chemical transformations are initiated by light absorption. High yields of chemical products, however, are obtained from these intermediates only if energy-dissipative back electron transfer is suppressed. *See* REACTIVE INTERMEDIATES.

Excited-state reaction surfaces. There are three modes by which a photochemical reaction can proceed: (1) an adiabatic photoreaction, in which all conversions take place completely on the excited-state surface, producing an excited product that relaxes to the observed ground-state product (**Fig. 7***a*); (2) a diabatic photoreaction, in which the excited-state reactant relaxes directly to the ground state of the product (Fig. 7*b*); and (3) a hot thermal reaction in which the excited-state reactant is converted by internal conversion to a vibrationally excited ground state, which is converted in a thermal transformation to the ground state of the product (Fig. 7*c*). The surface crossing characteristic of an adiabatic photoreaction is facilitated when the energy difference between the two surfaces is minimized: such a minimum is called a reaction funnel, and many photoreactions take place through such funnels.

Photochemical mechanisms. As in all studies of mechanisms of chemical reactions, determining the structure of all products is the first step in the specification of a photochemical reaction. Spectroscopic (nuclear magnetic resonance spectroscopy, electron spin resonance spectroscopy, infrared spectroscopy, mass spectroscopy, x-ray analysis, absorption spectroscopy) and chromatographic (gas, liquid, or thin-layer chromatography) techniques are used to establish product structure and to determine product yields. Monitoring the effect of solvent polarity on reaction rate, the retention or loss of optical activity during the reaction, the positions of isotopic labels, and the success of intermediate trapping experiments can distinguish step-wise chemical reactions (those that proceed through one or more intermediates) from concerted reactions (those that proceed without intermediates). In addition to these mechanistic approaches, the identity and lifetime of the reactive excited state (singlet, triplet, excimer, exciplex, and so forth) and the quantum yields for both product formation and for other competing photophysical processes are required for a full photochemical mechanistic characterization. Time-resolved flash photolysis and pulse radiolysis measurements can, in addition, be sometimes used for direct spectroscopic detection of absorptive or emissive intermediates encountered in a photochemical mechanism, as well as for their kinetic characterization. The addition of specific reactive quenchers or traps, conducting a photoreaction in a low-temperature matrix in which diffusion processes are stopped, and sensitization experiments are effective means for assigning the observed transient absorptions or emissions. The energetics of a well-defined photochemical reaction can be obtained by photoacoustic calorimetry measurements. *See* CHROMATOGRAPHY; MATRIX ISOLATION; PHOTOLYSIS; SPECTROSCOPY.

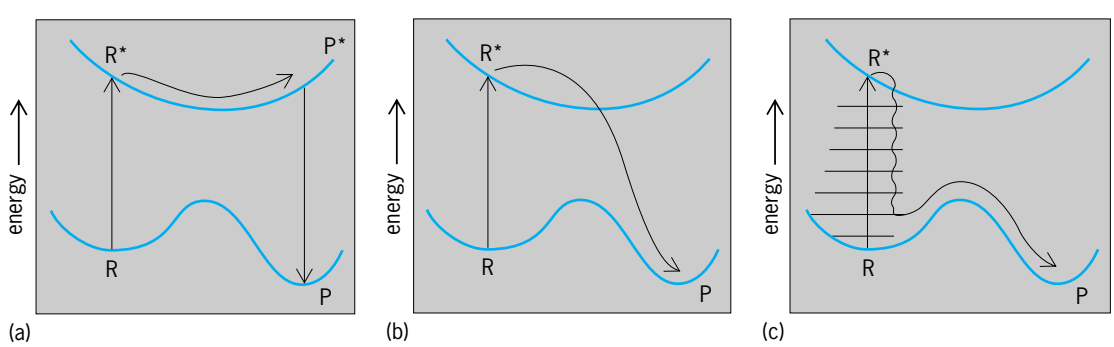

Fig. 7. Modes for photochemical reaction in which reactant R is converted to product P: (*a*) diabatic reaction; (*b*) adiabatic reaction; (*c*) vibrationally excited (hot) ground-state reaction. Excited states are indicated by asterisks.

Concerted photoreactions. There are four common types of concerted photoreactions: (1) cycloaddition reactions, in which the π systems of two separate molecules are joined to produce two new σ bonds in a cyclic product [reaction (11)]; (2) electrocyclic

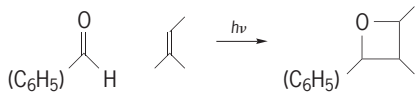

(11)

reactions, in which the ends of a π system within a single molecule are joined by a new σ bond in a cyclic product [reaction (12)]; (3) sigmatropic shifts,

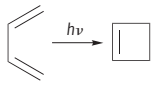

(12)

in which a σ bond migrates along a π system to a new position [reaction (13)]; and (4)

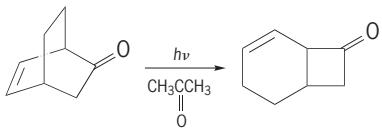

(13)

cheletropic reactions, in which an atom or group of atoms becomes fixed to the ends of a π system [reaction (14)]. The reverse reactions of each of

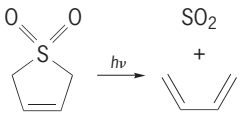

(14)

these types are also concerted and can often be induced by photoexcitation. These reactions are called pericyclic in that they all proceed without intermediates through a cyclic transition state.

The allowedness or forbiddenness of a given pericyclic reaction is dictated by orbital symmetry constraints, expressed in orbital or state correlation diagrams; and the stereochemical sense of allowed pericyclic reactions are defined for molecules containing varying numbers of π electrons by the Woodward-Hoffmann rules. *See* PERICYCLIC REACTION; WOODWARD-HOFFMANN RULE.

Chemiluminescence. Excited states can be produced by thermolysis of highly strained compounds or by highly exothermic electron transfer from a radical anion to a cation. For example, thermolysis of the dioxetane derived from singlet oxygenation of 2,3-dimethylbutane decomposes to two equivalents of acetone, one of which is produced in the excited state [reaction (15)]. The energy released in

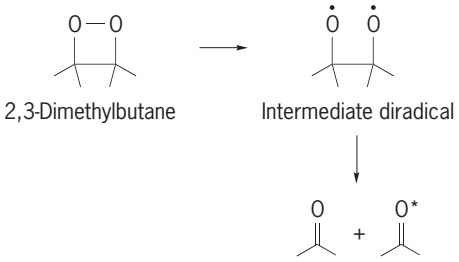

(15)

the cleavage to the intermediate diradical and in the subsequent step to form the two carbonyl groups (—C═O) is substantially more than the triplet energy of acetone. Similarly, back-electron transfer between two electrochemically generated radical ions also releases sufficient energy to produce the luminescent state of the lower-lying singlet. When the radical ions are produced electrochemically, this process is called electrochemiluminescence. *See* CHEMILUMINESCENCE; ELECTROLUMINESCENCE.

Nonhomogeneous environments. Photoreactivity can be substantially affected by the local environment in which an excited state is generated, as has been clearly demonstrated in nature's effect to construct a highly organized reaction center in which the primary events of photosynthesis take place. Since the late 1980s there has been a focus on characterizing altered kinetics and chemical properties that are observed in excited states produced in other than homogeneous solutions in the liquid, gas, or solid state. For example, altered properties are observed when molecules are excited in Langmuir-Blodgett monolayer films, bilayer or multilayer membranes, solid-state crystal lattices, zeolites, normal or reverse micelles, vesicles, colloids, polymers, proteins, enzymes, or antibodies. Similarly, excited-state chemistry is influenced by preadsorption on photoinert surfaces such as silica or alumina or on photoactive ones such as suspended semiconductor particles or metals. For example, environmental detoxification can be accomplished when a pollutant is brought into contact with an excited semiconductor particle. The excited semiconductor often oxidizes the pollutant and reduces oxygen; and in many cases complete breakdown of organic compounds to carbon dioxide, water, and inorganic compounds takes place. Precomplexation with cyclodextrins and other molecular recognition complexants has also been shown to alter chemical reactivity and decay pathways for included molecules. The study of altered photochemical properties induced by noncovalent association with another reagent is called supramolecular photochemistry. *See* CHEMICAL DYNAMICS; INORGANIC PHOTOCHEMISTRY; LASER PHOTOCHEMISTRY; ORGANIC PHOTOCHEMISTRY; PHOTOSYNTHESIS.

Marye Anne Fox

Bibliography. M. Montalti et al., *Handbook of Photochemistry*, 3d ed., 2006; V. Ramamurthy and K. S. Schanze, *Organic Photochemistry and Photophysics*, 2005; N. J. Turro, *Modern Molecular Photochemistry*, 1991; C. E. Wayne and R. P. Wayne, *Photochemistry*, 1996.

Photoclinometer

A term applied to directional surveying instruments which record photographically the direction and magnitude of well deviations from the vertical. Two instruments of this type are in wide use, the Schlumberger photoclinometer and the Surwell clinograph. Both instruments record a series of deviation measurements on one trip into and out of the well. From

this series of data it is possible to plot quite accurately the course of the well.

In the Schlumberger photoclinometer (**Fig. 1**) the deviation from the vertical is indicated by a small metal ball which rolls in a transparent glass bowl graduated in circular degrees. The direction of the deviation in azimuth is indicated by a magnetic compass. With the instrument suspended by an electrical cable, the positions of the compass and steel ball are photographed on a 35-mm film by operation of electrical controls at the surface which turn on lights in the instrument and snap the camera shutter. After the picture is taken, the film is moved to a new position. Pictures can be taken at a rate of about one per minute. Correlation of the pictures with the depths at which they are taken (known by the length of the suspending cable) yields a measure of the magnitude and direction of deviation of the hole as a function of depth.

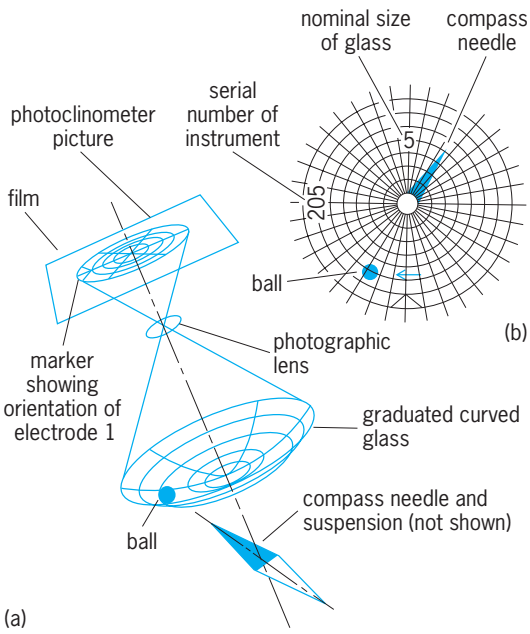

Fig. 1. Schlumberger photoclinometer. (*a*) Principal features. (*b*) Type of record obtained. (*Schlumberger Well Surveying Corp.*)

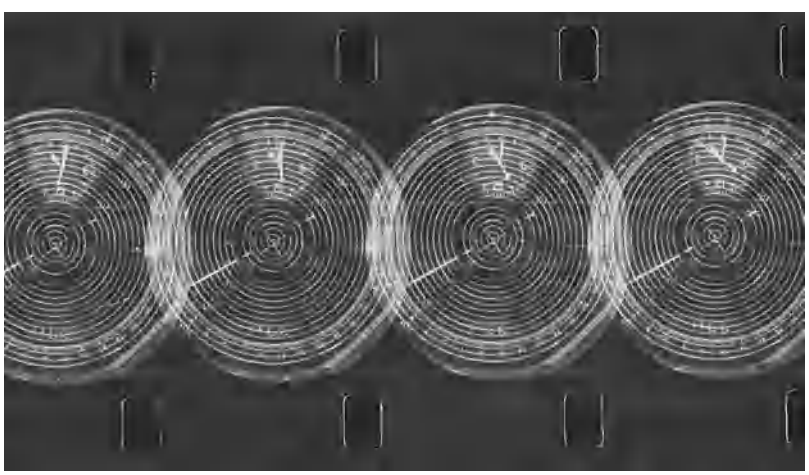

Fig. 2. Motion-picture film records which were made by the Surwell clinograph. (*Sperry-Sun Well Surveying Co.*)

Fig. 3. Vertical section through Surwell clinograph. (*Sperry-Sun Well Surveying Co.*)

The Surwell clinograph (**Figs. 2** and **3**) also operates electrically but is powered by batteries contained in the instrument. The deviation from the vertical is indicated by a box level gage and the direction in azimuth by a gyroscopic compass, permitting its use inside steel pipe. This operation is not possible when a magnetic compass is used, unless the pipe is made of special nonmagnetic steel. Since the instrument also contains a watch and a dial thermometer, a simultaneous record of amount and direction of deviation, temperature, and time can be made on 16-mm film. Readings are taken, both descending and ascending, at regular intervals which are preset on the instrument before it is lowered on a wire line into the well, thus providing a check on accuracy. Level gages having maximum inclinations of 20, 40, and 55°, respectively, are provided to be used according to the magnitude of deviation. Holbrook G. Botset

Photoconductive cell

A device for detecting electromagnetic radiation (photons) by variation of the electrical conductivity of a substance (a photoconductor) upon absorption of the radiation by this substance. During operation the cell is connected in series with an electrical source and current-sensitive meter, or in series with an electrical source and resistor. Current in the cell, as indicated by the meter, is a measure of the photon intensity, as in the voltage drop across the series resistor.

Photoconductive cells are made from a variety of semiconducting materials in the single-crystal or polycrystalline form. There are elemental types such as germanium, silicon, and tellurium; binaries such as lead sulfide, cadmium sulfide, and indium arsenide; ternaries such as mercury cadmium telluride, lead tin telluride, indium arsenide antimonide, and other combinations of elements. The semiconductor detectors respond to photons that exceed a material-related threshold energy. The bolometer is another important detector and consists of a blackened material having a temperature-sensitive conductivity. This device detects photons over a broad spectrum, utilizing the heating effect of photons. The cells are prepared by growth of the semiconducting materials as nearly pure single crystals, or as polycrystalline films deposited chemically or by evaporation on suitable substrates. *See* BOLOMETER.

Cadmium sulfide cells are used in the visible spectrum for street lighting control and camera exposure meters. Lead sulfide and mercury cadmium telluride, sensitive to infrared radiation, are used for night vision. Infrared photoconductive cells are used for detecting energy loss from buildings and for early detection of breast cancer.

The choice of cell type depends on the application requirements, which include operating temperature, wavelength to be detected, and response time. *See* PHOTOCONDUCTIVITY; PHOTOELECTRIC DEVICES.

Sebastian R. Borrello

Bibliography. N. V. Joshi, *Photoconductivity: Art, Science and Technology*, 1990; W. L. Wolfe and G. J. Zissis (eds.), *The Infrared Handbook,* Environmental Research Institute of Michigan, 1978, reprint 1985.

Photoconductivity

The increase in electrical conductivity caused by the excitation of additional free charge carriers by light of sufficiently high energy in semiconductors and insulators. Effectively a radiation-controlled electrical resistance, a photoconductor can be used for a variety of light- and particle-detection applications, as well as a light-controlled switch. Other major applications in which photoconductivity plays a central role are television cameras (vidicons), normal silver halide emulsion photography, and the very large field of electrophotographic reproduction. The phenomena related to photoconductivity have also played a large part in the understanding of electronic behavior and crystalline imperfections in a variety of different materials. *See* OPTICAL DETECTORS; OPTICAL MODULATORS; PARTICLE DETECTOR; PHOTOCONDUCTIVE CELL; PHOTOGRAPHY; TELEVISION CAMERA TUBE.

Since the electrical conductivity σ of a material is given by the product of the carrier density n, its charge q, and its mobility μ [Eq. (1)], an increase in

$$\sigma = nq\mu \tag{1}$$

the conductivity $\Delta\sigma$ can be formally due to either an increase in n, Δn, or an increase in μ, $\Delta\mu$. Although cases are found in which both types of effects are observable, photoconductivity ($\Delta\sigma$) in single-crystal materials is due primarily to Δn, with only small effects at low temperatures due to $\Delta\mu$ if photoexcitation decreases the density of charged impurities that scatter charge carriers. In polycrystalline materials, on the other hand, where transport may be limited by potential barriers between the crystalline grains, an increase in mobility $\Delta\mu$ due to photoexcitation effects on these intergrain barriers may dominate the photoconductivity.

The increase in carrier density Δn can be conveniently related to the photoexcitation density f (excitations per unit volume per second) by the simple relation (2), where τ is the lifetime of the

$$\Delta n = f\tau \tag{2}$$

photoexcited carrier, that is, the length of time that this carrier stays free and able to contribute to the conductivity before it loses energy and returns to its initial state via recombination with another carrier of opposite type (that is, electrons with holes, or holes with electrons). In Eq. (2) the photoexcitation term f includes all the processes of optical absorption (excitation across the band gap of the material, excitation from or to imperfection states in the material, generation of excitons that are thermally dissociated to form the free carriers), and the lifetime τ includes all the processes of recombination (free electron with free hole, free electron with trapped hole, free hole with trapped electron). An understanding of the detailed processes of photoconductivity therefore requires a comprehensive understanding of the variation of optical absorption with photon energy, and of the dependence of recombination on imperfection density, capture cross section, photoexcitation intensity, and temperature. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; BAND THEORY OF SOLIDS; ELECTRON-HOLE RECOMBINATION; EXCITON; HOLE STATES IN SOLIDS; TRAPS IN SOLIDS.

Photosensitivity. Although all insulators and semiconductors may be said to be photoconductive, that is, they show some increase in electrical conductivity when illuminated by light of sufficiently high energy to create free carriers, only a few materials show a large enough change, that is, show a large enough photosensitivity, to be practically useful in applications of photoconductors. There are several ways that the magnitude of the photosensitivity can be defined, depending on the application in mind.

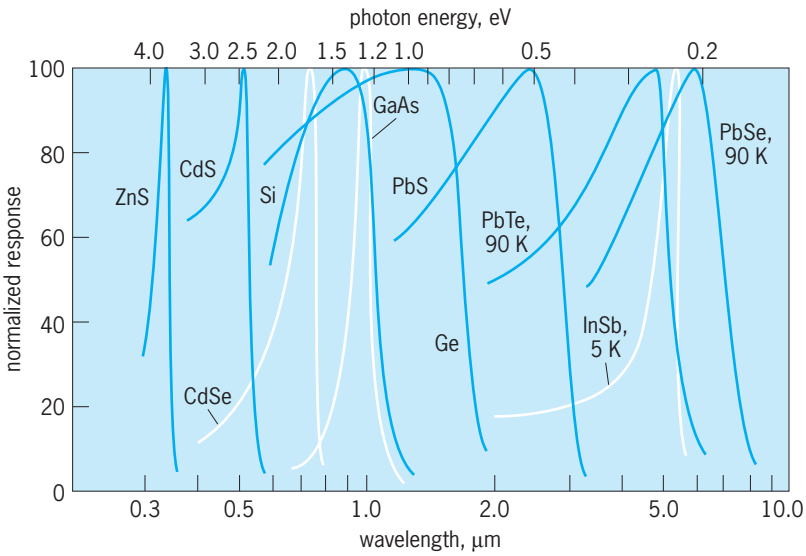

Spectral response of photoconductivity for 10 common photoconducting materials. (*After R. H. Bube, Photoconductivity of Solids, John Wiley and Sons, 1960; reprint, Krieger, 1978*)

Lifetime-mobility product. Comparison of Eqs. (1) and (2) shows that the basic measure of material photosensitivity is given by the $\tau\mu$ product in the common case where $\Delta\sigma$ results primarily from Δn. The mobility does vary from material to material, but in most practical photoconductors μ has values between 10^2 and 10^4 cm^2/V-s at room temperature, and of course the choice of a particular material is usually dominated by its desirable optical absorption characteristics. The free carrier lifetime τ, on the other hand, can take on a wide range of values from 10^{-9} to 10^{-2} s, depending on the particular density and properties of imperfections present in the material.

Detectivity. One of the major applications for photoconductors has been in the detection of small signals in the infrared portion of the spectrum, where the principal objective is to be able to detect the smallest signal possible with the detecting system. Since in this case the photoconductivity is usually much smaller than the dark conductivity, an ac technique is used in which the light signal is chopped and ac amplification stages are used. The limit to detectability is reached when the light-generated signal is comparable to the electrical noise in the photoconductor. Thus the photosensitivity in this particular case is often defined as a detectivity, which is a normalized radiation power required to give a signal equal to the noise. *See* ELECTRICAL NOISE.

Gain. A third device-oriented definition of photosensitivity is that of photoconductivity gain. The gain is defined as the number of charge carriers that circulate through the circuit involving the photoconductor for each charge carrier generated by the light. The time required for a charge carrier to pass through the photoconductor from one electrode to the other, called the transit time, t_r, is given by Eq. (3), where

$$t_r = \frac{L^2}{\mu V} \tag{3}$$

L is the distance between electrodes and V is the applied voltage. The gain is given then by τ/t_r for each type of possible charge carrier, giving Eq. (4)

$$\text{Gain} = \frac{(\tau_e\mu_e + \tau_b\mu_b)V}{L^2} \tag{4}$$

for the total gain if both electrons and holes contribute. Gains of hundreds or thousands can be readily achieved if the lifetimes are sufficiently long. Gains greater than unity require electrical contacts to the photoconductor that are able to replenish charge carriers that pass out of the opposite contact in order to maintain charge neutrality; such contacts are called ohmic contacts. If nonohmic contacts are used, so that charge carriers cannot be replenished, the maximum gain is simply unity, since only the initially created charge carrier contributes to the current flow. Historically, unity-gain currents of this latter type have been called primary photocurrents, whereas high-gain currents described by Eq. (4) have been called secondary photocurrents.

Spectral response. The variation of photoconductivity with photon energy is called the spectral response of the photoconductor. The curves in the **illustration** typically show a fairly well-defined maximum at a photon energy close to that of the band gap of the material, that is, the minimum energy required to excite an electron from a bond in the material into a higher-lying conduction band where it is free to contribute to the conductivity. This energy ranges from 3.7 eV, in the ultraviolet, for zinc sulfide (ZnS) to 0.2 eV, in the infrared, for cooled lead selenide (PbSe). Photoconductivity associated with excitation across the band gap of the material is called intrinsic photoconductivity. For photon energies smaller than the band gap, the light is not strongly absorbed by the material, and the photoconductivity decreases. For photon energies larger than the band gap, the optical absorption is large, and absorption takes place close to the surface of the material; since the surface has in general more imperfections than the bulk, the carrier lifetime at the surface is generally smaller, and hence the photoconductivity decreases. If the bulk of the material contains a sufficiently high density of imperfections contributing localized levels within the band gap of the material, it is often possible to detect photoconductivity corresponding to optical excitation from an occupied imperfection level to the conduction band or to an unoccupied imperfection level from the valence band of the material. This photoconductivity occurs for photon energies smaller than the band gap, and is called extrinsic photoconductivity.

Speed of response. A third major characteristic of a photoconductor of practical concern is the rate at which the conductivity changes with changes in photoexcitation intensity. If a steady photoexcitation is turned off at some time, for example, the length of time required for the current to decrease to $1/e$ of its initial value is called the decay time of photoconductivity, t_d. The magnitude of the decay time is determined by the lifetime τ and by the density of carriers trapped in imperfections as a result of

the previous photoexcitation, which must now also be released in order to return to the thermal equilibrium situation. If the photoexcitation intensity is high, or the density of imperfections is small, the decay time t_d approaches the lifetime τ as a minimum limiting value. For low light intensities or high imperfection densities, where the density of trapped carriers is much larger than the density of free carriers, the decay of photoconductivity is controlled not by the free carrier recombination rate but by the rate of thermal freeing of trapped carriers, and can be many orders of magnitude larger than the lifetime.

Device forms. Photoconductivity detectors may be made in the form of single-crystal devices or as polycrystalline films, in which the bulk material is homogeneous, or in the form of semiconductor junction diodes. Such junction diodes may be prepared in the form of Schottky barriers, *np* homojunctions, *np* heterojunctions, or more complex *npn* or *pnp* double-junction devices with gain greater than unity. *See* JUNCTION DIODE; MICROWAVE SOLID-STATE DEVICES; SEMICONDUCTOR HETEROSTRUCTURES.

Polycrystalline film photodetectors can be made from a variety of methods involving vacuum evaporation, powder sintering, and chemical solution deposition. The photoconductive behavior can be dominated by quite different effects in different materials systems. For example, the photoconductivity in cadmium sulfide (CdS) films deposited by spray pyrolysis is usually controlled by the modulation of intergrain barriers by photoexcitation, so that $\Delta\mu \gg \Delta n$. On the other hand, standard infrared detecting films of lead sulfide (PbS), deposited from chemical solution, exhibit a photoconductivity for which $\Delta n \gg \Delta\mu$, even though the effects of intergrain barriers are clearly measurable in the mobility.

The television camera vidicon and electrophotography are two applications of photoconductivity in which the device form is dictated by the specific nature of the information processing system involved. In both cases an electrical charge is deposited on one side of a high-resistivity photoconducting material; subsequent illumination of the material increases the conductivity locally and allows the charge to leak off through the material. The local absences of charge are then detected and used to produce or reproduce the original light pattern. The material involved must have special characteristics: it must have a high enough dark resistivity so that the deposited charge does not leak off by ordinary dark conduction, and a high enough photosensitivity so that the charge will leak off as quickly as desired. In the vidicon the charge is deposited by scanning by an electron beam; absence of charge is detected by current flow in a subsequent scanning by the beam. In electrophotography the charge is deposited by a corona discharge; absence or presence of charge is fixed by a subsequent printing process. Richard H. Bube

Bibliography. R. H. Bube, *Photoconductivity of Solids*, 1960, reprint 1978; R. H. Bube, *Photoelectronic Properties of Semiconductors*, 1992; N. V. Joshi, *Photoconductivity: Art, Science and Technology*, 1990; J. Mort and D. M. Pai (eds.), *Photoconductivity and Related Phenomena*, 1976; A. Rose, *Concepts in Photoconductivity and Allied Problems*, 1963, reprint 1978.

Photocopying processes

Processes that use light to generate copies directly from original paper documents. Light is employed to examine an original document and to detect the presence or absence of an image. This examination provides the intelligence required to make each copy. In most cases, light reflected from the original subject directly exposes the medium used to produce an image on a copy. However, in an increasing number of cases, the intelligence-bearing light is converted into an electrical signal, which is later converted back to light to expose the copying medium.

Xerography, the most popular of the photocopying processes, relies on photoconductors and toner powders to create copies. The toner is fused to copy paper by heat or pressure. Microfilming employs silver halide films which must normally be developed and fixed by using wet chemicals. Electrofax, diazo, thermography, and nonmicrofilm forms of silver halide photocopying, although not used as widely as in the past, also have continuing applications.

Various photocopying processes can reproduce black-and-white and color document originals, containing text, line, halftone, and continuous-tone images. In creating copies, users often can elect to reduce or enlarge the size of original document images. In some cases, contrast and exposure controls permit creation of photocopies with image quality superior to original documents.

However, photocopying processes represent just one technology for document duplication. Printing processes, which can produce many copies of a document from one exposure, may be more economical for long-run work. Electronic printers offer another route to paper copies of text and graphic information. However, while such printers yield paper images resembling photocopies, digital input sets them apart from photocopiers which receive their input directly from a paper original. The emergence of dual-personality imaging systems that accept both optical and digital input is making it increasingly difficult to characterize a given piece of equipment as a photocopier or electronic printer. A number of electronic printers equipped with scanner attachments can function as either, and some employ both photocopying and digital imaging processes in generating an individual copy.

This melding of photocopying and printing technologies is expected to continue. Increased future reliance on document storage media such as optical disk plus projected growth of multipurpose local-area networks in offices may hasten the entry of hybrid, multipurpose imaging systems. *See* LOCAL-AREA NETWORKS; OPTICAL RECORDING.

Xerography. There are two classes of xerographic copiers: low-volume units, often called personal or

convenience copiers, and high-speed, high-volume units, located in business offices and reproduction centers. Recirculating document feeders, finishers for stapling and binding, sorters, and other peripherals may be added to high-volume units to simplify and speed operations in delivering copy sets. In addition, microprocessor controls enhance their reliability and convenience. Many high-volume units also offer duplexing, the creation of double-sided copies from single-sided originals, in a single machine operation. In addition, some copiers can produce color copies of color originals. *See* MICROPROCESSOR.

The major components of a typical high-volume xerographic copier (**Fig. 1**) include the photoconductor, a primary charger, and systems for exposure, toning, transfer, erasing, and cleaning.

Standard process. The process begins at the primary charger. In this machine, the photoconductor belt is charged to −600 volts. The photoconductor retains the charge on its surface for significant periods of time if it is not exposed to light. Corona charging is the most common charging method. A small-diameter wire is stretched across the photoconductor approximately $^1/_4$ to $^1/_2$ in. (0.64 to 1.3 cm) above its surface. To charge negatively, the wire is raised to a negative potential of several thousand volts, creating a strong electrical field that separates the air molecules into negative and positive ions. The negative ions are repelled from the negatively charged wire to the photoconductor layer which is coated on a conducting surface. The conducting surface is normally held at ground potential.

When the photoconductor belt has moved into the exposure position, the flash lamps illuminate an original document positioned on the platen. With a black-and-white document, black image areas do not reflect light, but white nonimage areas do. The optical system focuses this reflected light on the negatively charged photoconductor, and where light strikes the photoconductor the negative charges are erased. Only areas corresponding to document images retain negative charges.

At the toning station, positively charged toner particles are wiped across the photoconductor surface and are attracted to the negatively charged areas. The greater the amount of negative charge in each area of the photoconductor, the greater the amount of toner attracted to it. This principle allows the creation of copies with gradations of gray tones approaching those of the original.

Next the photoconductor's negative field is relaxed by the postdevelopment erase, and the copy paper is brought into contact with the photoconductor. A heavy negative charge is laid down on the back of the paper, by the transfer charger, to attract the positive toner particles to the paper.

The paper then separates from the film and is carried to the hot roller fuser which softens the toner particles and presses them into the paper surface. Meanwhile, the photoconductor is erased, cleaned, and returned to the primary charger, where it is ready to start another cycle.

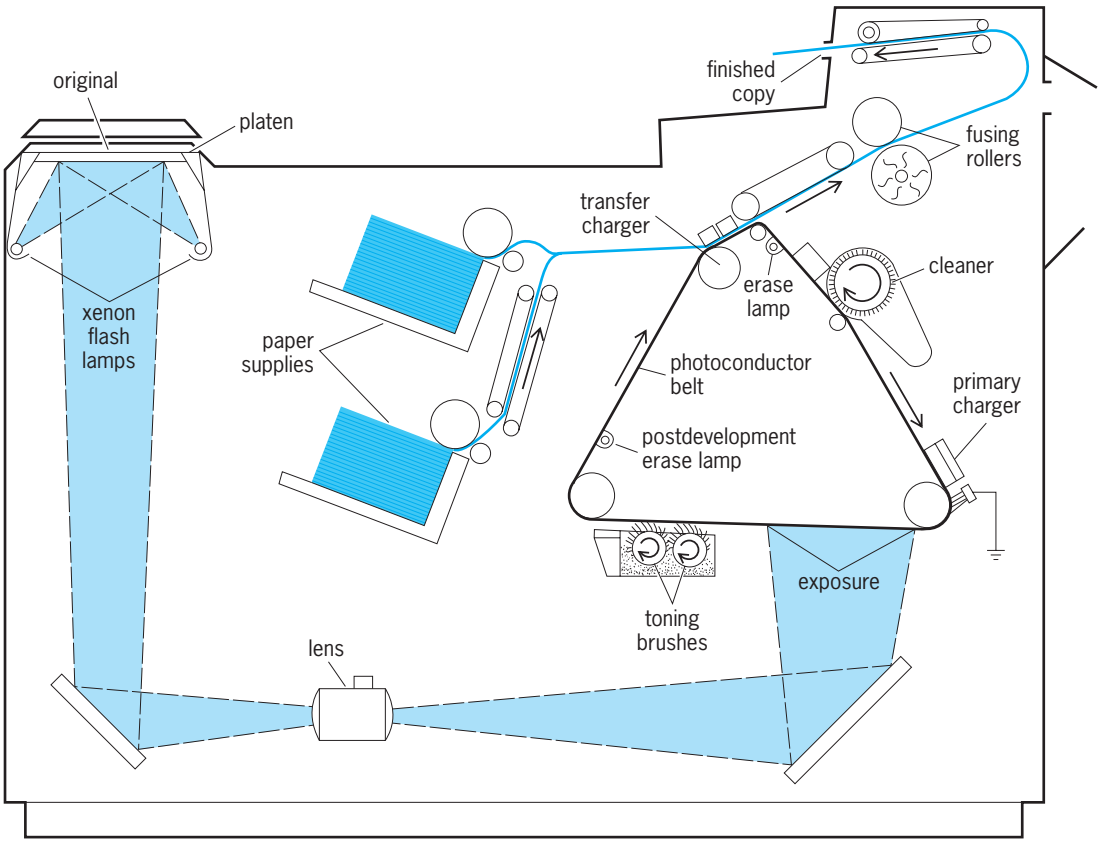

Fig. 1. Diagram of a copier-duplicator. (*Kodak*)

Process variations. Commercially available machines feature a number of process variations. For instance, some systems use a photoconductor which performs better if it is positively charged. In this case, a negatively charged toner is used. Most organic photoconductors, such as those incorporating zinc oxide and cadmium sulfide, are charged negatively. In contrast, most photoconductors relying on silicon or selenium and its alloys are charged positively.

Some machines transport the photoconductor on a rigid cylinder (**Fig. 2**) instead of using a flexible belt. This simplifies the mechanics of the photoconductor system, but poses optical system complexities. To focus light on a revolving drum, a scanning optical system is needed to project one section of the original at a time onto the appropriate cylinder section. This is done during a continuous exposure by synchronized lamp and scanner movement. Since a belt design can present an extended flat portion, or a virtual frame, for imaging, a simpler optical system can project an entire document on a frame in one brief exposure. A drum must be used for many inorganic photoconductor materials which develop artifacts if they are flexed.

Use of a scanning or stationary optical system is strongly influenced by the choice of photoconductor configuration. Both systems may use a variety of lenses, or lens and mirror configurations, to project the image of the document on the platen onto the photoconductor. Auxiliary lenses of varying focal lengths or zoom lenses are used to project smaller or larger images for copy reductions or enlargements. However, in a scanning system, the scanner's speed must be altered relative to photoconductor rotation speed any time there is a change in copy reduction or enlargement requirements.

Toning systems. Most high-volume machines rely on dry, two-component toning systems. They feature black-pigmented toner particles that form copy images and reusable carrier particles that provide their transport. Chemicals are added to the toner and carrier particles to control the magnitude and polarity of the electrical charging action. Roller-type magnetic brushes move the iron carrier particles and piggybacking toners across the photoconductor surface. Toner-carrier interaction gives toner particles their electrical charge. Toner is added automatically as required to maintain the proper toner-carrier mix.

Many low-volume copiers use monocomponent toning systems. In such systems, very small magnetic particles are incorporated into the toner particles themselves. Rotating magnets in the brushes directly attract the toner particles to wipe them across the photoconductor.

Most monocomponent toners also include agents such as carbon to increase their electrical conductivity. When the brush brings them near a charged image on the photoconductor, the electrical field of the image induces an opposite charge in the conducting toner and the toner is attracted to the image.

Liquid-dispersed toners are used in some machines. Very small toner particles are suspended in an insulating hydrocarbon liquid where they take on an electrical charge. The photoconductor is toned by bringing it into contact with the liquid. Various techniques are used to minimize the amount of liquid

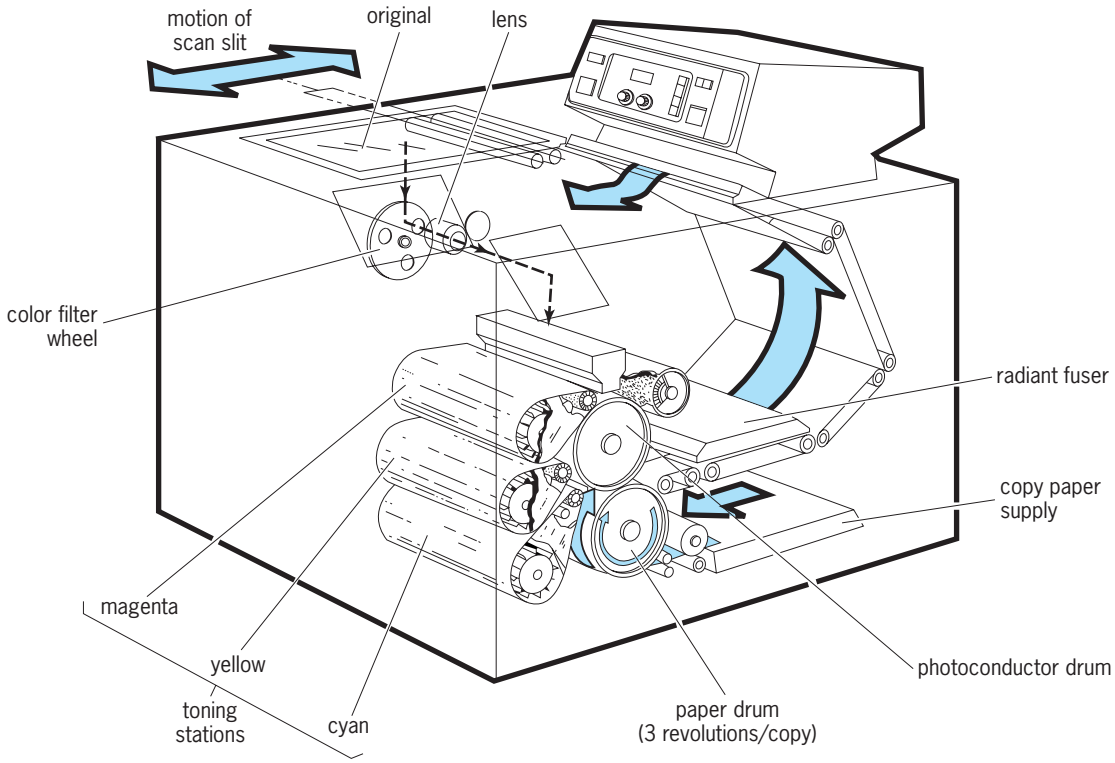

Fig. 2. Diagram of a color copier. (*Xerox Corp.*)

which is transferred to the photoconductor with the toner particles. The liquid is usually a highly refined, low-volatility, petroleum distillate. Drying and fixing are normally accomplished by heating the paper.

The heated fusing rollers used with dry toner processes are covered with silicone rubber or Teflon to prevent the toner from sticking to their surfaces. A thin layer of silicone oil often is wicked onto the surface as further insurance against sticking.

Some machines use unheated, hard-surfaced fusing rollers which rely on pressure alone to fix the toner to the paper. The design is simpler, but the resulting copies are glossy and the process can cause changes in the thickness and stiffness of the paper used for copying.

Two types of cleaners are used to remove untransferred toner from photoconductors. Photoconductors with a hard surface can be scraped with a plastic wiper blade. Rotating soft fur brushes are used in other machines. With either technique, the reclaimed toner can be collected in a waste container or recirculated into the toning system.

Color systems. Color xerographic copiers use the same process steps as black-and-white units. However, exposing and toning operations must be repeated three times to produce each of the three overlaid colors on every copy (Fig. 2). Blue, green, and red filters are rotated in front of the lens to make separate exposures for the original document's light reflectance of the primary colors. Then, toner particles of yellow, magenta, and cyan hues are used to develop the resulting photoconductor latent images.

Normally, the three color images are generated in sequence, and are transferred in register onto one sheet of paper. Since this requirement impacts machine speed, color copiers can produce only about one-third as many copies as a corresponding black-and-white copier using the same process velocity. This speed-volume penalty, the extra complexity entailed in color photocopiers, and additional material requirements make color copies considerably more expensive than black-and-white copies.

Microfilming. This is the only black-and-white, silver halide photocopying method used on a large scale to reproduce document originals (**Fig. 3**). John Danzer is credited with making the first microphotograph of a document in 1839, after substituting a form of a microscope lens for his camera's normal lens.

With source microfilming, original paper documents are imaged on very small areas of silver halide film. Fine-grain films and the special lenses in source document microfilmers allow documents to be filmed at reduction ratios of up to 250:1, although 50:1 is the upper limit for most commercial machines. So, for example, a document microfilmer operating at a reduction ratio of 50:1 is capable of imaging more than 25,000 $8^1/_2 \times 11$-in. (22 × 28-cm) documents on a single roll of 16-mm microfilm contained in a retrieval magazine.

While some source microfilmers use roll film, others use microfiche, or index-sized film cards (typically 105 × 148 mm or 4.1 × 5.8 in.). If paper documents are discarded after microfilming, users can obtain large file space savings. The microfilm permanence makes this photocopying process popular for archival and document searching purposes.

Since it is impossible to read text reduced to this size on a microfilm frame, projection display units are needed to enlarge the microfilmed images for viewing. When such viewers are combined with printers, they are called microfilm reader-printers. These units may incorporate xerographic or Electrofax printers, or may use printers that employ heat to develop coated papers containing silver behenate. Many microimage terminals not only project microfilmed images for viewing but offer automated computer-assisted retrieval capabilities.

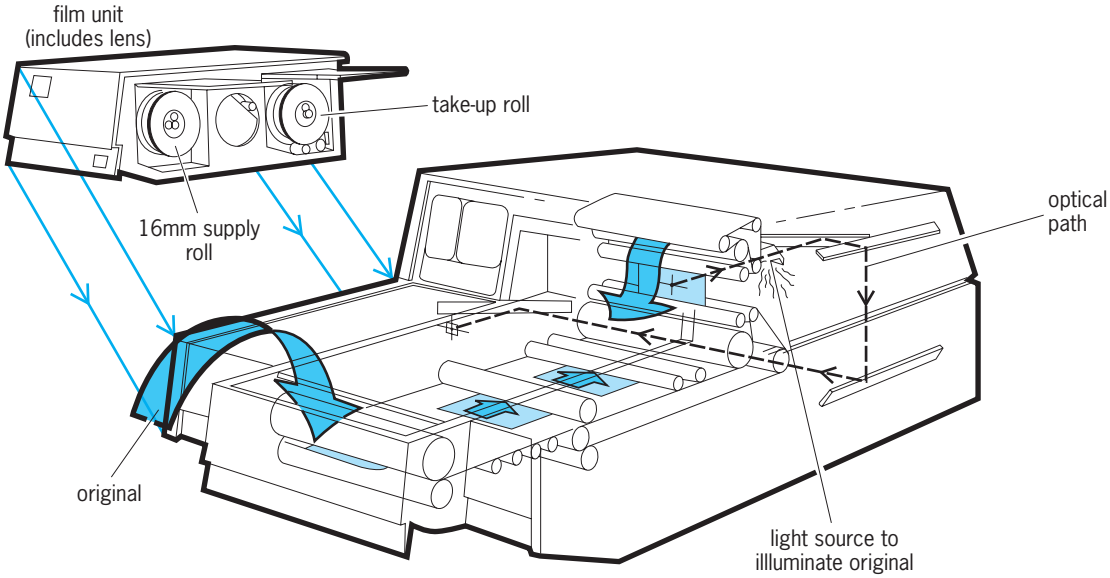

Fig. 3. Diagram of a Kodak Reliant 800 rotary microfilmer.

Computer-assisted retrieval systems link microimage terminals to mainframe or minicomputer systems. The addresses of microfilmed documents are stored in computer memory and may be accessed readily from a cathode-ray-tube computer terminal. Image marks placed on the film during microfilming enable the microimage terminals to search for the requested film frames automatically. Some microimage terminals directly linked to computers totally automate such searches. The speed of accessing microfilm photocopies enhances microfilm's value as an inexpensive mass storage medium.

Progress in digital input microfilm technology parallels electronic printing innovation. Computer output microfilmer devices are, in fact, electronic printers which translate digital input into light to create images on film rather than paper. Combination microfilmers/scanners allow images to be stored on optical disk, CD, and/or magnetic disk, in addition to microfilm, and digital workstations allow analog-to-digital microfilm conversion.

Electrofax. The Electrofax process is a simplified xerographic process which eliminates the need for a reusable photoconductor and its associated transfer, cleaning, and erase stations (**Fig. 4**). The photoconductor is an integral part of every sheet of copy paper.

The photoconductive layer on the Electrofax paper is nontoxic zinc oxide dispersed in a flexible resin binder. The paper is negatively charged by a corona charging method. After exposure to light reflected from a document, toner is applied. In early models the toner was dispersed in a liquid and was fixed by evaporating the liquid, often with a heat assist. Later, dry toner particles were applied by a magnetic brush and were heat-fused to the paper.

The need for specially coated papers eventually led to declining use of Electrofax units. The heavy paper tended to mark, did not feel like plain paper, and was more expensive to mail. However, some manufacturers have adapted the Electrofax process to lithographic platemaking, using large process cameras.

Diazo. Contact copying of documents onto chemically treated papers and films is a photocopying process that predates electrophotographic methods by decades, having been developed in the 1920s, and continues to find application niches. For example, engineering firms and other organizations that need to copy large drawings, maps, and other oversize documents continue to use diazo copies or whiteprints.

Diazo copies are produced by placing the original document in direct contact with a diazo paper or film and then exposing it to an ultraviolet light source (**Fig. 5**). The paper is coated with a diazonium compound and an activator. The exposed paper usually is developed by ammonia vapors or anhydrous ammonia gas.

With equipment needs limited to a large light box and a strong light, the process is inexpensive. However, only single-sided documents on a rather thin translucent material can be reproduced, and copy images also are somewhat unstable.

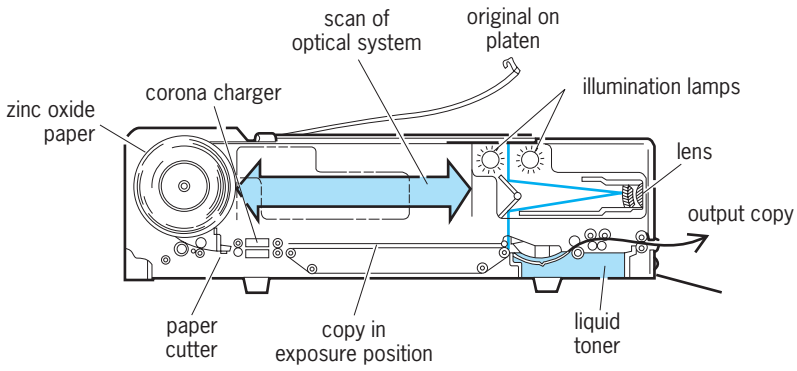

Fig. 4. Diagram of an Electrofax copier.

Diazo predecessors include blueprint and brownprint copy methods. The blueprint first appeared in 1842. After this blueprint paper, sensitized with a ferric salt and potassium ferricyanide, is contacted to an original and exposed to a high-intensity light, it is developed in water and dried. Brownprints, intermediate negatives for platemaking, were first used in 1895. Ferric iron and silver salts make the paper light-sensitive. The exposed paper is washed in a hypo solution to remove undeveloped silver.

Thermography. Thermography is a dry, heat-based photocopying process that was introduced in 1949. Updated versions of the process have continuing applications.

The process uses infrared radiation to reproduce an original on a thin, translucent paper coated with a plastic that contains materials that chemically react and darken when subjected to heat. When infrared rays are directed at the original document, they are absorbed by carbon-containing image areas and converted into heat. The heat is then conducted to the treated paper, which turns black in the radiated image areas.

A process limitation is that original document inks must contain carbon or metallic elements, since

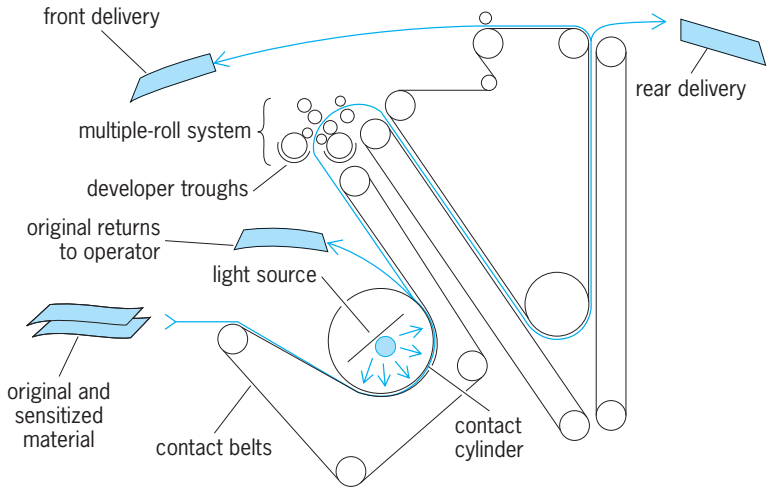

Fig. 5. Flow diagram of whiteprint machine, involving a diazo process. (*Charles Bruning Co.*)

organic inks are less efficient in converting radiant energy to heat. Also, the thin paper tends to curl and offers limited tone gradation capabilities.

In 1961 the dual-spectrum process was introduced. It combines photographic (light) exposure with thermal (heat) development. With this process, a thin chemically treated intermediate sheet receives a latent image and transfers it by heat to the final copy paper, where a chemical reaction renders it visible. These inexpensive copiers are used by individuals with limited copying needs.

Silver halide. Over the years, many photocopying processes have used the high (super) light sensitivity and versatility of silver halide materials, which produce high-quality archival copies. However, the relatively high cost of silver-based photographic materials, coupled with wet chemistry processing demands, has limited the value of this technology for routine office copying applications. Except for microfilming, black-and-white silver halide photocopying processes are limited to lithographic platemaking. *See* PHOTOGRAPHIC MATERIALS; PRINTING.

Robert I. Edelman; J. Gordon Jarvis; Thomas Destree

Bibliography. P. M. Borsenberger and D. S. Weiss, *Organic Photoreceptors for Xerography*, 1998; L. B. Schein, *Electrophotography*, 2d ed., 2001; E. M. Williams, *The Physics and Technology of Xerographic Processes*, 1992.

Photodegradation

Reduction in the useful properties of materials because of chemical changes resulting from the absorption of light. The chemical changes can include bond scission (especially of the molecular backbone), color formation, cross-linking, and chemical rearrangements. All organic materials can photodegrade, but the process has greatest practical relevance for polymers where scission of the polymer backbone is particularly important. Photodegradations of polymers in the absence of oxygen (photolysis) or using wavelengths shorter (more energetic) than those at the Earth's surface (<280 nanometers) have been studied extensively, but only the more practical situation of polymers exposed to terrestrial sunlight (or its equivalent) in air is discussed in this article.

Rate of degradation. Although all organic polymers can be degraded by light, the rate of degradation varies enormously from polymer to polymer, and is also dependent on the incident wavelengths. Light containing ultraviolet (uv; shorter-wavelength) components is much more destructive than visible light, so that polymers exposed indoors, behind window glass (transmitting >330 nm), will degrade much more slowly than samples exposed outdoors. The sensitivity of various polymers to direct sunlight (wavelength 280 nm) is shown in the **table**. Some aromatic polymers such as aramids absorb light in the visible region (380 nm) and are degraded even by these long wavelengths. In general, most aromatic polymers absorb in the uv regions of terrestrial sunlight (~280–380 nm), whereas all pure, aliphatic polymers do not. Nevertheless, the latter may still photodegrade rapidly in practice when exposed outdoors (weathering conditions). This results from the presence of impurities introduced during polymerization, processing, or storage, impurities which do absorb in the near uv and can trigger the degradation of these "nonabsorbing" polymers.

Mechanisms. The sequence of events following absorption by a light-absorbing group (chromophore) in a polymer is shown in the **illustration**. For many aromatic polymers, such as polyester and the aramids, in which the polymer itself is the chromophore, backbone scission results predominantly from this direct absorption of light energy. For many other polymers, including polyolefins, and poly(vinyl chloride) where only impurities absorb energy from sunlight, scission of a chemical bond by light to give free radicals is followed by reaction of these highly reactive free radicals with atmospheric oxygen. Oxygen can diffuse quite readily through the noncrystalline (amorphous) domains of a polymer and combine with radical sites to promote an oxidative chain reaction. This chain process generates oxidation products but also repeatedly regenerates the reactive radical center. The process may directly cause

Weathering lifetimes and absorption limits of common polymers

Polymer	Outdoor lifetime,* years	Cutoff, nm†	Absorption relative to solar limit‡
Polyethylene	0.5–1.0	180	Nonabsorber
Polypropylene	0.2	180	Nonabsorber
Poly(vinyl chloride)	0.5	220	Nonabsorber
Poly(methyl methacrylate)	>20	240	Nonabsorber
Polyamides	3–4	240	Nonabsorber
Polystyrene	0.1§	270	Nonabsorber
(Terrestrial solar limit)		(>280)	
Polycarbonate	0.5	280	Inherent absorber
Polyurethane (MDI-based)	2	280	Inherent absorber
Poly(ethylene terephthalate)	3	310	Inherent absorber
Aramid [poly(paraphenylene terephthalamide)]	0.3	350	Inherent absorber

*Unstabilized, normal impurity levels. Time to 50% loss in tensile, elongation, or impact properties.
†Wavelength at which absorbance of 10-micrometer film reaches 1.0.
‡Solar limit 280–285 μm depending upon location.
§Yellowing.

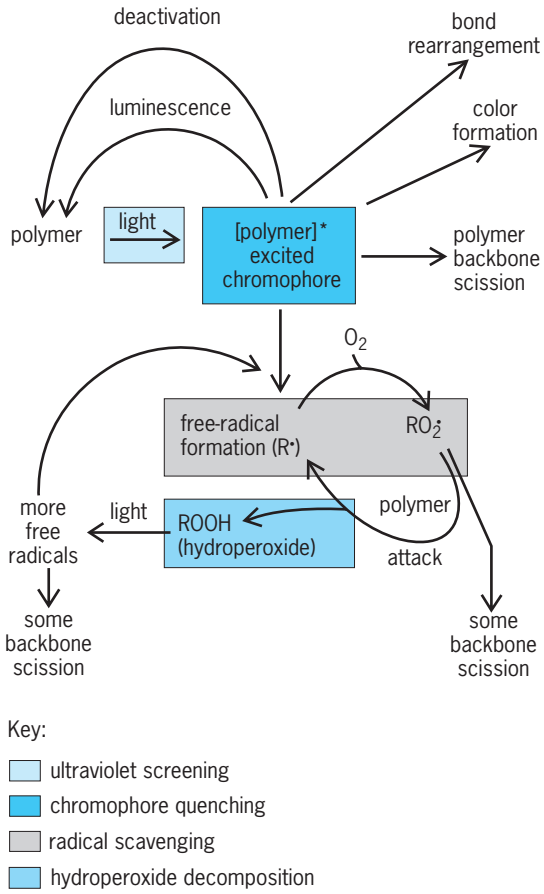

Mechanisms of polymer photodegradation (pathways) and photostabilization (boxes).

some backbone scission, but in addition the main oxidation products (hydroperoxide groups) are themselves subsequently destroyed by light to trigger further oxidation and backbone scission (see illus.). Thus a low concentration (1–100 ppm) of light-absorbing impurity groups can trigger extremely long-chain oxidations to produce a host of photolabile products. *See* FREE RADICAL.

Vulnerable materials. Although numerous organic materials will undergo photodegradation, hydrocarbon polymers are particularly vulnerable because their useful properties depend entirely on their high molecular weights, in the tens or hundreds of thousands. Anything that reduces the molecular weight of polymeric systems will alter the characteristics of these systems and limit their service life. In fact, the scission of as few as one carbon-carbon bond in a thousand in a polymer molecule can completely destroy its useful physical properties. This sensitivity is not observed in lower-molecular-weight substances such as liquid hydrocarbons.

Polymers which lack C-H links and which contain bonds of high strength such as C-F in polytetrafluoroethylene are extremely stable to photodegradation. Ironically, among commercially significant polymers, those with chromophores in the repeat unit, poly(ethylene terephthalate), polycarbonate, and such, tend to photodegrade less rapidly than those that formally have no chromophores and should be transparent to terrestrial sunlight, for example, the polyolefins, polystyrene, and poly(vinyl chloride) [see table]. In the former case, although enormous quantities of photons are absorbed during exposure to sunlight, efficient photophysical processes including fluorescence and phosphorescence (see illus.) harmlessly dissipate the absorbed energy without bond scission occurring. Furthermore, even when bond scission does occur, the rigidity of the aromatic polymers aids backreaction to reform the initial bonds before oxygen interception or rearrangement can occur.

The photodegradation of polymers is by no means uniform throughout a molded article or an extruded fiber or film. In the case of a strongly self-absorbing polymer, much of the incident sunlight is absorbed near the "front" surfaces so that fewer of the damaging photons are left to photooxidize the polymers at the back surface. Light absorption in "nonabsorbing" polymers will occur heterogeneously as well because the chromophoric impurities are not located randomly throughout such solids. Furthermore, since neither oxygen nor impurity chromophores are found in the very highly ordered (crystalline) regions of solid polymers, photodegradation occurs in the amorphous zones. Finally, owing to the restrictions to mobility (high viscosity) in solid polymers, the propagation of oxidizing species is restricted to relatively small distances away from the original sites of photon absorption. *See* PHOTON.

Photostabilization. To prevent or reduce the photodegradation of "nonabsorbing" polymers, exclusion of oxygen and other precautions can be taken during synthesis and thermal processing to reduce the number of chromophoric impurities in the finished articles.

A more universal approach to reducing the rates of photodegradation for all types of polymers, however, is the use of low levels of additives. These additives, known as photostabilizers or uv stabilizers, are effective at fractions of a weight percent. These nonpolymeric compounds are normally added to resin pellets prior to fabrication. They can be classified into four categories according to their modes of action (see illus.). (1) Ultraviolet screeners (absorbers or opaque particles) reduce the amount of incident light reaching the chromophores. This is believed to be the most effective practical method of protecting polymers which absorb strongly in the near uv and are damaged by the primary absorption process. (2) Chromophore "quenchers" remove by collision the electronic excitation energy from a chromophore after the absorption of a photon but before an irreversible chemical reaction can occur. (3) Hydroperoxide decomposers react with polymeric hydroperoxide groups, which are both photochemically and thermally unstable, before the relatively weak oxygen-oxygen bond is broken. (4) Radical scavengers react with (deactivate) radicals so that they cannot perpetuate degradative oxidation of the polymer molecules.

Photostabilizers which operate by hydrogen decomposition and radical scavenging are highly

effective in extending the lifetime of the "nonabsorbing" polymers which degrade primarily by an oxidative chain reaction. However, the more effective uv stabilizers operate by more than one of the four mechanisms. Thus 2-hydroxybenzophenones operate by both uv screening and radical scavenging, and hindered amine light stabilizers operate by both hydroperoxide decomposition and radical scavenging. Combinations of these two types of additives are particularly effective in prolonging the useful life of polyolefin materials during exposure to sunlight. *See* PHOTOCHEMISTRY; POLYMER; STABILIZER (CHEMISTRY).

D. M. Wiles; D. J. Carlsson

Bibliography. N. Allen, *Current Trends in Polymer Photochemistry*, 1995; S. H. Hamid, *Handbook of Polymer Degradation*, 2d ed., 2000.

Photodetachment microscopy

Detachment is the removal of an electron from a negative ion, which produces a neutral atom or molecule and a free electron. Photodetachment is detachment in which the necessary energy is provided by light, more precisely by a photon. Photodetachment microscopy consists in observing accurately the motion of the free electron produced when this process takes place in the presence of an external electric field. In such a field, the electron undergoes constant acceleration, as in free fall, and its possible trajectories are parabolas. With the magnitude of the initial electron velocity set by the photon energy, every point within reach on the detection plane can be reached by two, and only two, such parabolas. The ability of a single electron to follow several trajectories simultaneously is an essential feature of quantum mechanics, which is revealed visually in this case by the observation of an electron interference pattern. The sensitivity of the pattern to variations of the electron's ejection energy makes it possible to measure that energy with an accuracy 1000 times better than by classical techniques. Photodetachment microscopy is used to measure the electron affinities of atoms and molecules with unprecedented accuracy. *See* ELECTRON; ELECTRON AFFINITY; INTERFERENCE OF WAVES; NEGATIVE ION; PHOTON; QUANTUM MECHANICS; SUPERPOSITION PRINCIPLE.

Photodetachment. Since lasers became widely tunable in the 1970s, they have become a popular tool for the excitation of atoms, molecules, and their ions. Photons emitted by a laser are essentially monoenergetic, since the laser wavelength can be set very precisely. This makes it possible to selectively excite the electrons bound in atoms and molecules to higher energy levels. If the photon energy is sufficiently high, it can eject an electron from an atom or molecule to form a positive ion. This process is called photoionization. Negative ions are systems that are weakly bound due to the fact that the outermost electron moves in the short- range field of a neutral atom or molecule. The binding energies of these electrons in negative ions are significantly smaller than those in atoms or molecules. Photodetachment is a process analogous to photoionization that ejects an electron from a negative ion. The energetics of photodetachment is shown in **Fig. 1**. It can be seen that for monoenergetic photons the detached electron has a well-defined energy. *See* LASER; PHOTOIONIZATION.

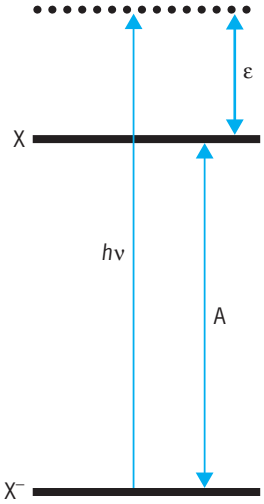

Fig. 1. Energetics of photodetachment. Absorption of a photon of energy $h\nu$ by a negative ion X^- leads to ejection of an electron of energy ε. The detachment energy of the negative ion X^-, which is the electron affinity of the neutral atom X, can be measured as the difference $A = h\nu - \varepsilon$.

Motion of the detached electron. In a typical photodetachment experiment, the total detachment probability is measured as a function of the laser wavelength. Important information can also be determined by measuring the kinetic energies and angular distributions of the detached electrons. In general, the electrons are emitted over a range of angles. In the presence of an electric field, however, the electrons can be guided into a detector. The effect of an electric field is to cause pairs of trajectories to rejoin at the detector (**Fig. 2**). In 1981 the theorists Yu. N. Demkov, V. D. Kondratovich, and V. N. Ostrovskii pointed out that electron interference

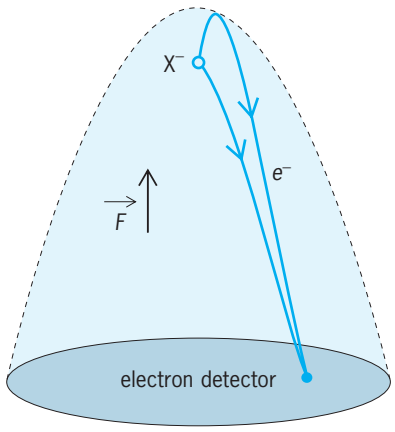

Fig. 2. Trajectories of an electron detached from a negative ion X^- in the presence of an electric field *F*. The detection probability is distributed within a circle on the electron detector. Since every point within this circle can be reached by a pair of trajectories, interference between the trajectories is expected to occur.

fringes, or more precisely rings, should be observable provided a detector with high spatial resolution was used. The authors coined the name "photoionization microscope" for such an apparatus.

Properties. It would appear that making a microscope based on the photodetachment of electrons from negative ions would be easier than making one based on the photoionization of electrons from neutral atoms or molecules. This is because the outermost electron in a negative ion is less tightly bound than in an atom or molecule and can be detached by relatively low-energy photons. Once free, the motion of the detached electron is relatively unperturbed, since it moves in the field of a neutral particle. As a result, the trajectories of the detached electron in the presence of an electric field are the simple parabolas shown in Fig. 2. In contrast, the motion of an ionized electron is more complicated due to the presence of a long-range Coulomb field, and the trajectories of such an electron are no longer parabolas.

There are some practical conditions that must be met in order to observe interference rings in a photodetachment microscope. For example, the interval between adjacent rings must be larger than the resolution limit of the detector. The maximum phase difference must also be large enough to make the number of rings greater than unity. The spectral width of the excitation source must also be sufficiently narrow so as not to wash out the ring interference pattern. The combination of these conditions requires that the detached electron energy be of the order of 100 microelectronvolts and the electric field strength be a few hundred volts per meter.

Figure 3 shows an example of an interference pattern arising from the photodetachment of a negative oxygen ion (O^-) taken under these conditions. The intense outer ring is due to the accumulation of classical trajectories at the maximum radius that the electron can reach on the detector. The inner dark rings and the central dark spot are the interference maxima that are associated with the quantum nature of the electron motion. The deviation from a perfect circular shape is due to inhomogeneities in the electric field. These long-range inhomogeneities do not, however, influence the number of rings in the interferogram.

Qualitatively, the interference patterns show that the coherence of the detached electron wave can survive propagation over macroscopic distances (about 0.5 m or 1.5 ft). The fringes are also among the largest ever obtained using matter waves, with a period up to 0.3 mm (0.012 in.). The measured interferograms are exactly what would be expected in the case of the propagation of an electron wave from a pointlike source in the presence of an electric field. The agreement between the measured data and a theoretical propagator formula gives confidence that the interference patterns can be used for quantitative measurements.

Photodetachment spectrometry. The most important parameter that comes out of the fitting of the experimental data with a propagator formula is the exact energy ε at which the electron is ejected. This energy determines both the radius of the interference pattern (proportional to $\varepsilon^{1/2}$) and the number of interference fringes (proportional to $\varepsilon^{3/2}$). Because of the latter dependence, ε is essentially constrained by the number of rings observed in the interferogram. This is fortunate since the number of rings can be measured more accurately than the absolute radius of the spot. The measurement of a radius entails determining the pixel size in a separate measurement. The energy of the detached electron, ε, can thus be determined with an interferometric accuracy to less than 1 μeV. This makes the photodetachment microscope the most accurate electron spectrometer ever built, as far as the absolute uncertainty is concerned. This technique has been applied to measurements of the electron affinities of fluorine, oxygen, silicon, and sulfur, with an accuracy of a few microelectronvolts. Part of the uncertainty arises in determining the laser wavelength. Thus, photodetachment microscopy is able to accurately measure electron affinities without scanning the wavelength of the laser, which is necessary in the competing technique of laser photodetachment threshold spectrometry.

Photodetachment microscopy also has the advantage of being a remarkably sensitive technique. Electrons produced at energies higher than those of the slow electron detached from the ground-state negative ion are effectively discriminated against. Such electrons could arise from parasitic sources or from internally excited ions. The slow electrons of interest concentrate in an area proportional to ε, which

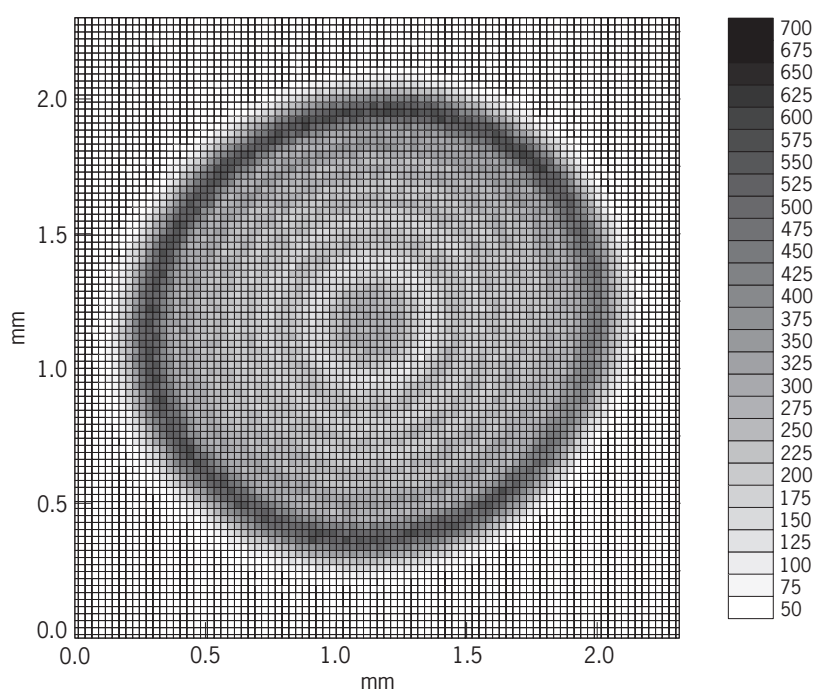

Fig. 3. Interference pattern obtained by photodetaching O^- ions in the presence of a 423-V/m electric field. By counting the observed interference rings, one can determine that the electron exited from the negative ion with a kinetic energy of 162 μeV. The intensity scale on the right indicates the number of electrons received per pixel (after 15 minutes of accumulation). 1 mm = 0.04 in.

more than counteracts the tendency of the detachment cross section to vary as $\varepsilon^{1/2}$ in the vicinity of an s-wave detachment threshold. Faster electrons get spread over a larger surface. The signal-to-noise ratio is sufficiently good that it is possible to detect the electrons detached from minor $^{17}O^-$ and $^{18}O^-$ isotopes and to measure the electron affinities of all their parent atoms, even though the $^{17}O^-$ beam current is not detectable, due to masking by a large unwanted $^{16}OH^-$ ion current.

Molecular photodetachment microscopy. The fact that all the photodetachment microscopy images for atomic negative ions or anions were similar encouraged researchers to extend the same imaging technique to a molecular anion, OH^-. A typical result is shown in **Fig. 4**. With molecular anions there is an additional level of complexity. They can rotate and store internal energy in integral multiples of a quantum of rotation, which is inversely proportional to the moment of inertia of the molecular system. The series of large rings in Fig. 4 corresponds to increases in the rotational quantum number associated with both the initial OH^- and the final OH molecule. Since the moments of inertia of these two systems are almost the same, the detachment threshold is nearly independent of the rotational state, and the rings, which are classical features, can merge together in a band head, such as the one observed as the darkest large ring, about 9 mm (0.35 in.) in diameter, in Fig. 4.

The small spot in the middle of Fig. 4 is analogous to the spot in Fig. 3. All rings contained in the spot are quantum interference rings, which are produced by electrons of the same detachment energy ε.

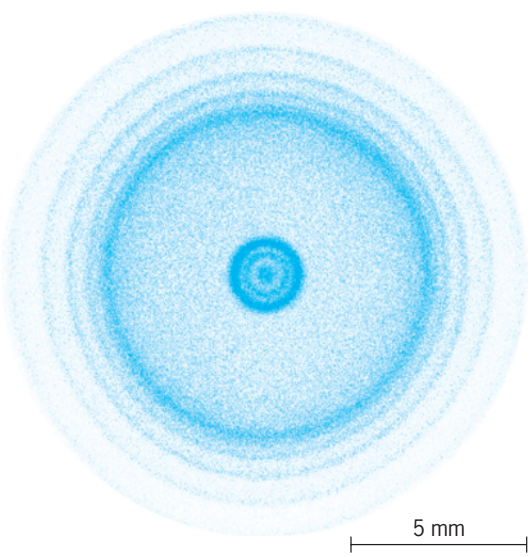

Fig. 4. Photodetachment of OH^- ions in the presence of a 396-V/m electric field. The interferogram of interest is the central spot, which corresponds to the true detachment threshold of OH^- (transition from the ground state of OH^- to that of OH). The interference pattern appears quite similar to the atomic one. The external rings are not interference fringes, but classical rings corresponding to different energies associated with the rotational excitation of the molecule. 5 mm = 0.2 in.

The detachment energy is significantly smaller than that corresponding to the outer rings due to the fact that the rotational quantum number increases by one unit during the detachment process. Remarkably, despite the rotational structure of the molecule and its polar character (it is not known whether this polar nature influences the motion of the outgoing electron), the interference pattern for molecular anions remains similar to that obtained using atomic anions. This similarity has made it possible to measure the electron affinities of OH and SH.

Prospects. Photodetachment microscopy can be applied to many atomic anions whose detachment energies are either unknown or known with accuracies no better than a few millielectronvolts. It is expected that negative ions will continue to serve as unique atomic systems for testing short-range effects, either in collisions or in detachment processes, free of the overwhelming long-range influence of an ionic core. Remarkably, the influence of a weak magnetic field on photodetachment microscopy images appears to be only a spatial shift of the electron spot, without any phase shift. This confirms the robustness of the method and the reliability of the electron affinity measures it provides, even though a residual magnetic field will always be present in a real experiment.

Christophe Blondel

Bibliography. T. Andersen, H. K. Haugen, and H. Hotop, Binding energies in atomic negative ions: III, *J. Phys. Chem. Ref. Data*, 28:1511–1533, 1999; C. Blondel et al., The electron affinities of O, Si, and S revisited with the photodetachment microscope, *Eur. Phys. J. D*, 33:335–342, 2005; C. Blondel et al., The fine structure of S and S^- measured with the photodetachment microscope, *J. Phys. B: At. Mol. Opt. Phys.*, 39:1409–1416, 2006; C. Blondel, C. Delsart, and F. Dulieu, The photodetachment microscope, *Phys. Rev. Lett.*, 77:3755–3758, 1996; C. Bracher et al., Three-dimensional tunneling in quantum ballistic motion, *Amer. J. Phys.*, 66:38–48, 1998; Yu. N. Demkov, V. D. Kondratovich, and V. N. Ostrovskii, Interference of electrons resulting from the photoionization of an atom in an electric field, *JETP Lett.*, 34:403–405, 1981.

Photodiode

A semiconductor two-terminal component with electrical characteristics that are light-sensitive. All semiconductor diodes are light-sensitive to some degree, unless enclosed in opaque packages, but only those designed specifically to enhance the light sensitivity are called photodiodes.

Most photodiodes consist of semiconductor *pn* junctions housed in a container designed to collect and focus the ambient light close to the junction. They are normally biased in the reverse, or blocking, direction; the current therefore is quite small in the dark. When they are illuminated, the current is proportional to the amount of light falling on the photodiode. For a discussion of the properties of *pn* junctions *see* JUNCTION DIODE.

Photodiodes are used both to detect the presence of light and to measure light intensity. *See* PHOTOELECTRIC DEVICES. W. R. Sittner

Photoelasticity

An experimental technique for the measurement of stresses and strains in material objects by means of the phenomenon of mechanical birefringence. Photoelasticity is especially useful for the study of objects with irregular boundaries and stress concentrations, such as pieces of machinery with notches or curves, structural components with slits or holes, and materials with cracks. The method provides a visual means of observing overall stress characteristics of an object by means of light patterns projected on a screen or photographic film. Regions of stress concentrations can be determined in general by simple observation. However, precise analysis of tension, compression, and shear stresses and strains at any point in an object requires more involved techniques. Photoelasticity is generally used to study objects stressed in two planar directions (biaxial), but with refinements it can be used for objects stressed in three spatial directions (triaxial). *See* NONDESTRUCTIVE EVALUATION.

For biaxial studies, a model geometrically similar to the object to be analyzed is prepared from a sheet of special transparent material and loaded as the object would be loaded.

Use of birefringent phenomenon. Model materials commonly used for photoelasticity are Bakelite, celluloid, gelatin, synthetic resins, glass, and other commercial products that are optically sensitive to stress and strain. The materials must have the optical properties of polarizing light when under stress (optical sensitivity) and of transmitting it on the principal stress planes with velocities dependent on the stresses (birefringence or double refraction). In addition, the material should be clear, elastic, homogeneous, optically isotropic when under no stress or strain, and reasonably free from creep, aging, and edge disturbances. *See* BIREFRINGENCE.

When the stressed model is subjected to monochromatic polarized light in a polariscope, the birefringence of the model causes the light to emerge refracted into two orthogonal planes. Because the velocities of light propagation are different in each direction, there occurs a phase shifting of the light waves. When the waves are recombined with the polariscope, regions of stress where the wave phases cancel appear black, and regions of stress where the wave phases combine appear light. Therefore, in models of complex stress distribution, light and dark fringe patterns (isochromatic fringes) are projected from the model. These fringes are related to the stresses. *See* POLARIZED LIGHT.

When white light is used in place of monochromatic light, the relative retardation of the model causes the fringes to appear in colors of the spectrum. White light is often used for demonstration, and monochromatic light is used for precise measurements.

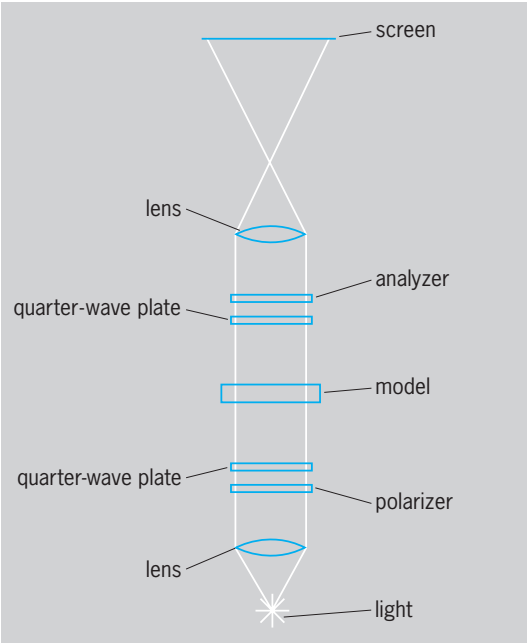

Fig. 1. Basic photoelastic polariscope.

Polariscope. A basic polariscope used in photoelasticity has a light source (generally monochromatic), a collimating lens, a polarizer, and a quarter-wave plate (**Fig. 1**). This plate is a birefringent material that causes the relative retardation of light to be exactly one-quarter the wavelength of the light. Next in the optical path is a planar model of the object under test and stressed in the direction of the plane. Finally there are a second quarter-wave plate, a polarizer called the analyzer, a focusing lens, and a viewing screen or film. Many variations of this basic transmission-type apparatus are in use. Other lenses may be added and components rearranged. If appropriate mirrors are added, the polariscope converts to a doubling type which is useful for the study of thin models under low stress, as the number of fringes doubles.

A typical isochromatic fringe pattern shows the effect on a flat plate with a central hole, pulled at the upper and lower ends (**Fig. 2**). The congestion of fringes at the boundary of the hole indicates a region of stress concentration, typical of stress behavior at cutouts. To study the exact stress at a given point in the model, the model is gradually loaded (from a condition of no load) and the number of fringe changes (fringe order) at that point is observed. Special equipment is sometimes employed to obtain partial fringe orders and to sharpen vague fringe boundaries. The fringe order is directly related by a calibrated constant to the difference of the principal stresses at that point.

Determination of principal stresses. Shear stresses can be related mathematically to the difference of the principal stresses, thereby relating shear stresses directly to fringe order. High shear stresses often cause the material to yield or fail, so that a point of large fringe order indicates a point of potential failure. In many applications of photoelasticity a knowledge of

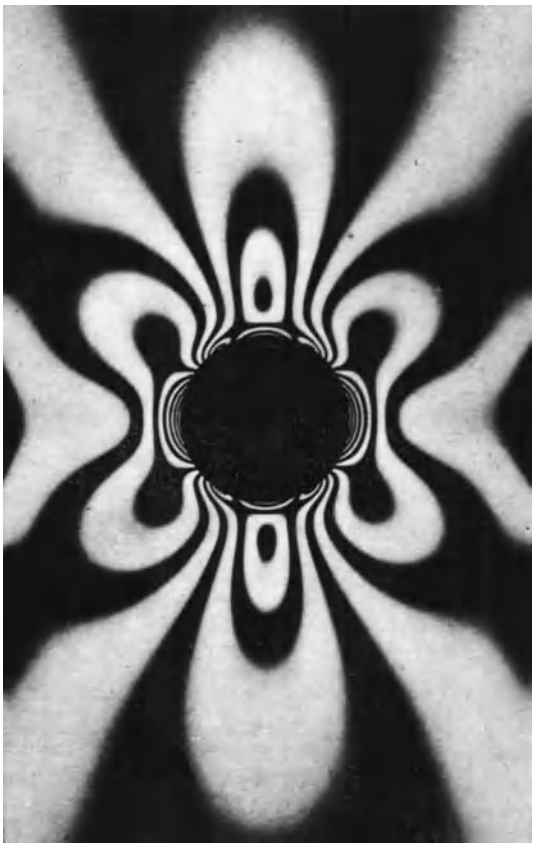

Fig. 2. Isochromatic fringe pattern for plate with hole. (*From M. M. Frocht, Photoelasticity, vol. 2, John Wiley and Sons, 1948*)

shear stress is all that is needed. This fact makes photoelasticity a simple and direct tool for investigation of complex stress systems. However, if principal stresses and their directions are required, additional experimentation is necessary, as described later. *See* FRINGE (OPTICS).

Isoclinic fringes are a different set of interference patterns made by using white light, removing the quarter-wave plates and rotating the polarizer and analyzer a fixed number of degrees. These fringes represent lines making known angles with the principal planes of stress.

Stress trajectories are lines of principal stress directions over the model, obtained graphically from the isoclinic fringes. Stress trajectories are not lines of constant stress.

The determination of principal stresses requires additional information, which may be obtained in several ways. Principal stresses are determined analytically by differences of the shear stresses based on equilibrium equations. This procedure requires a numerical point by point study of the model, utilizing the shear stresses and stress trajectories. Principal stresses can be found analytically or experimentally by solution of Laplace's equation of elasticity. In principle, this procedure supplies equations pertaining to the sum of the principal stresses at any point in the model. Utilizing equations for the difference of the principal stresses from the isochromatic fringe orders, the stresses may be found by solving the two equations for the two principal stresses. Principal stresses are found experimentally by measuring the thickness of the model under stress. Because thickness changes caused by the Poisson's ratio effect are minute, a sensitive measuring device such as an optical interferometer is needed, although direct-reading thickness gages are sometimes used. The interferometer produces fringe patterns called isopachic fringes. This method essentially provides information regarding the sum of the principal stresses as with Laplace's equation. Another experimental method is to pass polarized light obliquely to the surface of the model. The relative retardations of the light produce interference fringes. These oblique fringe orders can be related to the principal stresses differently from those obtained by isochromatic fringes. Using the information on stresses from the isochromatics in conjunction with the oblique relations, the principal stresses may be obtained. *See* ELASTICITY.

With care, stresses determined by photoelasticity are 98% accurate. With stresses determined, strains may be computed by elastic relations.

Three-dimensional measurements. Three-dimensional photoelasticity is also possible, although the techniques and stress-strain relationships are more involved than for planar objects.

The frozen stress method is well suited for three-dimension studies. Certain optically sensitive materials, such as Bakelite, when annealed in a stressed condition retain the deformation and birefringent characteristics of the initially stressed state when the load is removed. A three-dimensional model may therefore be cut into slices that may then be analyzed individually on somewhat the same principles as are the planar models.

The scattered light technique may also be used for three-dimensional models, such as torsion bars. The scattered light principle is based on the fact that polarized light passing through birefringent materials scatters in a predictable manner, acting as an optical analyzer in a polariscope.

Measurements on actual objects. The reflective polariscope method is essentially a variation of the normal polariscope. A sheet of birefringent material is bonded to the polished surface of the actual object to be studied, with a polarizer and quarter-wave plate interposed between the light and the object. The light (usually white light) passes through the stressed birefringent material and is reflected back through the quarter-wave plate and the polarizer (which now acts as an analyzer). Isochromatic fringes are projected as in a normal polariscope. Dependent on good bonding between the birefringent material and the stressed object, the reflective polariscope has the advantage that it can be used on actual structures under service loads without being reduced to model form. It may be used to find surface stresses and strains on curved or three-dimensional objects. Although photoelasticity is generally limited to elastic behavior, reflective techniques may be used to determine the nonelastic strains (but not stresses) of objects, provided the bonded optical material remains elastic. The technique may also be used with

nonheterogeneous materials such as wood and concrete.

Dynamically induced stresses and strains may also be studied by photoelasticity when high-speed motion picture cameras are used to photograph the fringe patterns. Another development in photoelasticity is the use of coherent light to produce holographic interference patterns, instead of the normal fringe patterns. *See* HOLOGRAPHY; STRESS AND STRAIN. William Zuk

Bibliography. J. W. Dally and W. F. Riley, *Experimental Stress Analysis*, 3d ed., 1991; S. A. Paipetis and G. S. Hollister (eds.), *Photoelasticity in Engineering Practice*, 1985; P. S. Theocaris and E. E. Gdoutos, *Matrix Theory of Photoelasticity*, 1979.

Photoelectric devices

Devices which give an electrical signal in response to visible, infrared, or ultraviolet radiation. They are often used in systems which sense objects or encoded data by a change in transmitted or reflected light. Photoelectric devices which generate a voltage can be used as solar cells to produce useful electric power. The operation of photoelectric devices is based on any of the several photoelectric effects in which the absorption of light quanta liberates electrons in or from the absorbing material. *See* PHOTOVOLTAIC CELL; SOLAR CELL.

Photoconductive devices are photoelectric devices which utilize the photoinduced change in electrical conductivity to provide an electrical signal. Thin-film devices made from cadmium sulfide, cadmium selenide, lead sulfide, or similar materials have been utilized in this application. Single-crystal semiconductors such as indium antimonide or doped germanium are used as photoelectric devices for the infrared spectrum. The operation of these devices requires the application of an external voltage or current bias of relatively low magnitude. *See* PHOTOCONDUCTIVE CELL.

Photoemissive systems have also been used in photoelectric applications. These vacuum-tube devices utilize the photoemission of electrons from a photocathode and collection at an anode. Photoemissive devices require the use of a relatively large bias voltage. *See* PHOTOEMISSION.

Many photoelectric systems now utilize silicon photodiodes or phototransistors. These devices utilize the photovoltaic effect, which generates a voltage due to the photoabsorption of light quanta near a *pn* junction. Modern solid-state integrated-circuit fabrication techniques can be used to create arrays of photodiodes which can be used to read printed information. *See* PHOTODIODE; PHOTOTRANSISTOR.

Photoelectric devices can be used in systems which read coded or printed information on data cards and packages. Similar systems are used to sense and control the movement of objects in perimeter guard systems which sense an intruder by the interruption of a light beam. *See* CHARACTER RECOGNITION.

Most visible and ultraviolet photoelectric devices for use in the infrared must be cooled with the longer-wavelength response devices requiring the most cooling. Richard A. Chapman

Bibliography. R. Bube, *Photoelectron Properties of Semiconductors*, 1992; S. Juds, *Photoelectric Sensors and Controls: Selection and Applications*, 1988.

Photoemission

The ejection of electrons from a solid (or less commonly, a liquid) by incident electromagnetic radiation. Photoemission is also referred to as the external photoelectric effect. The visible and ultraviolet regions of the electromagnetic spectrum are most often involved, although the infrared and x-ray regions are also of interest. Photoemission has several distinguishing experimental features: (1) the process of photon absorption and photoelectron generation is instantaneous, that is, there is no detectable time lag; (2) at a given frequency the number of photoelectrons ejected per second is proportional to the intensity of the incident radiation; and (3) the kinetic energies of the photoelectrons depend on the incident photon frequency and the work function of the surface, but are independent of the incident intensity.

It is important to distinguish between the external and internal photoelectric effects. In the case of the external photoelectric effect, the photoelectron is ejected from the surface to the vacuum energy level. It is the fundamental process in photomultiplier tubes and vacuum photodiodes, which are referred to as photoemissive devices. *See* PHOTODIODE; PHOTOMULTIPUER.

In the case of the internal photoelectric effect, the absorbed photon is converted into photogenerated electron-hole pairs in semiconductors. It is the fundamental process enabling semiconductor photodetectors and solar cells. The photogenerated electron-hole pairs in semiconductor photodetectors lead to photoconduction and photocurrent. The electromagnetic radiation of sunlight absorbed by semiconductor photovoltaic solar cells also leads to the creation of electron-hole pairs, which can be used for the generation of electrical power. In addition to photodetectors and solar cells, the internal photoelectric effect plays a fundamental role in charge-coupled devices and image sensors. *See* CHARGE-COUPLED DEVICES; IMAGE TUBE (ASTRONOMY); INFRARED IMAGING DEVICES; OPTICAL DETECTORS; PHOTOCONDUCTIVE CELL; PHOTOCONDUCTIVITY; PHOTOVOLTAIC CELL; PHOTOVOLTAIC EFFECT; SOLAR CELL.

Einstein's photoelectric law. The fundamental physical process of photoemission stems from the duality of light as both waves and particles. One of the most important concepts of the quantization of light into particles called photons emerged from the measurements of blackbody thermal radiation, for which Max Planck successfully provided the correct explanation by using discrete light energy in his analysis

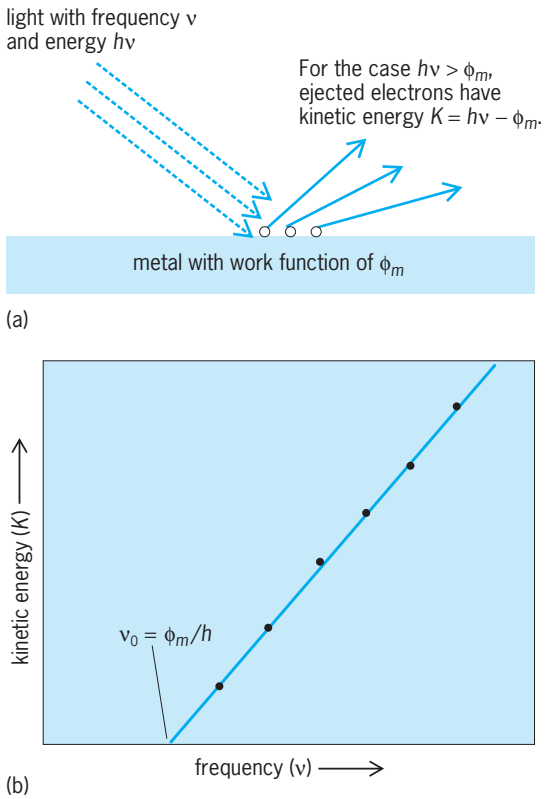

Fig. 1. Photoelectric experiment, (*a*) Photoelectric effect with incident light of energy $h\nu$ and work function of ϕ_m. For the case of $h\nu > \phi_m$, electrons are ejected with kinetic energy of $K = h\nu - \phi_m$. (*b*) Typical relationship of kinetic energy K of electrons with the frequency ν of the incident light. For the case of $\nu < \nu_0$, no electrons are ejected from the surface of the metal. For the case of $\nu > \nu_0$, electrons are ejected with a certain kinetic energy.

in 1900. By extending Planck's idea of quantization of electromagnetic radiation, Albert Einstein in 1905 proposed a theory to successfully explain the photoelectric effect, for which he was awarded the Nobel Prize in Physics in 1921. The photoelectric effect itself was first observed in 1887, but classical physics failed to provide the correct description for the experimental results. *See* DUALITY (PHYSICS); PHOTON; QUANTUM MECHANICS.

A schematic of the photoelectric experiment is shown in **Fig. 1**. It is based on shining incident light onto the surface of a metal and measuring the kinetic energy K of the electrons ejected from the surface of the metal, as shown in Fig. 1*b*.

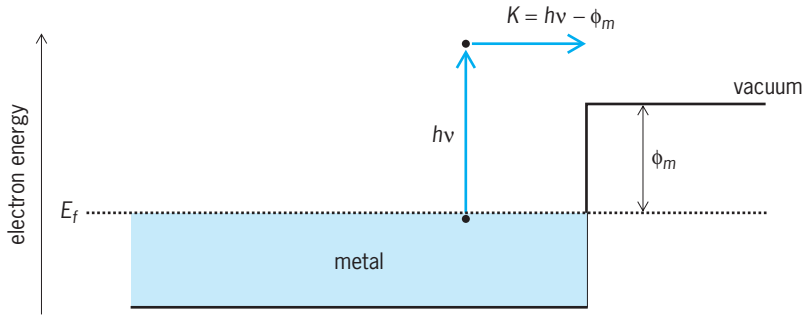

Fig. 2. Energy diagram for external photoelectric effect in metal.

The experiment reveals several important discoveries about the photoelectric effect:

1. If the frequency of incident light is smaller than ν_0, no electron can be ejected from the surface of the metal (regardless of how high the incident light intensity is). Note that $\nu_0 = \phi_m/h$, where ϕ_m corresponds to the work function of the metal and h is Planck's constant.
2. If the frequency of incident light is higher than ν_0, electrons will be ejected instanteously (regardless of how low the incident light intensity is).
3. For $\nu > \nu_0$, the number of ejected electrons is proportional to the light intensity.
4. The kinetic energy K of the ejected electrons depends on the frequency/energy of the incident light, as expressed by $K = h\nu - \phi_m$.

In 1905, Einstein provided the explanation for the photoelectric effect by treating light as a particle with quantized energy of $h\nu$ (in agreement with Planck's work). This quantized light particle is called a quantum or photon. K is the maximum kinetic energy of the ejected electrons; it is also referred to as the Einstein maximum energy. Careful photoelectric experiments by Robert A. Millikan in 1916 led to the measurement of Planck's constant h in the photoelectric relation with considerable precision, and furthered its identification with the constant that Planck had used in his theory of blackbody radiation. *See* HEAT RADIATION; PLANCK'S CONSTANT.

Classifications of photoemission devices. The primary use of photoemission devices is in photodetectors. There are two broad classes of photoemission detectors: metal-based and semiconductor-based. Both the metal- and semiconductor-based photoemission detectors utilize the external photoelectric effect from the surface of the metal and semiconductor, respectively. Electrons are ejected into the vacuum energy level (E_{vac}) from the surface of either the metal or semiconductor. The energy separation of the Fermi energy level of electrons from the vacuum energy level is defined as the work function (ϕ_m), which can be expressed as $\phi_m = E_{vac} - E_f$. *See* WORK FUNCTION (ELECTRONICS).

Metal-based photoemission detectors. In metal-based photoemission detectors, the electron is ejected from the surface of the metal into the vacuum level, if the incident photon energy $h\nu$ is higher than the work function ϕ_m (**Fig. 2**). In a metal, the minimum energy required by the photon to eject an electron is also defined as the photoelectric threshold energy, $\varphi = \phi_m$. The electron energy is located at the Fermi energy level E_f, and the electron should gain enough energy from the absorbed photon to surmount the barrier provided by the work function ϕ_m. The kinetic energy of the electron will then be $K = h\nu - \phi_m$.

As an example of metal-based photoemission devices, **Fig. 3** shows an energy diagram of the electrons in the metal sodium. The work function ϕ_m for metallic sodium is 2.3 eV, which is equal to the photoelectric threshold energy. Inside the metal the electrons occupy a band of energies about 3 eV wide. These electrons are said to be quasi-free. This means

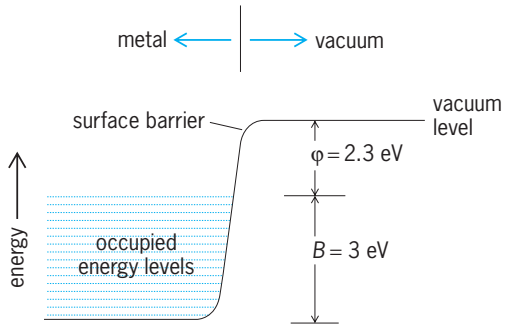

Fig. 3. Energy diagram for electrons in sodium. The photoelectric threshold energy is φ; In a metal it is equal to the electronic work function ϕ_m. The band of energy levels occupied by almost free electrons has a width *B*.

that they behave in many ways like a gas of free, noninteracting electrons; nevertheless, they move in the periodic potential due to the positive sodium ions, and in this sense they are bound. *See* FREE-ELECTRON THEORY OF METALS.

Thus far, the photoelectric threshold has been treated as a sharply defined quantity. This is precisely true for metals only at temperatures near absolute zero. At higher temperatures the upper edge of the band of occupied electron energy states in **Fig. 3** is no longer sharp. It becomes diffuse because of thermal agitation. Electrons may be then emitted for photon energies less than the photoelectric threshold value φ. At ordinary room temperatures, for example, measurable photoemission appears for photon energies as much as 0.2 eV below the threshold. Ralph Fowler developed a convenient graphical technique, known as a Fowler plot, for determining the absolute-zero threshold from data taken at higher temperatures on the spectral dependence of the photoelectric yield, which is the number of photoelectrons ejected per incident photon. Lee A. DuBridge developed a similar technique using either the temperature dependence of the photoelectric yield or the distribution of photoelectrons in energy at fixed frequency. These treatments show that the photoelectric yield is approximately proportional to the quantity $(h\nu - \varphi)^2$ when the photon energy $h\nu$ is within about 1 eV of the photoelectric threshold energy φ. **Figure 4** shows a graph of the spectral dependence of photoelectric yield for some typical emitters. Photoelectric yields from metals are of the order of 10^{-3} electron per incident photon when $h\nu - \varphi$ is 1 eV. **Figure 5** shows typical energy distributions of the photoelectrons from Te and Pt metals, with the same work functions. Photoelectric threshold energies range from 2 eV for cesium to values such as 5 eV for platinum. They vary for different types of crystal faces on the same crystal, and they are exceedingly sensitive to small traces of adsorbed gases.

Semiconductor-based photoemission detectors. In semiconductor-based photoemission devices, the electrons are ejected from the allowed band to the vacuum energy level. The electrons are only allowed to fill either the conduction or valence bands, depending on the distribution of electrons and holes at thermal equilibrium. For semiconductors, it is important to note that the difference between the vacuum energy level and the bottom of the conduction band is referred to as electron affinity (χ, in eV). *See* ELECTRON AFFINITY.

Based on the degeneracy of semiconductors, the following classifications can be obtained:

(1) Nondegenerate semiconductor. In the nondegenerate semiconductor (**Fig. 6***a*), the Fermi energy level E_f lies near the mid gap level, resulting in relatively low electron density in the conduction band. The majority of the electrons are available in the

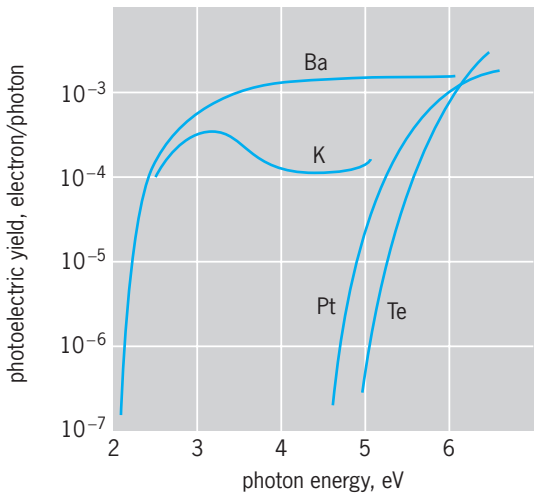

Fig. 4. Spectral distribution of the photoelectric yield from typical samples of barium, Ba; potassium, K; platinum, Pt; and tellurium, Te. Platinum and tellurium have practically the same electronic work function. Note the higher threshold and more steeply rising curve for tellurium, which is a typical elemental semiconductor.

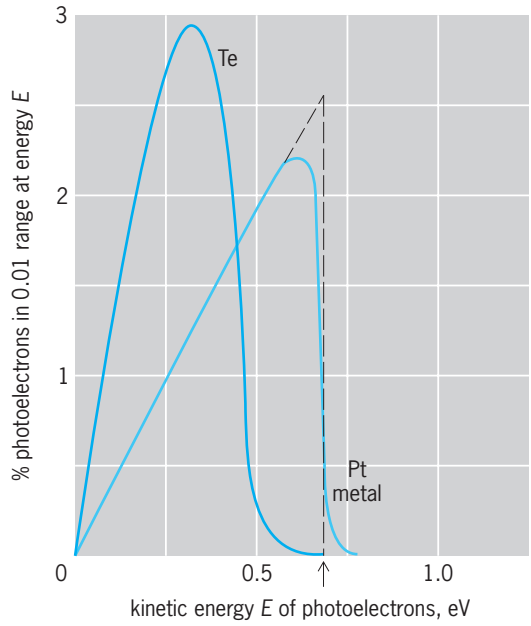

Fig. 5. Energy distributions of photoelectrons from tellurium and a metal having the same work function. The solid curve for the metal shows results for room temperature, and the broken curve shows results for absolute zero. The arrow marks the Einstein maximum energy; photon energy is 5.42 eV.

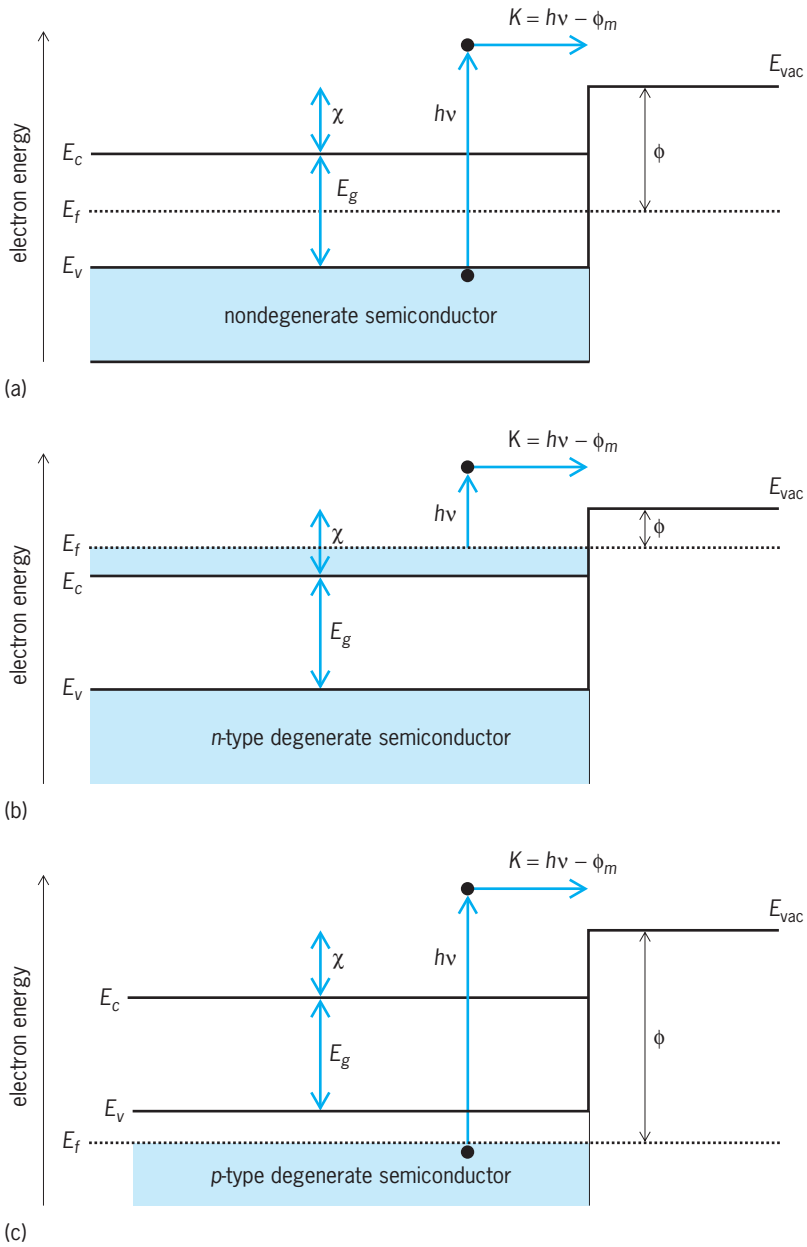

Fig. 6. Energy diagrams. (*a*) For external photoelectric effect in nondegenerate semiconductor. (*b*) For external photoelectric effect in *n*-type degenerate semiconductor. (*c*) For external photoelectric effect in *p*-type degenerate semiconductor.

valence band; thus the photoelectric threshold energy φ can be written as $\varphi = E_g + \chi$, where E_g is the energy gas between the conduction and valence bands.

(2) *n*-Type degenerate semiconductor. In the *n*-type degenerate semiconductor (Fig. 6*b*), the Fermi energy level E_f lies within the conduction band, resulting in a high electron density in the conduction band. The majority of the electrons are available inside the conduction band; thus the photoelectric threshold energy φ can be written as $\varphi = \phi$.

(3) *p*-Type degenerate semiconductor. In the *p*-type degenerate semiconductor (Fig. 6*c*), the Fermi energy level E_f lies within the valence band, resulting in a relatively low electron density in the conduction band. The majority of the electrons reside deep inside the valence band; thus the photoelectric threshold energy φ can be written as $\varphi = \phi$.

The advantage of semiconductor technology is the mature control of the composition and doping level of the semiconductor, and the ability to select appropriate semiconductors ranging from wide-bandgap semiconductors (AlN, GaN) to narrow-bandgap semiconductors (InSb, InAs). In particular for Si, Ge, GaAs, GaN, and InP-based technology, the doping and compositional doping are very mature, thus allowing the accurate control of the material properties for device implementations. *See* BAND THEORY OF SOLIDS; SEMICONDUCTOR.

Alkali halides. Three basically different kinds of photoemission are possible for alkali halides: intrinsic, extrinsic, and exciton-induced photoemission.

Intrinsic photoemission. The intrinsic photoemission process is characteristic of the ideally pure and perfect crystal. It is thus analogous to the emission already described for metals and semiconductors. It appears only for photon energies higher than the intrinsic threshold. For example, potassium iodide, KI, is an alkali halide having this intrinsic threshold in the far ultraviolet near 7 eV. Apparently the width of the forbidden electron energy band in KI is only about 1 eV less than this. The photoelectric yields are high, in excess of 0.1 electron per incident photon, as shown by section C of the curve in **Fig. 7**.

Extrinsic photoemission. The extrinsic photoemission occurs when a KI crystal contains imperfections in the form of negative ion vacancies (lattice sites from which negative iodine ions are missing). These vacancies can be filled by electrons. Color centers, which absorb visible light, are formed. They may reach concentrations as high as 10^{20} per cubic centimeter. External photoelectrons may be ejected

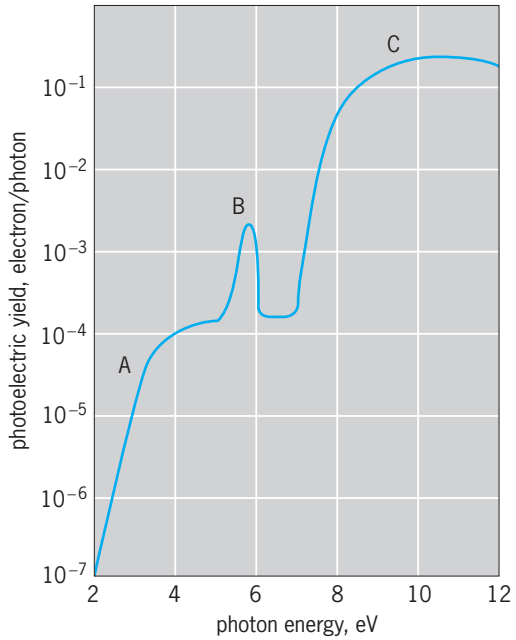

Fig. 7. Spectral distribution of photoelectric yield from potassium iodide containing color centers. Region A of the curve is due to direct ejection of photoelectrons from color centers; the peak B is due to exciton-induced emission; C is due to intrinsic emission.

directly from these centers by photons. It is termed an extrinsic process since the light is absorbed by a crystal defect; it is also called direct ionization. The threshold energy for this process is about 2.5 eV. The yields can reach values of the order of 10^{-4} electron per incident photon, as shown in section A of the curve in Fig. 7. The exact value of the yield depends on the concentration of color centers. Most of the incident radiation is lost because it is not intercepted by the centers, which present a limited cross section to the incident photons. *See* COLOR CENTERS.

Exciton-induced photoemission. When color centers are present, another photoelectric process called exciton-induced photoemission takes place in two stages. Potassium iodide has a sharp optical absorption band, peaking at a photon energy of 5.6 eV. This is the first fundamental or intrinsic optical absorption band. Energy absorbed in this peak does not release free electric charges in the crystal. Rather, it leads to a kind of nonconducting excited state called an exciton state. The exciton can transfer enough energy to color centers to eject photoelectrons from the crystal. This two-stage process is termed exciton-induced photoemission. It appears in the peak B on the curve in Fig. 7. It is more efficient than direct ejection of photoelectrons from color centers. The entire crystal is capable of the primary photon absorption, and the energy can be transferred rather efficiently to color centers. Thus the process avoids much of the loss in incident energy that arises from the limited cross section of color centers when they absorb photons directly. *See* EXCITON.

Other compounds. Other ionic crystals, such as barium oxide, behave much like the alkali halides. Direct ejection of photoelectrons from chemical impurities and from energy levels or defects localized at the crystal surface can be important. In addition to these extrinsic processes, exciton-induced emission and intrinsic photoemission both occur. *See* IONIC CRYSTALS.

Compounds such as zinc sulfide behave somewhat like germanium but have higher intrinsic threshold energies, of the order of 7 eV. The photoelectric yields are comparatively low, as for germanium. Extrinsic processes, such as direct ejection of electrons from chemical impurities (or defects), are sometimes detectable but are usually weak.

A complex photoemitter with high quantum efficiency and low photoelectric threshold energy is cesium-silver-oxide. The cesium-silver-oxide can be formed by letting cesium react with silver oxide to form the compound. The photoelectric threshold energy of this compound is below ~1 eV; thus it is sensitive in the infrared regime. The photoelectrons appear to be directly ejected either from cesium adsorbed on the oxide surface or from discrete energy levels in the cesium oxide. The yields are about 10^{-3} electron per incident photon. Intrinsic emission from cesium oxide (with yields above 0.01) occurs for photon energies above the intrinsic threshold at about 4 eV.

L. Apker; Nelson Tansu

Bibliography. M. Cardona and L. Ley (eds.), *Photoemission in Solids*, 2 vols., 1978, 1979; B. Feuerbacher, B. Fitton, and R. F. Willis, *Photoemission and the Electronic Properties of Solids*, 1978; A. Sommer, *Photoemissive Materials*, 1968, reprint 1980; J.-M. Liu, *Photonic Devices*, Cambridge University Press, 2005.

Photoferroelectric imaging

The process of storing an image in a ferroelectric material by utilizing either the intrinsic or extrinsic photosensitivity in conjunction with the ferroelectric properties of the material. Specifically, photoferroelectric (PFE) imaging refers to storing photographic images or other optical information in transparent lead-lanthanum-zirconate-titanate (PLZT) ferroelectric ceramics.

Imaging devices. The photoferroelectric imaging device consists simply of a thin (0.1–0.3 mm), flat plate of optically polished PLZT ceramic with transparent electrodes deposited on the two major faces. The image to be stored is made to illuminate one of the electroded faces by using near-ultraviolet illumination in the intrinsic photosensitivity region (corresponding to a band-gap energy of approximately 3.35 eV) of the PLZT. Simultaneously, a voltage pulse is applied across the electrodes to switch the ferroelectric polarization from one stable remanent state to another. Images are stored both as spatial distributions of light-scattering centers in the bulk of the PLZT and as surface deformation strains which form a relief pattern of the image on the exposed surface. The light scattering and surface strains are related to spatial distributions of ferroelectric domain orientations introduced during the image-storage process. These spatial distributions replicate the brightness variations in the image to which the PLZT is exposed.

Fig. 1. Photographic image with high resolution and gray scale stored in a photoferroelectric imaging device.

Stored images may be viewed directly or may be projected by using either transmitted or reflected light. For projection, the image contrast is usually improved by using collimated light and a schlieren optical system. *See* SCHLIEREN PHOTOGRAPHY.

Either total or spatially selective erasure of stored images is accomplished by uniformly illuminating the area to be erased with near-ultraviolet light and simultaneously applying a voltage pulse to switch the ferroelectric polarization back to its initial remanent state, that is, to switch it to the polarization state prior to image storage.

The solid-solution PLZT ceramics can be fabricated with a wide range of compositions and sintering conditions to produce an associated wide range of ferroelectric, dielectric, and electrooptic properties. For example, an important factor in determining the maximum resolution of stored images is the ceramic grain size, which can be controlled to a large extent by the sintering temperature and pressure. Images with resolutions as high as 2500 line pairs per inch (100 line pairs per millimeter) can be stored in plates with 2-micrometer grain size, while the maximum image resolution achieved in 5-micrometer grain size plates is about 1000 line pairs per inch (40 line pairs per millimeter). *See* CERAMICS; DIELECTRIC MATERIALS; ELECTROOPTICS; SINTERING.

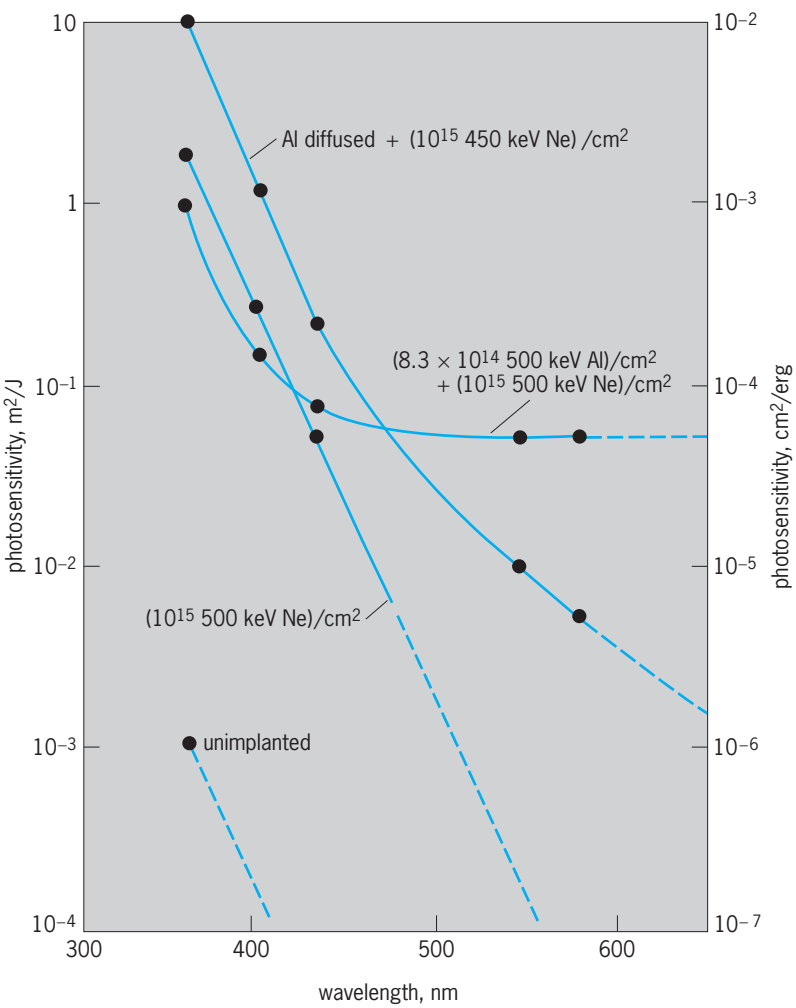

Fig. 2. Wavelength dependence of the photoresponse for aluminum-diffused-plus-neon-implanted PLZT compared to neon-implanted and neon-plus-aluminum-implanted PLZT. The near-ultraviolet response for unimplanted PLZT is also shown.

Important potential applications of photoferroelectric imaging devices include temporary image storage with periodic update, projection-type display, spatial-temporal light modulation, and high-density optical information storage. Various types of image processing, including image contrast enhancement, are also offered by the capability of switching from a positive to a negative stored image in discrete steps. Nonvolatile photographic images with gray scale ranges extending from an optical density of about 0.15 to more than 2.0 and with resolutions as high as 1000 line pairs per inch (40 line pairs per millimeter) are routinely stored in PLZT plates with 5-μm grain size (**Fig. 1**). *See* IMAGE PROCESSING; OPTICAL INFORMATION SYSTEMS; OPTICAL MODULATORS.

Intrinsic PFE effect. The intrinsic photoferroelectric effect is characterized by a photoinduced reduction of the coercive voltage V_C (the externally applied voltage required to switch the ferroelectric polarization from a saturation remanent state to an average value of zero) produced by irradiating the PLZT surface with band-gap or higher energy light. The reduction in V_C occurs because the absorbed near-ultraviolet light photoexcites charge carriers into the conduction state, and these carriers are transported, under the influence of the applied field, to the ceramic underlying the absorption region, where they contribute to ferroelectric domain nucleation and reorientation. Retrapped carriers establish space charge fields which, in combination with the photocurrent, provide the mechanisms for the photoinduced reduction of V_C. In any given localized area of the ceramic surface, V_C is reduced by an amount proportional to the exposure energy W_{ex} (product of near-ultraviolet light intensity and image exposure time) in that area. As a result, the stored images faithfully reproduce the gray scale of the input image. *See* BAND THEORY OF SOLIDS.

Photosensitivity enhancement. The exposure energy W_{ex} required to store photographic images in unmodified PLZT is 1–5 $\times$ 10^3joules/m^2 (1–5 $\times$ 10^6 ergs/cm^2). The relatively high value of W_{ex} (compared to electrophotographic processes or photographic film) tends to limit the scope of practical applications of photoferroelectric imaging devices. Because of the many attractive features of photoferroelectric imaging devices, including erasable and reusable image storage, image contrast enhancement, high-resolution optical information processing, and other optical storage and processing capabilities, various techniques to improve the photosensitivity and to extend the photoresponse from the near ultraviolet throughout the visible spectral region have been explored. Dramatic improvement of the near-ultraviolet photosensitivity, involving reduction of threshold W_{ex} from about 10^3 to 10^{-1} J/m^2 (10^6 to 10^2 ergs/cm^2), has been achieved by implantation of inert ions, such as argon, neon, and helium, and significant photoresponse throughout

the visible spectral region has been achieved by coimplantation of chemically active (for example, aluminum or chromium) and inert ions. It has also been found that the combination of thermal diffusion of aluminum followed by neon implantation yields photosensitivity improvement in the near-ultraviolet and blue spectral regions comparable to or better than that obtained by the ion implantation alone. *See* ION IMPLANTATION.

The primary effect of the surface modifications described above is to decrease the dark conductivity of the near-surface region, relative to that of unmodified PLZT, due to implantation-induced defects. These defect states result from atomic displacements produced by energy deposited in nuclear collision processes. In the surface-modified PLZT, a significant fraction of the externally applied voltage can be dropped in the thin ($\sim 1\ \mu$m), low-conductivity, damaged layer to produce very high drift fields in this near-surface region. Photogenerated charge carriers in this region experience extremely high electric fields and have a high probability of being swept into the underlying ceramic to nucleate and reorient ferroelectric domains. The resulting high photocurrents provide the basis of the ion-implantation enhancement of the photosensitivity.

The photosensitivity S (inverse of W_{ex}) versus wavelength is compared for aluminum-diffused plus neon-implanted and aluminum-plus-neon-implanted PLZT in **Fig. 2**. Also included in Fig. 2 are data for neon-implanted and unimplanted PLZT. Comparison of these data indicates that the major factor contributing to the visible photoresponse is the absorption states associated with aluminum rather than disorder-related defect states produced by the neon implantation. However, the neon implantation is the major contributor to the near-ultraviolet photosensitivity enhancement. This latter conclusion is supported by other published data for PLZT coimplanted with the inert ions argon, neon, and helium which shows that the near-ultraviolet photosensitivity approaches that of aluminum-diffused PLZT, but the visible photoresponse is similar to that of the neon-implanted PLZT.

The dramatic photosensitivity enhancements in both the near-ultraviolet and the visible spectral regions have greatly expanded the variety of possible applications of PLZT as an erasable and reusable photoferroelectric imaging device. *See* ELECTRONIC DISPLAY; FERROELECTRICS; PHOTOCONDUCTIVITY.

Cecil E. Land

Photogrammetry

The practice of obtaining surveys by means of photography. The camera commonly is airborne with its axis vertical, but oblique and horizontal (ground-based) photographs also are applicable. Many industrial and laboratory measurement problems are solved by photogrammetry. Data reduction is accomplished by stereoscopic line-of-sight geometry with use of both analytical and analog methods. *See* AERIAL PHOTOGRAPH; REMOTE SENSING; SURVEYING.

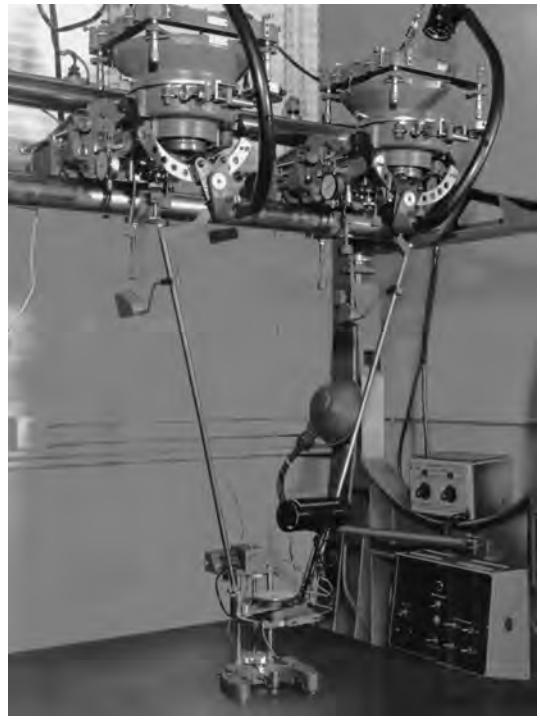

Fig. 1. Stereoplotting instrument, showing projectors and plotting table. (*Kelsh Instrument Co., Inc.*)

In vertical aerial surveys adjacent photos are overlapped. The two images of the same terrain are then superimposed for three-dimensional viewing by human operators or automated sensors.

In a widely used analog procedure the two photos are placed in the projectors of a stereoplotting instrument. With the aid of visible ground-control points the photos are oriented to the relative positions they had at the instants of exposure. Projections are aimed with the aid of space bars at a plotting table (**Fig. 1**), which can be raised or lowered and moved horizontally. The image is in focus at the center of the disk only when it is at the correct scalar elevation for a given horizontal position. The center of the disk is indicated by a small point of light called the floating dot. To trace a contour, the operator sets the table at its scalar elevation; he moves the table laterally until the elevation's focus is encountered by the floating dot, then lowers the plotting pencil and follows the contour by keeping the image in focus at the floating dot. Planimetry is plotted by moving the floating dot along visible lines (roads or building outlines).

In a typical automated stereoplotting system (**Fig. 2**) scanning devices substitute for human eyes to sense model-surface slope and thus to control servomechanisms that raise and lower the plotting table (z dimension) and translate it along a succession of closely spaced parallel horizontal dimensions, or ground profiles. Output is available in three forms: digitized x-y-z-coordinate representation of the terrain, contour plotting, and orthophotographs. Each element of the orthophoto represents a vertical view of an element of the stereoscopic image, thereby

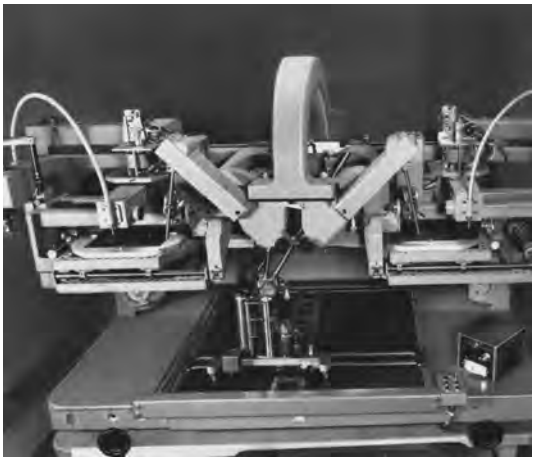

Fig. 2. Typical automated stereoplotting system. (*Lockwood, Kessler, and Bartlett, Inc.*)

eliminating the height-displacement distortion characteristic of normal aerials. Robert H. Dodds

Bibliography. E. M. Mikhail, J. S. Bethel, and J. C. McGlone, *Introduction to Modern Photogrammetry*, 2001; P. R. Wolf and B. A. DeWitt, *Elements of Photogrammetry with Applications in GIS*, 3d ed., 2000.

Photographic materials

The light-sensitive recording materials of photography, that is, photographic films, plates, and papers. They consist primarily of a support of plastic sheeting, glass, or paper, respectively, and a thin, light-sensitive layer, commonly called the emulsion, in which the image will be formed and stored. The material will usually embody additional layers to enhance its photographic or physical properties.

Supports. Film support, for many years made mostly of flammable cellulose nitrate, is now exclusively made of slow-burning "safety" materials, usually cellulose triacetate or poly(ethylene terephthalate) [polyester], which are manufactured to provide thin, flexible, transparent, colorless, optically uniform, tear-resistant sheeting. Polyester supports, which offer added advantages of toughness and dimensional stability, are widely used for films intended for technical applications. Film supports usually range in thickness from 0.0025 to 0.009 in. (0.06 to 0.23 mm) and are made in rolls up to 60 in. (1.5 m) wide and 6000 ft (1800 m) long. *See* ESTER; POLYESTER RESINS.

Glass is the predominant substrate for photographic plates, though methacrylate sheet, fused quartz, and other rigid materials are sometimes used. Plate supports are selected for optical clarity and flatness. Thickness, ranging usually from 0.04 to 0.25 in. (1 to 6 mm), is increased with plate size as needed to resist breakage and retain flatness. The edges of some plates are specially ground to facilitate precise registration. *See* GLASS.

Photographic paper is made from bleached wood pulp of high α-cellulose content, free from ground wood and chemical impurities. It is often coated with a suspension of baryta (barium sulfate) in gelatin for improved reflectance and may be calendered for high smoothness. Fluorescent brighteners may be added to increase the appearance of whiteness. Many paper supports are coated on both sides with water-repellent synthetic polymers to preclude wetting of the paper fibers during processing. This treatment hastens drying after processing and provides improved dimensional stability and flatness. *See* PAPER.

Before an emulsion is coated on the support, a "sub" (substratum) is applied to ensure good adhesion of the emulsion layer (**Fig. 1**).

Emulsions. Most emulsions are basically a suspension of silver halide crystals in gelatin. The crystals, ranging in size from 2.0 to less than 0.05 micrometers, are formed by precipitation by mixing a solution of silver nitrate with a solution containing one or more soluble halides in the presence of a protective colloid. The salts used in these emulsions are chlorides, bromides, and iodides. During manufacture, the emulsion is ripened to control crystal size and structure. Chemicals are added in small but significant amounts to control speed, image tone, contrast, spectral sensitivity, keeping qualities, fog, and hardness; to facilitate uniform coating; and, in the case of color films and papers, to participate in the eventual formation of dye instead of metallic silver images upon development. The gelatin, sometimes modified by the addition of synthetic polymers, is more than a simple vehicle for the silver halide crystals. It interacts with the silver halide crystals during manufacture, exposure, and processing and contributes to the stability of the latent image. *See* EMULSION; GELATIN; SILVER.

After being coated on a support, the emulsion is chilled so that it will set, then dried to a specific moisture content. Many films receive more than one light-sensitive coating, with individual layers as thin as 1.0 μm. Overall thickness of the coatings may range from 5 to 25 μm, depending upon the product. Most x-ray films are sensitized on both sides, and some black-and-white films are double-coated on one side. Color films and papers are coated with at least three emulsion layers and sometimes six or more plus filter and barrier layers. A thin, nonsensitized gelatin layer is commonly placed over film emulsions to protect against abrasion during handling. A thicker gelatin layer is coated on the back of most sheet films and

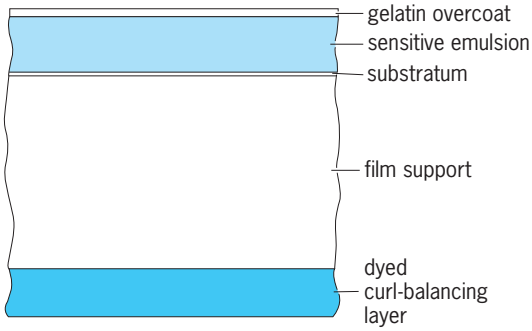

Fig. 1. Schematic cross section of film.

some roll films to counteract the tendency to curl, which is caused by the effect of changes in relative humidity on the gelatin emulsion. Certain films are treated to reduce electrification by friction because static discharges can expose the emulsion. The emulsion coatings on photographic papers are generally thinner and more highly hardened than those on film products.

Another class of silver-based emulsions relies on silver-behenate compounds. These materials require roughly 10 times more exposure than silver halide emulsions having comparable image-structure properties (resolving power, granularity); are less versatile in terms of contrast, maximum density, and spectral sensitivity; and are less stable both before exposure and after development. However, they have the distinct advantage of being processed through the application of heat (typically at 240 to 260°F or 116 to 127°C) rather than a sequence of wet chemicals. Hence, products of this type are called Dry Silver films and papers.

Spectral sensitivity. The silver halides (and silver behenates) are normally sensitive only to x-radiation and to ultraviolet, violet, and blue wavelengths, but they can be made sensitive to longer wavelengths by adding special dyes, predominantly polymethines, to the emulsion. The process is known as spectral sensitizing to distinguish it from the chemical sensitizing used to raise the overall or inherent sensitivity of the grains.

Spectrally nonsensitized emulsions are generally termed blue-sensitive or ordinary and exhibit little response beyond 450 nanometers. Emulsions treated with dyes to extend their sensitivity to 500 nm are identified as extended blue, while those with sensitivity extended through the green (550 nm) are termed orthochromatic. Some panchromatic films are sensitized out to 650 nm; others, with extended red sensitivity, respond efficiently out to 700 nm. The upper limit for spectral sensitizing is about 1200 nm, which is in the near-infrared region, but conventional infrared films exhibit little sensitivity beyond 900 nm. Ultraviolet sensitivity is diminished somewhat between 250 and 280 nm and sharply below 250 nm by the strong optical absorption by gelatin. Materials sensitive to much shorter wavelengths can be prepared by several means which result in a high concentration of silver halide crystals at the top surface of the emulsion layer. Films are selected for spectral characteristics to achieve efficient response to various exposing sources and, whenever possible, to provide convenient handling under safelight conditions. Spectral sensitivity is an important consideration in assuring visually satisfying tonal reproduction of colored objects with black-and-white materials or intentionally altering the tonal reproduction to provide greater spectral discrimination than the human eye. The spectral sensitivity of a recording material is altered, in effect, by placing color filters over the camera lens.

Halation. During exposure of a film or plate, part of the radiation entering the emulsion is absorbed; the rest is transmitted into the support. A small fraction of the transmitted light is reflected at the back surface and returned to the emulsion at a point displaced from the incoming radiation, depending upon the angle of incidence and the thickness of the support (**Fig. 2**). This secondary exposure is called halation because a bright point in a scene is recorded as a point surrounded by a halo. Halation can be suppressed to negligible levels by incorporating carbon black or selected dyes in the support, a layer on the back of the support, a layer between the emulsion and the support, or in rare instances, the emulsion. The absorbing layer need not be opaque to be effective since the offending radiation must make two passes through the absorbing layer and is strongly attenuated upon reflection at the back surface. Carbon black is added to some polyester supports to suppress light-piping as well as halation.

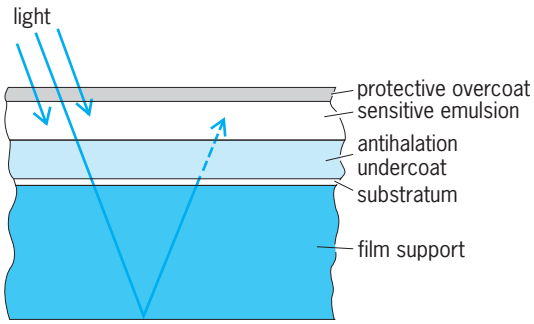

Fig. 2. Suppression of halation by an absorbing medium, shown here as an antihalation undercoat.

Dyes in layers under the emulsion or in curl-balancing layers on the back of some films are bleached or dissolved during processing. The antihalation backings on some plates and films are dissolved or stripped away during processing.

Photographic characteristics. For a discussion of characteristics of sensitive materials, as well as the theory of the photographic process, latent image, sensitometry, and image structure, *see* PHOTOGRAPHY.

All of these characteristics are related to the combination of emulsion, developer, and exposure. The structure of the developed image is of importance in determining the quality of the photograph, especially its ability to reproduce fine detail.

Photographic products. Hundreds of types of films, plates, and papers are available in numerous sizes and configurations for a wide range of applications. Each field of use generally requires special properties. Films for amateur, professional, and technical photography typically range in speed from an exposure index (ISO/ASA) of less than 10 to over 1000, and from very fine grain and high sharpness to rather coarse grain with corresponding loss in definition. They can be obtained in a wide range of contrasts and spectral sensitizations, especially for scientific and technical photography. While most materials are negative-working, a number of products provide positive-working emulsions which result in direct positive reproductions with conventional (negative) processing. Materials of extremely

high contrast and density are widely used for graphic reproduction and in the printing industry to obtain high-contrast line and halftone images. Many types of x-ray films are made, most sensitized on both sides and used in combination with fluorescent intensifying screens to minimize patient exposure. Several types of color film are available, designed primarily for pictorial photography.

Photographic papers are offered in both contact and enlarging speeds and with a range of contrasts for producing continuous-tone prints from camera negatives. A second class of enlarging papers provides selective control of contrast through the use of filters during exposure of two-component emulsions. One component is blue-sensitive only and has high contrast, while the other has sensitivity extended to the yellow-green and relatively low contrast. Papers are also distinguished by surface texture and tint of the paper support. Developing agents are incorporated in the emulsions of some photographic papers so that internally controlled development can be accomplished quickly by immersing the paper in an alkaline bath. Color printing papers are made for printing from color negative or positive films. High-contrast negative and direct-positive papers are used for photographic reproduction of documents, engineering drawings, or line copy. Several papers are designed for oscillographic recording, photographic typesetting, and other specialized applications.

Image-transfer systems. The image resulting from exposure and processing of most photographic materials is retained in the light-sensitive layer or layers. In another and very useful set of products, the image is transferred by chemical means to a receiver sheet which has been placed in contact with the basic photographic material during part or all of the processing step. Careful design of each component in these image-transfer systems generally results in convenient processing and relatively rapid access to the final image. Most widely known are camera-based systems in which continuous-tone silver or dye images are chemically transferred to a receiver sheet or layer from the light-sensitive and sometimes complex image-forming layers. In the original Polaroid-Land process, unexposed silver halide crystals in the emulsion are transferred by a solvent and reduced to form a positive black-and-white image on the receiver in as little as 10 s. Several newer systems form color images through the transfer of dyes to the receiver within 1 to 4 min (**Fig. 3**). These "instant" products are popular for amateur photography but are widely used for professional and technical applications as well, including the photography of cathode-ray-tube images. *See* CAMERA; CATHODE-RAY TUBE.

Other examples of image-transfer systems include materials for producing large color prints or transparencies from either negative or positive color film originals and several high-contrast black-and-white materials used for preparing short-run printing plates. While the image-forming component of most image-transfer systems is usually discarded after being pulled apart from the receiver, it is sometimes retained as a useful film negative. In some instant products intended primarily for amateur photography, all layers are retained as a unit to avoid litter. These integral film structures, including image-forming layers, an image-receiving layer, and chemicals, are handled in sheet form, while two-part structures are offered in both roll (Fig. 3) and sheet configurations.

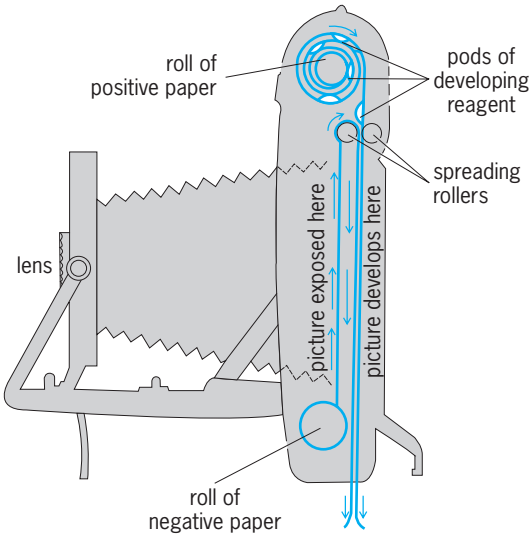

Fig. 3. Cross-sectional view of Polaroid Instant Camera and film. (*Polaroid Corp.*)

Storage of materials. Unexposed films and papers may be damaged by exposure to elevated temperatures and high relative humidities. Protection is provided by packing the materials hermetically at controlled humidities (30–50% relative humidity). Storage at a temperature of 40 to 55°F (4 to 13°C) is sufficient for most materials, but some products require deep-freeze conditions.

The life of processed films and prints depends strongly on processing conditions as well as environmental factors. Thorough fixing and washing are usually a primary requirement. Storage at 30–50% relative humidity and at temperatures not exceeding 70°F (21°C) in carefully selected enclosures is commonly recommended.

Nonsilver processes. Many special-purpose photographic materials use active compounds other than silver salts. In general, they are less sensitive by factors of 100 or more than comparable silver-based products, and they lack the amplification factor associated with the development of silver compounds. Accordingly, they are used mainly in reproduction processes where speed is not essential or in certain recording applications where intense exposures from laser sources are possible. They include inorganic compounds which are photosensitive, especially iron salts and dichromates in colloid layers such as gelatin, glue, albumen, shellac, and poly(vinyl alcohol) or other synthetic resins which become insoluble on exposure. Other important processes are based on light-sensitive organic compounds, especially diazo compounds; unsaturated compounds such as cinnamic acid derivatives, which

are insolubilized by cross-linking on exposure; systems in which polymerization occurs as a result of free-radical formation or exposure; dyes which are formed, bleached, or destroyed directly by exposure to light or electrons; and thermographic systems in which a thermal pattern resulting from intense exposure produces an image by means of melting a composition or initiating a chemical reaction.

Electrophotographic systems rely on the imagewise elimination of a charge by exposure of a uniformly charged photoconductive layer and subsequent development of the remaining charge distribution with charged powders or liquid toners, or by electrolytic deposition of metal on the charge image. Toned images may be fixed by heating so the toner particles are fused either to the photoconductive material or, following physical transfer, to a receiver sheet. Because electrophotographic systems commonly exhibit strong edge effects, they are especially useful in document-copying applications, but continuous-tone systems which rely on careful design of the charging and toning elements have been demonstrated. Many xerographic copiers use coatings of amorphous selenium on rigid supports as the photoconductive layer. The preferred medium in modern copiers of this type consists of an organic photoconductor coated on a flexible support similar to film base. *See* PHOTOCOPYING PROCESSES.

Robert D. Anwyl

Bibliography. N. R. Eldred, *Chemistry for the Graphic Arts*, 3d ed., 2001; T. Grimm, *Basic Book of Photography*, 4th ed., 1997; T. Grimm, *The Basic Darkroom*, 3d ed., 1999; M. J. Langford, *Advanced Photography*, 6th ed., 1998; M. J. Langford, *Basic Photography*, 6th ed., 1997; S. F. Ray, *Photographic Chemistry and Processing*, 1994; S. F. Ray, *Photographic Data (Technical Pocket Book)*, 1994; S. F. Ray, *Photographic Imaging and Electronic Photography (Technical Pocket Book)*, 1994; L. D. Stroebel (ed.), *Basic Photographic Materials and Processes*, 2d ed., 2000; L. D. Stroebel (ed.), *The Focal Encyclopedia of Photography*, 3d ed., 1996.

Photography

The process of forming stable or permanent visible images directly or indirectly by the action of light or other forms of radiation on sensitive surfaces. Traditionally, photography utilizes the action of light to bring about changes in silver halide crystals. These changes may be invisible, necessitating a development step to reveal the image, or they may cause directly visible darkening (the print-out effect). Most photography is of the first kind, in which development converts exposed silver halide to (nonsensitive) metallic silver.

Since the bright parts of the subject normally lead to formation of the darkest parts of the reproduction, a negative image results. A positive image, in which the relation between light and dark areas corresponds to that of the subject, is obtained when the negative is printed onto a second negative-working material. Positive images may be achieved directly in the original material in two principal ways: (1) In reversal processing, a positive image results when the developed metallic silver is removed chemically and the remaining silver halide is then reexposed and developed. (2) In certain specialized materials, all of the silver halide crystals are rendered developable by chemical means during manufacture. Exposure to light alters the crystals so that development will not proceed. As a result, conventional "negative" development affects only the unexposed crystals and a positive image is formed directly.

The common materials of photography consist of an emulsion of finely dispersed silver halide crystals in gelatin coated in a thin layer (usually less than 25 micrometers) on glass, flexible transparent film, or paper (**Fig. 1**). The halides are chloride, bromide, or bromoiodide, depending upon the intended usage. The most sensitive materials, used for producing camera negatives, consist of silver bromide containing some silver iodide; the slow materials used for printing are usually of silver chloride; materials of intermediate speed are of silver bromide or silver bromide and chloride.

Following exposure of a sensitized recording material in a camera or other exposing device such as a line-scan recorder, oscillograph, plotter, or spectrograph, the film or plate is developed, fixed in a solution which dissolves the undeveloped silver halide, washed to remove the soluble salts, and dried. Printing from the original, if required, is done by contact or optical projection onto a second emulsion-coated

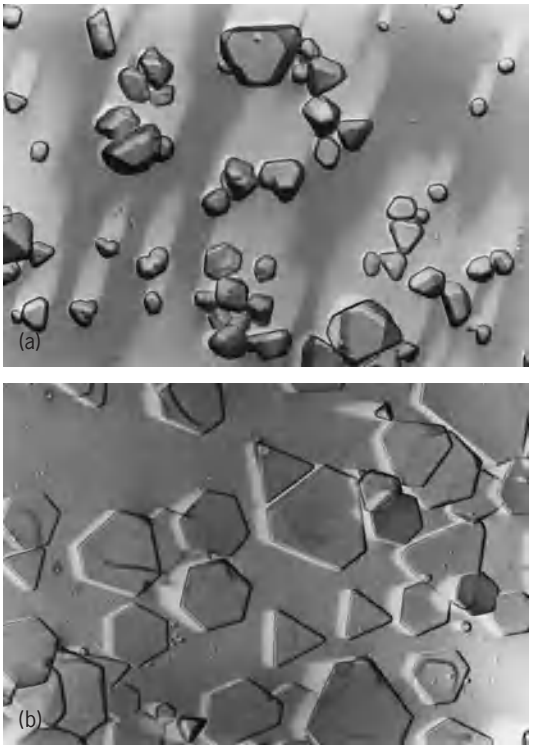

Fig. 1. Silver halide crystals, shown highly magnified. (*a*) Conventional pebblelike crystals. (*b*) Tablet-shaped crystals, used in some modern emulsions, that absorb light more efficiently, resulting in higher film speed.

material, and a similar sequence of processing steps is followed.

For about 100 years, the results of practical photography were almost exclusively in black and white or, more precisely, shades of gray. With the introduction of the Kodachrome process and materials in 1935 and of a variety of other systems over the succeeding years, a large and increasing fraction of photography has been done in color. In most color photography, development is basically the same as in black-and-white photography, except that the chemical reactions result in the formation of dye images and subsequent removal of the image-forming silver. Both negative and positive (reversal) color systems are employed. A special class of print materials relies on bleachable dyes instead of conventional silver halide processes. *See* PHOTOGRAPHIC MATERIALS.

Branches and Applications

The several branches of photography may be grouped according to who takes, processes, and uses the photographs. Within the branches, numerous forms or applications of the photographic art may be distinguished by the exposing conditions and the manner in which the photographic materials, methods, and processes have been designed to optimize the results. The divisions are not always distinct, and the subtleties or strengths of one application are often employed to advantage in another.

Amateur photography refers to the taking of photographs by the general public for purposes of recollection, record, or amusement. Historically, the cameras were bulky, film formats were little or no smaller than the prints, and good lighting was required to produce black-and-white photographs of still subjects. Advancements in film technology and sophistication of amateur cameras have enabled the amateur photographer to record color photographs of moving objects on progressively smaller film formats in lightweight pocketable cameras under a wide range of lighting conditions. Processing and printing are provided by commercial photofinishers. Alternatively, the amateur photographer can obtain color prints within minutes by using so-called instant cameras and films which embody their own processing means. Advanced amateurs generally use more costly cameras and a broader range of film materials to produce photographs of esthetic merit for display. They may do their own processing and printing or rely on the assistance of professional finishers.

Photography is practiced on a professional level for portraiture and for various commercial and industrial applications, including the preparation of photographs for advertising, illustration, display, and record-keeping. Press photography is for newspaper and magazine illustrations of topical events and objects. Photography is used at several levels in the graphic arts to convert original photographs or other illustrations into printing plates for high-quality reproduction in quantity. Industrial photography includes the generation and reproduction of engineering drawings, high-speed photography, schlieren photography, metallography, and many other forms of technical photography which can aid in the development, design, and manufacture of various products. Aerial photography is used for military reconnaissance and mapping, civilian mapping, urban and highway planning, and surveys of material resources. Microfilming, which is used to store documents and radiographs in significantly reduced space, represents one of several applications of a more general category, microphotography. Biomedical photography is used to reveal or record biological structures, often of significance in medical research, diagnosis, or treatment. Photography is widely applied to preparing projection slides and other displays for teaching through visual education. Nonamateur uses of photography far exceed the amateur snapshot and home movie fields in volume or commercial significance. *See* PRINTING; SCHLIEREN PHOTOGRAPHY.

Photography is one of the most important tools in scientific and technical fields. It extends the range of vision, allowing records to be made of things or events which are difficult or impossible to see because they are too faint, too brief, too small, or too distant, or associated with radiation to which the eye is insensitive. Technical photographs can be studied at leisure, measured, and stored for reference or security. The acquisition and interpretation of images in scientific and technical photography usually requires direct participation by the scientist or skilled technicians.

Infrared photography. Emulsions made with special sensitizing dyes can respond to radiation at wavelengths up to 1200 nanometers, though the most common infrared films exhibit little sensitivity beyond 900 nm. One specialized color film incorporates a layer sensitive in the 700–900-nm region and is developed to false colors to show infrared-reflecting subjects as bright red. *See* INFRARED RADIATION.

Photographs can thus be made of subjects which radiate in the near-infrared, such as stars, certain lasers and light-emitting diodes, and hot objects with surface temperatures greater than 500°F (260°C). Infrared films are more commonly used to photograph subjects which selectively transmit or reflect near-infrared radiation, especially in a manner different from visible radiation. Infrared photographs taken from long distances or high altitudes usually show improved clarity of detail because atmospheric scatter (haze) is diminished with increasing wavelength and because the contrast of ground objects may be higher as a result of their different reflectances in the near-infrared. Grass and foliage appear white because chlorophyll is transparent in the near-infrared, while water is rendered black because it is an efficient absorber of infrared radiation.

Infrared photography is used in a wide range of scientific and technical disciplines because it permits or greatly enhances the detection of features not obvious to the eye. For example, it has been used for camouflage detection because most green paints absorb infrared more strongly than the foliage they are meant to match, to distinguish between healthy and diseased trees or plants, to detect alterations in documents and paintings, to reveal subcutaneous

veins (because the skin is somewhat transparent to infrared radiation), and to record the presence of certain substances which fluoresce in the near-infrared when exposed to visible radiation. While infrared photography can be used to map variations in the surface temperature of hot objects which are not quite incandescent, it cannot be used for detecting heat loss from buildings or other structures whose surface temperatures are usually close to that of the surrounding atmosphere.

One technique for detecting radiation at wavelengths near 10 μm relies on localized transient changes in the sensitivity of conventional photographic materials when exposed briefly to intense radiation from certain infrared lasers, and simultaneously or immediately thereafter to a general low-level flashing with light. Another technique for detecting intense infrared exposures of very short duration with films having no conventional sensitivity at the exposing wavelength is thought to rely on two-photon effects. *See* INFRARED IMAGING DEVICES.

Ultraviolet photography. Two distinct classes of photography rely on ultraviolet radiation. In the first, the recording material is exposed directly with ultraviolet radiation emitted, reflected, or transmitted by the subject; in the other, exposure is made solely with visible radiation resulting from the fluorescence of certain materials when irradiated in the ultraviolet. In the direct case, the wavelength region is usually restricted by the camera lens and filtration to 350–400 nm, which is readily detected with conventional black-and-white films. Ultraviolet photography is accomplished at shorter wavelengths in spectrographs and cameras fitted with ultraviolet-transmitting or reflecting optics, usually with specialized films. In ultraviolet-fluorescence photography, ultraviolet radiation is blocked from the film by filtration over the camera lens and the fluorescing subject is recorded readily with conventional color or panchromatic films. Both forms of ultraviolet photography are used in close-up photography and photomicrography by mineralogists, museums, art galleries, and forensic photographers. *See* ULTRAVIOLET RADIATION.

High-speed photography. Photography at exposure durations shorter than those possible with conventional shutters or at frequencies (frame rates) greater than those achievable with motion picture cameras with intermittent film movements is useful in a wide range of technical applications.

Single-exposure photography. The best conventional between-the-lens shutters rarely yield exposures shorter than 1/500 s. Some focal plane shutters are rated at 1/2000 or 1/4000 s but may take 1/100 s to traverse the film format. Substantially shorter exposures are possible with magnetooptical shutters (using the Faraday effect), with electrooptical shutters (using the Kerr effect), or with pulsed electron image tubes. Alternatively, a capping shutter may be used in combination with various pulsed light sources which provide intense illumination for very short durations, including pulsed xenon arcs (electronic flash), electric arcs, exploding wires, pulsed lasers, and argon flash bombs. Flash durations ranging from 1 millisecond to less than 1 nanosecond are possible. Similarly, high-speed radiographs have been made by discharging a short-duration high-potential electrical pulse through the x-ray tube. *See* FARADAY EFFECT; KERR EFFECT; LASER; STROBOSCOPIC PHOTOGRAPHY.

High-speed serial photographs. The classical foundation for serial frame separation is the motion picture camera. Intermittent movement of the film in such cameras is usually limited to 128 frames/s (standard rates are 16 and 24). For higher rates (up to 10,000 frames/s or more) continuous film movement is combined with optical compensation, as with a rotating plane-parallel glass block, to avoid image smear. Pictures made at these frequencies but projected at normal rates slow down (stretch) the motion according to the ratio of taking and projection rates. Higher rates, up to 10^7 frames/s, have been achieved with a variety of ingenious special-purpose cameras. In some, the sequence of photographs is obtained with a rapidly rotating mirror at the center of an arcuate array of lenses, and a stationary strip of film. In others, the optics are stationary and the film strip is moved at high speed by mounting it around the outside or inside of a rapidly rotating cylinder. To overcome mechanical limitations on the rotation of mirrors or cylindrical film holders at high speeds, image dissection methods have been employed, that is, an image is split into slender sections and rearranged to fill a narrow slit at the film. The image is unscrambled by printing back through the dissecting optics. *See* CINEMATOGRAPHY.

Photomicrography. Recording the greatly enlarged images formed by optical microscopes is indispensable in the life sciences, in metallurgy (where it is known as metallography), and in other fields requiring examination of very small specimens.

Photomicrographic apparatus consists of an optical system in which are aligned an illuminating subsystem, a compound microscope, a simple achromat, and a film holder or transport, all mounted to be free of vibration. Photomicrographs can be made by attaching a simple camera to a complete compound microscope, but a simple achromat, properly designed and located, is superior in this application to a high-quality camera lens of similar focal length. Resolving power is fully as important as magnification, but dependent mainly on the quality of the microscope objective and illumination system rather than the film. Resolving power increases as the exposing wavelength decreases. Because the luminance range of microscopic specimens is generally very low, discrimination of specimen detail is often enhanced spectrally by staining the specimen and inserting suitable filters in the illumination system. Other means for revealing specimen detail rely on controlled illumination, polarization effects, and selecting and processing films to provide higher contrast than is appropriate for general pictorial photography. The practical limitations on resolving power and magnification (approximately

2000×) in optical microscopy can be overcome by using a transmission electron microscope (an electronic analog of the optical microscope), a scanning electron microscope, or an x-ray microscope. *See* ELECTRON MICROSCOPE; LENS (OPTICS); METALLOGRAPHY; OPTICAL MICROSCOPE; RESOLVING POWER (OPTICS); SCANNING ELECTRON MICROSCOPE; X-RAY MICROSCOPE.

Close-up photography. This is the photography of specimens, usually three-dimensional objects, which are little smaller or larger than the film format. By convention, the photography of small objects at magnifications ranging from 1 to 50× is called photomacrography, while the photography of gross specimens at magnifications ranging from 1 to 0.1× is termed close-up photography. Both may be accomplished with conventional camera lenses or with so-called macro lenses, which are specially designed for use at finite object and image distances. Lens apertures must be selected to achieve maximum image sharpness consistent with the depth or thickness of the specimen.

Document copying. Photography is used for reproducing documents of all kinds because (1) it does not introduce errors in the copying process; (2) it provides an inexpensive means for retrieving and disseminating information rapidly; (3) it offers security against loss or disaster because copies can be stored in multiple locations; and (4) it effectively extends the life of perishable records.

Documents are reproduced to essentially the same size as the original or, in some copiers, with modest changes in size. They may also be photographed at reduced size (microfilming) on 16-, 35-, or 105-mm film and printed, as needed, at that size for distribution in microform or with photographic enlargement to provide reproductions at any desired scale. Historically, a variety of silver and nonsilver processes have been used for producing paper copies of business documents at or close to original size. Except for the use of conventional photographic materials in reproducing engineering drawings and other large originals, most copiers now rely on the electrostatic methods (xerography). *See* PHOTOCOPYING PROCESSES.

Microphotography. This is the process of making photographs on a greatly reduced scale, usually from two-dimensional originals. Microfilming is the special technique of copying documents on film, usually in rolls 15, 35, or 105 mm wide. Reduction factors of 10 to 45× are common, though specialized systems employing reductions ranging from 60 to more than 200× have been used. Reductions greater than 60× are customarily made in two steps to overcome physical and optical limitations. In computer output microfilming, the original "document" is displayed on a cathode-ray tube or written on the film with a laser beam. Silver-based microfilms are usually processed to meet archival standards. Microfilm negatives (or prints made on silver, diazo, or vesicular films) may be retained in roll form, cut into strips, assembled into plastic sheaths holding several strips, mounted as individual images in apertures in business machine cards, or printed as an array of images on a sheet of film (microfiche) or, at further reduction, on a film chip (ultrafiche). When retrieved from storage, these microforms are viewed on enlarging readers or readerprinters, the latter incorporating means for making an enlarged paper print quickly from the image in the viewing station. Microfilm readers often include a system for locating selected frames rapidly and automatically by reading indexing information placed alongside the microimage during the recording process.

Microphotography is used also for producing scales, reticles, and, most notably, the sets of masks employed in manufacturing a wide range of microelectronic devices. Reductions from the original artwork by factors of 100 to 1000 are made in two or three steps, depending upon the size of the originals, using materials with progressively higher resolving powers. Historically, artwork was produced manually on large translucent sheets, but high-speed, computer-driven plotters, many using laser beams to expose special films or plates, yield highly complex patterns with high quality at smaller scales. This eliminates one or two photoreduction steps and results in masks of correspondingly higher quality. The plates used in the final step are characterized not only by their high resolving power (greater than 2000 lines/mm) but by their freedom from defects or inclusions in the emulsion layer. As the demands of the microelectronics industry have pushed classical microphotography to fundamental optical limits, it has been necessary to generate patterns at final size by using electron-beam plotters, in some cases directly on resist-coated silicon wafers. *See* COMPUTER GRAPHICS; INTEGRATED CIRCUITS.

Photographic radiometry. The intensity (brightness) and spectral distribution (color) of radiation emitted or reflected by a surface can be measured by photography by comparing the densities resulting from recording a test object with those obtained from photographing a standard source or surface, preferably on adjacent portions of the same film or plate. The method is capable of reasonable precision if the characteristics of the photographic material are accurately known, all of the other photographic conditions are closely or exactly matched, and the standard source is well calibrated. In practice, it is difficult to meet all of these conditions, so photographic radiometry is used mainly when direct radiometric or photometric methods are not practicable. It is even more difficult, though not impossible, when using color films. When photographic radiometry is applied to objects at elevated temperatures which self-radiate in the visible or near-infrared region and the emissivity of the materials is known, the surface temperature of the object may be determined and the process is known as photographic pyrometry. Reference objects at known temperatures must be included in the photograph. *See* PHOTOMETRY; PYROMETER; RADIOMETRY.

Remote sensing. The art of aerial photography, in which photographs of the Earth's surface are made with specialized roll-film cameras carried aloft on balloons, airplanes, and spacecraft, is an important

segment of a broader generic technology, remote sensing. The film is often replaced with an electronic sensor, the sensor system may be mounted on an aircraft or spacecraft, and the subject may be the surface of a distant planet instead of Earth. Remote sensing is used to gather military intelligence; to provide most of the information for plotting maps; for evaluating natural resources (minerals, petroleum, soils, crops, water) and natural disasters; and for planning cities, highways, dams, pipelines, and airfields. Aerial photography normally provides higher ground resolution and geometric accuracy than the imagery obtained with electronic sensors, especially when covering small areas, so it continues as the foundation for mapmaking, urban planning, and some other applications. Films designed for aerial photography, both black-and-white and color, have somewhat higher contrast than conventional products because the luminance range of the Earth's surface as seen from altitudes of 5000 ft (1500 m) or more is roughly 100 times lower than that of landscapes photographed horizontally. *See* AERIAL PHOTOGRAPH; PHOTOGRAMMETRY; TOPOGRAPHIC SURVEYING AND MAPPING.

The acquisition of image information with scanning sensors mounted on spacecraft provides an inexpensive means for gathering photographs of large areas of the Earth or the whole Earth at regular intervals (minutes or hours for meteorological satellites, days for Earth resources satellites) or for photographing subjects which cannot be reached with aircraft or approached with spacecraft. Some sensors operate at wavelengths beyond those detected by infrared films. The image information is transmitted to receiving stations on Earth, usually processed electronically to correct for geometric and atmospheric factors, and recorded on a variety of image recorders. Scanning sensors, as well as film cameras, are employed in aerial reconnaissance because they can transmit tactical information to ground stations for evaluation before the aircraft returns to base or is shot down. Synthetic aperture radar, which maps the reflectance of microwaves from the surface of the Earth and other planets, represents another form of remote sensing for both military and commercial purposes in which the information is returned to Earth and reconstructed in photographic form for study. *See* AERIAL PHOTOGRAPH; METEOROLOGICAL SATELLITES; MILITARY SATELLITES; REMOTE SENSING.

Stereoscopic photography. This technique simulates stereoscopic vision. It presents to the two eyes individually two aspects of the subject made from slightly different viewpoints. Relief can be distinguished visually only over moderate distances, although the appearance of relief can be introduced in more distant objects by making photographs at greater distances apart. Stereoscopic photography is possible with a single camera if two separate photographs are made, one after the other, from viewpoints separated by the interocular or other appropriate distance, by displacing the lens of a single camera for the two exposures, or by rotating the object or the camera so as to give a pair of exposures on one film. These methods can be used only with relatively stationary subjects. Most stereoscopic photography is done either by the simultaneous method, in which two photographs are made at the same time with two separate cameras; with a stereoscopic camera, essentially two cameras in one body with matched optical systems and coupled focusing movements; or with single cameras using beam splitters to give two photographs side by side on the film. *See* STEREOSCOPY.

In aerial photography, the stereoscopic effect is achieved by making successive overlapping photographs along the line of flight; in terrestrial photogrammetry, photographs are made from each end of a selected base line; in stereoscopic radiography, two x-ray photographs are made in rapid succession from separate viewpoints.

Stereoscopic photographs are viewed in equipment which presents the right-eye image to the right eye only and the left-eye image to the left eye only. This can be done by separate boxes, each provided with a lens; by open-type viewers with a pair of shielded lenses, sometimes prismatic; by cabinet viewers, with pairs of optics and devices for changing the stereo slides; by grids or lenticular elements permitting each eye to see only its appropriate field; by anaglyphs, in which two images are printed in ink in complementary colors and viewed through spectacles of similar colors such that each filter extinguishes one image; and by the vectograph print, which consists of a reflecting support on which the stereoscopic pairs are placed one over the other in plastic polarizing layers with polarization planes at right angles. Polarizing spectacles are used, the eyepieces arranged so that each eye sees only the appropriate image.

Color stereo prints are made commercially by exposing color film through a fine parallel-slotted grid while swinging the camera in an arc about a principal point in the scene; processing and printing the picture; and then coating it with a plastic and embossing this with parallel lenses, registering closely with the grid pattern of the original photograph.

Nuclear-particle recording. Charged atomic particles give records on photographic emulsions, and the method of recording provides an important adjunct to such detectors as the ionization chamber and Geiger counter. The first studies were done with alpha particles, which produce a track of silver grains in the developed emulsion. Protons were also found to produce tracks, with grains of different spacing. Cosmic rays cause nuclear disintegrations on collision with the atoms in the emulsion; starlike patterns result, consisting of tracks made by particles from the atomic nuclei. Tracks of charged particles can be recorded in special thick emulsions having high silver bromide content, small grains, minimum fog, and appropriate speed or in blocks of emulsions made up of many sheets or pellicles. The grains are made developable by ionization within them caused by impact of the particles. Protons and neutrons may react with atoms of the emulsion, giving rise to charged particles, and so are indirectly recorded.

The cloud chamber and the bubble chamber provide means for photographing particle tracks through the use of high-intensity flash. In the cloud chamber the tracks are formed by condensation of moisture from supersaturated vapor on ion trails produced by the particles. In the bubble chamber the medium is a superheated liquid, such as liquid hydrogen, which gives trails of gas bubbles. The chambers are photographed with three cameras whose axes are aligned orthogonally or with a holographic camera to permit accurate determination of track lengths and trajectories. *See* COSMIC RAYS; PARTICLE DETECTOR.

Theory of the Photographic Process

The normal photographic image consists of a large number of small grains or clumps of silver. They are the end product of exposure and development of the original silver halide crystals in the emulsion and may be two to ten times larger than the crystals, which range in size up to a few micrometers (**Fig. 2**). The shape, size, and size distribution of the crystals are determined by the way the emulsions are made, and are closely related to their photographic properties. In general, high-speed, negative-type emulsions have crystals in a wide range of sizes, while emulsions with low speed and high contrast have smaller crystals that are more uniform in size.

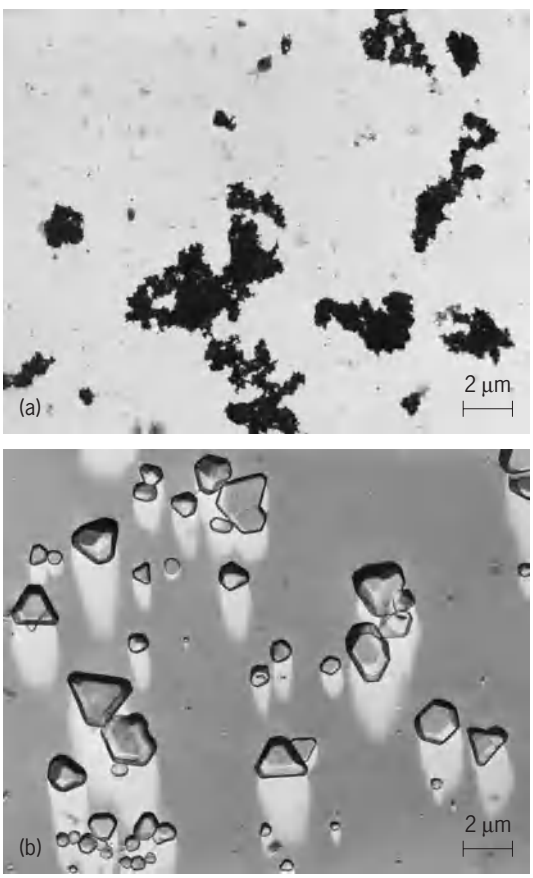

Fig. 2. Photographic silver. (*a*) Developed grains of a photographic emulsion. (*b*) Original silver halide crystals from which the silver grains were developed. Both views are highly magnified. (*Eastman Kodak Co.*)

Exposure of the crystals to light causes a normally invisible change called the latent image. The latent image is made visible by development in a chemical reducing solution which converts the exposed crystals to metallic silver, leaving unreduced virtually all of the crystals either unexposed or unaffected by exposure. The darkening which results is determined by the amount of exposure, the type of emulsion, and the extent of development. Density, a physical measure of the darkening, depends on the amount of silver developed in a particular area and the covering power of the silver. With very high exposures, darkening may occur directly, without development, as a result of photolysis. This is known as the print-out effect. Its use is confined mostly to making proof prints in black-and-white portraiture and papers for direct-trace recording instruments.

Latent image. As noted above, the term latent image refers to a change occurring in the individual crystals of photographic emulsions whereby they can be developed upon exposure to light. Early research indicated that development starts at localized aggregates of silver, mainly on the surfaces of the crystals. The existence of these discrete development centers suggested that the latent image is concentrated at specific sites. Subsequent research by numerous investigators relating to the chemical and physical properties of silver halide crystals during irradiation led eventually to a logical and comprehensive explanation for the process of latent image formation, formulated in 1938 by R. W. Gurney and N. F. Mott.

According to the Gurney-Mott theory, absorption of photon energy by a crystal liberates electrons which are able to move freely to sensitivity sites, presumably associated with localized imperfections in the crystal structure and concentrations of sensitizing impurities, where they are trapped and establish a negative charge. This charge attracts nearby positively charged silver ions, which are also mobile. The electrons and ions combine to form metallic silver, and the process continues until sufficient silver has accumulated to form a stable, developable aggregate. The evidence indicates that a minimum of three or four silver atoms is required.

At the exposure levels required to form a latent image, the absorption of one photon by a silver halide crystal can, in principle, form one atom of silver. However, many photons which reach a crystal are not absorbed, and many that are absorbed fail to contribute to the formation of a development center. The overall efficiency of latent image formation is dependent upon the structure of the crystal, the presence of sensitizing dyes and other addenda in the emulsion, temperature, and several other factors, but is essentially independent of grain size.

The efficiency of latent image formation is dependent on the rate of exposure. With low-intensity, long-duration exposure, the first silver atom formed may dissociate before succeeding ions are attracted to the site. Hence, exposing energy is wasted and the emulsion will exhibit low-intensity reciprocity failure. With very intense exposures of short

duration, an abundance of sites compete for the available silver atoms with the result that many of the sites fail to grow sufficiently to become developable. In addition, such exposures frequently result in the formation of internal latent images which cannot be reached by conventional surface-active developers.

The Gurney-Mott theory was not universally accepted at first, and there is still uncertainty with regard to some details of the mechanism, but it prompted extensive research and continues to be the best available explanation of latent image formation and behavior. For additional information on the Gurney-Mott theory *see* PHOTOCONDUCTIVITY; PHOTOLYSIS.

Development. Development is of two kinds, physical and chemical. Both physical and chemical developers contain chemical reducing agents, but a physical developer also contains silver compounds in solution (directly added or derived from the silver halide by a solvent in the developer) and works by depositing silver on the latent image. Physical development as such is little used, although it usually plays some part in chemical development. A chemical developer contains no silver and is basically a source of reducing agents which distinguish between exposed and unexposed silver halide and convert the exposed halide to silver. Developers in general use are compounded from organic reducing compounds, an alkali to give desired activity, sodium sulfite which acts as a preservative, and potassium bromide or other compounds used as antifoggants (fog is the term used to indicate the development of unexposed crystals; it is usually desirable to suppress fog).

Most developing agents used in normal practice are phenols or amines, and a classical rule which still applies, although not exclusively, states that developers must contain at least two hydroxyl groups or two amino groups, or one hydroxyl and one amino group attached ortho or para to each other on a benzene nucleus. Some developing agents do not follow this rule. *See* AMINE; PHENOL.

Alkalies generally used are sodium carbonate, sodium hydroxide, and sodium metaborate. Sulfite in a developer lowers the tendency for oxidation by the air. Oxidation products of developers have an undesirable influence on the course of development and may result in stain.

Developing agents and formulas are selected for use with specific emulsions and purposes. Modern color photography relies mainly on paraphenylenediamine derivatives. So-called fine-grain developers are made to reduce the apparent graininess of negatives. They generally contain the conventional components but are adjusted to low activity and contain a solvent for silver bromide. One fine-grain developer is based on *p*-phenylenediamine, which itself is a silver halide solvent. Some developers are compounded for hardening the gelatin where development occurs, the unhardened areas being washed out to give relief images for photomechanical reproduction and imbibition color printing. In a number of products, the developing agent is incorporated in the emulsion, development being initiated by placing the film or paper in an alkaline solution called an activator. Hardening agents may be included in developers to permit processing at high temperatures.

Monobaths are developers containing agents which dissolve or form insoluble complexes with the undeveloped silver halide during developing.

When film is developed, there is usually a period during which no visible effect appears; after this the density increases rapidly at first and then more slowly, eventually reaching a maximum. In the simplest case, the relation between density and development time is given by Eq. (1), where D is the density

$$D = D_x(1 - e^{-kt}) \tag{1}$$

attained in time t, D_x is the maximum developable density, and k is a constant called the velocity constant of development.

After processes. These include fixing, washing, drying, reduction, intensification, toning, and bleaching.

Fixing. After the image is developed, the unchanged halide is removed, usually in water solutions of sodium or ammonium thiosulfate (known as hypo and ammonium hypo, respectively). This procedure is called fixing. Prior to fixing, a dilute acid stop bath is often used to neutralize the alkali carried over from the developer. The rate of fixing depends on the concentration of the fixing agent (typically 20–40% for hypo) and its temperature (60–75°F or 16–24°C). Hardeners may be added to the fixer to toughen the gelatin, though they tend to reduce the fixing rate. In so-called stabilization processes, the undeveloped silver salts are converted into more or less stable complexes and retained in the emulsion to minimize the processing time associated with washing. Solutions of organic compounds containing sulfur are often used for this purpose.

Washing and drying. Negatives and prints are washed in water after fixing to remove the soluble silver halide-fixing agent complexes, which would render the images unstable or cause stain. The rate of removal declines exponentially and can be increased by raising temperature and increasing agitation. The rate can be accelerated with neutral salt solutions known as hypo clearing aids, thus reducing washing times. Print permanence can be enhanced with toning solutions (especially selenium toners) too dilute to cause a change in image coloration. After washing, the materials must be dried as uniformly as possible, preferably with dust-free moving warm air, though some papers are often dried on heated metal drums. Drying uniformity is aided by adding wetting agents to the final wash water.

Reduction and intensification. Reduction refers to methods of decreasing the density of images by chemically dissolving part of the silver with oxidizers. According to their composition, oxidizers may remove equal amounts of silver from all densities, remove silver in proportion to the amount of silver present, or remove substantially more silver from the higher

densities than from the lower. Intensification refers to methods of increasing the density of an image, usually by deposition of silver, mercury, or other compound, the composition being selected according to the nature of the intensification required.

Toning. The materials and processes for black-and-white photography are normally designed to be neutral in color. The coloration can be modified by special treatments known as toning. This may be done with special developers or additions to conventional developers, but toning solutions used after fixing and washing are more common. These solutions convert the silver image into compounds such as silver sulfide or silver selenide, or precipitate colored metallic salts with the silver image. Dye images can be obtained by various methods.

Bleaching. In reduction and in color processes and other reversal processes where direct positives are formed, the negative silver is dissolved out by an oxidizing solution, such as acid permanganate or dichromate, which converts it to a soluble silver salt. If a soluble halide, ferricyanide, or similar agent is present, the corresponding silver salt is formed and can be fixed out or used as a basis for toning and intensifying processes. Bleach-fixing solutions (blixes) are used to dissolve out silver and fix out silver halide simultaneously.

Sensitometry and Image Structure

Sensitometry refers to the measurement of the sensitivity or response to light of photographic materials. Conventional methods of measurement tend to mask the dependence of this response on factors which affect the structure of the photographic image. The simplest method of determining sensitivity is to give a graded series of exposures and to find the exposure required to produce the lowest visible density. Modern sensitometry relies on plotting curves showing the relation between the logarithm of the exposure H and the density of the silver image for specified conditions of exposure, development, and evaluation. Exposure is usually defined as the product of intensity and duration of illumination incident on the material, though it is more correct to define exposure as the time integral of illuminance I ($H = \int I dt$). Density D is defined as the logarithm of the opacity O, which in turn is defined as the reciprocal of transmittance T. If light of intensity I falls on a negative and intensity I' is transmitted, Eqs. (2) hold.

$$T = \frac{I'}{I} \qquad O = \frac{1}{T} = \frac{I}{I'} \qquad D = \log \frac{I}{I'} \qquad (2)$$

Illuminance is the time rate at which light is directed at a unit area of film, exposure is the amount of energy per unit area received by the film, and density is a measure of the extent to which that energy is first converted to a latent image and then, with an infusion of chemical energy from the developer, is amplified to form a stable or permanent silver image. Since density is dependent mainly on the amount of silver per unit area of image, and since the amount of energy required to form a latent image in each crystal is independent of crystal size, it takes more exposure to obtain a specified density with a fine-grain emulsion than a coarse-grain one. However, through advances in emulsion and developer chemistry, modern picture-taking films are roughly 25,000 times more sensitive than the gelatin dry plates introduced in 1880.

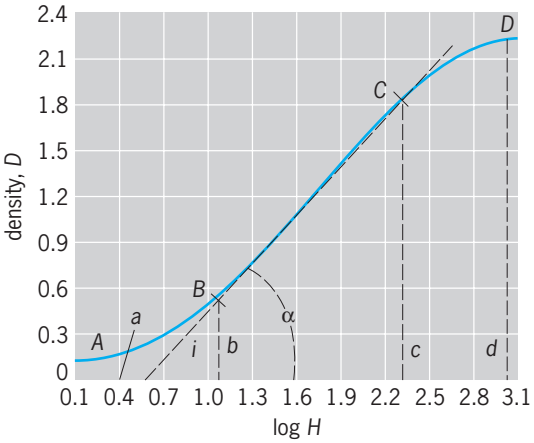

Fig. 3. Hurter and Driffield characteristic curve.

Hurter and Driffield curve. The relationship between density and the logarithm of exposure is shown by the Hurter and Driffield characteristic curve, which has three recognizable sections, the largest corresponding usually to the straight line B to C in **Fig. 3**. In this region, density is directly proportional to the logarithm of exposure. In the lower and upper sections of the curve, this proportionality does not apply. These nonlinear regions, AB and CD, are commonly called the toe and shoulder, respectively.

The characteristic curve is used to determine the contrast, exposure latitude, and tone reproduction of a material as well as its sensitivity or speed. It is obtained by using a light source of known intensity and spectral characteristics, a modulator which gives a series of graded exposures of known values, precise conditions for development, and an accurate means of measuring the resulting densities.

Sensitometric exposure. The internationally adopted light source is a tungsten-filament electric lamp operated at a color temperature close to 2850 K combined with a filter to give spectral quality corresponding to standard sensitometric daylight (approximately 5500 K). The light source and exposure modulator are combined to form a sensitometer, which gives a series of exposures increasing stepwise or in a continuous manner. Sensitometers are either intensity-scale or time-scale instruments, depending on whether the steps of exposure are of fixed duration and varying intensity or are made at constant intensity with varying duration. Special sensitometers are used to determine the response of materials to the short exposures encountered in electronic flash photography or in scanning recorders. The exposure may also be intermittent, that is, a series of short exposures adding up to the desired total, though continuous exposures are preferred to avoid intermittency effects.

(Below a critical high frequency, the photographic material does not add up separate short exposures arithmetically.) The best sensitometers are of the continuous-exposure, intensity-scale variety, and the exposure time used should approximate that which would prevail in practice. *See* EXPOSURE METER; HEAT RADIATION.

Development. Development is carried out in a standard developer or in a developer specifically recommended for use with the material under test, for the desired time at a prescribed standard temperature, and with agitation that will ensure uniform development and reproducibility. The procedure may be repeated for several development times.

Densitometers. Photographic density is measured with a densitometer. Early densitometers were based on visual photometers in which two beams are taken from a standard light source and brought together in a photometer head. One beam passes through the density to be measured, the other through a device for modulating the intensity so that the two fields can be matched. Visual instruments are seldom used now except as primary reference standards. Most densitometers are photoelectric devices which rely on one or more standard densities which have been calibrated on another instrument.

The simplest photoelectric transmission densitometers are of the deflection type, consisting of a light source, a photosensitive cell, and a microammeter or digital readout, the density being obtained from the relative meter deflections or from readings with and without the sample in place. Other photographic densitometers employ a null system in which the current resulting from the transmission of the test sample is balanced by an equivalent current derived from light transmitted through a calibrated modulator. In another null method, the sample and modulator are placed in one beam and the modulator controlled so that the net transmittance of the combination is constant. Some densitometers are equipped with devices for plotting the characteristic curves automatically. The densitometry of colored images is done through a suitable set of color filters and presents special problems in interpretation.

When light passes through a silver image, some is specularly transmitted and some is scattered. If all the transmitted light is used in densitometry, the result is called the diffuse density. If only the light passing directly through is measured, the density is called specular. The specular density is higher and related to diffuse density by a ratio known as the Callier Q factor. The Q of color films is negligible.

The density of paper prints or other images on reflective supports is measured by reflection. The reflection density is $D_R = \log(1/R)$, where R is the ratio of light reflected by the image to that reflected by an unexposed area of the paper itself or, better, by a reference diffuse light reflector. As with transmission densitometry, the angular relationships between incident and reflected light must be standardized.

Contrast. If extended, the straight-line portion of the characteristic curve BC intersects the exposure axis at the inertia point i on Fig. 3, and the angle α is a measure of contrast. The tangent of α, that is, the slope of the line BC, is known as gamma (γ). As development time or temperature increase, gamma increases, eventually reaching a maximum (**Figs. 4** and **5**). The projection of the straight line BC of Fig. 3 onto the exposure axis at bc is a measure of latitude, and the distance ad represents the total scale, that is, the whole exposure range in which brightness differences in the subject can be recorded.

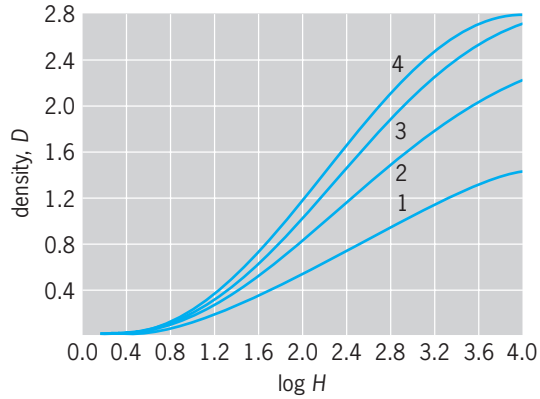

Fig. 4. Characteristic curves for development times increasing in the order 1, 2, 3, 4; horizontal scale is in lux seconds.

Another measure of contrast, contrast index, is more useful than gamma in pictorial photography because it is the average gradient over the part of the characteristic curve used in practice. It is determined by measuring the slope of the line joining a point in the toe of the curve and one at a higher density that are separated by a distance related to a specified difference in exposure or density.

Sensitivity. The sensitivity or speed of a photographic material is not an intrinsic or uniquely defined property. However, it is possible to define a speed value which is a numerical expression of the material's sensitivity and is useful for controlling camera settings when the average scene brightness is known. Speed values are usually determined from an equation of the form $S = k/H$, where S is the speed, k is an arbitrary constant, and H is the exposure corresponding to a defined point on the characteristic curve.

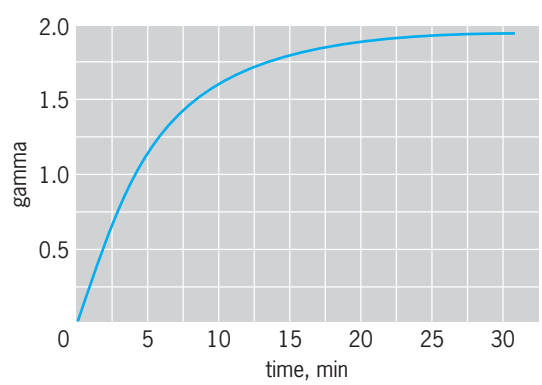

Fig. 5. Gamma versus time of development curve.

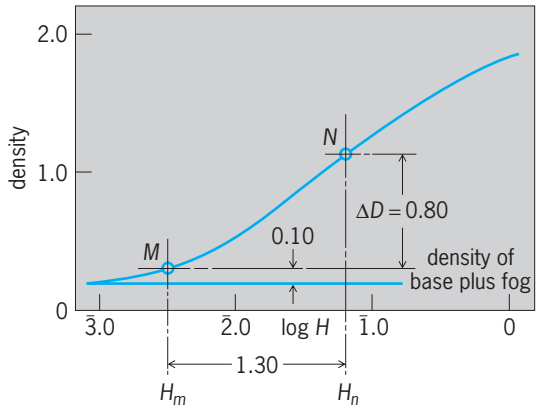

Fig. 6. Determination of the speed of a photographic film according to the 1979 American Standard Method, now adopted internationally. ISO/speed is $0.8/H_M$ when film is processed to achieve gradient defined by points *M* and *N*.

Historically, Hurter and Driffield speeds were obtained by dividing 34 by the exposure corresponding to the inertia point, but most modern definitions use much smaller constants and the exposure corresponds to a point in the toe region which is at some arbitrary density above base plus fog or where the gradient of the curve has reached some specified value. The goal is always to establish speed values which lead either to correctly exposed transparencies or to negatives that can be printed readily to form pictures of high quality. The constant k usually incorporates a safety factor to assure good exposures even with imprecise exposure calculations.

The present international standard for determining the speeds of black-and-white pictorial films is shown in **Fig. 6**. This standard has been adopted by the American National Standards Institute (ANSI) and the International Organization for Standardization (ISO). Film speeds determined by this method are called ISO speeds and correspond closely to values long known as ASA speeds. Point M is at a density 0.10 above base-plus-fog density. Point N lies 1.3 log H units from M in the direction of increased exposure. Developing time in a standard developer is adjusted so that N also lies at a density 0.8 above the density of M. Then, the exposure H_M (in lux seconds) corresponding to M is the value from which the ISO speed is computed according to $S = 0.8/H_M$. Exposure indexes or speeds for films and papers used in printing and for films and plates used in various technical applications are often determined in the same general manner, but the criteria for required exposure, the processing conditions, and the constant k are usually quite different, reflecting the nature of the application for which the material is intended.

Spectral sensitivity measures the response of photographic emulsions as a function of the exposing wavelength. It is usually determined with a spectrosensitometer, that is, a spectrograph with an optically graded wedge or step tablet over the slit. The contour of the spectrogram for a constant density is related to the spectral response of the material and spectral output of the light source. Ideally, spectral sensitivity curves are adjusted to indicate the response to a source of constant output at all wavelengths. These are known as equal-energy curves.

Reciprocity effects. Defining exposure as the product of illuminance and time implies a simple reciprocal relationship between these factors for any given material and development process. As noted above, the efficiency of latent image formation is, in fact, affected by very long, very short, or intermittent exposures. Variations in the speed and contrast of an emulsion with changes in the rate of exposure are therefore called reciprocity effects. In general, these effects are not dependent upon the exposing wavelength for a single emulsion, but they may vary among the individual emulsions of color materials, thereby affecting color rendition.

Tone reproduction. The relation between the distribution of luminance in a subject and the corresponding distribution of density in a final print is referred to as tone reproduction. It is affected by the sensitometric characteristics of the materials, including their nonlinear aspects, and by subjective factors as well. The exposure and processing of both the negative and print materials are manipulated by skilled photographers to achieve visually appealing results.

Image structure. The granular nature of photographic images manifests itself in a nonhomogeneous grainy appearance which may be visible directly and is always visible under magnification. Granularity, an objective measure of the sensation of graininess, is generally greater in high-speed emulsions and depends also on both the nature of the development and the mean density of the image area being measured. Producing emulsions with increased speed without a corresponding increase in granularity is a constant goal. Graininess in prints increases with increases in magnification, contrast of the print material, and (in black-and-white materials) with mean density of the negative.

The ability of a film to resolve fine detail or to record sharp demarcations in exposure is limited by the turbidity (light-scattering property), grain size, and contrast of the emulsion, and is affected by several processing effects. Resolving power is a measure of the ability to record periodic patterns (bars and spaces of equal width) and is usually expressed as the number of line pairs (bar and space pairs) per millimeter which can just be separated visually when viewed through a microscope. Resolving power values of 100 to 150 line pairs/mm are common but may range from less than 50 to more than 2000.

Sharpness refers to the ability of an emulsion layer to show a sharp line of demarcation between adjacent areas receiving different exposures. Usually such an edge is not a sharp step but graded in density to an extent governed by light scattering (turbidity) and development effects. Subjective assessments of sharpness correlate with a numerical value known as acutance, which is a function of both the rate of change in density and the overall difference in density across a boundary.

A different and generally more useful measure of the ability of a photographic material to reproduce image detail is called the modulation transfer function (MTF). If a sinusoidal distribution of illuminance is applied to the material, the resulting variation in density in the processed image will not, in general, coincide with that predicted by the characteristic curve, again because of localized scattering and development effects. (These effects are excluded in conventional sensitometry.) Modulation transfer is the ratio of the exposure modulation that in an ideal emulsion would have caused the observed density variation to the exposure modulation that actually produced it. This ratio is a function of the spatial frequency of the exposure distribution, expressed in cycles per millimeter. The MTF of a film may be combined with a similar function for a lens to predict the performance of a lens-film combination.

Photographic Apparatus

The camera is the basic exposing instrument of photography. In cameras having shutters within the lens, the entire image is recorded simultaneously; in those with focal plane shutters, exposure is governed by a slit which moves rapidly across the film so that the time required to expose the whole image is necessarily greater than the exposure duration for any line element parallel to the slit. Another important class of exposing instruments includes recorders and plotters in which individual picture elements (spots) are exposed sequentially in some orderly fashion. Other important instrumentation in photography includes means for lighting the subject, devices for handling the photographic materials during processing, equipment for printing by contact or projection, and means for viewing, storing, and retrieving photographs. *See* CAMERA.

Recorders and plotters. The information associated with an image (normally a two-dimensional array of spots of varying density and color) may be acquired in electronic form either directly, as with an electronic camera, or by computation from mathematical descriptions of an object or pattern that has not existed in fact, or indirectly by scanning an original photograph with an electronic sensor. In these cases the information may be retained in electronic or magnetic memory until recalled for viewing. It is often required that the image then be converted to conventional photographic form. Continuous-tone images are usually reconstructed by writing with a single spot which moves along the lines of a raster. Both the number of spots along each line and the number of lines in the height of the image are set by design and may range from several hundred to more than 10,000, depending upon the quality desired. In certain specialized recorders, several independently modulated spots are caused to scan and expose in parallel.

One form of raster imaging recorder consists of a cathode-ray tube or electronic panel display and a camera designed to relay the image to the film or paper, which usually is held stationary. In others, a small spot of light derived from a laser, light-emitting diode, xenon arc, or other intense source is modulated and caused to scan across the photographic material, often with the use of a rotating mirror. In these devices, one of the scanning motions may be achieved by moving the film with respect to the exposing system.

When one is recording maps or engineering drawings, it is often advantageous to use plotters which operate in vector rather than raster mode so that curved lines or straight lines not parallel with the scan lines of a raster will be continuous and time will not be wasted scanning large areas of the photographic materials that are to receive no exposure. Cathode-ray tubes and electron beam recorders (in which the electrons of a specialized cathode-ray tube impinge directly on the film instead of a phosphor screen) are especially useful for vector plotting because there is no mechanical inertia to be overcome when deflecting the electron beam.

The photographic materials used in image recorders and plotters are properly selected with regard for the short exposures which prevail at individual picture elements, the wavelength of the exposing source, the exposure range determined by the modulator, and the application or end use of the recorded image.

Darkroom equipment. Other important instrumentation includes devices for handling photographic materials during processing, printing, viewing, storage, and retrieval. Until processing is virtually completed, most films, plates, and papers must be handled in a darkroom or light-tight container. The darkroom may range from total darkness to subdued illumination from safelights, that is, lamphouses with filters to provide light of a color that will not fog the material in a reasonable time. Other equipment used for manual processing includes sinks with running water and drains; tanks and hangers for processing plates or sheet films vertically; light-tight tanks for processing roll films held in spiral reels; flat trays for processing materials in sheet form individually; thermometers, mixing valves, and other devices for controlling solution temperatures; special containers for washing negatives and prints; dryers, ranging from simple clips for hanging negatives in open air to cabinets with forced warm air and heated drums; clocks and preset timers; printers and enlargers; exposure meters and focusing devices for enlargers; and print trimmers.

Many photographic materials (especially color films and papers) require strict adherence to carefully designed processing cycles. In many technical applications, the quality of the recorded image and the information it represents are critically dependent upon uniform processing conditions. When processing films and prints in commercial quantities, as in the photofinishing, motion picture, and microfilm industries, economic considerations preclude manual darkroom operations. All of these needs are met through the use of various processing machines and automatic or semiautomatic printers. Processing machines commonly use rollers to transport the material through tanks or other stations where

developers, fixers, monobaths, activators, bleaches, or other processing chemicals are brought into contact with the emulsion and backing layers by immersion, application of a thin layer, or spraying. Solution temperatures are generally elevated to achieve reduced times. Washing and drying stations complete the processes.

Contact and projection printers. Negatives provide the primary records in most black-and-white photography and a large fraction of color photography. In many applications, the negative is more useful than a print derived from it. However, positive prints from original negatives or positive color transparencies are often required. Such prints are made in contact printers or projection printers (enlargers).

Contact printers are boxes containing lamps, diffusing screens, and a glass top on which a negative is placed. A sheet of printing paper is placed directly over the negative, usually emulsion to emulsion, and a lid with felt or other spongy material on its face is closed to assure good contact through the application of uniform pressure. Alternatively, the negative and print material may be placed in a vacuum printing frame and exposed to a distant point source to retain edge sharpness.

Projection printers, or enlargers, are optical projectors consisting of a lamphouse and light source, a condenser or diffusing sheet to assure uniform illumination, a holder for the negative, a projection lens designed to give optimum quality at relatively low magnifications, a board or easel to hold the sensitized paper, and means for holding these components when aligned and spaced to provide sharp focus at the desired magnification. Color filters are included for color printing. *See* OPTICAL PROJECTION SYSTEMS.

Color negatives are usually analyzed by sophisticated optoelectronic systems to determine the exposures needed for optimum color balance and mean density. In some systems the original is scanned and digitized, the information is processed electronically to achieve effects not possible with conventional printers, and the result is reconstructed as a new negative, from which prints can be made in quantity, or directly as a final print. In either case, the prints are intended to have enhanced esthetic or technical qualities.

Robert D. Anwyl

Color Photography

Color photography is the photographic reproduction of natural colors in terms of two or more spectral components of white light. In 1665 Isaac Newton observed that white light could be separated into a spectrum of colors, which then could be reunited to form white light. In 1861 James Clerk Maxwell prepared positive transparencies representing red-, green-, and blue-filter exposures, projected each through the respective "taking" filter, and superimposed the three images to form a full-color image on a screen.

Maxwell's demonstration was the first example of additive color photography. Red, green, and blue are called additive primary colors because they add to form white light. Their complements are the subtractive primary colors, cyan, magenta, and yellow. A cyan, or minus-red, filter absorbs red light and transmits blue and green; a magenta, or minus-green, filter absorbs green and transmits red and blue; and a yellow, or minus-blue, filter absorbs blue and transmits red and green. Subtractive color processes produce color images in terms of cyan, magenta, and yellow dyes or pigments (**Fig. 7**). Two-color subtractive processes (cyan and red-orange), which provide a less complete color gamut, have also been used. *See* COLOR; COLOR FILTER.

Additive color photography. Additive color is most widely represented by color television, which presents its images as arrays of minuscule red, green, and blue elements.

The screen plate system of additive photography uses a fine array of red, green, and blue filters

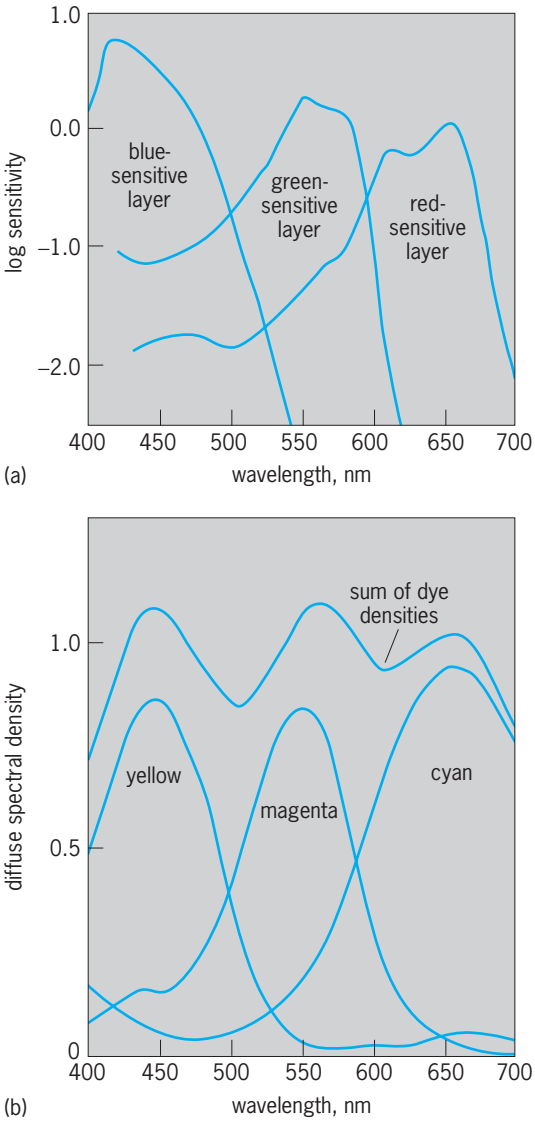

Fig. 7. Subtractive color material. (*a*) Spectral sensitivities of blue-, green-, and red-sensitive emulsions. (*b*) Corresponding spectral densities of dyes that form yellow, magenta, and cyan images. Sum of dye densities, corresponding to uniform exposure at all visible wavelengths, is not flat but yields visually neutral density.

in contact with a single panchromatic emulsion. Reversal processing produces a positive color transparency comprising an array of positive black-and-white image elements in register with the filter elements. Some early screen processes have a hyperfine color stripe screen integral with a silver transfer film.

Lenticular additive photography is based on black-and-white film embossed with lenticules. While the camera lens photographs the scene, the lenticules image onto the film a banded tricolor filter placed over the camera lens. The resulting composite image is reversal-processed and projected through the same tricolor filter to reconstruct the original scene in color.

Subtractive color photography. Many types of subtractive color negatives, color prints, and positive transparencies are made from multilayer films with separate red-, green-, and blue-sensitive emulsion layers.

In chromogenic or color development processes, dye images are formed by reactions between the oxidation product of a developing agent and color couplers contained either in the developing solution or within layers of the film. For positive color transparencies, exposed silver halide grains undergo nonchromogenic development, and the remaining silver halide is later developed with chromogenic developers to form positive dye images. Residual silver is removed by bleaching and fixing.

In the first commercially successful subtractive color transparency film (**Fig. 8**), reversal dye images are formed in successive color development steps, using soluble couplers. Following black-and-white development of exposed grains, the remaining red-sensitive silver halide is exposed to red light, then developed to form the positive cyan image. Next, the blue-sensitive layer is exposed to blue light and the yellow image developed, and finally the green-sensitive emulsion is fogged and the magenta image developed.

Simpler reversal processing is afforded by films with incorporated nondiffusing couplers that produce three dye images in a single color development step.

Color negative films with incorporated couplers are developed with a color developer to form cyan, magenta, and yellow negative images for use in printing positive color prints or transparencies. Colored couplers in color negative films provide automatic masking to compensate for the imperfect absorption characteristics of image dyes. For example, a magenta dye, ideally minus-green, absorbs not only green but also some blue light. The magenta-forming coupler is colored yellow, and in undeveloped areas it absorbs as much blue light as the magenta dye absorbs in developed areas, so that blue exposure through the negative is unaffected by the blue absorption of the magenta dye.

Chromogenic films designed for daylight exposure commonly have an ultraviolet-absorbing dye in the protective overcoat to prevent excessive exposure of the blue-sensitive emulsion. (Ultraviolet exposure of the remaining layers is blocked by the yellow filter layer.) Advances in silver halide, dye, and processing chemistry have resulted in chromogenic films with progressive improvements in speed, color rendition, image quality, and dye stability over early products of this type.

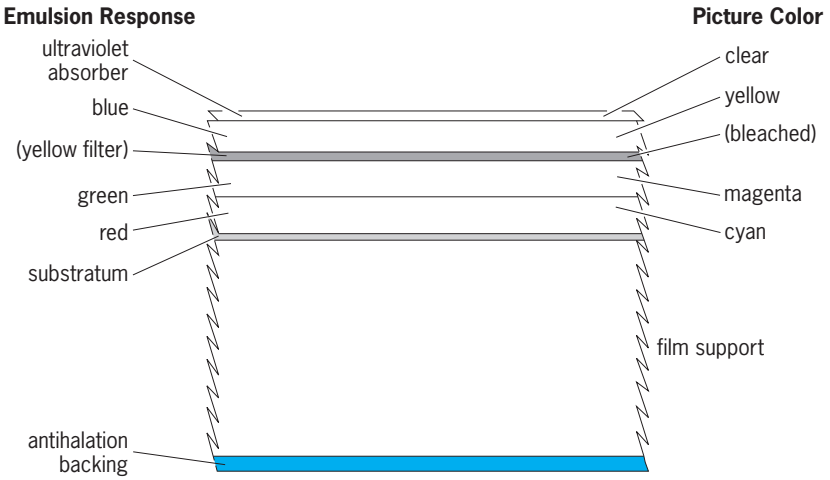

Fig. 8. Schematic cross section of subtractive color transparency film, showing configuration of three emulsion layers with (left) the spectral response of each emulsion and (right) the color of the dye image formed in each layer through chromogenic development. Protective overcoat contains ultraviolet absorbing dye to prevent overexposure of blue-sensitive emulsion with daylight or electronic flash illuminants.

One-step color photography. The introduction in 1963 of a single-structure film and a single-step process marked a significant branching in the history and technology of color photography. The availability of instant color photographs directly from the camera provided a useful alternative to conventional color films and print materials, which require a printing step and separate darkroom processing of each material, particularly in applications where enlargement and multiple prints are not primary requirements.

In the Polacolor process and later one-step, or instant, processes, color prints are produced under ambient conditions shortly after exposure of the film. Processing takes place rapidly, using a reagent—sometimes called an activator—provided as part of the film unit. Important concepts include the design of cameras incorporating processing rollers, the use of viscous reagents sealed in foil pods to provide processes that are outwardly dry, and traps to capture and neutralize excess reagent. The viscous reagent makes possible accurately controlled, uniform distribution. The closed system provides fresh reagent for each picture unit and permits use of more reactive chemicals than are practical in conventional darkroom processes.

Rather than forming dyes during processing, the first one-step process introduced preformed dye developers, compounds that are both image dyes and developers. Development of exposed silver halide crystals causes immobilization of the associated dye developers while dye developers in unexposed regions diffuse to a polymeric image-receiving layer. The positive receiving sheet includes a polymeric

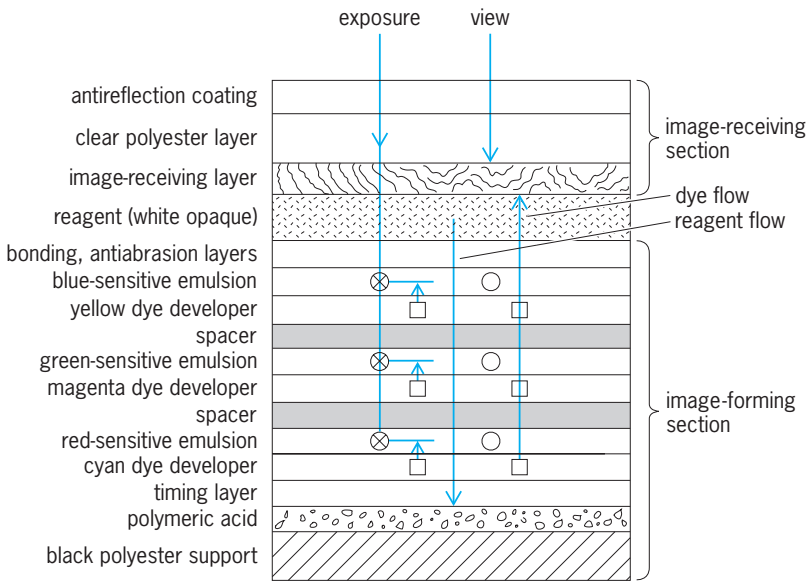

Fig. 9. Schematic cross section of integral one-step film units that provide full-color prints within 2 min after exposure. The dye image diffuses from exposed areas to the image-receiving layer and is viewed against the reflective white pigment of the reagent layer.

acid layer that neutralizes residual alkali. The full-color print is ready for viewing when the negative and positive sheets are stripped apart after 60 s. The process, like most peel-apart systems, is adaptable to large-format photography. A film based on similar principles yields color transparencies suitable for overhead projection.

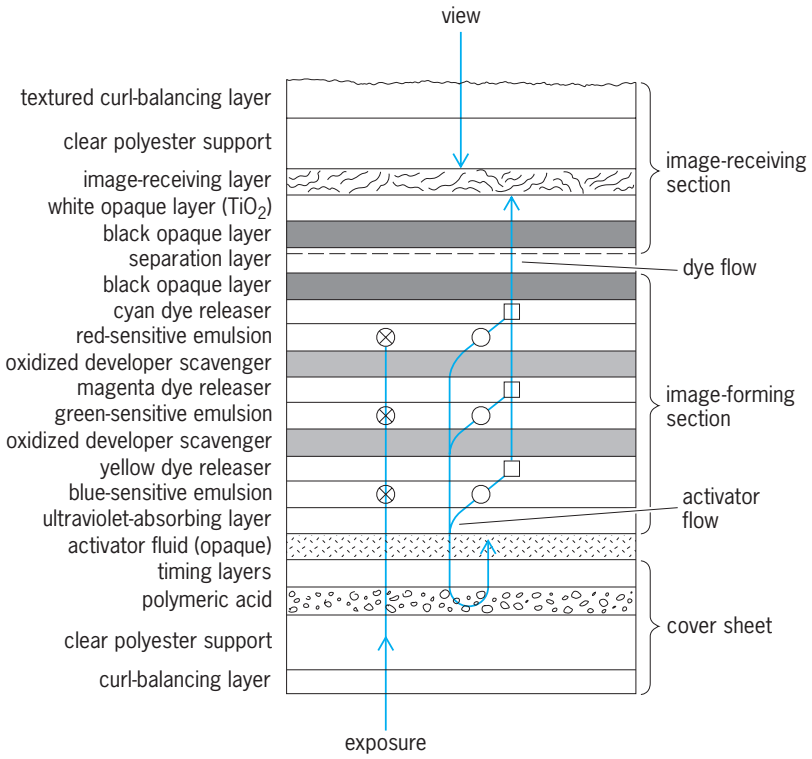

Fig. 10. Schematic cross section of integral one-step film. Dyes released from unexposed areas of image flow to image-receiving layer for viewing against white opaque layer. Image-receiving section may then be separated (at broken line) from image-forming section to obtain a thin print.

The first fully integral one-step color film was also based on dye developers (**Fig. 9**). The film unit comprises two polyester sheets, one transparent and one opaque, with all of the image-forming layers between them. The light-sensitive layers are exposed through the transparent sheet, and the final image is viewed through the same sheet against a reflective white-pigmented layer that is formed by the reagent. Cameras for the system automatically eject the film immediately, driving the film unit between rollers to spread the reagent within the film. Although still light-sensitive when ejected, the emulsion layers are protected from exposure to ambient light by the combined effects of the white pigment and special opacifying dyes included in the reagent. Timing of the process is self-controlled and there is no separation of the sheets. Images on materials which are exposed and viewed from the same side will be laterally reversed unless an odd number of image inversions is introduced in the exposing process. Hence, cameras intended for use with the film incorporate one or three mirrors. The use of three mirrors permits cameras of compact design.

Other instant color films incorporate imageforming and -receiving layers between two polyester sheets, both clear (**Fig. 10**). Direct positive emulsions are selectively exposed through the clear cover sheet and an ultraviolet-absorbing layer. Following exposure, an activator containing potassium hydroxide, a developing agent, and other components is distributed uniformly between the cover sheet and the integral imaging sections as the film unit is ejected from the camera. Carbon black in the activator blocks exposure of the emulsion layers while the activator penetrates them and develops the unexposed silver halide crystals. Oxidized developer resulting from development reacts with dye releasers in an adjacent layer and, in the presence of an alkali, liberates preformed dyes. (Scavenger layers prevent dye release from the wrong layers.) The released dyes migrate through two or more opaque layers to an image-receiving layer. Timing of the process is governed by the neutralizing action of a polymeric acid when the activator penetrates the timing layers.

Because the images formed in the film are viewed through the second clear plastic sheet, they are right-reading. Thus, cameras designed to accept these products need contain no mirrors, though two are incorporated in many such cameras to reduce overall camera size. All layers are retained as an integral unit in the film, but the film (Fig. 10) embodies two extra layers and is partially slit during manufacture so that the picture and border section of the image-receiving section may be separated from the unit after processing is complete. The image-forming section, including the activator pod and trap, is discarded, leaving a thin print. A color slide film produces transparencies instead of thin prints. Following separation from the image-forming section, the transparencies are trimmed and inserted automatically into 2 by 2 in. (5 by 5 cm) slide mounts with a slide-mounting accessory.

Color print processes. Color prints from color negatives are made primarily on chromogenic print materials by using simplified processes similar to those used for the original negatives. Color transparencies are made from such negatives using special print films which are processed like the original negatives. Chromogenic reversal materials, such as specially treated paper and duplicating films, are used to make prints or duplicate transparencies from positive color originals. Alternatively, a color negative intermediate film may be derived from the original positive transparency and used to make multiple prints with negative-working papers.

Color prints and transparencies can be made from original color negatives and transparencies in an image transfer system which uses a single, strongly alkaline activator for processing. As appropriate, a negative- or positive-working film is exposed to the original. After 20-s immersion in the activator, the film is laminated to a reflective (paper) or transparent (plastic) receiver sheet. The two parts are separated minutes later. Other color print films based on release and diffusion of preformed dyes in the presence of alkaline activators include a single-sheet material and a two-sheet graphic arts product.

Color prints and transparencies are also made from original positive transparencies by silver dye bleach processes. The processes are based on the destruction of dyes rather than their formation or release. Following development of silver images in red-, green-, and blue-sensitive emulsion layers, associated subtractive dyes are selectively destroyed by a bleach reaction catalyzed by the negative silver, leaving a positive color image in place. Residual silver salts are removed by fixing.

Dye transfer prints are made by sequential imbibition printing of cyan, magenta, and yellow images from dyed gelatin relief matrices onto gelatin-coated paper or film. Panchromatic matrix films are exposed from color negatives through appropriate color separation filters; blue-sensitive matrix films are printed from black-and-white separation negatives. The images formed by dye transfer and dye bleach exhibit good dye stability.

Color motion pictures. Many of the techniques described above have been used in motion picture processes. Since their introduction in the 1950s, chromogenic negative and positive films have become the most widely used motion picture production materials. Vivian K. Walworth; Robert D. Anwyl

Electronic Devices and Technology

The photographic art has been influenced progressively, especially since 1960, by electronic devices and technology. This influence has affected virtually every element of the photographic process, including exposure control, automatic focusing of cameras, process control, image enhancement through intervention in the printing step, and even acquisition of the original image information.

Early forms of electronic acquisition include the scanning of photographs for transmittal to another location (facsimile) or producing color separations in graphic arts scanners. The scanning of original subjects with television cameras, with point detectors (and oscillating mirrors) mounted in remote-sensing aircraft or spacecraft, or with scanning electron microscopes were forerunners to the charge-coupled arrays and other image-sensing devices that take the place of films or plates in some cameras. The image information may be stored on magnetic media or electronically and presented for initial viewing on a video monitor (cathode-ray tube), but it may also be converted to conventional photographic form by photographing the monitor or with other types of line-scan image recorders. Whenever the image information is in electronic form, it can be manipulated to alter the tonal and chromatic relationships in the resulting image. This manipulation may be introduced to create a print that appears to be more like the original scene than could be achieved without electronic intervention, or it may be used to create images clearly different from the original. The result may be visually arousing in an esthetic sense or reveal image information of scientific or technical significance. At the same time, it is possible to suppress grain noise and enhance edge sharpness. *See* CATHODE-RAY TUBE; CHARGE-COUPLED DEVICES; FACSIMILE; IMAGE PROCESSING; IMAGE TUBE (ASTRONOMY); TELEVISION CAMERA.

The rapid growth in electronic technology and its impact on virtually every other scientific or technical discipline were made possible by the development of integrated circuits, sensor arrays, and other microelectronic devices. These devices, in turn, were made possible by extending the art of microphotography and by using high-resolution photographic plates, photosensitive resists, and other specialized photographic products. Robert D. Anwyl

Digital Photography

Digital still photography has experienced exponential growth. This is primarily due to a series of technical improvements and price reductions that have made the technology more appealing to the mass market and professionals alike. Parallel growth has been taking place in digital video, and the most recent cameras can be used to capture high-resolution still images or to make lower-resolution video recordings. In addition, the use of digital photography has been increasing in a host of other imaging applications such as surveillance, astronomy, and medical imaging.

Digital photography has advantages over film photography in productivity, quality control, and profitability. Instead of shooting, processing, printing, and scanning, digital photography allows the capture of images directly as digital files, thus saving time and materials.

Sensors. Digital photography is based on the photoelectric effect of silicon. When photons strike silicon, they excite electrons from the valence band to the conduction band. The number of electrons freed is proportional to the size of the light-sensing element, the amount of light exposure, and the wavelength of the light. Silicon light-sensing elements are

arranged in a grid format with each photosite comprising one picture element (pixel). The electrons resulting from light exposure are collected and measured from the pixel array. The charge voltages are converted from analog to digital (A/D) format and amplified to create digital picture files. The cameras have random-access-memory (RAM) buffers to hold images while they are being downloaded or recorded on disk.

Two types of sensors—charge-coupled device (CCD) and complementary metal-oxide semiconductor (CMOS)—are used for the image-sensing elements in digital cameras. CCDs were developed by Bell Labs in the early 1970s for use in satellites and astronomy. By 1985 they were starting to appear in consumer products. CMOS sensors were developed soon after CCDs, but the technology was not extensively used until the early 1990s when products with lower noise and leakage current became available. CCD fabrication relies on older, less efficient metal-oxide semiconductor production methods. Today, CCDs produce higher-quality images and are used in most digital cameras. However, CMOS sensors have inherent size and cost advantages over CCDs and are making inroads on the low end of the image sensor market. *See* CHARGE-COUPLED DEVICES; SILICON.

CCD sensors. CCD photosites are light-receptive over most of the pixel surface, with photodiodes collecting electrons based on the intensity of exposure. Electrons are transferred to the adjacent charge transfer cells in the same column with the photosites. For readout, each row of data is transferred to horizontal charge transfer registers. Charged packets for each row are then are read out serially and sensed by a charge-to-voltage conversion and amplifier section.

CCD sensors have wide dynamic ranges and high light sensitivity, and can operate at 30 frames/second for video applications. They have low noise when correlated double sampling (a process by which a pixel's dark or reference output level is subtracted from its light-induced output) is used to cancel offset variations. They also have good uniformity across photosites because all charges are read through a single readout circuit. CCD pixel elements generally range between 9 and 13 micrometers. Larger pixel elements have greater light sensitivity but lower resolution. Arrays of over 2 million pixels are needed to generate 4 by 6 in. (10 by 15 cm) prints that would be evaluated as excellent.

The CCD architecture makes it impractical to integrate other electronics onto the same chip with the light sensor. CCD cameras require supplemental integrated circuits to support the imaging function. For example, they need a separate chip to generate precisely timed multiple phases of a clock to transfer the photon-generated image from the photosites to the amplifier. Other chips are needed to accomplish the A/D conversions or to supply the several bias voltages that CCDs require. CCDs use 15-volt power supplies, while CMOSs use as little as 2.5 volts. In total, CCD cameras typically require five or more supporting chips. *See* INTEGRATED CIRCUITS.

CCDs are sensitive to infrared radiation, which can cause overheating. Thus, an infrared filter is incorporated in the camera design. Even with these filters, CCDs can overheat with long or multiple exposures, which produces blooming, a defect caused by the overflow of electrons to adjacent photosites. To counteract overheating, manufacturers use special alloys, fans, and circuits to cool CCDs.

CMOS sensors. These have the advantage that many imaging functions are manufactured on the same chip with the light-sensing elements. This leads to smaller, less expensive product designs. Each photosite contains a photodiode that converts light to electrons, a charge-to-voltage conversion section, select/reset transistor, and an amplification section. A grid of metal interconnects overlays the sensor to supply timing and readout signals. Also overlaying the sensor is an array of column output signal interconnects.

The column lines connect to a set of decode and readout (multiplexing) electronics that are arranged by column outside the pixel array. Unlike CCDs, CMOS signals from the entire array, subsections, or a single pixel can be read out by an X-Y addressing technique. Also, unlike CCDs, nondestructive readout and random pixel access are possible with CMOS sensors. These techniques allow noise reduction through signal averaging and adaptive exposure. Contrast can be improved by allowing pixel collecting from dark scenes to have longer exposures than other pixels. CMOS cameras produce no smear because the charge stays within each pixel. They also are not prone to blooming, as CCD cameras are, because the CMOS sensor pixel is a diode held in reverse bias and precharged to a set charge. Photons reduce the charge of the pixel; when it reaches zero, it stops (the pixel is white).

CMOS sensors are of two types: active-pixel sensor (APS) CMOS and active-column sensor (ACS) CMOS. APS CMOS sensors use active pixels, where each photosite contains an amplifier that converts the charge to a voltage. This prevents the percentage of photosite that is light-receptive (fill factor) from being 100%, which reduces the magnitude of the output signal in relation to noise. Furthermore, mismatched amplifiers at the photosites cause differences in the responses of pixels. APS CMOS sensors also suffer from the dark leakage of current, which results in a thermally generated charge that contaminates images.

ACS CMOS, a more recent design, uses a unity gain amplifier where only the input element is present at each photosite. This is physically smaller and allows for higher fill factors. The ACS CMOS design provides the same gain for every photosite, thus improving image quality. ACS CMOS sensors do not convert photon-generated charges to voltages to drive column buses; instead they produce high-proportional currents, which allow them to run faster than APS CMOS sensors.

Image capture. The arrangements of CCDs in digital cameras are divided between linear sensors that scan the image plane and area arrays that capture

the entire image with a single exposure. Linear arrays are capable of capturing higher-resolution images, but they are restricted to still-life photography and are not compatible with strobe lighting, which is commonly used in photo studios. Area arrays, by contrast, produce lower-resolution images, but they can be used for live action photography and are strobe-light-compatible.

Capturing color images requires that three separate signals (blue, green, and red) be generated from a scene. This is generally accomplished by placing separation filters in the light path during exposure. Three approaches are used: filter wheels, dyed photo sites, and silicon filtering.

Filter wheels. These systems require three successive exposures after blue, green, and red filters have been placed in the image path. They have the advantage that each pixel records an image for each primary color and the disadvantage that capture times are greatly increased and live action cannot be captured.

Dyed photosites. These are systems where each pixel has a filter applied to its surface. When this approach is used with linear arrays, it results in a trilinear array configuration where blue, green, and red linear arrays scan the image plane. This has the advantage that each pixel measures all three primary components of light, but the cost of the sensor is greatly increased because three times as many photosites are required.

When dyed photosites are used with area arrays, each pixel is dyed with a primary color. This is usually done in a Bayer configuration where there are two green pixels for each blue and red pixel. This configuration has the advantage that a color image can be captured with a single exposure, thus making it capable of capturing live action scenes. Dyed photosite area arrays have the disadvantage that each pixel receives exposure from only one of the primary colors, but the image file requires picture information for all three primary colors at each pixel. To provide this missing information, the camera's firmware calculates the missing color values from the responses of neighboring pixels. This causes problems with aliasing in areas of fine detail or image edges, and it also complicates and slows the image processing function of cameras.

Dyed photosite area arrays are the most commonly used configuration in current digital cameras due to their ability to capture images in a single short exposure. The image processing algorithms have improved, and the number of photosites in area array sensors has increased to over 6 million in professional-grade cameras. This improves, but does not eliminate, the aliasing and pixelization problems caused by calculating rather than measuring color values for photosites.

An expensive approach to solving this dilemma is the tri-area array configuration, which uses a prism inside the camera to split the incoming light into three images that are individually filtered and focused on three different area array CCDs. This configuration allows single-exposure image capture where each pixel has measured blue, green, and red values. However, the resulting cameras are large and awkward to use in the field, and the engineering demands are daunting.

Silicon filtering. A solution to the problems of dyed photosites is the use of silicon filtering. It has long been known that light will penetrate into a silicon surface to a distance proportional to the wavelength of the light before it creates a hole-electron pair. Longer wavelengths penetrate deeper. A single pixel can capture all three primary colors of light if two critical conditions are met. First, the photodiode needs to be precisely controlled so that it will collect only the desired wavelengths. Second, an optically transparent insulating layer must separate the light-sensing layers. A sensor that has three layers of photoreceptive sites at different depths captures blue, green, and red information at each pixel without the need for external filtration. This eliminates the need for complex algorithms to interpolate the color information from neighboring pixels, thus avoiding errors, reducing camera costs, and reducing the click-to-capture time of cameras.

These sensors are also designed with a unique capability called variable pixel size. The pixels can be instantaneously grouped to form larger sizes that have greater light sensitivity and are desirable for video, auto focusing, and low-light situations. Another advantage of silicon filtering is the reduction of noise, which occurs with dyed photosites due to adjacent pixels having different color filters. Silicon filters are inherently more uniform than dyed photosites because they are precisely grown, flat, crystalline structures. Conventional filters are coated onto the uneven surface of the sensor in the last stages of production and are subject to subsequent fading.

The downside of silicon filtering is that each pixel requires three times the number of support circuits (reset, column access, exposure setting). This increases the size of the pixel, decreases the fill factor, and increases the incidence of defective pixels. Silicon filter designs need to have three separate pixel processing chains including amplifiers and A/D converters. Anthony Stanton

Digital cameras. Like conventional cameras, digital cameras come in compact, single-lens reflex, and large-format varieties. Low-resolution compacts are useful for producing classified advertisements and tend to have relatively simple optics, image-sensing electronics, and controlling software. *See* CAMERA.

Digital cameras are often based on existing single-lens reflex camera designs with the addition of CCD backs and storage subsystems. The capture resolution of these cameras is ideal for news photography and other applications with similar quality requirements.

Large-format systems can take the form of digital backs that can be used in conjunction with existing large-format cameras, providing an interchangeable alternative to a film back. Or they can be designed from the start as digital cameras. These large-format systems have adequate resolution for most graphic

arts purposes, and in most cases do not require their own storage subsystems since the image is downloaded directly to a host computer. They are most suitable for studio work and are widely used in catalog production, where the volume of product shots means that large savings can be made on film and processing.

When using a digital camera, the level and color temperatures of the lighting will inevitably change from shot to shot. A camera will have a fixed ASA/ISO speed equivalent, and if the brightness of the scene is higher than the camera is designed for, blooming can result where the charge induced in one element spills into adjacent elements. Thomas Destree

Bibliography. A. A. Blaker, *Photography: Art and Technique*, 2d ed., 1988; G. Craven, *How Photography Works*, 1986; G. T. Eaton, *Photographic Chemistry*, 2d ed., 1984; P. Green, *Understanding Digital Color*, 2d ed., 1999; T. Grimm, *Basic Book of Photography*, 5th ed., 2003; T. Grimm, *The Basic Darkroom Book*, 3d ed., 1999; R. Jacobson et al., *Manual of Photography: Photographic and Digital Imaging*, 9th ed., 2000; S. Kelby, *The Digital Photography Book*, 2006; M. J. Langford, *Advanced Photography*, 6th ed., 1998; M. J. Langford, *Basic Photography*, 7th ed., 2000; *Photographic Imaging and Electronic Photography* (Technical Pocket Book), 1994; S. F. Ray, *Photographic Chemistry and Processing*, 1994; S. F. Ray, *Photographic Data* (Technical Pocket Book), 1994; S. F. Ray, *Photographic Technology and Imaging Science*, 1994; F. Rost and R. Oldfield, *Photography with a Microscope*, 2000; L. D. Streobel (ed.), *Basic Photographic Materials and Processes*, 2d. ed., 2000; L. D. Stroebel (ed.), *The Focal Encyclopedia of Photography*, 3d ed., 1996.

Photoionization

The ejection on one or more electrons from an atom, molecule, or positive ion following the absorption of one or more photons. The process of electron ejection from matter following the absorption of electromagnetic radiation has been under investigation for over a century. The earliest measurements involved the irradiation of metal surfaces by ultraviolet radiation. The theoretical interpretation of this phenomenon, known as the photoelectric effect, played an important role in establishing quantum mechanics. It was shown that, contrary to classical ideas, energy exchanges between radiation and matter are mediated by integral numbers of photons. The ejection of electrons from a metal requires that the photon energy exceed the binding energy of electrons in the metal (that the wavelength of light be sufficiently small), whatever the photon flux. *See* PHOTOEMISSION.

Photons (light) can interact with atoms or molecules which are free, that is, not bound to a solid or liquid. In the gas phase the photoeffect is called either photoionization (atoms, molecules, and their positive ions) or photodetachment (atomic and molecular negative ions). *See* PHOTODETACHMENT SPECTROSCOPY.

Photoelectron spectroscopy. Photoionization involves a radiative bound-free transition from an initial state consisting of n photons and an atom, molecule, or ion in a bound state to a final continuum state consisting of a residual ion (or an atom in the case of photodetachment) and m free electrons, that is, reaction (1). This means that one or more photons

$$nh\nu + X \rightarrow X^{m+} + me^- \qquad (1)$$

(n) interact with a target atom or ion, resulting in the emission of one or more electrons (m).

In the simplest atomic photoionization process a single electron is ejected from an atom following the absorption of a single photon (**Fig. 1**). Each mode of photoionization defines a final-state channel that is characterized by the energy and angular momentum of the outgoing electron as well as the excitation state of the residual ion. Since the photoionization process is endoergic, each channel has a well-defined threshold energy below which the channel is energetically closed. The threshold photon energy for a particular channel is equal to the binding energy of the electron that is to be ejected plus the excitation energy, if any, of the residual ion.

Above threshold, the energy carried off by the outgoing electron represents the balance between the energy supplied by the photon and the binding energy of the electron plus the excitation energy of the residual ion (neglecting the small recoil of the heavy ion). A photoelectron energy spectrum is characterized by a discrete set of peaks at different photon energies, each peak being associated with a particular state of the residual ion. Information on the excitation state of the ion following photoionization can also be obtained by monitoring the fluorescence emitted in the subsequent radiative decay of the state. One of the earliest applications of photoionization measurements was the investigation of the structure of atoms by determining the binding energies of both outer- and inner-shell electrons by means of photoelectron spectroscopy. An example of a photoionization measurement (in fact a

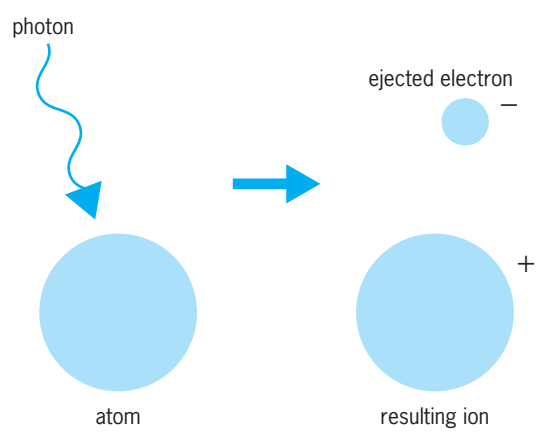

Fig. 1. Ionization of an atom by a photon results in emission of an electron and leaves a residual ion.

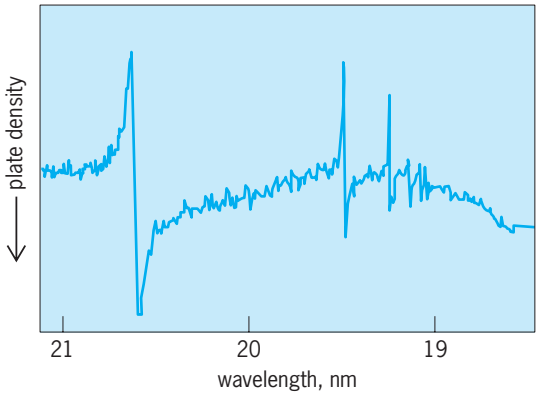

Fig. 2. Photoionization spectrum for helium atoms measured using synchrotron radiation and a photographic method (electronic detection would be used in a modern experiment). Each peak as a function of decreasing photon wavelength (increasing photon energy) corresponds to a photoexcitation process in which both electrons in a helium atom are excited, resulting in photoionization (emission of an electron) with the residual ion left in different states of excitation. (*After R. P. Madden and K. Codling, New autoionizing atomic energy levels in He, Ne, and Ar, Phys. Rev. Lett., 10:516–518, 1963*)

very early measurement from 1963 performed using synchrotron radiation) is the excitation of the two electrons of a helium atom followed by autoionization (**Fig. 2**). Absorption peaks are shown at different photon energies (wavelengths), with each peak corresponding to the helium-plus ion in a different state of excitation (one electron is photoionized). Similar but much more detailed photoionization spectra are obtained using more modern methods. *See* ELECTRON SPECTROSCOPY; FLUORESCENCE.

Measurements. Photoionization measurements involve the illumination of a target of atoms, molecules, or ions with a beam of photons from a light source. The binding energies of electrons in atoms and positive ions are sufficiently large that ionization thresholds lie in the ultraviolet or x-ray region of the electromagnetic spectrum. This generally limits the use of lasers as light sources for photoionization measurements. Traditionally ultraviolet lamps and, subsequently, synchrotron radiation sources have been used in photoionization studies. Lasers, however, are the most frequently used light source for photodetachment studies since the weakly bound electrons in negative ions correspond to thresholds in the infrared and visible. Sources of atoms and molecules are usually static or flowing gas targets, while positive and negative ions are most often produced in the form of beams by using an accelerator. Electromagnetic traps have also been used in photodetachment measurements. The photoionization process can be monitored by studying the absorption of the incident light or by using a more sensitive method based upon detecting one of the products of the interaction, that is, photoelectrons or residual ions. *See* LASER; MOLECULAR BEAMS; PARTICLE TRAP; SYNCHROTRON RADIATION.

Electron emission pattern. The selection rules for bound-free transitions specify the allowed changes in the quantum numbers characterizing the initial and final states of photoionization. In particular, for an electric dipole-induced process, the orbital angular momentum of the ejected electron is limited to the values $l = l_0 \pm 1$, where l_0 refers to the angular momentum of the bound electron prior to ejection. The outgoing electron carries off this angular momentum, which is manifested in its angular distribution and spin polarization. The partial waves associated with $l = l_0 \pm 1$ interfere as a function of photon energy to determine the shape of the photoelectron emission pattern. It has been demonstrated by using symmetry arguments that the form of the angular distribution for linearly polarized radiation in the electric dipole approximation, assuming an initially randomly oriented target, is given by Eq. (2).

$$f(\theta) = \frac{1 + \beta(3\cos^2\theta - 1)}{2} \qquad (2)$$

Here θ is the angle between the photoelectron emission direction and the polarization vector of the radiation, and β represents the asymmetry parameter (a measure of the deviation from isotropy, $\beta = 0$). This parameter, which characterizes the angular distribution, is restricted to the range $-1 \le \beta \le 2$. The value of $\beta = 2$ corresponds to the classical dipolar distribution ($\cos^2\theta$), where the electrons are preferentially ejected in the direction of the polarization vector. Values of $\beta \neq 2$ are the result of interferences between the two partial waves representing the electron. Angular distributions are studied by using angle-resolved photoelectron spectroscopy. *See* ANGULAR CORRELATIONS.

Cross section. The probability that an atom becomes ionized by the absorption of a photon is measured by the photoionization cross section, which is typically 0.1–10 megabarns (1 Mb = 10^{-18} cm^2). To understand the detailed manner in which an atom interacts with radiation requires an investigation of this cross section as a function of photon energy. Above threshold, the partial cross section for a particular channel decreases monotonically with photon energy. New channels open when the photon energy exceeds excited-state threshold values.

A photoionization cross section can also exhibit sharp changes over a narrow range of photon energies. This structure is the result of interference between resonant and direct (nonresonant) pathways to the same final state. The resonant pathway involves the transient photoexcitation of a doubly excited intermediate state which, in turn, rapidly decays by the spontaneous process of autoionization. Resonance parameters such as the energy and linewidth (inversely related to the lifetime against autoionization) can be found by fitting the experimental data to an analytical form called the Fano-Beutler function. The example shown in Fig. 2 illustrates a lineshape typical of interference between direct and indirect photoexcitation pathways.

The threshold dependence of the cross section is fundamentally different for photoionization and photodetachment. In the final state the ejected electron moves in a strongly attractive and long-range

Coulomb field of an ion in the case of photoionization and a weakly attractive and short-range field of a neutral atom or molecule in the case of photodetachment. The enhanced role of correlation in photodetachment is due to the absence of the normally dominant final-state Coulomb interaction. The short-range nature of the final-state interaction in photodetachment causes the cross section in the near vicinity of a threshold to exhibit a dependence on the orbital angular momentum, l, of the outgoing electron. The cross section is zero at threshold and rises to a maximum just beyond threshold. In contrast, because of the long-range final-state interaction, photoionization cross sections are finite at threshold, independent of angular momentum of the outgoing electron.

Use of synchrotron radiation sources. The availability of intense beams of high-energy photons from advanced synchrotron radiation light sources facilitates the investigation of photoionization of inner-shell electrons in atoms and inner- and outer-shell electrons in positive ions. The high brightness of modern light sources also allows studies with very high spectral resolution. Spectroscopic studies of ejected photoelectrons as well as subsequently emitted Auger electrons lead to a better understanding of the final-state dynamics of the photoionization process when an inner-shell vacancy relaxes. *See* AUGER EFFECT.

Resonance ionization spectroscopy. Resonance ionization spectroscopy is a sensitive and selective analytical method based on the photoionization of atoms. Typically it involves two or more resonant photoexcitation steps followed by an ionization step. Similar measurements designed to determine cross sections for photoionization from excited atoms have also employed a sequential photoabsorption scheme. The atom is prepared in an excited state by laser photoexcitation and subsequently photoionized by the absorption of synchrotron radiation. *See* RESONANCE IONIZATION SPECTROSCOPY.

Multiphoton and multielectron processes. There is increasing interest in photoionization processes that go beyond the most probable process of single-photon absorption and single-electron ejection. The high intensity of lasers has made it possible, for example, to study multiphoton absorption processes in which a single electron is ionized by the absorption of two or more photons, each with an energy less than the ionization energy. Interestingly, it has been observed in a process called above-threshold ionization that under certain conditions more photons are absorbed than the minimum number needed to reach the ionization threshold. If the intensity of the laser beam becomes very high, it is possible to observe effects associated with the electron being strongly driven by the oscillating laser field. This quiver motion can modify the energy of the ejected electron.

The availability of intense beams of high-energy photons from lasers and synchrotron sources has facilitated the investigation of ionization processes involving the simultaneous ejection of more than one electron following the absorption of a single photon. Double photodetachment, for example, is particularly interesting near threshold since it involves a three-body final state consisting of two electrons essentially at rest in the field of a singly charged core. The electron correlations that develop in this system are of fundamental interest in atomic structure theory. *See* ATOMIC STRUCTURE AND SPECTRA; LASER SPECTROSCOPY.

David J. Pegg; Alfred S. Schlachter

Bibliography. J. Berkowitz, *Photoabsorption, Photoionization and Photoelectron Spectroscopy*, 1979; R. R. Cordermann and W. C. Lineberger, Negative ion spectroscopy, *Annu. Rev. Phys. Chem.*, 30:347–376, 1979; V. S. Letokhov, *Laser Photoionization Spectroscopy*, 1987; S. T. Manson and A. F. Starace, Photoionization angular distributions, *Rev. Mod. Phys.*, 54:389–405, 1982; J. A. R. Samson, Photoionization of atoms and molecules, *Phys. Rep.*, 28:303–354, 1976; V. Schmidt, Photoionization of atoms using synchrotron radiation, *Rep. Prog. Phys.*, 55:1483–1659, 1992.

Photoluminescence

A luminescence excited in a body by some form of electromagnetic radiation incident on the body. The term photoluminescence is generally limited to cases in which the incident radiation is in the ultraviolet, visible, or infrared regions of the electromagnetic spectrum; luminescences excited by x-rays or gamma rays are generally characterized by special names. The graph of luminous efficiency per unit energy of the exciting light absorbed versus the frequency of the exciting light is called the excitation spectrum. The excitation spectrum is determined by the absorption spectrum of the luminescent body, which it often closely resembles, and by the efficiency with which the absorbed energy is transformed into luminescence.

Photoluminescence may be either a fluorescence or a phosphorescence, or both. Energy can be stored in certain luminescent materials by subjecting them to light or some other exciting agent, and can be released by subsequent illumination of the material with light of certain wavelengths. This type of photoluminescence is called stimulated photoluminescence. In contrast to normal photoluminescence, which is constant in intensity as long as the intensity of the exciting light does not vary, stimulated photoluminescence decreases in intensity as the stored energy is released. *See* FLUORESCENCE; LUMINESCENCE; PHOSPHORESCENCE.

Clifford C Klick; James H. Schulman

Photolysis

The decomposition of matter due to absorption of incident light. When photolysis occurs, it causes definite changes in the chemical composition of the illuminated material. For example, illumination of microcrystals of silver bromide embedded in gelatin

results in formation of metallic silver, and is the basis of the photographic process. *See* PHOTOGRAPHIC MATERIALS.

Reactors and features. Numerous metal complexes (for example, oxalate salts), azides, nitrides, and sulfides, and most organometallic compounds undergo decomposition upon illumination, often with concomitant evolution of a gaseous product. In the presence of water and oxygen, illumination of many semiconductors results in their corrosion; for example, an aqueous suspension of cadmium sulfide (CdS) undergoes rapid decomposition upon irradiation with sunlight. Such reactions are especially pronounced if the semiconductor is present in a colloidal form. The outcome of these photochemical processes can often be controlled by addition of adventitious materials. For example, illumination of cadmium sulfide in aqueous solution containing hydrogen sulfide (H_2S) and colloidal platinum results in evolution of hydrogen gas. Here, the semiconductor photosensitizes (or photocatalyzes) decomposition of hydrogen sulfide into its elements, and the reaction can be used to remove sulfides from industrial waste. The same reaction with water used in place of hydrogen sulfide represents the ideal goal of many attempts at the conversion and storage of solar energy. Thus, photolysis of inorganic materials plays an important role in both effluent treatment and energy conversion. *See* COORDINATION COMPLEXES; SEMICONDUCTOR.

Organic materials. The photolysis of organic materials is a particularly rich subject. Many ketones, for example, acetone, abstract hydrogen atoms from adjacent organic matter under illumination. According to the circumstances, the resultant free radicals may be used to initiate polymerization of a suitable monomer or to cause decomposition of a plastic film. Both reactions have important commercial applications, and photoinitiators are commonly used for emulsion paints, inks, polymers, explosives, and fillings for teeth and for development of photodegradable plastics. Photolysis of carbonyl compounds, released into the atmosphere by combustion of fossil fuels, is responsible for the onset of photochemical smog. Many other types of photochemical transformation of organic molecules are known, including isomerization of unsaturated bonds, cleavage of carbon-halogen bonds, olefin addition reactions, halogenation of aromatic species, hydroxylation, and oxygenation processes. Indeed, photochemistry is often used to produce novel pharmaceutical products that are difficult to synthesize by conventional methods. *See* FREE RADICAL; PHOTODEGRADATION; SMOG.

Photolysis causes a photochemical reaction, as distinguished from a thermal reaction. The most important photochemical reaction is green plant photosynthesis. Here, chlorophyll that is present in the leaves absorbs incident sunlight and catalyzes reduction of carbon dioxide to carbohydrate. This is the origin of all fossilized fuels. The reaction also results in oxidation of water to molecular oxygen and, therefore, can be said to be primarily responsible for all life on Earth. There are other photochemical reactions that serve important functions in biology and medicine. One demonstration of the power of photochemistry concerns the treatment of neonatal jaundice. Newborn infants suffering from this illness, which is caused by proliferation of bilirubin, are exposed to sunlight. This causes conversion of bilirubin to a form that is more soluble in water and is rapidly excreted from the body. *See* BILIRUBIN; PHOTOSYNTHESIS.

An important feature of all photochemical reactions is that only absorbed light is effective in causing photochemical change. This fact permits photolysis of particular materials simply by optimum selection of the wavelength of incident light. Thus, the ozone layer prevents ultraviolet light from reaching the Earth's surface and thereby acts as a protection against extensive damage that would result from light absorption by biological tissue. One dye in a mixture of similar reagents can be illuminated selectively by using monochromatic light, allowing control of the ensuing chemistry. Furthermore, photonic energy absorbed by one molecule can be transferred, according to certain laws, to similar or different molecules. This process, known as photosensitization, allows transparent materials, such as water, to be photolyzed.

Singlet molecular oxygen. Excitation of an organic molecule results in spontaneous generation of the singlet excited state. In most molecules, this highly unstable excited singlet state may undergo an intersystem-crossing process that results in population of the corresponding (less energetic) excited triplet state, in competition to fluorescence. The excited triplet state, because of spin restriction rules, retains a significantly longer lifetime than is found for the corresponding excited singlet state, and may be formed in high yield.

Almost without exception, these triplet states react quantitatively with molecular oxygen (O_2) present in the system via a triplet energy-transfer process. The resulting product is singlet molecular oxygen. This species is a potent and promiscuous reactant, and it is responsible for widespread damage to both synthetic and natural environments. Indeed, plants and photosynthetic bacteria contain carotenoids to protect the organism against attack by singlet oxygen. The same species is known to be responsible, at least in part, for photodegradation of paint, plastic, fabric, colored paper, and dyed wool. Secondary reactions follow from attack on a substrate by singlet oxygen, resulting in initiation of chain reactions involving free radicals. However, modern technological processes have evolved in which singlet molecular oxygen is used to destroy unwanted organic matter, such as tumors, viruses, and bacteria, in a controlled and specific manner. In photodynamic therapy a dye is injected into a tumor and selectively illuminated with laser light. The resultant singlet oxygen destroys the tumor. Similar methodology can be used to produce photoactive soap powders, bleaches, bactericides, and pest-control reagents. Mosquito larvae, for example, can be destroyed by exposure to sunlight after staining with a

photoactive dye. *See* CHAIN REACTION (CHEMISTRY); FLUORESCENCE; TRIPLET STATE.

Flash photolysis. Irradiation of a material with a high-intensity, short-duration pulse of light, delivered from a flash lamp or pulsed laser, can be used to create a high concentration of transient species. This experimental technique, known as flash photolysis, is used to characterize reactive intermediates formed during photolysis. The sample, which may be in the form of a solution, vapor, film, or powder, is irradiated with a single pulse of light of appropriate wavelength in order to generate the initial excited state. After this excitation pulse, a much weaker pulse of light is directed onto the sample in order to record an ultraviolet-visible absorption spectrum of the excited state. By delaying the monitoring pulse with respect to the excitation pulse, a series of spectra can be recorded at different delay times. In this manner the course of reaction can be followed, and important intermediates can be identified. By monitoring formation and deactivation of particular intermediates, kinetic information can be elicited, and the overall mechanism of the photoprocess can be evaluated. In addition to ultraviolet-visible absorption spectra, intermediates may be detected by infrared, Raman, resonance Raman, electron paramagnetic resonance, circular dichroism, magnetic circular dichroism, luminescence, and diffuse reflectance spectroscopic techniques. Time-resolved conductivity, thermal lensing, and optoacoustic calorimetry can be used to gain further information about the photochemistry of the system and to detect heat released into the system during nonradiative decay of transient species. By including compounds possessing certain properties (that is, easily abstractable hydrogen atoms, low oxidation potential, molecular oxygen) in the system, it becomes possible to explore secondary processes and to measure rates of reaction. Transient spectral and kinetic data can be collected on time scales as short as 100 femtoseconds by using readily available laser technology. *See* LASER SPECTROSCOPY.

Bacterial photosynthesis. The most impressive illustration of laser flash photolysis concerns the elucidation of the chemical sequence occurring within bacterial photosynthetic reaction center complexes. Thus, incident sunlight is harvested by arrays of chlorophylls, xanthophylls, and related pigments and transported throughout the array until reaching a trap. Energy migration occurs by way of rapid transfer between closely spaced (1.0–1.5-nanometer) pigment molecules, each transfer taking about 1 picosecond. The trap is a pair of bacteriochlorophyll molecules held in an approximate face-to-face configuration. Upon excitation, an electron is transferred to a bacteriopheophytin molecule, situated some 1.7 nm distant, on a time scale of a few picoseconds. The electron is subsequently transferred to secondary (quinones) acceptors, while secondary (hemes) electron donors serve to reduce the oxidized primary donor. In this manner, charge can be transferred across a membrane and used, via a series of biochemical processes, to produce valuable fuels. The rates and efficiencies of each of the electron transfer steps have been elucidated for various types of photosynthetic bacteria and, with less success, for many of the events occurring in the reaction centers of green plants. *See* PHOTOCHEMISTRY. A. Harriman

Bibliography. V. Balzani and F. Scandola, *Supramolecular Photochemistry*, 1990; J. A. Barltrop and J. D. Coyle, *Principles of Photochemistry*, 1978; G. J. Ferraudi, *Elements of Inorganic Photochemistry*, 1988; K. Kalyanasundaram, *Photochemistry of Polypyridine and Porphyrin Complexes*, 1992; K. Kalyanasundaram and M. Grätzel (eds.), *Photosensitization and Photocatalysis Using Inorganic and Organometallic Compounds*, 1993; N. J. Turro, *Modern Molecular Photochemistry*, 1991.

Photometer

An instrument used for making measurements of light, or electromagnetic radiation, in the visible range. In general, photometers may be divided into two classifications: laboratory photometers, which are usually fixed in position and yield results of high accuracy; and portable photometers, which are used in the field or outside the laboratory and yield results of lower accuracy. Each class may be subdivided into visual (subjective) photometers and photoelectric (objective or physical) photometers. These in turn may be grouped according to function, such as photometers to measure luminous intensity (candelas or candlepower), luminous flux, illumination (illuminance), luminance (photometric brightness), light distribution, light reflectance and transmittance, color, spectral distribution, and visibility. Since visual photometric methods have largely been supplanted commercially by physical methods, they receive only casual attention here, but because of their simplicity, visual methods are still used in educational laboratories to demonstrate photometric principles. *See* ILLUMINANCE; LUMINANCE; LUMINOUS FLUX; LUMINOUS INTENSITY.

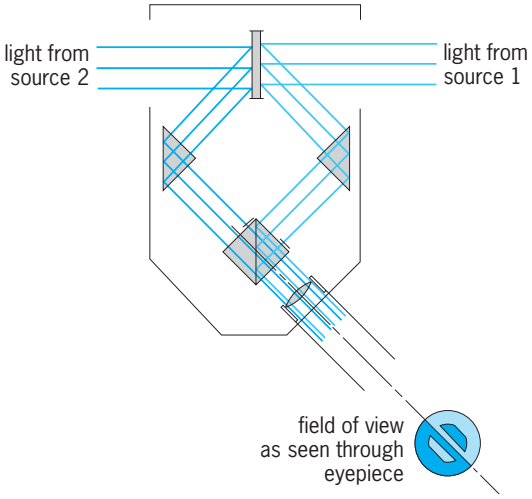

Fig. 1. Lummer-Brodhun contrast sight box.

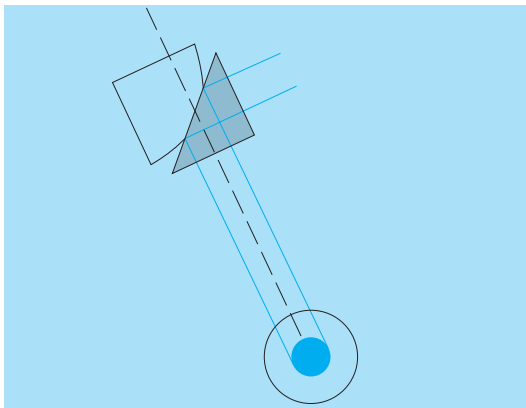

Fig. 2. Diagrammatic sketch of Lummer-Brodhun equality of luminance prism arrangement and appearance of field of view.

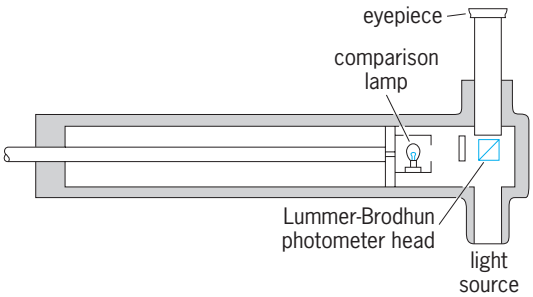

Fig. 3. Essentials of a Macbeth illuminometer.

Visual photometers. These are luminance comparison devices, and most of them utilize the Lummer-Brodhun sight box or some adaptation of its principles. The sight box is an optical device for viewing simultaneously the two sides of a white diffuse plaster screen illuminated by the light sources that are being compared. The general arrangement of the contrast sight box and the field of view as seen through the eyepiece are shown in **Fig. 1**. The contrast sight box gives greater sensitivity of observation than the equality of luminance prism arrangement and appearance of field of view (**Fig. 2**). If there is a color difference in the two light sources being compared, it may be difficult to make a photometric balance. In this case, methods of heterochromatic photometry, for example, the flicker photometer, may be used.

Flicker photometer. In this photometer a single field of view is alternately illuminated by the light sources to be compared. The rate of alternation is set at such a speed that the color differences disappear; the photometric balance is obtained by moving the sight box back and forth, the criterion of equality of luminance of the two sources being the disappearance of flicker. Other systems of heterochromatic photometry are the cascade method, the compensation of mixture method, and the use of color-matching or color-equalizing optical filters.

Illuminometer. An illuminometer is a portable visual photometer, as shown in cross section in **Fig. 3**. The light to be measured enters the instrument and is balanced by a Lummer-Brodhun sight box of the equality of luminance type against a comparison lamp, which is moved along the tube. A control box supplies a calibrated current to the comparison lamp, and calibrated optical filters can be placed in the light paths to correct for color differences in the comparison and measured sources and to extend the range of the instrument.

Luminance meter. This instrument measures photometric brightness and may be of the visual or photoelectric type; it is a major tool for research scientists in visual psychology and illuminating engineering. **Figure 4** shows a photoelectric luminance photometer; a separate unit combines the power supply with the controls and readout meter. The photometer has a telescopic viewing system for imaging the bright surface to be measured on the cathode of a photoemissive tube.

The visual type of illuminometer can also be used to measure the brightness of any surface which is large enough to fill the field of view of the instrument.

Photoelectric photometers. Photometers that use barrier-layer (photovoltaic) photocells and photoemissive electronic tubes as light receptors are classified as photoelectric photometers. Since barrier-layer cells generate their own current when

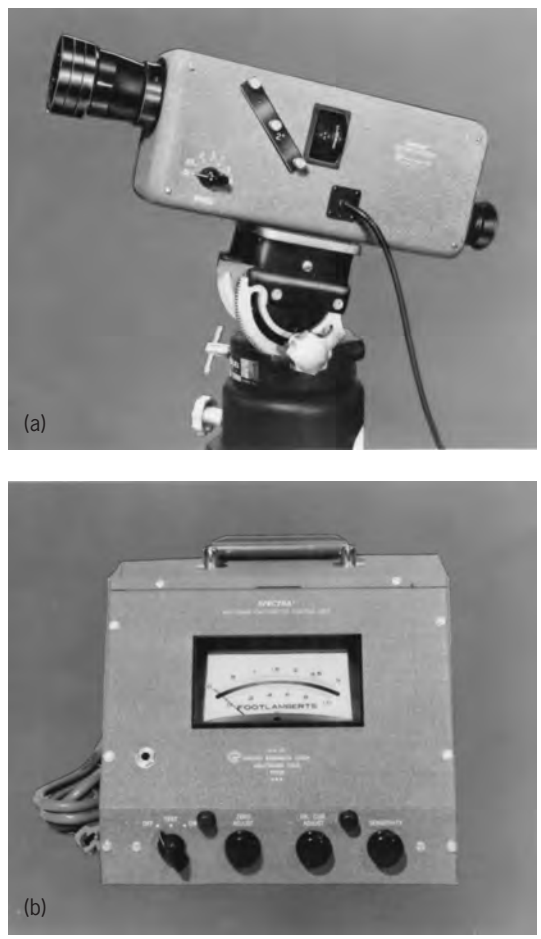

Fig. 4. Spectra Pritchard photoelectric luminance photometer. (*a*) Viewing system. (*b*) Readout meter. (*Photo Research Corp.*)

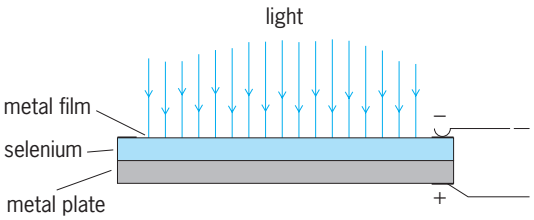

Fig. 5. Barrier-layer photocell.

illuminated and therefore do not require a power supply for operation, they may be used either in the laboratory or in the portable instruments. The type generally used consists of an iron plate, often circular, coated with a thin layer of selenium, which in turn is covered with a very thin, transparent film of metal such as gold or platinum (**Fig. 5**). A ring of metal is sprayed around the edge of this film; when this metal ring is connected to one terminal of a sensitive microammeter or galvanometer and the other terminal is connected to the iron plate, a current flows through the instrument when light penetrates the metal film to the selenium layer. The photoemissive tube photometers, on the other hand, require an external power supply, and so they are used mainly in laboratory work. Both types of cells should be fitted with optical filters to correct approximately their spectral response to the standard luminosity curve of the eye.

With the advent of semiconductor electronics, it is possible to construct current amplifiers for use with barrier-layer photocells, since both devices have inherent low-resistance characteristics. The linearity of these cells is also a function of the voltage developed across its terminals. To obtain maximum linearity, the load on the cell should approach zero. Semiconductor operational amplifiers have also been found to fulfill this requirement. Transistor amplifiers can be used with barrier-layer cells to drive digital voltmeters, which have made possible the semiautomation of most observational work in photometry.

Photoelectric illuminometer. The measurement of illumination has been revolutionized since 1931 by the introduction of photoelectric portable photometers employing rugged barrier-layer cells. These instruments eliminate the tedious visual comparison required with previous portable visual illuminometers and have considerably simplified illumination measurements. Barrier-layer cells generate a small electric current of about 6 or 7 microamperes/footcandle when light in the visible range falls upon them. The electric instrument used to measure the cell output should have low resistance (100 ohms or less), making the cell current nearly proportional to the luminous flux falling upon it. The photocell should be corrected to correspond to Lambert's cosine law of incidence. *See* PHOTOMETRY.

In this type of cell the current attains its final value after a short time because of the effects of fatigue and temperature. These effects are minimized by low resistance in the instrument circuit.

Barrier-layer cell photometers are also used extensively in photographic light meters to determine proper lighting and camera diaphragm openings. For such applications the instrument can be calibrated on an arbitrary scale. A simple manual computer is often included to convert the meter reading into the proper diaphragm opening and exposure time for a given type of film. *See* PHOTOGRAPHY.

Photoemissive tube photometer. This photometer utilizes a photoelectric tube whose sensitivity is much lower than that of a barrier-layer cell. This device is more accurate than the barrier-layer cell, and the linearity and stability are sufficient for highest-precision photometric work. Since electronic amplification is required to increase the sensitivity to a practicable level, this type of photometer is not usually used outside the laboratory. In calibration, the dark current which flows when the tube is not illuminated is an important factor and should be subtracted or eliminated by circuitry.

Integrating-sphere photometer. Any of the above photometers can be used with an integrating sphere to measure the total luminous flux of a lamp or luminaire. This is ordinarily made as an Uhlbricht sphere (**Fig. 6**), although other geometrical shapes can be used. The inside surface has a diffusely reflecting white finish which integrates the light from the source. A white opaque screen prevents direct light from the source from falling on a diffusing window, which is illuminated only by diffusely reflected light from the sphere walls. The brightness of this window, which is measured by the photometer outside the sphere, is then directly proportional to the flux emitted by the source if the inside finish of the sphere is perfectly diffusing.

Reflectometer. This instrument serves the dual purpose of a reflectometer and transmissometer and combines integrating spheres and barrier-layer photocells (**Fig. 7**). The absolute reflectance of test surfaces can be determined by measuring the total reflected light when a beam of light strikes the

Fig. 6. Uhlbricht sphere for measuring luminous flux and efficiency. (*Westinghouse Electric Corp.*)

surface. Transmittance can be measured by placing a flat sample of the material in the opening between the two spheres, with the bottom sphere containing the light source for transmittance measurements, and the upper one the light-measuring cells and a collimated beam of light for reflectance measurements.

Light-distribution photometer. One of the most frequently used photometric devices, this measures luminous intensity at various angles from lamps, luminaires, floodlights, searchlights, and the like. Although the most direct method is to tilt the light source so that measurements can be made in various directions, this may be impracticable with sources whose light output is position-sensitive. It is therefore sometimes necessary to use mirror arrangements similar to the one in **Fig. 8**. The source to be measured is placed at a fixed point on the photometric axis (the horizontal line passing through the centers of the source and photocell). The mirror system can then be moved bodily about the axis so that the light emitted by the source in any direction in a vertical plane perpendicular to the photometric axis is reflected to the photocell.

Fig. 7. Dual-purpose reflectometer transmissometer, in which integrating spheres and barrier-layer cell photometers are combined. (*General Electric Co.*)

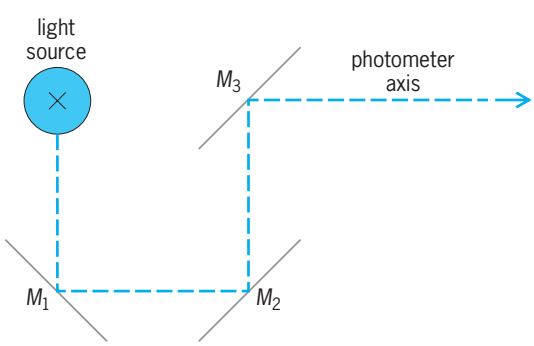

Fig. 8. Mirror system to measure light distribution.

Fig. 9. Modern automated distribution photometer. (*Westinghouse Electric Corp.*)

Fig. 10. Luckiesh-Moss visibility meter. (*General Electric Co.*)

Modern methods of automation have been applied to distribution photometers. They may be controlled by punched cards or tape to give complete programming of the photometer so that no human attention is necessary during the test. The photometric data can be automatically recorded, and curves of luminous intensity and footcandle intensity can be automatically drawn. A modern automated distribution photometer is shown in **Fig. 9**.

Spectrophotometer. Measurements of spectral energy distribution from a light source are made by this device. It measures the energy in small wavelength bands by means of a scanning slit, and the results are presented as a spectral distribution curve.

Visibility meters. These operate on the principle of artificially reducing the visibility of objects to threshold values (borderline of seeing and not seeing) and measuring the amount of that reduction on an appropriate scale. The Luckiesh-Moss visibility meter is probably the best known and easiest to use. The instrument (**Fig. 10**) consists of two variable-density filters (one for each eye) so adjusted that a visual task seen through them is just barely discernible. Readings are on a scale of relative visibility related to a standard task.

G. A. Horton

Photometry

That branch of science which deals with measurements of light (visible electromagnetic radiation) according to its ability to produce visual sensation. Specifically, photometry deals with the attribute of light that is perceived as intensity, while the related attribute of light that is perceived as color is treated in colorimetry. *See* COLOR.

The purely physical attributes of light such as energy content and spectral distribution are treated in radiometry. Sometimes the word photometry is used to denote measurements that have nothing to do with human vision, but this is a mistake according to modern usage. Such measurements are properly referred to as radiometry, even if they are performed in the visible spectral region. *See* RADIOMETRY.

Relative visibility. The relative visibility of a fixed power level of monochromatic electromagnetic radiation varies with wavelength over the visible spectral region (**Fig. 1**). The relative visibility of radiation also depends upon the illumination level that is being observed. The cone cells in the retina determine the visual response at high levels of illumination, while the rod cells dominate in the dark-adapted eye at very low levels (such as starlight). Cone-controlled vision is called photopic, and rod-controlled vision is called scotopic, while the intermediate region where both rods and cones play a role is called mesopic. *See* VISION.

Originally, photometry was carried out by using the human visual sense as the detector of light. As a result, photometric measurements were subjective. That is, two observers with different relative visibility functions, starting with the same standard and following identical procedures, would assign different values to a test radiation having a different spectral distribution than that of the standard.

In order to put photometric measurements on an objective basis, and to allow convenient electronic detectors to replace the eye in photometric measurements, the Commission Internationale de l'Eclairage (CIE; International Commission on Illumination) has adopted two relative visibility functions as standards. These internationally accepted functions are called the spectral luminous efficiency functions for photopic and scotopic vision, and are denoted by $V(\lambda)$ and $V'(\lambda)$, respectively. *See* LUMINOUS EFFICIENCY.

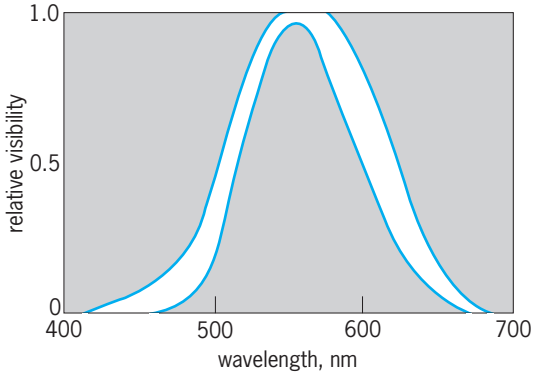

Fig. 1. Range of the relative visibility functions of 125 observers as measured by W. W. Coblentz.

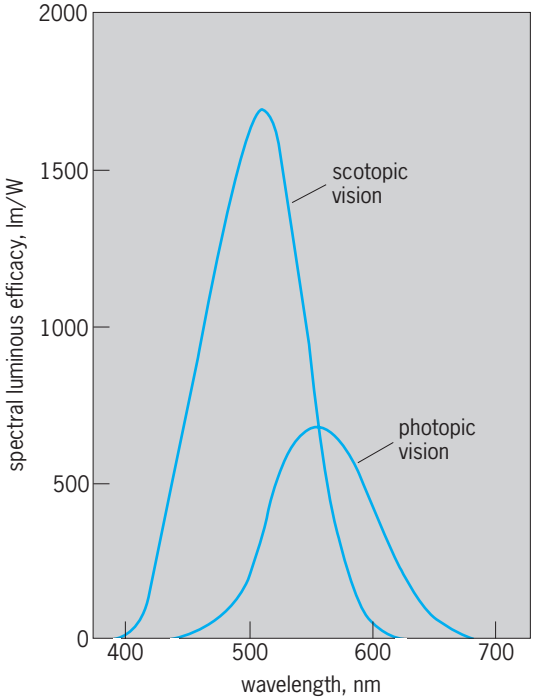

Fig. 2. Internationally accepted standard values for the spectral luminous efficacy of monochromatic radiation for human scotopic (dark-adapted) and photopic (light-adapted) vision.

Thus photopic and scotopic (but not mesopic) photometric quantities have objective definitions, just as do the purely physical quantities. However, there is a difference. The purely physical quantities are defined in terms of physical laws, whereas the photometric quantities are defined by convention. In recognition of this difference the photometric quantities are called psychophysical quantities.

Photometric units. According to the International System of Units, SI, the photometric units are related to the purely physical units through a defined constant called the maximum spectral luminous efficacy. This quantity, which is denoted by K_m, is the number of lumens per watt at the maximum of the $V(\lambda)$ function. K_m is defined in SI to be 683 lm/W for monochromatic radiation whose wavelength is 555 nanometers, and this defines the photometric units with which the photometric quantities are to be measured.

At various times, the photometric units have been defined in terms of the light from different standard sources, such as candles made according to specified procedures, and blackbodies at the freezing point of platinum. According to these definitions, K_m was a derived, rather than defined, quantity. *See* ILLUMINATION; PHYSICAL MEASUREMENT; UNITS OF MEASUREMENT.

Spectral luminous efficacy. The products of K_m and the spectral luminous efficiencies are called the spectral luminous efficacies denoted by $K(\lambda)$. The spectral luminous efficacy functions for photopic and scotopic vision are shown in **Fig. 2**. Currently, there are no internationally accepted spectral luminous efficiency functions for mesopic vision. Thus there are no objectively defined mesopic quantities in use in photometry. However, research suggests the usefulness of linear combinations of the curves shown in Fig. 2, where the relative weight assigned to each curve depends upon the level of illumination. *See* LUMINOUS EFFICACY.

Photometric measurements. Photometric measurements are made by a number of different means. The human eye is still used occasionally, particularly in the mesopic region where no other internationally accepted standard exists. It is also common to make photometric measurements with a detector-optical filter combination in which the filter has been chosen to make the relative spectral response of the combination closely resemble the $V(\lambda)$ function. It is sometimes difficult to get a good enough fit of the spectral response function to the $V(\lambda)$ function over a sufficiently broad spectral region to make accurate measurements on exotic light sources. It is becoming accepted practice to measure the spectral distribution of these sources by conventional spectral radiometric techniques, and to calculate the photometric quantities by multiplying the spectral quantities by $K(\lambda)$ and integrating with respect to wavelength. *See* LIGHT.

Jon Geist

Bibliography. R. McCluney, *Introduction to Radiometry and Photometry*, 1994.

Photomorphogenesis

The regulatory effect of light on plant form involving growth, development, and differentiation of cells, tissues, and organs. Morphogenic influences of light on plant form are quite different from light effects that nourish the plant through photosynthesis, since the former usually occur at much lower energy levels than are necessary for photosynthesis. While photosynthetic processes are energy-transducing, converting light energy into chemical energy for running the machinery of the plant, light serves only as a trigger in photomorphogenesis, frequently resulting in energy expenditure orders of magnitude larger than the amount required to induce a given response. Photomorphogenic processes determine the nature and direction of a plant's growth and thus play a key role in its ecological adaptations to various environmental changes. *See* PHOTOSYNTHESIS.

Morphogenically active radiation is known to control seed and spore germination, growth and development of stems and leaves, lateral root initiation, opening of the hypocotyl or epicotyl hook in seedlings, differentiation of tracheary elements in the stem, and form changes in the gametophytic phase of ferns, to mention but a few of such known phenomena. Many nonmorphogenic processes in plants are also basically controlled by light independent of photosynthesis. Among these are chloroplast movement, biochemical reactions involved in the synthesis of flavonoids, anthocyanins, chlorophyll, and carotenoids, and leaf movements in certain legumes such as *Albizzia*, *Mimosa*, and *Samanea*. A large number of these processes are affected only by wavelengths of light at the red end of the spectrum. The similarities in the action spectra for such different systems led to the concept that these divergent sets of developmental and physiological processes might be controlled by a single pigment. This has proved to be the case, the pigment being phytochrome. *See* CHLOROPHYLL; LEAF; ROOT; SEED; STEM.

Others of these processes, such as rapid stem and internode growth inhibition, directional organ growth (phototropism), and stomatal aperture regulation, are controlled primarily by wavelengths in the near-ultraviolet (UV-A) and blue regions of the spectrum. Although many of their action spectra are similar, suggesting a single pigment as with responses controlled by phytochrome, at least two distinct classes of nonphytochrome pigments, the cryptochromes and the phototropins, are now known to control these latter responses.

Pigment characterization. Phytochrome itself is a blue or blue-green pigment consisting of a protein (component) moiety and a light-absorbing group (chromophore). The chromophore, which is responsible for the sensitivity of the pigment to light, is a linear tetrapyrrole of the bilin type—that is, a molecule containing four pyrrole rings linked together in a linear fashion by single-atom bridges, whose parental state is fully oxidized. Phytochrome is closely related to the phycobilins, photosynthetic accessory pigments found in certain algae, as well as to the pigment in human bile. Its molecule exists in two forms, each of which can be converted to the other by irradiation with different regions of the spectrum. One form, P_r, absorbs maximally in the red (665-nanometer) region of the spectrum, and the other, P_{fr}, absorbs maximally in the far-red (730 nm). The latter is considered to be the biologically active form. The diagnostic test for phytochrome action generally involves its ability to induce a particular developmental or physiological response after exposure to red light, followed by the cancellation of the induction when the process is repeated with far-red light within a given short interval. Such reversibility of the phototrans-formation cannot be shown for every response involving phytochrome, however. *See* PLANT PIGMENT; PHYTOCHROME.

Phytochrome has been extracted and characterized from dark-grown (etiolated) seedlings of several plant species, the bulk of the early work being with oat and rye. The protein has a monomer molecular weight of 120,000, and the most common form found in solution is a dimer. Phototransformation to P_{fr} increases the susceptibility of the chromophore to bleaching. It also increases the accessibility of

the functional groups of various amino acids to externally supplied reagents. All plants contain multiple phytochrome subtypes, each encoded by a different gene. In the model plant *Arabidopsis thaliana*, there are five distinct phytochrome subtypes, designated phytochrome A–E. As will be discussed shortly, studies of mutants of *Arabidopsis*, as well as those of tomato and pea, have demonstrated the clear association of particular phytochrome species with particular physiological responses, sometimes uniquely and sometimes redundantly.

In etiolated seedlings, phytochrome behaves as a soluble cytoplasmic protein, independent of the subtype being examined, though a small fraction appears to be membrane-bound. Exposure of seedlings to light (red light being generally most effective) results in a translocation of the phytochrome pigment to the nucleus. Following this, the relative abundance of the pigment in the cytoplasm versus its abundance in the nucleus depends upon several factors. These include the particular subtype of phytochrome being examined; the quality, intensity and duration of the incident light; and the influences of tissue- or organ-specific components of the process. Formation of P_{fr} is thought to expose an active site that is recognized by a particular binding partner which is either unique to, or shared between, particular phytochrome subtypes. While differential gene activation appears to be involved in responses occurring in both short- and long-term irradiation conditions, many rapid responses occurring within seconds or minutes are likely caused by membrane changes.

The conversion of P_r to P_{fr}, and vice versa, does not take place in a single step, but through complex sets of intermediates. Although the precise contributions of the protein component and chromophore in phytochrome photoconversion have not been definitively established, recent studies of P_r molecular structure (via x-ray crystallography) suggest that the conversion of P_r to P_{fr} involves a light-driven rotation of one of the four aromatic rings of the tetrapyrrole. Proton transfers between the protein and chromophore follow, which result in conformational changes within the protein to stabilize the P_{fr} conformer. The reverse conversion—the generation of P_r from P_{fr}, whether through photoconversion or dark conversion—appears to occur via distinct sets of reactions. Establishment of the molecular details, however, awaits similar molecular structure studies of P_{fr} and the interconvertible forms of phytochrome.

Although the molecular identification of blue-light receptors lagged behind that of phytochrome by some 40 years, the understanding of molecular details by which cryptochrome and phototropin absorb light and transduce their signals is now considerable. In fact, the molecular structures of both cryptochrome and phototropin were established before that of phytochrome, and currently there is a far deeper understanding of the photocycle of phototrophin at the structural level than that of either phytochrome or cryptochrome.

Due to the low abundance of cryptochromes and phototropins in cells, essentially all studies of the purified pigments have been performed with recombinant proteins isolated from nonplant hosts (such as bacteria, yeast, and insect cells). Results from such studies have shown that while the cryptochrome (a yellow pigment protein with ~70,000 monomer molecular weight) and phototropin proteins (another yellow pigment protein with ~120,000 monomer molecular weight) are distinct and apparently evolutionarily unrelated, the two photoactive pigment proteins share several features. These shared features are also quite different from the phytochromes' protein-chromophore features. *See* RECOMBINATION (GENETICS).

First, whereas phytochrome uses a bilin as its chromophore, both cryptochrome and phototropin use a flavin as their primary chromophore. Flavins also have a yellow pigment but are coenzymes of the flavoproteins. [Cryptochrome uses flavin adenine dinucleotide (FAD) and phototropin uses flavin mononucleotide (FMN).] Phytochrome also binds a single chromophore molecule, whereas both cryptochrome and phototropin bind two chromophores. (The chromophores of phototropins are both FMN molecules, while cryptochromes have a chomophore called pterin (methenyltetrahydrofolate) as a second chromophore. Finally, a covalent bond joins phytochrome's chromophore and protein components, whereas the same attachment in the cryptochromes and phototropins is a non-covalent bond, at least in their nonphotoactivated states. It is clear, however, that despite their similarities in chromosome usage, the cryptochromes and phototropins have quite distinct evolutionary origins.

Cryptochrome is evolutionarily related to DNA photolyases, enzymes that catalyze the light-dependent repair of UV-damaged DNA. While the cryptochromes do not appear to catalyze DNA repair as photolyase does, their structure and chromophore usage is similar to that of photolyase, leading to the proposal that cryptochrome and photolyase share a conserved photoactivation mechanism involving redox changes. The FAD in cryptochromes appears to be reduced (deoxidized) by electrons coming from the protein backbone itself. However, there is currently no evidence that electrons ever leave the cryptochrome pigment as occurs with photolyase. *See* ENZYME.

Not surprisingly, given their divergent gene evolution and differing pigment usage, the phototropins undergo a much different photoactivation process from that observed with cryptochromes and photolyases. Photoexcitation of the noncovalently bound FMN chromophore of phototropin by blue light results in abstraction (or removal) of a proton from a conserved cysteine amino acid within the flavin-binding pocket of the protein, which in turn promotes the formation of a covalent linkage between the FMN and cysteine residue. This covalent linkage is completely reversible, either by incubation in darkness or exposure to a second photon

from the near-UV portion of the spectrum. This process, or photocycle, has recently been shown to represent a conserved photoactivation mechanism in a number of proteins from a variety of organisms. In each case, these proteins use an FMN chromophore and a flavin-binding region, the LOV (light, oxygen, and voltage) domain, similar to that of phototropin.

As has occurred with the phytochromes, more than one cryptochrome evolved among the higher plants; at least three cryptochrome subtypes exist in all plants thus examined. Two of the subtypes (cryptochromes 1 and 2) are clearly distinguishable both physically and functionally from each other, as well as from the third cryptochrome (cryptochrome 3), and from photolyase by the presence of carboxyl-terminal amino acid sequence extensions. Mutational studies in *Arabidopsis* have shown that these carboxyl-terminal sequences, also called the CCTs (cryptochrome carboxyl-terminus), are necessary for the function of cryptochromes 1 and 2, with the variance in CCT amino acid sequence between these subtypes defining in part the output pathways of each, presumably by interacting with different sets of transducing proteins.

Differences in photosensitivity and protein stabilities of cryptochromes 1 and 2 also contribute to their differing physiological functions. At present it is not known what processes cryptochrome 3 regulates in plants, although the fact that in the cell it is located within mitochondria and chloroplasts suggests that it regulates processes in those organelles. Cryptochromes 1 and 2 are nuclear residents in etiolated seedlings. Even when plants are exposed to light, the location of cryptochrome 2 does not change for the most part, although this cryptochrome undergoes rapid degradation after exposure to blue light. Cryptochrome 1, on the other hand, appears to primarily partition to, or locate itself in, the cytoplasm in light-grown plants. It also appears to remain relatively stable under all light conditions. Thus, none of the cryptochrome subtypes exhibit identical partitioning, and these patterns of partitioning differ from those observed with phytochrome.

All plant species examined to date contain at least two phototropin subtypes: phototropins 1 and 2. The phototropin subtypes are distinguishable from each other, not by differences in amino acid sequences per se, but more by differences in patterns of expression of the two subtypes. In particular, phototropin 1 appears to be fairly ubiquitously expressed in *Arabidopsis* plants with limited regulation over long periods of irradiation being known; phototropin 2, in contrast, is not expressed in dark-grown seedlings or is expressed only weakly, but expression is rapidly and strongly induced upon exposure to blue light. Thus, in etiolated seedlings phototropin 1 is the primary phototropin present, while in light-grown seedlings phototropin 1 and 2 are both present. This observation helps to explain why phototropin 1 represents the predominant receptor controlling responses in etiolated seedlings, while both phototropins 1 and 2 appear to redundantly mediate a common set of responses in light-grown plants.

Additionally, the two phototropin subtypes differ in their sensitivity to light intensity. Phototropin 1 can function over a very broad range of light intensities, whereas phototropin 2 appears to function exclusively in high light conditions. Unlike both the phytochromes and cryptochromes, the phototropins are predominantly localized to the plasma membrane, although some portion of phototropin moves to the cytoplasm in response to blue light, probably within membrane vesicles. It is not currently known how this light-dependent intracellular relocation of phototropin influences signaling. Relative to signal output from activated phototropins, both subtypes have a protein kinase domain located at the end of the carboxyl group of their photosensory region; activity of this protein kinase domain appears essential for phototropin functions. Moreover, the kinase domain, independent of whether it is biochemically functional or not, has been shown to be necessary for plasma membrane localization of phototropin in etiolated seedlings.

Physiological roles of pigments. Since phytochromes, cryptochromes, and phototropins play regulatory roles in a wide variety of physiological and developmental processes, considerable effort has been taken to elucidate the mechanisms of the various responses. It is usually assumed (although far from proven) that the primary action of the activated pigments (P_{fr}, photoreduced cryptochrome, and phototropin that contain a covalently linked flavin) is the same for a given individual pigment type across various responses. Subsequent reaction chains differ widely for a given pigment depending upon the system, the stage of development, or prior environmental conditions.

Seedling development. In order to survive and reproduce, most plants need to produce a light-capturing apparatus equipped to carry on photosynthesis efficiently; this apparatus is the leaf (or frond in the case of ferns) with its chloroplasts. Higher plants have evolved mechanisms to ensure that once a seed has germinated and growth has commenced, little stored energy will be expended on leaf development until the potentially green leaf encounters light conditions adequate for photosynthesis. Rather, the energy is expended in elongation processes that maximize chances that the undeveloped leaves will reach such light conditions. Each of the pigments introduced above represents key components of this system for regulating seedling growth and development.

Some seeds, frequently those with only small reserves of food, require light (red being most effective) for germination. The requirement for an adequate concentration of P_{fr} before germination progresses ensures that a seed that is buried too deeply in soil or is too heavily shaded by other plants will not germinate and, thus, exhaust its stored resources before light is sufficient for photosynthesis.

Dicotyledenous seedlings generally have broad expanded leaves, except when grown in darkness.

Without generation of P_{fr} and photoreduced cryptochrome by light treatment, the leaves are small, devoid of chlorophyll, and usually folded up and tucked beneath a protective hook of the stem which helps prevent leaf damage as the stem tip pushes up through the soil or litter. With the first light exposure, the hook straightens out to orient the leaves favorably for photosynthesis, and they then begin rapid expansion. Rapid elongation of the cells on the inside of the hook causes it to straighten out. The seedling continuously synthesizes ethylene, a gaseous plant hormone. As it escape from the etiolated hook, it regulates elongation of cells to maintain the hook, in combination with another plant hormone, auxin. Photoactivation of phytochrome and cryptochrome alter the synthesis of ethylene and the response to auxin, releasing cells on the inner side of the hook from the restraint on growth, while suppressing the elongation of cells on the opposing side, thus resulting in hook opening. In many dicotyledonous plants, this coincides with the beginning of a dramatic decline in the rate of stem elongation. Study of mutants has shown that this growth inhibition is mediated not by phytochrome but by phototropin 1 and cryptochrome 1. Phototropin 1 mediates this growth inhibition within minutes, and cryptochrome 1 does it in about an hour. Persistent stem growth inhibition is modulated by both cryptochrome and phytochrome, whereas phototropin appears to play no role in the latter response.

Etiolated dicotyledonous leaves are normally a small fraction of the size they attain in light. Not only is enlargement triggered by light, but continued exposure to light is required to maintain growth to normal size. Among monocotyledonous leaves, those of grasses remain tightly rolled in the dark and unfurl upon exposure to light. The tightly rolled leaves can be induced to unroll either by P_{fr} formation or by application of the plant hormone gibberellin. Several workers have shown that isolated etioplasts (precursors to chloroplasts) from etiolated grass leaves will release biologically active gibberellin within a few minutes upon exposure to red light, with the release prevented by subsequent far-red. In this case, many of the steps between phytochrome phototransformation and biological response are relatively well known.

To ensure upward growth, seedling stems are especially sensitive in their geotropic curvature response to the stimulus of gravity. To facilitate growth toward a source of light, they are also responsive phototropically. The phototropic response is mediated by light absorbed by the phototropins, but the responsivity to blue light can be dramatically altered through P_{fr} formation. Thus, a delicate balance is achieved between the response designed to raise a plant vertically above its neighbors and one designed to direct it toward light from some particular direction. These balanced, often competing responses represent the coordinated interactions between all three receptor types and at least four plant hormones: ethylene, auxin, gibberellin, and brasinosteroid. *See* AUXIN; BRASINOSTEROID; ETHYLENE; GIBBERELLIN.

Many roots if allowed to develop in darkness are insensitive to gravity. Such roots are free to follow moisture and nutrient gradients as they penetrate the soil. Formation of a minute amount of P_{fr} induces geotropic sensitivity such that the root rapidly turns downward. This response probably protects a root against breaking through the soil and into the air, where it could become desiccated and die. Root growth in general is controlled in large part by light. For example, root elongation and development is quite restricted in etiolated seedlings of dicotyledonous plants, leading to a short unbranched primary root. Subsequent light exposure stimulates root elongation and alters developmental programs to give rise to the highly branched root system typically associated with plants. The phytochromes and cryptochromes both play a role in regulating these latter responses. As with so many other photomorphogenetic responses, these responses represent an energy conservation mechanism by which the plant expends energy only on root growth once the aerial portion of the plant is established and can sustain growth via photosynthesis.

In line with the aforementioned adaptive mechanisms, etiolated plants make only a minimal investment of food reserves for the formation of chloroplasts in their leaves. The chloroplast precursors (etioplasts) have no chlorophyll in angiosperm species; in virtually all plants the etioplasts possess only a fraction of the necessary internal membrane system; and these immature organelles lack the full complement of enzymes needed to fix carbon dioxide into sugar. Thus, etiolplasts are incapable of performing photosynthesis. Light absorbed by protochlorophyll (a chlorophyll precursor) and later by chlorophyll itself initiates development and sustains it to produce mature, functioning chloroplasts. The light absorbed by another chlorophyll precursor, protochlorophyllide, converts it directly to chlorophyll, the principal photosynthetic pigment. The rate of subsequent chloroplast development can be markedly affected by light pretreatments that form small amounts of P_{fr}. In fact, P_{fr} formation may advance chlorophyll formation and chloroplast development by several hours. Both P_{fr} and photoreduced cryptochrome influence the expression of nuclear genes encoding proteins located in the chloroplast that are necessary for its development and function. Mutations in either class of receptor, or signal components associated with these receptors, can dramatically alter chloroplast development. This includes completely blocking the transition from etioplast to chloroplast. In addition to chloroplast autonomous and receptor-driven nuclear events, there is now considerable evidence for complex communication between the developing chloroplast and the nucleus in the regulation of these processes. *See* CHLOROPLAST; PLANT CELL; PLANT METABOLISM.

Continuing vegetative growth. Light-grown plants are known to contain phytochrome and cryptochrome

2; thus the destruction (decay) of the more labile pigments formed in the light does not result in complete disappearance of the pigments. Both can be detected spectrophotometrically and immunologically in extracts from green leaf tissue, and their presence is shown by innumerable experiments on light-grown plants, although cryptochrome 1 is certainly contributing to blue light responses in light-grown plants as well. As an example of persistent phytochrome presence, consider the growth responses of plants grown on day/night cycles in white light. After several days (or potentially weeks) of regular day/night cycles, plants that are treated with far-red light at the beginning of a dark period (to convert phytochrome to P_r) will produce longer stems and petioles than both their untreated and red-light treated siblings. Thus the presence of P_{fr} during the dark period inhibits elongation in some manner. A large portion of this inhibition can be reversed with the growth-regulating hormone gibberellin. Many other plants, including woody species such as pines and various broadleaf trees, respond similarly. *See* CIRCADIAN CLOCK (PLANTS).

Light that has passed through green leaves is much richer in far-red light than unobstructed sunlight is. This fact provides a natural counterpart to the experiments above. Rapid elongation of plants growing in the shade of other plants has a clear survival advantage in aiding them to reach photosynthetically adequate light intensities. Considerable research continues to be devoted to the study of such responses (called the shade-avoidance syndrome) since they have obvious agricultural and horticultural implications.

Response to daylength (photoperiodism). In regions with distinct seasons of alternating favorable and unfavorable growing conditions, the precise change in daylength is a more dependable warning of an oncoming seasonal change than are the vagaries of temperature or rainfall. Plants have evolved mechanisms of responding to the shortening days of late summer and fall (short-day plants) or the lengthening days of the spring (long-day plants). *See* PHOTOPERIODISM.

The most obvious physiological response to a change in daylength is flower initiation. Flowering is timed to occur soon enough for seed production to be complete prior to the onset of unfavorable conditions. It is also delayed enough to allow adequate vegetative growth to support the formation of fruit and seeds. The protection of growth centers, the meristems of the shoot, is often ensured by photoperiodic induction of dormant buds in trees and shrubs. For example, shortening days halt stem elongation and induce the formation of bud scales and dormant buds. The buds are reactivated only after progression through the cold conditions of winter and exposure to the warmth of spring. *See* ABSCISIC ACID; DORMANCY.

Current evidence suggests that the central players in photoperiodism are conserved in all plants, and that slight modifications of these components results in short- and long-day dependencies. For example, biochemical processes that occur in leaves during the dark period lead to the formation of one or more flower-inducing compounds that move to the meristems of shoot tips. This factor changes the course of development from vegetative to reproductive-growth—the formation of flowers. Possibly a reduction in a flowering inhibitor is also required. Take as an example an average short-day plant: if P_{fr} is formed in the leaves by a brief flash of light during the dark period, formation of the floral stimulus is interrupted. If neither of the two parts of the interrupted night is as long as the critical dark period, no flowering ensues. If far-red light treatments immediately follow the interrupting light, thus converting the P_{fr} back to P_r, the short-lived existence of P_{fr} may be insufficient to interrupt synthesis of the flowering stimulus. P_{fr} present for 30 min is sufficient to inhibit flowering in the cocklebur (*Xanthium*); its presence for as little as 30 s is sufficient for the Japanese morning glory (*Pharbitis nil*).

Many plants are capable of differentiating between dark periods differing by only 30 min, and one variety of rice responds very differently to nights varying by only 10 min in length. The nature of this precise measurement of night length is not clear at present. However, strong evidence implicates the diurnal "biological clock" involved in circadian rhythms, in combination with phytochromes and cryptochromes.

Long-day plants, in contrast to short-day plants, require a dark period shorter than some critical length in order to flower, and may flower under continuous light. Within a normal 24-h daily light/dark cycle, an uninterrupted dark period greater than the critical length will inhibit flower induction. If such an inhibitory dark period is interrupted in these plants by P_{fr} formation, the inhibitory effect of the dark period is prevented and the plants flower. Prompt removal of the P_{fr} by far-red light reverses the effect, and the inhibitory state is maintained.

Phytochromes and cryptochromes are both involved in photoperiodic flowering, acting to entrain the biological clock that informs the plant about the time of day, as well as to modify clock outputs in response to fluctuations in light intensity and quality. A phototropin-related protein, FLAVIN-BINDING KELCH REPEAT F-BOX 1 (FKF1), has also been shown to influence the biological clock in the presence of (blue) light to provide additional control over photoperiodism. The biological clock appears to regulate the transcription of the gene *CONSTANS (CO)*, with *CO* expression being highest at dawn and dusk in long-day plants, and highest in the subjective night in short-day plants. Receptor setting of the clock to influence timing of *CO* expression is regulated primarily by phytochromes A and B, cryptochrome 1, and FKF1. *CO* encodes an apparent transcription factor that directly regulates expression of a number of genes, such as *FLOWERING LOCUS T (FT)*.

Beyond clock influences on *CO* gene expression, the levels of CO protein in the nucleus are

regulated by direct light-dependent influences on both cytoplasmic-nuclear partitioning and proteosome-dependent degradation of the CO protein. In *Arabidopsis*, phytochrome A and cryptochrome 2 appear to stabilize nuclear CO protein in response to far-red and blue light, respectively, while phytochrome B appears to promote the degradation of CO in red light. Together these regulatory properties allow *Arabidopsis* to exhibit a long-day flowering phenotype. Results from studies of various mutants provide considerable support for such a model. For example, mutants lacking phytochrome A or cryptochrome 2 have less CO protein and exhibit quantitatively late flowering phenotypes (that is, it takes more long-days to induce flowering), while mutants lacking phytochrome B have more CO protein and flower early (after exposure to fewer long-days). In harmony with these observations, mutants lacking CO have been shown to be late-flowering, while plants overexpressing CO are early-flowering.

It is interesting that the encoded protein for *FT* messenger RNA causes a shift in the plant developmental program from vegetative to reproductive when it moves from the sites of CO-dependent transcription in the leaves to meristems. These properties suggest that FT represents one of the long-sought floral-inducing compounds. Consistent with this notion is the high correlation between the presence of FT in a plant's meristems and the amount of active CO in the plant—the presence of FT in a meristem thus conferring upon it the ability to switch from vegetative to reproductive development. *See* FLORIGEN; PHOTOPERIODISM; PLANT GROWTH; RIBONUCLEIC ACID (RNA).

Emmanuel Linscum; Winslow R. Briggs

Bibliography. L. C. Sage, *Pigment of the Imagination*, Academic Press, San Diego, 1992; E. Schäfer and F. Nagy (eds.), *Photomorphogenesis in Plants and Bacteria: Function and Signal Transduction Mechanisms*, 3d ed., Springer, Dordrecht, 2006.

Photomultiplier

A very sensitive vacuum-tube detector of light or radiant flux containing a photocathode which converts the light to photoelectrons; one or more secondary-electron-emitting electrodes or dynodes which amplify the number of photoelectrons; and an output electrode or anode which collects the secondary electrons and provides the electrical output signal. It is less frequently known as a multiplier phototube or photomutiplier tube. Because of the very large amplification provided by the secondary-emission mechanism, and the very short time variation associated with the passage of the electrons within the device, the photomultiplier is applied to the detection and measurement of very low light levels, especially if very high speed of response is required.

The first photomultipliers were developed during the late 1930s; applications were generally in astronomy and spectrometry. Development of photomultipliers was stimulated beginning in the late 1940s

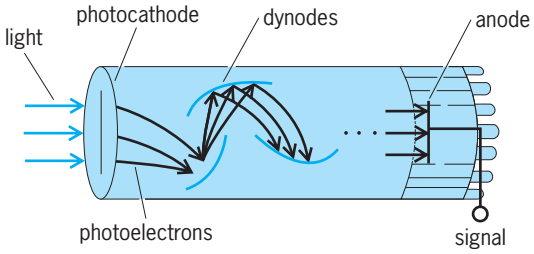

Fig. 1. Schematic of a photomultiplier. (*After R. W. Engstrom, Photomultipliers—then and now, RCA Eng., 24(1):18–26, June–July 1978*)

by their application to scintillation counting. Variations in electron-optical design and the introduction of a variety of photocathode, dynode, and envelope materials led to the use of photomultipliers in a wide variety of applications. Active design programs continue despite the competition from various solid-state detectors. Photomultiplier tubes are available in sizes ranging from $^1/_2$ to 5 in. (13 to 127 mm) in diameter, with photocathodes useful for detecting radiation having wavelengths from 110 nanometers, the deep ultraviolet (using a lithium fluoride window), to 1100 nm, the infrared (with special photocathodes), and with amplification factors ranging from 10^3 to 10^9.

Operation and design. The schematic of a typical photomultiplier in **Fig. 1** illustrates its operation. Light incident on a semitransparent photocathode located inside an evacuated envelope causes photoelectron emission from the opposite side of the photocathode. (Some photomultipliers are designed with the photocathode mounted inside the vacuum envelope to provide for photoemission from the same side of the photocathode on which the light is incident.) The efficiency of the photoemission process is called the quantum efficiency, and is the ratio of emitted photoelectrons to incident photons (light particles). Photoelectrons are directed by an accelerating electric field to the first dynode, where from 3 to 30 secondary electrons are emitted from each incident electron, depending upon the dynode material and the applied voltage. These secondaries are directed to the second dynode, where the process is repeated and so on until the multiplied electrons from the last dynode are collected by the anode.

A typical photomultiplier may have 10 stages of secondary emission and may be operated with an overall applied voltage of 2000 V. In most photomultipliers the focusing of the electron streams is done by electrostatic fields shaped by the design of the electrodes. Some special photomultipliers designed for very high speed utilize crossed electrostatic and magnetic fields which direct the electrons in approximate cycloidal paths between electrodes. Because the transit times between stages are nearly the same, anode pulse rise times as short as 150 picoseconds are achieved.

Another special photomultiplier design is based on the use of microchannel plates. A single-channel

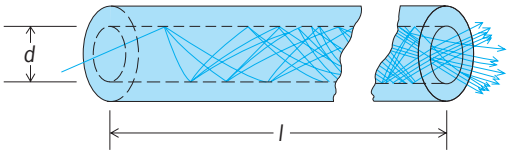

Fig. 2. Continuous-channel multiplier structure. (*After RCA Photomultiplier Manual, Tech. Ser. PT-61, 1970*)

multiplier is schematically shown in **Fig. 2**. The channel is coated on the inside with a resistive secondary-emitting layer. Gain is achieved by multiple electron impacts on the inner surface as the electrons are directed down the channel by an applied voltage over the length of the channel. The gain of the channel multiplier depends upon the ratio of its length l to its diameter d; a typical ratio is 50:1. A microchannel plate is formed by combining a large number of the channels in parallel with spacings on the order of 40 micrometers. A very high-speed photomultiplier utilizes a microchannel plate mounted in a closely spaced parallel arrangement between a photocathode and an anode. Time resolutions of less than 100 ps are achieved.

Dynode materials. Common dynode materials used in photomultiplier tubes are cesium antimonide (Cs_3Sb), magnesium oxide, and beryllium oxide. The surface of the metal oxide dynodes is modified by activation with an alkali metal such as cesium which lowers the surface potential barrier, permitting a more efficient escape of the secondary electrons into the vacuum.

Materials characterized as negative-electron affinity (NEA) have been developed and utilized in photomultipliers for both dynodes and photocathodes. A typical NEA secondary-emission material is gallium phosphide (GaP) whose surface has been treated with cesium (GaP:Cs). **Figure 3** shows the energy band model of an NEA material. The surface barrier is reduced by the electropositive cesium layer, and band-bending occurs so that the conduction band may actually lie above the vacuum level as indicated.

When a primary electron impacts an NEA material, secondary electrons are created within the material, and even low-energy electrons escape to the vacuum. The result is a much higher secondary-emission ratio than that achieved in other secondary emitters, especially at higher primary energies when the primary electrons penetrate more deeply into the material. Secondary electrons within an emitter lose energy before reaching the surface, so that in ordinary materials the secondary emission actually decreases with increase in voltage above an energy of perhaps 500 eV. For example. Cs_3Sb reaches a maximum secondary emission of about 8:1 at 500 V. At the same primary voltage GaP:Cs has a secondary emission of 25:1, and secondary emission increases linearly with voltage to a ratio of at least 130:1. *See* SECONDARY EMISSION; SEMICONDUCTOR.

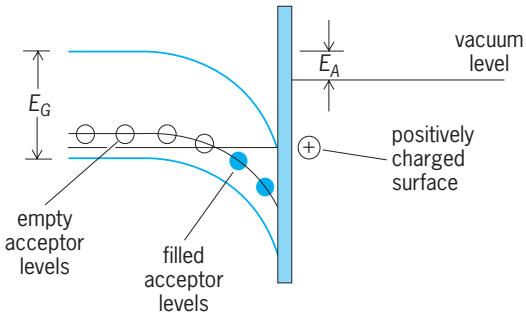

Fig. 3. Semiconductor energy-band model showing negative electron affinity, E_A. Band-gap energy = E_G. (*After RCA Photomultiplier Manual, Tech. Ser. PT-61, 1970*)

Photocathode materials. The first photomultipliers developed utilized a photocathode of silver oxide activated with cesium (Ag-O-Cs). Although the quantum efficiency of this photocathode is quite low (less than 1% at the wavelength of its maximum response), it is still used because of its near-infrared response out to 1.1 μm.

Most photomultipliers today have photocathodes classed as alkali-antimonides: Cs_3Sb, K_2CsSb, Na_2KSb, Na_2KSb:Cs, Rb-Cs-Sb. These photocathodes have good quantum efficiency in the visible-wavelength range—in some cases exceeding 20%. The NEA type of photocathodes such as GaAs:Cs and InGaAs:Cs have even higher quantum efficiency and response through the visible and into the near infrared. However, they are difficult to fabricate and are more readily damaged by excessive photocurrents. *See* PHOTOTUBE.

Detection limits. Because there is very little noise associated with the amplification process in a photomultiplier, detection limits are primarily determined by the quantum efficiency of the photocathode, by the thermionic or dark emission of electrons from the photocathode, and by the bandwidth of the observation. A typical photomultiplier limited by dark emission is capable of detecting 10^{-12} lumen—the equivalent of the flux on 1 cm^2 from a l-candlepower lamp at a distance of 6 mi (10 km). By selecting a photomultiplier having a photocathode with low dark emission, restricting the effective photocathode area magnetically, and cooling the tube to reduce dark emission further, a background count as low as one or two electrons emitted from the photocathode per second can be achieved. Against this background, with a photocathode having 20% quantum efficiency, the incidence of a flux of only 10 photons per second can be detected in an observation time of perhaps 10 s. In this technique, referred to as photon counting, each electron originating at the photocathode results in an output pulse which is counted. An evaluation must be made by statistically comparing the cases of source-on and source-off. Such techniques were applied successfully in 1969 in the laser-ranging experiments by using a retroreflector placed on the Moon by the *Apollo 11* astronauts. *See* THERMIONIC EMISSION.

Applications. Ever since their invention, photomultipliers have been found useful in low-level photometry and spectrometry. The many applications

of photomulipliers include high-speed inspection of small objects such as fruits, seeds, toys, and other industrial products; pollution monitoring; laser ranging; and process control with transmitted or reflected light to detect flaws in various solid, liquid, or gaseous manufacturing operations. *See* LASER; NONDESTRUCTIVE EVALUATION; PHOTOMETRY; SPECTROSCOPY.

Scintillation counting. The most important applications of photomultipliers are related to scintillation counting. In a scintillation counter, gamma rays produce light flashes in a material such as NaI:Tl. A photomultiplier is optically coupled to the scintillator and provides a count of the flashes and a measure of their magnitude. The magnitude of the scintillation flash is proportional to the energy of the gamma ray, thus enabling the identification of particular isotopes. A whole science of tracer chemistry has developed using this technique and is applied to agriculture, medicine, and industrial problems. *See* GAMMA-RAY DETECTORS; SCINTILLATION COUNTER.

Gamma-ray camera. The gamma-ray camera developed in the late 1950s has proved to be a very valuable medical tool. In this application, a radioactive isotope combined in a suitable compound is injected into or fed to a patient. Certain compounds and elements concentrate preferentially in particular organs, glands, or tumors: for example ^{131}I concentrates in the thyroid gland. Gamma rays which are then emitted are detected by a large scintillator, and an array of photomulitiplier tubes provides spatial data on the gamma-ray source. The technique is widely used in diagnostic medicine, especially in locating tumors. *See* NUCLEAR MEDICINE; RADIOLOGY.

Computerized tomography. The CT or CAT (computerized axial tomographic) scanner is a similar medical instrument developed during the 1970s. This instrument uses a pencil or fan-beam of x-rays that rotates around the patient. Several hundred photomultiplier-scintillator combinations surround the patient and record density from many different positions. These data are analyzed by a computer, thus providing a cross-section density map of the patient. The CT scanner is particularly useful because of the two-dimensional data provided, although it does not supply the functional information of the gamma-ray camera. Since 1980, silicon photocells have been replacing photomultipliers as the detectors of the scintillations. *See* COMPUTERIZED TOMOGRAPHY.

Positron camera. A similar development is the positron camera. This device uses radioisotopes, such as ^{11}C, ^{13}N, and ^{15}O, that have short half-lives and emit positrons. When a disintegration occurs, the positron is annihilated by an electron, and the result is a pair of oppositely directed gamma rays detected in coincidence with photomultiplier-scintillator detectors on opposite sides of the patient. Functional information is provided and analyzed in a manner similar to that of a CT scanner. Ralph W. Engstrom

Bibliography. E. L. Dereniak and D. G. Crowe, *Optical Radiation Detectors*, 1984; J. Schanda and I. Ungvari (eds.), *Photon Detectors*, 1987.

Photon

An entity that can be loosely described as a quantum of energy of electromagnetic radiation. According to classical electromagnetic theory, an electromagnetic wave can transfer arbitrarily small amounts of energy to matter. According to the quantum theory of radiation, however, the energy is transferred in discrete amounts. The energy of a photon is the product of Planck's constant and the frequency of the electromagnetic field. In addition to energy, the photon possesses momentum and also possesses angular momentum corresponding to a spin of unity. The interaction of radiation with matter involves the absorption, scattering, and emission of photons. Consequently, the energy interchange is inherently quantized. *See* ANGULAR MOMENTUM; ENERGY; MOMENTUM; SPIN (QUANTUM MECHANICS).

Particlelike behavior. For many purposes, the photon behaves like a particle of zero rest mass moving at the speed of light. Unlike spin-1 particles of finite mass, the photon has only two, rather than three, polarization states. This corresponds to the transverse nature of a classical electromagnetic wave in free space. The particlelike nature of the photon is vividly exhibited by the photoelectric effect, predicted by A. Einstein, in which light is absorbed in a metal, causing electrons to be ejected. An electron absorbs a photon, gaining its energy. In leaving the metal, it loses energy because of interactions with the surface; the energy loss equals the product of the so-called work function of the surface and the charge of the electron. The final kinetic energy of the electron therefore equals the energy of the incident photon minus this energy loss. *See* PHOTOEMISSION.

A second demonstration of the particlelike behavior of photons is provided by the scattering of an x-ray photon from an electron bound in an atom. The electron recoils because of the momentum of the photon, thereby gaining energy. As a result, the frequency, and hence the wavelength of the scattered x-ray, is altered. If the x-ray is scattered through a certain angle, the wavelength is shifted by an amount determined by this scattering angle and the mass of an electron, according to the laws of conservation of energy and momentum. *See* COMPTON EFFECT.

Quantum theory. From a more fundamental view, the photon is the quantum of excitation of a single mode of a radiation field. The dynamical equations for the electric and magnetic energy in such a field are identical to those of a harmonic oscillator. According to quantum theory, the allowed energies of a harmonic oscillator are given by $E = (j + {}^1/_2)hf$, where h is Planck's constant, f is the frequency of the oscillator, and the quantum number $j = 0, 1, 2, \ldots$, describes the state of excitation of the oscillator. This quantum relation was first postulated by M. Planck for the material oscillators in the walls of a thermal enclosure in order to obtain the correct form for the density of radiation in a thermal field, but it was quickly applied by Einstein to describe the state of the radiation field itself. In this picture,

j describes the number of photons in the field. *See* HARMONIC OSCILLATOR.

Photon distributions. Different sources of light are characterized by photon distributions with fundamentally different statistical properties. If light from a monochromatic source at high intensity, for instance a laser source, impinges on a photomultiplier, then each signal count corresponds to the absorption of a single photon. If repeated measurements are made, then the fractional root-mean-square (rms) scatter of the readings is exactly as would be expected for independent particles. In the limit of large numbers of absorbed photons, the fluctuations are unimportant, and the photon distribution describes a field with well-defined amplitude and phase, a classical field.

In contrast, for thermal radiation the fractional fluctuation is much larger than for independent particles. The larger fluctuation expresses the fact that photons, having integral spin, obey Bose-Einstein statistics. Such particles tend to clump together, producing large fluctuations. Finally, a radiation field can be considered in which the photon number is precisely determined, called a Fock state. In such a case, the phase of the corresponding electromagnetic field is completely random. Such a state has no classical analog. *See* BOSE-EINSTEIN STATISTICS; SQUEEZED QUANTUM STATES.

Need for quantized fields. Many of the processes by which radiation interacts with matter can actually be described without employing a quantized description of the radiation field or the concept of photons, that is, by treating the radiation field classically. However, the fundamental process of spontaneous emission, in which an excited atom or molecule emits a photon and makes a transition to its ground state, can be described only on the basis of a quantized field. Also, many statistical properties of the radiation field, for instance correlations between photons that are emitted simultaneously in a nonlinear process, are inexplicable in terms of classical fields but have a natural explanation in the quantum theory of radiation and the concept of photons. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; QUANTUM MECHANICS.

Daniel Kleppner

Bibliography. C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Atom-Photon Interactions*, 1992, reprint 1998; R. Loudon, *The Quantum Theory of Light,*, 3d ed., 2000; P. Meystre and M. Sargent III, *Elements of Quantum Optics*, 3d ed., 1999; P. W. Milonni, *The Quantum Vacuum*, 1997.

Photoperiodism

The growth, development, or other responses of organisms to the length of night or day or both. Photoperiodism has been observed in plants and animals, but not in bacteria (prokaryotic organisms), other single-celled organisms, or fungi.

Parameters. A true photoperiodism response is a response to the changing duration of day or night. Responses to the quantity of light energy in photoperiodism are secondary complications. Some species respond to increasing day lengths and decreasing night lengths (for example, by forming flowers or developing larger gonads); this is called a long-day response. Other species may exhibit the same response, or the same species may respond in some different way, to decreasing days and increasing nights; this is a short-day response. Sometimes a response is independent or nearly independent of day length and is said to be day-neutral.

Plant responses. **Figure 1** illustrates these phenomena as they apply to the initiation of flowers in plants, perhaps the most extensively studied photoperiodism response. Flowering can be measured in various ways, such as counting the number of flowers on each plant, classifying the size of the buds according to a series of arbitrary stages, or taking the inverse of the number of days until the first flower appears. A truly day-neutral plant flowers about the same amount on all day lengths (curve 1). (With many species, there is little or no flowering when days are unusually short, for example, from 2 to 6 h long.) Curve 2 represents a plant that is slightly but probably insignificantly promoted in its flowering by long days. Plants represented by curves 3 and 4 are quantitatively promoted by long days (to different degrees), although they flower on any day length. A qualitative or absolute long-day plant (5) such as henbane flowers only when days are longer than some minimum, 12 h in the example. A qualitative short-day plant (6) such as cocklebur in this example flowers only when days are shorter than about 15.6 h and nights are longer than 8.3 h. Note that cocklebur also fails to flower if days are shorter than 3–5 h. (In that sense, it is also a long-day plant.) A quantitative short-day plant (7) flowers on any day length but

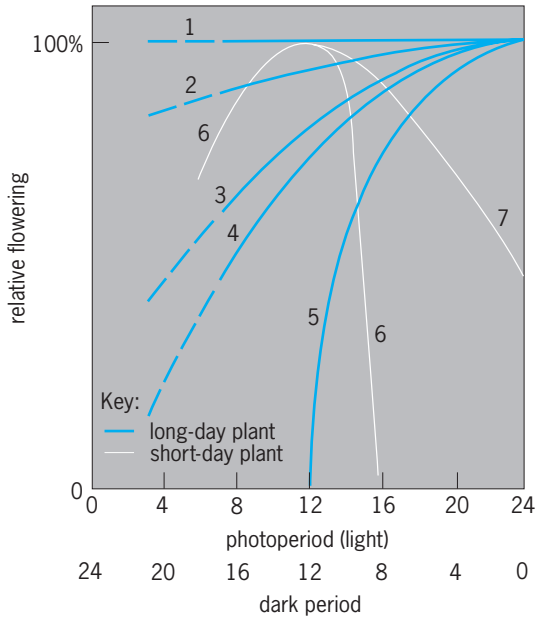

Fig. 1. Flowering (and other) responses to various day lengths. Each line represents a different hypothetical plant.

Selected examples of photoperiodism in plants and animals*

Organism	Short-day	Long-day	Response promoted by indicated day length
Algae			
Porphyra tenera (Rhodophyta)	X		Production of filamentous phase of spores that gives rise to leafy phase
Scytosiphon lomentaria (Phaeophyta)		X	Production of erect, cylindrical thalli from crustose phase
Bryophytes			
Sphagnum plumulosum	X(?)		Development of reproductive structure
Marchantia polymorpha		X	Development of reproductive structure
Gymnosperms			
Loblolly pine (*Pinus taeda*)		X	Stem growth
Angiosperms (flowering plants)			
Monocots			
Bluestem grass (*Andropogon gerardi*)	X		Flower formation
Wild oats (*Avena sativa*)		X	Flower formation
Onion (*Allium cepa*)		X	Bulb formation
Orchid (*Cattleya trianae*)	X		Flower formation and bud elongation
Yam (*Dioscorea alata*)	X		Tuber development
Dicots			
Soybean (*Glycine max*)	X		Flower formation
Sugarbeet (*Beta vulgaris*)		X	Flower formation
Radish (*Raphanus sativus*)	X		Root thickening
		X	Flower formation
Strawberry (*Fragaria chiloensis*)		X	Runner development
	X		Flower formation
Coneflower (*Rudbeckia* spp.)		X	Stem elongation and flowering
Cucumber (*Cucumis sativus*)	X		Female flowers increased
		X	Male flowers increased
Alfalfa (*Medicago sativa*)	X		Winter hardening
Potato (*Solanum tuberosum*)	X		Tuber development
Dahlia (*Dahlia* spp.)	X		Storage root development
		X	Fibrous root development
Bryophyllum (*Bryophyllum pinnatum*)		X	Development of foliar plantlets
	X		Flower formation
Kalanchoë (*Kalanchoë blossfeldiana*)	X		Dark fixation of CO_2
White mustard (*Sinapis alba*)	X		Flower development
Begonia (*Begonia evansiana*)		X	Formation of aerial tubers
Tulip poplar (*Liriodendron tulipifera*)	X		Cessation of vegetative growth, start of bud dormancy
Insects			
Mite (*Metatetranychus ulmi*)		X	Uninterrupted development
	X		Initiation of diapause
Mulberry silkworm (*Bombyx mori*)		X	Female lays diapause eggs
	X		Female lays developing eggs
Arthropods			
Water flea (*Daphnia pulex*)		X	Hatching of diapaused embryos
Green vetch aphid (*Megoura viciae*)		X	Production of parthenogenetic females instead of sexual females
Pink bollworm of cotton (*Pectinophora gossypiella*)	X		Diapause
Fish			
European minnow (*Phoxinus phoxinus*)		X	Egg maturation
Stickleback (*Gasterosteus aculeatus*)		X	Sexual maturation; male nest-building, spermatogenesis
Reptile			
Lizard (*Xantusia vigilis*)		X	Accelerated gonadal growth
Birds†			
Junco (*Junco hyemalis*)		X	Increase in body weight, marked deposition of fat, gonadal renewal, migratory behavior
Willow ptarmigan (*Logopus logopus*)	X		White winter plumage
Turkey (*Meleogris gallopavo*) and female chicken		X	Acceleration of sexual maturation and
(*Gallus gallus*)		X	stimulation of egg production
White-crowned sparrow (*Zonotrichia leucophrys gambelii*)		X	Testicle development
White-crowned sparrow (*Z. leucophrys gambelii*)	X		Elimination of refractory period so sexual cycle can recur
Mammals			
Sheep (*Ovis* spp.) and goat (*Capra* spp.)	X		Estrus; spermatogenic activity, hair growth
Snowshoe hare (*Lepus americanus*); also ermine		X	Spring molt, summer pelage, estrus, testicular
(*Mustela erminea cicognani*)		X	development
Snowshoe hare (*Lepus americanus*)	X		Autumn molt, winter pelage
Raccoon (*Procyon lotor*)		X	Estrus, male reproductive activity
Red fox (*Vulpes fulva*)	X		Autumn molt, winter pelage spermatogenesis, estrus
Golden hamster (*Mesocricetus aureus*)		X	Testicle development
Horse (*Equus equus*)		X	Estrus and associated hormonal changes in mare

*A given response is sometimes limited to a single genetic strain within a species.
† Rate of gonadal development is proportional to day length in many species of birds with no critical day length.

better under short days. Note that different species have different critical day and night lengths, not just the 12 and 15.6 h shown for henbane and cocklebur; yet as the curves show, the long-day and the short-day responses are qualitatively opposite to each other. But they can overlap so that both long-day and short-day plants flower on some day lengths (curves 5 and 6).

It is important to emphasize that a true photoperiodism response is to the *duration* of day or night and not to the total quantity of light received by the organisms. Typically, 8 h at a high light level followed by 16 h of darkness is perceived by the organism as a short day, whereas 8 h at the same high level, plus 8 h of extremely dim light (perhaps a thousandth as bright as the high level), followed by 8 h of darkness, is perceived as a long day.

There are many plant responses to photoperiod (see **table**). These include development of reproductive structures in lower plants (mosses) and in flowering plants; rate of flower and fruit development; stem elongation in many herbaceous species as well as coniferous and deciduous trees (usually a long-day response and possibly the most widespread photoperiodism response in higher plants); autumn leaf drop and formation of winter dormant buds (short days); development of frost hardiness (short days); formation of roots on cuttings; formation of many underground storage organs such as bulbs (onion, long days), tubers (potato, short days), and storage roots (radish, short days); runner development (strawberry, long days); flower formation (strawberry, short days); balance of male to female flowers or flower parts (especially in cucumbers); aging of leaves and other plant parts; and even such obscure responses as the formation of foliar plantlets (such as the minute plants formed on edges of *Bryophyllum* leaves), and the quality and quantity of essential oils (such as those produced by jasmine plants). Note that a single plant, for example, the strawberry, might be a short-day plant for one response and a long-day plant for another response.

Animal responses. There are also many responses to photoperiod in animals (see table), including control of several stages in the life cycle of insects (for example, diapause) and the long-day promotion in birds of molting, development of gonads, deposition of body fat, and migratory behavior. Even feather color may be influenced by photoperiod (as in the ptarmigan). In several mammals the induction of estrus and spermatogenic activity is controlled by photoperiod (sheep, goat, snowshoe hare), as is fur color in certain species (snowshoe hare). Growth of antlers in American elk and deer can be controlled by controlling day length. Increasing day length causes antlers to grow, whereas decreasing day length causes them to fall off. By changing day lengths rapidly, a cycle of antler growth can be completed in as little as 4 months; slow changes can extend the cycle to as long as 2 years. When attempts are made to shorten or extend these limits even more, the cycle slips out of photoperiodic control and reverts to a 10–12 month cycle, apparently controlled by an internal annual "clock." *See* MIGRATORY BEHAVIOR.

Seasonal responses. Response to photoperiod means that a given manifestation will occur at some specific time during the year. Day lengths are calculated as the time from the moment the uppermost part of the Sun (rather than its center) first touches the eastern astronomical horizon, that is, the horizon as it would appear on the open ocean, in the morning until the upper part of the Sun just touches the western astronomical horizon in the evening (**Fig. 2***a*). Thus, the day length at the Equator is slightly longer than 12 h. At the Arctic Circle, 66.5°N latitude, the center of the Sun just touches the astronomical horizon at midnight on the summer solstice, but part of the Sun is visible above the horizon at midnight for a few days before and after the summer solstice, as is evident from the day-length curve for 66.5°N latitude. Because of twilight, the day length perceived by an organism is somewhat longer than the curves show. In addition, different organisms respond to different light levels during twilight.

Response to long days (shortening nights) normally occurs during the spring, and response to short days (lengthening nights) usually occurs in late summer or autumn. Since day length is accurately determined by the Earth's rotation on its tilted axis as it revolves in its orbit around the Sun, detection of day length and its rate of change (Fig. 2*b*) provides an extremely accurate means of determining the season at a given latitude. Such other environmental factors as temperature and light levels also vary with the seasons but are clearly much less dependable from year to year. To an extent that depends on the species, day-length responses do interact with temperature and other factors, so a given plant or animal response may not be exactly the same from year to year.

Ecological relationships. All this is clearly important in the distribution of plants and animals. Wild plants, for example, growing in northern latitudes (of the Northern Hemisphere) usually flower only when days are long, while plants from more southerly climes respond when days are shorter. The response type may be either long-day or short-day, but the critical day length (the day length that must be exceeded for long-day plants or not exceeded for short-day plants) is nearly always shorter for plants in southern locations than for those to the north. This is also true for responses other than flowering. A given species of tree or shrub in Scandinavia, for example, has a much longer critical day length for spring shoot elongation when it grows in the far north than when it grows farther south. Plants growing higher on Scandinavian mountains also have longer critical day lengths for stem elongation, meaning that they will begin growth later in the spring.

An examination of the list of responses in both plants and animals confirms these ecological relationships. It is especially evident that flowering of plants and breeding times of animals are often

synchronized with the seasons by response to photoperiod. Most important is that the breeding times of members of a population are synchronized with each other by photoperiod. Actual flowering or breeding can occur at almost any time during the year other than coldest winter, but it would not be very successful if members of the same population were not able to breed at the same time. Humans are day-neutral for breeding time and can breed during any season. Indeed, it has been difficult to demonstrate photoperiodism responses in humans although a few studies suggest that humans do respond to day length. In a sense, seasonal affective disorder (SAD) may be an indication of human responses to daylength. Some individuals experience depression during winter in climates distant from the equator, and often this can be overcome by exposure to high light intensities for an hour each day. It seems possible that the high light levels early in the morning are extending the perception of day length. Certainly, artificial lights have extended the day length experienced by humans during the past century. We wonder what effect this might have had. For one thing, humans are much taller on average now than they were a century ago. Could this be a day-length response? Perhaps someone should study human growth hormone as influenced by day lengths.

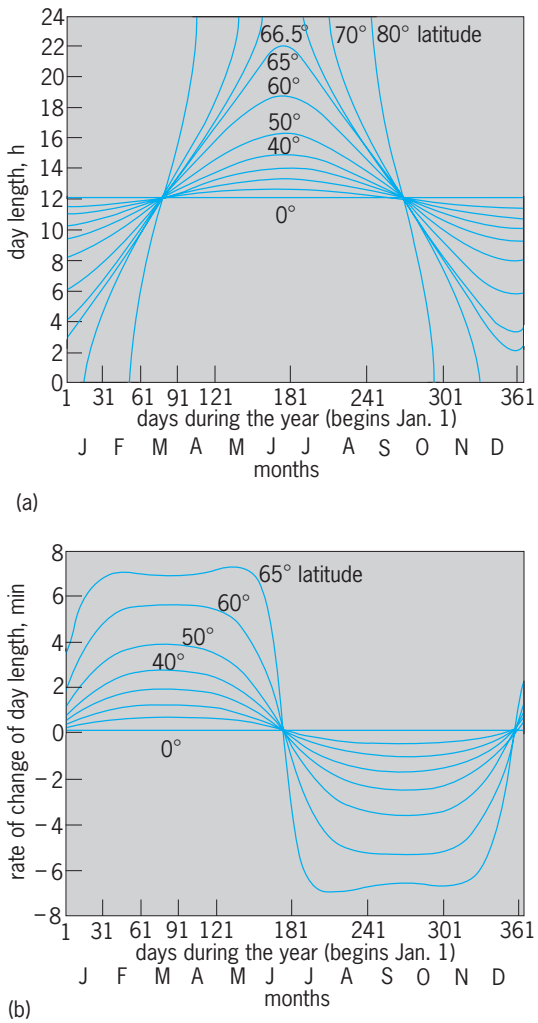

Fig. 2. Relationship of latitude and time of year on the length of day. (*a*) Day length as a function of time of year for various northern latitudes. (*b*) Rate of change of day length during the year. The more rapid rates of change occur during spring and autumn. (*After F. B. Salisbury and C. W. Ross, Plant Physiology, 4th ed., Wadsworth, 1991*)

Basis for anticipation. Photoperiodism provides for anticipation in the development of organisms. Thus many plants drop their leaves, becoming frost-resistant and dormant, in response to the shortening days of autumn, and insects enter diapause (a dormant condition) in much the same way. Some birds and mammals change their feather or fur color as winter approaches and then again with the coming of spring, both in response to changing day lengths. Birds also prepare for their migrations by storing body fat and developing restlessness.

Agricultural applications. There are also important applications of photoperiodism in agriculture. Day length, in addition to the limiting effects of temperature, moisture, or other factors, may restrict the area where a given crop can be grown. Northern cultivars of short-day soybeans, for example, can sometimes be grown successfully only in a belt of latitude about 50 mi (80 km) wide. If they are grown farther north, too much foliage is produced and too few flowers and therefore too few seeds; too far south means an excess of flowers and not enough foliage. Plant breeders who need to have different cultivars flower at the same time to facilitate crossbreeding can often control flowering by controlling day length. This is also done in commercial production of ornamental plants. For example, chrysanthemums, poinsettias, and orchids are caused to flower by shortening days with black curtains over the greenhouse benches, or they are kept vegetative with lights hung above the benches. Crops can also be bred for appropriate critical day lengths, or day-neutral types can be selected so that flowering and fruiting are independent of day length and thus latitude (except as other factors may limit). On the animal side, turkey production is routinely controlled by controlling day length, and chickens lay more eggs on long days, or when the night is interrupted with light.

Discovery. The importance of day length was overlooked until relatively recently, although A. Henfrey had suggested in 1852 that day length might influence plant distribution. The measurement of daily time by plants or animals must have seemed unlikely to nineteenth-century biologists. Nevertheless, Julian Tournois in Paris clearly discovered the short-day flowering responses of hops and hemp, publishing papers in 1910 and 1914. He observed that flowering in these plants occurred when days were shortened, but not when light levels were lowered and days lengthened to give the same quantity of light as with the short days. At the same time, in Heidelberg, Georg Klebs observed flowering of houseleek plants that were exposed to several days of continuous illumination with electric lights, but

he explained this long-day response in terms of nutrition and only suspected that the length of day was involved. Actually, the papers of Tournois and of Klebs had virtually no impact on the study of photoperiodism.

W. W. Garner and H. A. Allard discovered photoperiodism also and realized the importance of their finding; they gave the phenomenon its name in 1920. These workers observed, among other things, that Maryland mammoth tobacco flowered in their greenhouses in the winter but not during the summer in the fields. They investigated many of the environmental factors that might differ between summer fields and winter greenhouses, one of these being day length. Their tobacco plants bloomed profusely when subjected to days shorter than about 12 h by placing them in light-tight cabinets during summer. The plants remained vegetative when days were longer, extended in winter by artificial lights. They soon tried a number of other species, documenting and naming the short-day, long-day, and day-neutral response types for flowering (**Fig. 3**). The work of Garner and Allard gained much attention, and photoperiodism has been a field of study ever since.

W. Rowan exposed juncos in the middle of winter to artificial light that extended the day, publishing his results in 1925. These birds normally migrate northward only in spring, but when Rowan released them in the middle of winter after their long-day exposure, they immediately flew northward. Since that time the responses listed in the table as well as many others have also been documented in animals.

Complications. Many organisms that are day-neutral, in the sense that they will eventually flower or otherwise respond at virtually any day length, are nonetheless promoted in their response by some specific day length. Tomato is often cited as a typical day-neutral plant for flowering, but at proper temperatures its flowering is significantly promoted by short days. Plants that require a given day length are said to be qualitative, or absolute short-day or long-day plants, while those that would respond anyway but are nonetheless promoted by proper day length are said to be quantitative, or facultative, in response (Fig. 1). (In the absolute response, the quality, that is, the vegetative or reproductive condition, depends solely on photoperiod.) The majority of plants probably respond quantitatively for flowering; this is also true for stem elongation and many other responses.

Most animal responses are also of the quantitative type, but white-crowned sparrows have remained asexual for up to 2.5 years under short days, and their capacity to become sexual in response to long days remained unmodified during that time. The reproductive capacity of golden hamsters and sheep can be under complete control of day length for a long period of time, but eventually this control disappears. Hamster testes, for example, shrink to a small size on 12-h days but are full size on 13-h days—a sharp critical day. Placed under short-day regimens, the gonads soon shrink and cease to function, but after 25 to 30 weeks under short days, they begin to grow again, indicating the loss of photoperiodic control.

When, after a given period of time, an animal becomes unresponsive to further photoperiodic treatment of the same type, the animal is said to have entered the refractory period. Often the refractory period can be broken by treatment with an opposite day length. For example, gonadal development in certain sparrows is caused by daily 16-h light periods, but after several months the gonads begin to regress, even on these long days, and further long days will not cause them to develop again. Six weeks of 8-h days (16-h nights) restores the ability of a 16-h day to cause sexual development; that is, short days terminate the refractory period.

A somewhat similar phenomenon can be observed in certain species of plants, although the term refractory is not used by botanists. Some species flower only when short days are followed by long days (as happens in the spring) or when long days are followed by short days (as happens in autumn).

A few plant species (such as sugarcane and purple nutsedge) apparently require an intermediate day length; they do not flower if days are either too short or too long. Their counterparts are also known: plants that flower on short or long days, but not on days of intermediate length.

Frequently there is a strong interaction between temperature and day length. Some species (such as Japanese morning glory) have an absolute short-day requirement at high temperatures but are day-neutral at lower temperatures (a few degrees above freezing). Opposite situations are known where plants have absolute or quantitative photoperiod requirements at low temperatures but are day-neutral at higher temperatures. Many plants require or are promoted in their flowering by days to weeks of temperatures just above freezing. Such a low-temperature promotion of flowering is called vernalization. Typically, a vernalization requirement is followed by a long-day requirement or promotion (as with black henbane and many cereals), but sometimes the low-temperature treatment must be followed by short days (as for certain cultivars of chrysanthemum and various grasses). *See* VERNALIZATION.

The complexity of response may be surprisingly intricate. Sweet William (*Silena armeria*) requires long days at 82°F (28°C) to flower but is completely day-neutral at 41 or 90°F (5 or 32°C). It will flower at any temperature when treated with GA_7 (one of the gibberellic acids). *See* GIBBERELLIN.

Species also differ in sensitivity to photoperiod. A common weed, lamb's-quarters (*Chenopodium rubrum*), responds as a minute seedling in a petri dish, and the seed leaves (cotyledons) of Japanese morning glory are highly sensitive to photoperiod. In cocklebur the seed leaves are not sensitive, but the first true leaf becomes highly sensitive when it is half expanded (plants about 30 days old). The bud also

Fig. 3. Plant responses to photoperiodism. Short-day responses on left and long-day on right for each plant. (*a*) Tomato, a day-neutral plant (110 days). Both plants are flowering, although flowers are difficult to see. Note taller stem under long days, a typical response. (*b*) Cocklebur, a short-day plant for flowering (60 days). (*c*) Japanese morning glory, a short-day plant (35 days). (*d*) Spinach, a long-day plant (35 days). (*e*) Barley, a long-day plant (35 days). (*f*) Radish, a long-day plant (54 days). (*From F. B. Salisbury, The Flowering Process, Pergamon, 1963*)

becomes competent to flower. Some plants flower only when they are several years old, although it is not known whether they are actually responding to photoperiod.

Frequently, as plants age, they become less dependent on photoperiod for flower induction. Some species show changes in sensitivity to photoperiod within a few days after their seeds sprout; others require many weeks to many months for such changes. This is reminiscent of the eventual loss of photoperiod control over gonadal development in hamsters.

Plants also vary in the number of photoperiodic cycles required for response. Cocklebur, Japanese morning glory, perennial ryegrass, and a few other species respond to a single short day or long day (long day refers to ryegrass). Biloxi soybean requires at least three cycles, while chrysanthemums and many other plants require many more. Indeed, nearly mature flowers will abort from chrysanthemum plants moved from inductive short-day cycles to long days. Such plants as cocklebur that respond to a single cycle become induced; that is, flowering continues even though the conditions that were necessary to cause it are no longer present.

Mechanisms. It has long been the goal of researchers on photoperiodism to understand the plant or animal mechanisms that account for the responses. Light must be detected, the duration of light or darkness must be measured, and this time measurement must be metabolically translated into the observed response: flowering, stem elongation, gonad development, fur color, and so forth. Many of the basic mechanisms appear to differ not only between plants and animals, but among different species as well, although the roles (synchronization, anticipation, and so on) are similar in all organisms that exhibit photoperiodism. Although relatively little is known about many of these responses, the flowering response of a few plants has yielded to modern molecular techniques so that much is now known about the genes, proteins, and pigments that are involved in conversion to the flowering state.

Most of this work has been done with the model laboratory plant, *Arabidopsis thaliana*, a quantitative long-day plant. (Common names are "thale cress" or "mouse-eared cress," but these names are seldom used in scientific discussions.) This small mustard plant is only 6–12 inches tall after a flowering stock has grown from a rosette of leaves. *Arabidopsis* can

be grown from seed in test tubes. Treatments of seeds with x-rays or mutation-causing chemicals produces mutations, and then thousands of plants can be screened to find the desired mutations. Genes can also be transferred easily from one plant to another. These techniques have been used with this plant in numerous studies of functions besides flowering. Some photoperiodism findings with *Arabidopsis* have been extended to other species (for example, rice, a short-day plant), and in some cases the information learned from *Arabidopsis* has been found also to apply to insects, mice, and even humans. Many pages could now be filled with the details of how plants detect the light, how they measure time, and how they signal the vegetative bud to change from the production of leaves to the production of flowers. These mechanisms are only briefly discussed in the sections that follow.

Seeing light and dark. In 1938 Karl Hamner and James Bonner investigated whether day or night was more important in photoperiodism. Among other things, they interrupted the day with a short period of darkness and the night with a brief interval of light. Darkness during the day had essentially no effect on short-day cocklebur plants, but the plants completely failed to flower when the long dark period was interrupted by light. Other workers found that long-day plants were promoted in their flowering by a night interruption (**Fig. 4**). Again, the opposite nature of short-day and long-day responses is demonstrated. It was concluded from these experiments that day length was of secondary importance to night length, but the important role of photoperiod as well as darkness is now recognized. In any event, discovery of the effects of a night interruption led to an unexplored area of research in plant physiology in general and photoperiodism in particular.

Light interruption of the dark period. The results of studies on the effective times of light interruption are shown in Fig. 4. Strongest inhibition of flowering in short-day plants comes when the light interruption occurs around the time of the critical night (about 7–9 h for cocklebur plants), but actual effectiveness also depends on the length of the dark period. With short-day cockleburs, the shorter the night, the less the flowering and the longer the time that light inhibits flowering.

The irradiance or brightness level of an effective light interruption depends upon the duration of illumination. Higher levels mean shorter exposure times to produce a given inhibition or promotion—much as brighter light means shorter exposure time in photography. With cocklebur, a few seconds of light from four photo floodlamps held about a meter away are sufficient to saturate the inhibitory process, but complete inhibition can be produced by a naked 7.5-W incandescent bulb hanging all night a little over a meter away from the plants. Thus streetlights sometimes influence the photoperiod responses of nearby plants (for example, leaves might fail to fall in autumn).

Orange-red wavelengths (660 nm) used as a night interruption are by far the most effective part of the spectrum in inhibition of short-day responses, and the effects of orange-red light can be reversed by subsequent exposure of plants to light of somewhat longer wavelengths (730 nm), called far-red light. With some long-day plants, irradiation with far-red light and blue light are more effective in promotion of flowering than irradiation with orange-red light.

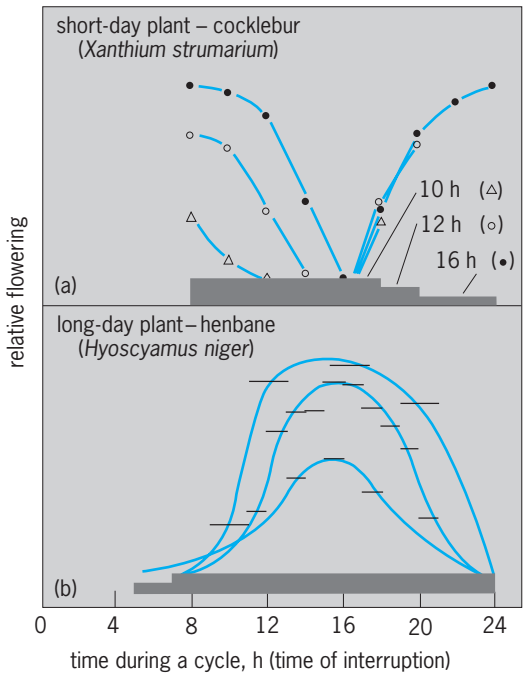

Fig. 4. Effects of a brief interruption of light given to plants during long dark periods. (*a*) Short-day plant inhibited in flowering by light interruptions (most effectively about 8 h after beginning of dark period) (*after F. B. Salisbury and J. Bonner, The reactions of the photoinductive dark period, Plant Physiol., 31:141–147, 1956*). (*b*) Long-day plant flowering promoted by a light interruption (1 or 2 h long as indicated by length of horizontal bars). Note maximum promotion near middle of dark period (*after H. Claes and A. Lang, Die Bltenbildung von Hyoscyamus niger in 48 Stndigen Licht-Dunkel-Zyklien und in Zyklien mit aufgeteilten Lichtphasen, Z. Naturforsch., 26:56–63, 1947*).

Some of the initial work on effectiveness of various wavelengths was done with photoperiodism, but many other responses were also studied. Examples include germination of many seeds, which is promoted by orange-red and inhibited by far-red wavelengths. These effects can be reversed many times; the last irradiation determines whether the seeds will germinate or not. Other examples include formation of pigments in apple skins and other plant parts and profound effects on seedlings that have been grown in the dark. Orange-red light inhibits stem elongation of dark-grown seedlings but promotes leaf expansion and straightening of a hook commonly seen at the tip of the stem in such seedlings. Far-red light will reverse all these effects, which are usually but not always independent of photoperiod.

These observations led to discovery in the 1950s of the phytochrome pigment system, which plays

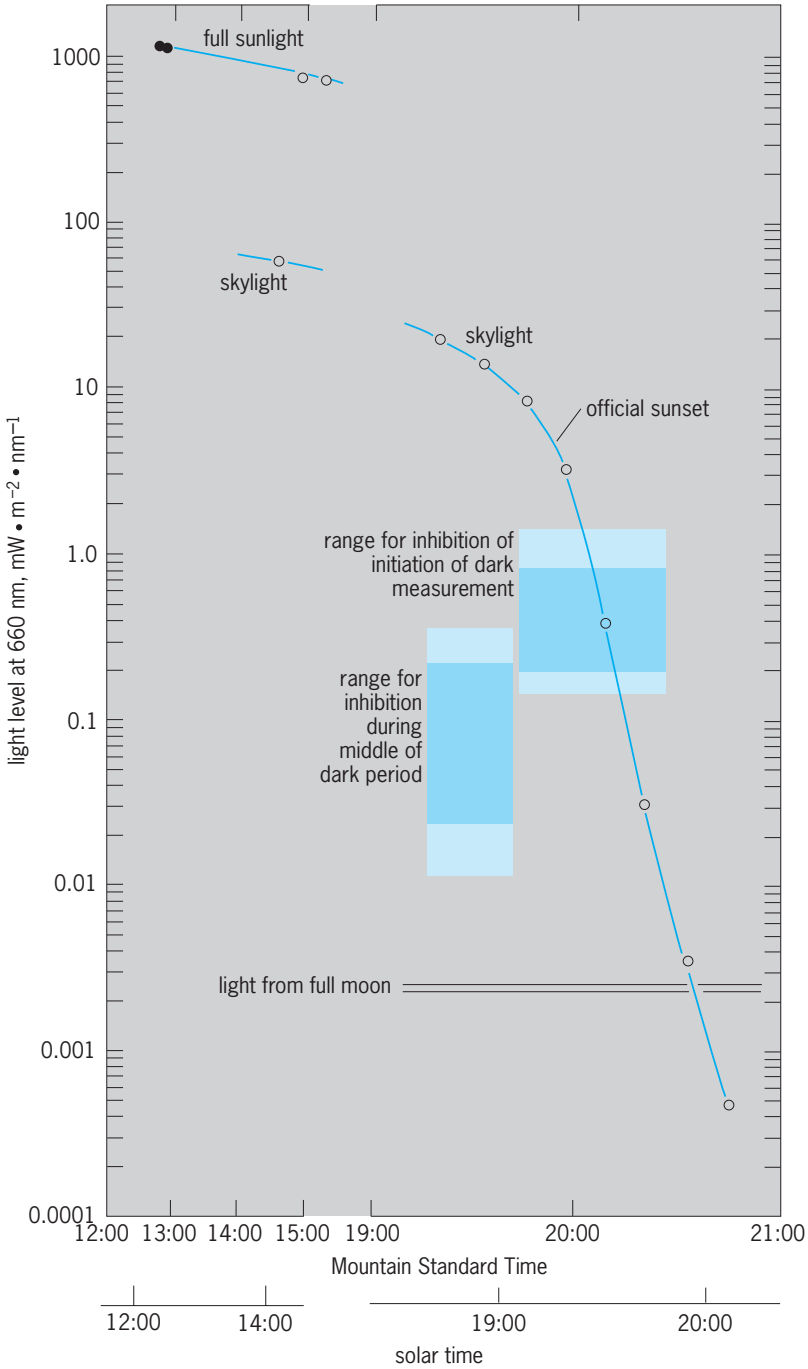

Fig. 5. Light levels at a wavelength of 660 nm (orange-red light) as a function of time on July 26, 1980, at Logan, Utah, and also showing light from a nearly full moon. The shaded areas represent ranges of light levels that inhibit the initiation of measurement of the dark period in cocklebur plants or inhibit flowering when applied for 2 h during the middle of a 16-h dark period. (*After F. B. Salisbury, The twilight effect: Initiating dark measurement in photoperiodism of Xanthium, Plant Physiol., June 1981*)

an extremely important role in plant responses to light, including in photoperiodism. (Phytochrome is also present in a few green algae and mosses but not in other algae, in fungi, or in animals.) *See* PHYTOCHROME.

Blue light has an effect on photoperiodism of many plants, and there are other important plant responses to blue light. For example, those dark-grown seedlings, as well as mature plants, will grow toward a source of light coming from the side, and blue light is by far the most effective in this response. This has been known for a least a century and a half, but for a hundred years the pigment responsible for the effect (called phototropism) remained "hidden, secret, mysterious." These words form part of the definition of the word *cryptic*; hence, long before it was isolated the pigment was named cryptochrome. It was finally isolated during the 1990s. Its role in plants called cryptogams (nonseed plants) also contributed to the name.

Other pigments are involved in the response (for example, phototropins), and some of them also play roles in absorbing the effective light in photoperiodism, but they are thought to be relatively minor compared with phytochrome and cryptochrome.

Thanks to the techniques of molecular genetics, we now know that there are five different genes controlling five different phytochrome molecules, which are referred to as phyA, phyB, phyC, phyD, and phyE. Each phytochrome molecule is a protein, and all have the same "chromophore" attached to a protein. It is the chromophore that actually absorbs the light, but the effects of this absorption depend on the proteins, and the various phytochromes play different roles in the varied light responses of plants. For example, phyB, phyD, and phyE are light stable and mediate responses to red light, but phyA is the main pigment that discriminates far-red light from darkness.

There are two genes for cryptochrome that control two cryptochrome proteins with their chromophores: cry1 and cry2. These pigments are especially important in setting the biological clock, as we'll note below.

Detecting dusk and dawn. In photoperiodism of short-day plants, an optimum response is usually obtained when the phytochromes are in the far-red-receptive form during the day and the red-receptive form during the night. Although normal daylight contains a balance of red and far-red wavelengths, the red-receptive forms are most sensitive, and so the phytochromes under normal daylight conditions are driven mostly to the far-red receptive form. At dusk this form is changed metabolically, and the red-receptive form builds up. It is apparently this shift in the forms of phytochromes that initiates measurement of the dark period. This is how a plant "sees." When the far-red-sensitive form of the pigment is abundant, the plant "knows" it is in the light; the red-sensitive form (or lack of far-red form) provides a biochemical indication that it is in the dark.

Figure 5 shows the changing intensity of red light during one cloudless twilight. As the sun began to set, the light level dropped at an increasingly faster rate until after official sunset, at which point light was decreasing by about a factor of ten each 10 minutes. When does the plant begin to "know" it is in the dark and should begin to time the length of the night? Experiments showed that the range of light

during which the plant changed from a light mode to a dark mode was surprisingly narrow (blue area in the figure labeled "range for inhibition of initiation of dark measurement"). Light intensity goes through this range in a relatively short time. Thus, as far as photoperiodism is concerned, a cocklebur plant responds almost as if it were controlled by a switch. During a period of about 5–11 min during twilight, the plant's mode of response changes from that of full daylight to that of complete darkness. This means that changing light levels during the day, including those caused by clouds, have no effect on the plant's photoperiodism responses. Light during the day and early twilight is always too bright for the plant to switch to the dark mode.

Detection of moonlight. Figure 5 also shows the range of light levels to which a cocklebur plant is sensitive in the middle of the dark period (such as light from the 7.5-W incandescent bulb mentioned above). The range of sensitivity levels during the middle of a long dark period is about 10 times lower on the logarithmic scale than the range of sensitivity at the beginning of the dark period. Nevertheless, this range is somewhat above the level of light from a full moon. Based on these rather meager and preliminary data, it appears not only that cocklebur plants ignore (as far as photoperiodism goes) changing light levels during the day, they are also insensitive to moonlight during the night. A very few experiments with other species in which plants were exposed directly to moonlight suggest that those plants responded to moonlight at a barely detectable level.

Light detection in animals. No phytochrome is found in animals although cryptochrome is often present, and genetic evidence indicates that cryptochrome does play a role in the photoperiod response in many animals. As a rule, however, animals are responsive to several wavelengths of light besides blue, suggesting that pigments in addition to cryptochrome must also play a role in the animal's response to daylength, and again evidence from molecular genetics has been used to identify some of these pigments.

A striking discovery is that the response to light can be independent or at least nearly independent of vision. In one experiment the eyes were removed from sparrows. From half of the test animals, all the feathers were removed on the top of the head, while the other birds were injected with India ink just below the skin on the top of the head to keep light from reaching the brain. The sparrows were then subjected to extremely dim green light, given for 8 or 16 h each day. Sparrows with the plucked heads responded to the long days as usual with increased gonadal development, but those with the India ink did not. Apparently, the light penetrating the skull of plucked birds was sufficient to induce a photoperiod response in those eyeless birds.

The eyes of mammals must detect the light used in photoperiodism, since blinding several species of rodents held under long days caused the same response as did a transfer to short days. Even in this case, however, signals from the eyes appeared to be carried by nerve tracts other than the primary optic pathways, indicating that even in mammals vision may be separated from other light effects. As with the cocklebur plants, this could allow the animals' photoperiodism mechanism to "ignore" changing light levels during the day and perhaps at night. It is essential for many animals to be able to see at night, but if they used the same light for both vision and photoperiodism, measurement of day or night might be unreliable. Bypassing the visual system in animal photoperiodism may be analogous to bypassing photosynthesis in plants.

Time measurement. The measurement of time—durations of the day or night—is the very essence of photoperiodism. With the discovery of the reversible nature of phytochrome in higher plants, it was suggested that time measurement might simply be the time required for the pigment to change in the dark from the far-red- to the red-receptive form. If timing were this simple, it would resemble an hourglass, in which the sand runs through after the hourglass has been inverted until it has all run out, just as the phytochrome pigment might change from one form to the other until the change is complete. There are elements of hourglass timing in plants (for example, the longer the dark period, within limits, the more the flowering in various short-day species), and the aphid *Megoura* seems to be completely controlled by an hourglass timer. At a certain point in this aphid's development, short days lead to production of fertile eggs when males are present; long days indefinitely maintain the virginopara condition (production by virgin females of eggs that develop into aphids). Results of all experiments designed to test the hourglass nature of the clock in this aphid agree with such a clock.

In most cases, however, the clock does not have simple hourglass characteristics. Rather than phytochrome requiring many hours to convert from one form to another, thereby accounting for the critical dark period of a short-day plant, phytochrome conversion appears to be complete in less than 1–3 h. Conversion only initiates a more subtle form of time measurement based on a clock that oscillates from one condition to another, each complete oscillation requiring about 24 h.

The discovery of such a biological clock in living organisms was strongly confirmed in the late 1920s. It was shown that the movement of leaves on a bean plant (from horizontal at noon to vertical at midnight) continued uninterruptedly for several days, even when plants were placed in total darkness and at a constant temperature, and that the time between given points in the cycle (such as the most vertical leaf position) was almost but not exactly 24 h. In the case of bean leaves, it was about 25.4 h. Many other cycles have now been found with similar characteristics in virtually all groups of plants and animals. Daily periods of animal activity under constant conditions provide other good examples. In all cases, the cycling continues, but under constant conditions, where the

cycles are "free-running" (not influenced by the usual external day/night cycle), the cycle lengths usually differ from 24 h by a few minutes to a few hours. This is strong evidence that the clocks are internal and not driven by some subtle daily change in the environment. Such rhythms are called circadian (from Latin *circa* = approximately, and *dies* = day).

Circadian rhythms usually have period lengths that are remarkably temperature-insensitive, which is also true of time measurement in photoperiodism. Furthermore, the rhythms are normally highly sensitive to light, which may shift the cycle to an extent that depends upon when the light is given. Thus, daily rhythms in nature are normally synchronized with the daily cycle as the Sun rises and sets each day. Their circadian nature appears only when they are allowed to manifest themselves under constant conditions of light (or darkness) and temperature, so that their free-running periods can appear.

Many lines of evidence show that the photoperiodism clock is dependent on the circadian clock. For example, it has been shown with several species that plants have not only an optimum night length for flowering but also an optimum day length. If, say, a constant night length is given but days are allowed to vary, the best flowering occurs when day plus night equals 24 h or multiples thereof; flowering is poor when day plus night totals multiples of about 36 h. This is also true with a constant day and varying nights. Thus, there is a cycling response with an optimum period length of about 24 h, just as in the circadian rhythms. Numerous other experiments have demonstrated the similarities between the photoperiodism clock and the clock controlling circadian rhythms. It is now accepted that the photoperiodism and the circadian clock are either the same or they are different clocks that have the same mechanisms of action. Nevertheless, although the clock controlling circadian leaf movements may be the same as the photoperiodism clock, the manifestations of the two can be separated. Thus leaf position bears no dependable relationship to the photoperiodism response of cockleburs, Japanese morning glories, or lamb's-quarters.

All of this supports a basic mechanism suggested by Erwin Bünning in Germany in 1936. (In 1930, Bünning and coworkers were the first to demonstrate the circadian nature of the clock controlling leaf movements in bean plants; that is, that under constant conditions the periods *approximately* but not *exactly* equaled 24 h.) Bünning suggested that flowering was promoted when the plant experienced light during one phase of the circadian clock and was inhibited when light occurred during the opposite phase. This is called the "external coincidence model" of photoperiodism.

Molecular mechanisms of the circadian clock are now well understood in various plant and animal species. They prove to be too complex for discussion here, but suffice it to say that in all organisms studied so far, the circadian clock is based on genes (DNA, now isolated and known) that are copied at a certain rate, forming messenger RNA, which is then "translated" into proteins that negatively feed back to inhibit the copying of the DNA. This produces a cycle in which the time it takes for the proteins to build up to sufficient concentrations to inhibit the copying, plus the time required for destruction of the proteins, accounts for the circadian period. *See* BIOLOGICAL CLOCKS.

Note that light, via the pigments discussed above, has two effects in photoperiodism. On the one hand, light "entrains" the circadian clock (for example, resetting the clock at dawn); on the other hand, light *interacts* with the phase of the circadian clock in controlling the photoperiodism response. Over all, the molecular and other evidence supports Bünning's external coincidence model of photoperiodism—with some minor modifications to the model.

Flowering hormone. The location of the photoperiodic flowering response in plants is the leaf. Exposing the stem and the bud to appropriate day lengths seldom has any effect. Apparently, some substance or other signal is produced in the leaf and moved to the bud where flowers are initiated. If such a substance is not a nutrient, and if it is effective in small quantities, then it would fit the definition of a hormone. In 1936, M. Chailakhyan named the hypothetical hormone florigen. He also showed that plants that have not been exposed to suitable day lengths can be made to flower by grafting them to plants that have been so exposed. Furthermore, day-neutral sunflower plants can be grafted to a vegetative cocklebur plant under long days, causing the short-day cocklebur to flower. Movement of this hypothetical florigen can be studied by removing the leaves of cocklebur or Japanese morning glory at various times following exposure to a single dark period. If leaves are removed immediately after the dark period, plants remain vegetative, as though the hormone had not had time to move out of the leaves. If leaves are removed several hours later, however, the plants flower nearly as well as when leaves are not removed at all. By then, the hormone has left the leaves.

There is a long history of attempts to extract florigen. Extracts from flowering plants would occasionally cause nonflowering plants to flower, but they induced only a minimal level of flowering, and the results could seldom be repeated by other investigators and usually not even by the investigator making the original report.

There is also evidence for a negative-acting inhibitor of flowering produced by photoperiodism. In a few species (such as short-day strawberries and long-day henbane), removing the leaves results in flowering under noninductive day lengths, as though the leaves normally produce an inhibitor under such day lengths. More than 40 years after his original experiments, Chailakhyan (working with I. A. Frolova and A. Lang) grafted short-day and long-day cultivars of tobacco to day-neutral cultivars and then grew them under various light conditions. When exposed

to short days, the short-day graft partners caused the day-neutral plants to flower earlier, and under long days, long-day partners also promoted flowering of day-neutral partners. These results confirm the presence of a positive- acting florigen. But when the short-day graft partners were maintained under long days, flowering in the day-neutral species was slightly or not at all retarded; under short days, the long-day tobacco completely prevented flowering of the day-neutral partner. Thus, the evidence for flowering inhibitors was nearly as impressive in this experiment as evidence for florigen itself.

Various chemicals known to influence plant growth in different ways have been applied to plants to see if flowering and photoperiodism could be affected. Although numerous effects have been observed, no clear picture has emerged, with the possible exception of the effects of various gibberellins (plant hormones that cause rapid elongation of plant stems and other responses). These substances will often substitute for long days or for cold treatment (vernalization), causing flowering of long-day plants held under short days or of vernalization-requiring plants never exposed to cold. Gibberellins seldom influence flowering of short-day plants.

As with the pigments and the clock involved in photoperiodism, identification of florigen has finally yielded to the techniques of molecular genetics—although tests still need to be done to confirm and pin down the identification. Again using the long-day *Arabidopsis*, a clock gene called *CONSTANS* (*CO*) varies in its activity with a circadian cycle. According to the coincidence model, flowering occurs in this long-day plant when *CO* is activated ("expressed") during the light period. This occurs in the phloem cells in the leaf. (Phloem cells make up the part of a plant's vascular system that transports dissolved sugars and other compounds from the leaf to other parts of the plant. Even fairly large proteins can move through the phloem.) The *CO* leads to production of a protein, CO, which activates another gene called *FLOWERING LOCUS T* (*FT*). *FT* is also expressed in the phloem of leaves. When *FT* is activated, it produces a messenger RNA (*FT* mRNA) that will be translated to the protein FLOWERING LOCUS T (FT). That protein is known to combine with another protein in the bud (meristem) of the plant, and the combination can activate several genes, all or some of which have been identified, to cause the bud to produce one or more flowers instead of only leaves. Evidence clearly indicates that the long-sought florigen that moves from the leaf to the bud is either the *FT* mRNA (most likely) or the protein FT itself.

There is clear evidence that photoperiodism affects hormones known to affect sexual reproduction in animals, particularly in vertebrates. The Japanese quail, for example, becomes sexually active when transferred from short to long days. Follicle-stimulating hormone (FSH), luteinizing hormone (LH), and testosterone increase in the quail under these conditions. Certain varieties of sheep, on the other hand, breed in the fall and are thus short-day animals for sexual activity. But again FSH and LH increase concurrently with the developing sexual activity. The situation is clearly highly complex in animals as well as plants. For example, in several mammals the pineal body (attached to the third ventricle in the brain) secretes an antigonadotropin that may cause testicular regression under unfavorable light schedules. Melatonin is known to be secreted by the pineal body in the golden hamster and the grasshopper mouse, and administration of this material suppresses testicle growth in both animals. Both are highly sensitive to photoperiod, but melatonin has no effect in two day-neutral species, the common house mouse and the laboratory rat. Thus there seems to be a relation of melatonin to photoperiodism, but since melatonin does not cause testicular regression in hamsters with their pineals removed, it may well act indirectly by causing the pineal to produce an unknown antigonadatropic material. *See* ENDOCRINE MECHANISMS; FLOWER; PLANT GROWTH; PLANT HORMONES; PLANT METABOLISM.

Frank B. Salisbury

Bibliography. S. D. Beck, *Insect Photoperiodism*, 2d ed., 1980; L. Corbesier and G. Coupland, *Plant, Cell and Environment*, 28:54–66, 2005; L. N. Edmunds, *Cellular and Molecular Bases of Biological Clocks*, 1987; F. B. Salisbury and C. W. Ross, *Plant Physiology*, 4th ed., 1992; B. M. Sweeney, *Rhythmic Phenomena in Plants*, 2d ed., 1987; L. Taiz and E. Zeiger, *Plant Physiology*, 4th ed., 2006; A. T. Winfree, *The Timing of Biological Clocks*, 1986; M. J. Yanovsky and S. A. Kay, *Nature Reviews/Molecular Cell Biology*, 4:265–275, 2003.

Photophore gland

A highly modified integumentary gland which arises from an epithelial invagination into the dermis. It becomes cut off from its site of origin and develops into a luminous organ composed of a lens and a light-emitting gland, at the back of which is a pigmented reflector of probable dermal-cell origin. These luminous bodies occur in deep-sea teleosts and elasmobranchs which live in areas of total darkness. *See* CHROMATOPHORE; EPITHELIUM; GLAND.

Olin E. Nelsen

Photoreception

The process of absorption of light energy by plants and animals and its utilization for biologically important purposes. In plants photoreception plays an essential role in photosynthesis and an important role in orientation. Photoreception in animals is the initial process in vision. *See* PHOTOSYNTHESIS; TAXIS.

The photoreceptors of animals are highly specialized cells or cell groups which are light-sensitive because they contain pigments which are unstable in the presence of light of appropriate wavelengths.

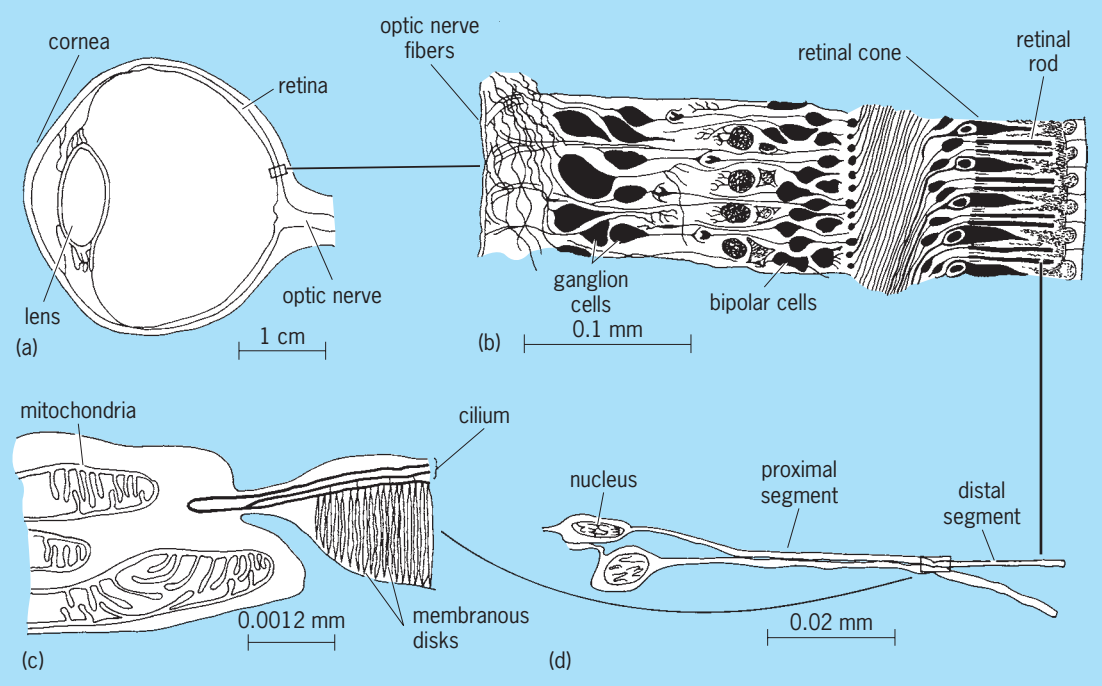

Fig. 1. Diagrammatic representation of gross and fine structure of photoreceptors of a typical vertebrate. (*a*) Section through eye; (*b*) layers of retina; (*c*) retinal rods (*after S. L. Polyak, The Retina, University of Chicago Press, 1941*). (*d*) Rod (*after F. DeRobertis, J. Biophys. Biochem. Cytol., 2(3):319–329, 1956*)

These light-sensitive receptor pigments absorb radiant energy and subsequently undergo physicochemical changes, which lead to the initiation of nerve impulses that are conducted to the central nervous system.

Morphology of photoreceptors. While their gross structures differ widely, photoreceptors of a variety of animals (protozoa, coelenterates, flatworms, annelids, mollusks, echinoderms, arthropods, and chordates) examined with the high-resolution electron microscope appear to have in common a fine structure featuring the presence of membranous organelles with appreciable surface area.

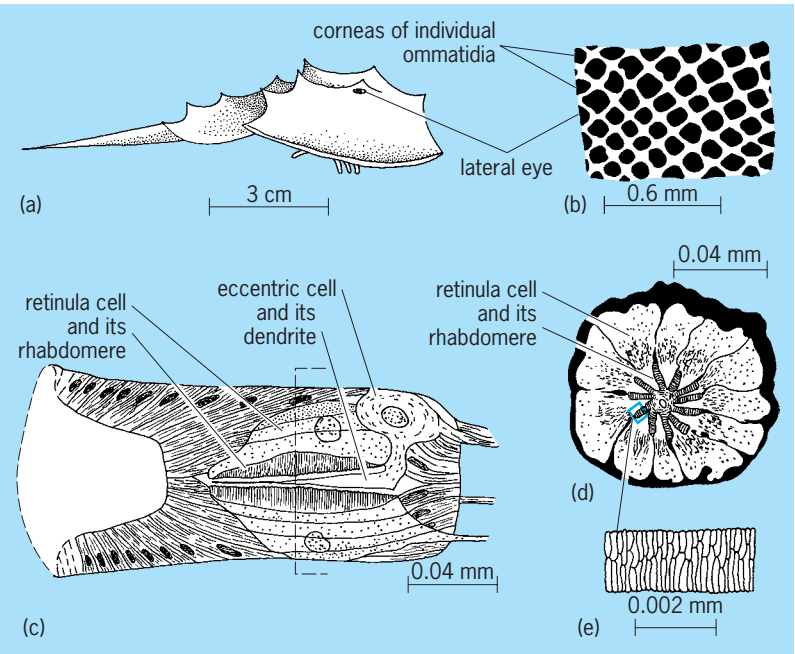

Fig. 2. Gross and fine structure of photoreceptors of the horseshoe crab (*Limulus*). (*a*) The organism. (*b*) The eye. (*c*) Longitudinal section through ommatidium. (*d*) Cross section through ommatidium; (*e*) microvilli of rhabdomere (*after W. H. Miller, Ann. N.Y. Acad. Sci., 74(2):204–209, 1958*).

Vertebrates. The structure of the vertebrate eye, exemplified by the human eye, is illustrated in **Fig. 1**. The retina is inverted so that the light-sensitive (photoreceptor) cells point toward the back of the eye (Fig. 1*b*). The visual pigments are located within the distal segments of the two types of photoreceptors, the rods (Fig. 1*b* and *c*) and the cones (Fig. 1*b*). The distal segment of the rod cell (Fig. 1*c* and *d*) consists of a stack of disks enclosed by a membrane. Electron microscopic investigation of rod distal segments fixed with osmium tetroxide or potassium permanganate or a freeze-drying procedure reveals each disk to be composed of three electron-dense membranes separated by two clear spaces of lower electron density; the outer two membranes are approximately 2 nanometers thick, the middle membrane about 4 nm thick, and the spaces about 3.5 nm thick. These findings, plus the results of investigations of the morphogenesis of rods, indicate that each disk is formed by the juxtaposition of two membranes caused by the invagination of the surface membrane of the distal segment. The disks are spaced 10–20 nm apart, and each distal segment contains many disks oriented perpendicularly to the light path; in the frog the estimated number of disks per rod distal segment is 1000. The disks are usually attached to a cilium which runs the length of the distal segment and connects the distal and proximal segments of the rod cell (Fig. 1*d*). The structure of distal segments of cones

is in general similar to that of the rod distal segment, but the disk membranes are somewhat thicker in the cones of some vertebrates, and in some cones the cilium appears to be absent. *See* ELECTRON MICROSCOPE.

Examination of the rod distal segments with polarized light reveals a marked dichroism, suggesting that the membranes contain light-absorbing pigment molecules whose long axes are oriented at right angles to the longitudinal axis of the distal segment and parallel to the transverse plane of the disks, but in this plane there is no evidence of further orientation. It has been demonstrated that if vitamin A deficiency is prolonged to the point where loss of the visual protein opsin occurs, there is widespread degeneration, that is, swelling and dissolution of the disks in distal segments of rods. These and other considerations have led to the view that the disk membranes contain the visual pigment and that the spaces between disk membranes contain a double layer of lipoid molecules. *See* DICHROISM (BIOLOGY); EYE (VERTEBRATE).

Arthropods. The compound eyes and ocelli of arthropods are morphologically different from vertebrate eyes in many respects, but the presence of a membranous organelle situated in the light path is a common characteristic. The lateral eye of the horseshoe crab (*Limulus polyphemus*) illustrates that the compound eye is an aggregate of ommatidia (**Fig. 2***a*), each ommatidium being a packet of photoreceptor cells with its own lens system (Fig. 2*b* and *c*). The structures of particular interest are those parts of the retinula cells in each ommatidium that form the rhabdomeres (Fig. 2*c*), which are the membranous organelles of appreciable surface area. Rhabdomeres of arthropods are composed of microvilli, tiny tubular extensions of the membrane of retinula cells, which in *Limulus* interdigitate with microvilli from the apposed surfaces of the membranes of adjacent retinula cells (Fig. 2*d* and *e*). Collectively the rhabdomeres constitute the rhabdome in the center of which is the dendrite of the eccentric sense cell (Fig. 2*c*), the membrane of which also contributes some microvilli to the rhabdome. (Eccentric cells of the *Limulus* type are not characteristic of arthropod compound eyes; similar cells have been found thus far only in the compound eye of the silkworm moth.) The membranes of the microvilli are approximately 7 nm thick and under high magnification with the electron microscope consist of two electron-dense boundaries separating a space of lower electron density approximately 3.5 nm wide. In those regions where microvilli make contact with each other, the outer boundaries fuse and form five layered membranes approximately 15 nm thick. In all arthropod eyes examined, the rhabdome is situated at or near the center of each ommatidium beneath the crystalline cone and directly in the light path. The microvilli are oriented in regular array with their long axes perpendicular to the long axis of the ommatidium. The diameters of the microvilli range from approximately 40 to 150 nm.

Although the basic structural plan of those arthropod eyes which have been examined is similar, differences in the architecture of the rhabdome exist, and in many arthropods and some cephalopods totaling more than 90 species these differences endow the eyes with the ability to discriminate among lights polarized in different planes. In the compound eye of the land crab (*Cardiosoma quanhumi*) this ability is in all probability conferred by the unique arrangement of the microvilli in the rhabdome (**Fig. 3***d*). The microvilli are arranged in plates or layers along the long axis of the ommatidium (Fig. 3*a*). In each plate the microvilli are straight and parallel but are perpendicularly oriented in alternate layers. The microvilli contributing to the alternate layers originate from different retinula cells (Fig. 3*b* and *c*). It has been demonstrated that the electrical response of any one retinula cell within an ommatidium of the *Cardiosoma* eye illuminated by flashes of polarized light rotated through 360° was modulated in amplitude with a period of 180°. Similar electrical observations have been made on the eyes of the crab (*Carcinus maenas*), two species of wolf spiders (*Arctosa variana*

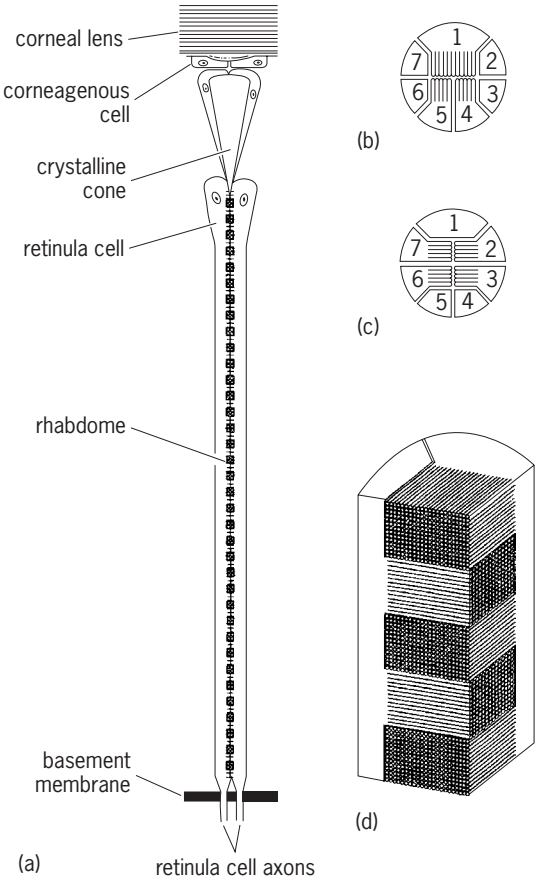

Fig. 3. Schematic diagram of the structure of a typical ommatidium from the crab compound eye. (*a*) Longitudinal section through the optic axis of an ommatidium. (*b*, *c*) Cross sections through different levels of the ommatidium, showing retinula cells from which the microvilli in the alternating layers originate. (*d*) Three-dimensional representation of fine structure of the rhabdome, showing the alternating layers of the tightly packed microvilli. (*After C. G. Bernhard, ed., Functional Organization of the Compound Eye, Pergamon Press, 1966*)

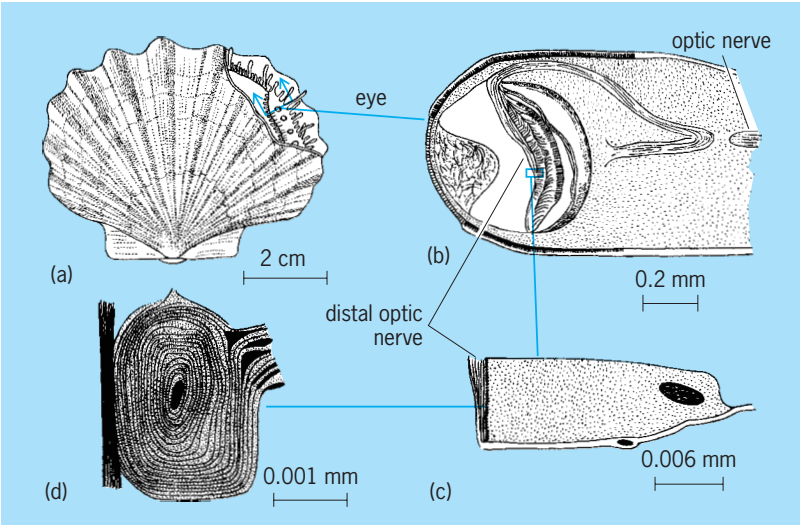

Fig. 4. Gross and fine structure of photoreceptors of the scallop (*Pecten*). (*a*) The organism. (*b*) Longitudinal section through eye; (*c*) distal sense cell (*after W. J. Dakin, Quart. J. Microsc. Sci., 4:49–112, 1910–1911*). (*d*) Appendage (*after W. H. Miller, Ann. N. Y. Acad. Sci., 74(2):204–209, 1958*).

and *Lycosa tarentula*), and the lobster (*Homarus vulgaris*). Compound eyes of worker honeybees, which are known to orient their flight by the plane of polarization of light reflected from the sky, possess a rhabdome which is roughly circular in cross section and consists of four quadrants in which the microvilli are parallel to those of the opposite quadrant but perpendicular to the microvilli of the adjacent quadrants.

There is consensus among investigators that visual pigments located within the rhabdomeres are responsible for the absorption of radiant energy, as well as the analysis of polarized light in arthropod eyes. Studies with polarized monochromatic light on the eye of the chalky mutant of the blowfly (*Calliphora erythrocephala*) indicate that the visual pigment is indeed located within the rhabdomeres and that the pigment molecules are oriented with their major axes parallel to the long axes of the microvilli. By virtue of this orientation, photons vibrating in a direction parallel to the microvilli are more readily absorbed.

Mollusca. Molluscan eyes, some of which are as sophisticated in gross structure as the vertebrate eye, contain membranous organelles of varying form. The eye of the squid and octopus contain rhabdomeres which are combined into rhabdomes and are composed of microtubules. The photoreceptor of *Pecten* (**Fig. 4**) has a unique membranous organelle, consisting of concentrically arranged light and dark bands which together constitute an oval or spherical body about 1 micrometer in diameter (Fig. 4*d*). The dark bands of this organelle are about 5 nm thick and appear to be continuous with the stalks of cilia which are attached to basal bodies within the cytoplasm of the receptor cell. *See* EYE (INVERTEBRATE).

Chemical behavior of photosensitive pigments. The more than 100 light-sensitive pigments of vertebrate and invertebrate photoreceptors which have been investigated appear to have a similar chemical constitution. Rhodopsin, the visual pigment in the distal segments of rod cells of many vertebrates and in the rhabdomeres of the invertebrates so far studied, consists of a protein conjugated with a carotenoid. The latter has been identified as the aldehyde of vitamin A, called retinal. Two types of retinal exist: retinal$_1$ (**Fig. 5**), and retinal$_2$, characterized by an additional double bond in the ring between carbon atoms 3 and 4. (Hereafter retinal without a subscript refers to retinal$_1$.) The protein moiety of rhodopsin, called opsin, has not been extensively characterized, and it is widely held that characteristics of opsins from different species vary and that this variance imparts species-specific properties to the rhodopsins. The absorption maximum of free retinal is at 387 nm; when it is combined with the protein opsin (molecular weight about 40,000), the visual pigment rhodopsin is formed and the absorption maximum shifts to higher wavelengths (498 nm in cattle rhodopsin, a visual pigment extensively investigated). The agreement of the absorption spectrum of many rhodopsins with the spectral sensitivity of the activity of photoreceptors from which the pigments were extracted clearly implicates these pigments in light reception.

Ultrastructure and rhodopsin. Evidence already cited suggests that rhodopsin forms an integral and highly organized part of the ultrastructure of vertebrate retinal rod distal segments and of the microstructure of invertebrate membranous organelles. In the frog, rhodopsin contributes about 40% of the weight of the outer segments of rods. Studies with polarized light indicate that rhodopsin molecules are highly oriented in vertebrate rods and in the dipteran rhabdomere; in both, the major axes of the pigment molecules are perpendicular to the long axis of the

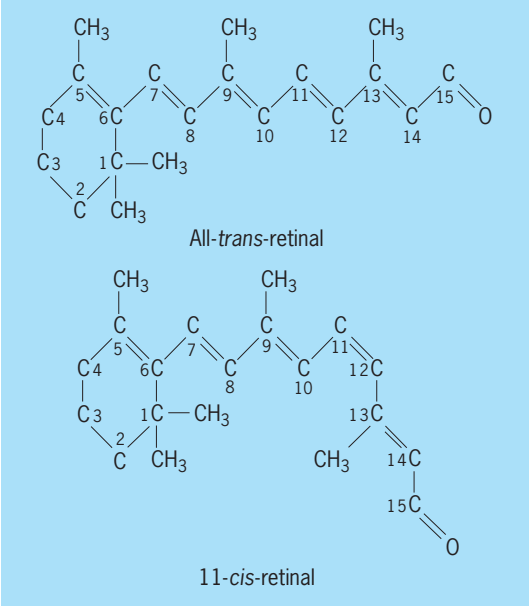

Fig. 5. Structure of two retinal isomers.

photoreceptor and, in the dipteran rhabdomere, parallel to the microvilli. Cattle rhodopsin has a phospholipid component which, if totally removed, destroys the stability of the pigment. The possibility exists that this phospholipid fraction represents a part of the structural fabric of the rod distal segment to which the rhodopsin molecule is bound.

Isomers of retinal. The bleaching and resynthesis of rhodopsin depend upon the existence of several isomers of retinal. The structural formulas of the most important pair are illustrated in Fig. 5. Only 11-*cis*-retinal, when incubated with opsin, forms rhodopsin. Illumination of a solution of rhodopsin results eventually in a release of retinal from the opsin in the all-*trans* form. Between these two events a number of intermediate transitory stages occur. The mechanism of bleaching of rhodopsin involves the absorption of a quantum of light energy by the molecule, the isomerization of 11-*cis*-retinal to all-*trans*-retinal, and via intermediate stages the hydrolysis of this incompatible isomer from its site on the protein.

Dark adaptation. After exposure to light, recovery of sensitivity by the receptor cell takes place; this process is referred to as dark adaptation. It is dependent upon an adequate supply of 11-*cis*-retinal and the rate of recombination of opsin with this active retinal isomer. The 11-*cis*-retinal apparently is supplied partly by the action of an enzyme, retinal isomerase, and upon the all-*trans*-retinal released from bleached rhodopsin. In addition, an equilibrium exists between vitamin A and retinal that is catalyzed by alcohol dehydrogenase and diphosphopyridine nucleotide. Opsin traps the 11-*cis*-retinal which is formed to reconstitute the visual pigment.

Other visual pigments. In addition to the rhodopsins, several other visual pigments have been identified in the retinas of vertebrates. Fresh-water fishes, some amphibians, and certain reptiles possess visual pigments which contain retinal_2. Retinal_2, when combined with rod opsin, forms a visual pigment called porphyropsin with an absorption maximum at 522 nm. Cones presumably contain a different opsin which combines with retinal to form a pigment called idopsin. This pigment, which has been isolated from chicken retinas, has an absorption maximum at 562 nm and regenerates more rapidly than rhodopsin. These properties of this cone pigment presumably account for the demonstrable shift in sensitivity to longer wavelengths in going from rod to cone vision in a variety of animals (the Purkinje shift) and the obviously faster recovery of cone sensitivity known to occur during dark adaptation. A fourth pigment, the existence of which is indicated by spectral sensitivity data, has been synthesized by incubating retinal_2 with cone opsin. The resulting pigment is called cyanopsin and has an absorption maximum at 620 nm. Microspectrophotometric measurements of absorption spectra of single cones in the tadpole and chicken suggest that a cone pigment having the absorption characteristics of cyanopsin exists in nature.

The properties of visual pigments are thus determined by the type of retinal, as well as the visual protein. Within the rhodopsin class, for example, the range of absorption maxima is 562–430 nm in the retinal_1 group and 620–510 nm in the retinal_2 group. This considerable range would seem to indicate that the visual proteins make important contributions to the characteristics of the visual pigments.

Considerable evidence indicates that cone pigments exist which have not yet been successfully isolated by extraction. The techniques of reflection photometry and microspectrophotometry of individual cones have produced results which indicate three populations of cones containing pigments with absorption maxima in ranges 445–450, 525–535, and 555–570 nm. The absence of specific cone sensitivities in certain human retinas has been shown to correlate well with the major types of color blindness. *See* COLOR VISION.

Invertebrate visual pigments. The visual pigments of invertebrates have not been as extensively studied as those of vertebrates, but those that are known contain retinal_1 (range in absorption maxima from 420 nm in the honeybee to 520 nm in *Limulus*) and seem to be chemically similar to the vertebrate pigments. There is evidence that the visual pigments in all rhabdomeres of arthropod eyes need not be identical; in the fly, rhabdomeres 1–6 contain a pigment with an absorption maximum at 510 nm, and rhabdome number 7 contains a pigment with a maximum at 470 nm. Photoreceptors of the squid eye contain rhodopsin which, following photoexcitation, exhibits isomerization of 11-*cis*-retinal to the all-*trans* form, but the retinal moiety remains attached to the protein. Photosensitive pigments based upon retinal have now been isolated from the eyes of insects, crustaceans, and arachnids.

Electrical activity of photoreceptors. Despite much information about the structure of photoreceptors and the chemistry of visual pigments, the mechanism which links photochemical events to the initiation of nerve impulses is still obscure.

Vertebrates. It has been known since the late 1800s that the vertebrate retina generates an electrical response upon illumination. The record of this electrical response, called the electroretinogram (ERG), is polyphasic in waveform with an initial cornea-negative component (the *a* wave), a large cornea-positive component (the *b* wave), and an off-response following cessation of illumination. In many cases a third relatively slow cornea-positive response (the *c* wave) is recorded which originates presumably in the pigment epithelium. It is established that the initial limb of the *a* wave and the off-response are of receptor origin and that the *b* wave is contributed by neural and perhaps glial cells located in front of the receptor layer. Nevertheless the ERG is an accurate index of visual sensitivity; it has been shown that human dark adaptation and spectral sensitivity curves determined electroretinographically agree well with those obtained using subjective criteria, and both methods yield results which

agree with the absorption spectra of visual pigments measured in place or in extracts. Measurement of the ERG is therefore a useful method for obtaining data in a variety of animals and has often been used in preference to behavioral techniques. The ERG has also been used clinically in the early diagnosis of diseases involving the retina.

By using glass micropipets filled with solutions of potassium chloride, intracellular receptor potentials have been recorded from the proximal (inner) segments of sense cells in the retina of the carp and the salamander (*Necturus maculosus*). The light-initiated intracellular potentials are sustained and of polarity such that the cell interior becomes increasingly negative. That is, the membrane potential increases and the electrical response is a hyperpolarizing one rather than the depolarizing type of response characteristic of nerve and other types of excitable tissues. By histological localization of a dye deposited by the recording micropipet, it has been determined that these receptor potentials were recorded from the proximal segments of the sense cells. In experiments with the carp retina the intracellular potentials were recorded under conditions permitting determination of the spectral sensitivities of single cones. The results of many experiments indicate three populations of cones of decreasing frequency of occurrence and with different spectral sensitivity maxima as follows: 611 ± 23 nm, 462 ± 15 nm, and 529 ± 14 nm. These maxima agree reasonably well with absorption maxima of single cones in the goldfish retina determined by microspectrophotometry, but the red maximum is displaced toward longer wavelength than in the retinas of higher vertebrates.

An electrical response of relatively small magnitude without detectable latency (less than 25 microseconds) has been recorded from both vertebrate and invertebrate photoreceptors in response to brief but intense light flashes. This response, aptly termed the early receptor potential, is mediated by the visual pigments but is resistant to agents such as low temperature, drugs, and removal of Na^+ and H_2O, which depress or abolish the late receptor potential. Similar rapid (also slower) light-evoked potentials have been recorded from the pigment epithelium of the toad eye, from green leaves, and from human skin. It is held that the early response of pigmented systems is the result of a charge displacement, which in the case of rhodopsin is coincident with the initial stages of bleaching, and that the early receptor potential is not directly involved in visual excitation.

The technique of using micropipets to record from single units in the vertebrate retina has been extensively applied, yielding considerable information on the responses of retinal elements removed from the primary photoreceptors by one, two, or more synapses. Such studies have been primarily directed toward an understanding of the neural organization of the retina.

Invertebrates. Invertebrate photoreceptors, especially those of the compound eye of *Limulus*, have been used to study the responses in nerve fibers from primary photoreceptor cells. Experiments in which impulse discharges from a single primary nerve fiber are studied in response to various intensities and durations of illumination delivered to the sense cell have demonstrated that (1) the frequency of impulse discharge up to some maximum is linearly related to the logarithm of the stimulus intensity; (2) within a certain critical duration, usually somewhere between 0.1 and 1.0 s, stimuli in which the product of intensity and duration is constant produce discharges of similar frequency and number; and (3) latency of the impulse discharge decreases as the intensity is increased. In addition, these electrophysiological recording techniques from primary sensory fibers have been used to measure spectral sensitivity, dark adaptation, interaction between ommatidia in the same eye, and the effect of manipulation of the environment in the immediate vicinity of photoreceptors.

The electrical response to illumination of arthropod ommatidia and many other invertebrate photoreceptors has been recorded from single light-sensitive cells by using micropipet electrodes. In the majority of cases the response is a sustained depolarization (that is, the cell interior becomes less negative) of gradually diminishing magnitude which may, depending upon the previous history of illumination and the stimulus parameters, exhibit an initial maximum.

In the retinula cells of *Limulus* and other arthropods spike potentials generally associated with the existence of nerve impulses are rarely recorded. However, responses usually consisting of sustained depolarization on which spike potentials are superimposed have been recorded from the eccentric cell in the lateral eye of *Limulus*. The light-induced depolarization of the eccentric cell is regarded as a generator potential, which initiates the nerve impulses in the axon. Spike potentials of variable but small amplitude have been recorded from retinula cells in the eye of *Limulus* but seem to be reflections of the spike potentials generated by the eccentric cell. The visual significance of the retinula cell in the *Limulus* lateral eye, which exhibits an electrical response to illumination that is independent of the eccentric cell response (except for the spike potentials), is not yet known. In the eyes of other arthropods nerve impulses are presumably initiated in structures behind the eye by virtue of the potential changes which occur in the retinula cells.

Light and electrical changes. It may thus be generally true that energy, absorbed by a photosensitive pigment located in the membranous organelles of photoreceptor cells, acts to produce a sustained electrical change, depolarization in invertebrates and hyperpolarization in vertebrates, by increasing the permeability of the cell membrane, and that this potential change acts as a generator to initiate nerve impulses in the visual pathway. *See* VISION.

Verner J. Wulff

Bibliography. D. J. Cosens and D. Vince-Pru (eds.), *The Biology of Photoreception*, 1984; A. Fein and

J. S. Levine, *The Visual System*, 1985; M. G. Holmes (ed.), *Photoreceptor Evolution and Function*, 1997; L. M. Hurvich, *Color Vision: An Introduction*, 1981; K. N. Leibovic, *Science of Vision*, 1990; J. Marvullo, *Color Vision*, 1989; S. P. McKee (ed.), *Optics, Physiology, and Vision*, 1990; H. Steave (ed.), *The Molecular Mechanism of Photoreception*, 1986.

Photorespiration

Light-dependent carbon dioxide release and oxygen uptake in photosynthetic organisms caused by the fixation of oxygen instead of carbon dioxide during photosynthesis. When the reactions of photosynthesis use oxygen instead of carbon dioxide, phosphoglycolate is formed and there is a loss of carbon from the photosynthetic pathway. Phosphoglycolate inhibits photosynthesis if it is allowed to accumulate in the plant. The reactions of photorespiration break down phosphoglycolate and recover 75% of the carbon for the photosynthetic reaction sequence. The remaining 25% of the carbon is released as carbon dioxide. Photorespiration reduces the rate of photosynthesis in plants in three ways: carbon dioxide is released; energy is diverted from photosynthetic reactions to photorespiratory reactions; and competition between oxygen and carbon dioxide reduces the efficiency of the important photosynthetic enzyme ribulose-bisphosphate (RuBP) carboxylase. There is no known function of the oxygenation reaction; most scientists believe it is an unavoidable side reaction of photosynthesis, but it is possible that photorespiration improves nitrate reduction. *See* CARBON DIOXIDE; OXYGEN; PHOTOSYNTHESIS.

Significance. Photorespiration reduces the potential photosynthetic rate by as much as 50%. Since photosynthesis provides the material necessary for plant growth, photorespiration inhibits plant growth by reducing the net rate of carbon dioxide assimilation (photosynthesis). Plants grow faster and larger when photorespiration is reduced by either raising the carbon dioxide concentration or lowering the oxygen concentration in the air surrounding a plant. Most of the beneficial effects on plant growth achieved by increasing CO_2 may result from the reduced rate of photorespiration. *See* PLANT GROWTH.

Precise measurements of the rate of photorespiration are difficult since carbon dioxide released in photorespiration is subject to refixation in photosynthesis during the measurement. A simple method has been devised for estimating the rate of photorespiration based on the properties of the photosynthetic enzyme RuBP carboxylase/oxygenase. The ratio of oxygenation (starting photorespiration) to carboxylation (the start of photosynthesis) is directly proportional to oxygen concentration and inversely proportional to carbon dioxide concentration, as expressed by the equation below, where Φ is the ratio of oxygenation to carboxylation at 77°F (25°C), O is the concentration of oxygen (in percent), and C is the concentration of CO_2 (in parts per million). The factor x, which varies directly with temperature, accounts for how well the enzyme distinguishes carbon dioxide from oxygen, and is called the specificity of the enzyme. Photorespiration is stimulated by high temperature and inhibited by low temperature. For example, for 77°F (25°C) the factor x is 6.7, for 59°F (15°C) the factor is 4.1, and for 95°F (35°C) the factor is 9.4. The effect of oxygen and carbon dioxide on photorespiration (**Fig. 1**) is greater at low carbon dioxide concentration and is less at high carbon dioxide concentration.

$$\Phi = x\frac{O}{C}$$

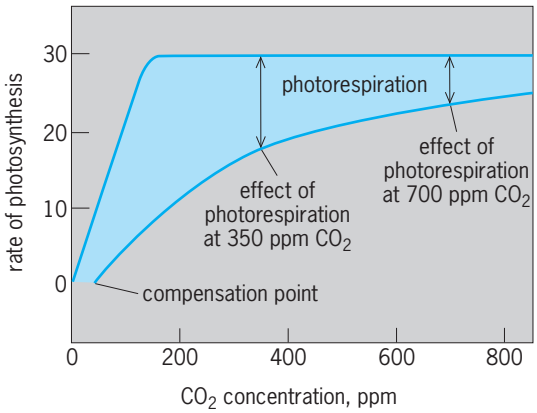

Fig. 1. Relation between photorespiration, photosynthesis, and carbon dioxide concentration. The top curve is the response of photosynthesis of a leaf to carbon dioxide when photorespiration is eliminated by reducing the oxygen concentration to 2%. The bottom curve is the response of photosynthesis to carbon dioxide when photorespiration occurs in the presence of 21% oxygen.

Treatments that cause the stomata (microscopic pores) of leaves to close often cause the carbon dioxide concentration inside leaves to fall, thus stimulating photorespiration. One example of this phenomenon is the effect of water stress (drought). The stomata on leaves close when water is in short supply. This reduces the rate of loss of water from leaves but also stimulates photorespiration by reducing the concentration of carbon dioxide inside the leaf. *See* PLANT-WATER RELATIONS.

Path of carbon. Photorespiration begins when a molecule of RuBP is oxygenated (**Fig. 2**). For each oxygenation event, one molecule of phosphoglycerate, the normal product of photosynthesis, is formed. The other part of the RuBP molecule forms phosphoglycolate, which is not a normal part of photosynthesis. Phosphoglycolate undergoes a series of reactions in three different compartments of the cell: the chloroplast, the peroxisome, and the mitochondrion. These reactions lead to the release of one carbon dioxide and one ammonia molecule for every two oxygenations. Most of the carbon (75%) is returned to the phosphoglycerate pool. The conversion of phosphoglycolate to phosphoglycerate and the recovery of the ammonia consume energy from the light reactions of photosynthesis, leaving less energy available for carbon dioxide fixation. *See* MITOCHONDRIA; PLANT CELL.

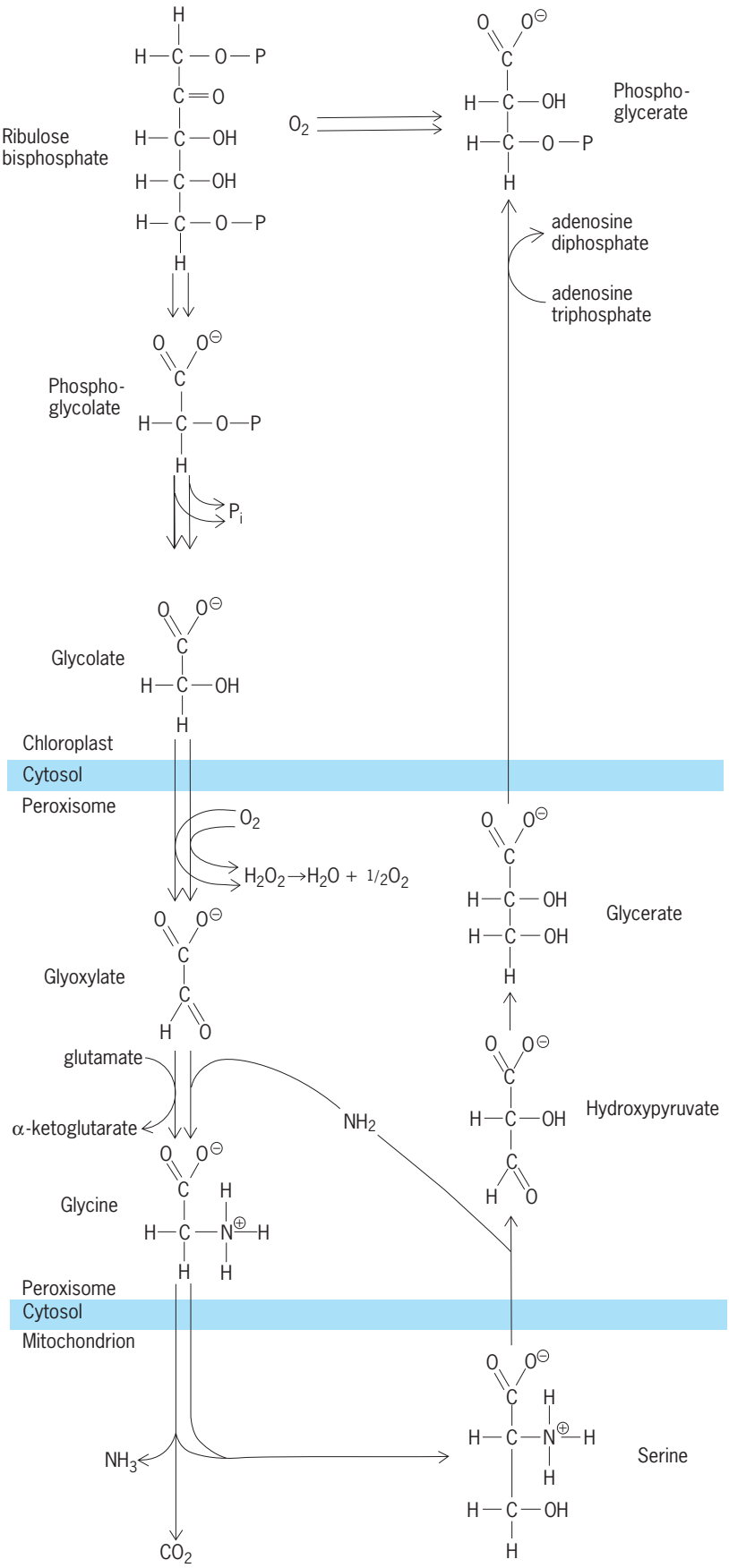

Fig. 2. Path of carbon in photorespiration. Phosphate (P) is normally a 50:50 mixture of HPO_3^- and PO_3^{2-}. Each arrow indicates a molecule following the pathway for every two oxygenations.

The first step in the metabolism of phosphoglycolate is removal of the phosphate group. This is catalyzed by the chloroplast enzyme phosphoglycolate phosphatase. The role of this enzyme in photorespiration was confirmed by discovery of a mutant plant having no phosphoglycolate phosphatase. This plant grows perfectly well under nonphotorespiratory conditions; however, in normal air, which causes photorespiration, it will not grow because it accumulates phosphoglycolate, which inhibits one of the reactions necessary for photosynthesis. Seven other classes of mutants have been found, and in each case the mutant plants can complete their life cycle when grown under nonphotorespiratory conditions but die when grown in normal air. This indicates that there are at least eight steps in plant metabolism that are essential for photorespiration but are not necessary for normal plant growth. *See* ENZYME; PLANT METABOLISM.

The next step in the path of carbon in photorespiration is the oxidation of glycolate to glyoxylate, which yields hydrogen peroxide as a by-product. The reaction occurs in the peroxisome, a specialized organelle that contains catalase, an enzyme that breaks down the hydrogen peroxide to water and oxygen. *See* PEROXISOME.

Amino groups are next added to the glyoxylate to make the amino acid glycine. One-half of the amino groups come from the conversion of serine to hydroxypyruvate (later in the pathway), with the other half coming from glutamate. The glycine molecules have two fates. Half of the glycine will be taken apart, releasing a carbon dioxide molecule and an ammonia molecule. Because of this release of carbon dioxide, the observed rate of photosynthesis is less than the actual rate of carbon dioxide fixation by one-half of the rate of oxygenation. The remaining carbon is attached to a derivative of folic acid and then transferred to a second glycine molecule to form serine, another amino acid. The serine transfers the amino group back to glyoxylate and becomes hydroxypyruvate, which is next reduced to glycerate and then phosphorylated to phosphoglycerate. *See* AMINO ACIDS.

Compensation point. During photorespiration, carbon dioxide is released at one-half of the rate of oxygenation. When the rate of oxygenation is twice the rate of carboxylation ($\Phi = 2$), the carbon dioxide released by photorespiration equals the carbon dioxide taken up by carboxylation, and so the net rate of photosynthesis is zero. This is called the compensation point (Fig. 1). The compensation point in most plants at 77°F (25°C) corresponds to a carbon dioxide concentration of about 50 ppm. This differs from the value calculated from the equation primarily because of a small effect of mitochondrial respiration.

Path of nitrogen. Fixed nitrogen is lost as ammonia, a volatile gas, during photorespiration. Most of the ammonia is recovered by the GS-GOGAT pathway, named after the major enzymes involved [glutamine synthetase (GS) and glutaryl:oxoglutarate aminotransferase (GOGAT), also called glutamate synthase (**Fig. 3**)]. Under normal conditions of

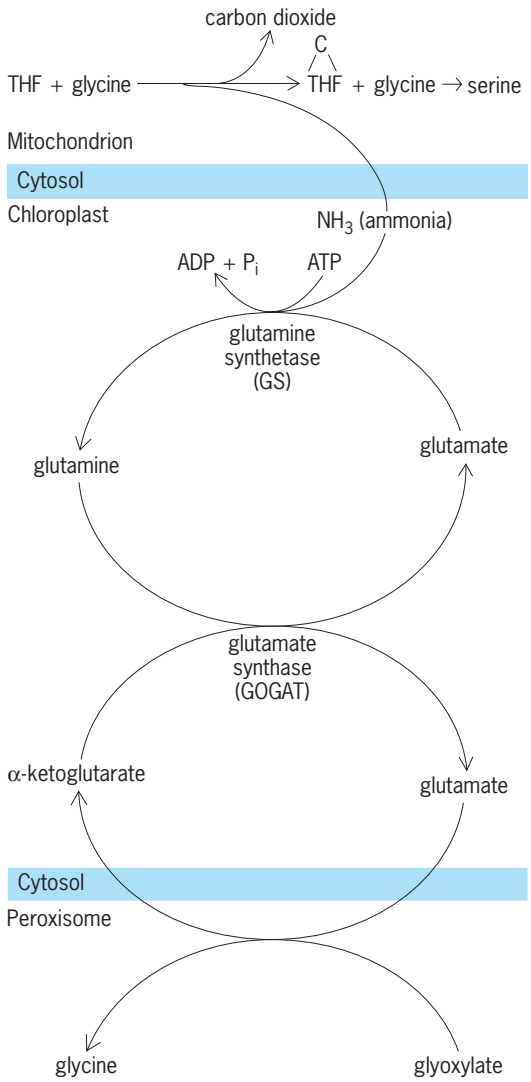

Fig. 3. GS-GOGAT ammonia cycle, which recovers the ammonia lost as a result of glycine metabolism during photorespiration. Tetrahydrofolate (THF) is a compound that can transfer single carbon atoms, in this case from one glycine to another.

photosynthesis, 5–20 ammonia molecules from photorespiration are fixed by the GS-GOGAT pathway for every one ammonia fixed for protein synthesis. The GS-GOGAT pathway consumes energy and therefore contributes to the loss of energy attributed to photorespiration.

Energetics. Three molecules of adenosine triphosphate (ATP) are required for each carboxylation in photosynthesis: one is used for converting ribulose phosphate to ribulose bisphosphate, and one is needed to convert each of the two phosphoglycerate molecules produced by carboxylation to sugars. For oxygenation, 3.5 ATP molecules are required: one molecule for converting ribulose phosphate to ribulose bisphosphate and one for converting the phosphoglycerate to sugar. In addition, three other reactions require ATP: conversion of glycerate to phosphoglycerate, reduction of phosphoglycerate to sugar, and recovery of ammonia. These three reactions occur once for every two oxygenation events and so are required at one-half of the rate of oxygenation. Therefore, the rate of ATP (energy) use is $3.5\ [2 + (3 \times 0.5)]$ per oxygenation. *See* ADENOSINE TRIPHOSPHATE (ATP).

Effect of carbon dioxide. The increasing level of carbon dioxide in the atmosphere is reducing the rate of photorespiration, an effect that can be calculated by using the equation above. In 1958, when precise carbon dioxide monitoring was begun at Mauna Loa, Hawaii, the carbon dioxide level was 315 ppm, which causes a ratio of oxygenation to carboxylation of $\Phi = 0.45$. By 2006, the carbon dioxide concentration had increased to 380 ppm, which reduced the ratio to $\Phi = 0.37$. Therefore, the rate of photorespiration has fallen by 18% in almost 50 years.

This decline in photorespiration may be responsible for an increase in the yearly oscillation in carbon dioxide concentration in the atmosphere. In the Northern Hemisphere, the carbon dioxide concentration is highest in May, then falls through the summer, reaching a minimum in October. This oscillation averaged 5.5 ppm (May to October decline) in the mid-1960s but averaged over 6 ppm in the mid-1980s. Stimulation of photosynthesis by the increased carbon dioxide concentration is believed to cause this change, and the stimulation of photosynthesis is caused, in large measure, by the inhibition of photorespiration.

Photorespiration in nature. Photorespiration is a process that is wasteful to the plant and may not be essential for plant survival. It is, therefore, surprising that a mutation eliminating photorespiration has not arisen. Many scientists believe that this indicates that oxygenation and the subsequent photorespiration are unavoidable side reactions of photosynthesis.

There are, however, some plants that avoid photorespiration by actively accumulating carbon dioxide inside the cells that have RuBP carboxylase/oxygenase. Many cacti do this by taking up carbon dioxide at night and then releasing it during the day to allow normal photosynthesis. These plants are said to have crassulacean acid metabolism (CAM). Another group of plants, including corn (*Zea mays*), take up carbon dioxide and attach it to a molecule in one part of the leaf, and then transport it to another part of the leaf for release and fixation by normal photosynthesis. The compound used to transport the carbon dioxide has four carbon atoms: thus these plants are called C_4 plants. Plants that have no mechanism for accumulating carbon dioxide produce the three-carbon compound phosphoglycerate first instead of the C_4 molecules that are seen first in C_4 photosynthesis. These plants are therefore called C_3 plants. Most species of plants are C_3.

The C_4 plants, CAM plants, and many algae avoid photorespiration by maintaining a very high concentration of carbon dioxide in the part of the cell containing RuBP carboxylase/oxygenase, which suppresses photorespiration. The RuBP carboxylase/oxygenase in carbon dioxide–accumulating plants often exhibits less specificity for carbon dioxide over

oxygen than the enzyme of C_3 plants, but photorespiration is still reduced because the carbon dioxide concentration is kept high by the carbon dioxide–accumulating mechanism.

The cost of actively accumulating carbon dioxide in many C_4 plants is about 2 ATP molecules, less than the cost of photorespiration in C_3 plants at 77°F (25°C) but more than the cost of photorespiration at 68°F (20°C). However, as the carbon dioxide level in the atmosphere increases, the cost of photorespiration in C_3 plants will fall, while the cost of carbon dioxide pumping in C_4 plants will remain constant. Hence, C_3 plants benefit more than C_4 plants from the increasing CO_2 concentration. *See* PHYSIOLOGICAL ECOLOGY (PLANT); PLANT RESPIRATION.

Thomas D. Sharkey

Bibliography. R. Boldt et al., D-Glycerate 3-kinase, the last unknown enzyme in the photorespiratory cycle in *Arabidopsis*, belongs to a novel kinase family, *Plant Cell*, 17:2413–2420, 2005; A. J. Keys et al., Photorespiratory nitrogen cycle, *Nature*, 275:741–743, 1978; W. L. Ogren, Photorespiration: Pathways, regulation, and modification, *Annu. Rev. Plant Physiol.*, 35:415–442, 1984; S. Rachmilevitch, A. B. Cousins, and A. J. Bloom, Nitrate assimilation in plant shoots depends on photorespiration, *Proc. Nat. Acad. Sci. USA*, 101:11506–11510, 2004; T. D. Sharkey, Estimating the rate of photorespiration in leaves, *Physiol. Plant*, 73:147–152, 1988; C. R. Somerville, An early *Arabidopsis* demonstration: Resolving a few issues concerning photorespiration, *Plant Physiol.*, 125:20–24, 2001; L. M. Voll et al., The photorespiratory *Arabidopsis shm1* mutant is deficient in *SHM1*, *Plant Physiol.*, 140:59–66, 2006.

Photosphere

The apparent, visible surface of the Sun. The photosphere is a gaseous atmospheric layer a few hundred miles deep with a diameter of 864,000 mi (1,391,000 km; usually considered the diameter of the Sun) and an average temperature of approximately 5800 K (10,500°F). Radiation emitted from the photosphere accounts for most of the solar energy flux at the Earth. The solar photosphere provides a natural laboratory for the study of dynamical and magnetic processes in a hot, highly ionized gas.

When studied at high resolution, the photosphere displays a rich structure. Just below it, convective motions in the Sun's gas transport most of the solar energy flux. Convective cells penetrate into the stable photosphere, giving it a granular appearance (see **illus.**) with bright cells (hot rising gas) surrounded by dark intergranular lanes (cool descending gas). A typical granule is approximately 600 mi (1000 km) in diameter. Measurements of horizontal velocity reveal a larger convective pattern, the supergranulation, with a scale of approximately 20,000 mi (30,000 km); the horizontal motion of individual granules reveals intermediate-scale (3000-mi or 5000-km) convective flows.

Magnetic fields also play an important role in shaping photospheric structure. Magnetic flux tubes, smaller than the intergranular lanes, penetrate the photosphere vertically. When many of these tubes are forced together, probably by convective motions deep within the Sun, they suppress the convective motions near the surface, cool the atmosphere, and form dark magnetic pores and sunspots.

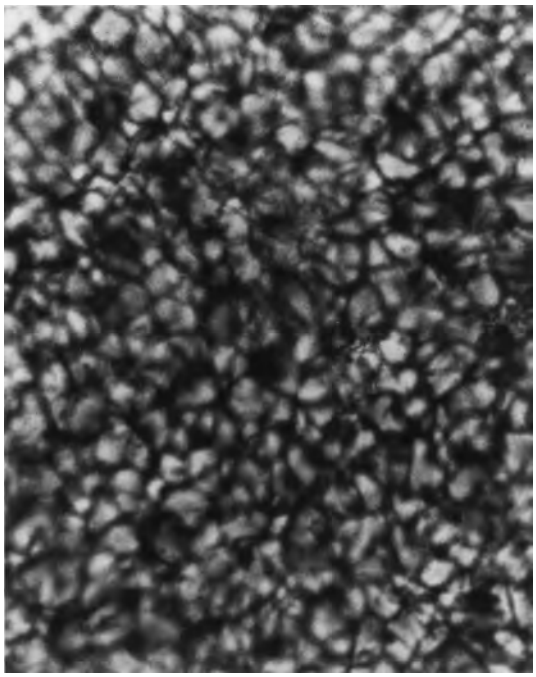

Image of the solar photosphere made in white light near 500 nm, using the 75-cm (30-in.) Vacuum Tower Telescope of the National Solar Observatory at Sacramento Peak, Sunspot, New Mexico. (*National Solar Observatory*)

Gravitational, acoustic, and magnetic waves are propagated through, and at some frequencies trapped in, the photosphere. The trapped waves, most of which have periods of about 5 min, represent global oscillations of the Sun and provide information about its interior when analyzed by seismological techniques. *See* HELIOSEISMOLOGY; SUN.

Stephen L. Keil

Bibliography. R. J. Bray, R. E. Loughhead, and C. J. Durrant, *The Solar Granulation*, 2d ed., 1984; R. Giovanelli, *Secrets of the Sun*, 1984; R. W. Noyes, *The Sun, Our Star*, 1982; H. Zirin, *Astrophysics of the Sun*, 1988.

Photosynthesis

Literally, the synthesis of chemical compounds using light. The term photosynthesis, however, is used almost exclusively to designate one particularly important natural process: the use of light in the manufacture of organic compounds (primarily certain carbohydrates) from inorganic materials by chlorophyll- or bacteriochlorophyll-containing cells. This process requires a supply of energy in the form of light, since its products contain much more chemical energy than its raw materials. This is clearly shown by the liberation of energy in the reverse process,

namely the combustion of organic material with oxygen, which takes place during respiration. *See* CHLOROPHYLL; PLANT RESPIRATION; RESPIRATION.

Among chlorophyll-containing plant and algal cells, and in cyanobacteria (formerly known as blue-green algae), photosynthesis involves the oxidation of water (H_2O) to produce oxygen molecules, which are then released into the environment. This is called oxygenic photosynthesis. In contrast, bacterial photosynthesis does not involve O_2 evolution (production). In this case, other electron donors, such as H_2S, are used instead of H_2O. This process is called anoxygenic photosynthesis. Both types of photosynthesis are discussed below.

The light energy absorbed by the pigments of photosynthesizing cells, especially by chlorophyll or bacteriochlorophyll pigments, is efficiently converted into stored chemical energy. Together, the two aspects of photosynthesis—the conversion of inorganic matter into organic matter, and the conversion of light energy into chemical energy—make it the fundamental process of life on Earth: it is the ultimate source of all living matter and of almost all of the energy of life. All types of photosynthesis require light; therefore, photosynthetic organisms are generally restricted to the photic zone, that is, the narrow region of the Earth near the surface that receives sunlight. The only known exceptions are the anoxygenic photosynthetic bacteria that live near deep-sea hydrothermal vents and utilize the very weak light given out by the hot vents.

The total productivity of photosynthesis on Earth has been estimated in various ways. The most accurate modern method uses satellite remote-sensing data to estimate both land-based and oceanic photosynthesis. These estimates take into consideration the fact that the vegetation in these areas make use of only 1% of the entire solar spectrum on average, or 2% if only visible light is considered. The net primary productivity of photosynthesis on Earth has been determined to be about 115 petagrams of carbon per year [1 petagram (Pg) equals 10^{15} grams or 10^9 metric tons]. The contributions of land- and ocean-based primary production are roughly equal, although the accumulation of biomass on land is much higher than in the ocean. This is due to the fact that organisms rapidly recycle the fixed carbon in the oceans by consuming it.

Oxygenic photosynthesis. The net overall chemical reaction of oxygenic photosynthesis (by plants, algae and cyanobacteria) is shown in Eq. (1), where

$$H_2O + CO_2 + \text{light energy} \xrightarrow[\text{enzymes}]{\text{chlorophyll}} \{CH_2O\} + O_2 \qquad (1)$$

$\{CH_2O\}$ stands for a carbohydrate (sugar). The photochemical reaction in photosynthesis belongs to the type known as oxidation-reduction, with CO_2 acting as the oxidant (electron acceptor) and water as the reductant (electron donor). The unique characteristic of this particular oxidation-reduction is that it is energetically unfavorable; that is, it converts chemically stable materials into chemically unstable products. Light energy is used to make this "uphill" reaction possible. A considerable part of the light energy used for this process is stored as chemical energy.

Temporal phases of photosynthesis. Photosynthesis is a complex multistage process that consists of nearly a hundred physical processes and chemical reactions. To make this complex process more understandable, it is useful to divide it into four temporal stages. Each phase is based roughly on the time scale in which it occurs. These phases are (1) photon absorption and energy transfer processes in antennas (or antenna chlorophylls, molecules which collect light quanta); (2) primary electron transfer in photochemical reaction centers; (3) electron transport and adenosine triphosphate (ATP) formation; and (4) carbon fixation and export of stable products.

Phase 1 takes place in the femtosecond [1 femtosecond (fs) = 10^{-15} s] to picosecond [1 picoseond (ps) = 10^{-12} s] time scale. Phase 2 takes place in the picosecond to nanosecond time scale [1 nanosecond (ns) = 10^{-9} s]. Phase 3 takes place on the microsecond [1 microsecond (μs) = 10^{-6} s] to millisecond [1 millisecond (ms) = 10^{-3} s) time scale. Phase 4 takes place in the millisecond to second time scale. We will consider each of these phases of photosynthesis in more detail below. The time scales are shown schematically in **Fig. 1**.

Sites of photosynthesis. The photosynthetic process takes place within certain structural elements found in plant cells. All algae, as well as all higher plants, contain pigment-bearing subcellular organelles called chloroplasts. In the leaves of the higher land plants, these are usually flat ellipsoids about 5 micrometers in diameter and 2.3 μm in thickness. Ten to 100 of them may be present in the average parenchyma cell of a leaf. Under the electron microscope, all chloroplasts show a layered structure with alternating lighter and darker layers of roughly 0.01 μm in thickness. These layers are membranes that contain proteins. These membranes are called thylakoid membranes (thylakoid

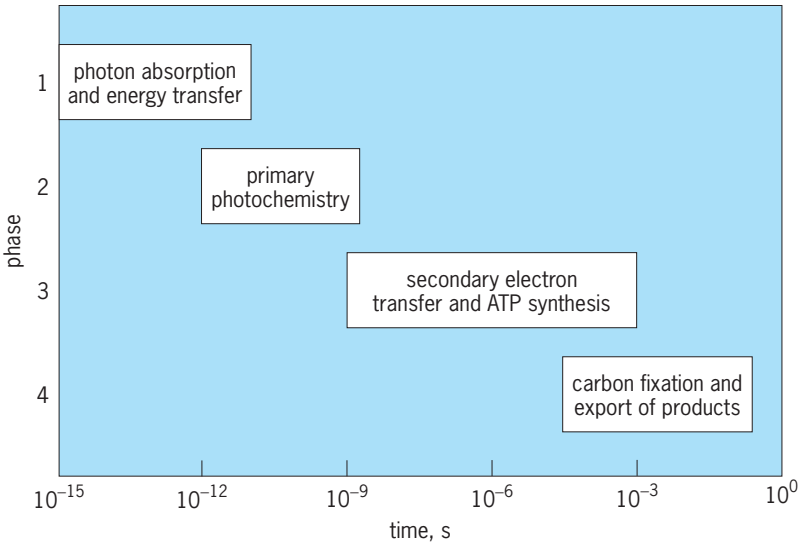

Fig. 1. Four temporal phases of photosynthesis.

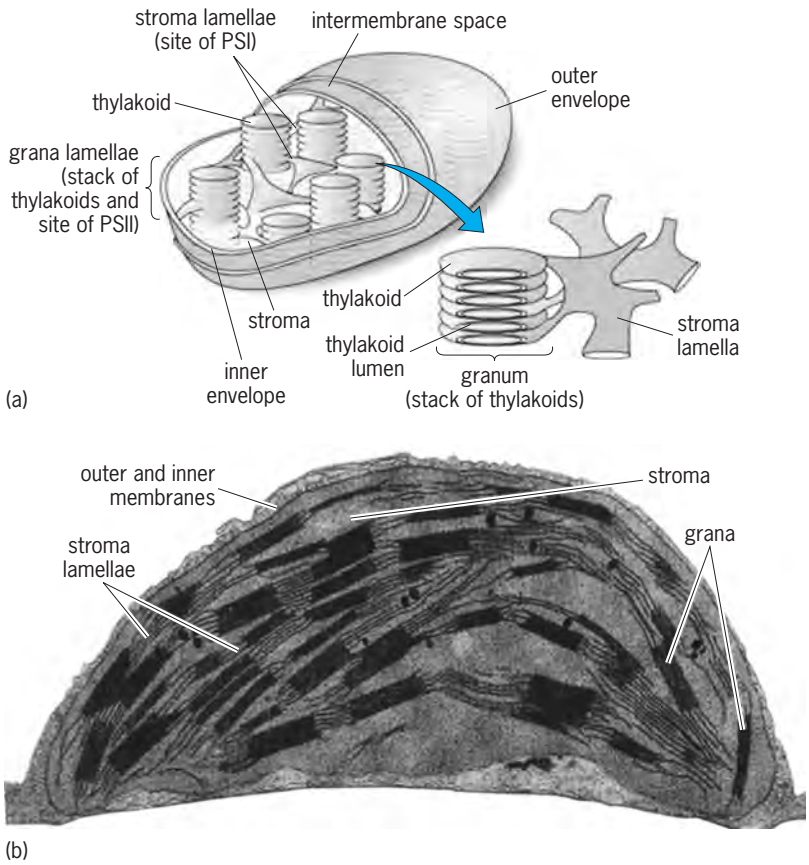

Fig. 2. A typical chloroplast. (*a*) Schematic diagram of a higher plant chloroplast. (*b*) Electron micrograph of a chloroplast. (*After L. Taiz and E. Zeiger, 2006*)

stands for a membrane sac). The proteins bind all of the chlorophyll. The membranes are the sites of the first three phases of photosynthesis. In algae, the number and shape of chloroplasts vary much more. The green unicellular alga *Chlorella* contains only one bell-shaped chloroplast. A typical chloroplast is shown in **Fig. 2**. *See* CELL PLASTIDS.

The photochemical apparatus is less complex in cyanobacteria. These cells are prokaryotes and therefore lack a nucleus and organelles such as chloroplasts and mitochondria. The early phases of photosynthesis take place on thylakoid membranes, which extend throughout the interior of the cell.

Quantum energy storage process. In photosynthesis, the energy of light quanta is converted into chemical energy. As seen in Eq. (1), the conversion of 1 mole of CO_2 and 1 mole of H_2O produces 1 mole of a carbohydrate group and 1 mole of oxygen. This reaction results in the storage of about 469 kilojoules of total energy per mole or, under natural conditions, about 502 kJ of potential chemical energy (free energy) per mole. Pigments absorb light in the form of quanta (photons). *See* ABSORPTION OF ELECTROMAGNETIC RADIATION.; PHOTON.

Discovery of two photosystems. A specific mechanism in which two photochemical events cooperate to carry out oxygenic photosynthesis was suggested by experiments of Robert Emerson. Emerson discovered that the "maximum quantum yield of photosynthesis" (number of O_2 molecules evolved per absorbed quantum), while constant at the shorter wavelengths of light (red, orange, yellow, green), declines in the far-red above 680 nanometers (the "red drop"). Later, Emerson and coworkers showed that this low yield could be enhanced if both chlorophyll *a* and *b* are simultaneously excited (only chlorophyll *a* absorbs above 680 nm). This effect, now known as the Emerson enhancement effect, suggests that two different pigments must be excited to perform efficient photosynthesis, thus indicating that two light reactions are involved in photosynthesis. Further, it suggests that one of these pigments is sensitized by light absorption in chlorophyll *a* and one by absorption in another pigment (for example, chlorophyll *b*). Experiments by others enlarged Emerson's observation by suggesting that plants contained two pigment systems. One (called photosystem I, or PSI; sensitizing reaction I) is primarily composed of chlorophyll *a*; the other (called photosystem II, or PSII; sensitizing reaction II) is also composed of chlorophyll *a*, but includes most of chlorophyll *b* or other auxiliary pigments. These other auxiliary pigments include red and blue pigments, called phycobilins, in red algae and cyanobacteria, for example, and the brown pigment fucoxanthol in brown algae and diatoms. Efficient photosynthesis requires the absorption of an equal number of quanta in PSI and in PSII. This ensures that within both systems excitation energy is absorbed by the antenna system and partitioned to each photosystem, where the energy drives the chemical reactions.

Robert Hill proposed that one of these reactions involves the transfer of an electron from an intermediate to cytochrome b_6 (shown later to be a plastoquinone molecule) during the conversion of water to oxygen). The other reaction, he proposed, involves the transfer of an electron from cytochrome *f* to an intermediate in the conversion of CO_2 to carbohydrate. The intermediate transfer of an electron from reduced plastoquinone (plastoquinol) to cytochrome *f* can occur without energy input because the former is a stronger reductant than the latter. There is a great deal of experimental evidence for the existence of two pigment systems and for the key role of plastoquinone and cytochrome *f* in this sequence. Louis N. M. Duysens and Jan Amesz conducted a crucial early experiment showing the existence of two photosystems. In this, they showed that light 1, absorbed by chlorophyll *a*, oxidizes cytochrome *f*, and that light 2 (absorbed in other photosynthetic pigments, including chlorophyll *a*) reduces this oxidized cytochrome *f*. The results suggest that the two light reactions are arranged in series.

Z scheme and the proteins that carry it out. The two photochemical events that were discovered at that time are now known to take place on large protein complexes embedded in the thylakoid membrane. **Figure 3** shows the four large protein complexes

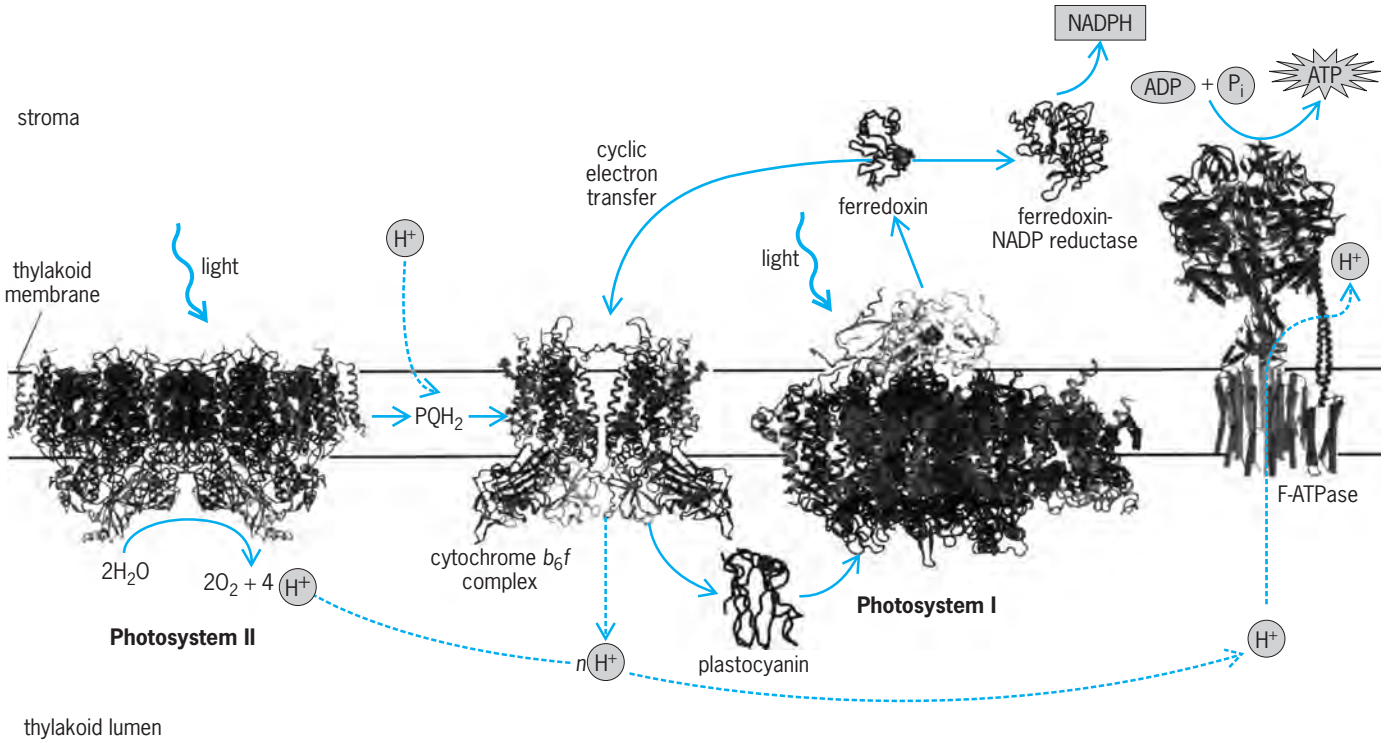

Fig. 3. Structural diagram of the four protein complexes present in the thylakoid membrane, as characterized by x-ray crystallography. Photosystem II (PSII), the cytochrome b_6f complex (cyt b_6f), photosystem I (PSI), and the ATP synthase are shown in the figure. Electron transfer steps are shown as solid arrows while proton flow is indicated by dashed arrows.

found in the thylakoid membrane, and demonstrates the interactions between photosystems I and II; the cytochrome b_6f complex functions between the photosystems, and the ATP synthase which makes ATP. Arrows show the pathway of electrons and protons (H^+), the latter is essential for the formation of ATP. Note that the membrane has an inherent asymmetry in that the protein complexes are oriented in a particular way in the membrane. This orientation is essential to the proper functioning of photosynthesis.

Figure 4 shows the Z scheme, summarizing the way in which the two photosystems cooperate to carry out electron transfer reactions involved in photosynthesis. It is an energetic diagram, in that the energy of the component, which is measured as the

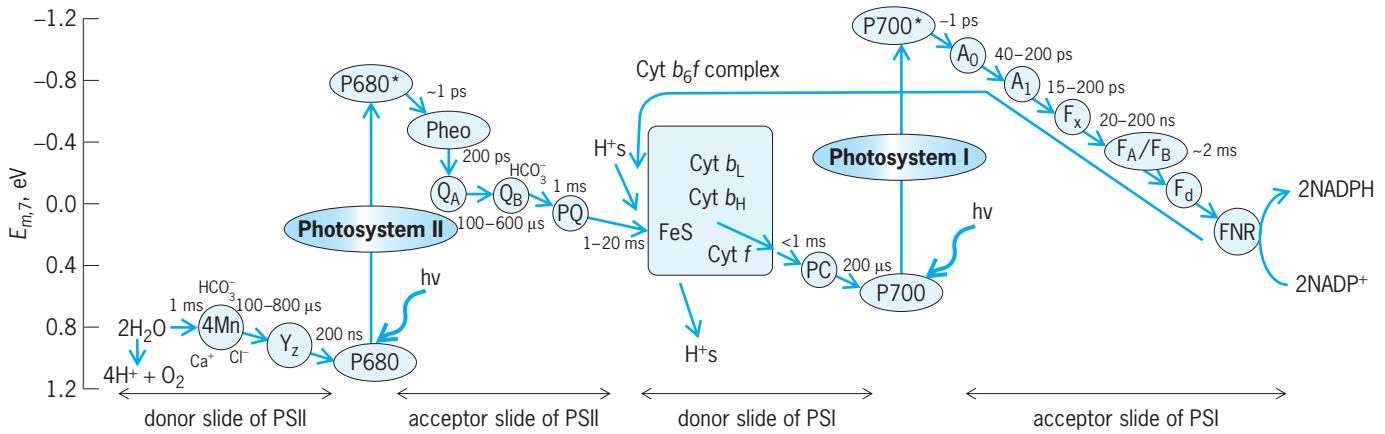

Fig. 4. Z scheme of photosynthesis, representating electron flow and chemical reactions in the two light steps in photosynthesis. Y_z (a tyrosine residue, also referred to as Z) is the electron donor to the oxidized chlorophyll *a* P680 of PSII, and Mn (manganese-containing protein) is the charge accumulator that leads to O_2 evolution. Pheo (pheophytin) and Chl (a A_0 chlorophyll *a*) are the primary electron acceptors of PSII and PSI, respectively. Q_A and F_X (an iron-sulfur center) are the stable electron acceptors of PSII and PSI, respectively. Q_B is another bound plastoquinone, while PQ is plastoquinone, Cyt is cytochrome, and PC is plastocyanin. A_1 is an electron acceptor of PSI, F_B and F_A are iron-sulfur centers, Fd is ferredoxin, and FNR is ferredoxin $NADP^+$ reductase. E_m is the reduction-oxidation (redox) potential.

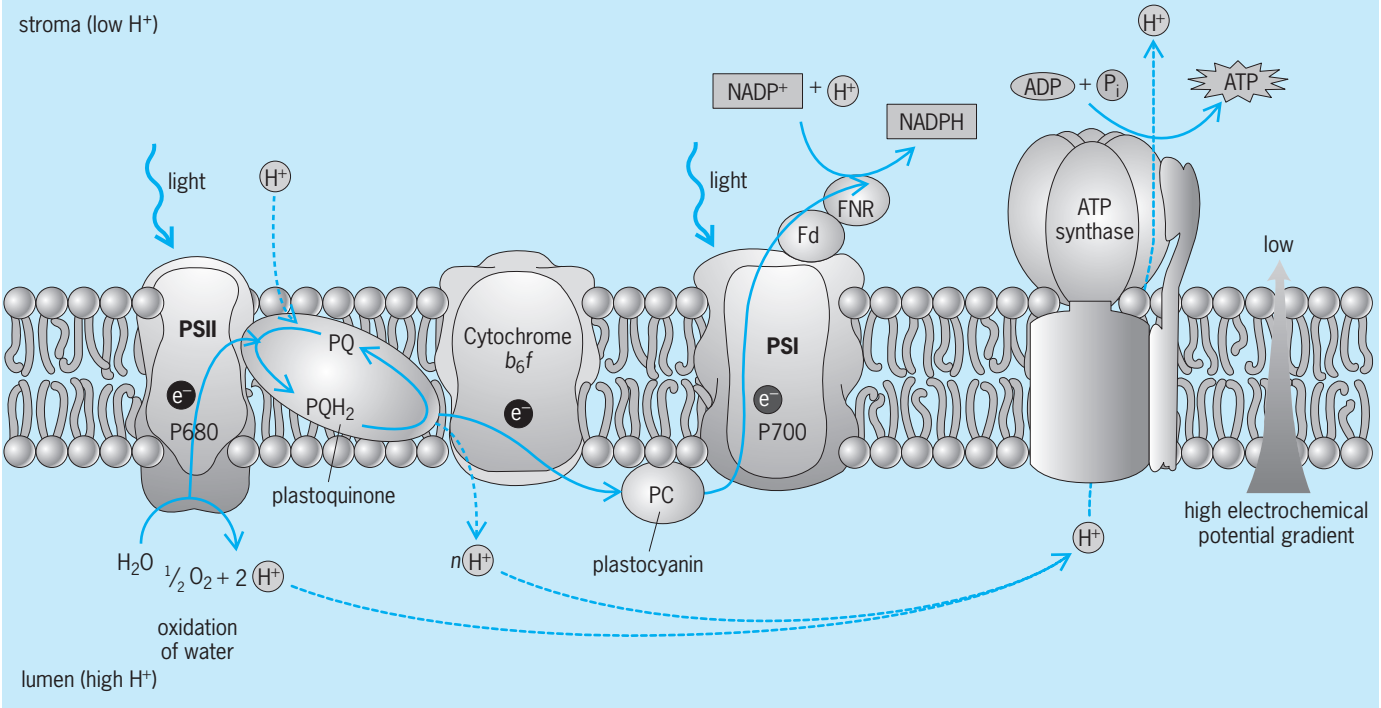

Fig. 5. A diagram of the four major protein complexes of the thylakoid membrane.

midpoint redox potential E_m, is shown on the y axis and the progress of the reaction is shown on the x axis. The two vertical arrows represent energy input to the system due to photon absorption.

The structures of all four of the large protein complexes are now known due to x-ray crystallography studies. The combination of the structural pictures in Fig. 3 and **Fig. 5** and the energetic picture in Fig. 4 provides an understanding of the process that neither diagram individually can portray.

Chemical role of chlorophyll. How does the chlorophyll *a* molecule, after excitation by the absorbed quantum of energy, utilize it for an energy-storing photochemical process, such as the transfer of an electron from a reluctant donor, H_2O, to a reluctant acceptor (perhaps $NADP^+$)? The chemical properties of chlorophyll are very different in the excited state from the unexcited or ground state. The excited chlorophyll acts as a very strong reducing agent, so that it easily loses an electron to a nearby acceptor. This is the primary photochemical process in photosynthesis. Support for this concept is provided by observations of reversible photochemical oxidation and of reversible photochemical reduction of chlorophyll in solution, and the comparison of these data with those in more intact systems. Studies of changes in the absorption spectrum of photosynthetic cells in light show that a small fraction of a special form of chlorophyll *a* (including P700), absorbing maximally at 700 and 430 nm, is in an oxidized state during illumination. This is at the heart of the reaction center of PS I. The reaction center Chl's of PS II (including P680) have been suggested also to undergo oxidation-reduction. Furthermore, a chlorophyll *a* molecule appears to be chemically reduced when P700 is oxidized to $P700^+$ in PSI. P700 has been shown to be a pair of Chl *a* and Chl *a*′ molecules. In the above, the P stands for pigment and the numerical designation gives the location, in nanometers, of the maximum decrease in the absorption in the red or the infrared region of the spectrum.

Detailed reactions of PSI and II. The PSII reaction is the one most closely associated with O_2 evolution. The final result of this set of reactions is the oxidation of water to O_2 and the reduction of a plastoquinone (an oxidation-reduction catalyst). Current evidence suggests that light absorbed by the major part of the accessory pigments is ultimately transferred to a special chlorophyll *a* molecule in the PSII reaction center, which is in a favorable position to act as an energy trap. There is a charge separation within 3 picoseconds; this special chlorophyll molecule is oxidized, and a specific pheophytin (Pheo) molecule is reduced (Chl^+ $Pheo^-$). The positive charge is transferred again within a few picoseconds to a pair of special Chl's nearby (P680). The P680 recovers by accepting an electron from a tyrosine amino acid residue that is represented by Y_Z (referred to as Z below), within 20–400 nanoseconds, as shown in reactions (2).

$$\mathrm{Chl} + h\nu_{\parallel} \rightarrow \mathrm{Chl}^* \qquad (2a)$$

$$\mathrm{Chl}^* + \mathrm{Pheo} \rightarrow \mathrm{Chl}^+ + \mathrm{Pheo}^- \qquad (2b)$$

$$\mathrm{Chl}^+\mathrm{P680} \rightarrow \mathrm{P680}^+ + \mathrm{Chl} \qquad (2c)$$

$$\mathrm{P680}^+ + \mathrm{Z} \rightarrow \mathrm{P680} + \mathrm{Z}^+ \qquad (2d)$$

The components of reactions (2), that is, the special chlorophyll, P680, pheophytin, and Z, and

the next intermediates, Q_A and Q_B (bound plastoquinones), are located on a protein complex consisting of two polypeptides, D-1 and D-2 (approximately 32,000 molecular weight each). The oxidation product, the strong oxidant Z^+, is utilized to oxidize water and liberate O_2 and protons (Figs. 4 and 5).

The PS I reaction is the one most closely associated with the reduction of $NADP^+$. Light absorbed by most of the chlorophyll *a* molecules is transferred to a special chlorophyll *a* molecule P700, where the primary charge separation occurs: P700 is oxidized, and another special Chl *a* (A_0) is reduced in about a picosecond, as shown in reactions (3).

$$\text{P700} + h\nu_l \rightarrow \text{P700}^* \qquad (3a)$$

$$\text{P700}^* + \text{Chl} \rightarrow \text{P700}^+ + \text{Chl}^- \qquad (3b)$$

After the primary photochemical processes, several additional steps take place. In PSII, the electrons flow on the electron donor side from H_2O to oxidized tyrosine Z^+ via a manganese (Mn) complex. On the electron acceptor side, electrons flow from $Pheo^-$ to two special plastoquinones called Q_A and Q_B. A bound bicarbonate ion is essential for the electron and proton transfer from photosystem II to the plastoquinone pool. Additional reactions and cofactors connect the two photosystems, including a pool of PQ, the cytochrome b_6f complex which contains cytochrome b_6, cytochrome *f* and the Rieske iron-sulfur protein, and the soluble copper protein plastocyanin. Plastocyanin finally rereduces $P700^+$ in PSI. On the electron acceptor side of PSI, the electron that is lost from P700 flows from Chl a^- (A_0^-), through an intermediate A_1 (phylloquinone), the iron-sulfur centers (F_X, F_A, F_B), and soluble ferredoxin (Fd) and a flavoprotein called ferredoxin-NADP-reductase, eventually to $NADP^+$, reducing it (Figs. 4 and 5).

Mechanism of O_2 evolution. The mechanism of O_2 evolution is the least known part of photosynthesis. All oxygen liberated in photosynthesis originates in water. Based on the measurements, by Pierre Joliot and Bessel Kok, of the amount of O_2 evolved in single brief (10-μs) saturating light flashes, it has been shown that four oxidizing equivalents must accumulate on the O_2-evolving complex before it can oxidize H_2O to O_2. There is a period 4 oscillation in the pattern of O_2/flash as a function of flash number (**Fig. 6***a*, left). The system is an oxygen "clock" in which the state (S) of the O_2-evolving complex undergoes sequential reactions, and the subscripts on S represent the oxidizing equivalents accumulated on the complex; the system starts in the dark in S_1 state (Fig. 6*b*, right).

Manganese (Mn) is required for O_2 evolution. Plants grown in a manganese-deficient medium lose their capacity to evolve O_2. The charge accumulator of the reaction mechanism is a manganese complex; manganese is known to exist in several valence states. In addition to manganese, chloride and calcium ions also function at the O_2-evolving site, though the mechanism of action is not known. The O_2-evolving complex includes at least four "intrinsic" proteins having molecular weights of 34,000, 32,000, 9000, and 4500 (D- 1, D-2, and two subunits of cytochrome *b*-559). In addition, three "extrinsic" polypeptides of molecular weights 33,000, 24,000, and 18,000 are also involved. In cyanobacteria, the 24,000 and the 18,000 molecular weight proteins are replaced by a cytochrome *c* and another low molecular weight protein. Further research is needed to understand the detailed chemical mechanism of O_2 evolution.

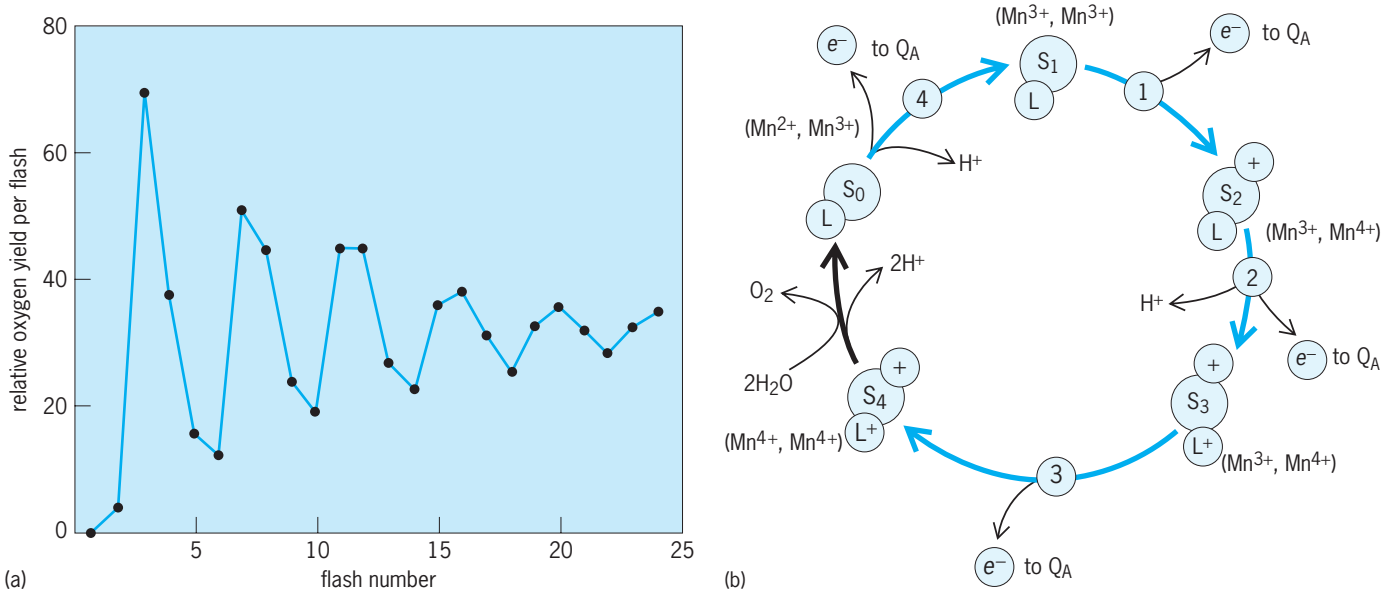

Fig. 6. Oxygen clock. (*a*) Pattern of oxygen evolution from dark-adapted chloroplasts when illuminated by a series of intense brief flashes. (*b*) S-state mechanism to describe oxygen evolution pattern. The oxygen-evolving complex can exist in five distinct states, S_0–S_4. Light activates PSII, which in turn advances the system through the states one step for each photochemical turnover of the PSII reaction center. When state S_4 is reached, oxygen is produced and the clock is reset to S_0.

Photophosphorylation. Inner cytoplasmic membranes from photosynthetic bacteria, cyanobacterial cells, and chloroplasts from green plants and algae—when illuminated in the presence of adenosine diphosphate (ADP) and inorganic phosphate (P_i)—use light energy to synthesize adenosine triphosphate (ATP). About 42 kJ of converted light energy in this reaction is stored in each mole of the high-energy phosphate, ATP. This photophosphorylation is coupled with energy-releasing steps in photosynthesis, such as the electron flow from PSII to PSI. When phosphorylation is associated with noncyclic electron flow from H_2O to $NADP^+$, it is called noncyclic photophosphorylation. *See* ADENOSINE DIPHOSPHATE (ADP); ADENOSINE TRIPHOSPHATE (ATP).

Under certain conditions, electrons in PSI, instead of going to $NADP^+$, may return to an intermediate (such as a cytochrome, plastoquinone, plastocyanin, or P700) and thus close the cycle. This type of cyclic electron flow, mediated by added cofactors and ADP and inorganic phosphate, also leads to the production of ATP and has been termed cyclic phosphorylation; it has been shown to exist under certain experimental conditions in vivo.

The mechanism of ATP formation is very different from the electron transfer processes described earlier. Light produces a high-energy state, and then the actual phosphorylation occurs in the dark. André Jagendorf and coworkers found that if chloroplasts are first suspended in an acidic medium and then transferred to an alkaline medium in the presence of ADP and P_i, ATP formation occurs without the need of light. All of these experiments were explained in terms of a hypothesis proposed by Peter Mitchell, in which light produces a H^+ ion gradient (ΔpH) across the lamellar membranes in the chloroplasts, and the energy dissipation of this H^+ ion gradient via the ATP synthase (sometimes called the coupling factor, a protein complex present in the lamellae) leads to phosphorylation. In addition, an electric field ($\Delta\Psi$) generated across the thylakoid membrane as a result of the initial light-induced charge separation can also drive photophosphorylation. Reagents that dissipate the electric field or the H^+ gradient also inhibit phosphorylation. The two together are referred to as proton motive force.

The ATP synthase is shown schematically in Fig. 5, and its structure is shown in Fig. 3. It has two parts, a section that is embedded in the thylakoid membrane and a section that protrudes into the stroma. Recent evidence demonstrates that the ATP synthase is a rotary motor, and most of the membrane-associated part rotates, while H^+'s are translocated across the membrane. Mechanical energy in this rotary motion is coupled to the formation of high-energy phosphate bonds in ATP. The chemical mechanism of this process is not yet well understood.

Photosynthetic unit. The concentration of the special chlorophyll *a* molecules (P700 or P680) that engage in the chemical reactions is very low, only one in several hundred chlorophyll molecules; energy absorbed by other pigments is effectively transferred to the reaction centers. The collection of antenna molecules along with their associated photosystems are often referred to as photosynthetic units.

Robert Emerson and William Arnold showed how the light reaction in photosynthesis can be separated from the dark reaction by the use of brief, intense light flashes, separated by intervals of darkness of variable duration. They found that the yield of a single flash was maximum when the interval between the consecutive light flashes was at least 0.04 s at 1°C (33°F). This, then, is the minimum time required for the efficient utilization of the products from the light reactions. Emerson and Arnold further observed that, under optimal conditions, the maximum yield from a single flash was one O_2 molecule per 2400 chlorophyll molecules present. However, it is now known that the two light-reaction mechanism of photosynthesis requires the transfer of four electrons twice for every molecule of oxygen evolved. Thus there are at least eight photoacts leading to the evolution of one O_2 molecule. Therefore, the ratio of one O_2 per 2400 chlorophylls means one photoact per 300 chlorophyll molecules. By using spinach grown in moderate to high light intensity, we know that there is one active PSII (P680) and one active PSI (P700) for a total of about 600 chlorophyll molecules present. If these chlorophyll molecules are equally divided in PSI and II, there is one reaction center per 300 chlorophyll molecules in each system. This is the commonly accepted size of one physical unit in higher plants, algae, and cyanobacteria.

Accessory antenna pigments. In addition to chlorophyll *a* (the one pigment present in nearly all oxygenic photosynthetic organisms), there are other chlorophylls, such as chlorophyll *b* in the green algae and higher plants. In brown algae, chlorophyll *c* replaces chlorophyll *b*. In a marine cyanobacterium *Acaryochloris marina*, most of Chl *a* is replaced by Chl *d*. There are also nonchlorophyllous pigments belonging to two groups: (1) the carotenoids, so called because of similarity to the orange pigment of carrots, are a variable assortment of pigments found in all photosynthetic higher plants and in algae. (2) The phycobilins, or vegetable bile pigments, are chemically related to animal bile pigments. The phycobilins are either red (phycoerythrin) or blue (phycocyanin). Both types are present in special granules called phycobilisomes in red algae (Rhodophyta) and cyanobacteria: the red pigment in red algae, and the blue pigment in cyanobacteria. Another phycocyanin called allophycocyanin is also present in cyanobacteria. All of these pigments are associated with specific proteins to form so-called antenna pigment proteins. The structures of many of these antenna complexes are known, and the energy transfer processes have been investigated using ultrafast laser techniques. *See* CAROTENOID; CHLOROPHYLL; PHYCOBILIN.

Light absorbed by accessory pigments does contribute to energy storage in photosynthesis. This is simply demonstrated from measurements of the so-called action spectra of photosynthesis. In

such measurements, photosynthesis is excited by monochromatic light, and the production of oxygen per incident quantum of light is measured as a function of wavelength. The observed spectral variations in the yield of photosynthesis can be related to the proportion of light absorbed at each wavelength by the different pigments in the cells. Measurements of this kind have led to the conclusion that quanta absorbed by most carotenoids are 50–80% as effective as those absorbed by chlorophyll *a* in contributing energy to photosynthesis. An exception is fucoxanthol, the carotenoid that accounts for the color of brown algae (Phaeophyta) and that of the diatoms; it supplies light energy to photosynthesis about as effectively as chlorophyll *a*. The red and blue pigments of the Rhodophyta and cyanobacteria are also highly effective. They can be as effective as chlorophyll or somewhat less, depending, among other things, on the physiological status of the organism and the color of the light to which they have become adapted. The primary function of all these pigments is to harvest the light energy and transfer it to reaction-center chlorophyll molecules. However, in addition, several xanthophylls (violaxanthin, antheraxanthin, and zeaxanthin) and lutein are involved in photoprotecting photosynthetic organisms against excess light. In many cases, excess light energy is lost as "heat" via deexcitation of chlorophyll directly or via transfer to zeaxanthin.

Energy transfer between pigment molecules in antenna system. Chlorophyll *a* in vivo is weakly fluorescent, that is, some of the light quanta absorbed by it (up to 6%) are reemitted as light (**Fig. 7**, white curve). Observations of the action spectrum of chlorophyll *a* fluorescence in different oxygenic organisms closely parallel the action spectrum of photosynthesis. In other words, fluorescence of chlorophyll *a* is excited also by light absorbed by the accessory pigments. Excitation of chlorophyll *a* fluorescence by light quanta absorbed by phycoerythrin requires transfer of the excitation energy from the excited phycoerythrin molecule to a nearby chlorophyll molecule (somewhat as in acoustic resonance, where striking one bell causes a nearby bell to ring). Therefore, light quanta absorbed by accessory pigments, such as carotenoids and phycobilins, contribute to photosynthesis by being transferred to chlorophyll *a*. By this mechanism, red algae, growing relatively deep under the sea where only green light penetrates, can supply the energy of this light to chlorophyll *a*, which has very weak absorption in the green region of the spectrum.

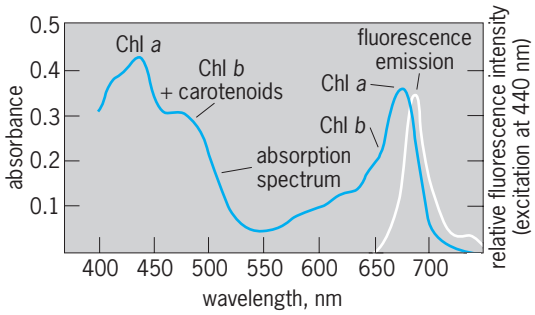

Fig. 7. Absorption spectrum of a maize (*Zea mays*) chloroplast suspension. Pigments responsible for specific bands are shown. Also shown is the fluorescence emission of chloroplasts from a maize chloroplast.

Excitation energy is transferred efficiently in the chloroplasts from accessory pigments to chlorophyll *a*. A similiar transfer (often referred to as energy migration) occurs also between different chlorophyll *a* molecules themselves. Excitation-energy transfer among chlorophyll *a* molecules or among phycobilin molecules, and excitation-energy transfer from accessory pigments (donor molecules) to chlorophyll *a* (acceptor molecules) or from various short-wavelength forms of chlorophyll *a* to the long-wavelength forms of chlorophyll *a*, has been demonstrated. The most widely accepted hypothesis, Förster's hypothesis, is that energy transfer is preceded by thermal relaxation in the donor molecules. The efficiency of energy transfer depends upon three basic factors: orientation of acceptor molecules with respect to the donor molecule; overlap of the fluorescence spectrum of the donor molecule with the absorption spectrum of the acceptor molecule; and the distance between the two molecules. The function of most of the pigments (including most of the chlorophyll *a* molecules) is to act as an antenna, harvest the energy, and transfer to very few (1 in 300) reaction-center Chl molecules, depending upon the pigment system. Energy is thus trapped and used for photochemistry. *See* CHLOROPHYLL; PLANT PIGMENT.

[Contributions of Rajni Govindjee to this article are acknowledged.] Robert E. Blankenship; Govindjee

Carbon Dioxide Fixation

The light-dependent conversion of radiant energy into chemical energy as adenosine triphosphate (ATP) and reduced nicotinamide adenine dinucleotide phosphate (NADPH) serves as a prelude to the utilization of these compounds for the reductive fixation of CO_2 into organic molecules. Such molecules, broadly designated as photosynthates, are usually but not invariably in the form of carbohydrates such as glucose polymers or sucrose, and form the base for the nutrition of all living things, as well as serving as the starting material for fuel, fiber, animal feed, oil, and other compounds used by people. Collectively, the biochemical processes by which CO_2 is assimilated into organic molecules are known as the photosynthetic dark reactions, not because they must occur in darkness but because, in contrast to the photosynthetic light reactions, light is not required. (We do recognize, however, that several enzymes need to be light-activated before they can function; see "Regulation of C_3 cycle enzymes" below.) CO_2 fixation by photosynthetic organisms is an important mechanism by which this "greenhouse" gaseous molecule is removed from the atmosphere during carbon cycling on Earth. Approximately 100 pentagrams of carbon (1 pentagram equals 10^9 metric tons) as CO_2 is assimilated annually into organic molecules by photosynthesis (about

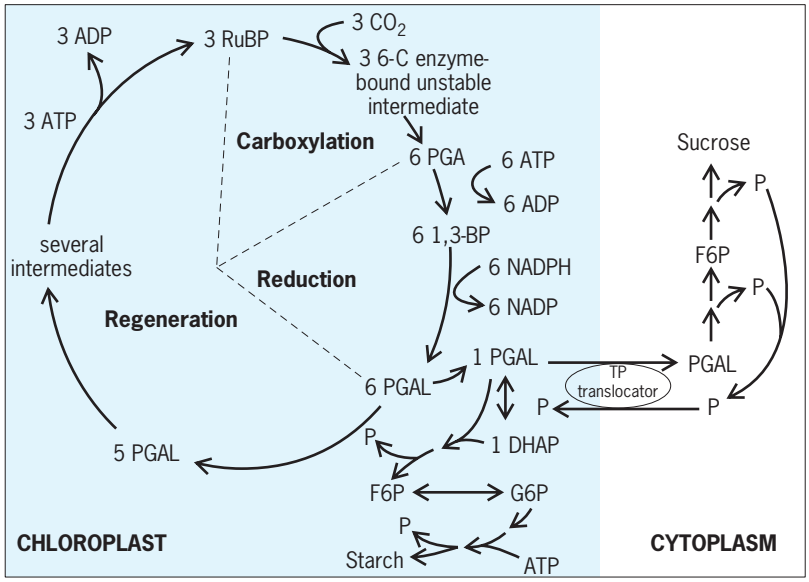

Fig. 8. Schematic outline of the C_3 (Calvin-Benson-Bassham) CO_2 assimilation cycle (the three phases are noted), showing the partitioning of assimilated carbon into starch within the chloroplast [via the phosphorylated 6-C intermediates, fructose-6-phosphate (F6P) and glucose-6-phosphate (G6P)], and the assimilate efflux (through the three-carbon triose phosphate [TP] transporter) across the chloroplast inner envelope to the cytoplasm (in exchange for inorganic phosphate) leading to sucrose synthesis. For the reactions of the C_3 cycle, shown in the chloroplast, the relative numbers of molecules involved in each specific step are shown to the left of the substrate and/or product.

half of this amount is assimilated by photosynthetic marine algae).

The route by which CO_2 is assimilated had been studied for over a century when it was discovered that photosynthesis leads to the accumulation of sugars and starch. Details of the biochemical pathway leading to CO_2 assimilation were worked out in the 1950s, when the availability of paper chromatographic techniques and $^{14}CO_2$ allowed Melvin Calvin, Andrew A. Benson, and James A. Bassham to develop the outline of the reductive pentose phosphate cycle (Calvin was awarded a Nobel Prize in Chemistry in 1961), now usually called the C_3 cycle. The name C_3 cycle refers to the first stable product generated from CO_2, which is a three-carbon organic molecule. The C_3 cycle forms the primary, or basic (with other, feeder pathways occurring in some plant types), route for the formation of photosynthate from CO_2.

C_3 photosynthesis. The essential details of C_3 photosynthesis can be seen in **Fig. 8**. The entire cycle can be separated into three phases—carboxylation, reduction, and regeneration. Since the smallest intermediate in the cycle consists of three carbons, we will start with three molecules of CO_2. During the initial carboxylation phase, these three molecules of CO_2 are combined with three molecules of the five-carbon compound ribulose 1,5-bisphosphate (RuBP) in a reaction catalyzed by the enzyme RuBP carboxylase/oxygenase (rubisco) to form three molecules of an intermediate, unstable enzyme-bound six-carbon compound. These unstable molecules are hydrolyzed by the enzyme into six molecules of the three-carbon compound phosphoglyceric acid (PGA). These products of the carboxylation phase, the six (three-carbon) PGA molecules, are phosphorylated by six molecules of ATP (releasing ADP to be used for photophosphorylation via the light reactions) to form six 1,3-bisphosphoglycerate (1,3-BP) molecules. The resulting compounds are reduced (that is, in the reduction phase of the C_3 cycle) by the NADPH formed in photosynthetic light reactions to form six molecules of the three-carbon compound phosphoglyceraldehyde (PGAL). PGAL is isomerized to form another three-carbon compound, dihydroxyacetone phosphate (DHAP). PGAL, the aldehyde, and DHAP, the ketone, are energetically equivalent, reduced compounds and can be considered the products of the reductive phase of the C_3 photosynthetic cycle. PGAL and DHAP together form the triose phosphate (TP) pool of the chloroplast. The chloroplast TP pool is primarily composed of PGAL; the isomerase responsible for PGAL:DHAP interconversion favors PGAL formation.

The rest of the C_3 photosynthetic cycle (the regeneration phase) involves enzymatic steps that allow regeneration of RuBP, the initial carboxylation substrate. One molecule of PGAL is made available for combination with DHAP isomerized from a second PGAL (requiring a second "turn" of the Calvin-Benson-Bassham cycle wheel) to form a six-carbon sugar. The other five PGAL molecules, through a complex series of enzymatic reactions, are rearranged into three molecules of RuBP, which can again be carboxylated with CO_2 to continue the cycle.

It should be noted that the enzyme that incorporates CO_2 into an organic compound, RuBP carboxylase/oxygenase (rubisco), also allows oxygen (O_2) to react with RuBP, hence the "oxygenase" in the name. This reaction initiates the process called photorespiration, which results in the release of one previously incorporated molecule of CO_2 for every two molecules of O_2 that are allowed to react. *See* PHOTORESPIRATION.

Due to its low catalytic efficiency, rubisco (**Fig. 9**) can be up to half of the soluble protein in C_3 chloroplasts, and most likely it is the most abundant protein found in nature. Rubisco is a large and

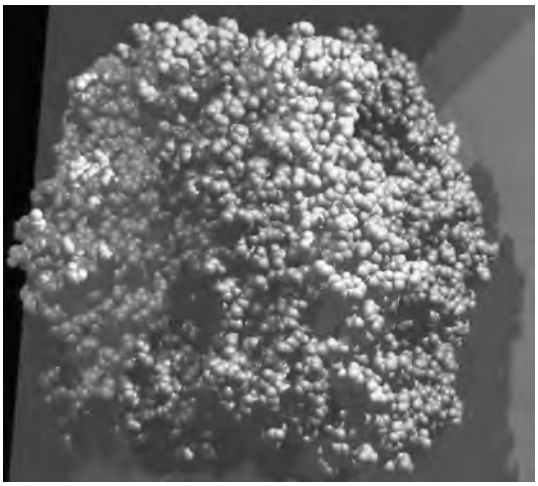

Fig. 9. Structural model of Rubisco.

complex enzyme, comprising eight large polypeptide subunits and eight small subunits. Interestingly, the small subunit polypeptide is produced (as a larger precursor form) in the cytoplasm from mRNA which is encoded in the nucleus. The precursor polypeptide is then transported across the chloroplast membrane (the mature form of this polypeptide cannot be transported in this manner); processed into the shorter, mature polypeptide; and then combined with large subunits (encoded in the chloroplast DNA and produced in the stroma) to form the mature enzyme.

The net product of two "turns" of the cycle, a six-carbon sugar (G6P or F6P), is formed either within the chloroplast in a pathway leading to starch (a polymer of many glucose molecules) or externally in the cytoplasm in a pathway leading to sucrose (condensed from two six-carbon sugars, glucose and fructose). This partitioning of newly formed photosynthate leads to two distinct pools; starch is stored in the photosynthesizing "source" leaf cells, and sucrose is available either for immediate metabolic requirements within the cell or for export to "sinks" such as developing reproductive structures, roots, or other leaves. Factors within the photosynthesizing cell, such as energy requirements in different compartments (mitochondria, cytoplasm, and chloroplasts), along with energy needs of the plant (such as increased sink requirements during different developmental stages) and external, environmental factors (such as light intensity and duration) ultimately regulate the partitioning of newly formed photosynthetic product (PGAL) into starch or sucrose. *See* PLANT METABOLISM.

This profound control of photosynthate partitioning is accomplished through regulation of PGAL export from the chloroplast to the cytoplasm, as well as by regulation of the enzymes that convert PGAL to sucrose in the cytoplasm and starch in the chloroplast. Under conditions where sink demand is low (and sucrose is not transported through the phloem away from source leaf cells), metabolic effectors accumulate in the cytoplasm that lower the activities of the sucrose-forming enzymes and increase the activities of the starch-forming enzymes. This results in a condition that reduces PGAL export from the chloroplast, and hence more PGAL is retained in the chloroplast for starch formation. Also, under conditions which cause low chloroplast PGAL levels (such as low light), PGAL transport out of the chloroplast is restricted, resulting in decreased substrate for sucrose formation, increasing the relative amount of starch production. The energy status of the cell affects sucrose formation (and therefore photosynthate partitioning) because cytoplasmic uridine triphosphate (used in the formation of sucrose) level is dependent on ATP generation, and also because PGAL export to the cytoplasm is coupled obligatorily to inorganic phosphate (formed when ATP is metabolized in the cytoplasm) import into the chloroplast.

In addition to providing carbon skeletons for starch and sucrose synthesis, PGAL is fed back into the C_3 cycle to allow for the regenerative phase of the reactions to synthesize more RuBP, the carboxylation substrate. We have known for quite some time that the maximal measurable activities of fructose 1,6-bisphophatase (FBPase) and sedoheptulose 1,7-bisphosphatase (SBPase), two enzymes involved in RuBP regeneration, are not much greater than the rate of photosynthetic carbon assimilation and concomitant carbon flow through the C_3 cycle. Thus, carbon flow through these enzymes might contribute to rate limitation of photosynthetic carbon assimilation. Studies with transgenic plants overexpressing these enzymes support this contention; increasing the amount of either enzyme led to a higher level of the carboxylation substrate RuBP as well as higher photosynthetic rates.

The autocatalytic nature of the cycle (that is, more substrate for initial carboxylation can be generated as carbon flows through the steps of the pathway) can be best understood by considering that the net product of one "turn" of the cycle (representing three carboxylations), that is, a PGAL molecule, can be fed back into the cycle. Thus the rate of carboxylation during an initial lag phase (as chloroplasts are initially illuminated) is dependent on the level of newly formed RuBP. If all newly fixed carbon were fed back into the cycle, the level of RuBP would double after five carboxylations. Since the rate of photosynthetic carbon fixation is initially dependent on the level of intermediates such as the substrate (RuBP) for the carboxylation reaction, the next five carboxylations would occur in a shorter amount of time, resulting in an exponential increase in the rate of photosynthesis until factors other than intermediate levels become limiting.

Regulation of C_3 cycle enzymes. The photosynthetic carbon assimilation cycle is regulated at a number of enzymatic steps. The initial carboxylation catalyst, rubisco, as well as some of the enzymes involved in the regeneration phase, including glyceraldehyde-3-phosphate dehydrogenase, phosphoribulose kinase, SBPase, and FBPase, require activation. These enzymes are inactivated in the dark and activated in the light. Several conditions are required for activation, including high concentrations of Mg^{2+}, high pH, and a reductant (supplied in the chloroplast by the enzyme thioredoxin). Thioredoxin is reduced by NADPH generated in the light. Thioredoxin acts as a protein disulfide oxidoreductase, converting disulfide (S-S) bonds of the target proteins (all of the above enzymes except rubisco) to a reduced (−SH) form. Rubisco activity is also modulated by a specific mechanism involving another enzyme, called rubisco activase. All of the aforementioned activating conditions within the chloroplast stroma are facilitated by light-dependent processes but are reversed in darkness. This regulatory mechanism conveniently allows for the synthesis pathway to be "shut off," preventing a futile cycle during the night, when starch reserves are mobilized to meet cell energy requirements via intermediates which, if C_3 cycle enzymes were activated, would be reconverted to starch.

C_4 photosynthesis. Initially, the C_3 cycle was thought to be the only route for CO_2 assimilation,

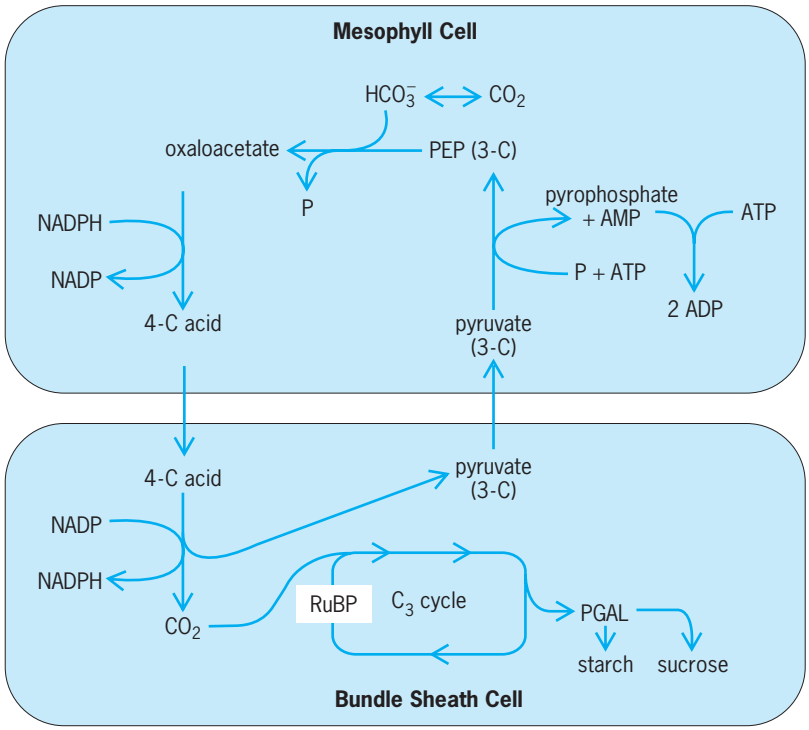

Fig. 10. Schematic outline of the C_4 carbon dioxide assimilation process in the two cell types of a NADP-ME-type plant.

although it was recognized by plant anatomists that some rapidly growing plants (such as maize, sugarcane, and sorghum) possessed an unusual organization of the photosynthetic tissues in their leaves (Kranz morphology). Work by Hugo Kortschak, Constance Hartt, and colleagues in Hawaii as well as that of M. D. (Hal) Hatch and Roger Slack in Australia demonstrated that plants having the Kranz anatomy utilized an additional CO_2 assimilation route now known as the C_4-dicarboxylic acid pathway (**Fig 10**). Carbon dioxide enters a mesophyll cell where it is combined (in the form of bicarbonate) with the three-carbon compound phosphoenolpyruvate (PEP) via the enzyme PEP carboxylase to form a four-carbon acid, oxaloacetate, which is reduced to malic acid or transaminated to aspartic acid. The four-carbon acid moves into bundle sheath cells where the acid is decarboxylated and the CO_2 reassimilated via the C_3 cycle. The resulting three-carbon compound, pyruvic acid, moves back into the mesophyll cell and is transformed into PEP (at the cost of 2 ATP molecules) via the enzyme pyruvate phosphate dikinase located in the mesophyll chloroplasts to complete the cycle. The net effect of this cycle is to increase the CO_2 concentration around rubisco, thereby reducing photorespiration via the competing oxygenase activity of this enzyme.

As depicted in Fig. 10, extensive transport of metabolites must occur between the two cell types that are found in most C_4 plants. The diffusion of metabolites between two cell types is facilitated by the presence of plasmodesmata connecting the cells to form a cytoplasmic continuum. However, in some cases the two cell types, mesophyll and bundle sheath, are not necessarily adjacent (sedges are an example). Exotic plants have now been found that perform C_4 photosynthesis in single cells in which the chloroplasts that initially combine carbon dioxide with PEP are spatially separated from the chloroplasts where the carbon dioxide is reassimilated via the C_3 cycle.

C_4 metabolism is classified into three types, depending on the primary decarboxylation reaction used with the four-carbon acid in the bundle sheath cells. The majority of C_4 species (exemplified by sugarcane, maize, crabgrass, and sorghum) are of type 1 (see below), and employ NADP-malic enzyme (NADP-ME) for decarboxylation. NAD-malic enzyme (NAD-ME) C_4 plants (type 2) include amaranthus, atriplex, millet, pigweed, and purslane. Type 3 C_4 types use phosphoenol pyruvate carboxykinase (PCK) for decarboxylation and include Panicum grasses. The decarboxylases are also located in different intracellular compartments as indicated:

1. NADP-ME type,

$$\mathrm{NADP^+ + malic\ acid \xrightarrow{NADP\text{-}malic\ enzyme\ (chloroplasts)} pyruvic\ acid + CO_2 + NADPH} \quad (6)$$

2. NAD-ME type,

$$\mathrm{NAD^+ + malic\ acid \xrightarrow{NAD\text{-}malic\ enzyme\ (mitochondria)} pyruvic\ acid + CO_2 + NADH} \quad (7)$$

3. PCK type,

$$\mathrm{Oxaloacetic\ acid + ATP \xrightarrow{phosphoenolpyruvate\ carboxykinase\ (cytosol)} PEP + CO_2 + ADP} \quad (8)$$

In addition to differing decarboxylation reactions, the particulars of the CO_2 fixation pathway in NAD-ME and PCK plant types differ from those depicted in Fig. 10 with respect to the three-carbon compound transported from bundle sheath to mesophyll cells. With NAD-ME types, the three-carbon compound can be either pyruvic acid or alanine, and in PCK types this compound is PEP. Therefore, the three variations in the C_4 pathway necessarily predicate different energy (ATP and NADPH) usage in the two cell types. The generation of ATP from ADP, and NADPH from NADP via noncyclic electron flow through photosystem I (PSI) and photosystem II (PSII), is tightly coupled: neither compound can be produced without sufficient substrate for both. Therefore, the different usage of ATP and NADPH in the mesophyll and bundle sheath chloroplasts of the three C_4 plant types (due both to variations in the pathway of carbon flow in the photosynthetic cycle and to variations in partitioning of portions of the pathway between cell types) is supported by variations in the photochemical apparatus which allow for differing ability to produce ATP without concomitant NADPH production. These alternative pathways of ATP production (which result in different ratios of ATP:NADPH produced) are cyclic and pseudocyclic

photophosphorylation, with the cyclic pathway considered the major pathway of uncoupled ATP production in chloroplasts, and the pseudocyclic pathway possibly acting as a "fine-tuning" modulator.

Variations in the photochemical apparatus that indicate enhanced cyclic photophosphorylation capacity (utilizing only PSI) are a high chlorophyll *a/b* ratio, low Chl/P700 ratio, and a low PSII reaction. These characteristics are found in bundle sheath chloroplasts of NADP-ME-type plants, indicating that the primary function of the photochemical apparatus in these chloroplasts is the generation of ATP. NADPH is supplied via the decarboxylation of malic acid to support the C_3 cycle activity (PGA conversion to PGAL) in these chloroplasts. Assays of chlorophyll *a/b* ratio, Chl/P700 ratio, and PS II activity indicate that NAD-ME mesophyll chloroplasts also have a primary role of cyclic photophosphorylation, while NAD-ME bundle sheath chloroplasts have a primary role of noncyclic electron flow. In PCK-type plants, mesophyll chloroplasts appear to have a photochemical apparatus similar to C_3 chloroplasts, while bundle sheath chloroplasts appear to have a low PSII activity. The enhanced ability of PCK bundle sheath chloroplasts to produce ATP via cyclic photophosphorylation supplies the extra ATP needed to convert pyruvic acid to PEP. These variations in the C_4 pathway and photochemical apparatus among the C_4 plant types demonstrate the close relationship that has evolved between light reactions and the biochemical processes of carbon dioxide assimilation, and show the highly integrated cooperation between the cell types involved.

Benefits of C_4 cycle. The concentration of CO_2 in air is presently about 0.037% by volume (and increasing with time due to burning of fossil fuels), a concentration that does not fully saturate the C_3 cycle when it is operating at capacity due to the low affinity of rubisco for CO_2. It would be necessary to have about 0.1% CO_2 to saturate photosynthesis in C_3 plants, which can be achieved only under controlled conditions (CO_2-enriched greenhouses or growth chambers). Leaf photosynthesis in C_4 plants, however, is fully saturated at air CO_2 concentrations. Thus, C_4 photosynthesis may be considered to be an evolutionary adaptation to current-day CO_2 levels in air. During the C_4 cycle, CO_2 is rapidly captured via biochemical reactions in mesophyll chloroplasts and released near rubisco in bundle sheath chloroplasts. This serves to increase the CO_2 concentration around the enzyme, increasing its catalytic efficiency and decreasing its reaction with O_2 and thus photorespiration; thus the ambient concentration of CO_2 in air is not rate-limiting. The spatial compartmentalization of portions of CO_2 assimilation into the two cell types not only allows C_4 plants to assimilate air CO_2 rapidly with minimal photorespiration, but also partly explains other physiological characteristics and responses to the external environment of C_4 plants. C_4 plants have a higher efficiency of water use. Water vapor exits from leaves through the same stomatal pores through which CO_2 enters the leaf. Since the C_4 plant is more efficient at fixing CO_2 than C_3 plants, more CO_2 is incorporated per unit water lost. C_4 plants have a greater efficiency of nitrogen usage. Since rubisco is produced only in bundle sheath cells in C_4 plants, only 10–35% of the leaf nitrogen is tied up in this enzyme, as opposed to 40–60% in C_3 plants. Since C_4 plants have to "expend" less carbon on producing the protein rubisco, they have higher rates of sugar formation, which can facilitate the rapid growth rates seen in such C_4 plants as maize, sugarcane, sorghurn, and crabgrass. Other differences in response to the environment between C_3 and C_4 plants are as follows: C_4 plants exhibit a nonsaturating response curve of leaf photosynthesis to light levels found in nature. In addition, C_4 plants tolerate more salinity and higher temperatures than do C_3 plants. The higher energy requirements of C_4 plants (2 ATPs per CO_2 assimilated) are also reflected by the fact that quantum yields of photosynthesis for C_3 plants are higher than for those possessing the auxiliary C_4 system. At 2% oxygen partial pressure and 30°C, maximum quantum yield for C_3 plants is about 0.073 mole CO_2 assimilated per absorbed Einstein (an Einstein is a mole of photons) of light, while for C_4 plants the maximum quantum yield is 0.054. However, at normal O_2 partial pressures (21% O_2), quantum yields are almost identical. This is due to the presence of high photorespiration in C_3 plants, and thus represents a net quantum yield rather than a true photosynthetic yield. *See* PHOTORESPIRATION.

Crassulacean acid metabolism photosynthesis. Under arid and desert conditions, where soil water is in short supply, transpiration during the day when temperatures are high and humidity is low may rapidly deplete the plant of water, leading to desiccation and death. By keeping stomata closed during the day, water can be conserved, but the uptake of CO_2, which occurs entirely through the stomata, is prevented. Desert plants in the Crassulaceae, Cactaceae, Euphorbiaceae, and 15 other families have evolved, apparently independently of C_4 plants, a similar strategy of concentrating and assimilating CO_2 by which the CO_2 is taken in at night when the stomata open; water loss is low because of the reduced temperatures and correspondingly higher humidities. Although these succulent plants with thick, fleshy leaves were known since the nineteenth century as being unusual, the biochemical understanding of the process did not occur until the 1960s and 1970s when the details of C_4 photosynthesis were being worked out. It was first studied in plants of the Crassulaceae; thus, the process has been called crassulacean acid metabolism (CAM).

In contrast to C_4, where two cell types usually cooperate, the entire CAM process occurs within an individual cell; the separation of C_4 and C_3 is thus temporal rather than spatial. At night, CO_2 combines with PEP through the action of PEP carboxylase, resulting in the formation of oxaloacetic acid and its conversion into malic acid. The PEP is formed from starch or sugar via the glycolytic route

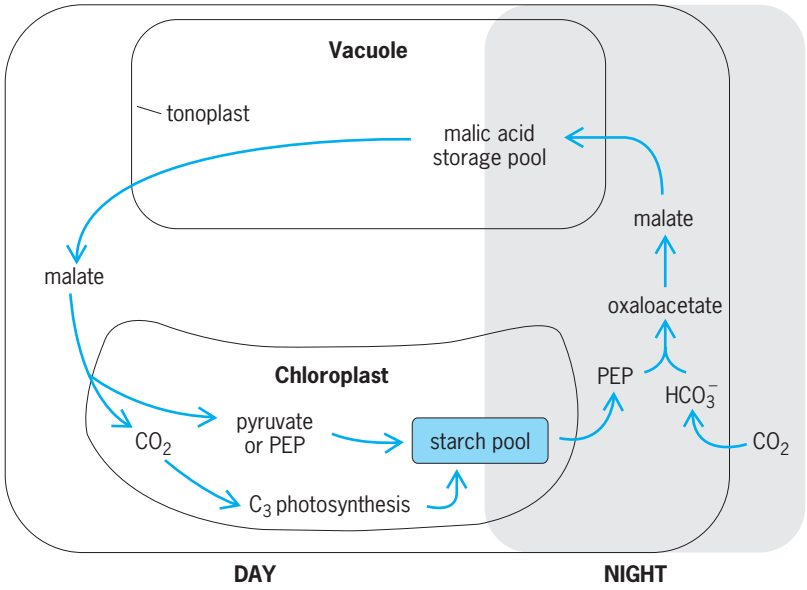

Fig. 11. Scheme for the flow of CO_2 within a single crassulacean acid metabolism (CAM) cell over a day, showing initial dark CO_2 fixation, malic acid storage in the vacuole at night, followed by decarboxylation and the C_3 cycle the next day.

of respiration. Thus, there is a daily reciprocal relationship between starch (a storage product of C_3 photosynthesis) and the accumulation of malic acid (the terminal product of nighttime CO_2 assimilation; **Fig. 11**).

As in C_4 plants, there may be variations in the decarboxylase that provides the CO_2 for assimilation via the C_3 cycle. In some CAM plants (such as pineapple) PCK is used, while in others (cactus) the decarboxylase is the NADP-malic enzyme (NADP-ME type). A few CAM species use NAD-ME for decarboxylation. Since the stomata are closed most of the day, decarboxylation of the stored malate (or oxaloacetate) results in an elevation of its concentration around rubisco. The **table** summarizes the major physiological differences between C_3, C_4, and CAM plants.

Other CO_2 assimilation mechanisms. Both the C_4 cycle and CAM involve the synthesis of oxaloacetic acid, which is also one of the intermediates in the tricarboxylic acid (TCA) cycle of respiration. In the late 1960s a light-driven reversal of the TCA cycle was discovered. This CO_2 fixation cycle, called the reductive carboxylic acid cycle, results in the net synthesis of pyruvic acid via the reversal of the three decarboxylation steps in the TCA cycle (pyruvic acid to acetyl coenzyme A, isocitric acid to α-ketoglutaric acid, and succinyl CoA to succinic acid). The pathway has been detected in some photosynthetic bacteria. *See* CITRIC ACID CYCLE.

In most photosynthetic bacteria, the C_3 cycle is functional despite some differences in detail. The green sulfur bacteria, however, carry out C_3 photosynthesis poorly or not at all. *Chlorobium thiosulfatophilum* (alternate name: *Chlorobium limicola*), lacking the key enzyme rubisco, utilizes a reductive carboxylic acid cycle in which reduced ferredoxin drives the TCA cycle in reverse, resulting in carboxylation reactions much like those of the reductive carboxylic acid cycle. Heterocysts of cyanobacteria do not have a functional C_3 cycle because, in contrast to the normal cells of these bacteria, the heterocyst cell (implicated in nitrogen fixation) lacks the key enzyme rubisco. Here, CO_2 fixation in heterocysts may occur through PEP carboxylase as in C_4 and CAM photosynthesis. Guard cells in

Some characteristics of the three major plant groups

Characteristics	C_3	C_4	CAM
Leaf anatomy in cross section	Diffuse distribution of organelles in mesophyll and palisade cells with less chloroplasts in bundle sheath cells if present	Layer of bundle sheath cells around vascular tissue with a high concentration of chloroplasts; layers of mesophyll cells around bundle sheath	Spongy, often lacking palisade cells; mesophyll cells have large vacuoles
Theoretical energy requirement for net CO_2 fixation (CO_2:ATP:NADPH)	1:3:2	1:5:2	1:6.5:2
Carboxylating enzyme	Rubisco	PEP carboxylase, then rubisco	Darkness: PEP carboxylase; light: mainly rubisco
CO_2 compensation concentration, ppm CO_2	30–70	0–10	0–5 in dark
Transpiration ratio, g H_2O/g dry weight increase	450–950	250–350	50–55
Maximum net photosynthetic rate, mg CO_2/(dm^2 leaf)(h)	15–40	40–80	1–4
Photosynthesis sensitive to high O_2	Yes	No	Yes
Photorespiration detectable	Yes	Only in bundle sheath	Difficult to detect
Leaf chlorophyll *a*/*b* ratio	2.8 ± 0.4	3.9 ± 0.6	2.5–3
Maximum growth rate, g dry wt/(dm^2 leaf)(day)	0.5–2	4–5	0.015–0.018
Optimum temperature for photosynthesis	15–25° C (59–77° F)	30–40° C (86–104° F)	About 35° C (95° F)

C_3 plants, which regulate the opening of stomatal pores for gas exchange in leaves, also lack rubisco and apparently use PEP carboxylase exclusively to fix CO_2.

Contributions of the late Martin Gibbs to this article are acknowledged.

Gerald A. Berkowitz; Archie R. Portis, Jr.; Govindjee

Bacterial Photosynthesis

Certain bacteria have the ability to perform photosynthesis. This was first noticed by Sergey Vinogradsky in 1889 and was later extensively investigated by Cornelis B. Van Niel, who gave a general equation for bacterial photosynthesis. This is shown in reaction (9).

$$2H_2A + CO_2 + \text{light} \xrightarrow[\text{enzymes}]{\text{bacteriochlorophyll}} \{CH_2O\} + 2A + H_2O \quad (9)$$

where A represents any one of a number of reductants, most commonly S (sulfur).

Photosynthetic bacteria cannot use water as the hydrogen donor and are incapable of evolving oxygen. They are therefore called anoxygenic photosynthetic bacteria. The prokaryotic cyanobacteria (formerly called blue-green algae) are excluded in this discussion of bacterial photosynthesis, since their photosynthetic system closely resembles that found in eukaryotic algae and higher plants discussed above. Anoxygenic photosynthetic bacteria can be classified in four major groups:

1. *Proteobacteria*. Two groups with somewhat different properties are known.

(A) Nonsulfur purple bacteria (Rhodospirillaceae). In these bacteria, H_2A is usually an organic H_2 donor, such as succinate or malate; however, these bacteria can be adapted to use hydrogen gas as the reductant. They require vitamins for their growth and usually grow anaerobically in light, but they can also grow aerobically in the dark by using respiration to utilize organic compounds from the environment. They are thus facultative photoheterotrophs. Examples of this group are *Rhodospirillum rubrum* and *Rhodobacter sphaeroides*.

(B) Sulfur purple bacteria (Chromatiaceae). These cannot grow aerobically, and H_2A is an inorganic sulfur compound, such as hydrogen sulfide, H_2S; the carbon source can be CO_2. These bacteria are called obligate photoautotrophic anaerobes. An example is *Chromatium vinosum* (alternate name: *Allochromatium vinosum*).

2. *Green sulfur bacteria* (Chlorobiaceae). These bacteria are capable of using the same chemicals as Chromatiaceae but, in addition, use other organic H_2 donors. They may then be called photoautotrophic and photoheterotrophic obligate anaerobes. An example of the green sulfur bacteria is *Chlorobium tepidum*.

3. *Green gliding bacteria* (Chloroflexaceae) [also known as filamentous anoxygenic phototrophs, FAP]. These are primarily photoorganotrophic bacteria which can grow under anaerobic conditions in light by photosynthesis or in aerobic conditions in the dark by using respiration to utilize organic compounds from the environment. They are thermophilic bacteria found in hot springs around the world. They also distinguish themselves among the photosynthetic bacteria by possessing mobility. An example is *Chloroflexus aurantiacus*.

4. *Heliobacteria* (Heliobacteriaceae). These are strictly anaerobic bacteria that contain bacteriochlorophyll *g*. They grow primarily using organic substrates and have not been shown to carry out autotrophic growth using only light and inorganic substrates. An example is *Heliobacterium chlorum*.

Like plants, algae, and cyanobacteria, anoxygenic photosynthetic bacteria are capable of photophosphorylation, which is the production of adenosine triphosphate (ATP) from adenosine diphosphate (ADP) and inorganic phosphate (P_i) using light as the primary energy source. Several investigators have suggested that the sole function of the light reaction in bacteria is to make ATP from ADP and Pi. The hydrolysis energy of ATP (or the proton-motive force that precedes ATP formation) can then be used to drive the reduction of CO_2 to carbohydrate by H_2A in reaction (9).

Photochemical apparatus. Photosynthetic bacteria do not have specialized organelles such as the chloroplasts of green plants. Electron micrographs of certain photosynthetic bacteria show tiny spherical sacs, with double-layered walls, as a result of invaginations which form stacks of membranes (**Fig. 12***a*). Other photosynthetic bacteria have invaginations which form thylakoids (Fig. 12*b*). These intracytoplasmic membranes, often called chromatophores, contain the photosynthetic apparatus and can be isolated easily by mechanical disruption of bacteria followed by differential centrifugation. Isolated chromatophores are often used for biochemical and biophysical studies of bacterial photosynthesis.

Reaction centers. The pigment bacteriochlorophyll (BChl) is a necessary component for bacterial

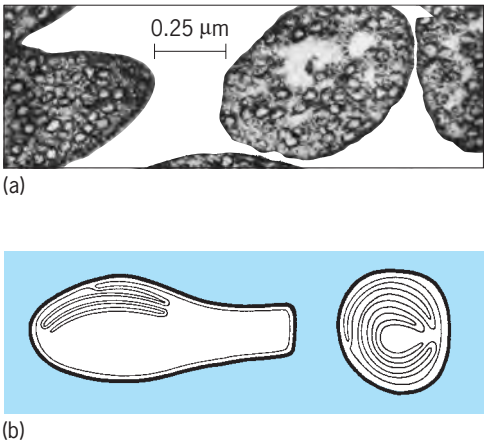

Fig. 12. Photosynthetic bacteria. (*a*) Electron micrograph of *Rhodobacter sphaeroides* with vesicle-like invaginations (*from T. W. Goodwin, ed., Biochemistry of Chloroplasts, vol. 1, Academic Press, 1966*). (*b*) Pictorial representation of a stacked invagination in a photosynthetic bacterium; at left is a longitudinal section and at right is a transverse section (*after R. Whittenbury and A. G. McLee, Archiv. für Mikrobiologie, 59:324–334, 1967*).

photosynthesis. There are specialized BChl molecules in bacteria which engage in the primary chemical reactions of photosynthesis. In addition to these specialized molecules, there are 40–50 BChl molecules referred to as antenna pigments, whose sole function is to harvest light energy and transfer it to reaction center molecules. This is similar to the photosynthetic unit of plants, algae, and cyanobacteria. Each reaction center contains a special pair (dimer) of BChl molecules that engage in chemical reactions after they trap the absorbed light energy. They are also called the energy traps of bacterial photosynthesis.

The energy trap in *Rhodobacter sphaeroides* has been identified as P870. Such identification is carried out with a difference (absorption) spectrophotometer. In this instrument a weak monochromatic measuring beam monitors the absorption of the sample; a brief but bright actinic light given at right angles to the measuring beam initiates photosynthesis. When photosynthesis occurs, changes in absorption take place. **Figure 13***a* shows the absorption spectrum of reaction centers isolated from *R. sphaeroides*. These changes are measured as a function of the wavelength of measuring light. A plot of the change induced in *R. sphaeroides* reaction centers by an actinic light flash, as a function of the wavelength of measuring light, is the difference absorption spectrum (Fig. 13*b*). This spectrum is due largely to the photooxidation of the BChl dimer, P870.

If P870 is the energy trap, then the following criteria must be met: (1) It must undergo a reduction or oxidation reaction, since this is the essential reaction of photosynthesis. The decrease in absorption at 870 nm (Fig. 13) is an oxidation reaction since chemical oxidants cause a similar change. (2) The quantum yield (number of trap molecules oxidized per absorbed photon) must be very high (close to 1.0). (3) The primary light reaction should occur at very low temperatures, down to 1 K (−460°F or −273°C). (4) The above photochemical reaction should be extremely fast, that is, in the picosecond range.

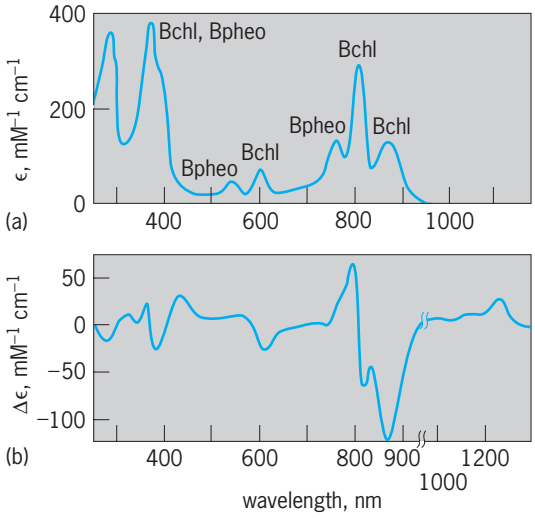

Fig. 13. Plots of (*a*) absorption spectrum and (*b*) the light-induced absorption changes in it, as occurring in reaction centers isolated from carotenoidless mutant R-26 of *Rhodobacter sphaeroides*. In *a*, bands attributed to bacteriochlorophyll and bacteriopheophytin are labeled BChl and BPheo, respectively. The ordinate in a is the millimolar extinction coefficient; in *b*, it is the differential extinction coefficient. (*After R. K. Clayton, Photosynthesis: Physical Mechanisms and Chemical Patterns, Cambridge University Press, 1980*)

All the above criteria are fulfilled by P870, and thus it is the reaction center of bacterial photosynthesis in *Rhodobacter sphaeroides*. Among other reaction centers that have been identified and studied extensively are P890 in *Chromatium vinosum*, P960 in *Rhodopseudomonas viridis* (also called *Blastochloris viridis*), P840 in *Chlorobium limicola* and P798 *Helicobacterium chlorum*. Each species of bacteria has only one type of reaction center, unlike plants, algae, and cyanobacteria, which utilize two types of reaction centers, PSI and PSII, that include P680 and P700, respectively. The reaction centers from oxygenic photosynthetic organisms have been identified by means similar to those used for bacterial reaction centers. Reaction centers have been isolated as pure proteins, which has served the important function of providing a well-defined system in which primary reactions of photosynthesis can be studied. A milestone in bacterial photosynthesis was reached in the early 1980s by the crystallization and determination of the three-dimensional structure of *Rhodopseudomonas viridis* reaction centers by Hartmut Michel, Johann Deisenhofer, and Robert Huber, who received the 1988 Nobel Prize in Chemistry for their work. These crystals enabled an atomic resolution of the molecular structure of the reaction center to be obtained.

Although isolated reaction centers are able to absorb light and convert it to chemical energy, the antenna pigment system in chromatophores (or in whole cells) absorbs most (>90%) of the light. The antenna transfers this energy to the reaction center. Antenna BChl molecules are bound to protein in a specific manner; this binding and pigment-pigment interactions modify the properties of the pigment and define the absorption maxima and the width of the absorption band. An example is B800 (B represents BChl, and the number indicates the wavelength of one of the absorption peaks in nanometers) found in *Rhodobacter sphaeroides* (**Fig. 14**).

Components of photosynthetic bacteria. These bacteria contain the usual components of living material: proteins, lipids, carbohydrates, deoxyribonucleic acid (DNA), ribonucleic acid (RNA), and various metals. However, the specific components of interest to the electron transport system of bacterial photosynthesis are quinones, pyridine nucleotides, and various iron-containing proteins (cytochromes, ferredoxins, Rieske iron-sulfur centers, and others) in addition to the photosynthetic pigments which capture light energy.

In contrast to plastoquinones found in plants, bacteria contain substituted benzoquinones called ubiquinones (UQ or coenzyme Q) and substituted naphthoquinones called menaquinones (MK or vitamin K_2) which act as electron acceptors. The purple bacteria have a pool of UQ (about 25 UQ per reaction center) which mediates transfer of electrons and protons between protein complexes in the

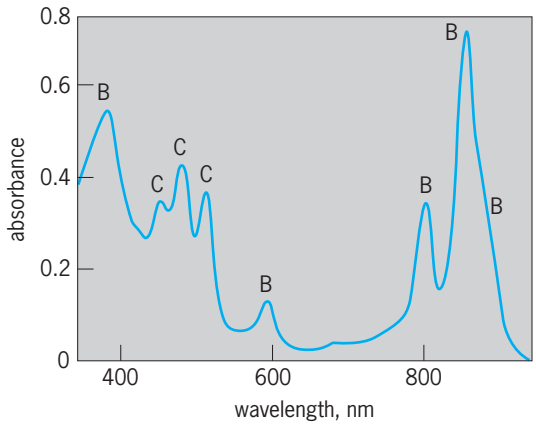

Fig. 14. Absorption spectrum of chromatophores from the bacterium *Rhodobacter sphaeroides*. Absorption bands attributed to BChl *a* are labeled as B, and those attributed to carotenoids as C. (*After R. K. Clayton, Photosynthesis: Physical Mechanisms and Chemical Patterns, Cambridge University Press, 1981*)

chromatophore membrane. However, MK is found only in some bacteria, sometimes in a smaller quantity (about 1–2 MK molecules per reaction center) than the more plentiful UQ. In these organisms, menaquinone's function is probably limited to electron transfer within the reaction center. Other organisms contain only menaquinone. In contrast to plants which contain NADP, the major pyridine nucleotide in bacteria is nicotinamide adenine dinucleotide (NAD); it is present in large quantities and seems to be active in photosynthesis. Among the various cytochromes, the *c*-type cytochromes and the *b*-type cytochromes are the important ones for bacterial photosynthesis.

Pigments. Most photosynthetic bacteria contain BChl *a*, a tetrahydroporphyrin. The chlorophyll of green plants, algae, and cyanobacteria, by contrast, is a dihydroporphyrin. In diethyl ether, BChl *a* has absorption maxima at 365, 605, and 770 nm. The infrared band of various antenna BChl *a* has maxima at 800 (B800), 850 (B850), or 890 nm (B890). These antenna absorption bands in the bacterial cell are due to the formation of complexes of BChl *a* with different proteins. *See* CHLOROPHYLL.

The reaction center protein (composed of L, M, and H subunits) from *Rhodobacter sphaeroides* binds four BChl *a* and two bacteriopheophytin (BPh; similar to BChl but does not contain magnesium). Two of the BChl form the energy trap P870. Another BChl and a BPh are involved in the transfer of electrons within the protein. The exact locations of these chromophores in the reaction center protein was first established in the crystals of *Rhodopseudomonas viridis* reaction centers (**Fig. 15***a*). Similar information is now available for *Rhodobacter sphaeroides* reaction centers (Fig. 15*b*).

The bacterium *Rhodopseudomonas viridis* utilizes an antenna with an infrared absorption band at 1015 nm. The isolated BChl from this species has absorption maxima at 368, 582, and 795 nm in diethyl ether, and has been designated BChl *b*. The reaction center of *R. viridis*, P960, uses BChl *b* and BPh *b* much in the same way as P870 in other bacteria utilizes Bchl *a*. Here, there is an "uphill" energy transfer from the antenna to its reaction center.

The green bacterium *Chlorobium* sp. contains a small amount of BChl *a* but a large quantity of another type of chlorophyll called chlorobium chlorophylls, BChl *c*, *d* or *e*, depending on the species. The BChl *a* has been shown to be associated with the reaction center, and some antenna complexes while the BChl *c* acts only as antenna. It is located in

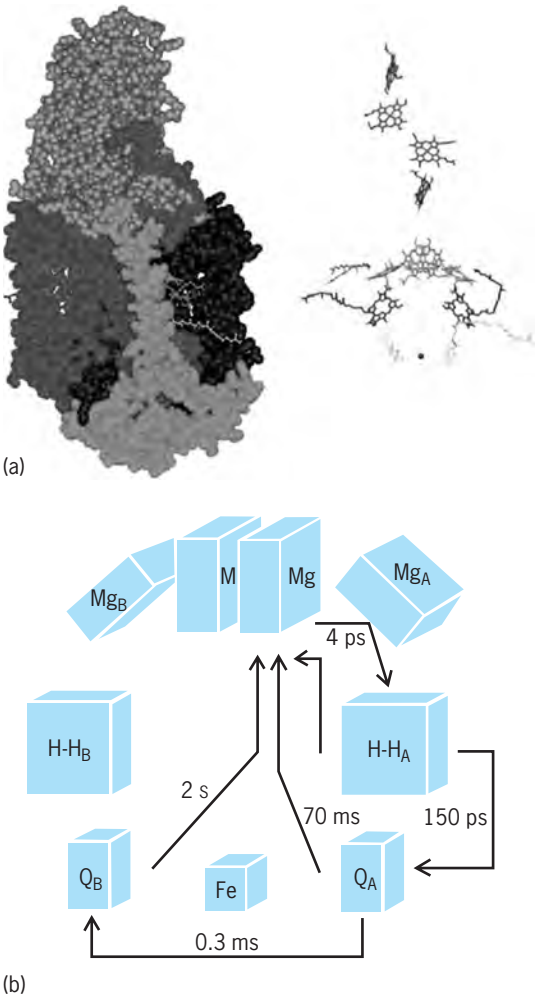

Fig. 15. Reaction centers of bacteria. (*a*) Structure of the reaction center of the purple nonsulfur bacterium *Rhodopseudomonas viridis* as determined by x-ray analysis of the crystalline preparation. In this diagram, the left side shows the complete structure in a space-filling representation. On the right side, the protein has been removed for clarity and only the components of the electron transport chains are shown (*after R. E. Blankenship, Molecular Mechanisms of Photosynthesis, Blackwell Science, Oxford, 2002*). (*b*) A simplified representation of the donor-acceptor complex based on the x-ray data and on spectroscopic data for *Rhodobacter sphaeroides*. The blocks define the aromatic ring systems of bacteriochlorophyll (M-Mg and Mg), bacteriopheophytin (H-H), the quinones (Q), which are ubiquinone and menaquinone, and Fe^{2+}. M-Mg is the primary electron donor, a dimer of bacteriochlorophyll *a* (*Rhodobacter sphaeroides*) or *b* (*Rhodopseudomonas viridis*). Subscripts A and B label the two potential electron transfer pathways, of which only pathway A appears active. The arrows show the various electron transfer reactions with their half-times. Note that Q_B is absent in the crystal of *Rhodopseudomonas viridis* (*after J. F. Norris and G. Van Brakel, Photosynthesis, in Govindjee, J. Amesz, and D. C. Fork, eds., Light Emission by Plants and Bacteria, Academic Press, 1986*).

chlorosome complexes, which are appressed to the cytoplasmic side of the cell membrane and contain about 200,000 molecules of chlorobium chlorophyll.

The second group of pigments is the carotenoids, which have absorption peaks from 450 to 550 nm. The carotenoids of photosynthetic bacteria are of great variety and include some which are found in green plants, for example, the lycopenes. However, some are typical only of bacteria: γ-carotene, which is found in large quantities in green sulfur bacteria, and spirilloxanthol, which is found mainly in purple bacteria. Carotenoids function to prevent photooxidation and destruction of antenna bacteriochlorophyll. They also function in bacterial photosynthesis by transferring their absorbed energy to bacteriochlorophyll. Similar roles are found for carotenoids in plants and cyanobacteria.

Transfer of excitation energy. Light energy absorbed by the carotenoids is transferred to BChl with varying efficiency (30–90%), as demonstrated by the method of sensitized fluorescence. (Similar methods have been used for demonstrating energy transfer from carotenoids, chlorophyll *b*, and phycobilins to chlorophyll *a* in oxygenic photosynthesizers.) When light energy is absorbed by carotenoids, only the fluorescence of bacteriochlorophyll (B875) is observed. By the same method, energy transfer with efficiencies approaching 100% has been demonstrated from B800 to B850 to B875. The high quantum yield (almost 1.0) of P870 oxidation, when bacteria are excited in the antenna pigments, is a clear demonstration of an extremely efficient excitation energy transfer by antenna pigments and trapping in reaction centers.

The lifetime of the excited state of antenna BChl in the bacterial cell is of the order of 30–50 ps. The excitation energy must be channeled from the antenna pigments to the energy traps within this time for efficient photosynthesis to occur. In reaction center preparations, it takes only 3 ps to create a definitively stable charge separation (see below) after the absorption of light. Moreover, the lifetime of the physical state or states preceding P870 oxidation is <3 ps. Thus, it appears that within a few picoseconds of receiving excitation energy, the reaction center has converted the absorbed light energy into chemical energy. Similar reactions occur in plants, algae, and cyanobacteria.

Mechanisms of electron transport. The first act of photosynthesis is the absorption of light by various pigments. As discussed above, light energy absorbed by the carotenoids B800 and B850 is transferred to B875 and finally to the reaction centers, where the primary reaction occurs: the oxidation of the reaction center BChl dimer leads to bleaching of P870 and reduction of an acceptor (**Fig. 16**). In the current model, P (short for P870 and so on) is oxidized to P^+ and an intermediate I is reduced to I^- within a few picoseconds; I includes a BChl monomer and a BPh molecule. The reduced I^- transfers the electron to an iron-quinone complex, reducing the primary quinone (Q_A) to a semiquinone within 100–200 ps. For most anoxygenic bacteria, Q_A is ubiquinone, though for those containing both menaquinone and ubiquinone the menaquinone functions as Q_A. Although an iron atom is in this complex and is within 0.5–1.0 nm of the quinone, its presence is not necessary for the reduction of Q_A, nor does the iron undergo redox changes. The function of this nonheme iron in the reaction center is unknown. In plants, algae, and cyanobacteria, PSII contains Q_A, which is a bound plastoquinone; the function of the iron there is also unknown.

The photooxidized donor BChl dimer, P^+, can be

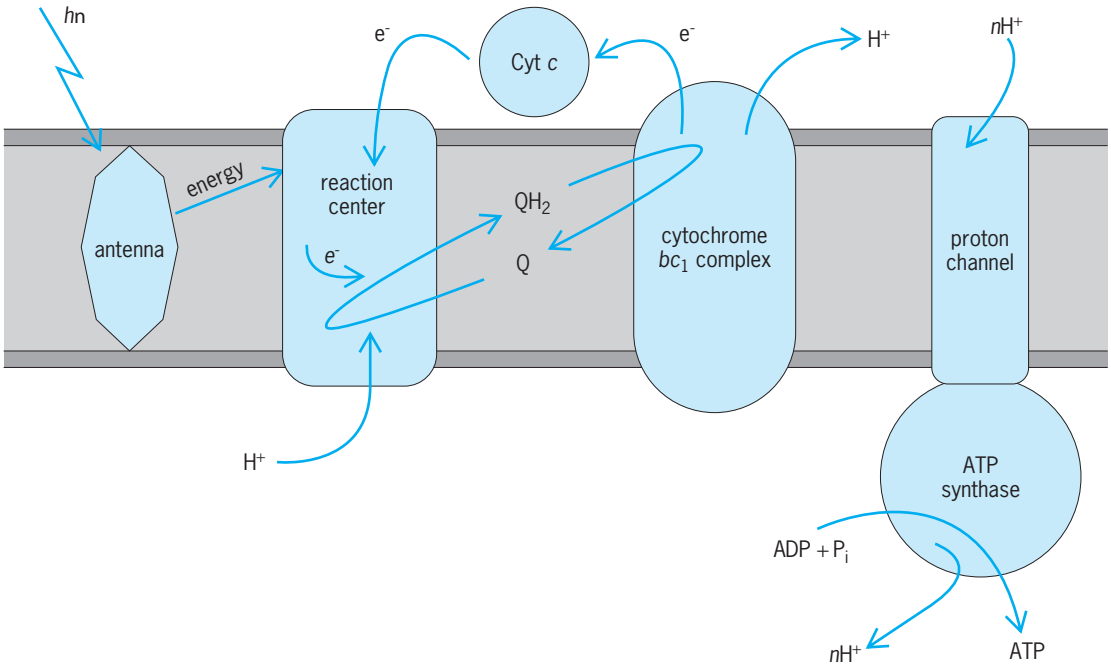

Fig. 16. Electron and proton transport in purple photosynthetic bacteria. For details and explanation of symbols, see the text. The shapes of the proteins are largely hypothetical.

re-reduced by a cytochrome *c* in 1–30 μs, thus oxidizing the cytochrome. In *Rhodobacter sphaeroides* and a number of other species, this cytochrome is soluble Cyt c_2. In other bacteria (for example, *Rhodopseudomonas viridis*) the cytochrome that donates electrons to P^+ is an integral part of the reaction center. The photochemical reactions and the electron transfers in the reaction center are summarized in reaction (10). After this set of reactions, the

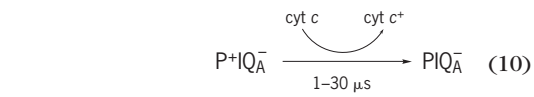

(10)

electron is transferred from Q_A^- to Q_B (a bound UQ), producing $Q_AQ_B^-$. In a subsequent absorption of a photon, the $Q_A^-Q_B^-$ state is created, which is followed by electron transfer from Q_A^- to Q_B^-, forming Q_BH_2 with the uptake of two protons. The bound quinol (Q_BH_2) is replaced by a UQ molecule. This cycle is known as the two-electron gate and is summarized in reaction (11) [omitting the early photochemical steps illustrated in reaction (10)]. The same cycle occurs in photosystem II, except that the electron donor to P^+ is tyrosine Z, and Q_BH_2 is another plastoquinol instead of ubiquinol. The molecular detail is so similar in oxygenic and anoxygenic photosynthesizers that many of the herbicides which act to inhibit PSII electron transfer from Q_A^- to Q_B are also potent inhibitors of electron transfer from Q_A^- to Q_B in photosynthetic bacteria. However, one difference involves a unique role of CO_2/HCO_3^- in reaction (11) in PSII of plants and cyanobacteria, but not in photosynthetic bacteria. Bicarbonate has been shown to be bound on the electron acceptor side of PSII, but not in photosynthetic bacteria.

Protons are taken up from the cytoplasm at the same time as the electrons reduce the quinones. The first proton does not bind directly to the semiquinone (Q_A^- or Q_B^-), but instead it binds to a protonatable amino acid of the reaction center. The net result from the absorption of two photons is the formation of a ubiquinol in the membrane, the oxidation of two Cyt *c*, and the removal of two protons from the cytoplasm of the bacterial cell. In plants and cyanobacteria, the quinol is plastoquinol, and it is water that is ultimately oxidized.

The doubly reduced ubiquinone (QH_2, quinol) through a cyclic pathway serves to re-reduce the oxidized cytochrome (cyt c^+). This cyclic reaction (Fig. 16) is coupled to the production of ATP via the creation of a proton gradient (more accurately a proton motive force) across the membrane. Just as in plants, the proton motive force (which includes two components: a membrane potential, and a proton gradient) is used to drive ATP synthesis. Protons move down the potential gradient through the ATPase to contribute energy to drive the ADP + $P_i \rightarrow$ ATP reaction.

This overall mechanism is consistent with P. Mitchell's chemiosmotic theory. The quinol produced by the two-electron gate mechanism binds

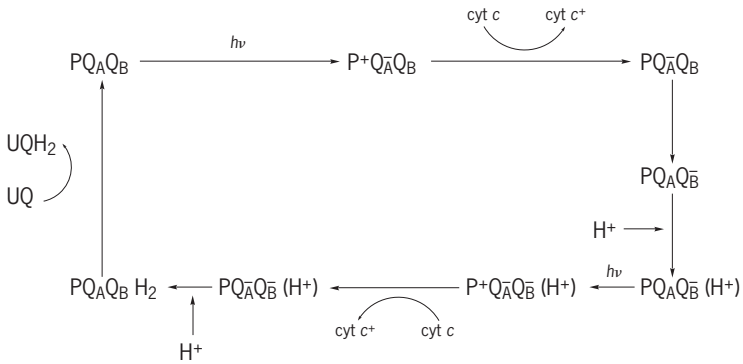

(11)

to the cytochrome *b-c* complex (an integral membrane protein) which contains two *b* cytochromes, a *c* cytochrome, a Rieske iron-sulfur center, and two quinone-binding sites. Plants also contain a similar complex, where cytochrome *b* is replaced by cytochrome b_6, and cytochrome *c* is replaced by cytochrome *f*. The mechanism is strikingly similar, on a molecular level, to that of noncyclic electron transfer from photosystem II to plastocyanin via the plastoquinone pool and the cytochrome *b-f* complex. In all likelihood, it includes a pathway called a Q cycle by Mitchell; this cycle incorporates two different redox-linked pathways for the electrons. For each quinol oxidized by this complex, two molecules of Cyt *c* are reduced, two protons are removed from the quinol, an additional two protons are removed from the cytoplasm, and these four protons are released into the intermembrane space. Absorption of two photons leads to the translocation of four protons across the membrane. Structural and mechanistic data on ATPases indicate that 3–5 H^+'s are needed to make an ATP (Fig. 16).

The mechanism described here for the generation of ATP from light energy is largely from studies on *Rhodobacter sphaeroides* and is generally valid for other purple photosynthetic bacteria.

The mechanisms for oxidizing the reduced substrate H_2A [reaction (9)] are known in much less detail than those for photophosphorylation. Most substrates feed electrons into the quinone pool, and the resulting quinol can be used by the cytochrome *b-c* complex. An example is succinate, which reduces quinone via a succinate dehydrogenase. In bacteria that have a low potential cytochrome *c* bound to the reaction center (such as Cyt c_{551} in *Chromatium vinosum*), electrons from some substrates can possibly be fed into the reaction center through this cytochrome. The electrons for the reduction of NAD^+ in purple photosynthetic bacteria are from the quinone pool, but these electrons require additional energy gained from the hydrolysis of ATP or action of the proton motive force.

Alternatively, especially in some green bacteria, the primary stable acceptor of electrons in the reaction center may not be a quinone but an acceptor with a negative enough oxidation-reduction potential to directly reduce NAD^+. In several green bacteria, this electron acceptor has been shown to be an iron-sulfur (Fe·S) center instead of a quinone. The midpoint redox potential of this Fe·S center is much lower than that for the quinone acceptor in

the purple photosynthetic bacteria. This Fe·S center can then directly reduce a ferredoxin, and this can drive the $NAD^+ \rightarrow NADH$ reaction. The reduced ferredoxin may also feed electrons into a cytochrome *b* complex from which a soluble Cyt *c* could be reduced, thus allowing cyclic electron transfer to occur. This scheme is very reminiscent of PSI-driven reactions in oxygenic photosynthesis. However, not all green photosynthetic bacteria follow the above pattern, but instead they resemble more the purple photosynthetic bacteria.

The reduced pyridine nucleotide NADH and the ATP made in the light reactions are then utilized to convert carbon sources into carbohydrates. The pathway of carbon in anoxygenic photosynthetic bacteria involves a reversed tricarboxylic acid (Krebs) cycle or another cycle called the hydroxypropionate cycle. *See* BACTERIAL PHYSIOLOGY AND METABOLISM.

Govindjee; Robert E. Blankenship; R. J. Shopes

Bibliography. R. E. Blankenship, *Molecular Mechanisms of Photosynthesis*, Blackwell Science, Oxford, 2002; B. Demmig-Adams et al. (eds.), *Photoprotection, Photoinhibition, Gene Regulation, and Environment*, Springer, Dordrecht, 2005; P. G. Falkowski and J. A. Raven, *Aquatic Photosynthesis*, 2d ed., Princeton University Press, Princeton, 2007; J. H. Golbeck (ed.), *Photosystem I: The Light-Driven Plastocyanin: Ferredoxin Oxidoreductase*, Springer, Dordrecht, 2006; Govindjee et al. (eds.), *Discoveries in Photosynthesis*, Springer, Dordrecht, 2005; B. R. Green and W. W. Parson (eds.), *Light-Harvesting Antennas in Photosynthesis*, Springer, Dordrecht, 2003; B. Grimm et al. (eds.), *Chlorophylls and Bacteriochlorophylls: Biochemistry, Biophysics, Functions and Applications*, Springer, Dordrecht, 2006; B. Ke, *Photosynthsis: Photobiochemistry and Photobiophysics*, Springer, Dordrecht, 2001; A. W. D. Larkum et al. (eds.), *Photosynthesis in Algae*, Springer, Dordrecht, 2003; R. C. Leegood et al. (eds.), *Photosynthesis: Physiology and Metabolism*, Springer, Dordrecht, 2000; G. Papageorgiou and Govindjee (eds.), *Chlorophyll a Fluorescence: A Signature of Photosynthesis*, Springer, Dordrecht, 2005; R. R. Wise and J. K. Hoober (eds.), *The Structure and Function of Plastids*, Springer, Dordrecht, 2006; T. Wydrzynski and K. Satoh (eds.), *Photosystem II: The Light-Driven Water-Plastoquinone Oxidoreductase*, Springer, Dordrecht, 2005.

Phototransistor

A semiconductor device with electrical characteristics that are light-sensitive. Phototransistors differ from photodiodes in that the primary photoelectric current is multiplied internally in the device, thus increasing the sensitivity to light. For a discussion of this property *see* TRANSISTOR.

Some types of phototransistors are supplied with a third, or base, lead. This lead enables the phototransistor to be used as a switching, or bistable, device. The application of a small amount of light causes the device to switch from a low current to a high current condition. *See* PHOTOELECTRIC DEVICES.

W. R. Sittner

Phototube

An electron tube comprising a photocathode and an anode mounted within an evacuated glass envelope through which radiant energy is transmitted to the photocathode. A gas phototube contains, in addition, argon or other inert gas which provides amplification of the photoelectric current by partial ionization of the gas. The photocathode emits electrons when it is exposed to ultraviolet, visible, or near-infrared radiation. The anode is operated at a positive potential with respect to the photocathode. *See* ELECTRICAL CONDUCTION IN GASES; ELECTRON TUBE.

Characteristics. A phototube responds to radiation over a limited range of the spectrum that is determined by the photocathode material. Radiant sensitivity, shown in the **illustration** as a function of wavelength, is the photoelectric current emitted per unit of incident monochromatic radiant power. Sensitivity on the short-wavelength side of the curves is limited by the transmittance of the glass envelope. Electron affinity of the photocathode determines the long-wavelength threshold of sensitivity. *See* PHOTOEMISSION.

Typical phototube characteristics are summarized in the **table**. Quantum efficiency, or photoelectron yield, is the number of electrons emitted per incident photon. It is tabulated at the wavelength of maximum response. For photometric applications a useful parameter is luminous sensitivity: the photoelectric current per lumen incident from a specified source of light. A source commonly used

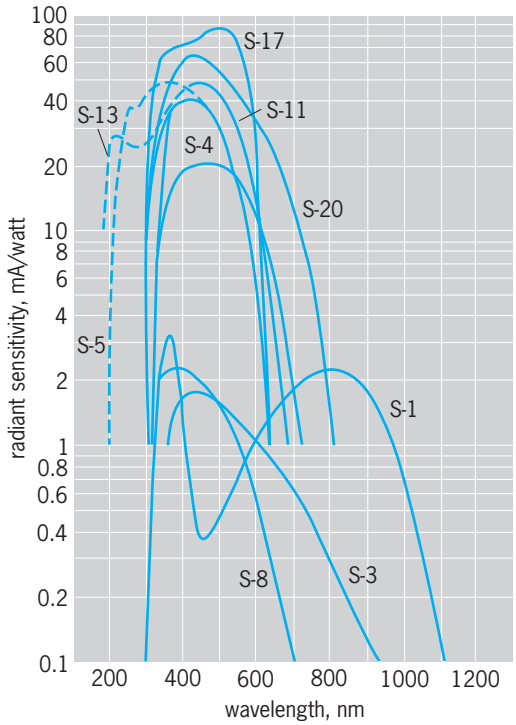

Curves of the average spectral sensitivity characteristics of some typical phototubes.

Average cathode characteristics

Spectral sensitivity characteristic*	Cathode material	Wavelength of maximum response, nm	Peak radiant sensitivity, mA/W	Peak cathode quantum efficiency, %	Luminous sensitivity, μA/lumen†	Remarks
S-1	Cs_2O, Ag	800	2.2	0.3	25	
S-3	Rb_2O, Ag	420	1.8	0.5	6.5	
S-4	Cs_3Sb	400	40	12.4	40	
S-5	Cs_3Sb	340	49	17.8	40	Ultraviolet transmitting window
S-8	Cs_3Bi	365	2.3	0.8		
S-10	Bi, Ag, O, Cs	450	20.3	5.6	40	Semitransparent
S-11	Cs_3Sb	440	48	13.5	60	Semitransparent
S-13	Cs_3Sb	440	47	13.2	60	Semitransparent; ultraviolet transmitting window
S-17	Cs_3Sb	490	85	21.4	125	Semitransparent, on reflecting substrate
S-20	(NaKCs)Sb	420	64	18.8	150	Semitransparent

*These characteristics, shown in the illustration, refer to typical phototubes rather than to photocathodes.
†Light source is a tungsten-filament lamp operated at a color temperature of 4700°F (2870 K).

is a tungsten-filament lamp operated at a color temperature of 4700°F (2870 K). *See* INCANDESCENCE; LUMINOUS FLUX; PHOTON.

Photocathodes. Photocathodes are semiconductors which contain one or more of the alkali metals sodium, potassium, rubidium, or cesium chemically combined with bismuth, antimony, or silver oxide. The cathode surface contains a critical excess of the alkali metal which enhances photoelectric emission by decreasing the affinity of the surface for electrons. Negative affinity for electrons is achieved with the gallium arsenide:cesium (GaAs:Cs) and indium gallium arsenide:cesium (InGaAs:Cs) photocathodes used in photomultipliers. Phototubes also emit electrons thermionically at ambient temperatures. This "dark current," observed in the absence of all irradiance, increases almost exponentially with temperature. Thermionic emission from the cesium antimonide (CsSb) photocathode is about 10^{-15} A/cm^2 at 68°F (20°C). *See* PHOTOMULTIPLIER; SEMICONDUCTOR.

Applications. Vacuum phototubes are used as detectors of radiant energy in the spectral range from 200 to 1100 nanometers. Since the photoelectric current is directly proportional to the intensity of the radiation, these tubes are used in radiometers, photometers, and colorimeters. By virtue of their narrow pulse response, vacuum phototubes are also used to measure the intensity of very short pulses of light generated by lasers and visible nuclear radiation. Gas phototubes can be used in light-operated relays and for the reproduction of sound from motion picture film, although their response to intensity-modulated light is limited to frequencies below 15 kHz. Vacuum as well as phototubes have been replaced in many applications by semiconductor photodiodes and photovoltaic cells. *See* LASER; PHOTODIODE; PHOTOMETER; PHOTOVOLTAIC CELL; RADIOMETRY.

James L. Weaver

Bibliography. M. Cardona and L. Ley (eds.), *Photoemission in Solids*, 1978 and 1979; E. L. Dereniak and D. G. Crowe, *Optical Radiation Detectors*, 1984; A. Sommer, *Photoemissive Materials*, 1968, reprint 1980.

Photovoltaic cell

A device that detects or measures electromagnetic radiation by generating a current or a voltage, or both, upon absorption of radiant energy. Specially designed photovoltaic cells are used for power generation, as in solar batteries or solar cells, and for sensitive detection of electromagnetic radiation in radiometry, optical communications, spectroscopy, and other applications. An important advantage of the photovoltaic cell in these particular applications is that no separate bias supply is needed—the device generates a signal (voltage or current) simply by the absorption of radiation.

Most photovoltaic cells consist of a semiconductor *pn* junction or Schottky barrier in which electron-hole pairs produced by absorbed radiation are separated by the internal electric field in the junction to generate a current, a voltage, or both, at the device terminals. The influence of the incident radiation on the current-voltage characteristics of a photovoltaic cell is to add a negative current, equal to the generated photocurrent, to the dark current-voltage characteristic of the diode and thus shift the current-voltage characteristic downward by the magnitude of the photogenerated current, as shown in the **illustration**. Under open-circuit conditions (current $I = 0$) the terminal voltage increases with increasing light intensity (points A), and under short-circuit conditions (voltage $V = 0$) the magnitude of the current increases with increasing light intensity (points B). When the current is negative and the voltage is positive (point C, for example), the photovoltaic cell delivers power to the external circuit. In this case, if the source of radiation is the Sun, the photovoltaic cell is referred to as a solar battery or solar cell. When a photovoltaic cell is used as a photographic exposure meter, it produces a current proportional to the light intensity (points B), which is indicated by a low-impedance galvanometer or microammeter. For use as sensitive detectors of infrared radiation, specially designed photovoltaic cells can be operated with either low-impedance (current) or high-impedance (voltage) amplifiers,

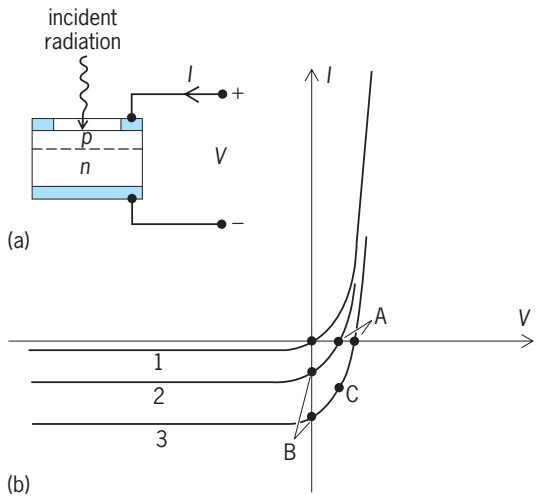

The *pn* junction photovoltaic cell. (*a*) Cross section. (*b*) Current-voltage characteristics. Curve 1 is for no incident radiation or the dark current-voltage characteristic, and curves 2 and 3 for increasing incident radiation.

although the lowest noise and highest sensitivity are achieved in the current or short-circuit mode. Another mode of operation of a *pn* junction diode as a photodetector involves the application of a reverse bias voltage to the diode. In this case, the photogenerated current is directly proportional to the incident power, and the diode is said to be operated in the photodiode mode rather than the photovoltaic mode. *See* EXPOSURE METER; JUNCTION DIODE; OPTICAL DETECTORS; PHOTODIODE; PHOTOELECTRIC DEVICES; PHOTOVOLTAIC EFFECT; RADIOMETRY; SEMICONDUCTOR; SEMICONDUCTOR DIODE; SOLAR CELL.

Gregory E. Stillman

Bibliography. B. K. Das and S. N. Singh (eds.), *Photovoltaic Materials and Devices*, 1985; H. J. Moller, *Semiconductors for Solar Cells*, 1993; L. D. Partain (ed.), *Solar Cells and Their Applications*, 1995; R. J. Van Overstraeten and R. P. Mertens, *Physics, Technology and Use of Photovoltaics*, 1986; C. M. Wolfe, N. Holonyak, Jr., and G. E. Stillman, *Physical Properties of Semiconductors*, 1989; K. Zweibel, *Harnessing Solar Power: The Photovoltaics Challenge*, 1990.

Photovoltaic effect

The conversion of electromagnetic radiation into electric power through absorption by a semiconducting material. Devices based on this effect serve as power sources in remote terrestrial locations and for satellites and other space applications. Photovoltaic powered calculators and other consumer electronic products are widely available, and solar photovoltaic automobiles and aircraft have been demonstrated.

Principles. The basic requirements for the photovoltaic effect are (1) the absorption of photons through the creation of electron-hole pairs in a semiconductor; (2) the separation of the electron and hole so that their recombination is inhibited and the electric field within the semiconductor is altered; and (3) the collection of the electrons and holes, separately, by each of two current-collecting electrodes so that current can be induced to flow in a circuit external to the semiconductor itself.

There are many approaches to achieving these three requirements simultaneously. A very common approach for separating the electrons from the holes is to use a single-crystal semiconductor, for example, silicon, into which a *pn* junction has been diffused. Silicon is often chosen because its optical band gap permits the absorption of a substantial portion of solar photons via the generation of electron-hole pairs. The fabrication of such a device structure causes a local transfer of negative charges from the *n* layer into the *p* layer, bending the conduction and valence bands in the vicinity of the *p-n* boundary, and thereby creating a rectifying junction. Electrons generated in the *p* region can lower their energy by migrating into the *n* region, which they will do by a random walk process in the electric-field-free region far from the junction, or by drift induced by the electric field in the junction region. Holes created in the *n* region, conversely, lose energy by migrating into the *p* region. Thus the presence of such a junction leads to the spontaneous spatial separation of the photogenerated carriers, thereby inducing a voltage difference between current-carrying electrodes connected to the *p* and *n* regions. This process will continue until the difference in potential between the two electrodes is large enough to flatten the bands in the vicinity of the junction, canceling out the internal electric field existing there and so eliminating the source of carrier separation. The resulting voltage is termed the open-circuit voltage V_{oc}, and approximates the built-in voltage associated with the *pn* junction in the dark, a value which cannot exceed the band gap of the semiconductor. *See* HOLE STATES IN SOLIDS; SEMICONDUCTOR; SEMICONDUCTOR DIODE.

In the limit when the device is short-circuited by the external circuit, no such buildup of potential can occur. In this case, one electron flows in the external circuit for each electron or hole which crosses the junction, that is, for each optically generated electron-hole pair which is successfully separated by the junction. The resulting current is termed the short-circuit current J_{sc} and, in most practical photovoltaic devices, approaches numerically the rate at which photons are being absorbed within the device. Losses can arise from the recombination of minority carriers (for example, electrons in the *p*-type region, holes in the *n*-type region) with majority carriers. *See* ELECTRON-HOLE RECOMBINATION.

Power generation. For a photovoltaic device to generate power, it is necessary to provide a load in the external circuit which is sufficiently resistive to avoid short-circuiting the device. In this case, the voltage will be reduced compared to the open-circuit voltage because a continuing requirement exists for carrier separation at the junction; thus some band bending and its associated internal field must be retained. A

maximum power conversion point in the available current-voltage (*I-V*) output plane exists for any photovoltaic device operating in a given photon flux. Device efficiency is defined as this maximum power output divided by the incident radiative power. The maximum power is usually expressed as the product of V_{oc}, J_{sc}, and F.F., the fill factor, which is the ratio of the area in the *I-V* plane defined by the maximum power output divided by the area defined by the product of V_{oc} and J_{sc}.

Alternative materials. While this explanation of the photovoltaic effect has been given in terms of the familiar crystalline silicon *pn* photovoltaic device, many alternative materials and device structures have been explored since the discovery of the photovoltaic effect in the nineteenth century. In fact, the first all-solid-state photovoltaic device appears to have used a semiconducting layer of glassy selenium illuminated through a semitransparent gold electrode in what would now be considered a metal-semiconductor Schottky barrier configuration.

Amorphous materials tend to have low carrier mobilities and, therefore, short diffusion lengths of minority carriers prior to recombination. Thus, devices utilizing amorphous materials tend to require internal electric field–induced carrier drift in order to achieve efficient carrier separation. Fortunately, the optical absorption of amorphous films can be matched so well to the incident solar radiation that devices as thin as 0.5 micrometer can be utilized, thus permitting the presence of electric fields throughout most of the absorbing layer thickness. *See* AMORPHOUS SOLID.

Multiple-layered devices. Various multiple-layered device configurations based on doped and undoped alloys of amorphous silicon have been developed for photovoltaic devices used in applications ranging from solar watches and calculators to remote power generators. The photovoltaic effect in these devices is particularly intriguing since it is possible to build up so-called tandem devices by stacking one device electrically and optically in series above another. In addition to the increased voltage and concomitant reduction in the required current-carrying capability of electrode grid structures, such devices permit, in principle, an increased efficiency of solar photovoltaic energy conversion. This is achieved by matching the band gap in the upper device with the higher-photon-energy portion of the solar spectrum, thereby achieving a higher voltage, while matching the lower device with the lower-photon-energy portion, thereby utilizing a higher portion of these lower-energy photons than could be efficiently achieved in a single device.

Other configurations and phenomena. The junction in photovoltaic devices, normally configured to lie in the plane normal to the incident radiation, can be contoured to promote significant lateral drift of carriers to point-source contracts on the back surface of the device, where they do not occlude any of the incident radiation. Such device structures, when they also include surface texturing to enhance trapping of incident photons, can increase solar conversion efficiency. Finally, while band bending and internal fields are implicitly combined to create rectifying junctions in a single material (homojunctions), these two charge-separating mechanisms can be utilized independently, in principle, in heterojunctions, which are junctions created between two dissimilar semiconductor materials. Extension of the thin-film deposition techniques used to create tandem amorphous devices appears to have the potential for creating structures, including expitaxial structures, in which the Fermi level and the conduction and valence band energies are independently controllable throughout the device thickness. Such control could permit the consideration of practical photovoltaic device structures which today exist only as theoretical possibilities. *See* SEMICONDUCTOR HETEROSTRUCTURES; SOLAR CELL.

John P. de Neufville

Bibliography. B. K. Das and S. N. Singh (eds.), *Photovoltaic Materials and Devices*, 1985; A. L. Fahrenbruch and R. H. Bube, *Fundamentals of Solar Cells*, 1983; H. J. Moller, *Semiconductors for Solar Cells*, 1993; L. D. Partain (ed.), *Solar Cells and Their Applications*, 1995; S. M. Sze, *Physics of Semiconducting Devices*, 2d ed., 1981; R. J. Van Overstraeten and R. P. Mertens, *Physics, Technology and Use of Photovoltaics*, 1986; K. Zweibel, *Harnessing Solar Power: The Photovoltaics Challenge*, 1990.

Phthiraptera

An order of small wingless insects, permanently parasitic on mammals and birds. The Phthiraptera are one of approximately 30 major groups of insects and can be divided into four suborders, three of which (the Amblycera, Ischnocera, and Rhynchophthirina) are known as chewing or biting lice, and the fourth (the Anoplura) as sucking lice. Amblycera and Ischnocera are found on both mammals and birds, whereas species of Rhynchophthirina and Anoplura are confined to mammals. In total there are approximately 5000 known species, of which just under 70% are recorded from a single host species. *See* ANOPLURA; INSECTA; PARASITOLOGY.

Characteristics. Lice spend their entire life cycle in the microhabitat provided by the hosts' skin, fur, or feathers, and are highly adapted to this environment. The louse body (see **illustration**) is dorsoventrally flattened, and their legs are modified for clinging to feathers or fur. Their adult size ranges from 3 to 11 mm (0.1 to 0.4 in.) and frequently correlates with the size of the host species on which they are normally found. Coloration varies from pale white through shades of yellow and brown to black. Patterns are evident on some species, and cryptic coloration is often employed to match the coloration of the host. Chewing lice have mandibles that have been variously modified in each of the suborders. The mandibles are involved in cutting up the hair or feather material before it is digested, and play a vital secondary role in anchoring the louse to the host. The mandibles of Rhynchophthirina (the elephant and warthog lice) are present at the end of

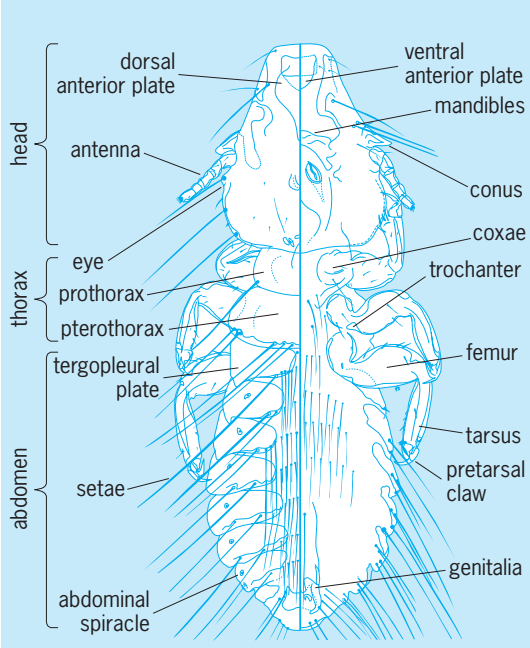

(a) (b)

Ischnoceran bird louse *Strigiphilus vapidus*, a male from an owl (*Ninox novaeseelandiae*), shown in (*a*) dorsal view and (*b*) ventral view.

a long rostrum and, unlike most chewing lice, articulate outward rather than oppose each other. This enables them to tear at the hosts' thick skin and feed from the pool of blood that forms. In most species of sucking lice, the mandibles have been completely lost. Instead, sharp stylets protrude from the head and are used to pierce small vessels located close to the skin's surface. Anoplura draw up blood using these stylets—hence their name sucking lice.

Phylogeny. Parasitic lice are closely related to their nonparasitic free-living ancestors, the Psocoptera (book- or barklice). Just one louse is known from the fossil record—a spectacularly preserved specimen that has been dated at approximately 44 million years old. The evolutionary position of this fossil louse in a relatively contemporary group of living lice suggests that the origin of parasitic lice significantly predates the age of this fossil. Based on the latest molecular evidence, the data suggest that lice diversified from their psocopteran ancestors at least 100 million years ago, long before the extinction of the dinosaurs. Given the common ancestry between birds and dinosaurs, and the more recent diversification of mammal lice compared to those infesting birds, it has been suggested that the first hosts for parasitic lice might have been the feathered theropod dinosaurs known from this period. The close association between hosts and lice may parallel similar patterns of evolutionary relationship, with particular events in the evolutionary history of lice being matched to events in the evolutionary history of their hosts. In some cases, these patterns have been used to resolve controversies in host evolutionary history, such as those surrounding the affinities of modern and archaic humans, and the relationships of flamingos with storks and ducks. *See* ORGANIC EVOLUTION; PSOCOPTERA.

Distribution. The present distribution of lice essentially mirrors that of their hosts, although true geographic distributions of lice are known to occur within the range of certain host species. Consequently lice are found worldwide on every continent and in virtually every habitat occupied by mammals and birds. All bird orders and most bird families have records of host-specific lice. Likewise, most mammals have lice with the exception of species belonging to the orders Chiroptera (bats), Edentata (anteaters, tree sloths, and armadillos), Pholidota (pangolins), Cetacea (whales and dolphins), and Sirenia (dugongs and manatees). The host specificity of most louse species can be attributed to the infrequency of contact between different host species; however, some lice have physical adaptations to particular host species and are not capable of living on another species. Because lice are unable to survive for more than short periods off the hosts' body, behavioral adaptations in the lice reinforce host specificity. For example, lice are attracted by body heat and repelled by light, preventing them from leaving the warmth and darkness of the hosts' plumage or pelage. Most birds and some mammals are typically infected with more than a single species of louse, with many having three to five species. These usually inhabit different parts of the hosts' body, and are adapted to the particular microhabitat of the host in which they normally live.

Life cycle. With the exception of the human body louse, all species complete their entire life cycle from egg to adult on the body of their host. Females are typically larger than males and often outnumber them on any one host. In some species, males are rarely found, and reproduction can occur without male fertilization (parthenogenesis). Eggs are laid singularly or in clumps, and in most cases are cemented to the feather or hair shaft. Once hatched, lice pass through three nymphal instars (developmental stages between molts) before reaching adult stage. This typically takes 20–30 days, depending upon temperature and the louse species. Often the life cycles of lice are carefully timed to coincide with the breeding activities of their host.

Infestation. Lice are important pests of domesticated mammals and birds. Heavy infestation may cause intense irritation, and scratching may lead to secondary infections. Rubbing and damage to hides and wool may also occur, and meat and egg production can be badly affected. Louse infestations on most animals are usually very low in number, but sick or injured animals, especially those whose ability to preen or groom has been impaired, may have infestations numbering thousands of individual lice. The human body louse (*Pediculus humanus humanus*) is the vector of *Rickettsia prowazekii*, which causes louse-borne typhus. During the seventeenth to early twentieth century, this louse was responsible for the deaths of millions of people during the plagues, famines, and wars that beset people

living at the time. Today, louse-borne typhus, along with louse-borne relapsing fever and trench fever are rare occurrences, usually confined to displaced persons during times of civil or political unrest. However, the human head louse (*Pediculus humanus capitis*) is commonplace in the developed world, and infests millions of schoolchildren every year. The only other louse known to infect humans is the pubic louse (*Pthirus pubis*), which is normally transmitted through sexual contact. *See* PEDICULOSIS; RELAPSING FEVER; RICKETTSIOSES; SEXUALLY TRANSMITTED DISEASES. Vincent S. Smith

Bibliography. L. A. Durden and G. G. Musser, The sucking lice (Insecta, Anoplura) of the world: A taxonomic checklist with records of mammalian hosts and geographic distributions, *Bull. Amer. Mus. Nat. His.*, 218:1–90, 1994; D. Grimaldi and M. S. Engel, *Evolution of the Insects*, Cambridge University Press, 2005; K. C. Kim, H. D. Pratt, and C. J. Stojanovich, *The Sucking Lice of North America*, Pennsylvania State University, 1986; M. A. Price and O. H. Graham, Chewing and sucking lice as parasites of mammals and birds, *USDA Agr. Res. Ser. Tech. Bull.*, no. 1849, 1997; R. D. Price et al., *The Chewing Lice: World Checklist and Biological Overview*, Illinois Natural History Survey, Champaign-Urbana, 2003; M. Rothschild and T. Clay, *Fleas, Flukes & Cuckoos: A Study of Bird Ectoparasites*, Collins, London, 1952; H. Zinsser, *Rats, Lice and History*, Routledge, London, 1935.

Phycobilin

Any member of a class of intensely colored pigments found in some algae that absorb light for photosynthesis. Phycobilins are open-chain tetrapyrroles structurally related to mammalian bile pigments, and they are unique among photosynthetic pigments in being covalently bound to proteins (phycobiliproteins). In at least two groups of algae, phycobiliproteins are aggregated in a highly ordered protein complex called a phycobilisome.

Occurrence. Phycobilins occur only in three groups of algae: cyanobacteria (blue-green algae), Rhodophyta (red algae), and Cryptophyceae (cryptophytes), and are largely responsible for their distinctive colors, including blue-green, yellow, and red. Five different phycobilins have been identified to date, but the two most common are phycocyanobilin [structure (**1**)], a blue pigment, and phycoerythrobilin (**2**), a red pigment. In the cell, these pigments absorb light maximally in the orange (620-nanometers) and green (550-nm) portion of the visible light spectrum, respectively. A blue-green light (495-nm) absorbing pigment, phycourobilin, is found in some cyanobacteria and red algae. A yellow light (575-nm) absorbing pigment, phycobiliviolin (also called cryptoviolin) is apparently found in all cryptophytes but in only a few cyanobacteria. A fifth phycobilin, which absorbs deep-red light (697 nm), has been identified spectrally in some cryptophytes, but its chemical properties are

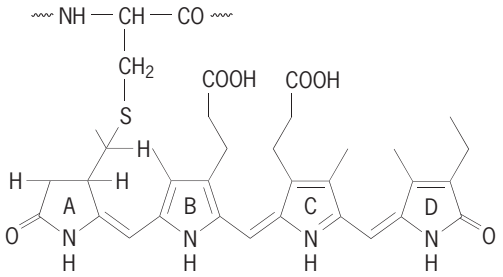

(**1**)

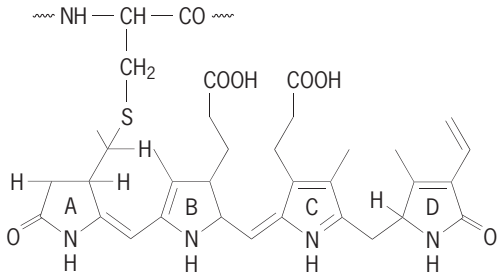

(**2**)

unknown. *See* CRYPTOPHYCEAE; CYANOPHYCEAE; RHODOPHYCEAE.

Phycobilins are associated with the photosynthetic light-harvesting system in chloroplasts of red algae and cryptophytes and with the photosynthetic membranes of cyanobacteria, which lack chloroplasts. Phycobilins are covalently bound to a water-soluble protein that aggregates on the surface of the photosynthetic membrane. All other photosynthetic pigments (for example, chlorophylls and carotenoids) are bound to photosynthetic membrane proteins by hydrophobic attraction. Phycobiliprotein can constitute a major fraction of an alga. In some cyanobacteria, for example, fresh-water or marine *Synechococcus*, phycobiliproteins can account for more than 50% of the soluble protein and one-quarter of the dry weight of the cell. *See* CELL PLASTIDS.

Phycobiliproteins are classified primarily by their absorption spectrum (**Fig. 1**), which depends on the protein structure and the number and kind of phycobilins that are attached. There are three major classes, allophycocyanin (APC), phycocyanin (PC), and phycoerythrin (PE) [see **table**]. All cyanobacteria and red algae contain allophycocyanin and phycocyanin, while some species of these groups also contain phycoerythrin. Cryptophytes possess either phycocyanin or phycoerythrin, but not both. All three biliproteins are composed of at least two polypeptide subunits, α and β, each having a molecular weight in the range of 16,000–22,000. The two subunits, which are always present in equal amounts, form stable aggregates of typically two, three, or six pairs. Some phycoerythrins contain a third polypeptide, designated γ, which is usually found singly with six α-β pairs. Each subunit may contain one to four phycobilin chromophores, depending on the biliprotein. Because of the tendency to aggregate to varying degrees, molecular weights of phycobiliproteins range widely between 35,000 and 240,000.

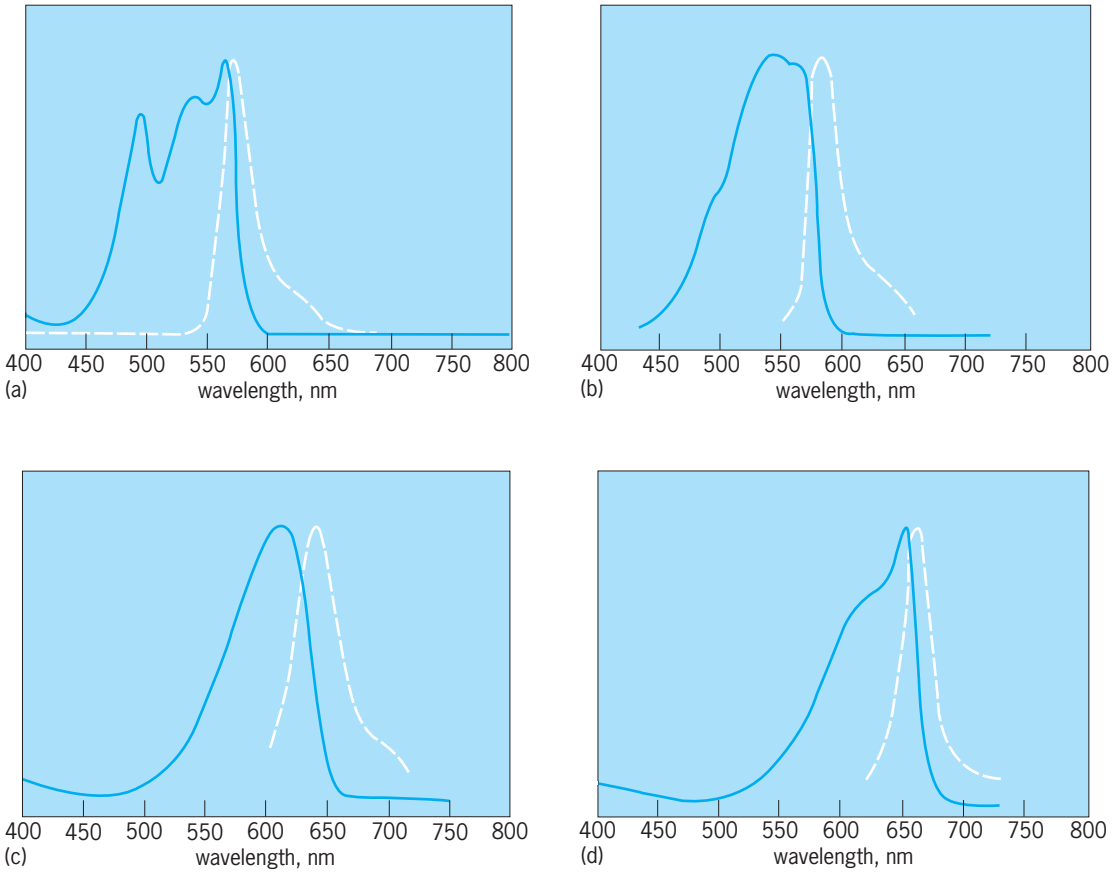

Fig. 1. Absorption spectra (colored lines) and fluorescence emission spectra (white lines) for several isolated phycobiliproteins. (*a*) R-phycoerythrin. (*b*) B-phycoerythrin. (*c*) C-phycocyanin. (*d*) Allophycocyanin. (*After M. N. Kronick, The use of phycobiliproteins as fluorescent labels in immunoassay, J. Immunol. Meth., 92:1–13, 1986*)

Function. Phycobilins are photosynthetic accessory pigments that absorb light efficiently in the yellow, green, orange, or red portion of the light spectrum, where chlorophyll *a* only weakly absorbs. Differences among species in phycobiliprotein may be related to the light environment in which the algae grow. For example, cyanobacteria that grow in shallow fresh-water environments generally contain red-light-absorbing phycocyanin as the primary accessory pigment, whereas cyanobacteria from the open

Some phycobiliproteins and their characteristics*

Phycobiliprotein	Distribution	Wavelength of absorption peaks in visible light spectrum, nm	Wavelength of maximum fluorescence emission, nm	Presence of phycobilins		
				α subunit	β subunit	γ subunit
Allophycocyanin	Cyanobacteria, Rhodophyta	650	660	1 phycocyanobilin	1 phycocyanobilin	
Allophycocyanin B	Cyanobacteria, Rhodophyta	671, 618	680	1 phycocyanobilin	1 phycocyanobilin	
C-phycocyanin	Cyanobacteria, Rhodophyta	620	637	1 phycocyanobilin	2 phycocyanobilin	
R-phycocyanin	Rhodophyta	617, 555	636	1 phycocyanobilin	1 phycocyanobilin, 1 phycoerythrobilin	
Phycoerythrocyanin	Cyanobacteria	568, 590	619	1 phycobiliviolin	2 phycocyanobilin	
Phycocyanin 645	Crytophyceae	645, 585	660	1 697†	2 phycocyanobilin 1 phycobiliviolin	
C-phycoerythrin	Cyanobacteria	565, 540	577	2 phycoerythrobilin	4 phycoerythrobilin	
R-phycoerythrin‡	Rhodophyta	568, 545, 498	578	2 phycoerythrobilin	2 phycoerythrobilin, 1 phycourobilin	1 phycoerythrobilin, 3 phycourobilin
B-phycoerythrin	Rhodophyta	545, 563, 498	575	2 phycoerythrobilin	3–4 phycoerythrobilin	2 phycoerythrobilin, 2 phycourobilin
b-Phycoerythrin	Rhodophyta	545, 563	570	2 phycoerythrobilin	4 phycoerythrobilin	
Phycoerythrin 555	Cryptophyceae	545	585	? phycoerythrobilin	? phycoerythrobilin	

*Not included are two known phycoerythrins and phycocyanins from cryptophytes, as well as several phycoerythrins from marine cyanobacteria, which contain unusually high amounts of phycourobilin.
†Unidentified chromophore with peak absorption at 697 nm.
‡Variable phycobilin composition.

ocean also possess phycoerythrin, which absorbs the blue-green and green light that penetrates to greater depths.

Energy transfer. Light energy absorbed by phycobilins is transferred with greater than 90% efficiency to chlorophyll *a*, where it is used for photosynthesis. Phycobiliproteins are associated with phycobilisomes, which are organelles attached to the surface of thylakoid membranes. Phycobilisomes are found in all cyanobacteria and red algae (but not cryptophytes) and contain allophycocyanin, phycocyanin, and sometimes phycoerythrin. The morphology of phycobilisomes is somewhat variable, but the most prevalent form is hemidiscoidal, that is, in the shape of a semicircular disk standing with its straight edge facing the membrane (**Fig. 2**). When this type of phycobilisome is isolated from the cell and viewed with an electron microscope, it is possible to see a core of two or three small disks with usually six radiating rods. The rods resemble stacked disks and vary in number depending on species and growth conditions. The individual disks contain phycobiliprotein, usually in a dimeric, trimeric, or heximeric α-β aggregate. Additional proteins, called linkers, which may have phycobilins attached, are also present and have

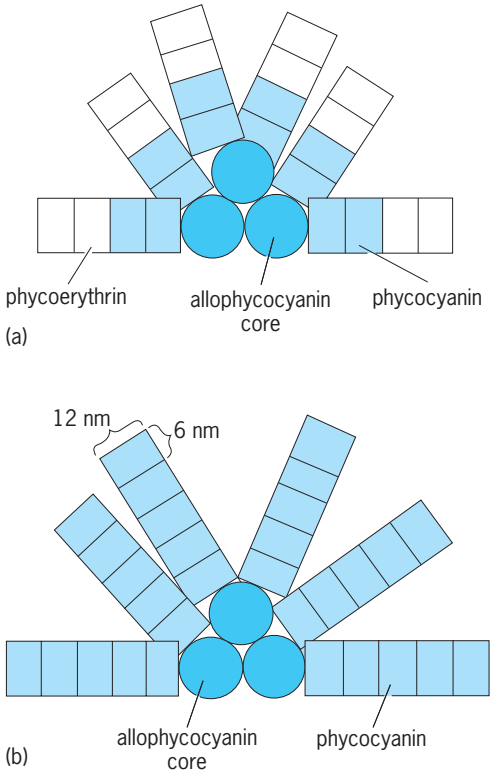

Fig. 2. Structure of a hemidiscoidal phycobilisome of *Tolypothrix tenuis* under different light conditions. (*a*) When illuminated by white light, the phycobilisome contains phycoerythrin, phycocyanin, and allophycocyanin. Energy absorbed by phycoerythrin is transferred to phycocyanin and allophycocyanin. The allophycocyanin core proteins are attached, via a linker protein, to the photosynthetic membrane, which is not shown. (*b*) When illuminated by red light, the phycobilisome undergoes complementary chromatic adaptation, in which phycoerythrin is no longer produced but additional phycocyanin is produced. (*After R. MacColl and D. Guard-Friar, Phycobiliproteins, CRC Press, 1987*)

been shown to be essential for coordinating the assembly of the phycobilisome and determining the aggregation state and spectroscopic properties of the phycobiliproteins. Complete phycobilisomes have a molecular weight ranging 7–15 $\times$ 10^6 and contain 300–800 phycobilin chromophores.

Phycoerythrin, phycocyanin, and allophycocyanin are aligned in the phycobilisome in a way that supports directional energy transfer to the photosynthetic reaction center inside the membrane. Allophycocyanin is always confined to the core of the phycobilisome, while phycocyanin and (when present) phycoerythrin are located in the rods. If phycoerythrin is present, it is always located on the end of the rod opposite the attachment to the core. This arrangement allows energy to flow "downhill" from phycoerythrin to phycocyanin to allophycocyanin to chlorophyll *a*. Energy transfer occurs by dipole-dipole interaction between donor and acceptor chromophores and depends on the precise alignment of the phycobilins within the rod and core structures. Light energy absorbed by the phycobilisome is passed primarily to photosystem II, the oxygen-evolving complex of the photosynthetic light reactions.

Physiology. Several environmental factors can influence the phycobiliprotein content. Algal cells grown under low light intensity may have up to 20 times more phycobiliprotein than those grown under high light intensity. This response increases the ability of the alga to absorb light when it is in limited supply. Some cyanobacteria and red algae are also influenced by the color of the growth light and exhibit a phenomenon called complementary chromatic adaptation. For example, when the cyanobacterium *Tolypothrix tenuis* is grown under red light, it produces red-light-absorbing phycocyanin as its accessory pigment. But when grown under green light, it produces green-light-absorbing phycoerythrin along with small amounts of phycocyanin (Fig. 2). This response is controlled by an unidentified photoreversible pigment in a manner similar to the action of phytochrome but with absorption maxima near 545 and 645 nm. This process is regulated at the level of deoxyribonucleic acid (DNA) transcription. *See* PHYTOCHROME.

Phycobiliprotein concentration also depends on the availability of nutrients, including nitrogen, carbon dioxide, phosphorus, sulfur, and iron. Nutrient starvation generally causes a loss of phycobiliprotein, with the rod proteins of the phycobilisomes being lost more rapidly than the core proteins. In cyanobacteria, this reduction is due to specific proteolytic degradation of biliprotein present in the cell and repression of synthesis of new biliprotein.

Isolation. Phycobiliproteins are highly water-soluble and readily leach out of broken cells in low-ionic-strength buffer. They can be purified by using ammonium sulfate precipitation and a wide variety of chromatographic techniques. Phycobilins can be cleaved from the apoprotein by refluxing in methanol or other alcohols or by incubating in concentrated acid, but the two techniques yield different

chromophore structures. Intact phycobilisomes have been isolated from a number of algae by using a higher-ionic-strength buffer that stabilizes the structure. They are then purified by using sucrose gradient centrifugation.

Chemistry. Phycoerythrobilin and phycocyanobilin are structural isomers but differ in the number of conjugated double bonds. This has a primary influence on the color of light absorbed by the molecules. Phycoerythrobilin, with seven alternating single and double bonds, absorbs light of shorter wavelength than phycocyanobilin, which has nine alternating single and double bonds. Free phycobilins in solution generally exhibit broad absorption peaks and weak fluorescence, whereas phycobilins attached to the apoprotein exhibit narrow absorption peaks and strong fluorescence. Binding of the phycobilin to the apoprotein also affects the wavelength of maximum absorption, reduces the ability of the chromophore to complex metals, and increases the stability of the chromophore.

Both phycocyanobilin and phycoerythrobilin are linked to their respective apoproteins through a sulfur atom of the amino acid cysteine. It is generally accepted that at least one binding site on the chromophore is through ring A of the tetrapyrrole. Phycobilins in some PEs may be doubly linked through rings A and D. The phycobilins are derived from 5-aminolevulinic acid, a precursor of cyclic tetrapyrroles such as heme and chlorophyll. Heme, an iron-containing porphyrin that is also found in hemoglobin, undergoes ring opening to form biliverdin. Biliverdin is then enzymatically isomerized and reduced to form phycocyanobilin.

Evolution. Immunochemical and amino acid sequence data have been used to investigate the evolutionary relationships between the different phycobiliproteins. Immunological reactivity analysis indicates consistent relatedness between allophycocyanin, phycocyanin, and phycoerythrin derived from cyanobacteria and red algae. Further, there is no cross reaction between any of the three types of phycobiliproteins. This evidence supports the idea that eukaryotic red algae evolved from the more primitive prokaryotic cyanobacteria. Amino acid sequences of the three phycobiliproteins, however, do exhibit some degree of homology, suggesting that they evolved from a common biliprotein.

Phycobiliproteins from cryptophytes do not show the same high degree of relatedness either to cyanobacteria or to red algae. Cross reactivity is generally greater with red algae than with cyanobacteria, suggesting a closer evolutionary link between cryptophytes and red algae than between cryptophytes and cyanobacteria. Also, cryptophyte phycoerythrins and phycocyanins exhibit cross reactivity, unlike the phycoerythrins and phycocyanins from the other two algal groups.

Biotechnology. Phycobiliproteins can be used as fluorescent labels of specific biological molecules or cells by covalently binding biliproteins to an antibody or other molecule. This is done by placing a reactive site on the antibody using a heterobifunctional reagent, then reacting with a derivatized sulfhydryl group on the biliprotein to form a sulfur bridge. This procedure yields a specific antibody with high fluorescence yield similar to free phycobiliprotein. Different phycobiliproteins, such as B-phycoerythrin, R-phycoerythrin, and C-phycocyanin, can be used to obtain specific absorption and fluorescence emission characteristics. This technique has become particularly important in studying blood disorders, such as acquired immune deficiency syndrome (AIDS). In the latter application, blood cells are treated with specific fluorescent antibodies and screened for the presence of antigen by using fluorescence microscopy or flow cytometry. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); IMMUNOFLUORESCENCE; PHOTOSYNTHESIS. Todd M. Kana

Bibliography. A. N. Glazer, Light guides directional energy transfer in a photosynthetic antenna, *J. Biol. Chem.*, 264:1–4, 1989; T. W. Goodwin, *Plant Pigments*, 1988; R. P. F. Gregory, *Biochemistry of Photosynthesis*, 3d ed., 1989; R. MacColl and D. Guard-Friar, *Phycobiliproteins*, 1987.

Phylactolaemata

A class of the phylum Bryozoa. Phylactolaemates have lophophores (crowns of tentacles surrounding the mouth) which are markedly U-shaped (or rarely nearly circular but still kidney-shaped) in basal outline, and relatively short wide zooecia (exoskeletal living chambers); these animals dwell only in freshwater. *See* BRYOZOA.

Morphology. Phylactolaemate colonies (**Fig. 1***a* and *b*) are either encrusting threadlike networks (*Plumatella*, *Fredericella*, *Hyalinella*) of relatively isolated zooecia with solid chitinous uncalcified walls, or small (*Lophopus*) to large (*Pectinatella*) masses of gelatinous material in which the individual zooids are embedded side by side without definite separating zooecial walls. Stolons (elongated projections of the body wall) are not present. The colonies lack ovicells (globular brood chambers) or monticules (hummocks on the colony surface), they are not polymorphic, and they are not divisible into distinct endozone and exozone regions (the inner and outer zones, respectively, of a massive or erect stenolaemate colony).

Quite simple in construction, phylactolaemate zooids (**Fig. 2**) and zooecia lack diaphragms; the aperture, round in outline and as wide as the zooecium, lacks any operculum (lid)-like covering. Just inside the zooecial wall or the gelatinous material lies an epidermis made of tall thick cells; beneath this is found a thin layer of longitudinal and circular muscles; finally, a peritoneum of flat thin cells is the innermost body-wall layer.

The mouth of a phylactolaemate zooid is overhung by a small lip (epistome). Large passageways directly connect the proximal portions of the visceral cavities of adjacent zooids. The lophophore bears moderate to large numbers of tentacles (20–106, where counted). Individual phylactolaemate zooids are

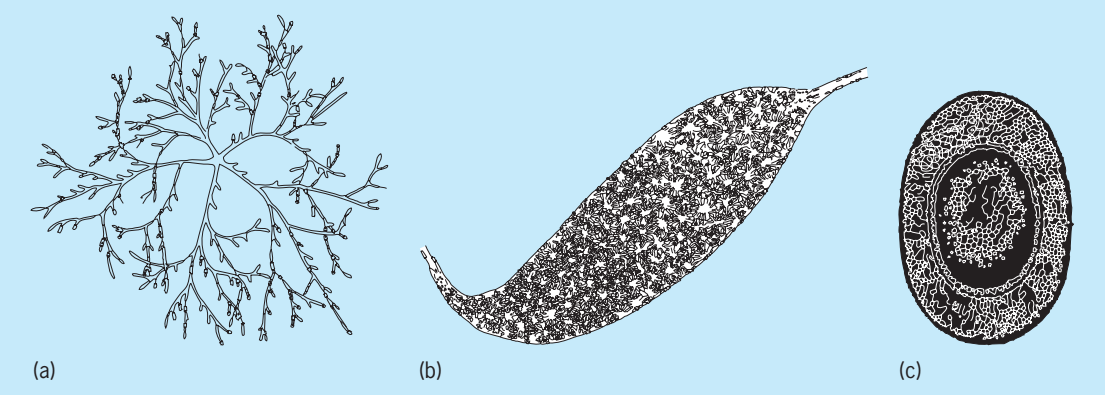

Fig. 1. Phylactolaemate colonies and resting bodies. (*a*) *Plumatella*; encrusting chitinous threadlike network (30 mm or 1.2 in. across) on underside of water-lily leaf/pad. (*b*) *Pectinatella*; large gelatinous mass (180 mm or 7 in. long) on submerged branch. (*c*) *Hyalinella*; tiny free-floating statoblast (1 mm or 0.04 in. long). (*Modified from L. H. Hyman, The Invertebrates, vol. 5, McGraw-Hill, New York, 1959*)

always hermaphroditic, as are consequently the individual colonies. *See* LOPHOPHORE.

Life cycle. Each phylactolaemate zygote develops into only one larva, which is sheltered within a soft internal sac (embryo sac) on the body wall. Freely swimming for less than a day, the larva, after attaching to a substrate, develops directly into an ancestrular zooid without degeneration of its larval structures. New zooids produced by asexual budding begin as sac-like invaginations on the inner side of the body wall of the parent zooid; rudiments of internal soft-part organs (polypide) appear first, followed immediately by formation of new body wall (cystid) materials. Although never degenerating to form brown bodies, phylactolaemate zooids do produce disk-shaped, chitinous-shelled, cold- and drought-resistant resting bodies (statoblasts; Fig. 1*c*), which can germinate or hatch to form new colonies asexually when favorable environmental conditions return. *See* STATOBLASTS.

History and classification. Only about 50–100 phylactolaemate species exist, all classified in a single order, the Plumatellida. Exclusively freshwater, phylactolaemate colonies occur on shallowly submerged tree branches, roots, floating leaves, pond weeds, stones, and debris (like automobile tires), in streams, ponds, and lakes, both natural and artificial. Isolated statoblasts may be found locally along shorelines. Some gelatinous phylactolaemate colonies when broken release neurotoxins which adversely affect fishes nearby. Like other freshwater organisms, many of the particular species of phylactolaemates are widely distributed around the world.

The phylactolaemates may have evolved relatively recently from ctenostomes, although some workers have suggested that phylactolaemates might be very primitive ectoprocts surviving as evolutionary relics. In recent years, the known geologic history of the Phylactolaemata has been extended as far back as the Late Triassic, although fossil statoblasts have long been reported from Pleistocene and postglacial pond (lacustrine) deposits along the glacial boundary in North America. In contrast to the other

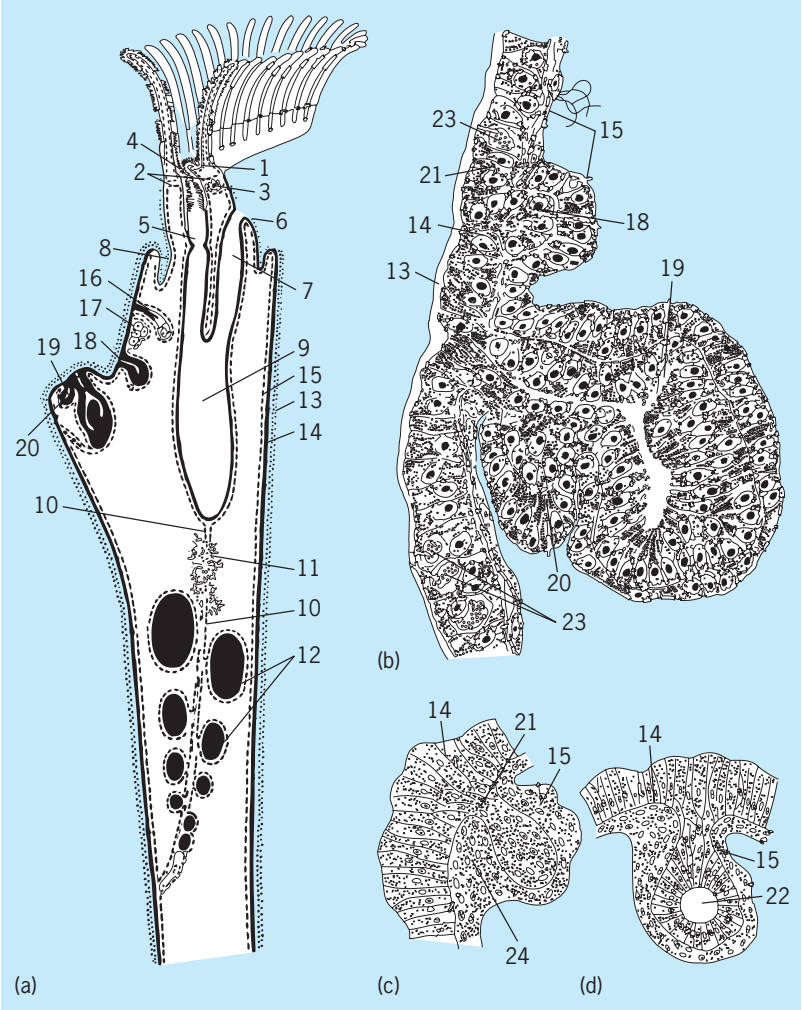

Fig. 2. Asexual reproduction. (*a*) Scheme of a phylatolaemate, showing budding zone and site of statoblast formation. (*b*) Stages of asexual buds of *Plumatella fungosa*. (*c*) Beginning bud of *Pectinatella*. (*d*) Slightly later stage of *c*. 1, epistome; 2, coelomic septum; 3, ganglion; 4, pharynx; 5, esophagus; 6, anus; 7, intestine; 8, vestibule; 9, stomach caecum; 10, funiculus; 11, testes; 12, statoblasts forming; 13, cuticle; 14, epidermis; 15, peritoneum; 16, embryo sac; 17, ovary; 18, adventitious bud; 19, main bud; 20, duplicate bud; 21, epidermal ingrowth; 22, epidermal vesicle; 23, vesicular cells; 24, peritoneal sheet to form funiculus. (*Parts a and b after Brien, 1953*; *parts c and d after Oka, 1891*)

classes of bryozoans (gymnolaemates and stenolaemates), the Phylactolaemata have always been a relatively minor group. *See* CTENOSTOMATA; FRESHWATER ECOSYSTEM; GYMNOLAEMATA; STENOLAEMATA.

Roger J. Cuffey

Bibliography. R. J. Cuffey and J. E. Utgaard, Bryozoans, pp. 204–216, vol. 1, in R. Singer (ed.), *Encyclopedia of Paleontology*, Fitzroy Dearborn, Chicago, 1999; L. H. Hyman, Class Phylactolaemata, pp. 433–498 in L. H. Hyman, *The Invertebrates*, vol. 5: *Smaller Coelomate Groups*, McGraw-Hill, New York, 1959; J. S. Ryland, *Bryozoans*, Hutchinson University Library, London, 1970; P. D. Taylor, Bryozoa, pp. 465–488 in M. J. Benton (ed.), *The Fossil Record 2*, Chapman & Hall, London, 1993; T. S. Wood, General features of the class Phylactolaemata, pp. 287–303 in R. A. Robison (ed.), *Treatise on Invertebrate Paleontology*, pt. G: *Bryozoa*, revised, vol. 1, Geological Society of America & University of Kansas, 1983.

Phyllite

A type of metamorphic rock formed during low-grade metamorphism of clay-rich sediments called pelites. Phyllites are very fine grained rocks with a grain size barely visible in a hand specimen. They have a well-developed planar element called cleavage defined by alignment of mica grains and interlayering of quartz-rich and mica-rich domains. Typically, mica grains show the greater alignment, although other mineral components (quartz, carbonate, and feldspars) may show a preferred shape orientation. Where all minerals of a particular type show the same degree of alignment and the fabric is well developed throughout the rock, the fabric is termed a penetrative fabric. Cleavage surfaces in phyllites have a glittery, lustrous sheen due to light reflecting off grains of chlorite and muscovite. The mineralogy of phyllites is dependent on chemical composition; typical minerals in phyllites are chlorite, muscovite, and quartz. Other minerals that may be present in phyllites formed during low-grade metamorphism include chlorotoid, garnet (rarely), sodium-mica, and sulfide minerals. These minerals are typically millimeter in scale and resolvable in a hand specimen by using a 10× hand lens. *See* CHLORITE; MUSCOVITE; QUARTZ.

Phyllites form during intermediate stages of progressive metamorphism of pelites within the greenschist facies. Initial stages of metamorphism transform pelites into slates, which are characterized by very fine grains, scarcely visible with a microscope. These rocks have a strong cleavage along which the rock is easily split. Phyllites are slightly coarser grained, and the rock cleavage may be slightly irregular because of mineral grains protruding across the cleavage planes. Schists are still coarser and are distinguished by mineral grains that are millimeter to centimeter in scale. Cleavage planes in phyllites usually cut across primary sedimentary layers and, where the rocks are folded, cleavage planes are subparallel to the axial surface of folds. *See* FACIES (GEOLOGY); SCHIST; SLATE.

Cleavage formation in phyllites involves the interaction between several different processes, including mechanical rotation of minerals; mechanical rotation of minerals crystallizing during metamorphism; preferred growth of minerals at the time of crystallization; and change in shape of mineral grains during deformation and metamorphism, possibly as a result of mineral dissolution.

Phyllite is found in most regionally metamorphosed terranes in the world, including the Appalachians of eastern North America, the Scottish Highlands, and the Alps. *See* METAMORPHIC ROCKS.

Matthew W. Nyman

Phyllocarida

A subclass of the crustacean class Malacostraca containing the extant order Leptostraca and the fossil order Archiostraca (see **illus.**). The Phyllocarida has a long fossil record, and many early fossil taxa were referred to this subclass. However, studies of presumed phyllocarids from the Burgess Shale have shown that only the archiostracans agree with the definition of the Phyllocarida. The apparent antiquity of the Phyllocarida, together with the foliacious structure of their thoracic appendages, gave rise to the hypothesis that fossil phyllocarids encompassed the ancestral stem of the Eumalacostraca. However, eumalacostracan stocks never show such phyllocarid features as thoracic filter-feeding mechanisms and a brood pouch formed with the setose tips of the thoracic endopods. Thus a consensus has been reached that the Phyllocarida represents an early, and perhaps dead-end, offshoot of the Malacostraca rather than the stem group for eumalacostracan evolution.

Phyllocarids are distinct from other malacostracan crustaceans because of two other characteristics considered to reflect the primitive condition, which strengthen the hypothesis of early separation from the main evolutionary line. The first is the presence of a bivalve carapace. The dorsal carapace hinge seen in archiostracans is absent in leptostracans, but both orders have the characteristic adductor muscle that holds the two sides together. The second is an abdomen consisting of seven fully formed somites and terminating in a telson that bears caudal rami. Only the anterior six somites bear

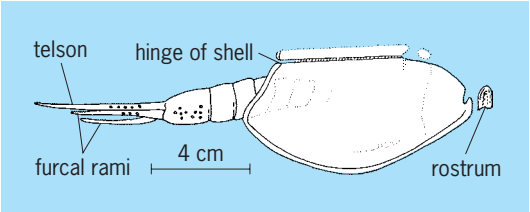

***Ceratiocaris stygia*, a fossil leptostracan of the subclass Phyllocarida. (*After T. R. Jones and H. Woodward, A Monograph of British Palaeozoic Phyllopoda (Phyllocarida Packard), Palaeontographical Society, 1888–1889*)**

appendages. *See* CRUSTACEA; LEPTOSTRACA; MALACOSTRACA. Patsy A. McLaughlin

Bibliography. E. Dahl, The subclass Phyllocarida (Crustacea) and the status of some early fossils: A neontologist's view, *Vidensk. Meddr. Dansk Naturh. Foren.*, 145:61–76, 1984; J. K. Lowry (ed.), *Papers from the Conference on the Biology and Evolution of Crustacea*, Austral. Mus. Mem. 18, 1983; F. R. Schram, Crustacean phylogeny, *Crustacean Issues*, vol. 1, 1983.

Phylogeny

The genealogical history of organisms, both living and extinct. Phylogeny represents the historical pattern of relationships among organisms which has resulted from the actions of many different evolutionary processes. Phylogenetic relationships are depicted by branching diagrams called cladograms, or phylogenetic trees. Cladograms show relative affinities of groups of organisms called taxa. Such groups of organisms have some genealogical unity, and are given a taxonomic rank such as species, genera, families, or orders. For example, two species of cats—say, the lion (*Panthera leo*) and the tiger (*Panthera tigris*)—are more closely related to each other than either is to the gray wolf (*Canis latrans*). The family including all cats, Felidae, is more closely related to the family including all dogs, Canidae, than either is to the family that includes giraffes, Giraffidae. The lion and tiger, and the Felidae and Canidae, are called sister taxa because of their close relationship relative to the gray wolf, or to the Giraffidae, respectively.

Cladograms thus depict a hierarchy of relationships among a group of taxa (**Fig. 1***a*). Branch points, or nodes, of a cladogram represent hypothetical common ancestors (not specific real ancestors), and the branches connect descendant sister taxa. If the taxa being considered are species, nodes are taken to signify speciation events. The goal of the science of cladistics, or phylogenetic analysis, is to discover these sister-group (cladistic) relationships and to identify what are termed monophyletic groups—two or more taxa postulated to have a single, common origin.

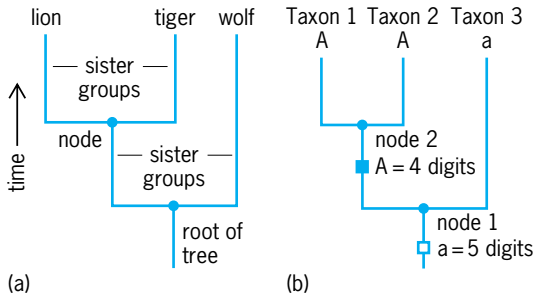

Fig. 1. Phylogenetic trees. (*a*) A tree representing a hierarchical pattern of sister-group relationships (those taxa descended from a common ancestor). (*b*) Relationships are determined by identifying derived characters; in this case the condition of five digits is primitive, whereas the loss of a digit is derived and unites taxa 1 and 2.

Methodology. The acceptance of a cladogram depends on the empirical evidence that supports it relative to alternative hypotheses of relationship for those same taxa. Evidence for or against alternative phylogenetic hypotheses comes from the comparative study of the characteristics of those taxa. Similarities and differences are determined by comparison of the anatomical, behavioral, physiological, or molecular [such as deoxyribonucleic acid (DNA) sequences] attributes among the taxa. A statement that two features in two or more taxa are similar and thus constitute a shared character is, in essence, a preliminary hypothesis that they are homologous; that is, the taxa inherited the specific form of the feature from their common ancestor. However, not all similarities are homologs; some are developed independently through convergent or parallel evolution, and although they may be similar in appearance, they had different histories and thus are not really the same feature. In cladistic theory, shared homologous similarities are either primitive (plesiomorphic condition) or derived (apomorphic condition), whereas nonhomologous similarities are termed homoplasies (or sometimes, parallelisms or convergences). This distinction over concepts and terminology is important because only derived characters constitute evidence that groups are actually related.

As evolutionary lineages diversify, some characters will become modified. Examples include the enlargement of forelimbs or the loss of digits on the hand. Thus, during evolution the foot of a mammal might transform from a primitive condition of having five digits to a derived form with only four digits (Fig. 1*b*). Following branching at node 1, the foot in one lineage undergoes an evolutionary modification involving the loss of a digit (expressed as character state A). A subsequent branching event then produced taxa 1 and 2, which inherited that derived character. The lineage leading to taxon 3, however, retained the primitive condition of five digits (character state a). The presence of the shared derived character, A, is called a synapomorphy, and identifies taxa 1 and 2 as being more closely related to each other than either is to taxon 3. Distinguishing between the primitive and derived conditions of a character within a group of taxa (the ingroup) is usually accomplished by comparisons to groups postulated to have more distant relationships (outgroups). Character states that are present in ingroups but not outgroups are postulated to be derived. To determine relationships among the species of cats (family Felidae), comparisons to characters in other carnivore outgroups such as dogs (Canidae) or bears (Ursidae) would help distinguish primitive and derived conditions within cats.

Systematists have developed computer programs that attempt to identify shared derived characters (synapomorphies) and, at the same time, use them to construct the best phylogenetic trees for the available data. Conceptually, the two procedures—determining the polarity of character transformations and finding the best phylogenetic tree—are logically related (**Fig. 2**). For example, three alternative

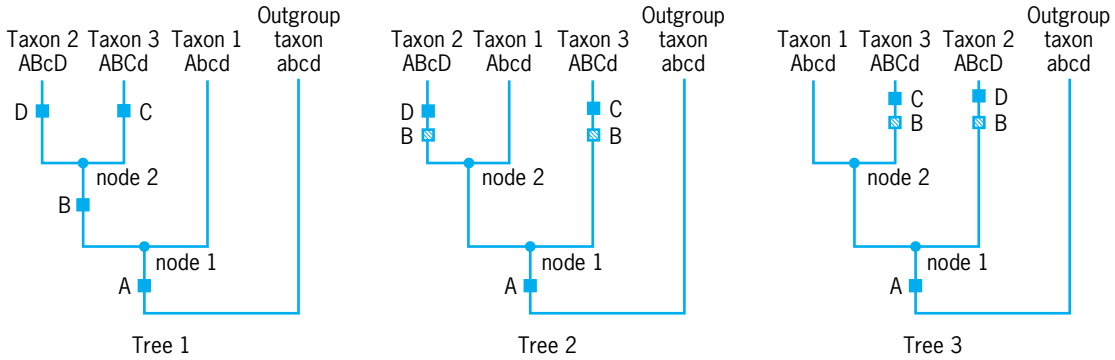

Fig. 2. Alternative trees for three taxa. On tree 1, taxa 2 and 3 are united by a postulated synapomorphy, B. On trees 2 and 3, no taxa are united by a synapomorphy and B evolves in parallel (cross-hatched). Tree 1 assumes fewer character-state transformations for all the data and is thus preferred.

phylogenetic trees are possible for three taxa. A comparative analysis of the three might reveal that they share four characters, A–D, each of which has two alternative states. Comparisons to one or more outgroup taxa help the systematist interpret the polarity of these character states (in this example, uppercase letters are taken to be the derived conditions). Given the distributions of the character states shown in Fig. 2, postulated character transformations can be assigned to the branches of each alternative tree. In all three trees, the derived condition, A, unites the three ingroup taxa but cannot help choose among the three alternatives for their interrelationships. When the second character is examined, it is seen that the postulated derived condition, B, unites taxa 2 and 3 to the exclusion of taxon 1, which has the primitive condition, a. The derived conditions of the third and fourth characters are each found in one taxon only: C in taxon 3 and D in taxon 2, respectively. Thus, only the transformation from b to B carries phylogenetic information. In addition, tree 1 also better explains the distributions of the different character states because it requires that b transform to B only once, whereas trees 2 and 3 require two independent origins of B from b. Thus, tree 1 is the simplest hypothesis given the available data and is to be preferred. Similarly, computer programs can process data sets having large numbers of taxa and characters, examine all possible character transformations on a set of trees, and then choose that tree which accounts for the fewest number of changes.

The interpretation of the status of a character shared among taxa depends upon an assessment of its history on a given tree. Thus, in tree 1 the derived character B is interpreted as a synapomorphy (it unites taxa 2 and 3); that is, B is inferred to be homologous in the two taxa. Within the framework of threes 2 and 3, however, the derived character B is interpreted as having evolved independently in taxa 2 and 3. In this case, character state B is interpreted as being homoplasious (parallel) and thus not homologous.

Value of phylogenies. Understanding how species and groups of species are interrelated is fundamental to interpreting observations across all of biology. Knowledge of phylogenetic relationships provides the basis for classifying organisms. Ideally, all taxonomic groups—genera, families, orders, and so on—should be monophyletic. Some groups, such as birds and mammals, are monophyletic; that is, phylogenetic analysis suggests they are all more closely related to each other than to other vertebrates. However, other traditional groups, such as reptiles, have been demonstrated to be nonmonophyletic (some so-called reptiles, such as dinosaurs and their relatives, are more closely related to birds than they are to other reptiles such as snakes). A major task of the science of systematics is to search for monophyletic groups. Classifications based on monophyletic groups are termed natural classifications.

Phylogenies are also essential for understanding the distributional history, or biogeography, of organisms. Knowing how organisms are related to one another helps the biogeographer to decipher relationships among areas and to reconstruct the spatial histories of groups and their biotas. *See* BIOGEOGRAPHY.

Biologists use phylogenies to make predictions about the characteristics of organisms that are still poorly studied. If one species produces a particular compound that has useful pharmacological properties, it is probable that its close relatives will share an identical or similar compound. The same is true for other structural, behavioral, or physiological characters.

History of life. Life has diversified on Earth for about 4 billion years. Systematists have described nearly 1.4 million living species, most of which are arthropods (insects, spiders, and their relatives), as well as tens of thousands of fossil species. It is estimated that from 3 million to perhaps as many as 100 million living species remain to be discovered, and the number of unknown fossil species is no doubt equally vast.

Much knowledge about the history of life has been derived from studies of comparative anatomy. Advances in understanding the history life have been made by using information derived from comparisons of DNA sequences, which largely involves diversification within the microbial world. It is often believed that fossils represent the direct ancestors

of other fossils or of living organisms and, therefore, that fossils are necessary to reconstruct phylogeny. Because of the vast panoply of taxonomic and structural diversity seen throughout deep evolutionary time, reconstructing its history is a complex undertaking, and many aspects are still not well understood. Many major groups of organisms appear to have diversified in an evolutionary burst about 600 million years ago. Rapid radiations, especially those taking place long ago, pose particular difficulties for reconstructing phylogeny, for often relatively few morphological and molecular characters evolve between branching events that are so close to one another in time. However, there has been an upsurge in interest in reconstructing phylogeny, so knowledge of the history of life should improve substantially. *See* ANIMAL EVOLUTION; SYSTEMATICS; TAXONOMIC CATEGORIES.

Joel Cracraft

Bibliography. D. R. Brooks and D. H. McLennan, *Phylogeny, Ecology, and Behavior*, 1991; N. Eldredge and J. Cracraft, *Phylogenetic Patterns and the Evolutionary Process*, 1980; B. Fernholm, K. Bremer, and H. Jornvall (eds.), *The Hierarchy of Life*, 1989; P. L. Forey et al., *Cladistics: A Practical Course in Systematics*, 1992; A. B. Smith, *Systematics and the Fossil Record: Documenting Evolutionary Patterns*, 1994.

Phymosomatoida

An order of regular sea urchins, class Echinoidea, characterized by imperforate tubercles, complex ambulacral compounding in which one or more elements are occluded from the perradial suture, and a stirodont lantern with unfused epiphyses. They comprise two families, one with crenulate and the other with noncrenulate tubercles. They first appeared in the Lower Jurassic and are probably paraphyletic, since they include the ancestors of camarodonts. The two extant genera are each known from a single species: *Glyptocidaris*, from a depth of 30–490 ft (10–150 m) around northern Japan, and *Stomechinus*, a common inhabitant of rocky shores around the Indo-West Pacific. They are epifaunal grazers. *See* ECHINODERMATA; ECHINOIDEA.

Andrew Smith

Physical anthropology

The subfield of anthropology that deals with human and nonhuman primate evolution, the biological bases of human behavior, and human biological variability and its significance. Some refer to the field as biological anthropology to signal the close links with other biological sciences. The term physical anthropology is largely an American and British invention; in most European and many other countries physical anthropologists are the only anthropologists, while persons who study behavioral aspects of the human condition are known as archeologists, ethnologists, linguists, or prehistorians. In Great Britain, and increasingly in the United States and elsewhere, the term human biology is also used instead of physical anthropology, especially in relation to curricula in the health sciences.

Paleoanthropology. Paleoanthropology is the multidisciplinary study of human evolution as evidenced by fossils, artifacts, and their geological and burial site contexts. Physical anthropologists organize expeditions and direct excavations that lead to discoveries of fossil Hominidae, and then engage in the painstaking repair and reconstruction of specimens, their anatomical description, comparison with other specimens, and placement in hominoid phylogeny. (Hominids are any species of the primate family Hominidae, consisting of the humanlike species. Hominoids are any species of the primate superfamily Hominoidae, consisting of hominids plus apes.)

Morphological paleoanthropologists must have detailed knowledge of human and other primate anatomy and the principles of taxomony in order to restore and interpret their discoveries. In addition to traditional anatomical descriptions and measurements, it is requisite that the variations in samples be presented and, when possible, be tested for statistical significance. Because of the fragmentary nature of many fossils, one of the most difficult problems is to decide whether the new discoveries belong with previously described species or represent new undescribed ones. For instance, D. C. Johanson initially announced that several hominid species were represented by a small sample of specimens from the Pliocene Hadar Formation in northern Ethiopia. But after many more complete specimens were collected and compared, he and T. D. White concluded that only one species, *Australopithecus afarensis*, inhabited the site. Contrarily, P. V. Tobias concluded that the collection could be accommodated in *A. africanus*, which had been described previously from South Africa. Y. Coppens suggested that there are two species of Hominidae at Hadar, including one of *Homo*. Currently, *A. afarensis* is generally accepted to comprise most of the 2.9–3.6-Ma (million-year-old) skeletal specimens from Laetoli, Tanzania, Hadar, and several other sites in eastern Africa. The 3.5-Ma trails of humanoid footprints might have been made by a species other than *A. afarensis*, in which case there were at least two species in the area then.

A trove of more recent discoveries has extended the variety, temporal depth, and geographic range of potential and actual early hominids, including *Sahelanthropus tchadensis* [approximately 7 Ma, Chad], *Orrorin tugenensis* [approximately 6 Ma, Kenya], *Ardipithecus ramidus* [4.4–5.8 Ma, Ethiopia], *Australopithecus bahrelghazali* [3–3.5 Ma], *A. anamensis* [3.9–4.2 Ma, Kenya], *A. garhi* [2.5 Ma, Ethiopia], *Kenyanthropus platyops* [3.5 Ma, Kenya], and *Paranthropus aethiopicus* [2.2–2.8 Ma, Ethiopia and Kenya]. Further discoveries of *A. africanus* [2.3–3.0 Ma, South Africa], *Paranthropus robustus* [1.2–2 Ma, South Africa], and *P. boisei* [1.4–2.5 Ma, eastern Africa] have added to knowledge of the physical variety of early hominids, and new species and expanded samples

of *Homo* have been forthcoming, including *Homo habilis* [1.6–2.3 Ma, eastern Africa], *H. rudolfensis* [1.8–2.4 Ma, eastern Africa and Malawi], *H. ergaster* [1.5–1.9 Ma, Kenya, Georgia?], *H. antecessor* [.78–1 Ma, Spain, Eritrea], *H. heidelbergensis* and *H. erectus* [≤.1?–1.9 Ma, Asia and perhaps Europe and Africa], and *H. neanderthalensis* [≥.025–≥.13 Ma, Europe and western Asia].

Proper geophysical dating is essential to place fossils in the larger picture of human evolution. If there is no datable mineral associated with a hominid fossil, paleontologists may be able to link associated animal species from new sites with those of geochemically dated sites elsewhere.

In order to depict the ways of life of human ancestors, carefully mapped contextual information from fossil sites is needed, including plant and animal remains that indicate the climate, habitat, and availability of food and water, and the artifacts and archeological features that show how the hominids might have processed food, protected themselves from predators, and sheltered themselves from daily and seasonal vicissitudes.

The densities, age profiles, and morphology of individuals at hominid sites provide clues to how successful they were reproductively in comparison with modern humans and other primates. The hominid fossils may show evidence of violence, cannibalism, or gnawing by animals. The teeth and other cranial structures may indicate crudely the texture of ingested foods, and fossil feces (coprolites) may provide direct dietary information. The content and ratios of stable isotopes of certain elements in skeletons provide clues to the paleoenvironment and diet of prehistoric people and the animals with which they lived. For example, the amount of carbon isotope 13 [^{13}C] in bone collagen and tooth enamel may indicate whether an individual ingested tropical or savanna grasses (or animals that ate them) versus trees, shrubs, or temperate grasses. Further, nitrogen isotope ratios may indicate the extent to which an individual was carnivorous and whether the protein came from marine or terrestrial animals. *See* DATING METHODS; FOSSIL; FOSSIL HUMANS; RADIOCARBON DATING; PALEOECOLOGY.

Paleoprimatology. Physical anthropologists also look to nonhuman primates for clues to human physical history and status as mammals, and for analogies to the behavior and cognitive abilities of human ancestors. Like paleoanthropologists, paleoprimatologists employ methods of other paleobiologists to collect, describe, and interpret fossil specimens phylogenetically and functionally. Contextual information, including the manner in which specimens were deposited and their possible alteration over time, is vital. Paleoprimatologists usually do not have to contend with archeological materials but are faced with an enormous time span, from the Late Cretaceous period, 75 Ma, to the beginning of the Holocene epoch, 10,000 years ago.

Numerous fossils from Eocene and Oligocene deposits (33–36 Ma) in the Egyptian Fayum show that the earliest ancestors of catarrhine primates (apes, hominids, and Old World monkeys) were similar to the highly arboreal New World monkeys in many features. Further, a good case has been made that Old World monkeys do not represent a stage in the evolution of the Hominoidea but instead derived from a hominoidlike common ancestor. *See* FOSSIL PRIMATES.

Comparative primate morphology. Recent primates, including humans, are the current end products of evolution. Anatomical studies of modern primates are essential to model the morphologies of fossil forms and to reveal their singularity. Carefully controlled comparative studies on extant primates hold the greatest promise for modeling the functional morphology, physiology, and habitat preferences of extinct forms. Since 1970, advanced techniques and concepts from rehabilitation medicine, kinesiology, orthopedics, orthodontics, radiology, and neurology have been used to establish a truly functional morphology of the primates. *See* PRIMATES.

Locomotor anatomy. Electromyography is used to document the activities of specific muscles in nonhuman primates as they engage in a variety of positional behaviors (postures and locomotion). Classic inferences from dissections of cadavers have had to be revised because of electromyographic research. For instance, it has been shown that the bulky upper limb muscles of apes are not active during quiet suspension from overhead supports; instead, special features of their joints are the chief supporting structures. *See* ELECTROMYOGRAPHY.

Cineradiography (x-ray motion pictures) reveals otherwise inaccessible details about the movements of bones and other structures in joints during regular and facultative locomotion. For example, cineradiographic films of bipedal chimpanzees have been used to interpret features of australopithecine (early hominid) limb bones. *See* RADIOGRAPHY.

Force plates record the directional forces of primate feet and hands as they walk over them, which can be compared interspecifically to show, for instance, how closely nonhuman primates approach human patterns of walking. High-speed motion pictures and computer analyses of limb segment positions from films and video tapes are providing details on the ways that primates move on different substrates. Comprehensive gait and kinematic analyses of many living species may serve as a solid basis for modeling hypothetical locomotor patterns of fossil species.

Cranial and dental anatomy. It is no longer common to assign each new fossil primate tooth or bit of skull to a new species, largely because studies on the crania, jaws, and teeth of living primates have shown that considerable variation is the rule even within a single species.

Electromyography and the use of strain gages to study cranial structures in several primate species have helped to explain the mechanics of chewing and other features related to processing foods. These studies help to identify probable (and to rule out unlikely) feeding habits of fossil forms. They also assist in the selection of animal models for experiments

that will contribute to the development of more efficient human orthodontic devices and procedures. *See* BIOMECHANICS.

Scanning electron microscopy, x-rays, serial tomographs, and enlarged photographs reveal details of dental cusp, root, and crest size and shape and abrasive patterns that are correlated with particular diets in living primates and other mammals. This information is valuable for inferring the diets of fossil species.

Neuroanatomy. Although bipedal habit and manipulatory hands set humans apart from other primates, it is the large brain that is usually held to be the great advancement of humankind. Hence considerable research is devoted to discerning the differences between human and nonhuman central nervous systems and to documenting the changes in brain size and shape during anthropoid phylogeny.

Like other areas of primatology, neuroanatomy is not the sole province of anthropologists. Indeed only since about 1960 have physical anthropologists centered research careers within it. The steady accumulation of fossil hominid endocasts (produced by minerals that filled the braincase after death and hardened to produce a mold of it) and crania from which latex endocasts can be made have attracted anthropologists to join other scientists in the quest to elucidate human brain evolution. A theory was developed indicating that the major change from ape to hominid brains was one of internal reorganization of the central nervous system and that this preceded the dramatic increase in gross size. A few workers also believe that speech areas can be identified on the endocasts of *Homo habilis*, a species that dates to approximately 2.3 Ma.

Anthropologists are especially keen to learn the extent to which nonhuman primates have cerebral asymmetries that correlate with behavioral traits, whether any of these may provide clues for the evolution of human language, and the extent to which expansion of the brain is a mere consequence of increased body size in the Hominidae. Apes and some monkeys also exhibit differences in size between the left and right hemispheres, but links to handedness and other behaviors remain elusive. Because of their costliness and for sound humanitarian and conservational reasons, the closest primates to humans (chimpanzees, gorillas, and orangutans) are not likely to be used in drastic experiments that could inform questions about the meanings of their asymmetries. The failure of apes to learn to speak, despite intensive efforts to teach them to do so, indicates that they lack the neurological basis for human speech production, though they can comprehend human utterances. *See* APES.

Molecular anthropology. The explosion of molecular biology that followed the cracking of the genetic code attracted physical anthropologists who wished to test hypotheses about the propinquity of humans with the apes, and the relationships of other primates to one another and to other creatures. Some have endeavored to find "molecular clocks" that could tell when species diverged from one another. Assuming that genetic changes occur at fairly steady rates and given a few well-dated fossils, the time when the living species may have branched from one another can be estimated. A molecular clock has been used to judge which fossils were part of hominid and other lineages. The studies have been invaluable for excluding monkeys and prosimians from immediate hominid ancestry, and for pinpointing the time interval during which the hominids diverged from apes (12–5 Ma). However, they have inspired a variety of heterodox taxonomic schemes which have African apes or even all apes in the Hominidae. It will probably be some time before compelling paleontological and genetic discoveries will lead to resolution of the controversies that surround these studies. For example, M. Goodman has steadfastly argued that molecular evolution slowed down in the hominid lineage, and hence that the clocks of V. Sarich and other scientists are unreliable. Further, little is known about the relationships between molecules and the morphological features that distinguish the apes and humans. *See* MOLECULAR ANTHROPOLOGY; PROTEINS, EVOLUTION OF.

Primate behavior and ecology. Although firmly rooted in comparative psychology, behavioral primatology has also become a major section of physical anthropology; anthropologists have contributed mainly through field studies.

An original goal of primate field studies by anthropologists was to find predictable relationships between specific habitats and the patterns of sociality of the primates that inhabit them. This, it was thought, would allow the identification of the social organization of fossil species for which there is information of the paleohabitats. Perhaps it could be learned whether ancestral hominids were monogamous or polygamous and whether one sex dominated the other and served as the stable core in social groups. As commonly occurs in behavioral science, however, the greater the variety of focal studies and the longer the span of observations, the fewer generalizations pertaining to the human career and condition that could be sustained.

Because of their close genetic relationships to humans, chimpanzees have emerged as the most popular model for early hominid sociality. But theorists emphasizing different aspects of their behavior arrive at markedly different models for early hominid behavior. For instance, while one worker has stressed the importance of female choice and maternal behavior in a flexible social system, another has placed the males fully in charge of early hominid monogamous units.

Because many primate species develop relatively slowly and can live for a long time, the most informative studies have notable temporal depth at specific field sites or in captive research settings, where the subjects are somewhat free-ranging or at least socially housed. Only then can one gain clear understanding of the overall social structure and relationships among males, among females, and between the sexes; how youngsters come to exhibit the traits of their elders; whether there are behaviors that vary between populations of a species that emulate

human cultural variation; how mating decisions are made; how individuals communicate with one another vocally and with gestures; how disputes are resolved and whether there are mechanisms for reconciliation among antagonists; why some species commit infanticide and lethal attacks against conspecifics and perhaps engage in cannibalism; whether foods are shared in a group and among which individuals; and how each species fits into the overall ecology of the regions where they are natural components of particular habitats.

The efforts of primate field behavioralists, and ecologists in general, will have to be directed increasingly toward conservation and gaining appreciation for their subjects by local people and government officials, because habitat destruction is occurring at alarming rates in the tropics. Because primates compete for natural resources and serve so well as models for many human diseases and physiological processes, hard choices have to be made between their conservation and the improvement of human lives. Their utility as models in medical science increases greatly as knowledge of their natural adaptations accrues. *See* BEHAVIORAL ECOLOGY; SOCIOBIOLOGY.

Human variation. The term human variation is rapidly replacing its historical predecessor "race" in anthropology because the latter carries so much negative connotation. Many scientists believe that the concept of race should be abandoned. Instead, researchers should simply record the gene frequencies and biological traits of human populations that are otherwise identified only by their geographic localities. This genotypic and phenotypic information would be interpreted in terms of historical and proximate selective forces in each environment. For example, the Asian sources of Paleoindian populations that immigrated to the New World, and the deployment of their successors in North America and South America over the past 12,000 years, may be sought by the modern population biological approach. *See* HUMAN GENETICS.

According to the modern perspective, people should no longer be labeled Caucasoid, Mongoloid, Negroid, Australian, Nordic, Black, Brown, White, and so forth. While this may cause problems in communication as anthropologists attempt to apply their knowledge practically to forensic and medical problems, in the long run humanity will benefit from the exclusion of racial typologies. *See* HUMAN BIOLOGICAL VARIATION.

Skeletal biology. Though this field became moribund because of its largely descriptive nature and long history of abuse by racial typologists, some creative skeletal biologists have adopted functional approaches from primatological anthropology, and have been discerning the genetic determinants of nonnumeric (discrete) and measurable (continuous) traits that would allow them to document the local history of burial populations in an area. For example, in a study of one macaque population, the history of which was well documented, the validity of nonmetric skeletal traits for revealing genetic links within and between populations was established. This permitted others to use cremated and other fragmentary human cemetery remains to establish the history of Native Americans in the Lower Illinois River Valley.

Paleopathology. Human bones and teeth sometimes reflect diseases and mechanical trauma during the lifetime of an individual. Rehydrated soft tissues of mummies may show evidence of parasitic infestations and lesions, such as those from cancer and tuberculosis. Such studies allow reasonable correlations of disease features with demography, ecology, diet, and social factors.

Unfortunately, most diseases cannot be specifically identified from skeletal remains, though improved chemical and microscopic techniques now reduce the number of possible conditions in many cases. The original hope of skeletal biologists that the places or origin and subsequent spread of ancient diseases, such as syphilis, could be documented now seems illusory.

The roots of medical practice itself are found in prehistory via evidence such as set fractures and trephined (surgically perforated) skulls. Whether the wounds were fresh, infected, or healed when the person died can be determined by the state of bones from the injured area. It is also possible to detect caretaking behavior by other group members toward afflicted individuals from the extent of their treatment and evidence of survival despite severe or repeated injuries. *See* PALEOPATHOLOGY.

Skin and hair. Racial typologies were not built from bones alone. Skin color and hair texture were also fundamental to historical arrangements of humankind into subgroups. In the modern view, however, skin color is chiefly of medical concern, such as for predicting susceptibility to melanomas. Further, the evolutionary problem of how and why light and dark epidermal pigmentation arose is of interest, and has produced hypotheses relating to human physiology and ecology. Dark skin, including that produced by tanning, may prevent hypervitaminosis D from overexposure to sunlight. Light skin may allow sufficient vitamin D to be metabolized by people exposed to low sunlight. There has also been the suggestion that epidermal structures functioned as sexually attractive features in human and prehuman reproductive behavior, but the true extent of such effects is not known.

Fingerprints failed to confirm racial typologies because they are too individualistic. Instead, they have been a boon to forensic and pediatric scientists. *See* FINGERPRINT.

Growth, physique, and aging. The field of growth studies has developed robustly as part of physical anthropology since the 1940s. Standards for the appearance and ossification of bones and for sexual maturation have been established so that congenital, nutritional, and other environmental effects can be detected and often corrected clinically in children and adolescents. It has been established from global nutritional surveys that small adult size is correlated with dietary insufficiency. The field promises to grow further as longitudinal studies are established with

more populations and as the nutritional status and general medical condition of people improve. Since 1975, physical anthropologists have begun to apply anthropometric and microscopic techniques to the study of aging in an effort to understand why some people have greater longevity. *See* AGING.

Nutritional anthropology. Refined chemical assays promise to reveal past diets from hominid skeletal remains. Hence paleonutrition has emerged as a special focus for technical research. Trace elements, such as strontium (Sr), sodium (Na), zinc (Zn), and calcium (Ca), may indicate the nutrients that were incorporated in bones, so that classes of food items can be inferred. For instance, because strontium from plants is taken up by bones much more than by muscles, the extent to which people were herbivorous (high Sr) versus carnivorous (low Sr) can be assessed in burial populations. Unlike strontium, zinc occurs at higher levels in meat than in leafy plants. Thus it can be used to check results from strontium assays, provided that there is no evidence for dietary nuts or mollusks, which also have relatively high levels of zinc. $^{13}C/^{12}C$ ratios in human bones have been an important tool to document the spread of maize agriculture in the New World. Protein and vitamin D deficiencies are indicated by flatness of the skull base relative to the skulls of better-nourished people. *See* NUTRITION.

Blood group genetics and disease. Many human features such as stature, head shape, epidermal pigmentation, and fingerprints are caused by undetermined multifactorial genetic systems whose phenotypic expression is commonly affected by environmental factors. In contrast, the genetics of blood group systems (such as ABO, MNS, and Rh) are based on single gene loci and are well understood.

Susceptibility to a specific disease has not been definitely linked to a particular human blood type, but there are numerous, usually deadly, single-locus diseases that afflict humans. Cogent arguments have been advanced that the gene for sickle-cell anemia was maintained in human populations because, when paired with a nonsickling gene, producing a heterozygous state, it conferred resistance to malaria. If the gene is paired with its counterpart, the resulting homozygous condition is lethal. The relationship of other genetic polymorphisms with diseases is much less well documented. Diseases are obvious candidates to act as selective agents for human blood groups. Alternatively, the distribution of blood groups in human populations could be the result of random factors, such as genetic drift and recombination. Compelling evidence that would resolve the selection-drift controversy is not available. *See* BLOOD GROUPS; POPULATION GENETICS.

Human adaptability and ecology. *Homo sapiens* is one of the most versatile species on Earth. The invention and ramification of culture has permitted people to survive on impoverished islands and in climatic extremes of high altitude, deserts, and polar regions. In addition to many clever technological and social conventions, people show physiological and ontogenetic characteristics that enhance their survival and ability to work in harsh environments. The extent to which these features are genetically determined is still largely unknown.

Human ecologists need not focus on exotic people in extreme climates. Urban settings offer challenges to anthropologists who would understand the interaction of human biology and culture in adapting to dietary insufficiency and oversufficiency; new and unconquered diseases; air, noise, and water pollution; and the sheer number and variety of fellow urbanites and their pets. Nutritional factors and different histories of disease, however, make control of comparisons between populations problematic. *See* HUMAN ECOLOGY.

Forensic anthropology. Forensic anthropology is growing quickly, as anthropologists are called as expert witnesses regarding not only classic sorts of criminal evidence (fingerprints, blood types, and skeletal remains) but also grisly exhibits such as bloody footprints and bite marks on murder victims. Forensic anthropology grew out of the need to identify the war dead and individuals who are killed in domestic disasters such as airplane crashes. In criminal cases, forensic anthropologists operate best if contextual evidence is preserved, particularly when the victims were hidden in graves. Here the techniques of field archeology can be as important as those of physical anthropology. *See* ANTHROPOLOGY; ARCHEOLOGY; FORENSIC ANTHROPOLOGY. Russell H. Tuttle

Bibliography. L. Aiello and C. Dean (eds.), *An Introduction to Human Evolutionary Anatomy*, 1990; B. G. Campbell and T. D. Loy, *Humankind Emerging, The Concise Edition*, Allyn & Bacon, Boston, 2002; L. L. Cavalli-Sfordza, P. Menozzi, and A. Piazza, *The History and Geography of Human Genes*, 1994; R. L. Ciochon and J. G. Fleagle, *Human Evolution Source Book*, 1993; F. E. Johnston (ed.), 1930–1980 Jubilee Issue, *Amer. J. Phys. Anthropol.*, 56:327–557, 1981; R. Jurmain et al., *Introduction to Physical Anthropology*, 9th ed., Wadsworth/Thomson, 2003; S. Mader, *Human Biology*, 7th ed., 2001; P. L. Stein and B. M. Rowe, *Physical Anthropology*, 7th ed., 1999; R. H. Tuttle, *Apes of the World*, Noyes, Park Ridge, NJ, 1986; R. H. Tuttle, D. P. Buxhoeveden, and G. W. Cortright, Anthropology on the move: Progress in experimental studies of nonhuman primate positional behavior, *Yearbook of Physical Anthropology*, 22:187–214, 1979.

Physical chemistry

The branch of chemistry that deals with the interpretation of chemical phenomena and properties in terms of the underlying physical processes, and with the development of techniques for their investigation. Chemical physics refers to a branch of physical chemistry in which the emphasis is on the interpretation and analysis of the physical properties of individual molecules and bulk systems, instead of their reactions. Theoretical chemistry is another major branch, in which the emphasis is on the calculation of the properties of molecules and systems,

and which uses the techniques of quantum mechanics and statistical thermodynamics. For the present purpose it is convenient to regard physical chemistry as dealing with three aspects of matter: its equilibrium properties, the structure of molecules and bulk matter, and matter's ability to change. *See* PHYSICAL ORGANIC CHEMISTRY; QUANTUM MECHANICS.

Chemical thermodynamics. The study of matter in a state of equilibrium constitutes the field of chemical thermodynamics. In particular, chemical thermodynamics provides a technique for discussing the response of a system to a change in the external conditions (such as the change in the boiling and freezing point of a substance when the applied pressure is changed, or a mixture when the composition is modified), and for rationalizing the energy changes that occur in the course of a chemical reaction. The branch of thermodynamics dealing with the latter is called thermochemistry. Chemical thermodynamics also provides a framework for the determination of the maximum work that may be generated by a system undergoing a specified change, and it therefore provides a way of establishing bounds for the efficiencies of a variety of devices, including engines, refrigerators, and electrochemical cells. Thermodynamics is used in chemistry to assess the position of equilibrium of a chemical reaction (that is, how far it will proceed before appearing to stop), and to determine what conditions are necessary in order to optimize the yield of a particular product. The branch of chemical thermodynamics dealing with ionic reactions occurring in the presence of electrodes constitutes the field of equilibrium electrochemistry. *See* THERMOCHEMISTRY.

Chemical thermodynamics is based on the laws of thermodynamics. In chemistry the most important thermodynamic properties are the enthalpy H and the Gibbs energy G (which is also called the free energy). The change in enthalpy may be identified as the heat transferred to a system during a specified change at constant pressure. Although absolute enthalpies cannot be measured, the enthalpies of compounds relative to their elements can be determined, and extensive tables of these quantities have been compiled. From them it is possible to predict the heat available (or required) for a particular reaction. Not the whole of the heat output of a reaction is available to do work (such as mechanical work, or the work to drive other reactions toward a particular desired product). The Gibbs energy expresses the maximum work, other than work of expansion, that a specified process may generate at constant pressure, and as such it may be used to assess whether one reaction may be used to drive another in an unnatural direction. For example, the assessment of the changes in the Gibbs energy that accompany biochemical reactions may be used to discuss the processes that occur in living cells, where the ingestion of food leads ultimately to growth, mechanical work, and nervous activity. *See* ENTHALPY.

The basis of the Gibbs energy is the tendency of energy to attain a condition of greatest dispersal. This tendency is expressed generally in terms of the entropy S of a system: as the dispersal of energy increases, so does the entropy. The concept of entropy may be given a sharp and quantitative definition, and entropies of substances have been determined (through measurements of heat capacity or by making use of spectroscopic data) and tabulated. The crucial feature for the present purpose, though, is that whereas the entropy of the universe increases whenever a spontaneous change occurs, a chemist is normally interested in the changes that occur in the system under investigation. The Gibbs energy focuses attention on changes in the properties of the system itself because, provided the pressure and temperature are constant, G automatically takes into account the changes in entropy of the surroundings. To determine whether a specified change has a natural tendency to occur at constant temperature and pressure, it is necessary only to assess whether that change is accompanied by a decrease in the Gibbs energy of the system. If a change is accompanied by an increase of Gibbs energy, it may still be achieved by coupling the system to another in which a change is occurring such that the overall change in the Gibbs energy of the combined system is negative. Because the Gibbs energy depends on the composition of the system, it is possible to use the Gibbs energy to predict the composition of the system when the system has attained equilibrium (such as when some reaction mixture attains some constant composition).

The crucial equation is $\ln K = \Delta_r G^{\ominus}/RT$, where K is the equilibrium constant of the reaction (a function of the concentration of the components at equilibrium), R the gas constant, T the absolute temperature, and $\Delta_r G^{\ominus}$, the standard reaction Gibbs energy (the change in the Gibbs energy under certain specified, standard conditions). This equation lies at the heart of chemical thermodynamics and is of great practical importance. *See* CHEMICAL EQUILIBRIUM; CHEMICAL THERMODYNAMICS; ENTROPY; FREE ENERGY.

Equilibrium electrochemistry. In equilibrium electrochemistry, attention is focused on systems in which a chemical reaction may release energy by driving electrons through an external circuit. The potential difference between the two electrodes immersed in the reaction mixture is related to the Gibbs energy for the reaction by $E = -\Delta_r G/F$, where F is the Faraday constant, the magnitude of the charge per mole of electrons (the product of the Avogadro constant and the fundamental charge; $F = 96{,}485$ coulombs mol^{-1}). The potential difference under these conditions is denoted E, and is called the electromotive force (emf) of the cell. One object of electrochemistry is to measure the emf of cells, and then to use the tabulated results to discuss the position of equilibria in reactions in aqueous solution (via ΔG) and their response to changes of conditions. Practical applications include chemical analysis, the assessment of the power generation and storage capabilities of electrochemical cells and fuel cells, the discussion of tendencies to corrosion, and the analysis of potential differences across biological membranes (such as those that are responsible for the

propagation of nerve impulses). *See* ELECTROCHEMISTRY; ELECTROMOTIVE FORCE (EMF).

Quantum chemistry. The principal role of quantum mechanics in chemistry is in the discussion of atomic and molecular structure and in the interpretation of spectroscopic data. In the branch of physical chemistry known as computational quantum chemistry, interest centers on the numerical solution of the Schrödinger equation in order to obtain the wavefunctions of electrons in molecules and to predict and understand the shapes of molecules. In ab initio calculations the computations are done without any appeal to experimental data, and an attempt is made to predict properties from first principles (that is, from the masses and charges of the electrons and nuclei constituting the molecule). In semiempirical calculations (which are identified by initials such as AM1 and MINDO), some of the computational difficulties are circumvented (but with some loss of reliability) by incorporating experimental data and adjusting parameters so that certain experimental quantities (such as enthalpies of reaction) are reproduced. It is now possible to calculate the shape, electron distribution, and spectroscopic properties of large molecules. One application is referred to as quantum pharmacology, where a pharmacologically active molecule is first screened by calculating and analyzing its electron distribution. To some extent, this type of screening calculation eliminates lengthy and expensive laboratory trials. Computational quantum chemistry is so developed that it is capable of being used to map the changes in the structures of molecules while they are in the course of reaction, when atoms and groups of atoms are being transferred from one molecule to another. *See* COMPUTATIONAL CHEMISTRY; QUANTUM CHEMISTRY; SCHRÖDINGER'S WAVE EQUATION.

Spectroscopy. Spectroscopic techniques are used not only to identify molecules present in a sample but also to determine their shape, size, and electron distribution. The techniques fall into four categories as follows.

Absorption spectroscopy. In absorption spectroscopy, the basic observation is the intensity of radiation absorbed at different frequencies (or wavelengths). When the exciting frequency lies in the microwave region (wavelengths in the vicinity of 1 cm), the absorption is due to the excitation of rotational motion of the molecule. Hence the technique may be used to assess the moment of inertia, and therefore the shape and size of the molecule. When the exciting radiation lies in the infrared region (wavelengths in the region of 1000 nm), the vibrational modes of the molecule are excited, and hence the absorption spectrum gives information about the stiffness of bonds. Furthermore, because different groups of atoms absorb in characteristic regions, infrared spectroscopy is also used to identify groups of atoms present in a molecule, and therefore to identify the molecule (this application is called fingerprinting). Absorption in the visible and ultraviolet region of the spectrum indicates the occurrence of an electronic transition, that is, a shift of electron density from one region of the molecule to another. *See* INFRARED SPECTROSCOPY; MICROWAVE SPECTROSCOPY.

Electronic spectroscopy is used for identification, and provides a basis for the discussion of the processes involved in photochemistry (chemical reactions induced by the absorption of radiation), in fluorescence and phosphorescence, and in laser action. Sometimes the incident radiation is so energetic that its absorption results in the ejection of electrons from the molecule. The analysis of the energies of these electrons leads to information about the energy levels of electrons in the molecules under investigation; these photoelectron spectroscopy techniques are designated uv-PES when ultraviolet radiation expels outer electrons, and X-PES when x-rays are used to eject the more tightly bound electrons (formerly known as ESCA, standing for electron scattering for chemical analysis).

Emission spectroscopy. In emission spectroscopy, the spectrum of frequencies present in electromagnetic radiation emitted from excited atoms is monitored. Apart from its use to identify the composition of mixtures, emission spectroscopy is used to determine the state of molecules immediately after they have reacted, and hence gives valuable information about the processes occurring during reaction.

Raman spectroscopy. In Raman spectroscopy, electromagnetic radiation is scattered from molecules. In the process, some energy may be transferred from the molecule to the radiation, or vice versa, with the result that additional frequencies appear in the radiation emerging from the sample. Rotational vibration Raman spectra give information that complements that obtained from the corresponding absorption techniques. *See* RAMAN EFFECT.

Resonance techniques. These techniques constitute the fourth type of spectroscopy. They depend on excitation frequencies of the molecule being brought into resonance with the surrounding electromagnetic radiation by modifying the external conditions. In nuclear magnetic resonance (NMR), the principal example of this technique, the relevant energy levels are those arising from the different possible orientations of the magnetic moments of the nuclei in the molecules, and resonance occurs at radio frequencies. Nuclear magnetic resonance is a powerful technique for structural analysis and identification. Other resonance techniques include electron spin resonance (ESR or EPR) and Mössbauer spectroscopy. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; ELECTRON SPECTROSCOPY; MOLECULAR STRUCTURE AND SPECTRA; MÖSSBAUER EFFECT; NUCLEAR MAGNETIC RESONANCE (NMR); PHOTOCHEMISTRY; SPECTROSCOPY.

Diffraction techniques. Other major techniques for the investigation of molecular structure are based on diffraction, the interference between waves (of electromagnetic radiation) leading to patterns of constructive and destructive interference. These techniques depend on the observation of the direction through which radiation and particles are scattered when they impinge on a sample. The principal example of these techniques is x-ray diffraction, for

through it detailed information may be obtained regarding the arrangement of atoms in crystals and very complex, large, biologically important molecules, such as proteins and deoxyribonucleic acid (DNA). Also, there are diffraction techniques based on electrons and neutrons. *See* DEOXYRIBONUCLEIC ACID (DNA); PROTEIN; X-RAY DIFFRACTION.

Other techniques. Other techniques for investigating structure include the electric and magnetic properties of molecules, in particular, the determination of electric polarizabilities and dipole moments (which give information about electron distribution, and are important for the discussion of the dielectric properties), magnetic properties, and the properties based on optical birefringence, such as optical activity and the Faraday effect. In the case of macromolecules and colloids, important techniques used for structural analysis include x-ray diffraction, osmometry, mass spectrometry, and viscosity measurements. *See* DIPOLE MOMENT; FARADAY EFFECT; MASS SPECTROMETRY; OPTICAL ACTIVITY; OSMOSIS; VISCOSITY.

Statistical thermodynamics. Structural properties and thermodynamic properties are brought together by statistical thermodynamics. This major theoretical procedure gives a way of predicting the thermodynamic properties of substances in terms of the properties of their individual molecules. Statistical thermodynamics represents a grand synthesis of the two major aspects of physical chemistry. It provides an understanding of the bulk, macroscopic properties of molecules and is an important technique for calculating otherwise inaccessible properties from readily available spectroscopic data.

Transport processes. The third major branch of physical chemistry is concerned with physical and chemical change. In particular, it is concerned with the rate of change (whereas thermodynamics is concerned with the possibility of change, and the criteria of spontaneous change). Physical change includes the diffusion of one substance into another, or the migration of ions in an electrolyte solution. The simplest version of the former are the transport properties of gases. These include thermal conductivity and viscosity. They are treated most simply in terms of the kinetic theory of gases, in which a gas is regarded as a swarm of noninteracting points; modern physical chemistry is concerned with the role of intermolecular forces in determining the transport properties. Ion migration gives rise to questions of the mobility of ions in a variety of solvents, and to the description of diffusion processes in general. The application of thermodynamics to change in general constitutes the field of nonequilibrium thermodynamics. *See* DIFFUSION; GAS; TRANSPORT PROCESSES.

Chemical kinetics. The major aspect of chemistry is change, and change is therefore also a major aspect of physical chemistry. Chemical change may be studied at a variety of levels. Empirical chemical kinetics is the study of reactions in order to determine how their rates depend on the concentrations of the participants in the reaction and on the conditions (mainly the temperature). Investigation of the time dependence of reactions (a time dependence that can be observed from the order of minutes to femtoseconds) yields a detailed picture of the sequence of molecular transformations involved in a complex chemical reaction. Each step in the sequence depends on the concentration of the participants and an empirical constant referred to as the rate constant. The aim of theoretical chemistry is to account for the value of the rate constant and its dependence on the temperature. The earliest broadly successful approach led to the Arrhenius law, which is still used as a general guide to the way that the temperature affects the rate of a reaction. The law asserts that the rate constant k varies with temperature as $k = A \exp(-E_a/RT)$, where A and E_a are empirical parameters, the latter being known as the activation energy of the reaction. A wide variety of simple reactions conform to this law, and it can be interpreted on the basis that only molecules possessing sufficient energy are able to undergo reaction when they encounter other species; the exponential factor indicates the fraction of collisions that satisfy the energy requirement. *See* CHEMICAL DYNAMICS; SHOCK TUBE; ULTRAFAST MOLECULAR PROCESSES.

Molecular reaction dynamics. Modern approaches to the prediction and explanation of observed reaction rates have been based either on a statistical view of the reaction (the activated complex theory or transition state theory), where the principles of statistical thermodynamics are applied to a system evolving with time; or on a fundamental molecular dynamics approach, where the trajectories of molecules are calculated as they undergo reaction—techniques then being used to convert these trajectories to values of the rate coefficients. The latter approach constitutes the field of molecular reaction dynamics. Experimental techniques have also been developed for observing individual molecular encounters and reactions. These are based on molecular beams, where a diffuse beam of one reactant is directed into the path of another, and the pattern of molecular scattering, including the electronic and vibrational states of the products, is interpreted in terms of the forces acting between the reactants during the reactive encounter. These techniques bring investigations closer to an atomistic and molecular interpretation of chemistry than anything that preceded them, but the relation of trajectories and rate coefficients and the relation of gas-phase events to those in solution remain problems. *See* MOLECULAR BEAMS.

Surface chemistry. Chemical kinetics is so important that it is rich in applications and extensions. An important extension is to the reactions that occur on surfaces; these are the processes involved in heterogeneous catalysis. The study of surface chemistry breaks down into the analysis of the steps that lead to the affixation of the species to the surface (that is, the study of adsorption processes), the determination of the structure of the adsorbed species, and finally the reactions, and escape, of the adsorbed species. The entire subject, including the related processes occurring at liquid surfaces, constitutes the field of surface chemistry. This field has been

transformed by the introduction of scanning tunneling microscopy, which gives a detailed portrayal of the surface on an atomic scale. A special application of surface chemistry is to the stability of colloidal suspensions of species in fluids, and another is to the processes that occur at the interface between an electrode and the solution in which it is immersed. Electrode reactions are governed largely by the rate of electron transfer between the ions in the solution and the electrode. The field of dynamical electrochemistry finds important applications in electrochemical power generation and storage, in corrosion, in electrodeposition, and in electrocatalysis. *See* ADSORPTION; COLLOID; ELECTRON MICROSCOPE; HETEROGENEOUS CATALYSIS; SCANNING TUNNELING MICROSCOPE; SURFACE AND INTERFACIAL CHEMISTRY; SURFACE PHYSICS. P. W. Atkins

Bibliography. P. W. Atkins, *Concepts in Physical Chemistry*, 1995; P. W. Atkins, *Physical Chemistry*, 6th ed., 1998; D. A. McQuarrie and J. D. Simon, *Physical Chemistry: A Molecular Approach*, 1997; I. Prigogine and S. A. Rice (eds.), *Advances in Chemical Physics*, ongoing series.

Physical geography

The study of the Earth's surface features and associated processes. Physical geography aims to explain the geographic patterns of climate, vegetation, soils, hydrology, and landforms, and the physical environments that result from their interactions. Physical geography merges with human geography to provide a synthesis of the complex interactions between nature and society.

The roots of physical geography may be traced to classical Greek, Roman, and Arab scholars who provided basic insights into the many facets of mathematical geography, the study pertaining to the configuration, measurement, and motion of the Earth. Observations from the great voyages and explorations in the fifteenth to eighteenth centuries afforded additional, sometimes fanciful, descriptive information about the Earth. However, most explanations of the Earth remained for geographers of the nineteenth century, as in the rigorous and methodical work of Alexander von Humboldt (1769–1859) in South America and the perceptive exploration of the American west by John Wesley Powell (1834–1902).

By the beginning of the twentieth century the basic content of physical geography was well established and comprised a number of areas of specialization. Climatology, the scientific study of climates, concerns the total complex of weather conditions at a given location over an extended time period; it deals not only with average conditions but with extremes and variations. Geomorphology is the interpretive description and explanation of landforms and the fluvial, glacial, coastal, and eolian process that operate on them. The forms, processes, and patterns within the biosphere, including vegetation and animal distributions, are studied as biogeography. With strong ties to fluvial geomorphology, geographic hydrology concerns the scientific study of water from the aspects of distribution, movement, and utilization. Soil geography, with emphasis on the origin, characteristics, classification, and utilization potential of soils, provides an area of specialization with links to land use. Ultimately, the physical geography of a region is understood through an integration of the multiple aspects.

The methods used to derive and present the findings of both specialized and integrative studies changed appreciably during the twentieth century. Initially, emphasis was on the organization of newly acquired information about the Earth's surface. Simple models of landscape evolution (for example, William Morris Davis's geomorphic cycle), procedural classification of climate (such as that formulated by Vladimir Köppen) and soils (for example, Curtis F. Marbut's great soil groups), and regional descriptions were typical areas of research and study. New methods and interpretive systems were introduced around the mid-1950s.

Of considerable importance was the change of methodological focus from form (describing the distribution of phenomena) to studies of processes that created the geographic patterns. This change was aided by the use of deductive scientific enquiry employing mathematical and statistical tools. The advent of computer technology enabled large spatial data sets to be analyzed and physical and stochastic models to be constructed. Significant research on a variety of topics, from the dynamics of climates to quantitative models of stream flow, and from systems science thinking to analysis of biogeochemical cycles, reflect the emphasis on physical geographic process.

Maps have always been important tools of the physical geographer. Maps, used or derived, range in scale from those providing intimate details of hill slope erosion to world maps of vegetation biomass, from explorer maps of landscapes to global wind systems. Modern maps have been greatly enhanced by technological advances, especially the development of remote sensing and geographic information systems (GIS).

Remote sensing, the acquisition of data from a distant location, was for many years based upon aerial photography. The development of nonphotographic imagery provides the physical geographer with data and new analytic techniques. Earth-observing satellites, such as NASA's *Landsat* series and the French *SPOT* series, permit, for example, analysis of sequential spatial changes in remote global regions; maps of global cloud cover and sea surface temperatures are examples. Remotely sensed and other derived data are used to construct a geographic information system, an assemblage of computer software and hardware that enables spatial data to be manipulated and displayed. Typically, the system uses a variety of spatial sets to be overlain in order to produce a map composite.

While some physical geographers still complete research dealing with broad theoretical problems on such subjects as chaos or systems theory, the

majority of practitioners deal with more specialized topics. Although physical geographers may study aspects of hydroclimatology or glacial geomorphology, they retain the inherent spatial, process-oriented context associated with the discipline. Underlying this philosophy are a number of basic tenets. First, the linkages between the patterns and processes that operate at different scales, from the microscale to the global level, need to be defined and explained. Second, the careful accumulations of field data and the remotely sensed data need to be reconciled. Third, application of basic research to environmentally sensitive issues is necessary to produce a better understanding of nature-society interrelationships. The integrative perspective of geography is at the basis of such endeavors.

Physical geography is of considerable importance in dealing with complex environmental issues. Applied physical geographers use their expertise to complete research on a variety of topics. Contributions toward understanding the impacts of global warming, flood and drought analysis, coastal erosion, and loss of biodiversity typify the research.

Further considerations of physical geography appear in topical and regional articles. For principal topical discussions *see* AERIAL PHOTOGRAPHY; CARTOGRAPHY; CLIMATOLOGY; CONTINENT; EARTH RESOURCE PATTERNS; ECOLOGY; GEOGRAPHIC INFORMATION SYSTEMS; GEOLOGY; GEOMORPHOLOGY; HILL AND MOUNTAIN TERRAIN; HYDROLOGY; MATHEMATICAL GEOGRAPHY; METEOROLOGY; PHYSICAL GEOGRAPHY; PLAINS; TERRAIN AREAS.

For physical geography of the continents and major island areas *see* AFRICA; ANTARCTICA; ARCTIC AND SUBARCTIC ISLANDS; ASIA; AUSTRALIA; EAST INDIES; EUROPE; NEW ZEALAND; NORTH AMERICA; PACIFIC ISLANDS; SOUTH AMERICA; WEST INDIES.

John E. Oliver

Bibliography. M. N. DeMers, *Fundamentals of Geographic Information Systems*, John Wiley, 1997; A. Goudie (ed.), *The Encyclopedic Dictionary of Physical Geography*, 2d ed., Blackwell, 1994; K. J. Gregory, *The Nature of Physical Geography*, Edward Arnold, 1985; A. N. Strahler and A. Strahler, *Introducing Physical Geography*, John Wiley, 1998.

Physical law

A highly ambiguous term that designates four different concepts: (1) objective pattern (or natural regularity), (2) formula purporting to represent an objective pattern, (3) law-based rule (or uniform procedure), and (4) principle concerning any of the preceding. To avoid confusion, these concepts will be called laws of types 1, 2, 3, and 4, respectively.

Four types of laws. For example, Newton's second law of motion, $ma = F$, is a law of type 2. It represents, to a good approximation, the actual behavior (law of type 1) of medium-size particles moving slowly relative to the speed of light. Alternative laws of motion, such as the relativistic and quantum-mechanical ones, are different laws of type 2 representing the same objective pattern or law of type 1 to even better approximations. One of the rules (laws of type 3) associated with Newton's second law of motion is: In order to set in motion a stationary particle, exert a force on it. Another is: In order to stop a moving particle, exert on it a force in the opposite direction. An example of a law of type 4 is: Newton's laws of motion are invariant under a Galileo transformation. *See* GALILEAN TRANSFORMATIONS; NEWTON'S LAWS OF MOTION.

A physical law of type 1, or objective pattern, is a constant relation among two or more properties of a physical entity. In principle, any such pattern can be conceptualized in different ways, that is, as alternative laws of type 2. The history of theoretical physics is to a large extent a sequence of laws of type 2. Every one of these is hoped to constitute a more accurate representation of the corresponding objective pattern or law of type 1, which is assumed to be constant and, in particular, untouched by human efforts to grasp it. Likewise, the history of engineering is to some extent a sequence of laws of type 3, or law-based rules of action, of which there are least two for every law of type 2. As for the laws of type 4, or laws of laws, they are of two kinds: scientific and philosophical. The general covariance principle is of the first kind, whereas the hypothesis that all events are lawful is a philosophical thesis. Unlike the former, whose truth can be checked with pencil and paper, the principle of lawfulness is irrefutable, though extremely fertile, for it encourages the search for pattern. *See* ENGINEERING; THEORETICAL PHYSICS.

Basic and derived laws. Not all formulas are called physical laws. For example, the regularities found by curve fitting are called empirical formulas. In physics a formula is called a law if and only if it meets the following conditions: it is part of a theory (hypothetico-deductive system), and it has been satisfactorily confirmed by measurement or experiment at least within a certain domain (for example, for small mass densities or high field intensities). Thus, the basic assumptions of all the standard physical theories are laws, and so are their logical consequences. In particular, the usual variational principles, such as Hamilton's, are basic laws. However, the equations of motion and field equations entailed by such principles are derived laws (theorems); so are the conservation laws entailed by the equations of motion and field equations. However, the distinction between basic and derived laws is contextual: what is a principle in one theory may be a theorem in another. For example, Newton's second law of motion is a theorem in analytical dynamics, and the first principle of thermodynamics is a theorem of statistical mechanics. *See* CONSERVATION LAWS (PHYSICS); CURVE FITTING; HAMILTON'S PRINCIPLE; PHYSICAL THEORY; SCIENTIFIC METHODS; STATISTICAL MECHANICS; THERMODYNAMIC PRINCIPLES; VARIATIONAL METHODS (PHYSICS).

Mario Bunge

Bibliography. M. Bunge, *Philosophy of Physics*, 1973; M. Bunge, *Scientific Research*, 1967; L. Sklar, *Philosophy of Physics*, 1992.

Physical measurement

The determination of the size of a physical quantity by comparison with a standard. All physical measurement is ultimately based on measurement unit standards defined by the SI-metric system. Metric units are the foundation from which all other measurement units are derived, even the conventional nonmetric units used by the United States. *See* METRIC SYSTEM; UNITS OF MEASUREMENT.

The SI-metric system consists of seven base units representing length, mass, time, electric current, thermodynamic temperature, amount of substance, and luminous intensity (**Table 1**). Each of these quantities and its unit will be explained in the following sections.

Length (meter). One of the most basic measurements is that of length. Throughout history various standards of length have been chosen, leading to a confusing array of units, some of which had the same name but which varied widely in size. In the absence of one set of measurement standards, there was ample opportunity for misunderstanding between buyers and sellers as to the quantity of goods being exchanged. This problem led to the need for standard measures that could be used not only within a nation but throughout the world to promote both domestic and international trade.

The first attempt at a scientific and logical set of measures was made in 1793, when the meter (metre) was defined as one ten-millionth of the distance from the pole to the equator of the Earth. This determination was made by measuring a portion of the Earth's quadrant, in particular a survey of the meridian from Dunkirk to Barcelona, and then extrapolating the result to include the entire distance of the meridian from pole to equator.

The meter so defined was first standardized as the distance between two lines etched on a platinum-iridium bar that is kept in a vault (at constant temperature to avoid expansion and contraction) at the International Bureau of Weights and Measures (BIPM) outside Paris, France. Copies of this prototype meter were distributed as standards of length to various countries signing on to the Treaty of the Metre in 1875. The meter was therefore accepted internationally and formed the basis for virtually all other length units, even those not part of the metric system.

As technology improved, it was possible to devise more precise definitions of the meter, and in particular, a definition was sought that was reproducible by physical experiment, and not defined solely by an artifact. The first proposed definition was based on the wavelength of the red line of cadmium. However, when the SI-metric system was formally established in 1960, the definition of the meter utilized a transition between two quantum levels of the krypton-86 atom (**Table 2**). That was followed in 1983 by a definition of the meter that fixed the speed of light as a constant and defined the meter as the length traveled in a very precise interval of time. This was possible because the second, as the base unit of time (discussed below), is defined even more precisely than the meter. *See* LIGHT.

Although the definition of the meter has changed several times over the years, the length of the meter has not changed, but the precision by which it is measured has improved. Currently the meter has an estimated relative standard uncertainty of $\pm 2.5 \times 10^{-11}$. *See* LENGTH.

Mass (kilogram). Another very basic measure is that of mass, commonly called weight. The distinction between mass and weight is an important one that is not generally understood by the layperson, whose perspective comes from a frame of reference where the acceleration of gravity is constant. In the metric system, mass and weight are distinct quantities and are assigned different unit names (kilogram and newton, respectively). That distinction is in contrast to the pound, which can be either a unit of mass or a unit of weight, with consequences that can be confusing and dangerous if these units are used improperly. *See* WEIGHT.

The kilogram is the only base unit of the SI-metric system whose name contains a prefix. That is because the gram originated as the unit of mass in the centimeter-gram-second (cgs) metric system, being defined as the mass of 1 cubic centimeter of water at standard temperature and pressure conditions. The cgs system was in some sense in competition with another metric system: the meter-kilogram-second (MKS) system. But the two systems were superseded by the SI-metric system in 1960, which adopted the kilogram as the base unit of mass in spite of its having a prefix; it was decided not to coin a new name for the mass unit, since the kilogram was already in wide use.

The kilogram remains defined by a platinum-iridium cylinder kept in a vault at the BIPM along with the original prototype meter, and is the only unit still defined by an artifact (Table 2). Competing studies are underway to create a definition of the kilogram, like that of the meter, which is reproducible in laboratories. A change in how this unit is defined may be expected in the near future. *See* MASS.

Time (second). The second originated with the Babylonian system of measuring time that is still used today, a system of 24-hour days divided into 60 minutes of 60 seconds each. However, the day cannot be used to precisely define the second, as the length of the day is not constant, nor is the length of the year. An early attempt to make the second more precise resulted in the second being defined in terms of a specific year, the tropical year 1900 (Table 2).

TABLE 1. SI-metric base units

Quantity	Unit	Symbol
Length	meter	m
Mass	kilogram	kg
Time	second	s
Electric current	ampere	A
Thermodynamic temperature	kelvin	K
Amount of substance	mole	mol
Luminous intensity	candela	cd

TABLE 2. Definitions of the SI-metric base units

Unit	Year	Definition*
Meter	1793	1/10,000,000 of the distance from the pole to the equator.
	1795	Provisional meter bar constructed in brass.
	1799	Definitive prototype meter bars constructed in platinum.
	1889	International prototype meter bar in platinum-iridium, cross-section X.
	1906	1,000,000/0.643,846,96 wavelengths in air of the red line of the cadmium spectrum.
	1960	1,650,763.73 wavelengths in vacuum of the radiation corresponding to the transition between levels $2p_{10}$ and $5d_5$ of the krypton-86 atom.
	1983	**The meter is the length traveled by light in vacuum during 1/299,792,458 of a second.**
Kilogram	1901	**The kilogram is the unit of mass; it is equal to the mass of the international prototype kilogram.**
Second	1956	The fraction 1/31,556,925.9746 of the tropical year for 1900 January 0 at 12 hours ephemeris time.
	1967	**The second is the duration of 9,192,631,770 periods of the radiation corresponding to the transition between two hyperfine levels of the ground state of the cesium-133 atom.**
Ampere	1948	**The ampere is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross section, and placed 1 meter apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newton per meter of length.**
Kelvin	1967	**The kelvin, unit of thermodynamic temperature, is the fraction 1/273.16 of the thermodynamic temperature of the triple point of water.**
Mole	1971	**The mole is the amount of substance of a system which contains as many elementary entities as there are atoms in 0.012 kilogram of carbon-12.**
Candela	1979	**The candela is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} hertz and that has a radiant intensity in that direction of 1/683 watt per steradian.**

*Present definition is in boldface.

As with other SI base units, a more precise definition of the second was sought, one based on a physically reproducible experiment. The result is that the second is currently defined in terms of a transition between hyperfine levels of the ground state of the cesium-133 atom (Table 2). The frequencies of atomic transitions such as this are known with an uncertainty of a few parts in 10^{13} to a few parts in 10^{14}. Donald W. Hillger

Advances in time and frequency measurement. Specifically, the second (s) is the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom. Frequency is the reciprocal of the period for regularly repetitive events, and the unit of frequency is the hertz (Hz). A time standard (clock) is the combination of a constant frequency generator and a period (or cycle) counter.

Advances in laser technology have made it possible to slow the motion of atoms (laser cooling), and this has resulted in a significant improvement in the performance of cesium frequency (time interval) standards. The low velocity of the atom after laser cooling (about 1 cm per second) allows a fundamentally different approach to building a cesium frequency standard. All atomic clocks involve one or more microwave regions where the atom is subjected to a microwave field. The microwave energy causes a transition in the state of the cesium atoms when the microwave frequency corresponds to the frequency of the atomic resonance, and this transition is detected. In older technology a thermal beam of cesium atoms passed through two aligned microwave cavities in an interval of 1 to 10 milliseconds. Such standards are referred to as thermal beam standards. With laser cooling, the cesium atoms can be gathered and launched upward as in a fountain at a low velocity (about 4 meters per second) such that the atoms pass through a microwave cavity on the way up and again through the same cavity on the way down as much as a second later. This type of standard is referred to as a cesium fountain. The longer interrogation time (time interval between passages through a microwave cavity) of the cesium fountain makes it inherently a more accurate and stable frequency (time interval) standard. In cesium fountains the frequency accuracy is now better than 1 in 10^{15}.

Cesium fountains are almost always operated as frequency meters. In other words, the frequency from an external source is converted to the cesium frequency by use of a frequency synthesizer (a device that performs various combinations of frequency multiplication, division, addition, and subtraction) and is then fed to the fountain. By noting how much the synthesizer's frequency must be adjusted to match the cesium resonance, the fountain can measure the input frequency relative to the cesium atom. Thus the only output from the fountain is a number.

There are other frequency standards besides the cesium fountain. In these standards the frequency from an external source may be locked to the atomic resonance, and therefore the standard will output a signal with an atom-stabilized frequency. Periods of the output frequency can be counted to make a clock. Cesium thermal beam standards are such devices and are still used in commercial clocks and frequency standards. There are also commercial frequency standards and clocks based on hydrogen and rubidium atoms. The choice of atom depends on user requirements such as frequency stability, cost, size, and power consumption. All of these standards must be calibrated by an external source to obtain the highest level of accuracy. *See* ATOMIC CLOCK; FREQUENCY MEASUREMENT; LASER COOLING. Tom Parker

Coordinated Universal Time. The second is not only the base unit of time in the metric system; it is also the

basis for a system called Coordinated Universal Time (UTC). UTC in turn provides the basis of civil time, as in common use in the world. Donald W. Hillger

UTC was defined in its present form on January 1, 1972, by the International Radio Consultative Committee (CCIR). It is defined in such a manner that it differs from international atomic time (TAI) by an exact whole number of seconds. TAI is determined from about 200 commercial atomic clocks and a small number of primary frequency standards. The difference between UTC and TAI may be adjusted occasionally by the use of a positive or negative leap second at the end of certain months to keep UTC in agreement with the time defined by the rotation of the Earth with an approximation better than 9/10 of a second. From 1972 through 2006, there were 23 positive leap seconds, averaging about 0.7 leap second per year (that is a rate offset of about 2 parts in 10^8). The legal times of most countries are further offset by a whole or fractional number of hours (time zones and "summer time"). *See* ATOMIC TIME; DYNAMICAL TIME; EARTH ROTATION AND ORBITAL MOTION; TIME.

Tom Parker

Electric current (ampere). The "international" electrical units, the ampere and ohm, were standardized around the turn of the twentieth century. However, "absolute" definitions of the electrical units were desired. Of the various electrical units, the ampere was chosen as an SI-metric base unit, since it can be defined in terms of mechanical units using basic principles of physics (Table 2). This ties together mechanical and electrical units into a unified system. The realization of the ampere, however, is difficult and time-consuming, since it involves the definitions of several other units. This limits the relative uncertainty in the value of the ampere, which is estimated to be a few parts in 10^7. Other electrical units can be built upon the ampere, using principles of the SI-metric system for obtaining derived units from SI base units. *See* CURRENT BALANCE; CURRENT MEASUREMENT; ELECTRICAL UNITS AND STANDARDS.

Thermodynamic temperature (kelvin). The SI-metric unit of thermodynamic temperature is the kelvin, the basis for an absolute temperature scale. The definition of the kelvin depends on the triple point of water, a fundamental fixed point in the temperature scale to which is assigned the temperature of 273.16 K (0.01°C). The kelvin unit is then defined as a fraction of the difference between the temperature of the triple point of water and absolute zero (Table 2). The reason for this definition was to make the size of the unit kelvin equal to the degree Celsius, which is defined as 1/100 of the difference between the melting temperature of ice (0°C) and the boiling temperature of water (100°C). *See* ABSOLUTE ZERO; TRIPLE POINT.

The kelvin scale is related to the degree Celsius scale through the relationship degrees Celsius = kelvin − 273.15 K, the constant being the melting temperature of ice (**Table 3**). Temperatures on this scale are called kelvins, not degrees kelvin. Also, kelvin is not capitalized, and the symbol (capital K) stands alone with no degree symbol. These changes in the official name and symbol for kelvin were set by the 13th General Conference on Weights and Measures (CGPM) in 1967.

TABLE 3. Some baseline temperatures in two temperature scales

	kelvins, K	degrees Celsius, °C
Boiling point of water	373.15	100.
Melting point of ice	273.15	0.
Absolute zero	0.	−273.15

Degrees Celsius (°C) were at one time known as degrees centigrade. However, it is no longer correct to use that terminology. The official name was changed from centigrade degree to Celsius degree by the 9th CGPM in 1948. *See* TEMPERATURE; TEMPERATURE MEASUREMENT.

Amount of substance (mole). There is a need to define an amount of substance, as in chemical reactions. A standard amount of chemical element or compound was at one time called atomic weight or molecular weight, or more correctly atomic or molecular mass. These amounts were in fact masses relative to the atomic weight (or mass) of oxygen, which was agreed to be 16, associated with the number of protons and neutrons in the oxygen atom. Oxygen, however, has a slightly variable mixture of isotopes ranging from 16 to 18, to which chemists associated the number 16. But physicists associated 16 with the isotope of oxygen of mass 16 only. This duality of definition, a source of confusion, ended when physicists and chemists agreed to assign a value of 12, exactly, to the relative atomic mass of the isotope of carbon with mass number 12, carbon-12. In SI-metric the amount of substance, now called a mole, is defined as a fixed mass of the carbon-12 isotope (Table 2). *See* ATOMIC MASS; ATOMIC MASS UNIT; MOLE (CHEMISTRY); MOLECULAR WEIGHT.

Luminous intensity (candela). The final SI-metric base unit is used in the area of photometry. Originally a "new candle" was defined as the luminance of a Planck (blackbody) radiator at the temperature of freezing platinum. This unit was given a new international name, the candela, in 1948. However, difficulties in realizing this unit at the high temperatures specified in the definition led to a new definition of the candela in 1979 (Table 2). *See* CANDLEPOWER; LUMINOUS INTENSITY; PHOTOMETRY.

Other units and measurements. The number of units that can be derived from the base units of the SI-metric system is almost endless. It is sufficient to say that all other measurement units can be traced back to SI-metric base units or combinations of SI-metric base and derived units, through the logical principles of the metric system for creating derived units. Even nonmetric units that are still in common use are defined in terms of SI-metric units, as no other fundamental standards for these units remain, other than their legal equivalencies to SI-metric units.

In particular, the units of the metric system became legal for use in the United States in 1866. In 1893 the Mendenhall Order redefined conventional

units (such as the yard and the pound) in terms of the metric system, rather than maintaining measurement standards for those units. This followed the U.S. agreement to the Convention of the Metre signed in Paris, France, by 17 nations. The Meter Convention, often called the Treaty of the Meter, provided for improved metric weights and measures and the establishment of the CGPM, devoted to international agreement on matters of weights and measures.

The United States is the only remaining major holdout of nonmetric usage. However, the process of conversion was made official in 1975 with the passage of the Metric Conversion Act. That Act was strengthened in 1988 to designate the metric system as the preferred system of units for the U.S. government. As a result, the metric transition, although voluntary, continues as more products and services adopt metric units in order to come in line with the dominant use of metric in the rest of the world. The end result will be international standardization of all physical measurement on the SI-metric system.

Donald W. Hillger

Uncertainty in Practical Measurements

Although the seven base units, and others derived from them, are thus exactly defined, their practical availability requires the development and refinement of standard devices or apparatus to realize each of them with high precision. Extensive theoretical studies and laboratory experiments are involved in the selection and refinement of operating principles and in design, construction, and operation of these standards. Once compared with the base standard, other subordinate standards and reference instruments, specimen objects, signal sources (generators or modifiers of the quantity), and so on can be used for further calibrations or measurement.

Comparison of standards. As noted above, a kilogram mass standard can be calibrated only through a series of comparisons, starting from the International Prototype. The units for the other five base quantities, and all quantities derived solely from them, are in principle independently realizable; that is, the standard apparatus may be constructed equally well in many laboratories. In practice, however, inevitable minor differences between standards constructed independently, even with equal care, and among the instruments, environments, and operators individually are bound to introduce small discrepancies. Periodic comparison of standards and the resolution of these discrepancies are required for compatibility among domestic standards laboratories, as well as internationally. In the United States, the National Institute of Standards and Technology (NIST; formerly National Bureau of Standards) provides calibration services for industrial, educational, and other governmental standards laboratories and cooperates with them in conducting measurement agreement comparisons. Periodic comparisons of NIST standards with those of other countries are made through the International Bureau of Weights and Measures, international scientific organizations, or direct arrangement.

Frequency and time comparisons within the United States are made between the NIST, the U.S. Naval Observatory, and other users of high-precision frequency standards. The data from worldwide astronomical observations and from standards laboratories in many countries are coordinated by the International Bureau of the Hour, which coordinates differences between the International Atomic Time scale (a running count of atomic seconds) and the Coordinated Universal Time scale and announces the time when the UTC offset should be changed by ± 1 second to keep UTC in phase with the solar year. Most radio broadcast signals are measured on UTC.

Sources of error. In general, any measurement has less than perfect accuracy; there is some error or uncertainty as to the true, or exact, numerical ratio between the magnitude of the measured quantity and the unit. Associated with the measuring instrument itself are the errors made in its calibration, the uncertainty in the constancy or reproducibility of the standard by which it was calibrated, and the possible changes of its response after its calibration. Other uncertainties, somewhat more under the control of the observer making a measurement, include those in taking readings, in correcting for environmental effects, and in allowing for characteristics of the instrument as well as of the system or object undergoing measurement.

Material and structural characteristics. Inherent characteristics of materials and structures cause the phenomena of drift, lag, hysteresis, damping, and resonance in measuring instruments or systems, as well as in the systems undergoing measurement.

These general phenomena may affect the relationship of any quantity or condition (mechanical, electrical, thermal, and so on) to any parameter. In terms of instrument reading and measured quantity they may be defined as follows:

1. Drift is the gradual continued change of instrument reading after a change to a different but constant value of the measured quantity.

2. Lag is the failure of the instrument reading to follow changes in the measured quantity instantly. The time constant of an instrument is the time required for the indication to change by $1/e$ (0.37) of a sudden change in the measured quantity in response to this change.

3. Hysteresis results from lag or drift, and is the difference between readings of the measured quantity for corresponding actual magnitudes of that quantity when the quantity is increasing and when it is decreasing.

4. Damping is the dissipation of energy (electrical, magnetic, or mechanical) caused by a change in the measured quantity. With critical damping, the indication changes to its new value with minimum lag without overshoot, following a sudden change in the measured quantity. With less damping, the indication overshoots the new reading and may oscillate about it with decreasing amplitude; with greater damping, the indication changes to its new value more slowly.

5. Resonance is the condition of enhanced response or oscillation which results when the rapidity of change of the measured quantity is close to the natural rapidity of response of the instrument. *See* CIRCUIT (ELECTRICITY); DAMPING; HYSTERESIS; RESONANCE (ACOUSTICS AND MECHANICS); RESONANCE (ALTERNATING-CURRENT CIRCUITS); TIME CONSTANT.

Change of properties. The change of properties of materials or structures with temperature, pressure, humidity, radiation, vibration, or other environmental conditions may cause further uncertainties in the response of the instrument and in the characterization of the quantity or property being measured.

Instrument-measurand interaction. The interaction of the measuring instrument or process with the quantity being measured (measurand) is often not negligible, and uncertainties remain even after corrections; for example, measuring a voltage usually requires some power, which generally tends to lower the measured voltage. On the atomic scale, Heisenberg's uncertainty principle states that it is impossible to determine simultaneously both the position and the momentum of a particle with the product of the uncertainties less than about 10^{-34} joule second. *See* UNCERTAINTY PRINCIPLE.

Error analysis and reduction. Thus, the analysis of experiments to detect all possible sources of error, the design of experimental procedures to minimize them, and the development of mathematical techniques for estimating their probable magnitudes are all important parts of any measurement in which highest accuracy is essential. The effect of random disturbances and reading errors may be minimized by averaging repeated measurements. Averages with various observers may reduce the effects of systematic reading errors of a given observer.

From the dispersion of repeated observations a statistical estimate of the imprecision (often called precision or repeatability) of the measurement may be obtained. *See* STATISTICS.

The inaccuracy (often called accuracy) of a measurement (or calibration) is the total uncertainty, including both the imprecision of observation and the systematic uncertainties associated with the measuring instrument and with the measuring process.

The reduction of uncertainty in measurement is one of the continuing objectives in science; it is also one of the most fruitful contributions to science. Scientific theories live or die as they are tested by more precise measurements. New theories are evolved to fit phenomena revealed by more precise observations, and new advances in many fields follow an improved measurement capability for any one quantity.

Measurement Techniques

The comparison of quantities as to equality, the counting of units, and the determination of the coincidence of events—all involved in physical measurement—may be done by an observer (usually using visual, aural, or tactile faculties) or by instruments which display or record the results, or apply them to automatic computation or control. *See* INSTRUMENTATION AMPLIFIER.

Direct measurement involves comparison with a standard (such as a meter bar) or measurement by a calibrated instrument (such as a voltmeter). Indirect measurements are those derived from measurements of related quantities; for example, the mass of the electron can be derived from measurements on the bending of the path of electrons of a known velocity in a known magnetic field and from separate measurements of the electron charge.

Measuring instruments or systems may respond to the physical quantity to be measured by generating or modifying another quantity or series of quantities. The final (output) quantity, having a functional relationship to the quantity measured, may be compared with a standard of its own kind, appropriately scaled and labeled to represent units of the quantity being measured. For example, in a voltmeter, the voltage across a resistor induces a current which reacts with a magnetic field to generate a force or torque which deforms a mechanical spring and moves a pointer along a scale. Thus the final quantity generated is a length, the displacement of the pointer, which is compared with a scale graduated in intervals of length (or angle) appropriately proportioned to indicate numerically the applied voltage. The proportionality factor is the product of the successive transformation factors from volts to amperes to force to displacement along the scale. Devices that effect this conversion or transformation of one quantity to another are called transducers. *See* TRANSDUCER.

The hundreds of physical phenomena relating one quantity to another supply a rich reservoir of alternatives for the designer of measurement experiments or measuring instruments. For example, the determination of concentration or composition is widely important in both science and industry. An increasing variety of spectroscopies—radio-frequency, infrared, ultraviolet, x-ray, gamma-ray, and so on—utilize atomic and molecular characteristics for absorption, emission, scattering, or reemission of electromagnetic radiation or nuclear particles to determine not only composition but even molecular and crystalline structure. *See* ACTIVATION ANALYSIS; SPECTROSCOPY.

Many atomic phenomena involve several quantities and may serve as a basis for extending the range or accuracy of measurement of each of them, and others. *See* JOSEPHSON EFFECT; LASER; MASER; MICROWAVE SOLID-STATE DEVICES; MÖSSBAUER EFFECT.

Some techniques for reducing uncertainty, extending range, or providing flexibility and convenience in measurements in general are briefly mentioned below.

1. Many natural phenomena and the fundamental properties of matter, such as the speed of light in a vacuum or the charge of the electron, provide constant values of certain quantities or combinations of quantities which may be used in physical measurement or instrument design. For example, the precession frequency of the spin of the proton in the

hydrogen atom, or hydrogen compounds, is proportional to the magnetic field, or to a current generating it, which may be determined from this frequency. It in turn is indicated by varying the frequency of a high-frequency electric field, applied at right angles to the magnetic field, to achieve maximum coupling, through synchronized spins, with a similar passive circuit detecting an induced electric field at right angles to both the magnetic and applied electric fields.

2. The quantity to be measured or one related to it in a known manner may be allowed to modulate or attenuate the value of another quantity—as when the thickness of metal or paper is determined from its attenuation of an x-ray beam, which is converted by any of several types of radiation meters (transducers) to a change in voltage or current and then to a visual indication or record. As is common in spectroscopy, uncertainty due to variation of the source may be reduced by taking the ratio of signals with and without the sample in the beam.

3. The measured quantity itself, or its signal, may be time-modulated to permit better discrimination between the measured quantity and extraneous effects (noise) in the measuring instrument.

4. In weighing on a balance, known weights are substituted for unknown weights previously balanced by counterweights, so that such uncertainties as knife-edge placement and beam lengths do not affect the comparison.

5. The effect of a quantity to be measured may be nearly offset by a similar known and fixed quantity, and the differences measured by an instrument of lesser range and higher sensitivity.

6. If the standard offsetting quantity can be divided into sufficiently small or continuously variable fractions, it can be adjusted for equality to within the smallest limits detectable. This is the so-called null, or balancing, method of measurement.

7. A number of small and equal magnitudes of the same quantity may be combined so that the cumulative magnitude is appropriate for measurement with available methods.

8. The undesired effects of environmental factors, such as temperature and external magnetic fields, may be compensated for in the instrument system by elements which are responsive to the disturbing factor and interact with the measuring or indicating means to offset such effects.

9. Based on careful design of experiments, observations may be made under a wide variety of conditions, with controlled variation in the factors considered as possible sources of error, permitting statistical estimation of the magnitude of the various errors and appropriate correction for them.

10. In measuring properties of materials, uncertainties may be reduced by comparing measurements on the sample with those made on a sample having closely similar, and accurately known, properties. Hundreds of such reference materials or standard samples are available from government or industrial sources with an indication of the value (and the uncertainty) of the characteristic property, be it composition, purity, size, radioactivity, viscosity, or some other. Measurements on such reference materials obviously provide a calibration check of measuring instruments; if their properties are close to those of the unknowns, the differences may be measured by more sensitive methods or devices.

11. Published values for carefully measured properties of generally available materials may obviate the need for repetitive measurements, and may also provide a basis for checking calibration.

12. Data obtained by the "round-robin" circulation of two test pieces, of somewhat different (and undisclosed) values, for measurement by each of a number of individuals or laboratories, provide a simple and practical method for self—as well as group—evaluation. *See* MAGNETOMETER.. William A. Wildhack

Bibliography. L. Finkelstein and K. T. V. Grattan (eds.), *Concise Encyclopedia of Measurement and Instrumentation*, Pergamon Press, Oxford, 1994; *Guide for the Use of the International System of Units*, NIST Spec. Publ. 811, 1995; Institute of Electrical and Electronic Engineers and American Society for Testing and Materials, *Standard for Use of the International System of Units (SI): The Modern Metric System*, IEEE/ASTM SI 10, 2002; *The International System of Units* (*SI*), NBS Spec. Publ. 330, 2001; P. H. Sydenham (ed.), *Handbook of Measurement Science*, 3 vols., 1982, 1983, 1992; P. H. Sydenham, V. H. Hancock, and R. Thorn, *Introduction to Measurement Science and Engineering*, 1992. See also the periodicals *Analytical Chemistry*; *Journal of Physical and Chemical Reference Data*; *Journal of Research of the National Institute of Standards and Technology; Measurement Science and Technology* (British); *Metrologia*; *Review of Scientific Instruments*.

Physical optics

The study of the interaction of electromagnetic waves in the optical range with material systems. The optical range of wavelengths may be taken as the range from about 1 nanometer (4×10^{-8} in.) to about 1 millimeter (0.04 in.). More narrowly, physical optics deals with the relationship between the atomic structure of a system and the manner in which the system affects light sent into it. The chief founder of this branch of science was Michael Faraday, who in 1845 provided the first clue to the electromagnetic nature of light by showing the optical properties of glass could be altered by a magnetic field. *See* FARADAY EFFECT.

The explanation of the absorption, reflection, scattering, polarization, and dispersion of light by a material medium in terms of the properties of the atoms and molecules making up the medium is the objective of physical optics. In the course of seeking this objective, physicists have found that optical investigations are powerful methods of determining the structures of atoms and molecules and of larger systems composed thereof. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; ATOMIC STRUCTURE AND

SPECTRA; CRYSTAL OPTICS; DIFFRACTION; DISPERSION (RADIATION); ELECTROMAGNETIC RADIATION; ELECTROOPTICS; FLUORESCENCE; INTERFERENCE OF WAVES; LASER; LIGHT; MAGNETOOPTICS; MOLECULAR STRUCTURE AND SPECTRA; POLARIZED LIGHT; REFLECTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES; SCATTERING OF ELECTROMAGNETIC RADIATION; SPECTROSCOPY. Richard C. Lord

Bibliography. M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976; A. Mickelson, *Physical Optics*, 1992.

Physical organic chemistry

A branch of science concerned with the scope and limitations of the various rules, effects, and generalizations in use in organic chemistry by means of physical and mathematical methods. It includes, but is not limited to, the dynamics and energetics of organic chemical transformations, transient intermediates in these reactions, rate comparisons between families of reactions, dynamic stereochemistry, conservation of orbital symmetry, the least-motion principle, the isomer number for a given elemental composition, conformational analysis, nonexistent compounds, aromaticity, tautomerism, strain and steric hindrance, and the double-bond rule. Spectroscopy is the main tool employed, with nuclear magnetic resonance being the most widely used spectroscopic technique. With the advent of modern fast computers, computational chemistry has also become an important tool. *See* NUCLEAR MAGNETIC RESONANCE (NMR); SPECTROSCOPY.

Physical organic chemistry is traditionally distinguished from, yet totally intertwined with, synthetic organic chemistry, which deals with the question of how to obtain desired products from available compounds. This distinction can be illustrated with a diagram (**Fig. 1**) showing how the energy might vary during a chemical reaction in which the reactant R yields a product P (R → P). Whereas the synthetic organic chemist will be interested primarily in the practical problem of how to convert R into P, the physical organic chemist studies the curve or curves connecting R and P as well as the structure and physical properties at all extrema, including R and P. However, the demarcation between synthetic organic chemistry and physical organic chemistry is not sharp. Physical organic chemists have contributed greatly to the understanding of the chemistry of hydrocarbons and their derivatives and have enhanced the repertoire of the synthetic organic chemists. In turn, synthetic organic chemists have made possible the construction of the custom-made, often intricate molecules that physical organic chemists use for their studies. The efforts of both groups, moreover, have made possible the birth of such new fields as molecular biochemistry and computational chemistry. *See* COMPUTATIONAL CHEMISTRY; MOLECULAR BIOLOGY; ORGANIC SYNTHESIS.

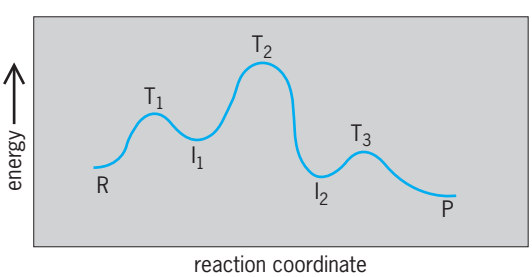

Fig. 1. Energy profile of an organic reaction R → P. The T terms indicate transition states and the I terms indicate intermediates.

Chemical reaction mechanisms. The diagram in Fig. 1 is also useful in discussions of the dynamics of chemical reactions. It is an attempt to portray how the atoms in the reactant molecule R may move in space to their final positions in the product molecule P, and how the potential energy of the system would vary as a function of these positions. A complete correlation would be multidimensional; what is normally shown is a cross section in which the maximum potential energy is in fact a minimum (saddle point). While an essentially infinite number of pathways between R and P can be imagined and followed, the vast majority of the molecules will in practice use the one that makes the least demand on energy to reach the next maximum; this pathway is known as the reaction mechanism. The maxima (T terms in Fig. 1) are known as transition states, and the minima (I terms in Fig. 1) as intermediates.

If a reaction has a single transition state (and, hence, no intermediate), it is known as concerted; alternatively, it is stepwise. A stepwise reaction is simply a succession of concerted steps in which the intermediates are not isolated. Criteria for concertedness have included stereospecificity; inability to detect intermediates either by spectroscopic means or by trapping them with auxiliary compounds to give different products; the rate law; the entropy and volume of activation (temperature and pressure dependence of the rate constant); comparison with reactions thought to be concerted; and isotope, solvent, and substituent effects. *See* ORGANIC REACTION MECHANISM.

In spite of all these criteria, it is exceedingly difficult to prove that a given reaction is concerted, and that there is not some elusive, high-energy intermediate present on the pathway. Indeed, diligent searches have turned up a variety of such species; they include carbanions, carbocations, free radicals and diradicals, radical anions and cations, carbenes, nitrenes, and highly crowded or strained molecules. *See* FREE RADICAL; PHOTOCHEMISTRY; REACTIVE INTERMEDIATES.

Chemical kinetics. The most important route to quantitative information about a reaction is the study of its kinetics. This must begin with an experimental determination of the rate law: the expression that shows how the rate of formation of product $d[P]/dt$ (or loss of reactant: $-d[R]/dt$) depends on the concentration of all species involved in the reaction other than the solvent. The differential might be found to equal $k[R_1]$, $k[R_1]\ [R_2]$, $k[R_1]^2$, and so

on; k is known as the rate constant, and the reaction is described as first-order, second-order, and so forth, depending on the total number of concentration terms. One important feature is that the reaction order equals the sum of all molecules that have participated in the formation of the transition state (T_2 in Fig. 1). Thus, for a reaction to be concerted, it is necessary that the order equal the sum of all reactant molecules involved in the stoichiometry; for stepwise reactions, this need not be the case, although it can be (if there is no intermediate following this state). Another important fact is that since at equilibrium the forward and reverse rates are equal, the ratio of the forward rate constant (k_f) to the reverse rate constant (k_r) equals the equilibrium constant K. When the rate law in one direction is established, the one for the reverse reaction is automatically established also, since this constant is defined by the reaction equation. It follows that the forward and reverse paths cannot be different (principle of microscopic reversibility). *See* CATALYSIS; CHEMICAL DYNAMICS.

To establish the rate law, the first step is to consider all the mechanisms that seem likely candidates. For each, a series of differentials is written expressing the change of concentrations with time; this includes the reactants and all intermediates up to the transition state. Since the concentrations of the reactive intermediates are generally not known, it is assumed that the differentials $d[\mathrm{I}]/dt$ equal zero (the steady-state approximation). The expressions are combined into a single differential equation relating the concentration of [R] and time, and this is integrated. The resulting equations are then tested in the laboratory; those that do not fit the experimental data represent mechanisms that are ruled out, and those that do fit correspond to mechanisms that remain as possibilities (there may be several). *See* DIFFERENTIAL EQUATION.

Entry into the thermodynamic quantities (energy, entropy, and volume) of the transition state can then be gained by means of the theory of absolute rates, which has led to Eq. (1), where k represents the rate

$$k = \frac{RT}{Nh} \exp(-\Delta G^{\neq}/RT) \qquad (1)$$

constant, R the gas constant, T the absolute temperature, N the Avogadro constant, h the Planck constant, and $\Delta G^{\neq}$ the difference between the free energies of the transition state and the reactants. Since the value for the entropy (S) is given by $(\partial G/\partial T)p = -S$ and the value for the volume (V) is given by $(\partial G/\partial P)_T = V$, the entropy and volume of the transition state can be derived from measurements of the temperature and pressure dependence of the rate constant, respectively. These quantities give very important insights into the structure of the transition state; thus, volume increases are often an indication of a breaking bond, or of the desolvation of ions. Similarly, solvent effects on the rate constant often reveal substantial polarity changes on the way to the transition state. *See* CHEMICAL THERMODYNAMICS; FREE ENERGY; HIGH-PRESSURE CHEMISTRY; HIGH-TEMPERATURE CHEMISTRY.

Substitution reactions. Substituent effects have played a dominant role in physical organic chemistry. Placing a substituent near the reactive center may have the effect of greatly altering the rate, that is, greatly changing the free energy (stability) of the transition state. Since the effect of a given substituent on the stability of a variety of structures can be studied independently, such a rate effect can provide significant information about the structure of the transition state. These effects are often of a steric nature, particularly if the substituent is placed very close to the reactive center. If the substituent is in conjugation with the reactive center, its influence may be felt electronically. *See* CONJUGATION AND HYPERCONJUGATION; STERIC EFFECT (CHEMISTRY).

In order to separate steric and electronic effects, chemists have resorted to the use of *meta*- and *para*-substituted phenyl groups; these substituents cannot affect the reactive center sterically. Studies of this sort produce the so-called linear free-energy relations, which may be of two types. In one of these, the rate constants of a reaction involving substituted phenyl compounds are plotted logarithmically against the difference in pK_a value of the similarly substituted benzoic acid and benzoic acid itself. The pK_a value is a quantitative measure of acid strength. A straight line is generally obtained. If the slope is positive, the substituent stabilizes or destabilizes the transition state in the same way that it affects the benzoate anion; this generally means that in the transition state the reactive center is developing negative charge. Conversely, if the slope is negative, the transition state involves positive charge, and the larger the slope, the larger the charge. Plots of this type have played an important role in the development of Brönsted theory. *See* ACID AND BASE; PK; SUPERACID.

The other type of linear free-energy relation results from plots of log k of one reaction against that of another. If the reactions are of the same type, a linear plot may be expected; if they are quite different, the correlation will be poor. An example is the plot that results when log k is measured for the reactions of a series of anions with methyl bromide (CH_3Br) to give bromide ion (Br^-) and with difluoramine (NF_2H) to give fluoride ion (F^-). The former is well known to be a simple, direct, and concerted displacement. A linear plot results, suggesting that the fluoride reaction is also a simple displacement; however, the point for hydroxide (OH^-) is far off this line, the reaction with difluoramine being much faster than expected on the basis of the reaction of this ion with methyl bromide. Thus, the hydroxide reactions are not similar. Subsequent investigation revealed that hydroxide reacting with difluoramine acts as a base rather than as a displacing agent, removing a proton (H) to leave an anion (NF_2^-) that then decomposes to give fluoride ion and fluoronitrene (NF·). The latter intermediate can be captured with a variety of trapping agents, and in fact, it has been observed spectroscopically. *See* MATRIX ISOLATION.

Reactivity-selectivity principle. Consideration of the reasons for rate differences in the reaction of

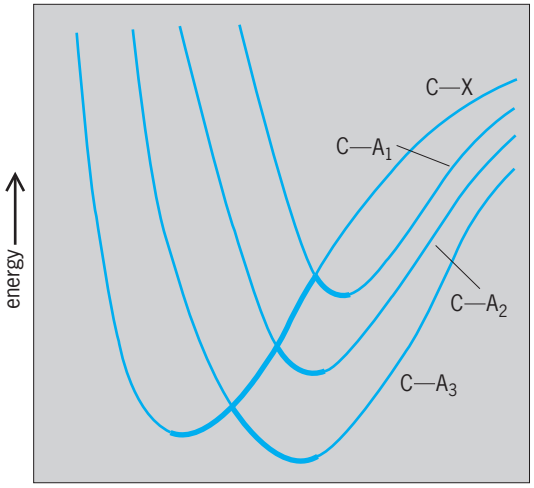

Fig. 2. Profile of the energetics of a series of displacement reactions of substrate C—X with anion A^-.

a given substrate C—X with a series of anions A^- led to a very useful insight into displacement reactions. The energy profile for the reaction is believed to be essentially the sum of the two bond energy curves of C—X and C—A_i (i = 1, or 2, or 3, etc.; **Fig. 2**). The transition states are presumed to lie near the crossover points of these curves. The energy profiles so obtained suggest that the more stable product is associated with a more stable transition state, and hence it is produced in a faster reaction, while the less stable product is associated with a less stable transition state and thus is produced in a slower reaction. This principle has been generalized to apply to all families of reactions; thus, if a series of organic halide derivatives of comparable stability solvolyzes at greatly different rates, it is believed that the fastest among them produce the most stable intermediate carbocations. The same picture, in fact, suggests another important generalization known as the Hammond postulate (originally stated in more succinct form): the faster a reaction, the "earlier" (more reactant-like) is the transition state, and the slower a reaction, the "later" is the transition state.

The reactivity-selectivity principle is also based on relative reaction rates. The basic assumption is that a highly reactive intermediate, when presented with two substrates or with two sites within a single substrate, will react at the first opportunity and hence be unselective (product ratio near unity); a much more stable species is likely to be more discriminating and lead to larger ratios. While this assumption is basically sound, its sometimes uncritical application to rather dissimilar species or reactions has led to many claims of exceptions that leave its usefulness in doubt.

Stereochemistry. This is also a powerful tool in physical organic chemistry. Experimental work demonstrated that when a chiral compound such as (−)HCR_1R_2X is converted into optically active HCR_1R_2Y by direct displacement with Y, the product obtained has an optical rotation that is opposite to that exhibited by the same material if it is produced in two steps, via initial displacement by A to give intermediate HCR_1R_2A [reaction scheme (2)]. This re-

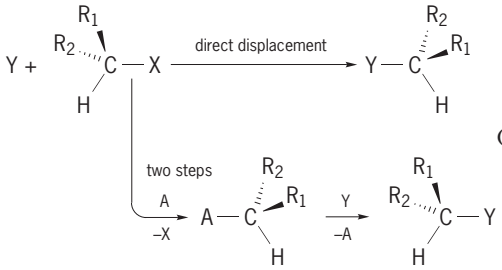

(2)

sult demonstrates that displacement reactions occur with inversion; the reagent approaches in front, and the leaving group departs in the back. *See* OPTICAL ACTIVITY.

Retention of configuration accordingly indicates a two-step reaction (the auxiliary displacing agent may be a solvent molecule or an internal part of the substrate itself). Sometimes racemization is encountered; this may be taken to mean that the substrate suffered initial cleavage to give a trivalent carbon intermediate (for example, a carbocation), which will either be planar or subject to rapid inversion, so that the final capture of the displacing agent can occur at either side with equal likelihood.

Another important area in which configuration of the product is reliably predictable from that of the reactant is the general pericyclic reaction, a type of organic reaction in which a transition can be drawn that possesses a complete cyclic array of parallel p orbitals. The stereochemistry is controlled by the tendency to maintain overall symmetry. This principle has overshadowed and in many cases supplanted an earlier one known as the principle of least motion, which held that the nature of the product is largely determined by a need to minimize nuclear displacements. *See* MOLECULAR ORBITAL THEORY; PERICYCLIC REACTION; STEREOCHEMISTRY.

Isomers. The isomer number has perhaps been the most important organizing principle in organic chemistry since its inception. Simply put, it means that for every elemental composition and molecular mass the isomer number can be predicted by writing all possible sequences of the atoms present, obeying the valence numbers of the atoms: four for carbon, three for nitrogen, two for oxygen, one for hydrogen and the halogens, and so forth. On this basis, it is possible to predict that there will be two structural isomers with the composition C_4H_{10} (**1**), two

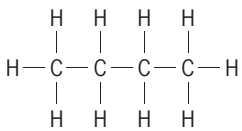

(**1*a***)

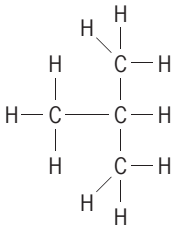

(**1*b***)

with C_2H_6O (**2**), four with C_4H_9Cl (**3**), and so forth.

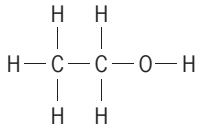

(*2a*)

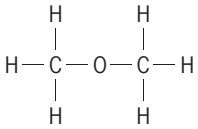

(*2b*)

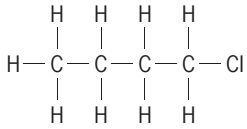

(*3a*)

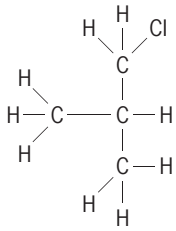

(*3b*)

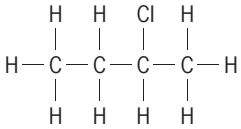

(*3c*)

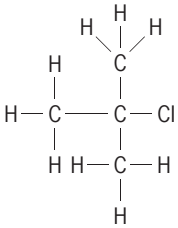

(*3d*)

Double bonds involving the elements carbon, nitrogen, and oxygen are allowed, as are triple bonds to carbon and nitrogen; thus structure (**4**) must be included for composition C_2H_4O, and structure (**5**) for composition C_2H_3N.

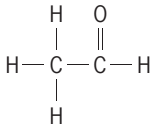

(**4**)

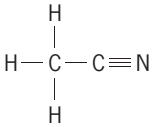

(**5**)

See CHEMICAL BONDING; ORGANIC CHEMISTRY; VALENCE.

Once reliable atomic weights became available so that elemental compositions could be reported with confidence, this simple rule proved remarkably successful. For example, the isomer rule predicts 35 isomers with the formula C_9H_{20}, and all are known. Furthermore, it is sometimes possible to correlate these structures with the properties of the known compounds. Thus, only one of the two known isomers C_2H_6O reacts with sodium to give hydrogen; since it resembles water in that sense, this must be the isomer with an O-H group.

In view of these successes, organic chemists tried to modify or expand the simple rule when exceptions were encountered; some of these came in the form of extra compounds, and others as missing compounds. Some of the extra compounds required only the revision that anomalous valences sometimes occurred; two for carbon, for instance (C═O), or four for nitrogen (ammonium salts).

The need to specify molecular mass has proved more troublesome: it requires a definition of the concept of a molecule. Such definitions usually refer to covalent bonds as the entities that hold the atoms together, to rule out ionic species such as sodium chloride as candidates. However, there are supramolecular aggregates. Examples are trihydrogen trifluoride (H_3F_3; **6**), which remains intact in the gas phase by virtue of the weak and electrostatic hydrogen bonds; diborane (B_2H_6; **7**), which

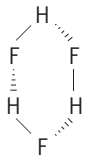

(**6**)

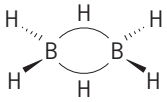

(**7**)

exists by virtue of hyperconjugation or no-bond resonance; and charge-transfer complexes (**8**), in which

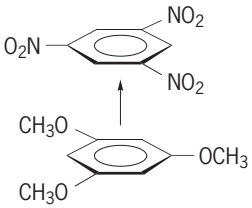

(**8**)

an electron-deficient molecule polarizes and attracts an electron-rich partner. *See* RESONANCE (ACOUSTICS AND MECHANICS).

The synthesis of the so-called catenanes has shown that there are molecules in which component parts are held together by repulsive forces rather than bonds—whether covalent or electrostatic. To mention a specific example: $C_{50}H_{100}$ should, among its isomers, include one in which two macrocyclic $C_{25}H_{50}$ rings are looped together like two links of a chain (**9**).

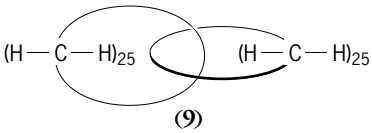

(**9**)

See MACROCYCLIC COMPOUND; STRUCTURAL CHEMISTRY.

Among the extra compounds, none have affected organic chemistry have drastically than the stereoisomers. It turns out that for all but the simplest compounds a given sequence of the atoms may represent two, more than two, or even many more isomers. An early insight into the real nature of molecules was provided by L. Pasteur, who studies of optical activity of solutions and of crystal structures led him to the conclusion that there must be chiral (and hence, three-dimensional) molecules. The origin of this phenomenon was revealed principally by J. H.

van't Hoff, who postulated that the four valences of carbon were directed to the corners of a tetrahedron. *See* MOLECULAR ISOMERISM.

Van't Hoff's work led to correct predictions of far larger numbers of isomers than had been realized before. Further work with stereoisomers revealed that rings containing six carbon atoms or more must be puckered, since in planar form the carbon-carbon bond angles would be 120° or more; the tetrahedral value is slightly less than 110°. *See* CONFORMATIONAL ANALYSIS.

Research has shown that rotation about single bonds if not totally free; the barriers are low so that the interconversion at room temperature is so fast that the isomers (conformers) cannot be isolated. Isolable conformers are possible only in certain especially designed molecules such as the triptycene derivative (**10**). In this case, rotation about bond a

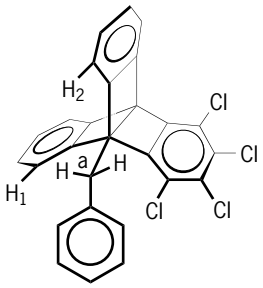

(**10**)

is not free but hindered by the protruding atoms Cl_1 and H_1. In such instances, the conformations must be counted as isomers.

In small rings, bond-angle deviations from the tetrahedral value cannot be relieved, and high-energy, strained molecules result; nevertheless, many are capable of existence. One of the activities of physical organic chemists has been to search for the limits of strain and crowding that molecules can endure and survive. A few examples of such structures are shown in **Fig. 3***a*; several structures that have so far resisted all efforts at synthesis are shown in Fig. 3*b*.

Many of the nonexistent compounds (allowed by the simple rules of isomer numbers) in fact are transient intermediates in various reactions. Thus, reaction (3) can be shown to produce the strained cy-

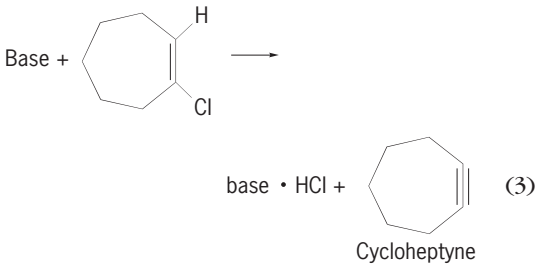

(3)

cloheptyne (bond angles at triply bound carbon are normally 180°), but only if it is conducted at the temperature of dry ice (solidified carbon dioxide); this species can then be detected spectroscopically. As it does not survive warming, the other products observed may be said to form via cyclopheptyne as the intermediate.

There are several additional groups of nonexistent compounds for which other causes have been identified. Thus, enols such as structure (**11**) can usually

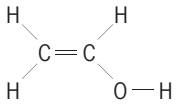

(**11**)

not be isolated, although they play the role of intermediates in many reactions. The reason for their non existence is not strain or crowding, but simply the fact that the barrier to a more stable isomer is low. The so-called tetrahedral intermediates fall in a similar category, for example, structure (**12**). These

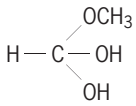

(**12**)

intermediates invariably eliminate water or alcohol very rapidly and hence cannot be isolated.

There are instances in which neither of two isomers can be isolated, but mixtures of the two can. In other words, the barrier between the two is low, and the equilibrium constant is close to unity. An example is acetoacetic ester, which normally contains about 15% of the enol isomer. Such isomers are known as tautomers. *See* TAUTOMERISM.

William J. LeNoble

Bibliography. F. A. Carey and R. J. Sundber, *Advanced Organic Chemistry*, *Pt. A*: *Structure and Mechanisms*, 4th ed., 2000; N. S. Isaacs, *Physical Organic Chemistry*, 2d ed., 1996; W. J. LeNoble, *Highlights of Organic Chemistry*, 1974; J. March, *Advanced Organic Chemistry: Reactions, Mechanisms, and Structures*, 5th ed., 2001.

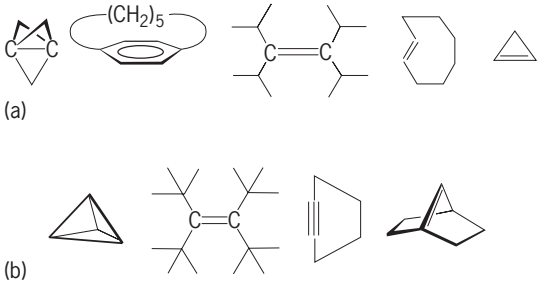

Fig. 3. Examples of high-energy, strained molecules. (*a*) Structures that have been synthesized. (*b*) Structures that have not been synthesized.

Physical science

The fields of inquiry to which the general designation science may be appropriately applied are broadly divided into social science and natural science. The

latter is further subdivided into biology and physical science. Physical science is generally considered to include astronomy, chemistry, geology, mineralogy, meteorology, and physics. These overlap more or less, as illustrated by astrophysics, chemical physics, physical chemistry, and geophysics. There is overlap, likewise, between the physical and biological sciences, as seen in biochemistry, biophysics, virology, and the close relation between geology and paleontology. The boundaries implied in all such classifications are artificial and consist of regions where one field shades into another. *See* ASTRONOMY; CHEMISTRY; GEOLOGY; METEOROLOGY; MINERALOGY; PHYSICS.

Chemistry and physics differ from astronomy, meteorology, and geology in that they are concerned with the properties of matter and energy encountered upon and within the Earth, the planets, and stars. For this reason, chemistry and physics are not set apart from but rather pervade the other sciences.

To regard the several areas of scientific inquiry as separated by sharp definable boundaries is unrealistic. Cross-fertilization of the different fields has produced some of the most notable advances in science, and an artificial barrier can advantageously be accepted as a challenge by a scientist with an adventurous mind. *See* SCIENCE. Joel H. Hildebrand

Physical theory

A physical theory usually involves the attempt to explain a certain class of physical phenomena by deducing them as necessary consequences of other phenomena regarded as more primitive and less in need of explanation. These more primitive phenomena may at the time the theory is formulated be undiscovered, so that part of the proof of the correctness of the theory consists in demonstrating the existence of the unknown assumed primitive phenomena. A classic example is the kinetic theory of gases, in which the pressure of a gas is explained as arising from the kinetic reactions of colliding molecules, the reality of which was established only later by the discovery of phenomena such as the brownian fluctuations.

The value of a theory depends on both the success with which it coordinates a wide range of presently known facts and its fertility in suggesting places to look for presently unknown phenomena. *See* SCIENTIFIC METHODS. Percy W. Bridgman; Gerald Holton

Physical vapor deposition

Production of a film of material often on a heated surface and in a vacuum. Physical vapor deposition technology is used in a variety of applications. Coatings are produced from a wide range of materials, including metals, alloys, compound, cermets, and composites. The two basic processes for physical vapor deposition are evaporation deposition and sputter deposition. In evaporation, thermal energy converts a solid or liquid target material to the vapor phase. In sputtering, the target is biased to a negative potential and bombarded by positive ions of the working gas from the plasma, which knock out the target atoms and convert them to vapor by momentum transfer. *See* ALLOY; CERMET; COMPOSITE MATERIAL; METAL.

Process steps. Physical vapor deposition processes consist of three major steps: generation of the depositing species, transport of the species from source to substrate, and film growth on the substrate. The processes are versatile because the steps occur sequentially and can be controlled independently.

The first step involves generating the depositing species by evaporation using resistance, induction, electron-beam, or laser-beam heating (**Fig. 1***a*), or by sputtering using direct-current or radio-frequency plasma generation (Fig. 1*b*).

In the second step, transport from source to substrate, various flow regimes are used. In molecular flow, the mean free path is larger than the source-to-substrate distance. Molecular flow occurs at low partial pressures of the depositing species and residual gas in the system, and is responsible for the line-of-sight feature typical of evaporation deposition processes and low-pressure magnetron-type sputter deposition processes. Viscous flow occurs at higher partial pressures (2.7–16 pascals or 20–120 millitorr) typical of diode sputter deposition; it also is intentionally added in the evaporation deposition process to cause gas scattering of the depositing species and thereby to increase the throwing power of the process.

An additional feature in the second step is the absence or presence of a plasma in the source-to-substrate region and the mode by which the plasma is excited, for example, direct current, radio frequency, or microwave. The mode of plasma excitation is an important aspect, since it controls the electron energy and distribution function and thereby the chemical reactions that take place in the plasma. The plot of ionization cross section versus electron energy peaks at low electron energies (50–100 eV). Processes that involve low electron energies, such as plasma-assisted evaporation, in which the electron energies generating the plasma can be controlled independently, offer a more versatile and a broader range of chemical reactions in the plasma than processes such as sputtering, in which the electron energies are dictated by other considerations such as target voltage (500–1000 eV), which controls the rate of sputtering. In sputter deposition, the electron energies cannot be controlled independently of other process parameters. The presence of a plasma is optional in the evaporation process but integral to the sputtering process.

The third step in physical vapor deposition is film growth on the substrate. The process depends on the energy of the incident species, which is typically 0.5 eV for evaporation and 10–100 eV for sputtering, and on the substrate temperature. The structure, composition, and residual stress in the film can be changed substantially by bombarding the growing film with energetic ions or neutral atoms. They can

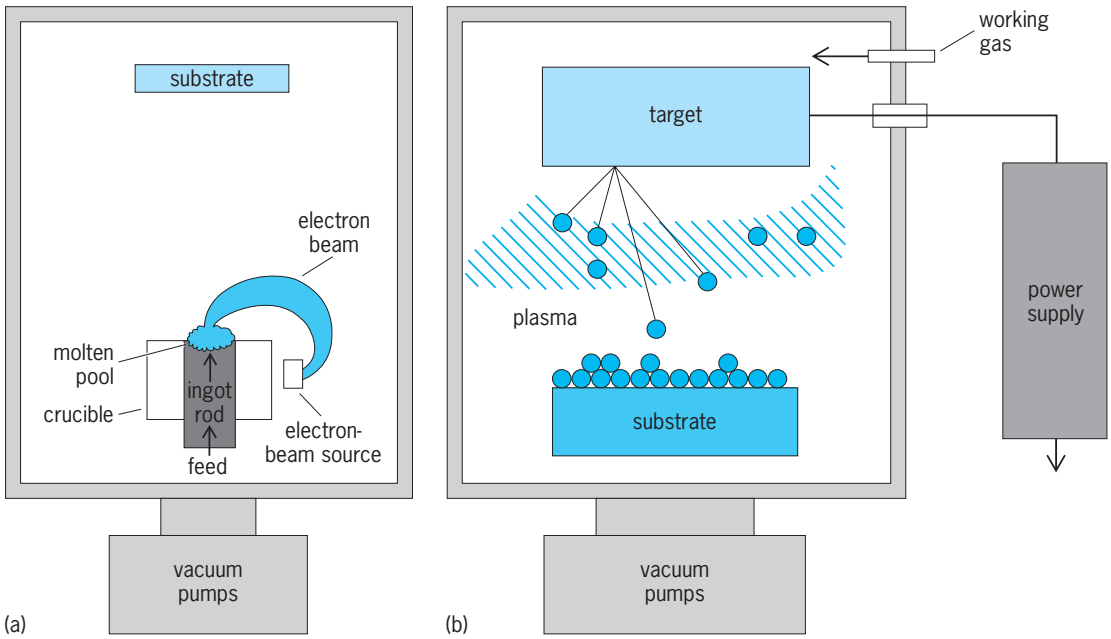

Fig. 1. Processes for generation of depositing species within vacuum chambers. (*a*) Evaporation deposition process using an electron-beam evaporation source. (*b*) Sputter deposition process.

be generated by a separate ion source, or they can be attracted to the film from the plasma by electrical biasing of the substrate/film. Thus, the location of the substrate inside or outside the plasma can substantially change the nature and amount of ion bombardment. In magnetron sputtering, the plasma is confined to a narrow zone near the target. Therefore, if the film is to be bombarded, a second plasma has to be created near the substrate by a suitable method. *See* MAGNETRON.

Single elemental species deposition. For single elemental species, the rate of evaporation is controlled by the temperature of the source, and this in turn governs the vapor pressure. At a given temperature, the rate of evaporation can vary over 10 orders of magnitude for the various elements in the periodic table. The sputtering rate for a given sputtering gas at a fixed pressure depends on the sputtering yield. The difference in sputtering yield for various elements is about 100 (ratio of the highest to the lowest sputtering yield), which is considerably smaller than the vapor pressure ratio for evaporation. In general, sputtering rates are 10–5000 times lower than evaporation rates.

Alloy deposition. Alloy deposition by evaporation can be accomplished in three ways: coevaporation, evaporation from a single source, and flash evaporation. Coevaporation involves the use of a number of sources corresponding to the number of elements in the alloy (**Fig. 2***a*). Its advantage is the ability to control the evaporation rate from each source. Its disadvantage is that the alloy composition varies along the plane of the substrate.

Evaporation from a single source (Fig. 2*b*) operates under steady-state conditions where the composition and volume of the liquid pool on top of a solid rod are kept constant. For a thermodynamically ideal solution, the elemental composition of the vapor above the liquid alloy is proportional to the product of the mole fraction and the vapor pressure of the pure elements. For the example in Fig. 2*b*, if the alloy has a composition A_1B_1 and the vapor pressure of B is 10 times the vapor pressure of A, a pool of the composition $A_{10}B_1$ will generate a vapor species A_1B_1. The composition and volume of the molten pool is kept constant by raising the rod at a rate that allows the same amount of the solid alloy to be melted into the pool as is lost from the pool by evaporation. This technique is used commercially for coating a five-component alloy of nickel-cobalt-chromium-aluminum-yttrium (Ni-Co-Cr-Al-Y) onto turbine blades. Operationally, the maximum difference in vapor pressures between the elements in the alloy cannot be greater than 5000, due to the relationship between the rate of diffusion in the liquid and the rate of evaporation in maintaining the compositional homogeneity of the molten pool.

In the flash evaporation process, pellets of an alloy are dropped onto a very hot strip and are vaporized completely, thus maintaining the composition of the alloy in the deposit. This process works very well for elements with high vapor pressures. In a version of flash evaporation known as pulsed-laser or electron-beam ablation, a high-energy pulse vaporizes a small chunk of the alloy, thus retaining the composition of the alloy in the deposit. Cathodic arc evaporation, which evaporates small volumes of the target, may also be used. However, there is a problem of droplet ejection from the target that causes major flaws in the deposited film; this impedes its use in some applications of thin-film deposition. *See* LASER.

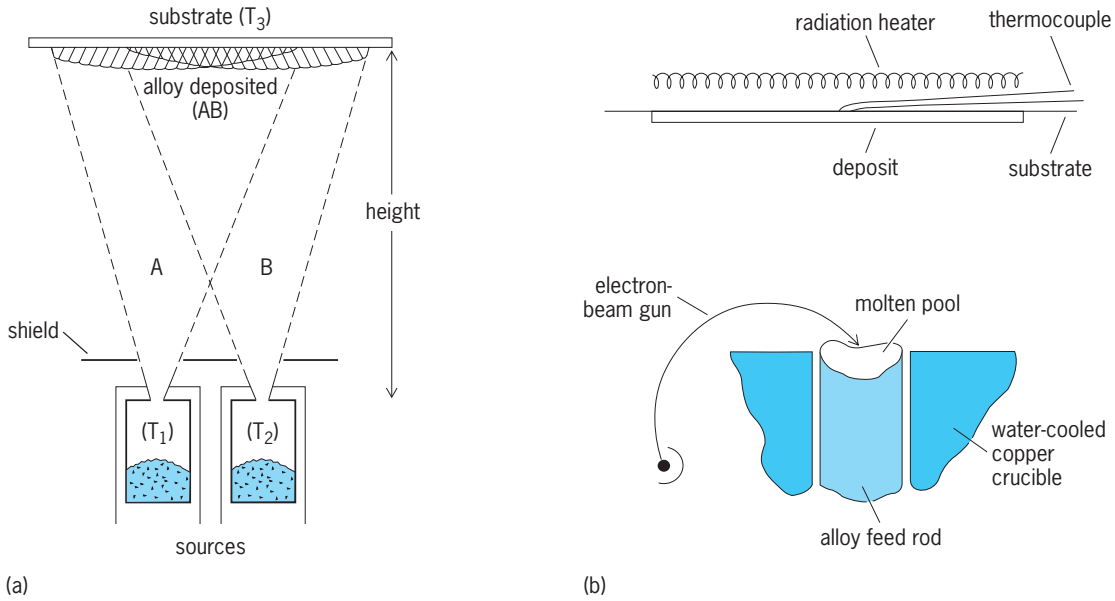

Fig. 2. Alloy deposition by evaporation. (*a*) Two-source (A and B) evaporation arrangement yielding variable film composition. T1, T2, and T3 are temperatures of A, B, and substrate, respectively. (*b*) Direct evaporation of an alloy from a single rod-fed source.

Sputter deposition. Alloy deposition by sputtering is analogous to the evaporation method. In cosputtering, several sputter cathodes or sputter guns are used to generate the various vapor species. The same composition variation in the plane of the substrate occurs as in evaporation. Due to the low sputtering rates, there is not as much latitude to increase the source-to-substrate distance as with evaporation without making the deposition rates impractically low. A variation of cosputtering uses a segmented target. In this case, the sputtering target is built up of segments of the elements composing the alloy, the surface area of each element being inversely proportional to the sputtering yield. Increasing the number of segments increases the compositional homogeneity in the film.

An alloy can be sputtered from an alloy target if the modified surface composition is inversely proportional to the sputtering yield of the various elements. This is accomplished by a conditioning period of the target to allow progress from the alloy composition to the modified surface composition. *See* SPUTTERING.

Deposition of compounds. Deposition of compounds can be performed in two ways. In direct evaporation/sputtering, the composition of the target is the same as that of the compound that is to be deposited; for example, both the target and film composition may be the oxide of yttrium (Y), barium (Ba), and copper (Cu), $YBa_2Cu_3O_{7-x}$. In reactive evaporation/sputtering, the element or elements of the compound are evaporated/sputtered and react with the gas to form the compound. For example, Y, Ba, and Cu are evaporated/sputtered and react with oxygen to form $YBa_2Cu_3O_{7-x}$. Plasma-assisted reactive processes are often used, since they activate the reactions leading to compound formation. The two plasma-assisted reactive processes for physical vapor deposition are activated reactive evaporation and reactive sputtering. These processes are of great importance for deposition of compound films in the manufacture of various devices.

Applications. Coatings provided by vapor deposition have many applications. For optical functions they are used as reflective or transmitting coatings for lenses, mirrors, and headlamps; for energy transmission and control they are used in glass, optical, or selective solar absorbers and in heat blankets; for electrical and magnetic functions they are used for resistors, conductors, capacitors, active solid-state devices, photovoltaic solar cells, and magnetic recording devices; for mechanical functions they are used for solid-state lubricants, tribological coatings for wear and erosion resistance in cutting and forming tools and other engineering surfaces; and for chemical functions they are used to provide resistance to chemical and galvanic corrosion in high-temperature oxidation, corrosion, and catalytic applications as well as to afford protection in the marine environment. Decorative applications include toys, costume jewelry, eyeglass frames, watchcases and bezels, and medallions. Coatings also act as moisture barriers for paper and polymers. They can be used for bulk or free-standing shapes such as sheet or foil tubing, and they can be formulated into submicrometer powders.

In some cases, a coating may serve multiple requirements. For example, a transparent conducting oxide may be a component of a *pn* junction as well as a current collector in a solar cell, or a gold-colored titanium nitride (TiN) coating may give a pleasing color as well as protect against wear. *See* VACUUM METALLURGY.

Rointan F. Bunshah

Bibliography. R. F. Bunshah (ed.), *Handbook of Deposition Technologies for Films and Coatings: Science, Applications and Technology*, 2d ed., 1994;

A. Eishabini-Riad and F. D. Barlow, *Thin Film Technology Handbook*, 1997; J. E. Mahan, *Physical Vapor Deposition of Thin Films*, 2000; D. M. Mattox, *Handbook of Physical Vapor Deposition (PVD) Processing*, 1998; M. Ohring, *The Materials Science of Thin Films*, 2d ed., 2001; D. L. Smith, *Thin-Film Deposition: Principles and Practice*, 1995; K. Wasa, M. Kitabatake, and H. Adachi, *Thin Film Materials Technology: Sputtering of Compound Materials*, 2003.

Physics

Formerly called natural philosophy, physics is concerned with those aspects of nature which can be understood in a fundamental way in terms of elementary principles and laws. In the course of time, various specialized sciences broke away from physics to form autonomous fields of investigation. In this process physics retained its original aim of understanding the structure of the natural world and explaining natural phenomena.

Basic parts. The most basic parts of physics are mechanics and field theory. Mechanics is concerned with the motion of particles or bodies under the action of given forces. The physics of fields is concerned with the origin, nature, and properties of gravitational, electromagnetic, nuclear, and other force fields. Taken together, mechanics and field theory constitute the most fundamental approach to an understanding of natural phenomena which science offers. The ultimate aim is to understand all natural phenomena in these terms. *See* CLASSICAL FIELD THEORY; MECHANICS; QUANTUM FIELD THEORY.

The older, or classical, divisions of physics were based on certain general classes of natural phenomena to which the methods of physics had been found particularly applicable. These consisted of classical mechanics with branches in celestial mechanics, hydrodynamics, and ballistics; heat and thermodynamics; kinetic theory of gases and statistical mechanics; optics; acoustics; and electricity and electromagnetism. These divisions are all still current, but many of them tend more and more to designate branches of applied physics or technology, and less and less inherent divisions in physics itself.

Branches. The divisions, or branches, of modern physics are made in accordance with particular types of structures in nature with which each branch is concerned. Thus particle physics, or high-energy physics, is the most recent branch and is concerned with understanding the properties and behavior of elementary particles, and more particularly of the heavy particles—mesons, baryons, and their antiparticles—which are produced in collisions involving energies in a range measured in billions of electronvolts. The next branch in this classification is nuclear physics, which is concerned with associations of neutrons and protons forming the nuclei of atoms; their structure, properties, and energy states; reactions between nuclei, including scattering processes and radioactivity; and related phenomena, such as the interaction of high-speed nuclear particles with matter. Atomic physics is concerned with the structure and properties of atoms as determined by the electrons outside the nucleus; the states of motion of these electrons, including such topics as energy levels, angular momentum properties, and magnetic moments; and the absorption and emission of radiation by atoms.

Continuing with this classification in ascending complexity there is molecular physics, which is concerned with systems of atoms formed into molecules, the nature of intermolecular forces, chemical binding, vibration and rotation spectra of molecules, and the like. Next in order are solid-state physics; physics of liquids; physics of gases; and plasma physics, which deals with properties of highly ionized atoms forming a mixture of bare nuclei and electrons called an ion plasma.

In this same classification could also be included biophysics, which deals with the application of physical methods and types of explanation to biological systems and structures.

Other more specialized classifications may be made in accordance with particular instruments or techniques, such as x-ray diffraction, neutron diffraction, mass spectrometry, infrared spectroscopy, and seismology. The special field of low-temperature physics is characterized not only by special instruments involved in the production and measurement of low temperatures in the range of liquid helium but also by the phenomena of superconductivity and superfluidity which occur only in this temperature range. Other fields, such as astrophysics and geophysics, are concerned with aspects of other sciences to which physics is applicable.

Mathematical physics is the study of physical phenomena by means of mathematics, and includes the more mathematical parts of all branches of physics, as well as most of the content of statistical mechanics, quantum mechanics, relativity, and field theory. A distinction is often made between mathematical physics and theoretical physics, in which the latter, although still entirely mathematical in form, is thought of as being more closely related to experimental physics. Neither mathematical nor theoretical physics can really be separated from experimental physics, since a complete understanding of nature can only be obtained by the application of both theory and experiment.

Aim. In every area physics is characterized not so much by its subject-matter content as by the precision and depth of understanding which it seeks. The aim of physics is the construction of a unified theoretical scheme in mathematical terms whose structure and behavior duplicates that of the whole natural world in the most comprehensive manner possible. Where other sciences are content to describe and relate phenomena in terms of restricted concepts peculiar to their own disciplines, physics always seeks to understand the same phenomena as a special manifestation of the underlying uniform structure of nature as a whole. In line with this objective, physics is characterized by accurate instrumentation,

precision of measurement, and the expression of its results in mathematical terms.

For the major areas of physics and for additional listings of articles in physics *see* ACOUSTICS; ATOMIC PHYSICS; BIOPHYSICS; CLASSICAL MECHANICS; ELECTRICITY; ELECTROMAGNETISM; ELEMENTARY PARTICLE; FLUID MECHANICS; HEAT; LOW-TEMPERATURE PHYSICS; MOLECULAR PHYSICS; NUCLEAR PHYSICS; OPTICS; SOLID-STATE PHYSICS; STATISTICAL MECHANICS; THEORETICAL PHYSICS; THERMODYNAMIC PRINCIPLES.

William G. Pollard

Physiological acoustics

The study of specific responses that take place in the ear or in the associated central (neural) auditory pathways. The ear is the receptor organ of mechanoacoustic energy, which occurs in the form of sound pressure waves. The responses are elicited by appropriate stimuli of well-defined parameters presented at any level of the auditory system. Such responses may be registered with the aid of various, usually invasive, recording techniques which make use of mechanical, electrical, optical, radiological, or biochemical phenomena, or their combinations. The approach employed by physiological acoustics therefore is purely analytical. This is in contrast to the noninvasive, holistic approach employed by psychoacoustics, which lends itself well to experiments on human subjects. Systematic physiological experiments can be performed only in animals. *See* PSYCHOACOUSTICS.

To bridge the gap between the two fields of inquiry, psychoacoustic techniques are employed in experimental animals that have received appropriate training. Such behavioral studies are combined with physiological methods, for example by setting lesions at selected places in the auditory system and testing various aspects of the animal's hearing function before and after. Postmortem histological and other controls may then be obtained at a time of the experimenter's choosing. Such experiments have shown clearly that the differences in results obtained in humans and in animals are usually in degree, and rarely in principle.

Anatomy

A schematic outline of the anatomy of the central auditory system is given in **Fig. 1**. The fibers of Corti's ganglion cells (or spiral ganglion cells), approximately 25,000 to 30,000 in number, connect peripherally to the hair cells. Centrally, these fibers terminate on a second set of nerve cells in either of the two cochlear nuclei, forming synaptic junctions with them. Fibers of the latter cell continue on a central course, most of them crossing over onto the opposite side of the head, following a general pattern of the nervous system. Additional synaptic junctions, each connecting with the next higher set of neurons, are encountered sequentially: first in the superior olivary complex, where for the first time fibers stemming from both sides are brought together on the same cells; in the nucleus of the lateral lemniscus, an upward running tract; in the inferior colliculus; and in the medial geniculate body. Fibers projecting from here terminate in the auditory cortex, which in humans is almost hidden in a deep furrow, known as the Sylvian fissure.

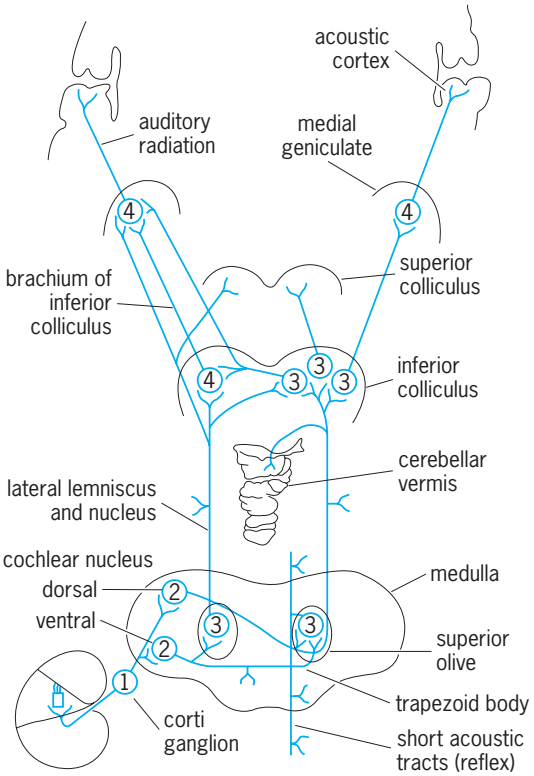

Fig. 1. Schematic outline of the central auditory pathways. Numbers indicate the number of neurons from the cochlea. (*After S. S. Stevens, ed., Handbook of Experimental Physiology, John Wiley and Sons, 1951*)

These various nuclei are much more than mere switching stations, and signal processing is not limited to the cortex. Some specific data handling goes on at each nucleus. What the cortex receives has already been extensively processed at lower levels. The cross-connection between the inferior colliculi on both sides (Fig. 1), for example, obviously aids binaural processing at that level.

Auditory Function

Mechanoacoustic energy received by the ear elicits a number of specific, interlinked events, first in the ear itself, then along the neural pathways, and finally in the auditory cortex. It is only after this sequence of events is completed that a sensation of sound commensurate with the applied signal is experienced by the listener.

Peripheral system. Sound pressure waves originating in the environment enter the external ear and are transmitted through middle and inner ears in the form of mechanoacoustic vibrations.

External ear. The pinna serves to collect sound from the environment in a directional, frequency-dependent manner, reception for most frequencies being maximal for normal incidence. This selectivity

aids directional hearing, even of one ear alone. Because of its shape and length, the ear canal resonates at approximately 2700 Hz. The resonance is rather broad (indicating moderate damping of the sound vibration in the ear canal). Hence, the pressure acting on the tympanic membrane is increased over a range of almost two octaves, maximally by about 20 dB at the resonant frequency and for normal incidence. *See* RESONANCE (ACOUSTICS AND MECHANICS).

Middle ear. The middle ear serves to match impedances between the air in the ear canal which has low impedance, and the fluid of the inner ear, which has high impedance. To this end, it employs the principle of a hydraulic lever (in addition to the small mechanical leverage of the ossicular chain): the area of the tympanic membrane is about 20 times larger than that of the stapes footplate. However, this mechanical transformer is optimally efficient over only a limited frequency range from about 1000 to 4000 Hz. One reason for this is that the tympanic membrane is not displaced like a stiff plate; its displacement pattern changes drastically at higher frequencies, at which point it becomes a relatively inefficient driver of the lever system. A second reason is that at frequencies below 1000 Hz the input impedance of the system increases at 12 dB per octave with inverse frequency. It has been shown that the shape of the auditory threshold curve is entirely determined by the middle ear.

In the midfrequency range, the middle ear optimizes the transfer of mechanoacoustic energy into the inner ear, but only at low-input levels. At higher input levels, it acts as a complex, mechanical, nonlinear network, absorbing some of the incident energy and thus to some degree protecting the inner ear against overloading. The middle-ear muscles, the tensor tympani and the stapedius, play an important role in this. By contracting reflexly, they reduce transmission when the input exceeds 70–75 dB SPL. Furthermore, displacements of the tympanic membrane serve to pump air into and out of the mastoid air cells. Hence, these cells serve as a mechanical damping system in exactly the same manner as the holes in the backplate of a condenser microphone. *See* LOUDNESS; MICROPHONE.

The performance of the middle-ear transformer is optimal only when there is no static air-pressure difference across the tympanic membrane. Middle-ear pressure is equalized from time to time via the Eustachian tube, which is briefly forced open by swallowing, yawning, and a number of other activities.

Inner ear. Incident mechanoacoustic energy displaces the stapes footplate, and thus the perilymphatic fluid of the scala vestibuli. Since the stiffness of the basilar membrane varies exponentially over its length (it becomes more compliant toward the helicotrema), the cochlea acts like a hydrodynamic, tapered transmission line (or delay line). The cochlear duct undergoes alternating displacements, which progress in a wavelike manner from base to apex. In this process, the round window at the basal end of the scala tympani acts as a relief valve. The traveling waves form localized displacement maxima at frequency-dependent places; they are located near the base for high frequencies and shift systematically toward the apex with inverse frequency (**Fig. 2**). This distribution of frequency converts frequency information into spatial information; such a frequency distribution is called tonotopic. As the delay-line property demands, there is a cochlear time delay that maximally (for low frequencies traveling all the way up to the apex) amounts to 4–5 ms. Both the tonotopic distribution and the spatial time delay (systematic with inverse frequency) are preserved throughout the entire auditory system. Seen from an individual location along the basilar membrane, the tuning is very sharp, a bandpath-filter property that is once more maintained throughout the mechanical system and part of the neural system as well. *See* ELECTRIC FILTER.

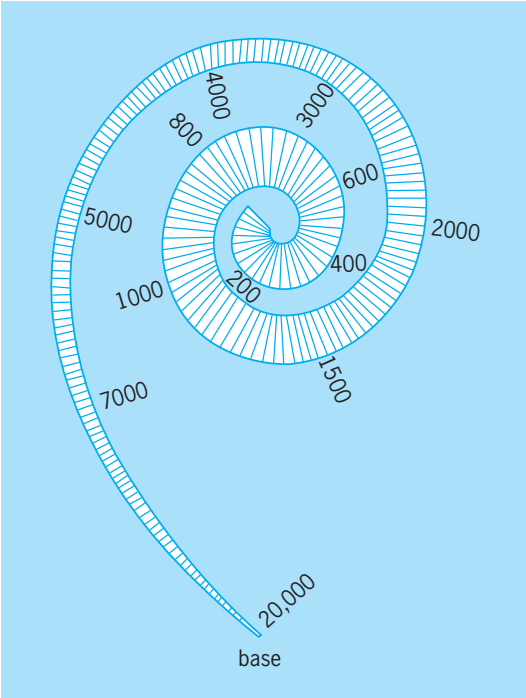

Fig. 2. Distribution of frequency maxima along the basilar membrane. Numbers give frequencies in hertz. Length of basilar membrane = 32 mm, width = 0.04 mm at base, 0.50 mm at apex. (*After O. Stuhlmann, An Introduction to Biophysics, John Wiley and Sons, 1943*)

Shearing motion. The traveling-wave displacements of the cochlear duct, in which the organ of Corti participates, produce tangential shearing motions between the organ proper and the tectorial membrane, to which the sensory hairs are attached. The shearing forces thus generated constitute the ultimate mechano-acoustic input to the hair cells. In this process, a second pressure transformation takes place (at the expense of displacement) and, perhaps, the tuning is also slightly sharpened.

Receptor potentials. At the sensory hair cells, the original mechanoacoustic energy is converted into a chemoelectric form of energy. Two kinds of electrical receptor potentials are generated in each cell being simulated, one ac and the other dc. The ac potential, an analog signal, has the general properties

of the applied acoustic stimulus as seen through a bandpass filter, sharply tuned for reasons already discussed. The dc signal, similarly, is mainly evoked by frequencies within the narrow pathband given by the cell's position (**Fig. 3**). The hair cells act as biological amplifiers; that is, their receptor potentials possess more energy than the applied mechanoacoustic signal.

Cochlear microphonics and summating potentials. The ac as well as the dc potentials spill over in the cochlear fluids, where the potentials from many cells are summed; for ac events this occurs according to their local phases, which are determined by the spatial delay of the traveling waves. Both potentials in their extracellular, summed forms can be recovered from the cochlear fluids or from the round-window membrane—they are called cochlear microphonics (ac) and summating potentials (dc) respectively. They are valuable research tools, even though they are rather broadly tuned.

The sensitivity of the ear is extremely high. In a quiet environment, a sound pressure of a mere 20 micropascals is sufficient for threshold stimulation in the most sensitive frequency region, around 3000 Hz. The resulting displacement of the cochlear duct is probably on the order of 100 nanometers. This extreme sensitivity is made possible, in addition to other factors, by the extremely low input of limiting background noise (essentially constituted by brownian motion) at the hair cell. *See* BROWNIAN MOVEMENT.

Hair cell–nerve fiber junction. A sequence of three events leads to the discharge of action potentials in the nerve fiber connected to a given hair cell (1) the receptor potentials of the hair cells; (2) the release of a chemical neurotransmitter, not yet clearly identified, into the synaptic cleft between hair cell and nerve fiber dendrites; and (3) dc generator potentials within the dendrites that, after spatial and temporal integration, finally trigger an action potential. Thus, the original acoustic signal is uniquely encoded as a series of brief electric discharges in a number of nerve fibers. Action potentials are the universal neural signals traveling in all nerve fibers. *See* BIOPOTENTIALS AND IONIC CURRENTS; SYNAPTIC TRANSMISSION.

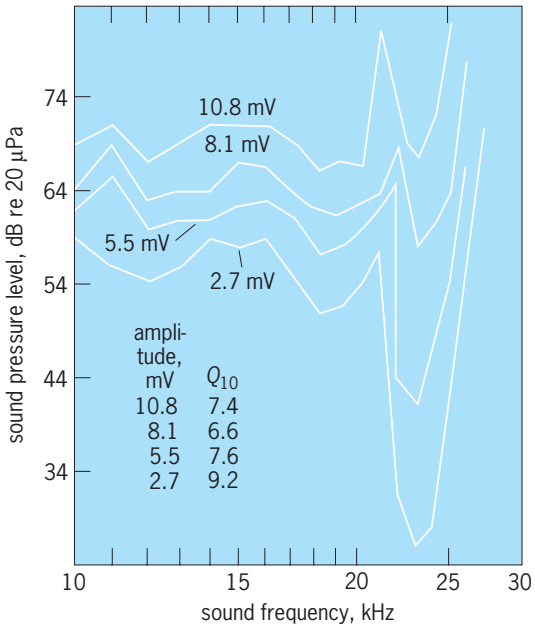

Fig. 3. Isoamplitude curves of the dc receptor potential within one guinea pig inner hair cell of the basal region. Q_{10} is a measure of the width of each curve, equal to the center frequency divided by the bandwidth at 10 dB above threshold. (*After I. J. Russell and P. M. Sellnick, Tuning properties of cochlear hair cells, Nature, 267:858–860, 1977*)

Neural system. The various synapses of the central auditory pathway (Fig. 1) are much more than mere switching stations—some specific data handling goes on at each of them. What the cortex of the brain receives has already been extensively processed at lower levels, aided by cross-connections between individual cells. For example, the bilateral innervation of superior olivary cells and the cross-linkage between the inferior colliculi on both sides (Fig. 1) aid binaural processing at those relatively low levels, which aids directional hearing, even though ultimate "decisions" are made at the cortical level.

Cochlear nerve fibers. About 90–95% of all cochlear nerve fibers connect exclusively with inner hair cells, each fiber linking up with only one cell, but each cell receiving up to 20 fibers (neural divergence). In contrast, each outer-hair-cell fiber connects to approximately 10 cells in an apparently random sequence, but each cell receives only one branch of one fiber (neural convergence). Since the outer-hair-cell fibers are so few in number, and since the diameter of each is rather small, they are probably rarely encountered by physiologist's electrodes; as a result, their response patterns have not been clearly identified. The following comments therefore apply only to inner-hair-cell fibers, which have been extensively studied.

One reason why the cochlear-nerve-fiber code has only been partly deciphered is that one outer-hair-cell fiber, as mentioned above, connects to more than one cell in random sequence. Furthermore, in each cochlear nerve fiber, there exists a "noise"; that is, random, spontaneous discharges occur that compete with the evoked potentials so that following any discharge, nerve fibers are unable to respond for approximately 1 ms, their so-called refractory period. For these reasons, the evoked firing of all nerve fibers is a highly complicated, probabilistic process. Cochlear nerve fibers increase their firing rate over the spontaneous level when excited by a signal. At higher levels, therefore, the spontaneous firings are reduced in number as the evoked discharges become dominant. However, the dynamic range of a single fiber is a mere 40 dB, and its linear portion is even smaller (**Fig. 4**). This is in marked contrast to the dynamic range of actual hearing, that is, the differentiation of loudness, which extends over at least 120 dB for frequencies around 3000 Hz.

The tuning of cochlear nerve fibers, based on an increased firing rate, is very sharp. The difference in Q-factor between high-frequency and low-frequency

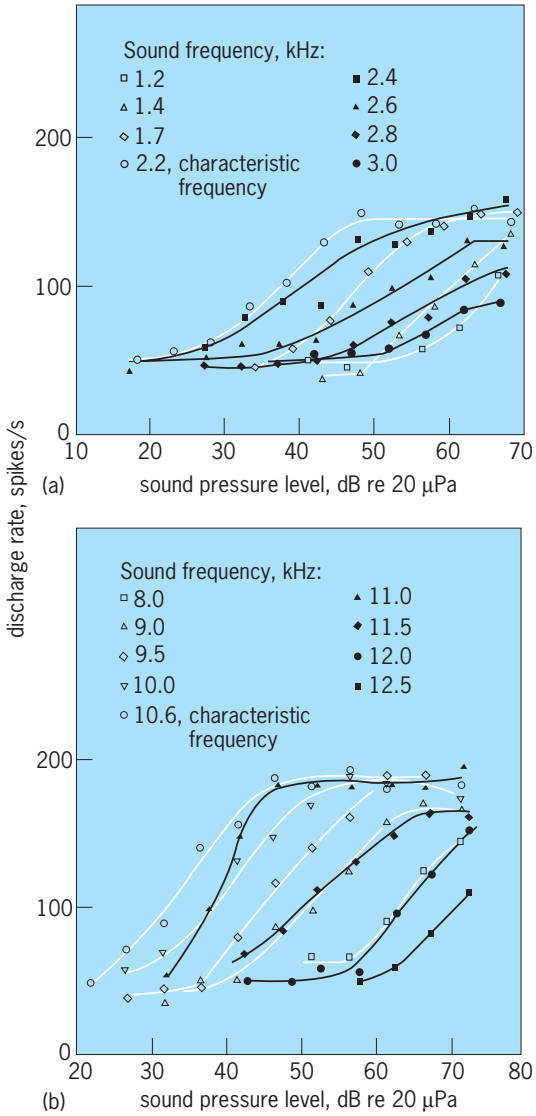

Fig. 4. Intensity/rate functions of two cat cochlear nerve fibers tuned to (*a*) 2.2 kHz and (*b*) 10.6 kHz. (*After W. D. Keidel and W. D. Neff, eds., Handbook of Sensory Physiology, Springer, 1975*)

fibers shown in **Fig. 5** is more apparent than real. It largely disappears when the data are plotted against a linear frequency scale instead of the conventional logarithmic one. The frequency to which each fiber is most sensitive is known as its best (or characteristic) frequency.

Because of the probabilistic nature of neural responses, most forms of assessment require statistical methods; the histogram is a frequently used tool. **Figure 6** presents some examples of post-stimulus-time (PST) histograms of a number of different fibers tuned to various frequencies. Responses are elicited by repeated and periodically timed signal presentations, usually 5000 in number. The height of each column at each time point represents the frequency of response at that point. Following each brief click signal, there is an extended period of responses, whose frequency of occurrence decays exponentially with time. In high-frequency fibers (best frequencies higher than 3–4 kHz), the firing occurs at random. In low-frequency fibers, however, the firing is periodic, synchronized with the period of the best frequency of the fiber.

The best frequency reflects the location along the cochlea at which the fiber is connected to its hair cell and the frequency of the cochlear traveling wave which has its maximum at that location. Thus, the responses of low-frequency fibers, including their exponential time decay, duplicate the mechanical "ringing" occurring locally in each (narrow) cochlear filter. The reason for the two different types of responses is that the firing of high-frequency fibers is triggered by the hair-cell dc receptor potential, with no possibility of time locking, whereas that of low-frequency fibers is triggered by the ac potential, to which responses may become time-locked.

Triggering by the ac potential also allows the firing of low-frequency fibers to be phase-locked to the period of a sine-wave signal. This phase locking can achieve frequency selectivity and sensitivity superior to that based on increased firing rate (Figs. 4 and 5). Signal frequency thus is encoded in two different ways: by phase-locked periodicity, mainly in low-frequency fibers, and by location, mainly in high-frequency fibers.

Intensity is also encoded in a twofold manner: by the firing rate in the best-frequency fiber and by the recruitment of additional fibers tuned to nearby frequencies. These latter fibers start to respond at higher levels, that is, when their tuning curves have become broad enough to include the frequency in question. Integration of such an intensity-staggered rate function over a number of fibers would offer a dynamic range wider than that of a single fiber, but there is no direct evidence for this notion as yet. Although the two codes are seen to overlap partially, there is hardly any confusion on the part of a listener

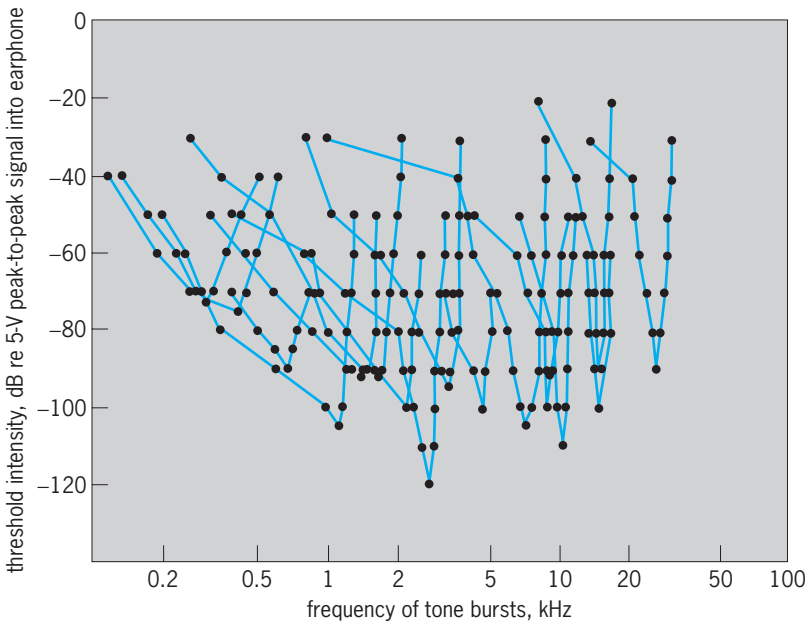

Fig. 5. Tuning curves of a number of nerve fibers in the cat cochlear nerve. (*After N. Y.-S. Kiang, Discharge Patterns of Single Fibers in the Cat's Auditory Nerve, MIT Res. Monogr. 35, MIT Press, 1965*)

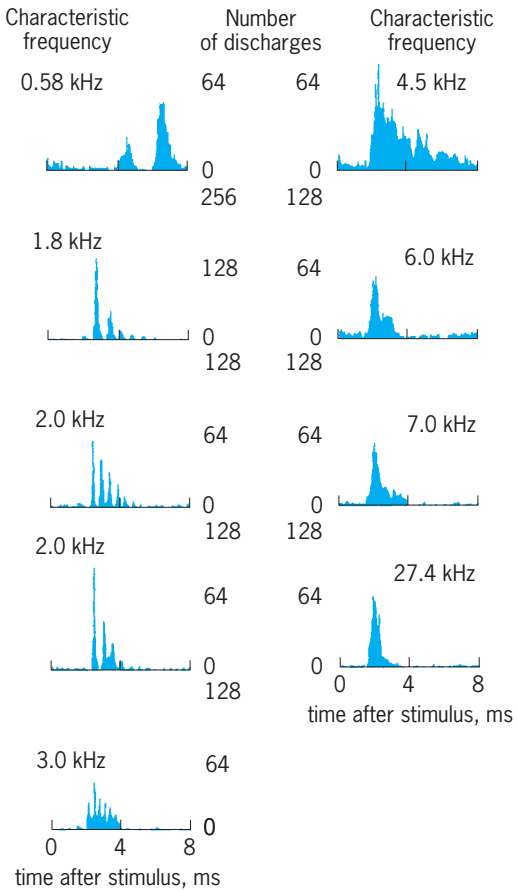

Fig. 6. PST histograms for a number of nerve fibers in a cat with various characteristic frequencies, showing response to click stimulation. Each histogram represents 1 minute of data (600 clicks at click rate of 10 per second). (*After N.Y.-S. Kiang, Discharge Patterns of Single Fibers in the Cat's Auditory Nerve, MIT Res. Monogr. 35, MIT Press, 1965*)

between intensity changes and frequency changes: there is only a very slight pitch change as signals become louder.

Except for the transduction process in the hair cell–nerve fiber junction, there is no evidence for signal processing in cochlear nerve fibers and their associated bipolar ganglion cells. In outer-hair-cell fibers, with their multiple inputs from different cells, there should be some integration, but proof for this assertion is lacking.

Cochlear nuclei. In the cochlear nuclei of the brainstem, there is the first real evidence for neural processing. For example, some cells, aptly named octopus cells, connect across approximately 10 parallel channels. Their tuning curves therefore become considerably broader. In addition to excitation (or no excitation) as found exclusively in cochlear nerve fibers, the outputs from cochlear nucleus cells may show inhibition (that is, a reduction in the rate of spontaneous discharges). Inhibition usually occurs in a frequency region slightly remote from the best frequency, where excitation is invariably found. Furthermore, in addition to the "primarylike" responses of cochlear nerve fibers (Fig. 6), there are three new patterns, descriptively called choppers, pausers, and "on" patterns, although their significance with respect to processing is not understood (**Fig. 7**).

Superior olivary complex. Cells in the superior olivary complex (Fig. 1) are the first to receive inputs from both cochleas; in general, those from the same side are inhibitory and those from the opposite side excitatory. Responses from such cells are controlled by differences in binaural inputs, that is, by interaural time or intensity differences. Their actions at this relatively low level presage psychoacoustic phenomena that have to do with binaural hearing, which aids in determining the direction of a given signal source.

Higher centers. The inferior colliculus is a large, conspicuous nucleus. In some lower animals it is even larger, relatively, than in humans. In birds and reptiles, which have a very thin brain cortex, it is the highest nuclear mass of the auditory system. Significantly, even in higher mammals, a number of ascending fibers terminate here. Some auditory tasks, such as amplitude discrimination, are carried out at this level, as has been judged from experiments in which the auditory cortex is surgically ablated. Several acoustic reflexes are controlled in this nucleus, for example the reflex contraction of the middle-ear muscles in response to sounds above 70–75 dB SPL. Also, the inferior colliculus controls the turning of the eyes and of the head toward a sound source, reflexes that aid auditory localization, which is optimal near the midline of the head.

The medial geniculate body belongs to a group of large nuclei, collectively known as the thalamus. All ascending sensory signals (except those of the olfactory receptors) pass through thalamic centers before being distributed to specific cortical areas.

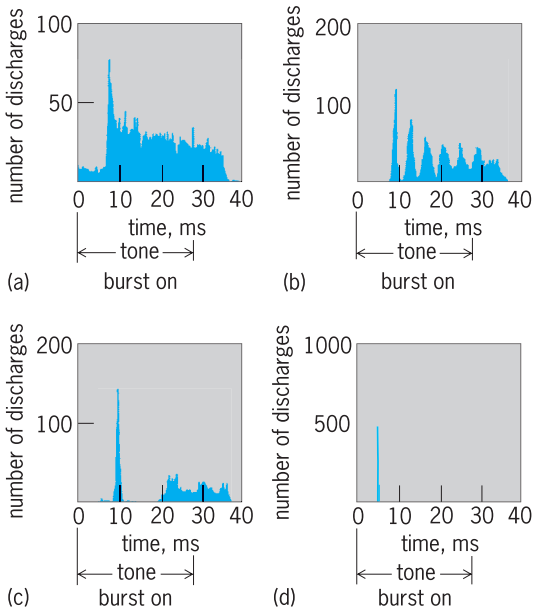

Fig. 7. Response patterns of cells in cat cochlear nuclei, with characteristic frequencies indicated. (*a*) Primarylike response, 3.5 kHz. (*b*) Chopper response, 7.0 kHz. (*c*) Pauser response, 6.6 kHz. (*d*) "On" pattern, 9.1 kHz. (*After R. R. Pfeiffer, Classification of response patterns of spike discharges for units in the cochlear nucleus: Tone burst stimulation, Exp. Brain Res., 1:220–235, 1966*)

How closely the thalamic centers are associated with the cortex is shown by the fact that, when a given cortical area is destroyed, its thalamic nucleus degenerates in a paradoxical, or retrograde, fashion—against the direction of signal flow. The medial geniculate bodies of both sides are interconnected by the commissure of Gudden, another structure that aids in binaural hearing. At this level, there are not only cells that respond to binaural stimulation as such, but also some that register the movement of an auditory target and the direction of such movement.

In all nuclei of the central auditory system, there are well-ordered tonotopic distributions, although multiple representations exist, for instance in the inferior colliculus. Their purpose remains unclear.

Auditory cortex. In contrast to those of the lower levels of the auditory system, cells of the auditory cortex show little, if any, spontaneous activity. A clear-cut tonotopic organization was finally established after considerable experimental difficulties were overcome. Cells of the "primary" auditory area respond directly to acoustic inputs, those of the "secondary" areas only after the primary areas have been paralyzed. The significance of this empirical fact is not yet clear.

With respect to the anatomical and physiological substrates of higher auditory functions, such as acoustic memory (both short-term and long-term), auditory associations, and the like, there is little known, except for some serendipitous findings made at the occasion of surgical interventions. Electrical stimulation of specific areas of the temporal lobe surface have given rise to sensations, such as voices known in the past or melodies or ditties long forgotten. However, no systematic studies have been possible on such memory releases, nor is anything known about the underlying mechanisms. The left auditory cortex, to which most of the fibers originating in the right ear project, fulfills mainly intellectual tasks, such as word recognition, whereas the right cortex fulfills emotional tasks, such as the appreciation of music.

Efferent system. Almost duplicating the afferent (ascending) auditory system (Fig. 1) is an efferent (descending) system which is just as elaborate (**Fig. 8**). The final projection of the efferent system to the cochlea is via the olivocochlear bundle of Rasmussen. Specialized nerve endings, different in appearance from their afferent counterparts, link with hair cells or their afferent nerve fibers.

The discovery of the olivocochlear bundle gave the first evidence for the existence of sensory efference, which has since been demonstrated in all other sensory systems. In anesthetized animals, central stimulation of the auditory efferent system was shown to decrease the discharges in cochlear nerve fibers, somewhat akin to the action of a governor controlling an engine; but it also increases the magnitude of the cochlear microphonics. Its function in awake and unrestrained animals has been difficult to

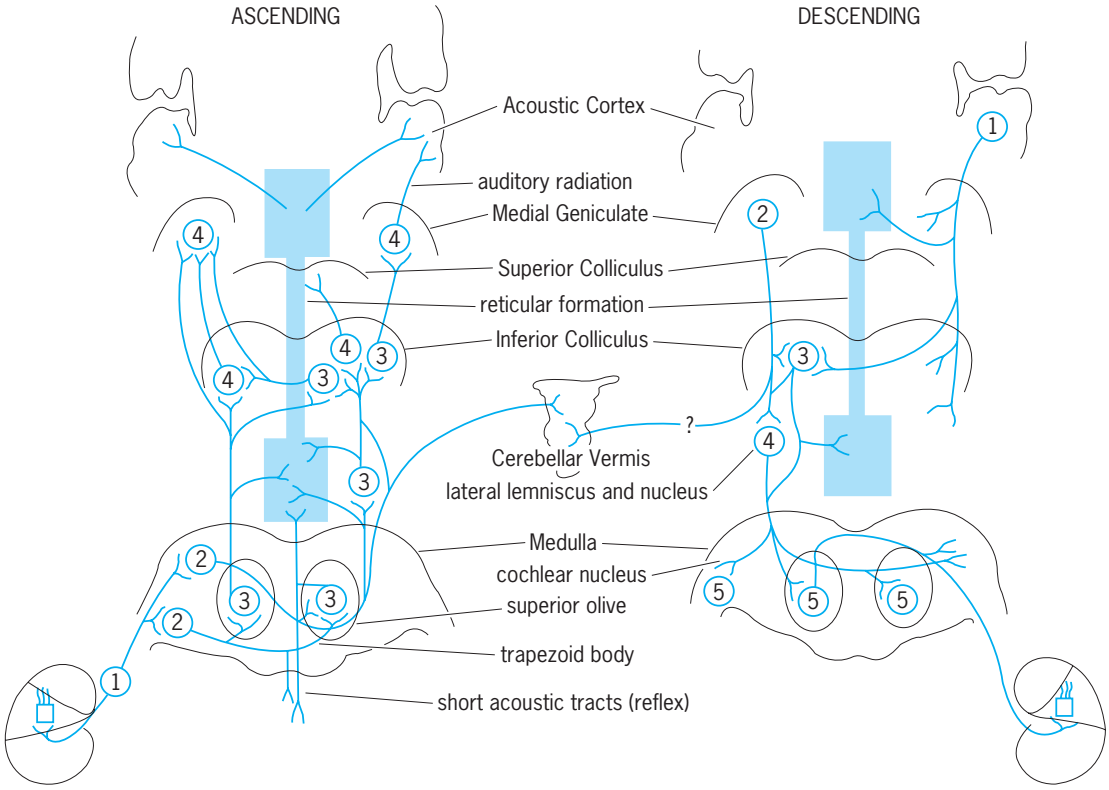

Fig. 8. Ascending (afferent) and descending (efferent) auditory nerve systems. Number labels in the ascending system indicate the number of neurons from the cochlea, while those in the descending system indicate the number of neurons from the auditory cortex. (*After R. Galambos, Some recent experiments on the neurophysiology of hearing, Ann. Oto-Rhino-Laryngol., 65:1053–1059, 1965*)

demonstrate, but it appears to improve the (physiological) signal-to-noise ratio in the presence of (external) interfering noises, to improve the resistance of the organ to noise-induced damage, and to aid frequency discrimination. *See* CONTROL SYSTEMS; GOVERNOR.

Reticular formation. The reticular formation in the brainstem is the general arousal center of the brain. Sensory inputs of any kind, in stimulating this center, will arouse all other senses, alerting them to the reception of specific inputs of their own so that the whole organism can take appropriate action. When this formation does not function, animals or humans remain in a comatose state, incapable of being aroused by any means. Figure 8, presenting the afferent and efferent systems side by side, shows their connections to the reticular formation. The study of the effects of the auditory system on the reticular formation, and vice versa, has greatly aided understanding of the general function of this formation. *See* RETICULAR FORMATION.

Overview. Understanding of peripheral auditory function is much more advanced than that of its central counterparts. This is because, in contrast to neural functions, peripheral functions can be readily modeled by mechanical and electrical devices, such as mechanical lever systems, electrical transformers, and delay lines, whose working principles are thoroughly understood. No such devices appear to exist that make large-scale use of the operating principles of the brain, for instance of parallel processing in multiple channels. Computers that are often cited as such devices and that indeed come closest in this respect cannot truly duplicate the action of the central nervous system. Existing models of neural functions are mostly phenomenological ones, and the lack of good analytical models makes their study difficult. *See* HEARING (HUMAN); NERVOUS SYSTEM (VERTEBRATE).

Juergen Tonndorf

Bibliography. W. L. Gulick, G. A. Gescheider, and R. D. Frisina, *Hearing: Physiological Acoustics, Neural Coding, and Psychoacoustics*, 1989; J. O. Pickles, *An Introduction to the Physiology of Hearing*, 2d ed., 1988; J. Tonndorf (ed.), *Physiological Acoustics*, 1981; G. von Békésy, *Experiments on Hearing*, 1961, reprint 1980.

Physiological ecology (animal)

Physiological ecology combines the study of physiological processes, the functions of living organisms and their parts, with ecological processes that connect the individual organism with population dynamics and community structure. In the past, physiological ecology focused principally on physiological processes in isolation from each other and without strong quantitative connections to ecological processes. That focus is now rapidly changing into a dynamic quantitative and more integrated discipline. Physiological ecology reaches up to population dynamics and community structure and down to genetic phenomena influencing key endocrine functions affecting behavior, energetics, survival, reproduction, and fitness in diverse environments. *See* POPULATION ECOLOGY.

Methodology and scope. Physiological ecologists focus on whole-animal function and adjustments to ever-changing environments. Short-term behavioral adjustments and longer-term physiological adjustments tend to maximize the fitness of animals, that is, their capacity to survive and reproduce successfully. Physiological ecologists need a breadth of technical competencies in physiology and ecology, in both theory and experiment, and in both laboratory and field. Quantitative and computer skills are essential, and interests that reach beyond these disciplines are needed to forge the quantitative links at higher and lower levels of biological organization.

Among the processes that physiological ecologists study are temperature regulation, energy metabolism and energetics, nutrition, respiratory gas exchange, water and osmotic balance, and responses to environmental stresses. These environmental stresses may include climate variation, nutrition, disease, and toxic exposure. **Figure 1** illustrates the interconnections between climate, microclimates, and individual properties of morphology, physiology, and temperature-dependent behavior. Climate affects animal heat and mass balances, and such changes affect body temperature regulation. Behavioral temperature regulation modifies mass and energy intake and expenditure, and the difference between intake and expenditure provides the discretionary mass for growth and reproduction. Mortality risk (survivorship) also depends on temperature-dependent behavior, which determines daily activity. Activity time constrains the time for foraging and habitat selection, which in turn influence not only mortality risk but also community composition.

Thus, growth, reproduction, and survivorship are the three variables that drive population dynamics. Population dynamics of multiple species strongly affect community attributes, such as species composition and numbers. Competitors in the community may alter temperature-dependent behavior of the species of interest, and thus their energetics and population dynamics. Community attributes may also affect the distribution of predators in space, which affects mortality risk of the species of interest. *See* METABOLISM.

Temperature regulation, energy metabolism, and energetics. Temperature regulation involves both behavioral and physiological components. The behavioral components typically involve avoiding temperature extremes. **Figure 2***a* and *b* illustrate the crucial importance of appropriate behavioral temperature regulation and show the energetic consequences to ectotherms (cold-blooded animals) and endotherms (warm-blooded animals) of deviations from optimal temperature. The optimal environmental temperature for each is the temperature that creates the maximum separation between the curves for absorbed mass (chemical energy) and expended mass (chemical energy). A temperature selected that is higher or lower than this optimum reduces the

separation, thereby reducing the potential fitness, as illustrated in Fig. 2*c* and *d*.

Endotherms have a maximum potential food absorption that is independent of environmental temperature from a physiological perspective. However, in reality food intake by endotherms is temperature-dependent. This phenomenon was described in domestic animals in the 1950s, but it has only recently begun to be quantified for animals in the wild. The broken curve that is concave downward in Fig. 2*b* represents temperature-dependent foraging behavior. The "absorbed" and "expended" curves show that the fitness of endotherms is as affected by environmental temperature as that of ectotherms, a fact that is not generally appreciated. Thus, temperature regulation, energy metabolism, and energetics are intimately interconnected for both ecto- and endotherms.

Physiological temperature regulation may take many forms. It may involve reduced or increased blood flow to appendages, depending on whether the environment is cold or hot, respectively. This method of modifying an organism's heat balance works best at intermediate and low temperatures. At high temperatures there is little heat transfer advantage to altered blood flow unless the appendage can be placed in an environment where high heat transfer rates are possible. Other mechanisms of physiological temperature regulation involve evaporative

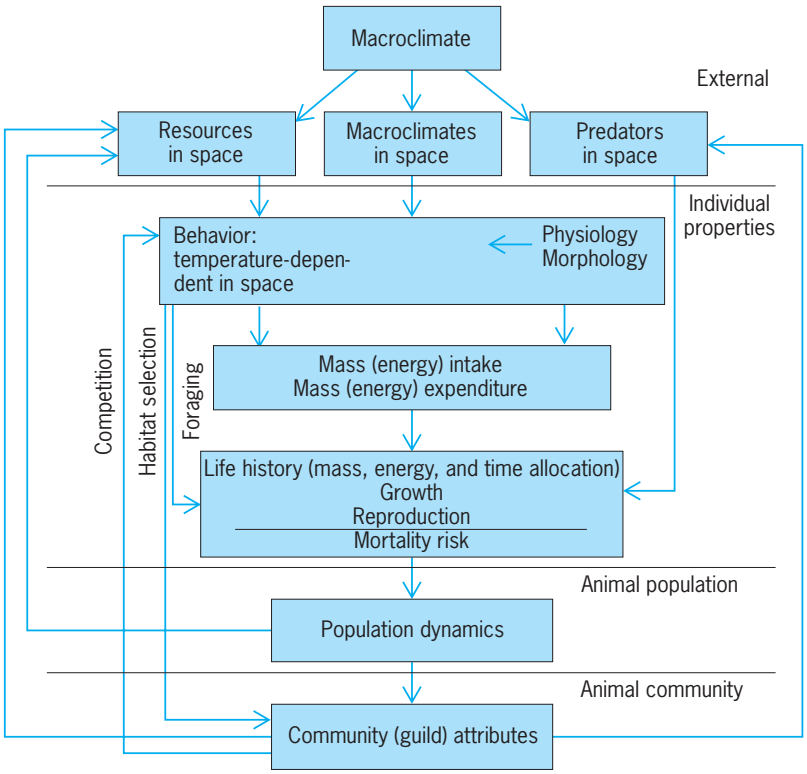

Fig. 1. Relationships between climate, individual animal properties, population dynamics, and community attributes.

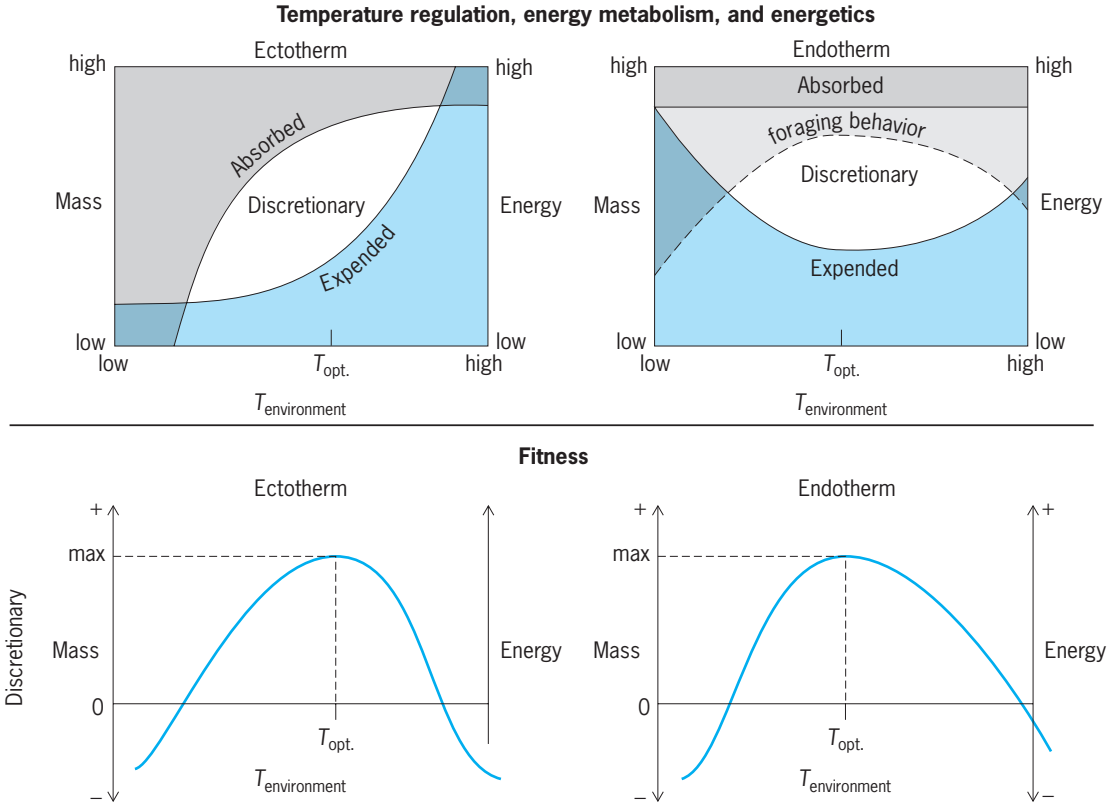

Fig. 2. Relationship between climate, intake and expenditure of mass and energy, and fitness associated with temperature regulation, energy metabolism, and energetics. For endotherms, the maximum potential for absorption is a straight line if body temperature remains constant. Behavior of the animal in cold or heat affects the amount of food ingested and therefore absorbed. The optimal temperature for achieving maximum discretionary mass and energy is different for each different body size. The temperature scale here assumes that air and radiant temperatures are equal.

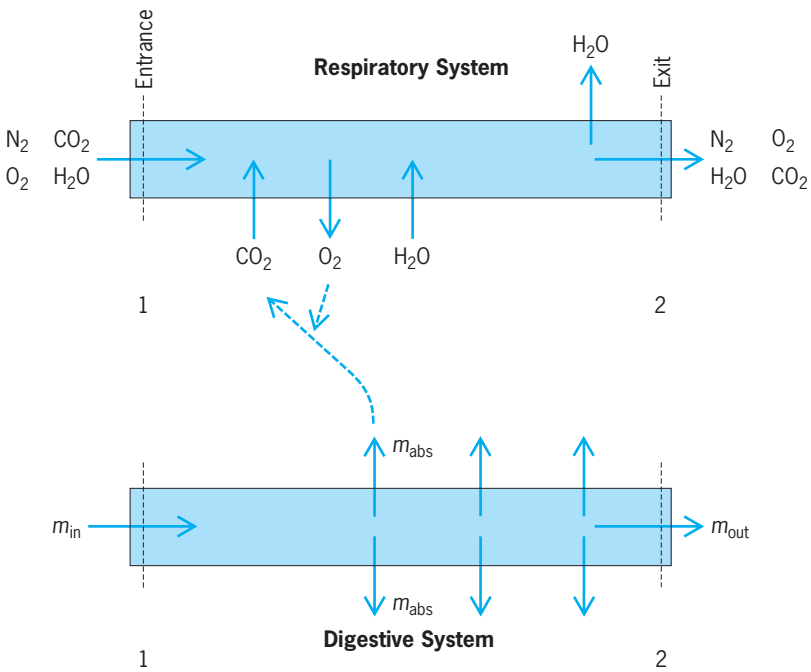

Fig. 3. System diagram models for a respiratory system and a digestive system of an organism. Mass (*m*) flows in at entrance 1 and flows out at exit 2. The type of food in the digestive system affects the ratio of carbon dioxide and oxygen exchanged in the respiratory system and therefore the mass flow and water balance associated with the respiratory system.

cooling by sweating and panting. Since sweating and panting affect water balance and respiratory gas exchange, there are strong constraints on when they may be used effectively. *See* ENERGY METABOLISM.

Respiration. Respiration is an exchange of gases between a living organism and its environment. For all but a few microorganisms, oxygen is absorbed and carbon dioxide is released to the environment. The top half of **Fig. 3** illustrates a respiratory system where gases enter at one location by diffusion, convection, or both. Gases may exit either at the same location or at another part of the body depending on the particular respiratory organ (gill, lung, or tracheae). Because water is added to air entering lungs and tracheae, respiration may affect the animal's overall water balance. Water may be recovered near the exit of these structures if the gas temperatures near the exit are close to external air temperature but below body temperature.

The ratio of oxygen absorbed to carbon dioxide released by a respiratory surface is affected by the substrates being oxidized for energy. For example, if carbohydrates are being oxidized, the moles of oxygen consumed are equal to the moles of carbon dioxide released. If protein or lipids are being oxidized, the respiratory quotient, R.Q. (moles carbon dioxide/moles oxygen), may be 0.80 or 0.71, respectively. Thus, as Fig. 3 indicates, the type of food in the digestive system can have an impact on gas transport in the respiratory system. *See* RESPIRATION.

Nutrition. The digestive system processes mass ingested by an organism within a gut or gut cavity. Very primitive organisms that lack a gut absorb nutrients across their cell surfaces, or engulf whole food particles into their cells. Quantitative models of gut function for a variety of organisms have emulated nutrient uptake very well. This is important because quantitative models of physiological processes allow us to understand much better the connections between subsystems in an organism. Good models allow us to explore how an organism's performance may vary as environmental condition changes, as an organism does in natural systems.

The amount of mass absorbed by a digestive system and the chemical energy that mass represents are of crucial importance for fitness (Fig. 2). The efficiency of food absorption by the gut in ectotherms is temperature-dependent. Thus, temperature regulation, energy metabolism, energetics, respiration, and nutrition are all interconnected to help maximize fitness. *See* NUTRITION.

Water and osmotic balance. Mass balances of water, ions, and absorbed nutrients are all involved in the maintenance of osmotic balance in living organisms. Proper osmotic balance is crucial for survival, since the enzymatic reactions of cells and communication between and within cells depends on ionic gradients. As **Fig. 4** indicates, water and ions may flow in across the outer integument, gills, or digestive system surfaces. Water and ions may exit the organism across respiratory surfaces immersed in water, from the kidneys, or from other excretory or digestive organs. Since ion transport under some circumstances may become very expensive energetically, water and ion balance are connected to the energetics of the organism and to the physical environment which modifies those energetics. *See* OSMOREGULATORY MECHANISMS.

Responses to stress. Organisms typically respond to stress in one of two ways. The first and less expensive option is to avoid stress behaviorally, if possible. If that does not work, a physiological response may be required. **Figure 5** illustrates a key behavioral response to thermal stress for ectotherms and endotherms. This particular response, change in activity time, is of central importance to linking individual energetics with population dynamics and community structure. This is because the total amount of time available for activity in a year simultaneously impacts the probability of survival from predation and the time available to gather food and water for growth, reproduction, and fat storage. Survivorship, growth, and reproduction are the key variables that

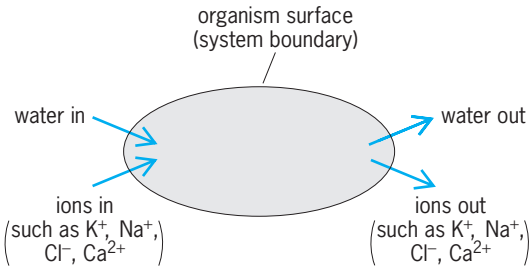

Fig. 4. System diagram for water balance and osmotic balance.

drive population dynamics and community structure. Ectotherms at high elevations have fewer hours of activity than in warmer climates. However, if the climate is too hot at lower elevations, midsummer activity may shift to early morning and late afternoon to avoid the hottest part of the day, also limiting the duration of activity. In contrast, endotherms at high elevations may be active on a year-round basis, but typically significantly restrict their activity in colder months during the daytime. At lower elevations, endotherms may be nocturnal during the hot months and diurnal during winter months. Such changes in activity patterns can have important implications for reducing costs of operation and maximizing fitness.

Changes in activity time directly affect discretionary mass and energy, and survivorship, the currencies associated with fitness. **Figure 6** illustrates how environmental temperatures can modify discretionary mass and energy because of the different times required in cold versus warm climates to first cover the maintenance costs from absorbed food. Ectotherms have lower metabolic rates at cold temperatures. Therefore the food they absorb and digest may meet maintenance costs sooner than if conditions are warm and metabolic rates higher. The opposite is true for endotherms that are typically maintaining high body temperatures. For them, warmer conditions that are not stressful mean that less time needs to be allocated to obtaining food to meet maintenance costs before they can begin to allocate absorbed food to discretionary mass and energy for growth and reproduction.

Survival probability for both ectotherms and endotherms is also influenced by temperature, as illustrated in the bottom half of Fig. 6. Ectotherms are typically inactive or minimally active under cold conditions, which likely reduces the probability of their being preyed on. In contrast, when warmer temperatures arrive and activity increases, predation may increase, resulting in lower survivorship. For endotherms the influence of temperature on survivorship could be higher or lower with changing temperature, depending upon body temperatures maintained and the types and activity levels of predators present at different temperatures. *See* BEHAVIORAL ECOLOGY; HOMEOSTASIS.

Issues of constraints and trade-offs. **Figure 7** illustrates stress responses to climate warming, inadequate nutrition, disease, and toxic stress. For ectotherms, a general warming trend may result in more time being spent at less than optimal higher temperatures for the organism. Because of reduced discretionary energy at a warmer temperature, fitness is reduced. Nutrition stress reduces the amount of available food for absorption relative to abundant food. Fighting disease may cause an animal to spend more time in warmer temperatures. This again reduces discretionary mass and energy, thereby compromising fitness. Many toxins, insecticides, herbicides, and fungicides are immunosuppressants; thus, pesticide exposure may result in a higher probability of contracting disease and increased duration of infection before recovery. Thus, climate warming, disease, and toxic exposure can move an animal to warmer temperatures and reduce discretionary mass, energy, and fitness.

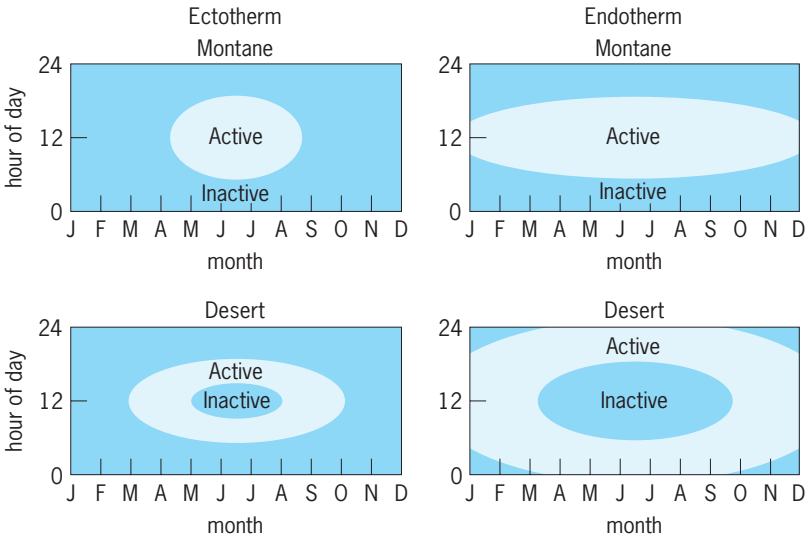

Fig. 5. Temperature-dependent changes in activity time in response to temperature stress. Effects of elevation difference (climate) are represented.

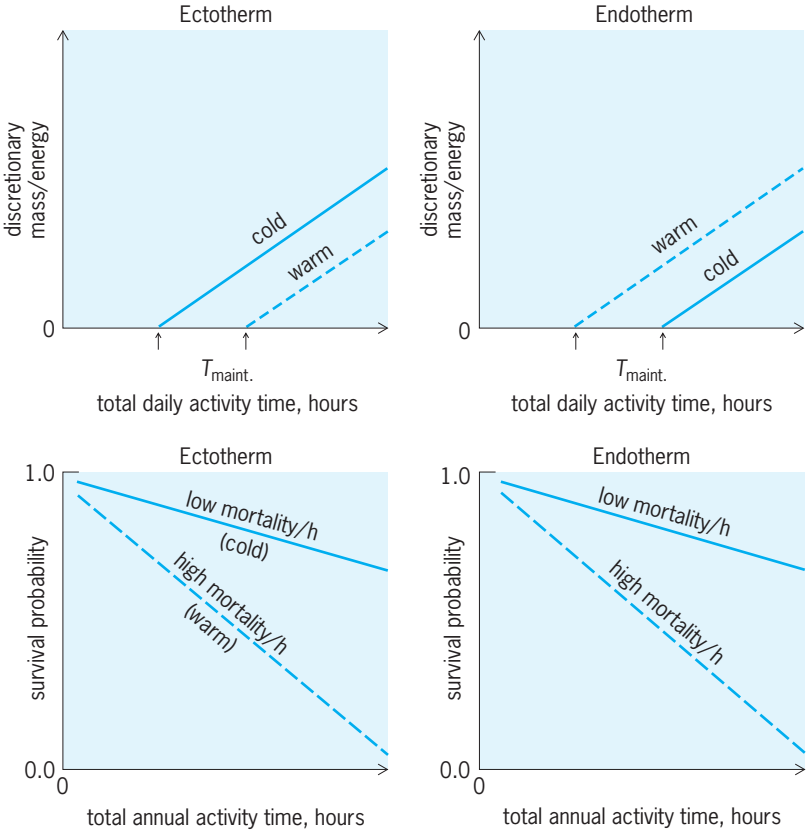

Fig. 6. Activity time connects individual energetics with population level variables of growth and reproduction potential, represented by discretionary mass/energy accumulated during a day. Length of time required to cover maintenance costs before allocating absorbed food to discretionary mass is illustrated for ectotherms and endotherms in different climates. Activity time also affects survival probability, which changes with climate.

Endotherms are similarly constrained in their discretionary mass and energy by reduction in nutrition, which decreases absorbed food, and by disease and toxins, which may elevate the costs to maintain a higher body temperature (fever). Sickness also induces reduced activity through suppression of neurological activity. Reduced activity can lead to reduced food intake, which could further constrain discretionary mass and energy. Thus endotherms, like ectotherms, respond in similar ways to climate warming, nutrition, disease, and toxic exposure. Figure 7 illustrates that if these organisms were to try to retain their optimal environmental temperature for maximum discretionary mass and energy, they could compromise their ability to fight disease or toxic exposure. Alternatively, they might spend an inordinate amount of time or resources trying to counter higher average temperatures by evaporation of water (sweating or panting), which could impact water and ion balance and other aspects of organism function.

Warren Porter

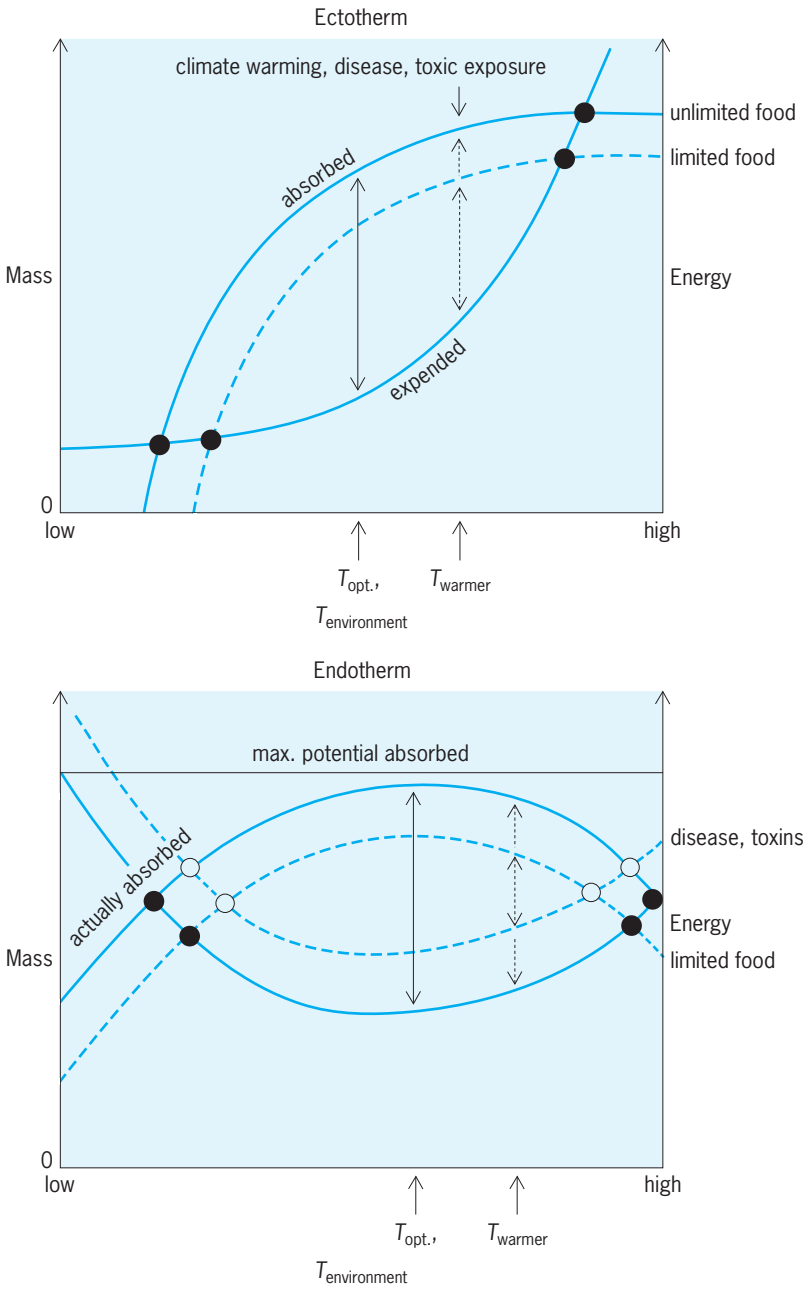

Fig. 7. Trade-offs in responding to environmental stresses of climate, disease, nutrition, and toxic stress. All stresses either affect the temperature that an animal selects in the environment or close the distance between the curves defining absorbed and expended mass. Climate warming which forces an animal to spend more time in a warmer environment decreases discretionary mass because of lower mass intake and higher mass expenditure. Nutritional stress narrows the allowable range of environmental temperatures that represent a positive mass balance for the animal. Disease and toxic exposure also narrow the range of environmental temperatures that represent a positive mass balance; this happens because of decreases in mass intake and increases in mass expenditure.

Physiological ecology (plant)

The branch of plant science that seeks physiological (mechanistic) explanations for ecological observations. Emphasis is placed on understanding how plants cope with environmental variation at the physiological level, and on the influence of resource limitations on growth, metabolism, and reproduction of individuals within and among plant populations, along environmental gradients, and across different communities and ecosystems. The responses of plants to natural, controlled, or manipulated conditions above and below ground provide a basis for understanding how the features of plants enable their survival, persistence, and spread. Information gathered is often used to identify the physiological and morphological features of a plant that permit adaptation to different sets of environmental conditions. The principal life forms studied are herbs, forbs, grasses, shrubs, trees, and nonvascular plants (such as mosses and lichens). Key life-cycle processes include the acquisition, transport, and loss of mass (CO_2 and H_2O), energy (heat and light), and mineral nutrients; the exchange of gases which influence photosynthesis (CO_2 and H_2O), transpiration (H_2O), and respiration (CO_2 and O_2); and the allocation of fixed energy in the form of carbohydrates (CHO) necessary for life-sustaining activities. Photosynthesis and respiration, in particular, involve a suite of complex physical and chemical reactions within the organelles, cells, tissues, and organs of a plant. Reaction rates are influenced by an array of environmental characteristics such as temperature, and water, nutrient, and light availability.

The environments that plants occupy are often subject to variation or change. The ecophysiological characteristics of these plants must be able to accommodate this or the plants face extinction. Given the right conditions, ample time, and genetic variation among a group of interbreeding individuals, plant populations and species can evolve to accommodate marked ecological change or habitat heterogeneity. If evolutionary changes in physiology or morphology occur on a local (for example, on an isolated mountain top) or regional scale, populations within a single species may diverge in their characteristics. Separate ecological races (ecotypes) arise in response to an identifiable, and often narrow, set of environmental conditions. Ecotypes are genetically distinct and are particularly well suited to the local or

regional environment they occupy (locally adapted). Such ecotypes can often increase the geographical range and amplitude of environmental conditions that the species occupies or tolerates. Ecotypes may also occur as a series of populations arrayed over a well-defined environmental gradient called an ecocline. In contrast, if ecotypes are not present, some plant species may still be able to accommodate a wide range of growth conditions through morphological and physiological adjustments, by acclimation to a single factor (such as light) or acclimatization to a complex suite of factors which define the entire habitat. Acclimatization can occur when individuals from several different regions or populations are grown in a common location and adjust, physiologically or morphologically, to this location. Acclimation and acclimatization can therefore be defined as the ability of a single genotype (individual) to express multiple phenotypes (outward appearances) in response to variable growing conditions. Neither requires underlying genetic changes, though some genetic change might occur which could mean that the response seen may itself evolve. Acclimation and acclimatization may also be called phenotypic plasticity. *See* PLANT EVOLUTION.

Studies of metabolic rates in relation to environmental conditions within populations, ecotypes, or species provide a way to measure the tolerance limits expressed at different scales. These data in turn help identify the scales at which different adaptations are expressed, and enhance an understanding of the evolution of physiological processes. Combining observations and measurements from the field with those obtained in laboratory and controlled environment experiments can help identify which conditions may be most influential on plant processes and therefore what may have shaped the physiological responses seen. Laboratory and controlled environment (common garden) experiments also assist in helping identify how much of the variation expressed in a particular metabolic process can be assigned to a particular environmental factor and how much to the plants themselves and the genetic and developmental plasticity they possess.

Methods and measurements. Accurate field measurements are difficult to obtain primarily because of the logistics involved in getting to suitable field locations. The vagaries of uncontrolled environments and their marked influence on rates of metabolism, and the limitations of using the appropriate equipment under unfavorable conditions contribute to the difficulty. Field measurements are, however, essential for providing information about how plants behave under natural conditions. The results obtained from field measurements provide the realistic context against which laboratory and controlled environments experiments can be designed and then compared. One of the most productive and instructive approaches for understanding the ecophysiology of plants combines both field and laboratory observations and experiments.

Solid-state electronics have led to computerized control systems and the miniaturization of key components, so that field instruments and their measurement capabilities have improved dramatically. Portable data acquisition systems which rapidly monitor and record both plant and environmental conditions have also increased the scope of the research and objectives and have extended understanding of how plants function under natural conditions. Satellite imagery and other remotely sensed geographical and botanical information have permitted ways of extending the interpretation of ecophysiological information gathered on individual plants or in well-defined locations. The findings have been extended to the community, ecosystem, and landscape scales. These field techniques have permitted more robust sampling of plants in the field, have added accuracy to the measurements obtained, and have significantly enhanced understanding of plant responses to the environment.

Obtaining information about the precise physiological mechanisms which best explain how plants respond to environmental conditions must often be done by collecting plants from natural environments and then growing them under precisely controlled conditions in the laboratory, growth chamber, or greenhouse. In this way, variables of interest (such as temperature, light, and soils) can be changed and monitored in a well-defined manner while holding all other conditions constant, and then plant responses to those variables can be measured. In addition, particular factors are manipulated, and plant responses in a large number of individuals or in a short span of time can be measured. This permits one to assign and partition the observed variation in a particular physiological behavior into its genetic, developmental, and environmental components. It permits understanding of how plants respond to conditions they may have experienced in the past or might experience in the future, allowing a broader interpretation of observed ecophysiological responses across a wide range of plant species or types. Laboratory analyses may also be essential for understanding how natural environments may have influenced plant function or growth. For example, the analysis of mineral nutrient concentrations of field-collected soil and plant materials may be essential for understanding patterns of plant nutrient acquisition under field conditions. The analysis of tissue characteristics such as the water transport system (xylem), root or leaf characteristics, or storage compounds, performed in the laboratory, are essential for providing detailed information about the ways in which plant form and function permit adaptation to natural conditions. New microscopic, isotopic, molecular, and biochemical methods have significantly expanded knowledge about the underlying basis of the variation seen in the ecophysiological characteristics and responses of plants.

The broad categories of processes studied in the field and in the laboratory under controlled conditions, which define the ecophysiology of a plant, are germination, growth, energy, water, carbon relations, mineral nutrition, and reproduction. These processes are often studied under both optimal and

stressful conditions or across the range of conditions that plants experience in the field. Many of the processes studied are clearly influenced by the environmental conditions, but they are also influenced by the growth form, morphology, and life history characteristics of the plant (such as herb, grass, tree, or shrub; woody or herbaceous; evergreen or deciduous; annual or perennial) as well as the shapes and characteristics of the roots, stems, and leaves. *See* PLANTS, LIFE FORMS OF.

Germination. The germination characteristics of plants are properties of the seed, of the seedling, and the factors which influence the transition between the seed and seedling stages. Seeds can possess a range of dormancy mechanisms which inhibit germination, and there are a number of environmental cues which serve to break the dormancy and permit or stimulate the embryonic plant to develop into a seedling and eventually a mature plant. Light, heat (fire), water, smoke, and other factors can influence germination in some species but not in others. Seed size and age (maturity) can also play a role in when and under what conditions seeds germinate. Seedling survival is highly variable and determined by species-specific properties (for example, the tolerance a species has for site conditions) and the variation and severity of growing conditions at the time when seedlings are establishing. Often it is alternating conditions (rather than constant conditions), such as temperature or moisture, which stimulate seeds to germinate, and result in higher germination percentages within a cohort of seeds. Assessing germination success is often estimated by marking seeds in the field or sowing a known number of seeds into plots or growth containers and following their fate over time. Once germination has occurred, there may be physiological and morphological characteristics unique to the seedling or juvenile stage which determine survival and growth rate into the later vegetative stages of development (such as the sapling phase in trees).

Growth. Plant growth in relation to the environment is quantified as changes in dimension (height, length, width, depth), in mass (fresh or dry), or the production of biosynthetic substances such as carbohydrates. Assessments of aboveground plant growth can be made by measuring changes in stem diameters (such as with tree bands which record changes in the diameter of the main trunk) and in the mass, shape, or dimensions of plant canopies (leaves and stems) with hemispherical photographs taken from the ground or digitally recorded images of the canopy from which the leaf area can be estimated. Whole plants or plant parts are harvested or marked and subsequently remeasured at two or more points in time to obtain a measure of growth or growth rate. Growth of belowground parts (rhizomes and roots) can be determined in the field with excavations, with subsamples of soil cores into which roots have grown, or with wedge-shaped Plexiglas boxes or clear tubes embedded in the soil. The boxes can be removed from the soil at different time intervals, and root growth measured along the surface between the box and the soil. Clear boxes can also be used, and root growth determined by tracing roots with colored pens on the box surfaces. The clear tubes can accommodate a small, mobile video camera, which records root presence along set positions of the tube. Large pits can also be excavated, and glass walls installed below the soil surface against which plants are either planted or allowed to grow (called a rhizotron). This permits one to observe the entire soil profile containing roots of large plants and to follow changes in the form and rates of root growth over time. *See* PLANT GROWTH.

Energy relations. The balance between incoming and outgoing (loss) solar radiation determines the temperature of a plant leaf or canopy. The energy relations of a plant therefore require the quantification of both plant features and habitat characteristics, which influence this balance. Radiation absorption, reflection, and transmission of both visible light and infrared light (heat) by the surfaces of a plant and the environment in which it grows from the sun, sky, and ground must be measured, as well as the loss of heat gain by convection, metabolism, and storage, and the loss of heat by evaporation and plant transpiration. Quantification of radiation input and loss is measured as both light and heat. Light is measured with a range of sensors designed to specifically quantify the light used in autotrophic metabolism (photosynthesis) and the light which influences the thermal relations of a plant. Leaf and air temperatures can be measured with noncontact infrared thermometers (sensors) or with fine-wire thermocouples or thermistors, which are in direct contact with surfaces of interest. The amount of energy absorbed by, say, a leaf surface is determined by a balance between energy inputs, features of the leaf (surface texture, size, and color which influence the size and behavior of the boundary layer around the leaf), wind speed and the turbulent heat transfer away from the leaf, the thermal stability of the air surrounding the leaf, and the processes of connective and latent (evaporation and transpiration) heat losses. The temperature of the leaf will affect metabolic processes such as photosynthesis and respiration and leaf-level behaviors such as stomatal opening and closing.

Water relations. The water relations of a plant are determined by its ability to gain access to and absorb water, the status of the plant tissues once water is obtained, and the factors which influence rates of water loss from plant organs such as leaves, stems, and roots. Water absorption and use can be quantified by measuring the changes in plant mass over time by repeatedly weighing potted plants or large blocks of containered soil with plants growing in them (called weighing lysimeters). Water uptake has also been determined by adding a nontoxic chemical (bromide or dye) or isotopic tracers to the water that roots have access to and looking for its appearance in the plant over time and under different conditions. The gradient in the chemical (or water) potential between the soil and the plant root determines if and how quickly water absorption occurs. Water

potential is the difference in the free energy content (or chemical potential) between pure water and the water contained inside cells, tissues, or organs. The water potential of pure water is set at a value of zero. Water in the conducting tissues (xylem) of plants or in soils (which contain dissolved solutes, making them solutions) is at a lower water potential than pure water and thus has water potential values that are expressed as negative numbers. Water in live cells or tissues such as leaves or roots also contains dissolved solutes, so these solutions are contained and compartmentalized. This provides a restrictive barrier, and positive pressures build up so that water potentials can have positive values at times called turgor pressure (which prevents leaves from wilting). In actively transpiring plants, water is lost to the atmosphere through minute pores embedded in the leaves called stomata. Stomata can open and close and therefore regulate rates of water loss or transpiration. When stomata are open, water leaving the leaf is replaced at the surface from which it evaporates from the surrounding tissues. As transpiration continues, it creates a gradient toward the leaves from the soil through the roots and stems and into the leaves. Water is pulled up the plant along this gradient, which becomes progressively more negative, from the soil to the atmosphere. Along the gradient there can be resistance to the flow of water which can put the water column under greater tension. If these tensions become too great for the forces which allow water to adhere to itself, bubbles form (because air which forms the bubbles comes out of solution under tension) and the water column can break (become embolized). This happens in roots and stems. Embolized stems or roots can be refilled with water, but if they are not, tissue dysfunction or plant death may result. Instruments called pressure chambers, psychrometers, or hygrometers can measure water potential (the water status) of plants and plant tissues. Other techniques can be used to determine the levels of embolism.

Rates of water loss from leaves can also be quantified with an instrument called a porometer that encloses a leaf in a chamber (cuvette) and determines the transpiration rates by the difference between incoming and outgoing humidity, and the leaf area and flow rates of air through the chamber. Rates of water loss from whole plant canopies can be estimated from measurements of water-flow velocities in the xylem tissues (termed sapflow methods) or from the precise quantification of water vapor loss from entire canopies or stands of plants with instruments placed on a tower that remotely sense changes in the energy status or actual water vapor loss. Water loss can occur from stems, reproductive parts, and roots. Water loss from roots is generally determined either indirectly from changes in the water content of the soil surrounding the roots or by tracing water with added stable or radioactive isotopes from its point of absorption to where it is lost. The amount of water lost from plants is greatest from leaves, though loss of water from roots has been shown to be significant. *See* PLANT-WATER RELATIONS.

Mineral nutrition. All plants require nutrients. These nutrients can come in organic and inorganic (mineral) forms and, depending upon the element, can be required in either large (macro) or very small (micro) quantities. Macronutrients include nitrogen, phosphorus, potassium, sulfur, calcium, magnesium, and iron. Micronutrients include the trace elements manganese, zinc, copper, molybdenum, boron, and cerium. Specific plant groups (such as legumes) may require specific elements (such as phosphorus) in greater amounts because of their importance in a specific metabolic process, such as biological nitrogen fixation; while other plants which inhabit particular substrate types may use or require elements not listed above. Nitrogen is often considered one of the most important elements in limiting plant performance and growth because it is required for so many different physiological processes (including photosynthesis, protein synthesis, and other enzyme-mediated processes). The nutrients required by plants become available by a variety of means, such as geological weathering; soil microbial processes; and biological nitrogen fixation, absorption, and transport through fungal symbionts (called mycorrhizal associations). Changes in soil pH affect the availability of nutrients for plants, as do inputs from anthropogenic sources from fossil fuel combustion (such as nitrogen and sulfur), mining activities (such as heavy metals like copper, magnesium, and iron), or natural disturbances like fire or floods.

Plants and vegetation play a critical role in the cycling of nutrients through the biosphere. The speed and magnitude of ecosystem element cycling is often determined by the interplay between the soils, the climate, the vegetation (form and species), and the microbial community which acts to decompose organic materials in soils and make them available for plant use. Measuring plant nutrient relations in the field is difficult and often requires experimental manipulation such as adding specific fertilizers or stable isotope tracers [for example, ammonium ($^{15}NH_4^+$) or nitrate ($^{15}NO_3^-$)] and following their uptake patterns and fates in the soils, roots, stems, leaves, and reproductive structures (flowers and fruits). Controlled environment experiments are often required to isolate and precisely track plant responses to nutrient additions or deficits and to follow the fate of specific nutrients into the plant and into specific cells, tissues, organs, or constituents. *See* BIOGEOCHEMISTRY; PLANT MINERAL NUTRITION; SYSTEMS ECOLOGY.

Carbon relations. Understanding the carbon relations of plants requires quantification of the processes which influence carbon gain, carbon loss, and carbon allocation to the activities which require the biochemical energy contained in the carbohydrate compounds made in plant photosynthesis. Precise measurements of plant carbon metabolism (photosynthesis and respiration) and the environmental conditions which influence it must be made on field- and laboratory-grown plants. To date, three main photosynthetic pathways have been described, and these show some relationship to the environmental

conditions where they are most common. The majority of plants possess the reductive pentose phosphate, or the C_3 pathway (also called the Calvin-Benson cycle; the 3-carbon molecules are the primary product of the biosynthetic reactions). A much smaller set of plants possess the C_4 pathway (or Hatch-Slack cycle), which is biochemically and anatomically more complex. The specialized anatomy (Kranz anatomy) of C_4 plants increases the efficiency of carbon dioxide (CO_2) fixation by virtue of the fact that the initial 4-carbon product (acid) produced in the mesophyll cells is physically transported and sequestered in a ring of cells called the bundle sheath cells. Here it is decarboxylated back to CO_2 and then refixed eventually into carbohydrates in the absence of photorespiration (which competes for the photosynthetic enzymes in C_3 plants). The increased efficiency of CO_2 metabolism often results in very high rates of productivity in C_4 plants; maize (corn) and sugarcane are common crop plant species that possess the C_4 pathway. The third, and least common, photosynthetic pathway, which can often be found in succulent or epiphytic plants, is called crassulacean acid metabolism (CAM; it was first discovered in the succulent plants in the Crassulaceae plant family). Unlike C_3 and C_4 plants which take up CO_2 from the atmosphere through their stomata during the day, CAM plants take up CO_2 at night and close their stomata during the day when temperatures and water loss from plant transpiration are highest. Carbon dioxide taken up at night by CAM plants is fixed into a 4-carbon acid (malic acid); in the daylight, when stomates are closed, the malic acid exits the plant vacuole where it is stored, and enters the chloroplasts where it is decarboxylated and refixed into carbohydrates in a way similar to C_4 plants. Regardless of which of the photosynthetic pathways a plant possesses, all require visible light which drives the light (photo) reactions on the membranes of green plant chloroplasts. The products of these reactions influence the synthesis (light-independent) reactions that convert CO_2 and the carbon compound ribulose bisphosphate into the carbon compounds which eventually are assembled into diverse carbohydrates such as sucrose, glucose, and cellulose (via the C_3 pathway).

At the physiological level, it is known that C_3, C_4, and CAM plants all possess dark respiration (in the mitochondria). In addition, photorespiration occurs, but at rates two to four times higher than dark respiration in C_3 plants. Photorespiration in C_4 plants occurs only in the bundle sheath cells in the presence of high CO_2 concentrations and is therefore very low. Rates of photorespiration in C_3 plants can be very high and can compete for the enzyme ribulose bisphosphate carboxylase-oxygenase (RuBisco), causing overall reductions in the amounts of carbon fixed compared with C_4 plants growing under similar conditions. High O_2 concentrations also cause a marked reduction in photosynthetic rates in C_3 plants because rates of photorespiration are enhanced; this does not occur in C_4 plants. The biochemical differences in plant photosynthetic pathways as well as in the spatial and temporal patterns by which CO_2 is eventually fixed into carbohydrates lead to marked differences in the photosynthetic efficiencies among photosynthetic types. The spatial separation between the mesophyll and bundle sheath in C_4 species and the temporal separation of CO_2 uptake (in the night) and fixation (in the day) in CAM plants commonly leads to greater photosynthetic efficiency compared to C_3 plants when expressed on the carbon income–to–water loss basis.

Differences in the efficiency of photosynthesis (carbon gain per water lost) can be detected by measurements of the actual rates of gas exchange at the leaf level or indirectly by analyzing the ratio of heavy (^{13}C) to light (^{12}C) carbon isotopes that compose the leaf, relative to a known standard; this is termed the carbon isotope ratio ($\delta^{13}C$) and is expressed in units of parts per thousand (‰). The CO_2 assimilated by plants comes in two stable forms, $^{13}CO_2$ and $^{12}CO_2$, and one radioactive form, $^{14}CO_2$. During C_3 photosynthesis, the two stable forms of CO_2 are not taken up equally. Due to differences in the rates of diffusion into the leaf through the stomata and isotopic discrimination against $^{13}CO_2$ by the principal carboxylating enzyme, RuBisco, C_3 plants end up assimilating more $^{12}CO_2$ relative to $^{13}CO_2$ than either C_4 or CAM plants. This means that when the $\delta^{13}C$ of plants with different photosynthetic pathways is measured, C_4 and CAM plants are isotopically heavier (contain more ^{13}C and therefore have higher $\delta^{13}C$ values) than C_3 plants. Because C_3 plants discriminate more than either C_4 or CAM plants, their $\delta^{13}C$ values generally range between -21 and -34‰, while C_4 plants or CAM plants have $\delta^{13}C$ values ranging between -10 and -14‰. CAM plants can have a larger range of $\delta^{13}C$ values than C_4 species depending upon whether they are obligate CAM (-10 to -14‰) or facultative CAM (-15 to -22‰). Negative numbers indicate less ^{13}C present in the plants relative to the source air that they have taken in. Variation in the $\delta^{13}C$ values of plants has been used not only to identify the different photosynthetic pathway types but also to examine photosynthetic efficiency within C_3 species; to trace the diets of animals which feed upon different plants (the $\delta^{13}C$ value of the plant material is transmitted, largely unchanged, to the carbon-containing compounds of the animals that ate it); and to reconstruct changes in the dominant plant species within vast regions (such as grasslands) over long periods of evolutionary time through the analysis of fossil carbon.

The type of photosynthetic pathway that a plant possesses can have a marked influence on its ecology and the environments that it can inhabit. C_3 plants are the most widely distributed on a global basis and can be found in every ecological habitat where plants grow. C_4 plants are commonly found in warmer climatic zones (for example, tropical and desert ecosystems) which can also experience periodic or chronic drought. Some C_4 species do occur in regions with cool or cold winters but hot and often very dry summers; they are not found in the Arctic and alpine regions of the world with short growing

seasons and long cold winters. There are also a number of C_4 plants which inhabit aquatic environments where access to CO_2 is restricted. The Earth's most primitive plants possess the C_3 pathway and have appeared most recently in the history of plants on Earth, and C_4 photosynthesis has evolved independently in many plant families that are only distantly related. A few genera of plants contain species with both C_3 and C_4 pathways (such as *Atriplex*) and also C_3–C_4 intermediates (such as *Flaveria*). CAM plants often occur in extremely arid regions or habitats such as deserts where water is often a limiting resource. Many CAM plants are succulents or belong to plant families comprising mostly succulent species; for example, most cacti are CAM plants, as are many members of the pineapple family. What led to the evolution of the C_4 and CAM photosynthetic pathways is still under investigation, but data support the notion that warm temperatures, marked periods of aridity, and global changes in the ambient carbon dioxide may have played a role, particularly in C_4 taxa. The optimum temperature for net photosynthesis in C_3 plants lies between 10 and 25°C (50 and 77°F), while for C_4 and CAM plants it is higher, generally between 30 and 40°C (86 and 105°F).

Several different methods measure rates of photosynthesis and respiration in the field and under controlled laboratory, growth chamber, or greenhouse conditions. They rely on determining the changes in either the CO_2 concentrations passing into and out of a chamber enclosing a leaf with an infrared gas analysis system (based in the laboratory or field portable), or the O_2 which is evolved from leaf material in the presence of nonlimiting CO_2 and light with a high-sensitivity oxygen electrode (in the laboratory only). The infrared gas analysis system allows measurements of both photosynthesis (in the light) and dark respiration (without light) under different conditions in the field and in the laboratory. Such systems allow different environmental variables to be held constant so that precise determinations of the factors which are known to influence rates of photosynthesis can be assessed one at a time. Infrared gas analysis systems are self-contained, computerized, and often automated for controlling conditions and collecting the data. An additional method for determining gross rates of carbon fixation uses radioactively labeled $^{14}CO_2$. Leaves are enclosed in a chamber which contains $^{14}CO_2$, and are allowed to take up the CO_2 for a short period of time. The exposed leaf sample is collected and chemically fixed or frozen and returned to the laboratory, where the determination of how much $^{14}CO_2$ was incorporated into the leaf is done using a liquid scintillation counter. The incorporation of labeled $^{14}CO_2$ into carbon-containing compounds can also be a useful method of determining how and where plants allocate carbon once it is fixed in the process of photosynthesis. Any of these methods can be used to investigate the carbon relation of plants under different environmental conditions, under changing conditions, or in a comparative framework using different species or genotypes (or ecotypes). *See* PHOTOSYNTHESIS; PLANT METABOLISM; PLANT RESPIRATION.

Reproduction. The processes of flowering, fruit production, and seed output are part of plant reproduction. These processes are studied in the field or laboratory by measuring how they change in relation to site or environmental conditions; to an array of physiological variables; and to annual, habitat, and genetic variation within and among plant populations. To understand what influences successful reproduction, these processes must be studied over time and in relation to specific factors such as light, temperature, photoperiod, plant water status, and plant carbon balance, in order to determine which factors are most important in governing flower, fruit, and seed production. This information can be further correlated to the hormonal status of the plants or its pollination ecology to arrive at a picture of what ultimately determines plant persistence or spread over time. Understanding the link between the physiological characteristics of plants, the environmental factors which shape these characteristics, and their influence on plant reproduction is critical for understanding what ecophysiological suites of traits may be favored and thus evolve in plant populations. *See* FLOWER; FRUIT.

Ecophysiology and biophysics. Some of the most promising areas for future research will merge plant physiological ecology with other areas of science. The relationship between plant physiological ecology and biophysics is seen on two scales; the subcellular scale, below the level of whole plants; and the global scale, where whole-plant behavior is related to the impinging environmental conditions. At the subcellular scale, there are a growing number of studies which have used biophysical information about the kinetics of important processes at the level of single electrons to their outcome at the whole-organ (such as leaf) or whole-plant level. At the environmental scale, organismal performance is related to the plant response to the physical environment; and the exchange and feedback of mass and energy between the plant biota and the environment. How temperature, gas exchange, water, and transport processes in the physical world affect or are affected by plants becomes critical to understanding atmosphere-biosphere interactions. This information is used in the formulation of global models of plant production and in predicting how changes in natural resources shape the strength and nature of atmosphere-biosphere interactions. *See* BIOPHYSICS; PLANT PHYSIOLOGY.

Agriculture and conservation. As contemporary agricultural practices begin to incorporate the notion of ecological sustainability, it is increasingly important to understand how the natural vegetation, as a model system, sustains itself in the face of environmental heterogeneity, disturbance, and change. Agriculture is moving away from the intensive cultivation of annual crop plants and toward the use of perennial crops and trees that are sustainable and more subject to natural variation in the environments where they grow, and therefore an understanding of

the physiological and morphological features which permit sustained growth will be necessary. This information will come by borrowing from the knowledge provided by plant physiological ecology. The outcome of more sustainable agricultural practices will be seen in the conservation of the Earth's soils and diversity of species and in the improvement of global water resources. *See* AGRICULTURE.

Ecosystem science. There is an increasing awareness that to fully understand the ecological behavior of large regions on Earth (such as ecosystems), an understanding of the behaviors of the organisms which compose those ecosystems could be critical. What determines the movement of energy, water, and nutrients into, through, and out of a grassland or forest, for example, can best be understood by knowing how the plants which fix, use, and recycle this energy and mass behave. Different plant species possess different chemical, morphological, and physiological features, which determine the speed, magnitude, and efficiency of mass and energy transfer between the atmosphere and the biosphere. In this way, they can influence the processes which occur at the ecosystem scale. The rates of nutrient cycling through an ecosystem, for example, can be dictated by the kinds of plants present; the requirements those plants have for nutrients; and if different species retain, return, or fix particular nutrients, such as nitrogen, within their areas of growth. As such, merging an understanding of the ecophysiological features and behaviors of plants with an investigation of what determines broad-scale ecosystem patterns and processes is yielding new insights.

An acknowledgement of the close relationship between plant physiological ecology and ecosystem science is becoming important in helping scientists to understand not only what may be dictating ecosystem function but how ecosystems will respond to large-scale, human-induced environmental changes now and in the future. A considerable fraction of global change research is focusing on the interface between plant physiological ecology, environmental biophysics, ecosystem science and biogeochemistry, and atmospheric science. Of particular importance is ongoing research focused on understanding the relationship between the Earth's biodiversity, the adaptations and features that the diverse organisms which inhabit the Earth possess, and the effects that these organisms have on ecosystem-level functions as environments continue to change. Loss of biodiversity caused by human-induced habitat destruction (from deforestation, desertification, mining, and pollution) or atmospheric change (from fossil fuel combustion and biomass burning) will mean the loss of species that play key roles in the function, integrity, and stability of ecosystems. For example, the loss of vast areas of tropical trees to timber harvesting and mining practices has changed the climate of the Amazonian basin and decreased the resilience of this ecosystem to additional impacts, either natural or human-induced. Likewise, increasing concentrations of carbon dioxide, soil, and atmospheric pollutants are changing the global carbon, nutrient, and water cycles because of changes in plant species composition or the response of different plant species to these unnatural changes.

The ability to scale up plant responses to the sorts of changes mentioned above to the ecosystem and global levels will therefore be critical to understanding and then mitigating the impacts imposed upon the biosphere in the future. This will require large-scale experiments under both natural and manipulated conditions in the field and under well-controlled conditions; long-term monitoring of plant, soil, water, and atmospheric conditions; and the extension of this information to predict future responses with the use of mathematical models of plant- and ecosystem-level performance in response to environmental change. *See* ECOLOGY; ECOSYSTEM.

Ecophysiology and evolutionary biology. Traditionally, plant physiological ecologists have studied the adaptive features of plants and the environmental characteristics that they surmise have shaped these features. However, these early studies were done with little or no knowledge of the evolutionary origins of the traits under study. Also, because "phsiology" is not preserved in the fossil records of plants, elucidating patterns of evolutionary change in physiology is problematic. Today, we can estimate, from indirect analyses of contemporary functional traits, the ancestral physiological states and possible environmental conditions that have led to them using the comparative (for example, features of plants in a group of related species that are known to share a common origin). Understanding the degree of the variation in plant form and function and knowing the evolutionary (phylogenetic) relationships among the taxa being compared is providing new insights into the causes underlying the origins of functional adaptations. Contemporary phylogenetic (evolutionary) methods applied to the study of the plant ecophysiological characteristics are permitting more robust analyses of the nature of character changes that have occurred over evolutionary time. This approach is also providing a way to determine if characters are products of history *or* newly evolved adaptations which have arisen in response to novel environmental challenges. Such an approach is also permitting a deeper understanding of what may have allowed speciation and diversification in some groups of plants. For example, the traits which may have permitted plants to cope with stressful and variable environments when mapped onto the evolutionary tree allows us to evaluate their current and/or past utility. From this, we might also better understand the origin of physiological adaptations, the nature of ecophysiological changes, and the probable trajectory of character and trait evolution. T. E. Dawson

Bibliography. F. A. Bazzaz, *Plants in Changing Environments*, 1996; G. S. Campbell and J. M. Norman, *An Introduction to Environmental Biophysics*, 2d ed., 1998; T. T. Koslowski and S. G. Pallardy, *Physiology of Woody Plants*, 1997; P. J. Kramer and J. S. Boyer, *Water Relations of Plants and Soils*, 1995; W. Larcher, *Physiological Plant Ecology*, 3d ed., 1995;

P. S. Nobel, *Physicochemical and Environmental Plant Physiology*, 1991; W. H. Schlesinger, *Biogeochemistry: An Analysis of Global Change*, 2d ed., 1995.

Phytamastigophorea

A class of the subphylum Sarcomastigophora, also known as the Phytomastigina. These are the plant flagellates which contain chlorophyll and other pigments, but colorless forms are also included. Grass green is the usually observed color, primarily because the green flagellates are the largest. Those containing an excess of yellow pigments generally are smaller, and fewer species have unusual colors, such as blue or red. Holophytic, saprophytic, and holozoic modes of nutrition occur, and specific chemical components may be demanded by individual species within the group.

Encystment is frequent among phytoflagellates, cyst composition being one method of determining relationships for some colorless species. Reproduction may occur within the cyst or while the organism is active. Gamete formation is largely restricted to Phytomonadina, but life cycles may include an alternation of flagellate with palmella or with ameboid generations. The Phytamastigophorea include 10 orders: Chrysomonadida, Silicoflagellida, Coccolithophora, Heterochlorida, Cryptomonadida, Dinoflagellida, Ebriida, Euglenida, Chloromonadida, and Volvocida. See articles on these groups. *See also* MASTIGOPHORA; PROTOZOA; SARCOMASTIGOPHORA.

James B. Lackey

Phytoalexins

Antibiotics produced by plants in response to microorganisms. Plants use physical and chemical barriers as a first line of defense. When these barriers are breached, however, the plant must actively protect itself by employing a variety of strategies. Plant cell walls are strengthened, and special cell layers are produced to block further penetration of the pathogen. These defenses can permanently stop a pathogen when fully implemented, but the pathogen must be slowed to gain time.

The rapid defenses available to plants include phytoalexin accumulation, which takes a few hours, and the hypersensitive reaction, which can occur in minutes. The hypersensitive reaction is the rapid death of plant cells in the immediate vicinity of the pathogen. Death of these cells is thought to create a toxic environment of released plant components that may in themselves interfere with pathogen growth, but more importantly, damaged cells probably release signals to surrounding cells and trigger a more comprehensive defense effort. Thus, phytoalexin accumulation is just one part of an integrated series of plant responses leading from early detection to eventual neutralization of a potentially lethal invading microorganism.

Toxicity. One of the primary features of phytoalexins is their ability to stop the growth of microorganisms. Early workers demonstrated that pieces of live plant tissue could block the growth of fungi nearby. If the tissue was killed, the fungus would grow unchecked. Further experiments showed that extracts of plant tissue that successfully blocked the growth of an invading fungus were equally effective in inhibiting the growth of other microorganisms. The extracts could be added to fungi growing under laboratory conditions, and inhibition of growth and in some cases destruction of the fungi were observed.

The toxicity of phytoalexins is not restricted to fungi. Phytoalexins can also interfere with the growth of other disease-producing organisms (pathogens) important to agriculture, such as bacteria and nematodes. In fact, phytoalexins are even toxic to the plants that produce them. They are carefully localized where they are used, and are produced only when needed. Since they are able to disrupt many basic life processes, it is not surprising that these compounds are able to inhibit a variety of microorganisms.

The method of inhibition by phytoalexins is not completely understood, but membranes appear to be a common target. Processes that require membranes, such as transport of nutrients into and out of cells, and electron transport needed for production of energy, are often affected by phytoalexins.

The effects of phytoalexins are complicated by the ability of some microorganisms to chemically modify or detoxify phytoalexins. As might be expected, organisms that routinely colonize a particular plant are better able to metabolize that plant's phytoalexins than are a randomly selected group of plant pathogens. It is likely that in some cases the ability of a pathogen to degrade its host's phytoalexins determines how successful it will be as a pathogen.

Chemical classification. The tremendous capacity of plants to produce complex chemical compounds is reflected in the structural diversity of phytoalexins (see **illus.**). Each plant species produces one or several phytoalexins, and the types of phytoalexins produced are similar in related species: Members of the potato family (Solanaceae), for example, produce phytoalexins of the terpenoid chemical group; members of the bean family (Leguminoseae), on the other hand, tend to produce phytoalexins which are isoflavonoids.

The chemical complexity of phytoalexins led to the use of simple common names for the compounds long before definitive structures had been obtained. The names were derived in many cases from the Latin names of plants in which they were first found. Peas (*Pisum sativum*), for example, produce pisatin as a primary phytoalexin. Other phytoalexins are phaseollin (bean, *Phaseolus*), glyceollin (soybean, *Glycine*), gossypol (cotton, *Gossypium*), orchinol (orchid, *Orchis*), viniferins (grape, *Vitis vinifera*), and ipomeamarone (sweet potato, *Ipomoea*). Of course, the names can also reflect other historical factors as in the case of rishitin (purified from potato

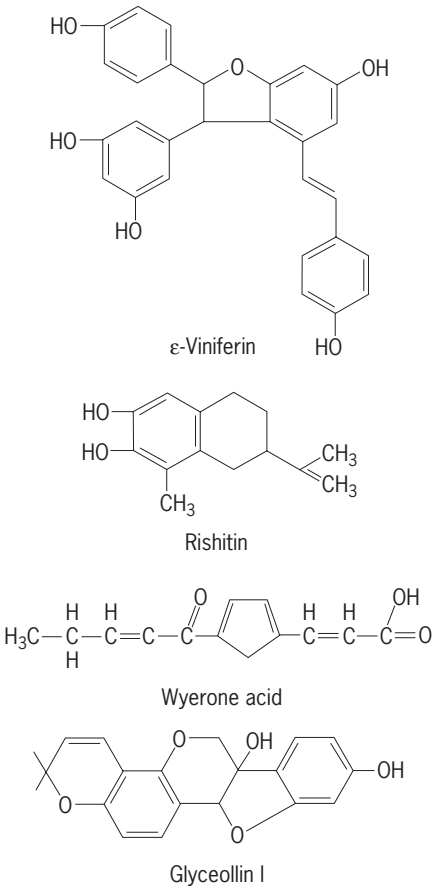

Structural formulas of four phytoalexins.

variety Rishiri) and wyerone acid (isolated from broad bean at Wye College, United Kingdom).

Role in disease resistance. The diversity, complexity, and toxicity of phytoalexins may provide clues about their function. The diversity of phytoalexins may reflect a plant survival strategy. That is, if a plant produces different phytoalexins from its neighbors, it is less likely to be successfully attacked by pathogens adapted to its neighbor's phytoalexins. Diversity and complexity, therefore, may reflect the benefits of using different deterrents from those found in other plants.

Phytoalexins appear to be in the right place at the right time. Healthy plant tissues have no detectable phytoalexins. In those cases in which a pathogen escapes detection and produces disease symptoms, phytoalexins may be produced, but it is too little too late.

Plants that exhibit only minimal damage by a fungal pathogen (those that are resistant) in many cases show a rapid accumulation of phytoalexins in excess of the concentrations needed to inhibit fungal growth. Accumulation is limited to very high concentrations in the cells directly adjacent to the pathogen. The rise in phytoalexins correlates in space and time with cessation of pathogen growth. Moreover, addition of chemicals that block phytoalexin synthesis permits otherwise impotent pathogens to develop and produce disease symptoms. Thus, in at least some host–pathogen interactions, phytoalexins play a critical role in resistance to disease.

Breeding for improved disease resistance has been essential to the success of modern agriculture, yet the function of resistance genes is not understood. Since the same defense mechanisms are triggered by different pathogens in the presence of different resistance genes, it is likely that resistance genes are involved in pathogen recognition. This is an important distinction. Disease resistance genes are not involved in making the biochemical machinery (enzymes) that is needed to produce a defensive response (such as phytoalexin accumulation), but rather they determine whether a particular pathogen will trigger the response. Therefore, it may be more appropriate to call these genes "pathogen recognition" genes.

Pathogen recognition. Plants detect pathogens in a variety of ways, but the most general detection system is based on damage to plant cell walls. Bacteria, fungi, nematodes, and other plant pathogens produce enzymes that degrade plant cell walls. Wall degradation releases nutrients used for pathogen growth and permits further invasion of plant tissues and also releases polysaccharide fragments, which (especially from the pectin portion of walls) stimulate phytoalexin accumulation in neighboring plant cells. Still smaller cell wall fragments are transported to other parts of the plant, making other leaves more resistant to attack. Thus, plants can express both local and systemic resistance.

Pathogens can also be directly detected through the release of components of their own cell walls. Plants produce enzymes that release polysaccharide fragments from fungal cell walls. These polysaccharides, termed elicitors, are produced by many different types of fungal pathogens and are in fact common constituents of fungi in general. Elicitors stimulate phytoalexin accumulation. Thus, the presence of elicitors may be a general signal of invasion that triggers the defensive response of antibiotic production. *See* PLANT PATHOLOGY. Arthur R. Ayers

Bibliography. J. A. Bailey and J. W. Mansfield (eds.), *Phytoalexins*, 1982; R. P. Sharma and D. K. Salunkhe, *Mycotoxins and Phytoalexins*, 1991.

Phytochrome

A pigment that controls most photomorphogenic responses in higher plants. As a result of the dependence of plants on the energy of sunlight and their sessile habit, mechanisms have evolved that allow them to adapt their growth and development to more efficiently seek and capture light and to tailor their life cycle to the climatic seasons. These mechanisms enable the plant to sense not only the presence of light but also its intensity, direction, duration, and spectral quality, which suggests a form of color vision. Plants thus regulate important developmental processes such as seed germination, growth direction, growth rate, chloroplast development, pigmentation, flowering, and senescence, collectively termed photomorphogenesis.

To perceive light signals, plants use several receptor systems that convert light absorbed by specific pigments into chemical or electrical signals to which the plants respond. This signal conversion is called photosensory transduction. The colors of light absorbed by the pigment determine the action spectrum, that is, the colors of light effective in initiating the photoresponse. Pigments used include cryptochrome, a blue light–absorbing pigment; an ultraviolet light–absorbing pigment; and phytochrome, a red/far-red light–absorbing pigment.

Spectral properties. Phytochrome consists of a compound that absorbs visible light (chromophore) bound to a protein. The chromophore is an open-chain (or linear) tetrapyrrole closely related to the photosynthetic pigments found in the cyanobacteria and similar in structure to the circular tetrapyrroles of chlorophyll and hemoglobin. Phytochrome is one of the most intensely colored pigments found in nature, enabling phytochrome in seeds to sense even the dim light present well beneath the surface of the soil and allowing leaves to perceive moonlight. *See* CHLOROPHYLL; HEMOGLOBIN.

Phytochrome can exist in two stable photointerconvertible forms, P_r or P_{fr}, with only P_{fr} being biologically active. Absorption of red light (near 666 nanometers) by inactive P_r converts it to active P_{fr}, while absorption of far-red light (near 730 nm) by active P_{fr} converts phytochrome back to inactive P_r (**Fig. 1**). Photoconversion of P_r to P_{fr} involves a *cis*-to-*trans* isomerization of one of the double bonds between the four pyrrole rings, a 31° reorientation of the chromophore relative to the protein moiety, and conformational changes within the protein; no additional factors are required for phytochrome phototransformations other than light. As a result, even the purified pigment can be repeatedly photointerconverted between P_r and P_{fr} in the test tube. During the conversion of phytochrome from one form to the other, the molecule passes through a series of intermediates lasting from nanoseconds to milliseconds, with the final P_r or P_{fr} form appearing within a second after light absorption.

Because the absorption of P_r and P_{fr} overlap in the red region of the spectrum, red light is not able to convert all P_r to P_{fr}. Thus even in saturating red light, a photoequilibrium of 87% P_{fr} and 13% P_r is established (**Fig. 2**). The residual P_r that remains after phytochrome solutions have been irradiated with red light can be observed in the absorbance spectrum as a shoulder at 665 nm. In contrast, because P_r does not absorb light of wavelengths greater than approximately 700 nm, far-red light is able to convert essentially all P_{fr} back to P_r. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION.

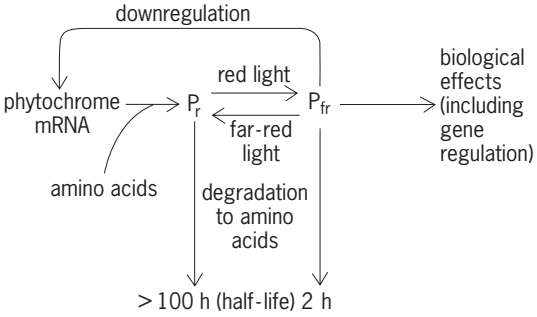

Fig. 1. Phytochrome pathway, showing the photointerconversion between the red-absorbing (P_r) and the far-red-absorbing (P_{fr}) forms, phytochrome degradation, and the control of phytochrome synthesis.

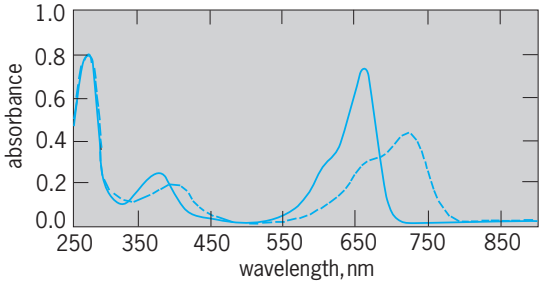

Fig. 2. Absorption spectra of phytochrome in the red-absorbing form (P_r; solid line) and after irradiation with red light producing mainly the far-red-absorbing form (P_{fr}; broken line).

Phytochrome protein. The protein moiety of phytochrome ranges in molecular mass from 116 to 127 kilodaltons, depending on the plant species. Genes encoding phytochrome proteins have been characterized from a variety of both monocot and dicot species. In all species studied, phytochrome is synthesized by a small gene family containing several members that encode related but not identical proteins. The complete amino acid sequence of phytochromes within the same plant species and among various plant species ranges in homology from about 60% to greater than 80%, with more closely related species displaying a greater degree of protein homology.

Structural analyses of phytochrome show that it consists of two large domains, an amino-terminal domain of approximately 74 kDa that contains the chromophore and a 55-kDa domain responsible for holding the phytochrome dimers together. The spectral properties of the 74-kDa fragment are identical to the intact molecule, indicating that it contains all the domains required for correct chromophore–protein interactions. The amino-terminal region, which is rich in serine or threonine, appears to extend significantly from the surface of the 74-kDa domain, which may explain its susceptibility to proteolytic cleavage.

A model for the quaternary structure of phytochrome (**Fig. 3**) proposes that phytochrome monomers are made of an ellipsoid and a cylinder comprising the 74-kDa and 55-kDa domains, respectively. The two 55-kDa cylinders interact, holding the phytochrome dimer together so that the 74-kDa ellipsoids are oriented in the same plane but 90° from each other. In this way, phytochrome has a cloverleaf shape when observed from above. This asymmetric shape explains why phytochrome behaves in solution as a protein with a molecular size (355 kDa) much larger than its actual mass (250 kDa).

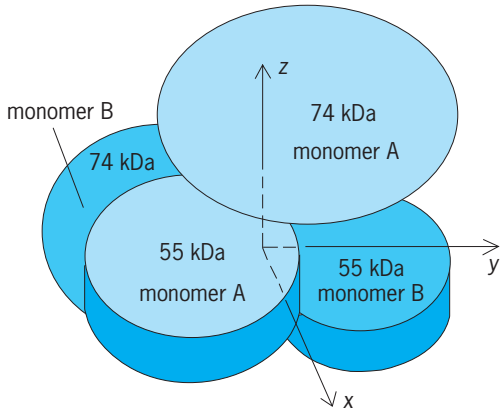

Fig. 3. Molecular model of a phytochrome dimer. The 74-kDa ellipsoids are oriented in the same plane but 90° from each other. (*After S. Tokutomi et al., A model for the dimeric molecular structure of phytochrome based on small-angle x-ray scattering, FEBS Lett., 247:139–142, 1989*)

Biosynthesis and degradation. Plants frequently respond quantitatively to light by detecting the amount of P_{fr} produced. As a result, the amount of P_{fr}—and hence phytochrome—must be strictly regulated nonphotochemically by precisely controlling both the synthesis and degradation of the pigment. Phytochrome also appears to be posttranslationally modified by phosphorylation and possibly by glycosylation.

In dark-grown (etiolated) seedlings, P_r is continually synthesized. Because its half-life is greater than 4 days, P_r accumulates to high levels within the tissue. However, following transfer of etiolated plants into the light, most phytochrome disappears after photoconversion to P_{fr}, which has a half-life of only 1–2 h. Rapid P_{fr} degradation is the result of the selective recognition of P_{fr} by a proteolytic pathway involving the small protein ubiquitin. Conjugated ubiquitin serves as a recognition signal for a specific protease that degrades only the target protein, releasing ubiquitin in a reusable form.

In addition to enhanced phytochrome degradation following P_{fr} conversion, a feedback mechanism exists whereby P_{fr} inhibits further synthesis of P_r. This inhibition involves not only the repression of continued phytochrome gene transcription into messenger ribonucleic acid (mRNA) but also an enhanced rate of phytochrome mRNA degradation. The repression is highly sensitive to P_{fr}, requiring only 1% of P_r to be converted to P_{fr} to maximally inhibit transcription. Upon the return of irradiated seedlings to the dark, synthesis of phytochrome mRNA slowly resumes.

Etiolated plants thus contain up to 100 times more phytochrome than light-grown plants, allowing a seedling with limited resources to sense even the dimmest light and thus reach the surface of the soil before depleting its energy. Once at the surface, however, detecting light is not as much a problem, and thus much less of the pigment is required to initiate photomorphogenesis.

Occurrence. Phytochrome has been detected in a variety of tissues from all higher plants examined. It also has been found in certain mosses, liverworts, and algae. In etiolated plants, concentrations of phytochrome are especially high in the root cap, crown meristem, and tip of monocot seedlings and the hook and cotyledons of etiolated dicot seedlings. Because of their high phytochrome content, etiolated seedlings are a convenient source from which the protein may be purified. Both the low levels of phytochrome and the presence of chlorophyll have substantially hampered the detection and isolation of the chromoprotein from light-grown tissues, but partially purified preparations of light-grown or "green" phytochrome have been obtained. Green phytochrome appears to differ slightly from that isolated from etiolated seedlings, suggesting that another type of phytochrome is synthesized in light-grown plants.

Subcellular localization. In etiolated seedlings, P_r is found primarily throughout the cytoplasm, with none detected in mitochrondria, chloroplasts, or nuclei. There is no evidence that a significant fraction of phytochrome is associated with membranes of higher plants. Once converted from P_r to P_{fr}, phytochrome changes from a disperse cytoplasmic distribution to one in which the molecules coalesce into many amorphous discrete areas. This sequestration is rapid, occurring within seconds, and requires metabolic energy. In monocots most phytochrome appears to aggregate, whereas in dicots only a portion aggregates while the rest remains dispersed. Following reconversion of P_{fr} to P_r, the pigment slowly reverts to a dispersed form. The function of phytochrome sequestering is unknown, but it may represent molecules that are awaiting degradation.

The failure of localization studies to find phytochrome on membranes or in the nucleus suggests that it does not act by directly altering membrane permeability or gene transcription. However, it is possible that the bulk of phytochrome detected in the cytoplasm is not active or that only a small percentage becomes active following interaction with the appropriate membrane or compartment. The filamentous alga, *Mougeotia*, has a single flat chloroplast that rotates to orient itself perpendicular to light; the movement is phytochrome-mediated. Experiments with microbeams of light demonstrate that the active P_{fr} is stably attached to a structure close to or on the plasma membrane.

Functions. Phytochrome has a variety of functions in plants. Initially, production of P_{fr} is required for many seeds to begin germination. This requirement prevents germination of seeds that are buried too deep in the soil to successfully reach the surface. In etiolated seedlings, phytochrome can measure an increase in light intensity and duration through the increased formation of P_{fr}. Light direction also can be deduced from the asymmetry of P_{fr} levels from one side of the plant to the other. Different phytochrome responses vary in their sensitivity to P_{fr}; some require very low levels of P_{fr} (less than 1% of total phytochrome) to elicit a maximal response, while others require almost all of the pigment to be converted to P_{fr}. Thus, as the seedling grows toward the soil surface, a cascade of photomorphogenic

responses are induced, with the more sensitive responses occurring first. This chain of events produces a plant that is mature and photosynthetically competent by the time it finally reaches the surface. Production of P_{fr} also makes the plant aware of gravity, inducing shoots to grow up and roots to grow down into the soil. *See* PLANT MOVEMENTS; SEED.

In light-grown plants, phytochrome allows for the perception of daylight intensity, day length, and spectral quality. Intensity is detected through a measurement of phytochrome shuttling between P_r and P_{fr}; the more intense the light, the more interconversion. This signal initiates changes in chloroplast morphology to allow shaded leaves to capture light more efficiently. If the light is too intense, phytochrome will also elicit the production of pigments to protect plants from photodamage.

Temperate plants use day length to tailor their development, a process called photoperiodism. How the plant measures day length is unknown, but it involves phytochrome and actually measures the length of night. A red light pulse given in the middle of a long night causes plants to believe that they have experienced a short night. Conversely, a far-red light pulse given before a short night makes plants believe that they have experienced a long night. These "night breaks" are widely used commercially to time the flowering of plants. *See* PHOTOPERIODISM.

Finally, phytochrome allows plants to detect the spectral quality of light, a form of color vision, by measuring the ratio of P_r to P_{fr}. When a plant is grown under direct sun, the amounts of red and far-red light are approximately equal, and the ratio of P_r to P_{fr} in the plant is about 1:1. Should the plant become shaded by another plant, the P_r/P_{fr} ratio changes dramatically to 5:1 or greater. This is because the shading plant's chlorophyll absorbs much of the red light needed to produce P_{fr} and absorbs almost none of the far-red light used to produce P_r. For a shade-intolerant plant, this change in P_r/P_{fr} ratio induces the plant to grow taller, allowing it to grow above the canopy.

Mechanisms of action. It is not known how phytochrome elicits the diverse array of photomorphogenic responses, but the regulatory action must result from discrete changes in the molecule following photoconversion of P_r to P_{fr}. These changes must then start a chain of events in the photosensory transduction chain leading to the photomorphogenic response. Many photosensory transduction chains probably begin by responding to P_{fr} or the P_r/P_{fr} ratio and branch off toward discrete end points.

To analyze phytochrome action, several avenues of research have been taken. One has been an attempt to assign enzymatic properties to the chromoprotein in the plant. Phytochrome may be a protein kinase capable of adding phosphate groups to other proteins, and given that protein phosphorylation is an important regulator of cell development and metabolism, such an enzymatic activity would be an attractive possibility. *See* ENZYME.

Another avenue has been to identify chemical differences between P_r and P_{fr}. Analyses indicate that in addition to changes in the chromophore, several protein domains change conformation following phototransformation. The most interesting and dramatic changes occur within 6–10 kDa of the amino terminus. This region appears more extended as P_r and becomes more protected as P_{fr}. It has been determined that the amino-terminal region interacts directly with the chromophore, potentially moving to shield the chromophore following photoconversion of P_r to P_{fr}. Other protein domains also undergo conformational changes during phototransformation. They exist within the molecule, and as opposed to the amino terminus, these regions are less exposed as P_r than as P_{fr}.

Which changes in conformation during phototransformation are responsible for phytochrome action remains to be determined. An approach to identifying these critical domains has been developed through the use of genetic engineering. It involves the transfer of an oat phytochrome gene into tobacco plants by using the bacterium *Agrobacterium tumefaciens* as the transfer agent. The tobacco plants then express both their own phytochrome and the inserted oat phytochrome, and as a result they have artifically higher levels of the chromoprotein. In light-grown plants, more than nine times more phytochrome is present in the genetically engineered plants, and their growth and development are altered. They exhibit a dwarfed bushy shape and have darker green leaves (**Fig. 4**). These changes indicate that oat phytochrome is morphogenically active in tobacco and can be used as an assay for the functional molecule in the plant. *See* GENETIC ENGINEERING.

Another possible strategy to elucidate phytochrome's mode of action is to trace back through the sensory transduction chain of a particular phytochrome response to identify the initial molecular events. The best-characterized response is phytochrome's control of greening, which involves the development of a photosynthetically competent chloroplast. Chloroplast development requires the synthesis of many proteins encoded by both the nuclear and chloroplast genomes. This synthesis is

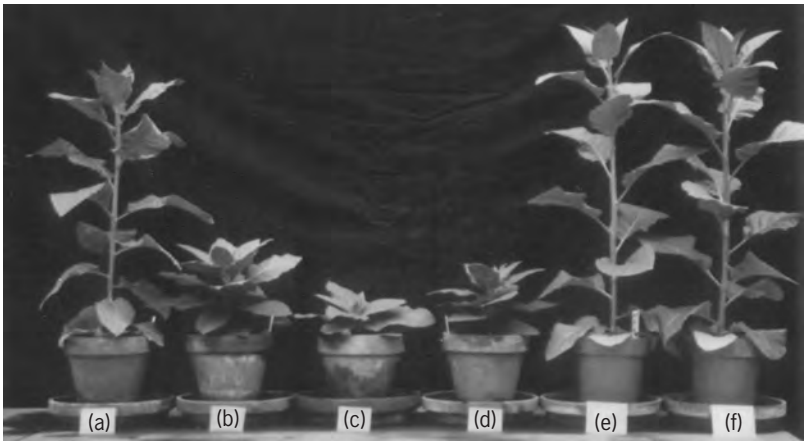

Fig. 4. Morphological effect of oat phytochrome synthesis in transgenic tobacco. (*a*) Transformed plant not expressing oat phytochrome. (*b–d*) Transformed plants expressing oat phytochrome. (*e, f*) Two nontransformed plants.

controlled directly or indirectly by phytochrome and involves transcriptional activation of the corresponding genes. How these genes are activated is unclear, but it requires the synthesis or activation of proteins, commonly called transcription or trans-acting factors, that bind to genes. This binding may facilitate gene transcription. Whereas many phytochrome responses require gene transcription, not all of them do. Phytochrome may function by stimulating a response similar to muscle contraction.

The third possible strategy is the isolation and characterization of mutants in phytochrome responses, including those with reduced levels of chromoprotein. Other mutants show attenuated phytochrome responses but have normal amounts of phytochrome, indicating a defect in the photosensory transduction chain. *See* PHOTOMORPHOGENESIS.

Richard D. Vierstra

Bibliography. P. Adamse, R. E. Kendrick, and M. Koornneef, Photomorphogenetic mutants of higher plants, *Photochem. Photobiol.*, 48:833–841, 1988; J. M. Keller et al., Expression of a functional monocotyledonous phytochrome in transgenic tobacco, *EMBO J.*, 8:1005–1012, 1989; R. E. Kendrick (ed.), *Photomorphogenesis in Plants*, 2d ed., 1993; H. McGee (ed.), *A Pigment of the Imagination: USDA and the Discovery of Phytochrome*, 1987; P. B. Moses and N.-H. Chua, Light switches for plant genes, *Sci. Amer.*, 258:88–93, 1988.

Phytoplankton

Mostly autotrophic microscopic algae which inhabit the illuminated surface waters of the sea, estuaries, lakes, and ponds. Many are motile. Some perform diel (diurnal) vertical migrations, others do not. Some nonmotile forms regulate their buoyancy. However, their locomotor abilities are limited, and they are largely transported by horizontal and vertical water motions.

Energy-nutrient cycle. All but a few phytoplankton are autotrophic—they manufacture carbohydrates, proteins, fats, and lipids in the presence of adequate sunlight by using predominantly inorganic compounds. Carbon dioxide, water, inorganic nutrients such as inorganic phosphorus and nitrogen compounds, trace elements, and the Sun's energy are the basic ingredients. This organic matter is used by other trophic levels in the food web for nourishment. The energy in this organic matter ultimately is released as heat in the environment. The inorganic materials associated with this organic matter, on the other hand, are recycled. They are released back into the aquatic environment in inorganic form as a result of metabolic processes at all levels in the food web, and in large part they again become available for reuse by the phytoplankton. Some structural materials, such as the silica in diatoms and silicoflagellates, and carbonates in the scales of certain chrysophytes, are returned to the aquatic environment by chemical dissolution rather than by biological processes. *See* FOOD WEB.

Varieties. A great variety of algae make up the phytoplankton. Diatoms (class Bacillariophyceae) are often conspicuous members of marine, estuarine, and fresh-water plankton. Reproducing mostly asexually by mitosis, they can divide rapidly under favorable conditions and produce blooms in a few days' time. Their external siliceous skeleton, termed a frustule, possesses slits, pores, and internal chambers that render them objects of great morphological complexity and beauty which have long attracted the attention of microscopists. *See* BACILLARIOPHYCEAE.

Dinoflagellates (class Dinophyceae) occur in both marine and fresh-water environments and are important primary producers in marine and estuarine environments. Dinoflagellates possess two flagella; one trails posteriorly and provides forward motion, while the other is positioned more or less transversely and often lies in a groove encircling the cell. This flagellum provides a rotary motion; hence the name dino- (whirling) flagellate. In some dinoflagellates the cell wall is thin, while in others it consists of a complicated array of rather thick cellulosic plates, the number and arrangement of which are characters used in identifying genera and species.

The dinoflagellates are often conspicuous members of the marine plankton. Some taxa are bioluminescent. Others are one of the causes of red water, often called red tides, although their occurrence is not related to the tides. Depending on the causative species, red water may be an innocuous discoloration of the water, or if *Gymnodinium breve* is dominant, extensive fish mortalities will be experienced. Several species of *Gonyaulax* occur in bloom proportions in inshore marine waters. They contain a toxin which can accumulate in shellfish feeding upon *Gonyaulax* and which is the ultimate cause of paralytic shellfish poisoning in humans.

Coccolithophorids (class Haptophyceae) are also marine primary producers of some importance. They do not occur in fresh water. This class of algae possesses two anterior flagella. A third flagellumlike structure, the haptonema, is located between the two flagella. Calcium-carbonate-impregnated scales, called coccoliths, occur on the surface of these algae, and are sometimes found in great abundance in recent and ancient marine sediments. *See* COCCOLITHOPHORIDA.

Under certain conditions in subtropical and tropical seas, members of the nitrogen-fixing blue-green algal genus *Trichodesmium* (class Cyanophyceae) can occur in sufficient concentrations to strongly discolor the surface of the sea. Other nitrogen-fixing blue-greens commonly occur in great abundance in eutrophic and hypereutrophic fresh-water lakes and ponds.

Members of still other algal classes occur in marine and estuarine plankton. Their abundance will vary in different environments at different times, and they may even occasionally dominate the standing crop. Much remains to be learned about the identity, physiology, and ecology of the very small (<8 micrometers) flagellates which commonly occur in almost every marine phytoplankton sample. *See* ALGAE.

Communities. Even though marine and fresh-water phytoplankton communities contain a number of algal classes in common (such as Bacillariophyceae, Chrysophyceae, and Dinophyceae), phytoplankton samples from these two environments will appear quite different. These habitats support different genera and species and groups of higher rank in these classes. Furthermore, fresh-water plankton contains algae belonging to additional algal classes either absent or rarely common in open ocean environments. These include the green algae (class Chlorophyceae), the euglenoid flagellates (class Euglenophyceae), and members of the Prasinophyceae. *See* FRESH-WATER ECOSYSTEM; MARINE ECOLOGY.

In estuarine environments in which salinities vary from essentially zero to those of the local inshore ocean salinities, a mixture of organisms characteristic of both fresh-water and marine environments will be encountered. In addition, giving the estuarine plankton a somewhat distinctive appearance, some euryhaline planktonic taxa will be present. *See* ESTUARINE OCEANOGRAPHY.

Samples of phytoplankton in shallow areas of the sea, lakes, and estuaries often contain benthic and epiphytic microalgae. These algae become suspended in the water as a result of strong turbulent mixing so often found in shallow aquatic environments, and are called meroplankton to distinguish them from the holoplankton organisms which are truly planktonic.

While phytoplankton community composition reflects some complicated and as yet rather poorly understood series of biotic and abiotic interaction within ecosystems, the chemical composition of water is recognized as an important factor affecting phytoplankton communities. Many of the differences between the marine, estuarine, and fresh-water phytoplankton communities are associated with changes in salinity, major ion concentrations and ratios, pH, nutrients, and quite possibly trace-element concentrations. *See* ECOSYSTEM.

Eutrophication. Society's common use of lakes as receptacles for wastes, coupled with nutrient-rich runoff from cultivated fertilized land, has had pronounced biological effects not only on the phytoplankton but on other levels in the food web. Domestic sewage and agricultural runoff are quite rich in inorganic phosphorus, which is generally in short supply in most inland bodies of water. The addition of a nutrient which is limiting phytoplankton production, such as phosphorus, to fresh-water environments increases primary production, alters phytoplankton composition, and can lead to dense algal blooms which adversely affect water quality and esthetic and recreational values. The effects of nutrient enrichment, generally known as eutrophication, have been understood by limnologists since the 1920s–1930s, but the American public became aware of the implication only in the 1960s with the widespread use of phosphate detergents and the associated and accelerated deterioration of the quality of lakes and ponds. Advanced techniques of sewage treatment are now available which remove phosphates sufficiently to greatly reduce the impact of sewage effluent upon receiving waters. Important in regulating these cycles are nutrient availability and the presence of a water density gradient, the pycnocline, at some relatively shallow depth (commonly 16.5–165 ft or 5–50 m) below the surface. Above the pycnocline, the location of which is usually well correlated with a zone of rapidly changing water temperature called the thermocline, is a well-mixed region of uniform density and temperature. This layer is called the mixed layer by oceanographers and the epilimnion by limnologists. Within this mixed layer, water motions caused by winds help keep the phytoplankton in suspension within the illuminated surface layer. The pycnocline also greatly reduces the rate of diffusion of nutrients upward into the mixed layer from nutrient-rich deeper waters and thus has a negative effect upon phytoplankton productivity. Any mechanism which introduces nutrients into the mixed layer that are limiting phytoplankton production will enhance primary production. *See* EUTROPHICATION; LIMNOLOGY; WATER POLLUTION.

Seasonal cycles. The phytoplankton in aquatic environments which have not been too drastically affected by human activity exhibit rather regular and predictable seasonal cycles. Coastal upwelling and divergences, zones where deeper water rises to the surface, are examples of naturally occurring phenomena which enrich the mixed layer with needed nutrients and greatly increase phytoplankton production. In the ocean these are the sites of the world's most productive fisheries. *See* UPWELLING.

In marine and fresh waters in temperate to arctic latitudes, the seasonal cycle of phytoplankton is to a large degree influenced by the pycnocline. During winter months a strong pycnocline is generally not present, and any planktonic plants in the water circulate to considerable depths. Even though nutrient concentrations are high during this period of deep vertical mixing, any phytoplankton present do not spend sufficient time in the illuminated surface layers for photosynthesis to exceed respiration. As spring approaches, incident solar radiation leads to the formation of a pycnocline. Phytoplankton then proliferate, giving rise to the annual spring phytoplankton bloom in which diatoms frequently dominate (thus it is often called the spring diatom bloom). The phytoplankton quickly exhaust the nutrients in the newly formed mixed layer, and throughout the summer phytoplankton concentrations generally remain low due to low mixed-layer nutrient concentrations resulting in part from the restricted upward movement of nutrients through the pycnocline. In the fall, incident solar radiation decreases and the pycnocline deepens. Nutrients or phytoplankton living in the pycnocline are incorporated into the deepening mixed layer, and a fall bloom occurs and then wanes as the pycnocline deepens, and finally disappears.

In the tropical oceans and lakes located near sea level, a permanent thermocline exists, and phytoplankton production remains low but may increase

if nutrients are injected into the mixed layer as a result of strong wind mixing or other advective processes.

Fossilization. Under favorable conditions, some phytoplankton are incorporated into the sediments and become part of the fossil record. Diatoms, silicoflagellates coccoliths, and the cysts of dinoflagellates are frequently sufficiently abundant and well preserved in lake and ocean sediments to permit paleontologists to reconstruct past environmental conditions and changes in environmental conditions through time. These fossilized remains are also used by stratigraphers to determine the age of sediments. *See* MARINE SEDIMENTS; PALEOECOLOGY; STRATIGRAPHY.

Robert W. Holmes

Bibliography. A. D. Boney, *Phytoplankton*, 2d ed., 1992; G. A. Cole, *Textbook of Limnology*, 4th ed., 1994; I. Morris (ed.), *The Physiological Ecology of Phytoplankton*, 1981; J. E. Raymont, *Plankton and Productivity in the Oceans*, 2 vols., 2d ed., 1980, 1983; C. D. Sandgren (ed.), *Growth and Reproduction Strategies of Freshwater Phytoplankton*, 1992.

Phytotronics

Research using whole plants and conducted under controlled environmental conditions to determine responses to a single or known combination of environmental elements. Originally, the term phytotronics was used to identify research conducted specifically in phytotrons where controlled plant growth units are available for simultaneous use. The name phytotronics is also often applied to any research conducted with whole plants in a controlled environment plant growth chamber or room.

The word phytotronics came into use after Fritz W. Went built the Earhart Plant Research Laboratory at the California Institute of Technology in 1949. This first phytotron consisted of glasshouses and artificially lighted plant growth facilities, each with controlled temperatures and lighting. The first research projects demonstrated the significance of a phytotron in advancing botanical knowledge, and resulted in the building and continuing operation of phytotrons throughout the world.

Scope. A phytotron is a research tool containing a large number of individually controlled environments (**Fig. 1**). These provide the means of studying the effect of each environmental factor such as temperature or light at many levels simultaneously. Thus, phytotronics is usually identified with research that involves the use of a range of environmental factors. In most phytotrons, plants are on carts that can be moved easily from one controlled environment to another. This provides the flexibility to greatly increase the number of environmental conditions available over the number of controlled units. For instance, using only five units and redistributing carts daily, 25 combinations of day/night conditions can be obtained (**Fig. 2**). The large number of controlled units in a phytotron also makes it possible to investigate the interaction of the various environmental factors on plants. However, before, or simultaneously with, the study on factor interactions, the effects of the individual factors must be determined if the interaction effects are to be understood. Effects of environmental factors are more often synergistic or antithetic than additive because plant growth is capable of adapting, within limits, to environmental changes.

Fig. 1. Large plant growth chamber with a solid bank of fluorescent incandescent lamps forming the ceiling. The mirrorized walls and the two rows of diffusers with their high volume of airflow provide uniform light, temperature, humidity, and carbon dioxide conditions.

Phytotronics has greatly extended the precision and scope of botanical research in the laboratory. Phytotronics is used also to complement field research, especially in the area of agriculture and ecology. The building of phytotrons became feasible only after the development of relatively low-cost

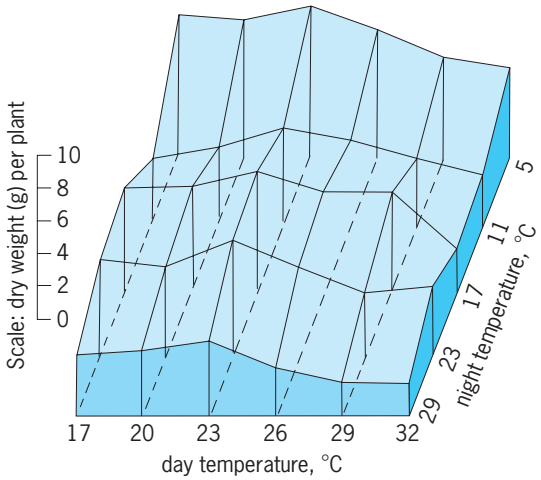

Fig. 2. Results of experiment using five growth chambers and moving the plants to obtain 25 combinations of day-night conditions. Increased total plant dry weight due to a 5°C (41°F) night with all day temperature was unexpected but later confirmed in the New Zealand *Pinus radiata* plantations. °F = (°C × 1. 8) + 32. (*After H. Hellmers and D. A. Rook, Air temperature and growth of radiata pine seedlings, N. Z. J. For. Sci., 3:271–285, 1973*)

air-conditioning systems and adequate lighting systems, utilizing high-output fluorescent lamps. Prior to this, whole plant research was conducted under field or slightly modified field or glasshouse conditions. In the field and in ordinary glasshouses the various environmental factors continually fluctuate, making it impossible to determine which factor or combination of factors produces an observed plant response. Phytotrons provide the means of dissecting the environment and studying the effects of the parts by controlling the individual components. Thus, while one factor is held at fixed but different levels in different plant growth units, all other factors are maintained constant or at a known rate of change in all units.

Precision in phytotronic research has increased with improvements in the ability to control the various environmental factors. While temperature and light were the factors studied in most early experiments, improvements in equipment not only provided facilities with greater accuracy of control of these factors, but made it feasible to work with different humidities and with various concentrations of atmospheric gases. Systems are available where the conditions surrounding the roots are controlled separately from the conditions to which the plant top are exposed. As equipment improves, phytotronics will increase in reliability, usefulness, and complexity. One equipment problem is that of lighting, namely, trying to provide artificial light approximating the intensity and wavelength distribution of sunlight received at the Earth's surface in summer.

The growth and development of plants is an expression of their genetic capability as affected by environmental factors. Thus, the potential for phytotronic studies is exceedingly great when one considers the large number of plant species in the world with their genetic variability and the wide range of each environmental factor.

Types of studies. Temperature, probably the most extensively and intensively studied environmental factor in phytotronics, affects the physical process of transpiration as well as all enzymatically controlled physiological processes. Thus, temperature is a dominant environmental factor affecting plant growth and development. Wide ranges of day and night temperature combinations have been used to determine the growth potential of many plant species. Phytotronic studies employing ranges of temperature conditions also have been used to determine the temperature requirements for the development of specific plant parts. These studies include among others the determination of the number of hours at temperatures below a set point, such as 45°F (7°C), required to break bud dormancy in deciduous fruit trees, and the number of heat units, termed degree-hours (temperature × time) required for fruit maturation. Optimum temperature requirements are often different for the various phases of plant development, including seed germination, root, shoot, and leaf growth dry weight, flowering, fruit set, and seed filling. Results from studies of this type are used by plant breeders, glasshouse operators, and agriculturists. The latter two use the information to determine planting and harvesting dates and for selecting plant varieties that give the best product under their expected glasshouse or field temperatures.

While temperature is a basic factor in most phytotronic studies, considerable work has been conducted in which temperature is held at fixed levels, and other environmental factors such as light intensity, light duration (photoperiod), light quality, water stress, humidity, nutrients, carbon dioxide, or air pollutant concentrations are varied. The objective is usually to determine the effects of these environmental factors, either individually or in combination, on a specific phase of plant growth or development.

Results of studies. Examples from a few selected studies serve to illustrate the practical and theoretical values of phytotronics. At least one plant species, tobacco, has been studied to the extent that under controlled environmental conditions plants can now be produced virtually to match field-grown plants at any given age. This requires environmental manipulation which provides stress as well as optimum growth conditions at specific physiological stages of plant development. Thus, plants for experimental work can be provided at any given growth stage throughout the year. In addition, many experiments have been conducted to supply specific information and numbers for use in mathematical plant growth models that have been or are being developed for wild plants as well as agricultural crop species.

Research with temperature as the variable has been conducted on many species. Itchgrass (*Rottboellia exaltata*), a weed introduced into Louisiana and Florida, represents a serious potential weed problem in the warmer moist parts of the United States. Phytotronic studies showed that unless cool-climate ecotypes develop the weed will be less of a problem in the Corn Belt and will probably not be competitive with crops in the northern regions. Some of the results also illustrate the effect of an environmental factor on different constituents of plant growth. Total plant dry weight increased with increased day and night temperatures, but height growth peaked at a cooler night condition, 73°F (23°C). These plants were grown in all 36 temperature conditions simultaneously, thus avoiding any differences that would be introduced due to a time factor. In another study an unusually low night temperature, 41°F (5°C), was found to favor the growth of radiata pine (*Pinus radiata*), which accounts for the more rapid growth of these trees in New Zealand than in their native California. In contrast, growth of young redwood trees (*Sequoia sempervirens*) is severely retarded when night temperature falls below 52°F (11°C). It is probably this factor which determines the northern and eastern limit of its range. Other studies using multiple day/night temperature conditions have also been used to screen plants for introduction to specific areas or conversely to identify areas where exotic plants can be grown.

Phytotronic studies have played key roles in

breeding of rice, wheat, and numerous other agricultural crops as well as in matching crops to environmental conditions. For example, the bulb diameter of Creamgold onions was found to be less if the plants grew at temperatures below a 79°F (26°C) day with a 64°F (18°C) night, while the Braeside variety was unaffected, making the latter the better choice for early spring planting.

Many studies have added to or changed basic concepts of what controls plant growth and development. For instance, by alternating weak and intense light on arctic plants it was found that they could be induced to flower, a response previously thought to require darkness. Other studies have shown that enrichment of the atmospheric CO_2 suppresses flowering in some short-day plants under inductive conditions and induces flowering of some long-day plants under noninductive conditions. Such changes in flower induction could be of major ecological importance if the world's atmospheric CO_2 content increases. Also, CO_2-enriched air during the light period or low night temperature can cause a starch buildup to 40–50% of leaf dry weight in some plants. Differences in the physiological response to temperature and day length have been found in races from the northern and southern part of the range of plant species with a wide north-south range.

The identification and confirmation of the acclimation ecotypes concept was obtained by growing arctic and alpine plants simultaneously in the same environmental conditions. Leaf temperature was discovered to control the ratio of chlorophyll *a* to chlorophyll *b*, resulting in size differences of the chlorophyll unit frequently noted in chlorophyll-deficient mutants. Metabolic processes controlling growth, pigments, and secondary plant products are temperature-sensitive and thus have often been studied. Sugar content of sugar beets was found to be markedly increased by cool conditions, a finding now used by growers in timing harvests. The development and aftereffects of water stress have been shown to be affected by temperature and light regimes not only because of the energy aspects per se, but because temperature and light caused pronounced effects on stomatal guard cells and enzymatic systems.

In a study using 18 environmental combinations, three temperatures, three photoperiods, and two light intensities, it was determined that even though light affected CO_2 assimilation it was temperature which regulated the pathway of CO_2 uptake either via organic acids or ribulosediphosphate in crassulacean acid metabolism (CAM) plants. *See* BREEDING (PLANT); PHOTOPERIODISM; PHYSIOLOGICAL ECOLOGY (PLANT); PLANT GROWTH. Henry Hellmers

Bibliography. P. Chouard and N. de Bilderling, *Phytotronics in Agricultural and Horticultural Research*, 1975; P. Chouard, N. de Bilderling, and R. Jacques, *Phytotronic Newsl.*, nos. 1–19, 1971–1979; *Environment Control in Biology*, vols. 1–17, Journal of Japanese Society of Environment Control in Biology, 1962–1979; F. W. Went, *The Experimental Control of Plant Growth*, 1957.

Piciformes

A large order of land birds, second in number of species only to the Passeriformes, that is found throughout the world except for the Australian region and is concentrated in tropical areas. The relationships of the suborder Galbulae to the other piciforms is strongly disputed, with some researchers placing the Galbulae with the Coraciiformes or in a separate order, the Galbuliformes. The evidence for retaining the Galbulae in the Piciformes as well as for shifting them to the Coraciiformes or elsewhere within the birds is inconclusive; here the Galbulae will be kept with the Piciformes, for which there appears to be somewhat better evidence. The overall affinities of the Piciformes may, in fact, be with other land birds such as the Coraciiformes and Passeriformes. *See* CORACIIFORMES; PASSERIFORMES.

Classification. The Piciformes is divided into the following three suborders and eight families; the family Picidae is the largest, with 204 species of woodpeckers (see **illustration**).

Order Piciformes
- Suborder Galbulae
  - Family: Primobucconidae (fossil; Eocene of North America)
    - Galbulidae (jacamars; 17 species; Neotropics)
    - Bucconidae (puffbirds; 3 species; Neotropics)
- Suborder Zygodactyli
  - Family Zygodactylidae (fossil; Miocene of Germany)
- Suborder Pici
  - Family: Capitonidae (barbets; 81 species; pantropical)
    - Ramphastidae (toucans; 33 species; Neotropics)

Red-headed woodpecker (*Melanerpes erythrocephalus*) in Desoto National Wildlife Refuge, Iowa. (*Photo by Dave Menke/U.S. Fish and Wildlife Service*)

Indicatoridae (honeyguides; 16 species; Africa and southern Asia)
Picidae (woodpeckers; 204 species; worldwide except the Australasian region)

Fossil record. The fossil record of the piciforms is quite poor, as expected for a group of woodland-dwelling birds. The earliest fossil barbets and woodpeckers are known from the Miocene, and the earliest honeyguides from the Pliocene. The family Primobucconidae from the Eocene of North America constitutes a small, rather diverse group that will require considerable study before their relationships can be determined with any certainty; they may belong to the Coraciiformes. *Zygodactylus*, from the early Miocene of Bavaria, is a puzzling fossil and is only tentatively included in the Piciformes; possibly it is a parrot. *See* PSITTACIFORMES.

Characteristics. The piciforms are small to medium-sized, hole-nesting land birds. The bill is short to medium long, straight, and strong, and the wings are of medium length and rounded. Toucans have a huge colorful bill that dominates the bird and is used to reach for its food. The legs are short and strong, with the strong toes arranged in a zygodactylous (yoke) pattern, with two toes forward and two toes back; this two-toe arrangement is less developed in the Galbulae. The tail may have stiffened feathers serving as a prop when climbing. The plumage, which varies greatly in hue, is frequently brightly colored and boldly patterned. Piciforms are good fliers and can easily perch and climb, but they walk poorly. Most species feed on insects; toucans, on the other hand, are primarily fruit eaters. Most are solitary, although some barbets and woodpeckers live and breed in social groups. All breed in cavities dug in earthen banks or hollowed out in trees. The eggs are incubated by both sexes, and both parents care for the unfeathered young, which remain in the nest. Except for a few species of woodpeckers, the piciforms are nonmigratory.

Woodpeckers are specialized for climbing on and drilling into trees to excavate nesting cavities and to search for food. Specialized characteristics include the shapes of the bill, some jaw muscles, and the tongue; stiffened tail feathers; and the arrangements of the toes and shape of the claws. Contrary to common belief, the zygodactyl arrangement of the toes is not an adaptation for climbing, but for perching. Specializations for climbing and drilling into trees vary considerably among the woodpeckers according to their feeding habits.

The honeyguides of Africa and India are nest parasites; they lay their eggs in the nests of other piciforms, usually barbets. They feed commonly at bee nests, most often consuming the wax combs rather than the insects because they have the digestive enzymes necessary to break down the wax into sugars. At least one species leads mammals, including humans, to beehives. After the beehive has been opened by the mammal to retrieve the honey, the birds feed on any remaining bits of honey and especially the beeswax.

The neotropical toucans appear to have evolved from the New World barbets and have been placed in the same family as these barbets. However, the distinctive features of the toucans argue for keeping them in a distinct family.

Economic significance. Although woodpeckers consume insects that damage trees, it is still not clear whether they actually save trees. Several species of woodpeckers have become extinct or seriously endangered because of habitat destruction, and woodpeckers have become relatively uncommon in Europe because of intensive forest management. The ivory-billed woodpecker (*Campephilus principalis*) of North America, which had been considered to be extinct, was possibly rediscovered in 2004 in large wooded swamps of Arkansas. A few woodpeckers are considered pests because they drill into wooden poles; also, they have begun to damage the polystyrene plastic insulation on the outside of houses in central Europe. Sapsuckers (*Sphyrapicus*) can kill trees by drilling large numbers of holes in the trunk and major branches for collecting the sap, on which they feed. *See* AVES; ENDANGERED SPECIES; FOREST MANAGEMENT. Walter J. Bock

Bibliography. J. del Hoyo et al. (eds.), Order Galbuliformes, *Handbook of Birds of the World*, vol. 9, pp. 74–138, Lynx Edicions, 2002; J. del Hoyo et al. (eds.), Order Piciformes, *Handbook of Birds of the World*, vol. 9, pp. 140–555, Lynx Edicions, 2002; L. L. Short, *Woodpeckers of the World*, Delaware Museum of Natural History, 1982; L. L. Short and J. Horne, *Toucans, Barbets and Honeyguides*, Oxford University Press, 2001; H. Winkler, D. A. Christie, and D. Nurney, *Woodpeckers: A Guide to the Woodpeckers, Piculets and Wrynecks of the World*, Pica Press, 1995.

Picornaviridae

A viral family made up of the small (18–30 nanometers) ether-sensitive viruses that lack an envelope and have a ribonucleic acid (RNA) genome. The name is derived from "pico" meaning very small, and RNA for the nucleic acid type. Most of the picornaviruses are stabilized by magnesium chloride against thermal inactivation. The virion is made up of a nucleic acid core surrounded by a capsid of 32 subunits (capsomeres) arranged in the icosahedral form of cubic symmetry. Within the infected cell, virus particles are assembled in the cytoplasm, where they tend to aggregate in crystalline array.

Picornaviruses of human origin include the following genera: *Enterovirus* (polioviruses, 3 types; coxsackieviruses A, 24 types; coxsackieviruses B, 6 types; echoviruses, 33 types) and *Rhinovirus* (more than 70 types). There are also 2 genera of picornaviruses of lower animals: *Aphtovirus* (bovine foot-and-mouth disease) and *Cardiovirus*. Certain plant viruses have characteristics similar to those of the picornaviruses; some RNA-containing

bacteriophages also have similar properties. *See* COXSACKIEVIRUS; ECHOVIRUS; ENTEROVIRUS; FOOT-AND-MOUTH DISEASE; POLIOMYELITIS; RHINOVIRUS.

Many picornaviruses cause diseases in humans ranging from severe paralysis to aseptic meningitis, pleurodynia, myocarditis, skin rashes, and common colds. However, subclinical (silent) infection is far more common than clinically manifest disease. Different viruses may produce the same syndrome; on the other hand, the same virus may cause more than a single syndrome.

The host range of the picornaviruses varies greatly from one type to the next, and even among strains of the same type. They may readily be induced by laboratory manipulation to yield variants which have host ranges and target organs different from those of certain of the naturally occurring "wild" strains; this variation has led to the development of attenuated poliovaccine strains.

The two major subgroups of picornaviruses, the enteroviruses and the rhinoviruses, differ in density of the virus particle; the site commonly inhabited in the human host; acid lability; and cell susceptibility and other conditions for optimal growth in tissue culture. The enteroviruses have a density of 1.34 g/ml, and the molecular weight of their RNA core is about 2×10^6; in contrast, the density of rhinoviruses is 1.38–1.40 g/ml but the RNA core still has a molecular weight of about 2×10^6. The enteroviruses multiply in the alimentary tract and are isolated chiefly from feces; the rhinoviruses are almost always isolated from the nose or throat and seldom if ever from feces. While enteroviruses are stable under acid conditions (pH 3–5), the rhinoviruses are inactivated by 1–3 h at this pH level.

Certain of the enteroviruses (coxsackieviruses of the A group) cannot be grown in tissue cultures, and their cultivation in the laboratory must be carried out in mice. Of those enteroviruses which will grow and produce cytopathic effects in cultured cells, most grow preferentially in primary cultures of human and monkey kidney cells (or in certain continuous human heteroploid cell lines). In contrast, rhinoviruses are more readily isolated in embryonic human kidney or human diploid cell strains than in monkey cells; most strains can be recovered only in cells of human origin. *See* ANIMAL VIRUS; TISSUE CULTURE; VIRUS CLASSIFICATION. Joseph L. Melnick

Bibliography. R. W. Compans et al. (eds.), *Current Topics in Microbiology and Immunology*, vol. 161: *Picornaviruses*, 1990; S. A. Evans (ed.), *Viral Infections of Humans: Epidemiology and Control*, 4th ed., 1997; H. H. Malherbe, *Viral Cytopathology*, 1980; V. M. Zhdanov and D. K. Lvov, *Etiology and Pathogenesis of Viral Infections*, 1989.

Picrite

The term picrite has been used with several different meanings. It is generally considered to include certain medium- to fine-grained ultramafic igneous rocks composed chiefly of olivine with smaller amounts of pyroxene, hornblende, and plagioclase feldspar (labradorite).

Its feldspar content is slightly higher than that of peridotite and lower than that of gabbro. Certain analcite-bearing types, associated with teschenite, have also been included under the term picrite. A characteristic feature is poikilitic texture in which large pyroxene or hornblende crystals enclose numerous small grains of olivine. *See* GABBRO; PERIDOTITE.

Picrite is rare and is found in small intrusives (sills and dikes) at or very near the surface. It may also occur in the lower portions of basaltic lava flows where olivine and pyroxene crystals have accumulated under the influence of gravity. *See* IGNEOUS ROCKS. Carleton A. Chapman

Bibliography. H. Sigurdsson, *Encyclopedia of Volcanoes*, 1999; W. Schumann, *Handbook of Rocks, Minerals, and Gemstones*, 1993.

Pictorial drawing

A view of an object (actual or imagined) as it would be seen by an observer who looks at the object either in a chosen direction or from a selected point of view. One such view often suffices to give the reader a clear picture of the shape and details of the object. Pictorial sketches often are more readily made and more clearly understood than are front, top, and side views of an object. Pictorial drawings, either sketched freehand or made with drawing instruments, are frequently used by engineers and architects to convey ideas to their assistants and clients. *See* DESCRIPTIVE GEOMETRY; ENGINEERING DRAWING.

In making a pictorial drawing, it is important to select the viewing direction that shows the object and its details to the best advantage. The resultant drawing is orthographic if the viewing rays are considered as parallel, or perspective if the rays are considered as meeting at the eye of the observer. Making perspective drawings with instruments is time-consuming and requires considerable knowledge and skill. There are, however, commercially available devices which make this chore easier and quicker than is the case when conventional instruments are used. Perspective drawings provide the most realistic, and usually the most pleasing, likeness when compared with other types of pictorial views (**Fig. 1**).

Several types of nonperspective pictorial views can be sketched, or drawn with instruments. Although each type has some distortion, all provide a good picture of what the object looks like. They are easier and quicker to make than perspective drawings. The isometric pictorial is especially popular because of the direction of its axes and the fact that all measurements along these axes are made with one scale (**Fig. 2**). In addition to isometric representation, two other forms—dimetric and trimetric—are sometimes used.

Oblique pictorial drawings, while not true orthographic views, offer a convenient method for

Fig. 1. Perspective drawing of a residence. (*Home Planners Inc.*)

drawing circles and other curves in their true shape (**Fig. 3**). In order to reduce the distortion in an oblique drawing, measurements along the receding axis may be foreshortened. When they are halved, the method is called cabinet drawing.

An effective freehand sketch of an object can be made if proper attention is given to viewing direction, proportions, orientation of ellipses, and location of tangent points. Shading is sometimes added to enhance the pictorial drawing.

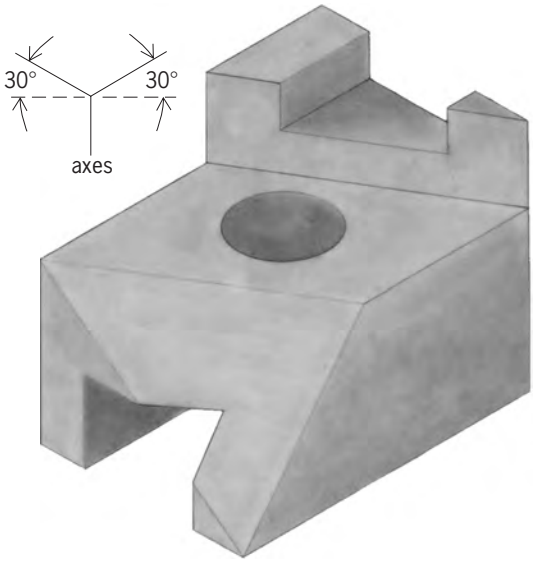

Fig. 2. Isometric drawing; measurements along each axis are made with the same scale.

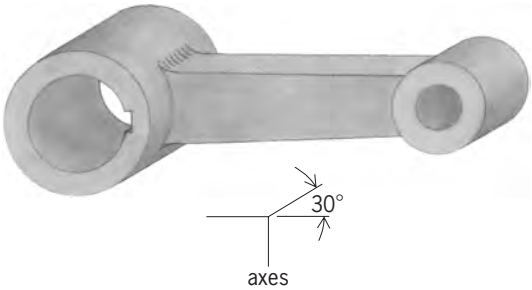

Fig. 3. Oblique pictorial drawing.

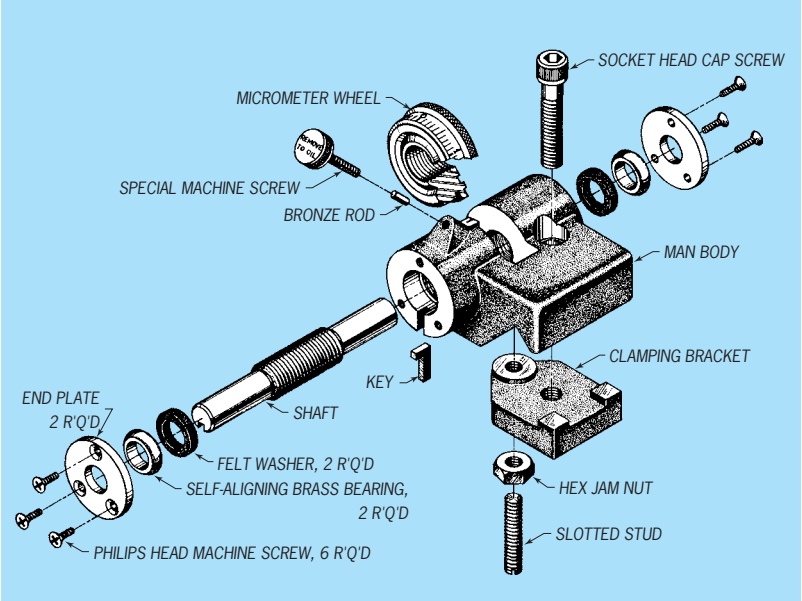

Fig. 4. Exploded-view production illustration; this drawing is isometric.

Shaded exploded-view production illustrations greatly facilitate the learning process in assembly of machines and devices (**Fig. 4**). When this type of illustration is used, the initial assembly of parts into a machine has been found to be three or four times faster than if a conventional assembly drawing is used. Photodrawings can be used to achieve the same visual results. Charles J. Baer

Bibliography. T. E. French et al., *Mechanical Drawing*, 11th ed., 1989; T. E. French, C. J. Vierk, and R. J. Foster, *Engineering Drawing and Graphic Technology*, 14th ed., 1993; C. H. Jensen and J. D. Helsel,

Engineering Drawing and Design, 5th ed., 1997; D. A. Madsen et al., *Fundamentals of Drafting Technology*, rev. ed., 1994.

Picture tube

A cathode-ray tube, or CRT, used as a television display. A cathode ray tube has an electron gun, electromagnetic deflection yoke, shadow mask or grill assembly, and a phosphor screen that is the light emitter (see **illustration**). All of this is contained in a glass envelope. Television picture tubes are largely being replaced by flat-panel displays, such as liquid-crystal displays. *See* CATHODE-RAY TUBE; ELECTRONIC DISPLAY.

Operation. An amplified waveform from the input video signal is sent to the cathode in the gun. As the gun is heated, it emits electrons at a current density in relation to the input signal. Grids in the gun shape and accelerate the beam toward the deflection yoke area, while magnetic fields generated by the yoke move the beam over the CRT face as a raster scan. The beam passes through the shadow mask and lands on the appropriate phosphor, causing it to emit the desired color. *See* TELEVISION.

Glass envelope. The CRT is made of several parts, all of which are mounted inside a glass envelope, sometimes also called the bulb. Normally this is made up of a funnel piece and face panel. The separate face panel, made of optical glass, allows for the coating of the phosphors and the mounting of the shadow mask. The funnel is made of lower quality glass coated with a graphite layer to control eddy currents. The two glass pieces are sealed together by a special frit to provide a vacuum-tight seal. The entire assembly is evacuated of atmospheric gases, and small amount of material, called a getter, is used to absorb the residual gas molecules inside the envelope. *See* EDDY CURRENT.

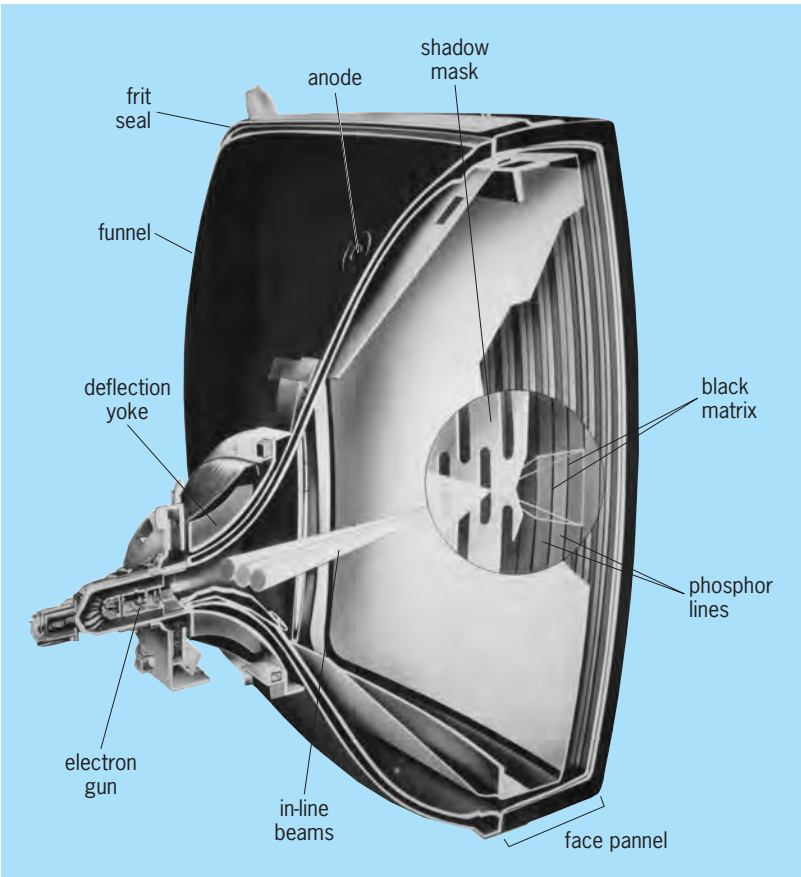

Components of a color picture tube.

Electron gun. The electron gun is made of a heater, a cathode, and plates, called grids, which accelerate the electron beam and focus it onto the phosphor.

There are three cathodes used in a color CRT, one for each phosphor color. These are coated with a special oxide to help increase electron emissions. A heater inserted into the base of the cathode at 430°C (800°F) further improves the electron emission so that current image content is fed from the display signal processing and is amplified to around −160 V and applied to the cathode. A scattered beam is emitted from the cathode face and is pulled toward the first grid "G1" which is at ground potential. Grid 2 is used to shape the beam using small holes. Grids 3, 4, and 5 build a set of static fields in the path of the beam. These provide a focus of the beam astigmatism using properties very similar to those used in lens optics. Grid 6 is set at a very high potential to accelerate the beam toward the face of the CRT and help in the alignment of the beam landing. The anode is then connected to the shadow mask to add a final potential to pull the beam toward the phosphor.

Deflection yoke. In order to display an image on the CRT, a raster scanning method is used. As the signal from the display electronics is decoded, ramp signals are amplified and sent to a horizontal and vertical set of coils. These coils are precisely placed at the part of the expanding area of the funnel and pull or push the electron beam to the appropriate area of the CRT face. Other magnets and coils are also included to help in shaping and aligning the beam.

Shadow mask. Once the beam exits the final grid of the gun, it is guided by the deflection yoke to the appropriate part of the CRT face. Since the beam is most always landing on the face at an angle, there must be control at the face to prevent the beam from landing on the wrong color phosphor. This is accomplished by placing a mask in front of the phosphor to block part of the beam. There are several different designs, but the most common use holes or slits.

Phosphor. All the previous processes have been electronic. However, there must be a means of converting the electronic signals to light. The process to make this conversion is an electrochemical process, whereby a phosphor material lights when an electron beam energizes it. The phosphors are applied to the face panel as a screen of lines or dots, on top of which a film (mirror) of aluminum is vacuum-evaporated. The reflection of light by the aluminum mirror increases the picture brightness and improves picture contrast by preventing stray light from illuminating the back of the phosphor screen. For color

CRTs, three different phosphors are used to generate red, green, and blue light. When these three primary colors are viewed, they are perceived as many colors depending on the balance between the primary colors. *See* PHOSPHORESCENCE. Gary Mandle

Bibliography. S. Sherr, *Electronic Displays*, 1993; J. Whitaker, *Electronic Displays: Technology, Design, and Applications*, 1994; J. Whitaker, *Standard Handbook of Video and Television Engineering*, 4th ed., 2003.

Piezoelectricity

Electricity, or electric polarity, resulting from the application of mechanical pressure on a dielectric crystal. The application of a mechanical stress produces in certain dielectric (electrically nonconducting) crystals an electric polarization (electric dipole moment per cubic meter) which is proportional to this stress. If the crystal is isolated, this polarization manifests itself as a voltage across the crystal, and if the crystal is short-circuited, a flow of charge can be observed during loading. Conversely, application of a voltage between certain faces of the crystal produces a mechanical distortion of the material. This reciprocal relationship is referred to as the piezoelectric effect. The phenomenon of generation of a voltage under mechanical stress is referred to as the direct piezoelectric effect, and the mechanical strain produced in the crystal under electric stress is called the converse piezoelectric effect. *See* POLARIZATION OF DIELECTRICS.

Piezoelectric materials are used extensively in transducers for converting a mechanical strain into an electrical signal. Such devices include microphones, phonograph pickups, vibration-sensing elements, and the like. The converse effect, in which a mechanical output is derived from an electrical signal input, is also widely used in such devices as sonic and ultrasonic transducers, headphones, loudspeakers, and cutting heads for disk recording. Both the direct and converse effects are employed in devices in which the mechanical resonance frequency of the crystal is of importance. Such devices include electric wave filters and frequency-control elements in electronic oscillator circuits. *See* MICROPHONE; ULTRASONICS.

Necessary condition. The necessary condition for the piezoelectric effect is the absence of a center of symmetry in the crystal structure. Of the 32 crystal classes, 21 lack a center of symmetry, and with the exception of one class, all of these are piezoelectric. In the crystal class of lowest symmetry, and type of stress generates an electric polarization, whereas in crystals of higher symmetry, only particular types of stress can produce a pieozoelectric polarization. For a given crystal, the axis of polarization depends upon the type of the stress. There is no crystal class in which the piezoelectric polarization is confined to a single axis. In several crystal classes, however, it is confined to a plane. Hydrostatic pressure produces a piezoelectric polarization in the crystals of those 10 classes that show pyroelectricity in addition to piezoelectricity. The pyroelectric axis is then the axis of polarization. *See* CRYSTALLOGRAPHY; PYROELECTRICITY.

The converse piezoelectric effect is a thermodynamic consequence of the direct piezoelectric effect. When a polarization P is induced in a piezoelectric crystal by an externally applied electric field E, the crystal suffers a small strain S which is proportional to the polarization P. In crystals with a normal dielectric behavior, the polarization P is proportional to the electric field E, and hence the strain is proportional to this field E. Superposed upon the piezoelectric strain S is a much smaller strain which is proportional to P^2 (or E^2). This strain is called the electrostrictive strain. It is present in any dielectric. *See* ELECTROSTRICTION.

Matrix formulation. The relation of the six components T_j of the stress tensor (three compressional components and three shear components) to the three components P_i of the polarization vector can be described by a scheme (matrix) of 18 piezoelectric moduli D_{ij}. The same scheme (d_{ij}) also relates the three components E_i of the electric field to the six components S_j of the strain:

		Compression			Shear		
		S_1	S_2	S_3	S_4	S_5	S_6
		T_1	T_2	T_3	T_4	T_5	T_6
E_1	P_1	d_{11}	d_{12}	d_{13}	d_{14}	d_{15}	d_{16}
E_2	P_2	d_{21}	d_{22}	d_{23}	d_{24}	d_{25}	d_{26}
E_3	P_3	d_{31}	d_{32}	d_{33}	d_{34}	d_{35}	d_{36}

The direct effect is obtained by reading this scheme in rows, as in Eq. (1). The converse effect is obtained by reading it in columns, as in Eq. (2).

$$P_i = -\sum_{j=1}^{6} d_{ij}T_j \qquad i = 1, 2, 3 \tag{1}$$

$$S_j = \sum_{i=1}^{3} d_{ij}E_i \qquad j = 1, 2, \ldots, 6 \tag{2}$$

An analogous matrix (e_{ij}) relates the strain to the polarization and the electric field to the stress, as in Eqs. (3).

$$\begin{aligned} P_i &= \sum_{j=1}^{6} e_{ij}S_j \qquad i = 1, 2, 3 \\ T_j &= -\sum_{i=1}^{3} e_{ij}E_j \qquad j = 1, 2, \ldots, 6 \end{aligned} \tag{3}$$

The matrices (d_{ij}) and (e_{ij}) are not independent, but are related by expressions involving the elasticity tensor$c_{jb}{}^E$ (for constant electric field E), as in Eq. (4).

$$e_{mb} = \sum_{j=1}^{6} d_{mj}c_{jb}^E \qquad \begin{aligned} m &= 1, 2, 3 \\ b &= 1, 2, \ldots, 6 \end{aligned} \tag{4}$$

Alternative formulations can be made by introducing the dielectric displacement D or visualizing the simultaneous action of electrical and mechanical stresses. *See* ELASTICITY.

The number of independent matrix elements d_{ij} or e_{ij} depends upon the symmetry elements of the crystal. For the lowest symmetry, all 18 matrix elements are independent, whereas piezoelectric classes of higher symmetry can have as few as one independent element in the matrix (d_{ij}). The matrix takes its simplest form if the natural symmetry axes of the crystal are chosen for the coordinate system.

Electromechanical coupling. The direct piezoelectric effect makes a crystal a generator, and the converse effect makes it a motor. Consequently, a piezoelectric crystal has many properties in common with a motor-generator. For example, the electrical properties, such as the dielectric constant, depend upon the mechanical load; conversely, the mechanical properties, such as the elastic constants, depend upon the electric boundary conditions. The electromechanical coupling factor k can be defined as follows. Suppose electrodes are attached to a piezoelectric crystal and connected to a battery. Then the ratio of the energy stored in mechanical form to the electrical energy delivered by the battery is equal to k^2. In general, k ranges from below 1 to about 30%. In quartz, for example, the coupling is roughly 10%. In ferroelectric crystals, k can approach unity in certain circumstances. *See* FERROELECTRICS.

In quartz, a stress of 1 newton/m applied along the diad axis produces a polarization of about 2×10^{-12} coulomb/m^2 along the same axis. Conversely, an electric field of 10^4 volts/m produces a strain of about 2×10^{-8}. In ferroelectric crystals, such as rochelle salt and potassium dihydrogen phosphate, and in certain antiferroelectrics, such as ammonium dihydrogen phosphate (ADP), these effects can be several orders of magnitude larger.

Molecular theory. Quantitative theories based on the detailed crystal structure are very involved. Qualitatively, however, the piezoelectric effect is readily understood for simple crystal structures. **Figure 1** illustrates this for a particular cubic crystal, zincblende (ZnS). Every Zn ion is positively charged and is located in the center of a regular tetrahedron $ABCD$, the corners of which are the centers of sulfur ions, which are negatively charged. When this system is subjected to a shear stress in the xy plane, the edge AB, for example, is elongated, and the edge CD of the tetrahedron becomes shorter. Consequently, these edges are no longer equivalent, and the Zn ion will be displaced along the z axis, thus giving rise to an electric dipole moment. The dipole moments arising from different octahedrons sum up because they all have the same orientation with respect to the axes x, y, and z.

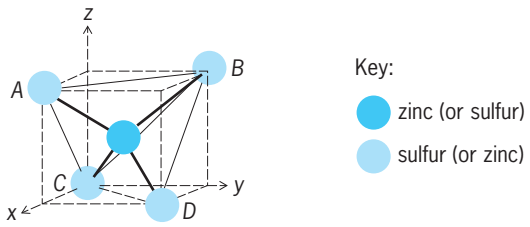

Fig. 1. Tetrahedral structure of zincblende, ZnS. Only part of unit cell is shown. Size of circles has no relation to size of ions.

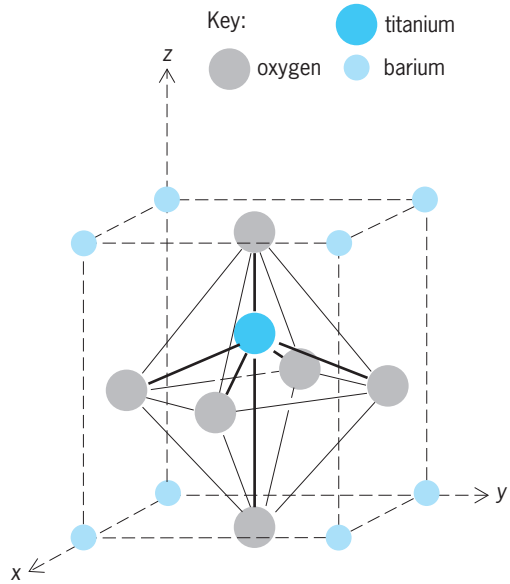

Fig. 2. Unit cell of tetragonal barium titanate, $BaTiO_3$. Deviation from cubic symmetry is exaggerated. Size of the circles has no relation to the size of the ions.

Another simple type of piezoelectric structure is encountered in barium titanate as shown in **Fig. 2**. The positive Ti ions are surrounded by an almost regular octahedron of negative oxygen ions. The Ti ions are not in the center of the octahedron, but somewhat displaced along the z axis. This structure already has a dipole moment or spontaneous polarization in the absence of externally applied stresses. It is clear from Fig. 2 that the Ti ion is pushed more off center when the crystal is mechanically compressed in the xy plane or elongated along z. The additional polarization associated with this deformation is the piezoelectric polarization.

Piezoelectric ceramics. Barium titanate and a few related compounds have the remarkable property that, by means of a sufficiently strong electric field, the direction of the spontaneous polarization can be switched to any one of the x, y, or z axes. This makes it possible to produce polycrystalline samples (ceramics) which are piezoelectric. The electromechanical coupling factors of such ceramics can reach about 50%.

Piezoelectric resonator. The piezoelectric strains that can be induced by a static electric field are very small, except in certain ferroelectrics. Larger strains can be obtained when a piezoelectric crystal is driven by an alternating voltage, the frequency of which is equal to a mechanical resonance frequency of the crystal. The vibrating crystal reacts back on the circuit through the direct piezoelectric effect. In the range of a mechanical resonance, this reaction is

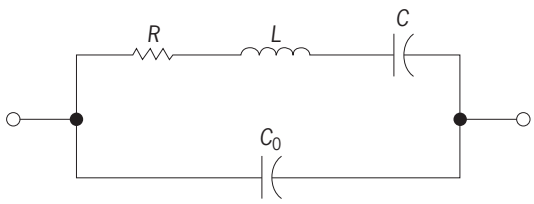

Fig. 3. Network equivalent to a piezoelectric resonator near and at a resonance frequency.

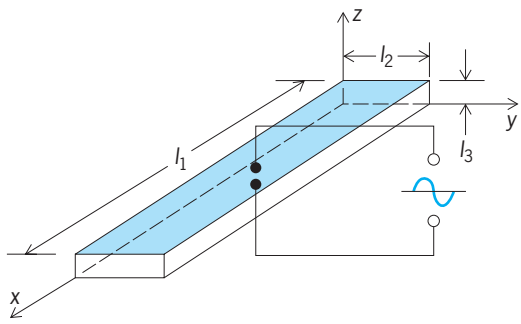

Fig. 4. Simple piezoelectric resonator. A voltage applied to the electrodes shortens or lengthens the bar, thus exciting longitudinal vibrations.

equivalent to the response of the network shown in **Fig. 3**, provided that the series resonance frequency of the network is equal to a mechanical resonance frequency of the crystal, as in Eq. (5). An important

$$f_R = \frac{1}{2\pi\sqrt{LC}} \tag{5}$$

difference between the network of Fig. 3 and the piezoelectric resonator is that the latter has many discrete modes of vibration, whereas the network has only one resonance frequency.

Network elements. The elements L, C, and C_0 of the equivalent network can be calculated from the physical constants of the crystal. Consider, for example, the simple resonator shown in **Fig. 4**. A rectangular crystal bar with the dimensions $l_1 \gg l_2 \gg l_3$ is excited to compressional lengthwise vibrations. The xy faces have adherent electrodes, and the bar is oriented with respect to the natural crystal axes so that an electric field E_3 along z causes a strain S_1 along the bar according to the equation $S_{1(\text{piezoel})} = d_{31}E_3$. A mechanical stress T_1 along the bar causes a strain $S_{1(\text{mech})} = s_{11}^E T_1$, where s_{11}^E is the elastic compliance measured at constant electric field E_3. The resonance frequency for the fundamental lengthwise compressional mode is then given by Eq. (6), where ρ is the

$$f_R = \frac{1}{2l_1\sqrt{\rho s_{11}^E}} \quad \text{Hz} \tag{6}$$

density of the crystal. The parallel capacitance C_0 is the static capacitance of the crystal, as in Eq. (7).

$$C_0 = \frac{8.85\epsilon l_1 l_2}{l_3} \quad \text{picofarads} \tag{7}$$

Here ϵ is the relative dielectric constant along z. For C and L, the analysis yields Eqs. (8) and (9). (All physical

$$C = \frac{70.8 d_{31}^2 \int_1 \int_2}{\pi^2 S_{11} E \int_3} \quad \text{picofarads} \tag{8}$$

$$L = \frac{\rho (S_{11}^E)^2 \int_1 \int_3}{8 d_{31}^2 \int_2} \quad \text{henrys} \tag{9}$$

constants are in mks units.) For the nth overtone, C_0 and L remain the same, whereas C must be divided by n^2. The losses (damping) represented by the resistance R in Fig. 3 arise, for example, from ultrasonic radiation, friction in the crystal mount, internal friction in the crystal originating in various imperfections, and dielectric relaxation.

At the mechanical resonance frequency f_R, the alternating current is maximum and is determined by R. At the antiresonant frequency, given by Eq. (10),

$$f_A = \sqrt{(C_0 + C)/LCC_0} \tag{10}$$

the current is minimum. The difference $\Delta f = f_A - f_R$ increases with increasing electromechanical coupling according to relation (11).

$$\Delta f \approx \frac{4k^2}{\pi^2} \tag{11}$$

The reactance depends upon frequency, as shown in **Fig. 5**. For a typical piezoelectric crystal such as quartz, resonating at about 10^5 Hz, the orders of magnitude given by relations (12) are typical for

$$\begin{aligned} L &\approx 10^2 \text{ henrys} \\ C &\approx 0.02 \text{ picofarad} \\ C_0 &\approx 5 \text{ picofarads} \end{aligned} \tag{12}$$

the elements of the equivalent network. The damping resistance R varies from about 10^2 to 10^4 ohms; that is, the Q factors, given by Eq. (13), are in the

$$Q = \frac{1}{R}\sqrt{\frac{L}{C}} \tag{13}$$

range between 10^6 and 10^4, and the resonances are very sharp. These characteristics cannot be achieved with conventional coils and condensors as circuit elements.

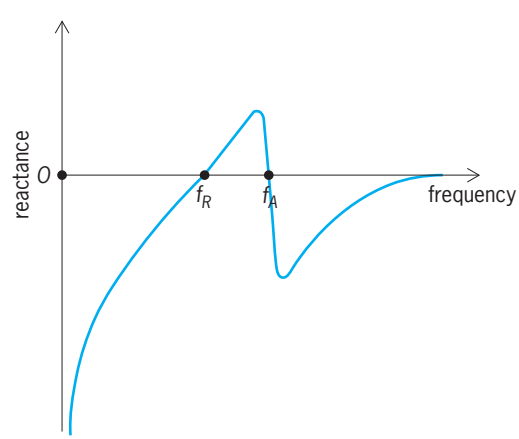

Fig. 5. Diagram showing reactance versus frequency for a piezoelectric resonator.

Vibration modes. With piezoelectric resonators of various types, the range from audio frequencies to many megahertz can be covered. The vibration modes frequently used are, in order of increasing frequency: (1) flexural vibrations of bars and plates, (2) longitudinal vibrations of bars and plates, (3) face shear vibrations of plates, and (4) thickness shear vibrations and compressional vibrations of plates. **Figure 6** illustrates some of these modes. The excitation of particular vibration modes can be achieved by proper orientation of the resonator with respect to the natural crystal axes, by proper positioning of the electrodes, and by proper mounting. A simple example is illustrated by **Fig. 7**. A bar is oriented so that an electric field along x causes an expansion or contraction along y. The electrodes are split and cross-connected so that the bar flexes in the yz plane when a voltage is applied. The fundamental flexure mode is easily excited with this arrangement; however, excitation of higher even-numbered flexural modes is also possible. Interesting resonators are possible with piezoelectric ceramics (barium titanate type) because different parts of the resonator can be polarized in different directions. *See* VIBRATION.

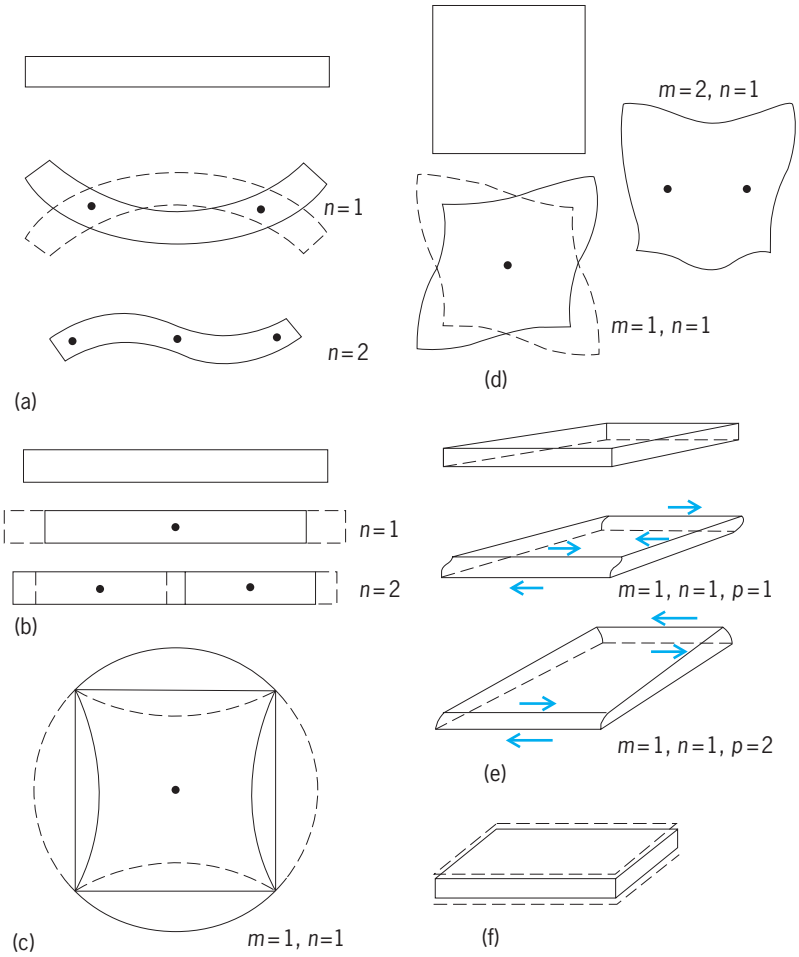

Fig. 6. Diagrammatic representation of examples of vibration modes of bars and plates. (*a*) Flexural vibrations of a bar. (*b*) Longitudinal vibrations of a bar. (*c*) Longitudinal vibration of a plate. (*d*) Face shear vibrations of a plate. (*e*) Thickness shear vibrations of a plate. (*f*) Thickness vibration of a plate.

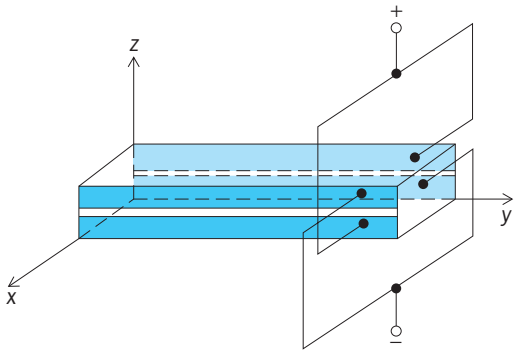

Fig. 7. Excitation of flexure mode by split electrodes.

Common applications. The sharp resonance curve of a piezoelectric resonator makes it useful in the stabilization of the frequency of radio oscillators. Quartz crystals are used almost exclusively in this application. The main advantages of quartz are high Q factor, stability with respect to aging, and the possibility of orienting the resonator with respect to the natural crystal axes so that the temperature coefficient of the resonance frequency vanishes near the operating temperature. **Figure 8** illustrates the orientation of commonly used cuts.

In vacuum-tube oscillators, the crystal generally is part of the feedback circuit. In the circuit proposed by G. W. Pierce, the conditions for oscillation are not satisfied unless the crystal reactance is positive. Hence, the oscillation frequency is between the resonant and antiresonant frequency of the crystal (Fig. 5). Circuits of this type hold the frequency within a few parts per million. Much greater stability can be achieved with the bridge circuit of L. A. Meacham. Here the oscillation conditions are fulfilled by zero phase shift in the feedback circuit, that is, at the exact series resonance frequency of the crystal. Long-term frequency stability of about 1 part in 10^8 and short-term stability of 1 part in 10^9 can be achieved with such oscillators; for an example *see* QUARTZ CLOCK. For detailed information on the Pierce and Meacham circuits *see* OSCILLATOR; QUARTZ CLOCK.

Selective band-pass filters with low losses can be built by using piezoelectric resonators as circuit elements. With a simple network consisting of resonating crystals only, a passband of twice the difference between resonant and antiresonant frequency can be obtained. For quartz resonators, this passband is about 0.8%. At relatively low operating frequencies, this band is too narrow, and combinations of crystal resonators with coils and condensors are generally used. A synthetic piezoelectric crystal which is often substituted for quartz in this application is ethylene diamine tartrate.

Piezoelectric crystals provide the most convenient means for generation and detection of vibrations in gases, liquids, and solids at frequencies above 10^4 Hz. Quartz, ammonium dihydrogen phosphate, rochelle salt, and barium titanate are frequently used in sonic and ultrasonic transducers. The mechanical impedances of liquids and solids

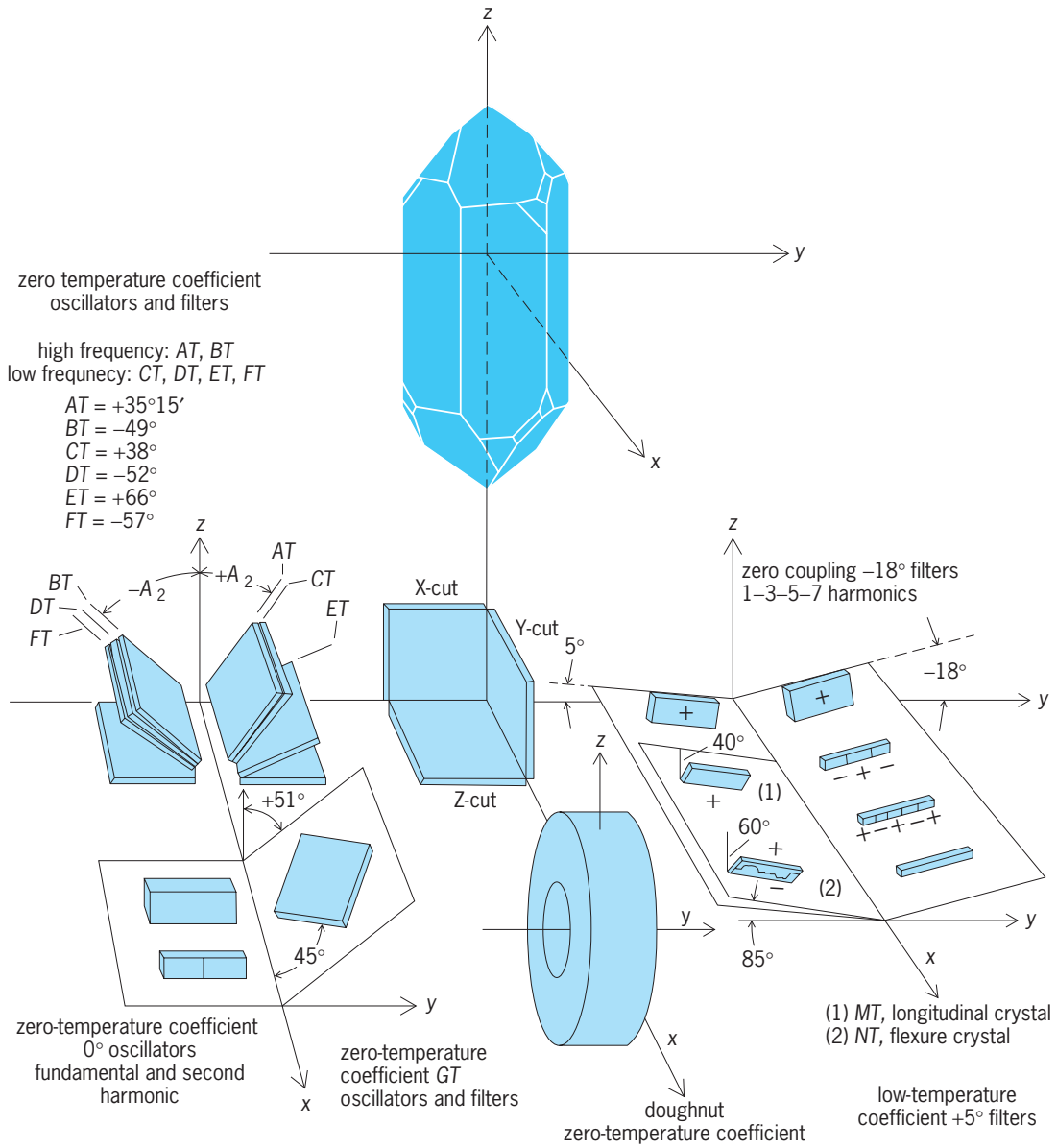

Fig. 8. Orientation with respect to the natural crystal axes of some of the more commonly used special cuts of quartz. (*After W. P. Mason, Piezoelectric Crystals and Their Application to Ultrasonics, Van Nostrand, 1950*)

are generally close enough to the mechanical impedance of the piezoelectric crystal so that efficient energy transfer is possible. The intensity of ultrasonic radiation that can be achieved is mainly limited by the mechanical strength of the piezoelectric crystal. The maximum ultrasonic intensity theoretically obtainable in water by means of quartz or ammonium dihydrogen phosphate is of the order 2000 watts/cm^2 and 200 watts/cm^2, respectively. For gases, the mechanical impedance match is so poor that the corresponding values are about 4000 times smaller. However, the mechanical impedance match can be greatly improved by using piezoelectric devices consisting of two differently oriented crystal cuts cemented together in such a way that a voltage applied to the electrodes causes the elements to deform in opposite directions, and a twisting or bending action results. Assemblies

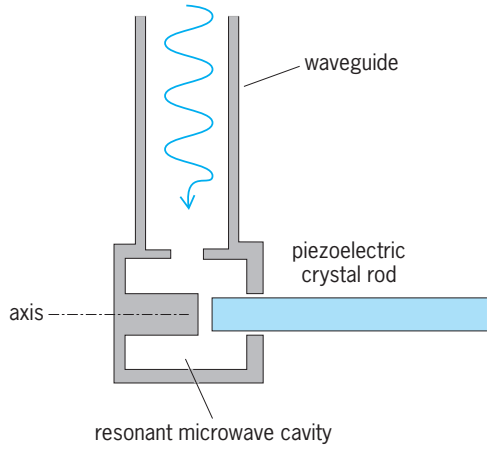

Fig. 9. Diagram showing experimental arrangement for generation of ultrasound at microwave frequencies by means of a piezoelectric crystal.

of this type (bimorphs) with barium titanate ceramics or rochelle salt are widely used in such devices as microphones, earphones, and phonograph pickup cartridges.

Ultrasonic waves at microwave frequencies up to 2.4×10^{10} Hz have been generated by means of the piezoelectric effect. The arrangement is shown in **Fig. 9**. The end surface of a piezoelectric crystal rod is exposed to a strong microwave electric field in a resonant reentrant cavity. The ultrasonic waves travel through the rod in a guided wave mode. The attenuation is low only at very low temperatures.

H. Granicher

Bibliography. T. Ikeda, *Fundamentals of Piezoelectricity*, 1990; W. P. Mason (ed.), *Physical Acoustics*, vol. 1, pt. A, 1964; J. F. Nye, *Physical Properties of Crystals: Their Representation by Tensors and Matrices*, 1957, paper 1987; C. Rosen, B. Hiremath, and R. Newnham (eds.), *Piezoelectricity*, 1992; G. W. Taylor et al. (eds.), *Piezoelectricity*, 1985; J. Zelenka, *Piezoelectric Resonators and Their Applications*, 1986.

Pigeonite

The name given to the monoclinic pyroxenes of the general formula $(Mg,Fe)SiO_3$ with some augite in solid solution. Pigeonite bears the same relation to the orthorhombic pyroxenes as augite does to the diopside-hedenbergite series. Pigeonite is the orthorhombic pyroxene equivalent in the volcanic rocks. Most high-temperature metamorphic and igneous orthorhombic pyroxenes were probably originally pigeonite. The small optic angle (2V) distinguishes the mineral from augite and the inclined extinction distinguishes it from the orthorhombic pyroxenes. Pigeonite forms black, brown, or dark-green short stubby crystals with the 87° pyroxene (110) cleavages. The slower cooling rates of the igneous and metamorphic rocks usually permit the augitic materials in solution to exsolve and the remaining monoclinic pyroxene to invert to the orthorhombic form. The original augitic material is evident by the oriented exsolution lamellae in the host orthorhombic pyroxene. The faster cooling rates of the volcanic rocks quenches in the augitic material and thereby preserves the metastable pigeonite at surface temperatures. *See* AUGITE; DIOPSIDE; ENSTATITE; ORTHORHOMBIC PYROXENE; PYROXENE.

George W. De Vore

Pigment (material)

A finely divided material which contributes to optical and other properties of paint, finishes, and coatings. Pigments are insoluble in the coating material, whereas dyes dissolve in and color the coating. Pigments are mechanically mixed with the coating and are deposited when the coating dries. Their physical properties generally are not changed by incorporation in and deposition from the vehicle. Pigments may be classified according to composition (inorganic or organic) or by source (natural or synthetic). However, the most useful classification is by color (white, transparent, or colored) and by function.

White pigments. These pigments are essentially transparent to visible light. Because of the difference in refractive index between the pigment particles and the vehicles, white pigments refract the light from a multitude of surfaces and return a substantial portion in the direction of illumination without significant change in the spectral composition of the light. *See* COLORIMETRY.

The common white pigments are titanium dioxide, derived from titanium ores; white lead, from corrosion of metallic lead; zinc oxide, from burning of zinc metal; and lithopone, a mixture of zinc sulfide and barium sulfate. Pure zinc sulfide and antimony oxide are less commonly used.

Titanium dioxide may be crystallized in the rutile or anatase form, depending on the method of production. It may be further modified by surface treatment to control the rate of chalking and other properties. Rutile titanium dioxide has a higher refractive index than anatase and therefore higher hiding power, but it has a somewhat yellow color. Anatase titanium dioxide provides a purer white.

White lead pigments are the oldest of white pigments and were used extensively to provide excellent hiding power, flexibility, and durability to interior and exterior paints and enamels. Consumer protection rulings have all but removed white lead paints from the market.

Zinc oxide and lithopone pigments were extensively used in paint formulation, but have been superseded by titanium dioxide. Pure zinc oxide pigment is rarely used. Antimony oxide pigment is used chiefly in certain fire-retardant paints.

Transparent pigments. The refractive indexes of these pigments are very close to the index of the paint vehicle (about 1.54). They are used to provide bulk, control setting, and contribute to the hardness, durability, and abrasion resistance of the paint film. Because they are commonly used to add bulk to other pigments, they are called extenders. Most transparent pigments are natural minerals reduced to pigment particle size. Among the most commonly used transparent pigments are calcium carbonate (ground limestone, whiting, or chalk), magnesium silicate, bentonite clay, silica, or barites (barium sulfate). Transparent pigments often constitute a substantial portion of a protective coating.

Colored pigments. These pigments are available in a wide variety of colors and properties, depending upon the end use. Several hundred have been used, the following being the most common. *See* COLOR.

Red. Iron oxides, often classified by color, include Indian red, Spanish red, Persian Gulf red, and Venetian red, a mixture of iron oxide and calcium sulfate. Other red pigments include cadmium red (cadmium selenide) and organic reds, which are usually coal tar derivatives either precipitated in pigment

form (toners) or deposited on a transparent pigment (lakes). Organic reds include toluidines and lithols.

Orange. Chrome orange (basic lead chromate), molybdate orange (lead chromate-molybdate), and various organic toners and lakes are the most common orange pigments.

Brown. Browns are nearly always iron oxides, although certain lakes and toners are used for special purposes.

Yellow. These pigments include natural iron oxides such as ocher or sienna, or synthetic iron oxides, which are stronger and brighter, such as chrome yellow (normal lead chromate) and cadmium yellow (cadmium sulfide), and organic toners and lakes such as Hansa yellow and benzidine yellow.

Green. The most important green pigments are chrome green, a mixture of chrome yellow and Prussian blue; chromium oxide, duller but more permanent; phthalocyanine green, an organic pigment containing copper; and various other organic toners or lakes, often precipitated with phosphotungstic or phosphomolybdic acid.

Blue. The blue pigments include Prussian blue (ferric ferrocyanide, sometimes called milori or Chinese blue, depending upon the shade); ultramarine, an inorganic pigment made by fusing soda, sulfur, and other materials under controlled conditions; phthalocyanine blue, an organic pigment containing copper; and numerous organic toners and lakes.

Purple and violet. These are nearly all organic toners or lakes. Manganese phosphate is a very weak, inorganic purple pigment.

Black. The vast majority of black pigments consist of finely divided carbon-carbon black, lampblack, and bone black—usually obtained by allowing a smoky flame to impinge on a cold surface. Black iron oxide and certain organic pigments are used where special properties are required.

Special pigments. Anticorrosive pigments are used to prevent the formation or spread of rust on iron when the metal is exposed by a break in the coating. The most common are red lead, an oxide of lead, and zinc yellow or zinc chromate, a basic chromate of zinc. Other colored chromates are sometimes used. The color of red leads fades rapidly, and the anticorrosive chromates are usually very weak in tinting strength. Metallic lead is sometimes used for anticorrosive paint.

Metallic pigments are small, usually flat particles of metal, prepared for dispersal in coatings. Aluminum is most commonly used because it leafs and forms a smooth, metallic film. The flakes are sometimes colored. Bronze, copper, lead, nickel, stainless steel, and silver appear occasionally. Zinc dust, or powdered zinc, is used more often because of its excellent adhesion to galvanized iron than because of its appearance.

Phosphorescent pigments continue to glow for a period after the exciting light has been removed; these are usually sulfides of zinc and other materials, with small amounts of additives which control the phosphorescent properties. Fluorescent pigments lose luminosity as soon as the exciting light is removed; these pigments may be sulfides, although many organic pigments have this property.

Other specialized pigments include pigments which change color at some predetermined temperature, and pigments which give a pearly appearance.

Coarse materials such as pumice are often added when a nonslippery coating is required. Glass beads give a very high degree of refractivity in the direction of illumination and are often used in center-line paints or for signs where night visibility is required. Intumescent pigments puff up under heat, giving a fire-resistant coating. *See* DYE; PAINT.

C. R. Martinson; C. W. Sisler

Bibliography. W. Freitag et al., *Paints, Coating and Solvents*, 2d ed., 1998; W. Herbst and K. Hunger, *Industrial Organic Pigments: Production, Properties, Applications*, 3d ed., 2003; R. Lambourne and T. A. Strivens (eds.), *Paint and Surface Coating: Theory and Practice*, 1999; H. M. Smith (ed.), *High Performance Pigments*, 2001; A. A. Tracton (ed.), *Coating Materials and Surface Coating*, 2006; Z. W. Wicks, Jr., Frank N. Jones, and S. Peter Pappas, *Organic Coatings: Science and Technology*, 2d ed., 1999.

Pigmentation

A property of biological materials that imparts coloration. Hence, pigmentation determines the quantity and quality of reflected visible light. The characteristics of light returning from living matter are a function of its chemical and physical properties and, therefore, not only are due to pigments proper but can be of structural origin as well.

The term visible light as used here refers to the spectral sensitivity of the human eye. Such a restricted application of the word pigmentation is useful in the description of colors seen by humans but is not appropriate in a wider biological context. For example, flowers that appear colorless may be "pigmented" to bees because the eyes of these insects are sensitive to the near-ultraviolet and can detect the attenuation of incident solar radiation by ultraviolet-absorbing flower pigments.

Structural colors. These colors may be due to reflection, scattering, or interference. Reflection occurs, for example, at the surface of tiny air pockets in white (or whitish) material such as certain hairs, feathers, butterfly wings, or flower petals. Small structural entities like cellular inclusions (biocolloids), especially purine crystals in cells called iridocytes, scatter preferentially the blue portion of the incoming light. The blue color of weakly pigmented eyes is based on such an effect, while in blue bird feathers the scattering is produced by air spaces of colloidal dimensions. Iridescent colors of many bird feathers or insect wings, finally, are the result of interference at successive layers of biological structures, that is, feather barbules or wing scales, respectively.

Role of pigments. Pigments are essential constituents of the living world. Their contribution to

the evolution and maintenance of life, and its manifold expressions, is most evident in the role of chlorophylls and the associated carotenoids of certain bacteria and most plants. These pigments harvest solar light energy for utilization in the photosynthesis of organic material from inorganic precursors. *See* CAROTENOID; CHLOROPHYLL; PHOTOSYNTHESIS.

The outermost structures on the animal skin are pigmented for many reasons, for example, to reduce the animal's visibility against a colored background or to provide optical signals to the other sex or to other species. Conspicuously pigmented flowers attract pollinators, and colored fruits are easily found by animals, which eat them and then disperse the undigested seeds.

The role of pigments in communication depends on the ability of organisms to discriminate between different regions of the solar spectrum. In animals with eyes, this is accomplished by differently colored visual pigments contained in specialized receptor cells. Microorganisms, fungi, and plants also have special pigment systems that permit these organisms to move or grow toward, or away from, light (positive and negative phototaxis and phototropism, respectively). *See* PLANT MOVEMENTS.

Since most organisms are totally dependent on light—at least indirectly—elaborate pigment systems have evolved which tune metabolic and activity patterns to the daily pattern of light and dark, and to the changes in the relative lengths of day and night in the course of a year. The phytochrome of plants and the pigments of the eye or of extraretinal photoreceptor organs of many vertebrates and invertebrates are typical representatives of pigments that correlate biological activity with light-dark cycles (photoperiodism). The pineal gland of higher animals plays a major role in transducing the light signal into regulatory actions of hormones such as melatonin. Simple on-off switches of light-regulated metabolic events, finally, may be provided by colored coenzymes such as flavins and cytochromes. *See* COLOR VISION; CYTOCHROME; PHOTOPERIODISM; PHOTORECEPTION.

In the examples listed above, pigments mediate, in various ways, the beneficial actions of light. Absorbed solar light energy may, however, also have detrimental effects by causing undesirable or even destructive reactions. Pigmentations can, in fact, provide a light-absorbing shield that protects the tissue below from such potentially damaging radiation of the Sun. *See* SKIN.

Chemical classes of pigments. Many pigments, particularly those participating in metabolic reactions, are associated with proteins. Such associations, called holochromes, comprise a colored chromophore bound to an apoprotein. Familiar examples of holochromes are the green pigments of the photosynthetic apparatus, rhodopsin in the retina, hemoglobin of the blood, and the phytochrome of plants which regulates photoperiodic and other light-mediated responses. *See* HEMOGLOBIN; PHYTOCHROME.

All organic biological pigments are polyunsaturated molecules. The main classes are hydrocarbons (polyenes), nitrogen- or oxygen-containing heterocyclic compounds, and substituted or condensed aromatic molecules. Typical representatives of the polyenes are the carotenoids with a long chain of alternating single and double bonds, and the structurally related visual pigment retinal and similar vitamin A derivatives.

Tetrapyrroles are the most common colored nitrogen-containing heterocycles, especially in their cyclic form as porphyrins. Indeed, in this latter form, and complexed with a metal ion such as magnesium or iron, the tetrapyrroles may be the oldest functional pigments in the evolution of the biosphere. As contributors to pigmentation, the chlorophylls (Mg-porphyrins) and the hemes (Fe-porphyrins) are the most important. Metabolic changes, either within the organism or after ingestion of other porphyrin-containing organisms, usually lead to the elimination of the complexed metal ion and often to a conversion of the cyclic arrangement of the four pyrroles to a linear one. Such linear tetrapyrroles are called bilins or bile pigments since they are typical components of the bile, a secretion of the gallbladder. An occlusion of the bile duct leads to an increased level of bile pigments in the blood and causes a yellow pigmentation of the skin (jaundice). Bilins and cyclic degradation products of porphyrins are also found in feces and urine and contribute to the pigmentation of skin, body fluids, and certain organs of many lower and higher animals. Interestingly, bilins (the phycobilins) are also responsible for the unique color of blue-green and red algae, in which they are constitutive pigments of the photosynthetic apparatus. Phytochrome, mentioned earlier, is a bilin as well. *See* BILIRUBIN; JAUNDICE; PHYCOBILIN; PORPHYRIN.

The melanins represent another type of widely distributed nitrogenous pigment. Typical eumelanins are more or less polymeric derivatives of indole formed through oxidative modifications of tyrosine. The simplest representative is the monomeric dopachrome. The lighter-colored phaeomelanins have cysteine residues attached to their indolic constituents and are found in the integuments of some birds and mammals, such as in red and yellow hair. In plants, melanins contribute to the coloration of the sap, and these pigments also occur in blackish or brown fungi and bacteria. But melanins are of greatest significance as skin and hair pigments in invertebrates and vertebrates.

Nitrogen-containing heterocyclic pigments also include the yellow flavins and the pteridines. The latter are best known as the coloring matter of butterfly wings, but they occur in many other insects as well. Pteridines are found also in crustaceans; and in vertebrates such as fishes, amphibians, and birds, pteridines are responsible for the brighter coloration of many irises.

The third class of substances listed above are the derivatives of aromatic hydrocarbons; examples are the quinoid and flavonoid pigments. Brownish quinoid pigments occur in bacteria, plants, fungi, and animals. Since such compounds are, in general, not synthesized by animals, quinoids are most likely

to be found in herbivores, notably insects, which ingest the pigment with their food. The red and purple coloring matter of the shells and spines of sea urchins and starfish are echinochromes, a special type of quinoid pigment.

The flavonoids are aromatic molecules comprising an oxygen-containing heterocycle bound to hydroxylated benzene rings. Flavonoid pigments include the flavones and anthocyanins; they have a widespread distribution in the plant kingdom, giving color to flowers, fruits, leaves, and other plant parts. They have also been isolated from animals such as insects. *See* FLAVONOID.

Pigment disposition. The disposition of the pigments is determined by two criteria: their function and their solubility. Photosynthetic and visual pigments are membrane-bound pigment-proteins, while hemoglobin is dissolved in the cytoplasm of blood cells. Flavonoid pigments are typically bound to carbohydrates and are water-soluble; they are usually dissolved in the agueous matrix of cell vacuoles.

Carotenoids are lipid-soluble pigments and, therefore, are often associated with the lipid bilayer of biological membranes, such as chloroplast membranes, the membranes of chloroplast-derived chromoplasts, and cell membranes (for example, bacteria, fungi, and various animal tissues, including nerve). Alternatively, they are dissolved in fat deposits or dispersed as lipid droplets or lipoproteins (as in egg yolk). Conjugated with proteins, the usually yellow or reddish carotenoids can be modified to attain a blue or purplish coloration, as in certain crustaceans. Carotenes of feathers are probably bound to feather-keratin; in the integuments of insects and crustaceans they may be associated with the polysaccharide chitin; and calcium-carbonate-linked carotenes occur in the skeleton of corals.

Animals are unable to synthesize polyenes. Their carotenoid deposits, therefore, are dependent on the amount of dietary polyenes. The best-known examples are the diet-dependent colors of bird feathers and egg yolks and of the skin of infants not exposed to extended periods of sunlight.

In vertebrates, the synthesis of melanins is largely regulated by the neurohormone melanocyte-stimulating-hormone (MSH) and occurs in organelles called melanosomes of special pigment cells referred to as melanocytes or melanophores. Melanocytes are often dendritic, penetrating between the intercellular spaces of tissues, especially in the skin. Melanosomes may be transferred from the melanocytes to other cells, such as to the keratinocytes of the epidermis.

Pteridine-containing cells are called xanthophores when they are yellow and erythrophores when they are red, and contain their pigments in organelles called pterinosomes. Iridophores are structural pigment cells that contain reflecting crystals in so-called reflecting platelets or iridosomes. The common developmental origin of all pigment cells explains their ability to interconvert, and the existence of mosaic cells with different kinds of pigmentary organelles.

Occurrence. The extent and type of pigmentation of organisms are genetically determined, but pigmentation can often be modified within wide limits by environmental factors. Photosynthetic organisms, whose existence depends on the amount of useful light energy absorbed, tend to become more strongly pigmented when their habitat receives little solar radiation. Selective attenuations of portions of the visible spectrum by other organisms causes red and blue-green algae to increase production of those pigments that are capable of absorbing the still-available light colors (chromatic adaptation). Carotenoids, which serve to protect many tissues from light-induced damage, are often (in the bread mold *Neurospora*, for example) synthesized in response to light. Temporary skin darkening observed during certain diseases and, for example, during the menstrual cycle presumably are the result of melanosome dispersing actions of hormones other than melanocyte-stimulating hormone. Ultraviolet-induced tanning is caused by complex sets of events leading to increased melanin production, melanosome formation, and melanocyte proliferation, and may involve a regulation by cellular messengers like prostaglandins. An uncontrolled proliferation of melanocytes with often abnormal melanosomes underlies the formation of melanomas. *See* RADIATION BIOLOGY.

Control. Neural and hormonal control of pigmentation is common in many animals. In humans, the dilation of subcutaneous blood vessels caused by the autonomic nervous system during a moment of embarrassment leads to blushing. Most spectacular, however, are the changes of skin colorations in certain species of reptiles, amphibians, and invertebrates. Sensations received by ocular or extraretinal light receptors can be transmitted to the melanocytes via hormones. The aforementioned melanocyte-stimulating hormone is involved in the regulation of the extent of dispersal of the melanosomes (skin darkening), occasionally in an antagonistic interplay with another hormone, the melanin-concentrating hormone (MCH). Different neurohormones acting in a corresponding fashion are found in invertebrates. The well-analyzed distribution of pigments in the skin of the Florida chameleon *Anolis carolinensis* gives a clue to its various assumed skin pigmentations. Beneath the epidermis occur xanthophores, which reflect yellow light. Below them are located iridophores with reflecting platelets that give rise to scattered blue light. Even farther down are the melanocytes, with long cellular extensions into the upper layers below the epidermis. An upward migration of dispersed melanosomes to the epidermis makes the skin appear brown. When the melanosomes are concentrated in the cell body of the melanocyte, they absorb most of the light that was transmitted by the xanthophore cells and not lost by scattering at iridocytes above them. The radiation returning from the skin of the anole will then contain mainly scattered blue light plus yellow light from the xanthophore. Such light may be all shades from blue to green,

depending on the relative abundance of iridocytes and xanthophore cells. *See* CHROMATOPHORE; PROTECTIVE COLORATION. Peter H. Homann

Bibliography. J. T. Bagnara (ed.), *Advances in Pigment Cell Research*, 1988; D. L. Fox (ed.), *Animal Biochromes and Structural Colors*, 1976; T. W. Goodwin (ed.), *Plant Pigments*, 1988.

Pike

Any of five species of fishes which compose the family Esocidae in the order Esociformes, known by a variety of names such as pickerel and muskellunge. These fishes are voracious predators with an elongated beaklike snout and sharp teeth. The head is partly scaled and the body is covered with cycloid scales that have deeply scalloped edges. The body is cylindrical and compressed; thus, these fishes are well adapted for rapid movements as they dart after prey. They prey upon each other, as well as other fishes, amphibians, small aquatic birds and mammals, and rats. All species are edible but are considered second-rate game fishes. *See* ESOCIFORMES; OSTEICHTHYES; PREDATOR-PREY INTERACTIONS; SCALE (ZOOLOGY).

Common pike. This species (*Esox lucius*) has a wide distribution and is known as the northern pike in the United States. It is also found in northern and central Europe, as well as in northern Asia. On average the fish weigh about 3 lb (1.4 kg); however, the record with rod and reel is 46 lb (20.7 kg) and the length over 4 ft (1.2 m). The teeth are directed backward and are constantly breaking and regenerating. When this species spawns near the muskellunge, cross-fertilization of the muskellunge eggs occurs, and the result is the striped tiger muskellunge. The female pike on spawning may release from 10,000 to 20,000 eggs which adhere to aquatic plants, are fertilized by one or more males, and hatch in about 2 weeks. The male matures at about the age of 2, and the female at 4. Large pike may be 20–30 years old and are relatively solitary in their habits.

The Amur pike (*Esox reicherti*), also known as the blackspotted pike, is endemic to the Amur River in eastern Asia. Although similar to the common pike, it is somewhat smaller, reaching lengths of 110 cm (3.6 ft), and has dark spots on a light background.

Fig. 1. Muskellunge (*Esox masquinongy*). (*Photo by Noel M. Burkhead; EFISH <http://www.cnr.vt.edu/efish/>*)

Fig. 2. Chain pickerel (*Esox niger*). (*Photo by Ken Hammond/U.S. Department of Agriculture*)

Muskellunge. This fish (*E. masquinongy*) [**Fig. 1**] lives in cooler and deeper waters than the common pike. It is common to the Great Lakes region and upper Mississippi. Its average weight is about 15 lb (7 kg), although the record is 70 lb (32 kg) and length over 5 ft (1.5 m). The muskie is prolific and lays as many as 100,000 eggs, with the fry hatching in about 2 weeks and quickly becoming predacious. Like many fishes at hatching, these fishes have a large yolk sac which provides food for growth during the first weeks. This is absorbed by the muskellunge in about 2 weeks.

Pickerel. The chain pickerel (*E. niger*) [**Fig. 2**] is found in the eastern United States, and averages 2.5 ft (0.75 m) in length and 2 lb (1 kg) in weight. Like its relatives, pickerel is a ferocious species which eats mice, birds, frogs, and crustaceans. This fish has a preference for clean, quiet weed areas of lakes. The smallest of the pikes is the grass pickerel (*E. americanus*) which reaches a length of 1 ft (30 cm); nevertheless it is as predacious as its relatives. Charles B. Curtin

Bibliography. J. F. Craig, *Pike: Biology and Exploitation*, Fish and Fisheries Ser. 19, Chapman & Hall, London, 1996; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; W. Senanan and A. R. Kapuscinski, Genetic relationships among populations of norther pike (*Esox lucius*), *Can. J. Fish. Aquat. Sci.*, 57(2):391–404, 2000.

Pile foundation

A special type of foundation that enables a structure to be supported by a layer of soil found at any depth below the ground surface. A pile foundation comprises two basic structural elements, the pile and the pile cap. A pile cap is a structural base, similar to a spread footing, that supports a structural column, wall, or slab, except that it bears on a single pile or group of piles. A pile can be descibed as a structural stilt hammered into the ground. Each pile carries a portion of the pile cap load and transfers it to the soil in the vicinity of the pile tip, located at the bottom of the pile (**Fig. 1**).

The pile and pile cap configuration has provided the basic design solution to the difficult problem of obtaining deep foundation support below areas where poor soil conditions prevail. Poor soil conditions may be difficult to excavate through, and are incapable of supporting structural loads. They are typically characterized by the presence of a soft, compressible layer of clay, high ground-water levels,

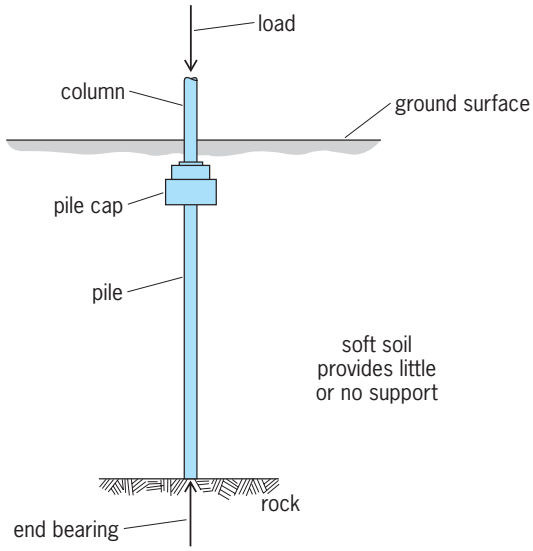

Fig. 1. End-bearing pile.

loosely filled soils, uncontrolled landfills, boulders, abandoned underground structures, and natural bodies of water. By supporting a structure on piles in lieu of spread footings, any adverse soil condition may be virtually bypassed, and adequate foundation support can be obtained at any depth, without the need to perform deep excavation, dewater, and install temporary sheeting and bracing.

While pile foundations offer a simple solution to obtaining deep foundation support through poor soil conditions, they also present major economic concerns. The type of pile, total number and length of piles required, means of pile installation, and size of pile caps are critical factors that may significantly increase foundation costs.

Installation of piles. Piles are commonly driven into the ground by means of an impact (pile) hammer. The pile is hung between a pair of parallel guides or leads suspended vertically from a standard lifting crane. The impact hammer is located above the top of the pile and guided axially by rails incorporated in the leads. To maintain vertical plumbness or to permit driving on a batter, an extendable horizontal member, known as a spotter, is connected from the bottom of the leads to the base of the crane boom.

Most impact hammers are of the steam or diesel type. Steam hammers lift the ram by means of steam pressure and allow the ram to fall by gravity. This is known as a single-acting hammer. If steam pressure is added to the gravitational downward energy, it is called a double-acting hammer. Compressed-air hammers are also used as a substitute for steam.

Diesel hammers lift the ram by explosion of fuel and compressed gas in a chamber between the bottom of the ram and the anvil block. An open-ended hammer is like a single-acting hammer in that it allows the ram to fall by gravity. A closed-end hammer is like a double-acting hammer in that it has a cylinder at the top that allows air to be compressed as the ram rises. The compressed air acts as a spring that limits the upstroke, but returns its stored energy to the ram on the downstroke.

Vibratory hammers are often used when piles are required to be installed to a specific depth, regardless of its load capacity. A vibrating device in the hammer generates a pulsating force in the pile and the surrounding soil, causing the pile to exhibit a controlled up-and-down motion. The weight of the vibratory hammer plus the weight of the pile provide enough downward force to allow the pile tip to penetrate the underlying soil. Vibratory hammers have also proven to be successful pile extractors, provided the upward-pulling capacity of the crane is greater than the weight of the pile hammer.

Determining pile capacity. Piles installed with an impact hammer derive their load capacity from the number of blows per inch of pile penetration. The higher the blow count, the greater the resistance to pile penetration and, therefore, the greater the pile capacity. This oversimplification of the pile-driving phenomenon has resulted in the derivation of various pile-driving formulas that express the relationship between the energy delivered by the hammer to the work done by the pile tip to penetrate a specific distance against soil resistance. Despite the unreliability of pile-driving formulas to accurately determine the load capacity of a driven pile, they have been used for many decades because of their great convenience, and in most cases they are adequate for evaluating low-capacity piles.

A more realistic analysis of the dynamics of pile driving models the driven pile as an elastic bar. A wave equation is used to evaluate the pile's ability to transmit stress waves, imposed by the impact hammer, to the tip of the pile. The capacity of a pile is the net force that develops at the tip of the pile during wave transmission down the length of the pile. The net force at the pile tip must be equivalent to the resistance of the surrounding soil to a downward displacement of the pile.

To verify the actual load capacity of a driven pile, a pile load test is performed. A load is imposed on a driven pile with the use of a jacking frame, a hydraulic jack, and a gage to measure the vertical displacement of the pile under load. The ultimate load capacity of the test pile is equivalent to the maximum load imposed on the pile before failure. Failure of the pile occurs when the anticipated downward displacement of the pile, resulting from elastic shortening of the pile under load, is exceeded. The allowable load assigned to the pile is usually not greater than one-half the load at failure. *See* LOADS, TRANSVERSE.

Development of pile capacity. A driven pile develops its load capacity from the ability of the surrounding soil to develop skin friction along its length and from the ability of the underlying soil to develop point resistance below the pile tip. For these reasons, piles are classified as either friction piles or point-bearing piles, but may actually be a combination of both (**Fig. 2**). The load capacity of a pure friction pile is measured only by the amount of adhesion or friction in the surrounding soil per unit area of pile,

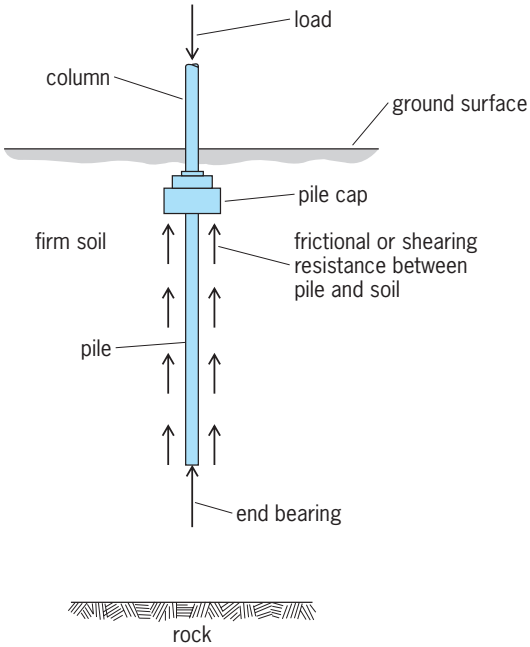

Fig. 2. Friction pile with some support from the end bearing.

multiplied by the actual contact area. The capacity of friction piles is not influenced significantly by the relative difficulty of penetration during driving or by slight differences in the embedded lengths. A pure point-bearing pile, by contrast, depends solely on the adequate contact with or sufficient penetration into a bearing layer to develop its required capacity. Thus, a few more inches of penetration into a firm layer of soil may greatly increase a bearing pile's load capacity.

In reality, most piles are driven through many layers of soils of different types. This would enable a driven pile to transfer load to the surrounding soil through any combination of skin friction from the pressure of sand grains against the pile, an adhesive bond from the clay particles in close contact with a comparatively rough pile surface, and point bearing from the pile tip embedded into a hard bearing stratum.

Precaution must be taken when designing large pile clusters. A stability analysis must be performed to ensure that the total load capacity of the pile group, which equals the sum of the capacity of each pile, does not exceed the bearing capacity of the underlying soil and the frictional resistance of the surrounding soil.

In addition, soft clays that are undergoing consolidation impose a downward force on a driven pile. This downward force tends to drag the pile down as a result of what is known as negative skin friction. This phenonemon imposes additional load on the pile. In effect, negative skin friction reduces the amount of load capacity in the pile available for support of a structure (**Fig. 3**). *See* SOIL MECHANICS.

Types of piles. Piles are available in a variety of sizes, shapes, and materials that enable a particular type of pile foundation to be viable both economically and structurally. For the purpose of identification and discussion, piles are generally classified by their principal materials such as timber, concrete, and steel.

Timber piles. Timber piles, using the trunks of trees with their limbs removed, date as far back as Roman times, and probably still represent the most commonly used type of pile throughout most of the world. Limited by the height of available trees, timber piles 40–60 ft (12–18 m) long are common, with the southern pine being the most common tree type of timber pile used in the United States.

Timber piles are limited to a capacity to no more than 30 tons (27 metric tons) each, because of their inability to be driven into stiff soils of high penetration resistance, without damage. Although damage to the pile tip may be somewhat reduced by using steel point called shoes, the danger of breaking the pile may still exist at the pile butt, the head of the pile that receives the impact from the pile-driving hammer. By limiting the energy rating of the pile hammer and the required number of blows of the hammer, the stress induced in the butt of the pile from the hammer can be significantly reduced.

The unique characteristic of wood is its ability to last indefinitely under water as long as it remains saturated and not allowed to dry. Most timber piles are pressure-treated with creosote to protect the wood above the water table from structural deterioration resulting from dry rot or attack by termites. Treating timber piles with creosote has been known to increase the life of the timber pile in excess of 40 years.

In a marine environment, where timber piles are subject to brackish or salt water and attack by various marine organisms, chemical treatment has not always proven to be an effective means of protection. Alternatively, providing concrete encasement or installing rubber jackets around the timber pile in

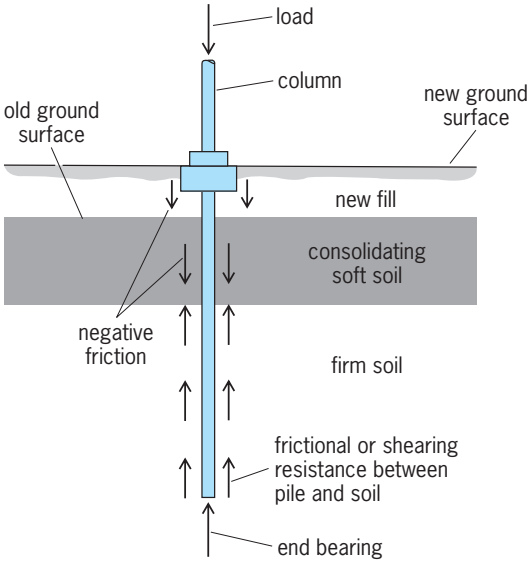

Fig. 3. Pile subjected to negative friction or downdrag. The pile must penetrate a sufficient distance in firm soil to support the external load plus the negative frictional forces.

the area exposed to the marine environment have proven to be a viable means of protection. *See* BORING BIVALVES; SHIPWORM.

Concrete piles. Concrete piles were introduced shortly before 1900. Although a wide variety of concrete piles are available, they are generally divided into two principal categories: cast-in-place and precast piles.

Cast-in-place. Cast-in-place piles entail pouring structural concrete into a hollow shaft of a predetermined size and depth created in the ground. The hollow shaft may be cased or uncased. The casing is usually a thin metal shell or pipe, which acts as a form that is left in the ground. Fundamentally, the casing must be able to resist collapse under lateral pressure from the surrounding soil before it is filled with concrete. Very thin walled shells may not be strong enough to resist damage when driven into the ground and may require the use of a mandrel. A mandrel is a heavy pipe inserted inside the thin shell to provide internal support of the shell and absorb the driving energy from the pile hammer during installation. The mandrel is removed from the thin shell casing prior to casting concrete inside. Very often, pipe casings with a heavier wall thickness are used to eliminate the need for a mandrel, and will in turn contribute to the load capacity of the pile.

Since the casing represents a material cost incorporated into the overall cost of the pile, uncased piles withdraw the casing for reuse as the hole is filled with concrete. Uncased piles are also installed without the use of a casing driven into the ground. This more elaborate method requires drilling a bored hole and filling it with a bentonite slurry to keep the hole from collapsing. Concrete is pumped through a hollow pipe, known as a tremie pipe, that is lowered to the bottom of the hole. As the level of the concrete rises in the bored hole, the tremie pipe is gradually lifted out of the hole to allow the volume of concrete to displace the volume of bentonite slurry. Eventually, all bentonite slurry is replaced by concrete, and the tremie pipe is removed, resulting in a solid column of concrete embedded in the ground. *See* BENTONITE.

The various types of cast-in-place piles differ in the shape, texture, and wall thickness of the casing; the method of casting the concrete in the hole; and the overall pile cost.

Precast. Precast concrete piles are reinforced concrete piles, cast prior to installation in a casting yard. Similar to timber piles, they must be reinforced to resist stresses in the pile due to driving. In addition, precast piles must be reinforced to prevent damage to the pile during handling prior to driving, especially when piles are stored, transported horizontally, and lifted into position. Prestressing of the steel reinforcement in the precast pile tends to reduce tension cracking during handling and driving, and provides an efficient way of withstanding bending stresses.

The primary difficulty in using precast concrete piles is in extending the length of the pile or cutting the excess length off, when there is a variation in the anticipated pile length. However, the advantages of using precast concrete piles are their high load-carrying capacity and their outstanding resistance to deterioration above the water table under ordinary conditions. When exposed to salt water, the precast concrete piles are subject to attack of the steel reinforcement through cracks in the concrete, causing rust to develop and the concrete to spall. Use of high-density concrete and prestressing to minimize tensions cracks serve as the best measures of protection against reinforcing steel corrosion. *See* CONCRETE; PRECAST CONCRETE; REINFORCED CONCRETE.

Steel piles. Steel is the only material used for piling that has a crushing strength comparable to that of hard rock. Available in a wide variety of sizes, steel piles are ideal when conditions call for hard driving, unusually great lengths, or high load capacity. Primarily, steel piles are either steel pipes, usually filled with concrete after being driven, or steel H-sections.

Steel-pipe piles are similar to cased concrete piles, except that the steel-pipe casing is the primary load-carrying component, an therefore it is designed with a heavy wall thickness. When pipe piles are driven open-ended, they must be cleaned out before they are filled with concrete. Often, pipe piles are closed at the bottom with a closure plate or conical point to eliminate the task of cleaning out the pipe and to allow for inspection of the inside of the pile to check for possible damage after driving.

Steel H-sections offer economic advantages not available in other types of piles. Manufactured as a rolled-steel section, the H-section or H-pile can be readily spliced in the field with full-strength welds, and they are available in extremely long lengths. H-piles do not require expensive field fabrication; they can be driven through unconsolidated soils containing timbers, boulders, and other debris with minimum difficulty; and they are designed to carry highly concentrated loads. Protection against marine organisms, termites, dry rot, spalling, and chipping are of no concern when using H-piles. *See* STRUCTURAL STEEL.

Corrosion of steel due to exposure to salt-water spray, wave action, and sand abrasion in waterfront structures is a major concern. The most severe corrosion will occur within the splash zone that extends 2 ft (0.6 m) above and below the water level, where the steel is wetted with a thin film of water saturated with oxygen. By comparison, the area below the water level is oxygen-starved, and the rate of corrosion decreases rapidly with water depth. Similarly, steel piles driven into dry soils or soils that maintain a water table at some depth below the ground surface will not experience significant corrosion, since the amount of oxygen contained below the ground level is negligible.

Cathodic protection is an effective method of preventing corrosion of unprotected steel exposed to seawater. Alternatives involve the use of protective coatings, including flame-sprayed aluminum and a blend of epoxy and polyamide resins. However, any protective coating will reduce corrosion as long as

the coating lasts. Maintaining the coating in the tidal region can be expensive and, in many cases, virtually impossible. Another effective means of protecting the steel pile from corrosion, similar to preventing timber piles from deterioration, is by installing a concrete jacket around the pile extending 2 ft (0.6 m) below low water level and 2 ft (0.6 m) above high tide. *See* CORROSION; METAL COATINGS.

Uses. A pile foundation can be used to support any type of structure through any type of adverse soil conditions. Obviously, pile foundations are used to support marine structures and offshore platforms, since they are located over bodies of water. On land, pile foundations are used primarily in locations where poor soil condition exist. Under extreme conditions, all structures including utility lines are supported on piles and pile caps.

Pile foundations have also been used as a means of underpinning. Bracket piles are used to extend the depth of foundation support for shallow footings adjacent to deep excavations. Piles are driven as close as possible to the edge of an existing footing. Pile caps, designed as cantilevered brackets, are installed at the top of the piles that extend beneath the existing footing. The space between the top of the bracket and the underside of the existing footing is wedged and dry-packed to create a positive load transfer from the existing footing to the bracket pile.

Special care must be taken when driving piles near existing structures that are not supported on pile foundations. Because of the displacement of soil that occurs at the pile tip during pile driving, adjacent spread footings are susceptible to settlement. The magnitude of the soil displacement is a function of the shape and cross-sectional area of the pile tip. Closed-end pipe piles, timber piles, and precast concrete piles yield a relatively greater soil displacement than steel H-piles. If the soil displacement directly below the footing can be kept to a minimum, the magnitude of the settlement will be negligible. Experience has shown that the optimum pile to be driven adjacent to a soil bearing structure is the steel H-pile. Open-end pipe piles, despite their relatively small cross-sectional area, are not effective in minimizing soil displacement because they are circular in shape. As the open-end pipe pile is driven into the ground, the soil at the pile tip displaces radially away from the outside face of the pipe and converges radially inward inside the pipe. After driving the pile a certain distance, the converging soil inside the pipe becomes extremely dense and forms an earth plug at the tip of the pile. Unless the earth plug is cleaned out intermittently during the course of driving the pile, the pile will essentially behave as a closed-end pipe pile.

Piles may also be installed by means of jacking. Jacked piles are used as a means of underpinning when job constraints prohibit the use of a pile-driving hammer. Short sections of open-end pipe are set up vertically below an existing footing and jacked down into the ground with a hydraulic jack installed between the top of the pipe and the bottom of the footing. Using the existing structure as a reaction, the hydraulic jack pushes the pipe downward. The next section of pipe is set on top of the previous pipe section and spliced by welding. The advantages of using an open-end pipe for jacking is the ability to remove an obstruction or excavate the soil inside the pipe to reduce the resistance by the soil to jacking. When the jacking operation is completed, the pile is wedged in place to ensure load transfer from the existing footing to the jacked pile.

Steel-sheet piling. The practice of driving structural members into the ground had led to the development of interlocking steel-sheet piling construction. Steel-sheet piles are rolled-steel shapes with interlocking joints along their edges. They are produced in three standard shapes: straight web, arch web, and Z type; and they are available in a graduated series of weights and strengths. The interlocks, which must be sufficiently loose to permit free sliding during installation, are designed to provide strength and watertightness in both longitudinal and transverse directions.

The Z-pile type has the highest bending resistance because of the favorable distribution of steel area away from the axis of bending. For this reason, Z-piles are most economical in the construction of permanent filled bulkheads, land walls, deeper braced cofferdams, and similar wall types that require high beam strength.

The arch web piling is used for similar but generally lighter applications than Z-piling, such as in the construction of low-head cofferdams and shallow bulkheads or walls. The advantages of using arch web piling instead of Z-piling is its relative ease in handling and installation and its adaptability to a wide variety of field conditions.

The straight web piling offers the least amount of bending strength, but the highest value of tension between the interlocks, which is the prime consideration in the construction of large, filled cellular-type structures. Most cellular structures are used as temporary cofferdams to aid in the construction of locks, dams, bridge piers, and land reclamation projects. Cellular cofferdams are also used in various types of permanent marine construction such as deep-water bulkheads, mooring or turning cells, artificial islands, and breakwaters. *See* COASTAL ENGINEERING; COFFERDAM; FOUNDATIONS. Anthony J. Mazzo

Bibliography. R. W. Brown, *Practical Foundation Engineering Handbook*, 2d ed., 2000; R. B. Peck, W. E. Hanson and T. H. Thornburn, *Foundation Engineering*, 2d ed., 1998.

Pilot production

The production of a product, process, or piece of equipment on a simulated factory basis. In mass-production industries where complicated products, processes, or equipment are being developed, a pilot plant often leads to the presentation of a better product to the customer, lower development and manufacturing costs, more efficient factory operations, and earlier introduction of the product. Following

the engineering development of a product, process, or complicated piece of equipment and its one-of-a-kind fabrication in the model shop, it becomes desirable and necessary to "prove out" the development on a simulated factory basis.

Facilities. To accomplish this aim, the pilot plant, an intermediate step between development laboratory and production factory, is established. It is provided with personnel and facilities that duplicate as nearly as possible actual manufacturing conditions but are free of the day-to-day necessity of meeting delivery schedules.

This requires personnel who are highly skilled in their field but are not necessarily experts. They must understand the thinking and objectives of development engineers and must also appreciate the day-to-day problems that arise in factory operation with unskilled or semiskilled operators.

Adequate equipment and services should be provided to accommodate the full range of factory requirements. Sometimes, however, it may be impractical to include certain facilities in a pilot plant because of floor space limitations, or because of specialized services required. In such cases, arrangements should be made to perform these pilot-plant functions on regular production equipment in the factory but under pilot plant supervision and financial control.

Objectives. Some of the objectives of a pilot plant operation are as follows.

Quality control. Pilot plant operation demonstrates the ability of processes or equipment to function consistently at the desired quality level under factory operating conditions.

Material use. To demonstrate economical use of materials during the pilot operation, the materials should be truly representative and contain the full range of variability permitted by material purchase specifications.

Process reliability. Pilot operation demonstrates whether the processes involved are practical and realistic and whether operating procedures can be followed in their entirety by the type of personnel used in the factory.

Equipment reliability. To demonstrate the capability of the equipment to produce at speeds and machine efficiencies, which will become standard in the factory, data are recorded during pilot production which show whether the equipment will perform at a satisfactory rate of defectives. Pilot plant experience establishes a list of spare parts which will be initially provided in the factor.

Personnel requirements. Pilot experience establishes the amount and labor grade of production and maintenance personnel required to operate the process or equipment. This is done with the cooperation of industrial engineering and industrial relations departments.

Safety. During pilot runs the safety department has an opportunity to review the process or equipment for compliance with safety requirements.

Factory acceptance. Perhaps the most important objective of a pilot plant is to afford representatives of the factory to which the process or equipment will be transferred an opportunity to witness the operation and to agree that it is ready for the factory. Thus, criticism of the development is minimized and its introduction into regular factory operation accomplished more quickly.

Manufacturing costs. Pilot plant operation provides the opportunity to verify engineering estimates of manufacturing cost and may indicate the necessity for pricing changes or for taking further steps to reduce the cost of the product in order to reach a satisfactory level of profit.

Development costs. The cost of designing and developing processes or equipment and of instructing personnel in their use is often only a fraction of the total cost involved in producing a final efficient process or machine. The debugging or working out of unforeseen problems can consume a great deal of time and add a large amount of expense before the process or machine is ready for factory operation. The pilot plant relieves the factory of the expense of a trial run both in direct cost and in lost production of established products. The close proximity of the pilot operation to engineering and other skilled personnel permits first-hand observation of operation and quicker decisions on required changes. It may also save the expense and time of such personnel traveling to factory locations and the return of equipment from the factory to its point of origin.

The financing of a pilot plant should be on a budgetary basis. Funds for its facilities, personnel, and operation should be provided as part of a manufacturing development budget. The pilot plant should not be expected to finance itself through manufactured products; otherwise management tends to keep the project in the pilot plant longer than necessary in order to create funds for its operation. The pilot plant then loses its value as a development function, and the desire and interest of its personnel to introduce improvements and to tackle new projects become subordinate to that of meeting production schedules. On the other hand, arrangements should be made to credit the operation with a fair value of the product manufactured so that the net amount remaining will represent development cost more accurately. *See* ACTIVITY-BASED COSTING; PRODUCT DESIGN; PRODUCTION ENGINEERING; QUALITY CONTROL.

James E. Woodall

Bibliography. D. M. Anderson, *Design for Manufacturability & Concurrent Engineering*, 2003; F. X. McConville, *The Pilot Plant Real Book*, 2002; G. H. Vogel, *Process Development: From the Initial Idea to the Chemical Production Plant*, 2006.

Piloting

The form of navigation in which position is determined relative to external reference points, usually fixed points on the Earth. It is the oldest form of navigation. The following remarks are formulated in the context of marine navigation. However, the use of landmarks to navigate terrestrial vehicles and

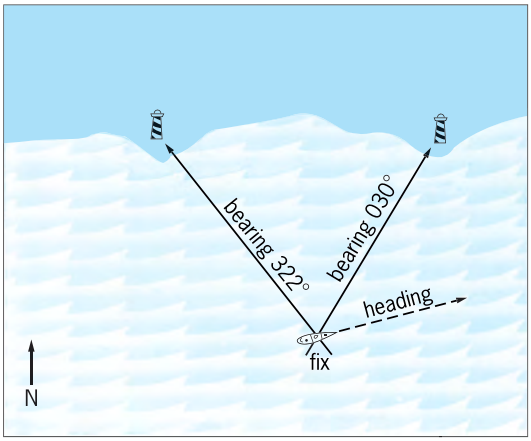

Fig. 1. A fix from two bearings.

aircraft is well established. Some authors refer to this means of navigating aircraft as pilotage.

With the development of electronic aids, piloting techniques have been extended far from shore. However, the term piloting is generally associated with nearness of land, where tidal and other currents may be strong, shoals and other underwater obstructions may be in near proximity, and maneuvering room is limited when other vessels are encountered.

Because of these limitations, piloting requires constant vigilance, frequent fixing of position, and a thorough knowledge of the characteristics and locations of landmarks, both natural and artificial, and local regulations. Thus it is not unusual for ships to employ the services of a local expert, called a pilot, to assist in the navigation of the vessel while it enters or leaves port.

Marks. A conspicuous object, structure, or light that serves as an indicator for establishing the position of a craft or otherwise assisting in its safe navigation is called a mark. To be useful, not only must a mark be identified, but its position must be known accurately.

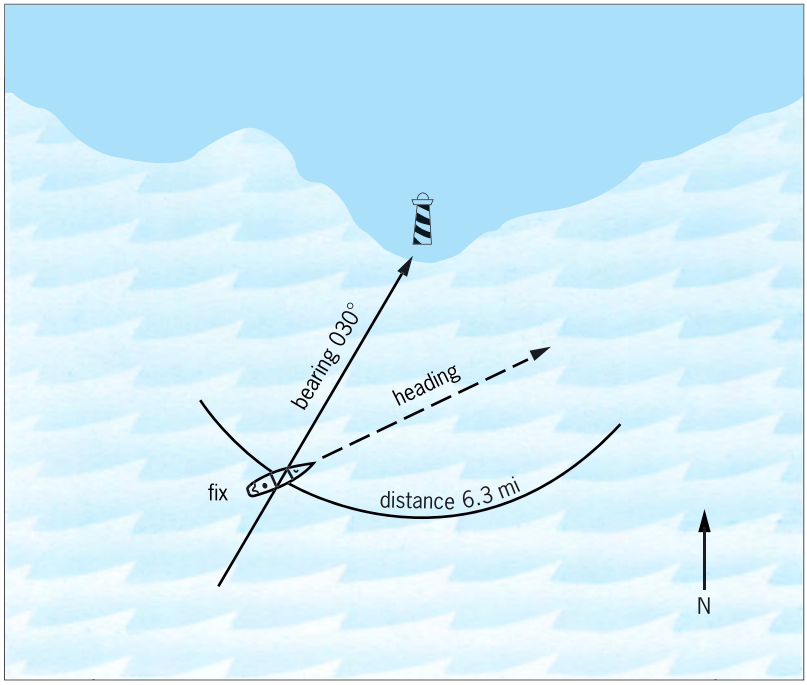

Fig. 2. A fix which is obtained from a bearing and the distance of a single object. 6.3 mi = 10 km.

Natural landmarks are prominent features of the landscape. Artificial marks might be prominent buildings or parts of buildings, towers, flagpoles, tanks, and so forth. Additionally, a number of artificial marks have been designed and erected specifically to assist the navigator. These marks are called aids to navigation, and are of several types.

Beacons, both lighted and unlighted, and lighthouses are at fixed positions on land or in the water. Buoys, both lighted and unlighted, and lightships float in the water, being moored at desired points. Unlighted aids are called daymarks. These are given distinctive shapes or coloring to assist in identification. A light installed as an aid to navigation is given a distinctive sequence and duration of light and dark periods and of color or colors, by which it can be identified. These characteristics are shown on the charts and given in the light lists available to navigators. *See* BUOY; LIGHTHOUSE.

In addition to visible aids to navigation, bottom topography can be of assistance in locating the position of a vessel. Sound signals transmitted through water or air may be used for navigation. Electronic beacons and positioning systems have been established at a number of places to assist in navigation. *See* ELECTRONIC NAVIGATION SYSTEMS.

Measurements. The measurements that are made for piloting purposes are direction, distance, differential distance between two points, and distance to bottom.

Direction. The direction of a landmark is called its bearing, commonly stated as the angular distance from some reference direction, usually true, magnetic, compass, or gyro north. Bearing values are generally given in integral degrees (although half and quarter degrees may be used), using three figures, from 000° at north clockwise through 360°. Relative bearings use the heading of the craft as the reference direction, and may be stated as above, right and left through 180°, or with reference to some part of the craft (as broad on the starboard bow, 10 degrees abaft the port beam, or two points on the port quarter).

Bearings are usually measured (1) by noting when two objects are in range (directly in line); (2) by means of a suitable attachment to a compass or compass repeater; (3) by pelorus, a compasslike instrument without directive properties; or (4) electronically, by radio direction finder, by radar, or by the indication of the receiver-indicator of an electronic system of navigation. *See* DIRECTION-FINDING EQUIPMENT.

Distance. This is now generally measured by radar in nautical miles or yards (meters in countries using the metric system). *See* RADAR.

Differential distance. The difference in distance from the ship to two points is usually measured electronically with the receiver-indicator of a hyperbolic navigation system. The reading is generally in some special unit such as a microsecond or one-hundredth

of a lane width. *See* HYPERBOLIC NAVIGATION SYSTEM.

Depth. Depth of water is measured in meters, feet (1 ft = 0.3 m), or fathoms (1 fathom = 6 ft = 1.8 m). Depth measurement is usually made by an echo sounder, a device that emits a sonic or ultrasonic signal downward in the water and measures the time interval until return of an echo from the bottom. *See* ECHO SOUNDER.

Position determination. Traditionally, position by piloting has been determined by lines of position, each indicating a series of possible positions of the craft at time of measurement. A measured bearing provides a straight line of position (actually part of a great circle) passing through the object sighted. A measured distance provides a circular line of position with the object as the center and the distance as the radius. A measured differential distance provides a hyperbolic line of position. A straight or circular line of position is usually plotted directly on the nautical chart. Depth of water generally does not provide a line of position unless the craft crosses a distinctive feature such as the 100-fathom curve. A position, called a fix, is determined by crossing two or more lines of position taken simultaneously or nearly so (**Figs. 1** and **2**).

If there is an appreciable period of time (more than several seconds) between observations, one of the lines of position is adjusted (advanced in the direction of the travel, or retired in the reciprocal direction) a distance equal to the run of the vessel between observations. The resulting position is called a running fix, which is generally somewhat less reliable than a fix because of uncertainty in the speed and direction of travel of the vessel between observations. However, if the objects observed are nearly in the same or reciprocal direction, the passage of time might result in a more favorable crossing angle (90° being ideal), thus more than offsetting the loss of accuracy because of uncertainty in the motion of the vessel. When only one mark is available for observation, there may be no alternative to a running fix.

Position of a vessel can be obtained in other ways than by crossing lines of position plotted on a chart. Certain combinations of successive bearings of the same mark provide immediate indication of position relative to that mark. As an example, bow and beam bearings (the first when the mark is 45° from the direction of travel of the vessel, and the second when the relative bearing, right or left, is 90°) indicate that the distance of the vessel from the mark at the time of the second bearing is equal to the run between bearings, and the direction is perpendicular to the direction of travel. Simultaneous horizontal angles between three objects can be set on a three-arm protractor (called a station pointer by British navigators) and the position determined by fitting the arms of the instrument to the chart symbols representing the objects. The horizontal angles are usually measured by marine sextant. This provides an accurate position sometimes used in hydrographic surveying or in locating the point of letting go an anchor, but seldom for other purposes. A line of soundings plotted on a transparency at the correct distance intervals at the scale of the chart and then matched by trial and error to the soundings shown on the chart may provide a reliable indication of position where the bottom topography has a distinctive pattern. Under favorable conditions, a bottom profile obtained by a recording echo sounder can provide a better position.

A vessel can sometimes be kept in safe water without a fix. In an area where the bottom shoals gradually, a well-chosen danger sounding can give warning in sufficient time to prevent grounding. A danger bearing of some object well ahead can provide an indication of a dangerous situation if the measured bearing exceeds the danger bearing. An off-lying rock or shoal can sometimes be successfully avoided by keeping the measured angle between two objects less (or more) than a preselected horizontal danger angle (**Fig. 3**). Similarly, a vertical danger angle can be established between the top and bottom of a single object.

Electronic charts. Digitization of a nautical chart provides the information needed for display of an electronic facsimile of a published chart. Selective determination of features to be shown makes possible the tailoring of the facsimile to individual requirements. If a radar image is added to the display, the position of the vessel relative to its surroundings is immediately apparent. A suitable electronic positioning system such as the Global Positioning System (GPS) or differential GPS (DGPS) can be used to display a symbol indicating the position of the vessel, thus providing a check on accuracy of the radar data. Other information can be added as desired. Color can be used to assist in identification of various types of information. The combined

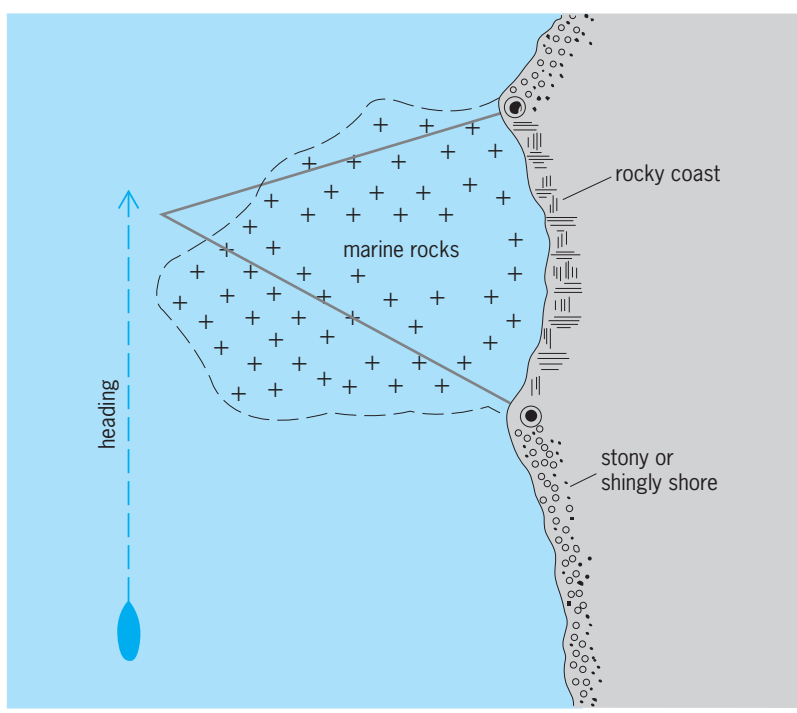

Fig. 3. A horizontal danger angle.

display, called an electronic chart display information system (ECDIS), is similar to that of a vessel traffic system (VTS), thus contributing to effective traffic coordination in a harbor. The electronic chart can be corrected by radio, so a vessel entering a harbor can utilize an up-to-date chart. *See* CELESTIAL NAVIGATION; DEAD RECKONING; MARINE NAVIGATION; NAVIGATION; POLAR NAVIGATION; SATELLITE NAVIGATION SYSTEMS. Alton B Moody

Bibliography. N. Bowditch, *American Practical Navigator*, U.S. Defense Map. Agency Hydrogr. Topogr. Center Publ. 9, Bicentennial Edition, 2002; T. J. Cutler, *Dutton's Nautical Navigation*, 15th ed., 2003; L. Melton, *Piloting with Electronics*, 1987; H. H. Shufeldt and G. D. Dunlap; *Piloting and Dead Reckoning*, 4th ed., 1999; U.S. Coast Guard Auxiliary Staff, *Coastal Piloting*, 1994.

Piltdown man

The Piltdown man hoax was the scientifically most successful fraud in the history of anthropology. Between 1910 and 1915, Charles Dawson, a lawyer and amateur prehistorian, claimed to have recovered craniodental remains of a fossil human ancestor, along with primitive stone tools and mammalian fossils, from a gravel pit at Barkham Manor near Piltdown in Sussex, southern England. The skull fragments included a partial braincase with relatively thick bone, but otherwise the skull was like modern humans in having a relatively large braincase, steep forehead, and poorly developed bony ridge above the eye socket (see **illus.**). The lower canine was apelike in being quite large, but it had a humanlike wear pattern. The fragmentary lower jaw was missing many of the critical parts useful for distinguishing humans from apes, such as the articular condyle (the region of the lower jaw that articulates with the cranium) and the chin region, but it was generally similar to that of great apes. The molar teeth had flat wear resembling humans. Overall, the skull provided an excellent intermediate stage or "missing link" between apes and humans, suggesting that brain expansion had preceded facial specialization during the course of human evolution. The associated fauna, the simple flaked stone tools, and the geology all indicated that the human finds were late Pliocene or early Pleistocene in age (in current terms, a period dating to 3–1 million years ago), older than any fossil human discovered prior to that time.

The Piltdown remains were published in 1912 by A. Smith Woodward, a leading British paleontologist working at the British Museum of Natural History in London. He formally named the remains *Eoanthropus dawsoni* (that is, Dawson's dawn man, in honor of its discoverer). This sensational discovery caused a great deal of excitement among the scientific community and the public at large. Not only was it an important piece of new evidence to help document the earliest stages of human evolution, but until the discovery at Piltdown, England had been a relative backwater for prehistoric finds compared with the wealth of discoveries from continental Europe and Asia. Although most scientists were quick to recognize the significance of the finds, especially given the credentials of the scholarly authorities that endorsed them, others were more skeptical. Some claimed, for example, that the Piltdown remains represented two different species—the cranium belonging to a human and the lower jaw belonging to an ape—that had been accidentally mixed together at the time of deposition of the Piltdown gravels. Other scientists recognized that the anatomy was anomalous, especially the unusual wear on the teeth. These doubts were largely dispelled, however, by the recovery in 1915 of a second individual at Sheffield Park, near Piltdown, that had the same combination of anatomical features as the original finds.

During the 1920s and 1930s, *Eoanthropus* was considered pivotal to understanding human evolution. However, with a growing number of fossil finds from Europe and Asia, and especially from Pliocene sites in Africa, the Piltdown finds became increasingly difficult to reconcile with the rest of the fossil record. The discoveries from Africa showed that humanlike teeth and jaws characterized the earliest

Three views of the reconstructed skull of Piltdown. The original preserved portions are shown as dark areas. The light portions are reconstructed only.

fossils, and that these specializations preceded brain expansion. In fact, when the first fossil human of Pliocene age from Africa was described from Taung in South Africa, the small-brained infant skull of *Australopithecus africanus* was claimed by Piltdown supporters to be that of an ape. By the 1940s and early 1950s, this incongruence was dubbed the "Piltdown enigma" by Arthur Keith, and Piltdown became increasingly ignored by mainstream human paleontologists. Eventually, logic led to the hypothesis of a deliberate hoax.

In 1953, J. S. Weiner (Oxford University), in collaboration with W. E. Le Gros Clark (Oxford University) and K. P. Oakley (British Museum of Natural History), published a detailed account exposing the fraud. Chemical tests showed that none of the human bones could be more than a few centuries old (this was later confirmed in 1959 with the advent of carbon-14 dating). Also, on reexamination, signs of forging were evident. All parts of the fossil human remains, as well as the tools and animal bones, were artificially stained to match the color of the deposits. The cheek teeth had been filed unnaturally flat, not worn by chewing, and the canine had also been reduced in size by filing. The thickened skull bones were from a modern human with Paget's disease (which is characterized by an increase in bone turnover). The mandible could be identified as that of an orangutan. The inevitable conclusion was that all of the evidence was fabricated and that it had been planted at Piltdown as part of a carefully orchestrated and deliberate hoax. *See* ARCHEOLOGICAL CHRONOLOGY.

The identity of the perpetrator of the hoax has been the subject of much speculation, but no firm evidence has been found to indisputably incriminate any particular individual. Few people doubt that Dawson was involved in the deception, but most believe that someone with greater expertise probably assisted him. Anti-evolutionists assert that the Piltdown affair exemplifies how easily scientists can be duped by falsified or misleading evidence, and that it represents a serious challenge to the objectivity and authority ascribed to the scientific method. In fact, to the contrary, Piltdown offers a valuable object lesson in how scientists are able to overthrow long-cherished hypotheses through advances in technology and the accumulation of contradictory evidence. *See* AUSTRALOPITHECUS; FOSSIL HUMAN. T. Harrison; W. W. Howells

Bibliography. C. Blinderman, *The Piltdown Inquest*, Prometheus Books, Amherst, NY, 1986; R. Millar, *The Piltdown Men*, Gollancz, London, 1972; F. Spencer, *Piltdown: A Scientific Forgery*, Oxford University Press, 1990; J. S. Weiner, *The Piltdown Forgery*, Oxford University Press, 1955.

Pimento

A type of pepper, *Capsicum annuum*, grown for its thick, sweet-fleshed red fruit. A member of the plant order Polemoniales, pimento is of American origin, and gets its name from the Spanish word designating all sweet peppers. In the United States, however, the term pimento generally refers to the heart-shaped varieties (cultivars) grown in the South for canning and used for stuffing olives and flavoring foods. Perfection is a popular variety. Harvesting begins when the fruits are fully red, usually $2^1/_2$–3 months after planting. Georgia is the only important pimento-producing state. *See* PEPPER; SOLANALES. H. John Carew

Piña

A fiber, also known as pineapple fiber, obtained from the large leaves of the pineapple plant grown in tropical countries. This natural fiber is white and especially soft and lustrous. In the Philippine Islands, it is woven into piña cloth, which is soft, durable, and resistant to moisture. Piña is also used in making coarse grass cloth and for mats, bags, and clothing. *See* NATURAL FIBER; PINEAPPLE. M. David Potter

Pinales

An order of the class Pinopsida (=Coniferopsida), of the division Pinophyta (Gymnospermae), with about 50 genera and 600 species still living. All are woody plants, as shrubs or trees, and are often the principal trees of the forests worldwide. Pine, spruce, fir, hemlock, cedar, larch, juniper, cypress, yew, redwood, big tree, kauri, podocarpus, araucaria, and others are all part of this order. The big tree (*Sequoia gigantea*) of California is the largest plant, reaching a height of 330 ft (100 m) and a diameter of 33 ft (10 m), and being over 3500 years old. Leaves are usually needlelike or scalelike, with a few exceptions such as *Agathis* (kauri), *Podocarpus*, *Phyllocladus*, and *Araucaria* (**Fig. 1***a*). Most species are evergreen, bearing their leaves year-round. Modern conifers form some of the most extensive forests in recent times, mainly occurring in the temperate regions or mountainous regions of the subtropics. The families of extant conifers are Araucariaceae, Pinaceae, Taxodiaceae, Cupressaceae, Podocarpaceae, Taxaceae, and Cephalotaxaceae. The phylogenetic relations among these families are practically unknown.

Reproduction. The cones are unisexual (Fig. 1*b* and *c*). In some species, both male and female cones are borne on the same plant (monoecious); in others, on separate plants (dioecious). The sperm is nonciliate and nonmotile. In the pine, the pollen grains (microspores) fall from the microsporangia of the staminate (male) cones and are carried by the wind to the pistillate (female) cones, where the microspores filter between the scales of the cone, settling on the megasporangia. At the end of each megasporangium is a tiny pore (micropyle) through which the microspore migrates to the nucellus, the central body of the megasporangium. Here the microspore germinates. Each megasporangium contains a megaspore which also germinates, producing a female

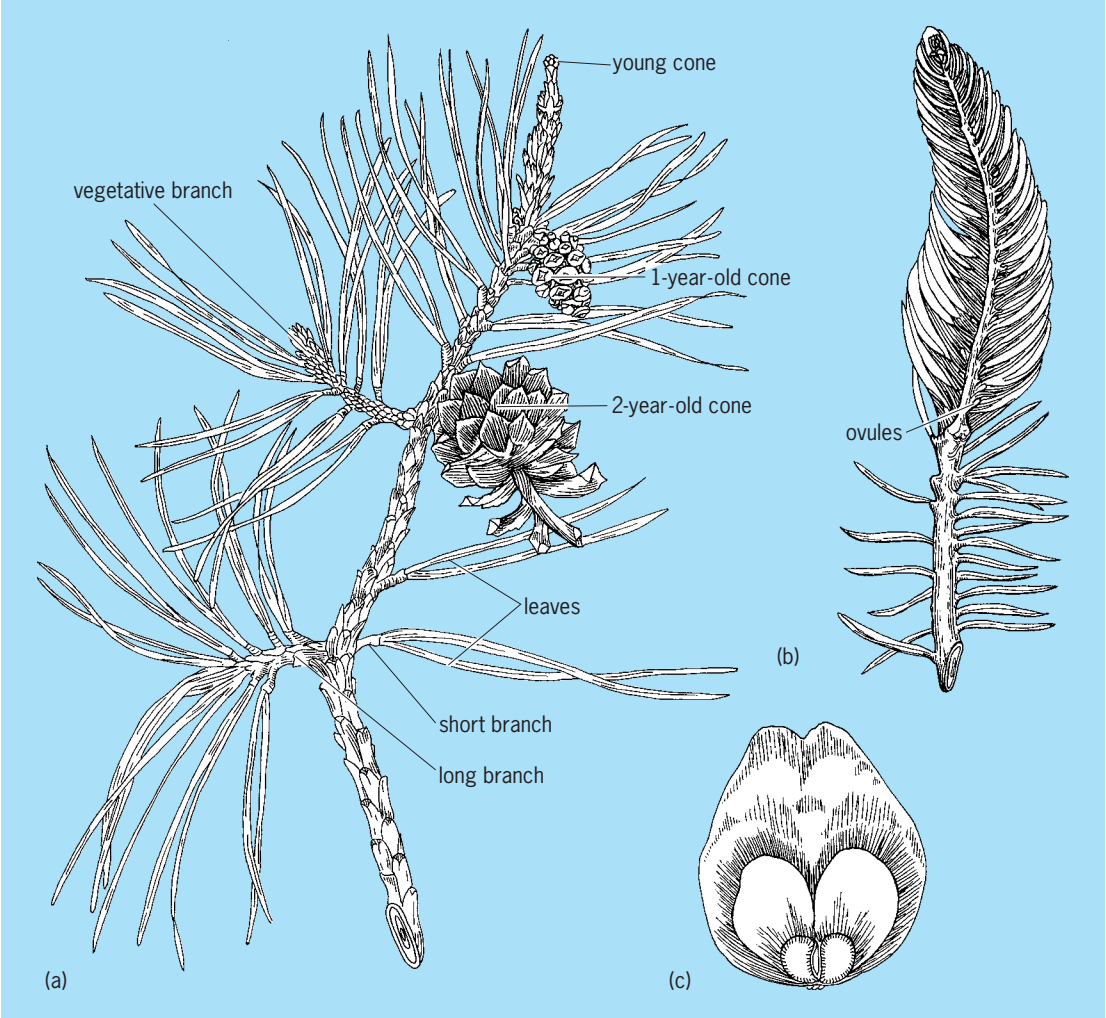

Fig. 1. Leaf, branch, and cone characteristics of conifers. (*a*) Branch of pine with leaves and ovulate cones. (*b*) Ovulate cone of spruce cut longitudinally to show scales with ovules. (*c*) Detail of one scale with two ovules. (*After J. B. Hill et al., Botany, 3d ed., McGraw-Hill, 1960*)

gametophyte with two or three archegonia, the egg-producing organs. The microspore produces the pollen tube (male gametophyte) with two sperms. The tube grows at the expense of the nucellus, carrying the sperm to the archegonia, each containing one egg. Here a sperm unites with an egg, forming a zygote which develops into an embryo. Often several eggs are fertilized in the same female gametophyte, so that for a time there are several embryos, but only one matures. Usually, a year intervenes between pollination and fertilization, and another year between fertilization and embryo formation. *See* PLANT REPRODUCTION.

Economic importance. The conifers are a principal source of lumber and pulp for paper and wood products. Turpentine, tar, resin, and essential oils are some by-products. The group yields little food for humans; some seeds are edible (pine, pinyon, or piñon nuts). *See* PINE NUT; PINE TERPENE. Thomas A. Zanoni

Fossils. Fossils of the earliest conifers consist of leafy shoots from the Middle Pennsylvanian of England. Their vegetative anatomy was similar to that of modern forms, and included uniseriate rays and resin ducts in the pith and leaves. Seed and pollen cones from the Late Pennsylvanian and early Permian of North America and Europe demonstrate that conifers were diverse by that time, although peak conifer diversity occurred during the Mesozoic. The modern families Araucariaceae, Podocarpaceae, Taxodiaceae, and Pinaceae occur in the Triassic; and some modern genera, including *Taxus*, *Torreya*, *Araucaria*, and *Sequoia*, were present by the Jurassic.

According to one theory, conifers are a sister group to the extinct Paleozoic order Cordaitales, from which they were previously believed to be descended. The Devonian progymnosperm *Archaeopteris* has been suggested as progenitor of the conifer-cordaite lineage because of its vegetative features, including wood anatomy and simple spirally arranged leaves. A pteridosperm ancestry is also possible, based on the coniferlike reproductive biology of the Callistophytales, which possess saccate pollen and platyspermic ovules. *See* CORDAITALES; PTERIDOSPERMS.

Fossil evidence suggests that the ovulate conifer cone was derived from a structure like that of the

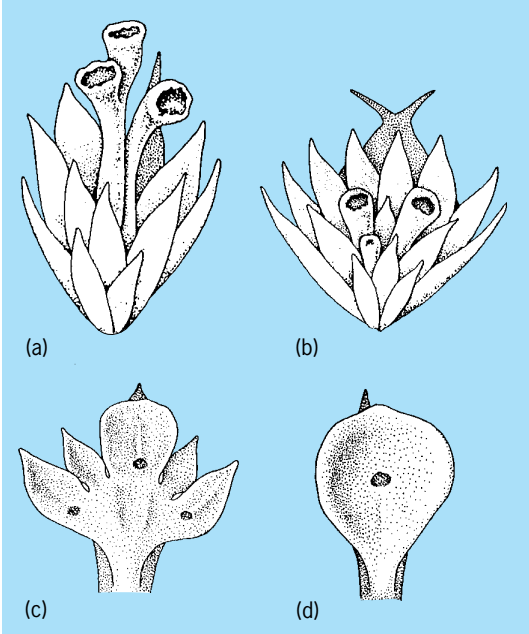

Fig. 2. Fertile dwarf shoots of a cordaite and select fossil conifers representing stages in cone scale evolution. A direct phylogenetic series is not implied. (*a*) In *Cordaites* the ovules were borne terminally on the megasporophylls. (*b*) *Emporia*. (*c*) The dwarf shoot is nearly flattened in one plane with scales fused basally. (*d*) The dwarf shoot of *Ullmania* is completely fused and flattened, forming a true cone scale. (*After G. Mapes and G. W. Rothwell, Structure and relationships of primitive conifers, N. Jb. Geol. Paleont. Abh. 183:269–287, 1991*)

Cordaitales, which consists of a long axis with bracts. In each bract axil was a fertile, dwarf shoot with spirally arranged sterile scales and elongate megasporophylls with terminal ovules (**Fig. 2***a*). Progressive flattening of these dwarf shoots and fusion of the scales and megasporophylls culminated in the cone scale of modern conifers. Fertile dwarf shoots of various fossil conifers are interpreted as intermediate stages in this transformation. The Pennsylvanian conifer *Emporia* (Fig. 2*b*) possessed fertile dwarf shoots similar to but more flattened than those of Cordaitales. Early conifers of the family Utrechtiaceae bore ovules laterally on fertile scales. Most early conifers retained sterile scales on the fertile dwarf shoot, although some had only fertile scales.

A more advanced condition occurred in the late Permian family Majonicaceae, in which the fertile dwarf shoot elements were fused basally. Dwarf shoot elements of *Pseudovoltzia* (Fig. 2*c*) were nearly flattened in one plane. The artificial family Voltziaceae includes many Paleozoic and Mesozoic forms, some of which have been interpreted as additional intermediates between Paleozoic and modern conifers. Although some of these are probably ancestors or early representatives of the modern families, others belong to extinct lineages.

By the late Permian, *Ullmania* (Fig. 2*d*) had achieved a flattened, fused cone scale like that of modern conifers. The suggestion is that the cone scale was derived at different times, independently in various lineages. The early Permian families Fergliocladaceae and Buriadiaceae had simple ovulate cones, and ovules borne singly among vegetative leaves, respectively.

Pollen cones of modern conifers are simple, with pollen sacs borne abaxially on microsporophylls. Fossil pollen cones exhibit greater diversity, and include Paleozoic forms with adaxial pollen sacs and Mesozoic forms with stalked pollen sacs surrounding the microsporophyll base. *See* PINOPHYTA; PINOPSIDA; PLANT KINGDOM. Brian J. Axsmith

Bibliography. C. B. Beck, *Origin and Evolution of Gymnosperms*, 1988; J. A. Clement-Westerhof, Aspects of Permian paleobotany and palynology, 8: The Majonicaceae, a new family of Late Permian conifers, *Rev. Palaeobot. Palynol.*, 52:375–402, 1987; G. Mapes and G. W. Rothwell, Structure and relationships of primitive conifers, *N. Jb. Geol. Paleont. Abh.*, 183:269–287, 1991; C. N. Miller, Mesozoic conifers, *Bot. Rev.*, 43:218–280, 1977.

Pinch effect

A name given to manifestations of the magnetic self-attraction of parallel electric currents having the same direction. The effect at modest current levels of a few amperes can usually be neglected, but when current levels approach a million amperes such as occur in electrochemistry, the effect can be damaging and must be taken into account by electrical engineers. Since the late 1940s the pinch effect in a gas discharge has been studied intensively in laboratories throughout the world, since it presents a possible way of achieving the magnetic confinement of a hot plasma (a highly ionized gas) necessary for the successful operation of a thermonuclear or fusion reactor.

Ampère's law. The law of attraction which describes the interaction between parallel electric currents was discovered by A. M. Ampère in 1820 and can be stated as follows: The force of attraction in newtons per meter of length between two thin straight wires r meters apart carrying currents I_1 and I_2 amperes, respectively, is $2 \times 10^{-7} I_1 I_2 / r$. The law applies equally to the attraction between the individual components of a current in a single wire, in which case, for a cylindrical wire of radius r meters carrying a total surface current of I amperes, it manifests itself as an inward pressure on the surface (**Fig. 1**) given by $I^2/2 \times 10^7 \pi r^2$ pascals.

For the electric currents of normal experience, this force is small and passes unnoticed but it

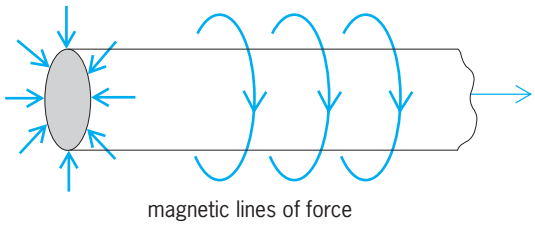

Fig. 1. Pinch pressure on a current-carrying conductor. Arrows at left show direction of pinch pressure.

is significant that the pressure increases with the square of the current, I^2. For example, at 25,000 A the pressure amounts to about 1 atm (100 kilopascals) for a wire of 1-cm radius, but at 10^6 A the pressure is about 1600 atm or about 12 tons per square inch (160 megapascals).

Manifestations. The pinch effect first showed up practically in certain early types of induction electric furnaces in which large low-frequency alternating currents of the order of 100,000 A were induced at low voltage in a horizontal ring-shaped fused-metal load (**Fig. 2**). At these currents, the pinch pressure can be larger than the hydrostatic pressure exerted by the fused metal, and as indicated above ($I^2/2 \times 10^7 \pi r^2$), the pinch pressure increases as the radius of the conductor decreases. Consequently, once the pinch process starts, the pressure at a narrow neck in the ring of fused metal can squeeze out the fluid metal until the neck pinches off completely, cutting off the current. This led to very uneven heating of the load. The term pinch effect was given to this process by C. Hering in 1907. The technical difficulty was eventually overcome by making the plane of the ring vertical and submerging it deeply below the surface of the fused metal. The force of the pinch effect also manifests itself by the crushing of tubular conductors exposed to large impulsive currents, such as occur in lightning strokes or high-power short circuits, and is used in certain metal-forming techniques.

Thermonuclear applications. One of the conditions for the attainment of a profitable balance between energy expended in heating and energy released in fusion from a thermonuclear reaction in a plasma composed of deuterium and tritium (DT, the most favorable case) is that the temperature shall be not less than about 10 keV (1.16×10^8 K). This enormous temperature can be reached and maintained only if the hot plasma is effectively isolated from the material walls of the container by vacuum. The isolation has the function of preventing cooling by contact with materials at normal temperatures and allowing the plasma to be heated. For the plasma to remain confined under these conditions, its outward pressure must be balanced by the inward pressure of nonmaterial origin, for example, a magnetic field. A profitable energy balance also depends on the density n of the confined plasma and on τ, the time it is confined. The product of $n\tau$ must exceed a certain minimum, which is 10^{14} ions cm^{-3} s for DT. *See* LAWSON CRITERION; MAGNETOHYDRODYNAMICS; PLASMA (PHYSICS); THERMONUCLEAR REACTION.

There are a number of ways in which a magnetic field can be arranged around the plasma to hold it together, and one of these methods is the pinch effect. A fusion reactor using this type of confinement would ideally be a toroidal tube in which the confined plasma would carry a large electric current induced in it by magnetic induction from a transformer core passing through the major axis of the torus. The current would have the double function of ohmically heating the plasma and compressing the plasma toward the center of the tube.

The fundamental equation for the pinch effect in a gas, derived theoretically by W. Bennett in 1934, gives the current required for the inward pinch pressure to balance the outward gas pressure, as shown in the equation below, where I is the total current

$$\frac{I^2}{2 \times 10^7} = Nk(T_i + T_e)$$

in amperes, N is the number of electrons (also the number of ions) per meter of length of the pinch, $k = 1.4 \times 10^{-23}$ J/K (Boltzmann's constant), and T_i and T_e are the temperatures in kelvins of the ions and electrons, respectively.

Experimental studies. In general, two types of apparatus were used in early studies of the pinch effect: (1) straight discharge tubes of quartz or porcelain, with metal electrodes at each end, intended for short-duration studies, in which the cooling of the plasma by the relatively cold electrodes was slight during the time of the experiment, and (2) toroidal

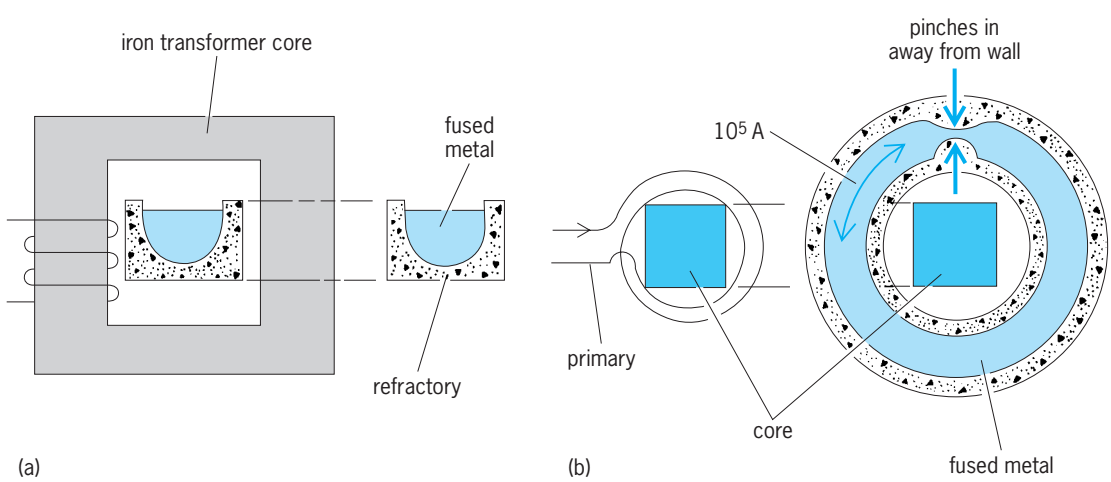

Fig. 2. Early type of ring induction electric furnace. (*a*) Side view. (*b*) Plan view.

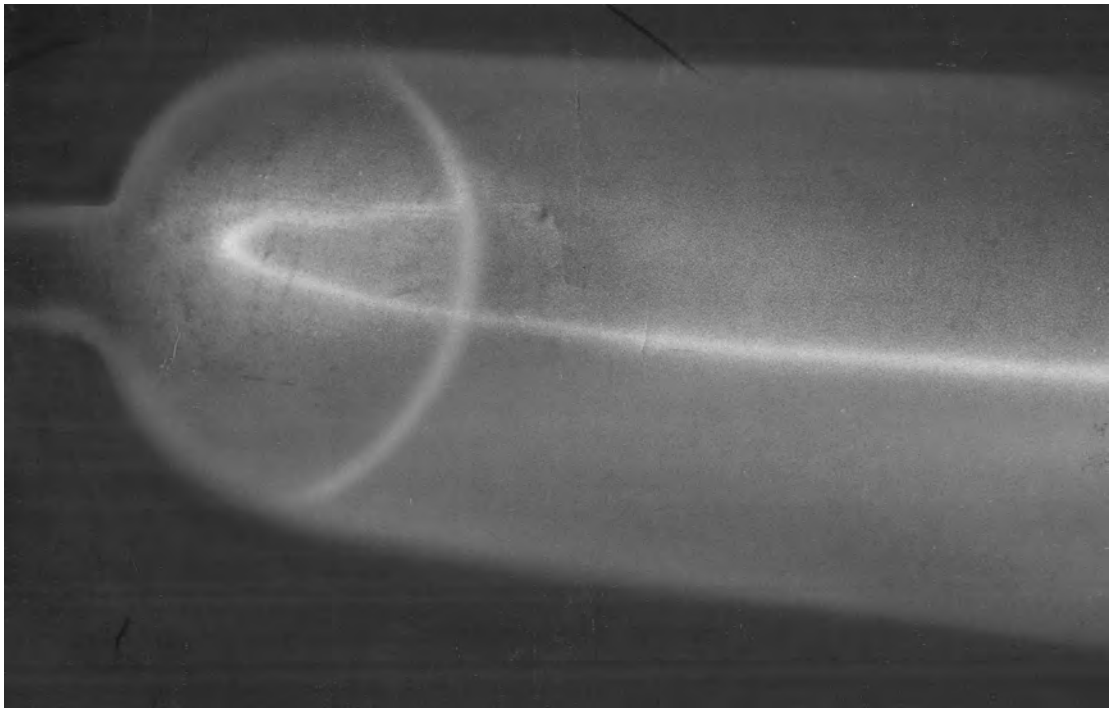

Fig. 3. Xenon pinched discharge in Perhapsatron torus.

discharge tubes, also composed of quartz or porcelain, in which the pinch was endless and consequently was more effectively confined than in the first type of apparatus, and the current was induced into the discharge by magnetic coupling to a primary winding. In both cases, currents of 50,000–500,000 A were obtained by electric fields of 10–100 V/cm along the pinch. The primary power sources in early experiments used charged capacitors with capacitances of 4–50 microfarads, charged to 10–100 kV.

Instability. Characteristically, as can be shown by high-speed photography, the pinch forms at the inner surface of the discharge tube wall and contracts radially inward, forming an intense line, the pinch, on the axis (**Fig. 3**); the pinch rebounds slightly; the contracted discharge rapidly develops necks and kinks; and in a few microseconds all structure is lost in an apparently turbulent glowing gas which fills the tube. Thus, the pinch turns out to be unstable, and plasma confinement is soon lost by contact with the wall. The cause of the instability is easily seen qualitatively: The pinch confinement can be described as being caused by the magnetic field lines encircling the pinch which are stretched longitudinally but which are in compression transversely (**Fig. 4**). For a uniform cylindrical pinch, the magnetic pinch pressure is everywhere equal to the outward plasma pressure, but at a neck or on the inward side of a kink, the magnetic field lines crowd together, creating a higher magnetic pressure than the outward gas pressure. Consequently, the neck contracts still further, the kink cuts in on the concave side and bulges out on the convex side, and both perturbations grow. The instability has a disastrous effect on τ, limiting it to 10^{-6} s or less in light atom plasmas such as DT.

Neutrons have been produced by deuterium pinches in large numbers. For a time (1952–1953), they were thought to be evidence of thermonuclear reaction, but it has since been shown that they are emitted preferentially in certain directions, and are associated with the instability of the pinch and the violent accelerations that are produced. Such neutrons

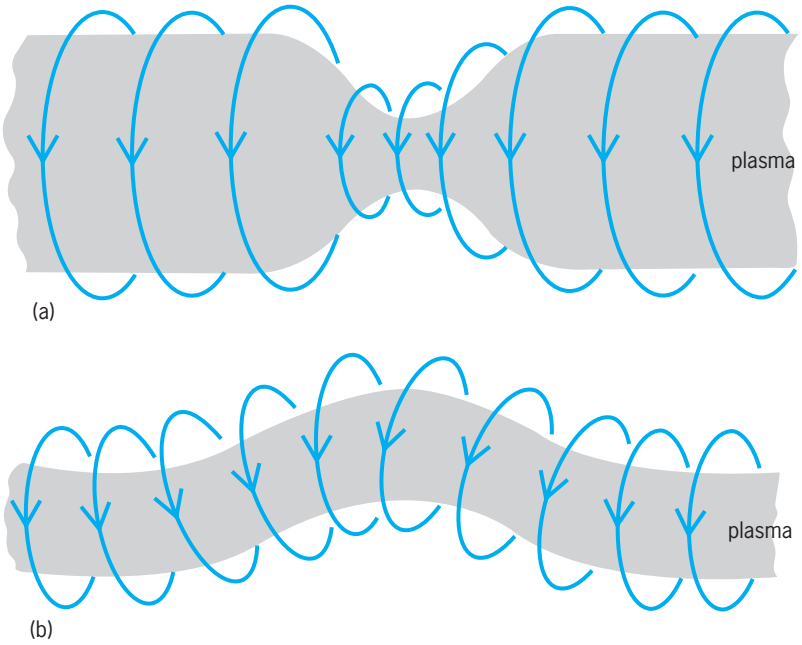

Fig. 4. Instability. (*a*) Sausage type. (*b*) Kink type.

are then not a product of thermal collisions and are not thermonuclear.

Great efforts have been devoted to overcoming the basic instability of the simple pinch. One such measure was to add an axial magnetic field by means of an external winding around the pinch tube. This might be expected to resist the neck and kink deformation by stiffening the discharge. Also, a conducting wall located close to the discharge tube has the effect of trapping the magnetic field between the pinch and the wall, cushioning and reflecting the moving pinch back to the center.

Studies which involved such modifications were in progress on a worldwide scale during the period 1955–1963, notably including Zeta (Harwell, United Kingdom), Alpha (Leningrad, Soviet Union), and the Perhapsatron (Los Alamos, United States). In general these measures were disappointing, and work on pinches of this type declined.

Reversed-field Z pinch. In the 1970s, encouraged by analysis of Zeta results and improved theoretical understanding of pinch instability, there was renewed interest in the pinch formed by adding another longitudinal field outside the pinch in the opposite direction to that inside the pinch described above. This geometry is known as the reversed-field Z pinch. Confirmation of improved Z-pinch stability was undertaken in ZT-40 (**Fig. 5**; Los Alamos, United States), ETA Beta II (Padua, Italy), HBTX (Culham, United Kingdom), and elsewhere.

High-power pulsed systems. The simple pinch is still of considerable interest because the reciprocal functional relation between n and τ for power production means that there is always the possibility of achieving a net power output no matter how much τ is reduced by instabilities, by the use of very high n. Such plasmas will require extreme high-pulsed electric power to heat and confine them, with perhaps the only limitation being that the resultant output burst of thermonuclear power may destroy the machine. For such pulsed systems, the pinch must always rank highly—it is uniquely the most efficient magnetic confinement system expressed in terms of magnetic energy expended per unit of plasma energy confined, and furthermore the pressure it confines can exceed the strength of any known material.

Theta pinch. The term theta pinch has come into wide usage to denote an important plasma confinement system which relies on the repulsion of oppositely directed currents and which is thus not in accord with the original definition of the pinch effect (self-attraction of currents in the same direction). Plasma confinement systems based on the original pinch effect are known as Z pinches. **Figure 6** defines the two geometries. The first laboratory thermonuclear reaction was in the small θ-pinch experiment, Scylla, at Los Alamos in 1957.

Tokamak. The stabilized Z pinch became the subject of intense interest when L. Artsimovich of the Soviet Union announced in 1969 the achievement of a 20-ms confinement and an ion temperature of 0.5 keV (5×10^6 K) in the T-3 tokamak experiment. Tokamak is essentially a low-density, slow Z pinch in

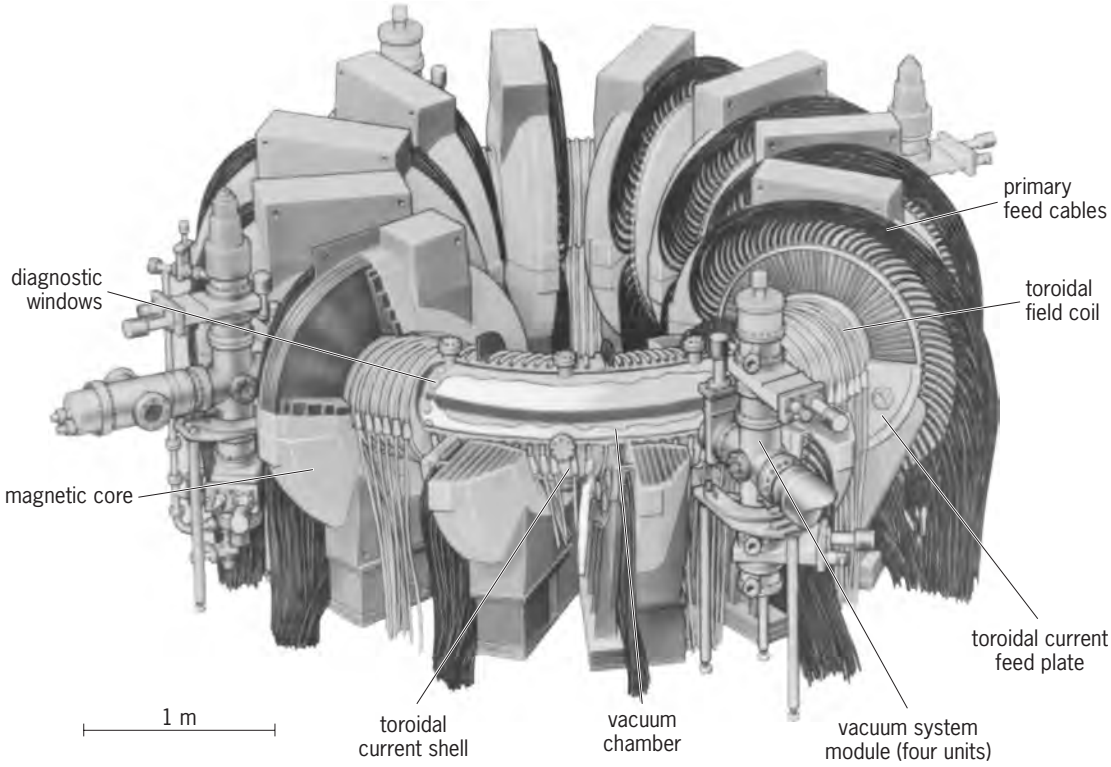

Fig. 5. ZT-40 reversed-field Z-pinch apparatus. 1 m = 3.3 ft. (*Los Alamos Scientific Laboratory*)

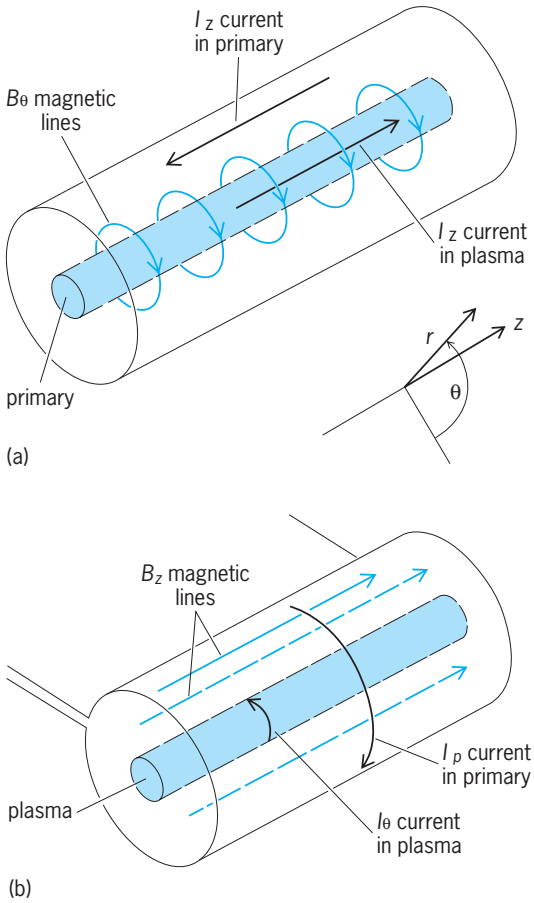

Fig. 6. Two geometries for plasma confinement systems. (*a*) Z pinch. (*b*) θ pinch.

a torus with a very strong longitudinal field such that the pinch current is below the Kruskal-Shafranov limit. This means that the helical magnetic field lines, resultant from the externally applied field and that of the pinch, do not close, that is, complete one revolution of the minor axis in going around the major axis of the torus once. This is known theoretically to prevent the growth of certain helical distortions of the plasma. The achievements of tokamak experiments are as follows, together with the target values for achieving a positive power balance: $T_i = 6$ keV for tokamak, target value 10; $n\tau\gamma = 3 \times 10^3$ ions cm^{-3} for tokamak, target value 10^{14}.

The temperatures achieved are higher than the 0.5 keV that could be reached by resistive heating of the plasma by the pinch current, such as was used in the early 1970s. Auxiliary heating is necessary, and the most successful has been the injection of energetic neutral beams.

Construction of several large tokamak installations has been undertaken, including the TFTR (Princeton, United States), JET (European Community at Culham, United Kingdom), JT-60 (Tokyo), and T-15 (Moscow). *See* NUCLEAR FUSION. James A. Phillips

Bibliography. K. Niu, *Nuclear Fusion*, 1989; E. Teller (ed.), *Fusion*, 2 pts., 1981.

Pine

The genus *Pinus*, of the pine family, characterized by evergreen leaves, usually in tight clusters (fascicles) of two to five, rarely single. There are about 80 known species distributed throughout the Northern Hemisphere. Botanically the leaves are of two kinds: (1) a scalelike form, the primary leaf, which subtends a much shortened and eventually deciduous shoot bearing (2) the secondary leaves or needles. The wood of pines is easily recognized by the numerous resin ducts and by the characteristic resinous odor. The pines may be divided into two classes according to the number of leaves in a cluster. *See* PINALES; SECRETORY STRUCTURES (PLANT).

Soft or white pines. Except for the nut pines, the white pines have five needles in a cluster. Eastern white pine (*P. strobus*), ranging 90–150 ft (27–45 m) in height, is found in the northeastern United States west to the Lake states, adjacent Canada, and the Appalachian Mountain region. Originally it was the most important timber tree of the eastern United States and Canada. The wood of eastern white pine is valuable because it can be easily worked, is light and soft, does not split when nailed, polishes well, and does not warp or swell appreciably. Almost everything—shipmasts, matches, doors, framing, finish, boxes, and crates—has been made from this wood. However, it is restricted to more particular uses because of its increasing scarcity and value.

Of the western white pines, the sugar pine (*P. lambertiana*) [**Fig. 1**] is the most important. This is a magnificent Pacific Coast tree attaining a height of about 250 ft (75 m) and having large cones, 10–20 in. (25–50 cm) long. It is lumbered only in Oregon and California, where it ranks in volume and value with redwood. Other white pines in the West are western white pine (*P. monticola*), a mountain species found almost entirely in Idaho, Montana, and Oregon; limber pine (*P. flexilis*), one of the smaller white pines of the Rockies; and whitebark pine (*P. albicaulis*), also a mountain species with a more northern range. *See* REDWOOD.

The nut pines or piñons are a subgroup of the Southwest with fewer needles, sometimes only one. *See* PINE NUT.

Hard or pitch pines. Red pine (*P. resinosa*), also known as Norway pine, reaches a height of 80 ft (24 m) or more, and is native in the northeastern United States from Maine to Minnesota and adjacent Canada and south along the mountains to West Virginia. The needles are in pairs, and the bark has a red-brown color, hence the name red pine. The wood is fairly soft, but a little harder than that of eastern white pine. The more dense red pine is also stronger and is important commercially for general construction, sash, door, and window frame manufacturing, flooring, boxes, crates, and shipmasts, but it is not durable in contact with the soil. Frequently it is sold in mixture with eastern white pine. About two-thirds of the total volume is in the Northeast and the remainder in

the Lake states. Technologically the red pine forms a sort of bridge between the soft white pines and the hard pines.

The hard yellow pines have two or three needles in a cluster. The longleaf pine (*P. palustris*), the loblolly (*P. taeda*), the shortleaf (*P. echinata*), and the slash pine (*P. elliottii*) are the principal coniferous or softwood trees of the southeastern United States and are of major economic importance. These four species are the primary producers of lumber and pulpwood in the South. Production of veneer for plywood is also important. The particularly hard and strong wood of the longleaf pine has long made it a preeminent species for structural timbers, flooring, railway car construction, and similar uses. The longleaf and slash pines are the principal trees for the production of turpentine, a distillate obtained from the resin when the trees are tapped. *See* GUM; PINE TERPENE; RESIN.

In the West the hard pines are represented chiefly by the ponderosa pine (*P. ponderosa*), which attains a height of 150–225 ft (45–70 m), and is found in the Rocky Mountain and Pacific Coast regions, including adjacent Canada.

The Austrian pine (*P. nigra*), which has two needles and closely resembles the red pine, has darker bark and is much planted in North America, as is the Scotch pine (*P. sylvestris*), with two shorter, bluish needles (**Fig. 2**). Other exotics cultivated are mountain pine (*P. mugo*), Japanese black pine (*P. thunbergi*), and Swiss stone pine (*P. cembra*), the last being a five-needled species. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Fig. 1. Mature sugar pine (*Pinus lambertiana*) in Sequoia National Park, California. (*National Park Service*)

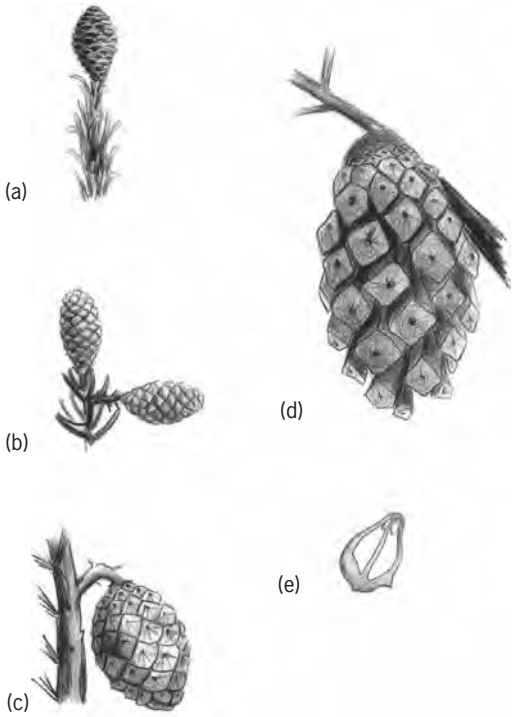

Fig. 2. Scotch pine (*Pinus sylvestris*). (*a*) Female (ovulate) cone, unfertilized. (*b*) Fertilized cones. (*c*) Advanced stage in cone development. (*d*) Cone opening to discharge seeds. (*e*) Seed scale showing two winged seeds.

Bibliography. A. Farjon, *Pines: Drawings and Descriptions of the Genus Pinus*, 2d ed., Brill Academic, 2005; A. Farjon and B. T. Styles, *Pinus* (*Pinaceae*), Flora Neotropica Monog. 75, New York Botanical Garden, 1997; N. T. Mirov, *The Genus Pinus*, Ronald Press, 1967; D. M. Richardson (ed.), *Ecology and Biogeography of Pinus*, Cambridge University Press, 1998.

Pine nut

The edible seed of more than a dozen species of evergreen cone-bearing trees in the genus *Pinus*, native to the temperate zone of the Northern Hemisphere. The important nut-producing species are the stone pine (*P. pinea*) of southern Europe; the Swiss stone pine (*P. cembra*), native to the Swiss Alps and eastward through Siberia to Mongolia; and the piñon pine (*P. cembroides* var. *edulis*) of the arid regions of the southwestern United States. The seeds or nuts, variable in size according to species, are borne in cones which take 3–4 years to develop.

The cones of the stone pine are 4–6 in. (10–15 cm) long, each containing a hundred or more nuts $^5/_8$–$^3/_4$ in. (16–19 mm) in length. Cones are picked from the trees before the cone scales separate and later dried in the sun to free the nuts. The nuts are removed from the cone by hand and cracked mechanically to free the kernels which are cleaned and

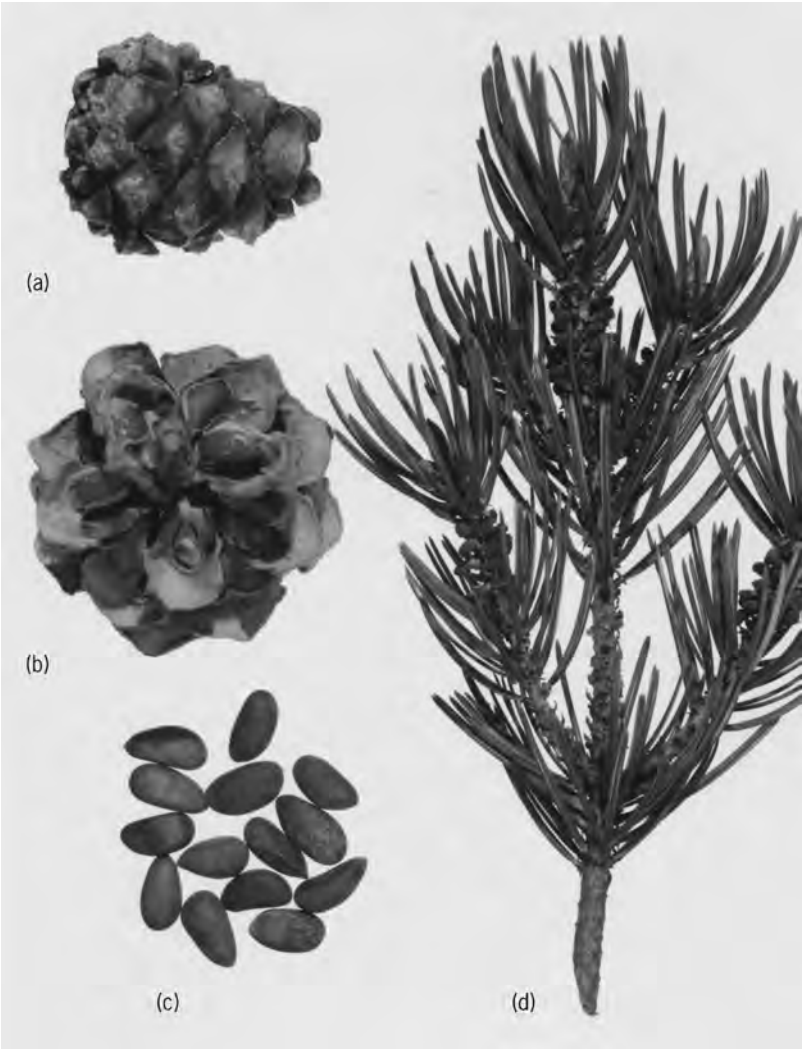

Piñon pine, *Pinus cembroides* var. *edulis*. (*a*) Unopened cone. (*b*) Opened cone. (*c*) Nuts in shell. (*d*) Branch with needles and old staminate cones.

packed for export. Imports to the United States are mostly from Italy and Spain.

The piñon pine (see **illus.**) of Colorado, New Mexico, Arizona, and northern Mexico grows in forests of scattered trees on arid land with only 12–14 in. (30–36 cm) annual rainfall. The trees are dwarf and grow slowly. The relatively small cones open on the trees, freeing the nuts, which are picked up, mostly by local Indians, and are either used as food or for commerce.

Pine nuts are an important staple food over large areas of the Northern Hemisphere. Only a small part of the nuts produced reaches local or world markets. The raw nuts of some species have a strong turpentine flavor which disappears when they are roasted. *See* PINE; SEED. Laurence H. MacDaniels

Pine terpene

A major component of the essential oils obtained from various *Pinus* species. The principal terpenes of the oil of southern pines [longleaf pine (*P. palustris*) and slash pine (*P. caribaea*)] are α- and β-pinene. Smaller amounts of moncyclic dienes and 3-carene are also present.

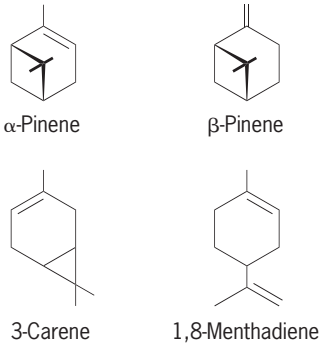

See ESSENTIAL OILS; PINE.

Turpentine. Gum turpentine (gum spirits) is the volatile fraction of the oleoresin that exudes from cuts made in the trunks of live trees. The resin is collected and distilled by a process that yields about 20% turpentine, mainly α- and β-pinene, and 70% rosin; it was the basis of the original naval stores industry.

Wood turpentine is obtained by steam distillation from stumps and other logging residues. The volatile material in this case consists of about 50% turpentine and 30–40% of higher-boiling-point alcohols; the latter fraction is known as pine oil. The bulk of the wood turpentine and pine oil produced by modern industrial processes is a by-product of the sulfate wood-pulping process (sulfate turpentine).

Important uses of turpentine or the purified pinenes derived from turpentine are in terpene resins, as a thinner in paints and varnishes, and as a starting material in the synthesis of other commercially valuable terpenes.

Pine oil. This is a mixture of monoterpene alcohols, mainly α-terpineol, obtained in large amounts mixed with wood turpentine or sulfate turpentine. The term pine oil is also used to designate the essential oil of various species of pine.

Much of the pine oil of commerce is prepared synthetically by acid-catalyzed hydration of α-pinene. This process involves a complex series of reactions that occur via cationic intermediates. A few of these steps are shown in the following reaction scheme.

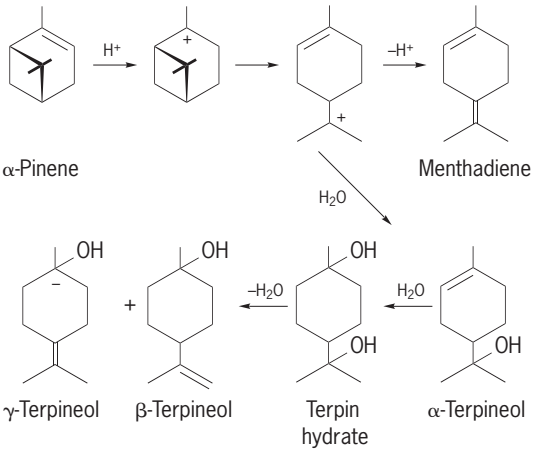

The composition of industrial-grade pine oil is approximately 65% α-terpineol, 20–25% of other monoterpene alcohols, and 10–15% hydrocarbons. Pine oil has surfactant and emulsifying properties and is also a disinfectant. Most of the pine oil manufactured is used in the manufacture of cleansers and textile penetrants. *See* TERPENE; WOOD CHEMICALS.

James A. Moore

Pineal gland

An actively functioning endocrine gland, located in the brain, which secretes melatonin, is strongly regulated by light stimuli, and is an important component of the circadian timing system. The pineal gland or pineal body has the Latin name *epiphysis cerebri*. It is virtually ubiquitous throughout the vertebrate animal kingdom. In nonmammalian vertebrates (lampreys, fishes, amphibians, reptiles, birds), it functions as a photoreceptive third eye and an endocrine organ. In mammals, it serves as an endocrine organ that is regulated by light entering the body via the paired eyes. In all vertebrates, including humans, the pineal gland produces the hormone melatonin, with the highest synthesis and secretion occurring during the night. Despite extensive species variation in anatomy and physiology, the pineal gland generally serves as an essential component of the circadian system which allows animals to internally measure time and coordinate physiological time-keeping with the external environment. *See* BIOLOGICAL CLOCKS; BRAIN.

Anatomy. The pineal gland is an unpaired organ attached by a stalk to the roof of the diencephalon. In frogs and lizards, one component of the pineal complex (the frontal organ or parietal eye) projects upward through the skull to lie under the skin; in all other vertebrates the pineal is located beneath the roof of the skull. No evidence of the pineal has been found in the simplest cyclostomes (myxinoids), and it is lacking in crocodiles, armadillos, and edentates. The size and shape of the pineal varies with species: it is relatively large in cocks, echidnas, rodents, ungulates, and humans. In anamniotes, the pineal has a prominent intrapineal neuronal apparatus transmitting light stimuli to the brain via neural mechanisms. This apparatus is lost in mammals, where endocrine cells predominate.

Across evolution, cells within the pineal gland have progressed from classic photoreceptor cells in the earliest vertebrates, to rudimentary photoreceptors in birds, to classic endocrine cells in mammals. In cyclostomes, two outgrowths project from the dorsal roof of the diencephalon in lampreys. One, derived from the right side of the brain, is the pineal body; the other outgrowth is the parapineal organ, which is more ventral and derived from the left side of the brain. The pineal gland forms a hollow, knoblike end vesicle directly beneath an area of skin devoid of pigment and lying dorsally between the paired eyes. The dorsal wall of the pineal forms a lenslike structure, while the ventral wall contains sensory cells and ganglion cells with processes passing down the stalk to various nerve centers in the brain. The parapineal of lampreys is essentially similar in structure. Both organs are sensitive to light. In some ganoids and larval teleosts, all tissues overlying the pineal are translucent, and light striking the pineal initiates nerve impulses which are transmitted to the brain. An intracranial parapineal also occurs in teleost fishes. Together, the pineal and parapineal are called the pineal complex. In certain reptiles (*Sphenodon, Sceloporus*, and some other lizards), a parapineal or parietal eye, which is highly specialized and endowed with a distinct lens, occurs along with a glandular pineal body. In snakes, galliform birds, and mammals, the pineal is more thoroughly a glandular structure. *See* PHOTORECEPTION; SENSE ORGAN.

Innervation and biochemistry. The pineal glands of all vertebrate species receive light input by indirect, retinal projections, while fish, amphibian, and bird

species also have directly photosensitive pineal organs. In mammals, nerve fibers extend from a variety of sources in the brain to the pineal gland. The best studied of these neural inputs is through the retinohypothalamic tract, which extends from the eyes to the pineal gland in mammals (see **illus.**). Originating in the retina, the majority of the retinohypothalamic fibers project to or around the bilateral suprachiasmatic nuclei in the hypothalamus. These nuclei serve as endogenous oscillators with period lengths close to 24 h. Thus, the suprachiasmatic nuclei function as pacemakers for the circadian system, which regulates daily physiological and behavioral rhythms. From the suprachiasmatic nuclei there are short projections to the paired paraventricular hypothalamic nuclei, and then long descending axons project from these nuclei to synapse on preganglionic sympathetic neurons in the upper thoracic spinal cord. These sympathetic neurons then extend out of the central nervous system to the superior cervical ganglia in the neck region. From there, postganglionic sympathetic axons reenter the cranium and ultimately innervate the pineal gland.

In mammals, information about environmental light and darkness is relayed from the eye to entrain circadian neural activity of the suprachiasmatic nuclei. In turn, the suprachiasmatic nuclei synchronize circadian rhythms in the pineal gland through its sympathetic innervation. One of the best-studied rhythms in the pineal gland is the biosynthesis of the hormone melatonin, with the structure below.

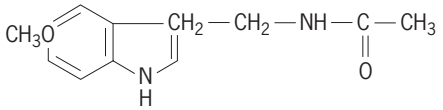

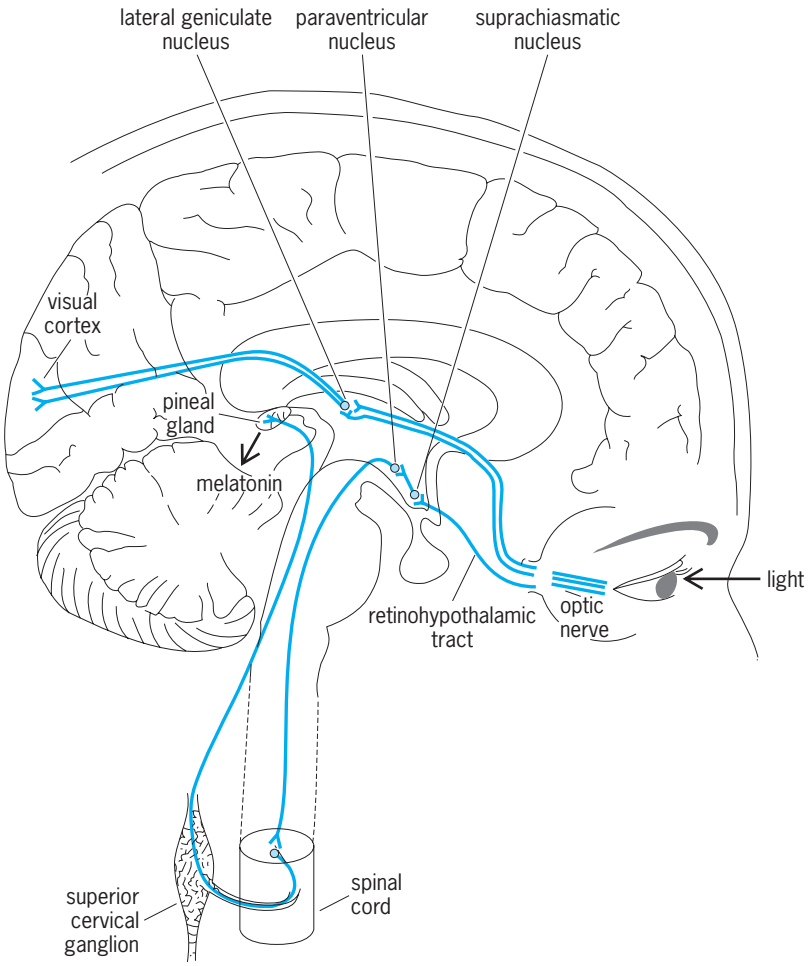

Two anatomically distinct pathways which convey light into the human brain. One pathway extends from the retina to the visual cortex and supports the sensory capacity of vision. The other pathway extends from the retina to the pineal gland for the circadian regulation of melatonin production.

Melatonin has been found in all vertebrate and nonvertebrate species studied, from single-celled plankton to humans. In vertebrates, pineal endocrine cells take up the amino acid tryptophan from the circulation and convert it to serotonin in two enzymatic steps. The conversion of serotonin to melatonin then takes two additional enzymatic steps which are controlled by adrenergic innervation. When the sympathetic nerves to the pineal release the neurotransmitter norepinephrine, it binds to receptors on the pinealocyte cell-surface membranes. This neural stimulation ultimately regulates melatonin synthesis within the pinealocytes by activating an intracellular second messenger. The second messenger initiates an intracellular cascade of biochemical changes which result in the conversion of serotonin to melatonin. For the most part, melatonin is not stored within the pineal gland. Once synthesized, melatonin is quickly released into the blood. From the circulation, melatonin can be transferred into virtually all of the body fluids such as saliva and cerebrospinal fluid. The primary site for melatonin metabolism is the liver, but some melatonin breakdown occurs in the kidney and brain. In addition to the synthesis of melatonin, pinealocytes have the necessary enzymes for converting tryptophan into a larger family of indole compounds. Furthermore, numerous polypeptides have been localized in the pineal gland. The biological functions of these other pineal indole and peptide constituents are currently unknown.

Light regulation of melatonin. In all vertebrate species studied, high levels of melatonin are produced and secreted during the night, while low levels are released during the day. The melatonin circadian rhythm is produced by the endogenous pacemaking activity of the suprachiasmatic nuclei, while the entrainment of this rhythm is coordinated by signals of light and darkness relayed from the eyes. In addition to synchronizing the daily circadian rhythm of pineal melatonin secretion, light has both an acute and longer-term effect on melatonin production. Specifically, unexpected exposure of the eyes to light during the night can induce an acute, rapid decrease in the high nocturnal synthesis and secretion of melatonin. This light-induced melatonin suppression has been documented repeatedly

across different vertebrate species and has been used in numerous studies to help determine the ocular, neural, and biochemical physiology of melatonin regulation. In general, the brighter the light stimulus, the greater the suppression of nocturnal melatonin. Different species, however, have widely different sensitivities for light suppression of melatonin.

In addition to light entraining the melatonin rhythm and acutely suppressing nocturnal melatonin, the day length or photoperiod can influence the duration that melatonin production is elevated during the night. This accounts for a seasonal effect of light on the pineal gland. Specifically, in the summer when days are longer and nights are shorter, the duration of increased nocturnal melatonin secretion is shorter than during the winter when nights are longer. This effect of photoperiod length influencing the duration of nighttime melatonin rise has been documented in many species, including humans.

Melatonin regulation of physiology. There is extensive species diversity in the capacity of melatonin to regulate physiology. Numerous species, ranging from insects to mammals, have yearly cycles of activity, morphology, reproduction, or development which are responsive to seasonal changes in day length. Biological regulation relative to annual cycles of daylight is called photoperiodism. Among many species that breed seasonally, melatonin has been shown to be a potent regulator of the reproductive axis in both males and females. Seasonal changes in day length influence melatonin secretion which, in turn, coordinates reproductive physiology in a way that allows for the birth of the young to occur when there is maximum possibility for their survival. Further, the effects of melatonin on the regulation of circadian physiology has been elucidated in many vertebrate species, including humans. In addition, melatonin has been studied in different species for its influence on retinal physiology, sleep, body temperature regulation, immune function, and cardiovascular regulation.

The pineal gland is the primary tissue that secretes melatonin, although the retina and perhaps portions of the intestinal tract have some capacity for melatonin synthesis. Once melatonin is secreted from the pineal gland into the blood, it circulates to target tissues where it binds to melatonin receptors. Melatonin may regulate circadian and reproductive responses by binding to specific membrane-bound receptors in discrete regions in the brain, including the hypothalamic suprachiasmatic nuclei and the pituitary pars tuberalis. A structurally related melatonin receptor also has been localized in the retina and, to a lesser extent, the brain. Outside the nervous system, melatonin binding has been revealed in various peripheral tissues, including the ovaries, gut, and blood vessels. This nonneural binding may be involved in temperature and cardiovascular regulation.

Melatonin has been used as a pharmacological agent for a variety of purposes. It has been used for timing and regulatation of reproduction in seasonally breeding livestock such as sheep. In humans, controlled clinical studies have demonstrated that melatonin can be a rapid, mild sleep-inducing compound. In addition, melatonin has a capacity to "reset" the human circadian system when it becomes inappropriately synchronized due to blindness, shift work, or east-west jet travel. Many other claims have been made for the therapeutic and health value of melatonin, but sufficient controlled clinical studies are lacking to support such claims. Although the toxicity of melatonin appears to be low, there are unresolved concerns about the long-term safety of taking it.

George C. Brainard

Bibliography. J. Arendt, *Melatonin and the Mammalian Pineal Gland*, Chapman and Hill, London, 1995; S. A. Binkley, *The Pineal: Endocrine and Nonendocrine Function*, Prentice Hall, Englewood Cliffs, NJ, 1988; G. C. Brainard, M. D. Rollag, and J. P. Hanifin, Photic regulation of melatonin in humans: Ocular and neural signal transduction, *J. Biol. Rhythms*, 12(6):537–546, 1997; D. C. Klein, R. Y. Moore, and S. M. Reppert, *Suprachiasmatic Nucleus: The Mind's Clock*, Oxford University Press, Oxford, 1991; R. J. Reiter, Pineal melatonin: Cell biology of its synthesis and of its physiological interactions, *Endocrine Rev.*, 12(2):151–180, 1991; S. M. Reppert, D. R. Weaver, and T. Ebisawa, Cloning and characterization of a mammalian melatonin receptor that mediates reproductive and circadian responses, *Neuron*, 13(5):1177–1185, 1994; L. Wetterberg, *Light and Biological Rhythms in Man*, Pergamon Press, Stockholm, 1993.

Pineapple

A low-growing perennial plant, indigenous to the Americas. The cultivated varieties (cultivars) belong to the species *Ananas sativus* of the plant order Bromeliales.

The edible portion of the pineapple develops from a mass of ovaries on a fleshy flower stock having persistent bracts (see **illus.**). On the cultivated types, the flowers are usually abortive. The leaves are long and swordlike, usually rough-edged, and grow to a height of 2–4 ft (0.6–1.2 m). Commercial plantings bear fruit at the age of 12–20 months, and may continue to be productive for as much as 8–10 years. Propagation is by suckers or offsets which may be rooted in sand, but are usually set directly in the field where they are to produce. The leafy crowns may also be used as cuttings, but because they are harvested with the fruit, other methods of propagation are more satisfactory. *See* BROMELIALES.

The pineapple, a warm-climate plant, is injured by temperatures below 32°F (0°C). It does best in a dry atmosphere and relatively poor soil, but responds well to fertilizers. The major producing area is Hawaii, where special methods of culture and harvesting have been developed. A paper mulch is used in the production of much of the Hawaiian pineapple

Pineapple (*Ananas sativus*). (*USDA*)

crop. The paper is laid in long strips by a special machine which pulls soil over the edges of the strips to hold them in place. The mulch aids in weed control and decreases moisture evaporation from the soil. Because of the tendency to iron chlorosis of plants growing on these soils, spraying with an iron salt is frequently a part of the fertilization program. Careful control of the nitrogen supply in relation to the hours of sunshine, together with the use of hormone sprays under certain conditions, has made possible considerable control over the time of ripening, an important consideration in obtaining maximum year-round use of processing facilities. Pineapples are also grown in the West Indies and other tropical areas, and to a limited extent in southern Florida.

Pineapples are consumed fresh in considerable quantity, but because of distance from markets and the problems of transporting fresh fruit, most of the crop is canned as sliced pineapple or as juice. *See* FRUIT.

J. Harold Clarke

Pinnipeds

Carnivorous mammals of the suborder Pinnipedia, which includes 32 species of seals, sea lions, and walrus in three families. All species of the order are found along coastal areas from the Antarctic to the Arctic regions. Although many species have restricted distribution, such as the Baikal seal (*Pusa sibirica*) of Lake Baikal in Russia, the group as a whole has worldwide distribution.

Pinnipeds are less modified both anatomically and behaviorally for marine life than are other marine mammals. Each year they return to land to breed, and they have retained their hindlimbs. They are primarily carnivorous mammals, with fish supplying the basic diet. They also eat crustaceans, mollusks, and in some instances sea birds. Their body is covered with a heavy coat of fur, the limbs are modified as flippers, the eyes are large, the external ear is small or lacking, and the tail is absent or very short. Dentition is variable among the families. The intestinal tract is simple, testes are internal in the phocids and external in the otariids, and the males have an os priapi or baculum. *See* COPULATORY ORGAN.

Odobenidae. The single species of this family, the walrus (*Odobenus rosmarus*) [**Fig. 1**], is large, growing to 10 ft (3 m) and weighing 3000 lb (1350 kg). It has no external ears but does have a distinct neck region. The upper canines of both sexes are prolonged as tusks which can be used defensively. The tusks may reach a length of 3 ft (0.9 m) and weigh 11 lb (5 kg). There are 18 teeth. Walrus populations are about equally distributed in shallow waters of the Atlantic and Pacific oceans in the Arctic polar ice regions. Population densities are continually reduced, threatening the walrus with possible extinction. The walrus feeds on marine invertebrates and fish. The teeth can crush the hardest shells.

Otariidae. This family, the eared seals, includes the sea lions and fur seals, which are characterized by an external ear. The neck is longer and more clearly defined than that of true seals (Phocidae), the digits lack nails, and there are 34 teeth with a dental formula I 3/2 C 1/1 Pm 4/4 M 1/1. *See* DENTITION.

The California sea lion (*Zalophus californianus*) occurs along the Pacific coast, and is commonly seen in zoos and performing in circuses. It can move along on land since the hindlimbs can be rotated to act as legs. The forelimbs, used for swimming, may have a span of 6 ft (1.8 m) when extended. These animals have been seen swimming at rates over 20 mi/h (9 m/s). An adult male may grow to 7 ft (2 m) long and weigh up to 600 lb (270 kg); the female is smaller. In the wild, males establish harems in the breeding locales with about 40 females. Young are born in the early summer. After the breeding season, a

Fig. 1. Walrus (*Odobenus rosmarus*). (*Photo from Donna Dewhurst Collection/U.S. Fish and Wildlife Service*)

Fig. 2. Steller sea lion and pup (*Eumetopias jubatus*). (*Photo from Alaska Image Library/U.S. Fish and Wildlife Service*)

migration takes place involving many immature males as well as adults. This permits more food for the mothers and young left behind in the rookery areas.

The southern sea lion (*Otaria byronia*) is found around the Galápagos Islands and along the South American coast. Steller's sea lion (*Eumetopias jubatus*) [**Fig. 2**] is a large species, the male sometimes reaching a length of 13 ft (4 m) and a weight of about 1300 lb (600 kg), and is found along the Pacific coast. Minor species are the Australian sea lion (*Neophoca cinerea*) and the New Zealand species (*Phocarctos hookeri*).

Phocidae. This is the largest family of pinnipeds, the true seals. It includes the monk seals, elephant seals, common seals, and other less well-known forms. These animals have 30 teeth. The family is unique in that the digits have nails, the soles and palms are covered with hair, and the necks are very short. Most species live in marine habitats; however, the Caspian seal (*Pusa caspica*) lives in brackish water. The only freshwater seal is the Baikal seal (*P. sibirica*), whose closest relative is the ringed seal (*P. hispida*). These two species perhaps derived from a single species, geographical isolation having caused the branching. It is estimated that there are between 40,000 and 100,000 Baikal seals in Lake Baikal. These animals are hunted for their valuable fur.

One of the best-known species is the Atlantic gray seal (*Halichoreus grypus*), a species fround in the North Atlantic along the coasts of Europe, Iceland, and Greenland. The leopard seal (*Hydrurga leptonyxi*) is found in southern seas around Antarctica, Australia, and South Africa, where they migrate during the postbreeding winter period. These solitary animals are voracious; they prey on penguins and frequently feed on carrion from whalers. They are able to swim rapidly. The southern elephant seal (*Mirounga leonina*) may weigh as much as 4 tons (3.5 metric tons) and reach more than 20 ft (0.6 m) in length. It is found in the southern Pacific, where it breeds on the rocky shores of the islands. The habits of these seals are among the best known of those of the pinnipeds. The cows arrive at the rookeries in early fall to give birth to a single calf. About 2 months later, the bulls fight for possession of the cows to form the harem. Mating occurs in November, after which the adults return to the sea for food. *See* CARNIVORA; MAMMALIA. Charles B. Curtin

Evolution. The major evolutionary lines of extant pinnipeds, and their close fossil allies, are relatively well documented in the fossil record. The earliest fossils showing unmistakable affinities to modern pinnipeds occur in latest Oligocene through early Miocene (approximately 25 million years old) sediments of the northeast Pacific margin.

The nearly complete skeleton of *Enaliarctos mealsi*, which was recovered in central California, represents the fullest available documentation of the evolutionary transition between modern pelagic pinnipeds and their terrestrial carnivoran ancestry. This taxon displays a combination of primitive and advanced skeletal features closely matching that expected for the most recent common ancestor of living pinnipeds. *Enaliarctos* was a capable swimmer; in addition to hands and feet that were modified into flippers, flexure of the vertebral column served as an important propulsive force.

Enaliarctos and several roughly contemporaneous forms were long considered related strictly to walruses and eared seals among living pinnipeds. True seals (Phocidae) were hypothesized to have originated from a terrestrial group separate from that which gave rise to *Enaliarctos*, walruses, and eared seals. The notion that similar environmental demands had shaped myriad pinniped resemblances repeatedly and independently through the process of convergent evolution was central to this dual (diphyletic) view of pinniped origins. More recent anatomical interpretations and biomolecular evidence, however, suggest that pinnipeds stem from a single and exclusive common origin (that is, they are monophyletic) and that *Enaliarctos* represents an outgroup to all modern forms, including true seals (**Fig. 3**).

Other early Pacific fossil pinnipeds include walruses, eared seals, and a distinctive group of extinct forms apparently closely related to true seals. The earliest walruses are approximately 15 million years in age; the morphological hallmark of the group's

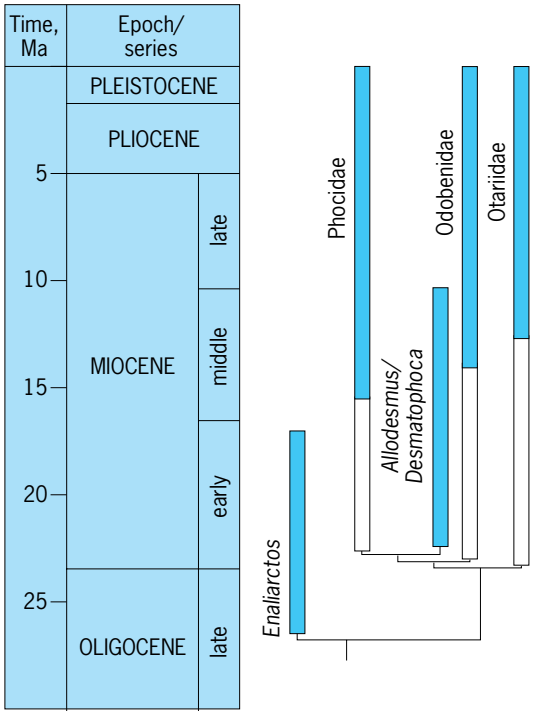

Fig. 3. Evolutionary relationships and stratigraphic ranges of major groups of pinnipeds. Color bars correspond to known ranges, white bars to inferred ranges. Ma = million years ago.

single modern species, tusks, appear only about 5 million years ago. Walruses previously occupied the midlatitudes; their present Arctic distribution is anomalous in this regard.

Allodesmus and *Desmatophoca*, large-bodied, 10–22-million-year-old Pacific forms, are the best-known representatives of an extinct group formerly considered related to walruses. These are currently viewed as more closely related to true seals than to walruses. Walruses, in turn, share a more recent common ancestor with this extinct group plus true seals, than with the eared seals, as was traditionally thought.

Eared seals (sea lions and fur seals) first appear in the fossil record about 10 million years ago, again in the Pacific. By this point the group had already adopted the specialized mode of forelimb swimming distinguishing modern forms.

In contrast to other major groups of pinnipeds, the early evolutionary diversification of phocids seems to have taken place largely in the Atlantic. Diverse lineages of true seals occur by about 15 million years ago, the earliest record of the group. *See* MIOCENE; OLIGOCENE.

Andre Wyss

Bibliography. A. Berta and A. R. Wyss, Pinniped phylogeny, *Proc. San Diego Nat. Hist. Mus.*, 29:33–56, 1993; T. A. Deméré, The family Odobenidae: A phylogenetic analysis of fossil and living taxa, *Proc. San Diego Nat. Hist. Mus.*, 29:99–123, 1993; R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999; D. Renouf (ed.), *Behavior of Pinnipeds*, 1991; M. Riedman, *The Pinnipeds: Seals, Sea Lions and Walruses*, 1991.

Pinophyta

One of the two divisions of the seed plants of about 600 to 700 species extant on all continents except Antarctica. The most familiar and common representatives are the evergreen, cone-bearing trees of the Pinales. Because the ovules (young seeds) are exposed directly to the air at the time of pollination, the Pinophyta are commonly known as the gymnosperms, in contrast to the other division of flowering plants, the angiosperms (division Magnoliophyta), which have the ovules enclosed in an ovary. The division Pinophyta consists of three classes: Ginkgoopsida, Cycadopsida, and Pinopsida. *See* CYCADOPSIDA; GINKGOOPSIDA; MAGNOLIOPHYTA; PINOPSIDA; PLANT KINGDOM.

Thomas A. Zanoni

Bibliography. C. B. Beck, Gymnosperm phylogeny: A commentary on the views of S. V. Meyen, *Bot. Rev.*, 51:273–294, 1985; S. V. Meyen, Basic features of gymnosperm systematics and phylogeny as evidenced by the fossil record, *Bot. Rev.*, 50:1–111, 1984; C. N. Miller, Jr., A critical review of S. V. Meyen's "Basic features of Gymnosperm systematics and phylogeny as evidenced by the fossil record," *Bot. Rev.*, 51:295–318, 1985; G. W. Rothwell, The role of comparative morphology and anatomy in interpreting the systematics of fossil Gymnosperms, *Bot. Rev.*, 51:319–327, 1985.

Pinopsida

The largest and most important class of the division Pinophyta (Gymnospermae), the other classes being Ginkgoopsida and Cycadopsida. There are two orders: the Cordaitanthales, with three extinct families, and the Pinales (Coniferales), with six extinct families and seven families with some extant genera.

The living Pinopsida are woody plants; most are trees with a central axis and excurrent branches. Leaves are simple, alternate, or opposite or in whorls, scalelike or needlelike or rarely planar. The wood lacks vessels and usually has resin canals. Male reproductive structures are aggregated on microsporophylls directly attached to the cone axis. The ovules are borne in compound cones or singly or paired at the end of a stalk (Taxaceae). The main seed plane is tangential to the cone axis (if the seed scale is regarded as a modified dwarf shoot). The embryo has two or more cotyledons. *See* CYCADOPSIDA; GINKGOOPSIDA; PINALES; PLANT KINGDOM.

Thomas A. Zanoni

Pionium

The electromagnetically bound or Coulomb-bound pion-pion system, the $\pi^+\pi^-$ atom, also called $A_{2\pi}$. Nowadays, this doubly exotic system is called pionium. In some older literature the term pionium

was used for atoms consisting of only one pion and a muon or an electron. Eighteen pion-muon ($\pi\mu$) atoms, formed in the decay $K_L^0 \rightarrow (\pi\mu)_{\text{atom}}\ \nu$ were detected in 1976 in an experiment at Brookhaven National Laboratoty in Upton, New York. *See* ELEMENTARY PARTICLE; HADRONIC ATOM; LEPTON; MESON; PARTICLE ACCELERATOR.

The $A_{2\pi}$ system is under extensive investigation in the DIRAC (DI-meson Relativistic Atom Complex) experiment at CERN (Conseil Européen pour la Recherche Nucléaire) in Geneva, Switzerland. The first $A_{2\pi}$ atoms were observed at the Serpukhov U-70 synchrotron in Russia in 1993. Although only about 270 $A_{2\pi}$ atoms were detected, this result made it possible to set a lower limit of 1.8 fs (1 femtosecond = 10^{-15} s) on the $A_{2\pi}$ lifetime (90% confidence limit).

In the DIRAC experiment (which includes physicists who were involved in the Serpukhov experiment), a 24-GeV proton beam from the Proton Synchrotron at CERN is directed onto a thin nickel target to produce pionium in a proton-nuclear reaction. A portion of these $A_{2\pi}$ atoms then break up (ionize) in the same target into positive and negative pions, which are subsequently detected by a double-arm magnetic spectrometer. This spectrometer has been specially designed to distinguish pairs of pions that originate in pionium breakup from the huge background of "free" pairs also produced in that target.

In 2001, the DIRAC collaboration was able to collect and analyze more than 6000 $A_{2\pi}$ breakups. This sample allowed the $A_{2\pi}$ lifetime to be determined to be 2.9 fs within a statistical error of 15%.

The strong pion-pion interaction in $A_{2\pi}$ is described by low-energy quantum chromodynamics (QCD), since the pions bound in the atom have low momenta. Theory predicts that pionium decays, mainly into two neutral pions, and has a lifetime of (2.9 ± 0.1) fs. This $A_{2\pi}$ lifetime depends on pion-pion scattering lengths. Low-energy QCD in the framework of chiral perturbation theory provides scattering lengths with a very high accuracy of less than 2% error. Therefore, a lifetime measurement such as that of the DIRAC experiment will check sensitively the understanding of chiral symmetry breaking in QCD. *See* QUANTUM CHROMODYNAMICS; SCATTERING EXPERIMENTS (NUCLEI).

The measured lifetime given above corresponds to an accuracy in the determination of the scattering length at the 10% error level. The DIRAC team hopes to reduce this error by a factor of 2 or even more.

Juerg Schacher

Bibliography. B. Adeva et al., First measurement of the $\pi^+\pi^-$ atom lifetime, *Phys. Lett.*, B619:50–60, 2005; L. G. Afanasyev et al., Observations of atoms consisting of π^+ and π^- mesons, *Phys. Lett.*, B308:200–206, 1993; G. Colangelo, J. Gasser, and H. Leutwyler, $\pi\pi$ scattering, *Nucl. Phys.*, B603:125–179, 2001; R. Coombes et al., Detection of π-μ Coulomb bound states, *Phys. Rev. Lett.*, 37:249–252, 1976; J. Gasser et al., Decays of the $\pi^+\pi^-$ atom, *Phys. Rev.*, D64:016008, 2001; L. L. Nemenov, Elementary relativistic atoms, *Sov. J. Nucl. Phys.*, 41:629, 1985 (*Yad. Fiz.*, 41:980–990, 1985).

Pipe flow

Conveyance of fluids in closed circular ducts. Pipes have been used for thousands of years, but the understanding of the flow details (velocity profile and flow rate) came about during the twentieth century. Equations that accurately predict energy losses have been in use since the midnineteenth century.

Flow in closed conduits is probably the most common way of transporting fluids. Crude oil and its components are moved through pipes in a refinery. Water in the home is transported through tubing. Heated and conditioned air is distributed to all parts of a dwelling in circular or rectangular ducts. *See* PIPELINE.

Flow details. Flow in a closed conduit (circular or otherwise) can be either laminar or turbulent. In laminar flow, the fluid particles move smoothly through the duct in layers called laminae. A fluid particle in one layer stays in that layer. In turbulent flow, flowing fluid particles move tortuously about the cross section, resulting in an effective mixing action. Eddies and vortices are responsible for the mixing, which does not occur in laminar flow. Turbulent flow exists at much higher flow rates than laminar flow.

The criterion for distinguishing between laminar and turbulent flow is this observed mixing action. When injected into a laminar flow in a duct, a dye moves downstream in a threadlike line. When injected into a turbulent flow, a dye disperses quickly. Experiments have shown that laminar flow exists when the dimensionless Reynolds number, $\text{Re} = VD/\nu$, is less than 2100; here V is the average velocity, D is the inside diameter of the pipe or tube, and ν is the kinematic viscosity (a property of the fluid). *See* LAMINAR FLOW; REYNOLDS NUMBER; TURBULENT FLOW.

Laminar flow. The axial velocity V_z in laminar flow is given by Eq. (1), in which Δp is the pressure drop

$$V_z = \frac{\Delta p r_0^2}{4\mu L}\left(1 - \frac{r^2}{r_0^2}\right) \tag{1}$$

over the length L, μ is the absolute viscosity (a property of the fluid), r_0 is the inside radius of the duct, and r is the radial coordinate. Equation (1) is of a parabola (**Fig. 1**). The volume flow rate, or discharge, Q, is given by Eq. (2), in which D is the inside diameter of the tube.

$$Q = \frac{\Delta p \pi D^4}{128\mu L} \tag{2}$$

Transition. When the Reynolds number is greater than 2100 but less than 4000, the flow can exhibit laminar, turbulent, or intermittent flow characteristics; the exact transition velocity depends on the nature of the piping system.

Turbulent flow. For very high Reynolds numbers (greater than 4000), instabilities are propagated and form eddies and vortices responsible for the observed mixing action in turbulent flow. Fluid particles move in random paths with large transverse velocity components. The velocity profile in

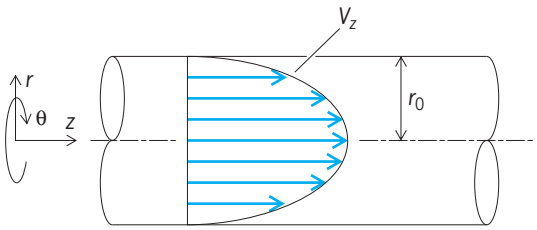

Fig. 1. Laminar velocity profile.

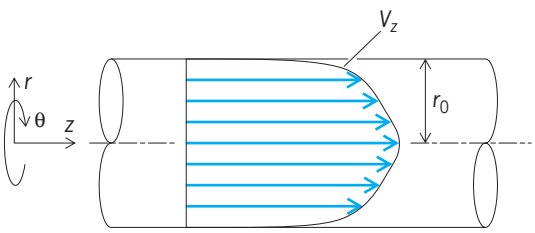

Fig. 2. Turbulent velocity distribution.

turbulent flow is more uniform than that for laminar flow due to the large transfer of momentum radially across the cross section of the conduit (**Fig. 2**). An equation that gives a fairly accurate representation of the profile is the one-seventh power law, Eq. (3), in

$$V_z \approx V_{\max}\left(1 - \frac{r}{r_0}\right)^{1/7} \tag{3}$$

which V_z is the axial velocity, $V_{\max}$ is the velocity at the pipe centerline, r_0 is the pipe radius, and r is the radial coordinate. The average (turbulent) velocity, V_{avg}, is given by Eq. (4).

$$V_{\text{avg}} \approx \frac{49}{60} V_{\max} \tag{4}$$

Energy loss and friction factor. The energy loss experienced by the fluid is manifested as a pressure drop Δp, which is found in terms of a dimensionless friction factor f as shown in Eq. (5). Here, the

$$\Delta p = \frac{fL}{D}\frac{\rho V^2}{2} \tag{5}$$

density ρ is a property of the fluid, L is the length over which the pressure drop occurs, D is the inside diameter of the pipe or tube, V is the average velocity, and the quantity $\rho V^2/2$ is the kinetic energy of the flow per unit volume. For laminar flow through a circular duct, the friction factor is found in terms of the Reynolds number, as indicated in Eq. (6).

$$f = \frac{64}{\text{Re}} \qquad \left(\begin{array}{c}\text{laminar flow,}\\ \text{circular duct}\end{array}\right) \tag{6}$$

For turbulent flow, the friction factor is dependent upon the wall roughness, the fluid properties, the average velocity, and the pipe diameter; that is, $f = f(V, D, \rho, \mu, \varepsilon)$, in which ε is a measure of the absolute roughness of the conduit wall, having the dimension of length. Values of the roughness ε have been measured for many commercial pipe materials (see **table**). The friction factor for turbulent flow may be obtained from a Moody diagram in which the friction factor f is graphed as a function of the Reynolds number Re, with the relative roughness ε/D as an independent parameter (**Fig. 3**). The Moody diagram is a result of many flow rate and pressure drop measurements made on commercial pipe and tube materials.

A number of equations have been devised to curve-fit the Moody diagram for Reynolds numbers greater than 2100. The older equations require an iterative process when trying to calculate the friction factor f for specified values of the Reynolds number Re and the relative roughness ε/D. More recently published equations overcome this difficulty. An example is the Haaland equation (7), which gives an explicit

$$f = \left\{-0.782 \ln\left[\frac{6.9}{\text{Re}} + \left(\frac{\varepsilon}{3.7D}\right)^{1.11}\right]\right\}^{-2} \quad \text{(turbulent flow)} \tag{7}$$

solution for f in terms of Re and ε/D.

For turbulent flow in a smooth pipe, $\varepsilon = 0$, and the expression for the friction factor f reduces to $f = f(\rho VD/\mu)$, in which $\rho VD/\mu$ is the Reynolds number. This relationship is shown as the bottom line on the Moody diagram. For turbulent flow in a rough pipe, the friction factor varies with the Reynolds number and the relative roughness ε/D, represented as all the other lines on the Moody diagram.

As a pipe ages, the wall roughness usually tends to increase linearly with time according to $\varepsilon' = \varepsilon + \alpha t$, in which ε is the roughness of the pipe material when new, ε' is the roughness after a certain time t has elapsed, and α is a constant that must be determined by experiments performed on the specific pipe material–fluid combination.

Calculation methods. Given values of pipe length, diameter, roughness, fluid properties, and volume flow rate, it is possible to find the head loss or the pressure drop using Eq. (5). The Reynolds number

Roughness factors for commercial pipe materials

Pipe material	ε, ft (cm)
Cast iron	0.00085 (0.025)
Asphalt-coated	0.0004 (0.012)
Cement-lined	0.000008 (0.00025)
Drawn tubing	0.000005 (0.00015)
Minerals	
Brick sewer	0.001–0.01 (0.03–0.3)
Cement-asbestos	
Clays	
Concrete	
Miscellaneous	
Brass	0.000005 (0.00015)
Copper	
Glass	
Lead	
Plastic	
Tin	
PVC	Smooth
Steel	
Commercial	0.00015 (0.0046)
Galvanized	0.0002–0.0008 (0.006–0.025)
Wrought iron	0.00015 (0.0046)

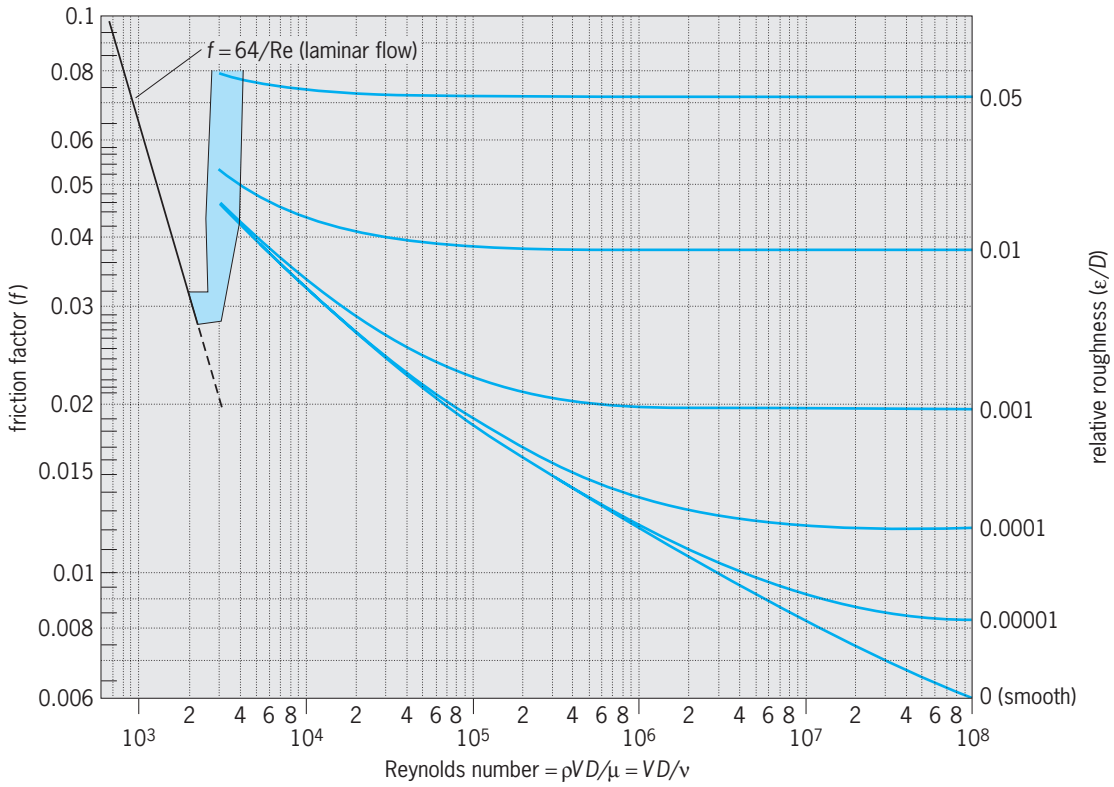

Fig. 3. Moody diagram, showing curves for laminar and turbulent flow.

and relative roughness are calculated and then used with the Moody diagram (or curve-fit equations) to compute the friction factor. Substituting into Eq. (5) gives the pressure drop.

When the pressure drop is known and the volume flow rate is to be determined, an iterative calculation scheme is required when using the Moody diagram to find the friction factor f. An assumed value for f is obtained from the Moody diagram for the given value of the relative roughness ε/D. It is used to find a trial value for V, which in turn is used to find a Reynolds number. The values of the Reynolds number Re and the relative roughness ε/D can then be used to find a better value for f, and the process is repeated until the calculations converge.

Gases. When a gas or a vapor is the flowing fluid, compressibility effects may have to be considered. For low gas velocities of less than 30 m/s, the calculation methods described in the preceding paragraphs can be used. Otherwise, an equation of state must be used to relate changes in density with changes in pressure and temperature.

Minor losses. Pipes and tubes are joined together with fittings such as elbows, T-joints, and valves. The loss of pressure when a fluid flows through a fitting is called a minor loss, which varies with the kinetic energy of the flow. William S. Janna

Bibliography. W. S. Janna, *Introduction to Fluid Mechanics*, 3d ed., 1993; R. W. Jeppson, *Analysis of Flow in Pipe Networks*, 1976; J. H. Roberson and C. T. Crowe, *Engineering Fluid Mechanics*, 6th ed., 1996; J. P. Tullis, *Hydraulics of Pipelines*, 1989.

Pipeline

A line of pipe and associated equipment for conveying materials. Pipelines are used for transportation of water, gases, vapors, slurries, sewage, sludge, petroleum, chemicals, foodstuffs, and other process media (**Fig. 1**). They comprise pipe, fittings, flanges, bolting, gaskets, valves, pressure-relieving devices, pumps, specialty items such as strainers, flow nozzles, expansion joints, and supports and hangers.

A tube has a cross section that may or may not be round and has a continuous periphery. A pipe is a tube with a round cross section, and can range in diameter from as little as $^1/_8$ in. (3 mm) to over 10 ft (3 m).

Wall thickness. Pipe is specified by nominal pipe size (NPS) and schedule or minimum wall thickness. The nominal pipesize identifies the nominal inside diameter for sizes from $^1/_8$ in. (3 mm) through 12 in. (305 mm) and the actual outside diameters for 14-in. (356-mm) and larger sizes. Along with the nominal pipe size, the pipe wall thickness is stated by schedule number, such as 8-in. (203-mm) nominal pipe size, schedule 40. The schedule number identifies the nominal wall thickness, which in case of 8-in. (203-mm) nominal pipe size, schedule 40, is 0.322 in. (8.2 mm).

When iron pipes were initially standardized, there was only one standard wall thickness for each nominal size of pipe. To satisfy the need for stronger and thicker wall pipes to transport fluids under higher operating pressure, larger uniform thicknesses were designed as extra strong (XS) and double extra strong

Fig. 1. Pipelines. (*a*) Construction of 300-mi, 24-in. (500-km, 60-cm) pipeline in Colombia that runs over the Andes to the Caribbean coast. (*b*) Construction of 30-in. (75-cm) gas pipeline in Pennsylvania in 1982. (c) Trans-Arabian oil pipeline, built in 1951 to link the Persian Gulf and the Mediterranean Sea. (*Bechtel Group, Inc.*)

(XXS), and original thicknesses were retained as standard (STD). These designations have been replaced or supplemented by schedule numbers. Schedule numbers begin with 5 and 5S, followed by 10 and 10S, then progress in increments of 10 through 40, and then in increments of 20 through 160.

Pipe made of special alloys for high-pressure and high-temperature service is very costly and, therefore, is specified by inside diameter and minimum wall thickness rather than nominal pipe size and schedule. Such pipes usually have a nonstandard outside diameter. The special pipe sizes are usually ordered when the material cost savings offset the additional cost of making a special setup for production.

Tube size is generally defined by outside diameter and wall thickness. However, it can be specified by outside diameter and inside diameter, or inside diameter and wall thickness.

Fittings and joints. Fittings, such as elbows, tees, crosses, reducers, returns, flanges, and couplings, are used to join different pipe segments; provide changes in direction or size; install branch lines and instrument connections; and connect pipes with in-line equipment.

Pipe sections, fittings, and in-line equipment are usually joined by welded joints, soldered or brazed joints, flanged joints, threaded joints, bell-and-spigot joints, grooved joints, coupling joints, caulked joints, and other mechanical means. Welded, soldered, or brazed joints are used where maintenance or repair is expected to be minimal. Mechanical joints, such as flanged, threaded, grooved, or coupling joints, offer ease of assembly and disassembly to permit maintenance and repair.

Materials. Metallic pipe and fittings made from cast iron, ductile iron, carbon steel, stainless steel, copper, brass, aluminum, and other metals and alloys are generally used for construction of pipelines or piping systems. Nonmetallic pipes, especially thermoplastic and thermosetting pipes, are excellent for handling corrosive fluids, chemicals, brackish water, and so forth. They are economical but are more susceptible to damage. Vitrified or unglazed clay pipes are used in drainage systems and to convey sewage, industrial waste, and storm water. Reinforced or unreinforced concrete pipe, made from cement, sand, gravel, and water, is used for conveying water, including storm water and cooling water. Because it has good resistance to chemical attack, glass pipe is used for laboratory drainage, chemicals, beverages, food, and pharmaceuticals. Metallic pipes lined with cement-mortar, rubber, epoxy, ceramic, or other materials are used to transport corrosive and abrasive materials. The lining provides protection for the metallic pipe against the corrosive, erosive, and abrasive effects of the flow medium. Also, the linings can be repaired or replaced.

Design. Safety is the primary objective in pipeline design. The codes, standards, or specifications that have jurisdiction over a pipeline or piping system depend on the site of installation, type of service, design considerations, and other requirements. For example, in the United States the water treatment and distribution piping systems in cities and towns are designed and constructed to the standards and specifications published by the American Water Works Association (AWWA), whereas the water-carrying piping systems at a pressure greater than 30 lb/in.^{2} (207 kilopascals) within the boundary of a coal-fired power plant must meet the design requirements of the American Society of Mechanical Engineers Pressure Piping Code. However, the water pipelines for the fire protection systems within the same power plant must be designed in accordance with the requirements of the applicable National Fire Protection Association (NFPA) standards.

Material selection. The materials selected for the pipe, fittings, and in-line equipment must be suitable for the application requirements, including pressure, temperature, and chemical properties of the conveyed fluid, and the environmental conditions. Furthermore, the materials must perform the service safely for the design life of the piping system.

Sizing. Depending upon the required flow rate and the velocity of flow, the cross-sectional area of the pipe is calculated from Eq. (1*a*) in British

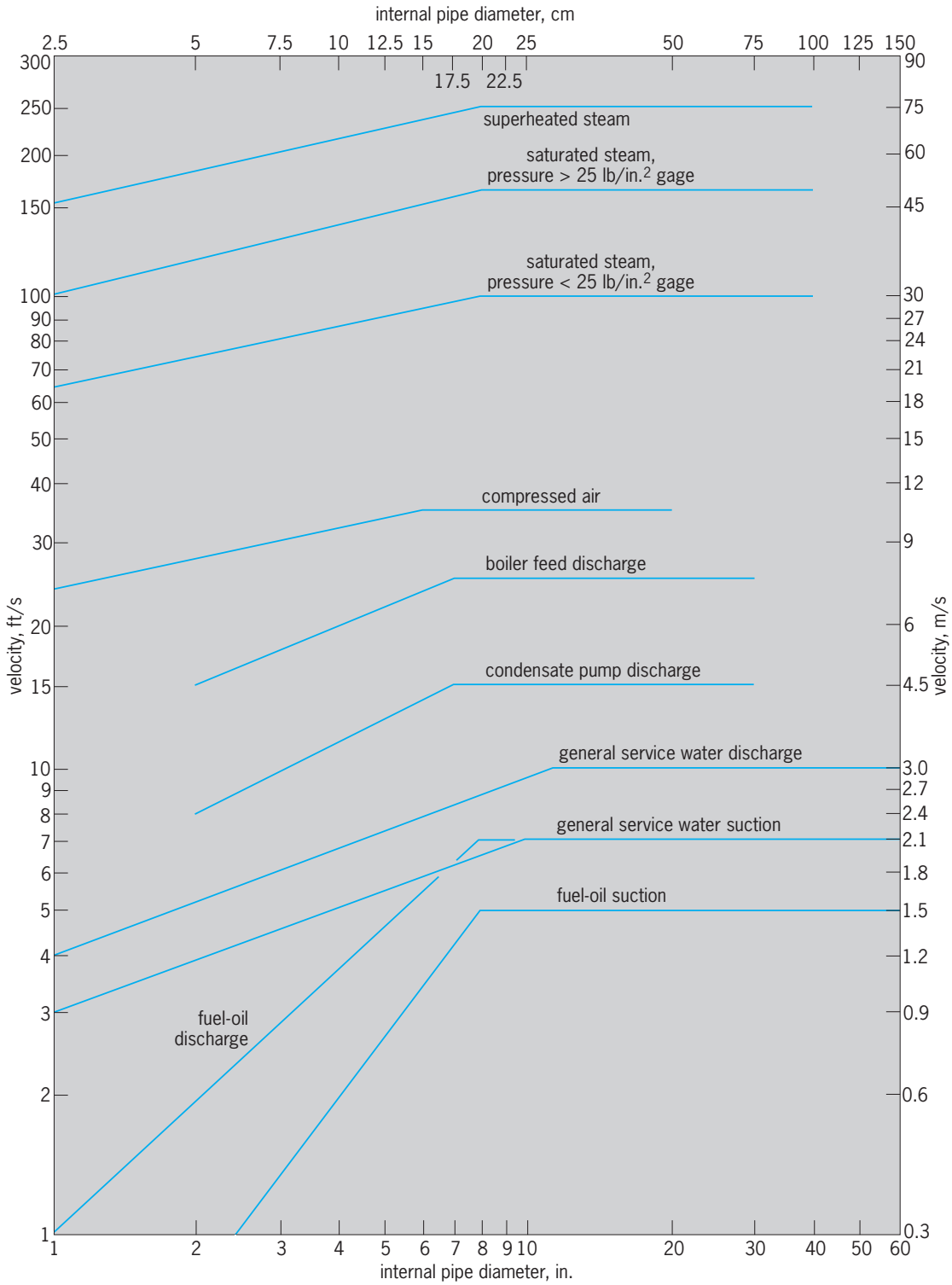

Fig. 2. Velocity of pipe flow as function of internal pipe diameter for various applications. 25 lbf/in.^{2}, gage pressure = 172 kPa above atmospheric pressure.

engineering units and Eq. (1*b*) in metric units. The

$$\text{Flow area (ft}^2) = \frac{\text{flow rate (ft}^3/\text{s)}}{\text{mean velocity of flow (ft/s)}} \quad (1a)$$

$$\text{Flow area (m}^2) = \frac{\text{flow rate (m}^3/\text{s)}}{\text{mean velocity of flow (m/s)}} \quad (1b)$$

standard pipe size that will give the required flow area is selected. **Figure 2** shows the ranges of velocities generally used. However, the actual velocity to be used for a pipeline may differ because of specific design requirements.

The minimum pipe wall thickness required to withstand the sustained fluid pressure can be calculated by Barlow's formula (2), where t_m is the

minimum wall thickness in inches; P is the fluid pres-

$$t_m = \frac{PD_o}{2S} \tag{2}$$

sure in pounds per square inch; D_o is the outside pipe diameter in inches; and S is the allowable stress for the pipe material at the fluid temperature in pounds per square inch. (Metric or other units may also be used in the equation, so long as t_m and D_o are in the same units, and P and S are also in the same units.)

Most piping codes contain a modified version of Barlow's formula or other guidelines that must be complied with to determine the minimum wall thickness. The calculated minimum wall thickness must be increased to provide for corrosion or erosion; threading or grooving; thinning caused by bending; mechanical strength to prevent damage, collapse, or buckling, of pipe due to external loads; and manufacturing process tolerance, whenever these factors are applicable.

Pressure drop. The flow in a pipe can be laminar, turbulent, or a combination of laminar and turbulent, depending upon the value of a dimensionless number known as the Reynolds number. This quantity can be determined by Eq. (3), where D is

$$\mathrm{Re} = \frac{Dv\rho}{\mu} \tag{3}$$

the internal diameter of the pipe in feet (meters); ρ is the weight density of fluid in pounds per cubic foot (kilograms per cubic meter); v is the mean velocity of flow in feet per second (meters per second); and μ is the absolute viscosity in pound-mass per foot-second or poundal seconds per square foot (kilograms per meter-second or pascal seconds). Usually the flow in a pipe is laminar if the Reynolds number is less than 2000, and turbulent if the Reynolds number is greater than 4000. Flows with Reynolds numbers between 2000 and 4000 are in a transition zone. *See* LAMINAR FLOW; PIPE FLOW; REYNOLDS NUMBER; TURBULENT FLOW.

The pressure drop due to friction can be determined by the Darcy equation (4) for laminar or

$$h_L = \frac{fL}{D}\frac{v^2}{2g} \tag{4}$$

turbulent flow of any liquid in a pipe. Here h_L is the pressure drop due to friction, in feet (meters)

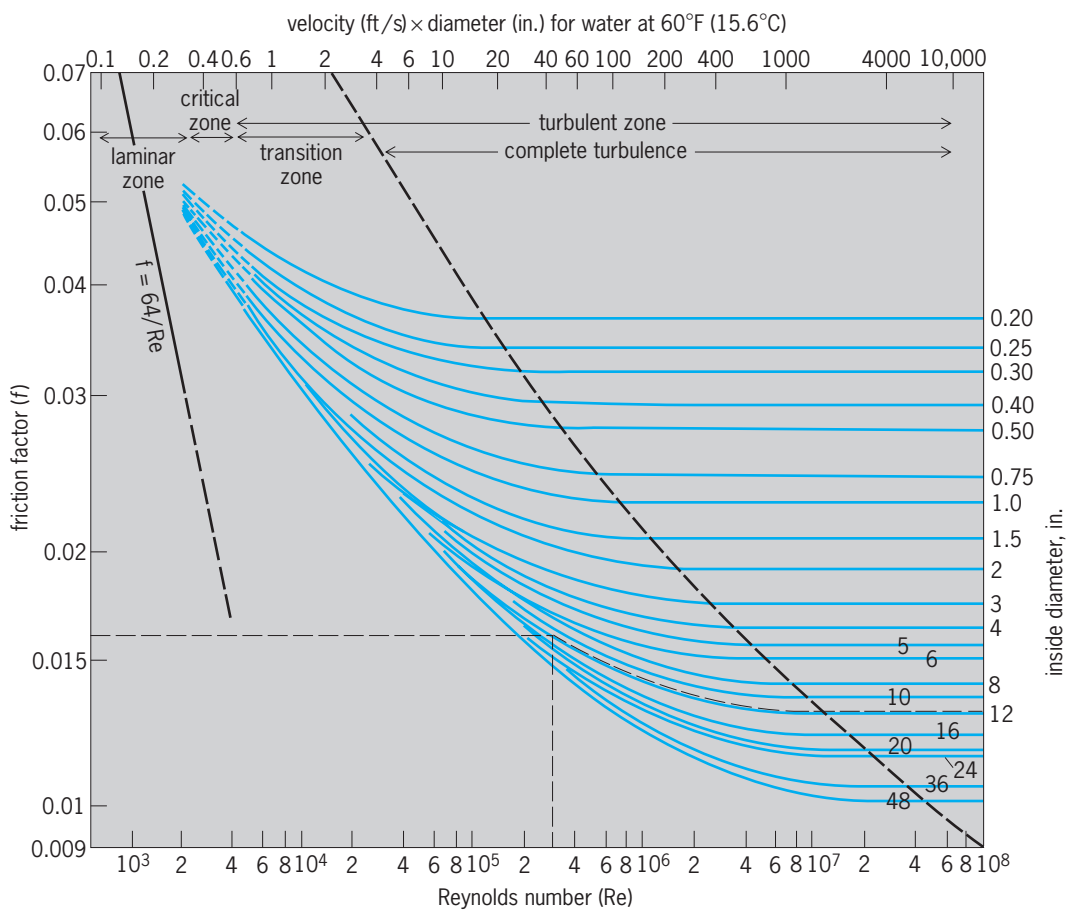

Fig. 3. Friction factor as function of Reynolds number for clean commercial steel pipes with various inside diameters. For example, pipe with inside diameter of 12 in. (305 mm) and flow having Reynolds number of 300,000 has friction factor of 0.016. Upper scale gives product of flow velocity and inside diameter that corresponds to Reynolds number for water at 60°F (15.6°C). 1 in. = 25.4 mm. Velocity in meters per second × diameter in centimeters = 0.77 × velocity in feet per second × diameter in inches.

of liquid; L is the length of pipe, in feet (meters); g is the acceleration of gravity, 32.2 ft/s^2 (9.8 m/s^2); and f is the friction factor, a dimensionless quantity.

The factor fL/D is referred to as the resistance coefficient, and can be determined for straight runs of pipe and various in-line fittings, such as elbows, reducers, and valves. Summation of individual pressure losses through straight runs of pipe, fittings, and other components provides the pressure drop between any two locations in a pipeline due to friction. In calculating pressure losses between two points in a pipeline, the increase or decrease in elevation must be accounted for as loss or gain in feet (meters) of liquid. The entrance and exit pressure losses, as applicable, must be included to determine the total pressure drop.

The friction factor f for laminar flows, that is, when the Reynolds number is less than 2000, can be calculated from Eq. (5). For turbulent flows, the friction

$$f = \frac{64}{\mathrm{Re}} \tag{5}$$

factor f is dependent upon Reynolds number and the relative roughness of the pipe. **Figure 3** can be used to determine the friction factor from the Reynolds number and the size of the pipe.

Pumping requirements. Pumps are utilized to cause flow against gravity; move liquids from one location to another; overcome pressure losses due to friction, changes in elevation, and other obstructions; and supply liquid at a desired pressure. The energy imparted by a pump is usually expressed in terms of head, in feet (or meters) of liquid. Many different types of pumps (centrifugal, reciprocating, and rotary) are available to add energy to liquids. Most commonly used is the centrifugal pump. A simple centrifugal pump consists of an impeller in a volute casing with an opening in the center for liquid to enter and a peripheral opening for liquid to exit. The impeller is rotated at high speed by an electric motor or turbine. The impeller vanes impart high velocity to the entering liquid, which is then discharged by the centrifugal force. *See* CENTRIFUGAL PUMP; PUMP.

When the liquid level in the source of supply is below the centerline or impeller eye of the pump, then the pump has to lift the liquid and overcome the friction and entrance losses from the point of supply to the centerline of the pump. This duty of the pump is termed suction lift. When the liquid supply level is above the pump centerline, then the suction head is equal to the static head above the pump centerline minus the suction-line losses.

The total head of a pump must be sufficient to overcome suction lift, friction losses in the discharge pipeline, and pressure losses due to other causes and equipment; and to supply liquid at the required pressure at the point of discharge. The hydraulic power required to develop the total head depends upon the flow rate and the total head. Since no pump is 100% efficient, the brake power of a pump must be greater than the hydraulic power; the brake power is the hydraulic power divided by the pump efficiency.

Other considerations. In designing pipelines so that pressure integrity is maintained, consideration must be given to weight effects, thermal expansion and contraction loads, relief and safety-valve thrust, vibrations, earthquake effects, and impact loads.

Construction. Subassemblies or pipe spools of large piping are generally fabricated in the shop and joined together in proper sequence in the field (Fig. 1*a* and *b*). Smaller piping is usually routed, fabricated, and erected in the field. Quality control is maintained by step-by-step implementation of preapproved fabrication, installation, and examination procedures. Piping joints may require examination by nondestructive methods.

Testing and inspection. After erection, pipelines are cleaned by flushing with water, steam, air, or water mixed with chemical cleaning agents. The pressure integrity of pipeline is then tested by hydrostatic or pneumatic testing. Leaks, if any, are repaired and the pipeline is retested.

Mohinder L. Nayyar

Bibliography. *Flow of Fluids*, Crane Tech. Pap. 410, 1988; Ingersoll-Rand Company, *Cameron Hydraulic Data*, 17th ed., 1992; K. K. Kienow (ed.), *Pipeline Design and Installation*, 1990; M. L. Nayyar (ed.), *The Piping Handbook*, 7th ed., 1999; M. P. Sharma and U. S. Rohatgi (eds.), *Multiphase Flow in Wells and Pipelines*, 1992; J. P. Tullis, *Hydraulics of Pipelines*, 1989.

Piperales

A small order of flowering plants (3600 species) in the eumagnoliid group, which is composed of three anomalously woody vines (shrubs) or herbaceous families—the pipeworts (Aristolochiaceae), the black pepper family (Piperaceae), and the lizard's tail family (Saururaceae). The last two families have reduced flowers in dense spikelike flower stems, and the first has medium-sized to enormous flowers that often trap insects for a period before releasing them, covered with pollen.

Black pepper comes from *Piper nigrum* and betel nuts from *P. betle*. Several species of *Aristolochia* have medicinal properties, and some genera in each of these families are commonly grown ornamentals in the temperate zones or house plants, such as *Asarum* (wild ginger), *Peperomia* (pepper elders), and *Houttuynia*. *See* EUMAGNOLIIDS; LAURALES; MAGNOLIALES; MONOCOTYLEDONS.

Mark Chase

Pisces (constellation)

The Fishes, a large zodiacal constellation in the northern sky. None of the stars in the constellation is very bright, but many are joined in the shape of a large V, making them easy to notice. The constellation also contains the Circlet of stars. *See* ZODIAC.

In Greek mythology, Venus and Cupid, or another mythical couple, escaped from a monster by disguising themselves as these fish. They jumped into the Euphrates River to make their escape.

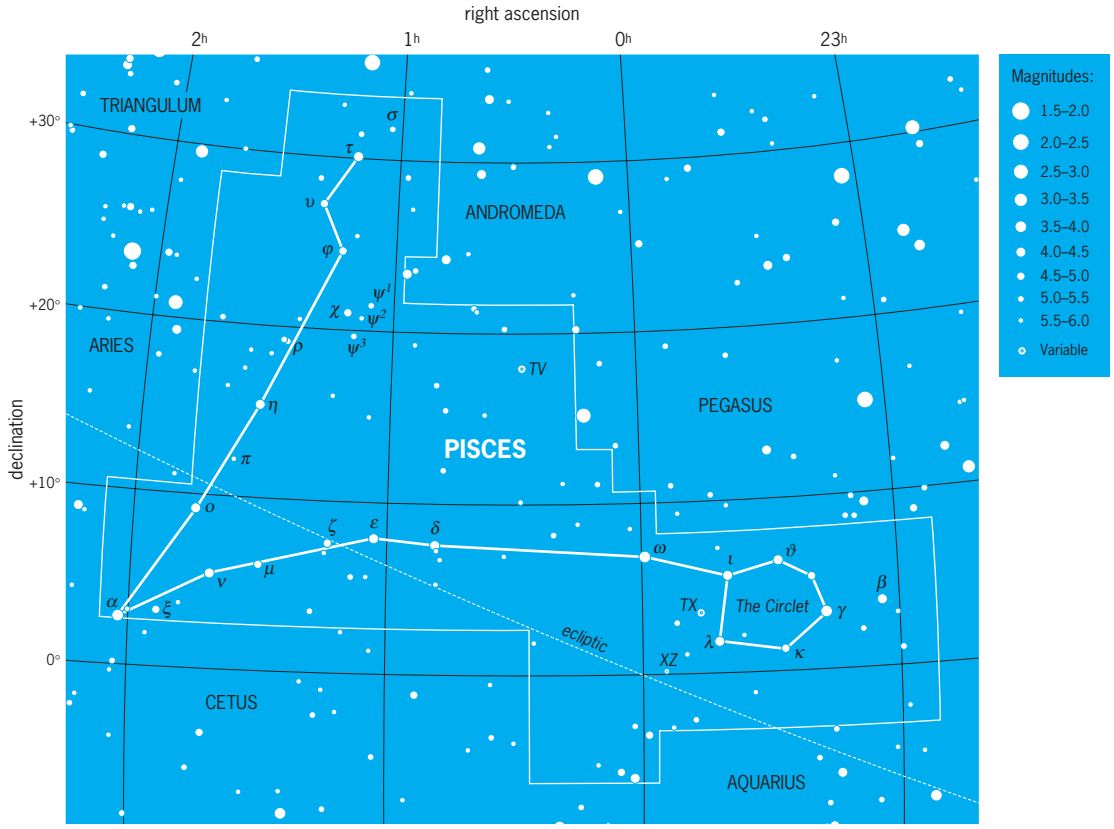

Modern boundaries of the constellation Pisces, the Fishes. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Pisces (zoology)

A term that embraces all fishes and fishlike vertebrates. In early zoological classifications fishes, like mammals, birds, reptiles, and amphibians, were ranked as a class of the vertebrates. As knowledge of fishes increased, it became apparent that, despite their common possession of gills and fins and their dependence on an aquatic environment, not all fishes were closely related. At least five groups of fishes with modern descendants were already established before the tetrapods appeared. Not only are these groups older, but some are decidedly more divergent structurally than are the four classes of tetrapods. For these reasons several classes of fishes are now recognized. The number of classes varies; one reputable but extreme classification recognizes 11 classes of fishes.

The primary cleavage in vertebrate classification is that separating the jawless fishes, or Agnatha, from those vertebrates with jaws, the Gnathostomata. After recognition of this split, the name Pisces was commonly restricted to the jawed fishes. When these in turn were divided into two or more classes, Pisces was further restricted by some authorities to the bony fishes. Another scheme involves assignment of class names to each of the major constituent groups of jawed fishes, and use of Pisces as a superclass name. In view of the confusion, it seems best to revert to early practice and to employ Pisces as a group name of convenience to embrace all classes of fishlike vertebrates, from jawless fishes to bony fishes. In this sense it has no actual taxonomic status because it cuts across natural classification, dividing the gnathostomes and grouping part of them with the agnaths. *See* GNATHOSTOMATA; JAWLESS VERTEBRATES.

The Pisces include four well-defined groups that merit recognition as classes: the Agnatha or jawless fishes, the most primitive; the Placodermi or armored fishes, known only as Paleozoic fossils; the Chondrichthyes or cartilaginous fishes; and the Osteichthyes or bony fishes. Furture research may demonstrate the need for further division, but this is most likely to involve Paleozoic groups. *See* CHONDRICHTHYES; OSTEICHTHYES; PLACODERMI.

Number of recent species. Present fish classification is not sufficiently precise to permit an accurate tabulation of the number of living species. New kinds are constantly being discovered, others are being synonymized as the result of new research, and the literature is scattered. Nevertheless,

estimates by competent ichthyologists are so diverse, ranging from 18,000 to 40,000, that an effort is here made to arrive at a reasonably acceptable approximation. There is indication that most previous estimates are too high. Counts for Recent groups and species include Agnatha, 2 families, about 11 genera, and approximately 45 species; Chondrichthyes, 31 families, some 132 genera, and roughly 575 species; and Osteichthyes, 32 orders, about 357 families, 3570 genera, and about 17,600 species. This last figure broken down gives about 5100 species in 29 orders, nearly 2000 species in the Siluriformes, 3000 species in the Cypriniformes, and 7500 species in the Perciformes. It is to be emphasized that the classification of many groups, especially those in the deep seas and in the tropics, is imperfectly known. Thus, further study will permit refinement of this enumeration of just under 18,300 recognized species of Recent fishes.

Ecology. Fishes live in almost all permanent waters to which they have been able to gain access. In general they have evolved a body conformation and specialized features that adapt them harmoniously to the world about them. Inhabitants of mountain torrents may have peculiar attachment organs; those living in Antarctic waters at a temperature below freezing have made needed physiological adjustments; fishes of the deep sea commonly carry their own light source, and the female anglerfish is assured a mate by the parasitism of the male on her body. In the East Indies some fishes skip with ease over mud flats, and others undertake nightly forays on land. Some fishes mature at extremely small size. A Philippine goby, *Mistichthys*, reaches a length of only 0.5 m (1.25 cm) and is commercially important as a food fish although it takes 31,500 fish to make 1 lb (70,000 fish to make 1 kg); and a Samoan fish, *Schindleria*, attains a weight of only 6 mg. At the other extreme, the whale shark is reputed to reach a length of 60 ft (20 m), and a 38-ft (13 m) individual weighed more than 13 tons (11.7 metric tons).

Adaptive radiation. Because their bodies are supported by water, fishes have been afforded the luxury of diversification in body form not possible for terrestrial animals. A deep, pancake-thin body is not uncommon, and an eellike form has been independently developed in many phyletic lines. Trunkfishes are enclosed in a boxlike casque, and some deep-sea fishes have eyes at the tips of elongate stalks. Long, trailing fins are frequent; sargassum fishes develop appendages that serve as holdfasts and for concealment.

Food habits. Most fishes are more or less carnivorous and predatory, but there is wide diversity in food habits. Many fishes have numerous slender gill rakers with which they strain microorganisms from the water; others have massive teeth and strongly muscled jaws to aid in crushing mollusks or crustaceans. Browsers, nibblers, and grazers employ specially adapted teeth and jaws to scrape vegetation or small attached animals. Some wrasses pluck parasites from larger fishes, and lampreys parasitize other fishes.

Reproductive habits. Reproductive habits are no less varied than feeding behavior. Most fishes are oviparous and scatter their eggs, but nest building and parental care assume a broad spectrum—from a prepared pile of pebbles, through a grassy spherical retreat, to oral incubation or development of a marsupial pouch on the underside of the male pipefish. Viviparity and ovoviviparity have originated along independent lines. Enormously complicated modifications of the anal fin have been evolved to effect insemination of some species in which the young are born alive.

Economics. Fishes play an important role in the lives of most people. Fishing is a way of life in most primitive cultures, and ranks high among recreational activities in highly civilized peoples. Maintenance and care of home aquarium fishes provide an avocation for probably millions of people. An occasional swimmer is killed by sharks; more people die from ciguatera contracted from eating poisonous fish flesh; and venomous fishes take a limited toll in human life and suffering. *See* CHORDATA; VERTEBRATA.

Reeve M. Bailey

Pistachio

A tree, *Pistacia vera*, of the Anacardiaceae family. It is native to central Asia and has been grown for its edible nuts throughout recorded history in Iran, Afghanistan, and Turkey and various other countries of the Mediterranean region. Extensive areas in the Sacramento and San Joaquin valleys of California were planted to pistachios in the 1970s, and the first commercial nut crop was harvested in 1977. *See* SAPINDALES.

The pistachio tree, relatively slow-growing, reaches a height and spread of 20–25 ft (6–8 m). It thrives under long, hot summers with low humidity, but needs moderately cold winters to satisfy its chilling requirement. Pistachio is deciduous and has imparipinnate leaves, most often 2-paired. It is dioecious, and both staminate and pistillate inflorescences are panicles that may have 150 or more individual flowers. They lack petals and nectaries and, consequently, are wind-pollinated. One male tree is provided in commercial orchards for every 8–12 female trees. Since the pistachio cannot be propagated from cuttings, seedling rootstocks of *P. atlantica*, *P. terbinthus*, and *P. integerrima* are T-budded with the desired cultivar. These stocks are more resistant to nematodes and other soil-borne organisms than is *P. vera*.

The fruit, a semidry drupe, is borne on 1-year-old wood in clusters similar to grapes and matures in September. The hull (exocarp and mesocarp) at that time slips easily from the shell (endocarp) which has already dehisced, exposing the kernel (see **illus.**).

Harvesting in California is done mechanically with tree shakers and catching frames. Mechanical hull removal, washing, and dehydration of the nuts soon after harvesting ensures unblemished, ivorylike shells, eliminating the necessity of dyeing them red, as is commonly done with imported nuts. A yield of 50 lb (22.5 kg) of dried-in-shell nuts per mature tree

Hull of the pistachio fruit at maturity separates easily from the dehisced shell that exposes the kernel.

may be expected in a good cropping year. The tree, however, is noted for its alternate bearing habit, producing a good crop one year followed by little or no crop the next.

Pistachio kernels contain only 5–10% sugars, but their protein and oil content of about 20 and 40%, respectively, make them high in food value. *See* NUT CROP CULTURE.

Julian C. Crane

Bibliography. J. C. Crane and B. T. Iwakiri, Morphology and reproduction of pistachio, *Hort. Rev.*, 3:376–393, 1981; R. A. Jaynes (ed.), *Nut Tree Culture in North America*, 1979; J. Maranto and J. C. Crane, *Pistachio Production*, Div. Agr. Sci. Univ. Calif. Leaflet 2279, 1982; W. E. Whitehouse, The pistachio nut: A new crop for the Western United States, *Econ. Bot.*, 11(4):281–321, 1957; J. G. Woodroof, *Tree Nuts*, 2d ed., 1979.

Pitch

The psychological property of sound characterized by highness or lowness. Pitch is one of the two major auditory attributes of simple sounds, the other being loudness.

Simple and complex sounds. A simple sound source, such as a tuning fork, produces an acoustic wave that approximates a perfect sinusoid, and the pitch of a sinusoid wave is almost completely determined by its frequency. Many sounds, however, are complex and contain a number of sinusoidal components. Complex sounds often appear to have a strong pitch, which is the frequency of a sinusoid that appears to match the complex sound. Hence, a tuning fork that vibrates at about 440 Hz will have a pitch very nearly equal to the note A above middle C on the piano. Loudness is determined by the amplitude of the sound vibrations. Loudness and pitch are independent attributes of sound quality. *See* LOUDNESS; TUNING FORK; WAVE (PHYSICS).

Sound sequences. A sequence of different sounds having definite pitches produces a musical tune, making pitch extremely important in music. Practically all musical conventions recognize that doubling the frequency of vibration produces a particular pitch interval, known as an octave. A conventional piano, for example, spans about seven octaves. Western music is based on the 12-tone scale. The frequencies representing the successive tones are separated by roughly equal steps on a logarithmic scale of frequency. The interval between successive notes has a frequency ratio of about $2^{1/12}$ and is called a semitone. Certain musical intervals, since they approximate simple fractions, are given special names. For example, there are seven intervals between C and G; the ratio of the frequencies of vibration between those two notes is nearly 3:2 and is called a perfect fifth. *See* MUSICAL ACOUSTICS; SCALE (MUSIC).

Human hearing. The human auditory system can hear frequencies in the range of 20–20,000 Hz. A hearing test consists of determining individual sensitivity to the frequencies 250, 500, 1000, 2000, 4000, and 8000 Hz. For frequencies between 100 and 4000 Hz, sinusoidal sounds have a clear pitch. Beyond these limits, the pitch of sound is not distinct. Sounds below 100 Hz may be described as rumbles, while those above 4000 Hz may be described as shrill and squeaky. Tunes can be recognized, however, when played in the octave between 6000 and 12,000 Hz. The ability to detect changes in pitch is remarkably acute. The just-detectable change in frequency is about 0.3% for the midfrequency range. Frequency changes are best detected when the sound is loud. Weaker sounds require greater changes in frequency to be detectable. *See* AUDIOMETRY; HEARING (HUMAN).

Mathematical description. Most musical tunes are not produced by simple acoustic sources but by string or reed instruments. Such instruments produce complex sounds—acoustic waveforms consisting of a number of sinusoid vibrations. Many musical sounds are periodic: the pressure waveform repeats itself after some fixed interval of time (period of the wave). According to Fourier's theorem, a periodic signal, such as a periodic sound, is a set of sinusoidal components with frequencies that can be expressed as successive integers. For a sound wave that repeats itself every 1/200 s, the fundamental of the wave is 200 Hz, the reciprocal of the wave's period. Its spectrum may have energy not only at the fundamental frequency but also at the harmonics—successive integral multiples of the fundamental. The pitch of such a periodic sound is generally considered to be equal to the fundamental, or 200 Hz. *See* FOURIER SERIES AND TRANSFORMS.

The missing fundamental. Even though energy at the fundamental as well as many of the lower frequencies might be absent, the pitch of the complex sound is still 200 Hz. This phenomenon is known as the missing fundamental. The fundamental of the human voice is about 100 Hz for males and 200 Hz for females. Since the conventional telephone passes energy only above about 300 Hz, the pitch of a human voice heard on the telephone is an example of the missing fundamental. During the middle of the twentieth century, J. Schouten conducted a number of important experiments on the pitch of the missing fundamental and other such periodic waveforms. His research has stimulated a major reevaluation of how the auditory mechanism is presumed to mediate the perception of pitch. *See* PSYCHOACOUSTICS; SPEECH.

David M. Green

Bibliography. D. M. Green, *Introduction to Hearing*, 1976; W. A. Yost and D. W. Nielsen, *Fundamentals of Hearing: An Introduction*, 4th ed., 2000; W. A. Yost and C. S. Watson (eds.), *Auditory Processing of Complex Sounds*, 1987.

Pitcher plant

Any member of the families Sarraceniaceae and Nepenthaceae. In these insectivorous plants the leaves form deep cups or pitchers in which water collects (**Fig. 1**). Visiting insects, falling into this water, are drowned and digested by the action of enzymes secreted by cells located in the walls of the pitcher-like structure of these plants. The Sarraceniaceae are divided into three genera; *Sarracenia* in eastern North America, *Darlingtonia* in northern California and southern Oregon, and *Heliamphora* endemic on high mountains in the northern part of South America. The Nepenthes family has only one genus, *Nepenthes*, which occurs in the Old World tropics from China to Australia, chiefly in Borneo. Often these plants climb by tendrils (prolongations of the midrib of the leaf). The end of a tendril may develop into a pitcher, which captures and digests insects (**Fig. 2**). *See* INSECTIVOROUS PLANTS; NEPENTHALES.

Perry D. Strausbaugh; Earl L. Core

Fig. 1. ***Sarracenia purpurea*, a pitcher plant.**

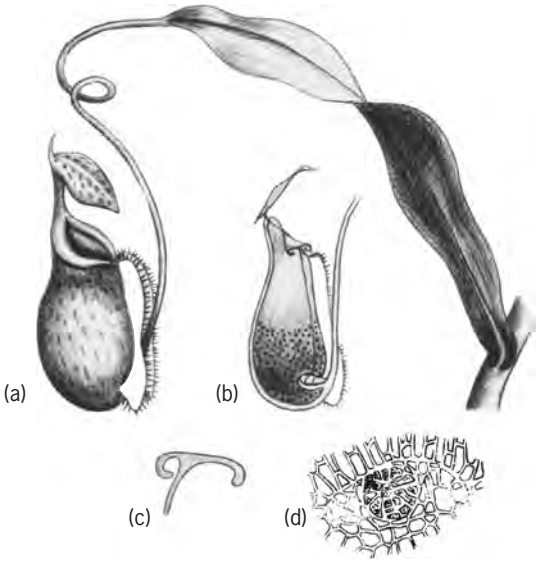

Fig. 2. ***Nepenthes*. (*a*) The complete leaf with its pitcher. (*b*) A vertical section through a pitcher. (*c*) A section through the margin of a pitcher. (*d*) Single gland from the lower part of a pitcher**

Pitchstone

A natural glass with dull or pitchy luster and generally brown, green, or gray color. It is extremely rich in microscopic, embryonic crystal growths (crystallites) which may cause its dull appearance. The water content of pitchstone is high and generally ranges from 4 to 10% by weight. Only a small proportion of this is primary; most is believed to have been absorbed from the surrounding regions after the glass developed. Pitchstone is formed by rapid cooling of molten rock material (lava or magma) and occurs most commonly as small dikes or as marginal portions of larger dikes. *See* IGNEOUS ROCKS; VOLCANIC GLASS.

Carleton A. Chapman

Pith

The central zone of tissue of an axis in which the vascular tissue is arranged as a hollow cylinder. Pith is present in most stems and in some roots (**Fig. 1**). Stems without pith rarely occur in angiosperms but are characteristic of psilopsids, lycopsids, *Sphenophyllum*, and some ferns. Roots of some ferns, many monocotyledons, and some dicotyledons include a pith although most roots have xylem tissue in the center. The pith may be present or absent in the same axis, depending upon size or vigor, the larger segments commonly containing pith.

Pith is composed usually of parenchyma cells often arranged in longitudinal files. This arrangement results from predominantly transverse division of pith mother cells near the apical meristem. The peripheral region consists of small cells with thick walls and remains alive longer than the central region. When the peripheral region is fairly well defined, it is called the medullary sheath or the perimedullary zone. The walls of pith cells may thicken in age and may become hard or remain soft. In some axes the pith may be composed principally of sclereids. In many fern stems, the inner pith is sclerenchymatous, whereas the outer is parenchymatous. In stems of some dicotyledons, plates or nests of sclerenchyma may be interspersed with the parenchyma. Such a pith is called diaphragmed pith (**Fig. 2**). If the parenchyma collapses or is torn during development, the sclerenchyma plates (diaphragms) alternate with hollow zones. Such a pith is said to be chambered. In many stems the entire pith becomes hollow except at the

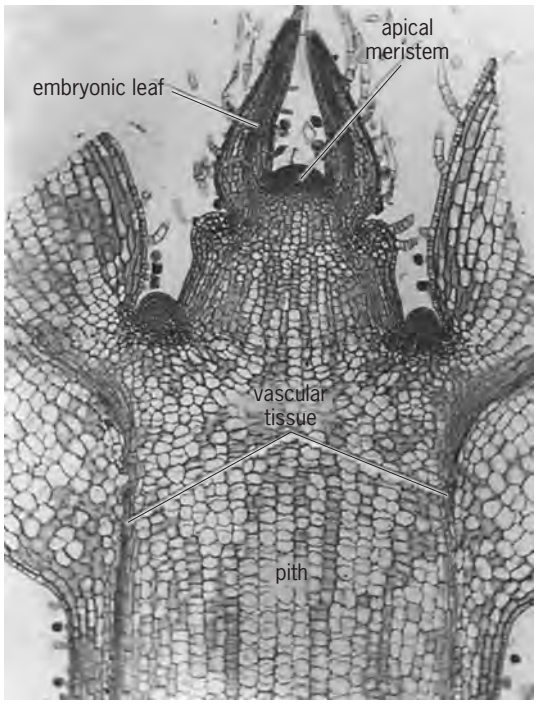

Fig. 1. Shoot apex of *Coleus blumei*, longitudinal section. (*After R. A. Popham, Laboratory Manual for Plant Anatomy, C. V. Mosby, 1966*)

nodes; the nodal diaphragms are sclerenchyma or parenchyma and, in monocotyledons, may contain vascular bundles.

Shape. The shape of the pith in stems of lower vascular plants and in roots of various plants is nearly cylindrical. In stems of higher vascular plants the pith is more or less angled or stellate in cross section. The shape is often characteristic of the plant groups, since it depends on phyllotaxy. In oaks, for example, the pith is five-angled and in alder, three-angled. In stems with cylinders of vascular bundles, the panels of ground tissue between bundles often are called medullary rays or pith rays. In stems in which the vascular bundles occur in a more complex arrangement than a simple cylinder, the limit of the pith is indefinite and when a major cylinder of vascular bundles can be distinguished, the internal bundles are called medullary bundles.

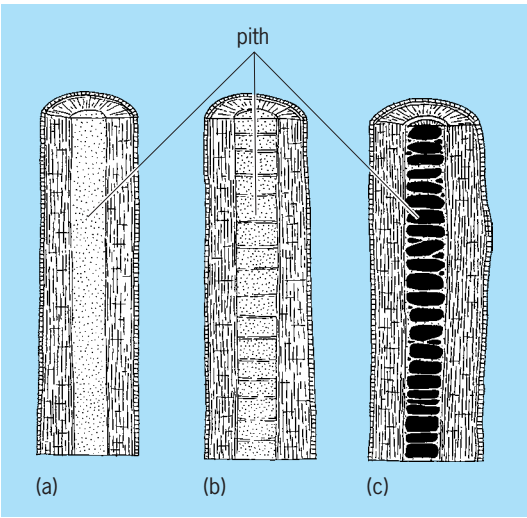

Fig. 2. Various types of pith in branchlets. (*a*) Continuous. (*b*) Diaphragmed. (*c*) Chambered. (*After E. L. Core, Plant Taxonomy, Prentice-Hall, 1955*)

Content. Ergastic materials often are stored in some or all cells of the pith. Secretory cells, or secretory canals, or laticifers may be present. In most stems with considerable secondary growth, the pith dies with the formation of heartwood, although the perimedullary zone may remain alive. In other stems the pith may consist partly or largely of dead cells by the end of the first year. *See* LEAF; PARENCHYMA; PERICYCLE; POLYPODIALES; ROOT (BOTANY); SCLERENCHYMA; SECRETORY STRUCTURES (PLANT); STEM.

H. Weston Blaser

Pitot tube

A device to measure the stagnation pressure due to isentropic deceleration of a flowing fluid. In its original form it was a glass tube bent at 90° and inserted in a stream flow, with its opening pointed upstream (**illus.** *a*). Water rises in the tube a distance, h, above the surface, and if friction losses are negligible, the velocity of the stream, V, is approximately $\sqrt{2gh}$, where g is the acceleration of gravity. However, there is a significant measurement error if the probe is misaligned at an angle α with respect to the stream (illus. *a*). For an open tube, the error is about 5% at $\alpha \approx 10°$.

Shielded probes. The misalignment error of a pitot tube is greatly reduced if the probe is shielded, as in the Kiel-type probe (illus. *b*). The Kiel probe is accurate up to $\alpha \approx 45°$.

Pitot-static probes. In most laboratory flows, there is no free surface, or the fluid is a gas; hence the visual water-rise effect of Pitot's invention is not possible. The modern application is a pitot-static probe, which measures both the stagnation pressure, with a hole in the front, and the static pressure in the moving stream, with holes on the sides. A pressure transducer or manometer records the difference between these two pressures. Pitot-static tubes are generally unshielded and must be carefully aligned with the flow to carry out accurate measurements. *See* BERNOULLI'S THEOREM.

Compressibility effects. When used with gases, estimate of the stream velocity is only valid for a low-speed or nearly incompressible flow, where the stream velocity is less than about 30% of the speed of sound of the fluid. For sea-level air, the stream velocity must be less than 100 m/s (225 mi/h). At higher velocities, estimate of the stream velocity must be replaced with a Bernoulli-type theory, which accounts for gas density and temperature changes. If the gas stream flow is supersonic, or the stream velocity is greater than the speed of sound of the gas, a shock wave forms in front of the probe and the measured static pressure and stagnation pressure are much different from the stream values upstream of the

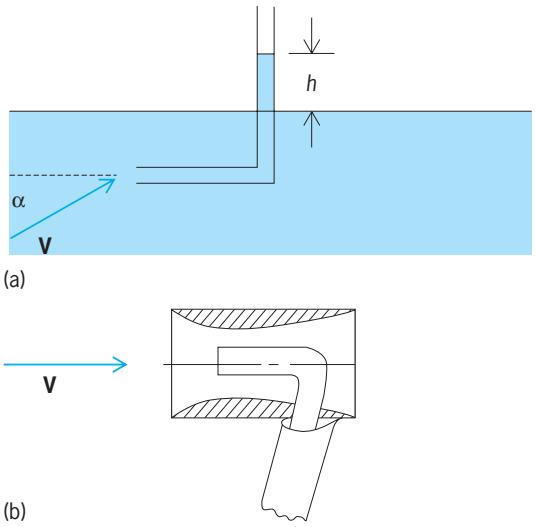

Pitot tubes. (*a*) Original form. Symbols explained in text. (*b*) Modern shielded Kiel probe.

shock. The theory must then be further corrected by complicated supersonic-flow algebraic relations. *See* COMPRESSIBLE FLOW; GAS DYNAMICS; SHOCK WAVE.

A disadvantage of pitot and pitot-static tubes is that they have substantial dynamic resistance to changing conditions and thus cannot accurately measure unsteady, accelerating, or fluctuating flows. Other fast-response devices are used for highly unsteady flows. *See* ANEMOMETER; FLOW MEASUREMENT.

Frank M. White

Bibliography. T. G. Beckwith, R. D. Marangoni, and J. H. Lienhard, *Mechanical Measurements*, 5th ed., 1993; J. W. Dally, W. F. Riley, and K. G. McConnell, *Instrumentation for Engineering Measurements*, 2d ed., 1993; R. J. Goldstein, *Fluid Mechanics Measurements*, 2d ed., 1996; R. W. Miller, *Flow Measurement Engineering Handbook*, 3d ed., 1996; S. Tavoularis, *Measurement in Fluid Mechanics*, 2005.

Pituitary gland

The most structurally and functionally complex organ of the endocrine system. Through its hormones, the pituitary, also known as the hypophysis, affects every physiological process of the body. All vertebrates, beginning with the cyclostomes, have a pituitary gland with a common basic structure and function. In addition to its endocrine functions, the pituitary may play a role in the immune response.

Developmental embryology. The hypophysis of all vertebrates has two major segments—the neurohypophysis (a neural component) and the adenohypophysis (an epithelial component)—each with a different embryological origin. The neurohypophysis develops from a downward process of the diencephalon (the base of the brain), whereas the adenohypophysis originates as an outpocketing of the primitive buccal epithelium, known as Rathke's pouch. An epithelial stalk connecting the embryonic adenohypophysis with the buccal epithelium commonly disappears during development, leaving the adenohypophysis attached to the neurohypophysis; however, the stalk may persist to form an open ciliated duct in some teleosts or a solid cord in some birds.

The adenohypophysis has three distinct subdivisions: the pars tuberalis, the pars distalis, and the pars intermedia. Development of the pituitary gland is a dynamic phenomenon that involves the formation of Rathke's pouch and its migration to a point of contact with the neural process that will become the neurohypophysis. The pars intermedia forms in the adenohypophysis at its zone of contact with the neurohypophysis. The failure of the pars intermedia to develop in birds, whales, dolphins, elephants, and armadillos can be explained by the failure of the adenohypophysis and neurohypophysis to make contact during embryogenesis. There is no such apparent embryological explanation for the regression of the human pars intermedia after it has formed.

The neurohypophysis comprises the pars nervosa and the infundibulum. The latter consists of the infundibular stalk and the median eminence of the tuber cinereum.

Physiology. The structural intimacy of neurohypophysis and adenohypophysis that is established early during embryogenesis reflects the direct functional interaction between the central nervous system and endocrine system. The extent of this anatomical intimacy varies considerably among the vertebrate classes, from limited contact in cyclostomes to intimate interdigitation in teleosts (**Fig. 1**). Vascular or neuronal pathways, or both, provide the means of exchanging chemical signals, thus enabling centers in the brain to exert control over the synthesis and release of adenohypophysial hormones.

Neurohormones, which are synthesized in specific regions of the brain, are conveyed to the neurohypophysis by way of axonal tracts, where they may be stored in distended axonal endings. Axons may also contact blood vessels and discharge their neurosecretory products into the systemic circulation or into a portal system leading to the adenohypophysis, or they may directly innervate pituitary gland cells.

In most animals, the vascular link is the prime route of information transfer between brain and pituitary gland. A description of this link begins in the tuber cinereum, the portion of the third ventricle floor that extends toward the infundibulum. The lower tuber cinereum, which is known as the median eminence, is well endowed with blood vessels that drain down into the pituitary stalk and ultimately empty into a secondary plexus in the anterior pituitary. The vascular link between the median eminence and the pituitary gland is known as the hypothalamo-hypophysial portal system. The median eminence in humans is vascularized by the paired superior hypophysial arteries. The pituitary gland is believed to have the highest blood flow rate of any organ in the body. However, its blood is received indirectly via the median eminence and the

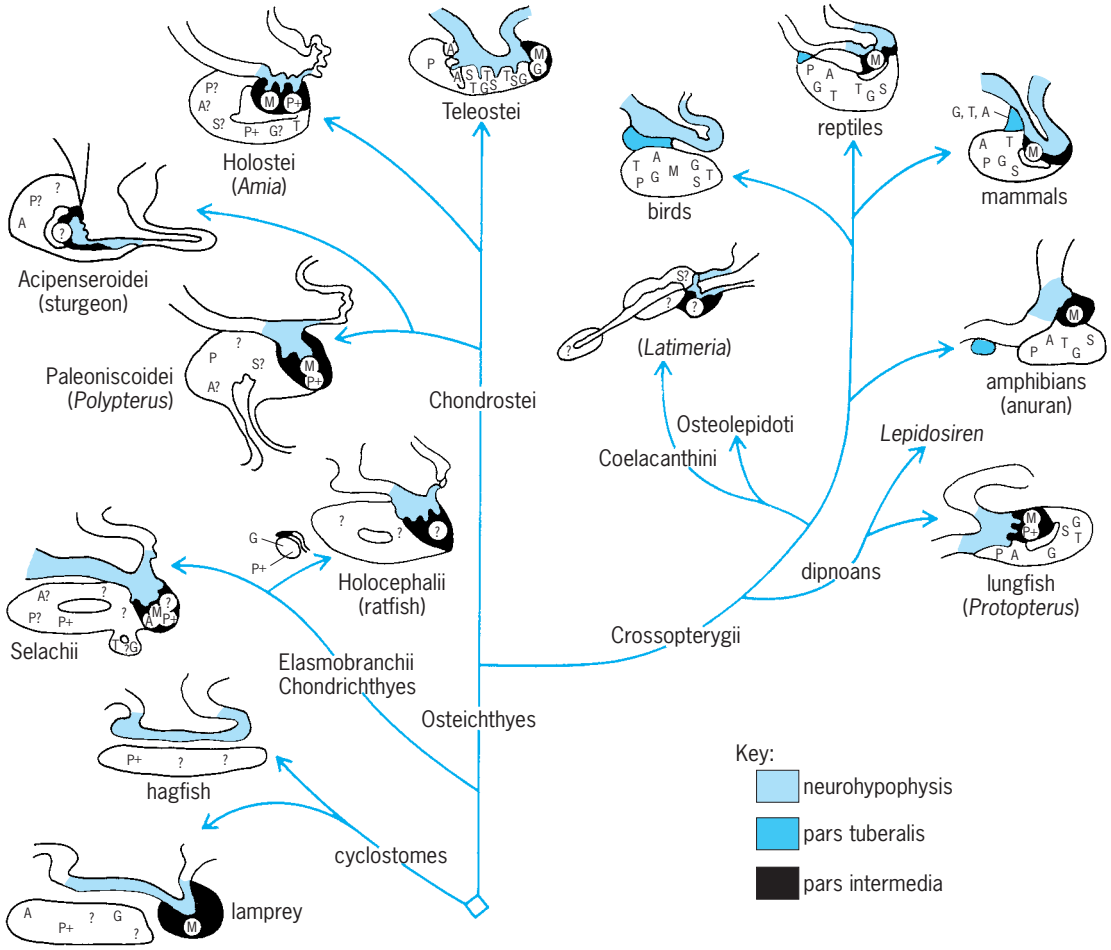

Fig. 1. Phylogenetic tree of generalized vertebrate pituitary glands (midsagittal section). P = prolactin; S = somatotropin; G = gonadotropin; T = thyrotropin; A = adrenocorticotropin; M = melanocyte-stimulating hormone; ? = unidentified function. (*After M. P. Schreibman, Pituitary gland, in P. K. T. Pang and M. P. Schreibman, eds., Vertebrate Endocrinology: Fundamentals and Biomedical Implications, vol. 1, Academic Press, 1986*)

hypothalamo-hypophysial portal system. Most of the blood flow is from the brain to the pituitary gland, with retrograde flow from the adenohypophysis to the hypothalamus, suggesting a two-way communication between nervous and endocrine systems. Although the brain is protected from the chemical substances in the circulatory system by the blood–brain barrier, the median eminence lies outside that protective mechanism and is therefore permeable to intravascular substances.

Histology and cytology. The adenohypophysis is a conglomeration of cell types, specific in appearance and function, that range in organization from masses that are highly segregated by cell type (in teleosts, the bony fishes) to more randomized mixtures of cells (in birds and mammals). Light microscopy with standard cytological stains and electron microscopy have revealed five or six different cell types within the pituitary. This correlates well with the number of hormones thought to originate from the pars distalis. Cytological analysis reveals cells to be identifiable on the basis of physiological function. Immunocytochemistry has facilitated an even more specific identification of pituitary cells and evaluation of their activity.

Comparative morphology. The pituitary gland varies somewhat in complexity and hormonal potential among the vertebrate classes.

Agnatha. Agnathans, the most primitive living vertebrates, are represented by the lamprey and the hagfish. The hagfish pituitary gland appears to be the more primitive of the two. Its neurohypophysis is highly developed, flattened, and saclike, whereas its hypothalamo-hypophysial portal system is absent or markedly reduced but may show species variation.

The adenohypophysis is separated from the neurohypophysis and is not differentiated into zones. Generally, islets of cells are embedded in a poorly vascularized, loose connective tissue that is continuous with the layer separating the adenohypophysis from the neurohypophysis. Several cell types appear to be present in the hagfish. A thyroid gland–stimulating substance may be present, but attempts to demonstrate a factor that stimulates the gonads in the hagfish have not been successful.

Hypothalamic gonadotropin-releasing hormone is delivered to target cells in the adenohypophysis of lampreys by means of simple diffusion between the neurohypophysis and adenohypophysis or by secretion of the hormone into the third ventricle of the

brain, where it can reach the pituitary gland. Levels of gonadotropin-releasing hormone are known to fluctuate with reproductive season, suggesting a typical regulatory role on gonad function.

The pars distalis in lampreys is separated from the infundibulum by connective tissue. The cell types display a regional distribution. Adrenocorticotropic hormone, luteinizing hormone, melanocyte-stimulating hormone, and perhaps other hormones are present in the lamprey pituitary gland. *See* JAWLESS VERTEBRATES.

Chondrichthyes. The two extant subclasses of chondrichthyes are the Elasmobranchii (sharks, skates, and rays) and the Holocephali (rat fishes or rabbit fishes and chimaeroids).

In elasmobranch fishes, the neurohypophysis consists of a thin-walled anterior portion—a true median eminence—connected to the pars distalis by a hypothalamo-hypophysial portal system. The adenohypophysis is divided into four regions: an elongated pars distalis (dorsal lobe), with rostral and caudal regions; a ventral lobe attached by a stalk to the caudal pars distalis; and a large pars intermedia that is heavily penetrated by neurohypophysial tissue to form a typical neurointermediate lobe. A pars tuberalis is not present. Follicles and spaces are characteristic of the chondrichthyian pituitary. They range from a highly developed system of vesicles and tubules that communicates with a hollow pars distalis (sharks and dogfish) to a hypophysial cavity that is small or lost entirely (skates and rays). The ventral lobe of the pars distalis is also hollow and often vesicular. Adrenocorticotropic hormone and prolactin are produced, presumably, in the rostral pars distalis; the former may also be found in the pars intermedia. The large cells of the ventral lobe produce thyrotropin (thyrotropes) and gonadotropin (gonadotropes).

The holocephalian neurohypophysis has a pars nervosa and a prominent median eminence that sends many short blood vessels into the rostral and caudal pars distalis, but the entire pars distalis is hollow and is not as clearly divisible into rostral and caudal regions as that of the elasmobranchs (Fig. 1). A ventral lobe is not present. Adult holocephalians have a large compact follicular structure called the Rachendach-hypophyse (pharyngeal lobe). Gonadotropic activity has been identified there in the rabbit fish, making this buccal lobe comparable to the elasmobranch ventral lobe. *See* CHONDRICHTHYES.

Osteichthyes. The subclass Actinopterygii (fish having a bony skeleton and paired ray fins) includes the infraclasses Teleostei, Chondrostei, and Holostei. Only the bony fishes (teleosts) will be considered here.

In those teleosts examined, the adenohypophysis is clearly divided into three regions, the rostral and caudal pars distalis and pars intermedia (Fig. 1). There is no pars tuberalis. The clarity of this separation is essentially due to a restriction of specific cell types to distinct regions of the gland.

In teleosts, there is a structural intimacy of the neurohypophysis with all three regions of the adenohypophysis, and there is generally no true median eminence or portal system. A unique feature is that the adenohypophysis cells are innervated directly, both neurohypophysial and adenohypophysial cells being in contact with neuronal endings from hypothalamic and extrahypothalamic nuclei. Axons may originate from perikarya as close as the hypothalamic nuclei or from distant extrahypothalamic nuclei.

In general, hormone-producing cells of a particular type are restricted to specific regions of the adenohypophysis. The rostral pars distalis is essentially a single mass of prolactin-producing cells. A narrow band of adrenocorticotropic hormone–producing cells (corticotropes) forms the posterior boundary of the rostral pars distalis. In the caudal pars distalis, cells producing thyroid gland–stimulating hormone (thyrotropes) and somatotropin- or growth hormone–producing cells (somatotropes) are intermingled in islets formed by pervading neurohypophysial tissue. The external boundary of the caudal pars distalis consists of several cell layers of gonadotropes, presumably producing at least two chemically different gonadotropins.

The pars intermedia generally contains two cell types. One is presumed to be the source of melanocyte-stimulating hormone, whereas the other contains glycoprotein and in some species is thought to be a source of a second type of gonadotropin.

The subclass Sarcopterygii consists of the infraclasses Dipnoi and Crossopterygii. The single coelacanth genus, *Latimeria*, and the six species of lungfish have been used to study the evolution of the tetrapod pituitary gland. The neurohypophysis of lungfish (Dipnoi) consists of a median eminence, an infundibular stalk, and a neural lobe. The adenohypophysis is more similar in structure to that of amphibians than to other fishes. Amphibian characteristics of the lungfish pituitary are a well-formed neural lobe, a prominent median eminence, the absence of a saccus vasculosus, and a less obvious regional distribution of cell types. Fishlike characteristics include direct aminergic innervation of the pars distalis cells, interdigitation of the neurohypophysis and pars intermedia, prominent hypophysial cleft between the pars distalis and the pars intermedia, absence of a pars tuberalis, and presence of follicles. The five pars distalis cell types are intermingled and lack the distinct zonation seen in teleosts.

In the infraclass Crossopterygii, *Latimeria chaulumnae* is the only living representative of the Coelacanthiformes. The orthodox features of its pituitary include a typically fishlike neurointermediate lobe ventral to the brain, direct contact between pars distalis and pars intermedia, follicles and tubules containing glycoprotein colloid as in many fishes, a well-developed portal system, and probable median eminence. An exceptional feature is a cylindrical, tubular hypophysial cavity up to 5 in. (12 cm) long that extends from the rostral lobe of a greatly extended, tripartite pars

distalis and that contains vascularized masses of adenohypophysial cells (the rostral islets or pars buccalis). These structures are comparable with the Rachendachhypophyse of holocephalians. *See* OSTEICHTHYES.

Amphibians. The amphibian pituitary gland is very much like the hypophysis of the other tetrapods. The adenohypophysis is generally composed of a flat, wide, compact pars distalis. Dorsally it continues with the pars intermedia without an intervening cleft, and anteriorly it is connected to the median eminence by a strand of connective tissue through which the portal vessels pass. The pars tuberalis is seen in amphibians for the first time in vertebrate phylogeny. Specific cell types are scattered, but the pars intermedia forms a distinct and separable band of tissue between the pars distalis and the pars nervosa. One dominant cell type in the pars intermedia predominates and is probably the source of melanocyte-stimulating hormone. *See* PIGMENTATION.

Reptilia. The morphology of the reptilian hypophysis strongly resembles both the amphibian and the avian types. It differs from the amphibian gland in that the pars distalis and the pars intermedia are structurally more intimate, the pars tuberalis may not be present in some groups, and the pars intermedia may be larger than in any other vertebrate. A median eminence is found on the infundibular stalk just anterior to the pars distalis. Portal vessels connect primary and secondary capillary plexuses.

As is typical in tetrapods, the reptilian pars distalis is made up of branching cords of cells that are highly vascularized by intervening sinusoids. Gonadotropes have a scattered distribution. In some snakes and lizards, they are ventrally located in the midline; in other lizards, they are restricted to the rostral lobe. The distribution of hypophysial cell types is similar in reptiles and amphibians. In some reptiles, such as lizards, an epithelial stalk (a remnant of Rathke's pouch) may persist. The pars tuberalis is the most variable component of the gland. Gonadotropes and thyrotropes are present in the pars tuberalis. *See* REPTILIA.

Aves. An epithelial stalk persists in some birds as a connection between the pars distalis and the buccal epithelium. The separation of the adenohypophysis from the neurohypophysis in birds makes the gland less compact than, for example, that of mammals. A well-developed pars tuberalis is located between the median eminence and the pars distalis, although in some species it may lie within the median eminence. The avian median eminence can be regionally differentiated into the "point-to-point" system: a rostral portion supplies blood to the rostral pars distalis, and a caudal area delivers blood to the caudal pars distalis. *See* AVES.

Mammalia. The mammalian pituitary gland shows only minor variations between the various genera (**Fig. 2**). In some mammals, including humans, a pharyngeal (craniopharyngeal) hypophysis develops from a remnant of Rathke's pouch. Found in the roof of the mouth, it is vascularized and has staining and immunoreactive properties similar to those of the sellar hypophysis.

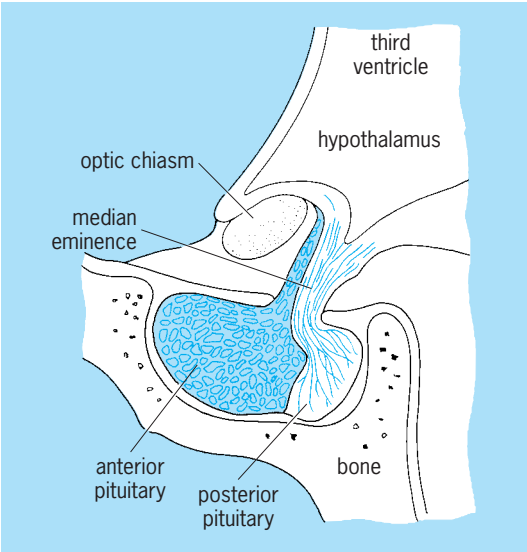

Fig. 2. Relationship of the pituitary gland to the brain and hypothalamus. (*After D. S. Luciano, A. J. Vander, and J. H. Sherman, Human Anatomy and Physiology: Structure and Function, 2d ed., McGraw-Hill, 1983*)

Another structural variation is the failure of the pars intermedia to form in cetaceans (whales and dolphins), elephants, and armadillos, presumably as a result of a connective tissue barrier that separates the neurohypophysis and the adenohypophysis during development. Although a pars tuberalis is characteristic of the mammalian gland, it may be absent in some animals, such as the sloth and pangolin.

The mammalian adenohypophysis is regulated primarily by its vascular supply. The adult mammalian neurohypophysis is a neurohemal organ that receives aminergic and peptidergic neurons. The median eminence is prominent and forms a primary capillary plexus around much of the infundibular stalk in most species. Most often, portal vessels pass through a well-developed pars tuberalis on their way to the pars distalis. *See* MAMMALIA.

Adenohypophysis hormones. The hormones of the adenohypophysis may be grouped into three categories based on chemical and functional similarities. The first category consists of growth hormone (also known as somatotropin) and prolactin, both of which are large, single, polypeptide chains; the second category consists of the glycoprotein hormones; this family of hormones contains the gonadotropins and thyrotropin. The gonadotropins in many species, including humans, can be segregated into two distinct hormones, follicle-stimulating hormone and luteinizing hormone. In other species only one gonadotropin molecule has been isolated, and in still other species two gonadotropins have been isolated. The third group comprises adrenocortiotropic hormone and melanotropin (MSH; melanocyte-stimulating hormone).

The regulation of the release of pituitary hormones is determined by precise monitoring of circulating hormone levels in the blood and by genetic and environmental factors that manifest their effect through the releasing and release-inhibiting factors of the hypothalamus. *See* ADENOHYPOPHYSIS HORMONE.

Hypothalamic and neuronal involvement. The hypothalamus is located at the base of the brain (the diencephalon) below the thalamus and above the pituitary gland, forming the walls and the lower portion of the third ventricle. It receives major neuronal inputs from the sense organs, hippocampus, thalamus, and lower brainstem structures, including the reticular formation and the spinal cord. Thus, the hypothalamus is designed and anatomically positioned to receive a diversity of messages from external and internal sources that can be transmitted by way of hypothalamic releasing factors to the pituitary gland, where they are translated into endocrine action. *See* NERVOUS SYSTEM (VERTEBRATE).

Neurohypophysis hormones. The neurohypophysis hormones, oxytocin and vasopressin, which are nonapeptides, are synthesized in different neurons of the paraventricular and supraoptic nuclei of the hypothalamus and travel by axonal flow to the terminals in the neurohypophysis for storage and ultimate release into the vascular system. Oxytocin is important in stimulating milk release through its contractile action on muscle elements in the mammary gland. It also stimulates uterine smooth muscle contraction at parturition. Vasopressin affects water retention by its action on certain kidney tubules. Thus, it also affects blood pressure (Fig. 2*a*). *See* LACTATION.

Birds, reptiles, amphibians, and lungfishes have arginine vasotocin and arginine mesotocin as the principal peptides of the posterior pituitary. In birds and reptiles, the nonapeptides primarily originate from the supraoptic and paraventricular nuclei; in poikilothermic (cold-blooded or exothermic) vertebrates, however, the preoptic nucleus is the homolog of the two nuclei found in homeotherms (endotherms). The principal nonapeptide found in various groups of fishes is arginine vasotocin, but a second neural peptide related to oxytocin and sometimes oxytocin itself may be found. Cyclostomes possess only arginine vasotocin.

In several species of birds, both arginine vasotocin and mesotocin are found in separate neurons in the supraoptic and paraventricular nuclei; however, only arginine vasotocin has been localized in the median eminence. *See* NEUROHYPOPHYSIS HORMONE.

Neurohormones. The five neurohormones in vertebrates are growth hormone–releasing hormone, somatostatin, corticotropin-releasing hormone, thyrotropin-releasing hormone, and gonadotropin-releasing hormone. In addition, other peptides found in the nervous system may be released into the median eminence or posterior pituitary and are probably capable of neurohormonal action.

Growth hormone–releasing hormone (GRH) consists of either 44 (human, ovine, porcine) or 43 (rat) amino acids. In some mammals, including humans, perikarya containing immunoreactive growth hormone–releasing hormone are located in the arcuate (infundibular) and ventromedial nuclei of the hypothalamus.

In mammals, somatostatin, which is also known as somatotropin release–inhibiting hormone (SRIH), contains 14 amino acids or 28 amino acids; both have been found to be identical in all mammals investigated. When released from the median eminence, somatostatin has a neurohormonal function, whereas in the arcuate nucleus it is most likely serving as a neurotransmitter. *See* SOMATOSTATIN.

In sheep, corticotropin-releasing hormone (CRH) is composed of 41 amino acids and affects the release of corticotropin, beta-endorphin, and melanotropin. In each major mammalian group examined, immunoreactive CRH has been localized in many brain areas, the spinal cord, and the gastrointestinal tract, with the greatest concentration in the hypothalamus.

Thyrotropin-releasing hormone (TRH), a tripeptide, controls the release of thyroid gland–stimulating hormone from the pituitary gland. Like somatostatin, thyrotropin-releasing hormone is widely distributed in the brain and in the gastrointestinal tract. Its levels are highest in the median eminence, and the densest collection of cell bodies containing thyrotropin-releasing hormone is in the medial, parvocellular division of the paraventricular nucleus. Studies of its distribution, which have been conducted mainly in mammals, reveal that the distribution of thyrotropin-releasing hormone in frogs and in rats is identical.

Gonadotropin-releasing hormone (GnRH), formerly known as luteinizing hormone–releasing hormone, is a decapeptide (10 amino acids) originally isolated from pig and cattle hypothalami. It regulates the release of luteinizing hormone and follicle-stimulating hormone from the adenohypophysis. There are several structural variants of mammalian gonadotropin-releasing hormone in nonmammalian vertebrates, and four have been sequenced, two from chicken, one from salmon, and one from lamprey. Additionally, several of the variants may exist within a single species, suggesting that there may be different functions for these peptides. Gonadotropin-releasing hormone or GnRH-like peptides have been localized not only in the brain and pituitary gland but also in other areas of the central nervous system and in the gonads, placenta, mammary glands, and pancreas. During evolution, the peptide has been subject to gene duplication and structural change and thus may serve a variety of regulatory functions. In addition to the release of pituitary gonadotropic hormone, gonadotropin-releasing hormone may also function as a neurotransmitter in the central and sympathetic nervous systems, as a paracrine regulator in the gonads and the placenta, and as an autocrine regulator in tumor cells.

The distribution of perikarya containing immunopositive gonadotropin-releasing hormone in

the brain has been studied in various mammalian species, particularly the rat. Perikarya that contain immunoreactive gonadotropin-releasing hormone are scattered in a continuum from the olfactory bulb and terminal nerve to the area of the premammillary nucleus. Similarly, immunoreactive gonadotropin-releasing hormone is distributed throughout the brain in nonmammalian vertebrates, although the specifics of this distribution remain controversial. Changes in the distribution of gonadotropin-releasing hormone may be related to the sex, genotype, physiological state, and age of the animals being studied.

In addition to those historically recognized as neurohormones, many of the peptides found in the nervous system are probably released into the median eminence or posterior pituitary and have neurohormonal action.

Neurotransmitters. The better-known neurotransmitters of the central nervous system include the catecholamines (dopamine, epinephrine, and norepinephrine), serotonin, acetylcholine, gamma-amino butyric acid (GABA), histamine, and the opioid peptides (enkephalins, endorphins, dynorphin, neoendorphin, rimorphin, and leumorphin). These substances are distributed widely in the central nervous system and, for most, also in the pituitary gland. As is seen with many of the peptides identified in the nervous system, if a particular amine or neurotransmitter is present in nerve fibers leading to the median eminence, it probably will influence pituitary gland activity via the portal system. Dopamine, serotonin, gamma-amino butyric acid, and acetylcholine are best known for such activity. These neurotransmitters play an important, but poorly understood, role in regulating pituitary function, either directly or by their action on neuropeptide-producing neurons. Understanding the pharmacology of neurotransmitters holds promise for the treatment of basic disorders of the hypothalamic-pituitary axis. *See* ACETYLCHOLINE; ENDOCRINE MECHANISMS; ENDOCRINE SYSTEM (VERTEBRATE); ENDORPHINS; HISTAMINE; HORMONE; NEUROBIOLOGY; NEUROIMMUNOLOGY; PITUITARY GLAND DISORDERS; SEROTONIN. Martin P. Schreibman

Bibliography. I. Chester-Jones and P. M. Ingleton (eds.), *Fundamentals of Comparative Vertebrate Endocrinology*, 1987; L. J. DeGroot (ed.), *Endocrinology*, 3 vols., 4th ed., 2001; H. Imura (ed.), *The Pituitary Gland*, 2d ed., 1994; E. E. Muller and N. Guiseppe, *Brain Messengers and the Pituitary*, 1988; P. K. T. Pang and M. P. Schreibman (eds.), *Vertebrate Endocrinology*, 4 vols., 1986–1991; W. F. Walker, *Functional Anatomy of the Vertebrates*, 2d ed., 1997.

Pituitary gland disorders

Inborn or acquired abnormalities in the structure or function of the human pituitary gland. Pituitary disorders can stem from any of five disease processes: tumors and other growths, intrinsic lesions of the anterior lobe, diseases affecting gland function, hypothalamic malfunction, and systemic disease that affects the adenohypophysis.

Tumors and other growths in and near the pituitary may cause failure of hormone secretion or impinge on nearby brain structures. The latter effect can give rise to neurological malfunctions, the most common of which is visual impairment such as narrowing of the visual fields. Intrinsic lesions of the anterior pituitary, usually benign, secrete excessive amounts of a single (rarely multiple) hormone, producing characteristic endocrine syndromes, the most dramatic of which are acromegaly and Cushing's disease. Infections, congenital anomalies, granulomas, vascular disorders, and, rarely, metastatic cancers may induce partial or total failure of one or more pituitary secretions, in which case growth hormone, follicle-stimulating hormone, and luteinizing hormone are the first to fail.

Diseases of the hypothalamus affect pituitary function through the mechanical effects of a mass or through disrupted secretion of the hypophysiotropic peptides (see **table**). Typical manifestations include precocious or delayed puberty; diabetes insipidus; and derangements of sleep, eating, and temperature regulation. The dysfunction may be of congenital, traumatic, inflammatory, or neoplastic origin. However, the most important hypothalamic-pituitary diseases are tertiary hypothyroidism, precocious or delayed puberty, diabetes insipidus, and Kallmann's syndrome, which consists of a deficient sense of smell and lack of sexual development due to inborn failure of the hypothalamus to secrete gonadotropin-releasing hormone. Hypothalamic disorders can cause abnormalities of growth hormone secretion. Whereas most individuals with acromegaly (an excess of growth hormone) have intrinsic pituitary tumors that secrete growth hormone, some have excessive secretion of hypothalamic growth hormone–releasing hormone or insufficient secretion of somatostatin. Pituitary dwarfism, caused by the failure to secrete growth hormone in childhood, is usually due to primary disease of the adenohypophysis, but in some individuals the fundamental dysfunction (opposite to that in acromegaly) lies in the hypothalamus. *See* DIABETES; DWARFISM AND GIGANTISM; SOMATOSTATIN; THYROID GLAND DISORDERS.

The adenohypophysis can be affected subtly by systemic diseases, or more obviously by metastases from breast cancer. Rarely, acute necrosis (death) of the gland occurs with diabetic crises, heart attack, generalized infection, or as a result of using mechanical respirators. Prolonged treatment with large doses of adrenal corticosteroids, as in cases of lupus, asthma, or acute leukemia, can result in the failure of the adenohypophysis to secrete adrenocorticotropic hormone in times of physical stress. A deficiency in adrenocorticotropic hormone is a potentially fatal condition. In addition, manic-depressive psychosis has been linked to an overactivity of the

Interrelationships between the hypothalamic hypophysiotropic hormones and the anterior pituitary hormones

Hypothalamic hypophysiotropic hormones			Anterior pituitary hormones			
Hormone	Structure	Function	Hormone	Structure	Cell of origin	Function
Growth hormone releasing hormone	Peptide	Stimulates growth hormone release	Growth hormone	Protein	Somatotrope	Stimulation of bodily growth
Somatostatin	Peptide	Inhibits growth hormone release				
Thyrotropin-releasing hormone	Small peptide	Stimulates thyrotropin release	Thyrotropin*	Glycoprotein	Thyrotrope	Regulation of thyroid gland
		Stimulates prolactin release	Prolactin	Protein	Mammotrope	Postpartum lactation
Dopamine	Amine	Inhibits prolactin release				
Gonadotropin-releasing hormone	Peptide	Stimulates release of follicle-stimulating hormone and luteinizing hormone	Follicle-stimulating hormone	Glycoprotein	Gonadotrope	Stimulation of ovarian follicle; spermatogenesis
			Luteinizing hormone	Glycoprotein	Gonadotrope	Stimulation of corpus luteum, and testosterone secretion by testis
Corticotropin-releasing hormone	Peptide	Stimulates release of adrenocorticotropin, β-endorphin, β-lipotropin, and melanotropin	Adrenocorticotropin	Peptide	Corticotrope	Regulation of adrenal cortex
			β-endorphin†	Small peptide	Corticotrope	Possible effect on stress reaction and memory
			β-lipotropin†	Peptide	Corticotrope	Breakdown of fat
			α, β-, and γ-melanotropins†	Peptide	Corticotrope	Pigmentation of skin

*Also known as thyroid-stimulating hormone.
† The role of these hormones in human physiology and disease is not known.

adrenocorticotropic hormone–adrenal cortex axis. *See* ADENOHYPOPHYSIS HORMONE; ADRENAL GLAND DISORDERS.

Anterior lobe malfunction. Malfunctions of the anterior lobe of the pituitary can cause hypopituitarism (pituitary insufficiency) or hyperpituitarism (pituitary excess).

Hypopituitarism. Adult hypopituitarism has many possible causes; more than 50% of cases result from benign or malignant tumors of the gland. Metastatic cancer originating in the breast, colon, or kidney rarely invades and destroys the pituitary. Tumors and cysts account for another 15–20%. Still another 10–15% can be traced to granulomas, as in sarcoid, tuberculosis, and syphilis; rare systemic disorders; autoimmune inflammation of the pituitary gland; defects in hypothalamic hypophysiotropic hormone secretion; and injury, including therapeutic x-radiation, surgery, or skull fracture. The most frequent cause of panhypopituitarism, that is, deficiency or absence of all pituitary hormones, is Sheehan's disease, which involves tissue death caused by insufficient blood supply to the gland due to copious blood loss or intravascular clotting and hemorrhage associated with premature separation of the placenta during childbirth.

The clinical picture of adult hypopituitarism depends on the cause. If a pituitary tumor or other mass is present, the first clinical symptoms are usually headache, visual impairment, and visual field defects. Deficiencies of thyrotropin, gonadotropins, and adrenocorticotropin are mild at first; hypothyroidism, hypogonadism, and adrenal insufficiency develop slowly. Still more gradual is the course of Sheehan's disease: full-blown diagnostic features may take 15–20 years to appear. Without treatment, death can follow severe hypoglycemia (low blood sugar) accompanying what would be a minor illness in a normal person. The causes of the hypoglycemia are gross deficiencies of pituitary adrenocorticotropin and growth hormone secretion.

Diagnosis is made by combining computerized tomographic scanning or magnetic resonance imaging with anterior pituitary hormone response tests. Treatment consists of hormone replacement therapy with estrogen or testosterone, thyroxine, and hydrocortisone. Tumors must be destroyed by radiation or surgery.

Chronic deficiencies of single pituitary hormones are usually due to failure of the corresponding specific hypothalamic hypophysiotropic hormone (see table). The most important deficiency for children is the failure to secrete growth hormone, with resulting dwarfism. The usual cause is intrinsic failure of the anterior pituitary somatotropes in the absence of a tumor. Children with hypopituitary dwarfism are small, normally proportioned, and slightly obese. Diagnosis can be established only by documenting subnormal growth over time and demonstrating inadequate secretion of growth hormone during deep sleep. Treatment by frequent injections of synthetic human growth hormone accelerates the rate of linear growth but does not result in full attainment of expected ultimate stature.

Hyperpituitarism. Hyperpituitarism is almost always due to excessive secretion of a single hormone by a tumor of the anterior pituitary. Somatotrope adenomas, which are benign tumors, secrete excessive growth hormone, producing acromegaly or gigantism. The acromegalic person is middle-aged and has enlarged bony prominences of the face, larger-than-normal lips, ears, hands, and feet, a deepened and broadened chest cage, and bowed legs. Diabetes mellitus is the chief metabolic abnormality. Treatment is

difficult and usually consists of x-radiation or microsurgical removal of the adenoma. After treatment, plasma growth hormone levels fall slowly.

The most common endocrinologically active tumor (30% of all secretory adenohypophyseal tumors) is the prolactin-secreting lactotrope adenoma. Milk secretion not associated with childbirth is often present in both sexes, and plasma prolactin levels are high. Treatment is controversial, as x-radiation alone is not very effective. The other two choices are removal of the tumor and daily administration over many years of bromocriptine, which attaches to the same cellular receptors as dopamine and thereby inhibits prolactin secretion. Corticotrope adenomas, secreting excessive adrenocorticotropin, occur in Cushing's disease or in those whose adrenal glands have been completely removed. Rare thyrotrope adenomas, which hypersecrete thyrotropin, are found in persons with inadequately treated hypothyroidism. Also, some people, for unknown reasons, lack the normal capacity of thyroid hormone to suppress pituitary secretion of thyrotropin (resistance syndrome). In addition, some have preexisting hyperthyroidism, and so treatment is extremely difficult. The hyperthyroidism can be fairly well controlled by standard endocrine methods, but often the thyrotrope adenoma can be brought under control only by giving a somatostatin analog. The uncommon gonadotrope adenomas secrete excessive follicle-stimulating hormone or, more rarely, luteinizing hormone. Curiously, the tumors produce no endocrine signs, and so they are undetected until they are quite large. Treatment is by microsurgery, x-radiation, or administration of synthetic chemical analogs of gonadotropin-releasing hormone, whose long-term effect is to inhibit the secretion of luteinizing and follicle-stimulating hormones.

These endocrinologically active tumors, together with the hormonally inert tumors of the adenohypophysis, constitute about 10% of all intracranial neoplasms. The hormonally inactive adenomas are noted clinically as neurological or visual problems; the endocrine consequences of pituitary destruction by a mass lesion occur much later. Tumor types are very numerous and can be either congenital or acquired. Diagnosis is made by neurological examination and radiography with computerized tomographic scanning or magnetic resonance imaging. Some individuals show an enlarged sella without pituitary tumor (empty sella syndrome), a developmental or acquired abnormality that can deform the gland. Treatment of pituitary tumors usually involves surgical removal, often combined with radiotherapy.

Posterior lobe malfunction. The two posterior lobe hormones that are secreted by the hypothalamus and stored in the pituitary are oxytocin and vasopressin. Oxytocin is secreted in large amounts at the onset of childbirth. It is not known whether hypersecretion of oxytocin triggers childbirth or is simply a response to it. Women with hypopituitarism and hypothalamic damage or destruction of the neurohypophysis can give birth normally. No known clinical disorders are attributable to abnormal secretion of oxytocin.

Absence or deficiency of vasopressin, the antidiuretic hormone, characterizes diabetes insipidus. Some cases are inherited, sometimes in association with other abnormalities, and are present at birth; some are of unknown cause; and some cases result from trauma, granuloma, tumor, meningitis, encephalitis, or disorders of the blood vessels. Because of the lack of vasopressin, the kidneys cannot excrete urine of normal concentration, and so as much as 3 gal (11 liters) of highly dilute urine are excreted each day. The water loss can lead to vascular collapse, shock, and death. Diagnosis is not simple; the principal disorder that must be ruled out is compulsive, excessive water drinking due to emotional disorder. Intrinsic resistance of the kidney to the action of vasopressin is exceedingly rare. Diagnostic accuracy requires carefully comparing values of plasma vasopressin with the electrolyte content of urine and of plasma under various standardized test conditions. Treatment with desmopressin, a synthetic, long-acting vasopressin analog in the form of a nasal spray, has proved satisfactory.

Vasopressin excess characterizes the syndrome of inappropriate antidiuretic hormone secretion. With this disorder, which has many causes, individuals are lethargic and have altered states of consciousness, seizures, very low serum concentrations of sodium, and variable abnormalities of blood volume. Hypopituitarism can produce the syndrome when deficient pituitary adrenocorticotropin secretion leads to subnormal production of hydrocortisone, a hormone that inhibits the release of vasopressin; hydrocortisone reduces plasma vasopressin levels and returns water balance to normal. *See* ENDOCRINE MECHANISMS; NEUROHYPOPHYSIS HORMONE; PITUITARY GLAND. Nicholas Christy

Bibliography. P. Belchetz, *Management of Pituitary Disease*, 1984; P. M. Black et al. (eds.), *Secretory Tumors of the Pituitary Gland*, 1984; L. J. DeGroot et al. (eds.), *Endrocrinology*, 4th ed., 3 vols., 2001; G. W. Harris, *Neural Control of the Pituitary Gland*, 1955; H. L. Sheehan and V. K. Summers, The syndrome of hypopituitarism, *Quart. J. Med.*, 42:319, 1949.

pK

The logarithm (to the base 10) of the reciprocal of the equilibrium constant for a specified reaction under specified conditions (for example, solvent and temperature). pK values are often more convenient to tabulate and use than the equilibrium constants themselves. The value of K for the dissociation of the HSO_4^- ion in aqueous solution at 77°F (25°C) is 0.0102 mole/liter. The logarithm is $0.008_6 - 2 = -1.991_4$. pK is therefore $+1.991_4$. The choice of algebraic sign, although arbitrary, results in positive values for most dissociation constants applicable to aqueous solutions. The concept of pK is especially valuable in the study of solutions. *See* CHEMICAL EQUILIBRIUM; IONIC EQUILIBRIUM; PH. Thomas F. Young

Placentation

The intimate association or fusion of a tissue or organ of the embryonic stage of an animal to its parent for physiological exchange to promote the growth and development of the young. Placentation enables the young, retained within the body or tissues of the mother, to respire, acquire nourishment, and eliminate wastes by bringing the bloodstreams of mother and young into close association but never into direct connection (**Fig. 1**). Placentation characterizes the early development of all mammals except the egg-laying duckbill platypus and spiny anteater. It occurs in some species of all other orders of vertebrates except the birds. In fact, in certain sharks and reptiles placentation is almost as well developed as in mammals. A few examples are also known among invertebrates (*Peripatus*, certain tunicates, and insects).

Placental modifications. With few exceptions the fetal structures used to establish placental relationships with the mother are modifications of organs present in kindred egg-laying (oviparous) species. In fishes, sharks and rays, and amphibians these include the gill filaments, tail fin, pericardium, and primitive yolk sac (midgut and adjacent ventral body wall). The essential placental modifications of all these are increased surface area and vascularity, and intimate contact with a highly vascularized and often secretory area of the mother. In fishes this highly vascularized area is usually the ovary; in sharks and rays, the uterus; in amphibians, the uterus or the skin.

Skin gestation is unique to certain oviparous South American frogs. As the eggs are extruded, they are fertilized by the male and placed on the back (in *Pipa*) or in a skin pouch on the back (in *Nototrema*). Here they become embedded in highly vascular compartments and receive both oxygen and nourishment from the mother's bloodstream. The live-bearing fishes Goodeidae have unique vascular rectal processes, trophotaeniae, which are known in no other group and which establish placental relationships with the ovarian tissue.

In amniotes (reptiles and mammals) the extraembryonic membranes utilized in placentation are specializations of basic membranes found in all oviparous amniotes. In fact, in the beginnings of placentation (portrayed by ovoviviparous lizards and snakes) the large shell-covered and yolk-laden eggs are simply retained within the uterus until hatched. Here apparently the only placental function is respiration; hence there is no obvious modification of the fetal membranes and only an increased and prolonged hypervascularity of the uterine lining.

In a few lizards and snakes and in all marsupial and eutherian mammals, complete placental function is necessary because the eggs are supplied with too little yolk to provide the needs of the embryo

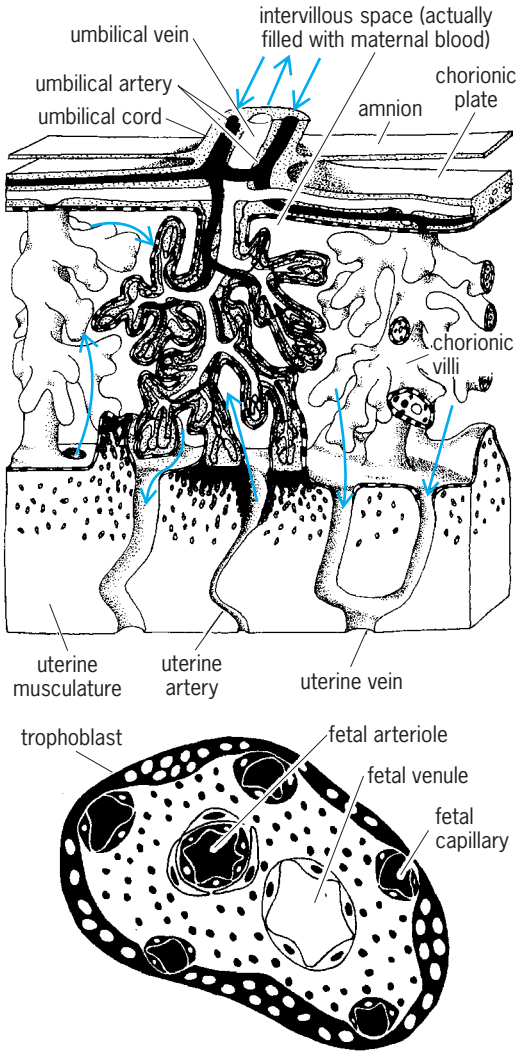

Fig. 2. Block removed from center of human placenta; cut end of branch villus is shown enlarged.

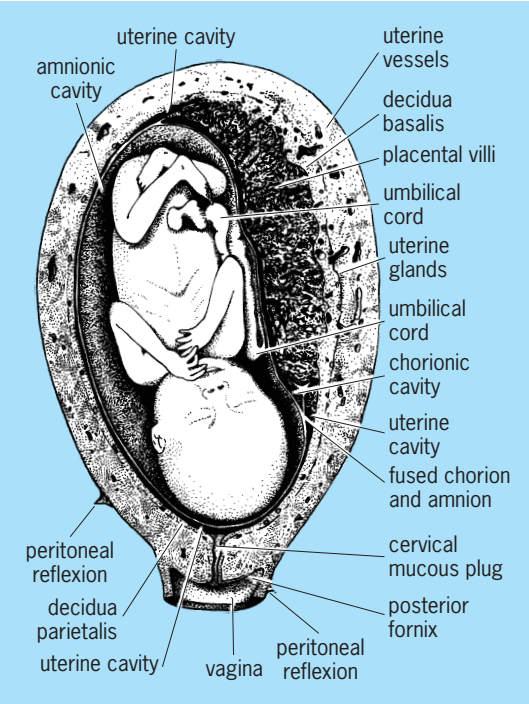

Fig. 1. Pregnant human uterus at 3½ months, split sagittally to show fetus, membranes, and placenta.

for nourishment until birth. Inadequate provision for elimination and storage of nitrogenous waste also occurs in most of these animals. Thus not only must these embryos interchange respiratory gases with the mother's blood, but they must also absorb nutriment and transfer wastes. Full placentation in these forms is provided mainly by anatomical and physiological specialization of three extraembryonic membranes: the chorion, yolk sac, and allantois. The amnion also probably plays a physiological role.

Physiological exchange. Efficient interchange depends on close proximity of large areas of fetal tissues to maternal blood and glandular areas. This is provided in mammals by a remarkable regulatory cooperation between the developing outer layer (trophoblast) of the chorion, together with the vascular yolk sac or allantois or both, and the mother's uterine lining (endometrium). In the typical mammalian placenta, which is always formed by the chorion and the allantoic vessels, the fetal and maternal bloodstreams are as close as a few thousandths of a millimeter from each other (**Fig. 2**). The surface area of the fetal villi which contain the functional fetal capillaries is probably several times larger than the body surface of the female. In humans this ratio is known to be about 8:1. Not only are the two vascular tissues closely

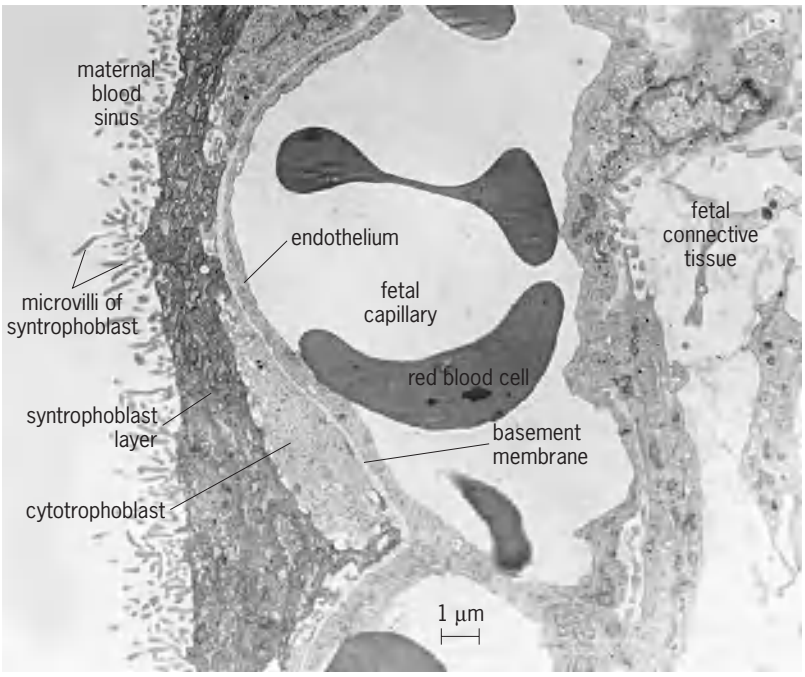

Fig. 3. Low-power electron micrograph of the surface of a section through the placental villus of a macaque monkey at 100 days of pregnancy (normal gestation period 164–168 days). This is very similar to the human placenta of similar stage.

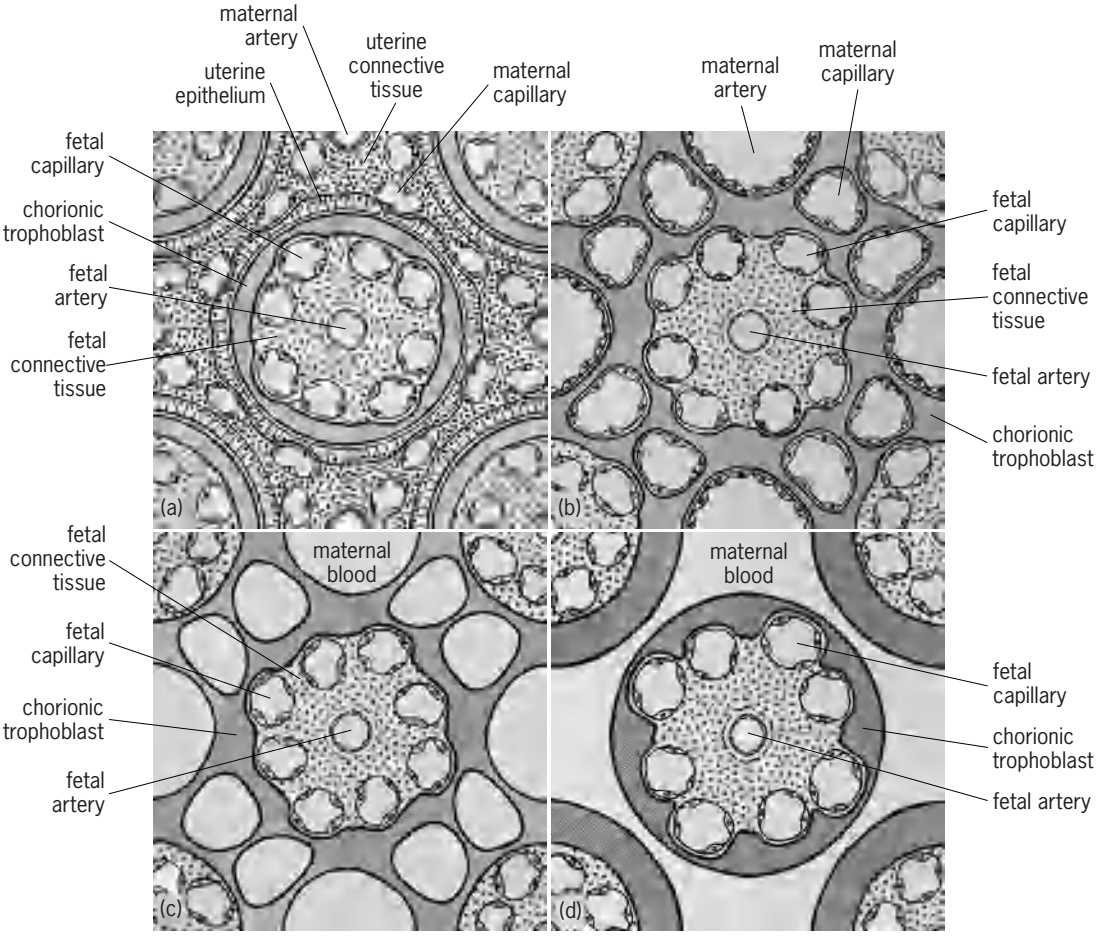

Fig. 4. Schemata of four fundamental types of placentation. (*a*) Epitheliochorial (villous); examples: hoofed animals, whales, lemurs. (*b*) Endotheliochorial (labyrinthine); examples: carnivores, bats, elephants, sloths. (*c*) Hemochorial (labyrinthine); examples: rodents, many insectivores, tarsiers. (*d*) Hemochorial (villous); examples: human, great apes, monkeys.

approximated, but in many mammals the arrangement of the vessels is such that in the area of interchange the two bloodstreams flow in opposite directions. This counterflow principle could greatly increase the efficiency of interchange, but clear physiological evidence that it actually does so is lacking.

Evidence from electron microscopy shows that the surface of the human villi exposed to maternal blood is covered by multitudes of microvilli (**Fig. 3**). These are minute fingerlike projections of the cytoplasm and its ultrathin covering, the plasmalemma (surface membrane of a cell). Obviously, microvilli increase the surface area manyfold beyond the 8:1 ratio mentioned above. However, the effect of microvilli on physiological interchange between mother and fetus is more complicated than it appears; although microvilli greatly increase the surface area for simple diffusion, they actually reduce it for such processes of absorption as endocytosis, which occurs only on the plasmalemma between the bases of the microvilli. *See* ENDOCYTOSIS.

The physiology of interchange through the placental membrane is a fertile field in research. Although simple physical diffusion and osmosis are factors, active membrane transport mechanisms are probably of greater importance. This transport through the placental separation membrane involves work done as the result of energy release within the plasmalemma and adjacent cytoplasm.

Types of placentation. Placentas are classified in various ways, but the most meaningful is based on the identity of the layers making up the separation membrane between the two bloodstreams. On this basis the human placenta is hemochorial; that is, the maternal blood is in direct contact with the chorionic trophoblast. The placenta of the dog is endotheliochorial; that is, the maternal blood is separated from the chorion by the maternal capillary endothelium. Another common condition is the epitheliochorial (**Fig. 4**).

Electron microscopy has shown the existence of ultrathin cellular layers previously missed by light microscopy. For instance, the trophoblastic layer may be single, double, or even triple. It is single in the mature human placenta, double in the rabbit, and triple in the rat. For these, A. C. Enders coined the terms hemomonochorial, hemodichorial, and hemotrichorial, respectively. *See* FETAL MEMBRANE.

Harland W. Mossman

Bibliography. P. Kaufmann and B. F. King (eds.), *Structural and Functional Organization of the Placenta*, 1982; K. R. Page, *The Physiology of the Human Placenta*, 1993; C. W. Redman, I. L. Sargent, and M. Starkey, *The Human Placenta*, 1993.

Placer mining

The exploitation of placer mineral deposits for their valuable heavy minerals. Placer mineral deposits consist of detrital natural material containing discrete mineral particles. They are formed by chemical and physical weathering of in-place heavy minerals, which are then concentrated through the action of wind or moving water. This concentration can be done through wave and current action in the ocean (beach and offshore placers), glacial action (moraine placers), wind action removing the lighter material (eolian placers), or the action of running water (stream placers). Stream placers are the most important of these deposits because of their common occurrence and their highly efficient concentration mechanisms. Marine placers, primarily beach placers, are the next most economically important, with the potential of offshore placers being the most recent to be recognized and developed. *See* MARINE MINING; ORE AND MINERAL DEPOSITS.

Specific gravities of several important placer minerals

Mineral	Specific gravity
Gold	15–19
Platinum	14–19
Diamond	3.5
Cassiterite (SnO_2)	6.8–7.1
Magnetite (Fe_3O_4)	5.2
Rutile (TiO_2)	4.2
Ilmenite ($FeTiO_3$)	4.7
Monazite $(Ce,La,Y,Th)PO_4$	5.0–5.3
Quartz	2.6
Water	1.0

Minerals that are concentrated in placer deposits are a result of differences in specific gravity and, therefore, the economically important deposits are for minerals with high specific gravities (see **table**).

Precious metals, primarily gold and platinum group metals, have been the most important product from placer mines. Their extremely high specific gravity coupled with their low chemical reactivity means that these minerals are efficiently concentrated in a placer environment and can be effectively recovered in a readily usable form. Historically, placer mines were civilization's first sources of gold and platinum and, until recent times, have continued to yield significant production. Although most modern gold is produced from lode, or "hard rock," deposits, the placer deposits of northern Canada, Alaska, and Siberia represent a virtually untapped source of the metal. *See* GOLD; PLATINUM.

Of more importance than gold are placer diamond deposits. Another important placer mineral is cassiterite, an ore of tin. Additionally, rutile and ilmenite, the principal ores of titanium, are found in commercial quantities only in beach placers. These same types of placers also yield monazite, a source of the rare earths yttrium, lanthanum, cerium, and thorium. *See* CASSITERITE; DIAMOND; ILMENITE; MONAZITE; RUTILE.

Although the importance of placer gold may have decreased over the years, placer mining still is an important component of modern industrial economy.

Methods. Most placer mining operations involve surface mining methods, although underground methods are sometimes used.

Underground mining. Underground mining of placer deposits is limited to some frozen auriferous gravels in the Arctic and the ancient fossilized gold bearing channels typically found in the "mother lode" country of California and the deep leads (placer gravels) of Australia. These types of deposits are typically worked by sinking a shaft to bedrock and then driving a series of drifts and crosscuts along bedrock to form a room-and-pillar mine. *See* UNDERGROUND MINING.

One unique aspect of drift mining is its application to frozen gravels in the Arctic. The temperature of the gravels ranges from 18 and 22°F (−8 and −5°C) with bedrock temperatures between 8 and 14°F (−13 and −10°C). The gravel contains 10% or more of ice by weight and while frozen provides competent support for the mining method. Rather than a conventional drilling and blasting of the frozen gravels (which are very resistant to this), steam points are driven into the faces and the cohesive ice is melted, allowing the loosened gravel to be loaded without fragmentation.

Surface mining. This proceeds in a manner similar to most trip mining operations. First the overburden is removed, and then the placer ground containing the valuable material is processed. This overburden may consist of material similar to the valuable deposit, such as the barren top section of a stream placer deposit where the gold is concentrated near or on bedrock. The character of the overburden may also be distinct from that of the value-containing deposit, for example, the windblown sands covering the diamond-bearing stream channels in South-West Africa. *See* SURFACE MINING.

In general, overburden is removed in a manner similar to the mining of the valuable material, as by mobile power equipment, dredges, and so on. One method unique to placer mining is the use of high-pressure water to remove the overburden. Hydraulicking involves shooting a stream of water against the overburden bank from a large nozzle (2–10 in. or 5–25 cm in diameter) on a device called a hydraulic monitor or hydraulic giant. This method is commonly used when the overburden contains a high percentage of clay or when it is necessary to strip muck from frozen gravels in arctic areas.

Muck consists of fine windblown material (loess), organic material, and water or ice. Hydraulicking has been the only effective method of removing frozen muck. All other methods require that the muck first be thawed. Common practice in Alaska and the Yukon Territory is to thaw with cold water, 50–60°F (10–16°C), pumped into boreholes through $^3/_8$-in.-diameter (10-mm) pipes called sweaters. Cold-water thawing has proven more cost-effective than either steam or hot water. *See* LOESS.

Techniques. For the actual mining of a placer deposit there are basically four different techniques: hand methods, hydraulic mining, use of mobile power equipment, and dredging.

Hand methods. In this technique, once the mainstay of placer operations, the mined material is shoveled into stockpile for hand washing or directly into a rocker box or "long tom" sluice box. This type of placer mining can be traced back to earliest recorded history, and was still a major factor through the Alaskan gold rush at the turn of the century. While hand mining now accounts for only a small portion of total placer production, it is often the most effective method for sampling prospective placer deposits.

Hydraulic mining. As mentioned above, hydraulic giants (monitors) have often been used to remove overburden. In addition, hydraulic pressure can be used to wash away the actual deposit. The resulting slurry is then channeled through sluices and the valuable material is recovered. Due to environmental regulations regarding water quality standards, the use of hydraulic giants in the United States has practically disappeared. They are still employed in the tin mines of Malaysia and Thailand, primarily for overburden removal prior to dredging. Another variation of hydraulic mining, called ground sluicing, uses water under natural pressure to wash the gravel through the sluices. Here, a stream is diverted and used to wash away the gravel bank. If the stream is first dammed and then released all at once, the process is called booming.

Mobile power equipment. The use of diesel powered equipment such as dozers, scrapers, front-end loaders, and hydraulic excavators has caused a revolution in placer mining technology. Not only does this equipment greatly increase the productivity of the miner for a relatively modest cost (compared to a floating dredge), but the mobility and flexibility of the equipment can be used for overburden removal, mining, and reclamation after mining. As a result, most modern placer mining operations will make use of mobile power equipment either in part or in total.

Dredging. Because of the intimate association of placer deposits and water, many of the more valuable deposits are located below the local water table level. If the water inflow is not severe, pumps can be used to dewater the cut and standard surface mining equipment can be used. However, if the water inflow is significant, such as with a modern beach placer or one located under a flowing river, then mining must be done with equipment that can dig below the waterline. This is called dredging.

There are three types of dredging operations used in placer mining: bucketline dredging, hydraulic dredging, and dragline excavation. The bucketline dredge (**Fig. 1**) is a floating excavator which uses a series of buckets attached to an endless rope. The bucketline travels from the floating dredge down to the lower tumbler at the end of the digging ladder, where it digs into the deposit, is pulled back to the surface, and dumps into the concentrating plant on the dredge. After concentration, a stacker conveyor deposits the waste material behind the dredge.

A cut is made by swinging the bucket ladder from side to side at a constant depth. At the end of each cut, the ladder is lowered and a new cut is begun. This continues until bedrock is reached or the maximum cutting depth is attained. Once a series of cuts

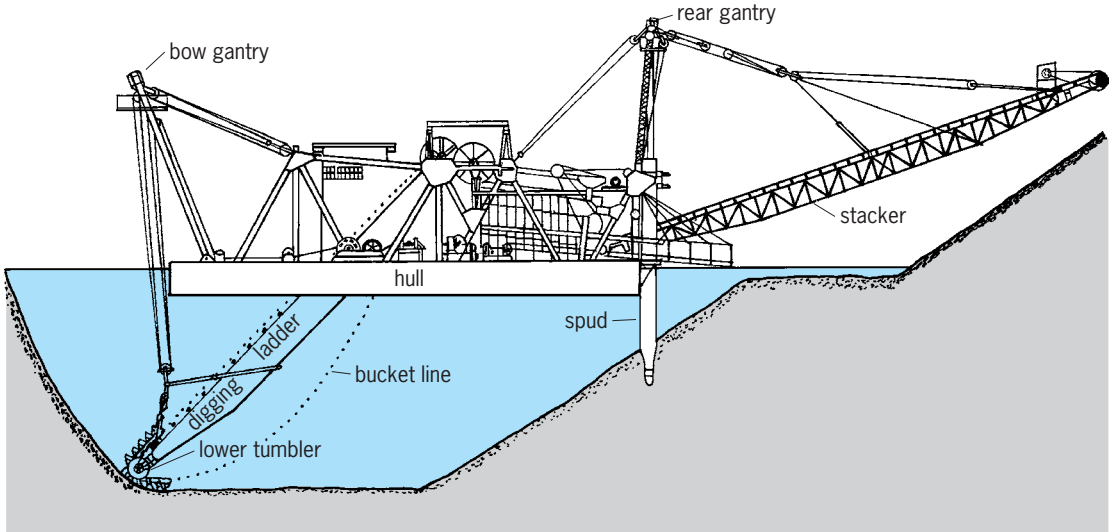

Fig. 1. California bucketline dredge.

has been carried to maximum depth, the dredge is moved forward into the newly extended dredge pond and the process is begun again. In this manner, the dredge moves its pond along in front of it, filling in behind with processed gravel (**Fig. 2**).

The size of the buckets ranges 2–20 ft^3 (0.06–0.60 m^3) each. The average speed of the bucketline is usually 20–24 buckets per minute. This implies that the modern large dredge will produce 10,000–15,000 yd^3 (7600–11,500 m^3) per day. A 20-ft^3 (0.6-m^3) bucket dredge, mining tin in Malaysia, produced nearly 900,000 yd^3 (690,000 m^3) in 1 month, digging 150 ft (46 m) below waterline.

Hydraulic dredges are similar to bucketline dredges in that a digging boom is suspended from a floating barge and material is discharged on shore. The principal difference is that, instead of buckets, the excavation is done with a suction pump. A stream of high-pressure water is directed at the digging face near the suction intake to break up the placer material, which is then pumped to the surface. A variation of this process is the cutterhead hydraulic dredge, which uses a rotating cutting head to churn the placer material into suspension. Material mined with hydraulic dredges can be processed onboard the barge, or the slurry can be pumped ashore for processing and disposal.

Hydraulic dredges have found extensive use in offshore and beach placers where sand is the primary material to be moved. Two cutterhead dredges moving sediments and glacial till in Ontario, Canada, each averaged over 2,000,000 yd^3 (1,500,000 m^3) per month for 3 years. On the other end of the scale, small, 2–4-in.-diameter (5–10-cm) suction dredges are often used in rivers and streams for hand operations and prospecting.

The final type of dredging operation uses a dragline or clamshell excavator to dig below water level (**Fig. 3**). Mined material is usually deposited directly into the concentrating plant, which is either crawler-mounted on the shore or barge-mounted offshore. This type of operation is used on smaller deposits where the capital cost of a large dredge cannot be justified.

Concentration. There are two basic principles used to recover valuable minerals from mined placer material: size classification and gravity separation. These apply to all sizes of operations from the smallest hand mining to the giant bucketline dredges, although in some special cases material adhesion or electromagnetic methods may also be used.

Size classification. This is the screening of the feed to the plant in order to discard oversized material which contains no values. In this manner, the gravity concentrating circuit not only has less material to process but works more efficiently because of the narrowed size range. This technique is effective because valuable placer minerals are discrete grains, separate from the gangue (waste) material and usually fairly small. For example, in a placer gold mining operation where the maximum expected nugget size is $^1/_4$ in. (6 mm), all material larger than 1 in. (25 mm) could be discarded without losing any gold. This would eliminate all of the boulders and cobbles as well as a significant amount of gravel.

Classification is usually done wet either over a series of vibrating screens or through a trommel, a long rotating, inclined cylinder with holes in its surface. In a trommel the material is fed into one end with sufficient water to wash the finer sands out through the holes, while the washed oversize is discharged out the other end, usually onto a stacking conveyor. *See* BULK-HANDLING MACHINES; MECHANICAL CLASSIFICATION; MECHANICAL SEPARATION TECHNIQUES.

Gravity concentration. The principle of gravity separation relies on the difference in specific gravity between the valuable minerals and the gangue.

1. *Panning.* The most simple gravity concentrator is the gold pan. In North America, these pans are flat-bottomed metal or hard plastic, 2–3 in. (5–7.5 cm) deep and 6–18 in. (15–23 cm) in diameter with sloping sides at a 30–40° angle. In some countries a

Fig. 2. Aerial oblique view upvalley over the site of a placer dredging, in which successive cuts, tailing patterns, and dredge pond are shown. The dredge appears in the left foreground. (*Pacific Aerial Survey, Inc.*)

12–30-in.-diameter (30–75-cm) conical wooden or metal bowl called a batea is used instead of a flat-bottomed pan. In either case the pan filled with the material is immersed in water, shaken to settle the heavy materials, and swirled to wash away the lighter minerals. The primary use of panning in modern placer mining is in prospecting and sampling. Some use is made of hand panning to separate gold from black sands (magnetite) where mercury amalgamation is not feasible.

2. *Sluice boxes.* A step up from hand panning, the sluice box is simply a trough with a series of riffles in the bottom. A mixture of placer material and water flows into one end and is washed down the inclined trough, which can vary in width from 12 to 60 in. (30 to 150 cm) and in length from a few feet to several hundred feet. The riffles, which can consist of rocks, wooden blocks, dowels, or angle irons, produce eddy currents in the flow, allowing the heavy material to settle out while the lighter material is washed away (**Fig. 4**). Plastic outdoor carpeting (artificial turf) is also sometimes used to produce these eddy currents. Mercury is often used in sluice boxes processing gold. The liquid mercury (specific gravity = 13.6) settles in behind riffles and is not washed away. When it comes in contact with gold, the gold particles adhere to the mercury, forming an amalgam. During cleanup of the sluice, the amalgam is collected and retorted to separate the gold; the cleaned mercury is recycled. Undercurrents, often used in conjunction with the primary sluice, are auxiliary sluices which handle specific

Fig. 3. Dragline dredge and pond. (*Bucyrus-Erie Co.*)

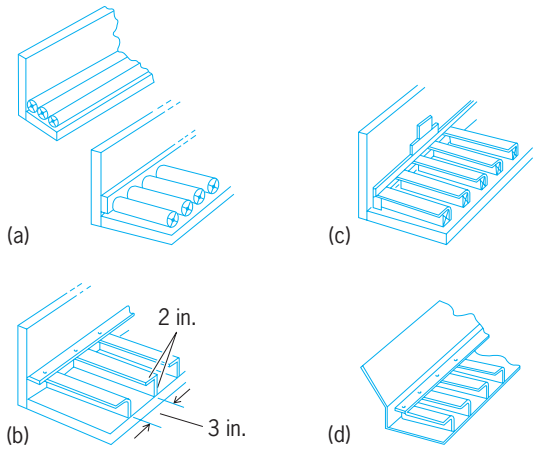

Fig. 4. Types of riffles in partial plan views. (*a*) Pole riffles. (*b*) Hungarian riffle (1 in. = 2.5 cm). (*c*) Oroville Hungarian riffle. (*d*) All-steel sluice. (*After G. J. Young, Elements of Mining, 4th ed., McGraw-Hill, 1946*)

size fractions of the sluice feed. Screens, gratings, and punch plates within the sluice separate these size fractions and increase the efficiency of the sluicing operation. Sluices are used to some extent in practically all placer operations. They are efficient, flexible and, because they have no moving parts, very low-cost concentrators.

3. *Spinning bowls.* Spinning bowls use principles similar to sluices. However, instead of using gravity as the settling force, they use the centrifugal force generated on a rapidly spinning surface. The sized placer material and water is fed into the center of the bowls. The heavy valuable minerals are trapped in baffles similar to riffles, and the lighter material is washed out over the rim. The advantage of the spinners is that a force higher than gravity can be generated, increasing the efficiency of the separation; this force can be easily adjusted through motor speed to "fine-tune" the recovery.

4. *Jigs.* Another mechanism used to recover heavy minerals, especially on the tin mining dredges in Malaysia, is the jig. A jig is basically a water-filled box into the top of which the slurried placer minerals are fed. A rapid series of upward pulses are generated in the water at the bottom. The heavy minerals sink through these pulses into a hutch at the bottom, while the lighter material is carried out over the edge.

5. *Other methods.* These include hand picking, especially in diamond processing; adhesion to specially prepared surfaces, as with the grease tables used to collect diamonds; magnetic separation, to separate magnetite and other iron-containing minerals; and other gravity techniques such as spiral classifiers, vibrating tables, and heavy-media separation.

Environmental concerns. The two major environmental problems associated with placer mining are water pollution and land disturbance. Both the mining and the processing of placer minerals require a great deal of water, and, once used, this water contains large amounts of suspended solids. If the water is allowed to run off into the rivers, these solids can have an adverse impact on the downstream environment. In suspension they can harm aquatic habitats, and when settled out can clog waterways and choke off irrigated crops. In 1884, legislation was passed in California banning hydraulic mining because sediments from the Marysville district were inundating crops in the Sacramento River Delta and causing navigational hazards in the San Francisco Bay.

In the United States, environmental laws, such as the Clean Water Act, set limits on the discharge of suspended solids from a mining operation. To operate within these limits, miners use a series of settling ponds downstream from the mine and recycle as much water as possible. *See* ENVIRONMENTAL ENGINEERING; WATER POLLUTION.

Since most placer mining is surface mining, surface disturbance is necessary, especially where dredging operations create vast piles of cobbles as mining progresses. In some arctic or subarctic areas the ground is very sensitive to disturbance. One method of land reclamation is a mining plan which stockpiles top soil (where possible), recontours the spoil (waste) piles, and then returns the land to useful status. This can be done by revegetation to match the original environment, creation of new recreational or agricultural areas or, in urban areas, creation of new industrial sites. These latter alternatives can be extremely beneficial to the local economy, especially in developing nations. *See* LAND RECLAMATION.

Danny L. Taylor

Bibliography. L. Cope and L. Rice (eds.), *Practical Placer Mining*, 1992; H. L. Hartman and J. M. Mutmansky, *Introductory Mining Engineering*, 2002; B. A. Kennedy (ed.), *Surface Mining*, 2d ed., 1990; R. L. Lowrie (ed.), *SME Mining Reference Handbook*, 2002; E. H. Macdonald, *Alluvial Mining: The Geology, Technology and Economics of Placers*, 1983.

Placodermi

A class of fishes (Pisces) known from the Devonian Period, and a few that survived into the base of the Carboniferous. They were true fishes or gnathostomes, and can be distinguished from other fishes by the following characters: the gill chamber extends far under the cranium and is covered laterally by opercula; there is a neck joint between the cranium and the fused anterior vertebrae; often there is also a coaxial joint developed between the dermal bones of the cranial roof and shoulder girdle; the head and shoulder girdle are covered with dermal bones composed typically of cellular bone and superficially of semidentine instead of true dentine; the bones are commonly ornamented with tubercles or ridges; the endoskeleton is cartilage and may be calcified in a globular fashion or perichondrally ossified; the notochord is persistent, and the vertebrae consist only of neural and hemal arches; the tail is diphycercal or slightly heterocercal, and an anal fin is lacking.

Most placoderms were bottom-dwelling fishes,

with the head and trunk dorsoventrally depressed; only the Stensioellida and some specialized arthrodires had laterally compressed and deepened bodies, suggesting a more nectic manner of life. Most of them were small or moderate-sized, but a few were, for their time, gigantic, reaching a length of as much as 20 ft (6 m). They were the dominant fishes of the Devonian, and are found in both marine and fresh-water deposits.

There have been many opinions regarding the relationships of the Placodermi. Some have considered them nearest to the Chondrichthyes, but their possession of dermal bones with unique histology as well as other characters makes this unlikely. Others have thought them to be relatives of the Osteichthyes, but this also is questionable. At present, it seems that they may best be considered a separate class of fishes, equivalent to the Chondrichthyes and Teleostomi.

The Placodermi are subdivided into nine orders, each with its own distinct specializations. The evolution and interrelations of these orders are matters of disagreement. According to the hypothesis of R. S. Miles and G. C. Young, primitive placoderms had well-developed head and trunk shields, the trunk shield consisting of a complete ring composed of two median and five paired plates: in different orders the trunk shield was either enlarged or reduced. According to R. H. Denison, the primitive placoderm shoulder girdle was short with few plates, and this was progressively ossified and enlarged in different orders, forming a trunk shield in some. The resultant differences in phylogenetic interpretations are illustrated in the two cladograms in **Fig. 1**. Arranged in order of increasing specialization according to Denison's hypothesis, the placoderm orders are as follows.

Class: Placodermi
- Order: Stensioellida
  - Pseudopetalichthyida
  - Rhenanida
  - Ptyctodontida
  - Anthracothoraci
  - Petalichthyida
  - Phyllolepida
  - Arthrodira
- Suborder: Actinolepina
  - Wootagoonaspina
  - Phlyctaeniina
  - Brachythoraci
- Order: Antiachi

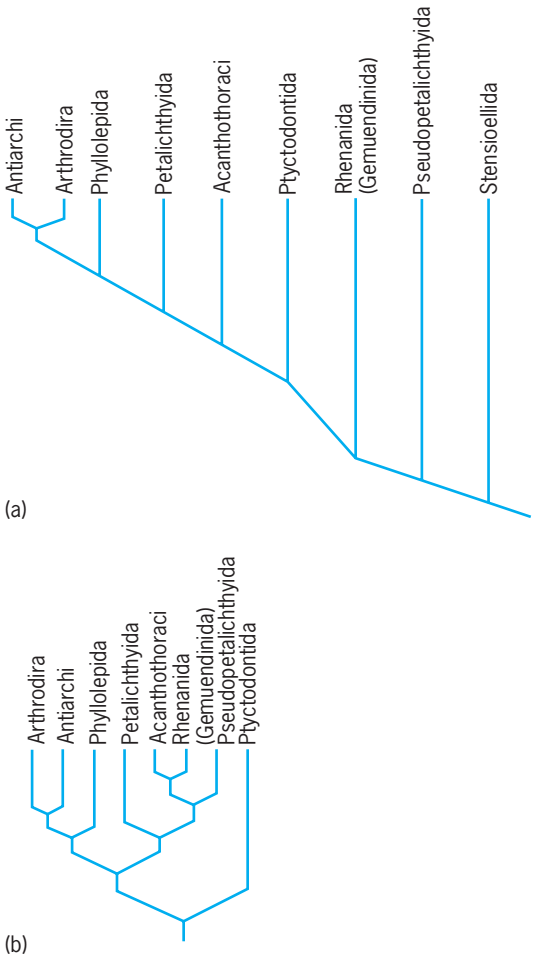

Fig. 1. Two cladograms portraying differing concepts of relationships of orders of placoderms: (*a*) according to Denison (*after R. H. Denison, Placodermi, vol. 2 of O. Kuhn and H. P. Schultz, eds., Handbook of Palaeoichthyology, 1978*); and (*b*) according to Miles and Young (*after R. S. Miles and G. C. Young, Placoderm relationships reconsidered in light of new ptyctodontids from Goga, Western Australia, Linn. Soc. Symp., 4:123–198, 1977*). Note particularly the absence of Stensioellida in *b*, the positions of Ptyctodontida, and alternative associations of some orders. See text for further explanation.

Stensioellida. This order is based on a *Stensioella heintzi* from the Lower Devonian Hunsrückschiefer of Germany. This is a slender, unarmored fish about 12 in. (30 cm) long (**Fig. 2**), quite different from other members of the class, but shown to be related by the presence of three small bones on its head, by the extension of the gill chamber far under the cranium, and by the articulation between the endocranium and the fused anterior vertebrae. The head is bluntly rounded and mostly covered with small denticles, the orbits are probably lateral, the mouth is ventral, and the jaws carry only small, pointed denticles. The shoulder girdle is short and not expanded into a trunk shield as in most placoderms. The body is gradually tapering and covered with small scales. The pectoral fins are long, slender, and flaplike, and the pelvic fins are small. This may be the most primitive known placoderm, but its relationships cannot be determined with certainty from the single known specimen.

Pseudopetalichthyida. This order is based on two genera, *Pseudopetalichthys* (**Fig. 3**) and *Paraplesiobatis*, from the Lower Devonian Hunsrückschiefer of Germany. They are slender fishes, 4 to 6 in. (10 to 15 cm) long, with prominent, dorsally placed orbits, a ventral mouth, and a skull covered with many small, bony plates and large, lateral gill covers. As in other placoderms, the gills extend far under the cranium, and there is a well-developed articulation between the cranium and fused anterior vertebrae.

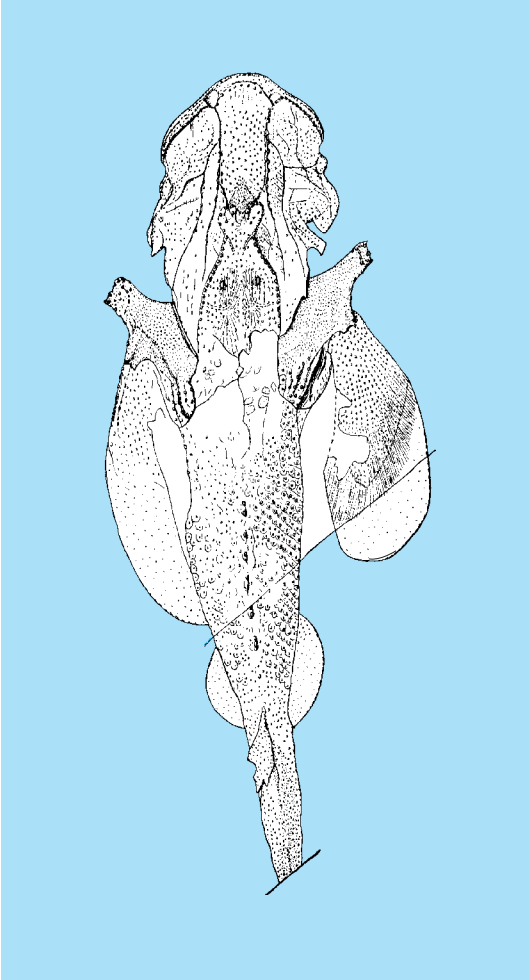

Fig. 2. *Stensioella heintzi* in dorsal view; length about 10 in. (25 cm).

The shoulder girdle is short and not expanded into a trunk shield, the pectoral fins are narrow-based, and the pelvic fins are relatively long-based. The trunk and tail may be covered with scales, and there is a series of crested, median dorsal scales. These fishes once were referred to the Stensioellida, but they differ in important respects.

Rhenanida. This order, also known as Gemuendinida, includes the genera *Gemuendina* from the Lower Devonian of Germany, *Asterosteus* from the Middle Devonian of Ohio, and *Jagorina* from the Upper Devonian of Germany. They are flattened, raylike fishes with large, semicircular pectoral fins, dorsal eyes and nostrils, and a terminal mouth (**Fig. 4**). The head shield consists of a few large plates separated by mosaics of small tesserae. The trunk shield is formed of only a single ring of plate, projecting ventrally far under the head.

The flat raylike braincase is not as broad as in arthrodires; it has small lateral processes and a pair of stout posterior processes for articulation with the anterior vertebrae, which are solidly fused into a single structure around the notochord. The jaws and gill covers were highly mobile, and the former could be protruded to open dorsally in front of the head. A stout hyomandibular suspended both the upper and lower jaw cartilages, which are well ossified and free of the braincase and dermal skeleton. There are no tooth plates, although the lips were probably armed with stellate tubercles. Small scales cover the whole body and the pectoral and pelvic fins in *Gemuendina*. The propulsive force in swimming came from the muscular body and not from the winglike pectorals, which were too stiffly armored to undulate as in rays.

The rhenanids are remarkably raylike and, like living skates and rays, were adapted for lying on the bottom of the sea and covering themselves with sand stirred up by the pectoral fins. It has been suggested that rhenanids were directly ancestral to rays, but this cannot be true as they differ from rays in the dorsal position of the nostrils and numerous other anatomical characters. The similarities are due to convergent evolution associated with closely similar modes of life. *See* CHONDRICHTHYES.

Ptyctodontida. These are small fishes from the Devonian and lowest Carboniferous of North America,

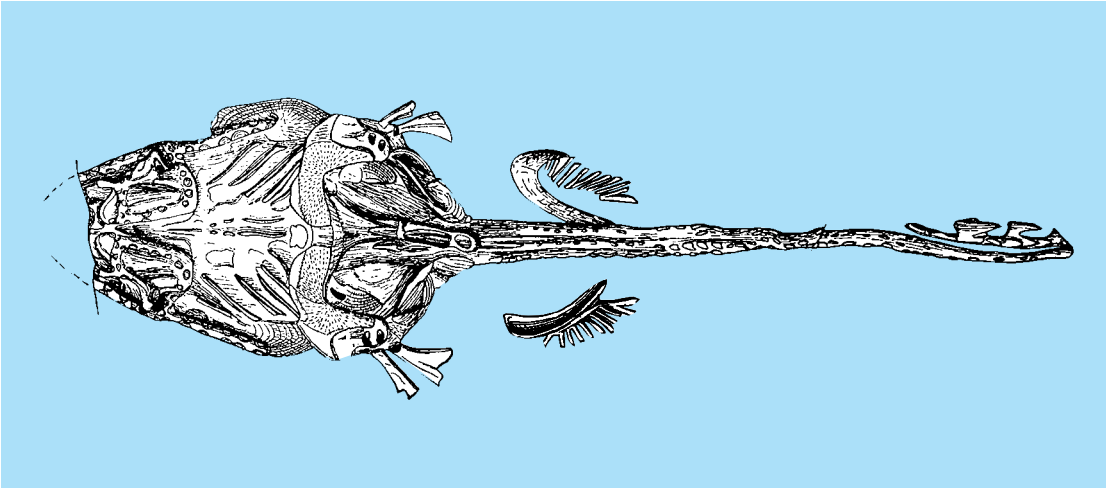

Fig. 3. *Pseudopetalichthys problematica*, ventral view; length about 6 in. (15 cm).

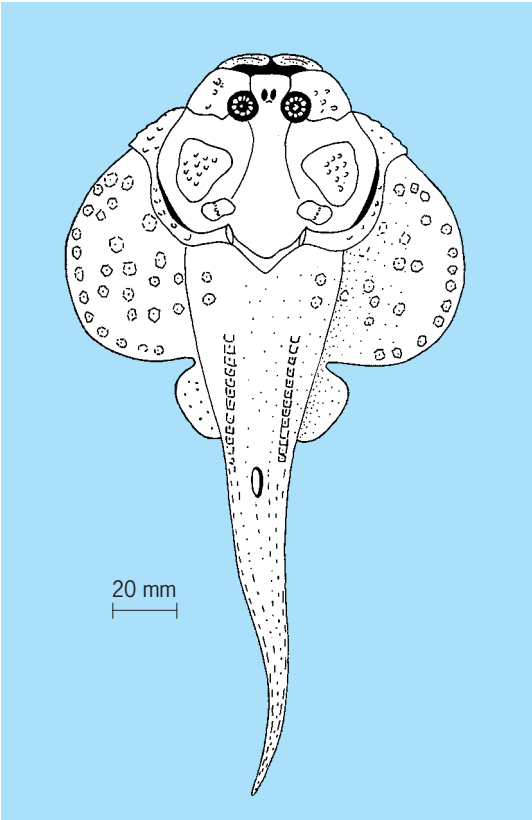

Fig. 4. ***Gemuendina stuertzi*** **from the Lower Devonian of Germany; restoration in dorsal aspect.**

Eurasia, Africa, and Australia. They include both marine and fresh-water species. Usually only the stout tooth plates are preserved, but ossified internal skeletons are found in the Scottish genus *Rhamphodopsis* (**Fig. 5**) and the German *Ctenurella*.

Both the head and trunk shields are present, and the joint between them is well differentiated and variable in structure. The head shield has a greatly reduced number of bones which show some similarity to those of petalichthyids in their arrangement and in the central X-anastomosis of pairs of the deeply situated laterosensory lines. The upper jaw cartilages are superficially ossified, large, and firmly attached to the braincase, whereby they gain support. They carry a single pair of large tooth plates which are opposed by a smaller pair of plates in the lower jaw. The tooth plates are variously modified, for shearing in *Rhamphodopsis* and *Ctenurella*, and for crushing with either a large central area (*Ptyctodus*) or separated cups (*Palaeomylus*) of a hard dentinal tissue. Rostral and labial cartilages supported the fleshy snout and lips. The trunk armor is short, and in *Rhamphodopsis* has long median dorsal and lateral fin spines, absent in *Ctenurella*. The vertebral column resembles that of arthrodires, with an unconstricted notochord and separate dorsal and ventral arches with long spines. There are two dorsal fins with an internal skeleton, and a small dorsal lobe in the caudal fin. The pectoral fins are large and mobile with a narrow base. The large pelvic fins are remarkable for their sexual dimorphism associated with internal fertilization, as in Chondrichthyes. In males, claspers bearing a laterally toothed dermal plate project ventrally from the root of the fins; they worked against a spiny plate on the ventral surface. In front of the fins are holocephalanlike prepelvic claspers, also bearing sharp spines. In females the fins are supported by large, immobile endoskeletal plates, and in *Rhamphodopsis* are clothed by numerous, overlapping ventral scales. The prepelvic claspers of males are represented by flat plates in the skin.

Ptyctodonts are of great interest as possible holocephalan ancestors, but because they resemble living rabbitfishes and are quite dissimilar from their fossil relatives, it seems possible that the resemblances merely reflect a similar benthic mode of life as a result of convergent evolution. According to R. S. Miles and G. C. Young, ptyctodonts are collateral descendants of all other placoderms, but Denison considers them primitive in retaining a short shoulder girdle.

Acanthothoraci. This order, alternatively called Palaeacanthaspidodea, is known from eight genera from the Lower Devonian of Eurasia and arctic Canada. Their relationship to other placoderms is uncertain. The skull is long, rather narrow, and covered with dermal bones and sometimes with a mosaic of small tesserae; it is characterized by the presence of two pairs of paranuchal plates (**Fig. 6**). The orbits are lateral or dorsolateral, and the nostrils dorsal and bounced in front by a premedian plate, unknown in other placoderms except Antiarcha. The shoulder girdle is short and high, and topped by a median dorsal plate that often has a high spine.

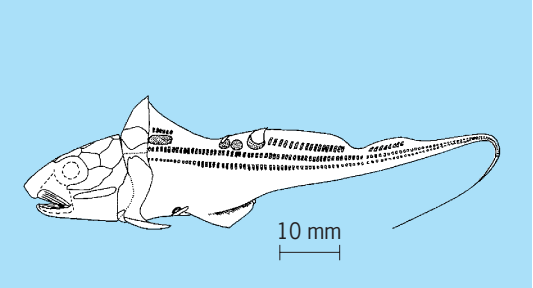

Fig. 5. ***Rhamphodopsis threiplandi*** **from the Middle Devonian of Scotland. Restoration of skeleton with outline of body in lateral aspect.**

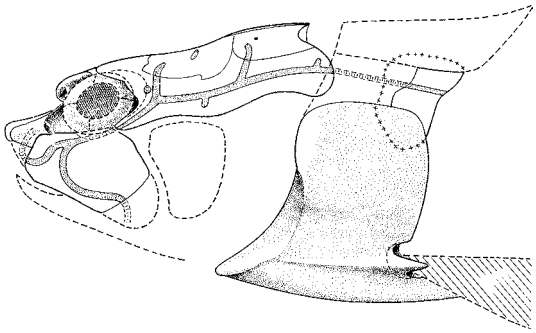

Fig. 6. ***Romundina stellina*****, head and shoulder girdle in lateral view; length about 1.2 in. (3 cm).**

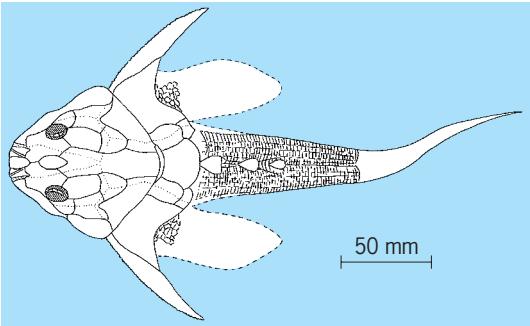

Fig. 7. ***Lunaspis broilii*** **from the Lower Devonian of Germany; restoration in dorsal aspect.**

Petalichthyida. This small order of dorsoventrally flattened placoderm fishes ranges from the Lower to the Upper Devonian. They are never common and are an exclusively marine group, seldom exceeding 1.7 ft (0.5 m) in length. Typical genera include *Lunaspis* (**Fig. 7**) from the Lower Devonian of Germany and *Macropetalichthys* from the Middle Devonian of North America and Germany.

The external armor is formed by a head and trunk shield, each composed of large plates. The head is characterized by the dorsal position of the orbits, which are completely enclosed in the shield, and the long occipital region which has a marked effect on the bone pattern in the hind end. The laterosensory lines are deeply sunk in the bones and open to the outside through small pores; further, a pair of canals meet in the center of the shield in a characteristic X-shaped junction. The braincase has a long, narrow occipital region, but is otherwise broad and flat with prominent lateral processes in the postorbital and occipital regions. The jaws and tooth plates are unknown. The trunk armor is short, with long spines in front of the pectoral fins and a poorly differentiated articulation with the head shield. In *Lunaspis* the body posterior to the armor is covered with scales, including three large dorsal ridge scales in place of dorsal fins.

Two genera from the Lower and Middle Devonian of China differ from typical members of the order, the Macropetalichthyidae, in having lateral orbits. They are believed to represent a more primitive family, the Xinanpetalichthyidae.

Although they survived through the Devonian, the evolutionary history of the petalichthyids is poorly known. Evidently they were benthic fishes that occupied a stable adaptive zone. Their adaptations are similar to those of some primitive Arthrodira, but their shorter trunk shield, absence of pectoral fenestrae, and heavily scaled tail suggest that they are more primitive than members of that order.

Phyllolepida. This order is known principally from a single genus *Phyllolepis* from the Upper Devonian. They were fresh-water fishes about 1 ft (0.3 m) long. The genus is widely distributed in Europe, Greenland, and Australia.

The armor is broad and flat; the component plates are easily recognized by their characteristic ornamentation of concentric and transverse ridges. The head shield includes a large central plate, surrounded by smaller plates which are not easily compared to the plates of other placoderms. Shallow grooves which converge on the growth center of the central plate mark the course of the laterosensory lines. There is no sign of the orbits, so unless these fishes were blind the eyes must have lain anterior or lateral to the bony shield (as shown in **Fig. 8**). The trunk shield has fewer plates than in Arthrodira, lacking most of the posterior series of plates of that order. The spinal plates are stout and the joint with the head shield poorly developed. The braincase, tooth plates, and jaws are not known. One specimen from Scotland has elements of the vertebral column. There are separate dorsal and ventral arches, which in life sat astride the notochord.

Phyllolepis was adapted for life on the bottom in streams and floodplains, where it probably fed on organic detritus. Its ancestry is unknown, although a Middle or Upper Devonian genus from Antarctica, *Antarctaspis*, may be a primitive member of the order. New phyllolepid material from Australia has led to the suggestion that phyllolepids may be related to primitive actinolepid arthrodires.

Arthrodira. These are the best-known, most varied, and most successful order of Placodermi, found in fresh-water and marine sediments of Devonian and lowest Carboniferous age in all continents except South America. The pattern of their skull bones is characteristic, the orbits are usually lateral, and the nostrils are anterior. Their jaw suspension is unique: the palatoquadrates attach anteriorly both to the suborbital dermal bones and to the neurocranium; posteriorly they attach to the postsuborbital dermal bones. The hyomandibuli do not aid in jaw suspension, but primitively hinge the gill covers to the endocranium. The dermal jawbones are two pairs of supragnathals and one pair of infragnathals. The dermal shoulder girdle is enlarged to form a trunk shield which typically consists of three median and eight paired bones. The dermal cranio-thoracic joint is primitively of a sliding type, but in most it is ginglymoid, with condyles on the trunk shield fitting in glenoid fossae on the skull. Long, low scapulocoracoids support the pectoral fins, which are primitively narrow-based

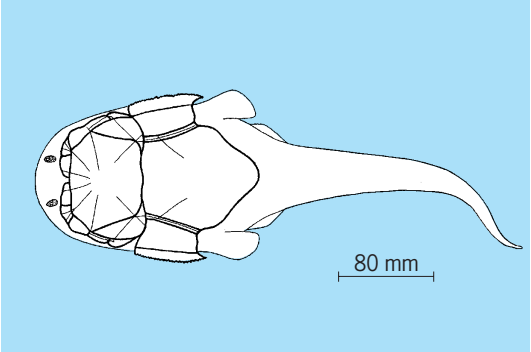

Fig. 8. ***Phyllolepis orvini*** **from the Upper Devonian of Greenland. The restoration of the dermal armor is shown in dorsal aspect. The outline of the front of the head and the body is hypothetical, but shows how the fish might have appeared in life.**

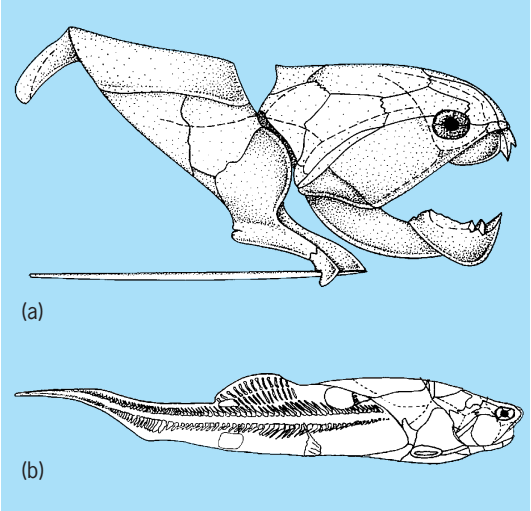

Fig. 9. Representative arthrodires. (*a*) Cranial and thoracic shields of giant Upper Devonian *Dunkelosteus* (*Dinichthys*), a brachythoracic arthrodire showing reduction of thoracic shield, large nuchal gap, and elevated cranio-thoracic joint. Length of originals 8–10 ft (about 3 m) (*after E. H. Colbert, Evolution of the Vertebrates, John Wiley and Sons, 1955*). (*b*) *Coccosteus*, a moderately derived Middle Devonian brachythoracic arthrodire, showing the structure of the postcranial axial skeleton and fins. Note fenestra for pectoral fin located in the basal-lateral portion of the thoracic shield. Original about 13.5 in. (35 cm) in length (*after R. S. Miles and T. S. Westoll, Trans. Roy. Soc. Edinburgh, vol. 67, 1968*).

and articulate through fenestrae in the trunk shield. In more advanced arthrodires, the lateral walls of the trunk shield are reduced so that the pectoral fenestrae can enlarge and open behind, permitting the bases of the pectoral fins to become long–based.

The four suborders of the order Arthrodira as listed in the classification of the Placodermi represent formal taxonomic units; as many as six such units are recognized in alternative classifications. There is no clear evidence that these suborders are monophyletic. Thus, the suborders may be treated informally as evolutionary grades, without an implication that the included species had a common specific base. This way of viewing the relationships of the arthrodires was suggested by Miles in 1969 in his study of their feeding mechanisms and is accepted in the following discussions.

The most primitive known arthrodires are from the Lower Devonian. Along with a number of Middle Devonian forms, they represent the organizational level of the Actinolepina. In the most primitive genera the cranial and thoracic shields are extensive, covering about 60% of the body. The remainder of the body was heavily scaled (**Fig. 9***a*). The intershield joint was not present. The plates tended to slide past one another, in contrast to those of less primitive arthrodires in which a joint was present. Altogether some 15 genera have been assigned to this group based on specimens from deposits in Europe, North America, eastern Asia, and Australia. Such primitive forms were at the base of a complex, adaptive radiation of arthrodires which climaxed in the Upper Devonian.

Among the probable derivatives of the Actinolepina were the distinctive form *Wuttagoonaspis* of the Middle Devonian of Australia, the sole genus of the suborder Wuttagoonaspina, and members of the Phlyctaeniina. Both groups are similar in structure to the Actinolepina, but a cranio-thoracic joint is present. The spines that lay anterior to the pectoral fins are strongly developed, and the median dorsal plates are elongated. *Dicksonosteus* from the Lower Devonian is a fairly representative example. A number of lines of specialized genera, predominantly Middle Devonian in age (Fig. 9*b*), arose from Lower Devonian forms. About 20 are assigned to the family Phlyctaenidae and another 10 to the family Holonematidae. In all of them the basic structural plan is preserved, but superimposed upon it are specializations related to various modes of feeding and locomotion, indicating considerable habitat diversification. The Brachythoraci include an extensive array of arthrodires grouped into five superfamilies. Approximately 60 genera of Brachythoraci have been assigned to the listed superfamilies. In addition, some 35 others are not currently placed in any of them, being either too specialized or too poorly known. Together these genera provide the tangible evidence of a massive adaptive radiation of the derived arthrodires.

As the Devonian radiations of the arthrodires were taking place, the cartilaginous fishes (Chondrichthyes) and the bony fishes (Osteichthyes) were undergoing somewhat similar expansions. Sharks in particular paralleled the predaceous arthrodires, but were more active. Members of both groups undoubtedly were direct competitors with the placoderms and probably played a role in what appears to have been a rather rapid placoderm extinction during the Late Devonian and very early lower Carboniferous. Changes in invertebrate animals and in the physical environment also contributed to the demise of

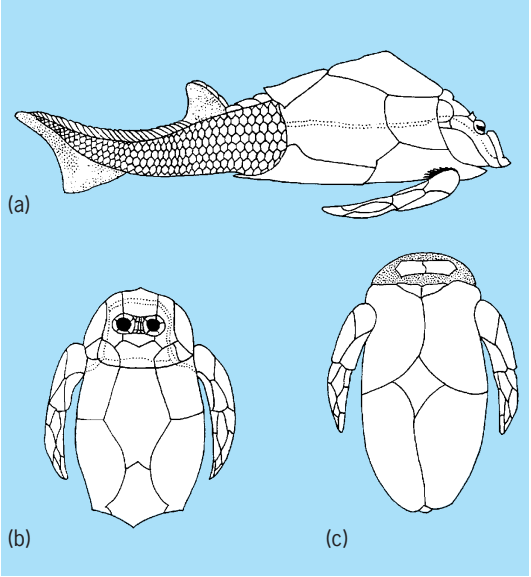

Fig. 10. *Pterichthyodes* (*Pterichthys*), a Middle Devonian antiarch. (*a*) Lateral view showing scale-covered tail; about 15 cm (6 in.) in length. (*b*) Dorsal and (*c*) ventral views of the armor. (*After Traquair*)

placoderms. Just how rapid this extinction actually was and its precise causes remain uncertain, masked by the incompleteness and biases of the fossil record as now known. What is clear is that while the last of the arthrodires died out, the cartilaginous and bony fishes remained successful and continued to expand their roles in the waters of the oceans and the continents. *See* OSTEICHTHYES.

Antiarcha. This is a highly specialized order, characterized by heavy armor consisting of a short head shield and a very long, flat-bottomed trunk shield (**Fig. 10**). Most antiarchs are small or moderate in size, the largest reaching a length of about 48 in. (120 cm). The head shield is so greatly modified that it is difficult to homologize its bones with those of other placoderms. The eyes and nasal openings occupy an orbital fenestra on top of the skull; in front of them is a premedian plate, unknown in most other placoderms. The mouth is ventral. There is an articulation between the dermal bones of the head and trunk shields, but contrary to the condition in Arthrodira, the condyles are on the head shield, and the glenoid fossae on the trunk shield. The trunk shield has both anterior and posterior median dorsal plates. There are peculiar pectoral appendages, completely covered by small plates, usually jointed near midlength, and articulating in axial fossae on the sides of the trunk shield. The body behind the shield is scaled or naked, and has one or two dorsal fins and a heterocercal caudal fin; pelvic fins are possibly represented by "frills" behind the shield of *Bothriolepis*.

Antiarcha were predominantly fresh-water stream and lake dwellers, though a few lived in the sea. They were benthic forms, propelling themselves along the bottom by their pectoral appendages, and surely feeding on the bottom; one genus, *Bothriolepis*, has been shown to have been a mud grubber by the preserved content of its spiral intestine. Antiarcha probably evolved from primitive Arthrodira. The most primitive, the Yunnanolepidae, come from the Lower Devonian of China. The other families, Asterolepidae, Bothriolepidae, and Sinolepidae, are found in the Middle and Upper Devonian and perhaps lowest Carboniferous of North America, Greenland, Eurasia, Australia, and Antarctica.

Robert H. Denison; Everett C. Olson

Bibliography. R. H. Denison, Further considerations of placoderm evolution, *J. Vert. Paleontol.*, 3:69–83, 1983; R. H. Denison, *Placodermi*, vol. 2 of *Handbook of Palaeoichthyology*, edited by O. Kuhn and H. P. Shultz, 1978; B. G. Gardiner, The relationships of the placoderms, *J. Vert. Paleontol.*, 4:379–395, 1984; R. S. Miles and G. C. Young, Placoderm relationships reconsidered in light of new ptyctodontids from Goga, Western Australia, *Linn. Soc. Symp.*, 4:123–198, 1977.

Placodontia

A small group of marine reptiles forming part of the Sauropterygia, known only from deposits of Triassic age in Europe, the Near East (Israel), North Africa (Tunisia), and China. The earliest occurrence of placodonts is in the uppermost Lower Triassic (upper Buntsandstein) of Upper Silesia (Poland); their latest occurrence is in the very late Triassic (Rhaetian) of the Southern Alps. Placodonts are reptiles with a grossly specialized dentition—flat-crowned teeth located in both the upper and lower jaws and on the palate—that functioned as a crushing device for hard-shelled invertebrate prey such as brachiopods, gastropods, clams, and cephalopods (see **illustration**). *See* SAUROPTERYGIA.

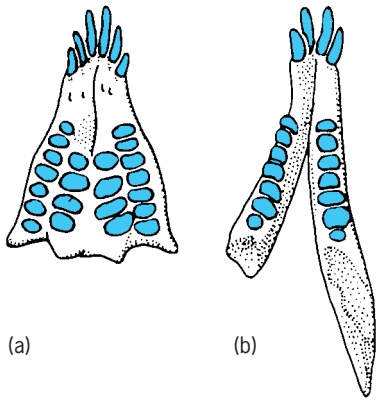

Dentition of *Paraplacodus broilii*, Triassic, Switzerland. (*a*) Upper jaw. (*b*) Lower jaw.

Morphology. Placodonts are divided into two clades: the unarmored Placodontoidea (*Placodus, Paraplacodus*) and the armored ("turtlelike") Cyamodontoidea. The skeleton of unarmored placodonts shows few adaptations to life in nearshore habitats or shallow epicontinental seas. The thorax was box-shaped, and the digits of hand and foot were probably webbed; the joint surfaces and muscle attachment sites on the limb bones suggest that the animals did little if any walking on land. The most serious modifications to an aquatic environment occurred in the skull, and are correlated with the development of the durophagous (crushing) dentition. The entire skull became massive in relation to the body as a result of pachyostosis (thickening of the bone) and partial fusion of the skull elements; the coronoid region of the mandible, to which most of the jaw-closing muscles were attached, became greatly extended upward, giving it a superficial similarity to a mammalian jaw. These changes provided optimal leverage for these muscles while limiting reaction forces acting on the jaw joint. *See* SKULL.

The advanced cyamodontoids, such as *Placochelys*, resemble sea turtles in overall appearance and in the fact that their body was encased in an armor of dermal bones similar to that of the extinct dermochelyid sea turtle *Psephophorus*. Such similarities with turtles are, however, entirely due to convergent evolution. Cyamodontoids were much more diverse than placodontoids, evolving a variety of skull shapes that indicate different feeding strategies. *Henodus* from the early Late Triassic of southern Germany had a much reduced dentition, consisting only of a pair of small crushing teeth located far back on the palate and lower jaws. Instead, the upper and lower jaws may have carried baleen-like substances for filtering

food particles. Unique among tetrapods, a series of small denticles was arranged on the outside of the anterior margin of the broad snout, possibly used for scraping off plant material from the substrate. *See* DENTITION.

Evolutionary trends. In the relatively brief span of Triassic time, during which differentiation of the placodonts took place, a number of striking evolutionary trends are evident. The most obvious are the acquisition of dermal armor and a progressive reduction in the number of teeth, coupled with an increase in the size (relative to skull size) of the remaining teeth from *Paraplacodus*, to *Placodus*, to *Cyamodus*, to *Placochelys*. Tooth reduction occurs in older individuals of the genus *Cyamodus*. In *Placochelys*, the anterior teeth have been lost and were probably replaced by a horny beak.

Although evolutionary trends are observable through time, generalized and fairly specialized forms, such as the placodontoid *Paraplacodus* and the cyamodontoid *Cyamodus*, occur together in the fossil record. This indicates that the placodonts had a considerable, though unknown, earlier history. Some cyamodontoid placodonts were the last representatives of the group to go extinct, as the climate became hotter and drier in the Upper Triassic. It is possible that their dermal armor provided an osmotic barrier that protected the animals from fluctuations in salinity levels that resulted from climate change. *See* ANIMAL EVOLUTION.

Phylogeny. The most primitive placodont known is *Paraplacodus* from the Middle Triassic of the Southern Alps. Its skull structure confirms the placement of placodonts among the advanced, diapsid reptiles. Placodonts are the sister group of all other sauropterygians, probably nested within crown-group (all taxa descended from a major cladogenesis event) diapsid reptiles (Sauria). Placodonts share with the other sauropterygians a number of key characters that indicate their relationships. These include closure of the palate and its fusion with the base of the braincase, enlarged upper temporal openings in the skull, and details of the anatomy of the pectoral and pelvic girdles. The more precise relationships of placodonts, and of other sauropterygians, within the Sauria remain poorly understood. *See* DIAPSIDA; NOTHOSAURIA; REPTILIA.

Olivier C. Rieppel; Rainer Zangerl

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988; O. Rieppel, *Encyclopedia of Paleoherpetology, Part 12A: Sauropterygia I. Placodontia, Pachypleurosauria, Nothosauroidea, Pistosauroidea*, F. Pfeil, Munich, 2000.

Placozoa

The most primitive metazoan animal phylum. At present, this phylum harbors a single named species, the enigmatic *Trichoplax adhaerens*. *See* ANIMAL KINGDOM; METAZOA.

Description. In 1883, the German zoologist Franz Eilhard Schulze discovered this microscopic marine animal on the glass walls of a seawater aquarium at the University of Graz in Austria. The animal, measuring less than 5 mm (0.2 in.) in diameter and 10–15 micrometers (0.0004–0.0006 in.) in height, looked like an irregular and thin hairy plate sticking to the glass surface (**Fig. 1**). Schulze named this animal *Trichoplax adhaerens* (Greek for "sticky hairy plate").

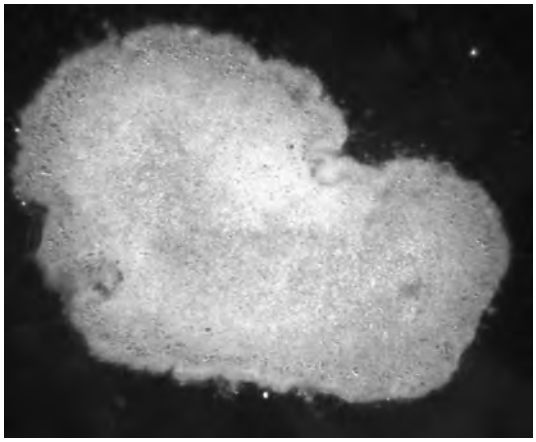

Fig. 1. The placozoan *Trichoplax adhaerens*, the only described species of Placozoa. The flat animal [~14 μm (0.0056 in.) in height] measures ~2–3 mm (0.08–0.12 in.) in diameter and changes its shape constantly. Specimens from different tropical and subtropical waters around the world look highly similar, but further molecular analysis may indicate different species or even genera and families of Placozoa. (*Photo by Wolgang Jakob and Bernd Schierwater*).

In contrast to typical multicellular animals (metazoans), *Trichoplax* does not have a head or tail, nor does the animal possess any organs, nerve or muscle cells, basal lamina, or extracellular matrix. *Trichoplax* lacks an axis and any type of symmetry. It almost looks like a giant ameba (a protozoan) when it crawls over the substrate.

The immediate and defining characteristic between placozoas and protozoans is the number of somatic (non–germ line) cell types. In contrast to Protozoa, which consist of either a single cell or several cells of the same somatic cell type. Placozoa have at least four defined somatic cell types: lower epithelial cells, upper epithelial cells, gland cells, and fiber cells. These cells are arranged in a sandwichlike manner, with the lower epithelial and gland cells at the bottom, the upper epithelial cells at the top, and the fiber cells in between (**Fig. 2**). Cells of the lower epithelium attach the animal to a solid substrate, enable the animal to crawl (with the aid of cilia), and allow feeding. During feeding, the animal lifts up the center region of its body to form an external digestive cavity between the substrate and lower epithelium. Interestingly, the upper epithelium is also capable of feeding. Algae and other food particles are trapped in a slime layer coating the upper epithelium and are subsequently taken up (phagocytized) by the inner fiber cells; this unique mode of feeding is called transepithelial cytophagy. Placozoa presumably harbor endosymbiontic bacteria in the endoplasmic reticulum of the fiber cells. A possible role for these endosymbionts in feeding is not yet understood.

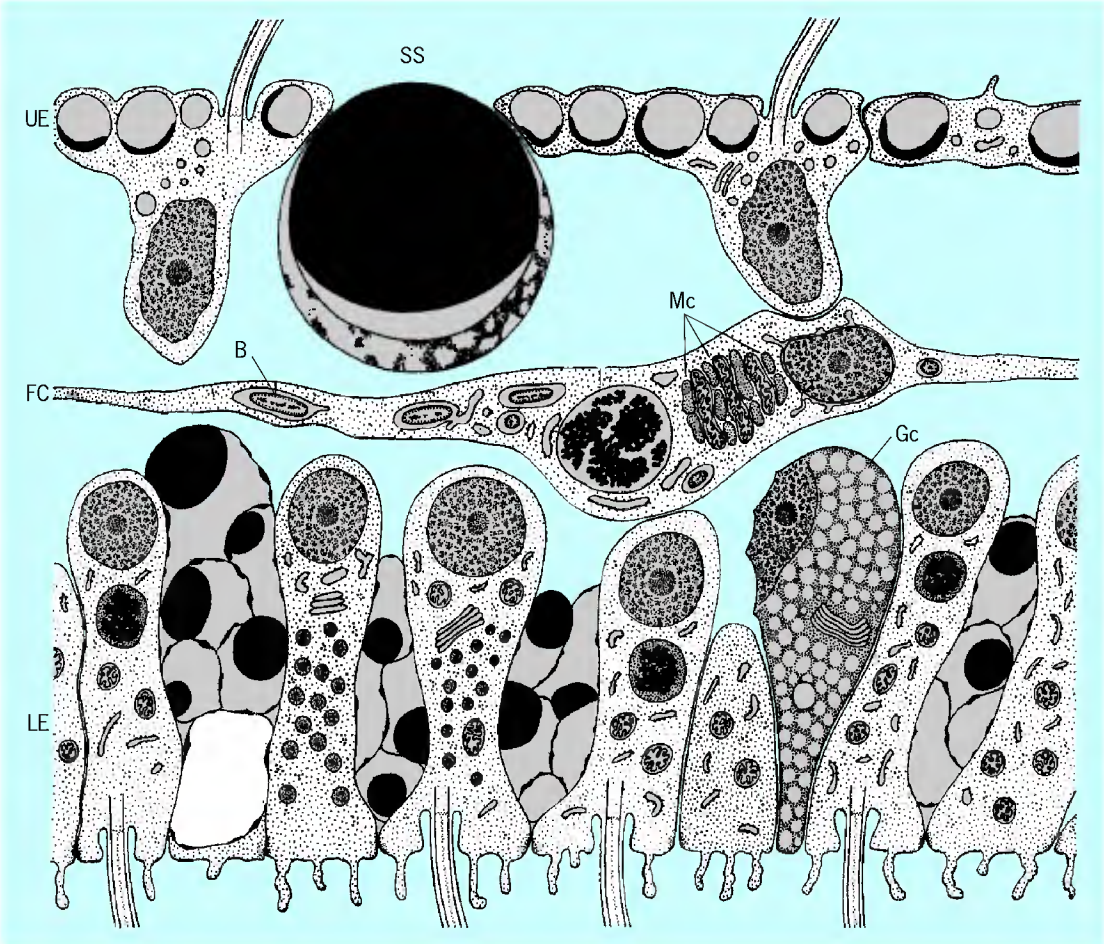

Fig. 2. Cross section of *Trichoplax adhaerens*. The body is organized into three tissue layers: the ciliated upper epithelium, the ciliated lower epithelium, and the fiber cells in between. The interior fiber cells form a contractile, three-dimensional meshwork of syncytial organization (multinucleated mass of cytoplasm that is not separated into individual cells). *Trichoplax* feeds upon different kinds of detritus and living protozoa (including algae). For digestion of the food outside the body (extracorporal digestion), nonciliary gland cells in the lower epithelium release digestive enzymes into the interspace between the lower epithelium and the substrate. The upper epithelium shows no specializations, with the exception of the "shiny spheres" (*Glanzkugeln*). These are lipid droplets, which are usually understood as residues of degenerated epithelial cells. A remarkable and exclusive feature of Placozoa is the lack of both an extracellular matrix and a basal lamina between the inner fiber cells and the enclosing epithelia. UE, upper epithelium; LE, lower epithelium; FC, contractile fiber cell; Gc, gland cell; SS, shiny sphere; Mc, mitochondrial complex; B, bacterium in endoplasmic reticulum. (*Modified from K. G. Grell and A. Ruthmann, 1991*)

In general, very little is known about the biology of Placozoa, and almost all current knowledge derives from laboratory observations. Field data are limited to records of finding *Trichoplax* on hard substrate surfaces from tropical and subtropical marine waters around the world. *See* FEEDING MECHANISMS (INVERTEBRATE); GERM LAYERS; PHAGOCYTOSIS; PROTOZOA.

Reproduction and development. The complete life cycle of Placozoa remains unknown, although three different modes of reproduction are known: (1) fission (normal type of vegetative reproduction in Placozoa), (2) swarmer formation (occasional type of vegetative reproduction), and (3) bisexual reproduction (the standard mode of reproduction in Metazoa). In fission, animals grow to a certain size and divide into two approximately equally sized daughter individuals; then each of these regrows to about the former size of the parent before it divides again. This process can go on ad infinitum and leads very efficiently to a rapid increase in population density. The second mode of vegetative reproduction occurs when environmental conditions become unfavorable (for example, food shortage) and transfer to a new place becomes advantageous. Under these conditions, *Trichoplax* may develop small spherical swarmers which are planktonic (free-floating) and thus are taken by water currents to new habitats. The transition from a swarmer into an adult *Trichoplax*, however, has not been observed. Similarly, the development of sexually produced embryos (third mode of reproduction) into an adult *Trichoplax* also remains unknown. Scientists have seen embryos only up to a stage where epithelial formation takes place, but have not seen any development into a known form, another generation, or a larval form of *Trichoplax*. To what degree Placozoa reproduce sexually in the field is unknown. *See* REPRODUCTION (ANIMAL).

Genetic analysis. Genetic analysis of *Trichoplax* specimens from different coral reefs around the world as well as from the subtropical Mediterranean revealed the presence of cryptic species, that is, species which are basically morphologically undistinguishable. While the phylum Placozoa presently harbors the single named species, *T. adhaerens*, the real biodiversity is estimated to include several dozen genetically and ecologically distinguishable species.

After its original description in 1883, *Trichoplax* attracted particular attention because it possibly mirrored the basic and ancestral state of metazoan organization. The German zoologist Karl Grell further highlighted this view in the 1970s and provided sufficient support for placing *T. adhaerens* in a new phylum, the Placozoa. The new phylum was named after one of several models of early metazoan evolution, the placula hypothesis, which sees a placozoan-like animal as the urmetazoan, that is, the very first metazoan bauplan (body plan). A variety of new molecular data support the traditional view that Placozoa are close to the very root of metazoan origin. Placozoa possess one of the smallest animal genomes, only 20 times the size of the bacterium *Escherichia coli.* Whole genome sequence data from Placozoa are expected to open new avenues for scientists to gain insights into the genetics underlying the origin and development of metazoan animals. *See* BIODIVERSITY.

Bernd Schierwater

Bibliography. S. L. Dellaporta et al., Mitochondrial genome of *Trichoplax adhaerens* supports Placozoa as the basal lower metazoan phylum, *Proc. Nat. Acad. Sci. USA*, 103:8751–8756, 2006; K. G. Grell and A. Ruthmann, Placozoa, in F. W. Harrison and J. A. Westfall (eds.), *Microscopic Anatomy of Invertebrates*, pp. 13–28, Wiley-Liss, New York, 1991; B. Schierwater, My favorite animal, *Trichoplax adhaerens, BioEssays*, 27:1294–1302, 2005.

Plague

An infectious disease of humans and rodents caused by the bacterium *Yersinia pestis*. The sylvatic (wild-animal) form persists today in more than 200 species of rodents throughout the world. At present, contact with wild rodents and their fleas leads to sporadic human disease. The explosive urban epidemics of the Middle Ages, known as the Black Death, resulted when the infection of dense populations of city rats living closely with humans introduced disease from the Near East. The disease then was spread both by rat fleas and by transmission between humans. During these outbreaks, as much as 50% of the European population died; occasionally, entire towns disappeared. *See* INFECTIOUS DISEASE.

In modern times, sylvatic plague has been concentrated in the southwestern United States, southern Russia, India, Indochina, and South Africa. In the United States, ground squirrels, mice, wood rats, prairie dogs, chipmunks, and, rarely, rabbits can be carriers. Rodent infection (and subsequent human disease) occurs most often in the spring and summer; its incidence varies year to year because of animal die-off, but it rarely completely disappears from a region. Feline predators of rodents, such as mountain lions and domestic cats, generally die when infected with *Y. pestis*; however, canines, such as foxes, coyotes, and dogs, often recover and may serve as serologic sentinels to indicate the presence of the disease in wild rodents. Humans can acquire the disease when domestic cats or dogs carrying plague-infected rodents or their fleas return to rural homes.

After infection by *Y. pestis*, fleas develop obstruction of the foregut, causing regurgitation of plague bacilli during the next blood meal. The rat flea, *Xenopsylla cheopsis*, is an especially efficient plague vector, both between rats and from rats to humans. Human (bubonic) plague is transmitted by the bite of an infected flea; after several days, a painful swelling (the bubo) of local lymph nodes occurs. Bacteria can then spread to other organ systems, especially the lung; fever, chills, prostration, and death may occur. Plague pneumonia develops in 10–20% of all bubonic infections.

Pneumonia starts as a secondary lung infection during bubonic/solpticemic disease; it may then be transmitted from human to human by bacterial aerosols that are spread by coughing. It has a rapid course in close quarters and is almost universally fatal if untreated. In some individuals, the skin may develop hemorrhages and necrosis (tissue death), probably the origin of the ancient name, the Black Death. The last primary pneumonic plague outbreak in the United States occurred in 1919, when 13 cases resulting in 12 deaths developed before the disease was recognized and halted by isolation of cases.

Since about 1960, the incidence of sporadic human plague in the western United States has been rising. Although urban rat-related human outbreaks are now rare, they represent a continual threat, which has generally been contained by public health rat-control measures. Spread from wild rodents into urban rats, however, occurred as recently as 1983 in Los Angeles.

Bubonic plaque is suspected when the characteristic painful, swollen glands develop in the groin, armpit, or neck of an individual who has possibly been exposed to wild-animal fleas in an area where the disease is endemic. A culture of blood or of material aspirated from a bubo will grow the organism in 2–5 days. Immediate identification is possible, however, by microscopic evaluation of bubo aspirate stained with fluorescent-tagged antibody. Serologic diagnosis, shown by a rise in antiplague antibody in the blood, takes 2–3 weeks. Antibiotics should be given if plague is suspected or confirmed. Such treatment is very effective if started early. The current overall death rate, approximately 15%, is reduced to less than 5% among patients treated at the onset of symptoms. *See* IMMUNOFLUORESCENCE; MEDICAL BACTERIOLOGY.

Darwin L. Palmer

Bibliography. T. Butler, *Plague and Other Yersina Infections*, 1983.

Plains

The relatively smooth sections of the continental surfaces, occupied largely by gentle rather than steep slopes and exhibiting only small local differences in elevation. Because of their smoothness, plains lands, if other conditions are favorable, are especially amenable to many human activities. Thus it is not surprising that the majority of the world's principal agricultural regions, close-meshed transportation networks, and concentrations of population are found on plains. Large parts of the Earth's plains, however, are hindered for human use by dryness, shortness of frost-free season, infertile soils, or poor drainage. Because of the absence of major differences in elevation or exposure or of obstacles to the free movement of air masses, extensive plains usually exhibit broad uniformity or gradual transition of climatic characteristics.

Distribution and varieties. Somewhat more than one-third of the Earth's land area is occupied by plains. With the exception of ice-sheathed Antarctica, each continent contains at least one major expanse of smooth land in addition to numerous smaller areas. The largest plains of North America, South America, and Eurasia lie in the continental interiors, with broad extensions reaching to the Atlantic (and Arctic) Coast. The most extensive plains of Africa occupy much of the Sahara and reach south into the Congo and Kalahari basins. Much of Australia is smooth, with only the eastern margin lacking extensive plains. *See* TERRAIN AREAS.

Surfaces that approach true flatness, while not rare, constitute a minor portion of the world's plains. Most commonly they occur along low-lying coastal margins, the lower sections of major river systems, or the floors of inland basins. Nearly all are the products of extensive deposition by streams or in lakes or shallow seas. The majority of plains, however, are distinctly irregular in surface form, as a result of valley-cutting by streams or of irregular erosion and deposition by continental glaciers.

Plains are sometimes designated by the situations in which they occur. In common speech a coastal plain is any strip of smooth land adjacent to the shoreline, though in geology the term is usually restricted to such a plain that was formerly a part of the shallow sea bottom. An example that fits both definitions is the South Atlantic and Gulf margin of the United States. Intermontane plains lie between mountain ranges, and basin plains are surrounded by higher and rougher land. Upland plains (sometimes loosely termed plateaus) lie at high elevations, or at least well above neighboring surfaces, while lowland plains are those lying near sea level, or distinctly below adjacent lands.

Types of plains are also sometimes designated according to the processes that have produced their distinctive surface features. These differences are discussed below.

Origin. The existence of plains terrain generally indicates for that area a dominance of the erosional and depositional processes over the forces that deform the crust itself. Most of the truly extensive plains, such as those of interior North America or that of northwestern Eurasia, represent areas which have experienced nothing more severe than slow, broad warping of the crust over tens or hundreds of millions of years. Throughout that time the gradational processes have been able to maintain a relatively subdued surface. Certain other areas, including the upland plains of central and south-central Africa and southern Brazil, have suffered moderate general uplift in late geologic time and have not yet been subjected to deep valley cutting.

Many plains of lesser extent, however, have been formed in areas where crustal deformation has been intense. Most of these represent depressed sections of the crust which have been partially filled by smooth-surfaced deposits of debris carried in by streams from the surrounding mountains. Examples are the Central Valley of California, the Po Plain of northern Italy, the plain of Hungary, the Mesopotamian plain, the Tarim Basin of central Asia, and the Indo-Gangetic plain of northern India and Pakistan.

Surface Characteristics

The detailed surface features of plains result mainly from local erosional and depositional activity in relatively recent geologic time. Each of the major gradational agents—running water, glacial ice, and the wind—produces its own characteristic set of features, and any given section of plains terrain is characterized predominantly by features typical of one particular agent.

Features associated with stream erosion. Plains sculptured largely by stream erosion, which are far more widespread than any other class, are normally irregular rather than flat. Integrated valleys and the divides between them are the hallmarks of stream sculpture, and the differences among stream-eroded plains are generally expressible in terms of the size, shape, spacing, and pattern of these features.

The shallow depth of valleys on plains usually indicates that the surface has not been uplifted, in late geologic times, far above the level to which the local streams can erode. Wide valleys commonly indicate weak valley-wall materials, active erosion on the valley sides by surface runoff and soil creep and, in many instances, a long period of development. Resistant materials and permeable upland surfaces, on the other hand, favor narrower valleys and steeper side slopes.

Differences in valley spacing or in the degree to which the plain has been dissected by the development of valley systems are especially striking. Ideally, tributary growth is considered to proceed progressively headward from the major streams into the intervening uplands. For this reason, plains on which tributary valleys are few, major valleys are widely spaced, and broad areas of uncut upland remain (**Fig. 1**) are often termed youthful. If tributary valleys have extended themselves into all parts of the surface so that the area is almost wholly occupied by valley-side slopes, the surface is called mature. If

valleys have so widened and coalesced that only a few narrow, low divide remnants are left, the surface is said to have reached old age. However, not all stream-eroded surfaces follow the ideal sequences, and many factors other than time affect their development.

Some plains, for example, remain persistently youthful because unusual permeability of the surface material, gentle or inadequate rainfall, excessive flatness, dense vegetation cover, or resistance of the rocks either fail to provide the surface runoff that could carve tributary valleys or else inhibit the erosive capacity of those smaller streams that do form. An example which combines several of these causal factors is provided by the High Plains, which stretch from southwestern Nebraska into the Panhandle of Texas. By way of contrast, some other plains appear to have reached maturity very quickly as a result of the generation of copious surface runoff by heavy rains and an initially rolling surface of low permeability, combined with readily erodible materials. The Dissected Till Plains of northern Missouri and southern Iowa are an example.

Although valleys of great breadth are not uncommon, extensive plains of ideal old age are rare. Valley widening is at best a slow process, and the reduction of low, gentle-sloped, well-vegetated divides requires immense lengths of time without interruption by uplift or environmental changes that initiate a renewal of valley deepening. Some modern scholars suggest that certain wide-valleyed, long-sloped, or gently undulating plains with an unbroken vegetation cover may in fact represent plains on which erosional development has practically ceased. *See* STREAM TRANSPORT AND DEPOSITION.

The pattern of valleys on plains depends chiefly upon the initial slope of surface and the pattern of outcrop of rock materials of contrasting resistance. In the absence of strong contrasts the pattern is usually branching and treelike as in Fig. 1. Where there is great differential resistance to erosion, the unusually resistant rocks form drainage divides or uplands, whereas weak rock belts are soon excavated into broad valleys or lowlands. Where erosional plains bevel across gently warped rock strata of varying resistance, the belts of outcrop of the more resistant strata form strips of higher, rougher country, with an abrupt escarpment on one margin and a more gradual dipslope in the direction toward which the strata are inclined (**Fig. 2**). These features, called cuestas, are common in the Middle West and Gulf Coastal Plain and in western Europe. The wolds and downs of England and the côtes of northeastern France are cuesta ridges.

Most plains that develop in dry climates are characterized predominantly by stream-produced landforms, in spite of infrequent rain. The development of valley systems and erosional features follows the same general rules as it does in humid regions. However, some differences in relative rates and relative significance of weathering and erosion produce distinctive landscape characteristics in arid lands. First, rock decomposition is very slow, so that the surface accumulation of weathered material is normally thin and coarse textured. Second, the sparse vegetation affords to the naked surface little protection against the battering and washing of the occasional torrential rains. As a result, the upper slopes become strongly gullied and often partially stripped of their covering material, leaving much bedrock exposed. Because of the short duration and local nature of the rains and hence the intermittent character of stream flow, however, most of the debris load is dropped in the neighboring basins and valley floors, "drowning" broad areas beneath plains of silt, sand, or gravel.

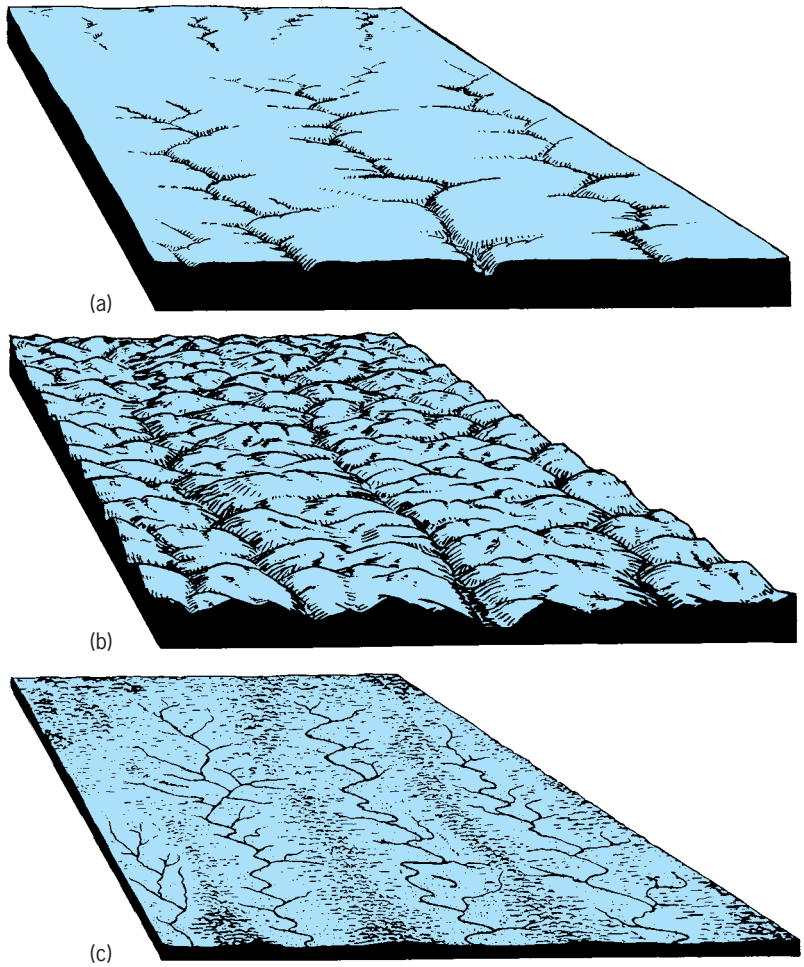

Fig. 1. Ideal stages in the progressive development of a stream-eroded landscape. (*a*) Youth. (*b*) Maturity. (*c*) Old age. (*After G. T. Trewartha, A. H. Robinson, and E. H. Hammond, Elements of Geography, 5th ed., McGraw-Hill, 1967*)

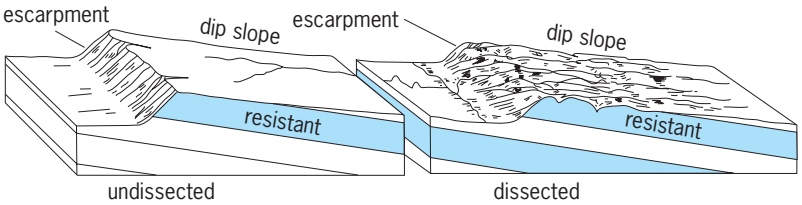

Fig. 2. Surface form and structure of cuestas. Dissected form at right is more typical, especially in humid regions. (*After G. T. Trewartha, A. H. Robinson, and E. H. Hammond, Elements of Geography, 5th ed., McGraw-Hill, 1967*)

Hence denuded and gullied upper slopes and broad depositional flats in the lowlands are characteristic features of desert plains.

Features produced by solution. Features resulting from underground solution characterize several rather extensive areas of plains. The principal features of this class are depressions, or sinks, produced by collapse of caverns underneath. Solution is also an active process in erosion by surface streams, but its effects upon valley form have been little studied.

Significant ground-water solution is largely confined to areas underlain by thick limestones. As subsurface cavities are progressively enlarged by solution, more and more drainage is diverted to subterranean channels. Surface streams become fewer, often disappearing into the ground after a short surface run. Eventually solution cavities near the surface collapse, forming surficial depressions of various sizes. Some are shallow and inconspicuous; others are great steep-walled pits or elongated enclosed valleys. Some of the small depressions contain lakes, because their outlets are plugged with clay. The most extensive areas of solution-featured plains in the United States are in central and northern Florida and in the Panhandle of Texas. In both of these areas shallow sink holes, some of them lake-filled, are numerous. Some of the areas of most active solution work have developed surfaces far too rough to be called plains. This is true of the Mammoth cave area of west-central Kentucky and especially so for the mountainous area of great sinks and solution valleys in the Dalmation Karst of Yugoslavia. *See* KARST TOPOGRAPHY.

Features associated with stream deposition. As a group, alluvial plains (so called from the term alluvium, which refers to any stream-deposited material) are among the smoothest and flattest land surfaces known. Stream-deposited plains fall into three classes: (1) floodplains, which are laid down along the floors of valleys; (2) deltas, formed by deposition at the stream mouths; and (3) alluvial fans, deposited at the foot of mountains or hills.

Floodplains. The flat bottomlands so common to valley floors develop wherever a given segment of a stream system is fed more sediment by its tributaries or upstream reaches than it is able to carry. Most of the excess sediment is deposited in the stream bed, usually in the form of bars. In flood time, some may be strewn across the whole width of the valley floor. Because of continued deposition and choking of the channel, repeated flooding, and the ease with which banks composed of loose alluvium collapse, streams continually shift their channels on floodplains. On silty plains the channel is usually highly sinuous or meandering, while on a coarse sandy or gravelly floodplain the channel is braided, that is, broad, shallow, and intricately subdivided by innumerable sandbars (**Fig. 3**). In either case, many loops or strands of abandoned channels, now mostly dry and partially filled with sediment, scar the floodplain surface. On broad silty floodplains slightly elevated strips, called natural levees, are formed by active deposition immediately adjacent to the channel when the velocity of flow is abruptly checked as the water leaves the swift current of the channel to spread thinly over the plain during floods.

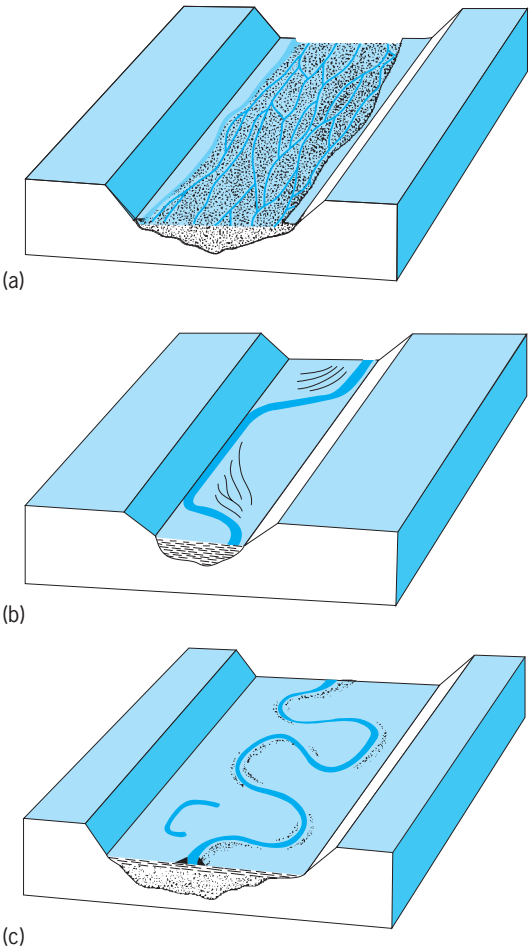

Fig. 3. Types of floodplains. (*a*) With braided channel. (*b*) With sinuous channel but too narrow for free meandering. (*c*) With freely meandering channel. (*After G. T. Trewartha, A. H. Robinson, and E. H. Hammond, Elements of Geography, 5th ed., McGraw-Hill, 1967*)

On major floodplains the ground-water table is everywhere close to the surface, and swampy land is common in the abandoned channels and shallow swales. For this reason, any slightly higher features, such as natural levees, are especially sought after for cultivation, town sites, and transportation routes. Though harassed by a high water table, recurrent floods, and shifting channels, floodplains, especially silty ones, are often prized agricultural lands because of the level, easily tilled surface. In some instances the alluvium is also more fertile than the soils of the surrounding uplands.

Here and there along the sides of valleys, benches and terraces of alluvium stand above the level of the present floodplain. These are remnants of earlier floodplains that have been largely removed because of a renewal of downcutting by the streams. These alluvial terraces may be valuable agricultural lands and serve well for town sites and transportation routes because they stand above flood levels. *See* FLOODPLAIN.

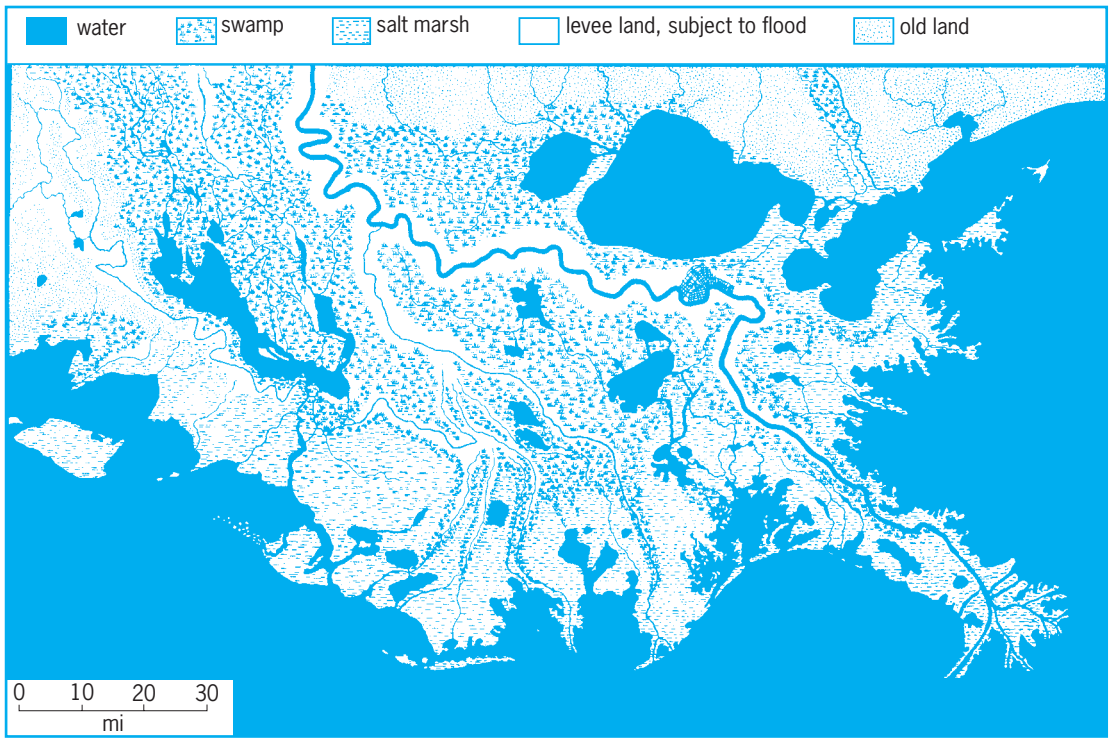

Fig. 4. Salt marsh, swamp, and natural levee lands in the Mississippi River delta. 1 mi = 1. 6 km. (*After G. T. Trewartha, A. H. Robinson, and E. H. Hammond, Elements of Geography, 5th ed., McGraw-Hill, 1967*)

Deltas. The surface features of deltas are essentially the same as those just described. As a rule, however, the delta surface is even less well drained than the floodplain surface and at its outer margin may merge with the sea through a broad belt of marshy land. As the delta grows, continual obstruction of the stream mouth produces repeated diversion and bifurcation of the channel, so that the stream commonly discharges through a spreading network of distributaries (**Fig. 4**).

Many deltas, like those of the Mississippi, the Nile, the Danube, or the Volga, are immense fan-shaped features that have produced broad coastal bulges by their growth. Others, like those of the Colorado, the Po, or the Tigris-Euphrates, though no less extensive, are less apparent on the map because they have been built in large coastal embayments. Some great rivers have no true deltas because their sediment load has been dropped in some interior settling basin. For example, the Great Lakes remove most of the sediment from the St. Lawrence system, and the Congo deposits most of its load in its broad upland basin.

Like floodplains, deltas are sometimes highly valued as agricultural lands, though they have even worse problems of poor drainage and frequent flooding. The Nile delta and the huge, silty delta plain of the Hwang Ho, in north China, are famous centers of cultivation. Many deltas, of which that of the Mississippi is a good example, are too swampy to permit tillage except along the natural levees. The Netherlands, occupying the combined deltas of the Rhine and Maas, stands as an example of what can be done toward reclamation when population pressure is great. *See* DELTA.

Alluvial fans. If a stream emerges upon a gentle plain from a steeply plunging mountain canyon, its velocity is abruptly checked, and it deposits most of its load at the mouth of the canyon. Because of the tendency toward repeated choking and diversion of the channel, the deposit assumes the form of a broad, spreading alluvial fan, essentially similar to a delta, even to the diverging distributary channels (**Fig. 5**). Usually, however, the gradients developed are steeper than those on a delta because sediments are coarser.

Small individual alluvial fans are common features in mountainous country, especially where the climate is dry except for occasional torrential showers. Particularly significant, however, are the rows of alluvial fans that have coalesced to form extensive, gently sloping piedmont alluvial plains at the bases of long, precipitous mountain fronts. The city of Los Angeles is built on such a plain. Still larger ones occupy much of the southern part of the Central Valley of California and stretch eastward from the Andes in northwestern Argentina, Paraguay, and eastern Bolivia.

Because of their smoothness and ease of tillage, alluvial fans, like other alluvial surfaces, are often especially amenable to cultivation. They are particularly significant in drier areas, partly because of the ease with which water may be conducted by gravity from the mountain canyon to any part of the fan, and partly because the thick, porous alluvium itself serves as a reservoir in which ground water is naturally stored.

Lake plains and coastal plains. Closely allied to stream-deposited plains are surfaces that represent

Fig. 5. Alluvial fans in the Mojave Desert, southeastern California. (*J. L. Balsley, U.S. Geological Survey*)

former lake bottoms or recently exposed sections of the former shallow sea floor. These nearly featureless plains have been formed by the deposition of sediments carried into the body of water by streams or wave erosion and further distributed and smoothed by the action of waves and currents. Former beach lines and other shore features often provide the only significant breaks in the monotonous flatness. In some places, shallow valleys have been cut by streams since the surface became exposed. The lower parts of lake plains and the outer margins of coastal plains are often poorly drained.

The flat surfaces upon which Detroit, Toledo, Chicago, and Winnipeg stand are all lake plains, as are the famed Bonneville Salt Flats of western Utah. The south Atlantic and Gulf margins of the United States and much of the Arctic fringe of Alaska and Siberia are examples of newly emerged coastal plains. Some of these plains represent valuable agricultural land; others are excessively swampy or sandy. *See* COASTAL LANDFORMS; COASTAL PLAIN.

Features due to continental glaciation. On several occasions during the last 1–2 $\times$ 10^6 years (Pleistocene Epoch), immense ice caps, comparable to the one which now covers Antarctica, developed in Canada and Scandinavia and spread over most of northern North America and northwestern Eurasia (**Fig. 6**). The last such ice sheet (called in North America the Wisconsin) reached its maximum extent about 18,000 years ago and did not disappear until 5000–6000 years ago. *See* GLACIATED TERRAIN.

These great glaciers significantly modified the land surfaces over which they moved. Except near their melting margins, they were able to remove and transport not only the soil but also much weathered, fractured, or weak bedrock. This material was subsequently deposited beneath the ice or at its edge, wherever melting released it. The marks of glacial sculpture are unsystematically distributed swells and depressions; numerous lakes, swamps, and aimless streams (**Fig. 7**); and an irregular mantle of debris (glacial drift), some of which is clearly not derived from the local bedrock upon which it now lies. Such surfaces often contrast strongly with the systematic valley and divide patterns characteristic of stream erosion.

Most of the area that was covered by ice during the Pleistocene is now more or less thickly covered with drift, the thickest cover generally occurring in regions of weak rocks. Characteristic glacial landforms, however, are largely confined to those areas that were occupied by the ice during middle and late Wisconsin times. Earlier glacial features have been modified beyond recognition by surface erosion and soil creep. Hence several large areas, including parts of the midwestern United States and southern Russia, though drift-mantled, display the surface features characteristic of stream erosion.

Drift-covered plains offer a wide range of potentialities for human use, the critical characteristics being chiefly the adequacy of drainage and the texture of the soils. Excessive roughness is rarely a problem.

Till. Much of the drift represents mixed rock and soil deposited directly by melting beneath the ice sheet or at its edge. This material, called till, is as a rule most thickly deposited in the valleys and

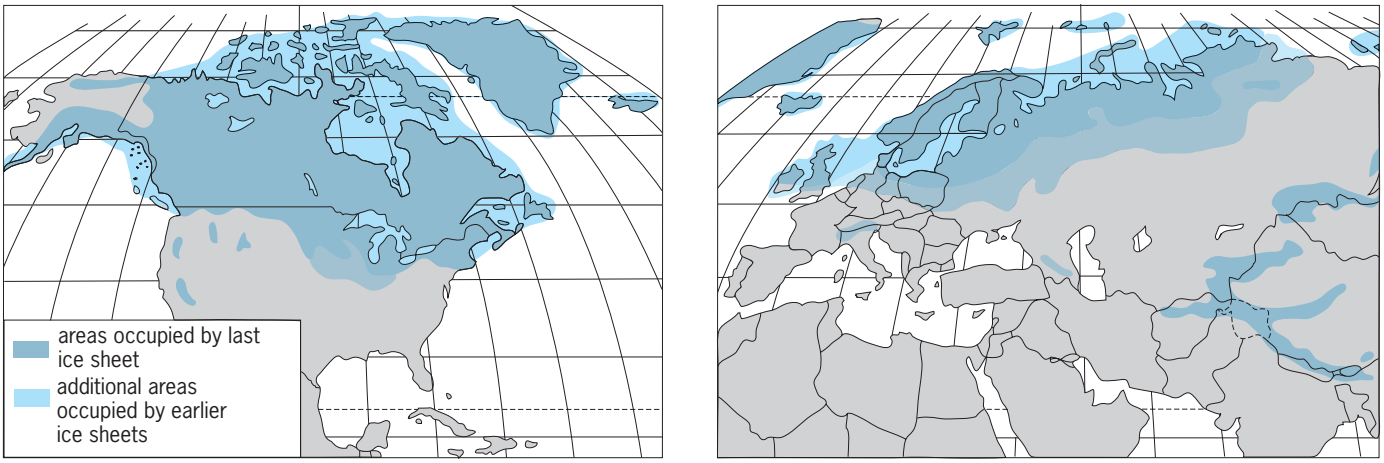

Fig. 6. Extent of former continental glaciers in North America and Eurasia. (*After G. T. Trewartha, A. H. Robinson, and E. H. Hammond, Elements of Geography, 5th ed., McGraw-Hill, 1967*)

thinly over the ridge tops, thus reducing terrain irregularity. The surface of the till sheet itself is usually gently rolling, with poor and unsystematically patterned drainage. Hummocky, often stony ridges, called marginal moraines, mark lines along which the fluctuating edge of the ice remained stationary for a considerable time. In several localities, notably in eastern Wisconsin and western New York, are swarms of smooth, low drift hills called drumlins, all elongated in the direction of ice movement.

The surfaces of stony till plains are usually more irregular than those on silt or clay till. Northeastern Illinois has a remarkably smooth surface developed on silt and clay till, much of it eroded from the Lake Michigan basin. Eastern Wisconsin, northern Michigan, western New York, and southern New England, on the other hand, have more rolling surfaces underlain by till having a high content of stone and sand. In a few areas, especially in southern New England and in the marginal moraines elsewhere, the till is so stony as to impede cultivation.

Outwash. Some of the debris transported by the ice is carried out beyond the glacial margin by streams of meltwater. This material, called outwash, may be deposited as a floodplain (here called a valley train) along a preexisting valley bottom, or it may be broadcast over a preexisting plain. In either case the surface will usually be smooth, with features typical of

Fig. 7. Diagram of the drainage pattern of a glaciated plain in eastern Wisconsin. 1 mi = 1.6 km. (*After G. T. Trewartha, A. H. Robinson, and E. H. Hammond, Elements of Geography, 5th ed., McGraw-Hill, 1967*)

alluvial plains. Unlike the heterogeneous, unsorted, and unstratified till, outwash material is usually layered and sorted in size. The fine material of silt and clay size is carried out downstream, leaving the coarser sands and gravels to form the outwash deposits. Most of the gravel and sand pits that abound in glaciated areas are developed in outwash plains.

Where outwash was laid down over already deposited till surfaces after the ice had melted back from its maximum extent, the surface of the outwash plain is sometimes pitted and lake-strewn, the depressions having formed as a result of the melting of relict ice masses that were buried by outwash deposition.

Patches and ribbons of outwash are common in glaciated areas, and in a few places, into which unusual quantities of meltwater were funneled, there are very extensive sandy plains. Noteworthy in this respect are the southern Michigan and northern Indiana area and Europe immediately south of the Baltic.

Lacustrine plains. Also present in and around the glaciated areas are numerous plains marking the beds of former lakes that resulted from the blocking of rivers by the glacial ice itself. During the melting of the last ice sheet, while the St. Lawrence and other northward-flowing rivers were still ice-dammed, the Great Lakes basins were much fuller than now and overflowed to the southward. When lake levels were eventually lowered, lacustrine plains were exposed, notably about Chicago, at the western end of Lake Erie, and about Saginaw Bay in Michigan. One of the most featureless plains of North America occupies the former bed of an immense lake (Lake Agassiz) that was present in late glacial time in southern Manitoba, northwestern Minnesota, and eastern North Dakota (**Fig. 8**).

Thin-drift areas. Considerable areas in central and eastern Canada and northern Fennoscandia have only a thin and discontinuous drift cover. These hard-rock regions yielded little debris and, especially in the rougher sections, the small drift supply was thinly strewn in the valleys and depressions. Lakes and swamps occupy drift-blocked valleys and shallow erosional depressions, rapids and waterfalls are common, and broad expanses of scoured bedrock are exposed. Such areas offer little opportunity for agricultural use and usually support only an open, patchy, and often stunted forest growth.

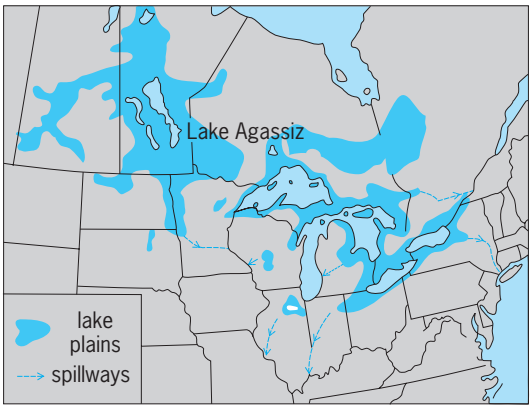

Fig. 8. Plains of former glacial lakes in North America. (*After G. T. Trewartha, A. H. Robinson, and E. H. Hammond, Elements of Geography, 5th ed., McGraw-Hill, 1967*)

Features reflecting wind action. Features formed by the wind are less widespread and, as a rule, less obtrusive than the forms produced by streams and glaciers. Since the wind can attack only where the surface is almost free of vegetation, its work is strongly evident only in arid regions and in the vicinity of beaches and occasionally exposed river beds in more humid lands. Today plowed fields are also important prey for wind erosion.

The most striking and significant wind-produced features are sand dunes. Where sand is exposed to strong winds, it is moved about for short distances and accumulates in heaps in the general vicinity of its place of origin. Sand dunes assume many forms, from irregular mounds and elongated ridges to crescent-shaped hills and various arrangements of sand waves, depending apparently upon the supply of sand, the nature of the underlying surface, and the strength and directional persistence of the wind.

Nearly all the truly extensive areas of sand dunes are found in the Eastern Hemisphere, especially in the Sahara and in Arabia, central Asia, and the interior of Australia. Most of the dunes have been whipped up from alluvium that has been deposited in desert basins and lowlands.

In several regions of the world, notably in north-central Nebraska and in the central and western Sudan south of the Sahara, extensive areas of dunes have become covered by vegetation and fixed in position since they were formed, suggesting the possibility of climatic change. *See* DUNE.

Other wind-formed features are polished and etched outcrops of bedrock that show the effects of natural sand-blasting; gravel "pavements" resulting from the winnowing out of finer material from mixed alluvium; and shallow blowouts, which are depressions formed by local wind erosion.

The finer silty material moved by the wind is spread as a mantle over broad expanses of country downwind from the place of origin. Though usually thin, this mantle in places reaches a thickness of a few tens of feet, thus somewhat modifying the form of the surface. The extensive deposits of unstratified, buff-colored, lime-rich silty material known as loess are believed to have originated from such wind-laid deposits. Loess is abundant in the central United States, eastern Europe and southern Russia, and in interior northern China. It yields productive soils but is subject to gullying erosion and has the facility of maintaining remarkably steep slopes, so that erosional terrain developed on deep loess is often unusually rough and angular. *See* LOESS.

Interrupted Plains

Plains interrupted by some features of considerable relief occur widely and merit independent treatment. They may be divided into two contrasting groups: (1) tablelands, which are upland plains deeply cut at intervals by steep-sided valleys or broken by escarpments, and (2) plains with hills or

Fig. 9. View of an ideal tableland: Canyon de Chelly National Monument, northeastern Arizona. (*Spence Air Photos*)

mountains, in which the excessive relief is afforded by steep-sided eminences that rise above the plain. Both of these types of surface, with their combination of plain and rough land, suggest histories of development that combined extensive gradation and strong tectonic activity.

Tablelands. These are essentially youthful plains that have been cut to unusual depth by valleys (**Fig. 9**). The plain first had to be brought to a level hundreds or even a few thousands of feet above the level to which streams can erode. In most cases, this elevation was accomplished by the broad uplifting of an erosional or alluvial plain, but in a few instances the plain was built to high level by the deposition of many thick sheets of lava.

It is also necessary that while a few major streams have cut deep canyons, large areas of the upland plain remained largely uncut by tributary development, a condition requiring special circumstances. Such inhibited tributary growth is usually the result of either (1) slight local surface runoff of water because of aridity, extreme flatness of the upland, or highly permeable material; or (2) the presence, at the upland level, of very resistant rock strata that have permitted only the most powerful streams to cut through. Hence tablelands are predominantly a dry-land terrain type. Those that do occur in rainier climates usually indicate the presence of an exceptionally resistant and permeable cap-rock layer.

The cliffs and escarpments that are common features of tableland regions are sometimes fault scarps, produced by the breaking and vertical dislocation of the crust during uplift. More often, however, they are simply steep valley sides that have been worn back long distances from their original positions under the attack of weathering and erosion. Sometimes a once extensive tableland is so encroached upon by the retreat of its bordering escarpments that nothing is left but a small, flat-topped mesa or butte.

Tablelands are the least widespread of the major

terrain types, presumably because of the limited circumstances under which they can develop. The most extensive examples occur in the American continents. Especially noteworthy are the Colorado Plateaus, largely in northern Arizona and southern Utah, which represent a complex erosional plain that has been greatly uplifted and then deeply carved by the Colorado River and its major tributaries. Preservation of large sections of the upland plain has been favored by dryness and by the presence of nearly horizontal resistant rock strata. The few major streams that have cut deep canyons are all fed from moister mountainous areas round about.

The Columbia Plateau in eastern Washington is a porous lava plain, cut by the Columbia River and a few tributaries. The northern Great Plains, lying east of the Rocky Mountains from Nebraska northward into Alberta, are an old erosional plain on weak rocks, now crossed by valleys of moderate depth that have been cut by streams issuing from the Rockies. The Cumberland Plateau in Tennessee is maintained by a resident sandstone caprock. The Patagonian Plateau of southern South America is a somewhat similar surface, locally reinforced by extensive lava flows. Parts of the upland of interior Brazil, though in a moist environment, retain a tableland form because of the resistance of thick sandstone beds and, locally, lavas at the upland level.

Though there are significant exceptions, tablelands as a group suffer, in their economic development, from the difficulty of passage through their narrow gorges and across their many escarpments, and in most cases from the dryness of their upland surfaces.

Plains with hills or mountains. These are much more widespread than tablelands and are extensively represented in each continent. Surfaces included under the general heading vary from plains studded with scattered small hills and hill groups to mountain-and-plain country in which high, rugged ranges occupy almost as much space as the plains between them (Fig. 5).

Surfaces of the general type can be produced by two quite different lines of development. (1) They may be mountain, hill, or tablelands that have been brought to the erosional stage of early old age, in which case the isolated hills represent the only remnants of a once extensive highland. (2) They may represent areas in which separated hills or mountains have been constructed by volcanic eruption or by folding or buckling of the crust, the land between them having remained smooth from the outset, or

Fig. 10. Death Valley, California, a basin of interior drainage in the Basin-and-Range province. (*Courtesy of John S. Shelton*)

having been smoothed by processes of erosion and deposition since the mountains were formed.

Examples of the first course of development are limited in the United States to small areas in southern New England and in the Appalachian Piedmont just east of the Blue Ridge. Many patches occur in the glacially scoured sections of northern and eastern Canada and similarly in northern Sweden and Finland. Extensive areas are found in the southern regions of Venezuela and Guiana, in the upland of eastern Brazil, and especially on the uplands of western, central, and southern Africa.

In such areas the plains are typical late-stage erosional-depositional surfaces, usually gently undulating. The remnant hills (monadnocks) rise abruptly from the plain, like islands from the sea. Although monadnocks commonly represent outcrops of unusually resistant rocks that have withstood erosion most effectively, many owe their existence solely to their position at the headwaters of the major stream systems, where they are the last portions of the highland to be reduced.

Surfaces on which spaced mountains have been constructed by crustal deformation or volcanic activity occur extensively in the great cordilleran belts of the continents, and rarely outside of those belts. The largest of all such regions is the Basin-and-Range section of western North America, which extends without interruption from southeastern Oregon through the southwestern United States and northern Mexico to Mexico City. The majority of the rugged mountain ranges of this section are believed to represent blocks of the crust that have been uplifted or uptilted and then strongly eroded. The plains between them are combination erosional-depositional surfaces, some of which have clearly expanded at the expense of the adjacent mountains. As the mountains have been reduced by erosion, there have evolved at their bases smooth, gently sloping plains that are in part erosional pediments, closely akin to old-age plains, and in part piedmont alluvial plains. Many basins have interior drainage; in these the floor is likely to be especially thickly alluviated and to contain a shallow saline lake or alkali deposit in its lowest part (**Fig. 10**).

Other areas of somewhat similar terrain are found in the central Andes, in Turkey and the Middle East, and in Tibet and central Asia. The Tibetan and Andean sections are noteworthy for the extreme elevation (12,000–15,000 ft or 4–5 km) of their basin floors. In general the ranges and basins in the Asiatic areas are developed on a grander scale than those in North America.

It is a curious circumstance that practically all these regions are dry and exhibit the intensely eroded mountain slopes and alluvially drowned basin floors characteristic of that climatic realm. For this reason they are only locally useful to inhabitants, in spite of the large amount of smooth land that they afford. There are, however, many important oases, usually near the bases of the mountains or along the courses of the few streams that have wandered in from moister adjacent regions. Edwin H. Hammond

Bibliography. D. Briggs and P. Smithson, *Fundamentals of Physical Geography*, 1993; D. J. Easterbrook, *Surface Processes and Landforms*, 1992; A. Goudie, *The Changing Earth: Geomorphological Processes and Time*, 1995; A. N. Strahler and A. H. Strahler, *Elements of Physical Geography*, 4th ed., 1990.

Planck's constant

A quantity, designated by the letter h, which occurs repeatedly in those fundamental formulas of physics that describe the microworld, the world of atoms and nuclei and elementary particles. Historically, h occurred first in 1900 in work of Max Planck on the theory of blackbody radiation, but its physical significance was at first not clear. Some physicists would claim that it is still not clear. *See* HEAT RADIATION.

In 1905 Albert Einstein sought to clarify Planck's work, as well as certain paradoxes that had arisen concerning the nature of light, by suggesting that light has a dual nature—that in some way, which he made no attempt to explain, it is both wave and particle. The particle, called a light quantum or photon, is massless, travels always at the speed of light c, and has an energy E which is related to the light wave's vibrational frequency ν by Eq. (1), where the

$$E = h\nu \tag{1}$$

$$6.62607 \times 10^{-34} \text{ joule-second}$$

numerical value of h is also given. *See* FUNDAMENTAL CONSTANTS; PHOTON.

If a photon has an energy, it also has a momentum. Maxwell's theory of the electromagnetic field states that if a portion of the field traveling in a given direction has an energy E, its momentum p is given by E/c. Putting this into Eq. (1) one encounters the ratio ν/c, which can also be written in terms of the wavelength λ of the corresponding light wave through the universal relation $\lambda\nu = c$. This leads to Eq. (2).

$$p = \frac{E}{c} = \frac{h\nu}{c} = \frac{h}{\lambda} \tag{2}$$

See MAXWELL'S EQUATIONS.

By the early 1920s, Eqs. (1) and (2) had been tested and verified in several experimental situations, notably those involving the photoelectric effect and the Compton effect. In 1923 Louis de Broglie suggested that these relations govern the behavior of forms of matter other than photons. He proposed experiments in which electrons, previously thought of as particles, would appear as waves with wavelengths given by Eq. (2), and in a few years this too had been verified. *See* COMPTON EFFECT; ELECTRON DIFFRACTION; PHOTOEMISSION.

Other occurrences of h in fundamental physics stem from these early considerations. In Niels Bohr's provisional atomic theory of 1913, the angular momentum of an orbiting electron is an integral multiple of $h/2\pi$, a quantity now denoted by the symbol $\hbar$, and de Broglie was later able to relate this proposal

to Eq. (2). Out of these considerations grew the modern theory of quantum mechanics, in which h plays a fundamental role. It occurs also in Heisenberg's principle of uncertainty or indeterminacy, which places limits on the accuracy with which it is possible to know and describe the microworld. *See* ATOMIC STRUCTURE AND SPECTRA; QUANTUM MECHANICS; UNCERTAINTY PRINCIPLE. David Park

Planck's radiation law

A law of physics which gives the spectral energy distribution of the heat radiation emitted from a so-called blackbody at any temperature. Discovered by Max Planck early in the twentieth century, this law laid the foundation for the advent of the quantum theory because it was the first physical law to postulate that electromagnetic energy exists in discrete bundles, or quanta. *See* HEAT RADIATION; QUANTUM MECHANICS. Heinz G. Sell; Peter J. Walsh

Plane curve

The locus of points in the euclidean plane that satisfy some geometric or algebraic definition. Not all sets of points deserve to be called a curve, but the distinction is somewhat arbitrary. For most of this article, a curve is considered to be the locus of a set of points that satisfy an algebraic or transcendental equation in two variables.

In general, a curve can be defined by geometric properties; an equation in cartesian coordinates $h(x,y) = 0$; an equation in polar coordinates $p(r, \phi) = 0$; a pair of parametric equations $x = f(t)$, $y = g(t)$; or an equation involving invariants such as arc length and radius of curvature. For example, the deltoid may conveniently be defined as the locus of points on the circumference of a circle C_1 rolling inside another circle C_2 of three times its radius, as points satisfying the cartesian equation (1), as points satisfying the parametric equations (2), or as points satisfying the implicit equation (3), where s represents arc length

$$(x^2 + y^2)^2 - 8ax(x^2 - 3y^2) + 18a^2(x^2 + y^2) = 27a^4 \quad (1)$$

$$\begin{aligned} x &= r(2\cos t + \cos 2t) \\ y &= r(2\sin t - \sin 2t) \qquad -\pi \le t \le \pi \end{aligned} \quad (2)$$

$$9s^2 + \rho^2 = 64r^2 \quad (3)$$

and ρ the radius of curvature. *See* ANALYTIC GEOMETRY.

Properties. The most interesting geometric properties are those preserved by linear transformations, especially translations, rotations, reflections, and magnifications. Useful geometric properties include the number of branches into which the curve is divided; the number and degree of nodes, cusps, isolated points, and flex points; the number of loops; symmetries; branches that go to infinity; and asymptotes.

These terms can be defined informally as follows. A branch is a maximal smooth continuous portion of the curve. A multiple point is a point in the plane that lies on two or more branches; its degree is the number of branches involved. A node is a multiple point where the branches cross. A cusp is a multiple point where the branches meet but do not pass; that is, each of the branches ends at that point. An isolated point is a point of the curve through which no branches pass. Multiple and isolated points are collectively termed singular points. Any point that is not singular is termed ordinary. A flex (or point of inflection) is a point on the curve whose tangent cuts the curve. A smooth closed branch forms a loop. A curve is symmetric about a line L if every line perpendicular to L intersects the curve at equal distances from L on opposite sides of L; that is, portions of the curve form mirror images about L. The curve is symmetric about a point P if every line through P intersects the curve at equal distances from P in opposite directions. An asymptote is a line toward which a branch approaches as it moves to infinity from the origin; the curve and line are said to intersect at infinity.

In addition to these geometric properties, the form of the defining equation is of interest. This can be an algebraic (polynomial in x and y) or transcendental equation. In the former case, quadratic, cubic, and quartic equations are of special interest.

Once a coordinate system has been chosen, and the defining equation is known (in any of the forms, though the parametric form is usually the most useful), various properties of the curve can be defined in terms of the equation. These include the locations of x- and y-intercepts, local maxima and minima, flexes, nodes, and cusps. For any point P on the curve, it is generally possible to specify the equation for the tangent line, radial angle (between the x axis and a line between P and the origin), arc length (from some fixed point on the curve), curvature, and center of curvature. Not all of these features may exist for every point of the curve.

Sketching a curve. It is possible to approximate the curve C represented by Eq. (4) by selecting points

$$h(x, y) = 0 \quad (4)$$

on the x axis at increasing distances from the origin, solving for y for each point, and plotting the resulting (x,y) pairs. Some experimentation is required to decide how close together the points on the x axis must be. Iterative techniques are in general necessary to find approximations to y. If the parametric equations are known or can be found, the process becomes easier, since x and y can be evaluated directly for each value of t to yield the coordinates of the points to be plotted. The process is greatly facilitated by using the defining equation (4) and its partial derivatives to help determine singular points, symmetries, and asymptotes of the curve C.

Derived curves. Once the properties of some curve C are known, they may possibly be used to derive

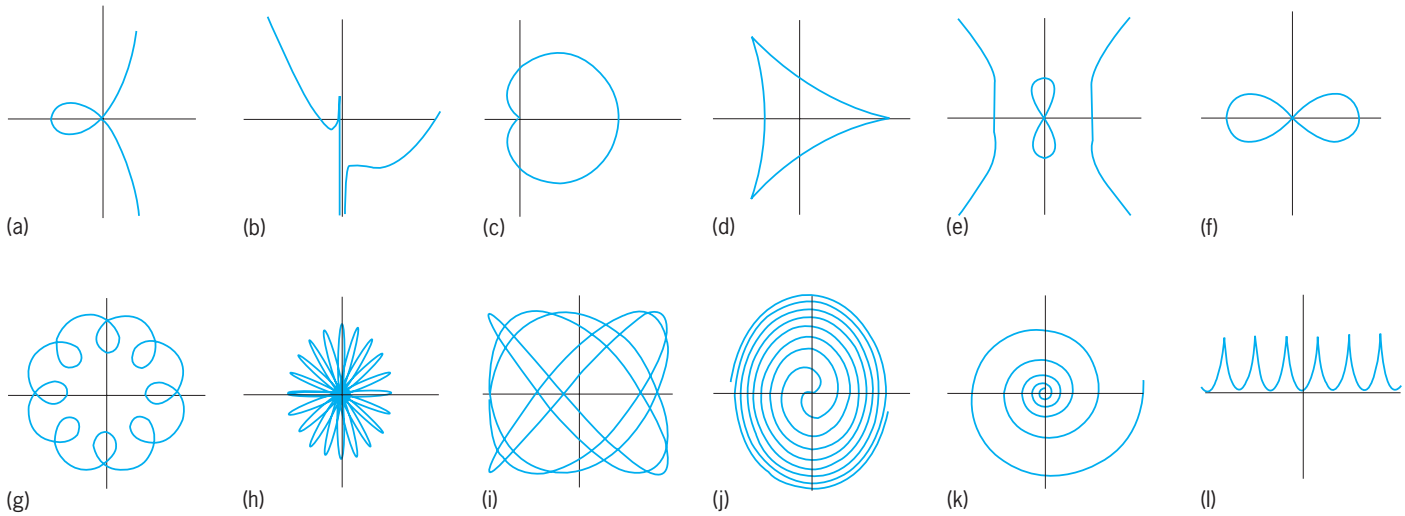

Plane curves. (*a*) Right strophoid. (*b*) Trident of Newton. (*c*) Cardioid. (*d*) Deltoid. (*e*) Devil on two sticks. (*f*) Lemniscate of Bernoulli. (*g*) Epitrochoid. (*h*) Rhodona. (*i*) Bowditch curve. (*j*) Fermat's spiral. (*k*) Logarithmic spiral. (*l*) Cycloid.

other curves. Auxiliary points, lines, and curves are commonly used.

The evolute of C is the locus of the center of curvature. An involute of C is the locus of a fixed point P on a tangent that rolls along C without slipping. A parallel to C is the locus of points a fixed distance from the points P of C, measured along a line perpendicular to the tangent to C at P. *See* INVOLUTE.

A pedal of C is defined with respect to a point O. For every point P on C, the intersection is found of the tangent at P with the line through O perpendicular to the tangent. These intersections form the pedal of C with respect to O.

To define the inverse of C, a point O and a constant k are chosen. For every point P on C, a line segment L is drawn to O, and a point P' on L is so selected that the distance from O to P, multiplied by the distance from O to P', is equal to k. The locus of such points forms an inverse of C with respect to O.

For strophoids, two points O_1 and O_2 are chosen. For each point P on C, the line L is drawn through O_1 and P, and two points Q_1 and Q_2 are found on L such that the distance from P to Q_1 and the distance from P to Q_2 are both equal to the distance from P to O_2. The locus of such points forms the strophoid of C with respect to the two points.

A roulette of C with respect to a second curve C' and a fixed point P on C is the locus of P as C rolls along C' without slipping.

Other types. Plane curves can be defined in many other ways than algebraic or transcendental equations, for example, solutions of differential or integral equations. Plane curves are formed by many special functions, such as trigonometric, exponential, hyperbolic, gamma, Bessel, elliptic, and probability functions. *See* DIFFERENTIAL EQUATION; INTEGRAL EQUATION; SPECIAL FUNCTIONS.

Limit processes of various kinds define curves that were once termed pathological. An example is the Koch curve. An equilateral triangle is drawn with sides of length l. The middle third of each side is replaced by an outward-pointing equilateral triangle of side $l/3$. The process is continued: the middle third of each line segment is replaced with such an extension. The Koch curve is the set of points formed by taking the limit of the construction. It is continuous everywhere, differentiable nowhere (no tangents exist), and has infinite length. A variation of the procedure can be used to define curves that pass through every point in a closed region of the plane. *See* FRACTALS.

Algebraic and transcendental curves. The **illustration** shows some of the many plane curves of sufficient historical interest to have received names.

The right strophoid (illus. *a*) is the strophoid of a line with respect to two points O and A, where A is the foot of the perpendicular from O to the line. It can also be defined as an inverse of a rectangular hyperbola, or as a pedal of a parabola. It is algebraic of degree 3 (cubic), has a single node at the origin, is symmetric about the x axis, and has asymptote $x = a$. It is expressed by Eq. (5).

$$y^2(a - x) = x^2(a + x) \tag{5}$$

The trident of Newton, (illus. *b*), also known as the parabola of Descartes, has a discontinuity at $x = 0$ and one flex; the y axis is an asymptote. It is expressed by Eq. (6).

$$xy = ax^3 + bx^2 + cx + d \qquad ad \neq 0 \tag{6}$$

The cardioid (illus. *c*) is an inverse of a parabola with respect to the focus. It can also be defined as the pedal of a circle with respect to a point on the circumference. The curve is a closed quartic with a cusp and is symmetric about the x axis. It is expressed by Eq. (7).

$$(x^2 + y^2 - 2ax)^2 = 4a^2(x^2 + y^2) \tag{7}$$

See CARDIOID.

The deltoid (illus. *d*), or tricuspid, has three cusps

and three lines of symmetry and forms a closed curve. It is expressed by Eq. (1).

The devil on two sticks (illus. *e*), or devil's curve, is a quartic with two infinite branches, a loop, and a node at the origin. It is symmetric about both axes and the origin. Asymptotes are $x \pm y = 0$. It is expressed by Eq. (8).

$$y^4 - a^2y^2 = x^4 - b^2x^2 \tag{8}$$

The lemniscate of Bernoulli (illus. *f*), or hyperbolic lemniscate, is a closed quartic with a single segment and one node (at the origin). There is one loop, and the curve is symmetric with respect to both axes and the origin. It is expressed by Eq. (9).

$$(x^2 + y^2)^2 = a^2(x^2 - y^2) \tag{9}$$

See LEMNISCATE OF BERNOULLI.

The epitrochoid (illus. *g*) is the roulette traced by a point attached to a circle rolling about the outside of another fixed circle. Depending on the ratio of the radii of the circles, the resulting curve may be algebraic or transcendental; there are many loops. The epitrochoid is best defined parametrically by Eqs. (10).

$$\begin{aligned} x &= m \cos t - b \cos \frac{m}{b} t \\ y &= m \sin t - b \sin \frac{m}{b} t \end{aligned} \tag{10}$$

The rhodona (illus. *h*), or rose, is the locus of the polar equation (11). If m is an even integer, there are

$$r = a \cos m\phi \tag{11}$$

$2m$ petals; if m is an odd integer, there are m petals. *See* ROSE CURVE.

The Bowditch curve (illus. *i*), or curve of Lissajous, is defined by Eqs. (12). The curve is algebraic if n

$$\begin{aligned} x &= a \sin (nt + d) \\ y &= b \sin t \end{aligned} \tag{12}$$

is rational; transcendental otherwise. Depending on the parameters, the curve may be closed or not, may have zero to many nodes, and generally forms a variety of complex shapes. *See* LISSAJOUS FIGURES.

Fermat's spiral (illus. *j*) is defined by the polar equation (13) with parameter $m = 2$. It is a special case

$$r^m = a^m\phi \tag{13}$$

of the archimedean spirals.

The sinusoidal spirals form a family given by the polar equation (14). Special cases are the logarithmic

$$r^n = a^n \cos n\phi \tag{14}$$

spiral (illus. *k*; $n = 0$), Cayley's sextic ($n = {}^1/_3$), cardioid ($n = {}^1/_2$), circle ($n = 1$), lemniscate of Bernoulli ($n = 2$), parabola ($n = -{}^1/_2$), and line ($n = -1$). *See* PARABOLA.

The cycloid (illus. *l*) is the locus of a point attached to a circle rolling on a line. The shape varies somewhat, depending on whether the point is inside or outside the circle or on the circumference. It is defined by the parametric equations (15).

$$\begin{aligned} x &= at - b \sin t \\ y &= a - b \cos t \end{aligned} \tag{15}$$

See CYCLOID. J. Dennis Lawrence

Bibliography. E. Brieskorn and H. Knorrer, *Plane Algebraic Curves*, 1986; W. Fulton, *Algebraic Curves*, 1989; J. D. Lawrence, *A Catalog of Special Plane Curves*, 1972; D. H. von Seggern, *CRC Handbook of Mathematical Curves and Surfaces*, 2d ed., 1992; R. J. Walker, *Algebraic Curves*, 1950, reprint 1991.

Plane geometry

The branch of mathematics that deals with geometric figures, that is, collections of points that all lie in the same plane (coplanar). Although the words "point" and "plane" are undefined concepts, for elementary applications the intuitive meanings will serve: a point is a location, and a plane is a flat surface. For similar definitions, together with a discussion of the postulates and axioms (assumed truths) used in plane geometry, *see* EUCLIDEAN GEOMETRY.

Dimensions and measures. There are three spatial dimensions; geometric figures are classified as being zero-, one-, two-, or three-dimensional. Plane geometry deals only with geometric figures having fewer than three dimensions. Three-dimensional geometric figures, called solids, are dealt with in another branch of euclidean geometry. *See* SOLID (GEOMETRY).

A dimension is any measurement associated with a geometric figure that has units of length. A length can be described by giving a number together with a standard unit of length (such as centimeter, foot, or mile), expressing a defined and often-used distance between two points. The measure of a geometric figure is a number multiplied by a power of a length, with the result giving information about the size of the figure. The power of the length will be 1, 2, or 3, depending on whether the figure is one-, two-, or three-dimensional. *See* UNITS OF MEASUREMENT.

A point is the only zero-dimensional geometric figure, because it has no size, only location. A point has no measure.

A curve is a one-dimensional geometric figure, since its measure is the first power of a length. A curve may be thought of as the path followed by a moving point (although, more precisely, a curve is a collection of different points). If the moving point never changes its direction of motion, the curve is a straight line (or merely line). The measure of a curve (its length) is the distance that the point moves. Some curves have finite length (for example, a circle) while others have infinite length (for example, a straight line).

A region is a two-dimensional geometric figure, since its measure (the area) is the second power of a length. A region is bounded if it can be surrounded by a curve of finite length, in which case the region has a finite area. The boundary of a region is the curve

Symbols commonly used in geometry

Example	Meaning	Comments
$\overline{AB}$	Line segment having end points A and B	Could be named $\overline{BA}$
AB	Length of $\overline{AB}$	$AB = 5$ means $\overline{AB}$ is 5 units long
$\overleftrightarrow{XY}$	Line containing points X and Y	Could be named $\overleftrightarrow{YX}$
$\overrightarrow{PQ}$	Ray with end point P and containing Q	
$\angle V$	Angle with vertex V	Use only if no more than two rays have end point V
$\angle RST$	Angle formed by $\overrightarrow{SR}$ and $\overrightarrow{ST}$	Could be named $\angle TSR$
$\angle x$	Angle named x	
$\angle x = 30^\circ$	Angle named x has measure 30°	
$\overset{\frown}{JK}$	Minor arc with end points J and K	
$\overset{\frown}{JLK}$	Major arc that contains point L	
$\odot C$	Circle with center C	
$\triangle XYZ$	Triangle with vertices X, Y, and Z	Could be named $\triangle YZX$, $\triangle XZY$, and so forth
$\square ABCD$	Rectangle with vertices A, B, C, and D	Adjacent letters must be adjacent vertices
$\square PQRS$	Square with vertices P, Q, R, and S	P and R are not adjacent vertices
$\square ABCD$	Parallelogram with vertices A, B, C, and D	
$ABCD\ldots X$	Polygon with vertices $A, B, \ldots, X$	Adjacent letters must be adjacent vertices
$\overline{AB} \perp j$	A segment ($\overline{AB}$) is perpendicular to a line (j)	
$\overrightarrow{ZP} \parallel \overline{TV}$	A ray ($\overrightarrow{ZP}$) is parallel to a segment ($\overline{TV}$)	
$\overline{AB} \cong \overline{XY}$	Two segments ($\overline{AB}$ and $\overline{XY}$) are congruent	Congruent segments have equal lengths
$\triangle ABC \sim \triangle XYZ$	Two triangles ($\triangle ABC$ and $\triangle XYZ$) are similar	

such that points on one side of the curve are points of the region and points on the other side are not. For bounded regions these points are, respectively, interior points and exterior points. Points on the boundary are normally considered part of the region. A region's shape often can be described by giving the name of the boundary curve. "Circle" and "square," for example, can be applied to either a curve or to the region enclosed by that curve.

A unit square is a square whose sides have a length of one unit. (Here "unit" can be any unit of length, but usually a standard unit is chosen.) The area enclosed by a unit square is the result of squaring the length of a side, giving 1 unit^2, or one square unit. Any region having finite area has a measure equal to the number of unit squares required to cover the region, counting whole squares and fractional parts of squares. many regions have areas that can be determined by using formulas.

Lines, line segments, and rays. Exactly one line passes through two given points. The part of a line between (and including) two points is called a line segment, with the two defining points being the end points of the segment. That part of a line that lies on one side of a point (together with that point) is called a ray. Individual points of interest in a discussion are commonly named (and labeled in a drawing) with uppercase letters. Names of segments, rays, and lines can be formed by using the names of two points of the figure, together with lines or arrows that overlie the names (see **table**).

Angles. An angle is the geometric figure formed by joining two rays having a common end point. Each ray is a side of the angle; the common end point is the vertex of the angle.

An angle can be named by using the angle symbol, $\angle$, followed by the names of three points: a point on one ray, the vertex, and a point on the other ray. The vertex must be the middle-named point. An angle might be named by using $\angle$ followed by a single symbol, often the uppercase-letter name of the vertex, or perhaps a lowercase Roman or Greek letter, such as alpha (α), theta (θ), or phi (ψ) [see table].

Although an angle is the union (joining) of two rays, at least one angle is formed whenever two lines, segments, or rays intersect. In a drawing, it may be necessary to identify a particular angle by drawing a small arc (part of a circle) whose ends lie on the sides of the angle. An angle identified by an arc is labeled by placing its name somewhere close to the

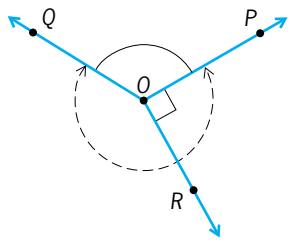

Fig. 1. Three rays with a common end point. Here, ∠*POQ* (solid arc) = $^1/_3$ revolution (an obtuse angle), and ∠*POQ* (dashed arc) = $^2/_3$ revolution (a reflex angle). Line segment $\overline{OP}$ is perpendicular to $\overline{OR}$, so ∠POR (square at vertex) = $^1/_4$ revolution = 90° (a right angle).

arc. Sometimes arrows are needed at the ends of the arc to avoid possible misinterpretation.

The concept of the measure of an angle may be understood by imagining that one ray of an angle is held fixed but the other ray is hinged at the vertex and allowed to rotate in the plane. A measure of the angle is a number, together with some unit of angular measure, that tells how much the hinged ray would need to be rotated so that it would overlie the fixed ray. If a ray were to be rotated exactly one revolution, it would return to its original position. The measure of an angle can be what fraction of one revolution would enable one side of the angle to become coincident with the other. (The measure of an angle is not a dimension, since it does not have units of length.)

In most applications of geometry that involve angles, the measure of an angle is of such importance that an angle is defined as the two rays together with the measure. In plane geometry, all rotations have positive measure, regardless of the direction of rotation (clockwise or counterclockwise). In some applications it is important which ray is rotated (the initial side) and which is held fixed (the terminal side), and whether the rotation of the initial side is clockwise (a negative rotation) or counterclockwise (a positive rotation). In some applications the measure of the angle is the only property of importance, with the locations of the two sides being immaterial. *See* TRIGONOMETRY.

If the measure of an angle has importance, it may be necessary to indicate the measure with an arc in a drawing (**Fig. 1**). If a rotation of one-third revolution is required to bring a ray, $\overrightarrow{OP}$, into coincidence with a second ray, $\overrightarrow{OQ}$, following a given arc, then the angle *POQ* that is identified with this arc has a measure of $^1/_3$ revolution. However, a second arc can be drawn between the rays $\overrightarrow{OP}$ and $\overrightarrow{OQ}$, which is followed by the ray $\overrightarrow{OP}$ when it is rotated in the opposite direction to bring it into coincidence with $\overrightarrow{OQ}$. The angle *POQ* that is identified with this second arc has a measure of $^2/_3$ revolution. If an angle is drawn without an arc, it normally is assumed that the measure is the smallest rotation that would bring one side into coincidence with the other.

The revolution is a convenient unit of angular measure for many applications. However, a more commonly used unit is the degree (°), which is defined by the equation 360° = 1 revolution. In the modern-day use of calculators and computers, fractions of degrees are most conveniently expressed by using decimal fractions; however, another subdivision of degrees is firmly entrenched in many applications: 1 degree = 60 minutes (1° = 60′), and 1 minute = 60 second (1′ = 60″).

An angle of measure $^1/_4$ revolution, or 90°, is called a right angle. Two lines (or rays or segments) that intersect so as to form a right angle are said to be perpendicular (see table). In a drawing, the fact that two geometric figures are perpendicular can be indicated by placing a small square near the intersection (Fig. 1). Two angles are complementary if their measures have a sum of 90°. An angle of measure $^1/_2$ revolution, or 180°, is called a straight angle. Two angles are supplementary if their measures have a sum of 180°.

Other angles can be classified according to how their measures compare with the measures of a right angle and a straight angle: an acute angle has a measure between 0 and 90°; an obtuse angle has a measure between 90 and 180°; and a reflex angle has a measure between 180 and 360° (Fig. 1).

Angles that share a side but whose arcs do not overlap are adjacent angles. Two nonperpendicular intersecting lines form two acute angles and two obtuse angles. Each acute angle and an obtuse angle form a pair of adjacent supplementary angles. Knowledge of the measure of one of these angles is sufficient to determine the measures of the other three.

If two coplanar lines do not intersect (no matter how far extended), those lines are parallel (see table). A transversal is a straight line that intersects two other lines, which themselves might intersect or be parallel. Parallel lines intersected by a transversal form two sets of angles, with each set containing four angles of equal measure. Certain words describe the relative positions of the arcs of angles in a set (**Fig. 2**). The arcs of vertical angles lie on opposite sides of the same vertex. Corresponding angles have different vertices, but their arcs have the same

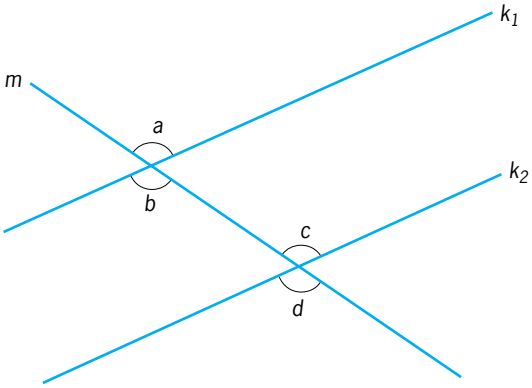

Fig. 2. Relationships between the obtuse angles formed when parallel lines, k_1 and k_2, are intersected by a transversal, *m*. The angles ∠*a* and ∠*b* are vertical angles, as are ∠*c* and ∠*d*. The angles ∠*a* and ∠*c* are corresponding angles, as are ∠*b* and ∠*d*. The angles ∠*b* and ∠*c* are alternate interior angles; ∠*a* and ∠*d* are alternate exterior angles. Similar relationships exist between the acute angles formed.

positions relative to the vertices. Alternate angles have different vertices, and their arcs lie on opposite sides of the transversal. The arcs of interior angles lie between the parallel lines, while those of exterior angles lie outside the parallel lines.

Polygons. A polygon is the geometric figure formed when line segments are joined end to end so as to enclose a region of the plane. The polygon having the least number of sides (three) is the triangle. Three lines, no two of which are parallel, generally intersect in three points, and the union of the three line segments having those points as end points forms a triangle. (Although a triangle is a one-dimensional geometric figure, sometimes the word means a region enclosed by a triangle.) An angle whose sides are two of those segments and whose arc lies inside the triangle is an interior angle of the triangle; an exterior angle of the triangle is any angle that is adjacent and supplementary to an interior angle (**Fig. 3**). The interior angles of any triangle have measures that sum to 180°. This fact allows the determination of the measures of all angles formed by the intersections of three line if the measures of only two angles with different vertices are known. *See* POLYGON.

Congruent and similar geometric figures. Two geometric figures are congruent (≅) if they have exactly the same shape and size. As examples, any two lines or any two rays are congruent; two segments are congruent only if they have the same length; two angles that have equal measures are congruent.

If two geometric figures are congruent, one figure could be made to overlie the other by a combination of these types of motion: translation (sliding), rotation (twisting), and reflection about a line (flipping over). The parts of the geometric figures that would then coincide are called corresponding parts, where a part of a geometric figure is any set of points associated with that figure.

Two triangles are congruent if any one of the following is satisfied:

1. Each side of one triangle is congruent to a corresponding side of the other triangle.
2. Two sides and the included angle of one triangle are congruent to the corresponding sides and included angle of the other triangle. (An angle is included if the two sides of the triangle are sides of the angle.)
3. Two angles and the included side of one triangle are congruent to the corresponding angles and included side of the other triangle. (A side of a triangle is included if it is a side of each of the two angles.)

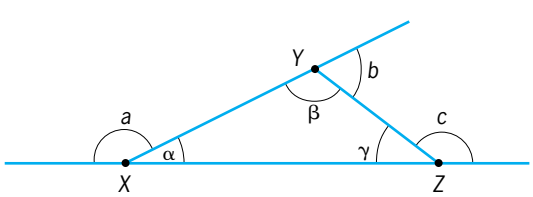

Fig. 3. Triangle *XYZ*, with interior angles $\angle\alpha$, $\angle\beta$, and $\angle\gamma$; and exterior angles $\angle a$, $\angle b$ and $>\angle c$.

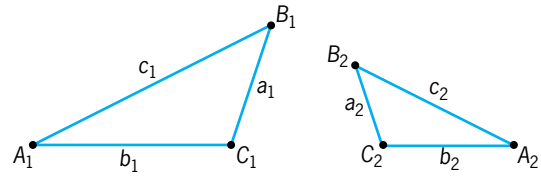

Fig. 4. Two similar triangles, $A_1B_1C_1$ and $A_2B_2C_2$, with the corresponding parts have the same letter names but different subscripts. Thus $\angle A_1 \cong \angle A_2$; $\angle B_1 \cong \angle B_2$; $\angle C_1 \cong \angle C_2$; $a_1/a_2 = b_1/b_2 = c_1/c_2 = 3/2$.

Two geometric figures are similar (∼) if they have the same shape but (perhaps) have different sizes. An examples, any two line segments are similar; any two squares are similar. If two geometric figures are similar, one figure could be made to overlie the other by a combination of translation, rotation, reflection, and either expanding or shrinking. The parts that would then coincide are corresponding. Two similar geometric figures will have corresponding angles that have equal measures, and will have corresponding dimensions that are proportional.

Most common applications of similar geometric figures involve triangles (**Fig. 4**). If the angles of one triangle have, respectively, measures that equal the measures of the angles of another triangle, those triangles are similar. If all pairs of corresponding sides have lengths that are proportional, those triangles are similar. (Neither of these two facts applies to polygons having more than three sides.)

Circles. A circle is a collection of points in the plane, all of which are the same distance from another point, called the center. The region bounded by a circle sometimes is called a disk, but usually the word "circle" is used, relying on context to distinguish between the curve and the region. Circumference means either the circle that is the boundary curve of a disk or the distance around that circle. A radius of a circle is any line segment that joins the center and a point of a circle. A chord is any line segment whose end points lie on the circle. A diameter is any chord that contains the center. Circles having the same center are concentric.

Any straight line and a given circle have two points, or one point, or on points of intersection. A secant is a line that intersects a circle in two points. A tangent is a line that intersects a circle in only one point, called the point of tangency. A tangent line is perpendicular to the radius that joins the center with the point of tangency (**Fig. 5**).

Associated with a circle are several important dimensions, including the lengths of a radius and a diameter. These lengths often are merely called the radius and the diameter of the circle, usually denoted, respectively, by r and d. These symbols appear in almost all formulas involving the length of the circumference C (called the circumference) or the area A enclosed by a circle (called the area of the circle). Relationships between these variables for any circle are given by Eqs. (1).

$$d = 2r \qquad C = \pi d \qquad A = \pi r^2 \qquad (1)$$

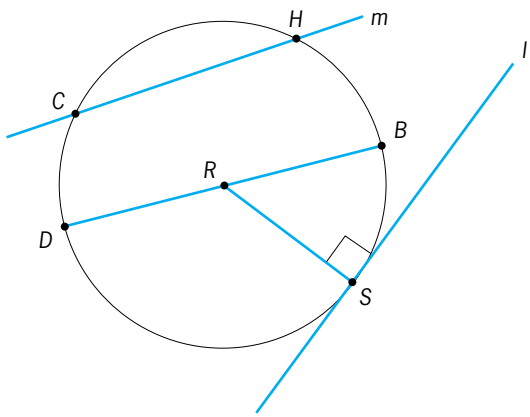

Fig. 5. Circle with center at *R*. Here $\overline{CH}$ and $\overline{DB}$ are chords; $\overline{RS}$, $\overline{DR}$, and $\overline{RB}$ are some radii; $\overline{DB}$ is a diameter; *m* is a secant line; and *l* is a tangent line with point of tangency *S*.

In these formulas, π (the lowercase Greek letter pi) represents the irrational number (value 3.141592. . .) that is usually defined as the ratio of the circumference to the diameter of any circle.

The end points of any chord divide a circle into two arcs: the larger major arc and the smaller minor arc. If the chord is a diameter, the arcs are of equal size and are semicircles. Rays from the center that pass through the end points define two central angles: a reflex angle that corresponds to the major arc, and an obtuse or acute angle that corresponds to the minor arc.

Any angle whose vertex lies on the circle and whose sides both intersect the circle at points other than the vertex is an inscribed angle. The points where the sides of an inscribed angle intersect the circle define a chord, which in turn defines two central angles corresponding to the major and minor arcs. The central angle corresponding to the arc that does not contain the vertex of the inscribed angle has twice the measure of the inscribed angle (**Fig. 6**). Therefore, any inscribed right angle intersects a circle at the end points of a diameter.

A unit circle is a circle having a radius of length 1 unit, and will have a circumference equal to 2π units. Since an arc's central angle is directly proportional to the length of that arc, an angle's measure can be given by the number of units of arc length on the unit circle. Such a unit of angular measure is called a radian, and one revolution equals 2π radians. All circles are similar, so the length s of a circular arc

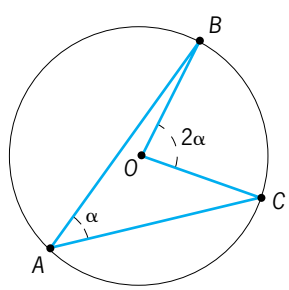

Fig. 6. Circle with inscribed angle and central angles, showing relationship. Here, $\angle BAC = \angle\alpha$ is an inscribed angle, $\angle BOC$ is the central angle of $\overset{\frown}{BC}$, and reflex$\angle BOC$ is the central angle of $\overset{\frown}{BAC}$. Then $\angle BOC = 2 \cdot \angle BAC$.

of a circle of radius r is given by Eq. (2), where θ is

$$s = r\theta \tag{2}$$

the measure of the central angle in radians.

A circular sector is the region bounded by two radii and an arc of a circle. Since the area, A_s, of a circular sector is directly proportional to the measure of the arc's central angle θ, Eq. (3) follows, where θ

$$\frac{\theta}{2\pi} = \frac{A_s}{A} = \frac{A_s}{\pi r^2} \tag{3}$$

is measured in radians, so that the area of the circular sector is given by Eq. (4).

$$A_s = {}^1\!/_2 r^2\theta \tag{4}$$

A circular segment is the region bounded by a circular arc and the chord that joins the end points of that arc. The area, A_g, of a circular segment of a circle having radius r and central angle θ is given by Eq. (5), where θ is measured in radians.

$$A_g = {}^1\!/_2 r^2(\theta - \sin\theta) \tag{5}$$

See CIRCLE. Harry L. Baldwin, Jr.

Bibliography. H. L. Baldwin, Jr., *Essentials of Geometry*, 1993; D. Hilbert and S. Cohn-Vossen, *Geometry and the Imagination*, 1952, reprint 1999; P. G. O'Daffer and S. R. Clemens, *Geometry: An Investigative Approach*, 2d ed., 1992; I. Todhunter, *The Elements of Euclid*, 1891.

Planer

A machine for the shaping of long, flat, or flat contoured surfaces by reciprocating the workpiece under a stationary single-point tool or tools. Usually the workpiece is too large to be handled on a shaper.

Planers are built in two general types, open-side or double-housing. The former is constructed with one upright or housing to support the crossrail and tools. The double-housing type has an upright on either side of the reciprocating table connected by an arch at the top.

Saddles on the crossrail carry the tools which feed across the work. A hinged clapper box, free to tilt, provides tool relief on the return stroke of the table. A variation is the milling planer; it uses a rotary cutter rather than single-point tools. *See* WOODWORKING.

Alan H. Tuttle

Planet

A relatively large celestial body moving in orbit around a star, in particular the Sun.

Solar system. Planets are by-products of the formation of the Sun, which condensed from a vast cloud of interstellar gas and dust. Initially the Sun was surrounded by a broad, flat disk of residual matter, from which the planets accumulated. *See* SOLAR SYSTEM.

The Sun is not the only star known to be encircled by planets. Astronomers now know of at least 180

other solar systems, more than 20 of which have multiple planets. *See* EXTRASOLAR PLANETS.

Definition. The word "planet" derives from the Greek word *planasthai*, meaning "to wander," and it was used by ancient stargazers to identify bright starlike objects that appeared to wander against the background of fixed stars.

Prior to 2006, there was no formal definition for a planet. However, in that year members of the International Astronomical Union (IAU) defined a planet in our solar system as an object orbiting the Sun that is massive enough to have assumed a round shape and to have cleared the neighborhood around its orbit of smaller objects through gravitational scattering. Smaller, less massive objects that are round in shape but unable to clear their orbital neighborhood are considered dwarf planets.

The IAU has not defined an upper size or mass limit for a planet. However, many astronomers informally favor a limit of 13 times the mass of Jupiter (presuming the object to have similar elemental ratios). Bodies with greater masses can generate energy through the fusion of deuterium in their cores and thus are considered to belong to a class of substellar objects called brown dwarfs. *See* BROWN DWARF.

Planets in our solar system. The solar system's eight (major) planets fall into two basic groups: the small, dense, terrestrial planets—Mercury, Venus, Earth, and Mars—and the giant or Jovian planets Jupiter, Saturn, Uranus, and Neptune. The terrestrial planets are all located relatively close to the Sun (**Fig. 1**). The lower-density giant planets extend outward from Jupiter to great distances. This distribution is not accidental, but is related to the fractionation of rocky, icy, and gaseous materials present in the circumstellar disk during the earliest stages of the solar system's formation. *See* JUPITER; MARS; MERCURY (PLANET); NEPTUNE; SATURN; URANUS; VENUS.

Minor bodies and dwarf planets. Each of the planets from Earth to Neptune is accompanied by one or more secondary bodies called satellites. Many of the smallest satellites are not observable from Earth, but were discovered during spacecraft visits. *See* SATELLITE (ASTRONOMY).

The gas-and-dust disk that surrounded the infant Sun also gave rise to a multitude of smaller bodies, which fall into two main groups. The asteroids, of which nearly 140,000 have well-determined orbits, are rocky and largely confined between the orbits of Mars and Jupiter; objects in the Kuiper Belt, hundreds of which are now known but which may number in the billions, are primarily icy and lie beyond the orbit of Neptune. *See* ASTEROID; KUIPER BELT.

Until recently, astronomers also considered Pluto to be a planet. However, the discovery of Kuiper Belt objects with orbits similar to Pluto's and in particular the discovery in 2003 of Eris, a very distant solar-system object larger than Pluto, triggered a protracted debate about whether Pluto truly qualified as a planet. In 2006, astronomers of the IAU approved a resolution that applies the term "planet" only to the solar system's eight largest worlds. In addition, Pluto, Eris, and the large asteroid Ceres became the first objects in a new class of "dwarf planets," a designation that may soon be applied to the largest asteroids and Kuiper Belt objects. *See* CERES; PLUTO.

Planetary orbits and motions. The motions of the planets in their orbits around the Sun are governed by

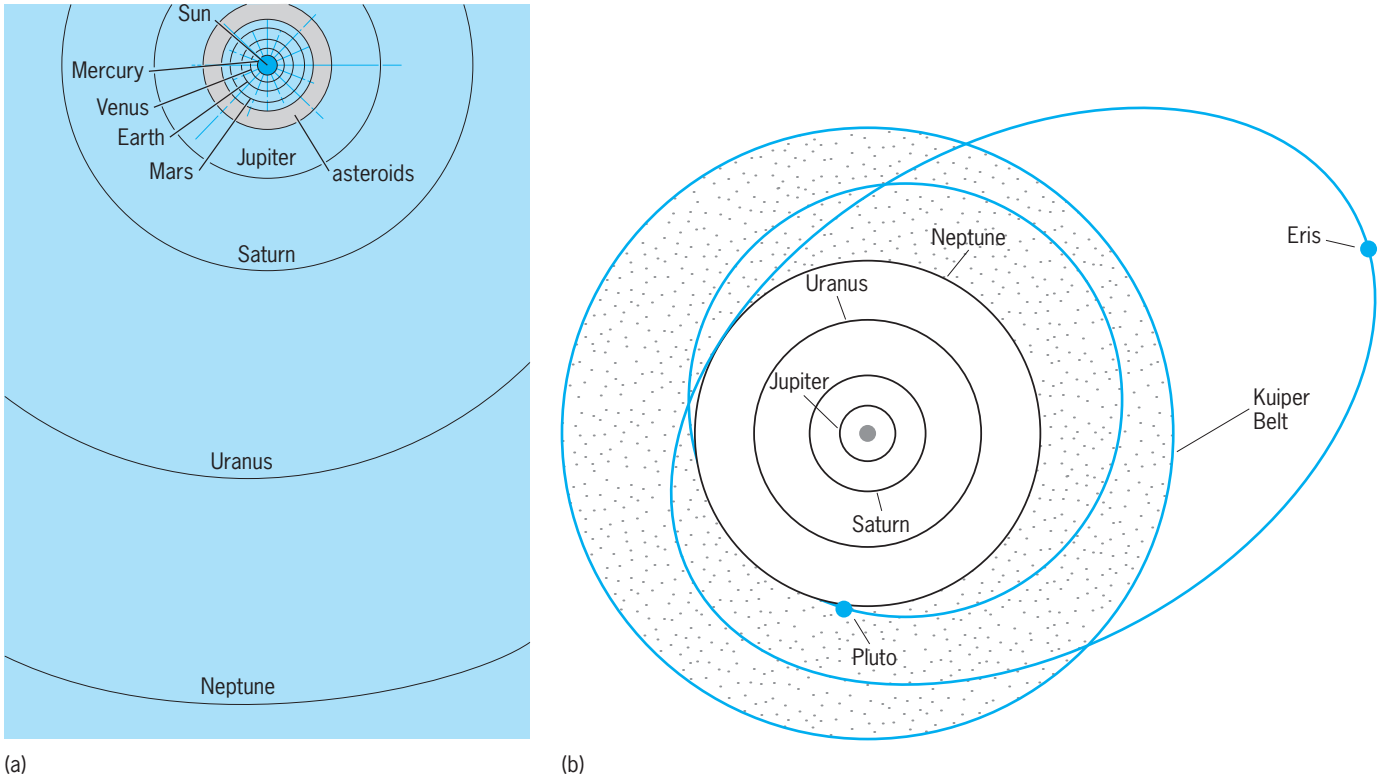

Fig. 1. Plan of the solar system. (*a*) Solar system out to Neptune. (*b*) Outer solar system. Current (2006) positions of Pluto and Eris in their orbits are indicated.

three laws of motion discovered by Johannes Kepler at the beginning of the seventeenth century. *See* CELESTIAL MECHANICS; KEPLER'S LAWS.

First law: The orbit of a planet is an ellipse, with the Sun at one of its foci.

Second law (the law of areas): As a planet revolves in its orbit, the radius vector (the line from the Sun to the planet) sweeps out equal areas in equal intervals of time.

Third law (the harmonic law): The square of the period of revolution P is proportional to the cube of the orbit's semimajor axis a; that is, for all planets the ratio P^2/a^3 is a constant.

If a is expressed in astronomical units and P in years, $P^2/a^3 = 1$. One astronomical unit (1 AU), the mean distance from Earth to the Sun, is approximately 92.96×10^6 mi (149.6×10^6 km). Otherwise, the constant of the harmonic law is given by Newton's law of gravitation as $G(M + m)/4\pi^2$, where M and m are the masses of the Sun and the planet, respectively, and G is the constant of gravitation (6.674×10^{-11} m^3/kg s^2). *See* ASTRONOMICAL UNIT; EARTH ROTATION AND ORBITAL MOTION; GRAVITATION; YEAR.

Planets revolve around the Sun along elliptical (not circular) paths (**Fig. 2**). For an elliptical orbit with the Sun at the focus S, a planet comes closest to the Sun on each revolution at perihelion q and lies farthest away at aphelion Q. The semimajor axis $a = (qS + QS)/2$, and the ellipse's eccentricity $e = CS/a$, where C is the ellipse's center. The perihelion distance $qS = a(1 - e)$, and the aphelion distance $QS = a(1 + e)$. At any other point, the Sun-planet distance $r = a(1 - e^2)/(1 + e \cos \theta)$, where θ is the angle qSA, termed the true anomaly. At that same point the planet's velocity $v = [2GM_{\odot}(1/r - 1/2a)]^{1/2}$, where $M_{\odot}$ is the mass of the Sun. *See* ELLIPSE.

Orbital elements. The position of a planet in its orbit and the orientation of the orbit in space are completely defined by seven orbital elements (**Fig. 3**). These are (1) the semimajor axis a, (2) the eccentricity e, (3) the inclination i of the plane of the orbit to the plane of the ecliptic, (4) the longitude Ω of the ascending node N, (5) the angle ω from the ascending node N to the perihelion q, (6) the sidereal period of revolution P, or the mean (daily) motion $n = 2\pi/P$, and (7) the date of perihelion passage T, or epoch E.

The intersection NN' of the plane of a planet's orbit

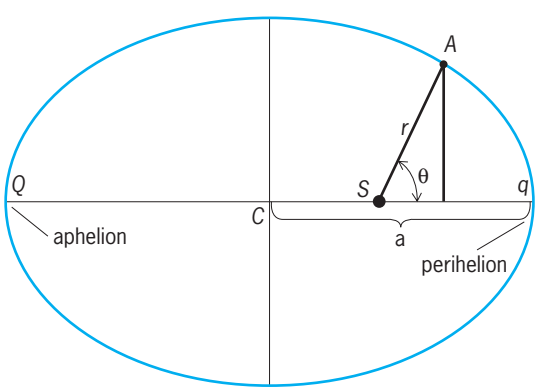

Fig. 2. Elliptic motion of a planet. Symbols explained in text.

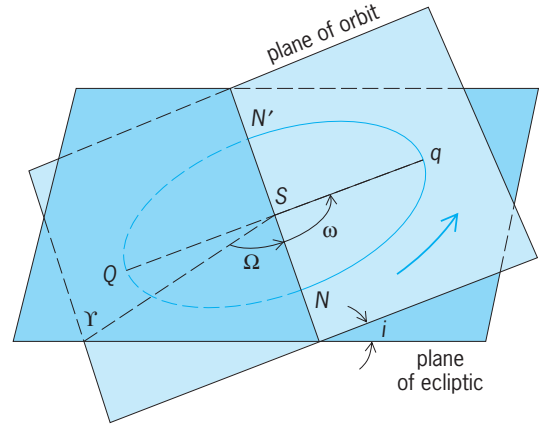

Fig. 3. Orbital elements, which determine the position of a planet in its orbit and the orientation of the orbit in space. Symbols explained in text.

and the plane of the ecliptic (defined by Earth's orbit) is the line of nodes. The planet crosses the plane of the ecliptic from south to north at the ascending node N and from north to south at the descending node N'. The longitude of the ascending node is the angle Ω measured in the plane of the ecliptic from the vernal equinox Υ. The longitude of perihelion $\overline{\omega} = \Omega + \omega$. The location of the plane of the orbit in space is defined by i and Ω, the orientation of the ellipse in this plane by ω, its form by e, its size by a, and the position of the planet on the ellipse by P and T (and by the time t). *See* ORBITAL MOTION.

Determination of orbital elements. The orbital elements of the planets are deduced from accurate telescopic observations of the planets' positions in the sky with respect to background stars, together with tracking data for spacecraft in the planets' vicinity and radar-derived range and velocity measurements. The orbital elements allow the motion of each body to be calculated into the past and future; a set of these calculated positions, published in tabular form, is called an ephemeris. *See* EPHEMERIS.

The main orbital characteristics of the planets and dwarf planets are given in **Table 1**.

Planetary configurations. In the course of their motions around the Sun, Earth and other planets occupy a variety of relative positions or configurations (**Fig. 4***a*). The configurations for the inferior planets, Mercury and Venus, which are located inside Earth's orbit, differ from those for the superior planets, from Mars to Neptune, circulating outside Earth's orbit.

Inferior planets. Venus and Mercury are in conjunction with the Sun when closest to the Earth-Sun line, either between Earth and the Sun (inferior conjunction) or beyond the Sun (superior conjunction). On rare occasions when the planet is very close to the plane of Earth's orbit at the time of an inferior conjunction, a transit in front of the Sun is observed. *See* TRANSIT (ASTRONOMY).

Between conjunctions, a planet's elongation, its angular distance from to the Sun as measured from Earth's center, varies up to a maximum value; the greatest elongations of Mercury and Venus are 28° and 47°, respectively. The superior planets are not so limited, and their elongations can reach up to 180°

TABLE 1. Elements of planetary orbits

Planets	Symbol	Mean distance from Sun (semimajor axis of orbit) AU	10^6 mi	10^6 km	Sidereal period of revolution Years	Days	Synodic period, days	Mean orbital velocity mi/s	km/s	Orbital eccentricity	Orbital inclination, degrees
Mercury	☿	0.387	36.0	57.9	0.241	87.97	115.88	29.8	47.9	0.206	7.00
Venus	♀	0.723	67.2	108.2	0.615	224.70	583.92	21.8	35.0	0.007	3.39
Earth	⊕	1.000	93.0	149.6	1.000	365.26		18.5	29.8	0.017	0.00
Mars	♂	1.524	141.6	227.9	1.881	686.98	779.94	15.0	24.1	0.093	1.85
Jupiter	♃	5.203	483.6	778.4	11.86	4,332.7	398.88	8.1	13.1	0.048	1.30
Saturn	♄	9.555	886.7	1426.7	29.46	10,759.5	378.09	6.0	9.7	0.054	2.49
Uranus	♅	19.19	1784.	2871.	84.00	30,685.0	369.66	4.2	6.8	0.047	0.77
Neptune	♆	30.07	2795.	4498.	164.79	60,190.0	367.49	3.4	5.5	0.009	1.77
Dwarf planets*	⯓										
Ceres		2.77	257.1	413.8	4.600	1,680.1	466.80	11.1	17.9	0.079	10.61
Pluto	♇	39.48	3670.	5906.	247.9	90,550.	366.73	2.9	4.7	0.249	17.14
Eris		67.76	6299	10137	557.7	203,720.	365.38	2.1	3.4	0.441	44.0

*No other bodies had been designated dwarf planets by late 2006, but several likely candidates exist.

when they are in opposition with the Sun; when the elongation is $\pm 90^\circ$, they are in quadrature (eastern or western) with the Sun.

Phases of planetary disks. The telescopic aspects of the disks of the planets vary according to the phase angle, the angle between the direction of illumination and observation. Between inferior conjunction and greatest elongations, the interior planets show crescent phases, like the Moon between new moon and first or last quarters (Fig. 4*b*); between greatest elongations and superior conjunction, they show gibbous phases, like the Moon between quarters and full moon. At superior conjunction they show a fully illuminated circular disk, while during transits their disks are seen in silhouette against the Sun. The superior planets show their full phase at both conjunction and opposition and a gibbous phase near quadrature, at which time the unilluminated portion of the disk is greatest. *See* PHASE (ASTRONOMY).

Apparent motions. Because the orbits of the main planets are only slightly inclined to the plane of the orbit of Earth, the apparent paths of the planets are restricted to the zodiac, a belt 16° wide centered on the ecliptic. The ecliptic is the path in the sky traced out by the Sun in its apparent annual journey as Earth revolves around it. Along this path, the apparent motions of the inferior planets with respect to the Sun are alternatively westward, from greatest elongation through inferior conjunction to greatest elongation, then eastward, from greatest elongation through superior conjunction to greatest elongation (Fig. 4). The mean motion of the superior planets is always westward. *See* ASTRONOMICAL COORDINATE SYSTEMS; ECLIPTIC.

For inferior planets, the apparent motions with respect to the celestial sphere, that is, to the fixed stars, appear as oscillations back and forth about the position of the Sun steadily moving eastward among the stars. For a superior planet, the apparent motion is generally eastward or direct, but for short periods near the time of opposition it is westward or retrograde (**Fig. 5**). At times when the direction of the apparent motion on the sphere reverses, the planet appears to be stationary.

The mean interval of time between successive returns to the same place with respect to the stars is the sidereal period, reflecting the true motion of revolution of the planet in its orbit around the Sun. The mean interval of time between successive returns of the same configuration with respect to the Sun as

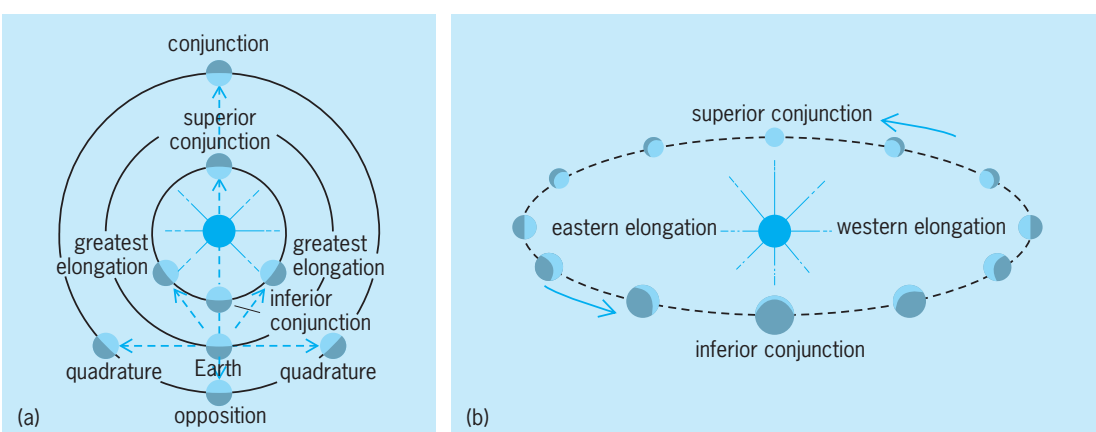

Fig. 4. Planetary configurations and phases. (*a*) Positions of Earth and other planets relative to the Sun. (*b*) Phases of inferior planets (Mercury and Venus). (*After L. Rudaux and G. de Vaucouleurs, Larousse Encyclopedia of Astronomy, Prometheus Press, 1959*)

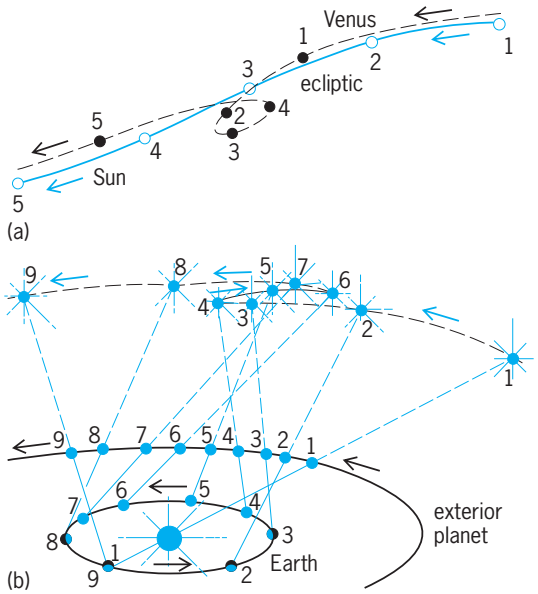

Fig. 5. Apparent motions as observed from Earth (*a*) of an interior planet with respect to the Sun and (*b*) of an exterior planet with respect to the fixed stars. (*After L. Rudaux and G. de Vaucouleurs, Larousse Encyclopedia of Astronomy, Prometheus Press, 1959*)

seen from Earth (for example, between conjunctions or oppositions) is the synodic period (Table 1).

Planetary characteristics. The apparent diameter of a planet may be determined telescopically or by planetary spacecraft. If the apparent diameter of a planet is d'' when its distance to Earth is Δ, the planet's diameter $D = \Delta \sin d'' = d''/206{,}265$, where d'' is measured in seconds of arc, and both Δ and D are expressed in astronomical units.

Polar flattening is apparent only on the planets Jupiter and Saturn. A planet's polar and equatorial radii, r_p and r_e, can be used to establish its mean radius $r = (r_p + 2r_e)/3$ and its ellipticity $\epsilon = 1 - (r_p/r_e)$. The mean radius may also be expressed in terms of the mean radius of Earth (3959 mi or 6371 km); the relative surface area is then very nearly equal to r^2 and the relative volume to r^3. For the nearer planets with a solid surface (Mercury, Venus, and Mars), their diameters can also be determined precisely by radar, using the time-delay technique. *See* RADAR; RADAR ASTRONOMY.

Masses, gravity, and density. The mass of a planet is deduced easily if it has one or more satellites. If a is the mean distance (semimajor axis) of the satellite's orbit and P its period of revolution expressed respectively in astronomical units and sidereal years, the mass m of a planet, expressed as a fraction of the Sun's mass, is given through Newton's law of gravitation by $m = a^3/P^2$. This assumes that the mass of the satellite relative to that of the planet may be neglected, which is nearly always the case within the accuracy of the data.

The masses of Mercury and Venus, which have no satellites, were initially estimated from the perturbations each caused in the motions of the other planets or of comets passing nearby. Since these perturbations are small, the masses so obtained are generally of low accuracy. However, the mass of Mercury was refined during the 1970s by its perturbations on passing spacecraft. The mass of Venus has now been determined precisely by its perturbations on spacecraft in orbit around it. *See* PERTURBATION (ASTRONOMY).

Once the mass m and the radius r of a planet are known in terms of Earth's mass and radius, its surface gravity and mean density relative to Earth are given by $g = m/r^2$ and $\rho = m/r^3$, respectively. Multiplication by 981 and 5.52 gives the corresponding values in cgs units (cm/s^2 and g/cm^3, respectively). From r and m follows also the escape velocity V_1 that permits a projectile (or a molecule) to leave the planet on a parabolic orbit: $V_1 = (2Gm/r)^{1/2}$; this is $2^{1/2}$ times the velocity of an hypothetical satellite moving in a circular orbit close to the surface of the planet. These elements are listed in **Table 2**. *See* ESCAPE VELOCITY.

TABLE 2. Physical characteristics of the Sun's planets

Planets	Equatorial radius (r_e) (Earth=1)	mi	km	Ellipticity	Volume (Earth=1)	Mass (Earth=1)	Density, g/cm^3	Escape velocity mi/s	km/s	Rotation period	Obliquity, degrees[1]
Mercury	0.38	1,515	2,440	0.000	0.055	0.055	5.43	2.5	4.4	58 d 15.5 h	0.1
Venus	0.95	3,761	6,052	0.000	0.854	0.815	5.20	6.5	10.4	243 d 0.5 h	177.4[2]
Earth	1.00	3,963	6,378	0.0034	1.000	1.000	5.52	7.0	11.2	23 h 56 m 23 s	23.45
Mars	0.53	2,110	3,397	0.0069	0.151	0.107	3.93	3.1	5.0	24 h 37 m 23 s	25.19
Jupiter	11.21	44,423	71,492	0.0649	1321.3	317.832	1.33	37.0	59.5	9 h 55 m 30 s[3,4]	3.12
Saturn	9.45	37,449	60,268	0.0980	763.6	95.159	0.69	22.1	35.5	10 h 39 m 22 s[3,5]	26.73
Uranus	4.01	15,882	25,559	0.0229	63.1	14.536	1.32	13.2	21.3	17 h 22.2 m[3,6]	97.86[2]
Neptune	3.88	15,389	24,764	0.017	57.7	17.147	1.64	14.6	23.5	16 h 6.6 m[3,7]	29.56
Dwarf planets[8]											
Ceres	0.08	303	487	0.066	0.0004	0.0002	2.08	0.3	0.5	9 h 4.5 m	5
Pluto	0.18	716	1153	?	0.006	0.002	2.1	0.8	1.3	6 d 9 h 17.6 m	119.6[2]
Eris	0.19	745	1,200	?	0.007	?	?	?	?	>8 h	?

[1]Obliquity is the tilt of the equator with respect to the orbit plane.
[2]Venus, Uranus, and Pluto are considered to have retrograde rotation.
[3]Internal (System III) rotation period, the rotation period of the planet's core, as deduced from its magnetic field.
[4]Jupiter's equatorial (System I) rotation period is 9 h 50.5 m.
[5]Saturn's equatorial rotation period is 10 h 14.0 m.
[6]Uranus's equatorial rotation period is about 18.0 h.
[7]Neptune's equatorial rotation period is about 18.8 h.
[8]No other bodies had been designated dwarf planets by late 2006, but several likely candidates exist.

Rotation periods. The rotation periods of all the main planets are now well determined (Table 2). A planet's rotation period can be determined by several methods: (1) Direct observation of permanent surface markings (Mars) or of long-duration atmospheric features (Jupiter, Saturn, Uranus, Neptune) is the classical method. (2) The line-of-sight velocity difference between the opposite equatorial limbs, determined either by means of radar sounding (Mercury, Venus) or spectroscopically can be combined with the diameter to yield a rotation period. (3) Since the giant planets lack solid surfaces, their rotation rates are found by timing the cyclic pattern of radio energy emanating from the planets' magnetospheres, which rotate in synchrony with their deep interiors. (4) When the apparent diameter of the disk is too small for any of these methods (Eris, asteroids), a determination of the periodicity of the light variations provides a fairly accurate value of the rotation period. These variations may be due to the changing presentation of bright and dark regions of the surface, variation in shape (for smaller objects), or a combination of both.

Planetary radiations. The electromagnetic radiation received from a planet is made up of three main components: visible reflected sunlight, including some ultraviolet and near-infrared radiation; thermal radiation due to the planet's heat, including both infrared radiation and ultrashort radio waves; and nonthermal radio emission due to electrical phenomena, if any, in the planet's atmosphere or in its radiation belts.

Planetary brightness. The apparent brightness of a planet, as measured by visual, photographic, and electronic means, is usually expressed as its magnitude, based on the logarithmic scale for stellar magnitudes. A planet's apparent brightness varies in inverse proportion to the squares of the distances r from the Sun and Δ from Earth. Consequently the apparent brightnesses of large solar-system bodies are typically cited when the bodies are at opposition. For smaller bodies, the usual convention is an absolute magnitude H, indicating the body's apparent brightness at a distance from both the Sun and the observer of 1 AU and a phase angle of 0°.

The fraction of the incident light reflected at full phase compared with the fraction that would be reflected under the same conditions by an equivalent perfectly diffusing reflecting disk is called the geometric albedo. It is a measure of the backscattering reflectivity of the planet's visible surface. The visual albedos of the planets vary between 5 and 70%. *See* ALBEDO; MAGNITUDE (ASTRONOMY).

Thermal radiation. The thermal radiation from a planet can be measured either with a radiometer at wavelengths of 8–14 micrometers, 17–25 μm, and 30–40 μm (which are partially transmitted by Earth's atmosphere) or with a radio telescope at wavelengths between 1 mm and 30 cm. In either case, the amount of energy corresponds to that which would be received under the same conditions from a perfect radiator of the same size at a certain temperature T, the blackbody temperature of the planet. Its relation to the actual temperature depends on the thermal and radiative properties of the atmosphere and surface of the planet. Jupiter radiates nearly twice as much energy as it receives from the Sun. From this it is inferred that the planet has an internal source, perhaps primordial heat continuing to escape long after Jupiter's formation 4.6×10^9 years ago. Saturn and Neptune (but not Uranus) also emit more energy than they receive from the Sun. While gravitational separation of the planet's hydrogen and helium has been suggested as one possible mechanism for this excess energy, no single solution appears to account for the entire energy outflow. *See* HEAT RADIATION; INFRARED ASTRONOMY; RADIO ASTRONOMY; RADIOMETRY.

Nonthermal radiation. Large radio telescopes have recorded nonthermal radio emission at decimeter and decameter wavelengths from Jupiter. The decameter emission takes the form of irregular bursts of noise originating within the planet's atmosphere. Voyager spacecraft revealed that powerful electric currents exist inside the Jovian magnetosphere, particularly one called a flux tube linking higher latitudes on the planet with satellite Io. Since the observed decametric radiation is modulated by the orbital position of Io, this current loop may be responsible for the outbursts.

Planetary atmospheres. The principal constituents of the atmospheres of the terrestrial planets are carbon dioxide, nitrogen, water, and (on Earth only) oxygen; Mercury has a very tenuous envelope dominated by atoms of sodium and potassium. The atmospheres of the giant planets are composed primarily of hydrogen and helium, with lesser amounts of methane, ammonia, and water. Atmospheric motions are driven by temperature gradients—in general, those existing between the warm equatorial regions and the cooler polar areas. An atmosphere thus tends to redistribute heat over the planetary surface, lessening the temperature extremes found on airless bodies.

On planets having relatively dense atmospheres, heat from the Sun is trapped by the greenhouse effect. That is, visible radiation from the Sun passes readily through the atmosphere to heat the planetary surface, but infrared radiation reemitted from the surface is constrained from escaping back to space by the lower transparency of certain atmospheric gases (especially carbon dioxide, water vapor, and methane) to longer wavelengths. For example, although Venus absorbs approximately the same amount of energy from the Sun as does Earth, the greenhouse effect is responsible for heating the surface of Venus to a much higher temperature, approximately 750 K (900°F). *See* GREENHOUSE EFFECT.

Directly through wind erosion or indirectly by carrying water vapor (which can precipitate as rain), the atmospheres of the terrestrial planets are a major factor in modifying surface structure and rearranging the distribution of surface materials. Mercury, being virtually airless, exhibits a relatively unmodified surface, very similar in appearance to that of the Moon.

Much of Pluto is covered with methane and nitrogen ice frost, and this dwarf planet maintains

a tenuous atmosphere of nitrogen gas (at least when in the part of its orbit nearest the Sun).

Possible unknown planets. During the nineteenth century, an unexplained irregularity in the motion of Mercury was thought by some investigators to be caused by an unknown planet circulating between the Sun and Mercury, called Vulcan, which was looked for in vain. This irregularity was satisfactorily explained in 1915 by Albert Einstein's general theory of relativity. It is now certain that no intra-Mercurial planet larger than 30 mi (50 km) can exist. *See* RELATIVITY.

The discovery of objects in the Kuiper Belt beginning in 1992 has spurred efforts to find large bodies beyond the orbit of Neptune. Eris was discovered in 2003 at a distance of 97 AU, nearly $2^1/_2$ times Pluto's mean heliocentric distance. Theoretical modeling of the formation of the Kuiper Belt and of outer-planet evolution suggests that even larger Kuiper Belt objects await discovery. J. Kelly Beatty

Origins and taxonomy of planets. The standard model of solar system formation assumes that the early solar nebula condensed from a diffuse interstellar cloud. The collapse of the cloud could have been initiated by cooling (such that thermal pressure in the cloud could no longer balance the cloud's self-gravity) or by perturbations that created a local high-density region. In any case, once self-gravity initiated condensation, the cloud's collapse would have been a runaway process. The cloud would have increased its rotational speed as it condensed (like a skater pulling in her arms), eventually forming a spinning disk with most of its mass in the center. This central mass became the basis for a proto-Sun, and the molecules and atoms in the rest of the disk settled into the central plane of the disk. Some molecules would have condensed into solid particles, and the particles would have accreted into planetesimals, the building blocks of planets. *See* MOLECULAR CLOUD; PROTOSTAR.

Hydrogen and helium are thought to have made up 98% of the solar nebula, followed by hydrogenated forms of carbon, nitrogen, and oxygen (CH_4, H_2O, and NH_3, respectively), silicates (rocks), and metals (iron, nickel, and aluminum). These various constituents condensed into solid particles in the central plane of the disk, but because metals, rocks, and ices solidify at different temperatures, there was a compositional gradient as a function of distance from the protosun. Beyond the frost line (thought to be around 2–5 AU), volatiles like water, methane, and ammonia could condense into ices, which meant that availability of solid material was several times higher outside the frost line than inside. The planetary cores outside the frost line were therefore more massive, enough to attract and retain atmospheres of hydrogen and helium.

Terrestrial planets (Mercury, Venus, Earth, and Mars) are thought to have formed in a region dominated by metallic and rocky planetesimals, and their surface gravities were not sufficient to hold onto hydrogen and helium gases. In contrast, the gas giants (Jupiter and Saturn) accreted massive cores with extremely high escape velocities. They retained hydrogen and helium atmospheres. Their current compositional make-up is similar to that of early solar nebula: 98% hydrogen and helium with traces of other constituents. Uranus and Neptune are sometimes called "ice giants" because their composition is similar to that of comets. They have less hydrogen and helium than the gas giants, but they accreted from numerous volatile-rich icy planetesimals to become an order of magnitude more massive than each of the terrestrial planets. Finally, there are "ice dwarves" like Pluto or Eris, which are also the accretion products of icy planetesimals, but did not grow to the extent of a Neptune or Uranus.

Orbital motion. The mass of the Sun is more than a thousand times greater than that of Jupiter, which in turn is about 2.5 more massive than the rest of the planets combined. As a result, the motion of the planets is mainly two-body Keplerian motion about a massive central body, which means that all of the planets execute elliptical orbits with the Sun at one focus. However, the presence of other planets, small objects, and even photons can have significant effects on the motion of solar system objects.

Resonances. A resonance between two objects means that the same relative geometrical orientation occurs repeatedly. Even small interactions can build

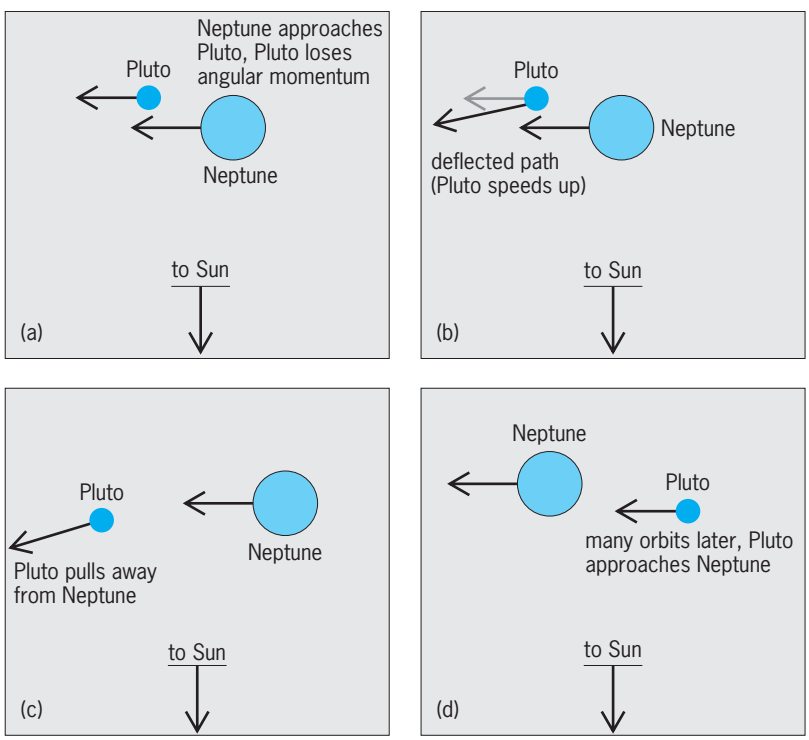

Fig. 6. Neptune-Pluto resonance. (*a*) Imagine that Neptune is approaching Pluto. Neptune is closer to the Sun and its tangential velocity is faster than Pluto's. Will the two collide? (*b*) As Neptune catches up to Pluto, it gains some of Pluto's angular momentum. Neptune's trajectory around the Sun moves to a slightly larger orbit as a result, and Pluto's heading is slightly more toward the Sun than before the encounter. Since Neptune is more massive than Pluto, most of the change is to Pluto's orbit. (*c*) Orbits with smaller semimajor axes have smaller periods. As Pluto loses angular momentum to Neptune, it dives inward toward the Sun on a new orbit whose period is less than before the encounter. (*d*) Eventually, the encounter geometry between Pluto and Neptune is such that Pluto gains angular momentum from Neptune. In these situations, Pluto's orbit is diverted outward from the Sun and Pluto's orbital period increases. Pluto slows down with respect to Neptune, and the encounter distances between them get larger.

constructively over time, just as small pushes to a swing can produce large amplitudes if the pushes are applied at a consistent point in the swing cycle. Many of the objects in the solar system are constrained by gravitational resonances. One classic example is the 3:2 resonance between Neptune and Pluto (Neptune completes three orbits for every two Pluto orbits). *See* RESONANCE (ACOUSTICS AND MECHANICS).

Pluto's heliocentric distance is less than Neptune's for about 20 years out of Pluto's 248-year period. Why have Pluto and Neptune not collided? There is a resonance that prevents Pluto from colliding with Neptune. Whenever Pluto gets close to Neptune, it exchanges angular momentum with Neptune in such a way that it pulls away from Neptune before a collision occurs (**Fig. 6**).

The close approaches between Pluto and Neptune have been exaggerated in Fig. 6. In the present solar system, Neptune never gets closer than 15 AU to Pluto, yet the cyclic exchange of angular momenta still takes place.

Tides, spin-orbit coupling, and synchronous rotation. A nonrotating, planet without mechanical strength in free space would assume a spherical shape under the influence of self-gravity. The same object in orbit around the Sun would have two tidal bulges, one near the Sun, one opposite it. There are two bulges because the different parts of the planet are different distances from the Sun (and hence on slightly different orbits), and would fly apart from each other if it were not for self-gravity keeping the planet together (**Fig. 7**).

The tendency for the near-Sun and anti-Sun components of a planet to diverge is stronger for planets that are closer to the Sun. The Roche limit is the distance from the central body at which self-gravity and tidal dissipation are balanced. Consider Saturn's rings, which lie within Saturn's Roche limit. Although ring particles may collide frequently, there is no chance that they will clump together to form moonlets—they live in a region where tides rip apart any agglomeration of particles. *See* ROCHE LIMIT.

The Earth's Moon always shows the same face to the Earth. This state is called synchronous rotation, and it is common among satellites. Synchronous rotation is caused by spin-orbit coupling, a mechanism which typically slows down an object's spin rate at the same time as it increases the size of the object's orbit (**Fig. 8**).

Close encounters, migrating planets, and the Late Heavy Bombardment. When a planetesimal encounters a planet, the planetesimal either impacts the planet or is gravitationally scattered away from the planet. Unless the planetesimal encounters something that slows it down (for example, gas drag in the neighborhood of the planet), it will approach the planet with speeds equal to or greater than the planet's escape velocity. The planetesimal—if it misses the planet's surface—will approach on one asymptote of a hyperbola and exit on the other. The change in direction depends on how close the planetesimal gets to the planet. Spacecraft often exploit hyperbolic encounters, more commonly called gravity assists. The *New Horizons* mission's gravity assist at Jupiter follows this pattern: the spacecraft approaches Jupiter on one arm of a hyperbola with a closest approach chosen to yield an exit trajectory and velocity that will intercept Pluto, shaving 6 years off a direct Earth-to-Pluto flight plan. *See* SPACE NAVIGATION AND GUIDANCE.

Gravitational close encounters between planets and planetesimals have played a critical role in determining the current structure of the solar sytem.

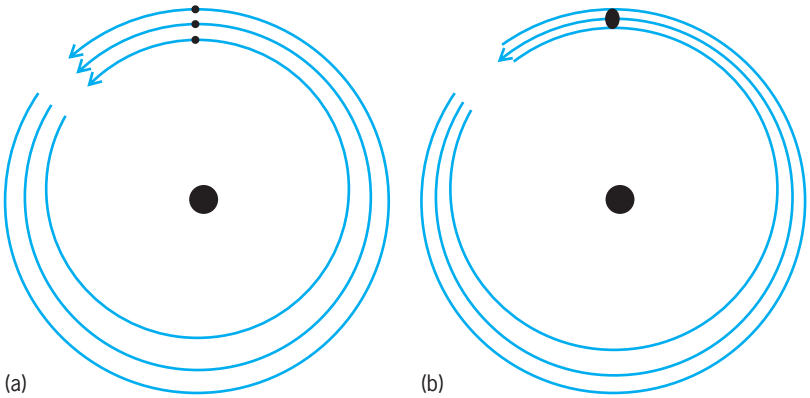

Fig. 7. Effects of tidal forces. (*a*) Imagine three separate point masses at slightly different distances from a central attractor, but all momentarily traveling with the same tangential velocity. The inner and outer point masses will soon diverge from the central point mass: the inner mass will diverge inward, the outer mass will diverge outward. (*b*) If the three points are actually part of a single mass, that mass will become elongated because its constituent parts want to follow diverging trajectories.

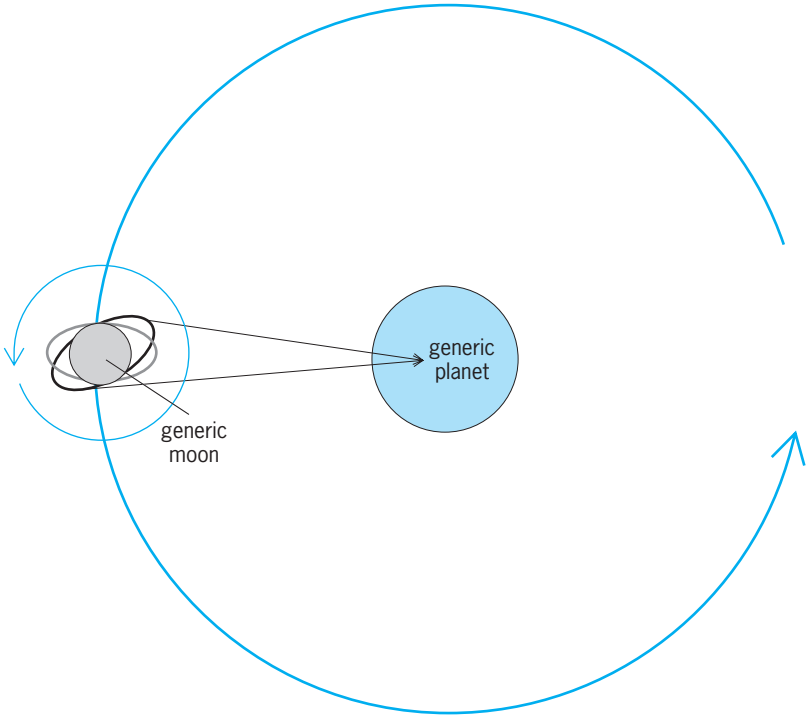

Fig. 8. Bird's-eye view of a hypothetical generic moon in orbit around a generic planet. The planet raises tidal bulges on this moon, but it will be supposed that the moon is not tidally locked; its spin period is a little shorter than its orbital period. Because the deformation of the moon's crust lags behind the instantaneous direction of the tidal disturbances, the planet exerts a torque on the moon which is continually slowing the moon's rotation rate until it is in a synchronous state. (The steady-state lag can be as great as 90° if the spin period is much faster than orbital period. This is an example of forced oscillations with the driving force being much higher in frequency than the deformation time scale of the moon's surface, resulting in a half-period lag in the observed oscillations.)

The planets Neptune, Uranus, and Saturn can scatter planetesimals all over the solar system, but they rarely send planetesimals out of the solar system. Only Jupiter is big enough and dense enough to regularly scatter planetesimals with enough velocity to escape the solar system. As a result, Neptune, Uranus, and Saturn have all migrated outward, and Jupiter has migrated inward. Here is how the migration takes place.

1. When a planet scatters a planetesimal outward, the planet loses a small amount of angular momentum to the planetesimal. Similarly, a planet gains angular momentum when it scatters a planetesimal inward.

2. If no planetesimals escape from the solar system, the net change in each planets' angular momentum will be zero. Scattered objects will eventually have subsequent close encounters with planets, and the average partitioning of angular momentum between each planet and the planetesimals will remain the same.

3. Consider what happens if some of the inward scattered planetesimals are lost from the solar system because they have close encounters with Jupiter. Neptune, Uranus, and Saturn all gain angular momentum from the inward-scattered planetesimals that they never see again. Jupiter loses angular momentun, because it launches planetesimals out of the solar system. Neptune, Uranus, and Saturn all migrate outward, while Jupiter migrates inward.

This migration scenario explains several peculiarities of our solar system. It helps overcome the difficulty in forming Neptune and Uranus at 30 and 18 AU, respectively, which has been a problem given the sparsity of planetesimals at those distances (it was theoretically difficult to form the ice giants in a timely manner). It explains the 3:2 resonance of Neptune and Pluto—Neptune migrated outwards until it encountered Pluto in the 3:2 resonance, then pushed Pluto outward as it migrated further. The migration scenario can even explain the phenomenon known as the Late Heavy Bombardment.

The Late Heavy Bombardment refers to a huge spike in the number of craters formed around 3.9 billion years ago, determined mainly from crater counts on the surface of the Moon. One migration model has initial positions for Neptune, Uranus, Saturn, and Jupiter inside 15 AU, about half of Neptune's present semimajor axis. In this model, Saturn is initially located inside the 1:2 resonance with Jupiter. As the planets scatter planetismals, Jupiter moves inward and Saturn moves outward until they hit their 1:2 resonance (that is, Jupiter makes two orbits for every Saturn orbit). The 1:2 resonance excites the eccentricities of Jupiter and Saturn, not to huge values, but enough to create a sudden increase in the number of planetesimals that are scattered by Jupiter and Saturn. The 1:2 resonance is thought to have occurred 3.9 billion years ago and thought to be the cause of the Late Heavy Bombardment.

Power of light. Although individual photons transfer miniscule amounts of momentum, the continued absorption and reemission of light can measurably affect the trajectories of dust and even asteroids in our solar system.

Radiation pressure refers to the pressure (force per unit area) exerted on objects by incident photons. In the context of the solar system, radiation pressure usually refers to photons emanating from the Sun. Very small particles have high surface area–to–mass ratios. For particles on the order of 0.1 μm, radiation pressure produces a force away from the Sun that exceeds the gravitational attraction toward the Sun. *See* RADIATION PRESSURE.

If an object is moving in a direction that cuts across the photon stream from the Sun, it experiences Poynting-Robertson drag. Just as a moving sailboat encounters an apparent wind that is a combination of the actual wind direction and the sailboat's own motion, an object moving through a radiation field will encounter a small vector component of the radiation that is due to the object's own motion and is always opposite to the object's direction of travel (**Fig. 9**). This tends to slow a particle down and, if the particle is small enough, causes it to spiral into the Sun. For very small particles, however, radiation pressure (which adds angular momentum to a particle orbiting the Sun) will overcome the effects of Poynting-Robertson drag (which takes away angular momentum).

Photons impart mass as well as momentum, but objects typically reradiate photons as blackbody radiation. In the case of small objects (pebbles or dust particles), the entire object is isothermal and emits thermal radiation uniformly in all directions. In the case of large objects, however, there is often a warm side and a cold side. The systematic asymmetry in thermal emission by a large rock or asteroid is called the Yarkovsky effect, and it produces a net thrust that can change the orbital parameters of small asteroids over long periods of time. How does the Yarkovsky effect work? Consider a typical asteroid, spinning on its axis. The sunward hemisphere receives the most energy from the Sun, but the asteroid surface takes some time to heat up and cool down. As a result,

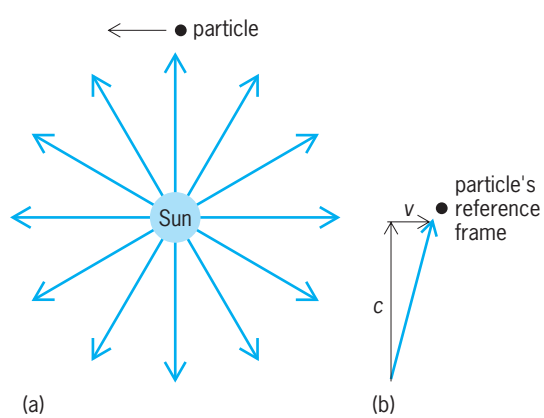

Fig. 9. Poynting-Robertson effect. (*a*) Even though photons from the Sun travel radially outward, (*b*) a particle with a tangential velocity component will perceive that the incident photons all have a head-on component, albeit a small one (proportional to the particle's tangetial velocity, v, over the speed of light, *c*).

the sunward hemisphere emits thermal photons for a period of time after it has rotated away from the Sun. Thermal photons emitted from the asteroid's "evening" side provide a weak but constant thrust toward the "morning" side of the asteroid. *See* HEAT RADIATION.

The net effect of this constant thrust depends on the orientation of an asteroid's pole with respect to the plane of its orbit around the Sun; the Yarkovsky effect can either add to or remove from an asteroid's angular momentum. The Yarkovsky effect has actually been measured for the asteroid 6489 Golevka, a half-kilometer-sized (0.3-mi) asteroid that has been accurately observed by radar since 1991. The force on Golevka from the Yarkovsky effect is only about an ounce (0.28 N), but that force displaced Golevka by 15 km (9 mi) over the period from 1991 to 2003.

Collisions. The most distinctive features on nearly all solid surfaces in the solar system are impact craters. Except for objects that have undergone recent resurfacing, such as the Earth or Europa, all solid surfaces reveal their history though impact craters.

Craters: energy scaling and gravity scaling. During an impact event, the kinetic energy of the impactor is rapidly partitioned into heating the impactor, heating the target, compaction of the target, comminuting the target (breaking it up into a powder or small pieces), and ejecting target material. Small impacts are in the "energy scaling" regime. Most of the energy goes into heating, compacting, or comminuting material. The craters formed in this regime are simple, bowl-shaped craters.

Above a certain threshold, about half of the impactor's energy goes into excavating and launching material from the impact site; these craters are in the "gravity regime." The transition from energy scaling to gravity scaling depends mainly on the acceleration due to gravity at the target's surface. For the Moon (with a surface gravity of 1.6 m/s^2), the transition occurs at 15 km (9 mi). Objects with higher surface gravities will have the energy scaling versus gravity scaling transition occur at smaller crater diameters. For the Earth and Mercury, the surface accelerations due to gravity are 9.8 m/s^2 and 3.7 m/s^2, respectively, and the transition diameters are about 2 and 5 km (1.2 and 3 mi).

Large craters, those formed in the gravity regime, are complex craters (**Fig. 10**). They often have central peaks made from rock that has frozen in midrebound. They have terraces and flat bottoms. The crater sides slumped and filled in the crater floor. The shallow slopes of complex crater walls are often substantially less than the angle of respose, evidence that the postimpact rock is—for a moment—more fluid than sand.

Catastrophic disruption and reaccretion. A catastrophic disruption is defined as a collision in which the largest fragment is less than half the mass of the parent object. Evidence for catastrophic disruption abounds in the solar system. For example:

1. The Earth-Moon system is thought to be the result of a Mars-sized impactor hitting the early Earth. Much debris was ejected into space, and some of that debris (outside the Roche limit) accreted to form the Moon.

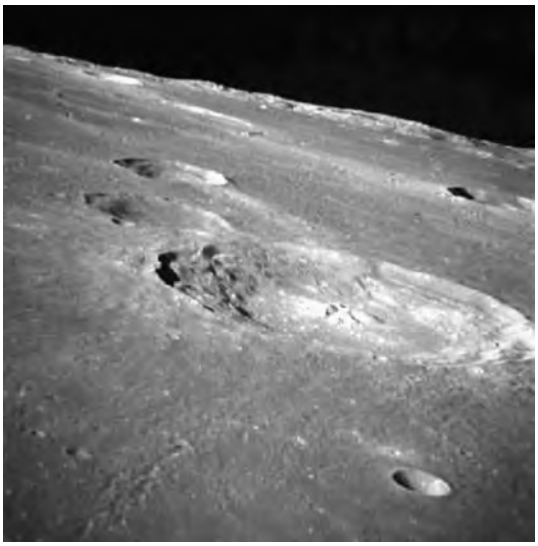

Fig. 10. Complex lunar crater Bullialdus, 60 km (37 mi) in diameter, shows terraced sides, a central peak, and a flat bottom. It may be compared to the small (less than 8 km or 5 mi) crater in the foreground, which is simple and bowl-shaped. (*NASA/Apollo Image Atlas AS16-119-19091*)

2. The planet Uranus spins on its side (its north pole lies almost in the plane of the solar system, whereas most planets' north poles point out of the solar system). This orientation must be the result of a massive impact to the early Uranus.

3. Pluto, like Uranus, spins on its side. In addition, Pluto's three satellites were almost certainly formed from postimpact debris that was orbiting Pluto. All of Pluto's satellites are outside Pluto's Roche limit (as expected, since they would not have reaccreted inside the Roche limit).

4. Asteroid families are groups of asteroids whose orbits indicate that there were once part of the same progenitor. In some cases, the orbits can be traced backward to determine the precise time of the breakup.

5. Meteorites are rocks that hit the Earth from space. There is a class of meteorites which are nearly completely iron. These are thought to have formed in the interior of a object that was large enough to undergo gravitational differentiation. Dense material (iron) settles to the core, and lighter material (rock) forms the the crust. The fact that there are iron meteorites implies that Mars-sized objects roamed the asteroid belt at one time, but underwent catastrophic disruption and their cores became the source for iron meteorites. *See* METEORITE.

Interiors. The interiors of planets are difficult to study directly. Even on Earth, core samples probe less than 1% of the Earth's radius. Seismic sensors can detect compression and transverse waves that propagate through an object. The speed and direction of these waves depend on density and density gradients. Sensors can detect waves that are reflected off interfaces (changes in state or composition within the planet). The Apollo and Viking missions deployed seismic sensors on the Moon

and Mars, but the Viking seismic experiments were contaminated by buffeting from Martian winds. For bodies besides the Earth and Moon, we must observe spin rates, oblateness, and gravitational fields to determine the distribution of mass in the planets' interiors. *See* SEISMOLOGY.

Moments of inertia and distribution of mass. A spherically symmetric planet has a spherically symmetric gravitational field. Most planets, however, are oblate—they are a little wider around the middle than they are from pole to pole. Saturn is the most oblate, with an equatorial radius that is 10% larger than its polar radius. As a result, Saturn's gravitational field is a little stronger in the plane of Saturn's equator than above either of Saturn's poles (at a given distance from the planet).

The gravitational field of an oblate planet depends on its distribution of mass. If the planet has a dense core, then less mass is distributed around its equator. The gravitational field around a hypothetical "dense-core" planet will be more spherical than for a "uniform-density" planet of the same oblateness. How does one measure the gravitational field in the neighborhood of a planet? Typically one needs an object near the planet—something whose course is measurably affected by its gravitational attraction to the planet, like a natural moon or an artificial spacecraft.

If we assume that a planet's interior is in hydrostatic equilibrium (that inward-pointing forces due to gravity are balanced by outward-pointing pressure gradient forces), then one can estimate the radial distribution of mass from two observables: an object's spin rate and its oblateness. To illustrate, consider a nonrotating, Jupiter-size planet. It would be a giant onion of spherical layers, with each layer growing denser toward the center of the planet. The density profile can be modeled—including changes of state (such as the onset of metallic hydrogen) where the pressure and temperature are high enough—but cannot be confirmed if the planet is spherical. Once we start spinning this gas giant, we can calculate the balance between self-gravity and centripetal acceleration at each point and compare the predicted shape of the planet to the oblateness we observe. In the case of Jupiter and Saturn, their oblatenesses would be even more extreme if they were composed of hydrogen and helium to their cores. We deduce that Jupiter and Saturn have dense, rocky cores (about 10–20 Earth masses in the case of Jupiter) from the fact that each of the objects is less oblate than they would be if they had no rocky cores.

Asteroids and small moons are generally not in hydrostatic equilibrium because their intrinsic strength prevails over self-gravity. The radius at which self-gravity overcomes intrinsic strength is roughly 200–400 km (125–250 mi).

Polytropes. Jupiter and Saturn are nearly the same size, but Jupiter is nearly three times as massive as Saturn. Both have similar compositions—why such a difference in densities? As a thought experiment, consider a small planet made of hydrogen to which we slowly add more and more hydrogen. For very small planets, the radius will scale with the cube root of the mass, since the density will be roughly constant in this regime. As the mass of the planet increases, however, the density of the core increases, and the radius will be proportional to the mass raised to a power less than 1/3. Eventually, the compression is so great that the addition of more hydrogen simply increases the average density, and the radius is unchanged. This is the case for hydrogen-based planets between 100 and 1000 Earth masses in size, including Saturn and Jupiter. Objects in this mass range have nearly constant radii, independent of their mass. The addition of even more hydrogen (beyond 1000 Earth masses) actually shrinks the radius; this is the regime of white dwarves and eventually black holes or neutron stars.

A polytrope is a hypothetical object in which the pressure is proportional to the density raised to a power. This power is usually written as $1 + 1/n$, where n is the polytropic index. For a hydrogen-dominated object like Jupiter, $n = 1$ is a reasonable approximation to its equation of state. Jupiter's radius is very close to the mass-independent radius of an $n = 1$ polytrope, but Saturn's radius is smaller. This is evidence that Saturn contains more elements (relative to Jupiter) that are heavier than hydrogen. *See* POLYTROPIC PROCESS.

Composition and phase changes. A polytrope is a useful approximation for examining the interiors of Jupiter and Saturn. In particular, it helps us determine where phase changes occur within the gas giants. At a pressure of about 2 Mbar (2×10^{11} N/m^2), hydrogen gas undergoes a phase change to state called metallic hydrogen. We can estimate the pressure from the density, and the density of an $n = 1$ polytrope has a radial dependence proportional to $\sin(Ar)/Ar$, where A is a constant and r is the radius. For Jupiter, we find that the pressure exceeds 2 Mbar inside 80% of Jupiter's radius, which means that just over half of Jupiter (by volume) is its metallic hydrogen core. For Saturn, the metallic transition takes place at the 50% radius, so Saturn's metallic hydrogen core occupies about 1/8 of its total volume.

Magnetic fields. Some planets have interiors that are both fluid and conducting. For the Earth and Venus, the interiors are composed of iron alloys that are denser than silicates (rocks) and have collected to form iron-rich cores. For Jupiter and Saturn, their metallic hydrogen cores are conductors. (Metallic hydrogen is so dense that some electrons are not bound to a proton but are free to move throughout the medium.) A body with a conducting fluid interior is a candidate for having dynamo.

A dynamo is mechanism that converts mechanical energy into electrical energy, including the specific example of a convecting planetary interior generating electrical currents and magnetic fields. The quantitative details of how dynamos work are an ongoing subject of study, but here are some general principles:

1. A planetary dynamo can exist when a conducting fluid moves across magnetic field lines. The magnetic field induces current loops in the fluid, which in turn generates a new magnetic field that is added to the ambient magnetic field. The process is self-exciting; all that is required is a sufficiently

conductive fluid in motion. *See* FARADAY'S LAW OF INDUCTION.

2. Planetary magnetic fields require energy sources, otherwise they decay. For the Earth, the free decay time scale of the core's magnetic field (in the absence of a dynamo) is on the order of 10,000 years. The implication is that the Earth must be generating its magnetic field; a primordial field would not last long. The free decay time scale is much shorter than the age of the solar system for all of the planets.

3. Dynamos can occur when the magnetic Reynolds number (equal to the product of the typical velocity and length scale of the motion in the conducting fluid, divided by the magnetic diffusivity) is greater than 10 or 100 and when fluid motions have certain characteristics. These fluid motions could likely be convection in the presence of a sufficiently large Coriolis force. *See* CONVECTION (HEAT); CORIOLIS ACCELERATION.

4. Contrary to what is written in many textbooks, Venus's slow rotation is not an explanation for its lack of a dynamo. Venus's 243-day period is sufficient to generate the necessary Coriolis forces. In fact, for a given temperature gradient, slower rotation rates lead to faster convective velocities. It is possible that Venus's liquid core is not undergoing convection, and that may be the reason that Venus has no dynamo.

5. The electrical conductivity of the fluid may not be high (although good conductors make the most efficient dynamos). Because high electrical conductivity correlates with high thermal conductivity, a liquid core with high electrical conductivity might cool by conduction as opposed to convection. It appears that terrestrial planets are close to the conduction-versus-convection threshold (that is, the Earth is barely in the convective regime). Gas giants are securely in the convective regime.

Aurorae. Charged particles (electrons, protons, ions) are constrained to move along magnetic field lines. To first order, the planetary magnetic fields are all dominated by a dipole term (that is, they look similar to the magnetic fields produced by bar magnets), although Neptune's dipole is offset from the center of the planet. Charged particles inside a planet's magnetosphere are guided along magnetic field lines until they intersect the planet's atmosphere, often near the north and south magnetic poles. Atoms in the atmosphere are then ionized and/or excited by collisions with the incoming charged particles. The excited atoms emit the auroral light when they return to their normal energy levels.

Since the source of charged particles is typically larger than the planet (on Earth, the solar wind is the main source of charged particles), the auroras are produced in symmetric pairs, aurora borealis and aurora australis, at both magnetic poles. On Jupiter, the source is particles trapped in the Jovian magnetosphere as opposed to the solar wind. Some of these particles are electrons that are sputtered or ejected off Jupiter's satellites, including Io, Ganymede, and Europa. The particles follow field lines all the way from the satellite surfaces to specific locations on Jupiter. Each satellite produces its own "auroral footprint" within Jupiter's broader auroral ring. The footprints show up as bright dots at the points where Jupiter's magnetic field lines from the satellites intersect Jupiter's atmosphere (**Fig. 11**). *See* AURORA; SOLAR WIND.

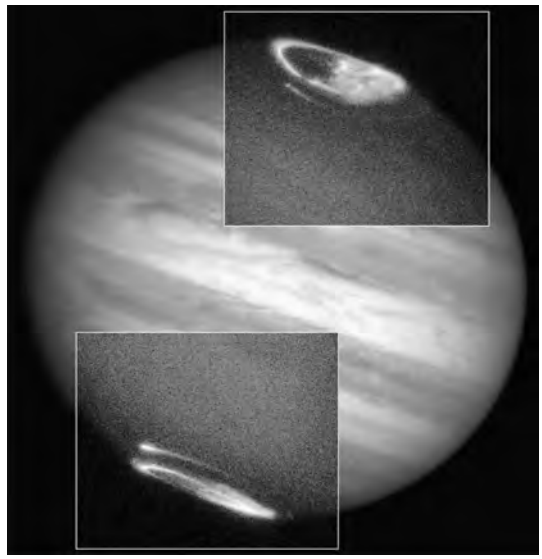

Fig. 11. Ultraviolet images from the *Hubble Space Telescope* of Jupiter's poles (superimposed on a visible-light image of the planet for context) show the mirror-image north and south polar aurorae on Jupiter. In this image the "Io footprints" are clearly visible. These are the white comet-shaped streaks that are visible outside the circular auroral curtains. They are caused by particles that were ejected from volcanoes on Io and traveled along magnetic field lines until they intersected Jupiter's atmosphere. (*John Clarke, University of Michigan; NASA*)

Surface-atmosphere interactions. There is significant coupling between the solid surfaces in the solar system and any atmosphere above those surfaces. On Earth, for example, the availability of surface water leads to atmospheric clouds, affects the amount of absorbed and reemited sunlight, and controls (in part) vertical temperature profiles.

Atmospheres supported by vapor-pressure equilibrium: Pluto, Mars. The Earth's atmosphere is mainly N_2, with oxygen, CO_2, water vapor, and other trace gases. The nitrogen gas will never condense onto the Earth's surface—it is far too warm for that. On Mars, however, the dominant gas is CO_2, which does condense onto the Martian surface. In fact, Mars experiences a global freeze-out of about 25% of its atmosphere when it is farthest from the Sun. The Martian column abundance is tied closely to the average temperature of the CO_2 polar caps. A small rise in the polar frost temperatures results in an increase in the bulk Martian atmosphere.

Pluto, like the Earth, has an atmosphere that is primarily composed of N_2. Unlike the Earth, the N_2 on Pluto regularly condenses onto its cold surface (about 40 K or −388°F). The nitrogen gas in Pluto's atmosphere maintains a nearly uniform surface temperature across the frost-covered regions of the entire planet through the latent heat of sublimation or condensation.

Parts of Pluto are often in perpetual darkness or sunlight for decades. How does the surface/atmosphere interaction of Pluto's N_2 gas force the

sunlit and dark regions to have nearly the same surface temperature? At any moment, half of Pluto is in sunlight, half is in shadow. As the sunlit hemisphere begins to warm, the vapor pressure of the nitrogen frost on the surface begins to rise. (Vapor pressure refers to the steady-state gas pressure that is maintained above a solid.) The vapor pressure of nitrogen (and most volatile ices) is extremely sensitive to changes in temperature—an increase of 1.5° K results in a doubling of the N_2 vapor pressure. As the temperature starts to rise over Pluto's sunny side, the local atmospheric pressure goes up and leads to winds from the subsolar point toward the dark side of the planet. Pluto's sunlit side is cooled by sublimation, the shaded side is warmed by the latent heat of condensation, and the entire frost-covered surface is maintained at nearly a uniform surface temperature. A temperature differential of even a fraction of a degree would lead to enormous, unsustainable pressure gradients. *See* VAPOR PRESSURE.

Greenhouse and anti-greenhouse effects. The "greenhouse effect" is a mechanism that can change the atmospheric temperature near a planet's surface. The analogy to ordinary glass greenhouses is a poor one: glass greenhouses trap heat by preventing warmed air from rising, whereas the greenhouse effect is due to the atmosphere's transmission of energy at one wavelength range and its opacity at another.

In the steady state, a planet reradiates the same amount of energy as it receives from the Sun. (The gas giants actually reradiate 2–3 times more energy than they receive from the Sun, essentially residual heat from their accretion stages. Terrestrial planets—even ones dominated by molten cores—lose heat so slowly by conduction through their solid crusts that they are basically in energy balance with the Sun.) A terrestrial surface will, on average, heat up until its heat output (its blackbody radiation) balances the heat input (the absorbed solar flux).

Greenhouse gases (like water vapor, carbon dioxide, and methane) are transparent in visible wavelengths where the solar spectrum has the most power. Like many molecules, however, these gases have significant absorption features at infrared wavelengths where a planet's surface is reemitting the bulk of its thermal photons. Since the density of greenhouse gases is typically highest near the surface, the greenhouse effect usually traps heat in the lowest layers of the atmosphere. On Earth, the greenhouse effect raises the surface temperature by about 40 K (72°F). The thin carbon dioxide atmosphere of Mars is about a 5 K (9°F) effect, while Venus's massive carbon dioxide atmosphere raises its surface temperature by about 500 K (900°F).

If the greenhouse effect makes a planet's surface hotter, why does it not reradiate more thermal photons to space and cool down? The warmer surface does radiate more heat, but the greenhouse gases prevent a significant fraction of those photons from escaping to space. The effective temperature of a planet, T_{eff}, is the temperature of the average layer that is visible from space, not necessarily the temperature of the surface.

The anti-greenhouse effect occurs in atmospheres that have bright particulates, like smog or haze particles or aerosols from a volcanic eruption. Many of these particles have small sizes, perhaps on the order of 0.1 to 0.5 μm. These particles are approximated as Rayleigh scattererers: their scattering efficiency is inversely proportional to the fourth power of the wavelength. At short wavelengths, small aerosols are very bright, very effective scatterers. At longer wavelengths, their scattering efficiency drops dramatically, and the aerosols are nearly transparent. Aerosols have the opposite transmissive properties from greenhouse gases: aerosols block visible light from the Sun from reaching a planet's surface, but they allow infrared light to escape to space. The famous "Year without a Summer" (1816) is thought to have been caused by the anti-greenhouse particulates spewed by Mount Tambora during its volcanic eruptions in April 1815. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

Atmospheric structure. In an isothermal atmosphere, density and pressure are both exponential functions of altitude. To see why this is so, consider a very thin layer in an isothermal atmosphere. The change in pressure across that thin layer is proportional to the mass per unit area in that layer (that is, the pressure is greater at the bottom of the layer because of the additional mass of the layer pressing down). The mass per unit area is proportional to the density, which (according to the ideal gas law at constant temperature) is proportional to the pressure. We see that the change in pressure across a thin layer is proportional to the pressure itself at that layer. This relationship (the rate of change being proportional to the variable itself) is the hallmark of an exponential.

Thermal profile. In real atmospheres, of course, there are changes in temperature, composition, and phase changes with respect to altitude, but to first order, every density profile in the solar system is essentially an exponential function. Can we say anything concrete about the expected themal profiles? Consider a parcel of air that we move slowly from the bottom of the atmosphere to a higher level. An adiabatic process is one in which no heat is exchanged. We will assume that the parcel exchanges no heat with its surroundings. Under that assumption, the parcel has to conserve energy within itself, which means that changes in the parcel's gravitational potential energy have to be offset by changes in the parcel's thermal energy (the internal kinetic energy of its molecules). *See* ADIABATIC PROCESS.

If we lift a parcel of mass m by a distance h against a gravitational acceleration g, the change in gravitational potential energy will be $P = mgh$. The equivalent change in thermal energy is $\Delta T m c_p$, where c_p is the specific heat of the parcel (the amount of energy needed to raise the parcel by 1 K at constant pressure) and ΔT is the parcel's change in temperature as it is raised adiabatically by a height h. We set the sum of the energy changes to zero (conservation of energy) and solve for $\Delta T/h$, which is simply $-g/c_p$. This quantity is called the adiabatic lapse rate, or the rate at which temperature changes with respect to altitude under adiabatic conditions.

Regardless of whether the adiabatic assumption is valid, the adiabatic lapse rate is a useful quantity. On Earth, $-g/c_p = -9.8$ K/km (that is, if you climb 1 km up a mountain, expect the temperature to be about 10 K colder), or $-28.4°$F/mi. If the actual lapse rate is greater that the adiabatic lapse, then the atmosphere is stable. If it is more negative than the adiabatic lapse rate, then the atmosphere is unstable and convection will often take place; these lapse rates are called superadiabatic or autoconvective. There are situations (in the Southwestern United States on clear summer days, for example) where the ground can get so hot that the local temperature gradient is superadiabatic.

Our simplistic energy conservation discussion of the adiabatic lapse rate neglected the energy associated with phase changes. If the parcel of air contained water vapor, for example, then some of the energy associated with raising or lowering the parcel might be balanced by the condensation or evaporation of water instead of the change in temperature of the parcel. This lapse rate is called the wet adiabatic lapse rate (as opposed to the dry adiabatic lapse rate for parcels devoid of condensable species). It is closer to zero than the dry lapse rate. *See* ATMOSPHERE; METEOROLOGY.

Escape processes and atmospheric upper boundaries. We expect the atmospheric pressure to decrease at higher altitudes. There is no sharp boundary between the atmosphere and space; the density and collision rates simply decrease to the point where individual molecules are on ballistic trajectories (in orbit around the planet) or they are escaping. A useful term is the exobase, the altitude above which molecules are more likely to escape than to collide with another molecule.

We can look at escape rates to determine which planetary bodies are capable of retaining an atmosphere. The Maxwell distribution (also called the Maxwell-Boltzmann distribution) describes the velocity distribution of molecules in a gas at a given temperature. The peak of the Maxwell distribution is sometimes called the thermal velocity; it represents the most common speed of all the molecules in a sample. In general, a planet is capable of retaining an atmosphere if the thermal velocity is smaller than the escape velocity. The Maxwell distribution has a long tail of fast-moving molecules, which means that the escape velocity ν_{esc} needs to be several times higher than the thermal velocity ν_{TH} to prevent the loss of an atmosphere over time. In practice, ν_{esc} needs to be about five times greater than ν_{TH} to retain an atmosphere over the age of the solar system. The two velocities are given by

$$\nu_{esc} = (2GM/R)^{1/2}$$

where G = gravitational constant, M = mass of planet, and R = radius of planet, and

$$\nu_{TH} = (2kT/m)^{1/2}$$

where k = Boltzmann's constant, T = temperature, and m = mass of particle.

The expression for the thermal velocity is inversely proportional to the square root of the particle's mass. Lighter particles have faster velocities than heavier ones: hydrogen molecules, for example, will have average velocities $\sqrt{14}$ times higher than nitrogen molecules. *See* BOLTZMANN STATISTICS; KINETIC THEORY OF MATTER; STATISTICAL MECHANICS.

Eliot Young

Bibliography. J. K. Beatty et al. (eds.), *The New Solar System*, 4th ed., Sky Publishing, Cambridge, MA, 1999; J. A. Burns, P. L. Lamy, and S. Soter, Radiation forces on small particles in the solar system, *Icarus*, 40:1, 1979; I. de Pater and J. L Lissauer, *Planetary Sciences*, Cambridge University Press, 2001; H. E. Levison et al., Planet migration in planetesimal disks, in B. Reipurth, D. Jewitt, and K. Keil (eds.), *Protostars and Planets V*, University of Arizona Press, Tucson, 2006; J. S. Lewis, *Physics and Chemistry of the Solar System*, 2d ed., Academic Press, 2004; N. McBride and I. Gilmour, *An Introduction to the Solar System*, Cambridge University Press, 2004; L. McFadden, P. R. Weissman, and T. V. Johnson (eds.), *Encyclopedia of the Solar System*, 2d ed., Academic Press, San Diego, 2006; C. D. Murray and S. F. Dermott, *Solar System Dynamics*, Cambridge University Press, 1999; M. A. Seeds, *The Solar System*, Thomson Brooks/Cole, Belmont, CA, 2006; J. H. Shirley and R. W. Fairbridge (eds.), *Encyclopedia of Planetary Sciences*, Chapman & Hall, London, 1997; D. J. Stevenson, *Earth Planet. Sci. Lett.*, 208:1, 2003.

Planetarium

An instrument that projects the stars, Sun, Moon, planets, and other celestial objects upon a large hemispherical dome, showing their motions as viewed from the Earth or space near the Earth. Days and years may be compressed into minutes. There are over 100 major planetariums around the world with domes 50 ft (15 m) or more in diameter; and there are also over 1000 smaller planetariums in communities, schools, and colleges.

The term planetarium originally applied to a mechanical model (also known as an orrery) that depicted the motions of the planets. Today the term refers to an optical projector. Most planetariums now have mechanical movements, but planetariums projecting computer-generated displays have also been developed. Additional optical devices and computer controls are common. The term planetarium also refers to the theater or building that houses the projector.

Development. Models of the sky date from early times. A celestial globe was made by Claudius Ptolemy of Alexandria around A.D. 150. The Gottorp Globe, a 10-ft (3-m) globe built in the midseventeenth century, is hollow, and up to 10 people may sit inside to see the sky which is painted on the interior. This globe is on display at the Lomonosov Museum, St. Petersburg, Russia. A larger globe, constructed in 1912, and which is on display at the Chicago Academy of Sciences, is 15 ft (4.5 m) in diameter and seats 17 spectators. The metal sphere has 692 holes to depict the stars.

Motions of the Earth, Moon, and planets have also

Fig. 1. Planetarium projectors. (*a*) Skymaster ZKP-3. (*b*) Starmaster. (*c*) Universarium M-IX. (*d*) Cosmoleap. (*e*) Geministar II. (*f*) Infinium αII. (*g*) GOTO Planetarium Projector GS. (*h*) GOTO Space Simulator Chronos. (*i*) GOTO Uranus. (*j*) GOTO Space Simulator S Hellos. (*Models a-c from Carl Zeiss Company, d-f from Konica Minolta Corp., g-j from GOTO*)

been represented from early times with various mechanical models. About 1682, C. Huygens designed an elaborate model that showed the planets out to Saturn. Around 1712, J. Rowley built the original orrery, in which, by turning a crank, the Moon's motion around the Earth could be seen as well as the Earth's motion around the Sun, thus explaining the lunar phases and the Earth's seasonal relations. *See* EARTH ROTATION AND ORBITAL MOTION; MOON; PLANET.

The invention of the projection planetarium by W. Bauersfeld in 1919 solved the problem of presenting stars and planetary motions in a realistic fashion in a domed theater. The first such planetarium, built by the Carl Zeiss Company of Germany, opened in Munich in 1923.

Types of projectors. Many projectors are patterned after the basic design of the Carl Zeiss Company. Star spheres at each end of the projector show

8900 stars down to magnitude 6.5; 32 lenses (16 located in each globe) are used to project the stars. Cages between the two star spheres contain projectors for the Sun, Moon, and planets. The center part of the machine houses the driving motors. Additional projectors show such effects as variable stars, solar and lunar eclipses, the Milky Way, comets, and various circles and coordinates. Depending on the manufacturer, there may be constellation outlines, clouds, and built-in zoom effects for the planets.

Of the 25 large-model projectors built by Zeiss prior to 1940, several are still in service and have been joined by many new models. Zeiss offers three models for various dome diameters. These are the Skymaster ZKP-3, the Starmaster ZMP, and the Universarium M-IX for dome diameters of 16–39 ft (5–12 m), 40–60 ft (12–18 m), and 60–100 ft (18–30 m), respectively (**Fig. 1***a–c*). The Skymaster, the Starmaster, and the Universarium project 6300, 8900, and over 9000 stars, respectively. All are computer-controlled and user-friendly. Both the Starmaster and the Universarium use fiber optics to increase star brilliance. A device also makes the stars twinkle.

Konica Minolta planetariums also come in several models: the Cosmoleap 8 and 10 for domes ranging 25–40 ft (7.5–12 m), the Geministar II for domes 44–66 ft (14–20 m), and the Infinium αII, βII, γII for domes 40–88 ft (12–27 m) [Fig. 1*d–f*]. The Cosmoleap series project from 3900 to 6700 stars, the Geministar II 10,000 to 15,000 stars, and the Infinium series 9200 to 26,000 stars. All series are computer-controlled. The Gerministar II is a starball utilizing fiber optics and the option of integration with a Digistar II CG projector from Evans & Sutherland. Konica Minolta also offers an all-sky CG (computer graphics) projection system for smaller domes of 10–20 ft (3–6 m). Called the Mediaglobe, the unit features a full-color fish-eye video projector with high resolution of Sun, Moon, planets, stars, constellation patterns, comets, and meteors. For larger domes, Konica Minolta offers Laserscan and Skymax (digital dome imaging). [See "Full dome imaging" on page 658.]

GOTO planetariums offer various models, including the GE II and GS for domes 20–27 ft (6–8.5 m) and projecting 2800 to 3500 stars; the Chronos for domes 26–40 ft (8–12 m) and projecting 8500 stars; the Uranus for domes 33–60 ft (10–18 m) and projecting 10,000 stars; the Super-Helios for domes 59–164 ft (18–50 m) and projecting 38,000 stars [Fig. 1*g–j*]. Each of the planetarium projectors of various manufacturers may be mounted on a hydraulic hoist, permitting it to descend below the floor for parts of the program.

Another design philosophy, begun in the late 1940s when A. Spitz designed a small planetarium for school classrooms and museums, was to manufacture a small and relatively inexpensive projector that would do for schools and small communities what the larger machines had done for the cities. Early models of Spitz Space Systems planetariums featured planetary motions through innovative planet analog mechanisms. One of these models became the most common projector for 24–30-ft (7–9-m) domes, while another was manufactured for dome diameters of 50 ft (15 m). Later models include System 1024, a medium-sized projector designed for the 24–58-ft (7–17-m) dome size. Spitz's Star Scape, designed for major theaters, has a starball which is 4 ft (1.2 m) in diameter and projects over 10,000 stars for dome diameters of 60–88 ft (18–27 m) [**Fig. 2***a–b*].

The Zeiss Universarium, the Minolta Infinium, the GOTO Super Helios, and the Spitz Star Scape are usually installed in tipped domes which may include all-sky 70mm motion picture projection systems. Another innovation in planetarium design, based on computer graphics television projection, is the Evans & Sutherland Digistar 3 SP (**Fig. 3**). The projector

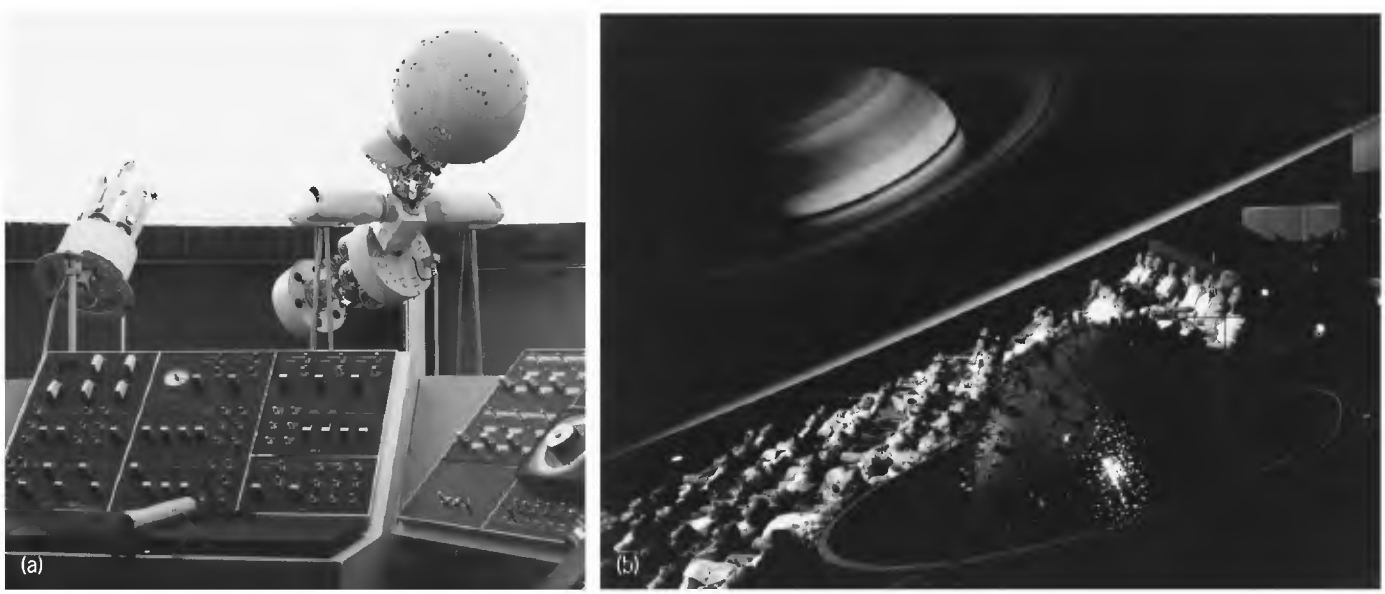

Fig. 2. Spitz planetariums. (a) System 1024 (*Spitz Space Systems, Inc.*). (*b*) Star Scape in a theater that has a 76-ft (23-m) dome tipped at a 26° angle and seats 344 people (*R. H. Fleet Space Theater, Balboa Park, San Diego*)

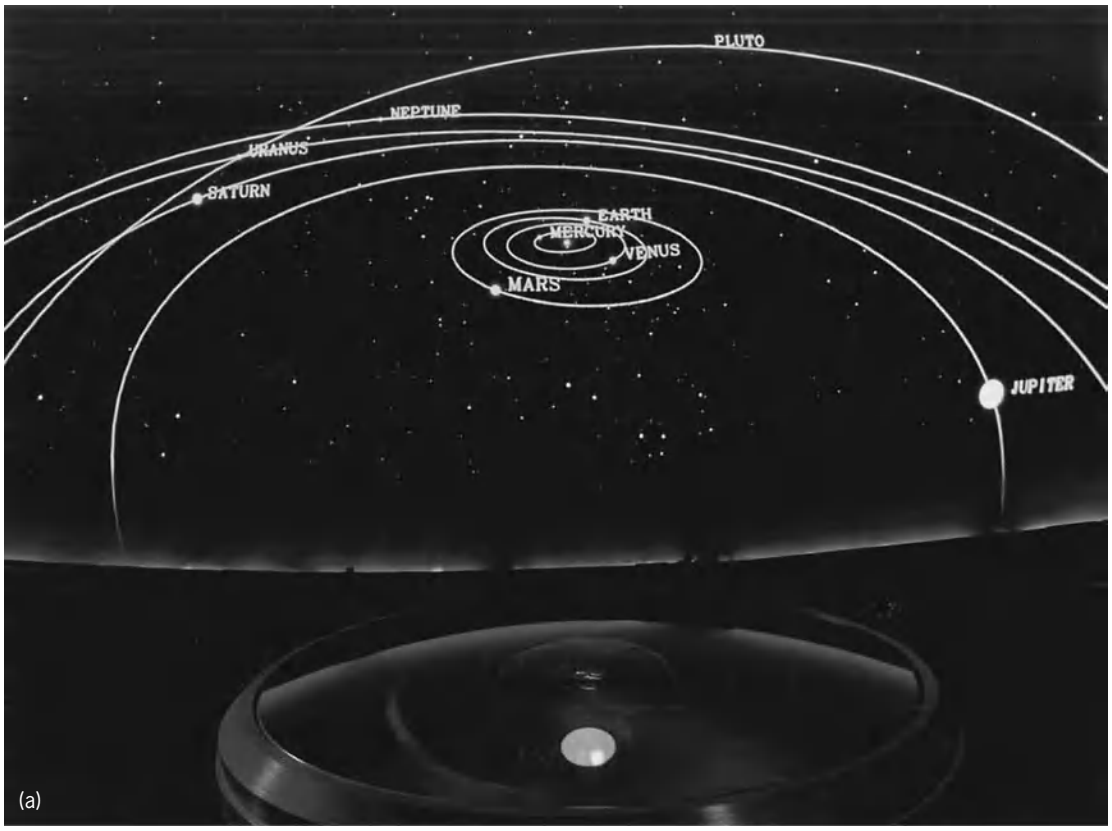
PLUTO
NEPTUNE
URANUS
SATURN
EARTH
MERCURY
VENUS
MARS
JUPITER
(a)

Digistar 3
(b)

(c)

(d)

(e)

utilizes a high-resolution cathode-ray tube with a special 160° wide-angle lens for projection onto domes up to a diameter of 33 ft (10 m). The planetarium may have a horizontal or tilted dome. Software and documentation include stellar, planetary, and constellation data files. The star and planet positions are fed into the high-intensity cathode-ray tube projector by a computer and thence onto the dome via a wide-angle lens.

Evans & Sutherland also offers the Digistar 3 for domes up to 60 ft (18 m) in diameter. Six high-resolution/brightness projectors are used providing all-dome coverage. Six-channel audio sound is also computer controlled to "follow" the audio source. All dome pictures can be displayed as well as stars, Sun, Moon, Planets, etc. Konica Minolta has developed Mediaglobe CG projectors for smaller domes and Skymax systems for larger domes. Spitz has a full-dome digital planetarium called SciDome utilizing Starry Night™ as software. Carl Zeiss offers all-dome laser projection in full color. *See* CATHODE-RAY TUBE; COMPUTER GRAPHICS.

Mechanical aspects. The Digistar 3 and 3 SP have no moving parts, whereas the other projectors described above all feature similar mechanical operation derived from the basic Zeiss design. These mechanical-optical projectors all show the apparent motions of the sky, which are, in reality, a reflection of the Earth's motions, or the motion of the observer. These mechanical projectors are geocentric, that is, they show the sky from an Earth-centered viewpoint. For example, daily motion depicts the turning of the sky (rising and setting of the Sun, Moon, planets, and stars) caused by the Earth's rotation on its axis. A variable-speed motor turns the entire projector. Latitude motion shows the changing aspect of the sky as the observer travels north or south over the Earth. Thus as one moves toward the North Pole, the North Star rises higher. This is accomplished by rotation of the entire projector around a horizontal (east-west) axis.

Annual motion uses auxiliary projectors to the star globe to depict the changing monthly phases of the Moon, the eastward motion of the Sun along the ecliptic, as well as the annual (yearly) motion of the planets in the zodiac. More complicated mechanical linkages are required to represent these motions, because of the more complex mathematical ratios of planetary periods. Numerous gears are required to produce the desired accuracy. Some instruments also approximate the variable planetary speeds. To project a planet, for example, Mars, in its correct position in the zodiac, requires not only a gear train to represent the period of Mars but also a gear train to represent the Earth's position (because the planet is observed from the moving Earth). A sliding rod links the Earth drive to the planet drive, which in turn aligns the planet projector in the correct direction in space. The Moon-phase projector not only shows where the Moon is among the stars each hour but also the correct phase. The projector also depicts the regression of the lunar nodes. Some planetariums also have built-in solar and lunar eclipse projectors as well as planet zoom effects. *See* GEAR TRAIN; LINKAGE (MECHANISM).

In addition to the rotation of the Earth upon its axis and its annual journey around the Sun, the Earth has a wobbling motion much like that of a top. The wobbling motion of the Earth's axis, called precession, is caused by the Moon and the Sun pulling on the equatorial bulge of the Earth. The Earth takes 25,800 years for one precession cycle. Planetariums reproduce this effect by rotating the projector around an axis which is inclined 23.5°; to the daily-motion axis, to show how the position of the pole star changes as well as the movement of zodiacal constellations in relation to the equinox. *See* PRECESSION OF EQUINOXES.

The Spitz system 1024 also utilizes mechanical analogs to reproduce the aforementioned motions. It does not follow the basic Zeiss design, however, because of the necessity for miniaturization. The daily, latitude, and precession motions are similar to those of the Zeiss, but the annual motion for the planets is based upon the Tychonic system rather than the Copernican or Keplerian. All Spitz projectors feature the ATM-4 Theater Auto motion System, which provides the automation of independent projection systems.

For the Spitz Star Scape Star Projector, mechanical linkages are replaced by computer control. The majority of projectors of the various manufacturers are computer-controlled. The "high-end" projectors that appear as large "starballs" are actually two hemispheres attached to a collar assembly that contains the servo motors. For example, the Zeiss Universarium, the Konica Minolta Infinium, and the GOTO Super Helios utilize separate zoom projectors to show the Sun, the Moon, and the planets (Fig. 1*b*,*c*,*e*,*f*,*i*,*j*). These projectors are located next to the base of the starball and are controlled by computer rather than mechanical linkages. Several manufacturers offer multiple images for these projectors—for example, comets, asteroids, and solar and lunar eclipses—in addition to the usual planets. These computer-controlled multiaxis projectors show planetary and solar motions not only from the Earth's surface but from any viewpoint in space near the solar system. For example, it is possible to view the planetary motions from Mars. The GOTO Chronos also has built-in x-y mirrors for planet motions that are computer controlled (Fig. 1*b*).

The Digistar 3 and 3 SP is also fully computer-controlled with no moving parts; it is not limited to geocentric projection, but may depict space views out to 650 light-years from the Sun (a limitation governed by knowledge of accurate star positions). The effect of a three-dimensional journey through our region of space may be simulated; for example, the

Fig. 3. Planetarium projector systems. (*a*) Digistar 3 computer graphics television projection planetarium. The projection lens is in the foreground, with dome projection above. (*b*) Digistar 3 SP projector. (*c*) Section of Digistar 3 SP theater. (*d*) Three Digistar 3 projectors. (*e*) Back view of Digistar 3 theater with image on screen. (*Evans & Sutherland*)

viewer can travel toward and then through the stars of Orion (Fig. 3).

Full dome imaging. We have discussed types of projectors used for planetariums that show the night sky: stars, Sun, Moon, planets, comets, etc. For many decades, planetariums have also incorporated into their projection displays a variety of special-effect projectors. These devices, often fabricated in-house as well as by outside vendors, have been used to enhance the dome displays. Such effects as panoramas, rotating galaxies, zoom effects to celestial objects, and exploding stars are common auxiliaries used un programs. The various vendors of planetarium projectors now offer a variety of full dome imaging systems. Here we list some of these by manufacturer, keeping in mind that additional vendors may offer similar products.

Carl Zeiss Jena GmbH. In ADLIP (All-Dome Laser Image Projection), six laser beams flood the planetarium dome in full color with any image that is fed into the system. It may be used in real-time or playback mode.

The ZULIP (Zeiss Universal Laser Images Projector) is a single laser projector, near the center of the dome on a slender lift incorporating a two-axis mount and motorized wide-angle zoom lens, that "writes" moving pictures on the dome in full color. The area of display is controlled in real-time or playback mode and may point in any direction.

Evans & Sutherland. The Digistar 3 and 3 SP are all-dome video systems of high resolution in full color for various-diameter domes. E&S also has a library of prerecorded planetarium programs made available to their customers. These projectors have been described elsewhere in this article.

GOTO. The Virtium II is an all-dome video system of high resolution in full-color display of computer-generated 3D images controlled by high-speed PC-type computers in real-time or playback mode. The Digicanvas uses up to four video sources to fill approximately half of a dome in full color.

The Ultra 70 and Astrovision 70 are 70/35mm film projectors using wide-angle lenses for various-sized domes. GOTO also has an extensive library of artworks, slides, films, and videos for the planetarium customer.

Konica Minolta. The Midiaglobe for smaller domes [10–20 ft (3–6 m)] and the Skymax for larger domes are video projection systems. The Mediaglobe features a 160° fisheye projection lens fed by an SXGA liquid-crystal color source in full color. The Skymax system for larger domes uses three to six multiscan video projectors for full-color all-sky coverage. The Laserscan is a full-color all-sky laser projection system. Konica Minolta also has a support program with a library of software for the various projection systems.

Spitz Inc. The Electricsky II and the Scidome are single-projection hemispheric video projectors. The eSky II is for domes up to 60 ft (18 m) and the SCIdome for up to 40 ft (12 m) in diameter. The eSKY II features a QXGA (2048 × 1536)–resolution video projector with fisheye lens. Playback uses QXGA video graphics in combination with a Spitz ATM-4 automation system. The SCIdome uses a D-ILA (1280 × 1024) digital projector with a wide-angle lens. Starry Night™ software is provided. Spitz also offers Lumisphere™ (by Laser Fantasy), a full-dome laser projector in full color for domes up to 50 ft (15 m). AVI Omniscan (by Audio Visual Imagineering) also creates full-color full-dome coverage. Spitz Inc. also offers an extensive library of full-dome video shows with a variety of titles 20 to 40 min in length as well as an extensive collection of panoramic space art. Charles F. Hagar

Bibliography. R. L. Beck and D. Schrader, *America's Planetariums and Observatories: A Sampling*, 1991; C. F. Hagar, *Planetarium*: *Window to the Universe*, Carl Zeiss, 1980; F. Haag and C. Haag, *Scientific American Guide to Science Centers, Museums, and Planetariums*, 2005; H. C. King, *Geared to the Stars*, University of Toronto Press, 1978.

Planetary gear train

An assembly of meshed gears consisting of a central or sun gear, a coaxial internal or ring gear, and one or more intermediate pinions supported on a revolving carrier. Sometimes the term planetary gear train is used broadly as a synonym for epicyclic gear train, or narrowly to indicate that the ring gear is the fixed member. In a simple planetary gear train the pinions mesh simultaneously with the two coaxial gears (see **illus.**). With the central gear fixed, a pinion rotates about it as a planet rotates about its sun, and the gears are named accordingly: the central gear is the sun, and the pinions are the planets.

In operation, input power drives one member of a planetary gear train, the second member is driven to provide the output, and the third member is fixed. If the third memberis not fixed, no power is delivered. This characteristic provides a convenient clutch action. A clutch or brake band positioned about the intermediate member and fixed to the gearbox housing serves to lock or free the third member. The holding device itself does not enter into the power path.

Any one of these three elements can be fixed: the sun gear, the carrier, or the ring gear. Either of the two remaining elements can be driven and the other one used to deliver the output. There are six possible combinations, although three of these provide velocity ratios that are reciprocals of the other three. The principal ratios for a simple planetary gear train are given in **Table 1**. The ratios are entirely independent of the number of teeth on each planet. However, to assemble the unit, Eq. (1), which does involve the

$$N_R = N_S + 2N_P \tag{1}$$

number of teeth on each planet, must be satisfied; N_R = number of teeth on the ring gear, N_S = number of teeth on the sun gear, and N_P = number of teeth on each planet. Typically more than one planet is used to distribute the load through more than one mesh and to achieve dynamic balance. If the several planets are to be equally spaced, assembly can be

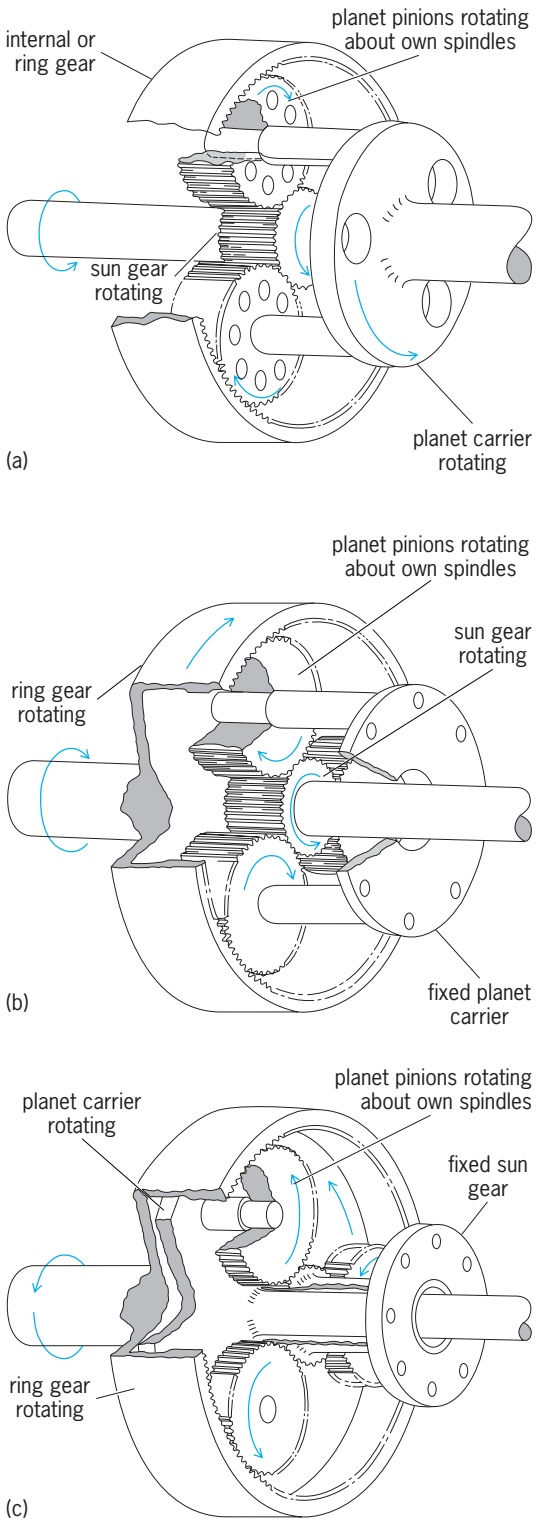

Cutaway isometrics of modes of operation for simple planetary gear train. (*a*) Ring gear locked. (*b*) Planet carrier locked. (*c*) Sun gear locked. (*After D. W. Dudley, ed., Gear Handbook, McGraw-Hill, 1962*)

made only if Eq. (2) is followed.

$$\frac{N_R + N_S}{\text{Number of planets}} = \text{an integer} \qquad (2)$$

Two simple planetary gear sets running on a common sun gear are known as a Simpson gear train. It is widely used in automotive automatic transmissions.

TABLE 1. Speed ratios for simple planetary train

Fixed member	Input member	Output member	Overall ratio*	Range of ratios normally used
Ring	Sun	Carrier	$\frac{N_R}{N_S} + 1$	3:1–12:1
Carrier	Sun	Ring	$\frac{N_R}{N_S}$	2:1–11:1
Sun	Ring	Carrier	$\frac{N_S}{N_R} + 1$	1.2:1–1.7:1

*N_S = number of sun teeth, N_R = number of ring teeth.

TABLE 2. Speed ratios for compound planetary train

Fixed member	Input member	Output member	Overall ratio*	Range of ratios normally used
Ring	Sun	Carrier	$\frac{N_R N_{P_1}}{N_S N_{P_2}} + 1$	6:1–25:1
Carrier	Sun	Ring	$\frac{N_{P_1} N_R}{N_S N_{P_2}}$	5:1–24:1
Sun	Ring	Carrier	$\frac{N_S N_{P_1}}{N_R N_{P_1}} + 1$	1.05:1–2.2:1

*N_S = number of sun teeth; N_{P1} = number of first-reduction planet teeth; N_{P2} = number of second-reduction plan et teeth; N_R = number of ring teeth.

In a compound planetary train, two planet gears are attached together on a common shaft. One planet meshes only with the central sun gear, the other only with the ring gear. As in simple planetary trains, there can be several of these planet pairs distributed around the train to distribute the load and achieve balance. **Table 2** shows the ratios obtained for the three principal arrangements.

To assemble the train, the pitch diameters of the gears must agree with Eq. (3), where D_R, D_s, and D_P

$$D_R = D_S + D_{P_1} + D_{P_2} \qquad (3)$$

designate ring gear, sun gear, and planet diameters, respectively. If the planets are to be equally spaced about the train, the number of teeth must agree with Eq. (4).

$$\frac{N_{P_1} N_R + N_{P_2} N_S}{\text{Number of planets}} = \text{an integer} \qquad (4)$$

See AUTOMOTIVE TRANSMISSION; GEAR TRAIN; RECIPROCATING AIRCRAFT ENGINE.

John R. Zimmerman; Donald L. Anglin

Bibliography. D. W. Dudley, *Handbook of Practical Gear Design*, 1984; J. E. Shigley and C. R. Mishke, *Mechanical Engineering Design*, 5th ed., 1989; J. E. Shigley and J. J. Uicker, *Theory of Machines and Mechanisms*, 2d ed., 1995.

Planetary nebula

An expanding nebula, or cloud, of gas thrown off by a dying star just before the star settles down to its ultimate endpoint as a white dwarf. Planetary nebulae

are among the brightest and best-studied nebular objects in the sky, even though they are generally a few thousand light-years from Earth. The many shapes of planetary nebulae reflect poorly understood processes that occur inside most stars late in their lives. The Sun is likely to eject a planetary nebula in about 5 billion years. Many of the most prominent planetary nebulae seen today formed and emerged from their dusty cocoons within just the past 1000–2000 years. *See* SUN; WHITE DWARF STAR.

Planetary nebulae appear small and greenish in color, like the planet Uranus, when seen through small telescopes—hence the origin of the name by their discoverers during the nineteenth century. However, planetary nebulae are not planets but large (0.1–10 light-years), expanding (10–100 km/s; 6–60 mi/s), highly symmetrical gaseous clouds of stellar ejecta. The shapes of planetary nebulae vary (see **colorplate**), and are denoted as round, elliptical, or bipolar depending on their outlines. Delicate filaments, knots, and bubbles of material characterize their interiors. These features evolve as stellar winds of increasing speed (from 100 to 1000 km/s; 60 to 600 mi/s) plow into older, much slower gas ahead of them, ejected about 1000 years earlier by the geriatric, unstable star. Inferences made from images of very young planetary nebulae suggest that their shapes are determined early in the ejection process, and evolve only slowly as the nebulae grow. The details of the shaping processes are not yet understood. However, the shaping of planetary nebulae is thought to be the result of strong stellar magnetic fields that impregnate the nebular gas as it is ejected from the rotating stellar surface, or gravitational and tidal forces from a closely orbiting companion star or Jupiter-sized planet in an orbit whose radius is roughly that of the Earth's orbit around the Sun.

Planetary nebulae emit electromagnetic radiation in many regions of the spectrum, including radio, infrared, visible, and ultraviolet. Complex patterns appear when the spectrum of light from a given nebula is analyzed. From these data the temperature, density, velocity, and chemical nature of the gas are measured. Such studies have shown that the gas ejected from the parent stars of planetary nebulae is enriched in carbon and nitrogen.

The light output of planetary nebulae can also be used to determine the distances to nearby galaxies, and then to estimate the cosmic distance scale. Just as the distance to a car is estimated from the apparent brightness of its headlights, the distance to a galaxy can be derived from the brightnesses of its planetary nebulae. *See* ASTRONOMICAL SPECTROSCOPY; GALAXY, EXTERNAL; HUBBLE CONSTANT.

A planetary nebula becomes visible only after the cool outer stellar layers have been ejected, exposing the hotter inner regions, or core, of the star to view. The fuel-starved core starts to cool, much like the filament of an electric oven when turned off, for billions of years after ejecting its nebula. Like the old stars from which they are believed to have evolved, planetary nebulae are distributed throughout the Milky Way Galaxy. They show weak signs of crowding in the spiral arms where dust clouds and new stars are found. *See* MILKY WAY GALAXY.

Many thousands of planetary nebulae have been cataloged. For every known planetary nebula, hunderds more are undoubtedly obscured by opaque dust clouds in the Milky Way. Planetary nebulae last for several thousand years before they expand and become too diffuse to be seen as discrete objects. Human lifetimes are far too short for the evolution of planetary nebulae to be noticed, although their existence is fleeting from a cosmic perspective.

Normal stars such as the Sun form planetary nebulae shortly after evolving through the red-giant phase of their lives. At the end of its life, the Sun will swell so much that it will engulf the planets Mercury and Venus. About half of the daytime sky on Earth will be filled with light from the Sun's huge ember-red surface, and the Earth's temperature will rise until its atmosphere and eventually its oceans evaporate into space. After a few thousand years, as the last of its available nuclear fuel is depleted, the Sun will eject its remaining outer layers, containing at least one-third of its mass. This gas will sweep past the Earth at 10 km/s (6 mi/s) and quickly exit the solar system. In another few thousand years, only the tiny, hot core of the Sun will still be visible. This final remnant of a star will have a temperature of about 100,000°C , and the winds that it will generate will approach speeds of 1000 km/s (600 mi/s). In spite of its tiny size, the solar remnant will be a hundred to a thousand times brighter than the present Sun. Its harsh ultraviolet light will ionize and sterilize the Earth's surface. *See* STAR; STELLAR EVOLUTION.

From stellar death comes future life, however. The winds from planetary nebulae cool as they sweep outward, making possible the condensation of dust particles of graphite (carbon), iron, and other types of cosmic dust particles. These carbon-rich dust particles merge with others from the detritus of other dead stars and become encrusted in water ice. These materials will condense into future generations of stars and planetesimals. Their masses, lifetimes, and large numbers suggest that planetary nebulae are the most prolific source enriching the interstellar medium with carbon and, to a lesser degree, nitrogen. Indeed, since no carbon could have survived the big bang, carbon-based life as we know it could not have occurred until generations of planetary nebuale seeded the clouds from which life formed with this element. *See* COSMOLOGY; INTERSTELLAR MATTER.

Bruce Balick

Bibliography. B. Balick and A. Frank, The extraordinary deaths of ordinary stars, *Sci. Amer.*, July 2004; S. Kwok, *Cosmic Butterflies*, Cambridge University Press, 2001; S. Kwok, *The Origin and Evolution of Planetary Nebulae*, Cambridge University Press, 2000; M. Livio, Hubble's top 10, *Sci. Amer.*, July 2006; P. Moore and L. Robinson, *Astronomy Encyclopedia*, Oxford University Press, 2002.

Plant

An organism that belongs to the Kingdom Plantae (plant kingdom) in biological classification. The term is also loosely used to indicate any organism that is not an animal. In the most commonly accepted modern classification (the five-kingdom system), fungi, bacteria, and algae belong to the kingdoms Fungi, Monera, and Protista, respectively, rather than to the plant kingdom as they did in the outmoded two-kingdom system. The study of plants is called botany. *See* BOTANY; CLASSIFICATION, BIOLOGICAL.

Characteristics. In the five-kingdom system, the Plantae share the characteristics of multicellularity, cellulose cell walls, and photosynthesis using chlorophylls *a* and *b* (except for a few plants that are secondarily heterotrophic). Most plants are also structurally differentiated, usually having organs specialized for anchorage, support, and photosynthesis. Tissue specialization for photosynthetic, conducting, and covering functions is also characteristic. Plants have a sporic (rather than gametic or zygotic) life cycle that involves both sporophytic and gametophytic phases, although the latter is evolutionarily reduced in the majority of species. Reproduction is sexual, but diversification of breeding systems is a prominent feature of many plant groups. *See* PLANT REPRODUCTION.

Diversity. A conservative estimate of the number of described species of plants is 250,000. There are possibly two or three times that many species as yet undiscovered, primarily in the Southern Hemisphere. The number of known species is, however, large enough that plants are further categorized into nonvascular and vascular groups, and the latter into seedless vascular plants and seed plants. The nonvascular plants include the liverworts, hornworts, and mosses. The vascular plants without seeds are the ground pines, horsetails, ferns, and whisk ferns; seed plants include cycads, ginkgos, conifers, gnetophytes, and flowering plants. Each of these groups constitutes a division in botanical nomenclature, which is equivalent to a phylum in the zoological system.

Function. Plants are essential to the survival of virtually all living things. Through photosynthesis, plants provide all of the food for human and animal consumption. Plants also are the source of such diverse products as building materials, textile fibers, gums, resins, waxes, rubber, perfumes, dyes, and tanning materials. Most of the drugs that are used in modern medicine were originally derived from plants, although many others came from fungi. In the past, medical doctors needed to be as knowledgeable about plants as they were about their patients. Plants also provide habitats for wildlife and birds, offer shelter, and contribute to soil-building processes. Plants of the Carboniferous era provided the energy that is now used as oil, coal, and gas. The importance of plants to the existence of humans makes the preservation and study of plants an essential component of environmental studies. *See* PHOTOSYNTHESIS; PLANT ANATOMY; PLANT KINGDOM; PLANT PHYSIOLOGY; PLANT TAXONOMY.

Meredith Lane

Bibliography. H. T. Hartmann et al., *Plant Science: Growth, Development and Utilization of Cultivated Plants*, 3d ed., 2000; R. M. Klein and D. T. Klein, *Fundamentals of Plant Science*, 1987; P. H. Raven, R. F. Evert, and H. Curtis, *Biology of Plants*, 6th ed., 1998.

Plant anatomy

The area of plant science concerned with the internal structure of plants. It deals both with mature structures and with their origin and development.

Plant anatomists dissect the plant and study it from different planes and at various levels of magnification. At highest magnification they examine the smallest units of plant structure, the cells; at intermediate magnification they observe the organized aggregations of these cells, the tissues; and at low magnification they determine the arrangement and interrelations of tissues in plant organs such as root, stem, leaf, and flower. At the level of the cell, anatomy overlaps plant cytology, which deals exclusively with the cell and its contents. Sometimes the name plant histology is applied to the area of plant anatomy directed toward the study of cellular details of tissues. *See* PLANT ORGANS; PLANT TISSUE SYSTEMS.

Katherine Esau

Bibliography. H. A. De Bary, *Comparative Anatomy of the Vegetative Organs of the Phanerogams and Ferns*, trans. by F. O. Bower and D. H. Scott, 1884; A. J. Eames, *An Introduction to Plant Anatomy*, 2d ed., 1947, reprint 1977; A. Fahn, *Plant Anatomy*, 4th ed., 1990; E. M. Gifford, Jr., *Morphology and Evolution of Vascular Plants*, 3d ed., 1989; G. F. Haberlandt, *Physiological Plant Anatomy*, 4th ed., 1928, reprint 1979; B. D. Jackson, *A Glossary of Botanic Terms, with Their Derivation and Accent*, 1986; P. J. Kramer, *Plant and Soil Water Relationships*, 1949; A. C. Leopold et al., *Plant Growth and Development*, rev. ed., 1975.

Plant-animal interactions

The examination of the ecology of interacting plants and animals by using an evolutionary, holistic perspective. For example, the chemistry of defensive compounds of a plant species may have been altered by natural-selection pressures resulting from the long-term impacts of herbivores. Also, the physiology of modern herbivores may be modified from that of thousands of years ago as adaptations for the detoxification or avoidance of plant defensive chemicals have arisen. Other examples of plant-animal interactions include pollination biology, fruit and seed dispersal ecology, herbivory on plants by invertebrates (such as insects) and vertebrates (such as ungulates), ant-plant interactions, the ecology and evolution of carnivorous plants, and community and ecosystem patterns of animal distributions resulting

Effects of interaction types for each species*

Interaction	Effect on species A	Effect on species B
Mutualism	+	+
Commensalism	+	0
Antagonism	+	−
Competition	−	−
Amensalism	0	−
Neutralism	0	0

*+ = beneficial, − = harmful, 0 = neutral.

from the availability of plant resources.

The application of the theories based on an understanding of plant-animal interactions provides an understanding of problems in modern agricultural ecosystems. For instance, the problem of limited genetic variation in crop species impacts on attempts to enhance a crop's genetically based resistance to pests. Similarly, plant-animal interactions have practical applications in medicine. For example, a number of plant chemicals, such as digitalin from the foxglove plant, that evolved as herbivore-defensive compounds have useful therapeutic effects on humans. *See* AGROECOSYSTEM.

Types of interactions. The evolutionary consequences of plant-animal interactions vary depending on the effects on each participant. Interaction types range from mutualisms, that is, relationships which are beneficial to both participating species, to antagonisms, in which the interaction benefits only one of the participating species and negatively impacts the other. Interaction types are defined on the basis of whether the impacts of the interaction are beneficial, harmful, or neutral for each interacting species (see **table**). In some cases, these distinctions of types can be unclear because the effects of interactions, whether beneficial, harmful, or neutral, can change in time and space. Amensalisms (that is, interactions that are neutral to one species but harmful to a second species) and neutralisms (that is, interactions that are neutral to both interacting species) are beyond the scope of most studies of plant-animal interaction.

Antagonism. The high frequency of antagonistic interactions in most environments can be attributed to the energy and nutrient needs of organisms. For example, plants are concentrated parcels of such resources for herbivores, as are insects for carnivorous plants. As a consequence of the strong influences that antagonisms have on individual plants, many plants have evolved physical defenses (for example, spines, thorns, or hairs) or chemicals that repel or poison potential herbivores. Some herbivores have counteradapted by evolving resistance to these defenses or mechanisms to avoid them entirely. There are several forms of antagonism, including herbivory, parasitism, and predation.

Mutualism. The numerous examples of mutualisms include many of the relationships between pollinators and flowers or between fruits and fruit-eating birds, mammals, or reptiles. The plant-produced reward (for example, pollen, nectar, or fruit) benefits the pollinator or disperser, while the animal's activities (for example, carrying pollen to another blossom or dispersing seeds) profit the plant. However, not all pollinator and fruit disperser and plant interactions are mutualisms. Some flowers and fruits deceive their pollinators or dispersers and provide no benefit to the animal. The frequency of mutualisms within a community depends on the richness of antagonistic interactions since mutualisms often evolve from antagonisms. Mutualisms themselves can create new resources (nectar, fruits) that promote the development of additional interactions. Thus, mutualisms contribute to the complex organization of communities and biodiversity.

Commensalism. The effects of commensalism are beneficial for one of the two interacting species but are neutral for the other. Examples include the use of trees as nesting places by birds or the temporal succession of mammalian herbivores on the Serengeti Plains of East Africa. In both cases, the relationship of animals to plants influences the animals' distribution and abundance. In the Serengeti, plants indirectly affect a commensalistic relationship involving the zebra, wildebeest, and Thomson's gazelle. The zebra, which migrates first, eats tall, protein-poor grass stems; the wildebeest follows and consumes shorter, more-nutritious grass leaves; while Thomson's gazelle, the last to migrate, feeds on the nutritious regrowth of grasses and herbs. The wildebeest is the commensal of the zebra, and Thomson's gazelle is the commensal of the wildebeest. However, the relationship between these herbivores and their plant foods is antagonistic.

Competition. This is an interaction in which one organism consumes a resource that would have been available to, and might have been consumed by, another. Within the context of plant-animal interactions, competition typically affects outcomes in indirect ways. For example, competition among herbivores for plant-food resources can strongly influence the distribution of competing animals. Variations in adaptations to feeding often lead to segregation of similar species within a habitat and consequently allow coexistence. *See* ECOLOGICAL SUCCESSION.

Warren G. Abrahamson

Fossil record. Plants and animals interact in a variety of ways within modern ecosystems. These may range from simple examples of herbivory (animals eating plants) to more complex interactions such as pollination or seed and fruit dispersal. Conversely, animals also rely on plants for food and shelter. For example, various invertebrates ingest algae, which provide a source of photosynthetically produced food in the absence of normally ingested microorganisms. The complex interactions that have existed between these organisms over geologic time not only have resulted in an abundance and diversity of organisms in time and space but have also contributed to many of the evolutionary adaptations found in the biological world.

Paleobiologists have attempted to decipher some of the interrelationships that existed between plants

and animals throughout geologic time. This field of study has been slow to emerge because there is a relative absence of detailed information about many plants and animals. In addition, the ecological setting in which the organisms lived in the geologic past is being analyzed in association with the fossils. Thus, as paleobiologists have increased their understanding of certain fossil organisms, it has become possible to consider some of the aspects of the ecosystems in which they lived, and in turn, how various types of organisms interacted.

Herbivory. Perhaps the most widespread interaction between plants and animals is herbivory, in which plants are utilized as food. A common example involves partially chewed leaves or areas along the margin where the lamina has been removed. One reason why there have not been many documented examples of partially chewed leaves is the bias of collecting complete, perfect specimens. In some cases, the outline of the area that has been removed provides insight into the type of organism responsible for the herbivory. Another method for determining the extent of herbivory in the fossil record is by analyzing the plant material that has passed through the digestive gut of the herbivore. For example, in Carboniferous coal balls there are numerous coprolites (fecal pellets) of various sizes and shapes. Some of these contain plant fragments of the same type, while others are made up of the leaves and stems of different types of plants. Still other coprolites contain only pollen grains, suggesting that the herbivores may have been more specific in their nutritional requirements. In addition, the size and shape of the coprolites can be used to determine their source.

The analysis of some coprolites in Carboniferous coal balls suggests that swamps were inhabited by mites, collembolans, millipedes, and cockroaches. Chemical systems that made leaves less palatable for herbivores allowed some of the plants to defend themselves against herbivory. Other plants evolved defense mechanisms in the form of hairs, thick seed coats reinforced with fibrous tissues, and various mechanical methods to protect the more nutritious reproductive organs. Other plants evolved the tree habit in which leaves and reproductive organs are positioned some distance from the forest floor. It is believed that the evolution of flight by certain invertebrates evolved in response to plants repositioning their reproductive organs in the upper canopy of the forest.

The stems of some fossil plants show tissue disruption similar to various types of wounds occurring in plant parts that have been pierced by animal feeding structures. As plants developed defense systems in the form of fibrous layers covering inner, succulent tissues, some animals evolved piercing mouthparts that allowed them to penetrate these thick-walled layers. In some fossil plants, it is also possible to see evidence of wound tissue that has grown over these penetration sites. *See* COAL BALLS; HERBIVORY.

Mimicry. Another example of the interactions between plants and animals that can be determined from the fossil record is mimicry. Certain fossil insects have wings that are morphologically identical to plant leaves, thus providing camouflage from predators as the insect rested on a seed fern frond.

Pollination. The transfer of pollen from the pollen sacs to the receptive stigma in angiosperms, or the seed in gymnosperms, is an example of an ancient interaction between plants and animals. It has been suggested that pollination in some groups initially occurred as a result of indiscriminate foraging behavior by certain animals, and later evolved specifically as a method to effect pollination. The rewards offered by plants can be documented in certain fossils in the form of specialized glands and cavities that secreted nectar, or of expendable pollen as sources of food. The particular structure and organization of flowers, size of floral parts, and nature of the anthers represent sources of information in fossil flowers that can be used to study ancient pollination systems. In some instances, the morphology of the flower can be used to determine the type of insect pollinator. In addition, the size, shape, and organization of fossil pollen grains provide insight into potential pollination vectors. Far less information is available about the role of animals in the dissemination of fossil fruits and seeds. However, the presence of spines and hooked projections on certain types of seeds suggests that they may have been transported by larger animals. In addition, many fossil species bore seeds designed to pass through the digestive guts of animals that ate fleshy reproductive organs such as fruits. *See* FOSSIL; PALEOBOTANY; POLLINATION. Thomas N. Taylor

Bibliography. W. G. Abrahamson (ed.), *Plant-Animal Interactions*, 1989; W. C. Ciepet, Insect pollination: A paleontological perspective, *BioScience*, 29:102–108, 1979; H. F. Howe and L. C. Westley, *Ecological Relationships of Plants and Animals*, 1988; A. C. Scott and T. N. Taylor, Plant/animal interactions during the Upper Carboniferous, *Bot. Rev.*, 49:259–307, 1983; J. N. Thompson, *Interaction and Coevolution*, 1982; S. L. Wing and B. H. Tiffney, The reciprocal interaction of angiosperm evolution and tetrapod herbivory, *Rev. Paleobot. Palynol.*, 50:179–210, 1987.

Plant cell

The basic unit of structure and function in nearly all plants. Although plant cells are variously modified in structure and function, they have many common features. The most distinctive feature of all plant cells is the rigid cell wall, which is absent in animal cells. The range of specialization and the character of association of plant cells is very wide. In the simplest plant forms a single cell constitutes a whole organism and carries out all the life functions. In just slightly more complex forms, cells are associated structurally, but each cell appears to carry out the fundamental life functions, although certain ones may be specialized for participation in reproductive processes. In the most advanced plants, cells are associated in

functionally specialized tissues, and associated tissues make up organs such as the leaves, stem, and root. Although a substantial body of knowledge exists concerning the features of various types of plant cells, there is a great gap in knowledge of how cells become specialized for particular functions and associations.

Plant and animal cells are composed of the same fundamental constituents—nucleic acids, proteins, carbohydrates, lipids, and various inorganic substances—and are organized in the same fundamental manner. A characteristic of their organization is the presence of unit membranes composed of phospholipids and associated proteins and in some instances nucleic acids.

For convenience this article first considers the structure of the cell in the interphase or nondividing stage and then in the dividing stage.

Interphase Cell

For many years it was usual to employ the term protoplasm to refer collectively to the cell components responsible for imparting the characteristics of life. The term has undergone some modification, and all the structures internal to the cell membrane are frequently referred to as protoplasts. However, techniques such as electron microscopy and cytochemistry have made it increasingly possible to define structures and to ascribe particular functions to the two major components of the protoplast: the nucleus and the cytoplasm (**Fig. 1**). Furthermore, it is possible to localize specific activities in given regions or substructures in these components. *See* CYTOCHEMISTRY; ELECTRON MICROSCOPE.

Nucleus. The nucleus of the undifferentiated plant cell is generally centrally located, whereas in mature plant cells it may be displaced to the edge of the cell by the presence of a large vacuole. The nuclear material consists of a ground substance and chromatin, which at the time of division is resolved into chromosomes. Nuclear components of key importance are deoxyribonucleic acid (DNA), ribonucleic acid (RNA), and proteins. Continuing syntheses of DNA, RNA, and proteins, notably histones, are predominant activities in the nucleus, although comparable processes also take place elsewhere in the cell. The singular importance of the relationships of DNA to RNA and proteins lies in the fact that the hereditary characteristics of the cell, encoded in DNA molecules, are transmitted in a complex sequence via RNA to proteins. *See* NUCLEIC ACID.

The same double-helix form characteristic of animal cell DNA is demonstrated for DNA isolated from plant cell nuclei, and the evidence indicates that the coding of proteins follows the same pattern in both cell types. *See* GENETIC CODE.

Different types of RNA are formed in the nucleus, transfer RNA (tRNA), ribosomal RNA (rRNA), and presumably messenger RNA (mRNA), all of which participate in the synthesis of proteins by ribosomes.

One or more nucleoli are found in the nuclei of all plant cells. Chemically, plant nucleoli contain various types of RNA and relatively large amounts of protein, components which presumably play distinctive roles in the transmission of hereditary information to various sites in the cell. Chromatin DNA is structurally intermingled with the components of the nucleoli. The appearance of the nucleolus changes through development, probably in relation to control of cellular activities.

As in animal cells, the nucleus of the plant cell is bounded by an envelope. The envelope is a complex structure, consisting of two membranes with a lumen (called the perinuclear space) between, and with discontinuities (called pores) fairly regularly spaced throughout. These pores are regular in outline and may be bounded by rings, or annuli. They are frequently seen to contain material that stains differently from the adjacent regions, and it has been suggested that interchanges, particularly of large molecules, between nucleus and cytoplasm take place through the pores rather than across the membranes of the envelope (**Fig. 2***a*). The perinuclear space has been relatively little studied, but some cytochemical studies suggest the presence of enzymes (Fig. 2*b*). The nuclear envelope, with a metabolically active lumen, apparently is a complex system for controlling nucleocytoplasmic exchanges.

The inner membrane of the nuclear envelope is frequently in close association with portions of the chromatin. The outer membrane may be studded with ribosomes. In some instances extensions of the nuclear envelope continue outward into the cytoplasm, often in specific association with particular organelles (**Fig. 3**). In instances in which segments of nuclear envelope have particular architectural modifications, and these same modifications are seen elsewhere in the cell, it has been suggested that the membrane may be transferred from the nuclear envelope, perhaps transporting perinuclear material with it. No substantial experimental evidence has been provided for this, but it does seem clear that the nuclear envelope is a dynamic structure from which vesicles may be evolved in certain stages of cellular activity. One example is the frequent appearance of what appear to be vesicles being pinched off from the nuclear envelope and transported to the Golgi apparatus. *See* CELL NUCLEUS.

Cytoplasm. The region of the cell between the nuclear envelope and the plasma membrane, or plasmalemma, bounding the cell is best considered in terms of a matrix throughout which are distributed various organelles and inclusions. Metabolic activities may be generally distributed throughout the matrix, confined to specific regions, or clearly carried out within the organelles. Although more research has been directed to the analysis of the chemical composition and activities of the organelles than of the matrix, it is apparent that in addition to water the matrix contains, at any given time, various minerals and the recognized organic components of the cell.

The organelles, which are increasingly interpreted as compartments in which certain metabolic

Fig. 1. Spinach (*Spinacia oleracea*) root epidermal cell showing the nucleus and the cytoplasm with inclusions. In plant cells, vacuole size often determines the position of the nucleus. (*Courtesy of M. Dauwalder*)

activities are localized, are bounded by unit membranes. The molecular components (phospholipids and proteins) of the membranes are subject to rapid turnover. The membranes act as sites for the synthesis or breakdown of materials and frequently, as in mitochondria, are structurally highly specialized for these activities. Therefore, far from being simply selective barriers to the movement of materials, the membranes of the plant cell are dynamic structures which play key roles in metabolism.

Ribosomes. Conspicuous among the components of the cytoplasmic matrix are particles, approximately 20 nanometers in diameter, composed of RNA and proteins; these particles are functionally related to similar ones in the nucleus, mitochondria, and plastids. Formerly called ribonucleoprotein bodies but now generally known as ribosomes, these particles function in the relay of the coding system from nucleic acids in the nucleus and as sites of protein synthesis. A characteristic feature of the plant cell is the occurrence in chloroplasts of ribosomes, distinguished by RNA components of different molecular size from those of the cytoplasmic matrix.

The RNA component of the cytoplasmic ribosomes of plant cells can be separated into two (sometimes three) molecular sizes, as can that of the animal cells that have been investigated. When studied by highly specialized techniques, each ribosome is seen to be composed of two distinguishable subunits, but whether this relates to the presence of rRNA molecules of different sizes and what its importance in protein synthesis may be remain subjects for further exploration.

In many instances the cytoplasmic ribosomes are seen in spiral or helical arrangements called polyribosomes or polysomes, in which the ribosomes are interconnected by thin strands of mRNA. The current interpretation is that these arrangements

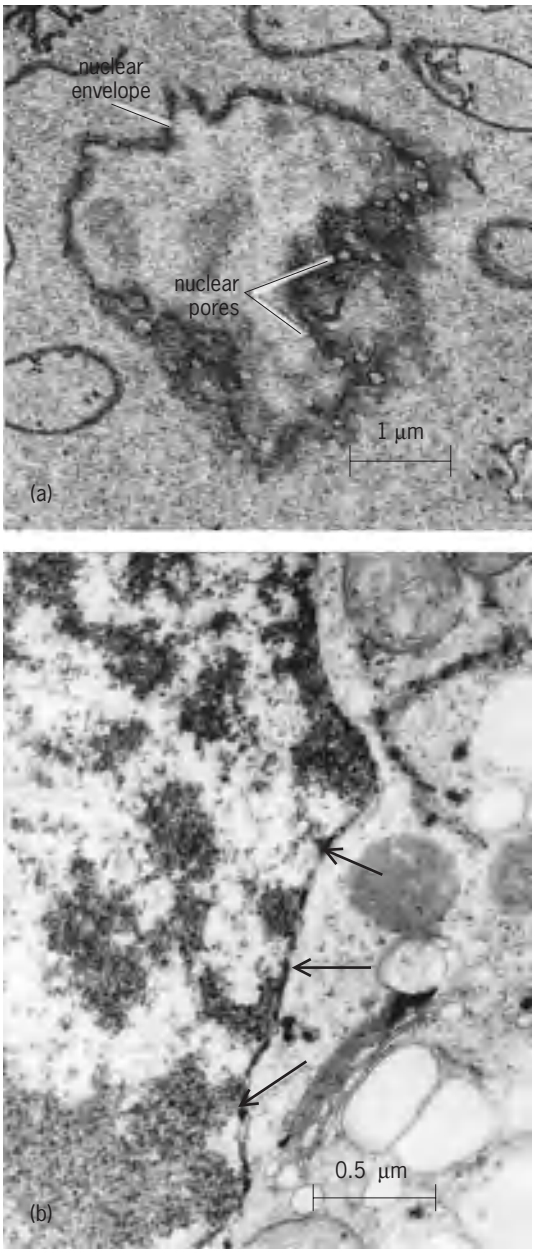

Fig. 2. Cell nucleus. (*a*) Somewhat tangential section (*courtesy of H. H. Mollenhauer*). (*b*) Section showing the nuclear envelope. The dark lead precipitate (arrows) indicates presence of acid phosphatase within the perinuclear space (*courtesy of C. W. Goff*).

represent ribosomes moving along strands of mRNA which impart information from the DNA and thus code protein synthesis. In all types of cells some of the ribosomes in the cytoplasm appear to be free, while others are attached to the surface of the membranes of the endoplasmic reticulum or to the outer membrane of the nuclear envelope. *See* RIBOSOMES.

Endoplasmic reticulum. The character of the cytoplasmic compartment known as the endoplasmic reticulum has been known only since electron microscopy has been adapted to the study of cells. This compartment may be tubular or lamellar; each form appears to be a profile of two membranes with a lumen between. In plant cells the endoplasmic reticulum is commonly lamellar; but frequently it appears to pass through the plasma membrane and sometimes even the cell wall, and it then becomes tubular. Extensions of the endoplasmic reticulum to and through the cell wall, which provide a form of continuity from one cell to another, form part of the cytoplasmic strands called plasmodesmata.

The endoplasmic reticulum is an architecturally regular structure only in a few types of plant cells. It frequently shows dilatations, sometimes containing crystals or other distinctively staining material. The endoplasmic reticulum is a labile structure, and the manner in which its profiles are associated differ with the stage of develoment and metabolic activity. In certain stages numbers of profiles are seen to be stacked, frequently parallel to the surface of the cell (**Fig. 4**). The profiles may also surround the nucleus or seem to encompass any of several types of organelles. The endoplasmic reticulum may be smooth or rough; that is, the outer surfaces of the membranes may be studded with ribosomes. Smooth endoplasmic reticulum occurs less abundantly in plant cells than in animal cells.

Experiments using animal cells indicate that the endoplasmic reticulum sequesters into its lumen proteins or protein precursors, presumably syntheized at the ribosomes, and that it is the site of synthesis of at least some of the lipid components of the cell. There is no reason to suppose that the same does not obtain in plant cells. Cytochemical analyses have demonstrated a number of enzymes either in the lumen or associated with the membranes of the endoplasmic reticulum. It is known that in animal cells certain substances are transferred from the endoplasmic reticulum to other sites in the cell. For example, in many secreting cells it can be shown

Fig. 3. Section of a *Chlamydomonas* cell showing extension of the nuclear envelope (arrow) to the vicinity of the Golgi apparatus. (*Courtesy of P. L. Walne*)

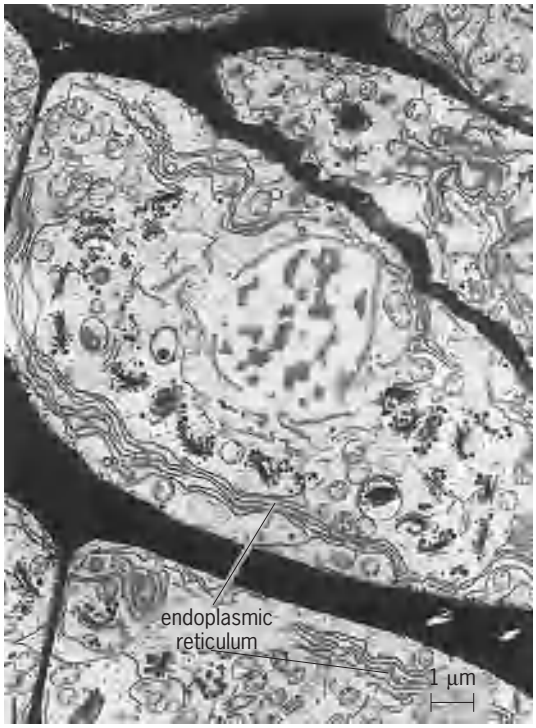

Fig. 4. Portions of maize root cap cells showing profiles of endoplasmic reticulum stacked more or less parallel to the cell surface. (*Courtesy of M. Dauwalder*)

by radioautography that proteins or protein precursors move from the endoplasmic reticulum to the Golgi apparatus, where they become part of protein-containing secretory products. The question of how material is transferred from the endoplasmic reticulum to the Golgi apparatus remains unanswered. The most common hypothesis, that the material is transferred in small membrane-bound vesicles, is lent some support by the appearance of blebs in the membrane of the endoplasmic reticulum near the Golgi apparatus and by the occasional appearance of numerous small vesicles between the two structures. Whatever the mechanisms, it seems likely that the patterns of transfer are comparable in both animal and plant cells, though there are obviously differences in physiology. *See* AUTORADIOGRAPHY.

The introduction of toxic substances or other modifications of the metabolic state of the cell sometimes causes extensive proliferation of the endoplasmic reticulum, usually of the rough type. This development has been noted in plant and animal cells and appears to be a compensatory response to deleterious conditions. It may well be that proliferation of the endoplasmic reticulum provides additional reaction surface and more enzymes which increase the capacity of the cell to detoxify substances or otherwise adjust to unfavorable circumstances. *See* ENDOPLASMIC RETICULUM.

Mitochondria. Much of the work involved with isolation and chemical analysis of mitochondria has been done with plant cells. The mitochondria are centers of energy conversion. They function in electron transport and oxidative phosphorylation and act as respiratory control centers of the cell.

Mitochondria are of the same fundamental structural pattern in plant and animal cells. They are bounded by a double-membrane system with the inner membrane projecting into the lumen to form cristae (**Fig. 5**), which may be either tubular or sheet-like, depending upon the type of cell and its activity. In general there is less extensive development of the cristae in the mitochondria of plant cells than in those of animals. This may reflect the fact that plant cells generally have substantially lower respiratory rates. In the few types of plant cells characterized by relatively high respiratory rates, the extent of the cristae more nearly resembles that in animal cells.

Cristae bear stalked structures in ordered arrangements. The integrity of these structures is essential for the conduct of certain metabolic activities involved in the respiratory cycle. In experiments with animal cells mitochondrial activity disappeared when this integrity was destroyed, and there was a recurrence of the activity when the integrity was restored. Inasmuch as both the structure and the indicated activities of mitochondria are similar in plant and animal cells, the indicated relationships presumably hold for plant mitochondria.

That other activities are also involved in the functioning of the mitochondria is suggested by the frequent development of crystalline inclusions (**Fig. 6**) and by the presence of nucleic acids. Mitochondria contain a small portion of the total DNA and RNA complement of the cell. Since ribosomelike particles are also present in the mitochondria, these organelles can probably synthesize some portion of their own proteins. In fact, there is evidence that amino acids are moved into the mitochondria. The presence of nucleic acids within the mitochondria provides the basis for a hypothesis, supported by studies of the biogenesis of mitochondria, that these organelles have at least some degree of autonomy in replication.

In the living cell, mitochondria are in constant motion, frequently in patterns that show the larger number of them shuttling back and forth between

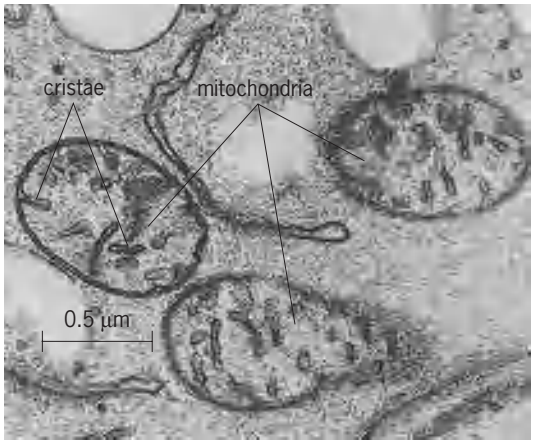

Fig. 5. Three mitochondria showing two bounding membranes and cristae formed by extension of inner membrane into lumen. (*Courtesy of H. H. Mollenhauer*)

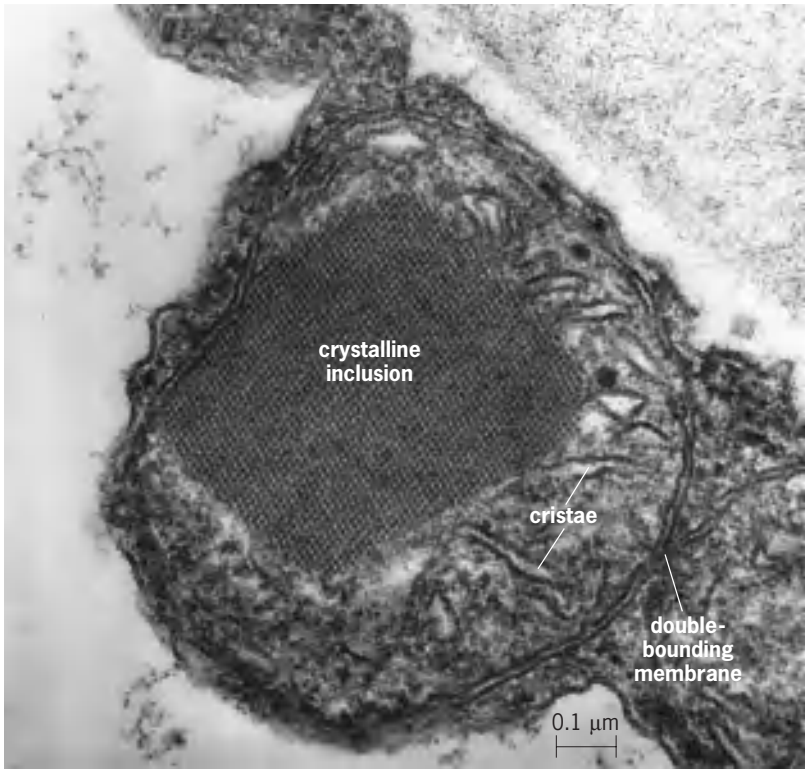

Fig. 6. A highly magnified mitochondrion. Components of crystalline inclusion are in an ordered array. (*Courtesy of H. J. Arnott*)

one cell site and another, presumably in relation to the transfer of energy. When in motion they tend to appear more elongate than in electron microscope preparations, but both light and electron microscope studies show them to be pleomorphic. *See* CITRIC ACID CYCLE; MITOCHONDRIA.

Golgi apparatus. The Golgi apparatus is a component of all plant cells, with the possible exception of certain fungi and some highly specialized cells such as mature sperm. This organelle has been most studied in relation to secretory activities of glandular cells in animals, and its presence in plant cells suggests that secretory activity is widespread.

In many plant cells the Golgi apparatus clearly functions in secretion, but the ubiquitous occurrence of the organelle suggests that it may have other roles in cellular activity. Although many aspects of its function are still obscure, it is apparent that certain materials are either sequestered into its saccules (**Fig. 7**) or synthesized there, or they are variously combined in the saccules to form complex secretion products. These products are then separated from the saccules as membrane-bound vesicles and transported to and through the plasmalemma. Not all cell secretion takes place, however, via the Golgi apparatus. In many types of secreting plant cells the Golgi apparatus is responsible for packaging and exporting a substantial part of the components of the wall that is built up around the cell. In such cells the direction of movement of the secretory product is well established.

The Golgi apparatus is a region characterized by membrane assembly and extension and to which, in animal cells at least, protein or protein precursors are transferred from the endoplasmic reticulum and either elaborated further or combined with carbohydrates. In both animal and plant cells the Golgi apparatus is also the site of synthesis of certain polysaccharides. It also functions in "complexing" activities, which may involve lipid-protein combinations and the addition to the materials assembled there of such elements as iron or sulfur.

The structural and functional similarity of plant- and animal-cell Golgi apparatus supports the concept of a high degree of comparability in their metabolic activities. Differences appear to be related to the presence of more highly specialized organ systems in the animal than in the plant. *See* GOLGI APPARATUS.

Microtubules. As in the case of animal cells, certain of the newer procedures for fixation, notably the use of glutaraldehyde, demonstrate the existence in plant cells of numerous microtubules; these are elongate elements usually with a diameter of 20 nanometers or somewhat more, and presumably composed of polymerized protein.

There is some evidence that there are distinct classes of microtubules distinguished by differences in diameter, location, and orientation. In interphase plant cells the most prominent type is found in conspicuous numbers close to the surface of the cell and in some instances seemingly parallel to it. The number appears to differ considerably from one cell type to another or one stage of activity to another and quite possibly with variations in preparation procedures.

It has been observed that cellulose microfibrils laid down in the wall frequently parallel microtubules inside the cytoplasm. This has led to the speculation that in some manner the orientation of the microtubules controls in part the orientation of the microfibrils. During growth, orientation of the two

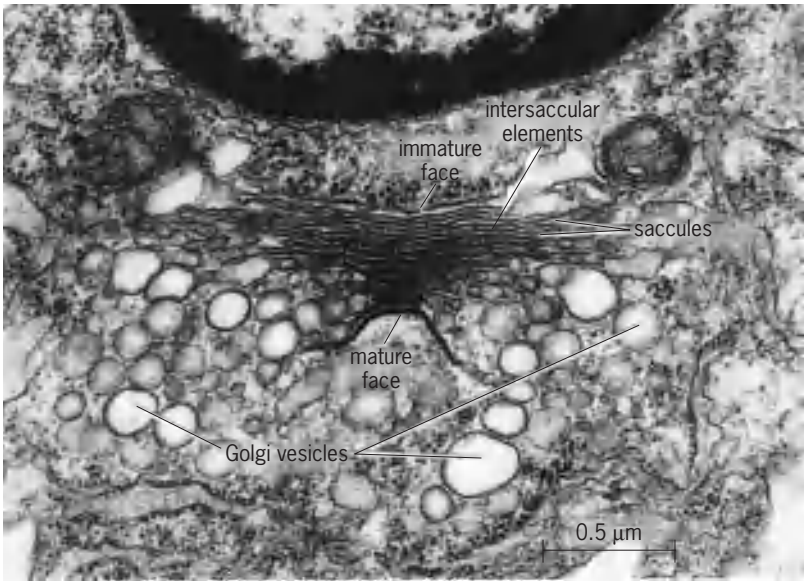

Fig. 7. Portion of a *Nitella* cell showing a Golgi apparatus composed of a stack of saccules, the contents of which stain differentially from the immature face to the mature face. Note also the intersaccular elements and the associated Golgi vesicles produced from the saccules. (*Courtesy of R. Turner*)

appears to be modified more or less simultaneously. The spindle fibers of the dividing cell and of the microtubules of the interphase cell are not only structurally similar but both seem to have some part in determining the direction of movement of certain vesicles within the cell. In some instances, as in motile sperm formed by algae and other lower plants, microtubules are arranged in a consistent pattern, which suggests that they may act as a sort of cytoskeleton. In these instances they appear to be related to, if not immediately derived from, the centrioles. Knowledge concerning the exact nature and functional role of the microtubules is extremely limited. However, it has been observed that they are composed of subunits in patterns similar to the basic structures of components of flagella of motile cells. *See* CILIA AND FLAGELLA.

Plastids. Perhaps the most conspicuous and certainly the most studied of the features peculiar to plant cells is the presence of plastids. Among the organisms generally classified as plants, only the bacteria, the blue-green algae, and some fungi lack plastids. The plastids are membrane-bound organelles with an inner membrane system. Chlorophylls and other pigments are associated, generally in noticeably regular molecular arrangements, with the inner membrane system. The extensiveness of the inner membrane varies greatly with the functional state of the plastid. The chlorophylls, or green pigments, have been the most studied of the plant pigments because of their conspicuous role in fundamental energy processes. The other pigments apparently sometimes play a part in energy processes, sometimes screen out particular wavelengths of light, and sometimes are degradation products. *See* CHLOROPHYLL.

The chloroplast was once thought to develop only from a structure termed a proplastid, and it was assumed that it might have a different origin than other types of plastids. It is fairly clear, however, that a membrane-bound structure, representing a compartment in which certain activities are localized, is characteristic of immature plant cells and that from this structure the several different types of plastids may develop, including the chloroplasts. Chloroplast membranes are characterized in most plants as associations of electron-dense disk-shaped grana and interconnecting stroma, which are lighter on electron micrographs (**Fig. 8**). Chloroplasts also contain what appear to be storage lipids.

The chloroplasts are responsible for the green color of many parts of plants and are functional in photosynthesis as the site where light energy is transformed to chemical energy; the chromoplasts are responsible for certain colors in flowers and fruits; in the amyloplasts starch is stored to provide a reservoir of energy for use in the plant's development; and in elaioplasts, the lipoidal material is the energy reservoir. Such colorless plastids as amyloplasts and elaioplasts are frequently termed leucoplasts.

Like the mitochondria, plastids contain DNA and RNA, which provide a basis for protein synthesis within the plastid as well as the possibility of some self-determination in replication. It has been known since the early 1900s that plastids have a genetic system that is to some extent independent of the genetic system of the nucleus; now it is known that these organelles replicate by division. Replicative ability and the presence of a genetic system which (like

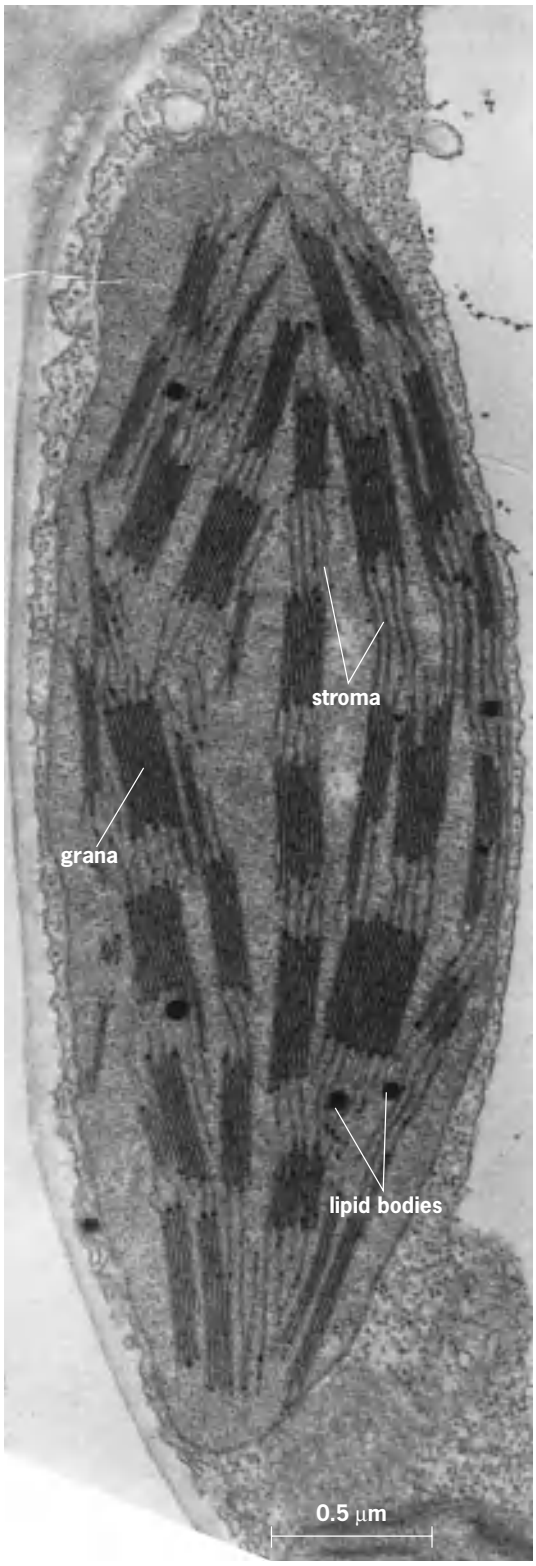

Fig. 8. Electron micrograph of chloroplast of a tomato fruit cell. (*Courtesy of S. W. Rosso*)

that of the mitochondria) may differ somewhat from that of the nucleus have given rise to the speculation that plastids may derive from organisms which invaded non-plastid-bearing cells with the result that an enduring symbiotic relationship was established. This idea finds support in that the RNA components of plastid ribosomes more nearly resemble those of bacteria than those of the nuclear, cytoplasmic, or mitochondrial ribosomal particles found in the plant cell. Whatever the origin of the plastids, these structures are important because of the carbohydrate-synthesizing ability of chloroplasts; color-producing function of the chromoplasts in attracting insects and other pollinators; and the energy reservoir for growth and development, as well as the contribution to plant nutritional value, provided by the storage plastids.

Changes in plastid form, distribution, and quite possibly metabolic efficiency have been conspicuous features of plant evolution. At any evolutionary level, however, the presence of plastids, and all that this implies metabolically, provides a sharp distinction between a plant and an animal cell. Plastids impart to plants their unique status in the biotic world. *See* CELL PLASTIDS; PLANT KINGDOM.

Vacuoles. Vacuoles of different types are characteristic of almost all sorts of cells, at least at certain stages of development. Many types of plant cells have large, centrally located vacuoles that make up the greatest part of the total volume of the cell. In meristematic cells vacuoles are generally small and are characterized by contents that stain darkly with certain procedures. The contents of these vacuoles seem to be utilized in the process of development and then are replaced by water. At a certain stage in this process the vacuoles, which are bounded by a membrane (the tonoplast) that seems to differ in thickness from the other membranes of the cell, fuse to form the large central vacuole. The increase in volume of this vacuole is important in the growth of the plant cell.

Since plants have no organ system to eliminate waste products, it has long been supposed that materials no longer usable in the metabolism of the cell are channeled into the vacuoles and there sealed off from the cytoplasm. This may happen, but it is now apparent that there may be several types of vacuoles in plant cells, some of which may function in a more complex fashion. An example is the cells of the roots of the water hyacinth (*Eichhornia*), in which the membrane of the vacuole may become greatly extended, intrude into the vacuole, and delimit the sites at which crystal formation takes place (**Fig. 9**). Another suggestion concerning metabolism within the vacuole is referred to below in lysosomal activity.

The colors of many flowers and in some instances tissues of other plant organs or structures result from the presence of pigments dissolved in the vacuolar fluid. Anthocyanins in the vacuoles produce red-blue colors, depending upon the relative acidity of the fluid in the vacuole. The origin of vacuoles in the plant cell has been inadequately studied. One common suggestion is that vacuoles arise as dilatations of the endoplasmic reticulum. In animal cells vacuoles

Fig. 9. ***Eichhornia*** **cell showing crystals developing within the vacuole. (*Courtesy of H. J. Arnott*)**

of different sorts are known to have different origins, with the endoplasmic reticulum, the plasmalemma, and perhaps the Golgi apparatus being involved. It seems likely that various vacuoles in plants also may have different origins. *See* VACUOLE.

Lipid droplets. Another conspicuous feature of many types of plant cells is the presence of large numbers of lipid droplets. These may be what some investigators have called spherosomes. They frequently are abundant in cells of embryos or in root or shoot apices, in which cell division and cell growth activities are building up. These lipids are apparently stored material which may support these activities, and they differ in chemical composition from the lipids of the membranes. Lipid droplets are less numerous in more mature plant cells. These droplets

are unique in having a structural boundary that is not the typical unit membrane.

Other organelles and inclusions. Specialized types of plant cells are characterized by other inclusions, some membrane-bound and some crystalline (crystalline inclusions may also occur within the membrane-bound organelles, for example, in the mitochondria). Some types of plant cells have what appear cytologically to be microbodies, though these structures have not been demonstrated to contain the same complement of enzymes as animal cell microbodies. "Multivesiculate" bodies, a catch-all category for bodies showing various types of internal membrane structure, are common in some plant cells, particularly degenerating ones; "dense" bodies, quite possibly representing accumulations of substances in dilatation of the endoplasmic reticulum, are sometimes seen also. Various types of storage bodies, some containing large amounts of protein, for example, are characteristic of the cells of grains and fruits. In addition, plant cells infected with viruses may show numerous bizarre structures that have not yet been interpreted. Normal plant cells may contain other structures at specific stages, such as the centrioles which are characteristic of division stages in some lower plants. Still other structures, such as the pyrenoid and the eyespot of cells of certain algae and the Woronin bodies of certain fungi, distinguish the cells of particular groups of plants.

Lysosomal activity. One apparent difference between higher plant cells and many animal cells is the absence of readily and distinctly definable lysosomes. Many types of animal cells are characterized by the presence of lysosomes: membrane-bound entities which contain a number of hydrolytic enzymes, among which acid phosphatase is generally used as a marker. It seems likely, however, that lysosomal activity occurs in plant cells just as it does in animal cells, because acid phosphatase, and quite possibly the other lysosomal enzymes as well, can sometimes be demonstrated in plant vacuoles, in certain storage bodies, and in other membranous structures, notably the endoplasmic reticulum, in the vicinity of the Golgi apparatus. In plant cells in which digestive activity is a predominant process, lysosomal enzymes are abundant, for example, in cells containing aleurone grains from which protein is released to support early growth of the germinating seedling. These and other storage bodies appear to have high lysosomal activity, though their functions are quite different from those of the digestive vacuoles with lysosomal activity in animal cells. Some lower plant cells, for example, *Euglena*, contain bodies with lysosomal activity that are more nearly comparable to those of the animal cell.

The apparent distinction between plant and animal cells with respect to definable lysosomes may rest in part upon the absence from plants of cells that function in such processes as detoxification and therefore require continuing destruction and renewal of their components, as is the case in liver and white blood cells. *See* LYSOSOME.

Plasmalemma. The cytoplasm is bounded externally by a membrane called the plasma membrane or plasmalemma. Whereas the membranes of the compartments in the cytoplasm separate certain activities from the matrix, the plasmalemma separates the activities of the protoplast from the surrounding environment, whether this is fluid or adjacent cells. The plasmalemma is a typical lipoprotein membrane and is known to be selective with respect to the passage of ions and small molecules into or out of the cytoplasm. While the nature of the exchanges through the plasmalemma is inadequately known, both specific ion exchanges and forms of active transport, energy-requiring exchanges, are involved.

However, neither ion exchange nor active transport explains the passage of all materials through plant cell membranes. These membranes, even those of fairly highly specialized cells in differentiated tissues, often allow the passage of large molecules; such movement into the cytoplasm constitutes endocytosis; movement outward, exocytosis. Exocytosis, such as is involved in some secretion, has been much studied in plant cells; endocytosis has been recognized in only a few instances. The plasma membrane is a very reactive structure. Studies (largely confined to animal cells) have indicated that, like other cellular membranes, there is a rapid turnover of phospholipid and protein molecules in the plasmalemma. Thus the membranous barrier between the cell contents and the environment, once thought to be a more or less passive structure, is now interpreted as a very dynamic one. *See* CELL MEMBRANES.

Motion within cells. Nearly all cells are characterized by several types of motion, ranging from a fundamental motion, so-called Brownian movement, to movements of cellular components along particular pathways to general cyclosis or protoplasmic streaming. All movement in the cell is arrested if the cell is fixed for study under a microscope. Thus the static picture of any cell fails to convey one of its most vital aspects. *See* CELL MOTILITY.

Cell wall. A distinctive feature of the plant cell is the wall surrounding the cell outside the plasma membrane (**Fig. 10**). In mature cells this wall may be variously thickened and sculptured. It is composed largely of cellulose, arranged in microfibrils which have distinctive patterns of orientation and hemicelluloses and pectic substances. The wall is a limiting factor with respect to cell growth beyond a certain stage in development, and it precludes cellular mobility.

The cell wall is formed in two stages, and because of this and differences in composition the botanist generally distinguishes between a primary and a secondary wall. The primary wall is laid down at the time of cell division and extended during cell growth. It is composed mostly of pectic substances and hemicelluloses but also contains cellulose and a small amount of protein. The secondary wall is laid down after the cell has ceased to grow and may vary greatly in structure and composition. It has more cellulose than the primary wall and sometimes contains

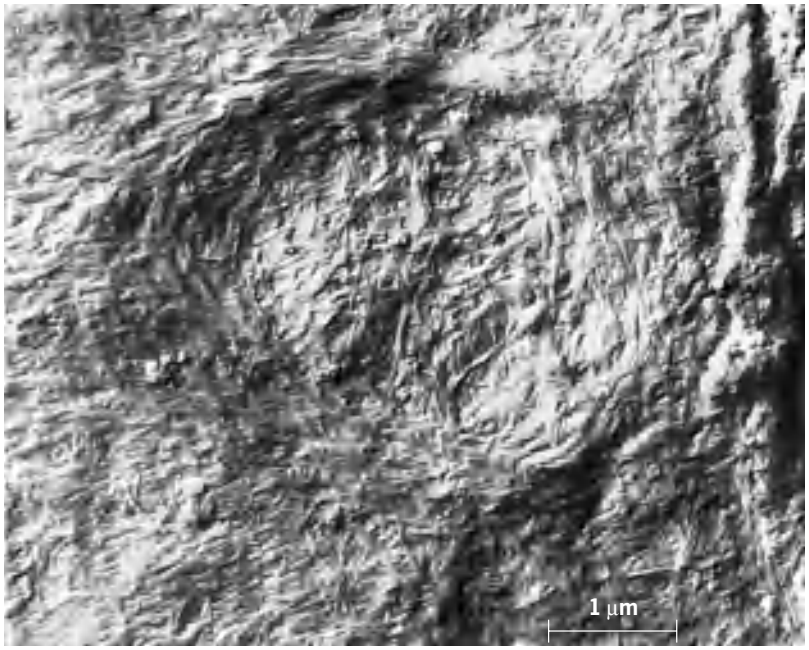

Fig. 10. Portion of a plant cell wall showing the fibrillar nature of cellulose. Microfibrils are oriented in a nearly circular arrangement surrounding intercellular connections (a pit field). (*Courtesy of J. A. S. Marshall*)

substantial amounts of lignin and other substances. In certain types of cells the secondary wall is uniformly thickened; in others it is thickened in bands or in spiral patterns; and in still others the thickenings are only at the corners. A significant portion of the primary wall is secreted by the Golgi apparatus. The largely cellulosic secondary wall appears, however, to be built from materials which move out of the cytoplasm by some other mechanism.

The walls of plant cells with their loosely arranged microfibril structure constitute intercellular pathways through which materials important in the economy of the cell can be moved from one cell to another or even from the external environment of the plant to the plasmalemma of nonsurface cells. *See* CELL WALLS (PLANT).

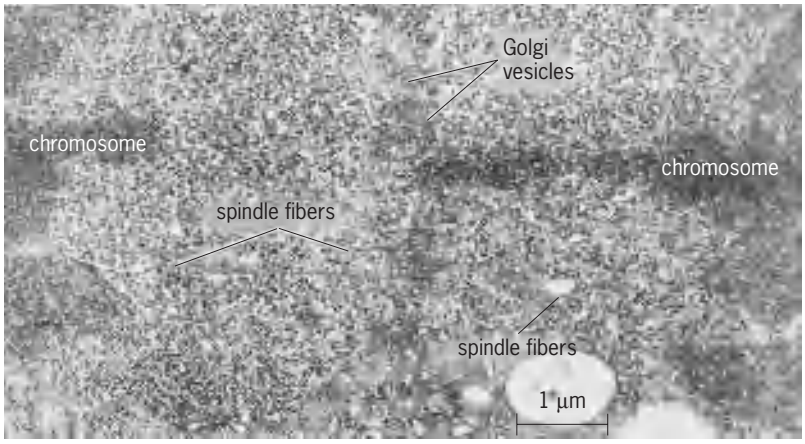

Fig. 11. Dividing cell showing aggregation of small Golgi vesicles in equatorial region and spindle fibers. Chromosomes have moved toward the poles. In right part of cell a chromosome still extends as far as the forming plate. (*Courtesy of H. J. Arnott*)

Variations. Any generalized consideration of the plant necessarily must omit detailed attention to variations characteristic of particular cell types. These variations are myriad and it is possible to touch on only a few. At maturity some cells, such as those of xylem vessels, lose all their cytoplasmic and nuclear contents. Others, such as the sieve-tube elements of the phloem, at maturity contain cytoplasm but no nucleus. The extent of the endoplasmic reticulum, the number of ribosomes, mitochondria, and plastids, and the Golgi apparatus vary widely with the type and functional state of the cell; as does the degree of vacuolation and the relative thickness and character of the walls. All these variations reflect differences in the functioning and therefore the metabolic activities of the cells. *See* PHLOEM; XYLEM.

Cell growth. The growth of the plant cell invariably involves four key features: the addition of more cytoplasmic material produced by metabolic reactions involving both the absorbed nutrients and the products of photosynthesis; intake of substantial amounts of water, usually greatly expanding the volume of the central vacuole; extension of the plasmalemma; and extension of the primary wall.

The details of accretion during growth are relatively obscure in all organisms, though the specific steps involved in many syntheses are well known. Presumably, most of the additional water enters by osmosis. Other mechanisms may be involved in the movement of water from the cytoplasm into the vacuoles. The plasmalemma apparently increases in extent by intussusception, in which more molecular units from the cytoplasmic pool are inserted between those already in the structure. How the primary wall is extended, though much studied, is still not definitely known. The interrelationship between the extension of the plasmalemma and the growth of the primary cell wall has yet to be elucidated.

The growing plant cell is distinguished by remarkable correlation between the increase in volume of the protoplast, which is significantly related to vacuolation, and the extension of the bounding wall. To what extent wall-growth processes are directly dependent upon influences from the cytoplasm and the nucleus is unknown. *See* PLANT GROWTH.

Differentiation. The range of differentiation of cells in a highly evolved plant is not as great as the comparable range in a given higher animal, and the complexity of the organization of cell types into tissues and organs is not as great. In part, the differences relate to the higher degree of localization of a great range of metabolic processes, including those of the stimulus-response, circulatory, and extracellular digestive systems of higher animals. There is nonetheless a vast number of different sorts of plants and an enormous degree of diversification in cell structure and metabolism and in the various end products of metabolism. *See* PLANT ANATOMY; PLANT MORPHOGENESIS.

Dividing Cell

Despite innumerable early theories to the contrary, it has long been apparent that plant and animal cells

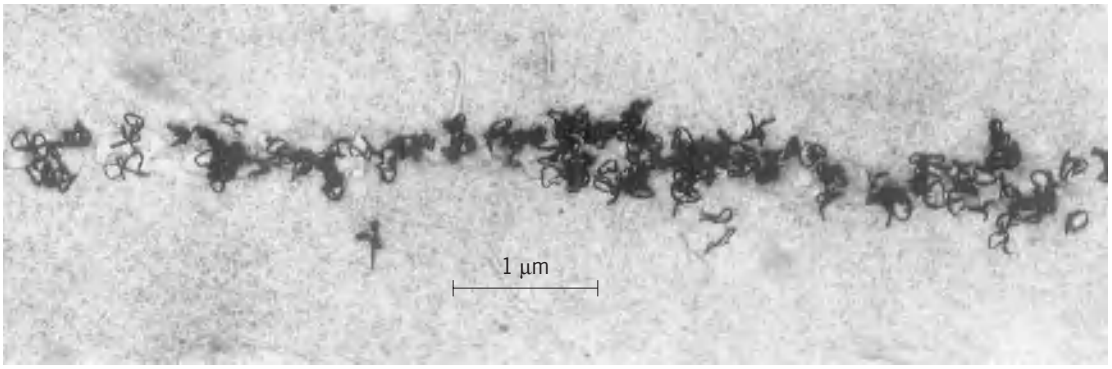

Fig. 12. Radioaudiograph of early stage in cell-plate formation in maize root cell treated with tritiated glucose. Silver grains at plate area result from glucose, transported via Golgi apparatus to the forming wall. (*Courtesy of T. P. Leffingwell*)

arise only by division of preexisting cells. This observation also pertains to some of the compartments of cells—the nucleus, the mitochondria, plastids, and probably the Golgi apparatus. The manner in which other components, such as the endoplasmic reticulum and vacuoles, behave in cell division is less clear.

Cell division begins with the constitution of the chromatin in the nucleus into the chromosomes. As the chromosomes are resolved, the nucleoli either disappear or are substantially reduced. There follows a breakdown of the nuclear envelope and the development of what is called a mitotic figure. *See* CELL DIVISION; CHROMOSOME.

The mitotic figure includes so-called spindle fibers, some of which are attached to the chromosomes and some of which pass between the chromosomes through the equator of the spindle (**Fig. 11**). The spindle fibers are of the same order of magnitude as, and seem to have a chemical composition comparable to, that of the microtubules. In mitosis the most conspicuous difference between plant and animal cells is the absence of centrioles in all but a few types of lower plant cells. In the animal cell two centrioles, one located at each end of the spindle, appear to represent "anchor points," because the spindle fibers seem to radiate from the centrioles. The spindle figure of plant cells appears to be firmly anchored without any demonstrable centrioles. *See* MITOSIS.

During mitosis the chromosomes are separated toward the poles of the spindle figure. At each pole a new nucleus is organized along with reconstitution of the nuclear envelope and redevelopment of the nucleoli in association with specific chromosome sites, known as nucleolar organizer regions. During mitosis the segmented nuclear envelope appears to surround the spindle figure. The manner in which the nuclear envelope is reconstituted is not known.

The partition of the cytoplasm into two masses, one surrounding each new nucleus, is a vastly different process in most plant cells from that in most animal cells. With few exceptions, this cytoplasmic division (cytokinesis) takes place in plants by a process known as cell-plate formation; a plate-shaped structure consisting of cell-wall material bounded by new plasma membrane appears in the mass of cytoplasm in the center of the cell and extends outward, ultimately connecting with the original walls. The initial stage in cell-plate formation is the movement of small vesicles produced by the Golgi apparatus into the spindle figure of the dividing cell and their clustering around spindle fibers. These small vesicles appear to form a group of islands so aligned as to determine the plane of cell division. Then larger vesicles produced by the Golgi apparatus move into this plane, and both types of vesicles fuse, with some as yet undefined involvement of the endoplasmic reticulum. The Golgi apparatus produces new plasmalemmas and initial wall components. The latter are composed largely of pectic substances and hemicelluloses. Radioautographic experiments have shown that glucose purposely incorporated into the cell during cell-plate formation collects in the Golgi apparatus, where it becomes part of one or more of these wall substances and from where it moves directly into the new wall (**Fig. 12**). Golgi apparatus are frequently seen clustered around the edges of the cell plate (**Fig. 13**). When the plate is fully extended to the existing walls of the cell, cytokinesis is complete, with the two new cells being separated by both plasma membrane and wall.

In most cases the mitochondria and plastids, and possibly the Golgi apparatus, appear to increase in number more or less concurrently with the division of the cell. Each is apparently also capable of increasing without cell division, for cell differentiation following the period of cell growth is often characterized by increases in the number of one or more of these organelles.

Segments of endoplasmic reticulum and the various small organelles and inclusions seem to be randomly distributed between the daughter cells, but not enough is known about developmental activities of these components to make any conclusive statements about the pattern of their distribution in cell division. There is substantial evidence that certain classes of compounds, including the so-called plant hormones, exert control over cell division, but the mechanism by which this control is exerted has not been clarified. There is little evidence concerning the means by which certain cells are caused to divide while others do not, and the determination of

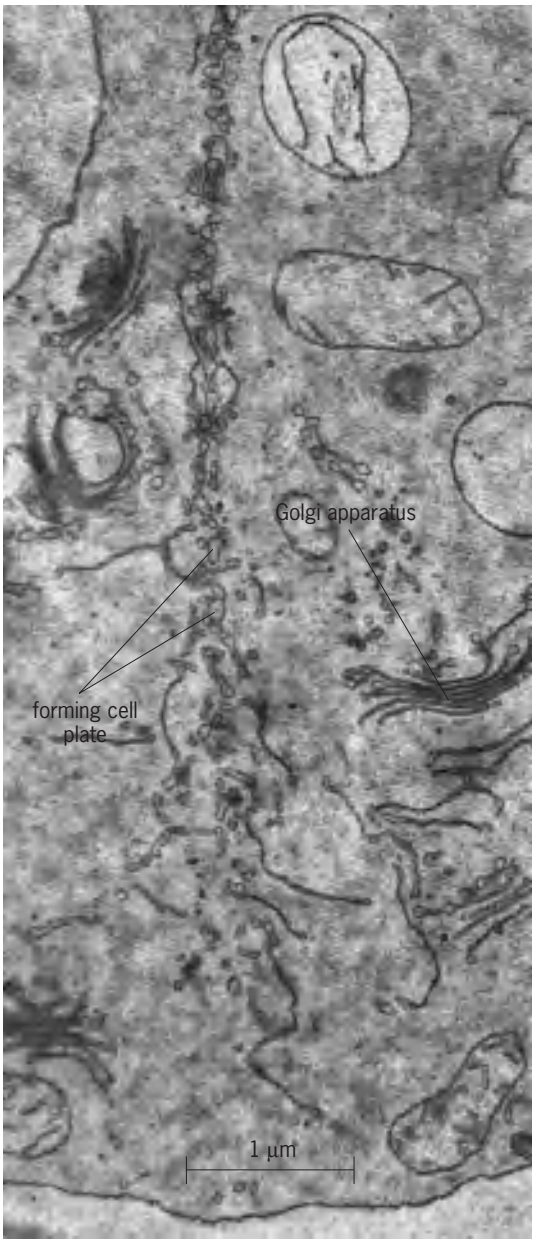

Fig. 13. A stage of cell-plate development showing the close association of the Golgi apparatus with the forming cell plate. (*Courtesy of H. H. Mollenhauer*)

the plant form. It is clear that cell reproduction involves not only the structural developments treated here but also significant metabolic changes.

W. Gordon Whaley

Plant Protoplasts

Although the plant cell is the basic unit of structure and function in nearly all plants, the term protoplast has been used to describe the organized entity of the living components of the plant cell which carry out active metabolism, biosynthesis, and energy transfer, in contrast to the extracellular, essentially metabolically inactive secondary plant matter. Plasmolysis studies on plant cells established that a functional semipermeable plasma membrane surrounds the protoplast.

For many years plant protoplasts remained largely cytological curiosities of little interest to physiologists, biochemists, and geneticists. However, the introduction of an enzymatic method for the isolation of protoplasts in 1960 stimulated renewed interest. Cell wall–degrading enzymes (cellulases) were employed, and large numbers of protoplasts could be isolated from various plant organs and also from plant tissue cultures. Pectinases were later additionally utilized to facilitate cell separation and wall degradation. A typical preparation of protoplasts isolated enzymatically from leaves of *Coleus* is shown in **Fig. 14**.

Experimental value. One of the most useful activities exhibited by protoplasts is that of cell wall regeneration. Studies on cell wall regeneration by protoplasts have provided a wealth of new information, particularly on the early stages of wall synthesis and the early organization of the wall. Protoplasts offer an important biological entity for many physiological and biochemical studies. For instance, chloroplasts isolated from protoplasts possess excellent structural and biochemical integrity, and fractionation of protoplasts has supplied large quantities of mature plant vacuoles. Plant protoplasts have also provided a system for plant virus research via laboratory culture.

Cell division. Once cell wall regeneration has become initiated, protoplasts cultured in suitable media under sterile conditions exhibit mitosis and cytokinesis similar to that normally found in cultured cells undergoing division. The behavior of protoplasts in this respect is extremely variable, ranging from failure even to divide following wall regeneration, to the ready establishment, at high frequency, of cell colonies and whole fertile plant regeneration. This is an area where much work is still required; ignorance of the control mechanisms for cell division in many species, particularly crop species, is delaying the use of protoplasts for crop improvement.

Genetic manipulation. The use of protoplasts for genetic manipulations centers on two major areas: the use of protoplasts in transformation studies,

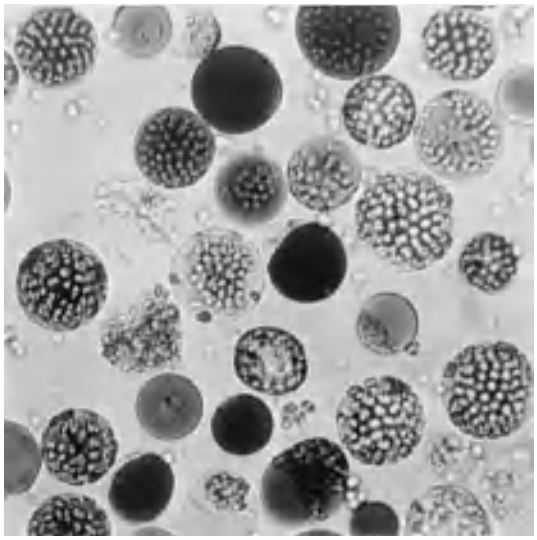

Fig. 14. Protoplasts isolated from *Coleus* leaves.

and obtaining new genetic variability by protoplast fusion.

Transformation. Protoplasts provide an ideal single-cell system for these basic studies in plant cell biology. Plasmids can be incubated with protoplast suspensions, and a microbiological approach can be used to follow plasmid interactions. If regeneration of whole plants from transformed protoplasts is possible, then a sexual analysis of somatic transformation will also be possible. It has been demonstrated that isolated *Agrobacterium* plasmids, incubated with *Petunia* protoplasts in the presence of poly(L-ornithine) [a known stimulator of virus uptake by protoplasts], are transformed and become capable of growing and dividing in the absence of growth substances, and of synthesizing crown gall tumor–specific opines, such as octopine.

Fusion. Fusion of protoplasts appears to be initiated when the plasma membranes are able to come sufficiently close to one another to enable membrane flow to take place. Membrane fusion requires that the membranes first be brought into apposition at molecular distances of 1 nm or less. Most protoplasts are negatively charged, and the addition of the fusion-inducing agent (or fusogen) probably eliminates this charge. They no longer repel each other, and apposition at molecular distances, with consequent membrane flow, can take place with resultant fusion of the plasma membranes. Calcium ions at high pH (9.5–10.5) and poly(ethylene glycol) are the most commonly used fusogens for protoplasts. Within the *Petunia* genus, sexually incompatible species can be crossed by this process of protoplast fusion and, coupled with the selective culture of somatic hybrid cells, into whole flowering hybrid plants. Horticulturally, this new somatic amphidiploid hybrid between *P. parodii* and *P. parviflora* is of interest, because it is suitable for the introduction of the hanging-basket habit of the sexually isolated *P. parviflora* species into the cultivated *Petunia*. *See* BREEDING (PLANT); CELL (BIOLOGY); SOMATIC CELL GENETICS.

E. C. Cocking

Bibliography. A. Fahn, *Plant Anatomy*, 4th ed., 1990; J. L. Hall, *Plant Cell Structure and Metabolism*, 2d ed., 1986; N. Harris and K. J. Oparka (eds.), *Plant Cell Biology: A Practical Approach*, 1994; B. King (ed.), *Cell Biology*, 1986; F. B. Salisbury and C. W. Ross, *Plant Physiology*, 4th ed., 1992.

Plant communication

Movement of signals or cues, presumably chemical, among individual plants or plant parts. These chemical cues are caused by damage to plant tissues or other stresses and stimulate physiological changes in the undamaged "receiving" plant or tissue. Communication among plants often involves induced plant defenses, which are chemicals produced actively by an individual plant in response to attack by insects or disease agents. *See* CHEMICAL ECOLOGY.

Plants produce a wealth of secondary metabolites that do not function in the main, or primary, metabolism of the plant, which includes photosynthesis, nutrient acquisition, and growth. Since many of these chemicals have very specific negative effects on animals or pathogens, ecologists speculate that they may be produced by plants as defenses. Plant chemical defenses either may be present all of the time (constitutive) or may be stimulated in response to attack (induced). Those produced in response to attack by pathogens are called phytoalexins. In order to demonstrate the presence of an induced defense, the chemistry of plant tissues or their suitability to some "enemy" (via a bioassay) must be compared before and after real or simulated attack. Changes found in the chemistry and suitability of the control or unattacked plants when nearby plants are damaged imply that some signal or cue has passed from damaged to undamaged plants. Responses in undamaged plants are related to the proximity of a damaged neighbor. *See* PHYTOALEXINS; PLANT GROWTH; PLANT METABOLISM.

Nature of the response. Responses by undamaged plants or plant tissues to other plants include altered metabolism, reduced quality for herbivores, and emission of damage-related volatiles that attract predators or parasites of the attacking herbivores. The main defensive change in plants exposed to volatiles is production of secondary metabolites that can act as defenses against insects or microbes. Evidence is accumulating to suggest that volatile signals often "prime" (accelerate or enhance) responses to subsequent attacks rather than stimulating defenses directly.

Nature of the cues. The most frequently proposed cue is methyl jasmonate (meJA), a volatile molecule produced by oxidation of lipids in almost all plants. The meJA influences a wide range of physiological activities in plants, including resistance to microbes and insects, senescence, growth, reproduction, and synthesis of many defensive chemicals. It is produced by most plant species in response to wounding, insect attack, pathogen infection, and other stresses. Blocking meJA production or perception can deprive plants of acquired immunity. The meJA regulates the expression of genes for enzymes in the biosynthetic pathways for secondary metabolites. Many other lipid oxidation products, as well as small, physiologically active isoprenoid and terpenoid compounds, are produced by plants; studies of damaged tissue have identified dozens of possibilities. How these volatile signals are perceived by plants has not been determined.

Ecological and evolutionary significance. Evidence that plant communication occurs frequently in the field or significantly protects plants from enemies remains inconclusive because of the difficulty in sampling volatile molecules in open air. Presumably, defensive responses cued in advance or primed by nearby plants could increase an individual plant's resistance to subsequent attack. If effective, such a response would reduce the spread or intensity of a widening infestation in a plant community.

Because wind currents dilute and redirect emitted volatiles, only a very large infestation, such as one that affects many plants or entire trees, seems likely to produce signals that are uniform or strong

enough to cue whole undamaged plants. It is more likely that cues may contact some tissues and not others, which may explain the broad chemical heterogeneity among plant parts and individuals. The nonuniformity of cue reception also suggests that communication-induced resistance is unlikely to be observed easily in nature.

The evolutionary origins and significance of the communication phenomenon are difficult to decipher. Plants control the composition of volatiles emitted when tissues are damaged; however, damage-specific emissions are most often directed at attracting protective predators or parasites. Volatiles such as methyl jasmonate or ethylene may be no more than necessary metabolic by-products of responses to attack and may not have evolved originally as signals among plants. There would be little or no advantage accruing to individual plants sending cues to unrelated neighbors unless those neighbors are genetic relatives. Hence, communication would be most useful to and frequent among clonal species or species with small genetic neighborhoods in which an individual is surrounded by relatives, including offspring. Yet, both ethylene and methyl jasmonate appear to influence the metabolism of many plants, regardless of sender or receiving species. *See* ETHYLENE; PLANT HORMONES.

There is obvious advantage accruing to plants that are able to respond to nearby damage in a way that increases their resistance to enemies. This is likely if the nearby damage is caused by some threat to the receiver, which would be true if neighbors are of the same species or share a common enemy. As a result, communication would tend to be beneficial in single-species stands or where generalist enemies (natural enemies with a broad range of diet, for example, aphids, snails, etc.) are important. Recent research has shown that the parasitic plant, dodder, locates host plants via host-emitted volatile cues.

Response to external cues may be more important within individual plants than between them. Although some induced (single individual) responses to attack are known to be cued by internal factors, it is possible that self-cuing could be external as well. Both ethylene and jasmonate are important signals coordinating within-plant responses to damage or infection. Communication among plant individuals thus may be an incidental consequence of contact with a self-cuing mechanism. *See* ALLELOPATHY; PHYSIOLOGICAL ECOLOGY (PLANT); PLANT PATHOLOGY; PLANT PHYSIOLOGY. Jack C. Schultz

Bibliography. A. A. Agrawal, S. Tuzun, and E. Bent (eds.), *Induced Plant Defenses Against Pathogens and Herbivores*, APS Press, St. Paul, MN, 1999; I. T. Baldwin et al., Volatile signaling in plant-plant interactions: "Talking trees" in the genomics era, *Science*, 311(5762):812–815, 2006; F. Baluska, S. Mancuso, and D. Volkmann (eds.), *Communication in Plants: Neuronal Aspects of Plant Life*, Springer-Verlag, Berlin, 2006; N. Dudareva et al., Plant volatiles: Recent advances and future perspectives, *Crit. Rev. Plant Sci.*, 25(5):417–440, 2006; M. Heil and C. Kost, Priming of indirect defences, *Ecol. Lett.*, 9(7):813–817, 2006; R. Karban and I. T. Baldwin, *Induced Responses to Herbivory*, University of Chicago Press, 1997; A. Kessler et al., Priming of plant defense responses in nature by airborne signaling between *Artemisia tridentata* and *Nicotiana attenuata*, *Oecologia*, 148(2):280–292, 2006; J. B. Runyon, M. C. Mescher, and C. M. De Moraes, Volatile chemical cues guide host location and host selection by parasitic plants, *Science*, 313(5795):1964–1967, 2006.

Plant evolution

The process of biological and organic change within the plant kingdom by which the characteristics of plants differ from generation to generation. Understanding of the course of plant evolution is based on several lines of evidence. The main levels (grades) of evolution have long been clear from comparisons among living plants, but the fossil record has been critical in dating evolutionary events and revealing extinct intermediates between modern groups, which are separated from each other by great morphological gaps (evolutionary changes in many characters). Plant evolution has been clarified by cladistic methods for estimating relationships among both living and fossil groups. These methods attempt to reconstruct the branching of evolutionary lines (phylogeny) by using shared evolutionary innovations (for example, presence of a structure not found in other groups) as evidence that particular organisms are descendants of the same ancestral lineage (a monophyletic group, or clade).

Many traditional groups are actually grades rather than clades; these are indicated below by names in quotes. Convergence, where the same new feature evolves in different lines, and reversal, where an older condition reappears, confuse the picture by producing conflicts among characters. These are resolved by using as many characters as possible and applying the principle of parsimony, that is, seeking the tree that requires the fewest total character changes, aided by a computer. Progress has also come from studies of molecular data [deoxyribonucleic acid (DNA) sequences of the same gene in different groups]; bases at any point along the DNA can change like any other kind of character and can be analyzed in the same way.

Most botanists restrict the term plants to land plants, which invaded the land after 90% of Earth history (see **illus.**). There is abundant evidence of photosynthetic life extending back 3.5 billion years to the early Precambrian, in the form of microfossils resembling cyanobacteria (prokaryotic blue-green algae) and limestone reefs (stromatolites) made by these organisms. Larger cells representing eukaryotic "algae" appear in the late Precambrian, followed by macroscopic "algae" and animals just before the Cambrian. *See* ALGAE; EUKARYOTAE; FOSSIL; PROKARYOTAE.

Origin of land plants. Cellular, biochemical, and molecular data place the land plants among the "green algae," specifically the "charophytes," which resemble land plants in their mode of cell division and differentiated male and female gametes

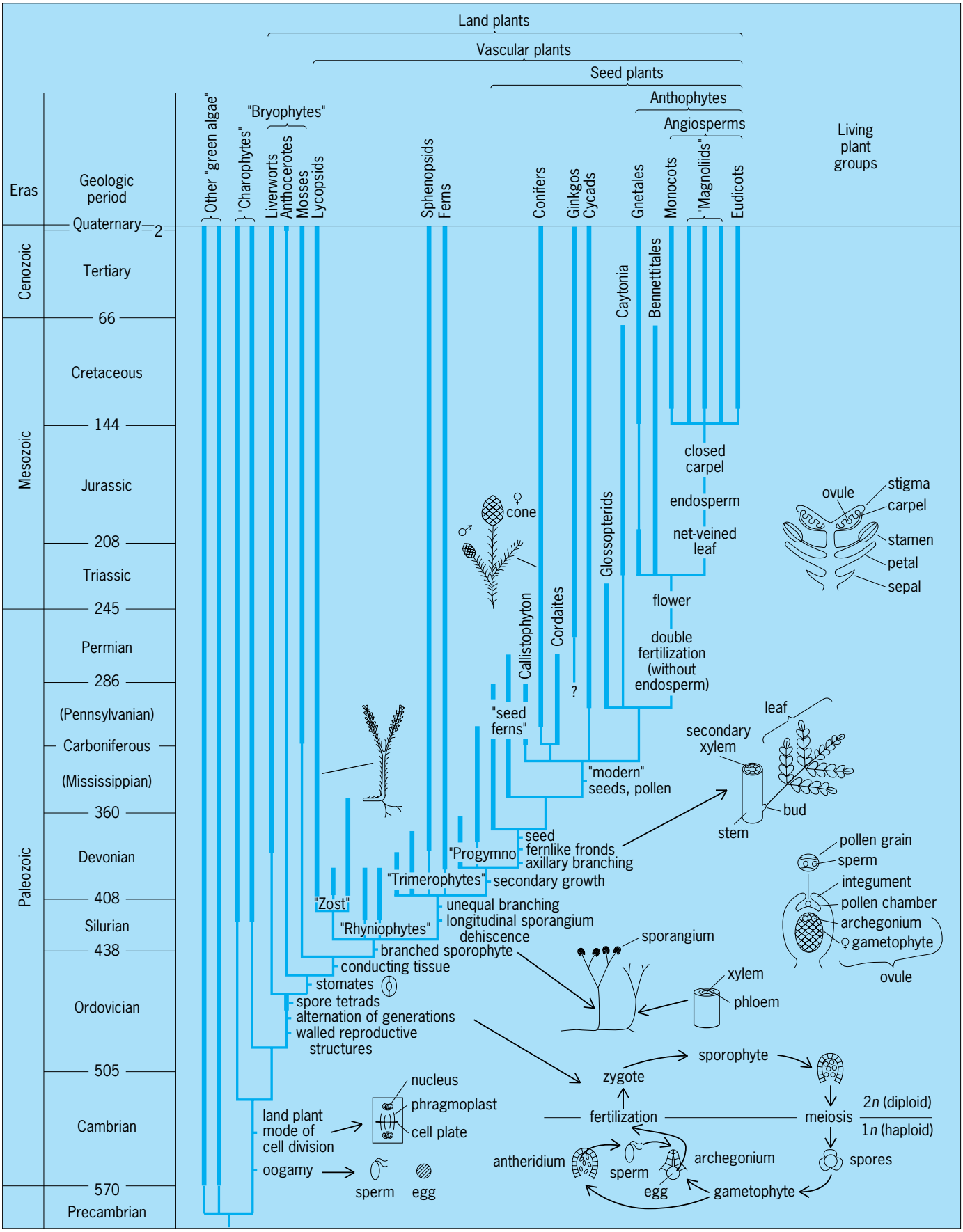

Inferred evolutionary relationships of major groups of plants, including some of their most important evolutionary innovations, and their distribution in geologic time (thick lines: known geologic range). Groups in quotes represent grades (levels of evolution).

(oogamy). Land plants themselves are united by a series of innovations not seen in "charophytes," many of them key adaptations required for life on land. "Charophytes" have a haploid life cycle, in which the zygote undergoes meiosis, and single-celled reproductive structures; but land plants have an alternation of generations, with a haploid, gamete-forming (gametophyte) and diploid, spore-forming (sporophyte) phase. Their reproductive organs (egg-producing archegonia, sperm-producing antheridia, and spore-producing sporangia) have a protective layer of sterile cells. The sporophyte, which develops from the zygote, begins its life inside the archegonium. The spores, produced in fours by meiosis, are air-dispersed, with a resistant outer wall that prevents desiccation. *See* CHAROPHYCEAE; PLANT REPRODUCTION.

Land plants have been traditionally divided into "bryophytes" and vascular plants (tracheophytes). These differ in the relative role of the sporophyte, which is subordinate and permanently attached to the gametophyte in "bryophytes" but dominant and independent in vascular plants. In vascular plants, tissues are differentiated into an epidermis with a waxy cuticle that retards water loss and stomates for gas exchange, parenchyma for photosynthesis and storage, and water- and nutrient-conducting cells (xylem, phloem). However, cladistic analyses imply that some "bryophytes," namely anthocerotes (hornworts) and mosses, are closer to vascular plants than others (liverworts, or hepatics). Both groups resemble vascular plants in having stomates, and mosses also have conducting cells, although of a more primitive type. This implies that the land-plant life cycle originated before the full suite of vegetative adaptations to land life, and that the sporophyte began small and underwent a trend toward elaboration and tissue specialization. *See* EPIDERMIS (PLANT); PHOTOSYNTHESIS; PRIMARY VASCULAR SYSTEM (PLANT).

In the fossil record, the first recognizable macroscopic remains of land plants are Middle Silurian vascular forms with a branched sporophyte, known as "rhyniophytes." These differed from modern plants in having no leaves or roots, only dichotomously branching stems with terminal sporangia. However, spore tetrads formed by meiosis are known from older beds (Middle Ordovician); these may represent more primitive, bryophytic plants. *See* BRYOPHYTA; RHYNIOPHYTA.

In one of the most spectacular adaptive radiations in the history of life, vascular plants diversified through the Devonian. At the beginning of this period, vegetation was low [perhaps at most 1 ft (30 cm) tall] and probably confined to wet areas, but by the Late Devonian, size had increased in many lines, resulting in large trees and forests with shaded understory habitats. Of the living groups of primitive vascular plants, the lycopsids (club mosses), with pointed, one-veined leaves (microphylls), branched off first, along with the extinct "zosterophyllopsids," which were leafless but resembled lycopsids in having lateral sporangia. A second line, the "trimerophytes," which showed incipient differentiation of a main trunk and side branches, and terminal sporangia with longitudinal dehiscence, gave rise to sphenopsids (horsetails), with simple leaves in a whorled arrangement, and ferns (filicopsids), with pinnately compound leaves (fronds) derived from whole branch systems. This radiation culminated in the coal swamp forests of the Late Carboniferous, with tree lycopsids (Lepidodendrales), sphenopsids (*Calamites*), and ferns (Marattiales). Remains of these plants make up much of the coal of Europe and eastern North America, which were then located on the Equator; these plants had many peculiar anatomical features apparently requiring wet and constant conditions. *See* LYCOPODIALES; MARATTIALES; SPHENOPHYTA.

Seed plants. Perhaps the most significant event after the origin of land plants was evolution of the seed. Primitive seed plants ("gymnosperms") differ from earlier groups in their reproduction, which is heterosporous (producing two sizes of spores), with separate male and female gametophytes packaged inside the pollen grain (microspore), and the ovule (a sporangium with one functional megaspore, surrounded by an integument, which develops into the seed). The transfer of sperm (two per pollen grain) from one sporophyte to another through the air, rather than by swimming between gametophytes living on or in the soil, represents a step toward independence from water for reproduction. This step is comparable to the evolution of the amniote egg in vertebrates, and it must have helped plants invade drier areas than they had previously occupied. In addition, seed plants have new vegetative features, particularly secondary growth, which allows production of a thick trunk made up of secondary xylem (wood) surrounded by secondary phloem and periderm (bark). Together, these innovations have made seed plants the dominant organisms in most terrestrial ecosystems ever since the disappearance of the Carboniferous coal swamps. *See* ECOSYSTEM; PTERIDOSPERMS; SEED.

A major breakthrough in understanding the origin of seed plants was recognition of the "progymnosperms" (for example, *Archaeopteris*) in the Middle and Late Devonian. These plants, which were the first forest-forming trees, had secondary xylem, phloem, and periderm, but they still reproduced by spores, implying that the anatomical advances of seed plants arose before the seed. Like sphenopsids and ferns, they were apparently derived from "trimerophytes." The earliest seed plants of the Late Devonian and Carboniferous, called "seed ferns" because of their frondlike leaves (a convergence with true ferns), show steps in origin of the seed, by fusion of branchlets surrounding the megasporangium into the integument. Origin of the typical mode of branching in seed plants, from buds in the axils of the leaves, occurred at about the same time. *See* PTERIDOSPERMS.

Among seed plants, coniferopsids (fossil cordaites, living conifers, and possibly ginkgos), with fan-shaped to needlelike leaves, have often been considered an independent line of evolution from "progymnosperms." However, cladistic analyses indicate

that coniferopsids were derived from "seed ferns" with coniferlike pollen and seeds, similar to the Late Carboniferous genus *Callistophyton*. This event may have involved the elimination of fronds and their replacement by scale leaves, which occur around the buds in all seed plants. Cordaites and conifers both appear in the Late Carboniferous; comparisons of their early members indicate that the seed-bearing cone of conifers is a complex structure, in which the cone scales are actually modified branches. *See* CORDAITALES; PINALES.

Seed plants became dominant in the Permian during a shift to drier climate and extinction of the coal swamp flora in the European-American tropical belt, and glaciation in the Southern Hemisphere Gondwana continents. Early conifers, with small, needlelike leaves, predominated in the tropics; extinct glossopterids, with simple, deciduous, net-veined leaves, inhabited Gondwana. Moderation of climate in the Triassic coincided with the appearance of new seed plant groups, including the living cycads and ginkgos (today reduced to one species, *Ginkgo biloba*) and the extinct Bennettitales (cycadeoids), with cycadlike pinnate leaves but flowerlike reproductive structures, as well as more modern ferns. Many Mesozoic groups show adaptations for protection of seeds against animal predation, while flowers of the Bennettitales constitute the first evidence for attraction of insects for cross-pollination, rather than transport of pollen by wind.

Angiosperms. The last major event in plant evolution was the origin of angiosperms (flowering plants), the seed plant group that dominates the modern flora. Angiosperms show numerous innovations that allow more rapid and efficient reproduction. The flower, typically made up of protective sepals, attractive petals, pollen-producing stamens, and ovule-producing carpels (all considered modified leaves), favors more efficient pollen transfer by insects. The ovules are enclosed in the carpel, so that pollen germinates on the sticky stigma of the carpel rather than in the pollen chamber of the ovule. The carpels (separate or fused) develop into fruits, which often show special adaptations for seed dispersal. Other advances include an extreme reduction of the gametophytes, and double fertilization whereby one sperm fuses with the egg, and the second sperm with two other gametophyte nuclei to produce a triploid, nourishing tissue called the endosperm. Angiosperms also developed improved vegetative features, such as more efficient water-conducting vessels in the wood and leaves with several orders of reticulate venation. These features may have contributed to their present dominance in tropical forests, previously occupied by conifers with scale leaves. *See* RAINFOREST.

The origin of angiosperms has been considered a great mystery of plant evolution, since angiosperms were thought to appear in diverse, modern forms during the Early Cretaceous, with no obvious links to other groups. However, most botanists believed that the most primitive living angiosperms are "magnoliid dicots," based on their "gymnosperm"-like pollen (with one aperture for germination), wood anatomy, and flower structure. Studies of Cretaceous fossil pollen, leaves, and flowers confirm this view by showing a rapid but orderly radiation beginning with "magnoliid"-like and monocotlike types, followed by primitive eudicots (with three pollen apertures), some related to sycamores and lotuses. *See* MAGNOLIOPHYTA.

There is still a gap between angiosperms and other groups, but both morphological and molecular data imply that angiosperms are monophyletic and most closely related to Bennettitales and Gnetales, a seed plant group that also radiated in the Early Cretaceous but later declined to three living genera. Gnetales also experience a type of double fertilization, but without formation of endosperm, an exclusively angiospermous innovation. Since all three groups have flowerlike structures, suggesting that the flower and insect pollination arose before the closed carpel, they have been called anthophytes. The closest relatives of anthophytes are controversial, whether advanced "seed ferns" (*Caytonia,* glossopterids) or coniferopsids. These relationships, plus problematical Triassic pollen grains and macrofossils with a mixture of angiospermlike and more primitive features, suggest that the angiosperm line goes back to the Triassic, although perhaps not as fully developed angiosperms. Within angiosperms, it is believed that "magnoliids" are relatively primitive, monocots and eudicots are derived clades, and wind-pollinated temperate trees such as oaks, birches, and walnuts (amentiferae) are advanced eudicots. However, "magnoliids" include both woody plants (for example, magnolias and laurels) and herbs (for example, waterlilies and peppers), and their flowers range from large, complex, and insect-pollinated to minute, simple, and wind-pollinated. These extremes are present among the earliest Cretaceous angiosperms, and cladistic analyses disagree on which is most primitive.

Although plant extinctions at the end of the Cretaceous have been linked with radiation of deciduous trees and proliferation of fruits dispersed by mammals and birds, they were less dramatic than extinctions in the animal kingdom. During the Early Tertiary, climates were still much milder than today, but mid-Tertiary cooling led to contraction of the tropical belt and expansion of seasonal temperate and arid zones. These changes led to the diversification of herbaceous angiosperms, such as composites and grasses, and the origin of open grassland vegetation, which stimulated the radiation of hoofed mammals, and ultimately the invention of human agriculture. *See* AGRICULTURE; FLOWER; PALEOBOTANY; PLANT KINGDOM.

James A. Doyle

Bibliography. C. B. Beck, *Origin and Evolution of Gymnosperms*, 1988; E. M. Friis, W. G. Chaloner, and P. R. Crane, *The Origins of Angiosperms and Their Biological Consequences*, 1987; L. E. Graham, *Origin of Land Plants*, 1993; P. M. Richardson, Origin and relationships of the major plant groups, *Ann. Missouri Bot. Grad.*, 81:403–567, 1994; W. N. Stewart and G. W. Rothwell, *Paleobotany and the Evolution of Plants*, 1993.

Plant geography

The study of the spatial distributions of plants and vegetation and of the environmental relationships which may influence these distributions. Plant geography (or certain aspects of it) is also known as phytogeography, phytochorology, geobotany, geographical botany, or vegetation science.

History. Plant geography arose as a result of the great botanical voyages of the early 1800s, especially those by Alexander von Humboldt, sometimes called the father of plant geography. Early work involved classification of physiognomically and ecologically similar plant types, and so plant geography and plant ecology remained somewhat indistinguishable. Around 1900 the study of vegetation began to split into plant ecology (which focused more on process) and plant geography (including vegetation geography). Two main perspectives had arisen: a historical perspective concerned with migration, dispersal, and the historical development of floras; and an environmental perspective concerned with environmental constraints and ecological relations influencing plant and vegetation distributions.

Floristic plant geography. A flora is the collection of all plant species in an area, or in a period of time, independent of their relative abundances and relationships to one another. The species can be grouped and regrouped into various kinds of floral elements based on some common feature. For example, a genetic element is a group of species with a common evolutionary origin; a migration element has a common route of entry into the territory; a historical element is distinct in terms of some past event; and an ecological element is related to an environmental preference. Aliens, escapes, and very widespread species are given special treatment. An endemic species is restricted to a particular area, which is usually small and of some special interest. The collection of all interacting individuals of a given species, in an area, is called a population.

The idea of area is fundamental to plant geography. An area is the entire region of distribution or occurrence of any species, element, or even an entire flora.

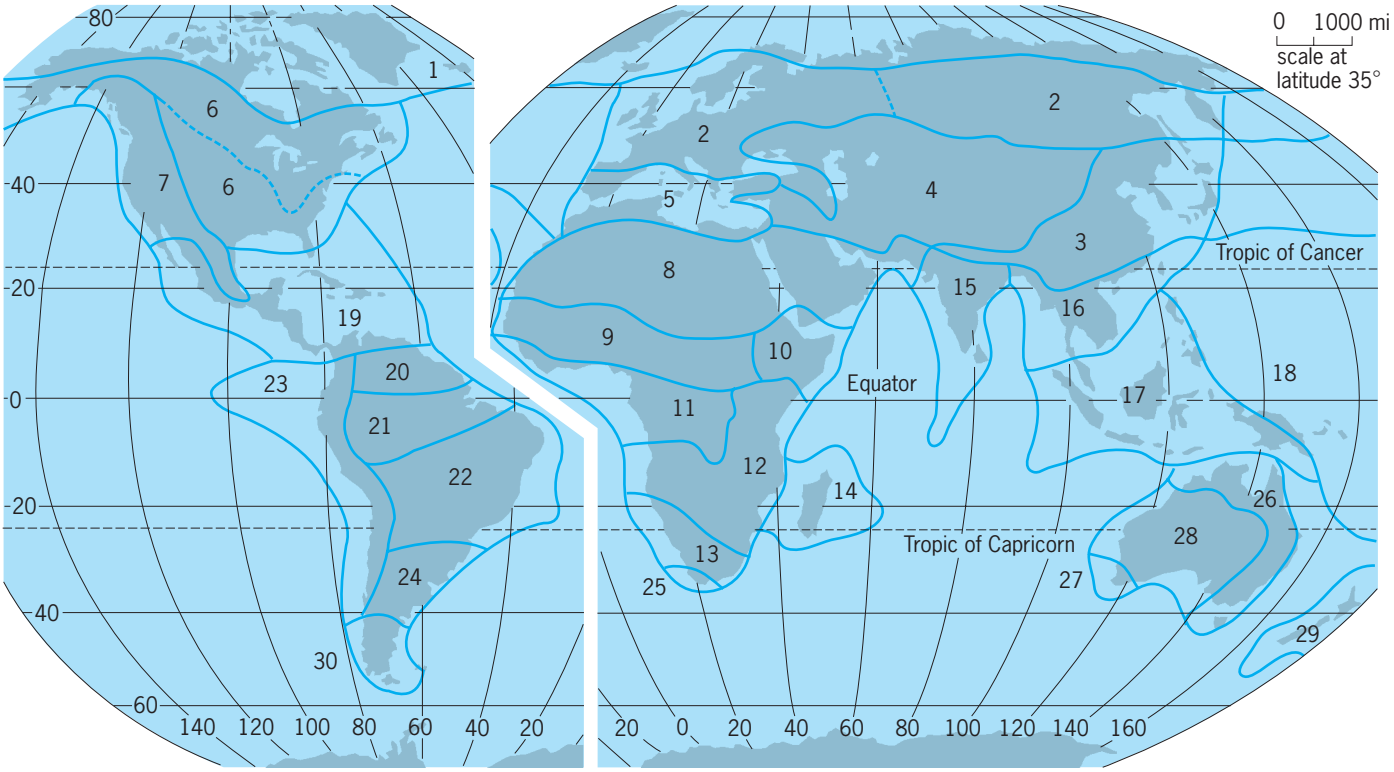

1 = Arctic and subarctic
2 = Euro-Siberian
a. Europe
b. Asia
3 = Sino-Japanese
4 = Western and central Asiatic
5 = Mediterranean
6 = Atlantic North America
a. northern
b. southern
7 = Pacific North America
8 = African–Indian desert
9 = Sudanese park steppe
10 = Northeastern African highland
11 = Western African rainforest
12 = Eastern African steppe
13 = South African
14 = Madagascar
15 = Indian
16 = Continental southeastern Asiatic
17 = Malaysian
18 = Melanesia and Micronesia
19 = Caribbean
20 = Venezuela and Guiana
21 = Amazon
22 = South Brazilian
23 = Andean
24 = Pampas
25 = Cape
26 = Northern and eastern Australian
27 = Southwestern Australian
28 = Central Australian
29 = New Zealand
30 = Patagonian

Fig. 1. Floristic regions of world. 1 mi = 1. 6 km. (*After R. Good, Geography of Flowering Plants, 3d ed., Longmans, 1964*)

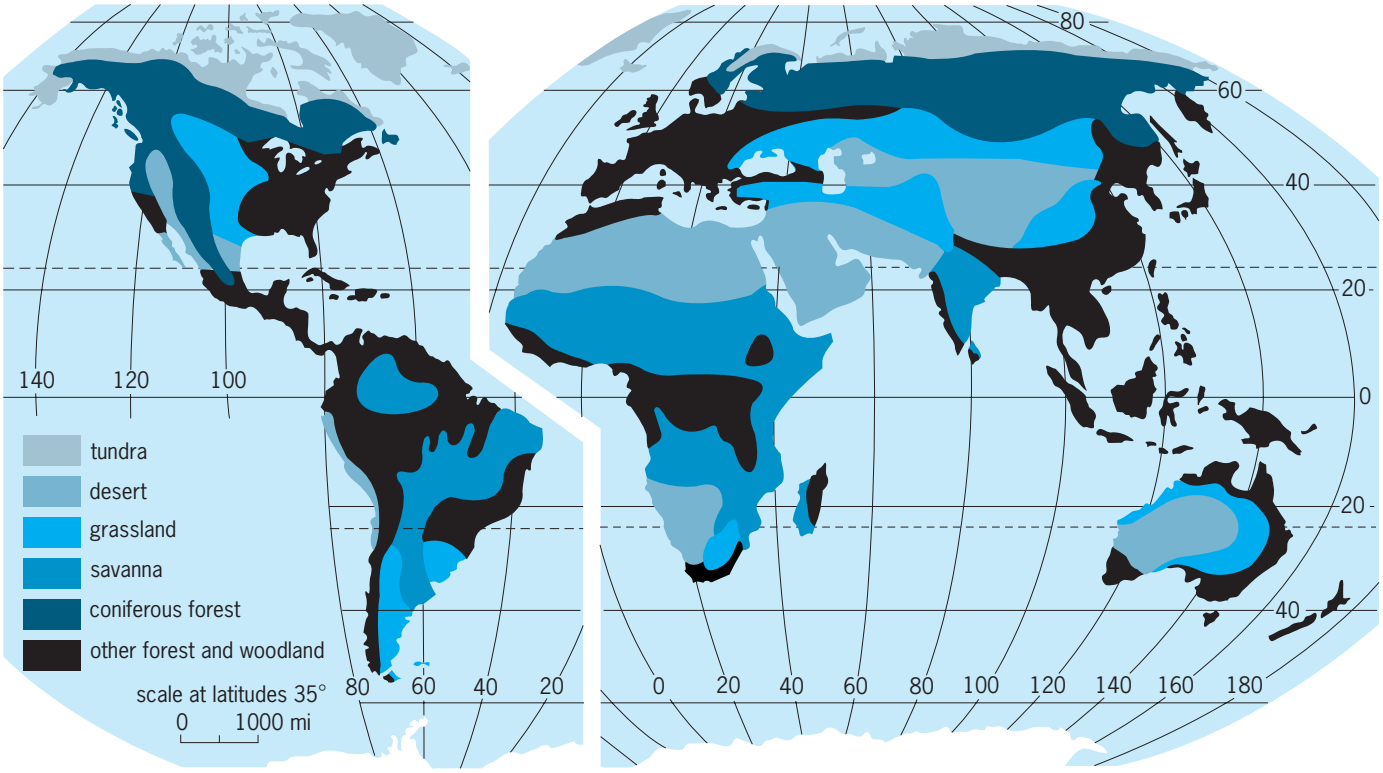

Fig. 2. Map of the world, showing the distribution of physiognomic vegetation types. 1 mi = 1. 6 km. (*After R. Good, Geography of Flowering Plants, 3d ed., Longmans, 1964*)

The description of areas is the subject of areography, while chorology studies their development. The local distribution within the area as a whole, as that of a swamp shrub, is the topography of that area. Areas are of interest in regard to their general size and shape, the nature of their margin, whether they are continuous or disjunct, and their relationships to other areas. Closely related plants that are mutually exclusive are said to be vicarious (areas containing such plants are also called vicarious). A relict area is one surviving from an earlier and more extensive occurrence. On the basis of areas and their floristic relationships, the Earth's surface is divided into floristic regions (**Fig. 1**), each with a distinctive flora.

Floras and their distribution have been interpreted mainly in terms of their history and ecology. Historical factors, in addition to the evolution of the species themselves, include consideration of theories of shifting continental masses, changing sea levels, and orographic and climatic variations in geologic time, as well as theories of island biogeography, all of which have affected migration and perpetuation of floras. The main ecological factors include the immediate and contemporary roles played by climate (temperature, water availability, wind), soil, animals, and humans. *See* ISLAND BIOGEOGRAPHY; PALEOBOTANY; PALEOECOLOGY.

Vegetational plant geography. Vegetation, a term of popular origin, refers to the mosaic of plant life found on the landscape. The vegetation of a region has developed from the numerous elements of the local flora but is shaped also by nonfloristic physiological and environmental influences. Vegetation is an organized whole, at a higher level of integration than the separate species, composed of those species and their populations. Vegetation may possess emergent properties not necessarily found in the species themselves. Sometimes vegetation is very weakly integrated, as pioneer plants of an abandoned field. Sometimes it is highly integrated, as in an undisturbed tropical rainforest. Vegetation provides the main structural and functional framework of ecosystems, which have been actively studied since 1950. *See* ECOSYSTEM.

Plant communities are an important part of vegetation and are studied in plant ecology and various subfields (such as phytosociology, plant sociology, vegetation science, or phytocoenology). No definition of the plant community has gained universal acceptance. In part, this is because of the high degree of independence of the species themselves. Thus, the community is often only a relative social continuity in nature, bounded by a relative discontinuity, as judged by competent botanists. *See* ECOLOGICAL COMMUNITIES.

In looking at vegetation patterns over larger areas, it is the basic physiognomic distinctions between grassland, forest, and desert, with such variants as woodland (open forest), savanna (scattered trees in grassland), and scrubland (dominantly shrubs), which are most often emphasized (**Fig. 2**). These general classes of vegetation structure can be broken down further by reference to leaf types (such as broad or needle) and seasonal habits (such as evergreen or deciduous). Geographic considerations may complete the names of the main vegetation

formation types, also called biomes (such as tropical rainforest, boreal coniferous forest, or temperate grasslands). Such natural vegetation regions are most closely related to climatic patterns and secondarily to soil or other environmental factors. *See* ALTITUDINAL VEGETATION ZONES.

Vegetational plant geography has emphasized the mapping of such vegetation regions and the interpretation of these in terms of environmental (ecological) influences. Distinction has been made between potential and actual vegetation, the latter becoming more important due to human influence. *See* VEGETATION AND ECOSYSTEM MAPPING.

Some plant geographers and other biologists are dissatisfied with these more general approaches and point to the effects of ancient human populations, natural disturbances, and the large-herbivore extinctions and climatic shifts of the Pleistocene on the species composition and dynamics of so-called virgin vegetation. On the other hand, it has been shown that the site occurrence and geographic distributions of plant and vegetation types (that is, above the species level) can be predicted surprisingly well from general climatic and other environmental patterns. Problems occur especially when substrates are unusually young or nutrient-poor and in marginal environments where disturbance can emphasize the stochastic nature of vegetation, permitting one or the other of competing types to gain the upper hand. Unlike floristic botany, where evolution provides a single unifying principle for taxonomic classification, vegetation structure and dynamics have no single dominant influence.

Plant growth forms. Basic plant growth forms (such as broad-leaved trees, stem-succulents, or forbs) have long represented convenient groups of species based on obvious similarities. When these forms are interpreted as ecologically significant adaptations to environmental factors, they are generally called life forms and may be interpreted as basic ecological types. These life forms not only may represent the basic building blocks of vegetation but may also provide a convenient way of describing vegetation structure without having to list each individual species. *See* PLANTS, LIFE FORMS OF.

In general, basic plant types may be seen as groups of plant taxa with similar form and ecological requirements, resulting from similar morphological responses to similar environmental conditions. For example, deciduous leaves are generally "softer" and photosynthesize more efficiently in favorable environments. On the other hand, they may lose more water and require more energy and nutrients for their construction (over the plant's life-span) than "harder," longer-lived evergreen leaves. Similarly, larger plants with larger total leaf area may be vulnerable to greater water loss, but may also have more extensive root systems for more effective water uptake. The vegetation of a particular site will be composed of plants with particular combinations of such form characters which permit the plant to function successfully in a particular environment. When similar morphological or physiognomic responses occur in unrelated taxa in similar but widely separated environments, they may be called convergent characteristics (for example, the occurrence of broad-sclerophyll shrubs in the world's five mediterranean-climate regions).

Research. Work in plant geography has involved a variety of approaches. J. Grime recognized three basic plant "strategies" (competitors, stress tolerators, and ruderals), their environmental relationships, and how these may interact. As a result of much work in ecology, there has been an increased emphasis on plant processes and ecophysiology. The focus especially is on plant water and energy budgets throughout the year and how these may influence or limit species distributions. Such work has been carried out, in particular, in nonforest vegetation of grassland, tundra, semidesert, and mediterranean-climate regions. *See* PHYSIOLOGICAL ECOLOGY (PLANT).

The revolutionary discovery of plate tectonics in the twentieth century raised many new questions about past plant migrations and the historical development of taxa and regional floras. The value (and validity) of cladistics and of vicariance theory in the study of plant geographic history has been strongly debated.

As human populations alter or destroy more and more of the world's natural vegetation, problems of species preservation, substitute vegetation, and succession have increased in importance. This is especially true in the tropics, where deforestation is proceeding rapidly. Probably over half the species in tropical rainforests have not yet even been identified. Because nutrients are quickly washed out of tropical rainforest soils, cleared areas can be used for only a few years before they must be abandoned to erosion and much degraded substitute vegetation. Perhaps the greatest current challenge in plant geography is to understand tropical vegetation and succession sufficiently well to design self-sustaining preserves of the great diversity of tropical vegetation. *See* BIOGEOGRAPHY; ECOLOGY; RAINFOREST. Elgene O. Box

Bibliography. E. O. Box, *Macroclimate and Plant Forms: An Introduction to Predictive Modeling in Phytogeography*, 1981; R. Daubenmire, *Plant Geography, with Special Reference to North America*, 1978; J. P. Grime, *Plant Strategies, Vegetation Processes, and Ecosystem Properties*, 2d ed., 2001; J. O. Rieley and S. Page, *Ecology of Plant Communities*, 1990; D. Tilman, *Plant Strategies and the Dynamics and Structure of Plant Communities*, 1988.

Plant growth

An irreversible increase in the size of the plant. As plants, like other organisms, are made up of cells, growth involves an increase in cell numbers by cell division and an increase in cell size. Cell division itself is not growth, as each new cell is exactly half the size of the cell from which it was formed. Only when it grows to the same size as its progenitor has growth

been realized. Nonetheless, as each cell has a maximum size, cell division is considered as providing the potential for growth. *See* CELL (BIOLOGY).

Comparison with animals. While growth in plants consists of an increase in both cell number and cell size, animal growth is almost wholly the result of an increase in cell numbers. Another important difference in growth between plants and animals is that animals are determinate in growth and reach a final size before they are mature and start to reproduce. Plants have indeterminate growth and, as long as they live, continue to add new organs and tissues. In animals growth of the different parts of the body is more or less simultaneous; in plants growth is restricted to the growing points or meristems. Therefore, in an animal most body cells attain about the same age, and the individual dies as a unit; but in a plant new cells are produced all the time, and some parts such as leaves and flowers may die, while the main body of the plant persists and continues to grow. The basic processes of cell division are similar in plants and animals, though the presence of a cell wall and vacuole in plant cells means that there are certain important differences. This is particularly true in plant cell enlargement, as plant cells, being restrained in size by a cellulose cell wall, cannot grow without an increase in the wall. Plant cell growth is thus largely a property of the cell wall. *See* ANIMAL GROWTH.

Changes in rate of growth. If the size of a plant (or any organism), plant part, or single plant cell is plotted against time, the resulting line produces an S-shaped, or sigmoid, growth curve (**Fig. 1**). This may have an initial lag phase, and then a phase of steadily increasing growth rate. During this latter phase, termed the log phase, the logarithm of the size plotted against time yields a straight line, since each unit of previous growth gains its own growth increment. The growth rate of an individual stem or root then stabilizes at an approximately constant rate. This is because a growing plant usually has the same number of dividing, growing, and maturing cells in its stem, and always maintains the same proportion of growing cells in its different stages. Thus, the regular growth rate of the whole plant is the integration of thousands of sigmoid growth curves of the individual cells. As the plant matures, the growth rate may slow down or stop completely, to be followed by senescence, or those changes which precede the death of the plant or plant part.

Cell division. Cell division in plants, as in animals, involves the division of the nucleus by mitosis. The various organelles, such as proplastids (precursors of chloroplasts), probably also divide. Following the formation of two daughter nuclei, a new cell wall starts to form between the two nuclei at a right angle to the plane of cell division. The wall starts as a coalescence of vesicles (membrane-bounded spheres) produced by an organelle in the cell termed the Golgi apparatus or dictyosome. These vesicles contain the necessary chemicals for cell wall formation, and as they coalesce the contents polymerize to form the initial layer of the new cell wall, which is termed the middle lamella and composed mainly of calcium pectate. (Calcium is an essential nutrient for plants, and lack of calcium prevents the formation of the middle lamella so that the apex dies.) The membranes of the vesicles become the membranes delimiting the new cell. The new cell wall starts to form across the plane between the two new cells and fuses with the old wall at the edges. Further deposition of cell wall polysaccharides into the new wall occurs through continued fusion of Golgi vesicles with the new wall, and perhaps also by direct synthesis on or just below the cell membrane lining the new wall. *See* CELL DIVISION; CELL WALLS (PLANT); MITOSIS; PLANT CELL.

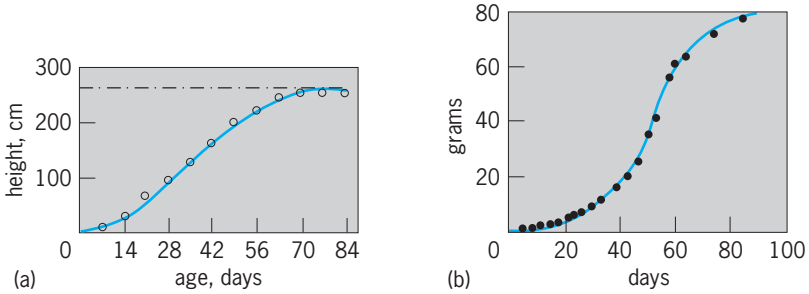

Fig. 1. Graphs showing growth curves. (*a*) Growth in height of sunflower stem; 1 cm = 0.4 in. (*after E. P. Odum, Fundamentals of Ecology, Saunders, 1953*). (*b*) Growth of entire corn plant as measured by increase in dry weight; 1 g = 0.035 oz (*after D. W. Thompson, On Growth and Form, Cambridge, 1942*).

Cell division in plants takes place in discrete zones called meristems (**Fig. 2**). In these meristems the cells divide and grow back to their original size before dividing again. Just beyond this region of cell division a much greater increase in cell size takes place. The principal meristems are at the tip (apex) of the stem and root (or each branch of the stem and root); in the vascular or conducting tissue of the stem and root (the vascular cambium) where the meristem is shaped like a hollow cylinder; and at the exterior surface of woody plants (the cork cambium). Cell division in the stem and root apices occupies only a small zone of about 0.04–0.08 in. (1–2 mm) at the tip of the stem or root of a herbaceous plant. As the cells are continuously produced at the stem and root apices, the daughter cells steadily emerge from the zone of cell division and start to elongate. Elongation itself is also confined to a region just below the tip and may extend for about 1.2–1.6 in. (3–4 cm) below the stem bud and only about 0.2–0.4 in. (0.5–1 cm) behind the root tip in herbaceous plants such as a pea plant (Fig. 2). *See* APICAL MERISTEM; LATERAL MERISTEM.

The stem and root apical meristems produce all the primary (initial) tissues of the stem and root. The cylindrical vascular cambium produces more conducting cells at the time when secondary thickening (the acquisition of a woody nature) begins. The vascular cambium is a sheet of elongated cells which divide to produce xylem or water-conducting cells on the inside, and phloem or sugar-conducting cells on the outside. Unlike the apical meristems whose cell division eventually leads to an increase in length of the stem and root, divisions of the vascular cambium occur when the part of the plant has reached a fixed length, and lead only to an increase in girth,

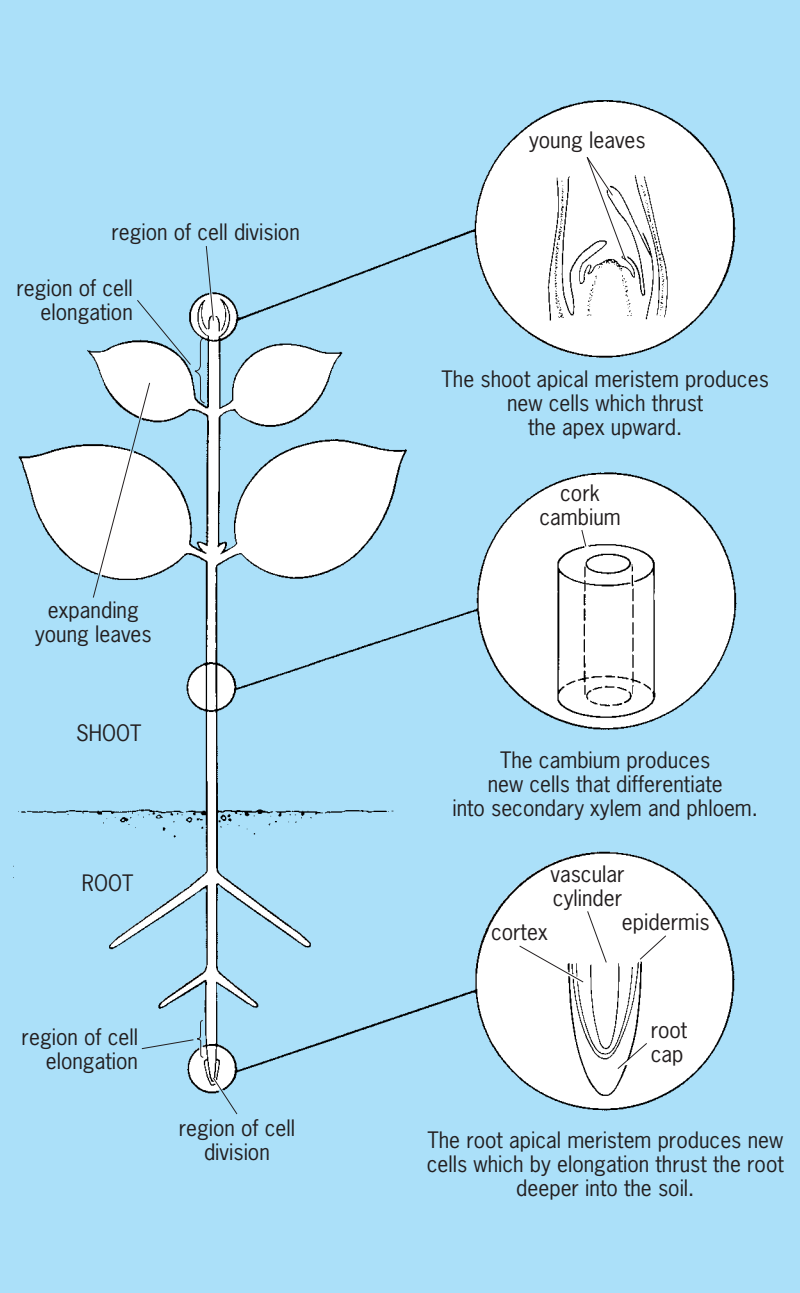

Fig. 2. Schematic drawing of a plant showing the regions of cell division and cell elongation.

not in length. The final meristematic zone, the cork cambium, is another cylindrical sheet of cells on the outer edge of older stems and roots of woody plants. It produces only new outer cells, which differentiate into the corky layers of the bark so that new protective layers are produced as the tree increases in circumference. *See* BUD; PERIDERM; ROOT (BOTANY); STEM.

Cell division, followed by enlargement, also takes place in certain areas of the plant not strictly regarded as meristems. Cell division continues for a short time in the newly formed leaves following their production on the stem apex. Fruits and seeds are also active areas of cell division and growth, and will be discussed later.

Cell enlargement. The young meristematic plant cell is approximately isodiametric. It has a thin cell wall composed of cellulose (made of glucose units) and other mixed polysaccharide polymers (containing xylose, glucose, galactose, arabinose, and other minor sugar constituents, as well as some protein). The interior contains dense protoplasmic contents indicative of intense metabolic activity and protein synthesis. Unlike mature plant cells, vacuoles are usually absent or very small. As this cell grows, it does not increase equally in all dimensions, but usually elongates to a final length many times its width and 20–150 times its original size. This cell expansion is primarily due to water uptake. Protein synthesis occurs during cell growth, but at a rate insufficient to keep up with the increase in cell size. As a result, the cytoplasm which originally filled the cell now comes to occupy only the periphery of the cell (**Fig. 3**), and the cell becomes largely filled with a huge central vacuole. This vacuole is bounded by a membrane called the tonoplast, and contains a water solution of salts, sugar, and various secondary metabolites that form a reserve of metabolites and a depository for waste products, as well as a reservoir for compounds involved in the defense of the plant against diseases and predators. The vacuole arises initially as many small vacuoles from either Golgi vesicles or the endoplasmic reticulum (a membrane system throughout the cytoplasm). These then coalesce to form the large central vacuole (**Fig. 4**). Movement of water into the cell increases the size of the vacuole.

While the length of the cell greatly increases, the thickness of the wall remains constant and its density does not change. This is because wall synthesis occurs at the same time as normal elongation. The dry weight of the cell, 80% of which can be accounted

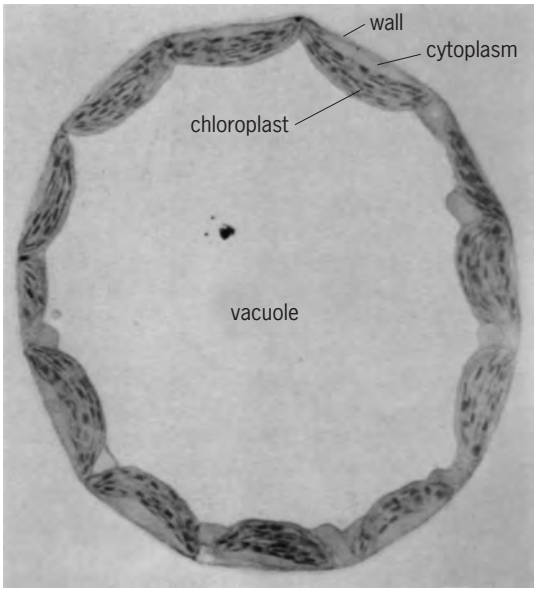

Fig. 3. Electron micrograph of a cell in a leaf. Expansion of the central vacuole is the principal change in cell content during growth. (*From A. W. Galston, P. J. Davies, and R. L. Satter, The Life of the Green Plant, 3d ed., Prentice-Hall, 1980*)

for by the wall, increases from the apex to a point where cell expansion ceases, and it is largely due to a synthesis of wall polysaccharides. After cell expansion is completed, there is a slow further increase in the wall through the deposition of secondary wall material.

The rate of cell extension can be very rapid. In a corn (*Zea*) root, increases of 40% per hour occur, while in the grass *Phleum* 100% per hour has been noted. Such rates occur, however, only in a narrow cell zone, and while the cells are expanding they do so at an average rate of about 20% per hour. The extent to which growth is visible in the whole plant depends on the rate of cell growth and the size of the extending zone. Corn plants are often noted to grow several centimeters per day, and leaf stalks of the water lily *Victoria regia* may increase from 3.6 to 27 in. (9 to 68 cm) in 24 h.

Elongation of a plant cell is dependent on the extension of the cell wall. Because the cell forms a rigid "wooden box" around the cell, there can be no increase in cell size unless the wall stretches. In plant cells there is always a stretching force pushing outward against the cell wall. This force, termed turgor pressure, results when the presence of dissolved solutes in the vacuole causes water uptake by the cell through osmosis. However, despite this pressure inside all cells, only cells in a narrow growing zone elongate. This is because wall extension is restricted and can occur only in the presence of certain growth hormones, most notably auxin and in some cases gibberellin.

During cell elongation the tangential stress on a cell is much greater than the longitudinal stress, yet the cell enlarges predominantly in a longitudinal direction. The reason for this unidirectional enlargement is the cellulose microfibril construction of the cell wall which, when stretchable, permits elongation with little lateral increase.

Factors affecting growth. Plant growth is affected by internal and external factors. The internal controls are all the product of the genetic instructions carried in the plant. These influence the extent and timing of growth and are mediated by signals of various types transmitted within the cell, between cells, or all around the plant. Intercellular communication in plants may take place via hormones (or chemical messengers) or by other forms of communication not well understood. Plants lack the nervous system of animals; yet transmembrane electrical potentials do occur, and action potentials may be transmitted from cell to cell, particularly in plants with rapidly moving parts, such as *Mimosa pudica* (the so-called sensitive plant) or insectivorous plants with actively moving traps, such as *Dionea* (Venus' flytrap). Cells may also be able to detect the presence of neighboring cells through pressure or gradients of various chemicals such as nutrients, though again little is known of the role of such signals. Something must, however, determine that central stem cells develop into pith, while those at the surface become epidermal cells. The genetic makeup of the plant, residing in the genes of the chromosomes in the nucleus, controls both the production of and the response to intercellular signals, as well as regulating development within the cell. Genes do not all operate at once; to do so would only produce chaos. They operate in both a predetermined and externally modulated sequence to control the growth and destiny of the cells, ultimately leading to the formation of the mature organism. *See* GENETICS; PLANT MORPHOGENESIS.

Fig. 4. Young parenchyma cells in culture which have begun to enlarge by development of one or more vacuoles in the cytoplasm. There are many small vacuoles in the upper cell, while vacuoles in the lower cell are coalescing to form a single large vacuole. Note the fusion of two vacuoles at X, and breakdown of the membranes and cytoplasm between them. (*From A. W. Galston, P. J. Davies, and R. L. Satter, The Life of the Green Plant, 3d ed., Prentice-Hall, 1980*)

Internal controls—hormones. The prime means of communication among the different parts of the plant is through the plant hormones. There are several hormones (or groups of hormones), each of which may be produced in a different location, that have a different target tissue and act in a different manner. The groups of hormones that are definitively known are briefly described below. Their role in the integration of growth will be described later. *See* PLANT HORMONES.

Auxin. This hormone is primarily represented in nature by one chemical, indoleacetic acid (IAA). Some other indole derivatives exist in plants (such as indoleacetonitrile in cabbages), but these are generally thought to act as precursors of IAA. Several synthetic derivatives also exist, and these have practical uses in agriculture and horticulture. IAA has numerous roles in plants as a stimulator of cell growth or division or both. *See* AUXIN.

Gibberellin. The gibberellins are a family of isoprenoid compounds with over 50 naturally occurring members which differ slightly in their chemical structure. Each plant species has several gibberellins, though usually no more than 6–10, and they are thought to represent different steps in an interconversion pathway, with some of the gibberellins being

biologically active, while some are inactive. Some gibberellins vary in their activity in different species. Gibberellins also enhance cell division and elongation, but in a different way and in different cells from auxin. *See* GIBBERELLIN.

Cytokinins. These hormones are adenine derivatives. There are two main natural compounds, isopentenyl adenine and zeatin (so named because it was discovered in *Zea*, or maize), and both of these also exist in their riboside and riboside phosphate forms in addition to the free base. Cytokinins are primarily active in stimulating cell division. They may also induce cell enlargement but only in leaves, and they also act to delay senescence. *See* CYTOKININS.

Ethylene. Although unusual in that it is a gas, ethelyene is very active in inducing processes normally associated with senescence, particularly fruit ripening, flower fading, and leaf abscission. *See* ETHYLENE.

Abscisic acid. This is a single isoprenoid compound. It has a general growth-inhibiting action and may be involved in the induction of dormancy. It is also synthesized in stomatal guard cells (cells surrounding pores in the leaf epidermis) under conditions of water shortage, when it causes the stomata to close, restricting further loss. *See* ABSCISIC ACID.

External controls—the environment. The external environments of the root and shoot place constraints on the extent to which the internal controls can permit the plant to grow and develop. Prime among these are the water and nutrient supply available in the soil. Because cell expansion is controlled by cell turgor, which depends on water, any deficit in the water supply of the plant reduces cell turgor and limits cell elongation, resulting in a smaller plant. As each cell can expand only for a certain interval of time before its new size is fixed by changes in the cell wall properties, any shortage of water, even if temporary, will result in reduced growth if it exceeds the time period over which any cell expands. Water shortage also reduces cell division, but this is resumed when the water supply returns, provided the shortage is not too prolonged.

It is, in fact, quite common for plants to experience a water deficit during the daytime, because the rate of water loss through transpiration will exceed water uptake on a hot dry day even if there is plenty of water available in the soil. This deficit is, however, made up at night when transpiration ceases, though water uptake continues. Over such a period the cell walls remain stretchable so that as the deficit is made up during the night, and the turgor pressure recovers, there is a spurt of growth to about the same extent as would have occurred in a 24-h period with no water deficit. *See* PLANT-WATER RELATIONS.

Mineral nutrients are needed for the biochemical processes of the plant. When these are in insufficient supply, growth is less vigorous, or in extreme cases it ceases altogether. *See* PLANT MINERAL NUTRITION.

An optimal temperature is needed for plant growth. The actual temperature range depends on the species. Temperate species often grow best at moderate temperatures (for example, 59–68°F or 15–20°C for peas), while tropical species are adapted to higher temperatures nearer 86°F (30°C). In general, metabolic reactions and growth increase with temperature, though high temperature becomes damaging. Most plants grow slowly at low temperatures (32–50°F or 0–10°C), and some tropical plants are damaged or even killed at low but above-freezing temperatures. Many temperate-zone plants can survive subfreezing temperatures if they are acclimated by slowly declining temperatures and short day lengths prior to the freezing period. When actively growing, however, virtually all plants are frost-sensitive. *See* COLD HARDINESS (PLANT).

Light. Light is important in the control of plant growth. It drives the process of photosynthesis which produces the carbohydrates that are needed to osmotically retain water in the cell for growth. These carbohydrates form the basis for all the organic compounds that are produced by the plant and comprise the protoplasm and cell walls. In addition, the carbohydrates are oxidized to produce the energy for metabolism and cell maintenance, both of which are needed for growth. Similarly, carbon dioxide, the raw material for photosynthesis, is needed for plant growth. Under conditions of low light or CO_2, enhancement of growth can be obtained by increasing the levels of either component. *See* PHOTOSYNTHESIS; PLANT RESPIRATION.

Phytochrome. The color of incident light is also important in determining plant growth. While chlorophyll, the pigment in photosynthesis, absorbs blue and red light, the plant uses another pigment called phytochrome to detect the light quality of its environment.

Phytochrome is a biliprotein (a protein with a straight-chain tetrapyrrole chromophore). It responds to red (660-nanometer) and far-red (731-nm) light by changing between two forms—a red-absorbing form (P_r) and a far-red-absorbing form (P_{fr}). While natural light does not consist of red and far-red flashes, the ratio of red to far-red in the incident light is changed by the chlorophyll in a leaf canopy. Sunlight is high in red light compared with far-red light and causes a high proportion of the phytochrome to exist in the P_{fr} form. By contrast, the chlorophyll in leaves absorbs the red while allowing the far-red to pass. Thus, under a leafy canopy the ratio of red to far-red is low, and the phytochrome in an understory plant will be predominantly in the P_r form. This change has profound effects on the plant. Many seeds need P_{fr} to germinate and will not do so beneath other plants because of the P_r level. A high P_r also causes etiolation so that the plant becomes tall and spindly (**Fig. 5**). Such enhanced stem growth, however, has adaptive advantage in speeding up the growth of the plant through the leafy canopy above into the sunlight needed for photosynthesis.

The photomorphogenetic effect of light absorbed by phytochrome is also intensity-dependent. The reason for this is thought to be the speed of cycling between the two forms of phytochrome which will increase as the light intensity increases. It is proposed that various intermediates between the two forms are also needed together with a high level of

Fig. 5. Effect of far-red light in promoting stem elongation in *Chenopodium album*. Both plants received an equal amount of white light, but the plant at the right received additional far-red light; therefore less of its phytochrome is in the P_{fr} form. (*From Morgan and Smith, Nature, 262:210–212, 1976*)

P_{fr} in order for there to be a maximum photomorphogenetic response. The typical plant form is therefore produced only in sunlight or in other high-intensity light that gives a high proportion of phytochrome in the P_{fr} form. The other role for phytochrome is the measurement of photoperiod in the control of numerous developmental phenomena which will be dealt with below. *See* PHOTOMORPHOGENESIS; PHYTOCHROME.

Control of cell division. In the absence of auxin and cytokinin, no cell division takes place. The function of auxin in cell division is unknown. However, there is some idea of the role of cytokinin. In the absence of cytokinins, DNA replication takes place in the nucleus, but the nucleus does not divide. When cytokinins are added, protein synthesis takes place rapidly and division ensues (possibly because proteins are needed for nuclear division). This protein synthesis appears to depend on preformed messenger ribonucleic acid (mRNA), and it has been proposed that cytokinins act by binding to ribosomes (the protein synthesis machinery), thereby allowing the translation of mRNA for cell division proteins.

Control of cell elongation. Under normal turgor pressure the cell wall is moderately elastic, but no permanent deformation takes place in the absence of auxin. If a dicotyledon (such as the pea) stem or grass (such as oat or corn) coleoptile segment is taken and placed in a flowing aerated solution in an apparatus such as that shown in **Fig. 6**, a slow growth will occur in the absence of external auxin. When auxin is added, there is often an immediate, slight, transitory decrease in the growth rate of the stem for unknown reasons. There is a lag period of 10–15 min following the application of the hormone, and then growth increases over a further 10–15 min to reach a new auxin-enhanced rate of growth 10–20 times the original rate.

Auxin does not work in all tissues in all species. In the case of mature grass stems, gibberellin enhances growth while auxin is without effect. In general, however, gibberellin produces hypoelongation of intact stems on a plant but has little effect on stem segments.

Auxin is needed for cell growth and acts to increase the plasticity (permanent stretchability) of the cell wall. Auxin does not, however, act directly on the wall, since it has no effect on dead cells which are mechanically held under tension. Active cellular metabolism is needed for auxin action, and the driving force for wall extension is the turgor pressure. If auxin is added to active stem cells in an osmotic medium which prevents cell elongation, and the stem cells are subsequently killed by freezing and thawing, then the cell walls will stretch when placed under tension. Thus some change has been brought about in the wall which is subsequently expressed

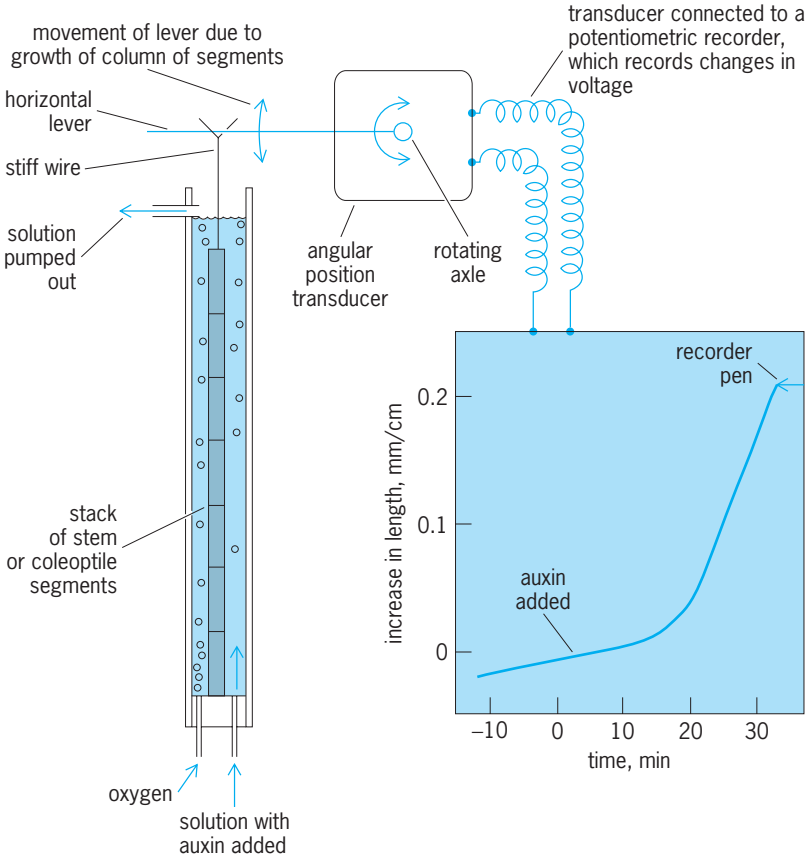

Fig. 6. Apparatus for recording rapid changes in growth, with a typical growth-tracing resulting from the addition of indoleacetic acid to a buffer at about pH 6.5. (*After A. W. Galston, P. J. Davies, and R. L. Satter, The Life of the Green Plant, 3d ed., Prentice-Hall, 1980*)

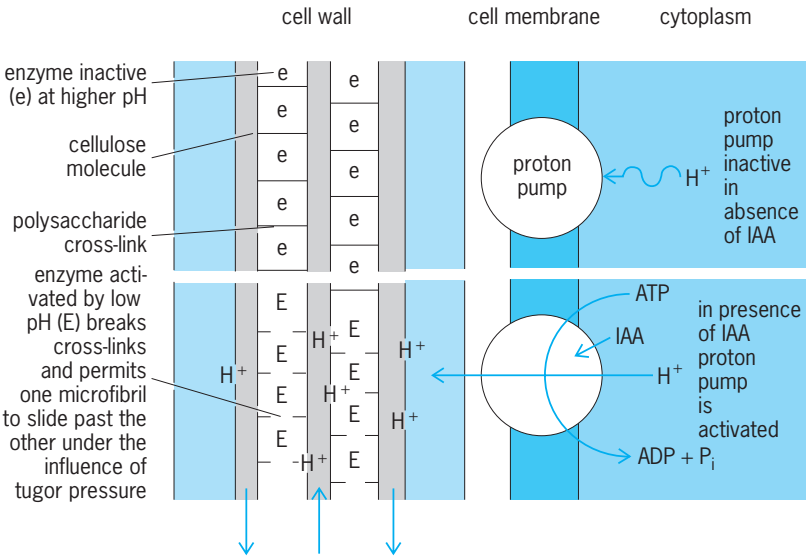

Fig. 7. Proton pump theory to explain the action of indoleacetic acid in inducing wall extensibility and cell elongation. (*After A. W. Galston, P. J. Davies, and R. L. Satter, The Life of the Green Plant, 3d ed., Prentice-Hall, 1980*)

as wall extension. The nature of this change became clear when it was found that the action of auxin, in both living cells and dead cells held under tension, can be mimicked by acid solution of pH 5. In addition, it was found that whan auxin is added to living cells, hydrogen ions (H^+) are pumped out of the cells into the cell walls, causing the walls to become acidic (**Fig. 7**). The increased acidity of the wall medium causes the enzymatic breaking of polysaccharide cross-links in the wall which normally hold the wall rigid. In this state the cellulose microfibrils can slide over one another so that the wall extends (**Fig. 8**). Resynthesis of the cross-links is subsequently needed to fix the new position of the microfibrils, and this is followed by auxin-stimulated cellulose synthesis so that the wall maintains its original thickness. The location of the auxin-promoted proton pump is currently under investigation. The cell membrane is the location of most such proton pumps, but auxin-binding studies seem to designate the endoplasmic reticulum membrane as the site of the auxin action.

RNA and protein synthesis are also needed for continued auxin-induced growth, and auxin enhances both these processes. The proteins needed for auxin-induced growth have been termed growth-limiting proteins, although their nature is as yet unknown. Possible roles for such proteins are as the proton pump or as enzymes needed for the breaking and resynthesis of the wall cross-links.

The action of gibberellin in inducing cell elongation may also involve wall acidification, although the acidification measured so far has been much less than that produced by auxin.

Vegetative growth. The pattern and extent of plant growth is the result of the various internal and external influences on the plant (**Fig. 9**). The roots supply water and nutrients to the shoots while the shoots supply photosynthates. Cytokinins are produced in the shoot tips, where they stimulate cell division, and also in the roots, from which they move to the shoot in the xylem stream. A supply of cytokinins is needed for leaf expansion so that when the root supply is reduced, by root pruning, for example, the resulting leaves are small, thus restricting the water loss. Cytokinins are also needed for the prevention of senescence. When a leaf ages, it can be rejuvenated by applying cytokinin. Since the cytokinins travel in the water stream from the root, it is likely that heavily shaded leaves, which lose less water, become deficient in cytokinins and thus senesce and are shed from the plant.

Auxin is produced primarily in the leaf primordia and young leaves of the bud. It moves basipetally downward through an active polar (unidirectional) transport mechanism primarily in the young xylem vascular cells and, in so doing, promotes the continued development of xylem cells up toward the

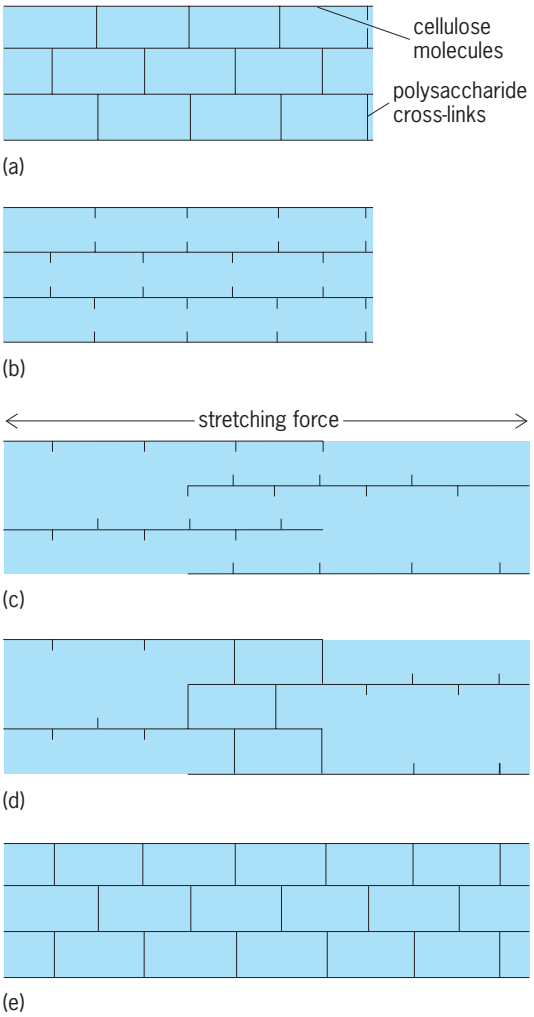

Fig. 8. Model suggesting mechanism of wall elongation under the influence of auxin. (*a*) Original wall structure. (*b*) Cross-links are broken and wall is loosened. (*c*) Under the influence of turgor pressure, cellulose molecules slide past each other, and the wall stretches. (*d*) Cross-links reform to fix the wall in its newly stretched position. (*e*) More wall material is synthesized. The wall is now elongated but still has its original thickness. The process is then repeated. (*After A. W. Galston, P. J. Davies, and R. L. Satter, The Life of the Green Plant, 3d ed., Prentice-Hall, 1980*)

growing bud. The auxin that reaches the root is mainly destroyed by an enzyme, IAA-oxidase, though some accumulates in the growing point of the root and the branch root initials where it promotes cell division.

Just below the bud, auxin promotes the elongation of the cells to produce stem growth. Gravity and light cause the transport of auxin toward the downward or dark side of the stem. The resultant higher concentration of auxin causes more rapid cell growth on that side and thus a bending toward the opposite side. Auxin transport may not be involved in tropistic stem movements, but evidence favors the possibility. *See* PLANT MOVEMENTS.

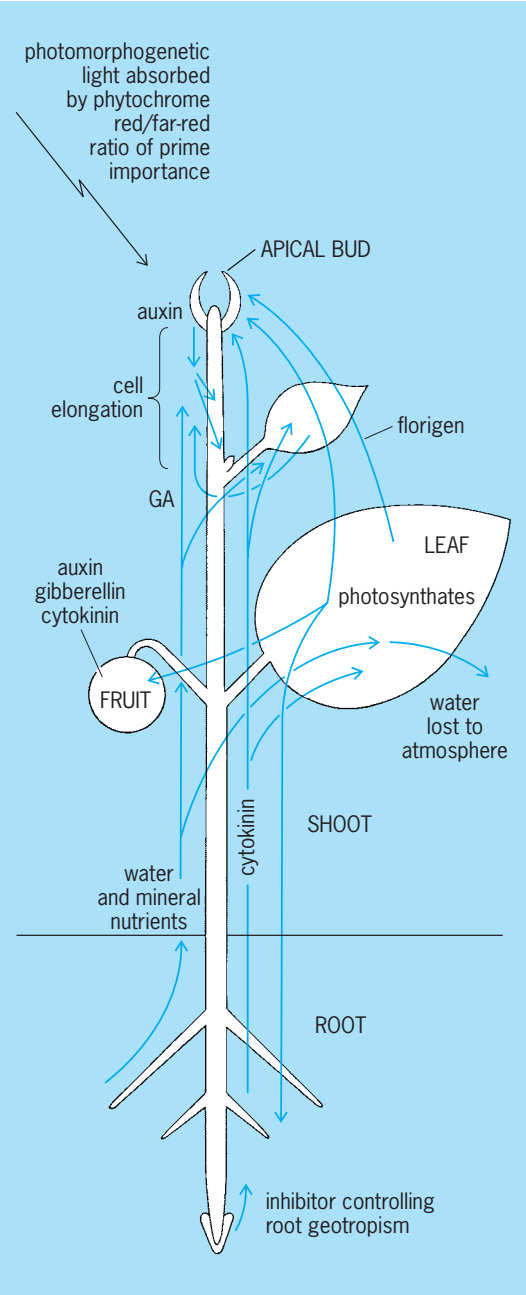

Fig. 9. Schematic drawing of a plant with the interrelationships of some factors involved in growth and development.

In the root, gravity causes a horizontal root tip to bend and grow downward. This is mediated by a growth inhibitor moving backward from the root cap to the growing zone of the root on the lower side of the horizontal root.

Another effect of auxin in the growing plant is the prevention of growth of lateral buds so that apical dominance exists. The lateral buds near the apical bud remain dominant, while those at a distance from the apical bud start to grow out. It is generally considered that it is auxin produced by the apical bud which inhibits the growth of the lateral buds and that it does so through the production of ethylene. As the distance between the apical and lateral bud increases through the growth of the leader, not only does the level of auxin at the lateral bud decrease, but cytokinins coming from the root oppose the auxin and cause the growth of the lateral branch. *See* APICAL DOMINANCE.

As auxin is transported morphologically downward, it accumulates at the base of stem cutttings. Here it promotes the formation of callus and root initials which then develop into roots so that a rooted cutting results. This can be enhanced by the application of auxin to the cut base of a shoot.

The site of gibberellin production is less definite. It is produced by young leaves and possibly by roots. It is involved primarily in stem growth. Many dwarf plants can be turned into tall plants by application of gibberellin. It also induces the bolting or rapid elongation of flowerstalks of biennial plants. Such plants grow as a rosette of leaves for most of their growth, and then the flowerstalk rapidly elongates just prior to flowering. It is probable that natural tall and bolting plants have higher levels of active gibberellins than dwarf or rosette plants, but because of the number of gibberellins present, each of which has a different biological activity, most of the exact gibberellin levels and changes have yet to be elucidated. In addition, the increase in gibberellins which lead to bolting is possibly of a transitory nature.

Growth of fruits and seeds. Fruits and seeds are rich sources of hormones. Initial hormone production starts upon pollination and is further promoted by ovule fertilization. These hormones promote the growth of both seed and fruit tissue. In some cases, application of the hormone to the deseeded or unfertilized fruit will cause the development of seedless or parthenocarpic fruit. The hormones vary with species: auxin induces fruit development in the strawberry, and gibberellins induce fruit development in grapes; pea seeds contain high levels of gibberellin; and developing corn grains are rich sources of cytokinins. Fruits grow initially by cell division, then by cell enlargement, and finally sometimes by an increase in air spaces. *See* FRUIT.

The growth of a seed starts at fertilization. A small undifferentiated cell mass is produced from the single-celled zygote. This proceeds to form a small embryo consisting of a stem tip bearing two or more leaf primordia at one end and a root primordium at the other. Either the endosperm or the cotyledons enlarge as a food store. Very specific seed storage

proteins are synthesized on endoplasmic reticulum–bound ribosomes, and the proteins are passed into the lumen of the endoplasmic reticulum to form storage-protein bodies. Messenger RNA for the enzymes of germination is also synthesized during the latter part of embryo development. The synthesis of storage proteins and the prevention of the translation of mRNA for the enzymes of germination are brought about by a rise in the level of abscisic acid. *See* PLANT REPRODUCTION; SEED.

Totipotency. It is possible to grow almost any plant species in tissue culture. This consists of taking a piece of sterile tissue and placing it on a jellylike agar medium containing sucrose, mineral nutrients, vitamins, and hormones. The piece of tissue can be meristematic, but fully differentiated parenchyma cells can also be made to grow. In general, growth occurs only if the medium contains both auxin and cytokinins. However, some tissues, notably tissues isolated from crown galls (whose cells have incorporated a plasmid from an infecting bacterium, though the bacterium may subsequently be absent), are able to synthesize their own hormones and grow vigorously with no further additions of hormones. Most tissues initially develop into an undifferentiated mass of cells, but manipulation of the hormone levels can cause the differentiation of roots and shoots, and ultimately a mass of small plantlets on the callus tissue. These plantlets can be removed and then grown into mature plants. This method is becoming important in the mass propagation of many horticultural species such as orchids and chrysanthemums.

If the callus tissue is transferred to a slowly agitated liquid medium, some single cells may slough off the callus, and in an appropriate hormonal regimen these cells start to develop as if they are fertilized zygotes going through all the stages of embryo development until they develop into plantlets.

The dedifferentiation of mature parenchyma, the subsequent differentiation into whole plants, and the formation of embryos and whole plants from single diploid cells without sexual fusion, show that during differentiation genes are simply turned on and off, rather than any loss of genetic material taking place. Under appropriate nutritional and hormonal signals, the genes can again be switched on and cause a previously mature cell to revert to an embryonic state. The ability of a mature cell to retain all the genetic instructions to build a whole plant is known as totipotency.

Flowering. At a certain time a vegetative plant ceases producing leaves and instead produces flowers. This often occurs at a particular season of the year. The determining factor for this event is day length (or photoperiod). Different species respond to different photoperiods.

Photoperiod classes. Three basic classes can be recognized, namely, short-day plants (such as *Chrysanthemum*, *Pharbitis*, and *Xanthium*), long-day plants (such as cereals and spinach), and day-neutral plants (such as tomato). Short-day plants flower if they receive a shorter duration of light than a certain critical photoperiod, while long-day plants flower if they receive more than the critical photoperiod. Day-neutral plants are not influenced by photoperiod and flower only when they reach a certain age. Some species have more complex day-length requirements, for example, a series of long days followed by a series of short days, which ensures that flowering occurs only in the autumn. Some long-day plants also have a prior cold requirement, termed vernalization, which is normally satisfied the previous winter.

It has been found that when plants respond to photoperiod, they are responding to the length of darkness rather than the length of light. If a short-day plant is given a long night, it flowers, but if that long night is interrupted by a short period of light, flowering is prevented. Experiments on the color of light which regulates flowering in darkness show that red is most effective and that the effect of red is inhibited by far-red. This demonstrates the involvement of the pigment phytochrome in the detection of photoperiod.

This detection occurs through the interaction of either the P_{fr} or the P_r with an endogenous rhythm in the plant. If a short-day plant is given a long dark period of several days interrupted at different times by a short period of light, the effect of the light depends on the time when it is given. If it is given at a time when the rhythm in the plant has, by previous exposures, entered a "light-anticipating" phase, then the light promotes flowering, but if given during the "dark-anticipating" phase of a 24-h cycle, flowering is inhibited, even if the plant has already received a longer dark period than normally required to induce flowering. This shows that phytochrome does not measure the period of darkness simply by converting from P_{fr} to P_r, which occurs in darkness. Instead the state of the phytochrome must coincide with the state of the plant rhythm. In short-day plants, for flowering to occur, the light phase of the rhythm and P_{fr} and the dark phase of the rhythm and P_r must coincide. (Normal white light converts phytochrome predominantly to the P_{fr} form.) In order for long-day plants to flower, the opposite must occur: the rhythm must be disrupted by the conversion of phytochrome to the P_{fr} form during the dark phase of the rhythm. *See* FLOWER; PHOTOPERIODISM.

Flowering hormone. The light signal for flowering is received by the leaves, but it is the stem apex that responds. Exposing even a single leaf to the correct photoperiod can induce flowering. Clearly, then, a signal must travel from the leaf to the apex. Grafting a plant that has been photoinduced to flower to one not so induced can cause the noninduced plant to flower. It has been proposed that a flower-inducing hormone travels from the leaf to the stem apex and there induces changes in the development of the cells such that the floral morphology results. This hypothetical hormone has been named florigen. In some long-day plants gibberellin will induce flowering, but gibberellin is clearly not the universal florigen. It may be that florigen is a mixture of substances.

Juvenility. Some perennial plants will not produce reproductive structures until they are several years old. This is termed juvenility and is a problem to the

growers of fruit-tree crops and to tree breeders particularly in the lumber industry. The causes appear to be hormonal. Many conifers have now been brought to a reproductive state while quite young by the application of nonpolar gibberellins.

Dormancy. At certain stages of the life cycle, most perennial plants cease growth and become dormant. Plants may cease growth at any time if the environmental conditions are unfavorable. When dormant, however, a plant will not grow even if the conditions are favorable. Dormancy is normally an adaptation that allows the plant to survive an adverse environmental situation, such as the cold of winter. In temperate zones many trees stop growing after the initial spurt of spring growth. This period is called summer dormancy and is distinguished from winter or true dormancy in that growth can be started again relatively easily by conditions such as defoliation by insects or overfertilization. True dormancy is induced by declining photoperiods and temperatures of autumn. A truly dormant tree can withstand temperatures of tens of degrees below 32°F but will not grow if simply placed in favorable growing conditions. Dormancy is broken by a period of low temperatures (6–12 weeks) just above freezing (about 41°F or 5°C. This has normally occurred by mid to late winter, and on the return of higher temperatures and longer photoperiods in spring, growth will resume.

Seeds may also show dormancy when shed from the plant. This dormancy may be due to immature embryos, in which case it is broken simply by time or by chemical or physical restrictions to germination. Seeds may have impermeable coats that prevent the entry of water and oxygen. Such coats are broken down by abrasion or microbial action over a period of time in the soil. Dormancy may also be imposed by growth inhibitors either in the seed coat or in the embryo itself. This chemical inhibition of growth can be removed or overcome by a period of low temperatures (similar to that required to break bud dormancy), light, or the leaching of the inhibitor by washing. Each of these represents adaptations to different natural conditions. Dormancy thus prevents germination prior to winter, under heavy shade, or during dry conditions in a desert.

The exact nature of the growth inhibitors controlling dormancy is uncertain. Abscisic acid is certainly involved in some species, but the presence of high levels of abscisic acid in vigorously growing shoots shows that it alone cannot be the dormancy inducer. Rather dormancy is brought about by a balance between abscisic acid (and perhaps other growth inhibitors) and growth-promoting hormones, particularly gibberellins and cytokinins. Following a dormancy-breaking cold or light treatment of seeds, gibberellin content often increases, and gibberellins or cytokinins or both increase in woody plants after the breaking of dormancy by a cold period. *See* DORMANCY.

Seed germination. A seed is an embryo plant usually with a food store in the cotyledons or the endosperm tissue. It is formed following sexual reproduction, and serves to disperse the species and, in some cases via a resistant stage, to carry the species through a period of cold or drought. Seeds may or may not be dormant as described above. A non- or postdormant seed will start to grow as soon as it is provided with suitable conditions of water, oxygen, and warmth. There is a rapid initial uptake of water primarily due to the imbibition of colloids. There is then a transitory stage during which no growth takes place. During this time cellular reorganization following the previous dry stage takes place, and the synthesis of proteins begins. The synthesis of proteins often precedes the synthesis of mRNA which codes for the synthesis of proteins. This is because seeds contain genetic instructions stored in previously synthesized mRNA. This information is translated into the enzymes needed during early germination as soon as germination begins. In cotton, abscisic acid has been found to prevent germination by blocking the translation of the information stored in mRNA. Following the initial protein synthesis, RNA synthesis also starts. After the lag period, growth recommences, primarily through the osmotic uptake of water. The embryonic axes (stem and root) may also increase in dry weight at the expense of the food reserves, but an absolute increase in dry weight cannot occur until the plant grows into the light and starts making is own carbohydrates through photosynthesis. Once the plant reaches the light, the stem straightens up and slows down its elongation rate (which is higher in darkness), the leaves expand, and chlorophyll is synthesized. All these changes are mediated through the pigment phytochrome.

Gibberellin has been found to play a specific role in cereal grains in addition to that of breaking dormancy. On imbibition gibberellin is produced by the embryo and diffuses to the aleurone layer—a protein-rich layer just below the surface of the grain. Here it induces the formation of various hydrolytic enzymes, notably amylase, which diffuses into the starchy endosperm and hydrolyzes the starch to soluble sugars for use by the embryo. The gibberellin has been shown to promote the transcription of mRNA specific for amylase.

Senescence, fruit ripening, and flower fading. Senescence refers to those processes which follow the cessation of growth and are a prelude to the death of the whole plant or plant part (such as a leaf, flower, or fruit). As with any other developmental stage in the plant, senescence appears to be under precise hormonal control. Whole plant senescence in annual plants follows flowering and fruiting and is prevented or delayed if the flowers are removed. Because it also occurs in male pollen-producing plants, and in some cases can be prevented by environmental manipulation or hormonal applications in plants with rapidly growing seeds, senescence cannot be caused by the nutrient drain put on the plant by developing seeds. The developing seeds cause senescence either by diverting growth-promoting hormones from the vegetative growing points or by exporting senescence-inducing substances to the rest of the plant. Which mechanism is operating is

not yet known. Senescence is prevented by growth hormones which vary with species (for example, gibberellins in peas or auxin and cytokinins in soybean).

The senescence of leaves is associated with a breakdown of chlorophyll and the associated enzymes of photosynthesis. At the same time there is a pronounced increase in protease, the enzyme responsible for protein breakdown. The yellowing of older leaves can often be prevented by cytokinin application.

Flower fading and fruit ripening represent two specialized cases of senescence. Both are promoted by ethylene applications. The fading of orchids or morning glory flowers, or the ripening of numerous fruits, is associated with the pronounced production of ethylene by the flowers or fruits. In certain fruits (such as the apple), this ethylene production combined with a drop in the auxin level triggers a sharp rise in the respiration rate, termed the respiratory climacteric, which is associated with ripening.

Leaf abscission. As a perennial plant grows, new leaves are continuously or seasonally produced. At the same time the older leaves are shed because newer leaves are metabolically more efficient in the production of photosynthates. A total shedding of tender leaves may enable the plant to withstand a cold period or drought. In temperate deciduous trees, leaf abscission is brought about by declining photoperiods and temperatures. Leaf abscission occurs at a zone of specialized cells at the base of the leaf stalk and is also under precise control. Before abscission can take place, an aging of the abscission-zone cells must occur. This is prevented in vigorous leaves by a continual supply of auxin by the leaf blade. When this auxin supply decreases, aging is hastened by ethylene produced by the abscission-zone cells. In the latter stages of abscission, this ethylene promotes the synthesis and release of cellulase by the abscission-zone cells. The cellulase degrades the cellulose of the cell wall of the abscission-zone cells. Thus the connections between the cells are weakened and break, causing the leaf to shed. *See* ABSCISSION; PLANT MORPHOGENESIS. Peter J. Davies

Bibliography. D. E. Fosket, *Plant Growth and Development: A Molecular Approach*, 1994; A. W. Galston, P. J. Davies, and R. L. Satter, *The Life of the Green Plant*, 3d ed., 1980; J. R. Porter and D. W. Lawler (eds.), *Plant Growth: Interactions with Nutrition and Environment*, 1991; T. A. Steeves and I. M. Sussex, *Patterns in Plant Development*, 2d ed., 1989; P. F. Wareing (ed.), *Plant Growth Substances*, 1983.

Plant hormones

A group of naturally occurring, organic substances that influence plant physiological processes at low concentrations. The processes consist mainly of growth, differentiation, and development, although other processes, such as stomatal movement, may also be affected. Typical effective concentrations [in the range of 1 μM externally, or about 1–100 nM internally (about one part per billion), depending on the compound] are far below those where either nutrients or vitamins would be effective.

Plant hormones have also been referred to as phytohormones or plant growth substances. The term "plant growth regulator" is mainly used by the agrichemical industry to denote synthetic growth regulators as distinct from endogenous growth regulators. Plant hormones include the auxins (represented principally by the compound indoleacetic acid), gibberellins, cytokinins, abscisic acid, ethylene, and brassinosteroids. Salicylic acid and jasmonic acid and some peptides are principally concerned with the defensive response to invasion by fungal diseases or insects. Polyamines can also be considered as plant hormones, although they operate at a higher concentration than those previously listed. The flowering hormone, florigen, is a phloem-transported gene product, probably a messenger ribonucleic acid (mRNA).

The original use of the term "plant hormone" in plant physiology was derived from the mammalian concept of a hormone, which is a signaling molecule transferred from one part of an organism to another. Auxin similarly produces a growth response at a distance from its site of synthesis and thus fits the definition of a *transported* chemical messenger. However, plant hormones do not fulfill the requirements of a hormone in the mammalian sense. The synthesis of plant hormones may be localized (as occurs for animal hormones), but may also occur in a wide range of tissues, or cells within tissues. While plant hormones may be transported and have their action at a distance, this is not always the case. They may also act in the tissue in which they are synthesized or even within the same cell. At one extreme we find the transport of cytokinins from roots to leaves, where they prevent senescence and maintain metabolic activity, while at the other extreme the production of the gas ethylene may bring about changes within the same tissue, or within the same cell, where it is synthesized. Thus, transport is not an essential property of a plant hormone. *See* HORMONE; PLANT GROWTH; PLANT MORPHOGENESIS; PLANT PHYSIOLOGY.

Early history. The plant hormone concept probably derives from observations of morphogenic and developmental correlations by the German botanist Julius von Sachs between 1880 and 1893, who postulated the existence of root-forming, flower-forming, and other substances that move in different directions through the plant. At about the same time, Charles Darwin was making his original observations on the phototropism of grass coleoptiles (the protective sheaths enclosing the shoot tip and embryonic leaves of grasses) that led him to propose the existence of a signal that was transported from the tip of the coleoptile to the bending regions lower down. After further characterizations, by several workers, of the way in which the signal was moved, Fritz Went in the Netherlands was able in 1927 to isolate the chemical by diffusion from coleoptile tips into agar blocks which, when replaced on the tips of

decapitated coleoptiles, resulted in the stimulation of the growth of the decapitated coleoptiles and their bending when placed asymmetrically on these tips. This demonstrated the existence of a growth-promoting chemical that was synthesized in the coleoptile tips, moved basipetally, and, when distributed asymmetrically, resulted in a bending of the coleoptile away from the side with the higher concentration. This substance was named auxin and was later identified from extracts of urine and fungal cultures as indoleacetic acid (IAA).

Other lines of investigation led to the discovery of the other hormones. Research in plant pathogenesis led to the elucidation of gibberellins in the 1950s, which were identified from fungal cultures, and efforts to culture plant tissues led to cytokinins, which were originally found in degraded DNA at about the same time. The effects of illuminating gas and smoke on plants had been known for many years, but it was not until about the 1950s that ethylene was realized to be a natural hormone. Investigations on the control of abscission (the natural shedding of a plant part, such as a leaf, flower, seed, or fruit) and dormancy in the 1960s led to the isolation of abscisic acid from plant tissues. Brassinosteroids were found in pollen in the 1970s, although it was not until about 20 years later that this class was recognized to be of universal occurrence.

Hormone levels, biosynthesis, and action. While early purifications from plant material utilized tens or even hundreds of kilograms of tissues, modern analyses can be performed on a few milligrams of tissue using high-performance liquid chromatography (HPLC) following solvent extraction and initial purification on solid-phase extraction (SPE) cartridges. Characterization and quantitation are done typically by gas chromatography–mass spectrometry (GC-MS) or by coupled HPLC–MS. Fully automated systems of extraction and analysis are being developed. Immunoassay is also used for hormone quantitation, although it is considered much less precise.

The analysis of the expression of genes for hormone biosynthesis is carried out by quantitative PCR (polymerase chain reaction) or on gene microarray chips, and localization of hormone biosynthesis by the expression in transgenic plants of marker genes driven by promoters of genes for one or more steps in the biosynthetic process. The location of hormone action in tissues and cells has also been investigated by examining the location of marker gene expression driven by promoters of genes known to be induced by the presence of hormone. *See* CLINICAL MICROBIOLOGY; EXTRACTION; GAS CHROMATOGRAPHY; IMMUNOASSAY; MASS SPECTROMETRY.

Nature, occurrence, and effects. The effects produced by each hormone were initially elucidated largely from exogenous applications. However, in more and more cases, we have evidence that the endogenous hormone also fulfills the originally designated roles, and new functions are being discovered. Such more recent evidence derives from correlations between hormone levels and growth of defined genotypes or mutants, particularly of the model plant *Arabidopsis*, or from transgenic plants in which hormone biosynthesis, reception, or signal transduction is enhanced by constitutive expression of genes involved in the pathway, or inhibited by the incorporation of antisense (a strand of DNA having a sequence identical to mRNA) or RNAi (RNA interference, whereby the production of double-stranded RNA renders homologous mRNA incapable of being translated) constructs targeted at the genes under investigation. It should, however, be emphasized that hormones do not act alone but in conjunction with or in opposition to each other, such that the final condition of growth or development represents the net effect of a hormonal balance.

How plant hormones work. The individual details of hormone function vary from hormone to hormone. However, a general pattern is clear. First the hormone binds to a receptor, either on the cell membrane or soluble within the cell. This causes the initiation of a signal transduction cascade that culminates in the activation of the transcription of a suite of genes in response to the biosynthesis or activation of one or more DNA-binding transcription factors. In all cases to date, at least one member of the signal transduction cascade is a negative regulator; that is, it is present in the absence of the hormone and blocks the hormone-type response(s). Thus, the null response in mutants in which this component is defective is the hormone-type response, but in the absence of the hormone. When the hormone binds to the receptor, the transmission of the signal, often via phosphorylation of the signal transduction components, leads to the destruction of the negative regulator so that the hormone-type response can be expressed. The negative regulator is tagged for destruction by the protein ubiquitin, and it is then broken down in organelles called proteasomes. Details will be found in the description of each hormone. *See* PROTEASOME; PROTEIN.

Auxin. Indole-3-acetic acid (IAA) [structure **1**] is the

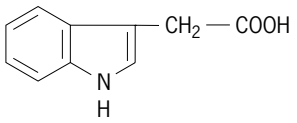

(1)

main auxin in most plants. Compounds that serve as IAA precursors may also have auxin activity (for example, indoleacetaldehyde). Some plants contain other compounds that display weak auxin activity (for example, indolebutyric acid). IAA may also be present as various conjugates such as indoleacetyl aspartate. 4-Chloro-IAA has also been reported in several species, although it is not clear to what extent the endogenous auxin activity in plants can be accounted for by 4-Cl-IAA. Several synthetic auxins are also used in commercial applications.

IAA is synthesized from tryptophan or indole primarily in leaf primordia and young leaves, and in developing seeds. IAA transport is cell to cell, mainly in the vascular cambium and the procambial strands but probably also in epidermal cells. Transport to the

root probably also involves the phloem. The transport mechanism involves carrier proteins located in the basal membranes of the transporting cells. *See* LATERAL MERISTEM; PHLOEM; PLANT TISSUE SYSTEMS; PRIMARY VASCULAR SYSTEM (PLANT).

Effects of auxin include the following: the stimulation of cell enlargement and stem growth; the stimulation of cell division in the cambium and, in combination with cytokinin, in tissue culture; the promotion of vascular tissue differentiation; the initiation of roots on stem cuttings, and also the development of branch roots and the differentiation of roots in tissue culture; the mediation of tropistic (bending) response of shoots and roots to gravity and light; the maintenance of apical dominance where the auxin supply from the apical bud represses the growth of lateral buds, though probably not directly; the direction of assimilate partitioning; the stimulation of the growth of flower parts; the induction of fruit setting and growth of some fruit; the inhibition or promotion of leaf and fruit abscission, depending on the timing and position of the source; delaying fruit ripening; and delaying leaf senescence. In several systems (for example, root growth), auxin, particularly at high concentrations, is inhibitory because of auxin-induced ethylene production. *See* APICAL DOMINANCE; FRUIT.

Commercial applications of auxin include the rooting of cuttings (indolebutyric acid), the induction of fruit set (naphthaleneacetic acid, or NAA), the promotion of fruit abscission after bloom, and the inhibition of fruit abscission before harvest (NAA). The synthetic auxin 2,4-dichlorophenoxyacetic acid is used as a weed killer. *See* AUXIN.

Gibberellins. The gibberellins (GAs) are a family of over 125 compounds based on the *ent*-gibberellane structure and numbered in the order of their discovery. While the most widely available compound is GA_3, or gibberellic acid, which is a fungal product, the most important gibberellin in plants is GA_1, which is the gibberellin primarily responsible for stem elongation (**2**). Many of the other gibberellins

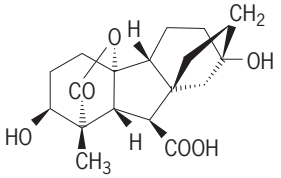

(2)

are precursors of the growth-active GA_1. Gibberellins are synthesized from glyceraldehyde-3-phosphate, via isopentenyl diphosphate, in young tissues of the shoot and developing seed. Their biosynthesis starts in the chloroplast and subsequently involves membrane and cytoplasmic steps. Some gibberellins are transported in the phloem and xylem, but the transport of the main bioactive, polar GA_1 is restricted. *See* XYLEM.

The effects of gibberellins include the following: the hyperelongation of stems by stimulating both cell division and cell elongation, producing tall, as opposed to dwarf, plants; bolting (stem elongation) in long-day plants, which respond to long days; the induction of seed germination, especially in seeds that normally require cold (stratification) or light to induce germination; the production of enzymes during germination, notably starch-degrading α-amylase in germinating cereal grains; the promotion of fruit setting and growth in some fruit (for example, grapes); the promotion of early embryo development; and the induction of maleness in dioecious (having the male and female reproductive organs on different individuals) flowers.

Commercially gibberellins are used to promote fruit and bunch size in seedless grapes, to increase the yield of malt in brewing, and to increase the production of sugar by sugarcane. Gibberellin biosynthesis inhibitors are important dwarfing agents in ornamentals and in cereals. Dwarf wheats have a defective gibberellin receptor. *See* GIBBERELLIN.

Cytokinins. Cytokinins (CKs) are adenine derivatives characterized by an ability to induce cell division in tissue culture (in the presence of auxin). The most common cytokinin base in plants is zeatin (**3**).

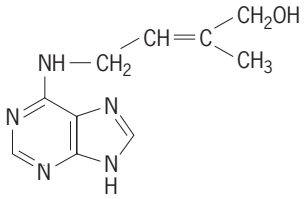

(3)

Cytokinins also occur as ribosides and ribotides. Cytokinin biosynthesis occurs in root tips and developing seeds through the biochemical modification of adenine. Transport from roots to shoots is via the xylem.

The effects of cytokinins include the following: the promotion of cell division in areas of actively dividing cells (for example, fruits and shoot tips), in tissue culture in the presence of auxin, and in crown gall tumors on plants; the induction of morphogenesis in tissue culture where cytokinins promote shoot initiation, whereas cytokinins in moss induce bud formation; the promotion of the growth of lateral buds; the promotion of leaf expansion; the induction of development of chloroplasts; the enhancement of stomatal opening in some species; and the delay of leaf senescence.

Cytokinins are not used directly in plant production, but the transgenic expression of cytokinin biosynthesis genes in aging tissues of leafy vegetables (such as lettuce) causes the prevention of senescence and the maintenance of greenness. This is not yet used commercially but could be in the future. *See* CYTOKININS.

Ethylene. The gas ethylene (C_2H_4) is synthesized from methionine via 1-aminocyclopropane-1-carboxylic acid (ACC) in many tissues in response to stress, and is the fruit-ripening hormone. It does not seem to be essential for normal mature vegetative growth, as ethylene-deficient transgenic

plants grow normally. However, as seedlings they cannot penetrate the soil because they lack the stem thickening and apical hook responses to ethylene, and they are susceptible to diseases because they lack the ethylene-induced disease resistance responses.

The effects of ethylene include the so-called triple response, when, prior to soil emergence, dark-grown seedlings maintain an apical hook and display a decrease in stem elongation, a thickening of the stem, and a transition to lateral growth, as might occur during the encounter of a stone in the soil. Other effects are an increase in the stem elongation rate in light-grown plants in response to shading; the stimulation of numerous defense responses in response to injury or disease; adventitious root formation; the induction of femaleness in dioecious flowers; flower and leaf senescence; fruit ripening; and leaf and fruit abscission.

Ethylene is used in the promotion of pineapple flowering so that flowering occurs more rapidly and mature fruits appear uniformly, thus permitting a single-harvest mechanical operation. Ethylene is used to accelerate fruit ripening after shipping. An ethylene-releasing compound, ethephon (4-chloroethyl phosphonic acid), is painted onto cuts on rubber trees to increase latex yield. *See* ETHYLENE.

Abscisic acid. Abscisic acid (**4**) is a single compound.

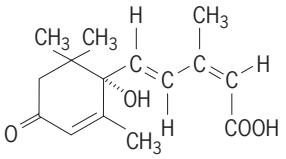

(4)

The first name given was abscisin II because it was thought to control the abscission of cotton bolls. At almost the same time, another group named it dormin for a purported role in bud dormancy. By a compromise the name abscisic acid (ABA) was coined. It now appears to have little role in either abscission (which is regulated by ethylene) or bud dormancy. ABA applications can inhibit growth, but ABA appears to act as much as a promoter as an inhibitor, such as in the promotion of storage protein synthesis in seeds. One of the main functions is the regulation of stomatal closure.

ABA is synthesized from glyceraldehyde-3-phosphate via isopentenyl diphosphate and carotenoids in roots and mature leaves, particularly in response to water stress. Seeds are also rich in ABA, which may be imported from the leaves or synthesized in situ. ABA is exported from roots in the xylem and from leaves in the phloem. There is some evidence that ABA may circulate to the roots in the phloem and then return to the shoots in the xylem.

Effects of abscisic acid include the following: stomatal closure in response to water stress; the inhibition of shoot growth, and to a lesser extent a promotion of root growth, which may also represent a response to water stress; the induction of storage protein synthesis in seeds; a counteraction of the effect of gibberellin on α-amylase synthesis in germinating cereal grains; and the induction and maintenance of some aspects of dormancy in seeds. ABA does not, however, appear to be the controlling factor in true dormancy or rest, which is dormancy that needs to be broken by low temperature or light. An increase in ABA that occurs in response to wounding induces gene transcription, notably for proteinase inhibitors, so it may be involved in defense against insect attack. *See* ABSCISIC ACID; ABSCISSION; DORMANCY.

Brassinosteroids. Brassinosteroids (BRs) are a group of over 60 steroidal compounds, typified by the compound brassinolide (**5**), which was first iso-

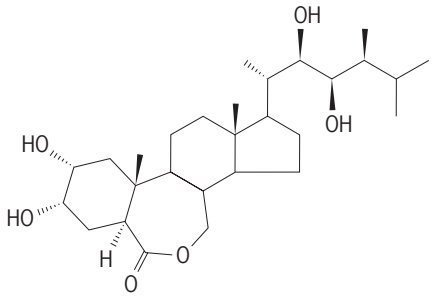

(5)

lated from *Brassica* pollen. They produce effects on growth and development at extremely low concentrations, and play a role in the endogenous regulation of growth and development. Their effects include cell division, possibly by increasing transcription of the gene encoding cyclin-D3, which regulates a step in the cell cycle; and cell elongation, where BRs promote the transcription of genes encoding xyloglucanases and expansins and promote wall loosening. This leads to stem elongation, vascular differentiation, and fertility.

Polyamines. The polyamines are a group of aliphatic amines. The main compounds are putrescine, spermidine (**6**), and spermine. They are derived from

$$H_2N — (CH_2)_3 — NH — (CH_2)_4 — NH_2$$

(6)

the decarboxylation of the amino acids arginine or ornithine. The conversion of the diamine putrescine to the triamine spermidine and the quaternary amine spermine involves the decarboxylation of *S*-adenosylmethionine, which also occurs in the pathway for the biosynthesis of ethylene. As a result, there are some complex interactions between the levels and effects of ethylene and the polyamines.

Polyamines have a wide range of effects on plants and appear to be essential for plant growth, particularly cell division and normal morphologies. In plants where the content of polyamines is genetically altered, development is affected. It is not possible to make an easy, distinct list of their effects. Polyamines are present in all cells.

Jasmonates. Jasmonates are represented by jasmonic acid (**7**) and its methyl ester. They are named

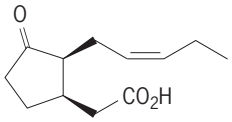

(7)

after the jasmine plant, in which the methyl ester is an important scent component. Jasmonic acid is synthesized from linolenic acid. Jasmonates play an important role in plant defense, where they induce the synthesis of proteinase inhibitors which deter insect feeding and, in this regard, act as intermediates in the response pathway induced by the peptide systemin. Jasmonates inhibit many plant processes such as growth and seed germination, and promote senescence, abscission, tuber formation, fruit ripening, pigment formation, and tendril coiling. *See* PLANT COMMUNICATION.

Salicylic acid. Salicylates have been known for a long time to be present in willow bark, but have only recently been recognized as potential regulatory compounds. Salicylic acid (**8**) is biosynthesized from

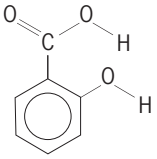

(8)

the amino acid phenylalanine, and plays a principal role in the resistance to pathogens by inducing the production of pathogenesis-related proteins. It is involved in the systemic acquired resistance response (SAR), in which a pathogenic attack on older leaves causes the development of resistance in younger leaves; however, whether salicylic acid is the transmitted signal is debatable. Salicylic acid is the calorigenic substance that causes thermogenesis in *Arum* flowers. It has also been reported to enhance flower longevity, inhibit ethylene biosynthesis and seed germination, block the wound response, and reverse the effects of ABA. *See* SALICYLATE.

Signal peptides. The discovery that small peptides could have regulatory properties in plants started with the discovery of systemin, an 18-amino-acid peptide that travels in the phloem from leaves under herbivore insect attack to increase the content of jasmonic acid and proteinase inhibitors in distant leaves, so protecting them from attack. Since then, over a dozen peptide hormones have been isolated from plants or have been identified by genetic approaches. These peptides regulate various processes involved in defense, cell division, growth, development, and reproduction. Among the effects caused by specific peptides are the activation of defense responses; the promotion of cell proliferation of suspension-cultured plant cells; the determination of cell fate during development of the shoot apical meristem; the modulation of root growth and leaf patterning in the presence of auxin and cytokinins; signals for self-incompatibility; and nodule formation in response to bacterial signals involved in nodulation in legumes.

Florigen. The flowering hormone, florigen, was proposed in 1937 but remained elusive for almost 70 years. In 2005 it was shown to be the phloem-transported product of the gene LEAFY, probably the LFY mRNA. The production of LFY mRNA occurs in leaves in response to inductive photoperiods. It is then transported in the phloem to the shoot apex, where its translation product brings about the transition to floral development. *See* NUCLEOPROTEIN; PHOTOPERIODISM; PLANT REPRODUCTION; RIBONUCLEIC ACID (RNA). Peter J. Davies

Bibliography. R. N. Arteca, *Plant Growth Substances: Principles and Applications*, Chapman and Hall, New York, 1995; P. J. Davies (ed.), *Plant Hormones: Biosynthesis, Signal Transduction, Action!*, Kluwer Academic, Dordrecht and Boston, 2004.

Plant keys

Artificial analytical constructs for identifying plants. The identification, nomenclature, and classification of plants are the domain of plant taxonomy, and one basic responsibility of taxonomists is determining if the plant at hand is identical to a known plant. The dichotomous key provides a shortcut for identifying plants that eliminates searching through numerous descriptions to find one that fits the unknown plant.

A key consists of series of pairs (couplets) of contradictory statements (leads). Each statement of a couplet must "lead" to another couplet or to a plant name (thus the term "lead"). Each couplet provides an either-or proposition wherein the user must accept one lead as including the unknown plant in question and must simultaneously reject the opposing lead. The user then proceeds from acceptable lead to acceptable lead of successive couplets until a name for the unknown plant is obtained. For confirmation, the newly identified plant should then be compared with other known specimens or with detailed descriptions of that species.

Keys are included in monographic or revisionary treatments of groups of plants, most often for a genus or family. Books containing extended keys, coupled with detailed descriptions of each kind of plant (taxon), for a given geographic region are called manuals or floras, though the latter technically refers to a simple listing of names of plants for a given region.

Types of keys. Several types of artificial keys may be constructed. Generally, modern keys are dichotomous, always composed of pairs of opposing statements as described above. Keys that are polychotomous (that is, have three or more optional lead statements simultaneously) are confusing and time-consuming and should be avoided. Most commonly, keys are presented in two forms, the indented or yoked key and the parallel or bracketed key.

Examples of each are presented below for the leaves of the six representative trees and shrubs of North America shown in the **illustration**.

Indented or Yoked Key

1. Leaf venation palmate	
2. Leaves once compound	Plant *e*
2′. Leaves simple	
3. Margins lobed; leaves opposite	Plant *a*
3′. Margins entire; leaves alternate	Plant *c*
1′. Leaf venation pinnate	
4. Leaves once compound	Plant *d*
4′. Leaves simple	
5. Margins serrate; stipules leafy	Plant *f*
5′. Margins lobed; stipules absent	Plant *b*

Parallel or bracketed key

1.	Stem bearing two leaves per node (opposite)	2
1′.	Stem bearing one leaf per node (alternate)	3
2.(1.)	Leaves simple	Plant *a*
2′.	Leaves once palmately compound	Plant *e*
3.(1)′.	Leaves once pinnately compound	Plant *d*
3′.	Leaves simple 4	4
4.(3′.)	Venation palmate	Plant *c*
4′.	Venation pinnate	5
5.(4′.)	Leaf margins lobed Plant *b*	Plant *b*
5′.	Leaf margins serrate	Plant *f*

In the indented or yoked key, the two leads of a couplet are arranged in yokes, the first lead separated from the second by all the former's subordinate yokes. Each subordinate yoke, in turn, is indented under its preceding lead. This type of key permits easy visualization of plant groups, each of which is based on similar morphology. However, the indented key consumes large amounts of page space because of progressive indentation, particularly in extended keys where the two leads of the couplet may be separated by several pages. In the parallel key, the two leads of every couplet are adjacent. The parallel key occupies much less space (a distinct advantage for long keys), but it is difficult to visualize relationships among the different groupings of plants.

In dichotomous keys, the couplets usually are consecutively numbered, lettered, or otherwise symbolized for convenience. The two leads of a given couplet either bear the same number or symbol (3 versus 3) or differentiating numbers or symbols (3 versus 3′, la versus lb, or a versus aa). The number or symbol of the previous couplet's lead may be placed parenthetically after the number or symbol in the first lead in a couplet in parallel keys. This enables the user to backtrack through the key, which could prove advantageous after taking a wrong choice or otherwise getting lost in the key. Leads will provide either a name of the plant or a number for the next couplet in the key that should be considered.

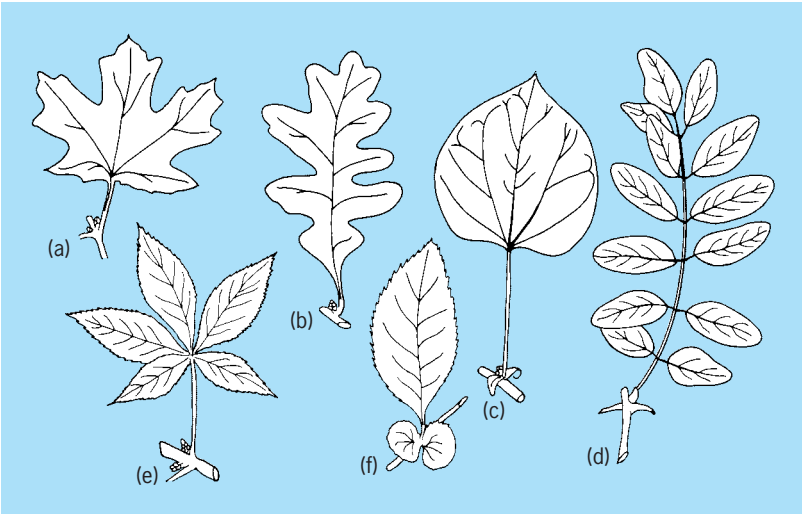

Leaves of six representative trees and shrubs of North America. (*a*) Leaves are simple and opposite; margins are lobed. (*b*) Leaves are simple, margins are lobed, and stipules are absent. (*c*) Leaves are simple and alternate; margins are entire. (*d*) Leaf venation is pinnate; leaves were once compound. (*e*) Leaf venation is palmate; leaves were once compound. (*f*) Leaves are simple, margins are serrate, and stipules are leafy.

A synoptic key, presented in either yoked or parallel key form, uses technical terms to emphasize evolutionary relationships (phylogenetic classification) among groups of plants, usually higher-ranking groups such as divisions, classes, or orders. Synoptic keys are not designed to identify individual plants but to illustrate general trends without concern for minor exceptions. Synoptic keys differ from keys designed specifically for identification of species, in that the latter use obvious, convenient characters regardless of evolutionary importance.

There are also several specialized keys. The pictorialized key includes illustrations to depict character states difficult to visualize as they arise in the key and to present for each plant an overall sketch. The card sort key consists of a set of cards, one for each kind of plant. Every card bears holes along its margins, each hole assigned to a specific character as written on the card face. If a plant possesses a certain character, the hole is cut away to the margin (notched), but if that plant does not possess a character, the hole is left intact. A specific notched hole may refer, for example, to "leaves simple"; if the same hole had been left intact, then it refers to "leaves not simple." A rod inserted through this hole position for the entire deck of cards will retain all cards with "leaves not simple" and cards that fall away from the deck will have notched holes, meaning "leaves simple." Those remaining on the rod are not considered further. The cards that fell away are then reassembled into a deck and the user goes to another character, and another, until the last card to fall away provides for the unknown plant its name, description, and illustration on the obverse face. Also available are computer programs for keys that will automatically generate dichotomous keys from data matrices prepared and coded for them.

Constructing keys. In making artificial keys, simultaneous visualization of differences and similarities

Comparison chart of character states for leaves of six trees and shrubs of North America shown in the illustration

Character	Character states (plants *a–f*)					
	a	*b*	*c*	*d*	*e*	*f*
Leaves per node	**two**	one	one	one	**two**	one
Venation	**palmate**	pinnate	**palmate**	pinnate	**palmate**	pinnate
Leaf type	simple	simple	simple	**compound**	**compound**	simple
Leaf (let) margin	**lobed**	**lobed**	entire	entire	*serrate*	*serrate*
Stipules	**absent**	**absent**	*leafy*	spines	**absent**	*leafy*

among plants is necessary. To accomplish this, a comparison chart (see **table**) may be used. For each plant type considered, character expressions (states) of all major contrasting characters need to be tabulated, such as "simple" or "compound" under the character "leaf type." Then it is possible to quickly ascertain and then indicate simultaneously on the table the differences and likenesses among all plant types.

In constructing the first couplet, one or more characters are selected that will divide all the kinds of plants (taxa) to be considered into two groups. The character expressions (states) of all members of the first group differ from the character-states of all members of the second group for each character used. Therefore, all members of the first group are alike in having the same character-state for each character used. At the same time, all members of the second group are alike in having the contrasting character-state for each character used. By using these differing character-states to separate the two groups, the first and second leads of the first couplet can be written. (The first lead refers to the first group, preferably the group with fewer members.) The process continues, subdividing the first group by the same procedures until all names of the first group are represented. The second lead of the first couplet is returned to, and is made to lead to the next consecutive number or symbol. No numbers or symbols used for the first group are repeated. The second group continues to be subdivided in like manner until all plant names of the second group are represented. Either the indented (yoked) or parallel key form can be used.

Rules. Certain general guidelines should be followed to construct good keys:

1. Place characters of a given lead in descending order of importance such that the best diagnostic character is first and the most variable last.

2. Make sure that both leads of a couplet are of parallel construction: the same characters in the same sequence and no characters omitted in either of the leads.

3. In opposing leads of each couplet, the character expressions (states) for all characters must be mutually exclusive. In difficult separations, a combination of characters may be necessary, or the same plant may be entered more than once in various portions of the key.

4. The first word of both leads of a couplet should be identical.

5. The first word of consecutive couplets should not be the same.

6. Avoid negative choices if possible. For example, use "leaves compound" versus "leaves simple" but do not use "leaves compound" versus "leaves not compound," or "leaves compound" versus "leaves not as above."

7. Place the name of the plant part before the description. For instance, use "fruits red" but do not use "red fruits."

8. Avoid overlapping variation limits. "Spine 1.5–5.5 cm long" should not be paired with "spine 4.5–8.6 cm long." Avoid vague descriptive parameters, such as "long" or "short," and "leaves narrowly lanceolate" or "leaves broader." Rather, use exact measurements. Avoid redundancy, such as having the same phrase in both leads of one couplet.

9. Use obvious characters where possible. Avoid characters such as chromosome numbers and physiological parameters that are not determinable from dried herbarium specimens. Also, avoid geographic distribution or habitat as a sole character of lead.

10. In dioecious species having separate staminate and pistillate plants, couplets must include characters for both flower types since the plant to be identified will have only one type of flower. Use flower and fruit characters for those plants that do not bear flowers and relatively mature fruits at the same time.

11. Use characters of reproductive plant parts in preference to characters of vegetative plant parts because the reproductive parts are usually less variable, because they have been exposed to influencing environmental factors for a shorter period of time.

12. For each dichotomy, it is most efficient to divide a group of plants into equal or subequal halves; do not splinter off one kind of plant at a time from the main group. An example of a key modified in a historic dichotomous key format is shown below.

Suggestions for the use of keys. Keys are valuable tools but require care in their use. Much material of the unknown plant should be examined to learn of its variability; to work most keys, both reproductive and vegetative materials are required. If possible, diseased, mutant, grazed, juvenile, or otherwise aberrant material should be avoided. Learn the terminology presented in the key. Read, in their entirety, both leads of each couplet before making a decision as to which lead to accept. Remember that the first character of a lead is the most diagnostic. If the user cannot accept either lead of a given couplet, perhaps because the specimen of the unknown plant is incomplete, then accept both leads and continue from

each until the unknown becomes identified. If neither lead is workable, then start over and seek out previous errors in judgment. If the unknown does not fit within the key, it may be new to that flora or new to science. *See* PLANT KINGDOM; PLANT TAXONOMY. Donald J. Pinkava

Bibliography. A. Cronquist, *Evolution and Classification of Flowering Plants*, 2d ed., 1988; M. J. Dallwitz, *User's Guide to the DELTA System: A General System for Coding Taxonomic Descriptions*, 2d ed., 1984; J. E. Forester, *Sort Guide to the Trees of the U.S. and Canada*, 1963; H. D. Harrington, *How to Identify Grasses and Grasslike Plants*, 1977; G. H. M. Lawrence, *Taxonomy of Vascular Plants*, 1951; G. W. Prescott, *How to Know the Aquatic Plants*, Picture Key Nature Series, 2d ed., 1980; A. E. Radford et al., *Vascular Plant Systematics*, 1974.

Plant kingdom

The worldwide array of plant life, including plants that have roots in the soil, plants that live on or within other plants and animals, plants that float on or swim in water, and plants that are carried in the air. Fungi used to be included in the plant kindom because they looked more like plants than animals and did not move about. It is now known that fungi are probably closer to animals in terms of their evolutionary relationships (although not similar enough to be considered animals). Also once included in plants were the "blue-green algae," which are now clearly seen to be bacteria, not plants, although they are photosynthetic (and presumably the group of organisms from which the chloroplasts present in true plants were derived). The advent of modern methods of phylogenetic DNA analysis has allowed such distinctions, but even so, what remains of the plantlike organisms is still remarkably divergent and difficult to classify.

Plants range in size from unicellular algae to giant redwoods. Some plants complete their life cycles in a matter of hours, whereas the bristlecone pines are known to be over 4000 years old. Plants collectively are among the most poorly understood of all forms of life, with even their most basic functions still inadequately known, including how they sense gravity and protect themselves from infection by bacteria, viruses, and fungi (they do not have immune systems like those of vertebrates but nonetheless ward off such attacks). Furthermore, new species are being recorded every year. The application of new approaches to the understanding of this diversity of form and function is yielding major insights, but this increased level of knowledge does not facilitate classification; the stream of new information means that anything suggested today may need to be modified tomorrow.

Classification of plants has relied historically on the use of hierarchy to express how the various named taxa are related. Although the relationships between the taxa are less than certain, there is still information in the levels of the traditional hierarchical scheme. The idea has been that species of plants are grouped into ever larger, inclusive categories to make a formal hierarchy. The white oak (*Quercus alba*), for example, may be classified as follows:

Kingdom Plantae
- Subkingdom Embryobionta
  - Division Magnoliophyta
    - Class Magnoliopsida
      - Subclass Rosidae
        - Order Fagales
          - Family Fagaceae
            - Genus *Quercus*
              - Species *alba*

Within the land plants, a great deal of progress has been made in sorting out phylogenetic (evolutionary) relationships of extant taxa based on DNA studies, and the system of classification listed below includes these changes. The angiosperms or flowering plants (Division Magnoliophyta) have recently been reclassified based on phylogenetic studies of DNA sequences. Within the angiosperms, several informal names are indicated in parentheses; these names may at some future point be formalized, but for the present they are indicated in lowercase letters because they have not been formally recognized under the Code of Botanical Nomenclature.

It is known that the bryophytes (Division Bryophyta) are not closely related to each other, but which of the three major groups is closest to the other land plants is not yet clear. Among the extant vascular plants, Lycophyta are the sister group to all the rest, with all of the fernlike groups forming a single monophyletic (natural) group, which is reflected here in the classification by putting them all under Polypodiophyta. This group is the sister to the extant seed plants, within which all gymnosperms form a group that is sister to the angiosperms. Therefore, if Division is taken as the highest category within Embryobionta (the embryo-forming plants), then the following scheme would reflect the present state of knowledge of relationships (an asterisk indicates that a group is known only from fossils). See separate articles on names marked by daggers.

Subkingdom Thallobionta (thallophytes)†
- Division Rhodophycota (red algae)
  - Class Rhodophyceae†
- Division Chromophycota†
  - Class: Chrysophyceae (golden or golden-brown algae)†
    - Prymnesiophyceae†
    - Xanthophyceae (yellow-green algae)†
    - Eustigmatophyceae†
    - Bacillariophyceae (diatoms)†
    - Dinophyceae (dinoflagellates)†
    - Phaeophyceae (brown algae)†
    - Raphidophyceae (chloromonads)†
    - Cryptophyceae (cryptomonads)†
- Division Euglenophycota (euglenoids)
  - Class Euglenophyceae†
- Division Chlorophycota (green algae)†

Class: Chlorophyceae†
Charophyceae†
Prasinophyceae
Subkingdom Embryobionta (embryophytes)†
Division Rhyniophyta*†
Class Rhyniopsida†
Division Bryophyta†
Class Hepaticopsida (liverworts)†
Subclass Jungermanniidae†
Order: Takakiales†
Calobryales†
Jungermanniales†
Metzgeriales†
Subclass Marchantiidae†
Order: Sphaerocarpales†
Monocleales†
Marchantiales†
Class: Anthocerotopsida (hornworts)†
Sphagnopsida (peatmosses)†
Andreaeopsida (granite mosses)†
Bryopsida (mosses)†
Subclass: Archidiidae†
Bryidae†
Order: Fissidentales†
Bryoxiphiales†
Schistostegales†
Dicranales†
Pottiales†
Grimmiales†
Seligeriales†
Encalyptales†
Funariales†
Splachnales†
Bryales†
Mitteniales†
Orthotrichales†
Isobryales†
Hookeriales†
Hypnales†
Subclass: Buxbaumiidae†
Tetraphididae†
Dawsoniidae†
Polytrichidae†
Division Lycophyta†
Class Lycopsida†
Order: Lycopodiales†
Asteroxylales*†
Protolepidodendrales*†
Selaginellales*
Lepidedendrales*†
Isoetales†
Class Zosterophyllopsida*†
Division Polypodiophyta†
Class Polypodopsida†
Order: Equisetales†
Marattiales†
Sphenophyllales*
Pseudoborniales*
Psilotales
Ophioglossales†
Noeggerathiales*
Protopteridales*
Polypodiales†
Class Progymnospermopsida*
Division Pinopsida
Class Ginkgoopsida†
Order: Calamopityales*
Callistophytales*
Peltaspermales*
Ginkgoales†
Leptostrobales*
Caytoniales†
Arberiales*
Pentoxylales*
Class Cycadopsida†
Order: Lagenostomales*
Trigonocarpales*
Cycadales†
Bennettiales*
Class Pinopsida†
Order: Cordaitales*†
Pinales†
Podocarpales
Gnetales†
Division Magnoliophyta (angiosperms, flowering plants)†
Class Magnoliopsida†
Unplaced groups: Amborellaceae, Ceratophyllaceae, Chloranthaceae, Nymphaeaceae, etc.
eumagnoliids†
Order: Magnoliales†
Laurales†
Piperales†
Winterales
monocotyledons†
Order: Acorales†
Alismatales†
Asparagales†
Dioscoreales†
Liliales†
Pandanales†
commelinids
Arecales†
Commelinales†
Poales†
Zingiberales†
eudicotyledons†
(basal eudicots)
Order: Ranunculales†
Proteales†
Buxales
Trochodendrales†
(core eudicots)
Order: Berberidopsidales
Gunnerales
Dilleniales†
Santalales†
Caryophyllales†
Saxifragales†
Rosidae†
Order: Vitales
Myrtales†
Geraniales†
Crossosomatales
(eurosid I)
Order: Celastrales†
Cucurbitales
Fabales†

Order: Fagales†
Malpighiales†
Oxalidales†
Rosales†
Zygophyllales†
(eurosid II)
Order: Brassicales†
Malvales†
Sapindales†
Asteridae†
Order: Cornales†
Ericales†
(euasterid I)
Order: Garryales
Gentianales†
Lamiales†
Solanales†
(euasterid II)
Order: Apiales†
Aquifoliales
Asterales†
Dipsacales†

See DEOXYRIBONUCLEIC ACID (DNA); PLANT EVOLUTION; PLANT PHYLOGENY; PLANT TAXONOMY.

Mark W. Chase; Michael F. Fay

Plant metabolism

The complex of physical and chemical events of photosynthesis, respiration, and the synthesis and degradation of organic compounds. Photosynthesis produces the substrates for respiration and the starting organic compounds used as building blocks for subsequent biosyntheses of nucleic acids, amino acids and proteins, carbohydrates and organic acids, lipids, and natural products.

Energy Metabolism

The processes involved in the utilization and generation of energy in plants are photosynthesis and respiration.

Photosynthesis. Photosynthesis, which forms all organic compounds that are used for growth, development, and maintenance processes, has two aspects: primary processes that use light energy to form stable, high-energy compounds [reaction (1)],

$$\mathrm{NADP} + \mathrm{ADP} + \mathrm{P_i} \xrightarrow[\text{light}]{\text{chlorophyll}} \mathrm{ATP} + \mathrm{NADPH} \qquad (1)$$

and carbon metabolism that produces organic compounds from CO_2 by using the energy of reduced nicotinamide adenine dinucleotide phosphate (NADPH) and adenosine triphosphate (ATP) produced by the primary processes [reaction (2)].

$$\mathrm{CO_2} + \mathrm{H_2O} + \mathrm{NADPH} + \mathrm{ATP} \rightarrow \mathrm{CH_2O} + \mathrm{NADP} + \mathrm{ADP} + \mathrm{P_i} \qquad (2)$$

Primary processes. A photosynthetic action spectrum shows major efficiency peaks in the blue and red. Quantum considerations of the chlorophyll molecule and light absorption indicate that blue light acts similarly to red. The electrochemical span for photosynthesis goes from hydrogen to oxygen, a potential difference of about 1.2 eV. Red light of 700 nanometers has 167 kilojoules einstein^{-1}, sufficient to drive the reaction over 1.2 eV (115 kJ) even with an efficiency of 70%.

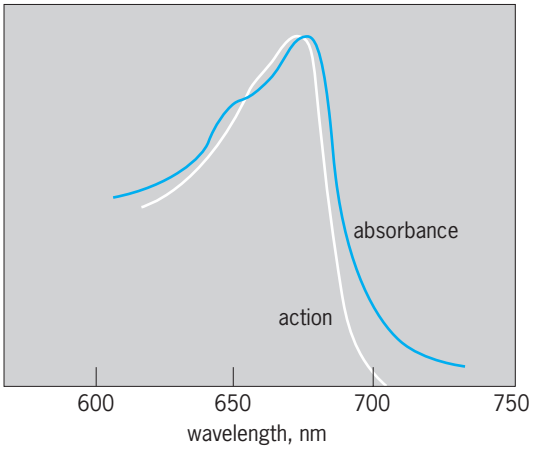

Fig. 1. Absorption spectrum for chlorophyll and action spectrum for photosynthesis, showing the red drop phenomenon in which absorption of red light occurs farther into the red than photosynthesis.

A photosynthetic unit is a minimum-sized particle that can absorb radiant energy and carry out the photochemical reactions. A unit is visualized to have about 200 chlorophyll molecules, carotenoids, cytochromes, and other components.

According to the modern concept of the photochemistry of photosynthesis, there are two different reaction centers, each composed of a specialized chlorophyll *a*. Associated with these reaction centers are accessory pigments, including chlorophyll *a* and chlorophyll *b*, that act as light-harvesting pigments or antenna pigments transferring energy to the centers. The reaction centers themselves make up only about 1% of the total pigments of the chlorophyll apparatus. The reaction center named photosystem I is a low-redox-potential system functioning in the reduction of reducing agents; the reaction center named photosystem II is a high-redox-potential system capable of oxidizing water.

The concept of two centers developed from the fact that the absorption spectrum for chlorophyll *a* extends further into the red than the action spectrum for photosynthesis (**Fig. 1**), and that low levels of red light (<680 nm) given simultaneously with low levels of far-red light (>680 nm) result in more photosynthesis than when the two lights are given separately. **Figure 2** illustrates the present two-light-system hypothesis for photosynthesis.

After light excitation of P700, a special form of chlorophyll *a*, electrons are passed to a primary acceptor. Rapid transfer of electrons to additional carriers ultimately results in reduction of NADP to NADPH [reaction (3)]. Thus light excitation of photosystem

$$\mathrm{PS\ I} \rightarrow \text{ferredoxin} \xrightarrow[\text{Fd-NADP-reductase}]{\text{NADP}} \mathrm{NADPH} \qquad (3)$$

I produces a strong reductant and a weak oxidant. There is some evidence that the primary acceptor may be a bound ferredoxin, or at least the ferredoxin is close to the primary electron acceptor. The primary acceptor appears to be an iron-sulfur protein. The donor of electrons to the oxidized P700, which come from photosystem II, is probably plastocyanin and cytochrome *f*.

Whereas the data concerning the redox agents of photosystem I are based primarily on electron spin resonance and absorption spectroscopy, data from photosystem II also come from studies of fluorescence. Thus, somewhat more is known about this system, which functions in the oxidation of water and consequent evolution of oxygen [reaction (4)]

$$2H_2O \rightarrow 4H^+ + 4e^- + O_2 \qquad (4)$$

and the transferring of electrons to photosystem I. The reaction center of photosystem II is named P680.

Coupling of the two centers through the weak reductant of photosystem II and the weak oxidant of photosystem I drives phosphorylation, producing two ATPs per electron pair. There are two sites for phosphorylation. Phosphorylation is driven by chemiosmotic coupling in which proton gradients are established across chloroplast membranes. The main features of the hypothesis are as follows: there is a vectorial transport of protons across the membrane; electron flow is coupled to proton transport; counterion flux across the membrane prevents positive-ion accumulation; and the proton gradient creates a proton-motive force sufficient to drive phosphorylation.

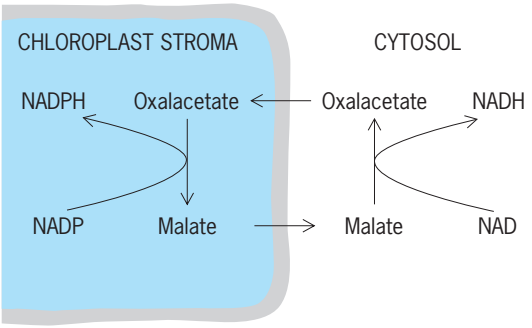

Fig. 3. Malate shuttle hypothesis; a possible mechanism transferring reducing power from NADPH in the chloroplast to NADH in the cytosol. Other possible schemes involve transport of triose phosphates.

Electron transport from water to NADP, forming ATP in the process and involving both photosystems, is called noncyclic electron transport. When photosystem II is uncoupled from photosystem I [as by 3(3,4-dichlorophenol)-1,1-dimethylurea, or DCMU], red light will still drive phosphorylation, but there is no NADP reduced. Electrons are cycled from photosystem I to the strong reductant and back through the electron transport system to the photosystem I center. This cyclic electron flow produces ATP but no NADPH.

Under special conditions in the absence of CO_2, chloroplasts consume oxygen and produce ATP but not NADPH. Electron transport is from water to oxygen, a process called pseudocyclic photophosphorylation [reaction (5)].

$$H_2O \rightarrow \text{PS II} \rightarrow \text{PQ} \rightarrow \text{cyt f/PC} \rightarrow \text{PS I} \rightarrow O_2 \qquad (5)$$

Cyclic photophosphorylation and perhaps pseudocyclic photophosphorylation may produce the extra ATP that is required for cellular activities other than the carbon metabolism of photosynthesis.

The chloroplast envelope is not permeable to ATP and NADPH, yet there is evidence that ATP and NADPH from photosynthesis are used in nonchloroplast metabolism. Glucose and sucrose do not pass readily across the envelope, as do malate and the triose phosphates. The latter are transported by a phosphate exchange reaction. It is visualized that shuttles carry ATP and reducing power across the envelope as shown in **Fig. 3**. *See* ADENOSINE TRIPHOSPHATE (ATP).

Carbon metabolism. The path of carbon in photosynthesis has been known since the 1950s (**Fig. 4**). Carboxylation catalyzed by ribulose bisphosphate carboxylase produces 2 moles of phosphoglycerate, the first product of photosynthesis. Reduction of the latter forms sugar, glyceraldehyde phosphate. Because

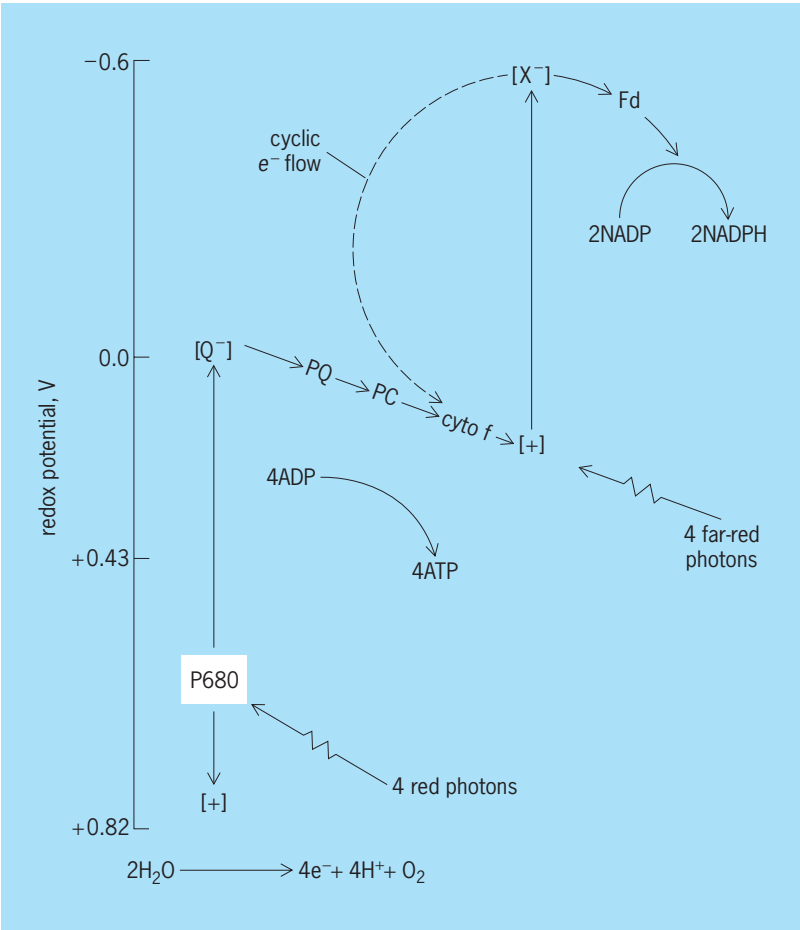

Fig. 2. Z scheme for electron transport in photosynthesis, showing noncyclic electron flow from water in photosystem II to NADP in photosystem I, and cyclic electron flow in which ATP is still formed, but not NADPH. PQ = plastoquinone, PC = plastocyanin, cyto f = cytochrome f, Fd = ferredoxin, Q = quinone.

the first product is a three-carbon compound, the process is referred to as C_3 photosynthesis.

In the 1960s, a modification of the pathway was discovered in sugarcane, wherein the first detectable product of carboxylation is a four-carbon organic acid, oxalacetate. Because of this, the process is referred to as C_4 photosynthesis. The oxalacetate is rapidly reduced to malate or aminated to form aspartate. The initial carboxylation of P-enolpyruvate (phosphoenolpyruvate, or PEP) carboxylase occurs in the outer mesophyll cells (the C_4 plants have Kranz anatomy). Malate or aspartate is transferred to the inner bundle sheath cells, where decarboxylation forms CO_2 and a three-carbon compound (such as pyruvate). The CO_2 is then assimilated through the ribulose bisphosphate carboxylase reaction. The three-carbon compound is recycled back to the mesophyll cells and converted back to the acceptor molecule, P-enolpyruvate, by a special enzyme called pyruvate inorganic phosphate dikinase. In some C_4 species, amino acids are transported rather than organic acids.

The chloroplasts of C_4 species have typical-looking chloroplasts in the mesophyll cells, but they lack ribulose bisphosphate carboxylase. The P-enolpyruvate carboxylase is in the cytosol. Both photosystems are present. Ribulose bisphosphate carboxylase is present in the large lamellar chloroplasts of the bundle sheath cells, but they have a reduced photosystem II.

In C_3 species there are 3 ATP and 2 NADPH required for CO_2 fixation (Fig. 4), whereas in C_4 species additional ATP is required, primarily to regenerate the P-enolpyruvate.

In most plants (C_3 species in particular) there is light-dependent oxygen consumption and CO_2 evolution called photorespiration that occurs in leaf microbodies. The enzyme ribulose bisphosphate carboxylase has oxygenase activity forming phosphoglycerate and glycolate phosphate (**Fig. 5**). The glycolate is oxidized in the microbodies by glycolate oxidase, consuming more oxygen and ultimately producing glycine and serine. Complete oxidation of the latter yields CO_2. In some C_3 plants, as much as 50% of the carbon fixed by photosynthesis can be lost through photorespiration with no conservation of energy, prompting researchers to search for plants with reduced photorespiration. Because C_4 photosynthesis keeps the CO_2 level high in the vicinity of ribulose bisphosphate carboxylase, there is less photorespiration in C_4 species.

Still another modification of photosynthesis occurs in succulent plants, called crassulacean acid metabolism (CAM) because it has been extensively studied in the Crassulaceae. In CAM plants, stomata are closed during the day and open at night. CO_2 entering the plant at night is fixed into malic acid through the P-enolpyruvate carboxylase/malate dehydrogenase couple. Malic acid accumulates in large water-storing vacuoles [100–200 meq kg^{-1} (fresh weight)]. During the subsequent day period, malic acid is decarboxylated, and the CO_2, which is trapped inside the leaf because of closed stomata,

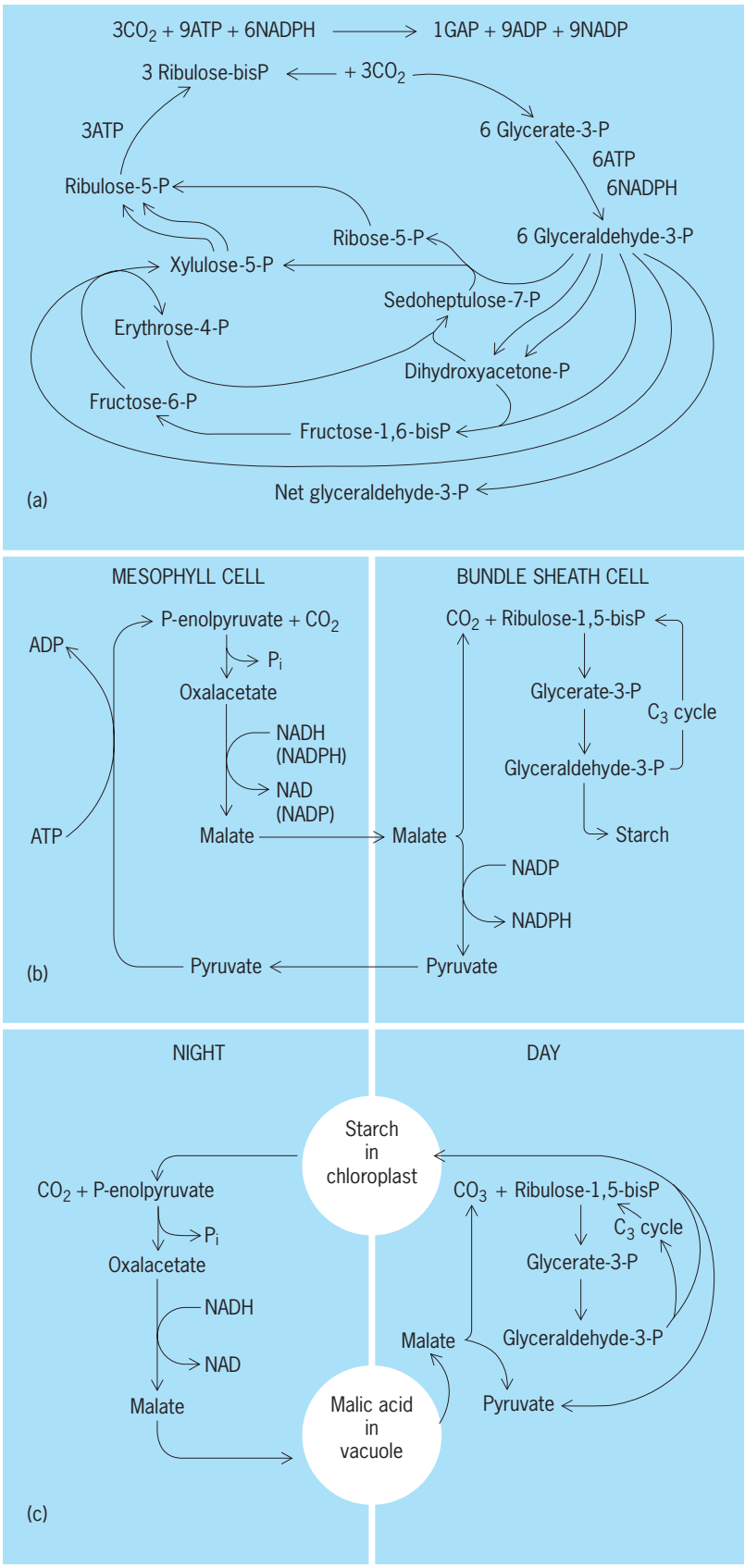

Fig. 4. Photosynthesis reactions. (*a*) C_3 photosynthesis. The reductive pentose cycle of photosynthesis in which 3 moles of CO_2 result in 1 net glyceraldehyde-3-P and the two modifications, C_4 photosynthesis and CAM photosynthesis. (*b*) C_4 photosynthesis. (*c*) CAM photosynthesis. In *b* and *c*, there is an initial carboxylation that forms malate or another four-carbon compound, followed by decarboxylation and subsequent assimilation to form glyceraldehyde-3-P as in the reductive pentose cycle.

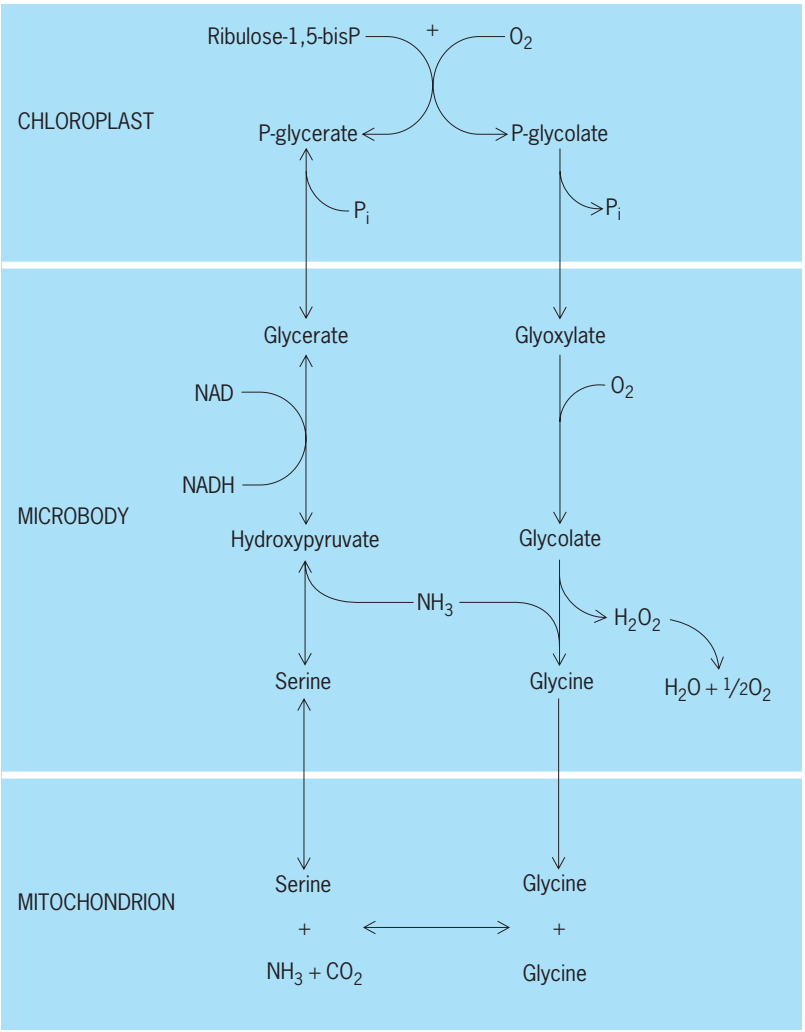

Fig. 5. Glycolate and glycerate pathways of photorespiration. There are two sites of oxygen consumption, one at ribulose bisphosphate carboxylase/oxygenase in chloroplasts and one at glycolate oxidase in the microbodies. Complete oxidation of glycine and serine results in CO_2 production. Glycerate could enter the pathway from the chloroplasts, or alternatively, it could be salvaged from the microbodies.

is assimilated by the C_3 process. Since CO_2 is taken up through open stomata at night when the evaporative demand is low, water loss is minimized. The transpiration ratio (water loss to carbon gained) is 5 to 10 times less in CAM than in C_3 and C_4. Thus CAM is an adaptation to drought and water-stressed environments.

Some CAM plants shift from C_3 to CAM in response to water stress. Thus CAM is adaptive and flexible, presenting interesting research possibilities. *See* PHOTOSYNTHESIS.

Respiration. Respiration is the oxidative degradation of organic compounds to yield usable energy usually in the form of reduced pyridine nucleotides (NADH and NADPH) and triphosphate nucleotides (such as ATP). It is a process of all living organisms, but in plants photosynthate is the immediate substrate. The process starting with hexose is shown in reaction (6). Complete combustion of hexose yields

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6H_2O + 6CO_2 + \text{energy} \qquad (6)$$

about 2900 kJ mol^{-1}. During respiration there is yielded about 36 ATP or ATP equivalents (free energy of hydrolysis of ATP = 30 kJ mol^{-1}) with an energy conservation efficiency of about 37%.

The oxidative process ordinarily starts with glycolysis utilizing hexose-phosphate as the substrate (**Fig. 6**), ultimately yielding P-enolpyruvate, which can be converted to pyruvate by catalysis with pyruvate kinase. There is a yield of about 6 ATP with an energy conservation of about 31%, because oxidation of hexose to 2 pyruvate yields about 6 ATP. If the complete oxidation of hexose is considered (2900 kJ mol^{-1}), glycolysis is only 6% efficient. To complete respiration, glycolysis couples to the tricarboxylic acid cycle in the mitochondria, conserving 37% of the energy as explained above.

An alternative oxidative pathway for hexose is through the pentose phosphate pathway (Fig. 6), yielding about 24 moles of ATP for an energy conservation of about 25%.

Organic acids are oxidized in the tricarboxylic acid cycle. Pyruvate in the mitochondria is decarboxylated by pyruvate dehydrogenase, a complex reaction involving several steps [reaction (7)].

$$\text{Pyruvate} + \text{CoA} + \text{NAD} \rightarrow \text{acetyl CoA} + \text{NADH} + CO_2 \qquad (7)$$

The acetyl CoA can act as a substrate. Alternatively, P-enolpyruvate still in the cytosol can be carboxylated by P-enolpyruvate carboxylase, forming oxalacetate. Oxalacetate is reduced by malate dehydrogenase, forming malate. In plants, malate is a better substrate for entry into the mitochondria than pyruvate.

Although the complete oxidation of acetate or other organic acid such as malate in the tricarboxylic acid cycle in the mitochondria yields some ATP (see Fig. 6), the reduced compounds flavins adenine dinucleotide ($FADH_2$) and NADH are usually formed. These two compounds feed electrons into a chain of oxidation-reduction electron transport compounds, collectively called the electron transport system, and ultimately to O_2 (**Fig. 7**). The oxidation of $FADH_2$ and NADH by O_2 is sufficient to form 2 or 3 ATP with a yield of about 24 moles ATP from the 2 moles of acetate.

Cyanide, carbon monoxide, azide, and antimycin A will inhibit respiration. In some plant tissues, however, such as aged potato tissue and in the flowers of Araceae, respiration continues in their presence, but only 2 ATP are formed rather than the usual 6 per electron pair transferred to oxygen. This observation indicates that two of the sites are bypassed presumably by an alternative oxidase system (Fig. 7). The excess energy is lost as heat.

Oxidative phosphorylation is driven by chemiosmotic coupling similar to photophosphorylation in chloroplasts.

The respiratory quotient (RQ) of plant tissues is close to 1, indicative of hexose oxidation, but during germination of seeds that store carbohydrate, RQs increase from about 1 to near 1.3, suggesting that the

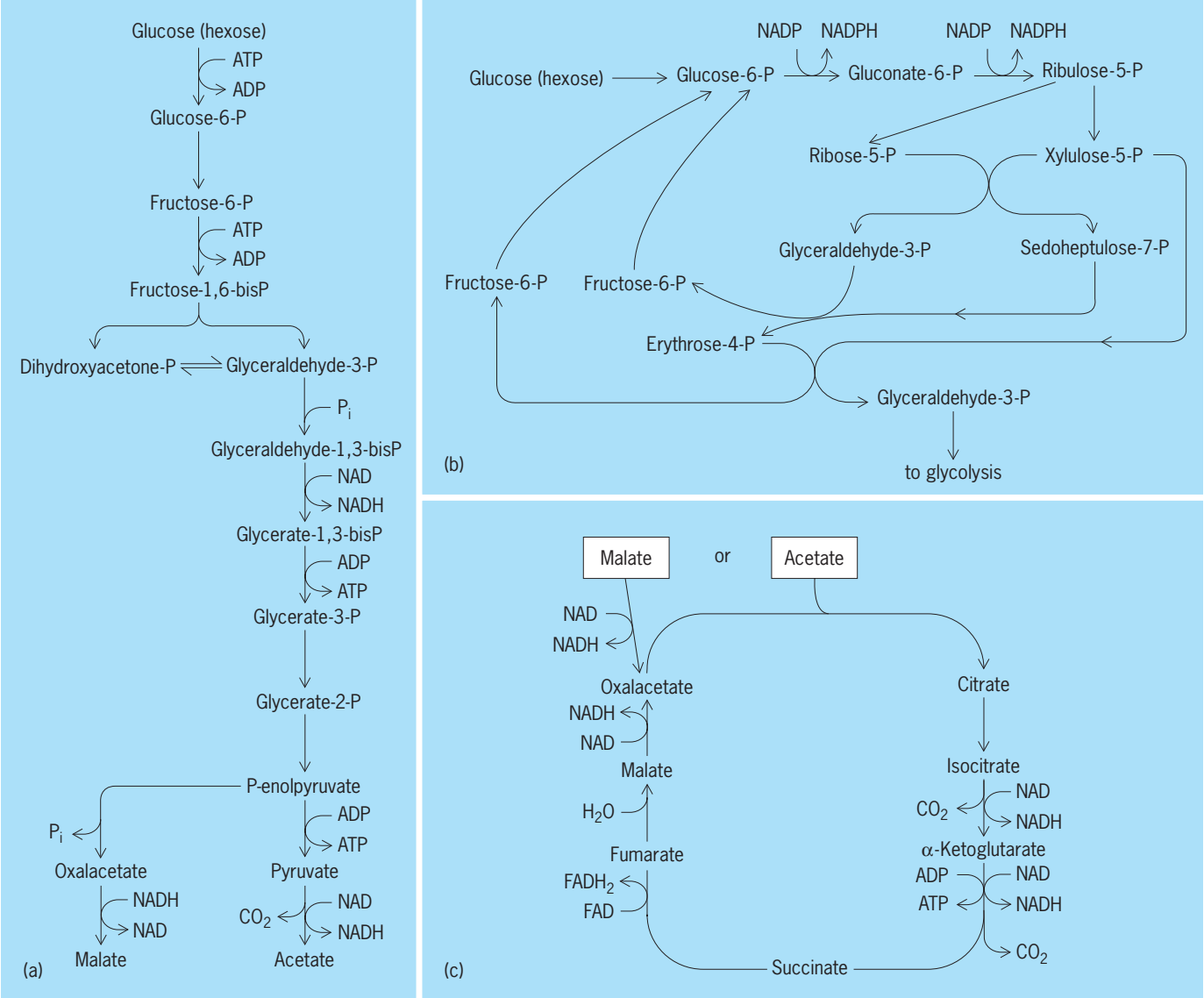

Fig. 6. Glycolysis and the generation of acetate and malate that act as (*a*) tricarboxylic acid cycle substrates, (*b*) the pentose phosphate pathway of hexose oxidation, and (*c*) the tricarboxylic acid cycle.

substrate shifts from carbohydrate to organic acids. During germination of seeds that store lipid, the RQ is low, indicative of fatty acid oxidation. *See* PHOTORESPIRATION; PLANT RESPIRATION.

Organic Constituents of Plants

The principal organic compounds of plants are nucleic acids, amino acids, proteins, carbohydrates, organic acids, lipids, and natural products.

Nucleic acids. Nucleic acids, found in virtually all plant cells, are complex polymers with masses of 10^7 daltons or more for native deoxyribonucleic acid (DNA) and 10^6 daltons for some kinds of ribonucleic acids (RNA). DNA contains the genetic code, and RNA functions in protein synthesis and regulation. *See* GENETIC CODE.

Nucleic acids are polymers of substituted purines and pyrimidines. DNA and RNA are composed primarily of adenine, guanine, cytosine, 5-methylcytosine, and thymine, except that RNA has uracil rather than thymine. Also, small quantities of rare, substituted bases probably function as recognition sites. In the nucleic polymer, the bases exist as nucleotides in which there is a ribose (in RNA) or a deoxyribose (in DNA) bridge between the base and a phosphate. DNA and RNA polymers are formed by polymerization of the deoxyribotides (DNA) or ribotides (RNA) through phosphate diester bonds. Ordinarily, the DNA exists in a double-helical form, whereas the RNA is single-stranded.

The central dogma of cell biology is diagrammed in **Fig. 8**. DNA of the genome (such as in the nucleus) has the information for protein synthesis. This information is transcribed into messenger RNA (mRNA). The mRNA migrates into the cytoplasm and in cooperation with ribosomal RNA and ribosomes is translated and proteins as synthesized. The process is aided by transfer RNAs acting as amino acid carriers. After peptide and protein synthesis, there may be further modification to form active enzymes.

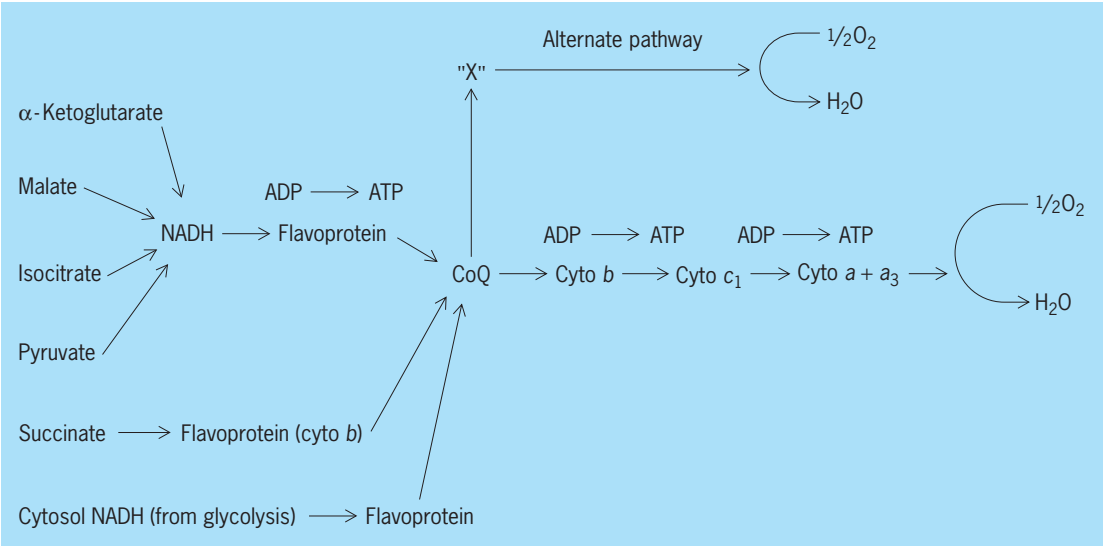

Fig. 7. Scheme for electron transport to oxygen in mitochondria during respiration. Entry of electrons from the tricarboxylic acid cycle intermediates—α-ketoglutarate, malate, isocitrate, succinate—and from pyruvate and NADH from the cytosol are shown. In addition, the alternative, cyanide resistance pathway is shown. Note that there are three "sites" where ATP is formed. Cyto = cytochrome. CoQ = a quinone.

Chloroplasts and mitochondria also have a genome composed of circular DNA not unlike prokaryote DNA. The chloroplast DNA has a mass of about 9×10^7 daltons. Chloroplasts and mitochondria also have RNA and ribosomes with the capacity for protein synthesis, but much of their protein is the result of the nuclear genome. Some organelle membrane proteins are coded for by the organelle genomes. In the special case of the chloroplast, the large subunit of the carboxylating protein, ribulose bisphosphate carboxylase, is coded for by the chloroplast-DNA, while the small subunit is coded for by the nuclear-DNA. *See* CELL PLASTIDS; MITOCHONDRIA.

Polypeptide synthesis systems have been isolated, primarily from wheat germ. Using plant DNA, nucleotides, RNA polymerase enzymes, and other factors, mRNA can be produced in culture. Polypeptides can be prepared with the purified mRNA, amino acids, and proper cofactors.

Nucleotides in plants are of importance to metabolism such as the oxidation-reduction factor, nicotinamide adenine dinucleotide (NAD) and its phosphate derivative (NADP), flavin mononucleotide (FMN), and flavin adenine dinucleotide (FAD). Phosphorylated nucleotides such as ATP and guanosine triphosphate (GTP) are important in cellular energetics driving reactions and processes requiring energy. *See* DEOXYRIBONUCLEIC ACID (DNA); NUCLEIC ACID; RIBONUCLEIC ACID (RNA).

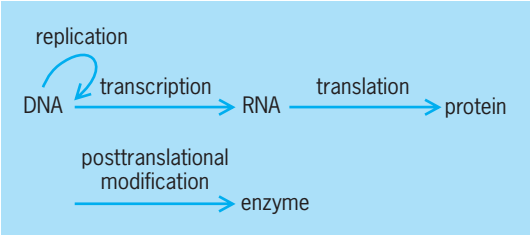

Fig. 8. Diagram of the central dogma of cell biology.

With the present state of knowledge about nucleic acid metabolism and genetics, genetic manipulation through recombinant DNA technology is possible. The ultimate goal of genetic engineering of plants through biotechnology is to transfer genetic information in the form of genes from one organism to another in order to alter the metabolism of the host plant. Specific tasks under intensive study using these techniques include the improvement of disease resistance; the development of agricultural plants that are resistant to herbicides; improved cold tolerance; increase in the quality of seed storage proteins for human consumption; increased general stress tolerance to salinity, drought, and heat; introduction of the capacity for nitrogen fixation; and increase in the efficiency of photosynthesis, perhaps by reducing photorespiration. In some cases, disease resistance and tolerance to herbicides are believed to be controlled by single genes, and so the introduction of a single gene may confer resistance of the recipient or transformed plant. Other objectives are more difficult to achieve. Nitrogen fixation in legumes, for example, is a multigene-regulated phenomenon that depends both on the host (legume plant) and the infectious symbiotic nitrogen-fixing bacterium (*Rhizobium* species). Thus the simple introduction of the genetic information for the synthesis of the nitrogen-fixing gene nitrogenase is unlikely to accomplish the goal. Similarly, tolerance to cold, drought, heat, and perhaps salt are likely multigene traits and are probably best handled through conventional plant breeding technology. *See* GENETIC ENGINEERING.

Amino acids and proteins. About 22 amino acids are found in proteins, and probably many others play some role in plant metabolism. Most are of the L configuration, but a few of the nonprotein amino acids

are D. Polymerization of amino acids by covalent linkage of an amino group of one and the carboxyl of another (CO—NH) forms peptides. A few small plant peptides such as glutathione (a tripeptide) are found, but they are not common. Larger polypeptides making up proteins are important as storage for nitrogen and carbon, as structural components of membranes and, to a lesser extent, cell walls, and as enzymes that regulate cellular reactions.

Seed storage proteins are important from the standpoint of energy reserves in the developing seedling. They are synthesized during seed development and stored in membrane-bound protein bodies. During germination, the proteins rapidly disappear as the nitrogen and amino acids are mobilized for growth and development of the young seedling. Because of considerations for human nutrition, most is known about storage proteins in cereals and legumes. Almost 70% of the world's protein supply comes from plants, with 35% from seed proteins. Although storage proteins are species-specific (such as zein in corn), they can readily be manipulated through traditional plant breeding technology. For the most part, the storage proteins of cereals are prolamines that are deficient in lysine and frequently also in tryptophan; hence, by themselves, they are poor protein sources from a human nutritional viewpoint. Rice is, however, high in lysine, and the artificial genus *Triticale*, a cross between wheat and rye, is also high in lysine. In legumes, which have substantially more storage protein than cereals, the proteins are mostly of the globulin type, which have adequate lysine but are deficient in sulfur amino acids such as the essential methionine. *See* CEREAL; LEGUME; SEED; TRITICALE.

Two kinds of membrane proteins can be differentiated: peripheral proteins that are easily removed by dilute salt treatment, and integral proteins that are tightly bound. Both membrane proteins play a role in metabolism as enzymes and may function in transport. The cell wall structural protein extensin is rich in hydroxyproline.

Enzyme proteins are of two types: Constitutive proteins that are always present include those regulating respiration, fatty acid synthesis, and protein synthesis. Adaptive enzymes appear during specific periods. As an example, enzymes of the glyoxylate cycle that couple fatty acid oxidation during lipid breakdown to carbohydrate synthesis are present just during germination.

Many membrane proteins evidently function in transport, although there is no direct evidence. Soluble proteins may play a role in the regulation of enzymic proteins; for example, there is an activating protein called activase for ribulose bisphosphate carboxylase. Many proteins are activated by inorganic elements such as calcium: environmental signals such as light bring about membrane transport of calcium, which acts as a secondary messenger binding to proteins and activating them. One such specific protein of plants is calmodulin, which is activated by calcium and in turn regulates metabolism at the level of protein activity.

Enzymes exist in more than one form. Multiple forms of one enzyme are called isozymes and have the same catalytic function but different properties. They are readily separated from each other by electrophoresis, usually on starch or acrylamide gels. Isozymes may function in different pathways that have a common reaction; they may be localized in different cellular compartments, or they may be redundant, acting in conservation during evolution, or simply be the result of heterozygosity. Isozymes may be nonallelic (different gene loci) or allelic (same locus). Most nonallelic isozymes will also exist as allelic isozymes since most gene loci are heterozygous. Also, isozymes may result from posttranslational modifications. The carbohydrate moiety of glycoproteins (such as peroxidases) may differ, resulting in isozymes.

There are three nonallelic isozymes of malate dehydrogenase, each of which occurs in a different compartment (microbodies, mitochondria, and the cytosol), and each functions in a different metabolic pathway (the glyoxylate cycle or photorespiration in microbodies, the tricarboxylic acid cycle in mitochondria, and organic acid synthesis in the cytosol; **Fig. 9**). These nonallelic isozymes are known to exist as allelic isozymes, the result of heterozygosity at individual gene loci. Isozymes are known for almost all enzymes that are found in plants.

Little is known about the regulation of protein synthesis, but it is assumed that inductive and repression-derepression systems are functional in plants similar to those in animals and bacteria. Light plays an important role in the regulation of protein synthesis through the blue pigment phytochrome. Phytochrome is a protein with an open tetrapyrrole chromophore that exists in two forms, red-absorbing form (P_r) that is converted to a far-red-absorbing form (P_{fr}). The P_{fr} form, which absorbs red light and is converted back to the P_r form, is considered to be the active form and is associated with changes at the genetic level, bringing about mRNA and protein biosynthesis. It plays an important role in the light-induced flowering process and in germination. In addition, phytochrome is believed to function at the membrane level, bringing about changes in transport, perhaps of calcium, that ultimately affect growth processes. The synthesis and regulation of many enzyme proteins of the chloroplast are light-mediated, also probably through the action of phytochrome. *See* PHYTOCHROME.

The mechanism of activity regulation occurs primarily by allosteric effects. Many enzymes have a site or sites other than the active catalytic site that binds small molecules (the allosteric site). Binding alters the catalytic activity, in some cases enhancing it and in others inhibiting it. The enzyme phosphofructokinase (PFK) catalyzes the conversion of fructose-6-P to fructose-1,6-bisphosphate. It is inhibited by ATP, phosphoenolpyruvate, and citrate. When phosphoenolpyruvate and citrate accumulate, their further synthesis is prevented by feedback inhibition of one of the enzymes (PFK) that leads to their synthesis.

P-enolpyruvate carboxylase, an enzyme involved

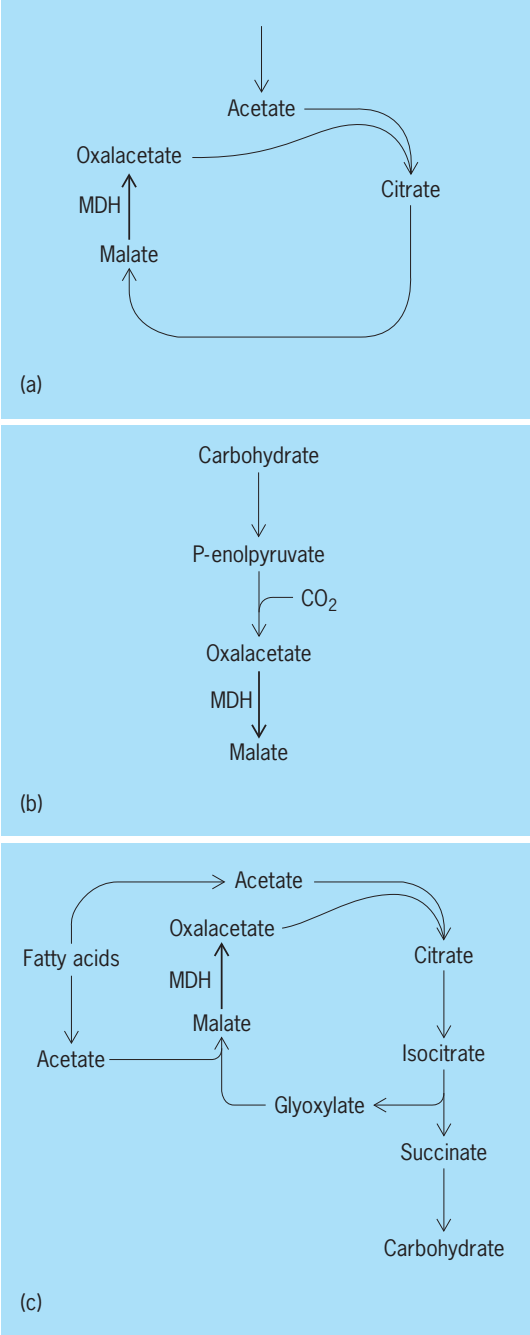

Fig. 9. Function of malate dehydrogenase (MDH) isozymes in different metabolic pathways. The abbreviated metabolic charts show the common steps of the (*a*) tricarboxylic acid cycle in mitochondria, (*b*) dark CO_2 fixation in cytosol, and (*c*) the glyoxylate cycle in microbodies called glyoxysomes that have MDH.

in malate synthesis, is also known to be an allosteric enzyme of plants. Coupled with malate dehydrogenase, malate is synthesized from P-enolpyruvate [reactions (8)]. Glucose-6-P, a precursor of PEP, activates

$$\begin{aligned} \text{PEP} + \text{HCO}_3^- &\rightarrow \text{oxalacetate} + \text{P}_i \\ \text{Oxalacetate} + \text{NADH} &\rightarrow \text{malate} + \text{NAD} \end{aligned} \tag{8}$$

PEP carboxylase and malate inhibits it. Thus PEP carboxylase is regulated by both a precursor and a product of the metabolic sequence. *See* AMINO ACIDS; ENZYME; PROTEIN; PROTEIN METABOLISM.

An interesting group of proteins is the agglutinins, which will agglutinate erythrocytes; they are also known as lectins. They have the capacity to recognize specific carbohydrate sites on macromolecules such as those of cell walls. Common examples include ricin from castor bean and concanavalin A from jack bean. Although their physiological function is not entirely clear, they may play a role in cell-to-cell recognition. For example, it is believed that they are responsible for the recognition of specific roots by the symbiotic nitrogen-fixing bacteria of the genus *Rhizobium*. Recognition accounts for the legume species' specificity of the rhizobia. They also may be important in protection against plant pathogens by specific binding and agglutination. *See* LECTINS.

Nitrogen. Despite being the most abundant element in the atmosphere, atmospheric nitrogen cannot be utilized directly for plant metabolism. The first step in the incorporation of nitrogen into plant organic matter is nitrogen fixation through symbiotic or free-living nitrogen-fixing prokaryotic organisms. Nitrogenase catalyzes the fixation of atmospheric nitrogen [reaction (9)]. Nitrogenase is a large complex

$$\text{N}_2 + 6e^- + n\text{MgATP} \rightarrow 2\text{NH}_4^+ + n\text{MgADP} + n\text{P}_i \tag{9}$$

protein containing both molybdenum and iron that is found only in prokaryotes; thus no plant is capable of nitrogen fixation. After NH_4^+ formation, there is a two-step enzymatic reaction incorporating the fixed nitrogen into amino acids. Alternatively, NO_3^- in the soil is available to plants. After uptake, the NO_3^- is reduced to ammonia (NH_3) or ammonium (NH_4^+), depending upon pH. The two-step reaction is catalyzed by nitrate reductase to form nitrite [reaction (10)]. The nitrite is then reduced to ammonia by nitrite reductase [reaction (11)]. Actual as-

$$\text{NO}_3^- + 2\text{H}^+ + 2e^- \rightarrow \text{NO}_2^- + \text{H}_2\text{O} \tag{10}$$

$$\text{NO}_2^+ + 6e^- + 6\text{H}^+ \rightarrow \text{NH}_3 + \text{H}_2\text{O} + \text{OH}^- \tag{11}$$

similation of the ammonia takes place by a second two-step reaction [(12) and (13)]. It is catalyzed first

$$\text{Glutamate} + \text{NH}_3 + \text{ATP} \rightarrow \text{glutamine} + \text{ADP} + \text{P}_i \tag{12}$$

$$\text{Glutamine} + \alpha\text{-ketoglutarate} \rightarrow 2\ \text{glutamate} \tag{13}$$

by glutamine synthetase and then by glutamate synthase, with the sum of the two reactions being glutamate, adenosine diphosphate, and inorganic phosphate [reaction (14)]. Once the nitrogen is an in-

$$\alpha\text{-Ketoglutarate} + \text{ATP} + \text{NH}_3 \rightarrow \text{glutamate} + \text{ADP} + \text{P}_i \tag{14}$$

tegral component of glutamate, other nitrogenous organic compounds can be synthesized by catalysis with transaminases and other enzymes of nitrogen metabolism. *See* NITROGEN FIXATION.

Carbohydrates and organic acids. Carbohydrates are among the most abundant biochemicals constituting the initial products of photosynthesis and the primary substrates for respiration. When phosphorylated, they are metabolically reactive. The six-carbon

hexose sugars glucose and fructose when phosphorylated are primary substrates for respiration, and the disaccharide sucrose is the primary organic compound transported, although other nonreducing sugars such as raffinose, stachyose, and verbascose are translocated in some species. The simple monosaccharides in plants may have an amino group (the amino sugars), a carboxyl group (the sugar acids), or exist as O-glycosides.

Oligosaccharides have six or fewer monosaccharides linked by an O-glycosidic bond between the aldehyde group of one to a hydroxyl of another. The larger polysaccharides are important from a structural and storage role. The most abundant carbohydrate is cellulose, a major constituent of cell walls. It is a straight-chain polymer of glucose linked through β-(1→4) O-glucosidic bonds and may be 200,000 to 2×10^6 daltons in mass. The common storage polysaccharide of plants, starch, is a mixture of branched (amylopectin) and straight chains (amylose). The linkages are α-(1→4) O-glucosidic bonds in amylose, and amylopectin and the branching of amylopectin is through α-(1→6) O-glucosidic bonds. The fructosans inulin with α-(2→1) linkages and levan with α-(2→6) linkages are other storage polysaccharides.

Their biosynthesis is poorly understood, but sucrose may illustrate a general mechanism. One enzyme that can catalyze sucrose synthesis is sucrose phosphate synthase. Sucrose phosphate is formed from uridine diphosphate glucose and fructose-6-phosphate [reaction (15)], followed by hydrolysis of

$$\text{UDP-glucose} + \text{fructose-6-P} \rightarrow \text{sucrose-P} + \text{UDP} \quad (15)$$

$$\text{Sucrose-P} \rightarrow \text{sucrose} + P_i \quad (16)$$

the sucrose-phosphate [reaction (16)]. The biosynthesis of other polysaccharides takes place by similar mechanisms in which the monosaccharides as nucleotide derivatives (usually of adenosine diphosphate or uridine diphosphate) are added to an existing saccharide, resulting in chain elongation [reaction (17)]. Breakdown of the polysaccharides is usually by hydrolysis [reaction (18)].

$$n\text{ADP-glucose} + \text{saccharide primer} \rightarrow \text{polysaccharide} + n\text{ADP} \quad (17)$$

$$\text{Starch} + nH_2O \rightarrow n\ \text{maltose} \quad (18a)$$

$$\text{Maltose} + H_2O \rightarrow 2\ \text{glucose} \quad (18b)$$

In seeds that store starch, germination is accompanied by hydrolysis of starch according to reactions (18) to produce the hexoses glucose and fructose. These hexoses are converted to sucrose according to reactions (15) and (16). The sucrose is transported to growing and developing sites within the young seedling and is used as the primary energy source, being metabolized to CO_2 and water and producing energy in the form of ATP.

Cell wall mannans (polymers of mannose), xylans (polymers of xylose), and mixed polysaccharides with more than one kind of sugar are called hemicelluloses, having properties similar to cellulose. The pectins and pectic acids are common polysaccharides of cell walls and are mixtures of arabans, galactans, and galacturonic acid.

The gums and mucilages are hydrophilic polymers of rhamnose, arabinose, xylose, galactose, glucuronic acid, and galacturonic acid. Because of their large water-holding capacity, they may play a part in plant-water relations or simply be storage carbohydrates.

Organic acids derived from the carbohydrates are substrates in the tricarboxylic acid cycle (such as citrate, malate, succinate, and isocitrate), play a role as counterions, in osmotic balance, and as buffers, and are products of photosynthesis. Plants accumulate specific organic acids. Thus apple has malic acid, citrus has citric acid, *Oxalis* has oxalic acid, *Fumaria* accumulates fumaric acid, and grapes accumulate tartaric acid.

Organic acids are highly compartmented in plant cells. Malic acid occurs in mitochondria as one of the substrates for respiration [reaction (19)], and

$$\text{Fumaric acid} + H_2O \rightarrow \text{malic acid} \quad (19a)$$

$$\text{Malic acid} + \text{NAD} \rightarrow \text{oxalacetic acid} + \text{NADH} \quad (19b)$$

there is another pool in the cytosol (and perhaps vacuole) that does not mix with the mitochondrial pool. This malic acid is synthesized by a carboxylation of P-enolpyruvate as mentioned previously. The enzymes catalyzing the common reaction sequences are isozymes. *See* CARBOHYDRATE; CARBOHYDRATE METABOLISM.

Lipids. Plant lipids are a heterogeneous group of compounds that include the neutral lipids (fats and oils), phospholipids, glycolipids, terpenoids, and waxes having the common characteristic of being hydrophobic. They are soluble in organic solvents and slightly soluble in water. They are storage compounds, components of membranes, and protective substances of exposed surfaces such as cutin and suberin.

Neutral lipids yield fatty acids and glycerol upon hydrolysis [reaction (20)]. They can be saponified

$$\begin{array}{l} CH_2OOR \\ | \\ CH\ OOR \\ | \\ CH_2OOR \end{array} + 3H_2O \longrightarrow \begin{array}{l} CH_2OH \\ | \\ CH\ OH \\ | \\ CH_2OH \end{array} + 3ROOH \quad (20)$$

Triglyceride Glycerol Fatty acids

with alkali such as NaOH or KOH to form glycerol and the salts of fatty acids (the soaps) [reaction (21)].

$$\begin{array}{l} CH_2OOR \\ | \\ CH\ OOR \\ | \\ CH_2OOR \end{array} + 3KOH \longrightarrow \begin{array}{l} CH_2OH \\ | \\ CH\ OH \\ | \\ CH_2OH \end{array} + 3ROOK \quad (21)$$

Fatty acids are usually even-numbered and include butyric acid (4 carbons), caproic acid (6), caprylic acid (8), capric acid (10), lauric acid (12), myristic acid (14), palmitic acid (16), stearic acid (18),

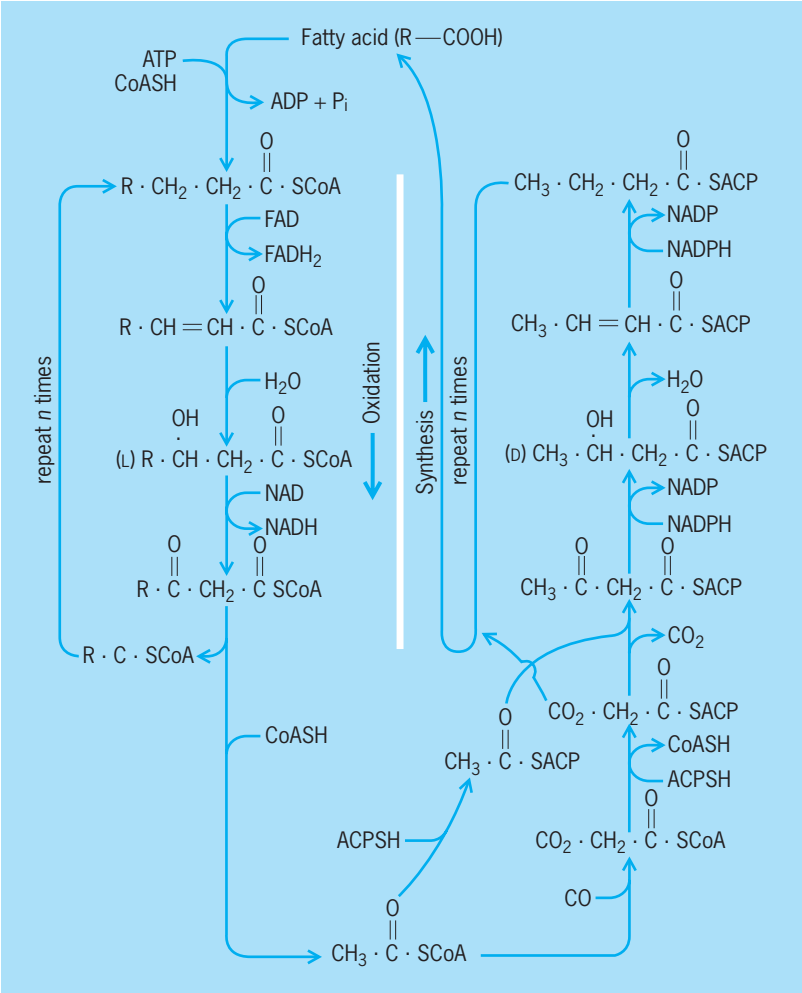

Fig. 10. Metabolic flow chart for fatty acid metabolism showing oxidation (β-oxidation) and biosynthesis. Note that in the synthetic pathway, NADPH is the reducing agent, the D form of the β-hydroxyketo acid is formed, and the active intermediates are derivatives of acyl carrier protein (ACP-SH), whereas in the oxidative direction, NAD is the oxidizing agent, the L form of the β-hydroxyketo acid is formed, and the active intermediates are derivatives of coenzyme A (CoA-SH).

arachidic acid (20), behenic acid (22), lignoceric acid (24), and cerotic acid (26). Palmitic acid is the most common, followed by lauric, myristic, and stearic acids. The unsaturated fatty acids oleic acid, linoleic acid, and linolenic acid—18-carbon fatty acids with one, two, and three double bonds—are also abundant.

The phospholipids of membranes are substitution products of phosphatidic acid (**1**). Free phosphatidic

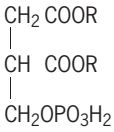

(**1**)

acid is rare in plants, but the phosphatides phosphatidylglycerol, phosphatidylethanolamine, phosphatidylcholine, phosphatidylserine, and phosphatidylinositol derived from phosphatidic acid by bonding to the phosphate are common.

Important constituents of chloroplast membranes are the glycolipids, monogalactosyl-diglyceride and digalactosyl-diglyceride. Unsaturated fatty acids such as linolenic acid are frequent components of the R groups.

The waxes that function primarily in protection differ from the triglycerides in that long-chain alcohols replace the glycerol.

Some plants store waxes rather than triglycerides. Jojoba (*Simmondsia chinensis*) stores a wax ester with an average content of 40 to 42 carbon atoms and is composed of a long-chain monounsaturated alcohol and a long-chain monounsaturated fatty acid. This wax, which is similar in properties to sperm whale oil, is the primary energy source for the germination and initial growth of the jojoba plant. The wax has attracted attention because of its potential use as a lubricant and a nonnutritive oil for human use. It is presently used extensively in cosmetics. *See* JOJOBA; WAX, ANIMAL AND VEGETABLE.

Biosynthesis of the saturated fatty acids begins with acetyl coenzyme A (CoA) adding to a small sulfur-containing carrier protein (ACP-SH; **Fig. 10**). Through the action of a specific enzyme the acetyl group is added to malonyl-S-ACP, forming acetoacetyl-S-ACP (**2**). Reduction of the beta keto

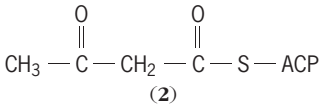

(**2**)

group results in butyric acid. Subsequent additions of acetyl groups cause chain elongation. Unsaturation of the final product would produce unsaturated fatty acids, but evidence indicates that unsaturation may occur before chain elongation is finished.

Oxidation of fatty acids occurs by a mechanism similar to the biosynthesis but in reverse (Fig. 9). First, the triglyceride is hydrolyzed to glycerol and free fatty acids. The β-oxidation pathway proceeds by formation of the coenzyme A derivative (R-SCoA), followed by oxidation to form a double bond between the alpha and beta carbons. Addition of water and then oxidation to form the beta keto group allows displacement with CoA-SH to form acetyl CoA and a new fatty acid CoA derivative with two less carbons. The process is continued until only acetyl CoA remains. Alpha oxidation removes one carbon at a time rather than two, and probably plays a role in the metabolism of odd-numbered fatty acids.

In seeds that store oil and lipids, such as castor bean and soybean, the mobilization of reserves is much more complicated than in starch-storing seeds and tubers. The triglycerides or wax esters are metabolized as discussed above to produce acetate. The acetate as the acetyl CoA derivative is metabolized in the glyoxylate cycle (Fig. 9*c*) to produce succinate and ultimately carbohydrate by reactions of gluconeogenesis. The latter is essentially a reversal of glycolysis (Fig. 6*a*). Succinate is converted to malate (Fig. 6*a*), and then the synthesis of P-enolpyruvate starts the pathway to hexose. The enzyme responsible for the catalysis is most likely P-enolpyruvate

carboxykinase [reaction (22)], effectively overcom-

$$\text{Oxalacetate} + \text{ATP} \rightarrow \text{P-enolpyruvate} + \text{ADP} + CO_2 \quad (22)$$

ing the energy barrier of the reaction from acetate to P-enolpyruvate. Subsequent reactions up to fructose-1, 6-bisP proceed with little difficulty. The conversion from fructose-1, 6-bisP to fructose-6-P is catalyzed by a phosphatase [reaction (23)]. Fructose

$$\text{Fructose-16-bisP} + H_2O \rightarrow \text{fructose-6-P} + P_i \quad (23)$$

is converted to glucose, and then glucose and fructose are converted to sucrose catalyzed by sucrose phosphate synthase [reactions (15) and (16)]. The sucrose is transported to sites of energy requirement. *See* LIPID; LIPID METABOLISM.

Natural products. Natural products are those compounds synthesized by secondary reactions from the primary carbohydrates and amino acids. They include many medicines, narcotics, poisons, stimulants, essential oils, resins, and compounds with teratogenic and carcinogenic properties. Such compounds as opium, quinine, digitoxin, cocaine, caffeine, and rubber are well known. They may be storage or end products. Many are allelochemics that affect the growth, health, or behavior of other organisms such as repellents, growth retardants, or attractants for pollinators and herbivores. *See* SECRETORY STRUCTURES (PLANT).

Natural products incllude phenolics, alkaloids, terpenoids, and porphyrins. Phenolics have an aromatic ring and at least one hydroxyl group such as the cinnamic acids, coumarins, ligini, flavonoids, and tannins. The polyketide phenolics are similar to the above except their biosynthesis is different. Rather than resulting from the shikimate pathway (**Fig. 11**), they are the result of cyclization of linear poly-β-keto chains (**3**). They frequently have hydroxyls in

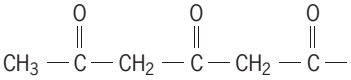

(**3**)

the meta positions, unlike those from the shikimate pathway. *See* PHENOL.

Alkaloids have basic properties because of a nitrogenous heterocyclic ring or rings. They include piperidine alkaloids, aromatics, betacyanins, isoquinoline and indole alkaloids, purine alkaloids, and aliphatic alkaloids. They are mainly synthesized from the amino acids tryptophan, phenylalanine, glutamate, orinithine, and lysine. *See* ALKALOID.

Terpenoids are lipids that may be acyclic or cyclic, with a few to several hundred carbon atoms. They include essential oils, steroids, and large polymers such as rubber. They are synthesized by condensation of isoprene (isopentenyl pyrophosphate), usually by a head-to-tail linkage. They are grouped according to number of isoprene units in their backbone: hemiterpenoids with one, monoterpenoids with two, sesquiterpenoids with three, diterpenoids with four, triterpenoids with six, tetraterpenoids with eight, and polyterpenoids with more than eight. *See* RUBBER; TERPENE.

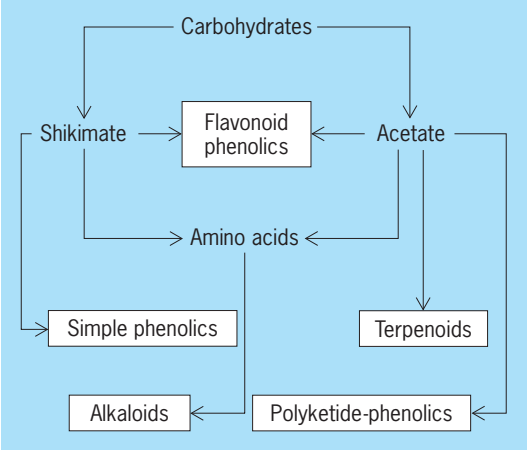

Fig. 11. Abbreviated flow chart for the synthesis of the various natural products found in plants.

Porphyrins are compounds synthesized from pyrrole usually by cyclization of four pyrroles to form a tetrapyrrole (**4**). Substitutions of the pyrrole rings

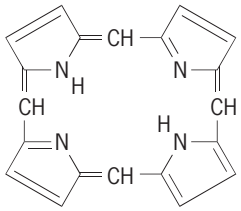

(**4**)

result in the different prophyrins. Magnesium porphyrin is chlorophyll, and the iron porphyrins of plants are the cytochromes. The pyrrole ring is synthesized from δ-aminolevulinic acid that is produced by condensation of succinate and glycine [reaction (24)]. In chlorophyll biosynthesis, however, there

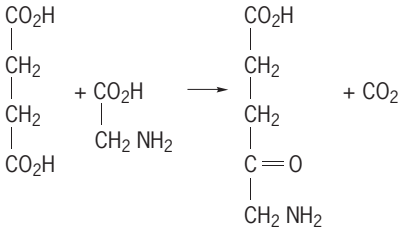

(24)

is evidence that the δ-aminolevulinic acid comes from glutamate rather than succinate and glycine. *See* CHLOROPHYLL; PLANT CELL; PLANT GROWTH; PLANT PHYSIOLOGY; PORPHYRIN.

Irwin P. Ting

Bibliography. G. Edwards and D. A. Walker, *C_3, C_4: Mechanisms, and Cellular and Environmental Regulations of Photosynthesis*, 1983; D. Grierson and S. N. Covey, *Plant Molecular Biology*, 1988; P. W. Ludden and J. E. Burris, *Nitrogen Fixation and CO_2 Metabolism*, 1985; T. Robinson, *The Organic Constituents of Higher Plants*, 5th ed., 1983; P. K. Stumpf and E. E. Conn, *The Biochemistry of Plants: A Comprehensive Treatise*, 16 vols., 1980–1988; I. P. Ting, *Plant Physiology*, 1982.

Plant mineral nutrition

The relationship between plants and all chemical elements other than carbon, hydrogen, and oxygen in the environment. Plants obtain most of their mineral nutrients by extracting them from solution in the soil or the aquatic environment. Mineral nutrients are so called because most have been derived from the weathering of minerals of the Earth's crust. Nitrogen is exceptional in that little occurs in minerals: the primary source is gaseous nitrogen of the atmosphere.

Mineral Nutrients in Relation to Growth

Some of the mineral nutrients are essential for plant growth; others are toxic, and some absorbed by plants may play no role in metabolism. Many are also essential or toxic for the health and growth of animals using plants as food.

Essential nutrients. In 1699 John Woodward published experiments describing how the growth of spearmint in water increased with the extent to which the water was contaminated by soil. His water culture, or hydroponics, technique has been progressively refined and used extensively in subsequent research aimed at identifying the specific components of his "peculiar terrestrial matter," and six basic facts have been established: (1) plants do not need any of the solid materials in the soil—they cannot even take them up; (2) plants do not need soil microorganisms; (3) plant roots must have a supply of oxygen; (4) all plants require at least 14 mineral nutrients; (5) all of the essential mineral nutrients may be supplied to plants as simple ions of inorganic salts in solution; and (6) all of the essential nutrients must be supplied in adequate but nontoxic quantities. *See* HYDROPONICS.

These facts provide a conceptually simple definition of and test for an essential mineral nutrient. A mineral nutrient is regarded as essential if, in its absence, a plant cannot complete its life cycle. The test for an essential mineral nutrient involves growing plants in water cultures containing a mixture of salts of all the elements thought to be essential: omitting any essential nutrient prevents growth (**Fig. 1**).

This test worked well for those nutrients which plants require in relatively large amounts, that is, the essential macronutrients—nitrogen, sulfur, phosphorus, calcium, potassium, and magnesium. It also worked well for iron which, although it is not required in large amounts and hence is regarded as an essential micronutrient or trace element, is rapidly precipitated from aerated solutions of inorganic salts by formation of insoluble ferric hydroxide. As a result, these seven chemical elements were all recognized as essential mineral nutrients before 1900.

However, the delineation of other essential mineral nutrients has proved much more difficult largely because of the extremely small quantities which plants require and the difficulties of removing them from the environment. With the progressive development of better techniques for purifying water and salts, the list of essential nutrients has continually expanded. By 1960, boron, manganese, zinc, copper, molybdenum, and chlorine were added, making 13 chemical elements which are universally recognized as essential mineral nutrients for all plants. Evidence has been accumulated in support of a fourteenth element, nickel, being essential. In addition, sodium and silicon have been shown to be essential for some plants, beneficial to some, and possibly of no benefit to others. Cobalt has also been shown to be essential for the growth of legumes when relying upon atmospheric nitrogen. Claims that two other chemical elements may be essential micronutrients (vanadium and selenium), while more firmly based than claims for the essentiality of yet other mineral nutrients, have still to be firmly established.

Toxicities of nutrients and other elements. Mineral nutrients may be toxic to plants either because the specific nutrient interferes with plant metabolism or because its concentration in combination with others in solution is excessive and interferes with the plant's water relations. Other chemical elements in the environment may also be toxic.

Salt toxicity. High concentrations of salts in soil solutions or aquatic environments may depress their water potential to such an extent that plants cannot obtain sufficient water to germinate or grow. Plants can offset this effect to some extent by accumulating organic molecules or salts within their own cells. Some desert plants growing in saline soils can accumulate salt concentrations of 20–50% dry weight in their leaves without damage. But salt concentrations of only 1–2% can damage the leaves of many species. *See* PLANT-WATER RELATIONS; PLANTS OF SALINE ENVIRONMENTS.

Specific ion toxicity. A number of elements interfere directly with other aspects of plant metabolism. For

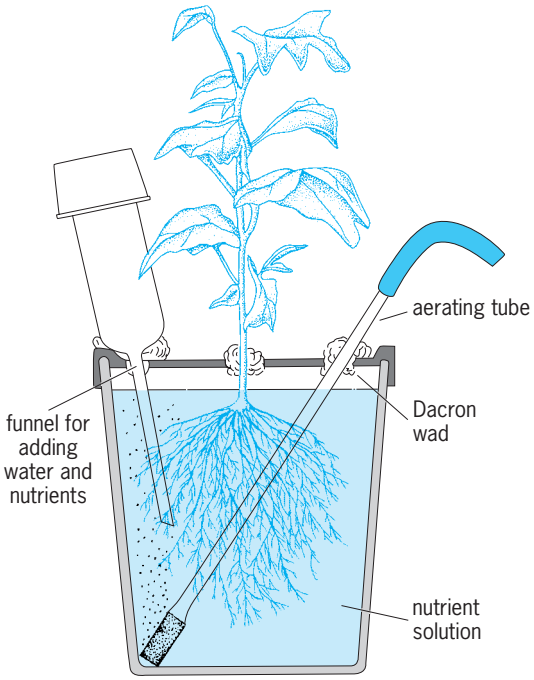

Fig. 1. Setup for growing plant in solution culture. (*After E. Epstein, Mineral Nutrition of Plants: Principles and Perspectives, John Wiley and Sons, 1972*)

example, sodium is thought to become toxic when it reaches concentrations in the cytoplasm that depress enzyme activity or damage the structure of organelles, while the toxicity of selenium is probably due to its interference in metabolism of amino acids and proteins as the result of its partial replacement of sulfur in their structure.

The ions of the heavy metals, cobalt, nickel, chromium, manganese, copper, and zinc are particularly toxic in low concentrations. The toxicities of these ions may result from their multivalent properties, which permit them to interfere with the function of proteins and other large molecules by forming cross linkages through adsorption to negatively charged sites. Plants are particularly sensitive to heavy-metal toxicities when the concentration of calcium in solution is low; increasing calcium increases the plant's tolerance. Aluminum and boron are also toxic at low concentrations. Aluminum is toxic only in acid soils for the reasons discussed below. Boron may be toxic in soils over a wide pH range, and is a serious problem for sensitive crops in regions where irrigation waters contain excessive boron, as in southern California, or where the soils contain unusually high levels of boron, as in parts of southern Australia.

Tolerance to toxic elements also varies with plant species and cultivars. In some cases, tolerance is accompanied by an unusually high accumulation of the normally toxic element in plant tops, where its toxic activity may be inhibited by excretion from the cell to the leaf surface, secretion into cell vacuoles, or combination with specific compounds synthesized by the cell.

Acid soils. All plants grow poorly on very acid soils ($pH \leqslant 3.5$); some plants may grow reasonably well on somewhat less acid soils. Several factors may be involved, and their interactions with plant species are complex. High concentrations of hydrogen ions may inhibit and distort root growth, but in soils other factors such as deficiencies of calcium, phosphorus, or molybdenum or toxicities of manganese and especially aluminum are usually more important. Many soil clays contain a high proportion of aluminum within their crystal structure, and in acid soils aluminate ions are released into solution. Aluminate ions are very toxic to root growth, stunting roots by inhibiting cell division. The harmful effects of soil acidity in some areas have been exacerbated by industrial emissions resulting in acid rain and in deposition of substances such as ammonium ions and sulfur dioxide, which increase the acidity on further reaction in the soil. The resulting acidity may release into the soil solution, streams, and lakes aluminate ions from soil clays, with consequent damage to plants and animals in these ecosystems. *See* ACID RAIN.

Nutrients in relation to animal growth. The elemental composition of plants is important to the health and productivity of animals which graze them. With the exception of boron, all elements which are essential for plant growth are also essential for herbivorous mammals. Animals also require sodium, iodine, and selenium and, in the case of ruminant herbivores, cobalt. As a result, animals may suffer deficiencies of any one of this latter group of elements when ingesting plants which are quite healthy but contain low concentrations of these elements.

Animal requirements for mineral nutrients may also differ quantitatively from plants. In addition, nutrients in forage may be rendered unavailable to animals through a variety of factors that prevent their absorption from the gut. It is thus possible for pastures to have a concentration of an element such as magnesium or copper that is adequate for their healthy growth but deficient for the animals grazing them.

Plants and animals differ also in their tolerance of high levels of nutrients, sometimes with deleterious results for grazing animals. For example, the toxicity of high concentrations of selenium in plants to animals grazing them, known as selenosis, was recognized when the puzzling and long-known "alkali disease" and "blind staggers" in grazing livestock in parts of the Great Plains of North America were shown to be symptoms of chronic and acute selenium toxicity. The soils of these areas had unusually high concentrations of selenium and supported several characteristic indigenous species with remarkably high concentrations of selenium that would be toxic to most other plants. Since that discovery, selenosis has been diagnosed in grazing animals on localized seleniferous soils in most other continents. Selenium in agricultural drainage waters has also been blamed for outbreaks of disorders in plants and animals in northern California.

Nutrients in Plant Metabolism

In order to utilize mineral nutrients for metabolism, plants must be able to absorb and distribute the nutrients within the plant body. In addition, plants require each nutrient in specific amounts for one or more unique functions. *See* PLANT METABOLISM.

Absorption of nutrients into cells. When transferred from water to mineral salt solutions, plant cells generally absorb nutrients rapidly. Two distinct phases are evident—an initial, rapid phase which reaches equilibrium within 20 or 30 min, and a slower phase which continues at a constant rate for a longer period (**Fig. 2**).

The first phase results from diffusion into spaces within and between the cell walls (the water free space) and adsorption of ions onto negative and positive charges within the cell wall and on the cytoplasmic surface (the Donnan free space). For cations, adsorption is usually very large owing to their ability to exchange with equivalent quantities of hydrogen ions which readily dissociate from the large amounts of pectic acids in the cell wall. Anions are adsorbed to a much smaller extent, there being few positively charged sites with anions available for exchange; proteins probably provide the reactive groups for anion exchange.

The second, slower phase of nutrient absorption results from the movement of nutrients across the outer membrane of the cytoplasm into the interior of the living cell. From the cytoplasm, nutrients may

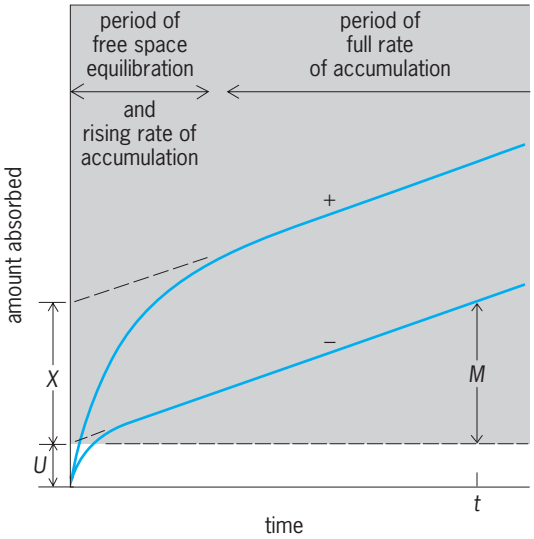

Fig. 2. Diagrammatic representation of sorption of anions and cations by a plant cell, plotted against time. *U* = amount of ion pairs in the free space; *X* = amount of cation exchange; *M* = amount absorbed within cells in time *t*. (*After G. E. Briggs, A. B. Hope, and R. N. Robertson, Electrolytes and Plant Cells, Blackwell, 1961*)

then move into organelles such as chloroplasts or mitochondria or move across the internal cellular membrane into the cell vacuole. The second phase of nutrient absorption is sensitive to temperature, anaerobiosis, and metabolic poisons, indicating that it is under metabolic control whereas the first phase is not. Metabolic absorption into the cell is clearly the significant process for the supply of nutrients for the metabolic activities of the cytoplasm.

The absorption of any one ion may be strongly inhibited or enhanced by other ions. The strongest ion antagonism is shown between nutrient ions which are chemically similar (for example, K^+ versus Rb^+, Ca^{2+} versus Sr^{2+}, Cl^- versus Br^-).

Plant cells can absorb nutrients to much higher concentrations than those of the surrounding solution, such as in the absorption by kelps of iodine to 10,000 times or more its concentration in seawater. Such absorption shows a strong selectivity for some nutrients. Accumulation of nutrients within the vacuoles of plant cells shows similar behavior. Nutrient selectivity varies greatly among plants and with environmental conditions. For example, all the ions of **Fig. 3** reach a higher concentration in the sap of *Nitella* (a fresh-water plant) than in the external solution. In *Valonia* (a seawater plant) the cell sap contains primarily potassium and chloride, whereas the seawater contains chiefly sodium and chloride.

To achieve these levels of selectivity and of concentration gradients, plant membranes must be efficient at discriminating between different nutrients and in keeping apart ions of the same nutrient. Most schemes envisage that membranes achieve these properties through specific sites within the membrane for binding anions and cations. Proteins with capacities to bind specific ions have been identified in animal, microbial, and plant cell membranes.

Differences in concentrations of mineral ions across membranes can result from either passive or active movement of the ions. Ions can accumulate in response to electrochemical potentials arising from differential permeability of the membrane or of active absorption of other ions. Ions can also accumulate from their active movement across membranes against electrochemical potentials. *See* ION TRANSPORT.

Energy supply for anion accumulation has been related to electron transfer processes, while that for cation accumulation has been related to mechanisms involving adenosine triphosphate (ATP) in

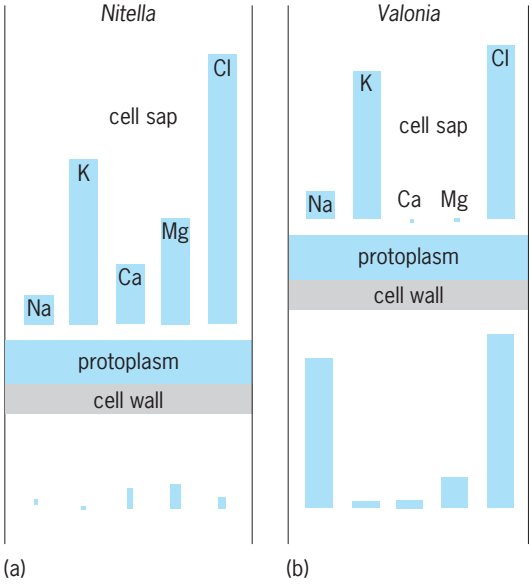

Fig. 3. Differences in absorption of ions by plants growing in (*a*) low-salt (pond-water) and (*b*) high-salt (seawater) environments. Relative concentrations of ions are expressed by the size of each bar. (*After P. J. Kramer, Plant and Soil Water Relationships, McGraw-Hill, 1969*)

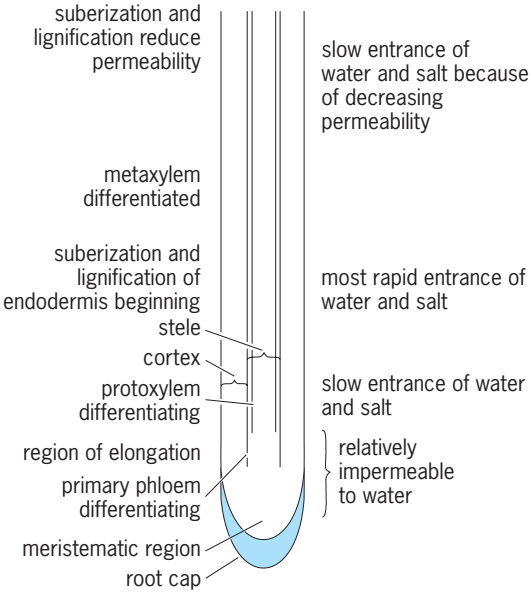

Fig. 4. Diagram of a young root, showing relation between anatomy (left) and regions that absorb (right) water and salt. (*After P. J. Kramer, Plant and Soil Water Relationships, McGraw-Hill, 1969*)

reactions with specific cation-activated ATPase enzymes within the membrane. Alternative proposals envisage that electron transport through the membrane is linked to a separation of charge which establishes a H^+ ion gradient across the membrane.

Absorption of nutrients into roots. Most of the ions which reach plant leaves are absorbed and transported to the xylem in a region of the root where xylem elements are mature and are not enclosed by impermeable outer layers (**Fig. 4**). This region contains a central stele with dead xylem cells surrounded by a continuous ring of living cells (**Fig. 5**). It is surrounded by a cortex consisting of a continuous cylinder of cells, the endodermis, and several layers of loosely arranged, thin-walled parenchyma. Outside the cortex, an epidermis forms a single compact layer of cells. Many of the epidermal cells develop long, fine root hairs which protrude for several millimeters into the medium. A thin skin of some material, possibly pectin, covers the outside of the hairs and the epidermal cells. *See* ROOT (BOTANY).

The cytoplasm of plant cells forms a continuous body (the symplasm) throughout the plant by virtue of interconnecting strands of cytoplasm which pass through pores in the walls of adjacent cells. The walls around individual cells also form a continuous fabric (the apoplast) throughout the plant. Water and nutrients can move readily through the free space of the apoplast except where, as at the endodermis, the cell walls are impregnated with waxes. The endodermis thus presents a barrier to the movement of water and nutrients, allowing passage only after absorption into the cytoplasm. To reach the xylem, nutrients are absorbed into the symplasm external to the endodermis. In the symplasm they can pass through the endodermis and into the parencyma cells surrounding the xylem. Here they are excreted into the dead xylem cells, where their concentrations may exceed those of the outside solution by more than a hundredfold.

Many of the characteristics of ion absorption by roots and excretion into root xylem are similar to those of absorption into individual cells. Thus, ion absorption by roots is metabolically controlled and shows similar patterns of selectivity and ion antagonism. As a result, many factors influence nutrient absorption by roots. Low oxygen and low temperature inhibit it. Low calcium or boron and high acidity or aluminum have marked and variable effects, probably resulting from deleterious effects of each of these environmental factors on membrane structures.

The concentration of the absorbing ion may also be very important (**Fig. 6**). However, the relationship of nutrient concentration to nutrient absorption into the root may vary with the form of the nutrient, the concentration and nature of other ions in solution (**Fig. 7**), the nutrient status of the plant (**Fig. 8**), and plant species. Moreover, the plant root interacts with and responds to its environment by excreting ions and organic substances and by generating reductive activity and enhanced capacities for absorption of nutrients (Fig. 8). The excretion of hydrogen or hydroxyl ions in response to the balance of cation and anion absorption is especially important in control

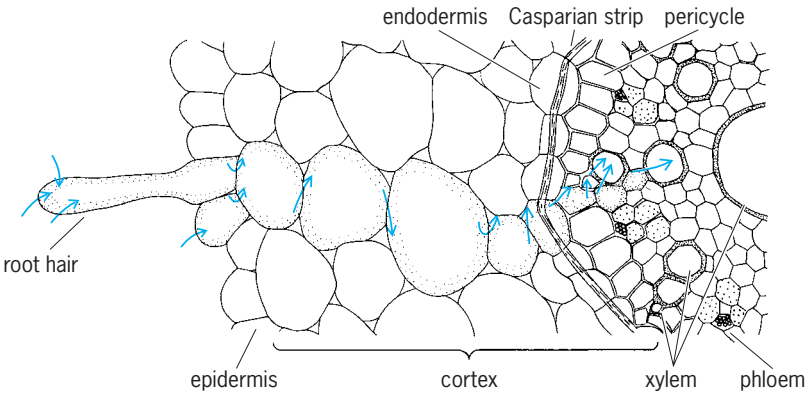

Fig. 5. Part of a wheat root transection in the region of mineral absorption and transport to xylem. (*After K. Esau, Plant Anatomy, John Wiley and Sons, 1965*)

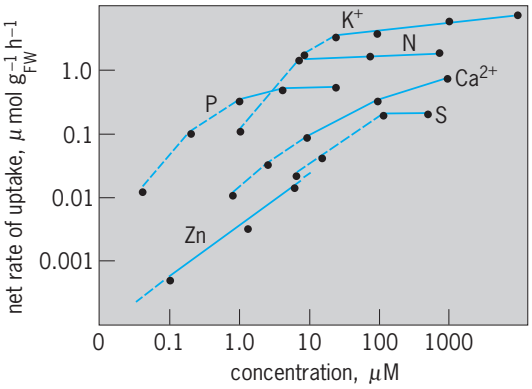

Fig. 6. Relation of the concentration of various nutrients in solution to their rates of absorption by plants. The data are plotted on a log/log basis to show the full range of concentrations used. Broken lines indicate that growth was depressed from nutrient deficiency or toxicity. (*After U. Luttge and M. G. Pitman, eds., Encyclopedia of Plant Physiology, new ser., Springer-Verlag, 1976*)

of solution pH in hydroponic systems. Reductive activity at the root surface is also important for the absorption of iron in most or all plants, while excretion of organic compounds with the capacity to form complexes with iron (siderophores) is a unique and valuable property of grass roots. *See* ABSORPTION (BIOLOGY).

Nutrient distribution. Within the plant body, nutrients move rapidly over long distances by being carried passively along with either water of the transpiration stream or food materials of the translocation stream. Nutrients may move from these streams to adjacent cells through either the cell walls or the cytoplasm. Appreciable quantities of nutrients may also be lost from leaves by leaching in rain. *See* PLANT TRANSPORT OF SOLUTES.

Xylem transport. Nutrients absorbed into the xylem sap pass to the leaves in the transpiration stream. During passage to the leaves, nutrients move rapidly into and out of cells of the stem. Transpiration rates have little effect on nutrient translocation to plant tops atlow levels of nutrient supply, but large effects at high supply when high transpiration strongly enhances transport to leaves. Within the leaves, nutrients move from the xylem sap into the leaf cells by

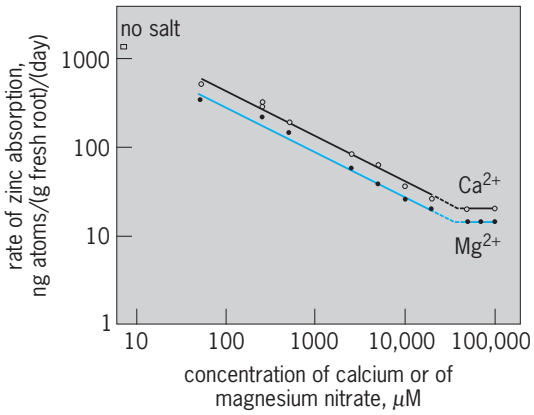

Fig. 7. Ion antagonism in wheat seedlings. As concentrations of calcium and magnesium nitrate increase, absorption of zinc decreases. The data are plotted on a log/log basis to show the full extent of the antagonism and the concentration of antagonizing ions. (*After F. M. Chaudhry and J. F. Loneragan, Zinc absorption by wheat seedlings and the nature of its inhibition by alkaline earth cations, J. Exp. Bot., 23:552, 1972*)

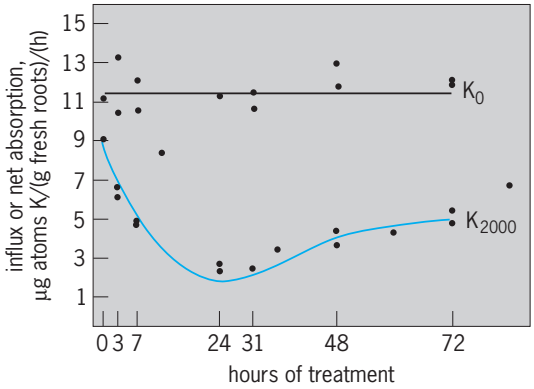

Fig. 8. Amount of potassium in plants influences the absorption of potassium. The rate of potassium absorption is constant in seedlings which receive no previous potassium (K_0). Treatment with potassium (K_{2000}) before the absorbing period quickly depressed subsequent absorption. (*After C. Johansen, D. G. Edwards, and J. F. Loneragan, Potassium fluxes during potassium absorption by intact barley plants of increasing potassium content, Plant Physiol., 45:601, 1970*)

the same mechanisms which govern their absorption by root cells. *See* XYLEM.

Phloem translocation. Phosphorus, potassium, and nitrogen move readily from leaves in the phloem to roots and other organs. By contrast, little if any calcium, boron, manganese, or iron moves from leaves to other organs. Nutrients like copper, zinc, and sulfur behave in an intermediate manner. The high mobility of nitrogen and phosphorus is associated with their incorporation into organic compounds in the phloem: in the case of phosphorus, phosphorylation of sugars is probably an important mechanism governing the movement of sugars into and out of phloem tissues. The reasons for the failure of elements to move out of leaves are not known—they could arise from precipitation as insoluble waste products outside or inside the phloem, from failure to move across cytoplasmic boundaries into the phloem, or by retention in metabolically active compounds or organelles.

When deficiencies develop, the concentrations of phloem-mobile nutrients generally drop to low values in old leaves but stay high in actively growing and metabolizing organs. By contrast, the concentrations of phloem-immobile nutrients remain high in old leaves but drop sharply in young organs. *See* PHLOEM.

Leaching from leaves. Mineral nutrients may be leached from leaves and stems in rain, fog, or dew. Leaching is favored in plants given an excessively high nutrient supply and from organs at maturity. As much as 80% of both potassium and boron and more than 50% of calcium and manganese have been reported to have leached from leaves.

Requirements for and function of nutrients. Plants are able to absorb nutrients from extremely low concentrations in solution. Figure 6 indicates the concentrations of some nutrients which are required in solution in order to sustain a reasonable growth rate. Plants do not need a continuous supply of those nutrients which are mobile in phloem; an excess supply stored in leaves or available to only part of the root system may be moved and used in any other part of the plant. By contrast, nutrients such as calcium, which are phloem-immobile, must be supplied continuously to the root system. If these nutrients become exhausted from the root environment, the root growth stops and both root and young shoots become deficient, no matter how great an excess is stored in old leaves or is available to other parts of the root system (**Fig. 9**).

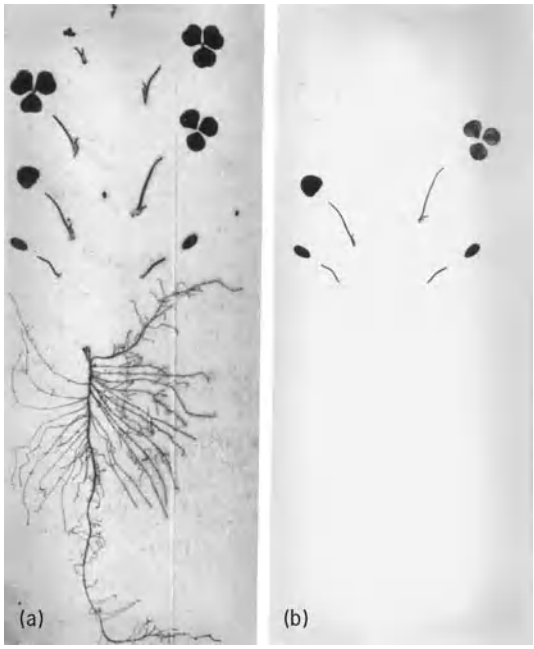

Fig. 9. Failure of nutrients to move from old leaves as shown by the behavior of radioactive calcium. (*a*) Plant grown in nutrient solution with ^{45}Ca and then transferred to solution without calcium. (*b*) Autoradiogram showing that high ^{45}Ca concentrations did not move from old leaves even though the young leaves had severe symptoms of calcium deficiency. (*After J. F. Loneragan and K. Snowball, Calcium requirements of plants, Austral. J. Agr. Res., 20:465, 1969*)

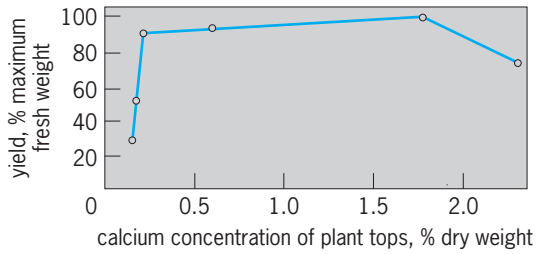

Fig. 10. Relation of yield to calcium concentration in tops of clover plants after growing in solutions with constant concentrations of calcium (open circles). When the plants were grown at a high calcium concentration and then transferred to a very low concentration, calcium deficiency developed in young organs and depressed yield of plant tops by 50% even though they contained 1.0% calcium. (*After J. F. Loneragan, Plant efficiencies in the use of B, Co, Cu, Mn, and Zn, in M. J. Wright, ed., Plant Adaptation to Mineral Stress in Problem Soils, Cornell University Press, 1976*)

The behavior of phloem-immobile nutrients creates difficulties in relating the nutrient content of plant shoots to their actual requirements or the onset of nutrient deficiency (**Fig. 10**). Hence plant composition is not always a definitive guide to the nutrient requirements or nutrient status of plant tissues. Indeed, the concentration of most nutrients can vary over a wide range without affecting growth. In addition, plants may contain appreciable concentrations of mineral elements such as aluminum which do not appear to be needed at all.

By analyzing plants grown under carefully controlled conditions, it is possible to estimate quantitative requirements for nutrients. For leaves, estimates of nutrient requirements have also been made by relating their nutrient concentrations to leaf function or growth or shoot growth. These relationships in leaves of defined physiological age, such as the youngest fully developed leaf, have been used as critical values for diagnosing nutrient deficiencies in plants. **Table 1** lists some values for green leaves. In total, leaves require only 4% of their dry weight as essential mineral nutrients, leaving the bulk of their weight as the nonmineral nutrient elements hydrogen, carbon, and oxygen. Since living leaves generally contain more than 90% water, they require for active metabolism less than 0.5% of their weight as essential mineral nutrients.

Historically, the essential mineral nutrients have been treated in two groups—the macronutrients and micronutrients, but the data of Table 1 show that the distinction is somewhat arbitrary. From the point of view of function, the essential mineral nutrients may be more appropriately grouped according to their chemical form and metabolic activities. On these bases, they may be placed into four broad groups, although the activities of some nutrients cannot be confined to a single group.

Group 1. These nutrients occur in plants as reduced atoms covalently bound into the structure of organic matter—namely, nitrogen and sulfur. Both are essential components of proteins, where they provide reactive groups which interact with other nutrients and metabolites. Nitrogen and sulfur are generally absorbed as the oxyanions nitrate and sufate which, apart from small quantities of sulfate in a few organic compounds, must be reduced before being used by the plant.

Group 2. These nutrients are covalently bound to carbon, hydrogen, and oxygen in organic structural compounds and metabolites but, unlike nitrogen and sulfur, in fully oxidized form—namely, phosphorus and possibly boron and silicon. Phosphorus is always present as mono- or polyorthophosphate; in the pH range of cells it functions as a weak acid. It occurs in phospholipids, which are basic components of plant membranes, in the nucleic acids of chromosomes, and in ATP. Phosphorus also occurs as an oxyanion in a large number of intermediary metabolites and is especially prominent in those concerned with transfer of energy.

Boron and silicon may behave in some ways analogous to phosphorus, functioning only in their fully oxidized forms. Their oxyacids are much weaker acids than the two acidic species of the phosphate ion. Consequently their oxyanions are present largely as undissociated acids. However, nothing is known of organic compounds which boron and silicon may form in plants. Boron deficiency has an immediate and marked effect on the function of cell membranes and cell extension, while silicon seems to exist mainly as opal in cells and cell walls where it acts as a structural stiffener of leaves.

Group 3. Nutrients of this group are present largely in ionic form either free in solution or reversibly bound by electrostatic forces to charged sites in the cell—namely, potassium, sodium, magnesium, calcium, manganese, and chloride. They are thought to function by osmotic adjustment of cell activities, as potassium does in the guard cells of stomata. They are also essential for some enzymic reactions where their ions are thought to act by modifying the shape

TABLE 1. Concentrations of nutrient elements in dry plant leaves at levels considered just adequate*

Element	Concentration in dry matter: μg/g or %	Concentration in dry matter: Relative number of atoms
Molybdenum	0.1 μg/g	1
Nickel	0.1	2
Copper	3	50
Zinc	20	300
Manganese	20	400
Boron	10	1,000
Iron	100	2,000
Chlorine	100	3,000
Sulfur	0.1%	30,000
Calcium	0.2	50,000
Phosphorus	0.2	60,000
Magnesium	0.2	80,000
Potassium	1.0	250,000
Nitrogen	2.5	1,300,000
Oxygen	45	30,000,000
Carbon	45	40,000,000
Hydrogen	6	60,000,000

*Modified from E. Epstein, *Mineral Nutrition of Plants: Principles and Perspectives,* John Wiley and Sons, 1972.

TABLE 2. Concentrations of ions in common soil and nutrient solutions and dominant forms of nutrients in some soil solutions

Nutrient element	Soil solution: Form*	Soil solution: Concentration	Nutrient solutions
Calcium	Ca^{2+}	1–3 mM	1–5 mM
Magnesium	Mg^{2+}	1–4	1–2
Potassium	K^{+}	0.2–2	2–10
Sodium	Na^{+}	0.5–20	0–1.4
Chlorine	Cl^{-}	0.2–30	0–10
Nitrogen	NO_3^{-}	2–10	3–15
	NH_4^{+}		1–11
Sulfur	SO_4^{2-}	1–3	1–7
Phosphorus	$H_2PO_4^{-}$, HPO_4^{2-}	0.05–5 μM	500–2000 μM
Molybdenum	$HMoO_4^{-}$, MoO_4^{2-}	0.02–0.08	0.03–0.5
Silicon	H_2SiO_4, $HSiO_4^{-}$	100–1000	0
Boron	H_3BO_3, $H_2BO_3^{-}$	3–1000	4–50
Cobalt	Co^{2+}, $Co(OH)^{+}$, Co^{2+} – OM	0.007–0.2	0–0.2
Zinc	Zn^{2+}, $Zn(OH)^{+}$, Zn^{2+} – OM	0.03–3	0.4–1
Manganese	Mn^{2+}, $Mn(OH)^{+}$, Mn^{2+} – OM	0.02–68	2–10
Copper	Cu^{2+} – OM	0.01–0.06	0.03–1
Iron	Fe^{3+} – OM	0.4–5	5–100

*OM indicates that the nutrient may be complexed with soluble organic matter: in nutrient solutions, only Fe^{3+} is added as an organic complex, all others being added as simple ions.

and orientation of enzymes and substrates; each is specifically required for the activity of one or more enzymes, and all of them can activate some enzymes in a general way. Calcium and, to a lesser extent, magnesium ions have the additional role of neutralizing and stabilizing cell structures, such as membranes and walls, probably by forming links between adjacent negative charges through cationic exchange.

In addition, magnesium has a structural role in chlorophyll, and manganese is tightly bound in chloroplasts, where it is required for the photolysis of water; in this activity, manganese behaves more like a member of group 4 than group 3.

Group 4. These nutrients are firmly bound into proteins as metalloproteins: copper, iron, molybdenum, and zinc. Metalloproteins govern a wide range of metabolic reactions, including many involving oxidation and reduction. All except zinc can act as electron carriers by undergoing reversible oxidation-reduction, such as iron in cytochrome enzymes, copper in a wide range of oxidase enzymes, and molybdenum in nitrate reductase. Zinc is also required by some oxidation-reduction enzymes but does not itself change valency. It also activates many other enzymes. Of the many physiological expressions of zinc deficiency, the best known is the disturbance of growth of the shoot apex resulting from a decreased concentration of the growth hormone indole acetic acid.

Mineral Nutrition of Plants in Soils

Although not required for growth, organic matter, solid mineral particles, and soil microorganisms strongly modify the behavior and concentration of mineral nutrients in the soil solution. In addition, root activities strongly modify soil properties at the root-soil boundary.

The concentration of mineral nutrients in the soil solution varies widely from values of less than $10^{-8}M$ for some micronutrients to $10^{-1}M$ for calcium in some soils (**Table 2**). The main forms of some nutrient ions in soil solutions also vary widely.

Nutrient reactions with soil materials. When plants are grown in water culture, the quantities of nutrients in solution are usually sufficient for 2 or 3 weeks' growth. This is generally achieved by having high concentrations of essential nutrients. Soil solutions, which have much lower concentrations of many nutrients (Table 2), can supply the plant's requirements only for very short periods of time. For plants to obtain an adequate supply, they rely on the solid phase of the soil to replenish nutrients rapidly as nutrients are withdrawn from the soil solution. The capacity of a soil to rapidly replenish a nutrient in solution is called buffering capacity, an important property governing nutrient response. Some of the many processes involved are (each is important for those nutrients in parentheses) chelation and complex formation by soil organic matter (Al, Cu, Fe, Ni, Zn); cation exchange on organic matter and clay minerals (K, Na, Ca, Mg); specific adsorption of cations on hydrous oxides of iron and aluminum (Co, Cu, Mn, Zn); ligand adsorption of anions on iron and aluminum hydroxide (P, Mo, B, Si, S); mineral decomposition and precipitation (all except N); and synthesis and decomposition of organic matter (N, S, P).

Microorganisms. Soil organisms influence plant nutrition in a variety of ways, including synthesis and degradation of soil organic matter, damage to roots, competition for nutrients, modification of soil properties, conversion of gaseous nitrogen to forms of nitrogen which plants can use, and formation of special structures in association with the root (mycorrhizal roots).

The conversion of atmospheric nitrogen into forms which plants can use for protein synthesis is especially important. A number of free-living soil bacteria can do this, the most notable being the

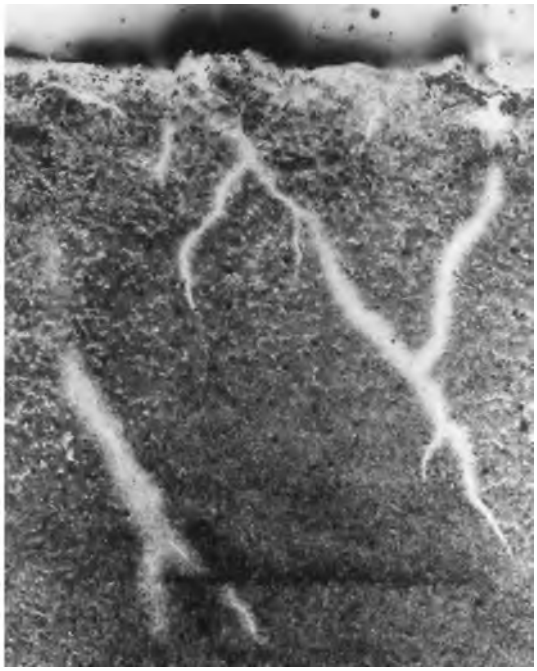

Fig. 11. Autoradiogram of wheat roots growing in soil to which radioactive calcium has been added. The white areas have had calcium removed from them, showing how the root has modified the soil in its environment. (*After H. F. Wilkinson, J. F. Loneragan, and J. P. Quirk, Calcium supply to plant roots, Science, 161:1245, 1968*)

blue-green photosynthetic cyanobacteria and certain free-living bacteria (*Azotobacter*, *Clostridium*). Other nitrogen-fixing bacteria form loose associations with plants, growing on leaf or root surfaces. The most important organisms are the *Rhizobium* bacteria, which form a close symbiotic association in special nodular structures on the roots of most leguminous species and of some other species. Plants having an effective symbiotic relationship with an appropriate *Rhizobium* can convert as much atmospheric nitrogen as they require for growth. After consumption and excretion by animals or death and decay of the plant, the fixed nitrogen is converted to soil organic matter or released to the soil solution by the activities of microorganisms in mineral forms suitable for general plant use. *See* NITROGEN CYCLE; NITROGEN FIXATION; SOIL MICROBIOLOGY.

Mycorrhizal root structures are also important, as they increase the effectiveness of roots in obtaining phosphorus and possibly other nutrients in short supply. *See* MYCORRHIZAE.

Soil–root interface. At the soil–root junction (the rhizosphere), soil properties are changed by plant activities. Such changes may be vital to nutrient absorption since young roots absorb plant nutrients from the soil solution in the rhizosphere. Young roots are coated with a mucilaginous material that extends some distance into the soil. This mucilaginous material and the many other organic materials known to exude from roots can modify the behavior of nutrients in the rhizosphere directly through complexation and cation exchange. The excretion of siderophores by grasses is especially important in making iron available from calcareous soils. Organic materials also act as food for microorganisms, encouraging growth and other activities. *See* RHIZOSPHERE.

Plants also influence the nutrient environment in the rhizosphere by absorbing both water and nutrients. Plants draw water to their roots from some distance away in the soil. Nutrients dissolved in the soil solution move passively with this mass flow of water. At the root surface they may either accumulate or be depleted by rapid absorption (**Fig. 11**). If depleted, they may be replaced by diffusion from the bulk soil. At the same time, any imbalance in absorption of anions and cations results in pH changes at the root surface that may impact on the availability of nutrients from the soil particles.

In soils, nutrients can move to roots only over limited distances so that the ability of root systems to explore the soil may become very important for the acquisition of nutrients. Thin roots with fine root hairs present a far larger surface area for interaction with soil nutrients than thick roots, and mycorrhizae extend the area of contact even further. The efficiency of root systems is also enhanced for several nutrients by their capacity to proliferate in zones of high nutrient supply. *See* PLANT GROWTH; PLANT PHYSIOLOGY; SOIL CHEMISTRY. J. F. Loneragan

Bibliography. D. T. Clarkson and J. B. Hanson, The mineral nutrition of higher plants, *Annu. Rev. Plant Physiol.*, 31:239, 1980; H. Marschner, *Mineral Nutrition in Higher Plants*, 2d ed., 1994; K. Mengel and E. A. Kirkby, *Principles of Plant Nutrition*, 4th ed., 1987; M. L. Van Beusichem, *Plant Nutrition: Physiology and Applications*, 1990; A. Wild (ed.), *Russell's Soil Conditions and Plant Growth*, 11th ed., 1988.

Plant morphogenesis

The origin and development of plant form and structure. Morphogenesis may be concerned with the whole plant or a plant part (for example, the shoot, root, leaf, or flower) or with the subcomponents of a structure (such as a tissue or surface hairs).

The inception and development of plant form is an orderly process which is controlled by internal (genetic, biochemical, and physiological) and external (environmental) factors. Techniques used in the study of plant morphogenesis include light microscopy, scanning and transmission electron microscopy, tissue and organ culture, surgical and grafting experiments, chemical (natural and synthetic) treatments and their analyses, and protein and enzyme determinations. Mutants have helped identify specific genes, and the molecular mechanisms by which they interact with other genes, in the morphogenesis of plant organs and tissues. Chimeras have proven valuable in determining the roles of specific tissue layers and the cellular interactions between different layers in the formation of a structure. Tissue cultures have provided insights into the regenerative ability and totipotency of plant cells, and the factors

that promote shoot, root, and embryo development. *See* CHIMERA; TISSUE CULTURE.

Basic processes. The development of a plant's form is primarily determined by its genetic constitution. At different stages of development, genes transcribe messages that are expressed through a complex network of processes, leading to the formation of a structure. The development of a structure follows a distinct pattern, but such patterns are not always rigid and can be modified in many cases by external stimuli. However, such stimuli are usually effective only until the particular pattern of development becomes stably fixed or developmentally committed. The following are some of the basic processes involved in plant morphogenesis.

Polarity. The establishment of differences at the two ends of a structure is called polarity. In plants, polar differences can be recognized very early in development, that is, at the zygote stage. In the zygote, cytological differences at the two ends of the cell establish the position of the first cell division, and thus the fate of structures produced from the two newly formed cells. During the development of a plant, polarity is also exhibited in the plant axis (in the shoot and root tips). If a portion of a shoot or root is excised and allowed to regenerate, the end toward the shoot tip always regenerates shoots whereas the opposite end forms roots. Polarity is also evident on the two sides of a plant organ, such as the upper and lower surface of a leaf, sepal, or petal.

Differential growth rates. The diversity in plant form is produced mainly because different parts of the plant grow at different rates. Furthermore, the growth of an individual structure is different in various dimensions. Thus the rate of cell division and cell elongation as well as the orientation of the plane of division and of the axis of cell elongation ultimately establish the form of a structure. Such differential growth rates are, however, not disorganized but rather are very well orchestrated by genetic factors. Although the absolute growth rates of various parts of a plant may be different, their relative growth rates, or the ratio of their growth rates, are always constant. This phenomenon is called allometry (or heterogony), and it supports the concept that there is an interrelationship between the growth of various organs of a plant body. Allometric studies are useful in analyzing different growth patterns in plants. *See* PLANT GROWTH.

Interdependence of plant organs. During development, either the removal of or changes in one part of the plant may drastically affect the morphogenesis of one or more other parts of the plant. This phenomenon is called correlation and is mediated primarily through chemical substances, such as nutrients and hormones. For example, at different stages of development, the various substances produced in the leaves affect the formation of flowers, the falling of leaves, the growth of lateral buds, and the differentiation of vascular tissues in the shoot. The roots produce substances which affect the growth of the shoot and lateral buds and the development of reproductive organs in flowers. Similarly, the terminal shoot apex inhibits the growth of lateral branches, presumably by either supplying or depriving lateral buds of certain growth substances. Thus, the mechanical removal of the shoot apex results in the growth of lateral buds—a practice commonly used by plant growers to produce bushy plants. *See* APICAL DOMINANCE.

Symmetric and asymmetric growth. The developmental pattern of plant structure follows a precise and definite order. In different plant species, however, the pattern may be variable. For example, the arrangement of leaves (phyllotaxy) on the shoot axis is symmetrical; it could be spiral, opposite, or whorled. In the case of spiral arrangement, the production of leaves follows a definite mathematical sequence which may vary in different plant species. Also, the pattern of leaf arrangement may vary at different stages of plant development. The development of plant structures (such as embryos) also follows a definite pattern, although it may vary in different species. *See* LEAF.

In the development of certain structures, however, asymmetric patterns occur. For example, asymmetric cell divisions are involved in the formation of stomata in grasses, and in the production of hairs of many plants. Similarly, asymmetric growth on the two sides of the shoot or a root results in the bending phenomenon (tropism) in plants. *See* PLANT MOVEMENTS.

The normal symmetric pattern of development in plants can be broken down by the invasion of pathogens (such as bacteria or nematodes), resulting in the formation of tumors, galls, and other abnormal forms. Likewise, when cut ends of stems or roots heal, or when isolated segments of plant tissues are cultured, disorganized growth often occurs, producing a mass of cells called a callus. *See* PLANT PATHOLOGY.

Regulation. The various processes involved in the morphogenesis of plants, described above, are affected by a number of internal and external factors, as described below.

Genetic factors. The ultimate factors controlling the form of a plant and its various organs are the genes. In general, several genes interact during the development of a structure, although each gene plays a significant role. Thus, a mutation in a single gene may affect the shape or size of a leaf, flower, or fruit, or the color of flower petals, or the type of hairs produced on stems and leaves. There are at least two classes of genes involved in plant morphogenesis: regulatory genes that control the activity of other genes, and effector genes that are directly involved in a developmental process. The effector genes may affect morphogenesis through a network of processes, including the synthesis and activity of proteins and enzymes, the metabolism of plant growth substances, changes in the cytoskeleton and the rates and planes of cell division, and cell enlargement.

The number of sets of chromosomes is also known to affect the size of structures. Polyploids ($4n$, $6n$, or $8n$) are generally larger in size than the diploids ($2n$),

which are usually larger than the experimentally induced haploids (1*n*). *See* POLYPLOIDY.

Biochemical factors. The development of plant form is regulated by the synthesis and activity of a large variety of proteins and enzymes that are encoded by genes. During the development of a structure, the occurrence and activity of certain proteins and enzymes, or different forms of the same enzyme (isoenzymes), are known to change at various stages. Also, stage-specific or organ-specific proteins and isoenzymes are known to occur. Mutants with differences in leaf form or in flower morphology exhibit different protein levels or enzymatic activity.

Plant growth substances. Nearly all aspects of plant form and structure are known to be affected by plant growth substances. Five classes of naturally produced plant growth substances, also called plant hormones or growth regulators, are auxins, gibberellins, cytokinins, ethylene, and abscisic acid. These hormones, along with the many synthetic compounds, form a long list of substances which affect plant growth and development.

Unlike animal hormones, each plant hormone is generally involved with a variety of developmental phenomena. For example, auxins affect the movement of plants toward light (phototropism), stimulate fruit growth in the absence of pollination (parthenocarpy), induce the production of roots on the main root and on stem cuttings, promote the differentiation of water-conducting tissue (xylem), and inhibit the growth of lateral buds. Gibberellins stimulate seed germination, stem elongation, and flowering of certain plants, promote the differentiation of male reproductive organs in flowers, and alter the leaf shape of many plants. Cytokinins play a role in bud growth, in fruit development, and in the development of male and female reproductive organs. They also delay the senescence of leaves. Ethylene causes an increase in the diameter of the shoot of young seedlings and promotes upright growth in horizontally growing plants. It also induces the formation of female flowers on male plants and enhances the ripening of fruits. Abscisic acid is generally inhibitory in nature, and is known to induce dormancy in seeds and buds, inhibit stem growth, enhance leaf abscission, and regulate stomatal opening in leaves. In some plants, it also affects leaf form. A number of synthetic growth regulators are used commercially for inducing rooting of stems and leaves, for producing short sturdier plants, and for enhancing fruit ripening. Inhibitors generally tend to retard the growth of shoots and are used by horticulturists in producing short and sturdier plants. *See* ABSCISIC ACID; AUXIN; CYTOKININS; ETHYLENE; GIBBERELLIN.

Each class of plant growth substances is produced in more than one region of the plant and transported to different regions. Some of the hormones are transported primarily in one direction (for example, auxins travel in the basal direction), but others (gibberellins and cytokinins) travel in all directions. Also, during the formation of a structure, the endogenous levels of growth substances change; some are higher at earlier stages, whereas others are more abundant at later stages. This has led to the concept that the formation of a structure involves a critical balance of a number of endogenous growth substances at various stages of development and that a change in this balance generally results in the modification of the form. *See* PLANT HORMONES.

Plant form is also known to be affected by nutritional factors, such as sugars or nitrogen levels. For example, leaf shape can be affected by different concentrations of sucrose, and the sexuality of flowers is related to the nitrogen levels in the soil in some species. Inorganic ions (such as silver and cobalt) have also been known to affect the type of flower produced. *See* PLANT MINERAL NUTRITION.

Environmental factors. Although genes are the ultimate controlling factors, they do not act alone, but interact with the existing environmental factors during plant development. Environmental factors, including light, temperature, moisture, and pressure, affect plant form.

Light has a dramatic effect on plant form. For example, a seedling grown in the dark shows considerable elongation of the shoot with little or no growth of leaves, in contrast to a light-grown seedling in which leaves expand and shoot growth is stunted. The intensity of light also strongly influences the development of a structure; under high light intensity, leaves are usually thicker and more pubescent than under low light. Similarly, the quality of light also affects the morphogenesis of plants. For example, under red light fern gametophytes grow as elongated filamentous structures, but under blue light normal flattened gametophytes are produced. Many of the developmental phenomena, such as seed germination and flowering of plants, are affected by red light. Red light effects are mediated through the pigment phytochrome. *See* PHYTOCHROME.

Temperature has a strong effect on a number of plant developmental processes, such as the growth of shoots and buds, the flowering of plants, and the type of flower produced. Some winter crops (for example, wheat and rye) require a chilling factor (vernalization) for flowering to occur in the next spring. Also, low temperatures tend to favor the development of female reproductive over male organs in some plants. *See* VERNALIZATION.

The availability of water also drastically affects plant form. Plants grown in dry-arid conditions produce small thick leaves which are often rolled on the upper surface, in contrast to plants grown underwater, which produce thin and highly dissected leaves. Furthermore, in some cases, if land plants with entire leaves are submerged in water, they too develop dissected leaves. *See* PLANT-WATER RELATIONS.

Many of the environmental effects described above are believed to be mediated through a change in the balance of endogenous plant hormones. *See* PHYSIOLOGICAL ECOLOGY (PLANT); PLANT PHYSIOLOGY.

V. K. Sawhney

Bibliography. D. E. Fosket, *Plant Growth and Development: A Molecular Approach*, 1994; R. F. Lyndon, *Plant Development: The Cellular Basis*, 1990;

T. A. Steeves and I. M. Sussex, *Patterns in Plant Development*, 2d ed., 1989.

Plant movements

The wide range of movements that allow whole plants or plant parts to transport or reorientate themselves in relation to changed surroundings in order to execute specific functions. For example, small free-floating aquatic plants migrate to regions optimal for their activities, and anchored plants move to facilitate spore or seed dispersal. There are two types of plant movement: abiogenic movements, which arise purely from the physical properties of the cells and therefore take place in nonliving tissues or organs; and biogenic movements, which occur in living cells or organs and require an energy input from metabolism.

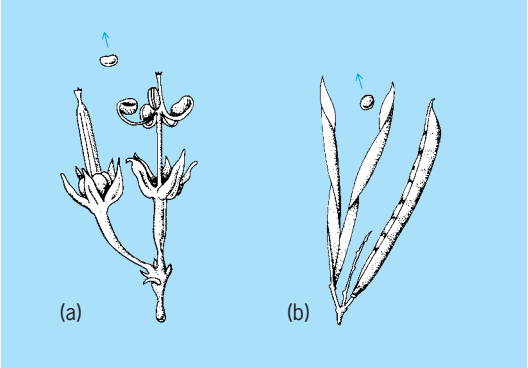

Fig. 1. Hygroscopic movements. (*a*) Fruits of *Geranium dissectum*: the "springs" of the fruit on the right have been released, and one seed is being slung out of the lobe of the capsule. (*b*) Seed pods of *Lathyrus vernus*: the two halves of the pod on the left have split apart, curled up explosively, and ejected one seed. (*After H. Straka, Encyclopedia of Plant Physiology, vol. 17, pt. 2, Springer-Verlag, 1962*)

Abiogenic Movements

Drying or moistening of certain structures causes differential contractions or expansions on the two sides of cells and hence causes movements of curvature. This differential behavior occurs because the composition and physical properties of cell walls vary from place to place in the organ and according to direction in the walls. Such movements are called hygroscopic and are usually associated with seed and spore liberation and dispersal. Examples of such movement occur in the "parachute" hairs of the fruit of dandelion (*Taraxacum officinale*), which are closed when damp but open when the air is dry to induce release from the heads and give buoyancy for wind dispersal. The elaborate teeth (peristome) on the rim of the capsule (spore case) of mosses respond to humidity changes, closing and opening the capsule mouth and allowing the periodic release of the spores. In some grasses (such as *Stipa pennata*), the fruit has a long thin spine (awn) at one end. This is straight when moist, but as it dries the half nearest the fruit twists into a tight coil, while the apical half stays straight and perpendicular to the coil. This acts as a lever when caught by a nearby object, allowing humidity changes to cause the hygroscopic coil to screw the fruit into the ground. In Queen Anne's lace (*Daucus carota*), under moist conditions the whole ripe umbrella-shaped umbel that makes up the infructescence contracts into a kind of ball by the curling-in of the individual rays. Under dry conditions the rays straighten, promoting the release of the small fruitlets. An extreme inverse form of this behavior is found in the rose of Jericho (*Anastatica hierochuntica*), where the whole plant forms a ball when it dries out and then spreads open again in rain to allow raindrops to disperse the seed.

Hygroscopic movements are relatively slow, but there are some which, after tensions increase due to water loss, take place suddenly by explosive release. An example is the "slingshot" fruit dispersal mechanisms of *Geranium* spp., where the five seed-containing parts in each flower are fixed to the base of a long five-partite beak, each part of the beak constituting a spring. When the spring has dried out sufficiently, it detaches the seed container at its base, coils up violently, and projects the seed from its now open container for a considerable distance (**Fig. 1***a*). Similar mechanisms are seen in members of the pea family, where the familiar pods, when dry, suddenly split apart into two valves which coil up, shooting out the seeds (Fig. 1*b*).

Another type of abiogenic movement is due to changes in volume of dead water-containing cells. In the absence of a gas phase, water will adhere to lignocellulose cell walls. As water is lost by evaporation from the surface of these cells, considerable tensions can build up inside, causing them to decrease in volume while remaining full of water. The effect is most commonly seen in some grasses of dry habitats, such as sand dunes, where longitudinal rows of cells on one side of the leaf act as spring hinges, contracting in a dry atmosphere and causing the leaf to roll up into a tight cylinder, thus minimizing water loss by transpiration. The most dramatic example is seen in sporangia of many ferns, which are encircled by a ring of special cells, the annulus. These cells are constructed with thick inner and thin outer walls. When water is lost and internal tensions increase, a tangential pull is exerted at a weak point in the ring (the stomium) which consequently ruptures, and the annulus, curling outward, opens up the sporangium, exposing the spores. This opening becomes larger as water loss continues until internal tensions in the annulus cells become so great that the water collapses from the walls, leaving the cells largely filled with water vapor. The outwardly directed tension is thus suddenly released, and the annulus springs back toward its original position, throwing the spores violently out into the air.

Yet other movements depend on changes in the turgidity of cells. Though not strictly abiogenic, since they occur in living cells and thus rely, at least to a small extent, on energy from metabolism, they are included here for convenience. The most important of such movements are those of guard cells which

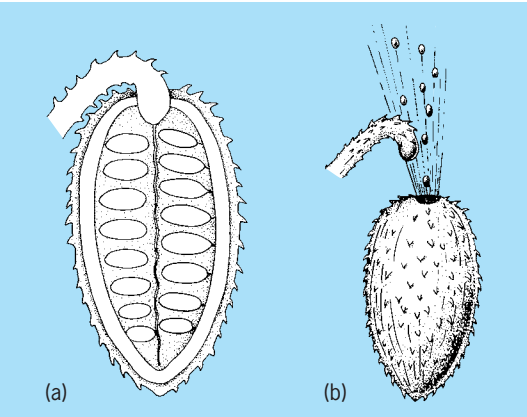

Fig. 2. Fruit of the squirting cucumber (*Ecballium elaterium*). (*a*) Section showing seeds embedded in a pulpy matrix inside a spiny outer covering. (*b*) Fruit in process of squirting as a result of hydrostatic pressure built up in the matrix. (*After H. Straka, Encyclopedia of Plant Physiology, vol. 17, pt. 2, Springer-Verlag, 1962*)

surround the stomatal pores in the epidermis of leaves and other green organs. Stomata allow the exchange of gases, such as carbon dioxide and oxygen in photosynthesis and respiration and the escape of water vapor in transpiration. Guard cells have differentially thickened walls. When the turgidity of these cells rises due to water absorption, their shape alters, causing the pore to widen; when turgidity falls, the cells collapse, closing the pore.

Other movements controlled in a similar way are the opening of grass flowers, which have special hinge cells at the base of the covering scales (glumes). A dramatic example where turgidity is used for seed dispersal is the squirting cucumber (*Ecballium elaterium*). Here increasing turgor, which can rise to 3–4 bars (300–400 kilopascals) inside the fruit as it becomes ripe, blows out the stalk like a champagne cork and with it the liquid seed containing pulp (**Fig. 2**).

A seed-dispersal mechanism, resembling in its action that of the fruit of *Geranium* (Fig. 1*a*), depends on tensions set up not by water loss from the seed-containing parts but by the differential water uptake by separate cell layers causing differential swelling of the affected tissue. An example of this type is found in touch-me-not (*Impatiens noli-tangere*). When the fruit is ripe and fully turgid, the slightest mechanical disturbance causes the detachment of the seed-containing valves at the base; explosive coiling of these valves then ejects the seed up to a distance of a meter (3.3 ft) or so.

Biogenic Movements

There are two types of biogenic movement. One of these is locomotion of the whole organism and is thus confined to small, simply organized units in an aqueous environment. The other involves the change in shape and orientation of whole organs of complex plants, usually in response to specific stimuli. These movements arise either from differential growth on opposite sides of the organ concerned or from local reversible changes in cell volume caused by the loss or gain of water.

Locomotion. In most live plant cells the cytoplasm can move by a streaming process known as cyclosis (**Fig. 3**). Energy for cyclosis is derived from the respiratory metabolism of the cell. The mechanism probably involves contractile proteins very similar to the actomyosin of animal muscles. How these contractile proteins act between the rigid outer cell wall and the fluid cytoplasm to produce this streaming is not yet known. In the slime molds the whole plant body moves across substrata by a similar mechanism. The separating movement of chromosomes during nuclear division, a phenomenon common to plants and animals, is also mediated by fibers constituted of similar contractile proteins.

Cell locomotion is a characteristic of many simple plants and of the gametes of more highly organized ones. Motility in such cells is produced by cilia, whiplike protein structures 10–25 micrometers long and about 0.2 μm thick, anchored in the peripheral layers of the cell and projecting into the surrounding medium. There is a complex internal structure of parallel threads (axonemes) running the length of the cilium. There are two axial axonemes and nine in an outer ring, and it is thought that the rhythmic and sequential contraction of the members of this outer ring gives the cilium either a whiplike or a screwlike motion which is responsible for propelling the organism through its aqueous environment. *See* CILIA AND FLAGELLA.

Cell locomotion is usually not random but is directed by some environmental gradient. Thus locomotion may be in response to specific chemicals; this is called chemotaxis. Light gradients induce phototaxis; temperature gradients induce thermotaxis; and gravity induces geotaxis. One or more of these environmental factors may operate to control movement to optimal living conditions. *See* CELL MOTILITY.

Movement of organs. In higher plants, organs may change shape and position in relation to the plant body. When bending or twisting of the organ is evoked spontaneously by some internal stimulus, it is termed autonomous movement. The most common

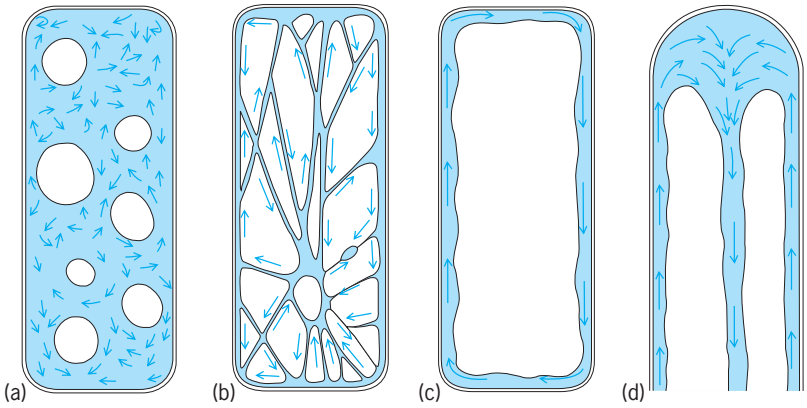

Fig. 3. Common types of protoplasmic streaming (cyclosis) in plant cells: (*a*) agitation; (*b*) circulation; (*c*) rotation; and (*d*) fountain streaming. (*After N. Kamiya, Encyclopedia of Plant Physiology, vol. 17, pt. 2, Springer-Verlag, 1962*)

movements, however, are those initiated by external stimuli such as light and the force of gravity. Of these there are two kinds. In nastic movements (nasties), the stimulus usually has no directional qualities (such as a change in temperature), and the movement is therefore not related to the direction from which the stimulus comes. In tropisms, the stimulus has a direction (for instance, gravitational pull), and the plant movement direction is related to it.

Autonomous movements. The most common movement is circumnutation, a slow, circular, sometimes waving movement of the tips of shoots, roots, and tendrils as they grow; one complete cycle usually takes from 1 to 3 h. In horizontal underground stems (rhizomes) there may be an up-and-down hunting movement that maintains the organ at a constant distance from the soil surface. Tendrils may show very extensive swings as they "search" for support. In twining plants this circumnutation is an essential part of the climbing mechanism, where the slenderness of the stem demands rigid external support. The direction of twining is characteristic of the species. It is usually anticlockwise as viewed from above, but a few species twine clockwise, for example, the hop (*Humulus lupulus*) and honeysuckle (*Lonicera periclymenum*).

These movements are due to differential growth, but some may be caused by turgor changes in the cells of special hinge organs and are thus reversible. Such is the case with the telegraph plant (*Desmodium gyrans*), whose leaflets show regular spontaneous gyrations, with cycle times as short as 30 s under warm conditions.

Nastic movements. There are two kinds of nastic movements, due either to differential growth or to differential changes in the turgidity of cells. They can be triggered by a wide variety of external stimuli.

Photonastic (light-dark trigger) movements are characteristic of many flowers and inflorescences, which usually open in the light and close in the dark. Night flowers, such as white campion (*Melandrium album*), close by day and open by night. Many photonastic movements are complicated by a superposed autonomous movement with a 24-h (circadian) rhythm. These rhythms are started by a change in light conditions, but may continue for days under conditions of constant illumination or darkness. A similar phenomenon is the "sleep movements" (nyctinasty) in the pinnate leaves of some plants (such as *Albizzia*), where the leaflets fold together at night. The latter movement is the result of changes in the turgor in special organs (pulvini) at the bases of the leaflets. The turgor changes are produced by the extrusion of potassium ions from cells on one side of the pulvinus, leading to a rise in water potential and associated loss of water from those cells. A special blue pigment, phytochrome, is the receptor for these light responses.

Thermonasty (temperature-change trigger) is seen in the tulip and crocus flowers, which open in a warm room and close again when cooled. Crocus is particularly sensitive, responding to a temperature change as small as 0.2–0.5°C (0.35–1°F). The movement is caused by differential growth responses at the bases of the tepals (floral leaves).

The most striking nastic movements are seen in the sensitive plant (*Mimosa pudica*). Its multipinnate leaves are very sensitive to touch or slight injury. Leaflets fold together, pinnae collapse downward, and the whole leaf sinks to hang limply. Only one leaflet needs to be touched or damaged for the stimulus to pass throughout the leaf, or even the whole plant, in a rapid sequence of collapse at a rate of 4–30 mm/s. The conducted stimulus is complex, having a rapid electrical action-potential component and a slower hormone-transport component. The turgor mechanism operates on the same principles as those involving pulvini in nyctinasty detailed above. The whole process is reversible, and recovery takes place in 15–20 min. The phenomenon is known as seismonasty.

Epinasty and hyponasty occur in leaves as downward and upward curvatures, respectively. They arise either spontaneously or as the result of an external stimulus, such as exposure to the gas ethylene in the case of epinasty; they are not induced by gravity.

Tropisms. Of these the most universal and important are geotropism (or more properly gravitropism) and phototropism; others include thigmotropism and chemotropism.

1. G*ravitropism*. In geotropism, the stimulus is gravity. The main axes of most plants grow in the direction of the plumb line with shoots upward (negative gravitropism) and roots downward (positive gravitropism). Lateral organs such as branches, leaves, and underground or surface stems grow normally at a specific angle to the plumb line (plagiogravitropism; **Fig. 4**). Angular displacement of such organs from these orientations induces curvatures to reestablish the normal position. In most organs curvatures are due to differential changes in growth rate in those parts of the organ which are still growing. In grass stems, however, where the curvatures are confined to the nodes, which do not grow in the vertical position, displacement by wind or heavy rain induces growth on the lower side of the node, making a "knee" which erects the stem to its normal upright position (**Fig. 5**).

One part of a plant may influence the gravitropic behavior of another. For example, the branches of many conifers are plagiogravitropic, but if the top of the leading shoot is removed, then, after a lag, neighboring branches will become negatively gravitropic and curve upward, indicating that the main shoot has in some way induced plagiogravitropic behavior on the laterals.

There have been many theories about the mechanism by which a plant senses the force of gravity, but the oldest and still most acceptable identifies the sensors as heavy starch-containing bodies known as amyloplasts, which can be observed to sediment to the lower internal surface of containing cells. These amyloplasts, called statoliths, occur, with one or two doubtful exceptions, in all higher plant organs which respond to gravity, even in species which do not otherwise manufacture storage starch (**Fig. 6**).

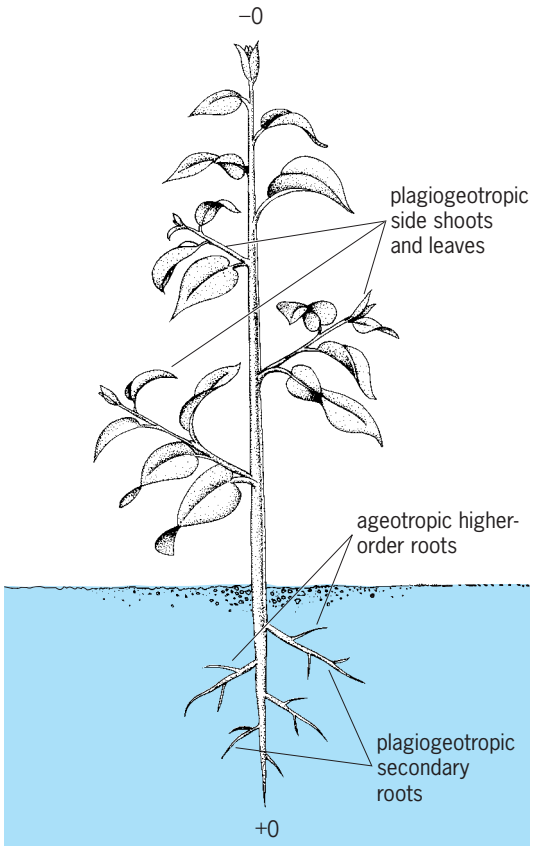

Fig. 4. Geotropic responses of plant organs: −0 = negatively geotropic main shoot; +0 = positively geotropic main root. (*After L. J. Andus, Physiology of Plant Growth and Development, ch. 6, Geotropism, McGraw-Hill, 1969*)

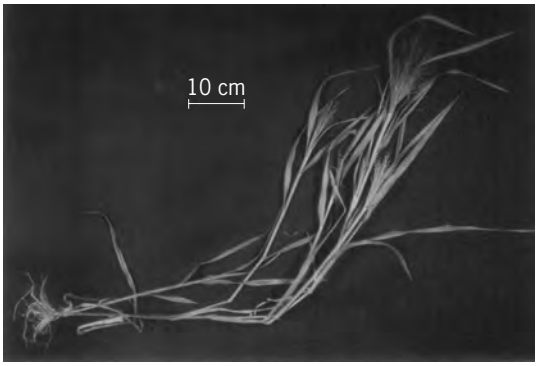

Fig. 5. Knee-forming ability of a witchgrass (*Panicum capillare*) shoot. (*Courtesy of P. B. Kaufman*)

Gravitropic curvatures are caused by differing concentrations of hormone on the two sides of the growth zone. In shoots this hormone appears to be the auxin indol-3-ylacetic acid, which is synthesized in the apex of growing shoots and in young expanding leaves; it is then transported to the zone of cell extension, where it controls elongation growth. A higher concentration on the lower side of the growth zone will cause faster growth there and consequently an upward curvature of the shoot. In most shoots this difference in auxin concentration seems to be due either to a downward migration from upper to lower side of the growth zone or to a more active transport from the tip down the lower side. The "knee-forming" nodes of grasses and cereals seem to have a different mechanism. Here the displacement of the stem into a slanting or horizontal position stimulates either the production of auxin or its release from inactive conjugates in the tissues on the underside of the node and, in consequence, its greatly accelerated growth. The sedimented starch statoliths in this lowermost part of the node are presumably in some way involved in this change of auxin status. In roots, where curvature is also due to the induction of differential growth rates, the nature of the hormone responsible is still in doubt since here the lower side grows more slowly than the upper. It is thought that the growth inhibitor abscisic acid or a close chemical relative is involved, although indol-3-ylacetic acid may also participate in some way.

2. *Phototropism.* In phototropism the stimulus is a light gradient, and unilateral light induces similar curvatures; those toward the source are positively phototropic; those away from the source are negatively phototropic. Main axes of shoots are usually positively phototropic, while the vast majority of roots are insensitive. Leaves usually respond by aligning their surfaces perpendicular to the incident rays and are said to be diaphototropic. Some plants change their responses during development, as can also happen in gravitropism. For example, ivy-leaved toadflax (*Linaria cymballaria*) grows on walls, and the flower stalks turn toward the light to facilitate pollination; when the fruit ripens, they turn toward the wall to bury the fruit in a crevice, ready for germination.

The first shoots (coleoptiles) of grasses and cereals have been used extensively for phototropic studies because of their great sensitivity and simple structure (**Fig. 7**). As for geotropism, differences in auxin concentration between the light and dark sides of the coleoptile seem to be responsible for the curvature; less auxin accumulates on the lighter side. However, in other plants, for example, sunflower (*Helianthus annus*) seedling hypocotyls, auxin seems not to be directly involved; instead more growth in-

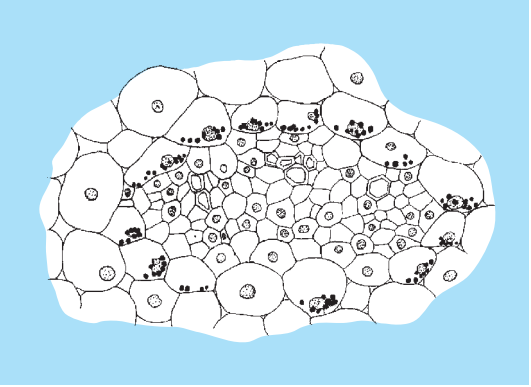

Fig. 6. Transverse section of a vascular strand in the first shoot of an onion (*Allium cepa*) seedling showing sheath of cells containing amyloplasts (small black granules) which have sedimented to the lower side of the cells. (*After H. E. Street and H. Öpik, The Physiology of Flowering Plants, Edward Arnold, 1970*)

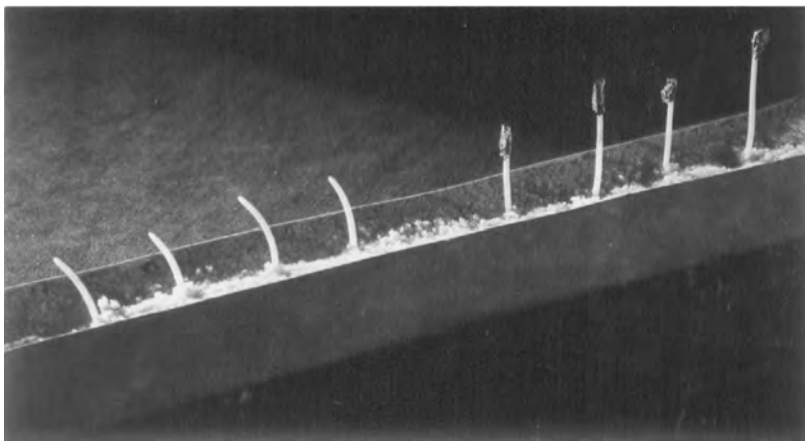

Fig. 7. Phototropism of oat (*Avena sativa*) coleoptile. The four coleoptiles on the left show the growth curvatures induced by light coming from the left. The four on the right have had their tips covered with small caps of aluminum foil. (*After L. J. Andus, Plant Growth Substances, 2d ed., Leonard Hill Ltd., 1959*)

hibitor (possibly abscisic acid) accumulates on the lighter side. The shorter wavelengths of the visible spectrum (that is, blue light) are the effective ones, implying that yellow pigments are concerned with detection of a light gradient across the organ. Whether these pigments are flavins or carotenoids or both remains to be established. The nature of light action in causing hormone concentration differences is also uncertain, but it does not seem to be photodestruction of auxin. A modification of hormone transport down the two sides may be involved.

Many fungi also show gravitropic and phototropic responses. However, since growth processes of these organisms differ greatly from those of higher plants, the mechanisms must be very different.

3. *Thigmotropism.* In thigmotropism (sometimes called haptotropism), the stimulus is touch; it occurs in climbing organs and is responsible for tendrils curling around a support. The curvature is due to a deceleration of growth on the contact side and an acceleration on the other. In many tendrils the response may spread from the contact area, causing the tight coiling of the basal part of the tendril into an elaborate and elastic spring.

4. *Chemotropism.* Chemotropism is induced by a chemical substance. Examples are the incurling of the stalked digestive glands of the insectivorous plant *Drosera* and incurling of the whole leaf of *Pinguicula* in response to the nitrogenous compounds in the insect prey. A special case of chemotropism concerns response to moisture gradients; for example, under artificial conditions in air, the primary roots of some plants will curve toward and grow along a moist surface. This is called hydrotropism and may be of importance under natural soil conditions in directing roots toward water sources. *See* ABSCISIC ACID; AUXIN; PLANT GROWTH; PLANT HORMONES; PLANT PHYSIOLOGY; ROOT (BOTANY). Leslie J. Audus

Bibliography. W. Haupt and M. E. Feinleib (eds.), *Encyclopedia of Plant Physiology*, new ser., vol. 7: *Physiology of Movements*, 1979; M. B. Wilkins, *Advanced Plant Physiology*, 1988.

Plant nomenclature

The generally accepted system of naming plants. The study of plant variation and kinds leads to classification—the grouping of individual plants, the grouping of the groups, and the organizing of the groups (taxa) into ranks. Nomenclature deals with the naming of the taxa, the goal being to establish one correct scientific name for a given taxon with a given definition.

There are two fundamentally different kinds of plant names, vernacular (common) and scientific. Vernacular names are culturally limited and differ from one society to another. For example, the beech tree is also known as *Buche* (German), *hêtre* (French), *faggio* (Italian), and *buna zoku* (Japanese). The scientific name of the beech, *Fagus*, is culturally neutral and universally recognized by botanists.

For most purposes, plants are considered in terms of three ranks: family, genus, and species. There are ranks above family (such as orders), ranks lower than species (such as subspecies), and even intercalated ranks (such as subgenera). Scientific family names are based on generic names and end in -aceae, such as Fagaceae (the oak family), which includes two major genera of trees, *Fagus* and *Quercus*.

The basic scientific plant name is a binomial, a combination of a Latin capitalized generic name and a Latin lowercase specific epithet, such as *Quercus alba*, the common American white oak. Scientific names may seem awkward, but to the extent that they are universally learned, they are also universally understood. The names also have meaning. *Quercus* is an ancient Roman word for the trees that provided the best and hardest woods, and *alba* means white, presumably a reference to the whitish undersides of the leaves.

History. The binomial system of nomenclature was first widely applied to plants by Carolus Linnaeus in his *Species Plantarum* (1753). Within a generation, binomial names displaced the previously used descriptive polynomials. For example, J. F. Gronovius (1739) called the white oak *Quercus foliis superne latioribus opposite sinuatis, sinubus angulisque obtusis* (oak with leaves broader upward, oppositely sinuate, [and] blunt sinuses and angles). Such polynomials had become increasingly cumbersome as more and more plant species were discovered in the New World during the eighteenth century.

The century following Linnaeus was an authoritarian period, in which plant nomenclature was dominated by botanical scholars such as J. B. P. de Lamarck, K. Willdenow, and their successors. As the numbers of taxa and workers increased, it became obvious that some rules must be established to prevent unnecessary name changes. This led to A. de Candolle's "Laws," ratified in 1867 and considered to be the best guide for botanical nomenclature.

Shortly thereafter, it became evident that the "Laws" were not consistent with past usage, and either the names or the rules had to change. Several individuals and institutions had devised their own

rules, which led to chaos. Eventually the first *International Rules of Botanical Nomenclature* was ratified by the botanical community in 1906. This introduced the concept of conservation of generic names in order to maintain well-known names that did not follow its guidelines. The third edition of the Rules (1935) codified the type method for determining the application of scientific names. The fourth edition was the first to be called the *International Code of Botanical Nomenclature* (ICBN), and has been followed by a new edition about every 6 years. Each new edition is produced after an International Botanical Congress, during each of which several hundred proposals to amend the ICBN are considered and voted upon by delegates to the Congress.

Name changes. There are three major reasons that scientific names should be changed: discovery that the type specimen of a name is not what had been assumed, discovery that a hitherto-overlooked name applies, or redefinition of the taxon.

Types control the application of names. Botanists describe their new taxa with at least one specimen in hand. Although the taxon can be redefined (usually as new material becomes available), it can never be redefined so as to exclude all the original (type) material. On occasion, however, the type material of a well-known name has been found to pertain to a completely different species. The ICBN has a formal procedure to maintain usage.

Priority controls when different names apply to the same taxon. Essentially the oldest valid name published after 1753 has rights that newer names do not. Relatively few names are changed because of overlooked earlier names.

Redefinition of taxa causes most name changing. Practically every new collection prompts a redefinition of a taxon. The new collection could establish a link between what had been thought to be separate species and result in two species becoming one. Conversely, careful study might show the opposite—what was thought to be a single, variable species is actually a mixture of two or more species. Taxonomy dictates which taxa are to be recognized and how they are to be defined, and nomenclature provides the names for the taxa. When the scope of the taxon is changed because of new knowledge, the name usually must be changed.

Refinements. Plant hybridization and horticulture lead to special problems that fit uneasily into the Linnaean binomial system. Plants more often hybridize than do animals. The ICBN devotes an appendix to hybrid nomenclature. Hybrid species, called nothospecies, can be named by one of two methods, a hybrid formula (giving the parents) or a binomial with a multiplication sign. *Verbascum lychnitis* L. × *V. nigrum* L. is a hybrid formula. (The "L." indicates that the plants were named by Linnaeus.) It is also correctly known as *Verbascum* × *schiedeanum* Koch. The appendix provides for naming more complex hybrids, but the multiplication sign is a constant feature.

The *International Code of Nomenclature for Cultivated Plants* deals with cultivated plants, generally known as cultivars. Cultivars are usually given fancy names or epithets that can be distinguished from scientific names in that they are capitalized and are not italicized. In addition, the names or epithets are either preceded by cv., the abbreviation of cultivar, or indicated by quotation marks: *Scilla hispanica* cv. Rose Queen and *Scilla hispanica* 'Rose Queen' are both correct for the "Rose Queen" cultivar of *Scilla hispanica*. *See* PLANT KINGDOM; ZOOLOGICAL NOMENCLATURE.

Dan H. Nicolson

Bibliography. C. D. Brickell et al. (eds.), International Code of Nomenclature for Cultivated Plants—1980, *Regnum Vegetabile*, 104:1–32, 1980; W. Greuter et al. (eds.), International Code of Botanical Nomenclature, *Regnum Vegetabile*, 118:1–328, 1988; E. E. Sherff et al., Symposium on Botanical Nomenclature, *Amer. J. Bot.*, 36:1–32, 1949; M. L. Sprague et al., A discussion on the differences in observance between zoological and botanical nomenclature, *Proc. Linn. Soc. Lond.*, 156:126–146, 1944.

Plant organs

Plant parts having rather distinct form, structure, and function. Organs, however, are interrelated through both evolution and development and are similar in many ways.

Roots, stems, and leaves are vegetative, or asexual, plant organs. They do not produce sex cells or play a direct role in sexual reproduction. In many species, nevertheless, these organs or parts of them (cuttings) may produce new plants asexually (vegetative reproduction). Sex organs are formed during the reproductive stage of plant development. In flowering plants, sex cells are produced in certain floral organs. The flower as a whole is sometimes called an organ, although it is more appropriate to consider it an assemblage of organs. *See* PLANT REPRODUCTION.

Root. The root is usually the underground part of the plant axis. It may consist of a dominant primary seedling root (taproot) with subordinate branch roots, as in carrots and beets; or it may be composed, as in grasses, of numerous branched roots of similar dimensions (fibrous roots). Collectively, all the roots of a plant are known as its root system. Roots anchor the plant, absorb water and mineral salts in solution from the soil, and conduct these to the stem. Organic food and growth substances received from the stem move to the areas of growth and storage in the roots. *See* ROOT (BOTANY).

Stem and leaves. The stem is usually the aerial part of the plant axis and bears leaves. The stem and leaves together constitute the shoot. In some species the major portion of the stem grows horizontally beneath the surface of the soil, and thus is called a rhizome, or underground stem. The stem conducts water and minerals from the roots to all parts of the shoot, and food materials and growth substances from the shoot to the root. The stem may also serve as a storage organ for water and food. Green leaves containing chlorophyll, when exposed to light and air, carry on photosynthesis. As a

by-product of this process, oxygen is returned to the atmosphere. Leaves also return large amounts of water vapor to the air through transpiration (evaporation). Some leaflike structures are protective (bud scales), others are fleshy types in which food and water may accumulate. The first leaves on a seed plant are called cotyledons. *See* LEAF; STEM.

Flower, fruit, and seed. The flower is often interpreted as a modified shoot bearing floral organs instead of leaves. These organs are the sepals, petals, stamens, and carpels. The sepals and petals are sterile leaflike appendages. Sepals, like leaves, are commonly green. Petals, containing little or no chlorophyll, are usually white or have some color other than green. The sepals collectively constitute the calyx, the petals the corolla. The calyx and corolla form the perianth or floral envelope. The stamens and carpels are the reproductive floral organs and produce sex cells. A stamen is usually composed of a slender stalk, called the filament, at the tip of which is an anther. The anther is divided into four or fewer pollen sacs in which pollen grains develop. When mature, the pollen sacs open, and the pollen is transferred by wind, water, insects, or man to the tips of the carpels, a process called pollination. The assemblage of stamens is called the androecium, a term implying the male nature of this part of the flower. The carpel assemblage is called the gynoecium to indicate its female nature. Pistil is another term used to designate the female part of the flower. A single carpel may form a pistil, or two or more may be combined into a compound pistil. The pistil generally consists of an enlarged basal part called the ovary. The apex of the ovary usually narrows into a stalk, called the style, that terminates in a sticky surface, the stigma. The ovary contains one or more ovules. The egg cell produced in each ovule becomes fertilized by the sperm brought to the ovule by the pollen tube. The latter is an outgrowth of a pollen grain that became attached to and germinated on the stigma. The pollen tube grows through the style to the ovule where it discharges the sperm. After the egg is fertilized by a sperm, the ovary, sometimes together with other floral parts, develops into a fruit. The ovules become seeds. *See* FLOWER; FRUIT; PLANT PHYSIOLOGY; PLANT TISSUE SYSTEMS; SEED. Katherine Esau

Bibliography. R. M. Klein and D. T. Klein, *Fundamentals of Plant Science*, 1987; J. D. Mauseth, *Plant Anatomy*, 1988.

Plant pathology

The study of disease in plants; it is an integration of many biological disciplines and bridges the basic and applied sciences.

Nature of Plant Disease

An understanding of the dynamic nature of disease in plants combines a knowledge of botanical sciences (for example, anatomy, cytology, genetics, and physiology); of disciplines that relate to causation of disease (for example, bacteriology, mycology, nematology, and virology); of disciplines such as ecology and meteorology which relate environmental factors to the severity of disease; and of those divisions of science that involve the impact of plant diseases upon society (for example, economics and sociology). As a science, plant pathology encompasses the theory and general concepts of the nature and cause of disease, and yet it also involves disease control strategies, with the ultimate goal being reduction of damage to the quantity and quality of food and fiber essential for human existence.

Fig. 1. Characteristic appearance of apple scab on leaves and fruit. (*H. H. Lyon, Cornell University*)

Diagnosis. Diagnosis of disease depends in part upon visible changes in color, form, and structure of the plant (**Fig. 1**). These changes, or symptoms, may involve death of cells and tissues (as in blights, rots); inhibition of plant growth and development (as in chlorosis, stunting); or increase in cell number and size (as in galls, witches'-brooms). Visible structures of the pathogenic agent, or signs, may also be present, as in bacterial ooze or powdery mildew. Identification of the cause of disease is included in diagnosis and becomes a basis of subsequent strategies in control. Many pathogenic agents may be isolated and grown in artificial culture. Such agents may then be used to inoculate healthy plants to test pathogenicity. These steps constitute Koch's Postulates, or Rules of Proof of Pathogenicity. Many other procedures are used in diagnosis: light and electron microscopy, biochemical tests (chromatography, electrophoresis), physical tests (determination of buoyant density, ultraviolet fluorescence), and serological tests (agar diffusion, enzyme-linked immunosorbent assay or ELISA), and molecular techniques.

Etiology. Etiology is the study of the causes of disease, and symptomology is the study of the symptoms expressed by the plant. The primary agent may be infectious, that is, transmissible from plant to plant (bacterium, fungus, or virus); or it may be noninfectious, that is, nontransmissible (for example, excess of nitrate). Infectious and noninfectious agents are considered pathogens because they cause disease. The term pathogen, however, is usually reserved for disease-inducing organisms, viroids, and viruses. The term abiotic agent is used for nonliving

causes of disease, such as excess or deficiency of a nutrient element. A parasite is an organism that derives food from its host. A pathogenic organism also usually derives food from its host, the susceptible plant, but it causes disease.

Epidemiolgy. Epidemiology is a study of how disease develops in populations. A sudden, widespread outbreak of disease is called an epidemic. The term epiphytotic is often used to denote such an outbreak in plants. The behavior and interactions of pathogens, plants, and environmental factors are considered quantitatively by epidemiologists.

Kinds of Plant Diseases

Diseases were first classified on the basis of symptoms. Three major categories of symptoms were recognized long before the causes of disease were known; necroses, destruction of cell protoplasts (rots, spots, wilts); hypoplases, failure in plant development (chlorosis, stunting); and hyperplases, overdevelopment in cell number and size (witches'-brooms, galls). This scheme remains useful for recognition and diagnosis.

When fungi, and then bacteria, nematodes, and viruses, were recognized as causes of disease, it became convenient to classify diseases according to the responsible agent. If the agents were infectious (biotic), the diseases were classified as being "caused by bacteria," "caused by nematodes," or "caused by viruses." To this list were added phanerogams and protozoans, and later mollicutes (mycoplasmas, spiroplasmas), rickettsias, and viroids. In a second group were those diseases caused by such noninfectious (abiotic) agents as air pollutants, inadequate oxygen, and nutrient excesses and deficiencies.

Other classifications of disease have been proposed, such as diseases of specific plant organs, diseases involving physiological processes, and diseases of specific crops or crop groups (for example, field crops, fruit crops, vegetable crops). Disease can be related to seven gross physiological functions of healthy plants: food storage, hydrolysis and utilization of stored food, absorption of water and nutrients by roots, growth and development of meristems, water translocation in the xylem, photosynthesis, and translocation of foods through the phloem. Diseases could then be grouped on the basis of their effect upon these essential functions. This system has much to offer. It directs attention to the disease phenomenon rather than to the pathogens, many different kinds of which could be responsible for a specific disease syndrome, for example, seedling blights. The system allows concentration on the environment as it affects the disease process, rather than on the plant-specific agents. It also forms a logical basis for exploration of possible control measures because it is related to specific stages of growth and development of the plant. Carl W. Boothroyd

Crop Losses to Disease

Diseases of agricultural crops are studied mainly because of the losses they cause; reliable crop loss data are therefore of primary importance in any plant protection program. It is not sufficient to state that diseases cause a loss; the magnitude of the loss must be quantified so that it can be related to the potential economic gain resulting from control measures. Information on crop losses is required by decision makers in all agencies involved in plant protection to allow formulation of sound policies, meaningful allocation of resources, and assignment of priorities. Similarly, the farmer must know the cost of the projected loss before it is decided how much to spend on control.

Assessment. The methodology to assess losses involves two phases. The first phase is an experimental study to develop a disease-loss model that describes the relationship between the amount of disease and the corresponding loss. The second phase involves large-scale surveys conducted to assess the amount of disease in commercial crops. The disease-loss models are then used to translate amount of disease into crop losses. This approach is time-consuming, but some countries, such as the United Kingdom and Canada, have demonstrated the usefulness of generating annual crop loss data for major crops. Crop loss studies in the United Kingdom in 1967 first demonstrated the importance of foliage diseases of barley and the need for fungicides, hitherto considered unnecessary but now applied to more than half of the total acreage of barley.

Such detailed information is generally not available, but there is a great need for reliable information on crop losses. Many countries, including the United States, have been involved in the development of programs on crop losses. The United States is the only country that has attempted to estimate losses due to disease for all crops. However, the reliability ranges from subjective estimates based on consensus of opinion to objectively generated experimental data.

Monocultures. The evolvement of crop monocultures with increasing genetic homogeneity increases the vulnerability of crops to diseases, and this could contribute to higher losses in the future. A leaf-blight epidemic on corn in the United States in 1970 dramatically demonstrated the susceptibility of modern varieties. This single epidemic caused a sudden and unpredicted $1 billion loss and seriously disrupted world food supplies. Substantial losses due to disease on a world basis continue to occur despite the continuous introduction of many resistant varieties and the large annual expenditures for fungicides. W. Clive James

Symptoms of Plant Diseases

Symptoms are expressions of pathological activity in plants. They are visible manifestations of changes in color, form, and structure: leaves may become spotted, turn yellow, and die (**Fig. 2**); fruits may rot on the plants or in storage (**Fig. 3**); cankers may form on stems (**Fig. 4**); and plants may blight and wilt. Diagnosticians learn how to associate certain symptoms with specific diseases, and they use this knowledge

in the identification and control of pathogens responsible for the diseases.

Those symptoms that are external and readily visible are considered morphological. Others are internal and primarily histological, for example, vascular discoloration of the xylem of wilting plants (**Fig. 5***a* and *b*). Microscopic examination of diseased plants may reveal additional symptoms at the cytological level, such as the formation of tyloses (extrusion of living parenchyma cells of the xylem of wilted tissues into vessel elements; **Fig. 6**).

It is important to make a distinction between the visible expression of the diseased condition in the plant, the symptom, and the visible manifestation of the agent which is responsible for that condition,

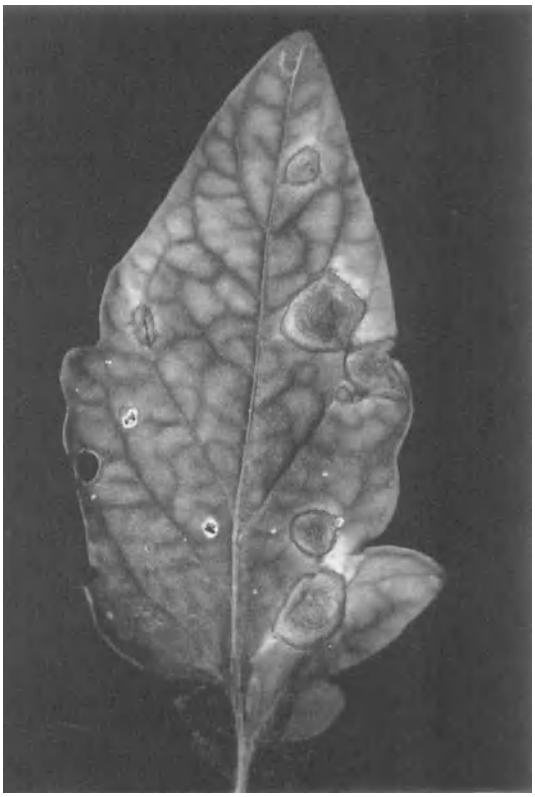

Fig. 2. Early blight of tomato. (*H. H. Lyon, Cornell University*)

Fig. 3. Black rot of squash. (*H. H. Lyon, Cornell University*)

Fig. 4. White pine blister rust. (*H. H. Lyon, Cornell University*)

the sign. The sign is the structure of the pathogen, and when present it is most helpful in diagnosis of the disease. For example, large galls may form on the ear of a corn plant. The gall is a symptom of disease and a typical expression of activity by the corn smut fungus. At a later stage of maturity that gall will break and dusty black spores will be exposed. The spore mass is the sign of that pathogen, and in association with the gall the problem is recognizable as corn smut (**Fig. 7**).

All symptoms may be conveniently classified into three major types because of the manner in which pathogens affect plants. Most pathogens produce dead and dying tissues, and the symptoms expressed are categorized as necroses. Early stages of necrosis are evident in such conditions as hydrosis, wilting, and yellowing. As cells and tissues die, the appearance of the plant or plant part is changed, and is recognizable in such common conditions as blight, canker, rot, and spot.

Many pathogens do not cause necrosis, but interfere with cell growth or development. Plants thus affected may eventually become necrotic, but the activity of the pathogen is primarily inhibitory or stimulatory. If there is a decrease in cell number or size, the expressions of pathological activity are classified as hypoplases; if cell number or size is increased, the symptoms are grouped as hyperplases. These activities are very specific and most helpful in diagnosis. In the former group are such symptoms as mosaic, rosetting, and stunting, with obvious reduction in plant color, structure, and size (**Fig. 8**). In

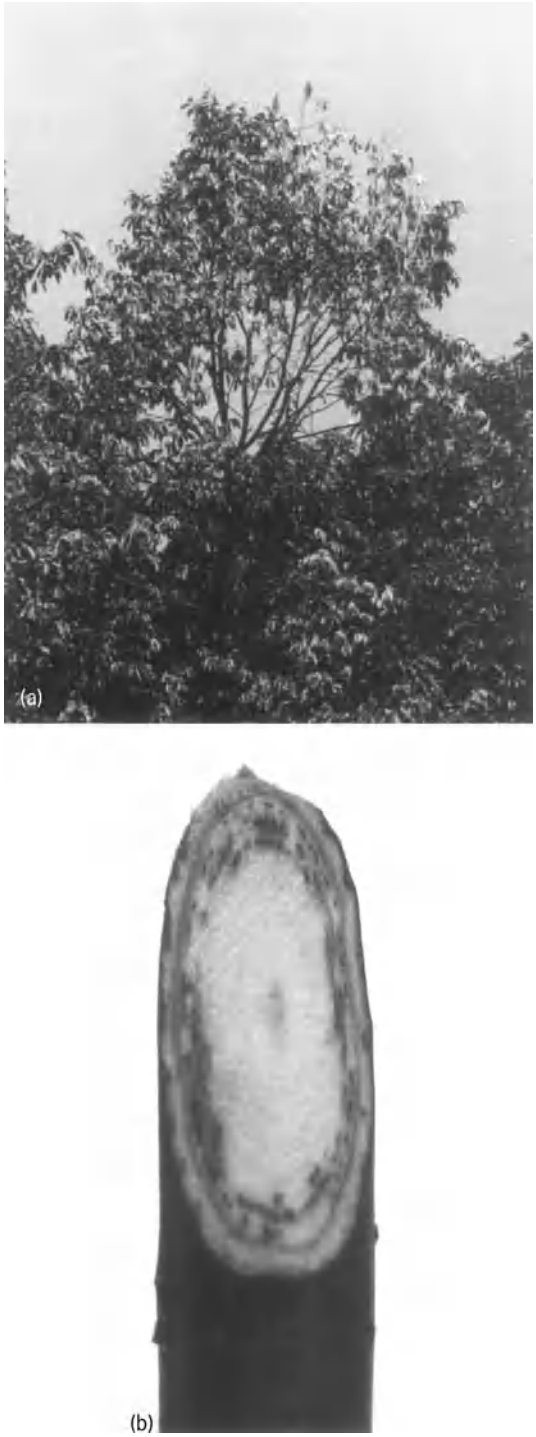

Fig. 5. Dutch elm disease. (*a*) Wilting of foliage. (*b*) Vascular discoloration in xylem. (*H. H. Lyon, Cornell University*)

the latter group are gall (**Fig. 9**), scab, and witches'-broom, all visible evidence of stimulation of growth and development of plant tissues. *See* CROWN GALL.

C. W. Boothroyd

Causative Agents of Disease

The primary agents of plant disease are fungi, bacteria, viruses and viroids, nematodes, parasitic seed plants, and a variety of noninfectious agents.

Fungi. More plant diseases are caused by fungi than by any other agent. The fungi include a vast variety of living organisms, all lacking chlorophyll. They are microscopic except for some species which at one stage in their life cycle produce visible fruiting bodies containing spores. The somatic phase (thallus) typically is threadlike or filamentous and grows

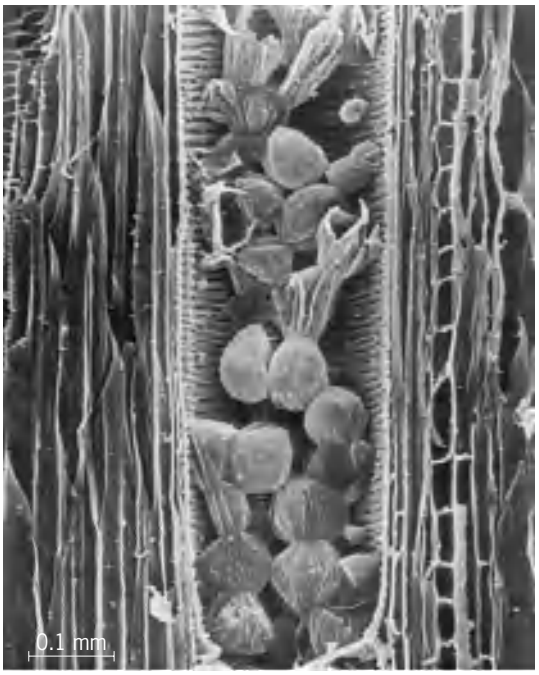

Fig. 6. Tylose formation in xylem vessel of banana with Fusarium wilt. (*G. E. VanderMolen, University of Rhode Island*)

Fig. 7. Gall formation in boil smut of corn. (*H. H. Lyon, Cornell University*)

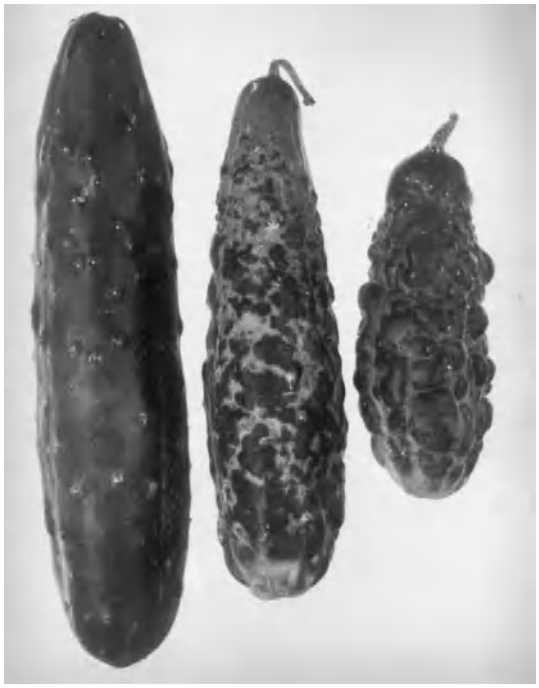

Fig. 8. Dwarfing and mosaic in cucumber mosaic; healthy fruit on left. (*H. H. Lyon, Cornell University*)

Fig. 9. Crown gall of tomato. (*H. H. Lyon, Cornell University*)

in a branching manner in all directions over or within plant tissues or other substrates. This thallus, known as mycelium, is usually septate, but septa or cross walls are lacking in some groups. Each filament of the fungal mycelium, called a hypha, is tubular and contains protoplasm, one or more nuclei, and most of the organelles, except chloroplasts, found in higher plant cells. Fungi typically reproduce by means of microscopic spores that may be formed by both asexual and sexual processes. The majority of fungi have spores that have one dimension from 5 to 50 micrometers. *See* FUNGI.

The fungi that cause plant disease derive their food from the plant (host) and are called parasites. Those that can live and grow only in association with living plant tissues are obligate parasites. Some fungi obtain their food from dead organic matter and are known as saprobes or saprophytes. Still others can utilize food from either dead organic matter or from living plant cells, and are referred to as either facultative parasites or facultative saprophytes.

Fungi are grouped into classes primarily on the basis of the morphology of their reproductive stages. The classes with organisms that cause plant disease are Plasmodiophoromycetes, Chytridiomycetes, Zygomycetes, Oomycetes, Ascomycetes, Basidiomycetes, and their asexual stages.

During the past three decades, the use of new technologies in cell biology, ultrastructure, and molecular systematics has resulted in a reevaluation of what are fungi. Currently, they are recognized as the Kingdom Fungi with four Phyla: the Chytridiomycota, Zygomycota, Ascomycota, Basidiomycota. The Form-class Deuteromycetes, which traditionally encompassed asexual states of fungi, has been largely abandoned in recent literature. The Plasmodiophoromycetes have been shown to align closely with Protista and the Oomycetes to the Stramenopila near the yellow-green algae (Chromista).

Plasmodiophoromycetes. The Plasmodiophoromycetes, called the endoparasitic slime molds, are a group of organisms related to protozoans along with slime molds. They do not have the threadlike mycelial phase of most fungi; the thallus is a naked multinucleate mass of protoplasm known as a plasmodium. Resting cells and thin-walled sporangia give rise to biflagellate zoospores. The zoospores penetrate root cells of the host and become ameboid, and the plasmodium develops. Only a few species in this class cause disease in cultivated plants. *Plasmodiophora brassicae*, the agent of club root of crucifers, is the best known.

Chytridiomycetes. Chytrid fungi produce motile cells (zoospores) with a single flagellum. Only a few species are economically important as causative agents (pathogens). A true mycelium is lacking in the species that cause plant disease. The simple thallus is entirely converted into reproductive structures called sporangia. Zoospores develop within the sporangia. One of the best-known disease-causing chytrids is *Synchytrium endobioticum*, the agent of black wart of potato.

Oomycetes. The Oomycetes include species of fungi that cause destructive plant diseases, most of which are in one order, Peronosporales. Included are the water molds, the downy mildews, and the white blisters (white rusts). *Plasmopara viticola*, the cause of downy mildew of grape, and *Phytophthora infestans*, the cause of late blight of potato and tomato, are examples of Oomycetes. All the plant parasitic species of the class have a filamentous, branched, coenocytic mycelium that grows between or within

the plant cells. In some species, haustoria, which are modified hyphal branches, penetrate the host cells. Asexual reproduction occurs by sporangia which, upon germination, produce biflagellate zoospores. In some species of downy mildew fungi the sporangia germinate by a germ tube.

Well-differentiated sex organs, antheridia and oogonia, are formed by these groups. Following fertilization, which is accomplished when a nucleus from the antheridium fuses with a nucleus in the oogonium, a thick-walled oospore is produced. The oospores germinate by germ tubes or by zoospores. The downy mildews derive their name from the abundance of mycelial branches (sporangiophores) bearing sporangia which grow out from the stomates of diseased leaves. This results in visible downy patches on the leaf. The species of *Albugo*, causative agents of white blisters, produce sporangia in compact masses under the host epidermis. A whitish blister results, and when the epidermis is ruptured a powdery white crust is evident.

Zygomycetes. Most of the Zygomycetes are saprobes, living on decaying plant and animal matter. One species, *Rhizopus stolonifer*, is a facultative parasite and can cause a serious rot disease of fleshy plant organs after harvest. Examples are soft rot of sweet potatoes and strawberries. *Rhizopus* has a filamentous mycelium without cross walls (nonseptate). Long aerial branches with spherical black spore bodies (sporangia) develop from the mycelium growing in or on the rotting fleshy roots or fruits. This cottony, whiskerlike growth speckled with black sporangia is the diagnostic sign of *Rhizopus*. The sporangia contain numerous small spores. A thick-walled resting spore called a zygospore develops as a result of sexual reproduction.

Ascomycetes. The Ascomycetes are the sac fungi, so named because the sexual or perfect-state spores are produced in a saclike cell, the ascus. Many such cells (asci), each typically containing eight ascospores, develop in differentiated fruiting bodies called ascocarps. Two kinds of ascocarps recognized for a long time are the perithecium and the apothecium. Perithecia are spherical or flask-shaped, completely closed or opening by a small pore. Apothecia are open and cup- or saucer-shaped. In later studies, two additional types of ascocarps, pseudothecia and ascostromata, have been described. They are similar in appearance to perithecia. Ascocarps may be single or aggregated and are visible signs of many of the plant-infecting Ascomycetes. However, in most instances the ascocarps must be studied microscopically to be certain of their identity. In addition to the ascospore state, most ascomycete fungi have an asexual or conidial spore state. The conidia are produced in tremendous numbers on branches of the septate mycelium or in specialized fruiting bodies.

The Ascomycetes are the largest class of fungal pathogens. Included here are the powdery mildew fungi, which are obligate parasites with a superficial mycelium but with haustoria penetrating the host epidermal cells. They produce enormous numbers of conidia that appear to the unaided eye as a whitish powdery coating on the host plant. The ascocarps have characteristic appendages that are of taxonomic value in separating the genera.

Other examples of Ascomycetes are *Claviceps purpurea*, causing the ergot disease of cereals; *Monilinia fructicola*, the cause of brown rot of stone fruits; *Venturia inaequalis*, the cause of apple scab; *Ceratocystis ulmi*, the cause of the Dutch elm disease; and *Endothia parasitica*, which has nearly eliminated the American chestnut in North America.

Basidiomycetes. Examples of Basidiomycetes that cause plant disease are the smuts, rusts, and the wood-decay fungi of forest trees. All fungi in this class produce basidiospores borne on the outside of a specialized cell or spore-bearing structure called the basidium. Fusion of two nuclei followed by meiosis occurs in the basidium prior to basidiospore formation. Four basidiospores on each basidium are usual. Four or five kinds of spores or reproductive cells may be produced by the rust fungi. These are aeciospores, urediniospores, teliospores, and basidiospores; spermatia (pycniospores), which are male sex cells, are also present. The various kinds of spores are usually produced in large numbers in pustules, sori, or other visible fructifications. The germinating teliospore of the rusts is the basidium. Some rust species have all the spore stages on one host (autoecious), whereas two different kinds of host plants are essential with heteroecious rust fungi. The rust fungal fructifications may be yellow, orange, red, brown, or black, and often powdery but are sometimes gelatinous or waxy. Rust fungi include *Puccinia graminis*, the causative agent of black stem rust of cereal grains; *Cronartium ribicola*, the cause of white pine blister rust; and *Gymnosporangium* species, which cause the cedar apple and cedar hawthorn rusts. About 4000 species of rust fungi are known, and all are obligate parasites.

Most of the economically important smut fungi are parasites of corn, sorghums, the cereal grains, and other grasses. The visible signs of these fungi are black, dusty masses of teliospores in host plant tissues. In the more destructive diseases, the spores develop in and destroy the flower parts so there is a total loss of grain yield. *Ustilago* species cause smuts of wheat, barley, corn, and oats. Species of *Tilletia* cause stinking smut or bunt of wheat.

The wood decay– and root rot–causing fungi of forest trees include *Armillaria mellea* and numerous species of *Fomes* and *Polyporus*. All of these fungi produce enormous numbers of basidiospores on gills (lamellae) or in pores or tubes on the underside of large complex fruiting bodies (basidiocarps). The basidiocarps are referred to as mushrooms, conks, or bracket and shelf fungi.

Deuteromycetes. There are several thousand species of fungi that so far as is known reproduce only by asexual spores, called conidia. Lacking a sexual (perfect) state, they have been referred to as imperfect fungi and placed in an artificial form-class Deuteromycetes. The classification of the Deuteromycetes is based on the morphology of the conidial states and the mode of development of the conidia.

Conidia of these species vary greatly in size, shape, and color, and may be one- to many-celled. They are formed singly or clustered on specialized hyphal branches, or in large numbers in specialized fructifications known as pycnidia, acervuli, or sporodochia. Many Deuteromycetes cause serious leaf spots and flower blights, as well as fruit rots, stem cankers, and wilts. Examples include species of *Septoria, Gloeosporium, Diplodia, Alternaria*, and *Verticillium*. C. Wayne Ellett; J. W. Kimbrough

Bacteria. Eleven Gram-negative and seven Gram-positive genera of bacteria are associated with diseases of flowering plants. Species affected by these bacterial diseases include the major cereal, forage, vegetable, and fruit crops. The bacterial pathogens of plants are almost exclusively associated with plants, and do not cause serious diseases of humans or animals.

Types of bacterial diseases. Each bacterial species produces distinctive symptoms on the host(s) that it attacks. Typical symptoms caused by bacterial plant pathogens, and the main genera of plant pathogenic bacteria are illustrated in **Fig. 10**.

The bacteria most commonly associated with plant diseases are Gram-negative, non-spore-forming, rod-shaped cells and include the genera *Agrobacterium, Acidovorax, Burkholderia, Erwinia, Ralstonia, Pantoea, Pseudomonas, Rhizomonas, Xanthomonas, Xylella*, and *Xylophilus*. Some of the bacterial species that are Gram-positive are in the genera of *Arthrobacter, Bacillus, Clavibacter, Clostridium, Cutrobacterium, Rhodococcus*, and *Streptomyces*. Species of *Bacillus* and *Clostridium* produce endospores that are survival mechanisms for the bacteria. Two additional groups of prokaryotes that are commonly associated with plants and are extremely difficult or virtually impossible to culture are the spiroplasmas and phytoplasmas (formerly termed mycoplasma-like organisms or MLOs). Both groups of these prokaryotes lack a true cell wall and are relatively recent members of the Mollicutes. These latter bacterial plant pathogens are associated with yellows-type diseases and various plant growth abnormalities; they inhabit phloem tissue in diseased plants. The spiroplasmas are helical cells that can be cultured on complex, artificial media.

Most of the typical bacterial plant pathogens are motile as a result of the rotation of threadlike flagella

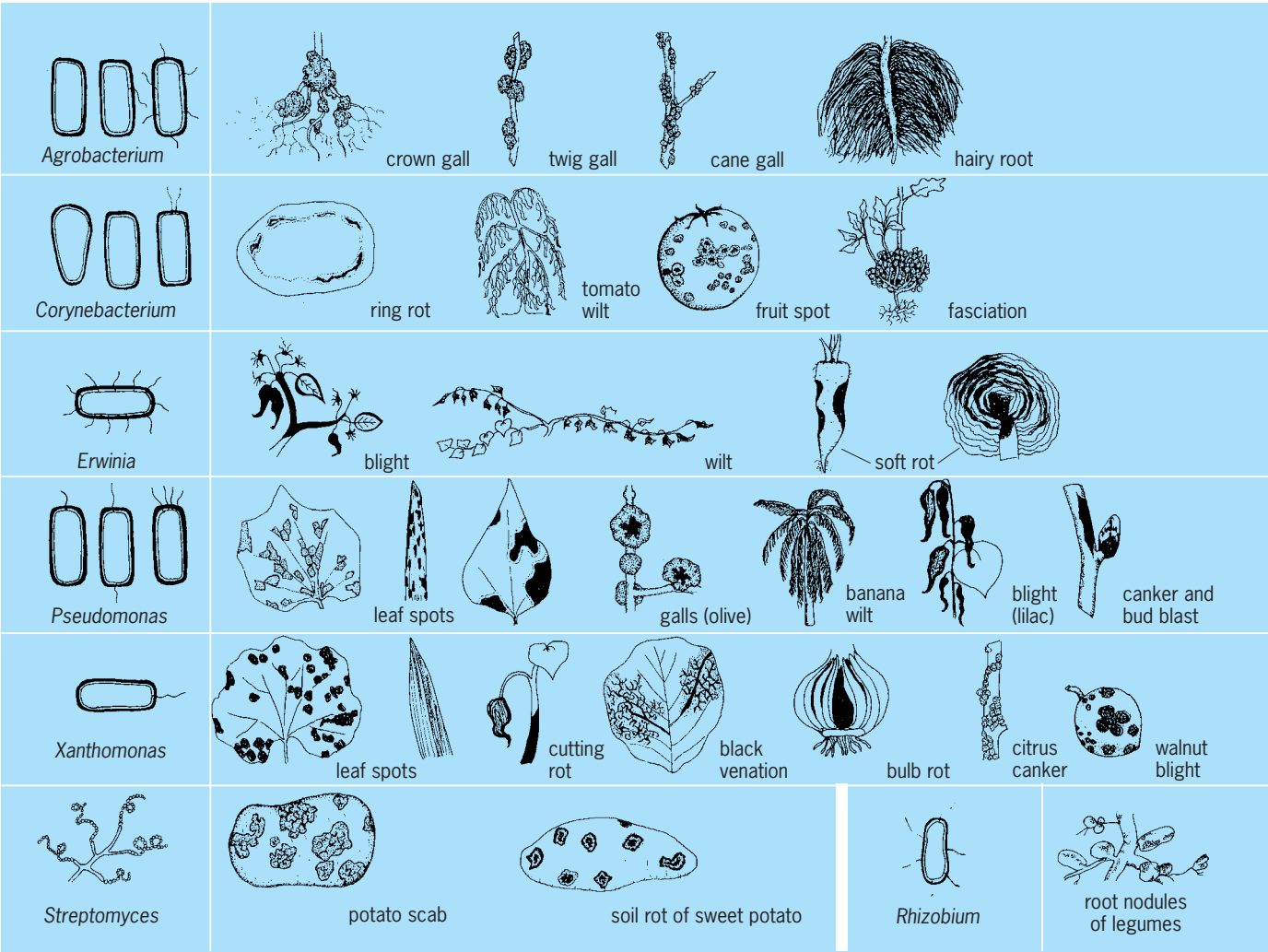

Fig. 10. Genera of bacteria and types of symptoms caused by species in these genera. (*After G. N. Agrios, Plant Pathology, 2d ed., Academic Press, 1978*)

attached to the inner and outer membranes. These flagella are involved for movement in water. The flagella are arranged randomly around the cell surface (peritrichous) in *Erwinia* and *Agrobacterium* species. Several other plant-pathogenic bacteria, such as *Xanthomonas*, have a single flagellum (monothrichous) attached to one end (polar). *Pseudomonas* species contain a tuft of flagella at one end (lophotrichous).

Bacterial cells divide by binary fission and produce millions of cells in a short period of time. Thus, a bacterium that multiplies every 30 min will produce more than 10^6 cells in less than 10 h. During periods of optimal conditions for bacterial multiplication after infection of plant tissue, disease can develop extensively and quickly.

Classification. Bacteria cannot be identified to species based on colony characteristics, cell morphology, or staining procedures. Therefore, biochemical, serological, genetic techniques, fatty acid analysis, protein profiles, and sensitivity to phages have been useful to classify species. An equally important test to identify bacteria is to demonstrate pathogenicity on a particular host or hosts.

Means of spread and survival. Many foliar bacterial pathogens are disseminated long distances by splashing rain and by wind-blown mist. When rain events occur, bacterial cells ooze from infected stem or leaf tissue or through lenticels or stomata. During rain events, tiny droplets and aerosols that contain bacteria are produced and carried to healthy plants. Bacteria may also be moved from plant to plant by insects, in irrigation water, and by cultural activities used during crop production (such as pruning). Several bacterial plant pathogens require an insect vector for dispersal. The bacterium that causes Pierce's disease of grapes requires a leafhopper insect to transmit, infect, and cause disease.

Most bacterial plant pathogens survive adverse conditions in the host plant. Those bacterial species that incite diseases of perennial hosts are particularly well adapted to long-term survival. Some bacteria have been reported to survive in the rhizosphere of grasses during the winter months. Few bacterial species are able to survive for long periods in the soil free of the host plant; however, the pathogen of bacterial wilt of tomato and potato is reported to survive for many years in fallow soil. Some species survive in the rhizosphere of nonhost plant species. Many bacteria are able to survive on or in seed. Plant pathogenic bacteria are also able to survive in irrigation water.

Pathways of entry. Bacterial plant pathogens are not able to penetrate directly through the leaf or stem; instead they enter through natural openings or wounds. Bacteria can enter through stomata or hydathodes in leaves, and through lenticels in corky tissue. Fresh wounds are excellent portals for bacterial ingress. Chewing insects and nematodes produce wounds that favor entry by bacteria. A leaf miner insect causes large wounds in the leaves of citrus trees which greatly accelerates the spread of bacterial canker. Cultural practices such as pruning and cultivation may create wounds and increase the likelihood of infection by bacteria such as the crown-gall bacterium.

Mechanisms of pathogenesis. The mechanisms associated with pathogenesis are quite variable and are not completely understood. One interesting mechanism has been elucidated recently. Many bacterial pathogens of animals as well as bacterial pathogens of plants have a unique secretion system that is associated with pathogenesis. It is termed the type III secretion system. Gram-negative bacteria use this system to secrete proteins across their inner and outer membranes into the host cell. This enables Gram-negative bacteria to inject pathogenicity-related proteins into the cytosol and nucleus of the eukaryotic host. It is unique among secretion pathways since it delivers virulence proteins into host cells.

Bacterial cells produce many compounds that are involved in disease development. Soft rot bacteria produce enzymes that hydrolyze pectic substances between the cell walls of plants to destroy the integrity of the tissue. This results in a mushy appearance of the host structure. Some bacteria produce phytotoxins that are low-molecular-weight compounds. These phytotoxins directly injure plant cells and influence the course of disease development or symptoms. One toxin, coronatine, is produced by several bacterial plant pathogens. Coronatine induces chlorosis, stunting, and hypertrophy.

Some bacteria that are associated with vascular-wilt diseases produce gumlike substances or extracellular polysaccharides (EPS). These appear to be particularly important in vascular wilt diseases. Extensive wilting occurs in these diseases partly from accumulation of bacteria and also from the release of extracellular polysaccharides. The extracellular polysaccharides pass into functional vessels where they block xylem cavities and accumulate in the pit membranes.

The Gram-negative, soil-inhabiting bacterium *Agrobacterium tumefaciens* genetically transforms host cells by inserting a piece of DNA, designated as T-DNA, into the host chromosome. The T-DNA carries genes involved in hormone synthesis, and these genes are then expressed in the host plant; the result is the induction of galls. This is the only known natural example of the transport of DNA between kingdoms of organisms. The *Agrobacterium* system has been an effective tool to introduce foreign genes into plants.

Control. The control of bacterial diseases relies heavily on the reduction of primary sources of inoculum. This includes bacterial pathogens in seeds, budwood, vegetative propagative material, young plants, crop residue that has not fully decomposed, alternative hosts, and various other sources. Therefore, control practices include the elimination of these sources of bacteria through seed certification and agronomic cultural practices to reduce or eliminate sources of primary inoculum. Plant varieties with resistance to bacterial diseases have been developed to control some of the foliar pathogens. Chemical control has been used quite extensively,

but with limited success. Antibiotics and copper bactericides have provided some control, but numerous bacterial species have developed resistance to these compounds. The strategies have also included the use of biological control with microbes antagonistic to the bacterial pathogen or microbes that induce a systemic-acquired resistance (SAR) in the plant. These microbes are applied prior to infection by the bacterial pathogens, and this results in reduced disease. Biological control has been used successfully to control crown gall, incited by *Agrobacterium tumefaciens*, with a nonpathogenic strain of *A. radiobacter*, known as K84. This strain is antagonistic to the causal agent because it produces an inhibitory compound termed a bacteriocin. More recently, biotechnological research has resulted in the cloning of resistance genes and the transfer of these genes to other genotypes to the same plant species or to related plant species. J. B. Jones

Viruses and viroids. Viruses and viroids are the simplest of the various causative agents of plant disease. The essential element of each of these two pathogens is an infective nucleic acid. The nucleic acid of viruses is covered by an exterior shell (coat) of protein, but that of viroids is not. *See* PLANT VIRUSES AND VIROIDS.

Approximately 400 plant viruses and about 10 viroids are known. Because of their small size, they can be seen only with the electron microscope. They are of two shapes, isometric ("spherical") and anisometric (tubular). Isometric viruses range from 25 to 50 nanometers in diameter, whereas most anisometric viruses are of 12–15-nm diameter and of varying length (200–2000 nm).

The nucleic acid of most plant viruses is a single-stranded RNA (mol wt 1.5–4 $\times$ 10^6); a number of isometric viruses have a double-stranded RNA (mol wt 10–20 $\times$ 10^6). A few viruses contain double-stranded DNA (mol wt 4–5 $\times$ 10^6), and several containing single-stranded DNA (mol wt 800,000) have been reported. The nucleic acid of viroids is a single-stranded RNA, but its molecular weight is much lower than that of viruses (about 120,000).

Replication. Viruses require living cells for their replication. Some viruses, such as tobacco mosaic virus (TMV) and cucumber mosaic virus, are found in many plant species; others, such as wheat streak mosaic virus, occur only in a few grasses. The replication of a plant virus appears to proceed according to the following general scheme: introduction of the virus to a plant through a wound, release of the nucleic acid from the protein coat, association of viral RNA (or messenger RNA of DNA viruses) with cellular ribosomes for its translation to the proteins required for virus synthesis, replication of the nucleic acid and production of coat protein, and assembly of the nucleic acid and coat protein to form complete virus particles. The replication of viroids is not clearly understood at present. Cell-to-cell spread of viruses usually occurs, and eventually the virus spreads throughout the plant. In some plants, the cells surrounding the initially infected cells die, and the virus usually does not spread further.

Transmission. Viruses are transmitted from plant to plant in several ways. Grafting scions to infected rootstocks often results in virus infection of the scion. Cuttings taken from infected mother plants will usually be diseased. Contact through leaf rubbing between healthy and diseased plants causes minute wounds through which viruses enter healthy plants. The majority of viruses are transmitted by vectors such as insects, mites, nematodes, and fungi which acquire viruses during feeding upon infected plants. Some viruses are transmitted to succeeding generations by infected seed. Viroids are spread mainly by contact between healthy and diseased plants or by the use of contaminated cutting tools.

Control. The control or prevention of virus diseases involves breeding for resistance, propagation of virus-free plants, use of virus-free seed, practices designed to reduce the spread by vectors, and, in some cases, the deliberate inoculation of plants with mild strains of a virus to protect them from the deleterious effects of severe strains. R. I. Hamilton

Nematodes. Nematodes (roundworms) are one of the most common groups of invertebrate animals and are found in a wide variety of aquatic habitats, both marine and fresh water, from the Antarctic sea bottom to thermal springs. All soils that support plant life contain nematodes living in the water films that surround soil particles. Most nematodes feed primarily on microscopic plants, animals, and bacteria, but a few, such as hookworms, pinworms, and ascarids, are parasites of animals; another relatively small group of nematodes parasitize plants. *See* NEMATA (NEMATODA).

Movement is serpentine, and there are no body appendages. Reproduction is by the laying of eggs. Males may or may not be present, and adults of some species are hermaphroditic. All species known have four larval (juvenile) stages and an adult stage separated by molts. The first molt takes place within the egg shell. In some species, adult females become swollen and sedentary. Plant-parasitic nematodes are distinguished by their small size, about 1 mm average length, and mouthparts that are modified to form a hollow stylet which is inserted into plant cells. Food intake of nematodes is by suction provided by a muscular esophagus.

All of the plant parasites are placed into two orders. The order Tylenchida have stylets with basal knobs, a three-parted esophagus (pharynx), and an ornamented integument (cuticle). The order Dorylaimida have a simple tooth, a two-parted esophagus, and a smooth cuticle.

In nature, all nematodes ingest living food, and plant feeders are considered obligate parasites in that they do not grow and reproduce in the absence of a living host plant. Feeding of ectoparasitic species is restricted to the root surface, whereas endoparasites move into plant tissues. Some species become sedentary, others remain migratory during feeding.

Injury to plants. Plant injury is of three general types and is related to feeding habits. Migratory endoparasites destroy tissues as they feed, producing necrotic

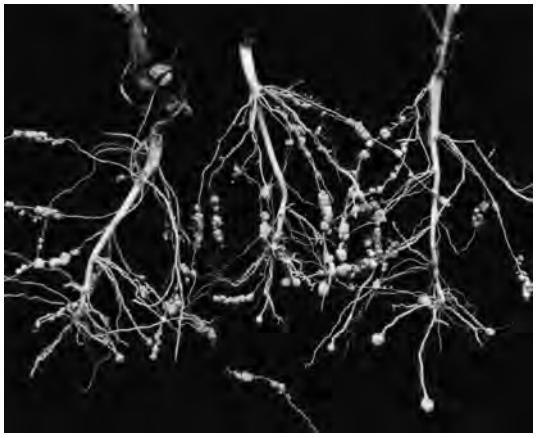

Fig. 11. Root knot nematode galls on roots of potato.

lesions in the root cortex. Other migratory endoparasites invade leaf tissues and produce extensive brown spots. Sedentary endoparasites do not kill host cells, but induce changes in host tissues, which lead to an elaborate feeding site or gall (**Fig. 11**). A unique feature of these galls is that the host plant cells surrounding the nematode's head become enlarged, with thickened walls, dense cytoplasm, and many enlarged nuclei. These syncytia, or giant cells, are concentrated centers of nutrients that have been diverted from normal metabolism of the host. A few species produce galls on stems or in seeds, and galls are also produced by some ectoparasites. The third general type of symptom is produced by certain migratory ectoparasites, where root tips are devitalized and cease to grow without any associated swelling or necrosis.

In addition to the plant injury that they cause directly, nematodes are important factors in disease complexes. Lesions and galls provide entrance courts for soil fungi and bacteria, and many diseases caused by soil-borne pathogens are more severe when nematodes are present. Plant varieties resistant to soil-borne fungi sometimes lose this resistance when previously parasitized by nematodes. Important viruses such as grapevine fan-leaf and peach stem pitting are transmitted by nematodes of the order Dorylaimida.

Important species. Root knot nematodes (*Meloidogyne*) are worldwide in distribution with many species that attack most kinds of plants (Fig. 11). The major symptoms are swollen galls on roots. Nematode eggs are produced in jellylike masses on the surface of galls.

Cyst nematodes (*Globodera* and *Heterodera*) are so called because eggs are retained inside the swollen female cuticle, which becomes tanned and leathery. There is relatively little root swelling, and cysts remain attached to roots by their embedded necks.

Lesion nematodes (*Pratylenchus*) and foliar nematodes (*Aphelenchoides*) produce brown necrotic spots on roots and leaves, respectively.

Stem and bulb nematodes (*Ditylenchus*) produce bloated swellings near the soil line. The swollen areas contain large numbers of nematodes which sometimes form resistant desiccated tufts of "narcissus wool." Clovers, alfalfa, and vegetable crops are also important hosts of stem and bulb nematodes.

Seed gall nematodes (*Anguina*) produce blackened "cockles," which are grains that have been largely replaced with thousands of nematodes. Seed gall nematodes remain viable for many years when dried, and the pathogen is transmitted when galls are planted with healthy seed.

Control. Control of plant-parasitic nematodes often is based on selection of nonhosts for crop rotations or nematode-resistant varieties when they are available. Some plant species, such as marigolds, release compounds into the soil that are toxic to nematodes. Some lessening of injury is achieved by irrigation and by fertilization, particularly with potash, since many nematode-induced diseases lead to potassium deficiency. Animal manures, compost, and other organic amendments enhance the buildup of natural enemies of nematodes, particularly the nematode-trapping fungi. Chemical control usually consists of treatment of soils with liquid chemicals that volatilize, producing toxic fumes which move through the soil mass. The high cost of soil fumigation and the requirement for special equipment and preparation restrict this practice to high-cash-value crops. Some of the systemic organophosphate and carbamate insecticides are effectively used, at higher dosages, to control nematodes in living plants. R. A. Rohde

Seed plants. Many parasitic seed plants (estimated at nearly 3000) attack other higher plants. They occur in at least 15 different families. In some families (for example, the mistletoes Loranthaceae and Viscaceae) all members are parasitic; in others, only a single genus is parasitic in an otherwise autotrophic family (for example, the dodder, genus *Cuscuta*, of the morning glory family Convolvulaceae).

Most parasitic plants are terrestrial; that is, the parasitic connection with the host plant is through the roots. Other parasitic plants grow on the aboveground parts of the host. Some plants are classed as semiparasites because they can live in the soil as independent plants for a time, but are not vigorous or may not flower if they do not become attached to a suitable host.

The nutritional status of parasitic plants ranges from total parasites with no chlorophyll (for example, the broomrapes, Orobanchaceae) to plants that are well supplied with chlorophyll and obtain primarily water and minerals from their hosts (many mistletoes).

Some parasitic plants cause serious damage in agricultural crops. Among the most damaging are the broomrapes and dodder. An Asian plant, the witchweed (*Striga asiatica*), has recently become established in the southeastern United States and is a serious threat to corn and other crops. The U.S. Department of Agriculture has begun an intensive research and control program to combat this damaging weed.

Parasitic plants, particularly the mistletoes, are also serious pests in the forest, orchard, and ornamental trees in many parts of the world. The dwarf

Fig. 12. Many small plants of a dwarf mistletoe (*Arceuthobium vaginatum*) parasitizing a ponderosa pine branch. This is the most damaging disease of ponderosa pine in many parts of the West.

mistletoes (genus *Arceuthobium*) are among the most damaging pests of coniferous forests in the western United States (**Fig. 12**); they cause losses of billions of board-feet of wood annually.

Parasitic higher plants of economic importance are being combatted in many ways: by breeding resistant crops, using less-susceptible species, applying chemicals, and by cultural means such as pruning out infected branches or removing diseased plants.

Frank G. Hawksworth

Noninfectious agents of disease. Plants with symptoms caused by noninfectious agents cannot serve as sources of further spread of the same disorder. Such noninfectious agents may be deficiencies or excesses of nutrients, anthropogenic pollutants, or biological effects by organisms external to the affected plants.

Deficiencies or excesses of nutrients. For proper growth, plants need periodically renewed supplies of nitrogen, phosphate, and potash, plus a number of essential minor elements. Nutrient availability may be affected by the acid-base reaction of the soil. *See* FERTILIZER; PLANT MINERAL NUTRITION.

Plants deficient in nitrogen appear chlorotic and develop yellowed older leaves. A phosphate-deficient plant may be darker green than normal with reddened or purpled leaf tips or edges. Plants deficient in potash develop light green leaves punctured with dead spots at the tips and along the margins. Deficiencies of some essential minor elements are associated with a variety of diseases (see **table**).

Naturally occurring excesses of sodium salts are the common cause of alkali injury in many arid regions of the world. A similar problem is caused by the application of deicing salts to highways. Polluting the soil adjacent to the travel lanes, this salt can kill affected plants.

Injury from excessive nitrogen is common on house plants and on seedlings in greenhouses. Plant damage has also been reported from soils excessive in aluminum, arsenic, boron, calcium, chromium, copper, fluoride, iodine, iron, lithium, magnesium, manganese, and nickel. Too much of some of these elements can unbalance the intake or metabolism of other essential nutrients.

Anthropogenic pollutants. The deicing salt, just mentioned, is a common example of an anthropogenic pollutant.

On the farm, plant-damaging pollution may be caused by careless use of pesticides. Although insecticides and fungicides can harm plants if not used properly, mishandled herbicides are by far the most damaging to plants. Herbicide sprays may drift unnoticed onto adjacent fields; herbicide residues may remain in soil to damage a succeeding crop; or herbicide residues may seep into wells used for irrigation. Herbicides may also contaminate other pesticide sprays or fertilizers when spraying or mixing equipment is not properly cleaned.

Another source of pollution on the farm is ethylene generated naturally by ripening fruit. Ethylene, a powerful gaseous plant hormone, affects the growth habits and metabolism of most plants. If apples are left in storage with carrots, for example, the ethylene from the apples will produce a bitter-tasting metabolite in the carrots.

Off the farm, anthropogenic air pollutants are generated by industrial processes, and by any heating or transportation method that uses fossil fuels. The most common air pollutants that damage plants are sulfur oxides and ozone. Sulfur oxides are produced when sulfur-containing fossil fuels are burned or metallic sulfides are refined. Human-generated ozone is produced by sunlight acting on clouds of nitrogen oxides and hydrocarbons that come primarily from automobile exhausts.

There are other plant-damaging air pollutants as well. Gaseous fluorides come from factories that manufacture aluminum, steel, ceramics, and phosphate fertilizers. Ethylene, mentioned before, also comes from waste hydrocarbons emitted by poorly functioning internal combustion engines and oil-burning space heaters. Ethylene is not found in natural gas, but is a component of manufactured gas. This source may again become a problem if there is a return to synthetic fuels. Other air pollutants that have damaged plants are the smog-component peroxyacetyl nitrate (PAN), chlorine escaping from

Examples of plant diseases associated with deficiencies of minor elements

Deficient minor element	Plant disease
Boron	Internal cork of apples
Calcium	Blossom-end rot of tomatoes
Copper	Exanthema of citrus
Iron	Maple chlorosis
Magnesium	Sand drown of tobacco
Manganese	Gray speck of oats
Molybdenum	Whiptail of cauliflower
Zinc	Little leaf of peach

water treatment plants, and cement dust. Methane, the principal component of natural gas, is not itself toxic to plants, but may act as food for methane-consuming bacteria in the soil. By proliferating, these bacteria can kill plant roots indirectly by depleting soil oxygen.

The symptoms of air pollution injury include leaf flecking, spotting, and mottling; red discoloration of leaves and fruits; withering of flower parts; stunting; and death of plant parts on the entire plant.

Some air pollutants, such as sulfur oxides and fluorides, accumulate within the exposed plant. Others, such as ozone and PAN, cannot be detected in the injured plant. The combination of sulfur oxides plus ozone is more toxic to some plants than either pollutant alone.

The severity of the damage depends on the specific pollutant or mixture of pollutants, the pollutant concentration, the length of exposure, the plant cultivar, the condition of the plant when it was exposed, and environmental conditions. Outdoors, air pollution episodes usually occur when temperature inversions confine and concentrate the pollutants. *See* AIR POLLUTION; WATER POLLUTION.

Biological effects by external organisms. The effect of apple ethylene on carrots, mentioned earlier, may be considered an example in this category.

Another group of such disorders may be classed as allelopathic. This term describes the harmful effect on one plant produced by toxic secretions from another plant or plant residue. For example, most plants grow poorly or not at all in the soil over the roots of black walnut trees. The alleged cause is the secretion of a quinonic compound by black walnut roots.

Tobacco frenching is a disease apparently caused by toxic diffusates from soil microorganisms, such as *Bacillus cereus*, that do not colonize tobacco plants. The suspected toxicants are isomers of isoleucine.

Saul Rich

Pathogenic Life Cycles of Agents

Plant pathogens initiate and induce disease in plants in a series of steps that make up the disease cycle and include overwintering of the pathogen, production or liberation of inoculum, dissemination of inoculum, inoculation, penetration, infection, and growth or multiplication of the pathogen leading to production of new inoculum. The new inoculum may cause one or several secondary disease cycles during the growing season, or it may overwinter and cause infection the following growing season. Disease cycles of a typical plant-pathogenic fungus, bacterium, mycoplasma, virus, and nematode are shown in **Figs. 13–17**.

Overwintering. Pathogens overwinter in infected plants (such as trees), in surviving belowground parts of plants, in some other perennial plant, in plant debris, in or on seeds, in some vector such as an insect, or in the soil. Viruses, mycoplasmas, and protozoa overwinter as such inside living cells. Bacteria usually overwinter when they are in contact with living cells. Fungi, parasitic higher plants, and nematodes overwinter, respectively, mainly as spores, seeds, or eggs either on, or usually apart from, the infected plant.

Production and liberation of inoculum. Inoculum, that is, the number of individual pathogens that can be liberated and disseminated to start new infections, may consist of the vegetative body of the pathogen, as in bacteria, viruses, and so on, or as with fungi and parasitic higher plants, it may consist of spores (Fig. 13) and seeds, respectively. The first inoculum of the growth season is called primary inoculum, and causes primary infections. Some pathogens (viruses, mycoplasmas, protozoa, some bacteria) never release inoculum outside the plant, but it is transferred instead with living cells (such as buds, grafts), or within vectors (such as insects; Fig. 15). Fungi, parasitic higher plants, most bacteria, and nematodes produce inoculum outside the host plant, or else it exudes or moves out of the plant, usually in the presence of moisture (Figs. 13, 14, and 17).

Dissemination of inoculum. Inoculum is disseminated with the help of wind (Fig. 13), running or splashing water (Figs. 14 and 17), insects (Fig. 15), other animals, and humans directly or indirectly by contaminated tools. A few pathogens move on their own for a few centimeters (bacteria, fungal zoospores) or up to a meter (nematodes).

Inoculation. Inoculation is the coming into contact of the pathogen with a susceptible part of the host plant in a way that could result in disease. With some pathogens inoculation is accomplished when the pathogen comes in contact with an intact surface, a wound, or a natural opening of the host plant (Figs. 13, 14, and 17); with others the pathogen must be brought into contact with a wounded living cell (Figs. 15 and 16).

Penetration. Viruses, mycoplasmas, protozoa, and some bacteria are inserted into plants by vectors and thereby accomplish penetration concurrently with inoculation (Figs. 15 and 16). Fungi, most bacteria, parasitic higher plants, and nematodes come into contact with the outer surface of the plant and enter it through wounds (Fig. 14), through natural openings such as stomata and lenticels, or through direct penetration of the cuticle (Fig. 13) and cell walls (Fig. 17). These pathogens penetrate susceptible plants directly by secreting enzymes that break down the plant cell walls and by exerting physical pressure on the plant.

Infection. Pathogens cause infection the moment they begin to obtain nutrients from their host plants. Infection continues either until the plant defends itself successfully and kills, or stops the effects of, the pathogen, or until the plant is finally killed by the pathogen. Infection takes place when suscept and pathogen are compatible with each other. In some instances, initial recognition of such compatibility seems to be made possible by the interaction of complementary molecular configurations of specific pathogen proteins called lectins with specific sugars contained in more complex molecules at the plant surface.

During infection, pathogens produce enzymes that disintegrate plant cells and neutralize the activity of cellular components, toxic substances (toxins)

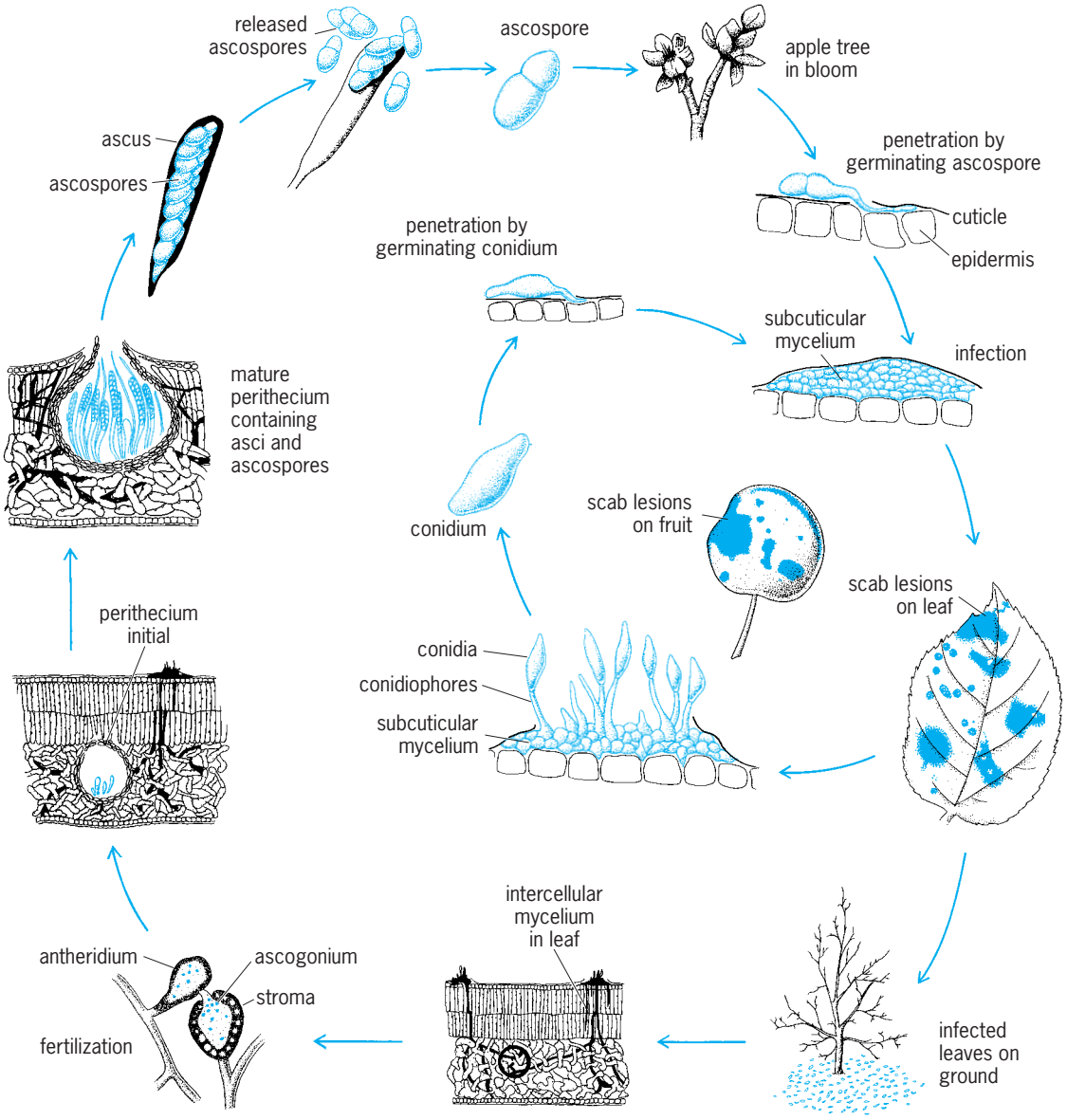

Fig. 13. Development of the disease cycle of apple scab caused by the fungus *Venturia inaequalis*. (*After G. N. Agrios, Plant Pathology, 2d ed., Academic Press, 1978*)

that inactivate plant proteins and destroy the function of plant membranes, and growth regulators that upset the hormonal balance in the plant. The plant also produces enzymes, toxic substances, and growth regulators that contribute to its defense effort against the pathogen.

The reaction between suscept and pathogen is determined by the DNA present in the chromosomes of each. Recent work, however, indicates that some diseases are determined by additional genetic material present in pathogens as relatively small circular strands of DNA called plasmids. Thus, plasmid DNA carries the genes responsible for production of toxins by some pathogens and for inducing uncontrolled cell division and gall formation (malignancy) by others (Fig. 14).

When infection occurs, parts of the plant are destroyed or altered in ways characteristic of the suscept-pathogen combination. These manifestations by the infected plant are the symptoms. The period of time between initiation of infection and appearance of symptoms (incubation period) is characteristic of the particular suscept-pathogen combination and the prevailing temperature; it varies from 2 days to 3 weeks in most diseases and up to 2–3 years in some diseases of trees.

Growth or multiplication of the pathogen. During infection most pathogens are inside the plant and increase only in numbers without increasing much in size; others, especially fungi and parasitic higher plants, increase primarily in size. Plant pathogenic bacteria, mycoplasmas, and protozoa multiply by fission, that is, each individual pathogen splits into two identical pathogens; viruses are replicated by the millions by the host cell (Fig. 16); and some nematodes lay eggs that hatch into more nematodes inside

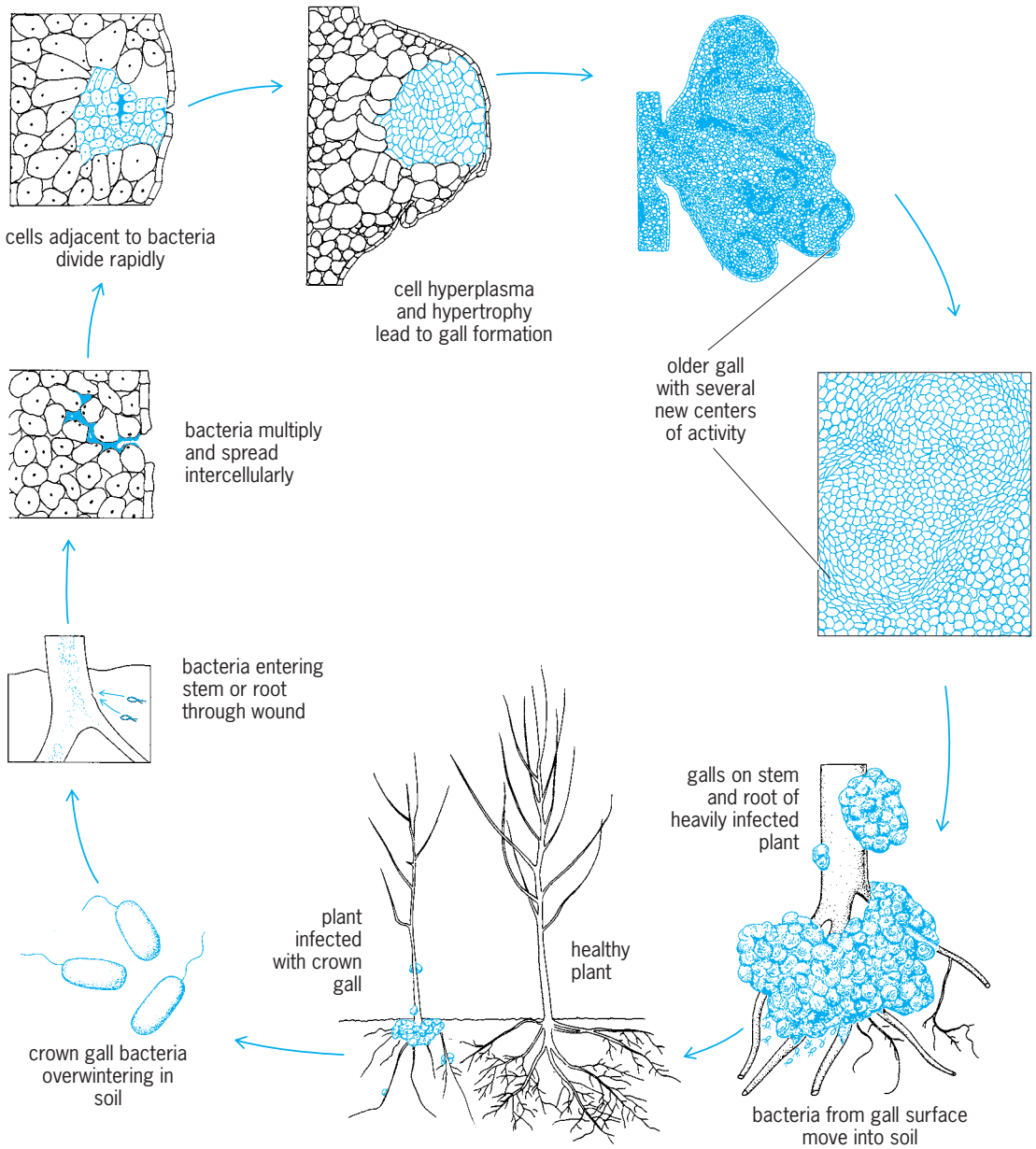

Fig. 14. Development of the disease cycle of crown gall caused by the bacterium *Agrobacterium tumefaciens*. (*After G. N. Agrios, Plant Pathology, 2d ed., Academic Press, 1978*)

the plants (Fig. 17). Viruses, mycoplasmas, protozoa, and some bacteria spread as such through the entire plant, particularly in the conductive tissues (phloem or xylem), whereas most bacteria and nematodes spread only into localized areas of plants. Parasitic fungi grow into their host plants as roots grow into soil, and almost all of them produce masses of countless spores at, above, or slightly below the plant surface (Fig. 13). Parasitic higher plants send food-absorbing structures into host plants and multiply by producing seeds outside the hosts.

Thus, in or on every infected plant, thousands, millions, or even billions of new inoculums are produced which, when disseminated to susceptible plants under favorable weather conditions, initiate new (secondary) infections. Secondary infections may be repeated from a few to 20 or more times during each growing season, depending on the pathogen. At the end of the growing season, though, pathogens go into dormancy or produce specialized hardy structures by which they overwinter and thus close the disease cycle. George N. Agrios

Epidemiology of Plant Disease

Epidemiology is the study of the intensification of disease over time and the spread of disease in space. The botanical epidemiologist is concerned with the interrelationships of the host plant (suscept), the pathogen, and the environment, which are the components of the disease triangle. With a thorough knowledge of these components, the outbreak of disease may be forecast in advance, the speed at which the epidemic will intensify may be determined,

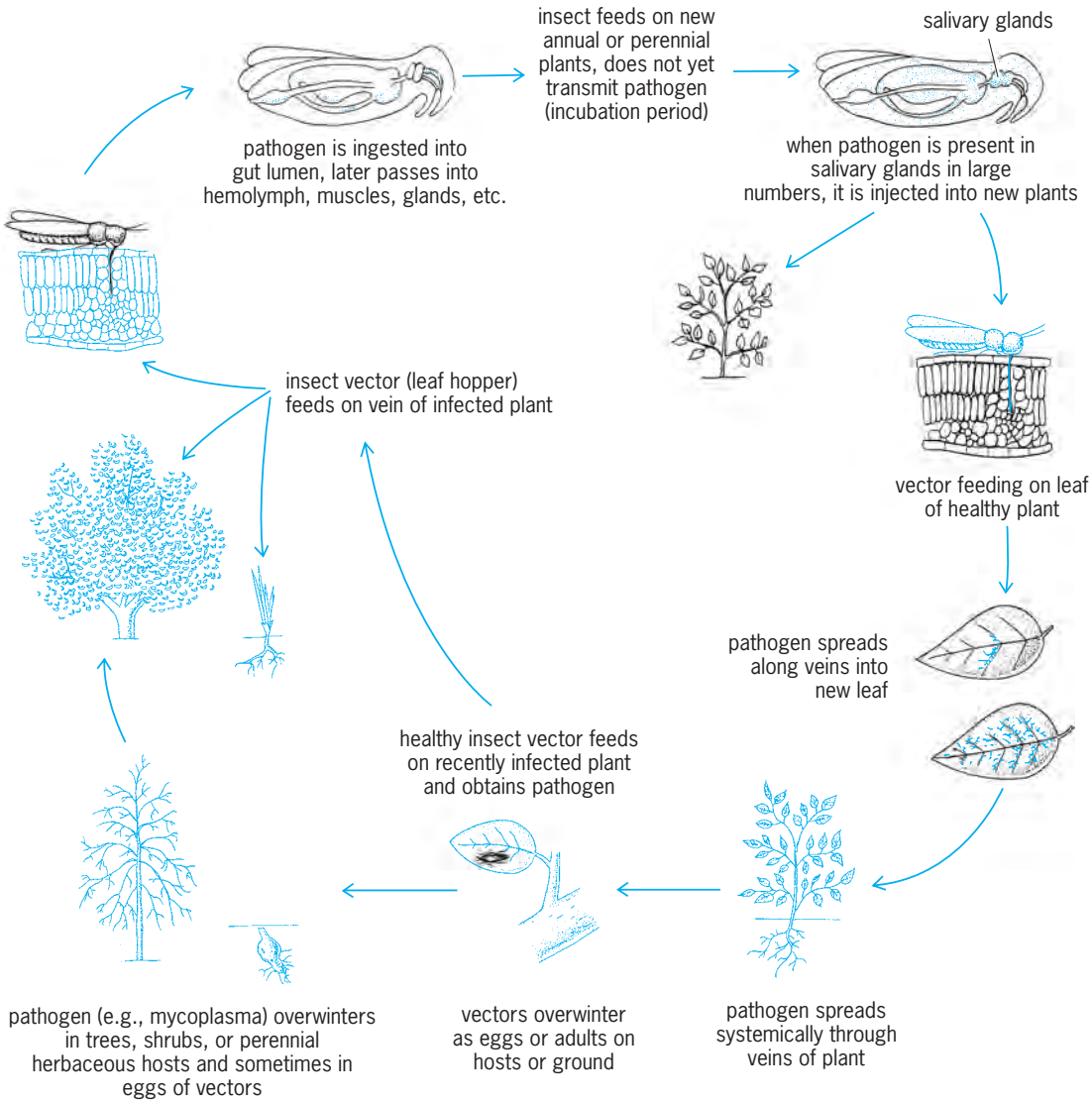

Fig. 15. Development of a typical disease cycle caused by mycoplasma, virus, and so on, transmitted by an insect vector such as a leafhopper. (*After G. N. Agrios, Plant Pathology, 2d ed., Academic Press, 1978*)

control measures can be applied at critical periods, and any yield loss to disease can be projected. The maximum amount of disease occurs when the host plant is susceptible, the pathogen is aggressive, and the environment is favorable.

Epidemiologically, there are two main types of diseases: monocyclic, those that have but a single infection cycle (with the rare possibility of a second or even third cycle) per crop season; and polycyclic, those that have many, overlapping, concatenated cycles of infection per crop season. For both epidemiological types, the increase of disease slows as the proportion of disease, y, approaches saturation or 100% ($y = 1.0$) [**Fig. 18**].

Monocyclic diseases. Fungal root rots, bacterial and fungal wilts, damping-off, basal-stem cankers, and other diseases caused by soil-borne pathogens (including nematodes) usually develop over time with a single infection cycle per crop season. Some foliar diseases that have only one infection cycle per season are also considered as monocyclic.

Epidemic development. The progress of monocyclic diseases when plotted against time resembles the inverted J-shaped curve of normal growth. The rate at which monocyclic diseases develop is determined primarily by the amount of pathogenic inoculum at the beginning of the season, and secondarily by the environmental conditions and host resistance.

The curves of monocyclic diseases are usually fit well by the monomolecular model, $y_t = 1 - \exp(-k_M t)$, where y_t is the proportion of disease at time t, and k_M is the monomolecular epidemic rate. Such curves can be linearized with the transformation $Y = \ln[1/(1 - y)]$. The rate parameter, k_M, can be used for comparative purposes.

Control. For many soil-borne pathogens, resistance of the plants to disease is not known or is incomplete. Therefore, the control of monocyclic diseases consists mostly in the reduction of the inoculum of the pathogen before the crop is planted. Sanitation measures such as burning, burying, or removal of crop debris, crop rotation, soil treatment with

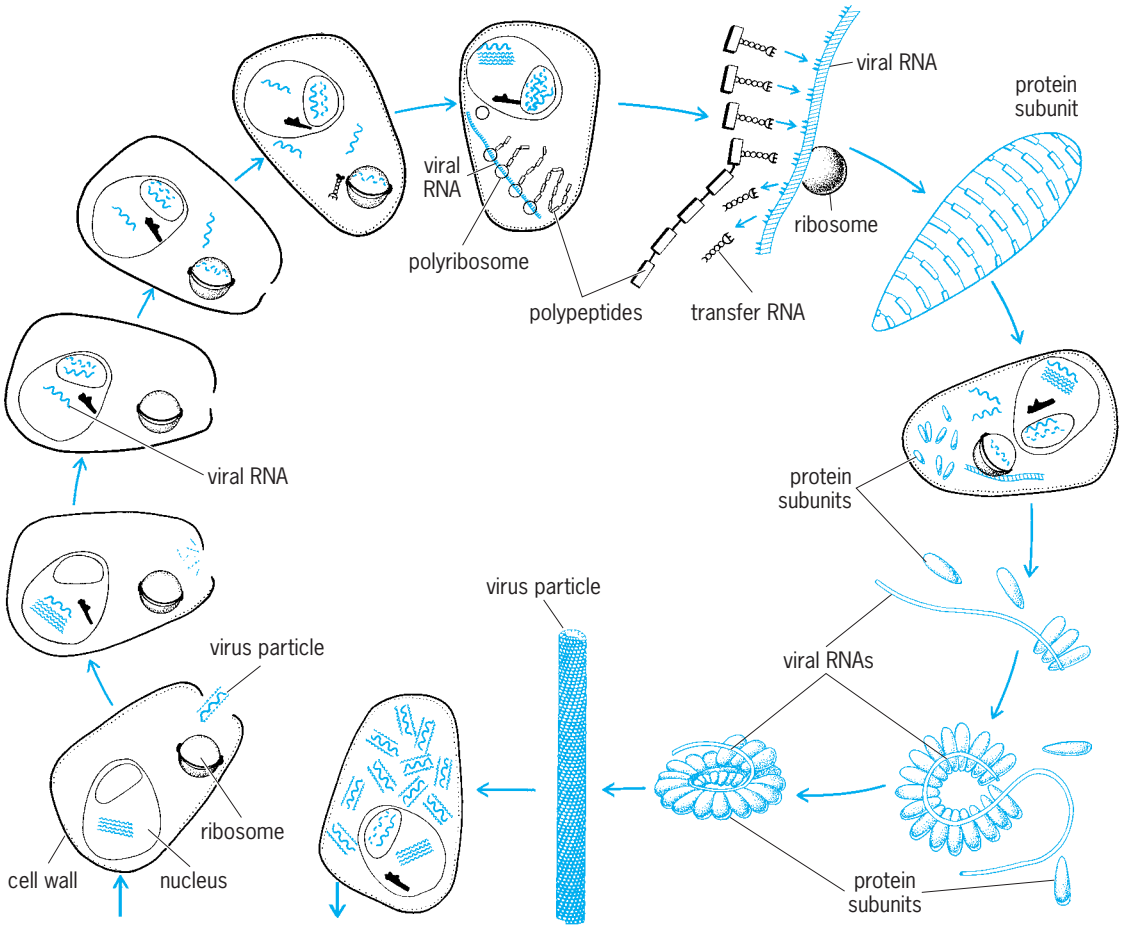

Fig. 16. Sequence of events in the infection of a plant cell by a virus. Two heavy arrows indicate the beginning (left) and the end (right) of the cell infection by the virus. (*After G. N. Agrios, Plant Pathology, 2d ed., Academic Press, 1978*)

mulches, fumigation, sterilization, or solarization are frequently used. Additionally, careful management of the soil moisture, pH, and fertility also reduces initial infections. A reduction in initial inoculum has the additional benefit of slowing the monomolecular rate, k_M. Some postharvest diseases develop monocyclically, and treatment with chlorinated water followed by cool temperatures in transit and storage slow the epidemic rate.

Polycyclic diseases. The rusts, mildews, leaf spots, blights, and other foliar diseases develop polycyclically. The pathogens have multiple infection cycles during a single crop season. Even from a single leaf spot or pustule, the propagules of the pathogen are spread by wind, rain, and vectors (including humans) to neighboring or distant plants. After a latent period that follows infection by the propagules, the pathogen produces inoculum, that is, more propagules, on the newly developed lesions and another cycle begins. These early cycles may produce "waves" of newly appearing disease. Inoculum continues to be borne on the old lesions as well, resulting in lesions of many different ages. Under conditions of favorable environment, infection may occur every day. For some plant pathosystems, such as the rusts, the lesions seldom enlarge in size or they expand only slowly. For other plant pathosystems, such as the blights, there is considerable lesion expansion, sometimes 1–25 mm^2 or more per lesion per day. With these latter pathosystems, the area of tissue involved with lesion expansion far exceeds the area occupied by new lesions. A plant pathosystem with short and frequent cycles of infection coupled with rapid lesion expansion (such as potato late blight) is characterized by a rapidly developing and devastating epidemic.

Epidemic development. The source of inoculum to begin polycyclic epidemics may be from infected seed, infested crop refuse, diseased volunteer plants, weeds, or alternate hosts, or it may come as influx of inoculum or as infested vectors from distant sources. From this beginning, the developmental rate of the polycyclic epidemic depends upon the magnitude of the components of the disease traingle. The plotted curve for the progress of polycyclic epidemics is more or less sigmoidal in shape. For convenience to analyze such curves, epidemiologists linearize them. Symmetrical-sigmoidal curves can be linearized with the logistic transformation, $Y = \ln[y/(1-y)]$. Some asymmetrical-sigmoidal curves are linearized with the Gompertz transformation, $Y = -\ln[-\ln(y)]$, or other appropriate transformation. The logistic model equation to describe symmetrical-sigmoidal curves is $y_t = 1/[1 + b(-k_L t)]$, where y is the proportion of

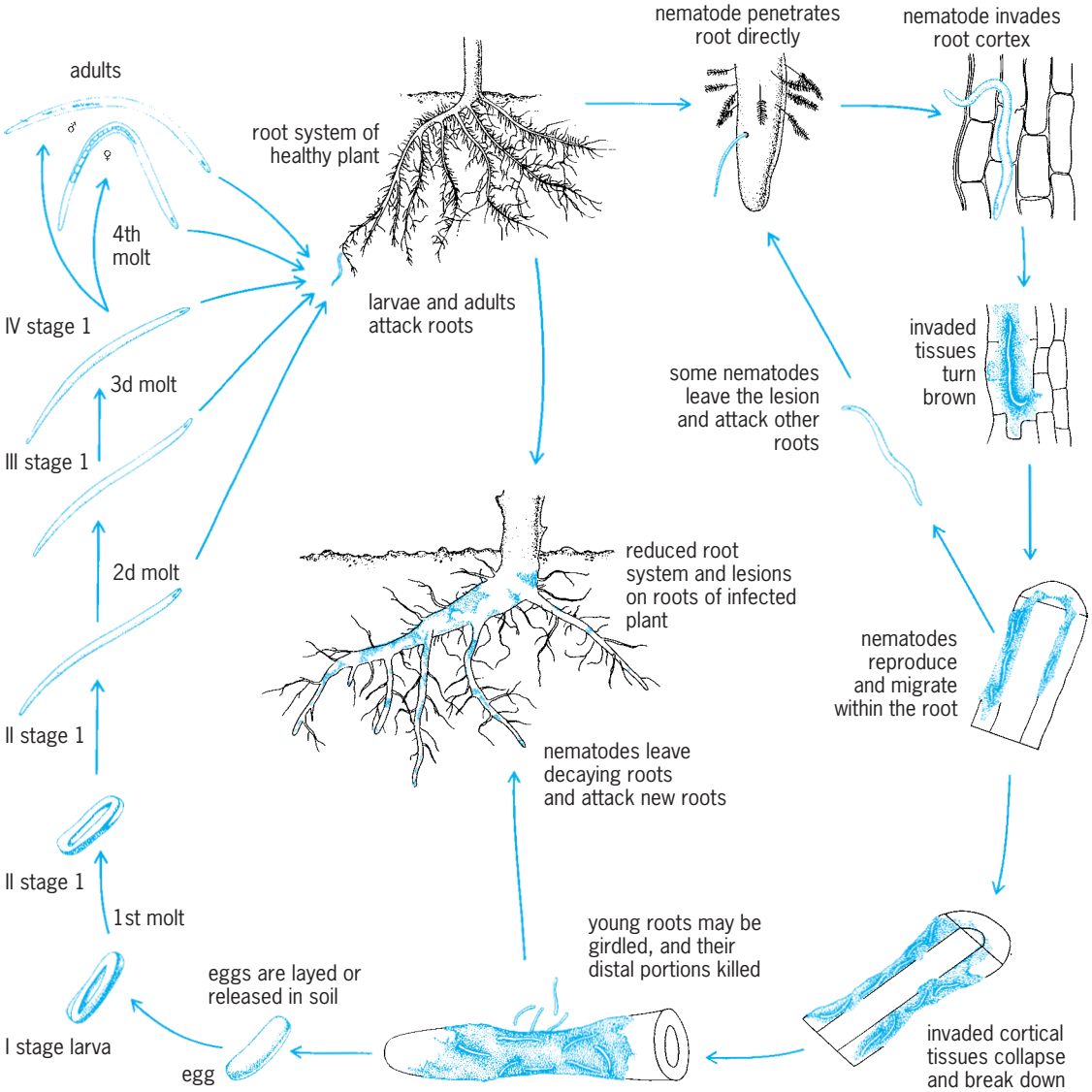

Fig. 17. Development of a typical nematode disease caused by the nematode *Pratylenchus*. (*After G. N. Agrios, Plant Pathology, 2d ed., Academic Press, 1978*)

disease at time t, k_L is the logistic rate parameter, and b is related to the amount of initial disease (y_0) as $b = 1/(1 - y_0)$. The Gompertz model equation to describe asymmetrical-sigmoidal curves is $y_t = \exp(-B \exp(-k_G t))$, where y_t is as before, k_G is the Gompertz rate parameter, and $B = -\ln(y_0)$.

The slope of the line fitted by linear regression through the transformed values of disease proportions is the epidemic rate. This rate parameter (k_L, k_G, etc.) is used to compare epidemics, for example, to determine the reduction in the rate of disease progress because of host resistance or the application of various measures of control. Epidemics that have occurred under different environmental conditions or in different geographical areas would be expected to have different rates of disease progress.

If the sigmoidal curve of disease progress is effectively linearized by one of the various models, disease at any future time can be predicted with reasonable accuracy. This prediction model is expressed as $Y_2 = Y_1 + k\Delta t$; where $\Delta t = (t_2 - t_1)$, Y_2 and Y_1 are the transformed disease proportions y_2 and y_1 for times t_2 and t_1, respectively. This equation is interpreted as disease y_2 at time t_2 is dependent upon current disease y_1 (at time t_1) plus the disease that will occur at epidemic rate k for the period of time between times t_1 and t_2.

Control. The general predictive equation with transformed values, $Y_t = Y_0 + k\Delta t$, can be used to develop strategies for disease control. A reduction in any of the three factors, y_0, k, or Δt, on the right side of the equation would cause lower disease y_t on the left side of the equation at any future point in time. The three epidemiological principles of disease control therefore are (1) to reduce initial inoculum (and initial disease, y_0); (2) to slow the rate (k) of the epidemic; and (3) to shorten the time (Δt) of exposure of the crop to the pathogen (**Fig. 19**). The control of disease would be maximized by a combination of the three principles.

The reduction in the amount of inoculum that would be available to start an epidemic is

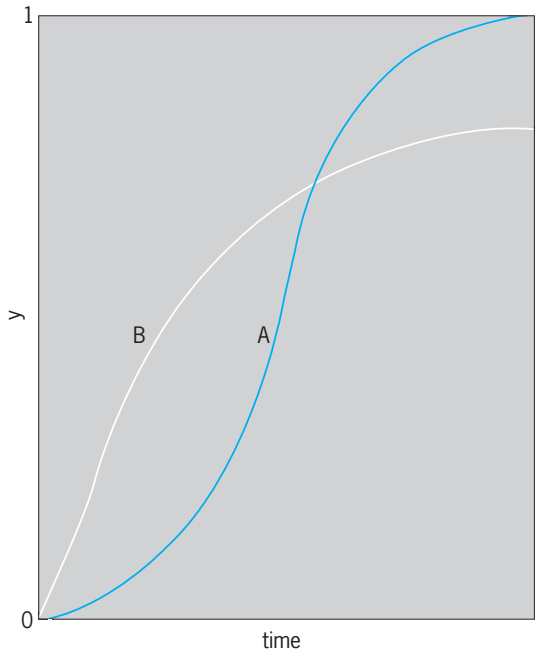

Fig. 18. Increase of proportion of plant disease (*y*) over time for a polycyclic disease (curve A) and a monocyclic disease (curve B).

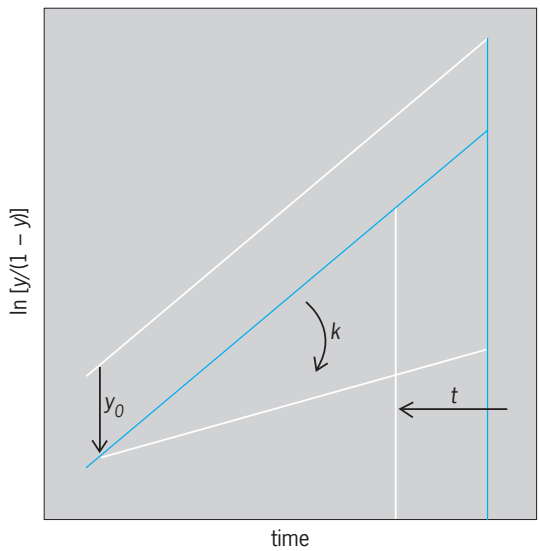

Fig. 19. Diagram of the three epidemiological principles of plant disease control: reduction of initial inoculum (y_0), slowing of epidemic rate (*k*), and shortening time (*t*) of exposure of a crop to the pathogen. Disease (*y*) is expressed as logistic transformation.

accomplished by procedures usually characterized as sanitation. Some of these effective measures are seed treatment with hot water and chemicals, seed indexing, disease-free transplants, removal of infested crop debris and infected volunteer plants, elimination of alternate hosts, and resistance against specific pathogen races.

The rate *k* of the developing epidemic can be slowed by removal of diseased tissue (pruning), spraying with fungicidal or bactericidal chemicals, nonspecific host resistance, barrier or trap crops, vector repellants, avoidance of dispersal of the pathogen, and management of the environment (irrigation, nutrition, wind, light, geographic orientation, and relative humidity). A good example of this management is the intentional thinning of leaves on grapevines to obtain better air movement through the vineyard. This practice will reduce the lengthy periods of high relative humidity that are favorable for infection of grape tissue by the pathogens.

The time of exposure to the pathogen can be shortened by planting short-season varieties, planting vigorous seed and transplants, maintaining optimum fertility and moisture, and controlling insect pests.

If the amount of initial inoculum is reduced to a very low level, then to slow the epidemic rate becomes an easier task. All control practices are more effective when the population of the pathogen is low. Once the proportion of visible disease approaches 0.03–0.05 (3–5%), the epidemics of most pathosystems become nearly impossible to control and some yield loss will likely occur. If the latent period of a pathogen is 7 days and the epidemic is proceeding at a conservative epidemic rate, say $k_L = 0.15$, then the 3–5% disease would increase to 8–15% disease in 7 days, solely from the infections that were latent.

Spatial components. In addition to the increase of disease over time, the diseases spread over distance. Frequently, disease begins as one or more distinct foci in the crop area. In time, the individual foci may overlap and their discreteness may be lost. A gradient of disease intensity exists from the focal center to the periphery. The steepness of the gradient is largely dependent upon the dispersal characteristics of the pathogen; that is, rain-splash dispersal of propagules results in a steep gradient (considerable change in disease intensity over a short distance), whereas wind dispersal of propagules results in a flat gradient (**Fig. 20**). The J-shaped curve of disease intensity over distance can usually be linearized by $\log_{10}(y)$ or $\ln[y/(1-y)]$ versus $\log_{10}(d)$, where $d =$ distance.

The spatial pattern of disease plants received considerable attention in the 1990s. The aggregation (or randomness) of diseased plants is interpreted by the epidemiologist to determine modes of dispersal of inoculum, to determine efficiency and movement of vectors, and to locate initial sources of inoculum. Some of the analytical methods employed are doublet analysis, ordinary runs, spatial autolag correlation, geostatistics, the fitting of disease incidences to various distance or cluster models, dispersion indices, and variance-to-mean ratio.

Disease estimation. The assessment of disease in time and space is accomplished by visual estimation of disease intensity. Plant pathologists use percentage keys, disease rating scales, pictorial keys, remote sensing (color or infrared photography), the digitization of color images of the crop, and the reflected radiation off the crop canopy. Sometimes the spores of fungal pathogens are monitored by volumetric air samplers. The populations of bacterial pathogens are

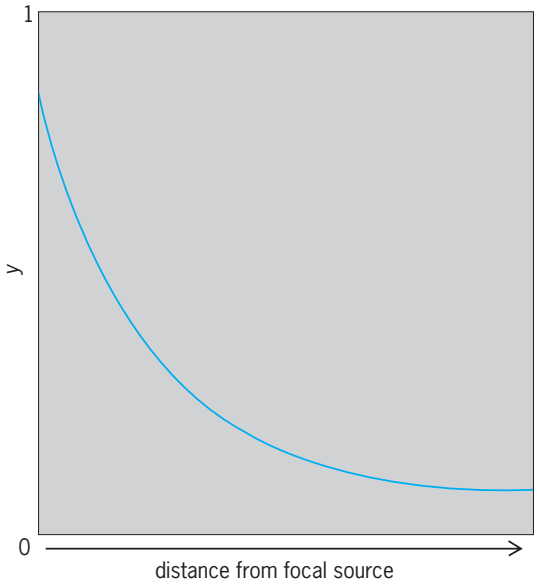

Fig. 20. Decreasing intensity of disease with increasing distance from source. Steepness of the gradient depends on dispersal characteristics of pathogen.

monitored by washing the bacterial cells off leaf surfaces and culturing on selective media. With these techniques, the buildup of pathogenic populations can be quantified and disease-favorable weather periods can be determined. From these estimates of disease and pathogenic propagules, the epidemic can be followed in its entirety.

Crop loss. The loss of yield to disease has been the primary reason to study plant pathology. In the past, plant pathologists have made the correlations of crop loss to some measure of disease intensity, for example, to the amount of disease at a certain crop stage, the maximum amount of disease observed, and the area-under-the-disease-progress curve. Sometimes these correlations are excellent for a given crop, but they have had very little predictive value because disease in another season may occur earlier or later, the growth of the crop will be different among seasons, considerable difference in environmental conditions will occur, and so on. A more recent approach is to correlate yield to the duration of the area of healthy leaves (HAD) and the amount of radiation intercepted by this area (HAA). The theory is that yield is determined by this healthy leaf area, rather than that loss is determined by the amount of disease. The HAD/HAA concept has been found to be a generally reliable determinant of crop yield for many plant pathosystems. The photosynthetic competence of leaves is well correlated to leaf color; therefore, crop yield in the future may be predicted from satellite-collected color images of crops.

Epidemic modeling. Models have been developed for many of the important plant pathosystems. These models may be simple in design to forecast likely disease occurrence based on certain weather criteria. Simulation models may be quite complex in which the various components of the pathogen's life cycle (inoculum production, dispersal, catch, infection, latent growth, and lesion expansion) are quantified (**Fig. 21**). Plant growth is carefully determined and mathematically described. By varying the parameters of the model, the experimenter can determine vulnerable places in the cycle where control measures would be most effective. Epidemic progress brought about by differences in environment, plant resistance, or control procedures can be characterized. The simulation exercise could lead to more effective strategies of disease control. The simulation models can also be used to examine epidemiological theories.

Emerging epidemics. Toward the end of the twentieth century, several plant disease epidemics emerged or reemerged to cause concern among the agriculturists of many continents. A new strain of potato late blight appeared that was resistant to the fungicide (metalaxyl) most commonly used for its control. This new race had other important features: with different mating types, there was the opportunity to have sexual recombination which could lead to more new races; with this sexual compatibility, more sexually derived oospores would be available for overseasoning to initiate new epidemics in succeeding crops. Additionally, this new strain was more aggressive pathogenically in that the haulms were attacked severely, resulting in more rapid plant death and more infection of the underground tubers. There was considerable controversy over the origin of this new strain. Nonetheless, it has spread rapidly to attack crops in most of the potato-growing regions of the world.

In the cereal-growing regions of the northern United States, epidemics of a fungus disease, fusarium head-blight of wheat and barley, have become severe, causing considerable losses in crop yield and grain quality. The agronomic change to no or low tillage over the past decades may have been largely responsible for the widespread occurrence of this disease.

In the citrus-growing regions of South America and Florida, epidemics of bacterial canker have been of great concern. In Florida, past epidemics have been controlled by the broad-scale eradication of infected and neighboring trees. However, outbreaks of bacterial canker keep reappearing in commercial groves and in dooryard citrus trees in urban Miami. In South America, canker is so prevalent in Argentina, Uruguay, Paraguay, and some states in Brazil that eradication is not an economical option. The departure from eradication as a control measure may occur in the future for those areas currently under threat.

Another epidemic of current interest is black sigatoka of banana. This disease is more difficult to control than gray sigatoka. Black sigatoka was found in the late 1990s at the edge of the tropical, banana-growing region of Brazil, and its presence is a severe threat to that crop. Another disease, karnel bunt of cereals, has been found occasionally in shipments of

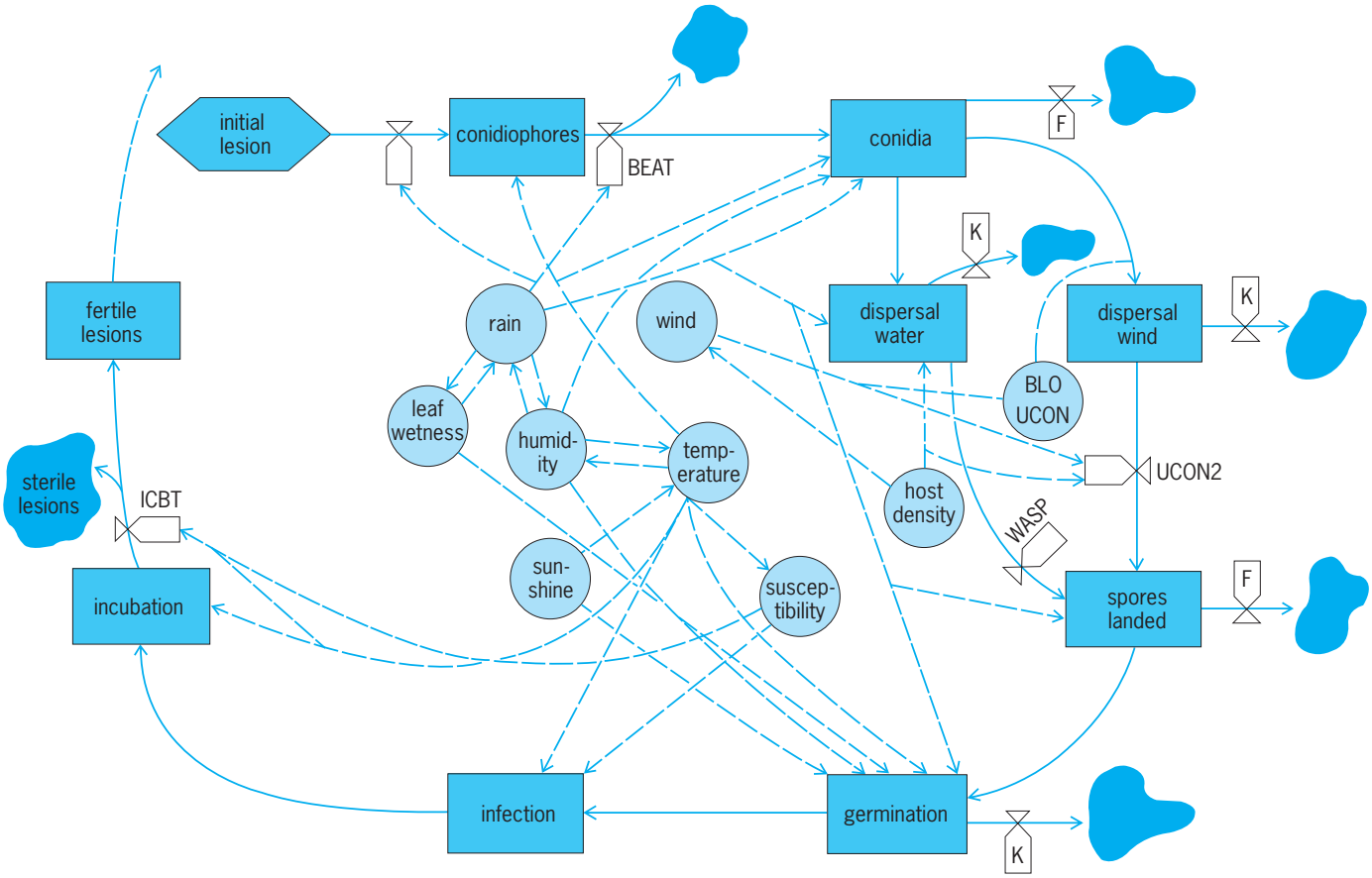

Fig. 21. Flow diagram for a plant disease simulator. The stages in the pathogen's life cycle are quantified with respect to environmental parameters. (*After J. Kranz et al., Z. Pflanzenkr., 80:181–187, 1973*)

grain to the United States. This pathogen is under strict quarantine in the United States, as an introduction of this fungal pathogen into the grain-growing regions may cause considerable losses. R. D. Berger

Control of Plant Disease

Control of plant disease is defined as the maintenance of disease severity below a certain threshold, which is determined by economic losses. Diseases may be high in incidence but low in severity, or low in incidence but high in severity, and are kept in check by preventing the development of epidemics. The principles of plant disease control form the basis for preventing epidemics. However, the practicing agriculturist uses three approaches to the control of plant disease: cultural practices, disease resistance, and chemical pesticides.

Cultural practices. Cultural practices affect the environmental requirement of the suscept-pathogen-environment triangle that is necessary for plant disease development. Changes in the atmospheric, edaphic, or biological environment favorable to the plant and unfavorable to the pathogen will result in less disease. Important atmospheric parameters include temperature, humidity, light, and wind; edaphic parameters are physical and chemical properties of soils, soil reaction (pH), temperature, and moisture; and the biological parameters are all of the organisms that affect plant and pathogen survival and growth. The environment that is important is the small volume near the plant surface, termed the microclimate, which may be much different than that measured at a weather station. Cultural practices may be very effective in altering a plant's microclimate, but generally have no effect on the climate of an area.

Crop rotation, crop location, mineral nutrition, cultivation, and irrigation practices are a few cultural practices that influence disease development. Each may exert its influence on a specific phase of the pathogen life cycle. Crop rotation may influence overwintering or survival of pathogens between crops, while crop location may be important for reducing dissemination of pathogen propagules to young plants. Land preparation may affect the production and liberation of inoculum. Mineral nutrition may reduce the amount of colonization by the pathogen once inside the plant, which in turn may reduce the amount of inoculum produced. Proper cultivation may reduce wounding and thus avenues of entrance for pathogens. Proper irrigation may reduce successful germination of fungal spores and penetration of plants by parasites by reducing the wet period of foliage and soils. Successful use of cultural

practices in disease control depends upon knowledge of how each practice affects the pathogen or plant. *See* AGRICULTURAL SOIL AND CROP PRACTICES.

Disease resistance. Resistant crop plant varieties have been used extensively for the control of plant diseases. For the farmer, the planting of a resistant cultivar is the easiest and often the most effective means of disease control. Resistant cultivars of many crops have been developed.

Resistant plants occur naturally in populations of crop species, but cannot be detected unless an epidemic occurs. One source of resistant plants is the large population of commercially grown plants that are subjected to natural epidemic. This source is the most desirable because selected plants contain advantageous yield and quality characteristics. The other main source of resistant plants is among natural populations. Progeny of resistant plants multiply more rapidly than progeny of susceptible plants in the wild, if a pathogenic organism is present. Most resistant plants come from this second source. The U.S. Department of Agriculture maintains a seed bank of wild plants of crop species collected from around the world.

Progeny of resistant plants and crosses of resistant plants with high-yield, high-quality susceptible plants must be screened before selections can again be made. For this screening, artificial epidemics are created in glasshouses or in the field, from which surviving plants are again crossed with commercial cultivars. Progeny are screened at each generation, and crosses are continued until a high-yield, high-quality resistant plant is obtained.

Two types of resistant plants may be selected from a population of a crop species. One type is plants with a very high level of resistance, determined by one or a few genes. This type is easily selected among progeny, but is vulnerable to development of more virulent strains of the pathogen. The second type of plants exhibit resistance at a lower level determined by several genes. However, once established in a high-yield, high-quality cultivar, this resistance tends to be durable to the development of more virulent strains of the pathogen.

A high level of resistance acts epidemiologically to reduce the initial inoculum, because it reduces the population of the pathogen that will cause disease to be near zero. The lower level of resistance acts epidemiologically to reduce the rate of disease progress, thus reducing the percentage of disease in each pathogen cycle. *See* BREEDING (PLANT); GENETIC ENGINEERING.

Chemical pesticides. Chemicals that are used to control plant diseases are used either as seed treatments, soil treatments, or protective sprays to plants. Although each use is important for plant disease control, the greatest use of chemicals is, by far, as protectants.

Seed treatment chemicals are used to control pathogenic organisms in and on seeds before they are planted, and around seeds when they are planted in soils. There is no chemical that will control pathogens at all locations, since not all chemicals enter seeds and none is effective against all organisms. Thiram and captan are the most common fungicides for treatment of fungi on seed and around seed in the soil. Sodium hypochlorite is used to control bacteria on seeds. Oxathiins are used on cereal seeds to control the smut fungi that occur in the seeds. Hot water treatments may be used to kill pathogens inside some seeds.

Certain volatile chemicals, such as methyl bromide, used to treat soils for eradication of pathogens have a broad spectrum of activity and kill all plant and animal life in the soil. Other volatile compounds may kill only animal life and thus control plant pathogenic nematodes—a mixture of dichloropropane and dichloropropene is such a nematocide. Many of these volatile compounds are toxic to crop plants, however, and treatment must be made before planting. Enough time must elapse between treatment and planting for fumigants to leave the soil.

Most chemicals applied to plants as protectants prevent germination of spores and invasion by pathogenic organisms since once the pathogen is inside the plant the chemicals are ineffective. Protective chemicals usually have very low solubility in water so they are not washed completely from plants by rains. Nevertheless, several applications must be made to keep new growth covered and to replace chemicals that may be inactivated by weathering processes.

The spectrum of pathogens controlled by protective sprays varies widely. For example, fixed copper compounds have broad activity and control most fungi and bacteria. Maneb controls most fungi, but is ineffective against bacteria. Benomyl controls only a few fungi. Level of control of a specific fungus may also vary. Copper, maneb, and benomyl all control *Cercospora apii*, the pathogen of early blight of celery, but benomyl provides the best control. Thus, specific compounds must sometimes be used for best control.

Chemical pesticides are strictly regulated to ensure against improper use. Pesticides must be registered with the federal government and must contain information on the label for safe use of the product. Chemicals must be tested for safety, and tolerances are set for limits of all pesticides in food products. *See* PESTICIDE.

Robert E. Stall

Bibliography. G. N. Agrios, *Plant Pathology*, 3d ed., 1988; G. W. Bruehl, *Soilborne Plant Pathology*, 1986; O. D. Dhingra and J. B. Sinclair, *Basic Plant Pathology Methods*, 2d ed., 1994; C. B. Kenaga, *Principles of Phytopathology*, 2d ed., 1986; G. B. Lucas, C. L. Campbell, and L. T. Lucas, *Introduction to Plant Diseases: Identification and Management*, 2d ed., 1992; O. C. Maloy, *Plant Disease Control: Principles and Practices*, 1993; D. C. Sigee, *Bacterial Plant Pathology: Cell and Molecular Aspects*, 1993; J. A. von Arx, *Plant Pathogenic Fungi*, 1993; B. M. Zuckerman, W. F. Mai, and R. A. Rohde (eds.), *Plant Parasitic Nematodes*, vols. 1 and 2, 1971.

Plant phylogeny

The evolutionary chronicle of plant life on the Earth. Understanding of this history is largely based on knowledge of extant plants, but the fossil record is playing an increasingly important role in refining and illuminating this picture. Study of deoxyribonucleic acid (DNA) sequences has also been revolutionizing this process in recent years. In spite of a remarkably improved situation, controversy over interpretation of the available data can be expected to continue. The molecular data (largely in the form of DNA sequences from several genes) summarized below have been demonstrated to be highly correlated with other information. *See* PHYLOGENY.

Algae. "Algae" was once a taxonomic designation uniting the lower photosynthetic organisms, but ultrastructural and molecular data have uncovered a bewildering diversity of species. Algae are now recognized as 10 divergent lineages on the tree of life that join organisms as distinct as bacteria and eukaryotic protozoans, ciliates, fungi, and embryophytes (including the land plants). In a biochemical context, the term "algae" defines species characterized by chlorophyll *a* photosynthesis (except Embryophyta); some of their descendants are heterotrophic (secondary chloroplast loss). Despite the variety of species it encompasses, the term "algae" also retains phylogenetic relevance.

Cyanophyta (blue-green algae or bacteria) include the prokaryotic algae and have a fossil record dating back 3 billion years. Cyanophytes flourished on the early Earth, their chlorophyll *a* photosynthesis increasing atmospheric oxygen and ultimately favoring eukaryote evolution. Eukaryotes differ fundamentally from prokaryotes by possessing a nuclear membrane and an endomembrane system that probably evolved by internalization of plasma membrane invaginations. Evolution of the eukaryotic cytoskeleton was another key event and enabled new types of movement, including phagotrophy (ability to engulf particles into food vacuoles). Phagotrophy had a profound impact on eukaryote evolution because entire prokaryotes could be ingested, or they could remain metabolically active in the host's food vacuole, leading to the establishment of an endosymbiosis (symbiosis in which one species lives inside another). Substantial evidence indicates that eukaryotic mitochondria and plastids evolved from such primary endosymbiont events. All eukaryotic plastids probably derive from a single primary endosymbiont event between an early eukaryote and a cyanophyte (producing the ancestral phototrophic eukaryote) and, therefore, cyanophytes and plastids share a common ancestor. Hence there is a paradox: algal plastids are closely related, but the host eukaryote lineages are not. The paradox is resolved by the secondary endosymbiont hypothesis, which concludes that six of the nine eukaryotic algal lineages have acquired plastids by engulfing and establishing endosymbioses with phototrophic eukaryotes. "Algae" in a phylogenetic context, therefore, refers to Cyanophyta and the diverse eukaryote lineages that have received their descendent plastids by vertical and horizontal transmission. *See* CHLOROPHYLL; PHOTOSYNTHESIS.

Three eukaryotic algal lineages (Glaucophyta, Chlorophyta, and Rhodophyta) have hypothetically evolved directly from the ancestral phototrophic eukaryote. All have plastids surrounded by two membranes: the inner derived from the cyanophyte plasma membrane; the outer, the plastid envelope, from the eukaryotic food vacuole membrane. Glaucophyta are a division of considerable evolutionary interest because their plastids retain many features of the ancestral cyanophyte. In fact, glaucophytes were considered to be eukaryotes with endosymbiotic cyanophytes living in vacuoles until recent evidence indicated that these structures are true plastids. If there was a single primary endosymbiont event for eukaryotic plastids, Glaucophyta must be related to Rhodophyta (red algae) and Chlorophyta (green algae, here including Charophyceae) that have also evolved from the ancestral phototrophic eukaryote. There is no evidence for such an association; Glaucophyta may have derived plastids from an independent primary endosymbiont event. Furthermore, evidence for a relationship between Rhodophyta and Chlorophyta is weak, and the hypothesis of a single primary endosymbiont event generating eukaryotic plastids remains tentative. Both Rhodophyta and Chlorophyta have additional phylogenetic significance. Rhodophyta completely lack flagella in all life history stages; and this distinction, as well as similarities between rhodophyte chloroplasts and cyanophytes, prompted the proposal that Rhodophyta were the earliest eukaryotes. This view has been rejected by contemporary systematic data. Chlorophyta is the division from which the embryophytes (including the land plants) evolved, and some workers recognize a division, Streptophyta, for embryophytes and their specific green algal relatives. *See* CYANOPHYCEAE; RHODOPHYCEAE.

Eukaryotic lineages derived from secondary endosymbiont events have four (sometimes three) plastid membranes (inner two from the plastid envelope and the third from the plasma membrane of the phototrophic eukaryote, the fourth from the food vacuole membrane of the secondary eukaryote host). Two divisions, Cryptophyta and Chlorarachniophyta, are of paramount phylogenetic importance because of a miniature nucleus (nucleomorph) in the compartment (derived from cytoplasm of the phototrophic endosymbiont) between the inner two and outer two plastid membranes. Cryptophyta are a lineage of flagellate species of unresolved phylogenetic position on the tree of life that nevertheless shares many plastid features with Rhodophyta. Although the host cryptophyte is not related to rhodophytes, the nucleomorph DNA is clearly of red algal nuclear origin, betraying the phylogenetic origin of cryptomonad plastids and

supporting the secondary endosymbiont hypothesis. Similarly, Chlorarachniophyta host cells are related to filose amebas, but their plastids are unequivocally of chlorophyte origin. Euglenophyta also have chlorophyte plastids that are considered to be derived by secondary endosymbiosis despite absence of a telltale nucleomorph. The plastids are so similar to the chlorophyte counterpart that euglenoids were included in Chlorophyta by some workers, but recent molecular data indicate that the host is related to a line of eukaryotic flagellates (Kinetoplastida). *See* CRYPTOPHYCEAE.

The final three lineages are loosely the chromophytic algae but, despite plastid similarities, are not close relatives and have probably derived their chloroplasts by independent secondary endosymbiosis events with related phototrophic eukaryotes. Contemporary phycologists lack consensus on a taxonomic label for the descriptively named heterokont chromophytes. This diverse assemblage includes the well-known Bacillariophyceae (diatoms) and Phaeophyceae (brown algae), to name but two of the many classes. These algae are now included in the Heterokonta (Stramenopila or Chromista) with exfungal lineages (such as Oomycota) and certain heterotrophic flagellates (Bicosoecida). These diverse organisms are united by a common flagellar arrangement (heterokont) in which the anterior flagellum has two rows of stiff tripartite hairs, whereas the posterior flagellum is smooth. Molecular data substantiate this feature for uniting this bewildering array of species. *See* BACILLARIOPHYCEAE; CHRYSOPHYCEAE; PHAEOPHYCEAE.

Haptophyta (Prymnesiophyta) is a group of unresolved phylogenetic affinity from the eukaryote crown on the tree of life. Although similar to the heterokont chromophytes in plastid attributes, they differ in flagellar arrangement and have an additional appendage (haptonema) unique to this group. *See* PRYMNESIOPHYCEAE.

Dinophyta also have a characteristic flagellar arrangement (dinokont), as well as plastid and pigment attributes, that distinguish them from other chromophytic algae. Dinoflagellates are close relatives of flagellate Apicomplexa and Ciliata in sharp contrast to earlier views. As an example, the mesokaryote hypothesis argued that dinoflagellates, with permanently condensed chromosomes that lack histones, were intermediate on the evolutionary path from prokaryotes to eukaryotes—a hypothesis since rejected. *See* CHROMOPHYCOTA; DINOPHYCEAE.

It is evident from the previous text that the Algae, a group of species linked phylogenetically through their plastid ancestry, represents a diverse assemblage of species branching from all points on the tree of life. Their importance to understanding eukaryotic evolution cannot be overstated, ensuring them a place of prominence in the continuing realm of phylogenetic investigation.

Gary W. Saunders

Embryobionta. Embryobionta, or embryophytes, are largely composed of the land plants that appear to have emerged 475 million years ago. The evidence indicates that land plants have not evolved from different groups of green algae (Chlorophyta) as suggested in the past, but instead share a common ancestor, which was a green alga. Land plants all have adaptations to the terrestrial environment, including an alternation of generations (sporophyte or diploid and gametophyte or haploid) with the sporophyte generation producing haploid spores that are capable of resisting desiccation and dispersing widely, a cuticle covering their outside surfaces, and separate male and female reproductive organs in the gametophyte stage. The life history strategies of land plants fall into two categories that do not reflect their phylogenetic relationships. The mosses, hornworts, and liverworts represent the first type, and they have expanded the haploid generation, upon which the sporophyte is dependent. Several recent analyses of DNA data as well as evidence from mitochondrial DNA structure have demonstrated that the liverworts alone are the remnants of the earliest land plants and that the mosses and hornworts are closer to the vascular plants (tracheophytes). The tracheophytes include a large number of extinct and relatively simple taxa, such as the rhiniophytes and horneophytes known only as Silurian and Devonian fossils. All tracheophytes are of the second category, and they have expanded the sporophyte generation. Among extant tracheophytes, the earliest branching are the lycopods or club mosses (*Lycopodium* and *Selaginella*), and there are still a diversity of other forms, including sphenophytes (horsetails, *Equisetum*) and ferns (a large and diverse group in which the positions of several families still are not clear). Within the tracheophytes, many groups have a much reduced but still free-living gametophyte and a dominant sporophyte. The sporophytes bear the haploid spores in sporangia that may be lateral (lycopods and psilophytes), terminal (horsetails and some ferns), or on the lower surfaces of leaves (most ferns). *See* EMBRYOBIONTA.

All seed plants take the reduction of the gametophyte generation a step further and make it dependent on the sporophyte, typically hiding it within reproductive structures, which are either cones or flowers. The first seed plants originated at least by the Devonian, and they are known to have a great diversity of extinct forms, including the seed ferns. There are two groups of extant seed-bearing plants, gymnosperms and angiosperms. In the gymnosperms, the seeds are not enclosed within tissue derived from the parent plant. There are four distinct groups of extant gymnosperms, often recognized as classes: Cycadopsida, Gnetopsida, Ginkgoopsida, and Pinopsida. For many years, the gnetophytes were considered to be the closest extant relatives of the angiosperms, but the accumulating molecular evidence indicates that they may instead be members of the gymnosperm group, which collectively are the closest relatives of the angiosperms.

The gnetophytes are composed of three families,

Ephedraceae (only the genus *Ephedra*, the source of ephedrine), Gnetaceae (*Gnetum*), and Welwitschiaceae (*Welwitschia*, known from southwestern Africa, and one of the most bizarre plants in existence). The cycads, which look like large ferns or palms, are not at all related to either; many are commonly grown ornamentals (sago palm), and all produce large conelike and sometimes leafy reproductive structures with large seeds. *Ginkgo* (the maidenhair tree) is often grown as a street tree. The most important of the gymnosperms, both ecologically and commercially (seeds eaten as food and timber production), are the conifers (Pinopsida), which include the pine (*Pinus*), spruce (*Picea*), fir (*Abies*), yew (*Taxus*), monkey puzzle (*Araucaria*), and cedar (*Cedrus* and *Juniperus*).

The angiosperms (also flowering plants or Magnoliopsida) are the dominant terrestrial plants, although the algae collectively must still be acknowledged as the most important in the maintenance of the Earth's ecological balance (fixation of carbon dioxide and production of oxygen). In angiosperms the seeds are covered by protective tissues derived from the parental plant. The nature of the ancestors of the flowering plants is still the object of a great deal of speculation, and a great deal more research will be required before their evolutionary history can be elucidated. Clearly, there were a large number of extinct lineages of seed plants, but which of these might have been close to the lineage that ultimately led to the angiosperms is highly uncertain. There are no generally accepted angiospermous fossils older than 120 million years, but the lineage is clearly much older based on DNA clocks and other circumstantial lines of evidence such as their current geographic distributions.

Traditionally the angiosperms have been divided into two groups, monocotyledons (monocots) and dicotyledons (dicots), based on the number of seed leaves. However, DNA sequence data have demonstrated that, although there are two groups, these are characterized by fundamentally different pollen organization, such that the monocots share with a group of dicots pollen with one pore whereas the rest of the dicots have pollen with three (or more) pores.

This difference in pollen organization places the monocots within the magnoliids or primitive dicots such as Laurales (avocado, cinnamon, and their relatives), Magnoliales (magnolias), and Piperales (black pepper, peperomias, and pipeworts). In general, magnoliids lack well-organized flowers, such that the number and orientation of parts are highly irregular, and these flowers are pollinated by animals with simple behavorial patterns, such as flies and beetles. The exceptions to this pattern are the more derivative groups of monocots, such as the orchids and gingers, in which the flowers are highly organized and pollinated by behaviorally complex animals (bees, butterflies, and mammals).

Of those groups with uniaperturate pollen, the monocots are by far the most numerous and important ecologically and economically (all the major grains). The most primitive monocots are *Acorus* (Acorales) and alismatids (pondweeds) and aroids (both Alismatales), all of which are predominantly wetland or aquatic plants. The lilioid or petaloid monocots are composed of generally the showiest families and include the lilies (Liliales), orchids and irises (both Asparagales), and yams (Dioscoreales). The most derivative group of monocots includes the palms (Arecales), gingers (Zingiberales), grasses and their relatives such as the sedges and rushes (Poales), and spiderworts, bloodroots, and pickerel weeds (Commelinales).

The triaperturate angiosperms or eudicots are more numerous than the magnoliids, and the details of their phylogenetic relationships, like those of the magnoliids, have only recently been resolved by the collection of large numbers of DNA sequences. The first branching groups include the buttercups and poppies (Ranunculales), lotuses, plane tree, and proteas (all Proteales), boxwoods (Buxaceae), and gunneras (Gunneraceae), which, like the magnoliids, generally lack well-organized flowers. There are four highly derivative eudicot lineages: the relatively small sandalwood and mistletoe families (Santalales), the much larger and more important caryophyllids (including some carnivorous families such as the sundews and tropical pitcher plants, as well as the amaranths, cacti, and carnations), and then the two largest sets of orders, the rosids and asterids. As opposed to the magnoliids and early branching eudicots, all four of these most derived groups have flowers arranged in well-defined whorls of largely definite numbers of parts. The rosids typically have two whorls of stamens, although some have amplified the numbers in each position, with unfused petals. The asterids have a single whorl of stamens fused to the fused petals, although some also have more stamens and may appear to have free petals due to extensive elongation of the petal apices late in floral development. The rosids include the familiar members of the rose family (apples, peaches, strawberries; Rosaceae of Rosales), euphorbs (Euphorbiaceae of Malpighiales), beans (Fabaceae of Fabales), beeches, oaks and birches (several families in Fagales), mustards (Brassicaceae of Brassicales), cotton and hibiscus (Malvaceae of Malvales), and maples and citrus (Sapindaceae and Rutaceae of Sapindales). Asterid orders and families include rhododendrons and ebenies (Ericaceae and Ebenaceae of Ebenales), dogwoods and hydrangeas (Cornaceae and Hydrangeaceae of Cornales), gentians and coffee (Gentianceae and Rubiaceae of Gentianales), mints and snapdragons (Lamiaceae and Scrophulariaceae of Lamiales), daisies and bellflowers (Asteraceae and Campanulaceae of Asterales), and carrots and ivies (Apiaceae and Araliaceae of Apiales).

Mark W. Chase

Bibliography. D. Bhattacharya (ed.), *Origin of Algae and Their Plastids*, Springer-Verlag, 1997; L. E. Graham and L. W. Wilcox, *Algae*, Prentice Hall, 2000.

Plant physiology

The branch of plant sciences that aims to understand how plants live, function, and grow. Its ultimate objective is to explain all life processes of plants by a minimal number of comprehensive principles founded in chemistry, physics, and mathematics. *See* BOTANY.

Breadth of branch. Organisms are distinguished from the inanimate world by three characteristics. First, they take up materials—water and inorganic and organic chemical compounds—from the environment and process them in various ways, usually producing more complex organic compounds. Second, while they may release part of the absorbed materials into the environment in the original or a modified form, they incorporate a major part into their bodies. This incorporation can be in the form of an increase in the size of the existing body and its parts, or of addition of more of the same structures that are already present in the organism (for example, more leaves, shoots, and roots); or the incorporation can be in the form of production of new or at least greatly modified structures (for example, flowers and fruits, or tubers). The former are quantitative changes and are known as growth; the latter are qualitative ones and are known as development. Third, organisms use part of the newly acquired substance for reproduction, that is, to produce offspring identical to the parent organisms.

In addition to these three, somewhat different but no less characteristic and important property of organisms is their continuous response to the changing environment. Plants, except for some relatively simple ones and for reproductive cells of some complex plants, spend their entire life at a single site and cannot escape their environment. Thus, interaction with both the physical and biotic environment is essential and elaborate. It endows plants with a considerable measure of adaptability, that is, ability to adjust to adverse environmental conditions, such as freezing temperatures in winter or periods of high temperatures and lack of water in summer, and enables them to use environmental signals to adjust their entire life to the given environment. For example, the form of growth will be altered by the light environment, including the intensity, the quality (or the color spectrum), and the duration of the light. Essential developmental processes, like flowering, may similarly be induced or delayed by certain environmental conditions. Plants may change their cellular chemistry to resist forthcoming adverse periods, such as freezing temperatures, with such changes initiated in response to environmental signals, such as declining daylengths and declining temperatures. Plants are also a factory for myriad chemicals which we recognize as scents and tastes. Most of these chemicals are part of the plant's defense against disease and predation, but some also serve as attractants for other organisms that aid in the plant's growth, survival, or reproduction. *See* COLD HARDINESS (PLANT); PHOTOMORPHOGENESIS; PHOTOPERIODISM.

Three major parts. Plant physiology seeks to understand all the aspects and manifestations of plant life summarized in the preceding section. In agreement with the major characteristics of organisms, it is usually divided into three major parts: (1) the physiology of nutrition and metabolism, which deals with the uptake, transformation, and release of materials, and also their movement within and between the cells and organs of the plant, which may be subdivided into the processes involved in metabolism or transport; (2) the physiology of growth, development, and reproduction; and (3) environmental physiology, which seeks to understand the manifold responses of plants to the physical and biotic environment (**Table 1**). The part of environmental physiology that deals with effects of and adaptations to adverse conditions is referred to as stress physiology. The interactions between plants and insects, fungi, and other microbes in the environment, both disease-causing and beneficial, is an increasingly important area of study, and it is now realized that plants are frequently host to an extensive microbial community.

It should be emphasized that these divisions are for convenience, and frequently interactions take place between the various processes. For example, photosynthesis is the source of transported sugars and all the other compounds in the plant, so it affects growth, development, and the response to environmental factors, whereas the environment and the developmental state of the plant very much influence photosynthesis. Even though many of the processes in Table 1 are now well understood, the controls are still under investigation. In addition, fine details, and sometimes surprising modifications and readjustments, are continually being made.

Levels of organization. Plant physiological research is carried out at various levels of organization and through the use of various methods. The main organizational levels are the genetic, molecular, subcellular, or organellar; the cellular; the organismal or whole-plant; and the population level.

At the molecular level, work is carried out on biochemical processes, such as those involved in the biosynthesis of a particular compound, often without consideration of the cellular or tissue location. Work at the subcellular level is aimed at understanding metabolic processes and their regulation, and also the localization of molecules in particular structures of the cell. This work is generally carried out with tissues or organs, the cellular structure of which has been disrupted, and often after separation (fractionation) of the cell contents into their components—the cell structures and organelles and the unstructured or soluble fraction.

Work at the cellular level often deals with the same processes, but is concerned with their integration in the cell as a whole. Since most cells are small, work is usually carried out with pieces of tissues. Tissues are selected which consist of cells that are as uniform as can be found, or with relatively undifferentiated tissues (so-called callus tissues) or cell suspensions.

TABLE 1. Some of the major topics covered by the subdisciplines of plant physiology

Nutrition and metabolism			
Metabolism	Transport	Growth and development	Environmental interactions
Cellular respiration Photosynthesis Light reactions CO_2 fixation Sugar biosynthesis Nitrogen metabolism Nitrogen fixation Biosynthetic pathways	Water uptake, movement, and loss to the atmosphere Sugar transport Mineral nutrition Mineral uptake	Plant hormones Photomorphogenesis Vegetative growth Flowering Fruiting Seed formation, dormancy, and germination	Stress physiology Water Salt Excess light Temperature Interactions with: Bacteria Fungi Insects Other plants

Research at the organismal level is concerned with the function of the plant as a whole and its different organs, and with the relationships between the organs. Research at the population level, which merges with experimental ecology, deals with physiological phenomena in plant associations which may consist either of one dominant species (like a field of corn), often in an agricultural context, or of numerous diverse species (such as a forest).

Work at the organismal level, and to some extent the population level, may be carried out in facilities permitting maintenance of controlled environmental conditions (light intensity and quality, temperature, water, humidity, air composition, and nutrient supply), often in specially constructed spaces referred to as growth chambers. The purpose is to obtain results that can be quantitatively reproduced, given that natural conditions cannot be exactly duplicated. However, results obtained under controlled environments may not be directly applied to the situation in the field, so attempts are also being made to experimentally modify large areas (for example, carbon dioxide level) under natural conditions to determine how plants function under the variable conditions of a natural environment. *See* PHYSIOLOGICAL ECOLOGY (PLANT); PLANT CELL.

Methods. The most elementary method of plant physiology is the descriptive method, in which a life process is studied and described under a certain set of conditions, at particular stages or throughout the life of the plant, with the goal of establishing its characteristics. The comparative method consists in the study of a process in plants belonging to different morphological, taxonomical, ecological, physiological, or genetic types but usually kept under identical conditions. The purpose is to establish variations of a process that are characteristic of different plant types; these variations may permit insights into the nature of the process in question. The correlational method seeks to recognize relationships between the course of a process and variations in the environment or in other processes. Such variations, too, often permit insights into the nature of a process, or the interdependence of two (or more) processes, although it is usually difficult to determine their precise causal relationship from correlational evidence alone. More definitive information is obtained from the experimental method, in which the process under study is modified by experimental treatments, and in this manner that insights are gained not only into its course but also into its causality. Experimental treatments are very diverse; they may be mechanical or surgical, physical (special conditions of light, temperature, and so on), chemical, or, increasingly, genetic.

Plant genotypes and mutants. The use of plant genotypes that vary in their form and function enables the development of correlations of endogenous differences between two or more genotypes and their chemical composition or metabolic functioning. Thus, insights are obtained as to those factors important in the life processes.

The use of genotypes has been extended through the use of purposely created mutants, particularly of the model plant *Arabidopsis thaliana*. The small size, rapid growth, and prolific seed production of this plant enable the screening of tens of thousands of randomly mutagenized plants for changes in their growth habit. As the entire genome (the DNA) of *Arabidopsis* has been sequenced, the exact nature of any mutation can be determined at the DNA level (**Table 2**). This enables particular genetic changes to be associated with metabolic, growth, developmental, or organismal interaction functions. The function of many of the genes has been elucidated, and is rapidly shared through vast data banks and the

TABLE 2. Classification of genes in *Arabidopsis* by function

Functional category	Percentage of genes
Biological process unknown	19
Other physiological processes	16.4
Other metabolic processes	15.7
Other cellular processes	15.4
Protein metabolism	6.9
Other biological processes	6.2
Response to abiotic or biotic stimulus	4
Transcription	3
Transport	2.8
Developmental processes	2.4
Response to stress	2.4
Cell organization and biogenesis	2.2
Electron transport or energy pathways	1.6
Signal transduction	1.4
DNA or RNA metabolism	0.7

SOURCE: From http://www.arabidopsis.org/tools/bulk/go/index.jsp or more generally http://www.arabidopsis.org.

Internet, facilitating an increased pace of understanding of plant function and growth. Once mutants are developed and characterized, cross-breeding of these mutants gives further characterization of the pathways and control of physiological processes through the interactions of specific combinations of genes.

Arabidopsis, however, is not a suitable model for all plant functions as, for example, it does not fix nitrogen through symbiotic bacteria, and it is not a host to symbiotic mycorrhizal fungi, which are important in mineral nutrient acquisition. For study of the latter functions, a small diploid alfalfa, *Medicago truncatula*, is emerging as the species of choice. For fruit development, the tomato (*Solanum lycopersicum*), whose genome is close to sequence completion, is being used extensively, and the rice (*Oryza sativa*) genome has been sequenced for the grass type of structure and growth functions.

Transgenic plants. Beyond the use of chance mutations, directed modification of the genome (the entire genetic content) is possible through the insertion of DNA sequences, or whole genes. This is done through the insertion of this DNA into the plant nuclei and its incorporation into the plant DNA, either (1) via the crown gall bacterium, *Agrobacterium faciens*, which has the unique ability to transfer specific portions of its own DNA into the DNA of plant cells; or (2) by shooting the DNA (coated onto minute tungsten particles) into plant cells, a technique referred to as microprojectile bombardment. The genes transferred are usually associated with a promoter (a gene switch) to ensure transcription; the promoter may be constitutive (active all the time) or may be one associated with expression in a certain cell type or tissue, under certain environmental conditions, or at a certain developmental stage. The gene or its promoter is frequently inserted attached to a reporter gene that produces a clearly visible marker (such as a green fluorescent protein) in the tissue where the gene is being transcribed.

The turning off of expression of specific genes to confirm the involvement of their products in a specific process is accomplished through antisense transformation, where the mirror-image DNA sequence of the targeted gene is inserted, or RNA interference, where a short, palindromic sequence of RNA, matching the target, results in the destruction of the target messenger RNA. *See* GENETIC ENGINEERING; NUCLEOPROTEIN; RECOMBINATION (GENETICS).

Gene expression. The expression of particular genes is studied under set environmental conditions, at specific growth stages, or under modified experimental conditions in order to understand the steps in a physiological process and its regulation. The genes may be of known or unknown function, though with time the function of more and more genes is becoming known. The expression of individual genes is studied by such techniques as quantitative PCR (polymerase chain reaction), but increasingly the expression of whole suites (thousands) of genes is being studied using gene microchips. The branch of genetics that studies organisms in terms of their genomes (their full DNA sequences) is called genomics, and the branch of genomics that determines the biological function of the genes and their products—in other words, that part related to plant physiology—is referred to as functional genomics. Further, researchers are now trying to understand the way in which the full set of proteins encoded by a genome (referred to as proteomics) directs the functioning and growth of the plant, either directly or through the production of the many chemicals in the plant. These chemicals range from sugars, amino acids, and all the compounds essential to life, to secondary metabolites, which are compounds involved in defense against fungal diseases and insects. Whereas metabolites used to be studied one or two at a time, instrumentation is now allowing the complete study of many proteins and metabolites at the same time, thus obtaining a much more complete picture of the makeup of the plant, and how it changes in response to environmental factors and during growth and development. The simultaneous study of the collection of all the small-molecule metabolites in a plant, which are the end products of gene expression, is called metabolomics. *See* GENOMICS.

Objectives. Plant physiology strives to be an exact science, which means essentially two things. First, it is a quantitative science which expresses its findings in numbers and in chemical, molecular, physical, and mathematical units. Because all organisms and hence also their life processes exhibit natural variation, this quantitative approach is subject to certain limitations. However, these limitations can in turn be defined by statistical methods, and a statistical evaluation of the results is often an essential part of research. Second, plant physiology uses the scientific method developed by other exact sciences; that is, the ultimate objective of plant physiology is to explain plant life by as few as possible comprehensive principles that are grounded in the principles of chemistry and physics, and the principles of mathematics that underlie physical and chemical principles. However, physiological phenomena exhibit far more diversity than purely physical and chemical phenomena. In addition to principles that are valid for all plants, some of which may be identical with principles of animal and human physiology, there will be principles valid only for certain kinds of plants or for certain kinds of life processes within specific plant types.

Relation to other biological disciplines. A life process may exhibit certain characteristic differences in different plant types. In addition, life processes do not proceed in a void but in organs, tissues, cells, and cell organelles, which in turn are produced and affected by other life processes. For these reasons, plant physiology depends upon information from other plant and life sciences (systematics, morphology, anatomy, cytology, molecular biology, microbiology). The ultrastructure of cells as studied with the electron microscope provides an extremely important backdrop to plant physiological research because, in conjunction with molecular studies, it permits the precise localization of a process within the

cell. Similarly, extensive attention is directed at the cellular localization of processes in different parts of the plant, as well as at the distribution of the products of these processes between these different parts. The precise spatial localization of processes within the plant and its cells is called compartmentation; the distribution of products is called partitioning.

Relation to other natural sciences. Because of the great importance of environmental conditions for the function of plants, plant physiology may draw on any science that deals with environmental phenomena and problems: meteorology, climatology, soil science, and others. However, the two natural sciences that have a unique position in relation to, and in fact within, plant physiology are physics and chemistry. The areas of life sciences which are based on the direct use of chemical and physical methods are called biochemistry and biophysics.

However, while such work is an essential, integral part of plant physiological work, plant physiology is a good deal more than the simple sum of plant biochemistry and plant biophysics. Biochemistry and biophysics are predominantly analytical disciplines which focus primarily on one process and seek to understand it with no concern for any surrounding or parallel processes; to accomplish this, they frequently destroy the integrity of the system. Physiology aims at understanding the function of the whole organism. It strives at preserving the latter's integrity, and to do this it needs more than the principles and methods of biochemistry and biophysics.

In addition, as mentioned above, plant physiology relies ever more intensively on the genetic sciences in explaining how an organism's genetic makeup directs the growth and functioning of the plant and its interaction with the world around it.

Global significance of plants. Plants—apart from some autotrophic microorganisms—are the only organisms capable of converting inorganic (mainly carbon and nitrogen) to organic matter by making use of solar energy. Thus, plants are the direct or indirect source of food for humans and all animals, and are also a source of fiber, fuel, and feedstock. While plants are profoundly dependent on the environment, they greatly affect the environment in turn. The peat bogs which cover wide areas of low-lying land in some parts of the world are products of plant life, and plants, especially trees, can be important factors in protecting land against erosion by wind and water. The most important general role of plants in relation to the environment is their function as a link in major natural cycles. Two examples may suffice. It is estimated that plants remove annually more than 2×10^{11} tons (1.8×10^{11} metric tons) of carbon dioxide from the Earth's atmosphere, transforming it into plant matter and other forms of life, which is an increasingly important function in the face of rising carbon dioxide levels and global warming. It has also been estimated that one-third of the water that is annually precipitated on the United States passes through plants, which take it up from the soil and release it into the air. Thus, disruptions of the plant cover may result in serious changes of the climate and hence of the condition of the land—changes which quite generally are of an unfavorable nature. Many regions of the Earth have become arid because of excessive exploitation and destruction of their original vegetation. Unfortunately, such uncontrolled and excessive exploitation of the natural vegetation is still continuing in many parts of the world. Plant physiology has helped greatly in understanding many of these plant-related phenomena, particularly in relation to agriculture. *See* BIOGEOCHEMISTRY; DESERTIFICATION; RAINFOREST; SOIL DEGRADATION.

Increasing plant yields. Breeding work having to do with crop plants was formerly directed primarily at increases of yield, with little attention to such aspects as the water and fertilizer demands of the plant, or little understanding of the reasons for the increases in yield. The result is that the most productive crop varieties are often wasteful in terms of the amounts of water and mineral nutrients which they need for optimal performance. Yet water is becoming a scarce commodity in some parts of the world, and the production of mineral fertilizers is a process requiring large amounts of energy, which is also getting scarce as well as expensive. Thus, efforts have been initiated to breed crop varieties for more efficient water and fertilizer use, and similarly varieties which can perform on land that is considered unsuitable for agriculture, such as in saline soils. The precise genetic engineering of crop plants and forest with more favorable growing characteristics can be solved only with the knowledge derived from the understanding of plant physiology. Plant physiology can also play an important function by providing the basis for the development of plants which can be used for reversal of the results of the excessive exploitation or pollution of the natural environment. *See* AGRICULTURAL SCIENCE (PLANT); BREEDING (PLANT); FERTILIZER; PHOTOSYNTHESIS; PLANT GROWTH; PLANT HORMONES; PLANT METABOLISM; PLANT MINERAL NUTRITION; PLANT RESPIRATION; PLANT-WATER RELATIONS.

Peter J. Davies; Anton Lang

Bibliography. B. B. Buchanan, W. Gruissem, and R. L. Jones, *Biochemistry and Molecular Biology of Plants*, American Society of Plant Biologists, Rockville, MD, 2000; L. Taiz and E. Zeiger, *Plant Physiology*, 4th ed., Sinauer, Sunderland, MA, 2006.

Plant pigment

A substance in a plant that imparts coloration. The photosynthetic pigments are involved in light harvesting and energy transfer in photosynthesis. This group of pigments consists of the tetrapyrroles, which include chlorophylls (chl) and phycobilins, and the carotenoids. The light-absorbing groups of these molecules, the chromophores, contain conjugated double bonds (alternating single and double bonds) which make them effective photoreceptors. Their effectiveness in harvesting light for photosynthesis is shown in **Fig. 1**. The sum of the absorption spectra of the chlorophylls and the carotenoids,

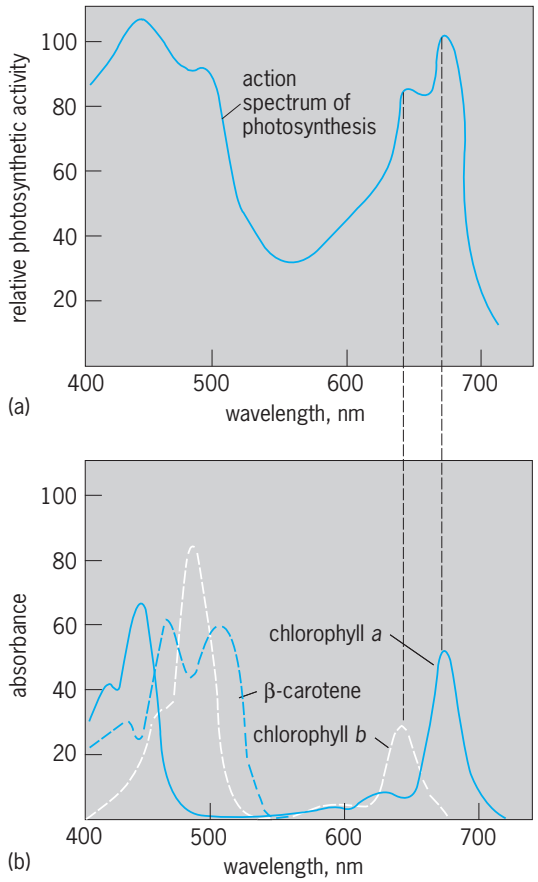

Fig. 1. Photosynthetic spectra of higher plants. (*a*) Action spectrum of photosynthesis showing the wavelengths of visible light that support photosynthesis. Measurements of photosynthetic activity include O_2 evolution or NADPH formation. (*b*) Absorption spectra of chlorophyll *a*, chlorophyll *b*, and β-carotene. (*After J. Darnell, H. Lodish, and D. Baltimore, Molecular Cell Biology, W. H. Freeman, 1986*)

evident in the absorption spectrum of a green leaf, is equivalent to the action spectrum of photosynthesis (Fig. 1*a*). *See* CHLOROPHYLL; LIGHT; PHOTOSYNTHESIS; PIGMENTATION.

The second major group of plant pigments comprises the anthocyanins, intensely colored plant pigments responsible for most scarlet, crimson, purple, mauve, and blue colors in higher plants. About 100 different anthocyanins are known. Unlike the chlorophylls and carotenoids, which are lipid-soluble chloroplast pigments, the anthocyanins are water-soluble and are located in the cell vacuole. Chemically, they are a class of flavonoids and are particularly closely related, both structurally and biosynthetically, to the flavonols. One of their major values to the plant lies in the contrasting colors they provide in flower and fruit against the green background of the leaf to attract insects and animals for pollination and seed dispersal. *See* FLAVONOID.

Photosynthetic Pigments

The conversion of light energy to chemical energy is initiated when light-harvesting pigments absorb light of appropriate wavelengths and a π electron in the conjugated double-bond system of each molecule is raised from the ground state to an excited, higher-energy state. The transition is very rapid, and the excited state has a very short lifetime (15 nanoseconds for chl *a*).

When an acceptor molecule is not available to the excited pigment, the excited electron may return to the ground state, with the energy reemitted as a quantum of light (photon). Because of the loss of some energy during the return from the excited to the ground state, the reemitted light, or fluorescence, has less energy and is therefore at a longer wavelength. The difference between the wavelength of maximal absorption and fluorescence is referred to as the Stokes shift and is approximately 8 nanometers for chlorophylls. *See* FLUORESCENCE.

When an acceptor molecule is available, the energy may be transferred from each excited pigment to an adjacent unexcited one by inductive resonance. This requires proximity (less than 3 nm) and appropriate orientation of the adjacent pigment molecules, as well as an overlap between the fluorescence emission spectrum of the excited molecule and the absorption spectrum of the acceptor molecule. Inductive resonance is random until the energy is transferred from the different antenna (light-harvesting) pigments to antenna chl *a* in reaction centers and from there to special chl *a* molecules in which the excited state is at a lower energy level. Each of these special molecules is oxidized by expelling a high-energy electron, which reduces an electron acceptor to initiate electron transport. Light energy thus induces the formation of a strong oxidant in the form of chl a^+ and a high-energy electron. The ultimate result of this process is chemical energy in the form of adenosine triphosphate (ATP) and reducing power in the form of reduced nicotinamide adenine dinucleotide phosphate (NADPH) or reduced nicotinamide adenine dinucleotide (NADH). *See* ADENOSINE TRIPHOSPHATE (ATP); NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD).

The range of absorption maxima of photosynthetic pigments in the visible spectrum (see **table**) is due to the different light-harvesting pigments that evolved from the ancestors of higher plants–photosynthetic bacteria, and algae which lived, as they still do, in aqueous environments. The absorption maxima thus represent the evolution of light-harvesting systems adapted to light available at different depths of a body of water. Many of these spectra are redshifted when they are compared to the spectra of the free pigments in solution. This is attributed to a polarizing effect on the conjugated double bonds by the protein to which the pigments are bound in the living cell. The specific functions of the various pigments, as well as the different functions of the same pigments, appear to be associated with their noncovalent or covalent binding to specific proteins in or on the thylakoid membrane. The proteins may localize and keep the chromophores in a proper orientation for efficient energy transfer.

Chlorophylls. Chlorophylls are cyclic tetrapyrroles, or porphyrins, in which the four pyrrole

Absorption and fluorescence maxima of photosynthetic pigments in living cells*

Pigment	Absorption maxima, nm										Fluorescence emission maxima, nm
Chlorophyll *a*	.	436	.	.	.	.	676	.	.	.	685
Chlorophyll *b*	.	480	.	.	.	.	650	.	.	.	—
Bacteriochlorophyll *a*	375	.	.	590	.	.	.	800	850	890	900
Bacteriochlorophyll *b*	.	400	.	605	.	.	.	.	835–850	1015–1035	1020–1050
Bacteriochlorophyll *c*	.	457	.	.	.	.	.	750	.	.	770
Bacteriochlorophyll *d*	.	446	.	.	.	.	.	730	.	.	771
Allophycocyanin	380	.	.	.	.	650	654	.	.	.	660–680
Cyanobacterial-phycocyanin	380	.	.	.	620	.	.	.	.	.	650
Red algal-phycocyanin	380	.	553	.	615	.	.	.	.	.	640
Cyanobacterial-phycoerythrin	380	.	.	565	.	.	.	.	.	.	578
Red algal-phycoerythrin	380	498	540	565	.	.	.	.	.	.	578
Bangiophycal-phycoerythrin	380	498	545	563	.	.	.	.	.		575
Carotenoids		430 →		560							
β-Carotene		463, 498									
	Ultraviolet	Blue	Green	Yellow	Orange	Red			Near-infrared		Red to near-infrared

*The absorption and fluorescence maxima are representative for each group of pigments and may vary slightly among organisms.

nitrogens are coordinated with a magnesium atom (**Fig. 2**). They are embedded in photosynthetic membranes, the thylakoids, in noncovalent association with specific intrinsic membrane proteins. This is believed to occur by ligand formation between the tetrapyrrole magnesium and a histidine, glutamine, or asparagine side-chain nitrogen or a side-chain oxygen of glutamate, aspartate, or tyrosine from the protein. Chlorophylls are an integral part of the photosynthetic apparatus of all higher plants, algae, cyanobacteria, and prochlorophytes. Chlorophyll *a* is found in all oxygen-evolving organisms, functioning both as a light-harvesting pigment and as the special energy-trapping reaction-center core pigment (P), either P680 for photosystem II (PS II) or P700 for PS I. It absorbs maximally in the blue and red regions of the visible spectrum (see table). There are several distinct species of chl *a*; certain fluorescent species are due to chl *a* in PS I, while others arise from chl *a* in PS II. The different chl *a* species reflect different chlorophyll-protein complexes. At room temperature, PS II is the only fluorescent photosystem. PS I quenches PS II fluorescence, and the efficiency of energy transfer from PS II to PS I under various conditions is usually measured by following the fluorescence yield, that is, the decrease in fluorescence with time.

Other chlorophylls that harvest and transfer the captured light energy to chl *a* may also be present. They differ from chl *a* primarily in side-group substitutions on the pyrrole rings. The antenna pigments do not fluoresce in the living cell because the energy is efficiently transferred to chl *a*. Chlorophyll *b* is found in higher plants, green algae, euglenoids, and prochlorophytes; brown algae, diatoms, and dinoflagellates contain chl *c*; and some red algae contain chl *d*. Cyanobacteria and most red algae contain only chl *a*. The antenna chlorophylls are arranged in light-harvesting complexes surrounding the photosynthetic reaction center.

Phycobilins. Phycobilins are linear, open-chain tetrapyrrole pigments which are covalently linked to specific proteins to form phycobiliproteins (Fig. 2). These phycobiliproteins are found on the surface of thylakoid membranes in highly ordered aggregates called phycobilisomes. Phycobilisomes function as antenna complexes, harvesting and transferring light energy to chl *a* of membrane-embedded PS II, in cyanobacteria, red algae, and cyanelles. The phycobiliproteins of cryptomonads have a similar function but contain one of only two phycobiliproteins, phycocyanin or phycoerythrin, which are located in the thylakoid lumen and are not organized into phycobilisomes.

The major phycobiliproteins composing the phycobilisomes fall into four classes based upon distinctive absorption spectra associated with their chromophores, which are covalently bound via ring A of the phycobilin in a thioether linkage to specific cysteinyl residues in the protein. These include phycoerythrin, which contains phycourobilin chromophores; phycoerythrocyanin, which contains phycocyanobilins and phycobiliviolins; phycocyanin, which contains phycocyanobilins; and allophycocyanin, which also contains phycocyanobilins. The distinct absorption spectra of the different phycobiliproteins is due to interaction of the chromophores with the protein to which they are bound as well as to the way in which they are aggregated to form phycobilisomes. All red algae and cyanobacteria contain allophycocyanin and phycocyanin, and in addition many have phycoerythrin or phycoerythrocyanin. There is a directed energy transfer from phycoerythrin or phycoerythrocyanin to phycocyanin, from phycocyanin to allophycocyanin, and from allophycocyanin to chl *a* in PS II. This is shown in

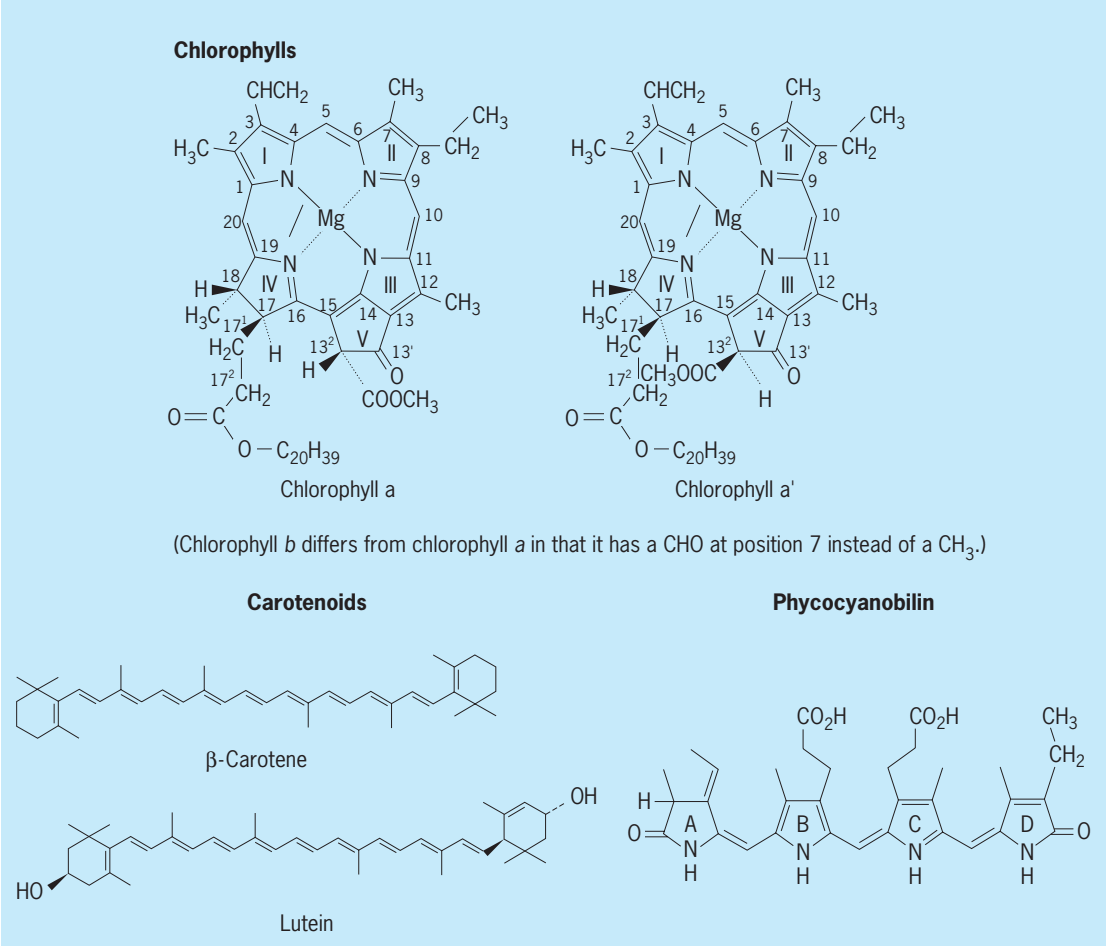

Fig. 2. Structures of three plant pigment types.

isolated phycobilisomes as fluorescence emission from allophycocyanin and little or no emission from the other excited phycobiliproteins, indicating efficient energy transfer. The table shows that phycobiliproteins exhibit the widest absorption range (490–620 nm), providing those organisms that contain them with great adaptability in harvesting light for photosynthesis. *See* PHYCOBILIN.

Carotenoids. Carotenoids are lipid-soluble orange and yellow pigments, noncovalently linked to specific integral thylakoid membrane proteins. They include the hydrocarbon carotenes and their oxygenated derivatives, the xanthophylls, with the latter being the most abundant naturally occurring carotenoids. They are necessary for photosynthesis in most photosynthetic organisms, functioning in both light harvesting and photoprotection of chlorophyll. Most carotenoids have 40-carbon isoprenoid (terpene) backbones with conjugated double bonds; a substituted cyclohexene is usually found at one or both ends of the molecule. The network of conjugated double bonds forms an extended π-electron system which accounts for the ability of carotenoids to absorb both ultraviolet (UV) and visible light. They absorb light in regions of the visible spectrum where chlorophylls absorb little, or not at all. Some carotenoids are reported to stabilize membrane structure, protecting membrane lipids from oxidative damage.

β-Carotene, the major carotene involved in photosynthesis, absorbs strongly in the blue region of the visible spectrum (Fig. 2). β-Carotene is found in the PS II reaction center and light-harvesting complexes. PS II is most sensitive to photodamage. In the absence of carotenoids (mutation analysis), chlorophyll is photooxidized by toxic forms of oxygen called reactive oxygen species (ROS) which can result in cell damage and death. β-Carotene protects chlorophyll from photooxidation by trapping and dissipating excess excitation energy from absorbed light before it reaches the reaction center. It quenches the triplet-excited chlorophyll (the state of chlorophyll that occurs when two of its electrons have their magnetic moments in parallel) and the highly reactive singlet oxygen which can be generated by the reaction of triplet chlorophyll with ground-state oxygen. *See* BIOLOGICAL OXIDATION.

Lutein (Fig. 2), violaxanthin, and neoxanthin are the major xanthophylls associated with the light-harvesting complexes of PS II in higher plants, absorbing light and transmitting the absorbed energy to chl *b* by resonance energy transfer. Under high light, some of the violaxanthin is enzymatically converted to zeaxanthin via the loss of two oxygen molecules

and the addition of two double bonds. In low light or in the dark, the reaction is reversed and zeaxanthin is converted back to violaxanthin. This is referred to as the xanthophyll cycle. The increase in conjugated carbon-carbon double bonds makes the singlet excited state of zeaxanthin lower than that of antenna chlorophyll and thus unable to transfer energy to chlorophyll by resonance energy transfer. However, it can accept a downhill transfer of energy from excited chlorophyll, so excess energy from chlorophyll can be harmlessly dissipated as heat. Thus, in high light the production of zeaxanthin results in photoprotection, and in low light the appearance of more violaxanthin allows for more light energy absorption and transfer to chlorophyll. *See* CAROTENOID.

Reaction center pigments. Reaction centers are the smallest multiprotein complexes isolated from thylakoid membranes functional in either PS I or PS II. They usually include small antenna chl *a* proteins. Reaction centers, embedded in the membranes in a highly ordered arrangement, are surrounded by light-harvesting complexes which, though probably not essential, ensure efficient energy transfer. The reaction-center cores are the smallest membrane complexes capable of specific energy transduction. Historically, each reaction center with its light-harvesting chlorophylls (approximately 300) is referred to as a photosynthetic unit. Special chl *a* molecules in the reaction center cores function as the primary electron donors in the two reaction centers of oxygen-evolving photosynthetic organisms. The special environment created by the association of these chl *a* molecules with certain proteins appears to result in a lower amount of energy being necessary to excite an electron out of the chlorophyll. P700 and P680 are the special chlorophylls of the PS I and PS II reaction centers, respectively, exhibiting the longest wavelength and thus the lowest energy when in the excited state. They are essentially energy sinks in their excited states, each donating an electron to a nearby molecule. Both P700 and P680 are each made up of a pair of chlorophyll molecules. The special pair in P700 consists of a chl *a* and a chl *a* epimer (a stereoisomer of another compound that has a different configuration at only one of several chiral carbon centers) which is called chl *a*′ (Fig. 2). The nature of the two chl *a* molecules in P680 has not been definitively determined.

Nonphotosynthetic Pigments

Anthocyanins are water-soluble plant pigments found in petals, flowers, and the lower epidermis of leaves. They are produced in the cytoplasm and usually transported into the cell vacuole. *See* PLANT CELL.

Anthocyanins are glycosylated flavonoids. They are the most abundant flavonoids found in flowers and petals. On acid hydrolysis, they yield sugars and colored aglycones called anthocyanidins. The most common aglycone is 3,5,7,3′,4′-pentahydroxyflavylium, or cyanidin, so called after the cornflower (*Centaurea cyanus*), from which it was first isolated; the structure is

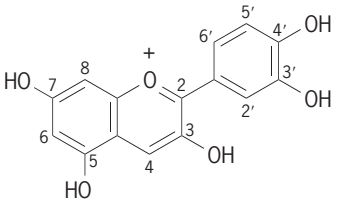

The other 15 known anthocyanidins have the same basic structure as cyanidin and differ only in having fewer or more hydroxyl groups or in being *O*-methylated. The other common anthocyanidins are pelargonidin (structure as for cyanidin without the 3′-hydroxyl), peonidin (3′-methyl ether of cyanidin), delphinidin (extra hydroxyl at 5′), and its 3′-monomethyl and 3′,5′-dimethyl ethers, petunidin and malvidin. The sugar present in anthocyanins may be glucose, galactose, rhamnose, xylose, or arabinose, a combination of two or three of these monosaccharides, or a combination of a monosaccharide with an aromatic hydroxy acid. Sugars are usually attached to the 3-hydroxyl, giving 3-glucosides, or to the 3- and 5-hydroxyl groups, giving 3,5-diglucosides. The presence of these sugars accounts not only for the sap solubility of anthocyanins but also for their natural stability, the free aglycones being rather water-insoluble and unstable.

Anthocyanins are cations, the positive charge being distributed among the various carbons and oxygens of the conjugated ring system. They are normally isolated as chlorides since, in extracting them from plant tissues into alcohol or water, it is necessary to add hydrochloric acid in order to retain the color. The pigments probably occur in plant tissue in association with organic acid anions, and the actual color in the plant may be determined by the acidity of the cell sap as well as by the presence or absence of other constituents in the sap. Anthocyanin–metal complexes have been isolated from the petals of a number of blue-flowered plants, such as the cornflower. These complexes have been shown to contain flavone and polysaccharide as well as metals such as iron and aluminum. The color of the isolated pigments fades above pH 4, and in alkali turns blue, with destruction of the pigment by aerial oxidation. Anthocyanins have characteristic absorption properties in the visible spectrum (**Fig. 3**), and are

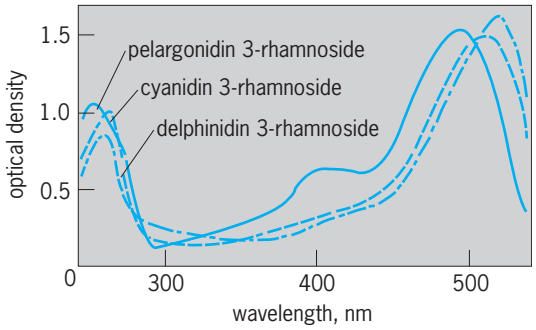

Fig. 3. Absorption spectra of three anthocyanins.

easily separated and identified by paper or thin-layer chromatography by using aqueous hydrochloric acid as a solvent. Characterization is facilitated by the fact that their mobilities on filter paper are closely correlated with the number of hydroxyl, methoxyl, and sugar groups present. The R_f values assigned to each pigment represent the observed migration of a chromatographic zone relative to the solvent front.

Numerous functions for anthocyanin pigmentation in leaf tissue have been reported. While permanent leaf coloration due to anthocyanin is not usual, transient anthocyanin synthesis in either young or autumnal leaves is quite common. Moreover, red coloration often builds up in the leaves of plants suffering from mineral deficiency, environmental stress, disease, or wounding. There is little doubt that under such circumstances anthocyanin synthesis is closely connected with carbohydrate metabolism. The concentration of anthocyanin in certain species of *Eucalyptus* has been found to increase with the altitude at which the tree grows, and in this case pigment production may well be a protective device against ultraviolet radiation. *See* LEAF.

Although in some plant species anthocyanins accumulate in healthy leaves, in other species they are induced in leaves in response to environmental stress such as low temperatures, high light, insect or pathogen attack, or deficiency in nutrients such as nitrogen and phosphate. Researchers have shown that anthocyanins function in photoprotection by modifying the quantity and quality of light reaching chloroplasts, thereby acting as light screens to reduce photoinhibition and photobleaching of chlorophyll under high light. Anthocyanins also function as antioxidants against harmful reactive oxygen species, modulate auxin transport, act as antimicrobial agents, and deter feeding by insects. In plant reproductive tissues, they attract pollinators and seed dispersal organisms. Shirley Raps; Jeffrey B. Harborne

Bibliography. K. M. Davies (ed.), *Annual Plant Reviews*, vol. 14: *Plant Pigments and Their Manipulation* [An introduction to plant pigments in biology and commerce, pp. 1–22; Chlorophylls, pp. 23–56; Carotenoids, pp. 57–91; Flavonoids, pp. 92–149], Blackwell/CRC Press, 2004; W. G. Hopkins and N. P. A. Huner, *Introduction to Plant Physiology*, 3d ed., Wiley, 2004; M. Lohr, C.-S. Im, and A. R. Grossman, Genome-based examination of chlorophyll and carotenoid biosynthesis in *Chlamydomonas reinhardtii*, *Plant Physiol.*, 138:490–515, 2005; M. Pessarakli (ed.), *Handbook of Photosynthesis*, 2d ed. [Photosystem I: Structures and functions, pp. 169–192; Effects of heavy metals on chlorophyll-protein complexes in higher plants: Causes and consequences, pp. 865–888], CRC Press, 2005; S. Raps, Differentiation between phycobiliprotein and colorless linker polypeptides by fluorescence in the presence of $ZnSO_4$, *Plant Physiol.*, 92:358–362, 1990; A. V. Ruban et al., Plasticity in the composition of the light harvesting antenna of higher plants preserves structural integrity and biological function, *J. Biol. Chem*, 281:14981–14990, 2006; C. Solfanelli et al., Sucrose-specific induction of the anthocyanin synthetic pathway in *Arabidopsis*, *Plant Physiol.*, 140:637–646, 2006; J. Veerman et al., The PsbU subunit of photosystem II stabilizes energy transfer and primary photochemistry in the phycobilisome-photosystem II assembly of *Synechocystis* sp. PCC 6803, *Biochemistry*, 44:16939–16948, 2005.

Plant propagation

The deliberate, directed reproduction of plants using plant cells, tissues, or organs.

Traditional method. Asexual propagation, also called vegetative propagation, is accomplished by taking cuttings, by grafting or budding, by layering, by division of plants, or by separation of specialized structures such as tubers, rhizomes, or bulbs. This method of propagation is used in agriculture, in scientific research, and in professional and recreational gardening. It has a number of advantages over seed propagation: it retains the genetic constitution of the plant type almost completely; it is faster than seed propagation; it may allow elimination of the nonfruiting, juvenile phase of the plant's life; it preserves unique, especially productive, or esthetically desirable plant forms; and it allows plants with roots well adapted for growth on poor soils to be combined with tops that produce superior fruits, nuts, or other products.

Cuttings. Propagation by cuttings involves removing branches, leaves, roots, or bud-bearing stems from a donor plant and causing each cutting to regenerate the other parts requisite to a new, complete plant. The plant part chosen for making cuttings varies with the plant species, because different species retain regenerative capacities to differing extents in their various parts. In propagating coleus plants, for example, branch cuttings are removed from the plant, caused to root, and then placed in soil. Coleus leaf cuttings also root, but they form only a bigger leaf, not a whole plant. Leaf cuttings are, however, used for propagating African violets. One or more buds are regenerated on the petiole (stalk) of the African violet leaf when the petiole is embedded in sand or vermiculite. The bud forms new leaves and the cutting roots, thus establishing a new plant. In horticultural practice, root formation on cuttings is often increased and accelerated by the application of auxin, a root-inducing chemical that mimics the plant's own root-forming hormone. *See* AUXIN; VERMICULITE.

In propagation by cuttings, periodic misting with water helps prevent desiccation. Shading and special disease-preventing precautions may also be necessary during this vulnerable period before wounds have healed and roots have been regenerated. In scientific research, the use of cuttings allows all test plants to be derived from one plant; thus a clone of plants is established. Such a procedure greatly diminishes biological variation, a source of random experimental error. However, if an infectious disease

organism infests the clone, it may cause a spontaneous change in test responses.

Grafting. Grafting, another method of vegetative propagation, follows a highly developed set of horticultural practices in which two different plants are combined into one plant that has some of the desirable traits of both. All grafts consist of at least two parts—the scion, which is the shoot-bearing or shoot-producing top part, and the stock, which is the root-bearing bottom part. In this way, a rootstock that tolerates unfavorable soil conditions or resists soil-borne disease organisms can be combined with a scion that produces pest-resistant fruit to yield a superior plant.

Grafting offers an opportunity to study the effects of one part of a plant on the functioning of its other parts. For example, evidence for a flowering hormone is based heavily on grafting experiments in which stock and scion are taken from different species or from different pretreatment environments. Evidence indicates that a hormone, florigen, almost universal in plants, is produced in leaves and is translocated to the shoot apex, where it evokes flower initiation.

In all grafting, the objective is to establish a firm junction between scion and stock. This junction must be traversed by new and fully functional vascular tissues before the graft is successful. Except at the growing tips of plants, vascular tissues are produced by a layer of rapidly dividing cells between xylem and phloem called the cambium, which is found in both stems and roots. The aim in grafting, to join the cambium of the scion to the cambium of the stock, can be accomplished by any of several grafting techniques (see **illus.**). In most grafts, wound tissue (callus) is formed by each of the graft partners at their juxtaposed surfaces. There the callus cells of one partner are made to adhere to those of the other, as in the whip graft (illus. *a*). Through cell divisions on both sides, the cambium of both stock and scion is extended into the callus. Both xylem and phloem are produced by this new cambium, thus establishing vascular continuity across the graft union. *See* PLANT TISSUE SYSTEMS.

Budding, or bud grafting, is a specific grafting technique of which T-budding or shield budding, the most common form, can be used successfully when bark can be easily removed from the wood (illus. *b*). Vertical and horizontal cuts are made through the bark of the stock plant in the shape of a T. The two cuts, each about $1^3/_4$ in. (4.5 cm) long, are made through the bark down to the underlying wood. The bud or scion is removed from a budstick by cutting out a shield- or tongue-shaped piece of wood and bark. The cut starts about $^1/_3$–$^1/_2$ in. (0.9–1.3 cm) below the bud and widens and deepens as it is continued upward. About $^1/_3$ in. (0.9 cm) above the bud, a horizontal cut is made so as to remove the bud and its underlying bark and wood as one piece. (If the wood can be removed from the bud-bearing piece without tearing the bark, it is done before inserting the bud into its grafted position.) The bark flaps at both sides of the T are spread open

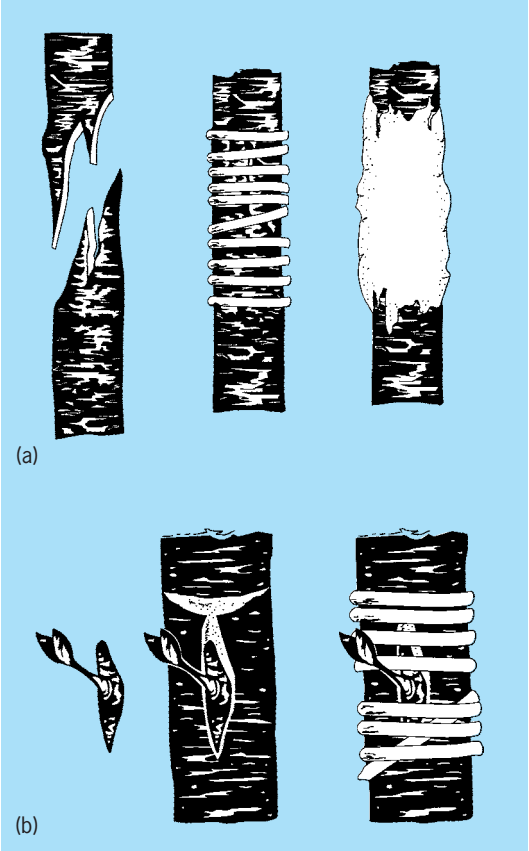

Two grafting techniques. (*a*) Whip graft, with paraffin to seal in moisture. (*b*) T-budding.

slightly, and the shield-shaped piece, with its original orientation maintained, is inserted into the T. The graft, but not the bud, is wrapped tightly with rubber budding strips or with strips of polyethylene or a paraffin-coated film to seal in water. When the graft union has completed its development, the original shoot on the stock above the grafted bud is cut off. T-budding is used by nurseries to propagate roses, most fruit-tree species, and many ornamental shrubs.

The extent to which successful grafts can be made is limited. Different varieties of the same species can almost always be grafted successfully, but different species in the same genus are usually not graft-compatible. In some cases, a graft between species succeeds in one direction only. For example, a Marianna plum scion can be grafted on a peach stock, but the reciprocal graft is not successful. Species from different genera are seldom compatible.

The physiological basis for graft failure or graft incompatibility is not fully understood. According to one theory, metabolites produced by one partner are toxic to tissues of the other or are converted to toxins by enzymes present in the other. The clearest example is that of pear grafted on quince. Quince produces cyanogenic glycoside that is capable of releasing poisonous cyanide when acted on by the enzyme glycosidase, which is produced in tissues of the pear. The cyanide kills pear cells which otherwise would have been involved in producing a graft

union. Another theory supposes the existence of sophisticated cellular mechanisms for recognizing foreign cells that are similar to the immune system reactions in animals.

Layering. Layering is used to propagate plants by causing roots to develop on stems, a phenomenon that occurs rarely in nature. This is accomplished by burying a portion of the stem, which is still attached to the plant, in order to keep it dark, warm, and moist. To further hasten root formation, the bark of the portion being buried may be cut or ringed.

In tip layering, only the stem tip is buried. Tip layering is used to propagate trailing blackberries and black raspberries, and duplicates the natural propagation methods of those plants. In simple layering, one portion of the stem, still attached to the plant, is buried. Filberts, for example, are routinely propagated by simple layering. In compound layering, more than one portion of the same stem is buried. Muscadine grapes are propagated by this method.

With air layering, a few inches of a stem are girdled or cut and then enclosed in sphagnum moss to provoke root formation. The moss is wrapped in transparent, nonporous film and is kept slightly damp until the roots form. The rooted piece of stem, including an axillary bud or a large leafy shoot at its upper end, is cut off and planted in its own pot. Air layering is used to propagate large plants such as *Dieffenbachia, Ficus*, pecan, and *Magnolia*. Stool layering involves pruning a plant so that it develops several branches near its base. The undersides of those branches are buried in soil until they form roots. The rooted branches are then cut off and set out as individuals. Plants that are propagated by stool layering include currants, gooseberries, and clonal rootstocks of pear and apple. *See* AGRICULTURAL SCIENCE (PLANT); BREEDING (PLANT); PLANT REPRODUCTION.

C. E. LaMotte

Culture technology. Tissue cultures and protoplast cultures are among the techniques that have been investigated for plant propagation; the success of a specific technique depends on a number of factors. Practical applications of such methods include the clonal propagation of desirable phenotypes and the commercial production of virus-free plants.

Tissue culture. Plant tissue cultures are initiated by excising tissue containing nucleated cells and placing it on an enriched sterile culture medium. To prevent growth of undesirable microorganisms, the plant tissue itself must first be surface-sterilized, usually by immersing it in a dilute bleach solution and then rinsing repeatedly with sterile water. Another approach is to surface-sterilize seeds, germinate them on sterile medium solidified with agar, and later remove axenic tissue from the seedlings. The response of a plant tissue to a culture medium depends on a number of factors: plant species, source of tissue, chronological age and physiological state of the tissue, ingredients of the culture medium, and physical culturing conditions, such as temperature, photoperiod, and aeration.

For unknown reasons, different plant species often exhibit unique requirements for successful growth in culture; it has been suggested that variation in products of secondary metabolism may explain observed variation in the responses. Within a single plant, different source tissues often show wide variation in response to a given culture condition, which may be due to different endogenous concentrations of either growth-promoting or growth-inhibiting substances.

Though most tissue culture media have similar compositions, recipes suitable for a given species and type of tissue must be determined by trial and error. Standard media include appropriate amounts of macronutrient and micronutrient elements, vitamins, a carbon source (usually glucose or sucrose), and plant hormones (usually an auxin and a cytokinin). In addition, many popular media contain complex additives such as coconut milk, yeast extract, and casein hydrolysate.

Identical plant tissues may respond quite differently when cultured on media with different compositions. Plant hormones have the strongest influence of all ingredients on the response of a given plant tissue. In general, relatively high concentrations of auxins stimulate callus proliferation and possibly root development, while relatively high concentrations of cytokinins induce development of shoots. In some plant species, somatic embryos that exhibit patterns of development similar to those exhibited by sexual embryos can be induced to form from callus tissue that is first cultured on a medium containing a high auxin concentration, then subcultured to a medium with little or no auxin. *See* PLANT HORMONES.

Differences in the physical environment used for culturing plant tissues can also profoundly affect their growth and development. For example, callus cultures grown in the light usually produce chlorophyll, while those grown in the dark do not. Light is usually required for shoot formation from callus tissue.

Protoplast culture. Though technically more demanding, successful culture of plant protoplasts involves the same basic principles as plant tissue culture. Empirical methods are used to determine detailed techniques for individual species; such factors as plant species, tissue source, age, culture medium, and physical culture conditions have to be considered.

Plant protoplasts are most often isolated by using fungal enzymes (usually a cellulase and a pectinase) that degrade the cell walls. Protoplasts have been isolated from leaves, roots, stems, petals, hypocotyls, cotyledons, coleoptiles, cell suspensions, and callus tissue, but most commonly they are isolated from leaf mesophyll tissue or rapidly growing cell suspensions. It is critical that both the isolation media and culture media contain the proper osmotic pressure to prevent lysis of the delicate protoplasts; mannitol, sorbitol, glucose, and sucrose are common osmotic stabilizers. *See* PLANT CELL.

Isolated protoplasts are injured cells, and they must replace their cell wall before they are capable

of sustained division. Trial-and-error techniques have not yet led to the discovery of procedures whereby protoplasts of many plant species can be successfully induced to resynthesize a cell wall, divide and form callus tissue, and then regenerate plantlets. Because recalcitrant species can be found in diverse taxonomic orders, it is assumed that epigenetic blocks (such as failing to recognize hormonal signals) or physiological blocks (such as not receiving requisite environmental signals) to cell division will eventually be discovered and overcome.

Applications of culture technology. Perhaps the most obvious use of plant tissue culture is in clonal propagation of desirable phenotypes. For example, elite hybrid *Cymbidium* orchids are rapidly propagated in tissue culture by hormonal stimulation of shoot meristem production from excised protocorms. In this way, changes that might cause loss of desirable flower traits during sexual reproduction can be avoided. However, tissue culture procedures have been shown to enhance mutation frequency, and true genetic clones are not always produced.

Meristem culture is also used in the commercial production of virus-free plants, including chrysanthemums, potatoes, and dahlias. Many viruses move through the plant's vascular tissue, and because vascular connections to the apical meristem are not continuous, excision and culture of apical meristems provides a means for eradicating viruses.

Embryos that form when two sexually incompatible species are crossed are often arrested in their development because of incompatibility between the developing hybrid embryo tissue and the maternal tissue of the ovule. However, if these embryos are excised from the ovule and cultured in a nutrient medium, they sometimes develop into mature hybrid plants that may exhibit desirable agronomic traits.

Protoplasts from two sexually incompatible species have been fused to form somatic hybrids. Hybrid plants have been regenerated from some somatic hybrid protoplasts. Commercially valuable hybrids have yet to be produced by this means, but the technique has considerable potential.

Both tissue and protoplast cultures have been explored as means for introducing limited genetic variation into agronomically useful cultivars. When individual mesophyll cells are isolated from a leaf and cultured as either protoplasts or tissue cultures, they often exhibit somatic variation. (Cytological studies show that chromosomal rearrangements and duplications are common in cultured plant cells.) This somaclonal variation may be useful in selecting cells that tolerate phytotoxins, herbicides, and environmental stresses, and the plants regenerated from these cells may exhibit the same desirable phenotypes.

Large-scale culture of plant cells for production of pharmaceuticals and agrochemicals has resulted in limited success. Though many patent applications have been filed, the technologies required for cost-effective cultivation of plant cells in large vats have not yet been developed.

Cryopreservation of plant cells has been proposed as an alternative to seed stores for preservation of genetic resources. The cells are stabilized by immersion in cryoprotectants such as dimethyl sulfoxide and glycerol, slowly frozen to −321°F (−196°C), and then stored in liquid nitrogen. They can be thawed in warm water, poured into a semisolid medium, and once again grown as a cell suspension in liquid within a week or two. This approach is especially relevant for storage of vegetatively propagated species and species whose seeds are short-lived. Both shoot meristem cultures and cell suspension cultures have been advocated as appropriate candidates for cryopreservation. The approach has been most successful when applied to cell suspension cultures. Cryopreservation works for years, but it is not known if it will be suitable for storage over decades. *See* CRYOBIOLOGY.

Several elaborate procedures have been devised to introduce foreign deoxyribonucleic acid (DNA) into plant cells. One widely applicable method involves cloning desirable genes into an *Agrobacterium*-derived plasmid vector, introducing this vector into *Agrobacterium*, and then allowing the *Agrobacterium* to infect an excised piece of plant tissue. As part of its normal infection process, the bacterium inserts the cloned genes into plant cells for integration into the plant's chromosomes. The transformed plant cells can be selected to form callus tissue from which transformed plants can be regenerated, allowing direct introduction of desirable genetic traits. *See* GENETIC ENGINEERING.

Finally, plant protoplasts have provided unique opportunities to study both the plasma membrane and relatively undamaged intracellular organelles. Purified plasma membranes, tonoplasts, and intact vacuoles are routinely isolated from protoplasts for study of their transport properties. *See* TISSUE CULTURE.

Karen Grady Ford

Bibliography. L. C. Fowke and F. Constabel (eds.), *Plant Protoplasts*, 1985; H. T. Hudson, D. E. Kester, and F. Davies, *Plant Propagation: Principles and Practices*, 5th ed., 1990; N. F. Jensen, *Plant Breeding Methodology*, 1988; N. C. Stoskoph, D. T. Tomes, and B. R. Christie, *Plant Breeding: Theory and Practice*, 1993; I. K. Vasil (ed.), *Cell Culture and Somatic Cell Genetics of Plants*, 3 vols., 1984–1986; I. K. Vasil and T. A. Thorpe (eds.), *Plant Cell and Tissue Culture*, 1994.

Plant reproduction

The formation by a plant of offspring that are either exact copies or reasonable likenesses. When the process is accomplished by a single individual without fusion of cells, it is referred to as asexual; when fusion of cells is involved, whether from an individual or from different donors, the process is sexual.

Asexual reproduction. The higher green plant has remarkable powers of regeneration, as shown by the fact that almost all of its approximately 10 trillion cells are capable of reproducing the entire plant

asexually. This can be demonstrated by the technique of tissue culture, in which individual cells of the plant are excised and grown in sealed containers on artificial media in the absence of contaminating microorganisms. When the experiment is properly done, a single isolated cell derived from root, stem, or leaf divides repeatedly to produce a mass of undifferentiated tissue. With proper chemical stimulation, such tissue can give rise to roots and buds, thus reconstituting a plant. Alternatively, the tissue may produce many embryos, each of which can yield a plant exactly like the original. Such plants regenerated from a single cell can usually flower and set seed normally when removed and placed in soil. This experiment shows that each cell of the plant body carries all the information required for formation of the entire organism. *See* TISSUE CULTURE.

The culture of isolated cells or bits of tissue thus constitutes a means of vegetative propagation of the plant and can provide unlimited copies identical to the organism from which the cells were derived. This use of isolated cells as a basis for producing asexual propagules in tissue culture has become the basis of the genetic engineering of plants. Genes isolated from the deoxyribonucleic acid (DNA) of any source, be it plant, microbe, or animal, can be inserted into an appropriate vector for transfer into a single cell to be cultured. Examples of vectors are viruses or bacteria of the genus *Agrobacterium*, which can enter cells readily and transfer part or all of their DNA [or ribonucleic acid (RNA)] into the genome of the receptor cell. In some of the infected cells, the foreign genes become stably incorporated and expressed in the host genome, thus yielding transformed or transgenic plants. Some valuable new plant types, carrying transgenes for resistance to disease or for greater hardiness, productivity, or nutritional quality, are already at hand in the laboratory and ready for introduction into commercial agriculture. In 2006, approximately 100 million hectares of transgenic corn, cotton, and soybeans were planted, mainly in the United States, China, Argentina, and Brazil. There is evidence that through *Agrobacterium* and perhaps some viruses, such gene transfer can occur naturally, without artificial intervention. *See* GENETIC ENGINEERING.

All other vegetative reproductive devices of higher plants are elaborations of this basic ability and tendency of plant cells to produce tissue masses that can organize into growing points (meristems) to yield the typical patterns of differentiated plant organs. For example, a stem severed at ground level may produce adventitious roots, a property used commercially in vegetative propagation. Similarly, the lateral buds formed along stems by action of the terminal meristem, can, if excised, give rise to entire plants. The "eyes" of the potato tuber, a specialized fleshy stem, are simply buds used in vegetative propagation of the crop. In many plants, cuttings made from fleshy roots can similarly form organized buds and reconstitute the plant by vegetative propagation. In a few specialized plants such as *Kalanchoë*, fully developed embryos are formed at the margins of leaves. When removed from the leaf and placed in soil, these embryos produce new plants. Thus, each of the vegetative organs of the plant (leaf, stem, and root) can give rise to new plants by asexual reproduction. Sometimes, the plant organ is structurally modified for more efficient vegetative reproduction: Stems can be modified into elongated, horizontal stolons that root at each node, into fleshy storage tubers, into horizontal rhizomes, or into fleshy, erect corms. Leaf buds are sometimes modified into bulbs (**Fig. 1**). Sometimes, as in tumbleweeds, entire plants or large portions of plants break from their original location to produce plants in a new location. *See* PLANT PROPAGATION.

Sexual reproduction. While in asexual reproduction, the genetic makeup of the progeny rarely differs greatly from that of the parent, the fusion of cells in sexual reproduction can give rise to new genetic combinations, resulting in new types of plants, some of which may be better able to cope with the daily environmental stresses. The life cycle of higher green plants consists of two distinct generations, based on the chromosomal complement of their cells. The sporophyte generation is independent and dominant in the flowering plants and ferns, but small, nongreen, and dependent in the mosses, and contains the $2n$ number of chromosomes. The diploidy results in each case from the fusion of sperm and egg to form the zygote, which then develops into an embryo and finally into the mature sporophyte, whatever its nature. The sporophyte generation ends with the formation of $1n$ spores by reduction division, or meiosis, in a spore mother cell. The spore then develops into the gametophyte generation, which in turn produces the sex cells, or gametes. The gametophyte generation ends when gametes fuse to form the zygote, restoring the $2n$ situation typical of sporophytes. *See* MEIOSIS.

In flowering plants (**Fig. 2**), the gametophyte or $1n$ generation is reduced to just a few cells (generally three for the male and eight for the female). The male gametophyte is formed after meiosis occurs in the microspore mother cells of the anther, yielding a tetrad of $1n$ microspores. Each of these microspores then divides mitotically at least twice. The first division produces the tube nucleus and the generative nucleus. The generative nucleus then divides again to produce two sperms. These nuclei are generally not separated by cell walls, but at this stage the outer wall of the spore becomes thickened and distinctively patterned—a stage typical of the mature male gametophyte, the pollen grain. *See* FLOWER; MITOSIS; POLLEN; POLLINATION.

Mature walls of pollen grains are constructed of a uniquely resistant material called sporopollenin, which can remain intact almost indefinitely while other materials are being degraded by physical forces of the environment or microorganisms. The patterns of sporopollenin walls are distinctive for each species, permitting ready identification. Cores taken from the bottoms of lakes can frequently be used to reconstruct ancient flora through analysis of their pollen residues.

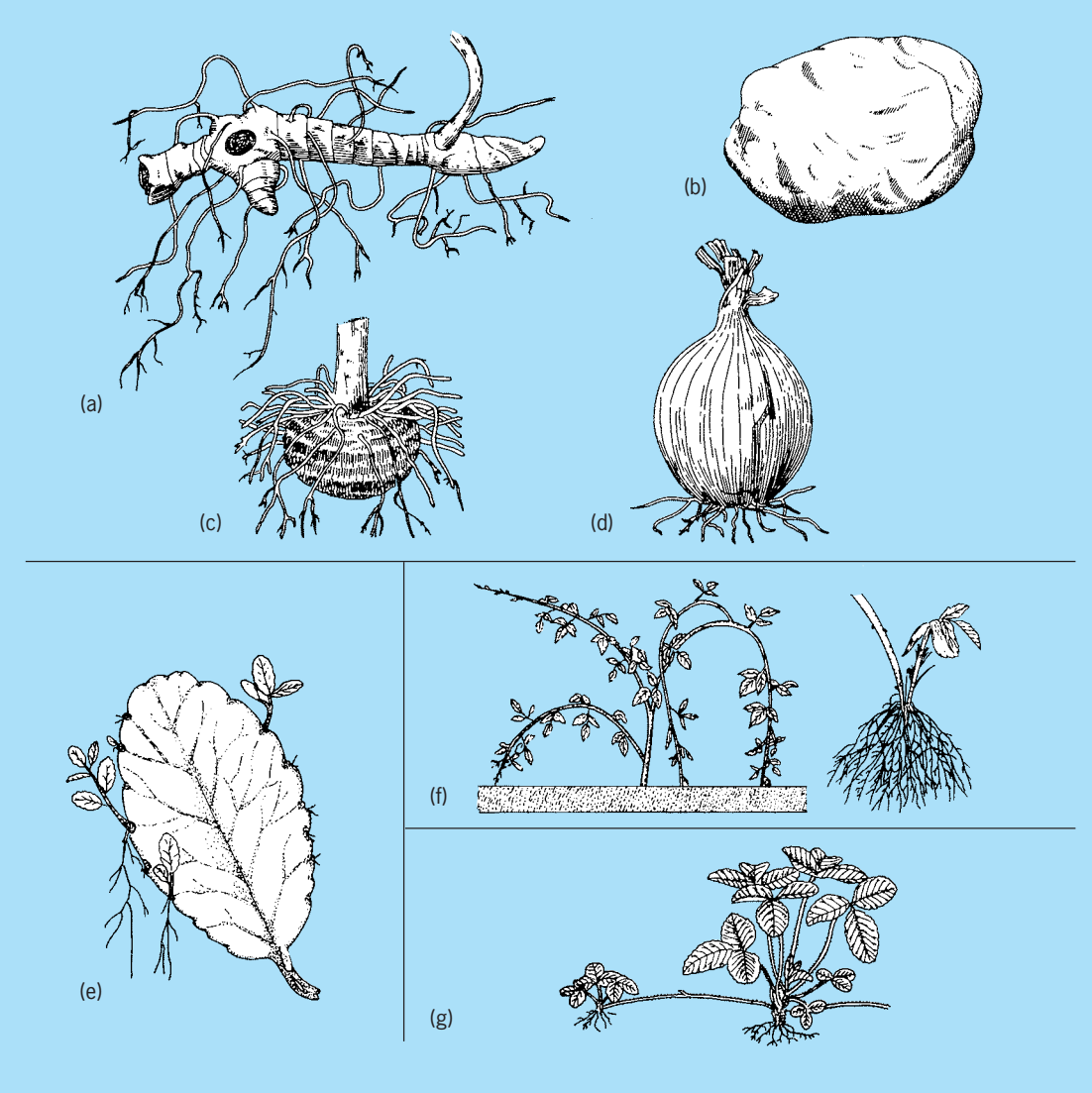

Fig. 1. Asexual reproduction by modification of the vegetative organs of the plant. (*a*) Rhizome (Solomon's seal); (*b*) tuber (potato); (*c*) corm (jack-in-the-pulpit); (*d*) bulb (onion) (*after E. L. Core, Plant Taxonomy, Prentice-Hall, 1955*). (*e*) Embryos formed on leaf margin (*Bryophyllum*) (*after J. B. Hill, L. O. Overholts, and H. W. Popp, Botany: A Textbook for Colleges, 2d ed., McGraw-Hill, 1950*). (*f*) Rooting of cuttings (black raspberry) (*after E. N. Transeau, H. C. Sampson, and L. H. Tiffany, Textbook of Botany, rev. ed., Harper, 1953*). (*g*) Stolon producing new plant at tip (strawberry) (*after E. L. Core, Plant Taxonomy, Prentice-Hall, 1955*).

Each pollen grain has a weak pore in the sporopollenin wall, through which the pollen tube emerges at the time of germination. Pollen germinates preferentially in the viscous secretion on the surface of the stigma, and its progress down the style to the ovary is guided through specific cell-to-cell recognition processes. Throughout its growth, which occurs through the deposition of new cell wall material at the advancing tip, the pollen tube is controlled by the tube nucleus, usually found at or near the tip. When the pollen tube, responding to chemical signals, enters the micropyle of the ovule, its growth ceases and the tip bursts, discharging the two sperms into the embryo sac, the female gametophyte of the ovary.

The female gametophyte generation, like the male, arises through meiotic division of a $2n$ megaspore mother cell. This division forms four $1n$ megaspores, of which three usually disintegrate, the fourth developing into an eight-nucleate embryo sac by means of three successive mitotic divisions. The eight nuclei arrange themselves into two groups of four, one at each pole of the embryo sac. Then one nucleus from each pole moves to the center of the embryo sac. One of the three nuclei at the micropylar end of the embryo sac is the female gamete, the egg, which fuses with one of the sperm nuclei to form the zygote, the first cell of the sporophyte generation, which produces the embryo. The second sperm fuses with the two polar nuclei at the center of the embryo sac to form a $3n$ cell that gives rise to the endosperm of the seed, the tissue in which food is stored. The entire ovule ripens into the seed, with the integuments forming the protective seed coat. The entire ovary ripens into a fruit, whose color, odor, and taste are attractive to animals, leading to dispersal of the seeds. The life cycle is completed

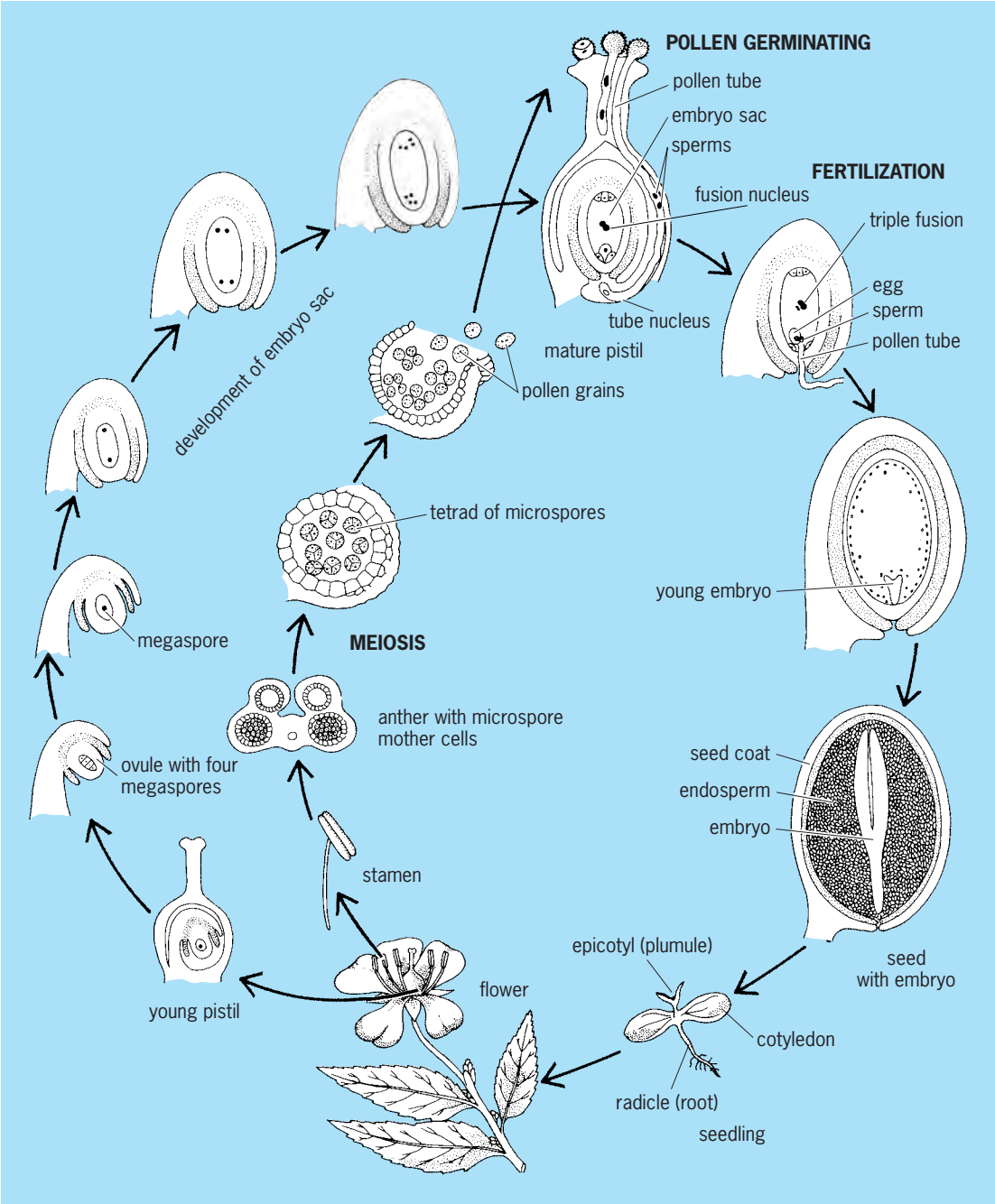

Fig. 2. Sexual cycle of a flowering plant. Self-pollination is depicted, but the pollen may also come from another individual. (*After E. W. Sinnott and K. S. Wilson, Botany: Principles and Problems, 6th ed., McGraw-Hill, 1962*)

when the seed germinates and grows into a mature sporophyte with flowers, in which meiotic divisions will once again produce $1n$ microspores and megaspores.

Nonflowering higher plants such as the ferns and mosses also show a distinct alternation of generations. The familiar fern plant of the field (**Fig. 3**) is the sporophyte generation. Meiosis occurs in sporangia located in special places on the leaves, generally the undersides or margins. A spore mother cell produces a tetrad of $1n$ spores, each of which can germinate to produce a free-living, green gametophyte called a prothallus. On the prothallus are produced male and female sex organs called antheridia and archegonia, which give rise to sperms and eggs, respectively. Sperms, motile because of their whip-like flagella, swim to the archegonium, where they fertilize the egg to produce the zygote that gives rise to the sporophyte generation again.

In mosses, by contrast, the dominant green generation is the gametophyte (**Fig. 4**). Antheridia or archegonia are borne at the tips of these gametophytes, where they produce sperms and eggs, respectively. When suitably wetted, sperms leave the antheridium, swim to a nearby archegonium, and fertilize the egg to produce a $2n$ zygote that gives rise

to a nongreen, simple, dependent sporophyte. The moss sporophyte consists mainly of a sporangium at the end of a long stalk, at the base of which is a mass of tissue called the foot, which absorbs nutrients from the green, photosynthetic gametophyte. Meiosis occurs in the sporangium when a spore mother cell gives rise to four reduced spores. Each spore can germinate, giving rise to a filamentous structure from which leafy gametophytic branches arise, completing the life cycle.

Various members of the algae that reproduce sexually also display alternation of generations, producing sperms and eggs in antheridia and oogonia. Sporophyte and gametophyte generations may each be free-living and independent, or one may be partially or totally dependent on the other. *See* FRUIT; POPULATION DISPERSAL; SEED.

Initiation of flowers. In many plants, the terminal meristem, which usually produces only vegetative organs, produces a flower through processes regulated by the length of day, a phenomenon known as photoperiodism. Because of the inclination of the Earth's axis of rotation with respect to the plane of its rotation about the Sun, the length of the day changes at all latitudes other than the Equator as the Earth rotates around the Sun, and the amplitude of the change is greater the farther the location is from the Equator. At the latitude of New York City, for example, the maximum daylength is about 15 h and the minimum about 9 h. This change of about 6 h is accomplished over a 6-month period (2 min per day on average). *See* PHOTOPERIODISM.

Photoperiod. There are three general photoperiodic classes of plants: short-day plants, long-day plants, and day-neutral plants. Short-day plants flower only when the number of hours of light each day is less than a certain critical value, and long-day plants only when this photoperiod exceeds a certain critical value. The flowering of day-neutral plants does not respond to photoperiod, and flowers form automatically after a certain number of nodes have been laid down. In short-day and long-day plants, the length of the uninterrupted dark period is the crucial determinant of behavior and even a brief flash of light in the middle of the night can interrupt flowering in short-day plants and promote flowering in long-day plants. The pigments that perceive this light are phytochrome, a bluish protein bound to an open-chain tetrapyrrole chromophore that resembles bile pigments and cryptochrome, a flavoprotein containing vitamin B_2. Phytochrome occurs in all higher green plants, where it regulates many important physiological processes, including seed germination, stem and leaf growth, chloroplast maturation, flowering, and senescence. *See* PHYTOCHROME.

The absorption of light by phytochrome produces many chemical and physical changes within the cell, including the properties of the membranes surrounding the cell and the activities of genes within the nucleus and chloroplast. It is not known which of these changes leads to flowering. There is, however, evidence that the signal from the light-activated phytochrome must interact with a rhythmic process

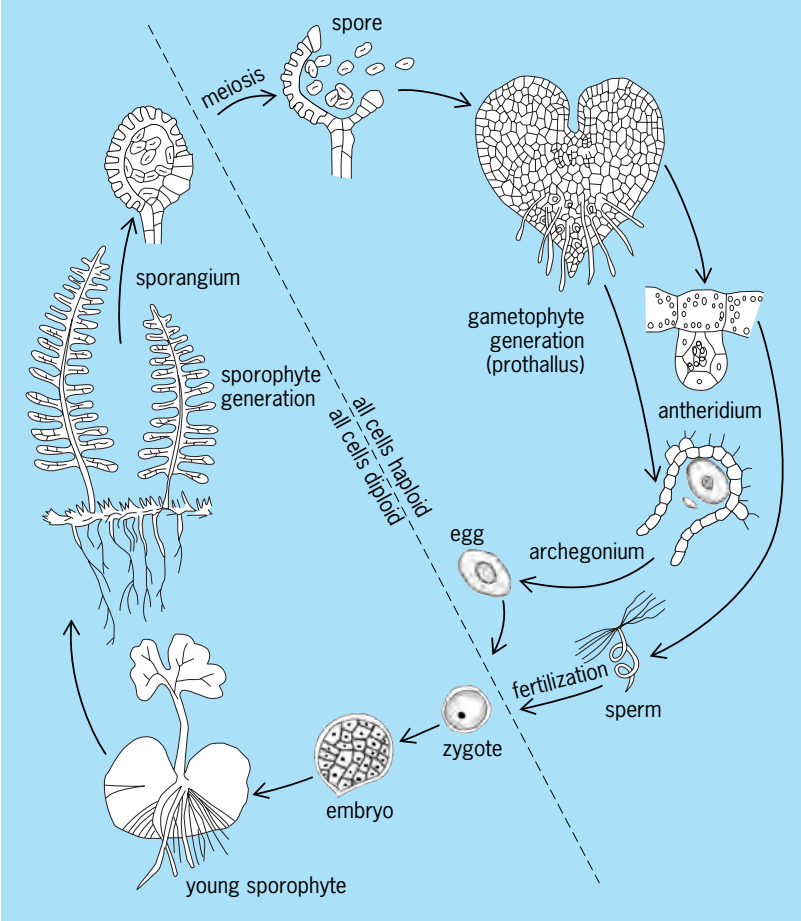

Fig. 3. Life cycle of a typical fern. Fertilization of the egg by sperm produces the first 2*n* cell, the zygote, which completes the life cycle by growing into the embryo and then the mature sporophyte. (*After A. W. Galston, P. J. Davies, and R. L. Satter, The Life of the Green Plant, 3d ed., Prentice-Hall, 1980*)

showing circadian (about 24 h) periodicity. The signal from the phytochrome probably initiates the rhythm and is also able to reset it, thus furnishing the elements of a self-timing system. The differences between long-day plants and short-day plants appear to involve the manner in which phytochrome signals interact with rhythmic processes. *See* BIOLOGICAL CLOCKS.

Some plants with a biennial habit are insensitive to photoperiod during their first year but become sensitive after they have experienced a winter season or have been exposed to low temperatures of about 41°F (5°C) for approximately 1 month in the laboratory. Plants that are made ready to respond to photoperiod by such low-temperature treatments are said to have been vernalized. Some important crops, such as winter wheat, must undergo vernalization in field or laboratory before they can initiate floral primordia in response to long days. *See* VERNALIZATION.

Gibberellins. One of the well-known plant hormones, gibberellin, can substitute for vernalization in some plants. Gibberellin-treated long-day plants not only are ready to receive the long-day stimulus but also sometimes "bolt," that is, they produce the elongated stem that characteristically bears the

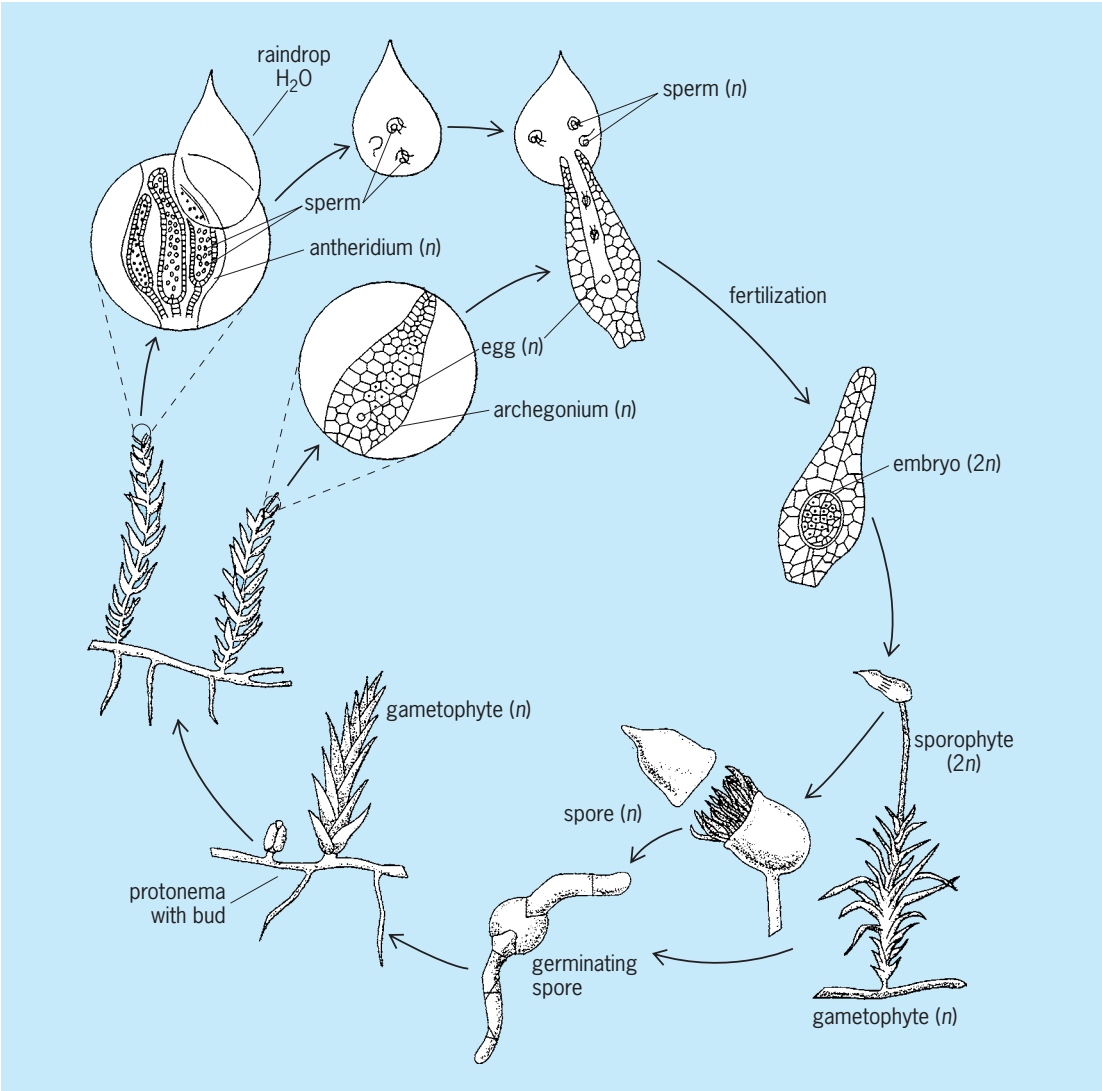

Fig. 4. Life cycle of a typical moss. Meiosis in the sporangium produces reduced spores whose germination produces a new gametophyte. (*After W. K. Purves, G. H. Orians, and H. C. Heller, Life, Sinauer, 1992*)

inflorescence. Low concentrations of gibberellins without subsequent long-day treatment may induce bolting without flowering, and higher concentrations sometimes induce flowering even without photoperiodic treatment. For this reason, gibberellins are presumed to be involved somehow in the transition to flowering in long-day plants, but there is no such evidence linking gibberellin to flowering in short-day plants. *See* GIBBERELLIN.

Florigen. More than a half-century ago, investigators proposed the existence of a specific floral hormone, which they named florigen. The proposal came from the observation that grafting a flowering plant to a nonflowering plant could induce the latter to flower (**Fig. 5**). The stimulus moving from the donor to the receptor seemed to travel through living vascular cells of the plant's phloem tissue. Even within a single photoperiodically sensitive plant, it is known that the leaf perceives the photoperiod whereas the bud responds to it. Some stimulus (florigen) must therefore travel from leaf to bud to induce the latter to make reproductive, rather than vegetative, organs. In some plants that are very sensitive to photoperiod, such as the short-day plants *Xanthium* (cocklebur) or *Pharbitis* (Japanese morning glory), a single unbroken dark period of sufficient length causes the plant to flower irrespective of subsequent treatment. Plants so altered are said to be induced, and leaves of induced plants continue to export the flowering stimulus even after transfer to photoperiods unfavorable to floral initiation.

The florigenic stimulus can be transferred from short-day plants to long-day plants, and vice versa, through grafts, indicating that florigen is the same or at least functionally equivalent in short-day and long-day plants. Although certain chemical substances, ranging from gibberellin to salicylic acid, have been reported to cause flowering in some plants under particular conditions, no known substance satisfies all of the requirements for florigen. However, investigators continue to explore reports of natural florigenic activity in plant extracts, including those from flowering short-day tobacco that seem to cause floral initiation in *Chenopodium*, a tiny short-day weed.

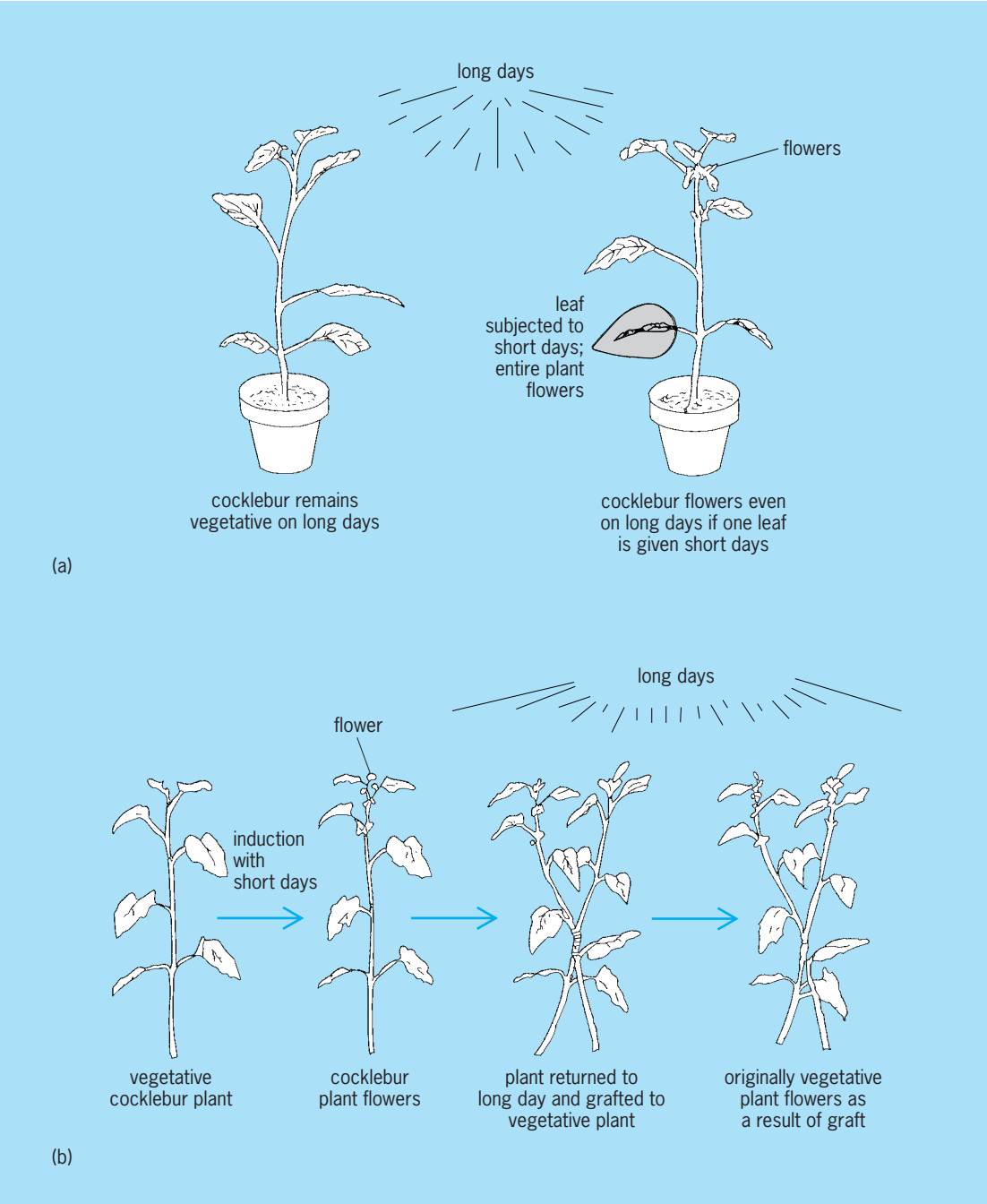

Fig. 5. Transport of the flowering stimulus. (*a*) From leaf to bud in a single plant. (*b*) Across a graft union between plants. (*After A. W. Galston, P. J. Davies, and R. L. Satter, The Life of the Green Plant, 3d ed., Prentice-Hall, 1980*)

Recent evidence indicates that florigen is a specific kind of mobile RNA.

Flowering threshold. Tissue culture techniques have been used to explore the control of flowering in tobacco. When thin surface layers of the flowering stalk of tobacco are grown in tissue culture, they can produce either vegetative or floral buds, depending on the hormonal composition of the nutrient medium. If the plant hormones auxin and cytokinin are both present in micromolar concentration, flower buds are produced, but if cytokinin is raised to 10 micromolar, only leafy buds are produced. Under these threshold conditions for flowering, the triamine spermidine, a naturally occurring plant constituent abundant in flower buds and flower parts, can cause flowering if added to the vegetative medium. Contrariwise, the prevention of spermidine accumulation on a flowering medium by a chemical inhibitor of spermidine synthesis prevents flowering, in a concentration-dependent fashion. While these observations compel an examination of the thesis that flowering in tobacco may be controlled by spermidine (or cytokinin), the internal biochemical events controlling flowering are still unknown. *See* AUXIN; CYTOKININS; PLANT PHYSIOLOGY.

Arthur W. Galston; Vivian F. Irish

Bibliography. A. W. Galston, *Life Processes of Plants*, 1994; A. H. Halevy, *Handbook of Flowering*, vols. 1–6, 1985–1988; W. S. Hillman, *The Physiology of Flowering*, 1962; E. Lord and G. Bernier, *Plant Reproduction: From Floral Induction to Pollination*, 1989; P. H. Raven, R. F. Evert, and H. Curtis, *Biology of Plants*, 6th ed., 1998; B. Thomas and D. Vince-Prue, *Photoperiodism in Plants*, 2d ed., 1996.

Plant respiration

The metabolism of organic molecules using enzymes to generate usable energy in the form of adenosine triphosphate (ATP). The ATP made by respiration is used in plant cells to move things within the cell, to make new molecules needed for growth, and to power active transport (the pumping of ions or other substances across a cell membrane against an osmotic gradient). Most plant respiration is aerobic—that is, it requires the presence of oxygen—and involves the oxidation of sugars to carbon dioxide. When oxygen is not present, anaerobic respiration, or fermentation, occurs, allowing some ATP synthesis. Fermentation is much less efficient in harvesting energy from organic molecules than aerobic respiration is, and most plants do not tolerate long periods of anaerobiosis. Plant respiration encompasses several pathways, including glycolysis, the Krebs cycle, and oxidative electron transport. It occurs in all plant cells and is generally similar to animal respiration, with a few notable exceptions such as a cyanide-insensitive respiration mechanism lacking in animals. *See* BIOLOGICAL OXIDATION; CARBOHYDRATE METABOLISM; CITRIC ACID CYCLE; ENERGY METABOLISM; PLANT METABOLISM.

A major misconception about plant respiration is that it occurs only in plant cells lacking photosynthesis. Moreover, photosynthesis is often thought of as the *only* means for plants to generate energy. All plant cells respire. During photosynthesis, respiration may be reduced and cells may gain some ATP from photosynthetic reactions, but measurements indicate that there still is a significant amount of respiration in photosynthesizing cells. *See* PHOTOSYNTHESIS; PLANT CELL.

Major steps. Glycolysis (**Fig. 1**) is a major component of both aerobic and anaerobic plant respiration. Glycolysis is the process by which glucose is broken down to form pyruvate, with energy captured in the form of ATP. It occurs in the cytosol of plant cells. Formally, glycolysis begins with glucose, but in many situations an intermediate of glycolysis is produced and the process begins with that product. For example, phosphorolytic breakdown of starch results in glucose 1-phosphate, which is converted to glucose 6-phosphate, the second compound of the glycolytic pathway. Similarly, during photosynthesis, dihydroxyacetone phosphate—one of the three carbon sugar phosphates (triose phosphates) that are part of glycolysis—is exported from chloroplasts and can undergo glycolytic reactions in the cytosol of photosynthetic plant cells. There are other pathways

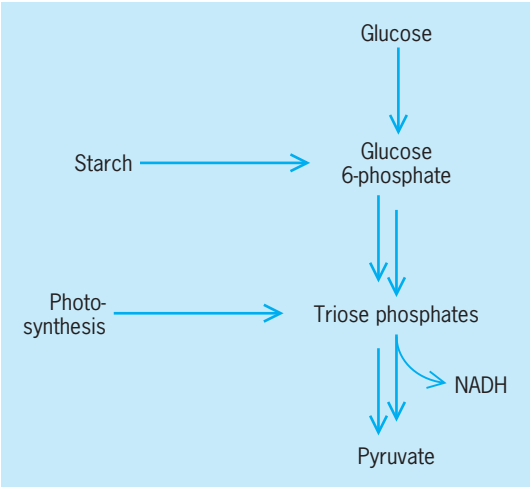

Fig. 1. Glycolysis is the conversion of glucose to pyruvate in 13 steps (not all are shown). Many metabolic processes can feed into the glycolytic pathway.

which can convert glucose to triose phosphates, notably the pentose phosphate pathway. In plants, however, it is likely that the pentose phosphate pathway is used primarily in photosynthesis and to generate intermediates for new plant constituents such as ribose for RNA, and erythrose 4-phosphate, a starting material for several essential amino acids, among other things.

Provided that oxygen is present, the pyruvate produced by glycolysis enters the mitochondria of plant cells (**Fig. 2**). There, it loses one carbon atom and is converted to acetyl coenzyme A. The two carbons of the acetate component of acetyl CoA enter the Krebs cycle (**Fig. 3**), also called the citric acid cycle or the tricarboxylic acid cycle. In one complete cycle, two carbon atoms are lost and significant reducing

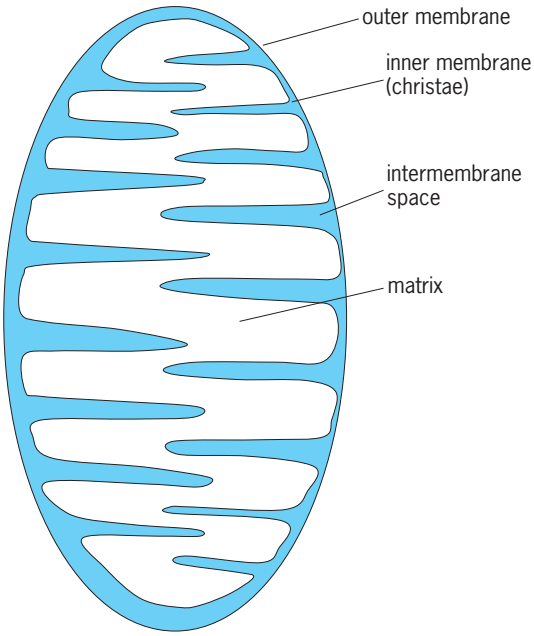

Fig. 2. Structural features of the mitochondrion.

power is stored as the reduced form of nicotinamide adenine dinucleotide (NADH) and the similar molecule flavin adenine dinucleotide ($FADH_2$). *See* MITOCHONDRIA.

The reducing power of NADH then can be used for oxidative electron transport on the inner membrane of the mitochondrion (the christae). Electrons from NADH are donated to carriers embedded in a large membrane-spanning complex (complex I). Next, a pair of electrons move to ubiquinone to make the reduced form ubihydroquinone (or ubiquinol). In this step, protons (H^+) are taken up from the matrix. Ubihydroquinone moves laterally through the inner mitochondrial membrane to the cytochrome *b/c* complex, also known as complex III. Here the electrons are donated to a cytochrome embedded in the protein complex, and the hydrogen ions are given off to the intermembrane space of the mitochondrion. At this stage every other electron cycles back through ubiquinone in the Q-cycle, which increases the number of hydrogen ions moved from the matrix to the intermembrane space. From the cytochrome complex, the electrons move to cytochrome *c*, a small protein that moves through the intermembrane space to the terminal oxidase, also known as cytochrome *a* oxidase, or complex IV. Oxidases are enzymes that catalyze oxidation reactions by the utilization using molecular oxygen as electron acceptors. Here the electron is combined with oxygen and hydrogen ions to form water. It is the high affinity for oxygen of electrons that provides the thermodynamic gradient for electron transport. *See* ENZYME.

Electron transport is coupled to ATP synthesis by the movement of protons from the matrix to the intermembrane space that happens at the ubiquinone step and during two additional steps. The energy gradient (proton motive force) resulting from the movement of hydrogen ions can be used to spin an ATP synthase complex on the inner mitochondrial membrane. This spinning forces adenosine diphosphate (ADP) to combine with inorganic phosphate (P_i) to make ATP.

Alternate oxidase. Plants have an alternate oxidase that allows electron transport directly from ubihydroquinone to oxygen. When electrons flow through this alternate oxidase, very few protons are moved across the inner mitochondrial membrane and so very little ATP is made. The energy normally captured as ATP is lost as heat. Plants use this ability to heat up. One of the best-known uses is by fly-pollinated flowers. To attract flies, these flowers use sulfur gases to simulate rotting flesh, which the flies favor for laying eggs. The compounds used include dimethyldisulfide and some volatile amines which have a strong unpleasant odor. (These plants have names that reflect this odor, such as skunk cabbage and the corpse flower.) The part of the flowering stalk that emits these chemicals can be as much as 10°C (18°F) above ambient air temperature. During the flowering process these plants are capable of high rates of alternate oxidase electron flow. Salicylic acid (the active agent of aspirin) signals the flower to start the thermogenic metabolism. However, all plants appear to have the capacity for alternate oxidase respiration. It is normally at low yet detectable levels, but can be as much as 50% of total respiration in some tissues. It is not clear why plants routinely use this less efficient respiratory pathway.

The alternate oxidase is not inhibited by cyanide, while the cytochrome oxidase is. The sensitivity of the cytochrome oxidase to cyanide accounts for the fatal consequences of cyanide for animals. In plants, because of the alternate oxidase, there is a capacity for cyanide-insensitive respiration. Many plants make

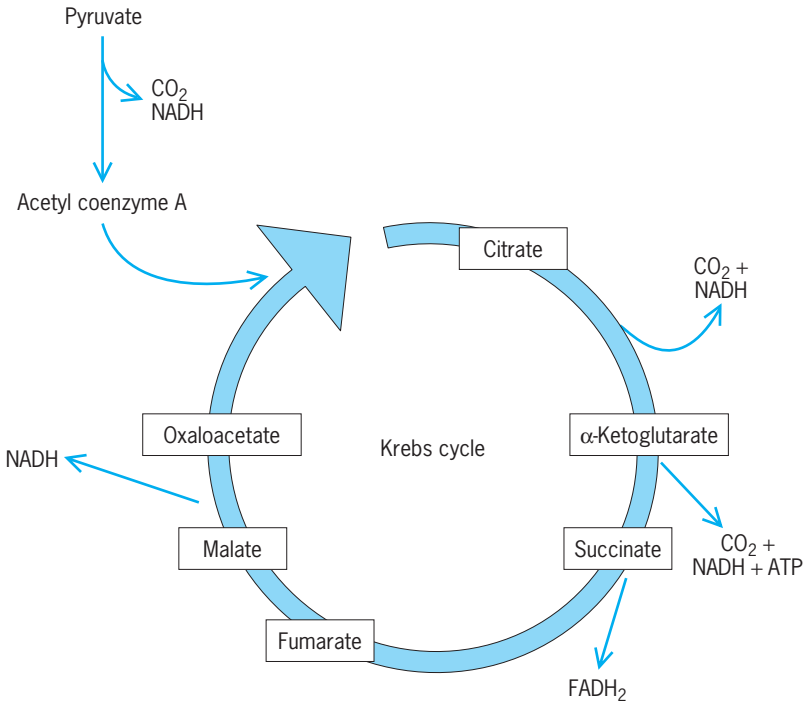

Fig. 3. Pyruvate is oxidized to CO_2 in the Krebs cycle.

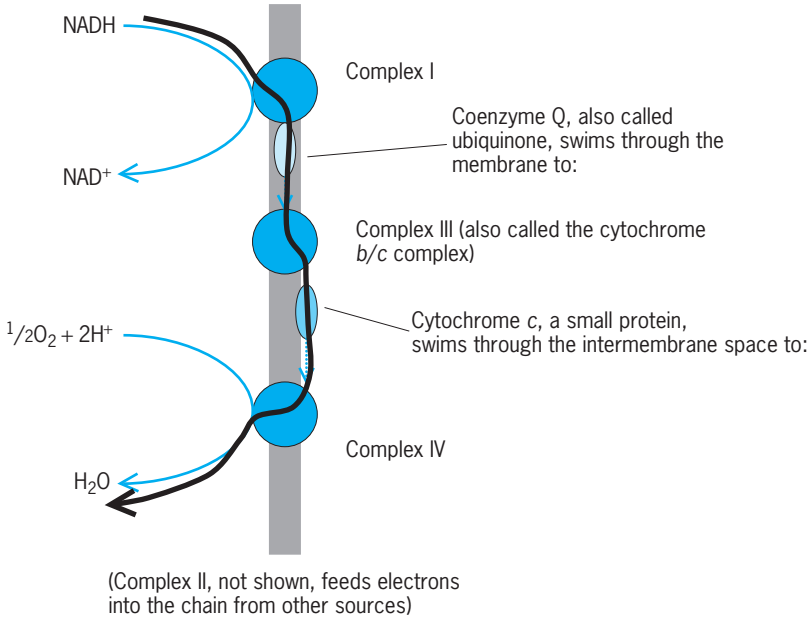

Fig. 4. Respiratory electron transport in plants can convert reducing power of NADH to ATP.

cyanogenic compounds that release cyanide under some conditions, for example when parts of the plant are eaten. This protects the plant from further herbivory, and any cyanide inadvertently produced in the plant from the cyanogenic compounds will not kill the plant because of its cyanide-insensitive respiration. *See* CYANIDE.

The alternate oxidase can be inhibited by salicylhydroxamic acid (SHAM). Early studies of the rates of the two major types of electron transport relied on inhibiting the cytochrome oxidase pathway with cyanide to measure the alternate oxidase pathway rate, and inhibiting the alternate oxidase pathway with SHAM to measure the cytochrome oxidase pathway. However, use of these inhibitors upsets the regulation of respiration, and such measurements are now considered unreliable. With these inhibitors, however, it has been possible to discover that the two oxidases have different effects on the oxygen isotope fractionation during respiration. The alternate oxidase discriminates against ^{18}O more than the cytochrome oxidase, and this can be used to estimate the proportion of alternate verses cytochrome oxidase activity at any given time. Generally, the more efficient cytochrome oxidase predominates.

Fermentation. When oxygen is not present, the reducing power harvested in glycolysis and the Krebs cycle cannot be used in electron transport and must be dissipated. The most common metabolic path in plants in this situation involves the enzymes pyruvate decarboxylase followed by alcohol dehydrogenase to make ethanol and CO_2, just as in yeast. A few plants, especially rice, are notable for their ability to make ethanol to survive periods of low oxygen. Oxygen is relatively insoluble in water and diffuses very slowly through water. The practice of flooding rice paddies allows for the killing of weeds because of the lack of respiration, while rice is tolerant of the low oxygen condition because of its high capacity for ethanol formation. Even so, rice will die if it is completely submerged because fermentation is not sufficient to meet the energetic needs even of rice. As long as part of the rice plant extends above the surface of the water in the paddy, a small layer of air lines the leaf (because of the hydrophobic nature of the leaf surface), and this allows a small amount of oxygen to diffuse to the roots of the rice plant. Often, plants that have their roots in water have specialized tissues called aerenchyma that provide air spaces inside the plant through which oxygen can diffuse. Thus, while fermentation can occur in plants, plants cannot live solely on fermentation the way that yeast can.

Climacteric. Many fruits undergo a burst of respiration as they ripen. This burst is called the climacteric and is triggered by the plant hormone ethylene (ethene). Climacteric fruits can stay just short of maturity for an extended period and then be induced to mature with a burst of respiration in response to the presence of ethylene. After the climacteric burst, fruits are often soft, easily damaged, and prone to rotting. Many fruits are shipped in containers scrubbed of ethylene and then ethylene is applied to induce ripening as needed.

Respiratory quotient. The number of oxygen molecules taken up for every CO_2 molecule released is called the respiratory quotient (RQ). For glucose, the RQ is 1, but the respiration of fats can lead to a RQ of more than 1. Organic acids such as malic acid and citric acid, on the other hand, produce RQs of less than 1. Measuring RQ can give some idea of the substrate sustaining respiration in a plant; however, there are many other ways that oxygen and CO_2 can be consumed or released. Thus, the RQ of intact plants is not a perfect measure of the source of carbon for respiration for that plant.

Glyoxylate cycle. Some plant parts, especially some seeds, convert fats and oils (triglycerides) to material for growth and energy to fuel that growth. During germination the triglycerides are broken apart, and the glycerol is converted to triose phosphate for metabolism in glycolysis. The three fatty acids are broken down by beta oxidation to yield acetyl CoA, and molecules that provide reducing power (one NADH and one $FADH_2$ for each acetyl CoA). The NADH and $FADH_2$ can be used to make ATP, but in some cases the reducing power is dissipated by donation to oxygen, eventually making hydrogen peroxide, which is broken down by catalase. This happens in specialized organelles within the plant cells called glyoxysomes. Glyoxysomes have substantial amounts of catalase to break down hydrogen peroxide and are similar to peroxisomes, which play a similar role in photorespiration.

The acetyl CoA can enter the Krebs cycle, but sometimes a Krebs cycle bypass occurs called the glyoxylate cycle. The glyoxylate cycle allows the carbon in acetyl CoA to be converted to sugars. *See* TRIGLYCERIDE.

Photorespiration. Photorespiration is the name given to the metabolism that occurs when oxygen acts as a substitute for CO_2 in the first step of photosynthesis. This results in an uptake of oxygen and release of CO_2, but consumes ATP rather than producing it. Therefore, photorespiration is a misnomer and is not a true form of respiration. Thomas D. Sharkey

Bibliography. J. S. Amthor, *Respiration and Crop Productivity*, Springer-Verlag, New York, 1989; H. Lambers and M. Ribas-Carbo, *Plant Respiration: From Cell to Ecosystem*, Springer, Dordrecht, 2005; H. Lambers and L. H. W. van der Plas, *Molecular, Biochemical and Physiological Aspects of Plant Respiration*, SPB Academic Publishing bv, The Hague, 1992; H. Öpik, *The Respiration of Higher Plants*, Edward Arnold Publishers, London, 1980.

Plant taxonomy

The area of study focusing on the development of a classification system, or taxonomy, for plants based on their evolutionary relationships (phylogeny). The assumption is that if classification reflects phylogeny, reference to the classification will help researchers focus their work in a more accurate manner. This principle is illustrated by the example of a mother and daughter, with a great deal being known about

the genetic traits of the mother. The genetic traits exhibited by the daughter can be largely predicted in this case, whereas for two people of unknown relationships no conjecture is possible regarding their genetic similarity. The task is to make phylogeny reconstruction as accurate as possible. How this phylogenetic system of classification is to be constructed has been subject to varying opinions over time. The basic unit of classification is generally accepted to be the species, but how a species should be recognized has been intensely debated. A historical perspective is useful in explaining the nature of this controversy. *See* PLANT KINGDOM; PLANT PHYLOGENY.

History. The earliest classifications of plants were those of the Greek philosophers such as Aristotle (384–322 B.C.) and Theophrastus (372–287 B.C.). The latter is often called the father of botany largely because he listed the names of over 500 species, some of which are still used as scientific names today. All of Theophrastus' plants were arranged into three categories: herbs, trees, and shrubs. This was a practical solution but one that was not based on an estimate of how closely related the species were. In the next 1600 years, as in most areas of science, little progress occurred in plant taxonomy. It was not until the fifteenth century that there was renewed interest in botany, much of which was propelled by the medical use of plants. Up to and including that time, medicines were taken directly from botanical sources (this is largely true even today). The Doctrine of Signatures stated that the appropriate medical use of plants was indicated by their shape or color; for example, plants such as the hepatic, of which the leaves were shaped like a liver, and the orchid, of which the underground tubers were shaped like testicles, were supposed to have been intended by God to treat liver disease and male infertility.

Historically if a plant was not of medical use, it was likely not classified at all. This was changed by Carolus Linnaeus, a Swedish botanist who in 1753 published his *Species Plantarum*, a classification of all plants known to Europeans at that time. Linnaeus's system was based on the arrangement and numbers of parts in flowers, and was intended to be used strictly for identification (a system now referred to as an artificial classification as opposed to a natural classification, based on how closely related the species are).

Binomial nomenclature. In *Species Plantarum*, Linnaeus made popular a system of binomial nomenclature developed by the French botanist Gaspard Bauhin (1560–1624). In this system, which is still in use, each species has a two-part name, the first being the genus and the second being the species epithet. For example, *Rosa alba* (italicized because it is Latin) is the scientific name of one species of rose; the genus is *Rosa* and the species epithet is *alba*, meaning white (it is not a requirement that scientific names be similar to common names or have real meaning, although such relevance is often the case). The genus name *Rosa* is shared by all species of roses, reflecting that they are thought to be more closely related to each other than to species in any other group.

Our ideas about classification have changed drastically since the time of Linnaeus, but we still use his system of binomial nomenclature today. The rules of botanical nomenclature (and those of zoology as well, although they are not identical) are part of an internationally accepted Code that is revised (minimally) at an international congress every 5 years. *See* PLANT NOMENCLATURE.

Natural classification system. In addition to his system of binomial nomenclature, Linnaeus introduced the Sexual System of classification, based on the number of male and female flower parts. The system was not widely used even in his day, with many of his French contemporaries preferring to use a system based more on assessments of how the plants were related (a more "natural" system). Their concept of relatedness was drastically different from that used today, due to the fact that Charles Darwin's theory of evolution (published in *Origin of Species*, 1859) came much later, but the eighteenth-century French botanists made major improvements in how plants were described and how the various categories above the level of genus were to be defined, and set us on a course to produce the natural taxonomic systems that took on its modern meaning after the time of Darwin. Today, we understand that the best classification system is one that reflects the patterns of the evolutionary processes that produced these plants. *See* PLANT EVOLUTION.

Modern taxonomic categories. Use of common names in science and horticulture is not practical because in each language and country common names will vary; for example, "bluebell" means something different in northern and southern North America as well as in England and Scotland. Scientific names are internationally agreed upon so that a consistent taxonomic name is used everywhere for a given organism. In addition to genus and species, plants are classified by belonging to a family (Rosaceae, the rose family, for example); related families are grouped into orders, and these are typically grouped into a number of yet higher and more encompassing categories. In general, higher categories are composed of many members of lower types—for example, a family may contain 350 genera, but some may be composed of a single genus with perhaps a single species if that species is distantly related to all others. An example is the genus of the sycamore or plane tree, *Platanus*, which is the only genus in its family, Platanaceae.

Many botanists use a number of intermediate categories between the level of genus and family, such as subfamilies, tribes, and subtribes, as well as some between species and genus, such as subgenera and sections, but none of these categories is formally mandated. They are useful nonetheless to reflect intermediate levels of relatedness, particularly in large families (composed of several hundreds or even thousands of species). Below the level of species, some botanists use the concept of subspecies (which is generally taken to mean a

geographically distinct form of a species) and variety (which is often a genetic form or genotype, for example a white-flowered form of a typically blue-flowered species, or a form that is ecologically distinct).

Phylogenetic classification. In the twentieth century, several approaches to constructing a natural system of classification were used. For many years, there were several competing systems that were generally similar in overall ideas of how plants were related but that differed in minor, though important, ways. Circumscription of many families has been highly consistent since the French botanists of the seventeenth century, but the grouping of families into orders has been much more variable. For example, the important agronomic family comprising the legumes (Fabaceae: beans, peas, lentils, and the like) has been presented as most closely related to several different families, such as Connaraceae, Polygalaceae, Rosaceae, and Saxifragaceae, depending on the author. The reason for these differing ideas was reliance on one or a few characteristics as the most important indicators of relationships. The different characters, such as pollen and floral morphology (shape and structure), growth form (tree, shrub, herb), and chemistry (whether alkaloids or steroids were present, for example), did not appear to reflect the same patterns of relatedness, and authors were likely to reach different conclusions based on their opinions of which characteristics were more important. *See* PLANT ANATOMY; PLANT PHYSIOLOGY.

The basic idea that plant classification should reflect evolutionary (genetic) relationships has been well accepted for some time, but the degree to which this could be assessed by the various means available (by studying morphology, chemistry, or chromosome numbers, for example) differed. It has only recently become possible to assess genetic patterns of relatedness directly by analyzing DNA sequences. Although most botanists have supported the idea that classification should summarize and synthesize all available information, this was rarely attempted, even though in the latter half of the twentieth century computerized methods became available to construct such syntheses. Part of the problem was the incompleteness of the information base available; some families of plants were much better studied than others (often due to the number of economically important species in a family), and completing these surveys was time-consuming and expensive. As a result, attempts to perform computerized analyses of large groups of plants, such as the flowering plants (angiosperms), were rare, and the missing data raised questions about the validity of these studies. The result was that plant taxonomy continued to be practiced as it had in the late nineteenth century, and the plethora of contradictory classifications frustrated scientists who wished to consider the evolutionary implications of their research.

DNA technology. In the 1990s, DNA technology became much more efficient and less costly, resulting in a dramatic upsurge in the availability of DNA sequence data for various genes from each of the three genetic compartments present in plants (nuclear, mitochondrial, and plastid or chloroplast). In 1998 a number of botanists collectively proposed the first DNA-based classification of a major group of organisms, the angiosperms or flowering plants. For the first time, a classification was directly founded on assessments of the degree of relatedness made with objective, computerized methods of phylogeny reconstruction. Other data, such as chemistry and morphology, were also incorporated into these analyses, but by far the largest percentage of information came from DNA sequences—that is, relatedness was determined mostly on the basis of similarities in plants' genetic codes. The advantages of such a classification were immediately obvious: (1) it was not based on intuition about which category of information best reflected natural relationships; (2) it ended competition between systems based on differing emphases; (3) the analysis could be repeated by other researchers using either the same or different data (other genes or categories of information); and (4) it could be updated as new data emerged, particularly from studies of how chromosomes are organized and how morphology and other traits are determined by the genes that code for them. *See* DEOXYRIBONUCLEIC ACID (DNA); GENETIC MAPPING.

Cladistics. At the same time that DNA data became more widely available as the basis for establishing a classification, a more explicit methodology for turning the results of a phylogenetic analysis into a formal classification became popular. This methodology, called cladistics, allowed a large number of botanists to share ideas of how the various taxonomic categories could be better defined. Although there remain a number of dissenting opinions about some minor matters of classification, it is now impossible for scientists to propose alternative ideas based solely on opinion. *See* TAXONOMY.

Mark W. Chase; Michael F. Fay

Bibliography. D. Bhattacharya (ed.), *Origin of Algae and Their Plastids*, Springer-Verlag, 1997; L. E. Graham and L. W. Wilcox, *Algae*, Prentice Hall, 2000.

Plant tissue systems

The tissues of a plant are organized into three systems: the dermal tissue system, the ground tissue system, and the vascular tissue system. Most plants are composed of coherent masses of cells called tissues. Large units of tissues having some features in common are called tissue systems. In actual usage, however, the terms tissue and tissue system are not strictly separated. A given tissue or a combination of tissues may be continuous throughout the plant or large parts of it. "Simple tissues" refer to tissues composed of one cell type. However, different cell types may be contained within a tissue, often referred to as a "complex tissue," and the same cell type may be part of different tissues.

Although classification of tissue systems may be based on structure or function, the two aspects

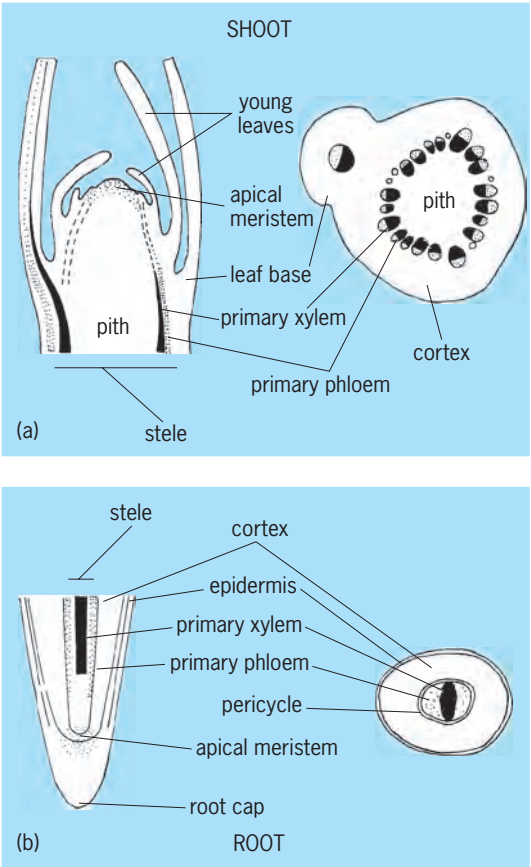

Fig. 1. Primary plant tissue systems in longitudinal and cross sections of (*a*) shoot and (*b*) root.

usually are combined. Plant tissues are primary or secondary in origin. The primary tissues arise from apical meristems, the perennially embryonic tissues at the tips of roots and shoots (**Fig. 1**). The primary tissues include the surface layer, or epidermis; the primary vascular tissues, xylem and phloem, which conduct water and food, respectively; and the ground tissues, comprising parenchyma, which is chiefly concerned with manufacture and storage of food, and collenchyma and sclerenchyma, which are the supporting tissues). *See* APICAL MERISTEM.

Properties and functions of a tissue are usually related to the structure of the cells present. A description of the structure of these primary tissues is detailed below. *See* PLANT ANATOMY; PLANT CELL; PRIMARY VASCULAR SYSTEM (PLANT).

Epidermis. The epidermis, usually one layer of closely packed cells, can vary enormously with many specializations of its cells, which usually relate to the environment. Many of these features reduce water loss from the plant (cuticle, thickened cell wall, and trichomes). Other structures, stomata (minute pores with two guard cells), regulate the passage of gases in and out. In the case of roots, epidermal areas just behind the root cap have small root hairs through which water and minerals from the soil are absorbed into the plant root. *See* EPIDERMIS (PLANT).

Ground tissues. In the ground tissue, parenchyma cells can be variable in shape and size, usually with thin cell walls, and sometimes arranged to maximize cell functions, forming a more specialized tissue. For example, they might be tightly or loosely packed cells (forming palisade or spongy mesophyll), or they may have large intercellular spaces (aerenchyma) as in aquatic plants. Where present, collenchyma and sclerenchyma have variably thickened cell walls, with differing wall chemistry, and these cells support the plant body. Collenchyma is usually found in bundles in younger plant tissues (prior to secondary growth occurring), whereas sclerenchyma is variable in form and lignified, found in discrete areas possibly associated with vascular tissues or more scattered, and is common in the secondary tissues. *See* CELL WALLS (PLANT); COLLENCHYMA; PARENCHYMA; SCLERENCHYMA.

Vascular tissues. Conduction of water and food through the xylem and phloem tissues is optimized by the predominant cell types. The water moves through groups of often long xylem elements which contain no cell contents and have pits in the cell walls. In some plants, the end walls of the xylem cells are broken down to form vessels, thereby aiding the water flow. The exact structure and size of these cells varies between plant groups. Food is transported by phloem tissue, which consists of pipelike conducting elements and smaller companion cells that actively regulate movement of food through the plant. *See* PHLOEM; XYLEM.

In the stem and root, the vascular tissues and some associated ground tissue are often treated as a unit, the stele. Ground tissue may be present in the center of the stele (pith) and on its periphery (pericycle). In the root, the meristematic pericycle produces lateral roots. The ground tissue system enclosing the stele on the outside is called the cortex. *See* CORTEX (PLANT); PERICYCLE; PITH.

Immediately outside the stele, the endodermis is the innermost part of the cortex consisting of cells that help regulate the passage of water into the vascular system, as a layer of cork-like material (casparian strip) compels water to enter the cells. At the cortex periphery, there may be a hypodermis derived from the cortex cells immediately to the inside of the epidermis. Where present, it often consists of sclerenchyma cells and is called an exodermis. *See* ENDODERMIS; HYPODERMIS.

Secondary tissues. The secondary tissues (**Fig. 2**) arise from lateral meristems (vascular and cork cambiums), and their formation is mainly responsible for the growth in thickness of stems and roots. They comprise secondary vascular tissues and the protective tissue called periderm. Secondary tissue development varies between the major plant groups and between species. In dicotyledons it is very extensive, but in monocotyledons lateral meristem activity is very limited; a few plants have either diffuse secondary growth or a secondary thickening meristem, which leads solely to new vascular tissue. *See* LATERAL MERISTEM.

In dicotyledons and gymnosperms, vascular

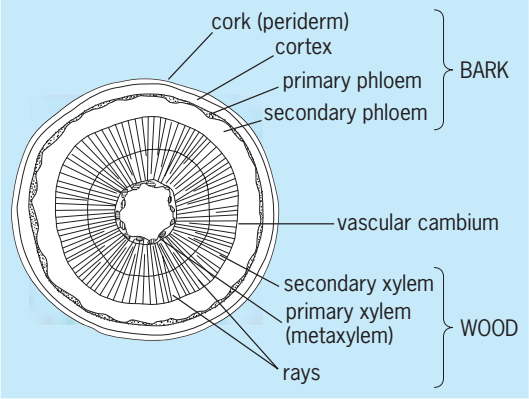

Fig. 2. Cross section of typical dicotyledon stem, showing secondary tissues that arise from lateral meristems.

cambium, a thin layer of cells, actively divides and develops into xylem tissue to the inside and phloem tissue to the outside. Over time, this secondary growth may build up a massive core of wood (secondary xylem), but the outer tissue system–the bark (secondary phloem and other cells)–remains relatively thin because its outer or older part becomes compressed and, in many species, is continuously sloughed off. In many "woody" plants, the rate of cambial division and growth is not constant and can produce the distinctive growth rings, which are often annually produced in temperate plants or relate to periods of flowering in tropical plants. *See* BARK; WOOD ANATOMY.

Periderm is part of the bark, produced from the lateral meristem (the cork cambium or phellogen). This layer of cells actively divides, though at a slower rate than the vascular cambium, producing phelloderm to the inside and phellem to the outside. The phellem cells have walls that are thick and waxy, protecting the tissues beneath from the external environment. *See* PERIDERM.

Within the large mass of secondary vascular tissues, there are often specialized "tissue systems" that operate or function within particular dimensions. In woods, radially arranged parenchyma and tracheids form rays, which transport water between the center and the outside of the trunk, perpendicular to the main axial water and food transport route (along the stem). Some plants have specialized cells that secrete substances. The specialized cells can be well developed, forming extensive tissue systems. For example, some conifers have parenchyma cells that produce resin which collects in canals, creating interconnected axial and radial systems.

Recent advances in techniques such as computerized tomography and nuclear magnetic resonance are leading to a better understanding of the three-dimensional structure and functionality of some tissue systems, especially those transporting water and haustoria parasitizing host plants (haustoria are outgrowths of certain parasitic plants that serve to absorb food from the host).

The production of flowers instead of vegetative shoots results from physiological and morphological changes in the apical meristem, which then becomes the flower meristem. The latter, however, produces tissue systems fundamentally similar to those in the vegetative body of the plant. *See* FLOWER.

Anna H. Lynch; Katherine Esau

Bibliography. C. B. Beck, *Plant Structure and Development: Plant Anatomy for the 21st Century*, Cambridge University Press, 2005; B. G. Bowes, *A Colour Atlas of Plant Structure*, Manson Publishing, London, 1996; R. F. Evert and S. E. Eichhorn, *Esau's Plant Anatomy: Meristems, Cells, and Tissues of the Plant Body: Their Structure, Function, and Development*, 3d ed., Wiley, Hoboken, NJ, 2006; A. Fahn, *Plant Anatomy*, 4th ed., Pergamon Press, New York, 1990; J. D. Mauseth, *Botany: An Introduction to Plant Biology*, 3d ed., Jones and Bartlett, Sudbury, MA, and London, 2003; P. J. Rudall, *Anatomy of Flowering Plants*, 3d ed., revised, Cambridge University Press, 2007.

Plant transport of solutes

The movement of organic and inorganic compounds through plant vascular tissues. Transport can take place over considerable distances; in tree species, transport distances are often 100–300 ft (30–100 m).

This long-distance transport is necessary for survival in higher land plants, in which specialized organs of uptake or synthesis are separated by a considerable distance from the organs of utilization. Diffusion is not rapid enough to account for the amount of material moved over such long distances. Rather, transport depends on a flowing stream of liquid in vascular tissues (phloem and xylem) that are highly developed structurally.

The movement of organic solutes occurs mainly in the phloem, where it is also known as translocation. In the phloem the direction of transport is from places of production, such as mature leaves, to places of utilization or storage, such as the shoot apex or developing storage roots. Organic materials translocated in the phloem include the direct products of photosynthesis (sugars) as well as compounds derived from them (nitrogenous compounds and plant hormones, for example). The phloem also transmits signals between various plant organs in the form of regulatory molecules (such as RNAs), which move with the organic solutes.

Some movement of organic solutes does occur in the xylem of certain species. However, the xylem generally moves water and inorganic solutes or mineral elements from sites of uptake in the roots to sites where water is lost from the plant, primarily the leaves. Some redistribution of the ions throughout the plant may then occur in the phloem. *See* PLANT TISSUE SYSTEMS.

Organic Solutes

Organic compounds are found in both the phloem and xylem of higher plants; the relative importance of these tissues as conducting channels in organic transport depends on both the nature of the organic

compound transported and the site of origin of the compound in the plant.

Importance of phloem. The phloem is the predominant tissue for the translocation of organic solutes, particularly sugars. If the continuity of the phloem is interrupted by removal of a narrow ring of bark external to the xylem (gindling), then the bark on the side of the girdle nearest the leaves swells somewhat and remains healthy and the bark on the side of the girdle which is separated from the leaves dries up and eventually dies due to a lack of organic nutrients. Radioactive tracers also show that the cells of the phloem called sieve elements are the major transport pathway for the conduction of sugars and other organic nutrients synthesized in the leaves. *See* PHLOEM.

The nature of the organic substances translocated in the phloem has been determined primarily by chromatographic analyses of sieve-element exudates collected from incisions in the bark or from severed stylets of phloem-feeding aphids (**Fig. 1**). Although the phloem sap obtained by incision or aphid stylet is probably not identical to the contents of the undisturbed sieve element, it does give data consistent with radioactive tracer experiments and thus appears to provide a valid picture of the compounds in transit.

In most plant species, sugars form the bulk of the substances moved in the phloem, with sieve-element sap containing 10–25% sugar by weight. Sucrose is the dominant sugar in the sap of most species analyzed. However, some species translocate nonreducing oligosaccharides or sugar alcohols, or both, as well as sucrose. The oligosaccharides belong to the raffinose family and include raffinose, stachyose, and verbascose. These sugars form a series consisting of sucrose (a disaccharide) with one, two, or three galactose residues joined to form the tri-, tetra-, and pentasaccharides. The transported sugar alcohols are generally mannitol or sorbitol. Although the hexoses glucose and fructose are frequently abundant throughout the remainder of the plant, analysis of sap from many different species reveals a consistent absence of hexose sugars in the sieve elements. Other organic substances are found in the sieve-element sap in addition to sugars, but these are generally much less concentrated than the carbohydrates. Included are nitrogenous substances, particularly amino acids and amides, as well as plant hormones, vitamins, and sugar phosphates. Phloem sap also contains specific proteins and RNA molecules. The proteins are structural or play roles in stress and defense reactions. There is evidence that many of these RNA molecules act as signaling molecules between plant organs. *See* CARBOHYDRATE; OLIGOSACCHARIDE.

Role of xylem. Xylem sap may also contain organic substances, including nitrogenous compounds, carbohydrates, and plant hormones, especially cytokinins and gibberellins. In some species, the xylem constitutes an important channel for transport to the shoots of organic nitrogenous compounds synthesized in the roots. Carbohydrates transported from the leaves via the phloem provide the carbon "skeletons" for these compounds; the necessary inorganic nitrogen either is absorbed from the soil or is made available to the plant by nitrogen-fixing microorganisms. *See* NITROGEN FIXATION; PLANT HORMONES.

In other species, however, organic nitrogenous

Fig. 1. Phloem-feeding aphid. (*a*) Droplet of sieve-tube sap being released by the aphid *Longistigma caryae*. (*b*) Transverse section through the bark of *Tilia americana*, showing the tip of an aphid stylet bundle that had been exuding prior to sectioning. Note that the tips of the stylets are inside a single sieve-tube member. (*M. H. Zimmermann, Harvard Forest, Harvard University*)

Fig. 2. Movement of ^{14}C sucrose through a bean plant. Sources are indicated by S1–S5. $^{14}CO_2$ was supplied to the youngest fully mature leave (S2) of a bean plant for 2 h in the light. Radioactive sucrose was synthesized from the $^{14}CO_2$ by the leaf and translocated throughout the plant. The plant was then dismembered, dried, mounted on a piece of paper, and placed next to x-ray film for a few days. (*a*) Mounted bean plant. The arrows indicate movement from the treated leaf S2 to the various sinks. (*b*) Developed and printed x-ray film, with radioactivity revealed (white). (*D. R. Geiger, University of Dayton*)

solutes are more concentrated in the phloem sap than in the xylem sap, and xylem nitrogen is primarily in the form of nitrate. In these species, nitrate that has arrived in the leaves via the xylem stream is then incorporated into organic compounds. Redistribution of nitrogenous compounds from the leaves occurs mainly in the phloem in all species. During senescence (yellowing and progressive deterioration) of leaves, large amounts of nitrogen- and phosphorus-containing organic compounds are exported from the leaves in the phloem tissue.

During translocation through the phloem of the stem of a plant, a considerable quantity of sugar may move into lateral tissues, such as the cortex, phloem parenchyma, or xylem parenchyma cells, where it is stored as starch available for later remobilization. In spite of the proximity of the phloem and of carbohydrate reserves, only negligible quantities of sugar appear in sap from the xylem elements during most of the growing season; the average maximal concentration seldom exceeds 0.05%. During the dormant season, however, when the flow of water in the xylem elements practically ceases, sugars often accumulate in readily detectable amounts. As an extreme example, in sugar maple (*Acer saccharum*) a concentration as high as 8% has been reported in the spring before the leaves emerge. This sugary sap is used to make maple syrup. Even in this species, however, the phloem is still the most important channel for transport of sugars during the growing season. *See* XYLEM.

Source-sink relationships in phloem. In general, organic compounds are translocated in the phloem from areas called sources, which are regions of synthesis or mobilization, to areas of utilization called sinks, such as young leaves, root tips, shoot tips, and developing fruits. Mature photosynthesizing leaves are the major sources; materials can be transported from a given source leaf (S) both acropetally and basipetally, that is, in the direction of the shoot tips and root tips, respectively (**Fig. 2**). When multiple source leaves are present, the upper sources usually supply the shoot tips, while the lower sources supply the root tips. The distribution pattern from a leaf also depends on the pattern of vascular connections with other organs (**Fig. 3**). *See* PHOTOSYNTHESIS.

Early in its development, a leaf is dependent on the import of a significant quantity of organic material via the phloem. Not until the leaf has expanded to about one-quarter of its mature size does it begin to export the products of photosynthesis. Export starts in the tip region of the leaf and extends basipetally (toward the base) until the entire leaf is exporting. For a time, the blade is both importing and exporting organic material. The transition from sink to source is complete when the leaf is approximately half expanded. A mature leaf may export nearly half of the carbohydrate produced during photosynthesis; much of the rest will be accumulated as starch, which will contribute to translocated materials during darkness. Only small amounts are used daily for the metabolic needs of the leaf cells. *See* LEAF.

Allocation and partitioning of organic solutes. Two processes, allocation and partitioning, are important in the utilization of photosynthates. Allocation refers to the fate of photosynthetic products in both source and sink leaves. For example, in source leaves some of the sugar formed during photosynthesis is used to synthesize storage materials, such as starch, for use by the plant during the night period, and some is used to synthesize transport sugars, including sucrose, for translocation out of the leaf. Partitioning refers to the relative distribution of transport sugars among the various sink tissues of the plant. Both of these processes are important because they determine the proportion of photosynthetic product that will reach edible portions of the plant and because they are a factor in improved crop productivity.

Allocation. Carbon dioxide that is converted into organic compounds (sugars) during photosynthesis is said to be fixed during the procedure. The mechanisms that determine the fate of this fixed carbon throughout the plant constitute a complex and finely tuned system. Rates of sucrose and starch synthesis must be coordinated in source leaves, and part of the allocation between translocation and storage occurs via a transport protein, called the phosphate translocator, located in the membrane of the chloroplast. The process of sucrose synthesis in the cytoplasm of a source cell results in the release of inorganic phosphate in the cytoplasm. The phosphate

translocator exchanges phosphate in the cytoplasm for one of the early products of photosynthesis, triose phosphate from the chloroplast. In this way, carbon destined for export from the leaf exits the chloroplast, and phosphate required during photosynthesis enters the chloroplast. Whenever the demand for sucrose by other parts of the plant is high, the rate of sucrose synthesis in the source tends to be high, and more phosphate is released and exchanged for carbon from the chloroplast. This sequence of events diverts fixed carbon away from the synthesis of starch when carbon is required in other parts of the plant. Other regulatory processes ensure that the rate of starch synthesis is maintained and that not all the available fixed carbon is drained away from the chloroplast. On the other hand, decreasing the activity of the phosphate translocator by various treatments results in increased retention of fixed carbon in the chloroplast and increased starch accumulation in source leaves. Interestingly, plants in which the phosphate translocator is inhibited generally show no obvious changes in growth. They simply bypass the phosphate translocator and utilize alternate means to move carbon out of the chloroplast. Allocation is an important process in sink leaves as well. In sink leaves, imported carbon compounds either are stored in various forms or are utilized for metabolic processes, depending on the type of sink involved.

Partitioning. Source and sink regions interact within a single plant to control rates of export and to regulate the distribution of photosynthate between various sinks. Increasing the photosynthetic production of sugars in a source leaf, by increasing the light intensity for example, increases the rate of sugar export by the source. Decreasing the activity of enzymes that metabolize sucrose in a given sink region decreases the proportion of photosynthate arriving at that sink from the source areas. Removing or shading all the source leaves except one on a plant usually results in an immediate redistribution of the assimilate exported from the remaining active source, with no change occurring in export rate. Over the short term, the sink tissues appear to compete for available sugars, with the strongest sinks diverting more of the translocation stream to themselves. Increased sink strength may depend in part on the ability of the stronger sink tissues to lower the turgor pressure in the translocating cells to a greater extent. Over the long term (on the order of 2 days or more), removal or shading of all the source leaves except one leads to compensatory increases in both photosynthesis and export by the remaining active source.

Transport of Signaling Molecules

Long-term source-sink interactions are probably mediated by plant hormones, such as cytokinins, and by nutrients, such as potassium, phosphate, and sucrose. Molecules such as sucrose can change gene expression in source-sink systems. Small RNAs in the phloem play roles in defense against viral nucleic acids and in developmental processes. As they link sources and sinks in this role, they are a powerful means of communication and integration within the plant.

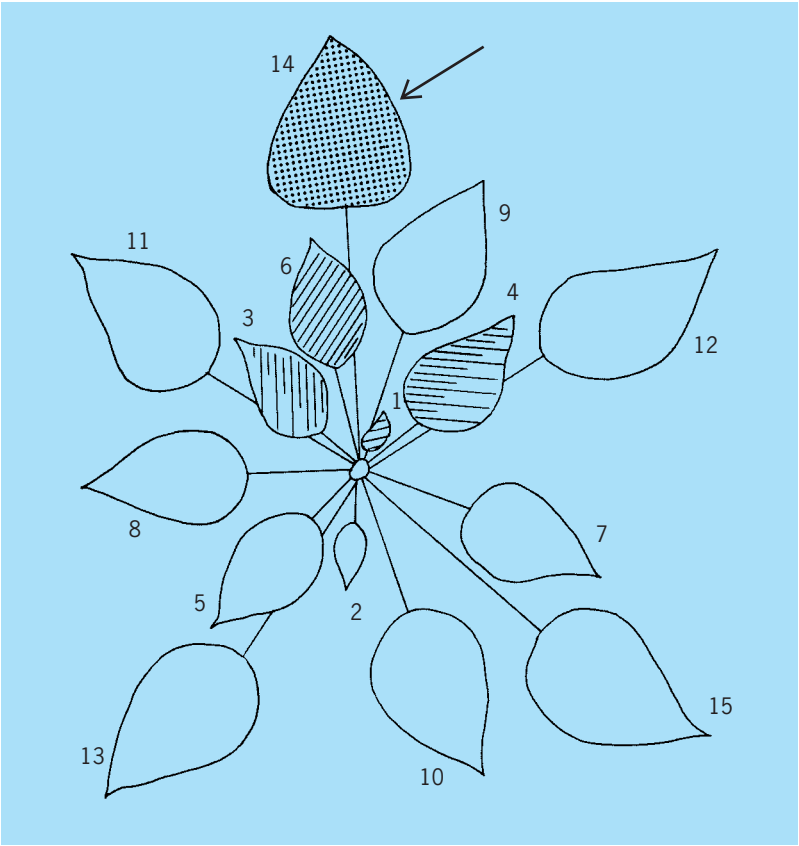

Fig. 3. Distribution of radioactivity in the leaves of a beet plant 1 week after administration of $^{14}CO_2$ for 4 h to a single leaf, indicated by arrow. Hatching indicates the approximate intensity of radioactivity from autoradiographs of leaves. Most of the ^{14}C is translocated to young leaves on the same side of the plant as the labeled source leaf. The leaves are numbered to show relative age, with 1 being the youngest.

Mechanism of phloem translocation. The mechanism of translocation must account for the following observations: polarity, bidirectional movement, velocity, energy requirement, turgor pressure, and phloem structure.

Polarity. Organic solutes move in the phloem from sources to metabolically active sinks, where growth or accumulation is occurring. The direction of translocation is determined more by physiological states of source and sink areas than by structural relations.

Velocity. Translocation is rapid and conveys large quantities of solutes. The velocity of translocation is about 0.4–0.8 in./min (~0.25 mm/s). The transport capacity for sugarbeet has been measured at 9 mg sucrose/h out of a 15.5-in.^{2} (100-cm^2) leaf. In the sausage tree (*Kigelia africana*), as much as 1.1 oz/day (31 g/day) of organic material must pass through the slender stem to account for the dry-weight gain of the cluster of four fruits.

Energy requirement. The high rates of directed transfer necessitate a considerable expenditure of energy both to move the stream of solutes and to maintain the structural integrity of the conducting tissue. Living phloem cells are necessary for translocation of

organic solutes in sieve elements. Killing the cells in a localized zone of the stem by heating is as effective a barrier to transport of phloem-limited solutes as is ringing. Treatments such as cooling lower the metabolism of a tissue and have been used to evaluate the role of energy in transport in the treated region. In cold-tolerant plants, translocation continues at an undiminished rate through a petiole held at 32°F (0°C) for many hours, indicating that energy expenditure in the conducting tissue is primarily required for maintenance of structure of the phloem rather than for moving the translocate stream. On the other hand, cooling source or sink regions slows translocation markedly. *See* COLD HARDINESS (PLANT).

Turgor pressure. Sieve-element contents are under positive turgor pressure. Pressures have been measured in the phloem of tree species up to 2.0 megapascals (20 atm) and in herbaceous species up to 1.0 MPa (10 atm).

Phloem structure. The sieve elements of flowering plants (angiosperms) lack a nucleus, have a specialized protoplast, and are interconnected by areas called sieve plates (**Fig. 4**). The pores in the sieve plates are essentially open channels that allow transport between cells. Because of their high degree of specialization, sieve elements are not a major source of energy for longitudinal transport. However, specialized phloem parenchyma cells (companion cells) associated with terminal regions of the phloem are large and have dense cytoplasm with numerous mitochondria. These cells appear to be well suited to controlling the concentration of solutes in the sieve elements and to supplying energy for loading and perhaps unloading as well.

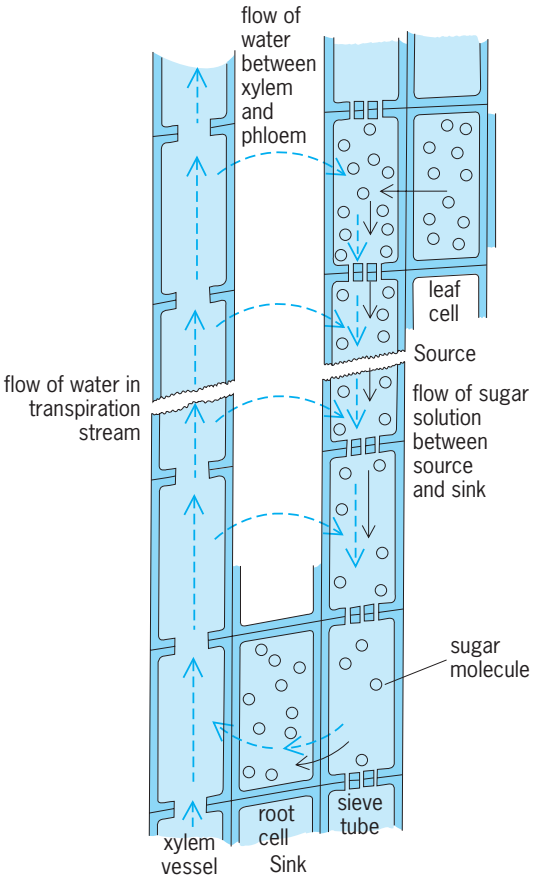

Fig. 5. Diagram of mass-flow theory; water flow is represented by colored arrows and sugar flow by black arrows. (*After P. H. Raven, R. F. Evert, and S. E. Eichhorn, Biology of Plants, 4th ed., Worth Publishers, 1986*)

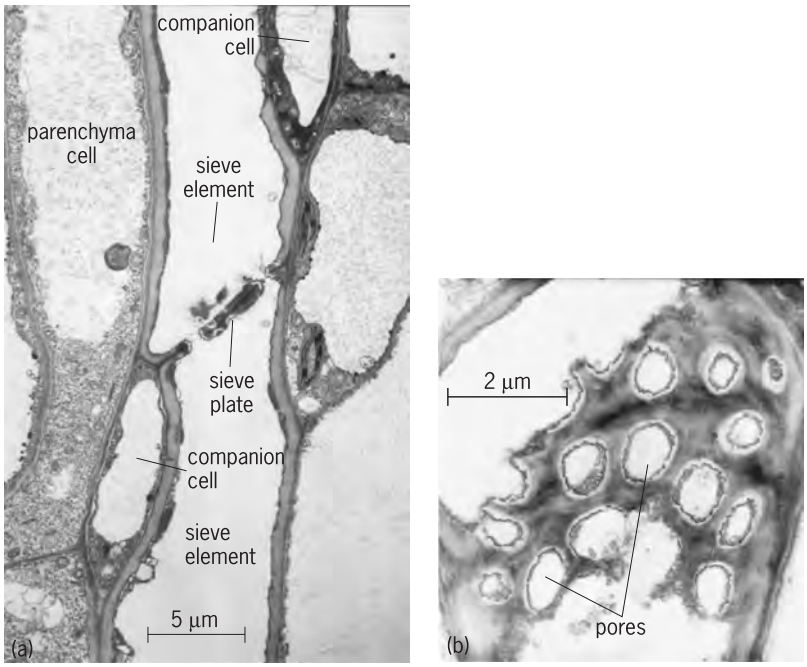

Fig. 4. Phloem in the stem of squash, as seen in electron micrographs. (*a*) Longitudinal view of parts of two mature sieve elements and a sieve plate (*from R. F. Evert, W. Eschrich, and S. E. Eichhorn, P-protein distribution in mature sieve elements of Cucurbita maxima, Planta, 103:193–210, 1968*). (*b*) Face view of a simple sieve plate (one sieve area per plate) between two mature sieve elements (*courtesy of R. F. Evert*).

Movement through the phloem sieve tubes consists of a mass flow of sugar solutes dissolved in water. Companion cells associated with the phloem endings in source regions accumulate solutes and deliver them to the sieve elements. The process, referred to as phloem-loading (described below), requires metabolic energy. At these sites, the sugar concentration in the sieve elements is high, and this causes the entry of water by osmosis from surrounding cells and from the xylem in particular; the resulting hydrostatic pressure causes water and dissolved solutes to flow along under the force of turgor pressure. At points of utilization or storage, solutes are removed from the sieve elements. The resultant lowering of solute concentration causes a lowering of the turgor pressure, also aiding translocation stream movement (**Fig. 5**).

Inorganic solutes. Inorganic solutes or mineral elements are found in both the phloem and xylem of higher plants. The initial distribution of minerals following uptake by the roots occurs within the vessels of the xylem and involves a general distribution to all transpiring surfaces. Transpiration is the loss of water from plants in the form of vapor, principally from the leaves, and is essentially an evaporation process. A more specific redistribution of some elements may then occur within the phloem.

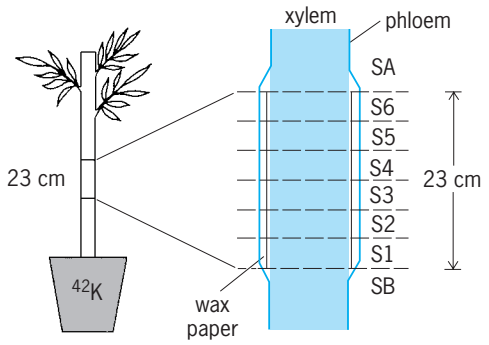

		stripped branch (xylem and phloem separated by wax paper)	
		ppm ^{42}K in phloem	ppm ^{42}K in xylem
above strip	SA	53	47
stripped section	S6	11.6	119
	S5	0.9	122
	S4	0.7	112
	S3	0.3	98
	S2	0.3	108
	S1	20	113
below strip	SB	84	58

Fig. 6. Demonstration, using radioactive potassium supplied to the roots, that xylem is the channel for the upward movement of water and minerals. Wax paper is inserted between the xylem and phloem to prevent lateral transport of the isotope. The relative amounts of radioactive potassium in each of the segments (S) is shown at the bottom. (*After P. H. Raven, R. F. Evert, and S. E. Eichhorn, Biology of Plants, 4th ed., Worth Publishers, 1986*)

Importance of phloem and xylem. Water and dissolved solutes move upward primarily in the tissues of the xylem (**Fig. 6**). The bark containing the phloem can be removed from a shoot with no immediate effects on the movement of water and dissolved mineral elements from the roots to the shoots.

Xylem sap can be collected for analysis in several ways. In herbaceous plants, xylem sap exudes from the root system of detopped plants, driven by a phenomenon called root pressure, and can be easily collected. Sap can also be extracted from shoots and twigs by suction or by pressure. Continuous measurements of xylem sap have been accomplished using a xylem-feeding insect, the meadow spittlebug. Sap collected by these methods contains relatively low concentrations of dissolved solutes (0.05–0.40% by weight) compared to phloem sap. Most of the solutes in xylem sap are inorganic cations and anions, with potassium, magnesium, calcium, sodium, and ammonium as the major cations, and with chloride, phosphate, sulfate, and nitrate as the major anions.

Analysis of phloem sap and studies of export of radioactively labeled ions from leaves indicate that some mineral ions move in the phloem as well. Potassium, magnesium, phosphate, and chloride are the major ions moving extensively in the phloem. In contrast, nitrate, calcium, sulfur, and iron move very little in phloem tissues. The low concentration of calcium in the phloem results primarily from its inability to enter the sieve elements. It is further maintained, at least in part, by the low solubility of calcium phosphate salts. Nitrogen does move in the phloem, but primarily in organic rather than inorganic forms.

One consequence of the differing conduction of elements in the phloem is that the symptoms of mineral deficiencies can differ according to the phloem mobility of the elements involved. Sink tissues, including young leaves, can withdraw nutrients from older leaves via the phloem. Thus, when a plant is deficient in an element which can be transported in the phloem, the symptoms appear earlier and are most pronounced in the older leaves. On the other hand, the deficiency symptoms of an ion which is not transported in the phloem initially affect young, actively growing tissues.

Entry into xylem. The absorption of inorganic nutrients by plants involves several sequential steps: movement of ions through the soil to root surfaces (by diffusion and mass flow); lateral movement across the roots through cell walls and root cells; release to the xylem elements; and movement to the aerial portions of the plant in the xylem stream (**Fig. 7**). The initial movement of mineral ions into the root can occur by diffusion through cell walls external to cell membranes. However, in young roots at least, there exists a barrier to free diffusion of ions across roots. This barrier, called the Casparian strip, consists of an impregnation of the cell wall of the endodermis of the root with suberin, which renders the cell wall impermeable. When they reach this strip, the inorganic solutes cannot continue to cross the root through cell walls but must enter the cytoplasm of the living cells. The solutes then pass from living cell to living cell until they reach the vascular tissues, where they are released from the cells into the xylem elements (which are dead at maturity). It is also possible for inorganic solutes to enter living cells closer to the root surface, at the epidermis or cortex; however, solutes must enter living cells at the endodermis if they have not already done so.

Although movement of ions within the cell walls and xylem elements is essentially a nonselective, passive process, passage across cell membranes in the root is selective and may require metabolic energy. Any movement of ions across membranes occurs via membrane proteins (carriers or channels), and these are more or less selective. The active transport step in the root, which requires a carrier, most likely occurs at the exit of solutes from the living cells rather than entry into them. This process is called xylem loading in the root. Because of these membrane transport processes, young roots show both selectivity in the uptake of inorganic solutes and a capacity to accumulate ions to levels higher than those in the soil.

Although it is generally recognized that most salt absorption occurs in the youngest region of the root (immediately behind the root tip), significant amounts of water and salt do enter through older

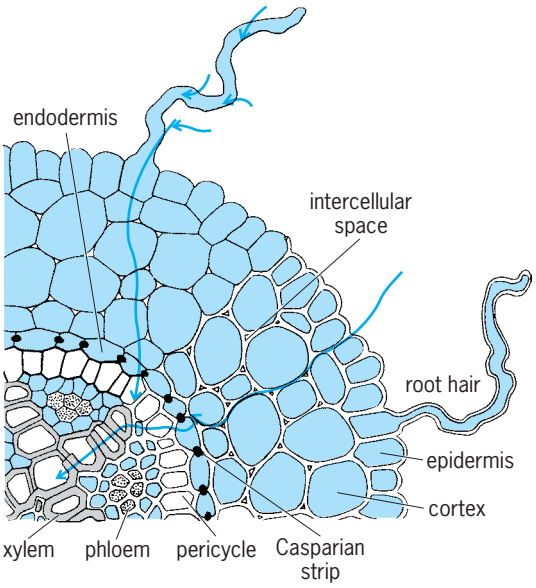

Fig. 7. Cross section of a root, showing the pathway of uptake of water and minerals. Note that water and ions may be taken up into the cytoplasm of the root cells even before they reach the barrier of the Casparian strip (upper part of figure). (*After P. H. Raven, R. F. Evert, and S. E. Eichhorn, Biology of Plants, 4th ed., Worth Publishers, 1986*)

roots in woody species. For the most part, these roots are rendered impermeable by the processes of secondary growth; however, it appears that water and solutes can enter through lenticels, fissures between plates of bark, and openings caused by the death of small branch roots. This nonselective movement occurs, as in the soil, both by diffusion and by mass flow. As a result of this nonselective component of ion uptake, practically all dissolved ions of the root environment are found in varying quantities in the aerial portions of plants.

Mechanism of xylem transport. Once in the xylem elements, inorganic solutes move with water to the transpiring surfaces. Water is lost in the leaves by evaporation from cell-wall surfaces; water vapor then diffuses into the atmosphere by way of small pores between two specialized cells (guard cells). The guard cells and the pore are collectively called a stomate. This loss of water from the leaf causes movement of water out of the xylem in the leaf to the surfaces where evaporation is occurring. Water has a high internal cohesive force, especially in small tubes with wettable walls. In addition, the xylem elements and the cell walls provide a continuous water-filled system in the plant. Thus the loss of water from the xylem elements in the leaves causes a tension or negative pressure in the xylem sap. This tension is transmitted all the way down the stem to the roots, so that a flow of water occurs up the plant from the roots and eventually from the soil. The velocity of this sap flow in tree species ranges from 3 to approximately 165 ft/h (1 to 50 m/h), depending on the diameter of the xylem vessels.

The movement of inorganic solutes within the transpiration stream might suggest that all mineral elements move upward at the same rate. However, some ions appear to move more slowly than others due to adsorption on the walls of the xylem elements and accumulation by living cells bordering the xylem elements. Some have suggested that sodium, for example, might be prevented from reaching toxic levels in leaves by uptake in upper parts of roots and lower parts of shoots. In addition, certain ions are accumulated more in the roots of some species and are thus prevented from reaching the xylem stream.

The rate of transpiration may affect the rate of ion uptake. Rapid water movement carries ions in the soil solution to the root surfaces and reduces the salt concentration in the root xylem, facilitating release of ions into the xylem sap.

In the absence of transpiration, the active uptake of ions by root cells may cause an osmotic influx of water into the roots, building up a positive pressure in the xylem elements and providing an alternate driving force for water and ion movement. This phenomenon, called root pressure, only occurs under special conditions and in some species. It is not sufficient to explain sap movement to the tops of tall plants. *See* PLANT MINERAL NUTRITION; PLANT-WATER RELATIONS.

S. Sovonick-Dunford

Bibliography. P. H. Raven, R. F. Evert, and S. E. Eichhorn, *Biology of Plants*, 2004; K. R. Stern, *Introductory Plant Biology*, 10th ed., 2006; L. Taiz and E. Zeiger, *Plant Physiology*, 2006.

Plant viruses and viroids

Plant viruses are pathogens which are composed mainly of a nucleic acid (genome) normally surrounded by a protein shell (coat); they replicate only in compatible cells, usually with the induction of symptoms in the affected plant. Viroids are among the smallest infections agents known. Their circular, single-stranded ribonucleic acid (RNA) molecule is less than one-tenth the size of the smallest viruses.

Viruses. Viruses can be seen only with an electron microscope (**Figs. 1** and **2**). Isometric (spherical) viruses range from 25 to 50 nanometers in diameter, whereas most anisometric (tubular) viruses are 12 to 25 nm in diameter and of various lengths (200–2000 nm), depending on the virus. The coat of a few viruses [for example, the large and complex rhabdoviruses such as lettuce necrotic yellows virus (LNYV) or the tomato spotted wilt virus] is covered by a membrane which is derived from its host.

Over 800 plant viruses have been recognized and characterized. The genomes of most of them, such as the tobacco mosaic virus (TMV) or the potato virus Y, are infective single-stranded RNAs; some RNA viruses (for instance, the wound tumor virus) have double-stranded RNA genomes. Cauliflower mosaic virus and bean golden mosaic virus are examples of viruses having double-stranded and single-stranded deoxyribonucleic acid (DNA). The genome of many plant viruses is a single polynucleotide and is contained in a single particle, whereas the genomes of brome

mosaic and some other viruses are segmented and distributed between several particles. There are also several low-molecular-weight RNAs (satellite RNAs) which depend on helper viruses for their replication. *See* DEOXYRIBONUCLEIC ACID (DNA); RIBONUCLEIC ACID (RNA).

Purified viruses suitable for characterization can be obtained from infected tissues by ultracentrifugation and other procedures. The diversity of particle size, chemical composition of nucleic acids and proteins, and other characteristics has resulted in the classification of plant viruses into 42 genera and 11 families, with more groups being formed as knowledge of the properties of viruses increases.

Replication. The natural hosts of plant viruses are widely distributed throughout the higher-plant kingdom, including some algae and fungi. Some viruses (TMV and cucumber mosaic virus) are capable of infecting over a hundred species in many families, whereas others, such as wheat streak mosaic virus, are restricted to a few species in the grass family. The replication of single-standard RNA viruses involves release of the virus genome from the coat protein; the association of the RNA with the ribosomes of the cell; translation of the genetic information of the RNA into specific proteins, including subunits of the coat protein and possibly viral RNA-synthesizing enzymes (replicases); transmission by vectors and diseases induction; synthesis of noninfective RNA using parental RNA as the template; and assembly of the protein subunits and viral RNA to form complete virus particles.

In other RNA viruses such as lettuce necrotic yellows virus, an enzyme which is contained in the virus must first make a complementary (infective) copy of the RNA; this is then translated into enzymes and coat protein subunits. The replication of double-stranded RNA viruses such as wound tumor virus is basically similar to that of lettuce necrotic yellows virus.

With double-stranded DNA viruses such as cauliflower mosaic virus, viral DNA is uncoated in a newly infected cell and transported to the nucleus, where it associates with histones to form a closed circular minichromosome. Two major RNA species (35S and 19S) are transcribed from the minichromosome by a host-encoded enzyme and are translated in the cytoplasm to produce virus-associated proteins. The 35S RNA serves as the template for a viral enzyme which transcribes it to viral DNA, which is then encapsidated to form virus particles. For example, the single-stranded DNA geminiviruses are usually found in nuclei of infected plants. Their circular genomes are encapsidated by coat protein arranged to form a couplet of particles. Synthesis of progeny DNA occurs in nuclei via a double-stranded DNA replicative structure associated with host-encoded DNA synthesizing enzymes.

Symptoms. Symptoms are the result of an alteration in cellular metabolism and are most obvious in newly developing tissues. Infected leaves often show a mosaic of light green-yellow (chlorotic) areas and areas of normal green color (**Fig. 3***a*); virus-induced chlorosis is associated with the development of abnormal chloroplasts. Infection of plants with two or more viruses is common, especially in perennial plants or those propagated by cuttings; specific diseases (for example, severe streak of tomato and rugose mosaic of potato) are caused by dual infection with tomato mosaic virus and potato virus X, and potato virus X and potato virus Y, respectively. In some plants, depending on the virus, the initial infection does not spread because cells surrounding the infected cells die, resulting in the formation of

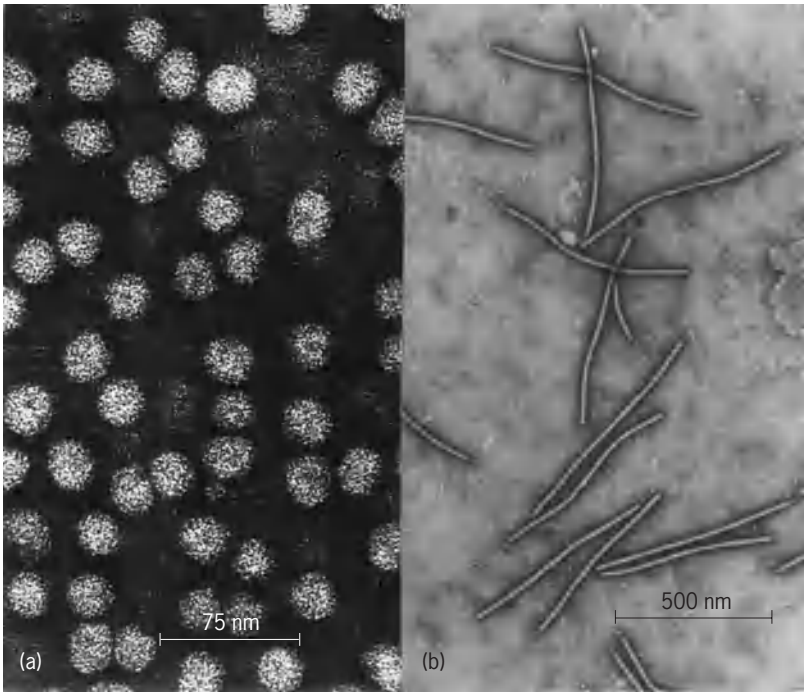

Fig. 1. Representative plant viruses in purified virus preparation obtained from infected leaves. (*a*) Tobacco streak virus (isometric). (*b*) Pea seed-borne mosaic virus (anisometric).

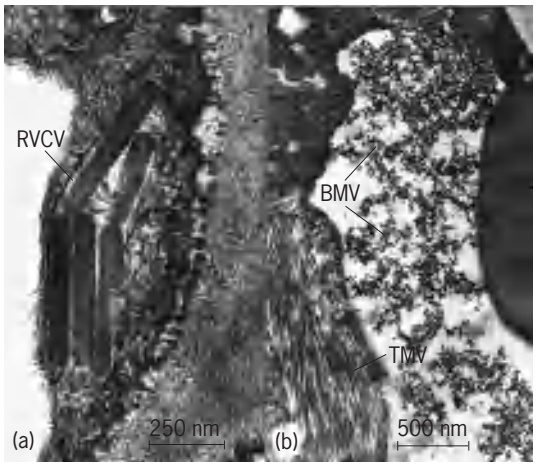

Fig. 2. Infected leaf cells. (*a*) Cell of raspberry showing particles of raspberry vein chlorosis virus (RVCV) (*courtesy of R. Stace-Smith*). (*b*) Leaf cell of barley infected with brome mosaic virus (BMV) and tobacco mosaic virus (TMV) (*from R. I. Hamilton and C. Nichols, The influence of bromegrass mosaic virus on the replication of tobacco mosaic virus in Hordeum vulgare, Phytopathology, 7:484, 1977*).

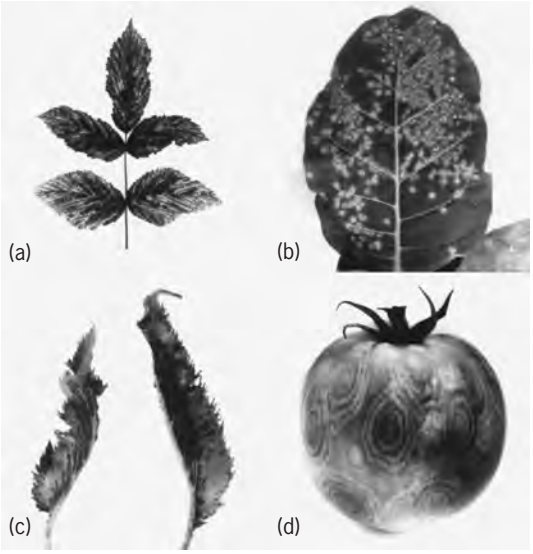

Fig. 3. Symptoms induced by virus infection. (*a*) Mosaic in leaves of raspberry infected with raspberry vein-banding mosaic disease induced by rubus yellow net virus in association with black raspberry necrosis virus. (*b*) Necrotic lesions in tobacco leaves induced by rub inoculation of cherry leaf roll virus. (*c*) Deformation of cherry leaves caused by infection with cherry raspberry leaf virus. (*d*) Necrotic rings on tomato fruit induced by infection with tomato mosaic virus. (*Courtesy of R. Stace-Smith*)

necrotic lesions (Fig. 3*b*). Such plants are termed hypersensitive. The size and shape of leaves (Fig. 3*c*) and fruit (Fig. 3*d*) may be adversely affected, and in some instances plants may be killed. Not all virus infections produce distinctive symptoms; potato infected with some North American strains of potato viruses S and X are symptomless.

Transmission. Some viruses, especially those infecting trees, can only be transmitted experimentally by grafting. The continued propagation of infected plants by cuttings also transmits viruses. Contact transmission occurs through wounds caused by leaves of infected plants rubbing those of adjacent healthy plants or by the handling of diseased and healthy plants.

The most common mode of transmission for many viruses is by means of vectors, mainly insects (predominantly aphids and leafhoppers), and to a lesser extent mites, soil-inhabiting fungi, and nematodes which acquire viruses by feeding on infected plants. Viruses transmitted by one class of vector are rarely transmitted by another, and there is often considerable specificity between strains of a virus and their vectors. Wound tumor virus infects its leafhopper vectors so that they transmit the virus for life, while the aphid vector of potato virus Y, which carries the virus on its mouthparts, loses the ability to transmit the virus within an hour or two unless it again feeds on a diseased plant.

Some viruses (such as barley stripe mosaic virus and soybean mosaic virus) are transmitted to succeeding generations mainly by embryos in seeds produced by infected plants; over 200 viruses are transmitted in this way. Some are also vectored by insects or nematodes. Infected pollen can transmit virus to healthy flowers, resulting in infected seed and, in some cases (ringspot in cherry and bushy dwarf disease in raspberry), infection of the seed-bearing plant.

Control of virus diseases. Virus diseases are difficult to control because of the complex interaction of virus strains, vectors, overwintering sources such as weeds, and the crop to be protected. Breeding for resistance utilizing genes for hypersensitivity to a particular virus may be useful because the virus is usually limited to the few cells bordering the point of infection. Some success is being obtained in breeding plants that are not suitable hosts for aphid development, thus reducing the probability that they will be infected by aphid-borne viruses. Plants genetically engineered to express specific viral proteins such as the coat protein are often highly resistant to infection by related viruses. *See* GENETIC ENGINEERING.

In vegetatively propagated plants (such as the potato or fruit trees), important virus diseases are being controlled by tissue culture of plants from virus-free stem tips selected from infected plants. Seed-transmitted viruses are controlled by the use of virus-free seed. Crop rotation and chemical treatment of fields aid in the control of soil-borne viruses, primarily by reducing the population of their soil-borne fungal and nematode vectors. Inoculation of young tomato and citrus plants with mild strains of tobacco mosaic virus and citrus tristeza virus, respectively, often protects such plants from damaging effects of severe strains.

Viroids. Only about 30 viroids are known, but they cause very serious diseases in such diverse plants as chrysanthemum, citrus, coconut, and potato. They can also be isolated from plants that do not exhibit symptoms. Viroids are mainly transmitted by vegetative propagation, but some, such as potato spindle tuber viroid, are transmitted by seed or by contact between infected and healthy plants. Tomato planta macho viroid is efficiently transmitted by aphids. *See* PLANT PATHOLOGY; VIROIDS; VIRUS.

Richard I. Hamilton

Bibliography. T. Hohn and J. Schell, *Plant DNA Infections Agents*, 1987; K. Maramorosch (ed.), *Viroids and Satellites: Molecular Parasites at the Frontier of Life*, 1991; R. E. F. Matthews, *Fundamentals of Plant Virology*, 1997; P. Pirone and J. G. Shaw (eds.), *Viral Genes and Plant Pathogenesis*, 1990; B. A. Voyles, *The Biology of Viruses*, 1993; T. Wilson, A. Michael, and J. W. Davies (eds.), *Genetic Engineering with Plant Viruses*, 1992.

Plant-water relations

Water is the most abundant constituent of all physiologically active plant cells. Leaves, for example, have water contents which lie mostly within a range of 55–85% of their fresh weight. Other relatively succulent parts of plants contain approximately the same proportion of water, and even such largely nonliving tissues as wood may be 30–60% water on a fresh-weight basis. The smallest water contents in living

parts of plants occur mostly in dormant structures, such as mature seeds and spores. The great bulk of the water in any plant constitutes a unit system. This water is not in a static condition. Rather it is part of a hydrodynamic system, which in terrestrial plants involves absorption of water from the soil, its translocation throughout the plant, and its loss to the environment, principally in the process known as transpiration.

Cellular water relations. The typical mature, vacuolate plant cell constitutes a tiny osmotic system, and this idea is central to any concept of cellular water dynamics. Although the cell walls of most living plant cells are quite freely permeable to water and solutes, the cytoplasmic layer that lines the cell wall is more permeable to some substances than to others. This property of differential permeability appears to reside principally in the layer of cytoplasm adjacent to the cell wall (plasma membrane, or plasmalemma) and in the layer in contact with the vacuole (vacuolar membrane, or tonoplast). This cytoplasmic system of membranes is usually relatively permeable to water, to dissolved gases, and to certain dissolved organic components. It is often much less permeable to sugars and mineral salts. The permeability of the cytoplasmic membranes is quite variable, however, and under certain metabolic conditions solutes that ordinarily penetrate through these membranes slowly or not at all may pass into or out of cells rapidly. *See* CELL (BIOLOGY); PLANT CELL.

Osmotic and turgor pressures. If a plant cell in a flaccid condition—one in which the cell sap exerts no pressure against the encompassing cytoplasm and cell wall—is immersed in pure water, inward osmosis of water into the cell sap ensues. Osmosis may be defined as the movement of solvent molecules, which in living organisms are always water, across a membrane that is more permeable to the solvent than to the solutes dissolved in it. The driving force in osmosis is the difference in free energies of the water on the two sides of the membrane. Pure water, as a result of its intrinsic properties, possesses free energy. Since there is no certain way of measuring the free energy of water, it is arbitrarily given a value of zero when exposed only to atmospheric pressure. The free energy of water is influenced by temperature, but this discussion will be restricted to isothermal systems. This free energy has been designated water potential and may be expressed in either energy or pressure units. It is more meaningful to use pressure units in considerations of the water potential of plants.

Inward osmosis of water takes place under the conditions specified above because the water potential of the cell sap is less than that of the surrounding pure water by the amount of its osmotic pressure. Introduction of solutes always lowers the potential of water by the amount of the resulting osmotic pressure. If the osmotic pressure of the cell sap is 15 atm (1.5 megapascals), then the water potential of the cell sap is 15 atm (1.5 MPa) less than that of pure water at the same temperature and under the same pressure.

The gain of water by the cell as a result of inward osmosis results in the exertion of a turgor pressure against the protoplasm, which in turn is transmitted to the cell wall. This pressure also prevails throughout the mass of solution within the cell. If the cell wall is elastic, some expansion in the volume of the cell occurs as a result of this pressure, although in many kinds of cells this is relatively small.

Because of the solutes invariably present, the cell sap possesses an osmotic pressure. The osmotic pressures of most plant cell saps lie within the range of 5–40 atm (0.5–4 MPa), although values as high as 200 atm (20 MPa) have been found in some halophytes (plants that can tolerate high-solute media). The osmotic pressures of the cells of a given plant tissue vary considerably with environmental conditions and intrinsic metabolic activities. More or less regular daily or seasonal variations occur in the magnitude of cell sap osmotic pressures in the cells of many tissues. It is the osmotic pressure of the cell sap, coupled with the differential permeability of the cytoplasmic membranes and the relative inelasticity of the cell walls, which permits the development of the more or less turgid condition characteristic of most plant cells.

With continued osmosis of water into the cell, its turgor pressure gradually increases until it is equal to the final osmotic pressure of the sap. Subjection of the water in the cell sap to pressure increases its water potential by the amount of the imposed pressure. In the example given above, disregarding the usually small amount of sap dilution as a result of cell expansion, the water potential of the cell sap is reduced 15 atm (1.5 MPa) because of the presence of solutes (the osmotic pressure is the index of this lowering of water potential) and raised 15 atm (1.5 MPa) as a result of turgor pressure when maximum turgor is reached. Hence, when a dynamic equilibrium is attained, the water potential of the cell sap is zero and thus is equal to that of the surrounding water, a necessary condition if equilibrium is to be achieved.

If the same cell in a flaccid condition is immersed in a solution with an osmotic pressure of 6 atm (0.6 MPa), inward osmosis of water occurs, but does not continue as long as when the cell is immersed in pure water. Disregarding sap dilution, a dynamic equilibrium will be attained under these circumstances when the turgor pressure of the cell sap has reached 9 atm (0.9 MPa), because at this point the water potential in the cell sap and the water potential in the surrounding solution will be equal. Since the water potential of the cell sap was originally diminished 15 atm (1.5 MPa) because of the presence of solutes and then raised 9 atm (0.9 MPa) because of turgor pressure, the net value of the water potential in the cell at equilibrium is −6 atm (−0.6 MPa). This is also the value of the water potential in the surrounding solution, as indexed by its osmotic pressure. At dynamic equilibrium the number of water molecules entering the cell and the number leaving per unit of time will be equal. *See* OSMOREGULATORY MECHANISMS.

Water potential. As the examples just given indicate, the effective physical quantity controlling the direction of osmotic movement of water from cell to cell in plants or between a cell and an external solution is the water potential (WP) of the water. In most plant cells this quantity is equal to the osmotic pressure (OP) of the water less the turgor pressure (TP) to which it is subjected. Since the osmotic pressure is the index of the amount by which the water potential of a solution is less than that of pure water under the same conditions, it must be treated as a negative quantity. Turgor pressures usually have a positive value, but if water passes into a state of tension, they are negative. Examples of the occurrence of water under tension in plants will be discussed later in this article. In an unconfined solution the water potential is equal to the osmotic pressure, since there is no turgor pressure. In a fully turgid plant cell the water potential is zero, since the turgor pressure is equal to the osmotic pressure. In a fully flaccid plant cell the turgor pressure is zero and the water potential is equal to the osmotic pressure. The interrelationships among these quantities may be expressed: $WP = (-OP) + TP$.

Since the osmotic pressure must be treated as a negative quantity and since turgor pressures rarely exceed osmotic pressures, the water potentials in plant cells are negative in value. The only well-known exception to this is root pressure, discussed later, during which small positive water potentials can be generated in the water-conductive tissues.

The relationships expressed in the equation are illustrated graphically in **Fig. 1**, in which allowance has been made for the effect of the volume changes which are characteristic of some kinds of cells with shifts in turgor pressure. The conditions which would prevail in the cell if the cell sap passed into a state of tension (negative pressure) are indicated by the dotted extension of curves to the left.

Cell-to-cell movement of water in plants always occurs from the cell of greater (less negative) water potential to the cell of lesser (more negative) water potential. Such movement of water in plant tissues apparently often occurs along water potential gradients in which the water potential of each cell in a series is less (more negative) than that of the preceding one.

Plasmolysis. If a turgid or partially turgid plant cell is immersed in a solution with a greater osmotic pressure than the cell sap, a gradual shrinkage in the volume of the cell ensues; the amount of shrinkage depends upon the kind of cell and its initial degree of turgidity. When the lower limit of cell wall elasticity is reached and there is continued loss of water from the cell sap, the protoplasmic layer begins to recede from the inner surface of the cell wall. Retreat of the protoplasm from the cell wall often continues until it has shrunk toward the center of the cell, the space between the protoplasm and the cell wall becoming occupied by the bathing solution. This phenomenon is called plasmolysis. If a cell is immersed in a solution with an osmotic pressure which just slightly exceeds that of the cell sap, withdrawal of the protoplasm from the cell wall would be just barely initiated. The stage of plasmolysis shown in Fig. 1 is called incipient plasmolysis, and it is the basis for one of the methods which is commonly used in measuring the osmotic pressure of plant cells.

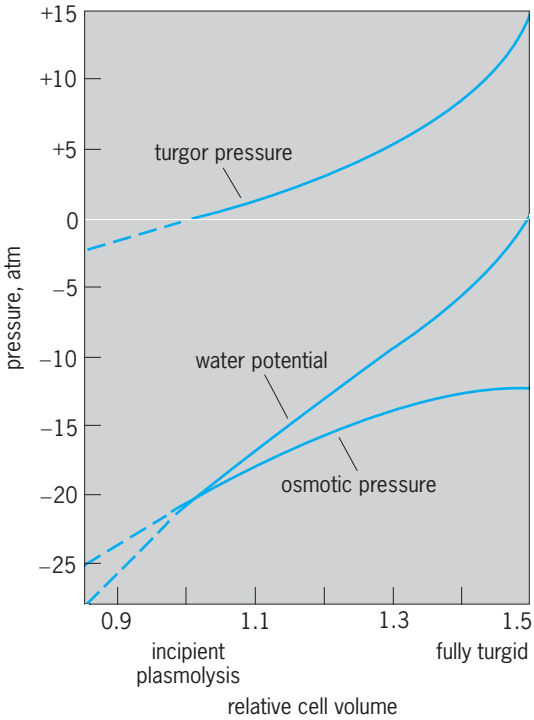

Fig. 1. Curves showing interrelationships among osmotic pressures, turgor pressures, water potentials, and volumes of a plant cell. 1 atm = 10^2 kPa.

Imbibition. In some kinds of plant cells movement of water occurs principally by the process of imbibition rather than osmosis. The swelling of dry seeds when immersed in water is a familiar example of this process. Imbibition occurs because of the more negative water potential in the imbibant as compared with the water potential in some contiguous part of the system. An equilibrium is reached only when the water potential in the two parts of the system has attained the same value. The water potential in a dry seed is extremely low, being equal in value, but with a negative sign, to its imbibition pressure. The water potential of pure water is zero; hence movement of water occurs into the seed. Even if the seeds are immersed in a solution of considerable osmotic pressure, which in an unconfined solution is an index of its negative water potential, imbibition occurs. However, if the osmotic pressure of the solution is high enough (of the order of 1000 atm or 10 MPa), the seed will not gain water from the solution and may even lose a little to the solution. In other words, if the water potential of the solution is negative enough, imbibition does not occur.

However, a more negative water potential in the imbibant, as compared with the surrounding or adjacent medium is not the only condition which must be fulfilled if imbibition is to occur. Many kinds of seeds swell readily if immersed in water but not when immersed in ether or other organic solvents.

On the other hand, rubber does not imbibe water but does imbibe measurable quantities of ether and other organic liquids. Certain specific attraction forces between the molecules of the imbibant and the imbibed liquid are therefore also a requisite for the occurrence of imbibition.

In an imbibitional system the imbibition pressure IP of the imbibant is the analog of the osmotic pressure in an osmotic system; hence in such a system water potential is related to the imbibition pressure and turgor pressure: $WP = (-IP) + TP$. The imbibition pressure may be regarded as an index of the reduction in water potential in an imbibant insofar as it results from attractions between the molecules of the imbibant and water molecules. For an unconfined imbibant immersed in water, the negative water potential initially equals the imbibition pressure, since there is no turgor pressure. The more nearly saturated such an imbibant becomes, the smaller its imbibition pressure and the less negative its water potentials. A fully saturated imbibant has a zero imbibition pressure and a zero water potential.

Stomatal mechanism. Various gases diffuse into and out of physiologically active plants. Those gases of greatest physiological significance are carbon dioxide, which diffuses into the plant during photosynthesis and is lost from the plant in respiration; oxygen, which diffuses in during respiration and is lost during photosynthesis; and water vapor, which is lost in the process of transpiration. The great bulk of the gaseous exchanges between a plant and its environment occurs through tiny pores in the epidermis called stomates. Although stomates occur on many aerial parts of plants, they are most characteristic of, and occur in greatest abundance in, leaves. *See* EPIDERMIS (PLANT); LEAF.

Each stomate or stoma (plural, stomates or stomata) consists of a minute elliptical pore surrounded by two distinctively shaped epidermal cells called guard cells. Stomates are sometimes open and sometimes closed; when closed, all gaseous exchanges between the plant and its environment are greatly retarded. The size of a fully open stomate differs greatly from one species of plant to another. Among the largest known are those of the wandering Jew (*Zebrina pendula*), whose axial dimensions average 31 by 12 micrometers. In most species the stomates are much smaller, but all of them afford portals of egress or ingress which are enormous relative to the size of the gas molecules that diffuse through them. The number of stomates per square centimeter of leaf surface ranges from a few thousand in some species to over a hundred thousand in others. In many species of plants stomates are present in both the upper and lower epidermises, usually being more abundant in the lower. In many species, especially of woody plants, they are present only in the lower epidermis. In floating-leaved aquatic species stomates occur only in the upper epidermis.

Rates of transpiration (loss of water vapor) from leaves of the expanded type often are 50% or even more of the rate of evaporation from a free water surface of equal area under the same environmental conditions. Loss of water vapor from leaves may occur at such relatively high rates despite the fact that the aggregate area of the fully open stomates is only 1–3% of the leaf area. Much more significant is the fact that the rate of diffusion of carbon dioxide, essential in photosynthesis, into the leaves through the stomates is much greater than through the equivalent area of an efficient carbon dioxide–absorbing surface.

Although some mass flow of gases undoubtedly occurs through the stomates under certain conditions, most movement of gases into or out of a leaf takes place by diffusion through the stomates. Diffusion is the physical process whereby molecules of a gas move from a region of their greater diffusion pressure to the region of their lesser diffusion pressure as a result of their own kinetic activity. Molecules of liquids and solids (to a limited extent), molecules and ions of solutes, and colloidal particles also diffuse whenever the appropriate circumstances prevail.

Diffusion of gases through small pores follows certain principles which account for the high diffusion capacity of the stomates. In the diffusion of a gas through a small pore, an overwhelming proportion of the molecules escape over the rim of the pore relative to those escaping through its center. Hence, diffusion rates through small apertures vary as their perimeter rather than as their area. The less the area of a pore, therefore, the greater is its diffusive capacity relative to its area. Therefore, a gas may diffuse nearly as rapidly through a septum pierced with a number of small orifices, whose aggregate area represents only a small proportion of the septum area, as through an open surface equal in area to the septum. The high diffusive capacity of the stomates can be accounted for in terms of these principles. Since diffusion of gases through stomates is proportional to the perimeter of the pore rather than to its area, the diffusion rate through a partially open stomate is almost as great as that through a fully open stomate.

In general, stomates are open in the daytime and closed at night, although there are many exceptions to this statement. The mechanism whereby stomates open in the light and close in the dark seems to be principally an osmotic one, although other factors are probably involved. Upon the advent of illumination, the hydrogen ion concentration of the guard cells decreases. This favors the action of the enzyme phosphorylase, which in the presence of phosphates causes transformation of insoluble starch into the soluble compound glucose-1-phosphate. The resulting increase in the solute concentration of the guard cells causes an increase in their osmotic pressure, and hence also in the negativity of their water potential. Osmotic movement of water takes place from contiguous epidermal cells, in which there is little daily variation in osmotic pressure, into the guard cells. The resulting increase in turgor pressure of the guard cells causes them to open. With the advent of darkness or of a relatively low light intensity, the reverse train of processes is

apparently set in operation, leading ultimately to stomatal closure.

Light of low intensity is, generally speaking, less effective than stronger illumination in inducing stomatal opening. Hence stomates often do not open as wide on cloudy as on clear days, and often do not remain open for as much of the daylight period. A deficiency of water within the plant also induces partial to complete closure of the stomates. During periods of drought, therefore, stomates remain shut continuously or, at most, are open for only short periods each day, regardless of the light intensity to which the plant is exposed. Opening of the stomates does not occur in most species at temperatures approaching freezing. Hence in cold or even cool weather stomates often remain closed even when other environmental conditions are favorable to their opening. Nocturnal opening occurs at times in some species, but the conditions which induce this pattern of stomatal reaction are not clearly understood.

Transpiration Process

The term transpiration is used to designate the process whereby water vapor is lost from plants. Although basically an evaporation process, transpiration is complicated by other physical and physiological conditions prevailing in the plant. Whereas loss of water vapor can occur from any part of the plant which is exposed to the atmosphere, the great bulk of all transpiration occurs from the leaves. There are two kinds of foliar transpiration: (1) stomatal transpiration, in which water vapor loss occurs through the stomates, and (2) cuticular transpiration, which occurs directly from the outside surface of epidermal walls through the cuticle. In most species 90% or more of all foliar transpiration is of the stomatal type.

Stomatal transpiration. The dynamics of stomatal transpiration is considerably more complex than that of cuticular transpiration. In the leaves of most kinds of plants the mesophyll cells do not fit together tightly, and the intercellular spaces between them are occupied by air. A veritable labyrinth of air-filled spaces is thus present within a leaf, bounded by the water-saturated walls of the mesophyll cells. Water evaporates readily from these wet cell walls into the intercellular spaces. If the stomates are closed, the only effect of such evaporation is to saturate the intercellular spaces with water vapor. If the stomates are open, however, diffusion of water vapor usually occurs through them into the surrounding atmosphere. Such diffusion always occurs unless the atmosphere has a vapor pressure equal to or greater than that within the intercellular spaces, a situation which seldom prevails during the daylight hours of clear days. The two physical processes of evaporation and diffusion of water vapor are both integral steps in stomatal transpiration. Physiological control of this component of transpiration is exerted through the opening and closing of the stomates, previously described.

Environmental factors. Light is one of the major factors influencing the rate of transpiration because of its controlling effect on the opening and closing of stomates. Since stomatal transpiration is largely restricted to the daylight hours, daytime rates of transpiration are usually many times greater than night rates, which largely or entirely represent cuticular transpiration. Since leaves in direct sunlight usually have temperatures from one to several degrees higher than that of the surrounding atmosphere, light also has a secondary accelerating effect on transpiration through its influence on leaf temperatures. Increase in leaf temperatures results in an increase in the pressure of the water vapor molecules within the leaf.

The rate of diffusion of water vapor through open stomates depends upon the steepness of the vapor pressure gradient between the intercellular spaces and the outside atmosphere. When the vapor pressure in that part of the intercellular spaces just below the stomatal pores is high relative to that of the atmosphere, diffusion of water vapor out of the leaf occurs rapidly; when it is low, water vapor diffusion occurs much more slowly.

Temperature has a marked effect upon rates of transpiration, principally because of its differential effect upon the vapor pressure of the intercellular spaces and atmosphere. Although leaf temperatures do not exactly parallel atmospheric temperatures, increase in atmospheric temperature in general results in a rise in leaf temperature and vice versa. On a warm, clear day such as would be typified by many summer days in temperate latitudes and with an adequate soil water supply, increase in temperature results in an increase in the vapor pressure of the intercellular spaces. Such a rise in vapor pressure occurs because the vapor pressure corresponding to a saturated condition of an atmosphere increases with rise in temperature, and the extensive evaporating surfaces of the cell walls bounding the intercellular spaces make it possible for the intercellular spaces to be maintained in an approximately saturated condition most of the time. An increase in temperature ordinarily has little or no effect on the vapor pressure of the atmosphere, and this is especially true of warm, bright days, when transpiration rates are the highest. Hence, as the temperature rises, the vapor pressure of the intercellular spaces increases relative to that of the external atmosphere, the vapor pressure gradient through the stomates is steepened, and the rate of outward diffusion of water vapor increases.

Wind velocity is another factor which influences the rate of transpiration. Generally speaking, a gentle breeze is relatively much more effective in increasing transpiration rates than are winds of greater velocity. In quiet air, localized zones of relatively high atmospheric vapor pressure may build up in the vicinity of transpiring leaves. Such zones retard transpiration unless there is sufficient air movement to prevent the accumulation of water vapor molecules. The bending, twisting, and fluttering of leaf blades and the swaying of stalks and branches in a wind also contribute to increasing the rate of transpiration.

Soil water conditions exert a major influence on

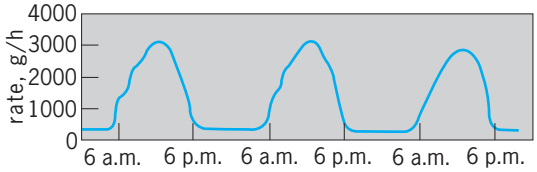

Fig. 2. Daily periodicity of transpiration of alfalfa on three successive clear, warm days with adequate soil water. Rate of transpiration expressed as grams per hour per 6-ft^2 (0.17-m^2) plot of alfalfa. 1 g = 0.035 oz. (*After M. D. Thomas and G. R. Hill, Plant Physiol., 12:285–307, 1937*)

the rate of transpiration. Whenever soil conditions are such that the rate of absorption of water is retarded, there is a corresponding diminution in the rate of transpiration.

Daily periodicity of transpiration. The rate of every major plant process, including transpiration, is measurably and often markedly influenced by the environmental conditions to which the plant is exposed. Many of the environmental factors exhibit more or less regular daily periodicities, which vary somewhat, of course, with the prevailing climatic conditions. This is especially true of the factors of light and temperature. Many plant processes, including transpiration, therefore exhibit daily periodicities in rate that are correlated with the daily periodicities of one or more environmental factors. *See* PHOTOPERIODISM.

The daily periodicity of transpiration in alfalfa, as exhibited on three clear, warm days with adequate soil water available, is illustrated in **Fig. 2**. A similar daily periodicity of transpiration is exhibited under comparable environmental conditions by many other species. During the hours of darkness, the transpiration rate is relatively low, and in most species water vapor loss during this period may be regarded as entirely cuticular or nearly so. The transpiration rate shows a steady rise during the morning hours, culminating in a peak rate which is attained in the early hours of the afternoon. The increase in transpiration rate during the forepart of the day results from gradual opening of the stomates, beginning with the advent of light, followed by a steady increase in the steepness of the vapor pressure gradient through the stomates, which occurs as a result of increasing atmospheric temperature during the morning and earlier afternoon hours.

In most plants, if transpiration is occurring rapidly, the rate of absorption of water does not keep pace with the rate at which water vapor is lost from the leaves. In other words, the plant is gradually being depleted of water during the daylight hours. In time the resulting decrease in the water content of the leaf cells results in a reduced vapor pressure within the intercellular spaces, and a diminution in the rate of transpiration begins. Stomates also start to close as a result of the diminished leaf water content, and their closure is accelerated during the later part of the afternoon by the waning light intensity. By nightfall complete closure of virtually all stomates has taken place, and water vapor loss during the hours of darkness is again restricted largely or entirely to the relatively low rate of cuticular transpiration. It is noteworthy that the peak rate of transpiration occurs during the early afternoon hours, correlating more closely with the daily temperature periodicity than with the daily periodicity of light intensity.

Under environmental conditions differing considerably from those postulated in the preceding discussion, patterns of transpiration periodicity may show a considerable variance from the one described. On cloudy days, for example, stomates generally open less completely than on clear days, and a curve for daily transpiration periodicity presents a greatly flattened appearance as compared with those shown in Fig. 2. A cool temperature, even in a range somewhat above freezing, greatly diminishes and may even cause cessation of stomatal transpiration, resulting in a pronounced modification in the daily march of transpiration periodicity. A deficient soil water supply is probably the most common cause of departures from the pattern of transpiration described above. A reduction in soil water content below the field capacity (optimum water availability) results not only in a general flattening of the transpiration periodicity curve, but frequently also in appearance of the peak of the curve somewhat earlier in the day. Since, even in temperate zone regions, drought periods of greater or less severity are of common occurrence during the summer months and in many habitats are the rule rather than the exception, transpiration periodicity curves of this flattened and skewed-peak type are undoubtedly of frequent occurrence.

Magnitude of transpiration. Transpiration of broad-leaved species of plants growing under temperate zone conditions may range up to about 0.0102 oz per square inch (5 g per square decimeter) of leaf area per hour. Sufficient quantities of water are often lost in transpiration by vegetation-covered areas of the Earth's surface to have important effects not only on soil water relations, but also on meteorological conditions. The quantities of water lost per acre by crops, grasslands, or forest are therefore a matter of basic interest. An acre of corn (maize), for example, transpires water equivalent to 15 in. (38 cm) of rainfall during a usual growing season. Transpiration of deciduous, largely oak, forest in the southern Appalachian Mountains has been estimated as equivalent to 17–22 in. (43–56 cm) of rainfall per year. Marked variations occur in such values from year to year, however, depending upon prevailing climatic conditions.

Significance of transpiration. Viewpoints regarding the significance of transpiration have ranged between the two extremes of considering it a process that is an unavoidable evil or a physiological necessity. Neither of these extreme views appears to be tenable. Some of the incidental effects of transpiration appear to be advantageous to the plant, but none of them is indispensable for its survival or even for its adequate physiological operation. Likewise, while some of the incidental effects of transpiration appear to be detrimental to the plant, none of them is so in such a critical fashion that survival of plants,

considered in the aggregate, is endangered.

Transpiration is a necessary consequence of the relation of water to the anatomy of the plant, and especially to the anatomy of the leaves. Terrestrial green plants are dependent upon atmospheric carbon dioxide for their survival. In terrestrial vascular plants the principal carbon dioxide–absorbing surfaces are the moist mesophyll cells walls which bound the intercellular spaces in leaves. Ingress of carbon dioxide into these spaces occurs mostly by diffusion through open stomates. When the stomates are open, outward diffusion of water vapor unavoidably occurs, and such stomatal transpiration accounts for most of the water vapor loss from plants. Although transpiration is thus, in effect, an incidental phenomenon, it frequently has marked indirect effects on other physiological processes which occur in the plant because of its effects on the internal water relations of the plant.

Water Translocation

In terrestrial rooted plants practically all of the water which enters a plant is absorbed from the soil by the roots. The water thus absorbed is translocated to all parts of the plant. In the tallest trees (specimens of the coast redwood, *Sequoia sempervirens*) the distance from the tips of the deepest roots to the tips of the topmost branches is nearly 400 ft (122 m), and water must be elevated for this distance through such trees. Although few plants are as tall as such redwoods, the same mechanisms of water movement are believed to operate in all vascular species. The mechanism of the "ascent of sap" (all translocated water contains at least traces of solutes) in plants, especially tall trees, was one of the first processes to excite the interest of plant physiologists.

Water-conductive tissues. The upward movement of water in plants occurs in the xylem, which, in the larger roots, trunks, and branches of trees and shrubs, is identical with the wood. In the trunks or larger branches of most kinds of trees, however, sap movement is restricted to a few of the outermost annual layers of wood. This explains why hollow trees, in which the central core of older wood has disintegrated, can remain alive for many years. The xylem of any plant is a unit and continuous system throughout the plant. Small strands of this tissue extend almost to the tip of every root. Other strands, the larger of which constitute important parts of the veins, ramify to all parts of each leaf. In angiosperms most translocation of water occurs through the xylem vessels, which are nonliving, elongated, tubelike structures. The vessels are formed by the end-to-end coalescence of many much smaller cells, death of these cells ensuing at about the same time that coalescence occurs. In trees the diameters of such vessels range from about 20 to 400 μm, and they may extend for many feet with no more interruption than an occasional incomplete cross wall. In gymnosperms no vessels are present, and movement of water occurs solely through spindle-shaped xylem cells called tracheids. Vertically contiguous tracheids always overlap along their tapering portions, resulting in a densely packed type of woody tissue. Individual tracheids may be as much as 0.2 in. (5 mm) in length. Like the vessels, they are nonliving while functional in the translocation of water. Small, more or less rounded, thin areas occur in the walls of vessels and tracheids that are contiguous with the walls of other tracheids, vessels, or cells. Structurally, three main types of such pits are recognized, but all of them appear to facilitate the passage of water from one xylem element to another. *See* XYLEM.

Root pressure. The exudation of xylem sap from the stump of a cutoff herbaceous plant is a commonly observed phenomenon. Sap exudation ("bleeding") from the cut ends of stems or from incisions into the wood also occurs in certain woody plants, such as birch, currant, and grape, especially in the spring. A single vigorous grapevine often loses a liter or more of sap per day through the cut ends of stems after spring pruning. This exudation of sap from the xylem tissue results from a pressure originating in the roots, called root pressure. A related phenomenon is that of guttation. This term refers to the exudation of drops of water from the tips or margins of leaves, which occurs in many species of herbaceous plants as well as in some woody species. Like sap exudation from cut stems, this phenomenon is observed most frequently in the spring, and especially during early morning hours. The water exuded in guttation is not pure, but contains traces of sugar and other solutes. Guttation occurs from special structures called hydathodes, which are similar in structure to but larger than stomates. In most species water loss by guttation is negligible in comparison with the water lost as vapor in transpiration. Like xylem sap exudation, guttation results from root pressure.

Root pressure is generally considered to be one of the mechanisms of upward transport of water in plants. While it is undoubtedly true that root pressure does account for some upward movement of water in certain species of plants at some seasons, various considerations indicate that it can be only a secondary mechanism of water transport. Among these considerations are the following: (1) There are many species in which the phenomenon of root pressure has not been observed. (2) The magnitude of measured root pressures seldom exceeds 2 atm (200 kPa), which could not activate a rise of water for more than about 60 ft, and many trees are much taller than this. (3) Known rates of xylem sap flow under the influence of root pressure are usually inadequate to compensate for known rates of transpiration. (4) Root pressures are usually operative in woody plants only during the early spring; during the summer months, when transpiration rates, and hence rates of xylem sap transport, are greatest, root pressures are negligible or nonexistent. *See* ROOT (BOTANY).

Water cohesion and ascent of sap. Although invariably in motion, as a result of their kinetic energy, water molecules are also strongly attracted to each other. In masses of liquid water the existence of such

intermolecular attractions is not obvious, but when water is confined in long tubes of small diameter, the reality of the mutual attractions among water molecules can be demonstrated. If the water at the top of such a tube is subjected to a pull, the resulting stress will be transmitted, because of the mutual attraction (cohesion) among water molecules, all the way down the column of water. Furthermore, because of the attraction between the water molecules and the wall of the tube (adhesion), subjecting the water column to a stress does not result in pulling it away from the wall.

The observations just mentioned have been made the basis of a widely entertained theory of the mechanism of water transport in plants, first clearly enunciated by H. H. Dixon in 1914. According to this theory, upward translocation of water (actually a very dilute sap) is engendered by an increase in the negativity of water potential in the cells of apical organs of plants. Such increases in the negativity of water potentials occur most commonly in the mesophyll cells of leaves as a result of transpiration.

Evaporation of water from the walls of the mesophyll cells abutting on the intercellular spaces reduces their turgor pressure and hence increases the negativity of the water potential in such cells. Consequent cell-to-cell movements of water cause the water potentials even of those mesophyll cells which are not directly exposed to the intercellular spaces to become more negative. The resulting decrease in the water potential of those cells directly in contact with the xylem elements in the veinlets of the leaf induces movement of water from the vessel or tracheids into those cells. Whenever transpiration is occurring at appreciable rates, water does not enter the lower ends of the xylem conduits in the roots as rapidly as it passes out of the vessels or tracheids into adjacent leaf cells at the upper ends of the water-conductive system; therefore the water in the xylem ducts is stretched into taut threads, that is, it passes into a state of tension. Each column of water behaves like a tiny stretched wire. The tension is transmitted along the entire length of the water columns to their terminations just back of the root tips. Subjection of the water in the xylem ducts to a tension (in effect, a negative pressure) increases the negativity of its water potential. The water potential in the xylem conduits is also subject to some lowering because of the presence of solutes, but xylem sap is usually so dilute that this osmotic effect is rarely more than a minor one. As a result of the increased negativity of the water potential in the xylem elements, movement of water is induced from adjacent root cells into those elements in the absorbing regions of roots.

The tension engendered in the water columns can be sustained by them because of the cohesion between the water molecules, acting in conjunction with the adhesion of the boundary layers of water molecules to the walls of the xylem ducts. The existence of water under tension in vessels has been verified in a number of species of plants by direct microscopic examination. There is some evidence that, under conditions of marked internal water deficiency, the tensions generated in the water columns are proliferated into the mesophyll cells of leaves and cells in the absorbing regions of roots. Conservative calculations indicate that a cohesion value of 30–50 atm (3–5 MPa) would be adequate to permit translocation of water to the top of the tallest known trees. However, under conditions of internal water deficiency, tensions considerably in excess of 50 atm (5 MPa) are probably engendered in the water columns of many plants, especially woody species. *See* PLANT TRANSPORT OF SOLUTES.

Water Absorption

This process will be discussed only from the standpoint of terrestrial, rooted plants. Consideration of the absorption of water by plants necessitates an understanding of the physical status of the water in soils as it exists under various conditions.

Soil water conditions. Even in the tightest of soils, the particles never fit together perfectly and a certain amount of space exists among them. This pore space of a soil ranges from about 30% of the soil volume in sandy soils to about 50% of the soil volume in heavy clay soils. In desiccated soils the pore space is occupied entirely by air, in saturated soils it is occupied entirely by water, but in moist, well-drained soils it is usually occupied partly by air and partly by water. In a soil in which a water table is located not too far below the surface, considerable quantities of water may rise into its upper layers by capillarity and become available to plants. In arid regions, however, there ordinarily is no water table. Even in many humid regions the water table is continuously or intermittently too far below the soil surface to be an appreciable source of water for most plants. In all soils lacking a water table or in which the water table is at a considerable depth, the only water available to plants is that which comes as natural precipitation or which is provided by artificial irrigation. If water falls on, or is applied to, a dry soil which is homogeneous to a considerable depth, it will become rapidly distributed to a depth which will depend on the quantity of water supplied per unit area and on the specific properties of that soil. After several days further deepening of the moist layer of soil extending downward from the surface virtually ceases, because within such a time interval capillary movement of water in a downward direction has become extremely slow or nonexistent. The boundary line between the moist layer of soil above and the drier zone below will be a distinct one. In this condition of field equilibrium the water content of the upper moist soil layer is, in homogeneous soils, essentially uniform throughout the layer. *See* SOIL.

The water content of a soil in this equilibrium condition is called the field capacity. Field capacities range from about 5% in coarse sandy soils to about 45% of the dry weight in clay soils. The moisture equivalent of a soil, often measured in the laboratory, is usually very close in value to the field capacity of the same soil. It is defined as the water content of a soil which is retained against a force

1000 times gravity as measured in a centrifuge. A soil at its field capacity is relatively moist, but is also well aerated. Soil water contents at or near the field capacity are the most favorable for growth of most kinds of plants.

A considerable proportion of the water in any soil is unavailable in the growth of plants. The permanent wilting percentage is the generally accepted index of this fraction of the soil water. This quantity is measured by allowing a plant to develop with its roots in soil enclosed in a waterproof pot until the plant passes into a state of permanent wilting. The water content of the soil when the plant just passes into this condition is the permanent wilting percentage. The range of permanent wilting percentages is from 2 to 3% of the dry weight in coarse sandy soils to about 20% in heavy clay loams. About the same value is obtained for the permanent wilting percentage of a given soil, regardless of the kind of test plant used.

The potential of soil water is always negative and has two major components. One is the osmotic pressure of the soil solution, which in most kinds of soils is only a fraction of 1 atm (0.1 MPa), although saline and alkali soils are marked exceptions. The other component is the attractive forces between the soil particles and water molecules, which may attain a very considerable magnitude, especially in dry soils. In moist soils, those at the field capacity or higher soil water content, the former of these two components is principally responsible for the negative soil water potential. In drier soils the latter component is almost solely responsible for the negative water potential. In the majority of soils the water potential is only slightly less than zero at saturation, not less than −1 atm (−0.1 MPa) at field capacity, and in the vicinity of −15 atm (−1.5 MPa) at the permanent wilting percentage. With further reduction in the water content of a soil below its permanent wilting percentage, its water potential increases rapidly in negativity at an accelerating rate. Almost no water can be absorbed by plants from soils with such extremely negative water potentials; for this reason the permanent wilting percentage is the index of soil water unavailable to plant growth (**Fig. 3**).

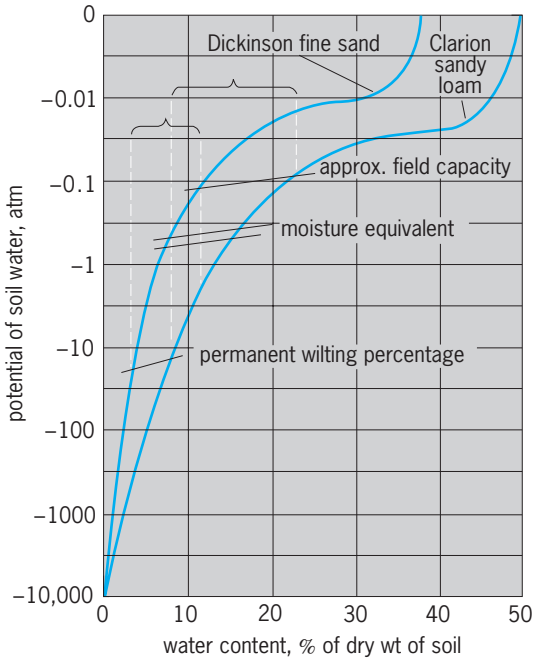

Fig. 3. Relation between soil water potentials and soil water contents of two soils over the entire range of soil water contents. Horizontal braces show the range for each soil over which water is readily available to its plants. Vertical scale is logarithmic. 1 atm = 10^3 kPa. (*After M. B. Russell, Amer. Soil Sci. Proc., 4:51–54, 1939*)

Root growth and water absorption. The successively smaller branches of the root system of any plant terminate ultimately in the root tips, of which there may be thousands and often millions on a single plant. As generally employed, the term root tip refers to the region extending back from the apex of the root for a distance of at least several centimeters. The terminal zone of a root tip is the root cap. Just back of this are the regions in which cell division and cell elongation occur and in which all growth in length of roots takes place. Just back of these regions, in the majority of species, is the zone of root hairs. Each root hair is a projection from the epidermal cell of which it is an integral part. A single root tip may bear thousands of root hairs, ranging in length from a few millimeters up to about a centimeter. In most species the root hairs are short-lived structures, but new ones are constantly developing just back of the growing region of the root as it elongates. Most absorption of water occurs in the root tip regions, and especially in the root hair zone. Older portions of most roots become covered with cutinized or suberized layers through which only very limited quantities of water can pass.

Whenever the potential of the water in the root hairs and other peripheral cells of a root tip is less than that of the soil water, movement of water takes place into the root cells. If the soil water content exceeds the field capacity, water may move by capillarity toward the region of absorption from portions of the soil not immediately contiguous with the root tips, and the supply of readily absorbable water is maintained in this way. Elongation of the roots, although slower in most species in relatively wet soils, also helps maintain contact between the root tips and untapped areas of soil water.

Many plants much of the time grow in soils with a water content in the range between the field capacity and the permanent wilting percentage. In this range of soil water contents, capillary movement of water through the soil is extremely slow or nonexistent, and an adequate supply of water cannot be maintained to rapidly absorbing root tips by this means. In such soils maintenance of contact between the root tips and available soil water is assured only by continued elongation of the roots through the soil. Mature root systems of many plants terminate in millions of root tips, each of which may be visualized as slowly growing through the soil, absorbing water from around or between the soil particles with which it comes in contact. The aggregate increase in the length of the root system of a rye plant averages 3.1 mi/day (5 km/day). Calculations indicate that the

daily aggregate root elongation of this plant is adequate to permit absorption of a sufficient quantity of water from soils at the field capacity to compensate for daily transpirational water loss.

Mechanisms of water absorption. As previously indicated, tension, generated in the water columns of a plant, most commonly as an indirect result of transpiration, is transmitted to the ultimate terminations of the xylem ducts in the root tips. The water potential in the water columns is diminished by the amount of the tension generated, which is equivalent to a negative pressure. As soon as the water potential in the water columns becomes less than that in contiguous cells in the root tip, water moves from those cells into the xylem. This activates further cell-to-cell movement of water in a lateral direction across the root and presumably in the establishment of gradients of water potentials, diminishing progressively in magnitude from the epidermal cells, including the root hairs, to the root xylem. Whenever the water potential in the peripheral root cells is less than that of the soil water, movement of water from the soil into the root cells occurs. There is some evidence that, under conditions of marked internal water stress, the tension generated in the xylem ducts will be propagated across the root to the peripheral cells. If this occurs, water potentials of greater negativity could develop in peripheral root cells than would otherwise be possible. The absorption mechanism would operate in fundamentally the same way whether or not the water in the root cells passed into a state of tension. The process just described, often called passive absorption, accounts for most of the absorption of water by terrestrial plants.

The phenomenon of root pressure, previously described as the basis for xylem sap exudation from cuts or wounds and for guttation, represents another mechanism of the absorption of water. This mechanism is localized in the roots and is often called active absorption. Water absorption of this type only occurs when the rate of transpiration is low and the soil is relatively moist. Although the xylem sap is a relatively dilute solution, its osmotic pressure is usually great enough to engender a more negative water potential than usually exists in the soil water when the soil is relatively moist. A gradient of water potentials can thus be established, increasing in negativity across the epidermis, cortex, and other root tissues, along which the water can move laterally from the soil to the xylem. There is evidence, however, that a respiration mechanism, as well as an osmotic one, may be involved in the correlated phenomena of active absorption, root pressure, and guttation.

Environment and absorption. Any factor which influences the rate of transpiration also influences the rate of absorption of water by plants and vice versa. Climatic conditions may therefore indirectly affect rates of water absorption, and soil conditions indirectly affect transpiration. Low soil temperatures, even in a range considerably above freezing, retard the rate of absorption of water by many species. The rate of water absorption by sunflower plants, for example, decreases rapidly as the soil temperature drops below 55°F (13°C).

Within limits, the greater the supply of available soil water, the greater is the possible rate of water absorption. High soil water contents, especially those approaching saturation, result in decreased water absorption rates in many species because of the accompanying deficient soil aeration. In the atmosphere of such soils the oxygen concentration is lower and the carbon dioxide concentration is higher than in the atmosphere proper. In general, the deficiency of oxygen in such soils appears to be a more significant factor in causing retarded rates of water absorption than the excess of carbon dioxide. This retarding effect on water absorption rate is correlated with a retarding effect on root respiration rate. *See* PLANT RESPIRATION.

Likewise, if the soil solution attains any considerable concentration of solutes, water absorption by the roots is retarded. In most soils the concentration of the soil solution is so low that it is a negligible factor in affecting rates of water absorption. In saline or alkali soils, however, the concentration of the soil solution may become equivalent to many atmospheres, and only a few species of plants are able to survive when rooted in such soils.

Wilting. Daily variations in the water content of plants, more marked in some organs than in others, are of frequent occurrence. The familiar phenomenon of wilting, exhibited by the leaves and sometimes other organs of plants, particularly herbaceous species, is direct visual evidence of this fact. In hot, bright weather the leaves of many species of plants often wilt during the afternoon, only to regain their turgidity during the night hours, even if no additional water is provided by rainfall or irrigation. This type of wilting reaction is referred to as a temporary, or transient, wilting and clearly results from a rate of transpiration in excess of the rate of water absorption during the daylight hours. As a result, the total volume of water in the plant shrinks, although not equally in all organs or tissues. In general, diminution in water content is greatest in the leaf cells, and wilting is induced whenever the turgor pressure of the leaf cells is reduced sufficiently.

Even on days when visible wilting is not discernible, incipient wilting is of frequent occurrence. Incipient wilting corresponds to only a partial loss of turgor by the leaf cells and does not result in visible drooping, folding, or rolling of the leaves. Leaves entering into the condition of transient wilting always pass first through the stage of incipient wilting. Occurrence of this invisible first stage of wilting is almost universal on bright, warm days on which environmental conditions are not severe enough to induce the more advanced stage of transient wilting.

Confirmation of the inferred cause of transient wilting has been furnished by investigations of the comparative daily periodicities of transpiration and absorption of water (**Fig. 4**). As illustrated in this figure, there is a distinct lag in the rate of absorption of water as compared with the rate of transpiration.

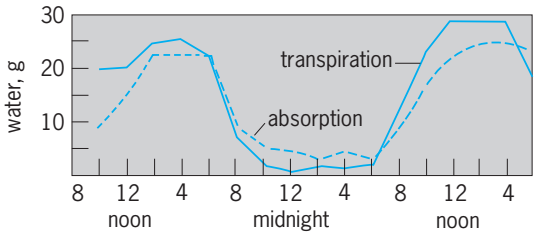

Fig. 4. Comparative daily periodicities of transpiration of water and absorption of water in the loblolly pine (*Pinus taeda*). 1 g = 0.035 oz. (*After P. J. Kramer, Plant and Soil Water Relationships, McGraw-Hill, 1937*)

During the daylight hours the tissues of the plant are being progressively depleted of water, whereas the store of water within the plant is being steadily replenished during the night hours. The lag in the rate of absorption behind the rate of transpiration during the daylight hours appears to result largely from the relatively high resistance of the living root cells to the passage of water across them.

Both incipient and transient wilting should be distinguished from the more drastic stage of permanent wilting. This stage of wilting is attained only when there is an actual deficiency of water in the soil, and a plant will not recover from permanent wilting unless the water content of the soil in which it is rooted increases. In a soil which is gradually drying out, transient wilting slowly grades over into permanent wilting. Each successive night recovery of the plant from temporary wilting takes longer and is less complete, until finally even the slightest recovery fails to take place during the night.

Although the stomates are generally closed during permanent wilting, cuticular transpiration continues. Plants in a state of permanent wilting continue to absorb water, but at a slow rate. Restoration of turgidity is not possible, however, because the rate of transpiration even from a wilted plant exceeds the rate of absorption of water from a soil at the permanent wilting percentage or lower water content. During permanent wilting, therefore, there is a slow but steady diminution in the total volume of water within the plant, and a gradual intensification of the stress in the hydrodynamic system. Tensions in the water columns of permanently wilted plants are relatively high and have been estimated to attain values of 200 atm (20 MPa) in some trees, although this is probably an extreme figure.

As previously mentioned, the permanent wilting percentage, an important index of soil water conditions, is defined as the soil water content when a plant just enters the condition of permanent wilting. The sunflower is the most commonly used test plant in making determinations of the permanent wilting percentage of a soil. Permanent wilting of the basal pair of leaves, judged to have occurred when they fail to recover if placed in a saturated atmosphere overnight, is taken as the critical point in the measurement. The range of soil water contents between the first permanent wilting of the basal leaves of sunflower plants and the permanent wilting of all the leaves is called the wilting range. The water content of the soil at the time all the sunflower leaves have become wilted is termed the ultimate wilting point. In general, the wilting range is narrower in coarse-textured soils than in fine-textured soils, and may be 10–30% of the soil water content between the field capacity and the ultimate wilting point. Although plants cannot grow while the soil in which they are rooted is in the wilting range, many kinds of plants can survive for considerable periods under such conditions. This is especially true of many shrubby species indigenous to semidesert areas.

Internal redistributions of water. For convenience, the processes of transpiration, translocation of water, and absorption of water are often discussed separately, although there is a close interrelationship among these three processes. The hydrodynamic system of a plant is essentially a unit in its operation, and changes in the status of the water in one part of a plant are bound to have effects on its status in other parts of the plant.

Whenever a plant is saturated, or nearly so, with water, differences in water potential from one organ or tissue to another are minimal. But whenever the rate of absorption of water lags behind the rate of transpiration, an internal water deficit develops in the hydrodynamic system of the plant, which in turn favors the establishment of marked differences in water potential from one part of the plant to another. Under such conditions redistribution of some of the water present from some tissues or organs of a plant to others generally occurs.

Internal movements of water from fruits to leaves and vice versa seem to be of common occurrence. Mature lemon fruits, while still attached to the tree, exhibit a daily cycle of expansion and contraction (**Fig. 5**). The lemon fruits begin to contract in volume early in the morning and continue to do so until late afternoon. Since transpirational loss from a lemon fruit is negligible, it is obvious that during this part of the day, corresponding to the period of high transpiration rates from the leaves, water is moving out of the fruits into other parts of the tree. Most of this movement probably occurs into leaves. During the daylight hours, the water potentials of the leaf cells presumably decrease until they soon are lower than those of the fruit cells, thus initiating movement of water from the fruits to the leaves. During the late afternoon and night hours, the volume of the fruits gradually increases, indicating that water is now moving back into the fruits. During this period, transpirational water loss from the leaves is small, leaf water contents increase, and the water potential of the leaf cells becomes less negative. Less of the absorbed water is translocated to the leaves than during the daylight hours, and more can move into the fruits, despite the relatively high water potential of the fruit cells. Marked daily variations take place in the diameters of lemon fruits, even under environmental conditions which result in no observable wilting of the leaves.

In growing cotton bolls, however, as long as enlargement is continuing, increase in diameter

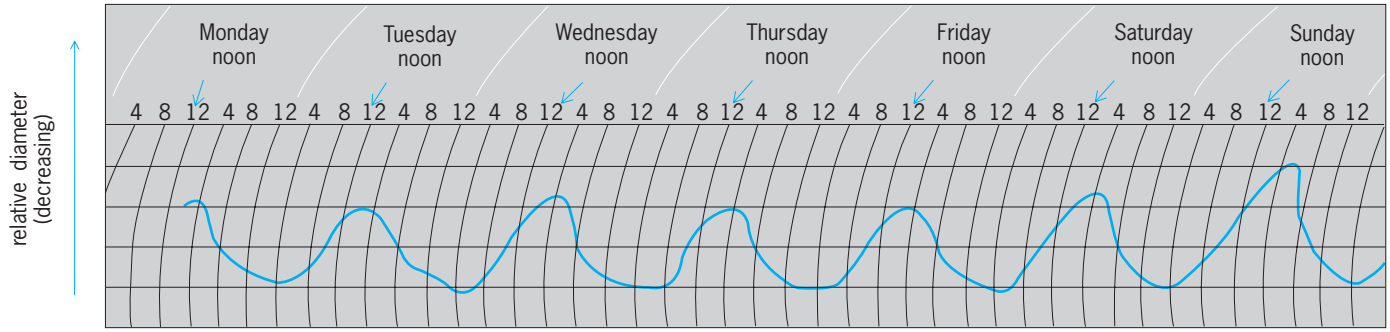

Fig. 5. Daily variations in diameter of lemons. (*After E. T. Bartholomew, Amer. J. Bot., 13:102–117, 1926*)

continues steadily both day and night and even during periods when the leaves are severely wilted. Movement of water is obviously occurring into the growing bolls without interruption during this period. Once the bolls cease enlarging, however, reversible daily changes occur in their diameter, similar in pattern to those which take place in mature lemon fruits (**Fig. 6**). Similarly, it has been shown that in a tomato plant the topmost node, within which growth in length occurs, continues to elongate at approximately the same rate both day and night. The stem below the first node, however, shrinks measurably in length during the daytime and elongates equally at night, undoubtedly as a result of reversible changes in the turgidity of the stem cells. The growing cells in the terminal node of the tomato stem continue to obtain water during the daylight hours while the rest of the stem is losing water, and some water utilized in their growth probably comes from the cells of the lower nodes.

In general, as the last two examples illustrate, actively meristematic regions such as growing stems and root tips and enlarging fruits, under conditions of internal water deficiency, apparently develop more negative water potentials than other tissues. Hence water often continues to move toward such regions even when an internal water deficiency of considerable magnitude has developed within the plant. However, under conditions of drastic internal water deficit, approaching or corresponding to a state of permanent wilting, growth of all meristems is greatly retarded or inhibited. *See* APICAL MERISTEM.

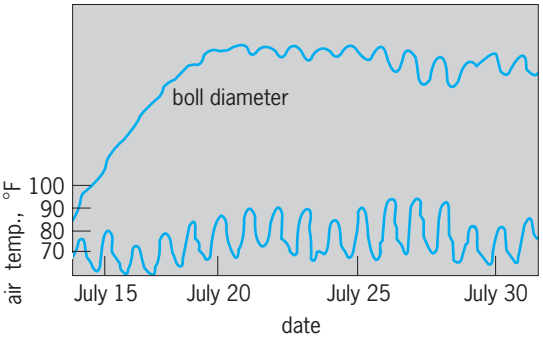

Fig. 6. Daily variations in diameter of a cotton boll (still growing for first 5 days). °C = (°F − 32)/1.8. (*After D. B. Anderson and T. Kerr, Plant Physiol., 18:261–269, 1943*)

Drought resistance. The term drought refers, in general, to periods during which the soil contains little or no water which is available to plants. In relatively humid climates such periods are infrequent and seldom of long duration except in certain local habitats. The more arid a climate, in general, the more frequent is the occurrence of droughts and the longer their duration. Most species of plants can survive short dry periods without serious injury, but a prolonged period of soil water deficiency is highly injurious or lethal to all except those species of plants that have a well-developed capacity for drought resistance.

Most species which grow in semidesert regions, such as those of the southwestern United States and adjacent Mexico, or in locally dry habitats must be drought resistant in one sense of the word or another. Annual species which grow in many arid regions during short rainy spells are exceptions to this statement. Such species complete their brief life cycle from seed to seed during a period when soil water is available, and they can survive in arid regions because they evade rather than endure drought.

Succulents, such as cacti, are found in most semidesert regions, and are also often indigenous to locally dry habitats such as sand dunes and beaches in humid climate regions. Succulents are able to survive long dry periods because of the relatively large quantities of water which accumulate, in some species in fleshy stems and in other species in fleshy leaves, during the occasional periods when soil water is available. Many succulents can live for months on such stored water.

Those plants which are drought resistant in the truest sense are those whose cells can tolerate a marked reduction in water content for extended periods of time without injury. Many shrubby species of semidesert regions have this property. Certain structural features undoubtedly aid in the survival of such plants for long periods of arid habitats. Many xerophytes (plants that can endure periods of drought) have extensive root systems in proportion to their tops; such a structural characteristic aids in maintaining a supply of water to the aerial portions of the plant longer than would otherwise be possible. Other drought-resistant species are characterized by having diminutive leaves; the transpiring surface of

the plant may thus be small relative to the absorptive capacity of the roots. In still other species the leaves abscise (fall) with the advent of the dry season, thus greatly reducing transpiration in the period of greatest internal stress in the hydrodynamic system.

Despite any structural features which may help maintain their internal water supply, shrubby plants of semiarid regions regularly undergo a gradual depletion in the store of water within them and a gradual intensification of the stress prevailing in the internal hydrodynamic system over dry periods often lasting for months. Only drought-resistant species can endure this condition, which is in essence a state of permanent wilting, for long periods of time without injury. A fundamental factor in the drought resistance of plants therefore appears to be a capacity of the cells to endure a substantial reduction in their water content without suffering injury. This capacity probably is based in part on structural features of the cells of such species and in part on the distinctive physiological properties of their protoplasm. *See* ECOLOGY; PLANT MINERAL NUTRITION.

Bernard S. Meyer

Bibliography. J. Kolek (ed.), *Physiology of the Plant Root System*, 1992; T. T. Kozlowski and A. J. Riker (eds.), *Flooding and Plant Growth*, 1984; T. T. Kozlowski and A. J. Riker (eds.), *Deficits and Plant Growth*, 1983; P. J. Kramer (ed.), *Water Relations of Plants*, 1983; E. Weiss, *Plants Tolerant to Arid and Semiarid Conditions*, 1988.

Plantaginales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Asteridae of the class Magnoliopsida (dicotyledons). The order consists of only the family Plantaginaceae, with about 250 species. Within its subclass the order is marked by its small, chiefly wind-pollinated flowers that have a persistent regular scarious corolla. The perianth and stamens of the flowers are attached directly to the receptacle (hypogynous) and there are typically four petals that are joined at the base to form a tube. The plants are herbs or seldom half-shrubs with mostly basal, alternate leaves. The common plantain (*Plantago major*) is a familiar lawn weed of this order. *See* ASTERIDAE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Plants, life forms of

A term for the vegetative (morphological) form of the plant body. A related term is growth form but a theoretical distinction is often made: life form is thought by some to represent a basic genetic adaptation to environment, whereas growth form carries with it no connotation of adaptation and is a more general term applicable to structural differences.

Life-form systems are based on differences in gross morphological features, and the categories bear no necessary relationship to reproductive structures, which form the basis for taxonomic classification. Features used in establishing life-form classes include deciduous versus evergreen leaves, broad versus needle leaves, size of leaves, degree of protection afforded the perennating tissue, succulence, and duration of life cycle (annual, biennial, or perennial). Thus the garden bean (family Leguminosae) and tomato (Solanaceae) belong to the same life form because each is an annual, finishing its entire life cycle in 1 year, while black locust (Leguminosae) and black walnut (Juglandaceae) are perennial trees with compound, deciduous leaves.

Climate and adaptation factors. There is a clear correlation between life forms and climates. For example, broad-leaved evergreen trees clearly dominate in the hot humid tropics, whereas broad-leaved deciduous trees prevail in temperature climates with cold winters and warm summers, and succulent cacti dominate American deserts. Although cacti are virtually absent from African deserts, members of the family Euphorbiaceae have evolved similar succulent life forms. Such adaptations are genetic, having arisen by natural selection. However, since there are no life forms confined only to a specific climate and since it is virtually impossible to prove that a given morphological feature represents an adaptation with survival value, some investigators are content to use life forms only as descriptive tools to portray the form of vegetation in different climates.

Raunkiaer system. Many life-form systems have been developed. Early systems which incorporated many different morphological features were difficult to use because of this inherent complexity. The most successful and widely used system is that of C. Raunkiaer, proposed in 1905; it succeeded where others failed because it was homogeneous and used only a few obvious morphological features representing important adaptations.

Reasoning that it was the perennating buds (the tips of shoots which renew growth after a dormant season, either of cold or drought) which permit a plant to survive in a specific climate, Raunkiaer's classes were based on the degree of protection afforded the bud and the position of the bud relative to the soil surface (see **illus.**). They applied to autotrophic, vascular, self-supporting plants. Raunkiaer's classificatory system is:

- Phanerophytes: bud-bearing shoots in the air, predominantly woody trees and shrubs; subclasses based on height and on presence or absence of bud scales
- Chamaephytes: bud within 10 in. (25 cm) of the surface, mostly prostrate or creeping shrubs
- Hemicryptophytes: buds at the soil surface, protected by scales, snow, and litter
- Cryptophytes: buds underneath the soil surface or under water
- Therophytes: annuals, the seed representing the only perennating tissue

Examples of life-form spectra for floras of different climates

Climate and vegetation	Area	Life form*				
		Ph	Ch	H	Cr	Th
Normal spectrum	World	46	9	26	6	13
Tropical rainforest	Queensland, Australia	96	2	0	2	0
	Brazil	95	1	3	1	0
Subtropical rainforest (monsoon)	India	63	17	2	5	10
Hot desert	Central Sahara	9	13	15	5	56
Mediterranean	Crete	9	13	27	12	38
Steppe (grassland)	Colorado, United States	0	19	58	8	15
Cool temperate (deciduous forest)	Connecticut, United States	15	2	49	22	12
Arctic tundra	Spitsbergen, Norway	1	22	60	15	2

*Ph = phanerophyte; Ch = chamaephyte; H = hemicryptophyte; Cr = geophyte (cryptophyte); and Th = therophyte.

Subclasses were established in several categories, and Raunkiaer later incorporated leaf-size classes into the system.

By determining the life forms of a sample of 1000 species from the world's floras, Raunkiaer showed a correlation between the percentage of species in each life-form class present in an area and the climate of the area. The results (see **table**) were expressed as a normal spectrum, and floras of other areas were then compared to this. Raunkiaer concluded that there were four main phytoclimates: phanerophyte-dominated flora of the hot humid tropics, hemicryptophyte-dominated flora in moist to humid temperate areas, therophyte-dominated flora in arid areas, and a chamaephyte-dominated flora of high latitudes and altitudes.

Subsequent studies modified Raunkiaer's views. (1) Phanerophytes dominate, to the virtual exclusion of other life forms, in true tropical rainforest floras, whereas other life forms become proportionately more important in tropical climates with a dry season, as in parts of India. (2) Therophytes are most abundant in arid climates and are prominent in temperate areas with an extended dry season, such as regions with Mediterranean climate (for example, Crete). (3) Other temperate floras have a predominance of hemicryptophytes with the percentage of phanerophytes decreasing from summer-green deciduous forest to grassland. (4) Arctic and alpine tundra are characterized by a flora which is often more than three-quarters chamaephytes and hemicryptophytes, the percentage of chamaephytes increasing with latitude and altitude.

Most life forms are present in every climate, suggesting that life form makes a limited contribution to adaptability. Determination of the life-form composition of a flora is not as meaningful as determination of the quantitative importance of a life form in vegetation within a climatic area. However, differences in evolutionary and land-use history may give rise to floras with quite different spectra, even though there is climatic similarity. Despite these problems, the Raunkiaer system remains widely used for vegetation description and for suggesting correlations between life forms, microclimate, and forest site index. *See* PLANT GEOGRAPHY.

Mapping systems. There has been interest in developing systems which describe important morphologic features of plants and which permit mapping and diagramming vegetation. Descriptive systems incorporate essential structural features of plants, such as stem architecture and height; deciduousness; leaf

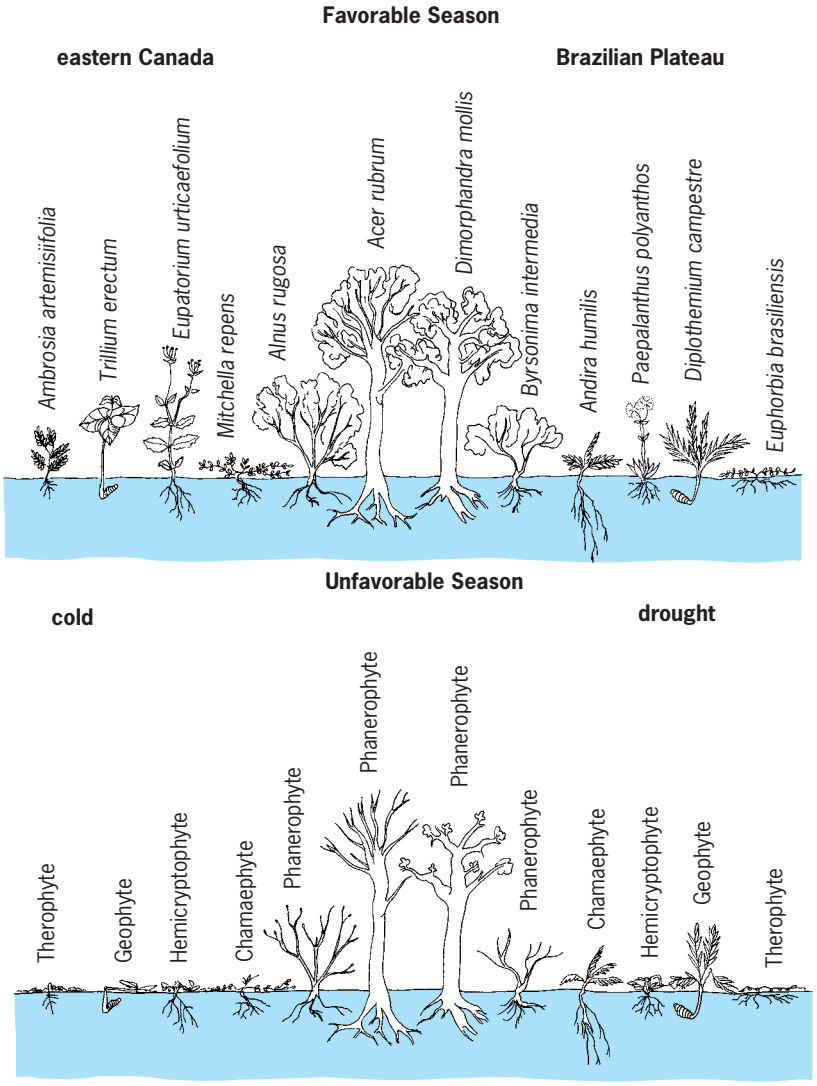

Life forms of plants according to C. Raunkiaer. (*After P. Dansereau, Biogeography: An Ecological Perspective, Ronald Press, 1957*)

texture, shape, and size; and mechanisms for dispersal. These systems are important in mapping vegetation because structural features generally provide the best criteria for recognition of major vegetation units. *See* ALTITUDINAL VEGETATION ZONES; VEGETATION AND ECOSYSTEM MAPPING. Arthur W. Cooper

Bibliography. H. C. Bold, *Morphology of Plants and Fungi*, 5th ed., 1990; S. A. Cain, Life forms and phytoclimate, *Bot. Rev.*, 16(1):1–32, 1950; P. Dansereau, *A Universal System for Recording Vegetation*, 1958; R. Daubenmire, *Plant Communities*, 1968; C. Raunkiaer, *The Life Forms of Plants and Statistical Geography*, 1934.

Plants of saline environments

Plants that have evolved to flourish where the salt concentrations (notably those of sodium chloride) are 400 mol m^{-3} and beyond. These plants are known as halophytes. The majority of plant species are unable to tolerate one-quarter of this salt concentration and are called glycophytes.

Saline environments. The greater proportion of the surface of the Earth is covered not by land but by water, and this water contains a considerable concentration of dissolved salts (see **table**). Although many different ions are present in seawater, the point here is the high concentration of sodium and chloride ions (about 500 mol m^{-3}). These ions predominate in most saline environments, since the sea can affect low-lying coastal regions and moreover may have had a historical influence on land no longer close to an ocean. Other ions that are commonly found, particularly in saline ground waters, are potassium, calcium, and magnesium, together with sulfates, carbonates, and borates. Two important categories of soil resulting from the presence of these salts in combinations are sodic soils, dominated by sodium carbonate (pH may be greater than 10), and acid sulfate soils, dominated by sulfates (pH may be as low as 3). Saline soils are dominated by sodium chloride itself. All the salt-affected soils tend to occur in lower latitudes where evapotranspiration rates are high.

Apart from deserts and oceans, the other most extensive naturally saline areas are coastal marshes which are periodically inundated by seawater; this process leads to mangrove swamps in tropical climates. *See* BOG; SALT MARSH; SWAMP, MARSH, AND BOG.

Chief cations and anions in ocean seawater with a salinity of 35%

Cations and anions	Concentration, mol m^{-3}
Sodium	483
Magnesium	55
Calcium	10
Potassium	10
Strontium	0.1
Chlorine	558
Sulfate	29
Bromine	0.8
Borate	0.4

Anthropogenically formed saline soils have become more common in areas where irrigation is practiced. Unfortunately, the most affected areas are also areas of high crop productivity, such as the Mediterranean basin, California, and southern Asia. These areas have optimal day length and temperatures, but rainfall is often limiting, necessitating irrigation (**Fig. 1***a*). Soluble salts in irrigation water plus those present in fertilizers contribute to the gradual accumulation of toxic levels of sodium ions (Na^+) in topsoil, adversely affecting many crops (Fig. 1*b*). Plants sensitive to salt may die or grow poorly. This problem has led to losses of otherwise good agricultural land since, once saline, the soil becomes relatively impermeable to water as the structure breaks down, and this condition is difficult to treat. Of note, all crops relevant for agriculture are glycophytes. *See* SALINE EVAPORITES.

Plants. There are many plant species adapted to saline environments. All those that normally grow in seawater are clearly halophytes. However, it is sometimes more difficult to classify terrestrial plants based on observations of their natural habitat, since the concentration of salts varies so much with the weather. It is possible to be more precise if the plants are grown in solution culture under controlled conditions of temperature, humidity, and light intensity. Under these circumstances, those which complete their life cycle in the presence of about 1% sodium chloride (say, 170 mol m^{-3}) are undoubtedly halophytes. Plants that would not survive in a concentration of 100 mol m^{-3} sodium chloride would be classed as glycophytes (**Fig. 2**); in between these concentrations is a gray area occupied by salt-tolerant glycophytes.

Marine flora. The flora of the oceans is dominated by algae, with few flowering plants adapted to this environment. The angiosperms and benthic algae are restricted to the ocean fringes, which make up about 2% of the total area. The remainder of the deep water is home to the phytoplankton, mainly diatoms, dinoflagellates, and silicoflagellates. The algae of the coastal regions are predominantly brown (Phaeophyta) or red (Rhodophyta), although the green algae (Chlorophyta) are well represented. The Cyanophyta and the Euglenophyta are largely confined to freshwater, although salt-tolerant genera are known. The distribution of the various algal species depends on the temperature of the water and the degree of exposure to air which the algal thallus (main body) can withstand. This gives rise to the zonation of algae seen on the seashore. *See* ALGAE.

Apart from the algae, the marine flora is restricted to a few angiosperms; bryophytes, pteridophytes, and gymmosperms are completely absent. As far as the angiosperms are concerned, there are no marine dicotyledonous plants and only one order of monocotyledons, the Helobiae, with some 12 genera and about 49 species. Most of these sea grasses grow in warmer waters, and since they cannot withstand long exposure to air, they occur at or about the level of the lowest tides. As far as is known, they appear to have evolved from terrestrial salt-tolerant plants.

Fig. 1. Pitfalls of artificial irrigation. (*a*) Agriculture in arid and semiarid regions depends on artificial irrigation. (*b*) Gradual accumulation of toxic levels of Na^+ in topsoil renders land useless.

Terrestrial flora. The distinction between terrestrial and aquatic plants is largely a matter of the degree to which the plant is predominantly in air or in water. For example, salt marsh plants are terrestrial but must withstand submergence, while algae of the upper shore are mostly submerged but must withstand periodic exposure to the atmosphere.

The terrestrial halophytes are nearly all flowering plants, and there are few, if any, salt-tolerant mosses, liverworts, or even gymnosperms. Among the ferns, there are just two families with salt-tolerant genera, the Pteridaceae and the Ophioglossaceae. Among the flowering plants, salt tolerance occurs in many different families and appears to have been polyphyletic in origin. Salt-tolerant species are to be found in about one-third of the families of flowering plants. However, about half of the approximately 500 genera known to contain halophytic species belong to just 20 families. Among the monocotyledonous plants, grasses dominate as halophytes with 109 species, followed by salt-tolerant sedges with 83 species. Within the dicotyledonous plants, the Chenopodiaceae is the preeminent family for salt tolerance with 44 of its 100 genera containing salt-tolerant species (there are 312 such species); examples are *Atriplex*, *Salicornia*, and *Suaeda*. Other dicotyledonous families with significant numbers of halophytes are the Asteraceae (53 species) and the Aizoaceae (48 species).

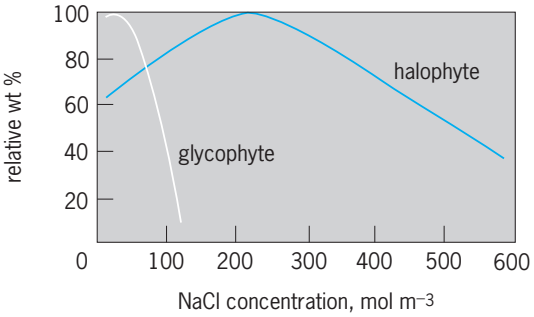

Fig. 2. Relationship between growth and sodium chloride concentration for halophytes and glycophytes. The diagram assumes that the plants are growing in a culture solution which accounts for about 10–20 mol m^{-3} of ions.

It is a notable feature of terrestrial saline environments, however, that many of the genera have worldwide distribution patterns, even though individual species and particular dominants vary with the mean temperature. For example, arctic salt marshes commonly have large populations of *Puccinellia phryganoides*, whereas *P. maritima* is often dominant in temperate marshes. *See* MAGNOLIOPHYTA.

Tropical salt marshes, or mangrove swamps, differ from those of temperate zones in being dominated by trees and shrubs rather than by herbs. There are at least 12 genera from eight different families, and besides their salt tolerance, these trees are remarkable for their root systems, which are adapted to provide stability in soft mud and facilitate gas exchange in an anaerobic substrate (**Fig. 3**). Apart from coastal areas, saline environments are also frequent inland,

Fig. 3. The mangrove *Avicennia marina* in a river estuary at Dangar Island, Australia. Aerial roots are growing out of the mud.

especially in the hotter and drier regions of the world. In such environments, plants must be adapted not only to tolerate the high concentrations of salts in the soil but also to be drought-resistant. Many of the adaptations which have evolved to combat the former stress are advantageous in surviving the latter (and vice versa). *See* MANGROVE.

Fig. 4. *Suaeda maritima*, the annual sea blite, a halophytic member of the Chenopodiaceae growing at a coastal site in northwest Spain.

Crops. Crops, like other plants, differ in the degree to which they will tolerate salt in the soil. Few, if any, reach the level of tolerance seen in halophytes. The growth of most crops is thus reduced significantly by the presence of salt in the soil, leading to serious reductions in yields. How much salt different crops will tolerate is frequently categorized by the effect of salinity on the relative yield, that is, the yield obtained in a saline soil compared with a control, nonsaline soil. Salinity is commonly quantified by the electrical conductivity of a saturated soil extract. One procedure is to determine the conductivity that reduces the yield to half of that in the nonsaline soil. Among the most tolerant crops are barley (*Hordeum vulgare*), cotton (*Gossypium hirsutum*), and sugarbeet (*Beta vulgaris*); the least tolerant are onion (*Allium cepa*), rice (*Oryza sativa*), and beans (*Phaseolus vulgaris*). A more elaborate approach to quantifying tolerance is to determine the maximum salinity (again characterized by the electrical conductivity of a saturated extract of the soil) that has no discernible effect on the yield (the threshold) and the rate of loss of yield at higher salinities.

Within crop species, there is also information about the relative tolerance of individual varieties. For example, although rice is generally rather sensitive to salt, the older variety Pokkali is much more salt-tolerant than modern varieties released by the International Rice Research Institute. The relative tolerance of varieties is often known to farmers or has been discovered by screening programs aimed at evaluating differences. Differential tolerance has rarely been developed by plant breeders, in spite of considerable effort in the case of some crops. It seems likely that breeders have not been successful because they have treated salt tolerance as a single genetic character such as plant height or flower color. Where it exists, salinity tolerance requires a complex of physiological characteristics. For rice, these characteristics include vigorous growth producing a large plant, the ability to exclude sodium (and chloride) ions from the shoots, the ability to maintain a lower salt concentration in younger as opposed to older leaves, and the ability to tolerate accumulated ions. In wheat, the balance between the uptake of sodium (one of the ions dominating saline soils and not required by the plant) and potassium (a chemically similar ion that is an essential plant nutrient) is such a physiological attribute. The regulation of this characteristic appears to be under the control of a gene or genes on the long arm of chromosome 4D, which is provided to the wheat hybrid by the species *Aegilops squarrosa*. However, much less is known about the genetics of most of the characteristics involved in tolerance, which accounts for the lack of success of planned breeding for tolerance within crops.

Adaptations. Many terrestrial halophytes, having small leaves which are often succulent, look somewhat different from glycophytes (**Fig. 4**). In many species the leaves bear glands which secrete salt. These visible differences are the external

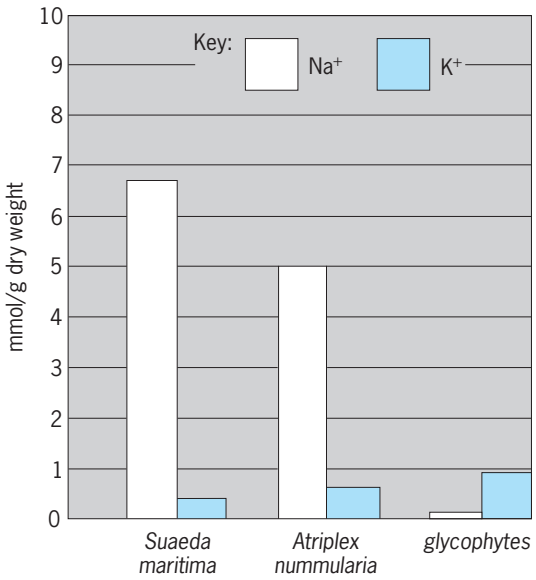

Fig. 5. Sodium and potassium contents in the leaves of two halophyte species growing in solution culture containing about 300 mol m^{-3} Na^+ and 6 mol m^{-3} K^+. Mean data are given for glycophytes growing in the presence of 1 mol m^{-3} Na^+ and 5 mol m^{-3} K^+.

manifestations of more fundamental differences in physiology.

Water potential. Perhaps the basic problem faced by all halophytes is to retain water within the plant in the face of large external osmotic forces. The concept of water potential has proved to be useful in understanding water movement. The water potential is a function of the free energy of water, and it determines the direction in which water moves into and out of the plant; just as for electricity, water flows from high to low potential. Halophytes must be adapted to take up and retain water from much lower potentials than glycophytes can normally cope with. Analysis shows that they achieve this by accumulating salts (**Fig. 5**) and thereby lowering the potential of the water in their cells, so that a positive gradient is maintained between plant and environment. However, this answer itself poses the further problem of living with a high salt content. *See* PLANT-WATER RELATIONS.

Ion content and distribution. The ion content of a halophyte growing in a saline environment may be some five times greater than that of a glycophyte in its natural surroundings. This may mean that the concentration of sodium ions is between 500 and 1000 mol m^{-3} when expressed on the basis of the plant water content. It has been hypothesized that to maintain the cytoplasm at a relatively lower ion content halophytes sequester sodium ions within their vacuoles. Advances in x-ray microanalysis have allowed the measurement of the subcellular distribution of ions within halophyte cells, and these confirm the earlier conclusions that the bulk of the ions which bring about osmotic adjustment are localized within the vacuole, leaving the cytoplasm at an ion concentration which is not very different from that of a glycophyte cell. This leads to the question of how the water potential of the cytoplasm is, in turn, adjusted.

Compatible solutes. Analysis of the leaves of halophytes has shown that they contain rather high concentrations of a number of organic compounds. Of these, perhaps the most common are the amino acid proline and the quaternary ammonium compound glycine betaine. Other important organic solutes found in salt-tolerant higher plants are polyols such as mannitol and sorbitol and the cyclitol, pinitol. Individual species of halophytes contain different compounds or combinations of compounds, and it has been postulated that these organic solutes are responsible for the osmotic adjustment of the cytoplasm.

Terrestrial halophyte. A terrestrial halophyte is a plant which must have a low water potential in order to maintain the water content of its cells. It achieves this by accumulating large quantities of ions, particularly sodium and often chloride, in its vacuoles and adjusting the cytoplasmic water potential by means of a suitable compatible solute. Although ions are freely available in the media in which halophytes grow, compatible solutes must be synthesized and consequently may be some drain on the energy resources of the plant. Restriction of such solutes to the cytoplasm rather than requiring their presence in the whole cell volume clearly lessens the demand, although the energy costs of salt tolerance are still poorly understood.

Ion accumulation must be carefully regulated so that all halophytes restrict, in some way, the entry of ions into the transpiration stream and hence their accumulation in the shoot. Xylem ion concentrations are frequently less then 10% of the external concentration. This accumulation must be carefully matched with the growth rate if the overall ion concentration in the leaf cells is to be maintained approximately constant.

Some halophytes have a secondary means of controlling the ion content of their shoots in that they can remove excess salts through the operation of specialized glands (**Fig. 6**) which are normally situated on the leaves. These glands, which generally have between 2 and about 40 cells, depending upon plant species, contain highly active cells with dense cytoplasmic contents and can excrete ions from the symplast to the exterior of the plant. *See* SECRETORY STRUCTURES (PLANT).

Marine halophyte. In general, marine halophytes appear to possess many of the adaptations seen in their terrestrial counterparts. Of course, they do not have to face the problems of transpiration, but many must be adapted to low light intensities and to the often violent movement of the sea. Many live in estuaries or other regions where the salinity of the water can vary considerably over relatively short periods of time. Such species utilize compatible solutes, often sugar alcohols, to adjust their water potential, and they can both synthesize and break down the compound they use in relation to the external salinity.

Molecular approaches to engineer salt-tolerant glycophytes. The most productive approaches to

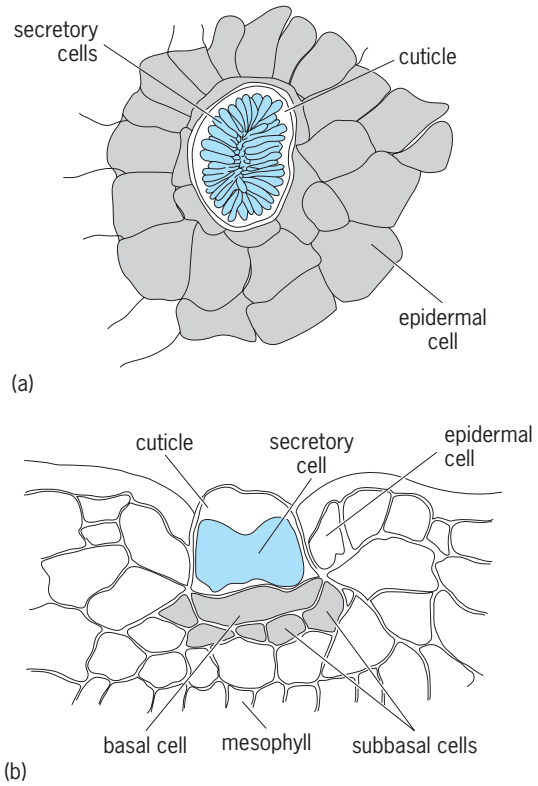

Fig. 6. Salt gland from the mangrove swamp species *Aegiceras corniculatum*. (*a*) Secretory cells surrounded by cuticle and epidermal cells (*after W. P. Anderson, ed., Ion Transport in Plants, Academic Press, 1973*). (*b*) Section through the gland.

identifying genes involved in salt tolerance have used two genetic model systems, namely the baker's yeast, *Saccharomyces cerevisiae*, and the tiny glycophyte weed *Arabidopsis thaliana*. The DNA sequences of the complete genomes for both *Arabidopsis* and yeast have been identified. Scientists also now have the transgenic technologies that allow them to transfer genes between these two organisms. Thus, the identification of relevant genes involved in the adaptive responses of these organisms to salt stress has been dramatically accelerated. Investigations with these two model systems have led to important insights on how to produce salt-tolerant crops.

Sodium toxicity. When a glycophyte plant encounters a high-salinity environment, there is a decrease in potassium (K^+) uptake and a simultaneous increase in sodium (Na^+) influx. High concentrations of sodium become toxic to important metabolic reactions that evolved in an intracellular environment where potassium is the major inorganic cation. Emerging evidence suggests that high intracellular sodium interferes with potassium uptake systems. So in order for a plant to grow in the presence of high sodium, it must deal with the cytotoxic effect of sodium and at the same time restore the required intracellular potassium levels. As mentioned above, some halophytes avoid toxicity by transporting sodium into their vacuoles and sequestering it there. This detoxification process has received a great deal of attention. An alternate, but not mutually exclusive, cellular strategy to avoid sodium toxicity is to increase sodium efflux systems or to reduce cellular sodium uptake.

Sodium sequestration. Plant physiologists described the mechanisms involved in the vacuolar Na^+ sequestration process more than 20 years ago. It was shown that the coordinated activity of a proton (H^+) pump and a proton/sodium (H^+/Na^+) exchanger is required to achieve the Na^+ sequestration into the vacuole. The role of the vacuolar H^+ pump is to accumulate H^+ into the vacuole; H^+ is then exchanged for Na^+ ions via a revolving-door-like protein, the Na^+/H^+ exchanger. There are two H^+ pumps at the vacuolar membrane involved in the generation of this H^+ gradient, one dependent on ATP (H^+-ATPase) and the other dependent on inorganic pyrophosphate (H^+-PPase). Inorganic pyrophosphate, or PPi, is produced as a by-product of a wide range of biosynthetic pathways. Its utilization as an energy donor for the activity of the vacuolar H^+-PPase allows ATP to be conserved and improves plant cell performance under a more demanding environment.

H^+/Na^+ exchanger. After the physiological characterization of the vacuolar detoxification mechanism, the search for the gene encoding the H^+/Na^+ exchanger began, but the initial search failed. A critical discovery was then made using yeast, which contributed to the final isolation of a plant gene encoding for an Na^+/H^+ exchanger. A gene involved in Na^+ sequestration into the yeast pre-vacuolar compartment was cloned and characterized. The sequence of this yeast Na^+/H^+ exchanger (*NHX1*) was then compared to the *Arabidopsis* genomic database, and the elusive plant Na^+/H^+ exchanger (*AtNHX1*) was finally discovered. Furthermore, the overexpression of this plant Na^+/H^+ exchanger was found to enhance salt tolerance in *Arabidopsis*, tomato, and canola. In each case, the plants accumulated more Na^+ in their vacuoles and actually prospered in high levels of salt. Na^+ accumulation occurred mainly in the green parts of the plant, but not in the fruit. These results imply that heightened activity of this Na^+/H^+ exchanger could enhance crop productivity.

H^+ pumps. An attractive alternative to improve vacuolar Na^+ sequestration would be to enhance the availability of H^+ in the vacuole. Salt-tolerant plants such as the ice plant *Mesembryanthemum crystallinum* naturally utilize high-level expression of both the vacuolar ATPase and the Na^+/H^+ exchanger to tolerate high-salt conditions. However, overexpression of the vacuolar ATPase is difficult mainly because this H^+ pump is composed of subunits encoded by at least 26 genes. Fortunately, the other vacuole H^+ pump, H^+-PPase, is a single subunit protein. Here again, yeast proved to be instrumental, as expression of a gain-of-function allele of the *Arabidopsis* vacuolar H^+-PPase gene (*AVP1-D*) suppressed the salt sensitivity of a yeast mutant lacking the plasma membrane sodium efflux pump. Transgenic *Arabidopsis* plants engineered to overexpress *AVP1* exhibited increased tolerance to salt, and were capable

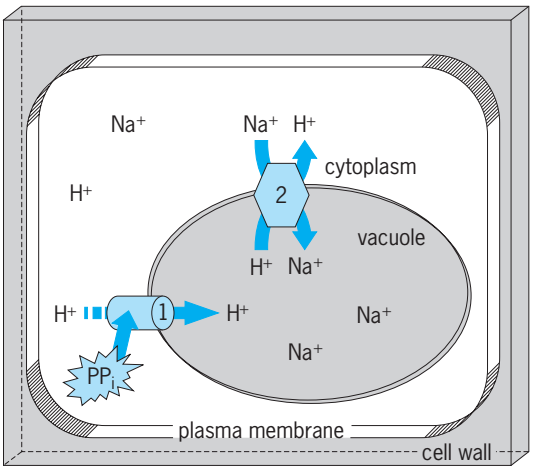

Fig. 7. ***Arabidopsis*** **vacuolar H^+-PPase (*AVP1*), an H^+ pump that utilizes inorganic pyrophosphate (PPi) as an energy source, provides increased H^+ to the vacuole (1). An H^+/Na^+ exchanger such as *AtNHX1* transports the H^+ out of the vacuole in exchange for Na^+ (2). By this mechanism Na^+ is removed from the cytoplasm and sequestered in the vacuole, and its cytotoxic effects are prevented.**

of flowering and producing viable seeds in the presence of 0.25 *M* NaCl.

Mechanism. The enhanced salt tolerance of *AVP1* transgenic plants is most easily explained by an enhanced proton-driven uptake of solutes into vacuoles. Presumably, greater AVP1 activity in vacuolar membranes provides increased H^+ to drive the uptake of cations via H^+/cation exchangers into the vacuole (**Fig. 7**).

Cellular sodium uptake. Attempts to further improve salt tolerance in plants should consider the importance of the cellular Na^+ uptake mechanisms.

SOS1 Na^+/H^+ exchanger. Genetic studies with *Arabidopsis* uncovered an Na^+-driven signal transduction pathway involved in sodium uptake by cells. The main interactions are a putative plasma membrane Na^+/H^+ exchanger, SOS1; a serine/threonine protein kinase, SOS2; and a Ca^{2+}-binding protein, SOS3. An elegant study in yeast demonstrated that the SOS1 Na^+/H^+ exchanger, the SOS2 protein kinase, and its associated Ca^{2+} sensor, SOS3, form a functional module involved in the control of cellular Na^+ homeostasis. SOS3's role is to shuttle SOS2 to the plasma membrane, where the SOS2/SOS3 complex activates the SOS1 exchanger. Interestingly, SOS1 expression has been localized in *Arabidopsis* to the epidermal cells at the root tip. This cell layer protects the root apical meristem, which consists of actively dividing cells lacking large vacuoles. Furthermore, the vacuolar Na^+/H^+ exchanger is not expressed at the root tip. So it appears that an important role of SOS1 is to prevent Na^+ uptake at the root tip level. SOS1 has also been localized to the cells surrounding the stele (the central vascular cylinder of roots and stems), suggesting that it is implicated in the Na^+ load of the vascular system.

Other systems. Another group of proteins involved in Na^+ uptake are the channel proteins (membrane proteins that catalyze the net movement of millions of ions per minute following their electrochemical potential). There is also a body of evidence suggesting that nonselective cation channels (NSCCs) play a major role in Na^+ influx into root cells. Interestingly, these NSCCs are inhibited by calcium, a behavior consistent with earlier observations that addition of calcium reduces the stringency of a sodium stress in plants.

Future directions. Recent research promoting salt tolerance in plants focuses on boosting the intracellular salt-sequestering mechanisms by transferring the gene for this phenotype from the model organism, *A. thaliana*, or from an halophyte plant to selected crop species. A complementary approach focuses on the difficult task of reducing the net input of salt into the plants via interfering with the function of channels and transporters involved in sodium uptake without disturbing potassium uptake. An ideal scenario contemplates the generation of transgenic plants with an enhanced vacuolar salt sequestration capability and a reduced salt uptake. One remaining concern is that if low-quality (high-salt) waters are used to irrigate this new generation of salt-tolerant crops, history will surely repeat itself. This practice will eventually result in soils with enough salt to outstrip the protective mechanisms described above. Therefore, research must also focus on alternative irrigation practices that optimize the use of water, and on treatments to optimize its quality. *See* IRRIGATION (AGRICULTURE); PLANT CELL; PLANT-WATER RELATIONS; VACUOLE.

Roberto A. Gaxiola; T. J. Flowers

Bibliography. P. Adam, *Saltmarsh Ecology*, 1990; J. H. Cherry (ed.), *Environmental Stress in Plants: Biochemical and Physiological Mechanisms*, 1989; T. J. Flowers, M. A. Hajibagheri, and N. J. W. Clipson, Halophytes, *Quart. Rev. Biol.*, 61:313–336, 1986; P. M. Hasegawa et al., Plant cellular and molecular responses to high salinity, in H. E. Gould (ed.), *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, pp. 463–499, 2000; H. G. Jones, T. J. Flowers, and M. B. Jones (eds.), *Plants under Stress*, 1989; E. V. Maas, Salt tolerance of plants, *Appl. Agr. Res.*, 1:12–26, 1986; D. Pasternak, Salt tolerance and crop production: A comprehensive approach, *Annu. Rev. Phytopathol.*, 25:271–291, 1987; R. Serrano and R. Gaxiola, Microbial models and salt-stress tolerance in plants, *Crit. Rev. Plant Sci.*, 13:121–138, 1994.

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

14 PLAS-QUI

www.MHEST.com

P

Plasma (physics) — Pythagorean theorem

Plasma (physics)

A gas of charged and neutral particles which exhibits collective behavior. All gases become ionized at sufficiently high temperatures, creating what has been called a fourth state of matter, together with solids, liquids, and gases. It has been estimated that more than 99% of the universe is in the plasma state. On the Earth, plasmas are much less common. Lightning is a familiar natural manifestation, and fluorescent lights are a practical application. Plasma applications and studies make use of an enormous range of plasma temperatures, densities, and neutral pressures (**Fig. 1**). They extend from plasma processing applications at relatively low temperatures (such as plasma etching of semiconductor chips at low pressure, or plasma cutting torches at atmospheric pressure) to studies of controlled fusion at very high temperatures. *See* FLUORESCENT LAMP.

Plasma physics is a many-body problem that can be described by a combination of Newton's laws and Maxwell's equations. The charged particles in plasmas are usually ions, both positive and negative, and electrons. Plasmas are normally quasineutral; that is, the net positive ion charge density approximately equals the net negative charge density everywhere in the bulk of the plasma. Quasi-neutrality refers to charge density and does not imply equal densities of electrons and ions since ions can be multiply charged and can also have negative charge. In space and fusion plasmas, plasmas are normally magnetized, while in application plasmas on Earth, such as plasma processing, both magnetized and unmagnetized plasmas are employed. *See* MAXWELL'S EQUATIONS; NEWTON'S LAWS OF MOTION.

Plasma parameters. It is convenient to keep track of plasma properties in terms of characteristic lengths, frequencies, and velocities. Among these are the Debye length, the electron and ion plasma frequencies, the electron and ion gyrofrequencies and gyroradii, the electron and ion thermal velocities, the

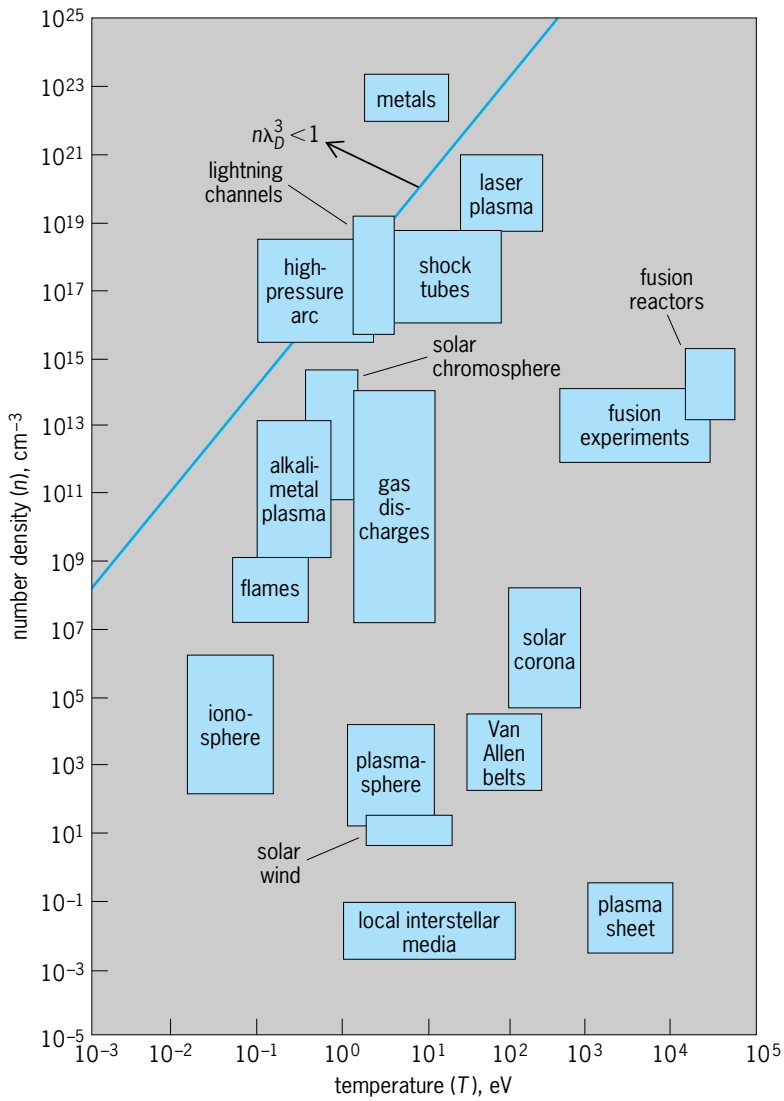

Fig. 1. Approximate number densities (n) and temperatures (T) of some typical plasmas. The temperature is measured in electronvolts (eV), with 1 eV = 11,605 K. The region above and to the left of the diagonal line represents strongly coupled plasmas ($n\lambda_D^3 < 1$, where λ_D is the Debye length). The figure uses logarithmic scales.

ion sound velocity, the Alfvén velocity, and various collision lengths. The definition of a plasma depends on several of these characteristic parameters, and the magnitude of ratios of these parameters to system size or applied frequencies determines most plasma behavior.

The simplest plasma is a collisionless, unmagnetized collection of ions and electrons with no significant currents. Such plasmas have quasineutral regions and nonneutral regions. The nonneutral regions are highly localized. They are usually located near boundaries (where they are known as sheaths), but are sometimes located within the plasma (where they are known as double layers).

The plasma potential, ϕ, is described by Poisson's equation. In a one-dimensional system, this equation reduces to Eq. (1), where ρ is the net charge density and ε_0 is the permittivity of free space. From Poisson's equation, it can be shown that the spatial dependence of the plasma potential in nonneutral regions scales with the Debye length, given by Eq. (2), regardless of the geometry of a particular situation (see **table**). Here n is the value of the electron density in the quasineutral region, e is the electron charge, and T_e is the electron temperature, measured in electronvolts (1 eV corresponds to 11,605 K, and Boltzmann's constant $k_B = 1$). The dimensions of nonneutral regions are the order of the Debye length when $e\delta\phi$ is of the order of the electron temperature (where $\delta\phi$ is a characteristic variation of the potential), and depend on the Debye length when $e\delta\phi$ is much greater than the electron temperature. In quasineutral regions where the charge density is approximately 0, the potential is normally constant or, at most, slowly varying with a constant electric field over distances comparable to the characteristic dimension, L, of a system. In terms of the Debye length, quasineutral means that electric fields must be much weaker than $T_e/(e\lambda D)$. *See* POTENTIALS.

$$\frac{d^2\phi}{dx^2} = -\frac{\rho}{\varepsilon_0} \tag{1}$$

$$\lambda_D = \sqrt{\frac{\varepsilon_0 T_e}{ne^2}} \tag{2}$$

Convenient expressions for evaluating plasma parameters

Parameter	Symbol	Value*
Debye length	λ_D	$7.4 \times 10^3 \sqrt{\frac{T_e}{n}}$
Electron plasma frequency	$f_{pe} \equiv \frac{\omega_{pe}}{2\pi}$	$9.0\sqrt{n}$
Electron cyclotron frequency	$f_{ce} \equiv \frac{\Omega_e}{2\pi}$	$2.8 \times 10^{10}\, B$
Electron thermal velocity	V_e	$4.2 \times 10^5 \sqrt{T_e}$
Ion sound velocity	C_s	$10^4 \sqrt{\frac{T_e}{\mu}}$
Alfvén velocity	V_A	$2.2 \times 10^{16} \frac{B}{\sqrt{n_\mu}}$

* Lengths are expressed in meters (m); frequencies are in hertz (Hz); velocities are in m/s; electron temperatures (T_e) are in electronvolts (eV); particle densities (n) are in m^{-3}; masses (μ) are in atomic mass units (u); and magnetic fields (B) are in tesla.

Plasma phenomena begin to be apparent only when the system's characteristic dimension is greater than the Debye length. Further, for a gas to be a plasma, there should be at least one electron within each Debye-length volume, λ_D^3.

The electron thermal velocity is given by Eq. (3),

$$v_e = \sqrt{\frac{T_e}{m_e}} \tag{3}$$

where m_e is the electron mass. Dividing this velocity by the Debye length yields the (electron) plasma frequency, given by Eq. (4), which is the characteristic frequency of electrons in unmagnetized plasmas.

$$\omega_{pe} = \frac{v_e}{\lambda_D} = \sqrt{\frac{ne^2}{\varepsilon_0 m_e}} \tag{4}$$

Similarly, the characteristic frequency of ions, called the ion plasma frequency, is given by Eq. (5), where m_i is the ion mass. Multiplying the ion plasma frequency by the Debye length gives a characteristic ion velocity, called the ion sound velocity and denoted c_s, which depends on the electron temperature rather than the ion temperature.

$$\omega_{pi} = \sqrt{\frac{ne^2}{\varepsilon_0 m_i}} \tag{5}$$

In the presence of collisions, it is still meaningful to consider the system to be a plasma so long as the collision frequency, ν, defined as the reciprocal of the mean collision time, is smaller than the plasma frequency. This idea can also be expressed by the statement that the mean free path for collisions is longer than the Debye length. If the neutral pressure is too high, electron and ion neutral collisions can dominate the properties of the charged gas.

In magnetized plasmas (with magnetic field B), electrons and ions gyrate with characteristic frequencies (cyclotron frequencies) given by Eqs. (6).

$$\Omega_e \equiv \frac{eB}{m_e} \qquad \Omega_i \equiv \frac{q_i B}{m_i} \tag{6}$$

Multiplying the ratio Ω_i/ω_{pi} by the speed of light yields another important characteristic velocity, the Alfvén velocity, denoted v_A.

The three criteria for the existence of a plasma are thus given by inequalities (7). While the charged particles in a typical room ($n = 10^6/\mathrm{m}^3$, $T_e = 0.025$ eV) satisfy inequalities (7*a*) and (7*b*), they are not a plasma because they do not satisfy inequality (7*c*).

$$L > \lambda_D \tag{7a}$$

$$n\lambda D^3 > 1 \tag{7b}$$

$$\nu < \omega_{pe} \tag{7c}$$

Collective behavior refers to the plasma properties not present in single-particle motion. Collective behavior is a distinguishing characteristic of a plasma.

It consists of flows, waves, instabilities, and so forth. Common examples are fluctuations in the aurora, generation of microwaves in devices such as magnetrons and klystrons, and reflection of electromagnetic waves from the ionosphere. *See* AURORA; KLYSTRON; MAGNETRON; RADIO-WAVE PROPAGATION.

Charged-particle gases. Curiously, very high density collections of equal numbers of ions and electrons are not plasmas. According to Eq. (2), the product $n\lambda_D^3$ is proportional to $1/\sqrt{n}$, so at high plasma density this product is less than 1, and inequality (7*b*) is no longer satisfied. Such systems are referred to as strongly coupled plasmas (even though, strictly speaking, they are not plasmas at all).

A collection of either electrons or ions can exhibit properties similar to those of an electrically neutral plasma if the charged-particle density is sufficiently large. For such so-called plasmas, the Debye length and the characteristic frequency of electrons or ions, ω_{pe} or ω_{pi}, can still be defined, and collective behavior is still exhibited when the Debye length is less than the system's characteristic dimension. So-called pure electron plasmas or pure ion plasmas are unconfined in an unmagnetized system. However, particle traps consisting of a combination of electric and magnetic fields can be used to confine the charges. *See* PARTICLE TRAP.

Plasmas based on carbon- or silicon-containing gases can give rise to polymer formation within the plasmas. As a consequence, macroscopic insulating particles can grow within the plasma, reaching sizes the order of micrometers. The insulating particles charge negative, with thousands of electron charges or more, so near the particles the plasma can consist of many positive ions and very few electrons. Unlike other plasmas, dusty plasmas consist of negative ions with a large variety of masses and charge-to-mass ratios.

Visual characteristics. The visual appearance of a plasma depends on the kind of ion present, the electron temperature, and the plasma density. Some plasmas are invisible. Curiously, if a plasma is present and not glowing, it is either very hot or very cold. For example, an H^+ plasma, or any other relatively hot plasma with fully stripped ions, contains atomic nuclei with no electrons, so there is no atomic physics and no optical emission or absorption. If plasma electrons and ions are very cold, there is insufficient energy to excite optical transitions. The glow often associated with plasmas indicates only where visible energy transitions are excited by energetic electrons or perhaps absorption of ultraviolet radiation, and may have little to do with the presence of bulk plasma. In fusion plasmas, the edges are often copious sources of emission associated with the dissociation and ionization of hydrogen and edge-generated impurities, while much of the hotter core plasma is fully ionized and invisible.

Direct-current glow-discharge plasmas originate from electrons created by secondary electron emission due to ion bombardment of a negatively biased cathode. The secondary electrons are accelerated through the cathode sheath potential (called the cathode fall) to energies the order of 1 keV, and partially ionize the neutral gas, releasing additional energetic electrons in a multiplicative process. The energetic electrons also undergo inelastic collisions with neutrals which result in optical emission that contributes to the so-called glow. *See* GLOW DISCHARGE; SECONDARY EMISSION.

Single-particle motion. The understanding of plasma physics begins with an understanding of the motion of single charged particles in a combination of electric and magnetic fields (*E* and *B*), produced by a combination of external fields and the motion of the charged particles themselves. The motion of a single particle, with mass m, charge q, and velocity $\mathbf{v}$, is governed by the Lorentz force, as given in Eq. (8).

$$m\frac{d\mathbf{v}}{dt} = q\,(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{8}$$

From the perpendicular component of Eq. (8), it can be shown that the charged particles gyrate about magnetic field lines with a characteristic frequency (the cyclotron frequency) given by Eqs. (6). Ions rotate about the magnetic field in the clockwise direction, while electrons rotate counterclockwise with the magnetic field pointing outward. *See* ELECTRIC FIELD; PARTICLE ACCELERATOR.

In addition to the motion parallel to the magnetic field and the gyromotion about the magnetic field, there are drifts perpendicular to the magnetic field. For a general force, $\mathbf{F}$, in the presence of a magnetic field, the perpendicular drift velocity is given by Eq. (9).

$$\mathbf{v}_D = \frac{\mathbf{F} \times \mathbf{B}}{qB^2} \tag{9}$$

Two specific examples of perpendicular drifts are the $\mathbf{E} \times \mathbf{B}$ and curvature drifts. Substituting $\mathbf{F} = q\mathbf{E}$ into Eq. (9) gives Eq. (10). This $\mathbf{E} \times \mathbf{B}$ drift is orthogonal to both the magnetic field and the electric field.

$$\mathbf{v}_D = \frac{\mathbf{E} \times \mathbf{B}}{B^2} \tag{10}$$

Curiously, it is independent of the particle's mass and charge. Given a perpendicular electric field, particles can walk across a magnetic field. Forces associated with magnetic-field curvature give rise to a curvature drift in the direction orthogonal to the magnetic field, and to the radius of curvature of the magnetic field lines.

Magnetic mirroring. For gyro motion in a slowly changing magnetic field, which is approximately periodic, it can be shown that the ratio of the perpendicular energy to the magnetic field is approximately constant. This means that a charged particle moving parallel to a magnetic field and gyrating about the field will gyrate faster as the magnetic field increases. If the magnetic field changes in space and is constant in time, the total energy is conserved. For a sufficiently large magnetic field, a point is reached where the total energy equals the perpendicular energy, so that the parallel energy goes to zero and the

particle reflects. This is known as magnetic mirroring.

Magnetic mirroring is the chief natural mechanism of charged-particle confinement. For example, this process confines charged particles in the ionosphere and magnetosphere. The magnetic field lines that connect the north and south magnetic poles of the Earth provide a mirror magnetic field which increases as either pole is approached. In the absence of collisions, a particle moving along and gyrating about such a magnetic field is magnetically confined, if it has a sufficiently large velocity perpendicular to the magnetic field. The Van Allen belts are composed of such mirror-trapped charged particles. The source of these particles is the solar wind, a stream of charged particles continuously emitted by the Sun.

Distribution functions. The number of particles per unit volume at position $\mathbf{x}$ and time t between velocities $\mathbf{v}$ and $\mathbf{v} + d\mathbf{v}$ can be specified by the distribution function $f(\mathbf{v},\mathbf{x},t)\, d\mathbf{v}$. This distribution function satisfies the Boltzmann transport equation, which expresses the rate of change of f as the sum of a term due to external forces, a term due to diffusion, and a term due to collisions of the particles. If the collision term is set equal to zero, the Vlasov equation results. *See* BOLTZMANN TRANSPORT EQUATION.

Under many situations the distribution functions of plasma electrons and ions are approximately maxwellian, which means that each distribution function that varies with $\mathbf{v}$ is proportion to the exponential of $-E/T$, where E is the kinetic energy of the particle in question. For maxwellian distribution functions, the average energy is equal to $^3/_2T$. *See* BOLTZMANN STATISTICS; KINETIC THEORY OF MATTER.

Fluid equations. It is often convenient to describe the plasma in terms of fluid variables such as the plasma species densities $n_j(\mathbf{x},t)$, the fluid velocity $\mathbf{u}(\mathbf{x},t)$, and the pressure tensor $\boldsymbol{P}(\mathbf{x},t)$. This procedure, together with Maxwell's equations, provides the simplest description of a plasma. The simplest fluid equations are the continuity equation and the equation of motion. The fluid equation of motion resembles the single-particle equation of motion with the addition of terms representing a pressure-gradient force on a single particle and a frictional force. *See* FLUID-FLOW PRINCIPLES.

For given electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$, determined by Maxwell's equations, the equations of continuity and motion are two equations with three unknowns, n, $\mathbf{u}$, and $\mathbf{P}$. Solutions to the fluid equations can be obtained by giving another equation for the extra unknown, an equation of state (11), where

$$p = C(nm)^{\gamma} \tag{11}$$

p is the pressure and C is a constant. For isothermal compression $\gamma = 1$, while for adiabatic compression γ is the ratio of specific heats, equal to $(2 + N)/N$, where N is the number of degrees of freedom. Even in this simple description, the equations are nonlinear. *See* POLYTROPIC PROCESS.

Nonlinear physics. Many of the equations describing plasma physics are nonlinear. The usual approach is to linearize the equations. This procedure simply ignores all nonlinear terms. However, solutions exist that are inherently nonlinear and not superpositions of linear solutions.

Sheaths are a simple example of a nonlinear structure. Ion acoustic solitons are another example of a nonlinear structure. Ion acoustic solitons are pulse-like modes which result from a balance of nonlinearity, which tends to steepen pulses (since large-amplitude components travel faster), and dispersion, which tends to broaden pulses. The velocity of solitons, measured in the frame moving at the ion acoustic velocity, is proportional to the soliton amplitude, h, while the soliton width, w, is proportional to $1/\sqrt{h}$, so the product hw^2 is a constant of motion. These solitons can survive collisions with other solitons with no change in shape. In this respect, they behave like linear pulses even though they are inherently nonlinear. *See* NONLINEAR PHYSICS; SOLITON.

MHD equations. For fully ionized plasmas, it is convenient to describe the plasma as a single fluid together with Maxwell's equations. This gives the magnetohydrodynamic (MHD) equations, which are used to describe plasma equilibria and plasma waves and instabilities. Their relative simplicity has made them ideal for solutions of fusion problems in complicated geometries, and they have been widely used to describe astrophysical plasmas and magnetohydrodynamic energy conversion. *See* MAGNETOHYDRODYNAMIC POWER GENERATOR; MAGNETOHYDRODYNAMICS.

Waves in plasmas. Plasmas can support an impressive variety of electrostatic and electromagnetic waves not present in the absence of plasma. The waves are distinguished by their frequency, the presence or absence of dc magnetic fields, and the plasma temperature and density. Wave characteristics also depend on whether the plasma is bounded, and if it is, how it is bounded. Relatively simple expressions exist only for unbounded (that is, infinite), unmagnetized plasmas or cold plasmas with zero temperature. Wave characteristics depend on the ratio of the wave frequency to the characteristic frequencies of the plasma, ω_{pi}, ω_{pe}, Ω_i, and Ω_e.

Unmagnetized plasmas. In unbounded, unmagnetized plasmas with finite electron temperature and zero ion temperature, only two modes can exist: ion acoustic waves with frequencies less than the ion-plasma frequency, ω_{pi}, and Langmuir waves with frequencies greater than the electron plasma frequency, ω_{pe}. Both modes are electrostatic with the fluctuating electric field parallel to the wave propagation direction. No waves can exist between the ion and electron plasma frequencies.

Alfvén waves. In the presence of a magnetic field, and at frequencies less than the ion cyclotron frequency, Ω_i, the solutions to the wave equation are known as Alfvén waves. They come in three varieties: magnetosonic, shear Alfvén, and slow Alfvén waves (ion acoustic waves).

For low frequencies (much less than Ω_i) the magnetosonic and shear Alfvén waves have similar values of frequency versus wave number. As the frequency is increased, the solutions to the wave equation split into two separate branches. Near Ω_i, one branch is known as the ion cyclotron waves, while the other continues to be known as the magnetosonic waves, or fast waves. *See* ALFVÉN WAVES.

Whistlers. At frequencies between $\sqrt{\Omega_i \Omega_e}$ and the electron cyclotron frequency, Ω_e, the electromagnetic wave solutions are known as whistlers (**Fig. 2**). Whistler waves are observed to propagate along the magnetic field lines connecting the Earth's magnetic poles. They are at audio frequencies and make a whistlinglike sound when the signals they generate are played over a loudspeaker. This phenomenon occurs because the phase and group velocities differ, with the higher-frequency components arriving first.

Bernstein waves. For frequencies greater than Ω_i, electrostatic waves propagating across the magnetic field are found to be associated with the harmonics of Ω_i and are known as ion Bernstein waves. These waves can exist over a range of frequencies between

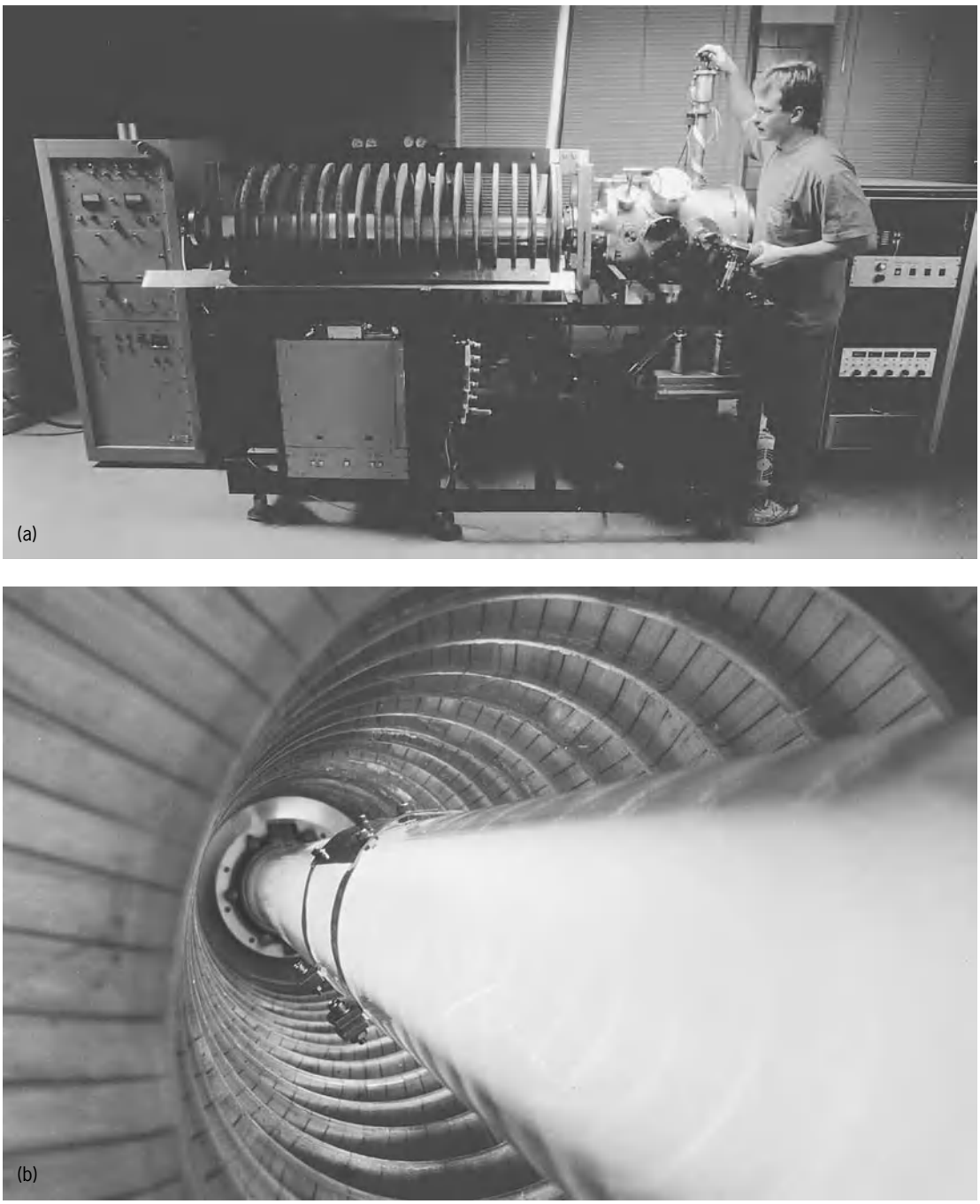

Fig. 2. Helicon plasma, which is generated by whistler waves, which then propagate in the plasma. (*a*) Plasma source. (*b*) Plasma in a Pyrex chamber mounted inside the magnetic field coils. (*Bruce Fritz, University of Wisconsin Engineering Center for Plasma Aided Physics*)

successive harmonics of Ω_i with new modes appearing at each harmonic. Electrostatic waves known as electron Bernstein waves are also found to be associated with the harmonics of Ω_e.

Collisionless Landau damping. It might be guessed that waves in a collisionless plasma would not damp, since there is no energy dissipation. This does not turn out to be true. Starting with the Vlasov equation, L. D. Landau showed that langmuir waves damp when the derivative of the distribution function evaluated at the phase velocity of the wave is negative. This condition is satisfied for the common case of a maxwellian electron distribution.

Plasma sources. Much of the progress in the understanding of plasma physics has been the result of the development of new plasma sources. Discharge tubes, Penning discharges, Q machines, multidipole confined plasmas, magnetron discharges, stellarators, and tokamaks are examples. The key elements shared by all plasma sources are ionization, confinement, equilibrium, stability, and heating.

Ionization. Ionization is the key to plasma production and can be accomplished in many different ways. The most common approach is to employ energetic electrons with energies greater than the ionization potential of the gas being ionized. In dc glow discharges, electrons produced by ion secondary electron emission are accelerated by the cathode sheath potential, as are electrons created by thermionic emission in hot-cathode plasmas. Electrons can also pick up energy by reflecting from oscillating radiofrequency sheath electric fields, or by cyclotron resonance in magnetic fields, or from collisions with other energetic electrons. *See* ELECTRICAL CONDUCTION IN GASES; GAS DISCHARGE; IONIZATION POTENTIAL; THERMIONIC EMISSION.

Several other approaches involving collisions, which do not require energetic electrons, also exist. These techniques include photoionization, ion-neutral charge exchange, surface ionization, and Penning ionization. Ions can also be produced in the dissociation of molecules. Yet another mechanism, called critical ionization velocity, is instability driven, and occurs when the kinetic energy of the neutral gas atoms streaming perpendicular to a magnetic field exceeds their ionization potential. *See* ION SOURCES; IONIZATION; PHOTOIONIZATION.

Confinement. A vacuum chamber provides the simplest approach to confinement. In an unmagnetized plasma, electrons are lost more rapidly than ions, and the plasma acquires a net positive charge. The excess positive charge appears in a sheath at the plasma boundary with the bulk plasma potential more positive than the boundary potential. The decrease in potential at the boundary provides plasma electron confinement, reducing their loss rate to balance the ion loss rate. The ion loss current is found to lead to the particle confinement time given by Eq. (12),

$$\tau = \frac{V}{A}\sqrt{\frac{m_i}{T_e + T_i}} \tag{12}$$

where A is the loss area and V is the plasma volume. For tabletop-size devices used in plasma processing, τ is of the order of 0.1 millisecond.

Addition of a uniform magnetic field reduces the loss rate of ions and electrons transverse to the magnetic field, but has no effect on losses parallel to the magnetic field because the Lorentz force has no components along this field. Effective confinement by magnetic fields requires that the ion and electron gyroradii be small compared to device dimensions. Plasma transport across the magnetic field can still occur as a result of collisions or of perpendicular drifts.

Open magnetic-field confinement devices, in which the magnetic field intersects a boundary, can take advantage of nonuniform magnetic fields to enhance axial confinement by providing magnetic mirrors. However, charged particles which do not have sufficiently large perpendicular velocities fall within a loss cone, in velocity space, and are not confined. Scattering in velocity space removes particles from the confined region.

In closed magnetic confinement devices, magnetic fields do not intersect a boundary, but charged particles can still flow across magnetic fields. A toroidal magnetic field can be produced by a set of toroidal field coils (**Fig. 3**). Perpendicular

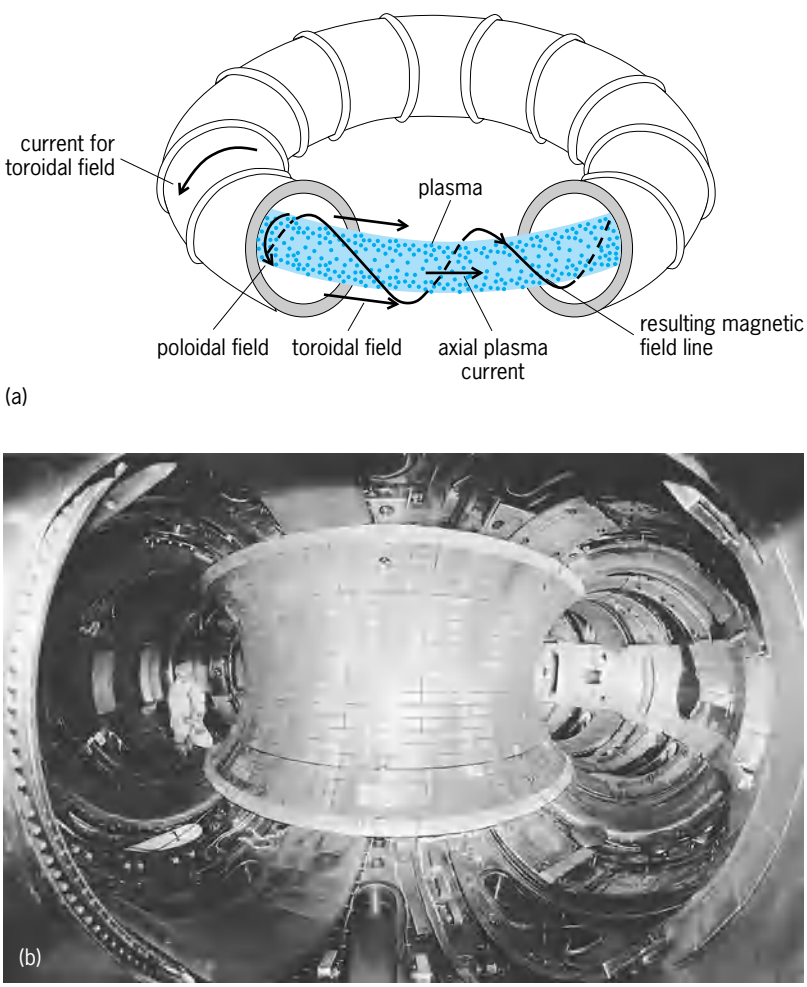

Fig. 3. Tokamaks. (*a*) Diagram of tokamak plasma and magnetic field. (*b*) Interior of the reaction chamber of the Tokamak Fusion Test Reactor (TFTR).

plasma drifts cause a vertical separation of electrons and ions and a corresponding vertical electric field. Combined with the toroidal magnetic field, this results in an outward $\mathbf{E} \times \mathbf{B}$ drift. Toroidal plasma confinement devices avoid this problem by twisting the magnetic field helically by adding a poloidal magnetic field. The $\mathbf{E} \times \mathbf{B}$ drift (which is always outward with respect to the major radius of the torus) changes from outward to inward (with respect to the minor radius of the torus) as the magnetic line of force goes from the low magnetic field, outside the torus, to the high field inside the torus and averages to zero. In stellarators, the twist in the magnetic field is accomplished by using external conductors. In tokamaks (Fig. 3), it is produced by current in the plasma itself.

Equilibrium. In the absence of magnetic fields (both inside and outside the plasma), an equilibrium can be achieved by establishing a pressure balance between plasma and edge walls or edge gas.

When a plasma column is confined by a magnetic field, it can be shown that Eq. (13) is satisfied, where

$$p_\perp + \frac{B^2}{2\mu_0} = \frac{B_0^2}{2\mu_0} \tag{13}$$

$p_\perp$ is the plasma pressure perpendicular to the magnetic field, B_0 is the magnetic field just outside the plasma, and μ_0 is the permeability of free space. Diamagnetic currents generated at the plasma edge reduce the magnetic field in the plasma. The quantity $B^2/(2\mu_0)$ can be regarded as the magnetic-field pressure. Its ratio to the plasma perpendicular pressure is called the plasma beta, given by Eq. (14). Equilibrium

$$\beta \equiv p_\perp \left(\frac{B_0^2}{2\mu_0} \right) \tag{14}$$

exists only when $\beta \leq 1$, and $\beta = 1$ corresponds to zero magnetic field inside the plasma. *See* DIAMAGNETISM.

Stability. The existence of an equilibrium does not guarantee that a particular configuration is stable. There are two classes of instabilities, macroscopic and microscopic. Macroscopic or magnetohydrodynamic instabilities are of relatively low frequency and involve large-scale motions of the plasma, such as ripples in the plasma edge. The energy for the instability comes from the plasma expanding in a magnetic field. Microinstabilities are of relatively high frequency and get their free energy from changes in the distribution function $f(\mathbf{v},\mathbf{x},t)$ toward a uniform maxwellian distribution function, the distribution function with the lowest free energy.

An important example of a magnetohydrodynamic instability is the Rayleigh-Taylor instability. The equilibrium between a uniform plasma column and a uniform magnetic field is not stable in the presence of a transverse force applied away from the plasma. The growth rate of the instability is proportional to the square root of the effective gravity associated with the transverse force. The gravitational force can be the result of curved magnetic field lines. The gradient of the magnetic field points away from the center of curvature of curved magnetic field lines, so there can be good or bad curvature. Plasmas are stable when this gradient points toward the plasma, and unstable when it points away from the plasma.

A similar instability results from an attempt to support a heavy fluid by a light fluid. When a glass of water is turned upside down, the upward air pressure on the water exceeds the downward force of the water so that an equilibrium exists, but the water falls out of the glass never the less. It is apparent that the equilibrium of a heavy fluid (water) supported by a light fluid (air) is not stable.

Plasma heating. Combined with collisions, dc electric fields can heat electrons. The electrons are accelerated by the electric field, and the energy is randomized in direction by collisions with other electrons, ions, or neutrals. Resonant heating can also be employed. In the presence of a circularly polarized electric field rotating with electrons at the electron cyclotron frequency, electrons can be spun up to high energies without the need for collisions. A similar approach can be used to heat ions. A third approach to heating plasmas is to use charge exchange on energetic neutral beams which converts energetic neutrals to energetic ions and cold ions to cold neutrals.

Ohmic heating, the combination of an applied electric field and Coulomb collisions, is a very successful heating technique in fusion plasmas. The resistivity averaged over a maxwellian electron distribution function is known as the Spitzer resistivity. For an electron temperature of 1 keV, this resistivity is 2.5×10^{-8} ohm-m, comparable to the resistivity of copper (2×10^{-8} ohm-m).

Ohmic heating has provided tokamaks with an easy way to heat plasmas and an easy path to relatively high density and high temperature, which may progress all the way to fusion reactors. However, the resistivity is proportional to the $-3/2$ power of the electron temperature, and thus becomes much less effective at high temperatures.

Plasma processing. Plasma processing can be defined as the collection of techniques which make use of plasmas to create new materials or to modify properties of existing materials. It is used in a large variety of applications including semiconductor etching, preparing plastic surfaces to accept ink, depositing polymers, depositing diamond films, and hardening artificial hip joints. The technique has its foundations in plasma physics, chemistry, electrical and chemical engineering, and materials science.

Plasma processing employs a wide range of plasma densities and neutral pressures, but a narrow range of electron temperatures, from 1 to 10 eV. At the lowest neutral pressures, less than 10^{-1} pascal (10^{-3} torr), plasma effects are most important and plasmas can generally be treated as collisionless. The most interesting application in this neutral pressure regime is ion implantation into objects immersed in plasmas. As the pressure is increased, through the range 10^{-2} to 10 Pa (10^{-4} to 10^{-1} torr), many-body effects begin to be important and the density of free

radicals (active chemical species) increases. Plasma etching of semiconductors is an important application in this pressure regime. At higher pressures, in the range from 10 Pa to 10 kilopascals (10^{-1} to 10^2 torr), a variety of plasma deposition techniques are important. Applications at higher pressure, from 10 kPa (10^2 torr) to several atmospheres (1 atm = 10^2 kPa), include plasma torches, plasma sprays, high-pressure deposition, and high-pressure arc lamps. *See* FREE RADICAL.

Plasma-source ion implantation. Low-energy ion beams produce secondary electron emission, radiation damage, and sputtering at material surfaces. Raising the ion energy results in ion implantation. Ions are embedded far into material surfaces, changing the surface properties. This process differs from coatings in that a layer 50–100 nanometers or more thick is transformed. Conventional ion implantation employs an accelerator to produce the energetic ions. Plasma-source ion implantation (PSII) is one of many names used to describe the process of ion implantation into objects immersed in a plasma. It provides ion acceleration in a sheath formed at the material surface.

A series of short, pulsed negative-bias voltages, δV, with $e\delta V$ much greater than the plasma electron temperature, are applied to an object immersed in the plasma. This results in a thick expanding sheath, with thickness much greater than the Debye length, during the pulses. Plasma-source ion implantation has the advantage that sheaths are conformal to objects, so ion implantation is as easy to use with objects with relatively complicated shapes as it is for planar objects. *See* ION IMPLANTATION.

Plasma etching. Chemical etching of semiconductor device structures undercuts masks by approximately 0.5 μm, so plasma etching has been required to achieve structures less than 1.0 μm in size. The directed ion acceleration at the plasma-wafer sheath boundary gives rise to the anisotropy in the etch. In many situations, a passivating polymer layer is deposited on the side walls of structures, preventing chemical etching. At the bottom of a trench, the passivating layer is sputtered away by the directed ion flux.

The most widely used etch tools have employed parallel-plate capacitive discharges operating at 13.56 MHz, a frequency that is designated for general use, at neutral pressures the order of 10 Pa (10^{-1} torr). This system is chemically etched, enhanced by the presence of ion beams. Several different approaches can be used to obtain high-density, low-pressure etching plasmas needed for finer linewidths. Among them are inductive discharges (**Fig. 4**) and helicon (bounded whistler wave) discharges (Fig. 2) operating at 13.56 MHz, and electron cyclotron resonant (ECR) discharges operating at 2.45 GHz, a frequency for which magnetron sources are widely available and relatively inexpensive because it is used in microwave ovens. *See* INTEGRATED CIRCUITS; MICROWAVE.

Plasma deposition. The walls of devices that employ plasmas are often found to be coated. The deposited species can originate from sputtering by plasma ions, accelerated by the plasma-object sheath at the edge of objects immersed in the plasma, or from deposition of molecular species created in the plasma. For example, gas-phase carbon difluoride (CF_2) can lead to poly(tetrafluoroethylene)-like polymer formation on silicon dioxide (SiO_2). Graphite can originate either by sputtering of carbon into a mostly noncarbon plasma or from a plasma based on a carbon-containing gas such as methane (CH_4). It is also possible to deposit diamond and diamondlike films. At atmospheric pressure, plasma sprays provide a way to deposit films as thick as several millimeters. Plasma sprays work by injecting powder into arc jets. The powders become molten and are ejected supersonically onto substrates, where they stick. *See* VAPOR DEPOSITION.

Fig. 4. Wafer mounted inside the inductive source used for plasma etching. The surface magnetic field produced by the permanent magnets helps to confine the plasma. (*Bruce Fritz, University of Wisconsin Engineering Center for Plasma Aided Physics*)

Thermal plasmas. At pressures greater than 10^2 Pa (1 torr), electron, ion, and neutral gas temperatures are almost in equilibrium and plasma properties are dominated by heat transport. Important applications involve arcs and plasma torches.

When the current in a discharge is increased by reducing the current-limiting resistance in external circuits, glow-discharge plasmas can make a transition to being an arc. Compared to glow discharges, which require voltages of several hundred volts, arcs are self-sustaining at relatively low voltages, the order of 10 V. This occurs because the cathode emission mechanism differs from those in glow discharges. Field emission of electrons combined with

thermionic emission result in cathode-spot current densities as high as 10^{10} A/m^2, compared to 10–100 A/m^2 for glow discharges. *See* ARC DISCHARGE; FIELD EMISSION.

Arcs are used in plasma torches, plasma sprays, lighting, gas heaters, and many other applications. Arcs often have a negative resistance characteristic, with current decreasing with increasing voltage. Arcs have high column luminosity at pressures above 1 kPa (10 torr). At pressures greater than or equal to atmospheric pressure, this luminosity is used in lighting applications. *See* ARC LAMP.

Arcs tend to be unstable, and the arc may wander about or move along an electrode. Several stabilization schemes are used to confine arcs to a restricted region. In arc lamps, the glass or quartz tube walls provides stabilization. In other applications, the glass tube is replaced by a stack of insulated, water-cooled copper disks or by an intense vortex of swirling gas or liquid. Magnetically stabilized arcs employ external magnetic fields to stabilize, deflect, and drive them.

Plasma torches have replaced chemical flames as heat sources in many applications and have many advantages. Temperatures determined by resistive heating in the arc plasma can be much higher than those found in chemical flames. Operation with inert gases avoids chemical reactions. In addition, torches can have very high thermal efficiencies, reaching as high as 92%. *See* ARC HEATING.

Controlled fusion. Controlled fusion aims at taking advantage of nuclear fusion reactions to generate net power. Cross sections for fusion reactions become significant only for energies greater than 10 keV and for light nuclei. The reaction that is likely to power the first fusion reactors, because it has the highest cross section for energies less than 30 keV, is reaction (15), employing deuterium and tritium plasmas.

$$\mathrm{D} + \mathrm{T} \rightarrow n\ (14.1\ \mathrm{MeV}) + \alpha\ (3.5\ \mathrm{MeV}) \qquad (15)$$

Plasma heating can be accomplished with ohmic heating, neutral beams, radio-frequency fields, and so forth. As the input power is increased, plasma is heated to higher temperatures, and the relative power from the fusion reaction also increases. In this state, the plasma is referred to as burning. For deuterium-tritium fusion, collisions with magnetically confined alpha particles convert power from the fusion reaction to heating power. At high enough temperatures, the plasma ignites; that is, the power provided by the alpha particles to the plasma is sufficient to balance power losses from the plasma. After ignition, the burning plasma no longer requires external heating.

At energies of the order of 10 keV or more, electrons and ions are essentially Coulomb collisionless, even though collisions leading to thermonuclear reactions are the basis of controlled fusion. Curiously, plasmas studied in controlled fusion are close relatives of low-density space plasmas. In spite of the great differences in parameter values, the ratios of characteristic lengths and time scales are comparable and many plasma phenomena are seen only in these two systems.

Advances in fusion studies have been tied to the techniques developed for plasma confinement and heating. Fusion experiments employ either magnetic confinement or inertial confinement, in which fusion reactions take place before the plasma has a chance to expand to chamber boundaries. Magnetic mirrors are an example of open systems, while tokamaks, stellarators, and reversed-field pinches are examples of closed toroidal systems. Most magnetic confinement research experiments are done on tokamaks (Fig. 3). *See* NUCLEAR FUSION. Noah Hershkowitz

Naturally occurring plasmas. Naturally occurring plasmas exist throughout the solar system and beyond. Above the atmosphere, most matter is ionized. The lower-density ionized materials are considered to be plasmas, and they behave in manners very different from the behavior of nonplasmas. Some dense materials, such as stellar matter or electrolytic solutions, are often not considered to be plasmas even though they are ionized; they behave, for the most part, as do ordinary fluids.

As discussed above, two important parameters of a plasma are its number density, n, and its temperature, T_e (Fig. 1). The combination of the two parameters determines whether the motions of the electrically charged particles that compose the plasma will be dominated by the large-scale electric and magnetic fields generated by the plasma or whether their motions will be dominated by interparticle collisions. Hot, low-density plasmas are termed collisionless, and cool, high-density plasmas are termed collisional. The behaviors of collisionless and collisional plasmas are quite different. In particular, for collisionless plasmas, a third parameter that is important is the strength of the magnetic field in the plasma. If the magnetic field is strong, the field acts to confine the plasma; if the magnetic field is weak, the movement of the plasma distorts the field and plasma motions can carry embedded magnetic-field structures.

The behavior of naturally occurring plasmas is complicated and often not well understood, and major gaps exist in knowledge about them. Several processes often operate simultaneously, some acting on large spatial scales and long time scales while others act on short spatial scales and fast time scales. Self-generated magnetic and electric fields which are difficult to predict are often important. Consequently, the fundamental dynamics of naturally occurring plasmas is an active area of research.

Ionosphere. The upper atmosphere of the Earth, known as the ionosphere, is in a plasma state because of the photoionization by sunlight of the atmospheric gas, and the impact ionization of the gas by particles associated with the aurora. Because plasmas are capable of reflecting electromagnetic waves, the ionosphere is studied for its role in enabling radio communication around the Earth and for its role in disrupting radio communication between the Earth and space satellites. *See* IONOSPHERE.

Solar atmosphere. The solar atmosphere (photosphere, chromosphere, and corona) is in a plasma

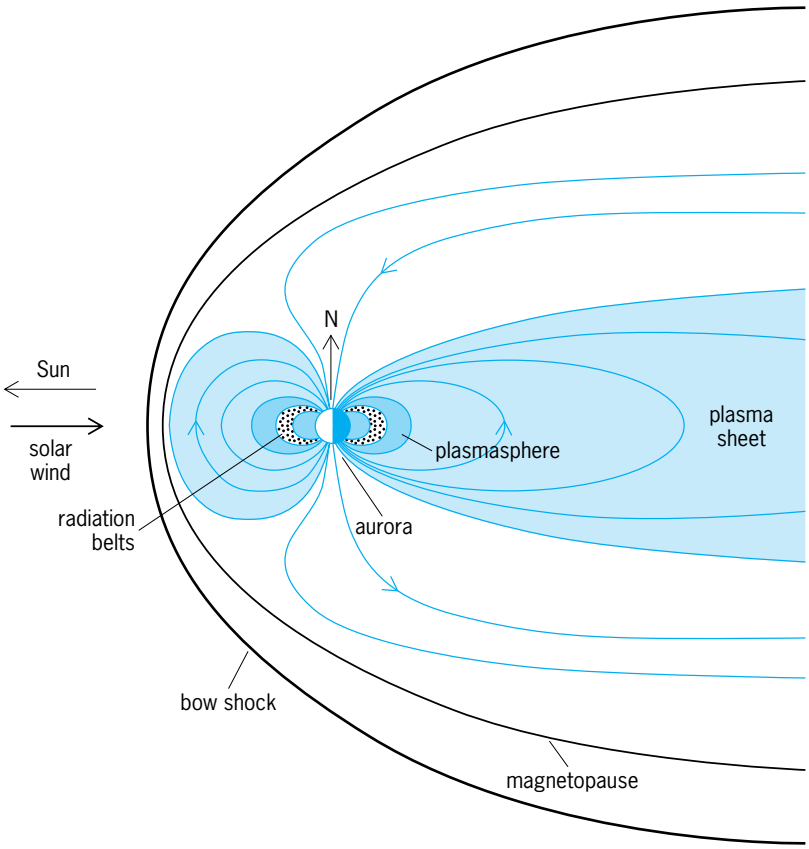

Fig. 5. Earth's magnetosphere. The magnetotail extends for a long distance off the figure to the right, away from the Sun.

state because of the heating of solar material by electromechanical processes driven perhaps by fluid motions at the solar surface. The study of the solar atmosphere is important to gain understanding of the transport of energy from the Sun, the production of solar wind (which can drive geomagnetic storms that disrupt communication), the production of energetic particles and energetic photons from solar flares in the atmosphere (which have environmental consequences), and the production of radio bursts (which interfere with communication). The insight gained from studying the solar atmosphere is used to increase and test the understanding of the atmospheres of other stars. *See* SUN.

Solar wind. The solar wind is a supersonic flow of plasma emanating from the Sun and extending far beyond the solar system, keeping the interstellar plasma pushed away from the Sun. The solar wind is turbulent, with embedded magnetic-field structures convected along with the flow. The self-generated electric field of the solar wind is the driver for geomagnetic disturbances at the Earth, and the embedded magnetic field of the solar wind modulates the ability of cosmic rays to enter the solar system. The solar wind is also the source of some intense radio bursts that interfere with communications.

Besides being important for understanding effects on the Earth, investigations of the solar wind serve to extend the understanding of stellar winds, stellar mass loss, and astrophysical-plasma turbulence. *See* COSMIC RAYS; SOLAR WIND.

Magnetospheric plasmas. The Earth's magnetosphere is the region of space where the magnetic field of the Earth controls the motion of the plasmas; the magnetosphere forms a barrier to the solar-wind flow. The interaction of the solar wind draws the magnetosphere out into a long tail (**Fig. 5**). The size, shape, and content of the magnetosphere varies with time; solar-wind-driven geomagnetic disturbances (substorms and storms) can reconfigure the magnetosphere on time scales of tens of minutes. Plasmas in the magnetosphere have two sources: solar-wind plasma that is captured by the magnetosphere, and ionospheric plasma that evaporates into the magnetosphere.

Two of the major plasmas of the magnetosphere are the plasmasphere and the plasma sheet (Figs. 1 and 5). The plasmasphere is a region of denser, cooler plasma that originates from the ionosphere. If the plasmaspheric plasma is removed by a geomagnetic disturbance, it refills from the ionosphere in a matter of days. Within the plasmasphere lie the Van Allen radiation belts, which are regions where very energetic electrons and ions are magnetically trapped for years. *See* VAN ALLEN RADIATION.

The plasma sheet is a region of lower-density, hotter plasma that has its source in the distant magnetotail where solar-wind plasma and evaporated ionospheric plasma are magnetically captured. The plasma-sheet plasma flows up the nightside magnetotail toward the Earth, then flows around the dawn and dusk sides of the Earth, through the dayside magnetosphere, then out into the solar wind, with the entire trip through the magnetosphere taking only hours.

Some of the major topics of study for magnetospheric plasmas are the energy flow from the solar wind to the Earth, the mechanics of geomagnetic disturbances, the sources and movements of energetic particles, the causes of the aurora and auroral radio emissions, and the effects of magnetospheric plasma on satellite operations. The plasmas of other planetary magnetospheres (Mercury, Jupiter, Saturn, Uranus, and Neptune) are studied to understand the sources of magnetospheric radio emissions, and the production of energetic charged particles by magnetospheres. *See* JUPITER; MAGNETOSPHERE; NEPTUNE; SATURN; URANUS.

Lightning. The hot, luminous channels of lightning strokes are dense plasmas. The plasma is produced by impact ionization of air molecules by the free electrons that carry the lightning currents. Typically, tens of thousands of amperes flow in channels that are only millimeters in radius. Lightning-channel plasmas are studied to determine the dynamics of the channels, to obtain current-versus-voltage information for lightning, to discern the effects of lightning on atmospheric chemistry, and to ascertain the sources of the intense radio emissions from lightning. *See* LIGHTNING.

Metals and semiconductors. The electrons that provide the electrical conduction of metals are in a plasma state, and the mobile electrons and electron holes in semiconductors exhibit plasma properties.

The reflection of light by a metal is an example of a plasma process (like the reflection of radio waves by the ionosphere). These solid-state plasmas differ from other plasmas because quantum-mechanical effects often are significant. The study of solid-state plasmas is important for the development of amplifiers and oscillators and for the comprehension of conductivity and superconductivity. *See* ELECTRICAL CONDUCTIVITY OF METALS; FREE-ELECTRON THEORY OF METALS; SEMICONDUCTOR; SUPERCONDUCTIVITY.

Interstellar medium. The hydrogen plasma of the interstellar medium (denoted in astronomy as H II regions) is produced by photoionization of interstellar gas by starlight and by the collisional ionization of interstellar gas by supernova-explosion plasmas, with an additional contribution from the ionized winds from stars. The interstellar plasma plays roles in the formation of stars from interstellar gas, in the dynamics and evolution of galaxies, and in the transport of cosmic rays. *See* INTERSTELLAR MATTER.

Other astrophysical plasmas. In addition to the interstellar medium, the study of plasma astrophysics includes (1) the high-velocity plasmas of supernova explosions, which are of interest for their interactions with interstellar gases and for their possible production of cosmic rays; (2) the electron-position plasmas believed to exist in pulsar magnetospheres, which are thought to give rise to the optical and radio emissions of pulsars; and (3) the radio-emitting extragalactic-jet plasmas, which emanate from the cores of large galaxies. Because of inaccessibility, the critical parameters of most astrophysical plasmas (temperatures, densities, and magnetic-field strengths) are not known. *See* ASTROPHYSICS, HIGH-ENERGY; GALAXY, EXTERNAL; PULSAR; SUPERNOVA.

Some of the major plasma-physics issues that are under study with naturally occurring plasmas are the energization of charged particles, the reconnection of magnetic fields (temporal changes in magnetic-field topology), the production of magnetic fields by dynamos, the production of electromagnetic waves, the interaction between waves and particles, and the transport of mass, momentum, and energy across magnetic fields.

Naturally occurring plasmas are in general difficult to measure. The solar-wind, ionospheric, and magnetospheric plasmas are diagnosed by single-point measurements by rockets and satellites; the solar atmosphere and all astrophysical plasmas are unreachable and must be diagnosed by the light and radio waves that they emit; and lightning is unpredictable and inhospitable to instruments and must be diagnosed primarily by the light that it emits. As a consequence of limited diagnostics, theoretical analysis and laboratory-plasma experiments play supporting roles in the investigations of naturally occurring plasmas.

Joseph E. Borovsky

Bibliography. F. F. Chen, *Introduction to Plasma Physics and Controlled Fusion*, 2d ed., 1984; H. Friedman, *Sun and Earth*, 1986; M. C. Kelley, *The Earth's Ionosphere*, 1989; D. M. Manos and D. L. Flamm (eds.), *Plasma Etching*, 1989; G. K. Parks, *Physics of Space Plasmas*, 1991; L. Spitzer, *Physical Processes in the Interstellar Medium*, 1978; H. Stix, *Waves in Plasmas*, 1992; M. A. Uman, *The Lightning Discharge*, 1987; J. Wesson, *Tokamaks*, 2d ed., 1997.

Plasma diagnostics

Techniques to measure the properties and parameters of ionized gases. Plasmas are gases in which a sufficient proportion of the atoms are ionized that the resulting mixture of free electrons and positively charged ions exhibits collective behavior resulting from electromagnetic interactions. *See* PLASMA (PHYSICS).

Plasma properties. The simplest properties of a plasma that one might wish to measure are the density (n) of particles per unit volume, the fluid velocity ($\langle \mathbf{v} \rangle$), and the temperature ($T$). These gaseous parameters may be different for the different particle species in the plasma. There are normally at least two such species, electrons (subscript e) and ions (subscript i); but often multiple ion species and neutral particles are present. (The brackets around $\mathbf{v}$ indicate that it is the average value of the particle velocities of all the particles, or of the particles of a particular species.) Because plasma temperatures are generally high, from a few thousand degrees Celsius upward, they are usually expressed in energy units such as electronvolts ($1\ \text{eV} = 1.6 \times 10^{19}\ \text{J} \simeq 11{,}600^\circ\text{C}$).

Plasmas often possess nonthermal particle distributions. That is, the velocity distribution function of electrons or ions, f_s ($s = e, i$), which measures the numbers of electrons or ions that have various velocities, has not simply the Maxwell-Boltzmann form corresponding to thermal equilibrium at some temperature but has some more complicated form $f_s(\mathbf{v}_s)$. In that case an ideal that may sometimes be approached is to measure the entire velocity distribution $f_s(\mathbf{v}_s)$ and not just the parameters n, $\langle \mathbf{v} \rangle$, and T, which are moments of this distribution. *See* BOLTZMANN STATISTICS; KINETIC THEORY OF MATTER.

The electromagnetic fields $\mathbf{E}$ and $\mathbf{B}$ are also essential parameters that affect the plasma. They are affected in turn by the charge and current densities, ρ and $\mathbf{j}$, which are simple sums of the densities and velocities of the plasma species weighted by their charge per particle, q: $\rho = \sum_s q_s n_s$; $\mathbf{j} = \sum_s q_s n_s \langle \mathbf{v}_s \rangle$.

Diagnostic techniques. Techniques to measure the plasma parameters, and their space- and time-variation, are based on a wide variety of physical principles. Direct invasion of the plasma using solid probes is possible only for relatively cold and tenuous plasmas; otherwise the probe will be damaged by excessive heat flux. However, powerful noncontact diagnostics are based on measurements external to the plasma of electromagnetic fields, waves propagated through or scattered by the plasma, photons emitted by plasma electrons or atoms, and particle emission, transmission, or scattering. The applicability of different techniques is determined by the plasma parameters, the accessibility of the plasma, and the acceptability of any perturbation of the plasma. For

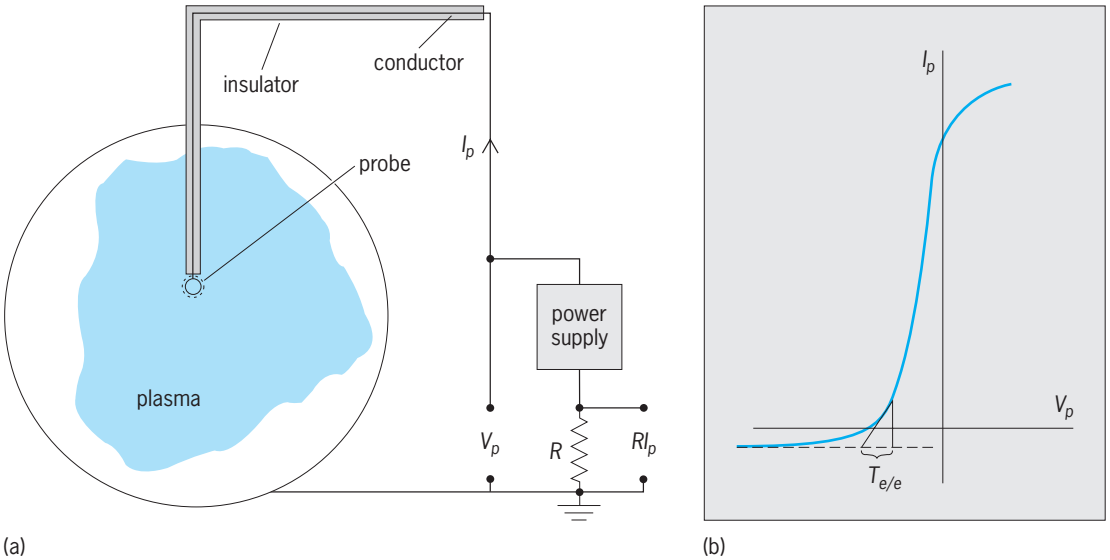

Fig. 1. Langmuir probe. (*a*) Probe setup. I_p = probe current; V_p = probe potential. (*b*) Current-voltage (I_p-V_p) characteristic. The base of the small right triangle, T_e/e, is the electron temperature in electronvolts.

example, direct plasma probing is feasible for cold laboratory plasmas and, by spacecraft, for solar-system plasmas, but not for the core of hot fusion research plasmas (because probes will not survive), for stellar or interstellar plasmas (because they are inaccessible), or for some industrial processing plasmas (because they might be contaminated by the probe). *See* NUCLEAR FUSION; SOLAR WIND.

Electric probes. Direct probing of plasma species is occasionally based on collection of specific atoms from the plasma as surface deposits that are subsequently analyzed by surface analysis techniques. By far the majority of probe measurements, however, are based on measuring the electric current carried to the probe by the plasma particles. The simplest and most direct such measurement, called the Langmuir probe, consists of an electrode of some well-defined geometry (plane, cylinder, and sphere are all used). The potential of this probe is varied and the resulting probe current is measured (**Fig. 1**).

The Debye length λ_D is an important parameter that determines the length scale over which significant electric charge density can exist in a plasma. In the advantageous situation where the Debye length is much smaller than the probe, the probe surface is surrounded by a thin "sheath" of thickness about $5\lambda_D$. Ions are accelerated toward the probe in the region beyond the sheath and enter the sheath at the sound speed, $c_s = [(ZT_e + T_i)/m_i]^{1/2}$, where m_i is the ion mass and Z its charge. Since the sheath is thin, the ion electric current density j_i to the probe is $Zen_{is}c_s$. Here, e is the electron charge and n_{is} is the ion density at the sheath edge, which is related to the electron density at the sheath edge n_{es} by $Zn_{is} = n_{es}$. This perturbed electron density is approximately half the unperturbed density n_e far from the probe, because of the dynamics of the acceleration up to the sheath. Therefore, the ion current density is $J_i = {}^1/_2 en_e c_s$. Assuming the probe potential V_p to be negative, the electron current to the probe is approximately the random one-way current of a maxwellian electron distribution, $-en_e(T_e/2\pi m_e)^{1/2}$, reduced by a Boltzmann factor, $\exp(eV_p/T_e)$. The total current in then given by Eq. (1). The electron temperature

$$j = {}^1/_2 en_e c_s - en_e(T_e/2\pi m_e)^{1/2}\exp(eV_p/T_e) \quad (1)$$

in electronvolts, T_e/e, can thus be deduced from the logarithmic slope of the probe current with respect to V_p. A simple graphical construction is shown in Fig. 1*b*. The density can then be deduced from the ion saturation current, which is the value of j at large negative values of V_p, where the second term in the right-hand side of Eq. (1) is small. Some appropriate value for T_i/T_e must be assumed.

The Langmuir probe cannot provide an independent measurement of the ion temperature because the ions are usually the attracted species and their current is almost independent of probe potential. More elaborate plasma flux measurements can be performed with devices called gridded energy analyzers, which can obtain energy distribution functions of both electrons and ions at the expense of far greater complexity, greater probe size and hence plasma perturbation, and often loss of absolute density calibration.

Electric and magnetic fields. Langmuir probes can provide measurement of the electrostatic potential in the plasma, and hence the electric field, although not quite directly, since the "floating potential" adopted by a probe at zero current is actually a few times T_e below the unperturbed space potential. The magnetic field can be measured by sensors based either on simple induction in a coil for pulsed fields, or typically on the Hall effect in a semiconductor for static fields. The former approach dominates fusion plasma research because of its linearity, noise insensitivity, and robustness, while the latter dominates space measurements in which sensitivity and dc response are essential. Magnetic probes suffer the same

limitations as Langmuir probes associated with overheating for internal plasma measurements. However, in plasmas that are sufficiently tenuous that they do not perturb the field, the internal magnetic field can often be deduced from external measurements by solving the governing magnetostatic equation, $\nabla \times \mathbf{B} = \mu_0 \mathbf{j}$, with $\mathbf{j} = 0$. If the plasma carries current, this equation can be used in the other way, to deduce the current density $\mathbf{j}$ from measurements of the magnetic field $\mathbf{B}$. Any plasma whose pressure, p (equal to $n_e T_e + n_i T_i$), is comparable to the internal magnetic field pressure, $B^2/2\mu_0$, carries appreciable current because it must satisfy the equilibrium equation $\mathbf{j} \times \mathbf{B} = \nabla p$. The plasma pressure thus tends to reduce the magnetic field inside it. A measurement of this "diamagnetic effect" provides a direct plasma pressure measurement. A more indirect estimate of plasma temperature can be obtained if both the current density and electric field are available, giving the plasma electrical conductivity $\sigma = j/E$ (parallel to $\mathbf{B}$), because the conductivity is a function of electron temperature, equal to $\sigma \simeq 1300\, T_e^{3/2}\, \Omega^{-1}\mathrm{m}^{-1}$ (for singly charged ions and negligible neutral collisions), where the electron temperature T_e is expressed in electronvolts. *See* ELECTRIC FIELD; HALL EFFECT; MAGNETIC INSTRUMENTS; MAGNETISM.

Refractive index measurements. Propagating an electromagnetic wave through the plasma and measuring the plasma's effect on it is a powerful nonperturbing diagnostic technique. The plasma's presence modifies the refractive index, kc/ω, where ω and $\mathbf{k}$ are the wave's frequency and wave vector and c the velocity of light. For the so-called ordinary wave, which propagates in a magnetic-field-free plasma, or propagates perpendicular to the ambient magnetic field $\mathbf{B}_0$ with its electric field $\mathbf{E}$ polarized along $\mathbf{B}_0$ (that is $\mathbf{k} \cdot \mathbf{B}_0 = 0$ and $\mathbf{B}_0 \parallel \mathbf{E}$), the refractive index at high frequency is given by Eq. (2). Here, ω_p, called

$$kc/\omega = \left(1 - \frac{\omega_p^2}{\omega^2}\right)^{1/2} \tag{2}$$

$$\omega_p = \sqrt{\frac{n_e e^2}{\epsilon_0 m_e}} \tag{3}$$

the plasma frequency, is given by Eq. (3), where ϵ_0 is the permittivity of vacuum. The measurement of the ordinary-wave refractive index thus provides a measure of the plasma electron density. The measurement requires a wave frequency larger than the plasma frequency for the plasma of interest; otherwise the index becomes imaginary and the wave is not transmitted. In fact, a wave frequency up to 20 times higher than ω_p is best for preventing excessive refraction but retaining adequate sensitivity. Coherent sources from microwave to x-ray are used, depending on the plasma to be measured, with far-infrared lasers dominating the magnetic fusion applications. *See* LASER; REFRACTION OF WAVES.

The index measurement is performed by measuring the phase difference caused by the plasma along an optical path, using some form of interferometer (**Fig. 2**); the Michelson or Mach-Zehnder configurations are most popular. The phase shift is approximately proportional to the average plasma density along the optical path. A variety of well-developed heterodyne techniques exist for obtaining the phase shift unambiguously in real time at high speed. Multiple optical paths are often used to obtain spatial profile information; and in the limit, a two-dimensional interferometric image of the plasma can be obtained, representing the plasma line density projected along the probing beam. *See* INTERFEROMETRY.

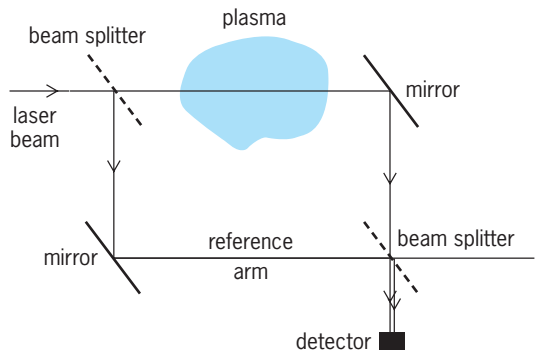

Fig. 2. Plasma interferometer with simple Mach-Zehnder configuration.

If the plasma density somewhere along the optical path rises to a value such that $\omega_p = \omega$, reflection occurs at that point. By detecting the reflected wave and measuring its phase or preferably group delay (that is, the round-trip propagation time), the location of the reflection point can be determined. This "reflectometer" measurement can also give density profile information if performed over an appropriate range of wave frequencies, but is generally less versatile and more sensitive to plasma fluctuations than interferometry.

At propagation angles other than perpendicular to $\mathbf{B}$, the magnetic field enters into the refractive index, which offers the diagnostic opportunity for measuring $\mathbf{B}$. The predominant effect is Faraday rotation: the progressive angular displacement of the angle of wave polarization as it propagates. The rotation is proportional to the path-integral of $n_e(\mathbf{B} \cdot \mathbf{k})$; so if the density is determined independently, the component of $\mathbf{B}$ along the propagation direction is obtained. This technique has been used in the laboratory, and also for interstellar radiation, providing one of the few measurements of the interstellar magnetic field.

Thomson scattering. In addition to causing the phase shifts measured by interferometry, plasma particles scatter electromagnetic waves by Thomson scattering. Collection of the scattered radiation provides an important diagnostic technique. The scattering can take place either incoherently (the powers of light scattered from individual particles add) or coherently (scattered electric fields add), depending on whether the quantity $2\lambda_D k \sin\theta$ is much greater than 1 (incoherent) or not (coherent). Here θ is the scattering angle.

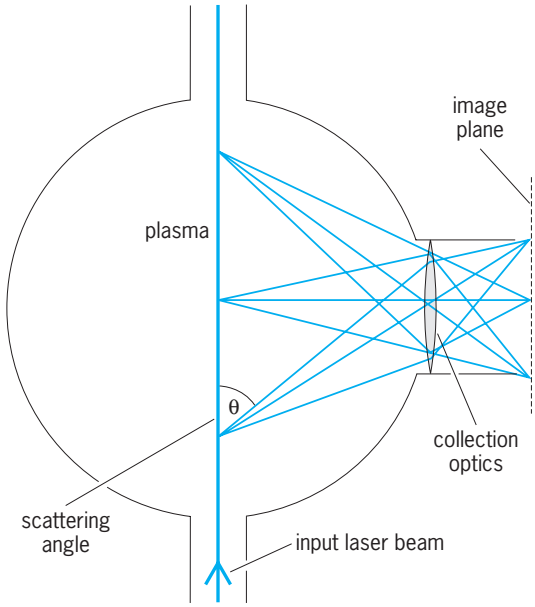

Fig. 3. Thomson scattering apparatus that uses imaging of the scattered light to give a spatial profile of the electron temperature T_e and the electron density n_e.

Incoherent scattering is purely from the electrons. Each scattered photon is Doppler-shifted by an amount corresponding to the velocity of the electron from which it scatters. Consequently there is a one-to-one correspondence between the scattered spectrum and the electron velocity distribution. Unfortunately the Thomson scattering cross section is so small (6.652×10^{-29} m^2) that typically only perhaps one photon in 10^{13} might be scattered into the collection optics. Therefore high-energy pulsed laser sources are essential, great care to eliminate spurious signals must be taken, and even so the measurement is difficult and the total detected spectrum is often only a few hundred photons. It is therefore very rare for details of the electron distribution other than temperature and density to be obtainable. However, for those parameters that are measured, the spatial localization defined by the input and collection crossed optical paths permits highly detailed and unambiguous profiles to be obtained (**Fig. 3**).

Coherent scattering samples the correlations between electrons, notably their density fluctuations. Therefore coherent scattering is used to measure plasma turbulence or to sense high-amplitude heating waves inside the plasma. Also, since electrons form correlated shielding clouds around the ions of the plasma, coherent scattering is able to sense the velocity of these clouds and hence measure ion parameters such as temperature. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

Plasma radiation. Electromagnetic wave emission from the free electrons in the plasma occurs predominantly from acceleration either by the ambient magnetic field (cyclotron or synchrotron emission) or by collisions with other plasma particles (bremsstrahlung).

Cyclotron emission in magnetically confined plasmas provides a spatially resolved measurement of electron temperature because the resonant frequency (eB/m_e) identifies the location of emission when B is a known function of position; and in the usual situation that the plasma is optically thick, the intensity is equal to the blackbody level, $\omega^2 T_e/8\pi^3 c^2$ (the Rayleigh-Jeans approximation), that is, it is proportional to the electron temperature. In astrophysical plasmas, cyclotron radiation from relativistic electrons is often intense enough for terrestrial detection, especially when the distribution function is such that collective amplification of the waves occurs. The radiation provides information about both the electrons and the magnetic field. *See* CYCLOTRON RESONANCE EXPERIMENTS; HEAT RADIATION; PARTICLE ACCELERATOR.

Bremsstrahlung from optically thin (hot and tenuous) plasmas has a photon energy spectrum that resembles the electron energy distribution; so, for example, its logarithmic slope gives the electron temperature. The spectrum for kiloelectronvolt-temperature plasmas is measured by x-ray techniques such as pulse-height analysis or approximately using foil filters. Large, dense, or cold plasmas are often optically thick to bremsstrahlung, in which case their spectrum is blackbody and provides electron temperature. Alternatively, by backlighting the plasma (for example with x-rays), their bremsstrahlung (collisional) absorption provides a measure of line-averaged density. This last technique is particularly important for laser-produced plasmas and inertial fusion research. *See* BREMSSTRAHLUNG.

Plasma spectroscopy. Line radiation arises from electronic transitions between bound levels of atomic species in the plasma. The characteristic wavelengths serve to identify the species, and the intensities are proportional to the excited-state densities. Therefore line radiation is the primary method for determining the atomic and molecular composition of plasmas. Excitation of the atom from the ground state usually occurs by electron impact, requiring an electron energy greater than the excitation energy. Therefore the ratio of the intensities of lines from different excited states whose excitation energies differ by an amount comparable to the electron temperature, is a measure of electron temperature. The rate of excitation is proportional to electron density; however, in order to eliminate from the measured intensity its dependence on the atomic density, a line ratio is generally chosen for measurement. If the population of one excited level of the line pair chosen is substantially affected by collisional deexcitation, because it is a metastable level for example, or is populated by recombination, then the ratio is sensitive to electron density. *See* ATOMIC STRUCTURE AND SPECTRA; SCATTERING EXPERIMENTS (ATOMS AND MOLECULES).

Lines are broadened by the Doppler effect arising from thermal atom velocities. So the spectral line shape reflects the atom velocity distribution and, especially, its width gives the atom temperature. If the atoms are charged (because of being a partially ionized species), they are collisionally well coupled

to the other ions in the plasma and often, therefore, represent the overall ion temperature. Moreover, if the plasma is flowing, its ion velocity can be obtained from the spectral line shift. At higher densities, broadening can also arise from collisions. This Stark broadening gives a width approximately proportional to $n^{2/3}$ and provides a density measurement. Spectral lines are split by magnetic fields (Zeeman effect) and electric fields (Stark effect), and the different components of the resulting multiplets have specific polarization with respect to the field. This effect thus provides an opportunity to measure both the field's amplitude and direction. *See* DOPPLER EFFECT; SPECTROSCOPY; STARK EFFECT; ZEEMAN EFFECT.

Heavy-particle diagnostics. Neutral atoms are continually formed from plasma ions by charge exchange with other neutrals or recombination with electrons. The neutrals formed may escape promptly from the plasma, providing a means for detection of the ion distribution using a neutral-particle spectrometer. In fusion research plasmas, fusion reactions give rise to megaelectronvolt-energy neutrons and charged products. These too provide information on the ion temperature and density via the reaction rate and the emitted spectrum.

Energetic neutral or ion beams, generated specifically for the purpose, can be used for active plasma probing. Their attenuation is generally proportional to plasma density; but more important special-purpose applications include the internal measurement of potential (by local beam ionization in the plasma) and the provision of localized atomic species for other spectroscopic diagnostics. I. H. Hutchinson

Bibliography. O. Auciello and D. Flamm, *Plasma Diagnostics*, Academic Press, Boston, 1989; R. H. Huddlestone and S. L. Leonard, *Plasma Diagnostic Techniques*, Academic Press, New York, 1965; I. H. Hutchinson, *Principles of Plasma Diagnostics*, Cambridge University Press, New York, 1987; W. Lochte-Holtgreven, *Plasma Diagnostics*, North-Holland, Amsterdam, 1968.

Plasma propulsion

The imparting of thrust to a spacecraft through the acceleration of a plasma (ionized gas). A plasma can be accelerated by electrical means to exhaust velocities considerably higher than those attained by chemical rockets. The higher exhaust velocities (specific impulses) of plasma thrusters usually imply that, for a particular mission, the spacecraft would use less propellant than the amount required by conventional chemical rockets. This means that for the same amount of propellant a spacecraft propelled by a plasma rocket can increase in velocity over a set distance by an increment larger than that possible with a chemical propulsion system. An increasing number of near-Earth satellites have benefited from the propellant savings of plasma thrusters and rely on them for attitude control and station-keeping maneuvers. Generally, the larger the required velocity increment, the more beneficial is plasma propulsion compared to propulsion with rockets having lower exhaust velocities. Consequently, many ambitious missions—such as deep-space exploration including Moon tours and orbiters to outer planets, sample return missions, and heavy cargo and piloted missions—may be enabled by plasma propulsion. *See* PLASMA (PHYSICS); SPECIFIC IMPULSE.

Plasma propulsion is one of three major classes of electric propulsion, the others being electrothermal propulsion and ion (or electrostatic) propulsion. In electrothermal devices the working fluid is a hot nonionized or very weakly ionized gas. All three electric propulsion systems offer high specific impulses; the plasma devices are further distinguished by their high thrust density (thrust per unit exhaust area) as indicated in the **table**. *See* ELECTROTHERMAL PROPULSION; ION PROPULSION; SPACECRAFT PROPULSION.

Types of plasma thrusters. One common type of electric thruster in which a plasma is generated is the ion thruster. However, in the ion thruster only the plasma ions are accelerated, and therefore this variant of plasma propulsion is not discussed here. The three most mature types of plasma propulsion devices are the Hall thruster, the magnetoplasmadynamic thruster (MPDT), and the pulsed plasma thruster (PPT).

Although these three types of plasma thrusters share the distinction of having been in use, or in the case of the MPDT having been tested, in space, there are many other promising plasma propulsion concepts that have been advanced to various degrees of technological readiness over the past few decades. These concepts differ in [1] their mode of operation (pulsed or steady-state) [2] methods of generating

Plasma propulsion as a class of electrical propulsion*

Propulsion system	Working fluid	Acceleration	Exhaust velocity, km/s (mi/s)	Thrust density, N/m^2 (lbf/ft^2)	Efficiency, %
Plasma	Plasma	Electromagnetic, or electromagnetic with thermal expansion	10–60 (6–40)	10–1500 (0.2–31)	10–60
Electrothermal	Hot gas, weakly ionized	Thermal expansion	2–10 (1.2–6)	10–100 (0.2–2)	40
Ion or electrostatic	Ions	Electrostatic	30–100 (18.7–60)	0.1–4 (0.002–0.1)	60–90

*Numerical values are typical.

the plasma (electrode-based discharge, coil-based induction, or antenna-based radiation) [3] plasma acceleration mechanisms (Lorentz forces; entrainment by plasma sheet; or traveling waves, whereby an alternating magnetic field is synchronized with the electric-field pulses to drive the plasma), [4] expansion through magnetic nozzles (whereby a heated plasma is exhausted through an expanding magnetic field), [5] implementation of the magnetic field (applied or self-induced), [6] propellant-feeding schemes (gas-fed or solid-fed, primarily polytetrafluoroethylene), [7] geometry (rectangular, coaxial, pinch, parallel-plate, constant, or variable-area geometry), and other details of implementation.

Hall thruster. The most successful plasma thruster to date is the Hall thruster. It has been gaining acceptance in the West since the early 1990s, after three decades of improvement in the former Soviet Union. More than 200 Hall thrusters have been flown so far on about 30 spacecraft, mostly by the former Soviet Union and Russia. In a Hall thruster, (**Fig. 1**) a mostly radial magnetic field is applied inside the acceleration chamber. A hollow cathode, placed outside the chamber, is kept at a few hundred volts below the potential of an annular anode located upstream in the device, thus setting a mostly axial electric field inside the chamber. As the cathode is heated, it emits electrons, which drift in the opposite direction of the electric field and enter the device, where they are subjected to the radial magnetic field and the axial electric field. This leads to an azimuthal (that is, about the device's axis of symmetry) drift for the electrons, which constitutes an electric current called the Hall current. The drifting electrons collide with propellant gas molecules (typically xenon) injected in the device through the anode. These impacts produce a plasma that finds itself acted on by an electromagnetic force, called the Lorentz force, which results from the interaction of the radial magnetic field and the azimuthal Hall current that are perpendicular to each other. The plasma is thus accelerated and exhausted out of the thruster at high exhaust velocities (from 10 km/s or 6 mi/s to above 50 km/s or 30 mi/s depending on the available power). The acceleration mechanism can also be described, equivalently, to be the result of the action of the applied electric field on the newly born ions (whose mass is too large to allow them to be magnetized) with the electrons acting as a neutralizing cloud whose momentum is coupled to the ions via interparticle electrostatic fields. Since the entire plasma is accelerated, and not only the positive ions, there is no space-charge buildup as in the ion thruster, and the device can have a thrust density many times that of the ion engine. *See* HALL EFFECT; MAGNETISM; MAGNETOHYDRODYNAMICS.

Hall thrusters can attain thrust efficiencies (ratio of propulsive power of the exhaust jet to input electric power) in excess of 50% and are thus attractive for many space applications. They have been developed

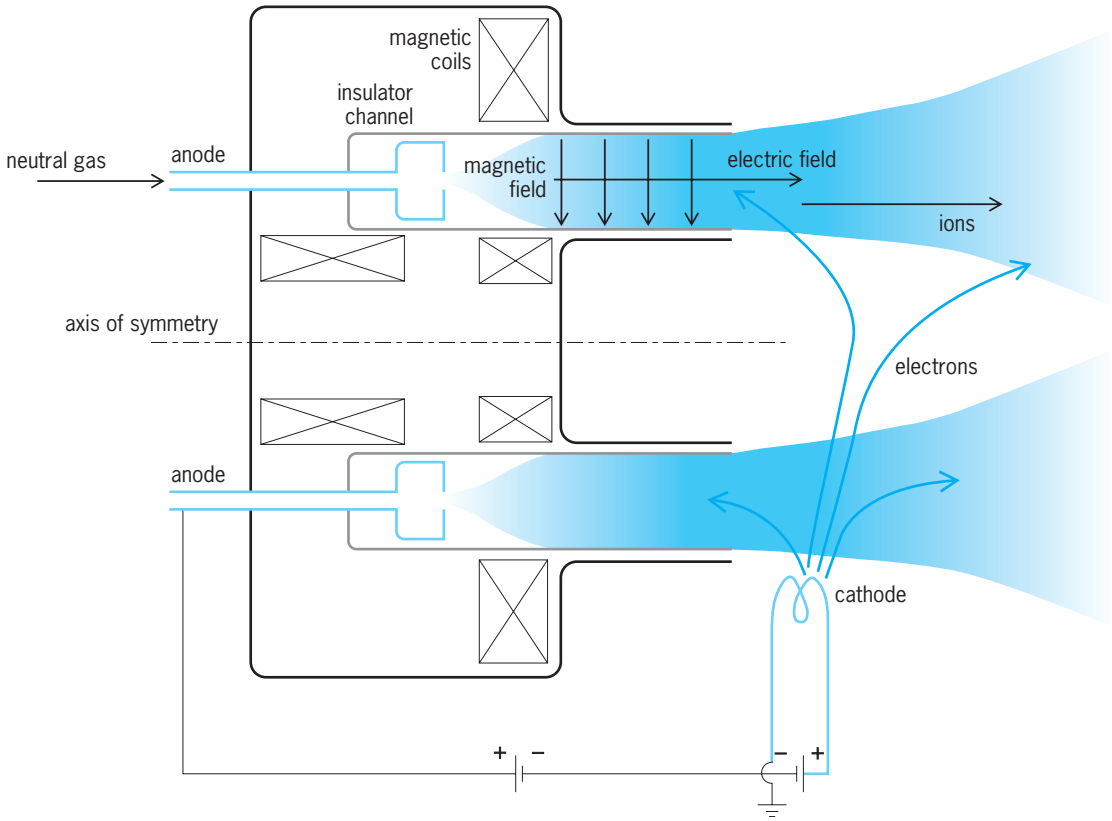

Fig. 1. Hall thruster schematic, showing the electric circuit, the applied electric and magnetic fields, and typical trajectories of ions and electrons. The electrons entering the cylindrical device execute azimuthal drifts (about the shown axis of symmetry) and collide with the injected neutral gas to produce the plasma.

for power levels of the order of 1 kW. Although they had considerable development in the former Soviet Union, a number of research problems remain, including scaling them to high power (in excess of 10 kW) for applications requiring high thrust, high specific impulse, or low power (below 50 W) for micropropulsion; lowering the divergence of the exhaust beam of plasma (which can adversely impact the performance of the device and the integrity of exposed spacecraft surfaces); and understanding the basic processes controlling the various oscillations, plasma waves, and instabilities that are excited in the plasma.

Magnetoplasmadynamic thruster. One of the most promising and thoroughly studied electromagnetic plasma accelerators is the magnetoplasmadynamic thruster. In this device the plasma is both created and accelerated by a high-current discharge between two coaxial electrodes. The discharge is due to the breakdown of the gas as it is injected in the interelectrode region. The acceleration process can be described as the result of the action of a body force acting on the plasma (**Fig. 2**). This body force is the Lorentz force created by the interaction between the mostly radial current conducted through the plasma and the magnetic field. The latter can be either externally applied by a magnet or self-induced by the discharge, if the current is sufficiently high. The reliance on the radial component of the current (which is aligned with the electric field) instead of the Hall component (perpendicular to both electric and magnetic fields) allows operation at much higher plasma densities (and thrust densities) than with the Hall thruster.

Microscopically, the acceleration process can be described as the momentum transfer from the electrons, which carry the current, to the heavy particles through collisions. Such collisions are responsible for the creation of the plasma (ionization) and its acceleration and heating (Joule heating). Even though charged particles (ions and electrons) in an electromagnetic field drift at the same velocity irrespective of their mass, the electrons recover faster from a collision because of their small mass and thus are the prime beneficiaries of the energy in the electromagnetic field. In other words, the electrons carry the current responsible for the acceleration of the plasma in the magnetoplasmadynamic thruster. Indeed, the thrust T of a purely electromagnetic coaxial and cylindrical self-field plasma accelerator is directly and strongly related to that current by Eq. (1), where b is a constant dependent on the

$$T = bJ^2 \tag{1}$$

ratio of the electrode radii and J is the total current conducted through the plasma. The thrust of a purely electromagnetic device is therefore independent of the mass-flow rate and type of propellant. The electromagnetic thrust of experimental prototypes of the magnetoplasmadynamic thruster obeys this law when operated at full ionization, which corresponds to operation at current levels at and above

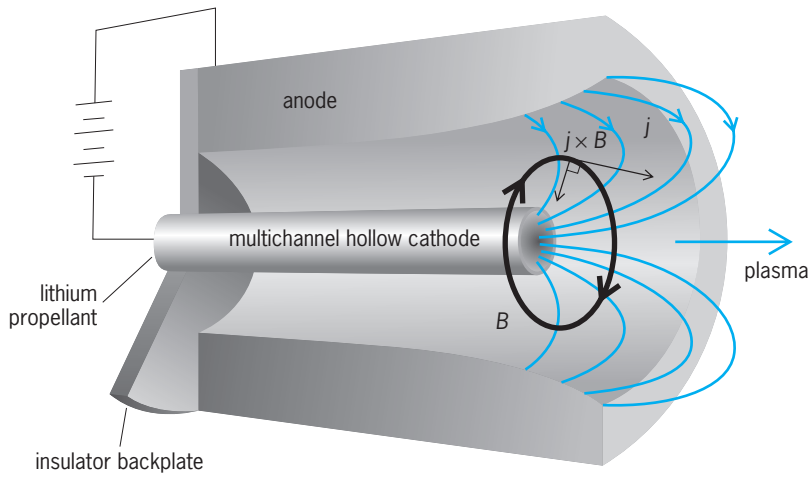

Fig. 2. Self-field, coaxial, magnetoplasmadynamic thruster. Here the current density *j* (color curves) and the induced azimuthal magnetic field *B* (black circle) interact to produce the mostly axial Lorentz force density. Electrodes are typically made of tungsten.

J_{fi}, the current required for effective full ionization of the plasma. This is the regime of $\xi \geq 1$, where ξ is the nondimensional current defined by Eq. (2) and J_{fi} in amperes is given by Eq. (3). Here $\dot{m}$ is

$$\xi \equiv \frac{J}{J_{fi}} \tag{2}$$

$$J_{fi} = \left(\frac{\dot{m}}{b}\right)^{1/2} \left(\frac{2\epsilon}{M}\right)^{1/4} \tag{3}$$

the mass-flow rate of propellant into the device in kilograms per second, and M and ϵ are the atomic mass and first ionization potential of the propellant in kilograms and joules, respectively. *See* IONIZATION POTENTIAL.

For typical magnetoplasmadynamic thruster geometry, b is about 0.17 N/kA^2 (0.038 lbf/kA^2), and for operation with 3 g/s of argon the nominal condition $\xi = 1$ is attained at $J = 12$ kA [from Eqs. (2) and (3)]. Furthermore, Eq. (1) implies an electromagnetic thrust of 25 newtons. At these conditions, the voltages associated with the back electromotive force, plasma heating, and electrode sheaths add up to about 80 V. Power levels exceeding 1 MW are thus required for Eq. (1) to hold at these operating conditions. It is, however, the universal scaling parameter ξ that determines the applicability of the electromagnetic thrust law [Eq. (1)] and not the power level. In other words, it is possible to attain the regime of electromagnetic acceleration at considerably more modest power levels as long as $\xi \geq 1$ is satisfied. This can be achieved by reducing the propellant mass-flow rate, and some devices are operated at power levels as low as 15 kW. However, the power losses associated with the electrodes and ionization dominate at lower power, and thrust efficiencies exceeding 20% are still attainable only with the megawatt-class devices or with the use of lithium as propellant. Below about 250 kW, an applied magnetic field is generally needed to improve the performance.

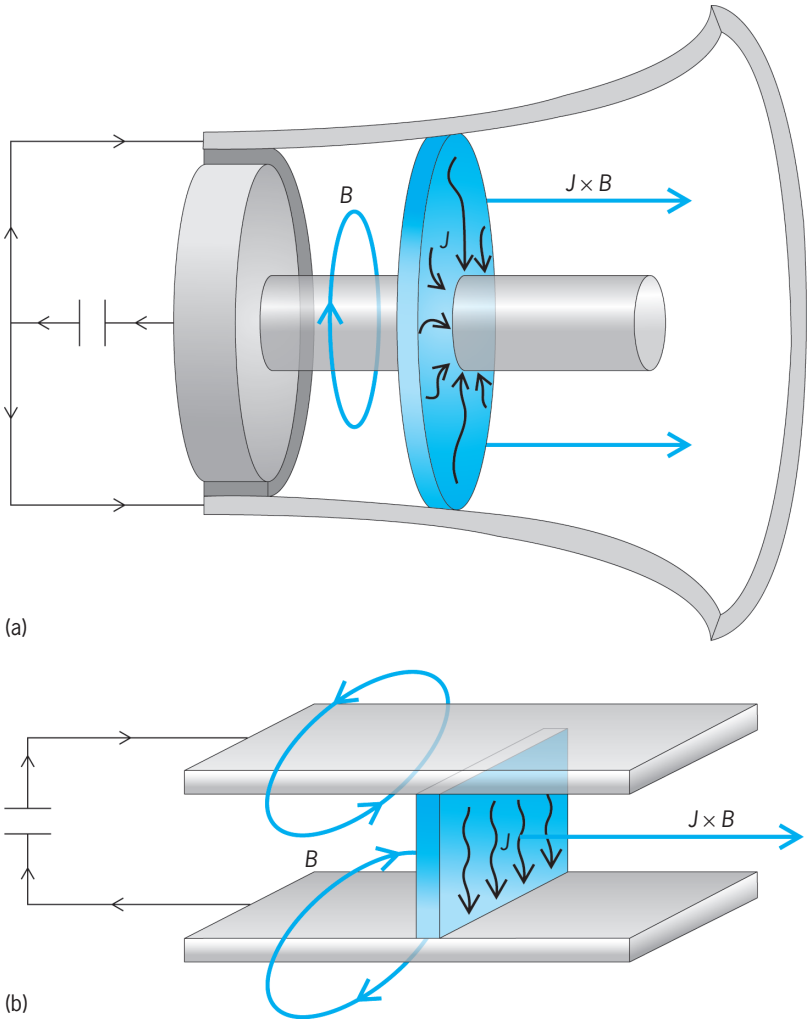

Fig. 3. Schematic of two configurations of the pulsed plasma thruster, showing the current sheet, the induced magnetic field, and the resulting electromagnetic force. (*a*) Coaxial configuration. (*b*) Parallel-plate configuration.

Research has demonsrated the advantage of using lithium vapor as propellant, injected, along with small amounts of barium vapor, from a multichannel hollow cathode (Fig. 2). Lithium (which has a low ionization potential) offers the advantage of low ionization losses, and barium, which has a low work function, ensures a low cathode operation temperature, which lowers the cathode evaporation rate, leading to longer lifetime. Such a magnetoplasmadynamic thruster variant is sometimes called the lithium Lorentz force accelerator (Li-LFA).

The scaling of the various power sinks with the input power is an area of active research, and of major concern for improving efficiency. Other research areas include the understanding of various mechanisms controlling the efficiency as well as an operation-limiting phenomenon called onset that occurs when ξ exceeds a critical value. Operation above onset can lead to severe damage to the anode and to strong voltage fluctuations.

Pulsed plasma thrusters. This type of plasma thruster, called the PPT, relies on pulsing a high-current gas discharge on a time scale of a few microseconds. This is done by quickly discharging the energy stored in a capacitor into a gas that is either injected (gas-fed PPT) or vaporized from a solid propellant (ablative PPT), typically polytetrafluoroethylene. The quick rise in current through the gas forces the formation of a plasma sheet through which the current flows as shown in **Fig 3**, where two possible configurations are illustrated. The current sheet is subjected to the Lorentz force produced by the interaction of the current it carries and the induced magnetic field. This causes the current sheet to accelerate downstream, ionizing and entraining ambient gas as in a snowplow. The exhausted plasma and gas give the propulsive impulse. The pulsed operation of such electric propulsion devices allows power-limited spacecraft to take advantage of high-current plasma acceleration at the low average power available on current spacecraft (hundreds of watts).

Depending on the gas pressure in the device, the gas injection mode, the rate of current rise, and other implementation details, acceleration by thermal effects can dominate over electromagnetic acceleration, or the acceleration can be due to a so-called deflagration effect where the current pattern spreads over the discharge chamber and exerts a body force that extends over a large volume of the plasma. This latter mode can lead to higher exhaust velocities and efficiencies.

Another type of plasma thruster that is an inherently pulsed device, the pulsed inductive thruster (PIT), relies on using coils placed outside the discharge to induce a current sheet in a gas injected from a downstream nozzle. The PIT has the advantage of foregoing the use of electrodes, which often are the life-limiting factor for plasma thrusters. Another pulsed inductive thruster variant, the Faraday accelerator with radio-frequency assisted discharge (FARAD), relies on injecting a plasma, instead of a gas, from an upstream pre-ionization radio-frequency plasma source and on an applied magnetic field to control the plasma injection and acceleration processes.

Propellants for plasma thrusters. Plasma acceleration allows a wider choice of propellants than any other type of propulsion. Plasma-propelled spacecraft thus have the potential of using propellants that are abundant in their environment rather than from Earth alone. There are, however, some constraints. For MPDTs, although Eq. (1) implies that the magnitude of the electromagnetic thrust is independent of the propellant, the attainability of the nominal condition [$\xi = 1$, Eqs. (2) and (3)], compatibility with thruster components, requirements of rate kinetics (for the case of a hybrid thruster), storability, and toxicity restrict the choice of appropriate propellants. Laboratory MPDTs and PPTs have been tested with argon, krypton, xenon, helium, hydrogen, air, deuterium, ammonia, hydrazine, lithium, and polytetrafluoroethylene, among other propellants and mixtures of propellants. For Hall thrusters, the plasma dynamics constrain the choice of propellants to gases of high atomic mass and reatively low ionization potential. Xenon and bismuth have been the most favored propellant.

Plasma thruster use in space. More than 200 plasma thrusters have been flown to date, mostly by the former Soviet Union and Russia, and mostly of the Hall thruster type. This excludes the dozens of ion thrusters presently used on geosynchronous spacecraft. A few PPTs of the solid-fed [polytetrafluoroethylene-ablative] type were used on Russian and U.S. spacecraft in the 1970s and 1980s. The first NASA spacecraft to use an ablative PPT was the EO-1 satellite launched in 2001. These ablative PPTs have a modest thrust and efficiency. The thrust level, however, is very reproducible and can be precisely controlled, thus making these systems ideal for fine-tuning spacecraft attitude control. A gas-fed quasi-steady pulsed magnetoplasmadynamic thruster was tested in space in 1998 onboard the Japanese spacecraft SFU; but MPDTs, which are inherently high-power devices, are still awaiting the availability of high electric power in space before their adoption in actual missions. With technical advancements, successful flight tests, and the advent of a new era of deep-space exploration, plasma propulsion is presently entering its era of wide application. *See* SPACE POWER SYSTEMS. Edgar Y. Choueiri

Bibliography. R. L. Burton and P. J. Turchi, Pulsed plasma thruster, *J. Propul. Power*, 14(5):716–735, 1998; E. Y. Choueiri, A critical history of electric propulsion: The first 50 years, *J. Propul. Power*, 20(2):193–203, 2004; E. Y. Choueiri, Fundamental difference between the two variants of Hall thrusters, *Phys. Plasmas*, 8(11):5025–5033, 2001; R. G. Jahn, *Physics of Electric Propulsion*, Dover Publications, 2006; H. R. Kaufman, Technology of closed-drift thrusters, *AIAA J.*, 23(1):78–80, 1985; A. Kodys and E. Y. Choueiri, *A Critical Review of the State-of-the-Art in the Performance of Applied-Field Magnetoplasmadynamic Thrusters*, AIAA Pap. 05-2005-4247, 2005; M. Martinez-Sanchez and J. E. Pollard, Spacecraft electric propulsion—An overview, *J. Propul. Power*, 14:(5), 688–699, 1998.

Plasmid

A circular extrachromosomal genetic element that is ubiquitous in prokaryotes and has also been identified in a number of eukaryotes. In general, bacterial plasmids can be classified into two groups on the basis of the number of genes and functions they carry. The larger plasmids are deoxyribonucleic acid (DNA) molecules of around 100 kilobase (kb) pairs, which is sufficient to code for approximately 100 genes. There is usually a small number of copies of these plasmids per host chromosome, so that their replication must be precisely coordinated with the cell division cycle. Their replication genes are clustered within a small segment of the genome. Larger plasmids usually can mediate their own transfer from cell to cell by bacterial conjugation. The genes mediating transmissibility are clustered in a segment of approximately 25 kb. The other regions (around 70 kb) of these larger plasmids can accommodate as many as 70 additional genes which determine a variety of properties (see **table**). The plasmids in the second group are smaller in size, about 6–10 kb. These plasmids may harbor 6–10 genes and are usually present in multiple copies (10–20 per chromosome). Although the smaller plasmids do not carry transfer genes, many of them can be mobilized to recipient cells if a larger transmissible plasmid is present in the same host cell.

Distribution. Plasmid has been identified in a large number of bacterial genera. Some bacterial species harbor plasmids wih no known functions (cryptic plasmids) which have been identified as small circular molecules present in the bacterial DNA. The host range of a particular plasmid is usually limited to closely related genera. Some plasmids, however, are much more promiscuous and have a much broader host range.

Functions. The functions specified by different bacterial plasmids are usually quite specialized in nature. Moreover, they are not essential for cell growth since the host bacteria are viable without a plasmid when the cells are cultured under conditions that do not select for plasmid-specified gene products. Plasmids thus introduce specialized functions to host cells which provide versatility and adaptability for growth and survival. Plasmids which confer antibiotic resistance (R plasmids) have been extensively characterized because of their medical importance.

Plasmid incompatibility. A fundamental characteristic of plasmids by which they have been classified into specific groups is the ability of different plasmids to coexist stably in a host cell. Some pairs of plasmids, isolated independently from nature, can coexist in the same host cell: this is referred to as compatibility. Other pairs of plasmids are called incompatible since they are not inherited stably when present in the same host cell. Based on these criteria, over 25 plasmid incompatibility groups have been identified in gram-negative bacteria, and more than a dozen incompatibility groups have been reported in gram-positive bacteria. R plasmids are represented in many different incompatibility groups, whereas other plasmid-specified functions are not as widely distributed among incompatibility groups. Experiments have shown that plasmid incompatibility is a result of the control of plasmid replication. Plasmids can be changed into a new incompatibility group as a result of a single base change in replication gene which controls the plasmid copy number.

R plasmids. The large increase in multiply drug-resistant bacterial pathogens has resulted from increasing use of antibiotics in chemotherapy and animal husbandry. Gram-negative and gram-positive bacterial species have acquired resistance to a variety of antibiotics such as penicillins and cephalosporins, chloramphenicol, aminoglycosides (amikacin, gentamycin, kanamycin, streptomycin, spectinomycin, tobramycin), tetracyclines, sulfonamides, trimethoprim, and MLS (macrolide, lincomycin, and streptogramin B-type) and to toxic heavy metals such as mercuric, cadmium, and silver ions. R plasmids do not confer resistance to all these antibiotics and

Some properties determined by plasmids

Property	Conferred by plasmids of	
	Gram-negative bacteria	Gram-positive bacteria
Autonomous replication	+	+
Genetic transfer	+	+
Bacteriocin production and resistance	+	+
Antibiotic resistance	+	+
Resistance to heavy metals	+	+
Resistance to agents that damage DNA	+	
Metabolism of hydrocarbons and aromatic compounds	+ (*Pseudomonas*)	
Sugar fermentation	+	+
Hydrogen sulfide production	+	
Protease and urease production		+
Toxin production	+	+
Virulence and colonization factors	+	
Hemolysin production	+	+
Restriction/modification	+	+
Antibiotic production	*Streptomyces coelicor*	
Tumorigenicity in plants	*Agrobacterium tumefaciens*	

heavy metals simultaneously, but transmissible R plasmids conferring resistance to four to six antibiotics are not uncommon. Smaller R plasmids usually confer resistance to one or two antibiotics or heavy metals. *See* ANTIBIOTIC.

The biochemical mechanisms of plasmid-determined antibiotic resistance have been elucidated in the majority of instances. At least six different types of resistance mechanisms have been described: (1) cleavage of a specific bond in an antibiotic, as in the inactivation of penicillins or cephalosporins by β-lactamases; (2) inactivation by chemical modification, as in the acetylation of chloramphenicol; (3) synthesis of antibiotic-insensitive substitute enzymes, as in the case of sulfonamide or trimethoprim resistance; (4) alteration of a macromolecular target site, as in methylation of 23S ribonucleic acid (RNA) to confer erythromycin or lincomycin resistance; (5) lower cell permeability, as in the case of tetracycline resistance; and (6) sequestering of the antibiotic so that it cannot bind to its target site as in the case of the binding of fusidic acid by chloramphenicol acetyltransferase (the enzyme which acetylates chloramphenicol). In this case the same gene confers resistance to two structurally unrelated antibiotics by two entirely different mechanisms. Many of the aminoglycoside antibiotics are inactivated by several different

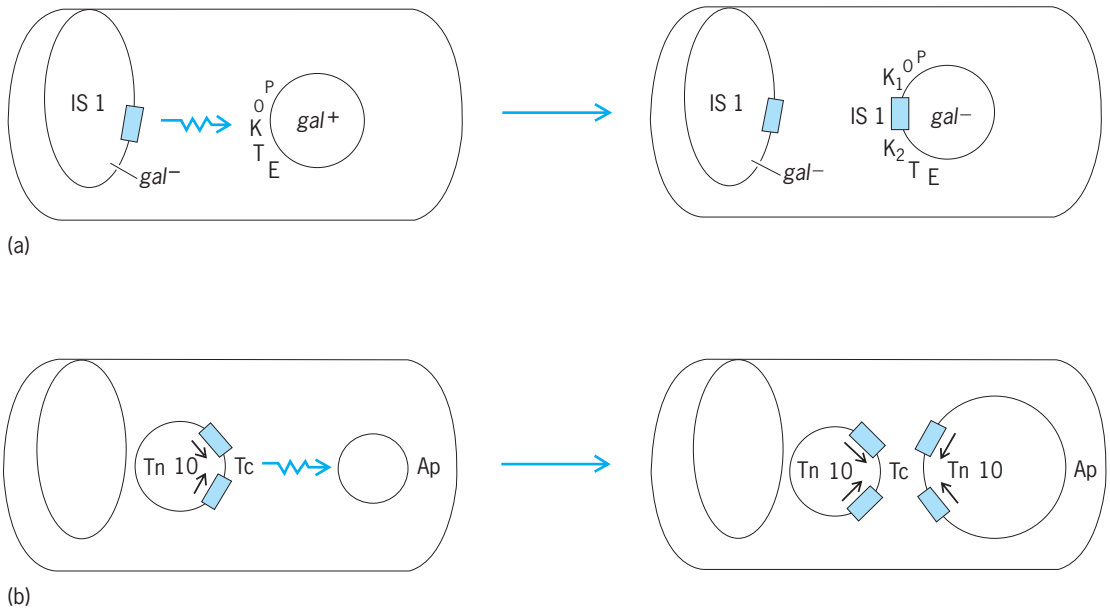

Fig. 1. Translocation of insertion sequences (IS) and transposable elements (Tn). (*a*) At left an IS 1 element is initially present in the host chromosome which carries a mutation in the *gal* genes which mediate galactose metabolism. The cell can metabolize galactose owing only to the presence of *gal* genes on a plasmid. After insertion of IS 1 into the *gal* genes on the plasmid, these genes are inactivated; the cell at right cannot metabolize galactose. (*b*) Tn 10, which mediates tetracycline resistance (Tc), translocates from one plasmid to another plasmid which also confers ampicillin resistance (Ap). In both insertion sequence and transposable element translocation, a copy of the element is retained in the donor genome.

types of chemical modification by specific enzymes. *See* DRUG RESISTANCE.

Insertion sequences and transposable elements. Many plasmids have been found to contain genetic elements called insertion sequences and transposable elements. Insertion sequences are specific DNA sequences that exist on bacterial genomes and have the remarkable property of being able to translocate (jump) spontaneously to new sites on the same or other genomes in a cell (chromosome, plasmid, or virus) by a mechanism which is independent of the normal recombination system (**Fig. 1**). The location of the insertion sequence elements on several bacterial plasmids suggests that they may play a seminal role in plasmid evolution. These insertion sequence elements appear to serve as recombination sites for the fusion of separate genomes, each of which carries a group of specific genes, to form more complex genomes specifying all of the fun of each of the components. Insertion sequence elements may thus serve as linkers in the assembly of different modules of genetic information in the evolution of plasmids and other genomes.

The alarming spread of multiple-drug resistance among bacterial pathogens also appears to be a manifestation of the genetic plasticity made possible by DNA insertions. Most drug resistance genes on bacterial plasmids are flanked by repetitious DNA sequences, usually in inverted orientation, which are analogous to insertion sequences or may even be insertion sequences. This entire unit, referred to as a transposable element, can translocate to new sites in a manner similar to insertion sequence elements with the important difference that the intervening drug resistance genes are transposed in addition to the flanking repetitious sequences. Transposable elements can transpose at low frequency from one genome to another and thus lead to a broad dissemination of the same drug resistance genes among different plasmids. *See* TRANSPOSONS.

Rearrangements and evolution. Insertion sequence and transposable elements also mediate various types of plasmid deletions and rearrangements, including the amplification of specific plasmid segments to increase the number of copies of specific drug resistance genes (**Fig. 2**). Analysis of the molecular properties of certain transmissible, multiple-drug-resistance plasmids have indicated that these plasmids have evolved by incorporation of genes mediating genetic transfer and various drug resistance genes into a small plasmid containing the replication functions. There are now many examples of the acquisition of specific functions by plasmids by transposition. This ensures the continued evolution of extrachromosomal genetic systems which provides extraordinary versatility and adaptibility to host cells.

Genetic engineering. Plasmids have played a seminal role in the spectacular advances in the area of genetic engineering. Individual genes can be inserted into specific sites on plasmids in cell cultures and the recombinant plasmid thus formed introduced into a living cell by the process of bacterial transformation.

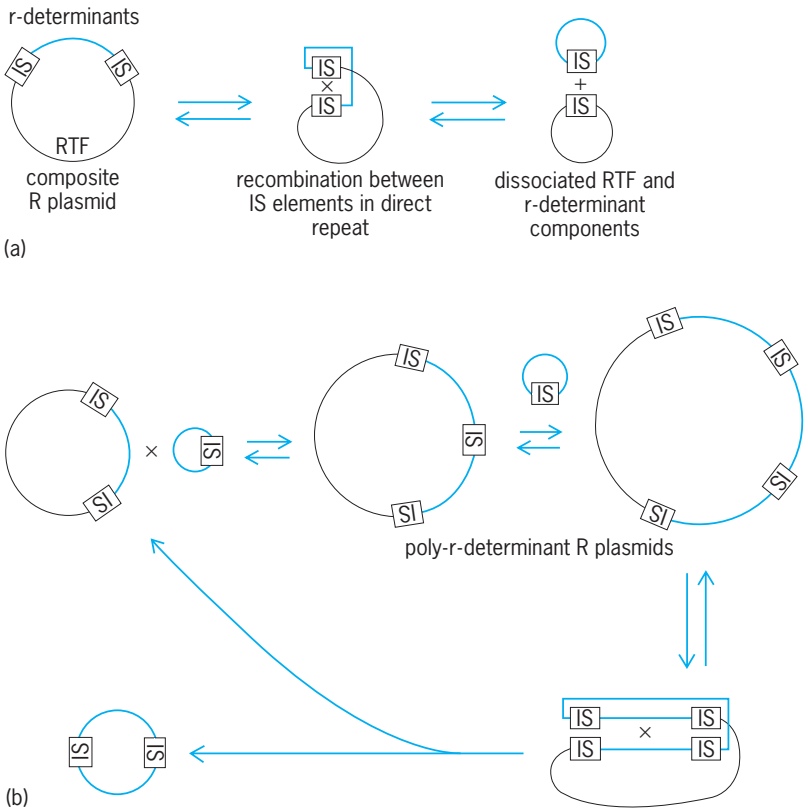

Fig. 2. Plasmid deletions and rearrangements mediated by insertion sequence (IS) elements. (*a*) The resistance (r-determinants) component of the plasmid is flanked by direct repeats of an IS element. By genetic recombination between the homologous IS elements, the composite plasmid can dissociate into the r-determinants and the resistance transfer factor (RTF) components. (*b*) Extra copies of the r-determinants can be incorporated into composite R plasmids by using any region of homology to form the multiple, tandem repeats. These poly-r-determinant R plasmids confer higher levels of antibiotic resistance. These molecules can also delete the extra copies of r-determinants to reform the basic composite structure.

Usually both the plasmid vector and the gene to be cloned have been cleaved by a restriction endonuclease which recognizes a specific DNA base sequence of four or six bases. The breaks introduced by most restriction endonucleases in the individual strands of duplex DNA are staggered with the result that the restriction fragments which are formed contain four-base protruding ends of complementary base sequence. These cohesive ends provide the specificity by which the fragment to be cloned is inserted into a single, specific site on the plasmid vector. The single-strand interruptions can then be sealed by the action of the enzyme ligase. Some restriction endonucleases cleave polynucleotide bonds of bases directly opposite to each other in duplex DNA; such "blunt" ends on restriction fragments can also be joined by ligase to form recombinant DNA molecules, although the ligation step is not as efficient. Plasmid and viral vectors have been developed in both prokaryotes and eukaryotes which make possible a high level of expression of cloned genes. In some cases the vector itself contains the regulatory elements controlling the expression of genes which are not normally expressed in the host cell so that a high level of expression of heterologous genes can be obtained. *See* GENETIC ENGINEERING.

Robert H. Rownd

Bibliography. J. Grinsted and P. M. Bennett (eds.), *Plasmid Technology*, 2d ed., 1990; K. G. Hardy, *Bacterial Plasmids*, 1984; K. G. Hardy, *Plasmids*, 1994; R. W. Old and S. B. Primrose, *Principles of Gene Manipulation*, 5th ed., 1994; R. H. Round (ed.), *Advances in Plasmid Molecular Biology*, 1988.

Plasmin

A proteolytic enzyme that can digest many proteins through the process of hydrolysis. Plasmin is found in plasma in the form of plasminogen, which is an inert precursor or zymogen; one site of synthesis is the liver. Plasma also contains inhibitors that limit either the transformation of plasminogen to plasmin (plasminogen activator inhibitors) or the action of plasmin (principally α_2-antiplasmin and α_2-macroglobulin).

Activation of plasminogen. In the laboratory, plasminogen can be converted to plasmin by extracts of bacteria, notably streptokinase, a protein that is formed by certain beta-hemolytic streptococci. Plasminogen can also be activated by factors that are present in the body. Tissue plasminogen activators, which are proteolytic enzymes derived from a variety of normal and neoplastic cells, can convert plasminogen to plasmin, as can urokinase, an enzyme found in urine and plasma. Exposure to organic solvents, particularly chloroform, will also lead to the formation of plasmin in human plasma, which is rich in plasminogen.

Plasma itself has the potentiality of activating plasminogen, because it contains intrinsic activators. A tissue plasminogen activator is produced in blood vessel walls, from which it is released into the bloodstream following vascular injury. Further, the process of blood coagulation may foster the activation of plasminogen. Activated Hageman factor (factor XII), activated plasma thromboplastin antecedent (factor XI), plasma kallikrein (an enzyme derived from plasma prekallikrein), and activated protein C, all evolving during blood clotting, bring about reactions leading to the formation of plasmin. Moreover, the presence of a fibrin clot enhances activation of plasminogen by many agents.

Physiologic actions of plasmin. Plasmin can act on many protein substrates. It liquefies coagulated blood by digesting fibrin, the insoluble meshwork of the clot; it also digests fibrinogen, the precursor of fibrin, rendering it incoagulable. The digestion products of fibrinogen and fibrin are anticoagulant substances that further interfere with the clotting process. Plasmin also inactivates several other protein procoagulant factors, particularly proaccelerin (factor V) and antihemophilic factor (factor VIII). Other substrates attacked by plasmin include gamma globulin and several protein or polypeptide hormones found in plasma.

The action of plasmin may be related to certain body defenses against injury. Its capacity to digest fibrin is thought to bring about liquefaction of intravascular clots (thrombi) that might otherwise cause damage by reducing blood flow to tissues. It can convert the first component of complement, a group of plasma proteins important in immune reactions, to the proteolytic enzyme C1 esterase ($C\overline{1}$). It can also liberate polypeptide kinins (for example, bradykinin) from plasma precursors. The kinins can reproduce such elements of the inflammatory process as pain, dilatation and increased permeability of small blood vessels, and the migration of leukocytes. In addition to these physiologic actions, plasmin hydrolyzes casein, gelatin, denatured hemoglobin, and certain synthetic esters of arginine and lysine. *See* COMPLEMENT; IMMUNOLOGY; INFLAMMATION.

In the test tube, clots formed from plasma or its euglobulin (water-insoluble) fraction dissolve spontaneously, through the action of plasmin. Ordinarily such fibrinolysis is slow, but if the plasma is prepared from the blood of individuals who have been under stress or have severe chronic liver disease, fibrinolysis may be much more rapid. Similarly, the administration of such substances as epinephrine (adrenalin) or des-amino-D-arginine vasopressin (DDAVP) may enhance fibrinolysis. These test-tube phenomena have not been related conclusively to events in the body, but fibrinolysis, although rarely significant, apparently does occur in the body, either as a primary event or secondary to widespread coagulation within the patient's blood vessels (disseminated intravascular coagulation) and may then be responsible for or contribute to hemorrhage. This fibrinolytic purpura has been treated with agents inhibiting either the activation of plasminogen, such as epsilon-aminocaproic acid or tranexamic acid, or the action of plasmin, such as aprotinin, a polypeptide that is derived from various tissues. A deficiency of α_2-plasmin inhibitor has been associated with a hemophilialike disorder, presumably because clots that might stop bleeding readily dissolve. Conversely, a tendency to form intravascular clots (thrombi) has been related to deficiencies or abnormalities of the proteins that foster fibrinolysis.

Fibrinolytic agents, principally tissue plasminogen activator, urokinase, streptokinase, and their derivatives, have been used therapeutically to dissolve thrombi and thereby restore blood flow in the affected vessels. Their chief application is in treating myocardial infarction, dissolving clots that occlude the coronary arteries that supply blood to cardiac muscle and thereby limiting damage by restoring circulation through these vessels. They have also been used to dissolve thrombi in other blood vessels and to dissolve pulmonary emboli, that is, thrombi that have broken off from peripheral veins and traveled to the pulmonary arteries so as to occlude these vessels and thereby damage pulmonary tissue (pulmonary infarction). *See* BLOOD; ENZYME; FIBRINOGEN; INFARCTION; RESPIRATORY SYSTEM DISORDERS; UROKINASE.

Oscar D. Ratnoff

Bibliography. V. J. Marder and S. Sherry, Thrombolytic therapy: Current status, *N. Engl. J. Med.,* 318:1512–1520, 1585–1600, 1988; O. D. Ratnoff and C. H. Forbes (eds.), *Disorders of Hemostasis,* 3d ed., 1996.

Plasmodiophorida

Protozoa composing an order of Mycetozoia, endoparasitic in plants and primarily causing, for example, club root of cabbage, powdery scab of potatoes, and galls in *Ruppia*. Underground portions of host plants are invaded by young parasites, often a flagellate which sometimes arises from a freshly excysted ameba. Becoming intracellular, the young parasite grows, apparently without phagotrophy, and develops into a plasmodium. At maturity, the plasmodium produces uninucleate cysts (spores) which are released upon degeneration of the damaged cell (see **illus.**). Under favorable conditions,

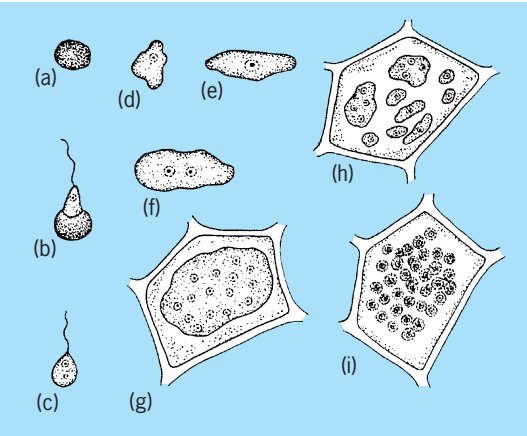

Typical life cycle of Plasmodiophorida. (*a*) Uninucleate cyst (spore). (*b*) Excystment. (*c*) Flagellate stage. (*d*) Ameboid stage. (*e*) Ameboid stage supposedly formed by syngamy. (*f*) Binucleate ameboid stage. (*g*) Growing plasmodium inside plant cell. (*h*) Products of plasmotomy. (*i*) Uninucleate stages formed from plasmodium (or from products of plasmotomy), prior to encystment. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)

the released spores hatch into uninucleate stages which become the infective forms. In the invaded area, the host's tissue commonly undergoes hypertrophy to form a gall. There are a number of genera, including *Plasmodiophora, Spongospora*, and *Sorosphaera*. *See* PROTOZOA; RHIZOPODEA; SARCODINA; SARCOMASTIGOPHORA. Richard P. Hall

Plasmodiophoromycota

Obligate endoparasites (parasites restricted to living inside a host) of flowering plants, green and yellow-green algae, seaweeds, diatoms, water molds, and ciliates. The Plasmodiophoromycota (plasmodiophorids) are prevalent in both soil and aquatic habitats, where their biflagellate zoospores (independently motile spores) disperse the pathogen and infect hosts. A plasmodium develops within the host cell, often causing cell enlargement and proliferation, resulting in hypertrophy (abnormal enlargement) of host tissue. These organisms include devastating pathogens of food crops, such as *Plasmodiophora brassicae*, which causes clubroot (a plant disease involving malformed and abnormally enlarged roots) of crucifers (for example, mustard, cabbage, radishes, turnips, cauliflower, and broccoli). Plasmodiophoromycota were once thought to be kin to fungi, but recent molecular systematic studies have revealed that they are actually protozoa, members of the phylum Cercozoa and class Phytomyxea. *See* PARASITOLOGY; PROTOZOA.

Life cycle. The life history of plasmodiophorids is complex and includes (1) a flagellated motile stage, the zoospore; (2) an infectious stage, the encysted zoospore (cyst); (3) an unwalled assimilative and feeding stage, the multinucleate plasmodium; (4) a zoospore proliferation stage, the zoosporangium; and (5) a resistant stage, the resting spore. Zoosporangia and resting spores, cleaved from their respective plasmodia, cluster together with a characteristic organization in sori (spore balls).

Uninucleate resting spores that survive adverse growth conditions in host tissue debris or directly in the soil germinate with primary zoospores when growth conditions are conducive. Zoospores are uninucleate and microscopic, 3–5 micrometers in diameter. They bear two apically inserted smooth-surfaced flagella. Flagellar basal bodies are especially long, with globules scattered in their core, and are positioned at an angle to each other. Two microtubular roots are associated with each basal body and extend underneath the cell membrane. The anteriorly directed flagellum does most of the work in zoospore motility, with the other, longer, posteriorly directed flagellum essentially serving as a rudder. Plasmodiophorid zoospores have a distinct method for penetration into the host. The zoospore swims to a suitable host, loses its flagella, crawls along the host surface, produces a wall around itself, and adheres to the host. Inside the cyst, a new structure shaped like a bullet (the stachel) forms within a channel (the rohr). A vacuole in the cyst expands and pushes the stachel into the host cell wall, piercing it. The parasite's protoplast is then injected into the host cytoplasm through the pore. *See* CILIA AND FLAGELLA.

The protoplast enlarges into a multinucleate plasmodium (**Fig. 1***a*) as it absorbs nutrients from the host. There have been several reports that plasmodia are phagotrophic, ingesting portions of host cytoplasm. Typically the infected host cell enlarges, resulting in swelling of host tissue, as in the distorted club-shaped roots of cabbage infected with *P. brassicae*.

The primary plasmodium then segments (Fig. 1*a*) into thin-walled zoosporangia, which release secondary zoospores. The site of cell and nuclear fusion (sexual reproduction) is controversial, but fusion between secondary zoospores has been reported. Secondary zoospores reinfect host cells with additional zoosporangial plasmodia or with secondary plasmodia (the resting spore producing plasmodia). After secondary plasmodial expansion, the plasmodium cleaves into uninucleate masses with thick chitinous cell walls, the resting spores (Fig. 1*b*). Transmission electron microscopy has revealed synaptonemal complexes (ribbonlike protein structures

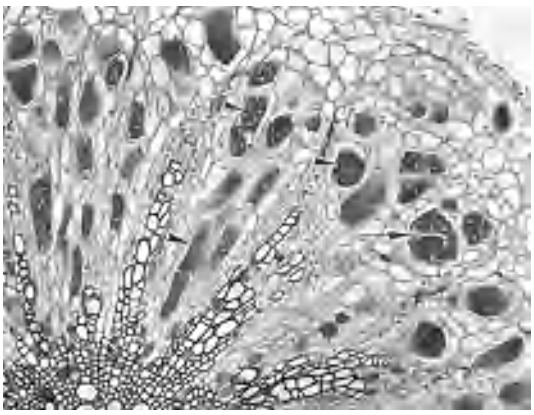

(a)

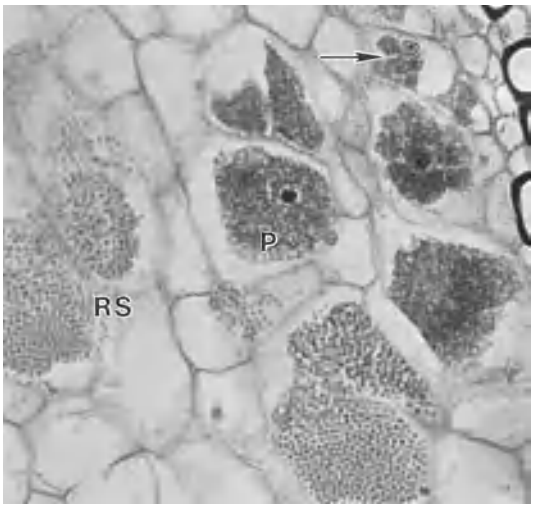

(b)

Fig. 1. Clubroot of cabbage. (*a*) Cross section of root infected with *Plasmodiophora brassicae*. The thick-walled, hollow cells radiating out like fingers are the water transport system of the host. Primary plasmodia (arrowheads) are scattered in cells of the root cortex. Notice the primary plasmodium (~50 μm in diameter) in an enlarged host cell beginning to segment (arrow) into masses that will become thin-walled zoosporangia. (*b*) More detailed cross section of infected root. The secondary plasmodium (P) [~50 μm in diameter] is beginning to segment into uninucleate masses (arrow) that will become resting spores. Characteristic of *Plasmodiophora*, resting spores (RS) are scattered in the host cell.

that extend the length of synapsing chromosomes and aid in exchange pairing) in resting-spore plasmodia, evidence that reduction division (meiosis) occurs in the secondary plasmodium. The lesions of powdery scab of potatoes and the deformed tissue of clubroot of crucifers are often invaded secondarily by bacteria, rotting host tissue and releasing plasmodiophorid resting spores. *See* PLANT PATHOLOGY.

Taxonomy. There are approximately 12 currently recognized genera, but exploration of new habitats is leading to discovery of new taxa. Genera are differentiated based on the organization of their resting spores, especially as spore balls. For example, *Plasmodiophora* resting spores lie free in host cells (Fig. 1*b*). *Spongospora* forms spongelike spore balls, traversed by pockets and channels (**Fig. 2**), and is distinguished from *Sorosphaera*, which forms hollow spore balls. Typically infections result in hypertrophy of host tissue, and hypertrophy in clubroot of crucifers is believed to be caused by the rerouting of the plant hormone precursor, glucobrassicin, into auxin (any organic compound promoting plant growth). However, some infections, as with *Ligniera*, there is little hypertrophy.

Several characters distinguish plasmodiophorids: (1) cruciform nuclear division, where the persistent nucleolus elongates toward the mitotic poles as chromosomes align along the equatorial plate, giving a crosslike configuration; (2) centrioles positioned end to end during interphase; (3) multiple globules in the core of flagellar basal bodies; and (4) relatively simple microtubular root systems in zoospores. *See* PLASMODIOPHORIDA.

Importance. Plasmodiophorids are economically important as pathogens of crop plants. For example, *Spongospora subterranea* causes powdery scab of potato, as well as crookneck disease of watercress. On potatoes, thick corky lesions become filled with thousands of spore balls containing resting spores. This disease has become more prevalent as a greater range of nonresistant potato varieties is produced. The economic loss to this disease has also increased with consumers' demands for high-quality potatoes. *Plasmodiophora brassicae* causes clubroot of crucifers and distorts root tissue, resulting in galls (abnormal growths) on twisted roots that cannot transport nutrients and water effectively to the aboveground portion of plants. Diseased crucifers, consequently, become stunted, yellowed, and wilted. *See* AGRICULTURAL SCIENCE (PLANT); AGRICULTURAL SOIL AND CROP PRACTICES.

In addition to being the causal agents of diseases, plasmodiophorids can serve as vectors (agents capable of transferring pathogens from one organism to another) for plant viruses. *Spongospora subterranea* zoospores transmit potato mop top virus. *Polymyxa graminis* is the vector for wheat spindle streak mosaic virus and barley yellow mosaic virus. *Polymyxa betae* is the vector for beet necrotic yellow vein virus. *See* PLANT VIRUSES AND VIROIDS.

Fig. 2. Powdery scab of potato. Longitudinal section through the outer portion of a potato infected with *Spongospora subterranea*. Notice the spongelike spore balls (sori) traversed by irregular channels (arrow). Each resting spore is ~4 μm in diameter. The spore balls are 45–80 μm in diameter.

The future. Plasmodiophorid diseases are difficult to control because of the longevity of their resting spores in the soil. Most chemicals used to control fungal diseases are not effective against plasmodiophorid, protozoan diseases of plants. Use of resistant plant varieties, long-cycle crop rotation, and sanitation are the three most effective controls of plasmodiophorid diseases.

The impact of plasmodiophorid parasites of algae and oomycetes on environmental balance is poorly studied. Modern molecular environmental studies are revealing new plasmodiophorids and their near relatives in newly explored habitats, such as oceanic thermal vents. Clearly, these modern methods will expand our understanding of the diversity of plasmodiophorids and of their role in the environment.

Martha J. Powell

Bibliography. C. J. Alexopoulos, C. W. Mims, and M. Blackwell, *Introductory Mycology*, 4th ed., 1996; D. J. S. Barr, The phylogenetic and taxonomic implications of flagellar rootlet morphology among zoosporic fungi, *BioSystems*, 14:359–370, 1981; J. P. Braselton, *Plasmodiophoromycota, The Mycota VII, Part A: Systematics and Evolution*, pp. 81–91, 2001; D. P. Dylewski, Phylum Plasmodiophoromycetes, *Handbook of Protoctista*, pp. 399–416, 1990; J. S. Karling, *The Plasmodiophorales*, 2d ed., 1968.

Plasmon

The quanta of waves produced by collective effects of large numbers of electrons in matter when the electrons are disturbed from equilibrium. Plasmon excitations are easily detected in metals because they have a high density of electrons that are free to move. The result of plasmon stimulation by energetic electrons is seen in **Fig. 1**. The graph shows the probability of energy losses by fast electrons transmitted through a thin aluminum foil. The number of detected electrons in a beam is plotted against the energy loss during transit through the foil. Each energy-loss peak corresponds to excitation of one or more plasmons. Within experimental error, the peaks occur at integral multiples of a fundamental loss quantum.

The name plasmon is derived from conventional plasma theory. A plasma is a state of matter in which the atoms are ionized and the electrons are free to move. At low particle densities this means an ionized gas, or classical plasma. In a metal the conduction electrons are also free to move, and there are many analogies to the classical low-density plasma. Similar to sound waves in a low-density gas of charged particles, the electron gas in a metal exhibits plasmon phenomena, that is, electron density waves. In both types of physical plasma, the frequency of plasma-wave oscillation is determined by the electronic density. Because of the high electron density in a metal (about 10^{23} electrons/cm^3), the frequencies are typically in the ultraviolet part of the electromagnetic spectrum. Quantum mechanics dictates that

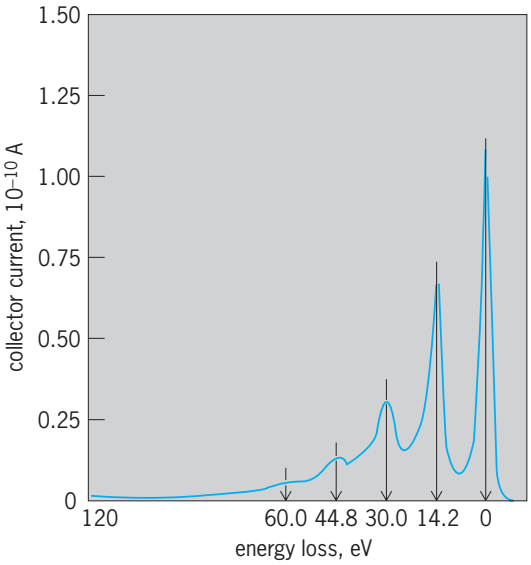

Fig. 1. Number of detected electrons in a beam versus their energy loss during transit through a thin aluminum foil. (Number of electrons is expressed as a current; 10^{-10} A = 6.7 × 10^8 electrons per second.) Peaks at approximate multiples of 14.2 eV correspond to energy donated to plasmons in the aluminum. (*After T. L. Ferrell, T. A. Callcott, and R. J. Warmack, Plasmons and surfaces, Amer. Sci., 73:344–353, 1985*)

the energy of a plasmon is quantized, and one quantum of energy of a plasmon is given by its frequency multiplied by Planck's constant. *See* FREE-ELECTRON THEORY OF METALS; PLASMA (PHYSICS); QUANTUM MECHANICS.

Bulk and surface plasmons. Metallic objects support two kinds of plasmons: bulk and surface plasmons. Bulk plasmons are longitudinal density waves (like sound waves) that propagate through the volume of a metal. Electron density waves can also propagate at the interface between a metal substrate and a dielectric (surface plasmons). This fact can be easily derived from the electromagnetic theory developed by James Clerk Maxwell. Such electronic surface waves are strongly coupled to the electromagnetic field (**Fig. 2**). They are officially termed

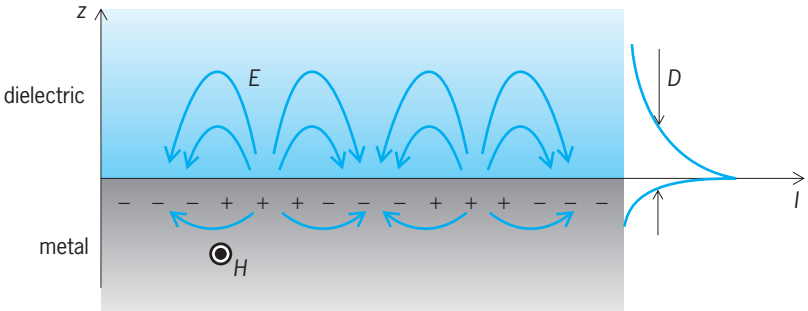

Fig. 2. Surface plasmon-polariton propagating at a metal-dielectric interface. Both the magnetic (*H*) and electric (*E*) field components of this wave are shown. The coupling of the electromagnetic fields to the electrons in the metal is also visualized by the charge density waves near the metal surface that creates a region of slightly negative charge (where electrons bunch up) and positive charge (where electrons have moved away). The electromagnetic field intensity, *I*, of a surface plasmon is highest at the surface and decays exponentially away from the interface. The diameter of the beam (optical mode), *D*, can be significantly smaller than the wavelength of light.

surface plasmon-polaritons, but the term surface plasmon is more commonly used by many scientists. The charge fluctuations are extremely localized in the z direction (perpendicular to the interface) within the Thomas-Fermi screening length of about 0.1 nm. The corresponding electromagnetic wave is transverse magnetic (TM) in nature and has a beam (mode) size, D, that is not limited by diffraction in the direction normal to the metal surface. The field intensity, I, of such a wave has a maximum at the surface and decays exponentially away from the interface. *See* ELECTROMAGNETIC WAVE TRANSMISSION; MAXWELL'S EQUATIONS.

History of plasmon research. Well before scientists set out to study the unique optical properties of metal nanostructures, these optical properties were employed by artists to generate vibrant colors in glass artifacts and to stain church windows. A famous example is the Lycurgus cup dating back to the Byzantine Empire (A.D. 4th century). Scientific studies in which surface plasmons were observed date back to the beginning of the twentieth century. In 1902 Robert W. Wood observed unexplained features in optical reflection measurements on metallic gratings. In 1904 Maxwell Garnett described the bright colors observed in metal-doped glasses using the newly developed Drude theory of metals, and the electromagnetic properties of small spheres as derived by Lord Rayleigh. In an effort to develop further understanding, in 1908 Gustav Mie developed his now widely used theory of light scattering by spherical particles. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

In 1956 David Pines theoretically described the characteristic energy losses experienced by fast electrons traveling through metals, and attributed these losses to collective oscillations of free electrons in the metal. In analogy to earlier work on plasma oscillations in gas discharges, he called these oscillations plasmons. In 1957 Rufus Ritchie studied electron energy losses in thin films and showed that plasmon modes can exist near the surfaces of metals. This study represents the first theoretical description of surface plasmons. In 1968 Ritchie and coworkers described the anomalous behavior of metal gratings in terms of surface plasmon resonances excited on the gratings. In the same year Andreas Otto as well as Erich Kretschmann and Heinz Raether presented methods for the optical excitation of surface plasmons on metal films, making experiments on surface plasmons easily accessible to many researchers.

In the late 1970s and 1980s, surface plasmons at optical frequencies attracted much attention because the high field intensity near the interface could be exploited in nonlinear optical experiments such as surface-enhanced Raman scattering, second-harmonic generation, and surface-enhanced fluorescence. This has been an active area of research for many years, and the physics of surface plasmon-polaritons on two-dimensional (2D) films is now fairly well understood. *See* NONLINEAR OPTICS.

Since the late 1990s, notable developments have taken place in the field of plasmon optics. These developments have come about primarily because of the ability to fabricate nanoscale metallic structures in a controlled fashion. Thomas Ebbesen showed that optical transmission through a subwavelength aperture in a metal film is strongly enhanced when the incident light resonantly couples to surface plasmons. John Pendry speculated that a thin silver film might act as a "perfect" lens capable of producing images with a spatial resolution beyond the diffraction limit. Furthermore, theory and experiments demonstrated that metal nanowires, ordered arrays of metal nanoparticles, and grooves in metal films can serve as light guides that can manipulate the flow of light at the nanoscale. These findings are expected to have a major impact on the semiconductor industry and on the fields of biology, medicine, energy, and the environment. *See* NANOSTRUCTURE; NANOTECHNOLOGY.

Optical excitation of surface plasmons. From basic electromagnetic theory, it follows that surface plasmon waves have a shorter wavelength than light waves of the same frequency. In the quantum picture, this means that surface plasmons possess a larger momentum than photons of the same energy. For this reason, free-space light beams in air cannot simply couple to surface plasmon waves and conserve both energy and momentum. Kretschmann realized that photons in a high-refractive-index material, such as glass, possess a larger momentum and can be used to excite surface plasmons at a metal-air interface. He showed that surface plasmons can be excited by directing light through a glass prism at a thin metal film deposited on the prism surface. He found that there was one specific, resonant angle for which significant energy from the light beam could be transferred to surface plasmons. At this angle the intensity of the reflected beam was significantly reduced, consistent with the fundamental law of energy conservation. From the early studies by Wood, it was also found that surface plasmons can be excited by means of a metal grating structure. When gratings are used, the momentum mismatch between photons and surface plasmons is taken up by the grating.

For the excitation of surface plasmons in nanoscale particles, a grating or a prism is not needed. Incident light typically induces an oscillating charge density on the surface of a particle or microstructure, with opposite charges at opposing sides. For particles small compared to the wavelength of the light, this is a purely dipolar charge distribution. The dipolar surface plasmon frequency for a sphere in vacuum is given by the bulk plasmon frequency divided by $\sqrt{3}$ (assuming a simple model for the electronic properties of the metal). The resonance frequencies for most metals are in the ultraviolet, visible, or near-infrared part of the electromagnetic spectrum. Metal nanoparticles suspended in a liquid or embedded in a glass film strongly absorb light at their surface plasmon frequency and therefore exhibit vibrant colors. The surface plasmon resonance frequency can be easily tuned by changing the shape of the particle

or the refractive index of the medium it is embedded in. Peter Nordlander and Naomi Halas provided a very intuitive model that can be used to predict the optical response of complex metallic nanostructures and systems. Called the plasmon hybridization model, it shows intriguing similarities to the well-established theories on molecular orbital formation from atomic orbitals. This model can predict the location of plasmon resonances in complex systems based on knowledge of the resonant behavior of elementary building blocks. It thus provides a powerful tool for optical engineers in the design of functional plasmonic nanostructures. *See* COLLOID; COLLOIDAL CRYSTALS; DIPOLE.

Applications of surface plasmons. One of the most well-established applications of surface plasmons is biochemical sensors that can detect the presence of specific biochemical compounds. Selectivity can be obtained by the use of compounds or molecular impressions that bind only the specific target compound. By coating a thin metal film on a prism with antibodies, and tracking the shift in the surface plasmon excitation angle as antigens from a solution bind to the antibodies, the concentration of antigens can be determined. Similarly, since the surface plasmon excitation angle is sensitive to the index of refraction of the overcoating, it is possible to distinguish monomers from dimers just by exposing the thin film to the solution containing both, the monomer form having a different index from that of the dimer. The indices must be known from other measurements, and the thin film must be protected from deleterious effects of the solution. *See* ANTIBODY; ANTIGEN; BIOELECTRONICS; TRANSDUCER.

An important contribution of surface plasmons to analytical chemistry and biology is in Raman scattering. A portion of visible light incident on a molecule is frequency-shifted in accordance with the vibrational and rotational spectrum of the molecule. This constitutes the Raman signal that is characteristic of the scattering molecule. When a molecule is sitting on the surface of a metallic nanostructure, the fields can be substantially enhanced due to the resonant excitation of surface plasmons. This in turn leads to a surface-enhanced Raman spectroscopy (SERS) signal. This signal can be many orders of magnitude stronger than the Raman signal obtained without the presence of the metal. This effect has enabled the successful implementation of Raman spectroscopy on trace amounts of chemical compounds and even single molecules. *See* RAMAN EFFECT.

Promising developments in biology and medicine followed the demonstration by Halas and Jennifer West in 2003 that the resonant properties of gold nanoshells could enable fast whole-blood immunoassays. It has been speculated that nanoshells could play a role in cancer treatments.

There is significant interest in the use of surface plasmon waves to transport information on a chip with nanoscale plasmon waveguides. Currently, two of the most daunting problems preventing significant increases in processor speed are thermal and resistance-capacitance (RC) delay time issues associated with electronic interconnection. Optical interconnects, on the other hand, possess an almost unimaginably large data-carrying capacity and may offer interesting solutions for circumventing these problems. Optical alternatives may be particularly attractive for chips with more distributed architectures in which a multitude of fast electronic computing units (cores) need to be connected by high-speed links. Unfortunately, dielectric photonic devices are limited in size by the fundamental laws of diffraction to about half a wavelength of light, and tend to be at least 1 or 2 orders of magnitude larger than their nanoscale electronic counterparts. This obvious size mismatch between electronic and photonic components presents a major challenge for interfacing these technologies. It thus appears that further progress will require the development of a radically new chip-scale device technology that can facilitate information transport between nanoscale devices at optical frequencies and bridge the gap between the world of nanoscale electronics and microscale photonics. Metallic or "plasmonic" interconnects may provide new solutions. The device technology based on plasmonic interconnects and components has been termed plasmonics. An increased synergy could be attained by integrating plasmonic, electronic, and conventional photonic devices on the same chip and taking advantage of the strengths of each technology. *See* INTEGRATED OPTICS.

One of the most promising developments in surface plasmon science is the superlens proposed by Pendry. A superlens is an optical element capable of imaging objects of subwavelength dimensions. The simplest version of a superlens consists of a thin slab of metal. Superlenses exhibit enormous potential for commercial applications, including ultrahigh-resolution imaging, high-density memory storage devices, and nanolithography.

Surface plasmons also play an important role in the extraordinary optical transmission (EOT) phenomenon observed in thin metallic films perforated with an array of subwavelength holes. Ebbesen demonstrated that significantly more light is transmitted through the tiny holes in the metal film than was previously predicted by classical diffraction theory. The reason is that the hole arrays scatter light into surface plasmon waves. First, these surface plasmons waves on the illuminated side of the metal couple to surface plasmon waves on the other side of the metal through the holes. In a second stage, these surface plasmons get scattered again to light waves propagating away from the metal. The exact transmission spectrum can be tuned by adjusting the period and the symmetry of the hole array. This phenomenon has attracted much attention to the field of plasmonics.

The energy losses imparted by materials on a charged particle beam can be analyzed in much the same way as in the original observations of plasmons. This analysis can be used to form high-resolution chemical maps in transmission electron microscopy. This technique has a number of applications in

materials analysis of inorganic, organic, and biomaterials. *See* ELECTRON MICROSCOPE.

Mark L. Brongersma; Pieter G. Kik

Bibliography. W. L. Barnes, A. Dereux, and T. Ebbesen, Surface plasmon subwavelength optics, *Nature*, 424:824–830, 2003; T. W. Ebbesen et al., Extraordinary optical transmission through subwavelength hole arrays, *Nature*, 391:667–669, 1998; J. B. Pendry, Negative refraction makes a perfect lens. *Phys. Rev. Lett.*, 85:3966–3969, 2000; E. Prodan et al., A hybridization model for the plasmon response of complex nanostructures, *Science*, 302:419–422, 2003; H. Raether, *Surface Plasmons*, 1988; J. Takahara et al., Guiding of a one-dimensional optical beam with nanometer diameter, *Opt. Lett.*, 22:475–477, 1997; R. W. Wood, On a remarkable case of uneven distribution of light in a diffraction grating spectrum, *Phil. Mag.*, 4:396–402, 1902; R. Zia, J. A. Schuller, and M. L. Brongersma, Plasmonics: The next chip-scale technology, *Mater. Today*, 9:20–27, 2006.

Plaster

A plastic mixture of solids and water that sets to a hard, coherent solid and that is used to line the interiors of buildings. A similar material of different composition, used to line the exteriors of buildings, is known is stucco. The term plaster is also used in the industry to designate plaster of paris.

Plaster is usually applied in one or more base (rough or scratch) coats up to $^3/_4$ in. (2 cm) thick, and also in a smooth, white, finish coat about $^1/_{16}$ in. (0.2 cm) thick. The solids in the base coats are hydrated (or slaked) lime, sand, fiber or hair (for bonding), and portland cement (the last may be omitted in some plasters). The finish coat consists of hydrated lime and gypsum plaster (in addition to the water). *See* LIME (INDUSTRY); MORTAR; PLASTER OF PARIS.

J. F. McMahon

Plaster of paris

The hemihydrate of calcium sulfate, composition $CaSO_4 \cdot {}^1/_2H_2O$, made by calcining the mineral gypsum, composition $CaSO_4 \cdot 2H_2O$, at temperatures up to 482°F (250°C). It is used for making plasters, molds, and models. For a description of the casting method of ceramic forming *see* CERAMICS.

When the powdered hemihydrate is mixed with water to form a paste or slurry, the calcining reaction is reversed and a solid mass of interlocking gypsum crystals with moderate strength is formed. Upon setting there is very little (a slight contraction) dimensional change, making the material suitable for accurate molds and models.

The chemical reaction of hydration requires 18.6 lb of water for each 100 lb of plaster of paris; any water in excess of this makes the mixture more fluid and it eventually evaporates, leaving a porous structure. In general, the greater the porosity, the lower the strength of the set plaster. Plaster of paris molds for ceramic casting must have a certain minimum porosity in order to absorb water from the slip; thus, for this application, the amount of water added represents a compromise between the conflicting properties of high strength and high porosity.

Diverse types of plaster, varying in the time taken to set, the amount of water needed to make a pourable slip, and the final hardness, are made for different uses. These characteristics are controlled by the calcination conditions (temperature and pressure) and by additions to the plaster. For example, hydrated calcium sulfate ($CaSO_4 \cdot 2H_2O$) greatly accelerates setting time and therefore the use of utensils contaminated with set plaster is to be avoided. *See* GYPSUM; PLASTER.

J. F. McMahon

Plastic deformation of metals

The permanent change in shape of a metal object as a result of applied or internal forces. This feature permits metals to be formed into pipe, wire, sheet, and so on, with an optimal combination of strength, ductility, and toughness. The onset of noticeable plastic deformation occurs when the applied tensile stress or applied effective stress reaches the tensile yield strength of the material. Continued plastic deformation usually requires a continued increase in the applied stress. The effects of plastic deformation may be removed by heat treatment. *See* MECHANICAL PROPERTIES OF METAL; PLASTICITY.

Mechanisms of plastic deformation in metals are understood in terms of crystal structure and incremental slip between atomic planes. These aspects are useful to understand: (1) regimes of stress and temperature for plastic deformation; (2) types of plastic stress-strain response; (3) reasons why metals exhibit plastic deformation; (4) mechanisms of plastic deformation in metal crystals and polycrystals; (5) strategies to strengthen metals against plastic deformation; (6) effect of strengthening, temperature, and deformation rate on ductility and toughness.

Deformation regimes and plastic stress-strain response. Plastic deformation in metals is observed macroscopically as a departure from linear-elastic stress-strain response. **Figure 1** shows that plastic deformation predominates over other deformation mechanisms when the applied stress is greater than the yield strength. Unlike creep deformation, plastic deformation usually does not increase with time unless the applied stress is also increased. When the applied stress is less than the yield strength, other deformation mechanisms may dominate. *See* CREEP; ELASTICITY; MECHANICAL PROPERTIES OF METAL; STRESS AND STRAIN.

Tensile-engineering stress-strain (σ-ε) curves exhibit several features of plastic deformation (**Fig. 2**). The response depends on single versus polycrystalline structure, alloy content, temperature, strain rate, loading geometry, as well as sample size, shape,

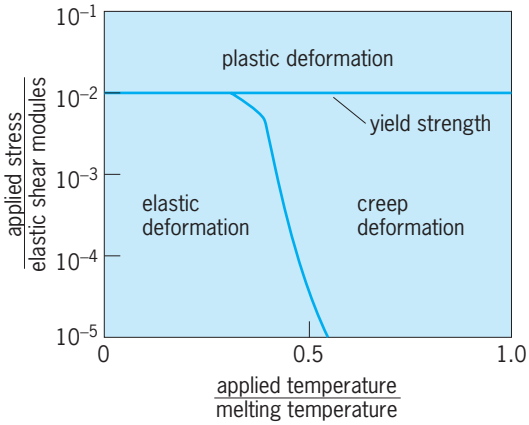

Fig. 1. Deformation mechanism map. (*After M. F. Ashby and D. R. H. Jones, Engineering Materials I: An Introduction to Their Properties and Applications, 2d ed., Butterworth-Heimann, 1996*)

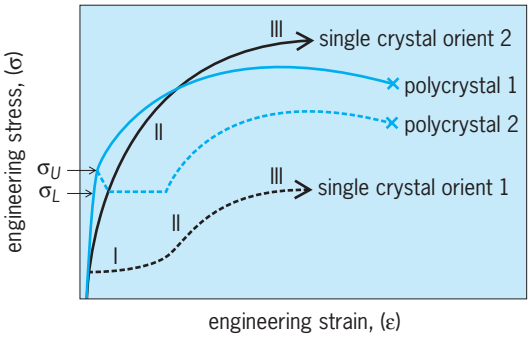

Fig. 2. Stress-strain curves for metals. (*P. M. Anderson, The Ohio State University, 2005*)

and orientation. For polycrystalline metals, departure from linear-elastic response can be gradual (polycrystal 1) or quite abrupt (polycrystal 2). Similarly, single-crystal samples can have an abrupt departure (single crystal orient 1) or a gradual one (single crystal orient 2). In the latter case, three distinct regimes of plastic deformation are noted. *See* SINGLE CRYSTAL.

Reasons for plastic deformation. Metals are unique in that they tend to form crystals of densely packed atoms. This occurs in metals since the valence or outer electrons have a continuous succession of possible energy levels. These electrons create a "sea" of negative charge that pulls positively charged nuclei into densely packed crystal structures such as face-centered cubic (fcc), hexagonal close-packed (hcp), and body-centered cubic (bcc). Metallic glasses are an exception, since they are noncrystalline or amorphous. *See* CRYSTAL STRUCTURE; METALLIC GLASSES; METALS.

Under stress, metal crystals are prone to shear between densely packed planes. **Figure 3*a*** shows two densely packed layers typical of metal crystals with fcc or hcp crystal structures. The upper (lighter) plane is positioned into depressions created by the lower (darker) plane. Plastic deformation occurs by shearing one plane relative to another, as depicted in Fig. 3*b*. Complete shearing by an amount **b** restores dense packing, since the upper plane moves into an equivalent set of depressions. During the process, atomic bonds are broken and made. Densely packed slip planes are preferred in part because the depressions are shallow and allow for "smoother" slip. In many metal crystals, the unit slip **b** is smoothened by a two-step process of partial slip **p1**, followed by partial slip **p2** (Fig. 3*b*). *See* SHEAR.

Experimental evidence confirms that slip produces plastic deformation in metals. **Figure 4** shows numerous slip traces along the surface of a copper-aluminum crystal deformed in tension. Since the length of unit slip **b** is only 2–3 angstroms (1 angstrom = 10^{-10} m), planes must shear repeatedly to produce macroscopically visible slip traces. Formation of slip traces during plastic deformation turns shiny metal surfaces to a dull state. In contrast, ceramic materials are brittle and lack plastic deformation at room temperature. This is due to the lack of closely packed planes and to the prohibitively

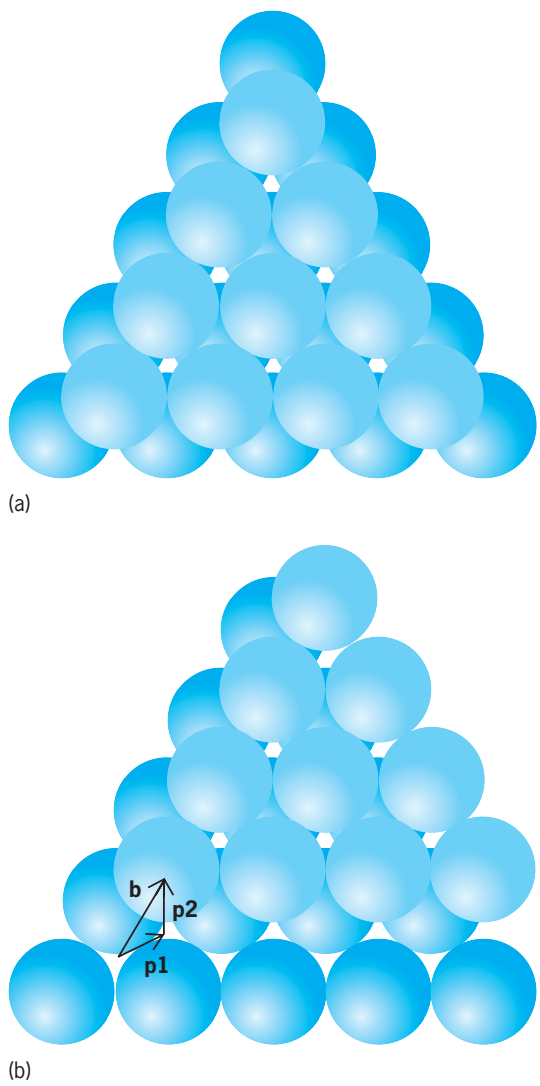

Fig. 3. Pair of closely packed planes in (*a*) initial and (*b*) sheared positions. (*P. M. Anderson, The Ohio State University, 2005*)

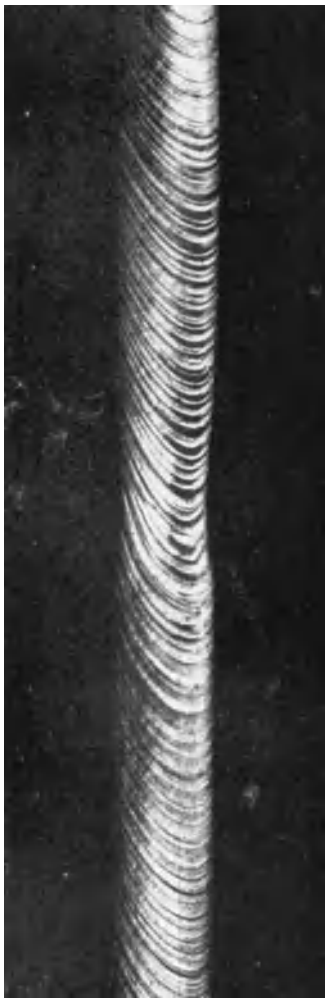

Fig. 4. Slip traces along the surface of a copper-aluminum crystal deformed in tension. (*From C. F. Elam, The Distortion of Metal Crystals, Oxford University Press, London, 1935*)

large energy needed to break bonds during shearing. *See* BRITTLENESS; CERAMICS.

Incremental slip and theoretical strength. Experimental evidence also demonstrates that atomic planes slip incrementally. The line ABCD in **Fig. 5** is known as a dislocation and identifies the boundary between slipped and unslipped portions. Dislocations are lines along which crystal packing is disrupted and atomic bonds are distorted. **Figure 6***a* shows several dislocations as viewed by a transmission electron microscope. Dislocations move under applied load, indicating the advance of slip between two planes. For example, movement of dislocation ABCD to the left (Fig. 5) corresponds to motion of additional atoms from unslipped to slipped positions. In reality, dislocation lines are not as atomically sharp as depicted in Fig. 5. Rather, atoms near dislocations displace to close gaps (for example, along lines BC and CD) and smoothen the transition. These displacements become negligible with increasing distance from the dislocation (Fig. 3). *See* CRYSTAL DEFECTS.

The existence of dislocations is also supported by the small applied tensile stress needed for plastic deformation. The yield strength of many metals varies from 10^{-3} to 10^{-5} times the elastic modulus and agrees well with theoretical estimates of stress to move dislocations along slip planes. In contrast, theoretical estimates of stress to simultaneously shear atomic planes are orders of magnitude larger than measured yield strengths. A convenient analogy is that sliding an entire rubber-backed carpet across a floor can be quite difficult. However, the carpet can be slid easily (and incrementally) by introducing a wrinkle (dislocation) along one edge and pushing the wrinkle to the opposite edge.

Some nanometer-scale crystals have nearly perfect atomic stacking and thus no dislocations. The premise is that dislocation-free crystals will be unable to slip incrementally and therefore approach theoretical strength. Indeed, single-crystal whiskers a few micrometers in diameter exhibit near-theoretical strength, and metal cylinders with diameters $<1\ \mu$m are observed to have extraordinary compressive yield strength. *See* CRYSTAL WHISKERS.

Single crystals and polycrystals. Several features of stress-strain curves (Fig. 2) can be understood qualitatively based on dislocation concepts. In single crystals, plastic stress-strain response is anisotropic, or dependent on crystal orientation. If crystal orientation favors slip on just one crystallographic plane and direction, three stages of deformation are often observed (single crystal orient 1). During stage I, a single slip plane and direction, called the primary slip system, operate. Primary slip dislocations move until arrested by other primary slip dislocations or until portions of dislocations (for example, AB in Fig. 5) move onto different crystallographic planes. The latter process, called cross slip, prevents movement of dislocations. Arrest of dislocations increases deformation stress and thus work hardens the metal. The slope $d\sigma/d\varepsilon$, or work-hardening rate, is small in stage I, and a moderately uniform distribution of interacting primary dislocations is produced. A second or conjugate slip system is activated when the

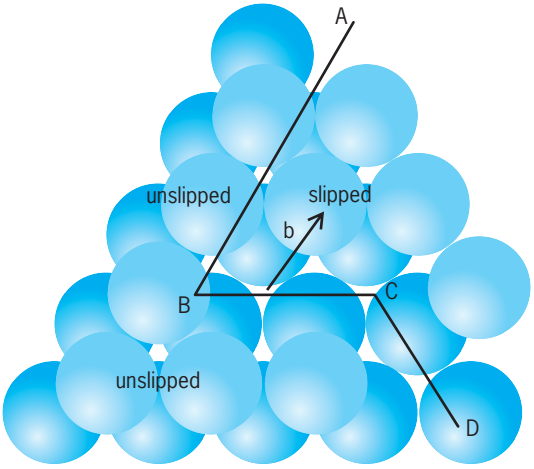

Fig. 5. Incremental slip between closely packed planes. Dislocation line ABCD separates slipped and unslipped regions. (*P. M. Anderson, Ohio State University, 2005*)

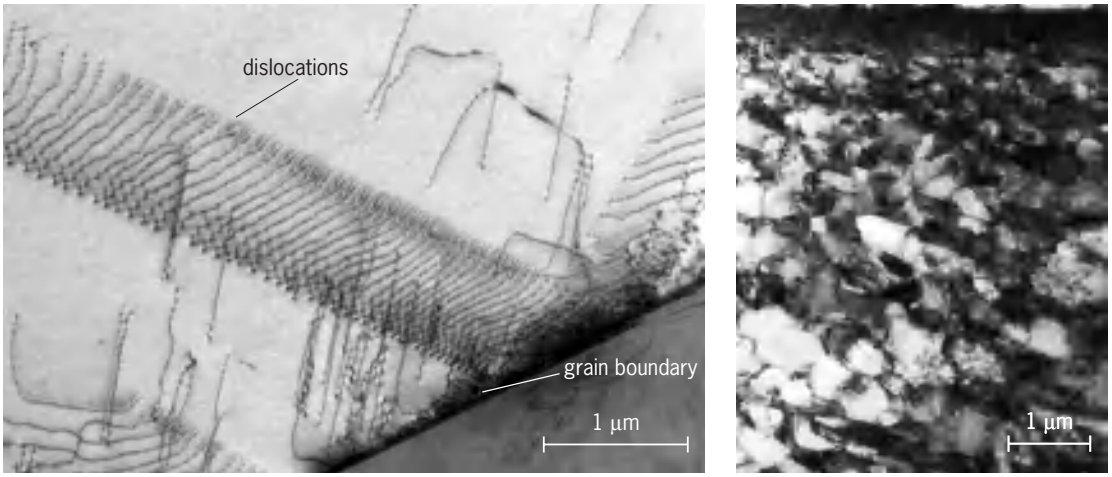

(a) (b)

Fig. 6. (*a*) Arrays of parallel dislocation lines positioned near a grain boundary (lower right) in α-Ti(Al) (*courtesy M. Mills and N. Thirumalai, The Ohio State University, 2005*). (*b*) Dislocation cell structure in oxygen-free high conductivity copper deformed by sliding the top surface against M2 tool steel (*R. Divakar and D. A. Rigney, The Ohio State University, 2005*).

primary slip system becomes sufficiently hardened. Primary and conjugate dislocations interact strongly during stage II so that the work-hardening rate is often large. Cells of entangled dislocations form (Fig. 6*b*), the size of the cells decreasing with continued deformation. During stage III, a minimum dislocation cell size is reached and remains constant with deformation, so that $d\sigma/d\varepsilon$ approaches 0. Single crystals may be oriented to activate two or more primary slip systems initially (single crystal orient 2). In this case, stage I deformation is absent. In special cases of severe deformation ($\varepsilon > 1$ or 2), additional stages are observed during which $d\sigma/d\varepsilon$ increases slightly or remains constant (stage IV) and then decreases and reaches 0 (stage V).

Maximum plastic strength is inversely proportional to the minimum dislocation cell size achieved during stage III. Minimum size is controlled, in part, by escape of dislocations from cell walls via cross slip. Cross slip is easier for metals with large stacking-fault energy, or energy associated with partial slip **p1** of a unit area of atoms (Fig. 3*b*). Consequently, metals with large stacking-fault energy, such as aluminum, exhibit easier cross slip and cannot achieve small stage III cell size or large maximum plastic strength.

Polycrystals tend to have larger initial yield strength (Fig. 2) since they contain internal boundaries between crystals that act as barriers to dislocation motion. Yield is often gradual due to a range of crystal (grain) orientations (polycrystal 1). The first grains to yield have larger values of stress to shear crystallographic planes. In some cases, abrupt yield occurs via breakaway of dislocations from impurity elements such as carbon and nitrogen in steel (polycrystal 2). The upper yield point (σ_U, Fig. 2) is the breakaway stress. Following breakaway, dislocations move at a reduced stress or lower yield point (σ_L, Fig. 2). This phenomenon can be repeated if the sample is unloaded and impurities are allowed to diffuse back to dislocations. *See* GRAIN BOUNDARIES; MECHANICAL PROPERTIES OF METAL.

Stress-strain curves exhibit work hardening due to increasing dislocation density ρ_d during plastic deformation. Density ranges from 10^3 to 10^5 cm of dislocation line length per cubic centimeter of material for carefully produced crystals to greater than 10^{10} cm/cm^3 after large plastic deformation. Dislocations generate new dislocations during plastic deformation and interact with one another via the internal stress they create. The increase in deformation stress is proportional to $Gb(\rho_d)^{1/2}$, where G is elastic shear modulus and b is the magnitude of unit slip vector **b** (Fig. 3*b*). Work hardening of pure metals can increase the deformation stress to more than ten times the initial yield strength.

Tensile engineering stress-strain curves exhibit a negative slope or softening near the end of a test due to a geometric instability called necking (Fig. 2). During testing, elongation in the tensile direction reduces the cross-sectional area, due to the volume-conserving nature of crystallographic slip. A neck, or region of smaller cross-sectional area, arises when softening due to reduction in cross-sectional area cannot be offset by work hardening. The neck becomes more pronounced until plastic failure occurs, usually by linking internal pores along a flat region of the neck center, coupled with angled shearing around the neck perimeter. Engineering stress peaks when true strain equals the work-hardening exponent n. However, true stress properly accounts for area reduction and continues to increase due to work hardening. *See* MECHANICAL PROPERTIES OF METAL.

Strategies to strengthen metals against plastic deformation. Four elementary strategies are used to make dislocation motion difficult and thereby increase the stress needed for plastic deformation. Grain-size strengthening involves the reduction of grain (crystal) size in polycrystals, so that grain boundaries provide more numerous trapping sites for dislocations. The contribution to deformation stress scales

as (grain size)$^{-1/2}$. Solution strengthening involves adding impurity atoms to pure metals (alloying). Impurity atoms interact with dislocations via the local distortion or strain they induce. The contribution to deformation stress scales as (impurity concentration)$^{1/2}$. Precipitate strengthening is also the result of alloying to the extent that a new phase is produced inside the host phase. Precipitates induce strain and provide interfaces that can trap dislocations. The contribution to deformation stress scales as (precipitate spacing)$^{-1}$. Cold work increases dislocation density via deformation at lower temperature and larger deformation rate. The contribution to deformation stress scales as (dislocation density)$^{1/2}$ or as (plastic strain)^n, where n is the work hardening exponent. *See* ALLOY.

Two additional strategies to strengthen metals against plastic deformation are to increase strain rate and decrease temperature. These are effective since dislocation motion involves kinetic processes related to bond breaking and forming during shear. In addition, motion can be assisted or suppressed by the kinetics of impurity or vacancy diffusion to or from dislocations. Deformation stress scales as (strain rate)^m, where m is the strain-rate sensitivity.

Effect of strengthening, temperature, and deformation rate. Cold work and other mechanisms to increase deformation stress often decrease ductility and toughness (**Fig. 7**). Fracture toughness is also reduced, since restricting dislocation motion often reduces the ability to blunt-out sharp cracks and dissipate energy. The effects of cold work are reversed by annealing, which can reduce dislocation density via dislocation annihilation or formation of new crystals, respectively. *See* HEAT TREATMENT (METALLURGY).

Dislocation motion and therefore ductility and toughness are affected by crystal structure. Metals with fcc structure, such as copper, have several orientations of closely packed planes and directions, so that they are easily deformed and require a large energy for failure under impact, regardless of test temperature (**Fig. 8**). In contrast, high-strength metals that use one or more strengthening strategies exhibit small impact energy. Certain metals with hcp structure such as zinc and bcc structure such as iron, chromium, tungsten, and molybdenum display a transition from brittle behavior at lower temperature to ductile behavior at higher temperature. The ductile-to-brittle transition temperature is positioned between the brittle and ductile regimes.

The unique characteristics of hcp and bcc metals stem in part from crystal structure. Hcp metals have only one orientation of closely packed planes and therefore rely on non–densely packed planes for general deformation. Bcc metals have no closely packed planes. Thus, the critical stress for dislocation motion in bcc metals (and some ceramics such as silicon) is typically larger than for fcc materials and is dependent on temperature. An outcome is that ductility, toughness, and deformation stress are also dependent on temperature. However, the transition temperature for each metal is not fixed, but a function of alloy cleanliness, grain size, prior cold work, rate of deformation, and even sample geometry. Thus, increasing the rate of deformation or including notches in the specimen geometry will shift the curve to the right, thereby creating a larger temperature regime over which the metal is brittle.

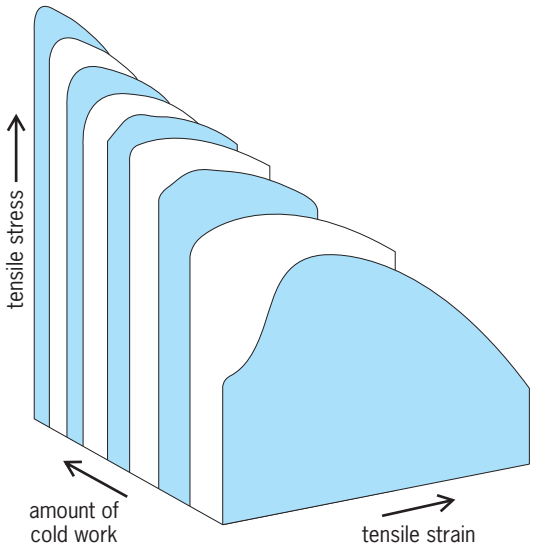

Fig. 7. Effect of cold work on the engineering stress-strain behavior of a low-carbon steel. [*ASM Handbook, vol. 1: Properties and Selection: Irons, Steels, and High Performance Alloys (Carbon and low-alloy steels: Cold-finished steel bars: Mechanical Properties), ASM International, 1990*]

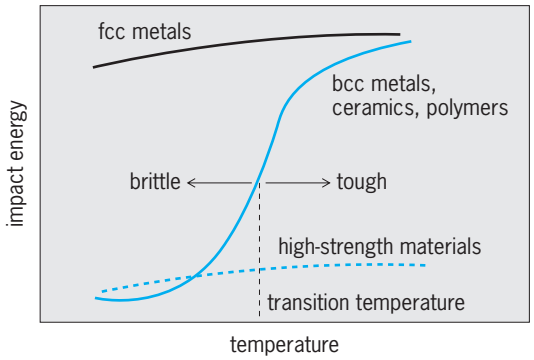

Fig. 8. Variation of impact energy with temperature for different materials. (*After C. R. Barrett, W. D. Nix, and A. S. Tetelman, The Principles of Engineering Materials, Prentice Hall, 1973*)

Peter M. Anderson

Bibliography. M. F. Ashby and D. R. H. Jones, *Engineering Materials 1: An Introduction to Their Properties and Applications*, 2d ed., Butterworth-Heinemann, Oxford, 1996; ASM International, *Atlas of Stress-Strain Curves*, 2d ed., ASM International, 2002; W. D. Callister, *Materials Science and Engineering: An Introduction*, 6th ed., Wiley, 2003; D. Hull and D. J. Bacon, *Introduction to Dislocations*, 4th ed., Butterworth-Heinemann, Oxford, 2001; J. Martin, *Materials for Engineering*, 2d ed., Maney, Leeds, 2002.

Plasticity

The ability of a solid body to permanently change shape (deform) in response to mechanical loads or forces. Deformation characteristics are dependent on the material from which a body is made, as well as the magnitude and type of the imposed forces. In addition to plastic, other types of deformation are possible for solid materials.

One common test for measuring the plastic deformation characteristics of materials is the tensile test, in which a tensile (stretching) load is applied along the axis of a cylindrical specimen, with deformation corresponding to specimen elongation. The load magnitude and degree of deformation are normally converted into meaningful parameters, which include the size and geometry of the body. The load is converted into stress; its units are megapascals (1 MPa = 10^6 newtons per square meter) or pounds per square inch (psi). Likewise, the amount of deformation is converted into strain, which is unitless. The test results are expressed as a plot of stress versus strain. *See* STRESS AND STRAIN.

Typical tensile stress-strain curves are shown for metal alloys and polymeric materials in **Figs. 1** and **2**, respectively. The inset plots show behavior at small strain values. For both materials, the initial regions of the curves are linear and relatively steep. Deformation that occurs within these regions is nonpermanent (nonplastic) or elastic. This means that the body springs back to its original dimensions once the stress is released, or that all of the deformation is recovered. In addition, stress is proportional to strain (Hooke's law), and the slope of this linear segment corresponds to the elastic (Young's) modulus. *See* ELASTICITY; HOOKE'S LAW; YOUNG'S MODULUS.

Plastic (permanent) deformation begins at the point where linearity ceases such that, upon removal of the load, not all deformation is recovered (the body does not assume its original or stress-free dimensions). The onset of plastic deformation is called yielding, and the corresponding stress value is called the yield strength. After yielding, all deformation is plastic and, until fracture, the curves are nonlinear. This behavior is characteristic of many metal alloys and polymeric materials, as shown in Figs. 1 and 2. *See* PLASTIC DEFORMATION OF METAL.

Qualitatively, the plasticity of a material depends on both the yield strength and fracture strain. That is, plasticity increases with a decrease in yield strength (as plastic deformation becomes easier) and an increase in fracture strain (as the degree of plastic deformation increases). Typical metal alloy and polymer yield strengths are indicated on the stress axis of Fig. 1, whereas the strain axis of Fig. 2 compares their fracture strains. Accordingly, the polymer has a higher plasticity since it has a lower yield strength and a greater failure strain. Many polymers are called plastics because of their relatively high degrees of plasticity.

Another aspect of plasticity deals with deformation mechanisms, the atomic or molecular processes that accompany deformation. Plastic deformation involves the net movement of large numbers of atoms or molecules.

The concept of plasticity does not normally relate to ceramic materials such as glasses and metal oxides (for example, aluminum oxide). At room

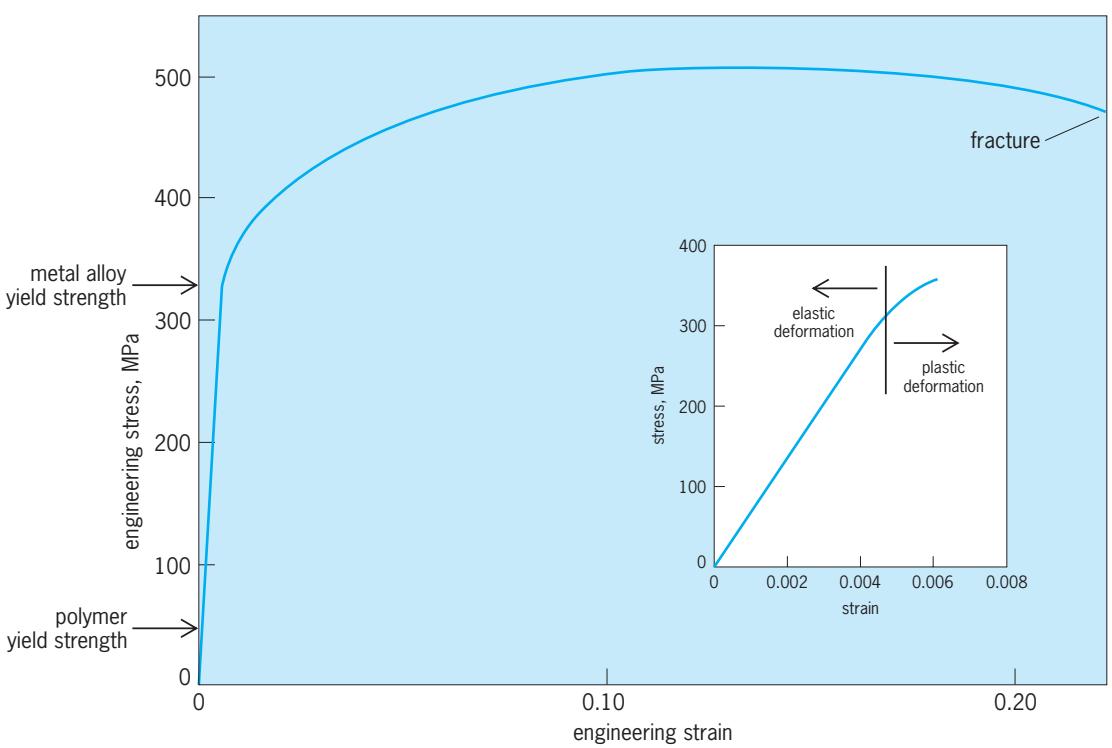

Fig. 1. Plot of stress versus strain to fracture for a typical metal alloy that is deformed using a tensile load.

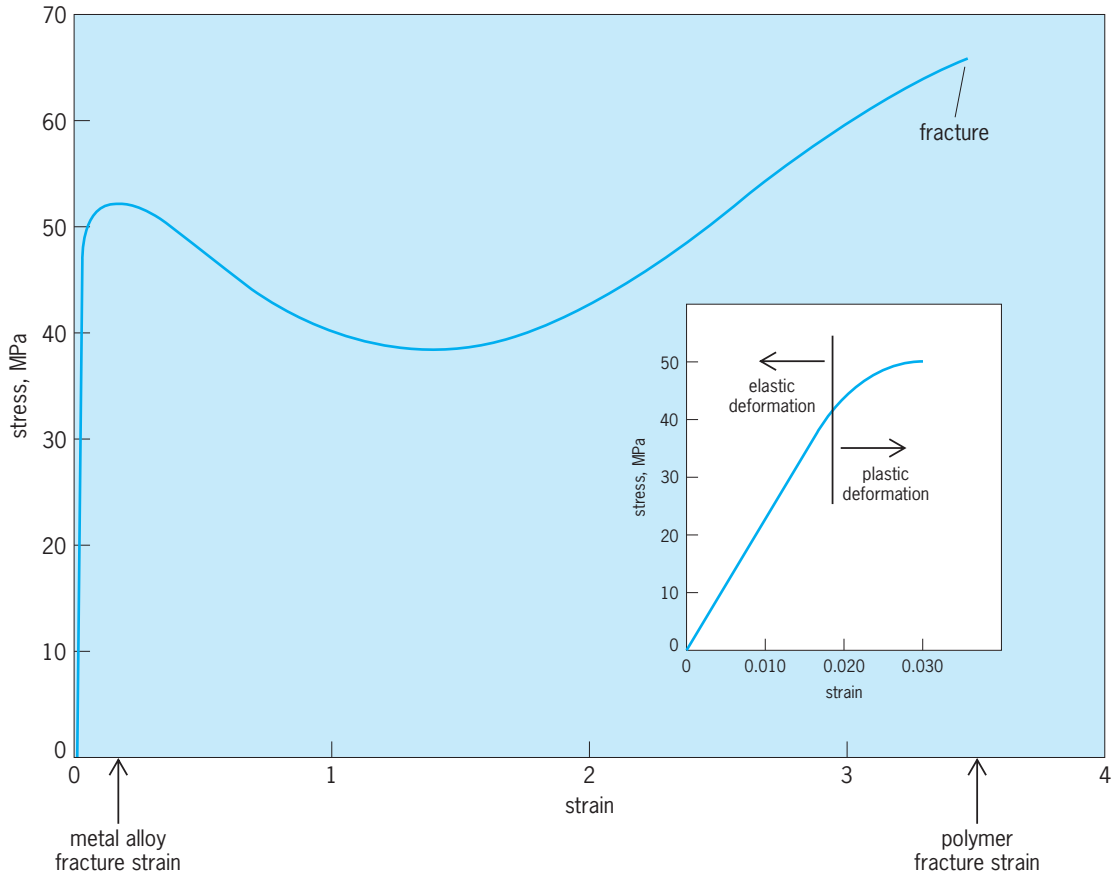

Fig. 2. Plot of stress versus strain to fracture for a typical plastic polymeric material that is deformed using a tensile load.

temperature, virtually all of these materials are brittle; that is, they fracture before deforming plastically. *See* BRITTLENESS. William D. Callister

Bibliography. T. H. Courtney, *Mechanical Behavior of Materials*, 2d ed., McGraw-Hill Higher Education, Burr Ridge, IL, 2000; N. E. Dowling, *Mechanical Behavior of Materials*, 2d ed., Prentice Hall, Upper Saddle River, NJ, 1998.

Plastics processing

The methods and techniques used to convert plastics materials in the form of pellets, granules, powders, sheets, fluids, or preforms into formed shapes or parts. Although the term plastics has been used loosely as a synonym for polymers and resins, plastics generally represent polymeric compounds that are formulated with plasticizers, stabilizers, fillers, and other additives for purposes of processability and performance. After forming, the part may be subjected to a variety of ancillary operations such as welding, adhesive bonding, machining, or surface decorating (painting, metallizing).

As with other construction materials, processing of plastics is only one step in the normal design-to-finished-part sequence. The choice of process is influenced by economic considerations, product specifications, number and size of finished parts, and the complexity of postfinishing operations, as well as the adaptability of the plastics to the process. A variety of processes have been used to produce the final plastic parts (**Fig. 1**). Thermoforming (not shown in Fig. 1) uses about 30% of the extruded sheets and films for primarily packaging applications. Among all plastics processing methods, injection molding ranks first in terms of the number of machines, employment, and total product value.

Injection molding. This process consists of heating and homogenizing plastics granules in an injection molding barrel until they are sufficiently fluid to allow for pressure injection into a relatively cold mold, where they solidify and take the shape of the mold cavity. For thermoplastics, no chemical changes occur within the plastic. Injection molding of thermosetting resins (thermosets) differs primarily in that the cylinder heating is designed to homogenize and preheat the reactive materials, and the mold is heated to complete the chemical crosslinking reaction to form an intractable solid. Solid particles, in the form of pellets or granules, constitute the main feed for injection moldable thermoplastics. The major advantages of the injection molding process are production speed, excellent dimensional tolerance, minimal requirements for postmolding operations, simultaneous multipart molding, and design freedom given by the complex geometries produced. In addition to thermoplastics,

thermosets, and elastomers, injection molding can be extended to materials such as fibers, ceramics, and powder metals, processed with polymeric binders. Among the polymer processing methods, injection molding accounts for 32% by weight of all plastics materials processed.

The early injection molding methods were quite similar to the old metal die casting machines. The pellets were fed through a hopper into one end of the machine, heated and fluidized in the cylinder, and injected by a ram into a mold, where cooling occurred. After solidification, the mold was opened, the part or parts were ejected, and the process was repeated.

Molding technology has since undergone an extensive evolution. Technology is now dominated by the reciprocating screw injection machine (**Fig. 2**), in which the screw acts as a material plasticizer as well as an injection ram. As the screw rotates, it is forced backward by a buildup of viscous plastic at the nozzle end of the cylinder. When an appropriate charge is accumulated, rotation stops and the screw moves forward, acting as an injection ram and forcing the melt into the mold. It remains forward to pack additional material into the mold cavity to compensate for the volumetric shrinkage of the material until it gains enough strength through solidification (thermoplastics) or chemical reaction (thermosets and elastomers) so that the molded part can be subsequently ejected without permanent deformation. When the material at the gate (a small opening through which the material enters the mold cavity from the melt delivery system, such as a sprue and runners in the mold) is frozen, the reciprocating screw starts to rotate and plasticate (soften) the material while it is moving backward, and the next cycle begins. The size of an injection molding machine is rated in one of two ways: by the clamping force (in tons) available to keep the mold closed during injection, or by the maximum amount (in ounces) of polystyrene that can be injected with one complete ram stroke.

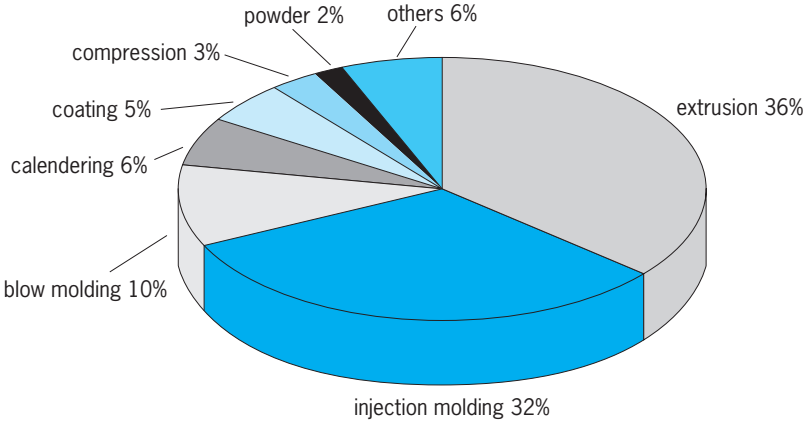

Fig. 1. Plastics consumption by process. (*After D. V. Rosato, M.,G. Rosato, and D. V. Rosato, Concise Encyclopedia of Plastics, Kluwer Academic, 2000*)

Since the introduction of the reciprocating screw injection molding machine, numerous attempts have been made to develop injection molding processes to produce parts with special design features and properties. Alternative processes derived from conventional injection molding have created additional applications, more design freedom, and special structural features. These efforts have resulted in a number of special injection molding processes, including:

- Co-injection (sandwich) molding
- Fusible (lost, soluble) core injection molding
- Gas-assisted and water-assisted injection molding
- Injection-compression molding
- In-mold decoration and in-mold lamination
- Insert and outsert molding
- Lamellar (microlayer) injection molding
- Liquid silicone injection molding
- Low-pressure injection molding
- Microcellular injection molding
- Multicomponent injection molding (overmolding)

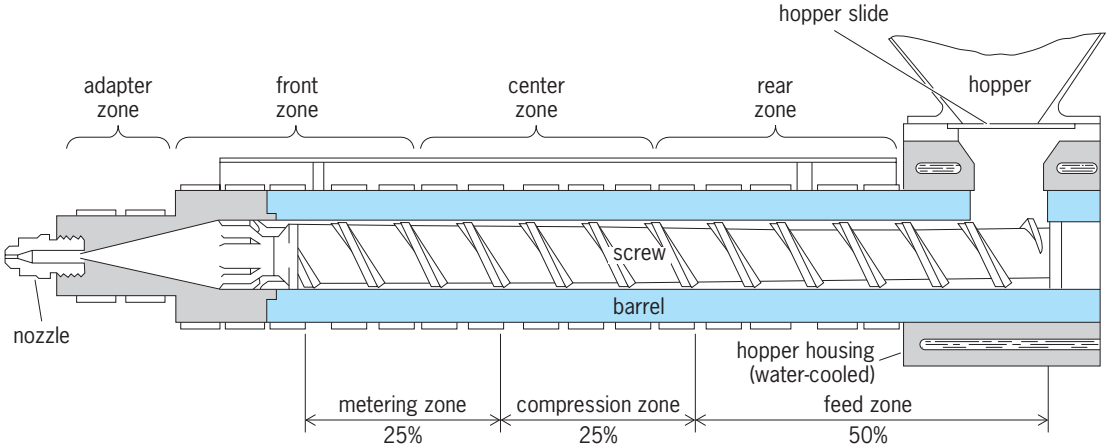

Fig. 2. Typical in-line reciprocating screw showing the three primary barrel zones and their percentages of total length. Note that the screw configuration differs in the three barrel zones. (*After W. O. Elliott, Injection molding, Modern Plastics Encyclopedia, vol. 45, no. 14a, pp. 644–661, McGraw-Hill, 1968–1969*)

Microinjection molding
Multiple live-feed injection molding
Powder injection molding
Push-pull injection molding
Rheomolding
Rubber injection molding
Structural foam injection molding
Thin-wall injection molding

Most processes listed above are for thermoplastics materials. Processes that use thermosetting resins include reactive injection molding, resin transfer molding, and structural reaction injection molding. The development of reaction injection molding allowed for the rapid molding of liquid materials. In these processes (cold or warm), two highly reactive, low-molecular-weight, low-viscosity resin systems are first injected into a mixing head and then into a heated mold, where the reaction to a solid is completed. Polymerization and cross-linking occur in the mold. This process has proven particularly effective for high-speed molding materials such as polyurethanes, epoxies, polyesters, and nylons. *See* POLYAMIDE RESINS; POLYESTER RESINS; POLYMERIZATION; POLYURETHANE RESINS.

Extrusion. In this typically continuous process, plastic pellets or granules are fluidized and homogenized through the rotating action of a screw (or screws) inside a barrel, and the melt is continuously pushed under pressure through a shaping die to form the final product. As material passes through the die, it initially acquires the shape of the die opening. The shape will change once the material exits from the die. The extruded product is called the extrudate, and the typical dimensional expansion exhibited by the extrudate is called extrudate swell (formerly, die swell). Extrudate swell is caused by the viscoelastic behavior of the plastic melt and other factors, such as the inertia effect and the change of velocity distribution. Products made this way include extruded tubing, pipe, film, sheet, wire and substrate coatings, and profile shapes. The process is used to form very long shapes or many small shapes, which can be cut from the long shapes. The homogenizing capability of extruders is used for plastics blending and compounding. Pellets used for other processing methods, such as injection molding, are made by chopping long filaments of extruded plastic.

The operation of an extruder provides a continuous length resembling that from a toothpaste tube. The great majority of machines are still single-screw extruders, although twin- or multiscrew extruders have been used as continuous mixing and compounding devices. Extruders consist of essentially seven machine elements (**Fig. 3**): the feed hopper into which the plastics pellets are charged; the barrel containing the screw and having heating and cooling units on its exterior face and (for some extruders) vent ports; the screw which plasticates, heats, fluidizes, homogenizes, and transports the plastic to the die; the motor and gear reduction system for rotating the screw; the screen pack and breaker plates for filtration of the plastic melt and for backpressure; the die which establishes the extrusion profile; and the ancillary elements and instrumentation for the monitoring of process variables such as pressure, temperature, and screw revolutions.

Single-screw extruders are often characterized by the length-to-diameter ratio of the screw, the number of stages (number of consecutive feed, compression, and metering sections in a screw), the compression ratio, and the meter ratio. These parameters provide a means of comparing different extruders. The extrusion process can be combined with a variety of postextruding operations for expanded versatility. Such postforming operations include altering round to oval shapes, blowing to different dimensions, machining, punching, and calendering (a finishing process in which the extruded sheet from

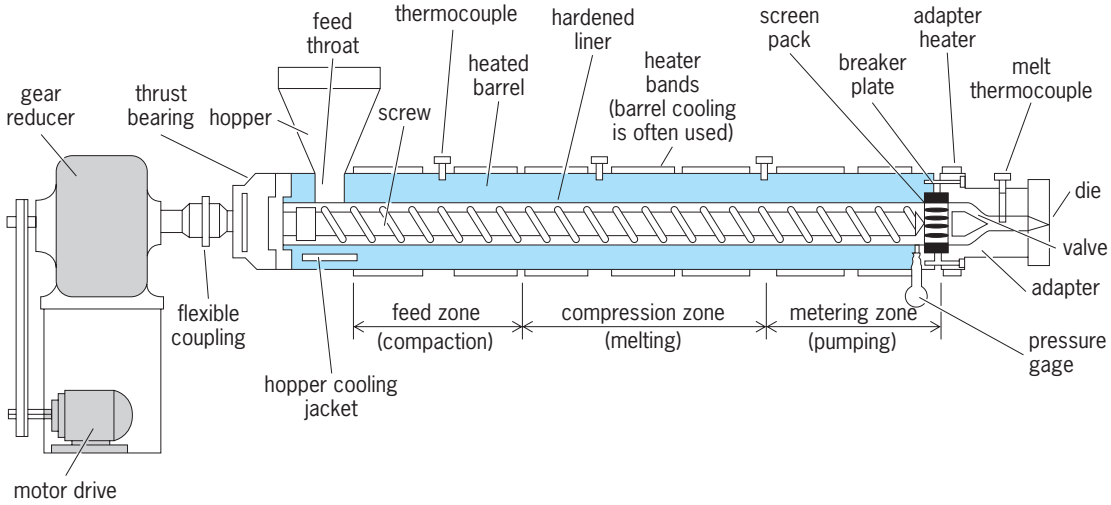

Fig. 3. Typical single-screw extruder. Thermoplastic material is fed into the hopper and emerges as a hot melt at the die. (*After R. T. Van Ness, G. R. DeHoff, and R. M. Bonner, Extrusion of plastics, Modern Plastics Encyclopedia, vol. 45, no. 14a, pp. 672–700, McGraw-Hill, 1968–1969*)

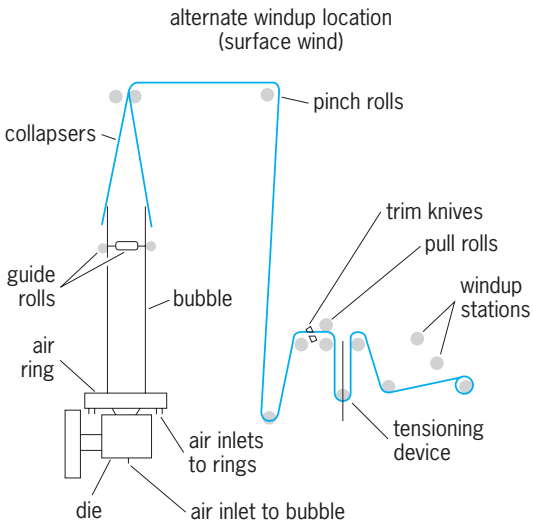

Fig. 4. Extrusion equipment (film blowing machine) for producing blown tubing. The tube is expanded by air (blown) and then is collapsed for windup. (*After R. T. Van Ness, G. R. DeHoff, and R. M. Bonner, Extrusion of plastics, Modern Plastics Encyclopedia, vol. 45, no. 14a, pp. 672–700, McGraw-Hill, 1968–1969*)

an internal batch mixer is passed through pairs of heated rolls to produce a smooth or textured surface). Extrusion can have the highest output rate of any plastics process.

The extrusion process produces pipe and tubing by forcing the melt through a cylindrical die. Wire is coated by running it through the center of an annular opening, so that the melt solidifies directly on the wire. Film can be produced by extruding a thin (less than 0.01 in. or 0.25 mm) sheet over chilled rolls (cast film process), which is then further drawn down to size by tension rolls. Blown film (generally stronger, tougher, and made more rapidly in less bulky equipment than cast film) is made by extruding a tube. In producing blown tubing, the melt flow is turned upward from the extruder and fed through an annular die. As this tube leaves the die, internal air pressure is introduced through the die mandrel, which expands the tube 1.5 to 2.5 times the die diameter and simultaneously draws the film, causing a reduction in thickness (**Fig. 4**). The resulting sleeve is subsequently slit along one side, making a larger film width than could be conveniently made via the cast film method. Sheet extrusion is similar to cast film, except that thicker sections (greater than 0.01 in. or 0.25 mm) are produced through an elongated die. In extrusion coating, the substrate (paper, foil, fabric, and so forth) is compressed together with the extruded plastic melt by means of pressure rollers so that the plastic impregnates the substrate for maximum adhesion. Profile extrusion is designed for rapid turnover of specially made dies because of the relatively short runs per shape. Control of shrinkage and resultant shape distortion is the most important consideration in profile extrusion. Several rules of thumb should be followed, such as avoiding thick sections, achieving uniform wall thickness, and minimizing the number of hollow sections.

Blow molding. This process consists of melting plastics pellets or granules similar to that of extrusion or injection molding, forming a tube (called a parison or preform) and introducing air or other gas to cause the tube to expand into a free-blown hollow object or, most commonly, against the surface of a spilt mold with a hollow cavity for forming into a hollow object with a definite size and shape. The parison is traditionally made by extrusion (extrusion blow molding), although injection molded preforms (injection blow molding) have gained prominence because they can be mass produced with molded-in features such as cap screws for bottles, do not require postfinishing (for example, trimming), have better dimensional tolerances and wall thicknesses, and can be made unsymmetrical and in higher volumes. In the injection blow molding process, the preforms can be transported from injection mold to blow molding machine directly, or they can undergo a reheating stage in between. In addition to extrusion blow molding and injection blow molding, a retrofitted stretch-blow system for stretch-blow bottle formation can be used in (**Fig. 5**).

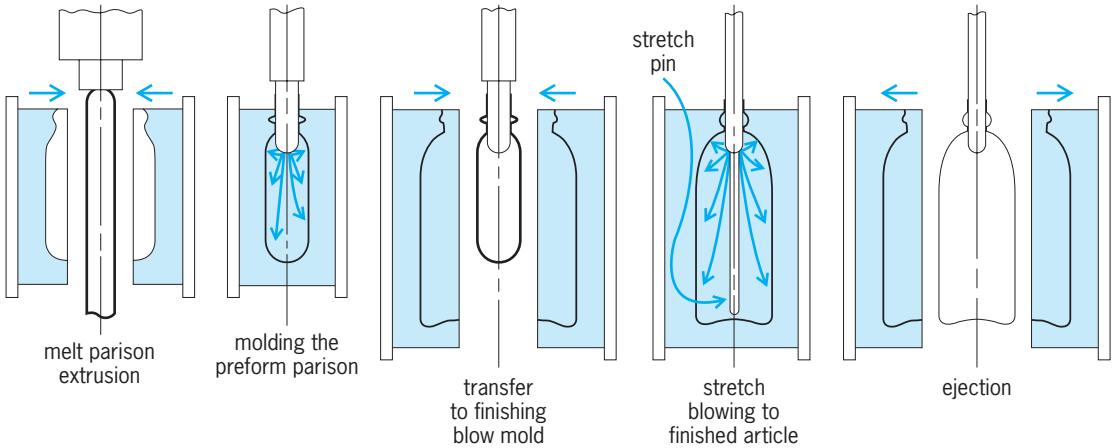

Fig. 5. Retrofitted stretch blow system that uses a preforming mold to condition a blow molded parison before final blow molding of a bottle. (*After G. R. Smoluk, Modern Plastics, p. 47, McGraw-Hill, February 1981*)

Fig. 6. Common blow-molded products with some preblown, injection molded preforms (tubes) shown in left of center near the foreground.

In blow molding, the mold is cooled with water for solidifying the expanded and thinned parison into the part. After the part is ejected, the pinch-off section of the parison is trimmed. One primary concern during blow molding is controlling the wall thickness distribution within the molded part, as it greatly influences the integrity, performance, and material cost of the final part. Unlike injection molding, which typically requires high-flow materials, resins used in blow molding have a high molecular weight to provide the proper melt viscosity for extrusion and melt strength in the parison for supporting its weight during extrusion as well as for expansion. A variety of blow-molded products are shown in **Fig. 6**, together with preblown, injection molded preforms (tubes). Bottles still represent a large share (80%) of the blow-molding market, with the remaining 20% considered as industrial blow molding (for example, fuel tanks).

Thermoforming. Plastics sheets into parts are thermoformed through the application of heat and pressure. Heat is used to soften the plastic sheet so that it can be stretched under pressure (vacuum) against one side of a cool mold surface. The pressure can be obtained through use of pneumatics (air) or compression (tooling) or vacuum (**Figs. 7** and **8**). Increasingly, prestretching of the softened sheet and/or a plug is used to improve the wall thickness uniformity of the final part. When the sheet has cooled enough to retain sufficient strength and the shape of the mold, the formed sheet is removed from the mold, and the excess material used to hold and connect the final articles (called trim, web, or skeleton) is cut off. Typically, thermoforming can be further classified into two subprocesses according to the thickness or gauge of the sheet used: thin-gauge thermoforming (sheet is less than 0.06 in. or 1.5 mm) and heavy-gauge thermoforming (sheet thickness greater than 0.12 in. or 3 mm). There is a gray area between thin-gauge and heavy-gauge thermoforming that depends on the way the sheet is presented to the thermoforming press (in rolls or discrete sheets). Tooling for thermoforming is the least expensive of the plastics processes, accounting for the method's popularity. It can also accommodate very large parts as well as small parts, the latter being useful in low-cost prototype fabrication. Thermoformed products are typically categorized as permanent, or industrial products, and disposable products (for example, packaging).

Rotational molding. Rotational molding is used to make hollow objects, such as those produced by blow molding, except that with a closed rotational mold this process is capable of producing articles without openings. In this process, finely ground powders are heated in a rotating mold until melting or fusion occurs. If liquid materials such as vinyl plastisols are used, the process is often called slush molding. With the rotating motion of the mold, the melted or fused resin uniformly coats the inner surface of the mold. When cooled, the mold opens and

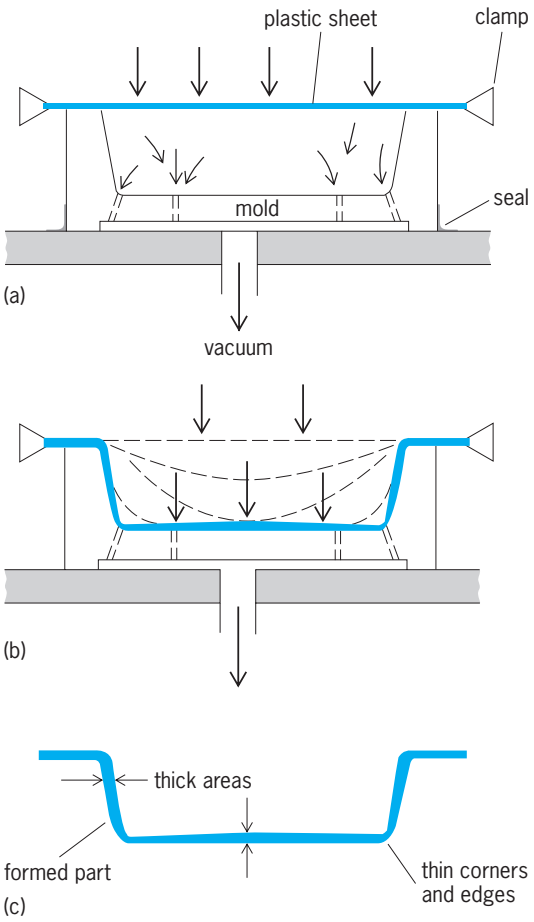

Fig. 7. Straight vacuum forming. (*a*) The plastic sheet is clamped and heated; the vacuum beneath the sheet causes atmospheric pressure to push the sheet down into the mold. (*b*) As the plastic contacts the mold, it cools. (*c*) In formed parts, areas of the sheet reaching the mold last are the thinnest. (*After W. K. McConnell, Jr., Thermoforming, Modern Plastics Encyclopedia, vol. 45, no. 14a, pp. 776–792, McGraw-Hill, 1968–1969*)

a hollow finished part is removed. The processes require relatively inexpensive tooling, are scrap-free, and are adaptable to large, double-walled, hollow parts that are strain-free and uniform in thickness. The processes can be performed by relatively unskilled labor, and the wall thickness is more uniform than for blow molding. However, the finely ground plastics powders are more expensive than pellets or sheets, the thin-walled parts cannot be easily made, and the process is not suited for large production runs of small parts due to long cycle times for heating and cooling of the mold and plastic material.

Compression and transfer molding. Compression molding is one of the oldest molding techniques and consists of charging a plastics powder or preformed plug into a mold cavity, closing a mating mold half, and applying pressure to compress, heat, and cause the plastic to flow and conform to the cavity shape (**Fig. 9**). The process is used primarily for thermosets, bulk molding compounds, and their composite counterparts, sheet molding compounds. Consequently, the mold is heated to accelerate the chemical cross-linking reaction.

Transfer molding is an adaptation of compression molding in that the molding powder or preform is charged to a separate preheating chamber and, when appropriately fluidized, injected into a closed mold (**Fig. 10**). The process predates, yet closely parallels, the early techniques of ram injection molding of thermoplastics. It is mostly used for thermosets and is somewhat faster than compression molding. In addition, parts are more uniform and more dimensionally accurate than those made by compression molding. Many integrated circuit chips are encapsulated by transfer molding of epoxy molding compounds.

Foam processes. Foamed plastics materials have achieved a high degree of importance in the plastics industry. Foams range from soft and flexible to hard and rigid. There are three types of cellular plastics: blown (expanded matrix such as a natural sponge), syntactic (the encapsulation of hollow organic or inorganic microspheres into the matrix), and structural (dense outer skin surrounding a foamed core).

The blowing agents used for producing porous or cellular plastics can be roughly classified into two

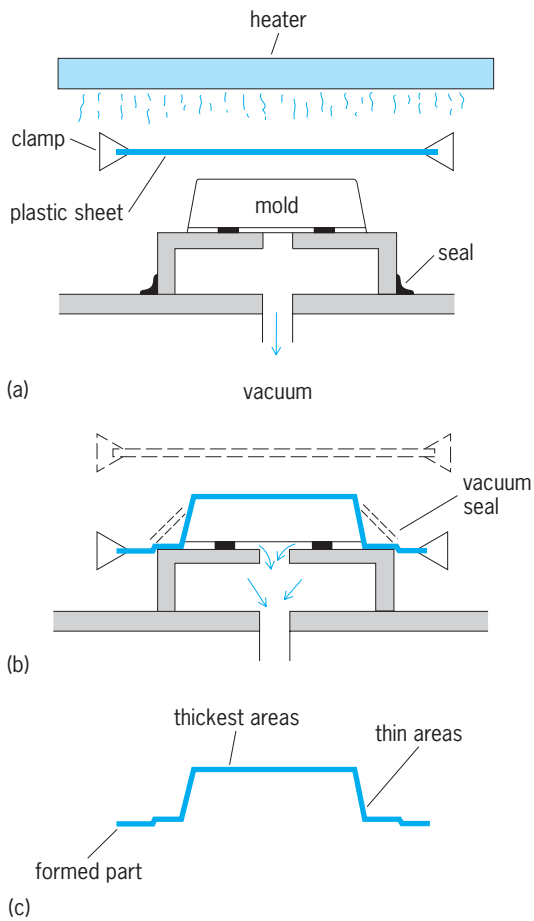

Fig. 8. Drape forming. (*a*) The plastic sheet is clamped and heated and then drawn over the mold. (*b*) When the mold has been forced into a sheet and a seal created, vacuum applied beneath the mold forces the sheet over the male mold. Side walls are formed from material draped between the top edges of the mold and the bottom seal area at the base. (*c*) Final wall thickness distribution. (*After W. K. McConnell, Jr., Thermoforming, Modern Plastics Encyclopedia, vol. 45, no. 14a, pp. 776–792, McGraw-Hill, 1968–1969*)

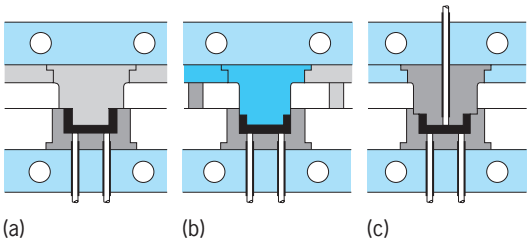

Fig. 9. Three types of compression molds. (*a*) Flash-type. (*b*) Positive. (*c*) Semipositive. (*After E. W. Vaill, Thermoset molding, Modern Plastics Encyclopedia, vol. 45, no. 14a, pp. 714–731, McGraw-Hill, 1968–1969*)

main categories: physical blowing agents and chemical blowing agents. The amount and type of blowing agent influences the densities of the finished products and their microstructures. Physical blowing agents are volatile liquids or compressed gases that change state during processing to form a cellular structure. Chemical blowing agents are solids that decompose above a certain temperature and release gaseous products. Depending on the blowing agents used and how they are being introduced, cellular plastics or plastics foams can be produced by a number of processes. These include the incorporation of a chemical blowing agent that generates gas through thermal decomposition in the polymer liquid or melt; gas injected into the melt which expands during pressure relief; gas generated as a by-product of a chemical condensation reaction during cross-linking; volatilization of a low-boiling liquid/solvent through the exothermic heat of reaction; dispersion of air by mechanical means (whipped cream); incorporation of nonchemical gas-liberating agents (adsorbed gas on finely divided carbon) into the resin mix which is released by heating; expansion of small beads of thermoplastic resin containing a blowing agent through the external application of heat; and dissolving atmosphere gases, such as nitrogen or carbon dioxide, at the supercritical state into the melt and subsequently introducing cell nucleation and growth during processing.

Structural foams differ from other foams in that they are produced with a hard integral skin on the outer surface and a cellular core in the interior. They are made by injection molding liquefied resins containing chemical blowing agents. The initial high injection pressure causes the skin to solidify against the mold surface without undergoing expansion. The subsequent reduction in pressure allows the remaining material to expand and fill the mold. Co-injection (sandwich) molding permits injection molding of parts containing a thermoplastic core within an integral skin of another thermoplastic material. When the core is foam, an advanced form of structural foam is produced. Microcellular injection molding (also known as the MuCell process) blends gas (usually nitrogen or carbon dioxide) in the supercritical fluid state with polymer melt in the machine barrel to create a single-phase, polymer-gas solution. To facilitate fast and homogeneous mixing of the gas and the polymer and to maintain a single-phase solution prior to injection, the melt pressure inside the screw barrel is typically one order of magnitude higher than the typical backpressure used with the conventional injection molding process. During injection, a sudden pressure drop in the material as it flows through the nozzle triggers the thermodynamic instability of the polymer-gas solution. As a result, the gas emerges from the melt, forming numerous microcells. The size (typically about 10 to 100 micrometers) and density of microcells (10^7 and above) depend strongly on the process conditions and the material system, and are much finer and denser than those of structural foam parts (250 μm to a few millimeters). This process is capable of producing parts with excellent dimensional stability, using lower injection pressures, shorter cycle times, and less material. It was applied to the extrusion process before seeing wider applications in injection molding. *See* FOAM.

Reinforced plastics/composites. These plastics have mechanical properties that are significantly improved due to the inclusion of fibrous reinforcements. The wide variety of resins and reinforcements that constitute this group of materials led to the more generalized use of the term composites.

Composites consist of two main components: a fibrous material in various physical forms, and a fluidized resin that will convert to a solid generally through a chemical reaction. Fiber-reinforced thermoplastic materials are typically processed in standard thermoplastic processing equipment.

For components using thermoset resins, the first step in the composite fabrication procedure is the impregnation of the reinforcement with the resin. The simplest method is to pass the reinforcement through a resin bath and use the wet impregnate directly. For easier handling and storage, the impregnated reinforcement can be subjected to heat to remove impregnating solvents or advance the resin cure to a slightly tacky or dry state. The composite in this form is called a prepreg (B stage). This B-stage condition allows the composite to be handled, yet

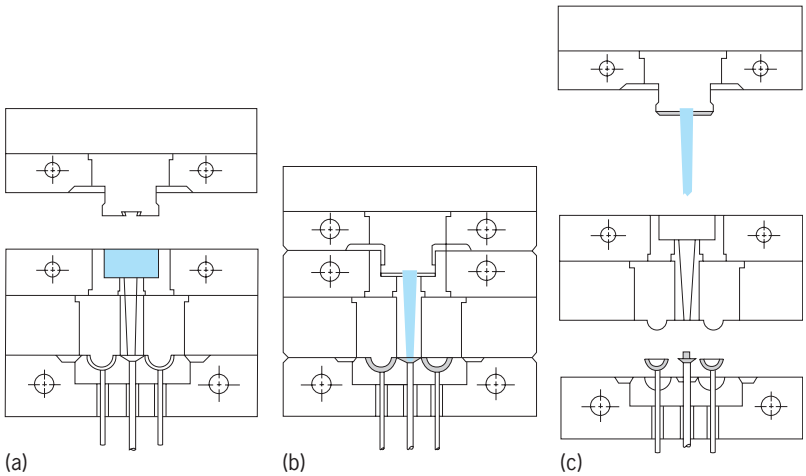

Fig. 10. Transfer molding. (*a*) In the molding cycle, material is first placed in the transfer pot. (*b*) Material is then forced through an orifice into the closed mold. (*c*) When the mold opens, the cull and sprue are removed as a unit, and the part is lifted out of the cavity by ejector pins. (*After E. W. Vaill, Thermoset molding, Modern Plastics Encyclopedia, vol. 45, no. 14a, pp. 714–731, McGraw-Hill, 1968–1969*)

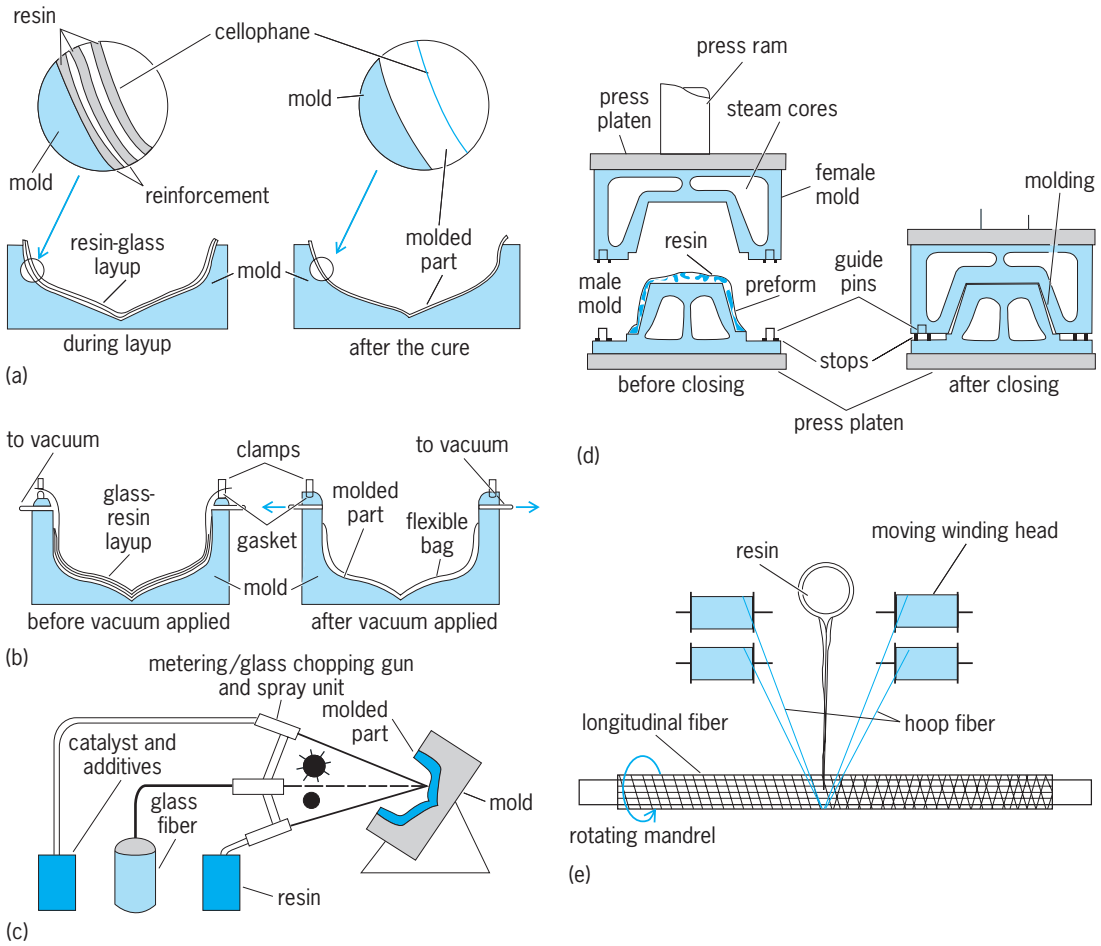

Fig. 11. Techniques for producing reinforced plastics and composites. (*a*) Hand lay-up technique for reinforced thermosets. (*b*) Vacuum bag molding method. (*c*) Spray-up method. (*d*) Matched metal die molding. (*e*) Filament winding. (*After R. S. Morrison, Reinforced thermosets, Modern Plastics Encyclopedia, vol. 45, no. 14a, pp. 628–637, McGraw-Hill, 1968–1969*)

the cross-linking reaction has not proceeded so far as to preclude final flow and conversion to a homogeneous part when further heat or pressure is applied.

Premixes, often called bulk molding compounds, are mixtures of resin, inert fillers, reinforcements, and other formulation additives, which form a putty-like rope, sheet, or preformed shape.

Converting these various forms of composite precursors to final part shapes is achieved in a number of ways. Hand lay-up techniques consist of an open mold onto which the impregnated reinforcement or prepreg is applied layer by layer until the desired thicknesses and contours are achieved (**Fig. 11***a*). The thermoset resin is then allowed to harden (cure). Often the entire configuration will be enclosed in a transparent sealed bag (vacuum bag) so that a vacuum can be applied to remove unwanted volatile ingredients and entrained air to improve the densification of the composite (Fig. 11*b*). External heat may be applied to accelerate the process. Often, a bagged laminate will be inserted into an autoclave so that the synergistic effects of heat, vacuum, and pressure can be obtained. At times, a specially designed spray apparatus is used which simultaneously mixes and applies a coating of resin and chopped reinforcement onto a mold surface (Fig. 11*c*). This technique is particularly useful for large structures, such as boat hulls and truck cabs, covering complex shapes as readily as simple configurations.

Resin transfer molding and structural reaction injection molding are two additional liquid composite molding processes that are well suited for medium to mass production of medium-to-large, complex, lightweight, and high-performance composite components primarily for the aerospace and automotive industries. In these processes, a reinforced fiber mat (preform) is preplaced in a closed mold to be impregnated by a low-viscosity, reactive liquid resin in a transfer or injection process. Sometimes, vacuum is used to assist the flow of the resin. Resin transfer molding and structural reaction injection molding are similar, with the main differences being the resin used, the mixing and injection set-up, the mold requirements, the cycle time, the fiber volume fraction, and the suitable production volume.

Matched-die compression molding resembles normal compression molding, although the pressures are considerably lower (Fig. 11*d*). Premix molding is essentially the same process, except that premix compounds are used. Pultrusion is a process for the

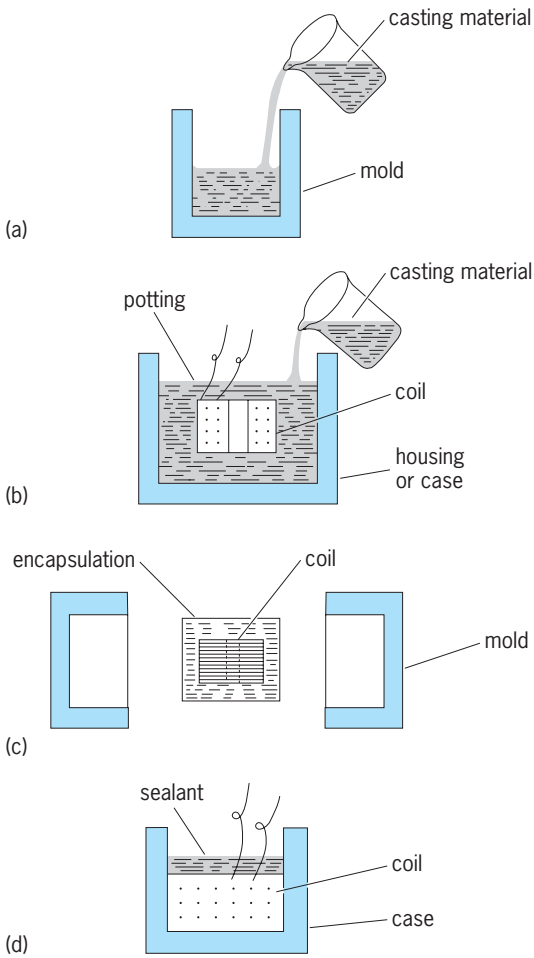

Fig. 12. Low-pressure plastics processes: (*a*) casting, (*b*) potting, (*c*) encapsulation, and (*d*) sealing. (*After H. Lee and K. Neville, Handbook of Epoxy Resins, McGraw-Hill, 1967*)

continuous extrusion of reinforced plastics profiles. Strands of reinforcement are drawn (pulled) through an impregnating tank, the forming die, and then a curing area (radio-frequency exposure). Filament winding is a process in which the continuous strands of reinforcement are drawn through an impregnating bath and then wound around a mandrel to form the part (Fig. 11*e*). This technique is most often used to form hollow objects such as chemical storage tanks or chemically resistant pipe. Advanced automated processes, such as ply cutting, tape laying and contouring, and ply lamination, provide improved parts and reduced costs, particularly in the aerospace industry. *See* COMPOSITE MATERIAL; POLYMERIC COMPOSITE.

Casting and encapsulation. Casting is a low-pressure process requiring nothing more than a container in the shape of the desired part. For thermoplastics, liquid monomer is poured into the mold and, with heat, allowed to polymerize in place to a solid mass. For vinyl plastisols, the liquid is fused with heat. Thermosets, usually composed of liquid resins with appropriate curatives and property-modifying additives, are poured into a heated mold, wherein the cross-linking reaction completes the conversion to a solid. Often a vacuum is applied to degas the resultant part for improved homogeneity.

Encapsulation and potting are terms for casting processes in which a unit or assembly is encased or impregnated, respectively, with a liquid plastic, which is subsequently hardened by fusion or chemical reaction (**Fig. 12**). These processes are predominant in the electrical and electronic industries for insulation and protection of components.

Calendering. In the calendering process (**Fig. 13**), a plastic is masticated between two rolls that squeeze it out into a film that passes around one or more additional rolls before being stripped off as a continuous film. Fabric or paper may be fed through the latter rolls so that they become impregnated with the plastic. *See* POLYMER.

Lih-Sheng Turng; Sidney H. Goodman

Bibliography. J. Avery, *Injection Molding Alternatives: A Guide for Designers and Product Engineers*, Carl Hanser Verlag, 1998; A. W. Birley, R. H. Heath, and M. J. Scott, *Plastics Materials: Properties and Applications*, 1988; J. Brydson, *Plastics Materials*, 7th ed., 1999; J. H. Dubois and W. I. Pribble, *Plastics Mold Design Engineering Handbook*, 4th ed., 1987; S. H. Goodman (ed.), *Handbook of Thermoset Plastics*, 2d ed., 1998; F. Hensen, W. Knappe, and H. Potente (eds.), *Plastics Extrusion Technology*, Carl Hanser Verlag, 1988; Plastics and Rubber Institute Staff, *Plastics and Polymer Processing Automation*, 1988; R. G. W. Pye, *Injection Mold Design*, 4th ed., 1989; C. Rauwendaal, *Polymer Extrusion*, 4th ed., Carl Hanser Verlag, 2001; L.-S. Turng, Special injection molding processes, in T. A. Osswald, L.-S. Turng, and P. J. Gramann (eds.), *Injection Molding Handbook*, Carl Hanser Verlag, 2002; A. Whelan, *Plastics Molding*, 4 vols., 1991; L. J. White, *Twin Screw Extrusion*, Carl Hanser Verlag, 1991.

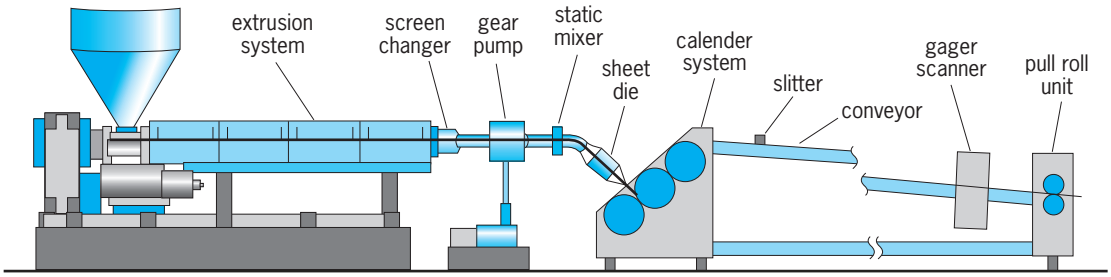

Fig. 13. Calendered sheet extrusion system. (*After D. Hanson, Modern Plastics, p. D-29, McGraw-Hill, Mid-November 1998*)

Plate girder

A beam assembled from steel plates which are welded or bolted to one another, used to support the horizontal surface of a bridge. Plate girders generally have I-shaped cross sections (**Fig. 1**). Cross sections with two or more webs that form a U- or box-shape can also be considered plate girders, but are generally classified separately as box girders. Plate girders are commonly used in bridge structures where large transverse loads and spans are encountered. On average, the height (measured from the top of the top flange to the bottom of the bottom flange) of plate girders is one-tenth to one-twelfth the length of the span, which can vary based on the loading and architectural requirements. The dimensions of plate girders are tailored to specific applications by adjusting the width and thickness of the web and flange plates along the length of the beam. By contrast, hot-rolled beams produced at steel mills are formed to specific dimensions according to a standardized list of structural shapes, and have limitations on their size. Plate girders can be produced with much greater depths than hot-rolled beams. The depth is generally considered to be the distance between the centroids of the flanges which defines the effective moment arm in bending. When different-grade steels are used for the flanges and webs, a plate girder is classified as a hybrid plate girder. *See* BRIDGE; LOADS, TRANSVERSE; STRUCTURAL STEEL.

Splices. Long-span plate girders are often fabricated in segments due to size or weight restrictions at the fabrication facility, during shipping, or at the construction site. These segments are spliced or joined together by either high-strength bolts or welding. Bolted splices are faster to install and easier to inspect than welded splices, but are readily visible to the observer, which can detract from the esthetics of the structure.

The individual flange and web plates that constitute plate girders commonly consist of multiple plates that are spliced together by welding. Individual plate splices are necessary whenever the plate length required exceeds what is available or when the width, thickness, or material grade of the plate changes. When width or thickness changes occur, the edges of the plates being spliced together are beveled prior to welding to minimize any sharp corners or notches that can cause fatigue cracks (Fig. 1). Since welding affects the material properties of the steel, splices in the flanges and webs should be staggered so two splice welds do not coincide.

Stiffeners. Plate girders commonly have stiffener plates which increase the strength of the web plate. Since plate girder webs are relatively slender, they are susceptible to buckling when subjected to shear and bending forces. The bending strength of webs can be enhanced using horizontal (longitudinal) stiffeners, while the shear strength of webs can be enhanced using vertical (transverse) stiffeners (Fig. 1). Increasing the thickness of the web plate increases the overall web strength and can eliminate the need for stiffeners altogether. The additional material cost of a thicker web plate is often less than the material, fabrication, and installation cost of using stiffeners. In some instances, the use or elimination of stiffeners is based on esthetic reasons alone. Bearing stiffeners are vertical stiffeners that are placed

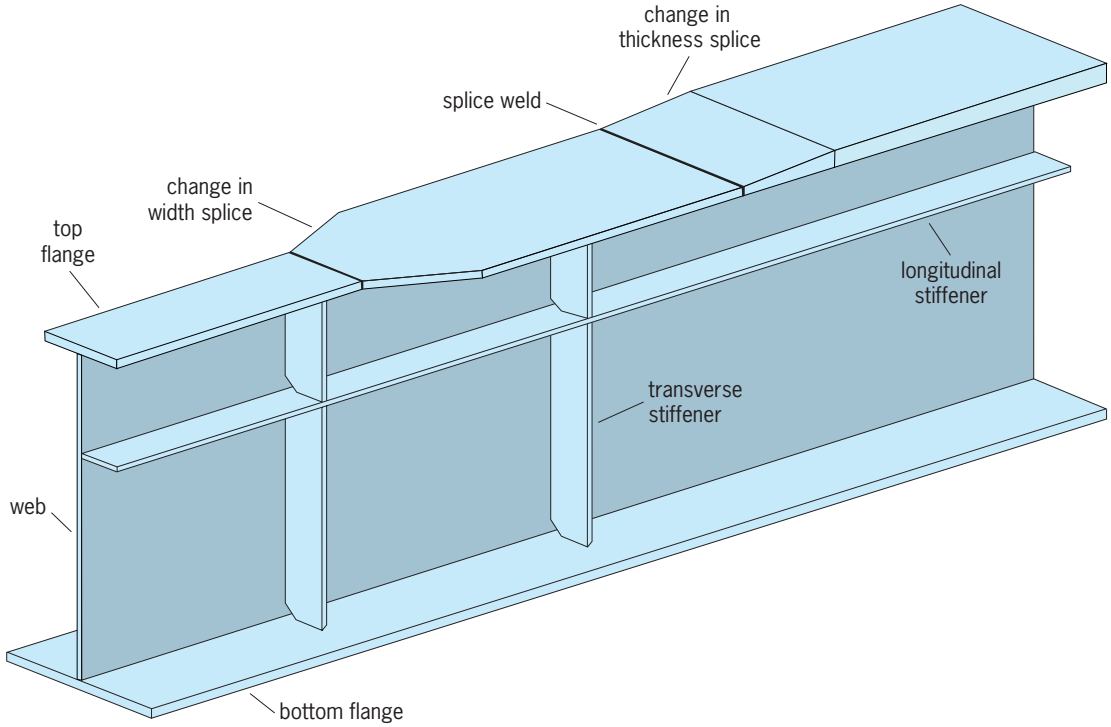

Fig. 1. Plate girder with stiffeners and plate splices.

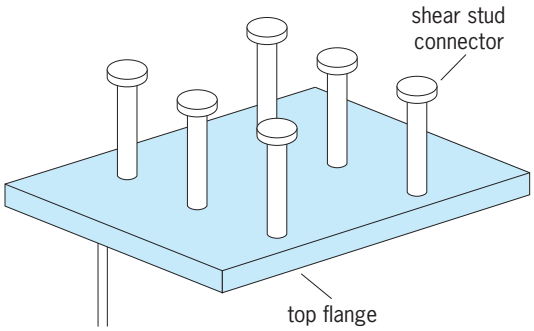

Fig. 2. Shear connectors for composite plate girders.

at locations where concentrated forces are present. Large concentrated forces, such as those at support locations, can cause local buckling of the web. *See* BEAM; SHEAR.

Composite girders. Plate girders that support concrete floors or roadways can be designed as composite or noncomposite girders. For noncomposite plate girders, the concrete is simply laid to rest upon the top flanges. Slip between the concrete and steel is not mechanically prevented, and the structural system behaves as two independent beams. Frictional forces between the concrete and steel are generally ignored as they are very small in comparison to the forces acting to produce slip. Composite plate girders use mechanical shear connectors to prevent slip at the interface (**Fig. 2**). These shear connectors are attached to the steel and embedded into the concrete. This allows the concrete and steel to act together as a single structural unit, producing a significantly stiffer and more efficient beam. The added expense of composite plate girders is generally offset by increased structural efficiency.

Curved girders. Plate girders can be fabricated with a horizontal curve to accommodate curved bridge structures such as highway interchanges. The narrow flange plates are usually cut and spliced as rectangular plates and then mechanically curved by applying heat. The curved flanges are then welded to the web to form the curved girder. Curved girders add a level of complexity to straight girders because forces due to gravity cause both bending and torsion. *See* HIGHWAY BRIDGES; TORSION. Brian Chen

Bibliography. W. F. Chen, *Handbook of Structural Engineering*, CRC Press, Boca Raton, FL, 1997; W. Schueller, *The Design of Building Structures*, Prentice Hall, Upper Saddle River, NJ, 1996.

Plate tectonics

Plate tectonics theory provides an explanation for the present-day tectonic behavior of the Earth, particularly the global distribution of mountain building, earthquake activity, and volcanism in a series of linear belts. Numerous other geological phenomena such as lateral variations in surface heat flow, the physiography and geology of ocean basins, and various associations of igneous, metamorphic, and sedimentary rocks can also be logically related by plate tectonics theory.

Theory and Evidence

The theory is based on a simple model of the Earth in which a rigid outer shell 30–90 mi (50–150 km) thick, the lithosphere, consisting of both oceanic and continental crust as well as the upper mantle, is considered to lie above a hotter, weaker semiplastic asthenosphere. The asthenosphere, or low-velocity zone, extends from the base of the lithosphere to a depth of about 400 mi (700 km). The brittle lithosphere is broken into a mosaic of internally rigid plates that move horizontally across the Earth's surface relative to one another. Only a small number of major lithospheric plates exist, which grind and scrape against each other as they move independently like rafts of ice on water. Most dynamic activity such as seismicity, deformation, and the generation of magma occurs only along plate boundaries, and it is on the basis of the global distribution of such tectonic phenomena that plates are delineated. *See* EARTH INTERIOR; LITHOSPHERE.

The plate tectonics model for the Earth is consistent with the occurrence of sea-floor spreading and continental drift. Convincing evidence exists that both these processes have been occurring for at least the last 600 million years (m.y.). This evidence includes the magnetic anomaly patterns of the sea floor, the paucity and youthful age of marine sediment in the ocean basins, the topographic features of the sea floor, and the indications of shifts in the position of continental blocks which can be inferred from paleomagnetic data on paleopole positions, paleontological and paleoclimatological observations, the match-up of continental margin and geological provinces across present-day oceans, and the structural style and rock types found in ancient mountain belts. *See* CONTINENTAL DRIFT; CONTINENTAL MARGIN.

Plate motion and boundaries. Geological observations, geophysical data, and theoretical considerations support the existence of three fundamentally distinct types of plate boundaries, named and classified on the basis of whether immediately adjacent plates move apart from one another (divergent plate margins), toward one another (convergent plate margins), or slip past one another in a direction parallel to their common boundary (transform plate margins). **Figure 1** shows the major plates of the lithosphere, the major plate margins, and the type of motion between plates. Plate margins are easily recognized because they coincide with zones of seismic and volcanic activity; little or no tectonic activity occurs away from plate margins. The boundaries of plates can, but need not, coincide with the contact between continental and oceanic crust. The nature of the crustal material capping a plate at its boundary may control the specific processes occurring there, particularly along convergent plate margins, but in general plate tectonics theory considers the continental crustal blocks as passive passengers riding on

the upper surface of fragmenting, diverging, and colliding plates.

The velocity at which plates move varies from plate to plate and within portions of the same plate, ranging between 0.8 and 8 in. (2 and 20 cm) per year. This rate is inferred from estimates for variations in the age of the sea floor as a function of distance from mid-oceanic ridge crests. Ocean-floor ages can be directly measured by using paleontological data or radiometric age-dating methods from borehole material, or can be inferred by identifying and correlating the magnetic anomaly belt with the paleomagnetic time scale.

Divergent plate margins. As the plates move apart from the axis of the mid-oceanic ridge system, the new volcanic material welling up into the void forms a ribbon of new material, which gradually splits down its center as the boundary of plate separation continues to develop. Each of the separating plates thus accretes one-half a ribbon of new lithosphere, and in this way new lithosphere and hence new surface area are added. The process is considered to be continuous, and the boundary at which separation is taking place always maintains itself in the center of the new material. *See* MID-OCEANIC RIDGE.

The accretion at any spreading boundary is usually bilaterally symmetric. The morphology of the ridges is also quite symmetric and systematic. The new material that wells up at the ridge axis is hot and therefore expanded and less dense than the surrounding older material. Consequently, the new material is topographically highest. As new material divides and moves away from the ridge axis, it cools, contracts, becomes denser, and subsides. The densification is caused by the combined effect of pure thermal contraction and thermally driven phase changes. Subsidence is fastest for newly generated oceanic crust and gradually decreases exponentially with time. This observation explains the fact that in cross section the shape of the slope of the ridges is steepest at the ridge axis and gradually decreases down the flanks beneath the abyssal sediments and to the bounding continents. Since all known oceanic lithosphere has been generated by the spreading at a ridge axis, all oceanic lithosphere is part of the mid-oceanic ridge system. Because of the systematic way in which the morphology of the ridges is formed, most oceanic crust follows the same time-dependent subsidence curve within an error range of about 300 ft (100 m). This means that the same age-versus-depth curve fits nearly all parts of the mid-oceanic ridge system. The ridge axis is found at a depth of 1.71 mi $\pm$ 300 ft (2.75 km $\pm$ 100 m), and oceanic lithosphere that is 30 m.y. old is found at a depth of 2.72 $\pm$ 300 ft (4.37 km $\pm$100 m). There are exceptional areas such as Iceland where the ridge axis is above sea level.

Magnetic lineations and age of oceans. Much of the evidence which leads to the development of the concept of sea-floor spreading, and in turn to an understanding of divergent plate margins and plate tectonics theory, came from analyses of the magnetic properties of the sea floor. Magnetometer surveys across the sea floor near mid-oceanic ridges reveal a pattern of alternating positive and negative magnetic anomalies (**Fig. 2**). The characteristics of these magnetic anomaly patterns (parallel to and in symmetrically matching widths across the ridge crest) and present-day general knowledge of the Earth's magnetism logically support the conclusion that sea-floor spreading occurs.

The Earth's magnetic field as measured at the surface may be approximately represented by an axial geocentric dipole—in other words, as if a simple two-pole magnet existed coincident with the Earth's rotational axis. The Earth's magnetic field reverses polarity episodically. Over an interval of time, the strength of the dipole field gradually decreases to zero and then gradually increases in the opposite direction. The total transit time from full strength in one direction (polarity) to full strength in the opposite direction may be less than 6000 years. The residence time in any one particular polarity ranges from 10,000 years up to tens of millions of years. Thus, the polarity behavior of the Earth's magnetic field is approximately analogous to a randomly triggered flip-flop circuit.

The top 600 to 1500 ft (200 to 500 m) of the lithosphere is the oceanic basalt, which is accreted at separating boundaries and which contains magnetizable iron minerals. As the newly injected material cools through its Curie point ($\sim$1100°F or 570°C), it becomes permanently magnetized in the direction of the Earth's magnetic field at that time and place. Because the oceanic crust is continuously being formed and transported in a bilaterally symmetric pattern away from the line of rifting, it is similar to a continuous magnetic tape record of the Earth's magnetic polarity events. At the mid-oceanic ridge axis, the present polarity of the field is recorded. Down the ridge flanks, progressively older oceanic crust is traversed; hence the crust contains in its remanent magnetism a continuously older record of the polarity of the field. *See* BASALT; CURIE TEMPERATURE.

As a result of sea-floor spreading and polar reversals, the oceanic crust appears magnetically as a set of alternately normally or reversely magnetized strips of basalt arranged in a bilaterally symmetrical pattern about the mid-oceanic ridge axes. Because the rate of separation at most segments of the mid-oceanic ridge axes has varied only slowly with time, the spatial distribution of the stripes is proportional to the temporal history of the polarity of the field. Thus, the polarity of the oceanic crust may be inferred indirectly from the pattern of magnetic anomalies.

Because the occurrence of polarity events is random, they form a unique sequence which is reflected directly in the magnetic anomaly patterns. A unique pattern of magnetic anomalies has been correlated throughout most of the world's oceans. From this pattern a magnetic reversal time scale from the present to 160 m.y. ago has been developed. This geomagnetic time scale has been calibrated by drilling on key anomalies in the oceans and paleontologically determining the age of the sediment–basalt

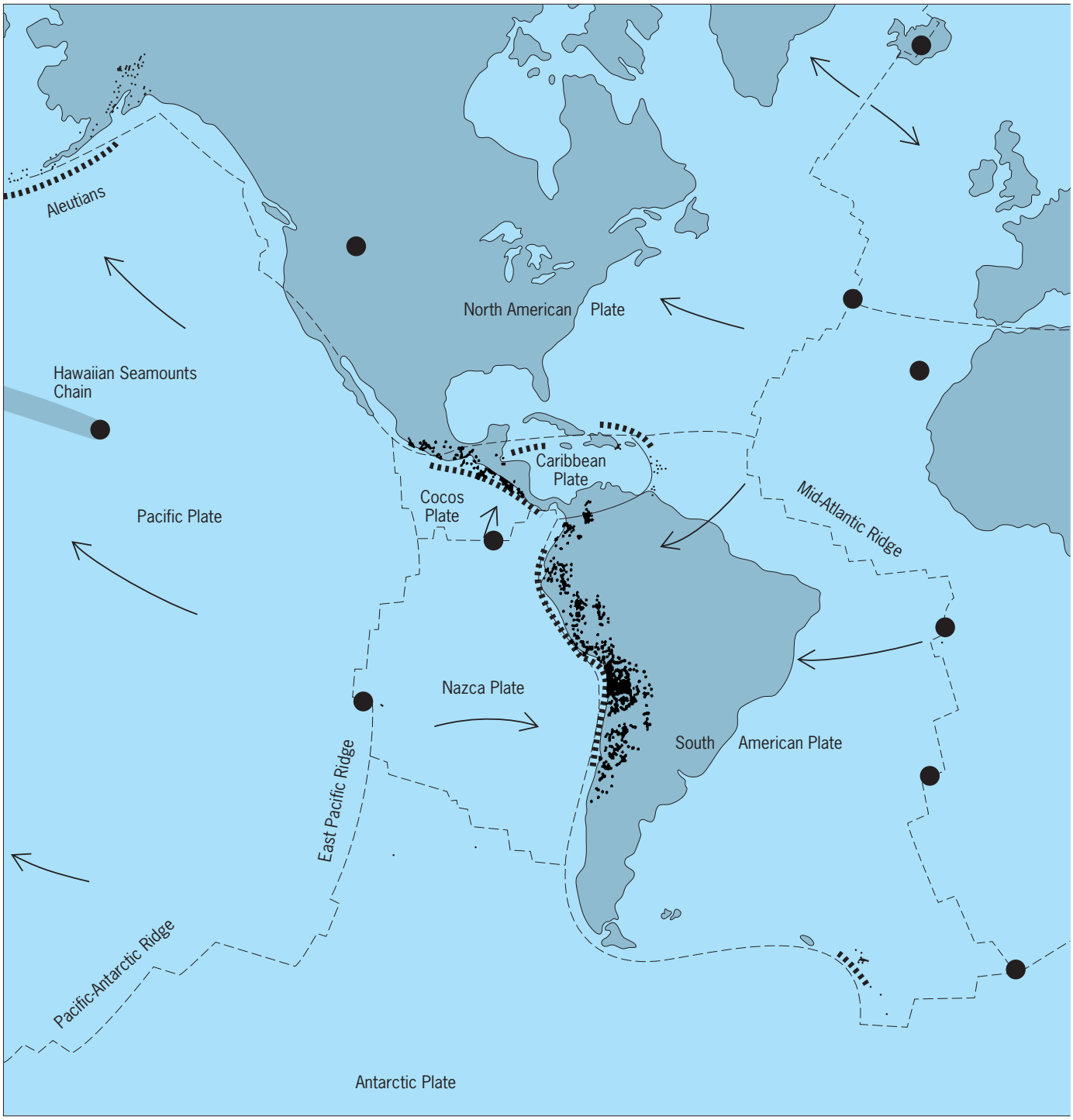

Fig. 1. Tectonic map of the Earth. (*After M. Nafi Toksoz, The subduction of the lithosphere, Sci. Amer., 233(5):88–98, November 1975*)

interface. Thus a magnetic polarity time scale has been developed, and with it identifiable magnetic anomalies that can be correlated with part of the time scale may be assigned an absolute age. In this way, the spatial distribution of the magnetic anomalies is used to calculate the rates of separation of the plates in the past. *See* GEOMAGNETISM; PALEOMAGNETISM.

Convergent (destructive) plate margins. Because the Earth is neither expanding nor contracting, the increase in lithosphere created along divergent boundaries must be compensated for by the destruction of lithosphere elsewhere. The rates of global lithosphere construction and destruction must be equal, or the radius of the Earth would change. Compensatory destruction or removal of lithosphere occurs along convergent plate margins (subduction zones) and is accomplished by plate subduction and continental collision. *See* SUBDUCTION ZONES.

Along subduction zones, one plate plunges

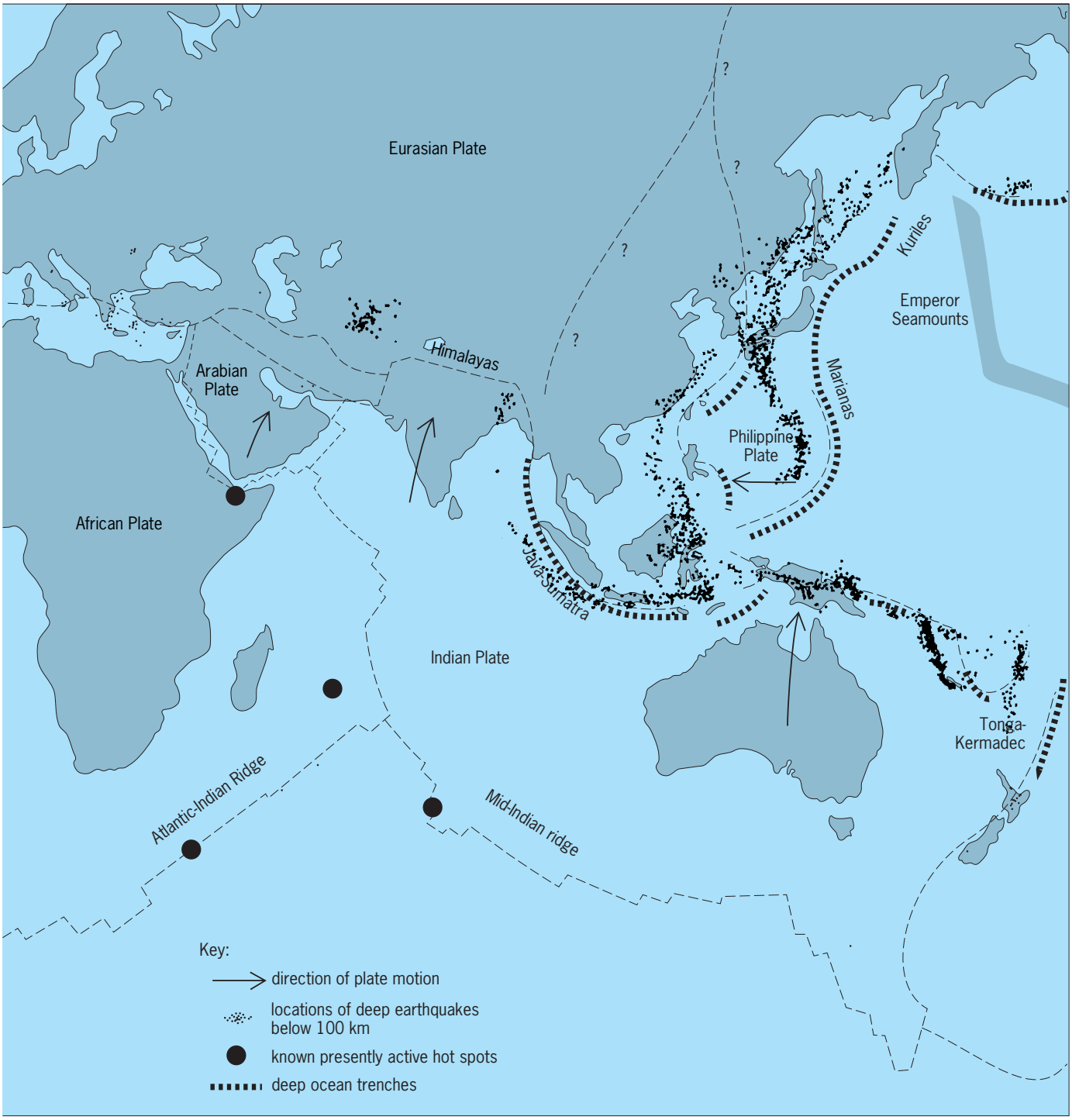

beneath another (**Fig. 3**). The downgoing slab is usually oceanic because the relatively buoyant continental lithosphere cannot be subducted beneath the relatively denser oceanic lithosphere. The upper or overriding plate may be continental or island arc lithosphere, and occasionally oceanic plateaus as in the case of Oregon and Alaska.

The dip of the downgoing underthrust slabs varies but averages 45°. Earthquake foci along individual subduction zones indicate the plate underthrust angle, often referred to as the Benioff plane or Wadati-Benioff zone. This plane dips away from oceanic trenches toward adjacent volcanic arcs and continents, and marks the surface of slippage between the overriding and descending lithospheric plates. Successive belts of shallow (less than 40 mi or 70 km), intermediate (40–200 mi or 70–300 km), and deep-focus (200–400 mi or 300–700 km) earthquakes are associated with subduction zones. Subduction zones are also associated with active

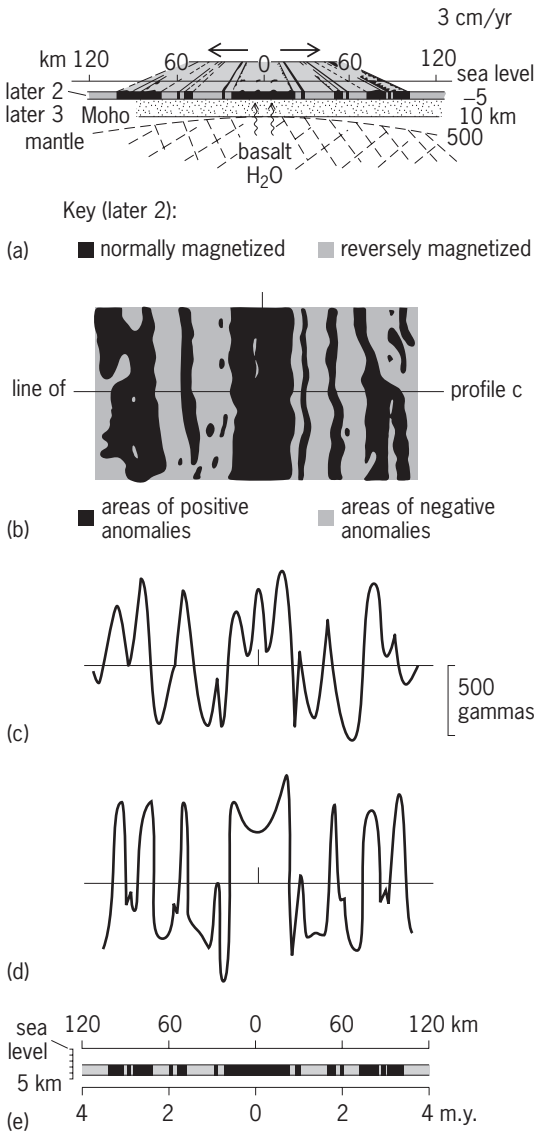

Fig. 2. Tape-recorder conveyor-belt scheme for sea-floor spreading. (*a*) Schematic of crustal model, applied to Juan de Fuca Ridge, southwest of Vancouver Island. (*b*) Part of summary map of magnetic anomalies recorded over Juan de Fuca Ridge. (*c*) Total-field magnetic anomaly profile along the line indicated in *b*. (*d*) Computed profile assuming the model and reversal time scale. Intensity and dip of Earth's magnetic field taken as 54,000 gammas (0.54 oersted) and +667°; magnetic bearing of profile 087°; 10 km horizontally is equivalent to 100 gammas vertically. Normal or reverse magnetization is with respect to an axial dipole vector, and the assumed effective susceptibility is 0.01 except for central block and ridge crest (+0.02). (*e*) Scale applicable to *b–d*. 1 cm = 0.4 in.; 1 km = 0.6 mi; 1 gamma = 1 nanotesla. (*After R. A. Phinney, ed., The History of the Earth's Crust, 1968*)

volcanism and the development of deep-ocean trenches. These features encircle the Pacific Ocean basin.

A line of andesitic volcanoes usually occurs on the upper plate, forming a chain that is parallel to the trench. The volcanism occurs at that point above the subduction zone where the upper surface of the downgoing lithosphere has reached a vertical depth of approximately 75 mi (120 km). The volcanic materials arise from partial melting of the downgoing slab. The volcanoes of the Andes chain of South America and island arcs in the Pacific such as Izu-Bonin and the Mariana have formed in this manner.

Convergent plate margins that develop island arc systems adjacent to trench systems are geologically complex. The region between a volcanic island arc and the trench consists of volcanic rock and sediments. Slumping and turbidity currents transport sediment into the trench axis. The trench sediment may be transferred from the downgoing plate to the continent or island arc. Conversely, a convergent margin without a large sediment supply may be eroded during subduction.

Because of density consideration, subduction requires at least one of the two converging plates to be oceanic. If both converging plates consist of continental lithosphere, continental collision occurs, but not subduction. Continental collision results in compensatory reduction in lithospheric width by folding and compressing the lithosphere into narrower, linear mobile belts. In such collisions, sediments deposited along the continental margins and within the closing ocean basins are compressed into a series of tight folds and thrusts. Fragments of oceanic crust may be thrust up onto adjacent continental rocks (obduction) as ophiolite successions. A classic example of a continental collision belt is the Himalayan belt, produced during the Cenozoic Era by the convergence of the Indian continent with Eurasia.

Like plate divergence, plate convergence produces a distinctive suite of igneous rock types. Subduction zones are marked by the belts of predominantly andesitic volcanoes either in island arcs located landward of the trench system (Japan and the Philippines) or along the rim of overriding continental blocks (the Andes belt). These andesitic volcanic terranes are commonly associated intimately with plutonic igneous rocks, mainly granodiorites. The origin of andesitic magmas, and the predominance of granodiorite plutons within continental blocks to the exclusion of most other igneous rock varieties, was perplexing until the development of plate tectonics theory. Now both seem to be directly related to the generation of parent magmas by the frictional melting of ocean-floor basalt and overlying sediment cover along subduction zones. Partial melting of the lower crust and upper mantle also occurs.

Plate subduction and continental collision can also explain the origin of two other puzzling rock sequences commonly found within mountain belts: mélange and blueschist terranes. Mélange, a heterogeneous assemblage of intensely sheared, poorly sorted, angular blocks set in a fine-grained matrix, is probably generated at shallow depths along subduction zones as the oceanic crust and overlying sediment cover of the descending plate are scraped and crushed against the overriding plate. Blueschist terranes (their dark blue color is imparted by the presence of various low-temperature–high-pressure metamorphic minerals such as glaucophane, lawsonite, and jadeite) occur in belts within mountain chains, parallel with but external to (toward the ocean) the more conventional paired metamorphic

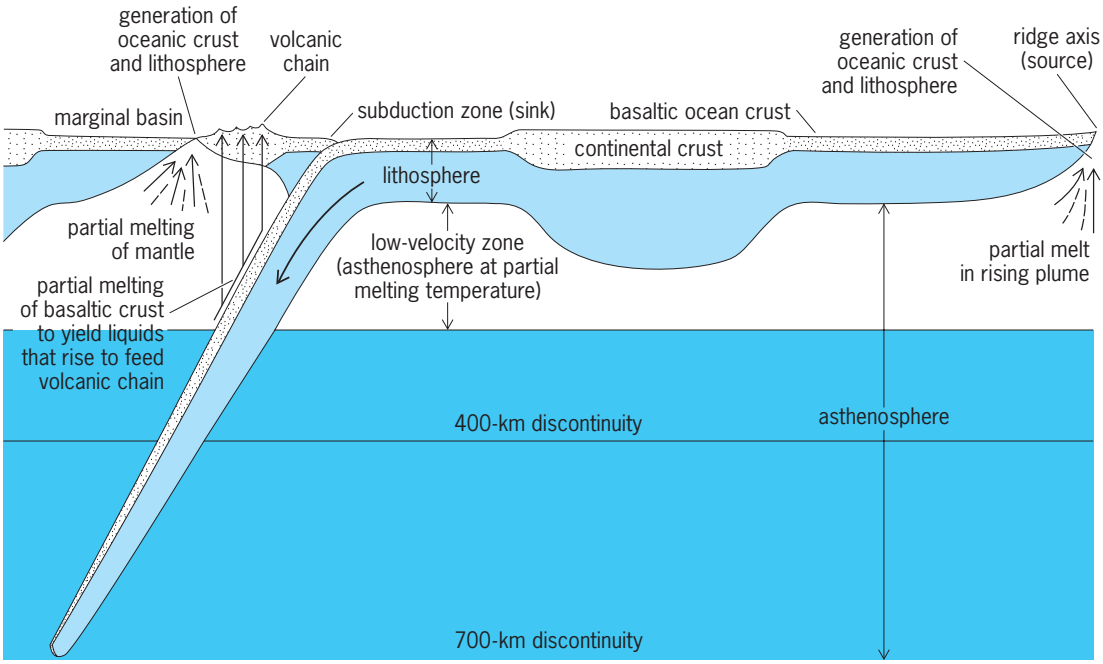

Fig. 3. Cross section of the upper mantle and crust showing a lithospheric plate riding on the asthenosphere. The continent is embedded in the plate and moves with it. Divergent plate margin with generation of oceanic crust and lithosphere is shown on the right; convergent plate margin with subduction is shown on the left. 1 km = 0. 6 mi. (*After J. F. Dewey, Plate tectonics, Sci. Amer., 226:56–68, May 1972*)

facies of the greenschist terrane. The peculiar physical conditions required by the blueschist facies, great burial depth (in excess of 10 mi or 16 km) but moderate temperature (390–800°F or 200–450°C), should be generated along subduction zones when the descending lithospheric plate is underthrust at a greater rate than the local geothermal gradient can heat it. *See* BLUESCHIST.

Transform plate margins. Transform faults are always strike-slip faults. They occur where the relative motion between the two plates is parallel to the boundary that separates the plates. They may join a ridge to a ridge, a ridge to a trench, or a trench to a trench. Ridge-trench transforms will always change length with time. A trench-trench transform may lengthen, shrink, or remain constant, depending on which of the plates that form the subduction system is the downgoing plate. A transform that joins two ridge axes will not change in length with time.

First, consider ridge-ridge transforms (**Fig. 4**). Earthquake epicenter data show that earthquakes occur only along the ridge axes and the connecting transform. Studies of earthquakes that occur at a ridge-ridge transform fault show that the first motion was strike-slip parallel to the direction of the transform and opposite to the sense of offset of the ridge axis. If the ridge axes are offset left-laterally, the relative motion across the transform that joins the ridges will be right-lateral as the plates separate. The motion of the fault takes place spasmodically. Deformation occurs on either side of the fault until the elastic limit is reached and the rupture occurs, causing an earthquake. *See* EARTHQUAKE.

Fracture zones are morphologic scars that are fossilized transform faults. They may be used to

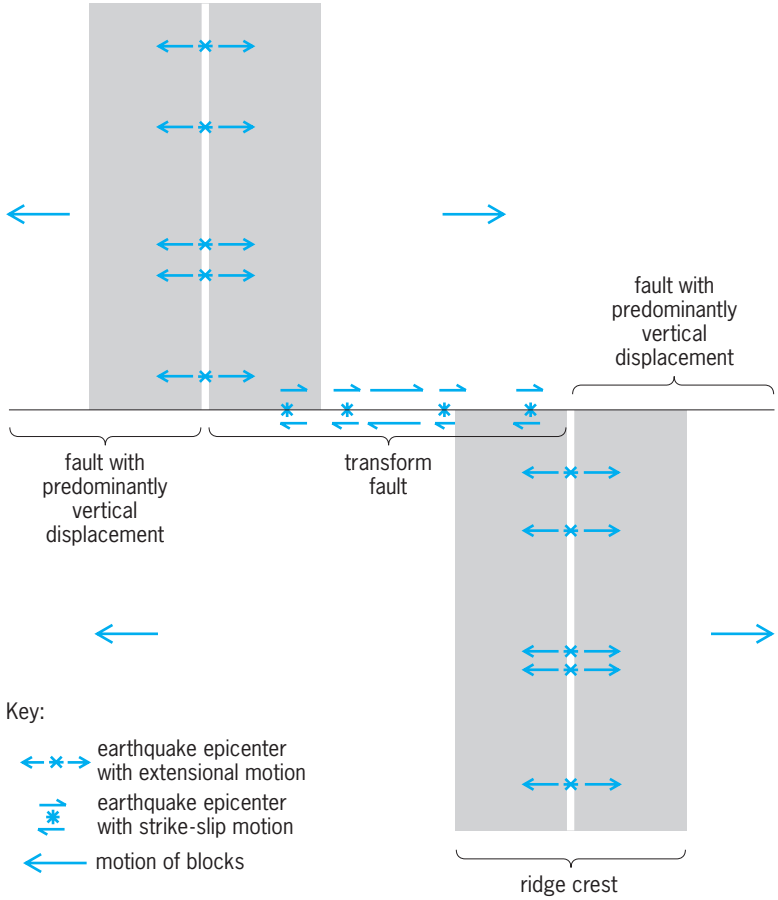

Fig. 4. Idealized map view of a transform fault. The ridge appears to be offset in a left-lateral sense, but the motion on the transform fault between the ridge crests is right-lateral. (*After C. K. Seyfert and L. A. Sirkin, Earth History and Plate Tectonics, Harper Row, 1973*)

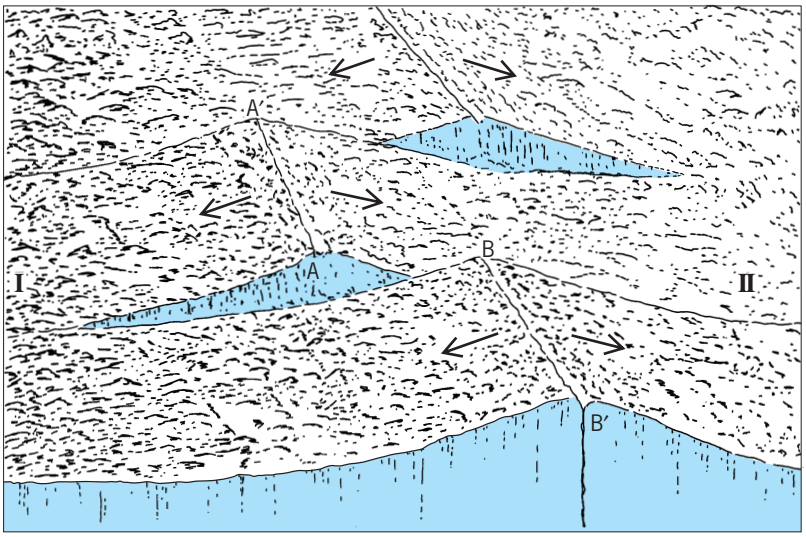

Fig. 5. Ridge-ridge transform fault appears between two segments of ridge that are displaced from each other. (*After J. Tuzo Wilson, ed., Continents Adrift and Continents Aground, W. H. Freeman, 1976*)

determine the past direction of relative motion of the bounding plates. In **Figs. 5** and **6**, the transform fault A-B offsets two ridge axes, and there is right-lateral strike-slip motion across the transform. All the material to the left of the ridge-transform-ridge boundary A′ABB′ belongs to the torsionally rigid plate I, and all the material to the right of this boundary is part of plate II. Thus, beyond the ends of the transform, to the left of A' and right of B, there is no relative horizontal motion.

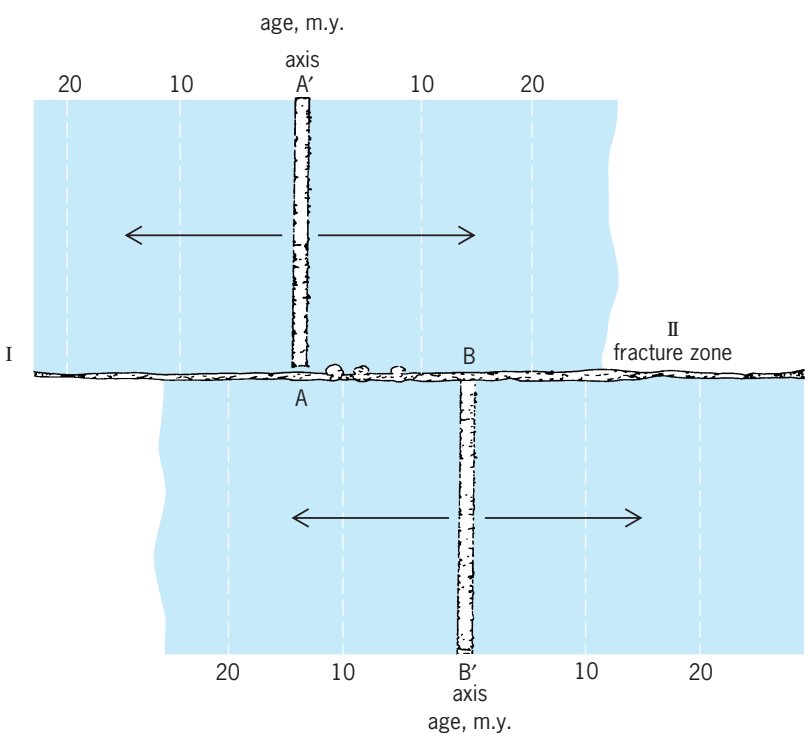

Fig. 6. Molten rock wells up from the deep Earth along a spreading axis, solidifies, and is moved out (shown by arrows). The axis is offset by a transform fault. Between two offset axes, material on each side of the transform fault moves in opposite directions, causing shallow earthquakes. (*After J. R. Heirtzler, Sea-floor spreading, Sci. Amer., vol. 219, no. 6, December 1968*)

Because of the way in which fracture zones are formed, they also preserve a record of the past direction of relative motion of the plates. The transform fault between plates I and II in Figs. 5 and 6 is in theory a vertical plane between these plates along which horizontal strike-slip motion occurs. The face of plate I that abuts against plate II in effect defined this transform plane. But as plate I moves to the left, this transform plane or face passes by the end of ridge A′A at point A, and new younger material is intruded against it. Together, the transform plane of plate I and the new younger material move off to the left. This process is continuous. The new younger material is more elevated and forms an escarpment whose horizontal trace records the direction of relative movement. Because the rate of subsidence decreases with time, the difference in elevation across a fracture zone decreases with age.

A very narrow and deep rift valley is also associated with the transform. However, because of time-dependent cooling, the oceanic lithosphere contracts horizontally. This contraction causes cracking of the lithospheric plate and widening and splaying of the rift associated with the fracture zone. Young fracture zones are characterized by an escarpment and a deep fissure along their length. The magnitude of the vertical offset decreases with age, and the fissure tends to widen. Down the lower flanks of the ridges, turbidities from the abyssal plains often finger into the fracture zone. The fracture zones form long curvilinear features throughout the oceans and are mappable because of morphology alone. *See* RIFT VALLEY.

The numerous strike-slip faults, which offset segments of the mid-oceanic ridge system, are classical examples of transform plate boundaries. The San Andreas system of California, which offsets portions of the East Pacific Rise, is probably the best-known example. Transform faults show apparent lateral displacements of many tens or even hundreds of kilometers. Occasionally extension, and consequent igneous activity, may occur at a transform fault.

Transform faults are parallel to the direction of relative motion between the bounding plates. According to a theorem of spherical geometry, if plate 1 is moving with respect to plate 2 on a sphere, the instantaneous relative motion may be represented by the rotation of one plate with respect to the other about a single stationary pole. Circles drawn on the sphere concentric to the pole of relative motion will be parallel to the direction of instantaneous relative motion. Great circles drawn perpendicular to these concentric circles will intersect at the pole of instantaneous relative motion. The transform faults that separate block 1 from block 2 are parallel to the direction of present-day relative motion and therefore must lie on circles that are concentric to a pole of relative motion of 1 with respect to 2. Great circles drawn perpendicular to the transform faults will intersect at the pole of relative motion. In this way, the pole of relative motion may be found for any two plates whose common boundary consists in part of transform faults.

The instantaneous rate of relative motion is an angular rate of rotation of 1 with respect to 2 about an axis through the pole of relative motion and the center of the Earth. The linear rate of relative motion R (in centimeters per year) will then be proportional to the sine of the colatitude with respect to the instantaneous pole,as shown in the equation below,

$$R = a \sin\theta \frac{d\phi}{dt}$$

where a = the radius of the Earth, 2.4×10^8 in. (6×10^8 cm); θ = the colatitude of the point at which the rate is to be determined with respect to the instantaneous pole; and $d\phi/dt$ = the angular rate of motion of 1 with respect to 2.

In this way, spreading-rate data calculated from the distribution of magnetic lineations near the axis of a spreading ridge-transform system may be used to compute a pole of relative motion. Note that in this case the pole is not strictly instantaneous because magnetic lineations covering a finite period must be used. It has been found that the poles and rates of relative motion tend to remain nearly constant, perhaps changing quite slowly over long time intervals. Thus, poles and rates computed in the above fashion may often be used to describe the relative motion for several tens of millions of years.

The fracture zones may often be treated as fossil transform faults. To obtain paleopoles of relative motion, segments of fracture zones spanning the same age and from the same plate must be used. Perpendicular great circles drawn to these segments give a pole of relative motion for that time interval in a reference frame fixed with respect to that plate. The magnetic anomaly lineations abutting the fracture zone are used to determine the rate of motion. By calculating a time sequence of poles and rates using time sequential segments of fracture zones, the spreading history of an entire ocean basin may be determined. *See* FAULT AND FAULT STRUCTURES; TRANSFORM FAULT.

Triple junctions. The point where the boundaries between three pairs of plates join is known as a triple junction. Because the Earth's lithosphere shell is segmented into a mosaic, the boundary between any two plates must end in a triple junction. Quadruple junctions are theoretically possible, but except under very unusual circumstances they degenerate immediately into two triple junctions. Figure 1 shows a number of triple junctions that vary as to the types of boundaries that meet and as to the geometry.

When three bodies are moving with respect to each other across a spherical surface, if the relative motion of each of two pairs of the plates is known, the relative motion of the third pair may be determined. In the case of a spherical surface, the relative motion between two plates may be represented by a vector through the pole of instantaneous relative motion, with the length of the vector proportional to the rate of rotation. The relative motion of three plates or bodies on a sphere may be represented by three such vectors. Each vector gives the instantaneous relative motion of two of the three plates. Thus, three such vectors describe completely the relative motion of the three plates and must add to zero.

The orientation of the boundary between these plates will determine the type of interaction at the boundary (that is, spreading, subduction, or transform). There are a number of possible triple junctions. The Cocos, Pacific, and Nazca plates join at a ridge-ridge-ridge triple junction (Fig. 1). This is an example of a stable triple junction, that is, the triple junction moves in a constant direction and at a constant speed with respect to each of the plates. Other triple junctions are unstable or transitional, such as a ridge-ridge-transform triple junction which will change immediately to a transform-ridge-transform triple junction. The most geologically important triple junctions are those that migrate along a boundary. Consequently, the boundary along which the triple junction moves will experience a change in the direction of relative motion. If the change in relative motion is radical, there may be an accompanying change in tectonic style. This type of plate interaction occurred at the western boundary of the North American Plate during the Cenozoic, and has resulted in a time transgressive change, from south to north, from subduction to strike slip. This is reflected in the geologic record as a change from compressive tectonics with thrusting, folding, and Andean or island arc volcanism to strike slip, as occurs along the San Andreas Fault.

Paleogeography. Because the accretion of new lithosphere is bilaterally symmetric, the associated magnetic lineations must also be bilaterally symmetric across a ridge axis. Once formed, oceanic crust is rarely disturbed by horizontal shear, but each magnetic lineation preserves the shape of the mid-oceanic ridge axis and offsetting transforms at the time of its formation. The fact that lineations of the same age from opposite sides of a ridge axis may be fitted together is considered evidence of the undisturbed nature of the oceanic crust.

The North and South Atlantic oceans were formed by the rifting and drifting of the continents that now surround the Atlantic. These continents have been rifted as part of the separating lithospheric plates since the Triassic as the Atlantic Ocean slowly opened. The magnetic lineation pattern in the North Atlantic records this history of separation of Africa, Europe, and North America. When magnetic lineations of the same age from opposite sides of the ridge axis are fitted together, the relative positions of the continents at the time the lineations formed are obtained. This is equivalent to reversing the process of sea-floor spreading. In effect, the material younger than the two lineations is removed, and the older portions of the plates with the continents move toward each other until the two lineations fit together.

This technique is applicable to any ocean where separation at ridge axes has led to the passive rifting of the surrounding continents, such as the southeast and central Indian Ocean, the South and North Atlantic, the Labrador and Norwegian seas, and the Arctic Ocean.

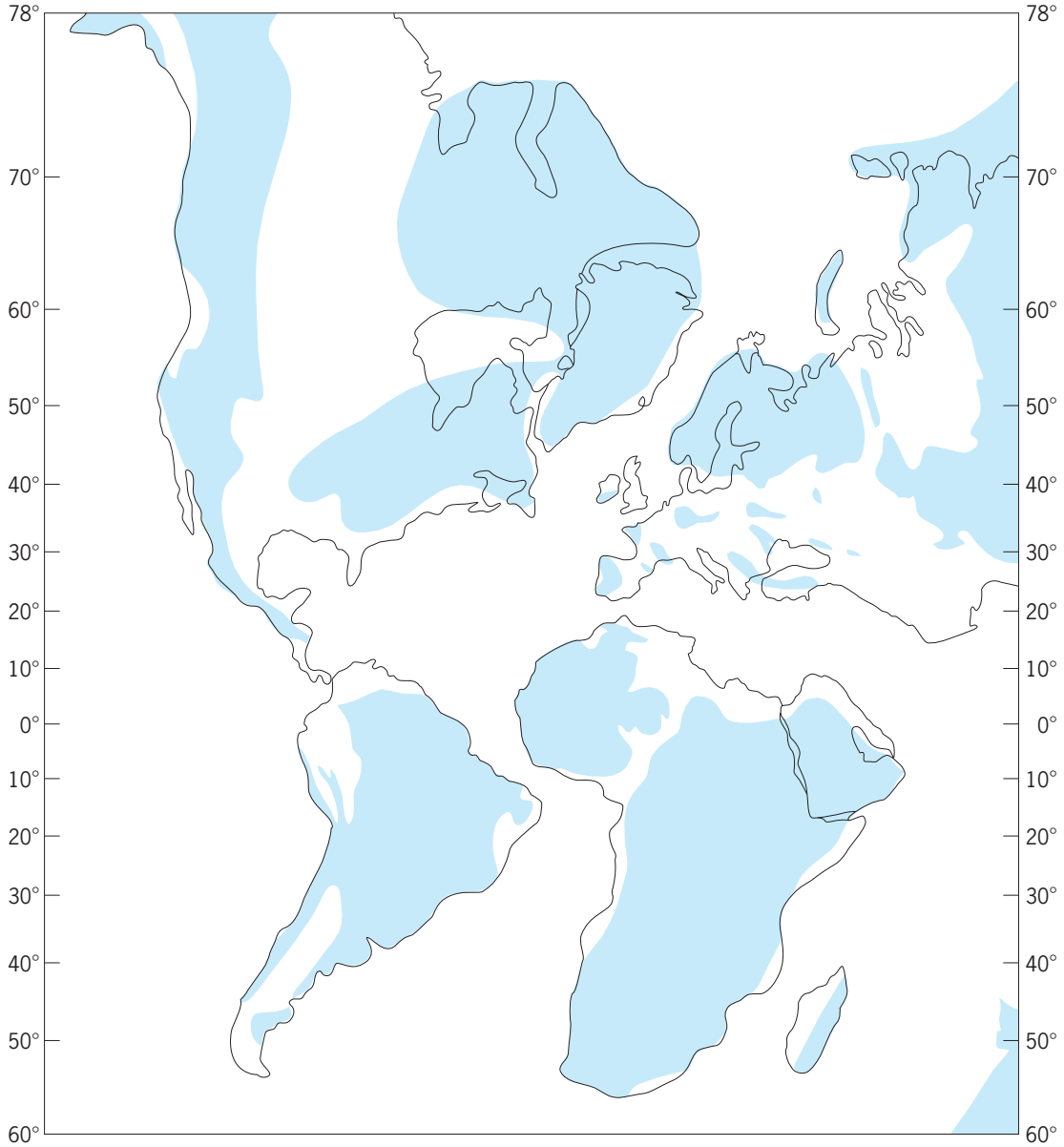

Fig. 7. Map illustrating the paleogeographic arrangement of the Atlantic continents during Coniacian time. Sea level may have been over 300 m (1000 ft) above present level at that time because of the volume of the mid-oceanic ridge system. Shaded areas are portions above sea level.

The above technique gives the position of the continental landmasses relative to each other, not their position relative to the rotational axis or paleomagnetic data. It is assumed that the Earth's magnetic field, now approximated by an axial geocentric dipole, has been similarly oriented throughout geologic time. Subaerial volcanic rocks acquire a remanent magnetism in the same way as oceanic crust, by cooling below the Curie point. Subaerial sedimentary rocks may also be remanently magnetized by the process of detrital remanent magnetization. As iron-bearing particles settle from moving currents, their orientation is influenced by the prevailing magnetic field. By measuring the direction of the remanent magnetism of both igneous and sedimentary rocks whose ages can be determined, the position of the paleomagnetic pole for that age can be found. From these data, the paleolatitude may be found for the set of continental landmasses whose positions relative to each other are known (**Fig. 7**). *See* PALEOGEOGRAPHY; ROCK MAGNETISM.

Sea-level changes. Many important geologic phenomena may in part be related to plate tectonics processes, for example, changes in sea level and climate and changing patterns of evolution. The geologic record contains evidence of almost continuous eustatic sea-level changes throughout the Phanerozoic. Several large and rapid sea-level changes can be attributed to fluctuations in the volume of massive continental glaciers. However, it is apparent that large sea-level changes (greater than 300 ft or 100 m) also occurred during periods when continental glaciers were small or nonexistent.

Several factors affect the volume of the ocean

basins or the ocean waters and can thus alter eustatic sea level. One of the most important of these in terms of magnitude and rate of sea-level change is variation in the volume of the mid-oceanic ridge system. As explained above, once formed, the sea floor subsides systematically with time, with all oceanic crust following the same exponential subsidence curve. The volume of the mid-oceanic ridge system is quite large. If sea-floor spreading were to cease today, 70 m.y. from now the ridges would have subsided sufficiently to decrease the depth of the water over the abyssal area by about 1500 ft (450 m). The continental freeboard would increase by 1100 ft (320 m).

The volume of the mid-oceanic ridge system may be altered in several ways, one of which is to change the spreading rate. Because all ridges follow the same subsidence curve (which is a function of time only), age-versus-depth relationships are the same for all ridges. This means that if two ridges have been spreading at a constant but different rate for 70 m.y., the ratio of their volumes per unit length will equal the ratio of their spreading rates. If the spreading rate of the faster ridge is reduced to that of the slower ridge, volume/unit length of the larger ridge will gradually be reduced to that of the slower ridge. If spreading rates decrease, ridge volume also decreases, increasing the freeboard of the continents. The converse is also true: by increasing spreading rates, ridge volume will increase and the freeboard of the continents will be reduced. Several other ways of changing the volume of the mid-oceanic ridge system exist: ridges may be destroyed; segments of the ridge system may be subducted; and new ridges may be created by continental rifting and new rifting within ocean basins.

Both the length of the mid-oceanic ridge system and the spreading rate at its various segments have varied considerably during Phanerozoic time. These changes have caused large variations in sea level. For example, by using magnetic anomaly data to calculate spreading rates and ridge lengths back to the Upper Cretaceous, it has been estimated that the sea level then may have been as much as 900 ft (300 m) above present (**Fig. 8**).

Hot spots and mantle plumes. The existence of convective plumes originating in the deep mantle from below the level of the asthenosphere and rising to the bottom of the lithosphere has been proposed. About 14 major convective plumes are believed to exist today (Fig. 1). These plumes are believed to be nearly stationary with respect to each other, and hence they may be used as reference points with respect to which all the plates are moving. Mantle plumes, rather than convection currents, may also be the driving mechanism for plate motion. The latter two hypotheses are not necessary conditions for

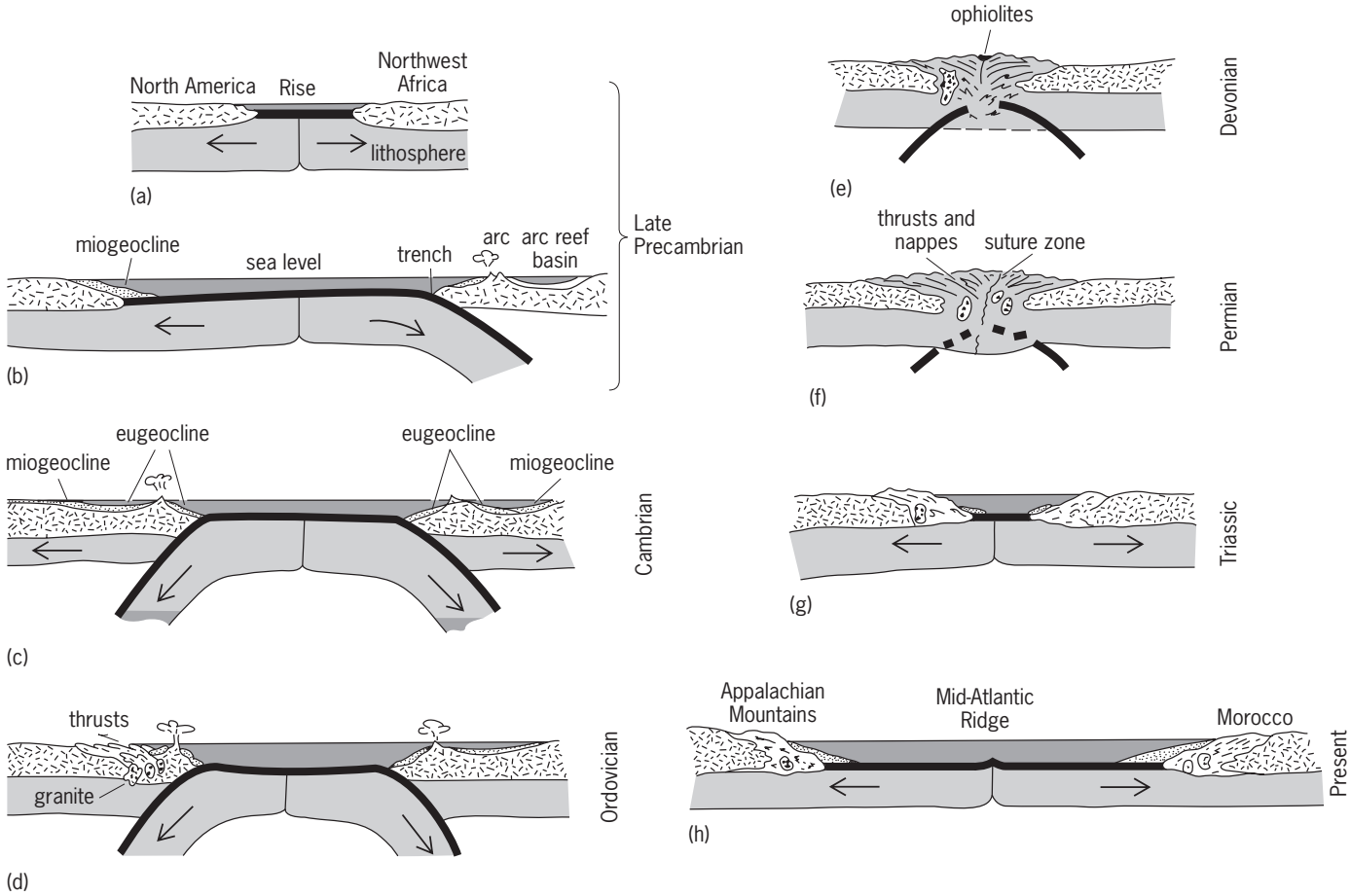

Fig. 8. Schematic representation of (*a–h*) the sequence of changes in the mid-oceanic ridge system in the plate tectonics history of the Appalachian orogenic belt. (*After K. C. Condie, Plate Tectonics and Crustal Evolution, Pergamon, 1976*)

the validity of the first. In other words, there may be mantle plumes that rise to the bottom of the lithosphere, but they may not be stationary with respect to each other and may not furnish the plate driving mechanism. The major clue to the existence of hot spots is the lines of intraplate volcanoes that are left as a trace of the passage.

As a plate passes over a hot spot, the hot spot burns its way through the plate. This releases volatile and eruptive magma on the plate surface. A classic example of the "volcanic" consequences of passage of a plate over a hot spot is the Hawaiian-Emperor Seamount chain. Hawaii and the islands immediate to it are volcanically active at present. The rest of the islands and seamounts of the Hawaiian and Emperor chains are inactive, but all are of volcanic origin. The islands and seamounts are found to be sequentially older west-northwestward along the Hawaiian chains and are progressively older north-northwestward along the Emperor Seamount chain. As an explanation, it has been hypothesized that the Pacific Plate moved first north-northwesterly and then west-northwesterly over a single hot spot. Other seamount chains in the Pacific, such as the Line Islands, the Tuamotu Archipelago, and the Austral Seamount chain, are presumed to be of similar origin. *See* HOT SPOTS (GEOLOGY); MAGMA; OCEANIC ISLANDS.

In the Atlantic, features that are regarded as major hot-spot traces are Iceland and the Iceland Faroes Ridge, the New England Seamount chain, the Columbia Seamount chain, the Rio Grande Rise, and Walvis Ridge; and in the Indian Ocean, the Ninety-East Ridge is an analogous feature. The age progression along each of these features has been estimated; the theory that these hot spots have remained fixed in position relative to each other has been tested. The results indicate that some relative motion may occur, but it appears to be an order of magnitude less than the generally observed rate of interplate motion.

The hot spot hypothesis is a significant complement to plate tectonics. There is little doubt that hot spots or plumes of some type do occur. Intraplate volcanism often seems to be a manifestation of their existence. If, in addition, the hot spots prove to be essentially stable with respect to the rotational axis, they may be useful as a reference frame. They could serve as a latitudinal constraint in addition to paleomagnetic data. They could also be very useful in finding the relative motion between plates that have been separated by subduction zones or transform faults (that is, Pacific–North America relative motion).

Earth history. Not only does plate tectonics theory explain the present-day distribution of seismic and volcanic activity around the globe and physiographic features of the ocean basins such as trenches and mid-oceanic rises, but most Mesozoic and Cenozoic mountain belts appear to be related to the convergence of lithospheric plates. Two different varieties of modern mobile belts have been recognized: cordilleran type and collision type. The Cordilleran range, which forms the western rim of North and South America (the Rocky Mountains, Pacific Coast ranges, and the Andes), has for the most part been created by the underthrusting of an ocean lithospheric plate beneath a continental plate. Underthrusting along the Pacific margin of South America is causing the continued formation of the Andes. The Alpine-Himalayan belt, formed where the collision of continental blocks buckled intervening volcanic belts and sedimentary strata into tight folds and faults, is an analog of the present tectonic situation in the Mediterranean, where the collision of Africa and Europe has begun. *See* CORDILLERAN BELT; MOUNTAIN SYSTEMS; OROGENY.

Abundant evidence suggests that sea-floor spreading, continental drift, and plate tectonics have occurred for at least the past 600 m.y., during Phanerozoic time. Furthermore, it is probable that plate tectonics phenomena have dominated geologic processes for at least 2.5 billion years (b.y.). Present-day ocean basins are young (post-Paleozoic) features, and their origin is adequately explained by sea-floor spreading. The origin and evolution of the continental blocks, substantial portions of which are Precambrian in age, probably may be explained by accretion at subduction zones. Continental blocks, though largely composed of granite and granodiorite, appear to be mosaics of ancient mobile belts that progressively accreted laterally, vertically, or both, through time. Many late Precambrian and Paleozoic mobile belts, such as the Appalachian-Caledonian belt of the North Atlantic region, are interpretable in terms of plate tectonics mechanisms (Fig. 8). Initial rifting breaks a preexisting supercontinental landmass into two or more fragments that drift apart at the flanking margins of the growing ocean basin. At the margin of the growing basin, continental shelves, continental rises, and abyssal plains accumulate sediments. Cessation of sea-floor spreading and the development of new convergent margins along one or both sides of the ocean basin lead first to cordilleran-type orogeny and subsequently to continental collision. These orogenies suture or stitch together the original continental fragments and laterally buckle the intervening sediments and volcanic belts into folds and faults, initiate metamorphism, and generate volcanic and plutonic magmas. Belts may later be bisected by the development of new divergent plate margins that produce ocean basins whose axes may cut across the older mountain chains. This seems to be the case with the Appalachian-Caledonian belt, where the Mesozoic-Cenozoic development of the Atlantic Ocean has separated the chain into the Appalachians in North America and the Caledonides in Greenland, Great Britain, and Scandinavia. Most Paleozoic mobile belts, and many Precambrian belts at least 2.5 b.y. old, can be interpreted in an analogous fashion, for they contain the various petrogenetic associations of ocean opening and suture: ophiolites, blueschist terranes, mélanges, and a suite of sedimentary rocks whose composition, texture, and distribution suggest deposition in trenches and inland seas, on the abyssal plains of ocean basins,

and along continental shelves and rises. Because the Earth itself has evolved, chemical and petrologic aspects of these more ancient mobile belts often differ in detail from more modern analogs. Their organization leaves little doubt that they were formed by plate tectonics processes.

Plate tectonics is considered to have been operative as far back as 2.5 b.y. Prior to that interval, evidence suggests that the plate tectonics may have occurred, although in a markedly different manner, with higher rates of global heat flow producing smaller convective cells or more densely distributed mantle plumes which fragmented the Earth's surface into numerous small, rapidly moving plates. Repetitious collision of plates may have accreted continental blocks by welding primitive "greenstone island arc terranes" to small granitic subcontinental masses. *See* CONTINENTS, EVOLUTION OF; GEODYNAMICS.

Walter C. Pitman III

Details of Processes at Convergent Margins

Convergent continental and island arc margins encircle the Pacific Ocean and occur locally in all other oceans (**Fig. 9**). At convergent margins, two of the Earth's lithospheric plates are directed against each other. Normally, one plate consists of dense oceanic lithosphere which underthrusts, or subducts, beneath less dense continental or island arc lithosphere (**Fig. 10**).

Structure and tectonic units. The thrust fault between the plates is the major tectonic feature of a subduction zone, and opposing plate motion causes great earthquakes that begin at 10–20 km (6–12 mi) depths in the seismogenic zone. As the plates shear past one another, sediment and rock of the subducting plate may be added to the upper plate and upper-plate material may be eroded, introducing terrigeneous material into the subduction zone. The dominant material transfer is a basis for dividing convergent margins into accretionary and erosive types. In the absence of a standard nomenclature, the following is used here. The upper plate of convergent margins is commonly formed by four units, consisting of a frontal prism that is sometimes backed against an accreted prism, which in turn backs against the core rock framework and is overlain by a slope apron. The lower plate consists of igneous rock overlain by oceanic and trench fill sediment. Sediment that stays attached to the lower plate when introduced into the subduction zone forms the subduction channel that is capped by the interplate thrust. *See* MARINE GEOLOGY.

Fig. 9. Plate tectonic map of the circum-Pacific region showing major zones of plate convergence and spreading ridges. Arrows and numbers indicate the direction and the rate (cm/yr; 1 cm = 0.4 in.) of plate movement. The thrusts of the zones of convergence are marked with teeth on the upper plate. (*After G. W. Moore, Plate-Tectonic Map of the Circum-Pacific Region: Explanatory Notes, American Association of Petroleum Geolgists, 1982*)

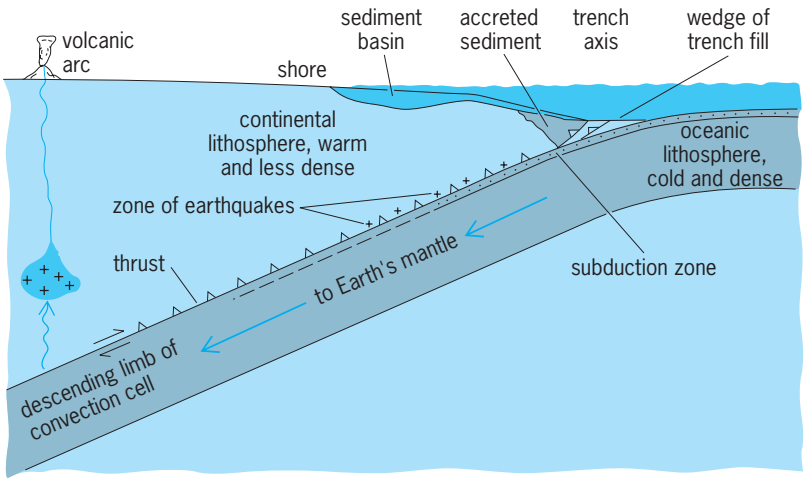

Fig. 10. Cross section through a typical convergent margin. The relatively dense oceanic lithosphere sinks beneath the relatively light continental lithosphere and forms the descending limb of a convection cell in the Earth's mantle.

The frontal prism begins at the initial contact between converging plates as represented by the first fault that displaces the trench axis sea floor, referred to as the deformation front. Rapid deformation characterizes the frontal prism. The prism extends to either a more rigid core framework or a well-consolidated accreted prism.

Compressional deformation tectonically thickens the prism in contrast to less prominent active tectonism landward. The frontal prism consists of material derived from off-scraped trench fill, slope apron sediment, or fragmented basement rock that was detached at the basement–prism contact. Along more than 50% of all convergent margins where trench fill is no thicker than 1 km (0.6 mi), frontal prisms are less than 20 km (12 mi) wide. Once they reach 25–30 km (15–20 mi) in width, material at the back of the prism has consolidated, and it becomes sufficiently rigid to become part of the accreted prism. Thus, the contact between frontal prism and accreted wedge migrates toward the trench axis. Frontal prisms widen continents only a small amount in contrast to accreted prisms.

The accreted prism, 30 to more than 100 km (20 to more than 60 mi) wide, is a body of oceanic and trench fill sediment that was transferred from the subducting lower plate at the frontal prism and became attached to the upper plate framework. Accreted prisms develop from the frontal prism and consist of consolidated and rigid sedimentary rock with porosities of about 15% or less. The accreted prism adds to the width of a continent or island arc and stores material for longer periods. Large accretionary prisms involve a margin configuration and plate convergence constant for tens of millions of years. They are common where sediment fill in the trench axis is more than 1 km (0.6 mi) thick and rates of orthogonal convergence are less than 5 km/m.y. Examples are found along the Lesser Antilles and the Nankai, Makran, and Indonesian margins.

The core framework is commonly referred to as basement, and it includes diverse igneous as well as metamorphic and occasional sedimentary rock bodies. The framework basement exposed along coasts is a strong and rigid rock mass that can be followed tens of kilometers seaward with geophysical methods and sampling. Along more than half of the global length [about 41,000 km (25,000 mi)] of convergent margins, the framework extends to within 10–15 km (6–9 mi) of the trench axis. Many of the sedimentary rock bodies of the framework basement are fossil accreted wedges, as along the Japan and Alaskan margins.

The slope apron is a sediment sequence that overlies the margin framework and accreted prism. It extends from the edge of the shelf to the frontal prism and can be from a few meters to 5 km (3 mi) thick. The steepness of the continental slope causes gravity sliding and a general downslope migration of sediment. Seismic imaging of slope-apron stratification and high-resolution sea-floor bathymetry [100 m (330 ft) or less], show the lack of active deformation that indicates the tectonic stability of the margin framework and accreted prism relative to the frontal prism. Stability is also indicated by a sea-floor morphology with integrated canyons that are disrupted and disappear over the frontal prism.

The plate interface is underlain by a layer from about 200–1200 m (about 660–4000 ft) in thickness called the subduction channel. This layer consists of subducted trench sediment and eroded clastic material on the igneous ocean crust and is commonly detected as a zone of low seismic velocity, relative to its bounding rock masses. The low velocity is interpreted as indicating a fluid-rich unit. Subduction beneath the frontal prism inflicts a load on the porous sediment in the subduction channel, which elevates the pore fluid pressure because the prism's permeability does not allow the fluid to escape as quickly as the load is applied. Fluid pressure can approach the weight of the overlying material. Such pressures greatly reduce friction and allow less consolidated materials to be subducted rather than scraped off at the deformation front. Farther down, the subduction zone temperature increases to levels where some clay minerals alter and release chemically bound water into the subduction channel. Elevated pressure in a fluid-rich layer is an explanation for reduced seismicity, and the manner of slip along the plate interface thrust is considered stable. The unstable slip released during large earthquakes and tsunamis occurs at ~10 km (6 mi) or greater depths in the portion of the subduction zone called the seismogenic zone. *See* SEISMOLOGY; TSUNAMI.

The subducting lower plate commonly bends into the trench axis and breaks into extensional or normal faults. This faulting produces voids that fill with sediment carried down the subduction zone on the lower plate. Some normal faults are imaged seismically through the entire oceanic crust, and could form aquifers that conduct seawater into the lower crust. Seawater in contact with lower crustal rock can change the mineralogy to weaker serpentine minerals and affect the mechanical behavior and

geochemical evolution of the subducting lithosphere. The cold, negatively buoyant subducting plate becomes a driving force for convection in the Earth's mantle, which is the layer between the lithosphere and the core. Ultimately, subduction of oceanic plates into the mantle is responsible for much of the tectonism expressed at the Earth's surface.

Tectonic processes. The dynamic character of convergent margins is evident from earthquakes and tsunamis. Convergent margins begin in the seismogenic zone at depths beyond the reach of direct observation even with scientific drilling. However, the mechanisms involved leave a geology to interpret that indicates the nature of dynamic processes involved. Accretionary mass addition has an erosional mass subtraction counterpart. The geology of material added to the upper plate is more easily observed than the record of massive material removed from a continent or island arc and carried out of sight. An important but difficult record to decipher from geologic evidence is the role of fluid in controlling tectonism and seismicity.

Erosional convergent margins. Subduction erosion is predominant along more than 60% of convergent margins globally. At erosional convergent margins, the net volume of the margin's upper plate decreases during subduction. Evidence of missing rock masses includes finding (1) the roots of ancient volcanic arcs along the coast or more than 100 km (60 mi) closer to the trench axis than where they formed; (2) large-scale and long-term subsidence to trench depths of unconformities that formed in shallow-water environments; and (3) the migration of volcanic arcs hundreds of kilometers inland, indicating that the continental slope retreated a similar distance. Drill cores at seven Pacific convergent margins show 3–5 km (2–3 mi) of subsidence during the past 15–20 m.y. This subsidence requires upper-plate thinning from erosion along the plate interface because sediment deposition during subsidence rules out regional seafloor erosion.

Erosional margins characteristically occur where plate convergence is relatively rapid (more than 4 km/m.y.) and sediment in the trench axis is less than 1 km (0.6 mi) thick. The lower plate of igneous oceanic crust and sediment bends downward into the trench axis, causing extensional structure diagrammed as tilted blocks or half-grabens (depressed, faulted blocks). The wedge-shaped upper plate consists of a continental crustal framework covered by a slope sediment apron. The margin framework is as strong at the coast, but downslope it fractures and develops extensional faults that break up the plate beneath the middle slope. Disruption of the framework structure results in mass wasting of the slope apron and basement. Debris from mass wasting migrates downslope and collects in a frontal prism beneath the lower slope. This thin weak apex of the frontal prism displays ridges and valleys at the sea floor that mimic the buried lower plate relief. A thrust fault bounds the underside of the frontal prism, but landward the plate interface becomes a subduction channel. In seismic images, the channel has a low-frequency–high-amplitude seismic reflectance that is 0.5–1.0 km (0.3–0.6 mi) wide, inferred to indicate a fluid-rich zone. Subduction beneath the frontal prism elevates fluid pressure in the lower plate and its sediment cover, and this pressure increases upward fluid migration along and across the plate interface. As fluid drains from the lower plate, its porosity is reduced, which consolidates and strengthens material below the interplate thrust, and at the same time fluid invading fractures in the upper plate weakens rock above the plate interface. Thus, the zone of minimum strength along the plate interface shifts upward. The shift results in a transfer of upper-plate material to the subducting lower plate, thereby thinning the plate. *See* GRABEN.

In the subduction erosional model, the upper-plate framework looses strength downslope as the plate thins and fractures dilate from invasion by pressured fluid. As the lower plate subducts, it continues bending, which increases lower-plate relief. Relief promotes fracturing in the upper plate from jostling as lower-plate relief subducts beneath it. As hundreds of kilometers of lower plate subducts, the jostling and dilation of fractures by pressurized fluid ultimately fragments the base of the upper plate. Detached fragments are dragged piece by piece onto the underthrusting lower plate, which carries them into the seismogenic zone.

Accretionary margins. Accretionary margins preferentially occur where convergence is less than 40 km/m.y. (25 mi/m.y.) and where sediment in the trench axis is thicker than 1 km (0.6 mi). They are distinguished from erosional margins by the accreted prism that develops from the frontal prism. The sediment wedge filling the trench axis is deposited rapidly and has a high initial porosity (about 70%). As the wedge converges against the upper plate, horizontal force creates small compressional faults, which do not extend to the sea floor, in the lower part of the trench-sediment wedge, and pore fluid pressure increases. Before faults break the sea floor, fluid drains from the trench-sediment wedge through intergranular permeability. Closer to the deformation, front faults reaching the sea floor provide more efficient dewatering channels. The deformation front develops when upper sediment layers consolidate sufficiently to break in thrust faults. Faults near the deformation front may extend through most of the wedge sediment section, but a horizontal fault above the igneous basement and parallel to bedding (a decollement) that forms the interplate thrust fault intercepts them (**Fig. 11**). Sediment below the thrust fault subducts beneath the upper plate as part of the lower plate. Where it has been drilled, subducted sediment is little deformed and indicates low interplate friction, along with other observations. Once faulting landward of the deformation front begins, fluid drains rapidly through fracture permeability in the faults. The majority of original pore fluid in accreted sediment (80%) is expelled in approximately the first 5–10 km (3–6 mi) of the accretionary frontal prism. Accretionary frontal prisms thicken rapidly by

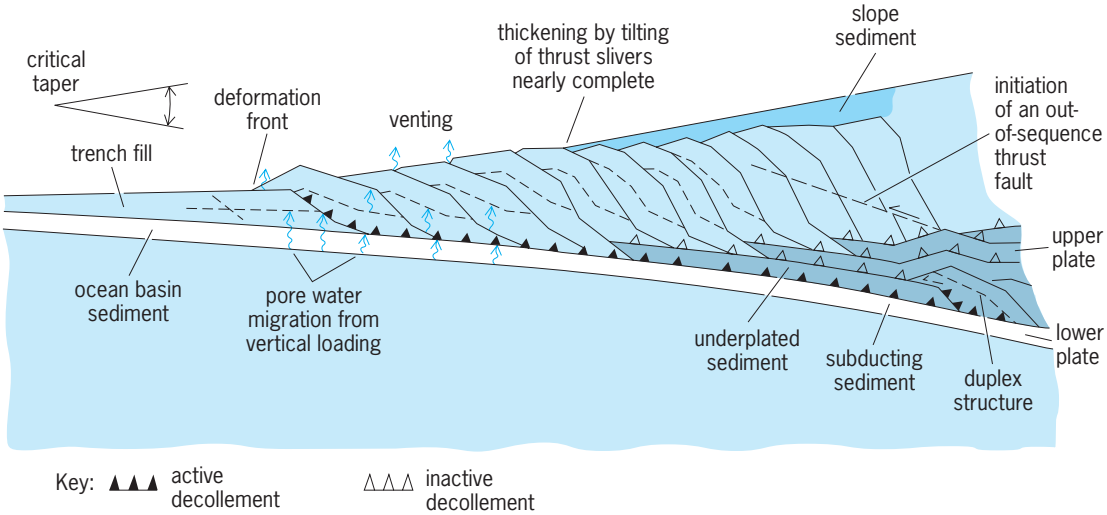

Fig. 11. Section across an accretionary convergent margin. The deformation front marks the partitioning of subducted and accreted sediment, a division that commonly occurs along the base of the trench fill. Rapid loading squeezes water from voids in the sediment, ultimately draining with the water from vents at the sea floor.

thrust faulting, and it is here that the greatest permanent strain deforms sedimentary bodies. In the subducted sediment and igneous crust, fluid is released farther down the subduction zone.

In many accretionary margins, the latest seismic and high-resolution bathymetric data show segmentation of tectonic activity not previously noted. A change in tectonic character separates the accretionary frontal prism from the accreted prism. The rapidly deforming frontal prism has a rough sea-floor morphology. The accreted prism deforms minimally, allowing deposition of a little deformed slope apron across it. The accreted prism is stronger and more rigid than the frontal prism, and the contact between them is a fault zone that commonly has sea-floor expression. This segmentation is not recognized along all accretionary margins, and it is difficult to resolve without the latest high-resolution geophysical data. Along the Nankai margin off Japan, the boundary between frontal prism and accreted prism is termed a splay fault, and similar features are resolved along the Sunda and Makran margins.

The concept of critical taper helps explain the thickening of an accretionary wedge. Taper refers to the angle between the upper and lower prism surfaces, and it depends on the strength of the material and the interplate friction. When the critical taper is exceeded, the slope becomes overcritical and fails by slumping to restore the optimum taper. When taper is undercritical, accretion restores it to the optimum taper for conditions at the time. To accomplish tectonic thickening, the accreted thrust packets are stacked and rotated, but only to a limiting angle. Once rotation and stacking of the detached packets can no longer thicken the wedge, thickening can be accomplished by the addition of material from below (underplating), by subsequent thrust faulting at a lower angle across previous structure (out-of-sequence thrusting), or by thrusting that is contained within a lower layer (duplexing). These processes occur at depths where clear seismic imaging of structures is at the limits of the current technology. They are currently modeled to explain taper related to upper plate thickening landward of the frontal prism.

The mass flux at convergent margins is important for understanding arc volcanism and the origin of continental crust. Subduction zones return crust to the mantle from which is once originated, but in an altered state. The past tendency to view subduction zones as primarily areas of growth has diminished as the extent of subduction zone erosion becomes known. The volume of continental crust returned to the mantle is small compared to its entire volume, but sufficient to impart the distinct continental geochemistry observed in mantle rock. Erosion is one argument for the paucity of very old exposed continental crust, functioning as a balancing process to have kept the growth of continents at their approximate size during much of the Earth's history (since pre-Cambrian time, about 2.5 billion years ago). Roland von Huene

Bibliography. A. W. Bally (ed.), *Seismic Expression of Structural Styles*, Amer. Ass. Petrol. Geol. Stud. Geol. 15, 1983; N. L. Bangs et al., Evolution of the Nankai Trough decollement from the trench into the seismogenic zone: Inferences from three-dimensional seismic reflection imaging, *Geology*, 32(4):273–276, 2004; K. C. Condie, *Plate Tectonics and Crustal Evolution*, 4th ed., 1997; *Continents Adrift and Continents Aground: Readings from Scientific American*, introduction by T. J. Wilson, 1976; R. von Huene, C. R. Ranero, and P. Vannucchi, Generic model of subduction erosion, *Geology*, 32(10):913–916, 2004; H. Kopp and N. Kukowski, Backstop geometry and accretionary mechanics of the Sunda margin, *Tectonics*, 22(6):1072, 2003; J. C. Moore et al., Tectonics and hydrogeology of the northern Barbados Ridge: Results from Ocean Drilling Program Leg 110, *Geol.*

Soc. Amer. Bull., 100(10):1578–1593, 1988; M. J. Selby, *Earth's Changing Surface*, 1985; C. Seyfert (ed.), *Encyclopedia of Structural Geology and Plate Tectonics*, 1987; B. F. Windley, *The Evolving Continents*, 3d ed., 1995.

Plateau

Any elevated area of relatively smooth land. Usually the term is used more specifically to denote an upland of subdued relief that on at least one side drops off abruptly to adjacent lower lands. In most instances the upland is cut by deep but widely separated valleys or canyons. Small plateaus that stand above their surroundings on all sides are often called tables, tablelands, or mesas. The abrupt edge of a plateau is an escarpment or, especially in the western United States, a rim. In the study of landform development the word plateau is commonly used to refer to any elevated area, especially one underlain by nearly horizontal rock strata, that once had a smooth surface at high level, even though that surface may since have been largely destroyed by valley cutting. An example is the now-hilly Appalachian Plateau of western Pennsylvania, West Virginia, and eastern Kentucky. *See* ESCARPMENT; PLAINS.

Among the extensive plateau lands of the world are the Colorado and Columbia plateaus of the western United States, the plateau of southeastern Brazil, the Patagonian Plateau of southern South America, the Central Siberian Plateau, and the Deccan Plateau of peninsular India.

Edwin H. Hammond

Platinum

A chemical element, Pt, atomic number 78, and atomic weight 195.09. Platinum is a soft, ductile, white noble metal. The platinum-group metals—platinum, palladium, iridium, rhodium, osmium, and ruthenium—are found widely distributed over the Earth. Their extreme dilution, however, precludes their recovery, except in special circumstances. For example, small amounts of the platinum metals, palladium in particular, are recovered during the electrolytic refining of copper. *See* IRIDIUM; OSMIUM; PALLADIUM; PERIODIC TABLE; RHODIUM; RUTHENIUM.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Physical properties of platinum

Properties	Value
Atomic weight ($^{12}C = 12.00000$)	195.09
Naturally occurring isotopes and % abundance	190, 0.0127%
	192, 0.78%
	194, 32.9%
	195, 33.8%
	196, 25.3%
	198, 7.21%
Crystal structure	Face-centered cubic
Lattice constant *a* at 25°C, nm	0.39231
Thermal neutron capture cross section, barns	8.8
Common chemical valence	2, 4
Density at 25°C, g/cm^3	21.46
Melting point	1772°C (3222°F)
Boiling point	3800°C (6900°F)
Specific heat at 0°C, cal/g	0.0314
Thermal conductivity, 0–100°C, cal cm/cm^2 s°C	0.17
Linear coefficient of thermal expansion, 20–100°C, μin./in./°C	9.1
Electrical resistivity at 0°C, microhm-cm	9.85
Temperature coefficient of electrical resistance, 0–100°C/°C	0.003927
Tensile strength, 1000 lb/in.^{2}	
Soft	18–24
Hard	30–35
Young's modulus at 20°C	
lb/in.^{2}, static	24.8×10^6
lb/in.^{2}, dynamic	24.5×10^6
Hardness, Diamond Pyramid Number (DPN)	
Soft	37–42
Hard	90–95

The platinum-group metals have wide chemical use because of their catalytic activity and chemical inertness. As a catalyst, platinum is used in hydrogenation, dehydrogenation, isomerization, cyclization, dehydration, dehalogenation, and oxidation reactions. *See* CATALYSIS; ELECTROCHEMICAL PROCESS.

Platinum is not affected by atmospheric exposure, even in sulfur-bearing industrial atmospheres. Platinum remains bright and does not visually exhibit an oxide film when heated, although a thin, adherent film forms below 450°C (840°F). Platinum may be worked to fine wire and thin sheet, and by special processes, to extremely fine wire. Important physical properties are given in the **table.**

Platinum can be made into a spongy form by thermally decomposing ammonium chloroplatinate or by reducing it from an aqueous solution. In this form it exhibits a high absorptive power for gases, especially oxygen, hydrogen, and carbon monoxide. The high catalytic activity of platinum is related directly to this property. *See* CRACKING; HYDROGENATION.

Platinum strongly tends to form coordination compounds. Platinum dioxide, PtO_2, is a dark-brown insoluble compound, commonly known as Adams catalyst. Platinum(II) chloride, $PtCl_2$, is an olive-green water-insoluble solid. Chloroplatinic acid, H_2PtCl_6, is the most important platinum compound. *See* COORDINATION CHEMISTRY.

In the glass industry, platinum is used at high temperatures to contain, stir, and convey molten glass.

In the electrical industry, platinum is used in contacts and resistance wires because of its low contact resistance and high reliability in contaminated atmospheres. Platinum is clad over tungsten for use in electron tube grid wires. In the medical field, the simple coordination compounds cisplatin and carboplatin are two of the most active clinical anticancer agents. In combination with other agents, cisplatin is potentially curative for all stages of testicular cancer. Both agents are used for advanced gynecologic malignancies, especially ovarian tumors, and for head and neck and lung cancers. Carboplatin was developed in attempts to alleviate the severe toxic side effects of the parent cisplatin, with which is shares a very similar spectrum of anticancer efficacy. *See* CANCER (MEDICINE); CHEMOTHERAPY.

Henry J. Albert; Nicholas Farrel

Bibliography. L. Cerveny, *Catalytic Hydrogenation*, 1986; J. R. Davis (ed.), *Metals Handbook: Desk Edition*, 2d ed., ASM International, 1998; F. R. Hartley (ed.), *Chemistry of the Platinum Group Metals: Recent Developments*, 1991; L. R. Kelland and N. Farrell (eds.), *Platinum-Based Drugs in Cancer Therapy*, 2000.

Platyasterida

A Paleozoic (Ordovician and Devonian) order of sea stars of the class Asteroidea; only two genera have been recognized. Platyasteridans have relatively elongate flat arms; broad, transversely aligned ventral arm plates; and paxilliform aboral plates (see **illus.**). The morphology suggests living members of the Luidiidae, and based on those similarities, some authors contend that luidiids are surviving platyasteridans. Fundamental differences in ambulacral system construction have been found between modern asteroids and all of the well-known early Paleozoic asteroids, including the platyasteridans. Similarities between platyasteridans and modern asteroids are atributable to evolutionary convergence. *See* ASTEROIDEA; ECHINODERMATA; PAXILLOSIDA.

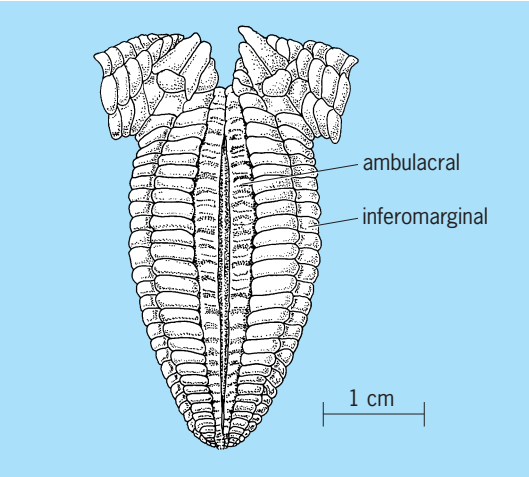

Ventral view of an arm and a portion of the disk of *Platanaster ordovicus*. Note lateral alignment of the arm plates. (*After W. K. Spencer and C. W. Wright, Asterozoans, in R. C. Moore, Treatise on Invertebrate Paleontology, Pt. U, Echinodermata 3, Geological Society of America and University of Kansas Press, 1966*)

Daniel B. Blake

Bibliography. D. B. Blake, A classification and phylogeny of post-Paleozoic sea stars (Asteroidea: Echinodermata), *J. Nat. Hist.*, 21:481–528, 1987.

Platycopida

An order of the Podocopa (Ostracoda) containing a single family, Cytherellidae, and two extinct superorders. Platycopids are small marine or brackish-water ostracods with an asymmetrical shell that is oblong, nearly rectangular, and laterally compressed; their six pairs of appendages are adapted for burrowing and filter feeding. Both fossil and living species have the inner lamella of each valve calcified only in the distal rim, if at all; if a narrow duplicature is developed, it is continuous with the outer lamella and does not form a vestibule (troughlike cavity) between the lamellae. Recent platycopids are distinct from all other ostracods: The larger valve is rabbeted to overlap the smaller around the free edge; the carapace lacks vertical lobation; none of the postoral appendages serve as legs; and the posterior end of the body is more clearly delimited and tapering than in other ostracods.

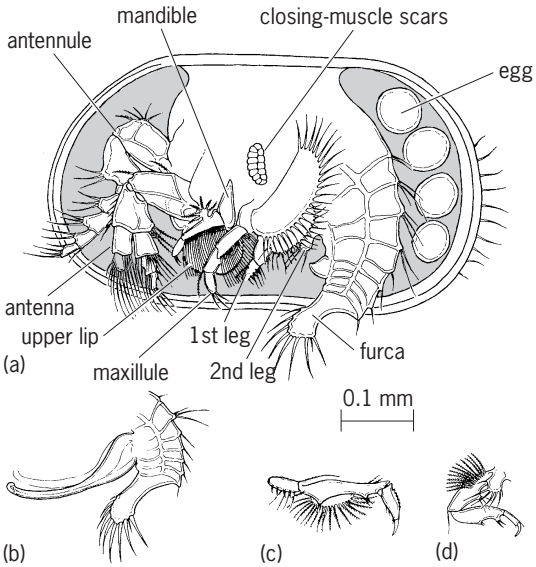

***Cytherella abyssorum*, an ostracod of the order Platycopida. (*a*) Female with left valve removed; eggs lie in posterior brood chamber outside body. (*b*) Posterior end of male body with paired furcae and long hemipenes. (*c, d*) Male first and second legs, both developed as claspers and very different from the corresponding appendages in the female. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. Q, 1961, and D. L. McGregor and R. V. Kesling, Contributions from the Museum of Paleontology, University of Michigan, 1969*)**

Species are dimorphic. The female carapace is wider near the rear and longer than that of the

male, providing a brood chamber behind the body. Scars of adductor muscles are arranged in a biserial aggregate (see **illus.**). The biramous antenna has both endopod and exopod developed for locomotion, and the paired lamelliform furcae may assist. All living species are benthic, and fossil forms probably were also. Some species live at great depths in the Atlantic Ocean, but others have been found in the Arabian, Caribbean, and Mediterranean seas in shallower water.

Unquestioned Platycopida are known from Jurassic to Recent. It has been proposed that ancestors of the platycopids were nondimorphic Paleocopida. The Cytherellidea combine a primitive, albeit elaborate, soft-body anatomy with specialized filter feeding. Neoteny may have played a significant role in the evolution of platycopids. *See* CRUSTACEA; OSTRACODA; PODOCOPA. Patsy A. McLaughlin

Bibliography. L. G. Abele (ed.), *The Biology of Crustacea*, vol. 1: *Systematics, the Fossil Record, and Biogeography*, 1982; H. Kozur, Einige Bermerkungen zur Systematik der Ostracoden und Beschreibung neuer Platycopida aus der Trias Ungarns und der Slowakei, *Geol. Palaeontol. Mitt. Innsbruk*, vol. 2, 1972; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; V. Pokorny, *Principles of Zoological Micropaleontology*, 2:219–237, 1965.

Platyctenida

An order of the phylum Ctenophora comprising four families (Ctenoplanidae, Coeloplanidae, Tjalfiellidae, Savangiidae) and six genera. All species are highly modified from the planktonic ctenophores.

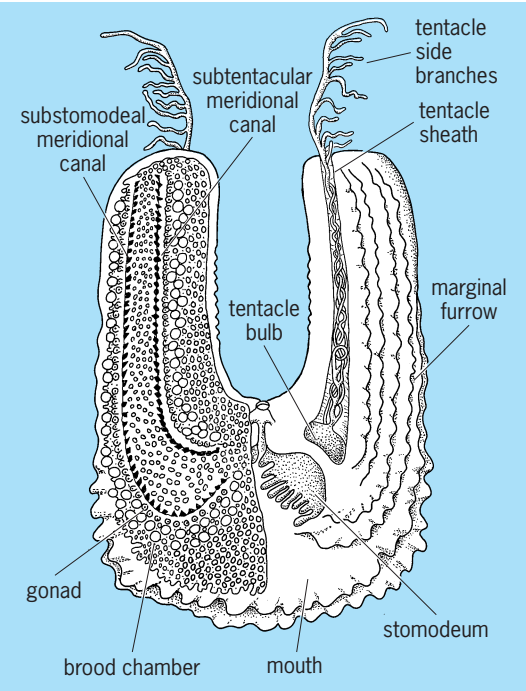

Comb jelly, *Lyrocteis imperatoris*.

The platyctenes are fairly small (1–6 cm or 0.4–2.4 in.) and brightly colored. They have adopted a variety of swimming, creeping, and sessile habits, with concomitant morphological changes and loss of typical ctenophoran characteristics. The body is compressed in the oral-aboral axis, and the oral part of the stomodeum is everted to form a creeping sole (see **illus.**). Platyctenids have a fairly typical cydippid larva, and comb rows are retained in adult ctenoplanids, which are capable of swimming or creeping on a substrate. The remaining families all have a benthic habit and have lost the comb rows, and even the statocysts as adults in the Savangiidae. Sexual reproduction involves internal fertilization in many species, with retention of the developing cydippid larvae in brood pouches. Some species also reproduce asexually by fission. Most platyctenids are found in tropical coastal waters, where many are ectocommensals on benthic organisms. *See* CTENOPHORA. Laurence P. Madin

Platyhelminthes

A phylum of the invertebrates, commonly called the flatworms. They are bilaterally symmetrical, nonsegmented worms characterized by lack of coelom, anus, circulatory and respiratory systems, and exo- and endoskeletons. Many species are dorsoventrally flattened. They possess a protonephridial (osmoregulatory-excretory) system, a complicated hermaphroditic reproductive system, and a solid mesenchyme which fills the interior of the body. Some parasitic species, that is, some trematodes, have secondarily acquired a lymphatic system resembling a true circulatory system. Some species of trematodes, the schistosomes, have separate sexes.

Traditionally, three classes were distinguished in the phylum: the Turbellaria, mainly free-living predacious worms; the Trematoda, or flukes, ecto- or endoparasites; and the Cestoda, or tapeworms, endoparasites found in the enteron (alimentary canal) of vertebrates, whose larvae are found in the tissues of invertebrates or vertebrates. However, recent cladistic analyses using morphology, including ultrastructure, as well as DNA analysis, have shown that the "Turbellaria" are an assemblage of taxa that are not monophyletic (that is, they are not a group containing all taxa with a common ancestor), and that the monogeneans, earlier included in the trematodes, do not belong to the trematodes. Most importantly, the Acoela do not belong to the Platyhelminthes, but are a very archaic group close to the base of the lower invertebrates; and all major groups of parasitic Platyhelminthes—that is, the Trematoda, Monogenea, and Cestoda—are monophyletic, constituting the Neodermata.

Morphology. Flatworm tissues and organs are derived from three germ layers. Most flatworms have adhesive or attachment devices. In turbellarians, these devices are glandular or muscular; in monogeneans, trematodes, and cestodes, they are suckers and/or hooks or spines. Muscular tissue occurs

in mesenchymal layers and permits rapid change in body form. Numerous receptor cells (mechanoreceptors and chemoreceptors) and sense organs (tentacles, statocysts for balance, and ocelli for photoreception) occur on the head and body. Some of these are reduced or absent in parasitic forms, although there is no general reduction in the complexity of the nervous system and sense receptors in the parasitic forms. Some parasitic species, in particular the Aspidogastrea, even have a nervous system of greater complexity and a greater variety of sensory receptors than free-living forms. The nervous system consists of longitudinal nerve cords (connectives) connected by transverse nerves (commissures). In some groups the nervous system has an irregular netlike appearance. Usually some of the anterior commissures are particularly well developed, forming a kind of brain (cephalization). *See* NERVOUS SYSTEM (INVERTEBRATE).

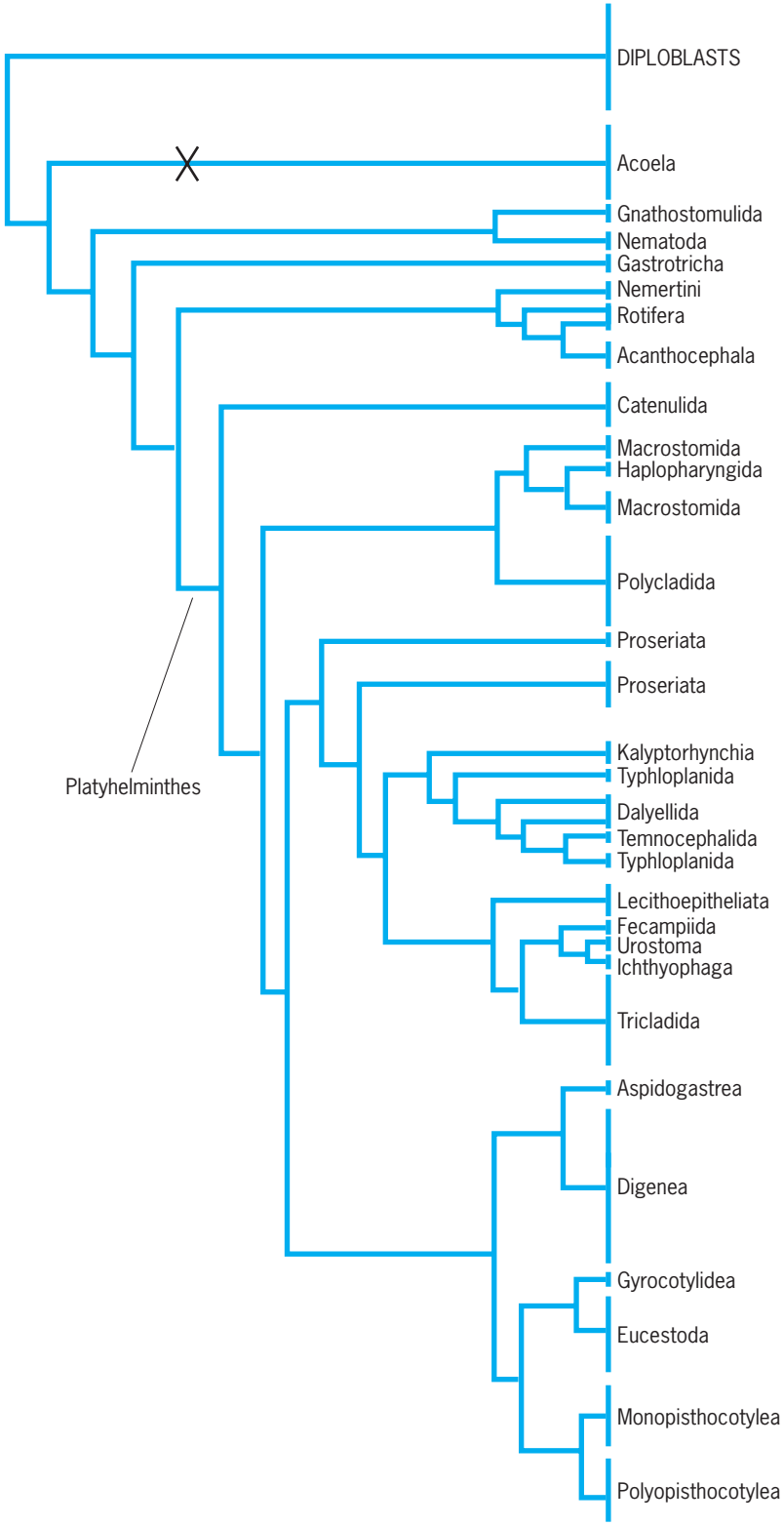

Fig. 1. Phylogeny of Platyhelminthes based on molecular (DNA) and morphological evidence. Note that the Acoela do not belong to the Platyhelminthes (X). (*Strongly modified from D. T. J. Littlewood, K. Rohde, and K. A. Clough, Biol. J. Linnean Soc., 66:75–114, 1999*)

A digestive system is present in turbellarians, monogeneans, and trematodes but is lacking in cestodes. A mouth, serving also for egestion, is situated either anteriorly or ventrally. It is often provided with an aspirating sucker or pharynx by which host tissues or other food particles are ingested. The gut is saclike or branched and lined with a single layer of cells often packed with granules. The protonephridial system is of uniform construction in the phylum. It consists of ramifying blind tubes capped with large cells, called flame cells, each bearing a tuft of cilia which projects into the lumen. The tubes course through the mesenchyme and discharge at the surface by means of one or more openings. The function of this system is not well understood but contributes to osmoregulation and excretion. *See* DIGESTION (INVERTEBRATE).

The reproductive system reaches very great morphological complexity. In each reproductive unit—which is the entire organism in the Turbellaria, Monogenea, Trematoda, Amphilinidea, and Gyrocotylidea, and which is the proglottid (segment) in the Eucestoda—one to several testes and ovaries occur as well as various accessory reproductive organs.

Reproduction. The Platyhelminthes reproduce both sexually and asexually. In sexual reproduction, fertilization is internal, following copulation or hypodermic insemination. Cross-fertilization is the rule but self-fertilization may occur. In most flatworms, the eggs are ectolecithal; that is, the ova are invested with yolk cells and the egg mass is enclosed in a capsule. In turbellarians, capsules are often deposited in cocoons, in which the juveniles develop. Eggs of monogeneans, trematodes, and cestodes emerge from the hosts. In the monogeneans, larvae infect the final vertebrate hosts. In the other two, they either are eaten or hatch in water as free-swimming larvae; then they either actively penetrate, or are passively ingested by, intermediate hosts. Alternatively, many species among the trematodes encyst on a substratum outside a host.

Asexual reproduction is a frequent occurrence in this phylum. Many turbellarians reproduce by fragmentation or binary fission; in some, chains of individuals are temporarily produced. The Digenea reproduce asexually in their larval stages, as do some tapeworms, such as *Echinococcus* and *Multiceps*.

Formation of proglottids is an asexual process, resulting from activity of a proliferating zone in a neck region. *See* REPRODUCTION (ANIMAL).

Economic and biological importance. Turbellaria are widespread in freshwater and the littoral zones (the biogeographic zones between the high- and low-water marks) of the sea, while one group of triclads occurs on land in moist habitats. In the meiofauna (small benthic animals) along the coasts, huge numbers of small turbellarians of many species occur, but only a very small proportion of species has been described. Turbellaria have been used to study regeneration, including the effects of chemicals and radiation upon the process. They have also been used in axial gradient research. Adults of parasitic flatworms occur on or in many tissues and cavities of the vertebrates on which they feed. They are responsible for troublesome diseases in humans and animals. Monogeneans are of foremost importance as agents of fish disease, especially in aquaculture, whereas digeneans are important agents of disease in domestic animals and humans (schistosomiasis). Larval flukes are frequent in mollusks, mainly gastropods, and occasionally occur in pelecypods (bivalves). Vector hosts, such as insects and fish, are often interpolated between the mollusk and vertebrate. Adult tapeworms, living in the gut or the biliary ducts, compete with the host for food and accessory food factors such as vitamins. Larval tapeworms reside chiefly in arthropods, but larvae of the Cyclophyllidea develop in mammals, which may be severely impaired or even killed by the infection. Investigations utilizing flukes and tapeworms have given clearer insight into the host–parasite relationship. *See* PARASITOLOGY.

Classification. **Figures 1** and **2** show the interrelationships of the major groups of Platyhelminthes according to cladistic analyses based on DNA analysis and morphology (gross morphology and ultrastructure). Note that the Acoela do not belong to the Platyhelminthes and that all major groups of parasitic flatworms, the Neodermata, form a monophylum including all groups derived from one common

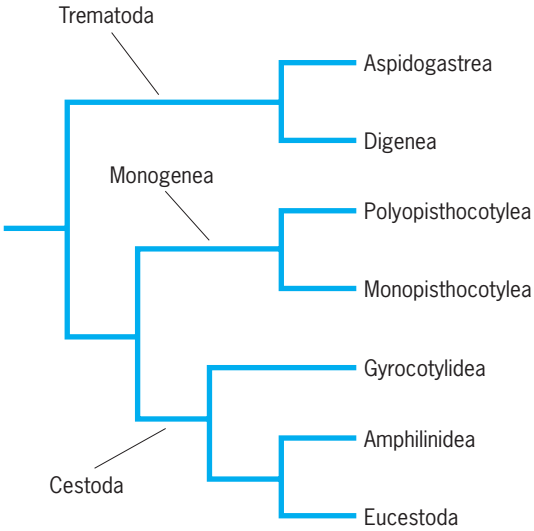

Fig. 2. Phylogeny of the Neodermata, the major groups of parasitic Platyhelminthes. Note that the Aspidogastrea plus Digenea, the Polyopisthocotylea plus Monopisthocotylea, and the Gyrocotylidea plus Amphilinidea plus Eucestoda form monophyletic groups, that is, the Trematoda, Monogenea, and Cestoda.

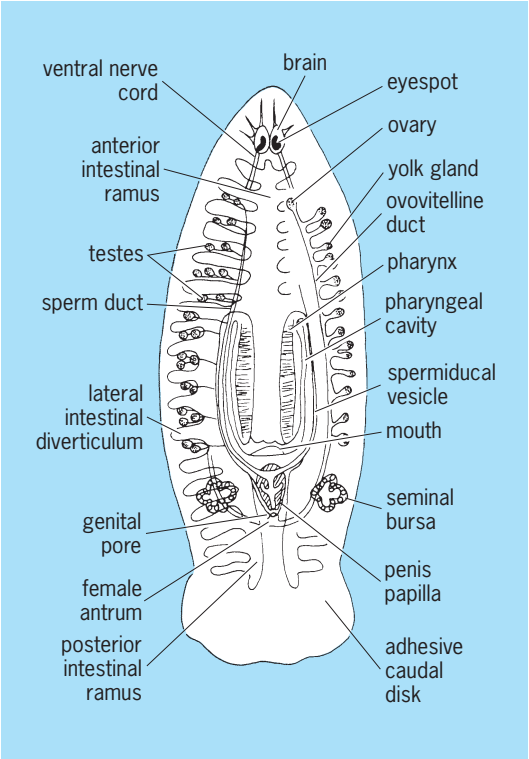

Fig. 3. *Bdelloura candida* (Tricladida), ectocommensal on the king crab, *Limulus*. Complete digestive and male systems are shown on the left, female systems on the right.

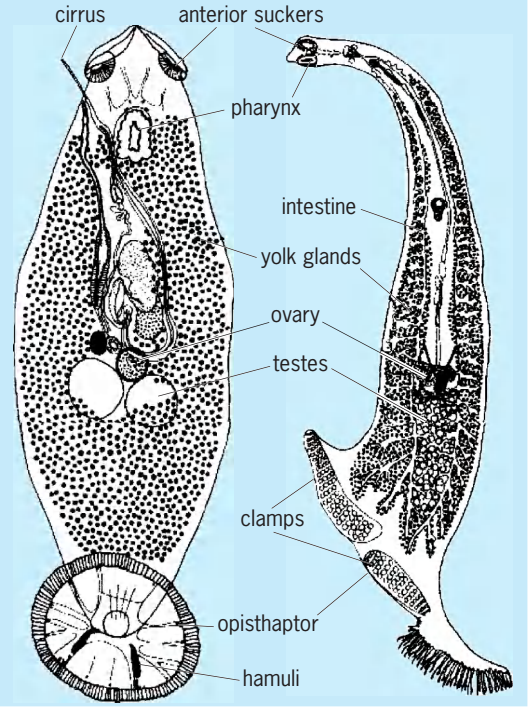

Fig. 4. A monopisthocotylean monogenean on the left and a polyopisthocotylean monogenean on the right. (*Based on K. Rohde, Ecology of Marine Parasites, 2d ed., CABI, Wallingford, Oxon, U.K., 1993*)

ancestor. The Catenulida is the sister group of all other Platyhelminthes, and there are many taxa of equal taxonomic rank.

"Turbellaria." This group is an assemblage of taxa which are not monophyletic. The body is cylindrical or flattened and may be less than 1 mm to over 50 cm long (0.04 to 20 in.). The members are chiefly free-living, but commensal and parasitic species occur. A ciliated epidermis may uniformly cover the body or may be restricted to the ventral or other surfaces. Peculiar rodlike rhabdoids are embedded in the epidermis. Characteristic color patterns are imparted to some species by chromatophores, while other species contain symbiotic algae and are green or brown. Species of Temnocephalida occur as ectocommensals (organisms living on the outer surface of the body of another organism, without affecting the host) mainly on freshwater crayfish. Tricladida consists of larger forms which are marine, freshwater, or terrestrial and have a three-branched intestine (**Fig. 3**). Polycladida contains marine forms of elongate or leaflike shape with few to many eyes near the tentacles and on the body margin. Many intestinal rami branch from a central digestive cavity above the pharynx. Many of the other taxa, in earlier taxonomic systems often grouped in the Rhabdocoela (taxa with a simple saclike gut), are either marine or freshwater forms. The Fecampiida and *Ichthyophaga* are entirely endoparasitic. *See* POLYCLADIDA; TRICLADIDA; TURBELLARIA.

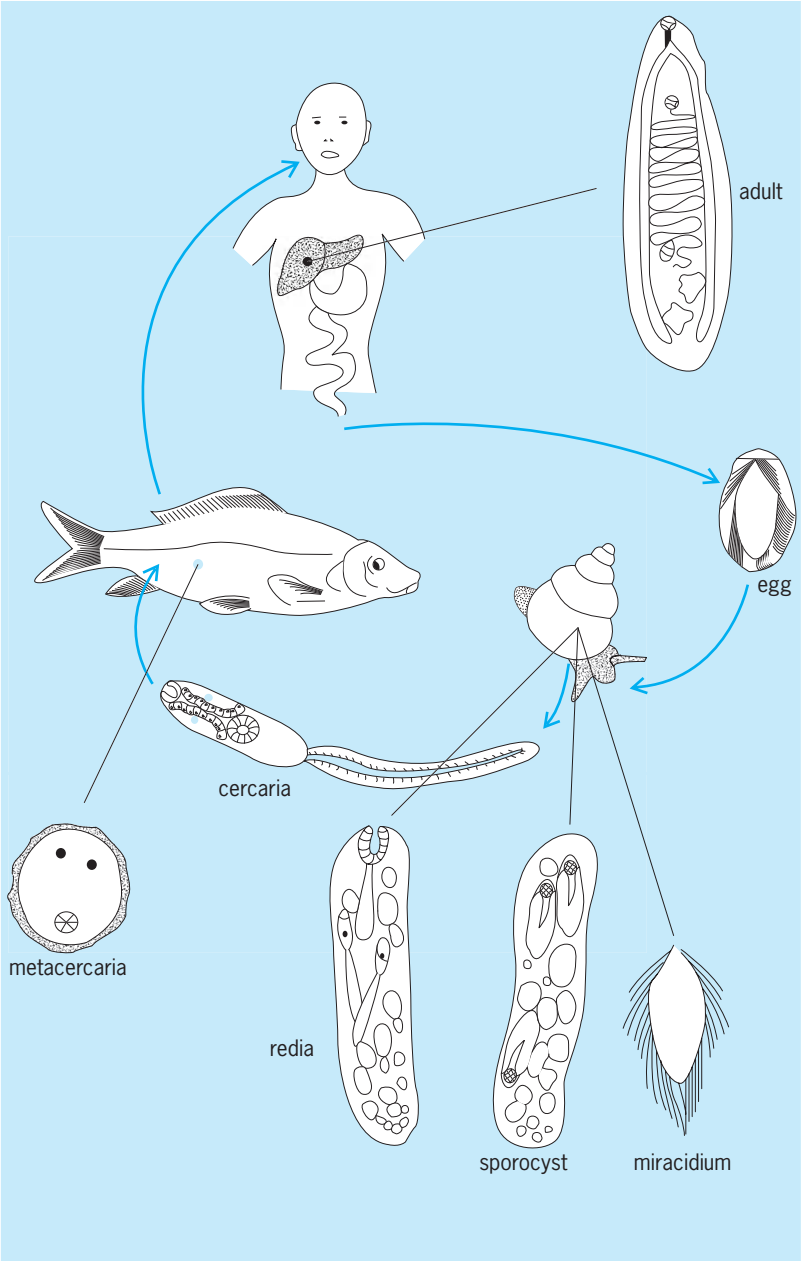

Fig. 5. Life cycle of a liver fluke. (*Based on K. Rohde, Ecology of Marine Parasites, 2d ed., CABI, Wallingford, Oxon, U.K., 1993*)

Neodermata. This large taxon consists of parasites covered with a surface syncytium (tegument) formed by underlying cells. It includes the Monogenea, Trematoda, and Cestoda.

Monogenea. The monogeneans are found mainly as ectoparasites on cold-blooded vertebrates, with direct and simple life cycles. They consist of two groups, the Polyopisthocotylea and Monopisthocotylea. The former have a posterior attachment organ (opisthaptor) with clamplike attachment sclerites and hamuli (hooks), whereas the latter have a posterior attachment organ consisting of a large sucker and/or hamuli (hooks) [**Fig. 4**]. Adult monogeneans produce eggs from which typically a ciliated larva (oncomiracidium) hatches. It attaches itself to the surface of the final host, usually the skin, gills, or fins of fish, and develops to the adult. *See* MONOGENEA.

Trematoda. The flukes consist of two groups, the Aspidogastrea (=Aspidobothria) and Digenea. The Aspidogastrea are endoparasites of invertebrates and vertebrates, characterized by their ventral attachment disk consisting of a row of suckers or many suckerlets, and by a simple life cycle which may, however, include two hosts. The Digenea are endoparasites of vertebrates, with two or more hosts involved in their life cycle (**Figs. 5** and **6**), and the intermediate host is almost always a mollusk. In the Aspidogastrea, there never is secondary multiplication of larval stages—that is, the larva is transformed into the adult. In the Digenea, larval stages multiply—that is, a miracidium (ciliated first larva) is formed in the egg, it develops to a sporocyst which gives rise to several or many daughter sporocysts or rediae, and they in turn give rise to many cercariae, which may encyst outside a host or in a second intermediate host, forming the metacercariae. Each cercaria is transformed to the adult after infection of the final (=definitive) host. *See* ASPIDOGASTREA; DIGENEA; TREMATODA.

Cestoda. These highly specialized animals are parasites of vertebrates and consist of the "segmented" Eucestoda and the nonsegmented Amphilinidea and Gyrocotylidea, formerly grouped in the Cestodaria, which, however, is unlikely to be a monophyletic group. A typical tapeworm in the Eucestoda consists of an anterior attachment organ, the scolex, and a series of similar reproductive segments, the proglottids. *See* CESTODA; EUCESTODA.

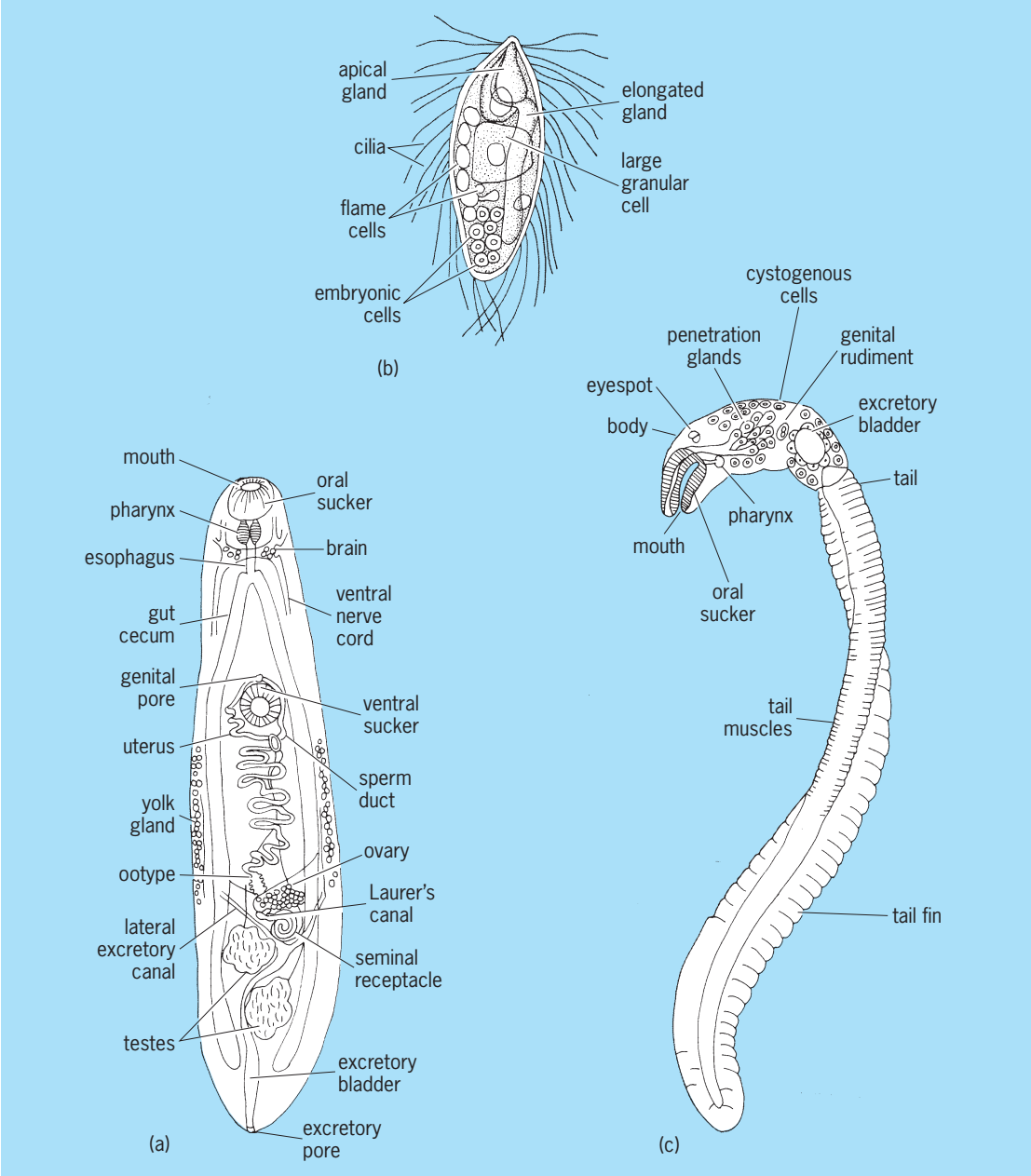

Fig. 6. Some of the stages in the life cycle of *Opisthorchis felineus*. (*a*) Adult from the bile ducts of a cat. (*b*) Swimming miracidium hatched from an egg (this stage penetrates the snail host). (*c*) Cercaria, view from the side (this stage encysts in a fish which is then eaten by the cat). (*After H. Vogel, Zoologica, 33:1–103, 1934*)

Amphilinidea infect the body cavity of teleost fishes and freshwater turtles; their larvae occur in crustaceans. The Gyrocotylidea infect chimaerid fishes; their life cycle is not known. *See* AMPHILINIDEA; GYROCOTYLIDEA.

In Eucestoda the ripe eggs or hatched ciliated larvae (coracidia) are ingested by intermediate hosts in which six-hooked embryos, the oncospheres, transform into larval forms such as procercoids, plerocercoids, cysticercoids, cysticerci, and other specialized types. Final hosts become infested by ingesting these larvae. Klaus Rohde; C. G. Goodchild

Bibliography. R. H. Barth and R. E. Broshears, *The Invertebrate World*, Sanders College Publishing, Philadelphia, 1982; K. Rohde (ed.), *Marine Parasitology*, CSIRO Publishing, Melbourne, and CABI Publishing, Wallingford, Oxon, U.K., 2005; G. D. Schmidt and L. S. Roberts, *Foundations of Parasitology*, 6th ed., McGraw-Hill, New York, 1999.

Platypus

The single species *Ornithorhynchus anatinus* of the family Ornithorhynchidae in the order Monotremata, which occurs in eastern Australia and Tasmania. This mammal is a monotreme, which lays eggs and incubates them in a manner similar to birds. The platypus is also known as the duckbill, duckmole, and water mole. One of the most primitive mammals,

Platypus, a primitive aquatic mammal. (*Photograph by Paddy Ryan/Ryan Photographic*)

it retains some reptilian characteristics. The female lacks a marsupium; however, marsupial or epipubic bones are well developed.

On the male's hindleg is a poison spur whose function is unknown, but which occurs in other monotremes. The skeletal system possesses several reptilianlike characteristics. In addition, as an oviparous species, reptilian features such as a cloaca, double uterus, and primitive reproductive organs are retained.

Ornithorhynchus is well adapted to aquatic life. It ranges from the tropical, sea-level streams to cold lakes at altitudes of 6000 ft (1800 m). It is covered with short dense fur, typical of many aquatic mammals; the external ears are lacking; and the tail is broad and flattened, resembling that of the beaver (see **illus.**). The rubbery bill is covered with a highly sensitive skin. There are no teeth, but horny ridges are used for crushing food and grubbing on the bottom of rivers. Food is principally crustaceans, worms, and tadpoles.

The adult is about 2 ft (0.6 m) long and weighs approximately 4 lb (1.8 kg). The duckbill is a burrowing animal, and digs a chamber in the side of a riverbank. The female lays and hatches two leathery shelled eggs in the chamber, which is lined with dry grass. Incubation takes about 2 weeks. The young are born naked and helpless, and the female feeds them on lacteal secretions produced by mammary glands which more closely resemble sweat glands. The young have temporary molar, deciduous teeth that are lost in the adult. *See* MONOTREMATA.

Charles B. Curtin

Bibliography. S. Jackson, M. Serena, and D. Middleton, Platypus, in S. Jackson (ed.), *Australian Mammals: Biology and Captive Management*, pp. 1–32, CSIRO, Collingwood, Victoria, Australia, 2003; R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999; M. Serena, Duck-billed platypus: Australia's urban oddity, *Nat. Geog.*, 197(4):118–129, 2000.

Playa

A nearly level, generally dry surface in the lowest part of a desert basin with internal drainage (see **illus.**). When its surface is covered by a shallow sheet of water, it is a playa lake. Playas and playa lakes are also called dry lakes, alkali flats, mud flats, saline lakes, salt pans, inland sabkhas, ephemeral lakes, salinas, and sinks. The original meaning in Spanish of shore or beach has been lost. An example is Death Valley in California.

A playa surface is built up by sandy mud that settles from floodwater when a playa is inundated by downslope runoff during a rainstorm. A smooth, hard playa occurs where ground-water discharge is small or lacking and the surface is flooded frequently. These mud surfaces are cut by extensive desiccation polygons caused by shrinkage of the drying clay. Puffy-ground playas form by crystallization of minerals as ground water evaporates in muds near the surface. Ground water moves upward because artesian conditions exist on many playas. Artesian conditions require that subsurface flow be confined within an aquifer, usually a sandy layer, so the piezometric surface created by the hydrologic head is near or above the ground surface. *See* AQUIFER; ARTESIAN SYSTEMS.

Subsurface brine is present beneath many playas. Brines evolve from surface waters that sink into alluvial fans adjacent to the playa basin. As ground water flows through the fans, through the adjoining sandflats, and under the playa, various minerals are precipitated progressively in the sediments, increasing the ionic concentration of the water and modifying the proportions of the ions. The type of brine depends on the original composition of the surface water and reflects the lithology of the rocks weathered in the surrounding mountains. *See* GROUNDWATER HYDROLOGY.

Numerous playas in the southwestern United States yield commercial quantities of evaporite minerals, commonly at shallow depths. Important are salt (NaCl) and the borates, particularly borax ($Na_2B_4O_7 \cdot 10H_2O$), kernite ($Na_2B_4O_7 \cdot 4H_2O$), ulexite

Light-colored playa in lowest part of Sarcobatus Flat in southern Nevada.

($NaCaB_5O_9 \cdot 8H_2O$), probertite ($NaCaB_5O_9 \cdot 5H_2O$), and colemanite ($Ca_2B_6O_{11} \cdot 5H_2O$). Soda ash (sodium carbonate; Na_2CO_3) is obtained from trona ($Na_3H(CO_3)_2 \cdot 2H_2O$) and gaylussite ($Na_2Ca(CO_3)_2 \cdot 5H_2O$). Lithium and bromine are produced from brine waters. *See* DESERT EROSION FEATURES; SALINE EVAPORITES. John F. Hubert

Bibliography. J. L. Melvin (ed.), Evaporites, Petroleum and Mineral Resources, *Dev. Sedimentol.*, no. 50, 1991; H. G. Reading (ed.), *Sedimentary Environments: Processes, Facies and Stratigraphy*, 3d. ed., 1996; R. W. Renault and W. M. Last (eds.), Sedimentology and Geochemistry of Modern and Ancient Saline Lakes, *Soc. Sed. Geol. Spec. Publ.*, no. 50, 1994.

Plecoptera

An order of primitive insects known as the stoneflies. The order comprises 16 families and more than 2000 species distributed on all continents except Antarctica. Stoneflies spend the majority of their lives aquatically as nymphs (immature larval forms); they exit the water and live only briefly as adults on land. There are relatively slight differences between the aquatic nymphal stage and the terrestrial adult form, except for wings and tracheal gills. In stoneflies, the soft and somewhat flattened or cylindrical body ending in two tail filaments, the strong legs with paired tarsal claws, the chewing mouthparts, and the rusty blacks, dull yellows, and browns are characteristic of both the nymph and adult (see **illustration**). The name Plecoptera literally means pleated wing, referring to the hindwings that are folded and hidden under the forewings, which the adult holds close to the abdomen when at rest or walking. *See* INSECTA.

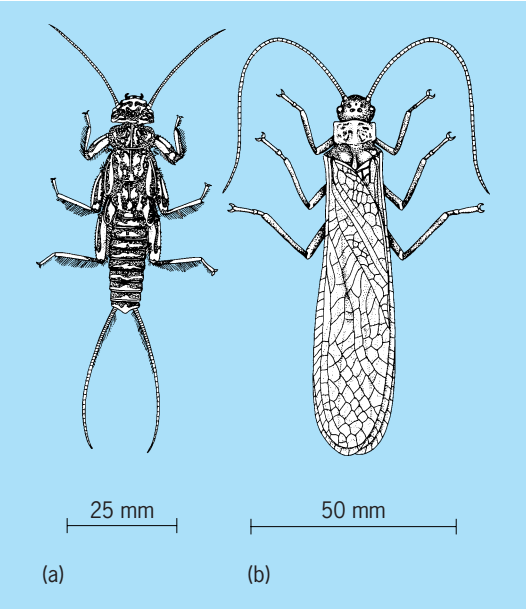

Plecoptera. (a) Nymph of *Paragnetina*. (b) Adult of *Pteronarcys*, one of the largest stoneflies. (*After A. H. Morgan, Field Book of Ponds and Streams, Putnam, 1930*)

Characteristics. Stonefly nymphs live in the rapid stony parts of clean swift streams, although a few species occur along the rocky shores of large temperate lakes. Filamentous tracheal gills are characteristic of all nymphs except those that live in highly aerated water and respire through the body wall. In various species, gills are attached at the neck, on the side of the thorax, and on the sides and top of the abdomen. In some species, such nymphal structures persist in the adults but do not function.

Most stonely species require 1 year for development, but a few may take as long as 3 years. In an observed complete life history of one species, the nymphs went through 22 instars, the form assumed by an insect between molts; the number of known instars ranges from 12 to 36. Nymphs have diverse feeding habits, including species that are carnivores and others that are herbivores and/or detritivores. Their diet can be determined by examining the stomach contents in newly collected specimens.

The adult life stage begins when full-grown nymphs climb onto rocks or plants, molt, and emerge as adults. In various parts of North America stonefly adults emerge during every month of the year. Numerous species of "winter stoneflies" reach their peaks of abundance from late winter to early spring. For these species, emergence typically takes place after openings appear in ice-covered streams. Adults are poor fliers and are often found creeping about on stream banks and shrubbery.

Reproductive behavior. Adults live from one to a few weeks, and the time is mainly spent engaged in reproductive activities. Mating occurs near the ground, never in flight. For most North American species of stoneflies, males and females locate each other through a complex system of vibrational communication called drumming. The male taps or rubs his abdomen on various substrates and waits for a receptive female to respond, whereby they are able to locate each other. Each species has its own unique pattern of drumming that varies in amplitude, frequency, and duration. Females deposit egg masses into the water using various means. Stoneflies that inhabit large streams deposit their eggs directly into the swirling water in clutches of more than 1000; other species place the eggs at the water's edge, drop them from the air, or even crawl underwater and deposit eggs onto the steam bottom. *See* INSECT PHYSIOLOGY. Andrew K. Rasmussen; Lewis Berner

Bibliography. E. Leiser and R. H. Boyle, *Stoneflies for the Angler*, Knopf, New York, 1982; W. P. McCafferty, *Aquatic Entomology: The Fishermen's and Ecologists' Illustrated Guide to Insects and Their Relatives*, Jones and Bartlett, Sudbury, MA, 1981; R. W. Merritt and K. W. Cummins (eds.), *An Introduction to the Aquatic Insects of North America*, 3d ed., Kendall Hunt, Dubuque, Iowa, 1996; K. W. Stewart and B. P. Stark, *Nymphs of North American Stonefly Genera (Plecoptera)*, 2d ed., Caddis Press, Columbus, Ohio, 2002; K. W. Stewart, S. W. Szczytko, and B. P. Start, The language of stoneflies, *Bioscience*, 33:117–118, 1983.

Pleiades

A typical open cluster of stars (see **illus.**), and the best example in the sky of an open cluster for the casual observer. (Open clusters are one of the two basic types of star clusters, distinguished from globular clusters by their smaller total populations and central densities, and by their younger age. Besides being a focus of astronomical research in the present era, the Pleiades has been considered important since ancient times in various cultures, including ancient civilizations of Pre-Columbian Mexico, Greece, Persia, India, and China. *See* STAR CLUSTERS.

The Pleiades, located in the constellation Taurus, contains several hundred stars within a radius of about 1°, centered on its brightest star, Alcyone. (In comparison, the Moon is 0.5° across.) The cluster is permeated with a faint diffuse nebulosity. Though the Pleiades is always referred to as having seven bright stars, only six are seen today. The "lost Pleiad" is likely to be Pleione, which is presently a 5th-magnitude variable star. The configuration of the brighter stars resembles a very small dipper. *See* TAURUS; VARIABLE STAR.

Several characteristics of the Pleiades make the cluster useful for research. It is a young cluster, about 100 million years old, and still contains bright hot blue stars (type B) on the main sequence (hydrogen-burning stars) with abundances of elements similar to that of the Sun. It is compact, and close enough to see its smaller fainter members (including brown dwarfs). Finally, although it has stars of widely varying masses and brightnesses, they are all approximately the same age and distance from Earth, about 130 parsecs (425 light-years). *See* LIGHT-YEAR; PARSEC; STAR; STELLAR EVOLUTION.

The Pleiades. (*Lick Observatory photograph*)

Until the *Hipparcos* satellite was launched in 1989, the Pleiades was just beyond the distance at which trigonometric parallaxes could be measured. While the distances measured by *Hipparcos* for the Hyades and other clusters were generally near the values expected from previous measurements, the Pleiades cluster was found to be closer than its expected distance. This implied that the actual light given off by the stars was only about three-quarters of what had been assumed from the pre-*Hipparcos* measured distance. After several years of investigation, the most likely resolution to the problem seems to be that the *Hipparcos* measurements for the Pleiades cluster contain unexpected systematic errors, in the complex solution of the extremely precise positional measurements obtained within a limited period of time, and the Pleiades is actually at its pre-*Hipparcos* measured distance. *See* PARALLAX (ASTRONOMY).

Because it is relatively nearby, the Pleiades is frequently studied. Brown dwarfs, whose masses are smaller than that of the faintest true star but larger than those of planets, have been found in the Pleiades. While they may make up a quarter of the objects in the cluster, they probably account for no more than 1% of the cluster mass. The Pleiades is also a useful basis for comparison with other clusters in order to obtain the distances of clusters much farther away. Other topics for research include the effects of rapid rotation (stars slow down as they age) and the fitting of theoretical models to the Pleiades. *See* BROWN DWARF; STELLAR ROTATION.

Nancy Remage Evans; Ronald Pitts

Bibliography. G. Basri, The discovery of brown dwarfs, *Sci. Amer.*, 282(4):76, April 2000; A. G. A. Brown, The Hyades—So close, and now, so familiar, *Mercury*, 27(3)17–21, May/June 1998; R. Burnham, Jr., *Burnham's Celestial Handbook*, vol. III, Dover Publications, New York, 1978; B. Paczynski, Astronomy: A problem of distance, *Nature*, 427:299–300, 2004; L. Kun Tan, The Pleiades, in Gallery, *Sky Telesc.*, 99(2):136, February, 2003.

Pleistocene

The older of the two epochs of the Quaternary Period. The Pleistocene Epoch represents the interval of geological time (and rocks accumulated during that time) extending from the end of the Pliocene Epoch (and the end of the Tertiary Period) to the start of the Holocene Epoch. Most recent time scales show the Pleistocene Epoch spanning the interval from 1.8 million years before present (m.y. B.P.) to 10,000 years B.P. The Pleistocene is commonly characterized as an epoch when the Earth entered its most recent phase of widespread glaciation. *See* QUATERNARY.

British naturalist Charles Lyell introduced the term Pleistocene in 1839. Earlier, he had subdivided the

CENOZOIC	QUATERNARY	Holocene
		Pleistocene
	TERTIARY	Pliocene
		Miocene
		Oligocene
		Eocene
		Paleocene
MESOZOIC	CRETACEOUS	

Tertiary Period on the basis of the degree of faunal similarity of the epochs with modern molluscan fauna. He identified an older Pliocene and a younger Pliocene, with 50% and 70% faunal similarity with modern fauna, respectively. In 1839 he introduced the concept of Pleistocene as a separate epoch to replace the younger Pliocene. Later he attempted to abandon this term in favor of a generic post-Pliocene, but this revisionism was largely ignored by the geological community as the use of the term Pleistocene had become popular in stratigraphic literature. The Holocene Epoch was proposed by the French naturalist P. Gervais in 1867 to cap the Pleistocene and the geological column. The Holocene represents the youngest of the geological epochs and spans the last 10,000 years of geological time. Its fossil content is characterized by 100% faunal similarity to modern assemblages. *See* HOLOCENE; PLIOCENE; TERTIARY.

Following the publication in 1840 of the work by Swiss zoologist Louis Agassiz, who envisioned a climatically induced continental ice sheet all the way from the North Pole to Central Asia and across the Alps, geologists for the first time began associating the Pleistocene as a geological epoch with widespread glaciation. Though Agassiz's notion was correct in its basic tenet, it was based on several ill-founded assumptions and was not taken seriously. By 1850, however, the concept of the Pleistocene ice age took hold, largely through the documentation of glacial till sediments by Scandinavian and Scottish geologists. The glacial deposits normally lack contemporaneous fossils (they may contain reworked older fossils) and are thus difficult to date. Nonetheless, the fossils contained in sediments below and above the glacial deposits showed considerable affinity to modern assemblages, thereby indicating to the early workers that these sediments were relatively recent, that is, of the Quaternary age. *See* TILL.

The early workers envisioned a single large glacial event in the Pleistocene, but observations soon led to evidence of more than one glaciation, separated by more equitable interglacial periods. As evidence has accumulated in recent decades from both the land and the sea, it clearly shows numerous glacial advances and retreats. At least 17 glacial events have been shown to have occurred during the Pleistocene. Modern paleoclimatic studies also indicate that the Pleistocene does not characterize the first glaciation. The first significant glaciation in the most recent phase of Earth's glacial history began as far back as the middle Miocene, some 15 m.y. ago. Nevertheless, extensive and high-frequency glacials events are characteristic of the Pleistocene Epoch. *See* CLIMATE HISTORY; GLACIAL EPOCH.

While geologists and glaciologists think widespread glaciation typifies the Pleistocene Epoch, in the past anthropologists considered the appearance of modern humans as a defining criterion, and vertebrate paleontologists used the fossiliferous occurrence of the elephant, modern horse, and cattle in their definition of this interval. Some of these events were later shown to be of greater antiquity and not restricted to the Pleistocene.

Subdivisions. In modern geological time scales, the Pleistocene is subdivided into a lower and an upper series. In Europe, the lower series is considered equivalent to the Calabrian Stage, while the upper series is equated with the Sicilian and Tyrrhenian stages (**Fig. 1**). The exact boundaries between these stages are, however, still a matter of debate among stratigraphers.

The Calabrian Stage was established by the French geologist M. Gignoux in 1913 based on shallow marine sediments in Calabria, the toe of Italy. The occurrence of the cold-water mollusk species *Arctica islandica* and the benthic forminifer *Hyalinea balthica* characterizes the Calabrian beds, which were assumed to correspond to the latest Pliocene cold period. The latter was elevated to represent the basal

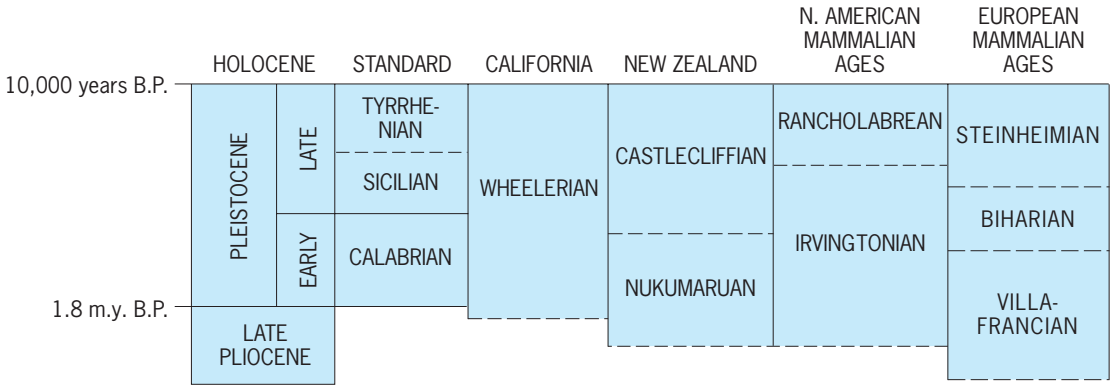

Fig. 1. Pleistocene stages and their temporal equivalents. (*After B. U. Haq and F. W. B. van Eysinga, Geological Time Table, Elsevier, Amsterdam, 1998*)

Pleistocene at the International Geological Congress in 1948. However, the base of the Pleistocene (or the Pliocene-Pleistocene boundary) has remained controversial. At present the Calabrian sections at La Castella and Vrica and near the village of Santa Maria di Catanzaro are variously used by stratigraphers to typify the Pliocene-Pleistocene boundary and the Calabrian Stage. In 1977 B. Haq and coworkers, in an attempt to reconcile the paleoclimatic and chronologic data, used multiple, fossil, radiometric and paleomagnetic criteria to define the boundary at the Calabrian sections. They equated the base of the Pleistocene with the top of the Olduvai magnetic event, and considered it as coeval with the first of the four major climatic deteriorations of the Pleistocene. This finding also reconciled the boundary with the arrival of the cold molluscan and foraminiferal fauna in the Mediterranean and a dramatic reduction in temperatures. More recently, Italian stratigraphers have preferred the Vrica section as the type section for the Pliocene-Pleistocene boundary. It is a thicker section, containing much the same events as the La Castella section but a better ash-bed chronology. Nevertheless, their placement of the boundary yielded the same conclusions as those reached by Haq and coworkers. *See* DATING METHODS; PALEOCLIMATOLOGY; PALEOMAGNETISM.

The marine Calabrian is replaced locally in Italy by fresh-water deposits. These were designated by L. Pareto in 1856 as the Villafranchian Stage based on a section near Villafranca di Asti that is rich in mammalian fauna. The occurrence of *Equus* and *Elephas* in these beds was considered to be their earliest, and to mark the beginning of the Quaternary according to vertebrate paleontologists. However, the time equivalency of the Villafranchian with Calabrian has been questioned, and the former has been shown to be at least in part older than the Calabrian, and therefore late Pliocene. The first appearances of modern horses and elephants have also been shown to date from the late Pliocene.

The Calabrian deposits are overlain by shallow marine sediments assigned to the Sicilian Stage, which was considered to characterize the second glacial episode marked by the return of *A. islandica* to the Mediterranean area. In the later part of the Sicilian, the climate once again ameliorated and warmer faunas returned. This was followed by the deposition of the sediments of the Tyrrhenian Stage, which represents another regressive-transgressive cycle. The Tyrrhenian sediments represent overall warming conditions with several smaller, colder intervening episodes.

Local Pleistocene stage equivalents (Fig. 1) in Italy include the Emilian (which straddles the early-late Pleistocene boundary) and the Milazzian (which is equivalent to the later part of Sicilian and early Tyrrhenian stages). In California, the Wheelarian Stage is considered largely to be equivalent to the Pleistocene, though its lower boundary is somewhat older and occurs in the latest Pliocene. In Japan the upper half of the Kechienjian Stage and the Yuzanjian Stages are considered to span the early Pleistocene. The Australian Werrikooian Stage also spans only the early Pleistocene. In New Zealand, stratigraphers recognize the Nukumaruan Stage as overlapping the latest Pliocene and the earlier Pleistocene, and the Castlecliffian as representing the later Pleistocene. They divide the Castlecliffian further into a lower, Okehuan substage, and an upper, Putikian substage. Students of vertebrate fossils, who often name their own stages based on mammalian fauna, equate the upper part of Villafrancian, the Biharian, and the Steinheimian (Oldenburgian) with the Pleistocene in the Old World, and use Irvingtonian and Rancholabrean in North America to subdivide the Pleistocene. Each of these ages is characterized by its own typical mammalian assemblage. *See* INDEX FOSSIL.

Glaciations and climate. Contrary to the belief of early workers who equated the Pleistocene with the ice age, evidence accumulated in recent years indicates that glaciation in the Northern Hemisphere actually began in the late Pliocene. Data from the stable oxygen isotopes from the marine record suggests that both Iceland and Greenland were glaciated at least 3.2 m.y. B.P. An especially cold interval also occurred around 3.0 m.y. B.P., when glacial ice reached the sea in both polar regions as indicated by substantial ice-rafted debris on the sea floor. The remainder of the Pliocene shows cyclic waxing and waning of the ice sheets. The cyclic nature of the glacial-interglacial variations becomes even more characteristic of the Pleistocene. Research in the 1970s and 1980s firmly tied this cyclicity to variations in astronomical parameters that control the distance between the Earth and the Sun (Milankovitch cyclicity, discussed below). The onset of the Pleistocene brought glaciations that were more widespread than those in the Pliocene. Mountain glaciers expanded and continental ice fields covered large areas of the temperate latitudes. Sea ice also became more widespread. Where there were no glaciers, such as in the lower latitudes, considerable cooling occurred.

Four or five major glacial events have been identified in the Pleistocene continental record of the temperate and high latitudes, and each of these episodes shows a higher frequency of advances and retreats of glacial ice. In the Alps, five major glacial episodes have been identified and are known as (from the oldest to the youngest) the Donau, Günz, Mindel, Riss, and Würm. On the North American continent, four major glacial advances have been identified in the last million years of the continental record. These are known as (from the oldest to the youngest) the Nebraskan, Kansan, Illinoian, and Wisconsinan. During each of these episodes, ice sheets covered much of the North American continent (**Fig. 2**). The synchroneity of the younger four glaciations in the Alps with those of North American is, however, not fully established due to difficulty in direct dating of glacial deposits. In New Zealand, four major Pleistocene glaciations are also evident when highland glaciers advanced to the lowlands, but their correlation with Northern Hemisphere glacial episodes have not been clearly established.

The most complete record of the Pleistocene glacial and climatic variations that can be tied to a

meaningful chronology comes from the deep sea. Deep-sea sediment cores, especially those of the North Atlantic, contain a record of climatic cyclicity that has been shown to be forced by Milankovitch modulation. In the early twentieth century, M. Milankovitch established curves for solar insolation (the amount of solar radiation impinging on the Earth's surface) as a factor of temporal variations in the Earth's distance from the Sun. This distance is controlled by three astronomical parameters with different periodicities. These include variations in (1) the stretch of the Earth's orbit around the Sun, or eccentricity, with a periodicity of about 95,000 years; (2) the tilt of the Earth's axis of rotation, or obliquity, with a periodicity of about 40,000 years; and (3) the wobble in the Earth's rotation, or precession, that has a periodicity of about 21,000 years. Eccentricity and precession directly define the distance of the Earth from the Sun, whereas obliquity influences the distribution of solar radiation in different latitudes. Milankovitch's curves implied that the variation in the solar insolation was the most important factor in modulating the Pleistocene climate. These time-dependent astronomical parameters have been directly tied to the observed periodicities in the stable-isotopic and faunal and floral record from the deep sea. *See* EARTH ROTATION AND ORBITAL MOTION; INSOLATION; MARINE SEDIMENTS.

Although the paucity of accurate time correlations remains a limiting factor, climatic data on a regional basis suggest that Pleistocene climates differed in different areas, and cooling and warming trends were not uniform. Generally, cooler climates followed the onset of glaciations, and warmer climates trailed glacial retreats. The interglacial periods, however, have been inferred to have become progressively cooler through the Pleistocene, with each successive interglacial as much as 1°C (1.8°F) cooler than the previous one. Vegetation and faunal distributions also followed these climatic fluctuations. For example, in North America, tundralike vegetation spread several thousand kilometers to the south during the peak of the Wisconsinan glaciation when compared to its present-day distribution. *See* GEOLOGIC THERMOMETRY; TUNDRA.

Oceans. The marine sedimentary record shows that in the oceans the equatorial regions seem to have suffered less cooling during glacial episodes than the temperate to higher latitudes. For example, near the end of the last glacial maximum, some 18,000 years ago, the Caribbean and Central Atlantic waters cooled only slightly, while the North Atlantic at 40°N cooled by as much as 10°C (18°F). This was a consequence of a significant southward shift of the paleo-Gulf Stream during the glacial maximum. The average temperature of the ocean waters was, however, only 2.3°C (4.1°F) cooler than at present. On the continents the temperatures were some 5°C (9°F) cooler than today. Temperature increases during the interglacials were also of greater magnitude in the continental interiors, where the average summer temperature may have increased by as much as 15°C (27°F).

The expansion and decay of the ice sheets had a direct effect on the global sea level. Global sea-level fluctuations of 50–150 m (170–500 ft) have been estimated for various glacial-interglacial episodes during the Pleistocene, following the sequestering of water as glacial ice on the continents and its subsequent melting. Since the last deglaciation, which began some 17,000 years ago, the sea level has risen by about 110 m (360 ft) worldwide, drowning all of the ancient lowstand shorelines. One important product of the sea-level drops was the migration of large river deltas to the edges of the continental shelves and to the deeper parts of the basins. These coarse-grained sandy sediment bodies constitute ideal reservoirs for the accumulation of hydrocarbons. Conversely, the last marine transgression that started in the late Pleistocene after rapid deglaciation and ended in the Holocene (6000 to 7000 years ago) resulted in new deltas that formed at the mouths of modern rivers. All present-day river deltas date from the cessation of the rapid sea-level rise and the attainment of relative eustatic stability in the mid-Holocene. *See* DELTA; PALEOCEANOGRAPHY; PETROLEUM GEOLOGY.

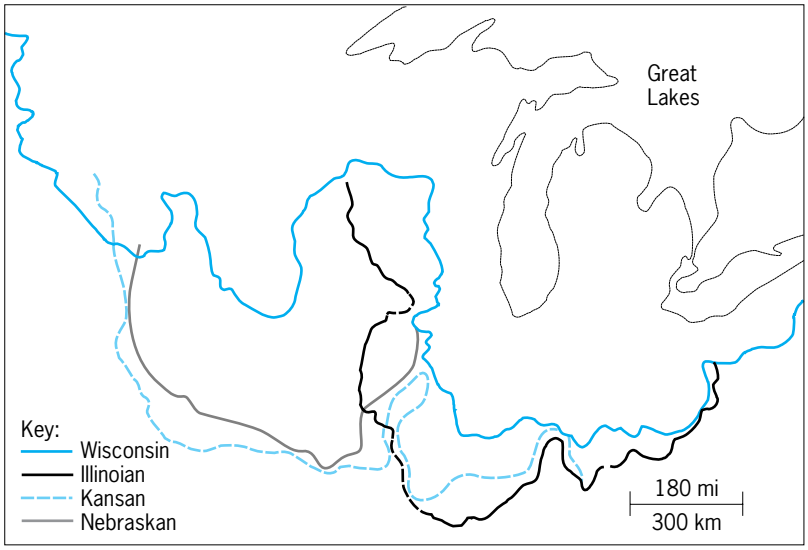

Fig. 2. Spread of the continental ice sheets during various Pleistocene ice ages over North America. (*After R. F. Flint, Glacial and Quaternary Geology, Wiley, New York, 1971*)

Life. The Miocene three-toed horse genus *Hipparion* survived into the latest Pliocene part of the Villafranchian. These and other survivor species, such as mastodonts, and premodern rhinos became largely extinct at the close of the Pliocene. The onset of cooler climate and Pleistocene glaciation is also approximated with a wave of mammalian migration from the east to the west. A relatively modern-looking fauna that included the first true oxen, elephants, and the first one-toed horse first appeared at the beginning of Pleistocene. The modern horse, *E. caballus*, made its first appearance some 250,000 years ago in the late Pleistocene in North America. From North America it migrated to Asia and then west to Europe. However, during the last glacial maximum, some 18,000 years ago, it became extinct in North America when it was unable to cross the deserts to migrate to South America. Oxen, deer and reindeer, large cats, mammoth, great elk, wolf,

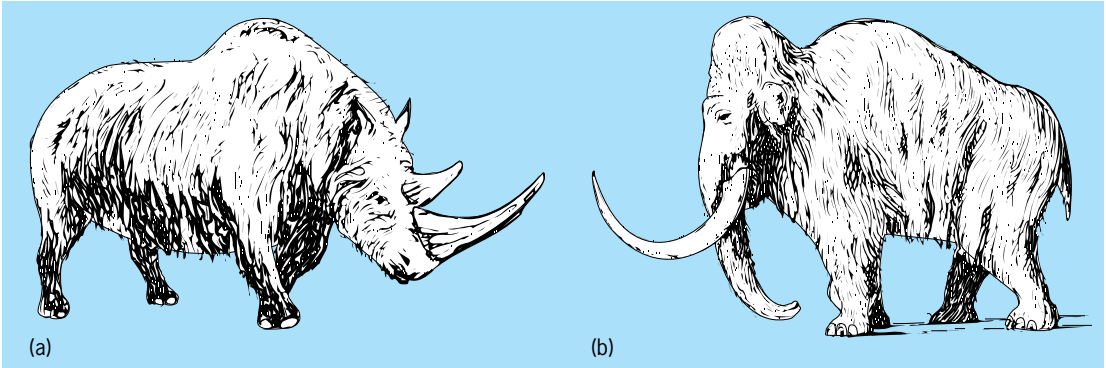

Fig. 3. Two European mammals of the Pleistocene that ranged widely and died out before the close of the epoch. (*a*) Woolly rhino (*Coelodonta antiquitatis*) (*after B. Kurten, Pleistocene Mammals of Europe, Weidenfeld and Nicolson, 1968*). (*b*) Woolly mammoth (*after O. Abel, Die vorzeitlichen Säugetiere, Gustav Fischer, 1914*)

hyena, and woolly rhinos (**Fig. 3**) proliferated during the middle and late Pleistocene. Mammoths, which have been found preserved nearly intact in frozen soils in Siberia, ranged over much of Europe during the glacial times. Thick fur and a substantial layer of protective fat gave these species an advantage in the cold climates. Their reign ended at the close of the glacial period. *See* GEOLOGIC TIME SCALE; MAMMALIA.

The unique Pleistocene mammalian faunas of some of the isolated islands, such as Madagascar, the Philippines, Taiwan, and the Japanese Archipelago, indicate restriction in the dispersal of species during the Pleistocene.

Before the onset of the Pleistocene glaciation, the Pliocene Epoch had already seen the evolution of a whole lineage of hominids in Africa that had given rise to the human precursor, *Homo habilis*, around 2 m.y. B.P. This was followed by the appearance of *H. erectus*, which came on the scene almost at the Plio-Pleistocene boundary, around 1.8 m.y. B.P. These species were followed by more recent human ancestors, and the Pleistocene Epoch is characterized by the final evolution of the Man. The first archaic *H. sapiens* are now considered to have arrived on the scene earlier than previously believed, around about a million years ago. The appearance of *H. neanderthalensis*, or the Neanderthal Man, whose tenure of the tropics and temperate regions partly overlapped with the archaic man, is now dated at least as far back as 250,000 years B.P. This species became extinct at approximately the same time as the rise of modern man. Recent datings have also pushed back the appearance of the first true modern *H. sapiens* (the Cro-Magnon Man) to around 100,000 years B.P. Even these dates are, however, quite likely to be revised (to greater antiquity) as new age data accumulate. *See* FOSSIL HUMANS; GEOLOGIC TIME SCALE.

Bilal U. Haq

Bibliography. B. U. Haq et al., The corrected age of the Pliocene-Pleistocene boundary, *Nature*, 269:483–488, 1977; B. U. Haq and F. W. B. van Eyinga, *Geological Time Table*, 6th ed., 2007; T. Nilsson, *The Pleistocene: Geology and Life in the Quaternary Ice Age*, 1982; C. Pomerol, *The Cenozoic Era: Tertiary and Quaternary*, 1982.

Pleochroic halos

Spherical or elliptical regions up to 40 micrometers in diameter in which there is a change in color from the surrounding mineral when viewed with a petrographic microscope. Pleochroic halos are found around small inclusions of radioactive minerals—for example, zircon, monazite, allanite, xenotime, and apatite—and in rock-forming minerals, principally quartz, micas, amphiboles, and pyroxenes. Halos have also been identified in coalified wood preserved in deposits on the Colorado Plateau. These small regions of discoloration were recognized by H. Rosenbusch as early as 1873 in cordierite, but it was not until the early twentieth century that J. Joly ascribed pleochroic halos to the effect of irradiation originating from the inclusions. *See* CORDIERITE; PETROFABRIC ANALYSIS.

The change in color is a result of radiation damage caused by alpha particles emitted during the radioactive decay of nuclides in the decay chains of uranium-238, uranium-235, and thorium-232. Typically, the alpha particle has an energy of 4.5–9 MeV and a range of 10–20 μm, dissipating most of its energy by ionization with only a few hundred atomic displacements at the end of its trajectory. The alpha recoil has an energy of 0.07–0.09 MeV and a range of only 0.02 μm. Thus, it is the range of the alpha particle and ionization effects that account for the size and color of the halos. The halos have a distinctive ring structure with varying degrees of discoloration between the rings: the coloration in the halos increases, saturates, and finally diminishes with increasing ion dose. A total dose over time of 10^{13} alpha particles per square centimeter is required to initiate discoloration in laboratory experiments. Natural fluences must be four to seven times greater in order to initiate the same effect in micas, suggesting annealing of radiation damage over geologic periods. Thus, the size of a halo is limited first by the range of the alpha particle and then by the dissipation of the fluence (inversely proportional to distance) below a critical threshold value required for discoloration. *See* ALPHA PARTICLES; METAMICT STATE.

The early interest in pleochroic halos was in their use for geologic age dating. Careful attempts

were made to correlate the halo color with the alpha-irradiation dose in order to estimate the age of the enclosing mineral. Additionally, the constant size of the rings of the uranium and thorium halos for minerals of different ages was taken as evidence that the decay constants for radionuclides used in age dating had remained constant throughout geologic time. Thermal annealing of the halos has been used to model the thermal histories of rock units. *See* FISSION TRACK DATING.

Of great interest are the giant halos (50–100 μm in radius) or the Y halos, which lack the distinctive ring structure of uranium and thorium halos. These halos were attributed to extinct or nearly extinct nuclides, such as plutonium-244. The much larger halos (up to 100 μm) might also have been caused by highly energetic alpha particles that accompany the spontaneous fission of plutonium-244 or by alpha decay of heavy elements (atomic number 126, 127, 116, or 124). Various interpretations suggest mechanical and chemical mechanisms for the enlargement of the discolored areas of the halo. *See* ATOMIC NUMBER; MICA; PLEOCHROISM; RADIOACTIVE MINERALS. Rod Ewing

Bibliography. R. V. Gentry, Radioactive halos, *Annu. Rev. Nucl. Sci.*, 23:347–362, 1973; C. H. Holbrow, Haloes from plutonium minerals and fission alphas?, *Nature*, 265:504–508, 1977; M. A. K. Lodhi (ed.), *Superheavy Elements*, 1978; C. Moazed, R. M. Spector, and R. F. Ward, Polonium radiohalos: An alternative interpretation, *Science*, 180:1272–1274, 1973.

Pleochroism

In some colored transparent crystals, the effect wherein the color is quite different in different directions through the crystals. In such a crystal the absorption of light is different for different polarization directions. Tourmaline offers one of the best-known examples of this phenomenon. In colored transparent tourmaline the effect may be so strong that one polarized component of a light beam is wholly absorbed, and the crystal can be used as a polarizer. For a fuller discussion of the effect *see* DICHROISM; POLARIZED LIGHT; TRICHROISM. Bruce H. Billings

Plesiosauria

An order of extinct aquatic diapsid reptiles within the infraclass Sauroptrygia, common throughout the world during the Jurassic and Cretaceous periods. These large carnivores are characterized by long, paddle-shaped limbs and short, dorsoventrally compressed bodies (see **illus.**). In contrast with nothosaurs (more primitive members of the Sauropterygia), plesiosaurs have greatly expanded ventral portions of both pectoral and pelvic girdles to provide large areas for the attachment of muscles to move the limbs anteriorly and posteriorly. Lateral undulation of the trunk was severely restricted by elaboration of the ventral scales. The great length of the limbs indicates that they must have been used symmetrically, rather than alternatively as in primitive reptiles. The configuration of the skeleton suggests that plesiosaurs swam in the manner of sea lions by horizontal retraction of the limbs, rather than like sea turtles and penguins which use the limbs like wings to fly through the water.

Approximately 40 genera of plesiosaurs are recognized. They are divided into two large groups: The plesiosauroids include the most primitive genera and others that have small heads and very long necks. Among the plesiosauroids, the elasmosaurids had as many as 76 cervical vertebrae. The pliosauroids had shorter necks but larger skulls. The skull of *Kronosaurus* exceeded 9 ft (2.7 m) in length. Both elasmosaurs and pliosaurs persisted until the end of the Mesozoic.

The common attributes of the loss of the lower temporal bar and the superficial position of the scapula relative to the clavicle support relationship between the plesiosaurs and the Triassic nothosaurs, but known members of these two groups differ significantly in limb proportions and the structure of the girdles. Intermediate forms might be expected in the latest Triassic, but few remains of marine reptiles are known from this time interval. *See* DIAPSIDA; NOTHOSAURIA; REPTILIA; SAUROPTERYGIA. Robert L. Carroll

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1987.

Pleuronectiformes

A very distinctive order of actinopterygian fishes, also known as the Heterosomata or commonly the flatfishes, which in turn are known by a variety of common names: halibut, turbots, plaice, flounders, soles, tonguefishes, and others. All adults are characterized by a flattened (laterally compressed) body with both eyes located on the same side of the head. Larval flatfishes are bilaterally symmetrical,

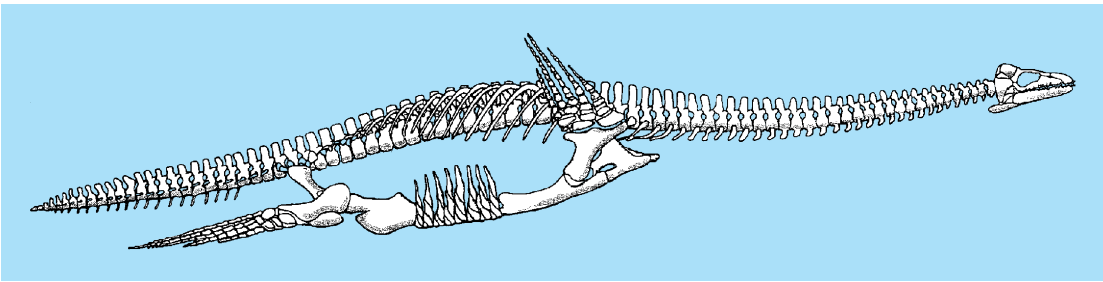

Skeleton of the plesiosauroid *Cryptocleidus*, about 10 ft (3 m) long.

but early in development one eye migrates across the top of the head to the other side. The asymmetry is reflected in complex modifications of the skull bones and associated muscles and nerves, as well as dentition, squamation (arrangement of scales), visceral anatomy, and paired fins. A given family is usually either dextral (eyes on the right side) or sinistral (eyes on the left side); however, some families have both dextral and sinistral members even at the species level. Flatfishes are further characterized by pigmented upper side; blind side usually lacking pigment; cycloid or ctenoid scales; tubercles in some species; long-based dorsal and anal fins, with the dorsal fin base extending to or beyond the eyes (except in the family Psettodidae); body cavity quite reduced in size; usually six or seven branchiostegal (gill bone) rays; and a swim bladder in adults of only a few species. *See* ACTINOPTERYGII; ANIMAL SYMMETRY; SWIM BLADDER.

Flatfishes are invariably benthic (bottom-dwelling) as postlarvae, spending their entire adult lives swimming on their side (except *Psettodes*, which occasionally swims upright), with the eyeless side down. They are carnivorous predators, keeping much to the bottom. The eyes are protrusible, allowing these fishes to see while buried in the substrate awaiting a wandering prey. In pursuit of prey, the flatfishes move far from the substrate. Worldwide many species of flatfish are valued as food and are important to the commercial fisheries industry. *See* MARINE FISHERIES.

The order Pleuronectiformes comprises about 678 extant species in about 134 genera, 14 families, and two suborders. About 10 species of soles are thought to be limited to freshwater habitats; a few more, primarily freshwater species enter estuaries or marine water, and another 20 that are normally marine species occasionally enter freshwater.

The families below were selected to represent the diversity of the pleuronectiforms.

Suborder Psettodoidei. The family Psettodidae (spiny turbots) is represented by a single genus, *Psettodes* (**Fig. 1**), and three species. It occurs in marine waters of tropical western Africa, the eastern Atlantic, eastern Africa and the Red Sea, and the western Pacific. It is distinguished from other flatfishes by a combination of the following characters: dorsal fin not extending onto the head to or past the eye; spinous anterior dorsal and anal rays; nearly symmetrical pelvic fins, with one spine and five soft rays; 15 branched caudal fin rays; sinistral or dextral eyes; very large mouth, with maxilla extending far beyond the eyes; large supramaxilla; barbed jaw teeth; gill arches with groups of teeth; palatine with teeth; and distinct preopercular margin, not covered with skin. Maximum length is about 60 cm (24 in.).

Fig. 1. Indian spiny turbot (*Psettodes erumei*) from the Indo-West Pacific. (*Photo by Jack E. Randall*)

Fig. 2. Ocellated flounder (*Pseudorhombus dupliciocellatus*) from the Indo-West Pacific. (*Photo by Jack E. Randall*)

Suborder Pleuronectoidei. In this group, the dorsal fin extends onto the head at least to the eyes, and the dorsal and anal fins lack spines.

Family Paralichthyidae (sand flounders). The paralichthyids (**Fig. 2**) are primarily marine fishes of the Atlantic, Indian, and Pacific oceans, and rarely enter freshwater. They can be distinguished from other flatfishes by the following combination of characters: sinistral eyes in most species; short and nearly symmetrical pelvic fin bases (but position of bases is variable between species); and branched pectoral rays. Maximum length is 75 to 152 cm (30 to 60 in.). There are about 16 genera and about 105 species.

Family Pleuronectidae (righteye flounders). These are marine (occasionally brackish-water, rarely freshwater) fishes of the Arctic, Atlantic, Indian, and Pacific oceans that are distinguished from other flatfishes by the following characters: eyes almost always dextral; origin of dorsal fin above the eyes; symmetrical pelvic fins; usually no oil globule in the egg yolk; and a well-developed lateral line on both sides. There are 23 genera with about 60 species. The commercially important and large halibuts belong to this family, including *Hippoglossus hippoglossus* (Atlantic halibut), with a maximum published weight of 320 kg (700 lb), and *H. stenolepis* (Pacific halibut), weighing in at 363 kg (800 lb).

Family Bothidae (lefteye flounders). Known from the tropical and temperate waters of the Atlantic, Indian, and Pacific oceans, these flatfishes are distinguished by sinistral eyes; pelvic fin base on eye side longer than on eyeless side, on midventral line, and origin anterior to the base on eyeless side; unbranched pectoral and pelvic fin rays; pelvic fins without a spine; and eggs with a single oil globule in the yolk. Twenty genera are known, with about 140 species, some of which are commercially important. Maximum total length is 3.5 to 46 cm (1.4 to 18 in.), depending on the species.

Sole families. Three families, Achiridae, Soleidae, and Cynoglossidae, form a monophyletic group, with

Achiridae being the primitive sister group to Soleidae and Cynoglossidae. In soles, the skin of the lower jaw and interopercle (a gill cover bone) is continuous ventrally and covers the isthmus and branchiostegals.

Achiridae (American soles) includes shore fishes in the eastern central Pacific and western Atlantic, and freshwater fishes of temperate and tropical waters of South America. They are characterized by the following combination of characters: dextral eyes; margin of preoperculum (another gill cover bone) detectable by a shallow groove; dorsal and anal fins free from the caudal fin; right pelvic fin joined to the anal fin; and eyed side of the lower lip having a conspicuous, variously fringed fleshy rim. American soles are carnivores, feeding primarily on fishes and benthic invertebrates. Maximum total length is 4 to 28 cm (1.6 to 11 in.), depending on the species. Seven genera are known, with about 33 species.

Soleidae (soles) occurs widely in cold temperate to tropical waters of the eastern Atlantic, the Indo-West Pacific, and Australia. The following combination of characters readily identifies the soles: dextral eyes; preoperculum completely concealed by skin and scales; dorsal fin origin above or before the eyes; dorsal and anal fins free from or united with the caudal fin; pectoral fins present or absent; and pelvic fins free from the anal fin. Maximum total length is about 2.5 to 110 cm (1 to 43 in.), depending on the species. Thirty-five genera have been identified, with 135 species.

Cynoglossidae (tonguefishes; also called tongue-soles) is primarily a marine family of the tropical and subtropical seas of the world. It is distinguished from other soles by the following characters: very small sinistral eyes that are usually close together; preopercle margin covered with skin and scales; dorsal and anal fins confluent with the pointed caudal fin; pelvic fin of eyeless side having four rays attached along midventral line, linked to the anal fin in some species; pelvic girdle and fin on the eyed side absent in some species; pectoral fins absent; and asymmetrical mouth. Most species occur in shallow water or estuaries; however, some species of the genus *Symphurus* may be found at depths to 1900 m (6240 ft). Five species are known primarily from rivers. Maximum total length of most species is from 4 to 30 cm (1.6 to 12 in.), rarely 40 to 48 cm (16 to 19 in.). Three genera are known, with about 127 species.

Herbert Boschung

Bibliography. F. Chapleau, Pleuronectiform relationships: A cladistic reassessment, *Bull. Mar. Sci.*, 52(1):516–540, 1993; R. Froese and D. Pauly (eds.), FishBase, World Wide Web electronic publication, <www.fishbase.org>, version (05/2005); T. A. Munroe, Bothidae (pp. 1885–1895), Scophthalmidae (1896–1897), Paralichthyidae (1898–1921), Poecilopsettidae (1922–1923), Achiridae (1925–1933), and Cynoglossidae (1934–1959), in K. E. Carpenter (ed.), *The Living Marine Resources of the Western Central Atlantic: FAO Species Identification Guide for Fishery Purposes*, vol. 3, pt. 2, FAO, Rome, 2003 (dated 2002); T. A. Munroe, Systematic diversity of the Pleuronectiformes (pp. 10–41), Distributions and biogeography (pp. 42–67), in R. N. Gibson (ed.), *Flatfishes: Biology and Exploitation*, Fish and Aquatic Resources Series 9, Blackwell Publishing, Oxford, England, 2005; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; R. T. C. Ramos, Achiridae (American soles), pp. 666–669 in R. E. Reis, S. O. Kullander, and C. J. Ferraris, Jr. (eds.), *Checklist of the Freshwater Fishes of South and Central America*, EDIPUCRS, Porto Alegre, Brasil, 2003.

Pliocene

The youngest of the five geological epochs of the Tertiary Period. The Pliocene represents the interval of geological time (and rocks deposited during that time) extending from the end of the Miocene Epoch to the beginning of the Pleistocene Epoch of the Quaternary Period. Modern time scales assign the duration of 5.0 to 1.8 million years ago (Ma) to the Pliocene Epoch. *See* MIOCENE; PLEISTOCENE; QUATERNARY; TERTIARY.

Charles Lyell introduced Pliocene in 1833 to characterize the youngest of the Tertiary strata. He based his subdivision of geological eras on the degree of faunal similarity of those times with the modern molluscan assemblages. For the Pliocene he identified an older Pliocene (with more than 50% extant species) and a newer Pliocene (with up to 95% modern species). In 1839 he substituted the term Pleistocene for the newer Pliocene, which he later tried to abandon in favor of the term post-Pliocene; but this suggestion was largely ignored, and the use of the term Pleistocene became widespread in the geological literature. Deposits in the hills of the lesser Apennines of northern Italy were shown to characterize the Pliocene type assemblages during Lyell's time.

Pliocene marine sediments are commonly distributed along relatively restricted areas of the continental margins and in the deep-sea basins. Continental margin sediments are most often terrigenous and range from coarser-grained sandstone to finer-grained mudstone and clay. Major rivers of the world, such as the Amazon, Indus, and Ganges, contain thick piles of Pliocene terrigenous sediments in their offshore fans. The Pliocene deep-sea sediments are carbonate-rich (commonly biogenic oozes) and are often very thick (up to 5000 m or 16,400 ft). *See* BASIN; CONTINENTAL MARGIN.

CENOZOIC	QUATERNARY	Holocene
		Pleistocene
	TERTIARY	Pliocene
		Miocene
		Oligocene
		Eocene
		Paleocene
MESOZOIC	CRETACEOUS	

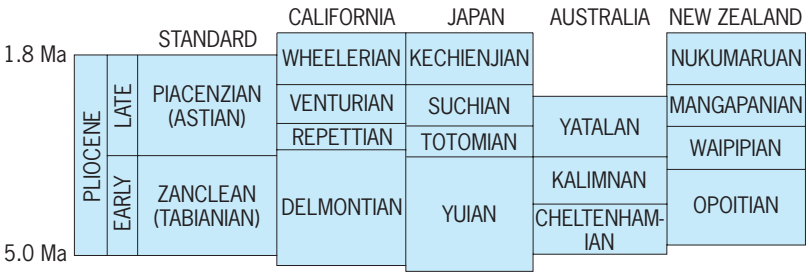

Fig. 1. Pliocene stages and their temporal equivalents in different regions.

Pliocene subdivisions. Modern stratigraphic usage subdivides the Pliocene Epoch into two standard stages, the lower, Zanclean stage and the upper, Piacenzian stage (**Fig. 1**).

In 1868 G. Seguenza proposed the Zanclean stage as the lower stage of the Pliocene, characterized by the deposits of Trubi Marl in Sicily that contained a diverse marine microfauna. The Zanclean is considered an equivalent of the Tabianian stage, introduced by K. Mayer-Eymer in that same year, with a type area in northern Italy. The Tabianian contained a less diverse and more restricted cooler-water fauna, less suitable for broader correlations, thus leading to the use of Zanclean as the standard stage for the lower Pliocene over the Tabianian. In 1973 M. Cita proposed a type section for the Zanclean in the Trubi Marl at Capo Rosello in Sicily.

Mayer-Eymer proposed the Piacenzian in 1858 with the type area in northern Italy. It was first proposed as a lower subdivision of the Astian stage that had been introduced by P. de Rouville in 1853 as an equivalent of Lyell's older Pliocene. The Piacenzian was later upgraded to a stage level by E. Renevier, with a type area designated near Castell' Arquato by M. Pareto in 1865. The sands and blue marls of type Piacenzian contain diverse benthic marine fauna as well as several planktonic foraminifer species. The equivalent interfingering strata of the Astian (type area near Asti, northern Italy) are largely calcareous sand, conglomerate, and marl, containing shallow-water benthos and only minor planktonic constituents. *See* MARL.

The Romanian and Dacian stages are sometimes used for the Pliocene in the Paratethys region (eastern and southeastern Europe) and are considered time-equivalent to Zanclean and Piacenzian, respectively. The Kimmerian and Akchagylian stages (Moldavian and Khaprovian substages for the latter) are often used as approximate equivalents for the lower and upper Pliocene in the former Soviet republics. Other regionally used stage terms of the Pliocene include the lower brackish-water Sarmatian stage and the upper nonmarine Pontian stage of the Vienna basin and the western Paratethys; the Piacenzian equivalent Redonian stage in Brittany; and the Anversian stage for the younger Pliocene in Belgium. In Japan, the Yuian, Totomian, Suchian, and younger part of the Kechienjian stages span the Pliocene, and in Australia it is subdivided into the Cheltenhamnian, Kalimnan, and Yatalan stages. New Zealand stratigraphers assign the Opoitian, Wapipian, Mangapanian, and the lower part of Nakumaruan stages to the Pliocene. In California, the regional Pliocene stages include the Delmontian, Repettian, Venturian, and the younger part of the Wheelarian.

Paleontologists working with vertebrate fauna often use their own age and stage terminology to characterize particular mammalian assemblages. In European Pliocene they identify the Ruscinian and the younger part of Villafranchian ages; while in North America they consider the oldest part of Hemphillian, the Blancan, and the youngest part of Irvingtonian ages as characteristic of the mammalian ages of the Pliocene.

Tectonics, seas, and climate. The most notable tectonic events in the Pliocene include the beginning of the third and last phase of the Himalayan uplift, the Attican orogeny that began in the late Miocene and continued into the Pliocene, and the Rhodanian and Walachian orogenies that occurred during the later Pliocene.

The latest Miocene is marked by a global cooling period that continued into the earliest Pliocene, and there is evidence that the East Antarctic ice sheet had reached the continental margins at this time. The Miocene Epoch terminated with the isolation of, and a salinity crisis in, the Mediterranean. At this time, the Mediterranean was cut off from the open ocean due to the emergence of the sill blocking the Gibraltar Straits, due to uplift or falling global sea level. The desiccated Mediterranean developed high salinities, and a large volume of evaporites was deposited. The onset of the Pliocene is marked by the reconnection with the Atlantic, when normal marine conditions returned to the Mediterranean. *See* SALINE EVAPORITES.

The global sea level had been falling through the late Miocene, and with the exception of a marked rise in the mid-Zanclean, the trend toward lowered sea levels continued through the Pliocene and Pleistocene. The mid-Zanclean sea-level rise (3.5–3 Ma) was also accompanied by a significant global warming event, which is evidenced by the paleontological and glacial records of both the Arctic and Antarctic regions and New Zealand. The oxygen isotopic data, which record the prevailing sea surface temperatures and total ice volume on the ice caps, show little variations in the Equatorial Pacific during the middle Pliocene. *See* GEOLOGIC THERMOMETRY.

By early Pliocene time, the major surface circulation patterns of the world ocean and the sources of supply of bottom waters were essentially similar to their modern counterparts. The only significant difference was the continuing connection of the Central Atlantic with the Pacific at the Isthmus of Panama, which did not close until the mid-Pliocene at about 3.2 Ma. The closure of the isthmus had a profound effect on the North Atlantic circulation patterns. From the marine oxygen isotopic record, the initiation of the Arctic glaciation has been dated at 3.5 Ma, but the continental record shows that land glaciation on North America did not start until 3.2 Ma. This time-lag effect may be linked to the closing of the Isthmus of Panama and the strengthening

of the North Atlantic Gulf Stream, which would then be displaced farther north. Until the closure of the isthmus, the terrestrial and marine ice mass generation systems were not coupled. But the combined effect of the northward displacement of the Gulf Stream and the initial development of the Arctic ice cap, which resulted in increased availability of moisture to northern high latitudes and lowered temperatures, led to the development of the first terrestrial ice sheet. From then on, the terrestrial and marine ice mass systems became coupled. A significant ice cap on the Arctic region, however, did not develop until 2.5 Ma. In the Southern Hemisphere, two major glacial advances of the East Antarctic ice sheet have been documented for the periods 3.2–2.7 Ma and 2.1–1.6 Ma.

The third phase of the uplift of the Himalayas that began in the Pliocene accentuated the monsoon-type circulation in the Arabian Sea and North Indian Ocean. The monsoonal system in this region had already been activated after the Miocene (second) phase of the Himalayan orogeny.

Thus, by mid-Pliocene time the conditions were in place that would lead to glacially dominated climates of the Quaternary Period.

Life. By Pliocene time, much of the marine and terrestrial biota had essentially evolved its modern characteristics. The late Pliocene cooling led to the expansion of cooler-water marine assemblages of the higher latitudes into lower latitudes, particularly the foraminifers, bivalves, and gastropods. At the onset of cooling, the warm-water-preferring calcareous nannoplankton group of discoasters began waning in the late Pliocene and became extinct at the close of the epoch. *See* BIVALVIA; COCCOLITHOPHORIDA; FORAMINIFERIDA; GASTROPODA.

The widespread grasslands of the Pliocene were conducive to the proliferation of mammals and increase in their average size. The mid-Zanclean sea-level rise led to the geographic isolation of many groups of mammals and the increase in endemism. But the late Pliocene-Pleistocene lowering of sea level facilitated land connections and allowed extensive mammalian migration between continents. The closing of the Panama isthmus in the late Pliocene also allowed interchanges between North and South America. Armadillos, capybaras, ground sloths, and porcupines, among others, migrated from south to north, and horses, elephants, cats, dogs, deer, and northern rodents went south. The arrival of the North American mammals led to increased competitive pressure and extinction of many typically South American groups, including the notoungulates and litopterns. Horses evolved and spread widely in the Pliocene (**Fig. 2**). The late Miocene North American horse genus, *Hipparion*, became extinct in the mid-Pliocene after spreading into the Old World during the latest Miocene and early Pliocene. The modern horse genus, *Equus*, had appeared before the demise of *Hipparion*. *See* MAMMALIA.

The Pliocene Epoch also saw the appearance of several hominid species that are considered to be directly related to modern human ancestry. The earliest hominid bones have been discovered from Baringo, Kenya, in sediments that are dated to be of earliest Pliocene age. After this first occurrence, a whole suite of australopithicine species made their appearance in the Pliocene. These include *Ardipithecus ramidus*; and *Australopithecus anamensis*, *A. afarensis*, *A. bahrelghazali*, *A. africanus*, *A. boisei*, and *A. robustus*. Before the beginning of the Pleistocene glaciation, the first *Homo* species, *Homo habilis*, had already evolved in Africa, first appearing around 2 Ma, and *Homo erectus* appeared on the scene near the Pliocene-Pleistocene boundary. Both of these species are considered to be the precursor of modern humans. *See* AUSTRALOPITHECINE; FOSSIL HUMANS; PALEOBOTANY; PALEONTOLOGY.

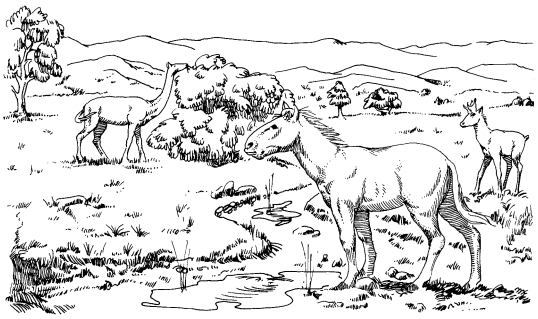

Fig. 2. Pliocene scene in the San Joaquin Valley, California, with ancestral species of camel (*Parocamelus*), horse (*Pliohippus*), and pronghorn (*Sphenophalos*). The spread of grasslands in the Pliocene was conducive to the proliferation of grazing mammalian species. (*After R. A. Stirton, Time, Life and Man, John Wiley, 1959*)

Bilal U. Haq

Bibliography. W. A. Berggren and J. A. van Couvering, The late Neogene, *Palaeogeog. Palaeoclimat. Palaeoecol.*, 16:1–216, 1974; B. U. Haq and F. W. B. van Eysinga, *Geological Time Table*, Elsevier, Amsterdam 1998.

Plum

Any of the smooth-skinned stone fruits grown on shrubs or small trees. Plums are widely distributed in all land areas of the North Temperate Zone, where many species and varieties are adapted to different climatic and soil conditions.

Types and varieties. There are four principal groups: (1) Domestica (*Prunus domestica*) of

Fig. 1. Branch of a Damson plum.

Fig. 2. Cluster of prunes decayed by brown rot. Fungus spore masses are on the fruit, and there is an exudation of gum at the base of the stems where the pathogen has entered the shoot. (*Courtesy of J. M. Ogawa*)

European or Southwest Asian origin, (2) Japanese or Salicina (*P. salicina*) of Chinese origin, (3) Insititia or Damson (*P. insititia*) of Eurasian origin (**Fig. 1**), and (4) American (*P. americana* and *P. hortulana*). The Domesticas are large, meaty, prune-type plums, the principal varieties being Agen or French prune (grown in California) and Italian or Italian hybrids (grown primarily in Oregon, Washington, Idaho, and Michigan).

A prune is a plum which dries without spoiling. Commonly but incorrectly the term "prune" is applied to all Domestica-type plums, whereas the term "plum" is used only for Japanese-type plums. Japanese plums are typically round, reddish or yellow, and juicy, and are represented by such varieties as Beauty, Santa Rosa, Duarte, Burbank, Wickson, and Satsuma. Insititias (Damson), represented by Shropshire, are small, meaty fruits grown sparingly for jam. American plums are small, watery fruits of low quality, represented by De Soto, Pottawattomi, and Golden Beauty, and valued chiefly for cold hardiness of the tree.

Propagation and culture. Plums are propagated by budding on seedlings of the myrobalan plum (*P. cerasifera*) and less commonly on the peach and certain strains of *P. domestica*. *See* PEACH; PLANT PROPAGATION.

The Domestica and Insititia plums are slightly hardier than the peach; the Japanese plums have about the same hardiness as the peach; and native American plums are considerably hardier than the other groups.

Trees of the Domestica and Damson types may grow 20–25 ft (6–7.5 m) high unless kept lower by training and pruning. They are usually planted about 20 ft (6 m) apart, and are adapted to a wide range of soils, from light to heavy. Trees of the Japanese type are smaller (15–20 ft or 4.5–6 m) and are best adapted to lighter soils. They are normally planted 18–20 ft (5.4–6 m) apart, but in the hands of skilled horticulturists they are adaptable to close spacings and controlled shapes. Most plums require cross-pollination. *See* REPRODUCTION (PLANT).

Production and economic importance. California is the leading producer (90%) of the prune-type Domestica plums, followed by Oregon, Washington, Michigan, and Idaho. In states other than California, where 100% are dried as prunes, 52% of the prune-type plums are sold fresh, 40% canned, and 8% dried.

R. Paul Larsen

Plum and prune diseases. Brown rot, a fungus disease caused by *Monilinia fructicola* and *M. laxa*, is probably the most widespread and serious disease of these crops (**Fig. 2**). Symptoms include blossom and twig blight and pre- and postharvest fruit rot. Injuries caused by the plum curculio in eastern states and twig borer in western states play an important role in fruit infection. Control of blossom blight is based on the application of a dormant eradicant spray and one or more protective sprays during bloom. Fruit decay can be prevented only by controlling the involved insects and the pathogens.

Other important fungus diseases include *Armillaria* root rot (caused by *A. mellea*), controlled mainly by soil fumigation and use of resistant rootstocks; black knot (caused by *Dibotryon morbosum*), a serious disease in eastern states that is controlled by gall removal and spraying; Valsa (Cytospora) canker (caused by *Valsa* ssp.), favored by freeze injury to the North and sunburning in the Southwest; and *Ceratocystis* canker (caused by *C. fimbriata*, an insect-vectored pathogen that invades mechanical-harvesting injuries).

Bacterial diseases of importance on plums and prunes are bacterial spot (caused by *Xanthomonas pruni*), bacterial canker (caused by *Pseudomonas*

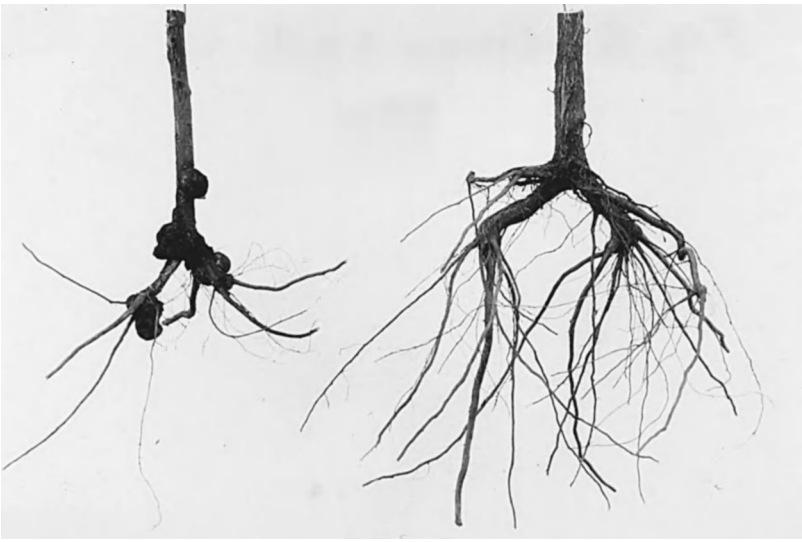

Fig. 3. Crown gall on plum rootstock at left. Tree at right was protected against infection by using the biocontrol agent *Agrobacterium radiobacter*. (*Courtesy of W. J. Moller*)

syringae), and crown gall caused by *Agrobacterium tumefaciens* (**Fig. 3**). Bacterial spot occurs only east of the Rocky Mountains, where it is especially damaging to Japanese plums. Bacterial canker is serious in the Pacific coast states on all types of plums and prunes. It is partially controlled through soil fumigation to kill nematodes that predispose trees to the disease, and by the use of peach rootstocks. Crown gall is biologically controlled in nurseries, where it is of greatest concern, by the use of *A. radiobacter*, which produces a bacteriostatic bacteriocin specific for the pathogen. *See* CROWN GALL.

Although plums and prunes are affected by several virus diseases—Prunus ring spot, line pattern, stem pitting, and diamond canker, to mention a few—these disorders mostly are of minor importance. The serious plum pox virus disease of Europe has not yet been found in North America. *See* FRUIT, TREE; PLANT PATHOLOGY.

Harley English

Bibliography. W. H. Chandler, *Deciduous Orchards*, 1957; N. F. Childers, *Modern Fruit Science*, 1973; U.S. Department of Agriculture, *Virus Diseases and Noninfectious Disorders of Stone Fruits in North America*, Agr. Handb. 437, 1976; E. E. Wilson and J. M. Ogawa, *Fungal, Bacterial, and Certain Nonparasitic Disease of Fruit and Nut Crops in California*, Division of Agricultural Sciences, University of California, 1979.

Plumbaginales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Caryophyllidae of the class Magnoliopsida (dicotyledons). The order consists of only the family Plumbaginaceae, with about 400 species. The plants are herbs or less often shrubs. The flowers are strictly pentamerous (that is, each floral whorl has five members); the petals are fused (sympetalous condition) and all alike in shape and size; the pollen is trinucleate; and there is a single basal ovule located in a compound ovary that has a single locule or cell. The Plumbaginaceae differ from most families of their subclass in their straight embryo, copious endosperm, absence of perisperm, and the presence of anthocyanin pigments instead of betalains. The family contains a few garden ornamentals, such as species of *Armeria*, known as thrift or sea-pink. *See* CARYOPHYLLIDAE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM; POLLEN.

Arthur Cronquist; T. M. Barkley

Pluto

Formerly considered the outermost planet, tiny Pluto is now recognized to be one of the largest and closest members of a disk of icy planetesimals that surrounds the solar system beyond the orbit of Neptune. For this reason, the International Astronomical Union resolved to reclassify Pluto in a new category, dwarf planet, in 2006.

Pluto was discovered on February 18, 1930, by C. W. Tombaugh at the Lowell Observatory, Flagstaff, Arizona, on photographic plates taken with a special astronomical camera as part of a systematic search. The presence of a planet beyond Neptune's orbit had been predicted independently by P. Lowell and W. H. Pickering (among others) on the basis of analyses similar to those that had led U. J. Leverrier and J. C. Adams to the prediction of Neptune. It was thought that there were perturbations in the motion of Uranus that Neptune alone could not explain. Pluto was found surprisingly near its predicted position, since it is now known that the mass of Pluto is far too small to have caused the (evidently spurious) perturbations. *See* NEPTUNE; PLANET.

Charon, Pluto's largest satellite, was discovered in 1978 by J. W. Christy and R. Harrington at the nearby Flagstaff station of the U.S. Naval Observatory. H. Weaver and a team of astronomers using the *Hubble Space Telescope* discovered two additional satellites, Hydra and Nix, in 2005 (**Fig. 1**).

Orbit. Pluto's orbit (**Fig. 2**) has a semimajor axis (mean distance to the Sun) of 5.9×10^9 km (3.7×10^9 mi), an eccentricity of 0.25, and an inclination of 17.1°. The inclination and eccentricity are larger than those of any of the planets. At a mean orbital velocity of 2.95 mi/s (4.7 km/s), it takes Pluto 248.0 years to make one revolution around the Sun. The large orbital eccentricity means that at perihelion Pluto is closer to the Sun than Neptune, but the orbits of the two bodies do not intersect.

Telescopic appearance. Pluto is visible only through fairly large telescopes since even at perihelion its visual magnitude is fainter than 13.5. This is thousands of times below the sensitivity of the unaided eye. Its visual magnitude at mean opposition is 15.1, thousands of times fainter than our visual threshold. Periodic variations in its brightness demonstrate that the surface of Pluto is covered with bright and dark markings and indicate a period of rotation of 6.4 days. This is longer than the rotational period of any of the planets except Mercury and Venus. *See* MAGNITUDE (ASTRONOMY).

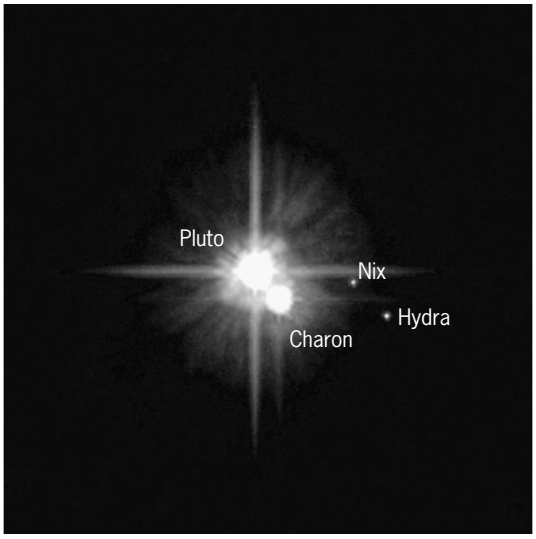

Fig. 1. Pluto and its three known moons, imaged by the *Hubble Space Telescope* on February 15, 2006. (*NASA, ESA, H. Weaver (JHUAPL), A. Stern (SwRI), HST Pluto Companion Search Team*)

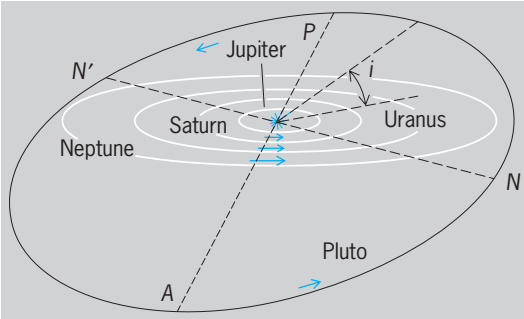

Fig. 2. Orbit of Pluto; a perspective view to show the inclination *I* and eccentricity of the orbit. *A*, aphelion; *P*, perihelion; *NN′*, line of nodes. (*After L. Rudaux and G. de Vaucouleurs, Astronomie, Larousse, 1948*)

Size and mass. The radius of Pluto is 1153±10 km (716±6 mi), making Pluto significantly smaller than the Earth's Moon (whose radius is 1738 km or 1080 mi), while Charon's radius is 603.5±1.5 km (375.0±0.9 mi). (The uncertainty in Pluto's radius includes the possibility that a lower atmosphere haze prevents the true surface of this dwarf planet from being measured by the occulation observation on which it is based.) Charon is thus about half the size of Pluto itself, making the most closely matched pair in the solar system. The density of Pluto is 2.0 g/cm^{-3}, while that of Charon is 1.65 g/cm^{-3}. These densities suggest compositions of ice and rock within the range exhibited by Saturn's regular satellites. Nix and Hydra are too small to permit determinations of size and mass.

Surface and atmosphere. The near-infrared spectrum of Pluto reveals absorption features of solid nitrogen (N_2), carbon monoxide (CO), and methane (CH_4) ice. From these absorptions it is possible to conclude that nitrogen is the dominant ice on Pluto's surface (some 20–100 times more abundant than the other ices), which means it must also be the major constituent of the planet's tenuous atmosphere. The nitrogen absorption in the spectrum indicates a surface temperature near 38 K (−391°F), consistent with an atmospheric surface pressure roughly 10^{-5} times that on Earth. Such a pressure was deduced directly from observations of a stellar occultation in 1988, which argues against the possible presence of a thick haze in Pluto's lower atmosphere. (If such a haze exists, it could affect the radius measurement and hence the deduction of the planet's mean density.) Its size, surface temperature, composition, and thin nitrogen atmosphere make Pluto very similar to Triton, Neptune's largest satellite. However, there are important differences such as the absence of frozen carbon dioxide (CO_2) on Pluto. *See* OCCULTATION.

Seasons on Pluto are modulated by the large eccentricity of the orbit. Pluto reached perihelion in 1989, which brought it closer to the Sun than Neptune. This means that during the 1988 occultation Pluto was observed at the warmest time in its seasonal cycle, when the atmospheric pressure might be expected to be at its maximum value.

However, observations of another stellar occultation by Pluto in 2002 revealed that the atmosphere was actually twice as massive as it had been in 1988. This result had been anticipated by some scientists, who had predicted that the changing orientation of Pluto as it moved around its orbit would expose polar deposits of frozen atmospheric gases to the Sun, if such deposits existed. The atmospheric pressure would then increase as the polar deposits sublimed. The same seasonal change in atmospheric pressure is observed on Mars, where the dominant gas is carbon dioxide (CO_2), which condenses out at the winter polar cap. On Pluto, this seasonal effect will soon be overwhelmed by the decrease in temperature as the dwarf planet moves farther from the Sun in its eccentric orbit, and the atmospheric pressure will gradually diminish as nitrogen again freezes out on Pluto's surface. Thus, to observe these processes it would be advantageous to explore the planet with a spacecraft as soon as possible, and such a mission is underway. Another stellar occultation, in 2006, found that the atmosphere had not changed further since 2002. *See* MARS.

Satellites. Pluto has three small satellites: Charon, Nix, and Hydra, with Charon far larger than the other two (see **table**; **Fig. 3**). Demonstrating another striking difference with Pluto, Charon apparently has little if any frozen methane on its surface, which instead exhibits the spectral signature of water ice. The orbit of Charon is unique in the solar system in that the satellite's period of revolution is identical to the rotational period of the planet. Thus an inhabitant of Pluto who lived on the appropriate hemisphere would see Charon hanging motionless in the sky. From the opposite hemisphere the satellite would be invisible. (On Earth, artificial satellites are put into such geostationary orbits for communications and television broadcasting.) *See* COMMUNICATIONS SATELLITE; SATELLITE (ASTRONOMY); SATELLITE (SPACECRAFT).

Pluto and the Kuiper Belt. In 1992, the first of what is proving to be a huge number of icy planetesimals

Satellites of Pluto

Name	Orbit radius, km (mi)	Period, days	Eccentricity	Inclination*	Radius, km (mi)	Magnitude at mean opposition
Charon	19,571 (12,161)	6.4	~0	~0°	603.5±1.5 (375.0±0.9)	17.3
Nix	48,675 (30,245)	24.9	~0	~0°	~45 (28)?†	24.6
Hydra	64,780 (40,252)	38.2	~0	~0°	~45 (28)?†	24.4

*With respect to Pluto's equatorial plane.
†Estimated from assumed reflectivity.

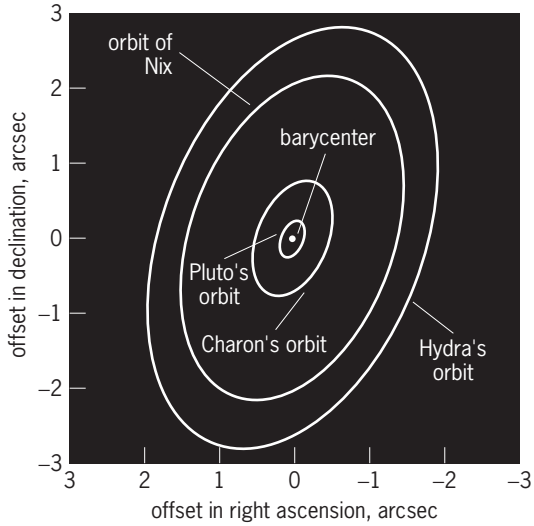

Fig. 3. Orbits of four bodies in the Pluto system (Pluto and its satellites, Charon, Nix, and Hydra) about the barycenter (center of mass) of the system, as seen from Earth. (*NASA; ESA; W. J. Merline, Southwest Research Institute; Pluto Companion Search Team*)

(200 km or 125 mi in diameter or less) was discovered orbiting the Sun at distances comparable to or greater than the orbit of Pluto. These are members of the Kuiper Belt, a disk-shaped distribution of material left over from the formation of the solar system. These objects appear to be representative of the solid components required to initiate the growth of giant planets. Held in the cosmic deep freeze of trans-Neptunian space, they must be among the most primitive, unprocessed objects still circulating in the solar system. Collisional fragments occasionally make their way into the inner solar system, where they become the short-period comets. In the course of this journey, they lose some of their most volatile constituents but still represent a unique resource for studying the early days of the solar system. The European Space Agency (ESA) mission *Rosetta* is scheduled to carry out a detailed investigation of a comet in 2012. *See* COMET; SOLAR SYSTEM; SPACE PROBE.

The existence of these small bodies (Kuiper Belt objects) finally provides a proper context for Pluto itself. Because of its small size and the high inclination and eccentricity of its orbit, Pluto has been an "ugly duckling" among the planets. As one of the largest members of the Kuiper Belt, however, it becomes a majestic "swan," swimming serenely at the boundary between the planets and this disk of debris left over from the formation of the giant planets.

By the middle of 2006, astronomers had discovered six Kuiper Belt objects with diameters of approximately 1000 km (625 mi), comparable in size to Charon, and at least one, Eris, that may be slightly larger than Pluto. They have also found that several of the Kuiper Belt objects, including Eris, have satellites. Of the larger bodies, Sedna and Eris are particularly interesting. The diameter of Sedna is 1000–1500 km (625–940 mi), and is still poorly known as it depends on an assumed reflectivity. (The more reflective the surface, the smaller the object needs to be to produce the same brightness.) Sedna's orbit has a very high eccentricity of 0.85 and an inclination of 12°, but it is the perihelion distance of 76 astronomical units that is most remarkable: Sedna's closest approach to the Sun is more than twice as distant as that of Pluto and is the farthest of any known Kuiper Belt object. This orbit is well beyond the gravitational influence of Neptune, and raises the questions of how Sedna got there and whether there are still more large bodies in such distant orbits. Eris is indeed larger, but its orbit lies somewhat closer to the Sun. Current models of solar system formation do not allow enough material at these enormous distances to form objects the size of these two Kuiper Belt objects, so they must have been scattered from closer orbits. Perhaps other large icy bodies suffered the same fate and ongoing searches will reveal them.

In 2003, the National Aeronautics and Space Administration (NASA) authorized the *New Horizons* Pluto-Kuiper Belt mission. A single spacecraft was launched on January 19, 2006, and is scheduled to fly past Pluto and Charon in 2015, making measurements and recording images. The spacecraft will then venture forth into the Kuiper Belt and attempt to fly past one or more of these small distant worlds. *See* KUIPER BELT.

Tobias C. Owen

Bibliography. J. K. Beatty, C. C. Petersen, and A. Chaikin (eds.), *The New Solar System*, 4th ed., Cambridge University Press, 1999; W. G. Hoyt, *Planets X and Pluto*, University of Arizona Press, 1980; D. Morrison and T. Owen, *The Planetary System*, 3d ed., Addison Wesley, 2003; S. A. Stern and J. Mitton, *Pluto and Charon: Ice Worlds on the Ragged Edge of the Solar System*, Wiley, 1998; S. A. Stern and D. J. Tholen (eds.), *Pluto and Charon*, University of Arizona Press, 1997; D. A. Weintraub, *Is Pluto a Planet?: A Historical Journey Through the Solar System*, Princeton University Press, 2006.

Pluton

A solid rock body that formed by cooling and crystallization of molten rock (magma) within the Earth. Most plutons, or plutonic bodies, are regarded as the product of crystallization of magma intruded into surrounding "country rocks" within the Earth (principally within the crust). During the middle part of the twentieth century some geologists believed that many igneous or plutonic rock bodies were the result of ultrametamorphism (granitization). However, that model has largely been abandoned.

Igneous rock bodies are referred to generally as either extrusive or volcanic on one hand, or as intrusive or plutonic on the other, although the term volcanic is sometimes also used to refer to small, shallow intrusive bodies associated with volcanoes. *See* IGNEOUS ROCKS; MAGMA.

Types. Plutons occur in a nearly infinite variety of shapes and sizes, so that definition of types is arbitrary in many cases. In general, two modes of emplacement can be recognized with regard to the country rock. Concordant plutons are intruded between layers of stratified rock, whereas the more

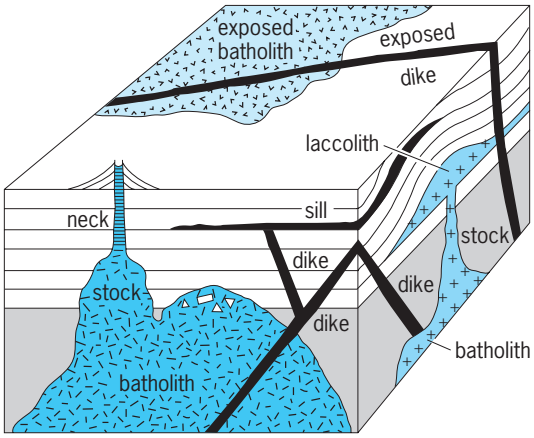

Schematic block diagram showing the occurrence of the more common types of plutons.

common discordant plutons are characterized by boundaries that cut across preexisting structures or layers in the country rock. The principal types of concordant plutons are sills, laccoliths, and lopoliths; the principal types of discordant plutons are dikes, volcanic necks or plugs, stocks, and batholiths. In complex plutons, characteristics of more than one type may be found from place to place. The more common types are shown in the **illustration**.

Dike. A dike is a tabular, discordant pluton. Dikes can range in size from a few millimeters wide by a few centimeters long to major structures several hundred meters wide and several hundred kilometers long. Many dikes tend to be sheetlike and nearly planar, whereas others can follow complex pathways through zones of weakness in country rocks. The concentric fracture systems found in many volcanic regions can also give rise to ring dikes.

Sill. While they are also tabular plutons, sills are by definition concordant with the stratification or layering in the surrounding country rocks. True sills are much less common than dikes.

Laccolith. Larger concordant plutons, particularly more viscous (granitic) ones, may tend to bulge upward, creating a domal shape. Such bodies are called laccoliths and may show well-defined feeder dikes or sills. Laccoliths tend to be several kilometers in diameter and several hundred meters thick in the center (1 km = 0.6 mi; 1 m = 3.3 ft).

Lopolith. A second type of large concordant pluton is the lopolith, which is characterized by a saucerlike or funnellike shape. They are associated primarily with large concordant intrusions of mafic (gabbroic) rock, and true lopoliths are relatively rare.

Plug (neck). The central conduits to most volcanoes become filled, generally, with solidified magma, and the resulting pipelike igneous bodies are referred to as volcanic plugs or necks. They tend to possess steep, discordant sides and roughly equidimensional horizontal sections; they may range from several meters to hundreds of meters in diameter.

Batholiths and stocks. A complete range exists in sizes of discordant plutons, and the designation of a particular pluton as stock or batholith may depend upon degree of exposure due to uplift and erosion. Plutons smaller than 40 mi^2 (100 km^2) and 6 mi (10 km) across are referred to generally as stocks, and larger ones are referred to as batholiths. Batholiths may be considerably larger than 40 mi^2 (100 km^2), and the term batholith has also come to refer to very large plutonic masses (up to several hundred kilometers in largest dimension) that are made up of many smaller, compositionally distinct individual plutons (stocks or batholiths). Batholiths do not extend indefinitely into the Earth, but possess a finite thickness, which ranges from 6 mi (10 km) or less to 18 mi (30 km). In some cases they show well-defined lower boundaries (floors), but in others they may gradually grade downward into root zones where the original magma was derived by melting of older crustal rocks. Unfortunately the deep zones of most batholiths are not exposed, so details of their deep structure must be inferred from petrologic and geophysical studies.

Origin. The modern view of most plutons is that they are the result of intrusion of molten rock (magma) into surrounding country rock, with subsequent cooling of the magma to form plutonic igneous rock. Several mechanisms of intrusion are known or proposed. The most simple ones, pertaining to smaller plutonic bodies, are forceful injection or passive migration into fractures. Larger plutons may form by several processes. For example, less dense magma may migrate upward along a myriad of channelways to accumulate as a large molten body within the upper crust. Further migration could occur by forceful injection, by stoping (a process where the magma rises as blocks of the roof of the magma chamber break off and sink), and by diapiric rise, where country rocks flow around the upward-moving magma body. The variations and interactions of such processes are complex, so that no two large plutons formed in exactly the same way. *See* PETROLOGY.

W. Randall Van Schmus

Bibliography. D. S. Barker, *Igneous Rocks*, 1983; M. G. Best, *Igneous and Metamorphic Petrology*, 1982; H. Blatt and R. Tracy, *Petrology (Igneous, Sedimentary, and Metamorphic)*, 2d ed., 1996; E. G. Ehlers and H. Blatt, *Petrology: Igneous, Sedimentary, and Metamorphic*, 1982; A. Hall, *Igneous Petrology*, 2d ed., 1996; F. Press and R. Siever, *Earth*, 4th ed., 1986; F. Press and R. Siever, *Understanding Earth*, 3d ed., 2001.

Plutonium

A chemical element, Pu, atomic number 94. Plutonium is a reactive, silvery metal in the actinide series of element elements. The principal isotope of chemical interest is ^{239}Pu, with a half-life of 24,131 years. It is formed in nuclear reactors by the process shown in the following reaction. Plutonium-239 is fissionable,

$$^{238}\mathrm{U} + n \longrightarrow {}^{239}\mathrm{U} \xrightarrow[23.5\ \mathrm{min}]{\beta^-} {}^{239}\mathrm{N_P} \xrightarrow[2.33\ \mathrm{days}]{\beta^-} {}^{239}\mathrm{Pu}$$

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

but may also capture neutrons to form higher plutonium isotopes. *See* PERIODIC TABLE.

Plutonium-238, with a half-life of 87.7 years, is utilized in heat sources for space application, and has been used for heart pacemakers. Plutonium-239 is used as a nuclear fuel, in the production of radioactive isotopes for research, and as the fissile agent in nuclear weapons.

Plutonium exhibits a variety of valence states in solution and in the solid state. Plutonium metal is highly electropositive. Numerous alloys of plutonium have been prepared, and a large number of intermetallic compounds have been characterized.

Reaction of the metal with hydrogen yields two hydrides. The hydrides are formed at temperatures as low as 150°C (300°F). Their decomposition above 750°C (1400°F) may be used to prepare reactive plutonium powder. The most common oxide is PuO_2, which is formed by ignition of hydroxides, oxalates, peroxides, and nitrates of any oxidation state in air of 870–1200°C (1600–2200°F). A very important class of plutonium compounds are the halides and oxyhalides. Plutonium hexafluoride, the most volatile plutonium compound known, is a strong fluorinating agent. A number of other binary compounds are known. Among these are the carbides, silicides, sulfides, and selenides, which are of particular interest because of their refractory nature.

Because of its radiotoxicity, plutonium and its compounds require special handling techniques to prevent ingestion or inhalation. Therefore, all work with plutonium and its compounds must be carried out inside glove boxes. For work with plutonium and its alloys, which are attacked by moisture and by atmospheric gases, these boxes may be filled with helium or argon. *See* ACTINIDE ELEMENTS; NEPTUNIUM; NUCLEAR CHEMISTRY; TRANSURANIUM ELEMENTS; URANIUM.

Fritz Weigel

Bibliography. W. T. Carnall and G. R. Choppin (eds.), *Plutonium Chemistry*, 1983; R. Guillaumont et al. (eds.), *Update on the Chemical Thermodynamics of Uranium, Neptunium, Plutonium, Americium and Technetium*, 2003; S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; J. J. Katz, G. T. Seaborg, and L. R. Morss (eds.), *The Chemistry of the Actinide Elements*, 2 vols., 2d ed., 1986.

Plywood

A wood product in which thin sheets of wood are glued together, grains of adjacent sheets being at right angles to each other in the principal plane. Because of this cross-grained orientation, mechanical properties are less directional than those of natural lumber and more dimensionally stable. Originally developed to replace wide wood boards that had been sawed directly from large logs, plywood was manufactured from logs not suitable for other purposes. It has since found such widespread use that tree farms are now cultivated specifically to yield logs suitable for processing into sheets for plywood. The softwood plywood capacity in the United States in 1997 was 12.5 million cubic meters. Southern pine represents the largest segment of this market.

Grades. The American Plywood Association identifies several grades of product. Plywood is designated group 1 when made from northern-grown Douglas-fir, western larch, and such southern pines as loblolly and longleaf, or other woods noted for their strength. Plywoods in groups 2, 3, and 4 are made from woods of successively lower strengths. Consequently, group 1 plywood offers the greatest stiffness, group 4 the least.

Plywood with waterproof glue is designated exterior type; it is also used interiorly where moisture is present. Plywood with nonmoisture-resistant glue is designated interior type; it can withstand an occasional soaking but neither repeated soakings nor continuous high humidity.

Veneer grades A through D extend from a smooth surface to a surface with occasional knotholes and limited splits. If the outer face of the plywood is cut from only heartwood or sapwood, free from open defects, the plywood is assigned veneer grade N, indicating that it will take a natural finish. *See* STEM.

Most commonly used plywoods are $^1/_4$-in. (0.6-cm) sanded interior paneling or $^1/_2$-in. (1.3-cm) exterior grade plywood sheeting. Other standard thicknesses extend to 1 in. (2.5 cm) for interior types and to $1^1/_8$ in. (2.8 cm) for exterior types. The most common panel size is 4 × 8 ft (1.2 × 2.4 m); larger sizes are manufactured for such special purposes as boat hulls.

Manufacture. In modern plywood plants, as described in *Forest Industries* magazine, logs are debarked and graded, and those of suitable diameter are trimmed and delivered to a lathe where they are turned into plys of required thickness. Automatic handling machines then stack the sheets into sandwiches, glue being spread as the sheets are moved into position. If a thin hardwood veneer face is called for, it may be conveyed by vacuum rollers. The sandwiches are either stacked in a hot press or placed in a multiopening press for adhesive curing. *See* VENEER.

Finished plywood may be unsanded, sanded, or overlayed with several types of coatings for decorative and specialty uses.

If other than a flat board is required, the plywood is built up ply by ply over or inside a form and molded

to the finished shape under heat and pressure as required by the glue that is used.

Plywood in appropriate grades is used in many different applications, such as furniture, wall facings, shelving, containers, crates, fences, forms, subflooring, and roof decking. *See* WOOD PRODUCTS.

Frank H. Rockett; Roger M. Rowell

Bibliography. H. N. Spelter, D. B. McKeever, and I. A. Durbak, *Review of Wood-based Panel Sector in United States and Canada*, Gen. Tech. Rep. FPL-GTR-99, USDA Forest Service, Forest Products Laboratory, 1997.

Pneumatolysis

The alteration of rocks or crystallization of minerals by gases or supercritical fluids (generically termed magmatic fluids) derived from solidifying magma. At surface conditions, magmatic fluids sampled from active volcanoes contain steam with lesser amounts of carbon dioxide, sulfur dioxide, hydrogen sulfide, hydrogen chloride, and hydrogen fluoride, and trace amounts of many other volatile constituents. Magmatic fluids may contain relatively high concentrations of light and heavy elements, particularly metals, that do not crystallize readily in common rockforming silicates constituting most of the solidifying magma; thus, valuable rare minerals and ores are sometimes deposited in rocks subjected to pneumatolysis. Magmatic fluids are acidic and may react extensively with rocks in the volcanic edifice or with wall rocks surrounding intrusions. Penetration of magmatic fluids into adjacent rocks is greatly aided by faults, fractures, and cracks developed during intrusion and eruption or created by earlier geologic events. *See* MAGMA; METAMORPHISM; METASOMATISM; ORE AND MINERAL DEPOSITS; VOLCANO.

Pneumatolysis describes specific mechanisms of mineral deposition, hydrothermal alteration, or metasomatism in which magmatic fluids play an extremely significant role (see **illus.**). For example, lavas and ejecta at volcanoes may contain blocks (xenoliths) of wall rock that react with magmatic fluids to form pneumatolytic minerals such as vesuvianite (idocrase). Many silicic lavas and pyroclastic flows, particularly rhyolites, contain pneumatolytic tridymite, cristobalite, alkali feldspar, and rarely, topaz, garnet, rutile, fluorite, cassiterite, hematite, among others. Gases streaming from volcanic fumaroles deposit sublimates of sulfur, sulfates, chlorides, fluorides, and oxides of many metals. Wall rocks surrounding volcanic conduits may be thoroughly altered to mixtures of quartz, alunite, anhydrite, pyrite, diaspore, kaolin, as well as other minerals by acidic fluids degassed from magma. Rarely, gold, silver, base-metal sulfides, arsenides, and tellurides are deposited by the fluids, making valuable ores. Physical conditions during shallow pneumatolysis range from 100 to 1000°C (200 to 1850°F) at relatively low pressure ($\leq$1 kbar or $\leq$14,000 lb/in.^{2}), depending on eruption temperature, cooling history, depth to magma, and wall rock configuration. In such

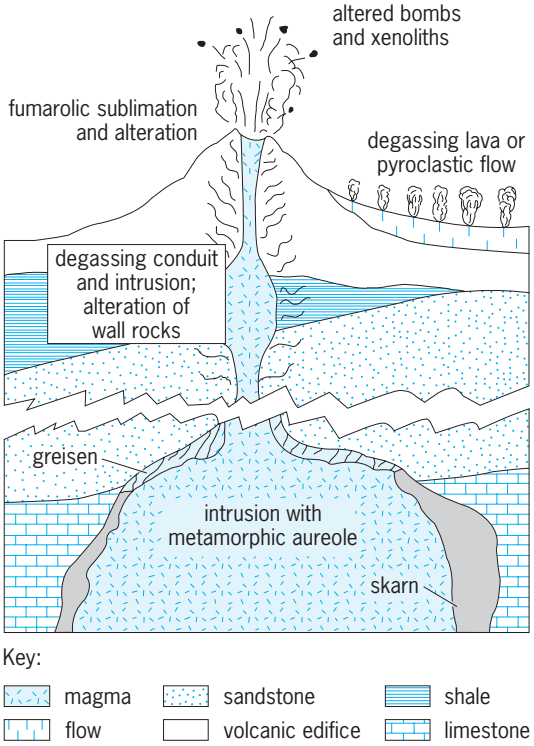

Various pneumatolytic environments. Shallow pneumatolysis occurs around volcanoes, flows, and shallow intrusions at depths of 0–3 km (0–2 mi), whereas deep pneumatolysis occurs around deep intrusions at depths of 3–15 km (2–10 mi).

environments, magmatic fluids are commonly subcritical, and near-surface ground waters and precipitation usually mix with degassing magmatic fluids. *See* LAVA; METASOMATISM; PYROCLASTIC ROCKS; RHYOLITE; SUBLIMATION; VESUVIANITE; XENOLITH.

When magma bodies intrude rocks at greater depths, lithostatic pressures are high (about 1–5 kbar or 14,000–70,000 lb/in.^{2}), and magmatic temperatures (650–1200°C or 1200–2200°F) are sustained for long periods of time ($\leq 10^6$ years). Released fluids are supercritical brines. Wall rocks undergo contact metamorphism, and a metamorphic aureole surrounds the intrusion. If highly reactive rocks such as limestones and dolomites are intruded, replacement deposits called skarns are formed that contain quartz, calcite, and a variety of iron, magnesium, and calcium silicates and oxides. Pneumatolytic minerals including ores of precious and base metals may be deposited from magmatic fluids during contact metamorphism. Silicic intrusions, especially granites, release substantial quantities of residual magmatic fluids that may alter wall rocks to form banded deposits and veinlike bodies called greisens. These deposits contain quartz, lithium-rich micas, and topaz, usually with accessory tourmaline, axinite, fluorite, apatite, cassiterite, wolframite, and others. An intrusion may undergo similar pneumatolytic alteration near contacts, displaying pegmatite-like mineralization. In these environments, pneumatolysis generally occurs at 400–600°C (750–1100°F). *See* CONTACT AUREOLE; GRANITE; GREISEN; PEGMATITE; SKARN.

Fraser Goff

Bibliography. M. R. Carroll and J. R. Holloway (eds.), Volatiles in Magmas, *Reviews in Mineralogy*, vol. 30, Mineralogical Society of America, 1994; J. S. Cline and R. J. Bodnar, Direct evolution of brine from a crystallizing silicic melt at the Questa, New Mexico, molybdenum deposit, *Econ. Geol.*, 89:1780–1802, 1994; W. D. Duffield et al., Origin of rhyolite-hosted tin mineralization: Evidence from the Taylor Creek rhyolite, New Mexico, *Econ. Geol.*, 85:392–398, 1990; F. Goff et al., Gold degassing and deposition at Galeras Volcano, Colombia, *GSA Today*, September 1994; D. M. Kerrick (ed.), *Contact Metamorphism, Reviews in Mineralogy*, vol. 26, Mineralogical Society of America, 1991; H. Sigurdsson et al. (eds.), *Encyclopedia of Volcanoes*, Academic Press, 2000.

Pneumonia

An acute or chronic inflammatory disease of the lungs. More specifically when inflammation is caused by an infectious agent, the condition is called pneumonia; and when the inflammatory process in the lung is not related to an infectious organism, it is called pneumonitis.

An estimated 45 million cases of infectious pneumonia occur annually in the United States, with up to 50,000 deaths directly attributable to it. Pneumonia is a common immediate cause of death in persons with a variety of underlying diseases. With the use of immunosuppressive and chemotherapeutic agents for treating transplant and cancer patients, pneumonia caused by infectious agents that usually do not cause infections in healthy persons (that is, pneumonia as an opportunistic infection) has become commonplace. Moreover, individuals with acquired immune deficiency syndrome (AIDS) usually die from an opportunistic infection, such as pneumocystis pneumonia or cytomegalovirus pneumonia. Concurrent with the variable and expanding etiology of pneumonia and the more frequent occurrence of opportunistic infections is the development of new antibiotics and other drugs used in the treatment of pneumonia. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); OPPORTUNISTIC INFECTIONS.

The incidence of noninfectious inflammatory conditions of the lung also seems to be increasing, or at least they are more frequently recognized. The number of drugs capable of inducing pneumonitis appears to be increasing as well. For example, cocaine derivatives have been shown to cause an inflammatory condition of the lung that may sometimes result in death.

Infectious pneumonia. Bacteria, as a group, are the most common cause of infectious pneumonia, although influenza virus has replaced *Streptococcus pneumoniae* (*Diplococcus pneumoniae*) as the most common single agent. Common bacterial agents causing pneumonia are *Streptococcus pneumoniae* (diplococci), other species of *Streptococcus*. *Staphylococcus* species, *Hemophilus influenzae*, *Klebsiella* species, *Pseudomonas aeruginosa*, *Legionella pneumophila*, *Mycobacterium tuberculosis*, and *Mycoplasma pneumoniae*. Some of these bacteria are normal inhabitants of the body and proliferate to cause disease only under certain conditions. Other bacteria are contaminants of food or water. *See* STEPTOCOCCUS PENUMONIAE (PNEUMOCOCCUS).

Most bacteria cause one of two main morphologic forms of inflammation in the lung. *Streptococcus pneumoniae* causes a type of pneumonia referred to as lobar pneumonia, in which an entire lobe of a lung or a large portion of a lobe, usually the lower lobe, becomes consolidated (firm, dense) and nonfunctional secondary to an influx of fluid and acute inflammatory cells (polymorphonuclear leukocytes) that represent a reaction to the bacteria. This type of pneumonia is uncommon today, usually occurring in people who have poor hygiene and are debilitated. If lobar pneumonia is treated adequately, the inflammatory process may entirely disappear, although in some instances it undergoes a process called organization, in which the inflammatory tissue changes into fibrous tissue. This results in a permanent obliteration of the normal lung architecture, usually rendering that portion of the lung nonfunctional.

The other morphologic form of pneumonia, which is caused by the majority of bacteria, is called bronchopneumonia. In this form there is patchy consolidation of lung tissue, usually around the small bronchi and bronchioles, again most frequent in the lower lobes (see **illus.**). Bronchopneumonia usually begins as an infection in the lumina of the small airways of the lower lobe and progresses to involve the surrounding alveoli, with an associated acute inflammatory response. This type of pneumonia may also undergo complete resolution if there is adequate treatment, although rarely it organizes.

Among the many viruses known to cause pneumonia are influenza A, influenza B, adenovirus, respiratory syncytial virus, cytomegalovirus, and herpes simplex virus. Viral pneumonia is usually a diffuse process throughout the lung and produces a different type of inflammatory reaction than is seen

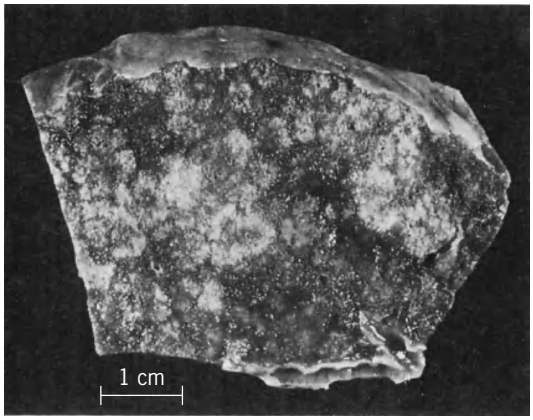

Representative region of lung tissue from the right lower lobe. The nodular whitish areas represent acute bronchopneumonia and appear white due to the influx of white blood cells, which represent an inflammatory reaction to the bacteria.

in bronchopneumonia or lobar pneumonia. Mycoplasma pneumonia, caused by *Mycoplasma pneumoniae*, is referred to as primary atypical pneumonia and causes an inflammatory reaction similar to that of viral pneumonia. Mycoplasma pneumonia usually occurs in families and groups and may become chronic and difficult to eradicate.

Fungal organisms relatively frequently causing pneumonia include *Coccidioides immitus*, *Cryptococcus neoformans*, and *Histoplasma capsulatum*. Pneumonia can be caused by a variety of other fungal organisms, especially in debilitated persons such as those with cancer or AIDS. *Mycobacterium tuberculosis*, the causative agent of pulmonary tuberculosis, produces an inflammatory reaction similar to fungal organisms. *See* MYCOBACTERIAL DISEASES; TUBERCULOSIS.

Legionella pneumonia. Legionella pneumonia, initially called Legionnaire's disease, is caused by bacteria of the genus *Legionella*. The condition is frequently referred to under the broader name of legionellosis. Although the illness was originally described as being caused by *Legionella pneumophila*, the organism that was isolated from patients who succumbed during the 1976 American Legion convention in Philadelphia, many other *Legionella* species are now recognized. Legionellosis is not uncommon, and two clinical syndromes are manifested: an acute pneumonia with a high mortality rate, and a self-limited flulike illness. Legionella pneumonia may also occur in compromised hosts as an opportunistic infection. The disease may be preceded by several days of fever, chills, malaise, and muscle aches, and it may be associated with nonpulmonary symptoms such as diarrhea and encephalopathy (brain dysfunction). The legionella organism is a coccobacillus that does not strain easily with the conventional (Gram) stain used to identify bacteria, and requires a special type of medium to grow in culture. *See* LEGIONNAIRES' DISEASE.

Pneumocystis pneumonia. Pneumocystis pneumonia is caused by *Pneumocystis carinii*, a putative protozoan parasite. Before the AIDS epidemic, this organism was a rare cause of pneumonia, occurring mainly in infants and young children and also in immunocompromised cancer patients. Between October 1980 and May 1981, five men in California were treated for pneumocystis pneumonia, and all five also had cytomegalovirus pneumonia. They were subsequently found to be immunodeficient, and pneumocystis pneumonia was quickly recognized as the primary cause of death in patients with AIDS. The pneumocystis organism, which is about the size of a red blood cell, causes a rapid consolidation of the lung. It is best identified with a silver stain in tissue sections or in special preparations made of fluid obtained from the lungs. The pneumonia can be effectively treated with erythromycin and pentamidine.

Hypersensitivity pneumonitis. Some persons are allergic to certain organic antigens, often the spores of various types of bacteria. When these antigens are inhaled into the lungs of susceptible persons, an inflammatory reaction occurs that can progress from acute to a more chronic disease. The most frequently recognized hypersensitivity pneumonitis is called farmer's lung: susceptible farmers or other persons dealing with such things as moldy hay inhale the spores of the bacteria *Thermoactinomyces vulgaris*. About 4–6 h later, they develop fever, cough, shortness of breath, and chest pain. X-ray shows a patchy infiltrate in their lungs. The disease can rapidly clear when the person is removed from the offending antigen or is treated with cortisone-type medicines.

Drug-induced pneumonitis. Many drugs can induce inflammation in the lungs. The mechanisms by which various drugs cause pneumonitis can either depend on the dose or be secondary to a hypersensitivity reaction. Alkylating agents such as busulfan, which is used in treating various types of solid tumors, are the drugs most frequently recognized to cause pneumonitis. The pneumonitis may occur weeks to months after the drug has been started and usually appears to depend on the dose. Gold sodium thiomalate, a drug used to treat rheumatoid arthritis, and amiodarone, a drug used in treating heartbeat irregularities, are agents that can cause a pneumonitis that may progress to fibrosis and that represents a hypersensitivity reaction. *See* CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS.

Cocaine and its derivatives have been associated with the development of a nonspecific type of pneumonitis and fibrosis, as well as pulmonary hemorrhage. The exact mechanism by which cocaine causes these types of change is uncertain but may be related to its effect on the contraction and relaxation of blood vessels.

Other types of pneumonitis and agents. Idiopathic interstitial pneumonitis, also referred to as idiopathic pulmonary fibrosis, is characterized by a chronic inflammation in the lung in association with the formation of fibrous tissue. The cause of most cases is uncertain, although some have been associated with diseases such as systemic lupus erythematosus. Bronchiolitis obliterans or organizing pneumonitis is characterized by inflammation of the bronchioles as well as lobar fibrosis. It usually is associated with cough, fever, and chest pain. Its cause is usually unknown, and most cases are reversible when treated with cortisone-type drugs.

Oxygen, when delivered at concentrations greater than 40% for more than 24 h, can cause inflammation in the lung. This usually begins as acute damage to the alveoli and can progress to pulmonary fibrosis. Radiation, used in the treatment of thoracic tumors, can also cause inflammation of the lung with subsequent fibrosis. Aspiration of blood or gastric contents can produce a localized region of severe inflammation and cellular death in the lung. The aspiration of an oily substance, such as oily nose drops, can result in the accumulation of lipid material in the lung, referred to as lipoid pneumonitis.

Signs, symptoms, and diagnosis. The signs and symptoms of pneumonia and pneumonitis are usually nonspecific, consisting of fever, chills, shortness of breath, and chest pain. Fever and chills are

more frequently associated with infectious pneumonias but may also be seen in pneumonitis. The physical examination of a person with pneumonia or pneumonitis may reveal abnormal lung sounds indicative of regions of consolidation of lung tissue. A chest x-ray also shows the consolidation, which appears as an area of increased opacity (white area). Cultures of sputum or bronchial secretions may identify an infectious organism capable of causing the pneumonia. In AIDS patients, fluid is frequently instilled into the pulmonary bronchial tree and then rapidly removed; examination of special preparations of this fluid may show such microorganisms as *Pneumocystis carinii*. In some cases of pneumonia or pneumonitis in which no definite cause can be determined from routine studies, the chest cavity must be opened surgically and the lung biopsied.

Treatment. The treatment of pneumonia and pneumonitis depends on the cause. Bacterial pneumonias are treated with antimicrobial agents such as penicillin. If the organisms can be cultured, the sensitivity of the organism to a specific antibiotic can be determined. Viral pneumonia is difficult to treat, as most drugs only help control the symptoms. The treatment of pneumonitis depends on identifying its cause; many cases are treated with cortisone-type medicines. Samuel P. Hammar

Bibliography. D. H. Dail and S. P. Hammar (eds.), *Pulmonary Pathology*, 2d ed., 1993; R. B. Fick and H. Y. Reynolds, The changing spectrum of pneumonia: News media creation or clinical reality?, *Amer. J. Med.*, 74:1–8, 1983; A. P. Fishman (ed.), *Update: Pulmonary Disorders and Diseases*, 1983.

Poales

An order of flowering plants (angiosperms) that includes the bromeliads, grasses, restios, rushes, and sedges. It comprises approximately 18 families and more than 18,000 species, over half of which are grasses (Poaceae). These families in general include plants without showy flowers (except for Bromeliaceae, Xyridaceae, and Rapateaceae) and with a grasslike form. These plants dominate vast areas of the Earth's surface and include some important grain crops, such as corn (*Zea mays*), rice (*Oryza sativa*), and wheat (*Triticum aestivum*), as well as rushes, sedges, grasses used for thatch, and bamboos used for construction. *See* DUNE VEGETATION; GRAIN CROPS; GRASS CROPS; PLANT EVOLUTION. Mark W. Chase; Michael F. Fay

Podicipediformes

A small order of aquatic birds that contains only a single living family, the Podicipedidae (grebes), with 20 species found throughout the world. They breed on freshwater, but some species winter along the coasts. Grebes have a number of convergent similarities with loons, but are not related to them or to the fossil hesperornithiforms. The close affinities of the grebes are still unresolved. *See* GAVIIFORMES; HESPERORNITHIFORMES.

Grebes are small to medium-sized birds that are specialized for aquatic life and can swim and dive skillfully. Their legs are placed far posteriorly with lobed—not webbed—toes that are used to propel the birds through the water, but their terrestrial locomotion is awkward. Their wings are small, and they rarely fly, preferring to dive to escape enemies; yet they migrate over long distances and have colonized many remote islands. *Rollandia microptera*, living on Lake Titicaca in the Andes, is flightless, as is *Podiceps taczanowskii*, living on the high-altitude Lake Junin in Peru; the latter species is endangered. Grebe bodies are compact with dense waterproof plumage. The neck is medium to long, the head is large, and the bill is short to medium and pointed. Grebes feed on fish and other aquatic animals that they pursue underwater. The sexes are alike in their plumage, which is dull except on the head, where it is bright and patterned and used in courtship displays. Grebes are monogamous, with an elaborate courtship and strong pair bond. The nest consists of a floating mat of vegetation and can be solitary or part of a large colony. The clutch of three to nine eggs is incubated by both parents. The downy young leave the nest after hatching and are cared for by both parents, which often carry the small young on their backs when swimming.

Grebes are found throughout the world except on some remote oceanic islands. They have a poor fossil record in spite of being aquatic birds. The earliest fossil is early Miocene, with only a handful of younger forms. Some researchers have suggested a gruiform affinity for grebes, but that remains speculative. Note that some taxonomists have enlarged the Ciconiiformes to include the podicipediform grebes, as in the Sibley-Ahlquist taxonomy a radical new approach to taxonomy, based on DNA hybridization studies. *See* CICONIIFORMES.

The western grebe (*Aechmophorus occidentalis*; see **illustration**) and Clark's grebe (*A. clarkii*) constitute an interesting pair of sibling species that are similar, not only in their nearly identical plumage coloration but also in their ecology and distribution. The giant pied-billed grebe (*Podilymbus gigas*),

***Aechmophorus occidentalis*, the western grebe. (© *California Academy of Sciences, photo by Dr. Lloyd Glenn Ingles*)**

found only on Guatemala's Lake Atitlan, probably became extinct during the past decade as a result of interbreeding with the closely related pied-billed grebe (*P. podiceps*). *See* AVES. Walter J. Bock

Bibliography. J. Fjeldä, *Grebes: Podicipedidae*, Oxford University Press, 2004.

Podocopa

One of two Recent subclasses of Ostracoda. The two (Podocopa and Myodocopa) may share many characteristics but differ significantly in others. Shared characteristics include a carapace (shell) with a straight-to-concave ventral margin and somewhat convex dorsal margin; adductor-muscle-scar patterns varying from numerous to few; the absence of an anterior notch in the bivalve shell; and the absence of lateral eyes, heart, and frontal organ (Bellonci organ). Podocopid ostracods are small, rarely exceeding 0.25 in. (7 mm). The carapace valves are unequal, weakly or strongly calcified, and may be heavily ornamented. The hinge varies from simple to complex. Six or seven pairs of appendages are present; the seventh is adapted for walking, for cleaning, or as a clasping leg in males. It is never annulated and wormlike, as in the Myodocopa. The caudal furca, when present, is lamelliform, flagellate, or footlike and ventral to the anus. In one podocopid genus the posterior part of the body is pronounced and has a pair of setae-bearing lobes, structures that have been interpreted by some researchers as representing a telson and uropods, respectively.

The exopod of the antenna, which consists of only one or two segments, is never adapted for swimming; it may be modified as a spinning structure or reduced to a simple bristle or scale. The maxillule is usually provided with a branchial epipod or plate. The first postoral appendage is lamelliform or sticklike.

The sexes are separate, and the male genital organs are large, highly specialized, and usually paired. Females of some fresh-water species are parthenogenetic and produce eggs capable of withstanding years of desiccation. Many podocopids brood their eggs; eight—or less frequently, seven—juvenile instars make up the developmental sequence. Podocopid ostracods are found in freshwater, marine, and brackish-water environments, and a few are terrestrial. The majority are burrowing or crawling inhabitants of the benthos and are filter or detritus feeders or herbivores. Some are capable of swimming, but none appear to ever lead a planktonic existence.

The subclass is represented by the two extant orders, Platycopida and Podocopida. Although morphologic differences between these two taxa are sufficiently distinct to have prompted some researchers to recommend taxonomic hierarchical separation, evidence supports the hypothesis of common ancestry. Despite differences in filter-feeding adaptations, some older representatives exhibit converging morphologic characteristics, and taxa from extinct lineages provide possible links between the modern representatives of the two orders. *See* CRUSTACEA; OSTRACODA; PLATYCOPIDA; PODOCOPIDA. Patsy A. McLaughlin

Bibliography. T. E. Bowman, The case of the nonubiquitous telson and the fraudulent furca, *Crustaceana*, 21, 1971; A. C. Cohen, Ostracoda, in S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, vol. 2, 1982; R. F. Maddocks, Evolution within the Crustacea, pt. 4: Ostracoda, in L. G. Abele (ed.), *The Biology of Crustacea*, vol. 1: *Systematics, the Fossil Record, and Biogeography*, 1982; R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, pt. Q: *Arthropoda*, 1961; H. K. Schminke, The ubiquitous telson and deceptive furca, *Crustaceana*, vol. 30, 1976.

Podocopida

An order of the subclass Podocopa, class Ostracoda, that consists of the extant suborder Podocopina and the extinct Metacopina. There is agreement on the assignment of the Cypridoidea, Bairdioidea, Cytheroidea, and Darwinuloidea to the Podocopida; however, some researchers also include the Sigillioidea, and others consider the Terrestrichtheroidea a podocopid taxon of equivalent rank. In all podocopids the two valves fit firmly, hermetically sealing the animal inside when closed. Although variously sculptured, none of the podocopids have strong vertical lobation like that found in the Paleocopa.

Appendages include antennules, antennae, mandibles, maxillules, and three pairs of thoracic legs. Antennules and antennae are both used in locomotion; swimming forms are provided with long, feathered setae, whereas crawling forms have only short, stout setae. Mandibles are strongly constructed and, with rare exceptions, used for mastication. The maxilla has setiferous endites and bears a large branchial plate to circulate water alongside the body. The furca is variously developed but never lamelliform. If eyes are present, the two lateral eyes and the median eye are joined into one central structure. No heart is developed.

Principal superfamilies. Differences in appendage specializations, male copulatory structures (Zenker's organs), furcae, and shell shape make a brief review of the four principal superfamilies appropriate.

Cypridoidea. These ecologically versatile ostracods are mostly fresh-water, but a few are marine; they are adapted to extremes in temperature, oxygen, pH, and postoral salinity. The legs are all different (**Fig. 1***a*). The first legs are accessory feeding structures, the second legs are for walking, and the third are modified for cleaning the interior of the carapace. Some species are parthenogenetic; if males are present, the Zenker's organs are large and well developed (Fig. 1*b*). The furca is typically an elongate ramus and two terminal claws (Fig. 1*c*), but in *Cypridopsis* the end is flagelliform. Basically, the adductor muscle scars consist of one elongate scar above and five others clustered below. The hinge is relatively

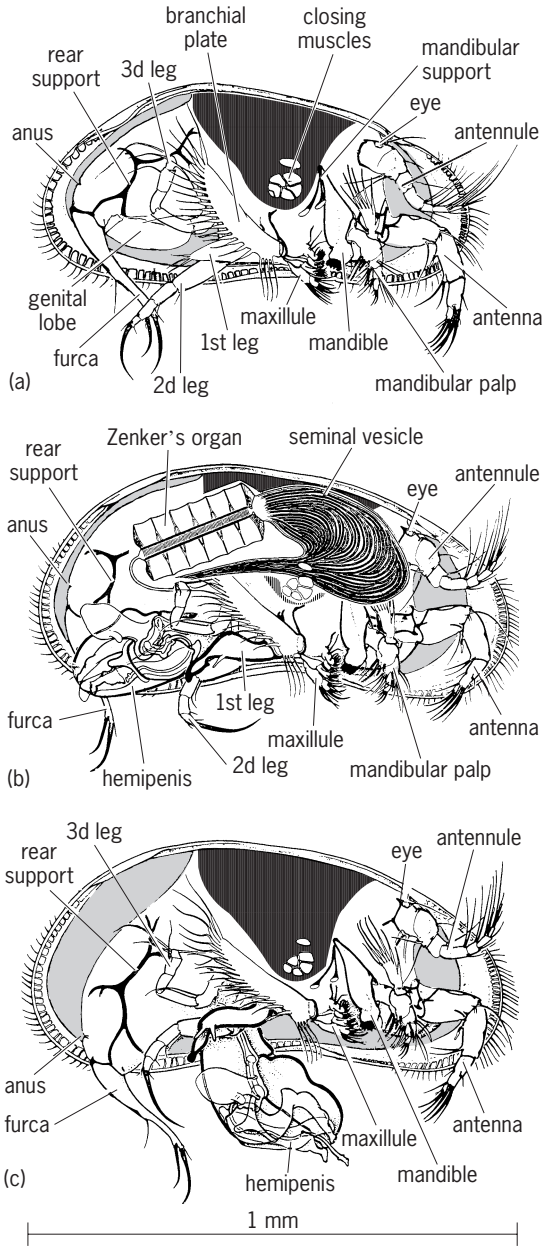

Fig. 1. ***Candona suburbana*, an ostracode of the Cypridoidea. Animals with right valves removed. (*a*) Female; note dimorphic differences in carapace outline and first leg. (*b*) Male with hemipenis rotated, folded, and retracted; part of the body wall is removed to show right Zenker's organ and seminal vesicle, the latter distended with spermatozoa. (*c*) Male with hemipenis erect. (*After D. J. McGregor and R. V. Kesling, Contributions from the Museum of Paleontology of the University of Michigan, 1969*)**

weak and simple, and teeth rarely occur.

Carapace shape varies from ovate to elongate and somewhat subtriangular; the carapace surface is usually smooth and weakly calcified, but may be strongly calcified and sculptured.

The first leg of the male has a pincerlike clasper. Ontogeny is brief, and some fresh-water species reach maturity in a month. In many species the eggs are encased and can withstand drying over periods of years, hatching only when returned to water; hence, many fresh-water species have spread over vast areas.

Bairdioidea. The Bairdioidea are exclusively marine ostracods with the three postoral appendages adapted for walking. The furca is rod-shaped, with several distal bristles. Males lack Zenker's organs. Muscle scars are irregular and loosely clustered. Hingement is relatively simple, although small teeth occur. The round-backed carapace is acuminate at the ends in many genera, and this lemon-seed form is known as the *Bairdia* shape. These ostracods are not as numerous as the Cypridoidea or Cytheroidea, but they are found throughout the world.

Cytheroidea. The superfamily contains the most genera and species of any ostracode group; in some faunas the cytheraceans are twice as numerous as all other ostracodes combined. Only a few live in fresh water. Most are crawling and digging forms on the ocean floor. Postoral legs are all alike, adapted for walking (**Fig. 2**). The furca is very small, without claws, and Zenker's organ is absent. The four closing-muscle scars are horizontal, arranged in a vertical row, although one or more scars may be interrupted. The hinge is strong, with complex interlocking structures in many genera. The carapace tends to be nearly straight-backed because of the long hinge; many are distinctly ornamented and colored. These are the only ostracodes with spinning glands, which are located in the antennae. The commensal ostracodes *Entocythere* and *Sphaeromicola* belong to this superfamily.

Darwinuloidea. This small superfamily, containing only two genera, is primarily fresh-water but widespread. The ostracodes crawl on bottom sediments. Males are very rare; most darwinuloids are parthenogenetic. The eggs and young animals are protected in a brood space behind the body (**Fig. 3**). The two postoral pairs of legs are alike and are used in walking. The rear end of the body is curiously tapered; the furca is rudimentary. The several muscle

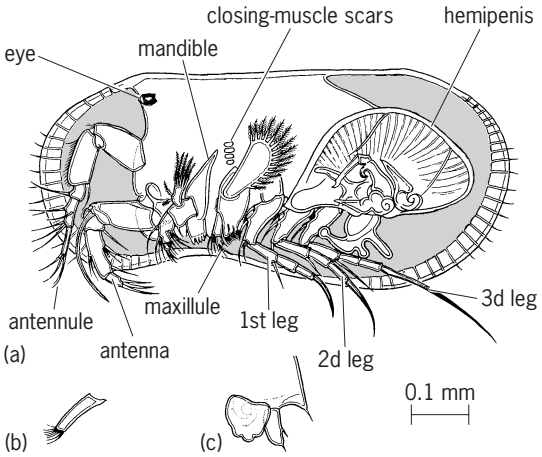

Fig. 2. ***Limnocythere sanctipatricii*, a podocopid ostracode of the Cytheroidea. (*a*) Male with left valve removed; the paired hemipenes occupy much of the space at the rear of the carapace. (*b*) Male left "brush-shaped organ," an unusual structure found only in adult males between the postoral legs, presumed to assist in copulation. (*c*) Female furca and genitalia. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. Q, 1961, and D. L. McGregor and R. V. Kesling, Contributions from the Museum of Paleontology of the University of Michigan, 1969*)**

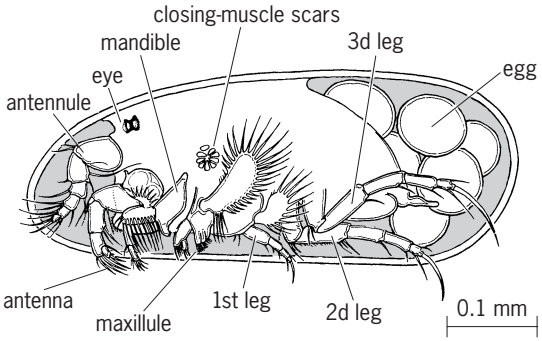

Fig. 3. ***Darwinula stevensoni*, a podocopid ostracode of the Darwinuloidea. In this parthenogenetic species, eggs and young are protected in a brood space behind the body. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. Q, 1961*)**

scars form a vertically elongate rosette. The hinge is relatively simple. The carapace is smooth and notably elongate.

Other groups. The Podocopida (Ordovician to Recent) have many fossil representatives, both marine and fresh-water. The superfamily Cytheroidea, in particular, is abundant in the Mesozoic and Cenozoic, and has many divisions. Nearly half of all fossil species described belong to this taxon. The Sigillioidea includes two extant species of *Saipanetta*; however, there is no universal agreement regarding the placement of the genus, or of the family in the Podocopida. Both are regarded by some as correctly being retained in the suborder Metacopina.

The Terrestrictheroidea is represented by a single genus (*Terrestricythere*) of terrestrial ostracods from the Kuril Islands of the northwestern Pacific Ocean. Other terrestrial ostracodes are assigned to the superfamily Cypridoidea, and it is the opinion of some specialists that the Terrestricytheroidea are no more than acutely specialized cypridoids.

Evolution. Within the Podocopida there appears to be two principal evolutionary pathways, one including the Cypridoidea and Darwinuloidea (also Sigillioidea), and the other the Bairdioidea and Cytheroidea. It is believed that the oldest podocopids may belong to the Bairdioidea and that those may approach the ancestral podocopid stem. The Cytheroidea are then assumed to have been descended from the Bairdioidea. The Darwinuloidea (and Sigilloidea) have retained many primitive characters, although that may be the result of the fact that most living species reproduce parthenogenetically. Although the Cypridoidea cannot be considered primitive, they do share several significant characteristics, such as appendage structure and genitalia, with the former taxa. *See* CRUSTACEA; PODOCOPA.

Patsy A. McLaughlin

Bibliography. A. C. Cohen, Ostracoda, in S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, vol. 2, 1982; P. D. Deckker, Terrestrial ostracods in Australia, in J. K. Lowry (ed.), *Papers from the Conference on the Biology and Evolution of Crustacea*, Austral. Mus. Mem. 18, 1983; R. F. Maddocks, Evolution within the Crustacea, pt. 4: Ostracoda, in L. G. Abele (ed.), *The Biology of Crustacea*, vol. 1: *Systematics, the Fossil Record, and Biogeography*, 1982; K. G. McKenzie, Saipanetta and the classification of podocopid Ostracoda: A reply to Schornikov and Gramm (1974), *Crustaceana*, vol. 29, 1975; R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, pt. Q, 1961; G. O. Sars, *Ostracoda: An Account of the Crustacea of Norway*, Bergen Museum, vol. 9, 1922–1928.

Podostemales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Rosidae of the class Magnoliopsida (dicotyledons). The order consists of only the family Podostemaceae, with about 200 species, the greatest number occurring in tropical America. They are submerged aquatics with modified, branching, often thalluslike shoots; small, perfect flowers with a much reduced perianth; a superior ovary with united carpels and numerous small seeds without endosperm. *See* MAGNOLIOPSIDA; PLANT KINGDOM; ROSIDAE.

Arthur Cronquist

Poecilosclerida

An order of sponges of the class Demospongiae in which the skeleton includes two or more types of megascleres, each localized in a particular part of the sponge colony. Frequently one type of megasclere is restricted to the dermis and another type occurs in the interior of the sponge (**Fig. 1**). Sometimes one category is embedded in spongin fibers while a second category, usually spinose, protrudes from the fibers at right angles. Spongin is always present but varies in amount from species to species. Microscleres are usually present; often several types occur in one species. A wide variety of microsclere categories is found in the order but asters are never present.

In shape, poecilosclerid sponges are encrusting, massive, lobate, or branching. Deep-sea species often have bizarre shapes (**Fig. 2**). Sponges of this order are

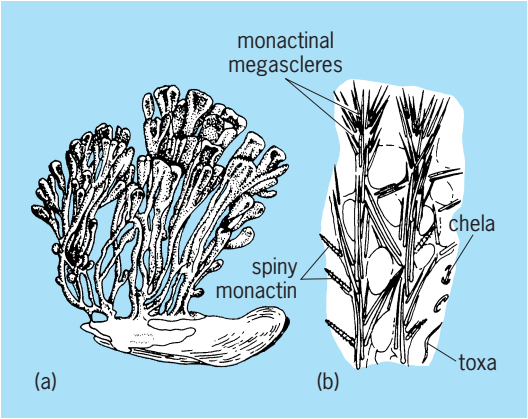

Fig. 1. ***Microciona prolifera*. (*a*) Whole organism. (*b*) Spiculation.**

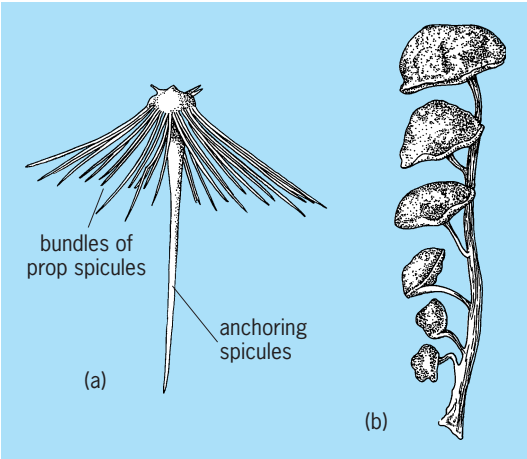

Fig. 2. Deep-sea poecilosclerids. (*a*) *Cladorhiza longipinna*, found at depth of 1800 ft (5500 m). (*b*) *Esperiopsis challengeri*, found at depth of 500 ft (1500 m).

found in all seas and range from tidal waters to depths of at least 1800 ft (5500 m). Undoubted poecilosclerid spicules are known from Tertiary deposits. *See* DEMOSPONGIAE. Willard D. Hartman

Poecilostomatoida

One of two major orders of parasitic Copepoda that were previously included in the Cyclopoida. The classification of parasitic copepods has been established on the basis of the structure of the mouth. In poecilostomatoids the mouth is represented by a transverse slit, partially covered by the overhanging labrum, which resembles an upper lip. Although there is variability in the form of the mandible among poecilostomatoids, it can be generalized as being falcate (see **illus.**). Body segmentation is typically podoplean, having prosome–urosome articulation between the fifth and sixth thoracic somites; however, this segmentation is often lost with the molt to adulthood. The antennules frequently are reduced in size and the antennae modified to terminate in small hooks or claws that are used in attachment to host organisms.

Most poecilostomatoid copepods (see illus.) are ectoparasites of marine fishes or invertebrates, usually attaching to the external surface of the host or in the branchial cavity on the walls or gill surfaces. Representatives of one family, however, have successfully made the transition to fresh-water habitats, and a second family has evolved an endoparasitic mode of life. The latter inhabit the mucous ducts, body wall, or viscera of the host teleost fish. The body structure of the parasites is reduced in females to a vermiform, saccular, or spherical mass usually lacking observable appendages; body segmentation in males is obsolete, but appendages are retained. Among ectoparasitic poecilostomatoids, external body segmentation is obsolete in those species that attach permanently to their hosts. More mobile species retain much of their body segmentation and attach to their hosts by large, subchelate antennae or by antennae provided with rows of small spinules or denticles. As a group, poecilostomatoids are usually less deleterious to their hosts than other parasitic copepods, since most species are small, tend to be less mobile, and thus are less likely to cause extensive tissue damage.

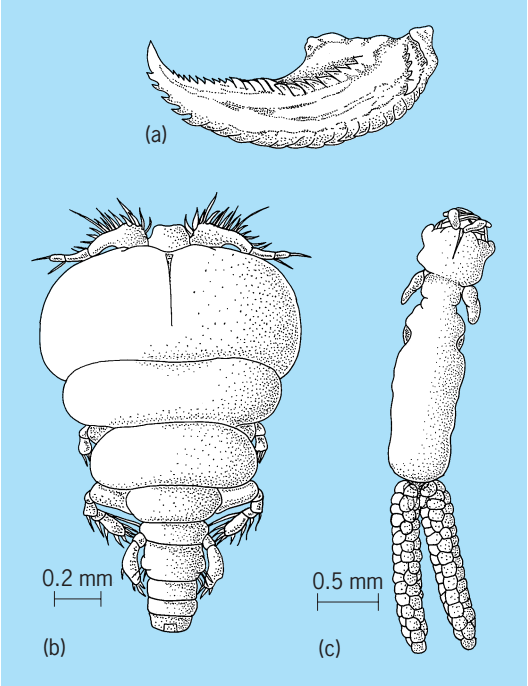

Representatives of poecilostomatoid copepods. (*a*) Poecilostomatoid mandible (*after L. Margolis and Z. Kabata, Guide to the Parasites of Fishes of Canada. pt. II: Crustacea, Can. Spec. Publ. Fisher. Aquat. Sci. 101, 1988*). (*b*) *Bomolchus concinnus* (*after W. Vervoort, Caribbean Bomolochidae* (*Copepoda: Cyclopoida*), *Studies on the Fauna of Curaçao and Other Caribbean Islands, vol. 28, 1969*). (*c*) *Heterochondria atypica* (*after J.-S. Ho, Four new parasitic copepods of the family Chondracanthidae from California inshore fishes, Proc. Biol. Soc. Wash., vol. 85, 1972*).

Poecilostomatoids are thought to have evolved rather recently from the Cyclopoida lineage. Perhaps the strongest argument for that hypothesis is found in the similarity of nauplii in the two orders. A possible link has been suggested through the genus *Hemicyclops*, which has a broad labrum that covers only the anterior edge of the mouth. Subsequent evolutionary steps might then have included hypertrophy and paedomorphosis, which led to the reduction in appendages and adult body segmentation. *See* COPEPODA; CRUSTACEA; CYCLOPOIDA. Patsy A. McLaughlin

Bibliography. Z. Kabata, *Parasitic Copepoda of British Fishes*, Ray Society, London, vol. 152, 1979; B. M. Marcotte, Evolution within the Crustacea, pt. 2: Copepoda, in L. G. Abele (ed.), *The Biology of Crustacea*, vol. 1: *Systematics, the Fossil Record, and Biogeography*, 1982; L. Margolis and Z. Kabata, *Guide to the Parasites of Fishes of Canada*, pt. II: *Crustacea*, Can. Spec. Publ. Fisher. Aquat. Sci. 101, 1988; W. Vervoort, *Caribbean Bomolochidae (Copepoda: Cyclopoida)*, Studies on the Fauna of Curacao and Other Caribbean Islands, vol. 28, 1969.

Pogonophora (Siboglinidae)

A family of sedentary marine worms (known as beard worms) that live in all the world's oceans, generally at depths between 330 and 13,200 ft (100 and 4000 m), shallower at higher latitudes and deeper in trenches. The first specimens of the Pogonophora, previously considered to be a phylum, but now known more accurately as Siboglinidae (a family of polychaete annelids), were dredged up late in the nineteenth century, but first investigated in the 1950s by the Russian zoologist A. V. Ivanov. Siboglinids construct a chitinous tube and have no mouth, gut, or anus in their postembryonic anatomy. Instead of actively feeding, they derive their nutrition from symbiotic bacteria hosted in specialized tissues or via absorption. With new technologies in deep-sea exploration, such as crewed and remotely operated submersibles, many new species of Siboglinidae have been discovered in unexpected habitats over the past three decades, notably around hydrothermal vents and cold-water hydrocarbon seeps, and on decaying whale bones. *See* DEEP-SEA FAUNA; HYDROTHERMAL VENT; MARINE ECOLOGY.

Anatomy. The body is generally divided into four regions (see **illus.**). The anterior region (cephalic lobe) bears from one to thousands of tentacles or branchial filaments and is followed by a short collar region known as the forepart or vestimentum. The largest section is usually an extended trunk region, which can have regional specializations such as papillae, cilia, and chaetae. Except in the whalebone-inhabiting *Osedax*, this region hosts symbiotic bacteria. In the frenulates, vestimentiferans, and moniliferans, the terminal region (opisthosoma) has the internal septa and external chaetae typical of polychaete annelids. No opisthosoma has been found in *Osedax*; instead this worm forms an extensive root system that reaches deeply into the whale bones and hosts the female gonads and symbiotic bacteria.

Physiology. Frenulates and moniliferans can absorb dissolved organic matter and other nutrients through their body wall, but symbiotic bacteria supplement their nutrition. Vestimentiferans seem to depend completely on their symbionts for nutrition. The bacteria in frenulates, moniliferans, and vestimentiferans are usually chemoautotrophic sulfur-oxidizers, but one frenulate species is known to host methanotrophic bacteria. The symbionts of *Osedax* are heterotrophic: they are capable of extracting and degrading complex carbohydrates from the whale bones that this genus inhabits. *See* FEEDING MECHANISMS (INVERTEBRATE).

Vestimentiferans have a unique type of hemoglobin which can bind sulfide and oxygen simultaneously. Sulfide and oxygen are absorbed through the bright red branchial filaments and transported to the symbionts, where they are released. The sulfide-binding capacity of vestimentiferan hemoglobin has several advantages: it prevents toxic effects of sulfide in the body as well as its spontaneous reaction with oxygen, while channeling the energy of the oxidation to the symbionts. *See* HEMOGLOBIN; VESTIMENTIFERA.

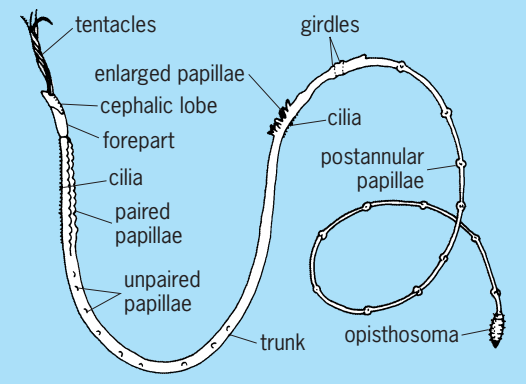

Diagram of a frenulate siboglinid; length has been greatly reduced. (*After J. D. George and E. C. Southward, A comparative study of the setae of Pogonophora and polychaetous annelida, J. Mar. Biol. Ass. U.K., 53(2):403–424, 1973*)

Reproduction and development. The sexes are separate in most siboglinids; only one hermaphroditic species has been reported. *Osedax* shows an extreme degree of sexual dimorphism in which only the females reach full size, while the males are essentially sexually mature larvae that live in the female tube, where they fertilize the eggs. Larval development has been studied in very few siboglinid species. Some frenulates brood their larvae in their tubes; in others, larvae have never been observed. Vestimentiferan larvae are planktonic (free-floating) and have a transient digestive tract just before settlement. They may spend several weeks in the water column, subsisting on yolk supplied with the egg, before metamorphosing into a juvenile. The larvae are symbiont-free, and the worms must acquire new symbionts from the environment with each generation. *See* SEXUAL DIMORPHISM.

Taxonomy. The phylogenetic position of the Siboglinidae has been the subject of much debate. The earliest specimens examined were incomplete. With the opisthosoma missing, the most obvious clue for their relationship with the segmented annelids was lacking. Due to the absence of the mouth and gut, some early researchers erroneously considered the nerve cord to be dorsal, suggesting deuterostome affinities. The frenulates and vestimentiferans were once considered separate phyla and were named Perviata and Obturata, respectively. However, recent molecular studies, supported by similarities in embryonic and larval development, the segmentation of the opisthosoma, and the ultrastructure of chaetae formation, clearly indicate that siboglinids are polychaete annelids. *See* ANNELIDA.

Anja Schulze; Edward B. Cutler

Bibliography. A. D. Nussbaumer et al., Horizontal endosymbiont transmission in hydrothermal vent tubeworms, *Nature*, 441:345–348, 2006; G. W. Rouse et al., *Osedax:* Bone-eating marine worms with dwarf males, *Science*, 305:668–671, 2004; A. Schulze and K. M. Halanych, Siboglinid evolution shaped by habitat preference and sulfide tolerance,

Hydrobiologia, 496:199–205, 2003; E. C. Southward et al., Pogonophora (Annelida): Form and function, *Hydrobiologia*, 535/536:227–251, 2005.

Poison

A substance which by chemical action and at low dosage can kill or injure living organisms. Broadly defined, poisons include chemicals toxic for any living form: microbes, plants, or animals. For example, antibiotics such as penicillin, although nontoxic for mammals, are poisons for bacteria. In common usage the word is limited to substances toxic for humans and mammals, particularly where toxicity is a substance's major property of medical interest. Because of their diversity in origin, chemistry, and toxic action, poisons defy any simple classification. Almost all chemicals with recognized physiological effects are toxic at sufficient dosage. The same compound may be considered a drug or a poison, depending on dosage, effect, or intended use.

Origin and chemistry. Many poisons are of natural origin. Some bacteria secrete toxic proteins (for example, botulinus, diphtheria, and tetanus toxins) that are among the most poisonous compounds known. Lower plants notorious for poisonous properties are ergot (*Claviceps purpurea*) and a variety of toxic mushrooms. Ergot, a fungal parasite of rye, has been the source of numerous epidemics of poisoning from the use of contaminated rye flour. The fungus contains many different alkaloids, some of which are also useful drugs. Among the best-known toxic mushrooms are the fly agaric (*Amanita muscaria*) containing muscarine, and the destroying angel (*A. phalloides*) whose toxic agents are phalloidin and aminitin. *See* ERGOT AND ERGOTISM; MUSHROOM; TOXIN.

Higher plants, which constitute the major natural source of drugs, contain a great variety of poisonous substances. Many of the plant alkaloids double as drugs or poisons, depending on dose. These include curare, quinine, atropine, mescaline, morphine, nicotine, cocaine, picrotoxin, strychnine, lysergic acid, and many others. Some of the alkaloids were used in classical antiquity (coniine was the toxic agent of the extract of spotted hemlock, *Conium maculatum*, drunk by Socrates). Some are of prehistoric antiquity (quinine and curare were used by South American Indians before the advent of Europeans), and some date from the earliest records of humans (the opium poppy is believed to have been cultivated in the Stone Age). *See* ATROPINE; COCAINE; MORPHINE ALKALOIDS; QUININE.

Poisons of animal origin (venoms) are similarly diverse. Toxic marine animals alone include examples of every phylum from Protozoa (dinoflagellates) to Chordata (a number of fishes). Insects and snakes represent the best-known venomous land animals, but on land, too, all phyla include poison-producing species. Among mammalian examples are certain shrews with poison-producing salivary glands. *See* POISON GLAND.

Poisons of nonliving origin vary in chemical complexity from the toxic elements, for example, the heavy metals, to complex synthetic organic molecules. Most of the heavy metals (gold, silver, mercury, arsenic, and lead) are poisons of high potency in the form of their soluble salts. Strong acids or bases are toxic largely because of corrosive local tissue injury; ingestion, for example, sometimes results in fatal perforation of the gastrointestinal tract. Many other elemental substances are toxic at low concentrations: selenium, beryllium, cadmium, manganese, phosphorus, and zinc. In general, it may be anticipated that any chemically reactive element is likely to be toxic to humans unless it represents one of the bulk elements in the body, in the ionic form in which it exists in the body. Na^+, K^+, Ca^{2+}, and Cl^-, for example, are relatively nontoxic.

The chemically reactive gases hydrogen sulfide, hydrocyanic acid, chlorine, bromine, and ammonia are also toxic, even at low concentration, both because of their corrosiveness and because of more subtle chemical interaction with enzymes or other cell constituents. An example is the toxic action of cyanide ion, probably explained by its inactivation of cytochrome oxidase.

Many organic substances of synthetic origin are highly toxic and represent a major source of industrial hazard. Most organic solvents are more or less toxic on ingestion or inhalation. Ethanol is a relatively innocuous exception, in part because available enzymatic reactions channel it into major normal metabolic pathways along which it can be oxidized to products familiar to mammalian metabolism, such as acetate. Many other alcohols, such as methanol, are much more toxic. Many solvents (for example, carbon tetrachloride, tetrachloroethane, dioxane, and ethylene glycol) produce severe chemical injury to the liver and other viscera, sometimes from rather low dosage. With certain poisons of this group, a high degree of variation in individual susceptibility exists.

Chemical correlations. Since poisons represent all chemical classes from the elements to complex alkaloids and large proteins, general chemical constitution has no defining value for toxicity. However, in limited instances chemical features correlate with toxic action. Many of the chlorinated hydrocarbons, for example (carbon tetrachloride), have similar toxicity for liver, heart, and kidneys. A number of alkyl phosphates (diisopropylfluorophosphate, tetraethylpyrophosphate, and related compounds) are very potent inhibitors of the enzyme acetylcholinesterase and produce a consistent set of physiological changes arising from the inactivation of this enzyme. A number of tertiary amino esters of aromatic acids (such as procaine and cocaine) are local anesthetics and also share marked cardiac and nervous system toxicity. The digitalis glycosides, all represented by steroids condensed with sugars and containing an additional lactone ring, share many pharmacological properties, notably characteristic effects on cardiac function which make these drugs cardiotoxins at low concentration.

Physiological actions. The action of poisons is generally described by the physiological or biochemical changes which they produce. For most poisons, a descriptive account can be given which indicates what organic system (for example, heart, kidney, liver, brain, and bone marrow) appears to be most critically involved and contributes most to seriously disordered body function or death. In many cases, however, organ effects are multiple, or functional derangements so generalized that a cause of death cannot be localized.

Although a comprehensive list cannot be presented here, some illustrations may be given. Anoxia, lack of oxygen for cellular respiration, can result from poisons acting at different sites. Phosgene or other inhaled corrosive vapors may produce massive flooding of the lungs with edema fluid and thus cause mechanical suffocation. Carbon monoxide acts by binding to hemoglobin with a much greater affinity than oxygen and so produces anoxia by interference with oxygen transport. Cyanide prevents tissue respiration at a terminal intracellular site by poisoning cytochrome oxidase which catalyzes the final step of oxygen utilization. Central nervous system depression, with coma and ultimate respiratory and circulatory failure, results from a large group of drugs: the general anesthetics (ether, chloroform, cyclopropane, and ethylene), the barbiturates, the opium alkaloids (morphine, codeine, and related compounds), and less regularly from many other poisons. Liver injury with jaundice (toxic hepatitis) is a prominent result of many poisons, characteristically the chlorinated hydrocarbons, elemental phosphorus, certain heavy metals (arsenic, antimony, barium, and copper), and a large diversity of organic compounds. Kidney damage results from many toxic chemicals, but is common after exposure to carbon tetrachloride, mercuric salts, and ethylene glycol. Convulsions, sometimes fatal, are a common response to acute poisoning by diverse agents, but are a particularly characteristic effect of strychnine, atropine, ergot, alkyl phosphates, picrotoxin, cocaine, and certain snake venoms. Among the medically important toxic agents are those that reduce or exaggerate effective transmission of the nerve impulse across synaptic junctions, especially in the autonomic nervous system. All such compounds, sometimes called autonomic agents, act so as to mimic or to block the physiological effects of either of the two known neurohumoral transmitter substances of the motor branches of autonomic nerves, acetylcholine and norepinephrine (**Fig. 1**). *See* ACETYLCHOLINE.

Poisons that act to exaggerate those physiological effects normally produced by acetylcholine release include muscarine and pilocarpine. The anticholinesterases (neostigmine, physostigmine, and the alkyl phosphates) have the same general physiological effects, although by a different mechanism. These compounds slow the heart, lower blood pressure, increase secretion of fluids into the respiratory tract, and narrow the respiratory passages. Toxic doses may produce death by suffocation or by failure of the circulation. Poisons that prevent acetylcholine from producing skeletal muscle contraction are curare and a number of related synthetic compounds used in anesthesia (succinylcholine, decamethonium). These drugs result in weakness of all voluntary muscle and may cause death through paralysis of respiratory muscles. Exaggeration of the normal effects of norepinephrine and epinephrine (adrenergic drugs) may produce dangerously high blood pressure, rapid heart action, and sometimes fatal disturbances of cardiac rhythm. *See* EPINEPHRINE.

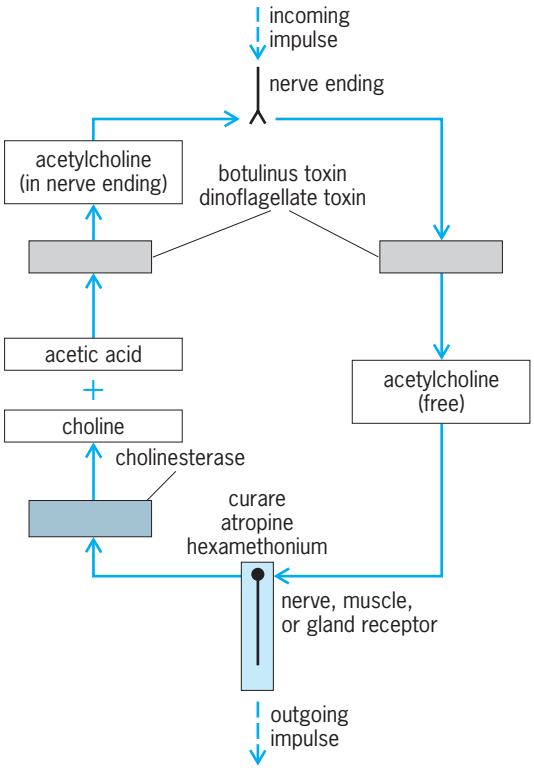

Fig. 1. Acetylcholine cycle, the target of many poisons. The "anticholinesterase" poisons prevent breakdown of acetylcholine by inactivating cholinesterase. Botulinus and dinoflagellate toxins hinder synthesis or release of acetylcholine. Curare and atropine make the receptor less sensitive to the chemical stimulus.

Many normal functions (glandular secretion and contraction of voluntary and smooth muscle) depend on the cyclic release, breakdown, and resynthesis of acetylcholine at the endings of cholinergic (acetylcholine-secreting) nerves. Figure 1 indicates the different ways in which several distinct poisons may interfere with the normal operation of the acetylcholine cycle. In the acetylcholine cycle, an impulse reaching a nerve ending liberates acetylcholine, which stimulates a receptor. The receptor is freed for further impulses by the enzyme cholinesterase, which breaks down acetylcholine into acetic acid and choline. These are resynthesized by other enzymes into new acetylcholine.

Mechanism of action. More precise understanding of the mechanism of poisons requires detailed knowledge of their action in chemical terms. Information of this kind is available for only a few

compounds, and then in only fragmentary detail. Poisons that inhibit acetylcholinesterase have toxic actions traceable to a single blocked enzyme reaction, hydrolysis of normally secreted acetylcholine. Detailed understanding of the mechanism of chemical inhibition of cholinesterase is not complete, but allows some prediction of chemical structures likely to act as inhibitors.

Carbon monoxide toxicity is also partly understood in chemical terms, since formation of carboxyhemoglobin, a form incapable of oxygen transport, is sufficient to explain the anoxic features of toxicity.

Heavy metal poisoning in many cases is thought to involve inhibition of enzymes by formation of metal mercaptides with enzyme sulfhydryl groups, the unsubstituted form of which is necessary for enzyme action. This is a general reaction that may occur with a variety of sulfhydryl-containing enzymes in the body. Specific susceptible enzymes whose inhibition explains toxicity have not yet been well documented.

Metabolic antagonists active as poisons function by competitive blocking of normal metabolic reactions. Some antagonists may act directly as enzyme inhibitors, others may be enzymatically altered to form derivatives which are even more potent inhibitors at a later metabolic step. An example of the latter is the biosynthetic incorporation of metabolite analogs into much more complex molecules, particularly the incorporation into nucleic acids of altered purine or pyrimidine bases, such as 8-azaguanine and 5-bromouracil, or the incorporation into proteins of altered amino acids, such as *p*-fluorophenylalanine or 7-azatryptophan. *See* ENZYME INHIBITION.

A well-studied example of biosynthetic production of a highly toxic metabolic analog arose from studies of fluoroacetate, a toxic component of plants of the genus *Dichapetalum*. Fluoroacetate is enzymatically converted to fluorocitrate, through reactions analogous to those normal steps leading from acetate to citrate. The fluorocitrate formed from fluoracetate is a potent inhibitor of aconitase, a citrate-utilizing enzyme in the tricarboxylic acid cycle, and this inhibition can explain the toxicity of fluoroacetate.

Where poison mechanisms are relatively well understood it has sometimes been possible to employ rationally selected antidotes. Thus the clarification of arsenic as a sulfhydryl poison led to the introduction of the effective metal-poisoning antidote, dimercaptopropanol (British antilewisite). Knowledge of the physiology of anticholinesterase poisons has permitted the use of rational physiological antagonists of excessive cholinergic stimulation, such as atropine and curare. On a more chemical basis, studies of the enzymatic mechanism of cholinesterase have led to more direct antidotes (pyridine aldoxime methiodide and other oximes) for the anticholinesterase poisons. These antidotes act by removing the inhibitor from the enzyme.

It should be noted that for the great majority of poisons the mechanism of toxic action is not as well understood as in the selected examples above. In only a few cases are poisons known to act directly by altering enzyme function. Other poisons, probably much more numerous, may involve reactions with less well-defined cellular components.

Toxicity rating	practically nontoxic	slightly toxic	moderately toxic	very toxic	extremely toxic	supertoxic	supremely toxic
Examples	glycerin, graphite, lanolin	ethyl alcohol, lysol, castor oil, soaps	methyl (wood) alcohol, kerosene, ether	tobacco, aspirin, boric acid, phenol, carbon tetrachloride	morphine, bichloride of mercury	potassium cyanide, heroin, atropine	botulinus toxin, some snake venoms
Probable lethal dose, mg/kg	more than 15,000	5000 to 15,000	500 to 5000	50 to 500	5 to 50	less than 5	less than 0.5
Probable lethal dose, for a 70-kg (155-lb) human	more than 1 qt	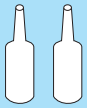 1 pt–1 qt	1 oz–1 pt	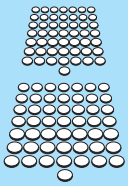 1 tsp–1 oz	7 drops–1 tsp	a taste (less than seven drops)	

Fig. 2. Toxicity-rate scale of substances according to the size of the probable lethal dose. Many drugs fall into the highly toxic categories. Drawings suggest the fatal doses of water, whiskey, ether, aspirin, morphine, and cyanide, the last three depicted in terms of aspirin-sized tablets. Scale suggested by M. N. Gleason, R. E. Gosselin, and H. C. Hodge.

Potency. The strength or potency of poisons is most frequently measured by the lethal dose, potency being inversely proportional to lethal dose (**Fig. 2**). Although killing dose represents a useful end point for potency appraisal, much smaller doses of a given poison may produce toxic changes in an organism. In estimating potency, a large test population should be used because of inherent individual variation in susceptibility. The relationship between percent of a population killed and dose of poison sometimes, but not often, conforms to a symmetrical sigmoid curve representing the integrated or cumulative form of the bell-shaped gaussian distribution curve. From statistically treated dose-response data, the dose killing 50% of the sample population can be determined, and is usually designated the MLD (median lethal dose) or LD_{50}. This is the commonest measure of toxic potency, although for special purposes the lethal dose for other fractions of the population may also be used (for example, LD_{10} and LD_{90}). The spread of the distribution (for example, dose difference between LD_1 and LD_{99}) is highly variable for different poisons and is of course also greatly influenced by the relative heterogeneity or homogeneity of the test population. *See* LETHAL DOSE 50.

Health aspects. The environment of modern industrial society abounds in toxic chemicals: drugs, industrial products, fuels, pesticides, and many common household products such as cleaning agents and paint products. Acute and chronic poisoning represent an appreciable public health hazard.

The multiplication of toxic chemicals in everyday use led, in 1953, to the establishment of information centers known as Poison Control Centers, which provide physicians with emergency telephone information concerning toxicity and antidotes for common poisons as well as for brand-named chemical formulations. These services were greatly assisted by the publication in 1957 of a compilation of brand-named products with information as to ingredients and quantitative toxicity. *See* TOXICOLOGY.

Elijah Adams

Bibliography. J. M. Arena, *Poisoning*: *Toxicology*, *Symptoms*, *Treatment*, 4th ed., 1979; R. H. Dreisbach, *Handbook of Poisoning*: *Prevention*, *Diagnosis*, *and Treatment*, 10th ed., 1982; M. N. Gleason, R. E. Gosselin, and H. C. Hodge, *Clinical Toxicology of Commercial Products*, 4th ed., 1976.

Poison gland

The specialized gland of certain fishes (**illus.** *a*), as well as the granular and some mucous glands of many aquatic and terrestrial Amphibia (illus. *b*). The poison glands of fishes are simple or slightly branched acinous structures which use the holocrine method of secreting a mucuslike substance. The poison glands of snakes are modified oral or salivary glands. Amphibian glands are simple, acinous, holocrine, with granular secretion. In some cases these amphibian poison glands produce mucus by a merocrine method of secretion. These glands function as protective devices. *See* EPITHELIUM; GLAND.

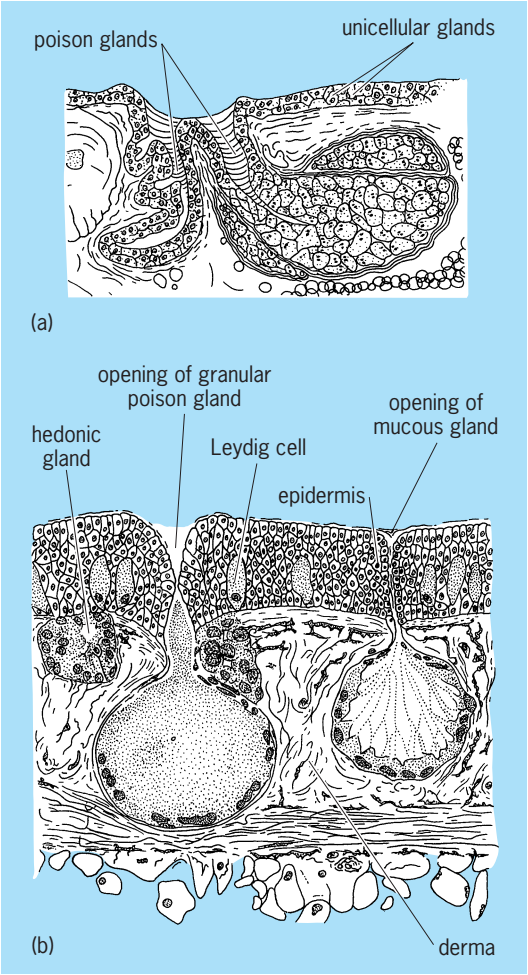

Poison glands. (*a*) Multicellular poison gland of catfish. (*b*) Amphibian skin, generalized.

Olin E. Nelsen

Poison ivy

A general name applied to certain species of the genus *Toxicodendron*, previously known as *Rhus*, in the sumac family (Anacardiaceae). *Toxicodendron radicans* is the poison ivy of eastern North America;

Three-foliolate leaf of poison ivy.

T. diversiloba is the poison oak of California. These plants are natives of North America. Both cause ivy poisoning, an annoying and often painful dermatitis.

Toxicodendron radicans, the most widespread species, is extremely variable. It has a bushy or climbing habit and three-foliolate leaves which are smooth and glossy or hairy and are entire, toothed, or lobed (see **illus.**). Poison ivy bears white fruits whereas the nonpoisonous sumacs bear red fruits. *See* SAPINDALES.

Perry D. Strausbaugh; Earl L. Core

Poison sumac

The plant *Toxicodendron vernix* (previously in the genus *Rhus*), a member of the sumac family (Anacardiaceae). It is an inhabitant of swamps ranging from Quebec to Minnesota, and southward to Florida, Louisiana, and Texas. It is a tall bush or small tree bearing pinnately compound leaves with 7–13 entire (without marginal teeth) leaflets, and drooping, axillary clusters of persisting white fruits (see **illus.**).

Poison sumac fruits and leaf (*Toxicodendron vernix*).

Like poison ivy, this plant is poisonous to touch, causing in many persons a severe inflammation of the skin, or dermatitis. The presence of white fruit separates this species from the nonpoisonous sumacs with their red fruits. *See* SAPINDALES.

Perry D. Strausbaugh; Earl L. Core

Poisonous plants

More than 700 species of seed plants, ferns, horsetails, and fungi that cause toxic, though rarely fatal, reactions in humans and animals (see **illus.**). Human allergic responses, including hay fever, asthma, and dermatitis, are widespread. Allergic responses are produced by many different plant species, but most common are poison ivy, poison sumac, and Pacific poison oak (all are species of *Toxicodendron*). Internal injury by toxic plants is less common but can be detrimental or lethal. *See* ASTHMA; POISON IVY; POISON SUMAC.

Glycoside-containing plants. Glycosides are common compounds in plants. They decompose to form one or more sugars, but sometimes the remaining compounds, aglycones, can be quite poisonous. Cyanogenic glycosides, which produce hydrocyanic acid, are found worldwide in many plant families; the best known are in the rose family (Rosaceae) and in the pea family (Fabaceae). Leaves, bark, and seeds of stone fruits such as cultivated and wild cherries, plums, peaches, bitter almonds, and apricots contain the glycoside amygdalin, which hydrolyzes to form hydrocyanic acid that can be fatally toxic to humans or animals. The same toxic substance is found in apple and pear seeds. Animals have been killed by high amounts of hydrocyanic acid in wilted leaves of wild species of cherry (*Prunus*). After ready absorption into the bloodstream, the acid causes death of body cells by blocking use of oxygen. Cattle poisoned by hydrocyanic acid develop convulsions, rapid breathing, coma, and a red froth at the muzzle. Cardiac glycosides are found in many unrelated species of

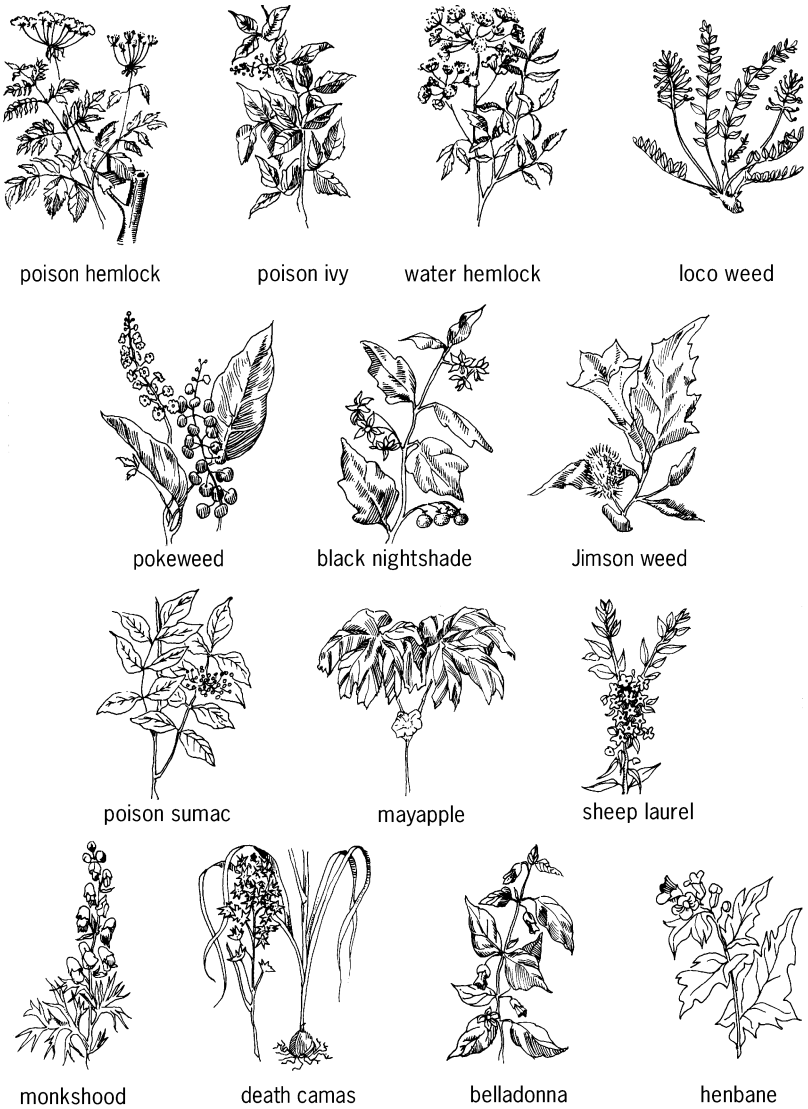

Poisonous plants found in the United States. (*After Webster's New International Dictionary, 2d ed., Merriam, 1959*)

plants. Those of the foxglove (*Digitalis purpurea*) contain a number of these glycosides used medicinally to slow and strengthen the heartbeat. Accidental poisoning or an overdose results in blurred vision, nausea, abdominal pain, and vomiting, and can be fatal. Oleander (*Nerium oleander*), which is cultivated in the warmer parts of the United States, contains a toxic glycoside that has an action similar to that of digitalis. The leaves have a very bitter taste that usually prevents their ingestion, but livestock have been poisoned from eating dried leaves mixed with edible feeds. *See* CYANIDE; DIGITALIS; GLYCOSIDE.

Alkaloid-containing plants. Alkaloids, compounds containing a nitrogen atom, have specific pharmacological effects on both humans and animals. Found in many different plant families, they have been used in drug therapy since ancient times, but misuse of these plants can produce poisonings. The potato family (Solanaceae) has many species that contain a number of alkaloids. Hyoscyamine and atropine are the alkaloids occurring in belladonna or deadly nightshade (*Atropa belladonna*), black henbane (*Hyoscyamus niger*), thornapples and jimsonweed (*Datura*), and tree daturas or angel's-trumpets (*Brugmansia*). Ingestion of these can produce pupil dilation, nausea, muscular incoordination, delirium, and finally death from respiratory paralysis and heart failure. The black nightshades, which comprise a number of species of *Solanum*, are common weeds in the United States. Plants of *Solanum* contain glycoalkaloids that break down to aglycones such as solanine and solanidine. Eating green berries of *Solanum* plants can result in dizziness, vomiting, diarrhea, and changes in heartbeat, but rarely causes death.

Plants of tobacco, *Nicotiana*, contain numerous alkaloids, principally the very toxic nicotine or its isomer anabasine. Humans can smoke tobacco because nicotine is detoxified in the body. Nicotine poisoning produces heart palpitations, clammy skin, nausea, vomiting, diarrhea, labored breathing, and even respiratory paralysis. In severe poisoning death may be rapid. Plants of poison hemlock (*Conium macalatum*) have several alkaloids similar to nicotine, which affect the central nervous system.

Plants of rattlebox (*Crotalaria*), groundsel (*Senecio*), and fiddleneck (*Amsinckia*) have alkaloids of similar molecular structure. Livestock losses have occurred from ingestion of hay containing seeds or dried plants of these genera. The alkaloids cause liver damage by constricting the smaller blood vessels. Symptoms, chronic or acute, vary in different animals. Pigs, cattle, horses, and poultry are particularly susceptible.

Anagyrine is a toxic alkaloid found in several species of lupine (*Lupinus*) in the western United States. It is especially abundant in young plants in spring, and in seeds and fruits during late summer. Although lupine poisoning is most common in sheep, other animals such as cattle, horses, swine, and deer are also affected. Acute poisoning produces nervousness, depression, loss of muscular control, convulsions, and coma. Anagyrine can move through the fetal membranes, producing deformed calves without affecting the pregnant dam.

Alkaloids present in species of monkshood (*Aconitum*) are extremely toxic. Ingestion of any part of the plant, but particularly roots and seeds, produces cardiac rhythmical disturbances and may be fatal to humans or livestock in 12 to 6 h. Symptoms include lowering of blood pressure and convulsions. Livestock may develop bloat. Larkspur plants (*Delphinium*) have similar toxic alkaloids affecting the central nervous system, causing excitability and muscular spasms.

Plants of false hellebores (*Veratrum*) and death camas (*Zigadenus*) have complex alkaloids of similar structure and cause livestock deaths in the western United States. Sheep that eat false hellebore plants in early pregnancy may produce deformed lambs. Other livestock and humans have also been poisoned by false hellebores, developing lowered blood pressure and slowed cardiac action. Nausea and vomiting may also occur. Although normally not eaten, death camas plants, the first to grow in the spring, may cause extensive loss of sheep. Dried death camas plants in hay have also caused death of livestock. Human poisonings have occurred by bulbs of this species mistakenly eaten for onions or edible camas (*Camassia*). Symptoms include salivation and vomiting, slowed heart action, lowered blood pressure, convulsions, and coma. *See* ALKALOID.

Heath plants. Toxic resins, andromedotoxins, formed by members of the heath family (Ericaceae) are derived from diterpenes. The most toxic species are mountain laurel, sheep laurel, and bog laurel, all in the genus *Kalmia*. Livestock develop symptoms within 6 h of ingestion, including copious salivation, vomiting, bloat, diarrhea, loss of muscular control, slowed heartbeat, coma, and death. Human toxicity has also been recorded.

Pokeweed. Pokeweed (*Phytolacca americana*) is a garden weed throughout the United States. but it is native to the eastern and central areas. The entire plant, especially the seeds and the large root, is poisonous. Young shoots are edible after boiling in three separate waters. Human poisonings have resulted from inadvertently including parts of the root along with the shoots. Symptoms include an immediate burning bitter taste, salivation, vomiting, diarrhea, and possible shock. Recovery is usually within 1 day.

Waterhemlock. The waterhemlocks (*Cicuta*) are widespread in North America. The large underground tubers, mistakenly considered edible, have caused human poisonings and death. Livestock usually do not eat waterhemlock, but have been fatally poisoned. The highly toxic principle, an unsaturated aliphatic alcohol which acts directly on the central nervous system, causes violent convulsions and in less than 1 h may be followed by paralysis and death.

Oxalate poisoning. Oxalic acid, as oxalate salts, accumulates in large amounts in some species of plants such as those of the genera *Halogeton*, *Bassia*, *Rumex*, and *Oxalis*. It reacts with calcium, removing it from the blood. Crystals of insoluble calcium

oxalate cause mechanical damage to the kidneys and other organs. Livestock may eat plants with large oxalate content; death of sheep and cattle can occur quickly as a result of calcium deficiency. Chronic poisoning produces kidney damage which may result in death accompanied by paralysis. Horses show bone degeneration from calcium deficiency.

Nitrate poisoning. Nitrate poisoning is widespread and results in many cattle deaths yearly. Any disruption of the normal synthesis of nitrates into amino acids and proteins causes large accumulations of nitrates in various species of plants, particularly in the goosefoot family (Chenopodiaceae). A nitrate content of 1.5% or more of the plant's dry weight is considered lethally toxic. Individual plants in the same pasture show great variations of nitrate content, but some plants accumulate as much as 10% or more. In the rumen of cattle, nitrates are converted to nitrites which are absorbed into the bloodstream and change hemoglobin into chocolate-colored methemoglobin which cannot carry oxygen. Animals die manifesting severe lethargy with increased cardiac and respiratory rates.

Allelopathic toxins. Allelopathic phytotoxins are chemical compounds produced by vascular plants that inhibit the growth of other vascular plants. Residues of grain sorghum (*Sorghum bicolor*) can markedly reduce the following year's growth of wheat and some weedy grass seedlings. Seeds of jimsonweed (*Datura stramonium*) will not germinate until the alkaloids are washed from the seed coat. However, these same alkaloids prevent the growth of seedlings of other plant species in the vicinity. Locoweeds, belonging to the genera *Oxytropis* and *Astragalus*, produce loss of livestock that become addicted to eating these unpalatable plants. The unknown toxin causes the animals to develop incoordination, erratic gaits, and excitability followed by depression. Other species of *Astragalus* produce toxic aliphatic nitro compounds. Acute nitro poisoning results in weakness and labored breathing with death within 20 h. Chronic poisoning is manifested as weakness in the hindquarters, a peculiar goose-stepping gait, and blurred vision. Poisoned animals that are forced to move may collapse and die. Still other species of *Astragalus* accumulate toxic quantities of selenium which usually causes chronic poisoning of livestock. They may develop excitability followed by depression, emaciation, loss of hooves, and death. *See* ALLELOPATHY.

Fungi. Every year human fatalities occur from ingestion of wild poisonous mushrooms. Those gathered in the wild require individual identification since toxic species may grow alongside edible ones.

Humans were poisoned by ergot fungal bodies in grain in past centuries, but modern grain-cleaning processes have eliminated this problem. Ergot fungus, infecting many species of grasses, causes widespread poisoning of livestock. Normal grass grains are replaced by hardened masses of ergot fungus (*Claviceps purpurea*) containing a number of toxic alkaloids and amines. Cattle show acute stimulation of the central nervous system, including stiff gait and lateral jumping. Chronic ergot poisoning produces contractions of the smaller blood vessels, leading to gangrene and sloughing of the tail, hooves, and ear tips. *See* ERGOT AND ERGOTISM; TOXICOLOGY.

Thomas C. Fuller

Bibliography. J. W. Hardin and J. M. Arena, *Human Poisoning from Native and Cultivated Plants*, 2d ed., 1974; R. F. Keeler, K. R. Van Kampen, and L. F. James (eds.), *Effects of Poisonous Plants on Livestock: Symposium on Poisonous Plants, Utah State University, 1977*, 1978; J. M. Kingsbury, *Poisonous Plants of the United States and Canada*, 1964; W. H. Lewis and M. P. F. Elvin-Lewis, *Medical Botany: Plants Affecting Man's Health*, 1977.

Polar meteorology

The science of weather and climate in the high latitudes of the Earth. In the polar regions the Sun never rises far above the horizon and remains below it continuously for part of the year, so that snow and ice can persist for long periods even at low elevations. The meteorological processes that result have distinctive local and large-scale characteristics in both polar regions.

The polar regions are shown in **Fig. 1** as mosaics of satellite photos taken in late spring. In winter the effective area of the Arctic region is enlarged threefold by snow covers that average 2.7×10^6 mi^2 (7×10^6 km^2) of the land surrounding the Arctic Ocean, and by an expansion of the area of the sea ice from 2×10^6 mi^2 (5×10^6 km^2) to 4.5×10^6 mi^2 (11.5×10^6 km^2). Similarly sea ice forming on the Southern Ocean each winter more than doubles the total ice-covered area of Antarctica, from 5.8×10^6 mi^2 (1.5×10^7 km^2). *See* ANTARCTIC CIRCLE; ANTARCTICA; ARCTIC CIRCLE.

Energy budgets and temperatures. A general feature of polar energy budgets is their small radiative income. Even in midsummer, when the height of the Sun above the horizon in latitude L varies in the range $23.5 \pm (90 - L)$ degrees, the average daily energy intake of a horizontal unit surface at the top of the atmosphere is reduced by about 60% ($= 1 - \sin 23.5$) to 550 W/m^2 from the 1370 W/m^2 received where the Sun is directly overhead. After further losses during its long slanted passage through the polar atmosphere, up to 80% of the remaining solar radiation is reflected before it can warm a surface covered by snow. In that case, the surface will become extremely cold unless it can draw upon the energy stored above it in warmer air advected from lower latitudes, and below it in the ground or ocean.

Representative energy budgets of typical polar surfaces are shown in the **table**. All have an excess of radiation in summer, but they use it differently. On the Arctic tundra the entire snow pack is melted; temperatures can then rise well above freezing, and so a great deal of the surplus energy is transferred to the atmosphere. Melting accounts for an even larger part of the radiational surplus on Arctic sea ice, but less

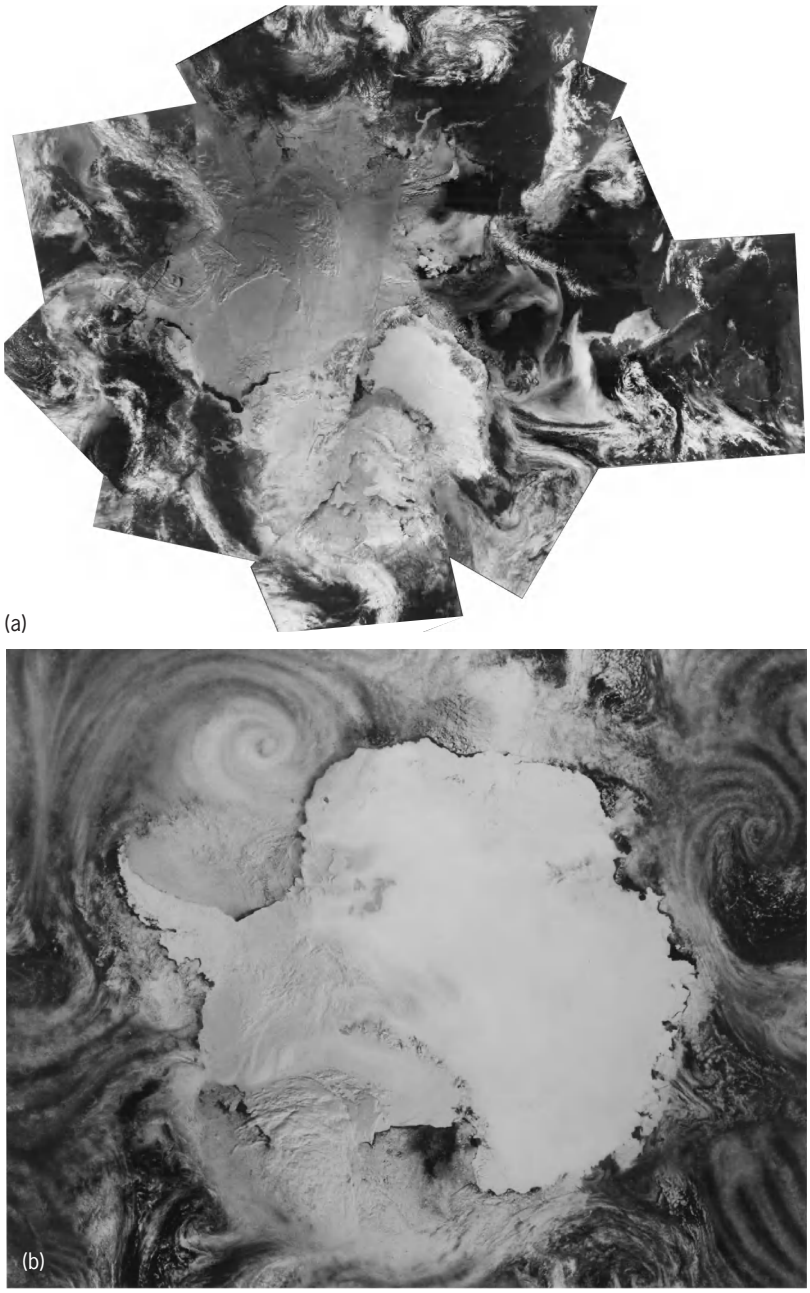

Fig. 1. Composites of polar regions as photographed by meteorological satellites. (*a*) Arctic, May 1989. (*b*) Antarctic, November 4–25, 1986. (*From U.S. Air Force Defense Meteorological Satellite Program film, NOAA/NESDIS, CIRES/National Snow and Ice Data Center, Boulder*)

heat goes into the atmosphere since the temperature of the ice surface cannot rise above 32°F (0°C). In Antarctica the various budget items have smaller values, and melting occurs only near the coast, where even midsummer monthly mean temperatures remain below freezing. In winter, radiational deficits prevail and produce temperatures that increase with height in the lower atmosphere (temperature inversions), which indicate that heat is being transferred from the air to the surface. The strongest inversions occur on the Antarctic plateau, where Vostok (elevation 11,440 ft or 3488 m) has a midwinter monthly mean temperature of −76°F (−60°C), and a lowest recorded temperature of −129.3°F (−89.6°C); even in midsummer the monthly mean temperature there is only −26°F (−32°C). *See* AIR MASS; SEA ICE; TEMPERATURE INVERSION; TUNDRA.

Fuller details of the temperatures and other meteorological parameters on the large ice sheets of Antarctica and Greenland are provided by automatic weather stations reporting via satellites (**Fig. 2**). The annual mean surface temperatures of the ice sheets lie between 32°F (0°C) and 14°F (−10°C) at their edges and decrease to −26°F (−32°C) in central Greenland and to below −76°F (−60°C) in the central Antarctic; they have been determined independently as snow temperatures some 33 ft (10 m) below the surface, where the seasonal air temperature changes are smoothed out by the heat transfer through the snow. *See* METEOROLOGICAL SATELLITES.

On the global scale the overall energy deficit of the polar regions can be seen in satellite measurements made during the Earth Radiation Budget (ERB) project. These have shown that the effective temperature (proportional to the fourth root of the long-wave radiation loss into space) of the Arctic is −25.6°F (−32°C), and that of the Antarctic, −65.2°F (−54°C), compared with −9.4°F (−23°C) for the Earth as a whole. These so-called planetary temperature differences between low and high latitudes find expression in the atmospheric circulations of the polar regions. *See* ATMOSPHERIC GENERAL CIRCULATION; HEAT BALANCE, TERRESTRIAL ATMOSPHERIC; INSOLATION; TERRESTRIAL RADIATION.

Pressure and wind. Frequent storms form along the boundary between polar and middle-latitude air masses (the polar front), and they create belts of low surface pressure (polar troughs) around both polar regions. In the north these storms move freely across the Arctic Basin, which they enter mainly via the Norwegian Sea; other small, intense polar lows often form rapidly in polar air outbreaks and accompany them into lower latitudes. By contrast, due to the Antarctic ice sheet [which has an average elevation exceeding 6560 ft (2 km) and a central plateau rising above 13,200 ft (4 km)], most Antarctic storms and their associated very strong westerly winds remain over the surrounding Southern Ocean, although a few storms move on to the ice sheet by way of the low-level Ross and Filchner-Ronne ice shelves. The atmospheric pressure along the ice sheet surface therefore depends primarily on elevation; however, it undergoes a general increase by 10–15 hectopascals (10–15 millibars) from October to December, when radiative heating expands the lower atmospheric layers around the continent. That permits some of the raised air to flood the ice sheet; the accompanying reduction in surface pressure over the rest of the globe was discovered a century ago and was later used to estimate correctly the average height of the ice sheet before its inner regions were explored. *See* FRONT; GLACIOLOGY; STORM.

The average circulation over the Antarctic ice sheet consists of inflow in the middle and upper troposphere and outflow in a shallow surface layer descending along the ice sheet; a similar katabatic drainage of cold surface air toward the coast occurs

Polar surface energy fluxes, in W/m^2*

Area	Net radiation, gain (+) or loss (−) by surface	Energy from above, gain (+) or loss (−) by surface	Energy from below, gain (+) or loss (−) [also energy used for melt]
Arctic tundra			
Summer	104	−75	−29
Winter	−17	15	2
Arctic sea ice			
Summer	74	−13	−61
Winter	−29	15	14
Antarctic coast			
Summer	31	−6	−25
Winter	−41	7	34
Antarctic plateau			
Summer	20	−18	−2
Winter	−19	17	2

*After G. E. Weller and G. Wendler, The energy budget over various terrain types in polar regions, *Ann. Glaciol.*, vol. 14, 1990.

also in Greenland. The Antarctic katabatic winds become extraordinarily strong in some coastal regions. Thus, at Port Martin, Adelie Land, in 1951 there was recorded a 24-h mean wind of 108.6 mi/h (48.5 m/s); the strongest monthly mean wind reached 65 mi/h (29.1 m/s). Terra Nova Bay, Victoria Land, is another strong-wind location where the katabatic outflow can actually be seen in infrared satellite photos as a warmer plume over the colder sea ice. The katabatic winds are sufficiently strong and persistent, even inland, to cut up the snow surface into sharp ridges. Known as sastrugi, from the Russian word for a wood plane, these winds have been mapped from the air and in satellite photos to establish the broad directional pattern of the surface flow (**Fig. 3***a*). Computer model simulations of the katabatic flow have refined these patterns and identified the major outflow concentrations (Fig. 3*b*). *See* TROPOSPHERE; WIND.

Fig. 2. Automatic weather station on the Antarctic ice sheet. (*Courtesy of G. Wendler*)

In the long polar winter night the upper layers of the atmosphere become very cold, and a persistent stratospheric jet stream is created at the boundary between the warmer sunlit and colder dark parts of the atmosphere. Inside the stable polar vortex thus formed, the stratospheric ozone previously created by solar radiation decays until the return of the Sun in spring destroys the vortex. Over Antarctica, that normal decay has become enhanced, apparently through reactions of the ozone with chemicals produced by human activities; if this process continues, this ozone hole could ultimately weaken the global ozone shield. *See* ATMOSPHERE; STRATOSPHERIC OZONE; JET STREAM; STRATOSPHERE.

Precipitation. Polar precipitation, like solar radiation, faces a physical limitation: the atmosphere's saturation limit for water vapor, which decreases from about 0.38% of atmospheric mass (6.1 hPa or 6.1 mbar vapor pressure) at 32°F (0°C) to 2.5×10^{-3}% (0.04 hPa or 0.04 mbar) at −58°F (−50°C). Both polar regions receive around 8 in. (20 cm) of water as annual precipitation, one-fifth of the average global precipitation rate of 37.4 in. (95 cm) per year. Since the area inside the two polar circles represents 8.3% of the global surface, the polar precipitation accounts for a mere 1.7% of the total global precipitation of 115,000 mi^3 (480,000 km^3) per year. Special interest attaches to the snow, which over millions of years has created the Antarctic and Greenland ice sheets; together they presently contain 6.5×10^6 mi^3 (2.7×10^7 km^3) of ice, which if melted would raise sea level by 203 ft (62 m); this figure takes into account a large portion of the Antarctic ice presently afloat or below sea level.

Since the yearly global precipitation equals 35 times the average atmospheric water-vapor content of some 3130 mi^3 (13,000 km^3) of water, the latter has

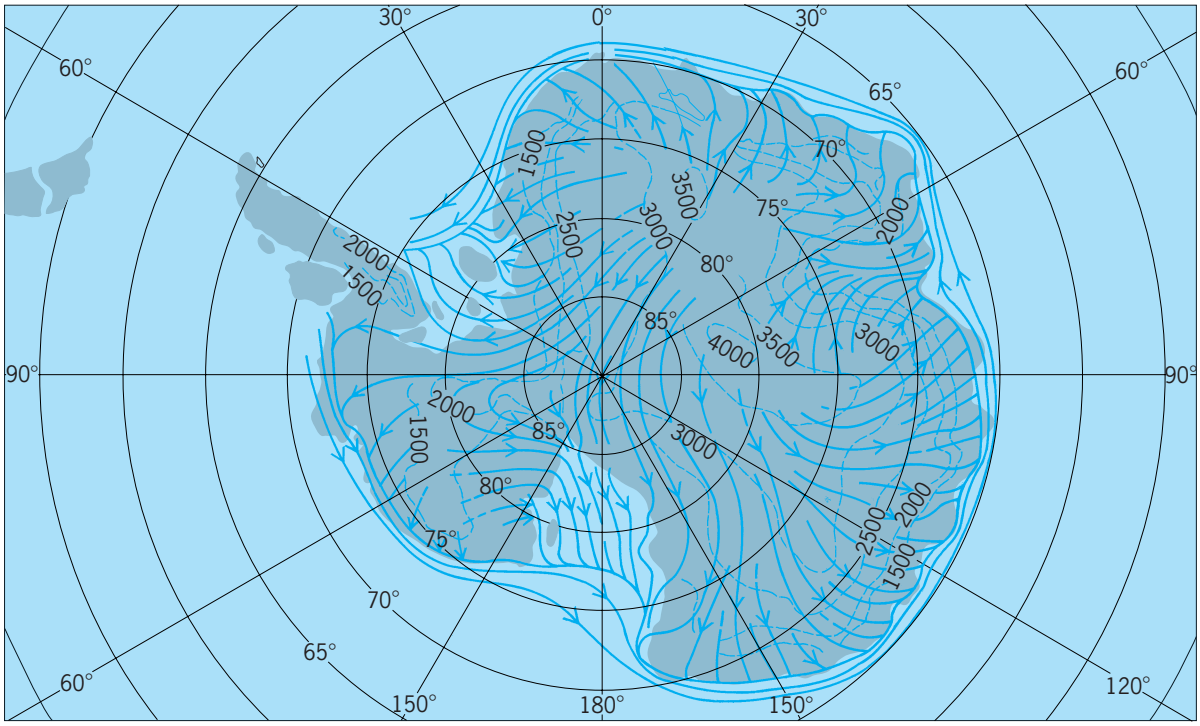

(a)

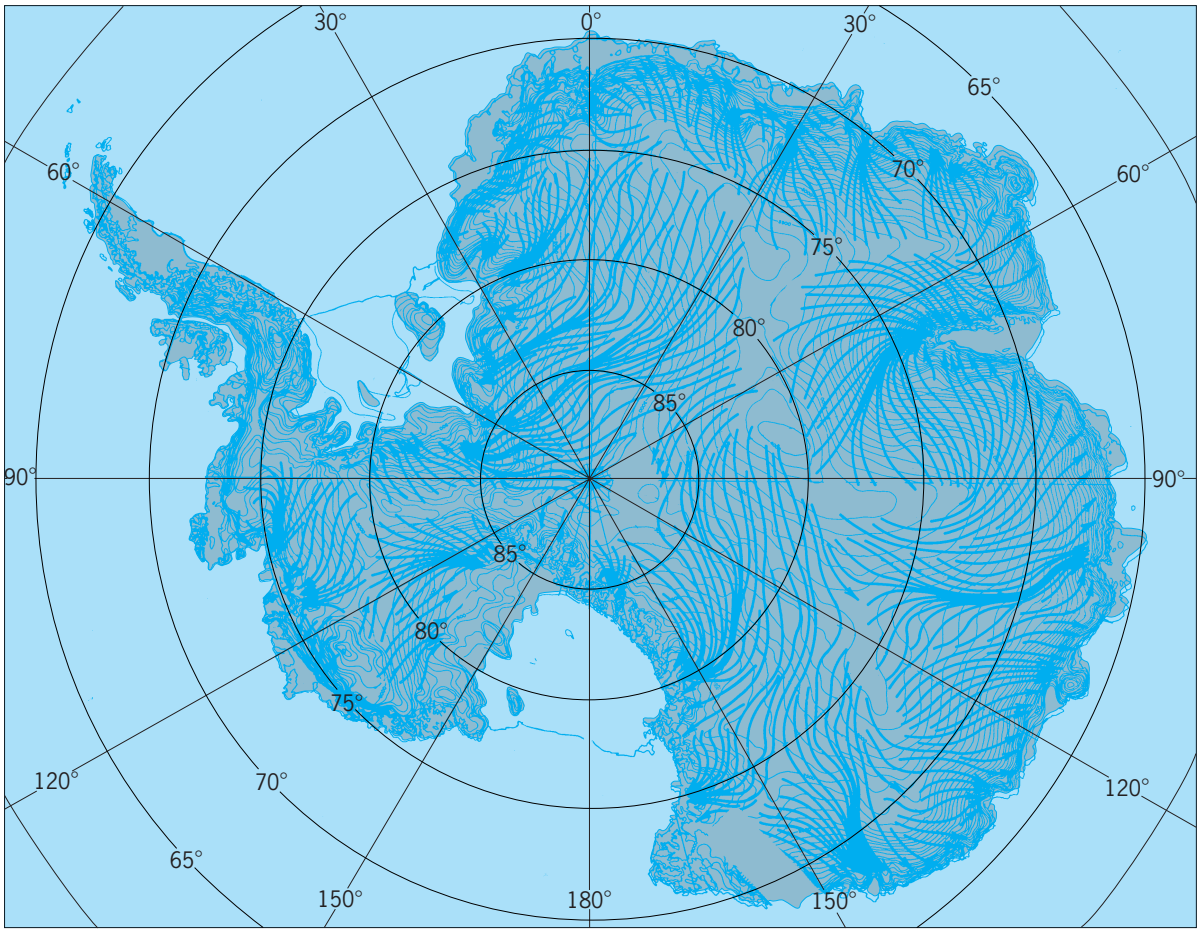

(b)

Fig. 3. Mean streamlines of the katabatic wind over the Antarctic ice sheet (*a*) as constructed from sastrugi directions and (*b*) as computed with an atmospheric model by T. Parish. Note the streamline concentrations indicating very strong winds.

an atmospheric residence time of only 11 days. By contrast, the small fraction deposited as snow on the ice sheets consolidates to ice and remains stored for thousands of years until it is discharged in icebergs. During that time the ice acquires a wealth of information on changes in polar and global climate, in the atmospheric composition, and even in the height of the ice sheet surface. The climatic information is provided by the isotopic composition of the ice, which contains fewer (more) heavy oxygen (^{18}O) and hydrogen (^{2}H) atoms in periods of lower (higher) temperatures; furthermore, the dust content of the ice has changed markedly with the changing strength of the prevailing winds and extent of deserts. Air bubbles trapped in the ice have revealed past changes in atmospheric composition [in particular, lower amounts of carbon dioxide (CO_2) and methane (CH_4)] during glacial periods. Moreover, the amounts of air enclosed have provided the atmospheric pressure (and hence the elevation) of the ice sheet surface at the time that its snow was compacted to ice. *See* GLACIAL EPOCH; PALEOCLIMATOLOGY.

The snow accumulating in the course of a year provides the best estimate of the current precipitation rate, which is difficult to measure directly in the windy conditions prevailing on the ice sheets. The winds are accompanied by almost incessant snow drift that has been calculated to carry during a year as much as 1×10^8 tons of snow across each 0.6 mi (1 km) of coastline at one of the coastal strong-wind sites. Inland transports are one to two orders of magnitude smaller and directed more nearly along the height contours (Fig. 3); consequently only about 5% of the annual accumulation on the Antarctic ice sheet and 1% of that on the Greenland ice sheet leave the ice sheets as drifting snow. The total amount of snow gained by the ice sheets each year and the ice losses due to melting and iceberg calving determine whether the ice sheets are tending to grow or shrink in the current polar climate; no firm answer has yet been reached to this question that has a strong bearing on future sea levels. *See* ICEBERG; PRECIPITATION (METEOROLOGY); SNOW.

Physiological aspects of polar climate. The wind-blown snow is a feature of the polar climate that can cause serious problems. It can markedly restrict low-level visibility and create a so-called white-out in which all contrasts are eliminated for the eye by a uniform scattering of light in all directions. In combination with low temperatures, the strong polar winds produce large body-heat losses. The windchill temperature in a 45 mi/h (20-m/s) wind at −40°F (−40°C) is −121°F (−85°C), and at −112°F (−80°C) even a mere 11.2-mi/h (5-m/s) wind feels like its temperature is −116°F (−110°C). A further effect is a loss of body moisture to the very dry polar air; just breathing that air, with a moisture content of around 10 oz/ft^3 (0.1 g/m), has been estimated to involve a body loss of 0.7 fluid oz (0.3 liter) of water or more in 24 h. *See* BIOMETEOROLOGY; COMFORT TEMPERATURES.

Uwe Radok

Bibliography. Atmospheric sciences in Antarctica: Special section, *Rev. Geophys.*, vol. 26, 1988; J. O. Fletcher and U. Radok, The role of polar regions in climate dynamics, Appendix A of Polar Research Board, *The Polar Regions and Climatic Change,* 1984; Y. Herman (ed.), *The Arctic Seas: Climatology, Oceanography, Biology, and Geology*, 1990; S. Orvig (ed.), *Climates of the Polar Regions,* World Survey of Climatology, vol. 14, 1971; Polar Group, Polar atmosphere-ice-ocean processes: A review of polar problems in climate research, *Rev. Geophys. Space Phys.*, 18:525–543, 1980; W. Schwerdtfeger, *Weather and Climate of the Antarctic*, 1984.

Polar molecule

A molecule possessing a permanent electric dipole moment. Molecules containing atoms of more than one element are polar except where forbidden by symmetry; molecules formed from atoms of a single element are nonpolar (except ozone). The dipole moments of polar molecules result in stronger intermolecular attraction, increased viscosities, higher melting and boiling points, and greater solubility in polar solvents than in nonpolar molecules. *See* DIPOLE MOMENT.

The electrical response of polar molecules depends in part on their partial alignment in an electric field, the alignment being opposed by thermal agitation forces. This orientation polarization is strongly temperature-dependent, in contrast to the induced polarization of nonpolar molecules. *See* FERROELECTRICS; MOLECULAR STRUCTURE AND SPECTRA; PERMITTIVITY; POLARIZATION OF DIELECTRICS.

Robert D. Waldron

Polar navigation

The complex of navigational techniques modified from those used in other areas to suit the distinctive regional character of polar areas. Although polar navigation has become routine to a rising number of navigators operating in and through such high-latitude parts of the word, their success continues to be based on a sound grasp of the regional differences and the developing adaptations of navigational principles and aids to suit these peculiar area needs. This article therefore singles out salient physical differences, from those of other areas, which demand navigational modifications and relates them to the best-known special techniques and applications.

Coordinates, directions, and charts. In polar regions the meridians radiate outward from the poles, and parallels are concentric circles. Thus the rectangular coordinates familiar to the navigator accustomed to using the Mercator projection are replaced by polar coordinates. A rhumb line is of little use in high latitudes. The usual significance of direction decreases as latitude increases, disappearing at the poles, where all directions are south (or north).

Polar navigational grid. For navigation, it is convenient to restore rectangular coordinates. This is done by placing a series of parallel lines on the chart, and

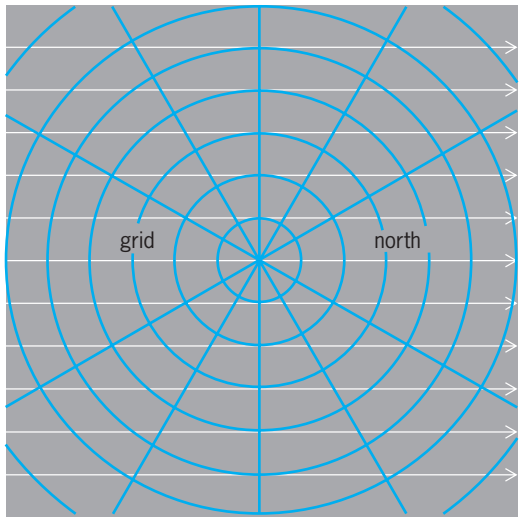

Fig. 1. Polar stereographic graticule with polar grid.

letting them represent fictitious meridians. This is called a navigational grid. Generally, these lines are parallel to the Greenwich meridian, and grid north is taken as north along this meridian (**Fig. 1**). In this way, grid and true directions near the poles are interconverted by the addition or subtraction of longitude. Variation of the compass (magnetic declination) is replaced by grid variation, or grivation, which is the angle between a magnetic meridian and a grid meridian. Thus grid and magnetic directions are interconverted by the addition or subtraction of grid variation.

Selection of chart projection. As latitude increases, the Mercator chart so familiar to the marine navigator of lower latitudes becomes of decreased utility. In subpolar regions—roughly in the area between the 60° and 70° parallels of latitude—the Lambert conformal projection is used increasingly for nautical charts. In higher latitudes both nautical and aeronautical charts are generally on a polar projection. The polar stereographic and transverse Mercator projections are used most commonly, but the modified Lambert conformal (Ney's), polar gnomonic, and polar azimuthal equidistant projections are also used. *See* MAP PROJECTIONS.

Light conditions. At the poles the day and year become synonymous as regards the proportion of daylight to darkness. The duration of twilight is lengthened proportionately, lasting several weeks. About 32 h elapse from the time the lower limb of the Sun touches the horizon until the upper limb disappears at sunset, during which time the Sun travels $1^1/_3$ times (480°) around the horizon. The period of several days that occurs each spring and autumn when the Sun is below the horizon but too close to it for other celestial bodies (except the Moon) to be visible is a critical time for navigation, because of the great dependence upon celestial bodies for both position and direction determination.

Celestial sphere and time patterns. The diurnal motion of celestial bodies is essentially horizontal. Half of the celestial sphere is always visible at each pole, and the other half is always below the horizon.

Time loses its usual significance in regions where all time zones come together. The time of any zone would be about equally satisfactory. It is common practice to keep Greenwich mean time. *See* TIME.

Meteorological factors. Temperatures are generally low, but not as low as might be supposed. It is true that the coldest atmospheric temperatures recorded have been on the Antarctic continent, but temperatures at the North Pole are never as low as some recorded in Yellowstone National Park. Although few places in Antarctica ever get above freezing, much of the Arctic does so regularly each summer. A temperature of 100°F (38°C) has been recorded on the Arctic Circle. *See* POLAR METEOROLOGY.

The polar air is relatively dry, but because of the low prevailing temperatures, fog and clouds are common. Visibility may also be limited by blowing snow. In most parts of the polar regions, precipitation is so light that these areas are sometimes classed as deserts. Nevertheless, over much of both polar areas the ice does not completely melt during the summer, and large quantities of ice are left both on the land and in the sea. When there is no fog, the atmosphere is exceptionally clear. In the clear, cold air of polar regions sounds travel great distances. It is not unusual to hear the bark of a dog 10 mi (16 km) away. *See* ATMOSPHERIC ACOUSTICS.

The strongest surface winds in the world are probably encountered in certain regions of Antarctica, where speeds of more than 200 knots (100 m/s) are not unusual. Over the Arctic Ocean, however, strong winds are not encountered except in some regions near land.

The conditions described are primarily responsible for the differences between navigation in polar regions and elsewhere. There is no well-defined line of demarcation between polar regions and subpolar regions, nor is there universal agreement as to the definition of polar regions. For the purposes of this discussion, however, the parallel of latitude 70° can conveniently be used as a general dividing line. Polar grid navigation is usually limited to the latitudes poleward of this parallel.

Piloting hazards. Piloting in polar regions is strongly affected by the absence of any great number of aids to navigation. Also, natural landmarks may not be shown on the chart, or may be difficult to identify. The appearance of some landmarks changes markedly under different ice conditions. When snow covers both the land and a wide ice foot attached to the shore and extending for miles seaward, even the shoreline is difficult to locate. Adjacent islands sometimes merge together as the straits between them fill up completely with ice. Along a rugged coast such as that of much of Greenland, snow-covered headlands may look alike. *See* PILOTING.

Unreliability of charts. Charts of polar regions are less reliable than those of other regions, because relatively little surveying has been done in the polar areas. Attempts to fix the position of a craft can be discouraging when the various landmarks used are not charted in correct position relative to each other. Marine navigators sometimes plot positions relative to land known to be shown in the wrong place,

rather than in the correct geographical positions, because it is the land and its attendant shoals that constitute danger to their vessels.

Because relatively few soundings are shown on charts, ships entering harbors often send small boats ahead to determine the depth of water available. The polar marine navigator finds a knowledge of geology useful in predicting safe areas and those areas in which rocks and shoals might be encountered. Ice concentrations and movement may also be an indication of the relative depth of the water.

The reliability of charts of polar regions has steadily improved as additional information has become available.

Limitations of electronic applications. Coverage of electronic navigation systems has also improved, but is still somewhat limited. Omega signals are available throughout both polar regions. However, there are several sources of error in the propagation of the very low frequencies utilized by the Omega system. Some of these are predictable, permitting corrections to be applied. But temporary sources of anomalous propagation identified as sudden ionospheric disturbances caused by influx of x-rays from the Sun and polar cap disturbances associated with proton enhancement by the Sun are more troublesome. In the Arctic, rapid attenuation of very low frequency signals is experienced when the signals cross the Greenland ice cap. Additionally, user equipment-installation problems can be severe, notably in aircraft, where electromagnetic interference problems are particularly troublesome. For maximum effective utilization of Omega signals, accurate absolute time is available, resulting in dependence on time signals from an external source. Loran C sky waves are available throughout the Arctic, and ground waves extend to some parts of this area, but neither ground waves nor sky waves are available in the Antarctic. *See* LORAN.

Radar is useful, but experience in interpretation of the scope in polar regions is essential for reliable results. This is particularly true in aircraft, where the relative appearance of water and land areas often reverse in winter and summer. Hummocked ice presents a different appearance from unbroken ice. A radio direction finder is useful, when radio signals are available. Except along the northern coast of Alaska and Russia, few radio beacons are available. The use of electronics in polar regions is further restricted by magnetic storms, which are particularly severe in the auroral zones. *See* ELECTRONIC NAVIGATION SYSTEMS; RADAR.

Difficulties of dead reckoning. Reliable dead reckoning depends upon the availability of accurate measurement of direction and distance (or speed and time). There are difficulties in meeting these requirements in polar regions. *See* DEAD RECKONING.

Compass limitations. Direction is measured largely by a compass. The compasses in common use are the magnetic, in which the directive element attempts to align itself with the horizontal component of the Earth's magnetic field, and the north-seeking gyro, in which the directive element attempts to align itself with the Earth's axis of rotation. The magnetic compass becomes unreliable in the vicinity of the magnetic poles of the Earth, and the north-seeking gyrocompass becomes unreliable in the vicinity of the geographical poles of the Earth. The magnetic and geographical poles are both in high latitudes, and the areas in which the compass errors are large and somewhat erratic overlap. In these areas frequent checks and comparisons are needed.

One solution to the directional problem in high latitudes is to use a directional gyro, a gyroscopic device that maintains its axis in a set direction but must be reset at frequent intervals, as every 15 min, because of gyroscopic drift. One of the devices discussed below for celestial direction determination can provide the directional data needed for setting or resetting a directional gyro. Directional gyros are not generally used aboard ship, and in aircraft they have been largely replaced by inertial navigators, which are also used by submarines and to a limited extent by surface ships. However, because inertial systems are self-contained, without reference to external signals, the insertion of incorrect data is always a possibility. Proposals to remedy this situation by means of geostationary satellites have encountered the problem that such satellites do not provide coverage beyond 65° latitude. *See* AIRCRAFT COMPASS SYSTEM; GYROSCOPE; INERTIAL GUIDANCE SYSTEM.

Celestial determination of direction. Direction can be determined by means of celestial bodies, but this is usually an instantaneous indication used to check the compasses, unless a device equipped with a clock mechanism is used to provide more continuous information. Several types of devices have been developed to facilitate the use of celestial bodies for determination of direction. The oldest is the sun compass, which utilizes the shadow of a shadow pin, or gnomon, and a suitable dial. This, of course, will not operate unless the Sun is visible. A sky compass indicates direction of the Sun by means of polarized light in the sky when the Sun is near the horizon, even though it may be below the horizon or otherwise obscured. This device may offer the only means of determining direction during the brighter part of the long polar twilight. A Canadian version is called a twilight compass because of the principal period of its use. An astrocompass (**Fig. 2**) can be set for the coordinates of any celestial body and the latitude of the observer, and then gives an indication of azimuth, true north, and heading.

Distance or speed problems. Such measurement in polar regions presents no problems in aircraft. When ships operate in ice, however, the sensing element in the water may be adversely affected or damaged by the ice. A method of determining speed or distance that has proved successful has been to track an iceberg or other prominent feature, either visually or by radar. If the feature is stationary, as a grounded iceberg, the result is speed or distance over the ground.

At best, dead reckoning is difficult aboard a ship operating in the ice, not only because of the difficulty encountered in measuring course and speed, but also because neither of these may be constant for very long. Land ice which flows down to the sea and breaks off in the form of icebergs or ice islands is not usually a problem, and may even prove beneficial.

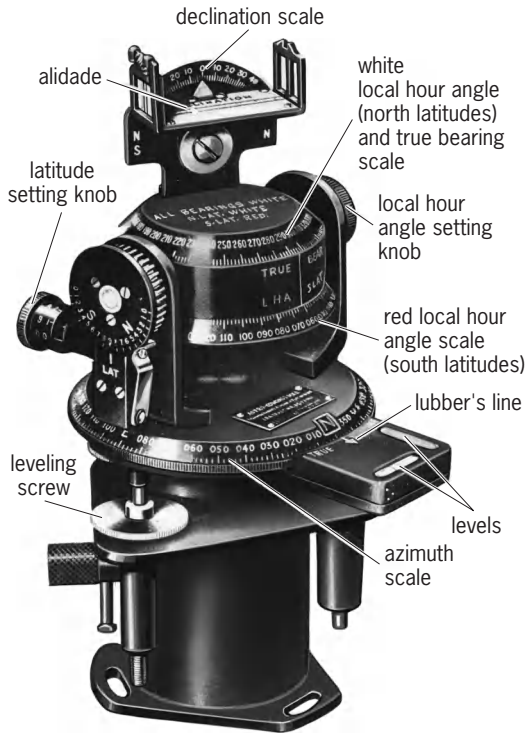

Fig. 2. An astrocompass.

Its use as an aid in the determination of speed has been mentioned. Individual pieces are usually so large that they move with deep-water currents, often in a direction differing from that of the sea ice, which moves mainly in the direction the wind blows. Thus an iceberg might clear a path in the desired direction of motion. One precaution is essential, however. It is dangerous to approach close to an iceberg because of possible underwater rams which might extend out for some distance from the berg and also because it is not uncommon for an iceberg to acquire unstable equilibrium, because of uneven melting, and capsize. *See* ICEBERG.

Ice formed at sea, called sea ice, is seldom an unbroken sheet over any very large area. The unequal pressure exerted by tides, currents, winds, and temperature changes produce stresses that break the ice and move different parts of it relative to each other, producing leads, long cracks that have opened wide enough to permit passage of a ship; polynyi, large areas, other than leads, relatively free from ice; or pressure ridges, ridges of ice piled up where two floes have come together under pressure. *See* SEA ICE.

A large field of floating pieces of ice which have drifted together is called a pack. If this is relatively loose (**Fig. 3**), a ship, skillfully handled, can negotiate it. If it is packed tightly under pressure, however, it is best avoided.

Successful negotiation of pack ice is the result of working with, not against, the pack, seeking out weak spots, ramming when appropriate, and retreating at other times, taking advantage of leads and polynyi, and avoiding heavy pressure that might stop the vessel or crush it. It is usually considered better to keep moving in the general direction desired than to risk complete stoppage with possible damage by staying exactly on the desired course. All this, plus a lack of detailed information on currents, with little opportunity to acquire it, makes marine dead reckoning difficult in polar regions.

Celestial navigation problems. Celestial navigation is of great importance in polar regions, sometimes providing the only means of determining position accurately, or establishing a directional reference.

In aircraft the use of a sextant is the same as it is in other latitudes. Aboard ship a marine sextant can be used when celestial bodies and a horizon are both available, an occurrence more frequent in most ocean areas at lower latitudes. Since the pack smooths the sea, resulting in very little rolling or pitching of a ship operating in ice, some marine navigators have found an artificial-horizon sextant of the type carried in aircraft useful. If the acceleration error is too great aboard ship, observations from a nearby ice cake may be possible. *See* SEXTANT.

When operating in lower latitudes, navigators generally avoid observation of bodies near the horizon because of the uncertainty of the refraction correction there. In polar regions, even though refraction is more uncertain, navigators often have no choice. Near the equinoxes the Sun may be the only body available for several weeks, and it remains close to the horizon. Under these conditions, the navigator observes the Sun and makes the most of the available information, usually with satisfactory results.

During the polar summer the Sun is often the only celestial body available. The Moon is above the horizon half the time, but is in a favorable position relative to the Sun during relatively few days each month. When only one body is available, it is observed at frequent intervals, perhaps hourly, and a series of running fixes is plotted.

Timing of celestial observations is less critical in polar regions than in lower latitudes because apparent motion of the bodies is nearly horizontal. A relatively large error in longitude, the effect of an error in time, is a relatively small error in miles in an area where the meridians are close together.

Sight reduction in polar regions is not attended by any unusual problems. The same methods in common use in lower latitudes are used in polar regions. Simplified tables might be prepared, but they are generally avoided in favor of familiar methods. *See* CELESTIAL NAVIGATION.

Fig. 3. Ship negotiating pack ice.

Impact of modern technology. Advances in technology have contributed significantly to the safety and reliability of navigation in polar regions. The availability of modern echo sounders has made possible a continuous plot of the bottom profile beneath a vessel while under way. Sonar, operating on the same principle as an echo sounder, indicates the presence of an underwater obstruction. Reliable inertial navigators having an error of less than 1 mi/h (0.5 m/s) are available to make aircraft navigation in polar regions almost routine, providing both positional and directional data. Ship inertial navigation systems are available and have made practicable submarine operations under sea ice. *See* ECHO SOUNDER; SONAR.

The U.S. Navy Navigation Satellite System (NNSS) provides means for intermittent position determination, with an error of a small fraction of a mile, for craft equipped with suitable instrumentation. Because of the need for knowledge of craft velocity vector to achieve such accuracy and intermittent availability, ships have benefitted from satellites more than aircraft. However, because all satellites of the system transmit similar signals on the same frequencies (150 and 400 MHz), both ships and aircraft encounter mutual interference between satellites when more than one is above the horizon at the same time, a frequent occurrence in high latitudes, where satellite orbits converge.

The NAVSTAR Global Positioning System (GPS) of the U.S. Department of Defense, when fully implemented operationally, will eliminate the limitations of the Navy Navigation Satellite System. Although all GPS satellites also operate on the same frequencies (1227 and 1575 MHz), signals are modulated with two codes, one of which provides for identification of and lock-on to the desired signal. The former Soviet Union advised member nations of the International Telecommunications Union (ITU) that it was developing a navigation satellite system similar to GPS, to be known as GLONASS. *See* SATELLITE NAVIGATION SYSTEMS.

The establishment of all eight Omega navigation system transmitters on an operational basis has provided an electronic means of continuously fixing the position of a craft anywhere in the world, including both polar regions.

The overall impact of modern technology has been to remove much of the uncertainty and hazard of polar navigation, which approaches the high reliability attained previously in lower latitudes. *See* NAVIGATION.

Alton B Moody

Bibliography. N. Bowditch, *American Practical Navigator*, Publ. 9, vol. 1, 1977; L. Logsdon, *The Navstar Global Positioning System*, 1992; A. B Moody, *Navigation Afloat—A Manual for the Seaman*, 1980.

Polarimetric analysis

A method of chemical analysis based on the optical activity of the substance being determined. Optically active materials are asymmetric; that is, their molecules or crystals have no plane or center of symmetry. These asymmetric molecules can occur in either of two forms, *d*- and *l*-, called optical isomers. Often a third, optically inactive form, called meso, also exists. Asymmetric substances possess the power of rotating the plane of polarization of plane-polarized light. Measurement of the extent of this rotation is called polarimetry. Polarimetry is applied to both organic and inorganic materials. *See* OPTICAL ACTIVITY.

The extent of the rotation depends on the character of the substance, the length of the light path, the temperature of the solution, the wavelength of the light which is being used, the solvent (if there is one), and the concentration of the substance. In most work, the yellow light of the D line of the sodium spectrum (589.3 nanometers) is used to determine the specific rotation, according to the equation below. Here α is the measured angle of rotation,

$$\text{Specific rotation} = [\alpha]_{\mathrm{D}}^{20} = \frac{\alpha}{l\rho}$$

l is the length of the column of liquid in decimeters, and ρ the density of the solution. In other words, the specific rotation is the rotation in degrees which this plane-polarized light of the sodium D line undergoes in passing through a 10-cm-long (4-in.) sample tube containing a solution of 1 g/ml concentration at 20°C (68°F).

In the **illustration**, light from the sodium lamp is polarized by the polarizer (a fixed nicol prism) before it passes through the cell containing the material being analyzed. After the light passes through the cell, it passes through the analyzer (another nicol prism) and then is detected by the eye or a photocell. A comparison of the angular orientation of the analyzer as measured on the scale with the cell empty and with the cell filled with solution serves to measure the rotation of the polarized light by the sample. This rotation may be either clockwise (+) or counterclockwise (−), depending on the substance in question.

Polarimetry may be used for either qualitative or quantitative analytical work. In qualitative applications, the presence of an optically active material is shown, and then a calculation of specific rotation often leads to the identification of the unknown. In quantitative work, the concentration of a given optically active material is determined either from a calibration curve of percentage of the constituent versus angular rotation or from the specific rotation, assuming the angular rotation to be a linear function of concentration. For this method of analysis to be useful, it is necessary that only one optically active material be present in solution.

Polarimetry is used in carbohydrate chemistry, especially in the analysis of sugar solutions. Polarimeters used for this work are specially designed and are called saccharimeters. Other materials often determined by polarimetry are tartaric acid, Rochelle salt (potassium sodium tartrate), various terpenes such as *d*- and *l*-pinene, many steroids, and other compounds of biological and biochemical importance. Since there is great difference between the biological

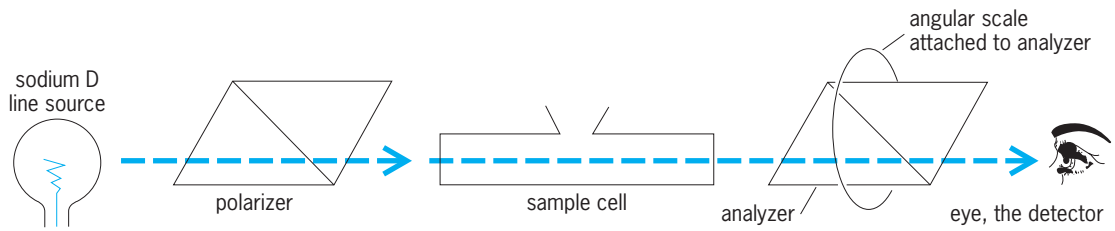

Simplified diagram of a polarimeter.

activities of the different optical forms of organic compounds, polarimetry is used in biochemical research to identify the molecular configurations.

Optical rotatory dispersion is the measurement of the specific rotation as a function of wavelength. To accomplish this, the sodium lamp is replaced by a monochromator, and a source of continuous radiation. A photocell circuit is substituted for the eye as a detector. In this way, the specific rotation may be determined in the ultraviolet or near-infrared, as well as in the visible region of the spectrum.

The information obtained by optical rotatory dispersion has shown that minor changes in configuration of a molecule have a marked effect on its dispersion properties. By using the properties of compounds of known configuration, it has been possible to determine the absolute configurations of many other molecules and to identify various isomers. Most of the applications have been to steroids, sugars, and other natural products, including amino acids, proteins, and polypeptides. *See* COTTON EFFECT; OPTICAL ROTATORY DISPERSION; POLARIZED LIGHT. Robert F. Goddu; James N. Little

Bibliography. E. Charney, *The Molecular Basis of Optical Activity*, 1979, reprint 1985; I. M. Koltoff and P. J. Elving, *Treatise on Analytical Chemistry*, pt. 1, vol. 6, 1965; S. F. Mason, *Molecular Optical Activity and the Chiral Discriminations*, 1982; G. Snatzke (ed.), *Optical Rotatory Dispersion and Circular Dichroism in Organic Chemistry*, 1976.

Polarimetry

The science of determining the polarization state of electromagnetic radiation (x-rays, light or radio waves). Radiation is said to be linearly polarized when the electric vector oscillates in only one plane. It is circularly polarized when the x-plane component of the electric vector oscillates 90° out of phase with the y-plane component. In experimental work, the polarization state is usually expressed in terms of the Stokes parameters I (the total intensity of the radiation), Q (the preference for light to be linearly polarized in the x plane or at 0° in the instrument reference system), U (the preference for linear polarization at 45°), and V (the preference for left-circular polarization). When $Q/I = 1$, for example, the radiation is 100% linearly polarized along an axis at 0°. Similarly, $V/I = -1$ means that the light is 100% right-circularly polarized.

To completely specify the polarization state, it is necessary to make six intensity measurements of the light passed by a quarter-wave retarder and a rotatable linear polarizer, such as a Polaroid or a Nicol prism.

The retarder converts circular light into linear light. It is placed upstream of the polarizer and is moved into the beam to determine Stokes V. Then I, Q, U, and V can be calculated from the six intensity measurements of the table. For fully polarized light, $I^2 = Q^2 + U^2 + V^2$. If the light is only partially polarized, the degree of polarization P is given by the equation below.

$$P = \sqrt{\frac{Q^2 + U^2 + V^2}{I^2}}$$

Polarization of radiation. An ensemble of atoms in a hot gas, such as a stellar atmosphere, usually radiates in no preferred state of polarization. Most starlight is unpolarized. However, atoms in the presence of a magnetic field align themselves at fixed, quantized angles to the field direction. Then the spectral lines they emit are circularly polarized when the magnetic field is parallel to the line of sight and linearly polarized when the field is perpendicular. In 1908, G. E. Hale used this property in solar spectral lines to prove that sunspots have magnetic fields of 1000–3000 gauss (0.1–0.3 tesla). For comparison, the Earth's magnetic field is 0.3 gauss (30 microteslas). Fields as high as 100 megagauss (10,000 T) have been measured in white dwarf stars. *See* SOLAR MAGNETIC FIELD; SUN; WHITE DWARF STAR; ZEEMAN EFFECT.

The light from sunspots is polarized because the magnetic fields impose some direction in the emitting gas. Other phenomena also remove isotropy and produce polarization. Sunlight scattered by electrons in the solar corona is partially linearly polarized because the radiation from the solar surface is much greater than the radiation into the surface (anisotropic radiation). Similarly, beams of electrons or protons will excite partially polarized light when they strike a gaseous cloud. Polarimetry can reveal the presence of the atomic beams.

Synchrotron emission and scattering are important polarizers of light. Lord Rayleigh showed that scattering by the barbell-shaped molecules of the Earth's atmosphere produces partial linear polarization of the blue sky. Similarly, needle-shaped grains in interstellar dust clouds polarize the starlight that passes through them whenever magnetic fields have

aligned the grains. By measuring the angle and degree of polarization in many stars, astronomers have been able to map the magnetic fields of the Milky Way Galaxy and many other galaxies. Galactic magnetic fields are typically only a few microgauss (a fraction of a nanotesla). *See* INTERSTELLAR MATTER; MILKY WAY GALAXY.

Synchrotron radiation is produced when electrons moving at nearly the speed of light spiral rapidly in a magnetic field. Because of the centripetal acceleration, the electrons emit linearly polarized light or radio waves. The Crab Nebula is a well-known source of highly polarized synchrotron radiation from the ultraviolet to the radio regime. The inferred field strength in the nebula is approximately 100 microgauss (10 nT). Radio galaxies, quasars, and active galactic nuclei also produce polarized synchrotron radiation. Sometimes the degree of polarization P is as high as 50%. *See* CRAB NEBULA; GALAXY, EXTERNAL; QUASAR; RADIO ASTRONOMY.

The *Wilkinson Microwave Anisotropy Probe* (*WMAP*) mission mapped the polarization of the cosmic microwave background radiation over the full sky. Polarization has been detected that arises from the anisotropic scattering of the photons off free electrons. This occurred both at the decoupling epoch and when the first stars reionized the universe. *WMAP* and other experiments are searching for a fainter polarization pattern that is predicted by inflation theory, arising not from scattering but from the gravitational waves produced during the sudden expansion of the universe. *See* COSMIC BACKGROUND RADIATION; GRAVITATIONAL RADIATION; INFLATIONARY UNIVERSE COSMOLOGY; WILKINSON MICROWAVE ANISOTROPY PROBE.

Magnetic fields penetrate every corner of the universe. They are generated by dynamos deep inside stars and planets and possibly in galaxies, but the dynamo mechanism is not well understood. And magnetic fields must have played some role in the evolution of the universe. These are good reasons to continue to develop more efficient and accurate polarimetry. *See* MAGNETOHYDRODYNAMICS.

Measurement of polarization. Electrooptical devices are rapidly replacing rotating polarizers and fixed retarders. To make daily maps of the weak magnetic fields on the Sun, H. W. Babcock invented the magnetograph, which consists of a spectrograph to isolate the atomic spectral line for study; a Pockels cell, an electrooptic crystal whose retardance depends on an applied voltage; a polarizing prism to isolate the polarization state passed by the retarder; a pair of photocells to detect the transmitted light; and a scanning mechanism to sweep the solar image across the spectrograph entrance slit. Two photocells are needed to simultaneously measure left- and right-circular polarization. *See* SPECTROGRAPH.

Since Babcock's magnetograph measured only circular polarization, it recorded only the line-of-sight, or longitudinal, component of the field. The sensitivity was about 5 gauss (500 μT). A magnetograph can be made sensitive to linear polarization, but the signal levels are about 100 times weaker for the inferred transverse fields than for longitudinal fields of comparable strength. In this case, longer measurement times are required to obtain useful signal-to-noise levels.

To improve signal-to-noise levels while measuring weak linear or circular polarization over a wide field, the spectrograph can be replaced with an optical filter having a narrow passband, and the photocells can be replaced with an array of photosensitive picture elements (pixels). In this manner, the narrow portion of the spectrum most sensitive to the effect (Zeeman effect) of magnetic fields can be isolated and recorded at many points simultaneously.

Most classical polarizing optical devices are effective over only narrow fields of view, sometimes as little as 1°. This characteristic severely limits the amount of light that can be analyzed at once. The introduction of polarizing dichroic films (Polaroids) was a great improvement over classical polarizing prisms (for example, the Nicol prism) because of their wide acceptance angle. It is more difficult to make a wide-angle retarder. The widest-angle device is the photoelastic modulator, invented by J. Kemp. The photoelastic modulator can be made to alternately pass right- and left-circular light over a field with a half-angle of 20°. The photoelastic modulator aperture is generally limited to about 5 cm (2 in.), but this dimension is about twice that of the aperture of classical prisms.

Some devices for detecting linear and circular polarization depend on scattering. To detect circular polarization in sunlight at the wavelength of the sodium line (589 nanometers), a cell filled with sodium vapor is sometimes used. Such a resonance cell is placed inside a laboratory magnet. Then the atoms preferentially absorb and reemit the circularly polarized component of the incident light. The resonance cell is very stable and accepts light from a very large angle.

In the infrared, glass containing tiny aligned silver needles can discriminate efficiently between the two senses of linear polarization because the needles scatter electromagnetic waves aligned with their long axis far more efficiently than they scatter nonaligned waves. David M. Rust

Bibliography. A. Adamson et al., *Astronomical Polarimetry: Current Status and Future Directions*, 2005; J. O. Stenflo, *Solar Magnetic Fields: Polarized Radiation Diagnostics*, 1994; J. Tinbergen, *Astronomical Polarimetry*, 1996, paper 2005; J. C. del Toro Iniesta, *Astrophysical Spectropolarimetry*, 2003.

Polaris

The star α Ursae Minoris, also known as the North Star or Pole Star. It is perhaps the best-known star in the northern sky. Its location only 1 degree of arc from the north celestial pole, the point where the Earth's rotation axis intersects the celestial sphere, has made it a very useful reference point for navigation. It may easily be found by following the line

joining the two bright stars at the end of the bowl of the Big Dipper. *See* URSA MAJOR; URSA MINOR.

Polaris (apparent magnitude 1.99) is a supergiant of spectral type F7 or F8, with an intrinsic brightness about 1500 times that of the Sun. Its effective temperature is approximately 6000 K (10,300°F). It is accompanied by a 9th-magnitude main-sequence star that has the same proper motion, at an angular separation of 18″. In addition, its spectrum shows evidence of another, much closer companion in an eccentric orbit with a period of 30 years. *See* SPECTRAL TYPE; SUPERGIANT STAR.

Polaris is a variable star, displaying slight changes in brightness with a period close to 4 days. The radial velocity also changes with the same period indicating that the diameter of the star is changing periodically. Polaris is a member of an important group of stars known as the Cepheid variables, which are typically high-luminosity stars that pulsate because of internal energy imbalances. Cepheids serve as an important tool for measuring distances in the Milky Way Galaxy and beyond because they are easily recognized by their characteristic patterns of light variation. Polaris is the brightest and nearest of the stars of this class, at a distance of approximately 130 parsecs from the Sun (4×10^{15} km or 2.5×10^{15} mi). However, it is atypical in that the amplitude of the variations is very small compared to other Cepheids and has decreased steadily over 100 years to the point where the pulsation of the star has virtually stopped. *See* CEPHEIDS; STAR; VARIABLE STAR. David W. Latham

Polarization of dielectrics

A vector quantity representing the electric dipole moment per unit volume of a dielectric material. *See* DIELECTRIC MATERIALS; DIPOLE MOMENT.

The polarization **P**, measured in coulombs/m^2, is related to the macroscopic electric parameters by the equation shown below, where **D** is the electric

$$\mathbf{P} = \mathbf{D} - \epsilon_0\mathbf{E} = \chi\epsilon_0\mathbf{E} = \epsilon_0(\epsilon_r - 1)\mathbf{E}$$

displacement, **E** is the electric field strength, ϵ_0 is the permittivity of free space (8.854×10^{-12} farad/m), ϵ_r is the relative permittivity, and χ is the electric susceptibility of the material. *See* ELECTRICAL UNITS AND STANDARDS.

Dielectric polarization arises from the electrical response of individual molecules of a medium and may be classified as electronic, atomic, orientation, and space-charge or interfacial polarization, according to the mechanism involved.

Electronic polarization represents the distortion of the electron distribution or motion about the nuclei in an electric field. This polarization occurs for all materials and is nearly independent of temperature and frequency up to about 10^{14} Hz for insulators.

Atomic polarization arises from the change in dipole moment accompanying the stretching of chemical bonds between unlike atoms in molecules. This mechanism contributes to polarization at frequencies below those of the vibrational modes of molecules (about 10^{12} to 10^{14} Hz). For a discussion of molecular vibrations *see* MOLECULAR STRUCTURE AND SPECTRA

Orientation polarization is caused by the partial alignment of polar molecules (molecules possessing permanent dipole moments) in an electric field. This mechanism results in a temperature-dependent component of polarization at lower frequencies.

Space-charge or interfacial polarization occurs when charge carriers are present which can migrate an appreciable distance through a dielectric but which become trapped or cannot discharge at an electrode. This process always results in a distortion of a macroscopic field and is important only at low frequencies. *See* ELECTRIC FIELD; ELECTRIC SUSCEPTIBILITY; PERMITTIVITY. Robert D. Waldron

Polarization of waves

The directional dependence of certain wave phenomena, at right angles to the propagation direction of the wave. In particular, ordinary light may be regarded as composed of two such asymmetrical components, referred to as its two states of linear polarization.

These two components are refracted differently by doubly refracting crystals, such as calcite, or Iceland spar. Each state of linear polarization is refracted according to its own separate refractive index. On a subsequent refraction by the same crystal, but now rotated through an angle θ about the direction of the beam, each component appears as a mixture of the original two polarization components, according to the proportions $\cos^2\theta : \sin^2\theta$. *See* BIREFRINGENCE; CRYSTAL OPTICS; REFRACTION OF WAVES.

The most familiar effect of light polarization occurs with polaroid material, such as in polaroid sunglasses. Polaroid absorbs one component of the light and allows the other component to pass through. For a perfect polaroid, the above-mentioned mixing effect could be seen by holding two such pieces against one another. For an angle θ between their polarizing directions, the second polaroid reduces the total intensity transmitted by a factor $\cos^2\theta$.

Nature of light. In the early nineteenth century, T. Young suggested that light polarization arises from transverse oscillations. In J. C. Maxwell's theory of light as electromagnetic waves, visible light—and also other types of electromagnetic radiation such as radio waves, microwaves, and x-rays (distinguished from visible light only by wavelength)—consists of electric and magnetic fields, each oscillating in directions perpendicular to the propagation direction, the electric and magnetic field vectors being perpendicular to each other (**Fig. 1**). The plane of polarization of the wave contains the electric vector (or magnetic vector; there is no general agreement which) and the propagation direction. *See* ELECTROMAGNETIC RADIATION; LIGHT; MAXWELL'S EQUATIONS.

Circular polarization. If the plane of polarization remains constant along the wave (as in the case of each light component in a doubly refracting medium), the

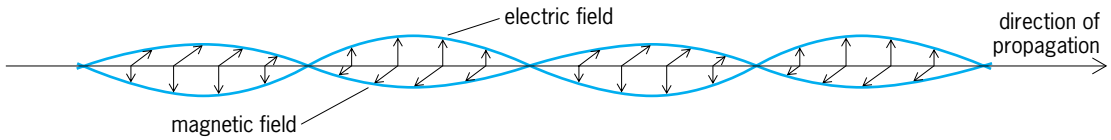

Fig. 1. Plane-polarized electromagnetic wave. (*After R. Penrose, The Emperor's New Mind, Oxford University Press, 1989*)

wave has linear (or plane) polarization. However, the plane of polarization can also rotate. If the rotation rate is constant, the intensity of the wave being also constant, a circularly polarized wave results (**Fig. 2**). These are of two types: right-handed and left-handed.

Any electromagnetic wave can be considered to be composed of monochromatic components, and each monochromatic component can be decomposed into a left-handed and a right-handed circularly polarized part. The states of linear polarization are each made up of equal magnitudes of the two circularly polarized parts, with differing phase relations to provide the different possible directions of plane polarization. Monochromatic waves composed of unequal magnitudes of the two circularly polarized parts are called elliptically polarized. This refers to the fact that the electric and magnetic vectors trace out ellipses in the plane perpendicular to the direction of motion.

Relation to spin. These descriptions refer to classical electromagnetic waves but apply equally well to the individual wave functions of single photons. Photons have quantum-mechanical spin, which refers to the angular momentum of the photon, necessarily about its direction of motion. A photon's spin has magnitude 1, in fundamental units. This spin can point along the direction of motion (positive helicity, right-handed spin) or opposite to it (negative helicity, left-handed spin), and this corresponds (depending on conventions used) to a classical electromagnetic wave of right- or left-handed circular polarization. *See* HELICITY (QUANTUM MECHANICS); PHOTON.

The wave functions of general, massless quantum particles of nonzero spin are similar, except that the magnitude of the spin can be some other integral or half-integral value, in fundamental units. Massless particles of half-integral spin do not give rise to classical fields, in the ordinary sense, since they are fermions (particles satisfying Pauli's exclusion principle). Particles of integral spin are the quanta of classical fields. In the half-integral case, there is the curious property that a rotation through 360° does not leave the wave function unaltered, but changes its sign. This occurs because this wave function is a spinor quantity, which indeed changes sign under a single complete rotation (about any axis). *See* BOSE-EINSTEIN STATISTICS; EXCLUSION PRINCIPLE; FERMI-DIRAC STATISTICS; SPIN (QUANTUM MECHANICS).

Classical waves. For a classical wave, corresponding to a massless quantum particle, with an integral spin, the symmetry of the field is the same as for the corresponding wave function. In particular, a gravitational wave (of Einstein's general relativity) has the same symmetry as the wave function of a graviton, a massless particle (as yet unobserved) of spin 2. For a gravitational wave traveling in one direction, the quantity describing the strength of the wave (namely the Riemann curvature tensor, describing tidal distortion) has the symmetry that a rotation through 180° about the direction of motion leaves it unaltered. If there were a classical field of spin 3, then for a wave of such a field propagating in one direction, a rotation through 120° about the propagation direction would leave the field unaltered. *See* GRAVITATIONAL RADIATION; GRAVITON.

In the case of zero spin, a rotation through any angle about the propagation direction will leave the field unaltered. This is the situation of a scalar wave, and it differs from all the other cases in that polarized

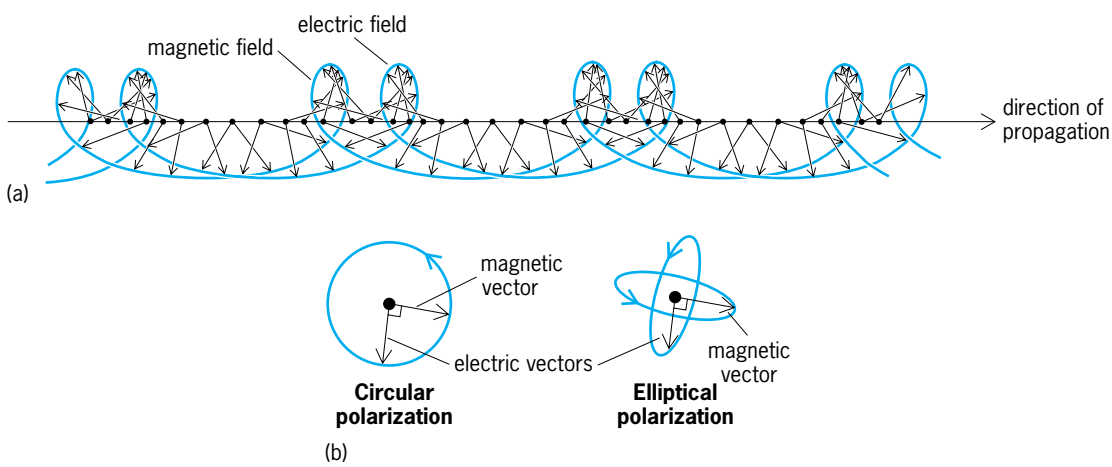

Fig. 2. Circularly polarized electromagnetic wave. (*a*) View along propagation direction. (*b*) Plane perpendicular to propagation direction for circular polarization and elliptical polarization (which is intermediate between plane polarization and circular polarization). (*After R. Penrose, The Emperor's New Mind, Oxford University Press, 1989*)

waves do not exist, and there is but a single state of polarization rather than two.

Orthogonal states of polarization. For a state of polarization of a general wave, changing the sign of the field (or wave function) leaves it unchanged. This applies to states of linear polarization as well as circular polarization. Thus, for an electromagnetic wave, a rotation about the propagation direction through 180° leaves a state of linear polarization unaltered. To pass to the opposite, or orthogonal state of linear polarization, it is necessary to rotate through half this angle, namely 90°. Thus, the two orthogonal states of linear polarization have polarization planes that are at right angles to each other. The plane of oscillation of the electric vector, for a particular state of linear polarization, would be at right angles to that for the orthogonal state of linear polarization.

For a gravitational wave, however, to pass to the orthogonal linear polarization, a rotation through 45° is needed. This can be understood in terms of the physical effect of a gravitational wave: inducing a tidal distortion on whatever matter it encounters. For a gravitational wave in a state of linear polarization, this tidal distortion can be pictured in terms of an ellipse which is perpendicular to the direction of propagation, this ellipse describing the distortion that is experienced by a circular pattern of particles in the path of the wave and parallel to its plane. For an oscillatory linearly polarized wave, the ellipse oscillates so that the axes of the ellipse remain constant, but the major and minor axes continually interchange with one another. The orthogonal state of linear polarization can be pictured in terms of an oscillating ellipse whose axes are at 45° to those of the original ellipse.

Elastic waves. Electromagnetic and gravitational waves both have the specific property that they are entirely transverse in character, which is a consequence of their speed of propagation being the absolute speed of relativity theory (the speed of light). This corresponds to the fact that their respective quanta, namely photons and gravitons, are massless particles. In the case of waves that travel at a smaller speed, as with fields whose quanta are massive rather than massless, there can be (unpolarized) longitudinal as well as transverse effects. Seismic waves traveling through the Earth's material, for example, can be transverse (polarized sideways oscillations) or longitudinal (unpolarized pressure waves). These waves travel at different speeds. Wave motion in other types of elastic material exhibits similar behavior. *See* SEISMOLOGY; SOUND.

Polarized and unpolarized light. In most situations encountered in practice, light (or gravitational waves) consists of an incoherent mixture of different polarization states, and is referred to as unpolarized. However, light reflected off a refracting surface (for example, glass or water) is polarized to some extent; that is, there is a certain preponderence of one state of linear polarization over the orthogonal possibility. Complete polarization occurs for a particular angle of incidence, known as the Brewster angle. Light from the sky is slightly polarized. Light from distant celestial objects is sometimes polarized (when produced by synchrotron radiation). This plane of polarization may become rotated en route to the Earth by an effect known as Faraday rotation, and this gives valuable astronomical information about the intervening magnetic fields. *See* FARADAY EFFECT; POLARIZED LIGHT; RADIO ASTRONOMY; REFLECTION OF ELECTROMAGNETIC RADIATION; SYNCHROTRON RADIATION; WAVE MOTION.

Roger Penrose

Polarized light

Light which has its electric vector oriented in a predictable fashion with respect to the propagation direction. In unpolarized light, the vector is oriented in a random, unpredictable fashion. Even in short time intervals, it appears to be oriented in all directions with equal probability. Most light sources seem to be partially polarized so that some fraction of the light is polarized and the remainder unpolarized. It is actually more difficult to produce a completely unpolarized beam of light than one which is completely polarized.

The polarization of light differs from its other properties in that human sense organs are essentially unable to detect the presence of polarization. The Polaroid Corporation with its polarizing sunglasses and camera filters has made millions of people conscious of phenomena associated with polarization. Light from a rainbow is completely linearly polarized; that is, the electric vector lies in a plane. The possessor of polarizing sunglasses discovers that with such glasses, the light from a section of the rainbow is extinguished.

According to all available theoretical and experimental evidence, it is the electric vector rather than the magnetic vector of a light wave that is responsible for all the effects of polarization and other observed phenomena associated with light. Therefore, the electric vector of a light wave, for all practical purposes, can be identified as the light vector. *See* CRYSTAL OPTICS; ELECTROMAGNETIC RADIATION; LIGHT; POLARIZATION OF WAVES.

Types of polarized light. Polarized light is classified according to the orientation of the electric vector. In linearly polarized light, the electric vector remains in a plane containing the propagation direction. For monochromatic light, the amplitude of the vector changes sinusoidally with time. In circularly polarized light, the tip of the electric vector describes a circular helix about the propagation direction. The amplitude of the vector is constant. The frequency of rotation is equal to the frequency of the light. In elliptically polarized light, the vector also rotates about the propagation direction, but the amplitude of the vector changes so that the projection of the vector on a plane at right angles to the propagation direction describes an ellipse.

These different types of polarized light can all be broken down into two linear components at right angles to each other. These are defined by Eqs. (1)

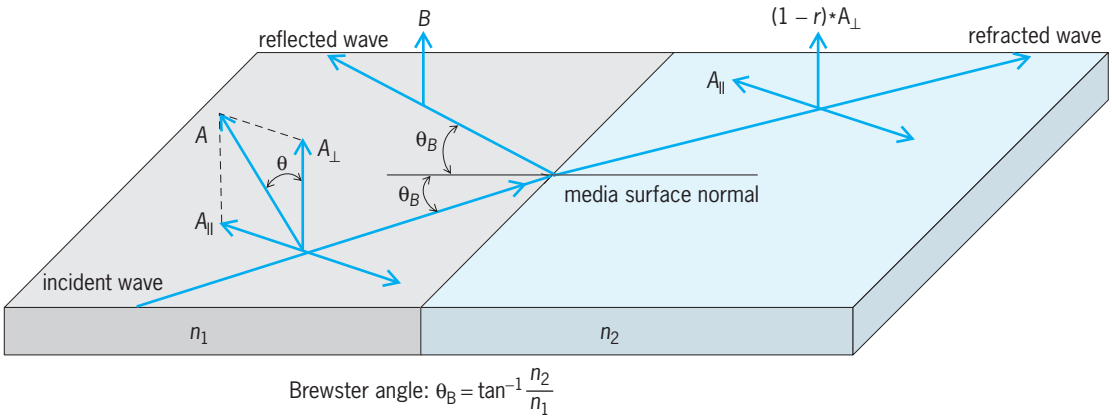

Fig. 1. Illustration of the law of Malus.

and (2), where A_x and A_y are the amplitudes, φ_x and

$$E_x = A_x \sin(\omega t + \varphi_x) \tag{1}$$

$$E_y = A_y \sin(\omega t + \varphi_y) \tag{2}$$

φ_y are the phases, ω is 2π times the frequency, and t is the time. For linearly polarized light Eqs. (3) hold. For circularly polarized light Eqs. (4) hold. For elliptically polarized light Eqs. (5) hold.

$$\varphi_x = \varphi_y \qquad A_x \neq A_y \tag{3}$$

$$\varphi_x = \varphi_y \pm \frac{\pi}{2} \qquad A_x = A_y \tag{4}$$

$$\varphi_x \neq \varphi_y \qquad A_x \neq A_y \tag{5}$$

In the last case, it is always possible to find a set of orthogonal axes inclined at an angle α to x and y along which the components will be E'_x and E'_y, such that Eqs. (6) hold.

$$\varphi'_x = \varphi'_y \pm \frac{\pi}{2} \qquad A'_x \neq A'_y \tag{6}$$

In this new system, the x' and y' amplitudes will be the major and minor axes a and b of the ellipse described by the light vector and α will be the angle of orientation of the ellipse axes with respect to the original coordinate system. The relationships between the different quantities can be written as in Eqs. (7) and (8). The terms are defined by Eqs. (9)–(12).

$$\tan 2\alpha = \tan 2\gamma \cos\varphi \tag{7}$$

$$\sin 2\beta = \sin 2\gamma \sin\varphi \tag{8}$$

$$\tan\gamma = \frac{A_y}{A_x} \tag{9}$$

$$\varphi = \varphi_x - \varphi_y \tag{10}$$

$$\tan\beta = \pm\frac{b}{a} \tag{11}$$

$$A_x^2 + A_y^2 = a^2 + b^2 \tag{12}$$

These same types of polarized light can also be broken down into right and left circular components or into two orthogonal elliptical components. These different vector bases are useful in different physical situations.

One of the simplest ways of producing linearly polarized light is by reflection from a dielectric surface. At a particular angle of incidence, the reflectivity for light whose electric vector is in the plane of incidence becomes zero. The reflected light is thus linearly polarized at right angles to the plane of incidence. This fact was discovered by E. Malus in 1808. Brewster's law shows that at the polarizing angle the refracted ray makes an angle of 90° with the reflected ray. By combining this relationship with Snell's law of refraction, it is found that Eq. (13) holds, where

$$\tan\theta_B = \frac{n_2}{n_1} \tag{13}$$

θ_B is the Brewster angle of incidence and n_1 and n_2 are the refractive indices of the two interfaced media (**Fig. 1**). This provides a simple way of measuring refractive indices. *See* REFRACTION OF WAVES.

Law of Malus. If linearly polarized light is incident on a dielectric surface at Brewster's angle (the polarizing angle), then the reflectivity of the surface will depend on the angle between the incident electric vector and the plane of incidence. When the vector is in the plane of incidence, the reflectivity will vanish. To compute the complete relationship, the incident light vector **A** is broken into components, one vibrating in the plane of incidence and one at right angles to the plane, as in Eqs. (14) and (15),

$$A_\parallel = A\sin\theta \tag{14}$$

$$A_\perp = A\cos\theta \tag{15}$$

where θ is the angle between the light vector and a plane perpendicular to the plane of incidence. Since the component in the plane of incidence is not reflected, the reflected ray can be written as Eq. (16),

$$B = Ar\cos\theta \tag{16}$$

where r is the reflectivity at Brewster's angle. The intensity is given by Eq. (17). This is the mathematical

$$I = B^2 = A^2 r^2 \cos^2\theta \tag{17}$$

statement of the law of Malus (Fig. 1).

Linear polarizing devices. The angle θ can be considered as the angle between the transmitting axes of a pair of linear polarizers. When the polarizers are parallel, they are transparent. When they are crossed, the combination is opaque. The first polarizers were glass plates inclined so that the incident light was at Brewster's angle. Such polarizers are quite inefficient since only a small percentage of the incident light is reflected as polarized light. More efficient polarizers can be constructed.

Dichroic crystals. Certain natural materials absorb linearly polarized light of one vibration direction much more strongly than light vibrating at right angles. Such materials are termed dichroic. For a description of them. *See* DICHROISM.

Tourmaline is one of the best-known dichroic crystals, and tourmaline plates were used as polarizers for many years. A pair was usually mounted in what were known as tourmaline tongs.

Birefringent crystals. Other natural materials exist in which the velocity of light depends on the vibration direction. These materials are called birefringent. The simplest of these structures are crystals in which there is one direction of propagation for which the light velocity is independent of its state of polarization. These are termed uniaxial crystals, and the propagation direction mentioned is called the optic axis. For all other propagation directions, the electric vector can be split into two components, one lying in a plane containing the optic axis and the other at right angles. The light velocity or refractive index for these two waves is different. *See* BIREFRINGENCE.

One of the best-known of these birefringent crystals is transparent calcite (Iceland spar), and a series of polarizers have been made from this substance. W. Nicol (1829) invented the Nicol prism, which is made of two pieces of calcite cemented together as in **Fig. 2**. The cement is Canada balsam, in which the wave velocity is intermediate between the velocity in calcite for the fast and the slow ray. The angle at which the light strikes the boundary is such that for one ray the angle of incidence is greater than the critical angle for total reflection. Thus the rhomb is transparent for only one polarization direction.

Canada balsam is not completely transparent in the ultraviolet at wavelengths shorter than 400 nanometers. Furthermore, large pieces of calcite material are exceedingly rare. A series of polarizers has been made using quartz, which is transparent in the ultraviolet and which is more commonly available in large pieces. Because of the small difference between the refractive indices of quartz and Canada balsam, a Nicol prism of quartz would be tremendously long for a given linear aperture.

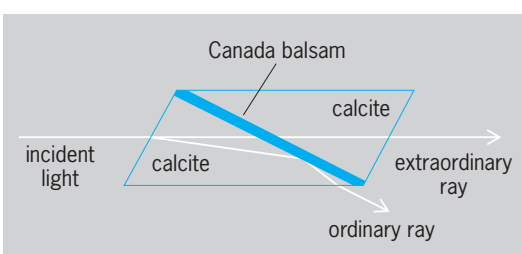

Fig. 2. Nicol prism. The ray for which Snell's law holds is called the ordinary ray.

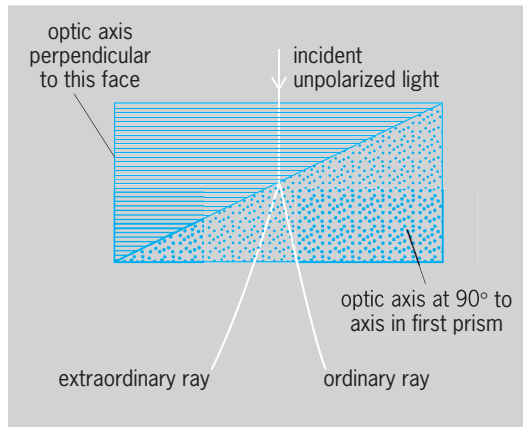

Fig. 3. Wollaston prism.

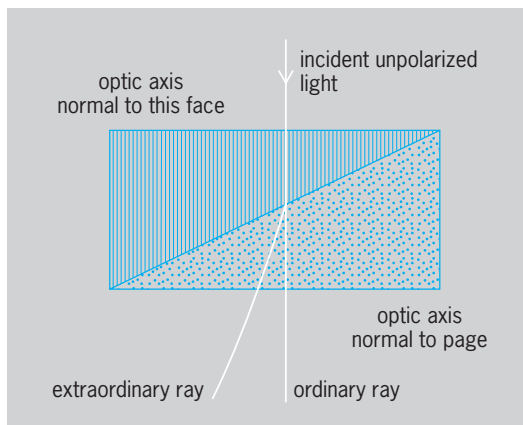

Fig. 4. Rochon prism.

A different type of polarizer, made of quartz, was invented by W. H. Wollaston and is shown in **Fig. 3**. Here the vibration directions are different in the two pieces so that the two rays are deviated as they pass through the material. The incoming light beam is thus separated into two oppositely linearly polarized beams which have an angular separation between them. By using appropriate optical stops (obstacles that restrict light rays) in the system, it is possible to select either beam.

In the Wollaston prism, both beams are deviated; since the quartz produces dispersion, each beam is spread into a spectrum. This is not the case of a prism which was invented by A. Rochon. Here the two pieces are arranged as in **Fig. 4**. One beam proceeds undeviated through the device and is thus achromatic. *See* OPTICAL PRISM.

Sheet polarizers. A third mechanism for obtaining polarized light is the Polaroid sheet polarizer invented by E. H. Land. Sheet polarizers fall into three types. The first is a microcrystalline polarizer in which small crystals of a dichroic material are oriented parallel to

each other in a plastic medium. Typical microcrystals, such as needle-shaped quinine iodosulfate, are embedded in a viscous plastic and are oriented by extruding the material through a slit.

The second type depends for its dichroism on a property of an iodine-in-water solution. The iodine appears to form a linear high polymer. If the iodine is put on a transparent oriented sheet of material such as polyvinyl alcohol (PVA), the iodine chains apparently line themselves parallel to the PVA molecules and the resulting dyed sheet is strongly dichroic. A third type of sheet polarizer depends for its dichroism directly on the molecules of the plastic itself. This plastic consists of oriented polyvinylene. Because these polarizers are commercially available and can be obtained in large sheets, many experiments involving polarized light have been performed which would have been quite difficult with the reflection polarizers or the birefringent crystal polarizer.

Characteristics. There are several characteristics of linear polarizers which are of interest to the experimenter. First is the transmission for light polarized parallel and perpendicular to the axis of the polarizer; second is the angular field; and third is the linear aperture. A typical sheet polarizer has a transmittance of 48% for light parallel to the axis and 2×10^{-4}% for light perpendicular to the axis at a wavelength of 550 nm. The angular field is 60°, and sheets can be many feet in diameter. The transmittance perpendicular to the axis varies over the angular field.

The Nicol prism has transmittance similar to that of the Polaroid sheeting but a much reduced linear and angular aperture.

Polarization by scattering. When an unpolarized light beam is scattered by molecules or small particles, the light observed at right angles to the original beam is polarized. The light vector in the original beam can be considered as driving the equivalent oscillators (nuclei and electrons) in the molecules. There is no longitudinal component in the original light beam. Accordingly, the scattered light observed at right angles to the beam can only be polarized with the electric vector at right angles to the propagation direction of the original beam. In most situations, the scattered light is only partially polarized because of multiple scattering. The best-known example of polarization by scattering is the light of the north sky. The percentage polarization can be quite high in clean country air. A technique was invented for using measurements of sky polarization to determine the position of the Sun when it is below the horizon. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

Production of polarized light. Linear polarizers have already been discussed. Circularly and elliptically polarized light are normally produced by combining a linear polarizer with a wave plate. A Fresnel rhomb can be used to produce circularly polarized light.

Wave plate. A plate of material which is linearly birefringent is called a wave plate or retardation sheet. Wave plates have a pair of orthogonal axes which are designated fast and slow. Polarized light with its electric vector parallel to the fast axis travels faster than light polarized parallel to the slow axis. The thickness of the material can be chosen so that for light traversing the plate, there is a definite phase shift between the fast component and the slow component. A plate with a 90° phase shift is termed a quarter-wave plate. The retardation in waves is given by Eq. (18), where $n_s - n_f$ is the birefringence; n_s is

$$\delta = \frac{(n_s - n_f)d}{\lambda} \tag{18}$$

the slow index at wavelength λ; n_f is the fast index; and d is the plate thickness.

Wave plates can be made by preparing X-cut sections of quartz, calcite, or other birefringent crystals. For retardations of less than a few waves, it is easiest to use sheets of oriented plastics or of split mica. A quarter-wave plate for the visible or infrared is easy to fabricate from mica. The plastic wrappers from many American cigarette packages seem to have almost exactly a half-wave retardation from green light. Since mica is not transparent in the ultraviolet, a small retardation in this region is most easily achieved by crossing two quartz plates which differ by the requisite thickness.

Linearly polarized light incident normally on a quarter-wave plate and oriented at 45° to the fast axis can be split into two equal components parallel to the fast and slow axes. These can be represented, before passing through the plate, by Eqs. (19) and (20), where x and y are parallel to the wave-plate

$$E_x = A_x \sin(\omega t + \varphi_x) \tag{19}$$

$$E_y = A_x \sin(\omega t + \varphi_x) \tag{20}$$

axes. After passing through the plate, the two components can be written as Eqs. (21) and (22), where

$$E_x = A_x \sin\left(\omega t + \varphi_x + \frac{\pi}{2}\right) \tag{21}$$

$$E_y = A_x \sin(\omega t + \varphi_x) \tag{22}$$

E_x is now advanced one quarter-wave with respect to E_y.

It is possible to visualize the behavior of the light by studying the sketches in **Fig. 5**, which show the projection on a plane $z = 0$ at various times. It is apparent that the light vector is of constant amplitude, and that the projection on a plane normal to the propagation direction is a circle. If the linearly polarized light is oriented at −45° to the fast axis, the light vector will revolve in the opposite direction. Thus it is possible with a quarter-wave plate and a linear polarizer to make either right or left circularly polarized light. If the linearly polarized light is at an angle other than 45° to the fastest axis, the transmitted radiation will be elliptically polarized. When circularly polarized light is incident on a quarter-wave plate, the transmitted light is linearly polarized at an angle of 45° to the wave-plate axes. This polarization is independent of the orientation of the wave-plate axis. For elliptically polarized light, the behavior of the quarter-wave plate is much more complicated. However, as was mentioned earlier, the elliptically polarized light can be considered as composed of two

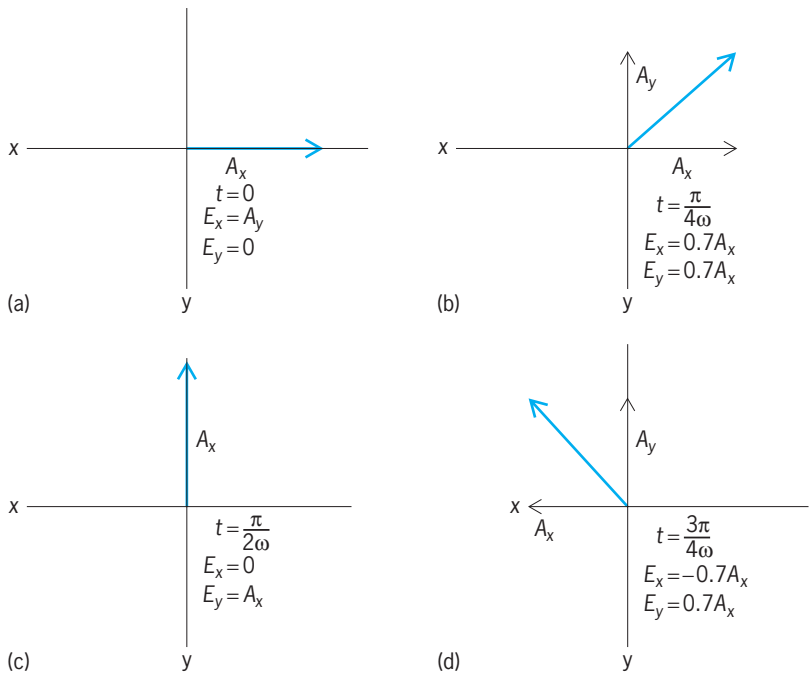

Fig. 5. Projection of light vector of constant amplitude on plane $z = 0$ for circularly polarized light. (*a*) $t = 0$. (*b*) $t = \pi/4\omega$. (*c*) $t = \pi/2\omega$. (*d*) $t = 3\pi/4\omega$.

linear components parallel to the major and minor axes of the ellipse and with a quarter-wave phase difference between them. If the quarter-wave plate is oriented parallel to the axes of the ellipse, the two transmitted components will either have zero phase difference or a 180° phase difference and will be linearly polarized. At other angles, the transmitted light will still be elliptically polarized, but with different major and minor axes. Similar treatment for a half-wave plate shows that linearly polarized light oriented at an angle of θ to the fast axis is transmitted as linearly polarized light oriented at an angle $-\theta$ to the fast axis.

Wave plates all possess a different retardation at each wavelength. This appears immediately from Eq. (18). It is conceivable that a substance could have dispersion of birefringence, such as to make the retardation of a plate independent of wavelength. However, no material having such a characteristic has as yet been found.

Fresnel rhomb. A quarter-wave retardation can be provided achromatically by the Fresnel rhomb. This device depends on the phase shift which occurs at total internal reflection. When linearly polarized light is totally internally reflected, it experiences a phase shift which depends on the angle of reflection, the refractive index of the material, and the orientation of the plane of polarization. Light polarized in the plane of incidence experiences a phase shift which is different from that of light polarized at right angles to the plane of incidence. Light polarized at an intermediate angle can be split into two components, parallel and at right angles to the plane of incidence, and the two components mathematically combined after reflection.

The phase shifts can be written Eqs. (23) and (24),

$$\tan\frac{\phi_\parallel}{2} = \frac{\sqrt{n^2\sin^2 i - 1}}{\cos i} \tag{23}$$

$$\tan\frac{\varphi_\perp}{2} = \frac{\sqrt{n^2\sin^2 i - 1}}{n\cos i} \tag{24}$$

where $\varphi_\parallel$ is the phase shift parallel to the plane of incidence; $\varphi_\perp$ is the phase shift at right angles to the plane of incidence; i is the angle of incidence on the totally reflecting internal surface; and n is the refractive index. The difference $\varphi_\parallel - \varphi_\perp$ reaches a value of about $\pi/4$ at an angle of 52° for $n = 1.50$. Two such reflections give a retardation of $\pi/2$. The Fresnel rhomb shown in **Fig. 6** is cut so that the incident light is reflected twice at 52°. Accordingly, light polarized at 45° to the principal plane will be split into two equal components which will be shifted a quarter-wave with respect to each other, and the transmitted light will be circularly polarized. Nearly achromatic wave plates can be made by using a series of wave plates in series with their axes oriented at different specific angles with respect to a coordinate system.

Analyzing devices. Polarized light is one of the most useful tools for studying the characteristics of materials. The absorption constant and refractive index of a metal can be calculated by measuring the effect of the metal on polarized light reflected from its surface. *See* REFLECTION OF ELECTROMAGNETIC RADIATION.

The analysis of polarized light can be performed with a variety of different devices. If the light is linearly polarized, it can be extinguished by a linear polarizer and the direction of polarization of the light determined directly from the orientation of the polarizer. If the light is elliptically polarized, it can be analyzed with the combination of a quarter-wave plate and a linear polarizer. Any such combination of polarizer and analyzer is called a polariscope. As explained previously, a quarter-wave plate oriented parallel to one of the axes of the ellipse will transform elliptically polarized light to linearly polarized light. Accordingly, the quarter-wave plate is rotated until the beam is extinguished by the linear polarizer. At this point, the orientation of the quarter-wave plate gives the orientation of the ellipse and the orientation of the polarizer gives the ratio of the major to

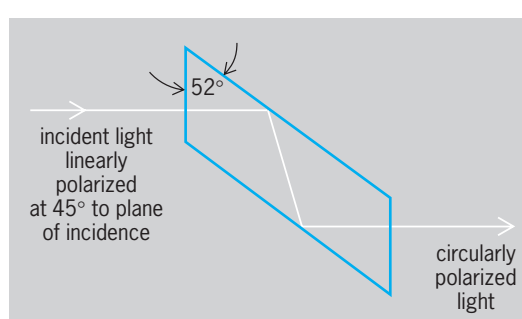

Fig. 6. Fresnel rhomb.

minor axis. Knowledge of the origin of the elliptically polarized light usually gives the orientation of the components which produced it, and from these various items, the phase shifts and attenuations produced by the experiment can be deduced.

One of the best-known tools for working with polarized light is the Babinet compensator. This device is normally made of two quartz prisms put together in a rhomb. One prism is cut with the optic axis in the plane of incidence of the prism, and the other with the optic axis perpendicular to the plane of incidence. The retardation is a function of distance along the rhomb; it will be zero at the center, varying to positive and negative values in opposite directions along the rhomb. It can be used to cancel the known or unknown retardation of any wave plate.

Retardation theory. It is difficult to see intuitively the effect of a series of retardation plates or even of a single plate of general retardation δ on light which is normally incident on the plate and which is polarized in a general fashion. This problem is most easily solved algebraically. The single-wave plate is assumed to be oriented normal to the direction of propagation of the light, which is taken to be the z direction of a set of cartesian coordinates. Its fast axis is at an angle α to the x axis. The incident light can be represented by Eqs. (25) and (26). A first step is to break

$$E_x = A_x \sin(\omega t + \varphi_x) \tag{25}$$

$$E_y = A_y \sin(\omega t + \varphi_y) \tag{26}$$

the light up into components $E_{x'}$ and $E_{y'}$ parallel to the axes of the plate. It is possible to write Eqs. (27) and (28). These components can also be written as Eqs. (29) and (30). After passing through the plate, the components become Eqs. (31) and (32).

$$E_{x'} = E_x \cos\alpha - E_y \sin\alpha \tag{27}$$

$$E_{y'} = E_x \sin\alpha + E_y \cos\alpha \tag{28}$$

$$E_{x'} = A_{x'} \sin(\omega t + \varphi_{x'}) \tag{29}$$

$$E_{y'} = A_{y'} \sin(\omega t + \varphi_{y'}) \tag{30}$$

$$E_{x''} = A_{x'} \sin(\omega t + \varphi_{x'} + \delta) \tag{31}$$

$$E_{y''} = A_{y'} \sin(\omega t + \varphi_{y'}) \tag{32}$$

In general, it is of interest to compare the output with the input. The transmitted light is thus broken down into components along the original axes. This results in Eqs. (33) and (34).

$$E_{x'''} = E_{x''} \cos\alpha + E_{y''} \sin\alpha \tag{33}$$

$$E_{y'''} = -E_{x''} \sin\alpha + E_{y''} \cos\alpha \tag{34}$$

With this set of equations, it is possible to compute the effect of a wave plate on any form of polarized light.

Jones calculus. Equations (33) and (34) still become overwhelmingly complicated in any system involving several optical elements. Various methods have been developed to simplify the problem and to make possible some generalizations about systems of elements. One of the most straightforward, proposed by R. C. Jones, involves reducing Eqs. (33) and (34) to matrix form. The Jones calculus for optical systems involves the polarized electric components of the light vector and is distinguished from other methods in that it takes cognizance of the absolute phase of the light wave.

The Jones calculus writes the light vector in the complex form as in Eqs. (35) and (36). Matrix operators are developed for different optical elements. From Eqs. (25)–(34), the operator for a wave plate can be derived directly. *See* MATRIX THEORY.

$$\begin{aligned} E_x &= A_x e^{i(\omega t + \varphi_x)} \\ E_y &= A_x e^{i(\varphi t + \varphi_y)} \end{aligned} \tag{35}$$

$$E = \begin{vmatrix} A_x e^{i\varphi_x} \\ A_y e^{i\varphi_y} \end{vmatrix} e^{iwt} \tag{36}$$

The Jones calculus is ordinarily used in a normalized form which simplifies the matrices to a considerable extent. In this form, the terms involving the actual amplitude and absolute phase of the vectors and operators are factored out of the expressions. The intensity of the light beam is reduced to unity in the normalized vector so that Eq. (37) holds. Under

$$A_x{}^2 + A_y{}^2 = 1 \tag{37}$$

this arrangement, the matrices for various types of operations can be written as Eq. (38). This is the

$$G(\delta) = \begin{vmatrix} e^{i(\delta/2)} & 0 \\ 0 & e^{-i(\delta/2)} \end{vmatrix} \tag{38}$$

operator for a wave plate of retardation δ and with axes along x and y. Equation (39) gives the operator

$$S(\alpha) = \begin{vmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{vmatrix} \tag{39}$$

for a rotator which rotates linearly polarized light through an angle α. Equation (40) gives the operator

$$P_b = \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} \tag{40}$$

for a perfect linear polarizer parallel to the x axis. A wave plate at an angle α can be represented by Eq. (41). A series of optical elements can be represented by the product of a series of matrices. This simplifies enormously the task of computing the effect of many elements. It is also possible with the Jones calculus to derive a series of general theorems concerning combinations of optical elements. Jones

$$G(\delta, \alpha) = S(\alpha)G(\delta)S(-\alpha) \tag{41}$$

has described three of these, all of which apply only for monochromatic light.

1. An optical system consisting of any number of retardation plates and rotators is optically equivalent to a system containing only two elements, a retardation plate and a rotator.

2. An optical system containing any number of partial polarizers and rotators is optically equivalent to a system containing only two elements—one a partial polarizer and the other a rotator.

3. An optical system containing any number of retardation plates, partial polarizers, and rotators is optically equivalent to a system containing four elements—two retardation plates, one partial polarizer, and one rotator.

As an example of the power of the calculus, a rather specific theorem can be proved. A rotator of any given angle α can be formed by a sequence of three retardation plates, a quarter-wave plate, a retardation plate at 45° to the quarter-wave plate, and a second quarter-wave plate crossed with the first, as in Eq. (42), where β is the angle between the axis

$$S(\alpha) = S(\beta)G\left(-\frac{\pi}{2}\right)S(-\beta)S\left(\beta+\frac{\pi}{4}\right) \cdot G(\delta)S\left(-\beta-\frac{\pi}{4}\right)S(\beta)G\left(\frac{\pi}{2}\right)S(-\beta) \quad (42)$$

of the first quarter-wave plate and the x axis, and δ is the retardation of the plate in the middle of the sandwich.

The first simplification arises from the fact that the axis rotations can be done in any order. This reduces Eq. (43) to Eq. (44). Now Eqs. (45) and (46) hold. When the multiplication is carried through, Eq. (47) is obtained.

$$S\left(\beta+\frac{\pi}{4}\right) = S(\beta)S\left(\frac{\pi}{4}\right) = S\left(\frac{\pi}{4}\right)S(\beta) \quad (43)$$

$$S(\alpha) = S(\beta)G\left(-\frac{\pi}{2}\right)S\left(\frac{\pi}{4}\right) \cdot G(\delta)S\left(-\frac{\pi}{4}\right)G\left(\frac{\pi}{2}\right) \quad (44)$$

$$S\left(-\frac{\pi}{4}\right)G\left(\frac{\pi}{2}\right) = \frac{1}{\sqrt{2}}\begin{vmatrix} 1 & -i \\ -i & 1 \end{vmatrix} \quad (45)$$

$$G\left(-\frac{\pi}{2}\right)S\left(\frac{\pi}{4}\right) = -\frac{1}{\sqrt{2}}\begin{vmatrix} 1 & i \\ i & 1 \end{vmatrix} \quad (46)$$

$$S(\alpha) = S(\beta) - \frac{1}{2}\begin{vmatrix} e^{i\delta/2}+e^{-\delta/2} & -ie^{i\delta/2}+ie^{-i\delta/2} \\ ie^{i\delta/2}-ie^{-i\delta/2} & e^{i\delta/2}+e^{-i\delta/2} \end{vmatrix} S(-\beta)$$

$$S(\alpha) = S(\beta) - \frac{1}{2}\begin{vmatrix} e^{i\delta/2}+e^{-\delta/2} & -ie^{i\delta/2}+ie^{-i\delta/2} \\ ie^{i\delta/2}-ie^{-i\delta/2} & e^{i\delta/2}+e^{-i\delta/2} \end{vmatrix} S(-\beta) \quad (47)$$

The rotation angle is therefore equal to one-half the phase angle of the retardation. This combination is a true rotator in that the rotation is independent of the azimuth angle of the incident polarized light.

A variable rotator can be made by using a Soleil compensator for the central element. This consists of two quartz wedges joined to form a plane parallel quartz plate.

Mueller matrices. In the Jones calculus, the intensity of the light passing through the system must be obtained by calculation from the components of the light vector. A second calculus is frequently used in which the light vector is split into four components. This also uses matrix operators which are termed Mueller matrices. In this calculus, the intensity I of the light is one component of the vector and thus is automatically calculated. The other components of the vector are given by Eqs. (48)–(50). The matrix

$$M = A_x^2 - A_y^2 \quad (48)$$

$$C = 2A_xA_y\cos(\varphi_x - \varphi_y) \quad (49)$$

$$S = 2A_xA_y\sin(\varphi_x - \varphi_y) \quad (50)$$

of a perfect polarizer parallel to the x axis can be written as Eq. (51).

$$P = \frac{1}{2}\begin{vmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix} \quad (51)$$

The calculus can treat unpolarized light directly. Such a light vector is given by Eq. (52). The vector for

$$\begin{vmatrix} I \\ M \\ C \\ S \end{vmatrix} = \begin{vmatrix} 1 \\ 0 \\ 0 \\ 0 \end{vmatrix} \quad (52)$$

light polarized parallel to the x axis is written as expression (53). In the same manner as in the Jones calculus, matrices can be derived for retardation plates, rotators, and partial polarizers. This calculus can also be used to derive various general theorems about various optical systems. *See* FARADAY EFFECT; INTERFERENCE OF WAVES; OPTICAL ACTIVITY; OPTICAL ROTATORY DISPERSION.

$$\begin{vmatrix} 1 \\ 1 \\ 0 \\ 0 \end{vmatrix} \quad (53)$$

Bruce H. Billings; Hong Hua

Bibliography. R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light*, 1977, paper 1987; B. Billings (ed.), *Selected Papers on Applications of Polarized Light*, 1992; M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; E. Collett, *Polarized Light*, 1993; F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976; D. S. Kliger, J. W. Lewis, and C. E. Randall, *Polarized Light in Optics and Spectroscopy*, 1990.

Polarized light microscope

A microscope that utilizes polarized light to form a highly magnified image of an object. Polarizing microscopes play an important role in crystallography, petrography, microchemistry, and biology. Although all light microscopes compare poorly with electron microscopes with respect to image resolution, polarized light microscopes have the unique ability to deliver information about the submicroscopic structure of the objects being examined. They also have the advantage of being relatively nondestructive, and may be used safely with living cells. Polarized light interactions with electromagnetically anisotropic structures, down to atomic dimensions, can be measured by polarized light microscopy. The sensitivity of polarized light microscopy as well as its importance to biology have been enhanced by the use of video technology. *See* POLARIZED LIGHT.

Optical anisotropy. Conventional light microscopes develop image contrast primarily as a result of light absorption in the specimen. Phase-contrast and interference microscopes develop contrast for nonabsorbing objects whose optical thickness differs from that of their surroundings. In turn, polarizing microscopes develop contrast for objects whose optical thickness has more than one value, depending on the orientation of the electric vector of the light relative to the axes of the object. *See* INTERFERENCE MICROSCOPE; PHASE-CONTRAST MICROSCOPE.

Just as materials such as wood have different mechanical properties in different directions, optically anisotropic materials have different electrical properties in different directions. A consistent directional difference in the alignment of chemical bonds, ions, or molecules results in materials that have different dielectric constants for electric fields applied in different directions. Because the velocity of light in a material is proportional to the square root of the dielectric constant in the direction of the electric-field displacement (electric-vector direction), the light velocity—and hence refractive index—in these materials differs for light polarized in different directions. For transparent materials, the refractive index is equal to the square root of the dielectric constant at optical frequencies.

Crystals, oriented polymers, and complex macromolecules tend to be optically anisotropic and can have up to three different refractive indices, one for each of the three orthogonal directions. The majority of crystals have two or three refractive indices; only regular cubic crystals and amorphous materials show the same refractive index in all directions. Light traveling through anisotropic materials experiences at most two refractive indices because its electric-vector direction is always normal to the direction of propagation. Hence, optically anisotropic materials are often referred to as birefringent, that is, having two refractive indices. The birefringence of such materials is a manifestation of their unique internal electrical structure and is independent of their environment. This is known as intrinsic birefringence. Stress birefringence and strain birefringence demonstrate the electric displacements, in initially isotropic solids, that result from elastic and plastic deformation, respectively. Stress birefringence is reversible, disappearing when the stress is removed, whereas strain birefringence is permanent. Stress and strain birefringences are also intrinsic in that they are strictly the result of internal electrical anisotropy.

Form birefringence, also called structural birefringence, occurs when asymmetric particles of one refractive index are aligned more or less parallel to each other in a surrounding medium of different refractive index. Form birefringence is dependent on the interparticle spacing, in at least one direction, being much less than the wavelength of light. Typical biological examples of form birefringence are the bundles of microtubules and microfilaments of the cytoskeleton. Flow birefringence is a special case of form birefringence in which submicroscopic particles in a fluid matrix are aligned by shear during flow. A fluid suspension of submicroscopic rodlets, such as tobacco mosaic virus particles in water, shows strong flow birefringence when squirted out of a pipette.

Form birefringence requires different refractive indices for the oriented structures and the intervening medium. The magnitude, but not the sign, of birefringence is a function of this difference. Consequently, form birefringence is indicated when changing the refractive index of the medium changes the birefringence. *See* BIREFRINGENCE; REFRACTION OF WAVES.

Polarizing microscopes. A polarizing microscope differs from a conventional light microscope in a number of ways (see **illus.**).

Polarizing devices. A polarizing microscope has a pair of polars (polarizing devices) in the optical train. The first polar (polarizer) defines the initial plane of polarization for light entering the microscope and is located between the illuminator and the condenser. The other polar (analyzer) is usually placed between the objective and the ocular tube and defines the plane of polarization of the light reaching the ocular. One or both must be accurately rotatable about the optical axis of the instrument. Usually the analyzer is also removable from the optical path.

A real polarizer is never totally efficient in rejecting light whose electric vector is at 90° to the desired direction. Thus, when two polars are placed in sequence in the optical path with their electric-vector transmission directions mutually crossed at 90°, most of the incident light is blocked. However, a small amount of unpolarized light $I_\perp$ leaks through. When the polars are oriented with their electric-vector transmission directions parallel, the intensity of light transmitted $I_\parallel$ is half the intensity of the incident unpolarized light minus any light lost by reflection or absorption in the polars, plus $I_\perp$. The figure of merit for polars is the extinction factor (EF), as given in Eq. (1).

$$\mathrm{EF} = \frac{I_\parallel}{I_\perp} \tag{1}$$

The most frequently used type of polar is a dichroic

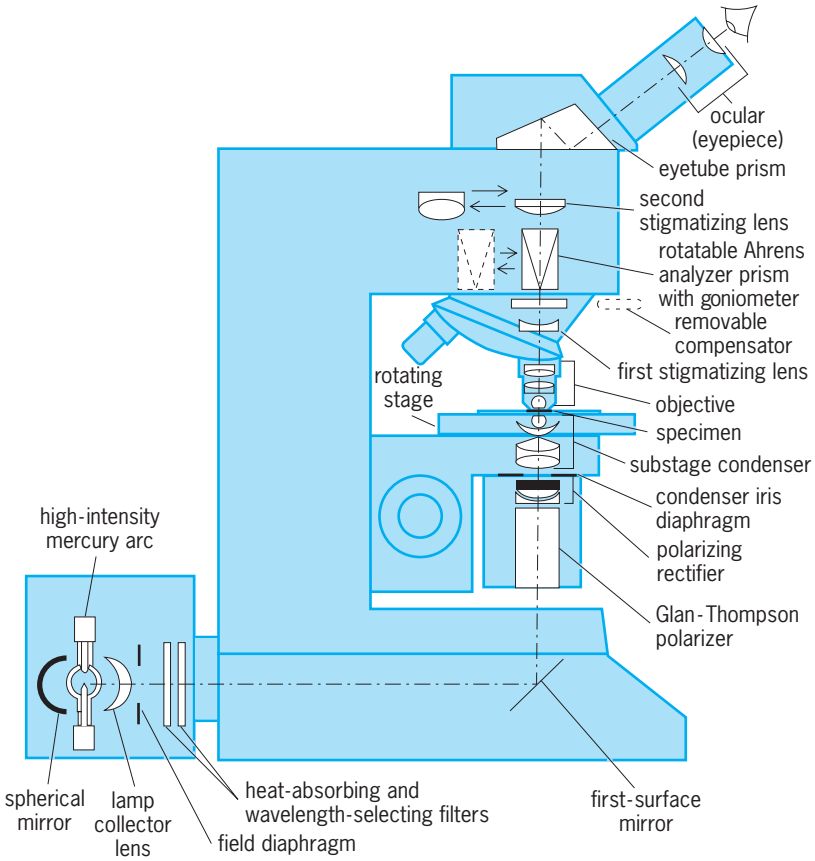

Diagram of a high-sensitivity biological polarizing microscope. Polarizing microscopes for petrography may not require the use of the expensive prism polarizer and analyzer, which can be replaced by dichroic polarizing filters. Some petrographic applications require the addition of a universal stage, which permits specimen rotation about all three axes. In all other respects the petrographic polarizing microscope has a similar optical layout. The parts of the optical train are indicated.

sheet polarizing filter. It can have an extinction factor as high as 500,000. Filter polars have wide acceptance angles, are easily accommodated without major changes in the optical system, and are significantly less expensive than prism polars. However, they do absorb as much as 56% of the light that should be transmitted. When used for both polarizer and analyzer, the maximum amount of light that filter polars can pass through the system is only 10% of the light input to the polarizer.

The sensitivity of a polarizing microscope is ultimately limited by the radiance of the illumination source and the extinction of the optical system. A high-quality prism polarizer can have an extinction factor in excess of 10^6, while transmitting virtually all of the properly polarized light. Other factors being equal, a microscope with high-quality filter polars requires a light source with more than five times the radiance needed to achieve the same sensitivity as a microscope using high-quality prism polars.

Unfortunately, the use of prism polarizers complicates the microscope design. Additionally, the polarizing prisms available share the common characteristic of transmitting the extraordinary component whose velocity varies with its angle relative to the crystal axes. This results in a slightly astigmatic image. *See* ABERRATION (OPTICS).

Thus, for petrographic, crystallographic, and most microchemical applications, dichroic filter polars are the better choice. For critical biological applications, such as investigating the weak birefringence of cytoskeletal structures in living cells, the expense and complication attendant to the use of prism polars can be justified by the attainment of sensitivity unobtainable by other means.

Other requirements and features. In addition to the polars, all polarizing microscopes need rotatable specimen stages and one or more removable birefringence compensators. Slots for insertion of the compensators are located near an aperture plane either between objective and analyzer or between polarizer and condenser.

The compensators are birefringent devices used to measure magnitude and sign of retardation due to specimen birefringence, to enhance specimen image contrast, and to manipulate the state of polarization of light passing through any point in the specimen. In general, compensators provide either a fixed amount of birefringent retardation, to which the specimen identifiably adds or subtracts, or a quantitatively variable amount of retardation that is added to or subtracted from the specimen retardation to bring the combination to an identifiable state.

Polarizing microscopes for different applications have some differences in construction. For petrography and crystallography, the microscope should be able to accept a universal specimen stage capable of rotating the specimen about three axes. A polarizing microscope for biological use has less rigid requirements for angular orientation but needs rectified optics, the extremely high extinction polars noted above, and a sensitive elliptic compensator because the birefringence of the typical biological specimen is much smaller. An ocular telescope or a Bertrand lens (a built-in lens that optionally converts the ocular into a telescope) is essential for some applications and useful for all.

Rectified optics. Even optics free of any form of birefringence have been found to depolarize light, degrading the extinction factor when large numerical apertures are involved. The cause of lens depolarization is the result of differential reflection at lens surfaces. Improved designs include polarization rectifiers consisting of a simple zero-power lens to duplicate the rotation caused by the lens to be rectified, and a half-wave retardation plate to turn the rotation into its mirror image at every point over the lens aperture. These eliminate or greatly reduce depolarization at high numerical apertures. It has also been shown that an image-degrading diffraction artifact is present in polarized light images produced with unrectified optics, and it has been demonstrated that rectification eliminates the artifact. Thus, for objects requiring both high extinction and high resolution, such as living cells, rectified optics greatly extend the power of polarization microscopy.

Modes of use. There are two traditional modes of use for the polarizing microscope, the orthoscopic mode and the conoscopic mode.

Orthoscopic mode. In the orthoscopic mode, the ocular projects an image of the specimen, as in conventional microscopy. Rotation of the specimen stage reveals the location and orientation of any anisotropic features in the specimen.

When rotated 360° between crossed polars, a birefringent object will go from dark to light four times, and all isotropic features will remain dark. A birefringent object is darkest whenever its slow axis is parallel to either polar, and brightest when the slow axis is at 45° to the polars. (The slow axis is the direction in the material for which the dielectric constant is highest.) This brightening is the result of the conversion by the specimen birefringence of the linearly polarized light from the polarizer to elliptically polarized light that has a component in the direction of the analyzer. The light intensity emerging from the analyzer I_a is related to the light intensity emerging from the analyzer when the polars are parallel. $I_\parallel$ by Eq. (2), where δ is the specimen retardation

$$I_a = I_\parallel \sin^2\left(\frac{\delta}{2}\right)\sin^2(2\theta) + I_\perp \qquad (2)$$

(expressed as the angular fraction of a wavelength) between ordinary and extraordinary waves, θ is the angle between the specimen slow axis and the polarizer, and $I_\perp$ is the unpolarized light leaking through the crossed polars.

Although the relative brightness of the specimen, placed with its axes at 45° to the polarizer, is a function of the specimen birefringence, it is not a linear function. Also birefringence is not the exclusive source of light between crossed polars. A compensator is therefore used to measure the birefringence retardation introduced by the specimen. For weakly retarding specimens, an elliptic compensator provides the greatest sensitivity. This compensator consists of a thin mica plate, providing one-tenth-wave or less retardation, mounted perpendicular to the microscope's axis and rotatable about that axis to 0.1° accuracy.

A measurement is made by setting the compensator to introduce the same ellipticity as the specimen but of opposite sign so that the image of the specimen is extinguished (maximally darkened). The procedure is as follows: (1) The compensator slow axis is set parallel to the polarizer (indicated by extinction of the nonbirefringent parts of the field of view). (2) The specimen is rotated to its extinction position, aligning its axes with the polars, and then it is rotated 45° to its brightest position (which the eye can less readily determine than its darkest position). (3) The compensator is rechecked for background extinction, in case the slide or cover glass has any residual birefringence, and the compensator angle is recorded. (4) The compensator is then rotated a small amount and the change in specimen brightness observed. If the specimen is brighter, the compensator is adding to the specimen birefringence and should be rotated in the opposite direction for compensation. If the specimen darkens, the compensator is compensating the specimen and should be rotated to the point of specimen extinction. (5) The compensator angle is recorded, and the difference between the two readings is used as the compensator azimuth angle in calculating the specimen retardation in degrees or in nanometers.

This procedure yields the value for the amount of the differential in light path through the specimen, but it does not yield directly either the specimen's extraordinary and ordinary refractive indices or its coefficient of birefringence. Because specimen retardation is the product of the thickness of the specimen times its coefficient of birefringence, any means that can yield the specimen thickness will give the coefficient of birefringence, but the actual value of either refractive index must come from some form of refractometry. In many cases, knowledge of specimen symmetry will allow a good estimate for specimen thickness. Estimates based on use of the microscope's fine focus calibration have poor accuracy but are at times the only available alternative.

A dichroic object is also birefringent and will show up as such in this test. To determine which birefringent objects are also dichroic, the analyzer is removed from the optical path, and the stage is rotated again. In the absence of the analyzer, a dichroic object darkens twice per revolution, being darkest when its slow axis is parallel to the polarizer. Optically active objects and scattering objects pass at least some light when placed between crossed polars, and do not darken with specimen rotation. An optically active object may be distinguished from a scattering object by rotating one polar relative to the other. At some offset from the crossed position between the polars, the optically active object darkens differentially relative to the background, while the scatterer does not darken.

Conoscopic mode. The conoscopic mode is used to characterize crystalline specimens. Here the ocular is not used to project an image of the specimen, but rather—with the aid of a special lens—the image of the objective exit pupil is examined to reveal the relative retardation experienced by polarized light as a function of its angle of incidence on the specimen.

For effective use of this mode, the specimen is mounted in a universal stage that allows it to be rotated about all three axes relative to the optical axis of the microscope. This permits the location of the crystal's optic axis (or axes) and the axial ratios of its index ellipsoid (indicatrix) to be determined, as well as the optic axial angle, in the case of biaxial crystals. Because the optic axis is the direction of travel through the crystal for which it appears isotropic, rotation of the specimen about the viewing axis will reveal the optic axes as spots that remain dark regardless of their orientation relative to the crossed polarizer and analyzer. *See* CRYSTAL OPTICS; MICROSCOPE.

Gordon W. Ellis

Bibliography. P. J. Duke and A. G. Michette (eds.), *Modern Microscopy: Techniques and Applications*, 1989; M. Pluta, *Advanced Light Microscopy: Specialized Methods*, vol. 2, 1989; E. M. Slayter and H. S. Slayter, *Light and Electron Microscopy*, 1993.

Polarographic analysis

An electrochemical technique used in analytical chemistry. Polarography involves measurements of current-voltage curves obtained when voltage is applied to electrodes (usually two) immersed in the solution being investigated. One of these electrodes is a reference electrode: its potential remains constant during the measurement. The second electrode is an indicator electrode. Its potential varies in the course of measurement of the current-voltage curve, because of the change of the applied voltage. In the simplest version, so-called dc polarography, the indicator electrode is a dropping-mercury electrode, consisting of a mercury drop hanging at the orifice of a fine-bore glass capillary (usually about 0.08 mm inner diameter). The capillary is connected to a mercury reservoir so that mercury flows through it at the rate of a few milligrams per second. The outflowing mercury forms a drop at the orifice, which grows until it falls off. The lifetime of each drop is several seconds (usually 2 to 5). Each drop forms a new electrode; its surface is practically unaffected by processes taking place on the previous drop. Hence each drop represents a well-reproducible electrode with a fresh, clean surface. *See* ELECTRODE; REFERENCE ELECTRODE.

Apparatus. The dropping-mercury electrode is immersed in the solution to be investigated and placed in a cell containing the reference electrode (**Fig. 1**). Polarographic current-voltage curves can be recorded with a simple instrument consisting of a potentiometer or another source of voltage and a current-measuring device (such as a sensitive galvanometer). The voltage can be varied by manually changing the applied voltage in finite increments, measuring current at each, and plotting current as a function of the voltage. Alternatively, commercial instruments are available in which voltage is increased linearly with time (a voltage ramp), and current variations are recorded automatically (**Fig. 2**).

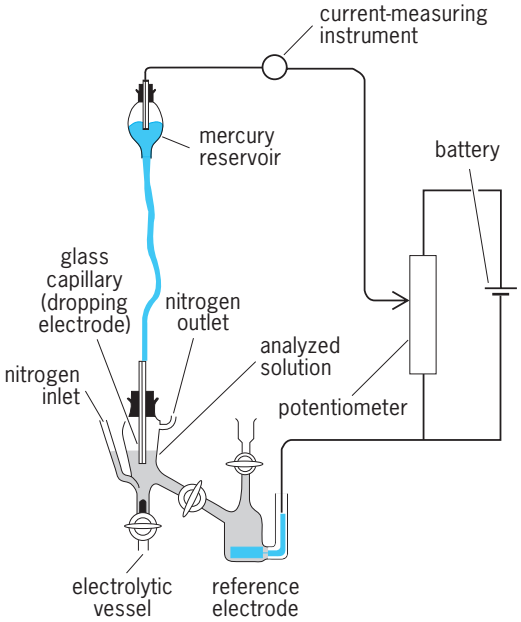

Fig. 1. Diagram showing components of a polarographic circuit.

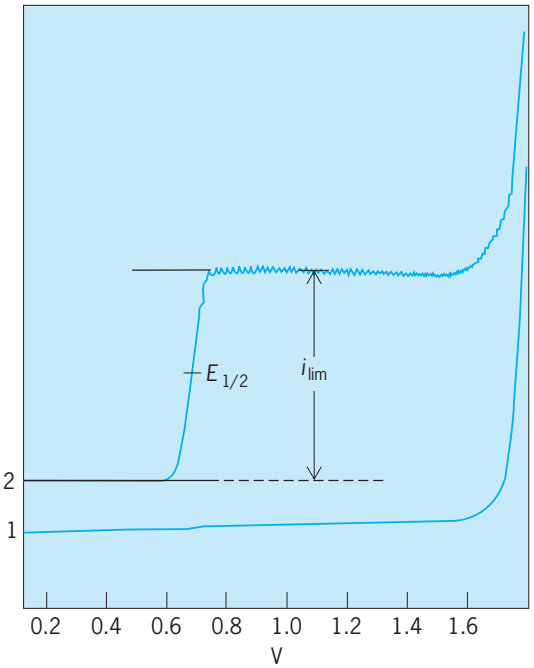

Fig. 2. Polarographic current-voltage curves: 1 = supporting electrolyte; 2 = curve in the presence of an electroactive (reducible) compound; $E_{1/2}$ = half-wave potential; i_{lim} = limiting current (wave height).

Polarographic curves. Oscillations of current result from growth of the individual mercury drops. Their mean value is usually measured. Three portions can be observed on a typical polarographic curve. At sufficiently positive potentials, only a small current flows. Then, in a region characteristic of the particular species and solution, an S-shaped current rise is observed. Finally, at sufficiently negative potentials, the current is again independent of applied voltage over a potential range. This current is called the limiting current and is usually a few microamperes or a fraction of a microampere. The limiting current (also called the height of the polarographic wave) increases, usually with concentration of the investigated species in the solution. The wave height can be thus used for determination of concentration (that is, how much of the species is present in the solution). The potential at the point of the polarographic wave, where the current reaches half of the limiting value, is called the half-wave potential (Fig. 2). This potential is characteristic of the species studied (that is, it can be used to confirm which species are present).

Solutions. The solutions investigated polarographically must contain the species to be studied in 10^{-3} to 10^{-6} M concentration. In addition, the solution must contain a large excess (50-fold or greater) of a supporting electrolyte which does not react at the electrode in the potential region of interest and which can be a neutral salt, acid, base, or a mixture (such as a buffer). The function of the supporting electrolyte is to reduce the resistance of the solution, to ensure that species are transported to the

surface of the electrode by diffusion rather than by migration in the electric field, to keep conditions at the electrode surface unchanged in the course of polarographic electrolysis, and to keep or convert the species into a form most suitable for electrolysis (for example, by complex formation or protonation).

Evaluations. The heights of polarographic waves are measured and compared with heights obtained with standard solutions of the species, the concentrations of which are known. Accuracy varies from 2 to 5%, according to the potential range, shape of polarographic waves, and presence of other components. Under strictly controlled conditions, even 1% accuracy can be achieved.

Polarographically active species. To be studied polarographically, a species must undergo reduction or oxidation at the surface of the mercury drop, must form compounds with mercury, or must catalytically affect electrode processes. From changes in the shape of polarographic curves, it is also possible to determine various surfactants.

To be able to follow a given reducible species, its reduction must occur at more positive potentials than reduction of the component of the supporting electrolyte or solvent. Oxidations can be followed if they occur at potentials more negative than that of oxidation of the solvent, supporting electrolyte, or electrode material. Reduction processes are indicated by cathodic waves (above the zero current line), and oxidation processes by anodic waves (below the zero current line). *See* OXIDATION-REDUCTION.

Polarography has been used for the determination of most metals (typically lead, copper, zinc, cadmium, iron, uranium, cobalt, nickel, manganese, potassium, sodium, and so forth), some inorganic acids and their anions (for example, iodic and periodic acids, nitrates, and nitrites), and some gases (for example, sulfur dioxide and oxygen). The wave of oxygen forms the basis of a convenient oxygen determination in technical gases, waters, or biological material. Alternatively the reducibility of oxygen means that oxygen must be removed (usually by purging with nitrogen) from solutions to be analyzed for other constituents. Reduction of inorganic species results in the lowering of the oxidation state, which may be accompanied by amalgam formation.

Electrode processes involving organic compounds at the dropping-mercury electrode result in a cleavage or formation of a chemical bond in the investigated organic molecule. Polarographic behavior depends primarily on the nature of the bond involved, even when the molecular environment (that is, the presence and kind of neighboring groups and other substituents, the type of the molecular frame to which the electroactive group is bound, and the spatial arrangement) also affects the electrode process. The reducibility of multiple bonds (for example, C═O, C═N, C═C, and N═N) or of strongly polarizable groups (such as NO_2, NO, and C-halogen), especially when conjugated, was recognized early. Recently the reducibility of some single bonds such as C—O, C—N, C—S, or C—P when the molecule contains activating groups was proved.

Because mercury ions react with various anions, the anodic current corresponding to the dissolution of mercury can be used, for example, for the determination of halides, thiocyanate, cyanide, and sulfide among inorganic species; and of mercaptans, urea and thiourea derivatives, and dithiocarbamates among organic compounds.

Most important among catalytic waves are those which correspond to catalytic hydrogen evolution. Among compounds giving catalytic waves in buffered ammoniacal solutions of cobalt salts are some types of proteins.

The protein molecule must contain a suitably situated thiol or disulfide group and must be able to form a complex with cobalt and ammonia. It is assumed that the proton transferred by this complex undergoes reduction more easily than the hydronium ion (H_3O^+). *See* PROTEIN.

When blood serum is alkali-denatured and the proteins with largest molecular weight are separated by precipitation with sulfosalicylic acid, the remaining filtrate contains proteins and their fragments, the contents of which differ in the serum of healthy and pathological individuals. Increase of the catalytic wave is observed for cancer and inflammatory diseases, and decrease of the catalytic wave for hepatitis and other liver diseases. In Europe the polarographic test is used in clinical analysis as a general screening test, in connection with nine other tests as a proof of malignancy, and in controlling cure of cancer and measuring the effectiveness of treatments such as surgery or irradiation. *See* CLINICAL PATHOLOGY.

Applications. Polarographic studies can be applied to investigation of electrochemical problems, to elucidation of some fundamental problems of inorganic and organic chemistry, and to solution of practical problems.

In electrochemistry polarography allows measurement of potentials, and yields information about the rate of the electrode process, adsorption-desorption phenomena, and fast chemical reactions accompanying the electron transfer. Since polarographic experiments are simple and not too time-consuming, polarography can be used in a preliminary test designed to find the most suitable model compound for detailed electrochemical investigations.

In fundamental applications, polarography makes it possible to distinguish the form and charge of the species (for example, inorganic complex or organic ion) in the solution. Polarography also permits the study of equilibria (complex formation, acid-base, tautomeric), rates, and mechanisms. For equilibria established in the bulk of the solution in more than 15 s, measurement of wave heights of individual components makes possible the evaluation of equilibrium constants. For equilibria that are very rapidly established at the electrode surface, equilibrium constants can be determined from shifts of half-wave potentials. Finally, for some equilibria between these two extremes, which are established in times comparable to the drop time (3 s), calculation

of rate constants is possible (for example, for dehydration of hydrated aldehydes, or for protonation of anions derived from organic acids, such as $C_6H_5COCH_2COCH_3$). In this way, rate constants of very fast reactions of the order 10^5–10^{10} liters mol^{-1} s^{-1} can be determined.

For slower reactions, rate constants can be found from changes of wave heights with time. Moreover, as some reaction intermediates giving separate waves can be detected, identified, and followed polarographically (if their half-lives are longer than 15 s), polarography can prove useful in mechanistic studies. Elimination of Mannich bases, hydration of multiple bonds in unsaturated ketones, and aldolization are examples studied. *See* ORGANIC REACTION MECHANISM.

Finally, polarography can be used for investigation of the relationship between electrochemical data and structure. Reduction of most organic systems, particularly in aqueous media, involves steps with a high activation energy and is therefore irreversible. Half-wave potentials of such systems are a function of the rate constant of the electrode process. This heterogeneous rate constant is frequently influenced by structural effects, as are, in a similar manner, rate constants of homogeneous reactions. Therefore, effects of substituents on half-wave potentials in aromatic systems and in aliphatic systems can be treated by the Hammett relations and the Taft substituent constants, respectively. Among steric effects, for example, steric hindrance of coplanarity or effects of cis-trans isomerisms affect polarographic curves. Polarography also makes it possible to distinguish between some epimers, for example, bearing axial or equatorial halogen. *See* MOLECULAR ISOMERISM.

Practical applications are predominantly analytical procedures. In inorganic analysis, polarography is used predominantly for trace-metal analysis (with increased sensitivity of differential pulse polarography and stripping analysis). In organic analysis, it is possible in principle to use polarography in elemental analysis, but such applications are deservedly infrequent. More frequent are applications in functional group analysis. Either the reaction product (for example, semicarbazone in the determination of carbonyl compounds, or *N*-nitrosoamine in determinations of secondary amines) is measured, or a decrease in concentration of a reagent (for example, chromic acid in determination of alcohols, or mercuric or silver ions in determination of thiols) is followed. Or, alternatively, the polarographic limiting current can be taken as a measure of the amount of a given electroactive group (for example, quinones, aliphatic nitro compounds, phenazines, or thiols).

More frequent is determination of individual compounds. If the species is electroactive, the analysis frequently consists of dissolving the material in a proper supporting electrolyte, recording the waves, and evaluating them. Electroinactive species can be determined indirectly, by converting them into electroactive by a suitable chemical reaction (for example, aromatics by nitration, secondary amines by nitrosation).

The most important fields of application of inorganic determinations are in metallurgy, environmental analysis (air, water, and seawater contaminants), food analysis, toxicology, and clinical analysis. The possibility of being able to determine vitamins, alkaloids, hormones, terpenoid substances, and natural coloring substances has made polarography useful in analysis of biological systems, analysis of drugs and pharmaceutical preparations, determination of pesticide or herbicide residues in foods, and so forth. Polarography also makes possible determination of monomers, catalysts, and some reactive groupings in polymers.

Other polarographic techniques. To eliminate unwanted charging current and to increase the sensitivity of polarography, the voltage is applied in regular pulses instead of gradually. When the current is measured only during the second half of the pulse, the technique is called pulse or differential pulse polarography. This technique is more sensitive by two orders of magnitude than dc polarography, and in inorganic trace analysis competes with atomic absorption and neutron activation analysis. With smaller dosage of more effective drugs, the sensitivity of differential pulse polarography has found application in drug analysis.

When an alternating voltage of small amplitude (a few millivolts) is superimposed on the dc voltage ramp, the technique is called ac polarography. It is particularly useful for obtaining information on adsorption-desorption processes at the surface of the dropping-mercury electrode. An increase in accuracy of polarographic methods can be achieved when the dropping-mercury electrode (or some other indicator electrode) is used in titrations. At constant applied voltage the current is measured as a function of volume of added titrant in amperometric titrations. The amperometric titration curve frequently has a shape of two linear sections. Their intersection corresponds to the titration end point.

Related methods. Methods in which a potential sweep is used (instead of the practically constant potential of the electrode during the life of a single drop in polarography) are frequently called voltammetry. Either an electrode with nonrenewed surface—solid electrode, mercury pool, or mercury drop—is used, or the whole scan is carried out during the life of a single drop when the dropping-mercury electrode is used. The use of solid electrodes (for example, platinum, gold, or various forms of graphite) makes it possible to extend the voltage range to more positive potentials, so that numerous substances which cannot be oxidized when mercury is used can be oxidized. Solid electrodes can be used when they are stationary, rotating, or vibrating, but the change at the electrode surface in the course of electrolysis often decreases the reproducibility of the current-voltage curve compared with polarographic curves.

When extremely small concentrations of amalgam-forming metals (such as lead, cadmium, or zinc) are to be determined, a preconcentration can be carried out. The metal is deposited into a hanging mercury

drop or into mercury plated on surfaces of solid electrodes. After a chosen time interval, a voltage sweep is applied, and the dissolution current of the metal ion from the amalgam is measured. Such anodic stripping, particularly when combined with the differential pulse technique, allows determination up to 1 part in 10^{12} parts of solution. Similarly, anions forming compounds with mercury can be preconcentrated by electrodeposition and dissolved in cathodic stripping. Alternatively, compounds absorbed at the mercury surface can be stripped.

When the applied voltage sweep is first increased and then decreased and the plot of dependence of voltage on time resembles a triangle, the technique is called cyclic voltammetry and offers, in particular, information about products and intermediates of electrode processes and their reactions.

Finally, when current is applied and the potential of an electrode measured, another group of techniques is developed. Most important among them is chronopotentiometry. In this technique, a constant current is applied to the indicator electrode in an unstirred solution, and its potential change is measured as a function of time. The time at which a sudden potential change is observed is known as the transition time, and its square root is proportional to the concentration of the reacting species. The method has found only limited practical application but has proven useful for the study of reactions of products of electrolysis. *See* ELECTROCHEMICAL TECHNIQUES.

Petr Zuman

Bibliography. A. Bard, *Electroanalytical Chemistry*, 1993; A. J. Bard (ed.), *Electroanalytical Chemistry*, vols. 1–16, 1966–1988; A. J. Bard, *Electrochemical Methods: Fundamentals and Applications*, 2d ed., 2000; A. M. Bond, *Modern Polarographic Methods in Analytical Chemistry*, 1980; Z. Brainina and E. Neyman, *Electroanalytical Stripping Methods*, 1993; G. W. Ewing, *Instrumental Methods of Chemical Analysis*, 5th ed., 1985; L. Meites and P. Zuman, *Handbook Series in Inorganic Electrochemistry*, vols. 1–8, 1979–1988; L. Meites and P. Zuman, *Handbook Series on Organic Electrochemistry*, vols. 1–6, 1977–1983; J. T. Stock and M. V. Orna (eds.), *Electrochemistry, Past and Present*, 1989; B. H. Vassos and G. W. Ewing, *Electroanalytical Chemistry*, 1983.

Polaron

A quasiparticle that forms from an electronic charge carrier interacting with a solid's vibrating displaceable atoms. In particular, a polaron comprises an electronic charge carrier together with the carrier-induced altered motions of surrounding atoms. Different classes of material have distinct types of polarons. Polarons may be divided into weak-coupling polarons and strong-coupling polarons. Strong-coupling polarons comprise large self-trapped polarons and small self-trapped polarons.

All types of polarons form because of the electron-lattice interaction. The electron-lattice, or electron-phonon, interaction describes the dependence of the energy of an electronic charge carrier on the positions of the surrounding atoms. As a result of this interaction neither electronic charge carriers nor atoms can move without affecting the other. *See* LATTICE VIBRATIONS; PHONON.

The electron-phonon interaction can be divided into two components. The long-range component of the electron-lattice interaction describes the dependence of the energy of an electronic charge carrier in an ionic (or polar) material on the positions of the surrounding ions. This interaction, like the underlying Coulomb interaction, is of long range. The short-range component of the electron-lattice interaction describes the dependence of the energy of an electronic charge carrier on the positions of nearby atoms with which the electronic carrier forms bonding or antibonding states. The short-range interaction predominates in covalent semiconductors. The term polaron was adopted in recognition of the electron-lattice interaction typically being especially strong in polar (ionic) materials. *See* COHESION (PHYSICS); COULOMB'S LAW; IONIC CRYSTALS; SEMICONDUCTOR.

Electrons are lighter and more rapidly moving than are the relatively heavy and sluggish atoms with which they interact. Thus, electronic carriers tend to follow atoms' motions. Polaron effects result as electronic charge carriers adjust to and modify atomic motion.

Weak-coupling polaron. In some instances the electron-lattice interaction only slightly alters the motions of an electronic carrier and the atoms with which it interacts. Then the range over which an electronic carrier and the surrounding atoms affect one another is termed the weak-coupling polaron radius. This radius can be viewed as the distance that an electronic carrier can diffuse in the time that it takes for atoms to execute a single vibration. In this weak-coupling regime the electron-lattice interaction simply acts to slightly lower the net energy and to slightly enhance the effective mass of an electronic charge carrier.

Strong coupling: self-trapping. More dramatic effects occur when the electron-lattice interaction is strong enough and the electronic carrier moves slowly enough for it to become self-trapped. To understand the physical origin of self-trapping, it is useful to consider an electronic charge held fixed in a condensed medium. The presence of a localized electronic charge forces a shift of the equilibrium positions of the surrounding atoms (**Fig. 1**). This change of atoms' equilibrium positions lowers the electron's potential energy. If the potential well enveloping the electron is deep enough, the electron will be bound within it. Then, without the imposition of external forces, the electron cannot escape the potential well unless it is altered by atoms changing their positions. This situation will be energetically stable if the lowering of the electron's energy due to its being bound exceeds the strain energy associated with displacing atoms' equilibrium positions. This situation will be dynamically stable if the bound electron can complete an orbit in its well before the surrounding atoms can move appreciably.

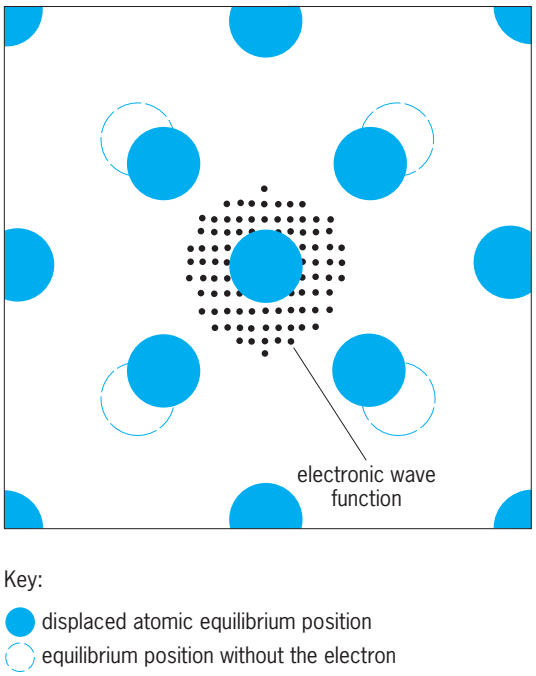

Fig. 1. An electron self-trapped by the displacements of equilibrium positions of atoms that surround it.

When these two stability conditions are fulfilled, an electronic carrier is said to be self-trapped. The term self-trapping emphasizes that the electronic carrier is bound within a potential well that its very presence induces. The associated polaron comprises the self-trapped carrier and the associated pattern of displaced atomic equilibrium positions.

Energies of self-trapped polarons lie below those for carriers that are not self-trapped (**Fig. 2**). The width of these polarons' energy band is less than that of a characteristic phonon, $h\nu$, where h is Planck's constant and ν is the characteristic atomic vibration frequency. The narrowness of the polaron band is indicative of the slowness of polaron motion. Self-trapped carriers cannot move without atoms altering their positions.

Polarons are also associated with the absorption of electromagnetic radiation. These absorption bands arise when self-trapped electrons are excited from their ground state. The electrons may thereby be freed from the self-trapped potential well or raised to higher states within it.

Polarons may also pair as bipolarons. Pairing is fostered because placing two carriers proximate to one another tends to deepen the self-trapping potential well that both carriers experience. This tendency of polarons to amalgamate competes with the carriers' mutual Coulomb repulsion. This Coulomb repulsion is reduced and bipolaron formation aided if the carriers can avoid one another while moving within a common potential well. The coordination of the two carriers' motion is termed an electronic correlation effect. A bipolaron forms a singlet when the electronic spins of its two electronic carriers are aligned in opposition to one another. Thus, a singlet bipolaron has no net electronic spin but twice the electronic charge of a single electron. These properties may be used to identify singlet bipolarons.

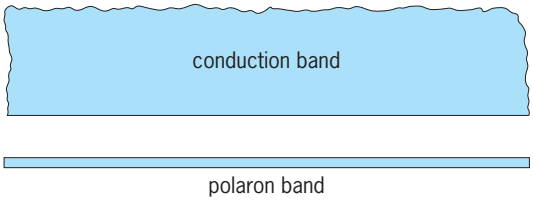

Fig. 2. Energy levels of quasifree and self-trapped states when the polaron is energetically stable.

Polaron formation is a nonlinear process. That is, self-trapping may be regarded as a feedback phenomenon. In particular, confining a charge carrier induces atomic displacements that deepen the potential well that confines the carrier, thereby enhancing the charge carrier's confinement. These feedback effects are subtle. Changes of the electron-lattice interaction qualitatively change the type of self-trapped polaron that is formed. *See* NONLINEAR PHYSICS.

Large (self-trapped) polaron. The self-trapping polaron that is formed with only the long-range component of the electron-lattice interaction is termed a large polaron since its self-trapped carrier generally extends over multiple atomic sites. The spatial extent of the self-trapped carrier decreases continuously as the strength of the electron-lattice interaction is increased.

Large polarons can move coherently through a solid. In particular, the centroid of a large polaron's self-trapped carrier shifts when the atoms responsible for the carrier's self-trapping alter their positions appropriately. Since atomic motion is required for the self-trapped carrier to move, large polarons move slowly with very large effective masses. Estimated effective masses are as high as several hundred times the rest mass of an electron. *See* BAND THEORY OF SOLIDS.

Large polarons' very high effective masses cause them to move with exceptionally large momenta. As a result, large polarons are not easily scattered. In particular, the rate with which a large polaron is scattered by atomic vibrations is often inversely proportional to its effective mass. Then a large polaron's mobility, its velocity divided by the electric field that drives its motion, is independent of its effective mass. The mobility of a large polaron also falls as its scattering by vibrating atoms increases upon raising the temperature. The large polaron mobility often varies inversely with temperature over a significant temperature regime (**Fig. 3**). Near room temperature, a large polaron's mobility is typically comparable to that of a conventional charge carrier that is not self-trapped. Thus, it is the relatively small scattering rate, not the mobility, that best distinguishes large-polaron motion from that of a conventional charge carrier, a weak-coupling polaron.

Small (self-trapped) polaron. The self-trapped carrier that is formed with just the short-range component of the electron-lattice interaction is termed a small polaron since its self-trapped carrier shrinks to

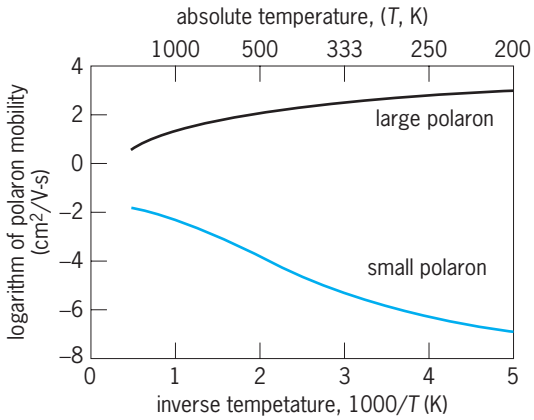

Fig. 3. Polaron mobilities versus inverse temperature.

the smallest relevant structural unit (such as an atom, bond, or molecule). Alternatively, if the electron-lattice interaction is too weak to induce a collapse into a small polaron, the charge carrier does not self-trap at all. The carrier's interaction with vibrating atoms then produces only a weak-coupling polaron. This dichotomy between a severely localized self-trapped carrier and a carrier that is not self-trapped is a manifestation of the nonlinear character of self-trapping.

In the presence of both long-range and short-range components of the electron-lattice interaction, a self-trapped carrier may form a large polaron or a small polaron, or both types of self-trapped polaron may coexist. Since imposition of disorder fosters a carrier's localization, disorder tends to trigger small-polaron formation. Thus small-polaron formation is more easily achieved in a disordered material than in a crystal.

The motion of a small polaron is generally described as proceeding by a succession of thermally assisted jumps between adjacent sites. Thus, a small polaron has an extremely low mobility that increases with rising temperature (Fig. 3). The temperature dependence of the small-polaron mobility becomes progressively milder as the temperature is lowered. This weakening of the mobility's temperature dependence results from atoms' motion being increasingly dominated by quantum-mechanical effects upon cooling. All told, small-polaron motion is qualitatively different from large-polaron motion. A large polaron has its free motion occasionally impeded by a scattering event. By contrast, a small polaron only occasionally departs from being static to jump to an adjacent site.

A magnetic field deflects a current of small polarons, thereby producing a Hall effect. However, the small-polaron Hall effect is unlike that of conventional charge carriers. In particular, the magnitude and temperature dependence of the mobility determined in a Hall-effect experiment, the Hall mobility, differs from the mobility that enters into the electrical conductivity (Fig. 3). The Hall mobility is generally larger and less temperature dependent than is the conductivity mobility. Most strikingly, the sign of the small-polaron Hall effect is often anomalous. Then, *p*-type small polarons in a semiconductor produce an *n*-type Hall effect, and similarly, *n*-type small polarons can generate a *p*-type Hall effect. These distinctive features provide a means of identifying the presence of small polarons. *See* HALL EFFECT. David Emin

Bibliography. C. L. Chien and C. R. Westgate (eds.), *The Hall Effect and Its Applications*, Plenum Press, New York, 1980; D. Emin, Small polarons, *Phys. Today*, 35(6):34–40, June 1982; C. Kittel, *Quantum Theory of Solids*, Wiley, New York, 1963; C. G. Kuper and G. D. Whitfield (eds.), *Polarons and Excitons*, Plenum, New York, 1963; E. K. H. Salje, A. S. Alexandrov, and W. Y. Liang (eds.), *Polarons and Bipolarons in High-T_c Superconductors and Related Materials*, Cambridge University Press, 1995, paper 2005.

Poliomyelitis

An acute infectious viral disease which in its serious form affects the central nervous system and, by destruction of motor neurons in the spinal cord, produces flaccid paralysis. However, about 99% of infections are either inapparent or very mild. *See* CENTRAL NERVOUS SYSTEM.

Infectious agent. Polioviruses are no longer thought to attack only neural tissue, because (1) they multiply in cultures of many nonnervous tissues; (2) their presence in the blood and antibody formation appear before the paralytic phase is reached and in cases where even transient signs of central nervous system involvement do not occur; and (3) the virus is regularly found in the throat and stools before the onset of disease, and after onset it is found for a week in the throat, and for several weeks in the stools. The infective virus particle is about 28 nanometers in diameter. Freezing preserves it for long periods. In contrast to the arboviruses, it is destroyed only slowly by alcohol and not at all by ether and deoxycholate. Poliovirus has a very restricted host range; most strains are limited to primates when cultured in the animal body and to primate tissue when in tissue culture. Three antigenic types of poliovirus are known. *See* ANIMAL VIRUS; ARBOVIRAL ENCEPHALITIDES; ENTEROVIRUS; TISSUE CULTURE.

Pathogenesis. The virus probably enters the body through the mouth; primary multiplication occurs in the throat and intestine. Transitory viremia occurs; the blood seems to be the most likely route to the central nervous system. The severity of the infection may range from a completely inapparent through minor influenzalike illness, or an aseptic meningitis syndrome (nonparalytic poliomyelitis) with stiff and painful back and neck, to the severe forms of paralytic and bulbar poliomyelitis. In all clinical types, virus is regularly present in the enteric tract. In paralytic poliomyelitis the usual course begins as a minor illness but progresses, sometimes with an intervening recession of symptoms (hence biphasic), to flaccid paralysis of varying degree and persistence. When the motor neurons affected are

those of the diaphragm or of the intercostal muscles, respiratory paralysis occurs. Bulbar poliomyelitis results from viral attack on the medulla (bulb of the brain) or higher brain centers, with respiratory, vasomotor, facial, palatal, or pharyngeal disturbances.

Diagnosis. Laboratory diagnosis is by isolation, usually from stools inoculated into tissue cultures, and subsequent identification by neutralization with specific antiserums in a test tube, and by complement-fixing and neutralizing serum antibody rises. Isolation of a virus which is cytopathogenic in tissue cultures, but which is not antigenically identified, is not sufficient for diagnosis, for many other cytopathogenic enteroviruses also inhabit the enteric tract and produce syndromes similar to mild or early poliomyelitis. *See* NEUTRALIZING ANTIBODY.

Epidemiology. Poliomyelitis occurs throughout the world. In temperate zones it appears chiefly in summer and fall, although winter outbreaks have been known. It occurs in all age groups, but less frequently in adults because of their acquired immunity. In crowded underdeveloped areas and in tropical countries, where conditions favor a constant widespread dissemination of virus, poliomyelitis continues to be a disease of infancy, and a high percentage of children over 4 years old are already immune. Before the introduction of vaccine, in some areas of the temperate zones the age incidence had tended to change in marked parallel with improving sanitation and hygiene; exposure had been postponed, sometimes so long that parents were without immunity and succumbed to infections transmitted from their children. In 1952, the year of maximum incidence in the United States, more than 57,000 cases were reported; more than half involved paralysis.

The virus is spread by human contact; the nature of the contact is not clear, but it appears to be associated with familial contact and with interfamily contact among young children. The virus may be present in flies.

Prevention and control. Inactivated poliovirus vaccine (Salk; IPV), prepared from virus grown in monkey kidney cultures, was developed and first used in the United States, but oral poliovirus vaccine (Sabin; OPV) is now generally used throughout the world. The oral vaccine is a living, attenuated virus, and like most viruses it is unstable except when held at very low temperatures in the frozen state. The use of magnesium chloride ($MgCl_2$) as a stabilizing agent for poliovirus overcomes this problem; in molar concentrations, $MgCl_2$ protects live poliovirus vaccines, so that they may be stored in an ordinary refrigerator for over a year with no loss in immunizing potency. Vaccines stabilized with $MgCl_2$ are apt to maintain potency over periods of several weeks at room or transit temperatures.

For inactivated poliovirus vaccine, at least four inoculations over a period of 1–2 years have been recommended in the primary series, followed by periodic booster immunizations to maintain immunity. Fewer injections seem necessary with the enhanced potency (eIPV) vaccine; killed vaccine induces humoral antibodies, but upon exposure, the virus is still able to multiply in the nonimmunized gut.

The live poliovirus vaccine multiplies, infects, and thus immunizes. In the process, infectious progeny of the vaccine virus are disseminated. Although the viruses, particularly types 2 and 3, mutate in the course of their multiplication in vaccinated children, only extremely rare cases of paralytic poliomyelitis have occurred in recipients of oral polio vaccine or their close contacts. Repeat vaccinations seem to be important to establish permanent immunity. The vaccine induces not only IgM and IgG antibodies in the blood but also secretory IgA antibodies in the intestine, which then becomes resistant to reinfection.

A potential limiting factor for oral vaccine is interference. If the alimentary tract of a child is infected with another enterovirus at the time the vaccine is given, the establishment of polio infection and immunity may be blocked. This may be an important problem in areas (particularly in tropical regions) where enterovirus infections are common.

Trivalent oral poliovirus vaccine is generally used in the United States. Primary immunization of infants begins at 2 months of age. The second and third doses should be given at 2-month intervals thereafter, and a fourth dose at $1^1/_2$ years of age. A trivalent vaccine booster is recommended for all children entering elementary school. Adults residing in the continental United States have only a small risk of exposure. However, those who are at increased risk because of contact with a person with the disease or who are anticipating travel to an endemic or epidemic area should be immunized.

Although in 1993 more than 150,000 paralytic cases occurred annually around the world, chiefly in developing countries, polio has become a rare disease in the United States because of vaccine effectiveness. In 1987, only five cases were reported. Wild virus, which is found in a naturally infected host, has not been isolated in the United States since 1979, and the disease has almost vanished in all industralized countries. The last case caused by wild poliovirus in the Western Hemisphere occurred in 1990. However, there is a continuing need for adequate vaccination programs in all population groups in order to limit the spread of wild viruses. This is particularly important when wild viruses are introduced from developing countries.

Both inactivated and oral poliovirus vaccines induce antibodies and protect the central nervous system from subsequent invasion by wild virus. However, upon subsequent exposure to wild virus, antibody resulting from the inactivated virus vaccine has little effect on intestinal carriage of virus. The gut develops a far greater degree of resistance after administration of the oral vaccine—resistance that seems to be dependent on the extent of initial vaccine virus multiplication in the alimentary tract rather than on serum antibody level.

Live vaccine should not be administered to immunodeficient or immunosuppressed individuals. On very rare occasions, the live vaccine induces paralytic disease in persons not known to be

immunodeficient. Such cases, which are carefully studied by public health agencies, are estimated to number about one vaccine-associated case for every million nonimmune persons vaccinated. *See* VACCINATION.

Postpolio syndrome. Paralysis and muscle wasting returns in many individuals decades after their experience with paralytic poliomyelitis. Many reported cases have been in the United States, where in 1993 there were approximately 300,000 persons with a history of polio disease. Although postpolio disease is rare, it is a specific syndrome. It does not appear to be a consequence of persistent virus infection, but seems a result of physiological and aging changes in paralytic individuals already burdened by loss of neuromuscular functions. The normal decline in anterior horn cells and consequently in muscle strength that occurs with aging is much more pronounced in persons who lost many anterior horn cells and perhaps 50% of their muscle strength during their acute bout with polio. However, loss of anterior horn cells as a result of aging seldom occurs before age 60, and most of the postpolio cases reported have occurred before this age. Also, some findings suggest that, rather than from loss of anterior horn cells, the syndrome may result from peripheral disintegration of individual nerve terminals in motor units that were reinnervated during recovery following the acute disease. *See* AGING. Joseph L. Melnick

Bibliography. E. Kcrstak, *Measles and Poliomyelitis: Vaccines, Immunization and Control*, 1993; S. M. Lemon and S. E. Robertson, Global eradication of poliomyelitis: Recent progress, future prospects, and new research priorities, *Prog. Med. Virol.*, 38:42–55, 1991; W. Melchers et al., The postpolio syndrome: No evidence for poliovirus persistence, *Ann. Neurol.*, 32:728–732, 1992; S. A Plotkin and E. A. Mortimer (eds.), *Vaccines*, 3d ed., 1999.

Pollen

The small male reproductive bodies produced in the pollen sacs of seed plants (gymnosperms and angiosperms). The term comes unchanged from Latin, meaning fine dust or flour, and was popularized by Carl Linnaeus in 1751 (but used since 1523). Pollen is usually carried by wind or insects to the sticky stigmas of angiosperms or to the naked ovules of gymnosperms. Fertilization, which follows pollination within a wide time range—minutes, hours, a year in *Agathis*, 15 months or more in juniper, and 2 years in stone pine (*Pinus pinea*)—is usually effected only by pollen of the same species. *See* FLOWER; PLANT REPRODUCTION.

Although the function of pollen has been intensively studied only in very recent times, Mesopotamians 5000 years ago understood the sexuality of date palms sufficiently to ensure their crops. Assyrian sculptures of about 1000 years later show ritual dusting by priests of pollen from male trees onto inflorescences of the female trees. Hand pollination is now a major factor in food production; for example, in the culture of hybrid maize (*Zea*). Where pesticides have eradicated pollinating insects, hand pollination has become essential. *See* AGRICULTURAL SCIENCE (PLANT); BREEDING (PLANT).

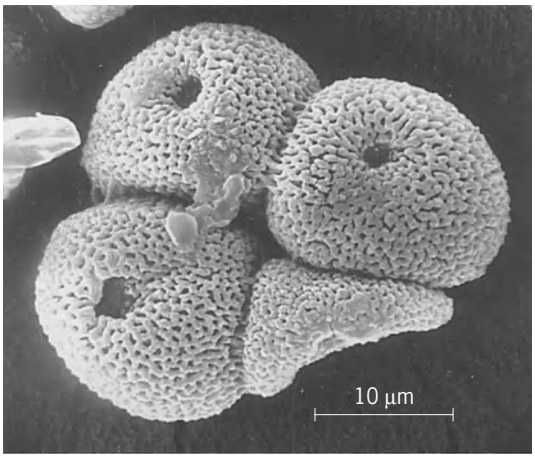

Fig. 1. Tetrad of cattail (*Typha latifolia*), grains cohering, one-pored, exine reticulate. (*Scanning electron micrograph by C. M. Drew, U.S. Naval Weapons Center, China Lake, California*)

Cattail (*Typha*) grains (**Fig. 1**) were made into cakes by natives of India and New Zealand, while cattail and maize pollen has long served in fertility rites of the Apache and Pueblo peoples, and in the Navaho search for the "Path of Peace." Unwittingly, humans have always ingested much pollen in honey and from the air. People may always have been allergic to the massive amounts shed by wind-pollinated trees such as olive and mulberry and by weed herbs in the amaranth, chenopod, daisy, and grass families, all of which should be combated as health hazards through selective planting or control of waste urban areas. *See* ALLERGY.

A palynologist's studies may involve many disciplines, from genetics and classification to aerobiology and the treatment of hay fever; to the study of pollen in pure, toxic, and adulterated honey; to the identification of remains in peats, coals, mineral sediments, feces, and in seasonal layers of ice. Most of these matrices yield datable stratigraphic records. Pollen analysis has become an indispensable tool in tracing vegetation history and climate, especially since the last glaciation. Macerated coals also yield to this technique since they are so rich in pollen and spores. Even the pollen and spore assemblages of Antarctic sediments are studied for correlation with those of other Gondwana segments.

Striking use of palynology has also been made by petroleum researchers in locating ancient shorelines and thus basins of deposition of oil and gas through recognition of land and shore algae, microspores, and other life forms, often present in high percentages in marine beds. *See* MICROPALEONTOLOGY; PALYNOLOGY; PETROLEUM MICROBIOLOGY.

The individual grain. On maturation in the pollen sac, a grain may reach 0.00007 mg as in spruce, or less than 1/20 of this weight. A grain usually has two waxy, durable outer walls—the ectexine and

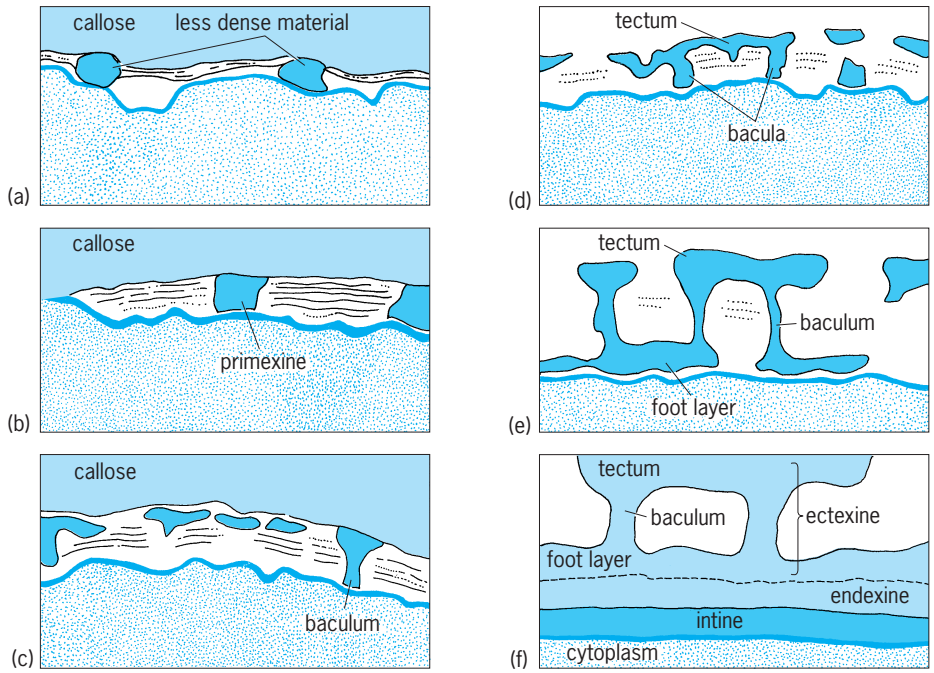

Fig. 2. Growth of hellebore (buttercup family) pollen walls: (*a*,*b*) matrix of sculpture material forms against callose wall; (*c*–*e*) ectexine (tectum, baculum, and foot layer) develops; and (*f*) endexine and intine appear. (*After P. Echlin, Pollen, Sci. Amer., 218:80–90, 1968*)

the endexine—composed of sporopollenin able to resist most acids and temperatures up to about 570°F (300°C). An inner fragile wall, the intine, surrounds the contents with its nuclei and reserves of starch and oil. Much of this detail can be seen in simple water or glycerin mounts at magnifications of 400× or less.

Only since the scanning electron microscope became commercially available have great advances been made with high resolution at very high magnifications, the range between 2000 and 12,000× being invaluable for pollen and spore work. However, because the scanning electron microscope registers only surface detail, the transmission electron microscope is used for interpretation of subsurface features far beyond the range of the light microscope. Used optimally in sequence on the same specimen, these two instruments have opened a new era of understanding of both living and fossil material.

The detail secured in studies of developing hellebore pollen grains is exemplified in **Fig. 2**, which should be compared with **Fig. 3** for a wider view of the callosic wall in which grains are embedded. In Fig. 2*a* and *b* is shown the initiation, after meiosis, of the waxy sporopollenin (biopolymer) walls with the deposition of the primexine matrix of fibrillar material in which the rod-layer elements develop. Growth of these little rods (known as columellae or bacula) extend below to form a "foot" layer and above (the limit of the ectexine) to form a "roof" (tectum), which may be perforate, as in Fig. 2*c*–*f*. **Figure 4** shows various polarity and surface patterns, as revealed by the scanning electron microscope. *See* SCANNING ELECTRON MICROSCOPE.

Diagnostic characters. Identification depends on interpretation of morphological features. Exine and aperture patterns are especially varied in the more highly evolved dicots, so that recognition at family, genus, or even species level may be possible despite the small surface area available on a grain. Since the morphological characters are conservative in the extreme, usually changing very slowly through geologic time, studies of fine detail serve to establish the lineal descent of many plants living today (**Fig. 5**). Ephemeral parts of the grain may also be useful in

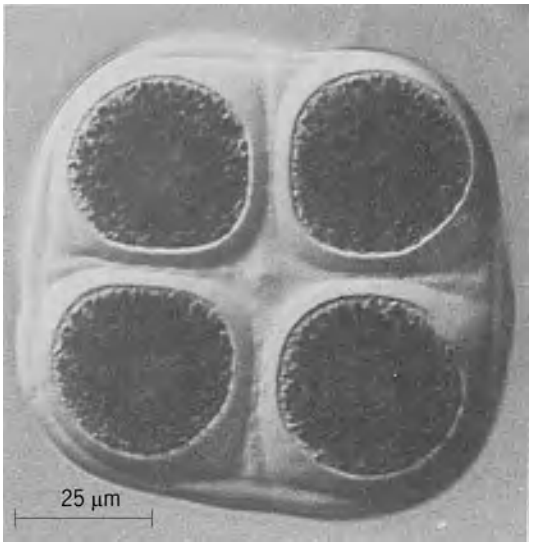

Fig. 3. Young tetrad of *Lavatera* (mallow family), with grains free in callose of pollen mother cell. (*Photomicrograph by Luc Waterkeyn*)

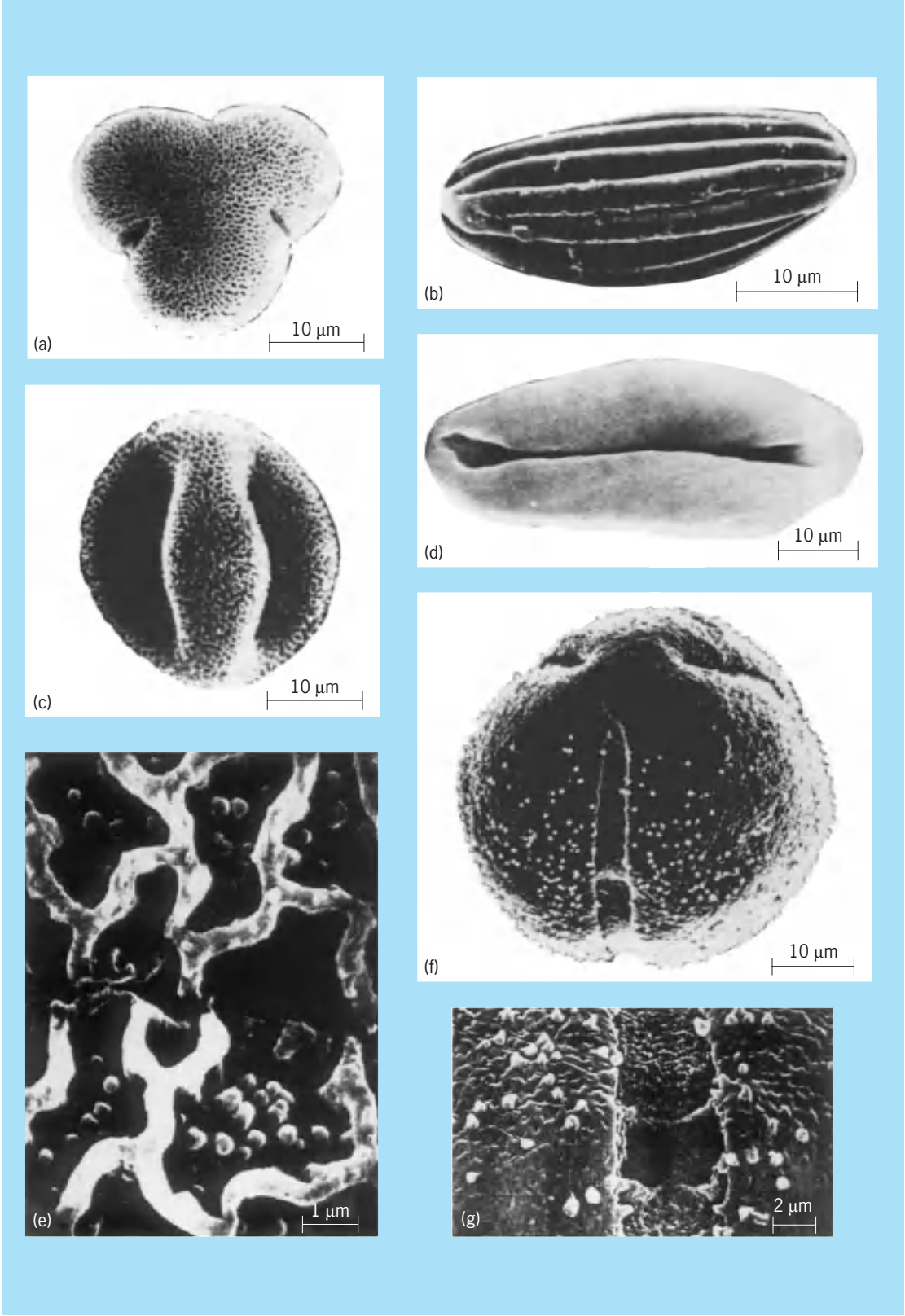

Fig. 4. Scanning electron micrographs of pollen grains, showing polarity and surface patterns. (*a*,*c*) Polar and side (equatorial) views of a reticulate, tricolpate *Homalanthus polyandrus* (spurge family) grain. (*b*,*d*) One-furrowed (monosulcate) grains with short polar axis: (*b*) ridged (plicate) grain of *Ephedra trifurca;* (*d*) smooth grain of nikau palm *(Rhopalostylis sapida)* with strongly defined sulcus. (*e*) Portion of exine of *Pelea barbigera* (orange family) showing an open network of walls (muri) supported by rods; free rods stand on the "floor" enclosed by the muri. (*f*) Tricolporate grain of *Scaevola kauaiensis* (Goodenia family), slightly oblique view showing the furrow, equatorial germ pore, smooth polar area, and scattering of spinules on perforate tectum. (*g*) Detail of sulpture and pore shape of *S. kauaiensis* pollen. (*Scanning electron micrographs a, c, and d by R. S. Thompson and L. M. Cranwell; b by C. M. Drew; e, f, and g by J. V. Ward*)

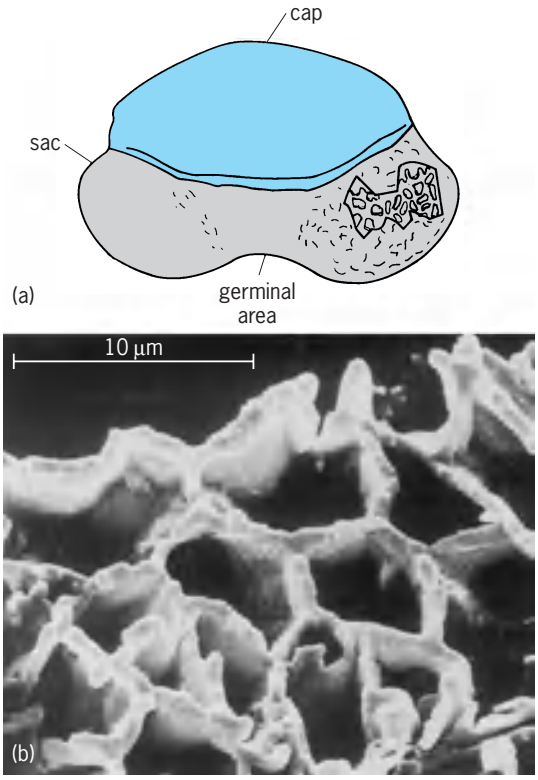

Fig. 5. Subfossil *Pinus* grain, corroded in soil. (*a*) Sketch showing cap and two sacs in side view, with right sac broken. (*b*) View into damaged right sac, which lacks its tough outer wall, to show honeycomb (alveolate) ectexine (sexine) structure in the cavity; note absence of rods (columellae). (*Scanning electron micrograph by C. M. Drew*)

diagnosis (**Fig. 6***a*), excessive thickening of the intine under the pores being a distinctive character of certain wind-pollinated types, as in hazel and birch, while hyaline plugs that swell to help split the aperture membranes are found in oak and beech. Color, though mainly yellow, may also be distinctive: the range from whitish to orange, green, or blackish is dramatic in the lily family. The walls of the grain, however, provide the most useful characters.

Ornamentation. Few grains are completely smooth (psilate), although they are so in some palms. Whether with or without a tectum, the sculptural elements occur on the surface. In subtectate grains, where narrow muri represent the tectum, a pattern of short rods may be exposed on the floor between the walls, whose longer rods (columellae) function as supports, like pilings under a wharf. Common patterns include echinate (*Hibiscus*); reticulate; microechinate (*Nothofagus*); smooth; foveolate (pitted); baculate (rods untapering, as in *Franklandia*); clavate or pilate (rods expanded at top, as in *Ilex*); and verrucate (warty). A striking stellate reticulate pattern of low muri topped by wedge-shaped clavae occurs in *Croton* and *Aleurites* (spurge family); and an identical pattern distinguishes *Stellatopollis*, which dates from the Early Cretaceous period.

Size. Extreme variations may occur within a family, but pollen grains range mainly from 15 to 50 micrometers, with the dicot range being from 2 μm in *Myosotis* to 250 μm in *Mirabilis*. A unique East African *Crossandra* has three-furrowed grains up to about 600 μm $\times$ 19 μm. Terrestrial monocots range from about 15 to 300 μm or more, and a few aquatic types are exceptionally long, with eelgrass (*Zostera*) having pollen measuring 2550 μm $\times$ 3.7 μm in a class of its own; living gymnosperms range from 15 μm (*Gnetum*) to about 180 μm in *Abies* (including sacs), while fossil types range from about 11 μm to 300 μm.

Symmetry and shape. In most pollen grains the polar axis runs from the inner (proximal) face to the outer (distal) face, as oriented during tetrad formation. The equator crosses it at right angles. Bilateral grains dominate in the gymnosperms and monocots, the polar

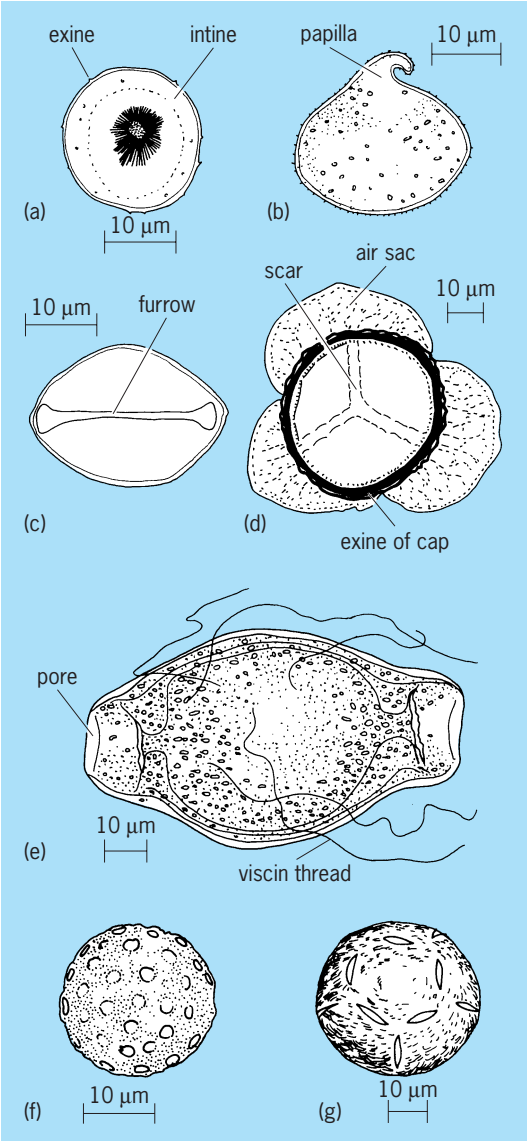

Fig. 6. Some varied pollen types of (*a–d*) gymnosperms and (*e–g*) angiosperms. (*a*) Inaperturate (*Cupressus arizonica*). (*b*) Papillate (*Sequoia sempervirens*). (*c*) Monosulcate (*Zamia floridiana*), a cycad. (*d*) Young three-saccate grain of kahikatea (*Dacrycarpus dacrydioides*) showing rounded germinal area (distal) and vestigial trilete scar on opposite (proximal) side of grain. (*e*) Diporate dicot type (*Fuchsia excorticata*) with viscin threads. (*f*) Many-pored grain of glasswort (*Salicornia australis,* goosefoot family). (*g*) Grain of *Heimerliodendron* (four-o'clock family) with 15 apertures.

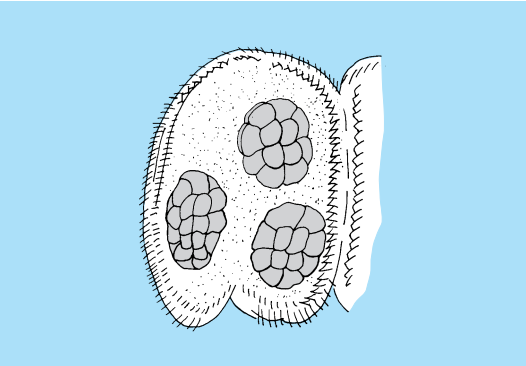

Fig. 7. Three polyads, each a symmetrical unit of 16 grains, lying free in an open anther lobe of *Albizzia julibrissin* (mimosa family).

axis usually being the shorter one, with the single aperture on the distal side. On the other hand, almost all dicot pollen is symmetrical around the polar axis (usually the long axis), with shapes ranging mainly from spheroidal to ellipsoidal, with rounded equatorial outlines, and sometimes "waisted" as in the Umbelliferae. Three-pored grains often have strikingly triangular outlines in polar view. Proximal tetrad scars (three-armed), so common in lower plants, occur freely in ancient conifer grains, but are only rarely traced in those of modern genera (Fig. 6*d*) and much more rarely in dicots, such as in the tropical *Cardiospermum*.

Aggregation. Most grains are free (monads) though often loosely grouped because of the spines, sticky oils, or viscin threads, as in *Fuchsia* (Fig. 6*e*), which are laid down in the pollen mother cell. It is suggested that the sticky "pollenkitt" may ensure germination of the pollen tube by enabling recognition of a compatible stigmatic surface. Compound grains (polyads) are richly developed in some angiosperm families, commonly occurring in fours (tetrads) as in *Catalpa*, many species of *Epilobium*, and most of the Ericales; or in multiples of four (polyads, **Fig. 7**) up to 64 or more (as in orchid pollinia). Twinned types (dyads) are rare. Permanent association in tetrads is rare in gymnosperms, but may occur freely in *Classopollis*, a fossil type.

Apertures. Rarely lacking, the flexible membranes of apertures are sometimes covered only by endexine. They allow for changes in volume and for the emergence of the pollen tube. They are classified as furrows (colpi), with elongate outlines, and pores usually more or less circular in shape (Figs. 1, 6*f*). Short slitlike intermediate forms occur (Fig. 6*g*). A few families have no apertures (for example, Lauraceae), some as a result of reduction as an adaptation to wind or water pollination involving loss of exine. Gymnosperm (for example, cypresses) may also lack openings, or have one small papilla as in *Sequoia* (Fig. 6*b*). Most have one furrow as in cycads (Fig. 6*c*), or a long weak sulcoid distal area (sulcus) as in some conifers (**Fig. 8***g* and *h*).

Angiosperm pollen exhibits a rich variety of furrow and pore types, with the association of pore and furrow (colporate forms) evolving in the dicots in the Albian. The first known tricolpate grains are Aptian in age.

Almost all grasses have one pore and almost all palms have one furrow. Dicot genera are typically tricolpate, triporate, or tricolporate, but may have numerous configurations (for example, 15 short apertures symmetrically arranged, as in Fig. 6*g*). The number of pores ranges from 1 to about 90. Rarely, paired pores occur, one above and one below the equator, and they are either free, as in *Anacolosa*, or set in a furrow, as in *Ribes* (gooseberry family). A few pore types, such as those of grasses (Fig. 8*f*) and gourds (**Fig. 9***e*), have an operculum (lid).

Dispersal. Grains of small to medium size (15 to 40 μm) travel highest and farthest; theoretically, they could circle in the stratosphere almost forever. In fact, most fall close to the parent plant or are soon washed out of the air by rain, hail, or snow. In studies of the settling velocity of certain Southwest desert types, it has been found that jojoba (*Simmondsia*) and *Ephedra* pollen falls, or is trapped by the female flowers, very close to the male plants. For *Ephedra*, it has been reported that 95% of the grains blow less than 11 ft (3.5 m) from their source. It is all the more remarkable, then, that some South American *Ephedra* and *Nothofagus* pollen does reach the subantarctic islands. The known limit of dispersal, 4700 mi (7500 km), was obtained from studies of peat on Marion Island in the Indian Ocean. In these studies, a few types blown 1100 mi (1800 km) from the southern tip of Africa were recovered.

Pollen also travels in rivers and ocean currents. J. G. Kölreuter in 1766 was the first to subdivide pollen according to windborne (anemophilous) and insect-borne (entomophilous) types. The former usually have small volume and little or no sculpture (Figs. 6, 8, and 9), while the latter range from very small to very large and heavy (Figs. 8*e* and 4*d* and *e*) and tend to "clump" because of their oily and sticky surfaces.

The number of grains per pollen sac is usually much lower in entomophilous (about 30 in *Mirabilis*) than in anemophilous types, which have

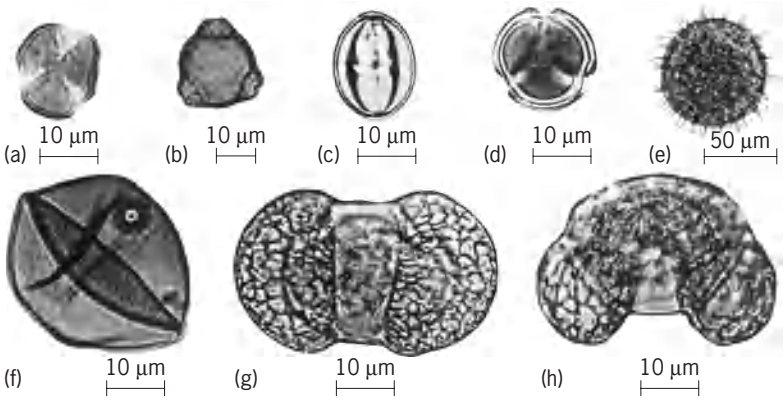

Fig. 8. Dicot aperture types, (*a*) Tricolpate (*Hebe rapensis*). (*b*) Triporate (*Casuarina stricta*) (*photomicrograph by Isobel C. Cookson*). (*c*) Tricolporate (*Griselinia littoralis*), side view, unexpanded; (*d*) polar view, intine present, expanded. (*e*) Periporate (*Hibiscus trionum*, mallow family). (*f*) Monocot type, one-pored (*Poa flabellata*, grass family). (*g*) Conifer type (*Podocarpus spicatus*), polar view, and (*h*) equatorial view.

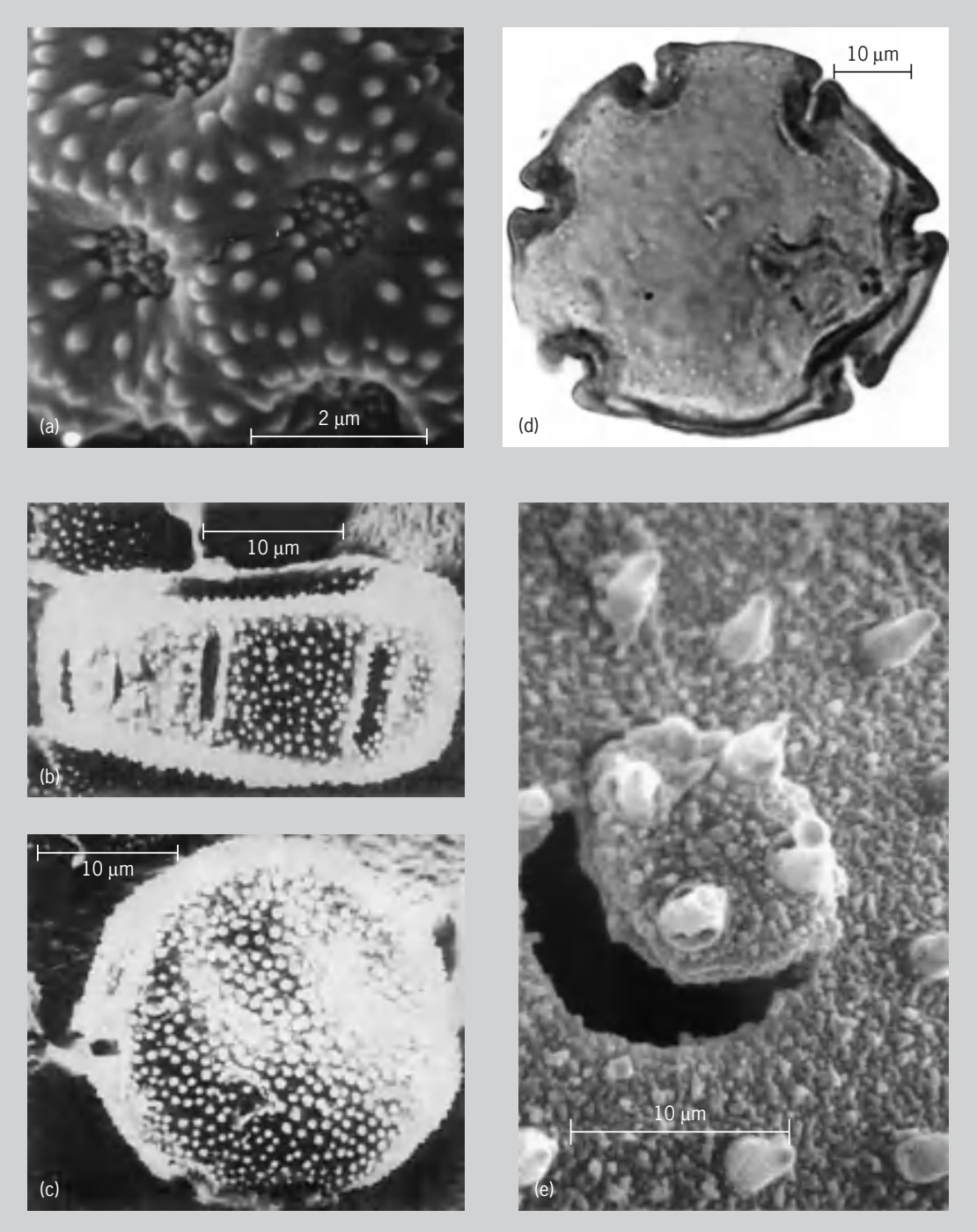

Fig. 9. Some anemophilous dicot types and an entomophilous type of pollen grains. (*a*) Surface of many-pored Russian thistle (*Salsola kali*), grain with very small spinules. (*b*, *c*) Modern and (*d*) fossil *Nothofagus fusca*–type grains notable for endexinal thickening around apertures; (*b*, *c*) *N. truncata*, side and polar views, with spinules dense and bold and apertures slitlike (thickenings hidden under surface); (*d*) fossil Antarctic grain with very thick rings from the Oligocene. (*e*) Spiny insect-carried grain of the coyote gourd (*Cucurbita palmata*) showing a portion of exine with pore lid detached. (*Stereoscan electron micrographs a–d and e by C. M. Drew*)

spendthrift yields that form "sulfur clouds" over forests and thick "blooms" on lakes. A single male cone of *Araucaria* may produce 10^9 grains; those of pine 1.5×10^6. Spruce forests of southern Sweden alone produce something like 75,000 tons (68,000 metric tons) of pollen each year, estimated at about 10^{21} grains.

Grasses yield vastly and are often geared to shed their pollen for a few minutes at a precise hour of day or in warm or cool weather only. In very cold, wet climates, as in South Georgia Island, the anthers may not open and may be preserved full of pollen in layers of peat. A few monocot species with thread-like pollen shed under water (for example, *Zostera*), while a few others raise their flowers to shed in the air (for example, *Ruppia*). Almost all gymnosperms

and most trees, shrubs, and herbs causing hay fever are wind-pollinated.

"Bat flowers" may also yield rich stores of pollen, well exposed on trunks and branches, as in *Kigelia pinnata*, the sausage tree. Such flowers are usually dusky or pallid, and may be odd-smelling; the bats are drawn to them after dark. Hummingbirds, seeking nectar, are also common pollinators, especially in tropical parts of the Americas.

Most flowering plants are solely either wind- or insect-pollinated, but there are odd deviants such as wormwood and sagebrush (*Artemisia* species) in the daisy family, whose pollen is otherwise almost exclusively carried by insects. Among the maples, box elder (*Acer negundo*) relies wholly on wind, while Norway maple (*A. platanoides*) depends on insects; the other species may combine both means of dispersal. The more ancient beech family, with fossil pollen dating back to the Upper Cretaceous, is about half wind- and half insect-pollinated, the oaks (*Quercus*) and southern beeches (*Nothofagus*), both wind-pollinated, now being widespread in the Northern and Southern hemispheres, respectively. Pollen of the latter genus is so distinctive (Fig. 9*b–d*) that it can be recognized with ease in deposits which are up to about 8×10^7 years old. Grasses and sedges, although now predominantly anemophilous, may produce nectar and thus may be reverting, in part, to an earlier condition. *See* POLLINATION.

Pollen record. First, in the Upper Carboniferous, over 3×10^8 years ago, appeared simple, often one-furrowed pollen, followed closely by types with one encircling sac, and later two separate sacs. Many were produced by gymnosperms, soon to supplant the giant lycopods and horsetails and dominate world vegetation for well over 2×10^8 years. Grains with three sacs had a Lower Cretaceous peak, especially in Gondwana countries. Today, pines and their allies (mainly in the Northern Hemisphere) and podocarps (mainly in the Southern) continue to shed vast crops of pollen, mainly bisaccate. Only a few western Pacific relict genera now have trisaccate grains, some large, as in the New Zealand *Dacrycarpus dacrydioides* (Fig. 6*d*), some small, as in the Tasmanian *Microcachrys*. Cycads, which are native to old land surfaces today, retain the shape and single furrow (Fig. 6*c*) of the ciliate sperm of their very ancient line of descent.

The question of the nature and rise of the earliest flowering plants (that is, the angiosperms) is being penetrated by modern and fossil studies, but it is still not known where their roots lie in the past.

Angiosperm characters typify *Clavatipollenites hughessi* (**Fig. 10**), a pollen type from Barremian and Aptian beds in England. It is now known that in their history of 1.15×10^8 years the four living genera of the chloranth family share the same diversity of aperture types and the same minute supratectate spinules and granules of their ancient forerunners in the Lower Cretaceous.

With the finding of varied monocot pollen in the eastern United States, the first firm dicot and monocot records have been proved to be more or less coeval. These ancient monocots (leaves associated) with larger, also reticulate grains have been described from basal Potomac beds. Their coarse and varied mesh patterns are typical of many modern insect-pollinated monocots, among them the modern lilies.

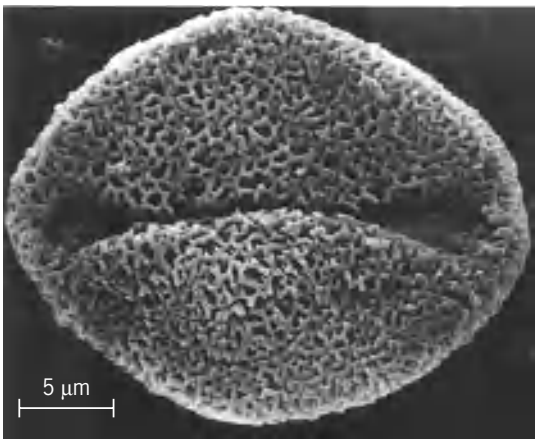

Fig. 10. Polar view of an Aptian *Clavatipollenites* grain, showing the long furrow and surface pattern (as in living *Ascarina*). From the Arundel clay of the Atlantic coastal plane, United States. (*Scanning electron micrograph by J. V. Ward*)

Also in the Aptian in a few countries appear a very few small dicot types with three furrows, much as in living *Homalanthus* (spurge family; Fig. 4*a* and *b*). It is clear, then, that while the flowering plants came in very gradually, by the Late Cretaceous they had richly diversified their aperture and sculpture types. However, many of the dominants of present landscapes, such as the grasses with their single pore (Fig. 8*f*), the chenopods with their many pores (Figs. 6*f* and 9*a*), and members of the daisy family (usually tricolporate), made slow starts in the early Tertiary before surging to their present importance, along with the more than 250,000 other species of flowering plants now known.

The structure of the exine walls has been found to fall into three basic categories: (1) columellar—restricted to angiosperms (except for *Classopollis*, an extinct gymnosperm); (2) alveolar (spongy)—characteristic of cycads and saccate conifers (Fig. 5*a* and *b*); and (3) granular—as in *Ephedra* and the spinulose *Gnetum*, but in only a few dicots, such as certain magnoliids. These advances mean that many of the earliest dicots can be identified as such, whereas their pollen might earlier have been confused with that of the smooth-walled gymnosperms. Unfortunately, the taxa of the third group can be separated only through transmission electron microscopy, which reveals a laminated endexine (below the granules) in the gymnosperms and an unlaminated endexine (or none at all) in the few known dicot examples. Lucy M. Cranwell

Bibliography. S. Blackmore and I. K. Ferguson (eds.), *Pollen and Spores: Form and Function*, 1986; K. Faegri and P. Kaland, *Textbook of Pollen*

Analysis, 4th ed., 1989, reprint, 2000; W. H. Lewis, P. Vinay, and V. E. Zenger, *Airborne and Allergenic Pollen of North America,* 1983; P. D. Moore, J. A. Webb, and M. E. Collinson, *Pollen Analysis,* 2d ed., 1992; J. Muller, Fossil pollen records of extant angiosperms, *Bot. Rev.* (Lancaster), 47:1–142, 1981; C. Newman, Breath of life and sneezes, *Nat. Geog.,* 166:490–520, 1984; K. J. Niklas, Wind pollination: A study in controlled chaos, *Amer. Sci.*, 73:462–470, 1985; A. Traverse, *Paleopalynology,* 1988.

Pollination

The transport of pollen grains from the plant parts that produce them to the ovule-bearing organs, or to the ovules (seed precursors) themselves. In gymnosperms ("plants with naked seeds"), including conifers and cycads, the pollen, usually dispersed by the wind, is simply caught by a drop of fluid excreted by each freely exposed ovule. In angiosperms ("flowering plants"), where the ovules are contained in the pistil, the pollen is deposited on the pistil's receptive end (the stigma), where it germinates. *See* FLOWER.

Ideally, each pollen grain forms a pollen tube which grows down through the slender style (stalk of the pistil) toward one of the ovules enclosed in the basal part or ovary. Here one of the two sperm cells within the pollen tube fuses with the egg cell of the ovule, making possible the development of an embryo, while the other sperm cell combines with two so-called polar nuclei of the ovule, thus starting the formation of endosperm, a tissue rich in reserve food. After this double fertilization, the ovule enlarges and becomes transformed into a seed.

Without pollination, there would be no fertilization; it is thus of crucial importance for the production of fruit crops and seed crops. Pollination also plays an important part in plant breeding experiments aimed at increasing crop production through the creation of genetically superior types. *See* BREEDING (PLANT); PLANT REPRODUCTION.

Self- and cross-pollination. In most plants, self-pollination is difficult or impossible, and there are various mechanisms which are responsible. For example, in dichogamous flowers, the pistils and stamens reach maturity at different times; in protogyny, the pistils mature first, and in protandry, the stamens mature before the pistils. Selfing is also impossible in dioecious species such as date palms and willows, where some plants bear flowers that have only pistils (pistillate or female flowers), while other individuals have flowers that produce only pollen (staminate or male flowers). In monoecious species such as hazel, where pistillate and staminate flowers are found on the same plant, self-breeding is at least reduced. Heterostyly, as exemplified by certain primroses, is another device that promotes outbreeding. Here some flowers (pins) possess a long pistil and short stamens, while others (thrums) exhibit the reverse condition; each plant individual bears only pins or only thrums. Pollen transport from stamens to pistils of the same length is easy, while in the other cases it is difficult.

Flower attractants. As immobile organisms, plants normally need external agents for pollen transport. These can be insects, wind, birds, mammals, or water, roughly in that order of importance. In some plants, for example, certain arum lilies, Dutchman's-pipes, and orchids, the pollinators are simply trapped; in the large majority of cases, however, the flowers offer one or more rewards, such as sugary nectar, oil, solid food bodies, perfume, sex, an opportunity to breed (as in figs), a place to sleep (*Serapias* orchids), or some of the pollen itself. For the attraction of pollinators, flowers provide either visual or olfactory signals. Color includes ultraviolet, which is perceived as a color by most insects and at least some hummingbird species. Fragrance is characteristic of flowers pollinated by bees, butterflies, or hawkmoths, while carrion or dung odors are produced by flowers catering to certain beetles and flies. A few orchids (such as *Ophrys* and *Cryptostylis* species), using a combination of olfactory and visual signals, mimic females of certain bees or wasps so successfully that the corresponding male insects will try to mate with them, thus achieving pollination (pseudocopulation).

At close range, pollinating insects are in many cases guided to the nectar by special color patterns, lines, or dots known as nectar or honey guides. Olfactory honey guides and osmophores (particular flower parts specialized for odor production) may be equally important. For better evaporation of odors, certain flowers may also develop considerable heat through a special process known as thermogenic respiration (certain arum lilies and water lilies).

Animal pollinators. While some flowers are "generalists," catering to a whole array of different animals, others, such as *Ophrys*, are highly specialized, being pollinated by a single species of insect only. Extreme pollinator specificity is an important factor in maintaining the purity of plant species in the field, even in those cases where hybridization can easily be achieved artificially in a greenhouse or laboratory, as in most orchids. The almost incredible mutual adaptation between pollinating animal and flower which can frequently be observed exemplifies the idea of coevolution.

Bird pollination. Hummingbirds have good red vision and feed on the wing. Flowers pollinated by these birds are often orange and red and hang down freely or are exposed (*Fuchsia*); they are sturdy, lack a landing platform, are open in the daytime, have little or no odor, and offer a not-too-concentrated nectar. Plants with flowers pollinated by birds other than hummingbirds (such as sunbirds in Africa) usually provide perches.

Bat pollination. Flowers pollinated by bats include those of the sausage tree, saguaro cactus, cup-and-saucer vine, and certain bananas. They are open in the evening, when they emit a rank or musty odor; their petal colors are either murky and nondescript or snow white; they are wide-mouthed and often

bell-shaped; and they normally offer both nectar and pollen in abundance. In connection with the poorly developed sonar sense of pollinating bats, these flowers often hang down freely from ropelike twigs (flagelliflory), or are on the main stem or limbs of the tree (cauliflory). These phenomena, however, may also be connected with the dispersal, by bats, of the fruits and seeds, even though the pollinators may have been different animals.

Hawkmoth pollination. Flowers pollinated by hawkmoths have a great deal in common with those pollinated by hummingbirds, but they are white or light-colored and very fragrant, especially in the evening when they are open.

Butterfly pollination. Flowers pollinated by butterflies open in the daytime, are usually colorful and fragrant (butterfly bush, or *Buddleia*; lilacs), and offer either a flat landing space (*Impatiens*) or, when small, are combined in inflorescences providing flattish surfaces to be walked on (red valerian, *Buddleia*). The narrow flower tubes allow the slender butterfly proboscis to go through, but discriminate against other insects.

Fly and beetle pollination. Both flies and beetles are highly diversified groups, which makes it difficult to describe "typical" fly- or beetle-pollinated flowers. Trap flowers exploit types such as carrion and dung beetles (or flies) which are not at all adapted to flowers. However, many modern flowers cater to animals with highly specialized mouthparts, such as certain longhorn beetles and the bee-fly *Bombylius*.

Bee pollination. By far the most important of all insect pollinators are the social bees, especially honeybees (*Apis*) and bumblebees (*Bombus*). While the former are subtropical in origin, bumblebees are adapted to cold climates. The larvae of both *Apis* and *Bombus* live on pollen and honey (a modified nectar); the adults drink nectar.

Among the physical and behavioral adaptations the adults possess to obtain these foods are long tongues, hairy bodies, special pollen baskets on the hindlegs, and a good color sense. The four broad color regions which bees distinguish, according to Karl von Frisch, are yellow, blue-green, blue, and ultraviolet. They learn very quickly to associate colors or minty odors with food, and remember the link for several weeks. These traits, and the bees' excellent time sense, which enables them to visit flowers when these have the most to offer, are important in the establishment of flower constancy: the bees stay with one, or very few, crop plants only, which is important for seed production.

Honeybees can distinguish various sugars. In contrast to bumblebees, which are "rugged individualists," honeybees have an excellent communication system: a worker bee can indicate both the direction and the distance of a rich food source by means of special dances, so that the resource can be exploited cooperatively. Flowers pollinated by bees, for example, true sage, bleeding heart, and Scotch broom, generally have bright colors, minty odors, and a landing platform; they often display bilateral symmetry. Their pollen grains are sticky, spiny, or highly sculptured for better adherence to the bees.

Wind pollination. Wind pollination is prevalent in the primitive cycads and in conifers, but it is also found in such familiar flowering plants as grasses, rushes, sedges, cattails, plantains, nettles, hemp, sorrel, lamb's-quarters, alder, birch, hazel, poplars, and temperate-region oaks. It offers advantages in regions where animal pollinators are scarce or where these cannot operate due to adverse environmental or seasonal conditions, such as strong winds. Wind-pollinated flowers are usually inconspicuous, being devoid of attractants and rewards for the pollinators, such as showy petals, fragrance, and nectar.

To facilitate exposure to the wind, blooming often takes place in early spring before the leaves have emerged, or the flowers may be placed very high on the plant. In a number of species, the anthers burst open in an explosive fashion, scattering the pollen widely; the stigmas often are long and divided into arms or lobes, or, like those of some willows and cereals, are feathery, which facilitates pollen catching. The latter process is also promoted because, in open areas, wind-pollinated plants of one species often grow together in dense populations. The pollen grains, which are shed singly, are dry, glueless, smooth-surfaced, and in some cases (pine, cedar) provided with air sacs. Wind pollination is a wasteful process because it lacks precision. For example, while one male plant of *Mercurialis annua*, a common weed, may produce 1.25×10^9 pollen grains, there is only a slim chance that an individual stigma will be hit by more than one or two pollen grains. Undoubtedly in connection with this, the number of ovules per pistil and (later) the number of seeds per fruit is very low (two ovules per flower in beech, hazel, oak, and walnut; one per flower in birch, grasses, elm, stinging nettle, and sweet gale).

Water pollination. Although the pollen grains of some plants can be made to germinate in aqueous sugar solutions, water alone in most cases will destroy them. Accordingly, water is seldom the true medium of pollen dispersal, even in aquatic plants. In ribbon weed (*Vallisneria*) the male and female flowers are released to the surface separately. There, male flowers are driven by the wind until they bump into female flowers; the shock causes the pollen to be catapulted over to the stigma. True water dispersal of pollen (hydrophily), in which the grains are wetted, is found only in the fresh-water hornworts (*Ceratophyllum*) and marine eelgrasses (*Zostera, Phyllospadix*).

Bastiaan J. D. Meeuse

Ecology. Traditionally pollination biology has considered mechanisms that promote or guarantee the proper transfer of pollen from one individual of a given plant species to another. Often this involves a careful description of the transfer mechanism coupled to a particular pollinating agent (such as hawkmoths, bumblebees, bats, or wind). Pollinator-plant interactions, however, have also been used as a model for understanding broader evolutionary and

ecological issues that pertain to all living systems. Pollination systems have proven to be an excellent model for studying many current issues in population biology.

Since plants have a variety of breeding structures, they provide an opportunity for assessing the reproductive value of different sexual strategies and for modeling the evolution of breeding systems and sexual allocation. A classic problem in animal populations is the number of males and females a mother should produce as offspring (the sex ratio). Hypotheses accounting for the evolution of the sex ratio can be tested by using dioecious plant species, and speculations on sex-ratio evolution can be far more complicated in plants due to their striking variety in methods of sexual allocation (such as gynodioecy, androdioecy, or hermaphrodism).

If pollen movement is considered as a measurable analog of gene movement, then pollination systems are ideal for unraveling patterns of gene flow and the concomitant genetic structuring of plant populations. Pollen transfer can be measured by numerous methods. Gene flow (or the lack of it) has been implicated in retarding (or promoting) local population adaptation to different habitats as well as affecting rates of species formation. The traditional and the ecological approaches to pollination studies are distinguishable by their different views on the important consequences of pollen transfer. The traditional view stresses the mechanism of transfer within individual flowers or plants, while the ecological view stresses the population consequences of transfer among many individuals, even across communities or habitats. *See* POPULATION ECOLOGY; POPULATION GENETICS.

Pollinators are often faced with having to make complicated decisions about when, where, and on what flowers to forage. They therefore have been used as models for testing hypotheses concerning animal behavior such as patterns of learning, the acquisition of information from the environment, or alteration of behavior associated with changes in the environment.

The complex interaction among plants and their pollinators has prompted research on community-level ecology as well as population-level approaches. The timing of flowering within given habitats may reflect competitive interactions among plants for pollination, and competition may also affect the kinds of pollinators found in a habitat and their pattern of resource use. The evolutionary ecology of many plants and pollinators may be best understood as a result of complex interactions among populations of many species. *See* ECOLOGICAL COMMUNITIES; ECOLOGY; POLLEN; POPULATION DISPERSAL. Leslie Real

Bibliography. S. Blackmore and I. K. Ferguson, *Pollen and Spores: Form and Function*, 5th ed., 1990; H. Hartmann, D. Kester, and F. Davies, *Plant Propagation: Principles and Practices*, 6th ed., 1996; R. P. Kapil (ed.), *Pollination Biology: An Analysis*, 1986; E. Ottaviano, M. S. Gorla, and G. Bergamini-Mulcahy (eds.), *Angiosperm Pollen and Ovules*, 1992.

Polonium

A chemical element, Po, atomic number 84. Marie Curie discovered the radioisotope ^{210}Po in pitchblende. This isotope is the penultimate member of the radium decay series. All polonium isotopes are radioactive, and all are shortlived except the three α-emitters, artificially produced ^{208}Po (2.9 years) and ^{209}Po (100 years), and natural ^{210}Po (138.4 days). *See* PERIODIC TABLE.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Polonium (^{210}Po) is used mainly for the production of neutron sources. It can also be used in static eliminators and, when incorporated in the electrode alloy of spark plugs, is said to improve the cold-starting properties of internal combustion engines.

Most of the chemistry of polonium has been determined using ^{210}Po, 1 curie of which weighs 222.2 micrograms; work with weighable amounts is hazardous, requiring special techniques. Polonium is more metallic than its lower homolog, tellurium. The metal is chemically similar to tellurium, forming the bright red compounds $SPoO_3$ and $SePoO_3$. The metal is soft, and its physical properties resemble those of thallium, lead, and bismuth. Valences of 2 and 4 are well established; there is some evidence of hexavalency. Polonium is positioned between silver and tellurium in the electrochemical series.

Two forms of the dioxide are known: low-temperature, yellow, face-centered cubic (UO_2 type), and high-temperature, red, tetragonal. The halides are covalent, volatile compounds, resembling their tellurium analogs. *See* RADIOACTIVITY; TELLURIUM.

Kenneth W. Bagnall

Bibliography. P. W. Atkins et al., *Inorganic Chemistry*, 4th ed., 2006; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; D. Sangeeta, *Inorganic Materials Chemistry Desk Reference*, 1997.

Polyacetal

A polyether derived from aldehydes ($R{=}CH_2$ or CHR') or ketones [$R{=}C(R'R'')$] and containing —O—R—O— groups in the main chain. Of the many possible polyacetals, the most common is a polymer or copolymer of formaldehyde, polyoxymethylene

$(—O—CH_2—)_n$. While the substance paraformaldehyde contains oligomers or low-molecular-weight polyoxymethylenes (n very small), high-molecular-weight, crystalline polyoxymethylenes constitute an important class of engineering plastics that, in commerce, is often simply referred to as polyacetal. Cellulose and its derivatives also have a polyacetal structure. For discussion of other polyethers and polymers containing acetal side groups. *See* ACETAL; CELLULOSE; FORMALDEHYDE; POLYETHER RESINS; POLYVINYL RESINS.

Chemistry. The carbonyl group is the only unsaturated linkage besides the carbon-carbon double bond that has been successfully polymerized by a chain reaction on a commercial scale. As shown in reaction (1), formaldehyde can be readily polymerized by

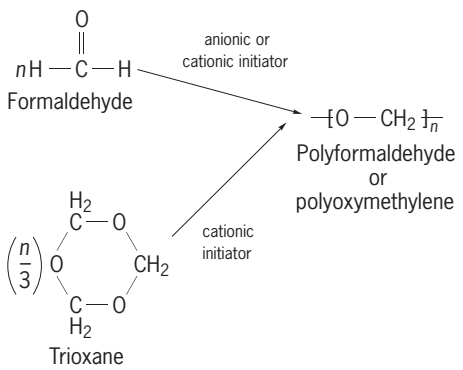

(1)

using anionic initiators such as triphenylphosphine and, somewhat less readily, by using cationic initiators such as protonic acids. Alternatively, a similar polymer can be obtained by the ring-opening polymerization of trioxane using, for example, a boron trifluoride complex as initiator.

At temperatures above ~110°C (230°F; the ceiling temperature, above which depolymerization becomes favored over polymerization), the polymers as formed degrade by an unzipping reaction to reform monomer. To prevent this, one of two approaches is commonly used: esterification of the hydroxyl end groups, or copolymerization with a small amount of a monomer such as ethylene oxide or 1,3-dioxolane. Polymerization using gamma radiation has also been reported.

In its linearity, polyoxymethylene resembles polymethylene; the molecules pack well and crystallize readily. However, the polarity associated with the —C—O— linkage results in a higher crystalline melting point than that of polymethylene [~195°C (383°F) versus ~135°C (275°F)]. Other aldehydes such as acetaldehyde or trichloroacetaldehyde can be polymerized at low temperatures (to avoid aldol condensation) and, preferably, high pressures by generally similar techniques, but the ceiling temperatures of the products are very low [−35°C (−31°F) and 13°C (55°F), respectively]. Even acetone has been polymerized on a small scale to yield the polyketal $[—O—C(CH_3)_2—]_n$, by using metallic or organometallic catalysts. The polymerization of sulfur analogs, for example, trithiane (the analog of trioxane), hexafluorothioacetone, and thiocarbonyl fluoride is also possible; these monomers are reactive to free-radical as well as ionic polymerization. *See* FREE RADICAL.

The successful synthesis of high-molecular-weight polyoxethylene on a commercial scale has been achieved after considerable earlier research on the synthesis of fiber-forming polyacetals by the classical reaction of formaldehyde (in the form of paraformaldehyde) with aliphatic diol, as in reaction (2), where R must contain more than five car-

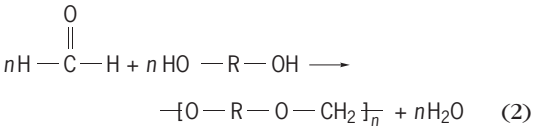

(2)

bons to avoid cyclization. An interchange reaction between a diol and an acetal yields a similar structure. The relatively low melting points obtained (60°C or 140°F) can be increased by use of a cycloaliphatic diol, for example, 1,4-cyclohexanediol.

The versatility of polyacetal formation is also evident in the polymerization of diverse monomers such as cyclic formals, as in reaction (3). Various

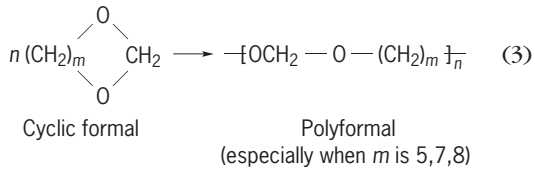

(3)

spiroacetals have been based on reactions of aldehydes or ketones with pentaerythritol, for example, reaction (4). Some of these polymers are unique in

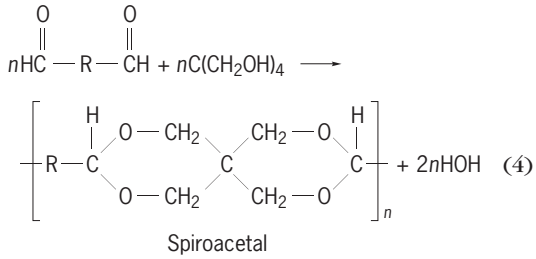

(4)

that expansion, rather than contraction, occurs on polymerization. *See* POLYMER; POLYMERIZATION.

Properties and uses. Polyacetals are typically strong and tough, resistant to fatigue, creep, organic chemicals (but not strong acids or bases), and have low coefficients of friction. Electrical properties are also good. Improved properties for particular applications may be attained by reinforcement with fibers of glass or polytetrafluoroethylene, and by incorporation of an elastomeric toughening phase. The combination of properties has led to many uses such as plumbing fittings, pump and valve components, bearings and gears, computer hardware, automobile body parts, and appliance housings.

John A. Manson

Bibliography. J. A. Brydson, *Plastics Materials,* 7th ed., 1999; G. W. Ehrenstein and R. P. Theriault,

Polymeric Materials: Structure, Properties, Applications, 2000; G. Odian, *Principles of Polymerization*, 4th ed., 2004; A. B. Strong, *Plastics: Materials and Processing*, 3d ed., 2005.

Polyacrylate resin

Polymers obtained from a variety of acrylic monomers, such as acrylic and methacrylic acids, their salts, esters, and amides, and the corresponding nitriles. The most important monomers are:

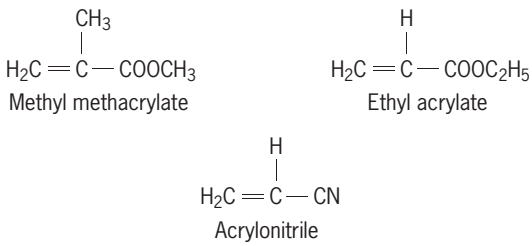

Poly(methyl methacrylate), poly(ethyl acrylate), and a few other derivatives are discussed in this article. *See* POLYACRYLONITRILE RESINS.

Poly(methyl methacrylate) is a hard, transparent polymer with high optical clarity, high refractive index, and good resistance to the effects of light and aging. It and its copolymers are useful for lenses, signs, indirect lighting fixtures, transparent domes and skylights, dentures, and protective coatings.

The monomer is usually made by the dehydration and methanolysis of acetone cyanohydrin as in reaction (1). Polymerization may be initiated

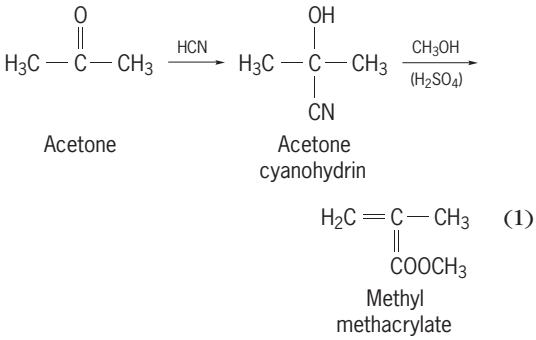

(1)

by free-radical catalysts such as peroxides, or by organometallic compounds such as butyl lithium. The free-radical polymerization, which is used for commercial production, can be carried out in bulk, in solution, and in aqueous emulsion or suspension. By changing the catalyst, temperature, and solvent in a polymerization it is possible to prepare polymers which are predominantly isotactic, syndiotactic, or atactic, or which have stereoregular blocks. Although solution polymerization is not commonly used, bulk polymerization, usually with comonomers, is frequently employed in various casting operations, as in the formation of sheets, rods, and tubes; in the mounting of biological, textile, and metallurgical test specimens; and in dental applications. A 20% reduction in volume accompanies the conversion of monomer to polymer. This makeit difficult to prepare articles to predetermined dimensions by the cast polymerization technique. The problem is minimized by using a syrup of polymer dissolved in monomer, or a dough of finely divided polymer dispersed in a relatively small amount of monomer.

Molding powders suitable for the injection molding of dials, ornamental fixtures, and lenses may be prepared from the granules produced by aqueous suspension polymerization or from the product of bulk polymerization.

Solutions of poly(methyl methacrylate) and its copolymers are useful as lacquers. Aqueous latexes formed by the emulsion polymerization of methyl methacrylate with other monomers are useful as water-based paints and in the treating of textiles and leather. *See* EMULSION POLYMERIZATION.

Poly(ethyl acrylate) is a tough, somewhat rubbery product. The monomer is used mainly as a plasticizing or softening component of copolymers. Ethyl acrylate is usually produced by the dehydration and ethanolysis of ethylene cyanohydrin, which can be obtained from ethylene oxide as shown by reaction (2).

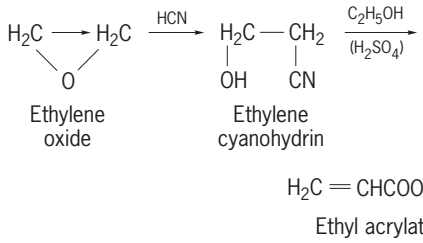

(2)

Polymerization may be affected by catalysts of the free-radical type. Copolymerizations with other monomers are frequently carried out in aqueous emulsion or suspension.

Modified acrylic resins with high impact strengths can be prepared. Blends or "alloys" with poly(vinyl chloride) are used for thermoforming impact-resistant sheets. Copolymers of ethyl acrylate with 2-chloro-ethyl vinyl ether are useful rubbers which can be vulcanized in the same way as ordinary polydienes. *See* COPOLYMER.

Methyl methacrylate is of interest as a polymerizable binder for sand or other aggregates, and as a polymerizable impregnant for concrete; usually a cross-linking acrylic monomer is also incorporated. The binder systems (polymer concrete) are used as overlays for bridge decks as well as for castings, while impregnation is used to restore concrete structures and protect bridge decks against corrosion by deicing salts.

Other acrylic derivatives are the subject of continuing interest. The butyl and octyl esters of acrylic acid yield rubbery materials, and fluorinated esters have been used as oil-resistant rubbers. Polymers of alkyl-2-cyano-acrylates make excellent adhesives that cure at low temperatures using adventitious moisture as a catalyst. Cross-linkable acrylic copolymers are often used in surface coatings. Polymers of methyl acrylate or acrylamide are water-soluble and useful for sizes and finishes. Addition of poly(lauryl methacrylate) to petroleum lubricating oil improves

the flowing properties of the oil at low temperatures and the resistance to thinning at high temperatures. *See* ACRYLONITRILE; PLASTICS PROCESSING; POLYMERIZATION; RUBBER. John A. Manson

Bibliography. H. Allcock, F. Lampe, and J. Mark, *Contemporary Polymer Chemistry*, 3d ed., 2003; R. G. Gilbert, *Emulsion Polymerization: A Mechanistic Approach*, 1995; G. Odian, *Principles of Polymerization*, 4th ed., 2004; M. P. Stevens, *Polymer Chemistry: An Introduction*, 3d ed., 1998.

Polyacrylonitrile resins

The polyacrylonitrile resins are hard, relatively insoluble, and high-melting materials. Polyacrylonitrile [poly(vinyl cyanide)] is used almost entirely in copolymers. The copolymers fall into three groups: fibers, plastics, and rubbers. The presence of acrylonitrile in a polymeric composition tends to increase its resistance to temperature, chemicals, impact, and flexing. *See* COPOLYMER.

Acrylonitrile is generally prepared by one of the following methods.

Catalyzed addition of hydrogen cyanide to acetylene as in reaction (1):

$$CH{\equiv}CH + HCN \xrightarrow{\text{catalyst}} H_2C{=}\underset{\underset{CN}{|}}{CH} \qquad (1)$$

Reaction between hydrogen cyanide and ethylene oxide as in (2):

$$\underset{\underset{O}{\diagdown\ \diagup}}{H_2C{-}CH_2} + HCN \longrightarrow H_2C{=}\underset{\underset{CN}{|}}{CH} + HOH \qquad (2)$$

Reaction between ammonia and propylene as in (3):

$$2H_3C{-}CH{=}CH_2 + 2NH_3 + 3O_2 \longrightarrow 2H_2C{=}\underset{\underset{CN}{|}}{CH} + 6H_2O \qquad (3)$$

The polymerization of acrylonitrile can be readily initiated by means of the conventional free-radical catalysts such as peroxides, by irradiation, or by the use of alkali metal catalysts. Although polymerization in bulk proceeds too rapidly to be commercially feasible, satisfactory control of a polymerization or copolymerization may be achieved in suspension and in emulsion, and in aqueous solutions from which the polymer precipitates. Copolymers containing acrylonitrile may be fabricated in the manner of thermoplastic resins.

The major use of acrylonitrile is in the form of fibers. By definition an acrylic fiber must contain at least 85% acrylonitrile; a modacrylic fiber may contain less, from 35 to 85% acrylonitrile. The high strength; high softening temperature; resistance to aging, chemicals, water, and cleaning solvents; and the soft woollike feel or fabrics have made the product popular for many uses such as sails, cordage, blankets, and various types of clothing. Commercial forms of the fiber probably are copolymers containing minor amounts of other vinyl derivatives, such as vinyl pyrrolidone, vinyl acetate, maleic anhydride, or acrylamide. The comonomers are included to produce specific effects, such as improvement of dyeing qualities. *See* MANUFACTURED FIBER.

Copolymers of vinylidene chloride with small proportions of acrylonitrile are useful as tough, impermeable, and heat-sealable packaging films.

Extensive use is made of copolymers of acrylonitrile with butadiene, often called NBR (formerly Buna N) rubbers, which contain 15–40% acrylonitrile. Minor amounts of other unsaturated esters, such as ethyl acrylate, which yield carboxyl groups on hydrolysis, may be incorporated to improve the curing properties. The NBR rubbers resist hydrocarbon solvents such as gasoline, abrasion, and in some cases show high flexibility at low temperatures. *See* RUBBER.

In the 1960s development of blends and interpolymers of acrylonitrile-containing resins and rubbers represented a significant advance in polymer technology. The products, usually called ABS resins, typically are made by blending acrylonitrile-styrene copolymers with a butadiene-acrylonitrile rubber, or by interpolymerizing polybutadiene with styrene and acrylonitrile. Specific properties depend upon the proportions of the comonomer, on the degree of graftings, and on molecular weight. In general, the ABS resins combine the advantages of hardness and strength of the vinyl resin component with toughness and impact resistance of the rubbery component. Certain grades of ABS resin are used for blending with brittle thermoplastic resins such as poly(vinyl chloride) in order to improve impact strength.

The combination of low cost, good mechanical properties, and ease of fabrication by a variety of methods, including typical metalworking methods such as cold stamping, led to the rapid development of new uses for ABS resins. Applications include products requiring high impact strength, such as pipe, and sheets for structural uses, such as industrial duct work and components of automobile bodies. ABS resins are also used for housewares and appliances and, because of their ability to be electroplated, for decorative items in general. *See* ACRYLONITRILE; PLASTICS PROCESSING; POLYMER; STYRENE. John A. Manson

Bibliography. H. Allcock, F. Lampe, and J. Mark, *Contemporary Polymer Chemistry*, 3d ed., 2003; A. W. Birley, R. H. Heath, and M. J. Scott, *Plastics Materials: Properties and Applications*, 2d ed., 1988; J. Brydson, *Plastics Materials*, 7th ed., 1999; G. Odian, *Principles of Polymerization*, 4th ed., 2004.

Polyamide resins

The product of polymerization of amino acids or the condensation of a diamine with a dicarboxylic acid. Polyamides are a class of polymers in which individual structral units are joined by amide bonds ($R{-}CO{-}NH_2$). Whitish, translucent,

and high-melting polyamide resins can be essentially transparent and amorphous when their melts are quenched. On cold drawing and annealing, most become quite crystalline and translucent. However, some polyamides based on bulky repeating units are inherently amorphous. The polymers are used for fibers, bristles, bearings, gears, molded objects, coatings, and adhesives. The term nylon formerly referred specifically to synthetic polyamides as a class. Because of many applications in mechanical engineering, nylons are considered engineering plastics. *See* MANUFACTURED FIBER.

Brief outlines of the preparations of commercial polyamides by the reaction of dicarboxylic acids with diamines and the condensation of amino acids are given in this article. The most common commercial aliphatic polyamides are nylons-6,6; -6; -6,10; -11; and -12. Other types find specialized application.

Nylon-6,6 and nylon-6,10 are products of the condensation reaction of hexamethylenediamine (6 carbon atoms) with adipic acid (6 carbon atoms), and with sebacic acid (10 carbon atoms), respectively. By heating equimolar proportions of the two reactants, a polymeric salt is formed, as in reaction (1),

$$\underset{\text{Adipic acid}}{n\mathrm{HOOC(CH_2)_4COOH}} + \underset{\text{Hexamethylenediamine}}{n\mathrm{H_2N(CH_2)_6NH_2}} \longrightarrow \underset{\text{"Nylon-6,6 salt"}}{[\overset{+}{\mathrm{N}}\mathrm{H_3(CH_2)_6}\overset{+}{\mathrm{N}}\mathrm{H_3}\overset{-}{\mathrm{O}}\mathrm{OC(CH_2)_4}\overset{-}{\mathrm{C}}\mathrm{OO}]_n} \xrightarrow{\text{heat}} [\mathrm{HN(CH_2)_6NHOC(CH_2)_4CO}]_n + 2n\mathrm{H_2O} \quad (1)$$

which, on further heating, yields the polyamide resin shown. Because the end groups on the polymer can react on further heating, as in melt spinning, it is desirable to add a very small amount of a monoacid or monoamine to the polymerizing mixture in order to prevent the formation of material of very high molecular weight.

Nylon-6,6, nylon-6,10, nylon-6,12, and nylon -6 are the most commonly used polyamides for general applications as molded or extruded parts; nylon-6,6 and nylon-6 find general application as fibers.

Nylon-6 and nylon-11 can be obtained by the self-condensation of ϵ-aminocaproic acid and ω-amino undecanoic acids, respectively, as shown by reactions (2) and (3).

$$\underset{\epsilon\text{-Aminocaproic acid}}{\mathrm{H_2N(CH_2)_5COOH}} \rightarrow \text{nylon-6} \quad (2)$$

$$\underset{\omega\text{-Aminoundecanoic acid}}{\mathrm{NH_2(CH_2)_{10}COOH}} \rightarrow \text{nylon-11} \quad (3)$$

Each molecule, containing both the amino and carboxylic groups, can condense to yield high polymers by reactions similar to those between the diacids and the diamines.

Nylon-6 is usually made by polymerization of the lactam of ϵ-aminocaproic acid, as in reaction (4).

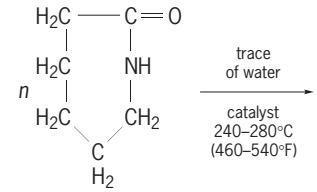

ϵ-Aminocaprolactam

$$\underset{\text{Nylon-6}}{(\mathrm{-\overset{H}{N}-(CH_2)_5-\overset{O}{\overset{\|}{C}}-})_n} \quad (4)$$

Caprolactam (a common term for ϵ-aminocaprolactam) is produced by rearrangement of cyclohexanone oxime, which is prepared by reaction of cyclohexanone with hydroxylamine sulfate.

The lactam process is generally preferred for commercial operation because it is easier to make and purify the lactam than the ϵ-aminocaproic acid. Nylon-12 is prepared by a similar self-condensation of the lactam of aminododecanoic acid, which is prepared by a series of reactions using butadiene as starting material. Lactams may also be polymerized by anionic catalysis. This process makes possible a one-step casting to form large articles, especially from nylon-6.

As a group, nylons are strong and tough. Mechanical properties depend in detail on the degree and distribution of crystallinity, and may be varied by appropriate thermal treatment or by nucleation techniques. Because of their generally good mechanical properties and adaptability to both molding and extrusion, the nylons described above are often used for gears, bearings, and electrical mountings. Nylon bearings and gears perform quietly and need little or no lubrication. Sintering (powder metallurgy) processes are used to make articles such as bearings, and gears which have controlled porosity, thus permitting retention of oils or inks. Nylon resins are used extensively as filaments, bristles, wire insulation, appliance parts, and film. Properties can be modified by copolymerization.

Reinforcement of nylons with glass fibers results in increased stiffness, lower creep and improved resistance to elevated temperatures. Such formulations, which can be readily injection-molded, can often replace metals in certain applications. The use of molybdenum sulfide and polytetrafluoroethylene as fillers increases wear resistance considerably. For uses requiring impact resistance, nylons can be blended with a second, toughening phase.

Other types of nylon are useful for specialty applications. Solubility may be increased by interference with the regularity and hence intermolecular packing. This may be accomplished by copolymerization, or by the introduction of branches on the amide nitrogen, for example, by treatment with formaldehyde. The latter type of resin may be subsequently cross-linked. Nylons incorporating aromatic structures, for example, based on isophthalic acid, are becoming more common for applications requiring resistance to very high temperatures.

Aromatic-nylon fibers based on all-para-substituted monomers can be spun from liquid-crystal suspensions to yield fibers that have exceptional moduli and strengths up to approximately 25,000,000 lb/in.^{2} (160 gigapascals) and 400,000 lb/in.^{2} (2.8 GPa), respectively, in the fiber direction. These values may be compared with theoretical moduli and strengths of approximately 40,000,000 lb/in.^{2} (270 GPa) and approximately 1,000,000–4,000,000 lb/in.^{2} (6.9–27 GPa), respectively.

Heterocyclic amide-imide copolymers are also available commercially, and a wide variety of other amide-containing polymers, such as polyesteramides and polysulfonamides, can be synthesized.

The unsaturated fatty acids in vegetable oil, for example, linoleic acid, may dimerize or polymerize to low polymers through their unsaturated groups. The di- or polycarboxylic acids obtained yield polyamides by condensation with di- or polyamines. These products are employed as coatings and adhesives, especially in epoxy and phenolic resin formulations. *See* HETEROCYCLIC POLYMER; PLASTICS PROCESSING; POLYETHER RESINS; POLYMERIZATION.

John A. Manson

Bibliography. H. Allcock, F. Lampe, and J. Mark, *Contemporary Polymer Chemistry*, 3d ed., 2003; J. Brydson, *Plastics Materials*, 7th ed., 1999; R. Puffr and V. Kubanek, *Lactam-based Polyamides*, vols. I and II, 1991; G. Odian, *Principles of Polymerization*, 4th ed., 2004.

Polychaeta

The largest class of the phylum Annelida (multisegmented worms) with about 13,000 described species. Polychaetes (**Fig. 1**) range from less than 1 mm (0.04 in.) to up to 3 m (10 ft) in length. Represented worldwide, they are predominantly marine organisms and occur from the intertidal zone to the greatest ocean depths at all latitudes. Some notable exceptions occupy nonmarine habitats such as supralittoral zones (beyond the high-tide mark) and freshwater lakes. The oldest known entire fossilized species dates back to the mid-Cambrian, but records on trace fossils suggest that the origin of the group predates this period. The only synapomorphy (a derived trait shared by two or more taxa that is believed to reflect their shared ancestry) that distinguishes them from other annelids is the pair of nuchal organs, chemosensory structures

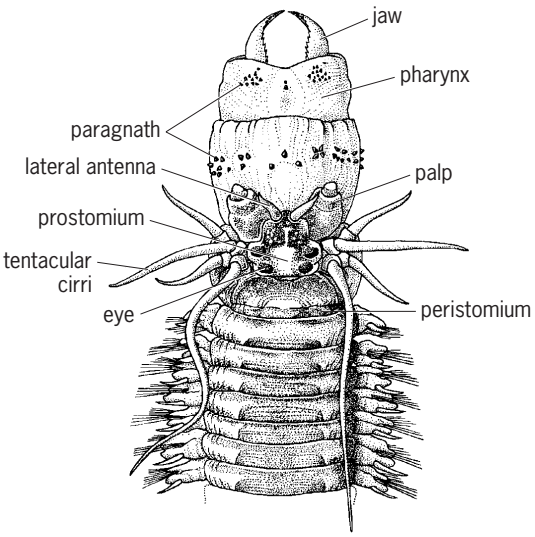

Fig. 2. ***Platynereis dumerilii antipoda*, showing the antero-dorsal end with the pharynx everted. (*After P. L. Beesley et al., 2000*)**

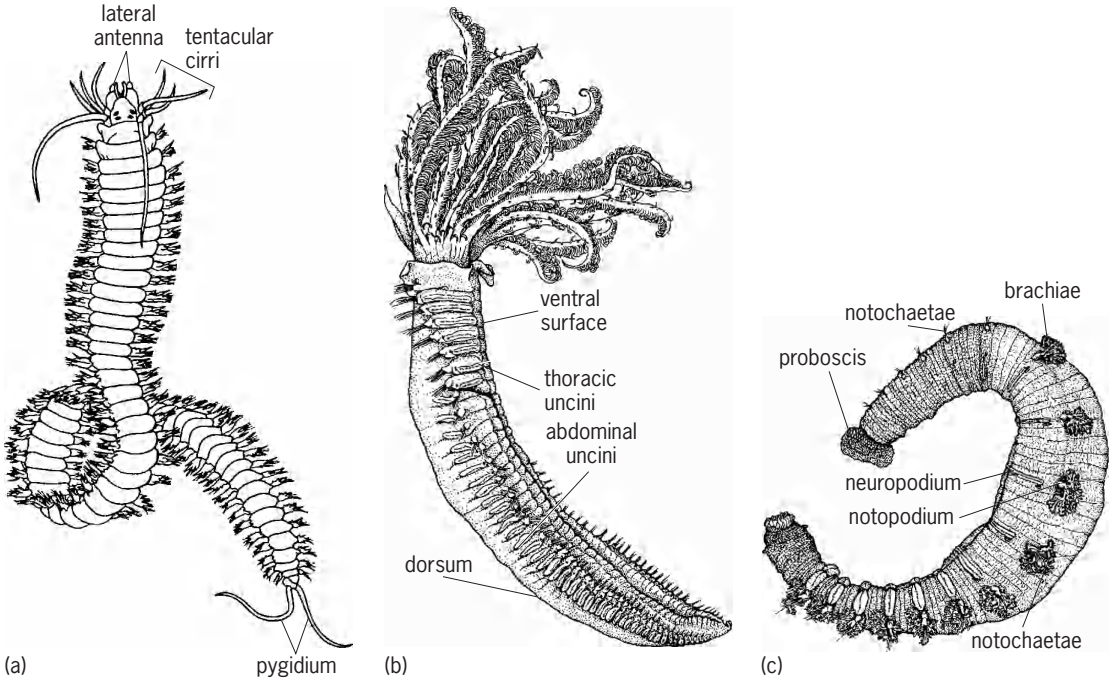

Fig. 1. Examples of polychaetes. (*a*) *Platynereis dumerilii antipoda*. (*b*) *Branchiomma* sp. (*c*) *Arenicola bombayensi*. (*After P. L. Beesley et al., 2000*)

located on the prostomium (the portion of the head anterior to the mouth). *See* ANNELIDA; DEEP-SEA FAUNA.

External morphology. As in other annelids, the polychaete body consists of two presegmental regions (the prostomium and the peristomium), a segmented trunk, and a postsegmental pygidium (tail region) [**Fig. 2**]. Often the prostomium and the peristomium are fused with the first anterior segments of the trunk, known as the cephalized segments, forming the head of the polychaete. The prostomium varies greatly in shape, from a simple conical structure to a simple T-shape, and may be fused to the peristomium. In many polychaetes, the only peristomial structure visible is the area surrounding the mouth; in a few others, the peristomium is seen as one complete ring separating the prostomium from the first segment.

Head appendages, which are present or absent according to the species considered, include antennae, palps (various sensory, usually fleshy appendages near the oral aperture), peristomial cirri (slender hairlike appendages), and tentacular cirri. Antennae (up to three) are always located on the prostomium. Palps emerge from the prostomium or the peristomium. They are structurally divided into grooved feeding palps and tapering sensory palps. A single pair of grooved palps is present in many polychaetes. In sabellids and serpulids, the grooved palps form a tentacular crown, with each tentacle being made of a long median radiole (spine) and numerous lateral pinnules (branchlets arranged in plumelike rows). Sensory palps are generally a pair of lateral, slender, tapering structures. Peristomial cirri are attached dorsolaterally at the level of the peristomium, while tentacular cirri emerge from the cephalized segments [**Figs. 3** and **4**].

The trunk of a polychaete comprises serially repeated segments, also called metameres, numbering from few to many. Segments may be similar to one another (homonomous metamerism) or different (heteronomous metamerism), resulting in the latter case in the elaboration of various trunk regions (thoracic, abdominal, caudal). Each segment basically bears a pair of lateral appendages, the parapodia (**Fig. 5**). A parapodium is equipped with musculature, and bears chaetae (setae) in most taxa. Parapodial structures vary, but they often consist of two rami (branches), a dorsal notopodium and a ventral neuropodium. In many polychaetes, the two rami differ in size, structure, and orientation. Dorsal and ventral cirri are distinctive parapodial features. Cirri are generally slender tapering structures, but in scaleworms some of the dorsal cirri are present as elytra (scales) that are held horizontally above the dorsal surface.

Chaetae are epidermal chitin-sclerotized (hardened) proteinaceous rods (like bristles) sometimes impregnated with calcium carbonate. Chaetae occur in bundles and emerge at the tip of neuropodia and notopodia. They are connected to internal muscles and can be retracted in the parapodium or protruded into the medium. They act in anchoring the parapodium to the substrate during locomotion. Many polychaetes have internal chaetae, the aciculae, which form the internal skeleton or supporting rods for the parapodia. Spines and hooks are thicker chaetae present in some polychaetes. The shape of the chaetae is highly diversified and used in specific diagnosis.

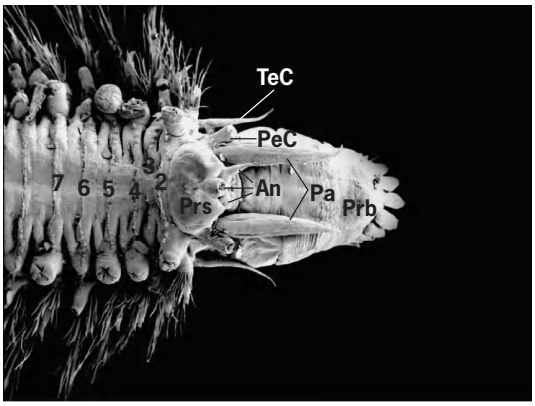

Fig. 3. ***Hololepidella laingensis*; scanning electron microscope view of the anterior end with the pharynx everted. An: antennae; Pa: palps; Prb: proboscis; Prs: prostomium; TeC: tentacular cirrus. Numbers refer to the segments of the trunk. 50×.**

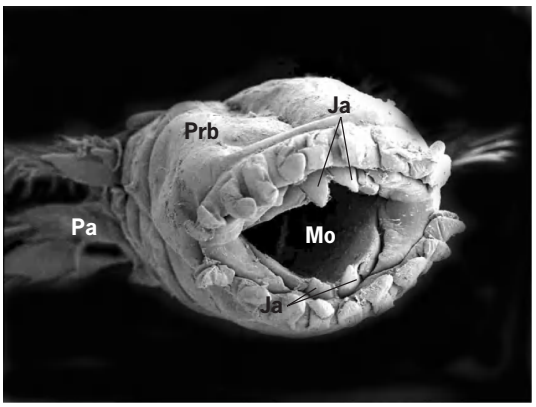

Fig. 4. ***Gastrolepidia clavigera*; scanning electron microscope view of the anterior end. Mo: mouth; Ja: jaw; Pa: palps; Prb: proboscis. 30×.**

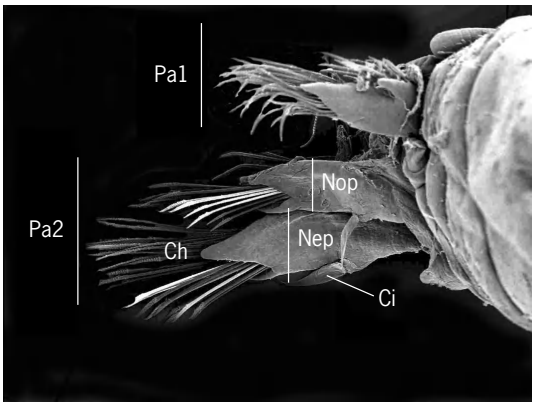

Fig. 5. ***Gastrolepidia clavigera*, scanning electron microscope view of the two first parapodia (Pa1, Pa2). Ch: chaetae; Ci: cirrus; Nop: notopodium; Nep: neuropodium. 40×.**

Internal morphology and physiology. Beneath the epidermis of polychaetes lies a nearly continuous sheath of circular muscle that is internally doubled by thick longitudinal muscles (four bands in most species). In addition, series of oblique muscles join the ventral area of the trunk to the midlateral regions. The innermost body wall component is a peritoneal layer lining the coelom. The coelom is segmented, and in most polychaetes, septa (partitions) separate internally adjacent metameres (**Fig. 6**).

The digestive tract is essentially a simple tube supported by a dorsal mesentery and partial-to-complete septa. The anterior part of the digestive tract is often differentiated in a buccal (mouth) organ that is axial or ventral, well developed or simple. The muscular region of the buccal organ is called the pharynx; the eversible nonmuscular region is often referred to as the proboscis. Feeding methods in polychaetes are diverse. Polychaetaes are raptorial (living on prey) feeders and use their buccal apparatus to seize food items; nonselective or selective deposit feeders; or filter feeders. The proboscis of raptorial feeders is well developed and often bears chitinous structures, the paragnaths (small, sharp jawlike structures) and jaws.

The great majority of polychaetes have a closed circulatory system made of medial dorsal and ventral longitudinal vessels linked by smaller capillaries and gut lacunae (cavities). Some taxa have more contractile hearts. Many species have elaborate branchiae (gills) developed as extensions of the body wall containing a loop of the vascular system and capillaries. Blood generally does not contain cells but respiratory pigments such as hemoglobins and chlorocruorins (green substances that are the probable cause of the color of the blood in some worms). Waste products of metabolism are excreted via nephridia (generally metanephridia). The distribution of nephridia varies, but most commonly nephridia are present as a pair per segment along most of the trunk. *See* EXCRETION.

The nervous system of polychaetes is composed of a brain, or cerebral ganglia, situated in the prostomium. The brain is dorsal and linked to a ventral nerve cord by two circumesophageal connectives. The ventral nerve cord is usually made up of a coupled pair of cords running the length of the trunk. Polychaetes have several sensory structures that include palps, antennae, cirri, eyes, statocysts, nuchal organs, and lateral organs. Statocysts are gravity receptors numbering from one pair to more than 20 pairs. They often appear as simple pits at the surface of the organisms. Nuchal organs are paired structures innervated from the dorsal part of the brain. They are most likely chemoreceptors, probably like the lateral organs that are present segmentally between the notopodia and the neuropodia. Most polychaetes have eyes or photoreceptors located on the prostomium. These structures range from simple pigmented cups with receptor cells to complex compound eyes analogous to those of arthropods. In certain families, eyes may be capable of forming images, but they are more probably structures integrating light intensity and direction.

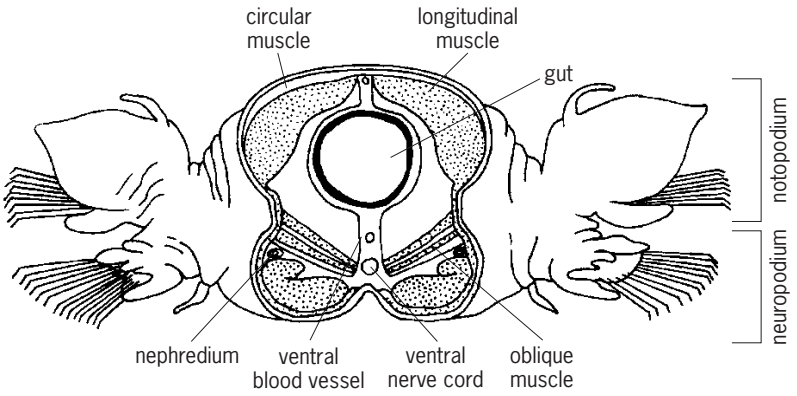

Fig. 6. Section through a polychaete. (*After E. E. Ruppert and R. D. Barnes, Invertebrate Zoology, Saunders College Publishing, Orlando, 1994*)

Gonochorism (separation of the sexes) is the most common sexual condition, although hermaphrodism (the presence of both male and female reproductive organs) is widespread. Most polychaetes have no permanent gonads. At some period of the year, spermatozoa and ova are differentiated from germ cells located in the peritoneum and are discharged into the coelom before being expulsed to the water medium through specialized gonoducts or nephridia (in a few cases by rupture of the body wall).

Life cycle and reproduction. Life cycle is either semelparous or iteroparous. Semelparity refers to breeding occurring once per lifetime. Iteroparity, where breeding occurs several times in a lifetime, can be divided into annual iteroparity (breeding yearly) and continuous iteroparity (breeding over an extended season). In most polychaetes, fertilization is external and the development gives rise to a pelagic trochophore (free-swimming) larva. Some obvious features of the larvae are the ciliary bands that are used for locomotion and feeding. After the pelagic larval stage, larvae of most species sink to the sea bottom and metamorphose into benthic juveniles that grow into adults. In a few cases, ova may be deposited in gelatinous or mucous balls or strands, sometimes covering the beaches during the spawning season. Ova may be attached to parapodia or laid in tubes, where brood care is displayed. Some onuphids construct special ovicells in which the young develop to a stage resembling adults. Special thoracic pouches or opercular chambers serve as ovaria in some serpulids.

Asexual reproduction is widespread and involves paratomy or architomy. Paratomy refers to the formation of a recognizable complete individual that separates from the parent piece; architomy (far more widespread) is simple fission or fragmentation of the body without prior cephalization (anterior specialization resulting in the concentration of sensory and neural organs in the head).

Synchronous spawning is widespread in polychaetes. In addition to temporal synchrony, many benthic species swarm to the surface of the water column to spawn. These species often undergo drastic morphological changes during swarming that

are called epitoky. One case of epitoky involves the Palolo worm, *Eunice viridis*, from the Samoa Islands, in which the worm separates from its posterior part during swarming.

Ecology. In marine habitats, polychaetes may burrow into the muddy or sandy substrates in both intertidal and subtidal areas. Some live naked in the sediments, burrowing freely into them; others live in tubes that range from semipermanent mucous types to well-constructed firm tubes of fine sediments. Tubicolous species are fairly immobile within sediments. Many others live under rocks and crevices, either free-living or with cemented tubes. Polychaetes such as the Christmas tree worm, *Spirobranchus giganteus*, are also found in live corals and can exhibit a wide range of colors even on a single coral.

Members of a few families are parasites. Some species actively bore into calcareous substrata: *Boccardia* and *Polydora* species are parasites of oysters and other bivalves, causing blisters to develop in the shell. All histriobdellids occur in the branchial chambers of marine and freshwater crustaceans. A species of ichthyomids are parasites found in the fins of eels. Several polychaetes are commensals (commensalism is a symbiotic relationship in which two different species are associated, wherein one is benefited and the other neither benefited nor harmed), especially polynoids that live in the tubes of terebellids or are found in association with holothuroids, asteroids, ophiuroids, antipatharian corals, or sponges. Some hesionids and flabelligerids are also found in echinoids.

Some nonmarine polychaetes, especially nereids, live in the supratidal zone. Others live in freshwater lakes and streams. Some ampharetids inhabit the inland waterways of central Europe, especially the Danube River and its tributaries. *Manayunkia speciosa* (sabellid) is frequent in the Great Lakes of the northeastern United States and their tributaries. Another species of *Manayunkia* has been observed in hypersaline ephemeral lakes (short-lived lakes which have no inlets/outlets and in which the water evaporates, is absorbed, or is consumed) in the wheat belt of Western Australia.

Geographic distribution. Polychaetes are distributed in all marine habitats and show remarkable specificity according to latitudes, depths, and kinds of substrata. Until recently, many polychaetes were regarded as being cosmopolitan (having worldwide distribution), and many species originally described from Arctic regions were also reported from Antarctica. Similarly, many species in Europe and in Australia were initially thought to be identical, but deeper analyses showed that many Australian species were new, suggesting that cosmopolitanism is less extended in polychaetes than previously thought.

In temperate intertidal zones, syllids, nereids, and polynoids predominate, whereas tropical intertidal zones have large numbers of amphinomids, eunicids, serpulids, and terebellids. Preponderant numbers of maldanids, spionids, and cirratulids are found in soft sediments in moderate depths, whereas abyssal depths of more than 6600 ft (2000 m) yield large numbers of ampharetids, flabelligerids, and other tubicolous kins. The diversity continues equally great from shallowest to greatest depths, but unit members diminish from littoral and upper slope, to deep slope and abyssal depths.

Any major zoogeographic area may expect to support from 350 to 800 species of polychaetes, coming from shallowest to greatest depth and in a variety of sediments. The number of endemic species may range from 25 to 40% of the total number; another 30% may be more widely distributed in adjacent areas; and 20 to 30% may be cosmopolitan. These percentages differ with latitude and place. The number of cosmopolitan species is lowest in high or polar latitudes and highest in warm to tropical latitudes, but such figures reflect the greater knowledge of tropical latitudes.

Economic significance. Some polychaetes are economically important to shipping. The calcareous tube-building *Hydroides norvegica* covers the bottoms of ships with thick layers of its tubes. Polychaetes are also used as bait by amateur anglers all over the world. In Britain, *Neanthes virens* is farmed for air-freighting around northern Europe.

With a few exceptions, polychaetes are harmless for humans, but the chaetae of amphinomids can injury people handling them. Some people develop local allergic reaction to these worms, called fireworms. Some other polychaetes can bite.

Classification. In the first half of the twentieth century, the taxon Polychaeta consisting of Archiannelida, Errantia, and Sedentaria was widely accepted. This system was used until recently, but various phylogenetic analyses made during the 1980s and 1990s provided evidence that the three groups are not reliable natural entities. Today, Polychaeta should be divided into two monophyletic groups, Scolecida and Palpata, with 81 accepted families. Palpata, which includes 72 families, are named for the presence of palps on the head. Palpata comprise two major clades, Canalipalpata and Aciculata. Aciculata are named for the presence of aciculae supporting well-developed parapodia. They include the Phyllodocida (27 families), Amphinomida (2 families), and Eunicida (8 families). Canalipalpata have grooved palps, usually peristomial in origin. They include the Sabellida (4 families), Spionida (8 families), and Terebellida (12 families). The Scolecida do not have palps but do have parapodia with similar-sized rami and two or more pairs of pygidial cirri; they include 9 families. Eleven other families are *incertae sedis* (of uncertain taxonomic position). Igor Eeckhaut; Olga Hartman

Bibliography. T. Bartolomaeus and G. Pürschke, *Morphology, Molecules, Evolution, and Phylogeny in Polychaeta and Related Taxa*, Springer, Dordrecht, 2005; P. L. Beesley, G. J. B. Ross, and C. J. Glasby, *Polychaetes and Allies: The Southern Synthesis—Fauna of Australia, Polychaeta, Myzostomida, Pogonophora, Echiura, Sipuncula*, CSIRO Publishing, Melbourne, 2000; G. W. Rouse and K. Fauchald, Cladistics and polychaetes, *Zool. Scripta*, 26:139–204, 1997; G. W. Rouse and F. Pleijel, *Polychaetes*, Oxford University Press, New York, 2002.

Polychlorinated biphenyls

A generic term for a family of 209 chlorinated isomers of biphenyl. The biphenyl molecule is composed of two six-sided carbon rings connected at one carbon site on each ring. Ten sites remain for chlorine atoms to join the biphenyl molecule. While the rules of nomenclature would indicate that the term polychlorinated biphenyl can refer only to molecules containing more than one chlorine atom, the term has been used to refer to the biphenyl molecule with one to ten chlorine substitutions.

Polychlorinated biphenyls (PCBs) were discovered in the late nineteenth century, were introduced into United States industry on a large scale in 1929, and have been in use since that time in most industrial nations. The qualities that made PCBs attractive to industry were chemical stability, resistance to heat, low flammability, and high dielectric constant. The PCB mixture is a colorless, viscous fluid, is relatively insoluble in water, and can withstand high temperatures without degradation. However, these characteristics are precisely the qualities that make polychlorinated biphenyl persistent in the environment; mono- and dichlorobiphenyl have half-lives in sunlight of about 6 days, but the higher-chlorinated isomers are not readily degraded in the environment.

Uses. The major use of PCBs has been as dielectric fluid in electrical equipment, particularly transformers (including transformers used on railway locomotives and self-propelled cars), capacitors, electromagnets, circuit breakers, voltage regulators, and switches. Electrical equipment containing PCBs is no longer produced in the United States, but in 1981 40% of all electrical equipment in use in the United States contained PCBs. PCBs have also been used in heat transfer systems and hydraulic systems, and as plasticizers and additives in lubricating and cutting oils. *See* DIELECTRIC MATERIALS.

Environmental distribution. PCBs have been reported in animals, plants, soil, and water all over the world. The fact that PCBs have been detected in the polar ice caps is an indication of the power of atmospheric transport as a major pathway of global contamination. PCBs are also transported through the biosphere in water and attached to sediments. It is not known what quantities of PCBs have been released to the environment, but major sources are industrial and municipal waste disposal, spills and leaks from PCB-containing equipment, and manufacture and handling of PCB mixtures. *See* ATMOSPHERIC GENERAL CIRCULATION; BIOSPHERE.

PCBs have been reported in fishes, eels, and benthic organisms in the waters of the North Sea near the Netherlands. Fish products from United States waters and the Baltic Sea have been found to contain several parts per million of PCB and have been declared unfit for human consumption. PCBs have been found in all organisms analyzed from the North and South Atlantic, even in animals living under 11,000 ft (3400 m) of water. These phenomena are the result of bioaccumulation and biomagnification in the food chain. Uptake of PCBs by plants has also been reported, but it is generally small and does not directly contribute to substantial human contamination. In a few instances, poultry products, cattle, and hogs have been found to contain high concentrations of PCBs after the animals have eaten feed contaminated with PCBs. *See* FOOD WEB.

Following the initial restriction of PCB production in the 1970s, contamination levels dropped rapidly in the Great Lakes. Since 1979, levels in the Great Lakes have ceased to decline, probably because of atmospheric transport, contaminated ground water, spills and leaks from authorized equipment, continued improper disposal, and the fact that the lower-chlorinated species degraded first, leaving the recalcitrant higher-chlorinated biphenyls. *See* EUTROPHICATION.

PCBs have been detected in human adipose tissues and in the milk of cows and humans. The estimated percentage of the United States population with detectable levels of PCBs was nearly 100% in 1981. The estimated percentage of the United States population with greater than 3 ppm PCBs in their tissue was 2.7% in 1972, and less than 1% in 1981. *See* HUMAN ECOLOGY.

PCBs have also been found as low-level, inadvertently generated contaminants in a wide range of chemical products such as certain paints, inks, plastics, and paper coatings. It has been estimated that fewer than 100,000 lb (45,000 kg) of PCBs are inadvertently generated in chemical processes each year in the United States.

Toxicity. PCBs can enter the body through the lungs, gastrointestinal tract, and skin, circulate throughout the body, and be stored in adipose tissue. Except for occupational contact, human exposure is mainly through food. PCBs can cause chloracne, skin discoloration, liver dysfunction, reproductive effects, development toxicity, and oncogenicity in exposed humans. Some PCBs have the ability to alter reproductive processes in mammals. Prenatal exposure in animals can result in various degrees of developmentally toxic effects. There is concern, based on extrapolation from animal studies, that PCBs may be carcinogenic in humans. The problem is complicated by the presence, in most PCB mixtures, of toxic impurities.

Disposal. Incineration of solid and liquid PCB materials is a common and highly efficient method of destruction. While laboratory tests have demonstrated PCB destruction at about 1470°F (800°C), the commercial incinerators operate at temperatures in excess of 1800°F (1000°C). High-efficiency destruction has been reported in large stationary incinerators and boilers, cement kilns, incinerator ships, and smaller mobile thermal destruction units.

PCB-contaminated fluids have been decontaminated by using chemical reagents to attack the chlorinated molecule. Generally, a metal, such as sodium, is used to remove the chlorine atoms from the biphenyl molecule. Some research has been conducted toward chemical treatment of PCB-contaminated soil by spraying the reagent on soil; the results and

environmental consequences are unclear. Microbial degradation of PCB-contaminated soils has been pursued by several investigators, but there is little evidence of commercial-scale success. *See* ECOLOGY; HAZARDOUS WASTE. Glenn Kuntz

Bibliography. M. Barros et al. (ed.), *PCB Seminar Proceedings*, Organization for Economic Cooperation and Development, 1983; W. Bouchard, PCBs in perspective, *Hazard. Mater. Manag. J.*, 1980; O. Hutzinger et al., *Chemistry of PCBs*, 1974, reprint 1983; A. Leifer et al., *Environmental Transport and Transformation of PCBs*, USEPA 560/5-83-025, 1983; R. Lucas et al., *PCBs in Human Adipose Tissue and Mother's Milk*, USEPA Contract 68-01-5848, 1982; S. Safe, *Polychlorinated Biphenyls (PCBs): Mammalian and Environmental Toxicology*, 1987.

Polycladida

An order of marine Turbellaria which are several millimeters to several centimeters in length and whose leaflike bodies have a central intestine with radiating branches. Most of these platyhelminth species live in the littoral zone on the bottom, on seaweed or on other objects, or as commensals in the shells of mollusks and hermit crabs. None are parasitic. Except in warm waters, they are seldom brightly colored.

Usually they have many eyes, and tentacles are frequently present. Frontal organs and statocysts are absent and adhesive organs are rare. The epidermis is covered with cilia and contains numerous rhabdite glands. Near the middle of the body is the mouth, followed by the plicate pharynx which opens into the central cavity of the intestine. The brain lies anterior to the pharynx, with a number of nerves radiating from it, but the two nerves which parallel the pharynx are usually the largest. Ovaries and testes are numerous and scattered, but yolk glands are lacking. Sperm ducts connect the testes with the copulatory apparatus which is variable in structure and often multiple rather than single. Insemination occurs through copulation or hypodermic impregnation. The entolecithal eggs usually accumulate in the oviducts or uteri and after fertilization pass to the exterior through the vagina. Lang's vesicle, a bursa with a long stalk, is generally present. Müller's larva, a free larval stage known in the Turbellaria, is found in some polyclads, but is lacking in most. This larva may indicate an evolutionary link between the Turbellaria and the Ctenophora, and between the Turbellaria and the Annelida.

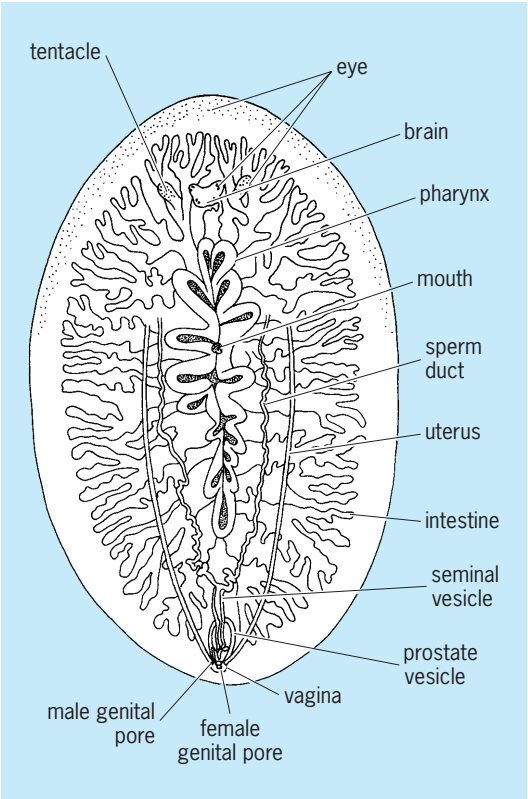

***Stylochus ellipticus*.**

Notoplana and *Stylochus* (see **illus.**) are two of the largest and best-known genera of polyclads with representatives from both coasts of North America. In particular, species of *Stylochus* are often large forms 2 in. (5 cm) or more in length. *See* TURBELLARIA. E. Ruffin Jones

Bibliography. H. G. Bronn (ed.), *Klassen und Ordnungen des Tierreichs*, vol. 4, 1930–1933.

Polyester resins

A large class of polymeric materials in which ester groups,

$$-\overset{\overset{\displaystyle O}{\|}}{C}-O-$$

are in the main chains. Generally, polyesters are prepared by the condensation reactions of glycols (or dialcohols) with dicarboxylic acids, and they range from relatively soft aliphatics to aromatic derivatives which are considered engineering materials. Their properties may be modified by cross-linking, crystallization, plasticizers, or fillers. *See* ESTER.

Commercial products include alkyds, which are used in paints and enamels and in molding compounds; unsaturated polyesters or unsaturated alkyds, which are used extensively with fiberglass for boat hulls and panels; aliphatic saturated polyesters; aromatic polyesters such as poly(ethylene terephthalate), which is used in the form of fibers and films, in blow-molded bottles, and in injection-moldable resins such as poly(butylene terephthalate); and aromatic polycarbonates.

The poly(diallyl esters) are frequently listed with the polyesters; however, they are not true polyesters as defined above.

Alkyds. The alkyds have been in common use as coatings since World War I. They were first commercialized by the General Electric Company as Glyptal, which became a common name. In the beginning, they consisted almost entirely of the reactions

products of *o*-phthalic anhydride and glycerol, and pigment. Because the functionality of the system is greater than two, a cross-linked insoluble polymer is formed as in reaction (1).

Fully cured alkyds are quite hard and brittle. Flexible and tough materials can be produced by incorporating monobasic acids or monohydroxy alcohols in proportions sufficient to increase flexibility but insufficient to prevent curing. Combinations of conventional vegetable drying oils and alkyd resins represent the basis of most of the oil-soluble paints. For example, by heating a mixture of dehydrated castor oil (the glycerol ester of linoleic acid) with suitable proportions of glycerol and phthalic anhydride, an oil-soluble polyester is formed. A common oil paint is produced by the addition of thinners, such as aromatic hydrocarbon solvents, a paint drier such as cobalt octoate, and pigments. By exposure to air in the presence of the paint dryer, the unsaturated diene groups of the linoleic ester polymerize to yield a tough, weather-resistant coating. *See* DRYING OIL.

The drying oil–alkyd described above may be further modified by the inclusion of a vinyl monomer, such as styrene, in the original esterification process. Some of the styrene polymerizes, probably as a graft polymer, and the remainder polymerizes and copolymerizes in the final drying or curing of the paint. Low-molecular-weight liquid condensates are useful as polymeric plasticizers for resins such as poly(vinyl chloride). Molding components are also used for such applications as appliance housings and electrical components. *See* COPOLYMER; PAINT AND COATINGS.

Unsaturated polyesters. The unsaturated polyesters were developed during and shortly after World War II. In combination with glass fiber, they found immediate applications as panels, roofing, radar domes, boat hulls, and protective armor for soldiers. Structural applications of this type continued to increase in importance over the years. The compositions are distinguished by ease of fabrication and high impact resistance. *See* POLYMERIC COMPOSITE.

A low-molecular-weight unsaturated polyester intermediate is first produced. The reaction of maleic anhydride with diethylene glycol is typical, as shown in reaction (2). The product is a viscous oil with a molecular weight of 2000–4000. Many other acids (or anhydrides) such as fumaric acid, and glycols such as propylene glycol may be reacted in a similar manner.

The low-molecular-weight unsaturated polyester will cross-link in the presence of a peroxide by copolymerization with styrene or other vinyl monomers. The unsaturated maleic group copolymerizes in essentially a 1:2 or 1:3 ratio with styrene. Therefore, several styrene molecules which react effectively join two ester chains together to yield an insoluble cross-linked structure, as in reaction (3).

The commercially available intermediate unsaturated polyesters usually contain about 30% styrene or other vinyl monomer. On addition of a peroxide or other free-radical catalyst and a paint drier, the copolymerization starts. In this stage, the resin may be handled as a viscous fluid for a few minutes to a few hours, depending upon the activity of the catalyst. The viscous liquid may be applied to glass fiber in the form of matt, tow, roving, or cloth, with precautions to eliminate air bubbles and to avoid bubbles that may be caused by overheating as a result of too rapid curing. The surface of the glass

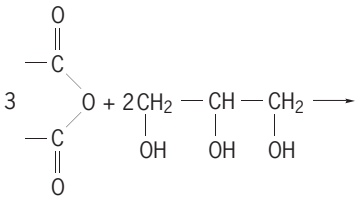

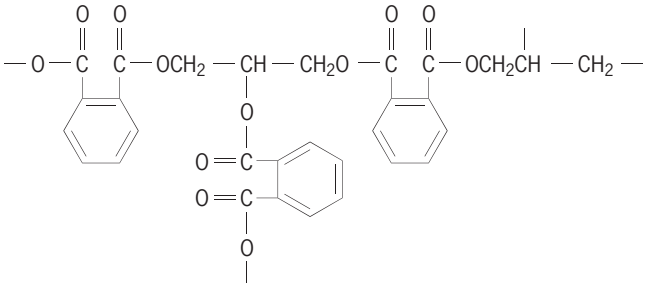

(1)

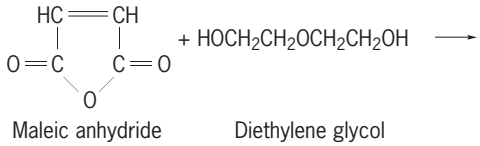

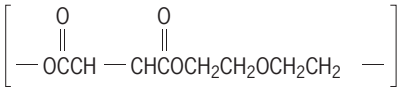

(2)

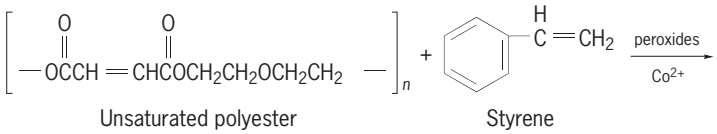

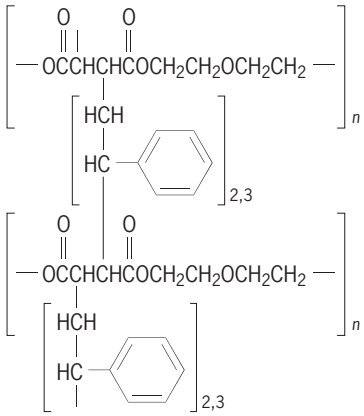

(3)

fiber must have been given a special finishing treatment in advance for the polyester to adhere strongly. Glass fibers treated with a vinyl silane or an organochrome complex are commercially available.

In the absence of the paint drier, oxygen of the air has an inhibiting effect on the curing process, with the result that the surface of the product remains soft after the inner portions have hardened. In the presence of a paint drier, such as cobalt naphthenate, this skinning effect is eliminated.

A number of modifications of the composition described above have been made. Other acids, other glycols, and various combinations may be used to vary properties of the final product, such as flexibility. The chlorinated derivatives have higher resistance to burning. By varying the free-radical initiator, the optimum temperature required for curing may be varied. There are thermosetting molding compositions which have glass fiber as a filler, and a catalyst which is relatively inactive at ordinary temperatures. When the mixture is cross-linked in the heated mold by the conventional process for thermosetting molding compounds, a fiber-reinforced, high-strength composite material results.

Saturated aliphatic polyesters. Linear polyesters were first studied by W. H. Carothers and coworkers at DuPont in the 1920s and 1930s. Made by the condensation of a diacid such as adipic acid with a diol such as diethylene glycol, the aliphatic polyesters have been frequently used as intermediates in the preparation of prepolymers for making segmented polyurethanes by reaction with diisocyanates. Lactone rings can also be opened to yield linear polyesters; for example, poly(ϵ-caprolactone) has been used as an intermediate in polyurethane technology, as a polymeric plasticizer for poly(vinyl chloride), and in other specialty applications. *See* RING-OPENING POLYMERIZATION.

A class of aliphatic polyesters that is growing in importance are those derived from bioresources. Poly(lactic acid), made by the polymerization of diolide, in turn derived from corn sugars, has become a commercial material, as in reaction (4). It

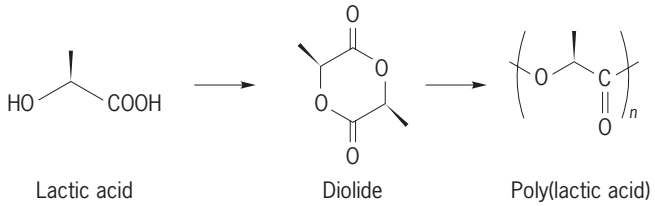

(4)

is used in applications demanding biodegradability such as packaging, compost bags, and medical applications. The poly(hydroxyalkanoates), especially poly(hydroxybutyrate), are synthesized within the cells of bacteria in very high yield. Genetic modification of the bacteria has been used to prepare a family of these materials, which have extremely high molecular weights.

One class of aliphatic polyesters with versatile properties is based on hydrogenated terephthalic acids or esters. Poly(cyclohexylidene cyclohexanedicarboxylate) [PCCD; structure **1**] is prepared

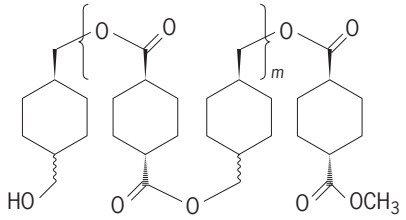

(**1**)

from 1,4-cyclohexanedicarboxylate (DMCD) and 1,4-cyclohexanedimethanol (CHDM). During reduction, a mixture of cis and trans isomers forms, and the properties of the resultant polymers depend strongly on the ratio of those isomers. For example, crystalline polymers with melting points of 230°C (450°F) or greater can be obtained from DMCD with >98% trans content, while transparent amorphous polymers are obtained when the trans content is less than 90%. CHDM can also be combined with terephthalic acid to make poly(cyclohexylidene terepthalate), a fast-crystallizing polymer, with a melting point of about 290°C (550°F). Similarly, combination of CHDM with PET provides polymers known as PCTG or PETG, depending on the ratio of CHDM to ethylene glycol units. These amorphous polymers are known for their transparency and toughness.

Aromatic polyesters. Several aromatic polyesters have achieved general importance as engineering materials, including poly(ethylene terephthalate) [PET; structure **2**a], poly(butylene terephthalate) [PBT; structure **2**b], poly(ethylene naphthoate) [PEN; structure **3**], and poly(trimethylene terephthalate) [PTT; structure **2**c]. PET is the most widely

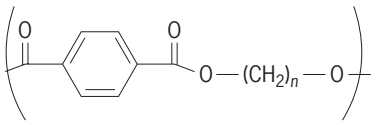

2a (n = 2), **2**b (n = 4), **2**c (n = 3)

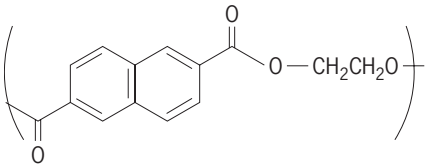

(**3**)

used material, in polyester fibers that are available in the United States and in Europe, in films, and in blow-molded bottles.

Typically, PET is prepared by a two-step process from ethylene glycol and terephthalic acid, as in reaction (5). The diacid is first reacted with excess ethylene glycol in the presence of a catalyst to form a hydroxyethyl terephthalate or an oligomer. In the

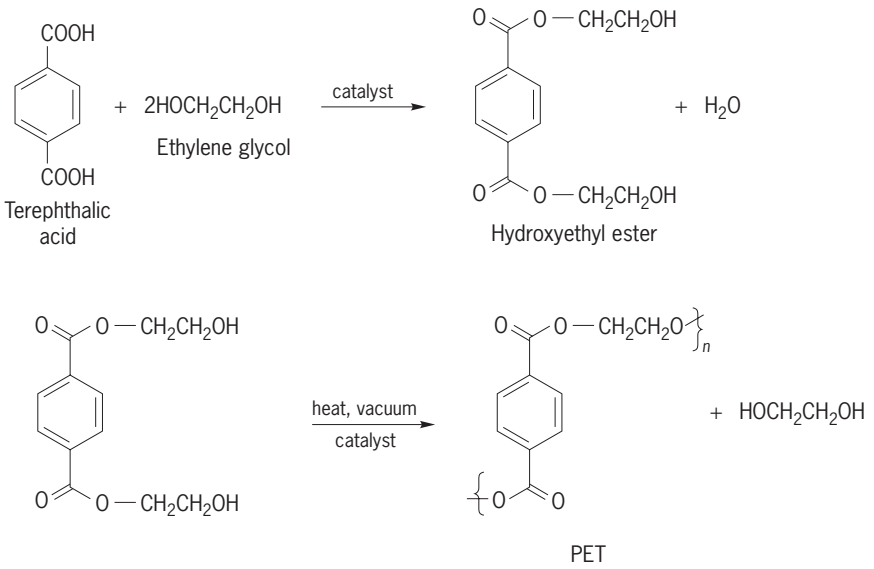

(5)

second step, polycondensation to high molecular weight is done at higher temperature under vacuum, eliminating excess ethylene glycol, and leading to an amorphous polymer. As produced in the melt, the polymer is usually amorphous, but it readily crystallizes on reheating, or on extension of the spun filaments as cast or extruded sheets. Molecular weight and crystallinity of PET can be increased by postpolmerization in the solid state as chips, at a temperature somewhat lower than the polymer's melting point of 250°C (482°F). Amorphous PET does crystallize readily, although the crystallization rate can be increased with the use of additives. Conversely, PET can be made completely amorphous with the addition of comonomers, such as diethylene glycol, which retard crystallization.

PET fiber is resistant to mildew and moths. It is used frequently in combination with cotton for clothing. Its chemical and heat resistance placed it in demand for sails and cordage. Thermoplastic molding compounds are used extensively to replace glass in beverage bottles, the recycle of which has become a major business.

PET film, which is stretched and oriented biaxially, is tough, strong, and insensitive to moisture. It is used for special packaging, as photographic film, in electrical transformers and capacitors, and in high-strength laminates. Useful films and fibers are also prepared using 1,4-cyclohexylene glycol as the glycol.

PBT is a thermoplastic, semicrystalline polyester with a melting point of about 225–230°C (437–446°F). Replacement of the ethylene glycol unit with the longer, more flexible 1,4-butanediol imparts an ability to crystallize rapidly, thus making injection-molding applications possible. PBT is often reinforced with glass fibers or other minerals, and its applications include electronic and communication equipment, automotive uses, appliances, and tools. The crystalline nature of the polymer offers good chemical and solvent resistance and good dimensional stability. PBT has low water uptake, providing consistent dimensional, impact, and electrical properties in a broad range of environments. Preparation of PBT is similar to that of PET, except that a diester, dimethyl terephthalate, is often used as the aromatic monomer, releasing methanol during the polymerization, as in reaction (6).

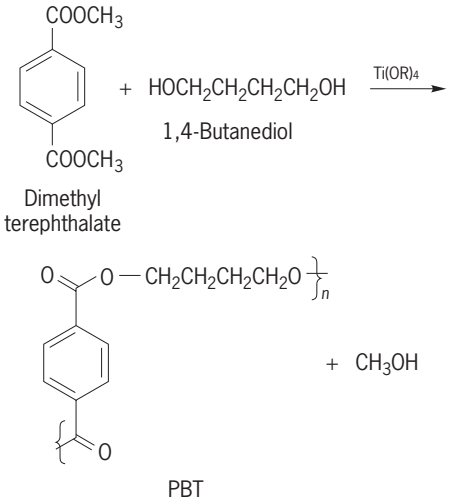

(6)

Other polyesters with commercial importance include PEN and PTT. PEN, based on 2,6-naphthalene dicarboxylates, is a high-temperature, melt-processible material with excellent barrier, strength, and thermal properties. Although it crystallizes more slowly than PET, its melting point is about 15°C (27°F) higher. PTT has been produced commercially only since 1998 due to the lack of availability of the monomer, 1,3-propanediol. With development of a new monomer process, PTT has been introduced specifically for fiber applications. The fibers have lower modulus and better bending

and recovery than PET fibers, opening applications for carpets and other materials. PTT has a melting point similar to PBT (228°C; 442°F) but crystallizes somewhat slower, although faster than PET does. *See* MANUFACTURED FIBER.

All aromatic polyesters include those commonly referred to as polyarylates, made from bisphenol A and mixtures of iso- and terephthalic acids (via their diacid chlorides). Although they may be crystalline, the commercial materials are amorphous, transparent polymers with glass transition temperatures in excess of 175°C (350°F). A special class of ultraviolet-resistant, weatherable polyarylates is based on resorcinol instead of bisphenol A as a monomer. Upon exposure to UV light, the polymer undergoes a backbone rearrangement, leading to a highly absorbent UV-resistant layer, which protects the underlying polymer.

An attractive class of aromatic polyesters are the liquid crystalline polyesters (LCPs), largely based on *p*-hydroxybenzoic acids in combination with other monomers such as terephthalic acid, hydroquinone, biphenol, or biphenol dicarboxylic acid. The LCP polyesters have excellent flow properties in the melt, but orient when spun to provide fibers with very high modulus (stiffness).

Aromatic polycarbonates. Polycarbonates (PC) are a strong, optically clear, tough group of thermoplastic polymers formed most frequently from bisphenol A. The products are noted for high glass-transition temperatures, usually greater than 145°C (284°F), high impact resistance, clarity, and resistance to creep.

Polycarbonates are prepared commercially by two methods: interfacial reaction of bisphenol A with phosgene, or transesterification with diphenyl carbonate in the melt as in reaction (7). The phosgene

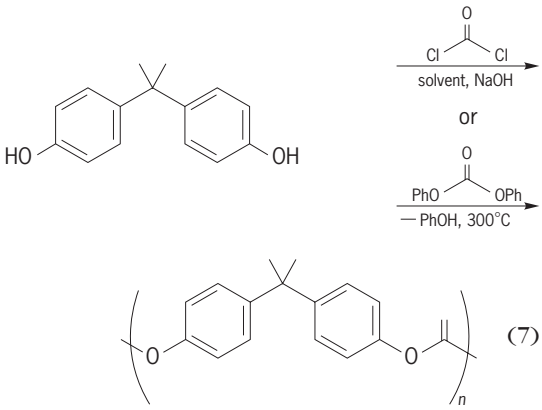

(7)

process is milder but requires removal of by-product NaCl, whereas the melt process produces resin pellets directly. Bisphenol A is obtained by the condensation of phenol and acetone.

The polymer is usually available as an injection-molding resin. Because of its high strength, toughness, and softening point, the resin, both by itself and as a glass-reinforced material, has found many electrical and engineering applications. It is often used to replace glass and metals. Examples include bottles, unbreakable windows, appliance parts, electrical housings, marine propellers, and shotgun shells. Flame-retardant grades are of interest because of low toxicity and smoke emission on burning. Films have excellent clarity and electrical characteristics, and have been used on capacitors and in solar collectors, as well as in electronic packaging. With excellent optical properties, polycarbonate has been the material of choice for optical media such as compact discs and DVDs.

Polydiallyl esters. These are polymers of diallyl esters, such as diallyl phthalate, diallyl carbonate, diallyl phenyl phosphonate, and diallyl succinate, in which cross-linked products are made by polymerization of the allyl groups, as in the case of diallyl phthalate, reaction (8).

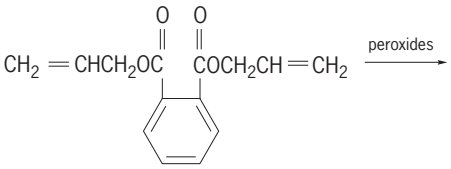

Diallyl phthalate

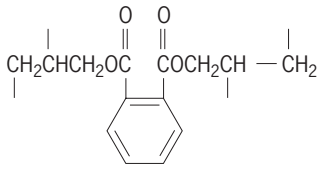

(8)

Cross-linked polymer

Thermosetting molding compounds may be produced by careful limitation of the initial polymerization to yield a product which is fusible. Then the polymerization and curing are completed in the final molding operation. Partially polymerized resins and the monomer are useful in making chemically resistant coatings. The monomer is often used as a cross-linking agent for polyester resins.

Major applications are in electronic components, sealants, coatings, and glass-fiber composites. *See* PLASTICS PROCESSING; POLYMERIZATION.

Daniel J. Brunelle

Bibliography. A.-C. Albertsson, *Degradable Alipathic Polyesters*, 2001; Y. Doi and A. Steinbüchel (eds.), *Polyesters III: Applications and Commercial Products*, vol. 4: *Biopolymers*, 2002; G. Odian, *Principles of Polymerization*, 4th ed., 2004; J. Scheirs and T. E. Long (eds.), *Modern Polyesters*, Wiley, West Sussex, England, 2003.

Polyether resins

Thermoplastic or thermosetting materials which contain ether-oxygen linkages, —C—O—C—, in the polymer chain. Depending upon the nature of the reactants and reaction conditions, a large number of polyethers with a wide range of properties may be prepared.

The main groups of polyethers in use are epoxy resins, prepared by the polymerization and

cross-linking of aromatic diepoxy compounds; phenoxy resins, high-molecular-weight epoxy resins; poly(ethylene oxide) and poly(propylene oxide) resins; polyoxymethylene, a high polymer of formaldehyde; and poly(phenylene oxides), polymers of xylenols.

Epoxy resins. The epoxy resins form an important and versatile class of cross-linked polyethers characterized by excellent chemical resistance, adhesion to glass and metals, electrical insulating properties, and ease and precision of fabrication.

In the preparation of a typical resin, a low-molecular-weight diepoxy compound is first mixed with cross-linking agents, fillers, and plasticizers and then allowed to cure either at room temperature or with the application of heat.

The intermediate diepoxy compounds are condensation products of epichlorohydrin and aliphatic or aromatic diols and are available commercially. An example is the product of the reaction of epichlorohydrin with bisphenol A [reaction (1)].

A wide variety of diols may be used in order to obtain a desired balance of properties. The liquid or solid intermediates obtained are quite stable, but do not have useful physical properties until they are polymerized further, by the addition of a curing or cross-linking agent, in a manner similar to addition polymerization. Polymerization may be initiated by catalysts such as boron trifluoride or tertiary amines. Alternatively, compounds containing a reactive hydrogen atom such as organic acids, alcohols, mercaptans, primary and secondary amines, and polyamides can also serve as catalysts. Reaction (2) represents this curing process with an amine compound. For special uses, polyols of greater functionalities are often used. These include epoxidized novalacs, epoxidized polyphenols, and glycidyl derivatives of amino-substituted aromatic compounds.

The type of curing agent employed has a marked effect on the optimum temperature of curing and has some influence on the final physical properties of the product. By judicious selection of the curing system, the curing operation can be carried out at almost any temperature from 0 to 200°C (32 to 390°F).

Various fillers such as calcium carbonate, metal fibers and powders, and glass fibers are commonly used in epoxy formulations in order to improve such properties as the strength and resistance to abrasion and high temperatures. Some reactive plasticizers, such as mercaptans and polyamides, act as curing agents, become permanently bound to the epoxy groups, and are usually called flexibilizers. Rubbery polymers such as carboxy-terminated butadienestyrene copolymers are added to improve toughness and impact strength. *See* COPOLYMER.

Because of their good adhesion to substrates and good physical properties, epoxies are commonly used in protective coatings. Because of the small density change on curing and because of their excellent electrical properties, the epoxy resins are used as potting or encapsulating compositions for the protection of delicate electronic assemblies from the thermal and mechanical shock of rocket flight. Because of their dimensional stability and toughness, the epoxies are used extensively as dies for stamping metal forms, such as automobile gasoline tanks, from metal sheeting. Foams are also made.

By combining epoxies, especially the higherperformance types based on polyfunctional monomers, with fibers such as glass or carbon, exceedingly high moduli and strengths may be obtained. Thus a typical carbon-fiber-reinforced epoxy may have a tensile modulus and strength of 75,000,000 lb/in.^{2} (500 gigapascals) and 300,000 lb/in.^{2} (2 GPa), respectively. Such composites are of importance in aerospace applications where high ratios of a property to density are desired, as well as in such domestic

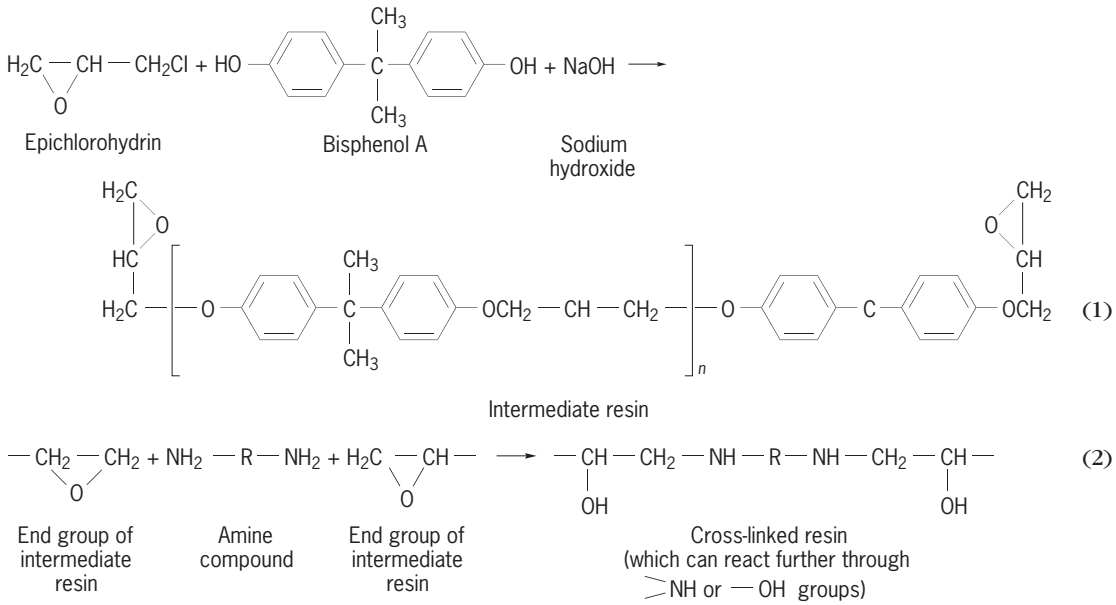

(1)

(2)

applications as sports equipment. *See* COMPOSITE MATERIAL; POLYMERIC COMPOSITE.

The adhesive properties of the resins for metals and other substrates and their relatively high resistance to heat and to chemicals have made the epoxy resins useful for protective coatings and for metal-to-metal bonding.

Poly(olefin oxide) resins. Poly(ethylene oxide) and poly(propylene oxide) are thermoplastic products whose properties are greatly influenced by molecular weight. The general method of preparation is represented by reactions (3) and (4). Other cyclic

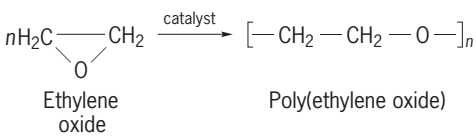

(3)

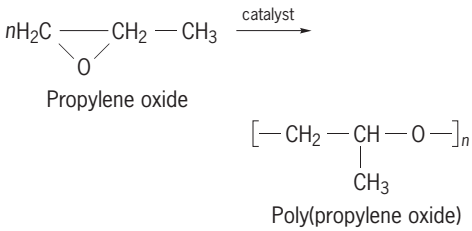

(4)

oxides, such as tetrahydrofuran, can also be polymerized to give polyethers.

Low-to-moderate molecular-weight poly(ethylene oxides) vary in form from oils to waxlike solids. They are relatively nonvolatile, are soluble in a variety of solvents, and have found many uses as thickening agents, plasticizers, lubricants for textile fibers, and components of various sizing, coating, and cosmetic preparations. The poly(propylene oxides) of similar molecular weight have somewhat similar properties, but tend to be more oil-soluble (hydrophobic) and less water-soluble (hydrophilic).

Nonionic surface-active agents can be prepared from C_{10}–C_{20} fatty alcohols and acids by the condensation of some 5–40 ethylene oxide groups, for example, reaction (5). An interesting commercial example

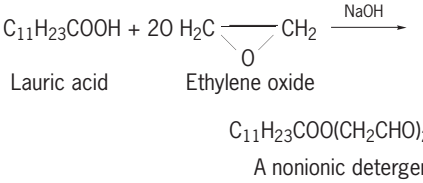

(5)

of a block copolymer has been produced by polymerization of propylene oxide onto poly(ethylene oxide) to yield a linear chain with sequences of the two compounds, as in reaction (6), in which x and

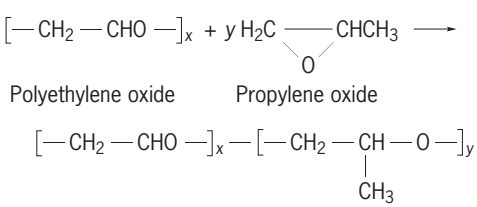

(6)

y may be in the range 10–100. The hydrophobic-hydrophilic balance necessary for surface activity is achieved because the polyethylene sequence is relatively hydrophilic and the polypropylene sequence is relatively hydrophobic.

Oil-water emulsions prepared by use of the product have remarkable stability to both hydrophilic and hydrophobic precipitating agents.

While poly(alkylene oxides) are not of interest as such in structural materials, poly(propylene oxides) are used extensively in the preparation of polyurethane forms.

Crystalline, high-molecular-weight poly(propylene oxides) with melting points up to about 74°C (165°F) have been prepared with three groups of catalysts: (1) solid potassium, (2) complexes of ferric or stannic chloride with propylene oxide, and (3) certain metallic alkyls such as aluminum triethyl. By starting with optically active propylene oxide, an optically active polymer is produced.

Phenoxy resins. Phenoxy resins differ from the structurally similar epoxy resins based on the reaction of epichlorohydrin with bisphenol A mainly by possessing a much higher molecular weight, about 25,000. The polymers are transparent, strong, ductile, and resistant to creep, and, in general, resemble polycarbonates in their behavior. Cross-linking may be effected by the use of curing agents which can react with the OH groups. Molding and extrusion may be used for fabrication. The major application is as a component in protective coatings, especially in metal primers.

Poly(phenylene oxide) resins. Poly(phenylene oxide) [PPO] is the basis for an engineering plastic characterized by chemical, thermal, and dimensional stability. The polymers are prepared by the oxidative coupling of 2,6-xylenol, using a cuprous salt and an amine as catalyst, as shown in reaction (7).

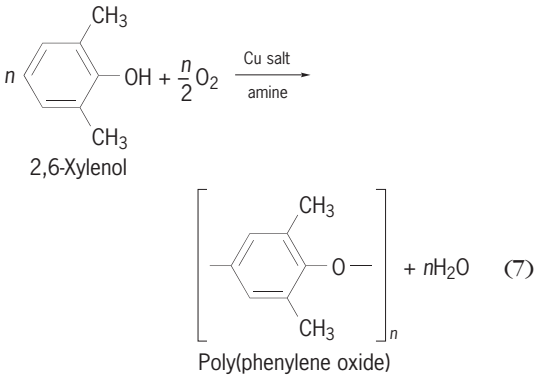

(7)

Similar reactions may be used to produce many varied structures. The poly(phenylene oxide) resin is normally blended with another compatible but cheaper resin such as high-impact polystyrene. The blends are cheaper and more processable than PPO and still retain many of the advantages of PPO by itself.

Poly(phenylene oxide is outstanding in its resistance to water, and in its maximum useful temperature range (about 170–300°C or 340–570°F). In spite of the high softening point, the resins, including glass-reinforced compositions, can be molded and extruded in conventional equipment. Uses include

medical instruments, pump parts, and insulation. Structurally modified resins are also of interest and the general family of resins may be expected to replace other materials in many applications.

An analogous polythioether, poly(phenylene sulfide), may be prepared by the reaction of *p*-dichlorobenzene with sulfur or sodium sulfide in a polar solvent yielding a branched polythioether, or by polymerization of metal salts of *p*-halothiophenols (I), yielding a linear polythioether (II) [reaction (8)].

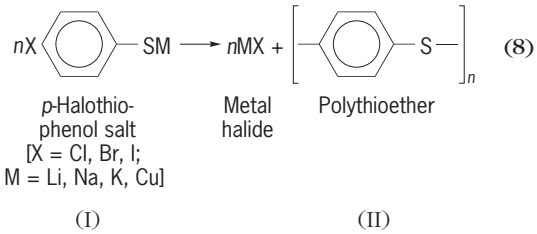

The combination of regular structure with the chemically and thermally stable ring-sulfur bond leads to a combination of a high melting point (~290°C or 550°F), thermal stability, inherent resistance to burning, and resistance to most chemicals. Commercial products can be processed by injection or compression molding, as well as by slurry or electrostatic coating; reinforcement with glass fibers is also commonly practiced. The combination of properties is useful in such applications as electronic pumps, and automotive components and coatings, especially where environmental stability is required.

Polyoxymethylene. Polyoxymethylene, or polyacetal, resins are polymers of formaldehyde. Having high molecular weights and high degrees of crystallinity, they are strong and tough and are established in the general class of engineering thermoplastics. Polymerization is accomplished by the use of a catalyst such as triphenyl phosphine, shown in reaction (9). Since the untreated polymers tend to de-

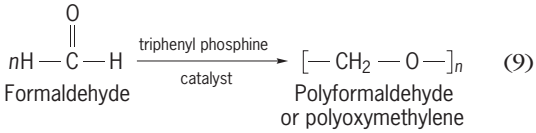

grade by an unzipping or depolymerization reaction, the end groups are frequently esterified to block the initiation of degradation. Alternatively, small amounts of a comonomer such as ethylene oxide or 1,3-dioxolane may be introduced during polymerization.

Though somewhat similar to polyethylene in general molecular structure, polyacetal molecules pack more closely, and attract each other to a much greater extent, so that the polymer is harder and higher-melting than polyethylene. Polyacetals are typically strong and tough, resistant to fatigue, creep, organic chemicals (but no strong acids or bases), and have low coefficients of friction. Electrical properties are also good. Improved properties for particular application may be attained by reinforcement with fibers of glass or polytetrafluoroethylene.

The combination of properties has led to many uses such as plumbing fittings, pump and valve components, bearings and gears, computer hardware, automobile body parts, and appliance housings. Other aldehydes may be polymerized in a similar way.

A chlorinated polyether based on the cationic polymerization of 3.3-bis(chloromethyl-1-oxacyclobutane) was developed as an engineering plastic. While it has excellent chemical resistance and chemical stability, other plastics can give as good performance at lower cost. *See* EPOXIDE; PLASTICS PROCESSING; POLYMERIZATION. John A. Manson

Bibliography. J. A. Brydson, *Plastics Materials*, 7th ed., 1999; G. W. Ehrenstein and R. P. Theriault, *Polymeric Materials: Structure, Properties, Applications*, 2000; G. Odian, *Principles of Polymerization*, 4th ed., 2004; A. B. Strong, *Plastics: Materials and Processing*, 3d ed., 2005.

Poly(ethylene glycol)

Any of a series of water-soluble polymers with the general formula $HO(CH_2CH_2O)_nH$. These colorless, odorless compounds range in appearance from viscous liquids for $n = 2$ to 13 (freezing point range of −15 to 0°C or 5 to 32°F); waxy solids and pastes for $n = 22$ to 90 (melting point range of 37 to 58°C or 98 to 136°F); and free flowing powder for $n = 113$ and above (melting point range of 60 to 67°C or 140 to 153°F).

Uses. The low-molecular-weight members, di(ethylene glycol) ($n = 2$) through tetra(ethylene glycol) ($n = 4$), are produced as pure compounds and find use as humectants, dehydrating solvents for natural gas, textile lubricants, heat-transfer fluids, solvents for aromatic hydrocarbon extractions, and intermediates for polyester resins and plasticizers. These lower poly(ethylene glycol)s are distilled by-products of ethylene glycol manufacture.

The intermediate members of the series with average molecular weights of 200 to 20,000 are produced as residue products by the sodium or potassium hydroxide-catalyzed batch polymerization of ethylene oxide onto water or mono- or di(ethylene glycol). These polymers, also known as poly(ethylene oxide) or polyoxyethylene, are formed by stepwise anionic addition polymerization and, therefore, possess a distribution of molecular weights. Examples of commercial uses for products in this range are in ceramic, metal-forming, and rubber-processing operations; as drug suppository bases and in cosmetic creams, lotions, and deodorants; as lubricants; as dispersants for casein, gelatins, and inks; and as antistatic agents. These poly(ethylene glycol)s generally have low human toxicity.

The highest members of the series have molecular weights from 100,000 to 10,000,000. They are produced by special anionic polymerization catalysts that incorporate metals such as aluminum, calcium, zinc, and iron, and coordinated ligands such as amides, nitriles, or ethers. These members of the poly(ethylene glycol) series are of interest because of their ability at very low concentrations to reduce

friction of flowing water. *See* ETHYLENE OXIDE; POLYMERIZATION; RING-OPENING POLYMERIZATION.

Robert K. Barnes; Tacey X. Viegas

Biotechnical applications. The properties of poly(ethylene glycol) make it suitable for medical and biotechnical applications. These polymers of ethylene oxide are frequently described as amphiphilic, meaning that they are soluble both in water and in most organic solvents. The medical and biotechnical applications derive from this amphiphilicity, from a lack of toxicity and immunogenicity, and from a tendency to avoid other polymers and particles also present in aqueous solution.

Poly(ethylene glycol)s are widely used in a variety of pharmaceutical formulations such as dermal ointments and creams, oral solvents in liquid-filled capsules, and in injectable solution concentrates where the low–molecular-weight polymers ($n < 9$) have been used in concentrations up to 30% by volume.

Attaching poly(ethylene glycol) to another molecule provides the latter with enhanced solubility in solvents in which poly(ethylene glycol) is soluble, and thus poly(ethylene glycol) is attached to drugs to enhance water and blood solubility. Similarly, poly(ethylene glycol) is attached to insoluble enzymes to impart solubility in organic solvents. These poly(ethylene glycol) enzymes are used as catalysts for industrial reactions in organic solvents. *See* CATALYSIS.

The tendency to avoid other polymers results in formation of two immiscible aqueous layers when a solution of a poly(ethylene glycol) is mixed with certain other polymer solutions. These aqueous two-phase systems are used to purify biological materials such as proteins, nucleic acids, and cells by partitioning of desired and undesired materials between the two phases.

The tendency of poly(ethylene glycol)s to avoid interaction with cellular and molecular components of the immune system results in the material being nonimmunogenic. Covalent attachment of a poly(ethylene glycol) to proteins results in greatly reduced immunogenicity of the poly(ethylene glycol) proteins. This property, along with reduced rate of kidney clearance because of larger size, leads to a greatly enhanced blood circulation lifetime of poly(ethylene glycol) proteins and to their application as pharmaceutical delivery systems. Similarly, adsorption of proteins and cells to surfaces is greatly reduced by attaching poly(ethylene glycol) to the surface, and such coated materials find wide application as biomaterials. *See* BIOMEDICAL CHEMICAL ENGINEERING.

J. Milton Harris; Tacey X. Viegas

Bibliography. J. Brydson, *Plastics Materials*, 7th ed., 1999; C. Delgado, G. E. Francis, and D. Fisher, The uses and properties of PEG-linked proteins, *Crit. Rev. Therap. Drug Carrier Sys.*, 9:249–304, 1992; J. Fried, *Polymer Science and Technology*, 2d ed., 2003; J. M. Harris (ed.), *Polyethylene Glycol Chemistry: Biotechnical and Biomedical Applications*, 1992; J. M. Harris and F. M. Veronese (eds.), *Adv. Drug Deliv. Rev.*, 55(10):1259–1350, September 26, 2003; F. M. Veronese and J. M. Harris (eds.), *Adv. Drug Deliv. Rev.*, 54(4):453–610, June 17, 2002.

Polyfluoroolefin resins

Resins distinguished by their resistance to heat and chemicals and by the ability to crystallize to a high degree. Several main products are based on tetrafluoroethylene, hexafluoropropylene, and monochlorotrifluoroethylene. Structural formulas of the monomers are shown below.

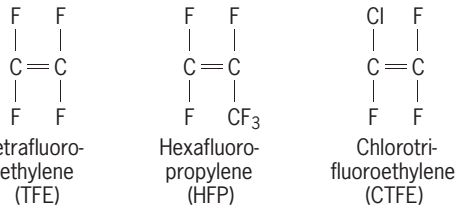

Copolymers of TFE and HFP with each other and of TFE and HEP with ethylene are available commercially. Homopolymers of vinylidene fluoride and copolymers with HFP are also in common use. For a description of poly(vinyl fluoride). *See* POLYVINYL RESINS.

Polytetrafluoroethylene. Tetrafluoroethylene can be obtained by the prolysis of monochlorodifluoromethane, which in turn is obtained from a rather complex reaction between anhydrous hydrogen fluoride and chloroform. This reaction is represented by (1), (2), and (3). Although poly-

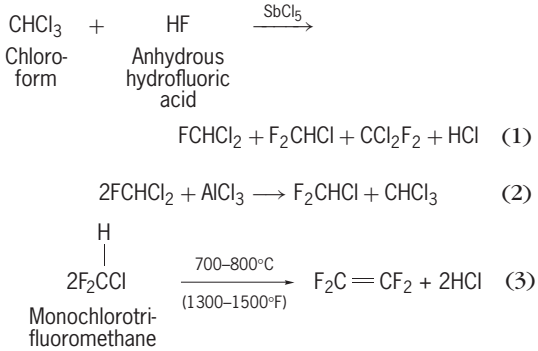

merization in bulk can proceed with violence, the monomer can be polymerized readily and conveniently in emulsion under pressure, using free-radical catalysts such as peroxides or persulfates. The polymer is insoluble, resistant to heat (up to 275°C or 527°F) and chemical attack, and in addition, has the lowest coefficient of friction of any solid. Because of its resistance to heat, the fabrication of polytetrafluoroethylene requires modification of conventional methods. After molding the powdered polymer using a cold press, the moldings are sintered at 360–400°C (680–750°F) by procedures similar to those used in powder metallurgy. The sintered product can be machined or punched. Extrusion is possible if the powder is compounded with a lubricating material. Aqueous suspensions of the polymer can also be used for

coating various articles. However, special surface treatments are required to ensure adhesion because polytetrafluoroethylene does not adhere well to anything.

Polytetrafluoroethylene (TFE resin) is useful for applications under extreme conditions of heat and chemical activity. Polytetrafluoroethylene bearings, valve seats, packings, gaskets, coatings and tubing can withstand relatively severe conditions. Fillers such as carbon, inorganic fibers, and metal powders may be incorporated to modify the mechanical and thermal properties.

Because of its excellent electrical properties, polytetrafluoroethylene is useful when a dielectric material is required for service at a high temperature. The nonadhesive quality is often turned to advantage in the use of polytetrafluoroethylene to coat articles such as rolls and cookware to which materials might otherwise adhere. A structural derivative containing perfluoroalkoxy side groups (PFA resin) is available. PFA resins behave generally like TFE resins, but can be processed by conventional molding, extrusion, and powder-coating methods.

Poly(chlorotrifluoroethylene). The monomer is prepared by the dechlorination of 1,1,2-trichloro-1,2,2-trifluoroethane as in reaction (4).

$$CCl_2FCClF_2 + \text{Zn dust} \xrightarrow{C_2H_5OH} ClFC{=}CF_2 + ZnCl_2 \quad (4)$$

Polymerization can be carried out in aqueous suspension by the free-radical process in which a combination of a persulfate and bisulfite is used as the initiator.

The properties of poly(chlorotrifluoroethylene) [CTFE resin] are generally similar to those of polytetrafluoroethylene; however, the presence of the chlorine atoms in the former causes the polymer to be a little less resistant to heat and to chemicals. The poly(chlorotrifluoroethylene) can be shaped by use of conventional molding and extrusion equipment, and it is obtained in a transparent, noncrystalline condition by quenching. Dispersions of the polymer in organic media may be used for coating.

The applications of poly(chlorotrifluoroethylene) are in general similar to those for polytetrafluoroethylene. Because of its stability and inertness, the polymer is useful in the manufacture of gaskets, linings, and valve seats that must withstand hot and corrosive conditions. It is also used as a dielectric material, as a vapor and liquid barrier, and for microporous filters.

Copolymers. Copolymers of TFE and HFP propylene (fluorinated ethylenepropylene, or FEP resins) and copolymers of TFE with ethylene (ETFE) are often used in cases where ease of fabrication is desirable. The copolymers can be processed by conventional thermoplastic techniques, and except for some diminution in the level of some properties, properties generally resemble those of the TFE homopolymer.

Copolymers of ethylene with CTFE can also be processed by conventional methods, and have better mechanical properties than PTFE, FEP, and PFA resins.

Poly(vinylidene fluoride). The monomer is prepared by the dehydrohalogenation of 1,1,1-chlorodifluoroethane or by the dechlorination of 1,2-dichloro-1,1-difluoroethane as in reactions (5) and (6). Polymerization is effected using free radical catalysis.

$$CH_3CClF_2 \xrightarrow{-HCl} H_2C{=}CF_2 \quad (5)$$

$$CH_2ClCClF_2 \xrightarrow{-Cl_2} H_2C{=}CF_2 \quad (6)$$

The properties are generally similar to those of the other fluorinated resins: relative inertness, low dielectric constant, and thermal stability (up to about 150°C or 300°F). The resins (PVF_2 resins) are, however, stronger and less susceptible to creep and abrasion than TFE and CTFE resins.

Applications of poly(vinylidene fluoride) are mainly as electrical insulation, piping, process equipment, and as a protective coating in the form of a liquid dispersion.

Fluorinated elastomers. Several types of fluorinated, noncrystallizing elastomers were developed in order to meet needs (usually military) for rubbers which possess good low-temperature behavior with a high degree of resistance to oils and to heat, radiation, and weathering. Fluorinated acrylic rubbers such as poly(heptafluorobutyl acrylate) and fluorine-containing silicone rubbers have been used to some extent for this purpose. Copolymers of hexafluoropropylene with vinylidene fluoride make up an important class with such applications as gaskets and seals. Copolymers of nitrosomethane with tetrafluoroethylene are showing considerable promise for similar applications. *See* HALOGENATED HYDROCARBON; PLASTICS PROCESSING; POLYMERIZATION.

John A. Manson

Bibliography. J. A. Brydson, *Plastics Materials*, 7th ed., 1999; G. G. Hougham et al. (eds.), *Fluoropolymers 1: Synthesis* and *Fluoropolymers 2: Properties*, 1999; G. Odian, *Principles of Polymerization*, 4th ed, 2004.

Polygalales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Rosidae of the class Magnoliopsida (dicotyledons). The order consists of 7 families and nearly 2300 species, mostly of tropical and subtropical regions. The vast majority of the species belong to only 3 families, the Malpighiaceae (about 1200 species), Polygalaceae (about 750 species), and Vochysiaceae (about 200 species). Within its subclass the order is distinguished by its simple leaves and usually irregular flowers, which have the perianth and stamens attached directly to the receptacle (hypogynous), and often have the anthers opening by terminal pores instead of longitudinal slits. The Barbados cherry (*Malpighia glabra*), noted for the high vitamin-C content of its fruits, is a well-known member of the

Polygalales. *See* MAGNOLIOPSIDA; PLANT KINGDOM; ROSIDAE. Arthur Cronquist; T. M. Barkley

Polygon

A geometric figure consisting of an ordered set of three or more (but a finite number of) points called vertices, each vertex connected by line segments called sides or edges to two other vertices. These two sides are said to be adjacent, and so are any two vertices that are end points of a side. The perimeter of the polygon is the sum of the lengths of the sides. The line segments that join two nonadjacent vertices of a polygon are called diagonals. A polygon is said to be directed, or oriented, if a preferred direction is assigned to each side so that at each vertex one of the adjacent sides is directed toward the vertex and the other away from it.

The angle between the two sides of a polygon at a vertex is called an angle of the polygon. Thus an n-sided polygon, called an n-gon, has n vertices and n angles. In particular, if n is 3, 4, 5, 6, 7, 8, 10, or 12, the polygon is called a triangle (3), quadrangle (4) [or quadrilateral, meaning four sides], pentagon (5), hexagon (6), heptagon (7), octagon (8), decagon (10), or dodecagon (12).

Plane and skew polygons. A plane polygon is one whose vertices all lie in the same plane. Other polygons are called skew polygons, except for the spherical polygons described below. Skew polygons can be constructed in any number of dimensions.

A plane polygon is called ordinary if no point belongs to more than two edges; it is proper if no two adjacent sides are collinear; it is simple if no two edges intersect each other except at vertices. A simple polygon divides the plane into two regions: an unbounded outside region and an inside region whose area is called the area of the polygon.

A simple polygon is called convex if its interior lies entirely on one side of each (infinite) line through any two adjacent vertices. A nonconvex polygon has at least one interior angle that exceeds 180°. The sum of the interior angles in a convex n-gon is $(n - 2)180°$. A plane polygon is called regular if all its sides are equal and all its angles are equal; it is semiregular if all its angles are equal. *See* REGULAR POLYTOPES.

Triangles. A triangle is equilateral if it has three equal sides, isosceles if two sides are equal, and scalene otherwise. It is a right triangle if one angle measures 90°, obtuse if one angle exceeds 90°, or acute otherwise. Its altitudes are the line segments (or their lengths) drawn from each vertex perpendicular to the opposite side (extended if necessary). These meet in a point H called the orthocenter (**Fig. 1**).

The circle passing through the three vertices A, B, and C of the triangle ABC is the circumscribed circle of radius R, having its center at the circumcenter P, which is the point of intersection of the perpendicular bisectors of the sides. Segments that join the midpoint of a side of the triangle to the opposite vertex are the triangle medians; they intersect at the center of gravity G that trisects both the medians and the segment PH, so that $HG = 2GP$ (Fig. 1).

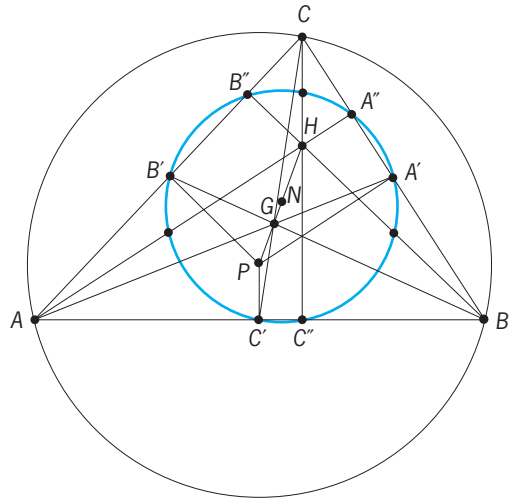

Fig. 1. Triangle *ABC* with nine-point circle. Symbols are explained in text.

The incenter of the triangle is the point of intersection of the interior angle bisectors; it is the center of the inscribed circle of radius r. Three escribed circles lying outside the triangle are each tangent to one side and to the other two sides extended; their radii are $rs/(s - a)$, $rs/(s - b)$, and $rs/(s - c)$, where $s = (a + b + c)/2$ is the semiperimeter of the triangle.

The area of the triangle ABC with sides of lengths a, b, and c is given by Eq. (1).

$$\text{Area} = rs = \sqrt{s(s-a)(s-b)(s-c)} = \frac{abc}{4R} \qquad (1)$$

Nine-point circle. An important theorem states that the so-called nine-point circle of radius $R/2$ with center N at the midpoint of HP (Fig. 1) passes through the midpoints of the sides (A', B', and C') and of the segments HA, HB, HC, as well as through the feet of the altitudes (A', B', and C') of the triangle ABC. Feuerbach's theorem states that this circle is also tangent to the inscribed and escribed circles. *See* ANGLE MODULATION; CIRCLE.

Right triangles. In a right triangle ABC, the side c opposite the right angle C is called the hypotenuse. The pythagorean theorem states that its square is equal to the sum of the squares of the other two sides: $a^2 + b^2 = c^2$. If the sides a, b, and c are integers with highest common factor 1 or 2, they form a pythagorean triple (a,b,c), expressible in the form of Eqs. (2),

$$a = m^2 - n^2 \qquad b = 2mn \qquad c = m^2 + n^2 \qquad (2)$$
$$m > n$$

where m and n are integers without a common factor. If $n < m < 2n$, the smallest angle A can be approximated in degrees to within 1 minute by Eq. (3). For example, for the pairs $(m,n) = (2,1)$,

$$A = \frac{86(m-n)(m^2-n^2)}{m^3-n^3} \qquad (3)$$

(8,5), and (10,7), this formula approximates by

36°51.4′, 26°0′, and 20°1.64′ the angles $A = 36°52.2'$, 25°59.4′, and 20°0.95′.

Quadrangles. Quadrangles or quadrilaterals are four-gons. They include (1) the trapezoid that has two parallel sides, whose area is equal to half the sum of these parallel sides times the distance between them; (2) the parallelogram, whose opposite sides are parallel and whose diagonals bisect each other; (3) the rhombus, whose four sides are equal; (4) the rectangle, whose four angles are right angles; (5) the inscribed quadrilateral, whose vertices lie on a circle so that opposite angles are supplementary; and (6) the square, which has equal sides and equal angles, and is included in all the types listed above.

If $s = (a + b + c + d)/2$ denotes the semiperimeter of an inscribed quadrilateral with sides a, b, c, and d, its area is given by Eq. (4).

$$\text{Area} = \sqrt{(s-a)(s-b)(s-c)(s-d)} \qquad (4)$$

Pentagons. The regular pentagon has five diagonals that form a polygon known as the star pentagon, or pentagram, a nonconvex polygon familiar in Christmas star decorations. The ratio of lengths of the diagonal and side of a regular pentagon, known as the golden ratio, is equal to $(1 + \sqrt{5})/2$, which is close to 144/89. If A, B, and C denote three consecutive vertices of a regular pentagon, and M is the midpoint of the diagonal AC, then the triangles ABM and CBM each have angles 36°, 54°, and 90° (**Fig. 2**). A triangle with sides 274, 377, and 466 has angles 36°0.84′, 53°59.98′, and 89°59.18′, close to 36°, 54°, and 90°, and is thus a building block for constructing a nearly regular pentagon without a protractor.

Hexagons. The regular hexagon is commonly found in floor tilings and is noticeable in the cross section of a beehive. When circular objects are closely packed, they tend to assume a regular hexagonal pattern. Crystalline solids having such a form are known as hexagonal close-packed (hcp) structures. *See* CRYSTAL STRUCTURE.

Spherical polygons. A spherical polygon consists of points on a spherical surface called vertices, each connected to two adjacent vertices by great-circle arcs, called sides, that are measured by their central angles. The n angles of a convex spherical n-gon are the angles at each vertex between the tangent lines to the two sides that meet there, and their sum exceeds $(n - 2)180°$ by an amount E, called the spherical excess, that is proportional to the spherical area enclosed by the polygon. If the spherical radius is R, this area is given by Eq. (5).

$$\text{Area} = \frac{E}{180°}\pi R^2 \qquad (5)$$

Spherical triangles and polar triangles. Any three distinct points, A, B, and C, on a spherical surface are the vertices of a spherical triangle, whose sides are the (shorter) great-circle arcs $a = BC$, $b = CA$, and $c = AB$, joining the vertices. The relations between the sides and angles of a spherical triangle are investigated in the subject of spherical trigonometry. *See* TRIGONOMETRY.

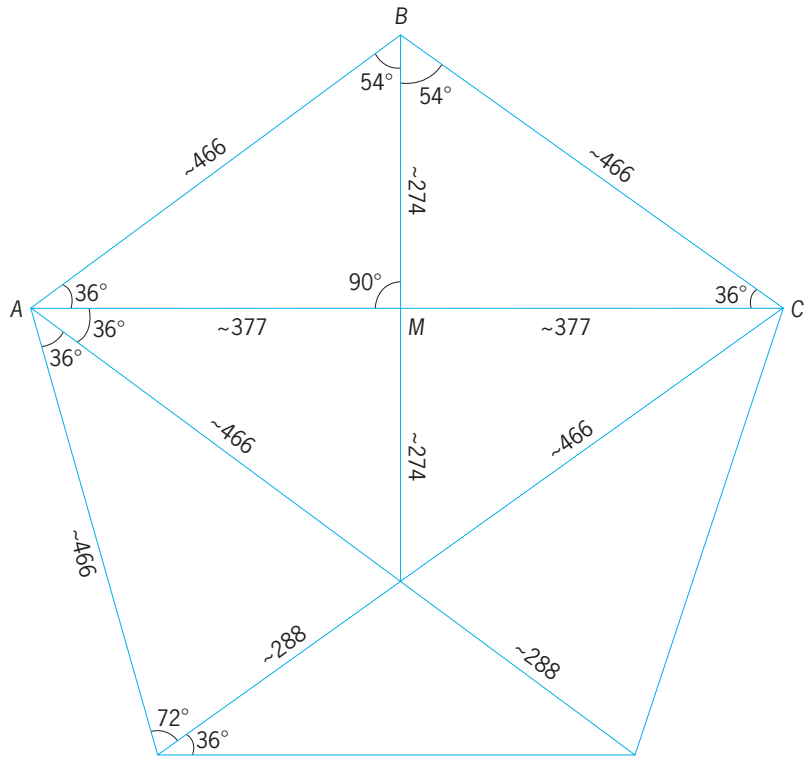

Fig. 2. Regular pentagon, showing approximate construction without a protractor. Symbols are explained in text.

The axis and poles of a great-circle arc are respectively the spherical diameter perpendicular to its plane and the points where this axis meets the spherical surface. Thus, on the Earth's surface (assumed to be spherical) the poles of the Equator are the North and South poles, lying on the Earth's axis of rotation. Given a spherical triangle ABC, the polar triangle $A'B'C'$ is the triangle whose vertices, A', B', and C', are those poles of the sides a, b, and c nearest to A, B, and C, respectively. The polar triangle of $A'B'C'$ is the original triangle ABC (**Fig. 3**). If the

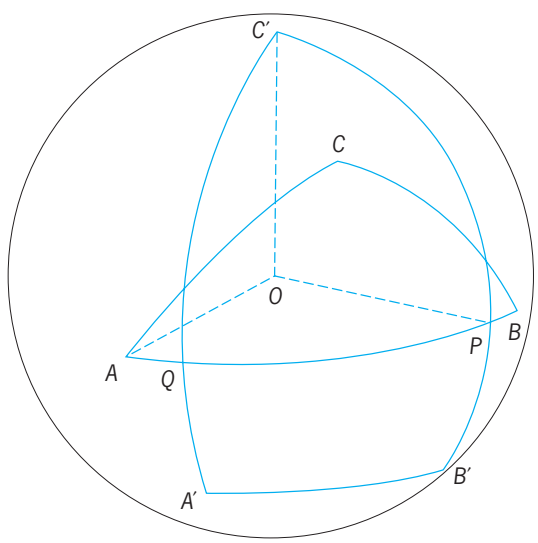

Fig. 3. Polar triangles *A*, *B*, *C* and *A′*, *B′*, *C′*. The center of the sphere is at *O*. *AQ* and *PB* are 90° arcs, so angle *C′* = arc *PQ* is the supplement of *AB*.

sides $A' = C'B'$ and $b' = C'A'$ of the polar triangle (extended if necessary) meet side c of ABC in points P and Q respectively, then arcs AP and QB measure 90°, so the angular measures of arc AB = side c and arc PQ = angle C' are supplementary. Thus, the angular measures of the sides of a spherical triangle are the supplements of the angles of the polar triangle.

Curve approximation. By taking a large number of consecutive points along a smooth closed curve as vertices of a polygon whose side lengths are sufficiently small, the length of the curve can be approximated by the perimeter of the polygon. The curve length is defined if these perimeter lengths converge when the largest side length approaches zero. If the curve forms a knot in space, the knot can be analyzed by projecting an oriented approximating polygon onto a suitable plane and indicating, at each point where two sides cross each other, both of their directions and which side is above the other in projection. J. Sutherland Frame

Bibliography. F. Bachmann and E. Schmidt, *n-Gons*, 1975; H. S. M. Coxeter, *Introduction to Geometry*, 2d ed., rev. printing, 1989; B. Grünbaum and G. C. Shephard, *Tilings and Patterns*, 1986; L. M. Kelly (ed.), *The Geometry of Metric and Linear Spaces*, 1975; K. Miyazaki, *An Adventure in Multidimensional Space: The Art and Geometry of Polygons, Polyhedra, and Polytopes*, 1986.

Polygonales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Caryophyllidae of the class Magnoliopsida (dicotyledons). The order consists only of the family Polygonaceae, with about 800 species, most abundant in north temperate regions (see **illus.**). Within its subclass the order is characterized by its well-developed endosperm; an ovary with but one chamber (unilocular), mostly tricarpellate with a single basal ovule which is usually straight and has the micropyle at the opposite end from the stalk (orthotropous); the flowers are often trimerous, that is, with floral parts in sets of three, and usually two or three sets of stamens. Rhubarb (*Rheum rhaponticum*) and buckwheat (*Fagopyrum esculentum*) are familiar members of the Polygonales. *See* BUCKWHEAT; CARYOPHYLLIDAE; MAGNOLIOPSIDA; PLANT KINGDOM; RHUBARB. Arthur Cronquist; T. M. Barkley

***Polygonum hydropiper*, eastern American smartweed, characteristic of order Polygonales. Sheathing stipules and many tiny flowers are typical of group. (*A. W. Ambler, National Audubon Society*)**

Polyhedron

A solid whose boundary consists of a finite number of polygonal faces, that is, planar regions that are bounded by polygons. The sides of the faces are edges of the polyhedron; the vertices of the faces also are vertices of the polyhedron. *See* PLANE GEOMETRY; POLYGON.

Most polyhedra met in applied geometry are convex and simply connected. A polyhedron is convex if it passes this test: if any face is placed coincident with a plane, then all other points of the polyhedron lie on the same side of that plane. A more informal test is to imagine enclosing the polyhedron within a stretched elastic membrane; the polyhedron is convex if all points on the boundary are in contact with the membrane. A simply connected polyhedron has a boundary that is topologically equivalent to a sphere: if the boundary were made of some perfectly elastic material, then the boundary could be distorted into a sphere without tearing or piercing the surface. A simply connected polyhedron is said to be eulerian, because the number of faces F, the number of edges E, and the number of vertices V satisfy Euler's formula (1). The remainder of this article will deal only with simply connected convex polyhedra.

$$F + V = E + 2 \tag{1}$$

A polyhedron has a volume, a measure of the amount of space enclosed by the boundary surface. As noted in cases below, often the volume can be determined if certain other measurements associated with the polyhedron are known.

The fewest number of faces that a polyhedron might have is four. "Poly" signifies "many," but since an n-hedron (a polyhedron having n faces) can be transformed into an $n + 1$-hedron by truncating the polyhedron near a vertex, polyhedra exist having any number of faces greater than three. Some polyhedra have names that convey the number of faces (but not the shape) of the polyhedron: tetrahedron, 4 faces; pentahedron, 5 faces; hexahedron, 6 faces; octahedron, 8 faces; dodecahedron, 12 faces; and icosahedron, 20 faces.

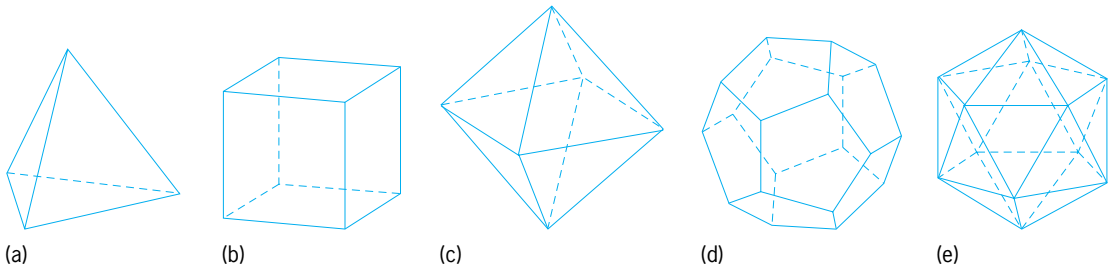

The five regular polyhedra: (*a*) Tetrahedron. (*b*) Cube. (*c*) Octahedron. (*d*) Dodecahedron. (*e*) Icosahedron.

Many polyhedra have properties that allow them to be placed in one or more of the categories below.

Prismatoid. A prismatoid is a polyhedron in which all vertices lie on exactly two parallel base planes. A base plane might contain only one or two vertices, but if it contains three or more, then that polygonal face is a base of the prismatoid. The height (or altitude) of a prismatoid is the perpendicular distance between the base planes.

Any prismatoid satisfies the necessary requirements for its volume to be given by the prismoidal volume formula: if h denotes the height of the prismatoid, L and U denote the areas of the bases (where one of these areas might equal zero), and M is the area of a plane section parallel to a base and midway between the base planes, then the volume, V, is given by Eq. (2). (The formula is not especially useful in

$$V = {}^1\!/_6 h(L + 4M + U) \tag{2}$$

practice because seldom is the area of the midsection known or easily determined.) Some solids other than prismatoids also have volumes given by this formula. *See* SOLID (GEOMETRY).

A prismoid is a prismatoid that satisfies some additional requirement. However, there is no agreement on precisely what that requirement is, so any use of the word should be accompanied by the user's choice from the many definitions appearing in dictionaries.

Prism. A prism is a prismatoid that has two congruent bases and whose other faces (the lateral faces) are all parallelograms. If the lateral faces are rectangles, then the prism is right; otherwise it is oblique. If B represents the area of either base and h represents the height, then the volume, V, of the prism is given by Eq. (3). A prism can be classified according

$$V = Bh \tag{3}$$

to the type of polygonal region forming the bases: a triangular prism, a square prism, a hexagonal prism, and so forth. *See* PRISM.

Parallelepiped. A parallelepiped is a prism in which the bases are parallelograms. If the bases and all lateral faces are rectangles, then the prism is a rectangular parallelepiped or, more informally, a rectangular box. If l and w denote the length and width of a base, and h is the height of a rectangular parallelepiped, then the volume, V, is given by Eq. (4).

$$V = lwh \tag{4}$$

Cube. A cube is a rectangular parallelepiped in which all six faces are squares. If s denotes the length of an edge of the cube, then the volume, V, is given by Eq. (5).

$$V = s^3 \tag{5}$$

See CUBE.

Pyramid. A pyramid is a prismatoid in which one base plane contains a single vertex: the vertex of the pyramid. The polygonal region defined by the vertices on the other base plane is the base of the pyramid. The base might contain any number of sides greater than two, but the other faces (the lateral faces) will all be triangular regions. A pyramid may be classified according to the shape of the base; thus a tetrahedron may also be called a triangular pyramid. A regular pyramid has a base that is a regular polygon, and its lateral faces are all congruent isosceles triangles. If B denotes the area of the base, and h is the height of the pyramid, then the volume, V, is given by Eq. (6).

$$V = {}^1\!/_3 Bh \tag{6}$$

Frustrum. A frustrum (alternative name, frustum) of a pyramid is the solid remaining after a pyramid is sliced by a plane parallel to the base and the part containing the vertex is discarded. (The discarded part is another pyramid, similar to the original.) The bases of the frustrum are the base of the original pyramid and the new face created by the slicing. If B and b denote the areas of the bases, and h is the height of the frustrum, then the volume, V, is given by Eq. (7).

$$V = {}^1\!/_3 h(B + \sqrt{Bb} + b) \tag{7}$$

Regular polyhedra. A polyhedron is regular (or platonic) if all faces are congruent and all dihedral

Regular polyhedra

Regular polyhedron	Number of elements*		
	F	V	E
Tetrahedron	4	4	6
Cube (hexahedron)	6	8	12
Octahedron	8	6	12
Dodecahedron	12	20	30
Icosahedron	20	12	30

* F, V, and E respectively represent the number of faces, vertices, and edges.

angles (the angles between adjacent faces) are equal. There are only five regular polyhedra: the tetrahedron, cube, octahedron, dodecahedron, and icosahedron (see **illus.** and **table**).

The center of each face of a cube is the vertex of a regular octahedron; the center of each face of a regular octahedron is the vertex of a cube. The cube and the regular octahedron, therefore, are said to be reciprocal polyhedra. The regular dodecahedron and the regular icosahedron also form a pair of reciprocal polyhedra. The reciprocal of a regular tetrahedron is another regular tetrahedron. *See* EUCLIDEAN GEOMETRY; REGULAR POLYTOPES. Harry L. Baldwin, Jr.

Bibliography. H. S. M. Coxeter, *Introduction to Geometry*, 2d ed., 1989.

Polymer

Polymers, macromolecules, high polymers, and giant molecules are high-molecular-weight materials composed of repeating subunits. These materials may be organic, inorganic, or organometallic, and synthetic or natural in origin. Polymers are essential materials for almost every industry as adhesives, building materials, paper, cloths, fibers, coatings, plastics, ceramics, concretes, liquid crystals, photoresists, and coatings. They are also major components in soils and plant and animal life. They are important in nutrition, engineering, biology, medicine, computers, space exploration, health, and the environment. *See* CERAMICS; MANUFACTURED FIBER; NATURAL FIBER; SURFACE COATING.

Natural inorganic polymers include diamonds, graphite, sand, asbestos, agates, chert, feldspars, mica, quartz, and talc. Natural organic polymers include polysaccharides (or polycarbohydrates) such as starch and cellulose, nucleic acids, and proteins. Synthetic inorganic polymers include boron nitride, concrete, many high-temperature superconductors, and a number of glasses. Siloxanes or polysiloxanes represent synthetic organometallic polymers. *See* SILICONE RESINS.

Synthetic polymers offer a savings in energy requirements over other materials such as metals. Their lightness reduces the consumption of fuel in vehicles and aircraft, and they outperform most metals when measured on a strength-per-weight basis. As properties have improved, polymers have been developed which can be readily and economically fabricated, and which can be used for engineering purposes such as gears, bearings, and structural members.

History. Most of the early discoveries in polymer science were empirical, with polymer chemists often converting naturally occurring materials into materials with enhanced properties. Charles Goodyear (1839) transformed natural rubber (from the rubber tree *Hevea brasiliensis*) to a less heatsensitive product using sulfur as a crosslinking agent, thus producing an elastic material with greater thermal stability. Alexander Parkes and John W. Hyatt (about 1868) added agents to natural materials such as cellulose and cellulose-derived materials to make them more flexible. Leo H. Baekeland (1907) produced crosslinked polymeric materials from formaldehyde and phenol, called Bakelite. *See* CELLULOSE; PHENOLIC RESIN; RUBBER.

In the 1920s, Hermann Staudinger demonstrated that natural and synthetic polymers are not aggregates such as colloids, or cyclic compounds such as cyclohexane, but are long chainlike molecules with specific terminal groups. His work, for which he was awarded the Nobel Prize in Chemistry (1953), is considered the basis for modern polymer science.

In 1928, Herman Mark and Kurt Meyer used x-ray techniques to determine the crystallographic structures of cellulose and natural rubber. The following year, Wallace H. Carothers synthesized and characterized linear (aliphatic) polymers. Shortly thereafter, his group at Dupont synthesized nylon, a fiberforming polyamide. Carothers also coined many of the terms used in polymer science. *See* POLYAMIDE RESINS.

Nomenclature. Since much of polymer science was developed in industry and within a wide range of disciplines, many polymers have both a common name and a structure-based name specified by the International Union of Pure and Applied Chemistry (IUPAC). Many polymers are commonly known by their acronyms. *See* ORGANIC NOMENCLATURE.

For example, polystyrene or PS is named poly(1-phenylethylene) by the IUPAC, while poly(methyl methacrylate) or PMMA is named poly[(1-methoxycarbonyl)-1-methylethylene]. *See* POLYACRYLATE RESIN; POLYSTYRENE RESIN.

Many companies use trade names to identify the specific polymeric products they manufacture. For example, Fortrel® polyester is a poly(ethylene terephthalate) or PET fiber. Many polymers are generically named, such as rayon, polyester, and nylon. *See* POLYESTER RESINS.

Composition

Generally, polymers have structures that can be represented by similar or identical repeat units. The repeat units are derived from smaller molecules, called monomers, which react to form the polymer. For example, propylene monomer (structure **1**) and the repeat unit (**2**) it forms in polypropylene are shown below.

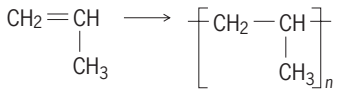

Monomer (**1**) Polymer repeat unit (**2**)

With the exception of its end groups, polypropylene is composed entirely of this repeat unit. The number of units (n) in a polymer chain is called the degree of polymerization or DP. *See* POLYVINYL RESINS.

Other polymers, such as proteins, can be described in terms of the approximate repeat unit (**3**)

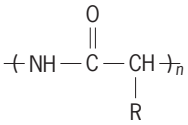

(**3**)

where the nature of R (a substituted atom or group of atoms) varies. *See* PROTEIN.

Primary structure. The sequence of repeat units within a polymer is called its primary structure. Unsymmetrical reactants, such as substituted vinyl monomers (**4**), react almost exclusively to give

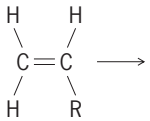

Vinyl monomer

(**4**)

a "head-to-tail" product (**5**), in which the R substituents occur on alternate carbon atoms. A variety of head-to-head structures (**6**) are also possible.

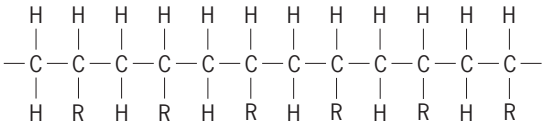

Head-to-tail

(**5**)

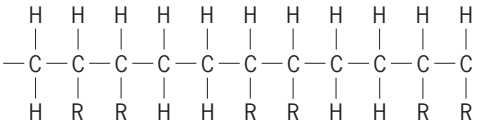

Head-to-head/tail-to-tail

(**6**)

Each R-substituted carbon atom is a chiral center (one atom in a molecule attached to four different groups) with different geometries possible. Arrangements where the substituents on the chiral carbon are random are referred to as atactic structures. Arrangements where the geometry about the chiral carbon alternates are said to be syndiotactic. Structures where the geometry about the chiral atom have the same geometry are said to be isotactic. Both syndiotactic and isotactic polymers are stereoregular. *See* POLYMER STEREOCHEMISTRY AND PROPERTIES.

Stereoregular polymers are produced using special stereoregulating catalyst systems. A series of soluble catalysts have been developed that yield products with high stereoregularity and low chain size disparity. As expected, polymers with regular structures—that is, isotactic and syndiotactic structures—tend to be more crystalline and stronger. *See* STEREOSPECIFIC CATALYST.

Polymers can be linear or branched with varying amounts and lengths of branching (**Fig. 1**). Most polymers contain some branching.

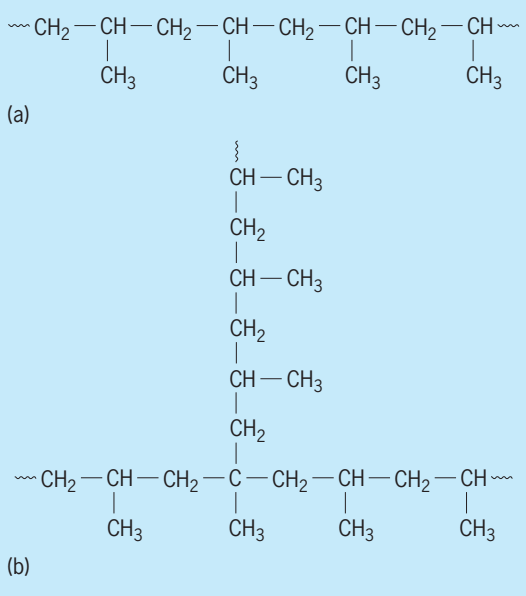

Fig. 1. Polymer structures: (*a*) linear, (*b*) branched.

Copolymers are derived from two different monomers, which may be represented as A and B. There exists a large variety of possible structures and, with each structure, specific properties. These varieties include alternating, random, block, and graft (**Fig. 2**). *See* COPOLYMER.

Secondary structure. Secondary structure refers to the localized shape of the polymer, which is often the consequence of hydrogen bonding and other hydrophilic and hydrophobic interactions. (Nonpolar regions tend to clump with other nonpolar regions, while polar regions tend to clump with other polar regions.) Most flexible to semiflexible linear polymer chains tend toward two structures—helical and pleated sheet/skirtlike. The pleated skirt arrangement is most prevalent for polar materials where hydrogen bonding can occur. In nature, protein tissue is often of a pleated skirt arrangement. For both polar and nonpolar polymer chains, there is a tendency

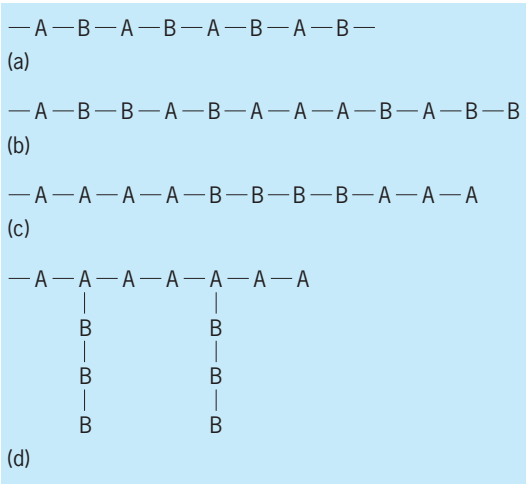

Fig. 2. Copolymer structures: (*a*) alternating, (*b*) random, (*c*) block, (*d*) graft.

toward helical formation with the inner core having "like" secondary bonding forces. *See* HYDROGEN BOND.

Tertiary structure. Tertiary structure refers to the overall shape of a polymer, such as in polypeptide folding. Globular proteins approximate rough spheres because of a complex combination of environmental and molecular constraints, and bonding opportunities. Many natural and synthetic polymers have "superstructures," such as the globular proteins and aggregates of polymer chains, forming bundles and groupings.

Quaternary structure. Quaternary structure refers to the arrangement in space of two or more polymer subunits, often a grouping together of tertiary structures. For example, hemoglobin (quaternary structure) is essentially the combination of four myoglobin (tertiary structure) units. Many crystalline synthetic polymers form spherulites. *See* HEMOGLOBIN.

Synthesis

For polymerization to occur, monomers must have at least two reaction points or functional groups. There are two main reaction routes to synthetic polymer formation—addition and condensation. In chain-type kinetics, initiation starts a series of monomer additions that result in the reaction mixture consisting mostly of unreacted monomer and polymer. Vinyl polymers, derived from vinyl monomers and containing only carbon in their backbone, are formed in this way. Examples of vinyl polymers include polystyrene, polyethylene, polybutadiene, polypropylene (**7**), and poly(vinyl chloride).

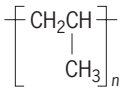

Polypropylene, PP

(**7**)

See POLYMERIZATION; POLYOLEFIN RESINS.

The second main route is a step-wise polymerization. Polymerization occurs in a step-wise fashion so that the average chain size within the reaction mixture may have an overall degree of polymerization of 2, then 5, then 10, and so on, until the entire mixture contains largely polymer with little or no monomer left. Polymers typically produced using the step-wise process are called condensation polymers, and include polyamides (**8**), polycarbonates (**9**), polyesters, and polyurethanes (**10**). Condensation polymer chains are characterized as having a noncarbon atom in their backbone. For polyamides the noncarbon is nitrogen (N), while for polycarbonates it is oxygen (O). Condensation polymers are synthesized using melt (the reactants are heated causing them to melt), solution (the reactants are dissolved), and interfacial (the reactants are dissolved in immiscible solvents) techniques. **Table 1** lists some addition and condensation polymers. *See* POLYURETHANE RESINS.

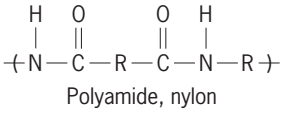

Polyamide, nylon

(**8**)

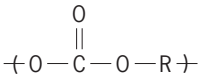

Polycarbonate

(**9**)

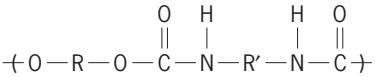

Polyurethane, PU

(**10**)

TABLE 1. Polymers and amounts synthesized annually in the United States

Polymer	Type of synthesis	Amount, 10^6 lb
Polyesters	Condensation	5,500
Ureas	Condensation	1,900
Phenolics	Condensation	3,300
Polyethylene	Addition	25,000
Polypropylene	Addition	10,000
Polystyrene	Addition	6,000
Poly(vinyl chloride)	Addition	11,000
Polyacrylonitrile	Addition	500
Polyamides, nylons	Condensation	30,000
Olefins	Addition	2,500
Polybutadiene	Addition	1,100
Styrene-butadiene	Addition	2,200

Properties

The properties of polymeric materials are determined by their molecular properties and morphology. Polymer morphology in turn depends on the polymerization conditions.

Molecular properties. These include molecular size and weight, molecular structure or architecture, molecular weight distribution, polarity, and flexibility of the polymeric chains (or chain segments between crosslinks in cured or vulcanized polymers). Molecular properties taken together determine the attractive forces between the molecules, the morphology or arrangement of masses of molecules, and the general behavior of the polymer.

Molecular weight. The molecular weight of a particular polymer chain is the product of the number of units times the molecular weight of the repeating unit. Polymerization reactions, both synthetic and natural, generally produce products with varying chain lengths. [The exception is the biochemical reactions that produce enzymes (proteins) and nucleic acids (such as DNA), which have precise molecular weights and chain orientations critical to their function.]

A polymer's varying chain lengths result in a product with a distribution of molecular weights. This molecular weight distribution may be narrow or broad, monomodal or polymodal, depending on the polymerization conditions.

Two statistical averages describe polymers, the number-average molecular weight $\bar{M}n$ and the weight-average molecular weight $\bar{M}w$. The $\bar{M}n$ is particularly dependent on the number of polymer chains with a specific chain length, while the $\bar{M}w$ is dependent on the size distribution of the polymer chains. The $\bar{M}n$ is determined using procedures that are related to the number of chains of a particular length. Colligative properties are dependent on the number of particles present, and can be used to determine $\bar{M}n$. These techniques include membrane osmometry, vapor-phase osmometry, ebulliometry (boiling-point elevation), cryoscopy (melting-point depression), end-group analysis, and osmodialysis. Light-scattering photometry is the most often used technique for measuring $\bar{M}w$.

Since polymers generally greatly increase the viscosity (resistance to flow) of solutions containing them, viscometry is often used to determine the relative molecular weight of polymers. Today, combinations of gel permeation chromatography and light-scattering photometry can be used to determine both average molecular weight and the molecular weight of narrow molecular weight groupings. *See* LIGHT-SCATTERING TECHNIQUES; MOLECULAR WEIGHT; SCATTERING OF ELECTROMAGNETIC RADIATION; VISCOSITY.

Size. Size is the most important property of polymers. Size allows storage of information (nucleic acids and proteins) and allows polymeric materials to "remember" when reversible movement occurs; in cases where movement is not reversible, polymer structures are a consequence of past treatment such as bending, stretching, melting—any action that distorts or moves polymer chains or segments). For large chains, entanglement occurs, introducing physical crosslinking and enhancing the ability of the polymers to remember and return to original molecular structures after applied stress/strain. An aspect of this remembering involves the need for segments of a polymer chain to accompany adjacent polymer segments. Thus, for space travel applications, even when the vacuum is enough to allow segmental "evaporation" of external polymer chains, some of the chain remains connected, trapped within the interior of the polymer matrix, preventing total removal. Polymeric materials are used extensively on the space shuttles.

Attractive forces. Size also accounts for an accumulation of the interchain (between molecules) and intrachain (within the same molecule) secondary attractive forces. These secondary forces, called van der Waals forces, are composed of London or dispersion forces, dipole-dipole forces, and hydrogen bonding. For a nonpolar polymer, such as polyethylene (**11**), the attractive forces for each

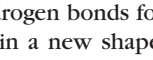

Polyethylene, PE

(**11**)

unit are about 2 kcal/mole per repeating unit. For a polyethylene chain 1000 units long, the cumulative dispersion attractive forces would be about 2000 kcal/mole, which is well above the carbon-carbon bond strengths so that even low-molecular-weight polymer chains are nonvolatile. As the temperature is increased, polymers will decompose prior to volatilization. *See* CHEMICAL BONDING; INTERMOLECULAR FORCES.

For polar polymers, such as poly(vinyl chloride) (**12**), attractive forces include both dispersion

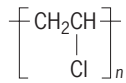

Poly(vinyl chloride), PVC

(**12**)

and dipole-dipole forces so that the total attractive forces are proportionally larger than those for polyethylene. Examples of polar polymers include poly(vinyl acetate) (**13**), poly(ethylene glycol) (**14**), and poly(methyl methacrylate) (**15**). For polymers that can take advantage of hydrogen bonding, the attractive forces are even greater. Such polymers include poly(vinyl alcohol), proteins, polysaccharides, nucleic acids, and nylons (**16**).

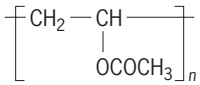

Poly(vinyl acetate)

(**13**)

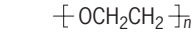

Poly(ethylene glycol), PEG

(**14**)

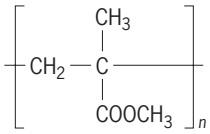

Poly(methyl methacrylate), PMMA

(**15**)

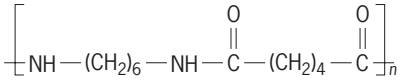

Nylon 6,6

(**16**)

Hydrogen bonding is so strong in cellulose that cellulose is not soluble in water until the inter- and intrachain hydrogen bonds are broken. Thus, "wood shaping" is often accomplished by breaking the hydrogen bonds by addition of sodium hydroxide. After the wood is shaped, the sodium hydroxide is washed away and the hydrogen bonds form again, allowing the wood to retain a new shape. This principle is similar to that of permanent press clothing, where crosslinks are introduced to retain fabric shape.

Morphology. As bonding forces and symmetry increase, the tendency to form regular structures increases, and since the inter-and intrachain forces are

inversely proportional to the square of the distance, the chains are able to reside closer together. Polymers often have a combination of ordered regions, called crystalline regions, and disordered or amorphous regions. The amorphous regions contribute to polymers' flexibility and diffusivity of gases and other materials. At low temperatures these regions are rigid, but as the temperature increases there is a temperature range where local or segmental chain mobility occurs. This temperature range is called the glass-phase transition temperature, T_g. Polymers are flexible only above their glass-phase transition temperature.

Regions where polymer chains are aligned in an orderly manner are called crystalline regions. Crystalline regions contribute to strength and resistance to external forces, including chemical attack. Crystalline regions can be melted. At the melting point or range, whole-chain movement can occur.

While crystalline regions add strength to a material, such regions are both rigid and brittle. The amorphous regions, where the chains are in a more random orientation, increase flexibility but decrease the strength of a material. Most commercial polymers have a balance between amorphous and crystalline regions, allowing a balance between flexibility and strength.

Crystallization is favored by the presence of regularly spaced symmetrical units. This is reinforced by the presence of regularly spaced polar groups that can form secondary dipole-dipole interactions. Polymers such as polyamides (nylons), polycarbonates (**17**), and polyesters tend to form crystalline

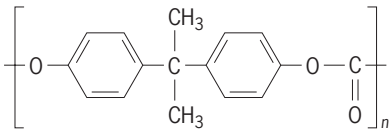

Polycarbonate, PC

(**17**)

polymers because of the presence of additional secondary forces. Polymers such as linear polyethylene (with little or no branching) tend to be crystalline because of their symmetry and the ability of their chains to closely approach one another.

Viscoelastic materials. Polymers are viscoelastic materials, meaning they can behave as liquids and solids. When they behave as liquids, their actions can be described by Newton's law, where the applied stress is proportional to the rate of strain. For materials behaving as solids, the mechanical properties can often be described in terms of a spring using Hooke's law, where the applied stress is proportional to the resultant strain. Stress is force per unit area, and strain or elongation is the extension per unit length. *See* STRESS AND STRAIN.

A reinforced or crystallizable elastomer exhibits a relatively low breaking stress but a high elongation. Ductile polymers, such as polyethylene and polypropylene, "give" or "yield," and at high elongations some strengthening and orientation occur.

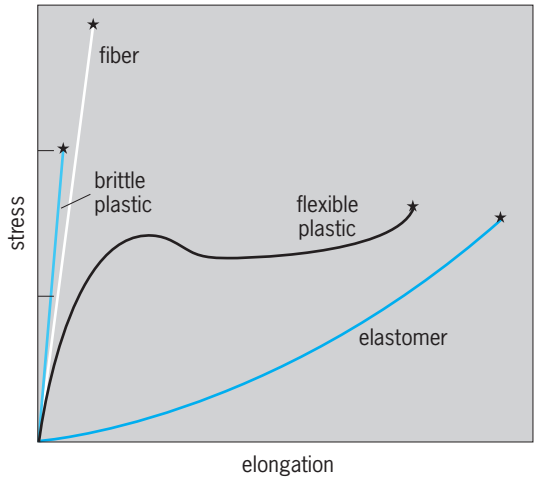

Fig. 3. Stress-elongation behavior of typical classes of polymer (not drawn to scale). The star represents the point at which failure occurs.

A brittle polymer, such as polystyrene, does not give much and breaks at a low elongation. A fiber exhibits high strength, high stiffness, and little elongation. In each case, the particular property is dependent on the temperature and the rate in which the stress is applied. Thus, at low temperatures where segmental and whole-chain mobility does not occur, the material can act as a solid. In addition, when a stress is rapidly applied so that chain or segmental mobility is not able to engage, the material will also act as a solid. Thus, a polymer that is brittle or "hookean" at low temperature may become ductile with newtonian behavior at higher temperatures where segmental or chain mobility can occur (**Fig. 3**). *See* ELASTICITY.

When an applied stress is removed, some or all of the deformation may be recovered, with the material returning to its original shape. The term "creep" is used to describe irreversible deformation. The term "elastic memory" is used to describe situations where a polymer is heated and then deformed and then cooled rapidly enough so that the polymer chains are "frozen" in place, with the material retaining the "deformed" shape. *See* CREEP (MATERIALS).

Industrial Polymer Processes

There are about 10,000 American companies active in the general area of synthetic polymers. These companies can be divided into manufacturers, processors, and fabricators. Over 200 companies produce the bulk polymers that are used by the other two groupings of companies. These polymers are generally sent to processor companies as solid pellets in a variety of shapes, including spheres and rice-grain-like. While most manufacturers produce bulk polymers in large quantities, more companies are producing specialty polymers for specific uses.

While some processors make their own polymers, most purchase the raw polymer material from manufacturers. Processors may specialize in a particular product or group of products (such as textiles or tires) or with certain polymers, such as nylons or polyethylenes, or the production of certain forms of material, such as pipes, adhesives, coatings,

composites, or sheets. They may also specialize in a particular type of processing, such as blow molding.

The majority of industrial processors are involved in the fabrication and finishing of polymer-containing products. Fabrication includes machining, forming, and fashioning.

The properties of processed polymers are dependent on the conditions of processing such as cooling rate, temperature, duration of application of heat, and pressure. The drawing or pulling of materials as they exit the processing machine can have a profound effect on the properties of a material. Materials with properties that are the same in all directions are called isotropic. Fibers and films that are pulled in one or two directions are stronger in the direction of the applied stress because of the preferential aligning of the polymer chains in the direction of the pull. Thus, fibers are much stronger in the "length" direction because this is the direction of the applied stress.

There are a number of fabrication techniques. In blow molding, air is blown into a hot tube held in a mold for the production of bottles and other hollow shapes.

In extrusion, a hot melt is forced through a hole, or die, with an opening shaped to produce the desired shape. Extrusion and molding techniques are combined to produce sheet foams. Films can also be made using extrusion when air is forced through a tube. Through control of the air pressure, films of varying thicknesses can be made. Films and sheets can also be produced using calendering, where the hot polymer is forced between tightly fitted rollers. These sheets can be formed into various shapes using thermoforming, where the heated sheet is forced against the contours of a mold by a positive pressure or vacuum.

Injection molding can be used to shape thermoplastics. The material is softened and then forced into a relatively cold mold. Pipes, films, rods, and fibers can be made using injection molding.

In casting, fluid materials are poured into molds of the desired shape and then allowed to cool or cure (crosslink). In some sense, painting can be considered to be a casting operation. *See* PLASTICS PROCESSING.

Applications

The usefulness of polymers depends on their properties, abundance, availability, and cost, including shipping or transportation, manufacture, fabrication, and environmentally related factors. Polymer properties are related to a variety of factors, including chain size, distribution of chain sizes, resistance to chemical attack, degradation (both biological and physical), flammability, dyeability, comfort, ease of fabrication, weatherability, and strength.

Fibers. Fibers are polymer materials that are strong in one direction, and they are much longer (>100 times) than they are wide. Fibers are usually formed by pulling the fiber in one direction. This orients the polymer chains, giving the fiber preferential strength in the direction of the pull. Fibers formed from symmetrical polymers allow the close approach of the polymer chains, resulting in more effective secondary bonding. The attractive forces between chains need to be relatively high, so polymers that engage in dipole-dipole seconding bonding are preferred. Fibers are to have little or no local mobility so that their use temperature is below the T_g and melting point (T_m). A small amount of crosslinking is generally introduced to lock in a preferred shape. **Table 2** contains selected synthetic fibers.

Elastomers. Elastomers (or rubbers) are polymeric materials that can be distorted through the application of force, and when the force is removed, the material returns to its original shape; this characteristic is called memory. Memory is a result of physical or, more normally, chemical crosslinks. The driving force for the return to the original shape is entropy. In the rest state, the polymer chains in the elastomer have a high degree of disorder. As the material is

TABLE 2. Industrially important synthetic fibers

Fiber	Definition	Uses
Acrylic	85% or more acrylonitrile units	Carpeting, skirts, socks, slacks, sweaters, blankets, draperies
Modacrylic	35–85% acrylonitrile units	Simulated fur, scatter rugs, stuffed toys, paint rollers, carpets, hairpieces
Polyester	85% or greater ester units	Permanent press wear, skirts, shirts, slacks, underwear, blouses, rope, fish nets, tire cord, sails, thread
Nylon	Polyamide	Carpeting, upholstery, tents, blouses, sails, suits, stretch fabrics, tire cord, curtains, rope, nets, parachutes
Polyurethane	85% or more urethane units	Girdles, bras, slacks, bathing suits, pillows
Rayon	Regenerated cellulose, with substituents replacing not more than 15% of the hydrogen's of the hydroxyl groups	Dresses, suits, slacks, tire cord, ties, curtains, blankets, blouses
(Rayon) acetate	More than 92% OH substituted	Dresses, shirts, slacks, cigarette filters, upholstery, draperies
Triacetate	Reaction of cellulose with acetic acid or acetic anhydride	Skirts, dresses, sportswear
Fibrous glass		Composites, insulation
Olefins	Polypropylene, polyethylene	Synthetic grass, rugs

TABLE 3. Synthetic elastomers

Elastomer	Use
Acrylonitrile-butadiene-styrene, ABS	Oil hoses, fuel tanks, pipe, appliance and automotive housings
Butadiene rubber, BR	Tire tread, hose, belts
Butyl rubber, IIR	Inner tubes, cable sheathing, roofing, seals, tank liners, coated fabrics
Chloroprene rubber, polychloroprene, CR	Tire and cable insulation, hoses, footwear, mechanical automotive products
Epichlorohydrine (epoxy copolymers)	Seals, gaskets, wire and cable insulation
Ethylene-propylene rubbers, EP, EPDM	Cable insulation, window strips
Fluoroelastomers	Wire and cable insulation, aerospace applications
Ionomers, mostly copolymers of ethylene and acid-containing monomers reacted with metal ions	Golf ball covers, shoe soles, weather stripping
Natural rubber, polyisoprene, NR	Tires, bushings, couplings, seals, footwear, belting
Nitrile rubber (random copolymer of butadiene and acrylonitrile), NBR	Seals, automotive parts that are in contact with oils and gas, footwear, hoses
Polysulfide	Adhesives, sealants, hose binders
Polyurethanes	Sealing and joints, printing rollers, tires, footwear, wire and cable covering
Silicons, mostly polydimethylsiloxanes	Medical-application body parts, outer space applications, flexible molds, gaskets, seals
Styrene-butadiene rubber, SBR	Tire tread, footwear, wire and cable covering

distorted, the polymer chains become more organized. When the distorting force is released, the elastomer returns to the original disordered state, taking advantage of the increase in disorder or entropy that results when it moves from a situation of greater order to a situation of lesser order. Products with low amounts of crosslinking, such as rubber bands, are more easily deformed or stretched. As the amount of crosslinking increases, the force needed to distort the elastomer increases and the material becomes harder. In order for elastomers to easily stretch, the attractive forces between chains must be low and the material must be above the temperature where local segmental mobility occurs—that is, it must be above the T_g. These properties are found in many hydrocarbon-intense polymers (**Table 3**).

Plastics. Plastics are materials that have properties between fibers and elastomers—they are hard and flexible. Polymers that may be fibers or elastomers may also be plastics (**Table 4**). Crystalline nylon is a good fiber, whereas less crystalline nylon is a plastic.

Coatings and adhesives. Coatings and adhesives are generally derived from polymers that are members of other groupings. For instance, polysiloxanes are elastomers, but also are used as adhesives. Poly(methyl methacrylate) is a plastic and coatings. Coatings or coverings are generally highly viscous (low-flowing) materials. The major use of coatings is for protection. Coatings resist normal wear and tear and degradation due to rain, sunlight, snow, ionizing radiation, and temperature. They protect housing, documents, chairs, stoves, cars, tables, bridges, and so on. Coatings must adhere to the surface to which they are applied. Coatings are generally a mixture of a liquid (vehicle), binder (adhesive), and one or more colorants (pigments). They generally contain a number of additives, which provide added protection against ionizing radiation and mildew, and decreased drying time.

In comparison to coatings that are required to adhere to only one surface, adhesives are used to join two surfaces together. Many good adhesives are also good coating materials. Adhesion for both coatings and adhesives can occur for a number of reasons, including physical interlocking, chemical adhesion where primary bonding occurs between the adhesive and the surfaces being joined, secondary

TABLE 4. Important plastics

Plastic	Use
Epoxies	Coatings, laminates, composites, molding, flooring
Urea-formaldehyde resins	Molding compounds, dinnerware
Phenol-formaldehyde resins	Molding compounds
Melamine-formaldehyde resins	Dinnerware, table tops
Polytetrafluoroethylene	Electrical components, nonsticking surfaces, gaskets
Polypropylene	Automotive parts, toys, housewares, appliance parts
Polystyrene	Containers, recreational equipment, housewares, appliance parts
Poly(vinyl chloride) and copolymers	Pipes, fittings, sheets, flooring materials, automotive parts
Polycarbonates	Tail light lens, bullet-resistant vests, appliance housings, signs, bottles
Polysulfones	Mechanical parts, small appliances, electrical connectors
Poly(phenylene sulfide)	Electrical and mechanical parts
Polyesters	Apparel, home furnishings
Polyethylenes	Containers, bottles, housewares, pipe and fittings
Styrene-acrylonitriles	Appliance housings, housewares
Poly(methyl methacrylate)	Signs, glazing, lighting, fixtures, automotive lenses, solar panels
Poly(phenylene oxide)	Business machine housings, electrical parts, automotive parts

bonding where hydrogen bonding or polar bonding occurs between the adhesive and bonded materials, and resistance to flow (viscosity) where movement is restricted because of the viscous nature of the adhesive material. *See* ADHESIVE.

Some industrially important adhesives and coatings are the following:

- Acrylic acid and acrylic ester polymers
- Acrylonitrile-butadiene copolymers
- Aromatic polyamides
- Butyl rubber
- Cellulose and starch derivatives
- Epoxy resins
- Phenol-formaldehyde
- Poly(alkyl cyanacrylates)
- Polyamides
- Polychloroprene
- Polyisobutylene
- Poly(methyl methacrylate)
- Polysiloxanes
- Polyurethane resins
- Poly(vinyl acetate)
- Poly(vinyl alcohol)
- Poly(vinyl butyral)
- Poly(vinyl chloride)
- Resorcinol-formaldehyde
- Styrene-butadiene copolymers
- Unsaturated polyester resins
- Urea-formaldehyde
- Vinyl acetate-ethylene copolymers

Laminates. The combination of an adhesive and adherent is a laminate. Laminates include products that are produced with wood as the adherent and phenolic-formaldehyde, urea-formaldehyde, epoxy, resorcinol-formaldehyde, and polyester resins as the adhesive. The major wood laminate is plywood. Laminates of paper or textile include Formica and Micarta. Laminates are also used as mechanical, electrical, and general-purpose structural materials. *See* COMPOSITE LAMINATES.

Sealants and caulks. Sealants and caulks provide barriers to gases, liquids, and solids; maintain pressure differences; and moderate thermal and mechanical shock. While adhesives are used for "load transfer" requiring high shear and tensile strengths, sealants act as insulators and shock attenuators and do not require high shear and tensile strengths.

Composites. Composites are materials that contain strong fibers (reinforcement) embedded in a matrix or resin. Composites are found in automobiles, fighter and bomber planes, the "reusable" space shuttle, shafts of golf clubs (graphite), synthetic body parts, and marine craft (fibrous glass). *See* POLYMERIC COMPOSITE.

Films. Films are two-dimensional forms of plastic that are thick enough to be coherent, but thin enough to be easily folded and creased without cracking. Sheeting is a two-dimensional form of plastic that is generally thicker than a film but generally does not easily fold or crease without cracking.

Polyblends. Polyblends are made by mixing polymeric components together. Most are heterogeneous materials in which one polymeric material is embedded in another polymer. The components of polyblends adhere using secondary forces.

Liquid crystals. Liquid crystals (LCs) are materials that undergo physical reorganization where at least one of the rearranged structures involve molecular alignment along a preferred direction, causing the material to exhibit behavior that is directionally dependent. LCs serve as one of the major building blocks of communications industries, including computers and active signs. *See* LIQUID CRYSTALS.

Ceramics. Most ceramics contain large amounts of inorganic polymeric materials. While traditional ceramics were based on natural clays, today's ceramics generally are based on synthetic materials. Ceramics are generally brittle, strong, resistant to chemicals, and high-melting. The new high-temperature superconductors are ceramic material.

Cements. The term "cement" is used to describe a wide variety of materials, such as the adhesive materials used by dentists and the concrete (portland cement) used in constructing roads and buildings. Most cements are based on polymeric materials. *See* CEMENT.

Smart materials. Most smart materials are polymeric. They are materials that react to an externally applied force (electrical, bending, pressure, temperature, light, magnetic), and then the interaction is used to achieve a defined engineering or scientific goal. Windows that regulate the amount of light that is transmitted through them as the amount of sunlight varies are an example of smart materials because the windows are part of a more complete energy-control system.

Additives

Processed polymeric materials are generally a combination of the polymer and the materials that are added to modify its properties, assist in processing, and introduce new properties. The latter materials are called additives.

Additives can be solids, liquids, or gases. Thus, for many latex paints, the polymer is poly(methyl methacrylate), titanium IV oxide is added as the pigment, and other materials are added to assist the paint to "cover" the surface, to protect the material from ultraviolet radiation degradation, and to help it resist mildew. *See* INHIBITOR (CHEMISTRY).

Some typical additives are:

- Antiblocking agents
- Antifoaming agents
- Antifogging agents
- Antimicrobial agents
- Antioxidants
- Antistatic agents
- Antizonates
- Blowing agents
- Coloring agents
- Coupling agents
- Curing agents
- Extenders
- Fillers
- Finishing aids
- Flame retardants
- Foaming agents
- Impact modifiers
- Low-profile materials
- Lubricants
- Mold release agents
- Odorants/fragrances
- Plasticizers
- Preservatives
- Reinforcements
- Slip agents
- Softeners
- Stabilizers—radiation and heat
- Viscosity modifiers

Plasticizers. Plasticizers are added to increase the flexibility and to reduce the melt temperature range and viscosity. Most plasticizers can be considered

to be near-solvents that allow segmental mobility to occur. In natural systems, water often acts as a plasticizer, allowing proteins, nucleic acids, and polysaccharides to be flexible. For synthetic polymers, most plasticizers are designed to be nonmobile and nonvolatile.

Antioxidants. Antioxidants are added to retard oxidative degradation generally brought about by the formation of free radicals through application of heat, ionizing radiation, or mechanical shear. *See* ANTIOXIDANT.

Colorants. Many materials contain colorants or coloring agents. Colorants are generally added to increase the esthetic value of a coating.

Fillers and reinforcements. Fillers are relatively inert materials, while reinforcements improve the properties of the material to which they have been added. For many composites, addition of reinforcing fibers produces dramatic improvements in the strength of the material. While most fillers are added to simply increase the bulk of the material, some provide needed increases in the mechanical properties of the material. Fillers are generally inexpensive, lightweight, and readily available in large quantity. Fillers include natural organic materials such as wood product discard materials, corncobs, and peanut shells. Glass of many forms as well as a number of minerals (including sand and diatomaceous earth, mica, and talc) are also used as fillers.

Recycling and Degradation

Polymers offer a wide variety of techniques for recycling and degradation. Recycling of simple polymeric materials such as bottles and containers is fairly straightforward. Many polymers are thermoplastics, that is, they can be reshaped through application of heat and pressure and used in the production of other thermoplastic materials. The recycling of thermosets, polymers that do not melt but degrade prior to softening, is more difficult. These materials are often ground into a fine powder, are blended with additives (often adhesives or binders), and then are reformed.

Most glass contains some reused glass, and much paper contains recycled paper.

The recycling of more complex objects such as tires is more difficult since they are composed of a number of different materials that must be separated before real recovery of the material is possible. *See* RECYCLING TECHNOLOGY.

The mechanism for the degradation of polymers can be divided into two general catagories—biological and physical. Many microorganisms contain enzymes that are "sent out" by the organism to locate food and render it usable to the organism. Some of these organisms have enzymes that degrade the polar linkages that are present in polyesters, polyamides, and other condensation polymers. Thus, these polymers can be degraded by microorganisms. *See* BIODEGRADATION.

Physical degradation includes "simple" hydrolysis (with catalysis by the acid found in acid rain), heat, and light. Again, condensation polymers are particularly sensitive to hydrolysis. Many vinyl polymers are degraded through exposure to light as well. *See* HYDROLYSIS.

The Future

Materials in today's society must be varied to perform many tasks. Often they must perform them repeatedly and in a special manner. Plastic hinges must be able to work thousands of times. The human heart, a complex muscle largely made up of protein polymers, provides about 2.5 billion beats within a lifetime, moving oxygen throughout the approximately 144,000 km of a circulatory system and delivering about 8000 liters of blood daily with little deterioration of the cell walls. Nerve impulses travel within the body about 300 m per minute, and macromolecules are the materials that allow this rapid and precise transfer of nerve impulses. Bones are made of polymeric materials, and have a strength about five times that of steel on a weight basis. In this century, synthetic polymers should approach the properties of these biological polymers.

Polymers are part of the so-called nanorevolution where single polymer strands, one atom thick, will be the connectors; polymer sheets one molecule thick will act as adhesives and insulators for communication (including computers); small motors the breadth of 1/20th of a human hair will go from zero to 500,000 revolutions per second in a real "blink of an eye." *See* NANOTECHNOLOGY.

Catalysis for production of polymers and the use of polymers as catalysts continue. The presence of soluble stereoregulating catalysts used in the production of polymers will allow better control and enhanced properties for these materials. The use of polymers for multisite synthesis will continue to be developed. *See* CATALYSIS.

The use of superstructures will continue, including in the mimicry of nature. For instance, the mucus of snails contains protein "ball-bearing-like" spheres: efforts are under way to mimic these proteins using synthetic polymeric materials. The synthesis of biomedical materials, including skin and tendons, will continue along with the creation of materials for replacement of structural body parts.

Manufacturing companies are dividing into those that produce bulk polymers on a large scale and others that produce specialty polymers on a smaller scale. The development, use, and production of specialty polymers for specific applications will increase. The use of, and improvement in properties of, the older "standby" polymers such as polyethylene will also continue. Charles E. Carraher, Jr.

Bibliography. H. Allcock, F. Lampe, and J. Mark, *Contemporary Polymer Chemistry*, 3d ed., 2003; E. A. Grulke et al. (eds.), *Polymer Handbook*, 2 vols., 4th ed., 2003; G. Odian, *Principles of Polymerization*, 4th ed., 2004; M. P. Stevens, *Polymer Chemistry: An Introduction*, 3d ed., 1998; A. N. Wilkinson and A. J. Ryan (eds.), *Polymer Processing and Structure Development*, 1999.

Polymer composites

Polymer materials or matrices that are combined with fillers or fibers. Fillers are intentionally placed in polymers to make them stronger, lighter, electrically conductive, or cheaper. Any filler will affect the mechanical behavior of a polymeric material. For example, long fibers will make the material stiffer but usually denser, whereas foaming will make it more compliant but much lighter. A filler such as calcium carbonate will decrease the polymer's toughness, while making it considerably cheaper (**Fig. 1**). Polymer composites are also found in nature, such as wood (cellulose fibers bonded with lignin) and bone (minerals bonded with collagen).

Polymer composites have become prevalent in automotive and aircraft design and manufacture. For example, newer commercial aircraft can have up to 80% of their structure made of polymer composites, including the whole fuselage. The automotive industry introduced composites in their body panel design in the early 1950s. Today, many automobiles not only have body panels that are compression- or injection-molded out of fiber-reinforced polymers, but also polymer-composite structural components as well as many under-the-hood applications such as valve covers and oil pans. Other applications of polymer composites include filament-wound pressure vessels and pipes for the chemical industry, braided fiber–reinforced epoxy tubes used in the construction of composite bicycles, and resin transfer–molded boats, skies, and snowboards, to name a few. *See* COMPOSITE MATERIAL.

Polymer matrix. In a composite structure, the task of the polymer matrix is to hold the fibers in place or bond them together into one integral structure. Polymer matrices can be either thermoplastic or cross-linked polymers. Thermoplastic polymers are divided into amorphous and semicrystalline thermoplastics, and cross-linked polymers can be either elastomers or thermosets. *See* POLYMER.

Amorphous thermoplastics. Amorphous thermoplastics, such as polystyrene, polycarbonate, and acrylics, exhibit "useful" behavior below a temperature called the glass transition temperature. An amorphous polymer will flow when subjected to temperatures above its glass transition. Amorphous polymers are brittle below their glass transition temperature and therefore fail at low strains. As the temperature increases, the strength of the amorphous thermoplastic decreases, since the thermoplastic becomes leathery in texture and is able to withstand larger deformations. Above the glass transition temperature, the strength decreases significantly, as the maximum strain continues to increase, until the flow properties have been reached, at which point the mechanical strength is negligible.

Some amorphous thermoplastics can be made highly impact resistant (less brittle) through copolymerization. The most common example is acrylonitrile-butadiene-styrene (ABS). Since butadiene chains vitrify (become glasslike) at temperatures below −50°C (−58°F), ABS is very tough at room temperature in contrast to polystyrene and acrylics by themselves. *See* BRITTLENESS; COPOLYMER; PLASTICITY; STRESS AND STRAIN.

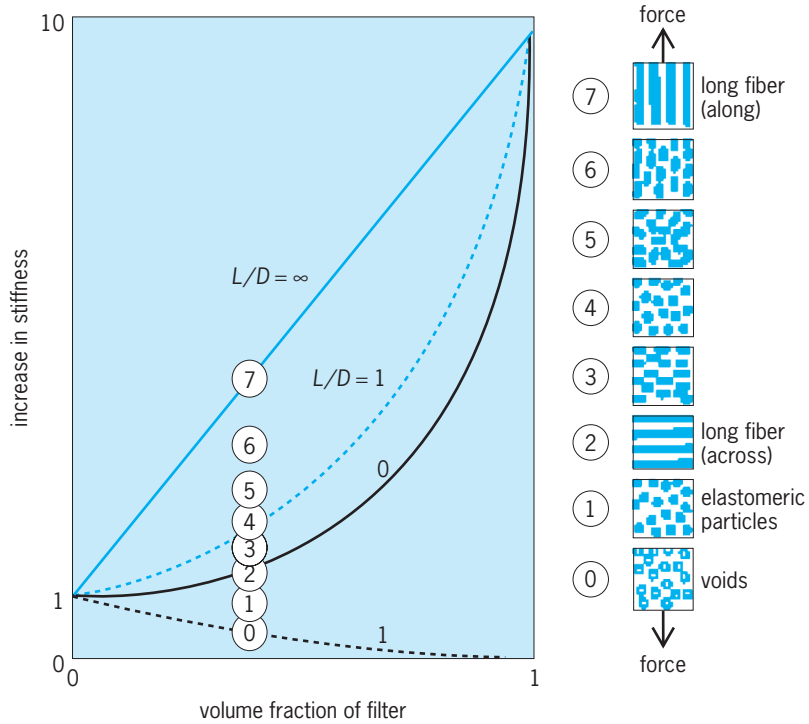

Fig. 1. Relation between stiffness and filler type and orientation in polymeric materials. *L/D* = length/diameter of fiber.

Semicrystalline thermoplastics. Semicrystalline thermoplastics, such as polyethylene, polypropylene, and polyamide, are leathery and tough at room temperature, since their amorphous regions usually vitrify at very low temperatures; for example, polypropylene has a typical glass transition temperature of −10°C (14°F). The temperature at which a semicrystalline polymer becomes useless is the melting temperature. The melting temperature and the strength and stiffness of semicrystalline plastics increase with the degree of crystallinity. The brittle behavior that polypropylene would exhibit below its glass transition temperature can sometimes pose a problem in design. However, this can be mitigated through copolymerization, as polypropylene is often copolymerized with ethylene. With semicrystalline polymers, an increase in toughness between the glass transition temperature and the melting temperature is observed. Room temperature usually falls in a range between the glass transition temperature and the melting point for most semicrystalline thermoplastics.

Cross-linked polymers. Cross-linked polymers, such as thermosets and elastomers, behave completely differently from thermoplastic polymers. A coarse cross-linked system, typical of elastomers such as natural rubber and styrene-butadiene rubber, renders a material that will no longer flow and that exhibits a low modulus (stiffness) above the glass transition temperature. The glass transition temperature of these materials is usually below −50°C (−58°F), so they are soft and flexible at room temperature. On the other hand, highly cross-linked

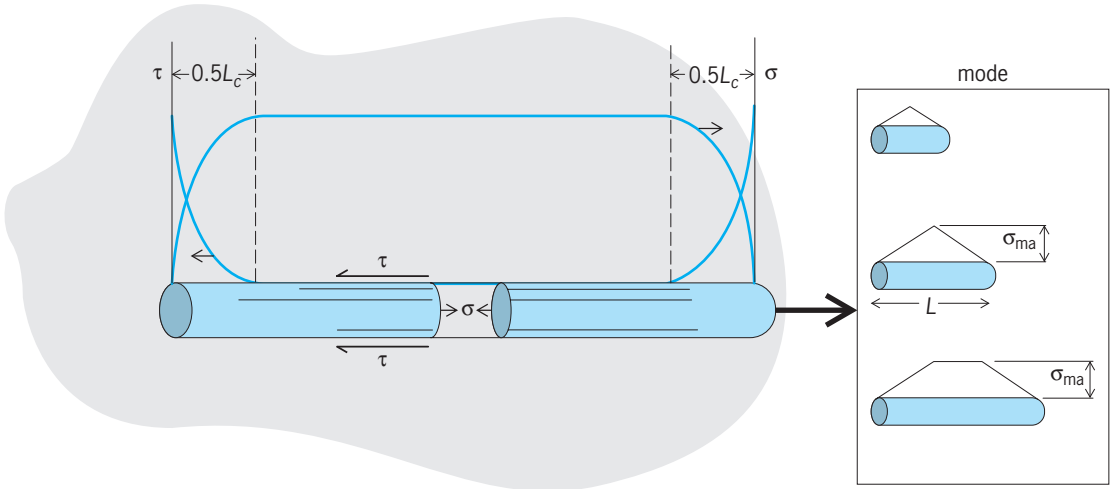

Fig. 2. Load transfer from matrix to fiber in a composite. L_C = critical length. σ_{ma} = matrix stress.

systems, typical in thermosets such as epoxy and unsaturated polyester, are very stiff and show only a small decrease in stiffness as the material is raised above the glass transition temperature; that is, the decrease in properties becomes smaller as the degree of cross-linking increases. The strength of a thermoset remains fairly constant up to the thermal degradation temperature of the material. *See* RUBBER.

Fiber reinforcement. In a well-designed composite material, the fiber is the load-bearing structure of a part. The purpose of introducing a fiber into a matrix is to transfer the load from the weaker material to the stronger one. This load transfer occurs over the length of the fiber (**Fig. 2**). The length necessary to complete the load transfer from the matrix to the fiber, without fiber or matrix fracture, usually is referred to as the critical length. The critical length is a function of fiber diameter, shear strength of the matrix, and tensile strength of the fiber. Experimental evidence suggests that fiber aspect ratios (fiber length to diameter) of 100 or higher are required to achieve maximum strength in a fiber-reinforced composite structure. If composites have fibers that are shorter than their critical length, they are referred to as short-fiber composites; and if the fibers are longer, they are referred to as long-fiber composites.

Traditionally, fibers were glass or carbon. However, fibers also can be made out of polymers by orienting or aligning the long molecular chains. The properties of polymeric fibers (polyethylene and aramid) are comparable to the properties of graphite and glass fibers, differing mostly in that polymers are less dense than inorganic fibers (see **table**). *See* CARBON; GLASS; GRAPHITE; MANUFACTURED FIBER.

Molecular orientation and polymeric fibers. If a thermoplastic is stretched at a temperature high enough that the polymer chains can slide past each other but low enough that stresses can build up, the orientation generated during stretching is retained within the polymer component. Any degree of molecular orientation results in property variations within thermoplastic polymers. The stretching will lead to decreased strength and stiffness properties perpendicular to the orientation, and increased properties parallel to the direction of deformation. In addition, highly oriented materials tend to split along the orientation direction under small loads. *See* STRENGTH OF MATERIALS.

If the material is stretched such that a needle-like, or fibrilic, morphological structure results, the stiffness of the material will be very high. The high strength and stiffness of fibrilic structures is taken advantage of when making synthetic fibers. **Figure 3** shows the theoretical and achievable elastic moduli of various synthetic fiber materials. *See* ELASTICITY; YOUNG'S MODULUS.

High-stiffness and high-strength synthetic fibers are increasingly important for lightweight high-strength applications. Extended-chain ultrahigh-molecular-weight polyethylene fibers have been available commercially only since the mid-1980s. They are manufactured by drawing or extending fibers of small diameters at temperatures below the melting point. The modulus and strength of the fibers increase with the drawing ratio or stretch. The tensile modulus of superdrawn ultrahigh-molecular-weight high-density polyethylene fibers reaches a maximum modulus of 200 GPa (the modulus of steel) at draw

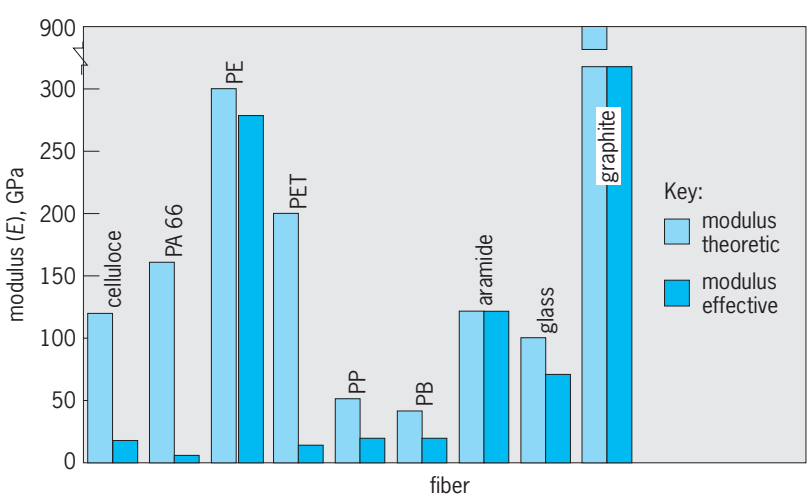

Fig. 3. Tensile modulus for various fibers.

ratios of 250. A draw ratio is the ratio of the final length to the initial length of the fiber. In addition to amorphous and semicrystalline thermoplastics, there is a whole family of thermoplastic materials whose molecules do not relax and thus retain their orientation, even in the melt state. This class of thermoplastics is the liquid crystalline polymers. One such material is aramid fiber, commonly known as Kevlar®, which has been available in the market for several years.

Mechanical properties. Fiber-reinforced polymers are often anisotropic; that is, they show different properties in different directions. As a result, the usual relations that are used to represent and analyze homogeneous materials are not valid. The most common model to analyze fiber-reinforced composites is the two-dimensional laminate structure (**Fig. 4**). In this simplified case, L and T define the longitudinal and transverse directions, respectively.

The longitudinal and transverse properties can be calculated using (among other models) the Halpin-Tsai model. The actual behavior of the composite is related not only to the individual properties of the matrix and the fibers but also to the volume fraction of fibers or fillers present in the composite, as well as the orientation of the loads relative to the fiber direction. In addition to the Halpin-Tsai model, several other models are in use today to predict the elastic properties of aligned fiber-reinforced laminates. Most models predict the longitudinal modulus quite accurately, but not the transverse. The longitudinal modulus is directly proportional to the fiber content for composites with unidirectional reinforcement, increasing linearly with respect to the fiber volume fraction. The increase of transverse modulus as a function of fiber content is not as dramatic, and in simple models often is approximated to be equal to the stiffness of the matrix.

The loading in a laminated structure is not always aligned with the transverse and longitudinal directions of the reinforcement. In other words, the laminate and its properties are often rotated by an angle θ. The stiffness of a laminate decreases as one rotates away from the longitudinal axis for an aligned fiber–reinforced composite with different volume-fraction fiber content (**Fig. 5**). For high volume-fraction fiber contents, only a slight misalignment of the fibers from the loading direction results in a drastic reduction of the properties. This is critical in composite structures that are constructed with aligned fibers, such as aircraft components.

With help of commercially available computer-aided technology (CAD), one can design and predict properties of composites with reinforcing fibers that have orientation distributions that result from flow and deformation during manufacturing. Such systems are typical of compression-molded automotive body panels, as well as injection-molded fiber-reinforced components. Tim A. Osswald

Mechanical properties of selected fibers

Fiber	Tensile strength, MPA	Tensile modulus, GPa	Elongation to break, %	Specific gravity
Polyethylene	3000	172	2.7	0.97
Aramid	2760	124	2.5	1.44
Graphite	2410	379	0.6	1.81
S-glass	4585	90	2.75	2.50

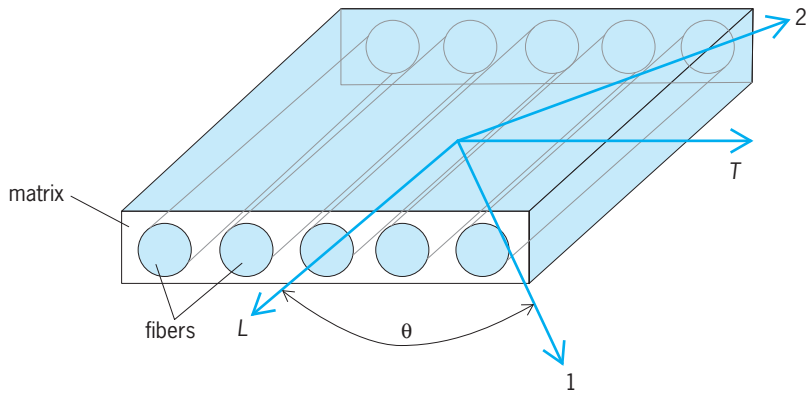

Fig. 4. Unidirectional continuous fiber–reinforced laminated structure. L = longitudinal direction. T = transverse direction.

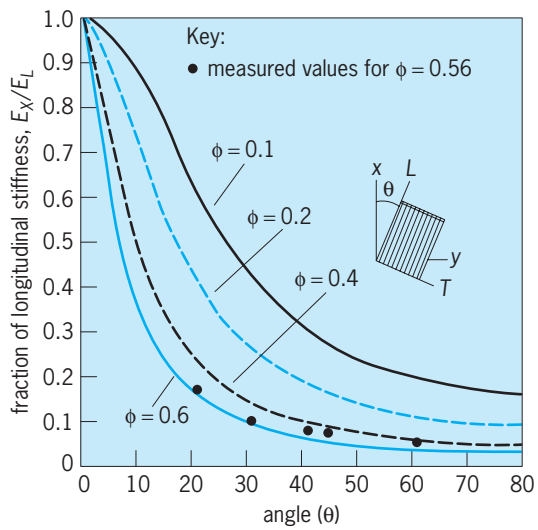

Fig. 5. Measured and predicted elastic modulus in a unidirectional fiber–reinforced laminate as a function of angle between the loading and fiber direction.

Bibliography. G. W. Ehrenstein, *Faserverbund-Kunststoffe*, 2d ed., Hanser, Munich, 2006; S. I. Krishnamachari, *Applied Stress Analysis of Plastics*, Van Nostrand Reinhold, 1993; L. E. Nielsen and R. F. Landel, *Mechanical Properties of Polymers and Composites*, 2d ed., Marcel Dekker, 1994; T. A. Osswald, E. Baur, and E. Schmachtenberg, *International Plastics Handbook*, Hanser, Munich, 2006; T. A. Osswald and G. Menges, *Materials Science of Polymer for Engineers*, 2d ed., Hanser, Munich, 2003; S. W. Tsai, J. C. Halpin, and N. J. Pagano, *Composite Materials Workshop*, Technomic, 1968.

Polymer stereochemistry and properties

The properties of polymers are significantly influenced by their stereochemistry. Cellulose (structure **1**) and amylose (**2**), the main components of

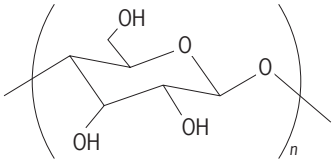

(1)

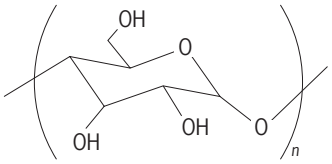

(2)

wood and starch, respectively, are among the most abundant polymers on Earth. Both are produced from D-glucose and are stereoisomers of each other. These two have different linkages between 1- and 4-carbons; the former has a 1,4-β-structure and the latter a 1,4-α-structure. This difference in the linkage of these polysaccharides leads to different higher-order structures of the polymer chains; cellulose has a stretchable polymer chain, which facilitates the regular arrangement of the polymer chains through the interchain hydrogen bond, whereas amylose has a helical chain, in which about seven glucose residues construct one turn of a helix. Because of their structural differences, their properties are also very different; for instance, cellulose is not soluble in water because of the strong interchain hydrogen bonds, but amylose is. Human beings cannot digest cellulose to D-glucose to use as an energy source, but can digest amylose or starch to D-glucose.

Vinyl polymers. When vinyl monomers ($CH_2{=}CH{-}X$) like propylene ($X = CH_3$), styrene ($X = C_6H_5$), vinyl acetate ($X = OCOCH_3$), and *N*-isopropylacrylamide ($X = CONHCH(CH_3)_2$) are polymerized, each step of the monomer addition to a propagating chain end yields an asymmetric *R* or *S* center. The continuing formation of the same center produces an isotactic polymer (**3**), and the alternating formation of the *R* and *S* centers produces a syndiotactic polymer (**4**). This kind of arrangement of side chains

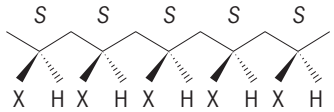

Isotactic

(3)

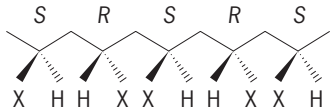

Syndiotactic

(4)

has been called stereoregularity or tacticity. The control of tacticity was first realized by G. Natta in 1956 for the stereospecific coordination polymerization of propylene using the Ziegler-Natta catalyst consisting of $TiCl_3{-}Al(C_2H_5)_3$. The isotactic polymer had properties superior to those of the atactic polypropylene, which has randomly arranged side-chain methyl groups. Isotactic polypropylene is a hard, durable material with a high melting point (~165°C or 330°F) and high mechanical strength, whereas the atactic polypropylene is a greasy material. Isotactic polypropylene is one of the most widely used polymers.

Nearly 100 million tons per year of the polyolefins polypropylene and polyethylene are produced worldwide. The production of these polyolefins is steadily increasing due to their low cost and high performance. They have been used for a number of products, such as grocery bags, containers, gasoline tanks, and automobile bumpers. The Ziegler-Natta catalyst is an insoluble and heterogeneous catalyst, and therefore control of the polymerization reaction with the catalyst is difficult. This defect has been improved by the discovery of new homogeneous catalyst systems composed of metallocenes (**5**, **6**).

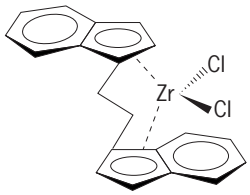

(5)

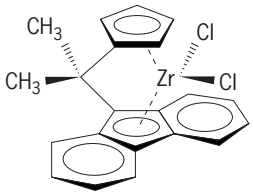

(6)

By changing the ligand structures on zirconium, the tacticity of polypropylene can be varied from highly isotactic to syndiotactic.

Polystyrene is also a very popular polymer and has been industrially prepared by radical polymerization, which usually produces an atactic polymer. The polymer is used as kitchen utensils, video cassettes, and machine enclosures. Styrene is also polymerized by coordination polymerization using a catalyst containing a half metallocene (**7**), which gives a highly

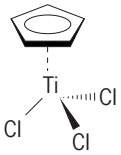

(7)

syndiotactic polymer. The properties of this stereoregular polymer significantly differ from those of the atactic polymer. Polystyrene obtained by radical polymerization becomes soft at around 100°C (212°F), while the syndiotactic polymer has high crystallinity and a melting point as high as 270°C (520°F). Syndiotactic polystyrene has both high thermal and chemical stability and therefore can be used as an engineering plastic.

The properties of poly(vinyl alcohol) (**8**) are also

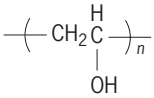

(8)

significantly influenced by its stereoregularity, and both the isotactic and syndiotactic polymers are known to have higher melting points than the atactic polymer. Poly(vinyl alcohol)s with different tacticities have been prepared by cutting the ether linkage of poly(vinyl ether)s obtained by cationic polymerization under various conditions. However, this method is not a practical process. The polymer is manufactured by the saponification (hydrolysis) of poly(vinyl acetate), which is obtainable only by the radical polymerization of vinyl acetate. For this polymerization, there exist almost no temperature and solvent effects on tacticity, making it is difficult to change the properties of poly(vinyl alcohol) by these processes.

N-isopropylacrylamide is also polymerized only by a radical process. The stereochemistry of this polymerization is difficult to change by the reaction conditions, as is the polymerization of vinyl acetate. Because the polymer exhibits a unique phase transition in an aqueous solution and becomes a crosslinked gel at around 32°C (90°F), it has been extensively studied without paying attention to the polymer structure, particularly the tacticity. However, recent progress in the stereocontrol during the radical polymerization of *N*-isopropylacrylamide using a catalytic amount of Lewis acids, such as the trifluoromethanesulfonates of yttrium (Y) and ytterbium (Yb), enables the production of isotactic-rich polymers. The phase-transition temperature of the isotactic-rich polymer decreases with an increase in the isotacticity.

Polyester and polyamide. A polyester, poly(L-lactic acid) (**9**), is prepared from L-lactic acid, which is man-

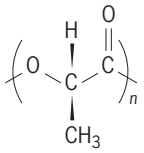

(**9**)

ufactured by the fermentation of glucose from starch. The polymer has the continuous L-configuration, and is a totally stereoregular and optically active polymer. This polymer has drawn much attention as being an environmentally friendly, biodegradable polymer. Its stereoisomer, poly(D-lactic acid), is prepared from D-lactic acid. Most properties of these two polymers are identical, but they differ in biodegradability. The degradation is catalyzed by enzymes, which often discriminate enantiomers like poly(L-lactic acid) and poly(D-lactic acid). To change the polymer properties, the preparation of the atactic polymer containing both L- and D-isomers is effective.

Protein analogue homopolymers, that is, poly(L-amino acids) (**10**), have been synthesized from

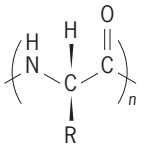

(**10**)

L-amino acids and are totally stereoregular polyamides. The incorporation of D-amino acids into the L-amino acid sequences yields an atactic polymer, the properties of which significantly differ from those of the corresponding homopolymers.

Higher-order structure: helicity. Most stereoregular polymers have a helical conformation at least in the solid state, and some of them can maintain the helicity in solution. It is also well known that most proteins have a partial helical conformation to exert their sophisticated functions in the living systems, and therefore some poly(L-amino acid)s also have a helical conformation in solution. Polymethacrylate (**11**), poly(*N*,*N*-disubstituted acrylamide) (**12**), and polychloral (**13**) are examples that can be synthe-

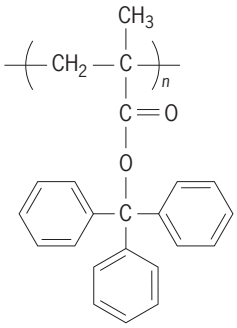

(**11**)

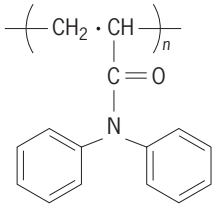

(**12**)

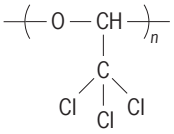

(**13**)

sized as one-handed helical polymers using optically active initiators. Due to the steric hindrance of the bulky side chains, these polymers can maintain the helical structure that is constructed via the polymerization processes. Some of the helical polymers exhibit high chiral recognition and have been used as the chiral stationary phases in high-performance liquid chromatography (HPLC) to separate enantiomers.

Polyisocyanates (**14**), stereoregular polyacetylenes (**15**) with the cis-transoid structure, and polysilanes (**16**) have a dynamic conformation,

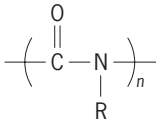

(**14**)

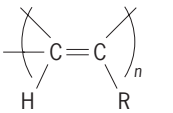

(**15**)

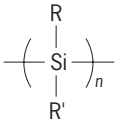

(**16**)

which can be changed through the incorporation of a tiny chiral component or through the interaction with chiral stimuli. For instance, poly(*N*-hexyl isocyanate) (**12**, R = n—C_6H_{13}—) takes a prevailing one-handed helical structure by introducing a 1% optically active R group. Poly(4-carboxyphenylacetylene) (**13**, R = —C_6H_4—COOH) can change its helicity through the interaction with

chiral amines. This change is readily detected by circular dichroism (CD) spectroscopy, which shows intense CD peaks in the ultraviolet (UV) and visible regions above 300 nanometers depending on the chirality of the amines. Because the amines with analogous structures induce the same CD pattern, the CD measurement can identify the absolute configuration of the chiral amines. The behavior of polysilanes (**16**) is similar to that of polyisocyanates (**14**), and the introduction of a very small amount of a chiral side chain forces the entire polymer chain to have the same helicity. *See* METALLOCENE CATALYST; POLYACRYLATE RESIN; POLYESTER RESINS; POLYMER; POLYMERIZATION; POLYOLEFIN RESINS; POLYSTYRENE RESIN; POLYURETHANE RESINS; POLYVINYL RESINS; STEREOCHEMISTRY. Yoshio Okamoto

Bibliography. M. Farina, Stereochemistry of linear macromolecules, *Top. Stereochem.*, vol. 17, no. 1, 1987; T. Nakano and Y. Okamoto, Synthetic helical polymers: Conformation and function, *Chem. Rev.*, 101:4013, 2001; Y. Okamoto et al., *Lewis Acid-catalyzed Tacticity Control during Radical Polymerization of (Meth)acrylamides*, ACS Symp. Ser. 854, ed. by K. Matyjaszewski, vol. 59, 2003.

Polymer-supported reaction

An organic chemical reaction where one of the species, for example, the substrate, the reagent, or a catalyst, is bound to a cross-linked, and therefore insoluble, polymer support. A major attraction of polymer-supported reactions is that at the end of the reaction period the polymer-supported species can be separated cleanly and easily, usually by filtration, from the soluble species. This easy separation can greatly simplify product isolation procedures, and it may even allow the polymer-supported reactions to be automated. Often it is possible to reuse or recycle polymer-supported reactants. Partly for this reason and partly because they are insoluble, involatile, easily handled, and easily recovered, polymer-supported reactants are also attractive from an environmental point of view.

Features. Polymer-supported reactants are usually prepared in the form of beads of about 50–100 micrometers' diameter. Such beads can have a practically useful loading only when their interiors are functionalized, that is, carry functional groups. With a typical polymer-supported reactant, more than 99% of the reactive groups are inside the beads. An important consequence is that for soluble species to react with polymer-supported species the former must be able to diffuse freely into the polymer beads.

Two major morphological types of polymer beads have been used: microporous and macroporous. The former are lightly cross-linked gels, and they are significantly porous only when swollen by solvent. This property makes the choice of reaction solvent for use with a microporous polymer-supported reactant crucial, especially as the swelling properties of the beads may change substantially as the supported functional groups undergo chemical change. Macroporous beads have a highly open structure that is maintained by a high degree of cross-linking. While functional groups within the pores are accessible in many solvents, functional groups in the highly cross-linked and relatively rigid structure surrounding the pores are much less accessible. Thus, macroporous polymer beads effectively have a lower maximum capacity than microporous polymer beads. *See* GEL; POLYMER.

The high concentration of functional groups in a polymer bead can result in a microenvironment for reaction that is significantly different from that in solution, and this result can lead to distribution effects and to increases or decreases in the rates of polymer-supported reactions compared with their nonsupported analogs. A rigid, very lightly loaded polymer-supported reactant may display a degree of site isolation; that is, reactions between supported groups may be reduced or, to a limited extent, prevented. The use of a spacer arm between the functional groups and the polymer backbone can often enhance reaction rates.

The most common polymer supports are cross-linked polystyrenes. They may be appropriately functionalized by electrophilic aromatic substitution. For less reactive electrophiles the polystyrene beads can be lithiated first. Chloromethylated polystyrene beads are easily functionalized by reaction with a wide range of nucleophiles. *See* ELECTROPHILIC AND NUCLEOPHILIC REAGENTS.

Substrates. In a typical application of reactions involving polymer-supported substrates, a substrate is first attached to an appropriately functionalized polymer support. The synthetic reactions of interest are then carried out on the supported species. Finally, the product is detached from the polymer and recovered. Often the polymer-supported species has served as a protecting group that is also a physical "handle" to facilitate separation. The fact that it is not possible to separate one polymer-supported species from another means, however, that in most cases it is usually necessary to carry out all the polymer-supported reactions cleanly and in high yields. As a result of much research effort, this goal has proved possible in the extremely important cases of "solid-phase" peptide and nucleotide syntheses, but such is not always the case.

Two other types of application of polymer-supported substrates have proved to be of interest. The first exploits easy attachment reactions, for example, the reactions of polymer-supported boronic acid (**1**) with diols to give cyclic boronate

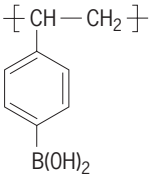

(**1**)

esters, and the reactions of polymer-supported diols such as polymer (**2**) with ketones to give ketals. Such

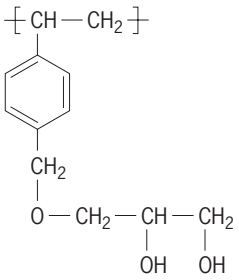

(**2**)

reactions can form the basis of a type of affinity chromatography. Thus, if the polymer-supported diol (**2**) is reacted with a mixture containing ketones, the latter will bind to the polymer through ketal formation. After washing away unreacted species, ketal formation is reversed and the ketones released, free of other types of species. *See* LIQUID CHROMATOGRAPHY.

The second application involves isolating (fishing out) interesting species formed in small amounts. The classical example is concerned with the synthesis of rotaxane (**3**). The capping of decane-1,10-diol

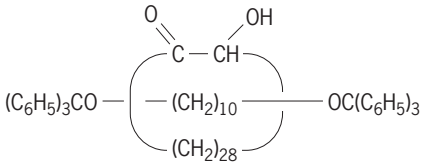

(**3**)

with trityl groups was carried out in the presence of the polymer-supported cyclic compound (**4**). At

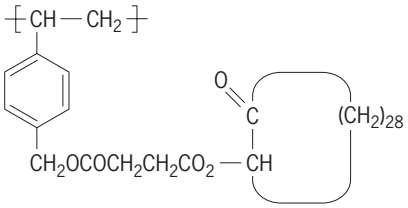

(**4**)

the time of the capping reaction, a small amount of the decane derivative was threaded through the 30-membered ring of compound (**4**); and because of the large terminal trityl groups, it was trapped, that is, a rotaxane was formed. The yield of rotaxane was less than 1%. However, by repeatedly reusing the polymer-supported species the amount of polymer-supported rotaxane was built up to a significant level. The ring was then detached from the support by hydrolysis of the ester linkages, and rotaxane (**3**) was isolated.

Polymer-supported reagents. Reactions involving polymer-supported reagents are generally more useful than those involving polymer-supported substrates, because no attachment or detachment reactions are needed, and it is not necessary for all the polymer-supported species to react in high yield. Indeed, polymer-supported reagents are often used in excess to drive reactions to high conversions.

The most accessible polymer-supported reagents are various anionic forms of commercial macroporous anion-exchange resins. An example is a compound with structure (**5**), where X^- represents

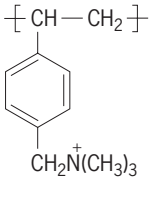

(**5**)

one of the following anions: hydrogen chromate ($HCrO_4^-$), borohydride (BH_4^-), fluoride (F^-), cyanide (CN^-), or periodate (IO_4^-). These can be considered as polymer-supported quaternary ammonium salts; since the support is macroporous, they can be used in a wide variety of solvents. *See* QUATERNARY AMMONIUM SALTS.

One of the most interesting applications of polymer-supported reagents involves the use of a mixture of two of them. For example, cleavage of certain 1,2-diols by IO_4^- gives rise to very sensitive dialdehydes. By using polymer-supported IO_4^- and BH_4^- anions together, the sensitive products of diol cleavage were immediately reduced to stable alcohols. Since the polymer-supported oxidizing and reducing agents are on separate beads, they do not come into contact and so cannot destroy each other. However, the species in solution can pass freely between the two reagents, and the overall reaction proceeds smoothly.

Polymer-supported reagents can also be covalently bound species, and approximately a hundred polymer-supported reagents of this type have been described. These include polymer-supported carbodiimides, peroxy acids, phosphorus and sulfur ylides, acylating agents, and bases such as polymer-supported dimethylaminopyridine (**6**).

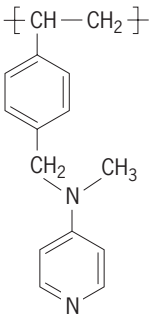

(**6**)

Polymer-supported catalysts. These are the most attractive type of polymer-supported reactants; in such

reactions, the loadings of catalytic sites need not be high, not all sites need be active and, in most cases, the polymer-supported catalyst is recovered in a suitable form for reuse. Accordingly a large number and variety of polymer-supported catalysts have been studied.

The earliest examples were polymer-supported sulfonic acids (**7**), that is, the acid forms of com-

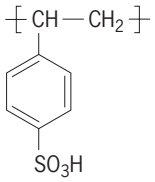

(**7**)

mercial strong acid cation-exchange resins. These have been used to catalyze, for example, esterifications, ester hydrolyses, ketalizations and deketalizations. Many variations have been prepared that are even stronger acids, for example, fluorinated acids and complexes formed from polymer-supported proton acids with Lewis acids. Polymer-supported acids are used in a flow system for the synthesis of methyl *t*-butyl ether, the favored antiknock additive to gasoline. Polymer-supported bases have also been widely studied. *See* GASOLINE.

An area of major interest is polymer-supported phase-transfer catalysts, because of the relative ease of studying some of the many factors involved. Polymer-supported quaternary ammonium and phosphonium salts, polymer-supported crown ethers, and polymer-supported poly(ethylene glycols) have been used to catalyze a wide range of nucleophilic displacements and other reactions such as dichlorocarbene generation. *See* MACROCYCLIC COMPOUND; POLY(ETHYLENE GLYCOL).

An area of great commercial interest involves polymer-supported transition-metal complex catalysts for hydrogenation, hydroformylation, olefin isomerization, or olefin oligomerization, and analogs of homogeneous transition-metal complex catalysts have been prepared by replacing some of the ligands with a polymer-supported ligand. The most common ligands are polymer-supported phosphines such as the polymer-supported analog of triphenylphosphine (**8**). The commercial interest arises mainly

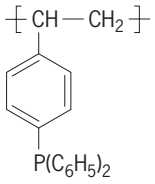

(**8**)

from the possibility of recovering the precious metals. The problem is that the metal tends to leach slowly from the support.

Other types of polymer-supported catalysts include polymer-supported arsonic acid (**9**), and polymer-supported selenic acid (**10**). These have

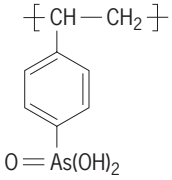

(**9**)

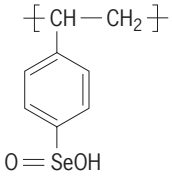

(**10**)

been used to catalyze epoxidations of olefins and the conversion of ketones into esters. Unlike the nonpolymeric analogs, these polymer-supported catalysts are nontoxic, because the polymers are insoluble and involatile. Polymer-supported porphyrin-manganese(III) complexes have also been used to catalyze epoxidations.

Some molecules, including many important compounds found in nature, exhibit a property known as handedness, or chirality. The structural feature commonly responsible for chirality is a tetrahedral carbon atom bearing four different substituents [indicated by an asterisk (*) in the products of reactions (1) and (2)]. Single molecules of this type can exist in two

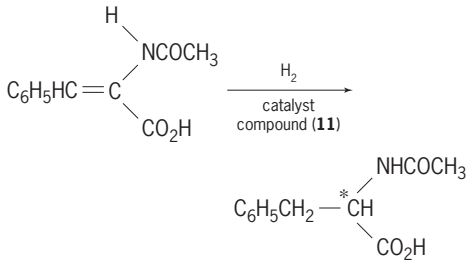

(1)

(*R*)-isomer in 91% *ee*

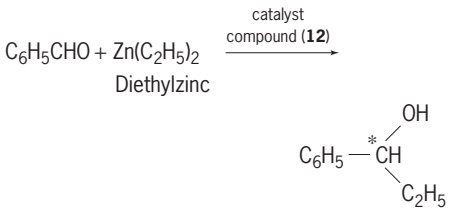

(2)

(*S*)-isomer in 92% *ee*

forms, left-handed and right-handed, known as enantiomers. In the absence of a chiral agent, enantiomers are formed in equal amounts.

Asymmetric synthesis, that is, the synthesis of enantiomers in other than 50:50 ratio, is important, especially as a source of homochiral intermediates in the synthesis of pharmaceuticals. The most efficient type of asymmetric syntheses are those achieved by chiral catalysts. One of the more successful ones

was a polymer containing the bisphosphine moieties [compound (**11**)], which had been treated with tris-

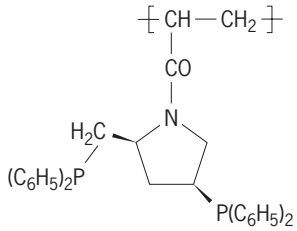

(**11**)

triphenylphosphinerhodium chloride ($[(C_6H_5)_3P]_3$-RhCl). This polymer was used to achieve the asymmetric hydrogenation in reaction (1). Other examples are the use of polymer-supported catalyst (**12**)

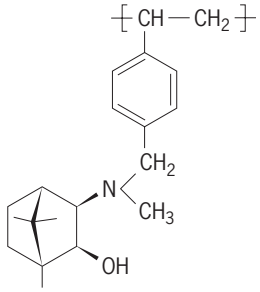

(**12**)

for reaction (2) and polymer-supported catalyst (**13**) for reaction (3).

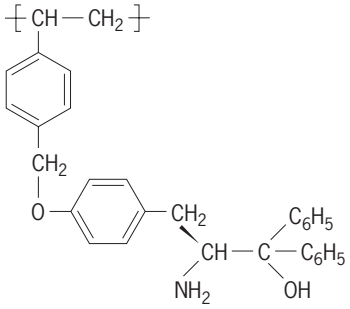

(**13**)

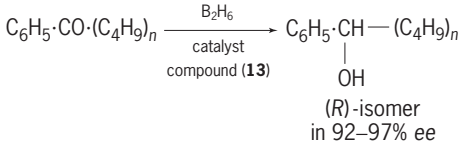

(3)

Reactions (2) and (3) have been carried out by using the polymer-supported catalysts in flow systems. In flow systems, the functional polymer beads are placed in a tube. Solutions of other reactants are caused to flow through the tube and to react under the influence of the polymer. Under optimum conditions, the solution leaving the tube contains just the chemical product, which is thus produced continually. *See* ASYMMETRIC SYNTHESIS; CATALYSIS; ORGANIC SYNTHESIS; STEREOCHEMISTRY.

Combinational synthesis. In combinatorial synthesis a short linear synthesis is carried out on polymer-supported substrates. This approach is being used in the pharmaceutical industry to identify lead compounds, which can thus be identified in weeks rather than years. Philip Hodge

Bibliography. A. Akelah and A. Moet, *Functionalised Polymers and Their Applications*, 1990; A. Kirsching (ed.), *Immobilized Catalysts: Solid Phases, Immobilization and Applications*, 2004; D. C. Sherrington and P. Hodge, *Syntheses and Separations Using Functional Polymers*, 1988.

Polymer thermodynamics

The relationships between properties of polymers, especially those affected by changes in temperature and the conversion of energy from one form to another. Like other materials, the bulk properties of polymers are determined by their chemical structure, but unlike many other classes of engineering materials, the range of properties that can be obtained is extremely wide. This is a consequence of the chemical and structural diversity of polymeric molecules.

Structure. Polymers are formed by linking together monomers (small molecules) to form chains of various "architectures," of which the linear chain is one example. Polymers formed from a single monomer species are called homopolymers, while those formed by the polymerization of more that one monomer species are called copolymers. In the copolymerization of two types of monomer species into a linear polymer, there are three key molecular features that may be varied independently: (1) the sequence in which the monomer units are joined, (2) the molecular weight distribution of the polymer, and (3) the tacticity (stereoregularity) of the polymer.

Possible sequences include block copolymers such as AAABBB, alternating copolymers such as ABAB, and random copolymers, such as ABBAABA, where A and B denote the two different monomers. The manner in which polymers are synthesized invariably leads to a distribution of molecular weights which are typically characterized by an average value and a quantity called the polydispersity (the width of the distribution). The tacticity of a polymer describes the orientation of functional groups attached to the backbone of the molecule. For example, the groups may be attached to one side only (isotactic); or they may alternate (syndiotactic); or they may be orientated randomly (atactic).

Phase behavior. Polymeric molecules typically have negligible vapor pressure and exist only in solid or liquid forms. However, the distinction between these phases is less clear than it is in typical substances composed of small molecules. Some polymers form a regular crystalline solid with a well-defined melting temperature T_m. However, irregularities of molecular size and structure typically lead to solid phases that are amorphous or, at most,

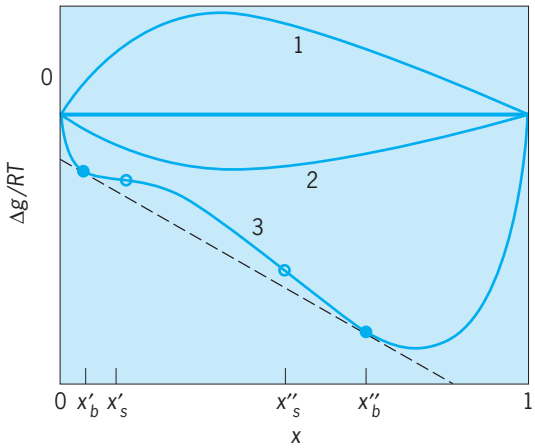

Fig. 1. Three possible forms of Δg for a binary-liquid mixture shown as a function of the mole fraction x of one component at constant temperature T and pressure p. (R = gas constant.) Solid curves show Δg for (1) a totally immiscible system, (2) a totally miscible system, and (3) a partially miscible system. On curve 3, filled circles identify the binodal points at mole fractions x_b' and x_b'', while open circles identify the spinodal points with mole fractions x_s' and x_s''. The broken line is the common tangent which connects the binodal points on curve 3.

semi-crystalline. Thermodynamic studies on such materials reveal a gradual transition from an amorphous glassy phase to a viscoelastic liquid as the temperature is raised above the glass-transition temperature T_g (the temperature above which the polymers becomes flexible). Above either T_m or T_g, the viscosity drops very rapidly with increasing temperature.

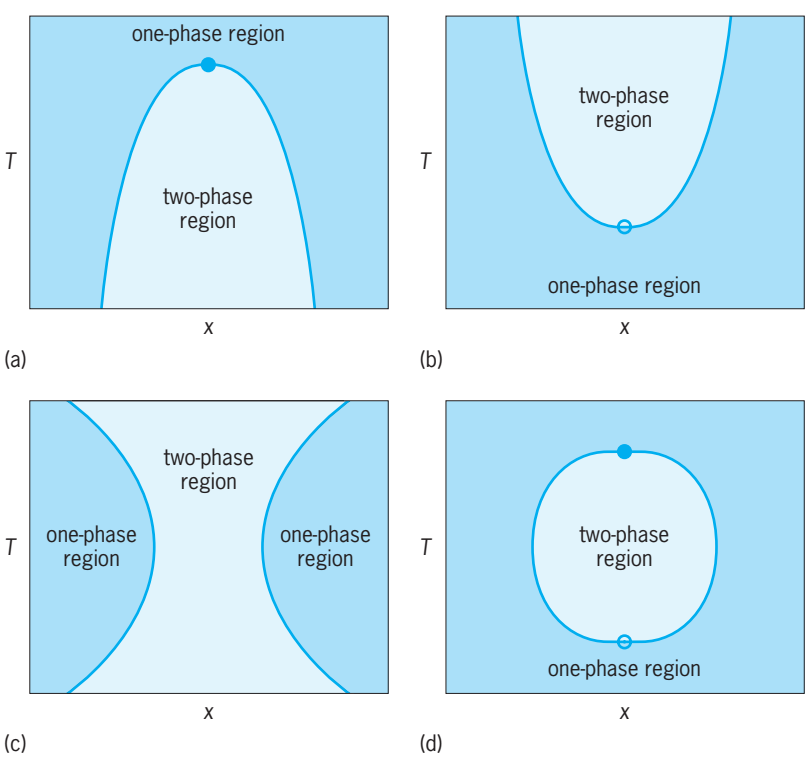

Fig. 2. Typical liquid-liquid miscibility diagrams for a binary mixture. (*a*) System with upper critical solution temperature (UCST). (*b*) System with lower critical solution temperature (LCST). (*c*) System without a critical solution point. (*d*) System with both a UCST and an LCST.

Miscibility of blends and solutions. Polymers are frequently blended in order to obtain materials with desired properties. In addition, polymers may be synthesized or processed in solution. In all these situations, the thermodynamics of mixing are of considerable importance. The most important issue is whether or not the components are miscible, and the key to understanding this is the behavior of the Gibbs free energy G of the system. For a system containing fixed amounts of the components under constant temperature T and pressure p, the fundamental thermodynamic criterion governing the number of phases present at equilibrium is that G should be minimized. Thus, a mixture held at constant T and p will split into different phases only if this leads to reduction in G. It follows that the phase behavior of a two-component (binary) mixture may be determined by examining the change of the molar Gibbs free energy upon mixing, Δg, as a function of the mole fraction x of one component. The fundamental criterion requires that, for stability as a single phase, $\Delta g(x)$ must be negative and have positive curvature with respect to the mole fraction x. Three possible scenarios are illustrated in **Fig. 1**. Curve 1 represents a system (for example, polymer A + polymer B or polymer + solvent) which is totally immiscible in all proportions: $\Delta g(x) \geq 0$ for all x. Curve 2 represents a system which is totally miscible in all proportions: $\Delta g(x) \leq 0$ and $(\partial^2 \Delta g/\partial x^2)_{T,p} \geq 0$ for all x. Curve 3 represents a system which is partially miscible, such that the two conditions for phase stability are satisfied only for $x \leq x_s'$ and $x \geq x_s''$, where x_s' and x_s'' are known as the spinodal compositions. Whenever this type of behavior is encountered, a miscibility gap exists within which a single homogeneous phase is not possible. Application of the fundamental criterion (that is, minimization of G) to this case shows that the miscibility gap spans the wider region $x_b'' \leq x \leq x_b'$ between the two binodal points. These are the points joined by a common tangent bounded from above by $\Delta g(x)$. A mixture prepared with overall composition such that $x_b'' \leq x \leq x_b'$ will, at equilibrium, exist as two separate phases with compositions x_b' and x_b''.

The molar Gibbs free energy change upon mixing is related to the corresponding changes Δh in molar enthalpy and Δs in molar entropy by the equation $\Delta g = \Delta h - T\Delta s$, where T is thermodynamic temperature. The entropy change is generally positive, and it does not vary greatly with molecular size. Thus the term containing Δs favors mixing, especially as the temperature is raised. The enthalpic term may be positive or negative, and its magnitude typically increases with molecular size so that, in systems containing polymers, Δh is important in determining the phase behavior. When considering how miscibility changes with temperature, a number of interesting possibilities arise (**Fig. 2**). As with mixtures of many ordinary liquids, the miscibility gap (that is, $x_b'' - x_b'$) may narrow with increasing temperature and eventually vanish at an upper critical solution temperature (UCST) above which the components are fully miscible (Fig. 2*a*). This is typical of systems in which

both Δb and Δs are positive and only weakly dependent upon T. Alternatively, the gap may narrow with decreasing temperature and vanish at a lower critical solution temperature (LCST; Fig. 2*b*). This situation is usually a consequence of specific interactions between the molecules (such as hydrogen bonding) which lead to negative values of Δb at low temperatures. In some cases, these two kinds of behavior combine, leading to a miscibility gap which first narrows with increasing temperature and then widens again without reaching a critical solution point (Fig. 2*c*). Occasionally, the binodal compositions define a closed loop within which the system is immiscible and one then has both a UCST and an LCST (Fig. 2*d*).

The phenomenon of partial miscibility is of course not unique to polymer blends and solutions. Many binary liquid mixtures exhibit a UCST if the system does not first freeze. The interesting features for polymer systems are that (1) partial miscibility is so common, and (2) both UCST and LCST behavior are frequently observed.

The actual process of dissolution in a polymer-solvent mixture is itself interesting. A lump of polymer immersed in the solvent usually first swells as the solvent is absorbed. Sometimes the mutual solubility is slight and the process stops there with a gellike polymer-rich phase, containing some solvent, coexisting with a solvent-rich phase which may contain only a vanishingly small concentration of polymer. In a miscible system, the polymer typically proceeds to dissolve after an initial period of swelling. Often the process is very slow, even with vigorous agitation.

Vapor-liquid equilibrium. Although polymers have negligible vapor pressure, situations of vapor-liquid equilibrium arise when polymers are dissolved in low-molecular-weight solvents. The question to be addressed in this case is the value of the saturated vapor pressure p_σ of the solvent above a polymer solution of specified concentration. This quantity varies smoothly from the vapor pressure of the pure solvent through to zero as the weight fraction of the polymer is increased. However, the relation is found to deviate greatly from the ideal-solution theory of elementary thermodynamics.

Thermodynamic models. Because of the huge number of possible blends and solutions of polymers, much of modern polymer thermodynamics is concerned with modeling the liquid-liquid and vapor-liquid equilibrium and other properties of such systems. The goal is to develop models that can predict the properties of mixtures containing polymers based on a minimum set of experimental data. Thermodynamic models fall into two broad categories: excess Gibbs free energy models and equation of state theories. The former seek to relate Δg to temperature, pressure, and composition and may be used to predict phase equilibrium, whereas the latter express the system pressure in terms of temperature, density, and composition. Equation-of-state theories also lead to Δg (and hence to predictions of phase behavior). They also permit other thermodynamic properties, such as compressibility and thermal expansivity, to be predicted. *See* COPOLYMER; ENTHALPY; FREE ENERGY; GLASS TRANSITION; POLYMER; POLYMERIZATION; THERMODYNAMIC PRINCIPLES.

J. P. Martin Trusler

Bibliography. F. W. Billmeyer, Jr., *Textbook of Polymer Science*, 3d ed., Wiley, Chichester, 1984; H.-G. Elias, *An Introduction to Polymer Science*, VCH, New York, 1997; J. I. Kroschwitz, *Concise Encyclopedia of Polymer Science and Engineering*, Wiley-Interscience, New York, 1990; F. Rodriguez, *Principles of Polymer Systems*, 4th ed., Taylor & Francis, Washington DC, 1996; S. I. Sandler, Models for thermodynamic and phase equilibria calculations, Dekker, New York, 1994.

Polymerization

The repetitive linking of small molecules (monomers) to make larger molecules (polymers). In general, polymers are named after the small molecules from which they are made. For example, the gaseous monomer ethylene can be polymerized to make solid polyethylene. A polymer molecule may contain hundreds of thousands repeat units, or monomers. Polymerization requires that each small molecule have at least two reaction points or functional groups. The polymerization of a monomer usually involves many steps. In each step, new bonds are formed, resulting in the extension and growth of the respective chain. Synthetic polymers are usually polydisperse; that is, in any given sample, the many polymer chains present will have varying lengths, while being of the same chemical composition. Consequently, the value describing a molecular weight or size of a polymer sample always refers to an average, such as a weight average or number average. Size and architecture of polymer molecules are of greatest importance to the polymer chemist, as they determine the ultimate material properties, and hence the necessary polymerization process.

Polymer Structures/Architectures

Depending on the particular chemistry of the polymerization process, the monomers can be combined in a variety of ways, resulting in polymers of different shapes and sizes. Possible shapes include linear, branched, ladderlike, star-shaped, arborescent, and cross-linked (**Fig. 1**).

Two samples of the same polymer, but with different architectures, will have distinctly different material properties, such as crystallinity, solubility, and mechanical strength, as well as the ability to swell, melt, or deform.

Homopolymers and copolymers. If the polymer synthesis involves only one monomer (except polycondensates from combinations of a-a/b-b type monomers; see below), the resulting polymer chains will feature only one structural repeat unit, and the polymer is called a homopolymer (**Fig. 2**).

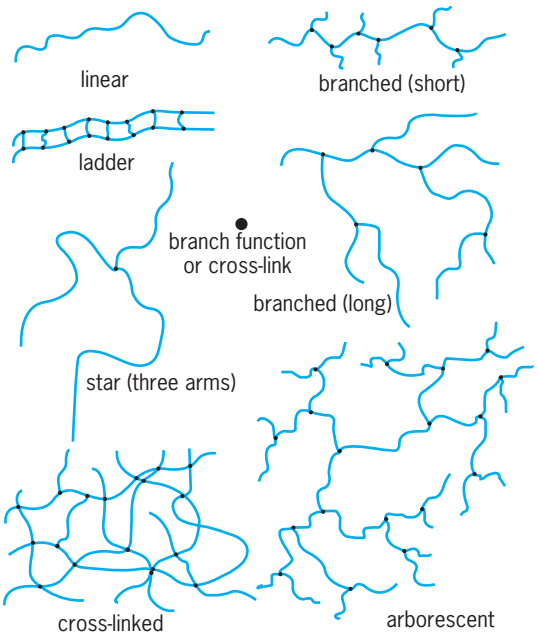

Fig. 1. Various polymer shapes/architectures.

In contrast, copolymers are derived from two or more monomers, resulting in two or more structural repeat units within the copolymer molecule (Fig. 2). Several types of copolymers are distinguished, depending on the sequential placement of the monomers, as well as the molecular architecture, for example, linear or branched.

In the case of copolymers with two different mers (repeat structures), A and B, the exact relative placement of A and B can be controlled by choosing the appropriate polymerization process, resulting in very different material properties. In the case of the random copolymer, A and B are distributed statistically along the chain. In alternating copolymers, A and B alternate along the chain, thus making it possible to view this copolymer as a homopolymer made of AB mers.

Block copolymers can be regarded as two block sequences of homopolymer, poly A and poly B, which are attached to each other. The properties of the resulting materials are a combination of the properties of the individual homopolymers, which offers a vast parameter space for macromolecular design of target materials. The number of block sequences in one chain is usually two (diblock) or three (triblock); the individual sequence length, however, can vary widely from several to several thousand mers. The synthesis of block copolymers has become increasingly important in recent years. Graft copolymers feature a homopolymer backbone poly A, laterally to which homopolymer block sequences of poly B are grafted. *See* COPOLYMER.

Orientation of repeat units. In a homopolymer or a homopolymer block sequence, the monomer molecules usually have two different sides in regard to their orientation within the polymer chain: head and tail. Depending on which side of the monomer adds to another, there are three general coupling modes, head-to-tail, head-to-head, or tail-to-tail (**Fig. 3**).

Usually, the placement is head-to-tail. However, depending on the monomer structure and reactivity, other possibilities may have to be taken into consideration during the polymer synthesis. In some instances, process conditions may be chosen that favor head-to-head or tail-to-tail coupling, should this be desirable with respect to the target material.

Another variable is that every R-substituted carbon in a vinyl monomer is a potential center of

Fig. 2. Structures of homopolymers and copolymers.

asymmetry, with two possible lateral orientations of R in respect to the polymers chain, resulting in a wide range of possible material characteristics (**Fig. 4**).

If the polymerization process offers no control during the monomer addition, the placement will be statistical, resulting in an atactic polymer. The reaction conditions, for example, using an appropriate catalyst system, will allow for precise control over the placement of the R substituents. If all remain on the same side of the chain, the polymer is called isotactic. If the placement alternates the sides of the chain, the polymer is called syndiotactic.

Polymer size. The size of a polymer molecule is commonly referred to as molecular weight and given in terms of g/mol. Unlike organic molecules with a defined structure and molecular weight, polymers of the identical chemical composition can vary greatly in size, depending on how many mers constitute one polymer molecule. The general influence of molecular weight on the mechanical strength is shown in **Fig. 5**.

Even a sample of a chemically pure polymer is composed of chains of various sizes. Molecular-weight characterization is an integral part of any polymerization process. Possible analytical approaches include using the colligative properties, viscosity, and light-scattering characteristics of polymer solutions. One has to distinguish between methods relative to known calibration standards and direct measurements. Methods based on the colligative properties enable the determination of the number-average molecular weight $\overline{M}_n$, essentially determining the number of polymer molecules in a given sample. Viscosity measurements yield the viscosity-average molecular weight $\overline{M}_v$, depending on the hydrodynamic volume of the polymer molecule. Modern developments in light-scattering instruments have enabled great advances in the direct determination of the weight-average molecular weight $\overline{M}_w$ on a routine basis, which are becoming more independent of measurements relative to known calibration standards. The molecular-weight uniformity of a sample, or polydispersity, is usually characterized by the polydispersity index,

$$\mathrm{PDI} = \frac{\overline{M}_w}{\overline{M}_n}$$

For comparison, the polydispersity of a small organic molecule, such as styrene, would have a value of PDI equal to unity, meaning the sample is monodisperse. Polymer PDI values may approach 1 in special cases of controlled/living polymerization, but usually are significantly larger than 1. *See* LIGHT-SCATTERING TECHNIQUES; MOLECULAR WEIGHT; VISCOSITY.

Polymer Synthesis

There are two distinct major types of polymerization processes: condensation polymerization, in which the chain growth is accompanied by elimination of small molecules such as water (H_2O) or methanol (CH_3OH), and addition polymerization, in which the polymer is formed without the loss of other materials.

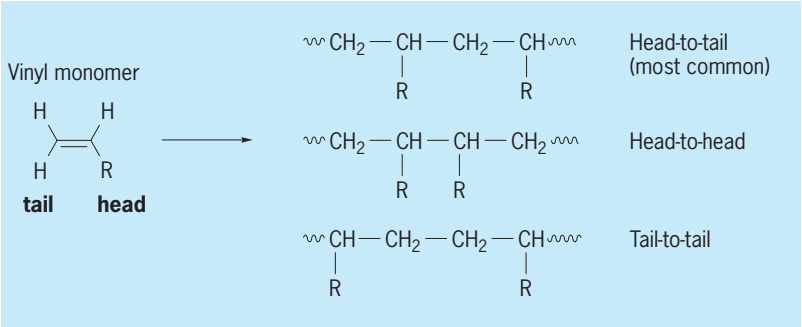

Fig. 3. Three general coupling modes of monomers.

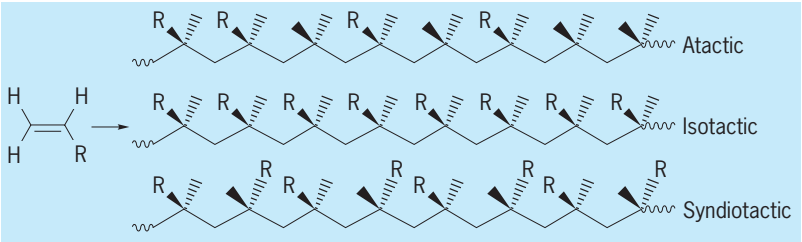

Fig. 4. Placement of the R substituents.

Condensation (step-growth) polymerization. The formation of macromolecules by the condensation process, as in the production of polyamides, polyesters, and polysulfides, requires the elimination of small molecules; at the same time, strongly polar and strongly attracting groups are produced. Most condensations proceed in a stepwise fashion. Common types of condensation polymers are polyamide, polyester, polyurethane, polysiloxane, phenol-formaldehyde, urea-formaldehyde, melamine-formaldehyde, polysulfide, polyacetal, protein, wool, and silk. *See* PHENOLIC RESIN; POLYACETAL; POLYAMIDE RESINS; POLYESTER RESINS; POLYSULFIDE RESINS; POLYURETHANE RESINS; PROTEIN; SILICONE RESINS; SILK; WOOL.

An example of the condensation process is reaction (1) of ε-aminocaproic acid in the presence of a catalyst to form the polyamide, nylon-6. The repeating structural unit is equivalent to the starting material minus H and OH, the elements of water. A similar product would be obtained by reaction (2) of a diamine and a dicarboxylic acid. In both cases, the

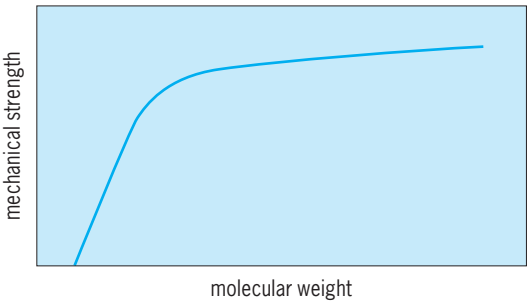

Fig. 5. Dependency of mechanical strength on molecular weight.

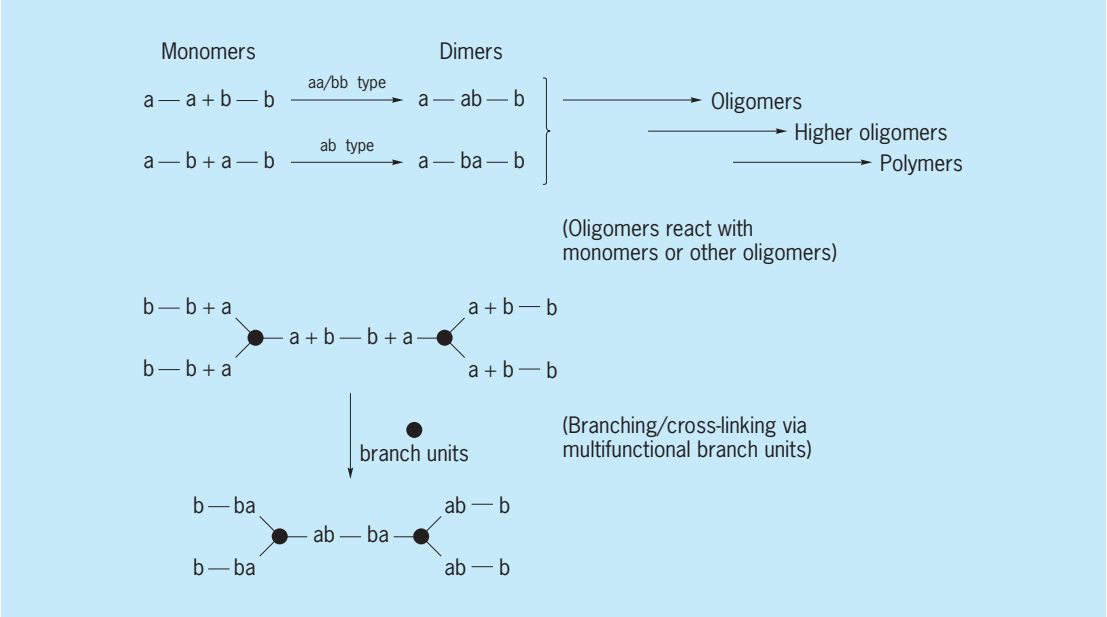

Fig. 6. General condensation behavior.

molecules formed are linear because the total functionality of the reaction system (functional groups per molecule) is always two (Fig. 1).

If a trifunctional material, such as a tricarboxylic acid, were added to the nylon-6,6 polymerizing mixture, a branched polymeric structure would result, because two of the carboxylic groups would participate in one polymer chain, and the third carboxylic group would start the growth of another (Fig. 1). Under appropriate conditions, these chains can become bridges between linear chains, and the polymer is cross-linked (Fig. 1).

This general condensation behavior is sketched in **Fig. 6**, where "a" could stand for a carboxylic acid function and "b" could correspond to the amine function discussed above.

The scheme also makes clear that all monomers, oligomers, and polymers in the reaction mixture possess equivalent chain end functionality, and as a consequence react with each other in a statistical fashion. This is in stark contrast to the mechanism of chain-growth polymerizations discussed below.

In the equimolar, stepwise reaction of a dicarboxylic acid a-a with a diamine b-b as an example, after 50% of the groups have reacted, the average degree of polymerization (degree of polymerization) of the polymer formed is 2; after 90% reaction, the degree of polymerization is 10; and after 99.5% reaction, the degree of polymerization is 200. If the molecular weight of the repeating unit is 100, then the average molecular weight of the polymer is 20,000. Polymers of considerable molecular weight thus are produced only at very high monomer conversions (**Fig. 7**). Furthermore, in order for the reaction to approach 100% completion and thus to

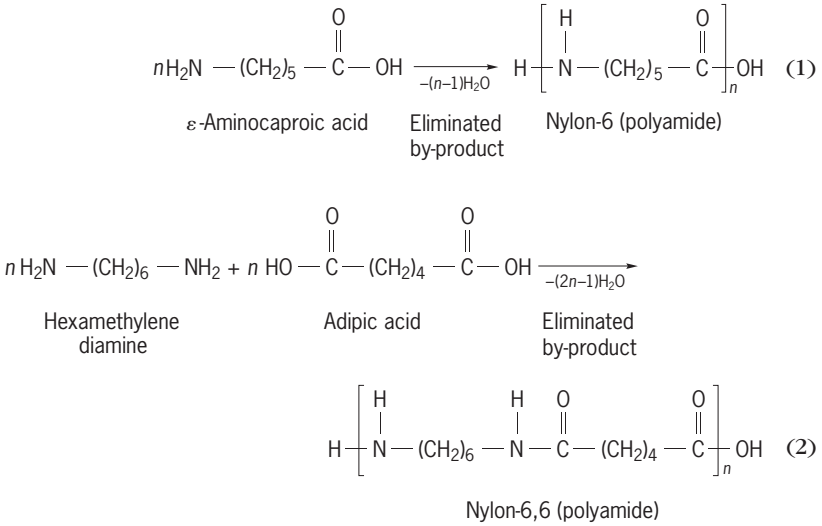

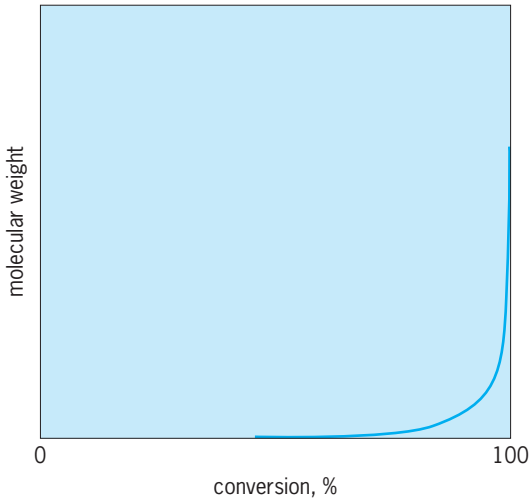

Fig. 7. Dependency of molecular weight on conversion (polycondensation).

obtain a high-molecular-weight product, the small molecules formed during the condensation must be removed from the reaction mixture.

In practice, condensations are usually started under moderate conditions of temperature and pressure and completed at high temperature (removal of by-product) and low pressure to yield linear products in the molecular-weight range of about 5000–30,000. In certain cases, interfacial polymerization is useful. If reactants can be dissolved in immiscible solvents, polymerization may take place at low temperatures and resources. The linear products, thermoplastic condensation resins, are used in fibers, films, coatings, molding compounds, and adhesives.

Useful condensation polymers that are highly cross-linked are prepared from low-molecular-weight polyfunctional reaction systems. The condensations of formaldehyde with phenol and with urea, and other condensations, are discussed in the Encyclopedia articles on specific products. The reaction of formaldehyde with urea is shown in (3).

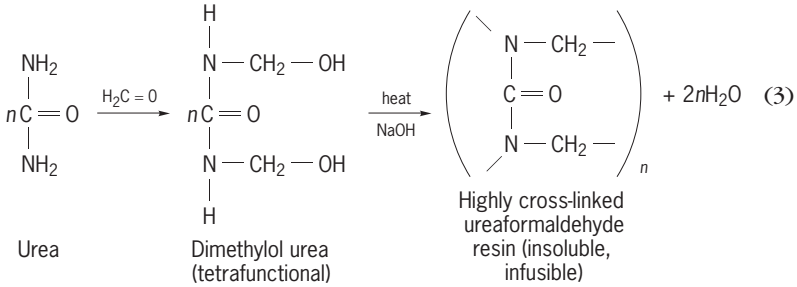

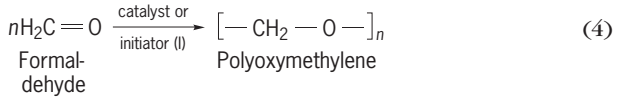

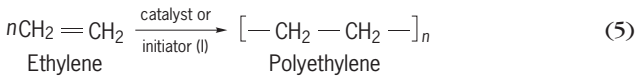

In the intermediate compound, dimethylol urea is tetrafunctional. The H— and HO— groups on one molecule react with the HO— and H— groups on other molecules to form water. Each urea unit finally becomes bound to other urea units via four methylene (—CH_2—) bridges. Because the final product is cross-linked and infusible, the final shaping operation usually must coincide with the final curing or cross-linking. In practice, the soluble, low-molecular-weight intermediate condensate can be isolated and mixed with strengthening fillers, coloring pigments, and curing catalysts to yield a molding powder. By subjecting the urea-formaldehyde molding powder to heat and pressure in a mold, the curing reaction takes place, some of the water is driven off as steam, and some is adsorbed by the filler. The molded object, such as a piece of dinnerware, is now insoluble and infusible. Because all the molecules are joined together, the molecular weight of a highly cross-linked polymer is a meaningless term since its value is effectively infinite.

The end groups of the polymer molecules are the functional groups that have not reacted at any stage. It is apparent that at exactly 100% conversion in a difunctional condensation, the reaction system would consist of only one molecule. The fact that the end groups of a condensation polymer can always undergo further reaction creates a difficulty in the high-temperature-melt spinning of polyamides and polyesters. To prevent subsequent changes in molecular weight, such as might occur in the melt during spinning at elevated temperatures, a monoacid or monoalcohol (molecular-weight modifier) is added to the original polymerization mixture. The excess of hydroxyl groups, for example, places a limit on the chain growth. That limit is reached when all the acid groups have reacted and all the end groups are hydroxyl.

Some examples of addition polymerization are reactions (4) and (5). The structure of the repeating unit is the difunctional monomeric unit, or mer. In the presence of catalysts or initiators, the monomer yields a polymer by the joining together of n mers. If n is a small number, 2–10, the products are dimers, trimers, tetramers, or oligomers, and the materials are usually gases, liquids, oils, or brittle solids. In most solid polymers, n has values ranging from a few score to several hundred thousand, and the corresponding molecular weights range from a few thousand to several million. The end groups of these two examples of addition polymers are shown to be fragments of the initiator.

Addition (chain-growth) polymerization. Unsaturated compounds such as olefins and dienes polymerize without the elimination of other products. Certain ring structures, for example, lactams or alkylene oxides, may polymerize by an opening of the ring. The molecular weight and structure of the polymer are determined by the reaction conditions, that is, the nature of the catalyst or initiator, the temperature, and the concentration of reactants, monomer, initiator, and modifying agents. In general, a ceiling temperature exists above which polymerization cannot occur. Most addition polymerizations proceed by a chain reaction in which the average chain length of the polymer formed initially is high and may increase further through secondary branching reactions as the polymerization approaches completion.

The molecular-weight range for many useful addition polymers is relatively high, typically from 20,000 to several million, as compared with the molecular weight range of 5000–30,000 for typical condensation polymers.

The types of catalysis or initiation which are effective for addition (chain-growth) polymerization fall into four groups: (1) free-radical catalysis by peroxides, persulfates, azo compounds, oxygen, and ultraviolet and other radiation; (2) acid (cationic) catalysis by the Lewis acids, such as boron trifluoride, sulfuric acid, aluminum chloride, and other Friedel-Crafts agents; (3) basic (anionic) catalysis, by chromic oxide on silica-alumina, nickel, or cobalt on carbon black, molybdenum on alumina, and complexes of aluminum alkyls with titanium chloride. (4) The precise mechanisms of the fourth group are

difficult to elucidate. However, evidence suggests that, at least in some cases, the mechanism involves an anionic process modified by coordination of the monomer and polymer with a surface.

It is convenient to discuss the mechanism and experimental methods of free-radical initiation as one subject, and to treat the remaining three types under the heading complex or ionic catalysts. *See* CATALYSIS.

Free-radical catalysis. Among the several kinds of polymerization catalysis, free-radical initiation has been most studied and is most widely employed. Atactic polymers are readily formed by free-radical polymerization, at moderate temperatures, of vinyl and diene monomers and some of their derivatives.

At an appropriate temperature, a peroxide decomposes to yield free radicals. In the presence of a monomer, the greater proportion of these radicals adds to the monomer and thereby initiates chain growth. The growing chains may terminate by coupling, by disproportionation, or by transfer to monomer, polymer, or added materials (transfer agents, retarders, and inhibitors). Reactions (6)–(10) illustrate the initiation, propagation, and termination by coupling in vinyl acetate.

If transfer occurs to the unreacted monomer or polymer already formed, higher-molecular-weight branched structures will be produced; and if branching is excessive, insoluble products may be formed. If the radical produced in the transfer process is not sufficiently active to initiate a new chain, the transfer agent is called an inhibitor or a retarder. Mercaptans (RSH), carbon tetrachloride, and various organic solvents are examples of transfer agents, whereas amines and phenols are often used as inhibitors or retarders.

The rate of chemically initiated free-radical polymerization is increased by raising the temperature or increasing the concentration of monomer and initiator, whereas the molecular weight of the polymer is increased by increasing the monomer concentration, by lowering the temperature, and by lowering the concentration of initiator and transfer agents. *See* FREE RADICAL.

Complex or ionic catalysis. Some polymerizations can be initiated by materials, often called ionic catalysts,

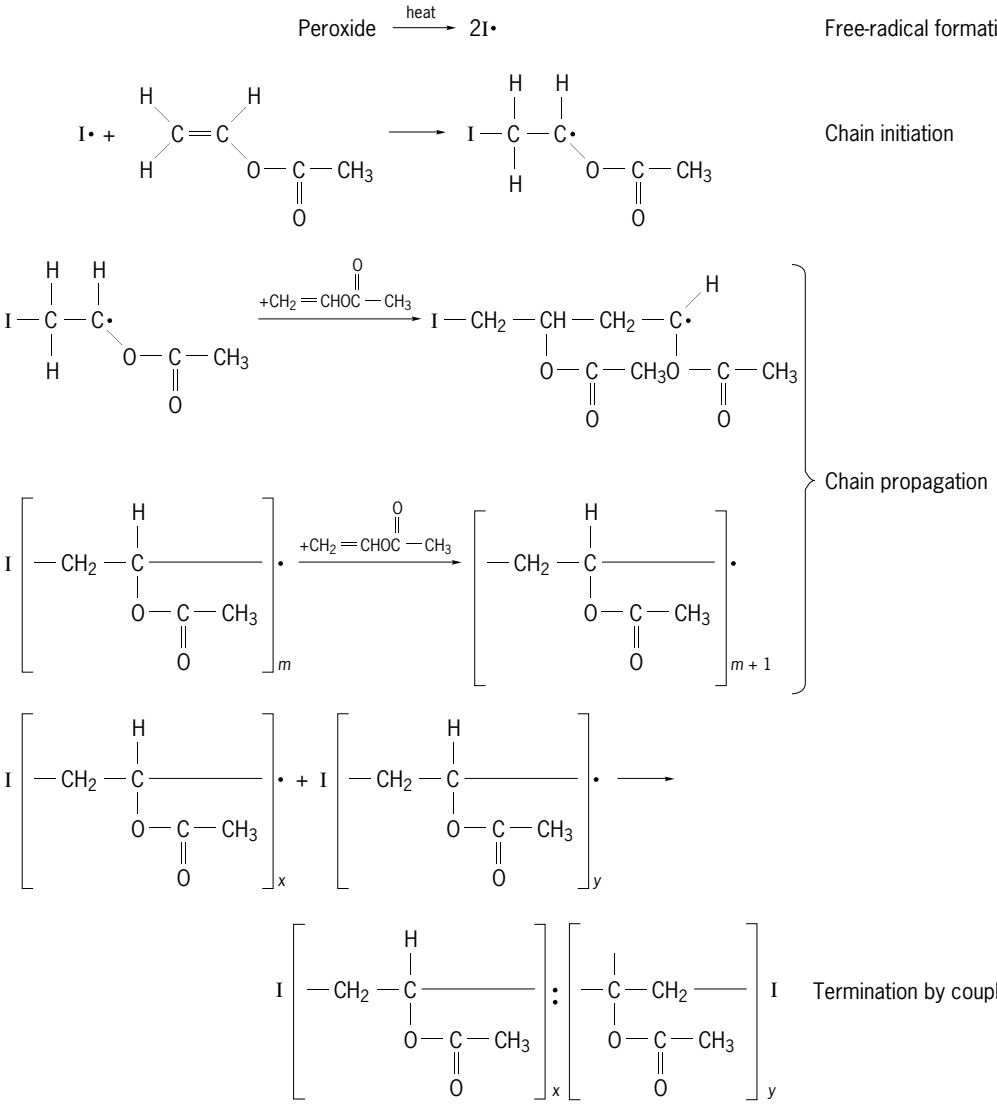

that contain highly polar reactive sites or complexes. The term "heterogeneous catalyst" is often applicable to these materials because many of the catalyst systems are insoluble in monomers and other solvents. These polymerizations are usually carried out in solution from which the polymer can be obtained by evaporation of the solvent or by precipitation on the addition of a nonsolvent.

A general mechanism is shown in reactions (11),

$$\begin{aligned}
&\text{Initiative complex:}\\
&\quad \text{Complex catalyst} + M \rightarrow CM^*\\
&\text{Initiation:}\\
&\quad CM^* + M \rightarrow M^*CM\\
&\text{Termination by decomposition of complex:}\\
&\quad M_x^*CM \rightarrow M_x + CM\\
&\text{Termination by transfer to monomer:}\\
&\quad M_x^*CM + M \rightarrow M_x + M^*CM
\end{aligned} \tag{11}$$

in which the growing chain is represented as an activated complex with the complex catalyst, without attempting to specify whether separate ions or free radicals are involved.

A distinguishing feature of complex catalysts is the ability of some representatives of each type to initiate stereoregular polymerization at ordinary temperatures or to cause the formation of polymers that can be crystallized. The polymerization process may often be visualized as the formation of an activated complex of the monomer with the complex catalyst. For stereoregular growth to take place, the entering monomer must collide with the complex and form a new, transient complex in which the new monomer molecule is held in a particular orientation. As reaction takes place, the new monomer assumes an activated condition within the complex catalyst and, at the same time, pushes the old monomer unit out. Chain growth is therefore similar to the growth of a hair from the skin. If conditions favor certain orientations of the old and new monomer units, a stereoregular polymer results. *See* SINGLE-SITE CATALYSTS (POLYMER).

The effect of conditions on rates of polymerization and on molecular size and structure is not yet fully understood. In general, the rate of polymerization is proportional to the concentrations of complex catalyst and monomer. The effect of temperature on the rate depends upon the stability and activity of the complex catalyst at the temperature under consideration. If the complex catalyst decomposes on increasing the temperature, the rate of polymerization will be reduced. The effect of temperature upon molecular weight also depends upon the stability of the complex catalyst and upon the relative rates of propagation and termination. In some cases, at an optimum temperature of polymerization, the molecular weight depends upon the product of the ratio of the rate of the propagation to termination and the monomer concentration, and in other cases, only upon that ratio of rates.

Examples of polymerization with the different types of complex catalysts are briefly described below.

1. *Lewis acids*. Carbonium-ion or cationic catalysts such as BF_3, $AlCl_3$, or H_2SO_4 usually require the presence of a promoter such as H_2O or HCl, as in reaction (12). The point of chain growth is a carbo-

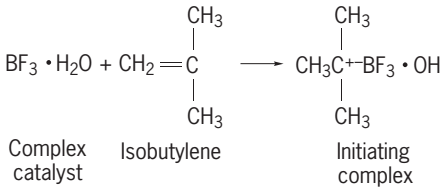

(12)

nium (positive) ion, which is associated with a negative counterion. Polymerization in the presence of Lewis acids takes place very rapidly at low temperatures, −100 to 0°C (−150° to +32°F). The order of reactivity of some olefins in Lewis acid catalysis is vinyl ethers > isobutylene > α-methyl styrene > isoprene > styrene > butadiene.

2. *Lewis bases*. Carbanions of anionic catalysts such as sodium, lithium, and lithium butyl function at moderate to low temperatures, −70 to +150°C (−94 to +300°F). Inert hydrocarbon or ether solvents are generally used as reaction media. The reaction is shown as (13).

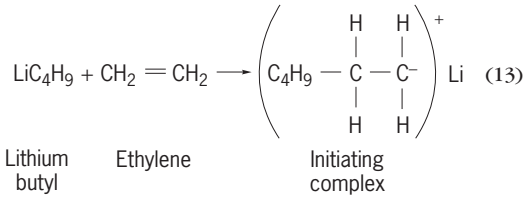

(13)

It has been suggested that the point of chain growth is a negative or anionic center, although not usually dissociated from its positive counterion. The order of reactivity of some monomers in carbanion or anionic polymerization is acrylonitrile > methacrylonitrile > methyl methacrylate > styrene > butadiene. *See* HOMOGENEOUS CATALYSIS; REACTIVE INTERMEDIATES.

3. *Heterogeneous catalysts*. Certain heavy metals or metal oxides on supports and complexes of aluminum alkyls with titanium chloride function as catalysts at moderate to high temperatures, 50–220°C (120–430°F). Inert hydrocarbon or ether solvents are generally used as reaction media. The catalysts may be used in a fixed bed or as a slurry. Two examples are shown as reactions (14) and (15).

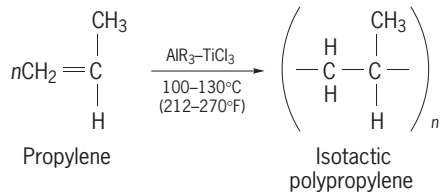

(14)

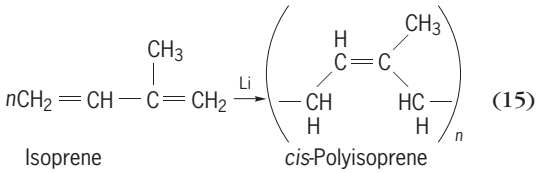

(15)

See HETEROGENEOUS CATALYSIS.

Polymerization processes. There are a number of processes for achieving polymerization, designed for various conditions and end products.

Bulk process. This consists of polymerization of the pure monomer in liquid form. On initiation by heat or light or very small amounts of azobisisobutyronitrile, a very pure polymer can be formed. The monomer and polymer are poor heat conductors; therefore the temperature of bulk polymerization is difficult to control. A further disadvantage is that small quantities of unreacted monomer are difficult to remove from the polymer. Polymerization in solution offers a means of carrying out the polymerization at lower monomer concentrations. Because solvents frequently act as transfer agents, polymerization in solution generally leads to the formation of lower-molecular-weight products.

Aqueous emulsion. This type of polymerization has the advantages of giving a high rate of polymerization, a high molecular weight, and ease of temperature control. A liquid monomer is emulsified in water by use of a surface-active agent, such as soap. The soap micelles provide the polymerization centers. The free radicals (from a water-soluble initiator or growing chains of low molecular weight) diffuse into the soap micelle, in which they react to form a relatively linear polymer of high molecular weight. The polymer particles of small diameter, 50–150 nanometers, are in stable suspension because the soap of the original micelle remains adsorbed in the outer layer of the polymer particle. The rate of emulsion polymerization and the molecular weight of the polymer increase with increasing numbers of micelle particles per unit volume. The product, a stable colloidal suspension of the polymer in water, usually is called a polymer latex or polymer emulsion. Polymer latexes are used directly for water-based paints, for adhesives, and for treating textiles. When polar solvents or electrolytes are added to the colloidal suspension, the polymer coagulates, and it can be separated and dried. To produce polymer emulsions which have the desired mechanical and thermal stability, it is frequently necessary to use moderately high concentrations of surface-active agents and protective colloids. Therefore, emulsion polymers are generally less pure than bulk ones. *See* EMULSION POLYMERIZATION.

Redox initiation. This system was developed for polymerization in aqueous emulsions. In the presence of a water-soluble reducing agent such as sodium bisulfite or ferrous sulfate, the peroxide decomposes more rapidly at a given temperature, and consequently polymerization at useful rates can take place at lower temperatures. By the use of the redox system, the temperature for the commercial emulsion polymerization of styrene and butadiene is lowered from 50 to 5°C (120 to 40°F) to form "cold" rubber, which is higher in molecular weight than the "50°" or "hot" synthetic rubber.

Suspension polymerization. This system offers several advantages. By the use of a very small amount of surface-active materials and mechanical aggregation, the monomer can be dispersed as droplets in the water. The monomer is not colloidally dispersed, but is temporarily broken up into droplets which would coagulate if the stirring were discontinued. In the presence of a peroxide which is soluble in the monomer, bulk polymerization takes place in the droplets. When the polymerization reaches some 15–40%, the droplets containing the dissolved polymer become sticky and may coalesce. Various agents, such as talc and metallic oxides, have been recommended for use in very small amounts to prevent coagulation in the sticky stage. At higher degrees of conversion, the droplets will be transformed to hard balls of polymer containing dissolved monomer. At the completion of polymerization, the balls or beads settle out. These may be dried easily and are ready for use. In the suspension system, the possibility of producing a pure polymer in bulk polymerization is combined with the ease of temperature control in the aqueous emulsion polymerization. There is the additional advantage that the product of suspension polymerization can be easily isolated for use.

Ring opening. In this type, some cyclic monomers can undergo ring opening to yield high polymers; the reaction may proceed by a chain- or step-growth mechanism. Commercially important products are the polymerization of caprolactam, caprolactone, and ethylene oxide. *See* ACID AND BASE; ALKENE; CHAIN REACTION (CHEMISTRY); CHEMICAL DYNAMICS; INHIBITOR (CHEMISTRY); MANUFACTURED FIBER; ORGANIC REACTION MECHANISM; OXIDATION-REDUCTION; POLYMER; RING-OPENING POLYMERIZATION.

Ralf M. Peetz; John A. Manson

Polymerizations in Supercritical Carbon Dioxide

The use of supercritical carbon dioxide as a solvent or dispersing medium for the production of polymers has the potential for offsetting the environmental concerns associated with conventional solvents. A supercritical fluid is typically a gas that above a certain temperature (referred to as the critical temperature; 88.3°F or 31.3°C for carbon dioxide) cannot be liquefied by an increase in pressure. The supercritical state exhibits properties intermediate between those of a gas and a liquid, and density-dependent properties such as solvent dielectric constant and viscosity can be tailored by simple changes in pressure or temperature. The unique properties of supercritical fluids (solvating ability similar to conventional liquids coupled with high diffusion rates and gaslike viscosities) offer many potential advantages in terms of reaction control. In addition, there are obvious environmental advantages of any technology with the potential to dramatically reduce waste streams. Supercritical carbon dioxide is compatible with a variety of chemistries, and is an attractive solvent

alternative in free-radical solution, precipitation, and dispersion processes; in cationic polymerizations; and for some metal-catalyzed polymerizations. A particularly attractive aspect of polymerizations in supercritical carbon dioxide involves the product-isolation step. Since carbon dioxide is a gas at standard pressure and temperature, simply venting carbon dioxide from the reaction vessel at the end of the polymerization leaves behind the solid or liquid polymer product. In addition, the employment of supercritical fluid extraction processes provides the potential for removal of residual reactants and other contaminants from the polymer when the presence of such impurities is undesirable.

Carbon dioxide can be employed in the solution polymerization of amorphous fluoropolymers and siloxanes. For solution processes, the monomer and initiator employed, as well as the resulting polymer, are soluble in carbon dioxide. After the polymerization is complete, carbon dioxide is vented, and the polymer remains in the reactor. Two key examples are the solution polymerization of 1,1′-dihydroperfluorooctyl acrylate (FOA), initiated by a free-radical initiator [reaction (16)], and the free-radical telomerization of tetrafluoroethylene [reaction (17)].

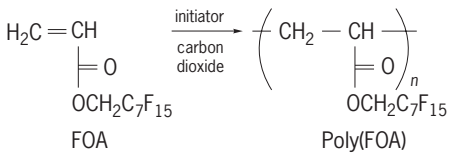

(16)

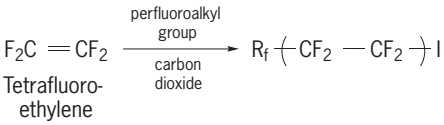

(17)

Amorphous fluoropolymers and siloxanes are the exception rather than the rule when considering polymer solubilities in supercritical carbon dioxide. Most polymers are insoluble, precluding the use of solution processes; thus application of heterogeneous methodologies such as precipitation and dispersion polymerization processes is appropriate.

Precipitation polymerization processes in supercritical carbon dioxide are useful for the preparation of many polymers. They have been employed most successfully in the preparation of water-soluble polymers such as poly(acrylic acid), which is widely used in preparation of superabsorbent materials. In this type of process, monomer and initiator are soluble in carbon dioxide, and the resulting polymer is not. Consequently, the polymer precipitates as it is formed. Venting the reaction vessel after polymerization is complete enables the isolation of dry free-flowing powders. For applications where even low levels of impurity are unacceptable, supercritical fluid extraction techniques may be applied to remove residual monomer, initiator, and other impurities. Carbon dioxide shows particular promise as a replacement for the toxic, environmentally detrimental solvents, such as benzene and methylene chloride, that have been used in the precipitation polymerization of acrylic acid.

An important development in this area involves dispersion polymerizations in supercritical carbon dioxide. In a dispersion polymerization, the monomer and initiator are soluble in carbon dioxide and the resulting polymer is insoluble, as in precipitation polymerization. However, in dispersion polymerization, specially designed surface-active agents (surfactants) stabilize the precipitated polymer as a colloid. This results in uniform spherical polymer particles similar to those achieved in conventional emulsion polymerizations. However, dispersion polymerizations conducted in supercritical carbon dioxide offer many unique opportunities resulting from carbon dioxide's ability to plasticize the growing polymer particles, thereby allowing for the formation of interpenetrating networks, core-shell materials, drug delivery systems, and so forth. The ease with which carbon dioxide can be vented provides a simple workup and isolation of the resulting polymer. In addition, carbon dioxide may find application as a propellant for aerosol coatings; thus a dispersion of polymer in carbon dioxide could be used directly in spray coating and similar applications. For applications where high purity is necessary or recovery of stabilizer is important, supercritical fluid extraction techniques may also prove useful. *See* CARBON DIOXIDE; COLLOID; CRITICAL PHENOMENA; SURFACTANT.

T. J. Romack; J. M. DeSimone

Living/Controlled Polymerizations (Block Copolymers)

As described, chain-growth polymerizations originate from an initiator fragment and involve the sequential addition of new monomer molecules to the growing chain end. The growth of any particular chain is stopped by transfer or termination reactions. In recent years, the toolbox of macromolecular engineering has grown to include a variety of polymerization chemistries that address the transfer and termination problem by either suppression or prevention, or by rendering them reversible. The result can be a polymer chain that stops growing when the monomer is depleted in the reaction mixture, but starts growing again when new monomer is added. Hence, the molecular weight $\overline{M}_n$ scales linearly with conversion (**Fig. 8**).

In such a living/controlled polymerization, a defined number of initiator fragments start a defined number of polymer chains, all of which ideally grow at the same rate, while no new chains are started. This results in a very uniform molecular weight distribution in the sample and a low polydispersity, often with PDI < 1.1. This enables the synthesis of polymers with defined molecular weights and high uniformity. Because any impurity in the reaction mixture has the potential to act as a transfer or termination agent, or an initiator site, particular attention has to be paid to the purity of the monomers and solvents.

Even after isolation of the polymer, the chain ends often remain functional for either a next

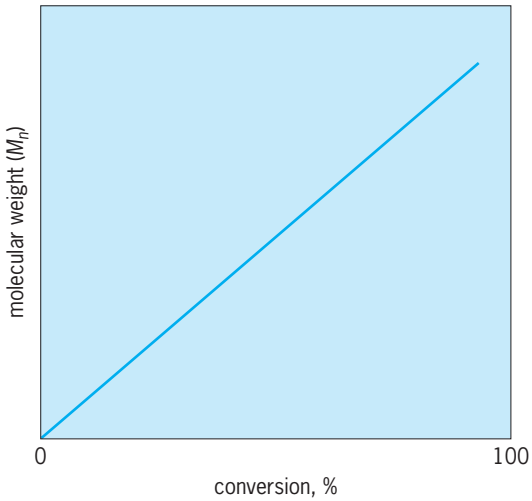

Fig. 8. Living/controlled polymerization, showing molecular weight as a function of monomer conversion.

polymerization or appropriate functionalization reactions. In the case of the living cationic polymerization of isobutylene, the initiator fragment may have two, three, or more equivalent functional groups that are able to start a growing chain. In the given example, starting from tricumyl chloride, the living cationic polymerization of isobutylene progresses in the presence of the Lewis acid $TiCl_4$ at a temperature of −80 to −60°C. Termination and transfer reaction are prevented by additives, such as *N,N*-dimethylacetamide and 2,6-di-*tert*-butylpyridine (**Fig. 9**).

After reaction completion, the result is a star-shaped polyisobutylene polymer with three functional ends, also called a tritelechelic polymer. As the functional groups at the chain end remain unchanged, they can be converted into a wide variety of alternative functional groups by conventional organic chemistry, depending on the end use of the telechelic polymer. This general principle has enabled the synthesis of polymers with precisely controlled molecular weights and various chain end functionalities. The polymers can be prepared and used as macromolecular building blocks in various applications later. In (Fig. 9), the resulting end-functional polyisobutylene polymer is subjected to another polymerization, in which the functional chain ends serve as initiating sites for the living polymerization of styrene at temperatures below −50°C. The overall result is a star-shaped block copolymer, in which three block copolymer arms emanate from the trifunctional core, with polyisobutylene on the inside of the star and polystyrene on the outside of the star, the corona. Such a combination of the rubbery polyisobutylene and the glassy polystyrene demonstrates the scope of this synthetic approach to generate a variety of high-tech materials, such as for biotechnological applications. The example shows the combination of a rubbery with a glassy polymer, the two of which are incompatible and phase separate in bulk, resulting in a soluble and fusible material called a thermoplastic elastomer (TPE). Thermoplastic elastomers are physically cross-linked rubbers, in which the cross-link can be reversed by dissolving the polymer in an appropriate solvent for both blocks, or heating above the glass transition temperature (T_g) of the glassy polymer. *See* RUBBER.

Living polymerization techniques exist for monomers that are polymerized radically or anionically. Historically, this field started with the living anionic polymerization of styrene, leading to the anionic block copolymers of the KRATON® type, a linear triblock copolymer of polystyrene(glassy)-*co*-polydiene(rubber)-*co*-polystyrene(glassy). Newer and commercially important developments include the industrially attractive radical techniques, which have developed into a wide array of synthetic polymerization methodologies, including atom-transfer radical polymerization (ATRP),

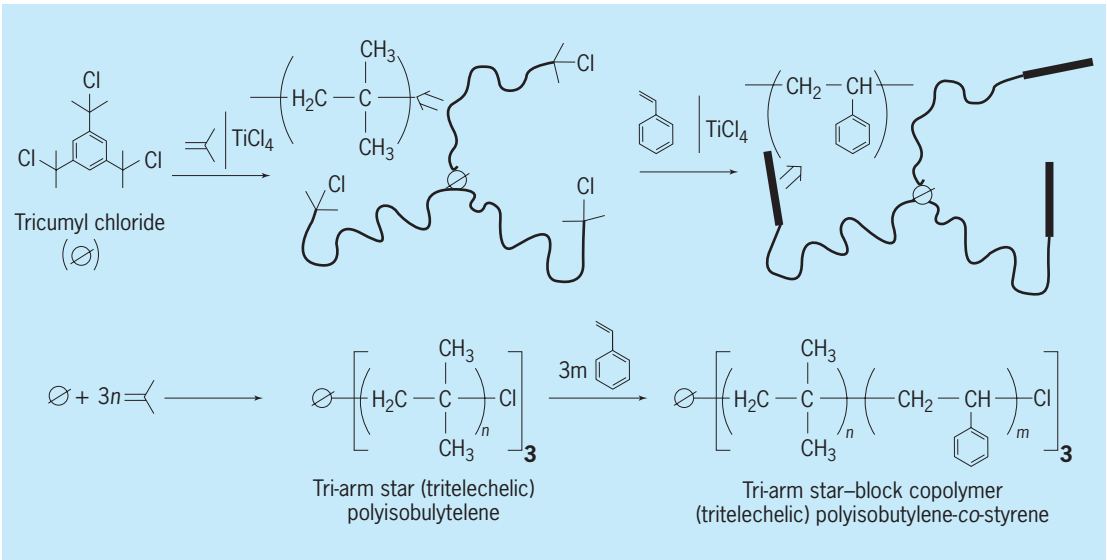

Fig. 9. Complex architectures via prevention of termination and transfer reaction.

stable free-radical polymerization (SFRP), and reversible addition-fragmentation transfer (RAFT). *See* CONTROLLED/LIVING RADICAL POLYMERIZATION; MACROMOLECULAR ENGINEERING.

Enzyme Catalysis

A new field has been receiving considerable attention both academically and industrially. Enzymes can by "recruited" for polymerization processes in different ways. One possibility is to use enzymatic catalysis to synthesize the monomer feedstock. An example is the synthesis of 1,3-propanediol from the renewable feedstock corn sugar in a fermentation reaction for the Sorona® process. The other possibility is to catalyze a wide variety of polymerization reactions under in-vitro conditions. The potential of enzymes spans from condensation polymerizations over various step-growth polymerizations. It includes the controlled synthesis of numerous examples of polysaccharides, polyesters, polyphenols, polyurethanes, and even the semiconducting polyaniline, as well as different vinyl polymers. In addition to novel polymer syntheses, enzymes are used for reactions involving the modification of natural polymers. It is even possible to synthesize block copolymer architectures.

Under in-vivo conditions, enzymes polymerize monomers with high precision in regard to the polymer structure, including molecular weight. This can be potentially exploited under in-vitro conditions, and polymer products can be obtained that would be very difficult or impossible to make by conventional polymerization methods. However, the development of such polymerization processes involve enzyme engineering—the targeted mutation of the respective enzyme by means of accelerated evolution strategies—in order to achieve the desired polymer structure. Combinatorial methods play an important role in the selection process. *See* ENZYME.

Ralf M. Peetz

Bibliography. R. F. Brady, *Comprehensive Desk Reference of Polymer Characterization and Analysis*, 2003; J. Brandrup and E. H. Immergut (eds.), *Polymer Handbook*, vol. 1: *Polymerization and Depolymerization*, 4th ed., 1999; J. A. Brydson, *Plastics Materials*, 7th ed., 1999; N. P. Cheremisinoff, *Handbook of Polymer Science and Technology*, vol. 1: *Synthesis and Properties*, 1989; R. T. Fenner, *Principles of Polymer Processing*, 1980; H. R. Kricheldorf, O. Nuyken, and G. Swift, *Handbook of Polymer Synthesis*, 2d ed., 2005; H. F. Mark et al. (eds.), *Encyclopedia of Polymer Science and Engineering*, 3d ed., 12 vols., 2003–2004; G. Odian, *Principles of Polymerization*, 4th ed., 2004; P. Rempp and E. W. Merrill, *Polymer Synthesis*, 2d ed., 1991.

Polymicrobial diseases

Contraction of multiple microbial infections. Such infections, which can occur in humans and animals, may be sequential or concurrent. For example, the initial infection may predispose the host to subsequent microbial infections, or the initial infecting microorganism may induce a state of immunosuppression. Concurrent infections are caused by various combinations of viruses, bacteria, fungi, and parasites. *See* BACTERIA; DISEASE; FUNGI; PARASITOLOGY; VIRUS.

Polymicrobial diseases are not new. Infections involving numerous pathogens were reported early in the twentieth century as complex infections, complicated infections, dual infections, mixed infections, secondary infections, co-infections, synergistic infections, or concurrent infections. In some instances, they were referred to as complexes or syndromes. In humans, acute necrotizing ulcerative gingivitis and respiratory diseases were recognized very early as having a polymicrobial etiology. In the 1920s, *Haemophilus influenzae* or *Streptococcus pneumoniae* were routinely found in individuals with viral respiratory disease. A similar relationship was seen during the influenza pandemics in 1918, 1957, and 1968–1969. In animals, foot rot and chronic nonprogressive pneumonia in sheep were among the first diseases found to have multiple etiologies. Respiratory disease in cattle was later shown to involve multiple infections by microorganisms. *See* INFECTION; INFECTIOUS DISEASE; PATHOGEN.

Interactions leading to diseases. Microorganisms can interact in a variety of ways, causing both acute and chronic diseases in humans and animals (**Tables 1** and **2**). One microorganism generates a niche for other pathogenic microorganisms to colonize; the presence of one microorganism predisposes the host to colonization by other microorganisms; two or more nonpathogenic microorganisms together cause pathology; or one microorganism interferes with infection or colonization of other microorganisms. Many of these polymicrobial interactions occur within biofilms that form on natural or artificial surfaces within the human host. Biofilms are sessile aggregate communities of bacteria attached to surfaces in hydrated polymeric matrices of exopolysaccharides. Organisms in biofilms are very hardy and resistant to antimicrobial agents as well as to host innate and acquired immune defenses. *See* BIOFILM.

Synergistic infections. Synergism results when one microorganism creates an environment that favors the infection and colonization of other, often pathogenic microorganisms. Recent examples are the concurrent infection of human metapneumovirus and coronavirus in severe acute respiratory syndrome (SARS) and concurrent infection of human metapneumovirus and respiratory syncytial virus in bronchiolitis and other respiratory infections. Other examples include measles virus, human T-lymphotropic virus type I (HTLV-I), HTLV-II, and human immunodeficiency virus (HIV), which induce immunosuppression favoring secondary infections with other viruses, bacteria, fungi, and protozoa. Concurrent Epstein-Barr virus and retrovirus infections may result in multiple sclerosis, but this is still to be confirmed.

Infections predisposing to disease. The presence of one microorganism may predispose the host to

TABLE 1. Combinations of microorganisms in polymicrobial diseases of humans

Disease	Etiologic agents
Synergistic polymicrobial infections	
Severe acute respiratory syndrome	Metapneumovirus and coronavirus
Bronchiolitis	Metapneumovirus and respiratory syncytial virus
Measles	Measles and *Mycobacterium tuberculosis* and *Staphylococcus aureus*
AIDS	HTLV-I, HTLV-II, and/or HIV-1, HIV-2; HIV and *Mycobacterium tuberculosis;* HBV or HCV and HIV-1; HIV and enteric viruses, *Acinetobacter radioresistens*, *Mycobacterium tuberculosis*, *Ehrlichia chaffeensis*, *Candida albicans*, *Histoplasma capsulatum*, *Cryptosporidium parvum*, *Trichomonas vaginalis*, and others
Multiple sclerosis	Epstein-Barr virus and retrovirus (unconfirmed)
Infections predisposing to polymicrobial disease	
Respiratory disease	Influenza viruses, parainfluenza viruses, respiratory syncytial viruses, adenoviruses, measles viruses, rhinoviruses and coronaviruses with *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Haemophilus influenzae*, *Staphylococcus aureus*, *Neisseria meningitidis*, *Mycobacterium tuberculosis*, or *Bordetella pertussis*
Otitis media	Respiratory tract viruses and bacterial infections
Invasive streptococcal disease	Varicella-zoster virus and *Streptococcus pyogenes*
Additive polymicrobial infections	
Periodontal infections	Over 500 bacterial species and fungi; some viruses are suspect
Pertussis	Nasopharynx can be colonized with species of pathogenic bacteria in addition to *Bordetella pertussis*
Lung abscesses	*Nocardia asteroides* and *Cryptococcus neoformans*
Tuberculosis	Herpes zoster and *Mycobacterium tuberculosis*
Cystic fibrosis	*Pseudomonas aeruginosa*, *Stenotrophomonas maltophilia*, *Prevotella oris*, *Fusobacterium gonidoformans*, *Bacteroides fragilis*, *Leptotrichia*-like spp., *Abiotrophia defecta*, *Citrobacter murliniae*, *Lautropia mirabilis*, and *Sarcina ventriculi*
Peritonitis	Aerobic and anaerobic gram-positive and gram-negative bacteria
Hepatitis	Any of hepatitis B, C, and D viruses; also hepatitis C virus and HIV
Gastroenteritis	Norwalk-like virus and *Aeromonas sobria* or *Escherichia coli*
Urinary tract infection	Combinations of *Corynebacterium urealyticum*, *Gardnerella vaginalis*, *Anaerococcus lactolyticus*, *Bacteroides vulgatus*, *Dialister invisus*, *Fusobacterium nucleatum*, *Lactobacillus iners*, *Leptotrichia amnionii*, *Prevotella buccalis*, *Prevotella ruminicola*, *Rahnella aquatilis*, and *Streptococcus intermedius*
Endocarditis	*Staphylococcus*, *Streptococcus*, and the HACEK group
Microbial interference	
AIDS	Flavivirus interferes with HIV infection
Staphylococcus aureus–related disease	*Steptococcus pneumoniae* suppresses the carriage of *S. aureus*

TABLE 2. Combinations of microorganisms in polymicrobial diseases of animals

Disease	Etiologic agents
Synergistic polymicrobial infections	
Immunosuppression in poultry	Infectious bursal disease virus, hemorrhagic enteritis virus, chicken anemia virus, avian leukosis virus, reticuloendotheliosis virus, and avian reovirus
Immunosuppression in cattle	Bovine respiratory syncytial virus and bovine diarrhea virus
Immunosuppression in pigs	Porcine reproductive and respiratory syndrome virus
Infections predisposing to polymicrobial disease	
Bovine respiratory disease complex	Respiratory infections in cattle involving 13 different viruses followed by secondary infection with one or more of 12 different bacteria
Porcine respiratory disease complex	Respiratory infections in swine involving 5 different viruses followed by secondary infection with one or more of 8 different bacteria
Additive polymicrobial infections	
Porcine atrophic rhinitis	Nasal infections in swine involving *Pasteurella multocida* and *Bordetella bronchiseptica*
Ovine conjunctivitis	Eye infections in sheep with *Branhamella ovis*, *Mycoplasma conjunctivae*, and *Chlamydia psittaci*
Ovine foot rot	Hoof infections in sheep with *Dichelobacter nodosus* and mixtures of other bacteria
Ovine pneumonia	Lung infections in sheep involving *Mycoplasma ovipneumoniae* often with mixtures of other viruses and bacteria
Bovine gastroenteritis	Enteric in cattle infections involving *Salmonella enterica* serovar Dublin, *S. enterica* serovar Typhimurium, coronavirus, and bovine diarrhea virus
Porcine gastroenteritis	Enteric infections involving bovine diarrhea virus and transmissible gastroenteritis virus
Poultry enteritis mortality syndrome	Enteric infections in turkeys involving coronavirus and *Escherichia coli*
Infectious coryza	Eye infections in poultry with *Haemophilus paragallinarum* and *Ornithobacterium rhinotracheale*

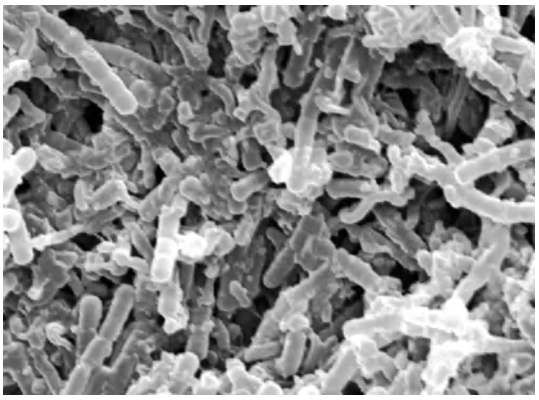

Scanning electron micrograph of human gingival plaque showing a mixed biofilm of rods and cocci. (*Courtesy of John E. Laffoon and Janet M. Guthmiller, Dows Institute for Dental Research, College of Dentistry, University of Iowa, Iowa City*)

colonization or infection by a second microorganism. Respiratory viral infections are notorious for predisposing humans and animals to secondary bacterial infections. In humans, influenza virus increases the incidence of bacterial pneumonia by *Staphylococcus aureus*, *Haemophilus influenzae*, and *Streptococcus pneumoniae*. In addition, influenza virus and respiratory syncytial virus increase the incidence of bacterial otitis media. In cattle, 13 different viruses can predispose the bovine respiratory tract to co-infection by any one of 12 different bacteria. In pigs, 5 different viruses can predispose the porcine respiratory tract to co-infection by any one of 8 different bacteria.

Additive infections. Multiple pathogens are involved in the etiology of bacteremia, abdominal abscess/secondary peritonitis, lung abscess, odontogenic infections, brain abscess/subdural empyema, chronic otitis media/mastoiditis, liver infections, and soft tissue infection/fasciitis. Human gingival plaque mixed biofilms (see **illustration**) contain interdependent microorganisms such as *Streptococcus gordonii*, *S. mitis*, *S. oralis*, *S. sanguis*, *Haemophilus parainfluenzae*, *Propionibacterium acnes*, *Veillonella atypica*, *Actinomyces naeslundii*, *Fusobacterium nucleatum*, *Actinobacillus actinomycetemcomitans*, *Porphyromonas gingivalis*, *Treponema denticola*, and *Prevotella intermedia*. Likewise, in many human abscesses there are extensive numbers of aerobic, anaerobic, gram-positive, and gram-negative bacteria.

Microbial interference. Some microorganisms can actually suppress the colonization of other microorganisms. GB virus C (GBV-C) is a flavivirus which is not known to be pathogenic in humans. However, infection with GBV-C is associated with prolonged survival among HIV-positive people. The reason for this association is not known, but likely is due to a concept called viral interference.

Microbial interference also occurs among colonized bacteria. For example, the presence or carriage of *Steptococcus pneumoniae* also suppresses the carriage of *Staphylococcus aureus* in children. Attempts to reduce the carriage of *S. pneumoniae* by vaccinating children with pneumococcal conjugate vaccines not only reduces the nasopharyngeal carriage of vaccine-type *S. pneumoniae* but results in a shift toward the carriage of nonvaccine *S. pneumoniae* serotypes or toward the presence of higher carriage rates of *Staphylococcus aureus*, including the methicillin-resistant form.

Future research areas. Polymicrobial diseases are serious diseases whose etiologic agents are sometimes difficult to diagnose and treat. Future work will undoubtedly focus on improving the identification of multiple microorganisms in polymicrobial diseases, defining microbe-to-microbe and microbe-to-host interactions, and improving prophylaxis and treatment of these diseases.

The diverse microbiota infecting and colonizing humans is currently being defined by metagenomic analyses, quantitative 16S ribosomal DNA hybridization assays, real-time polymerase chain reaction, terminal restriction fragment length polymorphism, and length heterogeneity PCR. Flow-cells are being used to study microbial interactions in biofilms. These techniques, however, cannot presently be utilized in clinical laboratories to assist clinicians in daily practice. Clinical microbiology laboratories presently identify polymicrobial intra-abdominal or soft-tissue infections as fecal flora or mixed gram-positive flora and struggle to identify viral agents in mixed respiratory tract infections. As a result, critical-care physicians have trouble justifying invasive sampling techniques such as broncho-alveolar lavage to identify the true etiology of mixed pulmonary infections. *See* MICROBIOTA (HUMANS).

Traditional therapies of many of these diseases are generally targeted at individual etiologic agents without consideration for effect on the polymicrobial communities. For example, treatment with broad-spectrum antibiotic agents may kill anaerobic flora in the intestinal tract, favoring infection and colonization of multiresistant bacteria such as the vancomycin-resistant enterococcus. In some cases, however, there are beneficial consequences in that treatment of the primary infectious agent may reduce or eliminate infection by a concurrent or sequential agent. In Africa, tuberculosis is prominent among patients with HIV infection. The use of antiretroviral drugs reduces the incidence of tuberculosis to the level observed immediately after HIV seroconversion. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); TUBERCULOSIS.

Probiotics are novel prophylactic or therapeutic strategies for managing polymicrobial diseases. *Saccharomyces boulardii*, lactobacilli, or enterococci prevent or reduce the duration of gastroenteritis caused by bacteria and rotaviruses, and treatment with lactobacilli such as *Lactobacillus fermentum*, *L. rhamnosus*, or *L. crispatus* have the potential to reduce vaginal infections. Future studies are needed to determine the probiotic mechanisms of microbial inhibition as well as the optimal conditions for use.

Summary. Polymicrobial diseases have a multifactorial etiology and develop as a result of complex

interactions among environmental factors, host factors (genetic predisposition, stress, nutritional state, etc.), and exposure to pathogens. Understanding the role of predisposing factors, interactions among etiologic agents, and host responses to infection may ultimately assist in better strategies for prophylaxis and/or treatment. Kim A. Brogden; Janet M. Guthmiller

Bibliography. L. O. Bakaletz, Developing animal models for polymicrobial diseases, *Nat. Rev. Microbiol.*, 2:552–568, 2004; K. A. Brogden and J. M. Guthmiller (eds.), *Polymicrobial Diseases*, ASM Press, Washington, DC, 2002; K. A. Brogden and J. M. Guthmiller, Polymicrobial diseases, a concept whose time has come, *ASM News*, 2003; K. A. Brogden, J. M. Guthmiller, and C. E. Taylor, Human polymicrobial infections, *Lancet*, 365:253–255, 2005; H. Smith, The role of microbial interactions in infectious disease, *Phil. Trans. Roy. Soc. Lond. B Biol. Sci.*, 297:551–561, 1982.

Polymixiiformes

The only order of the superorder Polymixiomorpha, comprising one family (Polymixiidae) with one genus (*Polymixia*) and ten extant species. Several fossil genera are known from the Upper Cretaceous. Polymixiidae was originally placed with the Beryciformes; however, recent evidence suggests that the family may be the sister group to all other acanthomorphs (spiny-rayed fishes). The family is known commonly as beardfishes. *See* BERYCIFORMES; OSTEICHTHYES; TELEOSTEI.

The beardfishes are so called for the pair of long hyoid (between the lower jaws) barbels (see **illustration**). They are further distinguished by a moderately elongate and compressed body; subocular shelf, orbitosphenoid, and basisphenoid present; two supramaxillae; four branchiostegal rays; a continuous dorsal fin, with 4–6 spines and 26–38 soft rays; an anal fin with 4 short spines and 13–17 soft rays; subabdominal pelvic fins with one spinelike ray and six soft rays; 16 branched caudal rays; about 33–38 lateral-line scales; and two sets of intermuscular bones, which is a condition otherwise unknown among the acanthomorphs.

Polymixiids are bathydemersal [living and feeding on the bottom below 200 m (660 ft)] fishes to depths of 700 m (2300 ft), occurring in warm temperate to tropical waters of the Atlantic, Indian, and Western Pacific oceans. They feed on other fishes, worms, and crustaceans and attain a maximum length of 38 cm (15 in.). Herbert Boschung

Pacific beardfish (*Polymixia berndti*). (*Photo by Jack E. Randall*)

Bibliography. R. Froese and D. Pauly (eds.), FishBase, World Wide Web electronic publication, version (05/2005); G. D. Johnson and C. Patterson, Percomorph phylogeny: A survey of acanthomorphs and a new proposal, *Bull. Mar. Sci.*, 52(1):554–626, 1993; J. A. Moore, Polymixiidae, pp. 960–962 in K. E. Carpenter (ed.), *The Living Marine Resources of the Western Central Atlantic*: *FAO Species Identification Guide for Fishery Purposes*, vol. 2, FAO, Rome, 2003 (dated 2002); J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; M. L. J. Stiassny, The limits and relationships of the acanthomorph teleosts, *J. Zool. Lond.* (B), 1:411–460, 1986.

Polymorphism (crystallography)

The existence of different crystal structures with the same chemical composition. If only one chemical element is present, the forms are called allotropes. Graphite and diamond are allotropes of carbon, whereas quartz and cristobalite are polymorphs of silica (silicon dioxide, SiO_2). Although properties are different in these forms, reversible transformations, which involve small shifts in atom positions and no bulk transport of material, are common. The quartz transformation at 1063°F (573°C) is a reversible, atom-displacement transformation. *See* SILICA MINERALS.

In metals and ceramics, similar transformations are called martensitic. Advantage is taken of the localized nature of reversible transformation in steel by controlling the melting atmosphere, temperature, composition, mechanical working (alloying), and tempering and quenching operations. *See* HEAT TREATMENT (METALLURGY).

Transformations. Transformation from one polymorphic form to another is a response to changed conditions, most commonly temperature. The best example is carbon. Diamond is made only under enormous pressures and, when heated to very high temperatures, is transformed irreversibly from a three-dimensional, tetrahedrally coordinated structure to the two-dimensional layered structure of graphite. The large spacing between layers accounts for the lubricity of graphite. The adjustability of this spacing reflects weak interlayer forces. This adjustability is exploited in the production of intercalation compounds. *See* DIAMOND; GRAPHITE; INTERCALATION COMPOUNDS.

With temperature as the only variable, tin undergoes a dramatic reversible transformation, resulting in spontaneous brittle fracture. Normally metallic, tin expands by 21 vol % and becomes nonmetallic below 64°F (18°C). *See* TIN.

Control of mechanical properties. Control over transformations to achieve desirable properties as either devices or structural materials in extreme environments is a frequent objective. In the case of tin, reversibility on the atomic scale can have devastating consequences for bulk properties. Similar transformations may be beneficial in the right place and in the desired degree. Such transformation is attempted with metals and ceramics.

As with tin, but at much higher temperatures, zirconium dioxide (ZrO_2) contracts on heating and expands on cooling. Addition of other oxides gives a new structure that does not transform. This structure can be toughened by small internal regions of zirconium oxide that do transform. In transforming during cooling, these regions expand, occupying more volume. As they push against the surrounding material, internal cracks are sealed off and do not propagate, which could result in fracture. This procedure is called transformation toughening. It makes it possible to modify ceramic materials so that they exhibit the desired mechanical properties at higher temperatures than metal alloys. *See* CERAMICS; CRYSTAL STRUCTURE. Doris Evans

Bibliography. W. L. Bragg and G. F. Claringbull, *Atomic Structures of Minerals,* 1964; J. Donahue, *The Structures of the Elements,* 1974, reprint 1982; H. Kammimura, Graphite intercalation compounds, *Phys. Today,* 40(12):64–71, 1987.

Polymorphism (genetics)

A form of genetic variation, specifically a discontinuous variation, occurring within plant and animal species in which distinct forms exist together in the same population, even the rarest of them being too common to be maintained solely by mutation. Thus the human blood groups are examples of polymorphism, while geographical races are not; nor is the diversity of height among humans, because height is "continuous" and does not fall into distinct tall, medium, and short types. *See* MUTATION.

Control mechanisms. Distinct forms must be controlled by some switch which can produce one form or the other without intermediates such as those arising from environmental differences. This clear-cut control is provided by the recombination of the genes. Each gene may have numerous effects and, in consequence, all genes are nearly always of importance to the organism by possessing an overall advantage or disadvantage. They are very seldom of neutral survival value, as minor individual variations in appearance often are. Thus a minute extra spot on the hindwings of a tiger moth is in itself unlikely to be of importance to the survival of the insect, but the gene controlling this spot is far from negligible since it also affects fertility. *See* RECOMBINATION (GENETICS).

Genes having considerable and discontinuous effects tend to be eliminated if harmful, and each gene of this kind is therefore rare. On the other hand, those that are advantageous and retain their advantage spread through the population so that the population becomes uniform with respect to these genes. Evidently, neither of these types of genes can provide the switch mechanism necessary to maintain a polymorphism. That can be achieved only by a gene which has an advantage when rare, yet loses that advantage as it becomes commoner.

Natural selection. Occasionally there is an environmental need for diversity within a species, as in butterfly mimicry. Mimicry is the resemblance of different species to one another for protective purposes, chiefly to avoid predation by birds. It takes two forms. (1) In Batesian mimicry, palatable species resemble those that are distasteful or poisonous. Such mimics generally copy several protected models and are therefore polymorphic. (2) In Mullerian mimicry, protected species resemble one another. These attempt to reduce the risk of predation by producing a uniform type with conspicuous warning patterns, such as those of bees and wasps, and are not therefore affected by polymorphism.

Sexual dimorphism falls within the definition of genetic polymorphism. In any species, males and females are balanced at optimum proportions which are generally near equality. Any tendency for one sex to increase relative to the other would be opposed by selection.

Batesian mimicry and sexual dimorphism, however, are rather special instances. In general, a gene having both advantageous and disadvantageous effects may gain some overall advantage and begin to spread because one of the features it controls becomes useful in a new environment. A balance is then struck between the advantages and disadvantages of such a gene, ensuring that a proportion of the species carry it, thus giving rise to permanent discontinuous variation, that is, to polymorphism. *See* PROTECTIVE COLORATION.

Occurrence in a population. Polymorphism is increasingly known to be a very common situation. Its existence is apparent whenever a single gene having a distinct recognizable effect occurs in a population too frequently to be due merely to mutation, certainly if present in over 1% of the individuals. Even if recognized by some trivial effect on the phenotype, it must in addition have important other effects. About 30% of the people in western Europe cannot taste as bitter the substance phenylthiourea. This is truly an insignificant matter; indeed, no one even had the opportunity of tasting it until the twentieth century. Yet this variation is important since it is already known that it can affect disease of the thyroid gland. *See* GENETICS; POPULATION GENETICS. E. B. Ford

Methods of study. Genetic polymorphisms can be of several types, depending upon the level of study: visible, chromosomal, protein (enzymes), and deoxyribonucleic acid (DNA).

Visible polymorphisms. Visible polymorphisms are easily studied by observing and counting the frequencies of the different morphs in populations. Examples of such polymorphisms are the color and banding patterns of the land snail (*Cepaea*), and eye color and hair color in human populations.

Chromosomal polymorphisms. Various cytological techniques are required for the study of chromosomal polymorphisms. Tissues with chromosomes particularly favorable for study are stained with a dye which makes the chromosomes easily visible through a microscope. The most favorable tissues for such studies are the salivary glands of *Drosophila* (and some other Diptera), which have "giant" polytene chromosomes. Such chromosomes have a characteristically complex banding pattern which allows identification of each part of each chromosome. For example, if a population is polymorphic for inversions, the order of genes in some sections of some chromosomes will be inverted with respect to the gene order of other chromosomes. Inversions will be evident by an inversion of the banding pattern in the inverted segment. Such inversion polymorphisms are common in many species of *Drosophila* and have been the subject of considerable research. *See* CHROMOSOME; CHROMOSOME ABERRATION.

Most organisms, however, do not have giant polytene chromosomes. To study chromosomal polymorphisms, it is necessary to study cells undergoing division, either mitosis or meiosis. It is possible to obtain particularly favorable tissue at the right stage of mitosis by tissue culture techniques. However, meiotic cell divisions of the gonadal tissues are often more informative to study than mitotic divisions. When homologous chromosomes pair at meiosis, it is possible to detect structural differences between the chromosomes. For example, an individual heterozygous for an inversion polymorphism will display a loop in the section of the chromosome involved in the inversion. *See* MEIOSIS.

In addition to inversion polymorphisms, other chromosomal polymorphisms are known. Translocation polymorphisms (involving an exchange of chromosomal material between nonhomologous chromosomes) are known in many plants. Aneuploid polymorphisms (addition or deletion of whole chromosomes) have been studied in several mammal species.

Techniques of studying chromosomes have made it easier to detect chromosomal polymorphisms. Chromosomes which are not polytene can be induced to show characteristic banding patterns. Thus it is becoming possible to detect chromosomal polymorphisms in almost any organism by looking at any tissue. It is not necessary to use cells in meiosis, nor is it necessary that the individual be a heterozygote in order to detect the variation. Such banding techniques have been a great advance in studying chromosome variation and have proved especially successful in studying human chromosomes.

Protein polymorphisms. There are three basic types of protein polymorphisms: deficiency, antigenic, and electrophoretic. Deficiency polymorphisms occur when certain genotypes in a population do not produce a certain enzyme or the enzyme produced is nonfunctional. By assaying individuals for the presence or absence of a certain enzyme activity, it is possible to detect such polymorphisms. Examples of this type of biochemical polymorphism in human populations are glucose-6-phosphate dehydrogenase deficiency and lactose deficiency in adults.

Antigenic polymorphisms are detected by preparing antibodies specific for different antigens. Thus if a population is polymorphic for an antigen, tissues from different genotypes will interact with different antibody preparations. For example, blood from humans with type A blood reacts with anti-A serum; blood from type B individuals reacts with anti-B serum; blood from type AB individuals reacts with both anti-A and anti-B sera; blood from type O individuals reacts with neither antiserum. A widespread polymorphism in vertebrates is that of transplantation antigens, mainly responsible for separation of grafts between individuals of a species. *See* ANTIBODY; ANTIGEN; BLOOD GROUPS; TRANSPLANTATION BIOLOGY.

Electrophoretic polymorphisms are very common and occur in high frequencies in plants, animals, and humans. Proteins from different individuals are extracted and applied to an electrophoresis gel; such gels can be made from starch or polyacrylamide and have a consistency somewhat like gelatin. An electrical field is applied across the gel. Since proteins usually have a net electrical charge, they migrate in the gel toward the pole opposite to the charge on the protein. The rate of migration is determined by the size of the electrical charge as well as the size and shape of the protein. Thus if a mutation which results in a change in the amino acid composition of a certain protein has occurred in the gene coding for that protein, the change is often detected by the change in electrophoretic mobility of the protein. Thus different genotypes in a population produce proteins which migrate differently in an electrophoretic gel. The easiest proteins to study by this method are enzymes because simple techniques are available for visualizing many enzymes in the gel. *See* ELECTROPHORESIS; PROTEIN.

Electrophoretic studies have been made on a variety of proteins in many species of organisms. Much genetic variation has been revealed. On the average, about one-half of all proteins studied are electrophoretically polymorphic in any population; of course, different populations and different species may vary in their degree of electrophoretic variation. About 30% of the more than 100 proteins studied in humans are polymorphic for electrophoretic variation. Such studies have had very important implications for population genetics and evolutionary studies.

DNA polymorphisms. Advances in deoxyribonucleic acid (DNA) technology have made it possible to study polymorphisms at the most basic level. Two methodologies have been developed. One involves the use of enzymes called restriction endonucleases, and the other determines actual DNA sequences. Restriction endonucleases recognize specific sets of DNA sequences, usually a contiguous set of four to eight bases. After DNA has been digested with such enzymes, the sizes of the DNA fragments produced can be determined by electrophoresis. Individuals polymorphic for the recognition sites of restriction

enzymes will produce differently sized fragments. Hence, such polymorphisms are called restriction fragment length polymorphisms (RFLP). They are ubiquitous and are extremely useful genetic markers for a large number of purposes, including identifying the genetic location of disease-causing genes. In addition, highly repetitive DNA sequences have been studied by using restriction enzymes. The resulting size patterns produced appear to be virtually unique for unrelated individuals. This technology, called DNA fingerprinting, has played a role in some legal cases. *See* FORENSIC MEDICINE; RESTRICTION ENZYME.

A direct way of studying DNA polymorphism is to determine actual DNA sequences of the same genetic region in several individuals. Results from DNA sequence data indicate that different copies of the same gene may differ by 0.5–2.0% of their base positions. The number of alleles (defined on the nucleotide level) is thus tremendous. However, most differences among alleles are silent; that is, they do not change the amino acid sequence of the protein being coded.

In addition to base-pair differences revealed by RFLP analysis and DNA sequencing, insertions or deletions of various lengths of DNA bases are also common. These may be a few base pairs in length or extend up to several thousand bases. These latter insertions or deletions may be caused by transposable elements, segments of DNA that have the ability to move around the genome. *See* HUMAN GENETICS; TRANSPOSONS.

Jeffrey R. Powell

Bibliography. J. A. Bryant and V. L. Dunham, *DNA Replication in Plants,* 1988; S. Heinemann and J. Patrick (eds.), *Molecular Neurobiology: Recombinant DNA Approaches,* 1987; J. B. Jenkins, *Human Genetics,* 1990; K. Kammermeyer and V. L. Clark, *Genetic Engineering Fundamentals: An Introduction to Principles,* 1989; M. Levitan, *Textbook of Human Genetics,* 1988.

Polynomial systems of equations

A collection of equations in several unknowns, expressed in terms of addition, subtraction, and multiplication. System (1) is a system of polynomial equations in the three unknowns (or variables) x, y, and z. Typically, the objective is to solve such a system, which means, for Eqs. (1), to find all choices of numbers to substitute for x, y, and z so that all three equations are simultaneously satisfied (in other words, all equal zero). For example, $x = 0$, $y = 1$, and $z = -1$ satisfies Eqs. (1). There are exactly six solutions to Eqs. (1). They are $(x,y,z) = (0,\pm1,-1)$, $(3,\pm2,-4)$, $(8,\pm3,-9)$.

$$\begin{aligned} x^2 + xz + y^2 - 1 &= 0 \\ x^3 - 11x^2 + 23x + y - 1 &= 0 \\ x + z + 1 &= 0 \end{aligned} \tag{1}$$

A system of polynomial equations may have any number of equations and any number of variables. Polynomial systems arise commonly in science and engineering, for example, as chemical equilibrium systems, as equations of kinematics (in robotics), as design criteria systems (in the design of mechanisms), and as systems expressing the intersection of curves and surfaces (in computer graphics and geometric modeling). Solving such systems is challenging, although certain special cases, for example, systems of linear equations and polynomials in only one variable, are more routine.

General system. Since a general polynomial system may have more than three variables, it is customary to denote the variables by using subscripts, as x_1, $x_2, \ldots, x_n$, where n denotes the number of variables. Similarly, the number of equations may be other than three; in this article, this number is denoted by m. (The most common case is $m = n$.) Then, a precise mathematical description of a general polynomial system is: a collection of mathematical equations of the form of system (2). Each $f_i(x_1, x_2, \ldots, x_n)$, $i = 1, 2, \ldots, m$, is a sum of terms of the form shown as expression (3), where the coefficient a

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= 0 \\ f_2(x_1, x_2, \ldots, x_n) &= 0 \\ &\ldots\ldots\ldots\ldots \\ f_m(x_1, x_2, \ldots, x_n) &= 0 \end{aligned} \tag{2}$$

$$a\, x_1^{i_1}\, x_2^{i_2} \ldots x_n^{i_n} \tag{3}$$

is a constant, and the exponent i_j of the variable x_j is a nonnegative whole number. (A zero exponent means the variable is omitted.) The values of the coefficients and exponents may change from term to term. In this notation, the second equation of system (1) is $1\, x_1^3 x_2^0 x_3^0 - 11\, x_1^2 x_2^0 x_3^0 + 23\, x_1^1 x_2^0 x_3^0 + 1\, x_1^0 x_2^1 x_3^0 - 1\, x_1^0 x_2^0 x_3^0 = 0$.

The expressions $f_i(x_1, x_2, \ldots, x_n)$ are called polynomials in several variables (or multivariable polynomials). The problem posed by system (2) is to characterize the set of solutions, where a solution is a choice of values for the variables $x_1 = a_1$, $x_2 = a_2, \ldots, x_n = a_n$, which simultaneously satisfies each equation of the system. For example, the objective might be to count the solutions or to compute some (or all) of them explicitly. Solution sets of polynomial systems can be geometrically complicated, and most solution methods are limited to finding subsets of solutions (for example, all nonsingular solutions). There is no practical reliable method for computing the full solution set to general polynomial systems. Characterizing the solution sets to polynomial systems is the subject matter of algebraic geometry, while techniques for solving polynomial systems are usually classified as part of numerical analysis. *See* ALGEBRAIC GEOMETRY; NUMERICAL ANALYSIS.

Polynomial equations are examples of general nonlinear equations, where transcendental as well as polynomial terms may appear. (For example, terms in general nonlinear equations may include exponential or trigonometric functions.) Thus, methods for

solving systems of general nonlinear equations are applicable to systems of polynomials.

Special cases. The most classical special case of polynomial systems is polynomials in one variable. [Thus, in terms of system (2), $m = 1$ and $n = 1$.] The degree of a polynomial in one variable is the largest degree occurring in any term. [In terms of expression (3), it is the largest i_1.] The fundamental theorem of algebra states that a polynomial has the same number of solutions as its degree. For example, a degree-10 polynomial has 10 solutions. (Solutions that are singular have to be counted by their multiplicity.) For systems with more than one variable, the fundamental theorem of algebra generalizes to results usually known as Bezout's theorem, which give the generic number of solutions in terms of the degrees of the equations. In one variable, equations of degree 2, 3, and 4 have solutions that can be expressed by elementary formulas (the quadratic, cubic, and quartic formulas, respectively). A famous consequence of Galois theory is that no such formulas exist for polynomials of degree greater than four. In computations, these formulas are tricky to use, and numerical techniques (rather than formulas) are often more reliable. No analog of these formulas for multivariable polynomial systems exists. *See* EQUATIONS, THEORY OF; GROUP THEORY.

The other special case of note occurs when the equations are linear. [In terms of expression (3), $i_1 + i_2 + \cdots = i_n \leq 1$.] Cramer's rule provides an elementary formula for the solutions of a linear system, which is computationally useful when $n = 2$ or $n = 3$. For larger linear systems, numerical methods, such as gaussian elimination, are faster and more reliable. Numerical methods for solving linear systems and polynomials in one variable are highly developed and reliable, compared to the general case, discussed below. *See* LINEAR SYSTEMS OF EQUATIONS.

Methods of solution. Methods for solving polynomial systems of equations include single-solution methods, polynomial continuation, the method of resultants, and generalized bisection.

Single-solution methods. These methods aim to find only a single solution. They are usually for systems of general nonlinear equations, not specifically for polynomial equations. Examples include Newton's method (and its variants), steepest-descent least-squares methods, and single-path continuation methods. These approaches require a start point, an estimate of the solution being sought. In some cases, the method can be shown to be globally convergent, in which case a solution will be found no matter where the start point is chosen. On the other hand (and more typically), these are local methods, in that convergence can be guaranteed only if the start point is sufficiently close to the target solution.

Polynomial continuation. This technique is a multiple-start-point method, guaranteed to find all geometrically isolated solutions but not the positive-dimensional solution components. A start system is generated that has certain generic properties of the target system (that is, the system to be solved). Then the start system is transformed by a continuous parameter (the continuation parameter) into the target system, and at the same time the solutions to the start system are transformed into the solutions to the target system.

Method of resultants. The classical method of resultants (also called elimination) has been revived and made more computationally feasible via the technique of Gröbner bases. The given polynomial system is transformed into triangular form (that is, the ith equation involves only the first i variables). Then the triangular system can be solved by using a one-variable method.

Generalized bisection. By using ideas from interval analysis (in particular, the interval Newton's method), a multivariable numerical method has been developed that is analogous to (classical) one-variable bisection techniques. Alexander P. Morgan

Bibliography. O. Balci, R. Sharda, and S. A. Zenios (eds.), *Computer Science and Operations Research: New Developments in Their Interfaces*, 1992; D. Bini and V. Pan, *Polynomial and Matrix Computations: Fundamental Algorithms*, vol. 1, 1994; I. N. Herstein, *Abstract Algebra*, 3d ed., 1996; D. Kahaner et al., *Numerical Methods and Software*, 1989; A. Morgan, *Solving Polynomial Systems Using Continuation for Scientific and Engineering Problems,* 1987; W. Rheinboldt, *Methods for Solving Systems of Nonlinear Equations*, 2d ed., 1998.

Polynuclear hydrocarbon

One of a class of hydrocarbons possessing more than one ring. The aromatic polynuclear hydrocarbons may be divided into two groups. In the first, the rings are fused, which means that at least two carbon atoms are shared between adjacent rings. Examples are naphthalene (**1**), which has two six-membered rings, and acenaphthene (**2**), which has two six-membered rings and one five-membered ring.

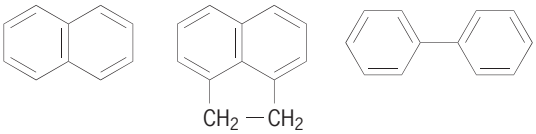

(**1**) (**2**) (**3**)

$C_6H_5CH_2CH_2C_6H_5$

(**4**)

In the second group of polynuclear hydrocarbons, the aromatic rings are joined either directly, as in the case of biphenyl (**3**), or through a chain of one or more carbon atoms, as in 1,2-diphenylethane (**4**).

The higher-boiling polynuclear hydrocarbons found in coal tar or in tars produced by the pyrolysis or incomplete combustion of carbon compounds are frequently fused ring hydrocarbons, some of which may be carcinogenic. *See* AROMATIC HYDROCARBON; STEROID. C. K. Bradsher

Polyol

Polyols or, as they are sometimes called, polyhydric alcohols, are compounds containing more than one hydroxyl group (—OH). Each hydroxyl is attached to separate carbon atoms of an aliphatic skeleton. This group includes glycols, glycerol, and pentaerythritol and also such products as trimethylolethane, trimethylolpropane, 1,2,6-hexanetriol, sorbitol, inositol, and poly(vinyl alcohol).

Polyols are obtained from many plant and animal sources and are synthesized by a variety of methods. 1,2-Glycols and glycerol are produced by the hydrolysis of epoxides or chlorohydrins. Formaldehyde reacts with acetaldehyde, propionaldehyde, and butyraldehyde to form, respectively, pentaerythritol, trimethylolethane, and trimethylolpropane. Catalytic hydrogenation of sugars produces sorbitol; 1,2,6-hexanetriol is obtained by the reduction of 2-hydroxyadipaldehyde. Saponification of poly(vinyl acetate) is employed in the industrial manufacture of poly(vinyl alcohol), which is shown below.

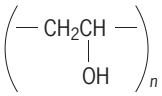

Polyols such as glycerol, pentaerythritol, trimethylolethane, and trimethylolpropane are used in making alkyd resins for decorative and protective coatings. Glycols, glycerol, 1,2,6-hexanetriol, and sorbitol find application as humectants and plasticizers for gelatin, glue, and cork. Explosives are made by the nitration of glycols, glycerol, and pentaerythritol.

The polymeric polyols used in manufacture of the urethane foams represent a series of synthetic polyols which came into prominence in the 1960s. Generally polyoxyethylene or polyoxypropylene adducts of di- to octahydric alcohols, these cover a molecular weight range of 400 to 6000. *See* GLYCEROL.

Philip C. Johnson

Polyolefin resins

Polymers derived from hydrocarbon molecules that possess one or more alkenyl (or olefinic) groups. The term polyolefin typically is applied to polymers derived from ethylene, propylene, and other alpha-olefins, and from isobutylene, cyclic olefins, and butadiene and other diolefins. Polymers produced from other olefinic monomers, such as styrene, vinyl chloride, and tetrafluoroethylene, generally are considered separately. *See* ALKENE.

Polyethylene. The term polyethylene generally refers to any homopolymer or copolymer in which ethylene is the major component monomer. Polyethylene is a semicrystalline polymer of low to moderate strength and high toughness. The melting point varies from approximately 105 to 125°C (220 to 257°F), depending on the degree of crystallinity. The density of polyethylene reflects the degree of crystallinity, and is generally used in classifying commercial grades. Low-density polyethylene, having densities ranging from 0.905 to 0.936 g/cm^3, can be subdivided into high-pressure low-density polyethylene (LDPE), and low-pressure low-density polyethylene copolymers, referred to as linear low-density polyethylene (LLDPE). This last term denotes a copolymer of polyethylene that is produced by the low-pressure polymerization process, as opposed to low-density polyethylene (or conventional polyethylene), which is produced by an older high-pressure polymerization process. In a typical linear low-density polyethylene, the percentage of comonomer would be on the order of 2–3 mol %, typically butene. High-density polyethylene (HDPE) covers the density range from 0.941 to 0.967 g/cm^3. *See* COPOLYMER.

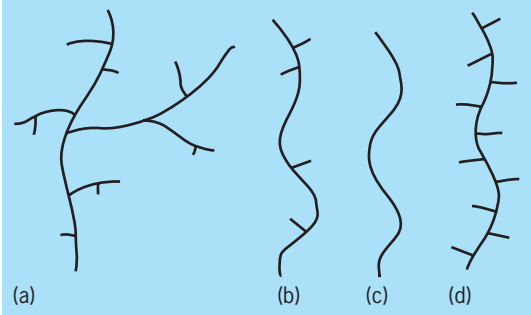

Fig. 1. Polyethylene structures. (*a*) Conventional low-density polyethylene. (*b*) Linear low-density polyethylene and high-density polyethylene with alpha-olefin comonomer. (*c*) Linear high-density polyethylene with no comonomer [a polymethylene chain]. (*d*) Metallocene low-density polyethylene.

The various types of polyethylene are characterized by differences in structure (**Fig. 1**). High-density polyethylene generally consists of a polymethylene $(CH_2)_n$ chain with no, or very few, side chains to disrupt crystallization, while linear low-density polyethylene contains side chains whose length depends on the comonomer used. The spacing and uniformity of the incorporation of the comonomer in linear low-density polyethylene depends on both the process and the type of comonomer, but the incorporation of comonomer is generally nonuniform. To some extent, the uniformity tends to decrease with the size of the comonomer. Polyethylene with structures different from those produced with previous-generation catalysts are being produced with metallocene catalysts. Metallocene materials can cover the density range of both the high-density polyethylene and the linear low-density polyethylene, with the important distinction that the structures are very uniform in both molecular weight and comonomer incorporation. *See* METALLOCENES.

Polymerization. The monomer for polyethylene is produced by cracking aliphatic hydrocarbons separated in the refining process. Low-density polyethylene is formed by the polymerization of ethylene at high pressures (15,000–50,000 lb/in.^{2} or 100–350 megapascals) and at temperatures of 150–300°C (300–570°F) in the presence of a small amount of organic peroxide or oxygen. The density is

controlled by regulating the temperature, where lower temperatures lead to higher densities. Linear low-density polyethylene is produced commercially by the copolymerization of ethylene with alpha-olefins ranging from C_3 to C_8 at temperatures up to 250°C (480°F) at low pressure (300–1100 lb/in.^{2} or 2–7.6 MPa) in the presence of a chromium catalyst or titanium-based Ziegler-Natta catalyst. *See* CATALYSIS; STEREOSPECIFIC CATALYST.

High-density polyethylene is formed at temperatures similar to those for linear low-density polyethylene and at pressures of generally less than 1000 lb/in.^{2} (7 MPa) by using either chromium or Ziegler catalysts. The molecular-weight distribution of high-density polyethylene is largely controlled by the type of catalyst, while the molecular weight is most often controlled with hydrogen level.

Metallocene, or more generally single-site, catalysts have a transition metal sandwiched between organic ring compounds. They allow molecular weight and comonomer distributions to be closely controlled, such that each polymer molecule closely resembles the others. Additionally, these catalyst systems permit the incorporation of comonomers once thought incompatible with previous catalyst generations. *See* POLYMERIZATION; SINGLE-SITE CATALYSTS (POLYMER SCIENCE).

Properties. The physical properties of polyethylene, regardless of process, are primarily dependent on the degree of crystallinity, which is related to the density and the molecular weight. Uniformity in both molecular weight and comonomer can also affect the observed properties. The stiffness, yield strength, and thermal and mechanical properties increase with the crystallinity, which can be systematically reduced by the addition of comonomer. Failure properties, such as toughness and ultimate tensile strength, increase with molecular weight, while to a lesser extent failure properties tend to decrease with broadening of the molecular-weight distribution for a given peak molecular weight. Polyethylene generally shows excellent toughness at low temperatures, because of the low glass transition temperature (generally believed to be between −80 and −120°C or −112 and −184°F). Environmental stress cracking, in the presence of detergents, is reduced for polyethylene copolymers as the size of the alpha-olefin comonomer is increased and as the density is decreased. Clarity is generally increased with decreasing crystallinity, although the morphology of the crystalline phase can also affect clarity.

The melt flow properties are affected by molecular weight and long-chain branching. Choice of fabrication technique to produce items of commerce from polyethylene generally determines the melt flow properties needed and therefore the type of polyethylene used. Among commodity polymers, polyethylene is unique with respect to the ease of formation of long-chain branching, the presence of which can dramatically affect melt fabrication options. Generally, the ease of flow for both extrusion and injection molding increases with molecular-weight distribution and long-chain branching, but decreases with molecular weight. Consequently, low-molecular-weight resins are favored for injection-molding applications. Processes requiring melt strength and strain hardening of the molten polymer, such as blown film and foam, favor high molecular weight, the presence of long-chain branching, and to a lesser extent broadening of the molecular-weight distribution. For a given molecular weight, long-chain branching can either enhance or diminish the melt viscosity at low shear rates (melt strength), with the high levels in most low-density polyethylenes yielding lower viscosity at a given molecular weight. *See* PLASTICS PROCESSING.

Use and fabrication. Polyethylene is used in greater volume worldwide than any other plastic, because it is relatively inexpensive, extremely versatile, and adaptable to a large array of fabrication techniques. Its advantages are chemical inertness (resistance to solvents, acids, and alkalis) and good dielectric and barrier properties. Major uses include packaging films, plastic containers, molded articles, foam, pipe, cable sheathing, and coatings.

Because of its broad melting range, low-density polyethylene can be used in heat-seal operations. Film applications include food packaging bags, shrink wrap, and garment bags. Blow-molded containers are used in milk and other household containers, while injection-molded items include housewares, pails, and toys. Because of its low degree of crystallinity, low-density polyethylene is limited for applications requiring high stiffness, temperature resistance, tensile strength, and good barrier properties.

Linear low-density polyethylene is a versatile material that finds wide application in plastic films. Films produced by extrusion, blown, and cast film processes are widely used as garbage bags, stretch cling films, and other items that require the properties of flexibility, tear and puncture resistance, and toughness. The toughness of linear low-density polyethylene is generally superior to low-density polyethylene and increases with the size of the comonomer. The very low density formulations of linear low-density polyethylene are also used for heat-seal applications, and are stronger than those using low-density polyethylene. While the relatively narrow molecular-weight distribution makes extrusion-coating applications somewhat difficult, specific resin grades are commercially available for this purpose. The molded items made from linear low-density polyethylene are similar to those made from low-density polyethylene, such as trash cans, food containers, and closures. Sheathing and flexible pipe are applications that take advantage of the flexibility and low-temperature toughness of linear low-density polyethylene.

High-density polyethylene is used in food packaging applications, because it gives very good moisture barrier properties. However, like all polyethylenes, it is not used in applications requiring an oxygen barrier. While high-density polyethylene has good chemical resistance, it is prone to environmental stress cracking in the presence of detergents and other surfactants. Longer-chain comonomers and

higher-molecular-weight resins improve the environmental stress-crack resistance of high-density polyethylene. This is an important consideration in applications where polyethylene contacts chemicals, such as dishwashing detergent and laundry bleach. Because of the higher crystallinity, high-density polyethylene is used in applications where higher stiffness is required, such as injection-molded beverage cases, thin-walled blow-molded milk bottles, grocery bags, and thermoformed sheet for pickup truck bedliners. Large blow-molded applications, which take advantage of the higher rigidity high-density polyethylene, include large industrial containers. Fibers, having very high tensile strengths and tensile moduli, have been produced by a process known as gel spinning from very high molecular weight high-density polyethylene. In gel spinning, very high molecular weight polyethylene that has been swollen with solvent is spun into an extremely high tenacity fiber. This fiber approaches the strength of spider silk, and is used in medical and other applications where a very high strength fiber is desirable, for example, in fishing line. *See* SPIDER SILK.

Ethylene copolymers. Copolymers of polyethylene are generally limited in comonomer content to preserve some load-bearing properties in the form of enough stiffness to retain their shape. Higher levels of comonomer form rubbery polymers. Nearly all of these rubbery ethylene copolymers are used in compounded mixtures, which include the impact modification of rigid polymers such as polyethylene and polypropylene. These materials generally range in comonomer content from 25 to 60% by weight, with propylene being the most widely used comonomer to form ethylene-propylene rubber. In addition to propylene, small amounts of a diene are sometimes included, forming a terpolymer, that is, a copolymer in which three monomers are used. The distinguishing attribute of this terpolymer is that it can be vulcanized by means of peroxides. Compounded products containing ethylene-propylene rubber and terpolymer have many automotive uses such as in bumpers, dashboard panels, steering wheels, and assorted interior trim.

Another major use of ethylene copolymers is to produce poly(ethylene-co-vinyl acetate). The properties of this polymer are governed by the percentages of ethylene and vinyl acetate. Poly(ethylene-co-vinyl acetate) is polymerized by using a high-pressure process similar to that used in producing low-density polyethylene. As with all polyethylenes, increasing the amount of vinyl acetate comonomer decreases crystallinity, reduces stiffness, increases clarity, and reduces the melting point. Chemical resistance is similar to low-density polyethylene, while the clarity, low-temperature flexibility, and impact strength are superior. These materials are more permeable to oxygen, carbon dioxide, and water vapor than low-density polyethylene. Applications include specialty film for heat sealing, adhesives, flexible hose and tubing, footwear components, bumper components, and gaskets. Foamed and cross-linked poly(ethylene-co-vinyl acetate) is used in energy-absorbing applications.

An interesting category of ethylene copolymers is ionomers. These copolymers are produced from the copolymerization of ethylene with a comonomer containing a carboxylic group (COOH). Most often, the comonomer is methyl acrylic acid. When this copolymer is reacted with a salt of sodium or zinc, a polymer with very high tensile strength is obtained having outstanding toughness and abrasion resistance. Ionomers exhibit good adhesion to metal, glass, and many fibers. They are melt-processible, with the ionic bonds acting like cross-links at lower temperatures, yet allowing flow at higher temperatures. Because of toughness, ionomers are widely used in the covers for golf balls. Bruce Bersted

Polypropylene. Commercial polypropylene homopolymers are high-molecular-weight, semicrystalline solids having melting points around 160–165°C (320–329°F), low density (0.90–0.91 g/cm^3), excellent stiffness and tensile strength, and moderate impact strength (toughness). Incorporating small amounts of ethylene or higher alpha-olefins in the isotactic polypropylene structure produces random copolymer resins with lower melting points and densities, reduced stiffness and tensile strengths, but higher impact strengths and clarity than homopolymer polypropylene resins. Intimate blending of homopolymer polypropylene with ethylene-propylene rubber produces impact copolymers that possess both high strength and excellent toughness.

Polypropylene was discovered in the 1950s and is now the second most important thermoplastic material, after polyethylene. The attractive combination of low density, excellent mechanical properties over a wide temperature range, low electrical conductivity, and low cost leads to the widespread use of polypropylene homopolymers and random and impact copolymers in products such as automotive parts, appliances, battery cases, carpeting, electrical insulation, fiber and fabrics, film wraps, food packaging and storage containers, clothing, geotextiles, luggage, medical equipment, rope, and tamper-resistant closures.

Production. Propylene, $CH_3CH{=}CH_2$, is produced with ethylene in large quantities at low cost from the cracking of oil and other hydrocarbon feedstocks. It can also be made by propane dehydrogenation. *See* PETROLEUM PRODUCTS.

The polymerization of propylene can yield the three basic types of polypropylene materials—isotactic, syndiotactic, and atactic—which differ by the orientation of the methyl groups (CH_3) along the polymer chain (**Fig. 2**). Idealized isotactic polypropylene has the same methyl group configuration at each tertiary carbon atom along the polymer backbone, while in perfectly syndiotactic polypropylene the methyl group configuration at the tertiary carbon positions alternate along the chain. Atactic polypropylene possesses a random orientation of methyl groups. These variations in polymer structure produce polypropylenes with widely

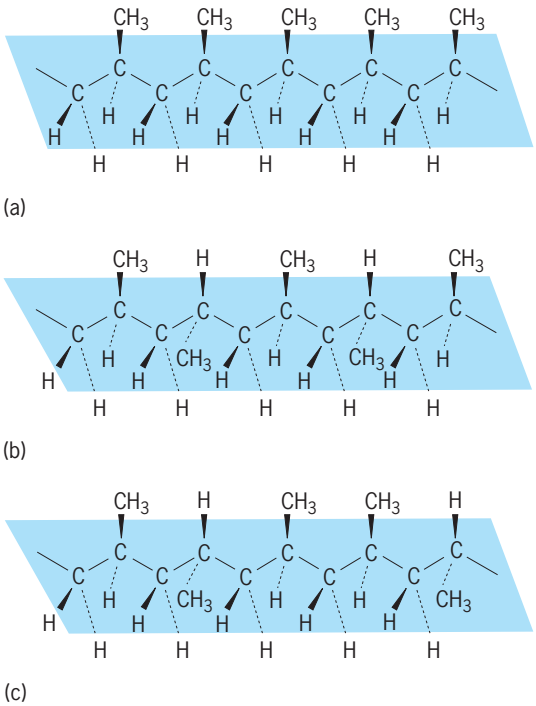

Fig. 2. Idealized structure of polypropylene materials. (*a*) Isotactic; methyl group always on same side of chain. (*b*) Syndiotactic; methyl group position alternating. (*c*) Atactic; methyl groups randomly distributed.

different properties. The stereoregular structures of the isotactic and syndiotactic polypropylene chains result in high-melting, crystalline solids, whereas atactic polypropylene is an amorphous, frequently sticky material. In isotactic and syndiotactic polypropylenes, occasional errors in the orientation of the methyl groups disrupt the crystallization of these polymers, lowering the melting point and decreasing stiffness. The choice of catalyst and the polymerization conditions determine the type of polypropylene produced (syndiotactic, isotactic, or atactic) and the number of defects contained in the polymer. Nearly all of the polypropylene resins that are commercially produced possess the isotactic chain structure, although a small amount of atactic polypropylene is manufactured for use in adhesives and roofing materials.

Sterospecific, titanium-based catalysts are used to produce most isotactic polypropylenes. Titanium trichloride catalysts, used with alkylaluminum cocatalysts, were the earliest and the simplest catalysts, yielding the desired isotactic polypropylene as well as a significant amount of the atactic form as a by-product. Organic compounds containing oxygen, sulfur, or nitrogen were later found to act as catalyst modifiers, increasing the stereoregularity of the isotactic polymer chains while decreasing the amount of atactic polypropylene produced. The most advanced titanium-based catalysts contain titanium tetrachloride and an organic diester modifier supported on magnesium chloride. These titanium/magnesium/diester catalysts, used with an additional silane-ether modifier and trialkylaluminum cocatalyst, produce highly isotactic polypropylene in very high yields with virtually no atactic polypropylene by-product.

Metallocene and other single-site catalysts have been developed for the stereospecific polymerization of propylene. These single-site catalysts are complexes of zirconium and hafnium that have very precise molecular structures. The ability to tailor the structures of these catalysts has allowed the structure and properties of the resulting polymers to be precisely engineered. Single-site catalysts are now used commercially to produce isotactic polypropylene resins having very narrow molecular-weight distributions, very high melt flow rates, and elastomeric properties.

Polypropylene is produced in a variety of slurry-phase and gas-phase processes. Slurry-phase polymerizations are done in hydrocarbon diluents, such as hexane or heptane, or in liquid propylene (also called a bulk slurry process). Gas-phase polymerizations are done in stirred-bed and fluid-bed processes. Hydrogen is used as a chain transfer agent to control the molecular weight of the polymer. More than one reactor in series may be required to produce polymer blends such as impact copolymers. The high-activity titanium/magnesium/diester catalysts are used in both slurry and gas-phase processes, and their development has allowed many processes to be substantially simplified. Single-site catalysts also may be used in the simplified slurry and gas-phase processes. In addition, the advent of the single-site catalyst has led to the commercialization of a high-temperature solution process for producing polypropylene.

Applications. Polypropylene homopolymer, random copolymer, and impact copolymer resins are tailored to fit specific polymer applications and fabrication methods and to achieve the desired product performance. Low-molecular-weight resins, used for melt spun and melt blown fibers and for injection-molding applications, are produced by single-site catalysts or by oxidative degradation of higher-molecular-weight polymers at elevated temperatures. The latter materials, often called controlled rheology resins, have narrower molecular-weight distributions and lower viscoelasticity. The brittleness of polypropylene homopolymer, particularly at temperatures below 0°C (32°F), is greatly reduced by blending it with ethylene-propylene rubber. Compounding with mineral fillers and glass fibers improves product stiffness and other properties. Higher-stiffness resins also are produced by increasing the stereoregularity of the polymer or by adding nucleating agents. Polypropylene resins are used in extrusion and blow-molding processes and to make cast, slit, and oriented films. Stabilizers are added to polypropylene to protect it from attack by oxygen, ultraviolet light, and thermal degradation; other additives improve resin clarity, flame retardancy, or radiation resistance. *See* RHEOLOGY; STABILIZER (CHEMISTRY).

Other poly(alpha-olefins). Isotactic polymers of 1-butene and other higher alpha-olefins are produced by using the same stereospecific catalyst systems

used for polypropylene. Poly(1-butene) is a tough and flexible resin that crystallizes from the melt in a metastable form (melting point 130°C or 266°F) before transforming to another, more stable crystalline form (melting point 138°C or 280°F). This polymer has been used in the manufacture of film and pipe. Poly(4-methyl-1-pentene) is a transparent polymer with a high melting point (240°C or 464°F) and exceedingly low density (0.83 g/cm^3). This material is clear and can be autoclaved, and is used in manufacture of chemical and medical equipment. Other poly(alpha-olefins) have been synthesized, including poly(1-pentene), poly(3-methyl-1-butene), and poly(vinylcyclohexane), but they have attracted little or no commercial interest.

Polyisobutylene. High-molecular-weight rubbery solids are produced when isobutylene is polymerized at low temperature via carbocationic intermediates using a Lewis acid catalyst such as aluminum trichloride. Copolymerization of isobutylene with around 3% isoprene produces butyl rubber. Chlorination or bromination of butyl rubber is used to produce halogenated butyl rubbers. All three materials can be vulcanized and have low permeability to gases, making them useful as sealants, inner tubes, and tubeless tire liners. Low-molecular-weight polyisobutylenes are prepared near 0°C (32°F) and have molecular weights of 200–3000. They are used in formulations for caulking, sealants, and lubricants. *See* RUBBER.

Polybutadiene and polyisoprene. Butadiene and isoprene can be polymerized by anionic, free-radical, and transition-metal catalysts to give a number of polymer structures. The commercially important forms of polybutadiene and polyisoprene are similar to the *cis*-1,4-poly(diene) structures possessed by natural rubber, as shown in the structure below.

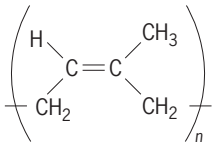

Synthetic *cis*-1,4-polyisoprene is produced by alkyllithium catalysts, and high *cis*-1,4-polybutadiene (that is, mostly *cis* double bonds) is obtained by using a catalyst containing titanium, iodine, and trialkylaluminum. The 1,4-polybutadiene produced with lithium catalysts contains roughly equal amounts of *cis* and trans structures, but it is suitable for a variety of compounding applications. Elastomeric copolymers of these dienes with styrene and acrylonitrile can also be produced. *See* POLYACRYLONITRILE RESINS; POLYMER; POLYSTYRENE RESIN.

Steven A. Cohen

Bibliography. B. H. Bersted, On the effects of very low levels of long chain branching on rheological behavior in polyethylene, *J. Appl. Polym. Sci.*, 30:3751–3765, 1985; J. Boor, Jr., *Ziegler-Natta Catalysts and Polymerizations*, 1979; J. Brandrup and E. H. Immergut (eds.), *Polymer Handbook: Polymerization and Depolymerization*, vol. 1, 4th ed., 1999; J. D. Chadwick, Ziegler-Natta catalysts, in *Encyclopedia of Polymer Science and Technology*, vol. 8, 3d ed., pp. 517–536, 2003; N. P. Cheremisinoff (ed.), *Handbook of Polymer Science and Technology*, vol. 1: *Synthesis and Properties*, 1989; Y. V. Kissin and R. B. Lieberman, Olefin polymers, in *Kirk-Othmer Encyclopedia of Chemical Technology*, vol. 17, 4th ed., pp. 702–839, 1996; R. Leino, Single-site catalysts, in *Encyclopedia of Polymer Science and Technology*, vol. 4, 3d ed., pp. 136–179, 2003; R. Lieberman, Propylene polymers, in *Encyclopedia of Polymer Science and Technology*, vol. 11, 3d ed., pp. 287–358, 2004; I. I. Rubin, *Handbook of Plastic Materials and Technology*, 1990; J. Scheirs and W. Kaminsky (eds.), *Metallocene Based Polyolefins*, 2000; J. L. Throne, *Plastics Process Engineering*, 1979; R. Wehrenberg, New organic fibers, *Plast. World*, pp. 60–62, October 1985.

Polyoma virus

A papovavirus (member of a DNA-containing group of animal viruses) that infects rodents. This virus can induce a wide variety of tumors when inoculated into newborn animals. The simple structure has made the virus a popular model for the study of lytic infection (a viral infection that results in the lysis of the host cell and release of viral progeny) by animal viruses, and viral-induced tumorigenesis (tumor formation). As a result, much is known of the biology of this virus. Discoveries made through studying its action have influenced the work on most other cancers. *See* TUMOR VIRUSES.

Polyoma virus is endemic in most wild populations of mice but causes few harmful effects. Tumors produced by this virus are unknown in the wild. Inoculation of large quantities of virus into newborn rodents, however, induces a variety of tumor types, particularly sarcomas and carcinomas. These tumors contain the viral genome but produce few infectious viral particles. Infected animals produce neutralizing antibodies directed against structural components of the virus; tumor-bearing animals produce antibodies to antigens (T antigens) which are present in tumor cells but not in the virus particle. *See* ANTIBODY; ANTIGEN.

The icosahedral viral particle consists of DNA and protein only. The genome is a small, double-stranded, closed circular DNA molecule approximately 5300 base pairs in length. It encodes the T antigens expressed early in the productive cycle and in transformed cells, and the viral capsid proteins, expressed late in the productive cycle. *See* DEOXYRIBONUCLEIC ACID (DNA).

The virus is easily propagated to high titer in mouse embryo tissue culture, resulting in cell lysis and the production of a hemagglutinin (a red blood cell-clumping antibody). This "productive infection" occurs only in mouse cells. Polyoma virus will infect cells of other rodents, but little viral proliferation occurs. Small numbers of "abortively infected" cells retain the viral genome and become transformed.

Transformed cells are tumorigenic when injected into syngeneic (having the same genotype and thus being genetically compatible for transplant purposes) animals, and contain the T antigens but produce no virus. The expression of the T antigens has been shown to be necessary and sufficient to induce cell transformation.

During lytic infection, the three T antigens (large, middle, and small T) are required for initiating replication of the viral genome through an interaction between the large T antigen (LT) and viral DNA, and for creating a suitable environment for high levels of DNA synthesis to occur (all three T antigens). Only one enzymatic activity has been shown to be a property of the T antigens, a DNA unwinding action of the LT molecule involved in initiating DNA synthesis. All the other effects of the T antigens on cells are caused by their ability to bind to, and change the activities of, cellular proteins. Major binding proteins to LT include members of the retinoblastoma associated protein (p105Rb) family, and a number of proteins involved in control of RNA transcription. Middle and small T antigens interact with the core heterodimer of protein phosphatase 2A. Middle T antigen (MT) alone also binds members of the SRC [the sarcoma-inducing (transforming) gene of Rous sarcoma virus] family of tyrosine kinases, type I phosphatidyl inositol 3′ hydroxyl kinase, and the adaptor proteins ShcA and Grb2. These MT binding proteins are normally involved in stimulating cell growth in response to extracellular signals. *See* IMMUNOLOGY; ONCOLOGY; VIRUS. S. M. Dilworth

Bibliography. T. L. Benjamin, Polyoma virus: Old findings and new challenges, *Virology*, 289:167–173, 2001; S. M. Dilworth, Polyoma virus middle T antigen and its role in identifying cancer-related molecules, *Nat. Rev. Cancer*, 2:951–956, 2002; J. L. Sever and D. L. Madden (eds.), *Polyomaviruses and Human Neurological Disease*, 1985; J. Tooze (ed.), *DNA Tumor Viruses*, 2d ed., 1981.

Polyplacophora

A class within the phylum Mollusca whose members are popularly called chitons, coat-of-mail shells, or sea cradles. Polyplacophora, or Loricata, was formerly combined with Aplacophora in the class Amphineura. All chitons are marine, and they typically live in the intertidal zone, although some live in deeper waters. They are found from subarctic to tropical latitudes but are most abundant in warmer waters. There are roughly 750 chiton species living today.

Polyplacophorans exhibit bilateral symmetry and are oval or elongate in outline. Chitons vary in length from a few millimeters ($^1/_{16}$ in.) to over 30 cm (12 in.), though most are a few centimeters (about 1 in.) long. They are flattened dorsoventrally and bear eight shell plates on their back. The plates, termed valves, are formed from crystals of the mineral aragonite (a type of calcium carbonate), and the upper layer of the valves contains canals filled with sensory organs called esthetes. The larger of these structures serve as photoreceptors and hence act as eyes.

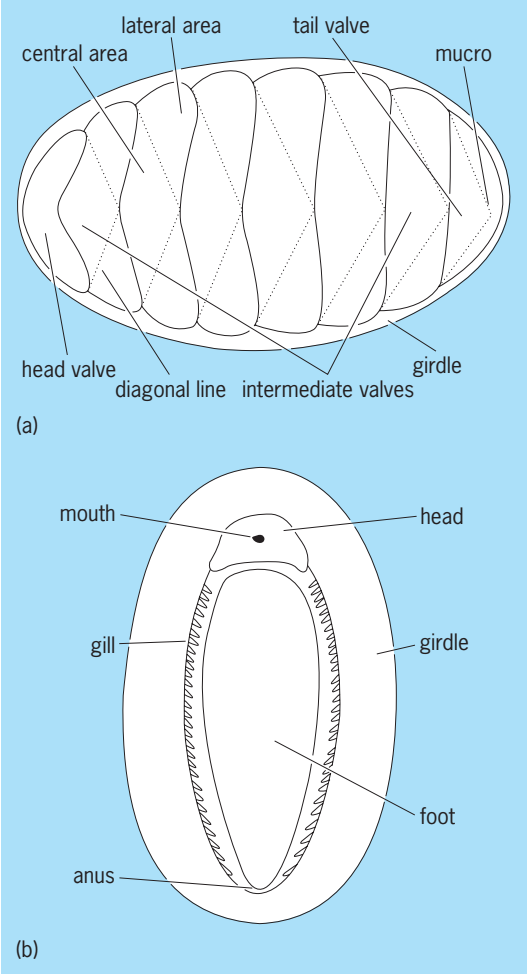

Polyplacophoran mollusk. (*a*) Dorsal view. (*b*) Ventral view.

Chitons possess three valve types—head, intermediate, and tail (see **illus.**). The valves are laterally embedded in the fleshy but tough girdle, which typically bears small aragonite spicules or scales of varying shapes. The width of the girdle differs from species to species, in some cases completely covering the valves (such as in the large Pacific gumboot chiton, *Cryptochiton*), but in most species covering only a small lateral region of the valves. Ventrally, polyplacophorans possess a large foot which they use for moving about and for attaching to rocks in a manner similar to a limpet gastropod. This allows the chitons to cling tenaciously to the substrate. The gills hang along both sides of the animal, in a deep groove between the foot and the shell. The anus is located near the posterior end of the foot. At the front of the foot is the head which, although without tentacles or eyes, bears the mouth with its associated rasping structure (the radula, which is characteristic of most mollusks). A unique feature of the chiton radula is that the largest teeth are hardened with magnetite, an iron mineral.

Most chitons feed by scraping algae and diatoms from the rock surface with their mineral-hardened

radula. Others feed on small crustaceans which they trap with their foot (*Placiphorella* and *Craspedochiton*), or on wood (*Xylochiton* and *Ferreiraella*). Chitons are usually mobile, although some, such as the California troglodyte chiton *Nuttalina fluxa*, remain within small depressions they have carved out of the rock. Chitons typically spend most of the daytime hidden in crevices or under rocks. They forage when submerged under water at night, returning to inactivity, and often to the same home location, by the time the tide recedes or dawn comes.

Most chiton species have separate sexes and spawn eggs and sperm. The eggs develop initially into a free-swimming form which looks like the trochophore larva characteristic of most annelids. Unlike clams and snails, chiton larvae do not pass through a veliger stage. More than 30 chiton species brood embryos until the young are sufficiently developed to remain attached to the substrate.

Although the fossil record of chitons is sparse, isolated valves are known from sediments as old as the Late Cambrian (about 500 million years ago). The radula is known from only one fossil species (*Glaphurochiton*), and girdle spicules are known from only a few fossils. Otherwise, the chiton fossil record consists of isolated or articulated shell plates. Chitons were affected by both the Permian-Triassic and the Cretaceous-Tertiary extinctions, when the diversity of this group dropped considerably. Modern chitons (the crown group) first appeared in the late Paleozoic (around 320 million years ago).

Chitons are usually allied with the aplacophorans on the basis of similarities in the nervous system, in the possession of spicules, and in a similar overall shape. One main difference between these two classes is that chitons bear a shell whereas aplacophorans do not. Determining the relationship between the shell of the chiton and that of other mollusks has not been straightforward. Although the general structure of the chiton shell is similar to that of other shell-bearing mollusks, there are some features unique to the chiton shell. These include the presence of eight valves and the perforation of the upper shell layer with esthete-containing canals. Characters of the valves, the gills, and the radula typically have been used to differentiate major groups within the Polyplacophora. *See* APLACOPHORA; LIMPET; MOLLUSCA; MONOPLACOPHORA. Michael Vendrasco

Bibliography. R. C. Brusca and G. J. Brusca, *Invertebrates*, 2d ed., pp. 701–770, Sinauer Associates, Sunderland, MA, 2002; G. Burghardt and L. Burghardt, *A Collector's Guide to West Coast Chitons*, San Francisco Aquarium Soc. Spec. Publ., no. 4, 1969; A. Solem, *The Shell Makers: Introducing Mollusks*, pp. 48–55, Wiley, New York, 1974.

Polyploidy

The occurrence of related forms possessing chromosome numbers which are multiples of a basic number (n), the haploid number. Forms having $3n$ chromosomes are triploids; $4n$, tetraploids; $5n$, pentaploids, and so on. Autopolyploids are forms derived by the multiplication of chromosomes from a single diploid organism. As a result the homologous chromosomes come from the same source. These are distinguished from allopolyploids, which are forms derived from a hybrid between two diploid organisms. As a result, the homologous chromosomes come from different sources. About one-third of the species of vascular plants have originated at least partly by polyploidy, and as many more appear to have ancestries which involve ancient occurrences of polyploidy. The condition can be induced artificially with the drug colchicine and the production of polyploid individuals has become a valuable tool for plant breeding.

Polyploid series are irregularly distributed through the plant kingdom. They are particularly common in some families, notably the Gramineae and Rosaceae, and are rare in others, such as the Fagaceae and Umbelliferae. Not infrequently there exist related genera of the same family, one with and one without polyploidy, such as *Thalictrum* and *Aquilegia* of the family Ranunculaceae, or *Salix* and *Populus* of the family Salicaceae. Polyploidy has a significantly higher frequency in perennial herbs than in annual herbs and woody plants. A polyploid series often cited is that of the wheats, in which the basic chromosome number is 7 and somatic chromosome numbers of 14, 28, and 42 occur.

In animals, most examples of polyploidy occur in groups which are parthenogenetic, such as crustaceans of the genus *Artemia*, certain earthworms, weevils of the family Curculionidae, moths of the genus *Solenobia*, sawflies of the genus *Diprion*, lizards of the genus *Cnemidophorus*, and fishes of the genera *Carassius* and *Poecilopsis*; or in species which reproduce asexually by fission, such as the flatworm *Dendrocoelum infernale*. Sexual polyploids have been found in the frog genera *Odontophrynus*, *Ceratophrys*, and *Phyllomedusa*, as indicated both by chromosome numbers and nuclear DNA content. In many animals the sex chromosome mechanism is so upset by polyploidy that sterile intersexes are produced. Because hybrid sterility in animals is usually developmental rather than segregational in nature and is not eliminated by chromosome doubling, allopolyploids can occur only rarely. Developmental hybrid sterility, as in the mule, is the result of genes, contributed by the parents, interacting disharmoniously in the hybrid to disturb gonad development, sex-cell differentiation, and the course of meiosis, irrespective of chromosome pairing. Segregational hybrid sterility is the result of nonhomology between parental chromosomes reflected by their inability or reduced ability to pair at meiosis, or of abnormal segregation of genes or chromosomal segments from meiotic pairs formed by chromosomes differing with respect to inversions, translocations, or other rearrangements.

Hybridization. Polyploidy and hybridization are usually associated with each other in evolution. When distantly related species are crossed, the sterile F_1 hybrid often has little or no meiotic chromosome

pairing. However, in some cells the chromosome complement undergoes a doubling, leading to the formation of a few seeds from which second generation hybrids are formed. This fertile polyploid derivative possesses only bivalents, so that it breeds true for the intermediate condition. Such a polyploid is designated allopolyploid. Well-known examples are bread wheat (*Triticum aestivum*, $2n = 42$), cultivated tobacco (*Nicotiana tabacum*, $2n = 48$), and *Raphanobrassica* ($2n = 36$), which is the hybrid between radish (*Raphanus sativus*, $2n = 18$) and cabbage (*Brassica oleracea*, $2n = 18$). Polyploids have also originated from hybrids between closely related species (*Primula kewensis*), or between subspecies of the same species (*Dactylis glomerata*). These are designated as segmental allopolyploids and autopolyploids of hybrid origin, respectively. Such polyploids may form varying numbers of trivalent and quadrivalent chromosome configurations and so may segregate in the direction of their parental types. Autopolyploids derived from a single diploid race have been produced many times artificially but are rare in nature since they are usually weaker and less fertile than their diploid progenitor. In many plant genera the simultaneous occurrence of several polyploids of these three types, plus others which combine characteristics of auto- and allopolyploidy, along with their diploid ancestors, has produced the polyploid complex. This consists of a series of usually distinct diploid species or subspecies, which represent extremes of morphological and ecological variation, plus a much larger number of polyploids, which form a network of connecting links between the diploids. Examples are *Bromus*, *Vaccinium*, *Galium*, and *Antennaria*.

Distribution. The geographic distribution of polyploids in relation to their diploid ancestors does not follow any consistent set of rules. Certain northern regions, such as Iceland and Spitsbergen, have floras with particularly high percentages of polyploidy, but so do some regions with subtropical or even tropical climates, such as New Zealand and Sri Lanka. Lowland and adjacent high alpine floras in western Europe do not differ significantly in percentage of polyploids. Diploid representatives of individual polyploid complexes tend to occupy geologically older habitats, and tetraploids are most prevalent in regions newly open to colonization. Consequently, the larger, older, and more stable landmasses have relatively high percentage of diploids, while islands with floras derived through immigration, and areas disturbed by glaciation, volcanic activity, or other causes have high percentages of polyploidy.

Plant breeding. Artificial polyploids have been produced in most of the major species of crop plants by treating seeds or cuttings with colchicine. Generally, they have been less useful than their diploid progenitors because of their slower growth and reduced fertility, but economically valuable autopolyploids have been produced in rye, sugarbeets, rapeseed oil, red clover, snapdragons, marigolds, various orchids, and some other plants. Success has been achieved only when the doubling has been accompanied by intervarietal hybridization and selection. Although several artificial allopolyploids such as wheat-rye and wheat-*Agropyron* or quack grass have achieved partial success, none has yet been grown on a commercial scale. Artificial autopolyploids have also made possible the transfer of genes for disease resistance from wild species to cultivated species, even in instances in which the species are so distantly related that the F_1 hybrids between their normal diploid forms are completely sterile. Examples are the transfer of rust resistance from goat grass (*Aegilops umbellulata*) to bread wheat and of resistance to both tobacco mosaic and black shank diseases from wild species of *Nicotiana* (*N. glutinosa*, *N. plumbaginifolia*) to cultivated tobacco. Polyploidy is, therefore, a useful tool in plant breeding when combined with hybridization and selection. *See* BREEDING (PLANT); CHROMOSOME ABERRATION; GENE; GENETICS; PLANT EVOLUTION; SPECIATION.

G. Ledyard Stebbins

Polyploid and polytene tissues. In addition to polyploid organisms in which all of the body cells contain multiples of the basic chromosome number, most plants and animals contain particular tissues that are polyploid or polytene. Both polyploid and polytene cells contain extra copies of the deoxyribonucleic acid (DNA), but they differ in the physical appearance of the chromosomes. In polytene cells the replicated copies of the DNA remain physically associated to produce giant chromosomes that are continuously visible and have a banded pattern. The term polyploid has been applied to several types of cells: multinucleate cells; cells in which the chromosomes cyclically condense but do not undergo nuclear or cellular division (this process is termed endomitosis); and cells in which the chromosomes appear to be continually in interphase, yet the replicated chromosomes are not associated in visible polytene chromosomes. *See* CHROMOSOME.

There are many examples of polyploid and polytene tissues. Binucleate cells are prevalent in vertebrates, for example in liver cells. The megakarocytes of vertebrates can attain ploidy levels of $64n$. Vertebrates also contain polytene cells such as the trophoblast cells of the mammalian placenta. These cells most commonly are $256–512n$ but can be $4096n$. The suspensor cells of plant embryos are also polytene. The Diptera exhibit an extreme use of polytenization; almost all of the larval tissues become polytene. The *Drosophila* polytene salivary gland chromosomes ($1024–2048n$) have been used extensively for cytologically mapping the position of genes and chromosomal rearrangements. Although the advantages that polyploid and polytene cells provide to the organism have not been tested directly, there is a good correlation between polytene cells and rapidly growing, nutritional organs. The fact that these cells are permanently in interphase may facilitate high metabolic activity, and the large size of the cells may be important in the physical structure of tissues. *See* CHROMOSOME ABERRATION; GENETICS; MITOSIS.

Terry L. Orr-Weaver

Bibliography. V. Y. Brodsky and I. V. Uryvaeva, *Genome Multiplication in Growth and Development: Biology of Polyploid and Polytene Cells*, 1984;

V. Grant, *Plant Speciation*, 2d ed., 1981; W. H. Lewis (ed.), *Polyploidy: Biological Relevance*, 1980.

Polypodiales

The largest order of modern ferns, commonly called the true ferns, with approximately 250 genera and 9000 species; also known as Filicales. Although well represented in the temperate regions, they reach their greatest development in the moist tropics. They vary in habit from small filmy structures to large tree-like plants (**Fig. 1**). Many are epiphytic (live perched on other plants) and a number are climbing species. A few are aquatic. Perhaps the most striking species are the tropical tree ferns with their upright, unbranched stems and terminal clusters of large graceful leaves. The Polypodiales differ from the other fern orders in being leptosporangiate—that is, their sporangium, or spore sac, arises from a single surface cell—and in having small sporangia with a definite number of spores. The wall of the sporangium

Fig. 1. Representative Polypodiales. (*a*) Polypody fern on a cliff edge (*photograph by C. H. Jones*). (*b*) Largest known specimen of the staghorn fern, an epiphyte on trees in the tropics. Photosynthesis occurs chiefly in the upright leaves. The rounded leaves pressed against the tree cover masses of roots. Reproductive structures develop on the pendant leaves (*photograph made on island near Brisbane, Australia, by C. J. Chamerlain, World Book Co.*). (*c*) Tree fern (*U.S. Forest Service*). (*d*) Cultivated Boston fern and five of its many mutant varieties, all originated from mutations in vegetative cells (*from E. N. Transeau, H. C. Sampson, and L. H. Tiffany, Textbook of Botany, rev. ed., Harper, 1953*).

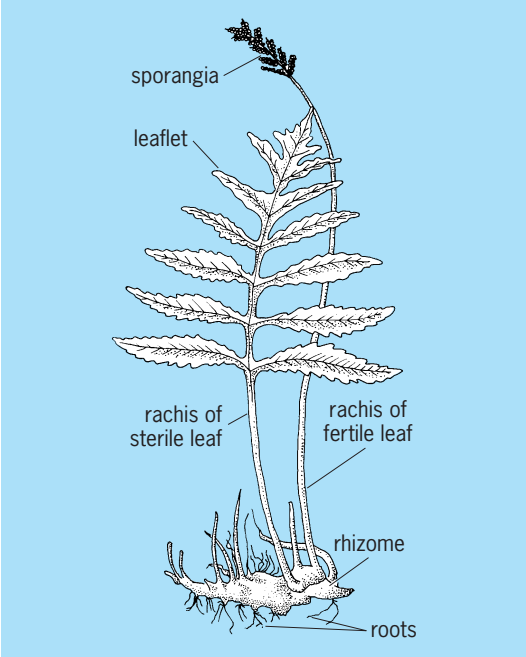

Fig. 2. The sensitive fern (*Onoclea sensibilis*), a representative of the Polypodiopsida. (*After W. W. Robbins, T. E. Weier, and C. R. Stocking, Botany: An Introduction to Plant Science, 3d ed., John Wiley and Sons, 1964*)

is almost encircled with a ring of cells having unevenly thickened walls. This ring is called the annulus. When the sporangium is mature, the annulus, acting as a spring, causes the sporangium wall to rupture, thus discharging the spores. These plants are valued for their beauty and for the clues they give to the evolutionary history of the Polypodiales which extends back through the coal measures of the Paleozoic. *See* PALEOBOTANY.

Structure. The sporophyte is the conspicuous phase of the true ferns, and like other vascular plants it has true roots, stems, and leaves (**Fig. 2**). In most ferns, especially those of the temperate regions, the mature stem is usually a creeping rhizome (underground stem) without aerial branches. However, in several species the stems are branched, and in some they are erect. Whereas in the tropics the leaves are usually persistent and evergreen, in temperate regions the leaves of most species die back each year and are replaced by new ones the next growing season. *See* LEAF; STEM.

Characteristic of this order is the apparent uncoiling of the leaves from the base toward the apex, a phenomenon known as circinate vernation (**Fig. 3**). Few ferns have simple leaves, the common type being the compound pinnately divided leaves known as fronds. The leaflets of the fronds are called pinnae, and if they are further divided, the subdivisions are known as pinnules. The branching of the veins in the leaf is characteristically dichotomous and open.

The branching of the adventitious root is always monopodial (having one main branch). This is usually true also of the stem, although branching occasionally may be dichotomous. Nearly all members of this order display a large apical meristematic cell in both stem and root. *See* APICAL MERISTEM; ROOT (BOTANY).

The internal structure of the blade of the leaf and of the root is very similar to that of these organs in the seed plants. The main difference is the presence of large intercellular spaces in the fern leaf and the frequent lack of apparent distinction between the spongy and palisade cells of the mesophyll, possibly because most ferns grow in the shade. The vascular structure of the ferns includes the four types of stele shown in **Fig. 4**: (1) the protostele, in which the phloem (food-conducting tissue) completely surrounds the xylem (water-conducting tissue), the latter being a solid strand with no central pith; (2) the amphiphloic siphonostele, the xylem of which forms a cylinder enclosing the pitch, with phloem bordering the xylem on both the inside and outside; (3) the ectophloic siphonostele, which has only external phloem surrounding the xylem that is around the pith; and (4) the dictyostele, in which the stele is broken up into a number of vascular bundles either scattered or arranged in a circle. The dictyostele is the common type, with the amphiphloic siphonostele occurring rather frequently and the other two types being comparatively rare. No

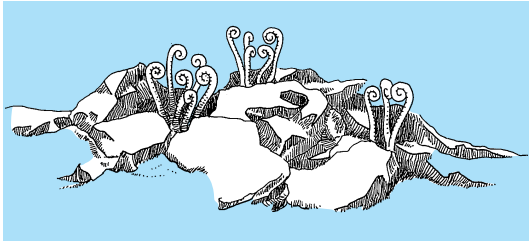

Fig. 3. Fern leaves pushing upward through a cement sidewalk. Growth pressure may amount to hundreds of pounds per square inch. (*After E. N. Transeau, H. C. Sampson, and L. H. Tiffany, Textbook of Botany, rev. ed., Harper, 1953*)

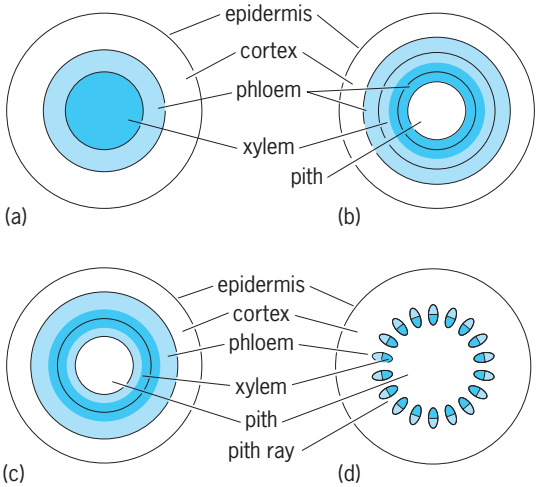

Fig. 4. Types of steles in the Polypodiales. (*a*) Protostele. (*b*) Amphiphloic siphonostele. (*c*) Ectophloic siphonostele. (*d*) Dictyostele with central pith surrounded by xylem and phloem in the form of discrete vascular bundles. Bands of parenchyma cells known as pith rays are located between the bundles. (*After H. J. Fuller and O. Tippo, College Botany, rev. ed., Holt, 1954*)

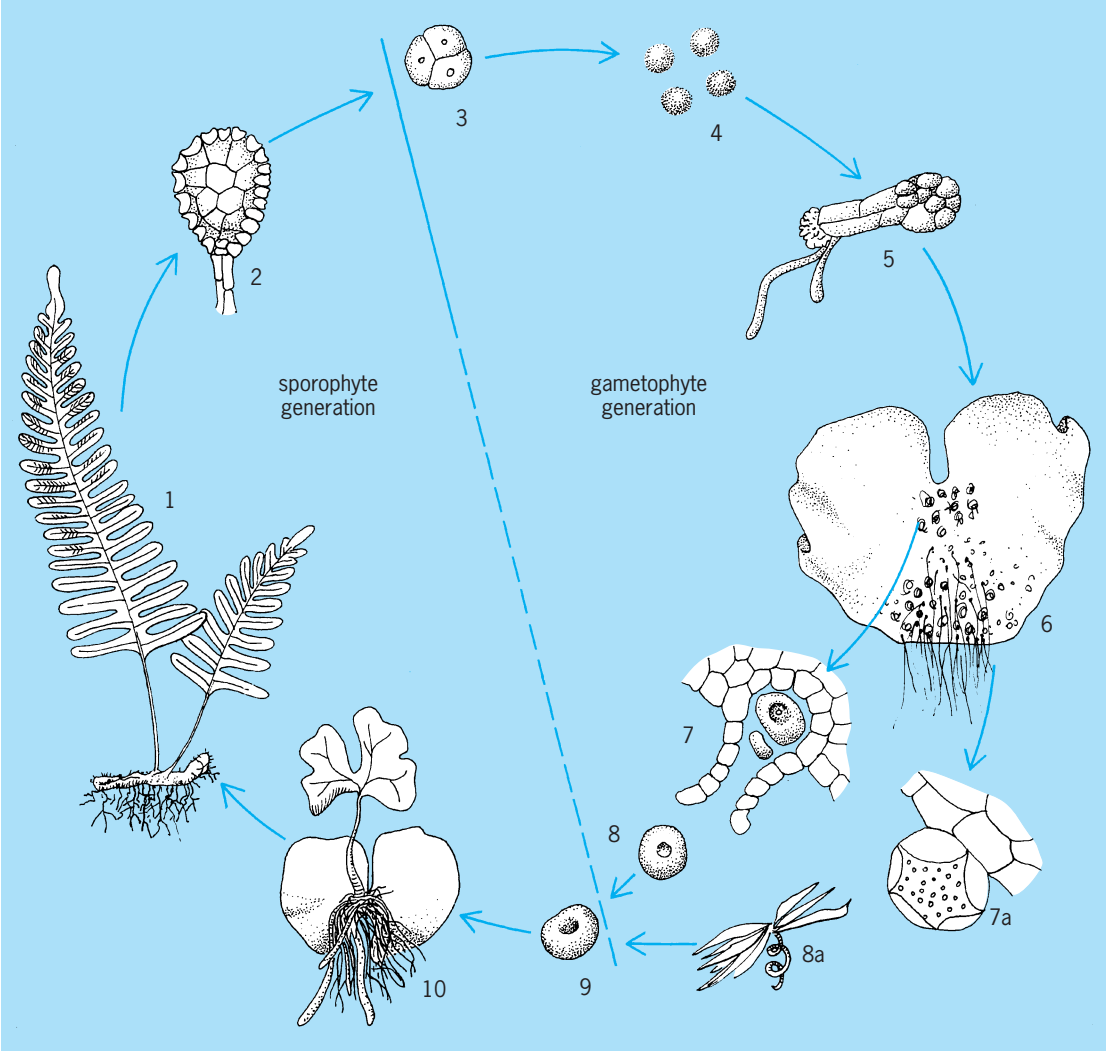

Fig. 5. Life history of a fern: (1) the fern plant, or sporophyte, bearing sori, or clusters of sporangia, on its leaves; (2) a sporangium; (3) a tetrad of young spores; (4) the four mature spores which have come from the tetrad shown in 3; (5) a spore germinating into a young gametophyte; (6) mature gametophyte, bearing sexual organs; (7) archegonium; (7*a*) antheridium; (8) egg cell, or female gamete; (8*a*) sperm or male gamete; (9) fertilized egg; (10) young sporophyte growing out of a fertilized egg, the whole still attached to the remains of the gametophyte. (*After E. W. Sinnott and K. S. Wilson, Botany: Principles and Problems, 6th ed., McGraw-Hill, 1962*)

cambium is present in the true ferns so that all the tissue is of primary origin. The development is mesarch (centripetal and centrifugal) since the metaxylem surrounds the protoxylem. The protoxylem consists of spirally thickened tracheids, the thickenings appearing as bands on the cell walls. Thick-walled sclerenchyma cells are numerous in most fern stems. Except in the few species which have a protostele, leaf gaps are present where the leaf (vascular) traces emerge from the stele. *See* PHLOEM; PITH; PLANT ANATOMY; PLANT TISSUE SYSTEMS; SCLERENCHYMA; XYLEM.

Alternation of generations. The life cycle of the fern consists of two independent (self-sustaining) alternating generations (**Fig. 5**). The common leafy fern plant is the sporophytic (spore-producing) generation. When the mature spores are discharged and reach a suitable substrate, they germinate and produce a small, flat, green, heart-shaped structure known as the prothallium or gametophytic (gamete-producing) generation. The gametophyte produces the sex organs antheridia (male) and archegonia (female). The gametes (sperm and egg) unite in fertilization and the resultant cell, or zygote, develops into the spore-bearing (sporophytic) fern plant.

In all ferns, the spores are produced in special multicellular organs known as sporangia. Except for a few genera, the sporangia are arranged in groups or clusters called sori (**Fig. 6**). These are on the lower (abaxial) surface of the leaves or fertile fronds, either along the midrib of the pinnae, near the leaf margins, or scattered. Usually each sorus is covered by a flap-like structure called the indusium, which may be of various shapes and sizes. However, a few ferns have naked sori, and others have a false indusium formed by the folding or inrolling of the leaf margin. Varying degrees of specialization occur between the vegetative, or sterile, leaf and the reproductive, or fertile, leaf, with most species having leaves able to perform both functions. As a result of meiotic divisions, each

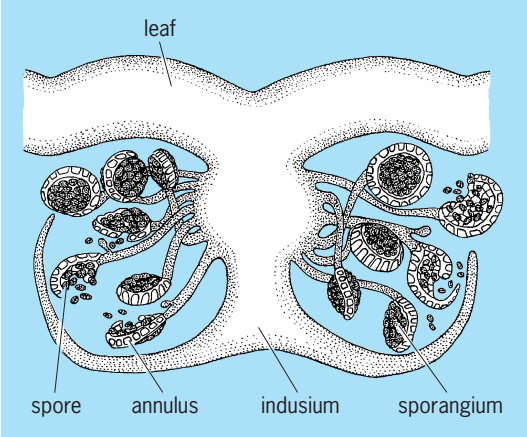

Fig. 6. Diagram of section through a fern leaf, showing details of a sorus. (*After W. W. Robbins, T. E. Weier, and C. R. Stocking, Botany: An Introduction to Plant Science, 3d ed., John Wiley and Sons, 1964*)

sporangium produces a number of tetrads of spores from the spore mother cells. In most species these spores are all alike (homosporous). While the number of spores produced in each sporangium is small, usually 48–64, though sometimes as low as 16 or as high as 512, the total number produced by a single individual may be surprisingly large. For one species of *Dryopteris* it has been estimated that 50 million spores may be produced in a single season. When the spores are mature and conditions are favorable, the annulus ruptures the sporangium, throwing the spores some distance. *See* MEIOSIS.

The spore is the first cell of the gametophytic generation. Each spore may germinate and produce a short filament of cells which eventually develops a triangular apical meristematic cell. By repeated cell division, the apical meristematic cell produces the heart-shaped or lobed prothallium or gametophyte (**Fig. 7**). In most species the gametophyte is monoecious; that is, both male and female sex organs are on the same thallus. While the gametophyte is still quite young, antheridia develop on the lower surface. The archegonia are produced later when the prothallium has matured and developed a thickened archegonial cushion just posterior to the growing point (meristem). Since gametophytes usually occur in various stages of development, it is likely that this difference in maturation time of the cells ensures the

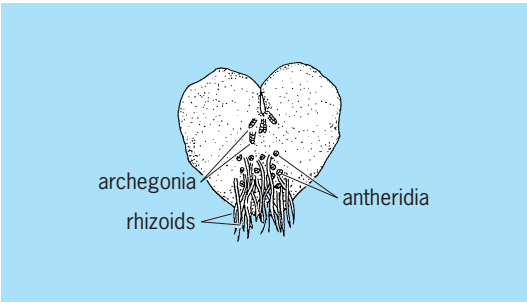

Fig. 7. Diagram of a fern prothallium. (*After W. W. Robbins, T. E. Weier, and C. R. Stocking, Botany: An Introduction to Plant Science, 3d ed., John Wiley and Sons, 1964*)

crossing of gametes from different prothalli or gametophytes. The form of the gametophyte and the number of rhizoids have been found to be strongly influenced by the quality and intensity of the light reaching them. Light seems to be needed for germination of the spores. With strong white or blue light the normal flat thallus with numerous rhizoids is formed, while with low intensity and red light the gametophyte remains filamentous, with few rhizoids formed.

The antheridium is a rounded structure directly attached to the thallus or sometimes borne on a short stalk. Although each antheridium produces only a few sperms, the presence of several antheridia on each prothallus ensures a large number of sperms. Each sperm is large and spirally coiled, and has many cilia.

The archegonia develop in such manner that when mature, the archegonial venter containing the egg is embedded in the prothallium, whereas the neck cells project above the surface. The mature archegonium liberates a substance into the water which elicits a chemotactic response from the sperms, attracting them to the venter and the egg cell therein where fertilization is accomplished and a zygote is formed.

The ferns are characterized by regularity in the development of the embryo. A vertical division of the zygote followed by a transverse division produces a quadrat of cells. Each of these cells then divides to produce a group of eight cells in four pairs. The anterior pair of inner cells gives rise to the underground stem or rhizome, whereas the posterior pair produces the foot, a short-lived absorbing organ. The anterior pair of outer cells forms the first leaf and the other pair produces the primary root. The first leaves usually do not suggest the species, but these are followed by a series of transitional forms until finally the characteristic type of leaf is produced. The production of spores by the adult plant completes the sporophytic generation of the life cycle.

Because of the fact that the ferns have long been of special interest to botanists, a great deal of morphological and paleobotanical evidence has been accumulated which strongly points to their origin from the Psilotopsida. There is also considerable evidence to show that some ancient group of the Polypodiopsida may have given rise to some of the seed plants (Magnoliophyta), as exemplified by the Cycadicae and their modern descendants. *See* POLYPODIOPHYTA; PSILOTOPHYTA.

Paul A. Vestal

Bibliography. H. C. Bold, *Morphology of Plants and Fungi*, 4th ed., 1980; F. Scagel et al., *An Evolutionary Survey of the Plant Kingdom*, 1965; G. M. Smith, *Cryptogamic Botany*, vol. 2, 2d ed., 1955; K. R. Sporne, *The Morphology of Pteridophytes*, 4th ed., 1975.

Polypodiophyta

A division of the plant kingdom, commonly called the ferns, which is widely distributed throughout the world but is most abundant and varied in moist,

tropical regions. The Polypodiophyta are sometimes treated as a class Polypodiopsida of a broadly defined division Tracheophyta (vascular plants). The group consists of five living orders (Ophioglossales, Marattiales, Polypodiales, Marsileales, and Salviniales), plus several orders represented only by Paleozoic fossils. The vast majority of the nearly 10,000 species belong to the single order Polypodiales, sometimes also called Filicales.

Vegetative features. The Polypodiophyta ordinarily have well-developed roots, stems, and leaves that contain xylem and phloem as conducting tissues. The central cylinder of vascular tissue in the stem usually has well-defined parenchymatous leaf gaps where the leaf traces depart from it. The leaves are spirally arranged on the stem and are usually relatively large, with an evidently branching vascular system. In most kinds of ferns the leaves, called fronds, are compound or dissected. *See* LEAF; PHLOEM; ROOT (BOTANY); STEM; XYLEM.

Life cycle. The Polypodiophyta show a well-developed alternation of generations, both the sporophyte and the gametophyte generation being detached and physiologically independent of each other at maturity. The sporophyte is much the more conspicuous, and is the generally recognized fern plant. On some or all of its leaves it produces tiny sporangia which in turn contain spores. *See* REPRODUCTION (PLANT).

The spores are produced by reduction division and represent the first step in the gametophyte generation. They are usually distributed by the wind. The spore gives rise to a small, thalloid, usually green and photosynthetic gametophyte that is seldom as large as the nail of a person's little finger. Sperms and eggs are produced in characteristic male and female sex organs (antheridia and archegonia) of the gametophyte. The egg in the archegonium is fertilized in place by the sperm, which swims through a film of water to reach it.

The fertilized egg is the first step in the sporophyte generation. The young sporophyte begins its growth as a parasite on the parent gametophyte, but it soon establishes its independence by putting forth a leaf and sending a root down into the substrate. The mature sporophyte with roots, stems, and leaves develops from this tiny plantlet produced by the gametophyte.

Evolutionary history. The Polypodiophyta originated from the Rhyniophyta during the Devonian Period. Some of the fossil connecting forms, such as *Protopteridium*, could be referred about as well to the one division as to the other. The Polypodiophyta became abundant and varied in the Carboniferous Period, but many of the fossils which have caused the Carboniferous to be called the age of ferns actually represent "seed ferns," belonging to a different division (Pinophyta). Although somewhat diminished from their Carboniferous abundance, the Polypodiophyta are still an important part of the flora of the Earth. *See* COENOPTERIDALES; MARATTIALES; OPHIOGLOSSALES; POLYPODIALES.

Arthur Cronquist

Polypodiopsida

A class (also known as Filicineae) of the plant division Polypodiophyta containing a large group of plants commonly called ferns. They are widely distributed throughout the world with their greatest development in the moist tropics. *See* POLYPODIOPHYTA.

The stems, which may be unbranched or branched, frequently bear several large, spirally arranged leaves which are usually compound. In the majority of the species, the green leaves, sometimes called fronds, are the conspicuous organs of the sporophyte. The sporangia (spore sacs) are grouped in special structures called sori on the under margin or abaxial (lower) side of the leaves. Most of the sporangia are borne in groups. The fern sporophyte possesses true stems, leaves, and roots. Leaf and branch gaps both are present in the conducting cylinder or stele. Antherozoids are multiflagellate. *See* LEAF; ROOT (BOTANY); STEM.

Polypodiopsida is an old class with a good representation in the Paleozoic flora. Some of the plant fossils resemble contemporary living species. *See* COENOPTERIDALES; MARATTIALES; OPHIOGLOSSALES; PALEOBOTANY; POLYPODIALES.

Paul A. Vestal

Polypteriformes

A distinctive and apparently ancient order of actinopterygian fishes comprising the subclass Cladistia, commonly known as bichirs and reedfishes. The following combination of characters distinguishes the order from all other actinopterygian fishes: elongate body covered with thick, rhombic, ganoid scales; slitlike spiracle behind each eye; well-ossified internal skeleton, including vertebrae with ossified centra and neural canals; maxillae firmly united with the skull; basically heterocercal tail, but with a symmetrical caudal fin, and with the upper part continuous with the dorsal fin; dorsal series of free sail-like finlets, each supported by a spine; lobed pectoral fin with its rays supported by ossified radials; pair of enlarged gular plates; very large intercalar bone; no branchiostegals; bilobed swim bladder (the left lobe less developed than the right) located ventral to and attached to the esophagus, serving as lungs; functional gills; young with a pair of feathery external gills, which are assimilated with age; spiral valve-type intestine; and four rather than five pairs of gill arches as is typical for actinopterygians (see **illustration**). *See* SCALE (ZOOLOGY); SWIM BLADDER.

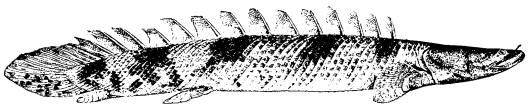

Bichir (*Polypterus endlicheri*), length to 3 ft (0.9 m). (*After G. A. Boulenger, Catalogue of the Fresh Water Fishes of Africa in the British Museum, vol. 1, 1909*)

For many years, Polypteriformes were regarded to be either members of the Sarcopterygii or a taxon more closely related to them than to the Actinopterygii. They are now regarded as the sister group of all other actinopterygians. The order is known from the Eocene and presently comprises a single extant family, Polypteridae, and two genera, *Polypterus*, with about 15 species, and *Erpetoichthys*, with one species. Both genera are confined to freshwaters of tropical Africa. They are common along lake shores and in swamps and are able to withstand anoxic (oxygen-deficient) conditions by breathing atmospheric oxygen. *See* ACTINOPTERYGII; SARCOPTERYGII.

Erpetoichthys. *Erpetoichthys* (synonym *Calamoichthys*) *calabaricus* differs from *Polypterus* in having a snakelike body and no pelvic fins versus an elongate body and pelvic fins. It is endemic to the coastal areas adjacent to the Gulf of Guinea, where it is restricted to reedy habitats in slow-flowing rivers and standing waters. It moves snakewise over the bottom, feeding at night on worms, crustaceans, and insects, but it may also sidewind quite rapidly through the water. The reedfish can breathe air and thus can tolerate low oxygen concentrations. Larvae have external gills and resemble salamander larvae. Maximum total length is 90 cm (36 in.).

Polypterus. Species of *Polypterus* are generally similar in habitat and habits in preferring shallow, vegetated, still or slow-moving waters of streams, flood zones, freshwater lagoons, and lakes. They are primarily carnivorous, with the adults feeding mostly on fishes and the young feeding more on worms, mollusks, insect larvae, and other crustaceans. They may lie quietly or glide snakelike through the water. In spite of their armorlike scale covering, they are very flexible and agile when pursuing prey. The majority of the species occur in the Zaire (Congo) River basin. Maximum total length is 97 cm (38 in.), with most species much smaller.

Herbert Boschung; Reeve M. Bailey

Bibliography. W. E. Burgess, Bichirs and rope fish, *Trop. Fish Hobbyist*, 31(5):14–17, 1983; R. Froese and D. Pauly (eds.), FishBase, World Wide Web electronic publication, version (05/2005); J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; R. Sacca and W. Burggren, Oxygen uptake in air and water in the air-breathing reedfish *Calamoichthys calabaricus*: Role of skin, gills and lungs, *J. Exp. Biol.*, 97:179–186, 1982; G. Sterba, *Freshwater Fishes of the World: A Studio Book*, Viking Press, New York, 1963.

Poly(p-xylylene) resins

Linear, crystallizable resins based on an unusual polymerization of *p*-xylene and derivatives. The polymers are tough and chemically resistant, and may be deposited as adherent coatings by a vacuum process.

Synthesis begins, as in the reaction below, with

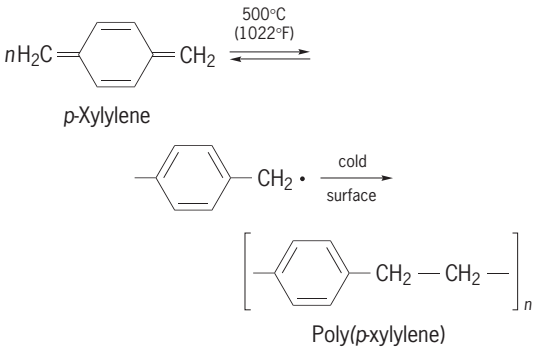

the pyrolysis of xylene to form di-*p*-xylylene. After purification, the di-*p*-xylylene is pyrolyzed to give the monomer *p*-xylylene, a diradical. Condensation of the monomer on cold substrates results in polymerization.

The vapor deposition process makes it possible to coat small microelectronic parts with a thin layer of the polymer. Other potential uses are as coatings for heat-exchange tubes and for encapsulation. Specific properties may be modified by varying the structure of the original xylene, for example, by chlorination. *See* XYLENE.

John A. Manson

Bibliography. G. Odian, *Principles of Polymerization*, 4th ed., 2004; M. P. Stevens, *Polymer Chemistry: An Introduction*, 3d ed., 1998.

Polysaccharide

A class of high-molecular-weight carbohydrates, colloidal complexes, which break down on hydrolysis to monosaccharides containing five or six carbon atoms. The polysaccharides are considered to be polymers in which monosaccharides have been glycosidically joined with the elimination of water. A polysaccharide consisting of hexose monosaccharide units may be represented by the reaction below.

$$nC_6H_{12}O_6 \rightarrow (C_6H_{10}O_5)_n + (n-1)H_2O$$

The term polysaccharide is limited to those polymers which contain 10 or more monosaccharide residues. Polysaccharides such as starch, glycogen, and dextran consist of several thousand D-glucose units. Polymers of relatively low molecular weight, consisting of two to nine monosaccharide residues, are referred to as oligosaccharides. *See* DEXTRAN; GLUCOSE; GLYCOGEN; MONOSACCHARIDE; STARCH.

Polysaccharides are either insoluble in water or, when soluble, form colloidal solutions. They are mostly amorphous substances. However, x-ray analysis indicates that a few of them, such as cellulose and chitin, possess a definite crystalline structure. As a class, polysaccharides are nonfermentable and are nonreducing, except for a trace of reducing power

due, presumably, to the free reducing group at the end of a chain. They are optically active but do not exhibit mutarotation, and are relatively stable in alkali. *See* CELLULOSE; OPTICAL ACTIVITY.

The polysaccharides serve either as reserve nutrients (glycogen and inulin) or as skeletal materials (cellulose and chitin) from which relatively rigid mechanical structures are built. Some polysaccharides, such as certain galactans and mannans, however, serve both functions. Through the action for acids or certain enzymes, the polysaccharides may be degraded to their constituent monosaccharide units. Some polysaccharides yield only simple sugars on hydrolysis; others yield not only sugars but also various sugar derivatives, such as D-glucuronic acid or galacturonic acid (known generally as uronic acids), hexosamines, and even nonsugar compounds such as acetic acid and sulfuric acid.

The constituent units of the polysaccharides molecule are arranged in the form of a long chain, either unbranched as in cellulose and amylose, or branched as in amylopectin and glycogen. The linkage between the monosaccharide units is generally the 1,4- or 1,6-glycosidic bond with either the α or β configuration, as the case may be. The branched glycogen and amylopectin contain both the 1,4 and 1,6 linkages. However, other types of linkage are known. In plant gum and mucilage polysaccharides, 1,2, 1,3, 1,5, and 1,6 linkages occur more commonly than the 1,4 type.

In an attempt to systematize the carbohydrate nomenclature, the generic name glycan was introduced as synonymous with the term polysaccharide. This term is evolved from the generic word glycose, meaning a simple sugar, and the ending, "an," signifying a sugar polymer. Examples of established usage of the "an" ending are xylan for polymers of xylose, mannan for polymers of mannose, and galactomannan for galactose-mannose copolymers. Cellulose and starch are both glucans or glucoglycans, since they are composed of glucose units.

Polysaccharides are often classified on the basis of the number of monosaccharide types present in the molecule. Polysaccharides, such as cellulose or starch, that produce only one monosaccharide type (D-glucose) on complete hydrolysis are termed homopolysaccharides. On the other hand, polysaccharides, such as hyaluronic acid, which produce on hydrolysis more than one monosaccharide type (*N*-acetylglucosamine and D-glucuronic acid) are named heteropolysaccharides. *See* CARBOHYDRATE.

William Z. Hassid

Bibliography. L. Stryer, *Biochemistry*, 4th ed., 1995; G. Zubayet et al., *Principles of Biochemistry*, 3 vols., 1994.

Polystyrene resin

A hard, transparent, glasslike thermoplastic resin. Polystyrene is characterized by excellent electrical insulation properties, relatively high resistance to water, high refractive index, clarity, and low softening temperature.

Styrene produced by the dehydrogenation of ethyl benzene, which in turn is obtained by the alkylation of benzene with ethylene as in reaction (1).

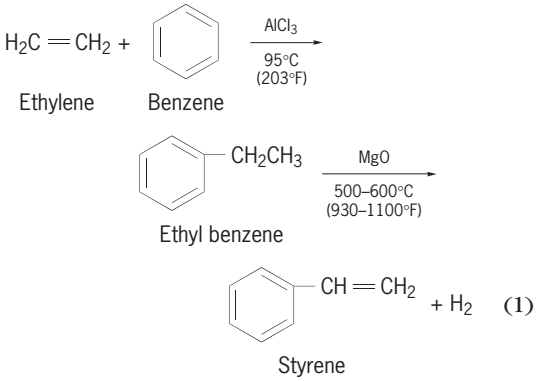

(1)

Free-radical catalysts such as peroxides are often used for polymerization and copolymerization in bulk, solution, and in aqueous emulsion and suspension, as in reaction (2).

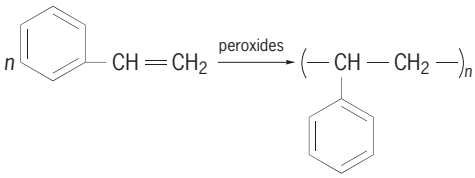

(2)

Ionic and complex catalysts may also be used in bulk and in solution.

The high-molecular-weight homopolymers, copolymers, and polyblends are used as extrusion and molding compounds for packaging, appliance and furniture components, toys, and insulating panels. For a discussion of copolymers of styrene with unsaturated polyesters and with drying oils. *See* POLYESTER RESINS.

The copolymer of styrene and butadiene was the major synthetic rubber of World War II. During the 1940s, the redox system of polymerization was developed in which the presence of a reducing agent caused the peroxide to yield free radicals more rapidly at lower temperatures. At the lower temperature of the redox polymerization, a more linear copolymer called cold rubber is obtained in high conversion with improved physical properties. Styrene-butadiene copolymers are still used for automobile tires and in various rubber articles. *See* POLYMERIZATION; RUBBER.

High-styrene-butadiene copolymers (containing more than 50% styrene) are resinous rather than rubbery. The latexes, as produced by emulsion polymerization, achieved wide usage in water-based paints.

Copolymers with 20–30% acrylonitrile may be used when a somewhat higher softening point is desired without sacrifice of transparency. Fabrication is, however, more difficult.

By sulfonation of the copolymer of styrene and

divinyl benzene, an insoluble polyelectrolyte is produced. This product in the form of its sodium salt is employed as a cationic-exchange resin which is used for water softening.

The effects of blending small amounts of a rubbery polymer, such as butadiene-styrene rubber, with a hard, brittle polymer are most dramatic when the latter is polystyrene. The polyblend may have impact strength greater than ten times that of polystyrene. Even better results may be obtained by graft polymerization of styrene in the presence of the rubber.

Other complex polyblends and interpolymers of styrene, butadiene, and acrylonitrile in various combinations (ABS resins) are important as molding resins which have an excellent balance of properties and low cost. *See* POLYACRYLONITRILE RESINS.

ABS resins and their acrylic counterparts (interpolymers of methyl methacylate, butadiene, and styrene, or MBS resins) are used as toughening agents for polymers such as polyvinyl chloride.

Block copolymers of styrene with butadiene, the so-called thermoplastic elastomers, are of interest as resins for the injection molding of such articles as toys, footwear components, and pharmaceutical items. These copolymers typically consist of rigid polystyrene end units connected by elastomeric polybutadiene units. They may be fabricated by conventional hot thermoplastic techniques, but on cooling behave like typical elastomers.

The strength of amorphous, atactic polystyrene may be increased by orientation induced by uniaxial or biaxial stretching even though crystallinity is not produced. Oriented products in the form of filaments, sheet, and film are available.

Crystallizable, isotactic polystyrene can be formed in the presence of (1) triphenylmethyl potassium in hexane solution, (2) the alfin catalyst (sodium allyl–sodium isopropoxide) in hexane and benzene solutions, and (3) a Ziegler catalyst (titanum tetrachloride–aluminum triethyl) in petroleum ether solvent. The softening temperature of the crystalline polymer is substantially greater than that of the amorphous product.

Besides the many applications of styrene in combination with other materials as in rubber and paints, the homopolymer and the polyblends are employed in the injection molding of toys, panels, and novelty items, and in the extrusion of sheets. The sheets are used for panels, or they may be further shaped by vacuum forming for uses such as liners for refrigerator doors.

Polystyrene may also be fabricated in the form of a rigid foam, which is used in packaging, foodservice articles, and insulating panels. *See* ACRYLONITRILE; COPOLYMER; PLASTICS PROCESSING; POLYMER; STYRENE. John A. Manson

Bibliography. H. Allcock, F. Lampe, and J. Mark, *Contemporary Polymer Chemistry*, 3d ed., 2003; J. A. Brydson, *Plastics Materials*, 7th ed., 1999; G. Odian, *Principles of Polymerization*, 4th ed., 2004; M. P. Stevens, *Polymer Chemistry: An Introduction*, 3d ed., 1998.

Polysulfide resins

Resins that vary in properties from viscous liquids to rubberlike solids. Organic polysulfide resins are prepared by the condensation of organic dihalides with a polysulfide as in the reaction below. By the

$$\underset{\text{1,2,-Dichloro-ethane}}{ClCH_2CH_2Cl} + \underset{\text{Sodium tetrasulfide}}{nNa_2S_4} \longrightarrow \underset{\text{Polysulfide resin}}{\left(-CH_2-CH_2-\overset{S}{\overset{\|}{S}}-\overset{S}{\overset{\|}{S}}-\right)_n} + 2nNaCl$$

use of other dichlorides, such as bis(2-chloroethyl) ether, $ClCH_2CH_2OCH_2CH_2Cl$, by changing the functionality of the sodium sulfide, or by incorporating a trifunctional halide to yield branching, the properties may be varied. Molecular weights can also be adjusted by treatment with chain-breaking disulfides. The condensation is usually conducted in an aqueous medium from which the product may be separated and dried. Many of the polysulfide resins have an odor which is generally characteristic of monomeric sulfur compounds but is usually milder in nature.

The linear, high-molecular-weight polymers can be cross-linked or cured by reaction with zinc oxide. Compounding and fabrication of the rubbery polymers can be handled on conventional rubber machinery. The polysulfide rubbers are distinguished by their resistance to solvents such as gasoline, and to oxygen and ozone. The polymers are relatively impermeable to gases. The products are used to form coatings which are chemically resistant and special rubber articles, such as gasoline bags.

Low-molecular-weight, liquid polymers prepared by cleavage of the —S—S— linkages in the conventional polysulfide are useful as curable ingredients in caulking compounds. In addition, they are used as reactive diluents to reduce the viscosity of epoxy resins and, following reaction with free epoxy groups, to flexibilize and toughen the resin.

The polysulfide rubbers were among the very first commercial synthetic rubbers. Although the products are not as strong as other rubbers, their chemical resistance makes them useful in various applications. The polysulfide rubbers were also among the first polymers to be used in solid-fuel compositions for rockets. *See* ORGANOSULFUR COMPOUND; POLYMERIZATION; PROPELLANT; RUBBER. John A. Manson

Bibliography. J. A. Brydson, *Plastics Materials*, 7th ed., 1999; J. Mark et al. (eds.), *Physical Properties of Polymers*, 3d ed., 2004; G. Odian, *Principles of Polymerization*, 4th ed., 2004.

Polysulfone resins

Plastics whose molecules contain sulfone groups ($—SO_2—$) in the main chain, as well as a variety of additional constituents such as ether or isopropylidene groups. Polysulfones based on aromatic backbones

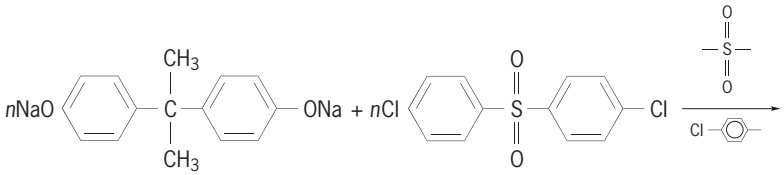

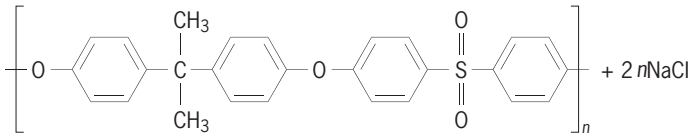

(1)

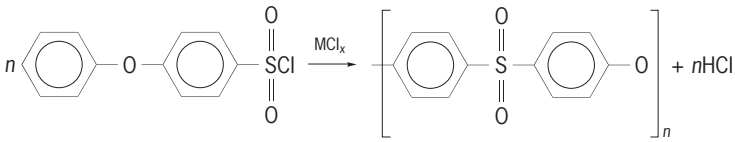

(2)

constitute a useful class of engineering plastics, owing to their high strength, stiffness, and toughness together with high thermal and oxidative stability, low creep, transparency, and the ability to be processed by standard techniques for thermoplastics. Aliphatic polysulfones are less stable to thermal degradation or hydrolysis, so they are not of commercial interest. *See* PLASTICS PROCESSING.

Three major types of polysulfones have been of commercial interest: polysulfone, polyethersulfone, and polyphenylsulfone. A typical synthesis for polysulfone involves the nucleophilic reaction of the disodium salt of bisphenol A with p,p′-dichlorodiphenylsulfone in dimethyl sulfoxide [reaction (**1**)]. Friedel-Crafts-type condensations of chlorosulfonyl goups with aromatic groups can also be used to prepare polyethersulfones [reaction (**2**)], where MCl_x is a metal halide such as $FeCl_3$ or $SbCl_5$, although this is not practiced commercially. Polyether sulfones can also be prepared by the self-condensation of 4-chlorophenyl-4′-hydroxyphenylsulfone. Many variations of these reactions are possible to yield a wide range of structures, including aromatic polysulfones with even higher thermal performance. *See* FRIEDEL-CRAFTS REACTION.

The aromatic structural elements and the presence of sulfone groups are responsible for the resistance to heat and oxidation; ether and isopropylidene groups contribute some chain flexibility. Aromatic polysulfones can be used over wide temperature ranges, to ~150°C (300°F) for polysulfone and to 200°C (390°F) for polyethersulfone and polyphenylsulfone.

Polyphenylsulfone is also known for its supertough impact strength. Resistance to hydrolysis at high temperatures and to most acids, alkalis, and nonpolar organic solvents is excellent; however, the resins may be attacked or dissolved by polar solvents, especially if the component is under stress. While resistance to ionizing radiation is high, protection against ultraviolet light is necessary for outdoor applications. Dimensional stability and electrical properties such as dielectric loss and strength are retained well during service, and flammability is low, especially for polyphenylsulfone. Polysulfones also withstand sterilization and exposure to sterilants better than most other polymers, and polyphenylsulfone can even replace metals in applications requiring repeated sterilization. *See* PHOTODEGRADATION.

Because of the combination of properties discussed, polysulfone resins find many applications in electronic and automotive parts, medical instrumentation subject to sterilization, chemical and food processing equipment, and various plumbing and home appliance items. The inherent low flammability of polyphenylsulfone makes it one of the few materials that can pass the stringent regulations for aircraft interiors. Coating formulations are also available, as well as grades reinforced with glass or carbon fibers. *See* COPOLYMER; HETEROCYCLIC POLYMER; ORGANOSULFUR COMPOUND.

Theodore T. Moore; John A. Manson

Bibliography. J. A. Brydson, *Plastics Materials*, 7th ed., Butterworth-Heinemann, Woburn, MA, 1999; R. J. Cotter, *Engineering Plastics: A Handbook of Polyarylethers*, Gordon and Breach, Basel, Switzerland, 1995; M. J. El-Hibri et al., Poly(arylethersulfones), in O. Olabisi (ed.), *Handbook of Thermoplastics*, Mercel Dekker, 1997; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., vol. 19, Wiley, 1996; V. Lakshmana Rao, Polyether sulfones *J. Macromol. Sci.-Rev. Macromol. Chem. Phys.*, C39:655–711, 1999; F. Parodi, Polusulfones, in *Comprehensive Polymer Science: The Synthesis, Characterization, Reactions, and Applications of Polymers*, vol. 5, ed. by G. C. Eastmond et al., Pergamon Press, Oxford, U.K., 1989.

Polytrichidae

A subclass of the mosses (class Bryopsida), consisting of one family and 19 genera. The plants are acrocarpous and perennial, and the leaf structure

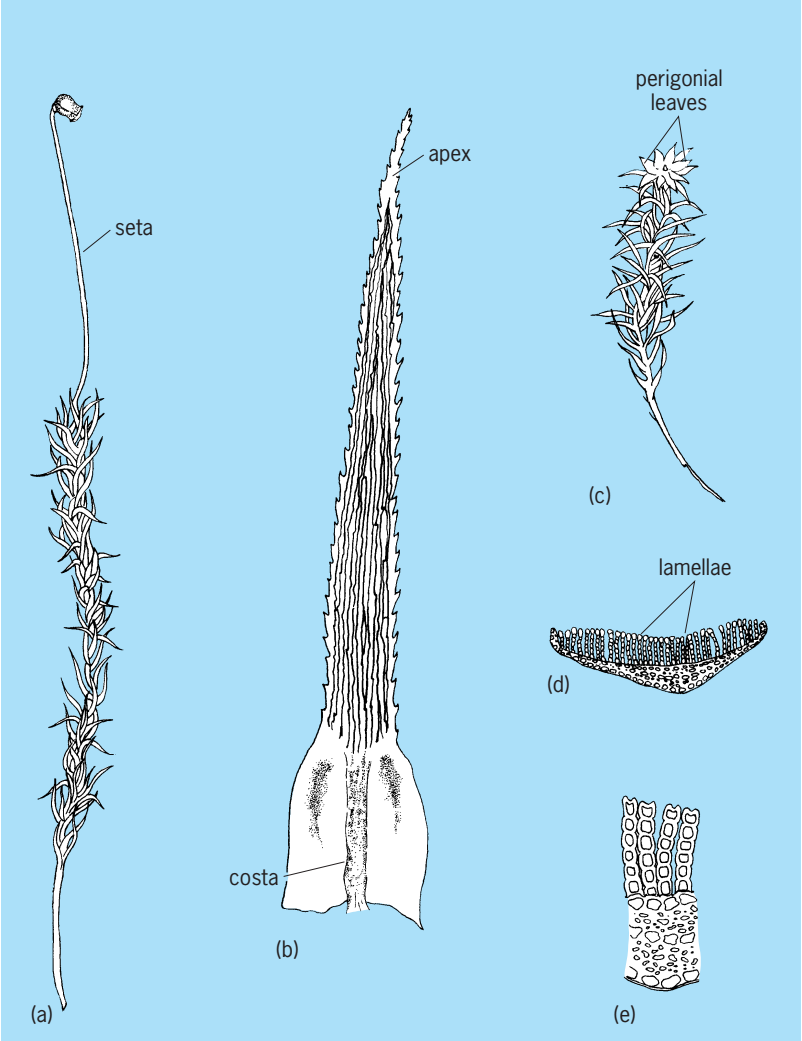

***Polytrichum commune*, member of the order Polytrichales, an acrocarpous perennial moss. (*a*) Female plant. (*b*) Serrate apex of leaf. (*c*) Male plant. (*d*) Cross section of leaf showing lamellae, which extend parallel to one another along the midrib of the leaf. (*e*) Lamellae enlarged; each of the lamellae is several cells high and one cell wide. (*After W. H. Welch, Mosses of Indiana, Indiana Department of Conservation, 1957*)**

(see **illus**.) is characteristic (though similar to that of the Dawsoniidae); the peristome structure is unique.

The plants grow on soil, often in dry habitats, and they are usually unbranched; annual increments often grow up through the male inflorescences, but do not give the appearance of forked stems. The plants are nearly always dioecious, with both male and female inflorescences terminal. The stems generally have well-developed vascular tissues. The long, narrow leaves are sheathed at the base; above the base the midrib often occupies most of the leaf and is covered with few to many longitudinal plates of green cells on the upper surface, sometimes also on the lower. Stomata are usually present at or near the base of the capsule. The peristome consists of a single series of 32–64 (or rarely 16) tongue-shaped teeth attached to a drumlike expansion of the tip of the columella. The tooth cells become long and fiber-like, thickened all around, and U-shaped, extending from the tip of one tooth to the tip of the next. The calyptrae are cucullate and often matted with hairs. The chromosome number is 7, sometimes in polyploid series. *See* BRYOPHYTA; BRYOPSIDA; DAWSONIIDAE.

Howard Crum

Polytropic process

A process which occurs with an interchange of both heat and work between the system and its surroundings. The nonadiabatic expansion or compression of a fluid is an example of a polytropic process. The interrelationship between the pressure (P) and volume (V) and pressure and temperature (T) for a gas undergoing a polytropic process are given by Eqs. (1) and (2), where a and b are the polytropic constants for

$$PV^a = \text{constant} \tag{1}$$

$$\frac{P^b}{T} = \text{constant} \tag{2}$$

the process of interest. These constants, which are usually determined from experiment, depend upon the equation of state of the gas, the amount of heat transferred, and the extent of irreversibility in the

process. Once these constants are known, Eqs. (1) and (2) can be used with the initial-state conditions (P_1 and T_1 or V_1) and one final-state condition (for example, P_2) to determine the temperature or specific volume of the final state. From this information, and other equations of thermodynamics, the entropy change and the enthalpy or internal energy change for the process can be evaluated, so that the power produced by a turbine or necessary to drive a compressor can be computed.

Equations (1) and (2) for a polytropic process may be compared with Eq. (3) for the isothermal expansion or compression of an ideal gas, and Eqs. (4)

$$PV = \text{constant} \tag{3}$$

$$\begin{aligned} PV^{-(C_p/C_v)} &= \text{constant} \\ \frac{P^{(R/C_p)}}{T} &= \text{constant} \end{aligned} \tag{4}$$

for the reversible adiabatic (isentropic) expansion or compression of an ideal gas. In Eqs. (3) and (4), R is the gas constant, and C_p and C_v the ideal gas heat capacity at constant pressure and volume, respectively. *See* GAS; ISENTROPIC PROCESS; ISOTHERMAL PROCESS; THERMODYNAMIC PROCESSES. Stanley I. Sandler

Polyurethane resins

Polymeric materials that are usually produced by reacting polyfunctional isocyanates [$R(N{=}C{=}O)_n$, where R is alkyl or aryl] with a macrodiol, or polyol, and a glycol extender in the presence of a catalyst. The macrodiols are based on polyethers or polyesters or a combination of both. Some of these have terminal hydroxyl groups, while those reacted with an excess of diisocyanate have terminal isocyanate groups, as in (structure **1**), the prepolymer of propylene oxide [with $R(NCO)_2$]. In recent years, diamines have been used as extenders to achieve higher reaction rates in molding or spray applications. The polymerization reactions are done in a mold (casting, injection molding) or continuously in an extruder or on a conveyor (block foam). *See* GLYCOL; ISOCYANATE; POLYMER; POLYESTER RESINS; POLYETHER RESINS; POLYMERIZATION; POLYOL.

Polyurethanes are the most versatile class of reaction polymers, providing products ranging from soft thermoplastic elastomers or elastomeric (spandex) fibers to flexible or rigid foams. The elastomers are used mainly in automotive bumper applications produced by the reaction injection-molding process. Spandex fibers are used in sport clothing applications. Flexible foams are used in upholstery materials for furniture and bedding, automotive seating, and carpet underlay. Rigid foams are used as insulation materials in building and construction, refrigerated appliances, and in vehicles. In 2000, the total polyurethane global consumption was about 8000 kilotons. *See* PLASTICS PROCESSING.

The commodity aromatic isocyanates are toluene diisocyanate (a distilled 80:20 mixture of the 2,4- and 2,6-isomers; TDI; structure **2**), 4,4′-diphenylmethane diisocyanate (MDI; **3**), and polymeric MDI (PMDI).

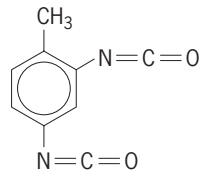

(2)

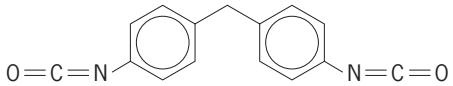

(3)

PMDI is obtained in quantitative yield in the phosgenation of aniline/formaldehyde resins. The diisocyanate, MDI, is separated from PMDI by thin-film vacuum distillation, and the distillation residue is sold as PMDI. The manufacture of TDI involves nitration of toluene, hydrogenation of the dinitrotoluenes to give diaminotoluenes, and phosgenation to give the diisocyanate.

When color stability is required in polyurethane resins, the more costly aliphatic diisocyanates are used. Examples are hexamethylene diisocyanate (HDI), isophorone diisocyanate (IPDI), and the reduced MDI (HMDI).

Foams. The diisocyanates used in polyurethanes are often prereacted with a small amount of the diols to form isocyanate-terminated prepolymers. Sometimes the lower- boiling aliphatic diisocyanates are converted into higher-boiling derivatives (allophanates, biurets, or trimers). In the manufacture of foams, water often is added as a comonomer. The water reacts with an isocyanate group to form carbon dioxide and an amine. The latter reacts with excess isocyanate to give a urea. In this manner, poly(urethane ureas) are generated. Poly(urethane ureas) are also obtained when diamines are used as extenders. The generated carbon dioxide is used as a co-blowing (foaming) agent in the production of foams. *See* AMINE; UREA.

Most commercial flexible foam is produced from TDI, a polyether triol, water, a surfactant, and tin/amine catalysts. Flexible foam blocks are produced continuously on a conveyor, while molded flexible foam is produced by reaction molding,

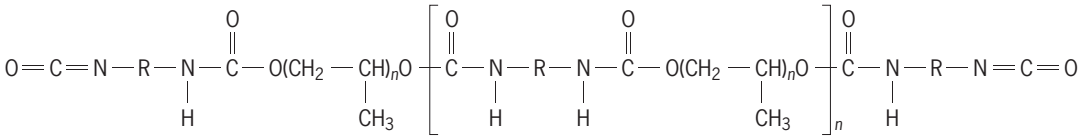

(1)

as in automotive seating applications. Semiflexible molded polyurethane foams are used in the construction of automotive instrument panels and dashboards. *See* CATALYSIS; SURFACTANT.

In the production of rigid polyurethane foams, chlorofluorocarbons (CFCs) were used as blowing agents. Because of the mandatory phase-out of CFCs caused by its effects on the ozone layer, new blowing agents, such as hydrogen-containing CFCs or hydrocarbons, are used today. In addition, poly(urethane isocyanurate) [PUIR] rigid foams are produced by using an excess of PMDI and a trimerization catalyst. The excess isocyanates undergo a cyclotrimerization reaction to form isocyanurate bonds, which have better thermal stability than polyurethane bonds. PUIR foams are used to insulate the main fuel tank of the space shuttle as well as liquefied natural-gas (LNG) tankers. *See* FOAM.

Elastomers. Polyurethane elastomers are segmented copolymers. The hard segments are formed in the reaction of the glycol extender with MDI, while the soft segments are obtained from the reaction of the macrodiol with MDI. Linear thermoplastic polyurethane elastomers (TPUs) are produced by reaction extrusion from MDI, polyols, and glycol extenders, while thermoset polyurethane elastomers are produced from TDI prepolymers and the diamine extender MOCA (3,3′-dichloro-4,4′-diaminodiphenylmethane), or from MDI prepolymers by reaction injection molding. In the latter reaction, diethyltoluenediamine (DETDA) is used as chain extender. Also, amine-terminated polyols are used in reaction injection molding and spray applications. Sometimes, monomers containing unsaturated groups are used, which can be cured with peroxides. Polyurethanes with the rheology of natural rubber, called millable gums, are made by mixing hydroxyl-terminated prepolymers and isocyanate cross-linking agents, such as the 2,4-TDI dimer, and other additives on a rubber mill. Soft cast elastomers are also used for puncture-free tires. Other tire applications include solid industrial tires for off-road machines, pneumatic tires for tractors, and skate board and roller skate wheels.

Segmented elastomeric fibers are produced from MDI, using hydrazine or ethylene diamine extenders in a dry spinning process, in which the polymer solution is isocyanate prepolymers extruded through a spinerette into a column of circulated hot air. Wet spinning processes are also used, in which isocyanate terminated prepolymers are extruded into a diamine bath. *See* MANUFACTURED FIBER; RUBBER.

Adhesives. Reconstructed wood products, such as particleboard, increasingly are using PMDI as a liquid adhesive. The use of PMDI as a binder for foundry cores, rubber waste products, or solid rocket fuels is also known. Athletic surfaces are sometimes constructed from ground rubber tire scrap bonded with isocyanate prepolymers. Also, flexible foam scrap is bonded with isocyanate prepolymers to form carpet underlay. *See* ADHESIVE; COMPOSITE MATERIAL; WOOD PRODUCTS.

Polyurethane sealant formulations use TDI or MDI prepolymers made from polyether polyols. The sealants contain 30 to 50% of the prepolymer; the remainder consists of pigments, fillers, plasticizers, adhesion promoters, and other additives.

Coatings. A wide variety of tough and abrasion-resistant polyurethane coatings are also available. Many are formed by reacting triols, such as castor oil, with excess diisocyante followed by moisture curing. Polyurethane alkyd coatings are produced by reacting an unsaturated drying oil with glycerol followed by curing with diisocyanates. Often polyurethane coatings are directly sprayed onto the surface of the substrate.

Solutions of fully reacted polyurethanes based on MDI or IPDI are also used in coating applications. Because of concern over volatile organic solvents, aqueous polyurethane dispersions with ionic groups in the polymer backbone are used. *See* PAINTS AND COATINGS.

Henri Ulrich

Bibliography. W. F. Gum, W. Riese, and H. Ulrich (eds.), *Reaction Polymers*, Carl Hanser, Munich, 1992; P. Herrington and K. Hock (eds.), *Flexible Polyurethane Foams*, Dow Chemical, 1991; N. M. K. Lamba, K. A. Woodhouse, and S. L. Cooper, *Polyurethanes in Biomedical Applications*, CRC Press, 1997; G. Oertel, *Polyurethane Handbook*, Carl Hanser, Munich, 1993; D. Randall and S. Lee (eds.), *The Polyurethane Book*, Wiley, 2002; H. Ulrich, *The Chemistry and Technology of Isocyanates*, Wiley, 1993; H. Ulrich, Polyurethanes, in *Encyclopedia of Polymer Science and Technology*, 3d ed., Wiley, 2003.

Polyvinyl resins

Polymeric materials generally considered to include polymers derived from monomers having the structure

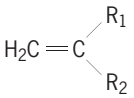

in which R_1 and R_2 represent hydrogen, alkyl, halogen, or other groups. This article refers to polymers whose names include the term vinyl. Of these polymers, several have been used for a number of years, such as poly(vinyl chloride), poly(vinyl acetate), poly(vinylidene chloride), poly(vinyl alcohol), poly(vinyl acetals), and poly(vinyl ethers). Indeed the terms vinyls and vinyl resins are frequently used to refer to the first three polymers of this group. Some polyvinyl resins of more recent origin are poly(vinyl fluoride), poly(vinylpyrrolidone), and poly(vinylcarbazole). For discussions of other vinyl-type polymers. *See* POLYACRYLATE RESIN; POLYACRYLONITRILE RESINS; POLYFLUOROOLEFIN RESINS; POLYOLEFIN RESINS; POLYSTYRENE RESIN.

Many of the monomers can be prepared by addition of the appropriate compound to acetylene. For example, vinyl chloride, vinyl fluoride, vinyl acetate, and vinyl methyl ether may be formed by the reactions of acetylene with HCl, HF, CH_3OOH, and CH_3OH, respectively. Processes based on ethylene

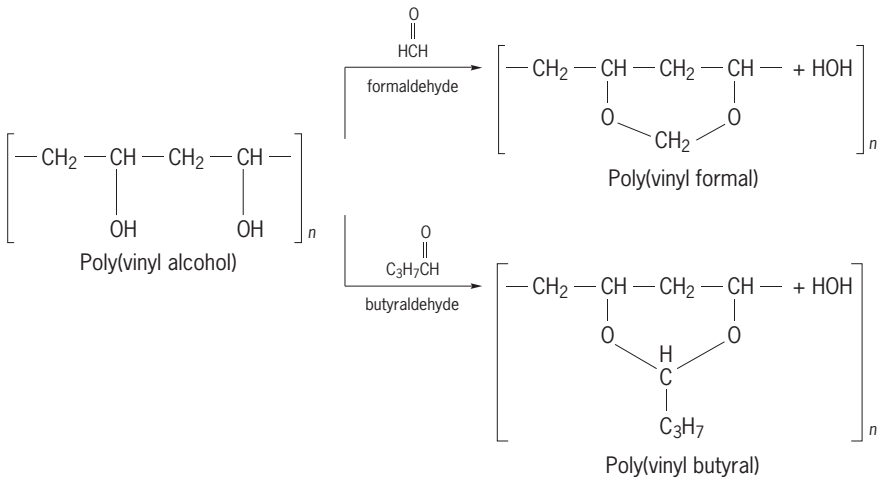

(1)

as a raw material have also become common for the preparation of vinyl chloride and vinyl acetate.

The polyvinyl resins may be characterized as a group of thermoplastics which, in many cases, are inexpensive and capable of being handled by solution, dispersion, injection molding, and extrusion techniques. The properties vary with chemical structure, crystallinity, and molecular weight.

Poly(vinyl acetals). These are relatively soft, water-insoluble thermoplastic products obtained by the reaction of poly(vinyl alcohol) with aldehydes [reaction (1)]. Properties depend on the extent to which alcohol groups are reacted. Poly(vinyl butyral) is rubber and tough and is used primarily in plasticized form as the inner layer and binder for safety glass. Poly(vinyl formal) is the hardest of the group; it is used mainly in adhesive, primer, and wire-coating formulations, especially when blended with a phenotic resin.

Poly(vinyl butyral) is usually obtained by the reaction of butyraldehyde with poly(vinyl alcohol). The formal can be produced by the same process, but is more conveniently obtained by the reaction of formaldehyde with poly(vinyl acetate) in acetic acid solution.

Poly(vinyl acetate). Poly(vinyl acetate) is a leathery, colorless thermoplastic material which softens at relatively low temperatures and which is relatively stable to light and oxygen. The polymers are clear and noncrystalline. The chief applications of poly(vinyl acetate) are as adhesives and binders for water-based or emulsion paints.

Vinyl acetate is conveniently prepared by the reaction of acetylene with acetic acid, as in reaction (2).

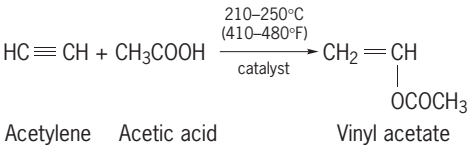

(2)

As acetaldehyde has become readily and cheaply available from the oxidation of olefins or hydrocarbons, it has been used to an increasing extent as a starting material. This is shown by reaction (3). Vinyl

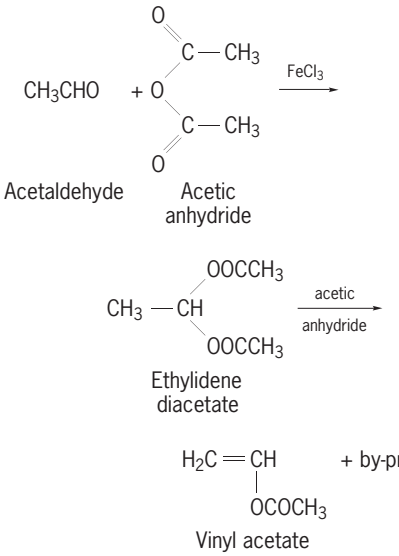

(3)

acetate is also prepared by the vapor-phase reaction of ethylene, oxygen, and acetic acid using a palladium chloride catalyst.

Polymerization and copolymerization may be conveniently effected by free-radical catalysis in aqueous emulsion and suspension systems. Vinyl acetate copolymerizes readily with various other vinyl monomers; however, it does not copolymerize with styrene by the free-radical process. The polymerization process is represented by reaction (4).

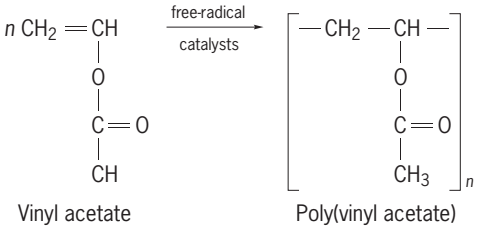

(4)

See COPOLYMER.

Anhydrous solid polymers and copolymers may be used directly in adhesives.

Aqueous dispersions, produced by the emulsion polymerization process and commonly called polymer emulsions, or latexes, are used for treating

textiles and paper, as adhesives, and as water-based paints. The water-based paints, prepared by pigmenting vinyl acetate polymer and copolymer emulsion achieved wide usage because of low cost of materials, ease of application, and resistance to weathering. As water is removed from the latex by evaporation or absorption, the suspended polymer particles coalesce into a tough film. The character of the film may be modified by the use of comonomers in the original polymerization or by the addition of plasticizers to the final emulsions.

Vinyl acetate is also frequently used as a comonomer with ethylene or vinyl chloride. *See* ETHYLENE.

Poly(vinyl alcohol). Poly(vinyl alcohol) is a tough, whitish polymer which can be formed into strong films, tubes, and fibers that are highly resistant to hydrocarbon solvents. Although poly(vinyl alcohol) is one of the few water-soluble polymers, it can be rendered insoluble in water by drawing or by the use of cross-linking agents.

So far, vinyl alcohol itself, $H_2C{=}CHOH$, has not been isolated; reactions designed to produce the monomer yield the tautomeric acetaldehyde instead. However, the polymer can be produced on a commercial scale by the hydrolysis of poly(vinyl acetate), reaction (5).

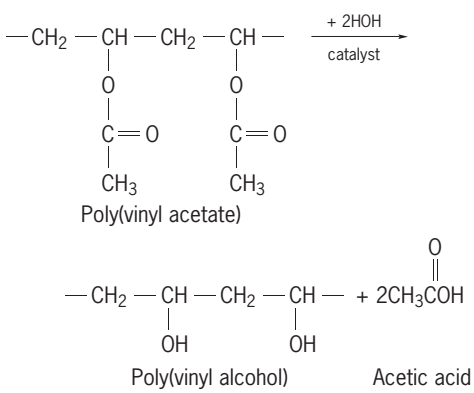

(5)

Two groups of products are available, those formed by the essentially complete hydrolysis (97% or greater) of poly(vinyl acetate), and those formed by incomplete hydrolysis (50–90%).

The former, the completely hydrolyzed products, may be plasticized with water or glycols and molded or extruded into films, tubes, and filaments which are resistant to hydrocarbons; these can be rendered insoluble to water by cold-drawing or heat or by the use of chemical cross-linking agents. On cold drawing, the degree of crystallinity is substantially increased. These products are used for liners in gasoline hoses, for grease-resistant coating and paper adhesives, for treating paper and textiles, and as emulsifiers and thickeners. Insolubilized fibers have found uses in Japan for clothing, industrial fabrics, and cordage. Films are strong, tough, and relatively impermeable to oxygen but tend to be sensitive to moisture.

The partially hydrolyzed products are generally more water-soluble and less subject to crystallization by drawing. These materials are used as emulsifying solutions, in adhesive formulations, and textile sizes, and in water-soluble films.

Poly(vinyl carbazole). Poly(vinyl carbazole) is a tough, glassy thermoplastic with excellent electrical properties and the relatively high softening temperature of 120–150°C (240–300°F). Polymerization, as represented by reaction (6), can be carried out in

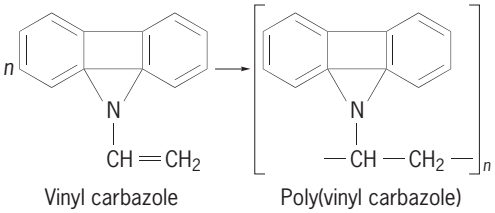

(6)

bulk by free-radical catalysis. Uses of the product have been limited to small-scale electrical applications requiring resistance to high temperatures.

Poly(vinyl chloride). Poly(vinyl chloride) (PVC) is a tough, strong thermoplastic material which has an excellent combination of physical and electrical properties. The products are usually characterized as plasticized or rigid types. Poly(vinyl chloride) [and copolymers] is the second most commonly used poly(vinyl resin) and one of the most versatile plastics.

The plasticized types, either soft copolymers or plasticized homopolymers, are somewhat elastic materials which are familiar in the form of shower curtains, floor coverings, raincoats, dishpans, dolls, bottle-top sealers, prosthetic forms, wire insulation, and films, among others.

Rigid poly(vinyl chloride) products, which may consist of the homopolymer, copolymer, or polyblends, are used in the manufacture of phonograph records, pipe, chemically resistant liners for chemical-reaction vessels, and siding and window sashes.

The monomer is frequently prepared from chlorine, acetylene, and ethylene by a combination of processes which affords complete utilization of the chlorine, as shown, for example, in reactions (7) and (8).

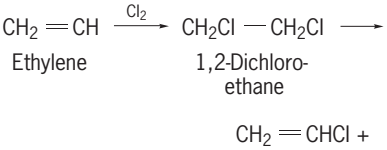

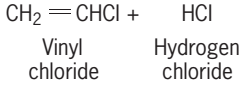

(7)

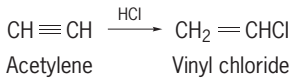

(8)

The polymerization of vinyle chloride, reaction (9), and its copolymerization with other vinyl

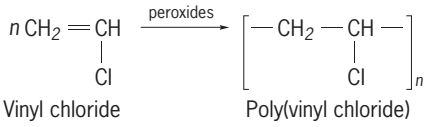

(9)

monomers may be initiated by peroxide or azo compounds and carried out in bulk or in aqueous emulsion or suspension systems. The structure and properties of the product are quite dependent on polymerization temperature; the lower the temperature, the higher the softening point of the resin. Polymers prepared in bulk or suspension are used for many applications as molded, extruded, or calendered objects. Polymers prepared in emulsion are often molded by dipping or pouring techniques.

Because poly(vinyl chloride) products have a tendency to lose hydrogen chloride at high temperatures, a stabilizer such as a tin or lead compound is included in the final composition. Fillers are also commonly incorporated.

Copolymers with monomers such as vinyl acetate or propylene can be processed at lower temperatures than the homopolymer. The acetate copolymer is especially useful in floor tiles and phonograph records, though homopolymers are used as well.

Blends or "alloys" of poly(vinyl chloride) with small amounts of rubbery materials such as the interpolymer of acrylonitrile, butadiene, and styrene (ABS) have been produced for applications such as panels and pipe in which impact resistance, as well as hardness and strength, is desired.

Chlorination of poly(vinyl chloride) is sometimes effected to obtain a stiffer resin, at the expense of processability.

Poly(vinylidene chloride). Poly(vinylidene chloride) is a tough, horny thermoplastic with properties generally similar to those of poly(vinyl chloride). In comparison with the latter, poly(vinylidene chloride) is softer and less soluble; it softens and decomposes at lower temperatures, crystallizes more readily, and is more resistant to burning.

Because of its relatively low solubility and decomposition temperature, the material is most widely used in the form of copolymers with other vinyl monomers, such as vinyl chloride. The copolymers are employed as packaging film, rigid pipe, and as filaments for upholstery and window screens.

Vinylidene chloride is normally prepared by the pyrolysis of 1,1,2-trichloroethane. The latter is obtained by the chlorination of 1,2-dichloroethane which, in turn, is formed by the addition of chlorine to ethylene. This preparation is shown by reactions (10) and (11).

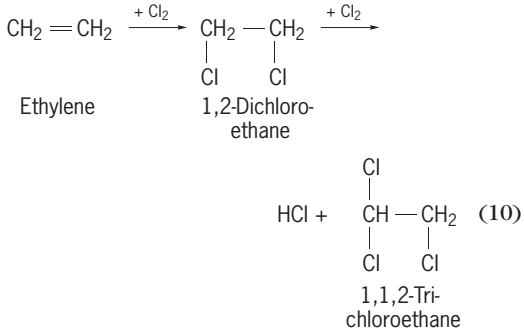

(10)

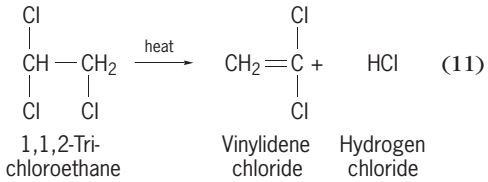

(11)

Polymerization as well as copolymerization may be initiated by peroxides and other free-radical catalysts and is most satisfactorily effected by emulsion and suspension techniques. Reaction (12) represents the polymerization.

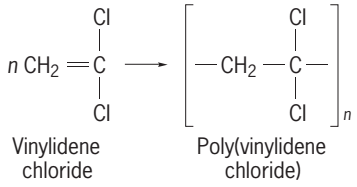

(12)

Because of the relatively low decomposition temperature of poly(vinylidene chloride), a stabilizer such as an amine is normally included in the composition.

Films of poly(vinylidene chloride), and especially the copolymer containing about 15% of vinyl chloride, are resistant to moisture and gases. Also, they can be heat sealed and, when oriented, have the property of shrinking on heating. By warming a food product wrapped loosely with a film of the polymer, a skintight, tough, resistant coating is produced.

By cold-drawing, the degree of crystallinity, strength, and chemical resistance of sheets, filaments, and even piping can be greatly increased.

Poly(vinyl ethers). Poly(vinyl ethers) exist in several forms varying from soft, balsamlike semisolids to tough, rubbery masses, all of which are readily soluble in organic solvents. Polymers of the alkyl vinyl ethers are used in adhesive formulations and as softening or flexibilizing agents for other polymers.

The monomers may be prepared by the reaction of alcohols with acetylene in the presence of alkali. Polymerization may be effected in bulk or solution at temperatures of −100 to 25°C (−148 to 77°F) by use of cationic initiators such as boron trifluoride, reaction (13). By careful choice of conditions, it is possi-

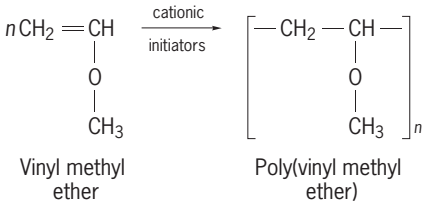

(13)

ble to achieve stereoregular polymerizations which yield partially crystalline polymers that are harder and tougher than the amorphous products.

Poly(vinyl methyl) ether is soluble in cold water but precipitates when the temperature is raised to about 35°C (95°F). The other alkyl vinyl ether polymers are insoluble in water. Copolymers of vinyl

methyl ether with maleic anhydride are useful in textile applications.

Poly(vinyl fluoride). Poly(0vinyl fluoride) is a tough, partially crystalline thermoplastic material which has a higher softening temperature than poly(vinyl chloride). Films and sheets are characterized by high resistance to weathering.

Polymerization can be effected in the presence of oxygen and peroxidic catalysts, reaction (14). Be-

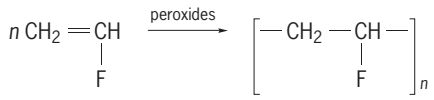

(14)

Vinyl fluoride Poly(vinyl fluoride)

cause of the low boiling point (−88°C or −126°F) and high critical temperature of the monomer, polymerization is accomplished by use of pressure techniques similar to those employed in the high-pressure process for polymerizing ethylene. Like other poly(vinyl halides), poly(vinyl fluoride) tends to lose the halogen acid at elevated temperatures.

Films are used in industrial and architectural applications. Coatings, for example, on pipe, are resistant to highly corrosive media.

Poly(vinyl pyrrolidone). Poly(vinyl pyrrolidone) is a water-soluble polymer of basic nature which has film-forming properties, strong absorptive or complexing qualities for various reagents, and the ability to form water-soluble salts which are polyelectrolytes. The polymer can be prepared by free-radical polymerization in bulk or aqueous solution, reaction (15). Iso-

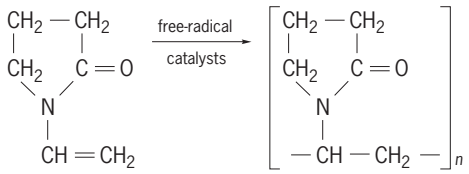

(15)

tonic solutions were used in Germany in World War II as an extender for blood plasma. The main uses today are as a water-solubilizing agent for medicinal agents such as iodine, and as a semipermanent setting agent in hair sprays. Certain synthetic textile fibers containing small amounts of vinylpyrrolidone as a copolymer have improved affinity for dyes. *See* PLASTICS PROCESSING; POLYMER; POLYMERIZATION.

John A. Manson

Bibliography. H. Allcock, F. Lampe, and J. Mark, *Contemporary Polymer Chemistry*, 3d ed., 2003; G. Odian, *Principles of Polymerization*, 4th ed., 2004; M. P. Stevens, *Polymer Chemistry: An Introduction*, 3d ed., 1998.

Pomegranate

A small deciduous tree, *Punica granatum*, belonging to the plant order Myrtales. Pomegranate is grown as an ornamental as well as for its fruit. Propagation is by cuttings and occasionally by layering. The pomegranate is a native of Asia, and was described by writers as early as 300 B.C. It was originally known for its medicinal qualities, and cures for various ills were attributed to the fruit juice, the rind, and the bark of the roots. The fruit is a reddish, pomelike berry, containing numerous seeds embedded in crimson pulp, from which an acidic, reddish juice may be obtained. Limited quantities are grown in California and the Gulf states. *See* FRUIT, TREE; MYRTALES; PLANT PROPAGATION.

J. H. Clarke

Poplar

Any tree of the genus *Populus*, family Salicaceae, marked by simple, alternate leaves which are usually broader than those of the willow, the other American representative of this family. Poplars have scaly buds, bitter bark, flowers and fruit in catkins, and a five-angled pith. *See* SALICALES; WILLOW.

Some species are commonly called cottonwood because of the cottony hairs attached to the seeds. Other species, called aspens, have weak, flattened leaf stalks which cause the leaves to flutter in the slightest breeze. One of the important species in the United States is the quaking, or trembling, aspen (*P. tremuloides*), which attains a height of 90 ft (27 m) in the Rockies but is smaller in the East. This tree is widely distributed in North America from Labrador to Alaska. In the West it extends south through the Rockies and California to New Mexico and Lower California, and in the eastern United States it grows as far south as West Virginia. It is readily recognized by its comparatively small, finely toothed leaves, and by the shiny pointed winter buds (**illus.** *a*). The soft wood of this species is used for paper pulp. The bigtooth aspen (*P. grandidentata*) attains a height of 60–70 ft (18–21 m), and occasionally a diameter of 2 ft (0.6 m). The bigtooth aspen has a more restricted range in the northeastern quarter of the United States. The leaves are usually larger, $2^1/_2$–4 in. (6–10 cm) long, and have larger teeth (illus. *b*), hence the common name. The buds are plumper and somewhat downy.

The European aspen (*P. nigra*; illus. *e*), which is similar to the quaking aspen, is sometimes planted, and its variety, *italica*, the Lombardy poplar of erect columnar habit, is used in landscape planting.

The black cottonwood (*P. trichocarpa*) is the largest American poplar and is also the largest broad-leaved tree in the forests of the Pacific Northwest. It attains a height of 175–225 ft (53–69 m) and a diameter of 7–8 ft (2.1–2.4 m). This tree ranges from southern Alaska to California and eastward through Washington and Oregon to Idaho, Montana, and Nebraska. The hairy fruit of the black cottonwood is a three-valved capsule.

The cottonwood or necklace poplar (*P. deltoides*; illus. *c*), native in the eastern half of the United States, is a fast-growing tree which usually attains 80–100 ft (24–30 m) in height and 3–4 ft (0.9–1.2 m) in diameter, but under favorable conditions in the

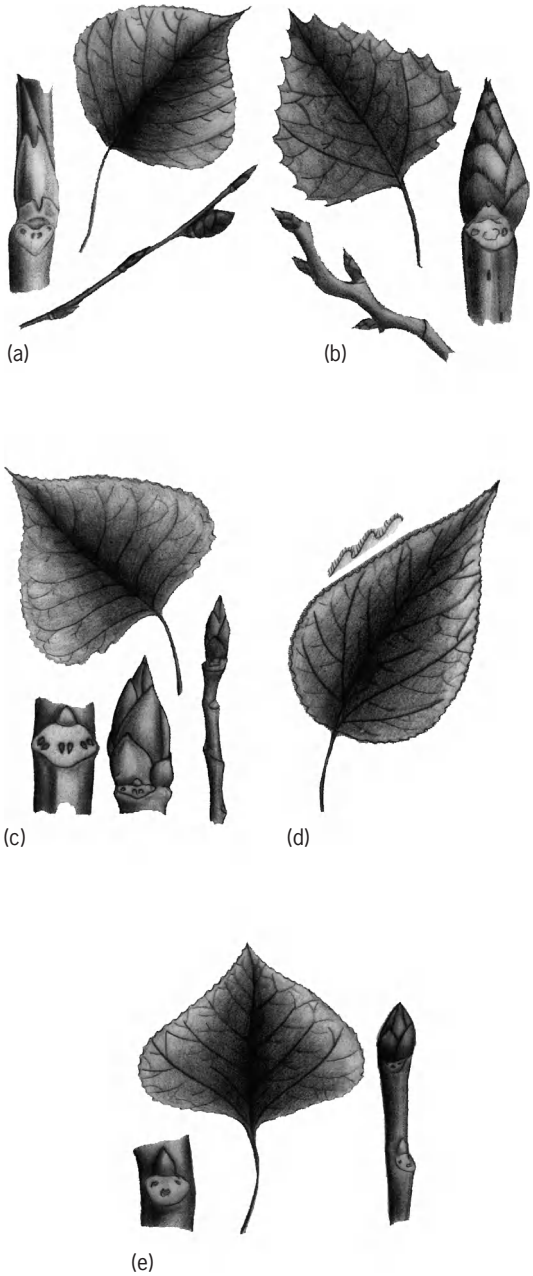

Common poplar species. (*a*) Quaking aspen (*Populus tremuloides*); leaf scar with an axial bud, leaf, and twig. (*b*) Bigtooth aspen (*P. grandidentata*); leaf, terminal bud, and twig. (*c*) Cottonwood poplar (*P. deltoides*); leaf, leaf scar, terminal bud, and twig. (*d*) Baldwin poplar (*P. balsamifera*); leaf showing detail of margin. (*e*) Lombardy poplar (*P. nigra* var. *italica*); leaf scar, leaf, and twig.

Mississippi Valley, it may attain a height of 150 ft (45 m) and a diameter of 7–8 ft (2.1–2.4 m). The leaves are broadly triangular (illus. *c*), hence the specific name, and the large terminal buds contain a pleasant-smelling balsamic resin. In the balsam or tacamahac poplar (*P. balsamifera*; illus. *d*), the resin is used in medicine as an expectorant. The wood is used for veneer, boxes, crates, furniture, paper pulp, and excelsior. It is also planted as a shade tree and used in shelter belts. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Poppy

A plant, *Papaver somniferum* (Papaveraceae), which is probably a native of Asia Minor. It is cultivated extensively in China, India, and elsewhere. This plant is the source of opium, obtained by cutting into the fruits (capsules) soon after the petals have fallen (see **illus.**). The white latex (juice) flows

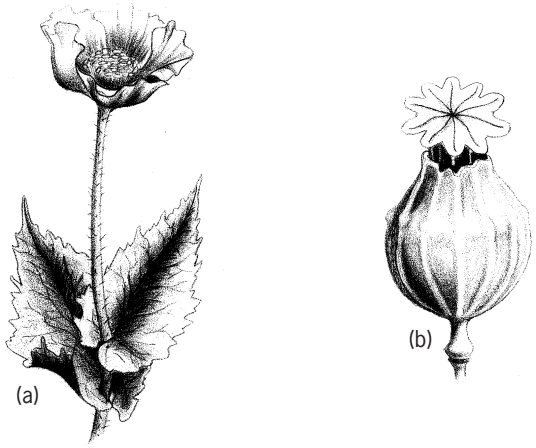

Opium poppy (*Papaver somniferum*). (*a*) Stem with flower. (*b*) Capsule of poppy.

from the cuts and hardens when exposed to the air. This solidified latex is collected, shaped into balls or wafers, and often wrapped in the flower petals. This is the crude opium, which contains at least 20 alkaloids, including morphine and codeine. These drugs are used in medicine to allay pain, induce sleep, and relax spasms. Opium is one of the most useful drugs, but it is habit-forming and consequently should be used with the utmost caution. The opium habit is deleterious physically, mentally, and morally, and misuse of the drug is an extremely serious problem. *See* ALKALOID; OPIATES; PAPAVERALES.

Perry D. Strausbaugh; Earl L. Core

Population dispersal

The process by which groups of living organisms expand the space or range within which they live. Dispersal operates when individual organisms leave the space that they have occupied previously, or in which they were born, and settle in new areas. Natal dispersal is the first movement of an organism from its birth site to the site in which it first attempts to breed. Adult dispersal is a subsequent movement when an adult organism changes its location in space. As individuals move across space and settle into new locations, the population to which they belong expands or contracts its overall distribution. Thus, dispersal is the process by which populations change the area they occupy.

Migration. This is the regular movement of organisms during different seasons. Many birds, butterflies, salmon, caribou, and other species migrate

between wintering and breeding ranges. Such migratory movement is marked by a regular return in future seasons to previously occupied regions, and so usually does not involve an expansion of population range. Some migratory species show astounding abilities to return to the exact locations used in previous seasons, even if experimentally transported away from their normal movement paths. Other species show no regular movements, but wander aimlessly without settling permanently into a new space. Wandering (called nomadism) is typical of species in regions where the availibility of food resources are unpredictable from year to year. Many birds in the harsh deserts of Australia are nomads. Neither migration nor nomadism is considered an example of true dispersal. *See* MIGRATORY BEHAVIOR.

Dispersal unit. Virtually all forms of animals and plants disperse. In most higher vertebrates, the dispersal unit is an entire organism, often a juvenile or a member of another young age class. In other vertebrates and many plants, especially those that are sessile (permanently attached to a surface), the dispersal unit is a specialized structure (disseminule). Seeds, spores, and fruits are disseminules of plants and fungi; trochophores and planula larvae are disseminules of sea worms and corals, respectively. Many disseminules are highly evolved structures specialized for movement by specific dispersal agents such as wind, water, or other animals.

Transportation. Disseminules or dispersing individuals can be transported passively (moved along with a dispersal agent) or actively (under their own efforts). Passive dispersal may involve movement by wind (anemochory) or water (hydrochory). Wind dispersal is common in plants and many small invertebrates such as spiders, and allows dispersers to be scattered over a large spatial area. Species that are associated with highly disturbed habitats (for example, dandelions) are often wind-dispersed, because patches of disturbed habitat may be widely scattered across a landscape, forcing dispersers to travel great distances.

Water dispersal is common in all marine and other aquatic organisms. Marine plankton usually contains many dispersers, such as the larval forms of organisms from the intertidal or benthic zones. Many terrestrial organisms found along shorelines, rivers, and streams also disperse passively by using water currents. Perhaps the best-known example of a disseminule that is evolved for dispersal using water currents is the coconut. Coconut palms are widely distributed on many tropical beaches, mainly because their disseminule, the coconut itself, can float across large expanses of ocean and still be viable when washed ashore.

Dispersal using animal agents (zoochory) involves attachment to skin, feathers, or scales of the animal agent with barbed or sticky surfaces on the disseminule, or ingestion of the disseminule by the dispersal agent. Many disseminules are highly specialized to be carried or eaten by a specific agent, and many agents are highly evolved to make use of specific disseminules. This close evolution between a disperser and its animal vector is an example of the process of coevolution. For example, phainopeplas, birds that inhabit the deserts of the southwestern United States, eat mistletoe berries for the thin layer of sugary pulp that surrounds the seed. The seeds pass through the digestive system of the bird and are excreted onto the branches of desert trees and shrubs, where the seeds stick and grow into new mistletoe plants. Thus, the phainopepla is an effective dispersal agent for the plant, while the plant is a reliable food source for the bird.

A special case of zoochory involves transport by humans. The movement of people and cargo by cart, car, train, plane, and boat has increased the potential dispersal of weedy species worldwide. Many foreign aquatic species have been introduced to coastal areas by accidental dispersal of disseminules in ship ballast water, which is usually dumped overboard when ships reach their destination. The zebra mussel is one exotic species that arrive in this manner and is now a major economic problem throughout the Great Lakes region of North America. Countries such as Australia and New Zealand take precautions to ensure that airplane passengers do not accidentally carry seeds of foreign plants into the countries in their luggage or on their shoes.

Some organisms have been deliberately introduced by humans into new areas. Domestic animals and plants have been released throughout the world by farmers. A few pest species were deliberately released by humans; European starlings, for example, were originally introduced to the United States via New York City. However, they have spread across most of North America and are considered to be one of the worst avian pests ever released on this continent.

Seed dispersal systems and coevolution. Some of the most highly coevolved dispersal systems are those in which the disseminule must be eaten by an animal. Such systems have often evolved a complex series of signals and investments by both the plant and the animal to ensure that the seeds are dispersed at an appropriate time and that the animal is a dependable dispersal agent. Such highly evolved systems are common in fruiting plants and their dispersal agents, which are animals called frugivores. Fruiting plants cover their seeds with an attractive, edible package (the fruit) to get the frugivore to eat the seed. To ensure that fruits are not eaten until the seeds are mature, plants change the color of their fruits as a signal to show that the fruits are ready for eating. Thus, when cherries change from green to dark red, potential frugivores (including humans) are alerted that the fruit is ready to eat.

When eaten by a frugivore, seeds must be able to survive passage through the animal's digestive system without being destroyed. These seeds are often covered with a tough outer coating to protect the seed from the digestion process. In some species, the seed must be scarred by the digestion process itself for the seed to be able to germinate. A loss of the frugivorous dispersal agent in these latter systems means that seeds will be unable to sprout because of

lack of scarification. A number of Central American trees produce large numbers of huge fruits, which fall to the ground and rot under the parent tree, without germination. It has been speculated that these trees were coevolved to have their fruits dispersed by the large frugivorous members of the Pleistocene mammal fauna such as the gomphotheres. Extinction of these large mammals left the trees without an effective dispersal agent. The range of many of these tree species has decreased until the trees have become quite rare—an example of how ineffective dispersal can limit populations. However, cattle have been introduced to fill the niche of the extinct frugivores, and these tropical trees have begun to expand their range again as cattle disperse their seeds.

Many plants in the tropical rainforests are coevolved to have their seeds dispersed by specific animal vectors, including birds, mammals, and ants. Many tropical trees, shrubs, and herbaceous plants are specialized to have their seeds dispersed by a single animal species. Temperate forest trees, in contrast, often depend on wind dispersal of both pollen and seeds.

Dispersal barriers. Dispersal barriers are physical structures that prevent organisms from crossing into new space. Oceans, rivers, roads, and mountains are examples of barriers for species whose disseminules cannot cross such features. These types of barriers are often regions of unsuitable temperature, humidity, or light conditions in which disseminules die before reentering suitable conditions. Dispersal barriers do not have to be large. For example, several species of rare forest butterflies refuse to cross open, nonforested space, no matter how small the open gap; some neotropical forest birds do not cross open water. Thus, populations of such species can quickly become isolated due to lack of successful dispersal.

It is believed that the creation of physical barriers such as mountains and rivers is the primary factor responsible for the evolution of new species. A widespread species can be broken into isolated fragments by the creation of a new physical barrier. With no dispersal linking the newly isolated populations, genetic differences that evolve in each population cannot be shared between populations. Eventually, the populations may become so different that no interbreeding occurs even if dispersal pathways are reconnected. The populations are then considered separate species. Therefore, an understanding of the role that dispersal plays in maintaining genetic variation among populations is an important subject in evolutionary biology. *See* SPECIATION.

Conservation importance. Dispersal is of major concern for scientists who work with rare and endangered animals. Extinction is known to be more prevalent in small, isolated populations. Conservation biologists believe that many species exist as a metapopulation, that is, a group of populations interconnected by the dispersal of individuals or disseminules between subpopulations. The interruption of dispersal in this system of isolated populations can increase the possibility of extinction of the whole metapopulation. Conservation plans sometimes propose the creation of corridors to link isolated patches of habitat as a way of increasing the probability of successful dispersal. Corridors have been proposed for species such as mountain lions in highly fragmented habitats in the western United States, but are controversial because little data usually exist on how the species involved actually disperses. *See* EXTINCTION (BIOLOGY); POPULATION DISPERSION.

John B. Dunning, Jr.

Bibliography. C. S. Elton, *The Ecology of Invasions by Animals and Plants*, 1958; B. Sleeper, *Animal Migrations: Wildlife in Motion*, 1994; N. C. Stenseth and W. Z. Lidicker (eds.), *Animal Dispersal: Small Mammals as a Model*, 1994; I. R. Swingland and P. J. Greenwood, *The Ecology of Animal Movement*, 1983.

Population dispersion

The spatial distribution at any particular moment of the individuals of a species of plant or animal. Under natural conditions organisms are distributed either by active movements, or migrations, or by passive transport by wind, water, or other organisms. The act or process of dissemination is usually termed dispersal, while the resulting pattern of distribution is best referred to as dispersion. Dispersion is a basic characteristic of populations, controlling various features of their structure and organization. It determines population density, that is, the number of individuals per unit of area, or volume, and its reciprocal relationship, mean area, or the average area per individual. It also determines the frequency, or chance of encountering one or more individuals of the population in a particular sample unit of area, or volume. The ecologist therefore studies not only the fluctuations in numbers of individuals in a population but also the changes in their distribution in space. *See* POPULATION DISPERSAL.

Principal types of dispersion. The dispersion pattern of individuals in a population may conform to any one of several broad types, such as random, uniform, or contagious (clumped). Any pattern is relative to the space being examined; a population may appear clumped when a large area is considered, but may prove to be distributed at random with respect to a much smaller area.

Random or haphazard. This implies that the individuals have been distributed by chance. In such a distribution, the probability of finding an individual at any point in the area is the same for all points (**Fig. 1***a*). Hence a truly random pattern will develop only if each individual has had an equal and independent opportunity to establish itself at any given point. In a randomly dispersed population, the relationship between frequency and density can be expressed by Eq. (1), where F is percentage frequency, D is

$$F = 100(1 - e^{-D}) \qquad (1)$$

density, and e is the base of natural or napierian logarithms. Thus when a series of randomly selected

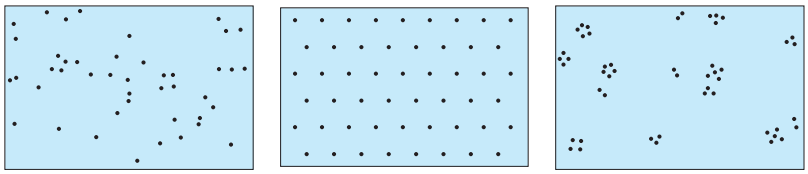

Fig. 1. Basic patterns of the dispersion of individuals in a population. (*a*) Random. (*b*) Uniform. (*c*) Clumped, but groups random. (*After E. P. Odum, Fundamentals of Ecology, Saunders, 1953*)

samples is taken from a population whose individuals are dispersed at random, the numbers of samples containing 0, 1, 2, 3,. . . , n individuals conform to the well-known Poisson distribution described by notation (2).

$$e^{-D}, De^{-D}, \frac{D^2}{2!}e^{-D}, \frac{D^3}{3!}e^{-D}, \ldots, \frac{D^n}{n!}e^{-D} \qquad (2)$$

Randomly dispersed populations have the further characteristic that their density, on a plane surface, is related to the distance between individuals within the population, as shown in Eq. (3), where $\bar{r}$ is the

$$D = \frac{1}{4\bar{r}^2} \qquad (3)$$

mean distance between an individual and its nearest neighbor. These mathematical properties of random distributions provide the principal basis for a quantitative study of population dispersion. Examples of approximately random dispersions can be found in the patterns of settlement by free-floating marine larvae and of colonization of bare ground by airborne disseminules of plants. Nevertheless, true randomness appears to be relatively rare in nature, and the majority of populations depart from it either in the direction of uniform spacing of individuals or more often in the direction of aggregation.

Uniform. This type of distribution implies a regularity of distance between and among the individuals of a population (Fig. 1*b*). Perfect uniformity exists when the distance from one individual to its nearest neighbor is the same for all individuals. This is achieved, on a plane surface, only when the individuals are arranged in a hexagonal pattern. Patterns approaching uniformity are most obvious in the dispersion of orchard trees and in other artificial plantings, but the tendency to a regular distribution is also found in nature, as for example in the relatively even spacing of trees in forest canopies, the arrangement of shrubs in deserts, and the distribution of territorial animals.

Contagious or clumped. The most frequent type of distribution encountered is contagious or clumped (Fig. 1*c*), indicating the existence of aggregations or groups in the population. Clusters and clones of plants, and families, flocks, and herds of animals are common phenomena. The degree of aggregation may range from loosely connected groups of two or three individuals to a large compact swarm composed of all the members of the local population. Furthermore, the formation of groups introduces a higher order of complexity in the dispersion pattern, since the several aggregations may themselves be distributed at random, evenly, or in clumps. An adequate description of dispersion, therefore, must include not only the determination of the type of distribution, but also an assessment of the extent of aggregation if the latter is present.

Analysis of dispersion. If the type or degree of dispersion is not sufficiently evident upon inspection, it can frequently be ascertained by use of sampling techniques. These are often based on counts of individuals in sample plots or quadrats. Departure from randomness can usually be demonstrated by taking a series of quadrats and testing the numbers of individuals found therein for their conformity to the calculated Poisson distribution which has been described above. The observed values can be compared with the calculated ones by a chi-square test for goodness of fit, and lack of agreement is an indication of nonrandom distribution. If the numbers of quadrats containing zero or few individuals, and of those with many individuals, are greater than expected, the population is clumped; if these values are less than expected, a tendency toward uniformity is indicated. Another measure of departure from randomness is provided by the variance:mean ratio, which is 1.00 in the case of the Poisson (random) distribution. If the ratio of variance to mean is less than 1.00, a regular dispersion is indicated; if the ratio is greater than 1.00, the dispersion is clumped.

In the case of obviously aggregated populations, quadrat data have been tested for their conformity to a number of other dispersion models, such as Neyman's contagious, Thomas' double Poisson, and the negative binomial distributions. However, the results of all procedures based on counts of individuals in quadrats depend upon the size of the quadrat employed. Many nonrandom distributions will seem to be random if sampled with very small or very large quadrats, but will appear clumped if quadrats of medium size are used. Therefore the employment of more than one size of quadrat is recommended.

A measure of aggregation that does not depend on quadrat size of the mean density of individuals per quadrat and that can be applied to patterns consisting of a mosaic of patches with different densities has been developed by Morisita. His index of dispersion is a ratio of the observed probability of drawing two individuals randomly from the same quadrat to the expected probability of the same event for individuals randomly dispersed over the set of quadrats being studied. Index values greater than 1.0 indicate clumping, and values between 0 and 1.0 point to regularity of dispersion.

The fact that plot size may influence the results of quadrat analysis has led to the development of a number of techniques based on plotless sampling. These commonly involve measurement of the distance between a randomly selected individual and its nearest neighbor, or between a randomly selected point and the closest individual. At least four different procedures have been used (**Fig. 2**). The

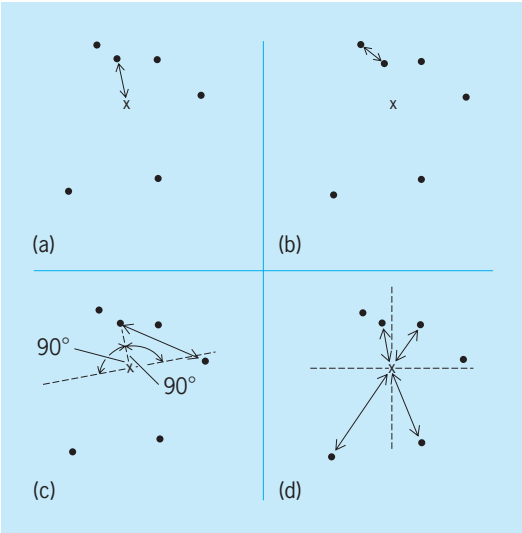

Fig. 2. Distances measured in four methods of plotless sampling. (*a*) Closest individual. (*b*) Nearest neighbor. (*c*) Random pairs, with 180° exclusion angle. (*d*) Point-centered quarter. The cross represents the sampling point in each case. (*After P. Greig-Smith, Quantitative Plant Ecology, Butterworths, 1987*)

closest-individual method (Fig. 2*a*) measures the distance from each sampling point to the nearest individual. The nearest-neighbor method (Fig. 2*b*) measures the distance from each individual to its nearest neighbor. The random-pairs method (Fig. 2*c*) establishes a base line from each sampling point to the nearest individual, and erects a 90° exclusion angle to either side of this line. The distance from the nearest individual lying outside the exclusion angle to the individual used in the base line is then measured. The point-centered quarter method (Fig. 2*d*) measures the distance from each sampling point to the nearest individual in each quadrant.

In each of these four methods of plotless sampling, a series of measurements is taken which can be used as a basis for evaluating the pattern of dispersion. In the case of the closest-individual and the nearest-neighbor methods, a population whose members are distributed at random will yield a mean distance value that can be calculated by use of the density-distance equation (3). In an aggregated distribution, the mean observed distance will be less than the one calculated on the assumption of randomness; in a uniform distribution it will be greater. Thus the ratio $\bar{r}_A/\bar{r}_E$, where $\bar{r}_A$ is the actual mean distance obtained from the measured population and $\bar{r}_E$ is the mean distance expected under random conditions, affords a measure of the degree of deviation from randomness.

Students of human geography have used the nearest-neighbor measure as a basis for a highly sophisticated methodology to analyze the dispersion of towns, department stores, and other features of land-use patterns.

Additional information about the spatial relations in a population can be secured by extending these procedures to measurement of the distance to the second and successive nearest neighbors, or by increasing the number of sectors about any chosen sampling point. However, since all of these methods assume that the individuals are small enough to be treated mathematically as points, they become less accurate when the individuals cover considerable space.

Factors affecting dispersion. The principal factors that determine patterns of population dispersion include (1) the action of environmental agencies of transport, (2) the distribution of soil types and other physical features of the habitat, (3) the influence of temporal changes in weather and climate, (4) the behavior pattern of the population in regard to reproductive processes and dispersal of the young, (5) the intensity of intra- and interspecific competition, and (6) the various social and antisocial forces that may develop among the members of the population. Although in certain cases the dispersion pattern may be due to the overriding effects of one factor, in general populations are subject to the collective and simultaneous action of numerous distributional forces and the dispersion pattern reflects their combined influence. When many small factors act together on the population, a more or less random distribution is to be expected, whereas the domination of a few major factors tends to produce departure from randomness.

Environmental agencies of transport. The transporting action of air masses, currents of water, and many kinds of animals produces both random and nonrandom types of dispersion. Airborne seeds, spores, and minute animals are often scattered in apparently haphazard fashion, but aggregation may result if the wind holds steadily from one direction. Wave action is frequently the cause of large concentrations of seeds and organisms along the drift line of lake shores. The habits of fruit-eating birds give rise to the clusters of seedling junipers and cherries found beneath such perching sites as trees and fencerows, as well as to the occurrence of isolated individuals far from the original source. Among plants, it seems to be a general principle that aggregation is inversely related to the capacity of the species for seed dispersal.

Physical features of the habitat. Responses of the individuals of the population to variations in the habitat also tend to give rise to local concentrations. Environments are rarely uniform throughout, some portions generally being more suitable for life than others, with the result that population density tends to be correlated directly with the favorability of the habitat. Oriented reactions, either positive or negative, to light intensities, moisture gradients, or to sources of food or shelter, often bring numbers of individuals into a restricted area. In these cases, aggregation results from a species-characteristic response to the environment and need not involve any social reactions to other members of the population. *See* ENVIRONMENT.

Influence of temporal changes. In most species of animal, daily and seasonal changes in weather evoke movements which modify existing patterns of dispersion. Many of these are associated with the

disbanding of groups as well as with their formation. Certain birds, bats, and even butterflies, for example, form roosting assemblages at one time of day and disperse at another. Some species tend to be uniformly dispersed during the summer, but flock together in winter. Hence temporal variation in the habitat may often be as effective in determining distribution patterns as spatial variation.

Behavior patterns in reproduction. Factors related to reproductive habits likewise influence the dispersion patterns of both plant and animal populations. Many plants reproduce vegetatively, new individuals arising from parent rootstocks and producing distinct clusters; others spread by means of rhizomes and runners and may thereby achieve a somewhat more random distribution. Among animals, congregations for mating purposes are common, as in frogs and toads and the breeding swarms of many insects. In contrast, the breeding territories of various fishes and birds exhibit a comparatively regular dispersion. *See* REPRODUCTIVE BEHAVIOR.

Intensity of competition. Competition for light, water, food, and other resources of the environment tends to produce uniform patterns of distribution. The rather regular spacing of trees in many forests is commonly attributed largely to competition for sunlight, and that of desert plants for soil moisture. Thus a uniform dispersion helps to reduce the intensity of competition, while aggregation increases it. *See* POPULATION ECOLOGY.

Social factors. Among many animals the most powerful forces determining the dispersion pattern are social ones. The social habit leads to the formation of groups or societies. Plant ecologists use the term society for various types of minor communities composed of several to many species, but when the word is applied to animals it is best confined to aggregations of individuals of the same species which cooperate in their life activities. Animal societies or social groups range in size from a pair to large bands, herds, or colonies. They can be classified functionally as mating societies (which in turn are monogamous or polygamous, depending on the habits of the species), family societies (one or both parents with their young), feeding societies (such as various flocks of birds or schools of fishes), and as migratory societies, defense societies, and other types. Sociality confers many advantages, including greater efficiency in securing food, conservation of body heat during cold weather, more thorough conditioning of the environment to increase its habitability, increased facilitation of mating, improved detection of, and defense against predators, decreased mortality of the young and a greater life expectancy, and the possibility of division of labor and specialization of activities. Disadvantages include increased competition, more rapid depletion of resources, greater attraction of enemies, and more rapid spread of parasites and disease. Despite these disadvantages, the development and persistence of social groups in a wide variety of animal species is ample evidence of its overall survival value. Some of the advantages of the society are also shared by aggregations that have no social basis. *See* ECOLOGICAL COMMUNITIES; SOCIAL MAMMALS.

Optimal population density. The degree of aggregation which promotes optimum population growth and survival, however, varies according to the species and the circumstances. Groups or organisms often flourish best if neither too few nor too many individuals are present; they have an optimal population density at some intermediate level. The condition of too few individuals, known as undercrowding, may prevent sufficient breeding contacts for a normal rate of reproduction. On the other hand, overcrowding, or too high a density, may result in severe competition and excessive interaction that will reduce fecundity and lower the growth rate of individuals. The concept of an intermediate optimal population density is sometimes known as Allee's principle. *See* POPULATION GENETICS.

Francis C. Evans

Bibliography. A. A. Berryman (ed.), *Dynamics of Forest Insect Populations*, 1988; A. A. Berryman (ed.), *Dynamics of Insect Populations: Patterns, Causes, Implications*, 1988; P. Greig-Smith, *Quantitative Plant Ecology,* 3d ed., 1983; S. K. Jain and L. W. Botsford (eds.), *Applied Population Biology*, 1992; S. A. Levin (ed.), *Applied Mathematical Ecology*, 1989; A. G. Pakes and R. A. Maller, *Mathematical Ecology of Plant Species Competition*, 1990; E. C. Pielou, *Population and Community Ecology: Principles and Methods,* 1974.

Population ecology

The study of spatial and temporal patterns in the abundance and distribution of organisms and of the mechanisms that produce those patterns. Species differ dramatically in their average abundance and geographical distributions, and they display a remarkable range of dynamical patterns of abundance over time, including relative constancy, cycles, irregular fluctuations, violent outbreaks, and extinctions. The aims of population ecology are threefold: (1) to elucidate general principles explaining these dynamic patterns; (2) to integrate these principles with mechanistic models and evolutionary interpretations of individual life-history tactics, physiology, and behavior as well as with theories of community and ecosystem dynamics; and (3) to apply these principles to the management and conservation of natural populations.

Definition of a population. A population is the total number of individuals of a given biological species found in one place at one time. In practice, ecologists often deal with density—numbers per unit area for land organisms and numbers per unit volume in aquatic systems—or even weight rather than raw numbers. What may be described as an "individual" depends on the kind of organism and the aim of the scientific enquiry. In most animals, the life cycle starts with a fertilized egg, passes through a largely irreversible process of coupled growth and differentiation, and ends in a tightly integrated,

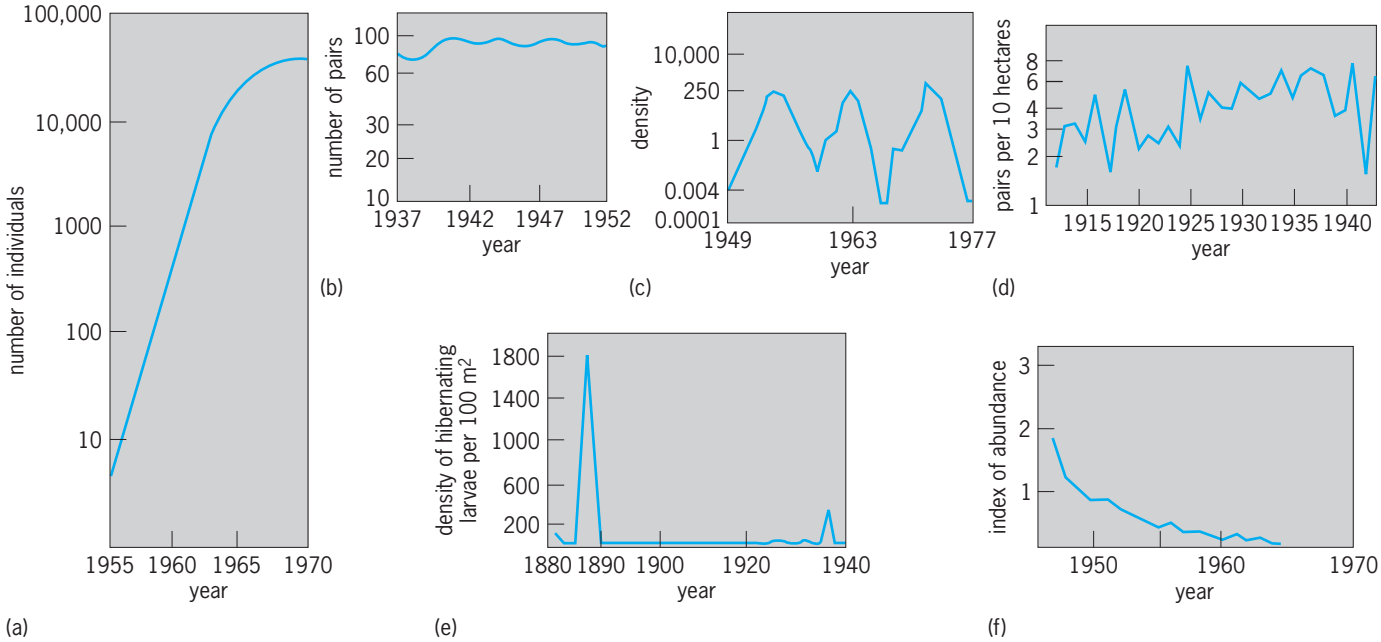

Fig. 1. Graphs of dynamical behaviors showing the diversity of population dynamics. (*a*) Collared dove; a phase of exponential growth asymptoting toward carrying capacity. (*b*) Yellow-eyed penguin; steady state. (*c*) Budmoth; regular cycles. (*d*) Songbird; irregular fluctuations. (*e*) *Dendrolimus* moth; outbreaks. (*f*) Blue whale; decline toward extinction.

unitary, adult organism. Population size can be measured by merely counting adult units and their juvenile prologues. But in most plants and some colonial animals, growth and differentiation proceed in a modular fashion; growth involves the replication of a basic body unit, so that a fertilized egg generates a spatially distributed "population" of connected modules. Modular organisms show tremendous plasticity in size and form. The connections between modules sometimes break, leading to a dispersed clone of physiologically independent units. For example, one clone of quaking aspen (which spreads by root buds) occupies 200 acres (80 hectares) and is more than 10,000 years old. To a geneticist, this clone constitutes a single huge individual; to an ecologist, each aspen trunk may be considered to be an individual. Modular organisms have an additional dimension of complexity that must be quantified to understand their dynamics. In this article, the population concepts presented apply broadly to both unitary and modular organisms.

There are only four ways a population can change in size: birth, death, immigration, and emigration. If immigration and emigration are negligible, the population is closed, and the difference between birth and death rates drives its dynamics. Terrestrial animals on islands often have closed populations. If immigration and emigration are important, however, the population is open, and its abundance may be substantially influenced by spatially distant events. For example, the number of barnacles that are found on a rocky coastline often reflects the density of setting larvae, which in turn is governed by events in offshore waters. If a population that is under study is found to be highly open, the spatial scale of the study may be too narrowly circumscribed to capture the important mechanisms of its population dynamics.

Basic population models. Populations exhibit a great variety of dynamical patterns, ranging from explosive outbreaks, to local extinctions, to regular cycles or relatively constant abundances (**Fig. 1**). To help describe and explain these patterns, ecologists rely on population models. Simple life cycles and closed populations provide a useful starting point in developing population models. Many temperate-zone insects have one annual generation, and so at any given time all individuals are at the same stage of life. For a population with discrete, nonoverlapping generations, if $N(t)$ is the number of adults censused in generation t, and $R(t)$ is the number of adult offspring in generation $t + 1$ produced per adult in generation t, the number of individuals in the next generation is given by Eq. (1). The quantity $R(t)$ is

$$N(t + 1) = N(t)R(t) \tag{1}$$

the growth rate of the population for generation t. Iterating this discrete time growth model for subsequent generations allows one to project population numbers through time. In the limit of very small census intervals, population growth is described by the differential equation (2), where $r(t)$, the instan-

$$\frac{dN}{dt} = r(t)N(t) \tag{2}$$

taneous per-capita growth rate, is the difference between per-capita birth and death rates. Equation (2) is literally true only if populations grow continuously with overlapping generations (as is approximately true for some microbes). Constant values for R or r imply exponential growth. If $R > 1$ ($r > 0$), a population grows without bounds; if $R < 1$ ($r < 0$), it

declines to extinction. The theoretical framework of population ecology largely consists of elaborations of these basic growth models, including extensions to more complicated life cycles and multiple species.

Exponential growth has a snowballing effect: if the growth rate is constant, then the more individuals there are, the faster the population grows. Even low growth rates eventually lead to populations of enormous sizes. Some natural populations show transient phases of exponential growth, particularly in colonizing episodes: for instance, the collared turtledove invaded Great Britain in 1955 and increased exponentially for nearly a decade (Fig. 1*a*). The per-capita rate of growth during exponential growth is called the intrinsic rate of increase, r_0. The value of r_0 quantitatively expresses the interplay of individual traits, such as life history strategies, with the environment.

Exponential growth during colonization often involves expansion across space as well as an increase through time. Equation (2) can be expanded to include immigration and emigration, as given in Eq. (3).

$$\frac{dN}{dt} = rN + (\text{net change due to dispersal}) \qquad (3)$$

When individuals move down spatial density gradients, which is analogous to chemical diffusion along concentration gradients, their rate of movement is characterized by a diffusion coefficient, D. In a homogeneous environment, this model predicts that the expanding population wave advances at a velocity $2(rD)^{1/2}$, implying a linear expansion in range area with time. Data from a number of colonizing populations show exactly this relationship.

Population heterogeneity. Models such as those given in Eqs. (1) and (2), when interpreted literally, assume that all members of a population are identical. This is rarely true. Birth and death rates typically vary as a function of age, body weight, and genotype. A great deal of work in population ecology is devoted to elucidating age-specific schedules of mortality and fecundity, using these patterns to predict population growth, and interpreting these patterns in the light of evolutionary theory. To study age-structured population dynamics, the number of individuals in each age class must be monitored. The two ingredients needed to project changes in population size and age structure are the mortality schedule or survivorship curve, which describes the fraction of newborns surviving to each age, and the fecundity schedule, which describes the rate of female births per female at each age. It is a formidable task to measure complete fecundity and mortality schedules in natural populations, but if these schedules are given, the geometric growth model of Eq. (1) can be generalized to a matrix model, as in Eq. (4), where $N(t)$ is a vector in

$$N(t+1) = A(t)N(t) \qquad (4)$$

which each element is the number of individuals in an age class, and $A(t)$ is a matrix incorporating the fecundity and mortality schedules. When individuals can be categorized into discrete stages such as body size (for example, number of connected modules in a clonal organism) in addition to age, more general matrices can describe population growth; the matrix elements are rates of transition between each pair of stages. [The continuous time model of Eq. (2) can be similarly generalized by using partial differential equations.] A fundamental principle of demographic theory is that if these schedules are constant, a population will (with rare exceptions) eventually settle into a stable age distribution in which each age class comprises a constant fraction of the total population. A population in its stable age distribution grows geometrically, as in Eq. (1), at a rate of increase uniquely determined from the mortality and fecundity schedules. A population displaced from its stable age distribution may exhibit transient phases of growth or decline, divergent from its long-term growth pattern.

All populations are genetically variable. If different genotypes have different fecundity or mortality schedules, genetic variation can influence population dynamics. Models that simultaneously incorporate changes in genetic composition and population growth can be quite complex; however, they may be important in describing some populations and are necessary for linking population ecology with evolutionary theory.

Population limitation and regulation. Populations cannot expand exponentially forever. For instance, the growth rate of the collared dove substantially declined in the second decade of its invasion, probably because mortality rates rose or birth rates declined as a result of competition for limited resources such as food or nest sites. Such mechanisms are called negatively density-dependent factors. The notion of density dependence is a specific example of the more general concept of feedback. An alternative hypothesis is that the environment worsened for reasons unrelated to dove density, such as a shift in the weather. Such causes for variation in birth or death rates are called density-independent factors. Ecologists have long disputed the relative importance of density-dependent and density-independent factors in determining population size. The current consensus is that both are important but to differing degrees in different species and environments. For a population to be regulated, it must tend to increase when below a certain size and decrease when above that size. If growth rates vary with time but in a fashion unrelated to density, closed populations will eventually fluctuate to extinction or expand without limit. If a closed population persists over long periods of time, it must be regulated to some degree.

As illustrated below, population regulation by no means implies population stability. In general, a population is stable if it returns to equilibrium following a perturbation. Moreover, many local populations may not be persistent over long periods of time and so may not be regulated in the usual sense. Open populations, by definition coupled by dispersal with other populations, can become reestablished by immigration following a local extinction. The total population of a species may persist, even though no

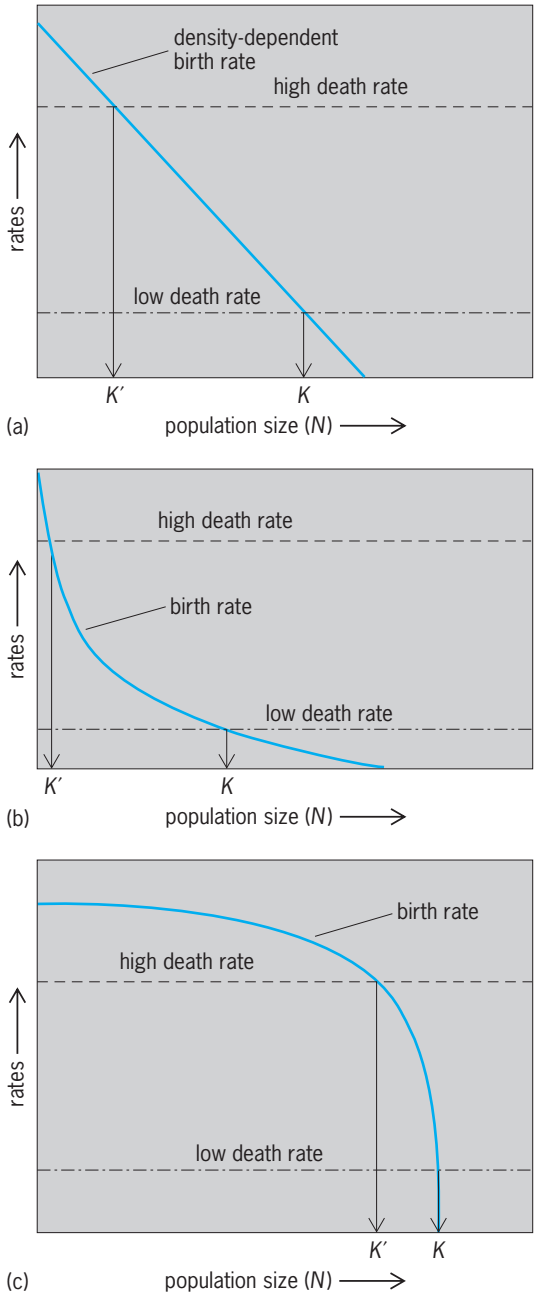

Fig. 2. Carrying capacity as a reflection of both density dependence and the intensity of density-independent mortality. Births are considered to be purely density-dependent, deaths purely density-independent. The three curves show the effects on population size of (*a*) an increase in density-independent mortality (a high death rate), (*b*) weak density dependence (near *K*), and (*c*) strong density dependence (near *K*).

single local population survives, because there is a spreading of risk among an ensemble of local populations (which experience somewhat different environmental conditions) that are loosely coupled by dispersal. *See* POPULATION DISPERSAL.

A useful method for considering the interplay of density-dependent and density-independent factors in determining population size is to plot birth and death rates as functions of density (**Fig. 2**). The carrying capacity K of a population in a given environment is defined to be the largest number of individuals for which the birth rate just matches the death rate. The population decreases above K and increases below K. A given change in density-independent death rates can produce very different changes in population size, depending on the form of the underlying density dependence. If density dependence is weak (Fig. 2*b*), fluctuations in mortality generate large oscillations in population size; if density dependence is strong (Fig. 2*c*), the population readily buffers such fluctuations. Density-dependent factors are necessary to regulate populations, but density-independent factors must also be considered to understand fully what limits populations to a given value of K.

Evidence for density dependence. Statistical analyses of time series of population data can suggest density dependence, but it is difficult to demonstrate density dependence conclusively by using such data. Sometimes density dependence can be shown for particular stages in the life cycle, and in general, density dependence may be observed in births, deaths, or dispersal. However, without examining density dependence at each stage in the life cycle, it is difficult to infer the regulatory importance of any single stage. Stronger evidence comes from manipulative experiments in which control populations are compared with artificially enhanced or depressed populations.

Density-dependent population models. The exponential growth model [Eq. (2)] can be modified to include density dependence by expressing r as a function of N. The simplest model that generalizes Eq. (2) is the logistic equation (5), in which per-

$$\frac{dN}{dt} = r_0 N \left(1 - \frac{N}{K}\right) \tag{5}$$

capita growth rate declines linearly with increasing density. Populations displaced from K converge smoothly back to it, without overshoots. Population growth is maximal when $N = K/2$. This model provides a good fit to some laboratory colonies and captures much of the qualitative behavior of more complex models; it is, however, a somewhat crude first approximation to accurate models of population growth.

The logistic model can be improved upon in numerous ways. The simplest method is to use a nonlinear function for the per-capita growth rate. Further refinements in the model can be achieved by incorporating threshold effects and explicitly incorporating submodels that encapsulate the mechanisms of density dependence. At densities far below carrying capacity, there may be an Allee effect—a positive relationship between density and per-capita growth rates. One intriguing possibility that arises in more complex models is that the population may exhibit alternative stable states; the one it actually occupies will depend upon accidents of history. For instance, insects may be regulated at low densities by bird predation, but at a higher threshold density the birds may be satiated and thus no longer regulate insect

density; the insect population will then grow until checked by some other factor.

The discrete-time model [Eq. (1)] can be similarly modified to incorporate density dependence by expressing R as a function of N. Analysis of such models has led ecologists to reevaluate their traditional assumptions about the causes of population fluctuations in nature and about the relationship between regulation and stability. For instance, an analog of Eq. (5) is Eq. (6), which, along with similar equations, reveals a rich array of dynamical patterns. If $r < 2$, the population equilibrates stably at K; if $2 < r < 2.7$, the population fluctuates cyclically; if $r > 2.7$, the population exhibits chaotic behavior, with cycles of arbitrary periodicity or even aperiodic fluctuations. Such fluctuations are intriguingly similar to the fluctuations in real-world data, which in the past were assumed to be produced by random environmental noise. The qualitative properties of the model that trigger pronounced population fluctuations are the time lag implicit in the discrete-time formulation, high growth rates at low densities, and strong density dependence at high densities. This suggests that a potential for complex dynamical behavior exists whenever there are time lags in the feedback between population size and population growth rates. Age structure is a ubiquitous source of time lags in populations, simply because time is required to reach reproductive maturity. Similarly, interactions between two, three, or more species can introduce long time lags, together with strong density dependence. It is an open question whether observed variability in natural populations reflects to any significant extent the complex dynamics latent in deterministic growth models instead of the force of fluctuations in the physical environment.

$$N(t+1) = N(t)\exp\left[r\left(1 - \frac{N}{K}\right)\right] \qquad (6)$$

Mechanisms of density dependence. Given that density dependence exists, the mechanisms generating it can be used both to predict the consequences of environmental change for population dynamics and to provide insight into systems where experimental manipulations are difficult. Density dependence often arises from competition, which is said to exist when organisms utilize common limiting resources and thereby negatively affect each other. (A resource is limiting if an increase in its supply increases per-capita growth rates.)

There are two principal sorts of competition, interference and exploitative. Interference competition occurs when one individual directly harms another. Interference may be dramatic, as in lethal aggression, or subtle, as when social interactions reduce the time available for gathering resources or increase the risk of predation. A surprising number of animal species are cannibalistic. Large scorpions, for instance, eat with relish their smaller-bodied conspecifics. Because encounter rate increases with increasing population size, cannibalism is likely to be a potent density-dependent factor in scorpion populations. Exploitative competition occurs when one individual consumes a resource such as food that otherwise would have been consumed by another individual. Because exploitative competition is mediated indirectly through a shared resource base, it can be more difficult to demonstrate than interference. In territorial animals, such as many songbirds, less space is available for additional territory holders as population size increases. As a result, competition for space can sharply cap population numbers.

Population regulation and interspecific interactions. Negative density dependence may arise from interspecific interactions. A schematic classification of interactions between two species comes from considering the positive (+) or negative (−) effect that individuals of one species have on the growth rate of the other. In interspecific competition the interaction is (−,−); in mutualism it is (+,+). Natural enemies, defined broadly to include predators, herbivores, and parasites, are often engaged in (+,−) relations with their prey or hosts. Most species are potential prey to one or more natural enemies; even top-level carnivores may be beset by parasites.

Obviously, competitors and mutualists can dramatically affect the size of a given population and so must be considered when studying population limitation. However, natural enemies are far more likely to be regulatory agents than are either competitors or mutualists. If two species are competing and one increases in density, the other will decrease. This will relax the interspecific competition on the first, which can then increase even more. Hence, competitive loops (and for similar reasons, mutualist loops) tend to produce positive feedback and so will not regulate population growth. By contrast, predator–prey interactions may produce negative density dependence acting across several time scales on both the predator and prey. As prey in one habitat patch become more numerous, predators may almost immediately become more active or switch over from other prey types or patches. Predators may also show an intergenerational numerical response to increased prey availability. Because an increase in predator numbers usually decreases prey numbers, this induces delayed density dependence in both the predator and its prey. In host–pathogen systems, these numerical responses may be pronounced, even within a host generation. For example, in tropical trees, seedling mortality from fungal pathogens increases with increasing seedling density, because the rate of spore dispersal increases as the distance between hosts declines. There is mounting evidence that parasites are significant regulatory factors in a wide range of natural populations, including economically important forest insects and game animals.

Several cautionary remarks about natural enemies and density dependence are in order. Predation and intraspecific competition can interact in complex ways. Compensatory responses by prey populations sometimes diminish the impact of enemies. For instance, plants typically produce many more seeds

than can possibly become reproductive adults. Predation on some seeds may simply enhance the survivorship of the nonconsumed seeds, leading to no net effect on adult numbers. Conversely, predation can intensify intraspecific competition in mobile animals by restricting them to limited refuges. Generalist natural enemies can greatly depress prey numbers without being important regulatory agents, because by depending on many prey types they are less likely to show a strong numerical response to any one. Indeed, generalist predators may often be destabilizing, driving local prey populations to extinction. Finally, density-dependent responses by natural enemies often involve time lags, setting up the possibility of oscillatory behavior. Host–pathogen systems seem particularly likely candidates for generating strongly cyclic or chaotic dynamics.

Population growth models can incorporate interspecific interactions by taking models such as those in Eqs. (5) and (6) and adding expressions that describe how competitors, mutualists, or natural enemies affect the growth rate of a given species. The dynamics of two or more coupled species is studied in theoretical community ecology, which among other things seeks to understand how the species richness and stability of communities is related to the pattern and strength of interspecific interactions. Analyses of models of interacting species suggest that strong intraspecific density dependence is required for community stability. For two competing species to persist at a stable equilibrium, the negative effect that each species exerts on its own growth must on average exceed the cross-species negative effects. In like manner, predator–prey interactions are most stable when one or both populations experience intraspecific density dependence.

Applied population ecology. In addition to its intrinsic conceptual appeal, population ecology has great practical utility. Control programs for agricultural pests or human diseases ideally attempt to reduce the intrinsic rate of increase of those organisms to very low values. Analyses of the population dynamics of infectious diseases have successfully guided the development of vaccination programs. In the exploitation of renewable resources, such as in forestry or fisheries biology, population models are required in order to devise sensible harvesting strategies that maximize the sustainable yield extracted from exploited populations. Conservation biology is increasingly concerned with the consequences of habitat fragmentation for species preservation. Population models can help characterize minimum viable population sizes below which a species is vulnerable to rapid extinction, and can help guide the development of interventionist policies to save endangered species. Finally, population ecology must be an integral part of any attempt to bring the world's burgeoning human population into harmonious balance with the environment. *See* ECOLOGY; MATHEMATICAL ECOLOGY; THEORETICAL ECOLOGY.

Robert Holt

Bibliography. M. Begon, J. L. Harper, and C. R. Townsend, *Ecology: Individuals, Populations, Communities*, 3d ed., 1998; M. Gilpin and I. Hanski, *Metapopulation Dynamics: Empirical and Theoretical Investigations*, 1991; C. J. Krebs, *Ecology: The Experimental Analysis of Abundance and Distribution*, 5th ed., 2000; T. Royama, *Analytical Population Dynamics*, 1992; I. A. Shilov, *Population Ecology and Ecophysiology*, 1989; J. W. Silvertown and J. L. Doust, *Introduction to Plant Population Biology*, 1993.

Population genetics

The study of both experimental and theoretical consequences of mendelian heredity on the population level, in contradistinction to classical genetics which deals with the offspring of specified parents on the familial level. The genetics of populations studies the frequencies of genes, genotypes, and phenotypes, and the mating systems. It also studies the forces that may alter the genetic composition of a population in time, such as recurrent mutation, migration, and intermixture between groups, selection resulting from genotypic differential fertility, and the random changes incurred by the sampling process in reproduction from generation to generation. This type of study contributes to an understanding of the elementary step in biological evolution. The principles of population genetics may be applied to plants and to other animals as well as humans. *See* MENDELISM.

Mendelian populations. A mendelian population is a group of individuals who interbreed among themselves according to a certain system of mating and form more or less a breeding community. These individuals share a common gene pool which is the total genic content of the group. A mendelian population is the unit of study in population genetics. The population may be very large or very small, and is to be distinguished from species or varieties, which may consist of numerous isolated or partially isolated mendelian populations. Mendelian population is a genetic rather than a taxonomic term. Mendelian populations differ from each other in their genic content or chromosomal organization, not necessarily in their taxonomic features. The term deme, originally defined as an assemblage of taxonomically closely related individuals, has been used as a synonym for mendelian population. Gamodeme, a deme forming a more or less isolated local intrabreeding community, would be a better substitute.

Mutation pressure. Gene mutation arises from time to time in nature. The causes for mutation are not fully known, and thus it can be said that mutations arise "spontaneously." The effect of a new mutant gene is unpredictable and the gene is therefore said to mutate "at random." One property of mutation has been established: It is recurrent. Each type of gene mutates at a certain rate per generation. The rate is usually low—about 1 mutant in 10^5–10^8 genes of a given sort, varying from locus to locus on the chromosomes, even under uniform conditions. Ionizing radiation, certain chemicals, heat, and some

other agents increase the rate of mutation. *See* MUTATION.

Let μ be the rate of mutation from an allele A to another form a per generation. If a fraction p of the genes of a population is A in one generation, then in the next generation the frequency of A will be diminished by the amount $p\mu$, so that the new frequency of A will be $p(1 - \mu)$. The amount of change, $p\mu$, is said to be due to the mutation pressure. If this pressure is unopposed generation after generation, the gene A will gradually disappear from the population, as $p_n = p_0(1 - \mu)^n \doteq p_0 e^{-n\mu}$, where p_0 is the initial gene frequency and p_n is the frequency after n generations. Therefore, for all existing genes there must be some kind of compensating mechanism which supports its continuing presence in nature. One important problem in population genetics is the mechanism of maintenance of a gene in a population or its change in frequency from generation to generation.

If, in addition to the mutation from A to a, there is reverse mutation from a to A at the rate ν per generation, then the net amount of change in the frequency of a is $\Delta q = p\mu - q\nu$. At the time when these opposing changes cancel each other, there will be no change in gene frequency despite the recurrent mutations. This state of affairs is said to be in equilibrium and is obtained when $\Delta q = 0$; that is, $\hat{p} = \nu/(\mu + \nu)$ and $\hat{q} = \mu/(\mu + \nu)$, where q is a frequency of a, $\hat{q}$ is the equilibrium point for a, and $\hat{p}$ is the equilibrium point for A. The equilibrium gene frequencies are determined by the opposing rates of mutation only and are independent of the initial frequencies of the genes in the population. The amount of change in gene frequency per generation is larger when the current q is far away from the equilibrium $\hat{q}$ than when q is close to $\hat{q}$. Substitution gives $\Delta q = -(\mu + \nu)(q - \hat{q})$, indicating that the amount of change per generation is proportional to the deviation $(q - \hat{q})$. It also shows that if $q > \hat{q}$, q decreases, and if $q < \hat{q}$, q increases, or that q will approach $\hat{q}$ from either side. Such an equilibrium is said to be stable. The changes in q described above are independent of the mating system practiced in the population.

In nature, and under artificial conditions, the mutation rates may not remain constant in all generations but may fluctuate within a certain range from time to time. In such cases, instead of a single fixed equilibrium point $\hat{q}$, there will be an equilibrium distribution of q within a certain range, and the apparent change in gene frequency from one generation to the next may be purely a stochastic phenomenon without necessarily having long-term significance. The same remark applies to all equilibria to be established in subsequent paragraphs. *See* STOCHASTIC PROCESS.

Migration and intermixture. If a fraction m of a population with a gene frequency q consists of immigrants from outside and the immigrant group has a gene frequency $\overline{q}$, the new gene frequency of the population will be $q_1 = (1 - m)q + m\overline{q} = q - m(q - \overline{q})$. The amount of change in gene frequency in one generation is thus $\Delta q = q_1 - q = -m(q - \overline{q})$, showing that the change is proportional to the deviation $(q - \overline{q})$. This expression for Δq is of the same form as that for mutation. If the immigrants have the same gene frequency as the population, there will be no change in gene frequency in spite of the migrations. The continued intermixture of neighboring populations will eventually make them homogeneous in terms of gene frequencies. Thus, if a large population is divided into a number of partially isolated subpopulations, migrations between the groups will eventually make all subpopulations have the same gene frequency $\bar{q}$, which then denotes the average for the entire population in the absence of other disturbing factors. If the local populations are differentiated genetically, there must be some mechanism (for example, local selection) to counteract the pooling effect of migrations so that an equilibrium condition may be reached. The change in gene frequency due to migration is independent of the mating system practiced in the population.

Mating systems. In a gene pool with respect to one locus, if a proportion p of the genes is A and a proportion q of the genes is a, the genotypic proportions in the population are still unknown until the mating pattern is specified. The mating pattern is a system by which the genes are associated into pairs to form the diploid genotypes. The mating systems vary widely in nature for different organisms and populations. Thus, wheat may have 1% cross-pollination and 99% self-fertilization, whereas maize practices just the reverse. One of the simplest and most extensively studied systems is random mating, also known as panmixis.

Panmixis. Random mating between individuals is equivalent to a random union of gametes. Thus, if the (pA, ga) gametes of one sex unite at random with the (pA, qa) gametes of the opposite sex, the resulting genotypic array will be p^2AA, $2pqAa$, q^2aa. This principle was first discovered by W. E. Castle in 1903. These genotypic proportions will be realized only in very large populations. *See* HARDY-WEINBERG FORMULA; HUMAN GENETICS.

Inbreeding. Inbreeding refers to mating between genetically related individuals; the frequency with which two A gametes unite will be greater than p^2; and a similar situation is true for gene a. Consequently, inbreeding leads to an increase of homozygosis at the expense of heterozygosis. Let H be the heterozygosis proportion in a population, H' that in the preceding generation, H' that two generations ago, and so on. On continued systematic inbreeding, the manner in which the value of H decreases is shown in **Table 1**.

Continued close inbreeding, such as those degrees indicated in Table 1 and many others, eventually leads to complete homozygosis. The population will then consist of pAA and $q\ aa$; that is, an A gamete will always unite with another A gamete, and an a gamete with another a. Inbreeding between remote relatives does not necessarily lead to complete homozygosis but only decreases the heterozygosis below the random mating level to a certain extent.

TABLE 1. Decrease in heterozygosis with systematic inbreeding

Inbreeding system	Manner of heterozygosis (H) decrease	Limiting situation
Self-fertilization	$H = \frac{1}{2}H'$	$H = 0.500H'$
Same purebred sire × successive daughters	$H = \frac{1}{2}H'$	$H = 0.500H'$
Brother × sister	$H = \frac{1}{2}H' + \frac{1}{4}H''$	$H = 0.809H'$
Younger parent × offspring	$H = \frac{1}{2}H' + \frac{1}{4}H''$	$H = 0.809H'$
Half brother × half sisters	$H = \frac{3}{4}H' + \frac{1}{8}H''$	$H = 0.890H'$
Half brother × full sisters	$H = \frac{1}{2}H' + \frac{1}{4}H'' + \frac{1}{16}H'''$	$H = 0.870H'$
Double first cousins	$H = \frac{1}{2}H + \frac{1}{4}H'' + \frac{1}{8}H'''$	$H = 0.920H'$

The inbreeding coefficient is an index intended to measure the amount or degree of inbreeding that has been accomplished in a population. Various indices may be constructed. One that has been proved highly useful in both theoretical investigation and practical breeding work is the inbreeding coefficient F defined as the correlation coefficient between the uniting gametes. The value of F ranges from 0 for random mating to 1 for inbreeding in a homozygous populations, as shown below.

Random mating:

	A	a	
A	p^2	pq	p
a	pq	q^2	q
	p	q	1

Correlation = 0

Inbred population:

	A	a	
A	$p^2 + Fpq$	$pq - Fpq$	p
a	$pq - Fpq$	$q^2 + Fpq$	q
	p	q	1

Correlation = F

Inbreeding in homozygous population:

	A	a	
A	p	0	p
a	0	q	q
	p	q	1

Correlation = 1

In an inbred population where the correlation between the uniting gametes is F, the genotypic array in the population will be as in Eqs. (1)–(3). These equations show that the population may be mathematically considered as having two separate components, $(1 - F)$ panmictic and F fixed. If the mating system is such that F remains constant (instead of increasing) from generation to generation, the population will reach an equilibrium state with the genotypic array shown above.

$$AAp^2 + Fpq = (1 - F)p^2 + Fp \quad (1)$$

$$Aa2pq - 2Fpq = 2(1 - F)pq \quad (2)$$

$$aaq^2 + Fpq = (1 - F)q^2 + Fq \quad (3)$$

The correlation between uniting gametes is due to the correlation between mating individuals. In an equilibrium population, if M denotes the correlation between mates, then $M = 2F/(1 + F)$ or $F = M/(2 - M)$.

Genotype selective values. Within a large population not all individuals produce the same number of offspring. In the situation to be considered, the average number of living offspring born to each of the genotypes in the population is studied, while the random fluctuation in the number of offspring from family to family is ignored. Furthermore, it is assumed that the population is so large that only the relative frequencies of the various genotypes and genes in the population are of interest. Suppose that the average number of offspring for each genotype is as follows: AA: 2.00; Aa: 2.50; aa: 1.50. Given these differential rates of reproduction, the new gene frequency of the next generation may be calculated. Inasmuch as it is only their relative magnitude that matters, these reproductive rates may be simplified into the ratio $W_{11}:W_{12}:W_{22} = 1:1.25:0.75$ or alternatively into 0.80:1:0.60. In order to standardize the description, it is convenient to take one of the three reproductive values as unity. In the previous example, depending upon whether the reproductive value of AA or Aa is taken as unity (the standard), that of aa is $0.75 = 1 - 0.25$ or $0.60 = 1 - 0.40 = 1 - s$ in general. The value of s is known as the selection coefficient against the genotype aa. When a selection coefficient is used, it should always be stated which genotype has been employed as the standard.

Natural selection. The doctrine of the survival of the fittest needs clarification from the genetic viewpoint. The relative genotypic reproductive values (W_{11}, W_{12}, W_{22} of the preceding paragraph) simply given an ex post facto description, by which the genetic composition of the offspring generation may be related to that of the parent generation. These W values include all causes for differential reproduction such as fecundity, fertility, sexual maturity and capacity, survival and viability, length of reproductive life, and many others depending on the details of the life cycle of the organism. The W value, sometimes briefly referred to as relative fitness, is not necessarily correlated with any observable morphological characteristics, no matter how desirable they may seem to humans. From the genetic viewpoint, only those who reproduce count. Thus, natural selection has no

TABLE 2. Effect of selection in a random mating population

Genotype	Frequency, f	Fitness, W	Frequency after selection, fW
AA	p^2	W_{11}	p^2W_{11}
Aa	$2pq$	W_{12}	$2pqW_{12}$
aa	q^2	W_{22}	q^2W_{22}
Total	1.00		$\overline{W}$

particular purpose except to perpetuate those who are fit to reproduce under the given conditions. Only when a characteristic lowers the organism's reproductive capacity does it have a genetic effect on the subsequent generations.

Selection pressure and equilibrium. The effect of selection may be described in terms of changes in gene frequency. In a random mating population with respect to one gene locus, the situation is as shown in **Table 2**.

The population after selection is the parental population of the next generation through random mating. The value $\bar{W}$ is the total of the selected parental population, but may also be regarded as the average fitness of the original unselected population. The new frequency of gene a among the selected is $q' = (pqW_{12} + q^2W_{22})/\overline{W}$ and therefore the amount of change per generations is $\Delta q = q' - q$, or more explicitly, Eq. (4). This represents the effect of se-

$$\Delta q = \frac{pq}{2\overline{W}} - \frac{d\overline{W}}{dq} \tag{4}$$

lection pressure on gene frequency. When p or q is zero, there is no change in gene frequency; there can be no selection in the absence of alternatives. Therefore, all selection effects involve the factor pq. Further, when p or q is very small, the selection is ineffective whether it is for or against a gene. Besides these terminal conditions, if there exists a q value such that $\Delta q = 0$, it is called the equilibrium value of gene frequency, because a population with that particular gene frequency will remain unchanged in spite of the selection pressure. When such a q value exists, it must be the solution of the relations $d\overline{W}/dq = 0$ shown in Eq. (5). In order that $\hat{q}$ be

$$\hat{q} = \frac{W_{11} - W_{12}}{(W_{11} - W_{12}) + (W_{22} - W_{12})} \tag{5}$$

a positive fraction, the differences $W_{11} - W_{12}$ and $W_{22} - W_{12}$ must be both positive or both negative; that is, the selective value of the heterozygote must be lower or higher than those of both homozygotes. For all other cases, there will be no equilibrium except when $q = 0$ or 1.

Stability of an equilibrium. The value of $\overline{W} = p^2W_{11} + 2pqW_{12} + q^2W_{22}$ may be plotted against the value of q or p. When W_{12} is greater than W_{11} and W_{22}, The $\overline{W}$ curve has a maximum point (**Fig. 1**). The q value corresponding to the maximum value of $\overline{W}$ is the stable equilibrium point. This means that whether it is smaller or larger than $\hat{q}$, the q value will approach $\hat{q}$ as selection proceeds from generation to generation. A stable equilibrium of this type leads to balanced genetic polymorphism, that is, to the coexistence of alleles in a population. Conversely, if q_{11} is lower than both W_{11} and W_{22}, the $\overline{W}$ curve has a minimum point yielding an unstable equilibrium; the selection pressure will make the q value move away from the equilibrium value toward either 0 or 1, depending upon which side of the equilibrium the q happens to be. Consequently, selection against the heterozygote leads to the elimination of one of the alleles. In more complicated situations, there could be more than one stable or unstable equilibrium value in a population, or both.

The genotype selective values W_{11}, W_{12}, W_{22} have been assumed to be fixed for each genotype, but in nature they may vary in a number of ways. In addition to the omnipresent random fluctuations, the selective values may vary with the gene frequency itself. For instance, a genotype favored by selection when it is rare in the population may suffer a disadvantage when it is too common. For such cases, there will be an equilibrium yielding genetic polymorphism. Let U_{11}, a function of q, be the varying selective value of genotype AA, and so on. Then the equilibrium value of gene frequency is given by the appropriate solution of Eq. (6). The study of selection effects may be

$$q = \frac{U_{11} - U_{12}}{(U_{11} - U_{12}) + (U_{22} - U_{12})} \tag{6}$$

extended to cases with multiple alleles, sex-linked alleles, autopolyploids, and inbreeding populations.

Gamete selection. The effective rate at which the A and a gametes function may not be the same; that is, selection may operate in the gametic stage instead of in the diploid genotypic stage. If the selective actions for the genotype aa, for example, and gamete a are in opposite directions, an equilibrium may result.

Balance between selection and mutation. There are many different types of genotypic selection. Two simple cases will illustrate the principle of balance between selection and mutation pressures.

Selection against recessives. Suppose that the selective values of AA, Aa, and aa are 1, 1, and $1 - s$, where s is a positive fraction known as the selection coefficient. Then the new gene frequency will be $q' = (q - sq^2)/(1 - q^2)$ in the next generation, so that the amount of change per generation is $q' - q = -sq^2(1 - q)/(1 - sq^2)$. At the same time, if μ is the

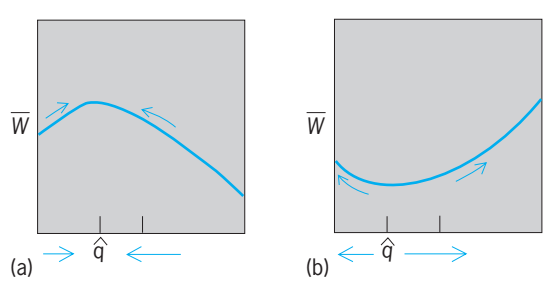

Fig. 1. Diagram of the relationship between the average fitness value $\bar{W}$ and $\hat{q}$, the equilibrium point for *a*. (*a*) Stable equilibrium. (*b*) Unstable equilibrium.

mutation rate from A to a, the value of q will be increased by the amount $\mu(1 - q)$ per generation. At equilibrium, the forces to increase and to decrease the gene frequency must cancel each other; that is, $\mu(1 - q) = sq^2(1 - q)/(1 - sq^2)$. By solving, $sq^2 = \mu/(1 + \mu)$ is obtained, which closely approximates μ. Hence, $a^2 = \mu/s$ and $q = \sqrt{\mu/s}$, which usually is a small quantity. When aa is lethal or unable to reproduce, $s = 1$ and $q = \sqrt{\mu}$. This explains the persistence of deleterious recessive genes in a population in spite of continuous selection.

Selection against dominants. If the selection is against homozygous dominants only, the situation is the same as in the previous instance except for substitution of p for q and ν (mutation rate from a to A) for μ. To bring out the distinction between selection against dominants and that against recessives, take the extreme case in which AA is lethal, the selective value of Aa is $1 - s$, and aa is the norm. The value of p will then be so low that the usual genotypic proportions p^2, $2pq$, q^2 will take the limiting form 0, $2p$, $1 - 2p$, as q is very close to unity. The increase in q through selection is approximately sp, whereas the loss through mutation from a to A is $q\nu = \nu$. Hence, at equilibrium, $p = \nu/s$. This value is much lower than $q = \sqrt{\mu/s}$ for selection against recessives. Thus, selection against dominant alleles is more effective than selection of the same intensity against recessives. Selection against heterozygotes will eventually lead to the same limiting situation. All the equilibrium values supported by mutation pressure are low but stable. Mutations prevent complete extinction of an allele.

Random drift. The random drift of gene frequencies in finite populations is often called the Sewall Wright effect because of his analysis of its significance. The gene frequency of any generation is determined by the uniting gametes produced by the parents of the preceding generation. If the number of parents is limited and constitutes a random sample of the entire population, the gene frequency of the next generation will not remain exactly the same as that of the previous generation but will be subject to a random fluctuation on account of the sampling process. In a random mating population of N individuals, one-half of whom are males and one-half females, and maintaining the same population size, the variance of the gene frequency based on $2N$ gametes is $q(1 - q)/2N$. The gene frequency may become a little higher or a little lower in the following generation. The smaller the population, the greater is the variance. This random process will continue to operate in all generations. In a sufficiently long time, the value of q will reach either the terminal value 0 or 1. Hence the random drift leads eventually to complete homozygosis for small populations. It can be shown that the limiting rate of reaching the state 0 or 1 is each $^1/_4$ per generation, so that the total rate of "decay" of genetic variability is $^1/_2N$ per generation. Naturalists have found numerous small isolated colonies (for example, snails in mountain valleys) with characteristics uncorrelated with the environmental conditions to substantiate the theory of random (nonadaptive) fixation.

The effective size of a population is the actual number of individuals producing offspring and thereby responsible for the genetic constitution of the next generation. The random mating population with one-half males and one-half females, and producing the same number of offspring, is an idealized model. Any deviation from the ideal situation will have a different sampling variance and a different rate of decay. Equating these to the "standard" variance $q(1 - q)/2N$ or the ideal decay rate $^1/_2N$, an equivalent N is obtained for the ideal population. The latter number is known as the effective size of a population. It is convenient to use in mathematical descriptions of the genetic behavior of a population. Some of the factors that tend to make the effective size smaller than the actual breeding size are given below.

1. Unequal number of males and females. If M and F are the respective numbers, the effective size N_e is not simply $M + F$ but is defined by $1/N_e = {}^1/_4M + {}^1/_4F$ and is equal to $N_e = 4MF/(M + F)$. The larger the difference between M and F, the smaller the number N_e as compared with $M + F$.

2. Unequal size of families. If the gametes are drawn wholly at random from the parents, the number of gametes k contributed by a parent will form a Poisson distribution. In such a case, the effective size is the same as the actual breeding size. However, without perfect random sampling, if the mean number of gametes per parent is $\bar{k} = 2$, the effective size is equal to N_e as expressed in Eq. (7), where σ is presumably larger than 2.

$$N_e = \frac{4N - 2}{\sigma_k^2 + 2} \tag{7}$$

3. Inbreeding. If F is the inbreeding coefficient of a population, then the effective size is $N_e = N/(1 + F)$.

Periodic change in population size. If $N_1, N_2, \ldots, N_t$ are the respective sizes of the t generations, the average effective size for the period is approximately equal to the harmonic mean of the t sizes. The harmonic mean is much closer to the smallest number of a series than to the largest one.

Distribution. A stationary distribution of gene frequencies results from two opposing forces: the systematic pressures (mutation, migration, selection) which tend to make the gene frequency attain a certain fixed value, and the random variation due to sampling which tends to make the gene frequency drift away from any fixed value. The result of these opposing tendencies is not a single equilibrium value of gene frequency but a stationary distribution of gene frequencies. This distribution may be viewed in three different ways: as the distribution of q for a particular locus in a population in a long period of time; as the distribution of the allelic frequencies of all loci subject to the same pressures in one population at any given time; and finally as the distribution of q of one locus among a large number of populations of the same size and with the same pressures at a given time. *See* ALLELE.

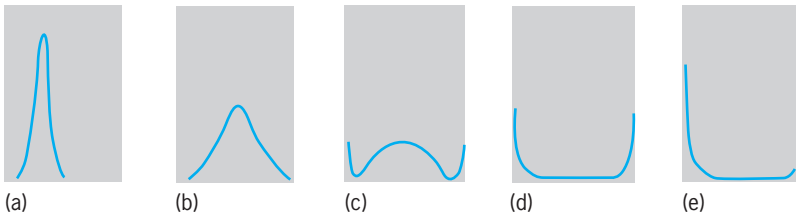

Fig. 2. Distribution of gene frequency in various populations under differing conditions. (*a*) Large population, distribution clustered close to equilibrium value. (*b*) Intermediate population, distribution of wider spread. (*c*) Intermediate small population, distribution with some terminal fixation. (*d*) Small population, most genes close to fixation or extinction in the absence of systematic pressures. (*e*) Small population with strong selection pressure.

Under selection pressure and the sampling variation, the distribution function $\phi(q)$ is $\phi(q) = C\bar{W}^{2N}/q(1-q)$, where C is a constant, $\bar{W}$ the average fitness of the population, and N the effective size of the population. The exact form of the distribution depends upon the value of $\bar{W}$, which is a function of the selection coefficients as well as the gene frequency.

If there is mutation, or migration pressure, or both, the distribution of gene frequency is simply a β distribution: $\phi(q) = Cq^{U-1}(1-q)^{V-1}$, where $U = 4N\mu$, $V = 4N\nu$ if there is only mutation pressure; $U = 4Nm\bar{q}$, $V = 4Nm\bar{p}$ if there is only migration pressure; and $U = 4N(\mu + m\bar{q})$, $V = 4N(\nu + m\bar{p})$ if there are both. The exact form of the β distribution depends upon the values of U and V. When they are smaller than unity, the population is considered small; when they are close to unity, the population is intermediate in size; when they are much larger than unity, the population is considered large. When the effects of mutation, migration, and selection are combined, the distribution function is as shown in Eq. (8). **Figure 2** shows some of the forms of the

$$\phi(q) = C\overline{W}^{2N} q^{U-1}(1-q)^{V-1} \tag{8}$$

distribution under various conditions. The joint distribution of more than one locus is naturally very complicated. Furthermore, if the mutation rates and selection coefficients also vary instead of being constant, the mathematical description of the stochastic process becomes very laborious and the forms shown in Fig. 2 give only a first approximation to the real situation.

The distribution forms of the gene frequencies depend upon the relative magnitudes of the various factors which bring about populational changes. In large random-mating populations all gene frequencies remain close to their stable equilibrium values, which are determined by the counteracting but systematic pressures of mutation, selection, and migration. There will be no further genetic change unless the environmental conditions change so as to define new equilibrium points. Evolution is such large populations is guided essentially by intragroup selection, and progress is very slow.

In small and completely isolated populations, most of the gene frequencies are close to 0 or 1 because of the random drift process which dominates the situation. Selection is ineffective. The loci are prevented from being completely fixed only by occasional mutations or immigrants. The ultimate fate of such small homozygous populations is probably extinction because they are nonadaptive and unable to respond to new conditions.

In populations of intermediate size, all factors, both random and systematic, come into play and the population is more responsive to evolutionary change. If a large population is subdivided into many partially isolated groups with migrations between them, there will be some differentiation among the groups, some of it adaptive and some nonadaptive, but there is very little fixation. The selection effect, varying from one locality to another, then operates largely on an intergroup basis which is more efficient than the intragroup selection within one single large population. If the groups are small, some of them will be eliminated by selection while others flourish. This provides the most favorable condition for evolutionary success for the species as a whole. The conclusion is that there is no one all-important factor in evolution. Evolutionary advance depends upon the interplay and balance of all factors.

Two-locus selection. All of the properties described above apply strictly to one locus, or to independently distributed loci. When selection acts upon a two-locus genotype directly, the genes of two loci, linked or unlinked, are in general not independently distributed, even in a large random mating population. In selection involving more than one locus, it is inadequate to consider the gene frequencies at the various loci separately and it is necessary to consider the gamente frequencies. As the simplest example, consider two unlinked loci (A,a and B,b). Let the four gamete frequencies (x) be as follows:

	B	b	
A	x_1	x_2	p
a	x_3	x_4	q
	u	v	1

Random union of these gametes yields the following offspring genotype frequencies:

	BB	Bb	bb	
AA	x_1^2	$2x_1x_2$	x_2^2	p_2
Aa	$2x_1x_3$	$2(x_1x_4 + x_2x_3)$	$2x_2x_4$	$2pq$
aa	x_3^2	$2x_3x_4$	x_4^2	q_2
	u^2	$2uv$	v^2	1

Suppose that the nine genotypes have the relative fitness values (W) indicated below:

		BB	Bb	bb
	AA	1	1	0
W:	Aa	1	2	1
	aa	0	1	1

The symmetry of the selection pattern ensures the

gene frequencies to be $p = q = {}^1/_2$ at the (A,a) locus and $u = v = {}^1/_2$ at the (B,b) locus. But it may be verified that the equilibrium gamete frequencies are

	B	b	
A	.28493	.21508	.50
a	.21508	.28492	.50
	.50000	.50000	1.00

and not

	B	b	
A	.25	.25	.50
a	.25	.25	.50
	.50	.50	1.00

The alleles of the two loci are positively correlated. The deviation from random distribution is $D = .28492 - .25000 = .03492$. If these two loci were linked, the deviation D would be even more pronounced. Thus, selection leads to association (positive or negative) among the alleles of different loci. Also, the ultimate result of selection is not necessarily the maximization of the average fitness ($\overline{W}$) of a random mating population as is the case with only one locus (Fig. 1). The lack of independence of genes at different loci renders the problem of natural selection and its evolutionary consequences far more complicated than is depicted by simple mathematical models. *See* BIOMETRICS; GENETICS; MUTATION.

C. C. Li

Bibliography. G. DeJong, *Population Genetics and Evolution*, 1988; J. S. Gale, *Theoretical Population Genetics*, 1989; J. B. S. Haldane, *The Causes of Evolution*, 1966, reprint 1990; N. D. Henderson and R. Benno (eds.), *Developmental Behavior Genetics: Neural, Biometrical and Evolutionary Approaches*, 1990; T. Nagylaki, *Introduction to Theoretical Population Genetics*, 1992.

Population viability

The ability of a population to persist and to avoid extinction. The viability of a population will increase or decrease in response to changes in the rates of birth, death, and growth of individuals. In natural populations, these rates are not stable, but undergo fluctuations due to external forces such as hurricanes and introduced species, and internal forces such as competition and genetic composition. Such factors can drive populations to extinction if they are severe or if several detrimental events occur before the population can recover. The concept of population viability gained importance with the National Forest Management Act of 1976, which charged the U.S. Forest Service with maintaining "viable populations" of all species of vertebrates found on its properties. *See* ECOLOGY; POPULATION ECOLOGY.

Factors affecting population viability. The viability of a population depends on the number of individuals in the population and on how the population size is changing.

Intrinsic rate of increase. Simple deterministic models can be used to assess how well a population is doing. In these models, the current number of individuals in the population and the expected number of births, deaths, immigrants, and emigrants determine the number of individuals found in the future. To evaluate a population's viability, ecologists calculate the intrinsic rate of increase (r), which is the change in population size per individual per unit time. When $r > 0$, the population is increasing in size, and when $r < 0$, it is decreasing in size. Viable populations are those with intrinsic rates of increase greater than or equal to 0. *See* DETERMINISM.

Stochasticity. In natural populations, viability depends both on the intrinsic rate of increase and on the sources and magnitude of stochasticity (chance variation) in this growth rate. Three different causes of stochasticity are usually recognized: environmental, catastrophic, and demographic.

1. Chance variation in a population's growth rate caused by random variation in temperature, rainfall, competitors, or other conditions is called environmental stochasticity. As a result of these factors, the population may grow faster or slower than the long-term average. The variation may be due to small-scale fluctuations in, for example, resources or predators, and runs of "bad" or "good" years can occur randomly.

2. Chance variation in a population's growth rate resulting from large-scale or severe disturbances such as hurricanes, fires, or floods is known as catastrophic stochasticity. In general, it involves factors that reduce a population's size. For example, in 1994 the golden toad went extinct because the ephemeral breeding pools in Costa Rica where the toads normally develop, dried up before metamorphosis took place.

3. Chance variation in a population's growth rate that arises from random variation in birth and death rates is known as demographic stochasticity. Traits such as survival and reproduction are not perfectly predictable. For example, a bird may have a 50% probability of producing female offspring, but 50% of her offspring will not necessarily be female. This example of predicted sex ratios had real consequences for the dusky seaside sparrow—all five remaining individuals were males, so the species went extinct in 1987. Demographic stochasticity is especially a problem for small populations, in which low population density can result in lower rates of mating and, therefore, lower birth rates and population growth. *See* ECOLOGICAL SUCCESSION; POPULATION DISPERSION; STOCHASTIC PROCESS.

Genetic variability. In addition to the sources of variability above, changes in the frequency, number, or distribution of genes in a population may put populations at risk of extinction. When populations are small, genetic drift (changes in genetic diversity due to chance mating and survival) can result in the loss of genetic diversity. Low genetic diversity often produces a number of deleterious effects, including inbreeding depression and a lack of potential for evolutionary response to changing environmental conditions. One example is the well-studied lion

population from the Ngorongoro Crater in Tanzania. This population was dramatically reduced by biting flies (catastrophic stochasticity), and the outcome was a dramatic loss of genetic diversity and increased inbreeding. As a result, reproductive success has declined, and males show high levels of sperm abnormalities.

Population viability analysis. One of the most important uses of population viability models comes from modern conservation biology, which uses these models to determine whether a population is in danger of extinction. This is called population viability analysis (PVA) and consists of demographic and genetic models that are used to make decisions on how to manage populations of threatened or endangered species. The National Research Council has called population viability analysis "the cornerstone, the obligatory tool by which recovery objectives and criteria [for endangered species] are identified." *See* ECOLOGICAL MODELING.

Population viability analysis treats extinction as a random process and therefore quantifies viability in terms of the probability of extinction. The viability of a population can be evaluated in two principal ways. Scientists can estimate (1) the mean time to extinction of a population of a given size, given specified stochastic factors, and (2) the minimum population size necessary for a population to persist for a specified period of time (for example, the population size necessary to be 95% confident that the population will persist for 100 years). This population estimate is referred to as the minimum viable population size (MVP). The MVP is the smallest population size that will persist for some specified length of time with a specified probability. In practice an MVP is defined as the smallest population size that has a 95% probability of persisting for 100 years.

The earliest population viability models were based on principles of population genetics. Scientists realized that the number of individuals present in a population was not equal to the number of individuals breeding in the population. Most estimates of the number of individuals breeding in a population suggest it is usually 10–33% of the true population size. The difference between the true and breeding population sizes can be as simple as the difference between the number of reproductively active individuals and the total number of individuals, or it can be due to more subtle factors such as variation in population size, sex ratio, or among-individual variance in reproduction. Early population-genetic models predicted that a population size of 50 would reduce the inbreeding rate to 1% per generation and would maintain genetic variation over the short term, whereas a population size of 500 would be large enough to maintain genetic variation in the population over the long term. These early benchmarks have been roundly criticized for a number of reasons, but particularly because they are based on the genetic data from a single population of fruit flies that may not be representative of the genetics of other species. *See* POPULATION GENETICS.

Deterministic models had been used since the early 1970s in the management of endangered species. The first true application of PVA was with stochastic models developed to investigate grizzly bear management in Yellowstone National Park. M. L. Shaffer developed a stochastic population simulation that incorporated chance events (demographic and environmental stochasticity) and produced estimates of extinction and population sizes needed to prevent extinction. *See* ENDANGERED SPECIES.

The components of PVA are not strictly defined; each PVA must be tailored to the biology of the study species. Therefore, it must incorporate the best available information on the study species. For example, many small organisms are restricted to a few habitat patches, but within those patches they often have high population densities. For these species, PVAs should focus on environmental and catastrophic stochasticity. A good PVA is designed to determine (1) whether the population is viable in both the short term and long term, (2) the factors that are currently putting the population at risk, and (3) how these risks can be reduced or eliminated so that the population can both survive and recover.

Population risk analysis. Over the past decade, the use of PVA for population risk analysis has become accepted. The development of easy-to-use computer programs has also contributed to the proliferation of PVA models. Although use of these models has increased, the method has come under increasing criticism. The two main sources of criticism are the data used to calibrate the models, and the methods used to confirm and structure models. For example, demographic data, the principal inputs of the models, are often based on studies too limited in duration to reasonably estimate the variance in birth rates and survival due to environmental or catastrophic stochasticity.

Population viability analysis has both strengths and weaknesses. It does allow a prediction of how long a population of a given size will persist, given the magnitude and frequency of different stochastic disturbances, which can, in turn, be useful in making predictions for endangered species or for designing preserves. However, any given population may go extinct within the first year, or it may persist for over 10 years. Therefore, by its very nature, the prediction of stochastic behavior of a particular population may not be very accurate.

Despite these methodological concerns, population viability will remain an important concept in conservation biology and population ecology. Habitat loss, habitat fragmentation, introduced species, and other recent environmental changes have resulted in the loss or serious decline in a number of plant and animal populations. Our ability to preserve and protect these species will depend on understanding the forces that determine population viability. Gretchen LeBuhn; Thomas E. Miller

Bibliography. M. S. Boyce, Population viability analysis, *Annu. Rev. Ecol. Systemat.*, 23:481–506, 1992; W. Morris et al., *A Practical Handbook for*

Population Viability Analysis, Nature Conservancy, Arlington, VA, 1999; M. L. Shaffer, Minimum population sizes for species conservation, *BioScience*, 31:131–134, 1981; M. E. Soulé, *Viable Populations for Conservation*, Cambridge University Press, 1987; B. A. Wilcox, Extinction models and conservation, *Trends Ecol. Evol.*, 1:46–48, 1986.

Porcelain

A high-grade ceramic ware characterized by high strength, a white color (under the glaze), very low absorption, good translucency, and a hard glaze. Equivalent terms are European porcelain, hard porcelain, true porcelain, and hard paste porcelain. *See* GLAZING; POTTERY.

Porcelain is distinguished from other fine ceramic ware, such as china, by the fact that the firing of the unglazed ware (the bisque firing) is done at a lower temperature (1800–2200°F or 1000–1200°C) than the final or glost firing, which may be as high as 2700°F (1500°C). In other words, the ware reaches its final state of maturity at the maturing temperature of the glaze.

The white color is obtained by using very pure white-firing kaolin or china clay and other pure materials, the low absorption results from the high firing temperature, and the translucency results from the glass phase.

The term porcelain has been applied to such items as electrical insulators and bathroom fixtures. Very often these are made in a one-fire process, the glaze being applied to the green or unfired ware; where this is the case and high-grade materials are used in compounding the body, the term porcelain may be correctly applied. However, the pieces have no translucency because of their great thickness. On the other hand, the term porcelain is often applied to quite different ware. For example, zircon porcelain is used to describe a material made largely of zircon ($ZrSiO_4$), with small amounts of fluxes to yield a low absorption. *See* CERAMICS.

J. F. McMahon

Bibliography. J. Doherty, *Porcelain*, 2002; J. S. Reed, *Principles of Ceramics Processing*, 2d ed., 1995.

Porcupine

A mammal with spines or quills in addition to regular hair. The 23 species are classified in the order Rodentia and consist of two families: the Hystricidae, or Old World porcupines (11 species), and the Erethizontidae, or New World porcupines (12 species). Although traditionally considered to be related, the two families of porcupines may have evolved independently from different, unrelated ancestors, The infraorbital foramen is usually larger than the foramen magnum, and the zygomatic arch is also very large. Members of both families have a dental formula of I 1/1, C 0/0, Pm 1/1, M 3/3 × 2 = 20. Porcupines are herbivorous, terrestrial, and generally nocturnal. *See* DENTITION; MAMMALIA; RODENTIA.

Quills. Quills are present on the back, sides, and tail of porcupines. Each quill consists of long sharp bristles of hair that are fused together. Porcupines defend themselves by striking attackers with their quilled tails. When a porcupine is disturbed, it rattles the quills on its tail. These produce a warning signal that resembles the noise created by a rattlesnake's tail. If the sound does not make the would-be aggressor retreat, the porcupine turns its back on it, erects its rapierlike quills, and charges backward. The quills come out easily and stick in the attacker's flesh. The porcupine grows new quills to replace the ones lost. Porcupines cannot "shoot" quills at their enemies. In New World porcupines, the tip of each quill is covered with tiny backward-pointing projections called barbs. The barbs hook into the flesh, and the quills are almost impossible to remove. Because of the barbed tip, the quill works continuously deeper into the wound and can even cause death. Quills progress at the rate of 1 millimeter (0.04 in.) or more per hour. Porcupine victims may die from infections caused by germs on the quills or from damage to a vital organ. Quills may even stick in an attacker's jaw, making the animal unable to open its mouth and causing starvation. Fishers (*Martes pennanti*), large members of the weasel family, attack porcupines by flipping them over onto their backs and attacking their unprotected ventral surface.

Hystricidae. Old World porcupines live in Africa, southeastern Asia, India, and southern Europe. Most species attain a length of about 3 ft (90 cm), including the tail. They are terrestrial or semifossorial (live both above and below ground) and make their homes in tunnels in the ground. Often six or seven porcupines share a communal burrow. They do not climb trees.

The crested porcupine (*Hystrix*) is the largest living porcupine. It may reach a length of over 3 ft (90 cm) and weight of 40–60 lb (18–27 kg). Its tail is just a few inches long. A mass of needle-pointed hollow quills extending from the nape of its neck and down its back is mixed with much longer and more slender flexible spines, or guards, measuring up to 21 in. (53 cm) in length. The guards protect the sharp points of the quills when they are not in use. This species lives in arid mountains, rocky outcrops, wooded hills, scrub-covered plains, and thick forest. These animals feed on bulbs, the bark of trees, and fallen fruit, and are destructive to farm crops, including root tubers, pumpkins, sweet potatoes, Irish potatoes, and maize. The annual litter usually consists of two young that are born after a gestation of 6–8 weeks. The nest consists of leaves, dry grass, and root fibers. The precocial young are well developed at birth with their eyes open and their bodies covered with soft flexible spines. The young remain in the den until the spines have grown and hardened; they are already quite dangerous at the end of 10 days. Average life expectancy in the wild is 12–15 years, although captive animals have lived for over 20 years.

Porcupines and their distribution

Name	Number of Species	Range
Hystricidae		
Long-tailed porcupine (*Trichys*)	1	Malaysia, Sumatra, Borneo
Brush-tailed porcupine (*Atherurus*)	2	China, Thailand, Malaysia, Gambia, Kenya, Zaire
Old World porcupine (*Hystrix*)	8	Philippines, Sumatra, Borneo, China, Malaysia, India, Africa
Erethizontidae		
Upper Amazonian porcupine (*Echinoprocta*)	1	Colombia
Prehensile-tailed porcupine (*Coendou*)	4	Central and South America
Hairy dwarf porcupine (*Sphiggurus*)	6	South America
North American porcupine (*Erethizon*)	1	North America

The brush-tailed porcupine (*Atherurus*) is only 20 in. (51 cm) in length, not including the 9-in. (23-cm) tail. It has short soft spines on the head, legs, and ventral surface, and very long grooved spines on the back. The long spiny tail terminates in a tuff of stiff bristles. The odd thing about this brush is that the bristles are composed of alternating thick and thin regions, appearing rather like a string of beads. The long-tailed porcupine (*Trichys*) has weak spines, or bristles, without quills and a long tail.

Erethizontidae. New World porcupines (Erethizontidae) are found in mixed coniferous forests, tropical forests, grasslands and deserts. Despite their generally chunky heavyset bodies, erethizontids can be arboreal unlike Old World porcupines. The North American porcupine (*Erethizon*) [see **illustration**], which weighs an average of 15 lb (7 kg) but may reach 40 lb (18 kg), is both terrestrial and semiarboreal, whereas South American species (*Coendou* and *Sphiggurus*) have prehensile tails and spend most of their time aboveground. *Erethizon* has a well-developed big toe on its hindfoot, whereas the hindfoot of the South American species has a broad pad opposable to the four toes. Heavy spines with a barbed tip, embedded singly and not in clusters as in Old World porcupines, occur over much of the dorsum and sides of the body. The ventral surface has coarse long hair and no spines. The claws are long and curved.

The staple winter food of North American porcupines is the cambium layer of many woody plants. During other seasons, they feed on buds, twigs, sedges, and young pine needles. *Erethizon* characteristically takes shelter in rock piles, beneath overhanging rocks, or in hollow logs, but does not dig burrows as do Old World porcupines. These porcupines are normally solitary animals except during the breeding season and in winter denning concentrations. Breeding occurs from September through November. After a gestation period of 205–215 days, exceptionally long for a rodent, a single young is born.

North American porcupine (*Erethizon dorsatum*). (*Photo from National Park Service*)

Prehensile-tailed porcupines, or coendous, are arboreal, with short limbs modified for climbing. They have a thick-set short body covered with long hairs mixed with spines, which increase in number toward the posterior and dorsal part of the body and the tail. These spines, or quills, can be neither "dropped" nor "shot" but can be made loose. Although not usually found great distances from trees, these porcupines can exist in areas of jumbled rock and shrubs.

Although porcupines are not endangered, they are being affected by loss of habitat, especially in South America. One species (*Sphiggurus pallidus*) may have become extinct in historic times. It was reported from the West Indies in the midnineteenth century, but no porcupines occur there now.

Donald W. Linzey

Bibliography. J. A. Chapman and G. A. Feldhamer (eds.), *Wild Mammals of North America: Biology, Management, and Conservation*, Johns Hopkins University Press, 2003; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; J. O. Whitaker, Jr., and W. J. Hamilton, Jr., *Mammals of the Eastern United States*, Comstock Publishing, 1998.

Porifera

The sponges, a phylum of the animal kingdom which includes about 5000 described species. The body plan of sponges is unique among animals. Currents of water are drawn through small pores, or ostia, in the sponge body and leave by way of larger openings called oscula. The beating of flagella on collar cells or choanocytes, localized in chambers on the interior of the sponge, maintains the water current.

Support for the sponge tissues is provided by calcareous or siliceous spicules, or by organic fibers, or by a combination of organic fibers and siliceous spicules. Some species have a compound skeleton of organic fibers, siliceous spicules, and a basal mass of aragonite or calcite. The skeletons of species with supporting networks of organic fibers have long been used for bathing and cleaning purposes. Because of their primitive organization, sponges are of interest to zoologists as an aid in understanding the origin of multicellular animals. *See* ANIMAL KINGDOM; PARAZOA.

The Porifera have a fossil record extending from the Precambrian to Recent times. More than 1000 genera of fossil sponges have been described from the Paleozoic, Mesozoic, and Cenozoic eras.

Taxonomy. The living Porifera are divided into four classes on the basis of their skeletal structures. A taxonomic scheme of the Porifera follows. See separate articles on each class.

Class Hexactinellida
  Subclass Amphidiscophora
    Order: Amphidiscosa
      Hemidiscosa
      Reticulosa
  Subclass Hexasterophora
    Order: Hexactinosa
      Lynchniscosa
      Lyssacinosa
Class Calcarea
  Subclass Calcinea
    Order: Clathrinida
      Leucettida
  Subclass Calcaronea
    Order: Leucosoleniida
      Sycettida
  Subclass Pharetronida
Class Demospongiae
  Subclass Tetractinomorpha
    Order: Homosclerophorida
      Choristida
      Spirophorida
      Hadromerida
      Axinellida
  Subclass Ceractinomorpha
    Order: Dendroceratida
      Dictyoceratida
      Halichondrida
      Haplosclerida
      Poecilosclerida
      (Permosphincta, extinct)
Class Sclerospongiae

General structure and cell types. The structure and functioning of a sponge is most easily studied in young fresh-water sponges as they develop from dormant bodies called gemmules. Such young sponges possess two discrete canal systems, an inhalant and an exhalant system, which are in communication with each other by way of numerous chambers lined with flagellated collar cells (**Fig. 1**). The outer or dermal epithelium is made up of flattened polygonal cells called pinacocytes, and is perforated at intervals by pores, or ostia, through which water enters the inhalant canal system, usually by way of large subdermal cavities. The exhalant canals are also lined by pinacocytes.

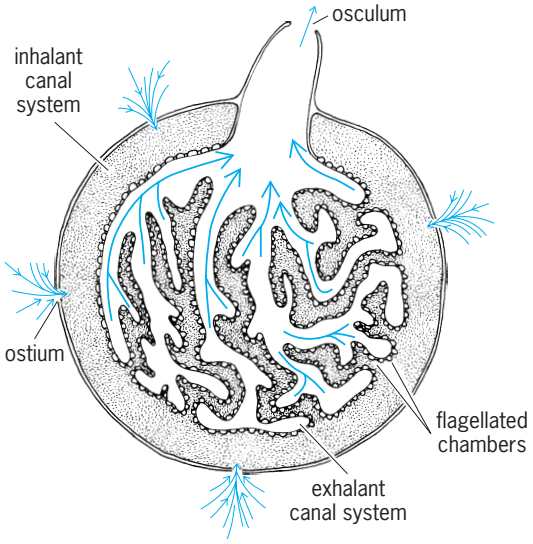

Fig. 1. Diagram of the canal system of a young fresh-water sponge which was cultured between cover slip and microscope slide.

All other cell types in the sponge (**Fig. 2**) are located in the inhalant canals. These cells may be divided into (1) those with large nucleolate nuclei and (2) those with smaller nuclei lacking nucleoli or with very small nucleoli. In the former category are the archeocytes, which are ovoid in shape and have

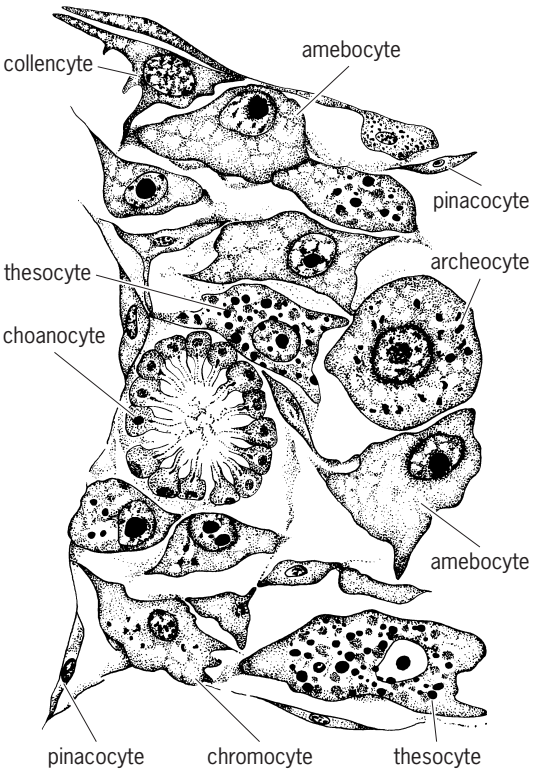

Fig. 2. Cell types found in a fresh-water sponge as seen in a cross section through the interior of the sponge.

blunt pseudopodia and hyaline cytoplasm, and many types of wandering amebocytes, which are named according to their functions. They are termed chromocytes if they bear either pigment granules or zoochlorellae (unicellular green algae); thesocytes or spherular cells if they are filled with food reserves; and trophocytes if they are acting as nutritive cells for developing eggs. Mucus-secreting gland cells, presumably derived from amebocytes, are of common occurrence in sponges. Scleroblasts, derived from archeocytes, secrete spicules, and spongioblasts secrete spongin fibers.

The anucleolate cells with small nuclei include the pinacocytes, collencytes, and choanocytes. In addition to their function as epithelial cells, pinacocytes may be differentiated into spindle-shaped contractile cells, called myocytes, which occur in circlets under the epithelium around ostia and oscula. Collencytes are called desmacytes or fiber cells and are found in abundance in the cortex of some species of sponges. Star-shaped collencytes with a vesicular or vacuolate cytoplasm are called cystencytes. The choanocytes are the most highly differentiated cells of the sponge. Each consists of a small spherical cell body surmounted by a collar. In freshwater sponges and certain marine Demospongiae and Calcarea, cells called lophocytes (Fig. 5*b*) occur beneath the dermal membrane. These cells bear a process terminating in a tuft of fine cytoplasmic filaments that secrete fibrils of collagen.

The mesenchymal cells and the skeleton of the Demospongiae and Calcarea are surrounded by a colloidal gel, called mesoglea, secreted by amebocytes and possibly choanocytes. Elastin fibers help provide support for the sponge and occur in the mesoglea of most species.

Water-current-system physiology. The flagellated chambers are hemispherical with a diameter of 40–60 micrometers in fresh-water sponges. They are composed of choanocytes firmly held together. The uncoordinated beating of the flagella of the choanocytes creates the flow of water through the sponge. Water enters the flagellated chambers from the inhalant canals by way of openings between choanocytes. Two or three such small openings (prosopyles) lead into each chamber. Water leaves each chamber and enters the exhalant system by way of a single larger pore (apopyle) through the epithelial lining of the exhalant canal. The lumen of an apopyle is 10 times greater than that of all the prosopyles leading into the chamber; thus water enters the chambers at a velocity 10 times greater than that at which it leaves. The collars of the choanocytes are so oriented that each directs its current toward the apopyle.

A fall in pressure takes place between apopyles and prosopyles, and therefore the prosopyles exert a suction effect on the water in the inhalant canal. The effect of this lowered pressure is transmitted to the subdermal cavity, and water is therefore drawn into the ostia forcibly. Food particles that may be at considerable distances from the pores are sucked in along with the water.

The incurrent canals are filled with a network of free cells which tend to decrease the velocity of the water, but the retarding effect of these frictional forces is compensated by the resultant decrease in diameter of the streams of water passing through the inhalant canals. The velocity of current flow through the inhalant canals is twice as great as that through the exhalant canals, which have smooth walls but are also greater in diameter. An excess pressure exists in the exhalant canals as a result of the pumping action of the flagellated chambers and the changing size of the oscular opening. Hence food-free water laden with excretory products is carried far away from the sponge as it leaves the osculum, and is unlikely to reenter through the ostia.

Feeding and digestion. Feeding activities are best known in fresh-water sponges in which the dermal ostia, the inhalant canals, the flagellated chambers, and the choanocyte collars constitute a set of sieves of decreasing mesh size. The dermal ostia are 50 μm in diameter, the prosopyles are 5 μm in diameter, and the spaces between the cytoplasmic tentacles of the collars vary from 0.10 to 0.15 μm (**Fig. 3**). Only the smallest particles can enter the flagellated chambers and come in contact with the collars of the choanocytes. The latter trap the small particles, which are passed to the cell body and ingested. Larger particles that have been able to pass through the dermal pores are ingested by the archeocytes and collencytes that form a reticulum in the inhalant canals or by the cell bases of the choanocytes (**Fig. 4**). Particles too large to enter the ostia can be ingested by the pinacocytes of the dermal epithelium. Studies of tropical, coral-reef-inhabiting demosponges indicate that they are highly efficient in filtering out bacteria and unarmored flagellates and moderately efficient in retaining diatoms, dinoflagellates, and

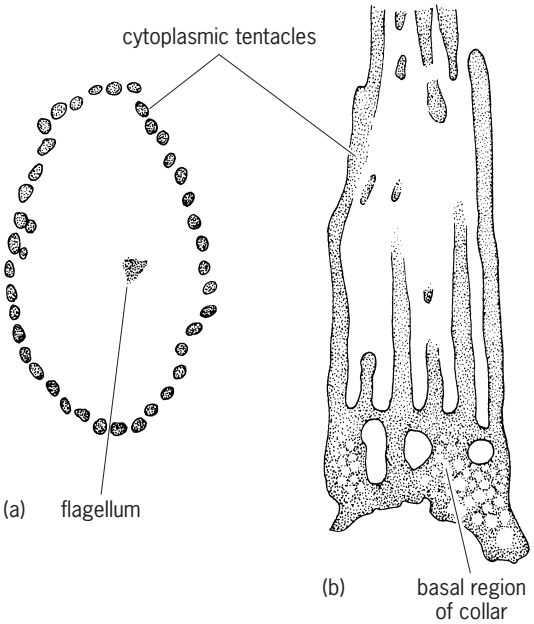

Fig. 3. Sections through the collar of a choanocyte of a fresh-water sponge. (*a*) Transverse. (*b*) Tangential.

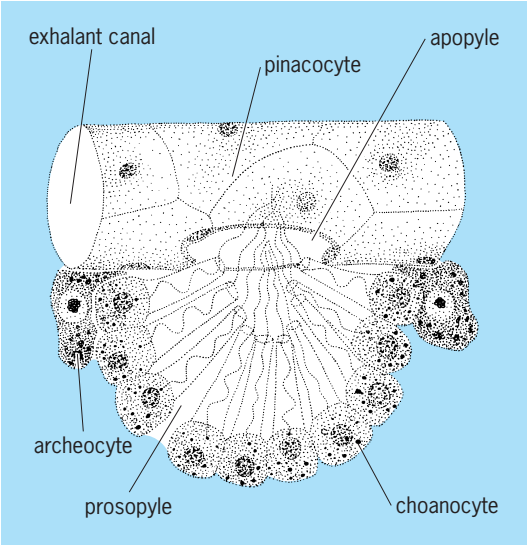

Fig. 4. Diagram of a flagellated chamber of a fresh-water sponge. The chamber is shown in optical section; exhalant canal, in three dimensions.

coccolithophores. Microscopically resolvable detritus and unresolvable particulate organic matter are additional sources of nutrients for these sponges, with the latter constituting by far the most important food source. Sponges are uniquely adapted to exploit food sources of minute particle size, and as a result can inhabit nutrient-poor waters of tropical reefs, polar shelves, and submarine caverns.

Direct observations of living fresh-water sponges indicate that the archeocytes are the chief cells responsible for digestive processes. After food has entered the sponge, choanocytes take up and then rapidly lose the particles which they have ingested, and it has been assumed that the particles are transferred to archeocytes and amebocytes for digestion. Among simple calcareous sponges and homosclerophorid demosponges with large choanocytes that make up a great proportion of the total mass of sponge cells, digestion appears to be confined to the choanocytes. Studies of the activity of digestive enzymes in the marine siliceous sponge *Halichondria* indicate higher proteolytic, lipolytic, and carbohydrate-decomposing activities in choanocytes than in archeocytes. Following the digestion of the particles in the archeocytes, they migrate to the walls of the exhalant canals or to the dermal epithelium, where indigestible material is voided from blisterlike vacuoles.

Reactions of sponges. Studies of a shallow-water marine siliceous sponge, *Hymeniacidon heliophila*, have shown that the oscula close upon exposure to air and when the sponge is in quiet seawater, although they may remain open when the sponge is in silt-laden water. Touching or stroking either oscula or ostia does not induce closure. The reactions of this sponge are thought to be solely the result of the contractions of myocytes reacting directly to external stimuli. Movements are slow and transmission of stimuli is limited. Reactions of oscula in the fresh-water sponge *Ephidatia fluviatilis* differ somewhat from those of the previous species. Oscula remain open in running or still water. Oscular tubes elongate in still water and flatten out in strong currents. If a needle is rubbed around the edge of the oscular opening, the orifice contracts immediately. Weak electrical stimuli applied to an osculum tip cause it to close, and a wave of contraction runs down the entire oscular tube. Transmission of the electric stimuli to neighboring regions of the sponge was not observed.

Although neither of the siliceous sponges described above reacted to light, it has been reported that in the calcareous sponge *Leucosolenia* bright light causes oscular constriction and a change in tube shape.

The marine demosponge *Tethya aurantia* reacts to needle pricks by general contraction of the body and closure of the oscules. The reactions begin within 30 s and are not completed until 3–7 min have elapsed. The slow behavioral responses of sponges fail to support the existence of nerve cells, although these have been reported by some zoologists from histological preparations (**Fig. 5***a*).

Studies of field populations of sponges demonstrate that in some species the rhythmic cessation of water pumping occurs. All oscules of such sponges close simultaneously when the water currents cease; a coordination mechanism for this is indicated.

Skeleton. Characteristic of sponges is the presence of a skeleton of sclerites or of organic fibers or both. Only a few genera, such as *Halisarca* and *Oscarella*, lack skeletal elements. The skeleton is of primary importance in the classification of sponges; indeed, the four classes of the Recent phylum are separated on the basis of the skeletal structure.

The shapes of the sclerites or spicules vary greatly (**Fig. 6**), and an elaborate terminology has been

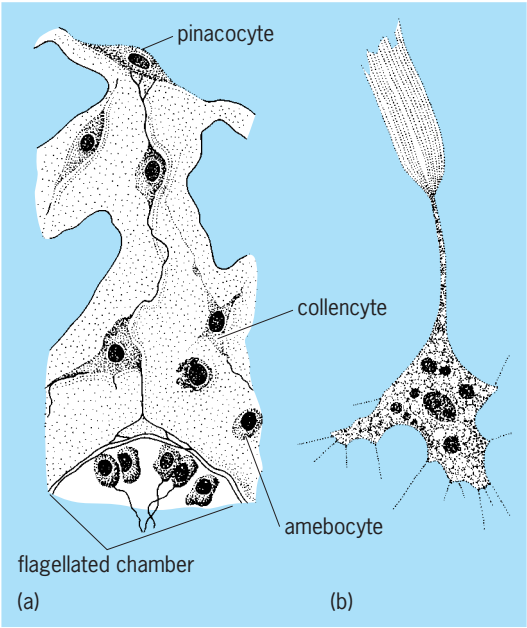

Fig. 5. Sponge cells. (*a*) Two-cell arc of supposed nerve cells joining pinacocyte and choanocyte chamber in a calcareous sponge. (*b*) Lophocyte of a fresh-water sponge.

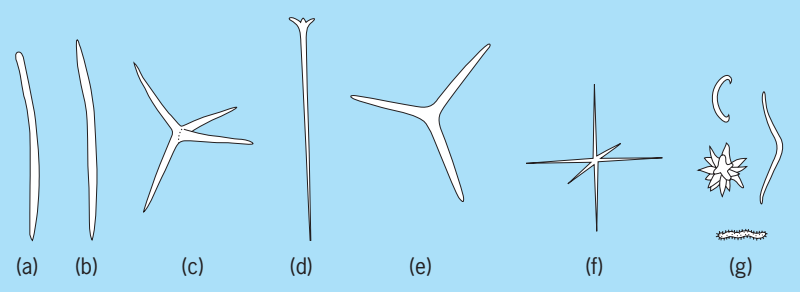

Fig. 6. Spicule types. Monaxons: (*a*) monactinal and (*b*) diactinal. Tetraxons: (*c*) equirayed and (*d*) triaene. (*e*) Triradiate. (*f*) Triaxon or hexactine. (*g*) Microscleres.

developed for them by taxonomists. An initial subdivision into megascleres and microscleres is made on the basis of size. A more detailed nomenclature is based on the number of axes or rays present; an appropriate numerical prefix is added to the endings -axon (referring to the number of axes) or -actine (referring to the numbers of rays or points). Spicules formed by growth along a single axis are known as monaxons. Such spicules are monactinal (pointed at one end only) if growth occurs in one direction, or diactinal (pointed at each end) if growth occurs in both directions. Tetraxons are spicules with four axes or rays, each pointing in a different direction. All four rays may be of equal length, but often one ray is longer than the other three, and those spicules are called triaenes. Triradiate (three-rayed) spicules are commonly found in calcareous sponges. Hexactinellid sponges are characterized by the presence of spicules with three axes (triaxons) meeting at right angles. Such spicules have six rays and are also called hexactines.

The reticulate fibrous skeletons of bath sponges are composed of collagenous protein called spongin, which also occurs as fine fibrils in the mesoglea. Spongin is also found in many species of Demospongiae with siliceous spicules, where it serves as an interspicular cement or forms fibers in which the spicules are embedded.

Reproduction. Both sexual and asexual reproduction occurs in sponges. The formation of asexual bodies, gemmules, is characteristic of most fresh-water species.

Sexual reproduction. The germ cells may arise from the flagellated cells of the larva, from choanocytes, from archeocytes, or from pinacocytes. The sperm enters the egg not directly, but through the intermediary of a choanocyte which loses its collar and flagellum and migrates from the flagellated chamber to a position adjacent to the egg. This process has been observed in several species of Calcarea and Demospongiae and is probably of general occurrence in these groups. The sperm, which loses its tail and enlarges, is eventually transferred to the egg, presumably following the fusion of the carrier cell with the egg (**Fig. 7**). *See* FERTILIZATION.

Subsequent development of the embryo and larva varies somewhat in the several classes. The free-swimming flagellated larvae are either hollow amphiblastulae or solid parenchymellae. *See* CALCAREA; DEMOSPONGIAE; HEXACTINELLIDA.

Asexual reproduction. Many fresh-water and marine sponges disintegrate with the onset of unfavorable environmental conditions, leaving behind reduction bodies which are compact masses of cells, chiefly amebocytes, covered by an epidermal layer. In some shallow-water marine sponges, the entire sponge remains alive during the winter months but changes internally into a mass of amebocytes. Both reduction bodies and the partially dedifferentiated adult sponges can develop into functional individuals upon return of favorable conditions.

Most fresh-water and many marine sponges produce asexual reproductive bodies, called gemmules, as part of their life cycle (**Fig. 8**). Gemmules are formed of masses of archeocytes laden with food reserves in the form of lipids and proteins. In fresh-water sponges the inner mass of cells is surrounded by a layer of columnar cells which secretes a double layer of spongin around the gemmule. Characteristic spicules are deposited between those layers. The gemmules of many species of marine sponges are also enclosed in a spongin coat, which may or may not be provided with spicules. The gemmules of fresh-water sponges and a few marine species carry those species over periods of unfavorable environmental conditions, such as drought or low temperatures, during which times the adults degenerate. Germination of the gemmules occurs upon return of suitable conditions. In many marine sponges, gemmules are formed at all seasons, and the adult sponges show no seasonal degeneration.

In a few species of Demospongiae and Hexactinellida, asexually produced gemmules, lacking spongin coats, have been described as developing into flagellated larvae identical in structure to those formed as a result of sexual processes.

Budding is another common type of asexual reproduction in sponges. In *Tethya*, for example, groups of archeocytes migrate to the tips of spicule bundles projecting from the surface (**Fig. 9**). These

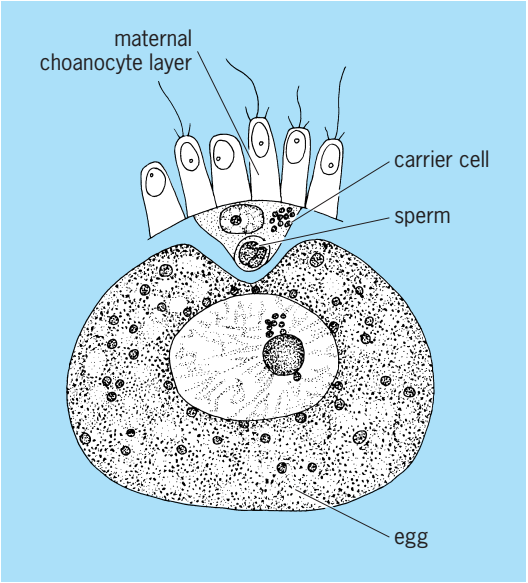

Fig. 7. Carrier cell transmitting sperm to egg in a calcareous sponge.

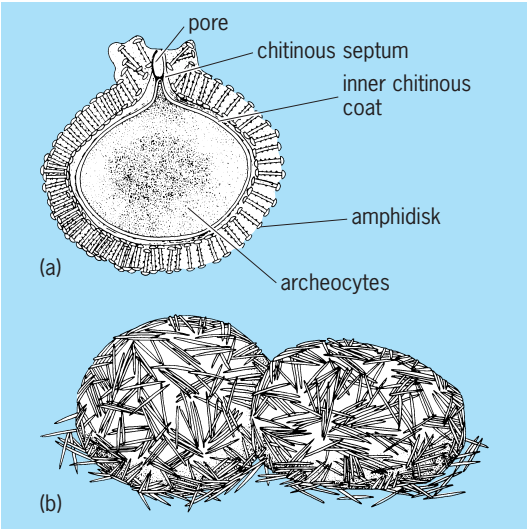

Fig. 8. Sponge gemmules (asexual) reproductive bodies. (*a*) Fresh-water sponge gemmules in optical section. (*b*) Marine siliceous sponge gemmules, surface view.

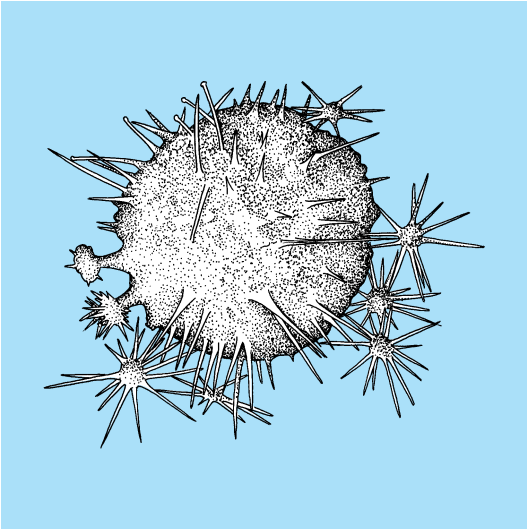

Fig. 9. Colony of *Tethya*, a marine siliceous sponge, with buds located at the tips of projecting spicule bundles.

buds fall off and form new individuals. *See* REPRODUCTION (ANIMAL).

Regeneration. Fragments cut from sponges or pieces which break off accidentally in nature can reattach to the substrate and reconstitute functional individuals. Suspensions of cells prepared by squeezing fragments of sponges through fine silk bolting cloth into seawater are also capable of reorganizing into functional individuals. All types of cells normally present in the adult sponge, except scleroblasts, have been identified in the cell suspensions. Reorganization of the cells to form functional sponges thus involves chiefly a migration of the cells to take up their appropriate positions in the reconstituted organism. The pinacocytes form a peripheral epithelium around the aggregated cells, and the various mesenchymal cells form the central mass. The collencytes are active in forming the inhalant and exhalant canals; the latter become lined with pinacocytes. Choanocytes group together to reconstitute flagellated chambers between the inhalant and exhalant canals. Scleroblasts form anew from archeocytes. Some workers have reported that collar cells dedifferentiate or are phagocytized by pinacocytes or amebocytes and are formed anew during the reorganization of the sponge.

The initial aggregation of the dissociated cells to form masses results from random movement of archeocytes, amebocytes, and other cell types, including choanocytes that have been observed to put forth pseudopodia and also move short distances by means of their flagella. As cells come into contact with one another, they adhere, possibly as a result of antigen-antibody-like forces. The presence of homologous antibodies inhibits reaggregation. Neighboring cell aggregates often adhere as well, if they come into contact with one another.

Mixed cell suspensions of two species of sponges of different colors may result in an initial intermixture of cells of the two species, but these later sort out so that the cell aggregates are eventually composed exclusively or predominantly of cells of one species or the other.

Commercial sponge fisheries. Although plastic sponges now offer competition, there is still a demand for natural sponges for use by various artisans, for surgical purposes, and for cleaning automobiles. Commercially valuable sponges are harvested in the eastern part of the Mediterranean Sea, off the western coast of Florida and off the Florida Keys, in the West Indies, and to a limited extent off the Philippines. They are gathered by hooking or harpooning from a boat in shallow waters; by nude diving, a method used especially for exploiting cave populations in the Aegean Sea; by machine diving with the aid of diving suits attached by a life line to an air pump; and by dredging, a wasteful method prohibited on most sponge grounds.

Artificial propagation of sponges has proven to be feasible when serious efforts have been made to do so. In the Bahamas, sponge cuttings planted on natural bottoms or on concrete disks in carefully selected areas yielded a harvest of 140,000 sponges during the years 1936–1939. Regrettably a disease, presumably caused by a fungus, spread through the West Indies in 1938–1939, put an end to the cultivation experiments, and killed a large percentage of the natural population of sponges.

Before World War II the Japanese had achieved considerable success in cultivating sponges in the Marshall and Caroline islands. Their methods included planting directly on concrete disks or stringing cuttings on aluminum wire anchored to rafts of various sorts or provided with floating buoys.

Organisms associated with sponges. Sponges are preyed upon regularly by chitons, snails, and nudibranchs, and occasionally by fish and turtles and probably other types of animals. The larvae of neuropteran insects of the family Sisyridae (spongilla flies) live in and feed upon colonies of fresh-water sponges. Sponges also play host to myriads of animals that live in the natural cavities of the sponge or form burrows of their own. Sea anemones, zoanthids (**Fig. 10**), polychaetes, octopuses, copepods,

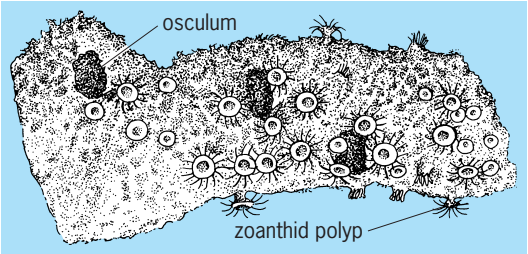

Fig. 10. Zoanthids living in a sponge.

barnacles, amphipods, shrimps, brittle stars, and fish live in sponges, but the natures of their relationships with their hosts have been little explored. The siliceous sponge *Suberites domunculus* always lives on shells occupied by hermit crabs and provides the crabs with a continually growing house. Certain crabs plant or hold pieces of sponge on their backs and thus blend into their environment. Other encrusting animals which compete for space with sponges, such as hydroids, bryozoans, and ascidians, often overgrow or are overgrown by sponges in the struggle.

Zoochlorellae commonly live in the amebocytes of fresh-water sponges and many shallow-water marine species. Because the algal cells fail to develop starch reserves when living in sponge cells, it seems likely that the sugars synthesized by the algae are utilized by the sponge cells. Dying algal cells are consumed by the amebocytes. Apparently both the algae and sponges can survive independently of each other. Young sponges are reinfected with algae from the water in which they live. Certain species of marine Demospongiae harbor algal symbionts in the form of zooxanthellae (dinoflagellates) or blue-green algae.

Certain marine sponges are regularly filled with filamentous green or coralline algae. The blue color of some sponges has been attributed to bacteria living in the cells. *See* ARCHAEOCYATHA. Willard D. Hartman

Fossils. Fossil sponges are relatively rare when compared with other invertebrate fossils, but they do occur in profusion in some areas. Upon death the spicular skeletons of sponges are commonly disarticulated or disjointed, except in those forms where the spicules are articulated or fused to one another to construct a resistant skeleton. These latter forms are most frequently preserved in the fossil record. Dissociated sponge spicules are common in marine sedimentary rocks, but once isolated from the skeletal fabric they are difficult to identify, and assignment to species or genera is usually impossible and certainly not recommended unless the spicule has unique characteristics. One sponge may have several kinds of spicules, and the same kind of spicule may be present in several genera.

Relatively solid skeletons occurred in lithistid and sphinctozoid Demospongiae; dictyid, brachiospongid, hexactinosid, and lychniscosid Hexactinellida; Sclerospongia; and pharetronid Calcarea. As a result, these sponges have left a moderately common fossil record.

Geologic record. Precambrian sponges are known from a few localities in the world, such as in the Ediacaran beds of Australia or roughly equivalent beds in China. Lower Cambrian sponges are also known from a few localities, such as Hunan Province in China. It was not until the Middle Cambrian, however, that sponges left even a modest record. The Burgess Shale of the Canadian Rocky Mountains has produced the most extensive Cambrian sponge fauna known to date, although scattered sponges of that general age are known from many other areas of the world. Hexactinellida, Demospongiae, Sclerospongea, and Heteractinida are known from Cambrian rocks and, except for the Heteractinida, range up to the Recent. Ordovician and Silurian sponge faunas are dominated by a great variety of demosponges and are moderately common throughout the Appalachian Mountains, in western Tennessee, Utah, Nevada, Texas, and Oklahoma, in Arctic Canada in North America, and in limited localities in northern Europe, Australia, and China. Devonian fossil sponges are dominated by hexactinellid forms in North America, but lithistid sponges helped produce reefs during that period in western Australia. Sclerosponges, represented by stromatoporoids, appear most common in Ordovician through Devonian rocks. Sponges of Mississippian and Pennsylvanian ages occur as scattered rare specimens in many areas, but it was not until the Permian that sponges became common fossils again. In Permian rocks, permosphinctid (sphinctozoan) demosponges and aspiculate inozoan Calcarea are major elements in fossil reefs and are particularly common in West Texas and southern New Mexico, Tunisia Sicily, Timor, and southern China.

Triassic permosphinctid sponges are common in reefs and associated beds in the Alpine region of southern Europe, but are rare in the Americas and most other parts of the world. Jurassic sponges, mainly lithistid and hexactinellid genera, are common in Europe, in reefs of Morocco, and in eastern Mediterranean countries, but are virtually unknown in North America. The Cretaceous was a time when hexactinosid and lychniscosid hexactinellid sponges were common in western and northern Europe, but fossil sponges are relatively rare in Cretaceous rocks in North America. Cenozoic fossil sponges are closely related to Recent sponges, but have left only a moderate record, particularly in North America.

Fossil types. Fossil sponges are divided into five classes: Calcarea (calcisponges), Hexactinellida, Demospongiae, Sclerospongia, and the extinct Heteractinida.

Calcarea. The Calcarea are sponges with spicules of calcite or aragonite. Such spicules in the pharetrone sponges are monaxons, triradiates, or quadriradiates and are commonly fused into a rather rigid network of fibers. Pharetrone sponges first appear in the Permian of West Texas and appear slightly later in Permian rocks in Sicily and Tunisia. They range to the Recent, but only locally are they common in modern seas. *See* CALCAREA.

Hexactinellida. These sponges have skeletons composed of opaline silica and have spicules composed of four or six, rarely more, rays which are at right angles to each other. They appeared in the late Precambrian, and by the Early Paleozoic two large subgroups

are recognizable. One has thin walls in which the spicules are regularly arranged, and the other has thick walls in which spicules are irregularly arranged. In neither early group are the spicules fused together, but the skeletons may have been held together by organic materials which were not preserved. Late Paleozoic forms that are descended from the regularly spiculed protosponges developed thick regular walls and appear transitional to the sponges with fused skeletons that became common Mesozoic and Cenozoic representatives of the class. Abundant and well-preserved hexactinellid sponges have been recovered from Devonian rocks of southern New York and northern Pennsylvania and from Mississippian rocks of Indiana. Oldest Hexactinosa with fused, rectangularly based, three-dimensional skeletons have been questionably reported from the Permian, but by the Triassic they were certainly present. These sponges reached their peak of diversity in the Cretaceous, along with the Lychniscosa. The latter are similar-appearing sponges and also have rectangular-based skeletons. In addition, the lychniscosid sponges have short diagonal braces that form the lanternlike lychnisc structure at the spicule nodes. These Mesozoic hexactinosid sponges reached their maximum diversity and abundance in Jurassic and Cretaceous rocks in Europe. *See* HEXACTINELLIDA.

Demospongiae. Demosponges have skeletons composed of spongin fibers, siliceous spicules, or mixed spicules and spongin. Only those forms with spicules have left a fossil record. Spicules range greatly in shape from monaxons, to tetraxons where rays do not meet at right angles, to irregular rootlike shapes. The vast majority of known fossil demosponges belong to the order Lithistida, a group that is well preserved because their articulated and fused spicules produced a solid cross-braced skeleton. Most lithistid spicules are irregular, with knobs and spines, or with complexly branching articulating ray tips that united to produce skeletons that could be rather easily preserved. Demosponges first appeared in the geologic record in the Middle Cambrian, but rapidly diversified and increased in numbers in the Early Ordovician to become one of the major reef builders of that period. They continued to diversify and remained common in the Silurian, but these families decreased in relative abundance and diversity in the later Paleozoic.

Permosphinctid (sphinctozoan) demosponges have calcareous beadlike or honeycomblike skeletons. They were considered as calcareous sponges until recently. The sphinctozoans are locally common and are composed of hollow segments arranged like a string of beads or as bubblelike clusters with solid walls in which the spicule structure is not preserved. Earliest known sphinctozoan sponges are from the Cambrian of southern Australia, and by the Ordovician these sponges are known from Australia, California, and Alaska. They remained as relatively minor elements of sponge faunas in the Devonian and through the Carboniferous, but rapidly reached their maximum diversity, abundance, and geographic distribution in tropical realms of the Permian and Triassic. Their diversity was greatly reduced in the Jurassic, and all but one species became extinct by the end of the Cretaceous.

Spiculate demosponges have another peak of abundance and diversity in Jurassic and Cretaceous rocks in Europe, and they are the major sponge elements in modern seas. *See* DEMOSPONGIAE.

Heteractinida. The Heteractinida have skeletons composed of large calcareous spicules. Early representatives of the class have spicules with six rays in one plane; later sponges of the class have octactine-based spicules in which six equally spaced rays occur in one plane that is approximately at right angles to the other two opposed rays. Two broad patterns of spicule arrangement are recognizable in the class: one group is characterized by thin walls composed of regularly arranged spicules, and the other is characterized by thick-walled sponges with irregularly arranged, almost felted spicules. Oldest known Heteractinida are from the Lower Cambrian, and sponges of the class are moderately common in Middle Cambrian rocks, but they are particularly common in Silurian and Carboniferous rocks of North America and Europe. The Heteractinida is the only class of sponges to become extinct, for they died out near the end of the Permian.

Sclerospongiae. The class Sclerospongiae was recently erected to include six genera of living sponges. These modern forms have laminated skeletons of calcium carbonate and are structurally similar to stromatoporoids, which are common fossils in Paleozoic and Mesozoic rocks. Stromatoporoids are now considered by some workers to belong in the Porifera as early sclerosponges. Their laminate skeletons are pierced by astrorhizal canals, and presumably the living tissue was confined to their upper surfaces, as in modern sclerosponges. Until recently, stromatoporoids were included as fossil cnidarians allied to hydrozoans. Sclerosponges are now considered to have first appeared in the geologic record in the Cambrian, and flourished during the Middle Paleozoic when they were important reef builders in the Silurian and Devonian. After that, particularly from near the end-Paleozoic, they became relatively minor elements, even into modern seas. *See* SCLEROSPONGE.

J. Keith Rigby

Bibliography. P. R. Bergquist, *Sponges*, University of California Press, Berkeley, 1978; D. Hill, Archaeocyatha, *Treatise on Invertebrate Paleontology*, pt. E, vol. 1, 2d ed., rev. 1972; J. N. A. Hooper and R. W. M. Van Soest (eds.), *Systema Porifera: A Guide to the Classification of Sponges*, vols. 1 and 2, Kluwer Academic/Plenum, New York, 2002; R. L. Kaesler (ed.), *Treatise on Invertebrate Paleontology*, pt. E (rev.), vols. 2 and 3, 2003, 2004.

Porocephalida

One of two orders in the class Pentastomida of the phylum Arthropoda. In this order the larvae have four legs. The hooks on the adult articulate on a chitinous base or fulcrum and are arranged in a flattened trapezoidal pattern, or in a curved line or straight line. There are no podial or parapodial

lobes. The two suborders are Porocephaloidea and Linguatuloidea. The Porocephaloidea comprise five families—Porocephalidae, Sebekidae, Armilliferidae, Sambonidae, and Subtriquetridae—and include most of the pentastomid species. These animals are cylindrical, with rounded ends. The hooks are sessile; the mouth is anterior between the hooks. The cephalothorax contains large cephalic and hook glands. The female genital pore is at or near the posterior end of the body; the uterus is long and tubular and disposed in loops. Eggs in the terminal part of the uterus contain fully formed larvae. There is a single testis; the male genital pore is ventral, at the anterior border of the abdomen. A cirrus sac is present; the cirrus is long.

Adults occur in the respiratory passages of snakes, lizards, turtles, and crocodiles; the larvae are encysted in fishes, snakes, lizards, crocodiles, and mammals. Linguatuloidea contains a single family, with a single genus, *Linguatula* Fröhlich, 1789. Adults live in the nasal cavities of carnivorous mammals; the larvae are encysted in herbivorous mammals. The body is elongate, flattened ventrally, annulate, and attenuated posteriorly. The females are much larger than the males and may exceed 4 in. (100 mm) in length. In the males the testis is double. The hooks of the adults are simple; those of the larvae are binate. *See* PENTASTOMIDA.

Horace W. Stunkard

Porphyrin

One of a class of cyclic compounds in which the parent macrocycle consists of four pyrrole-type units linked together by single carbon bridges. Several porphyrins with selected peripheral substitution and metal coordination carry out vital biochemical processes in living organisms. Chlorins (**1**), bacteriochlorins (**2**), and corrins (**3**) are related tetrapyrrolic

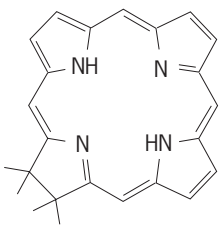

(**1**)

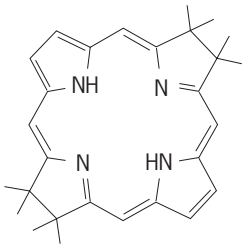

(**2**)

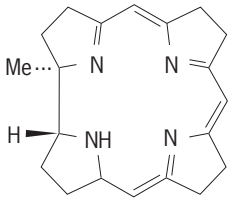

Me = CH_3

(**3**)

macrocycles that are also observed in biologically important compounds.

Nomenclature. The complexity of porphyrin nomenclature parallels the complex structures of the naturally occurring derivatives. Hans Fischer (who won a Nobel prize for his studies on porphyrins) used a simple numbering system for the porphyrin nucleus (**4**) and a set of common names to identify the different porphyrins and their isomers. A systematic naming based on the 1–24 numbering system for the porphyrin nucleus (**5**)

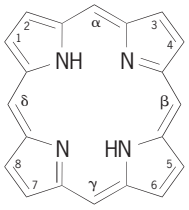

(**4**)

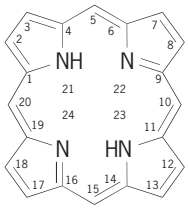

(**5**)

was later developed by the International Union of Pure and Applied Chemistry (IUPAC) and the International Union of Biochemistry (IUB), and this system has gained general acceptance. The need for common names (see **table**) is clear after examination of the systematic names; for example, protoporphyrin IX has the systematic name 2,7,12,18-tetramethyl-3,8-divinyl-13,17-dipropanoic acid.

Properties. The unique way in which some molecules can delocalize their own electrons make them especially stable. Such molecules (for example, benzene) are said to be aromatic. The aromatic character (hence stability) of porphyrins has been confirmed by measurements of their heats of combustion. In addition, x-ray crystallographic studies have established planarity of the porphyrin macrocycle; this is a basic requirement for aromatic character. *See* DELOCALIZATION; X-RAY CRYSTALLOGRAPHY.

Metal-free porphyrins are usually red-orange in solution. They exhibit characteristic optical spectra (**Fig. 1**), with a very strong absorption band around 400 nanometers (known as the Soret band) and usually four additional bands in the 500–700-nm range. In acidic solution, the inner ═N— groups are

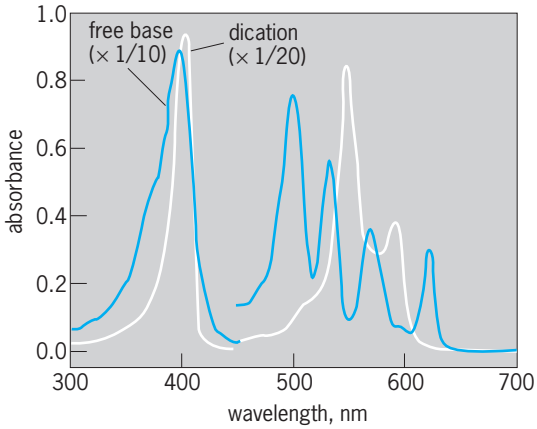

Fig. 1. Optical spectra of a porphyrin and its protonated dication.

Common names of some naturally occurring porphyrins

Name*	Substituents and locants†							
	2	3	7	8	12	13	15	17
Coproporphyrin I	Me	Cet	Me	Cet	Me	Cet	H	Me
Coproporphyrin II	Me	Cet	Me	Me	Me	Cet	H	Cet
Coproporphyrin III	Me	Cet	Me	Me	Me	Cet	H	Cet
Coproporphyrin IV	Me	Cet	Me	Cet	Cet	Me	H	Cet
Cytoporphyrin	Me	$-CH(OH)CH_2R'$	Me	Vn	Me	Cet	H	Cet
Deuteroporphyrin	Me	H	Me	H	Me	Cet	H	Cet
Etioporphyrin I	Me	Et	Me	Et	Me	Et	H	Me
Etioporphyrin II	Me	Et	Et	Me	Me	Et	H	Et
Etioporphyrin III	Me	Et	Me	Et	Me	Et	H	Et
Etioporphyrin IV	Me	Et	Me	Et	Et	Me	H	Et
Hematoporphyrin	Me	$-CH(OH)CH_3$	Me	$-CH(OH)CH_3$	Me	Cet	H	Cet
Mesoporphyrin	Me	Et	Me	Et	Me	Cet	H	Cet
Phylloporphyrin	Me	Et	Me	Et	Me	H	Me	Cet
Protoporphyrin	Me	Vn	Me	Vn	Me	Cet	H	Cet
Pyrroporphyrin	Me	Et	Me	Et	Me	H	H	Cet
Rhodoporphyrin	Me	Et	Me	Et	Me	$-CO_2H$	H	Cet
Uroporphyrin I	Cm	Cet	Cm	Cet	Cm	Cet	H	Cm
Uroporphyrin II	Cm	Cet	Cet	Cm	Cm	Cet	H	Cet
Uroporphyrin III	Cm	Cet	Cm	Cet	Cm	Cet	H	Cet
Uroporphyrin IV	Cm	Cet	Cm	Cet	Cet	Cm	H	Cet
Phytoporphyrin	Me	Et	Me	Et	Me	$-C(O)$	$-CH_2-$	Cet

*Roman numerals refer to an isomer with a particular order of peripheral substituents.
†Cet = carboxyethyl = $-CH_2CH_2CO_2H$; Cm = carboxymethyl = $-CH_2CO_2H$; Vn = vinyl = $-CH{=}CH_2$; Me = methyl = $-CH_3$; Et = ethyl = $-CH_2CH_3$; C = carbon; H = hydrogen; O = oxygen; R' = organic group

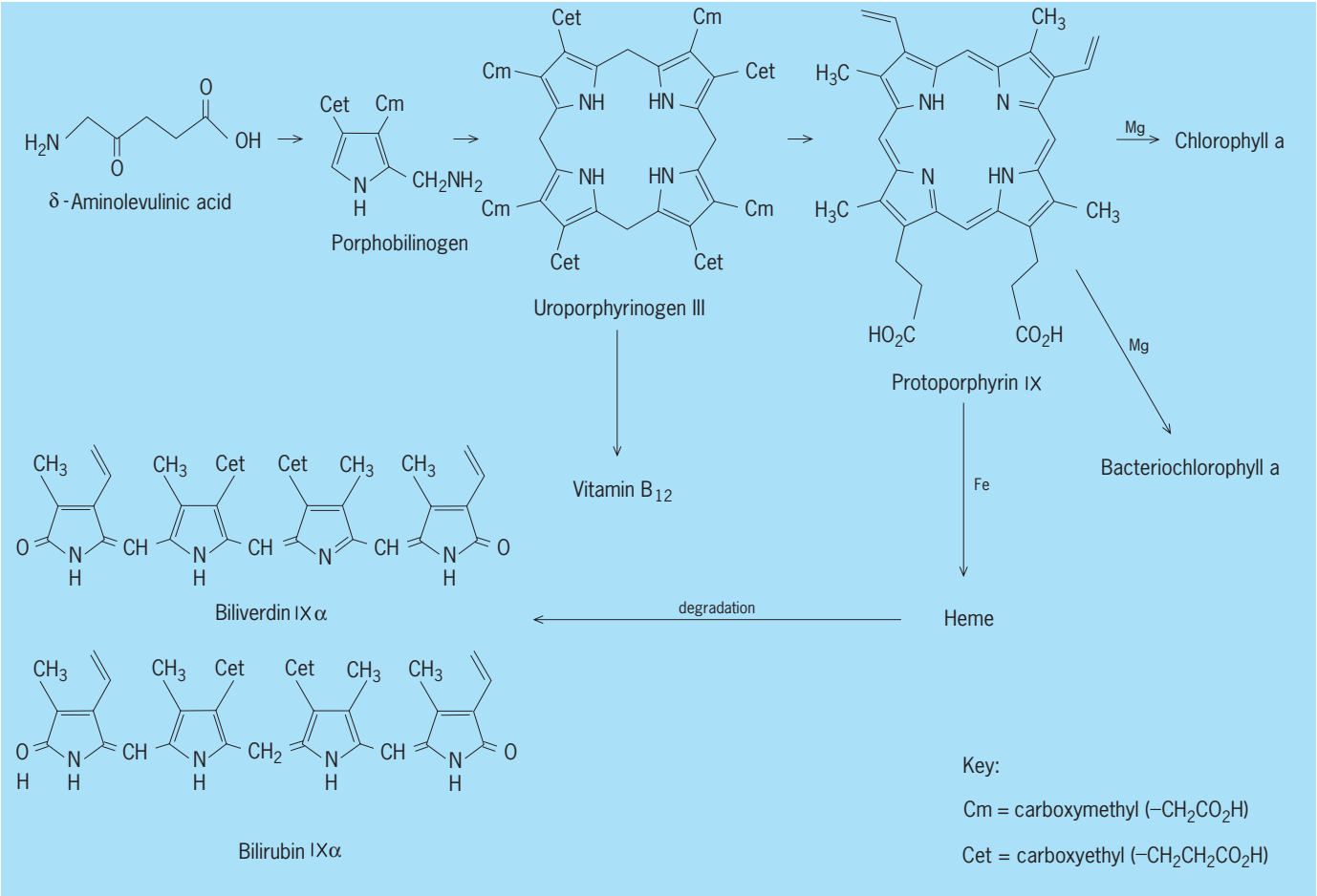

Fig. 2. Biosynthesis and metabolism of porphyrin-related macrocycles.

protonated, giving rise to the dication. This increases the symmetry of the molecule and simplifies the optical spectrum.

Porphyrins, metalloporphyrins, and their metabolic products, the bile pigments biliverdin and bilirubin (**Fig. 2**), are used as pigments in nature. Indeed, many of the rich colorations in birds' eggs are derived from porphyrin. Phthalocyanins (**6**), which

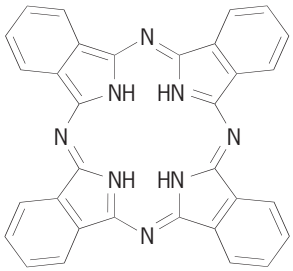

(**6**)

are closely related to porphyrins, are synthetic pigments. Like the porphyrins, they are both colorful and stable and are used industrially to color bluejeans and automobiles. *See* DYE.

Metalloporphyrins. Most metals and metalloids have been inserted into the central hole of the porphyrin macrocycle. The resulting metalloporphyrins are usually very stable and can bind a variety of small molecules (known as ligands) to the central metal atom. Metalloporphyrins also exhibit a strong Soret band, but have a simplified spectrum in the visible region of their absorption spectrum due to increased symmetry. In addition to the iron complexes (heme) and magnesium complexes (chlorophyll and bacteriochlorophyll), other metalloporphyrins occur in nature. In particular, nickel and vanadium complexes occur in oil, oil shales, and other geological deposits During the refining process involved in the conversion of oil to gasoline, these very stable metalloporphyrins pose many problems. *See* COORDINATION COMPLEXES.

Heme (**Fig. 3**), the iron complex of protoporphyrin IX, is the prosthetic group of a number of major proteins and enzymes that carry out diverse biological functions. These include binding, transport, and storage of oxygen (hemoglobin and myoglobin), electron-transfer processes (cytochromes), activation and transfer of oxygen to substrates (cytochromes P-450), and managing and using hydrogen peroxide (peroxidases and catalases). *See* CYTOCHROME; ELECTRON-TRANSFER REACTION; ENZYME; HEMOGLOBIN; IRON METABOLISM; PROTEIN.

Chlorophylls and bacteriochlorophylls are magnesium complexes of porphyrin derivatives known as chlorins and bacteriochlorins, respectively. They are the pigments responsible for photosynthesis. Several chlorophylls have been identified, the most common being chlorophyll *a* (**7**),

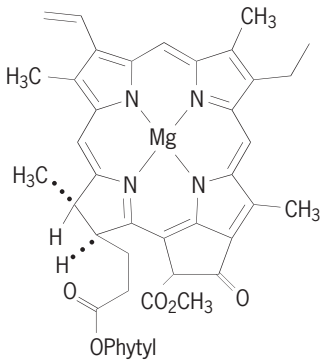

(**7**)

which is found in all oxygen-evolving photosynthetic plants. Bacteriochlorophyll *a* (**8**) is found in many photosynthetic bacteria. *See* CHLOROPHYLL; PHOTOSYNTHESIS.

Coenzyme B_{12} is the biologically active form of vitamin B_{12} (**9**). The parent macrocycle, referred to as corrin, is a highly reduced tetrapyrrole with one less bridging carbon atom than the porphyrin. The B_{12} coenzyme and related corrinoids represent the most complex nonpolymeric structures found in nature. In addition, with a unique cobalt-carbon bond, they are the only known naturally occurring organometallic complexes. Although many enzymatic reactions depend on coenzyme B_{12}, its mechanism of action is not yet fully understood. *See* COENZYME; VITAMIN B_{12}.

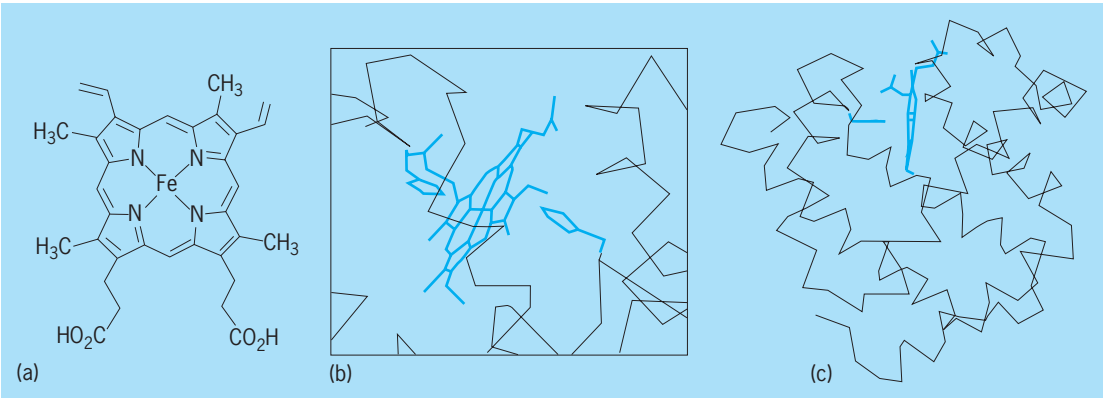

Fig. 3. Heme. (*a*) Structure. (*b*) Its bonding to myoglobin. (*c*) View from side showing that it is essentially flat.

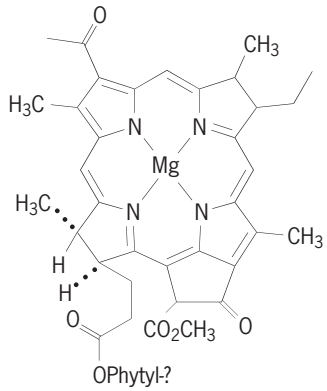

Phytyl =

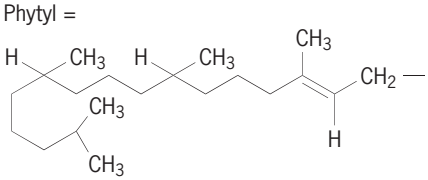

(8)

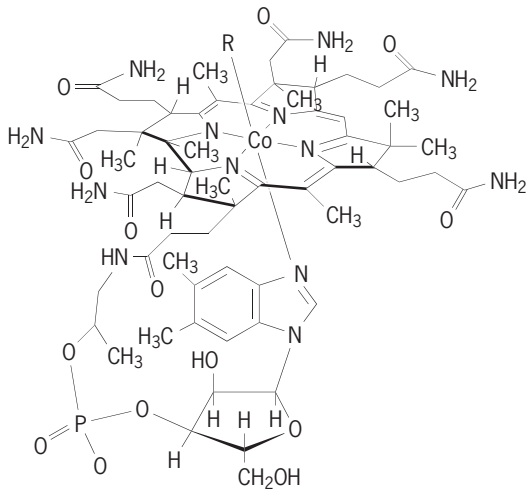

(9)

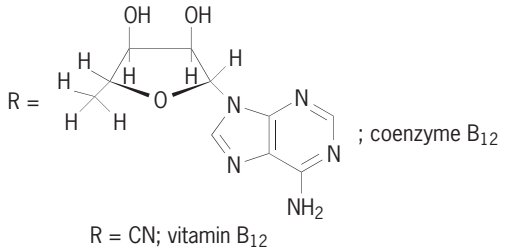

Biosynthesis. Porphyrins and their related macrocycles found in nature are derived by the same biosynthetic pathway (Fig. 2). Glycine and succinic acid are condensed to yield δ-aminolevulinic acid, which is dimerized to the pyrrole porphobilinogen. Four units of porphobilinogen are coupled to give uroporphyrinogen III (the Roman numerals here refer to the order of the peripheral substituents). Uroporphyrinogen III is the cornerstone from which all other naturally occurring tetrapyrrolic macrocycles are made, including heme, chlorophyll, bacteriochlorophyll, and vitamin B_{12}. Each of the biosynthetic steps shown in the reaction scheme in Fig. 2 is under enzymatic control. A number of genetic diseases known in humans, the porphyrias, result from malfunctioning of the heme biosynthetic pathway. These genetic disorders can result in both neurological and dermatological diseases.

In nature, porphyrins are synthesized by the tetramerization of a single pyrrolic unit (porphobilinogen; Fig. 2). Chemists have developed similar synthetic routes [reactions (1)–(3)]. Thus, the self-condensation of 2-hydroxymethylpyrrole (**7**) gives porphin (**7**) [the parent structure in this class of compounds]. Similarly, coupling of 5-aminomethyl-3,4-diethylpyrrole-2-carboxylic acid (**9**) gives octaethylporphin (**10**), while the condensation of pyrrole (**11**) and benzaldehyde (**12**) readily gives *meso*-tetraphenylporphyrin (**13**). Many additional synthetic routes to porphyrins have been developed. *See* BIOSYNTHESIS.

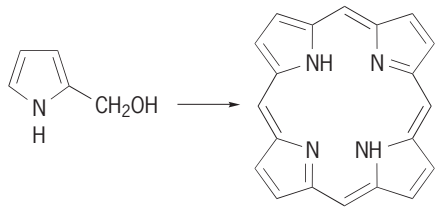

(1)

(10) **(11)**

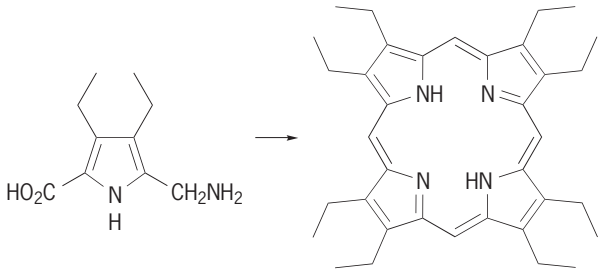

(2)

(12) **(13)**

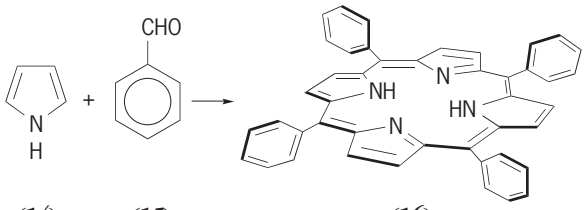

(3)

(14) **(15)** **(16)**

Applications. Porphyrins and metalloporphyrins exhibit many potentially important medicinal and industrial properties. Metalloporphyrins are being examined as potential catalysts for a variety of processes, including catalytic oxidations. They are also being examined as possible blood substitutes and as electrocatalysts for fuel cells and for the electrochemical generation of hydrogen peroxide. *See* CATALYSIS.

The unique optical properties of porphyrins make them likely candidates for photovoltaic devices and in photocopying and other optical devices. A major area where porphyrins are showing significant potential is in the treatment of a wide range of diseases, including cancer, using photodynamic therapy (PDT). Because of their unique optical and physical properties, many porphyrins have the capacity to produce singlet oxygen when they are illuminated by visible light. If the porphyrin could be made to accumulate in a tumor, then bathing the tumor in red light might result in its destruction. Photodynamic therapy using a porphyrin-based drug is in clinical trials and may eventually prove a powerful tool in the treatment of cancer and other diseases.

Photochemical techniques have been developed for use in neonatal clinics. Many infants are jaundiced and yellow at birth because of their inability to metabolize bilirubin (Fig. 3), a metabolite of heme. Bilirubin can be neurotoxic. However, by merely placing the jaundiced child under a fluorescent lamp, the buildup of bilirubin is diminished to safe levels until the infant's own bilirubin metabolism increases. *See* BILIRUBIN; COORDINATION CHEMISTRY; PYRROLE. T. Wijesekera; D. Dolphin

Bibliography. P. K. Bondy and L. E. Rosenberg (eds.), *Metabolic Control and Disease*, 8th ed., 1980; D. K. Dixon, *Advances in Porphyrin Chemistry*, I: *Cytochrome b5*, 1991; D. Dolphin (ed.), *The Porphyrins*, vols. 1–7, 1978–1979; IUPAC-IUB, Nomenclature of tetrapyrroles, *Eur. J. Biochem.*, 178:277–328, 1988.

Porphyroblast

A relatively large crystal formed in a metamorphic rock. The presence of abundant porphyroblasts gives the rock a porphyroblastic texture. Minerals found commonly as porphyroblasts include biotite, garnet, chloritoid, staurolite, kyanite, sillimanite, andalusite, cordierite, and feldspar.

Porphyroblasts are generally a few millimeters or centimeters across (see **illus.**), but some attain a diameter of over 1 ft (0.3 m). They may be bounded by well-defined crystal faces, or their outlines may be highly irregular or ragged. Very commonly they are crowded with tiny grains of other minerals that occur in the rock.

Some porphyroblasts appear to have shoved aside the rock layers (foliation) in an attempt to provide room for growth. Others clearly transect the foliation and appear to have replaced the rock. The presence of ghostlike traces of foliation, in the form of stringers and trails of mineral grains, passing uninterruptedly through a porphyroblast is further evidence of replacement.

Porphyroblasts have many features in common with phenocrysts but are to be distinguished from the latter by the fact that they have developed in solid rock in response to metamorphism. Most commonly they develop in schist and gneiss during the late stages of recrystallization. As the rock becomes reconstituted, certain components migrate to favored sites and combine there to develop the large crystals. *See* GNEISS; METAMORPHIC ROCKS; PHENOCRYST; SCHIST. Carleton A. Chapman

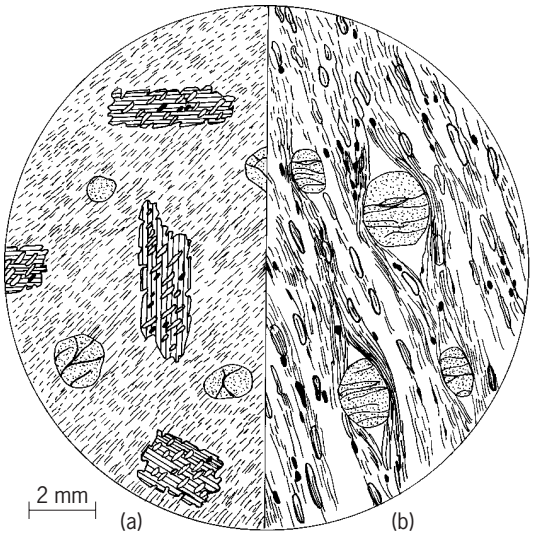

Porphyroblasts. (*a*) Porphyroblastic mica schist. Large crystals (porphyroblasts) of biotite mica formed late and replaced the rock except for the enclosed grains (quartz) aligned parallel to foliation. Smaller porphyroblasts of garnet replaced the rock completely. (*b*) Porphyroblastic quartz-mica schist. Large porphyroblasts of garnet have grown by spreading apart mica-rich layers of schist. Elongate grains (gray) are flakes of biotite mica. Finer flakes of muscovite mica are widespread and give a pronounced schistosity. They are closely packed where crowded by garnet crystals. Elongate grains of iron and titanium oxide (black) are also aligned.

Porphyry

An igneous rock characterized by porphyritic texture, in which large crystals (phenocrysts) are enclosed in a matrix of very fine-grained to aphanitic (not visibly crystalline) material.

Porphyries are generally distinguished from other porphyritic rocks by their abundance of phenocrysts and by their occurrence in small intrusive bodies (dikes and sills) formed at shallow depth within the earth. In this sense porphyries are hypabyssal rocks. *See* PHENOCRYST.

Compositionally, porphyries range widely, but varieties may be distinguished by prefixing to the term the common rock name which the porphyry most closely resembles (such as granite porphyry, rhyolite porphyry, syenite porphyry, and trachyte porphyry).

Porphyries are gradational to plutonic rocks on the one hand and gradational to volcanic rocks on the other. In the granite clan, for example, six porphyritic types may be recognized: porphyritic granite, granite porphyry, rhyolite porphyry, porphyritic rhyolite, vitrophyre, and porphyritic obsidian or porphyritic pitchstone.

A rock of granitic composition with abundant large phenocrysts of quartz and alkali feldspar in a very fine-grained matrix of similar composition is a porphyry, or more specifically a granite porphyry.

Granite porphyry passes into porphyritic granite as the grain size of the matrix increases and abundance of phenocrysts decreases. Thus the rock becomes a granite, or more specifically a porphyritic granite. Granite porphyry passes into rhyolite porphyry as the grain size of the matrix and abundance of phenocrysts decrease. The main distinction between rhyolite porphyry and porphyritic rhyolite is mode of occurrence. Rhyolite porphyry is intrusive; porphyritic rhyolite is extrusive. Porphyritic rocks with a glass matrix are known as vitrophyres. With decrease in number of phenocrysts, vitrophyre passes into porphyritic obsidian or porphyritic pitchstone. *See* GRANITE.

The phenocrysts of porphyries consist commonly of quartz and feldspar. Quartz occurs as well-formed (euhedral) hexagonal bipyramids, which in thin section under the microscope exhibit a diamond, square, or hexagonal outline. Individual phenocrysts may show more or less rounding or resorption with deep embayments. Alkali feldspar is usually euhedral sanidine, orthoclase, or microperthite. Plagioclase occurs more in association with phenocrysts of hornblende or other dark-colored (mafic) minerals. Porphyritic rocks with predominantly mafic phenocrysts (olivine, pyroxene, amphibole, and biotite) are commonly classed as lamprophyres. *See* LAMPROPHYRE.

Porphyries occur as marginal phases of medium-sized igneous bodies (stocks, laccoliths) or as apophyses (offshoots) projecting from such bodies into the surrounding rocks. They are also abundant as dikes cutting compositionally equivalent plutonic rock, or as dikes, sills, and laccoliths injected into the adjacent older rocks. *See* IGNEOUS ROCKS.

William Ingersoll Rose, Jr.

Positron

The elementary antiparticle of an electron with mass equal to that of an electron but with positive rather than negative charge. Its existence was theoretically predicted by P. A. M. Dirac in 1928. It was first observed by C. D. Anderson in 1932. Positrons and electrons have the same spin and statistics. Positrons are naturally produced by radioactive decay of many heavier particles; this is referred to as beta-plus decay. Electron-positron pairs are also produced when high-energy photons or electrons interact with matter. *See* ANTIMATTER; ELECTRON; ELECTRON-POSITRON PAIR PRODUCTION; ELECTRON SPIN; ELEMENTARY PARTICLE; FERMI-DIRAC STATISTICS; PARTICLE ACCELERATOR; QUANTUM STATISTICS; RADIOACTIVITY; SPIN (QUANTUM MECHANICS).

In addition, quantum field theory predicts that positrons can be created when electrons are subjected to very strong, static electric fields. For this to occur, the field strength must exceed the energy equivalent of twice the rest mass of an electron, $2m_0c^2$, where m_0 is the electron rest mass and c is the speed of light. Experimentally, such strong fields can be achieved by using heavy-ion accelerators to collide two heavy atoms and temporarily form a superheavy atom with an atomic number $Z > 173$. However, to date there has been no definitive observation of positrons produced via this method. *See* NUCLEAR MOLECULE; QUASIATOM; SUPERCRITICAL FIELDS.

A positron is, in itself, stable, but cannot exist indefinitely in the presence of matter where it will ultimately collide with an electron. The two particles will be annihilated as a result of this collision, and photons will be created. However, a positron can become bound to an electron to form a short-lived "atom" termed positronium. *See* POSITRONIUM.

Positrons are of interest or have practical uses in many fields. In the field of physics, they are used to study differences between particles and antiparticles and how particles and antiparticles interact. In 1995 at the CERN laboratory in Geneva, Switzerland, these studies took on a new role when positrons and antiprotons (the antiparticle of a proton) were combined to produce antihydrogen atoms. The original studies produced only a few antihydrogen atoms, but methods have improved and now it is routine to produce a hundred antihydrogen atoms per second. Present and future antihydrogen studies include investigating whether excited antihydrogen produces the same light emission as hydrogen, whether they have the same magnetic properties, and whether they interact with gravity in the same manner. Should differences be found, existing laws of physics will have to be altered, and answers as to why the observable universe appears to be made entirely of matter and not antimatter may emerge. Another interest in the field of physics is with respect to the virtual production of electron-positron pairs by an electromagnetic field. This produces a polarization of the vacuum and results in effects such as the scattering of light by light and modification of the electrostatic Coulomb field at short distances. *See* CPT THEOREM; QUANTUM ELECTRODYNAMICS.

In medicine, positron emission tomography (PET) is used to produce three-dimensional images of the inside of the body. It is primarily used to image the brain and to identify tumors. In this procedure, a radioisotope that undergoes positive beta decay is injected. After a wait period, images are produced by monitoring the two photons that are produced and emitted in opposite direactions (to conserve momentum) when an emitted positron annihilates with an electron. *See* MEDICAL IMAGING; NUCLEAR MEDICINE.

In the field of material science, positrons are routinely used to study defects in the crystalline structure of solids or the properties of electrons within the solid. To study defects, that is, either missing or replaced atoms in the solid, positrons are injected into the solid and the annihilation photons are measured as a function of time. To annihilate, a positron must spend sufficient time close to an electron; this occurs predominantly in voids in the crystal where the positron can become temporarily "trapped." Thus, shorter lifetimes imply more defects. By precisely measuring the directions of the annihilation photons that are emitted back-to-back, the location of the defects can be determined. This back-to-back

emission is also used to provide information about the electron density and momentum distribution within the solid. *See* ANGULAR CORRELATIONS; CRYSTAL DEFECTS; QUANTUM THEORY OF MATTER.

In astrophysics, positron-filled jets of gas have been observed at the center of the Milky Way Galaxy by National Aeronautics and Space Administration's *Compton Gamma-Ray Observatory* (*CGRO*). It is unclear how these antimatter jets are produced, but one suggestion is that they are related to the birth of stars occurring near the large black hole at the center of the Milky Way Galaxy. *See* BLACK HOLE; GAMMA-RAY ASTRONOMY; MILKY WAY GALAXY.

At some future date, positron-electron or other forms of antimatter-matter annihilation might be used as a power source or weapon. However, this is still far away. For example, currently 10^{10} times more energy is required to produce antiprotons than could be derived from them, and it would take all the antimatter produced each year at CERN to power a 100-watt light bulb for 15 minutes. Robert D. DuBois

Bibliography. K. Bethge et al., *Medical Applications of Nuclear Physics*, Springer Series: Biological and Medical Physics, Biomedical Engineering, Springer Verlag, 2004; D. A. Bromley (ed.), *Heavy Ion Science*, vol. 5, 1985; M. Charlton and J. Hangst, Probing the antiworld, *Phys. World*, 18(10):22–27, October 2005; P. G. Coleman and P. Coleman (eds.), *Positron Beams and Their Applications*, World Scientific Publishing, 2000; S. Gasiorowicz, *The Structure of Matter: A Survey of Modern Physics*, 1979; J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons*, rev. ed., 1975; C. M. Surko and F. A. Gianturco (eds.), *New Directions in Antimatter Chemistry and Physics*, Springer Verlag, 2001.

Positronium

The bound state of the electron and its antimatter counterpart, the positron. Positronium (abbreviated Ps) is perhaps the simplest bound state, since the constituents are structureless and the electromagnetic binding force is well understood. As in the hydrogen atom, the states of positronium are quantized. The spins of the two spin-$^1/_2$ constituents combine to yield total spin zero (singlet state) or one (triplet state). Singlet and triplet states are called parapositronium (p-Ps) and orthopositronium (o-Ps), respectively. States of positronium are labeled by principal quantum number n, orbital angular momentum, total spin, and total angular momentum (orbital plus spin). The energy levels are approximately -6.80 eV/n^2, with numerous small corrections due to relativity, virtual annihilation, and other electromagnetic effects. Additional corrections from strong and weak interactions are negligible. *See* ANGULAR MOMENTUM; ATOMIC STRUCTURE AND SPECTRA; ELECTRON; ELECTRON SPIN; POSITRON; SPIN (QUANTUM MECHANICS).

Positronium has a finite lifetime due to electron-positron annihilation into energy in the form of gamma-ray photons. The energetics of this reaction is governed by Albert Einstein's equation $E = mc^2$, where m is the total positronium mass (about twice the electron mass), c is the speed of light, and E is the total photon energy. A selection rule constrains the allowed decays, so that singlet positronium decays into an even number of photons (usually two), and triplet positronium into an odd number (usually three). Decay into a single photon is forbidden by energy-momentum conservation. The lifetimes are approximately 10^{-10} s for the singlet state and 10^{-7} s for the triplet state. *See* CONSERVATION LAWS (PHYSICS); GAMMA RAYS; REST MASS; SELECTION RULES (PHYSICS).

Positronium has been extensively studied since it was first produced by Martin Deutsch in 1951. High precision measurements have been done on the energy levels and decay rates of the lowest-energy states, and on energy levels of the first excited states. Because positronium is relatively free of strong and weak interaction complications, comparison between theory and experiment for these measured quantities provides a stringent check of quantum electrodynamics (QED). Positronium has been used to check for violations in fundamental symmetries and as a probe for physics beyond the standard model. Positronium is an analog for the more complicated quarkonium (quark-antiquark) bound states, particularly charmonium and bottomonium. *See* ELEMENTARY PARTICLE; MESON; QUANTUM ELECTRODYNAMICS; QUARKS; STANDARD MODEL; SYMMETRY LAWS (PHYSICS).

Annihilation radiation from positronium forms a component of the gamma-ray spectrum observed by astronomers, in particular from the galactic center. *See* GAMMA-RAY ASTRONOMY; MILKY WAY GALAXY.

Formation. Positronium is formed in the laboratory when a positron, usually obtained from the decay of a radioactive isotope such as sodium-22 (^{22}Na), captures an electron. The energetic positrons emitted in the nuclear decay must first be slowed to a few eV of kinetic energy before positronium will form. This slowing can be accomplished through inelastic collisions with the atoms or molecules of a material, often the same material from which the electron is obtained. Formation media include gases and powders. Positronium can also be formed when a beam of slow positrons is incident on a thin foil, a gas target, or the surface of a solid in vacuum. *See* CHARGED PARTICLE BEAMS.

Spectroscopy. Many of the energy intervals involving $n = 1$ and $n = 2$ states of positronium have been measured.

Hyperfine structure. The difference in energy between the $n = 1$ triplet and singlet states of positronium is traditionally referred to as the hyperfine splitting. The measured value of the corresponding frequency is 203,389 ± 1 MHz. Theoretical calculations based on QED are of comparable precision, and give a result that differs from this by about 3 MHz. This small difference is not thought to represent a significant problem.

Optical transition. The 5.10-eV interval between the

$n = 1$ and $n = 2$ triplet states with zero orbital angular momentum can be bridged by the simultaneous absorption of two photons of visible light. When the photons are arranged to move in exactly opposite directions, the usual frequency spread caused by the Doppler shift is greatly reduced. Measurements done using counterpropagating high-energy laser pulses incident on ground-state positronium in a vacuum give a result for the corresponding transition frequency of 1,233,607,216.4 ± 3.2 MHz. Theoretical results are in accord with this value.

Fine structure. Relativistic corrections, magnetic effects, and virtual annihilation split the $n = 2$ level into several sublevels. The frequencies of many of these sublevel differences have been measured by monitoring the absorption of microwave radiation by a positronium beam moving through a microwave cavity. Experiment and theory are in good agreement for these transitions. *See* MOLECULAR BEAMS.

Decay rates. The lifetime of the relatively long-lived $n = 1$ orthopositronium has been measured to high precision. A typical experimental configuration consists of a ^{22}Na positron source, a gas chamber in which the positronium forms, photomultiplier gamma detectors, and timing electronics. The emission of a position from the source is accompanied by a gamma ray which can be detected and used as a start signal. The stop signal is the detection of the final annihilation radiation. Following the prompt peak caused by direct positron-electron annihilation and the fast decay of parapositronium, there is an extended exponential distribution coming from orthopositronium decay. Since the presence of the gas increases the rate of decay, an extrapolation to zero gas pressure must be done to obtain the vacuum rate. High-precision measurements are also done for positronium formed in powders and in thin coatings in vacuum. The best lifetime measurements give a value of 142.05 ± 0.03 nanoseconds, in excellent agreement with theoretical calculations based on quantum electrodynamics. *See* GAMMA-RAY DETECTORS; PHOTOMULTIPLIER.

The lifetime of parapositronium is too short for a high-precision direct timing procedure. Instead, positronium is produced in a magnetic field, which mixes the triplet and singlet states and increases a component of the triplet decay rate by an amount that depends on the singlet rate and the field strength. The parapositronium lifetime, measured using this approach, is 125.14 ± 0.03 picoseconds, in good agreement with theoretical expectation.

Materials science probe. Since the formation of positronium requires the close approach of a positron and an electron, beams of slow positrons can be used as probes of the electron density in gases, in insulating solids, or near surfaces. Since the singlet and triplet forms of positronium have very different lifetimes, and transitions between the two states can be induced by environmental factors, study of the decay of positronium can also provide information about electron densities and magnetic fields on a microscopic scale. This is useful in the study of density fluctuations in gases near the critical point for condensation into liquids or solids. This technique is also used to study vacancies in metals and semiconductors, and is now being applied to high-temperature superconductors and biological materials. *See* CRITICAL PHENOMENA; CRYSTAL DEFECTS; SUPERCONDUCTIVITY.

Gregory S. Adkins

Bibliography. S. Berko and H. N. Pendleton, Positronium, *Annu. Rev. Nucl. Part. Sci.*, 30:543–581, 1980; M. Charlton and J. W. Humberston, *Positron Physics*, Cambridge University Press, 2001; T. Hyodo et al. (eds.), *Positron Annihilation ICPA-13*, Trans Tech Publications, Uetikon-Zuerich, 2004 (http://www.scientific.net/0-87849-936-9/); Y. C. Jean, P. E. Mallon, and D. M. Schrader (eds.) , *Positron and Positronium Chemistry*, World Scientific, Singapore, 2003; B. Maglic (ed.), Discovery of Positronium, *Adventures Exp. Phys.*, 4:64–127, 1975; A. Rich, Recent experimental advances in positronium research, *Rev. Mod. Phys.*, 53:127–165, 1981.

Postglacial vegetation and climate

The Pleistocene or Glacial Epoch is less remarkable for radical evolutionary changes in plant life than for repeated shifts of position because of climatic changes and the resulting ice movements. The influence of these changes is known to have extended far beyond the limits of glaciation, particularly in mountainous regions, whose belts or zones of vegetation migrated up and down in response to climate. *See* PLEISTOCENE.

Four main glacial stages are recognized in North America within the Pleistocene Epoch: Kansan, Nebraskan, Illinoian, and Wisconsin. The climatic change which initiated the last major ice advance of the late Wisconsin began more than 30,000 years ago. This ice advance reached its maximum extent in southern Illinois, Indiana, and Ohio about 18,000–20,000 years ago. The climate began to warm about 16,000–18,000 years ago and the glaciers melted more rapidly than they advanced. Glacier ice withdrew from the continental United States sometime prior to 12,000 years ago, with the exception of residual ice caps on higher elevations in mountains in New England and in the Rocky Mountains of the West. There was a brief readvance of the ice about 10,500 years ago, known as the Valders readvance.

The retreat of the ice was not uniform; it occurred with many pauses and some readvances, each producing an elevated belt of drift called moraines. As the ice retreated, broad expanses of sand, silt, and clay mixed with rocks and boulders of all sizes were uncovered. Bare mineral soil was colonized by plants which had been forced south in advance of the ice some thousands of years before. In New England, leaves and pollen of high-arctic plants have been radiocarbon-dated at more than 15,000 years, marking a time of extremely harsh climate similar to that which is today found north of the Arctic Circle. *See* GLACIOLOGY; RADIOCARBON DATING.

Palynological studies. Attempts to understand and interpret the climatic changes that have occurred since the ice retreat began have involved many different types of study. One of the most fruitful techniques for environmental reconstruction has been the field of pollen analysis, or palynology. The technique, pioneered by the Swedish geologist Lennart von Post in 1916, was first extensively applied in the United States by Paul B. Sears in the early 1930s. The method depends upon the fact that each year flowering plants produce enormous numbers of pollen grains; although many plants are pollinated by insects, most are pollinated by wind. Austrian pine (*Pinus sylvestris*) produces as many as 1,480,000 pollen grains per male strobilus, and there may be up to 20 strobili in a single inflorescence, with more than 100 inflorescences on a single medium-sized tree. The spruce forests of southern Sweden are estimated to produce up to 75,000 tons (67,500 metric tons) of pollen each spring. Small size (10–250 micrometers) and light weight permit pollen grains to be carried for long distances by wind. Stray pollen grains of alder (*Alnus viridis*) have been found in the middle of the North Atlantic Ocean, more than 400 mi (650 km) from the nearest source. *See* PALYNOLOGY.

Eventually, pollen grains fall to the ground, the majority remaining close to the source plants. Those which fall into lakes and bogs are not decomposed, but are preserved in the annual accumulation of sediment and form a record of the numbers and kinds of pollen of the local and regional flora. Pollen grains are distinctive and can be recognized at least to genus level in most cases. A pollen diagram showing the succession of dominant pollen types throughout a stratigraphic section of bog or lake sediments can then be used to interpret the vegetational and climatic changes that have occurred since the deposit was formed. *See* BOG; POLLEN.

Regional pollen diagram. Interpretation of a pollen diagram is facilitated by an understanding of the kinds of pollen records produced by different environmental types. Surface pollen samples from six different vegetational and climatic regions are shown in the **illustration**. Nichicun Lake, the northernmost lake, is at the northern edge of Boreal Parkland in northern Quebec, a region characterized by isolated stands of spruce and pine trees with extensive open areas of tundra. Just south of James Bay is Cedar Lake, Ontario, near the northern edge of the Boreal Forest, which consists of dense stands of spruce and pine together with considerable amounts of fir and birch. Further south, Leech Lake is in the Northern Great Lakes–St. Lawrence Forest with dense stands of spruce, pine, and birch. In the southern part of the Great Lakes–St. Lawrence Forest, Little Round Lake is surrounded by stands of spruce and pine with extensive upland areas in oak and hickory and considerable amounts of hemlock and northern white cedar around the lake border. Otter Pond, in the Maple-Basswood region of the Deciduous Forest in southern Wisconsin, is closely similar to Silver Lake, Ohio, which is in the Oak-Hickory region of the Deciduous Forest. The principal difference between these two pollen samples is that there is a larger amount of hickory pollen at Silver Lake and a larger amount of beech pollen at Otter Pond.

At Silver Lake, in sediments which are approximately 12,000 years old, the pollen spectra most closely resemble the environment represented by the Nichicun Lake surface samples. A computer program to calculate product-moment correlations (r) on 25 different pollen types shows a high correlation between the contemporary Nichicun Lake surface sample and 12,000-year-old Silver Lake sediments ($r = +.925$ for $N = 25$). This value is not different from a population coefficient of $r = +1.0$ at the .995 significance level. By 11,000 years ago the climate had warmed sufficiently to permit the establishment of a closed Boreal forest, and a computer-generated search of more than 200 surface pollen samples from all of eastern North America show highest correlation ($r = .896$) with Cedar Lake in the Boreal Forest. The environment of Silver Lake 6500 years ago was apparently very similar to the present environment around Otter Pond in the Maple-Basswood portion of the Deciduous Forest ($r = .912$). About 3500–5000 years ago, the climate around Silver Lake was apparently warmer and drier than at present, for the highest correlation with contemporary surface samples is found in southwestern Illinois near the boundary between the Deciduous Forest and the Tall Grass Prairie. This period was first identified in Midwestern pollen diagrams as the "Xerothermic Interval" by P. B. Sears in the early 1940s. A return to somewhat cooler and moister conditions is indicated following the Xerothermic Interval by increases in the amount of beech pollen and a corresponding decrease in oak and hickory pollen in sediment cores from Silver Lake.

A computer search of a matrix formed by 25 different pollen types in more than 200 lakes from eastern North America and 38 pollen spectra from different levels in sediment cores at Silver Lake indicates that at no time in the postglacial history of Silver Lake did the environment resemble very closely the modern pollen records of Little Round Lake or Leech Lake in Ontario.

The appearance of European colonists is indicated in the upper portions of sediment cores throughout North America by the presence of large amounts of the pollen of ragweed (*Ambrosia* sp.) and the presence of European weed pollen such as white-man's foot (*Plantago lanceolata*), dandelion (*Taraxacum* sp.), and species of *Rumex*. These plants were either brought in by the colonists or the permanent clearings made for houses and farms provided ideal conditions for the spread of these and native species.

Environmental reconstruction. Since the beginning of the Christian era, historical documents have recorded the effects of climatic change on the activities of humans. A period of benign climatic conditions around A.D. 1000 favored the westward explorations of the Vikings who colonized Iceland, Greenland, and briefly the coast of North America near L'Anse aux Meadows in Newfoundland. About A.D. 1150 the climate began to turn colder and

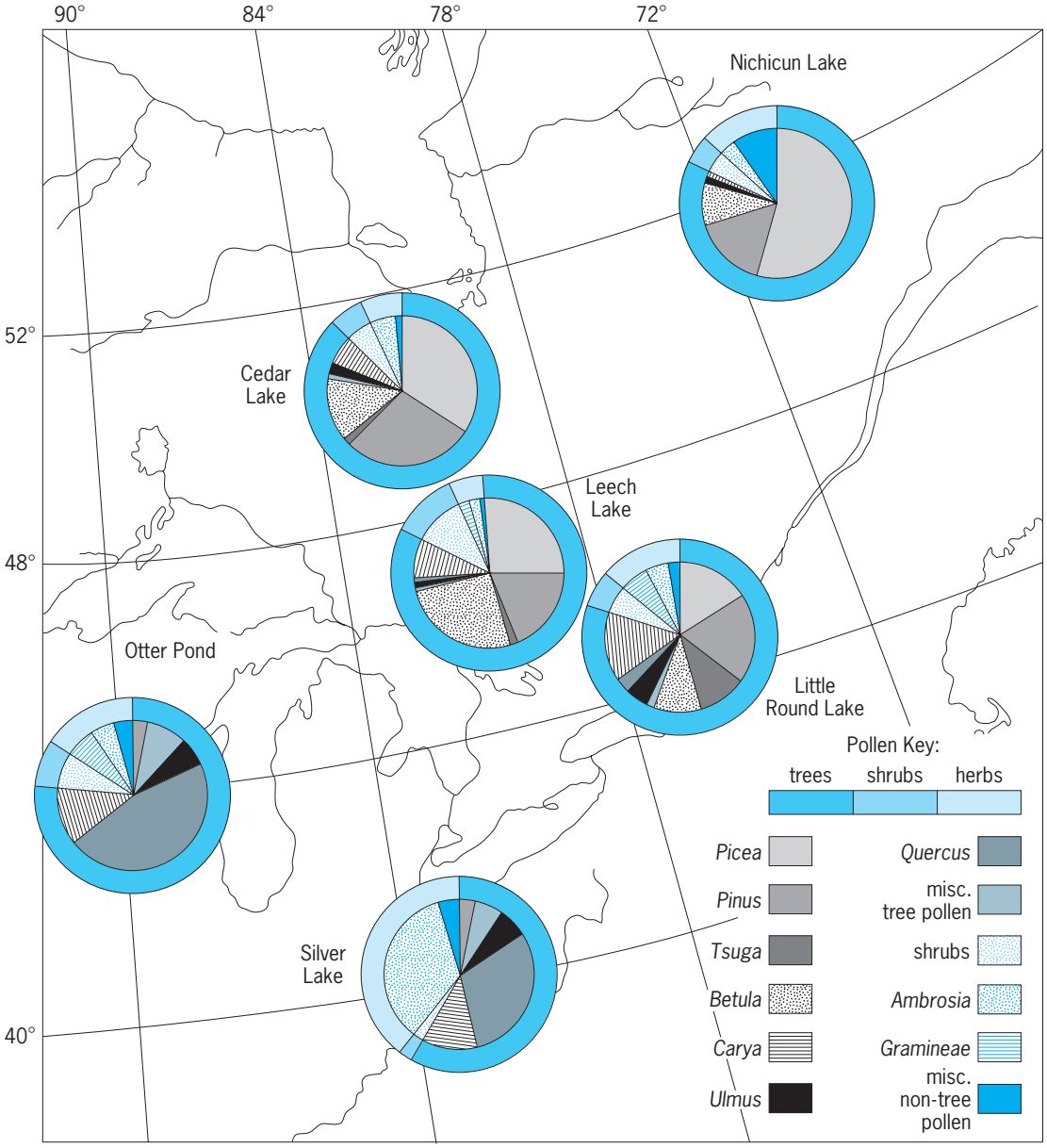

Pollen diagrams from six different vegetational and geographic regions.

winter storminess increased. The climatic change was not the least of the reasons for abandoning the settlements in North America and Greenland.

By the middle of the seventeenth century, glaciers in high mountains in North America and Europe had advanced dramatically, in some cases overwhelming farms or small settlements. Toward the close of the seventeenth century the climate again began to warm and glaciers around the world began to retreat. This warming trend is also recorded by dates of the opening of cherry blossoms in monastery gardens in Asia and has continued until the present. There is some evidence that the rate of warming has decreased since 1940 and that rainfall has been increasing north of about 40°N latitude.

Although the amount of vegetational change in the last 14,000 years has been great, the actual temperature changes seem rather small. Temperature measurements derived from measurement of the ratios of oxygen-16 to oxygen-18 in the carbonates deposited in marine Foraminiferida indicate that mean annual temperatures were not more than 12–14°F (6.5–8.5°C) lower than those during the maximum advance of the late Wisconsin ice. Similarly, measurements of ocean water temperature show an increase of only 0.9°F (0.5°C) over the last 200 years, covering the period of the retreat of the "Little Ice Age" maximum of the seventeenth century. *See* FORAMINIFERIDA.

Radiocarbon dating of late Pleistocene events indicate that environmental changes were nearly synchronous throughout the world at least for the last 20,000 years. Although the evidence is less clear for earlier periods of ice advance, four main periods of ice advance with similar substages are recognized in Europe and Russia.

Postglaciation environmental phases in North America

Time 10^3 years B.P.*	Environment: South	Environment: North	Pollen sequences: Midwest	Pollen sequences: New England	Pollen sequences: Alaska	Climate
0			Oak-hickory-beech-maple-ragweed	Oak-hemlock-chestnut-ragweed	Spruce-pine [Little Ice Age]	W
1						C-M
2	POSTGLACIAL		Oak-beech	Oak-chestnut		
3		LATE GLACIAL	Oak-hickory (XEROTHERMIC)	Oak-hickory (HYPSITHERMAL)	Spruce	W-D
4						
5						
6			Oak-beech	Oak-chestnut		
7					Park-tundra	
8			Oak-elm	Oak-hemlock	— — — — — —	W-M
9		FULL-GLACIAL				
10			Pine	Pine		C-D
11			Spruce-pine	Spruce-fir		C-M
12			Spruce-fir-birch	Spruce parkland		
13						C
14			Spruce parkland	Tundra		C
15						

*B.P. = before present (arbitrarily set at A.D. 1950).

In the reconstruction of climatic change within or near the glacier boundary, it is possible to distinguish three major environmental events: Full-glacial, which represents a period of time during which ice either covered an area or its presence prevented the establishment of a vegetational record; Late-glacial, an indefinite period of early colonization by plants following the retreat of the ice but prior to the establishment of closed forests; and Postglacial, marking the development of dense forests. In Europe, the Postglacial is arbitrarily taken as the period following the emplacement of the Fennoscandian end moraines about 10,000–11,000 years ago. In North America, their equivalent is the Valders end moraines in Wisconsin. The distinction between Full-glacial, Late-glacial, and Postglacial in North America has generally been recognized to depend upon particular localities. For example, Postglacial time begins about 12,500 years ago in southern Ohio and Indiana as well as in southern New England. These areas were ice-free several thousand years before upper Michigan, southern Ontario, and the Maritime Provinces. Much of Greenland and the arctic islands of Canada can be considered to be still in Full-glacial.

Within Postglacial time in North America, two subdivisions are generally recognized. The Hypsithermal period, about 9500–2500 years before present, marks a period in Postglacial time during which the climate was at least as warm as it is at present. The warm maximum of Postglacial time which is not everywhere recognizable in northeastern North America has been called the Xerothermic Interval and marks a time of about 5000–3500 years ago during which prairie, today restricted to the region west of Illinois, extended eastward into Ohio.

Postglaciation environmental changes are summarized in the **table**. The area included is limited to the major glaciated portions of North America partly because more work has been done there and partly because the record beyond the glacier boundary is not as clear. Studies in southwestern North America have demonstrated considerable shifts in woodland and forest belts moving up and down the sides of the Rocky Mountains, and these changes seem to be broadly synchronous with the changes observed inside the glacier boundary. Additional work on ancient lake basins will help to establish the environmental changes outside the glaciated regions.

J. Gordon Ogden, III

Bibliography. R. J. Dodd and R. J. Stanton, Jr., *Paleoecology: Concepts and Applications*, 2d. ed., 1990; *Symposium on Changes in Climate*, Rome, UNESCO, 1965; H. E. Wright, Jr., and D. G. Frey (eds.), *International Studies on the Quaternary*, 1965.

Posttraumatic stress disorder

An anxiety disorder in some individuals who have experienced an event that poses a direct threat to the individual's or another person's life. The characteristic features of anxiety disorders are fear, particularly in the ongoing absence of a real-life threat to safety, and avoidance behavior.

Diagnosis. A diagnosis of posttraumatic stress disorder requires that four criteria be met. First, the individual must have been exposed to an extremely stressful and traumatic event beyond the range of normal human experience (for example, combat, armed robbery, rape, or natural disasters such as earthquakes). Second, the individual must periodically and persistently reexperience the event. This reexperiencing can take different forms, such as recurrent dreams and nightmares, an inability to stop thinking about the event, flashbacks during which the individual relives the trauma, and auditory hallucinations. Third, there is persistent avoidance of events related to the trauma, and psychological numbing that was not present prior to the trauma.

Individuals often withdraw from others, feel emotionally deadened, lose interest in activities, and avoid contact with anything even remotely similar to the traumatic event. Some children with posttraumatic stress disorder regress to an earlier stage of development. Fourth, enduring symptoms of anxiety and arousal are present. These symptoms can be manifested in different forms, including anger, irritability, a very sensitive startle response, an inability to sleep well, and physiological evidence of fear when the individual is reexposed to a traumatic event. Posttraumatic stress disorder can occur at any age, and the symptoms typically begin soon after the cessation of the traumatic event.

Most knowledge of posttraumatic stress disorder comes from research on combat soldiers and other victims of war; other knowledge comes from rape and assault victims, and survivors of disasters or medical illness. There is a growing appreciation for the conceptualization of posttraumatic stress as a spectrum disorder. At one pole are individuals who are exposed to significant trauma but do not go on to develop symptoms of posttraumatic stress. At the opposite pole are individuals who manifest the full range of posttraumatic stress disorder symptoms. In between these extremes are individuals who develop mild forms of the disorder or manifest only certain symptom complexes which do not qualify as an actual diagnosis of posttraumatic stress disorder. Recent evidence suggests that there may be gender differences in how posttraumatic stress is manifested. Men appear to be significantly more likely than women to experience irritability and to use alcohol to excess. Within both genders, however, symptoms tend to follow an acute stress reaction, occur early, and persist for many months.

Causative factors. In the United States, most adult individuals have experienced at least one event that qualifies as sufficiently traumatic to stimulate the development of posttraumatic stress/posttraumatic stress disorder, and approximately 10% meet diagnostic criteria for posttraumatic stress disorder at some time in their lives. This incidence is probably influenced by the subjective nature of perceived stress. What is mildly stressful for some individuals may be moderately stressful or highly stressful for others. Thus, it can be seen that most individuals do not develop a pathological reaction to psychological trauma. Estimates of occurrence of the disorder in large groups of combat veterans have ranged from 15 to 50%. The degree of the posttraumatic stress response is likely to be influenced by a complex interaction of biological processes, personality, nature of the trauma, and posttraumatic events. High levels of avoidance behavior may actually potentiate and prolong the symptoms. Failure to expose oneself to stimuli that evoke fear allows the fear response to remain strong. Increased exposure, however, diminishes the fear response. In short, many symptoms of posttraumatic stress disorder may be self-perpetuating. Events that occur after the cessation of the traumatic event also seem to be important in whether symptoms become severe and intractable.

The prediction of posttraumatic stress disorder and symptoms of posttraumatic stress following a trauma are also complex and multidimensional. Biological predictors include lower baseline levels of cortisol, hippocampal atrophy, and deficiencies in hypothalamic-pituitary-adrenal activity. Cognitive and personality factors also play a role in predicting the severity of posttraumatic stress following a trauma. Individuals who tend to exaggerate the probability of future negative events, to become somatically focused, to exhibit cognitive biases, and to experience shame following the trauma are at an elevated risk of developing posttraumatic stress disorder. Tendencies to avoid aversive information are also related to the development of chronic posttraumatic stress. Thus, the development of posttraumatic stress disorder is likely influenced by a complex pattern of memory encoding, self-appraisal, coping strategies, cognitive styles, the nature of the trauma exposure, and biological responses that occur soon after the traumatic event and continue in a deteriorating cascade.

Physiological changes. There are also important biological events that accompany stress. Physiological arousal responses in individuals with posttraumatic stress disorder include increases in heart rate, respiration rate, and skin conductivity upon reexposure to traumatic stimuli. Posttraumatic stress disorder may also be associated with structural and physiological changes in the brain. For example, when individuals are placed under high levels of stress, the activity levels of two brain neurotransmitters, norepinephrine and serotonin, are reduced. Reductions in these chemical transmitters are not typically found under mild or moderate stress conditions. Stressful events also affect the activity level of the pituitary and adrenal glands. These fluctuations in adrenergic hormonal activity are likely to be associated with periodic alterations in the brain's endogenous opiates. Under conditions of chronic stress, hormones (glucocorticoids, in particular) contribute to impairment of cognitive functioning and can actually result in damage to brain structures, particularly the hippocampus. All these physiological changes are probably complexly related to the persistence and waxing and waning of symptoms in posttraumatic stress disorder. In addition, extreme and prolonged stress is associated with a variety of physical ailments, including heart attacks, ulcers, colitis, and decreases in immunological functioning. *See* ADRENAL CORTEX; BRAIN; NEUROBIOLOGY; NORADRENERGIC SYSTEM; SEROTONIN.

Associated disorders. Posttraumatic stress disorder is rarely the only psychological problem that an affected individual manifests. When an individual is diagnosed as having this disorder, particularly after it has been present for a number of years, it is common to also find significant depression, generalized anxiety, substance abuse and dependence, marital problems, and intense, almost debilitating anger. Although the primary symptoms of posttraumatic stress disorder (anxiety, fear, intrusion, and avoidance) are usually amenable to treatment efforts, these

secondary problems commonly associated with the chronic disorder are more difficult to treat. *See* ADDICTIVE DISORDERS; AFFECTIVE DISORDERS; ANXIETY DISORDERS.

Treatment. Posttraumatic stress disorder can be treated by pharmacological means and with psychotherapy.

Psychotherapeutic methods. Most psychological treatments for the disorder involve reexposure to the traumatic event. This reexposure is typically imaginal and can range from simply talking about the trauma to having the person vividly imagine reliving the traumatic event. This latter behavioral procedure is called implosion therapy or flooding. While flooding is not appropriate for all posttraumatic stress disorder cases, especially individuals with substance abuse and interpersonal difficulties, the procedure can dramatically decrease anxiety and arousal, intrusive thoughts, avoidance behavior, and emotional numbing. Along with specific behavior interventions, individuals with posttraumatic stress disorder should become involved in psychotherapeutic treatment for secondary problems. A wide variety of effective treatments for both posttraumatic stress disorder and associated psychiatric and psychological conditions have been demonstrated to be effective. These interventions include cognitive-behavioral therapy, group therapy, psychodynamic treatment, psychosocial rehabilitation, hyponosis, and marital and family therapy. *See* PSYCHOTHERAPY.

Pharmacological methods. Numerous medications have been demonstrated to be effective in treating posttraumatic stress disorder, including selective serotonin reuptake inhibitors, antiadrenergic drugs, and tricyclic antidepressants. Other medications that are showing promise include mood stabilizers and some antipsychotic drugs. *See* PSYCHOPHARMACOLOGY; STRESS (PSYCHOLOGY).

Robert W. Butler

Bibliography. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed., 1994; N. Breslau, Epidemiologic studies of trauma, posttraumatic stress disorder, and other psychiatric disorders, *Can. J. Psychiat.*, 47:923–929, 2002; C. R. Brewin and E. A. Holmes, Psychological therories of posttraumatic stress disorder, *Clin. Psychol. Rev.*, 23:339–376, 2003; E. B. Foa, T. M. Keanne, and M. J. Friedman (eds.), *Effective Treatments for PTSD*, 2000; B. S. McEwen, The neurobiology and neuroendocrinology of stress: Implications for post-traumatic stress disorder from a basic science perspective, *Psychiat. Clinics N. Amer.*, 25:469–494, 2002; C. Moreau and S. Zisook, Rationale for a posttraumatic stress spectrum disorder, *Psychiat. Clinics N. Amer.*, 25:775–790, 2002.

Postulate

In a formal deductive system, a proposition accepted without proof, from which other propositions are deduced by the conventional methods of formal logic. There is a certain arbitrariness as to which propositions are to be treated as postulates, because when certain proved propositions are treated as such, other propositions which were originally postulates often become proved propositions.

The term postulate is sometimes loosely used in connection with tentative assumptions with regard to matters of fact. In strict usage, the term postulate is nearly equivalent to axiom, although axiom is often loosely used to denote a truth supposed to be self-evident. *See* LOGIC.

Percy W. Bridgman; Henry Margenau

Postural equilibrium

A lifeless object is said to be in equilibrium, or in a state of balance, when all forces acting upon it cancel. The result is a state of rest. In an actively moving animal, internal as well as external forces have to be considered, and the maintenance of a balanced attitude in a body consisting of a number of parts that are loosely connected by movable joints is complex.

Maintenance of equilibrium. The maintenance of equilibrium is relatively easy in limbless animals, such as worms, snails, or starfish, which rest and move on an extensive surface of support. The tactile contact between the normal creeping surface and the substrate furnishes sensory information on the animal's orientation with respect to the plumb line. When the animal is turned on its back, the lack of contact pressure on the creeping surface and the tactile stimulation of the back initiate movements which return the animal to its normal position. This is known as the righting reflex. The accompanying change in the direction of illumination may also contribute to efficient righting.

Free-swimming and flying animals are often in a precariously poised state of equilibrium, so that the normal attitude can be maintained only by the continuous operation of corrective equilibrating mechanisms. The same applies to a lesser degree to long-legged quadrupeds, such as many mammals, and to bipeds, such as birds and some primates (including humans). In these animals the center of gravity lies some distance above the area of support which, in some cases, such as in humans, can be relatively small.

Sensory control. For most species, there is a specific orientation of the whole body or of body segments (such as the head) with respect to gravity. This orientation is based on multisensory inputs and on a set of inborn reflexes acting on the musculature. These postural reflexes also stabilize the genetically defined body orientation against external disturbances. The righting reflexes in vertebrates is an example of this basic organization. When a cat is falling upside down, the head is first reoriented toward the horizontal plane, and then the trunk axis and finally the legs are oriented vertically.

The sensory information relies on a number of sources: (1) static and dynamic information from the eyes; (2) static and dynamic mechanoreceptors

incorporated in the various types of statocysts in invertebrate animals and in the vestibular organ or labyrinth of vertebrates; (3) proprioceptor organs such as muscle spindles, Golgi endings in tendons, Pacinian corpuscles and similar encapsulated endings associated with tendons and joints, and other pressure receptors in supporting surfaces (for example, the soles of feet); and (4) sensory endings in the viscera, capable of being differentially stimulated by changes in the direction of visceral pull on mesenteries, and on other structures, *See* SENSATION; SENSE ORGAN.

Vision plays an important role in orientation in the gravitational field. In daylight, visual fixation of the horizon or reference to the bright sky as opposed to the ground can serve as a means of basic orientation. This is important in the equilibration of free-swimming and flying animals. Visual postural mechanisms contribute in part to balance in aquatic invertebrates such as shrimp and fish. These animals turn their back toward the light, showing the dorsal light reflex. Some shrimp swim belly up, showing the ventral light reflex. Among flying animals, insects in flight rely largely on visual orientation. The compound eye is well suited to the task of keeping the body in a horizontal or nearly horizontal plane. Its subdivision into separate ommatida makes it possible to hold the image of the horizon, or of other horizontal or vertical landmarks, fixed on a certain part of the retina. If the animal deviates from an even keel, the image leaves the fixation region, evoking corrective reflex movements designed to return the image to the same part of the eye. Like other arthropods, insects are equipped with hair receptors located between joints in the armor of limbs and body, monitoring their relative position. *See* EYE (INVERTEBRATE); INSECT PHYSIOLOGY.

Gravity reception. In many invertebrates and in all vertebrates, specialized receptor organs known as statocysts supply information on the animal's orientation with respect to the plumb line. A statocyst consists of a fluid-filled space, the floor of which is made up of sensory and supporting cells. The sensory cells have hair processes which project into the lumen of the statocyst. In the lumen the hair processes come into close contact with the statolith, a body of higher specific gravity than the surrounding fluid and tissues. The displacement of the statolith caused by the animal's deviation from an even keel deforms the hair processes, initiating afferent impulses in the sensory nerve. Thus, the central nervous system receives information regarding the direction of the gravity vector and the degree of deviation with respect to this vector. The otolith organ of the vertebrate labyrinth functions on the same principles and monitors the head position in space. The semicircular canal of the vertebrate labyrinth is arranged in three different planes of space and constitutes receptor organs for angular accelerations of the head, serving to stabilize the head orientation in space.

Graviception in vertebrates is also monitored by body graviceptors. For example, experiments under water or in microgravity show that afferent inputs related to body weight as well as receptors around the kidney play an important role in the body orientation with respect to gravity. *See* EAR (VERTEBRATE).

Stretch receptors. Invertebrate stretch receptors are usually not as closely associated with skeletal muscles as in vertebrates. Endings of sensory neurons are applied to connective tissue fibers in the vicinity of the joints, monitoring changes in their length brought about by flexion or extension of the joint. More complex types of stretch receptors are the chordotonal organs. Consisting of nerve endings capped by peglike structures, they are hung on an elastic chord suspended within the cavity of the limbs and trunk. The chordotonal organs are frequently involved in the control of posture and movement.

In vertebrates stretch receptors are mostly associated with muscle stretch, and in mammals they are inserted in specialized organs called the muscle spindles. They monitor muscle length and velocity. Muscle force is measured by Golgi tendon organ receptors. The stretch receptors are involved in the regulation of the segmental muscle tone and in the antigravity postural tone maintained by the extensor muscles opposing the gravity forces. They also contribute to a central representation of the body segment position.

Organization of posture. In higher terrestrial vertebrates, during stance the center of mass (CM) is usually situated high above the ground due to the support of the body by the limbs. A critical aspect of posture in quadrupedal and bipedal stance is equilibrium maintenance which is preserved only when under static conditions the projection of the center of mass remains inside the support base. This positioning of the center of mass is based on two main controls. A "bottom up" control is based on the afferent nerve impulses, cutaneous and proprioceptive, from the feet and the ankle joint muscles. These nerve impulse serve in building up posture from the feet to the head. A "top down" control starts from the head, and is predominant during dynamic activities such as locomotion. Due to labyrinthine afferent nerve impulses, the head axis orientation remains stable with respect to space. The movement-related visual afferents recorded by the retina monitor the head displacements with respect to space and adjust the body posture as a function of these inputs.

Internal representation. In order to explain the dual mode of postural organization, bottom up and top down, it was proposed that the control of posture by the central nervous system depended on a hierarchical organization. At a higher level, an internal representation of posture or postural body schema was organized. This representation was rather stable, partly genetically determined and partly acquired through learning. This level included (1) a representation of the body-segments geometry from the feet to the eyes mainly based on proprioceptive muscle

afferents; (2) a representation of the segments mass and center of inertia; (3) a representation of the body posture and mass with respect to the external world, based on multisensory inputs (visual, labyrinth, cutaneous, graviceptor). This level of representation has been illustrated, for example, by the illusory perceptions provoked by artificial sensory stimulations such as those of the spindle afferents under the action of tendon vibration. When holding the fingertip on the nose, the subject feels, during arm biceps tendon vibration, an elongation of the nose (Pinocchio nose). Biceps vibration stimulates the spindle afferents which are normally excited during stretching of the muscle (extension of the arm). As the arm is not moving, the subject's interpretation is the elongation of the nose.

A lower level in the central organization is that of the postural reactions. Balance disturbances are corrected by several mechanisms, depending on their intensity: (1) the joint stiffness related to the muscle contraction around the joints which minimize the angular displacements; (2) the postural reactions characterized by plurisegmental muscle activation patterns reestablishing the initial reference posture through a whole-body rotation around the ankle joint (ankle strategy) or through a hip flexion or extension (hip strategy); (3) stepping; (4) arm stretching during falling.

Central organization. Two levels of control are involved in maintaining balance. A first level includes the spinal cord and the brainstem, where a set of inborn reflexes are organized for stance regulation and head orientation. Most postural reflexes rely on networks at that level. The cerebellum is involved in the adaptation of these reflexes to the external constraints. *See* REFLEX.

A second level of control includes cortical areas involved in multisensory integration and control as well as the basal ganglia. The postural body schema and the body orientation with respect to the external world are organized mainly at that level, with a predominant role in the right hemisphere. Coordination between balance control and locomotion or movements also depends on these higher levels.

Jean Massion

Bibliography. C. Gans, *Biomechanics: An Approach to Vertebrate Biology*, 1980; F. Horak and J. M. Macpherson, *Postural Orientation and Equilibrium in Handbook of Physiology*, sec. 12, 1996; T. D. M. Roberts (ed.), *Neurophysiology of Postural Mechanisms*, 2d ed., 1978.

Potassium

A chemical element, K, atomic number 19, and atomic weight 39.102. It stands in the middle of the alkali metal family, below sodium and above rubidium. This lightweight, soft, low-melting, reactive metal (see **table**) is very similar to sodium in its behavior in metallic forms. *See* ALKALI METALS; PERIODIC TABLE; RUBIDIUM; SODIUM.

1												13	14	15	16	17	18
1 **H**	2																2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Potassium chloride, KCl, finds its main use in fertilizer mixtures. It also serves as the raw material for the manufacture of other potassium compounds. Potassium hydroxide, KOH, is used in the manufacture of liquid soaps, and potassium carbonate in making soft soaps. Potassium carbonate, K_2CO_3, is also an important raw material for the glass industry. Potassium nitrate, KNO_3, is used in matches, in pyrotechnics, and in similar items which require an oxidizing agent.

Marshall Sittig

Potassium is a very abundant element, ranking seventh among all the elements in the Earth's crust, 2.59% of which is potassium in combined form. Seawater contains 380 parts per million, making potassium the sixth most plentiful element in solution.

Potassium is even more reactive than sodium. It reacts vigorously with the oxygen in air to form the monoxide, K_2O, and the peroxide, K_2O_2. In the presence of excess oxygen, it readily forms the superoxide, KO_2.

Potassium does not react with nitrogen to form a nitride, even at elevated temperatures. With hydrogen, potassium reacts slowly at 200°C (392°F) and rapidly at 350–400°C (662–752°F). It forms the least stable hydride of all the alkali metals.

The reaction between potassium and water or ice is violent, even at temperatures as low as −100°C (−148°F). The hydrogen evolved is usually ignited in reaction at room temperature. Reactions with aqueous acids are even more violent and verge on being explosive.

The potassium ion (K^+) is the most common intracellular cation and is essential for maintaining osmotic pressure and electrodynamic cellular properties in organisms. The intracellular potassium ion concentrations are typically high for most cells, whereas the potassium ion concentrations present in extracellular fluids are significantly lower. The hydrolysis of the coenzyme adenosine triphosphate (ATP) is mediated by the membrane-bound enzyme Na^+, K^+-ATPase. This enzyme is called the sodium pump and it is activated by both potassium and sodium ions; however, many enzymes are activated by potassium ions alone (for example, pyruvate kinase, aldehyde dehydrogenase, and phosphofructokinase).

Physical properties of potassium metal

Property	Temperature °C	Temperature °F	SI units	Customary (engineering) units
Density	100	212	0.819 g/cm^3	51.1 lb/ft^3
	400	752	0.747 g/cm^3	46.7 lb/ft^3
	700	1292	0.676 g/cm^3	42.2 lb/ft^3
Melting point	63.7	147		
Boiling point	760	1400		
Heat of fusion	63.7	147	14.6 cal/g	26.3 Btu/lb
Heat of vaporization	760	1400	496 cal/g	893 Btu/lb
Viscosity	70	158	5.15 millipoises	6.5 kinetic units
	400	752	2.58 millipoises	3.5 kinetic units
	800	1472	1.36 millipoises	2 kinetic units
Vapor pressure	342	648	1 mm	0.019 lb/in.^{2}
	696	1285	400 mm	7.75 lb/in.^{2}
Thermal conductivity	200	392	0.017 cal/(s)(cm^2)(cm)(°C)	26.0 Btu(h)(ft^2)(°F)
	400	752	0.09 cal/(s)(cm^2)(cm)(°C)	21.7 Btu/(h)(ft^2)(°F)
Heat capacity	200	392	0.19 cal/(g)(°C)	0.19 Btu/(lb)(°F)
	800	1472	0.19 cal/(g)(°C)	0.19 Btu/(lb)(°F)
Electrical resistivity	150	302	18.7 microhm-cm	
	300	572	28.2 microhm-cm	
Surface tension	100–150	212–302	About 80 dynes/cm	

Potassium deficiency may occur in several conditions, including malnutrition and excessive vomiting or diarrhea, and in patients undergoing dialysis; supplementation with potassium salts is sometimes required. *See* OSMOREGULATORY MECHANISMS.

Duarte Mota de Freitas

Bibliography. H. J. Adrogue and D. E. Wesson, *Potassium*, 1994; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; C. A. Pasternak (ed.), *Monovalent Cations in Biological Systems*, 1990.

Potato, Irish

A plant of the genus *Solanum* in the nightshade family, Solanaceae; it is related to tomatoes and peppers. There are more than 2000 species of *Solanum*, of which about 150 bear tubers. The potato of commerce, *S. tuberosum*, originated in South America, probably in the highlands of Peru and Bolivia, where it has been cultivated for several thousand years. Potatoes were introduced into Europe by Spanish explorers in the late sixteenth century and into the United States from Ireland in 1719. The crop became a staple in Europe, was a primary source of food in Ireland, and is known even today as the Irish potato. The potato blight famine of 1845 and 1846, caused by *Phytophthora infestans*, was responsible for the death of more than a million Irish people, and the emigration of about 1.5 million others. *See* SOLANALES.

The potato plant (**Fig. 1**) is an annual, herbaceous dicotyledon that is grown primarily for its edible tubers, which are short, thick underground stems that form on the ends of stolons (lateral stems). Lateral buds (eyes) on the mature tuber are the growing points for a new crop, provided that whole tubers or pieces with at least one eye are planted. Because it can reproduce vegetatively by means of its tubers, it is sometimes considered to be a perennial.

Varieties. There are 95–100 varieties certified for seed production in the United States and Canada. In the United States, seven varieties account for more than 70% of the commercial acreage planted. These are Russet Burbank, Norchip, Atlantic, Russet Norkotah, Superior, Centennial Russet, and Kennebec.

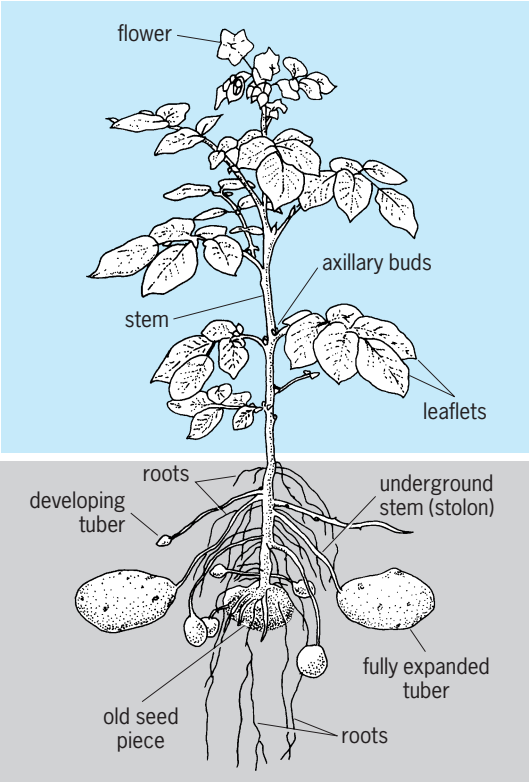

Fig. 1. Diagram of a potato plant. (*After Commercial Potato Production in North America, Potato Association of America, 1980*)

Production. Potatoes are grown in more countries than any crop except for corn (maize), and they are the fourth-most important food crop in supplying energy in the human diet, following rice, wheat, and corn. The leading potato producers are Russia, China, Poland, and the United States, in descending order. In the United States, potatoes are grown commercially in every state, but most production is in Idaho, Washington, Maine, Oregon, Colorado, Wisconsin, California, North Dakota, Minnesota, Michigan, Florida, New York, Pennsylvania, Texas, and New Mexico, in descending order.

Constituents. Chemical constituents can be affected by variety, production area, cultural practices, maturity at harvest, and storage conditions. The average composition of potatoes is 78–80% water, 14–18% starch, 2% protein, 1% minerals, 0.4% fiber, and 0.1% fat, with some sugars, organic acids, amino acids, and vitamins.

The potato is very nutritious, serving as a nearly complete food itself, and it produces more food per acre than any other crop. It ranks second in carbohydrate and protein production per unit area of land, behind sugarcane and soybean, respectively. The quality of potato protein is excellent. It has the best balance of the eight essential amino acids of any plant food. As a world average, 200 lb/acre (226 kg/hectare) of potato protein is produced, which is greater than that in wheat grain or in rice grain. Potatoes provide the world with more than 6.6×10^6 tons (6×10^6 metric tons) of protein. Starch constitutes about 65–80% of the dry weight of the potato and, calorically, is the most important component. The texture of a cooked potato is determined largely by its starch content.

Potatoes contain appreciable amounts of the B vitamins (niacin, thiamine, riboflavin, and pyridoxine), and they are an excellent source of vitamin D. One medium-sized baked potato yields 20 mg of vitamin C, about one-third of the recommended daily allowance; potatoes contribute more vitamin C to the United States food supply than any other single food source. Inorganic constituents or minerals of potatoes are predominantly potassium, phosphorus, sulfur, chlorine, magnesium, and calcium. Potatoes are the best nutritional source of potassium and contain sufficient quantities of iron to be a good nutritional source of this mineral as well. Since sodium content is low, potatoes are an ideal food for individuals on a low sodium diet. J. Creighton Miller, Jr.

Diseases. Diseases of potatoes are especially difficult to control for three reasons: potatoes are propagated vegetatively, using tubers for seed, and thus bacteria, viruses, fungi, and nematodes may easily be transmitted from generation to generation; clonal propagation results in genetically homogeneous populations which favor epidemic diseases; potato tubers are highly perishable and subject to storage diseases.

The most economically important group of diseases are those caused by viruses. Most common are mild, rugose, and latent mosaics, and leaf roll. Spindle tuber viroid is also widespread. Viruses reduce yield and sometimes lower the quality of the tubers. Because they are tuber-borne, viruses are controlled principally by roguing out diseased plants and growing seed crops in isolation. Some potato varieties are resistant or immune to certain viruses. *See* PLANT VIRUSES AND VIROIDS.

Several species of bacteria cause potato diseases. *Corynebacterium sepedonicum* (causing ring rot and wilt) is seed-borne and is controlled by the same methods used for viruses. *Pseudomonas solonacearum* (causing brown rot or southern wilt) is also seed-borne and in the southern United States survives in the soil over the winter, making potato culture unprofitable on fields that have become infested. *Erwinia carotovora* var. *atroseptica* (cause of blackleg) is also seed-borne and is similar to *E. carotovora* var. *carotovora*, which causes soft rot of tubers in storage. Proper storage of tubers and careful handling of seed pieces at planting reduce incidence of these diseases.

The most destructive fungus disease is late blight, caused by *Phytophthora infestans* (**Fig. 2**). Primary infection is from infected seed and typical epidemics occur when the weather is cool and wet. Genes for immunity have been bred into new varieties, but the genetic plasticity of the fungus has resulted in failure

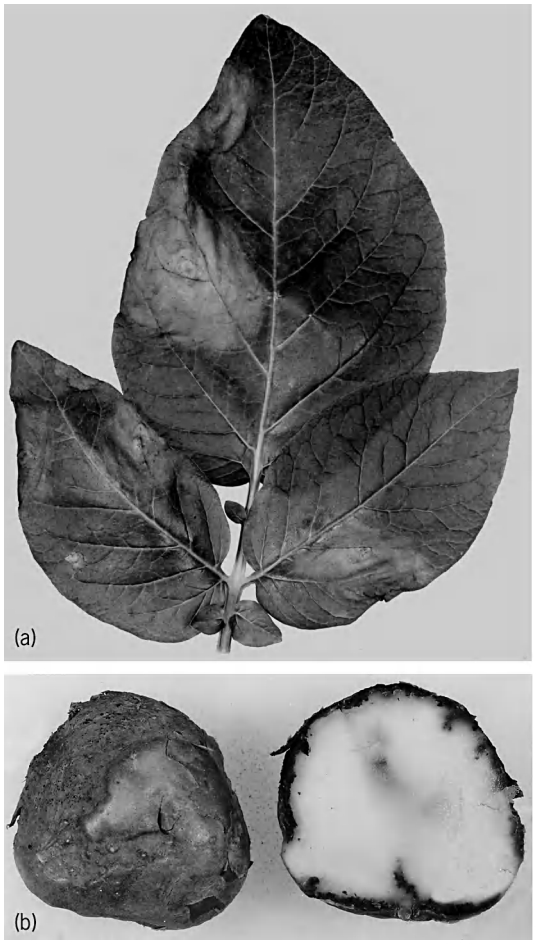

Fig. 2. Late blight (*Phytophthora infestans*) of Irish potato. (*a*) Foliar infection. (*b*) Tuber infection. (*USDA*)

of such resistance. There are, however, potato varieties with generalized but partial resistance which are useful in reducing the spread of epidemics. Fungicides are still the principal means of control. *See* FUNGISTAT AND FUNGICIDE.

Alternaria solani causes early blight, and various species of *Rhizoctonia, Verticillium,* and *Fusarium* cause stem cankers, wilts, and storage rots. Tubers are also subject to diseases that are principally blemishes, such as common scab, caused by *Streptomyces scabies*; silver scurf, caused by *Helminthosporium atrovirens*; and black scurf, caused by *Rhizoctonia*.

In addition to storage diseases caused by organisms, potato tubers are subject to nonparastic disorders such as black heart, hollow heart, and various kinds of of internal necroses and discolorations due to growing or storage conditions.

Nematodes cause root knots and tuber rots. The most destructive is the golden nematode (*Globodera rostochiensis*). At present it is not widespread, but once introduced into the soil, it persists indefinitely. Varieties resistant to *G. rostochiensis* and nematicides are used for control. The lesion nematode (*Pratylenchus penetrans*) and root-knot nematodes (*Meloidogyne* spp.) are serious in some potato-growing areas. *See* NEMATA (NEMATODA); PLANT PATHOLOGY. H. David Thurston

Bibliography. W. G. Burton, *The Potato,* 3d ed., 1989; G. Lisinska and W. Leszcynski (eds.), *Potato Science and Technology,* 1989; R. C. Rowe (ed.), *Potato Health Management*, 1993; J. A. Woolfe, *The Potato in the Human Diet,* 1987.

Potato, sweet

The fleshy root of the plant *Ipomoea batatas*. The sweet potato was mentioned as being grown in Virginia as early as 1648. In 1930 the selection of outstanding strains of the Porto Rico variety, which was introduced into Florida in 1908, was begun in Louisiana, and the best strain, Unit I Porto Rico, was released in 1934. The Unit I Porto Rico is now being replaced by the Centennial, a yam type which has three times its vitamin A content. More than 70% of the commercial crop in the United States is of the Centennial variety. Another variety, Nemagold, is used largely in the Virginia area because of its resistance to nematodes. Most varieties released now have multiple disease resistance, particularly resistance to wilt and soil rot.

In 1937 new techniques were developed for inducing the sweet potato to bloom and set seed. This stimulated a surge of research on breeding for higher yield, greater nutritional value, better shape, storage ability, market and canning quality, greater disease resistance, and new food products and industrial uses for the sweet potato throughout the Southern states and from New Jersey to California. Louisiana is the leading state for commercial production of both the canned and fresh products.

Fig. 1. Sweet potatoes. (*a*) Porto Rico, a yam variety. (*b*) Big-stem Jersey type. (*USDA*)

Types. There are two principal types of sweet potato, the kind erroneously called yam and the Jersey type (**Fig. 1**). The chief difference between the two is that in cooking or baking the yam, much of the starch is broken down into simple sugars (glucose and fructose) and an intermediate product, dextrin. This gives it a moist, syrupy consistency somewhat sweeter than that of the dry (Jersey) type. On cooking, the sugar in the dry type remains as sucrose.

The yam is produced largely in the Southern states; however, because of the breeding of more widely adapted varieties, it is now being grown farther north. The Jersey sweet potato is grown largely along

the eastern shore of Virginia. Maryland, Delaware, and New Jersey, and also in Iowa and Kansas. With the development of new varieties, the northern limits of sweet potato production have been extended to Canada and the northern United States. This has come about by breeding early and better varieties. In Michigan and Iowa a number of baby-food manufacturers are growing this crop for processing.

Julian C. Miller

Production practices. Technically, the sweet potato is a perennial plant, but in commerical production the growth is terminated by harvest due to impending change of seasons (cold or rainy) or by achievement of optimum storage root size. Plantings are established from sprouts obtained from stored roots or from vine cuttings. Production of storage roots does not require pollination, so that adverse conditions only delay or reduce harvests; crop failures are rare. A variety of soil types may be used; however, light sandy soils are preferred for ease of harvest. The harvested roots may be marketed at once or cured for a week at 85°F (29°C) and then stored for up to a year at 55°F (13°C).

Attributes. A number of factors contribute to the worldwide importance of sweet potatoes. Perhaps foremost is the high energy yield, which exceeds that of Irish potatoes and grains. Sweet potato roots are also high in nutritional quality (see **table**). For example, if the energy requirement of the average man between 23 and 50 years of age were met by sweet potato roots, 73% of the protein and 4160% of the vitamin A requirement would be met. The nutritional quality of the protein is high as measured by amino acid balance.

Utilization. The sweet potato is a versatile crop that is used as a food, an ingredient in other foods, an animal feed, and a feedstock for the production of starch and ethanol.

Food. Sweet potatoes can be baked, boiled, or strained and may be substituted in recipes for carrots, squash, or pumpkins to improve culinary and nutritional properties. They may be preserved by canning, freezing, or dehydration. Sweet potato flour produced from dried sweet potatoes can be substituted for 10–25% of wheat flour (depending on the product) with no loss of culinary quality and sometimes with a considerable increase in nutritional quality. Tender vine tips and leaves are consumed in many parts of the world as a pot herb, providing a good source of pro-vitamin A, calcium, iron, vitamin B_2, and vitamin C.

Nutritive value of sweet potato roots*

Nutrient	Value†
Protein	73
Vitamin A	4160
Vitamin C	825
Niacin	81
Riboflavin	91
Thiamin	167
Phosphorus	139
Iron	161
Calcium	95

*After *Composition of Foods*, Handb. 8, USDA, 1975; Recommended Daily Allowances, NAS/NRC, 1980.
† Expressed as the percentage ratio of nutrient supplied per unit energy supplied to nutrient required per unit energy required. Based on a male, age 23–50, weight 154 lb (70 kg), height 5 ft 6 in. (1.70 cm).

Feed. The high energy production and nutritional quality of the sweet potato also contribute to its importance as an animal feed, especially in the tropics, where many temperate-zone feed grains are poorly adapted. Dried roots and dried or fresh vines are appropriate feeds for dairy cows, beef cattle, and sheep. Dried roots are less suitable for swine and poultry and should not exceed 25% of the energy provided to them because of their low levels of trypsin inhibitors and poor digestion of starch. *See* ANIMAL FEEDS.

Industrial products. The energy content of sweet potato roots is largely due to starch, and the crop has been used as a feedstock for the production of starch and ethanol. The starch-extracting procedures are similar to those used for other root crops such as cassava or canna. The starch is of fairly good quality, with a amylose:amylopectin ratio between 20:80 and 30:70. It is suitable for food or for industrial products such as glues and sizing. *See* STARCH.

Subsistence agriculture. Sweet potatoes will generally produce a crop even under adverse conditions. In some cultures, sweet potatoes are considered a survival food—an important staple food in times of war or adverse weather. In the United States, land devoted to sweet potato production reached an all-time high of 1.06×10^6 acres (4.3×10^5 hectares) in 1932, more than 10 times the current acreage. In some countries, subsistence on sweet potatoes is associated with hard times, leading many people to avoid them as economic conditions improve. The most accepted utilization technologies in developing countries may eventually be in the areas of food ingredients, feeds and feeding systems, and improved efficiencies for industrial production.

Germ-plasm resources. Plant and human factors contribute greatly to a large germ-plasm base. Plants flower sparingly and usually require cross pollination, so progeny rarely resemble parents closely. Either chance seedlings or mutations, if superior to parents, can easily be preserved by vegetative reproduction, resulting in a wider gene pool. The germ plasm has not been adequately evaluated, but the work that has been done suggests considerable variability for most traits studied. The lack of importance of sweet potatoes in many temperate countries and the lack of status in tropical developing countries have contributed to low priorities for the collection and preservation of germ-plasm resources, a situation that has been addressed by the International Board for Plant Genetic Resources and the U.S. Department of Agriculture. Worldwide, over 6000 accessions of sweet potato and 500 of related species have been collected and are being maintained. *See* BREEDING (PLANT).

John C. Bouwkamp

Diseases. Sweet potato diseases are caused by fungi, bacteria, viruses, nematodes, and mycoplasmas, and by nutritional and environmental factors during plant growth and transit and storage of

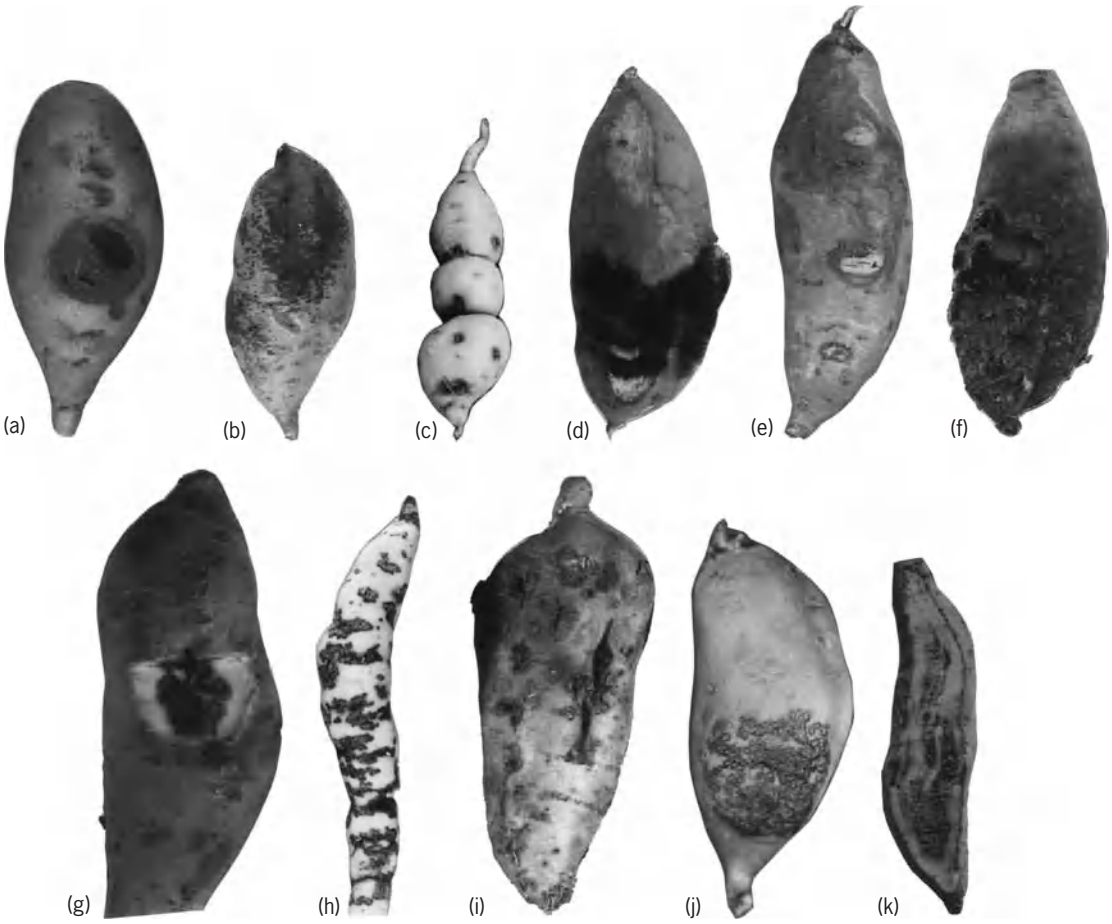

Fig. 2. Common diseases of sweet potato. (*a*) Black rot. (*b*) Scurf. (*c*) Soil pox. (*d*) Rhizopus soft rot. (*e*) Surface rot. (*f*) Java black rot. (*g*) Internal cork. (*h*) Russet crack. (*i*) Root knot. (*j*) Blister. (*k*) Internal breakdown.

edible roots (**Fig. 2**). Sweet potatoes are propagated vegetatively, and the biotic disease agents are transmitted from crop to crop by planting infected or contaminated plants. The fungi that infect roots in transit and storage gain entrance through mechanical wounds made in harvesting and packaging the potatoes. Major fungal diseases in the field are *Fusarium*-induced wilt, black rot, scurf, stem blight, and soil pox. Minor fungal diseases include mottle root necrosis, two leaf spots, and white rust. *Rhizopus*-induced soft rot and black rot are important transit diseases that also occur in storage. Other storage diseases are *Fusarium*-induced rot, charcoal rot, surface rot, Java black rot, and *Diaporthe*-induced rot. A bacterial wilt and soft rot occur in the United States and China. The viruses of internal cork, russet crack, and dwarf, and the mycoplasma of witches'-broom impair root quality and reduce plant growth and productivity. The importance of viruses such as those of mosaic, feathery mottle, and other diseases is uncertain. Root knot is the important nematode-induced disease. Deficient boron nutrition during plant growth stimulates blister in the stored roots. Exposing stored roots to temperatures below 50–55°F (10–13°C) induces internal breakdown.

Sweet potato diseases are controlled by using disease-free roots or plants; rotating other crops with sweet potatoes in infested land; chemically treating infested land, propagative roots, and storage containers and houses; curing freshly harvested roots; and planting disease-resistant varieties (cultivars). Field diseases are controlled by growing plants from disease-free roots treated with a fungicide. Sometimes soils infested with nematodes or the soil pox organism must be chemically treated to reduce losses. To control storage diseases, freshly harvested roots are placed in new or fungicide-treated containers and transferred immediately to a ventilated room at 86°F (30°C) and 90% relative humidity for 1 week to heal (cure) harvest wounds. Cured roots are stored at a minimum of 55°F (13°C). Cultivars with resistance or tolerance to *Fusarium*-induced wilt, internal cork, soil pox, and some root knot nematodes have been developed by breeding and selection. Some cultivars do not develop the destructive russet crack virus symptoms, and these are replacing those that develop symptoms. Sources of resistance to other disease are being sought in sweet potatoes from around the world to be used for developing other or improved disease-resistant cultivars. *See* PLANT PATHOLOGY; POTATO, IRISH.

L. W. Nielson

Bibliography. J. C. Bouwkamp (ed.), *Sweet Potato Products: A Natural Resource for the Tropics,* 1985; R. K. Jansson and K. V. Raman, *Sweet Potato Pest Management: A Global Perspective,* 1991; J. A. Woolfe, *Sweet Potato: An Untapped Food Resource*, 1992.

Potential flow

A fluid flow that is isentropic and that, if incompressible, can be mathematically described by Laplace's equation. For an ideal fluid, or a flow in which viscous effects are ignored, vorticity (defined as the curl of the velocity) cannot be produced, and any initial vorticity existing in the flow simply moves unchanged with the fluid. Ideal fluids, of course, do not exist since any actual fluid has some viscosity, and the effects of this viscosity will be important near a solid wall, in the region known as the boundary layer. Nevertheless, the study of potential flow is important in hydrodynamics, where the fluid is considered incompressible, and even in aerodynamics, where the fluid is considered compressible, as long as shock waves are not present. *See* BOUNDARY-LAYER FLOW; COMPRESSIBLE FLOW; ISENTROPIC FLOW.

In the absence of viscous effects, a flow starting from rest will be irrotational for all subsequent time, and according to Kelvin's theorem will be circulation-free. In the absence of circulation, solid bodies (such as a sphere or an airfoil) within the flow cannot experience any net forces. The paradox that an object in steady inviscid flow does not produce drag, whereas the same object in a real flow would do so, is attributed to Jean le Rond d'Alembert. *See* D'ALEMBERT'S PARADOX; KELVIN'S CIRCULATION THEOREM.

For an irrotational flow, the curl of the velocity is zero ($\nabla \times \boldsymbol{V} = 0$). The curl of the gradient of any scalar function is zero ($\nabla \times \nabla\phi = 0$). It then follows mathematically that the condition of irrotationality can be satisfied identically by choosing the scalar function, ϕ, such that $\boldsymbol{V} = \nabla\phi$. For this reason, this scalar function ϕ has been traditionally referred to as the velocity potential, and the flow as a potential flow. *See* CALCULUS OF VECTORS; POTENTIALS.

By applying the continuity equation to the definition of the potential function, it becomes possible to represent the flow by the well-known Laplace equation ($\nabla^2\phi = 0$), instead of the coupled system of the continuity and nonlinear Euler equations. The linearity of the Laplace equation, which also governs other important physical phenomena such as electricity and magnetism, makes it possible to use the principle of superposition to combine elementary solutions in solving more complex problems. In a potential flow problem, ϕ is calculated first by solving $\nabla^2\phi = 0$, subject to appropriate Dirichlet or Neumann boundary conditions. Then, $\boldsymbol{V}$ is determined by evaluating $\boldsymbol{V} = \nabla\phi$. Finally, the pressure field is determined by solving Bernoulli's equation. Many computer codes exist that can solve and demonstrate a general class of potential flow problems. A convenient way to visualize potential flows in two spatial dimensions is by drawing the orthogonal system of potential lines (lines of constant ϕ) versus streamlines (lines of constant stream function ψ). In two spatial dimensions, the potential function is related to the stream function by the Cauchy-Riemann conditions for a complex function $w = \phi + i\psi$. *See* BERNOULLI'S THEOREM; COMPLEX NUMBERS AND COMPLEX VARIABLES; EQUATION OF CONTINUITY; LAPLACE'S DIFFERENTIAL EQUATION; LAPLACE'S IRROTATIONAL MOTION; STREAM FUNCTION.

The numerical solution of potential flows, including those with sources and sinks, can be obtained easily by the use of standard finite-difference, finite-volume, and finite-element methods. Solutions can also be obtained by the use of general-purpose mathematical and symbolic software, as well as by spreadsheet software. Given the elliptical nature of the governing equations, the solution at one point in the flow field depends on the surrounding boundary values. Consequently, boundary integral methods can provide computationally efficient solutions, especially if the interest lies in only a few points within the domain. *See* DIFFERENTIAL EQUATION; FINITE ELEMENT METHOD; FLUID-FLOW PRINCIPLES; INTEGRAL EQUATION; MATHEMATICAL SOFTWARE; SYMBOLIC COMPUTING.

Peter E. Raad

Bibliography. G. K. Batchelor, *An Introduction to Fluid Dynamics*, Cambridge University Press, 1967, reprint 2000; I. G. Currie, *Fundamental Mechanics of Fluids*, 2d ed., McGraw-Hill, New York 1993; H. Lamb, *Hydrodynamics*, 6th ed., Cambridge University Press, 1932, reprint 1993; A. R. Paterson, *A First Course in Fluid Dynamics*, Cambridge University Press, 1983, reprint 1987; F. M. White, *Fluid Mechanics*, 4th ed., McGraw-Hill, Boston, 1999.

Potentials

Functions or sets of functions from whose first derivatives a vector can be formed. A vector is a quantity which has a magnitude and a direction, such as force.

When the potential is a single function (scalar potential), the vector is obtained by the gradient operation, as in Eq. (10). When the potential itself is a vector (vector potential), the vector is obtained by the curl operation, as in Eq. (21).

Potential theory. The mathematical theory of the potential is basically the study of differential equations (1) and (2), called Laplace's equation and Poisson's equation, respectively.

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = 0 \qquad (1)$$

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = -C\rho \qquad (2)$$

Here ϕ is a function defined in part or all of space, C is some constant, and ρ is a function characterizing the source of ϕ. The theory is best expressed in terms of vector calculus. The three vectors of unit magnitude which point in the positive directions of the x, y, and z axes of a right-handed cartesian coordinate system are denoted by $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$. The radius vector which points from the origin to an arbitrary point with coordinates x, y, z is then $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Suppose that to each point in a region of space (or all of space) there is

associated a number ϕ such that the function $\phi(\mathbf{r})$ has partial derivatives with respect to x, y, and z, of first and second order everywhere in that region; and suppose .further that the second derivatives satisfy $\partial^2\phi/\partial x\partial y = \partial^2\phi/\partial y\partial x$, and corresponding equations for x and z, and for y and z. Then $\phi(\mathbf{r})$ is called a scalar potential. The vector $\mathbf{V}$ given by Eq. (3) is defined by

$$\begin{aligned}\mathbf{V}(\mathbf{r}) &= \mathbf{i}V_x + \mathbf{j}V_y + \mathbf{k}V_z \\ &= -\mathbf{i}\frac{\partial\phi}{\partial x} - \mathbf{j}\frac{\partial\phi}{\partial y} - \mathbf{k}\frac{\partial\phi}{\partial z}\end{aligned} \tag{3}$$

this equation for every $\mathbf{r}$ in the region and is called the gradient vector field associated with this potential. Equation (3) is abbreviated as $\mathbf{V}(\mathbf{r}) = -\nabla\phi(\mathbf{r})$, in which ∇ is called the gradient operator. It is a vector and a differentiation operator. The minus sign is a matter of convention. The curl of $\mathbf{V}$, a vector denoted by $\nabla \times \mathbf{V}$ and defined by Eq. (4), satisfies $\nabla \times \mathbf{V} = 0$ if

$$\begin{aligned}\nabla \times \mathbf{V} = \mathbf{i}\left(\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}\right) + \mathbf{j}\left(\frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}\right) \\ + \mathbf{k}\left(\frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y}\right)\end{aligned} \tag{4}$$

$\mathbf{V}$ is defined by Eq. (3). Conversely, whenever a vector $\mathbf{V}$ satisfies $\nabla \times \mathbf{V} = 0$ in a certain region, there exists a scalar potential ϕ in that region such that $\mathbf{V} = -\nabla\phi$. Then $\mathbf{V}$ is called the (negative) potential gradient. If, however, the divergence of $\mathbf{V}$ vanishes, that is, if $\mathbf{V}$ satisfies Eq. (5), which can be written as

$$\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z} = 0 \tag{5}$$

$\nabla \cdot \mathbf{V} = 0$, then ϕ satisfies Laplace's equation, Eq. (1), that is, $\nabla^2\phi = 0$. In that case ϕ is called a harmonic function. *See* CALCULUS OF VECTORS.

If ϕ satisfies Laplace's equation $\nabla^2\phi = 0$, in a finite region of three-dimensional space enclosed by the closed surface S, then the solution ϕ inside S is uniquely determined by specifying either the values of ϕ on S or the values of the derivatives of ϕ normal to S, $\partial\phi/\partial n$, on S. In the former case, ϕ is given by Eq. (6), in the latter case by Eq. (7). Here the normal

$$\phi(\mathbf{r}) = -\frac{1}{4\pi}\int\int \phi(\mathbf{r}')\frac{\partial}{\partial n}\left(\frac{1}{R}\right) dS \tag{6}$$

$$\phi(\mathbf{r}) = \frac{1}{4\pi}\int\int \frac{\partial\phi(\mathbf{r}')}{\partial n}\frac{1}{R}\, dS \tag{7}$$

derivative is taken with the normal pointing out of the enclosed volume; the quantity R is the distance between the end point of $\mathbf{r}$ which is inside the surface, and the end point of $\mathbf{r}'$ which is on the surface,

$$R = \sqrt{(x-x')^2 + (y-y')^2 + (z-z')^2}$$

and the integration is extended over the whole closed surface S. The two cases are called the Dirichlet and Neumann boundary conditions, respectively. It is also possible to specify mixed boundary conditions such as ϕ on part of S and $\partial\phi/\partial n$ on the remaining parts of S.

If the vector field satisifes $\nabla \times \mathbf{V} = 0$ and $\nabla \cdot \mathbf{V} = C\rho$, where ρ is some function defined in the region and whose integral over the region exists, then ϕ satisfies Poisson's equation, $\nabla^2\phi = -C\rho$. A solution of this equation is given by Eq. (8). The integral ex-

$$\phi(\mathbf{r}) = \frac{C}{4\pi}\int\int\int \frac{\rho(\mathbf{r}')}{R}\, dx'\, dy'\, dz' \tag{8}$$

tends over the whole three-dimensional region and R is again the distance from $\mathbf{r}$ to $\mathbf{r}'$, $R = |\mathbf{r} - \mathbf{r}'|$. The most general solution of Poisson's equation is obtained by adding to the particular solution given by Eq. (8) the general solution of Laplace's equation.

Laplace's equation also has important applications in two dimensions for problems in a plane. *See* LAPLACE'S DIFFERENTIAL EQUATION; LAPLACE'S IRROTATIONAL MOTION.

Gravitational potential. According to Isaac Newton's theory of gravitation, every spherical body of mass M produces a force field $\mathbf{F}$. Any other body, of mass m, will experience an attractive force, directed toward M, and of the magnitude given in Eq. (9),

$$F = G\frac{Mm}{r^2} \tag{9}$$

where distance between the centers of gravity of the masses M and m is denoted by r; if M is at the origin, $\mathbf{r}$ is the position vector of m and r is its magnitude; G is Newton's constant of gravitation. Such a force is present at every point in the space surrounding the mass M at which one chooses to put the mass m. This gravitational force $\mathbf{F}$ can be expressed in terms of the gravitational field strength $\mathbf{f}$ defined by $\mathbf{f} = \mathbf{F}/m$, that is, the force per unit mass. The gravitational potential is that function ϕ whose negative gradient equals $\mathbf{f}$, $\mathbf{f} = -\nabla\phi$, or $\mathbf{f}_x = -\partial\phi/\partial x$, and so forth. This definition of ϕ is made unique by requiring ϕ to vanish as the distance from M goes to infinity. Then $\phi = -GM/r$ for the gravitational potential of the mass M, for all points outside the mass M. The gravitational force exerted by the mass M on the mass m can be written as Eqs. (10): the differentia-

$$\mathbf{F} = -m\nabla\phi \qquad \phi = \frac{-GM}{r} \tag{10}$$

tion is with respect to the position of the mass m; ϕ is negative because $\mathbf{F}$ is attractive. The quantity $V = m\phi$ is called the gravitational potential energy. The negative gradient of ϕ gives $\mathbf{f}$; the negative gradient of V gives $\mathbf{F}$. The potential energy of m in the field of M increases (becomes less negative) as the distance increases. *See* GRAVITATION.

If the body is not spherically symmetric, $\mathbf{F}$ in Eq. (9) and ϕ in Eqs. (10) will hold only at large distances. But for any distance and arbitrary shape, if the mass M is spread out over a certain volume with mass density $\rho(\mathbf{r})$, then ϕ satisfies Poisson's equation (2) in the form $\nabla^2\phi = +4\pi G\rho$. Its solution is given by Eq. (8) with $C = -4\pi G$.

When there are several masses present, M_1, $M_2, \ldots, M_n$, each of them will produce a force on a given mass m. The total force on m can be obtained by adding all the forces $\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_n$ by vector addition. However, it is obtained more easily by adding the potentials $\phi_1, \phi_2, \ldots, \phi_n$, produced by these masses and then finding the total force $\mathbf{F}$ from $\mathbf{F} = -m\Delta\phi$, where ϕ is the sum of the potentials. This is the principle of superposition. It is one reason why potentials are so useful.

Those sets of points at which ϕ has a fixed value form two-dimensional surfaces in space. These surfaces are called equipotential surfaces. For a spherically symmetric mass distribution these surfaces are concentric spheres whose centers are at the center of mass. Since the vector $\nabla\phi$ is parallel to the normal of the surface $\phi =$ constant, the lines of force are always perpendicular to the equipotential surfaces. *See* LINES OF FORCE.

Electrostatic potential. An electrically charged object with total electric charge Q will, when at rest, produce an electrostatic force field $\mathbf{F}$. If another charge q is placed at a distance r from it which is large compared to the size of the two charges, this force is given by Coulomb's law. Its magnitude is as given in Eq. (11), and it is repulsive or attractive de-

$$\mathbf{F} = \frac{1}{4\pi\epsilon}\frac{Qq}{r^2} \tag{11}$$

pending on whether the two charges Q and q have the same signs (both positive or both negative) or have opposite signs. The constant ϵ can be written as $K\epsilon_0$, where K is the relative permittivity (K equals 1 for vacuum) and ϵ_0 depends on the units. In SI units, $(4\pi\epsilon_0)^{-1} = c^2/10^7$, where c is the velocity of light. The analogy between Eqs. (9) and (11) is obvious; an important difference is that gravitational forces are always attractive. *See* COULOMB'S LAW.

If the charge Q is spherically symmetric, of radius a, and centered at the origin, the electrostatic potential ϕ at a distance $r > a$ from the origin is given by Eq. (12). This formula also holds when Q is not

$$\phi = \frac{1}{4\pi\epsilon}\frac{Q}{r} \tag{12}$$

spherically symmetric, but of characteristic size a, provided r is much larger than a. The electrostatic field strength is then $\mathbf{E} = -\nabla\phi$. It is the analog of $\mathbf{f}$ in the gravitational case. In general, the electrostatic force on a charge q is $\mathbf{F} = -q\nabla\phi$. In the special case when ϕ is given by Eq. (12), this force can be expressed by Eq. (13). The magnitude of this vector is

$$\mathbf{F} = -q\frac{\mathbf{r}}{r}\frac{d\phi}{dr} = \frac{1}{4\pi\epsilon}\frac{Qq}{r^3}\mathbf{r} \tag{13}$$

given by Eq. (11). *See* ELECTRIC FIELD.

If the charge Q is spread out over some volume with charge density $\rho(\mathbf{r})$ so that $\iiint \rho\, dx\, dy\, dz = Q$, then ϕ satisfies Poisson's equation (2) in the form $\nabla^2\phi = -\rho/\epsilon$. Its solution is given by Eq. (8) with $C = 1/\epsilon$. The Coulomb potential given by Eq. (12) is a special case of this result. If the charge is not distributed over a volume, but over a surface such as the surface of a sphere [surface charge density $\sigma(\mathbf{r})$], then the triple integral over $\rho dx' dy' dz'$ in Eq. (8) must be replaced by the double integral over σdS, where dS is the differential surface element.

As in the gravitational case, the total force acting on a charge q due to the presence of a set of charges $Q_1, \ldots, Q_n$ can be obtained by adding the potentials of these charges and finding $\mathbf{F} = -q\nabla\phi$; the principle of superposition holds. Equipotential surfaces are defined as in gravitation theory.

The electrostatic potential energy V of a charge q in a potential ϕ is given by the relation $V = q\phi$. *See* ELECTROSTATICS.

Magnetic scalar potential. If the magnetic induction vector B satisfies $\nabla \times \mathbf{B} = 0$, for example, in the absence of current densities and time-dependent electric fields, then $\mathbf{B}$ can be derived from a scalar potential ψ by the relation $\mathbf{B} = -\nabla\psi$. Maxwell's equations require $\nabla \cdot \mathbf{B} = 0$, so that the magnetic scalar potential must be a solution of Laplace's equation $\nabla^2\psi = 0$ which has the form of Eq. (1). *See* MAXWELL'S EQUATIONS.

Logarithmic potential. A straight electrically charged cylinder of circular cross section (typically a wire) and of effectively infinite length poses an electrostatic problem in two dimensions, in the plane perpendicular to the cylinder axis. If λ is the charge per unit length on the cylinder, then the electrostatic field at a distance r greater than the radius of the cylinder and measured from the cylinder axis is given by the relation $\mathbf{E} = \lambda/(2\pi\epsilon\mathbf{r})$. It is directed radially outward if λ is positive. The potential ϕ whose negative gradient gives this field $\mathbf{E}$ is called logarithmic potential and is given by Eq. (14). Here ln is the natural logarithm. The

$$\phi(r) = -\frac{\lambda}{2\pi\epsilon}\ln r + \text{constant} \tag{14}$$

additive constant can be so chosen that $\phi = 0$ for $r = a$, where a is the radius of the cylinder. Then Eq. (14) can be written as Eq. (15).

$$\phi(r) = -\frac{\lambda}{2\pi\epsilon}\ln\frac{r}{a} \tag{15}$$

Work and potential difference. When a charge q is moved in an electrostatic potential field, work is being done. This work W amounts to a loss of energy and is therefore negative. If q is moved from point 1 to point 2, W will be given by Eq. (16). The

$$W = -\int_1^2 \mathbf{F}\cdot d\mathbf{r} = q\int_1^2 \nabla\phi \cdot d\mathbf{r} = q(\phi_2 - \phi_1)$$
$$= V_2 - V_1 \tag{16}$$

work done equals the difference in potential energy between the final and initial positions. Thus, W does not depend on the path along which q is moved, but only on the initial and the final position. This is a direct consequence of the fact that $\mathbf{F}$ is a gradient vector field. It follows, in particular, that no work is being done if the charge q is moved in an arbitrary path which returns to the initial position, for example, a path which is closed. Mathematically this

means $\oint \mathbf{F} \cdot d\mathbf{r} = 0$ whenever $\mathbf{F}$ is a gradient field. Physical systems in which all forces can be derived from potentials are therefore called conservative systems. *See* WORK.

The potential difference between two points 1 and 2 in a potential field is the work per unit charge that needs to be done to move a charge from 1 to 2. The electrostatic potential at $\mathbf{r}$ given by Eq. (12), in which ϕ is chosen to vanish at infinite distance, is the work per unit charge necessary to move a positive charge from infinity to the position $\mathbf{r}$. In an electric network the potential difference between two points is the work per unit charge that must be done to move a positive charge through the network from one point to the other. It is usually measured in volts.

Neither potential nor potential energy has absolute physical meaning. They must be measured relative to some arbitrary reference potential or potential energy. For example, it can be chosen to be zero at infinite distance from a charged body; or in electric networks one point can be connected to the ground and defined to have zero potential. Only voltage difference is a physically meaningful quantity, not absolute voltage.

In gravitation theory the potential difference between a body located at a distance h above the Earth's surface and its location on the Earth's surface is, according to Eq. (10), expressed as $\phi(h) - \phi(0) = -GM/(R+h) + GM/R = GMh/R^2$, where the Earth's radius and mass are R and M, respectively, and h is much smaller than R. This potential difference is written gh where $g = GM/R^2$ is the gravitational acceleration on the Earth's surface. If the body has a mass m, the two locations differ in potential energy by mgh.

Vector potential. The electromagnetic vector potential $\mathbf{A}$ is a vector, that is, a set of three functions A_x, A_y, A_z, whose partial derivatives give the magnetic induction vector $\mathbf{B}$ through the relation $\mathbf{B} = \nabla \times \mathbf{A}$. The curl operation is defined in Eq. (4). This definition ensures that one of the Maxwell equations, $\nabla \cdot \mathbf{B} = 0$, is identically satisfied. $\mathbf{A}$ is not uniquely defined by $\nabla \times \mathbf{A}$. Since the curl of a gradient vanishes identically, a gradient added to $\mathbf{A}$ will not change $\mathbf{B}$. This nonuniqueness is called the gauge invariance of $\mathbf{B}$.

If the magnetic field strength $\mathbf{H}$ and the induction $\mathbf{B}$ are linked by a constant permeability μ such that $\mathbf{B} = \mu\mathbf{H}$, then the vector potential $\mathbf{A}$ satisfies Eq. (17),

$$\nabla^2 \mathbf{A} = -\mu \mathbf{j} \tag{17}$$

where $\mathbf{j}$ is the current density which produces the magnetic field. Each component of this vector equation is a Poisson equation of the type of Eq. (2). Its solution is given by Eq. (18), where R is the magnitude of the vector $\mathbf{r} - \mathbf{r}'$. If the current I does not have a density $\mathbf{j}$ but is a line current, Eq. (16) is replaced by Eq. (19), the integral being a line integral along the whole current loop. When this loop is small compared to R, Eq. (17) can be transformed by Stokes' integral theorem into the potential of a magnetic dipole, which is given by Eq. (20). Here $\mathbf{R}$ is the distance of the end point of $\mathbf{r}$ from the dipole, and $\mathbf{M}$ is the magnetic dipole moment. When $\mathbf{M}$ is produced by a current I forming a circle of radius a, $M = \pi a^2 I$. *See* STOKES' THEOREM.

$$\mathbf{A}(\mathbf{r}) = \frac{\mu}{4\pi} \int\int\int \frac{\mathbf{j}(\mathbf{r}')}{R}\, dx'\, dy'\, dz' \tag{18}$$

$$\mathbf{A}(\mathbf{r}) = \frac{\mu}{4\pi} \oint \frac{I d\mathbf{s}}{R} \tag{19}$$

$$\mathbf{A}(\mathbf{r}) = \frac{\mu}{4\pi} \frac{\mathbf{M} \times \mathbf{R}}{R^3} \tag{20}$$

Electromagnetic potentials. In the general time-dependent case, the electric field strength $\mathbf{E}$ and the magnetic induction $\mathbf{B}$ can be expressed in terms of the scalar potential ϕ and the vector potential $\mathbf{A}$ as in Eqs. (21). This ensures that the homogeneous Maxwell equations are identically satisfied. In free space, outside a source distribution of charge density ρ and current density $\mathbf{j}$, the potentials satisfy inhomogeneous wave equations (22). The constant ϵ_0 was defined above [after Eq. (11)] and μ_0 is in SI units $4\pi \times 10^{-7}$. When the potentials are time independent, Eqs. (22) reduce to the Poisson equations. *See* WAVE EQUATION.

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \qquad \mathbf{B} = \nabla \times \mathbf{A} \tag{21}$$

$$\begin{aligned} \nabla^2\phi - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} &= \frac{-\rho}{\epsilon_0} \\ \nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} &= -\mu_0\mathbf{j} \end{aligned} \tag{22}$$

The potentials ϕ and $\mathbf{A}$ are not uniquely determined by Eqs. (21) if the fields are known. One can always add a gradient $\nabla\Lambda$ to $\mathbf{A}$ and a term $-\partial\Lambda/c\partial t$ to ϕ without affecting $\mathbf{E}$ and $\mathbf{B}$. Such a transformation of the potentials is called a gauge transformation. The invariance of the fields under this transformation is gauge invariance. This freedom of ϕ and $\mathbf{A}$ can be restricted in a relativistically invariant way by the Lorentz condition on the potentials, Eq. (23).

$$\nabla \cdot \mathbf{A} + \frac{1}{c}\frac{\partial\phi}{\partial t} = 0 \tag{23}$$

In the special theory of relativity the four functions ϕ, A_x, A_y, A_z transform under Lorentz transformations like four components of a four-vector. They are sometimes collectively called four-potential. *See* RELATIVITY.

Retarded and advanced potentials. Electromagnetic fields propagate through empty space with a finite velocity c, the velocity of light, rather than with infinite velocity. Consequently, it takes a certain time for an electric or magnetic field produced by a charge or current at time t_Q and at position $\mathbf{r}_Q$ to reach another position r_P. This time is $|\mathbf{r}_P - \mathbf{r}_Q|/c$. The time t_Q is the retarded time of the fields which are observed at time $t_P = t_Q + |\mathbf{r}_P - \mathbf{r}_Q|/c$. The fields present at

t_P and produced at $t_Q < t_P$ are the retarded fields. The potentials from which they are derived are the retarded potentials. They are important in problems of electromagnetic radiation.

If time would run backward, the fields and potentials observed at time t_P at point $\mathbf{r}_P$ would have to be produced at some future time t_Q' and at a position $\mathbf{r}_Q'$ of the source such that $t_P = t_Q' - |r_Q' - r_P|/c$. The potentials at time t_P which are produced at the later time $t_Q' > t_P$ are the advanced potentials. While they have no direct physical meaning, they are important in the mathematical analysis of radiation theory.

Liénard-Wiechert potentials. The retarded and advanced electromagnetic scalar and vector potentials produced by a moving point charge can be expressed in terms of the (retarded or advanced) position and velocity of that charge. This was first done by A. Liénard in 1898 and E. Wiechert in 1900.

Polarization potentials. Two vectors $\mathbf{\Pi}_e$ and $\mathbf{\Pi}_m$ can be defined by $\mathbf{A} = (1/c)\,(\partial\mathbf{\Pi}_e/\partial t) + \nabla \times \mathbf{\Pi}_m$ and $\phi = -\nabla \cdot \mathbf{\Pi}_e$. The electromagnetic fields $\mathbf{E}$ and $\mathbf{B}$ can then be expressed in terms of $\mathbf{\Pi}_e$ and $\mathbf{\Pi}_m$. These are the polarization potentials. They are unusual because from them are derived not only a vector, $\mathbf{A}$, but also a scalar, ϕ. The Lorentz condition is identically fulfilled. The fields $\mathbf{E}$ and $\mathbf{B}$ are now given by second derivatives of $\mathbf{\Pi}_e$ and $\mathbf{\Pi}_m$, an exception to the usual definition of potentials, which involves only first derivatives. Heinrich Hertz in 1889 was the first to use such potentials. They are therefore often called Hertz potentials.

Debye potential. Radiation or scattering of electromagnetic waves by a distribution of localized sources in a homogeneous isotropic medium leads to fields $\mathbf{E}$ and $\mathbf{B}$ which can be expressed in terms of only two scalar potentials Π_e and Π_m. They were first introduced by Peter Debye in 1909 and are related to the polarization potentials.

Velocity potential. In the theory of fluid dynamics the absence of vortices is expressed by $\nabla \times \mathbf{v} = 0$, where $\mathbf{v}$ is the velocity vector which is a function of position inside the fluid. If that equation holds, there always exists a velocity potential, that is, a scalar function ϕ with the property $\mathbf{v} = -\nabla\phi$. One then has potential flow of that fluid. This potential satisfies Laplace's equation (1) if the fluid is incompressible and in steady motion, so that the continuity (conservation of mass) equation requires that $\nabla \cdot \mathbf{v} = 0$. *See* FLUID-FLOW PRINCIPLES.

Thermodynamic potential. In thermodynamics the Helmholtz function of free energy A (or F) and the Gibbs function G are both referred to as thermodynamic potentials. The reason for this name is the analogy to mechanical potential energy rather than to potential. When an isolated system is in equilibrium under various conditions, these quantities are a minimum. Thus thermodynamic equilibrium can be attained at constant temperature and volume, in which case A is a minimum; alternatively, it can be attained at constant temperature and pressure, in which case G is a minimum. *See* CHEMICAL THERMODYNAMICS; FREE ENERGY.

Chemical potential. A thermodynamic system in general consists of several chemical constituents. Let n_k be the number of moles of constituent k. Then the chemical potential of this constituent is the rate of change of the Gibbs function G with n_k, $\partial G/\partial n_k = \mu_k$. Even when constitutent k is not initially present, μ_k can be different from zero, so that its addition to the system will affect the Gibbs function. Chemical potentials play an essential role in phase transitions and chemical equilibrium. F. Rohrlich

Bibliography. J. Bliedtner, *Potential Theory*, 1986; M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; J. Kral et al. (eds.), *Potential Theory: Surveys and Problems*, 1988; F. Rohrlich, *Classical Charged Particles*, 1965, reprint 1989; W. R. Smythe, *Static and Dynamic Electricity*, 3d ed., 1968, reprint 1989; M. W. Zemansky and R. Dittman, *Heat and Thermodynamics*, 7th ed., 1996.

Potentiometer

An instrument that precisely measures an electromotive force (emf) by balancing it against a known potential drop established by a three-terminal resistive divider (**Fig. 1**), across which a known voltage exists. A divider can adjust and control the voltage applied to some device or part of a circuit, for example, a volume control. The ratio of output to input voltage is equal to that of the tapped to total divider resistance when the current required by the controlled device is negligible. Potentiometers as voltage-measuring instruments have largely been replaced by the ratio instrument known as a digital voltmeter. The Weston standard cells which supplied a known standard voltage have also been superseded by solid-state voltage reference devices which are usually incorporated in the digital voltmeter.

Another type of potentiometer (**Fig. 2**) involves a current comparator which senses and corrects inequality of ampere-turns in two windings on a magnetic core. Two matched toroidal cores, wound with an identical number of turns, are excited by a fixed-frequency oscillator. The fluxes induced in the cores are equal and oppositely directed so that they cancel with respect to a winding which links both cores. In the absence of any additional magnetomotive force (mmf), there is no signal in the detector winding which links both cores. If, in another winding A linking both cores, a direct current is injected, its mmf adds to the flux in one core and subtracts from the other at any instant. The resulting net flux in the detector winding induces a voltage in it. This signal is used to control direct current in another wind-

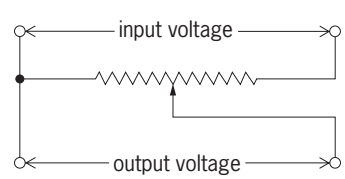

Fig. 1. Three-terminal resistive voltage divider.

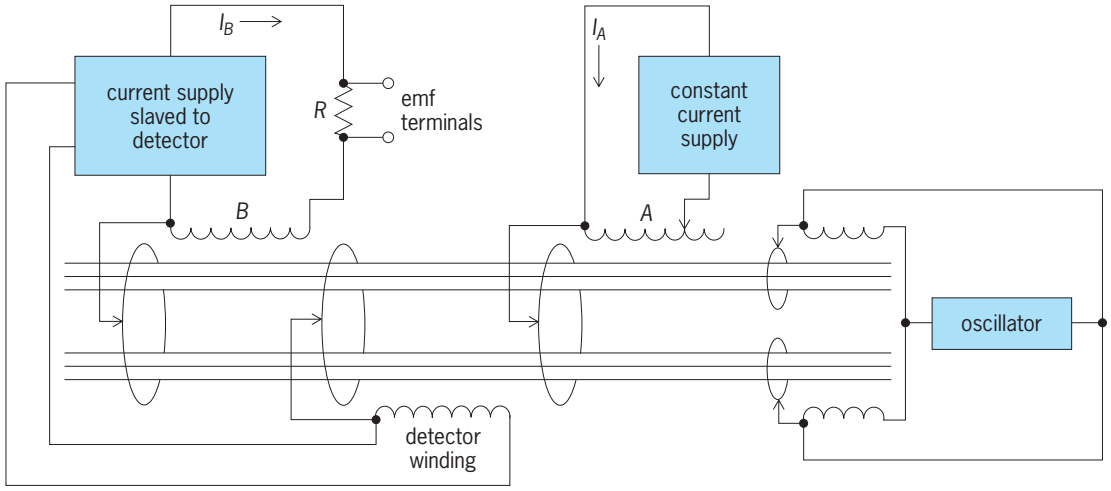

Fig. 2. Circuit diagram of a current-comparator potentiometer, a type of constant-resistance potentiometer.

ing B that also links both cores. Its mmf is opposed to that of A and when they are equal, there is no signal in the detector winding. Thus a constant current in an adjustable number of turns in winding A is matched to a variable current in a fixed number of turns in winding B; that is, their ampere-turns are equal and opposed. The voltage drop I_BR is used to oppose the emf to be measured. The potentiometer is made direct-reading in volts in terms of the turns-ratio B/A, by adjusting the constant-current source, using a voltage standard (omitted to simplify the figure). An advantage of this potentiometer over those whose long-term accuracy depends on the stability of a resistance ratio is that the ratio here is the turns-ratio of windings on a common magnetic core, dependent solely on conductor position and therefore not subject to drift with time. *See* CURRENT COMPARATOR; ELECTROMAGNETIC INDUCTION; MAGNETISM.

Volt box. Volt boxes, used to extend the range of potentiometers (in some instances to 1500 V), are resistive voltage dividers having a limited number of fixed-position taps (**Fig. 4**). Depending on its anticipated value, the unknown voltage is connected across appropriate input terminals (say 0–150 V), and the potentiometer is connected to the 0–1.5 V output terminals. The potentiometer is balanced in the usual way, and its reading is multiplied by the factor corresponding to the tap point used. While no current is taken by the potentiometer at balance, current is supplied to the volt box by the source, and it is the voltage drop in the divider resistance which is measured, not the open-circuit emf of the source. For a given voltage, power consumption and internal heating are greater in a low- than in a high-resistance circuit, with consequent greater change in resistance ratio. Thus a higher ohms-per-volt design is desirable. However, the quality of insulation becomes more critical in high-resistance designs, since the insulating structures that support the resistance elements are themselves high-resistance leakage paths. As a compromise, the resistance of working volt boxes ranges between 200 and 750 ohms/V. High-accuracy standard volt boxes (1000 ohms/V) have a guard circuit which maintains shields at appropriate potentials along the insulating structure.

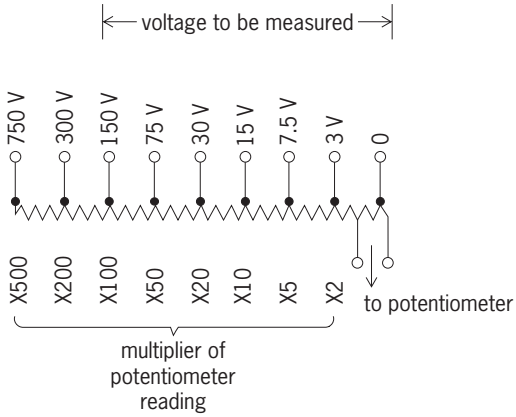

Fig. 3. Configuration of fixed-position taps in a volt box.

Current measurements. Potentiometer techniques may also be used for current measurement, the unknown current being sent through a known resistance and the IR drop opposed by balancing it at the voltage terminals of the potentiometer. Here, of course, internal heating and consequent resistance change of the current-carrying resistor (shunt) may be a critical factor in measurement accuracy; and the shunt design may require attention to dissipation of heat resulting from its I^2R power consumption. *See* CURRENT MEASUREMENT; JOULE'S LAW.

Alternating-voltage measurements. Potentiometer techniques have been extended to alternating-voltage measurements, but generally at a reduced accuracy level (usually 0.1% or so). Current is set on an ammeter which must have the same response on ac as on dc, where it may be calibrated with a potentiometer and shunt combination. Balance in opposing an unknown voltage is achieved in one of two ways: (1) a slide-wire and phase-adjustable supply; (2) separate in-phase and quadrature adjustments

on slide wires supplied from sources that have a 90° phase difference. Such potentiometers have limited use in magnetic testing. *See* ALTERNATING CURRENT; ELECTRICAL MEASUREMENTS; VOLTAGE MEASUREMENT.

Voltage divider. A three-terminal resistive voltage divider (sometimes called a potentiometer) is shown schematically in **Fig. 5**. Two terminals are the ends of the resistor (or a series combination of resistors) and the third terminal is an intermediate movable connection. This divider is used to adjust and control the voltage applied to some device or part of a circuit, for example, the volume control on a radio. The ratio of output to input voltage is equal to the tapped to total divider resistance when the current required by the output device is negligible.

Forest K. Harris; Ronald F. Dziuba; Bryan P. Kibble

Bibliography. C. F. Coombs, *Electronic Instrument Handbook*, 3d ed., 1999; B. P. Kibble et al., *A Guide To Measuring Direct and Alternating Current and Voltage Below 1 MHz*, Institute of Measurement and Control, U.K., 2003; M. V. Reissland, *Electrical Measurements*, 1989; L. Schnell (ed.), *Technology of Electrical Measurements*, 1993; L. M. Thompson, *Electrical Measurements and Calibration*, 2d ed., 1994.

Pottery

Vessels made entirely or partly of clay, and fired to a strong, hard product; occasionally, the term refers to just the lower grades of such ware. Alternatively, it refers to the manufacturing plant at which such ware is made. An older meaning is the art of making such ware; in this use it becomes synonymous with the older definition of ceramics. Pottery may be glazed or unglazed. *See* CERAMICS; GLAZING.

Grades of pottery are distinguished by their color, strength, absorption (the weight of water soaked up when the piece is submerged, expressed as a percentage of the original weight), and translucency (ability to transmit light). All these properties refer to the material or "body" under any glaze present. *See* PORCELAIN.

China is white in color, strong, has less than 2% absorption, is always glazed, and is translucent in thin ware. Special types are bone china, containing phosphates from calcined bones as a fluxing material; hotel china, made extra thick for maximum strength and therefore not translucent; frit china (also called frit porcelain), containing ground glass to give a translucent body maturing at a moderate firing temperature.

Stoneware has a characteristic cream or brown color, high strength, 0–5% absorption, and no translucency; it is frequently unglazed.

Earthenware (sometimes known as semivitreous china) is white or ivory, has less strength than china or porcelain, 3–10% absorption, and no translucency; it is usually glazed. Most everyday tableware is of this type. Faience is a special type with a soft, porous, red or yellow body covered by an opaque glaze. Majolica ware is a type having over 15% absorption and an opaque glaze over a relatively weak red or gray body.

Other special types of pottery are Parian ware, a body with a high flux content, usually unglazed, which fires to a smooth, marblelike finish, and terra cotta, a yellow, red, or brown earthenware with no glaze, used for art sculpture and similar purposes.

The firing of glazed ware is usually done in two steps; first the unglazed body is fired to give it strength, and then after the glaze is applied, it is refired at a lower temperature (except porcelains, when the second firing is at a higher temperature). *See* KILN.

Absorption is due to the presence of open pores or voids in the fired material into which water can penetrate; in general, the higher the firing temperature, the lower the absorption. Body color is determined mainly by raw-material purity. Strength depends on the porosity and also on the amount and type of glass and crystals developed in the body on firing. Translucency is obtained in products in which there is low porosity and little difference in index of refraction between the glass and crystals in the body. *See* CLAY.

J. F. McMahon

Bibliography. J. Colbeck, *Pottery Materials: Their Composition, Preparation, and Use*, 1988; J. Colbeck, *The Technique of Throwing*, 1988; E. Cooper and D. Royle, *Glazes for the Studio*, 1986; P. Rado, *Introduction to the Technology of Pottery*, 2d ed., 1988; C. F. Speight, *Hands in Clay: An Introduction to Ceramics*, 4th ed., 1999.

Pottiales

An order of the true mosses (subclass Bryidae), consisting of three families and about 91 genera. The order is characterized by short, papillose leaf cells and deeply divided, commonly twisted peristome teeth (see **illus.**). Members of the Pottiales typically

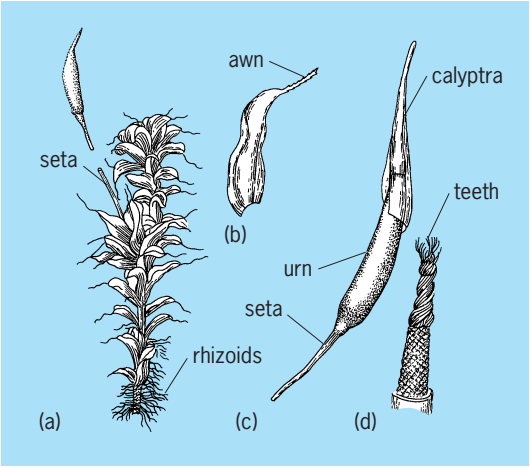

***Tortula ruralis*. (*a*) Entire plant. (*b*) Leaf. (*c*) Urn with calyptra. (*d*) Peristome. (*After W. H. Welch, Mosses of Indiana, Indiana Department of Conservation, 1957*)**

grow erect, with stems simple or forked by subfloral innovation, and they often produce gemmae (asexual reproductive bodies) on leaves, stems, or rhizoids. The leaves have revolute margins, and are arranged in many rows. The single costa often ends in a hyaline hairpoint. The upper cells are short and thick-walled and often roughened with numerous C-shaped papillae. The alar cells are not differentiated, although the basal cells sometimes extend up the margins as a border. The cylindric capsules are generally exserted and usually erect. An operculum is usually differentiated, and the peristome, usually present, consists of 16 teeth which are often deeply divided into two filiform segments and papillose-roughened. The calyptra is usually cucullate. The chromosome numbers are generally 12, 13, or multiples thereof. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA. Howard Crum

Poultry production

Poultry production comprises two major categories, meat production and egg production. Most poultry produced in North America is grown under close control on highly specialized farms. The evolution from small flocks to large commercial units after World War II was facilitated by rapid advances in the knowledge of nutrition, breeding, housing, disease control, and processing of poultry and eggs, and by improvements in transportation and refrigeration which made possible distant marketing of fresh products.

Incubation. Artificial incubation was a major advance in poultry production because it became possible to hatch large numbers of chicks of the same age for farmers to raise for meat or egg production. Poultry eggs can be successfully incubated in fan-ventilated incubators by maintaining a temperature of 99.5–100°F (37.5–37.8°C) and 55% relative humidity. Eggs must be turned during the first two-thirds of the incubation period to prevent adherence of the embryo to the embryonic membranes. The duration of incubation, under these standard conditions, for several types of poultry are: chicken, 21 days; turkey, guinea fowl, peafowl, and most ducks, 28 days; most geese, 30 days; pheasants, Chukar partridge, and Bobwhite quail, 24 days; Japanese quail, 17–18 days; and Muscovy ducks and Egyptian geese, 35 days.

Modern incubators are constructed of materials that can be effectively cleaned and disinfected and that provide good insulation of the chamber. Eggs are set in specially designed plastic flats which fit into channels in egg racks that move on wheels. The egg racks are equipped with mechanical systems to tilt the eggs 45°; turning is usually done hourly. Eggs are transferred from the setting trays into hatching trays 3 days before expected hatch. Hatching trays are held in specially designed hatching units much like setter cabinets. After the hatch is completed, the chicks are transferred to a conveyer belt for processing or directly into plastic boxes with absorbent paper pads or into disposable paper boxes with wood fiber pads. Chick servicing often involves sexing, vaccination, and beak trimming. A machine is available for injection of eggs at transfer; vaccines for protection of chicks against Marek's disease can be administered to the 18-day-old-embryos with improved early immunity. Usually, chicks are held at the hatchery prior to shipment to farms, starting early on the day after hatching. Specially designed delivery trucks or buses are used to provide adequate ventilation for chicks during shipment.

Hatcheries must be carefully designed for proper sanitation, and adequate air supply to incubators and hatchers. Ventilation air must be cooled in very hot weather and warmed in cold weather. Most hatcheries add moisture to air with humidifiers. Hatchery ventilation systems are designed to prevent movement of contaminated air from the hatching room, chick processing area, or equipment washing area into the clean egg storage or setting rooms.

Breeding. The genetic stock used for modern poultry production is produced by highly specialized breeding companies. Meat poultry is selected for good meat type, fast growth, disease resistance, and efficient conversion of feed to meat. Different strains of chickens are used for table egg production. These are selected for high egg production, large egg size, and small body weight for better conversion of feed to eggs and good livability. The body weights of meat and egg production strains are dramatically different (**Fig. 1**). The slow growth and low body weight of the commercial Leghorn demonstrate why these strains have low meat value.

Dual-purpose breeds or strains of poultry are no longer used extensively for commercial production. Instead, producers of specialty poultry and eggs (for example, free-range poultry) may still use traditional breeds such as New Hampshire Red or Barred Plymouth Rock which were used in the past for both meat and egg production.

Parent stock. Commercial poultry breeders usually sell parent stock chicks to hatcheries or integrated poultry companies. In most cases the commercial product consists of female and male parent lines which are crosses. The customer must continue to buy parent stock from the breeding company because the second generation from these stocks would not produce the same desirable product.

The parent stocks used to produce commercial chicks for meat or egg production require special

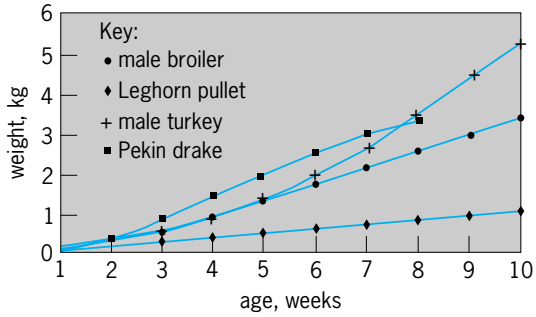

Fig. 1. Comparison of growth of commercial poultry.

care and management. Chicken parent stock is usually kept on litter floors or pens with part litter and part slatted floors. The litter areas are necessary for effective mating. Meat-type chickens must have feed restricted to control body weight, or they rapidly become overweight and mating, fertility, and egg production decline rapidly. Feed restriction is usually started by 3 weeks of age and continues throughout life. Males and females are raised separately for more effective weight control. At sexual maturity, males and females are housed together, but they feed from separate feeders. This system allows the manager to control the body weight of both sexes.

Turkey parent lines are also kept on litter because of their heavy body weight, but all mating is performed by artificial insemination. This procedure is more expensive, but high-quality males can be mated with larger numbers of hens, which is a distinct advantage. Also, the change to artificial mating has allowed geneticists to make more rapid progress in body weight gain and breast yield. Turkey semen cannot be successfully preserved by freezing, but short-term storage is possible. Therefore, toms may be raised on separate stud farms, where semen is harvested, processed, extended, and transported to the location of the hens for insemination. The duration of fertility after one mating in poultry can exceed 3 weeks, but commercial flocks are often inseminated weekly for high fertility.

Controlled lighting programs. Controlled lighting programs are necessary for breeding poultry. They start with short days during the growing phase to prevent the onset of egg production before females reach an appropriate body size. At the correct age for the genetic stock, the daylength is increased to provide a photostimulatory photoperiod (often 14–16 h of light per day). Males must be exposed to photostimulatory daylength at least a week prior to the females because their response to light is slower.

Poultry have a diurnal photosensitive period that occurs 11–16 h after dawn. A significant period of light must be provided within this period to stimulate a reproductive response. Light penetrates into the brain, where it stimulates cells in the hypothalamus to produce gonadothophin-releasing hormone. The latter stimulates the pituitary gland to produce follicle-stimulating hormone and luteinizing hormone, which initiate and maintain reproductive function. *See* BREEDING (ANIMAL).

Fig. 2. Commercial broiler house showing side curtain and ridge vents.

Brooding and rearing. Day-old chicks require an ambient temperature of 85–87°F (29–30°C) for normal growth and health during the first week of life. As the chicks grow and feather, they can tolerate lower temperatures. Brooding heat is often provided by a radiant gas brooder stove. The temperature is typically adjusted to 95°F (35°C) at the stove edge during the first week, and the chicks are allowed to find their comfort zone. The stove temperature is often decreased 5°F (3°C) per week until the prevailing house temperature is reached. This system provides an easy method of managing the environment for young chicks, with a low risk of overheating or chilling the birds.

If a controlled-environment house is available, the entire brooding area can be heated to the comfort zone, but it is important to provide uniform conditions to prevent piling or chilling mortality. If houses are well insulated, this system reduces the amount of energy required for brooding; with increasing energy costs and improved ventilation controls available, warm room brooding has increased on commercial farms.

Most chicks are started on floors that are covered with 2–4 in. (5–10 cm) of a litter material such as pine shavings, rice hulls, or peanut hulls. Litter materials should be dry and free from harmful chemicals or microbes. If meat flocks are healthy, the litter is often reconditioned by removal of cake, top dressed with new litter, and reused. Complete litter removal may occur annually, depending on flock health and economic conditions.

Feeding is usually done is small troughs or on plastic trays until the chicks learn to eat and drink. Chicks are quickly trained to eat from mechanical feeders and drink from closed water delivery systems to reduce labor. Watering systems have traditionally provided water in small troughs or cups, but nipple drinkers are also popular because they provide little opportunity for bacterial growth. Modern brooding houses have the feeders and waterers suspended from above with a cable-and-winch system. This system allows the height to be adjusted frequently for bird comfort and improved sanitation.

Feeding and nutrition. Poultry diets consist of common grains and protein sources with mineral and vitamin supplements. Animal or vegetable fats may be added to increase energy and reduce dustiness. Corn, grain sorghum, wheat, oats, and barley are often used for poultry feeding in the United States. Soybean meal is widely used as a protein supplement. Other important protein supplements are meat meal, fish meal, safflower meal, feather meal, and canola meal. By-product ingredients (for example, corn gluten meal or brewers grains) are also used in poultry diets when warranted by availability and price. Rations

for poultry are typically formulated by computer programs which compare ingredient costs with dietary requirements to provide least-cost formulas.

Ingredient quality is important in realizing optimum flock performance. Some feeds such as soybeans contain antinutritional factors which must be destroyed by heat treatment before they are useful for poultry feed. The ingredients must be free from mold toxins, harmful bacteria, and chemical contaminants. Moreover, quality control at the feed mill is extremely important to assure optimum flock performance and freedom of poultry and eggs from undesirable residues. *See* ANIMAL FEEDS.

Housing. The purpose of a poultry house is to confine the birds; to protect them from predators and environmental extremes which would cause mortality or reduce growth, feed efficiency, immunocompetence, fertility or egg production; to facilitate light control; and to facilitate bird management (**Fig. 2**).

Poultry houses can be constructed from locally available building materials. Smooth interior surfaces are preferred for effective sanitation. Houses are usually a maximum of 40 ft (12 m) wide to facilitate more uniform ventilation. House length is approximately 500 ft (152 m); most houses are constructed with a gable roof (**Fig. 3**).

In mild climates, poultry houses with open sides covered with poultry netting are often used. The side openings are usually covered with a curtain which can be raised or lowered to control the internal environment. If ambient temperatures rise above 95°F (35°C) or fall below 32°F (0°C), the roof of the house is insulated. When temperatures over 100°F (38°C) are expected, the house should also be equipped with a fogging system and fans (or both) to move air over the birds during high-temperature periods.

Mechanically ventilated houses allow more complete control of the environment within. For example, a well-constructed house in a mild climate might be able to maintain temperatures of 65–80°F (18–27°C), which reduce stress and provide efficient conversion of feed to meat or eggs. In hot climates, incoming air can be cooled with evaporative pad coolers or fogging systems.

Production systems. Chickens for table egg production are often housed in cages to provide cleaner eggs and protect the birds from disease agents which are recycled to birds from the manure (**Fig. 4**). Cages also allow more birds to be placed within a house because the cages can be arranged in tiers. Increasing the bird density in the house increases the overall bird heat which is available, and houses can be kept warmer in cold climates. Laying cages also allow eggs to be collected by belt conveyors.

Meat chickens and turkeys are usually grown in litter-floor houses because heavier poultry experience more lameness, breast blisters, and weaker bones and joints when grown in cages. Litter floors require careful management and ventilation to prevent problems with wet or caked litter. Meat chickens are typically grown in the same house from 1-day-old until market (6–9 weeks of age). Meat turkeys are typically kept in a brooder building until about 6 weeks of age and then transferred to a growing house until marketing (12–25 weeks of age).

Fig. 3. Interior of commercial broiler house showing pan feeder, bell-type waterers, and radiant gas brooders.

Ducks and geese. With some modification of husbandry, ducks and geese can be successfully raised in confinement. They do not require water for swimming and can be grown in litter-floor houses similar to meat chickens or turkeys. Young ducklings are sometimes started on slatted or raised wire floors because of the wetness of the droppings. Pekin and Muscovy ducks are the most common breeds used for meat production, but Indian Runner and Khaki Campbell breeds are used for both eggs and meat.

Embden geese are most often used for meat production. The White Chinese breed is often used for weeding in commercial crops or backyard gardens. Toulouse geese are common on small farms and are sometimes used for commercial meat production.

Health maintenance. The production of commercial poultry in large flocks requires well-designed disease-control programs. The first requirement is

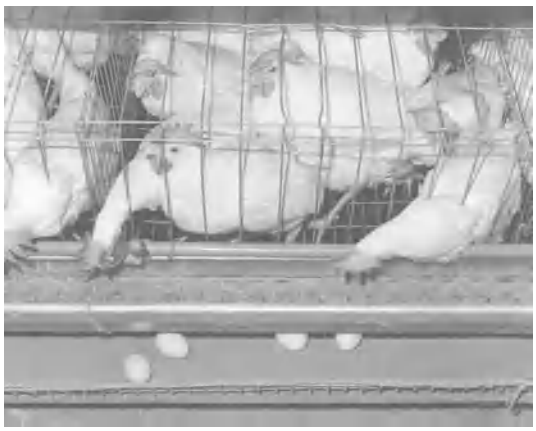

Fig. 4. Multiple-hen laying cage showing feeder, water cup, and egg collection belt.

for maintenance of biosecurity in production units. This means that entrance of contaminated workers and visitors, birds, feed, and equipment must be prevented. Equipment must be cleaned and disinfected before it is allowed into a production unit. Farms are operated on an all-in all-out basis with only one age of poultry at any given time. This prevents the vertical transmission of disease between age groups. Houses and equipment are often cleaned and disinfected between flocks, and always if a serious disease problem has been experienced.

Some poultry diseases are more effectively or economically controlled by vaccination. Examples are Marek's disease, Newcastle disease, infectious bronchitis, avian pox, infectious bursal disease, and often several others, depending on the disease history of the farm and the area where the poultry are raised. Some virus vaccines (for example, Newcastle, infectious bronchitis) can be administered by spray or in drinking water, while others (Marek's disease) must be injected. *See* NEWCASTLE DISEASE.

Processing and marketing. Shell eggs are often processed in plants located on the farm. With successful mechanization of egg collection, the eggs are often carried from the house directly into the processing machinery on a bar conveyor (in-line processing). This reduces processing labor and egg handling and breakage. By this procedure, eggs must be processed daily. Off-line processing provides more flexible processing hours, and eggs can be processed at a remote site; this system also allows one-age farms to be used for production.

A mechanical system can detect and separate cracked eggs. Egg processing machines can process up to 300 cases of eggs (360 eggs per case) per hour. Machines are also available to separate the yolk and albumen from the egg shell. Liquid yolk, albumen, or whole egg is pasteurized and sold for production of many food products. It can also be dried for use in cake mixes and similar products. Hard-cooked eggs are produced in mechanized plants and sold to restaurants and institutional food establishments. Poultry are usually processed at large central plants under inspection by the U.S. Department of Agriculture. The meat birds are loaded on trucks at farms during the night, and processing typically begins at midnight, continues for two shifts, and is followed by extensive cleaning and disinfection of the plant. Poultry are removed from transport racks or coops and hung on shackles. After stunning, bleeding, scalding, and feather removal, the carcasses must be transferred to a second line or table in a different room for evisceration, chilling, cutting, and packaging. Many poultry processing plants cut up poultry before packaging, and they may also separate meat from bone or skin. Special cuts include nuggets, breast strips, and cuts for fried chicken outlets. Meat may be mechanically removed from bone of less desirable cuts (for example, backs) to be used in production of hot dogs, sausages, and other products. *See* AGRICULTURAL SCIENCE (ANIMAL); EGG (FOWL).

Ralph A. Ernst

Bibliography. B. W. Calnek, (ed.), *Diseases of Poultry*, 10th ed., 1997; R. D. Crawford, ed., *Poultry Breeding and Genetics*, 1990; S. Leeson and J. D. Summers, *Commercial Poultry Nutrition*, 2d ed., 1997; L. S. Mercia, *Raising Poultry the Modern Way*, 1993; M. O. North and D. D. Bell, *Commercial Chicken Production Manual*, 4th ed., 1990; M. L. Scott and W. F. Dean, *Nutrition and Management of Ducks*, 1991.

Powder metallurgy

A metalworking process used to fabricate parts of simple or complex shape from a wide variety of metals and alloys in the form of powders. The process involves shaping of the powder and subsequent bonding of its individual particles by heating or mechanical working. Powder metallurgy is a highly flexible and automated process that is environmentally friendly, with a low relative energy consumption and a high level of materials utilization. Thus it is possible to fabricate high-quality parts to close tolerance at low cost. Powder metallurgy processing encompasses an extensive range of ferrous and nonferrous alloy powders, ceramic powders, and mixes of metallic and ceramic powders (composite powders). *See* METALLURGY.

Metal powders. Regardless of the processing route, all powder metallurgy methods of part fabrication start with the raw material in the form of a powder. A powder is a finely divided solid, smaller than about 1 mm (0.04 in.) in its maximum dimension. There are four major methods used to produce metal powders, involving mechanical comminution, chemical reactions, electrolytic deposition, and liquid-metal atomization. The method selected is dictated primarily by composition, intended application, and cost. Typically, metal powders for commercial usage range from 1 to 1200 micrometers. Depending on the method of production, metal powders exhibit a diversity of shapes ranging from spherical to acicular. Commercial atomized powders are shown in **Fig. 1**. Particle shape is an important property, since it influences the surface area of the powder, its permeability and flow, and its density after compaction. Chemical composition and purity also affect the compaction behavior of powders. For most applications, powder purity is higher than 99.5%.

Processes. Powder metallurgy processes include pressing and sintering, powder injection molding, and full-density processing.

Pressing and sintering. The basic powder metallurgy process is two-step involving powder compaction followed by sintering. First the rigid steel die (mold) is filled with powder. Pressure is then applied uniaxially at room temperature via steel punches located above and below the powder, after which the compact is ejected from the mold (**Fig. 2**). Commercial compaction pressures are normally in the range from 140 to 900 megapascals ($2–13 \times 10^4$ lb/in.^{2}). The powder compact is porous. Its density depends on compaction pressure and the resistance of the powder particles to deformation; compact densities about 90% of the theoretical level are common.

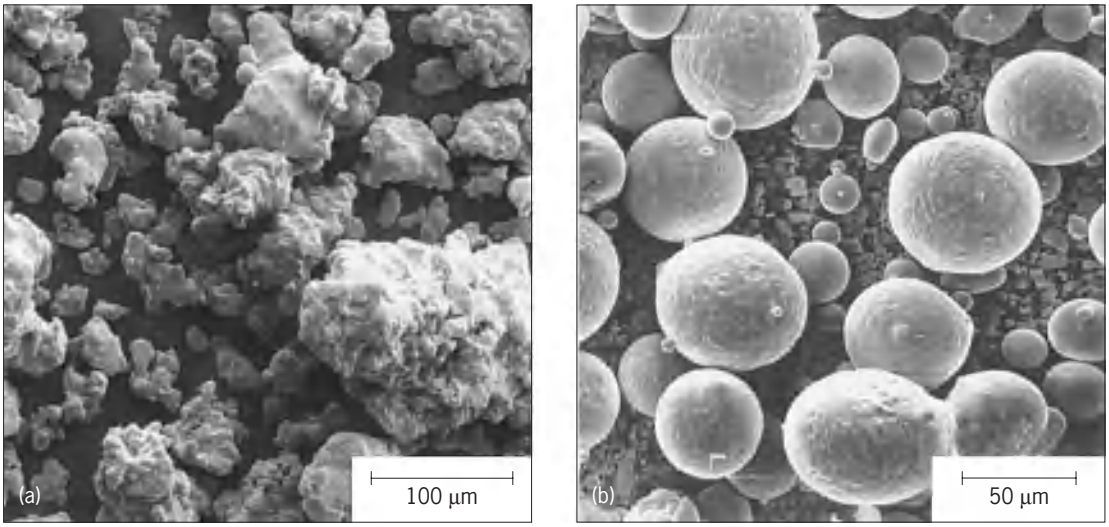

Fig. 1. Typical commercial iron-based powders: (*a*) water-atomized, (*b*) gas-atomized.

In the sintering process, the powder compact is heated below its melting point in order to promote bonding of the solid powder particles. The major purpose of sintering is to develop strength in the compact; normally an increase in the density of the compact also occurs, but the compact still contains pores. The internal architecture (microstructure) of the material is developed during sintering. As a rule of thumb, the sintering temperature must be higher than one-half the melting temperature (Kelvin scale) of the powder. This condition results in atomic diffusion and neck formation between the powder particles in the solid state. Commercial sintering furnaces are designed to provide controlled atmospheres during sintering. The atmosphere may be chemically neutral, reducing, or oxidizing with reference to the powder compact.

Normally, parts made by pressing and sintering require no further treatment. However, properties, tolerances, and surface finish can be enhanced by secondary operations. Examples of secondary or finishing operations are repressing, resintering, machining, heat treatment, and various surface treatments such as deburring, plating, and sealing. Warm compaction is a new process for enhancing powder density, whereby the powder is treated with an organic binder and then compacted at 100–150°C (212–302°F). *See* SINTERING.

Powder injection molding. This process builds on established injection molding technology used to fabricate plastics into complex shapes at low cost. The metal powder is first mixed with a binder consisting of waxes, polymers, oils, lubricants, and surfactants. After granulation, the powder-binder mix is injection-molded into the shape desired. Then the binder is removed and the remaining metal powder skeleton sintered. Secondary operations may be performed on the sintered part, similar to those for conventional press-plus-sinter parts. The viscosity of the powder-binder mix should be below 100 Pa · s (1000 poise) with about 40% by volume of binder, and spherical powders less than 20 μm in diameter are preferred. Powder injection molding produces parts which have the shape and precision of injection-molded plastics but which exhibit superior mechanical properties such as strength, toughness, and ductility. *See* PLASTICS PROCESSING.

Full-density powder processing. Parts fabricated by pressing and sintering are used in many applications. However, their performance is limited because of the presence of porosity. In order to increase properties and performance and to better compete with products manufactured by other metalworking methods (such as casting and forging), several powder metallurgy techniques have been developed that result in fully dense materials; that is, all porosity is eliminated. Examples of full-density processing are hot isostatic pressing, powder forging, and spray forming.

In the hot isostatic pressing process, the metal (or ceramic) powder is sealed in a metallic or glass container, which is then subjected to isostatic pressure at elevated temperature. Temperatures up to 2200°C (4000°F) and pressures up to 250 MPa (3.6×10^4 lb/in.^{2}) are possible in modern presses. After

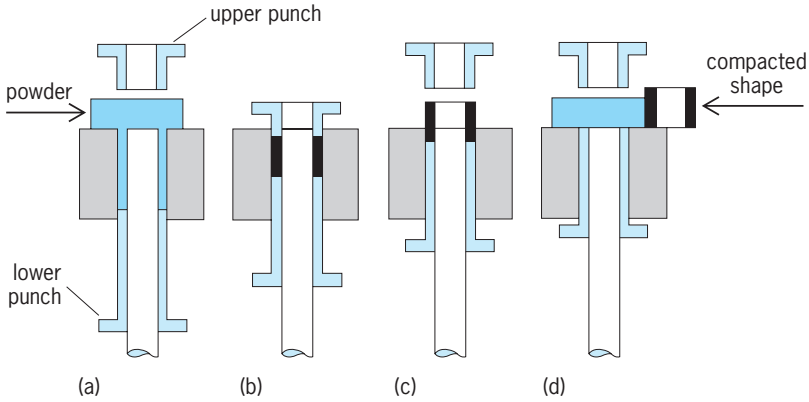

Fig. 2. Uniaxial powder compaction. (*a*) Die cavity is filled with powder. (*b*) Upper and lower punches simultaneously compact the powder in the die. (*c*) Upper punch is withdrawn and the powder compact is ejected by upward movement of the lower punch. (*d*) Compact is removed in order to repeat operating cycle.

complete densification of the powder, the compact is removed and the container stripped from the densified compact. Hot isostatic pressing is applied to the consolidation of nickel and titanium alloys, tool steels, and composites.

Powder forging is an adaptation of conventional forging in which a compact is prepared that is then sintered to retain about 20% by volume of porosity. This preform is transferred to a forging press and formed in a closed die cavity to its final shape with one stroke of the press. All porosity is removed in the forging operation. Powder forging is used primarily to fabricate components from low-alloy steels. *See* FORGING.

In the spray forming process, a stream of liquid metal is gas-atomized into a spray of droplets that impinge on a substrate. The spraying conditions are such that the droplets arrive at the substrate in the semisolid state. By control of the geometry and motion of the substrate, it is possible to fabricate sheet or plate, tubes, and circular billets from a wide range of ferrous and nonferrous alloys. It is a rapid process with a deposition rate up to about 2 kg (4 lb) per second.

Characteristics and applications. Powder metallurgy competes with several more conventional metalworking methods in the fabrication of parts, including casting, machining, and stamping. Characteristic advantages of powder metallurgy are close tolerances, low cost, net shaping, high production rates, and controlled properties. Other attractive features include compositional flexibility, low tooling costs, available shape complexity, and a relatively small number of steps in most powder metallurgy production operations. Control of the level of porosity is the intrinsic feature of powder metallurgy that enables the parts producer to predict and specify physical and mechanical properties during fabrication. Broad areas of usage for powder metallurgy parts include structural (load-bearing) components; and controlled-porosity, electrical, magnetic, thermal, friction, corrosion-resistant, and wear-resistant applications. Industries that make extensive use of powder metallurgy parts are aerospace, agriculture, automotive, biomedical, chemical processing, and electrical. In addition, the fabricators of domestic appliances and office equipment are dependent on the availability of a wide range of sizes and geometries of powder metallurgy parts exhibiting unique combinations of physical and mechanical properties.

Ferrous parts fabricated by pressing and sintering make up the largest segment of the powder metallurgy industry, and a majority of the applications are for the automotive market. In 1999 the typical United States passenger car contained more than 16 kg (35.2 lb) of powder metallurgy parts. Representative powder metallurgy parts include gears, bearings, rod guides, pistons, fuel filters, and valve plates. In the nonautomotive sector, examples of parts include gears, levers, and cams in lawn tractors and garden appliances; and gears, bearings, and sprockets in office equipment.

Press-plus-sinter fabrication techniques are used to create steel main bearing caps for 3.1- and 3.8-liter V-6 engines (**Fig. 3**). The 3.1-liter engine contains three caps weighing about 2.2 kg (4.8 lb), and the 3.8-liter engine contains four caps weighing slightly more than 3.6 kg (7.9 lb). With a density of 6.6 g/cm^3 (3.8 oz/in.^{3}), the caps have a tensile strength of 448 N/mm^2 (64,000 lb/in.^{2}), an apparent hardness of 70 units using the Rockwell method on the B scale (HRB), an elongation of 3%, a fatigue endurance limit of 159 N/mm^2 (23,000 lb/in.^{2}), and a Young's modulus of 107 kN/mm^2 (15.5 $\times$ 10^6 lb/in.^{2}). A recent development is the aluminum cam bearing cap.

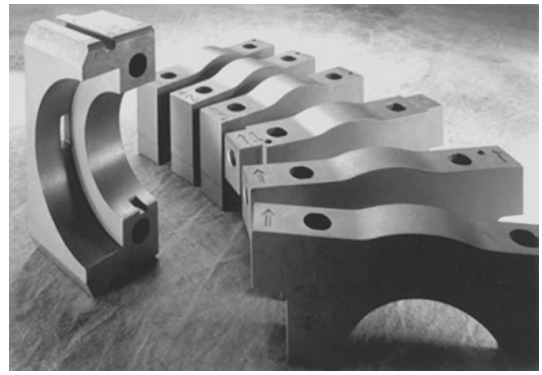

Fig. 3. Powder metallurgy main bearing end caps made by pressing and sintering. (*Metal Powder Industries Federation*)

Fig. 4. Automobile connecting rods made by powder forging. (*Metal Powder Industries Federation*)

Full-density processing has been used in the powder forging of steel automobile connecting rods (**Fig. 4**). The iron-copper-carbon rod is forged to a minimum density of 7.84 g/cm^3 (4.54 oz/in.^{3}), and has a tensile strength of 758 N/mm^2 (110,000 lb/in.^{2}), a yield strength of 551 N/mm^2 (79,900 lb/in.^{2}), and an endurance limit of 275 N/mm^2 (39,900 lb/in.^{2}). Important characteristics of powder forging include straightness, elimination of surface defects inherent in conventional forging, uniform microstructure, dimensional control, and superior machinability. Secondary operations include deflashing, shot peening, and machining.

An example of a part manufactured by powder injection molding is the miniature read/write latch arm made from stainless steel powder and used in a small high-capacity hard-disk drive (**Fig. 5**). The geometric form of the entire outside profile of the part is very critical relative to the pivot hole. Coining

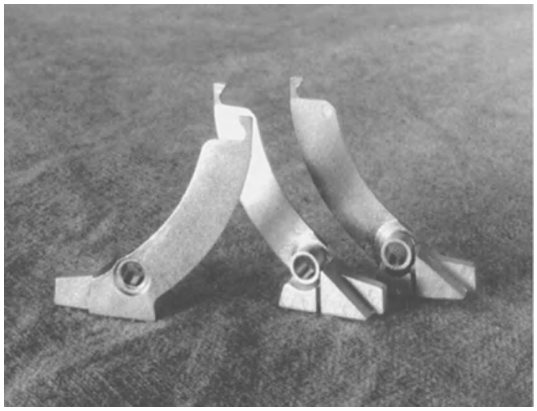

Fig. 5. Metal-powder injection-molded latch arms used in a hard-disk drive. (*Metal Powder Industries Federation*)

is the only secondary operation. *See* COINING.

Safety and health considerations. Because metal powders possess a high surface area per unit mass, they can be thermally unstable in the presence of oxygen. Very fine metal powders can burn in air (pyrophoricity) and are potentially explosive. Therefore, the clean handling of powder is essential; methods include venting, controlled oxidation to passivate particle surfaces, surface coating, and minimization of sparks or heat sources. Some respirable fine powders pose a health concern and can cause disease or lung dysfunction: the smaller the particle size, the greater the potential health hazard. Control is exercised by the use of protective equipment and safe handling systems such as glove boxes. There is no recognized hazard associated with the normal handling of the common grades of metal and alloy powders such as copper and iron. *See* INDUSTRIAL HEALTH AND SAFETY.

Alan Lawley

Bibliography. *ASM Handbook: Metal Powder Technologies and Applications*, vol. 7, 1998; R. German, *A-Z of Powder Metallurgy*, 2006; L. F. Pease III and W. G. West, *Fundamentals of Powder Metallurgy*, 2002; *Powder Metallurgy Design Manual*, 3d ed., 1998.

Power

The time rate of doing work. Like work, power is a scalar quantity, that is, a quantity which has magnitude but no direction. Some units often used for the measurement of power are the watt (1 joule of work per second) and the horsepower (550 ft-lb of work per second).

Usefulness of the concept. Power is a concept which can be used to describe the operation of any system or device in which a flow of energy occurs. In many problems of apparatus design, the power, rather than the total work to be done, determines the size of the component used. Any device can do a large amount of work by performing for a long time at a low rate of power, that is, by doing work slowly. However, if a large amount of work must be done rapidly, a high-power device is needed. High-power machines are usually larger, more complicated, and more expensive than equipment which need operate only at low power. A motor which must lift a certain weight will have to be larger and more powerful if it lifts the weight rapidly than if it raises it slowly. An electrical resistor must be large in size if it is to convert electrical energy into heat at a high rate without being damaged.

Electrical power. The power P developed in a direct-current electric circuit $P = VI$, where V is the applied potential difference and I is the current. The power is given in watts if V is in volts and I in amperes. In an alternating-current circuit, $P = VI \cos \phi$, where V and I are the effective values of the voltage and current and ϕ is the phase angle between the current and the voltage. *See* ALTERNATING CURRENT.

Power in mechanics. Consider a force F which does work W on a particle. Let the motion be restricted to one dimension, with the displacement in this dimension given by x. Then by definition the power at time t will be given by Eq. (1). In this equation W can be

$$P = \frac{dW}{dt} \tag{1}$$

considered as a function of either t or x. Treating W as a function of x gives Eq. (2). Now dx/dt represents

$$P = \frac{dW}{dt} = \frac{dW}{dx}\frac{dx}{dt} \tag{2}$$

the velocity v of the particle, and dW/dx is equal to the force F, according to the definition of work. Thus Eq. (3) holds. This often convenient expression for

$$P = Fv \tag{3}$$

power can be generalized to three-dimensional motion. In this case, if ϕ is the angle between the force $\mathbf{F}$ and the velocity $\mathbf{v}$, which have magnitudes F and v, respectively, Eq. (4) expresses quantitatively the observation that if a machine is to be powerful, it must run fast, exert a large force, or do both. *See* WORK.

$$P = \mathbf{F} \cdot \mathbf{v} = Fv \cos \phi \tag{4}$$

Paul W. Schmidt

Power amplifier

The final stage in multistage amplifiers, such as audio amplifiers and radio transmitters, designed to deliver appreciable power to the load. *See* AUDIO AMPLIFIER; RADIO TRANSMITTER.

Power amplifiers may be called upon to supply power ranging from a few watts in an audio amplifier to many thousands of watts in a radio transmitter. In audio amplifiers the load is usually the dynamic impedance presented to the amplifier by a loudspeaker, and the problem is to maximize the power delivered to the load over a wide range of frequencies. The power amplifier in a radio transmitter operates over a relatively narrow band of frequencies with the load essentially a constant impedance.

The usual mode of operation of power amplifiers is denoted by class A, AB, B, and C. Class C operation is limited to radio frequencies with a tuned load.

The other classes may be used for audio and high-frequency operation. For discussion of the modes of operation *see* AMPLIFIER

Class A. Class A operation is used when the amount of power transferred to the load is relatively small, say, less than 2 W. The amount of harmonic distortion introduced into the load voltage can be kept small by operating transistors within the nearly linear region of their characteristics. Class A operation is limited to low-power applications, such as small radios, cassette players, and compact-disk (CD) players, because the conversion efficiency (the efficiency of a power amplifier) is low. The maximum possible efficiency is 50%. However, for the usual operating conditions the efficiency is on the order of 10%. If the power amplifier were required to deliver 10 W with 10% efficiency, the transistor would have to be capable of dissipating an average power of 100 W. Furthermore, the power supply must be capable of supplying the power dissipated as heat plus the useful power delivered to the load. This poses an unnecessary burden upon the power supply. Other classes of operation have a higher conversion efficiency and are therefore used for higher-power applications.

Class AB. An improvement in the conversion efficiency can be obtained by using class AB operation. However, while a class A amplifier can be operated single-ended (one output transistor), a class AB amplifier must be operated push-pull. In class AB operation the transistor current does not flow for the complete cycle of the input voltage. In a single-ended circuit this would introduce excessive distortion. *See* PUSH-PULL AMPLIFIER.

Class B. This class is often used for the power amplifier in an audio amplifier. The amplifier in this class must be a push-pull circuit. Theoretically, with ideal transistors, the class B amplifier can have a conversion efficiency of 78.5%; practically, the efficiency is on the order of 50%, an appreciable improvement over that of class A operation.

Audio applications constitute a significant percentage of the applications of the class B amplifier. The requirements may range from a power output of a few watts for a small system, to 100 watts per channel for a home stereo system, to much more for a music concert sound system. The design of the loudspeaker loading is challenging, since the loading consists of three speakers (the woofer, midrange, and tweeter) operating over different frequency ranges. The rated loading is typically 8 Ω. The load may be transformer coupled to the output transistors, but it may also be connected directly in a complementary symmetry circuit. Since the operation for peak power output is over a wide range of the transistor characteristics, the design has to be done carefully to reduce the total harmonic distortion (THD) that can be generated by the nonlinear transistors. Typical figures for a high-quality home stereo system are 0.09% total harmonic distortion or less. If a load other than the rated value is used (say, a 4-Ω speaker system), the operating region of the output transistors is shifted, and increased total harmonic distortion may result. Operation at maximum power output with a load less than rated may also cause other damaging effects. *See* DISTORTION (ELECTRONIC CIRCUITS).

The amplifier bandwidth is another important parameter of an audio amplifier system. Typical figures for a home stereo system are 40 Hz for the lower half-power frequency to 20,000 Hz for the upper half-power frequency. At these limits, the power output is half the output at midrange for the same input signal. The frequency response characteristics of the speakers have to be considered in the overall picture (woofers for low frequencies and tweeters for high frequencies). High-quality systems give the user the ability to adjust with equalizers the system frequency response for personal preferences. *See* EQUALIZER; RESPONSE.

Class C. Because the collector current flows for less than one-half cycle of the input sinusoidal signal, this class of operation is restricted to radio-frequency operation where a tuned load is employed. The load is usually the input impedance of an antenna or of an antenna matching network. The load voltage will be nearly sinusoidal, even though the current flows in pulses, because of the relatively sharp tuning of the load. This phenomenon allows the amplification of large amounts of power at conversion efficiencies as high as 80%. This is extremely important for applications requiring delivery of large amounts of power to the load.

The driving source must usually be called upon to deliver power to the base circuit of the power amplifier, in many cases as much as 10% of the power delivered to the load. This requirement is not excessive. A class B power amplifier can be used in the driving circuit to obtain an efficient combination of driver and final amplifiers. Harold F. Klock

Bibliography. G. Ballou (ed.), *Handbook for Sound Engineers*, 2d ed., 1991; P. Chirlian, *Analysis and Design of Integrated Electronic Circuits*, 2d ed., 1987; A. Sedra and K. Smith, *Microelectronic Circuits*, 4th ed., 1997; R. Smith, *Electronics: Circuits and Devices*, 3d ed., 1987.

Power factor

In sinusoidal alternating-current theory, the cosine of the phase angle between the voltage across and the current through electrical circuitry. It equals the ratio of the mean power dissipated to the product of the root-mean-square voltage and current, and is unity for a circuit containing only pure resistances. It is less than unity if inductances and capacitances are involved. *See* ALTERNATING-CURRENT CIRCUIT THEORY.

Bryan P. Kibble

Power-factor meter

An instrument used to indicate whether load currents and voltages are in time-phase with one another.

In an alternating-current circuit, the current I may lead or lag the voltage E by some angle ϕ, and the average power P would be the average of either $(E_{max} \cos \omega t)$ $[I_{max} \cos (\omega t \pm \phi)]$ or $P = EI \cos \phi$. The factor $\cos \phi = P/EI$ is called the power factor of the load. *See* ALTERNATING-CURRENT CIRCUIT THEORY.

The formula $P = EI \cos \phi$ shows that, if two customers use equal currents, the energy loss in the line is the same for the two. If one of them has a lower power factor, that customer uses less energy, and a rate based only on energy delivered would involve an unequal charge for service.

Single-phase meter. This instrument contains a fixed coil that carries the load current, and crossed coils that are connected to the load voltage. There is no spring to restrain the moving system, which takes a position to indicate the angle between the current and voltage. The scale can be marked in degrees or in power factor.

The angle between the currents in the crossed coils is a function of frequency, and consequently each power-factor meter is designed for a single frequency and will be in error at all other frequencies. For low harmonic content the indication of the power-factor meter is usually accepted for contract purposes.

Polyphase meters. These meters are usually designed differently from single-phase meters. For balanced polyphase loads their application is straightforward. For a quarter-phase load, two voltages 90° out of phase are present and each can supply one of the crossed coils of the power-factor meter. If the voltages are unequal or their relative phase angle is different from 90°, the indication will no longer have meaning.

Three-phase power-factor meters connect the fixed coil in one line and the crossed coils between the other two lines. The meter is constructed with a 60° angle between the crossed coils. An unbalance in either voltage or angle will cause erroneous readings.

In the case of balanced four-phase and balanced six-phase loads, single-phase power-factor meters can be used directly without any correction.

In the case of unbalanced polyphase circuits with harmonics, various portions of the load may have different power factors and for the combined load there is no common phase angle; consequently the statement that the power factor is equal to $\cos \phi$ becomes meaningless. *See* NONSINUSOIDAL WAVEFORM. Harry Sohon; Edward C. Stevenson

Bibliography. S. D. Prensky and R. Castellucis, *Electronic Instrumentation*, 3d ed., 1982; M. B. Stout, *Basic Electrical Measurements*, 2d ed., 1960.

Power integrated circuits

Integrated circuits that are capable of driving a power load. The key feature of a power integrated circuit that differentiates it from other semiconductor technologies is its ability to handle high voltage, high current, or a combination of both.

In its simplest form, a power integrated circuit may consist of a level-shifting and drive circuit that translates logic-level input signals from a microprocessor to a voltage and current level sufficient to energize a load. For example, such a chip may be used to operate electronic displays, where the load is usually capacitive in nature but requires drive voltages above 100 V, which is much greater than the operating voltage of digital logic circuits (typically 5 V). At the other extreme, the power integrated circuit may be required to perform load monitoring, diagnostic functions, self-protection, and information feedback to the microprocessor, in addition to handling large amounts of power to actuate the load. An example of this is an automotive multiplexed bus system with distributed power integrated circuits for control of lights, motors, air conditioning, and so forth. *See* AUTOMOTIVE ELECTRICAL SYSTEM; ELECTRONIC DISPLAY; MICROPROCESSOR.

Technology. In addition to high-voltage and high-current operation, power integrated circuits are often also capable of performing many of the tasks associated with analog, digital, and sensor chip technologies.

Sensors. Due to the need to carry out a variety of diagnostic functions, power integrated circuits usually incorporate the current-, voltage-, and temperature-sensing functions that are normally performed by analog circuits. The sensor elements are placed on the power integrated circuit to provide local feedback at the point of use. Power modules can be made immune to short circuits or to misapplication of loads because the sensing function is performed locally. This enables local power interruption before the current in the system can rise to destructive levels.

Digital logic. Digital logic circuits incorporated on the power integrated circuit enable it to interface with microprocessors. The logic is implemented with complementary metal–oxide–semiconductor (CMOS) technology. Because of the high temperatures and rapid increase of voltage (voltage transients) encountered in typical applications for power integrated circuits, special measures are required during design and processing to make this CMOS logic immune from latch-up. The logic circuits may also need to be placed in special electically isolated silicon regions that can be floated up to high voltages. *See* LOGIC CIRCUITS.

Analog control. The control loop between the sensor and the protection circuit is usually implemented with bipolar analog circuits. Thus, bipolar and CMOS transistors must be fabricated simultaneously (BI-CMOS technology) during the manufacture of power integrated circuits. In some applications the bipolar transistors in the analog loop must have a very high frequency response to enable rapid implementation of control strategies for power regulation and integrated circuit protection.

High-voltage operation. Since the power integrated circuit usually operates at high voltages, unique process considerations become necessary. The design of

Applications of power integrated circuits

Application	Power requirements		Operating frequency*	Circuit configuration
	Voltage, V	Current, A		
Display drives	100–400	0.01–0.1	High	dc
Computer power supplies	50	10–50	High	dc
Motor drives	300–600	10–50	Medium	dc
Factory automation	300–600	10–50	Low	ac
Telecommunications	100–600	1–10	High	ac
Appliance control	300–600	1–20	Low	ac
Consumer electronics	50	0.1–10	Medium	dc
Lighting	300–600	1–10	Medium	ac
Smart house	300–600	5–50	Low	ac
Aircraft electronics	100–600	10–50	Medium	dc
Automotive electronics	50–100	1–20	Low	dc

*Low is less than 1 kHz; medium is 1–50 kHz; high is above 50 k Hz.

high-voltage transistors requires proper termination to achieve large blocking voltage capability. The need to bus high voltages around the chip requires special metal crossover design methodology and unique processes.

High-current operation. The fabrication of high-current devices on the same chip with high-density digital logic circuits requires the use of a double metal architecture in addition to the polysilicon gate technology. Further, if a high switching speed for the power device is desirable, the application of a refractory metal-silicide process becomes mandatory to obtain uniform turn-on and turn-off in the large-area power devices. Without the refractory gate process, the power dissipation in the power devices due to nonuniform switching can become prohibitive, leading to problems with reliability.

Dielectric isolation. The incorporation of multiple high-current bipolar devices on a monolithic chip becomes impossible with junction isolation technology because of the interaction between the adjacent devices. This problem can be solved by utilizing dielectric isolation technology. As the cost of dielectrically isolated silicon wafers is reduced by the development of improved techniques of wafer preparation, it is expected that most power integrated circuits will be fabricated with this starting material because it can simplify circuit design by eliminating circuit parasitics.

Requirements. The combination of digital logic circuits, analog sensing functions, high-voltage capability, and high-current devices must operate under harsh environmental conditions such as elevated ambient temperatures and high-voltage transients. All of these criteria must be satisfied under severe cost constraints. This demands continual advancements in device and process technology.

In most power integrated circuits, the power devices occupy a major portion of the chip area. Consequently, advances in the power-device structures, designs, and processes will dictate the pace of improvements in this technology.

Applications. Power integrated circuits are expected to have an impact on all areas in which power semiconductor devices are presently being used. In addition, they are expected to open up new applications based upon their added features. The wide spectrum of voltages and currents over which power semiconductor devices are utilized are summarized in the **table**, where the power requirements, system configuration, and operating frequency for the various applications are listed. At one extreme are display drives that require relatively low currents and moderate voltages. These applications are already served by smart power chips. At the other extreme are chips, under development for automotive electronics, that demand very high levels of current. *See* INTEGRATED CIRCUITS.

B. Jayant Baliga

Bibliography. P. Antognetti, *Power Integrated Circuits*, 1986; B. J. Baliga, *High Voltage Integrated Circuits*, 1988; B. J. Baliga, *Power Semiconductor Devices*, 1996.

Power line communication

A technology that uses electrical wiring as a medium to transmit digital signals from one device that has been rendered capable of employing this technology to another. Power line communication (PLC) enables the communication of data as well as multimedia messages by superimposing digital signals over the standard 50- or 60-Hz alternating current (AC). Low-speed power line communication has been used widely for some time in electrical utilities to provide voice communication, to control substations, and to protect high-voltage transmission lines. High-speed data transmission over power lines began in the late 1990s and was targeted to provide broadband Internet services and utility applications such as remote meter reading, as well as in home networking for data and multimedia communications. A short-range form of power line communication is also used for home automation and controls. *See* DATA COMMUNICATIONS; ELECTRIC POWER SYSTEMS; ELECTRIC POWER TRANSMISSION; INTERNET.

While high-speed power line communication indeed is used in developed countries to provide such

services, it is of greater interest in developing countries because of the lack of alternative broadband infrastructure such as cable television (CATV) or telephone lines [for digital subscriber lines (DSL)]. Using the existing electrical power distribution infrastructure, power line communication can provide areawide and even nationwide digital communication without the need to lay new cabling. The benefits are obvious, and thus many countries are testing and deploying power line communication technology, despite concerns that power line communication can introduce electromagnetic interference (EMI). Many countries are beginning to define regulations addressing spectrum usage and coexistence of power line communication with other existing and anticipated applications. *See* ELECTRICAL INTERFERENCE.

In-home power line communication provides another networking option in addition to the popular wireless communication technology and other wired technologies like Ethernet, Home PNA (Phone Networking Alliance), and fiber optics. The major benefit of in-home power line communication is that people are able to use existing power plugs as access points for computer communications and other applications like surveillance video and home automation.

Background. Traditional power circuits were designed for distributing electrical power to and within buildings. The existence of a plethora of sources of interference from nearby electrical devices, motors, and lights makes it a challenge to use the harsh environment of existing power lines as a high-frequency communication channel. The channel characteristics, such as noise, impedance, and attenuation of power lines, are found to be highly unpredictable and variable with time, frequency, and location. Fortunately, the rapid advances in digital communications, signal processing, and error-control coding enable quite successful power line communication systems.

Power line communication devices operate by injecting a carrier wave into the electrical wires at the transmitter. The carrier is modulated by digital signals. Each receiver in the system has a unique address and can receive and decode the signals transmitted over the wires. These devices may be either plugged into regular power outlets or permanently wired in place.

However, the harsh environment requires that power line communication devices use advanced signal modulation, digital signal processing, and error detection and correction techniques. These technologies convert the digital 0 and 1 into a useful carrier waveform that becomes the underlying physical-layer signals. The frequency band used in power line communication depends on the developer and technology used. Typically, the wider the frequency band, the higher the raw bit rate. For example, the X10 home automation system uses power line communication at the zero crossing voltage points in the AC wave (the two instants in each cycle when the AC voltage vanishes), and thus has a frequency of 120 Hz (for 60-Hz systems) and a relatively low data rate, good only for sending short bursts of data for controlling devices in the home. However, current (2006) high-speed power line communication devices use frequency bands in the range from 1 to 30 MHz and are able to provide raw data rates up to 200 megabits per second (Mbps). Power line communication chips operating at a raw data rate of 1 gigbit per second (Gbps) are contemplated, using higher bandwidths and more sophisticated communication techniques.

To provide simultaneous communication between power line communication devices, rules or protocols must be adopted to enable the effective use of the PLC channel for delivering digital data. These rules regulate how long and when a device can use the channel, and define the medium access (MAC) protocols of power line communication. Typical MAC protocols used in power line communication devices are carrier sense multiple access/collision avoidance (CSMA/CA), time-division multiple access (TDMA), and frequency-division multiple access (FDMA). While several proprietary systems (often backed by industry consortia) are available, there is an ongoing effort in international standards organizations to propose and adopt uniform standards for coexistence and interoperability of different power line communication products, for both in-home and Internet access domains.

Types of technology. Power line communication technologies may be distinguished by their voltage levels.

Low-voltage power line communication. Low-voltage power line communication devices usually are used in home networks. The benefit of using power lines in-home network infrastructures is the availability of power outlets almost everywhere in a home. Unlike wireless networks, power line communications do not require a specific access point since power outlets are the access points. Since the power line communication channel is essentially a broadcast channel, no switches are required in the communication; all devices communicate directly with each other. Because of these properties, to set up an in-home network based on power line communication requires no new wires, no drilling of walls, and no rewiring—the procedure is literally plug-and-play.

Low-voltage power line communication may also be used for temporary local-area network (LAN) setups for conferences, temporary medical centers, and other situations that require simple and fast setups. To set up a power line communication–capable computer, a user has only to find a power outlet and plug the device to it using either an Ethernet cable or a USB (Universal Serial Bus) connection, and then the computer is ready for communication. While most Ethernet or USB power line communication adapters are still separate from computer power supplies, manufacturers also embed the power line communication capabilities directly into device power

Comparison of throughputs of power line and wireless communication technologies

Technology	Power line	Wireless: 802.11n	Wireless: 802.11g
Theoretical maximum throughput	First-generation: 14 Mbps Turbo: 85 Mbps HomePlug AV/UPA: 200 Mbps	540 Mbps	54 Mbps
Actual expected throughput	First-generation: 10 Mbps Turbo: 40 Mbps HomePlug AV/UPA: 100 Mbps	150–200 Mbps	15–20 Mbps

supplies, so that the wires that provide electrical power to the device are also used to connect the power line communication system.

While low-voltage power line communication is also used for "last mile" Internet access, medium- and high-voltage power line communication systems are used to provide power line communication over much larger distances.

Medium- and high-voltage power line communication. Medium- and high-voltage power line communications usually are used for broadband over power line (BPL) or home electricity usage metering. Medium- and high-voltage power line communication systems can often cover longer distances with less attenuation and higher data rates because of the channel conditions and the marked absence of interfering sources common in low-voltage and in-home power line communication channels.

BPL provides broadband Internet access through ordinary power lines. It offers obvious benefits over regular cable or DSL connections; the extensive infrastructure already available should allow more people in more locations to have access to the Internet. The BPL Internet service provider (ISP) usually uses a coupler near a house to superimpose broadband signals into a low-voltage electric power line, delivering high-speed broadband capabilities to the end user. BPL can be used not only for computer communication, but also for voice over the Internet (VoIP) services and streaming video services such as Internet Protocol Television (IPTV). *See* VOICE OVER IP.

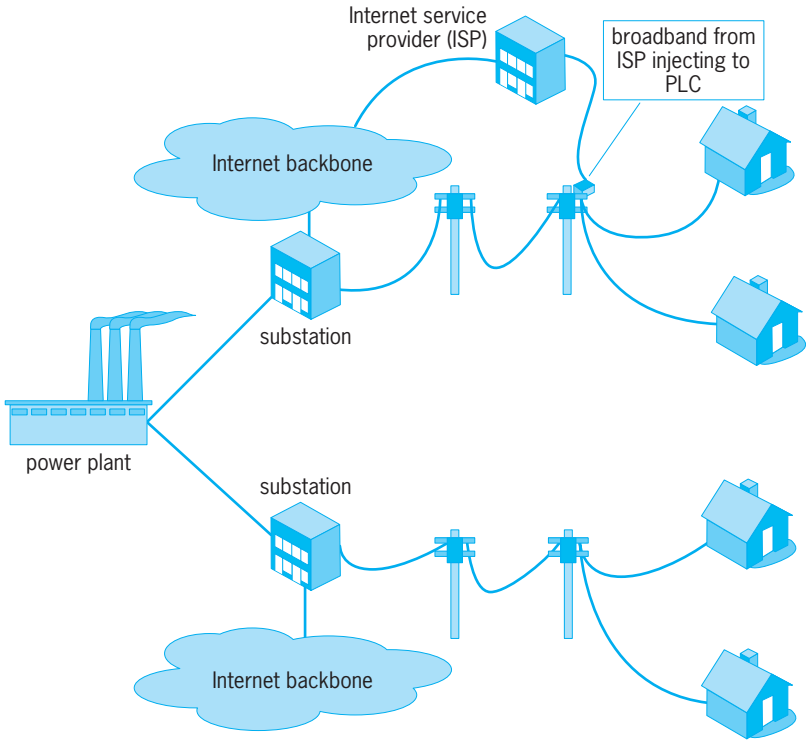

Broadband power line communication to homes.

Comparisons between technologies. Power line communication may be compared with alternative technologies.

Wireless and power line communication. Wireless communication and power line communication both make it possible to set up a local-area network with no new wires. The advantage of using one or the other technology is not as considerable as might be supposed. The two technologies are comparable in cost, and their throughputs are compared in the **table**. *See* WIRELESS FIDELITY (WIFI).

While power line communication may in some cases provide an alternative for wireless networking, or vice versa, the two technologies also can be used cooperatively. While power line communication has the advantage of using any power socket for power and communication, it is not operable at some locations that lack power sockets. On the other hand, while wireless certainly has the mobility advantage, it has been shown that wireless signals are not able to penetrate multiple concrete walls, and mobile devices eventually must connect to a power socket to recharge batteries. By combining power line networking technology with wireless access points, it is possible to reduce the "dead zone" problems and also provide another broadband access option to users.

BPL, DSL, and cable modem. Supporting broadband Internet access for remote areas has been a problem for a long time. Existing DSL and digital cable (CATV) providers are hesitant to set up infrastructure for these areas for fiscal reasons, and wireless access technology like satellite is expensive. The emerging WiMax technologies seem promising but are still questionable, especially in mountainous areas. Broadband over power line provides a viable solution in solving this "last mile" problem (see **illustration**).

Prospects. In-home power line communication is promising as it provides a convenient way for setting up a local-area network with good throughput and coverage. BPL for high-speed Internet access

is emerging as a viable option. However, BPL has several unsolved problems. Radio-frequency users reported severe interference during field tests of BPL equipment by some manufacturers. Regulations are being crafted to address these issues. In addition, international standards urgently need to be adopted to avoid incompatibilities in terms of both coexistence and interoperability for equipment from different manufacturers and for both in-home and access applications. The future of power line communications hinges on the ability of the power line communication community to resolve effectively these issues in the short term.

Haniph A. Latchman; Yu-Ju Lin; Sunguk Lee

Bibliography. H. A. Latchman and L. W. Yonge, Power line local area networking, *IEEE Commun. Mag.*, 41(4):32–33, April 2003; Y. Lin et al., A power line communication network infrastructure for the smart home, *IEEE Wireless Commun.*, 9(6):104–111, December 2002; Y. Lin et al., A comparative performance study of wireless and power line networks, *IEEE Commun. Mag.*, 41(4):54–63, April 2003; A. Majumder and J. Caffery, Jr., Power line communications, *IEEE Potentials*, 23(4):4–8, October–November 2004; D. Ruiz et al., In-home AV PLC MAC with neighboring networks support, pp. 17–21 in *2005 International Symposium on Power Line Communications and Its Applications*, April 6–8, 2005.

Power plant

A means for converting stored energy into work. Stationary power plants such as electric generating stations are located near sources of stored energy, such as coal fields or river dams, or are located near the places where the work is to be performed, as in cities or industrial sites. Mobile power plants for transportation service are located in vehicles, as the gasoline engines in automobiles and diesel locomotives for railroads. Power plants range in capacity from a fraction of a horsepower (hp) to over 10^6 kW in a single unit (**Table 1**). Large power plants are assembled, erected, and constructed on location from

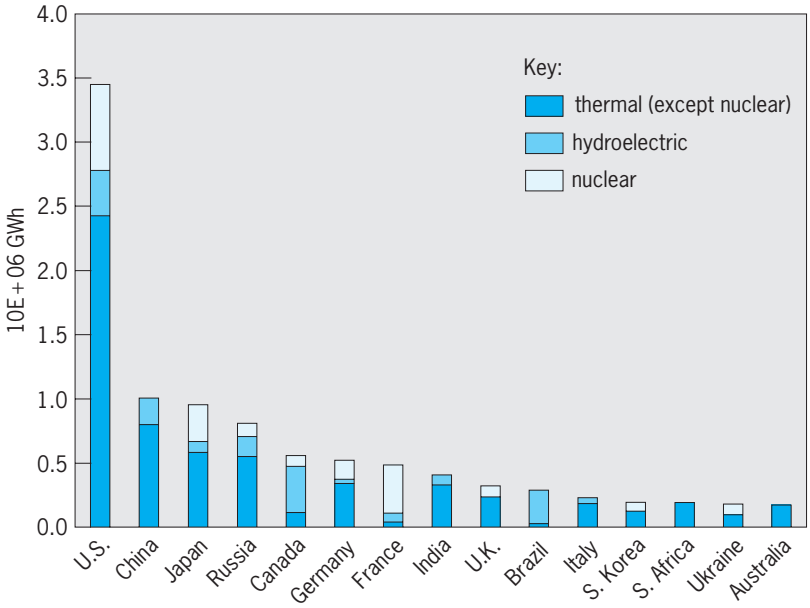

Fig. 1. Electric energy production by type in the 15 highest-producing countries, in gigawatt-hours. Nearly 27% of world electrical energy is generated in United States plants. (*U.S. DOE Publ. DOE/EIA-0219 (97), April 1999*)

TABLE 1. Representative design and performance data on power plants

Type	Unit size range, kW†	Fuel*	Plant weight, lb/kW†	Plant volume, ft³/kW‡	Heat rate, Btu/kWh§
Central station					
Hydro	10,000–700,000				
Steam (fossil-fuel-fired)	10,000–1,300,000	COG		20–50	8,200–15,000
Steam (nuclear)	500,000–1,300,000	N			10,000–12,000
Diesel	1,000–5,000	DG			10,000–15,000
Combustion turbine, simple cycle	2,000–250,000	D′G			8,600–15,000
Combustion turbine, simple cycle with recuperator	4,000–5,000	D′G			8,300–9,500
Combustion turbine, combined cycle	7,500–800,000	GD′			5,700–9,000
Industrial (by-product) steam	1,000–25,000	COGW		50–75	4,500–6,000
Diesel locomotive	1,000–5,000	D	100–200	2–3	10,000–15,000
Automobile	25–300	G′	5–10	0.1	15,000–20,000
Outboard motor	1–50	G′	2–5	0.1–0.5	15,000–20,000
Truck	50–500	D	10–20		12,000–18,000
Merchant ship, diesel	5,000–20,000	D	300–500		10,000–12,000
Naval vessel, steam	25,000–100,000	DON	25–50		12,000–18,000
Airplane, reciprocating engine	1,000–3,000	G′	1–3	0.05–0.10	12,000–15,000
Airplane, turbojet	3,000–10,000	D′	0.2–1		13,000–18,000

* C, coal; D, diesel fuel; D′, distillate; G, gas; G′, gasoline; N, nuclear; O, fuel oil (residuum); W, waste.
† 1 lb/kW = 0.45 kg/kW.
‡ 1 ft³/kW = 2.83×10^{-2} m³/kW.
§ 1 Btu/kWh = 1.055 kJ/kWh = 2.93×10^{-4} J(heat)/J(output).

equipment and systems made by different manufacturers. Smaller units are produced in manufacturing facilities.

Most power plants convert part of the stored raw energy of fossil fuels into kinetic energy of a spinning shaft. Some power plants harness nuclear energy. Elevated water supply or run-of-the-river energy is used in hydroelectric power plants. For transportation, the plant may produce a propulsive jet, as in some aircraft, instead of the rotary motion of a shaft. Other sources of energy, such as fuel cells, winds, tides, waves, geothermal, ocean thermal, nuclear fusion, photovoltaics, and solar thermal, have been of negligible commercial significance in the generation of power despite their potential in the future. *See* ENERGY SOURCES.

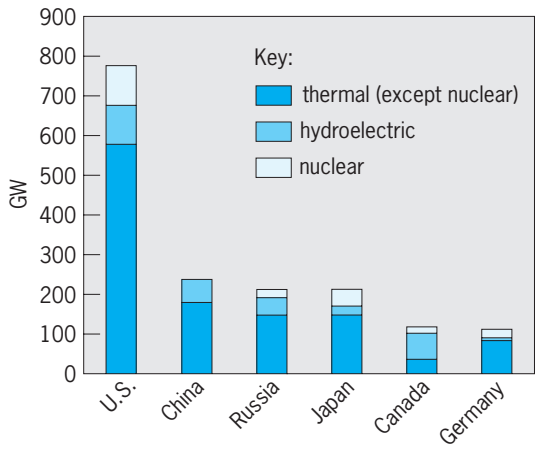

Fig. 2. The six countries with the highest capacities to generate electric power. (*U.S. DOE Publ. DOE/EIA-0219 (97), April 1999*)

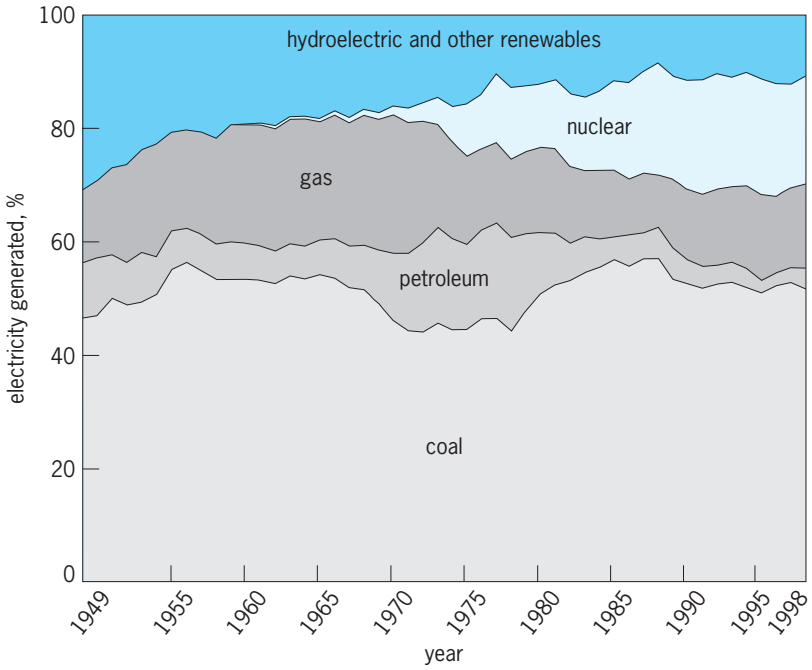

Fig. 3. Percentage of electrical energy generation in the United States utility industry (including Alaska and Hawaii since 1963), by fuel source. (*U.S. DOE Publ. DOE/EIA-0384 (98), July 1999*)

TABLE 2. Range of capacity factors for selected power plants

Power plants	Factor, %
Public-utility systems, in general	50–70
Industrial cogeneration, three-shift operation	80–95
Seagoing ships, long voyages	70–80
Seagoing ships, short voyages	30–40
Airplanes, commercial	20–30
Private passenger cars	1–3
Main-line locomotives	30–40
Interurban buses and trucks	5–10

If prices of conventional fuels (fossil and nuclear) rise faster than the general inflation rate, alternative energy sources become more significant. While they will not likely supply a large percentage of society's energy needs in the foreseeable future, they could someday become a significant part of the overall generating capacity (**Figs. 1–3**).

Figure 4 shows rudimentary flow or heat-balance diagrams for important types of practical power plants. **Figure 5** is a diagram for a by-product-type industrial steam plant (also known as cogeneration) which has the double purpose of generating electric power and simultaneously delivering heating steam by extraction or exhaust from the prime mover. *See* COGENERATION SYSTEMS; HEAT BALANCE.

Plant load. There is no practical way of storing the mechanical or electrical output of a power plant in the magnitudes encountered in power plant applications, although several small-scale concepts have been researched. As of now, however, the output must be generated at the instant of its use. This results in wide variations in the loads imposed upon a plant. The capacity, measured in kilowatts or horsepower, must be available when the load is imposed. Some of the capacity may be idle during extended periods when there is no demand for output. Hence a signifant portion of the potential output, measured as kilowatt-hours or horsepower-hours, may not be generated because there is no demand for output. Kilowatts cannot be traded for kilowatt-hours, and vice versa. *See* ENERGY STORAGE.

The ratios of average load to rated capacity or to peak load are expressed as the capacity factor and the load factor, Eqs. (1) and (2), respectively. The range

$$\text{Capacity factor} = \frac{\text{average load for the period}}{\text{rated or installed capacity}} \quad (1)$$

$$\text{Load factor} = \frac{\text{average load for the period}}{\text{peak load in the period}} \quad (2)$$

of capacity factors experienced for various types of power plants is given in **Table 2**.

Variations in loads can be conveniently shown on graphical bases as in **Figs. 6** and **7** for public utilities and in **Fig. 8** for air and marine propulsion. Rigorous definition of load factor is not possible with vehicles like tractors or automobiles because of variations in the character and condition of the running surface. In propulsion applications, power output

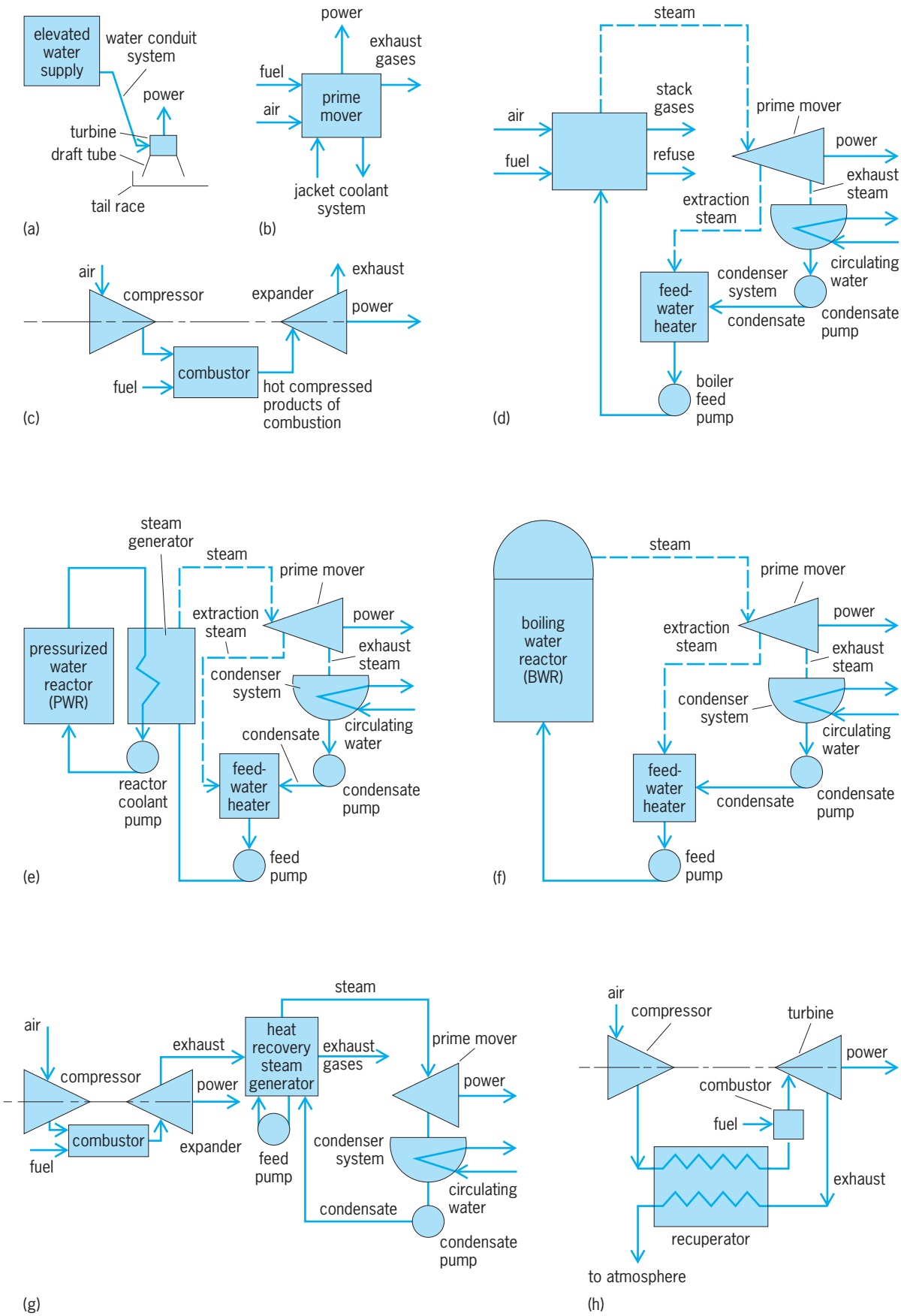

Fig. 4. Rudimentary flow or heat-balance diagrams for power plants. (*a*) Hydro. (*b*) Internal combustion. (*c*) Combustion-turbine. (*d*) Fossil-fuel–fired. (*e*) Nuclear steam (pressurized water reactor, PWR). (*f*) Nuclear steam (boiling water reactor, BWR). (*g*) Combustion turbine combined cycle. (*h*) Combustion turbine with recuperator.

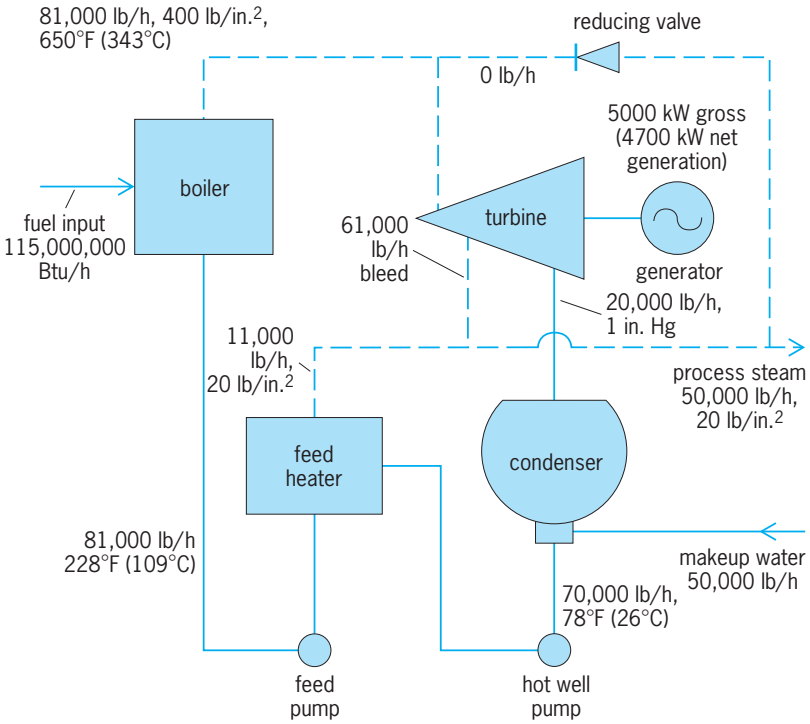

Fig. 5. Heat balance for a cogeneration industrial power plant delivering both electrical energy and process steam. 1 lb/h = 1.26 × 10^{-4} kg/s; 1 Btu/h = 0.293 W; 1 lb/in.^{2} = 6.895 kPa; 1 in. Hg = 3.386 kPa.

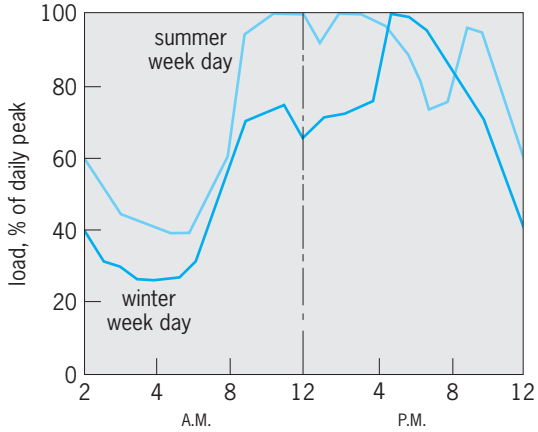

Fig. 6. Daily-load curves for urban utility plant.

may be of secondary import: performance may be based on tractive effort, drawbar pull, thrust, climb, and acceleration.

Plant efficiency. The efficiency of energy conversion is vital in most power plant installations. The theoretical power of a hydro plant in kilowatts is $QH/11.8$, where Q is the flow in cubic feet per second and H is the head (height of water intake above discharge level) at the site in feet. In metric units, the theoretical power of the plant in kilowatts is $9.8\,Q'H'$, where Q' is the flow in cubic meters per second and H' is the head in meters. Losses in headworks, penstocks, turbines, draft tubes, tailrace, bearings, generators, and auxiliaries will reduce the salable output 15–20% below the theoretical in modern installations. The selection of a particular type of waterwheel depends on experience with wheels at the planned speed and on the lowest water pressure in the water path. Runners of the reaction type (high specific speed) are suited to low heads (below

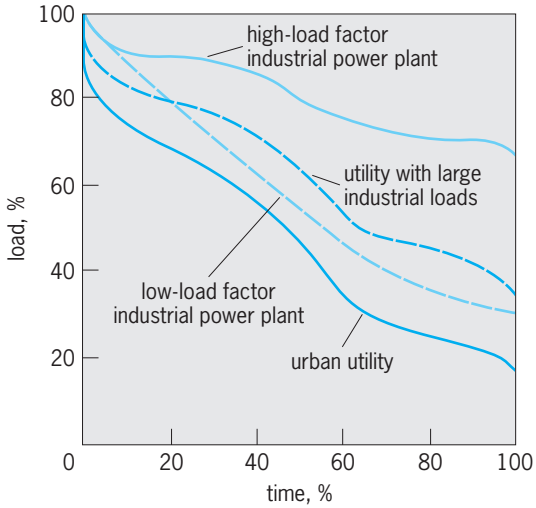

Fig. 7. Annual load-duration curves for selected stationary public utility power plants.

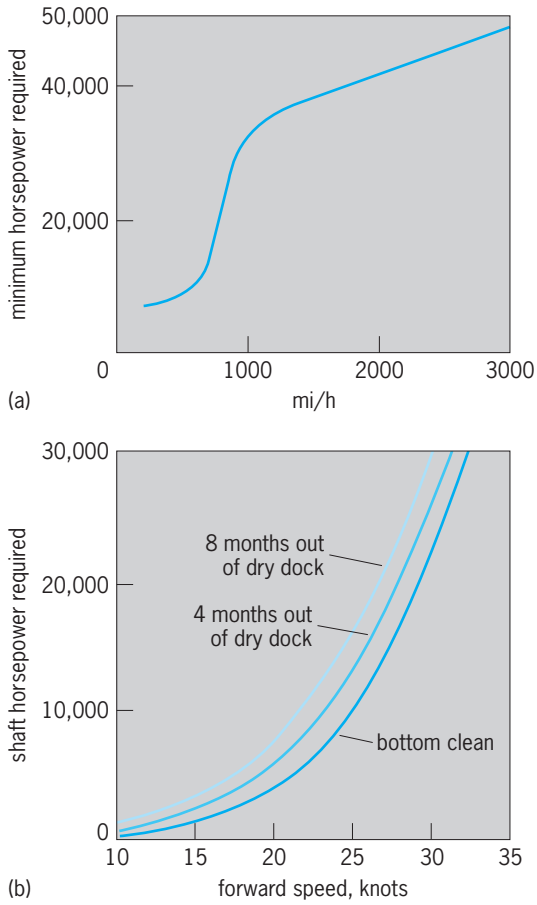

Fig. 8. Air and marine power. (*a*) Minimum power required to drive a 50-ton (45-metric-ton) well-designed airplane in straight level flight at 30,000 ft (9.1 km) altitude. (*b*) Power required to drive a ship, showing effect of fouling. 1 hp = 746 W; 1 mi/h = 0.447 m/s; 1 knot = 0.514 m/s.

500 ft or 150 m) and the impulse type (low specific speed) to high head service (about 1000 ft or 300 m). The lowest heads (below 1000 ft or 300 m) are best accommodated by reaction runners of the propeller or the adjustable blade types. Mixed-pressure runners are favored for the intermediate heads (50–500 ft or 15–150 m). Draft tubes, which permit the unit to be placed safely above flood water and without sacrifice of site head, are essential parts of reaction unit installations. *See* HYDRAULIC TURBINE;; WATERPOWER.

With thermal power plants the basic limitations of thermodynamics fix the efficiency of converting heat into work. The cyclic standards of Carnot, Rankine, Otto, Diesel, and Brayton are the usual criteria on which heat-power operations are variously judged. Performance of an assembled power plant, from fuel to net salable or usable output, may be expressed as thermal efficiency (%); fuel consumption (lb, pt, or gal per hp-h or per kWh); or heat rate (Btu supplied in fuel per hp-h or per kWh). American practice uses high or gross calorific value of the fuel for measuring heat rate or thermal efficiency and differs in this respect from European practice, which prefers the low or net calorific value.

Tables 1 and 2 give performances for selected operations. **Figures 9** and **10** reflect the improvement in fuel utilization of the United States electric power industry since 1900, although there have been some minor decreases since 1970. Figure 10 is especially significant, as it shows graphically the impact of technological improvements on the cost of producing electrical energy despite the harassing increases in the costs of fuel during the same period. **Figure 11** illustrates the variation in thermal performance as a function of load for an assortment of stationary and marine propulsion power plants. *See* BRAYTON CYCLE; CARNOT CYCLE; DIESEL CYCLE; OTTO CYCLE; RANKINE CYCLE; THERMODYNAMIC CYCLE.

Most of the improvements in heat rates were achieved by increasing steam temperatures and pressures, thus improving the Carnot cycle efficiency. The slight rise in heat rates since 1970 is due in large part to pollution control equipment, including the increased use of closed cooling water systems. The pollution control equipment is detrimental to heat rates because it uses large quantities of electric power. The closed cooling water systems raise heat rates by increasing the steam cycle condensing temperature.

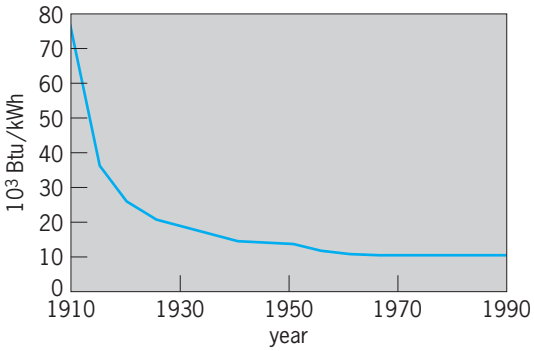

Fig. 9. Thermal performance of fuel-burning electric utility power plants in the United States. 1 Btu/kWh = 1.055 kJ/kWh = 2.93 × 10^{-4} J(heat)/J(output). (*Edison Electric Institute*)

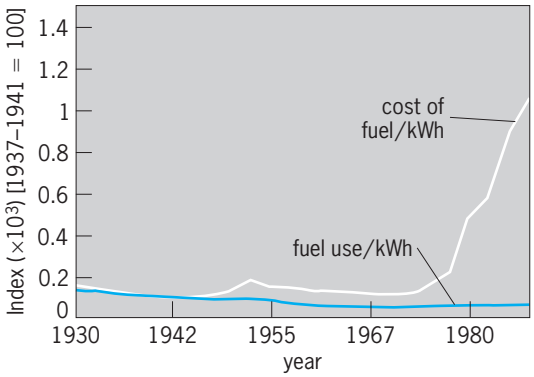

Fig. 10. Effect of efficiency of fuel use upon cost of fuel per kilowatthour generated in the United States utility industry (including Alaska and Hawaii since 1963). (*After Edison Electric Institute and Mon. Energy Rev., no. 4, December 1983*)

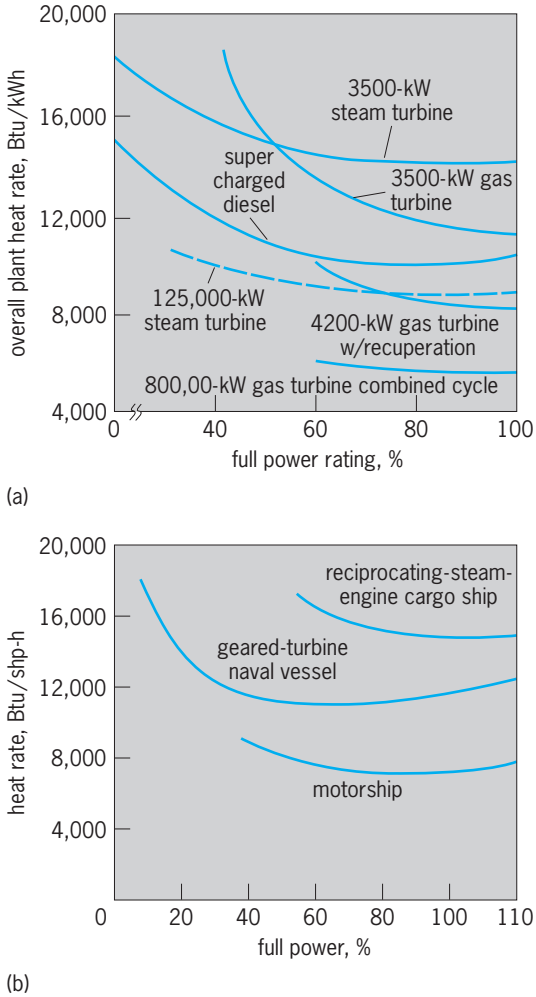

Fig. 11. Comparison of heat rates. (*a*) Stationary power plants. (*b*) Marine propulsion plants. 1 Btu/kWh = 1.055 kJ/kWh = 2.93 × 10^{-4} J(heat)/J(output); 1 Btu/hp-h = 1.415 kJ/kWh = 3.93 × 10^{-4} J(heat)/J(output).

In a steam plant the steam is condensed back to water after passing through the turbine (Fig. 4*d–g*). The lower the temperature at which this steam is condensed, the higher the Carnot efficiency. The circulating water used to condense the steam has traditionally come from large bodies of water such as rivers, lakes, and seas, but because of thermal pollution concerns many plants switched to a closed evaporative system with the circulating water being cooled by cooling towers or cooling ponds. The cooling towers and ponds are usually incapable of providing the low condensing temperatures of natural bodies of water. Further pollution requirements have dictated the use of plume-free cooling towers (where the saturated air exhausting from the tower has to be heated), with the corresponding loss of cycle efficiency, or the use of air-cooled condensers, resulting in usually higher back pressure and a considerable amount of electricity used in running their fans.

In scrutinizing data on thermal performance, it should be recalled that the mechanical equivalent of heat (100% thermal efficiency) is 2545 Btu/hp-h and 3413 Btu/kWh (3.6 megajoules/kWh). Modern steam plants in large sizes (75,000–1,300,000 kW units) and internal combustion plants in modest sizes (1000–20,000 kW) have little difficulty in delivering a kilowatt-hour for less than 10,000 Btu (10.55 MJ) in fuel (34% thermal efficiency). For condensing steam plants, the lowest fuel consumptions per unit output (8200–9000 Btu/kWh or 8.7–9.5 MJ/kWh) are obtained in plants with the best vacuums, regenerative-reheat cycles using eight stages of extraction feed heating, two stages of reheat, primary pressures of 4500 lb/in.^{2} gage or 31 megapascals gage (supercritical), and temperatures of 1150°F (620°C). An industrial plant cogenerating electric power with process steam is capable of having a thermal efficiency of 5000 Btu/kWh (5.3 MJ/kWh).

In the 1990s, combustion turbines used in combined cycle configurations began to take a dominant role in new power generation capacity. The primary reasons are the higher efficiency and lower emissions of the power plant in this arrangement. The rapid pace in advances in combustion turbine technology (such as higher firing temperatures that improve the Brayton cycle efficiency) has driven combined cycle efficiency to nearly 60% when using natural gas as fuel, while attaining low emission rates. Low fuel consumption (5700–6000 Btu/kWh or 6.0–6.3 MJ/kWh) is obtained by using higher firing temperatures, steam cooling on the combustor and gas turbine blades, a reheat steam cycle with a three-pressure heat recovery steam generator, and higher pressure and temperature of the steam cycle. These conditions are balanced with the need to keep the exhaust flue gas temperature as low as practical to achieve low emissions.

Gas turbines in simple cycle configuration are used mostly for peaking service due to their fast startup capabilities. The advances in the gas turbines have also increased the efficiency of simple cycle operations. Recuperation of the classic Brayton cycle gas turbine (simple cycle) is an accepted method of improving cycle efficiency that involves the addition of a heat exchanger to recover some portion of the exhaust heat that otherwise would be lost. *See* GAS TURBINE.

The nuclear power plant substitutes the heat of fission for the heat of combustion, and the consequent plant differs only in the method of preparing the thermodynamic fluid. It is otherwise similar to the usual thermal power plant. The pressure of a light-water reactor core is limited by material and safety considerations, while the temperature at which the steam is produced is determined by the core pressure. Because a nuclear reactor does not have the capability to superheat the steam above the core temperature, the steam temperature in a nuclear cycle is less than in a fossil cycle. *See* NUCLEAR REACTOR.

Power economy. Costs are a significant, and often controlling, factor in any commercial power plant application. Average costs have little significance because of the many variables, especially load factor. Some plants are short-lived and others long-lived. For example, in most automobiles, which have short-lived power plants, 100,000 mi (160,000 km) and 3000–4000 h constitute the approximate operating life; diesel locomotives, which run 20,000 mi (32,000 km) a month with complete overhauls every few years, and large seagoing ships, which register 1,000,000 mi (1,600,000 km) of travel and still give excellent service after 20 years of operation, have long-lived plants; electric central stations of the hydro type can remain in service 50 years or longer; and steam plants run around the clock and upward of 8000 h a year with very high reliability even when 25 years old. Such figures greatly influence costs. Furthermore, costs are open to wide differences of interpretation.

In the effort to minimize cost of electric power to the consumer it is essential to recognize the difference between investment and operating costs, and the difference between average and incremental costs. Plants with high investment (fixed) costs per kilowatt should run at high load factors to spread the burden. Plants with high operating costs (such as fuel) should be run only for the shortest periods to meet peak loads or emergencies. To meet these short operating periods various types of peaking plants have been built. Combustion (gas) turbines and pumped-storage plants serve this requirement. In the latter a hydro installation is operated off-peak to pump water from a lower reservoir to an elevated reservoir. On-peak the operation is reversed with water flowing downhill through the prime movers and returning electrical energy to the transmission system. High head sites (for example, 1000 ft or 300 m), proximity to transmission lines, and low incremental cost producers (such as nuclear or efficient fossil-fuel-fired plants) are necessary. If 2 kWh can thus be returned on-peak to the system, for an

input of 3 kWh off-peak, a pumped-storage installation is generally justifiable. *See* PUMPED STORAGE.

In any consideration of such power plant installations and operations it is imperative to recognize (1) the requirements of reliability of service and (2) the difference between average and incremental costs. Reliability entails the selection and operation of the proper number and capacity of redundant systems and components and of their location on the system network. Emergencies, breakdowns, and tripouts are bound to occur on the best systems. The demand for maximum continuity of electrical service in modern civilization dictates the clear recognition of the need to provide reserve capacity in all components making up the power system.

Within that framework the minimum cost to the consumer will be met by the incremental loading of equipment. Incremental loading dictates, typically, that any increase in load should be met by supplying that load with the unit then in service, which will give the minimum increase in out-of-pocket operating cost. Conversely, for any decrease in load, the unit with the highest incremental production cost should drop that decrease in load.

In the 1990s, an international trend toward privatization of generating plants and deregulation of the electric power industry emerged. In the United States, this is being adopted on a state-by-state basis and is expected to be widespread in the 2005–2010 time frame. The impact of these initiatives has been a significant increase in the construction of combustion turbine combined cycle plants relative to other types of power plants, such as coal and nuclear. Reasons for this selection include higher efficiency, shorter construction schedules, lower capital and operating costs, widespread fuel availability, and lower emissions. *See* ELECTRIC POWER GENERATION.

K. Keith Roe; Reginald S. Gagliardo

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; Babcock and Wilcox Co., *Steam: Its Generation and Use*, rev. 40th ed., 1992; Combustion Engineering, Inc., *Combustion: Fossil Power Systems*, 3d ed., 1981; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; M. C. Going, 1984 annual statistical report, *Elec. World*, 198(4):49–72, April 1984; R. L. Harrington, *Marine Engineering*, 1971; L. C. Lichty, *Combustion Engine Processes*, 7th ed., 1967; *U.S. Department of Energy Publication*, DOE/EIA-0219 (97), April 1999; *U.S. Department of Energy Publication*, DOE/EIA-0384 (98), July 1999.

Power shovel

A power-operated digging machine consisting of a lower frame and crawlers, a machinery frame, and a gantry supporting a boom which in turn supports a dipper handle and dipper (see **illus.**). Dipper sizes range from 1 to 180 yd^3 (0.7 to 137 m^3). The machines are powered by on-board diesel engines or by electric motors. Diesel-powered machines utilize a series of clutches and brakes that allow the operator to control various motions. Electric motor machines generally have individual motors for each motion, but occasionally clutches and brakes are used allowing one motor to drive two motions.

Machine components. The lower frame and crawlers serve to support the machine during digging and provide self-propulsion. The machinery frame rotates upon the lower frame by a large rolling-element thrust bearing. Machinery to raise and lower the dipper and to cause rotation are mounted on the machinery frame along with support points for the boom and gantry. The gantry protrudes above the cab and provides a means for connecting cables to support the boom. The boom is mounted on the front of the machinery frame. The dipper handle is mounted near the midpoint of the boom, where machinery is placed to cause the dipper handle to go in and out (crowd). The dipper is attached to the end of the dipper handle, and these two components are raised and lowered by a system of cables leading over a sheave at the point of the boom and back to the hoist machinery.

Machine operation. Four primary motions of the machine are propel, hoist, crowd, and swing. The controls for these motions are contained in the operator's station. The operation starts with the dipper positioned in front of the crawlers. The operator causes the dipper to enter the material to be dug by using the hoist and crowd motions. Once the dipper is full, it is retracted from the material and the machine is caused to rotate. To discharge the material, the bottom of the dipper is released. This digging cycle continues until no more material can be reached. The operator then propels the machine forward toward the bank.

Hydraulic excavators. A major variation of the power shovel is a machine whose motions are controlled and driven by hydraulic pumps, motors, and cylinders. The machine is operated similar to a mechanical machine, but the hoist and crowd motions are caused by hydraulic cylinders that activate the boom and arm. The boom is approximately half the length of a normal boom, and at its upper end the arm is pivotally connected (illus. *c*). At the other end of the arm, the dipper is pivotally connected. The operation of the machine is the same as that of a mechanical power shovel.

Backhoe machines. Mechanical power shovels and hydraulic excavators frequently use a backhoe front end (illus. *b*). The basic machine remains the same, but the lever arrangement of the front end is reversed so that the dipper movement starts extended away from the machine and is caused to dig downward into the material as it moves toward the machine.

Stripping shovels. Power shovels greater than 40 yd^3 (30 m^3) and equipped with long booms are used to remove overburden from underlying coal

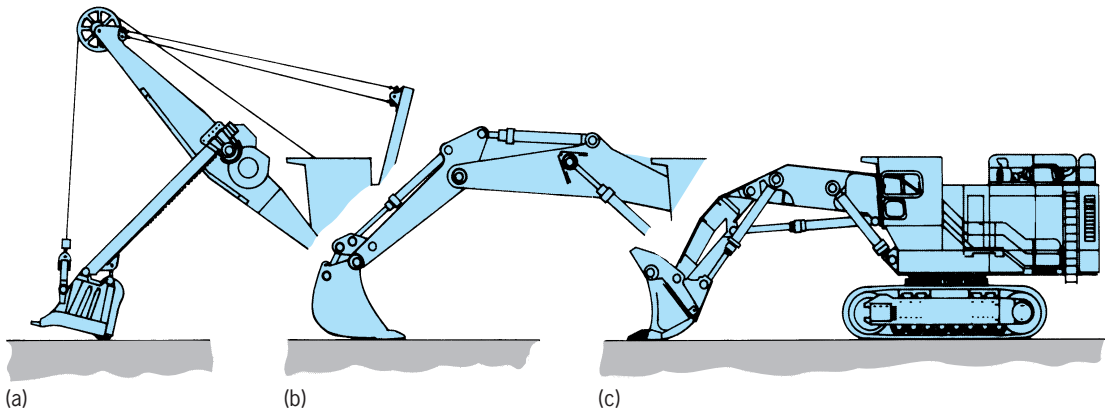

Comparison of the boom, dipper handle, and dipper for (*a*) conventional power shovel, (*b*) hydraulic backhoe, and (*c*) hydraulic front shovel. A typical lower frame, crawlers, and machinery frame are shown for the hydraulic shovel.

seams. These power shovels weigh up to 13,500 tons (12,000 metric tons). To compensate for this weight, four pairs of crawlers are used instead of the normal single pair. These machines are fully electrically operated.

Lattice boom machines. Occasionally the power shovel is converted to another material-handling device that uses a long lattice boom. Cranes, clamshells, grapples, orange peels, and dock or portal cranes are used but do not efficiently use the high-speed functions of the power shovel. When the power shovel is operated as a dragline, the speed of the motions are effectively used. From these draglines, a special device has evolved for operating on soft ground. The lower frames of the power shovel are replaced by a large circular frame that sits on the ground. To propel, the crawlers are replaced by two large shoes that are activated by a propel linkage and machinery mounted on the upper frame. *See* BULK-HANDLING MACHINES; CONSTRUCTION EQUIPMENT; HOISTING MACHINES. E. W. Sankey

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; K. Haddock, *A History of Marion Power Shovel's First 100 Years*, 1984; J. Hollingsworth, *History and Development of Strip Mining Machines*, Bucyrus Erie Co.; T. Shaw and V. Pavlovic, *Surface Mining and Quarrying: Mechanization, Technology, and Capacity*, 1992.

Poynting's vector

A vector, the outward normal component of which, when integrated over a closed surface in an electromagnetic field, represents the outward flow of energy through that surface. It is given by Eq. (1),

$$\Pi = \mathbf{E} \times \mathbf{H} = \mu^{-1}\mathbf{E} \times \mathbf{B} \qquad (1)$$

where $\mathbf{E}$ is the electric field strength, $\mathbf{H}$ the magnetic field strength, $\mathbf{B}$ magnetic flux density, and μ the permeability. This can be shown with the aid of Maxwell's equations (2), where $\mathbf{D}$ is the electric displacement and i the current density. Integration over any volume v and use of the divergence theorem to replace one volume integral by a surface integral give Eq. (3), where $\mathbf{n}$ is a unit vector normal to dS. In the volume integral, $^1/_2\mathbf{B} \cdot \mathbf{H}$ is the magnetostatic energy density, and $^1/_2\mathbf{D} \cdot \mathbf{E}$ is the electrostatic energy density, so the integral of the first two terms represents the rate of increase of energy stored in the magnetic and electric fields in v. The product of $\mathbf{E} \cdot \mathbf{i}$ is the rate of energy dissipation per unit volume as heat; or, if there is a motion of free charges so that $\mathbf{i}$ is replaced by ρv, ρ being the charge density, it is the energy per unit volume used in accelerating these charges. The net energy change must be supplied through the surface, which explains the interpretation of Poynting's vector.

$$\mathbf{H} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{H}) = \nabla \cdot (\mathbf{E} \times \mathbf{H}) = -\mathbf{i} \cdot \mathbf{E} - \mathbf{E} \cdot \frac{\delta \mathbf{D}}{\delta t} - \mathbf{H} \cdot \frac{\delta \mathbf{B}}{\delta t} \qquad (2)$$

$$-\int (\mathbf{E} \times \mathbf{H}) \cdot \mathbf{n}\, dS = \int_v \left[\frac{\delta}{\delta t}({}^1/_2\mathbf{B} \cdot \mathbf{H}) + \frac{\delta}{\delta t}({}^1/_2\mathbf{D} \cdot \mathbf{E}) + \mathbf{E} \cdot \mathbf{i} \right] dv \qquad (3)$$

It should be noted that this proof permits an interpretation of Poynting's vector only when it is integrated over a closed surface. In quantum theory, where the photons are localized, it could be interpreted as representing the statistical distribution of photons over the surface. Perhaps this justifies the common practice of using Poynting's vector to calculate the energy flow through a portion of a surface.

When an electromagnetic wave is incident on a conducting or absorbing surface, theory predicts that it should exert a force on the surface in the direction of the difference between the incident and the reflected Poynting's vector. *See* ELECTROMAGNETIC RADIATION; MAXWELL'S EQUATIONS; RADIATION PRESSURE; WAVE EQUATION. William R. Smythe

Prairie dog

A type of ground squirrel belonging to the family Sciuridae in the order Rodentia. These stout, short-tailed, short-legged ground squirrels inhabit open plains, short-grass prairies, and plateaus in the western part of North America from Canada to Mexico. The short coarse fur is grayish-brown. They have small beady eyes, pouched cheeks, and a short flat tail. Adults have a head-body length of 280–330 mm (11–13 in.), a tail length of 30–115 mm (1–4.5 in.), and weigh 0.7–1.4 kg (1.5–3 lb). *See* RODENTIA; SQUIRREL.

Classification. Prairie dogs are classified in the genus *Cynomys*. Obviously not dogs, they received this name because of their shrill doglike bark. Five species are currently recognized (see **table**).

Behavior. Prairie dogs are diurnal, terrestrial, social mammals. They once lived predominantly in colonies known as towns, with miles of well-worn tunnels and dens extending in every direction beneath plateaus and upland prairies. A large town may have had millions of inhabitants. In 1901, a prairie dog town north of San Angelo, Texas, was estimated to be 100 mi (160 km) in width and 250 mi (400 km) in length. If the population averaged 25 prairie dogs to the acre, then the expanse of 16,000,000 acres contained 400 million animals. Today, few dog towns of any size remain outside of refuges, parks, or sanctuaries. *See* SOCIAL MAMMALS.

The home is constructed by digging a tunnel straight down for 12 ft (3.7 m) or more. At the bottom of the tunnel, the prairie dog hollows out several rooms, including a sleeping room and a room for storing food. At the entrance to the hole, a mound of hard-packed dirt keeps water from entering the burrow. Most burrows have only one or two entrances, but some have as many as five or six. The prairie dog spends the winter in its home but often comes out on sunny days when the wind is not blowing. True hibernation has been reported for *C. leucurus*, *C. gunnisoni*, *and C. parvidens* but not for *C. mexicanus* or *C. ludovicianus*. At high altitudes and in the colder parts of their range, they sleep through the winter, but in southern localities they are active year-round. Prairie dogs feed on grasses, herbs, roots, and insects such as locusts and grasshoppers when they are abundant. Their feeding habits eliminate the taller plants. This deprives predators of cover, improves visibility around the burrow, and encourages fast-growing weeds without abundant seeds and fruits. Tall vegetation in the middle of the town is removed entirely. *See* BURROWING ANIMALS.

Breeding usually occurs during late winter and early spring. The single annual litter of four to six young are born in an underground nest in early May after a gestation period of 28–37 days. During the breeding season, females may defend individual territories around their burrows. The young are weaned at about 4 weeks of age. Sexual maturity is reached at 1 year of age. A black-tailed prairie dog has lived 8.5 years in captivity.

If an enemy, such as a coyote, should come into sight, prairie dogs warn each other with loud chirps or barks. They have a "vocabulary" of 10 different calls, ranging from a commonly used warning bark to a chuckle, a "fear" scream, and a fighting snarl. Each call is the result of a specific stimulus and elicits a characteristic action by nearby individuals. When the warning bark is sounded, they rush into their homes deep underground, where they are out of danger. Besides coyotes, prairie dogs are hunted by badgers, foxes, ferrets, and wild cats. Eagles and hawks swoop down and grab them. Rattlesnakes often hibernate in prairie dog burrows, and burrowing owls live in vacant burrows. *See* ANIMAL COMMUNICATION.

Habitat loss. Within their geographic range, suitable habitat for the prairie dog was maintained by grazing bisons and prairie fires. By the early 1900s, however, prairie dogs were considered a serious pest and economic menace to farmers and cattle ranchers in the West because of their feeding on crops and pastures. Also, their open burrows could lead to a running horse or cow breaking a leg. At one time, millions of prairie dogs lived in the western United States, but intensive public and private poisoning programs in the United States have greatly reduced their numbers. What was the largest single town remaining in the United States, covering approximately 110,000 hectares (425 mi^2) in South Dakota, was practically eradicated by poisoning in 1987–1988. The largest town now known, in northwestern Chihuahua, Mexico, occupies 35,000 ha (135 mi^2) and is part of a complex comprising about 55,000 ha (212 mi^2) and more than 1 million animals, but that population is jeopardized by habitat destruction and poisoning.

Relationship with black-footed ferret. The fate of the black-footed ferret (*Mustela nigripes*) is closely tied to prairie dog populations. At one time, the

Species and range of prairie dogs (genus *Cynomys*)

Species	Common name	Range
C. ludovicianus	Black-tailed prairie dog	Great Plains from Montana and southern Saskatchewan to northern Mexico
C. mexicanus	Mexican prairie dog	Northeastern Mexico
C. leucurus	White-tailed prairie dog	South-central Montana, western Wyoming, northeastern Utah, northwestern Colorado
C. parvidens	Utah prairie dog	South-central Utah
C. gunnisoni	Gunnison's prairie dog	Southeastern Utah, southwestern Colorado, Arizona, western New Mexico

Prairie dogs. (*Photo by Ron Singer/U.S. Fish and Wildlife Service*)

geographic range of the ferret coincided with the range of the prairie dog. The greatest cause of death of black-footed ferrets has been human activity in control of prairie dogs. The widespread use of sodium fluoroacetate ("1080") to poison prairie dogs has resulted in secondary poisoning of black-footed ferrets that fed on them. However, probably the greatest mortality results from the large-scale eradication of prairie dogs from the range of the black-footed ferret, thus greatly reducing the food supply. The black-footed ferret was never numerous enough to pose a serious threat to domestic stock, but it benefited early settlers by helping to keep populations of prairie dogs from expanding on grazing lands. Ironically, humans have driven the black-footed ferret to near extinction by reducing the habitat and number of prairie dog towns to levels where effective breeding populations no longer exist. This has resulted in the black-footed ferret becoming one of the most endangered mammals in North America. As of March 1997, the wild and captive ferret population numbered only approximately 500 animals. Since then ferrets have been reintroduced into suitable sites in Montana (1994), South Dakota (1994), and Arizona (1996). *See* ENDANGERED SPECIES; FERRET; POPULATION VIABILITY. Donald W. Linzey

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Praseodymium

A chemical element, Pr, atomic number 59, and atomic weight 140.91. Praseodymium is a metallic element of the rare-earth group. The stable isotope 140.907 makes up 100% of the naturally occurring element. The oxide is a black powder, the composition of which varies according to the method of

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

preparation. If oxidized under a high pressure of oxygen it can approach the composition PrO_2. The black oxide dissolves in acid with the liberation of oxygen to give green solutions or green salts which have found application in the ceramic industry for coloring glass and for glazes. *See* PERIODIC TABLE; RARE-EARTH ELEMENTS. Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

Prasinophyceae

A small class of mostly motile, photosynthetic, unicellular algae in the chlorophyll *a-b* phyletic line (Chlorophycota). It is segregated from the Chlorophyceae primarily on the basis of ultrastructural characters, especially the possession of one or more layers of polysaccharide scales outside the plasmalemma. Prasinophytes are mainly members of the marine plankton, but they are also found in brackish- and fresh-water habitats. A few are benthic, with both coccoid and colonial forms known, while others live symbiotically within dinoflagellates, radiolarians, and turbellarian worms. Approximately 180 species are known in 13 genera.

Cells of Prasinophyceae range in maximum dimension from 1 to more than 800 micrometers. Small cells are bean-shaped or ellipsoid, while large cells are spherical. Some genera lack cell walls, but others have thick walls composed of a noncellulosic material similar to sporopollenin or to pectin. Motile cells have from one to eight flagella usually covered with scales and hairs. The form of the flagellar and body scales is useful in classification. Flagella of uniflagellate and biflagellate cells usually arise laterally or apically, while those of multiflagellate cells usually arise from an apical depression. The flagellar root system is composed of a rhizoplast, an array of microtubular rootlets, and often a microbody. Flagellate cells have a nucleus located beneath the flagellar root system and a large, lobed, parietal chloroplast usually encompassing a thylakoid-traversed pyrenoid and an eyespot. Chloroplast ultrastructure and the photosynthetic pigments are similar to those of

other Chlorophycota. Some species have siphonein and siphonaxanthin, xanthophylls that otherwise are characteristic of the chlorophycean order Bryopsidales. Starch, which accumulates in the chloroplast, often as a sheath around the pyrenoid, is the main storage product, but mannitol, a sugar alcohol common in Chromophycota, also occurs in some species. Unlike other Chlorophycota, prasinophytes do not produce sucrose.

Sexual reproduction has not been observed with certainty. Binary fission is the only form of reproduction known for many species. Some benthic species and the plankton species that comprise very large, nonmotile, floating cells (phycomata) produce zoospores. Resting stages and cysts are known for some fresh-water species.

These scaly green flagellates are considered primitive in relation to other Chlorophycota. *See* ALGAE; PHYTOPLANKTON. Paul C. Silva; Richard L. Moe

Bibliography. R. E. Norris, Prasinophyceae, in S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, vol. 1, pp. 162–164, 1982; R. E. Norris, Prasinophytes, in E. R. Cox (ed.), *Phytoflagellates*, pp. 85–145, 1980.

Preamplifier

A voltage amplifier intended for amplifying a low-level input signal. The output is the input to another amplifier with a higher input level. For example, if the second amplifier needs a 500-mV signal to produce the desired power output and one has a microphone with a maximum output of 10 mV, one would use a preamplifier with a gain of 50 to achieve the desired power output. The term preamplifier is not precise; it has been applied to the portion of an amplifier preceding the power output stage and is not an additional system component. The term can also apply to a radio-frequency amplifier to amplify weak signals. The output is the radio receiver input. *See* AUDIO AMPLIFIER; RADIO-FREQUENCY AMPLIFIER; VOLTAGE AMPLIFIER.

The design of preamplifiers is critical because the input signal level is low. The noise introduced by transistors, hum from the power supply, and pickup from the leads must be minimized. Audio preamplifiers use bipolar or field-effect transistors, whereas radio-frequency preamplifiers use field-effect transistors only. The input impedance is generally high, and the output impedance is low. A frequency-compensation network may be incorporated into the design to compensate for the frequency characteristics of the input transducer to make the frequency response of the preamplifier–amplifier combination uniform. The bandwidth of the preamplifier must be sufficient to avoid reduction in performance. Operational-amplifier circuits can make good audio preamplifiers. The power output level is low since preamplifiers are not intended to drive loads such as that of a loudspeaker. *See* AMPLIFIER; ELECTRICAL NOISE; ELECTRONIC POWER SUPPLY; OPERATIONAL AMPLIFIER; TRANSISTOR. Harold F. Klock

Bibliography. G. Ballou (ed.), *Handbook for Sound Engineers*, 2d ed., 1991; N. J Sclater and S. Gibilisco, *Encyclopedia of Electronics*, 2d ed., 1990; A. Sedra and K. Smith, *Micro-electronic Circuits*, 4th ed., 1997.

Prebiotic organic synthesis

The plausible pathways by which the molecular precursors of life may have formed on the primitive Earth. Amino acids, the nitrogenous bases, and ribose phosphates can be prepared under conditions that might have prevailed on the primitive Earth. The linking together of amino acids to form polypeptides, and of nucleotides to form polynucleotides, has in principle been established.

Harold C. Urey's model of the primordial Earth postulates an atmosphere rich in methane, ammonia, water, and hydrogen. When this gas mixture is subjected to an electrical spark, analogous to the way that lightning may have initiated such syntheses 4 billion years ago, the identified products included several amino acids (glycine, alanine, and aspartic acid), the building blocks of proteins. This novel result lent credibility to a theory in which the origin of life was viewed as a cumulative, stepwise process, beginning with the gaseous synthesis of small molecules, which rained down into oceans, lagoons, and lakes. With water as a ubiquitous solvent, organic molecules could then react with one another to form larger molecules (biopolymers) and finally to assemble into primitive cells. This general scenario has guided the design of prebiotic simulations.

However, an accumulation of geophysical data and computational models has cast doubt on the relevance of the synthesis of amino acids to the primordial Earth. Hydrogen probably escaped rapidly as the Earth cooled, leaving an atmosphere in which methane and ammonia were virtually absent. As the input of hydrogen is diminished, the formation of biomolecules is inhibited. This problem has led some scientists to look for extraterrestrial sources of organic matter. For example, meteorites are known to contain a rich source of amino acids and other small biomolecules, and perhaps the infall of such cosmic bodies onto the young Earth gave life its start. Alternatively, there may have been localized environments on the Earth where methane and other hydrogen-rich precursors were abundant, such as deep-ocean hydrothermal vents, which would have been favorable for the formation of life. *See* AMINO ACIDS; HYDROTHERMAL VENT.

Role of hydrogen cyanide. Alternative routes to amino acids and other biomolecules begin with hydrogen cyanide (HCN), which can be synthesized from a variety of gas mixtures and energy sources. In a slightly alkaline solution of hydrogen cyanide in water; cyanide reacts with itself to form a pentamer known as adenine—an essential constituent of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). In fact, adenine can be synthesized from

hydrogen cyanide by two distinct routes, one occurring in the dark and the other proceeding by a light-induced process. *See* DEOXYRIBONUCLEIC ACID (DNA); RIBONUCLEIC ACID (RNA).

Hydrogen cyanide in water can provide a suite of molecular constituents found in proteins and nucleic acids. Following hydrolysis of the soluble products from hydrogen cyanide polymerization, the amino acids glycine, alanine, and aspartic acid were identified along with adenine and orotic acid (which undergoes a light-induced reaction to yield uracil, a component of RNA). Uracil can be formed by a direct pathway that does not require light. Guanine, another constituent of nucleic acids, can be formed by the reaction of urea with an imidazole derivative obtained from hydrogen cyanide. That such a diverse array of important biomolecules can be synthesized from such simple precursors has led many scientists to conclude that this compound may have played a central role in the origins of life. *See* CYANIDE.

Proteins and nucleic acids. Proteins consist of long chains of amino acids linked together by bonds known as peptides. Amino acids can be joined together by a variety of activating agents, among which the most plausibly prebiotic is a hydrogen cyanide derivative called cyanamide. Cyanamide has been used with amino acids to form small proteinlike structures with four peptide units, although the strategy could be applied to synthesize larger products. Polypeptides containing up to 50 amino acids have been prepared by using adenosine triphosphate (ATP) as the activating agent with clay minerals as catalysts. *See* ADENOSINE TRIPHOSPHATE (ATP).

A question regarding origins of life has been centered on whether proteins came before nucleic acids or vice versa. The sensible answer is "both," since each class of molecules was believed to serve a separate and unique purpose: proteins act as catalysts known as enzymes, accelerating biochemical reactions, while nucleic acids (polynucleotides) encode the information required to make those proteins. Ribonucleic acid can catalyze bond cleavage and bond formation. Thus, it is possible that nucleic acids could have served as both genes and enzymes in the first living things. *See* ENZYME.

If RNA was the first genetic material, there must have been a mechanism for it to make copies of itself without the aid of modern enzymes. In nucleic acid chemistry, the molecular copy is analogous to a photographic negative, since the pairing between nitrogen-containing bases requires a specific complementary sequence in the product: adenine directs the incorporation of uracil, guanine directs the incorporation of cytosine, and vice versa. Nonenzymatic replication is indeed possible, provided the procedure is begun with a set of activated precursors. Although there are some constraints on the range of polynucleotides that can serve as templates, replication proceeds with high fidelity—the products almost always contain the base expected from the pairing principle. The template can discriminate against nonbiological "left-handed" nucleotides, but these components also inhibit the reaction. However, the newly synthesized polynucleotide does contain a predominance of the same type of sugar-phosphate linkage as natural RNA strands. While these models rely on artificial starting materials that may not have been present on the primitive Earth, they do demonstrate that replication can proceed in the absence of highly evolved enzymes. *See* OLIGONUCLEOTIDE.

Polynucleotides. A more fundamental problem is where the first polynucleotide came from. By starting with an adenine nucleotide and a diphosphate-linked dimer, clay minerals can catalyze a synthesis up to the decanucleotide (10 bases long) containing predominantly the biologically correct ribose-phosphate bond. However, the assembly of a polynucleotide from the constituent nitrogenous bases, ribose and phosphate, is not an easy task. Baking the ingredients as a dry film can cause the ribose to join to the phosphate and to adenine or guanine, but the reaction is inefficient. Because ribose (a five-carbon sugar) is plagued by its instability and tendency to undergo side reactions, it has been proposed that a three-carbon alcohol, glycerol, fulfilled the linking role that was subsequently taken over by ribose; such nucleotide analogs can also undergo template-directed polymerization.

Six-carbon sugars can also serve the same structural function as ribose; a selective route was developed to phosphate-substituted sugars, starting from glycoaldehyde phosphate in alkaline solution. A major advantage to this pathway is that the phosphate group protects the products from decomposition. In the presence of formaldehyde, the major product is ribose 2,4-diphosphate (accounting for one-third conversion of the starting material), a remarkably efficient synthesis. More plausible routes to the constituent nucleotides will be articulated as ribose phosphates undergo greater scrutiny in the laboratory. *See* NUCLEIC ACID; PROTEIN.

William J. Hagan, Jr.

Bibliography. C. Chyba, The cosmic origins of life on Earth, *Astronomy*, 20:28–35, 1992; A. Eschenmoser and E. Lowenthal, Chemistry of potentially prebiological natural products, *Chem. Soc. Rev.*, 21:1–16, 1992; J. P. Ferris and W. J. Hagan, Jr., HCN and chemical evolution: The possible role of cyano compounds in prebiotic synthesis, *Tetrahedron*, 40:1093–1109, 1984; G. F. Joyce, RNA evolution and the origins of life, *Nature*, 338:217–224, 1989; S. F. Mason, *Chemical Evolution: Origins of the Elements, Molecules and Living Systems*, 1991; L. E. Orgel, Molecular replication, *Nature*, 358:203–209, 1992.

Precambrian

A major interval of geologic time between about 540 million years (Ma) and 3.8 billion years (Ga) ago, encompassing most of Earth history. The Earth probably formed around 4.6 Ga and was then subjected to a period of intense bombardment by

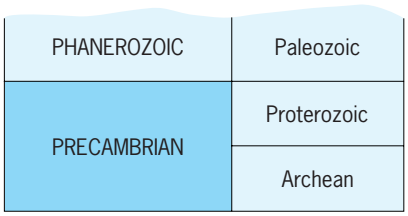

meteorites so that there are few surviving rocks older than about 3.8 billion years. Ancient rocks are preserved exclusively in continental areas. All existing oceanic crust is younger than about 200 million years, for it is constantly being recycled by the processes of sea-floor spreading and subduction. The name Hadean has been proposed for the earliest turbulent part of Earth's history. Development of techniques for accurate determination of the ages of rocks and minerals that are billions of years old has revolutionized the understanding of the early history of the Earth. *See* DATING METHODS; GEOLOGIC TIME SCALE; HADEAN; ROCK AGE DETERMINATION.

Detailed sedimentological and geochemical investigations of Precambrian sedimentary rocks and the study of organic remains have facilitated understanding of conditions on the ancient Earth. Microorganisms are known to have been abundant in the early part of Earth history. The metabolic activities of such organisms played a critical role in the evolution of the atmosphere and oceans. There have been attempts to apply the concepts of plate tectonics to Precambrian rocks. These diverse lines of investigation have led to a great leap in understanding the early history of the planet. *See* PLATE TECTONICS.

Subdivisions. Early attempts to subdivide the Precambrian were mainly based on geological characteristics such as rock structure and metamorphic grade. Meaningful subdivision of the Precambrian has become possible only with the development of sophisticated techniques for deciphering the ages of rocks. The Subcommission on Precambrian Stratigraphy of the International Union of Geological Sciences (IUGS) has suggested a subdivision based on specific time intervals (**Fig. 1**). Some have claimed that subdivisions of the Precambrian should correspond to specific evolutionary changes that can be observed in the rock record. Many of these changes occurred at widely different times in different parts of the planet so that "arbitrary" time subdivisions are somewhat artificial. To overcome some of these conceptual problems, a simple subdivision using geons, which are time units of 100 million years' duration, has been suggested (Fig. 1). The era names suggested by the IUGS have been widely accepted, whereas the smaller subdivisions (equivalent to periods in younger rocks) have not. Subdivision of the Precambrian has been contentious. Most agree on names for major subdivisions, but the concept of rigidly applicable global subdivisions based on arbitrary times poses philosophical problems for some.

Archean. Rocks of the Archean Eon (2.5–3.8 Ga) are preserved as scattered small "nuclei" in shield areas on various continents. The Canadian shield contains perhaps the biggest region of Archean rocks in the world, comprising the Superior province. Much of the Archean crust is typified by greenstone belts, which are elongate masses of volcanic and sedimentary rocks that are separated and intruded by greater areas of granitic rocks. The greenstones are generally slightly metamorphosed volcanic rocks, commonly extruded under water, as indicated by their characteristic pillow structures (**Fig. 2**). These structures develop when lava is extruded under water and small sac-like bodies form as the lava surface cools and they are expanded by pressure from lava within. Such structures are common in Archean greenstone assemblages in many parts of the world. *See* ARCHEAN; METAMORPHIC ROCKS.

Ultramafic magnesium-rich lavas, known as komatiites, are considered to have formed at high

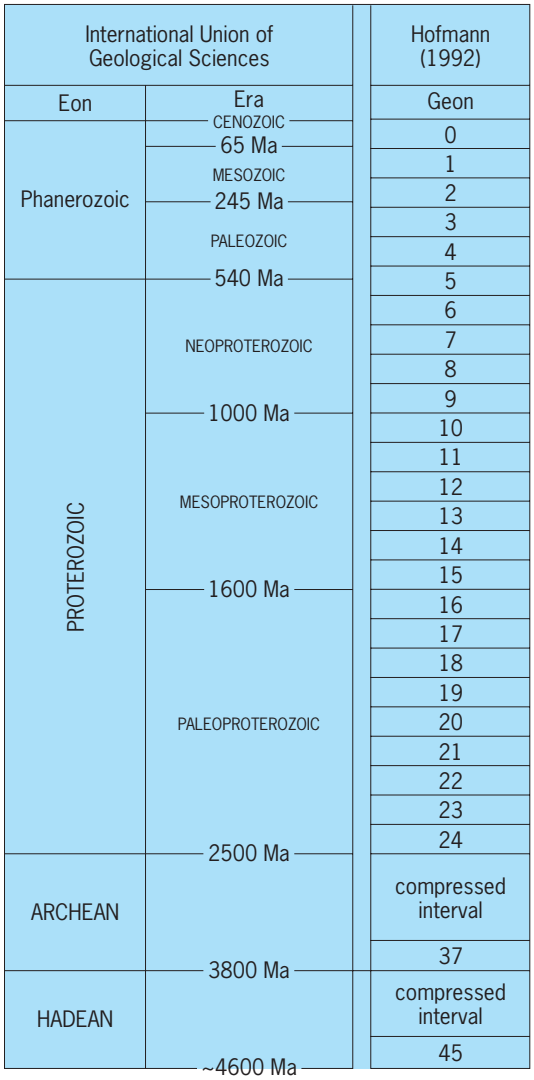

Fig. 1. Major subdivisions of geologic time. The Precambrian comprises the Archaean and Proterozoic eons.

Fig. 2. Neoproterozoic pillow lavas from Anglesey, North Wales.

temperatures (about 3000°F or 1650°C). These unusual lavas have been interpreted to mean that the interior of the Earth (mantle) was much hotter in the Archean than it is now. Geochemical and other investigations of lavas preserved in greenstone belts have led to the interpretation that they formed in a number of environments, analogous to present-day ocean floors, oceanic and continental volcanic arcs, and rifts.

Many of the associated sedimentary rocks are "immature" in nature, having formed by rapid erosion and deposition, from volcanic and, in some cases, older continental crustal sources. Such sedimentary rocks are commonly known as turbidites, formed as a result of gravity-controlled deposition in relatively deep water. Relatively rare "supermature" sediments formed in shallow-water environments as the result of extreme weathering. They are composed almost entirely of quartz and are known as orthoquartzites. In addition, iron-rich sediments formed by chemical precipitation from seawater (**Fig. 3**). These sediments indicate the existence of liquid water on the Archean Earth's surface and provide important clues concerning the nature of the ancient atmosphere and oceans. *See* QUARTZ; SEDIMENTARY ROCKS; TURBIDITE.

Many Archean shield areas also contain terranes composed of highly metamorphosed rocks (high-grade gneiss terranes). The origin of these gneissic terranes is controversial, but some may represent subduction zones located at ancient continental margins where sedimentary rocks were carried to great depths, metamorphosed, and intruded by igneous rocks. There has been a growing (although not universal) tendency to interpret Archean shield areas as the result of some form of plate tectonic activity, probably involving accretion of island arcs and collision of small continental nuclei. Theoretical considerations and the presence of magnesium-rich komatiitic lavas have led to the suggestion that the interior of the early Earth was much hotter than at present. The end of the Archean is marked in many places by amalgamation of large numbers of volcanic arcs and crustal fragments, which together formed the world's first large continental masses. This process appears to have occurred over a long period of time, beginning at about 3.0 Ga in South Africa and Australia and culminating at about 2.5 Ga in North America. The latter date has been adopted as the demarcation point between the Archean and the Proterozoic. *See* CONTINENTAL MARGIN; CONTINENTS, EVOLUTION OF; IGNEOUS ROCKS; LAVA.

Proterozoic. The Proterozoic Eon extends from 2.5 Ga until 540 Ma, the beginning of the Cambrian Period and Phanerozoic Eon. Proterozoic successions include new kinds of sedimentary rocks, display proliferation of primitive life forms such as stromatolites, and contain the first remains of complex organisms, including metazoans (the Ediacaran fauna). Sedimentary rocks of the Proterozoic Eon contain evidence of gradual oxidation of the atmosphere. Abundant and widespread chemical deposits known as banded iron formations (BIF) make their appearance in Paleoproterozoic sedimentary basins. *See* BANDED IRON FORMATION; PROTEROZOIC.

Paleoproterozoic. In most Proterozoic depositional basins there is a higher proportion of sedimentary than volcanic rocks. The world's first widespread glacial deposits are preserved in Paleoproterozoic successions. The existence of glacial conditions so long ago (~2.3 Ga) is indicated by the presence of unusual conglomerates (diamictites or tillites) with large rock fragments set in an abundant fine-grained matrix (**Fig. 4*a***). These peculiar conglomerates consist of a variety of rock fragments set in an abundant finer-grained matrix. Such conglomerates may form as a result of deposition of unsorted sediments from glaciers, in which case they are called tillites. Additional evidence of glaciation includes glacially scratched (striated) rock surfaces, striated and faceted rock fragments and "dropstones" (rock

Fig. 3. Banded iron formation (BIF) from the Archean of Finland. The light layers are silica-rich (chert), and the dark layers consist of iron oxides. The highly contorted nature of many such iron formations is due to early slump movements of these chemical sediments prior to complete consolidation.

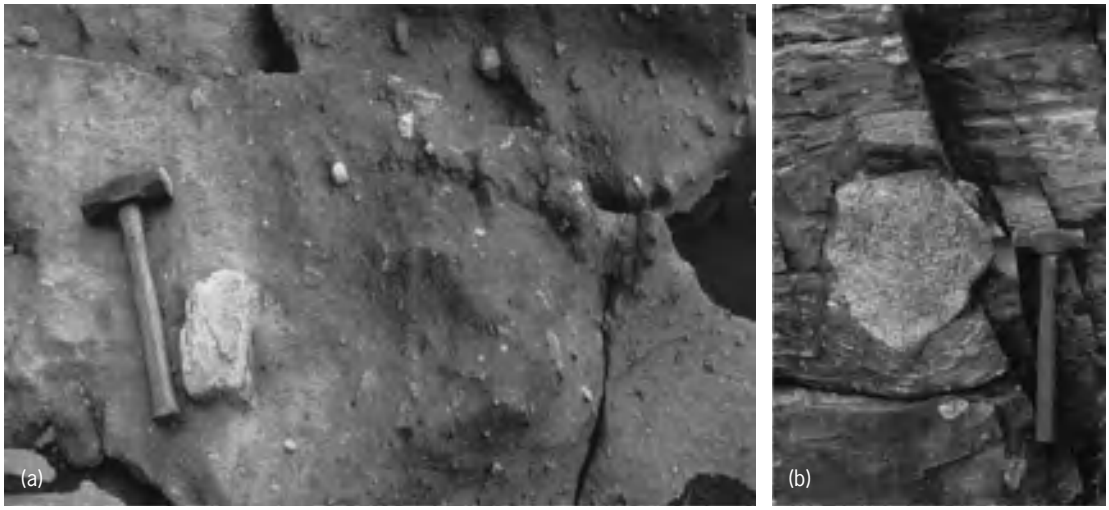

Fig. 4. Glacial sedimentary rocks. (*a*) Diamictites forming part of the Paleoproterozoic Gowganda Formation (~2.3 Ga) on the north shore of Lake Huron, Ontario, Canada. (*b*) Dropstones from the Gowganda Formation. Such large rock fragments in otherwise fine-grained bedded sedimentary rocks are thought to have been "vertically" emplaced from melting icebergs.

fragments that are interpreted as having been released, by melting, from floating icebergs into finely bedded sediments; Fig. 4*b*). Possibly contemporaneous Paleoproterozoic glacial rocks are known from five continents. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS.

Paleoproterozoic sedimentary rocks contain some of the world's greatest accumulations of valuable minerals, such as the uranium deposits of the Elliot Lake region in Ontario, Canada, and the great banded iron formations of the Lake Superior region. The uranium deposits are believed to have resulted from deposition of mineral grains in ancient river systems. Because uranium minerals (and associated pyrite) were not oxidized, it has been inferred that the Earth's atmosphere had a very low free oxygen content. By contrast, the somewhat younger (~2.0 Ga) banded iron formations are thought to indicate an oxygen-rich hydrosphere/atmosphere system.

The Paleoproterozoic also contains abundant carbonate rocks. Some of these contain spectacular organo-sedimentary structures called stromatolites (**Fig. 5**). The organo-sedimentary nature of these structures is confirmed by the preservation of microscopic primitive prokaryotic cells of photosynthetic cyanobacteria. These photosynthetic microorganisms trapped, or in some cases precipitated, carbonates from seawater, resulting in the construction of stromatolitic mounds of highly varied morphology. The metabolic activity of the microorganisms responsible for building these structures contributed to oxygenation of the atmosphere. Identical structures occur in some areas of the modern ocean, such as shallow-water saline environments of Shark Bay in Western Australia. Widespread development of stable continental shelves at the beginning of the Proterozoic (in contrast to the unstable, dominantly volcanic settings of much of the Archean) may have favored stromatolitic growth. *See* STROMATOLITE.

Two theories have emerged to explain oxygenation of the atmosphere. Increased photosynthetic activity, due to proliferation of stromatolites, is one. A second theory suggests that oxygen produced by marine photosynthetic microorganisms was consumed by reducing reactions related to widespread hydrothermal circulation of seawater at mid-ocean ridges and in other oceanic volcanic systems. Although there was significant production of photosynthetic oxygen, little was released to the atmosphere. This idea is supported by the preservation of different proportions of carbon isotopes (^{12}C and ^{13}C) in ancient organic carbon and carbonate rocks. Removal of ^{12}C during photosynthesis results in seawater that is enriched in the heavier isotope. The carbon-isotopic composition of seawater may, under certain conditions, be preserved in carbonate rocks. Differences in the ratios of carbon

Fig. 5. Stromatolites in the upper part of the Paleoproterozoic Snowy Pass Supergroup in southeastern Wyoming, United States.

isotopes of such carbonates and of organic carbon have been detected in many Archean sedimentary rock successions. These data suggest that photosynthetic activity was fairly widespread even during the early history of the Earth, but that release of significant amounts of free oxygen to the atmosphere was delayed until the tectonic regime changed from being ocean-dominated at the end of the Archean. *See* ATMOSPHERE, EVOLUTION OF; CYANOBACTERIA; MID-OCEANIC RIDGE; PHOTOSYNTHESIS; SEAWATER.

Another aspect of sedimentary geology that supports an increase in free oxygen during the Paleoproterozoic Era is the appearance of redbeds, sedimentary rocks that contain finely disseminated hematite. Many sedimentary rocks of this kind formed in subaerial or shallow-water settings where they were in close contact with the atmosphere, which must have contained at least some free oxygen, to impart the characteristic red or purple coloration of the highly oxidized iron. *See* REDBEDS.

Another indication of atmospheric change is the preservation of widespread banded iron formations at around 2.0 Ga. The major problem with these economically important and enigmatic chemical sedimentary rocks is that iron is virtually insoluble in present-day rivers and oceans. Under reducing conditions, such as those postulated for the Archean, the solubility of iron would have been greatly increased. The development of widespread banded iron formations in relatively shallow depositional settings in the Paleoproterozoic is thought to represent a great "flushing out" of iron from the oceans as a result of changes in the partial pressure of oxygen in the atmosphere/hydrosphere system. Paleoproterozoic limestones in various parts of the world display a period of exceptional ^{13}C enrichment at around 2.0 Ga, approximately the same time as the peak in banded iron formations production. One explanation for the high ^{13}C content of these carbonates is that photosynthetic organic activity was particularly high, so that considerable amounts of carbon were sequestered from atmospheric carbon dioxide (CO_2), leading to high ^{13}C content in the oceans and also to release of oxygen as a by-product.

By the end of the Paleoproterozoic, many of the oceans formed by breakup of the earliest(?) end-Archean "supercontinent," known as Kenorland, appear to have closed as a result of subduction and collisions to produce new large cratonic areas. For example, much of what is now the North American continent amalgamated at this time. There is plentiful evidence of widespread Mesoproterozoic igneous activity, in the form of intrusive bodies such as anorthosites and as abundant volcanic episodes, thought to be related to subduction and continental growth on the margins of large continental masses.

A notable feature of the Mesoproterozoic era is the dearth of evidence of glaciation during this long time interval. The Mesoproterozoic era was brought to a close by the Grenville orogeny, a widespread mountain-building episode that is named from the Grenville tectonic province in eastern North America. This orogenic period marks the construction of yet another supercontinent, known as Rodinia. Much of the subsequent Precambrian history in the Neoproterozoic is concerned with the complicated disintegration of Rodinia and culminates with the opening of a precursor to the present-day Atlantic Ocean, known as the Iapetus Ocean. *See* OROGENY.

Neoproterozoic. The period between the approximate end of the Grenville orogeny (~1.0 Ga) and the beginning of the Cambrian (~540 Ma) is known as the Neoproterozoic. This period is extremely important. It contains evidence of several glacial episodes that may have been the greatest that the world has known. Rocks belonging to this period preserve evidence of the first complex fossil forms, the controversial Ediacaran fauna, that is widely believed to represent the first animals with differentiated cells. These enigmatic fossils are believed by some to represent the ancestors of subsequent animal phyla. Others regard them as a "failed experiment" and assign them to completely separate, extinct taxa. Most Ediacaran forms occur in sandy rocks formed subsequent to deposition of the younger of two widespread glacial units, which is thought to have been deposited at about 620 Ma. *See* CAMBRIAN; EDIACARAN BIOTA.

As many as four or five glaciations have been postulated in the Neoproterozoic, but at most locations it is possible to document only two. The two most widespread glacial events are dated at about 750 Ma and 620 Ma and are respectively named Sturtian and Marinoan, from localities in Australia. The glacial deposits of the Neoproterozoic Era have recently come under close scrutiny for two main reasons. First, careful studies of the magnetic signature preserved in some of these glacial deposits (at least in parts of Australia), indicate deposition at near-equatorial latitudes. Second, some carbonate rocks associated with the glacial diamictites have very low $^{13}C/^{12}C$ ratios. This has been interpreted to mean a drastic reduction in photosynthetic activity, which would normally have led to preferential removal of the light carbon isotope from oceanic waters. It has been speculated that, during this period, life was almost eliminated from Earth and that ice extended into tropical latitudes—a condition referred to as the snowball earth. Alternatively, it has been suggested that increased obliquity of the Earth's ecliptic (greater inclination of the spin axis, relative to the orbital plane) could have led, during cold periods, to preferential buildup of snow and glacial ice in low latitudes. The ultimate cause of these great Proterozoic glaciations may lie in fluctuations in atmospheric greenhouse gases (particularly carbon dioxide). The reasons for such fluctuations remain obscure, but it has been proposed that increased weathering, as a result of uplift of the land surface during orogenic (mountain-building) episodes or greatly increased organic productivity, may have led to such climatic changes. *See* CLIMATE HISTORY.

Other aspects. Both water and life appear to have existed on Earth from at least 3.8 Ga. This evidence and the dearth of glacial deposits in the Archean have led to the "faint young sun paradox." In spite of much reduced solar radiation inferred by astrophysicists for the early part of Earth history, there is little evidence of low temperatures at the Earth's surface in Archean times. This has been explained by invoking much higher partial pressures of carbon dioxide during the early part of Earth history. The reasons for widespread Paleo- and Neoproterozoic glaciations, some of which may have occurred at low paleolatitudes, are not well understood. The theory of plate tectonics, which provides an elegant explanation of most of the features on the present-day surface of the Earth, appears to be applicable, perhaps in modified form, to rocks of both Archean and Proterozoic eons. The gradual buildup of free oxygen in the atmosphere/hydrosphere system was particularly important, for it permitted new metabolic pathways leading to the plethora of species that inherited the Neoproterozoic Earth and whose descendants survive to the present day. Grant M. Young

Bibliography. P. E. Cloud, *Oasis in Space*, 1988; P. F. Hoffman, United plates of America: Early Proterozoic assembly and growth of Laurentia, *Annu. Rev. Earth Planet. Sci.*, 16:543–603, 1988; H. J. Hofmann, New Precambrian time scale: Comments, *Episodes*, 15:122–123, 1992; A. H. Knoll, End of the Proterozoic Eon, *Sci. Amer.*, 262:64–72, 1991; S. R. Taylor and S. M. McLennan, The origin and evolution of the Earth's continental crust, *AGSO J. Austral. Geol. Geophys.*, 17:55–l62, 1997.

Precast concrete

Concrete that is cast into forms and cured at a location other than its final position in a structure. Products that are commonly fabricated in this manner include beams and joists, slab units, wall panels, and utility items such as pipe sections and access holes for workers. Cast stone that is made to simulate cut and dressed stone is a form of precast. Concrete block is also sometimes considered a form of precast. Precast concrete can be unreinforced or reinforced, and prestressed by the pretensioning or posttensioning method. Precast concrete is used with cast-in-place, structural steel, or load-bearing masonry construction. Field connections are more critical for precast than for monolithic cast-in-place concrete, because the connections must provide structural continuity and load transfer between elements.

Construction products can be fabricated at a precast concrete plant or in a convenient location on the project site. On-site methods of precasting are usually limited to tilt-up and lift-slab construction. Precast concrete is economical because it permits the efficient mass production of repetitious members and the reuse of forms; reduces the cost of on-site formwork, falsework, and shoring; and speeds construction. Precast concrete sections, particularly when they are prestressed, are often lighter weight than the equivalent cast-in-place section would be, representing a savings in materials and permitting a reduction in supporting foundations. Plant casting also provides better quality control over concrete mixing and curing procedures, facilitates assembly-line production methods, and permits work to proceed expeditiously in any weather. Some savings can be offset, however, by transportation and handling costs if hauling distances are long. *See* PRESTRESSED CONCRETE; REINFORCED CONCRETE.

Wall, floor, and roof systems. Complex concrete wall-panel shapes cannot be easily or economically cast in place. Precasting methods, however, make this cladding system competitive. Wall panels, sills, lintels, and decorative elements such as balusters and arches are often precast in special colors and finishes for exposed architectural applications. Wall panels are typically connected to structural supports by welded or bolted connections. Bolted connections and wedge inserts allow field adjustment for construction tolerances. Expansion joints are used to allow thermal expansion and contraction between panels, and exterior joints are protected against water penetration with elastomeric sealants.

Floor and roof slabs can be constructed with precast hollow-core planks, tees, double-tees, or channels. These are typically supported on masonry-bearing walls, on steel framing, or on cast-in-place or precast concrete beams and joists. Precast double-tees provide spans of 60 ft (18 m) or more, and are frequently used for parking structures. For shear transfer through diaphragm action, reinforcing steel from supporting elements can be tied into a relatively thin, reinforced or unreinforced topping slab cast in place on top of the precast members. *See* FLOOR CONSTRUCTION; ROOF CONSTRUCTION; WALL CONSTRUCTION.

Tilt-up walls. Large wall panels that are of simple design can be economically precast on site. The panels are cast in a horizontal position immediately adjacent to their wall location and then are tilted up into the final vertical position. This system is popular for warehouse and industrial buildings of simple rectangular plan because of its speed and economy. Doors and windows are easily installed by simply blocking out the proper size of opening. Walls are typically 5–8 in. (13–20 cm) thick and can function as deep, thin beams spanning horizontally between columns to eliminate the need for a continuous footing. The floor slab is frequently used as a casting platform, coated with a bond-breaking agent or covered with sheet materials to prevent bonding with the wall panel. Only side forms are needed, and the upper (outside) surface may be finished in the same way that a typical concrete slab is finished by floating or troweling. More often, however, tilt-up panels are given a decorative finish such as exposed aggregates.

When the wall panel has attained sufficient strength, it is lifted by embedded plates and temporarily braced until permanent attachments are

made. The lifting and handling process induces significant stress on the panels, so the lifting plates must be carefully located and installed to prevent damage. Permanent welded connections can be made to steel columns or to precast concrete columns with embedded steel plates. Alternatively, cast-in-place columns can be poured after the wall panels have been erected, with reinforcing steel from the panels extending into the column for structural connection.

Lift-slab construction. Floor and roof slabs for multistory buildings can be cast on a base slab at or near ground level and then lifted into their final position. The slabs are cast in succession one atop the other, with the roof slab cast last, on top. A bond-breaking compound or membrane is used to keep the slabs from bonding to one another. Columns are erected first, usually to the full height of the structure, and the slabs are cast around them. An embedded steel collar is placed around each column opening for subsequent lifting and attachment. Jacks mounted to the columns are used to lift the slabs into place, beginning with the roof slab, and followed by successive floor slabs, working from top to bottom. The slabs are permanently welded to steel supporting plates attached to each column.

Lift-slab construction eliminates most on-site formwork, reduces transportation and handling costs, and greatly reduces construction time. This method is also sometimes called the Youtz-Slick method after its two inventors. *See* BUILDINGS; CONCRETE SLAB.

Civil engineering applications. Precast concrete pipe sections, box culverts, access holes, and other shapes make the construction of utility, highway, and bridge projects faster and more economical. Precast items can be manufactured with steel or fiber reinforcing to produce the thinnest and lightest-weight sections to reduce transportation costs. Field connections of various types are required to provide structural or watertight jointing. *See* CONCRETE.

Christine Beall

Bibliography. G. Balaguru, *Concrete Handbook*, 1997; F. S. Merritt, *Building Construction Handbook*, 6th ed., 2000.

Precession

The motion of an axis fixed in a body around a direction fixed in space. If the angle between the two is constant so that the axis sweeps out a circular cone, the motion is pure precession; oscillation of the angle is called nutation. An example of precession is the motion of the Earth's polar axis around the normal to the plane of the ecliptic; this is the precession of the equinoxes. A fast-spinning top, with nonvertical axis, which precesses slowly around the vertical direction, is another example. In both examples the precession is due to torque acting on the body. Another kind of precession, called free or fast precession, with a rate which is comparable to the rotation rate of the body, is seen, for instance, in a coin spun into the air. *See* NUTATION (ASTRONOMY AND MECHANICS); PRECESSION OF EQUINOXES.

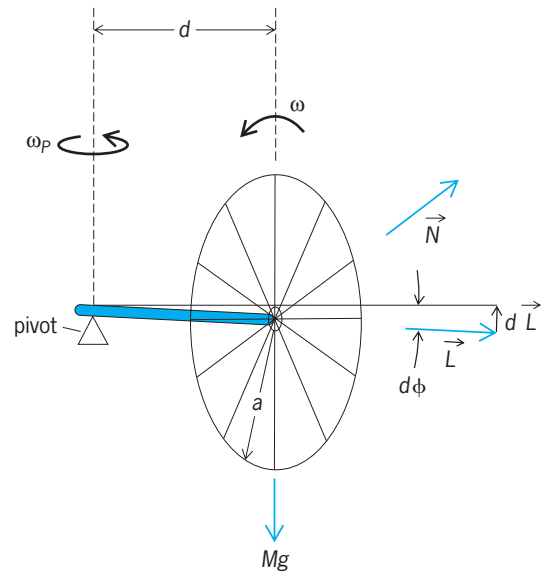

Fig. 1. Simple precession of a rapidly spinning wheel with a horizontal axis supported by a pivot.

Gyroscopic motion. As a simple example of gyroscopic motion, consider a rapidly spinning wheel with a horizontal axis supported at a distance d from the plane of the wheel (**Fig. 1**). The angular momentum $\vec{L}$ is along the wheel symmetry axis and is approximately given by the angular momentum of the wheel about this axis; in the simple precession approximation to the motion, the angular momentum associated with precessional motion is neglected. The external torque $\vec{n}$ due to the gravitational force is perpendicular to the wheel axis in the horizontal plane. The change in the angular momentum $\vec{L}$ in an infinitesimal time interval dt is given by the rotational equation of motion in Eq. (1).

$$d\vec{L} = \vec{N}\,dt \tag{1}$$

For the derivation of Eq. (1) *see* RIGID-BODY DYNAMICS.

Since $d\vec{L}$ and $\vec{L}$ are perpendicular, the length $\vec{L}$ is unchanged to first order in dt. The direction of $\vec{L}$ is rotated counterclockwise in the horizontal plane through an angle $d\phi$ given by Eq. (2). The angular

$$d\phi = \frac{dL}{L} = \frac{N}{L}dt \tag{2}$$

velocity of precession ω_P about the vertical axis is then given by Eq. (3).

$$\omega_P = \frac{d\phi}{dt} = \frac{N}{L} \tag{3}$$

The magnitude of the gravitation torque is Mgd, where M is the mass of the wheel and g is the gravitational acceleration. The spin angular momentum is $L = I\omega$, where ω is the spin angular velocity and

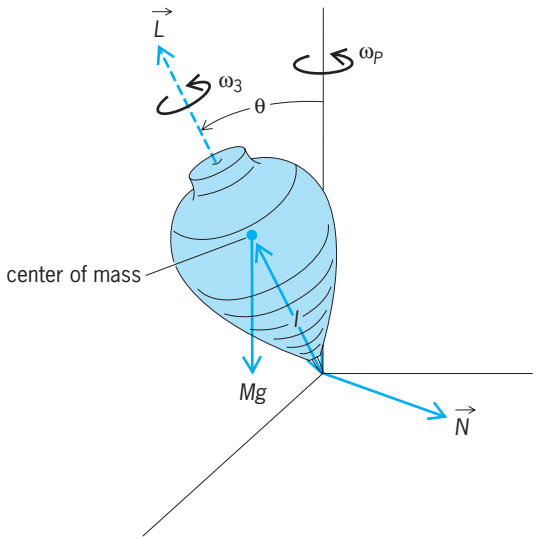

Fig. 2. Simple precession of a fast-spinning top about the vertical axis.

$I = Ma^2$ is the moment of inertia of the wheel of radius a. Thus ω_P, as given in Eq. (4), is independent

$$\omega_P = \frac{gd}{\omega a^2} \tag{4}$$

of M and inversely proportional to ω. For ω very large, the precession rate ω_P is quite slow. *See* GYROSCOPE.

Spinning top. The precessional motion of a top inclined at an angle θ from the vertical and having a fixed point of contact with a supporting surface (**Fig. 2**) can be readily deduced from the preceding discussion. For a fast-spinning top the angular momentum is nearly along the symmetry axis of the top, and is given by Eq. (5) where I_3 is the moment of

$$L \simeq I_3\omega_3 \tag{5}$$

inertia about this axis and ω_3 is the angular velocity. The change in $\vec{L}$ is given by Eq. (1) with the torque $\vec{n}$ along the azimuthal direction. Again the magnitude of $\vec{L}$ is unchanged to first order in dt, and from Fig. 2 the change in azimuthal angle ϕ is given by Eq. (6). The magnitude of N is $Mg\ell \sin\theta$, where ℓ

$$d\phi = \frac{dL}{L\sin\theta} = \frac{Ndt}{L\sin\theta} \tag{6}$$

is the distance of the center of mass from the pivot. Hence the precession angular velocity of Eq. (7) is

$$\omega_P = \frac{d\phi}{dt} \simeq \frac{Mg\ell}{I_3\omega_3} \tag{7}$$

obtained. In addition to the slow precession motion of a fast top, the polar angle θ of the symmetry axis may also undergo an oscillating motion, nutation.

Rising top. When a spinning top with a blunt peg end is set down on a rough surface, the top usually slips initially. Due to the frictional force at the peg, there is a frictional torque about the center of mass of the top of approximate magnitude $N \approx \mu Mg\ell$, which is perpendicular to the peg for a thin peg. Here μ is the coefficient of friction and ℓ is the distance from the center of mass to the peg. The torque is perpendicular to the symmetry axis, and hence to $\vec{L}$ for a fast top, in the direction of decreasing polar angle θ. Thus the frictional torque causes the top to precess toward the vertical with angular velocity θ given by Eq. (8).

$$\dot{\theta} = -\frac{N}{L} \approx -\frac{\mu Mg\ell}{I_3\omega_3} \tag{8}$$

Boomerang. The return of a boomerang is explained by precession. The cross-blade boomerang consists essentially of two perpendicular fan blades that intersect at the boomerang's center of mass. As the boomerang is thrown in a vertical plane, the blades rotate rapidly about the center of mass with the center of mass moving parallel to the ground. Due to its spin the boomerang has an angular momentum about the center of mass that is perpendicular to its plane.

The aerodynamic force on the airfoils of the blades gives a net force and torque on the boomerang. (The torque comes from the higher forces on the upper blades which are moving faster through the air.) The force accelerates the boomerang perpendicular to the plane of rotation, in the direction of $\vec{L}$. The torque causes the plane of the boomerang to precess about the vertical. The net effect of the force and torque is thus to make the center of mass of the boomerang move in a circle.

Free or fast precession. An axial symmetric body has two of its principal moments of inertia equal, say $I_1 = I_2$; $\vec{n}_3$ is chosen to be the principal axis which is the symmetry axis. Then the angular momentum $\vec{L}$ about the center of mass of the body takes the form Eq. (9), expressed in terms of the body axes $\vec{n}_1$,

$$\vec{L} = I_1(\omega_1\vec{n}_1 + \omega_2\vec{n}_2) + I_3\omega_3\vec{n}_3 \tag{9}$$

$\vec{n}_2$, $\vec{n}_3$ and the body-axis components of the angular velocity ω ($\omega \cdot \vec{n}_1 = \omega_1$, and so forth). The motion of $\vec{n}_3$ is given by Eq. (10), where the first equality is

$$\dot{\vec{n}}_3 = \vec{\omega} \times \vec{n}_3 = \left(\frac{\vec{L}}{I_1}\right) \times \vec{n}_3 \tag{10}$$

by definition of ω, and the second is true because ω and $\vec{L}/I_1$ differ only by a vector parallel to $\vec{n}_3$. If the body is free, $\vec{L}$ is constant, and so Eq. (10) shows that the symmetry axis $\vec{n}_3$ precesses around a fixed direction, namely the direction of $\vec{L}$, at a rate L/I_1. As an example consider a coin thrown in the air, spinning nearly in its plane. The precession of its symmetry axis, that is, the normal to the plane of the coin, is seen as a wobble; the ratio of the rate of the wobble to the rate of spin is $(L/I_1)/(L/I_3) = I_3/I_1 = 2$.

An ordinary top, when first thrown, may exhibit fast precession, similar to what its motion would have been if free, because the effect of gravity and the contact with the ground produces only an additional slow precession. So-called fast precession will not be fast if the spin of the body is not fast. An example is given by a rolling coin in the late stage of its motion when it is "lying down." It exhibits a fast wobble (infinitely fast in the limit of lying flat) which is not a fast precession, being much faster than the spin rate of the coin (as can be seen from the rotation of the emblem on its face), and is instead a precession driven by torque. *See* ANGULAR MOMENTUM; RIGID-BODY DYNAMICS. Vernon D. Barger

Bibliography. A. P. Arya, *Introduction to Classical Mechanics,* 2d ed., 1997; V. Barger and M. Olsson, *Classical Mechanics: A Modern Perspective*, 2d ed., 1995; H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; D. Halliday, R. Resnick, and K. Krane, *Physics*, 5th ed., 2002; K. Symon, *Mechanics*, 3d ed., 1971.

Precession of equinoxes

A slow change in the direction of the axis of rotation of the Earth and also a very small change in the orientation of the Earth's orbit around the Sun. Thus, both the celestial equator (the projection of the terrestrial equator onto the sky) and the ecliptic (the projection of the Earth's orbit onto the sky) move slowly in time and, consequently, so do their intersections, the vernal and autumnal equinoxes. *See* ECLIPTIC; EQUATOR; EQUINOX.

Earth's axis. The change in direction of the Earth's axis of rotation is caused primarily by the Moon and secondarily by the Sun. The Earth is an oblate spheroid, bulging outward around the Equator by about 1 part in 300. Because the Earth's axis of rotation is tilted by about 23.4° to the plane of its orbit about the Sun, this bulge is also tilted. The extra gravitational pull of the Moon and Sun on this bulge produces a revolution, or precession, of the axis of rotation about the perpendicular to the orbit (see **illustration**); one complete revolution requires just under 26,000 years. This has been recognized for over 2000 years and was formerly called lunisolar precession. *See* PRECESSION.

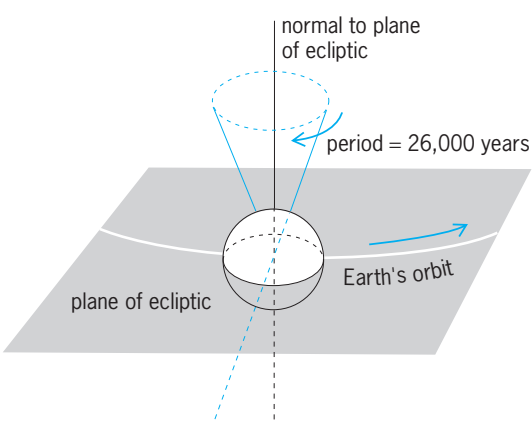

Motion of Earth's axis of rotation.

In addition to uniform revolution, there are very small amplitude periodic oscillations; the primary one has a period of 18.6 years. These oscillations are collectively called nutation. They are caused by the same forces as precession, but are treated observationally and computationally as separate phenomena. *See* NUTATION (ASTRONOMY AND MECHANICS).

Earth's orbit. The slow change in the Earth's orbital plane (ecliptic) is caused by the gravitational perturbations of the other planets in the solar system. The principal effect is to produce an extremely small rotation of the axis of revolution of the Earth around an axis parallel to the total-angular-momentum axis of the solar system. This therefore produces a very small motion of the ecliptic itself and thus an additional small change in the apparent position of the equinox. This precession was formerly called planetary precession.

Effects. Formulas for precession and nutation are used to give the direction of the Earth's spin axis (pole) in space, while the sidereal time formula is used to give the Earth's rotation about the pole. The combination orients the Earth in inertial space and thus determines the positions of observing telescopes, antennas, and so forth. *See* SIDEREAL TIME.

Such information is necessary for telescope pointing; it is crucial for spacecraft navigation. In fact, even the extensive formulations for precession, nutation, and sidereal time are not adequate for the most precise navigation. The orientation of the Earth is not fully predictable due to earthquakes, volcanoes, atmospheric loading, polar cap melting, and so forth. As a result, it is necessary to continually measure the Earth's exact orientation on a daily basis, collect the measurements, and immediately distribute the information worldwide.

In former times, the Earth's pole, equator, and mean orbital path (ecliptic) were extended out into space in order to define the basic celestial coordinate system. Nowadays, the basic coordinate system, though oriented close to the Earth's equator in the year 2000, is actually defined by the International Celestial Reference Frame (ICRF), a specific catalog of radio sources (quasars and so forth). *See* CELESTIAL REFERENCE SYSTEM.

People in the Northern Hemisphere have been lucky over the past few centuries, for they have had a "north star." The north celestial pole, the direction of the Earth's north pole in space, has been close to the star Polaris. It was 2.3° away in A.D. 1700, was 0.7° away in 2000, and will pass Polaris at about 0.5° in 2100. In the future, the north celestial pole will not come within 5° of any bright star (magnitude <3) until it completes its 26,000-year circuit and comes again near to Polaris. For the south pole, it is just the opposite. At present, the south celestial pole is not within 12° of any bright star, but it will

pass within 1° of bright stars during the (approximate) years 6720–6970, 7930–8290, 9070–9430, and 24160–24470. *See* ASTRONOMICAL COORDINATE SYSTEMS; EARTH ROTATION AND ORBITAL MOTION; POLARIS. E. Myles Standish; Robert S. Harrington

Bibliography. K. Lambeck, *The Earth's Variable Rotation*, 1980; H. Moritz and I. I. Mueller, *Earth Rotation*, 1987; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, 1992; F. R. Stphenson, *Historical Eclipses and the Earth's Rotation*, Cambridge University Press, 1997.

Precious stones

The materials found in nature that are used frequently as gemstones, including amber, beryl (emerald and aquamarine), chrysoberyl (cat's-eye and alexandrite), coral, corundum (ruby and sapphire), diamond, feldspar (moonstone and amazonite), garnet (almandite, demantoid and pyrope), jade (jadeite and nephrite), jet, lapis lazuli, malachite, opal, pearl, peridot, quartz (amethyst, citrine, and agate), spinel, spodumene (kunzite), topaz, tourmaline, turquois, and zircon.

The terms precious and semiprecious have been used to differentiate between gemstones on a basis of relative value. Because there is a continuous gradation of values from materials sold by the pound to those valued at many thousands of dollars per carat, and because the same mineral may furnish both, a division is essentially meaningless. All but a few of the gems listed cost many dollars per carat in fine quality. Thus, almost all gemstones merit the use of the term precious. *See* GEM. Richard T. Liddicoat, Jr.

Precipitation (chemistry)

The process of producing a separable solid phase within a liquid medium. In a broad sense, precipitation represents the formation of a new condensed phase, although other terms are often used to describe the process. Thus (1) a vapor or gas condenses to liquid droplets, or more specifically as in meteorology, water vapor in the atmosphere precipitates to form rain, snow, or ice; (2) a substance in the liquid state freezes or solidifies; (3) a dissolved component crystallizes from a supersaturated solution; (4) a new solid phase gradually precipitates within a solid alloy as the result of a slow, inner chemical reaction; or (5) a metal electrodeposits upon the passage of an electrical current through a solution.

In analytical chemistry, precipitation is widely used to effect the separation of a solid phase in an aqueous solution. For example, the addition of a water solution of silver nitrate to a water solution of sodium chloride results in the formation of insoluble silver chloride. Quite often, one of the components in the solution is thus virtually completely separated in a relatively pure form. It can then be isolated from the solution phase by filtration or centrifugation, and the substance determined by weighing. This procedure is known as gravimetric analysis. Precipitation may also be used merely to effect partial or complete separation of a substance for purposes other than that of gravimetric analysis. Such purposes might involve either the isolation of a relatively pure substance or the removal of undesirable components of the solution. *See* GRAVIMETRIC ANALYSIS.

Solubility product constant. The extent to which a component can be separated from solution can be determined from the solubility product constant obtained by determining the quantity of dissolved substance present in a known amount of saturated solution. This value is known as the solubility. The solubility can be drastically altered merely by adding to the solution any of the ions that make up the precipitate, for example, by adding varying quantities of either silver nitrate or sodium chloride to a saturated solution of silver chloride. Although solubility can be altered over a wide range, the solubility product itself remains practically constant over this same range.

The solubility product constant can be used to ascertain the quantity of dissolved component remaining unprecipitated in the presence of known concentrations of the ions common to the precipitate. By proper adjustment of the concentration of the added common ion, it is possible to reduce the quantity of dissolved component to a negligible value, although never to zero. It is an extremely important criterion of a method of gravimetric analysis that the quantity of unprecipitated component be negligible, particularly in comparison with the quantity of precipitate formed. The analytical chemist uses the word quantitative to describe such a chemical reaction. *See* SOLUBILITY PRODUCT CONSTANT.

Impurities. The purity of the precipitated solid phase is of major concern, both for the preparation of a desired chemical compound and for a quantitative method of gravimetric analysis. It is not possible for a precipitate to be formed as an absolutely pure compound by chemical reaction within the solution phase. Other soluble substances present in the solution, such as the ions not involved in the structure of the precipitate, tend to accompany the solid phase in varying amounts. This phenomenon is known as coprecipitation. The fraction of the total quantity of such foreign ions coprecipitating may be quite small. Although this fraction depends on experimental variables, it is highly dependent upon the relationship between the solubility characteristics of the desired chemical precipitate and the foreign substance. As a specific example, partial precipitation of iodide (as silver iodide) using silver nitrate would coprecipitate (as silver chloride) only a small fraction of any chloride present, whereas it would coprecipitate (as silver bromide) a larger fraction of any bromide present. The iodide is more insoluble than the bromide,

which is more insoluble than the chloride. Knowledge of the relative solubility characteristics of the chemical species present is thus extremely desirable to the chemist, who needs to know whether foreign substances are being collected with the solid or are being left in solution.

A precipitated phase may incorporate foreign ions within its structure in several ways. Best understood of these is isomorphous mixed-crystal formation. Radium and barium sulfates form isomorphous mixed crystals because the two compounds have the same crystal structure and the ionic radii of radium and barium are not greatly different. Thus, radium and barium ions are interchangeable within the crystal lattice to a considerable degree. Silver bromide and silver chloride also form isomorphous mixed crystals. Because of the ease with which interchange can take place within the crystal lattice, isomorphous mixed crystal systems should be avoided if a good separation of the two different ionic species is desired. On the other hand, the property of isomorphism may also be put to good use in concentrating minute traces. For example, barium sulfate precipitated in the presence of minute traces of radium carries with it almost all of the radium. Such procedures are frequently used in the collection of minute traces of radioactive species.

An ion present at high dilution may sometimes be incorporated, apparently by mixed crystal formation, even though such formation would not be predicted on the basis of crystallography and ionic radii. An example of this is the coprecipitation of traces of lead with potassium chloride. This phenomenon is known as anomalous mixed crystal formation, or isodimorphism.

When foreign-ion incorporation cannot be ascribed to isomorphism, coprecipitation may occur by adsorption. Residual charges at the surface of a precipitate attract charged ions in the solution. The adsorption process results in a greater concentration of foreign ions near the precipitate surface than exists in the main body of the solution. Adsorbed foreign ions may remain quite firmly attached to the solid. In fact, they may be covered as succeeding layers of the crystal are deposited and cause imperfections within it. This phenomenon is known as occlusion, although it arises as a result of adsorption. The term occlusion should be distinguished from inclusion, which refers to the mechanical trapping of pockets of solution (and solutes in it) within the precipitate.

Sometimes one substance in a mixture precipitates rapidly, but a second foreign substance then precipitates slowly as a second solid phase. This is not generally considered to be coprecipitation, but is referred to as postprecipitation. The postprecipitation of zinc sulfide with copper sulfide is a typical example.

Methods for reducing contamination. In an effort to reduce contamination by foreign ions, the chemist resorts to various techniques. Precipitation from dilute solution is often effective. Heating the reaction mixture, that is, digesting the solid in contact with the liquid phase, speeds recrystallization processes by which incorporated foreign ions may be returned to the solution phase. Precipitation from homogeneous solution, a technique in which the desired precipitating reagent is formed internally within the solution by chemical synthesis, results in the slow formation of large crystals of small surface area and hence lessens coprecipitation.

If all these methods fail to reduce adequately the quantity of foreign ions incorporated in the solid phase, then reprecipitation is applied. The precipitate is dissolved and reprecipitated by the previous procedure. As in the initial precipitation, most of the foreign ions remain in solution. The process of reprecipitation must be repeated until the quantity of foreign ions present in the precipitate can be disregarded. *See* ADSORPTION; CHEMICAL SEPARATION TECHNIQUES; CRYSTALLIZATION; NUCLEATION.

Precipitation from homogeneous solution. In this technique the precipitant is generated in place within the solution phase instead of being added directly, as in the conventional manner.

The substances required to effect precipitation from homogeneous solution can be generated within a solution phase in a variety of ways. For example, if a source of hydroxyl ions is needed to precipitate metallic ions, urea in solution can be hydrolyzed to produce ammonium hydroxide to act as the source. Other anions needed as precipitants can often be formed by the hydrolysis of appropriate esters. Thus, dimethyloxalate can serve as a source of oxalate ions to precipitate thorium, and dimethyl sulfate can be the source of sulfate ions to precipitate lead. There also are methods for the generation of cations. For example, silver ions can be slowly released from a complex ion such as $Ag(NH_3)_2^+$ to precipitate chloride ions. Various investigators have developed ingenious methods for producing the necessary precipitant to effect precipitation from homogeneous solution.

The formation of larger, more perfect, and purer crystals occurs almost without exception when precipitation from homogeneous solution is used in place of direct addition of the precipitant.

The **illustration** shows the difference in appearance of nickel dimethylglyoximate precipitated by two different methods. The precipitate shown in illus. *b* was formed by the direct addition of dimethylglyoxime to a solution containing nickel ions, as in reaction (1). The more perfect crystals shown in

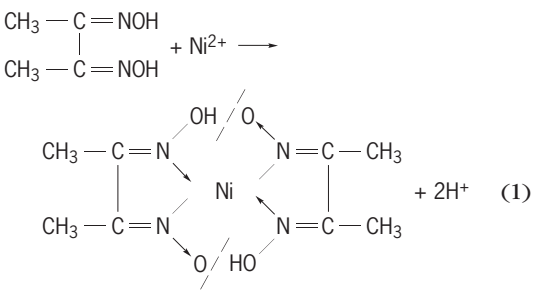

(1)

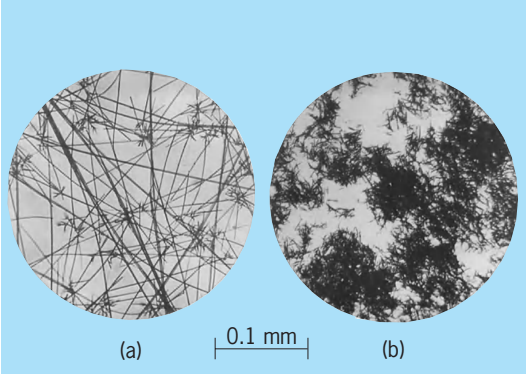

Photomicrographs of nickel dimethylglyoximate precipitated by two methods. (*a*) Precipitation from homogeneous solution. (*b*) Direct addition to solution. (*From L. Gordon and E. D. Salesin, Precipitation from homogeneous solution, J. Chem. Educ., 38(1):16, 1961*)

illus. *a* were precipitated from homogeneous solution by synthesizing the necessary dimethylglyoxime from biacetyl and hydroxylamine by reaction (2).

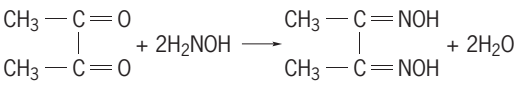

Because the rate at which a precipitant is generated can be controlled very closely, preciptation from homogeneous solution is used as a research technique in studies of the mechanisms of nucleation, precipitation, and coprecipitation. The technique has many applications in gravimetric analysis, in the production of pure chemicals, and in the control of particle size of crystals. Louis Gordon; Royce W. Murray

Bibliography. Center for Occupational Research and Development Staff, *Gravimetric and Volumetric Analysis*, 1985; A. G. Walton, *The Formation and Properties of Precipitates*, 1967, reprint 1979.

Precipitation (meteorology)

The fallout of waterdrops or frozen particles from the atmosphere. Liquid types are rain or drizzle, and frozen types are snow, hail, small hail, ice pellets (also called ice grains; in the United States, sleet), snow pellets (graupel, soft hail), snow grains, ice needles, and ice crystals. In England sleet is defined as a mixture of rain and snow, or melting snow. Deposits of dew, frost, or rime, and moisture collected from fog are occasionally also classed as precipitation.

All precipitation types are called hydrometeors, of which additional forms are clouds, fog, wet haze, mist, blowing snow, and spray. Whenever rain or drizzle freezes on contact with the ground to form a solid coating of ice, it is called freezing rain, freezing drizzle, or glazed frost; it is also called an ice storm or a glaze storm, and sometimes is popularly known as silver thaw or erroneously as a sleet storm. *See* CLOUD.

Most precipitation particles carry an electrostatic charge, either positive or negative, but the origin of the charge and its relation to other problems of atmospheric electricity are not completely understood. *See* ATMOSPHERIC ELECTRICITY; THUNDERSTORM.

Rain, snow, or ice pellets may fall steadily or in showers. Steady precipitation may be intermittent though lacking sudden bursts of intensity. Hail, small hail, and snow pellets occur only in showers; drizzle, snow grains, and ice crystals occur as steady precipitation. Showers originate from instability clouds of the cumulus family, whereas steady precipitation originates from stratiform clouds.

The amount of precipitation, often referred to as precipitation or simply as rainfall, is measured in a collection gage. It is the actual depth of liquid water which has fallen on the ground, after frozen forms have been melted, and is recorded in millimeters (mm) or hundredths of inches. A separate measurement is made of the depth of unmelted snow, hail, or other frozen forms. *See* PRECIPITATION MEASUREMENT.

Liquid types. Rain and drizzle are somewhat arbitrarily differentiated. Raindrops are generally over 0.02 in. (0.5 mm) in diameter and differ from drizzle mainly in having larger sizes. Drizzle may also be distinguished from rain by the meteorological conditions of its formation; drizzle falls usually from fog or thick stratus clouds, whereas rain comes from clouds of cumulus type or clouds extending well above the freezing level. Raindrops are rarely larger than 0.2 in. (5 mm) in diameter because larger sizes tend to break up.

The small raindrops are approximately spherical, but large falling drops are flattened (**Fig. 1**), especially on the bottom side. The following speeds of falling drops, in meters per second, were measured in still air by R. Gunn (drop diameters are in millimeters):

Speed	0.27	2.06	4.03	6.49	8.06	8.83
Diameter	0.1	0.5	1.0	2.0	3.0	4.0

Frozen precipitation. Snowflakes are branched, six-point star crystals, irregular forms such as matted ice needles, or combinations of both; often they are coated by rime. The speed of falling flakes is variable; it is greater when the flakes are rimed and is generally 2–4 mi/h (1–2 m/s). *See* SNOW.

Hail, mostly seen in thunderstorms, forms with the aid of warm, moist air when clouds build to great heights. Hailstones usually have concentric layers of rime and hard ice, typically 0.4 in. (1 cm) in diameter but ranging up to more than 4 in. (10 cm). *See* HAIL.

Ice pellets are of hard ice, either clear or opaque, and may be spherical or irregular. They are formed by the freezing of drops of rain or drizzle.

Snow pellets are of soft, opaque ice, irregularly shaped, sometimes with scalloped edges and often studded with a few oblong or branched crystals. Typical diameters are 0.08–0.2 in. (2–5 mm). They fall

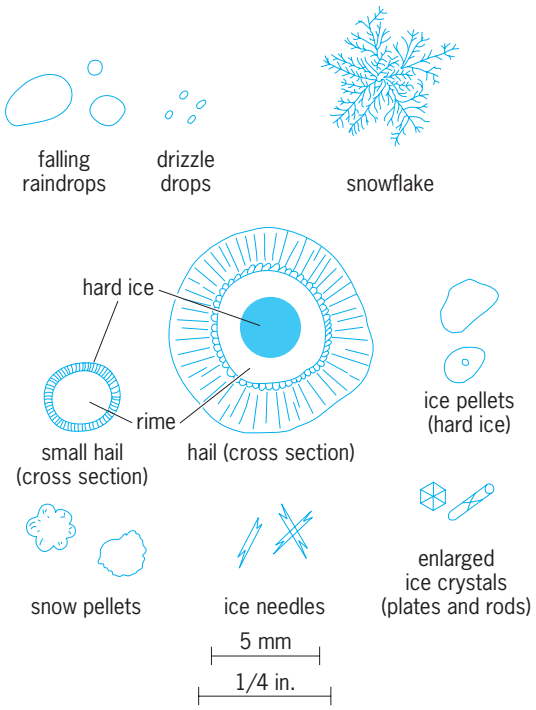

Fig. 1. Various forms of precipitation.

usually in showers, sometimes even when ground temperatures are a few degrees above freezing. U. Nakaya found their specific gravity to be about 0.12. Snow grains are similar but smaller and flatter, with diameters around 1 mm or less; they fall generally in small amounts from stratus clouds or fog.

Ice needles are narrow, pointed crystals, roughly 0.04–0.012 in. (1–3 mm) long and about 0.01 in. ($^1/_4$ mm) in diameter, which fall singly or in clusters, mostly at temperatures near or a little below freezing.

Ice crystals are small rods or plates. They fall in cold, stable air that is often without clouds, and glisten in light as they settle to the ground.

Geographical distribution. Precipitation is part of the hydrologic cycle, the continuous interchange of water between sea, land, and atmosphere, but its distribution over the Earth is uneven (**Fig. 2**).

Some world-record rainfall amounts are, for 1 min, 1.23 in., Unionville, Maryland, 1956; 1 h, 12.00 in. (305 mm), Holt, Missouri, 1947; 12 h, 52.76 in. (1340 mm), Belouve, La Réunion, 1964; 24 h, 73.62 in. (1870 mm), Cilaos, La Réunion, 1952; 2 days, 98.42 in. (2500 mm), Cilaos, La Réunion, 1952; 31 days, 366.14 in. (9300 mm), Cherrapunji, India, 1861; and 1 year, 1041.78 in. (26,461 mm), 2 years, 1605.05 in. (40,768 mm), both Cherrapunji, India, 1860–1861.

Precipitation and weather. Precipitation occurs most often in cyclones or tropical disturbances. In weather forecasting, several synoptic (weather map) types of precipitation are recognized: warm front, warm moist air rising over a wedge of cold air; cold front, cold air undercutting and lifting warmer air; convective, caused by local updrafts of moist air; convergent, general lifting of air caused by convergence at low levels; and orographic, moist air forced upward on mountain slopes. Any type may act alone; however, convection or convergence may occur with other types. Lake snow falls when air, at first extremely cold, blows off a bay or off lakes such as the Great Lakes. Minor causes of precipitation

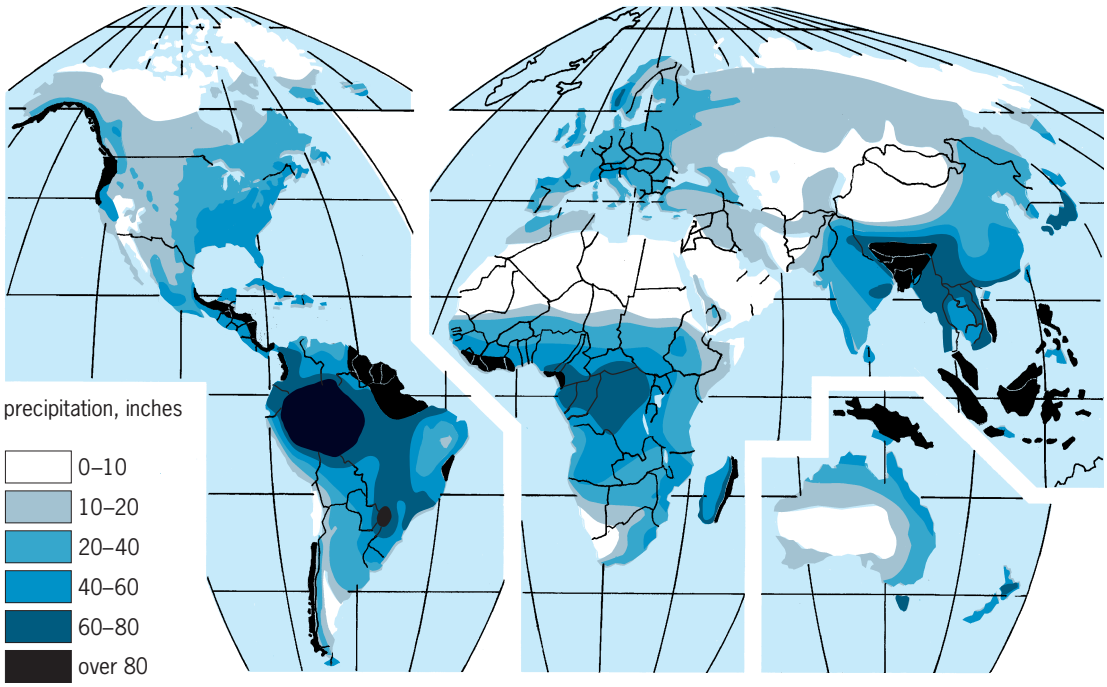

Fig. 2. Map showing the distribution of average annual precipitation over the land areas of the world. 1 in. = 25 mm. (*After N. A. Bengtson and W. Van Royen, Fundamentals of Economic Geography, 4th ed., Prentice-Hall, 1956*)

are turbulence, contact cooling (air moving over a colder surface), and nighttime radiation from cloud tops, but the amounts of precipitation from these sources are small.

Condensation in the atmosphere. Basically, this requires that air be cooled to and below its dew point. Then minute waterdroplets form by condensation, or ice crystals form by sublimation. These droplets, or ice crystals, are cloud particles from which, by a growth mechanism, the larger precipitation particles may form. The only cooling process sufficient to produce appreciable precipitation is adiabatic expansion, which occurs in air rising toward lower pressures. The rate of condensation in rising saturated air is greater the warmer the air and is directly proportional to the speed of ascent. For discussions of condensation nuclei, formation of precipitation in clouds, and artificial stimulation of precipitation *see* CLOUD PHYSICS; WEATHER MODIFICATION.

For discussions of other topics related to precipitation *see* ATMOSPHERIC GENERAL CIRCULATION; DEW; DEW POINT; FOG; HUMIDITY; HYDROLOGY; HYDROMETEOROLOGY; METEOROLOGY; RAIN SHADOW; TROPICAL METEOROLOGY; VAPOR PRESSURE.

J. R. Fulks

Bibliography. W. R. Cotton and R. A. Anthes, *Storm and Cloud Dynamics*, 1989; F. K. Lutgens and E. J. Tarbuck, *The Atmosphere: An Introduction to Meteorology*, 7th ed., 1997; B. J. Mason, *Clouds, Rain, and Rainmaking*, 2d ed., 1976; A. Miller and J. C. Thompson, *Elements of Meteorology*, 6th ed., 1999; J. M. Moran and M. D. Moran, *Meteorology: The Atmosphere and the Science of Weather*, 5th ed., 1996; H. R. Pruppacher and J. D. Klett, *Microphysics of Clouds and Precipitation*, 1980; C. Riegel, *Fundamentals of Atmospheric Dynamics and Thermodynamics*, 1993.

Precipitation measurement

Instruments used to measure the amount of rain or snow that falls on a level surface. Such measurements are made with instruments known as precipitation gages. A precipitation gage can be as simple as an open container on the ground to collect rain, snow, and hail; it is usually more complex, however, because of the need to avoid wind effects, enhance accuracy and resolution, and make a measurement representative of a large area. Precipitation is measured as the depth to which a flat horizontal surface would have been covered per unit time if no water were lost by runoff, evaporation, or percolation. Depth is expressed in inches or millimeters, typically per day. The unit of time is often understood and not stated explicitly. Snow and hail are converted to equivalent depth of liquid water.

Wind effects. In strong winds the gage catch is reduced by turbulence around the opening. The ideal exposure for a precipitation gage is an area that is free from obstructions that would create large eddies that deflect the flow, and where the winds are light to allow rain and snow to fall vertically. These conditions can be approximated reasonably well for large raindrops but are nearly impossible to achieve for snow. A gage in the lee of an obstruction could over- or undercollect precipitation, depending upon the obstruction, the wind speed, and the particle size. The usual exposure is to set the gage orifice a few feet above ground and well away from obstructions. Sometimes a special wind screen placed around the gage creates turbulence in a form tending to minimize systematic wind bias. Precipitation gages should never be placed around or on a building.

Fig. 1. Automatic gages. (a) Tipping-bucket rain gage. (*b*) Weighing rain gage. (*Belfort Instrument*)

Accuracy and resolution. The funnel of a manual gage directs the precipitation into a tube whose cross-sectional area is one-tenth that of the gage orifice. The depth of precipitation is exaggerated and can be measured more readily with a special dip stock having an expanded scale. Automatic gages use a weighing mechanism or a tipping bucket as shown in **Fig. 1**. The tipping bucket fills and tips for each 0.01 in. (0.25 mm). The greatest source of error, excluding representativeness of the sample, is exposure, that is, wind and obstruction effects that can reduce the gage catch 5–80%. Exposure problems are compounded with snow, which may block the gage orifice or freeze the mechanism. Gage accuracy specifications ignore problems of exposure and representativeness.

Representativeness. Rain and snow showers are local phenomena; a heavy rain shower may completely miss a nearby precipitation gage. The estimate of precipitation over a large area may be improved by building a dense network. An alternative is to use a radar to supplement ground-based gages. A weather radar can detect precipitation over a large area, typically over a circular area having a 190-mi (300-km) radius. The optimum combination is a weather radar operating in conjunction with a network of gages that report in real time. This would allow prompt reporting of representative precipitation over a large area, which is essential in

predicting flash floods. *See* METEOROLOGICAL INSTRUMENTATION; PRECIPITATION (METEOROLOGY); SNOW SURVEYING. Fred V. Brock

Radar. Accurate quantitative precipitation measurement is probably the most important weather radar application. It is extremely valuable for hydrological applications such as watershed management and flash flood warnings. Radar can make rapid and spatially contiguous measurements over vast areas of a watershed at relatively low cost. However, radar detection of two equal-diameter raindrops will return less echo power than one raindrop containing the same amount of water. Thus, even though echo power is correlated with precipitation amounts, there is no unique relation between the two; other data are needed to adjust this relation. For example, different precipitation types (such as stratiform rain, thunderstorm rain, and snow) require different relations between the precipitation rate Q and reflectivity factor Z, a radar-derived parameter calculated from measurements of echo power. The relation between rainfall rate $R \equiv Q$ (mm h^{-1}) and Z (mm^6 m^{-3}), expressed as $Z = 200\ R^{1.6}$, has proven to be quite useful for stratiform rain, where Z is proportional to the rain's backscatter cross section per unit volume. Other relations are used to relate radar measurements of Z to the fall rate of other precipitation types. Rain gages, sparsely deployed over a watershed, can also be used to tune these relations to improve the quantitative measurement of precipitation. *See* HYDROMETEOROLOGY; METEOROLOGICAL RADAR; PRECIPITATION (METEOROLOGY); RADAR METEOROLOGY; STORM DETECTION.

Nevertheless, Q, Z relations are subject to many sources of error, including (1) radar calibration, (2) attenuation of the echoes, (3) blockage of the radar beam (by buildings, trees), (4) echoes from ground objects, (5) variations in the pedicle size distribution, and (6) beam filling (the extent that the radar beam is uniformly filled with precipitation). Although considerable effort has been expended to improve radar estimates of precipitation, the basic way in which precipitation is measured has not changed. All of these error sources either go away or can be mitigated by the use of polarimetric radar.

Polarimetric weather radar transmits and receives power in two states of wave polarization (horizontal and vertical), whereas weather radars, operated by weather services into the twenty-first century, transmit and receive waves only in a single polarization state (for example, horizontal). The relation between the specific differential phase K_{DP} (the differential propagation phase shift per unit length between vertically and horizontally polarized waves) and rain rate measured with a polarimetric radar is relatively insensitive to variations in drop size distribution and can also form the basis of rain measurement using an R,K_{DP} relation.

The addition of polarimetric capability to operational radars will provide a revolutionary approach to the radar measurement of rainfall. Polarimetric radars may significantly increase the accuracy of moderate-to-heavy rainfall rate measurements, lead to issuance of more timely and accurate flash flood warnings, and provide other improved hydrological products to agricultural and commercial enterprises.

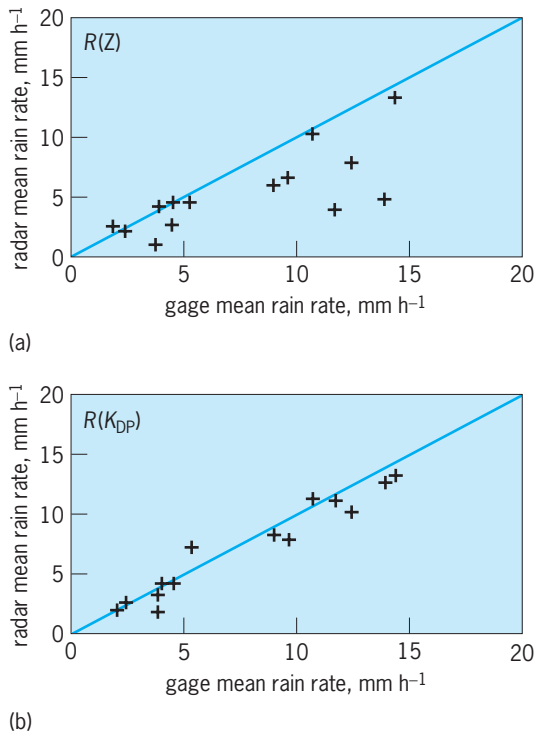

Fig. 2. Comparison of rainfall rates measured with gages and radar using (*a*) the *R*, *Z* relation and (*b*) the *R*,K_{DP} relation. Each data point (+) is the rainfall rate for a single storm.

An R, K_{DP} relation has been tested on data collected with a research polarimetric weather radar operated by the National Severe Storms Laboratory in Norman, Oklahoma. Comparisons were made with rainfall $R(G)$ measured by a dense network of 42 rain gages for 15 rain events; $R(G)$ is the rain gage estimate of the accumulated rainfall over the area of the gages. The rainfall rates were averaged over the duration of the events lasting from about 1 to 4 hours. Rainfall measured using the R,K_{DP} relation showed significant accuracy improvement over that measured using the R,Z relation (**Fig. 2**). Richard J. Doviak

Bibliography. F. Dobson, L. Hasse, and R. Davis (eds.), *Air-Sea Interaction: Instruments and Methods*, 1980; R. J. Doviak and D. S. Zrnić, *Doppler Radar and Weather Observations*, 2d ed., 1993; D. H. Lenschou (ed.), *Probing the Atmospheric Boundary Layer*, American Meteorological Society, 1986; Meteorological Office, *Handbook of Meteorological Instruments*, vol. 5: *Measurement of Precipitation and Evaporation*, 2d ed., 1981; M. Sachidananda and D. S. Zrnić, Differential propagation phase shift and rainfall rate estimation, *Radio Sci.*, 21:235–247, 1986; J. T. Snow and S. B. Harley, Basic meteorological observations for schools: Rainfall, *Bull. Amer. Meteorol. Soc.*, 69(5):497–507, 1988; D. S. Zrnić and A. Ryzhkov, Advantages of rain measurements using specific differential phase., *J. Atm. Oceanic Technol.*, 13(2):454–464, 1996.

Precipitin

The visible result of the chemical interaction of antigen and antibody. Such precipitate was first described in 1897, and it was later noted that the optimal ratio of antigen to antibody concentrations resulted in the visible precipitate. Not all antibodies will result in precipitation, yet they may participate in agglutination reactions or add onto particulate antigens, and evidence for their occurrence together with precipitating antibody can be obtained for most sera. Precipitins may be noted qualitatively or be quantified by noting the end-point dilution (titer) of serum required to give a precipitate at the threshold of visiblity, or the amount of antibody may be determined in milligrams or micrograms by analysis of the precipitate with correction for the antigen contained therein. *See* ANTIBODY; ANTIGEN; PRECIPITIN TEST. Donald Raum

Bibliography. J. M. Cruse and R. E. Lewis, Jr. (eds.), *The Year in Immunology 1989-90*, 1990; F. Dixon (ed.), *Advances in Immunology*, vol. 48, 1990; J. W. Kimball, *Introduction to Immunology*, 3d ed., 1990; D. M. Weir, *Immunology: An Outline for Students of Medicine and Biology*, 6th ed., 1987.

Precipitin test

A test to measure a specific reaction between antigen and antibody which results in a visible precipitate. *See* PRECIPITIN.

Qualitative tests. Two types of qualitative tests are commonly used. The first is the ring test. A solution of antigen is laid over a solution of antibody or antiserum so that a sharp interface is formed. A precipitate at the interface is a positive reaction and occurs as antigen and antibody diffuse toward one another, creating a zone in which the ratio of their concentrations is optimal for precipitation. As little as 1.0 microgram of protein may be detected. This is an easy, rapid test. The second qualitative test involves mixing solutions of antigen and antibody. This is done by varying either the concentration of antibody or the concentration of antigen. When the amount of antigen is varied, three zones of precipitation are noted: antigen excess, equivalence, and antibody excess.

Quantitative determination. The quantitative determination of precipitating antibodies depends upon the two-stage nature of antigen-antibody interaction. The first stage of antigen-antibody combination reaches completion within a few minutes, while the second stage, in which aggregation of complexes occurs, may take days. To standardize these tests, it is necessary to consider time, temperature, and hydrogen ion and electrolyte concentration. Precipitates are digested and analyzed for nitrogen content, and the amount of antibody calculated. Since complement solubilizes antigen-antibody complexes, spuriously low values will be obtained if the test is performed in the presence of fresh serum. *See* COMPLEMENT.

Modifications. A number of modifications of the precipitin test allow one to test multiple samples and to determine whether the antigen-antibody system is homogeneous.

In the Oudin method, antigen and antibody are each held in an agar matrix in a test tube and allowed to diffuse toward one another. The Oudin technique can be used as a measure of antigen concentration, and can be read as the position of the precipitin band after prolonged incubation or as the change in position early during diffusion.

In the Ouchterlony method, antigen and antibody are placed in multiple wells cut into agar on a flat slide and allowed to diffuse toward one another. If the system is homogeneous, one line results; if there are multiple components, there will be multiple lines when two samples are placed in juxtaposition to one another and an antiserum. The identity of the two antigens is indicated by a clear corner on the intersecting lines; a spur indicates partial identity; and crossing of the intersection indicates nonidentity.

The Mancini, or single-radial-immunodiffusion, method for quantitating soluble proteins involves placing the solution to be measured into a well cut into an agar or agarose gel containing antiserum and measuring the area subtended by the precipitin ring and relating this to known standards. To improve resolution, the antigen can be made to migrate through a gel matrix (agarose) by an electric current (electrophoresis) or in the presence of a pH gradient (isoelectric focusing). Then a trench can be cut in the gel, and specific antibody placed in the trough and allowed to diffuse into the gel. This procedure can be reversed to locate immunoglobulins.

In Laurell's rocket technique, antibody may be dissolved in the gel, and the antigen sample then driven through the gel by an electric field. The resulting height of the precipitin peak will bear a specific relation to the quantity of antigen in the sample, allowing accurate estimation of antigen concentration in many samples. *See* IMMUNOASSAY; IMMUNOCHEMISTRY; IMMUNOLOGY. Donald Raum

Bibliography. F. Dixon (ed.), *Advances in Immunology*, vol. 48, 1990; W. E. Paul (ed.), *Annual Review of Immunology*, vol. 5, 1987; D. M. Weir and J. Stewart, *Immunology*, 8th ed., 1997.

Precision agriculture

The application of technologies and agronomic principles to manage spatial and temporal variability associated with all aspects of agricultural production for the purpose of improving crop performance and environmental quality. The intent of precision agriculture is to match agricultural inputs and practices to localized conditions within a field (site-specific management) and to improve the accuracy of their application. The finer-scale management of precision agriculture is in contrast to whole-field or whole-farm management strategies, where management decisions and practices are uniformly

applied throughout a field or farmstead.

Successful implementation of precision agriculture requires three basic steps. First, farmers must obtain accurate maps of the spatial variability of factors (soils, plants, and pests) that determine crop yield and quality and/or factors that cause environmental degradation. Second, once known, variability can be managed using site-specific management recommendations and accurate input control technologies. Third, precision agriculture requires an evaluation component to understand the economic, environmental, and social impacts on the farm and adjacent ecosystems and to provide feedback on cropping system performance.

Precision agriculture is technology-enabled, information-based, and decision-focused, because it relies on an increasing level of detail in information acquired with technology to improve decision making in crop production. Consequently, precision agriculture will evolve as technology, information management, and decision tools emerge in this era of rapid technological advancement. *See* AGRICULTURAL SOIL AND CROP PRACTICES; AGRONOMY; DECISION SUPPORT SYSTEM; INFORMATION SYSTEMS ENGINEERING.

Enabling technologies. The major technologies collectively enabling precision agriculture include computers, geographic information systems (GIS), global positioning systems (GPS), sensors, and application control.

Computers, in terms of both speed and storage, paved the way for precision agriculture; and computer enhancements will drive significant technological development to enable precision agriculture for some time.

A geographic information system is an organized collection of computer hardware, software, geographic data, and personnel designed to efficiently capture, store, update, manipulate, analyze, and display all forms of geographically referenced information. It has been referred to as the brain of a precision farming system. A geographic information system is important because the value of precision agriculture is derived only when farmers use detailed information on spatial variability in their management decisions. The payoff to farmers is when these decisions increase profitability, benefit the environment, or provide some other value to the farm.

Location control is essential to precision agriculture for assessing spatial variability and for site-specific application control, and a global positioning system fulfills those needs for precision agriculture. The United States NAVSTAR (Navigation System with Time and Ranging) and the Russian GLONASS (Global Navigation Satellite System) global positioning systems have been fully operational since the mid-1990s; together they provide highly reliable and accurate positioning data. A differential correction to the GPS signal is usually required to achieve position accuracies needed in production agriculture. Differential GPS (DGPS) involves the transmission of a differential correction, that is, the difference between actual and predicted position at the base GPS receiver, to rover GPS receivers, which then apply the corrections to received GPS signals to solve for a more accurate position.

Sensors are devices that transmit an impulse in response to a physical stimulus such as heat, light, magnetism, motion, pressure, or sound. With computers to record the sensor impulse, a global positioning system to measure position, and a geographic information system to map and analyze the sensor data, any sensor output can be mapped at very fine scales. Sensors are critical because they have fixed costs, can sample at very small scales of space and time, and facilitate repeated measures. Sensors have been developed to measure machinery, soil, plants, pests, atmospheric properties, and water by sensing motion, sound, pressure, strain, heat, light, and magnetism, and relating these to properties such as reflectance, resistance, absorbance, capacitance, and conductance.

For precision agriculture, control must be achieved in space and time for varying single or multiple agronomic inputs at different rates, at varying soil depths, and in an accurate and location specific manner within fields. Control systems are available at varying degrees of precision for variable seeding, metering granular fertilizers and pesticides, changing crop varieties on-the-go, anhydrous application of nitrogen, spraying, irrigation, manure application, and various methods of tillage. *See* EXPERT SYSTEMS; GEOGRAPHIC INFORMATION SYSTEMS; SATELLITE NAVIGATION SYSTEMS.

Mapping. Mapping is a fundamental tool for understanding and managing variability, and three types of maps are useful in precision agriculture. Condition maps relate to assessing variability and represent the spatial state of a condition of interest, for example, soil pH or pest distribution. Management maps relate to managing variability and represent the prescription of inputs needed to manage a particular condition site-specifically, for example, application of lime or a pesticide. Management maps are used in application control systems to regulate application rates and products and to form the basis for variable-rate technology. Performance maps are related to evaluation and record either inputs (such as fertilizers, pesticides, seeds, and energy) or outputs (crop yield, quality) and include derivatives of performance maps, such as profit maps (outputs-inputs). Combinations of condition, management, and performance maps can be used to create other maps in each category, and a geographic information system is an excellent tool for this purpose.

Yield mapping serves all three map functions. It requires a yield-monitoring sensor, a measure of the area of harvest, some measures of crop quality, and the location for each sensed measurement. Sensors also are available to measure crop quality, including moisture, density, and protein content, with more sensors under development. A yield monitor measures crop yield by sensing crop flow. A crop-flow sensor measures either the volume or the mass of the harvested portion of a crop using a variety of engineering principles, including light interception,

radiation absorption, measurement of impact force, and directly weighing the crop. Yield sensors are available or under development for grain crops (such as corn, soybeans, and wheat), potatoes, sugarbeets, tomatoes, forages, and cotton. *See* CARTOGRAPHY.

Assessing variability. The basis for precision agriculture is found in the notion that the processes and properties that regulate crop performance and yield vary in space. For many variables, particularly pests but also soil variables such as nitrogen and water, the spatial variability can vary over short time periods. Adequately quantifying the variability of those processes and properties and determining when and where different combinations are responsible for the spatial and temporal variation in crop yield or environmental contamination is the challenge facing precision agriculture. Therefore, assessing variability is the first and most important step in precision agriculture, because farmers cannot manage what they do not know. Techniques for assessing spatial variability are readily available and have been applied extensively in precision agriculture, particularly geostatistics. Geostatistics is a major tool in the assessment of spatial variability, and a number of software packages are available. Commercial mapping and analysis software for use on the farm or in agribusiness offer a variety of interpolation and mapping routines of varying difficulty and quality. The most important issue for precision agriculture is that variability within a field is spatially structured (nonrandom) and conditions maps are accurate. *See* SOIL.

Managing variability. Once variation is adequately assessed, farmers must match agronomic inputs to known conditions using management recommendations that are site-specific, and use accurate application control equipment. Precision agriculture has operated for some time on the assumption that best management practices developed over decades of agronomic research are applicable at any scale of management. These assumptions will be tested as knowledge of spatial variability and its management evolves. While many aspects of production agriculture could respond to precision management, each aspect differs in the type and amount of variability and in the ways in which precision management might be applied. Major aspects of crop production include nutrient management, crop management, pest management, and water management.

Precision nutrient management. Significant opportunities exist for precision management of soil fertility, because the ability of soil to supply nutrients to plants and the plant demand vary in space and time, and nutrient losses through leaching, erosion, and runoff also vary temporally and spatially. One of the first applications of precision agriculture involved the management of fertilizers and lime according to fertilizer application maps, a practice called variable-rate technology (VRT). Soils are sampled within a field using various grid or zone sampling schemes, and are analyzed for fertility using established soil testing and analysis procedures. Fertilizer or lime recommendation maps are created by making recommendations for each sample location, interpolating values for unsampled locations within the field using various techniques (for example, geostatistics, inverse distance), and mapping the results. Fertilizers are applied using equipment capable of varying application rates based on a digitized version of the fertilizer application map as input to a computer-controlled application system. In some cases, sensors are available to assess a particular nutrient, and fertilizer applications can be made in real time in response to the sensed value as the applicator moves across the field. Another approach bases fertilizer applications on maps derived by monitoring the nutrient status of a crop. For nitrogen management, for example, the idea is to map plant concentrations of nitrogen by monitoring plant or canopy reflectance of light as a measure of plant nitrogen content, derive a nitrogen fertilizer application map using established relationships between reflectance and nitrogen content, and fertilize the crop using VRT equipment. Results with VRT and other precision nutrient management studies have been mixed. Success here relies to a large extent on the accuracy of the nutrient map, the appropriateness of the fertilizer recommendation, application control, and the extent to which nutrients limit crop yield or quality in a given growing season. *See* FERTILIZER; FERTILIZING; SOIL FERTILITY.

Precision crop management. Precision crop management could be achieved by varying species or cultivars, and/or by manipulating planting geometry (populations, row spacing, seeding depth) according to variation in environments within fields. The success of precision crop management depends on whether species or cultivars are available that are adapted to specific environments. In some cases, cultivars are bred to adapt to multiple environments, suggesting that these might not respond to precision management. The technologies needed to vary seed type or planting geometry are available. What appears to limit precision crop management are lack of accurate maps of the environments within fields and lack of agronomic decision rules for varying genetics or geometry across environments encountered within fields and across growing seasons.

Precision pest management. Pests are spatially aggregated and not randomly distributed within most agronomic fields. Furthermore, the efficacy, efficiency, and fate of pest-control inputs vary with pest and crop conditions and with physical and chemical properties of the soil, all of which can vary spatially and to some extent temporally. The application of precision agriculture to pest management is potentially beneficial to agriculture not only because it offers an opportunity to reduce chemical/nonchemical inputs into crop production and improve pest management, but also because the acquisition of spatial and temporal information on pest occurrence and distribution that has been made possible with precision-agriculture technologies will lead to an improved understanding of the pest biology and ecology needed to develop more effective pest management strategies. Thus, an important dimension of precision pest management is that new

pest management strategies will evolve because of it. A major difficulty with the management of some pests is that their populations are highly dynamic, and prediction of pest density is difficult or uncertain; both of these make it necessary to collect field estimates of pest density in order to monitor density over time. The difficulty and cost in mapping pest density is considered by some to be an insurmountable barrier to the use of precision pest management in established integrated pest management (IPM) programs. New techniques to measure and manage pest variability are needed to support precision pest management.

Precision water management. There are three approaches to precision water management: variable rate irrigation, matching agronomic inputs to water availability defined by soil and/or landscape properties, and drainage. Like other factors, variable irrigation management requires sufficient knowledge about the spatial and temporal variability of the factors that regulate water availability to plants and water use in plants in order to achieve precision management. Differences in water availability within a field are favored by the occurrence of dissimilar soil types, the presence of soil degradation processes (such as erosion), and variation in landscapes that affect watershed hydrology. Variable-rate irrigation systems are available, but cost-effective systems to measure plant water availability within fields will likely emerge with future advances in remote sensing. Poor drainage is often cited by farmers as a source of yield variability within fields. Since a number of options for drainage currently exist and can be applied site-specifically, there is little need to design site-specific drainage practices. Site-specifically, the cost of draining portions of fields or small isolated areas may be higher because costs to connect to drain outlets are proportional to the distance to the outlet. *See* IRRIGATION (AGRICULTURE).

Evaluation. Three important evaluation issues surround precision agriculture: economics, environment, and technology transfer. The economic evaluation focuses on whether the documented agronomic benefits—translated into value through market mechanisms—exceed the technological and service costs. The most important fact regarding the analysis of the profitability of precision agriculture is that the value comes from the application of the data, not from the use of the technology itself. This can be contrasted with traditional agricultural innovations, where the value is derived from the use of the new technology itself, for example, new seed genetics that increase yield or a new herbicide that reduces yield loss. If any value is to be derived from precision agriculture, it will come from the resulting management decisions, not through the use of the enabling technologies.

Environmental evaluation focuses on whether precision agriculture can improve soil, water, and the general ecological sustainability of agricultural systems. Reduced use of agrochemicals, more efficient use of nutrients, increased efficacy of managed inputs, and increased protection of soils from degradation are frequently cited as potential benefits to the environment. While the impacts of precision agriculture on the environment are assumed positive, proof that it is even benign in its environmental impacts is generally undocumented. Studies evaluating the environmental benefits of precision agriculture are limited, and it is quite possible that negative impacts of precision agriculture could derive from such things as increased aggression by farmers on site-specific potentials or problems. Furthermore, a number of factors may limit reductions in chemical applications in some situations and encourage increased use in others.

Finally, there is the question of whether this bundle of enabling technologies and agronomic principles will work on individual farms. The impacts of precision agricultural systems will extend beyond crop production to the environment and to the very structure of the agriculture system. Consequently, evaluation needs to involve all sectors of agriculture. *See* AGRICULTURE. F. J. Pierce; P. Nowak

Bibliography. P. A. Burrough and R. A. McDonnell, *Principles of Geographic Information Systems*, Oxford University Press, New York, 1998; Environmental Systems Research Institute (Redlands, CA), *Understanding GIS: The ARC/INFO Method*, John Wiley, New York, 1997; M. Morgan and D. Ess, *The Precision-Farming Guide for Agriculturalists*, John Deere Publications, Moline, IL, 1997; F. J. Pierce and P. Nowak, Aspects of precision agriculture, *Adv. Agron.*, 67:1–85, 1999; F. J. Pierce and E. J. Sadler (eds.), *The State of Site-Specific Management for Agriculture*, American Society of Agronomy, Crop Science Society of America, and Soil Science Society of America, Madison, WI, 1997.

Precision approach radar (PAR)

A tracking system that provides a ground control approach (GCA) air-traffic controller with a precise display of an aircraft's position relative to a runway final-approach course. To ensure absolute safety, precise information is displayed on a plan position indicator (PPI). This display provides the controller with aircraft position information for control of heading and rate of descent. To accomplish this and maintain the required precision for a final-approach aid, the display shows the aircraft position in relation to range, azimuth, and elevation. The information presented on the precision approach radar display allows an air-traffic controller to direct a pilot down along a runway approach course to a precision landing. Precision radar approaches are accomplished in most weather conditions and do not require any on-board avionics equipment, such as an instrument landing system (ILS).

Antenna systems. There are two ways to provide the required range, azimuth, and elevation information on the plan position indicator: use of two antennas, one scanning elevation and the other azimuth; and a single computer-controlled phased-array antenna that can provide pencil-beam

tracking for both elevation and azimuth positions. The dual-antenna system uses a single transmitting magnetron that supplies radio-frequency power at a frequency of approximately 9 GHz to both antennas. A revolving mechanical switch then alternately feeds radio-frequency energy to each antenna. The azimuth antenna beam width scans 20°, and the elevation beam scans 8°. For the single antenna, radio-frequency energy is developed by a traveling-wave tube–crossed-field amplifier. The phasing and direction of the transmitted beam are controlled by a computer through phase shifters. This scan coverage is also 8° in elevation and 20° in azimuth. *See* ANTENNA (ELECTROMAGNETISM).

Display. The plan position indicator is divided into two sections (see **illus.**). The upper section is used to display the aircraft's elevation on an elevation cursor. The lower display shows the aircraft's azimuth by using a course-line cursor. This combination of azimuth and elevation display is commonly referred to as the AZ-EL indicator. Range marks are also displayed on the AZ-EL indicator at 1-mi intervals. The range marks are spaced logarithmically to provide increased detail at critical close ranges. To increase safety even further, an elevation safety cursor is displayed. The elevation scan requires increased accuracy because of an aircraft's inherent instability while descending compared to steady course. This safety cursor ensures that an aircraft does not fly too low for a safe approach. The safety cursor is distinguished from the elevation cursor by dashing or blinking. A decision-height indicator cursor also appears on the elevation display. Decision height is the point at which the pilot must decide, based on the ability to acquire runway visual references, either to continue the approach to land, or to execute a missed approach.

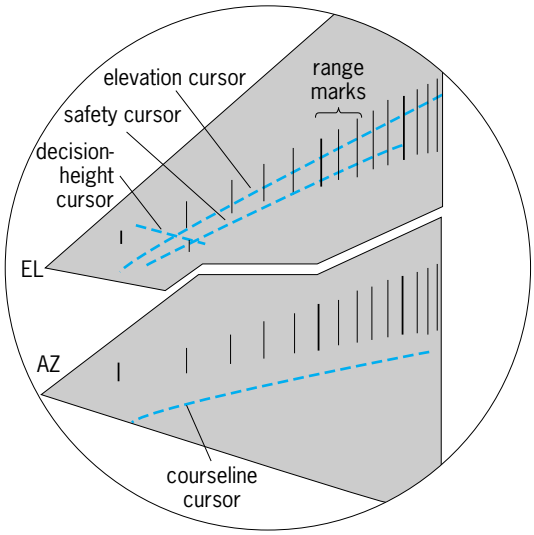

Standard plan position indicator (PPI) display for precision approach radar, with azimuth (AZ) and elevation (EL) displays. (*After Air Traffic Control Radar Specialist, PAR Indicators, Performance Monitoring and Assessment, and Auxiliary Systems, Career Development Course, vol. 4, Extension Course Institute Air University, Gunter Air Force Station, Alabama, 1982*)

Configuration and use. Precision approach radar systems are configured as both permanent and mobile facilities. Their use is particularly valuable for locations where terrain, excessive site preparation costs, and airspace restrictions prevent installation of an instrument landing system. The use of precision approach radar systems is heavily dependent on personnel and human factors. Differences in human reaction times can limit the effectiveness of such systems. After observing a target deviation, the controller issues a heading correction or an advisory in relation to the course or elevation cursor. Depending on the controller's reaction time and the pilot's response to instructions, there may be a several-second delay before the aircraft actually alters course. *See* INSTRUMENT LANDING SYSTEM (ILS).

Precision approach radar systems, which are located next to the runway, rotate to more than one runway to accommodate wind direction and changes to the runway in use. In some cases, the shelter that houses the air-traffic controllers and the plan position indicator display is in a location remote from the precision approach radar equipment. Also, some precision approach radar systems can data-link control instructions to the pilot without radio communication. *See* ELECTRONIC NAVIGATION SYSTEMS; RADAR.

Mark A. DePlasco

Bibliography. D. Rider (ed.), *Jane's Airport and ATC Equipment*, 12th ed., 1993–1994; M. I. Skolnik, *Radar Handbook*, 1990; G. J. Sonnenberg, *Radar and Electronic Navigation*, 6th ed., 1988.

Predator-prey interactions

Predation occurs when one animal (the predator) eats another living animal (the prey) to utilize the energy and nutrients from the body of the prey for growth, maintenance, or reproduction. In the special case in which both predator and prey are from the same species, predation is called cannibalism. Sometimes the prey is actually consumed by the predator's offspring. This is particularly prevalent in the insect world; for example, a wasp may paralyze and lay eggs on a butterfly caterpillar, leaving its offspring to consume the prey. Insect predators that follow this type of lifestyle are called parasitoids, since the offspring grow parasitically on the prey provided by their mother.

Predation is often distinguished from herbivory by requiring that the prey be an animal rather than a plant or other type of organism (bacteria). To distinguish predation from decomposition, the prey animal must be killed by the predator. Some organisms occupy a gray area between predator and parasite. For example, lamprey are generally considered parasites that attach to the sides of fish and suck blood. However, lamprey can reach sufficient size to kill their hosts, and then they act as predators. Finally, the requirement that both energy and nutrients be assimilated by the predator excludes carnivorous plants from being predators, since they assimilate only

nutrients from the animals they consume. *See* FOOD WEB.

Population dynamics. Population dynamics refers to changes in the sizes of populations of organisms through time, and predator-prey interactions may play an important role in explaining the population dynamics of many species. Predators affect the population dynamics of prey, because increases in the density of the predator population lead to increased predation on the prey and a consequent decrease in the prey's population growth rate. Conversely, increases in the density of prey mean more potential food for predators and a consequent increase in the predator's population growth rate. Predator-prey interactions are thus a type of antagonistic interaction, in which the population of one species has a negative effect on the population of a second, while the second has a positive effect on the first. For population dynamics, predator-prey interactions are similar to other types of antagonistic interactions, such as pathogen-host and herbivore-plant interactions.

The characteristic population dynamics associated with predator-prey interactions are predator-prey cycles. Simple mathematical models of predator-prey interactions exhibit cycles; cycles are also found in simple laboratory experiments. In nature, several examples of cycles have been observed, the most famous of which is the snowshoe hare cycle (see **illus.**). How cycles arise can be understood by considering what happens when both prey and predator densities are low. With few predators, the density of the prey increases. This in turn increases the population growth rate of the predator. The resulting increase in the predator's density decreases the population growth rate of the prey. When the prey's population growth rate decreases sufficiently, the prey density decreases, and this in turn leads to a decrease in predator density. The cycle then repeats itself.

Even though snowshoe hares and their predators show cyclic dynamics, extensive experiments and observations have shown that the cycles are not simply driven by interactions between hares and their predators. In concert with the predator-hare cycles, there is also a hare-plant cycle in which plant abundance and quality decrease when hare densities are high, resulting in reduced body condition and reproduction of hares. Furthermore, there is a synergistic effect between plant quality and predation, with reduced food availability leading to increased predation rates, possibly because low food availability forces hares to feed in areas that are more exposed to predation. Thus, the cyclic dynamics of snowshoe hares could be the product of simultaneous predator-prey and herbivore-plant interactions.

Stabilization of predator-prey dynamics. Although there are several well-documented cases of cyclic population dynamics, predator-prey cycles are very rare when considering all the possible predator-prey systems in nature. Nonetheless, predator-prey interactions are thought to increase fluctuations in the population dynamics of predator and prey, with these fluctuations having the potential to cause local extinctions. This has been shown by many studies on organisms that can be maintained in the laboratory; when predator and prey are placed together, the population dynamics show cycles of increasing amplitude, eventually leading to the extinction of the predator, or both prey and predator.

Given the contrast between persistent predator-prey systems in nature and the difficulty of constructing persistent predator-prey systems in the lab, much theoretical and experimental work has been

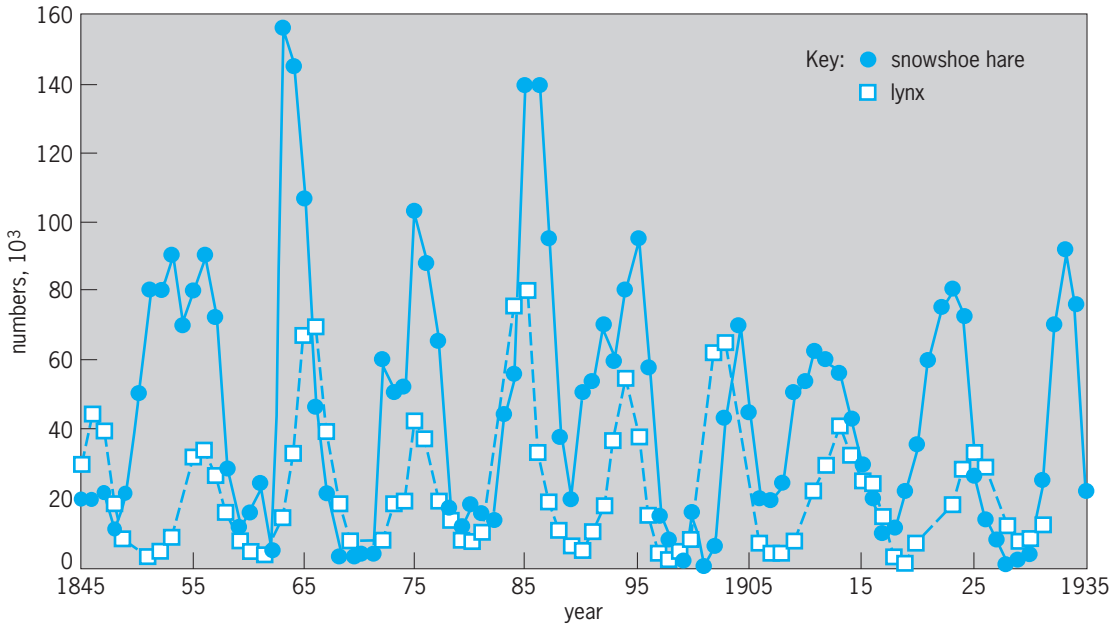

Snowshoe hare and lynx populations inferred from the numbers of pelts sold to the Hudson's Bay Company. (*After D. A. MacLuclich, 1937*)

devoted to explaining what stabilizes predator-prey dynamics in nature. Most of the numerous possible explanations involve heterogeneity, either in the habitat in which the species live or in the susceptibility of prey individuals to predation. If, for example, the habitat contains a refuge in which the prey are safe from predation, the prey population cannot be driven extinct, and this results in a stable predator-prey system. This appears to occur frequently in insect predator-prey systems, in which the herbivorous prey hide in the nooks and crannies of plants where they cannot be found by predators. Similarly, if a certain fraction of the prey population is invulnerable to attack, then again the predator cannot drive the prey population extinct. A possible example is rhinos, which are vulnerable to predators when they are young, but once they reach full size they can defend themselves against any predator. Research has provided a wealth of rich examples showing how heterogeneity may stabilize predator-prey interactions.

Biological control. Biological control is the use of natural enemies to control pests, and demonstrates how ecological knowledge can be used to address an applied problem. In classical biological control, an exotic natural enemy is introduced into a new country to control a pest, normally a pest that also is exotic. After introduction and establishment of the natural enemy, no further intervention is conducted. In contrast, biological control can be conducted actively, for example by continuously breeding and releasing natural enemies into targeted areas. Repeated, inundative releases are most often used in greenhouses to control pests on vegetable and flower crops.

The first biological control program was conducted in 1888, when the vedalia (a type of ladybird beetle) was released to control scale insects on citrus crops in California. Since then, many hundreds of biological control programs have targeted large numbers of insect and plant pests. For example, prickly pear cacti (*Opuntia stricta*) were introduced into Australia as garden plants in the nineteenth century, but in the absence of native flora similar to cacti, the prickly pear population exploded to cover much of the rangeland of eastern Australia. In 1926 Australian entomologists introduced the cactus moth (*Cactoblastus cactorum*) from Argentina, and within 15 years the prickly pear population was reduced to a few patches spread throughout the eastern part of the country. In practice, introduced natural enemies are selected to specialize only on the target pest, and for this reason many control agents are wasp parasitoids, since these tend to attack a narrow range of host species.

The basic problem of what stabilizes predator-prey dynamics figures prominently in understanding the success of classical biological control programs. In successful programs, the natural enemy suppresses the pest density to low levels, yet does not lead to total extinction. If extinction were to occur, the reintroduction of the pest would lead to an outbreak in the absence of the natural enemy. For example, the cactus moth has not completely eradicated the prickly pear from Australia, which still persists in small patches. Even though these patches provide a source of prickly pears that could spread to other areas, they also provide a source of cactus moths that can spread and control incipient prickly pear outbreaks.

Predation and community structure. Community structure refers generally to how species within an ecological community interact. The simplest conception of a community is as a food chain, with plants or other photosynthetic organisms at the bottom, followed by herbivores, predators that eat herbivores, and predators that eat other predators. This simple conception works well for some communities. For example, deer and moose eat plants, and wolves eat deer and moose. Nonetheless, the role of predator species in communities is often not clear. Many predators change their ecological roles over their lifetime. For example, pike may reach 50 lb (23 kg) and are generally viewed as predators at the top of the food chain in lakes, yet many are killed as young fingerlings by diving beetles. Many insect predators that share the same prey species are also quite likely to kill and devour each other. This is called intraguild predation, since it is predation within the guild of predators. Furthermore, many species are omnivores, feeding at different times as either predators or herbivores; for example, brown bears feed on both plant roots and salmon. Therefore, the role of particular predator species in a community is often complex.

Predator-prey interactions may have a large impact on the overall properties of a community. For example, most terrestrial communities are green, suggesting that predation on herbivores is great enough to stop them from consuming the majority of plant material. In contrast, the biomass of herbivorous zooplankton in many aquatic communities is greater than the biomass of the photosynthetic phytoplankton, suggesting that predation on zooplankton is not enough to keep these communities green. As a more specific example, the diversity of species occupying the rocky intertidal area along the coast of the Pacific Northwest is largely determined by the presence or absence of a large predatory starfish (*Pisaster ochraceus*). When starfish are absent, the community becomes dominated by mussels that are competitively superior to other sessile intertidal organisms. When starfish are present, however, they preferentially eat mussels, thereby creating vacant areas that are colonized by less competitive organisms such as barnacles. *See* POPULATION ECOLOGY.

Anthony R. Ives

Bibliography. M. Begon, J. L. Harper, and C. R. Townsend, *Ecology: Individuals, Populations and Communities*, Blackwell Scientific Publications, Boston, 1990; S. I. Dodson et al., *Ecology*, Oxford University Press, New York, 1998; D. A. MacLuclich, Fluctuations in numbers of the varying hare (*Lepus americanus*), *Univ. Toronto Studies Biol. Ser.*, 43:1–13, 1937; R. J. Taylor, *Predation*, Chapman and Hall, New York, 1984.

Pregnancy

The period during which a developing fetus is carried within the uterus. In humans, pregnancy averages 266 days (38 weeks) from conception to childbirth. Traditionally, pregnancy duration is counted from the woman's last menstrual period, which adds roughly 2 weeks to gestational age. This is how physicians arrive at a pregnancy length of 40 weeks (280 days).

Establishment of pregnancy. The critical events that establish pregnancy are ovulation, fertilization, and implantation. In women of childbearing age (15–44), the release of an egg from the ovary (ovulation) occurs approximately 14 days after the beginning of a menstrual period. Successful ovulation requires the coordinated interplay of hormones produced by the ovary, pituitary gland, and hypothalamus. Conception or fertilization, the union of a single sperm with the egg, usually takes place in the Fallopian tube within a day or two of ovulation. The fertilized egg (zygote) next passes into the uterine cavity and, on about day six, attaches (implants) into the lining tissue or endometrium. This tissue, which normally sloughs off each month in the nonpregnant female, is instead maintained and nurtured by placental hormones. As pregnancy progresses, the placenta mediates the inflow of nutrients and removal of waste products while synthesizing increasing amounts of estrogen and progesterone, hormones that influence nearly every tissue and organ system in the body. The 9 months of pregnancy are typically divided into three periods (trimesters) of 3 months. *See* FERTILIZATION (ANIMAL); MENSTRUATION; OVUM; PLACENTATION; SPERM CELL.

The first sign of pregnancy is often the absence of an expected menstrual period. Common symptoms include nausea, breast tenderness, fatigue, and frequent urination. The diagnosis of pregnancy can be made as early as 10 days after fertilization by means of blood tests. By 6 weeks (from the last menstrual period), the uterus feels soft and is palpably enlarged. Pregnancy can be positively confirmed by observing cardiac motion of the fetus by ultrasound scanning (8 weeks) or by hearing fetal heart "tones" by using a Doppler detection instrument (10–12 weeks).

Embryonic development. Early in the first trimester, the embryo's germ layers differentiate into organs and systems, a process that is nearly completed by the twelfth week. It is during this critical period of development that the fetus is most vulnerable to the adverse effects of drugs and other teratogenic influences. The second and third trimesters of pregnancy are characterized by increased fetal growth and gradual physiologic maturation of fetal organ systems. During this time, the maternal changes of pregnancy are greatest. The enlarging uterus encroaches on the abdominal region by the fourth month and at term nearly reaches the diaphragm. The breasts gradually enlarge in preparation for lactation. Striking cardiovascular changes, including nearly a 50% increase in cardiac output, provide the increased blood flow to accommodate the growing fetoplacental unit. Other changes in the renal, digestive, pulmonary, and endocrine systems reflect the numerous maternal adaptations that eventually must occur in a healthy pregnancy. *See* EMBRYOLOGY.

Prenatal care. Early, regular prenatal care is associated with improved pregnancy outcome and seeks to identify risk factors in the pregnancy that may apply to mother or fetus. At 6–8 weeks, a complete physical examination, along with blood and urine analyses, should be performed. In addition to undergoing traditional tests, patients are now routinely screened for hepatitis B at the beginning of pregnancy, for fetal neural-tube defects such as open spine at 16 weeks, and for gestational diabetes at about 28 weeks. In addition to these blood tests, many physicians offer a sonogram at 16–18 weeks to establish gestational age, check for a multiple pregnancy, and screen for birth defects. During prenatal visits, a physician can evaluate nutrition, blood pressure, and fetal growth. Weight gain in the last two trimesters should normally be 1 lb (0.45 kg) per week, for a total of 25–35 lb (11–16 kg). During the pregnancy, common symptoms are noted, such as edema, varicose veins, hemorrhoids, backache, fatigue, heartburn, and vaginal discharge. Ideally, at the end of the third trimester, the process of labor begins. The muscles of the uterus contract, dilating the cervix and allowing the baby to begin moving into the vagina or birth canal. Continued contractions push the baby out of the mother's body. In the final stage of labor, the placenta detaches from the uterine walls and is expelled as the afterbirth. An alternative to vaginal delivery is the cesarean section, in which the baby is removed surgically through an abdominal incision. *See* PRENATAL DIAGNOSIS.

Birth rates and postnatal care. The number of United States births in 1988—82 million—was a modern record, the largest number of births in any single year since 1954. This appears to be due to a slight increase in the fertility rate (the number of births in a given period in relation to the number of women of reproductive age) along with a 1% increase in the number of women of childbearing age. Pregnancy and childbirth in the United States have become safer. The rate of maternal deaths dropped from 37 per 100,000 live births in 1960 to 8 per 100,000 in 1985. During this time, perinatal mortality—the sum of stillbirths and neonatal deaths (deaths within the first month of life)—also decreased from 34 per 100,000 births in 1960 to 15 per 100,000 in 1985. Factors credited with improving pregnancy outcome usually include increased hospital births (now over 99%), improved childbirth pain relief (such as epidural anesthesia), and increased reliance on specialists for mother-to-be and baby, including perinatologists and neonatologists. Regionalization of obstetric facilities and referral of certain complicated pregnancies to centers with neonatal intensive care units have improved the outlook for infants of low birth weight, those who weigh less than 5 lb (2500 g). Survival of even extremely premature infants with birth weights of less than 2 lb (1000 g) is often possible, although these

infants have increased rates of mental and physical handicaps. Better monitoring of physicians through hospital quality assurance programs has been identified as another factor that has contributed to safer health care for mother and baby. *See* REPRODUCTIVE BEHAVIOR.

Termination of pregnancy. The legal status of pregnancy termination (therapeutic abortion) varies from country to country, but about two-thirds of women in the world have access to legal abortion. Abortion became legalized throughout the United States in 1973, and since that time, abortion rates have risen steadily, with some leveling off between 1980 and 1985, at about 1.3 million abortions per year. Over 90% of abortions in the United States are performed in the first trimester by suction curettage, a technique that uses suctioning and removal of the uterine contents through the vagina with surgical instruments. Later pregnancies are terminated by a procedure called dilatation and evacuation (D&E) or by administration of drugs to stimulate uterine contractions. Medical and psychological sequelae to abortion are few, and are fewest for terminations in the first trimester. *See* PREGNANCY DISORDERS.

Bruce D. Shephard

Bibliography. W. G. Birch, *Guide to Pregnancy and Childbirth*, 1991; F. G. Cunningham, P. C. McDonald, and N. F. Gant, *Williams Obstetrics*, 20th ed., 1997; N. F. Maklad, *Ultrasonic Diagnosis of Ectopic Pregnancy*, 1990; R. H. Petrie (ed.), *Perinatal Pharmacology*, 1989; B. Shephard, *The Complete Guide to Women's Health*, 3d ed., 1997.

Pregnancy disorders

Physical disorders that can arise as a consequence of pregnancy, ranging from mild to life-threatening. Extreme conditions can result in termination of the pregnancy or death of the mother.

The most common disorder of early pregnancy, persistent vomiting, is without known cause and usually subsides spontaneously within a few weeks. A more serious threat to pregnancy, vaginal bleeding (within the first 20 weeks), may be a sign of miscarriage (spontaneous abortion) or much less often, ectopic pregnancy (embryonic development outside the uterus, usually within the Fallopian tube). Approximately 10–15% of pregnancies reportedly end in miscarriage, although the actual rate is closer to 40%, as some early miscarriages are not recognized. Over 50% of miscarriages are due to a spontaneous chromosomal abnormality in the sperm, egg, or developing embryo. Ectopic pregnancy occurs about once in every 50 pregnancies and represents a leading cause of maternal mortality in the United States. Ectopics are nearly always removed surgically; the involved Fallopian tube usually can be preserved.

Some abnormal conditions in pregnancy are identified as a result of routine blood tests. Examples include rubella (German measles) and syphilis, both of which occur relatively infrequently and can cause birth defects. Rh disease is also uncommon because of prevention by prenatal blood tests and treatment with immunoglobulins in Rh-negative mothers. Routine screening is also recommended for hepatitis B, an increasingly prevalent liver condition that can cause maternal liver damage or infection in the newborn. Although no specific treatment for hepatitis B is available for the mother, it is possible to prevent infection in the newborn by using vaccine and immunoglobulin immediately after birth. *See* HEPATITIS; PRENATAL DIAGNOSIS; RH INCOMPATIBILITY; RUBELLA; SYPHILIS.

Maternal conditions that may worsen during pregnancy include some forms of heart disease, seizure disorders, hypertensive disease, and acquired immune deficiency syndrome (AIDS). Transmission of the AIDS virus to the newborn occurs in up to 50% of cases and can take place at any stage of pregnancy. Mortality among HIV-infected infants is about 45%, and many die before the age of 2. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS).

Among the most important disorders during the second half of pregnancy are those associated with low birth weight due to either premature labor or fetal growth problems. Low birth weight is associated with twins, hypertensive disorders, smoking, and inadequate nutrition. The hypertensive disorders include pregnancy-induced hypertension (formerly known as toxemia) as well as chronic hypertension that exists before the pregnancy. Pregnancy-induced hypertension occurs in about 5% of pregnancies, mostly among women who have not given birth previously. Signs and symptoms can include swelling of the hands and face, headaches, and a sudden weight gain of 5 lb (2 kg) or more in 1 week. Pregnancy-induced hypertension can be assessed by in-office fetal monitoring tests as well as by sonography to evaluate fetal growth. These tests are important for evaluating many other complications, including bleeding in late pregnancy, which is a potential emergency condition. Vaginal bleeding often results from a placental problem; either the placenta is abnormally located near or over the cervix (placenta previa), or the placenta separates prematurely from the uterus (abruptio placenta). *See* HYPERTENSION.

Far less dramatic than these conditions, but equally important as a cause of fetal distress, is the "post dates" pregnancy, which occurs once pregnancy has extended 2 weeks beyond the date of expected delivery. When a pregnancy reaches 42 weeks, delivery is attempted as soon as possible as the aging placenta may lose its ability to provide adequate oxygen and nutrition.

Increased rest and proper nutrition are especially important in pregnancies complicated by high blood pressure, fetal growth problems, and twins. More immediate management often centers on the timing of delivery. Certain abnormal test results may indicate the need for delivery before term, often by cesarean birth. Many potentially difficult pregnancies can be conservatively managed until near term, however. If preterm labor threatens, several promising drugs

are available to treat this condition, but this situation as well as other problems such as persistent bleeding may require prolonged hospitalization. *See* PREGNANCY. Bruce Shephard

Bibliography. R. S. Abrams, *Handbook of Medical Problems During Pregnancy*, 1988; G. N. Burrow and T. R. Ferris, *Medical Complications During Pregnancy*, 5th ed., 1999; D. F. Hawkins (ed.), *Drugs and Pregnancy: Human Teratogenesis and Related Disorders*, 2d ed., 1987; C. Stern (ed.), *Immunology of Pregnancy and Its Disorders* 1988.

Prehistoric technology

The set of ideas that prescribe the manufacture and use of implements before written history. Technology is the principal means through which the human species has succeeded in occupying most of the world. The prehistoric record documents that for over 2 million years the human lineage has been making and using implements. Archeologists tend to use the term "artifact" for any material that was modified by ancient humans, whether this material was used or not, and the term "tool" for any material that was used by ancient humans, whether it was modified or not.

The archeological evidence for prehistoric technologies may be biased because of the nature of the raw materials that were used and the conditions of burial and preservation of these materials. In general, artifacts of stone, bone, pottery, and metal preserve fairly well in many areas, whereas artifacts of wood and other vegetable materials, skin, and horn tend to decay fairly rapidly and are normally found only in exceptional prehistoric contexts.

In 1816, C. Thompsen, director of the Danish National Museum, began to chronologically order the museum's prehistoric collections into three major groups, based upon technology, and now called the Three Age System. The earliest was a Stone Age, followed by a Bronze Age, and finally an Iron Age. As time went on, the prehistory of Europe was further divided, based on regional sequences that could be documented through excavation. In other places, such as the Americas, Australia, Oceania, and sub-Saharan Africa, different nomenclature was often used, since the regional technological sequences differed from that of western Europe.

New technological traits can be introduced into a society in a number of ways: (1) innovation or invention, which is the development of new ideas, including new technological characteristics; (2) diffusion, which involves the spread of ideas, including technological knowledge, from one group to another; and (3) migration, which is the spread of peoples, often with new technologies, into new areas. One of the goals of a prehistorian is to try to ascertain, based on archeological evidence, which factors best explain the technological changes seen in the archeological record.

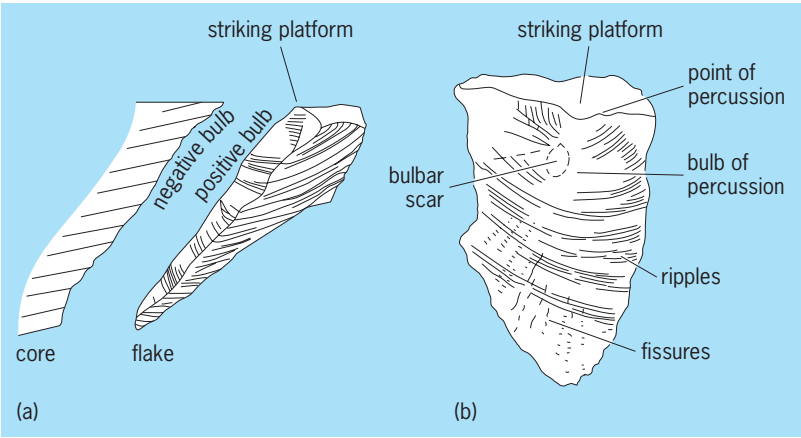

Fig. 1. Elements of stone flaking. (*a*) Relationship between the core and the flake. (*b*) Major feature of the ventral or release surface of the flake. (*After I. Tattersall, E. Delson, and J. A. Van Couvering, eds., Encyclopedia of Human Evolution and Prehistory, Garland, 1988*)

Stone Technology

Archeological evidence has demonstrated that the human lineage has been making and using identifiable tools for at least 2.5–2.6 million years, beginning with stone implements and associated prehistoric remains. Although animals may use unmodified objects as tools (for example, the California sea otter cracks open shellfish with a stone hammer or anvil), most of this behavior appears to be instinctive. It is likely that early hominids (proto-humans) went through rudimentary technological stages that predated the earliest stone artifacts. Materials that could have been used for tools include wood, bone, horn, and shell. Unfortunately, these materials either tend to be perishable or do not usually leave clear indications of modification or use in such a primitive technology.

Principles of stone fracture. For most of the Stone Age, rocks were fashioned into various forms by flaking, that is, removing spalls (flakes) from a stone by percussion or pressure. The ideal raw materials for such flaking are hard, fine-grained rocks that fracture well in any direction and have no preferential cleavage planes, such as flint, chert, chalcedony, jasper, agate, quartzite, lava, and obsidian. In flaking stone, a glancing blow at an angle of less than 90° with a hammer of stone or other material normally produces a controlled fracture pattern that results in a parent piece of rock (the core) from which sharp fragments are detached (flakes and fragments). Such conchoidal fracture normally produces flakes that exhibit a bulb of percussion, ripples, and fissures or hackle marks (**Fig. 1**). Such features help identify rocks that have been modified by humans. A number of techniques can be used for working stone, each of which produces diagnostic flaking characteristics and by-products (**Fig. 2**). These include hard-hammer percussion, anvil technique, bipolar technique, soft-hammer percussion, indirect percussion (punch technique), and pressure flaking.

Cultural-historical outline. Much of the nomenclature used in Stone Age studies dates from the nineteenth century and applies to western Europe. It

later became clear that the technological progression of populations in prehistoric Europe was not necessarily universal and that any given geographical region has its own prehistoric sequence of development.

Early Paleolithic. The first definite stone tools come from African prehistoric sites, the earliest in Ethiopia dated to approximately 2.5–2.6 million years ago (Ma). These technologies have been named Oldowan, after the famous site of Olduvai Gorge in northern Tanzania, and consist of simple flaked cobbles or blocks of stone made into a range of core forms, battered stones, retouched flake scrapers, and flakes and fragments produced from these core forms as well. The primary techniques used were hard-hammer percussion (Fig. 2*a*), which consisted of direct blows to the core with a stone hammer, and bipolar technique (Fig. 2*c*), which involved setting the core on a stone anvil and striking the core from above with a hammerstone.

These earliest industries are known to be contemporaneous with at least two major fossil hominid lineages. The first is the smaller-brained *Australopithecus boisei* in eastern Africa and *robustus* (also known as *Paranthropus robustus*) in southern Africa. The second lineage encompasses the larger-brained forms, which are believed to be ancestral to modern humans, *Homo habilis* and *rudolfensis* and subsequently early representatives of *H. erectus* (*H. ergaster* in Africa). Although it cannot be conclusively demonstrated which hominid forms were the major toolmakers, the larger-brained genus *Homo* was probably the more habitual toolmaker. *See* FOSSIL HUMANS.

Early archeological sites tend to be situated in tropical or subtropical grassland environments, often near streams or lakes, and are associated with animal bones, some of which bear cut marks from sharp-edged stone tools. Early stone tools may have served as hammers for cracking open long bones to obtain marrow and as chopping or scraping tools for making wooden spears or digging sticks.

A new industrial stage, the Acheulean, began about 1.5 Ma in Africa and appears to have spread to many parts of Europe, the Near East, and the Indian subcontinent. Artifacts include hand axes, cleavers, and picks, either made on large flake blanks struck from boulder cores or worked down from large cobbles or nodules of rock. These large flakes could have been detached by hard-hammer percussion, with a large hand-held stone hammer or thrown boulder hammer, or by anvil technique (Fig. 2*b*), in which a large core was struck against a stationary stone anvil to detach flakes. Later Acheulean industries often exhibit finer, more controlled flaking and thinner, more refined hand axes and other tool forms, which suggests that soft-hammer percussion (Fig. 2*d*) was employed. By this technique, a material that is softer than the stone being worked, such as bone, antler, wood, or ivory, is used as a percussor.

Acheulean and contemporaneous industries are associated with the fossil forms of *H. erectus* and a larger-brained hominid, often designated as archaic *H. sapiens*. Between 1.7 Ma and 200,000 years ago, these populations spread over much of the Old World, occupying tropical and temperate environments. In most of eastern and southeastern Asia, stone technologies were more similar to Oldowan sites, perhaps because of the extensive use of perishable cutting materials such as split bamboo. Some prehistoric sites of this technological stage also show evidence of the use of fire.

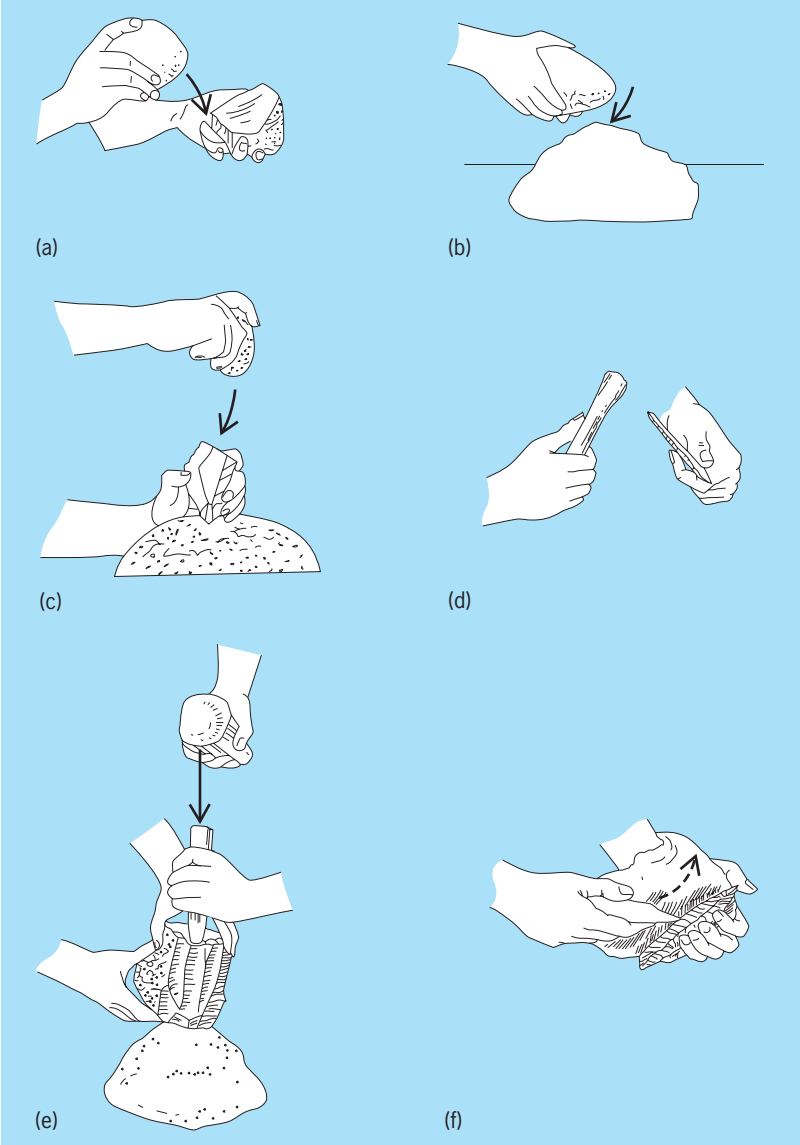

Fig. 2. Principal techniques of stone-tool manufacture. (*a*) Hard-hammer, with free-hand percussion. (*b*) Anvil technique. (*c*) Bipolar technique. (*d*) Soft-hammer percussion. (*e*) Indirect percussion or punch technique. (*f*) Pressure flaking. (*After I. Tattersall, E. Delson, and J. A. Van Couvering, eds., Encyclopedia of Human Evolution and Prehistory, Garland, 1988*)

It is important to remember that a direct correspondence does not exist between stone industries and human species. The earliest Oldowan tools are generally and controversially attributed only to early members of the genus *Homo*, but continued to be made by all subsequent human species in various parts of the world. *Homo ergaster* had appeared in Africa by 1.8 Ma, but Acheulean tools are

currently thought to have emerged 300,000 years later. Archaic and anatomically modern *H. sapiens* are both associated with Middle Paleolithic technologies.

Middle Paleolithic. Beginning over 100,000 years ago, the large hand axes and cleavers so characteristic of the Acheulean were replaced in many areas by smaller, standardized flake tools such as side scrapers, which were probably used for hide working and woodworking, and unifacial points for projectile points or butchery knives. Hard- and soft-hammer techniques were common during this industrial stage. Some Early and Middle Paleolithic technologies included prepared-core or Levallois methods that served to produce tool blanks of predictable size and shape.

In Europe and the Middle East, most of these industries are associated with the fossil hominid *H. sapiens neanderthalensis* (Neandertal). In some sites in Africa and Israel, however, Middle Stone Age industries have been found with anatomically modern human fossils of *H. s. sapiens*. Compared with previous hominids, Neandertals appear to have been adapted to a wider range of environments, including the severe cold of Europe during the last glaciation, and hunting may have become more common. *See* NEANDERTALS.

Later Paleolithic. Major changes in lithic technology are observed after 40,000 years ago. Blade blanks have been found on Early Paleolithic sites and were struck from prepared cores during the Middle Paleolithic. However, blades were produced from different core technologies during the Late Paleolithic and were made into a variety of tool forms, such as end scrapers, burins, and awls. In blade technology, blanks at least twice as long as they are wide appeared. Blades can be produced by hard- and soft-hammer percussion, but in the finest blade industries a new technique, indirect percussion or punch technique (Fig. 2*e*), was employed. A punch of antler, horn, or other material was set against the edge of a blade core and struck at the opposite end with a hammer. This technique allowed for more precise flaking and control. However, at the present time indirect techniques are clearly indicated only during the postglacial Mesolithic Period.

In many places, bones and antler tools become common for the first time as well, as can be observed in the form of spear points, barbed harpoon heads, needles, spear throwers, and the earliest representational sculpture and engraving. The emphasis on stone tools to make other tools in a range of materials seems to have accelerated markedly during this period, and it is likely that wooden tools were more common as well.

More geographical variation and faster technological changes occurred, which may suggest different cognitive and language skills compared with earlier hominid forms. These later Paleolithic technologies often take one of three forms: a large blade industry (Late Paleolithic); a small blade industry with small geometric microliths (Mesolithic, Epipaleolithic, or Later Stone Age); or a technology based on recurrent projectile point forms, often bifacially worked. A new technique, pressure flaking (Fig. 2*f*), was employed as early as the Late Paleolithic in some areas, in which small flakes or bladelets were removed by pressing them off an implement with a pointed piece of suitable material, such as antler, bone, or horn. The finely executed projectile points that are characteristic of such places as western Europe, the Horn of Africa, and the Americas were often fashioned by this technique.

Later Paleolithic hunters and gatherers appear to have become more adept at predicting the location of resources and exploiting potential foods than were earlier populations. Big-game hunting was characteristic of some areas; the spear thrower (atlatl) and the bow and arrow appear to have been invented during the last 25,000 years, with single stone points or tips made of composite microliths. Harvesting of fresh-water and marine resources became more common, and the concentration on certain plant resources is first documented during this time period. All of these industries subsequent to 30,000 years ago are associated with anatomically modern humans, *H. s. sapiens*.

Food-producing technologies. With the advent of food-producing communities (called Neolithic or New Stone Age in the Old World), beginning around 9000 years ago, new stone implements became common, including ground stone tools (especially axes and adzes) and seed-grinding equipment (querns and grindstones). Ground stone tools could be fashioned by first flaking or pecking and then final grinding and polishing, usually using an abrasive grit or sand with water against another piece of rock. They permitted the use of coarser-grained raw material than was previously employed. Most of these ground stone tools were hafted onto a handle of some sort. Desirable lithic raw materials or tools often became a major commodity of trade networks.

The advent of metalworking technologies, especially in the Old World, gradually brought about a decline in stone artifacts. By 2000 B.C., stone tools had been replaced by metal implements in much of Eurasia. Subsequent migrations of peoples and diffusion of ideas accelerated this replacement.

Nicholas Toth; Brooke Blades

Ceramics

Ceramics refers to a class of materials made of clay mixed with a binding liquid, usually water, which are shaped and hardened through exposure to heat. Depending upon the types of clays used, porosity, and temperatures reached in kilns, they are classified as earthenware or terracotta, stoneware, and porcelain. Earthenware ceramics found in archeological contexts are fired at relatively low temperatures, about 1100–1832°F (600–1000°C); stoneware and porcelains, about 2200–2600°F (1200–1400°C). Some early ceramics were sun-baked.

The earliest known earthenwares are figurines discovered at Dolni Vestonice in former

Czechoslovakia, dating to approximately 28,000 years ago. Containers of fired clay are found somewhat later, occurring in archeological contexts by 10,000 B.C.E. (Before the Common Era) in the Near East and Japan (the Jomon) and slightly later in other parts of the Old World. In the New World, they appear around 3000 B.C.E. Prehistoric peoples used ceramics for a variety of purposes. In addition to containers and figurines, clay was made into amulets, ornaments, bricks, tiles, house models, storage pits, and mudbrick architecture. Clay objects or tokens may even have been the earliest precursors to writing. Hand-modeled tokens first appear in 8000 B.C.E. in the form of geometrics and other shapes. Each token represented a particular good and was part of the first known abstract numerical system. This system was largely replaced around 3200 B.C.E., when clay tablets impressed with shapes like the older tokens began to be utilized. Other innovations in ceramic technology appeared later. Stoneware and porcelain are known from the fourteenth or fifteenth century B.C.E. in the Far East. However, the technology of stoneware production had developed earlier in the Indus Valley in Pakistan at around 2500 B.C.E., when it was used to produce bangles. The technique apparently was lost until its reinvention.

Manufacture. Prehistoric potters probably used many of the same procedures adopted by present-day potters throughout the world (**Fig. 3**). Several basic steps must be considered in producing ceramic objects: procuring raw materials, refining clays, shaping and finishing objects, applying decoration to them, drying and heating them.

Obtaining raw materials involves gathering or mining from natural deposits or through trade or purchase. Raw materials include clay, the binding liquid, and the inclusion of minerals or organic matter, such as sand, shell, fiber, crushed rock, or reused pottery bits. Their inclusion improves the workability of clays and minimizes cracking during drying or firing. Potters often have preferences for different types of clay, especially as to color. Many prehistoric potters preferred clays rich in iron oxides that imparted an earthy tone to the final products.

Refining clays involves extracting foreign matter, as well as drying, crushing, sieving, or soaking the clay in water in order to settle out heavy particles.

Shaping ceramics can be accomplished using a variety of hand-building methods, such as pinched, coiled, slabbed, or molded techniques. The most rudimentary is the pinched pot that involves stroking a lump of clay into a desired shape. Coiled and slabbed constructions are additive techniques in which clay coils or slabs are successively added to a base. Molds are forms onto which clay is pressed. The mold provides a structure for complex shapes or relief decoration. Finally, the potter's wheel is a rotating device that, when spun, allows the potter to use centrifugal force and hand pressure to build up vessel walls out of lumps of clay. It appears not to have been utilized in the prehistoric New World, but is known by 4000 B.C.E. in the Near East, South Asia, and China.

Finishing objects involves scraping and smoothing to remove excess materials or irregularities. It may

Fig. 3. Potters from the village of Harappa, Pakistan, display their wares in front of a replica of a two-chambered kiln used by ancient potters in the Indus Valley civilization. (*Courtesy of R. P. Wright*)

also include adding details such as handles.

Decorative effects may be applied either before or after drying the ceramic. Materials such as nets or basketry may be impressed onto the wet ceramic to create a textured surface. Techniques employed after the clays are dried include cutting or incising the surface and applying slips or painted designs. Slips are clay-based coatings to which pigments or colorants have been added to enhance the overall color of the surface of the ceramic. They also may be applied as painted designs to parts of the object when the ceramic is leather-hard.

Drying is a critical step in the production process and may occur at various stages of manufacture. The amount of time necessary is dependent upon a variety of factors, especially weather conditions. Cold or damp weather increases the amount of drying time required.

Firing ceramics makes them hard and relatively watertight. The simplest firing method is bonfires in which ceramics are stacked in an open clearing and fuel (bark, straw, or dung, for example) is placed under, over, and around the objects. Firing temperatures attained in bonfires are relatively low (about 1100–1700°F or 600–900°C), and firing time is of short duration. The kiln is an important technological development because it protects objects from drafts and allows heat to be contained. The two most common types are the pit kiln and double-chambered kiln. Pit kilns are dug into the ground and fuel and ceramics are stacked together, sometimes within an enclosing wall. In double-chambered kilns the fuel and ceramics are placed in separate parts of the kiln. The advantages of double-chambered kilns are the longer duration of firing, higher temperatures reached (2400°F or 1300°C), and added protection during cooling. Updraft kilns were developed in the Near East, South Asia, and China some time around the fifth millennium B.C.E. *See* CLAY, COMMERCIAL; KILN.

Ceramic invention. There is no direct evidence for how ceramic technology was invented. One theory proposes that the techniques involved in cooking were the precursors to pottery production. Processes such as grinding, mixing, kneading, shaping, and firing occur in both cooking and pottery production. In addition, the proximity of clay-lined depressions, ovens, and hearths in household courtyards may have resulted in the accidental firing of mud clay. Other technologies linked to the invention of pottery are basketry, since the technique of coiling was used in both. Mud architecture also has been proposed as a precursor because the building up of vessels in slab construction bears similarities to early mud architecture construction.

Function. Changes in storage and cooking technology were among the many improvements that occurred as a result of the invention of ceramic containers. In contrast to basketry and nonceramic containers in which stone boiling was the major cooking method, the refractory properties of ceramics made it possible to cook food at higher temperatures with sustained heat. In the Near East, the appearance of ceramics generally coincides with broad changes in diet and strategies for obtaining food. Several of the earliest domesticated foods, such as barley, wheat, oats, peas, and sorghum, contain toxins that can be broken down through heating at sufficiently high temperatures. Ceramics also were effective containers for use in brewing and fermentation of grains or fruits because of their durability during fermentation and storage. *See* CERAMICS; POTTERY.

Rita P. Wright

Metals

Scholarly speculation on the origins, development, and economic consequences of metallurgy can be traced back 200 years, but systematic empirical research did not begin until the 1960s with the emergence of the field of archeometallurgy. Archeometallurgists apply techniques from archeology, anthropology, mining engineering, geology, chemistry, and materials science to the study of past mining and metallurgy. Radiocarbon dating, which first became widely available in the 1950s, is of particular importance to archeometallurgy, since prehistoric mines, furnaces, and slag heaps often cannot be dated by other means.

Origins of metallurgy. The empirical record does not support the popular view that the invention of metallurgy rapidly transformed agriculture and warfare. The earliest known metal artifacts, from Cayönü Tepesi in Anatolia, date to about 7000 B.C. and are ornaments of hammered native copper, not functional tools or weapons. Native copper is found in the oxidized upper zones of many copper deposits together with brightly colored copper minerals such as malachite, azurite, dioptase, and turquoise. All of these were used for decorative effect long before the first use of metals, and it appears that copper was initially valued for its color, luster, and malleability.

Metals are very scarce in archeological assemblages dated between 7000 and 5000 B.C. Toward the end of this interval, the first unequivocal evidence of the smelting of metals from ores appears (as small ornaments of metallic lead) in Anatolia, the Near East, and the Aegean. The earliest appearance of metals in quantity is in the Balkans, where the spectacular gold and native copper burial goods in the Varna cemetery and the substantial copper mines of Ai Bunar and Rudna Glava are all dated around 4000 B.C. It is now accepted that metallurgy in the Balkans developed independently from metallurgy in the Near East.

Between 4000 and 2000 B.C., the archeology of the lands around the Black Sea (the Balkans, the Caucasus, modern Turkey, and southern Ukraine) is remarkable for the extraordinary quantities of gold (geochemically the scarcest of the metals used in antiquity) in burial and votive deposits, of which those of Troy II/III (2600–2100 B.C.) are the most familiar to western readers. In Mesopotamia and the Levant, gold was much less common than silver, which was recovered from smelted lead metal by cupellation from about 3300 B.C.

In both the Balkans and the Near East, the smelting of copper from ore developed between 4000 and 3000 B.C. from a crude form of crucible smelting, using relatively pure oxide or carbonate ores, to the use of low shaft furnaces. The latter could attain the temperatures (1150–1300°C; 2100–2370°F) required to form a fully molten slag from the nonmetallic gangue (quartz, clay minerals, and so on) in ores of lower grade, and could produce much more metal per smelt than crucible smelting. These advances were presumably made possible by innovations in the supply of air to the charcoal fuel, probably a transition from lung-powered blowpipes to bag bellows of animal hide.

By 3000 B.C., copper was widely used for tools and weapons, and its use spread rapidly east and west, reaching both Britain and China between 2500 and 2000 B.C. It is still unclear whether the technology itself was spread by migrant metalworkers, or whether the appearance of traded copper objects stimulated local invention of smelting processes. The only regions of sub-Saharan Africa that had metallurgy before 1000 B.C. were the Nile Valley and the Red Sea coast.

In the New World, metallurgy was independently invented in two regions. Native copper was used from about 3000 B.C. in midwestern and eastern North America for ornaments, burial goods, and display of status, but was not smelted before the arrival of European colonists after A.D. 1500. The earliest known metallurgy in the Andes dates to about 1300 B.C. The smelting and alloying of copper by lung-powered blowpipes spread through much of South America, but did not reach Mexico until about A.D. 400, and appears never to have reached southwestern North America. Prehistoric South American metallurgy is notable for its sophisticated treatment of gold, silver, and copper, both as pure metals and as alloys. *See* GOLD METALLURGY.

Copper alloys. In all regions but China, the earliest widespread use of copper is marked by the use of copper-arsenic or copper-antimony alloys. At concentrations of 2–5 wt % arsenic or antimony, these alloys make harder and more durable cutting edges than pure copper; at higher concentrations, they have attractive colors and produce clear musical tones when struck. It is not yet clear whether these alloys were made by deliberate co-smelting of the parent minerals, or whether they were the unintended consequence of smelting complex ores. Either way, it is clear that metalworkers recognized variation in the smelted product and selected appropriate alloys for specific uses.

After 2000 B.C., in the Near East and the Balkans these alloys were replaced by bronze, the alloy of copper with tin. Tin is rarely found in copper ores, so this alloy was deliberately produced by the addition of tin oxide or tin metal to molten copper. There are no major tin deposits in these regions, so the sources of the tin used are still a matter of controversy. *See* COPPER METALLURGY.

Iron and steel. Iron ores are much more abundant than those of copper or tin but are more difficult to reduce to metal. Copper smelters around the Mediterranean typically added iron ore to their furnaces to form a fluid slag; and as smelting techniques improved after 2000 B.C., small amounts of iron metal were sometimes produced with the copper. Yet bronze was preferred to iron, even though tin had to be obtained by long-distance trade. It was not until after the collapse of many Mediterranean civilizations and long-distance trade, around 1200 B.C., that iron began to be widely used.

Unlike copper, iron could not be fully molten in these small furnaces, so the product was small solid lumps (blooms) of iron or steel, which were then forged by the blacksmith to shape. Steel (iron with 0.3–1.5 wt % carbon) was made as early as 1100 B.C., but was not consistently produced until after 500 B.C. in western Eurasia. Iron was first smelted in West and East Africa about 500 B.C., but did not reach southern Africa until about A.D. 400. It was never smelted in the New World before European colonization, though forged iron meteorites have been recovered from Native American sites.

Iron metallurgy in China took a very different turn with the invention about 600 B.C. of the blast furnace to produce liquid cast iron (1.5–4 wt % carbon). The earliest known blast furnaces in Europe are in Sweden and Germany and date to about A.D. 1100. While the former may be a case of diffusion from China, the latter developed from the local bloomery iron smelting tradition. *See* IRON METALLURGY; METALLURGY.

David Killick

Bibliography. D. E. Arnold, *Ceramic Theory and Cultural Process*, 1985; P. T. Craddock, *Early Metal Mining and Production*, 1995; J. Gowlett, *Ascent to Civilization*, 1984; M. Inizan, H. Roche, and J. Tixier, *Technology of Knapped Stone*, 1992; P. M. Rice, *Pottery Analysis: A Sourcebook*, 1987; O. S. Rye, *Pottery Technology: Principles and Reconstruction*, 1981: A. O. Shepard, *Ceramics for the Archaeologist*, 1976; I. Tattersall et al. (eds.), *Encyclopedia of Human Evolution and Prehistory*, vol. 2, 2000; R. F. Tylecote, *The Early History of Metallurgy in Europe*, 1987; R. J. Wenke, *Patterns in Prehistory: Humankind's First Three Million Years*, 3d ed., 1989.

Prehnite

A mineral, $Ca_2(Al,Fe^{3+})(OH)_2[Si_3AlO_{10}]$, with Al in parentheses in octahedral and Al in brackets in tetrahedral coordination by oxygens. The $[Si_3AlO_{10}]$ tetrahedral sheet reflects its perfect platy cleavage. Owing to difficulty in securing suitable single crystals, the atomic arrangement of the mineral had not been known until recently. Prehnite crystallizes in the orthorhombic system, but various space groups have been assigned to it owing to Al-Si ordering, such as Pncm, P2cm, and monoclinic (pseudo-orthorhombic) P2/n.

The mineral usually occurs as stalactitic aggregates

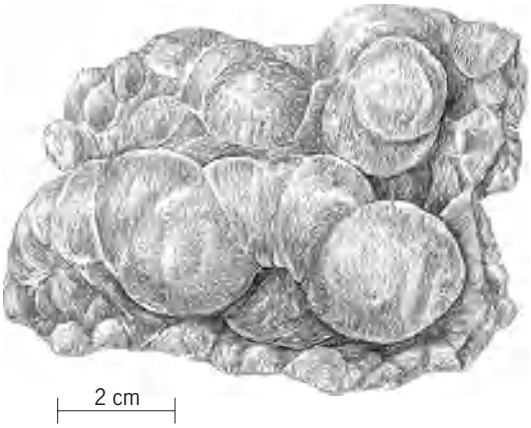

Prehnite in reniform aggregates, Paterson, New Jersey. (*Specimen from Department of Geology, Bryn Mawr College*)

or as curved crystals (see **illus.**), has a vitreous luster, and is yellowish green to pale green in color. Hardness is 6–6$^1/_2$ on Mohs scale; specific gravity 2.8–2.9. Common occurrences include vesicular basalts such as the Keweenaw basalts in the Upper Peninsula of Michigan, and the Watchung basalts in New Jersey. Associated minerals include zeolites, calcite, drusy quartz, datolite, and pectolite. *See* SILICATE MINERALS.

Paul B. Moore

Prenatal diagnosis

The identification of disease before the birth of a fetus. It often implies genetic diagnosis, but identification of anatomical defects as well as assessment of fetal functions and maturity are also considered. Some of the relatively common diseases that can be diagnosed prenatally are Tay-Sachs disease, cystic fibrosis, Duchenne's muscular dystrophy, hemophilia A, congenital adrenal hyperplasia, thalassemia, and sickle cell anemia.

Ultrasound imaging. Ultrasound is a method of tissue imaging that is based on analysis of reflected high-frequency sound waves directed at the tissue. This method has revolutionized obstetrics perhaps more than any other technology. Ultrasonic data have vastly improved the understanding of normal growth and development, thus permitting earlier and more accurate diagnosis of fetal disease. Fetal movements can also be observed, allowing assessment of functional well-being. Such information has enabled physicians to evaluate the fetus itself rather than relying on indirect signs that can be detected by examining the mother. This change, in turn, has made it practical to consider the fetus as a patient for whom prompt newborn treatment can be planned. The fetus can even be considered a candidate for surgical therapy or medical therapy before birth.

The benefits of ultrasound imaging of the fetus for patient care include the identification of the number of fetuses and their position in the womb; a more accurate determination of fetal (gestational) age, so that fetal growth retardation and postterm (overdue) gestation are diagnosed more accurately; the ability to diagnose most fetal anomalies with obvious structural or functional implications; and the improved assessment of associated structures, such as the placenta, umbilical cord, and amniotic fluid. *See* MEDICAL IMAGING; MEDICAL ULTRASONIC TOMOGRAPHY.

Assessment of fetal well-being. A traditional way of monitoring the baby's tolerance of labor was to listen to the mother's abdomen with a stethoscope. Modern technology has made possible the direct measurement of fetal heart rate, using techniques such as electrocardiography. Monitoring during labor consists of analyzing fetal heart rate changes occurring in response to uterine contractions. Certain abnormal patterns in the fetal heart rate may suggest possible asphyxia (inadequate delivery of oxygen to tissues and inadequate removal of carbon dioxide).

Basic surveillance of fetal status before labor is often accomplished in part by having pregnant women count the number of fetal movements over a given time. In addition, nonstress testing is performed by checking for fetal heart rate accelerations following fetal movements. Contraction-stress testing looks at fetal heart rate changes following uterine contractions, which can occur either spontaneously or in response to the hormone oxytocin.

Another measure of fetal well-being is the biophysical profile. It is an index of chronic (long-standing) asphyxia and combines nonstress testing with ultrasonic measurement of fetal body tone and movements, fetal breathing measurements, and amniotic fluid volume. Such tests may indicate a high likelihood of fetal asphyxia and may prompt the clinician to use more direct methods of analysis such as a sample of fetal blood obtained from the fetal skin during labor or from the umbilical cord.

Assessment of fetal lung maturity. Respiratory distress syndrome of the newborn is the major cause of illness of preterm infants. Many cases result from an inadequacy of surfactant, a fatty substance that is normally produced by mature fetal lungs at term and that permits lung expansion after birth. Unfortunately, gestational age is not always a good predictor of lung maturity. However, chemical analysis of amniotic fluid for surfactant produced by the fetus is possible. Certain values for the surfactant phospholipids strongly suggest functional lung maturity, thus making respiratory distress syndrome very unlikely. If for medical reasons the fetus must be delivered before the lungs have reached maturity, steroid hormones may be used to induce surfactant production. *See* INFANT RESPIRATORY DISTRESS SYNDROME.

Genetic diagnosis. Obstetricians typically review the health history of the pregnant woman, the father-to-be, and their families in order to identify any possible heritable disorders in the families. Certain risks related to the patient's age, race, and geographic origin may be noted. Based on these assessments, genetic testing of the unborn child may be discussed or recommended.

Using cells from the fetus, a prenatal genetic diagnosis can be made for at least 10% of those disorders that are known or assumed to result from a gene mutation. Such diagnosis is based on deoxyribonucleic acid (DNA) analysis or on detection of an abnormal enzyme or other protein produced by the defective gene. Expert counseling must be available to the parents to ensure complete understanding of the strengths, limitations, and results of any testing procedure on the fetus. *See* HUMAN GENETICS.

Amniocentesis. Once the feasibility of withdrawing a small amount of amniotic fluid was demonstrated for the management of Rh incompatibility, it was realized that genetic diagnosis could be based on analysis either of fetal biochemical products in the fluid or of living fetal cells contained therein while posing little or no risk to mother or fetus. In amniocentesis, 20–30 cm^3 of amniotic fluid is aspirated with a long needle and syringe. This procedure is performed at about 15–17 weeks' gestation. If multiple fetuses are present, the gestational sacs can be sampled separately. *See* RH INCOMPATIBILITY.

Chorionic villus sampling. At 9–11 weeks' gestational age, a thin catheter may be inserted under ultrasound guidance through the uterine cervix and into the placenta, where about 10–25 mg of chorionic tissue is aspirated. An alternate method for obtaining the specimen involves the insertion of the catheter via a needle through the abdominal and uterine walls before obtaining the placental specimen. This technique permits a biopsy of the placenta, which is primarily of fetal origin, thus providing living fetal cells for genetic analysis.

Applications. As more precise prenatal diagnosis becomes available, the opportunities for fetal therapy continue to expand. Imaging techniques, such as magnetic resonance imaging, may facilitate visualization of elusive structures, such as the fetal brain, and may even permit assessment of metabolic well-being. Refinements in ultrasound techniques will permit more refined fetal surgery. Fetal distress may become classifiable by its cause, permitting rational intrauterine therapy, such as nutritional treatment for growth-retarded fetuses. Placental blood-flow compromise in maternal conditions such as hypertension, and fetal circulatory disturbances accompanying chronic asphyxia, may be detectable with Doppler ultrasound, and noninvasive assessment of fetal lung maturity by ultrasound or other techniques may become possible. Finally, very early prenatal diagnosis may become feasible by biopsying the preimplantation embryo when genetic disease is likely. Gene therapy of abnormal embryos may also become a clinical possibility. *See* CONGENITAL ANOMALIES; PREGNANCY.

Donald McNellis

Bibliography. R. K. Creasy and R. Resnik (eds.), *Maternal-Fetal Medicine*: *Principles and Practice*, 4th ed., 1999; K. C. Due, *Fetal Echocardiography*: *Principles and Applications*, 1988; R. G. Edwards and P. Hollands, New advances in human embryology: Implications of the preimplantation diagnosis of genetic disease, *Hum. Reprod.*, 3(4):549–556, 1988; S. G. Gabbe, J. F. Niebyl, and J. L. Simpson (eds.), *Obstetrics*: *Normal and Problem Pregnancies*, 3d ed., 1996; D. Maulik and D. McNellis, *Doppler Ultrasound Measurement of Maternal-Fetal Hemodynamics*, 1987.

Press fit

A force fit that has negative allowance; that is, the bore in the fitted member is smaller than the shaft which is pressed into the bore. Tight fits have slight negative allowance so that light pressure is required to assemble the parts; they are used for gears, pulleys, cranks, and rocker arms. Medium force fits have somewhat greater negative allowance and require considerable pressure for assembly; they are used for fastening locomotive wheels, car wheels, and motor armatures. *See* ALLOWANCE; FORCE FIT; SHRINK FIT.

Paul H. Black

Pressure

Pressure on a surface immersed in a fluid is defined as the normal force per unit area. Pressure always tends to compress the object on which it acts. This definition identifies the so-called static pressure p.

Another distinction is made between the absolute pressure, which is measured with respect to a perfect vacuum (that is, zero pressure), and the gauge pressure, which is measured with respect to the atmospheric value. In the International System of Units (SI), the pressure unit is 1 Pa (pascal), corresponding to a force of 1 N (1 newton = 1 kg m/s^2) acting on a surface area of 1 m^2. In the English system, the pressure unit is 1 psi (pound per square inch), corresponding to a force of 1 lbf acting on a surface area of 1 $in.^2$; 1 psi $\approx$ 6895 Pa. Other common units of pressure are 1 bar = 100,000 Pa, 1 atmosphere (atm) = 101,325 Pa, 1 in. $H_2O \approx 249$ Pa, and 1 torr (1 mm Hg) $\approx$ 133 Pa. The standard atmospheric pressure at sea level is 101,325 Pa, whereas the pressure in interstellar space approaches zero. When considering mixtures of ideal gases (that is, such that each component does not influence the others), one may define the partial pressure of each component by assuming that it occupies the entire volume separately from the others; then the pressure of the mixture would be the sum of the partial pressures of all components. *See* DALTON'S LAW; FORCE; GAS; METRIC SYSTEM; PHYSICAL MEASUREMENT; UNITS OF MEASUREMENT.

In static fluids with uniform density, the pressure difference between two points at different elevations is called the hydrostatic pressure difference; it increases downward and is equal to $\rho g \Delta h$ ($g \approx 9.81$ $m/s^2 \approx 32.2$ ft/s^2 is the gravitational acceleration and Δh is the elevation difference). To define a "static" pressure within a moving fluid, one must average the normal stresses in three orthogonal directions. In addition to their static pressure, flowing fluids are characterized by their total pressure $p_o = p + 1/2\rho V^2$ (ρ is the density and V is the flow

velocity). This is the pressure that would be achieved if the fluid were brought to rest, as would be the case if a fluid jet impacted on a wall. The term $^1/_2\rho V^2$ is referred to as dynamic pressure. *See* PRESSURE MEASUREMENT.

Stavros Tavoularis

Pressure measurement

The determination of the magnitude of the normal force applied by a fluid to a unit area. Pressure measurements are generally classified as absolute pressure, gauge pressure, or differential pressure. Absolute pressure is the pressure referenced to perfect vacuum, which by definition has zero pressure. Gauge pressure is the difference between an absolute pressure and the pressure of the atmosphere. Differential pressure is the difference between any two pressures, neither of which is atmospheric. *See* PRESSURE.

Pressure measurement is one of the most common activities in the laboratory, whether for educational purposes or for research, because the pressure level has a significant effect on most physical, chemical, and biological processes. Pressure levels of a great variety of fluids, including air, water, steam, and oil, are measured routinely in all power, manufacturing, and process industries as well as in powered vehicles of most types. In the process industries, pressure is measured and controlled to maintain uniformity of product, to guide in safe plant operation, to determine pumping head for fluid transfer, and to measure other variables indirectly, including weight, liquid level, temperature, flow and density of fluids, and hydraulic forces.

Pressure can be measured by a variety of instruments whose operation is based on different principles and according to which they can be classified as liquid-column manometers, deadweight pressure gauges, mechanical pressure gauges, and electrical pressure transducers. Each of these general categories is subdivided into more specific types. With the exception of the liquid-column manometers and the deadweight pressure gauges, whose output can be related to pressure from first principles, all other pressure gauges and transducers have an empirical response, which means that they need to be calibrated versus a standard (usually a manometer or a deadweight pressure gauge). Factory calibration is normally provided, but it may be necessary to check the instrument's response at regular intervals, as it may be affected by temperature changes, humidity, vibrations, and other disturbances.

The term pressure gauge commonly signifies a self-contained instrument that senses pressure and indicates its value, usually by some mechanical linkage on a graduated scale. The term pressure transducer may also be used to identify some pressure gauges, but usually applies to more complex instruments, which comprise a pressure-sensing element; a device that converts motion or force produced by the sensing element to a change of an electrical, mechanical, or pneumatic parameter; and sometimes components for the conditioning of the output, such as temperature compensation. A transducer may include an indicating or recording device or may generate a pneumatic output or an electric signal that can be used as input to other devices for recording or control purposes. A related term is pressure transmitter. Although pneumatic and mechanical transducers are commonly used, electrical measurement of pressure is often preferred because of a need for long-distance transmission, higher accuracy requirements, more favorable economics, or quicker response. Especially for control applications, pneumatic pressure signal transmission may be desirable over electrical where flammable materials are present. *See* PROCESS CONTROL.

Liquid-column manometers. A liquid-column manometer consists of a vertically oriented U-shaped tube, which is partly filled with a nonvolatile liquid. Manometric fluids are color-coded and most commonly consist of water, alcohol, or various mineral oils, while the previously common use of mercury is currently restricted by safety regulations.

If one leg is left open to the atmosphere and the other one is connected to a gas (**Fig. 1**), the difference h (head) in liquid level between the two legs is a direct measure of gauge pressure p_g in the gas, according to the relationship

$$P_g = \rho g h,$$

where ρ is the density of the liquid and $g \approx 9.81$ m/s^2 ≈ 32.2 ft/s^2 is the gravitational acceleration. If each leg is connected to a gas, the manometer measures the differential, rather than the gauge, pressure. The level rises in the low-pressure leg and drops in the high-pressure leg.

If the manometer is connected to a liquid (which must be immiscible with the manometric fluid) instead of a gas, the relationship $p_g = \rho g h$ has to be corrected for the pressure corresponding to the column of the experimental fluid. The smaller is the difference between the densities of the manometric and experimental fluids, the higher is the sensitivity of the manometer, which means that a higher column is generated by a given pressure difference. Thus, for measuring air pressure, an alcohol manometer is more sensitive than a water manometer, which in turn is more sensitive than a mercury manometer. Other arrangements have been devised to increase the sensitivity of the manometer or the ease of reading its output. A common type is the inclined

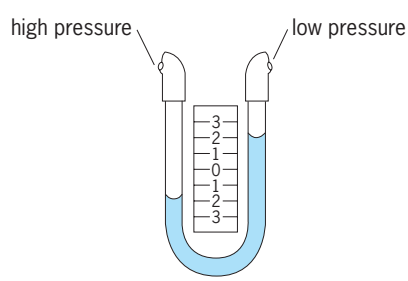

Fig. 1. Liquid-column (U-tube) manometer.

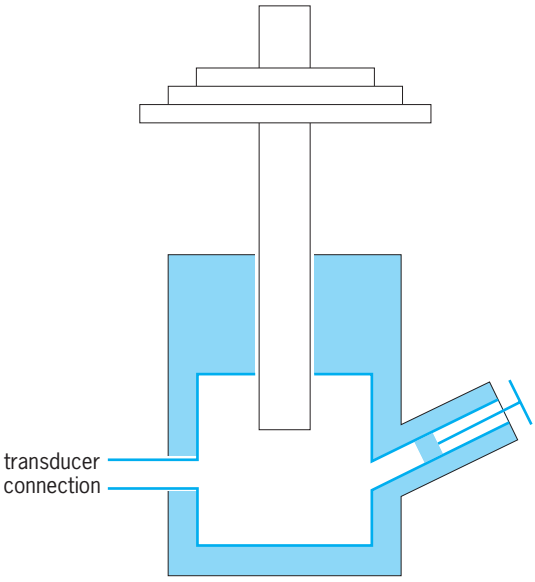

Fig. 2. Deadweight pressure gauge.

manometer, in which the tube is oriented by an angle to the vertical direction, such that the elevation of the liquid surface is read along an inclined scale.

A special type of liquid-column manometer is the mercury barometer. It consists of a single vertical glass tube, whose top part is evacuated and sealed, while its bottom end is immersed in a small tank of mercury, which is open to the atmosphere. The height of the rising column of mercury measures the absolute atmospheric pressure. *See* BAROMETER.

Deadweight pressure gauges. These gauges (**Fig. 2**) are not normally used to monitor pressure, but serve as standards for the calibration of other pressure instrumentation because they are very precise. The pressure is measured directly as the ratio of a precisely known weight required to balance a piston, whose end is immersed in a pressurized hydraulic fluid, and the piston's precisely known area. The pressure gauge or transducer that is to be checked or calibrated is connected to the hydraulic reservoir. Weights corresponding to the desired pressure are placed on the piston, and the hydraulic pressure is adjusted with the use of a pump. The piston and weights are rotated to reduce the effect of friction.

Mechanical pressure gauges. These are in wide use throughout industry due to their low cost, sturdiness, and freedom from the operational limitations of liquid manometers. There are three common elements that serve as sensors of pressure differences: Bourdon, diaphragm, and bellows. These elements may be designed to produce either motion or force under applied pressure. The more common motion type may directly position the pointer of a concentric indicating gauge; position a linkage to operate a recording pen, or a pneumatic relaying system to convert the measurement into a pneumatic signal; or position an electrical transducer to convert to an electrical signal. Depending on its design, it may measure absolute, gauge, or differential pressures. Accuracies vary depending on materials, design, and precision of components.

Bourdon-spring gages (**Fig. 3**), in which pressure acts on a shaped, flattened, elastic tube, are by far the most widely used type of instrument. When the tube is pressurized, it deforms such that its cross-sectional shape tends toward a circular one. These gauges are simple, rugged, and inexpensive. Common types have a C-, spiral-, helix-, or twisted-tube element.

In diaphragm-element gauges, pressure applied to one or more contoured diaphragm disks acts against a spring or against the elastic deformation of the diaphragms, producing a measurable motion (**Fig. 4**). Size, number, and thickness of the disks determine the range.

In bellows-element gauges, pressure in or around the bellows moves the end plate of the bellows against a calibrated spring, producing a measurable motion (**Fig. 5**).

Electrical pressure transducers. Such transducers convert the pressure sensed by an element to an electrical signal by a resulting change of an electrical property, including resistance, capacitance, or inductance.

Variable-resistance pressure transducers. Pressure is measured in these transducers by an element that changes its electrical resistance as a function of pressure.

Many types of such transducers use a movable contact, positioned by the pressure-sensing element. One form is a contact sliding along a continuous resistor, which may be straight-wire, wire-wound, or nonmetallic such as carbon. If the cross section of the resistor is constant, the change in resistance is proportional to the motion of the contact. The cross

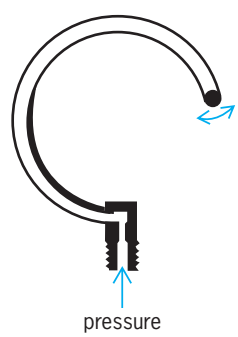

Fig. 3. Bourdon-spring gauge with C-tube element.

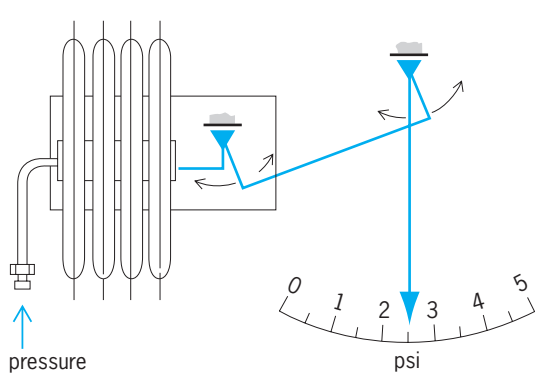

Fig. 4. Diaphragm-element gauge.

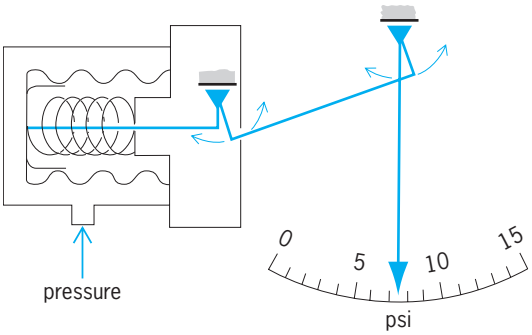

Fig. 5. Bellows-element gauge.

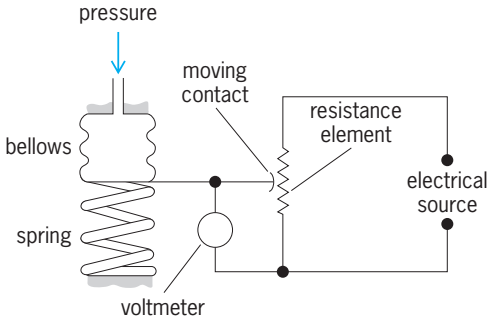

Fig. 6. Variable-resistance pressure transducer.

section may be made nonuniform to give a nonlinear relation between motion and change of resistance. The resistance element may be curved or part of an arc for convenience in measuring angular motion.

Figure 6 shows one type of resistive pressure transducer. A bellows opposed by a precisely designed spring senses the pressure and converts the pressure to a linear motion of the plate between the bellows and the spring. The plate bears a contact which wipes the surface of the precision wire-wound resistor. If a constant potential (AC or DC) is maintained across the resistor and if the resistance of the voltmeter is high with respect to the resistor, the measured voltage is a precise measure of the pressure. *See* ELECTRICAL RESISTANCE; RESISTANCE MEASUREMENT.

Strain-gauge pressure transducers. These might be considered to be resistive transducers but are usually classified separately. Strain-gauge pressure transducers convert a physical displacement into an electrical signal. When a wire is placed in tension, its electrical resistance increases. The change in resistance is a measure of the displacement, hence of the pressure. Its advantages include high resolution and small size. The strain gauge is used in conjunction with a bridge circuit. Often four resistors are connected in a bridge circuit so that an applied pressure leaves one pair of resistors in compression and the other pair in tension, yielding the maximum output change. *See* BRIDGE CIRCUIT; STRAIN GAUGE.

Variable-reluctance pressure transducers. In this type, a change of pressure is converted into change of magnetic reluctance when one part of a magnetic circuit is moved by a pressure-sensing element—Bourdon tube, bellows, or diaphragm. Motion produces a change of magnetic reluctance (reluctance is the resistance of some material to the transmission of a magnetic field) in a magnetic circuit, which is directly related to pressure. The change of reluctance is usually within one or two coils, wound about the magnetic core in the magnetic circuit. *See* MAGNETISM; RELUCTANCE.

A representative reluctance-changing device is shown in **Fig. 7**. A Bourdon-type or other pressure-sensing device tilts the armature. The reluctances in the magnetic paths A and B are determined chiefly by the lengths of the air gaps between the armature and the core. The inductance and inductive reactance of each winding depend on the reluctance in its magnetic path. *See* INDUCTANCE; REACTANCE.

If the armature is at a neutral symmetrical position, the air gaps are equal, and the inductive reactances X_{LA} and X_{LB} are equal. Change of pressure decreases one air gap and increases the other, thus changing the ratio of the inductive reactances X_{LA} and X_{LB}. These changes can be used in a variety of circuits to produce an electrical signal that is a measure of pressure.

Variable-inductance pressure transducer. A change in inductance and inductive reactance of one or more windings is produced by the movement of a magnetic core that is positioned by a Bourdon tube or other pressure-sensing element. Unlike the action of a reluctance-type transducer, the inductance change is caused by a change in air gap within the winding, rather than in a relatively remote portion of the magnetic circuit.

Figure 8 shows a representative ratio-type inductive device. The pressure-sensing element moves the core in response to changes of pressure. When the core is in a central position, the inductances of the two coils are equal. When a pressure change moves the core, the ratio of the two inductances is changed. Energy is supplied to the coils by the same bridge circuit that measures the ratio of inductances.

Another form of inductive pressure transducer is

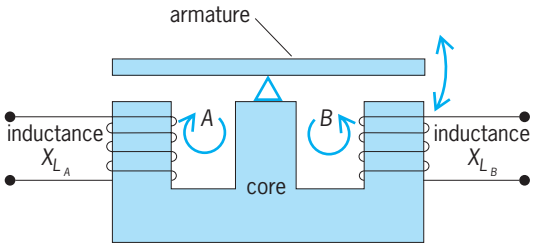

Fig. 7. Variable-reluctance pressure transducer.

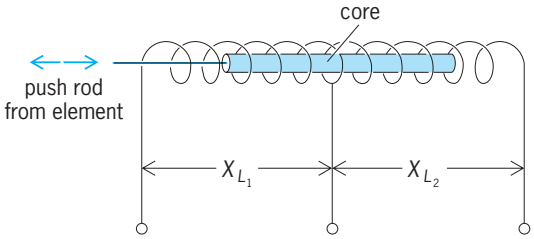

Fig. 8. Inductive pressure transducer.

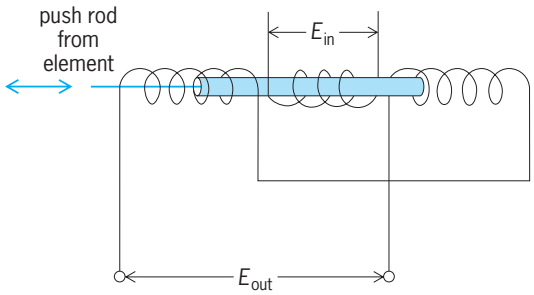

Fig. 9. Linear variable differential transformer.

the linear variable differential transformer (LVDT; **Fig. 9**). When the core is centered, equal voltages are induced in two oppositely wound secondary windings and the output voltage is zero. A change of pressure moves the core, increasing the voltage induced in one secondary and decreasing the voltage induced in the other. The change in output (differential) voltage is thus a measure of the pressure.

Piezoelectric pressure transducers. Some crystals produce an electric potential when placed under stress by a pressure-sensing element. The stress must be carefully oriented with respect to a specific axis of the crystal. *See* PIEZOELECTRICITY.

Suitable crystals include naturally occurring quartz and tourmaline, and synthetic crystals such as Rochelle salts and barium titanate. The natural crystals are more rugged and less subject to drift. Although the synthetic crystals offer much higher voltage output, an amplifier is usually required for both types. Crystal transducers offer a high speed of response and are widely used for dynamic pressure measurements in such applications as ballistics and engine pressures.

Variable-capacitance pressure transducers. Almost invariably, these sense pressure by means of a metallic diaphragm, which is also used as one plate of a capacitor. Any variation in pressure changes the distance between the diaphragm and the other plate or plates, thereby changing the electrical capacitance of the system. The change in capacitance can be used to modify the amplitude of an electrical signal. *See* CAPACITANCE.

Semiconductor pressure transducers. Technologies, such as photolithography, developed for the manufacturing of integrated circuits have also been applied to the fabrication of silicon-based pressure transducers. Pressure is measured by its relationship to a piezoelectric potential or to a change in resistance or capacitance affected by the deformation of a semiconductor film. The great advantage of these transducers is their small size: miniature transducers are available with sizes as small as 1 mm (0.04 in.), while much smaller ones, with sizes comparable to 1 μm, are referred to as MEMS (microelectro-mechanical) transducers. The latter can be manufactured in large, relatively inexpensive arrays, and they are very sensitive to pressure fluctuations. *See* MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS); MICROLITHOGRAPHY.

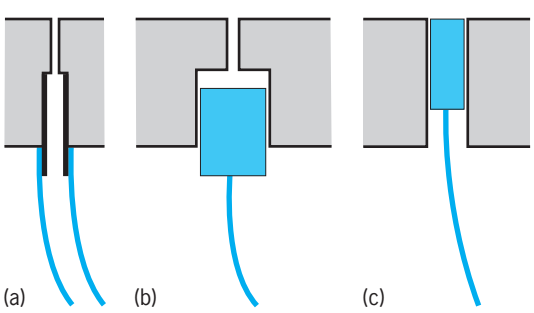

Fig. 10. Pressure taps and types of connections to transducers. (*a*) Plastic tubing connected to transducer. (*b*) Transducer in cavity. (*d*) Transducer flush with wall.

Pressure-sensing in pipes. When measuring the pressure of fluids flowing in pipes and channels, it is important not to disturb the fluid by the measuring instrumentation. Most commonly, small orifices [pressure taps, 1 mm (0.04 in.) or less, in diameter] are drilled through the wall, and the flow is connected to a manometer or pressure transducer via plastic tubing (**Fig. 10*a***). The orifice has to be free of burrs and perpendicular to the wall. This method cannot measure accurately pressure transients of relatively fast pressure fluctuations. To improve the response, a small transducer may be embedded in a small cavity near the orifice (Fig. 10*b*) or flush with the internal wall (Fig. 10*c*). *See* PHYSICAL MEASUREMENT; VACUUM MEASUREMENT.

Stavros Tavoularis; John H. Zifcak

Bibliography. R. P. Benedict, *Fundamentals of Temperature, Pressure and Flow Measurement*, 3d ed., 1984; E. O. Doebelin, *Measurement Systems*, 5th ed., McGraw-Hill, 2004; P. W. Harland (ed.), *Pressure Gauge Handbook*, Marcel Dekker, 1985; E. Ower and R. C. Pankhurst, *The Measurement of Air Flow*, 5th ed., Pergamon Press, 1977; S. Tavoularis, *Measurement in Fluid Mechanics*, Cambridge University Press, 2005; J. G. Webster (ed.), *Mechanical Variables Measurement*, CRC Press, 2000.

Pressure seal

A seal is used to make pressure-proof the interface (contacting surfaces) between two parts that have frequent or continual relative rotational or translational motion; such seals are known as dynamic seals, as compared with static seals. While the pressure in seals is lower than that in gaskets, the motion hinders their effectiveness so that there are more types of seals than gaskets, each type attempting to serve its environment. The materials are leather, rubber, cotton, and flax, and for piston rings, cast iron. The forms of nonmetallic seals are rectangular, V ring, and O ring. Cartridge seals are available for rolling-contact bearings. Special seals include carbon ring and labyrinth seals for turbines and mechanical seals for pumps. *See* GASKET; STEAM TURBINE.

Paul H. Black

Pressure vessel

A cylindrical or spherical shell structure capable of withstanding internal pressure and external loadings at nozzles, supports, and attachments. The pressure vessel is designed with consideration for all possible modes of failures (such as fatigue, fracture, or collapse) and with provision of rational margins of safety against each type of failure. The Boiler and Pressure Vessel Code has been developed by the American Society of Mechanical Engineers (ASME), and the Code has been widely accepted as a standard for designing pressure vessels.

The pressure vessel may be subdivided into four regions: closure (upper) head, nozzle-shell course, beltline, and lower head regions (see **illus.**).

Heads. There are three basic types of heads: flat heads, conical and toriconical heads, and dished heads (hemispherical, ellipsoidal, or torispherical, a configuration consisting of one spherical part and one toroidal part). The highest elastic stress in an integral flat head occurs at the junction between the cylinder and the head. When the head is thicker than the cylinder (the usual case), this high stress is always in the cylinder. The highest stress in a cylindrical pressure vessel with a conical head is the axial stress at the junction. This is similar to the situation in a flat head. The conical head, unlike the flat head, also has large hoop stresses at the junction due to the tendency of the base of the cone to deflect inward and the cylinder to deflect outward when internal pressure is applied. The toroidal transition will help to reduce the critical stresses at the junction. The hemispherical head may be considered as a special case of the ellipsoidal head, and the torispherical head may be designed as an approximately ellipsoidal head. As the head becomes flatter and the shell becomes thinner, the possible weakness of the dished head begins to appear at the knuckle and at the junction to the shell, but even so, the stresses are markedly lower than those found in flat or conical heads. An important characteristic of dished heads is that for the flatter shapes and thinner walls there is a tendency to develop relatively high circumferential compression in the knuckle. Since this is accompanied by axial tension, there is a strong tendency toward yielding at this location, which has an important effect on the stability of the head.

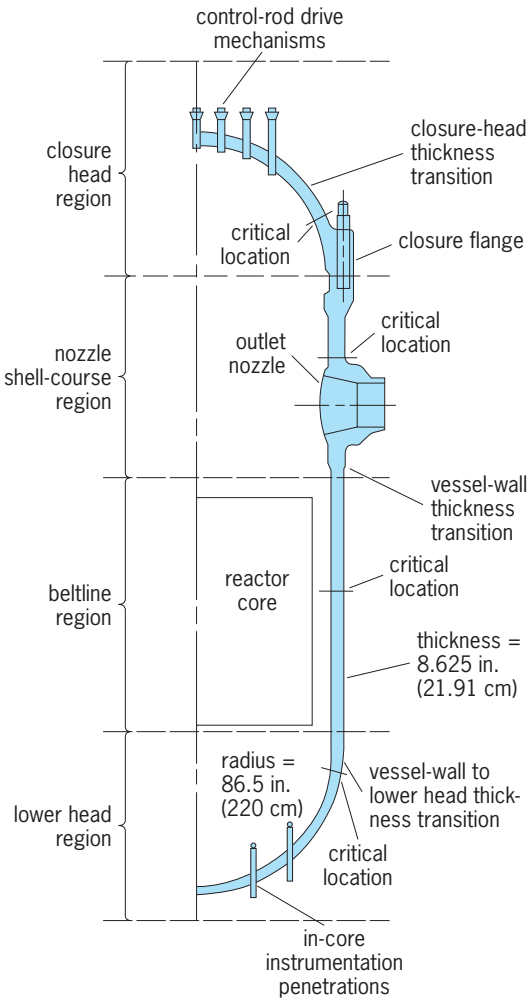

Diagram of a typical pressurized water reactor vessel. (*After Pressure Vessel and Piping: Design and Analysis, A Decade of Progress, vol. 1, p. 326, American Society of Mechanical Engineers, 1972*)

Beltline. The required thickness of a cylindrical or spherical vessel under internal pressure loading is based on the criterion that the maximum hoop stress must not exceed the alllowable value at operating temperatures. Since the allowable stress never exceeds 25% of the ultimate tensile strength of the material, the implication is that there is a safety factor of 4 against bursting due to excess pressure. Since ductile metals are used for the majority of pressure vessels, and the bursting of a ductile pressure vessel involves large plastic strains, it is recommended that plastic analysis be performed to determine the wall thickness.

Nozzle-shell course. Most pressure vessels are elements of systems and, as such, are subjected to loadings at the supports, nozzles, and other attachments. These loadings may sometimes produce stresses of greater importance than those produced by pressure. The higher stress concentration occurs at the junction between the nozzle and the shell. Finite element method of analysis provides good stress and deformation estimates around the nozzle. *See* BOILER; FINITE ELEMENT METHOD; HEAT EXCHANGER; HIGH-PRESSURE PROCESSES; NUCLEAR REACTOR. Y. S. Shin

Bibliography. J. S. Cheung and L. S. Ong, *Pressure Vessel and Piping Technology*, 1993.

Pressurized blast furnace

A blast furnace operated under higher than normal pressure. The pressure is obtained by throttling the off-gas line, which permits a greater volume of air to be passed through the furnace at lower velocity and results in an increasing smelting rate. *See* IRON METALLURGY; PYROMETALLURGY, NONFERROUS.

The process advanced markedly after its introduction in the United States in 1943, and is now

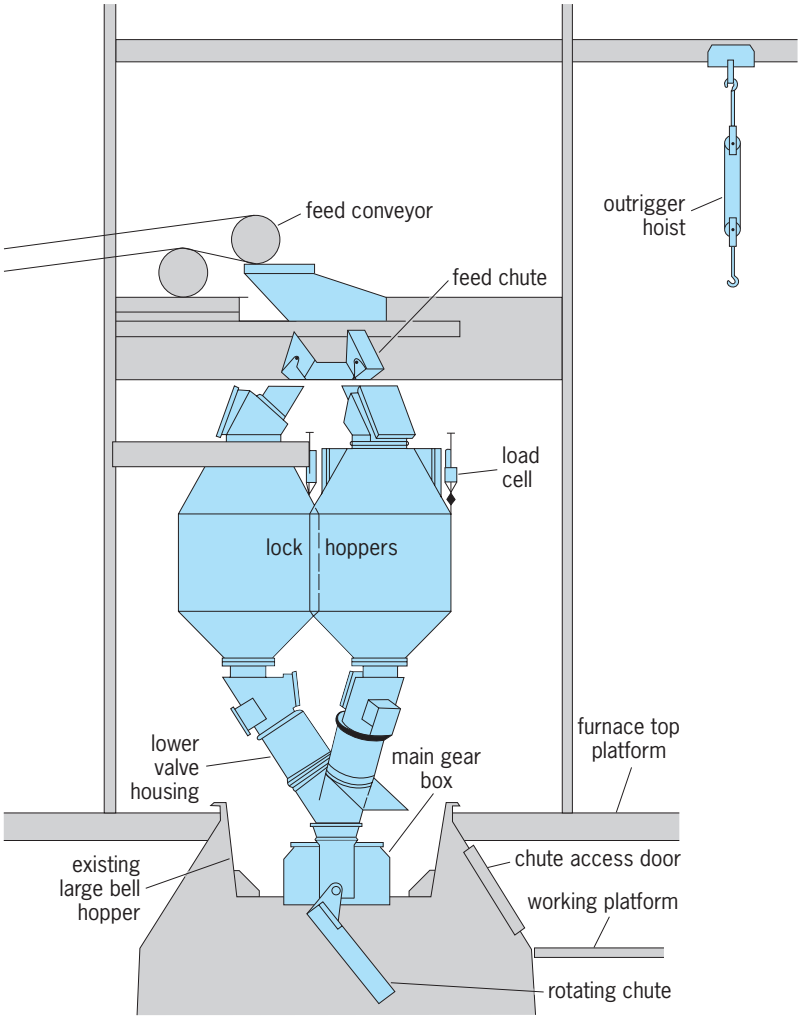

Diagram of a pressurized blast furnace with a Wurth top. (*After R. A. Powell, The blast furnace top of the future, Iron Steel Eng., 50(1):86–90, January 1973*)

utilized on all modern blast furnaces in the world. High-top-pressure operation has made possible the design of mammoth blast furnaces having over 140,000 ft^3 (4000 m^3) inner volume and hearth diameters of over 45 ft (14 m); they produce over 11,000 tons/day (10,000 metric tons/day) compared with 1540 tons/day (1400 metric tons/day) in 1943. The process permits large increases in the weight of high-temperature air blown into the bottom of the furnace at lower gas velocities, thus increasing the rate of smelting and decreasing the rate of coke consumption, and also permitting smoother operation with less flue dust production through decreased pressure drop between bottom and top pressures.

Many improvements have also been made in the mechanical features of the flow diagram in the **illustration.** Novel charging mechanisms, such as the Wurth top (lockhoppers with rotating shute), have essentially eliminated wear maintenance problems and improved the distribution of ore, coke, and limestone charged into the furnace. Also, at top pressures of 43 lb/in.^{2} (3 kg/cm^2), the throttling valve has been replaced by an expansion turbine to recover power from the top gases. Instrumentation and controls have also been vastly improved. *See* FURNACE CONSTRUCTION.

Bruce S. Old

Bibliography. A. Chatterjee, *Beyond the Blast Furnace*, 1993; W. G. Davenport, *Flash Smelting: Analysis, Control, and Optimization*, 1987; M. Higuchi, M. Ilzuka, and T. Shibuya, High top pressure operation of blast furnaces at Nippon Kokan K. K., *J. Iron Steel Inst.*, September 1973; Iron and Steel Institute of Japan (ed.), *Blast Furnace Phenomena and Modelling*, 1987; B. S. Old, E. L. Pepper, and E. R. Poor, Progress in high pressure operations of blast furnaces, *Iron Steel Eng.*, 25(5):37–43, 1948; O. D. Rice, New blast furnace design boosts iron yields, *Iron Age*, 185(4):96–98, 1960.

Prestressed concrete

Concrete with stresses induced in it before use so as to counteract stresses that will be produced by loads. Prestress is most effective with concrete, which is weak in tension, when the stresses induced are compressive. One way to produce compressive prestress is to place a concrete member between two

abutments, with jacks between its ends and the abutments, and to apply pressure with the jacks. The most common way is to stretch steel bars or wires, called tendons, and to anchor them to the concrete; when they try to regain their initial length, the concrete resists and is prestressed. The tendons may be stretched with jacks or by electrical heating.

Prestressed concrete is particularly advantageous for beams. It permits steel to be used at stresses several times larger than those permitted for reinforcing bars. It permits high-strength concrete to be used economically, for in designing a member with reinforced concrete, all concrete below the neutral axis is considered to be in tension and cracked, and therefore ineffective, whereas the full cross section of a prestressed concrete beam is effective in bending. *See* CONCRETE BEAM.

An especially desirable characteristic of prestressed concrete is that as long as the material is maintained in compression it cannot crack. If cracks should appear under small overload, they generally will close when the load is removed. Sometimes concrete is prestressed principally to prevent cracking.

Basic principles. The effect of compressive prestresses is like picking up a pile of books by applying pressure to the end pair. If the pressure is large enough, none of the books will slip out.

If this concept were applied to a concrete beam in actual practice, steel tendons would be tensioned and placed along the centroidal axis of the beam. The resulting prestress would result in a uniform compression at every section (**Fig. 1**). Loads would produce both tensile and compressive stresses at the middle of the span. The prestress would combine with these to increase the compression and cancel out the tension. The whole concrete section would be effective in resisting bending, and there would be no cracks. *See* CONCRETE BEAM.

In practice, however, tendons are rarely placed along the centroidal axis. A smaller prestressing force is required, and therefore less steel for the tendons, if the steel is placed below the centroidal axis of the beam. With the eccentric prestress, stresses at each section of the unloaded beam may vary from tension at the top to compression at the bottom (**Fig. 2**).

When loads are applied to the beam, they produce both tensile and compressive stresses at the middle of the span. At the top of the beam they cause compressive stresses, which are reduced by the tensile

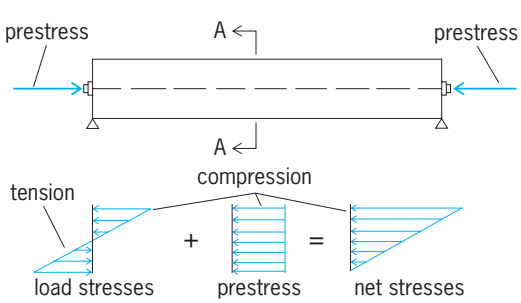

Fig. 1. Diagram showing stresses at section A-A of a concrete beam with uniform prestress.

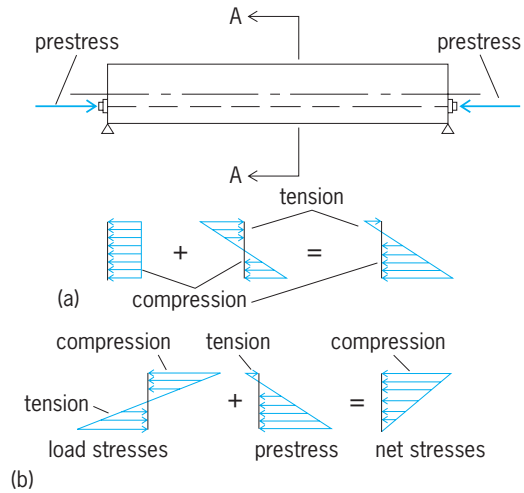

Fig. 2. Stresses at section A-A of a beam with eccentric prestress. (*a*) In unloaded beam, simple-bending stress component is largely counteracted by uniform compression component. (*b*) In loaded beam, tension components are counteracted and only compression remains.

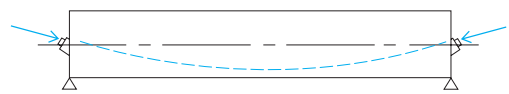

Fig. 3. Beam with tendons draped in a vertical curve.

prestress there. Elsewhere, the tensile stresses produced by the loads are counteracted by the compressive prestress.

With this arrangement of the tendons, there is a possibility that near the ends of the beam the tensile prestress may exceed the compressive stresses produced by the loads. The net tension may be undesirable, even though very small. To avoid this condition, the tendons may be draped in a vertical curve (**Fig. 3**). The distribution of prestress at any section of a beam so prestressed is similar to that for straight tendons applying an eccentric prestress except that the stresses decrease from midspan toward the ends, as do the bending stresses due to the loads. The draped arrangement of the tendons also is advantageous in counteracting diagonal tension near the beam ends.

Continuous beams may be prestressed in a similar manner. The tendons may be placed near the bottom of the beams near midspan and near the top over supports.

Tendons. Tendons generally are made of high-strength steel so that they can serve at high working stresses. The reason for this is that losses in stress due to shrinkage and plastic flow of the concrete are relatively high. If the tendons were given a small tension, they might lose nearly all the prestress in a few months. But at a high tension, the loss might be only about 15%, because the increase in stress loss is smaller than the increase in prestress.

Wires are used for prestressing much more frequently than bars because of their greater strength. Wires may be used singly, in pairs, in cables composed of several parallel wires, or in strands. They may be stretched by electric heating, but by far the

most common method of tensioning is with jacks. Various devices are used for gripping or anchoring tendons, including swaged fittings on strands, bolts threaded on bars, wedges, and buttonheads on wires.

Prestress losses. One reason for loss of prestress is elastic shortening of concrete and steel due to compression of the concrete. Another reason is creep, or plastic flow, on concrete. It is an inelastic deformation dependent on time that occurs under constant stress.

Concrete shrinks when it dries and chemical changes take place. Shrinkage is dependent on time but not on stresses due to external loading.

Other possible losses in prestress that should be considered include those due to creep of the steel and to friction when the tendons rub against the concrete. *See* CREEP (MATERIALS).

Procedures for estimating the loss of prestress due to shrinkage, creep, friction, and other causes are given in design codes, such as the American Concrete Institute *Building Code Requirements for Reinforced Concrete*, ACI 318, and the American Association of State Highway and Transportation Officials *Standard Specifications for Highway Bridges*.

Pretensioning and posttensioning. Two methods are used in fabricating prestressed beams. In one method, the concrete is bonded to the stretched steel before the prestress is applied. This is called pretensioning. In the other method, posttensioning, the prestress is applied initially through end anchorages and the concrete may or may not be bonded later to the steel.

In pretensioning, the steel is laid through the beam forms and stretched between external abutments. Next, concrete is placed in the forms and allowed to set. When it has gained sufficient strength, the external pull on the tendons is relieved, transferring the prestress to the concrete through bond. The pretensioning method can be used on casting beds several hundred feet long to mass-produce many beams simultane- ously.

In posttensioning, the tendons are prevented initially from bonding to the concrete, usually by encasement in sheaths. The concrete is placed in the beam forms around the sheathed tendons and allowed to set. When it has gained sufficient strength, jacks are used to tension the tendons, and in so doing, the jacks react against the ends of the beam. The tendons then are anchored to the concrete to apply the prestress, and the jacks are released and removed.

Frequently, grout is forced into the sheaths to bond the tendons to the beam concrete. This gives the prestressed beam greater reserve strength and better crack control under overload. Posttensioning appears to be most advantageous for long-span beams and for assembling precast beam components in the field.

Circular prestress. Circular tanks, pipe, or the ring girder of domes may be prestressed, in contrast to the linear prestressing used for beams, by wrapping with steel bars or wires under high tension. Special machines have been developed for rapid circular prestressing with wire. *See* REINFORCED CONCRETE; STORAGE TANK.

Frederick S. Merritt

Bibliography. P. W. Abeles et al., *Prestressed Concrete Designers Handbook*, 3d ed., 1981; M. Fintel, *Handbook of Concrete Engineering*, 2d ed., 1985; M. K. Hurst, *Prestressed Concrete Design*, 2d ed., 1998; P. Jones, *Concrete and Masonry*, 1983; A. H. Nilson, *Design of Prestressed Concrete*, 2d ed., 1987.

Priapulida

One of the minor groups of wormlike marine animals. Once linked with the Echiurida and the Sipunculida under the term Gephyrea, it has since been recognized as a separate phylum of the animal kingdom with uncertain zoological affinities. *See* ECHIURA; SIPUNCULA.

Priapulida inhabit the colder waters of both hemispheres. They burrow in mud and sand of the sea floor, from the intertidal region to depths of 14,800 ft (4500 m).

The phylum is a small one with only two genera, *Priapulus* and *Halicryptus*. In the former there are seven known species: *P. caudatus* and *P. bicaudatus*, which occur in the northern temperate zone and Arctic waters; *P. australis*, found around the southern tip of South America; *P. tuberculatospinosus* from Antarctic seas; *P. horridus* from the coast of Uruguay; and *P. atlantisi* and *P. profundus* from bathyal and abyssal depths, respectively, off the northeastern coast of the United States. There is but one species of *Halicryptus, H. spinulosus*, found in northern waters.

Priapulids are small to medium-sized animals, the largest specimen attaining 6 in. (15 cm) in length. The body of *Priapulus* is made up of three distinct portions: proboscis, trunk, and caudal appendage (**illus.** *a*). Separated by a constriction from the trunk, the bulbous, introversible proboscis usually constitutes the anterior third of the body and is marked by 25 longitudinal ridges of papillae or spines. The

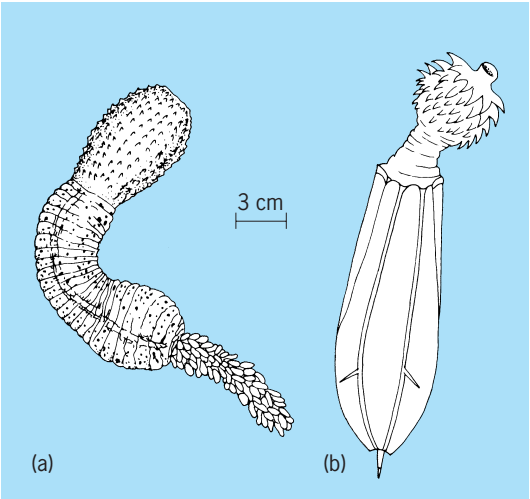

Priapulida. (*a*) *Priapulus* adult and (*b*) larva.

mouth is located at the anterior end of the proboscis and is surrounded by concentric rows of teeth; during feeding the mouth region is everted for the capture of prey. The cylindrical trunk is annulated, but not segmented, and is often covered with irregularly dispersed spines and tubercles. At the posterior end of the trunk there are three openings: the anus and two urogenital apertures. Continuous with the body cavity, the caudal appendages consist of hollow stalks with numerous lateral vesicles. Proposed functions of the caudal appendage are respiration and chemoreception. *Halicryptus* has no caudal appendage and the proboscis is only one-sixth of the body length.

The body wall of priapulids is composed of a cuticle (periodically molted), epidermis, circular and longitudinal muscle layers, and a thin, nucleated membrane which lines the entire body cavity. Originating on the body wall and attached to the mouth region and pharynx are retractor muscles which control the introversion of the mouth and proboscis.

The digestive tract consists of a muscular pharynx, a long, straight intestinal tube, and a short rectum which leads to the terminal anus. Priapulids are predacious, feeding on polychaetes, small crustaceans, and other priapulids which they swallow whole.

The simple nervous system includes a nerve ring encircling the pharynx and a midventral nerve cord situated within the epidermis.

A pair of elongate urogenital organs lies in the body cavity, one on either side of the intestine. Each organ is made up of a central duct into which the gonad opens on one side and the excretory cells (solenocytes) on the other side. The duct carries both excretory products and gametes to the exterior through a terminal urogenital pore.

Sexes are separate. Gametes are shed into the seawater where fertilization takes place. The larva, similar in morphology to the adult, possesses a retractable proboscis armed with spines (illus. *b*). Differing from the adult, the larval trunk is encased in a cuticularized lorica, composed of several plates, which is eventually shed. Mary E. Rice

Bibliography. L. H. Hyman, *The Invertebrates*: *Acanthocephala, Aschelminthes, and Entoprocta*, vol. 3, 1951; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; H. L. Sanders and R. R. Hessler, *Priapulus atlantisi* and *Priapulus profundus*: Two new species of priapulids from bathyal and abyssal depths of the North Atlantic, *Deep Res.*, 9:125–130, 1962.

Prilling

Solidification of droplets of molten material free-falling against an upward stream of air in a tower. It is a process used extensively in nitrogen fertilizer manufacturing. Melt is dispersed in the top of the tower at a temperature just above the solidification point of the material being processed. The dispersion arrangement, air flow, tower dimensions, and feed material are selected so the droplets approach a spherical shape and solidify before reaching the bottom of the tower.

Prilling has long been a major way of agglomerating ammonium nitrate and urea because of its relative simplicity and economy. It has been estimated that prilling urea is slightly less costly than pan granulation. Prills of ammonium nitrate and urea are smooth, spherical, dust-free, and moderately rugged, but usually are smaller and more fragile than granulated fertilizers.

Tower heights for prilling ammonium nitrate and urea usually range from 72 to 240 ft (22 to 72 m). A process for prilling ammonium nitrate has been developed in which the melt droplets fall through an upward-moving cloud of suitable airborne dust in a tower only 23 ft (7 m) tall, dust conditioning of the prills is accomplished simultaneously, and control of dust emissions can be accomplished economically.

Prills, mostly <0.1 in. (3 mm) in diameter, are too small for effective use in blends of granulated fertilizers, and their production in larger sizes increases cost and lowers product quality. Because of the large volume and high velocity of air through prilling towers, abatement of dust, and especially of ammonium nitrate fume, can present problems and increase production cost. Despite these disadvantages, prilling remains in wide use. There have been reports of economical production of N-P and N-P-K fertilizers in a prilling process said to be cleaner than granulation. However, in production of urea there is a trend toward preference for granulation in new installations because of ease of granule size control and of superior granule hardness. *See* DRYING; FERTILIZER.

E. O. Huffman

Bibliography. G. M. Blouin, *Urea Costs, Granulation vs Prilling*, TVA Bull. Y-92, April 1975; R. W. R. Carter and A. G. Roberts, *Proc. No. 110, The Fertiliser Society*, October 23, 1969; G. Drake, *Proc. No. 136, The Fertiliser Society*, October 25, 1973; F. E. Steenwinkel and J. W. Hoogendonk, *Proc. No. 109, The Fertiliser Society*, April 24, 1969.

Primary vascular system (plant)

The arrangement of conducting elements which serves for two-way transportation of substances between different parts of a plant. The conducting elements are of two principal kinds: xylem, which is mainly responsible for the conduction of water together with dissolved inorganic substances upward from the root to other plant organs; and phloem, which is mainly responsible for the conduction of food materials (assimilates), a flow which may take place in either direction. In the shoot region of the plant, xylem and phloem are usually associated into vascular bundles (**Fig. 1**). In the root, however, they usually alternate with one another on different radii. *See* PHLOEM; XYLEM.

This system of vascular tissue functions while the young shoots and roots are growing in length and

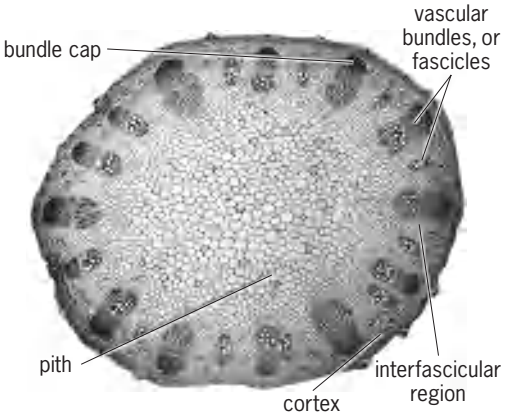

Fig. 1. Cross section of sunflower (*Helianthus*) stem.

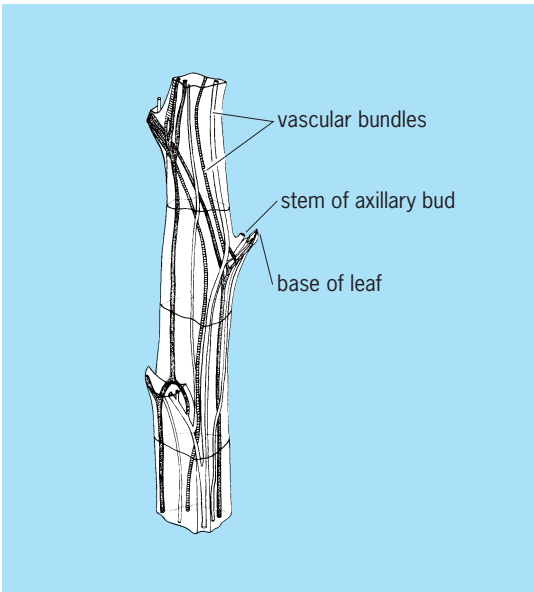

Fig. 2. Primary vascular system of the stem of the potato (*Solanum tuberosum*). (*After A. J. Eames and L. H. MacDaniels, An Introduction to Plant Anatomy, 2d ed., McGraw-Hill, 1947*)

while their tissues are undergoing differentiation and maturation. For this reason it is known as the primary vasular system, and is the only vascular tissue found in many plants. However, in most plants that increase in girth after elongation has ceased, much of the additional tissue concerned in this secondary growth is vascular. This secondary vascular tissue is formed by the activity of a cambium, which in nearly all cases develops within the primary vascular bundles.

Course of bundles in the stem. The vascular bundles in the shoot of most higher plants are interconnected to form a complex system (**Fig. 2**). The arrangement of the bundles in this system is subject to much variation in different plants, but in most dicotyledons it consists, as seen in cross section, of a ring of bundles toward the periphery of the stem. The ground tissue of the stem outside this ring is the cortex, and that to the inside of it is the pith. Cortex and pith are connected between the bundles by ground tissue known as medullary rays.

The bundles of the ring usually run up the stem more or less vertically without branching except at the positions where leaves are attached (nodes). Here the bundles frequently branch and become interconnected. Each leaf receives a supply of one or more vascular bundles which are continuous with bundles of the stem. In most monocotyledons the bundles are not arranged in a single ring, but are scattered through much of the ground tissue. This appearance is due to some of the individual bundles not running vertically in the stem, but inclining toward the center of the stem and then at a higher level returning to the stem periphery.

Concept of the stele. The term stele has been used for all tissues inside the cortex, that is, for the primary vascular system together with the pith, if any. Three major types of stele are commonly recognized. (1) The protostele consists of a solid core of xylem surrounded by phloem, but with no pith. In general, protosteles are found in the stems of most primitive vascular plants (lycopods, psilophytes, and some ferns) and in the roots of most vascular plants. (2) In the siphonostele, a cylinder of vascular tissue encloses a pith. The cylinder may be pierced by gaps at the insertions of the leaves. This stele is characteristic of many ferns. (3) In the dictyostele, the vascular cylinder is composed of bundles separated by medullary rays as well as by leaf gaps. This is the type of stele characteristic of seed plants.

Types of bundle. Vascular bundles vary in structure chiefly with reference to the arrangement of xylem and phloem. The most common type of bundle in seed plants is the collateral, in which the phloem appears on one side of the xylem (**Fig. 3**). The phloem may be rather deeply embedded in the xylem in a collateral bundle; then the xylem assumes the shape of a V in cross section (**Fig. 4**). In a bicollateral vascular bundle the phloem occurs on both sides of the xylem (**Fig. 5**). The third main type of vascular bundle is the concentric, which occurs in two forms. In the amphivasal bundle (**Fig. 6**), the xylem surrounds

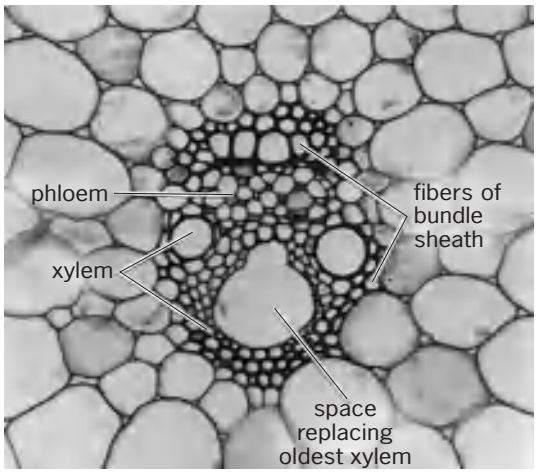

Fig. 3. Cross section of a closed collateral vascular bundle from stem of maize (*Zea mays*). (*After K. Esau, Plant Anatomy, 2d ed., John Wiley and Sons, 1967*)

the phloem, whereas in the amphicribral (**Fig. 7**) the phloem surrounds the xylem.

In addition to the conducting elements, vascular bundles contain various amounts of supporting tissue, usually fibers. These may form a complete sheath enclosing the vascular tissues, or they may appear as strands outside the phloem or the xylem, or on both sides. As seen in cross section, such strands are often called bundle caps.

Origin and development. Shoots and roots grow in length by additional tissue being formed at their extremities by growing points, or apical meristems. At first the tissue consists of more or less uniform undifferentiated cells, but gradually the arrangement of tissues found in the mature stem or root is acquired. The future vascular system can first be distinguished as strands of cells, narrower in cross section, known as procambial strands, that run through the general ground tissue. In the shoot, these arise in relation to the primordia of the leaves and develop into the pattern of vascular bundles characteristic of the plant. In the root also, procambial strands arise in positions corresponding to the pattern of xylem and phloem strands found in the mature root.

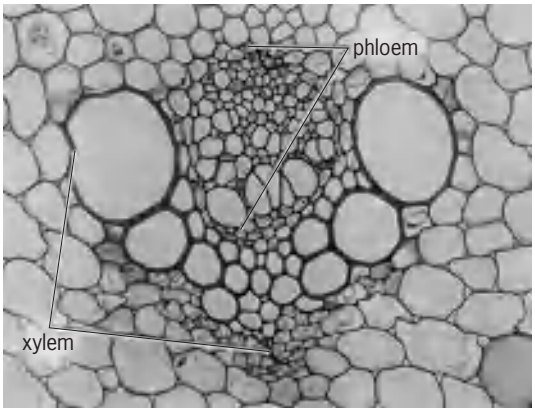

Fig. 4. Cross section of a closed collateral vascular bundle from stem of *Asparagus*.

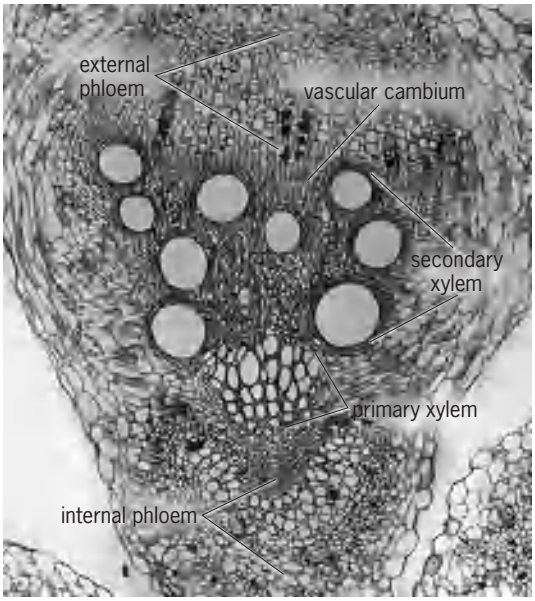

Fig. 5. Cross section of an open bicollateral vascular bundle from stem of squash (*Cucurbita*). (*After K. Esau, Plant Anatomy, 2d ed., John Wiley and Sons, 1967*)

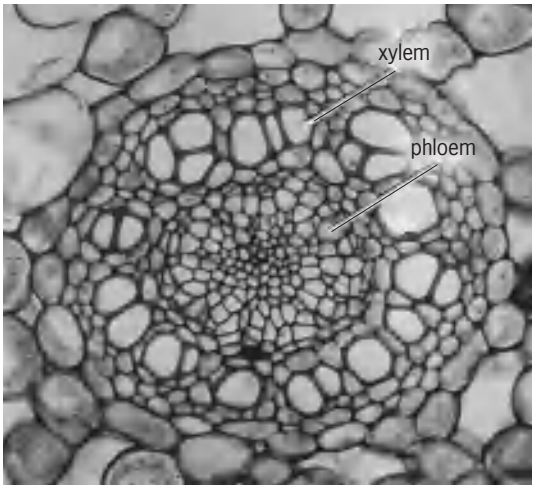

Fig. 6. Cross section of a concentric amphivasal vascular bundle from rhizome of sweet flag (*Acorus calamus*). Here xylem surrounds phloem.

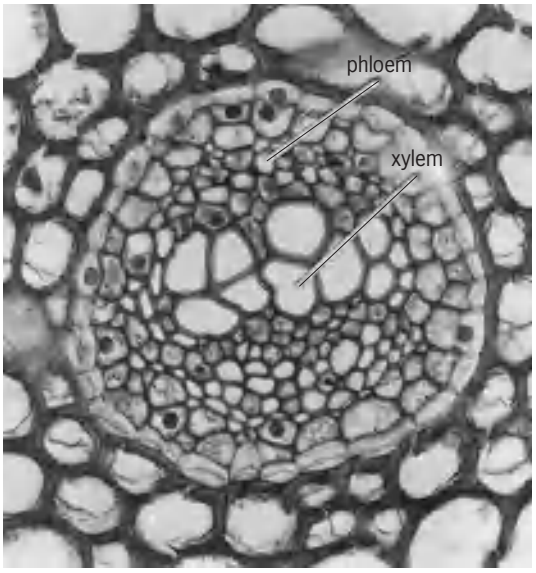

Fig. 7. Cross section of a concentric amphicribal vascular bundle from rhizome of *Polypodium*, a fern. (*After K. Esau, Plant Anatomy, 2d ed., John Wiley and Sons, 1967*)

During differentiation of a procambial strand into a vascular bundle in the shoot of a dicotyledonous plant, the cells lying next to the pith usually mature as xylem elements, and those toward the cortex as phloem. Differentiation then proceeds toward the center of the strand, and in this way a typical collateral bundle is formed. If all the cambial cells develop into xylem or phloem, the resulting bundle is incapable of further growth and is commonly called a closed vascular bundle. Such bundles are characteristic of lower vascular plants, monocotyledons, and those dicotyledons that have no secondary growth. If secondary growth occurs, the last procambial cells become cells of the vascular cambium. These cambial cells are capable of continued

growth and division, their products differentiating as further xylem and phloem elements. Such enlargement of the vascular bundle by a cambium is referred to as secondary growth. The sectors of cambium within each vascular bundle usually become linked into a complete cambial cylinder by the development of similar cambial growth in the medullary rays. These linking sectors are known as interfascicular cambium and also produce secondary xylem and phloem. As a result, the secondary vascular tissues form a continuous cylinder and not a system of distinct strands. *See* PLANT GROWTH; PLANT TISSUE SYSTEMS.

William B. Philipson

Primates

The mammalian order to which humans belong. Primates are generally arboreal mammals (living in trees) with a geographic distribution largely restricted to the Tropics. Unlike most other mammalian orders, the primates cannot be defined by a diagnostic suite of specializations, but are characterized by a combination of primitive features and progressive trends. These include.

1. Increased dominance of vision over olfaction, with eyes more frontally directed, development of stereoscopic vision, and reduction in the length of the snout.
2. Eye sockets of the skull completely encircled by bone.
3. Loss of an incisor and premolar from each half of the upper and lower jaws with respect to primitive placental mammals.
4. Increased size and complexity of the brain, especially those centers involving vision, memory, and learning.
5. Development of grasping hands and feet, with a tendency to use the hands rather than the snout as the primary exploratory and manipulative organ.
6. Progressive elaboration of the placenta in conjunction with longer gestation period, small litter size (only one or two infants), and precocial young.
7. Increased period of infant dependency and more intensive parenting.
8. A tendency to live in complex, long-lasting social groups.

It has been recognized for a long time that many of these features are adaptations for living in trees. However, it has been proposed more recently that primates may have developed their specializations as a consequence of being visually directed predators, living among the smaller branches of the forest canopy or undergrowth, that captured insects with their hands.

Classification of the primates is as follows:

Order Primates
- Suborder Strepsirhini
  - Infraorder Lorisiformes
    - Superfamily Lorisoidea
      - Family: Lorisidae (lorises)
        - Galagidae (bushbabies)
  - Infraorder Lemuriformes
    - Superfamily Lemuroidea
      - Family: Cheirogaleidae (dwarf lemurs)
        - Lepilemuridae (sportive lemur)
        - Lemuridae (true lemurs)
        - Indriidae (sifakas, indri, woolly lemur)
        - Daubentoniidae (aye-aye)
- Suborder Haplorhini
  - Hyporder Tarsiiformes
    - Superfamily Tarsioidea
      - Family Tarsiidae (tarsiers)
  - Hyporder Anthropoidea
    - Infraorder Platyrrhini
      - Superfamily Ceboidea
        - Family: Callitrichidae (marmosets, tamarins)
        - Cebidae (capuchins, squirrel monkeys, douroucoulis, titis)
        - Atelidae (sakis, uakaris, howler monkeys, spider monkeys, woolly monkeys)
    - Infraorder Catarrhini
      - Superfamily Cercopithecoidea
        - Family Cercopithecidae (Old World monkeys)
      - Superfamily Hominoidea
        - Family: Hylobatidae (gibbons, siamang)
        - Hominidae (orangutan, gorilla, chimpanzees, humans)

Strepsirhines. There are two major groups of primates: the strepsirhines or "lower" primates, and the haplorhines or "higher" primates. Strepsirhines have elongated and forwardly projecting lower front teeth that form a toothcomb, used for grooming the fur and for obtaining resins and gums from trees as a source of food. The digits of the hands and feet bear flattened nails, rather than claws, except for the second toe, which retains a sharp toilet claw for grooming. They also have a moist, naked rhinarium and cleft upper lip (similar to the wet noses of dogs). Most strepsirhines are nocturnal, with large eyes and a special reflective layer (the tapetum lucidum) behind the retina that intensifies images in low light. Compared with haplorhines, the brain size is relatively small and the snout tends to be longer.

The strepsirhines are subdivided into two major groups: the lorisoids, which are found throughout tropical Africa and Asia, and the lemuroids, which are restricted to Madagascar.

Lorisoids. These include the galagids or bushbabies (*Galago*, *Otolemur*, *Euoticus*, and *Galagoides*) and the lorisids or lorises (*Loris*, *Nycticebus*, *Perodicticus*, *Pseudopotto*, and *Arctocebus*). They are small nocturnal primates, in which the largest species, the greater bushbaby, weighs only about 1 kg (2 lb). Their diet consists mainly of a combination of insects, fruits, and gums. Lorisoids are semisolitary, living in small, dispersed social groups. Each male has a territory that overlaps with that of several females, but individuals usually associate only during courtship and mating. At other times,

Fig. 1. Mouse lemur (*Microcebus murinus*). (*Photo by David Haring/Duke Lemur Center*)

communication is maintained by scent marking or vocalizations. Bushbabies are widespread in sub-Saharan Africa, ranging from tropical rainforests to woodland savanna. They are specialized vertical clingers and leapers, with elongated lower limbs. Lorises progress much more cautiously, employing a slow deliberate mode of quadrupedal locomotion. They are specialized for this type of movement by having powerful grasping hands with a large thumb and highly reduced second finger, a long flexible body, and a tail that is very short or absent.

Lemuroids. The greatest diversity of strepsirhines is found on Madagascar, where more than 30 species are represented, belonging to five different families. The cheirogaleids or dwarf lemurs include eight species of small lemuroids belonging to *Microcebus*, *Mirza*, *Allocebus*, *Cheirogaleus*, and *Phaner* (**Fig. 1**). This group includes the smallest known primate, the pygmy mouse lemur (*Microcebus myoxinus*), which weighs as little as 25 g. All cheirogaleids move quadrupedally. They are nocturnal, sleeping during the day in nests or holes in trees, often in small groups. *Cheirogaleus* is unusual in that it hibernates during the dry season, utilizing reserves of fat stored in its tail. Cheirogaleids are omnivorous, feeding mainly on fruits, flowers, gums, and insects. Like lorisoids, semisolitary males have ranges that overlap that of several females. Litters can be as large as four, but one or two infants are most common.

Lepilemuridae contains a single living genus, *Lepilemur* (the sportive or weasel lemur). They are medium-sized nocturnal lemurs, weighing less than 1 kg (2 lb), and are specialized for vertical clinging and leaping. Their diet consists almost exclusively of leaves and flowers. They are unique among primates in reingesting their feces in order to extract adequate nourishment from their poor-quality diet. True lemurs (*Lemur*, *Eulemur*, and *Varecia*) and bamboo lemurs (*Hapalemur*) are included in the family Lemuridae. The best-known member of this group is the ringtailed lemur, which is common in captivity. True lemurs are cat-sized primates (about 2–4 kg or 4–9 lb) that move quadrupedally in the trees and on the ground. They are mostly active during the day, and eat a variety of leaves, fruits, and flowers. Lemurs are gregarious and live in large social groups. Bamboo lemurs specialize in eating bamboo shoots and pith. They are generally smaller than true lemurs, and live in one-male groups.

The indriids include the indri (*Indri*), the sifaka (*Propithecus*), and the woolly lemur (*Avahi*). All are specialized for vertical clinging and leaping. The indri and sifaka are diurnal, and among the largest strepsirhines (3–7 kg or 7–15 lb), while the woolly lemur is nocturnal and somewhat smaller (about 1 kg or 2 lb). The indri is unusual among lemuroids in having a vestigial tail. All indriids subsist primarily on a diet of leaves. The indri and woolly lemur live in groups consisting of an adult pair and their offspring, while sifakas live in small, mixed groups. The aye-aye (*Daubentonia madagascariensis*) is a rare and specialized lemuroid belonging to its own family, the Daubentoniidae. It has a suite of unique specializations for obtaining insect larvae from dead wood. These include large ears for locating the larvae; continuously growing incisors, like a rodent, for gnawing at the bark of trees; and a long slender wire-like middle finger to extract the larvae from holes. Aye-ayes are the largest of the nocturnal lemuroids (2–3 kg or 4–7 lb).

Tarsiers. These tiny primates (weighing only about 120 g) from the islands of Southeast Asia all belong to a single genus, *Tarsius* (**Fig. 2**). They are nocturnal with the largest eyes of any primate, and unlike strepsirhines they lack the tapetum lucidum. They have legs that are much longer than the forelimbs, a long almost hairless tail with a tuft at the tip, and fleshy pads on the ends of their fingers—all adaptations for a specialized lifestyle as vertical clingers and leapers. Unlike strepsirhines, which only have a toilet claw on their second toe, tarsiers are unique among primates in having toilet claws on both their second and third toes. Their diet consists primarily

Fig. 2. Philippine tarsier (*Tarsius syrichta*). (*Photo © 2006 William Flaxington*)

of insects. In the past, tarsiers have been grouped together with the strepsirhines as prosimians, because they retain many primitive features lost in higher primates. However, tarsiers share a number of distinctive specializations with anthropoids (for example, the development of a bony partition that partially separates the eyeball from the chewing muscles behind, lack of a toothcomb, a dry hair-covered nose and undivided upper lip, a short snout with reduced facial whiskers, and the structure of the placenta) that suggest that they are more closely related to each other than either is to the strepsirhines. For this reason, tarsiers and anthropoids are classified together as haplorhines.

Anthropoids. This group contains the platyrrhines or New World monkeys and the catarrhines or Old World monkeys, apes, and humans. Anthropoids are distinguished from strepsirhines and tarsiers in having a larger brain, relatively small eyes (all anthropoids are diurnal, active by day, except for the nocturnal douroucouli from South America), eye sockets almost completely enclosed by a bony septum, the two halves of the lower jaw fused in the midline rather than separated by a cartilage, small and immobile ears, the hands and feet bearing nails with no toilet claws (except for the callitrichids that have secondarily evolved claws on all fingers and toes), a single-chambered uterus rather than two-horned, and a more advanced placenta.

New World monkeys. The platyrrhines from South and Central America are a diverse group of primates comprising more than 50 species and 16 genera. Primatologists have had a difficult time establishing a classification of platyrrhines that reflects their evolutionary interrelationships, and no consensus has been reached. There is agreement, however, that three distinct clusters can be defined: the callitrichids, the pitheciines, and the atelines. The last two groups appear to be closely related and are commonly included together in the family Atelidae. The relationships of the remaining platyrrhines are uncertain, and they are often placed together for convenience in the Cebidae.

All platyrrhines are arboreal, and they are widely distributed throughout tropical forests extending from Mexico to northern Argentina. They are small to medium-sized primates ranging from 100 g to 15 kg (0.2 to 33 lb). Platyrrhines exhibit a variety of quadrupedal locomotor types ranging from squirrel-like scrambling to leaping and forelimb suspension. Atelines and capuchin monkeys are unique among primates in having a specialized prehensile tail that can grasp around branches for extra support.

The smallest New World monkeys are the callitrichids (100–700 g). This group includes the marmosets (*Cebuella* and *Callithrix*), tamarins (*Saguinus* and *Leontopithecus*), and Goeldi's monkey (*Callimico*). The callitrichids are distinguished from other platyrrhines in having claws on all of their digits, associated with their squirrellike ability to cling to and scamper up large tree trunks, and in the absence (except for *Callimico*) of the third molar. Their diet consists mainly of gums, fruits, and insects. Marmosets have large and specialized incisors for gouging trees to obtain exudates. Callitrichids generally live in small, extended family groups containing a single adult female and one or more adult males. They typically produce twins (except for *Callimico*) that are cared for primarily by the males.

Cebids include the capuchin monkey (*Cebus*), which is a medium-sized platyrrhine, weighing about 3 kg (7 lb). Their teeth and dextrous hands are specialized for opening and processing hard seeds and fruits, which form an important part of their diet. Capuchins are highly intelligent primates, and tool-using behaviors have been reported. They have a semiprehensile tail which, unlike the larger atelines, is fully hair-covered. Capuchins live in small groups with several adult males and females. The squirrel monkey (*Saimiri*) is smaller than the capuchin and more lightly built. Squirrel monkeys travel quadrupedally through the forest, often leaping between branches and trees. Their diet consists mainly of fruits, supplemented by insects. They live in large social groups, sometimes of up to 200 individuals, dominated by a hierarchy of adult females. The cebids also include the douroucoulis (*Aotus*) and titis (*Callicebus*), which are small (about 1–2 kg or 2–4 lb), short-faced, long-tailed monkeys that live in monogamous family groups. They are agile quadrupeds, adept at leaping. Their diet consists of fruits, leaves, and insects. The douroucouli is exceptional in being the only nocturnal anthropoid. Like many strepsirhines, douroucoulis have relatively large eyes, but lack the tapetum lucidum and retain cones in the retina for color vision (typical of anthropoids).

The pitheciines are medium-sized platyrrhines (2–3 kg or 4–7 lb) that include the sakis (*Pithecia* and *Chiropotes*) and uakaris (*Cacajao*). They are striking monkeys with highly distinctive coats and facial hair patterns. Pitheciines have a specialized dentition associated with predation on hard seeds and nuts. They often live in large social groups, although *Pithecia* generally travels in family units. The atelines are the largest platyrrhines (5–15 kg or 11–33 lb), and are characterized as a group by having a long prehensile tail that bears a naked grasping pad on its underside. Howler monkeys (*Alouatta*) are heavy-set, slow-climbing arboreal quadrupeds that subsist mainly on fruits and leaves. Most groups contain one adult male and several females with their offspring. Howlers advertise their presence to neighboring groups by emitting loud roars. They have an enormously inflated hyoid bone below the chin that acts as a resonating chamber to enhance these vocalizations. The spider monkeys and woolly spider monkeys (*Ateles* and *Brachyteles*) have long and slender limbs (**Fig. 3**). The thumb is reduced to a small tubercle. They are much more agile and acrobatic than howler monkeys. During feeding on ripe fruits and young leaves, they hang below branches using different combinations of their limbs and prehensile tail.

Catarrhines. The catarrhines include all anthropoid primates from Africa, Asia, and Europe. There are two main groups: the cercopithecids or Old World monkeys, and the hominoids or apes

Fig. 3. Spider monkey (*Ateles geoffroyi*) from Panama. (*Photo by Roy Fontaine/WPRC AV Archives*)

and humans. Catarrhines are distinguished from platyrrhines by a reduction in the number of premolars from three to two in each half of the upper and lower jaw, and the development of a tubelike (rather than ringlike) tympanic bone that supports the eardrum.

Cercopithecoidea. Old World monkeys are widely distributed throughout sub-Saharan Africa and tropical Asia, and also occur in the extreme southwestern tip of the Arabian Peninsula, northwest Africa, Gibraltar (their only European record), and East Asia. They are a highly successful group comprising more than 80 species. They are distinguished from other anthropoids in having bilophodont molar teeth that bear a pair of transverse crests. They also have naked, roughened sitting pads on their rumps, called ischial callosities—a feature that they share with hylobatids. In addition, most Old World monkeys are highly sexually dimorphic, with males considerably larger than females.

Cercopithecids are subdivided into two subfamilies: the Cercopithecinae and Colobinae. Both groups have developed distinctive dietary specializations. The cercopithecines are predominantly fruit eaters, and have cheek pouches for storing food, broad incisors, and low cusps on their molars. In contrast, colobines have complex sacculated stomachs for the digestion of leaves, narrow incisors, and high molar cusps. Colobines are also distinguished from cercopithecines in having a shorter face, a broader distance between the eye sockets, longer hindlimbs for leaping, and thumbs that are highly reduced.

Cercopithecines are exclusively African, except for the highly diverse and successful macaques (*Macaca*), which are widely distributed from Morocco and Gibraltar to Southeast Asia. The 19 species of macaques (**Fig. 4**) occupy a wide range of habitats from temperate woodland to tropical forest, and they exhibit different degrees of preference for traveling on arboreal or terrestrial substrates. They are medium-sized (3–10 kg or 7–22 lb) quadrupedal monkeys with omnivorous diets. Macaques are closely related to the mangabeys (*Cercocebus* and *Lophocebus*) and baboons (*Papio*, *Mandrillus*, and *Theropithecus*) of Africa, which are grouped together in the tribe Papionini. The mangabeys are large, forest-dwelling monkeys with slender limbs and long tails. Baboons are the largest Old World monkeys (with males exceeding 30 kg or 66 lb). The savanna baboons (*Papio*) are widespread throughout sub-Saharan Africa, and they even occur in the Arabian Peninsula. They prefer woodlands and savannas, where they travel mainly on the ground, although they do require trees or rocky cliffs for sleeping in at night. Baboons, like most papionins, live in large, complex social groups with up to 80 individuals, dominated by a hierarchy of adult males. The hamadryas baboon from the dry savannas of Ethiopia and Arabia is unusual in living in small one-male groups. Mandrills and drills (*Mandrillus*) live in the dense tropical rainforests of western Central Africa. Adult males are distinctive in having long muzzles with brightly colored faces and rumps. The gelada

Fig. 4. Long-tailed macaque (*Macaca fascicularis*). (*Photo by Roy Fontaine/Monkey Jungle*)

Fig. 5. Guenon species, De Brazza's monkey (*Cercopithecus neglectus*). (*Photo by Aaron Logan/LIGHTmatter Photography*)

baboon (*Theropithecus*) is restricted to the treeless highlands of Ethiopia. It is a ground-living quadruped that specializes in feeding on grasses. The hands and feet have short digits as a specialization for terrestrial locomotion, while the thumbs are long and dextrous to facilitate the manipulation and harvesting of small food objects. Geladas, like hamadryas baboons, live in one-male groups with up to 20 individuals.

The guenons (*Cercopithecus*) and their allies (*Miopithecus*, *Allenopithecus*, and *Erythrocebus*) are mostly forest monkeys. They are a diverse and widespread group from sub-Saharan Africa represented by more than 20 species. Guenons (**Fig. 5**) are medium-sized (3–7 kg or 7–15 lb), colorful, arboreal monkeys that live in large gregarious social groups. The vervet (*Cercopithecus aethiops*) is unusual in spending a good deal of time on the ground (**Fig. 6**). The diet of guenons consists of a combination of fruits, insects, and young leaves. The closely related talapoin (*Miopithecus*), from western Central Africa, is the smallest Old World monkey (weighing in at just over 1 kg or 2 lb). The patas (*Erythrocebus*) is a graceful, slender-limbed, fast-running monkey that lives on the grassy plains, savannas, and dry woodlands of equatorial Africa.

There are two major groups of colobines or leaf monkeys: the African colobus monkeys and the Asian langurs. Colobus monkeys are subdivided into three genera, mainly based on their coat color: the black-and-white colobus (*Colobus*), the red colobus (*Piliocolobus*), and the olive colobus (*Procolobus*). The black-and-white colobus and red colobus are large monkeys (8–10 kg or 18–22 lb), while the olive colobus is smaller (3–4 kg or 7–9 lb). Their diet consists of leaves, with some fruits, although the black-and-white colobus specializes in eating hard seeds. Most African colobines live in large multimale groups, but black-and-white colobus monkeys live in small groups with a single adult male.

Fig. 6. Vervet monkey, or green monkey (*Cercopithecus aethiops*). (*Photo by Gary M. Stolz/U.S. Fish and Wildlife Service*)

The colobines from Asia are more diverse than those from Africa, with over 20 species belonging to six genera. *Presbytis* and *Trachypithecus* are found throughout the forests of South and Southeast Asia. The most familiar and most adaptable representative is the large, semiterrestrial Hanuman langur from India and Sri Lanka, which lives in a variety of habitats ranging from forest to semidesert, and even occurs in urban centers. The other members of this group are arboreal, forest-dwelling monkeys ranging in size from 5 to 12 kg (11 to 26 lb). They mainly eat leaves, but fruits and seeds form an important part of their diets.

The other Asian colobines are commonly referred to as odd-nosed monkeys because of the peculiar shape of the nose. The best-known species is the proboscis monkey (*Nasalis*) from the mangrove swamps of Borneo. These large monkeys are remarkable leapers and excellent swimmers. Adult males, which are more than twice as large as females, have an enormous pendulous nose, while females have a petite upturned nose. The related simakobu (*Simias*) is found only on the Mentawai Islands, just off the coast of Sumatra. It is a distinctive colobine with a very short tail and relatively long forelimbs. The douc langur (*Pygathrix*) and snub-nosed monkeys (*Rhinopithecus*) from China and Indochina are large, handsome monkeys with flattened or upturned noses. The Golden snub-nosed monkey (*Rhinopithecus roxellana*) from the mountain forests of southern China is the largest living colobine, with males weighing almost 40 kg (88 lb).

Hominoidea. This is the superfamily to which apes and humans belong. Hominoids are distinguished from cercopithecoids in having primitive non-bilophodont molars, larger brains, longer arms than legs (except in humans), a broader chest, a shorter and less flexible lower back, and no tail. Many of these specializations relate to a more upright posture in apes, associated with a greater emphasis on vertical climbing and forelimb suspension.

Hominoids can be classified into two families: the Hylobatidae, which includes the gibbons and siamang, and the Hominidae, which includes the great apes and humans. The gibbons and siamang (*Hylobates*) are the smallest of the hominoids (4–11 kg or 9–24 lb), and for this reason they are sometimes referred to as the lesser apes. The nine or so species are common throughout the tropical forests of Asia. They are remarkable in having the longest arms of any primates, which are 30–50% longer than their

Fig. 7. Orangutan (*Pongo pygmaeus pygmaeus*) from Borneo.

legs. This is related to their highly specialized mode of locomotion, called brachiation, in which they swing below branches using only their forelimbs. Gibbons are fruit eaters, while the larger siamang incorporates a higher proportion of leaves in its diet. Hylobatids live in monogamous family groups in which males and females are similar in size (unlike most catarrhines). Both sexes defend a territory, mainly through complex vocalizations or songs, which serve to advertise their presence to neighboring pairs. As in Old World monkeys, hylobatids sit upright on branches to sleep, resting on their ischial callosities, rather than sleeping in nests like great apes.

The great apes include the orangutan (*Pongo*) from Asia and the gorilla (*Gorilla*) and chimpanzees (*Pan*) from Africa. These were formerly included together in their own family, the Pongidae, to distinguish them from humans, who were placed in the Hominidae. However, recent anatomical, molecular, and behavioral evidence has confirmed that humans are closely related to the great apes, especially to the African apes, and for this reason most scientists now classify them together in a single family, the Hominidae. The orangutan is restricted to the tropical rainforests of Borneo and northern Sumatra (**Fig. 7**). They are large, arboreal primates that climb cautiously through the trees using all four limbs for support. Orangutans subsist mainly on fruits. As is typical of other great apes, they lack ischial callosities and build nests to sleep in at night. Adult male orangutans are about twice as heavy as females, and the sides of their faces are adorned with huge fatty flaps. Orangutans are the least sociable of the hominoids and tend to lead solitary lives. Males aggressively defend their territory from other males, which encompasses the ranges of several female consorts.

The gorilla, the largest of the hominoids, has a disjunct distribution in tropical Africa. Because of their great size, gorillas are almost entirely terrestrial, although females and young individuals frequently climb trees. Nests are often built on the ground. Gorillas move quadrupedally, and like chimpanzees, the hands are specialized for knuckle walking in which the weight of the animal is borne on the upper surface of the middle joints of the fingers. Mountain gorillas eat a variety of leaves, stems, and roots, while lowland gorillas eat a greater proportion of fruits. Groups consist of a dominant male, called a silverback, as well as several adult females, subadults, and infants.

There are two species of chimpanzees, the common chimpanzee (*Pan troglodytes*) and the bonobo or pygmy chimpanzee (*Pan panicus*). The common chimpanzee is widely distributed in the forests and woodlands stretching across equatorial Africa, while the pygmy chimpanzee is restricted to the tropical rainforests of the Congo. Both species nest and feed in trees, but they mostly travel on the ground. Common chimpanzees have eclectic diets, including meat, which they obtain by hunting small to medium-sized mammals. Tool-using behaviors are common, and more than a dozen simple tool types have been identified. Chimpanzees are gregarious and sociable, and they live in large multimale communities that divide into smaller subgroups for foraging. *See* APES; FOSSIL APES; FOSSIL HUMANS; FOSSIL PRIMATES; MAMMALIA; MONKEY.

Terry Harrison

Bibliography. J. G. Fleagle, *Primate Adaptation and Evolution*, 2d ed., 1998; R. A. Mittermeier et al., *Lemurs of Madagascar*, 1994; J. R. Napier and P. H. Napier, *Natural History of the Primates*, 1985; R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999; A. Richard, *Primates in Nature*, 1985; B. B. Smuts et al. (eds.), *Primate Societies*, 1987; R. Tuttle, *Apes of the World*, 1987.

Prime mover

The component of a power plant that transforms energy from the thermal or the pressure form to the mechanical form. Mechanical energy may be in the form of a rotating or a reciprocating shaft, or a jet for thrust or propulsion. The prime mover is frequently called an engine or turbine and is represented by such machines as waterwheels, hydraulic turbines, steam engines, steam turbines, windmills, gas turbines, internal combustion engines, and jet engines. These prime movers operate by either of two principles (**Fig. 1**): (1) balanced expansion, positive displacement, intermittent flow of a working fluid into and out of a piston and cylinder mechanism so that by pressure difference on the opposite sides of the piston, or its equivalent, there is relative motion

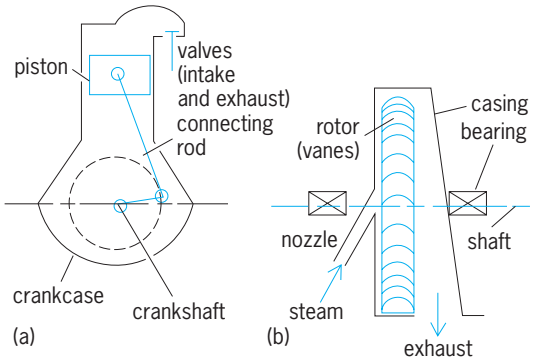

Fig. 1. Representative prime movers. (*a*) Single-acting four-cycle, automotive-type internal combustion engine. (*b*) Single-stage, impulse-type steam turbine.

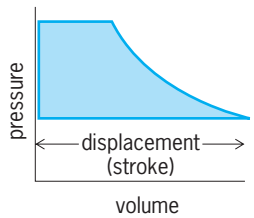

Fig. 2. Pressure-volume diagram (indicator card) for ideal, no-clearance, fluid-displacement types of prime mover.

of the machine parts; or (2) free continuous flow through a nozzle where fluid acceleration in a jet (and vane) mechanism gives relative motion to the machine parts by impulse, reaction, or both. *See* IMPULSE TURBINE; INTERNAL COMBUSTION ENGINE; REACTION TURBINE; STEAM ENGINE.

Displacement prime mover. Power output of a fluid-displacement prime mover is conveniently determined by pressure-volume measurement recorded on an indicator card (**Fig. 2**). The area of the indicator card divided by its length is the mean effective pressure (mep) in pounds per square inch, and horsepower of the prime mover is given by Eq. (1), where

$$\text{Horsepower} = \frac{\text{mep} \times Lan}{33,000} \qquad (1)$$

L is stroke in feet, a is piston area in square inches, and n is number of cycles completed per minute. Actual mep is smaller than the theoretical mep and may be related to the theoretical value by diagram factor or engine efficiency (**Table 1**). *See* MEAN EFFECTIVE PRESSURE.

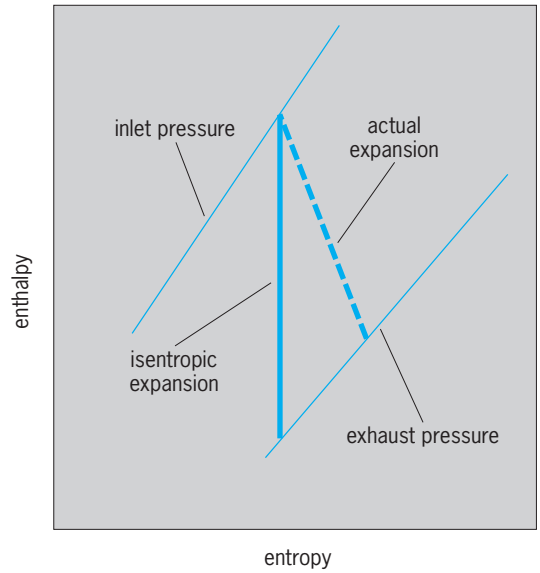

Fig. 3. Enthalpy-entropy (Mollier) chart of performance of steam- or gas-turbine type of prime mover.

Acceleration prime mover. Performance of fluid acceleration (hydraulic) prime movers is given by Eq. (2), where Q is water flow rate in cubic feet per

$$\text{Horsepower} = \frac{QH}{8.8} \times \text{efficiency} \qquad (2)$$

second, and H is head in feet. For heat-power prime movers of the fluid acceleration type, actual properties of the thermodynamic fluid, as given in tables and graphs, especially the Mollier chart, permit the rapid evaluation of the work or power output from the general energy equation which resolves to the form of Eq. (3), where h is the enthalpy in Btu/lb, and the

$$\Delta W, \ \text{Btu/lb of fluid} = h_{\text{inlet}} - h_{\text{exhaust}} \qquad (3)$$

inlet and exhaust conditions can be connected by an isentropic expansion for ideal conditions, or modified for irreversibility to a lesser difference by engine efficiency (**Fig. 3** and **Table 2**). Fluid consumption follows from Eq. (4) or Eq. (5).

$$\text{Fluid consumption, lb/hph} = 2545/\Delta W \qquad (4)$$

$$\text{lb/kWh} = 3413/\Delta W \qquad (5)$$

In the fluid acceleration type of prime mover, jet velocities experienced in the nozzles can be found in feet per second, for nonexpansive fluids,

TABLE 1. Dimensional and performance criteria of some selected fluid displacement-type prime movers*

Type	Size, hp	rpm	Stroke, in.	Bore: stroke ratio	Piston speed, ft/min	Brake mep, lb/in.^{2}	Diagram factor, or engine efficiency
Steam engine	25–500	100–300	6–24	0.8–1.2	400–600	50–100	0.6–0.8
Automobile engine	10–300	2000–4000	3–5	0.9–1.1	1000–2000	50–100	0.4–0.6
Aircraft engine	100–3000	2500–3500	4–7	0.8–1.1	1500–3000	100–300	0.4–0.6
Diesel, low-speed	100–5000	100–300	10–24	0.8–1.0	500–1000	40–80	0.4–0.7
Diesel, high-speed	25–1000	1500–2000	3–6	0.8–1.0	800–1500	50–100	0.4–0.6

*1 hp = 0.75 kW. 1 in. = 25 mm. 1 ft/min = 5×10^{-3} m/s. 1 lb/in.^{2} = 6.9 kPa.

TABLE 2. Dimensional and performance criteria of some selected fluid acceleration-type prime movers*

Type	Rating, kW	Number of stages	Head, ft; or pressure, lb/in.^{2}	Temperature, °F (°C)	Exhaust pressure, in. Hg abs	rpm	Tip speed, ft/s	Efficiency
Pelton waterwheel	1000–200,000	1	500–5000 ft	Ambient	atm	100–1200	100–250	0.75–0.85
Francis hydraulic turbine	1000–200,000	1	50–1000 ft	Ambient	atm†	72–360	50–200	0.8–0.9
Propeller (and Kaplan) hydraulic turbine	5000–200,000	1	20–100 ft	Ambient	atm†	72–180	70–150	0.8–0.9
Small condensing steam turbine	100–5000	1–12	100–400 lb/in.^{2}	400–700 (200–370)	1–5	1800–10,000	200–800	0.5–0.8
Large condensing steam turbine	100,000–1,000,000	20–50	1400–4000 lb/in.^{2}	900–1100 (480–590)	1–3	1800–3600	500–1500	0.8–0.9
Gas turbine	500–20,000	10–20	70–100 lb/in.^{2}	1200–1500 (650–820)	atm	3600–10,000	500–1500	0.8–0.9

*1 ft = 0.3 m. 1 lb/in.^{2} = 6.9 kPa. 1 in. Hg = 3.4 kPa. 1 ft/s = 0.3 m/s.

by Eq. (6). For expansive fluids, they may be found by Eq. (7), where H and ΔW are as given above and C

$$\text{Jet velocity} = C\sqrt{2gH} = 8.02\ C\sqrt{H} \qquad (6)$$

$$\text{Jet velocity} = C\sqrt{2g\Delta W} = 223.7\ C\sqrt{\Delta W} \qquad (7)$$

is the velocity coefficient, seldom less than 0.95 and usually from 0.98 to 0.99.

Selected representative performance values of some prime movers are presented in Tables 1 and 2. *See* GAS TURBINE; HYDRAULIC TURBINE; POWER PLANT; STEAM; STEAM TURBINE; TURBINE.

Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996.

Primitive gut

The tubular structure in embryos which differentiates into the alimentary canal. The method by which the primitive gut arises depends chiefly on the yolk content of the egg.

Two-layered blastoderms. Eggs with small or moderate amounts of yolk usually develop into spherical blastulae which invaginate at the vegetative pole to form double-walled gastrulae. The invaginated sac extends in length to become the primitive gut. In some groups, such as the echinoderms, chaetognaths, chordates, and amphibians, its external opening, the blastopore, persists as the anus. The mouth forms as a new opening by contact of the opposite end of the gut with the skin, which then perforates. In other groups, such as the annelids, arthropods, and molluscs, the blastopore becomes the mouth, and a new anal opening is formed.

Three-layered blastoderms. Animals such as fish, reptiles, and birds, having more yolk than can be cleaved, form flattened gastrulae consisting of three-layered blastoderms surmounting the yolk. Mammals also belong in this group, although the yolk has been lost secondarily in all except the monotremes. The head is formed by a folding of the blastoderm upon itself. The endodermal layer within the head fold becomes the pharynx. This foregut is extended by an anterior growth of the whole head and by the union of lateral endodermal folds at its posterior boundary. In most forms, the hindgut arises by a similar folding in the opposite direction, the tail fold, at the posterior end of the blastoderm. The shark is an exception in which the hindgut is formed by the union of two lateral endodermal ridges beneath the notochord, much as the medullary plate folds into neural tube above the notochord. The midgut remains open until the endodermal layer of the blastoderm surrounds the yolk. The communication of the midgut with the yolk sac gradually narrows to an umbilicus as the foregut, hindgut, and lateral body folds press in from all sides, but the midgut floor is brought into position only with the absorption of the yolk sac. Mouth and anus form by contact of the pharynx and hindgut with the ectoderm at their respective ends of the body. Biochemical incompatibilities of ectoderm and endoderm cause such areas of contact to perforate and form the body orifices.

Teleost fish are an exception to the above account in that the true endoderm does not spread beneath the whole blastoderm, but is confined to the embryonic area. Along the axis of the body beneath the notochord, the endoderm condenses into a solid rod which then hollows out secondarily to form the primitive gut. *See* CLEAVAGE (DEVELOPMENTAL BIOLOGY); GASTRULATION; OVUM.

Howard L. Hamilton

Primulales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Dilleniidae of the class Magnoliopsida (dicotyledons). The order consists of three families: the Myrsinaceae, with about 1000 species; the Primulaceae, with about 1000 species; and the Theophrastaceae, with a little more than 100 species. These are plants with sympetalous flowers; that is, the petals are fused by their margins to form a corolla with a basal tube and

terminal lobes. The functional stamens are opposite the corolla lobes, and there is a compound ovary that has a single style and two to numerous ovules which usually have two integuments and are on a free-central or basal placenta. The Myrsinaceae and Theophrastaceae are chiefly tropical and subtropical woody plants, but the Primulaceae are mostly herbaceous and are best developed in north temperate regions. Primrose (*Primula*) and cyclamen are familiar members of the Primulaceae. *See* DILLENIIDAE; MAGNOLIOPSIDA; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Printed circuit board

The printed circuit board (PCB) has been in use for over a century, but it is still one of the most important parts of electronic devices, as the circuitry links all components together into an economical and efficiently integrated system. Once called printed wiring since they were first used to replace hand wiring in early telephone switchboards and radios, printed circuit boards are now ubiquitous. The printed circuit board consists of patterned arrays of conductors, such as copper, that are held onto and within dielectric materials, such as epoxy. But many other materials are used with dozens of processes to produce a myriad of circuit types that enable thousands of products.

History. The earliest documentation of circuitry is a 1902 patent by inventor Albert Hanson, who wanted to help solve the "density dilemma" of the telephone exchanges. The telephone switchboard required countless telephone lines to be brought into proximity to allow an operator to manually connect the correct parties together. Hanson recognized that individual wires were not the best answer and conceived of an invention where cut-out strips of metal were glued to a flexible nonconductive material, such as paper or cloth. This flexible wiring array would allow conductors to be placed close together in an orderly and more easily manufactured system. Hanson thus invented a circuit that increased density (more signals per area), which could reduce cost by mass production. Hanson's designs included conductors on both sides and multilayers. **Figure 1** shows the Hanson circuits from the patent figures.

Other early circuit inventors, including Thomas Edison, suggested using materials that could be painted on. Ideas moved from painting to printing, and the printed circuit concept was well on its way by the 1920s. World War II drove printed circuit technology to a new level, as electronics engineers sought robust and highly reliable assemblies for bomb and torpedo proximity detonators. The result was ceramic thick-film (CTF) circuitry, where ceramic-metal (cermet) inks were screen-printed onto alumina and fired in a high-temperature oven to fuse the conductor pattern onto substrate. This basic thick-film method, while improved, is still used today to fabricate ceramic printed circuits. Later, chemical etching (an adaptation of photoengraving) gained favor over direct printing methods. Most modern circuits are not really printed.

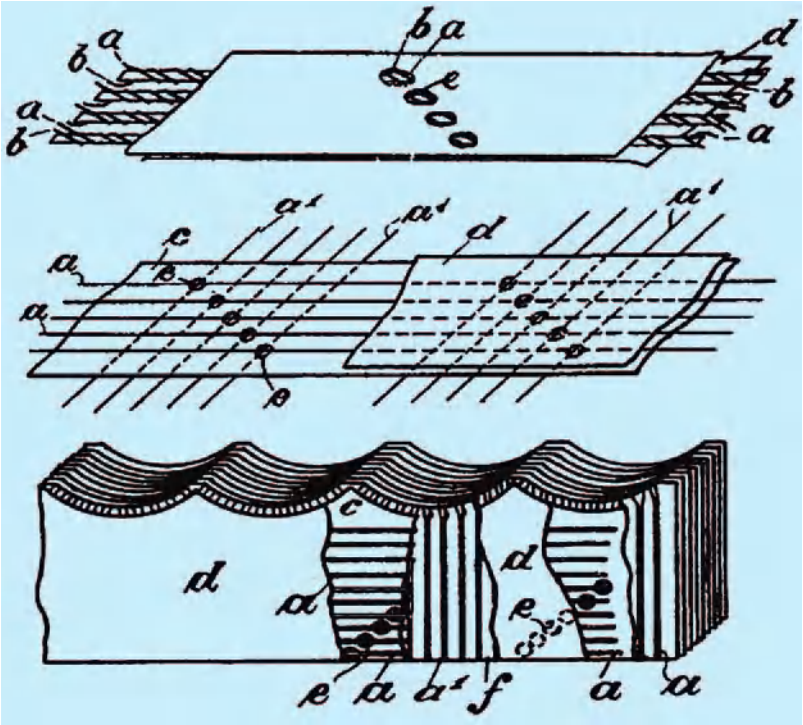

Fig. 1. First documented printed circuit; British patent 4,681, filed in 1903 by Albert Parker Hanson.

Importance. The printed circuit board is at the center of modern electronics products, and its value cannot be overstated. Every computer, Internet server, digital camera, cellular phone, television, pacemaker, and satellite has one or more printed circuit boards. Printed circuit boards have become specialized during a century of progress, and products range from tiny circuits for watches and hearing aids to extremely large ones for telecommunications and space stations. Today's circuits are efficiently mass-produced using technology from many disciplines, especially chemistry, physics, and mechanical engineering.

Materials. The printed circuit board uses two basic materials: dielectrics and electrical conductors.

Dielectrics. The dielectric must isolate the conductor array, while serving as the mechanical support platform. Although thousands of dielectrics exist, only a few are used by the printed circuit board industry. Dielectric materials can be divided into inorganic, including glass and ceramic, and organic, a much larger group of polymer-based compositions. *See* DIELECTRIC MATERIALS; POLYMER.

1. *Ceramics*. Most ceramic materials have excellent electrical insulation resistance and mechanical characteristics. Several fabrication methods have been developed to produce ceramic substrates and circuitry. Ceramics are distinguished by very high temperature performance that exceeds any organic substrate. Alumina (Al_2O_3), a strong hard material, is the most common ceramic circuit substrate used.

Aluminum nitride (AlN) is also used, it has good thermal and mechanical properties, and exceptional thermal conductivity for heat dissipation. *See* CERAMICS; INSULATION RESISTANCE TESTING.

2. *Glass*. Circuits are fabricated onto glass by the direct deposition of metals or other conductors. Glass is used for thin-film circuits, optical products, and flat-panel displays. Once the glass is metallized, electronic components can be attached by soldering and other methods. Glass circuits are much less common than ceramic and organic types. *See* SOLDERING.

3. *Epoxy-glass*. Thermoset (cross-linked) resins have long been used to make printed circuit board dielectrics, and nearly all plastics are good electrical insulators. Early plastics, such as Bakelite, amino resins, and phenolics, were used on some of the first printed circuit boards, but epoxy resins have become the de facto standard. Pure epoxy resin does not have the right thermal and mechanical properties for good printed circuit boards, but reinforcing it with inorganic fillers provides an adequate material. Electronic-grade glass weave (E-glass) is combined with epoxy resin and cured to produce a composite with higher strength and lower thermomechanical expansion than the polymer alone. Brominated epoxy is often added to reduce flammability (to pass Underwriters Laboratories ratings); this product is generally referred to as FR4. Different grades and types of epoxy-glass laminates are used for the many printed circuit board applications. Thickness can range from a few thousandths of an inch to very thick boards for large printed circuit boards that support heavy components. *See* COMPOSITE MATERIAL; GLASS; POLYETHER RESINS; POLYMERIC COMPOSITE.

4. *Polymer films*. Unreinforced polymer film is used to make flexible printed circuits (FPC)—a special, but extremely valuable subset of organic printed circuits. Flexible printed circuits must bend and are therefore made from thin polymer films without reinforcing glass. Polyimide (PI), one of the highest-temperature organic polymers, is inherently flame-retardant, and is the most common flexible substrate. It was introduced in the 1960s as DuPont's Kapton®. Today many suppliers offer polyimide circuit products. The next most important film is polyester, which is lower in cost but temperature-limited. Polyester generally is not used when soldering is required.

5. *Other materials*. Polyimide rigid composites are used for high-reliability rigid printed circuit boards, such as military assemblies. Although properties are excellent, this is one of the most costly printed circuit board dielectrics. Many specialty or lower-volume substrates are offered to meet special needs, such as fluoropolymers for very high frequency applications.

Conductors. All printed circuits use electrical conductors to carry power, ground, and electronic signals. Copper is by far the most common conductor and the industry standard. But alloys, other metals, metal composites, and even nonmetals are used as conductors. *See* CONDUCTOR (ELECTRICITY); COPPER.

1. *Copper foil*. Two basic types of copper are used, common electrodeposited (ED) and special rolled/annealed (RA), which has high fatigue resistance making it useful for flexible printed circuits. Copper thickness can have a wide range, but 3–37 micrometers is common. Thicker copper is used for high-current circuitry. Very thin copper cannot carry much current and is used with build-up methods discussed in "Manufacturing process" below.

2. *Cermet (ceramic/metal composite)*. Ceramic circuits often use conductive inks made from metals, such as silver/palladium mixtures, and a smaller amount of ceramic binder. These ceramic/metal mixtures are referred to as cermets. The cermet is printed onto dielectric and fired at high temperatures (800–1200°C; 1500–2200°F) to fuse the ink. *See* CERMET.

3. *Thin film*. Circuit conductors can be placed on glass, ceramic, and other nonporous dielectrics by vacuum deposition. Metal can be deposited as atoms and atom clusters by energetic processes in a vacuum chamber. Thin films of copper, silver, tungsten, gold, and many other metals can become applied circuitry as thin as 1 μm.

4. *Polymer thick film*. Polymer thick film (PTF) is the organic analog of cermet and uses a polymer binder to hold the metal particles instead of ceramic. Silver-filled ink is the most common, but carbon inks can also be used alone or mixed with silver ink. Thermoset epoxies, as well as solvent-soluble thermoplastics, are used to make printable inks, which are dried or polymerized to durable circuit conductors. Polymer thick film is a low-temperature system that can be used on nearly any dielectric, including paper.

Environmental regulations. Recent regulations, especially Reduction of Hazardous Substances (RoHS), restrict certain metals such as lead and cadmium. In the future, regulations could restrict halogens, including bromine. Printed circuit boards can no longer use the once popular tin-lead solder finish, and brominated epoxy may need to be replaced in epoxy laminates in the future.

Designing. The simplest design strategy is to involve printed circuit board designers in the product concept from the very beginning and make sure that the design team understands the technology. For example, a rigid board designer may not be qualified to design flexible printed circuits.

Computer-aided design. Computer-aided design (CAD) has been used for printed circuit boards for several decades. It would be nearly impossible to design and layout a complex printed circuit board without it. A printed circuit board can have dozens of conductor layers interconnected by vertical pathways (vias) numbering in the tens of thousands. CAD is an efficient routing tool (best wiring paths), and is used to check for design-rule compliance. CAD software is used to generate artwork for photoimaged circuits, as well as for "direct write" systems which are used

mostly for prototyping. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Computer simulations. Many modern circuits are more than wires. For example, high-frequency circuits have become an integral part of the electronics and circuit tuning may be required. Simulation software can virtually run a circuit design and measure electrical characteristics to modify and adjust the printed circuit board before it is built.

Tool design. The printed circuit board industry uses several hard tools such as dies for cutting out panels, but the trend is to move toward a data-driven "no tool" environment. Tools include artwork, a photographic film with clear and opaque (black) areas to define the exposure pattern for the photoimaging process. For tool generation, CAD files are typically electronically transmitted to laser photo-plotters that print the artwork. Metal cutting/blanking tools are also generated from CAD files.

Prototyping, testing, and qualifying. Designers recommend building prototype printed circuit boards, except for relatively standard products and simple modifications of existing parts. However, errors are made even with advanced CAD and simulation systems. A small run of printed circuit boards assembled with components, followed by testing, is a time-proven method. Even a "correct" printed circuit board may experience production and assembly problems that can be eliminated by adjustments to the design and artwork. The prototype, or a short production run on the factory line, can be used to generate parts that can be tested per the relevant specifications in order to pass or qualify the product.

Types (material-based classification). Printed circuit boards can be divided into two categories, organic and inorganic, but three common types are recognized: organic rigid (rigid board), organic flexible (flex circuit or FPC), and ceramic.

Ceramic/inorganic. These are specialty circuit boards that are very robust, extremely reliable, and heat-tolerant, but they are more expensive than other types (**Fig. 2**).

1. *Thick film*. The original ceramic circuits were produced by screen-printing cermet ink onto the substrate and firing it in a high-temperature oven. This is still a popular method. Although silver-palladium cermet ink is commonly used, copper and gold inks are available. Copper, since it can oxidize and become poorly conductive, may require an inert atmosphere, such as nitrogen, in the oven, which somewhat offsets the material savings.

2. *Thin film*. Thin-film circuits can use ceramic, glass, or other nonporous durable dielectric materials. Metal, including silver, platinum, palladium, gold, copper, titanium, and others, can be deposited through a mask to define a conductor pattern. Alternatively, the entire surface can be metallized and undesired metal etched away.

3. *Copper layer*. Copper may be bonded to the substrate surface by plating or vacuum deposition, followed by etching to generate the conductor array.

Organic. Circuits made from carbon-containing materials, especially plastics, were the earliest printed circuit board substrates and are the dominant materials today. They can be divided into rigid and flexible printed circuit boards.

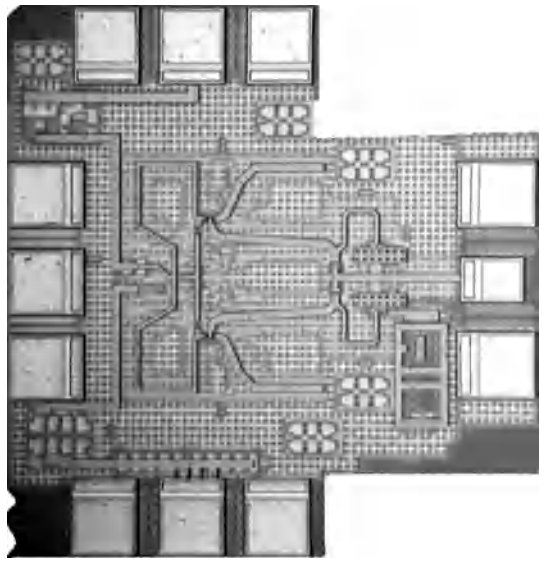

Fig. 2. Ceramic circuit for high-frequency radio.

1. *Rigid*. Typified by epoxy-glass FR4, rigid circuits are the largest class of printed circuit boards (**Fig. 3**). Rigid circuits are used in nearly every application unless special properties, such as flexibility or harsh environment capability, are required. Copper foil is applied to a lay-up of epoxy resin and glass weave, followed by heating in a press to polymerize the epoxy and bond the foil. Copper may be bonded to one side (single-sided) or both sides of the (double-sided) dielectric. The conductor patterns are generally formed by chemically etching.

2. *Flexible*. Flexible circuits comprise a thin (25- to 125-μm) dielectric plastic film with copper on one or both sides. Processing may be similar to methods used for rigid printed circuit boards, but with special handling to accommodate the flexibility. Continuous-roll (roll-to-roll) processes can be used to increase productivity and throughput. Flex is an enabling technology used for products

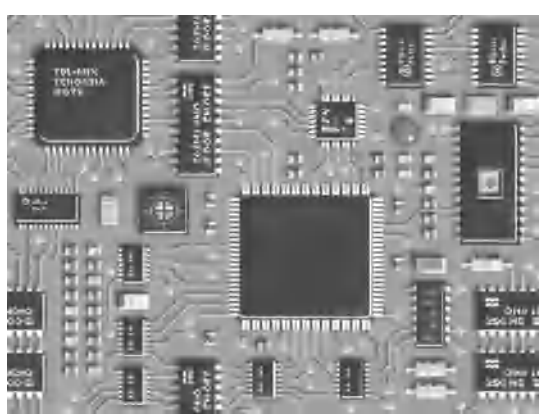

Fig. 3. Common printed circuit board made from etched copper on epoxy-glass substrate, shown with assembled components.

that combine motion and electronics, such as disk drives.

3. *Rigid-flex*. Rigid and flexible circuits can be combined by lamination to form printed circuit boards that are used in military, avionics, and some commercial products. Rigid-flex provides the maximum volumetric density, making it valuable for avionics and other complex systems. Designs typically bond the rigid substrate only to selective areas of the flex so that the final printed circuit board can be efficiently folded into a compact volume (**Fig. 4**). Some of the most complex and expensive printed circuit boards are rigid-flex designs.

4. *Molded*. The idea of molding a circuit from plastic resin has been tried with a modest level of success since the 1950s. Plastic resin can be molded into the required three-dimensional (3D) shape, followed by the addition of appropriate conductor patterns by several proven methods. The resulting printed circuit board can serve as the circuit board, interconnect, and housing. Three-dimensional molded printed circuit boards are produced today but mostly for specialty applications.

Special types. There are numerous specialty printed circuit boards for high- and low-volume products. For example, calculators use a specialty printed circuit board in which cost is paramount. Low-cost calculator circuits with integrated keypads can be made from thin polyester film printed with conductive carbon ink to avoid the cost of metal and chemical etching.

Thermal control. Two basic methods are used for increasing thermal dissipation in printed circuit boards used for power modules and other high-heat-producing products. Ceramics, such as high thermal aluminum nitride, can be used. Organics typically have poor thermal conductivity but can be modified with thermally conductive inner layers or cores. Printed circuit boards with cores are more difficult and costly to produce, and are used only where excessive thermal energy must be controlled.

Manufacturing processes. Dozens of circuit processes have been developed, but the basic processes deal with conductor pattern formation, or image placement. The pattern-generating method is often used to classify the printed circuit board production and can determine board performance. Conductor patterns must be accurately manufactured, but vertical (z-axis) interconnects must also be fabricated for printed circuit boards with more than a one conductor layer.

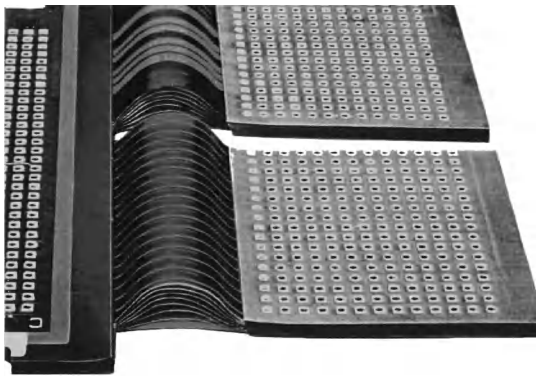

Fig. 4. Rigid-flex circuit commonly used for military and avionics.

Material preparation. The starting point for printed circuit board manufacturing is the substrate, and there are two classes: dielectric only, and laminates and clads with conductive material, especially metal foil, already mated to the dielectric. A clad is directly bonded metal with no adhesive. The first step is to prepare the substrate, laminate, or clad by cutting to the desired size for the printed circuit board line, if necessary. Secondary steps include cleaning, oxide removal, and surface preparation for the next process step.

Metallization. A laminate or clad does not require metallization since the foil layer provides the necessary metal for the conductor pattern. But some printed circuit board processes start with dielectric only and add metal during fabrication. Metal can be added by a mass process such as vacuum metallization or by a selective method that produces the conductor pattern directly. Metal plating is the most common process for creating patterned conductor arrays. Printing of conductive cermet or polymer thick film (PTF) inks can also be viewed as metallization, but many use the term only for mass deposition of metal over the entire substrate surface. *See* ELECTROPLATING OF METALS.

Through-hole formation. Double-sided and multilayer printed circuit boards generally require vertical connectors between layers. A hole must be formed through the substrate as the first step. Drilling has been the standard method for rigid (organic) printed circuit boards, while thin flexible circuitry has mainly relied on mechanical punching. Laser drilling, or machining, is increasingly popular as equipment costs drop and power efficiency increases. While ceramic substrates can be laser-drilled and cut, it is more common to punch holes in "green" ceramic, a pliable and workable material that has not been fired in an oven (kiln).

Conductor patterning (image placement). The conductive pattern-generating process determines the circuit density (conductors per unit width) of the printed circuit board and can significantly influence the cost. While dozens of patterning processes have been proposed and evaluated, printed circuit board manufacturing uses only a few.

1. *Fully additive*. In the additive process, the conductors are added to the dielectric. Many of the early circuit processes were additive, especially conductive-ink printing, and conductors were placed exactly where required. Additive copper methods fell out of favor early, but have become important, especially for high-density fine-line circuitry. Conductors can be printed for ceramic circuits, making this a true printed circuit method. Cermet ink is screen-printed with a high-precision screen of woven metal (or polymer) selected to give the desired ink thickness and feature resolution. The printing screen is coated with an ink-blocking emulsion,

having an open pattern corresponding to the intended conductor image. The pattern in the emulsion is generated by a photoimaging method using artwork produced by a laser plotter from the CAD files for the particular printed circuit board. Once printed, the ink is finished by heating to drive off volatile organics, and then fired at temperatures near 1000°C (1800°F) to fuse and form the ink into a durable metallic conductor. In the co-fire process, cermet ink is printed onto green (uncured) ceramic, and the substrate and ink are fired concurrently. The co-fire process also allows holes to be punched in the substrate prior to firing, so that the cermet ink can be printed into the holes to connect the top and bottom conductors.

Polymer thick film uses the same screen-printing methods, including through-hole printing; the primary difference is in the materials. It uses polymer binders and substrates to yield substantially lower cost, but with lower performance.

Metal, especially copper, can be selectively plated onto a substrate that has been sensitized with a plating catalyst. While the catalyst can be applied selectively, a more desirable method is to apply a plating resist, or mask, over the catalyzed surface to restrict and contain the plated metal. A plating resist also prevents the copper from plating outward from the catalyzed area. While fully additive copper can generate very high density circuits, the process is rather slow and expensive.

2. *Subtractive etching*. The subtractive patterning process begins with a copper foil laminate and removes, or selectively subtracts, the metal, leaving the desired conductor pattern. The most common method is to apply an etch resist and chemically etch away the exposed copper. The resist can be printed directly onto the copper laminate, then hardened with heat or ultraviolet (UV) energy, and finally submerged in an etching solution. Once the exposed metal is removed, the resist is stripped off the conductor pattern using another chemical solution. While the printing resist is low in cost, the circuit pattern resolution is limited to the capability of the printing method, usually screen printing. Although carefully controlled screen printing can resolve features down to 50 μm, this print-and-etch process is used for high-volume production at 150- to 200-μm center-to-center spacing (pitch) for better yields. Print-and-etch is one of the lowest-cost methods and is used for making a large number of low- to moderate-pitch printed circuit boards.

The subtractive etching process is also used with photoresists, where the initial pattern is produced in the resist using UV radiation. The resist is applied to copper laminate as a dry film or as a liquid, which is printed, sprayed, or otherwise applied, followed by drying. An artwork film of the conductor pattern can be placed against the resist and exposed to high-energy UV. The exposed resist areas become insoluble (negative-acting) or soluble (positive-acting) depending on the resist chemistry. The resist is then developed to remove undesired material, thus allowing etchant to contact and remove metal to produce the conductor pattern. The resist is stripped after etching. Photoetching is a popular printed circuit board process capable of fine-line geometries approaching 25 μm, but 50- to 100-μm minimum trace widths are more typical.

3. *Semiadditive*. The semiadditive process combines additive and subtractive in an efficient process capable of very high circuit density (less than 10-μm conductor widths). The method is increasingly popular for flexible printed circuits, which often require very high density. Metal, usually thin (2- to 5-μm) copper known as seed or bussing, is deposited over the entire surface of the dielectric by a vacuum deposition method such as sputtering. A plating resist is applied over the thin metal, and then imaged and developed to form a pattern. Metal, usually copper, is electrolytically plated using the thin layer of copper as the electrical buss. Once the desired thickness of copper is plated, the resist is stripped away. The remaining thin copper is easily etched away without damaging the relatively thick plated-up conductors. The method is more common with flexible circuits, since thin copper can be applied in a cost-effective and continuous-roll fashion. This is more difficult for rigid printed circuit board substrate. *See* SPUTTERING.

Connecting conductor layers. A double-sided circuit has a top and bottom conductor layer produced by one of the above processes. Most double-sided circuit designs require connections between the two layers at specific locations. The most common vertical connection method is the plated through-hole (PTH) [**Fig. 5**]. First, holes are formed through the substrate by drilling, punching, or lasing. The holes are catalyzed and then plated with a thin layer of electroless copper (no current needed). The electroless copper allows more efficient electrolytic copper plating to complete the process. The resulting hollow conductor looks like an eyelet but is called a barrel. More recently, smaller holes have been used that will plate shut to form solid through-hole posts, or solid vias. *See* ELECTROLESS PLATING.

Multilayer fabrication. There are two strategies for increasing circuit density: more lines per width using finer geometries, and additional conductor layers. Since circuit-patterning density is near its practical limit, multilayer printed circuit board fabrication is used to increase density. The two common multilayer (ML) methods are lamination of double-sided circuit

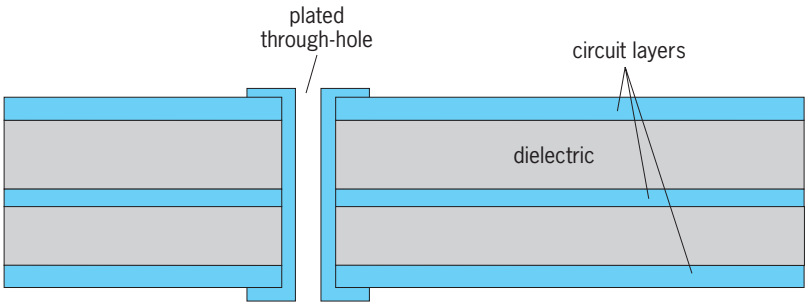

Fig. 5. Cross section of a plated through-hole for connecting circuit layers.

stacks and build-up alternate layers of conductors and dielectrics.

1. *Conventional lamination*. The conventional common multilayer process starts with double-sided circuits described earlier, but without any holes. The layers are carefully aligned, followed by lamination in a heated press that may also have vacuum capabilities to help eliminate air voids between layers. The laminated stack-up is then drilled to create through-holes, which are plated with copper similar to double-sided printed circuit boards, to connect the desired layers. This time-tested method is preferred. Rigid printed circuit boards, flex, and rigid-flex all use this process.

2. *Co-fired and co-lamination multilayers*. A commonly used process for ceramic circuits is to form the top and bottom conductor patterns, and the vertical through-hole connections, before combining the individual double-sided circuits. Circuit layers can be printed on green ceramic while simultaneously screen-printing into the through-holes. The patterned green layers are stacked, registered, pressed together, and then fired to cure the ceramic substrate, fuse the ink, and bond the layers into one interconnected multilayer circuit.

A somewhat analogous process has been applied to flexible circuitry, called co-lamination. Double-sided circuits are manufactured, including plated through-holes. The circuit layers are aligned together, and an interposer layer is inserted between each circuit pair. The interposer is a special laminating adhesive with conductive material that interconnects the adjacent circuit layers. The stack-up is heated in a press, whereby the adhesive bonds the layers together and creates the vertical connections between the layers. Co-lamination and co-firing offer higher density as well as assembly efficiency. While co-fired multilayer ceramic printed circuit boards are popular, the organic analog, co-lamination, has not yet caught on.

3. *Build-up multilayer*. The build-up process can use ceramic thick film, polymer thick film, or rigid printed circuit boards. In all cases, the layers are formed sequentially, one at a time. Ceramic and polymer thick film circuits both start by forming the first layer of patterned conductors. Next, dielectric is applied over the conductors by screen printing, or a coating method. Through-holes can be formed by screen printing or by using a photoimageable dielectric. Once the dielectric layer is finished, a second circuit layer is applied by screen printing, or some other means. The steps can be repeated until the desired number of layers is formed. Ceramic build-up has been replaced largely by co-firing, but the method is still used with polymer thick film.

Build-up can also be used with organic copper printed circuit boards. The first circuit layer is conventionally produced, and then a dielectric layer is applied and photoimaged to enable through-hole formation. Metal is next applied to the dielectric surface and into the holes by sputtering or plating, followed by patterning. The steps can be repeated to add more layers but is often limited to just the top layer, which can be a high-density pattern. The process can also start with a conventional multilayer circuit, using the build-up just for the high-density top and bottom layers.

Surface finishes. While copper is an excellent electrical conductor with good mechanical properties, it is also chemically active and will react with oxygen, carbon dioxide (slowly), and other airborne materials in the atmosphere. Some bare-copper printed circuit boards are sold, but most exposed copper is protected with a surface finish. The simplest approach is treatment with a solution containing an organic solderability preservative (OSP), which provides moderate protection. Although the organic solderability preservative is the simplest and lowest-cost method, protection will diminish over time and the preservative is easily abraded. A more durable finish is provided by plating metal over the copper. Gold has long been one of the most popular circuit finishes because it is inert and is suitable for wire bonding, which allows integrated circuit chips (die) to be directly connected to the printed circuit board. A barrier layer of nickel is first applied to the copper conductors to prevent the gold from diffusing into the copper. More recently, immersion silver has become a popular finish that is easily applied by dipping the printed circuit board into a silver plating bath. The immersion plating process is self-limiting and stops when no more copper is available at the surface. Modern immersion silver baths contain inhibitors to prevent tarnishing and general corrosion. Silver, gold, and organic solderability preservatives are the most common lead-free finish for solder assembly, but gold must be used for wire bonding and durable edge connectors.

Final testing. Although modern printed circuit board processes minimize defects, dozens of problems can arise during the many process steps. Contamination as small as a particle can ruin a circuit by producing a defect in the photoresist pattern. Essentially all printed circuit board makers electrically test finished circuits, and many use automatic optical inspection (AOI). The methods, especially automatic optical inspection, can also detect trends and thus help adjust the process before out-of-spec printed circuit boards are produced.

Depanelization (singulation). Manufacturing efficiency requires that the processes use their full area capacity, and this often means that there are many circuit boards per panel. Once a printed circuit board panel is completed and tested, the individual circuit boards are cut out or otherwise singulated; depanelize is a more common term in the rigid circuit industry. Rigid circuits can be sawed, punched out, cut with a water jet, or laser-cut. Thinner flexible circuits can be laser-cut or die-cut with a blanking tool. The printed circuit boards may also be partially precut to allow easy separation at this stage. Ceramic circuits can be precut, sawed, or laser-scribed to allow mechanical separation (snapstrate) later.

Basic applications. Wherever there is electronics, there will be printed circuit boards, but circuitry is also used for electrical-only interconnects including

light-emitting diode strip lighting. The range of circuit applications, and the corresponding designs, is amazing when one considers that a circuit can fit inside an "invisible" hearing aid, or fly over the Earth as a 100-ft-long (33-m) flexible power bus for the solar arms in the *International Space Station*. While high-volume products, such as personal computers, use standard printed circuit boards, tens of thousands of other applications require unique circuitry with special properties. Electronics connects the world and the printed circuit board connects electronics.

Interconnect only. Printed circuits, especially flexible types, may be used to connect electronic subsystems together as flex cables with no components. Flex circuits for inkjet printers are a good example of interconnect-only, where electrical power and signals are conveyed to a moving assembly. Stationary circuits, such as the behind-the-dash instrument cluster circuits, may also be connect-only products. Most flex cables are single-sided, but higher density can require double-sided cables or several flexes bonded together. However, the introduction of surface mount technology (SMT) provided the ideal component form factor for flex, and components may be easily added to cables.

Component assembly platform. From the very beginning, the printed circuit has been used as a platform for vacuum tubes, coils, and other heavy components. Large and heavy components is one of the reasons that the industry moved from thin flexible substrates to thick rigid boards to better support the weight. The solid-state electronics revolution, catalyzed by the invention of the transistor, still required an assembly platform for attaching transistors, preferably by soldering. Although lighter in weight, packaged transistors still used a somewhat rigid platform, and the domination of rigid printed circuit boards continued. The next electronic device revolution, solid-state integration in the form of the integrated circuit (IC), increased the demands for component platforms since the integrated circuit enabled much more powerful and complex products such as the personal computer. Today, most integrated circuits and passive components are packaged for surface-mount assembly, which enables much higher densities, full automation, and smaller lighter products. *See* ELECTRONIC PACKAGING; INTEGRATED CIRCUITS.

Chip carrier. Although most integrated circuit chips are generally provided in a package format that allows easy solder assembly to printed circuit boards, chips can be directly attached to printed circuit boards. Unpackaged chips (bare die) may be wire-bonded, connected using a tape carrier package (tape automated bonding; TAB), or directly attached using flip-chip technology. Some refer to bare-die attachment as chip-on-board (COB). Flexible circuits were the first printed circuit boards used as chip carriers, beginning in the mid-1960s. Small rigid printed circuit boards, both ceramic and organic, became popular in the early 1990s when packaging designs went from perimeter connections to the area array to increase the number of connections but keep size small. The plastic ball grid array (PBGA), a small printed circuit board with solder balls on the bottom, has been increasingly popular. Today, printed-circuit technology is used to make organic and ceramic packages, including the newer chip-scale package (CSP), which is only slightly larger than the integrated circuit chips.

End uses. Since it is not practical to list all the printed circuit board uses, the most important and more interesting ones are covered.

Aerospace. Aircraft and space vehicles generally have sophisticated electronics for sensing, control, guidance, and communications. Some may have very complex military electronics for defense and offense. Space and weight are typically limited for most of the craft, and reliability requirements must be extraordinary even in the harsh environments encountered in flight. Some of the most complex and unusual printed circuit boards are used in aerospace and often combine flex, rigid, and even ceramic types into a fully integrated system that can fit into a restricted volume.

Automotive. The automotive industry began using printed circuits about a half-century ago, and all types are used today. Rigid printed circuit boards are used for entertainment, safety, security, and engine control systems, although more heat-tolerant ceramics are often preferred for "under-the-hood" applications where temperatures can exceed 150°C (302°F). Flexible circuits can be used for the same applications as rigid, but flex uniquely solves instrument connections, where the circuit must be shaped to fit around the back of the instrument cluster (**Fig. 6**). Most flexible cluster circuits are connection-only types made from low-cost polyester film that is not readily soldered, but the trend is to attach components.

Computers and peripherals. Computers have always used printed circuit boards. The early mainframes developed by IBM in the 1960s used high-density, high-layer-count ceramic multilayer to achieve extraordinary reliability and thermal management. Personal computers use mostly organic printed circuit boards, but flex is used where there is motion and as the interconnect to flat-panel displays (FPDs). Disk drives use very thin and tiny flex circuits that allow the read-write head to move over the spinning media disk. Remarkably, these flexible disk-drive circuits may be flexed a billion times without damage. Computers remain one of the largest printed circuit board markets.

Consumer electronics. Consumer electronics is a very large and diversified end-use area for mostly organic printed circuit boards, but ceramics may be used for radio-frequency applications and some packages. Cameras, players, educational toys, and hundreds of other products use printed circuit boards. Even rechargeable smart batteries use circuitry. Rigid printed circuit boards are the most common, but small products often use flex circuitry to reduce the size and weight of the final product.

Industrial. Thousands of industrial machines use printed circuit boards since nearly all have

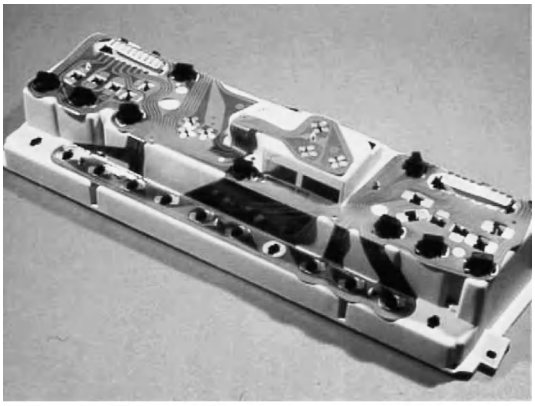

(a)

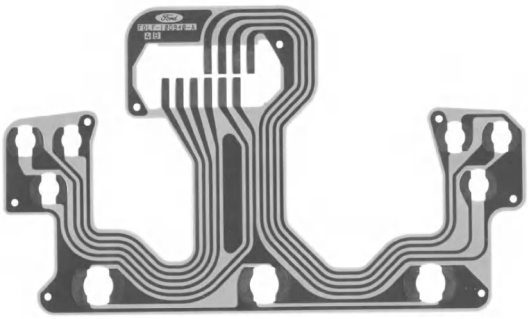

(b)

Fig. 6. Flexible circuits. (*a*) Assembled automobile instrument cluster. (*b*) Individual flexible circuit.

electronics or electrical devices that need sensing, monitoring, and control.

Medical. Medical electronics is a large and growing sector. Medicine has advanced substantially by the adoption of electronics, and this trend is accelerating. Medical electronics circuitry applications can be as simple as a home-use blood pressure meter or as complex as a CAT scan machine, requiring the power of a supercomputer. All types of printed circuit boards are used. Sensors and electrodes can use flexible circuits, including those made with polymer thick film technology. Many control panels also use flex in the form of membrane switches, which permit sealing of the unit for safety. Wearable and implantable medical products with specialized printed circuit boards include hearing aids and pacemakers.

Military. Military electronics remain an important market for printed circuit boards, with stringent specifications (MIL-STD-883F) for harsh-environment reliability. Avionics is one important area of military electronics. All military vehicles, from large ships to humvees, add more electronics of higher sophistication are adopted and advanced. Portable electronics is also an important sector for communications, detection, sensing, and control. Portable military electronics is increasing as the military strive to be a more flexible and nimble force.

Telecommunications. Communications is now the driver for electronics, a position once held by computers. Communications continues to advance at an accelerated pace, and each product needs one or more printed circuit boards. All types of circuits are used, and each can play a unique role. Ceramic circuit modules are ideal for higher radio-frequency applications. Rigid printed circuit boards continue to be the standard for nearly all products, but may be assisted by flex when a small-form factor requires a thinner substrate, conformability, or motion. Cell phones are the highest-volume telecom market, but newer personal products, such as Apple's iPod, are now a large outlet for rigid and flexible printed circuit boards.

Special applications. The above applications cover a very wide range, but there are others worth noting. The *International Space Station's* solar array power bus circuit may be the world's longest circuit and the most space-resistant. These circuits, exceeding 30 m (100 ft) in length, are made of very robust polyimide, coated with silica to prevent damage by the energetic solar wind. Tiny flexible circuits are being tested in experimental site-restoration devices that will eventually be connected to the retina or optic nerve. In a typical office, school, or library, there are hundreds of printed circuit boards, mostly hidden inside products. While often unseen, printed circuit boards enable the world of technology, a requisite for modern existence.

Future directions. Circuits may switch to nonmetals in the future. Nanotechnology has been able to produce carbon nanotubes, which can be used for transistors and even circuitry. Nanoimprinting and ink-jetting of materials is advancing. While metal conductors have been used for 100 years, a change is almost certain during the coming century.

Ken Gilleo

Bibliography. R. Cadenhead and D. DeCoursey, The history of microelectronics, Part One, *Int. J. Microelectr.*, 8(3):14–30, September 1985; C. F. Coombs, *Coombs' Printed Circuits Handbook*, 2001; A. Hanson, British Patent 4,681, filed, 1903; K. Gilleo, *Handbook of Flexible Circuits*, 1992; K. Gilleo, The history of the printed circuit, *Print. Circ. Fab.*, 22(1):18–21, January 1999; R. S. Khandpur, *Printed Circuit Boards*, 2005.

Printing

The processes for reproducing text and images using ink on paper. These include all the processes from typesetting through finishing of the printed product. Prepress operations for conventional plate processes consist of layout and design, typesetting, and platemaking. Traditionally, these were manual labor-intensive operations, but now most are digital, using desktop publishing systems.

There are five general printing categories: (1) relief printing, including letterpress and flexography; (2) planographic printing, including offset lithography, screenless lithography, collotype, and waterless printing; (3) intaglio, including gravure and engraving; (4) stencil and screen printing; and (5) digital

printing, including electrostatic, inkjet, and laser printing.

Each printing process has specific requirements for its printing plates or image carriers. The printing press unit has a cylinder for mounting the plate and an inking system to feed ink to the plate, as well as a means for feeding paper or other substrates into the printing units and a delivery device for collecting the printed product. Digital printers differ, depending on the process; for example, electrophotographic printers are more like high-speed copiers.

After the sheets are printed, most are put through some postpress finishing operations to make a functional product. Postpress operations include folding, binding, embossing, die cutting, foil stamping, and so on.

Evolution of Modern Printing

The modern printing industry is the result of six major inventions and developments. (1) About 1550, Johannes Gutenberg invented the process of duplicating movable type and printing from it using special inks and a press designed from a wine press. The printing industry remained essentially unchanged for over 350 years. (2) In 1799, Alois Senefelder invented lithography, which provided a practical and convenient means of producing illustrative materials. Previously, woodcuts and copperplate engravings, both requiring highly skilled artists and craftspersons, were the only means of producing illustrations in printing. (3) About the same time, the industrial revolution introduced mechanization to the printing press, first by steam power and later by electricity. The design of multicolor sheet-fed, web-fed (rolls of paper), and offset presses followed. (4) Later in the nineteenth century came the invention of photography, first by the silver halide process of wet plates, followed by dry plates and film, the glass screen for halftone photography which made it possible to produce illustrations photographically, and color films in the twentieth century; and then, photomechanical processes for making printing plates were developed. These techniques stimulated the invention of gravure, collotype, and other processes such as photoengraving for letterpress, albumen and other bichromated coatings, and photopolymers. (5) The Linotype machine, invented by Ottmar Mergenthaler in 1886, mechanized typesetting; teletypesetting followed in the 1920s and phototypesetting around 1950. (6) The introduction of electronics and computers in the early 1960s revolutionized the printing industry and many of its processes and techniques.

To put these developments into perspective, it was 440 years from the time that Gutenberg invented the process of duplicating movable type to Mergenthaler's invention of the Linotype machine. However, it was only 68 years from the introduction of the Linotype machine to the invention of phototypesetting by the Harris Corporation in 1954. In the graphic arts, the trend of accelerated technological change continued. Fourteen years later, in 1968, satellite data were transmitted simultaneously to multiple sites to produce *The Wall Street Journal*. Computer graphics on cathode-ray tubes took hold in 1974, first for engineering, architectural rendering, automobile design, and then for graphic arts applications. In 1977, the Mitsushita Corporation in Japan produced the first prototype of a color monitor that printed out the content on its display in full color. Mitsushita built an inkjet printing system into a television set to demonstrate the potential of integrated media. This was the forerunner of a desktop publishing system. In 1979, the first practical combination of multiple graphic arts technologies from Scitex in Israel demonstrated how nearly all prepress functions (page layout, retouching, typesetting, graphic arts photography, and imagesetting) could be integrated in one system, called the Response 300. In 1980, Gannett demonstrated that satellites could transmit color data to produce the news paper *USA Today*. *See* COMPUTER GRAPHICS; MICROCOMPUTER.

By 1980, technological transitions in the printing industry were occurring so rapidly that it became impossible to point to individual developments. From 1980 to 1985, deregulation of telecommunication in the United States opened the airwaves for broad commercial applications of communication media. From 1985 to 1990, desktop publishing technology evolved and matured. Multimedia applications grew and matured between 1990 and 1995. And on-demand digital color printing was the focus during 1995 to 2000.

The use of electronics and computers has rendered obsolete many techniques and led to the creation of new systems, equipment, and processes that have increased the speed precision of production, including scheduling and control; have simplified the operation of complicated systems in prepress, press, and postpress; and have improved the consistency and quality of the output of the printing processes.

Digital imaging and printing has resulted in the development of three specialized printing markets: on-demand printing, variable-data printing, and short-run color printing. On-demand printing produces documents at the time they are needed and usually within 24 hours. Variable-data printing allows for customized and personalized text or graphics. Short-run color printing allows for full-color printing in runs from 100 to 5000 impressions, which previously had been too expensive to print by conventional processes.

Composition

This is the assembly of all type, art, illustrations, and other elements into the final page format or layout for the press. This was traditionally a manual operation, requiring paste-up of page elements as paper prints or stripping of films on a flat, but now it is done digitally by desktop publishing.

Typesetting. Typography is concerned with the visual impact, effectiveness, and readability of the printed piece. Typesetting is the basic method for setting type, including cast metal or hot type (set either by hand or machine) and phototypesetting.

These methods have given way to electronic imagesetting. The term "typesetting" is independent of its output technology. *See* TYPE (PRINTING).

Metal composition. Hand typesetting, similar to Gutenberg's method, is done with individual metal characters assembled into lines. The individual characters or letters are placed in a composing stick, which is held in one hand, while the letters are selected from a type case with the other hand until a full line is set. Blank spaces are inserted between words to justify the line, that is, fill it out to the desired width. This is repeated line by line until all the copy has been set. For line spacing, or leading, metal strips or slugs called leads are inserted between the lines. The process of hand composition is very slow, about two characters per second, and is used mainly for fine art printing.

Machine composition, which is now obsolete, was a major advance in typesetting, as it increased speed to about five characters per second. Machine-set type was produced on any of several machines. Linotype and Intertype machines cast one line of type at a time. They cast a one-piece line or slug of type to a predetermined length. On both machines, at the touch of a key, a matrix, which is a mold for each character, was released from a magazine or storage case. When all the matrices or characters for one line were assembled, the line was automatically justified by spacebands or wedges, which mechanically expand the space between words. The line moved into the casting mechanism, where molten metal was forced into the character molds. Some line-casting machines set type semiautomatically from perforated tape. These tape-controlled systems sped the typesetting operation to about 10 characters per second, allowing one operator to handle several linecasters.

In all these methods, the lines of type were locked up in a chase, which is a heavy rectangular steel frame, on a large table known as a stone. Empty spaces were filled up with "furniture" of wood or metal blocks, and quoins (steel wedge-shaped devices) were placed on the sides of the form and tightened slightly until all the parts were made level, after which they were tightened to hold the form securely in place.

Phototypesetting. Metal typesetting gave way to phototypesetting, which reduced the number of production steps leading to the completed printing plate. Phototypesetters varied in electronic sophistication, automation, imaging speed, and typographic versatility. They produced composition from simple to full-page multicolumn formats combining text and images to be output on photographic paper, film, or polymer-based plates.

All photographic typesetters consisted of three elements: a master character image, a light source, and light-sensitive material. Phototypesetting systems have evolved through three generations of changes. First-generation units were adapted from linecasting. In these machines, the master character image was carried on the matrix, which contained a film negative of the character that was then photographed instead of cast. Second-generation units used electromechanical principles. Third-generation units were electronic, using cathode-ray-tube (CRT) or laser technology. Also known as digitized typesetting, this process produced characters as a series of minute dots or lines, from which the resultant image was transferred to photographic material. Because third-generation typesetters employed a beam of electrons, they were capable of setting type at extremely high speeds (500–1000 characters/s).

All second-generation phototypesetters had master character sets that included all of the characters and symbols for a particular type style that were stored as disks, film strips, grids, or film segments mounted on a drum. In most phototypesetters, the master character image was enlarged by individual lenses or by a zoom lens to achieve the final type size.

In using a second-generation phototypesetter, characters were selected by referencing certain marks on the photomatrix. When the proper mark was reached, a burst of light from a xenon flash lamp exposed the character through the master font image. The beam of light was directed to the correct optical path by mirrors or prisms; the character was enlarged or reduced by lenses and then positioned properly on the photosensitive output material. Because newspapers were pioneers in the use of phototypesetting, its speed has been measured in newspaper lines per minute (lpm). A standard line is 8 point in type size, 11 picas in width, with 30 characters. Most second-generation phototypesetters averaged 50 lpm, with a high of 150 lpm (25–75 characters/s).

Third-generation phototypesetters were capable of greatly increased speeds (over 1000 lpm) and created complete pages of text, headlines, and line-art graphics in position. Characters were stored as digital information that was converted into dot images. In operation, patterns of the required characters were called from memory and imaged in the correct size at precise locations on the page or format. The finished composition was then exposed from the tube to the photographic paper or film.

Although character images appear solid with accurately contoured shapes after imaging and printing, they are actually patterns of characters consisting of adjoining sweeps of scan lines, similar to the images created on television screens. The line density or number of scan lines to the inch (resolution) determines the typographic quality of the imaged characters and the imaging speed of any particular digitized system. The more scan lines to the inch, the higher the resolution, or sharpness, of the characters and the slower the imaging speed.

Electronic imagesetting. Typesetting now is often done using an electronic imagesetter to produce the films for platemaking. In imagesetting, photographic paper or film is exposed using raster or vector techniques. Imagesetting uses electronic files instead of conventional camera technology. It yields higher quality and less expensive printing.

Imagesetter users are under constant pressure to produce quickly and inexpensively, not only because

electronic media are raising customers' expectations on throughput times, but also because print runs are becoming shorter.

Some imagesetters use multiple lasers and a 30,000–40,000 rpm spinning mirror that traverses the inside of a drum. The lasers sit next to each other and use the same optical path. Multiple lasers increase speed. To keep costs down, some imagesetters use a single beam that is split, with the beams being deflected. A five-faceted spinning disk with a holographic deflection surface is used to direct the laser beam on to the film surface. Each rotation of the deflector produces five scan lines, making it up to five times faster than a conventional mirror-based system at the same rotation speed.

Platesetters are high-resolution imaging devices that use a laser to expose a photosensitive plate. This process is also known as computer-to-plate (CTP). It produces better quality than the traditional film-based process. CTP can cost less and allows for quicker turnaround times, but requires investment in equipment, training, and changing workflows.

Desktop publishing. Desktop prepress has evolved as a process for preparing pages for reproduction on printing presses, copying machines, or electronic printers. The multipurpose nature of the personal computer and the increasing pagination ability of word-processing programs have advanced desktop publishing. The essential part of this approach is the use of typographic characters and the integration of boxes, shapes, line art, low-resolution pictures, and other graphic elements into page form.

It is more appropriate to define this application as page processing, since its approach to putting pages and documents together is applied at all levels of workstation and system configurations. The use of high-level graphics-oriented workstations with multitasking functions permits increased productivity.

A further step in the evolution of imagesetting, incorporating the advances that have been made in digitization and scanning, is the integration of text with pictures, illustrations, diagrams, and so on. This has brought typesetting head to head with developments in page composition systems. Once a page is in a digitized form, it can be used in many ways. It can remain in storage, or it can be transmitted to remote locations where it can be output to an imagesetter, a photographic film, a platemaker, an inkjet printer, or an electronic reproduction system.

Page makeup. Most programs for page makeup can position type so that manual assembly of the page elements is reduced. Heads and text can be output according to their page position. To accomplish this, commands must be given to the output device in coded form (by menu or keystroke sequences) that translates the positioning information into machine instructions. The term "computer-assisted makeup" refers to workstations that show positions of image elements and actual or simulated typeface, point size, line length, leading, and so forth as they will appear in final pages. Passive computer-assisted makeup devices are sometimes known as soft typesetters or preview screens, since they show only how the typesetter will set the information according to the commands. Interactive computer-assisted makeup devices allow the operator to make changes in the typographic format on the screen and to see instantly how the changes will appear on the page.

Almost all modern systems include the ability to view type, graphics, and pictures on-screen as a simulated page with all the items in position. This can be in a preview mode, where coded text is translated into a screen image, or interactive, which involves typographic context at all times. The interactive integration of text and graphics has no single descriptive name; hence, the term "publishing" is often used as a synonym.

Page description language. A page description language is a computer language or software that describes an entire page, including text, graphics, lines, and halftones, as a series of codes and allows for viewing or output on any device capable of decoding the language.

PostScript is the most popular page description language. Invented by Adobe Systems, PostScript consists of software that forms the desired image on an output device such as a laser printer or an imagesetter. PostScript is commonly used for text and line art. In the latter case, it is referred to as encapsulated PostScript (EPS).

PostScript's advantage is that the code is converted to images in an interpreter in the output device, not the computer. Therefore, the images are readable or printable regardless of the computer platform being used.

Essentially, PostScript allows for device independence, or the ability to generate virtually identical output on devices made by different manufacturers, so long as they can interpret PostScript commands. The original version of PostScript is known as PostScript Level 1. It was one of the most important elements in the invention of desktop publishing. The latest version of PostScript, called PostScript 3, provides support for digital printing, image setting, and color separation and correction.

Raster image processing (ripping). Imagesetter raster image processors (RIPs) are software-based running on standard, fast, and easily replaceable workstations. A RIP is a microcomputer that uses a laser to create images out of lines of dots (rasters). It translates commands in a page description language, such as PostScript, or in other formats, such as PDF, TIFF, EPS, and JPEG into a precise image of a page. Everything a RIP produces is in pixels. The RIP then sends the rasterized data to a recording device (imagesetter) for output on paper, film, or plates.

To create a halftone dot, the RIP divides the area to be imaged into halftone cells. Each halftone dot then is created out of the pixels available within each halftone cell. The speed of a RIP generally determines how fast a printer can produce the first copy of a page.

Preflighting. Preflight is the process of checking an electronic document before it is moved into the

printing process. The process makes sure that all the elements are correct and within the specifications for printing. Since nearly all work submitted to printing companies is in electronic form, files sent from a remote site or delivered on a disk must be verified for consistency, presence of graphics and font files, and proper selection of colors. Preflight checks include color mode, cropping, bleeds, photo sizes, resolution, retouching, masking, gray balance, proper spelling, page order, type size and style, trapping, dot gain, flattening of layers, screen lines per inch, files saved in proper format, and if the file has the most recent changes.

Process Photography

Process, or graphic arts, photography comprises the photographic techniques used in the traditional graphic arts processes. Until the introduction of electronic scanners, color electronic prepress systems, and desktop publishing systems, all the conventional printing processes used photography for the production of negatives or positives to make the printing plates by a photomechanical process. Today, process photography is obsolete, as film is rarely used in the printing processes.

Copy preparation. The original or copy is the starting point in the printing processes. It can be in many forms, including text and/or pictorial matter such as line drawings, diagrams, black-and-white and color photographs, color transparencies, paintings, other art subjects, and subjects created on computers. In some cases, the original image can be produced directly on the plate. In all other cases, the original is converted for reproduction by photomechanical or digital means.

In printing that uses traditional photographic techniques, originals are photographed to convert them into the proper positive or negative films for the plate processes by which they will be reproduced. The copy must be clean, sharp, and in focus. If the final printed job consists of type and pictorial matter and everything is at the correct size, the typeset copy and pictures can be pasted-up in position before photography. On complicated jobs, such as color advertising, magazines, and books, in which the copy usually comes from a number of sources, the separate pieces of copy are photographed at the correct size and assembled in the film form. This completed form is known as a flat, which is ready for platemaking. Digital copy preparation, where all copy is assembled on a monitor, has rendered these manual and film techniques obsolete. Hence, copy can also be prepared electronically. Computers can be used to generate graphics, and special software and design stations have been developed to produce layouts, design artwork, and compose pages on a monitor and to output fully composed pages directly on a plate or printing press cylinder.

Films. Photomechanical processes use three types of photographic emulsions. High-contrast orthochromatic or panchromatic films are used for line and halftone photography. Low-contrast orthochromatic commercial types of films are used for single-color gravure and for other continuous-tone photography. Low-contrast panchromatic films are used for color-separation photography. Special films are made for scanners and imagesetters, and special daylight films are available to make contact films for image assembly. Rapid-access processing using continuous-tone developers is used to speed development of laser-exposed films. The favored film substrate is polyester film, which has good mechanical properties and dimensional stability. *See* EMULSION; PHOTOGRAPHIC MATERIALS.

Halftone photography. Halftones are an essential part of the print production process. They are optical illusions that are necessary to compensate for the inability of most printing processes to reproduce photographs and art subjects.

Printing processes, such as lithography, letterpress, and screen printing, cannot print continuous-tone reproductions such as photographs and art subjects that have many intermediate tones between the highlights and shadows. These binary printing processes can print only a solid impression of the ink on the press or no ink at all. For these processes to reproduce photographs and art subjects, the appearance of continuous tones must be simulated by the optical illusions produced by images using the halftone principle.

The human eye, at a normal reading distance of 10–16 in. (25–40 cm), is not capable of distinguishing two points separated by about 0.004 in. (0.0016 cm). This is equivalent to a linear screen ruling of about 125 lines per inch (lpi) and 125 blank spaces with the same width as the lines. When two such rulings are superimposed at right angles, a cross-line screen of 125 lpi is produced.

Halftone screens. When a halftone screen is positioned in front of a photographic film and an image is photographed through the screen, the dots formed on the film are proportional in size to the amount of light reflected from the corresponding area on the original. These dots are amplitude-modulated; that is, they have equal spacing between the centers of the dots and vary in size or area. Screens with finer rulings than 125 lpi (such as 133, 150, up to 500 lpi) are not resolved by the normal unaided eye, reproduce finer detail, and produce higher-quality reproductions. Screens with lower rulings (such as 60, 80, 100, and 120 lpi) are visible to the naked eye, show less detail, are easier to print, and are used for newspaper and less critical printing.

Tone reproduction. Tone reproduction (**Fig. 1**) and contrast are two important conditions that determine the quality of the reproduction. By using a stepped gray scale, good tone reproduction in halftone photography is achieved when the darkest area (shadow) of the subject prints as a solid and the lightest area (highlight) as a white, with no dots in either area. The intermediate tones of the gray scale will have varying sizes of dots, from about 3% dot area in the light or highlight end to about 95% dot area in the dark or shadow end of the scale. High contrast is achieved when two or three steps of the gray scale in the shadow end print solid or when several

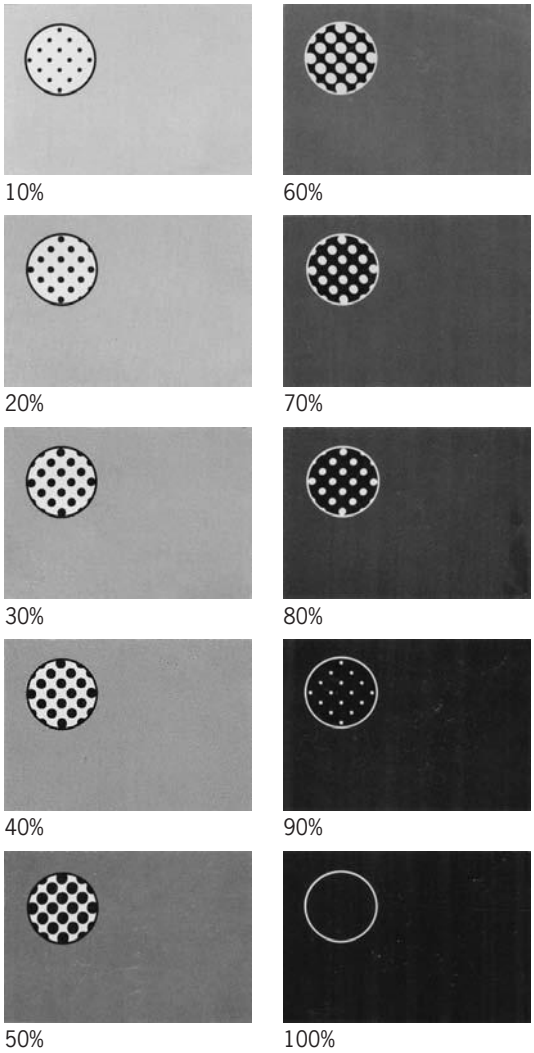

Fig. 1. Gradation and magnification of tones. (*From The Pocket Pal, 18th ed., International Paper Company, 2000*)

steps in the highlight end print white, with a corresponding increase in density difference between the other steps of the scale. Low contrast is achieved when the solid contains 80–90% dots or the white end of the scale has 10–20% dots, with a corresponding decrease in density in the rest of the gray scale. A number of special techniques, such as highlighting or dropout, posterization, and flashing, can be used to change the contrast of the reproduction in local areas.

Digital halftones. Most halftones now are produced digitally in raster image processors or in imagesetters. Digital halftone algorithms have been developed to simulate photographic halftone reproduction with screen rulings, screen angles, and amplitude modulation. These have the same limitations as photographic halftones, with the additional problem that 15 and 75° angles have irrational tangents because the tangent cannot be expressed as a ratio at two integers (horizontal and vertical spots), which results in very long computation times. Some algorithms avoid this problem by using rational tangent angles of 18.4 and 71.6°, but the use of these angles requires varying the screen rulings for each angle to avoid moiré patterns. Conventional halftone screening is called amplitude-modulation (AM) screening. Amplitude means size, so AM screening breaks up an image into dots of varying sizes to simulate the original image. The use of frequency-modulated (FM; also called stochastic) screens, in which all the dots are the same size but vary in spacing, eliminates many of the problems of amplitude-modulated screens as they are independent of screen rulings and screen angles. *See* MOIRÉ PATTERN.

Amplitude-modulated screening. During the 1980s, many advances were made in computer hardware and software that allowed for increased electronic image alterations such as retouching, color correction, and gray component replacement. At the same time, advances were made in computer storage and raster image processors for more efficient image data processing. Along with the ability to electronically scan, digitize, store, retrieve, and manipulate images, there were advances in digital screening. However, because color electronic prepress systems remained linked to high-end electronic color scanners, digital screening centered on the simulation of the traditional amplitude-modulated (AM) halftone dot. *See* COMPUTER STORAGE TECHNOLOGY; IMAGE PROCESSING.

Although it was recognized that dot shape played a significant role in image detail, the electronic digital imaging techniques of the mid-1980s still centered on a fixed screen ruling (lines per inch or lines per centimeter), and the amplitude (size) of the dots was modulated (increased or decreased) relative to the gray levels being reproduced. This type of digital screening is known as a deterministic or order dither, which uses an error distribution process to assign halftone dots to a bitmap in a coherent area (**Fig. 2**). That is, each halftone dot is usually produced from four scanned pixels. The pixel is divided into a matrix of single recording dots. For example, a typical matrix of 6×6 single dots would produce a halftone dot consisting of $4 \times 6 \times 6$ or 144 single dots.

The number of scanned pixels and the selected recording dot matrix depend on the imagesetter

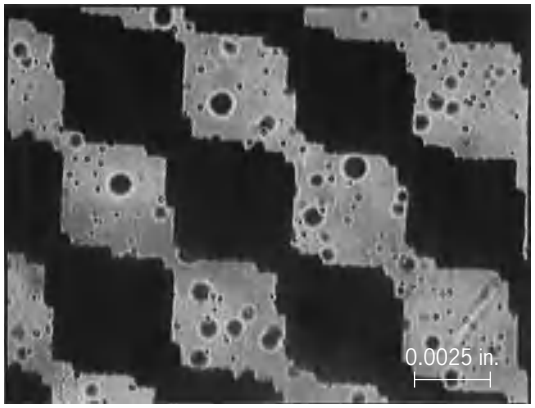

Fig. 2. Photomicrograph showing a typical AM deterministic digital halftone, 50% tone at 150 lines per inch.

used, but 6 × 6, 8 × 8, and 12 × 12 matrixes are common. The recording dots form a coherent area on a bitmap that consists of several contiguous single recording dots to produce an amplitude-modulated halftone dot. In general, the coherent areas are selected to produce symmetrical dot shapes, such as square, round, or elliptical, but they can also be arranged to produce irrational dot shapes. Thus, a dot arrangement in the form of a bitmap can be defined for each tone value and stored in memory.

Although the ordered dithering process and digital AM screening have proven to be an excellent method for producing halftones, there are limitations. There is a physical and mathematical relationship between the number of gray levels, the dots per inch, and the screening ruling, such that there is always a trade-off between resolution and the number of reproducible gray values. Also, the algorithms for conventional screen angles of 15 and 75° can never be exactly calculated, and require long and burdensome computational times because of their irrational tangents.

Stochastic screening and frequency-modulated screening. Stochastic and frequency-modulated (FM) screening allows outputting high-quality color images at lower resolutions by using a different approach to screening technology. Unlike amplitude-modulated (AM) screening, which varies the size of the dot to simulate different shades of color, stochastic screening keeps the dot size the same and varies the number of dots used. Stochastic screening uses randomly placed dots instead of the traditional halftone dots aligned along visible screen angles (**Fig. 3**). Stochastic and FM screening have become synonymous, although FM screening is not as random as the word "stochastic" implies. In effect, most stochastic screening products available are variations of FM screening.

The concept of stochastic screening is that dots placed randomly will not cause moiré patterns. Completely random positioning does not work well, however. It may prevent interference patterns, but it can cause graininess and other visual artifacts. FM screening products apply advanced algorithms to determine the optimum placement of dots to appear random, yet not cause visual artifacts.

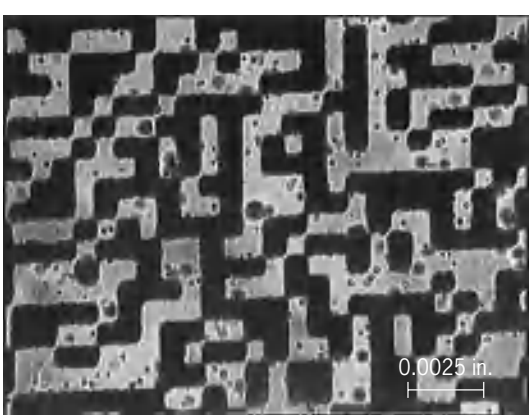

Fig. 3. Photomicrograph showing a typical FM stochastic digital halftone, 50% tone at 2400 dots per inch.

In FM screening, the concepts of screen angle and frequency no longer apply. Because the dots are randomly placed, there is no direction to the dots. The variable spacing of the dots means there is no fixed spacing and therefore no screen frequency.

Applications for FM screening span the industry, including newspapers, magazines, and commercial printers. Commercial printers who do high-end color work, such as brochures, art reproductions, and annual reports, are interested in FM screening as a way to achieve the 600-lpi-equivalent. Commercial printers with imagesetting operations are good candidates for FM screening because they can control the entire output production cycle. With the ability to generate custom FM screens for any combination of paper, ink, and imagesetter parameters, printers can set up a completely controlled environment for their full range of work.

Magazine printers with imagesetting operations can use FM screening to produce high-quality color. Once the dot gain calibration has been performed in collaboration with the printer for the press, paper, and ink, the magazine can use these settings for every production run.

Newspapers can also use FM screening to good advantage. Newspaper industry tests with frequency-modulated screening show higher-quality color and image detail than with conventional screening because of the relative coarseness of the conventional screen used with newsprint. Newspapers can output pages more quickly at lower resolutions, and the color images will be of much higher quality than if conventional halftone screening were used.

Color Reproduction

The basic principle of color reproduction in modern color printing of pictures lies in the use of only three primary colors to create the illusion of seeing in the reproduction the hundreds or more colors the eye sees in the original. Color printing with inks is the controlled proportioning of red, green, and blue light, and this is accomplished in an indirect way called the subtractive process. Printing is done with yellow, magenta, and cyan inks because these colors ideally are complementary to blue, green, and red respectively, and each can individually shut off the reflectance of red, green, and blue light from the paper surface.

Until 1893, colors were commonly printed separately with ink on paper for commercial printing, book illustrations, and art picture reproductions. As many as 40 colors were sometimes printed to facsimile-match an oil painting. While some copper- or steel-plate engravings were used, plus many woodblocks, 95% of all color printing up to 1900 was done from hand-drawn images on stone. On a litho stone, an artist would draw (crayon) or paint (tusche) a separate printing image for every important color in the original. The opaque mineral ink pigments used were similar to those used in the original oil painting.

Primary colors. In modern color printing, only three printing ink colors are needed. This came about by bringing together the invention of photography, the halftone screen, coal tar, and synthetic dyes to make transparent color inks and color separation filters. All of these technologies were coupled with a better understanding of color vision.

The human retina has photoreceptors called cones that are separately stimulated by long red wavelengths, medium green wavelengths, and short blue wavelengths of the light spectrum. The brain interprets thousands of hues, tints, shades, and intermediates from different proportions and amounts of the three-color signals sent through the optic nerves. When the three printing ink colors are controlled in the same proportion and amounts as in the original, the brain accepts these as matching. Even with imperfect photographic plates, James Maxwell in the 1860s demonstrated this principle by using three lantern slide projectors whose lenses were separately covered with red, green, and blue filters. He simulated the color appearance of objects he had photographed through red, green, and blue filters. It took almost 40 more years to improve the color sensitivity of photographic plates, invent the halftone screen, and produce inks transparent enough to make this a practical reality in printing.

Practically all commercial color printing also uses a fourth color, black. Technically, black is not a color but the absence of color. In its use, black creates no new hues beyond those in the basic three-color printing but adds depth, contrast, and stability to the colors of the image. When quality transparent inks are printed on low-absorbency, high-gloss, coated papers, the three colors overprint to an intense black, making a separate printing of black ink unnecessary. However, black ink is printed for type, as well as to give more intense blacks and contrast when lower-quality inks and papers are used. *See* COLOR.

Color separation. Originally, all input copy for color printing consisted of large oil paintings, watercolors, pastels, or dye-transfer or photographic color prints. These were reflection copy, and huge process cameras with four arc lamps were usually used to reduce the lens image size. However, most modern input copy consists of small color transparencies. Smaller enlargers are used instead of huge reducing process cameras. Larger color transparencies, groups, and page layouts are also same-size color-separated by contact in vacuum frames. Even this revolution of downsizing color separation input and equipment is being superseded by the electronic age of scanners, in which no full-size picture image is color-separated at once. Copy input of color transparencies, photographic color prints, or reflective hand-drawn art is mounted on a rotating transparent cylinder and color-separated spot by spot in lines only a few thousandths of an inch wide and in only a few millionths of a second per spot. Faster flatbed charge-coupled scanners can make four-color separations in a minute or less. *See* CAMERA; CHARGE-COUPLED DEVICES.

The first step in any color reproduction system is known as color separation. The photographic image is exposed separately through red, green, and blue filters. Each filtered separation image must also be exposed through a halftone screen to create different-size dots that control the hues and tone values. When the filtered color separation and halftone screen exposures are made simultaneously to the same photographic film, it is known as the direct screen process. For many years, the indirect process was most commonly used. In this, the filtered color-separation negatives are normal continuous-tone photographic images, and the halftone dot screening is introduced in a second photographic step with a high-contrast film to create positive dot images. These positive dots could be more conveniently color- and tone-corrected in size by hand dot-etching steps. Improved and better-controlled photographic masking steps minimized the need for hand dot etching so that the direct screen method became more popular.

Primary color inks. The actual final press plate images can be exposed through either halftone negatives or halftone positives, but the printing-plate image for the press will be positive. Ideally, this image should be transferred to the printing paper with a process ink that is complementary to the color of the filter that made each color separation image. Yellow ink is used to print from the blue-filtered image because yellow ink absorbs the blue light; magenta ink (not red) is used to print the green-filtered image; and cyan ink (not blue) is used to print from the red-filtered image. This is the first step in the subtractive process of color printing. Each ink, printed as 100% color, acts like a transparent color filter and subtracts one of the additive primary colors of light from the spectrum of the white light reflecting from the unprinted white paper surface. Each of the three colored process inks used in printing absorbs approximately one-third of the white light spectrum. However, being imperfect, they each also absorb more than their proper third of the spectrum. This prevents the reproduction of some of the brightest reds, oranges, greens, purples, and blues. The color from printing that the eye sees is what is left after the ink's subtraction; the secondary and intermediate colors mostly required in reproduction are always less pure and bright than those available from the white paper.

Good yellow inks are the most efficient (about 95%) of the three colors. They usually reflect 100% of the red light, but are typically warmer in hue than they should be theoretically because they also absorb some green light. Cyan ink is the major handicap to reproducing bright pure greens since its hue is bluer than correct because it absorbs considerable amounts of green light. Magenta ink is typically the most inefficient because it absorbs substantial amounts of the blue light that it should reflect 100%. Printers usually refer to this primary ink as red or process red because most commercial magentas have a hue error of more than 50%. The theoretically correct magenta ink would have

a more purple hue, reflecting red and blue equally. Therefore, commercial magentas make acceptable reds, but their blues and purples are less pure and bright.

The three primary colors seldom print alone in most pictures, and their performance in two- and three-color mixtures becomes pictorially more important to the reproduction. If a four-color-ink printed picture is examined with a 10-power or stronger magnifying glass, no more than eight actual colors will be found: yellow, magenta, and cyan appear as separate dots; two-dot overlaps make red, green, and blue; three-dot overlaps make black; and then there are some areas of unprinted white paper between dots. The illusion of hundreds of different colors seen visually without the magnifying glass is the additive color mixture averaging of these eight different areas.

Neutral gray. Perfect process ink would require equal halftone dot sizes in the three colors to reproduce neutral gray. Because the yellow, magenta, and cyan hues of process inks are not equally pure, an important production control is to create a specific inequality of the color dot sizes in a gray scale. This three-color inequality will vary with the use of different sets of process inks, but the cyan dot must always be larger than the other two.

Variables. The theory of color reproduction by printing inks is well understood, but in practical and economic terms process inks and paper are less than the optimum quality. And while the separate inks individually may transfer well to the paper, they do not always all print fully on each other (a problem known as trapping). Another press variable may be that different pressures, ink consistencies, and paper absorptiveness result in larger or different dot sizes. Thus, not all printed color reproduction is automatically of optimum production quality.

Proofing

For most high-quality printers, a key part of the workflow has been the contract proof. The proof is a document presented to a client for approval prior to printing. Printers wishing to provide high value to their customers now have an expanding range of proofing options.

Contract proofs have changed dramatically. Yet as the developments of laser scanning, desktop publishing, and computer-to-plate systems have improved the quality of the printing process, trends in contract proofing is often driven more by cost than by quality.

Virtual proofs. Virtual proofs, sometimes called soft proofs, are images on a monitor, as opposed to hard proofs on paper, film, or some other material. A virtual proof is used to check the pages ready for transmission to a printing plant as well as to detect errors (in the RIP, lack of fonts, colors). In virtual proofing, the cyan, magenta, yellow, and black bitmaps are converted into red, green, and blue on the monitor. Virtual proofing does not eliminate the need for a hard-copy proof, but it does allow the correction cycle to occur faster and reduces the number of hard proofs that are needed.

Virtual proofing systems are increasingly used for comparisons at the press. However, the process demands a lot from the monitor and the lighting environment. All parts of the system have to be thoroughly calibrated. Virtual proofing provides press operators with a general idea of what the printed page should look like. One problem is that virtual proofing is not user-friendly for customers or production operators for making changes or corrections, and for transporting proofs between operators and to customers. There is also the issue of two or more monitors on which the proof is viewed being calibrated differently and, hence, displaying images that appear different in color. However, advances in color management tools, improved electronic file transfer methods, and upgraded computer monitor technology are prompting printers to pursue virtual proofing methods.

Calibrating the monitor requires determining its color gamut and adjusting it to match what the printing press will produce. This must be consistent among all parties viewing the proof. For example, when proofs are transmitted to a magazine's art director, advertiser, and printer, all parties must simultaneously review the page files and indicate any corrections to the soft proofs with accuracy in minutes.

Virtual-proofing methods can range from PDF files via e-mail or FTP sites, to systems that allow users to review content and layout, to more complex color-managed systems that seek to emulate hard-copy proofs. While users may adopt soft proofing as a workflow tool to speed production, whether they embrace it as a replacement for the hard-copy contract proof depends on the comfort level of printers and print buyers.

While a large amount of current online soft proofing simply involves posting PDF and JPEG files on password-protected areas of a Web site, many e-commerce vendors and even some makers of prepress-workflow systems have added online virtual proofing capabilities to their products. Online virtual proofing is especially valuable for projects with tight deadlines.

However, the process is not for all situations and particularly not for those involving color-critical projects because calibrating displays can be difficult. The process can be valuable for applications that warrant it, such as repeat jobs, projects involving Pantone or spot colors, or clients who trust the production vendor to produce accurate color.

Remote hard-copy and remote proofing. Traditional analog and digital halftone proofing systems have a high acquisition cost, making them inaccessible to many smaller printers and impractical for low-volume remote-proofing applications. Lower-cost inkjet proofing systems can deliver prints with six colors at high print resolution. These printers are powerful proofing tools, capable of output that simulates traditional analog proofs. They have color profiling capabilities and the ability to print on special media.

Remote proofing involves sending files to clients (or vice versa) for output on a proofing device, or soft proofing via a computer monitor. Either or both can be integrated into a printer's workflow depending on client needs and expectations.

The availability of lower-cost alternatives, which simulate the output of analog and digital proofing systems, lowers the price of proofing and places hard-copy color proofing at more points in the workflow. To ensure consistent print quality between systems, some manufacturers have integrated densitometers that automatically make color measurements and adjustments to keep color values within specifications.

Electrophotographic color copiers and printers are also being used as proofing devices. They may be more expensive to acquire than inkjet proofing systems. However, they usually have a lower color-page cost. The lower page cost and a larger paper capacity make electrophotographic units more suitable than inkjet for printing short-run color work. Hence, a color copier can be a dedicated proofing system and handle other graphic and office printing.

A challenge to printers is getting their customers to adopt remote proofing. The long-standing, stable accounts that generate high dollar volume are more receptive to remote proofing than newer customers or customers with infrequent print work. Wide acceptance of remote hard-copy proofing also depends on the development of inexpensive systems that are easier to implement and maintain.

Not all remote proofing workflows rely on the transmission of job files to a client for output on a remote proofing system. Some printers install color-calibrated digital proofing systems at customer locations and let them produce their own proofs, which are submitted along with the job files for print production.

Also aiding remote proofing are inkjets with improved color gamut, light-fastness, and stability. In addition, the cost of consumables on these inkjet printers is lower than that for high-end digital proofing systems.

Remote proofing in general has presented challenges for printers with critical-color clients, but emerging new technologies and products are addressing this issue. For color-managed remote proofing to succeed, there has to be a relationship between the designer, the print buyer, and the printer so that all involved agree on paper and standard ink sets, and all of the components that affect color. One of the most important considerations is that the lighting under which the proof is viewed must be controlled.

Image Assembly and Imposition

Rarely used any longer, the image assembly, or stripping, operation is a skilled manual procedure for constructing and assembling negative or positive films in such a precise manner as to provide the platemaker with complete film assemblies for producing printing press plates. Image film assembly was a simple manual procedure involving the positioning of camera films on a predrawn support flat (**Fig. 4**). The most important skill was the planning and drawing of the layout arrangement, known as an imposition. Computerized systems are now used for the layout and assembly of images.

Machines have been developed that automatically control image placement, practically eliminating hand assembly methods. In these step-and-repeat machines, computer programs control films and their positions on the plate, and optical imposition cameras that photograph pages onto a film mounted on a movable back whose movement is in the x, y plane. Advances in the technology have allowed plates to be exposed digitally in platesetters by completely imposed computer programs.

Platemaking

Platemaking is the final prepress process. Without good plates, even a sophisticated printing system cannot produce optimum results. Therefore, the methods by which plates are made are crucial to the outcome of the printing process.

Platemaking processes. Printing plates or cylinders are made in one or a combination of seven different ways: (1) manually by using hand tools, engravers, knives, and so forth, to produce plates for the relief, intaglio, and stencil processes; greasy crayon or tusche can be used for making hand-drawn lithographic plates; (2) mechanically by engraving and geometric lathes, ruling machines, pantographs, and so on, for the relief and intaglio processes; (3) electrochemically by electrodeposition of metals to produce longer-wearing images; (4) electromechanically by electronic engravers to make relief and intaglio image carriers; (5) electrostatically by using xerographic principles to make lithographic plates; (6) photomechanically by combining the photographic process to produce images and a mechanical process to make plates; (7) digitally by laser exposure or other radiant energy from direct digital information in storage or prepress systems. The photomechanical method is the most widely used, while digital platemaking is increasing and will dominate in the future. *See* DIGITAL CONTROL.

Each of the printing processes has different platemaking or image-forming characteristics and requirements. Letterpress image carriers (plates) consist of engravings of single subjects locked in position with other engravings and cast-metal type into an imposition form for printing on a flatbed or platen press. Plates for rotary presses are duplicate plates made from curved molds of the original imposition forms. These can be of cast type metal called stereotypes, electroplated metal called electrotypes, and cast plastic or rubber plates. Rotary letterpresses have been developed that use photopolymer-coated wraparound metal plates exclusively. Flexographic plates can be cast rubber plates made from engraving forms, but flexography uses mainly photopolymer-coated wraparound metal plates. Lithographic plates are wraparound metal (usually aluminum) plates with light-sensitive

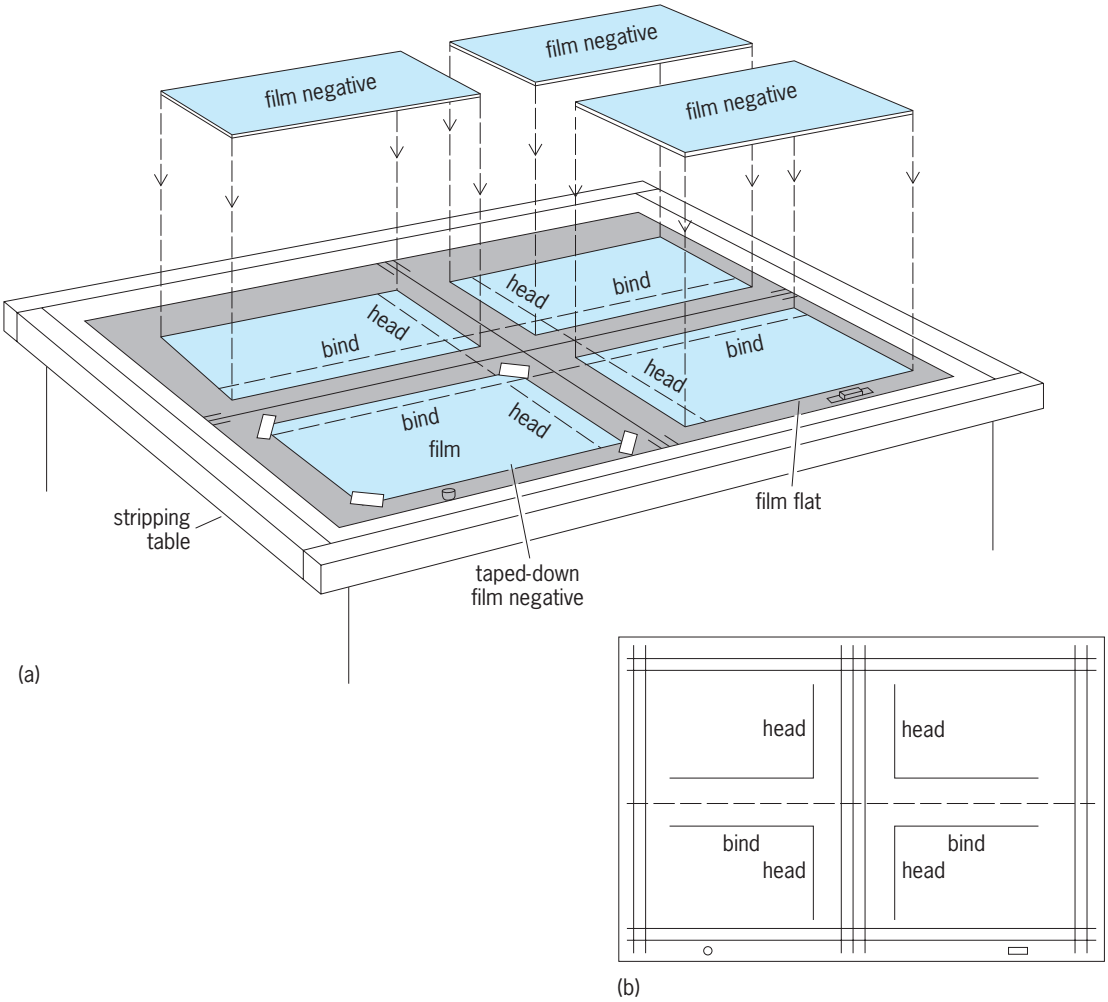

Fig. 4. Image film assembly. (*a*) Assembly setup. (*b*) Page position layout.

coatings. Most gravure uses copper-plated steel cylinders on which intaglio cells are etched by electromechanical engraving machines. Screen-printing stencils are made on polyester, nylon, or stainless steel fine-mesh screens, using band-cut stencils, transfer films, or light-sensitive coatings. Electrolytically produced rotary screens with perforated images and internal ink squeegees are also used for screen-printing.

Most of these processes use photomechanical methods that consist of photographic films exposed on light-sensitive coatings either directly coated on the printing surface or transferred to it. Since most photomechanical coatings are of slow speed, high-energy pulsed xenon or metal halide light sources are used for exposures, which are usually controlled by light-integrating devices. Step-and-repeat machines are used for exposing large plates for lithography, flexography, and wraparound letterpress where multiple exposures, as for label printing, are required. Automatic processing machines are used for making lithographic plates. The important feature of light-sensitive coatings is that on exposure to light, physical characteristics, such as solubility in water or other chemicals, change so that the coatings can be developed to produce images that serve as the printing images or as resists for producing the printing images. The present trend is toward computer-to-plate systems, whereby plates are laser-imaged directly from a computer. Some of these systems use thermal exposing as opposed to light exposure.

Photoengraving. This is the photomechanical process used to produce relief plates for letterpress, letterset, and flexographic printing. There are two general types of engravings: line and halftone. Zinc, magnesium, and copper are the metals commonly used for photoengraving. In the United States, zinc and magnesium have been used almost exclusively for line engravings, and copper for halftones and fine-line engravings.

Photopolymer plates. Polymeric materials are used to produce relief images by chemical means. Several types of photopolymers are used, such as poly(methacrylates), nylon, polyamide, urethanes, and photosensitive synthetic rubbers. The preferred combination has been nylon-type photopolymers on steel plates used on the press with magnetic plate cylinders and ultraviolet curable ink. *See* POLYMER.

Letterpress declined in use because hard relief plates (metal and most photopolymers) require very precise preparations and adjustments (makeready) prior to printing, especially to compensate for the variable pressures exerted by the different-size image elements. To transfer ink, more pressure per unit area is exerted on small highlight dots than on large shadow dots or solid areas. To achieve equivalent ink transfer, considerable pressure adjustment is required.

Flexographic platemaking. Flexography uses flexible or resilient relief plates that alleviate the differential pressure/makeready problems of the inflexible relief plates used in letterpress. As a result, flexography has become the dominant relief printing process and has replaced letterpress in many newspaper, label, packaging, and corrugated-box printing applications. It uses rubber duplicate plates made from relief engravings for printing corrugated boxes and some packaging, and special resilient photopolymer relief plates for newspapers, pressure-sensitive and other labels, and milk cartons and other packaging materials. Seamless rubber-coated cylinders for printing gift-wrapping papers and wallpapers are made digitally by laser-engraving the rubber.

Gravure platemaking and cylinder making. Gravure is printed principally on web presses from etched cylinders of copper-plated steel and is generally called rotogravure. A small amount of sheet-fed gravure is printed mostly from thin copper plates wrapped around the plate cylinders of special single-color and multicolor presses. Sheet-fed gravure is capable of very high quality pictorial reproduction and is used for proofing and short runs of art and photographic reproductions, as well as annual reports.

Gravure differs from other printing processes in that it can print varying amounts and densities of ink to simulate continuous-tone imaging, as in photographic prints. The intaglio cells, cut or engraved and etched to various sizes or depths below the surface of a plate or cylinder, vary in volume and transfer different amounts of ink corresponding to the tonal values of the image areas being printed. Two processes for producing the cells represent two different types of cells. Conventional gravure, which is used for making most of the plates for sheet-fed gravure, produces cells all with the same areas but variable depths. Electromechanical engraving, the process used for almost all rotogravure printing, produces cells with variable areas and variable depths.

Conventional gravure. In this method, a sensitized carbon tissue (paper coated on one side with a pigmented gelatin layer sensitized by bichromates) or transfer film is contact-printed through a continuous-tone positive film and given a second exposure by contact with a gravure screen consisting of a grid of 150 or 175 transparent lines and opaque square dots to the inch with a ratio of line to dot width of 1:3. The exposed tissue or film is moistened and squeegeed into contact with the clean copper surface of the plate or cylinder. The image is developed in water until a gelatin relief resist is produced. The image is etched with ferric chloride solutions, which etch the copper to different depths corresponding to the thickness of the gelatin resist in the different tone areas. The fully exposed areas such as the clear lines in the screen remain unetched and provide lands or bridges of common height to support the doctor blade during printing. The image areas consist of cells, all with the same area and variable depths depending on the tone values. This process is used mainly for high-quality single-color and process-color printing. It suffers from short life on the press, as the shallow highlight dots wear prematurely, causing undesirable changes in tone reproduction.

Electromechanically engraved cylinders. Electromechanical engraving systems consist of three parts (**Fig. 5**): a scanning unit with a rotating drum on which photographic prints are mounted in position for printing and scanned by one or more reading heads; a computer that processes the image density data recorded by the scanning heads and converts the data into electrical impulses; and an engraving unit that uses the electrical impulses to drive diamond styli on one or more heads that engrave gravure cells in the rotating copper-plated steel printing cylinder. The depth of the cells varies with the strength of the electrical impulses, which correspond to the densities of the same areas on the originals. The electromechanical engraving system produces 4000 cells per second. The cells vary in area, depth, and volume, corresponding to the tone values they represent (**Fig. 6**). After engraving and before mounting on the press, the cylinder is polished to remove all burrs and chrome-plated to resist wear from the doctor blade on long runs.

Halftone gravure. The use of continuous-tone images for electromechanical engraving required high skills for retouching and correcting the films, prints, and engraved cylinders, and the production of expensive press proofs. The shift to the use of the same halftone images as produced for lithography has resulted in many improvements in cylinder making. Color separation and corrections have been simplified, and off-press and offset press color proofs can be used. Reading heads are used to defocus the halftone images on electromechanical engraving machines, and the hues of gravure color inks for magazine printing have been revised to match the inks formulated to the *Specifications for Web Offset*

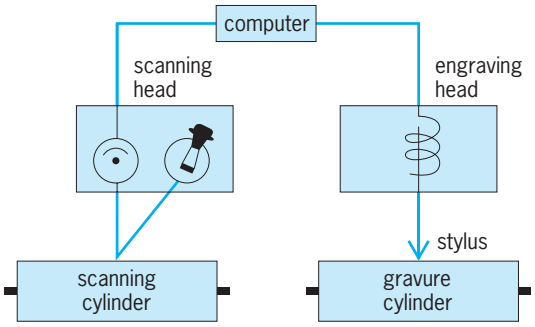

Fig. 5. Electromechanical engraving system.

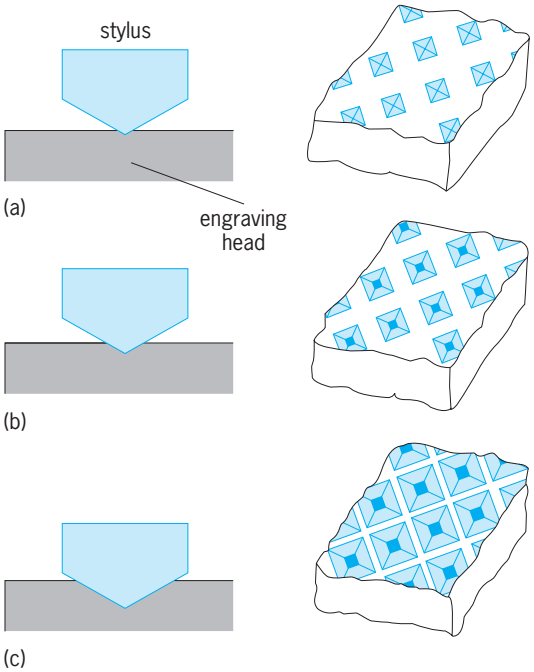

Fig. 6. Electromechanical engraving. Depth of cells and shapes and sizes as engraved by the stylus corresponding to tone values: (*a*) highlights, (*b*) middle tones, and (*c*) shadows.

Publications, which in turn match most off-press and offset press color proofs. There are also cost savings, since the same films can be used to produce the proofs and images for printing by lithography or gravure.

Filmless gravure. Gravure was the first process to use computer-to-plate (cylinder) digital technology to eliminate the use of films for making the printing cylinder. Interfaces have been developed between prepress color systems and electromechanical engraving systems, so that the digital data are fed directly to the electromechanical engraving computer to produce the electrical impulses to drive the engraving styli, thus making the scanning cylinder, prints, and reading heads unnecessary. The use of filmless gravure also eliminates the need for halftone gravure.

Photopolymer gravure. Photopolymer plates are also available for gravure. They are similar to letterpress photopolymer plates except that the image or printing areas are intaglio (recessed) instead of relief. Hard polymers, such as nylon, are used to resist wear from the doctor blade, which scrapes the excess ink off the surface of the plate. Such plates make gravure competitive with lithography and flexography for run lengths under 100,000, especially in packaging.

Lithographic platemaking. The main metal used for platemaking in lithography is aluminum with a mechanically, chemically, or electrolytically roughened surface. The image areas of lithographic plates are ink-receptive (nonwettable by water), and nonimage areas are water-receptive (nonwettable by ink). The wider the difference in ink and water receptivity of the two areas, the easier it is to control printing and the longer the plate will run on the press. The chemistry of the production of the images is related to the changes that occur in the photosensitive coatings on exposure to light and the development process.

Chemistry is also involved in making the nonimage areas water-receptive and desensitized to ink. One way to improve water wettability of the nonprinting areas is by roughening the plate surface. Another method employs surface treatments such as anodizing or treating with silicates or phosphates. Wettability is also improved and maintained by the use of hydrophilic (water-receptive) materials such as gum arabic or synthetic gums, which have good adhesion to metal surfaces. Gums are used in combination with phosphoric acid and salts such as diammonium monohydrogen phosphate and zinc or magnesium nitrate. Known as a densitizing etch, the mixture improves the water wettability of the metal surface to which it is applied.

Surface plates. The light-exposed coating on these plates becomes the image that accepts ink on the press. There are two types of surface plates: additive and subtractive plates. For additive plates, an ink-receptive lacquer or ink is applied to the image during development. Subtractive plates are precoated with an ink-receptive lacquer that is removed from the nonimage areas during development. Until the introduction of presensitized plates in 1950 and "wipe-on" plates in 1957, the most popular surface plates were made with bichromated albumin or casein. These have become obsolete because of pollution problems and difficulties in printing, with image areas blinding (refusing to take ink) and nonimage areas scumming (wetting with ink) because of poor densensitization.

Additive plates are easy to make but are not as consistent and durable as subtractive plates. Diazo presensitized and wipe-on plates are simple to make. They print well because their surface treatments prior to coating are very water receptive, allowing printing with a minimum of water and reducing the tendencies for the images to blind and the nonimage areas to scum. Presensitized plates are precoated by the manufacturer; wipe-on plates are coated in-plant with diazo resin, usually by applying the coating with a simple roller. After exposure with a negative film, presensitized and wipe-on plates are handled in the same way. An emulsion developer, consisting of a pigmented lacquer, acidified gum solution, and a suitable emulsifier, is rubbed over the whole plate. The gum dissolves the coating in the unexposed (nonimage) areas and makes them water-receptive, while the lacquer deposits on the exposed (image) areas, making them visible and ink-receptive. The plate is washed and gummed, and it is ready for use.

Additive plates are subject to abrasive wear on the press, so they are not suitable for runs much over 50,000 on sheet-fed presses or about 100,000 on newspaper and other web presses. The prelacquered subtractive presensitized plates are capable of runs almost double this.

Subtractive plates are coated by the manufacturer with a light-sensitive layer. After exposure, development removes the coating from the nonprinting areas, leaving the exposed coating in the image areas. Subtractive plates are generally more consistent and capable of longer press runs.

Positive plates are made from coatings that use naphthoquinone diazides. After exposure, alkaline developing solutions remove the exposed areas and leave the unexposed ones as the printing areas. These image areas are inked to protect them from light degradation, and the plate is gummed to protect the nonimage areas.

Photopolymer plates are the longest-running of lithographic surface plates. Some use polymers that can be developed in aqueous solutions, thus eliminating problems of water pollution. They are all subtractive plates for which processing is very simple. All are organic compounds that are very inert and abrasion-resistant. Light exposure changes the molecular size and molecular weight of some, and cross-links others to produce very durable images that are capable of runs in excess of 250,000 impressions. An important feature of some photopolymer plates is that run life on the press can be further increased by baking the plates in an oven at 450–550°F (220–270°C) for about 4–5 min. This treatment increases life on web presses to over a million impressions.

Bimetal plates. Unlike surface plates, the coating on bimetal plates serves as a stencil for etching the unexposed areas. The plate consists of copper plated on aluminum or stainless steel, on top of which a photosensitive coating is applied. After the coating is exposed and developed, the plate is etched with a ferric chloride solution to remove copper in the unexposed areas. The final printing plate consists of copper in the image areas and aluminum or stainless steel in the nonimage areas. These very rugged plates can take considerable abuse on the press from excessive pressure and papers with gritty surfaces, and are capable of runs of 1–5 million impressions.

Bimetal plates became popular because enlargement in the size of the halftone dot, or dot gain, can be controlled by etching rather than exposure. Photopolymer plates made from positives are preferred because the images have less dot gain than those on plates made from negatives. However, because the dot gain correction is erratic, many web-offset publication printers turned to bimetal plates, where the dot gain correction is proportional and controllable. Disposal of etching wastes is a problem, however.

Waterless plates. These plates differ from surface plates in that they use silicone rubber in the nonprinting areas and the image is printed from ink on the metal base. Silicone rubber has very low surface energy, preventing it from being wet easily with ink. Therefore, plates with silicone rubber in the nonimage areas can be printed without dampening solution, or water, eliminating the need for an ink-water balance that is the source of many printing problems, including long setup time and excessive paper waste. The process is sensitive to increases in temperature, which can cause printing in the nonimage area, called toning. Presses printing waterless plates need special inks and a temperature-control inking system. With these controls, waterless plates have shorter setup time, higher press speeds, and exceptional print quality in screen rulings up to 500 lpi. *See* SILICONE RESINS.

Digital plates. These plates are used in computer-to-plate digital printing systems. They are made directly from digital data from high-end color electronic prepress systems or imagesetters interfaced with color desktop publishing systems, and they are used for printing on conventional lithographic presses. There are three types: high-speed light-sensitive plates, ablation plates, and heat-sensitive polymer plates.

High-speed light-sensitive plates. These plates are exposed by lasers in imagesetters. There are four types: silver halide, electrophotographic, high-speed dye-sensitized photopolymer, and hybrid.

Silver halide plates on polyester bases are used for short-run single-color and spot-color printing. They consist of silver halide emulsions processed in special developers. Silver halide plates on metal bases for process color printing were introduced in 1990.

For newspapers that receive satellite transmissions, electrophotographic plates are coated with organic photoconductors, charged by corona discharge, and imaged by digital-data-driven visible-light lasers and liquid toners.

High-speed dye-sensitized photopolymers are used to produce long-run lithographic plates. The photopolymers used for conventional plates are mainly sensitive to ultraviolet radiation. As is done with photographic emulsions to produce orthochromatic, panchromatic, and infrared-sensitive films, special dyes are added to the photopolymers to extend their light sensitivity into the visible spectrum, thus increasing their exposure speed. A number of such plates are available for use. Most have light sensitivities for exposure by a medium-power 30–50-milliwatt (blue) argon ion (488-nanometer) laser.

Hybrid plates use two photosensitive coatings on a metal plate. The top coating is a silver halide emulsion whose light sensitivity can be varied to handle a full range of speeds, and a full gamut of spectral responses from ultraviolet to blue, green, red, and infrared lasers. The bottom coating is a photopolymer with good performance characteristics on the press. The top coating provides a wide range of light sensitivities to accomplish imaging for contact, projection, camera, or digital computer-to-plate imaging, with controlled dot gain and image contrast. The bottom coating provides ease of printing characteristics that result in high productivity and consistent quality such as plate life, durability, ease of makeready, and maintenance of ink-water balance, ink transfer, register, and other running controls on the press. *See* LASER; ULTRAVIOLET RADIATION.

The top layer (silver halide) is exposed by a laser in the platesetter, and then the plate is developed to form an opaque image. Next, the plate's photopolymer coating is exposed to ultraviolet light through the clear (top) image areas and then developed. Plate life can range from 150,000 impressions to over 1 million if the coating is baked after processing.

Ablation plates. These are lithographic printing plates with slightly depressed (ablated) wells in the image areas. They were introduced as waterless plates with silicone rubber in the nonimage areas, and the printing wells were produced by spark discharge, driven by digital data. Because ablation is a dry process (no liquid development), four-color plates can be imaged simultaneously directly on the press, with an additional feature that the pixel distribution for each color is counted as the plates are being exposed and the data are used to automatically set the ink fountains. The spark discharge system has been replaced by laser diode technology with considerable improvement in printed image quality (**Fig. 7**). Laser diode technology can also be used to make plates that can run in a conventional lithographic ink-water balance mode.

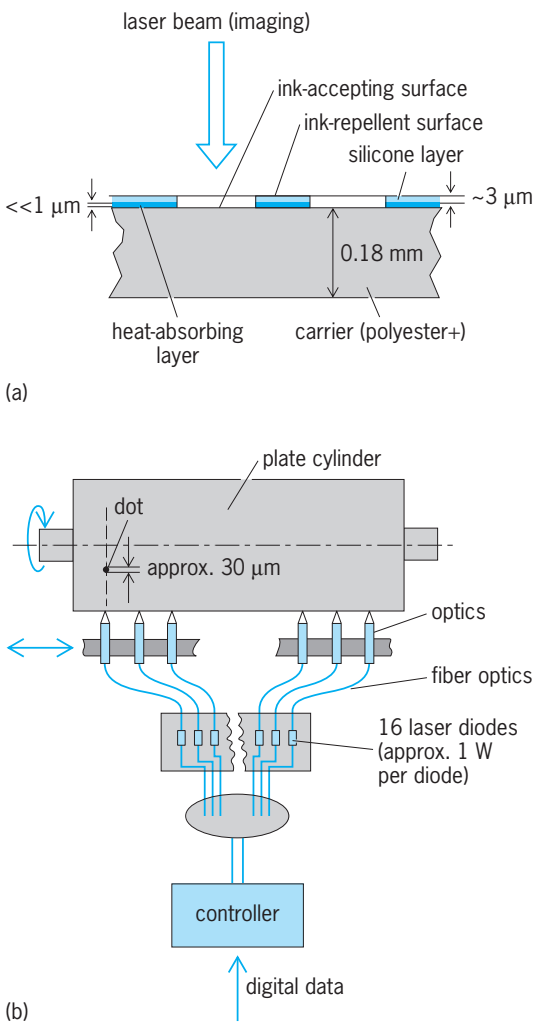

Fig. 7. On-press plate imaging for waterless offset printing with thermal laser ablation technology. (*a*) Printing plate (thermal plate) structure for waterless offset and thermal ablation process. (*b*) Principle of the imaging head as used for plate imaging in the system.

Heat-sensitive polymer plates. A thermally sensitive polymer responds to heat but not light, so it can be handled in ordinary room light. It is imaged in a proprietary platesetter, fitted with a laser diode (heat source). The plate is processed with an aqueous solution. Run lengths of over 1 million impressions, without baking, have been achieved, as well as reproductions of fine image details and resolutions of 4.8 μm, the equivalent of a 1% dot on a 600-lpi halftone screen.

Printing Processes

Two categories of printing presses are sheet-fed presses and web presses. Sheet-fed presses print on individual sheets of paper, as opposed to continuous rolls of paper used on web presses. Some types of sheet-fed and web presses can print on both sides of the paper at the same time; this is called perfecting. Sheet-fed presses range in size from small presses that print on letter-size paper to presses that handle large sheets of paper. In addition, sheet-fed presses can print on cardstock, plastics, metal, and other substrates. A sheet-fed press may consist of multiple print units, with each printing a different color.

Web presses print on continuous rolls of paper or other substrates. Web presses are used for high-volume printing of magazines and newspapers. They typically are faster than sheet-fed presses. Web presses can consist of several connected units for printing different colors and providing finishing processes such as cutting, folding, and punching.

High-speed commercial web presses use wide rolls of paper for printing magazines and books, and use heat to set the ink (heat-set web). Cold-set web presses are used for printing on newsprint or other paper on which ink is dried by absorption and oxidation.

Offset lithography. Offset lithography is the most popular and widely used process and represents approximately 45% of all printing. In offset lithography, an image is transferred to the paper using inks, specially treated plates, and an offset cylinder. The image and nonimage areas of the printing plate are on the same plane, as opposed to plates with raised or etched type or images.

Lithographic plates' image areas are ink-receptive and water-repellent, and the nonimage areas are water-receptive and ink-repellent. The term "offset" refers to the image being not printed directly on the paper but transferred (offset) from the plate to another surface that makes contact with the paper. An important function of the lithographic printing press is to provide a means for transferring ink and water to the plate and maintaining the proper ink-water balance during printing.

In offset lithography, an image is produced on a thin aluminum plate, although other metals, plastics, and paper can be used. Rollers apply ink (usually oil-based) and fountain solution (mostly water) to the plate. Since oil and water do not mix, the oil-based

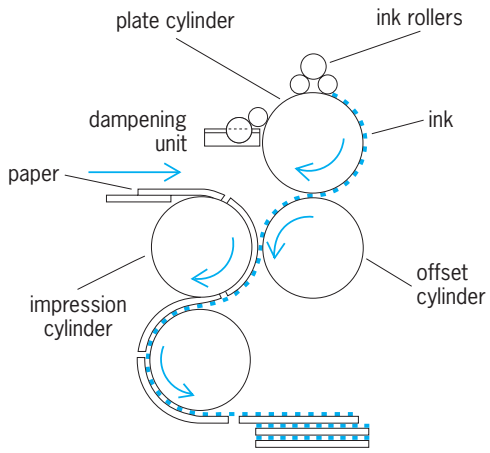

Fig. 8. Lithographic printing process.

ink adheres only to the image areas. The inked image is transferred to a rubber blanket mounted on a cylinder and then transferred to the paper as it passes between the blanket and an impression cylinder beneath the paper (**Fig. 8**).

Lithographic printing press. Lithographic presses are rotary presses; that is, the plate is mounted on a cylinder that rotates during printing. Because the plate is comparatively thin, from 0.005 in. (0.125 mm) for a small press to 0.02 in. (0.5 mm) for a large press, uniformity is very consistent, which reduces makeready time.

For sheet-fed presses, the plate cylinder is provided with the means to rotate it independent of its gear drive, making possible the desired front-to-back positioning of the image on the sheet (**Fig. 9**). Systems have been developed that accurately position the image on the plate to minimize or eliminate the need to change the position of the plate cylinder. Some plants install pins in the plate cylinder that correspond to the pin system used to coordinate image positioning in platemaking operations. Newer electronic printing presses automatically position the plates.

Dampening system. When the lithographic press is running, the plate cylinder carries the plate under two different roller systems. The first system that contacts the plate is the dampening system. In a conventional dampening system, the dampener is separate from the inking system and uses cloth-covered form rollers and a ductor roller that alternately contacts the fountain (dampener reservoir) roller and the dampening oscillator roller. Some dampening systems feed the solution to the first ink form roller, which transfers it to the plate. Others spray the solution on the plate. The surface of the plate is slightly dampened, but the image areas, being hydrophobic (water-repellent), resist wetting. The dampening (fountain) solution is about 98% water, with the other 2% made up of a variety of chemicals, which in most systems make the solution acidic. Alkaline dampening solutions are also used, primarily in lithographic newspaper printing. Automatic devices are used for mixing, supplying, and replenishing fountain solutions, as well as measuring their pH and conductivity and controlling their temperature.

Inking system. After dampening, the plate passes under the ink form rollers. The nonimage areas of the plate do not accept the ink if properly dampened, but the image areas do. Lithographic ink is thixotropic, as well as very viscous. These attributes necessitate a complex inking system, consisting of drum, oscillator, riders, and fountain rollers made of hard materials such as steel or hard rubber (**Fig. 10**). With the exception of rider and fountain rollers, the steel rollers are usually covered with copper, ebonite, or nylon to resist the tendency of acids in the dampening solution to desensitize them. Another reason for this complex inking is the fact that some of the

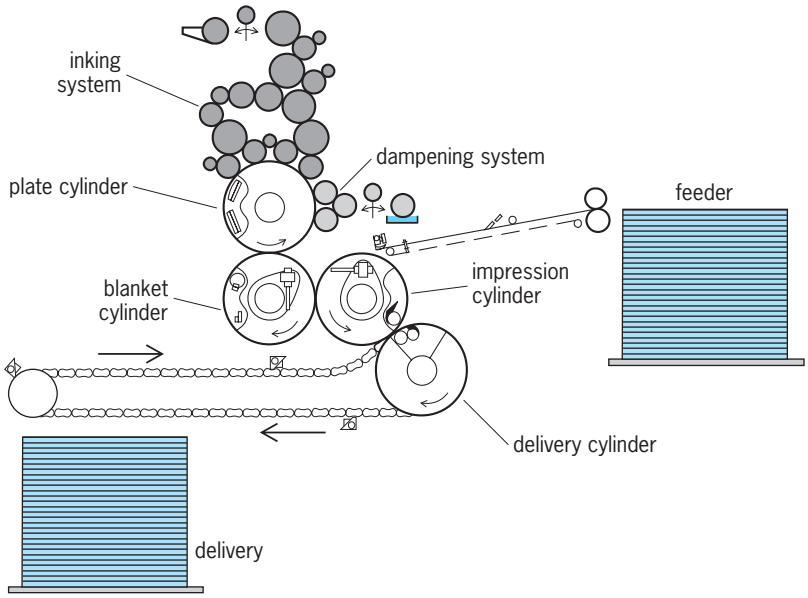

Fig. 9. Single-color sheetfed press. (*After L. DeJidas and T. Destree, Sheetfed Offset Press Operating, 2d ed., Graphic Arts Technical Foundation, 1994*)

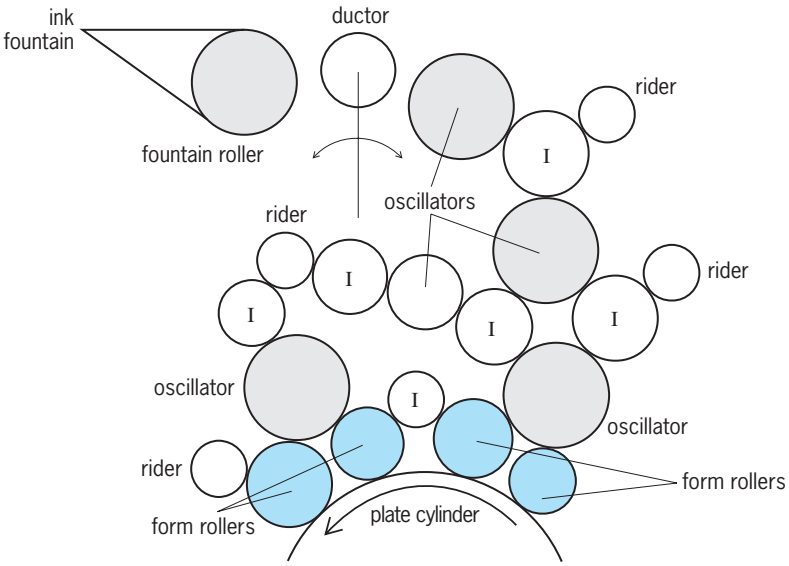

Fig. 10. Typical lithographic press inking system. I = intermediate roller. (*After L. DeJidas and T. Destree, Sheetfed Offset Press Operating, 2d ed., Graphic Arts Technical Foundation, 1994*)

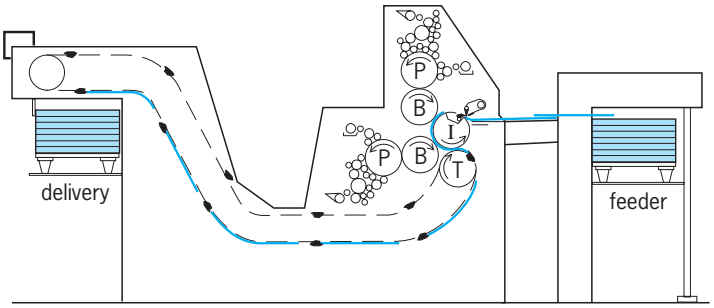

Fig. 11. Two-color press that has a common impression cylinder. P = plate cylinder. B = blanket cylinder. I = impression cylinder. T = transfer cylinder. (*After L. DeJidas and T. Destree, Sheetfed Offset Press Operating, 2d ed., Graphic Arts Technical Foundation, 1994*)

dampening solution is picked up by the ink, and the higher the surface area of the rollers, or the more rollers, the faster the dampening solution is evaporated from the inking system to prevent emulsification. Ink fountains have keys for adjusting the flow of ink to the ink train. Setting the ink keys manually is a tedious and time-consuming operation. Electronic plate scanners are available for setting ink fountains automatically. There are also means to calculate the ink demands of digital plates to set the ink fountains automatically.

Waterless systems. Special presses are built for the printing of waterless plates. They have no dampening system, but temperature-controlled inking rollers. Most systems use refrigerated water circulation in the ink distributor or vibrator rollers. Others use special fans to blow cool air over the inking rollers. Inks for waterless printing have a control temperature range of about 10°F (5.6°C) between good ink transfer and toning, or scumming. Since the blanket is dry, waterless plates require the use of higher grades of paper to avoid the accumulation of lint and other paper debris on the blanket during printing.

Blanket cylinder. After inking, the plate contacts the blanket cylinder. The cylinder is equipped with blanket mounting and tightening devices, just as the plate cylinder has similar devices to mount and tighten the plate. A principal difference is that the plate cylinder can be rotated independently of its driving gear, while the blanket cylinder cannot. The position of the plate cylinder in the frame of the press is fixed, but the blanket cylinder of all but some small-sized presses is so mounted that its position with respect to the impression cylinder can be changed. This feature makes it possible to change the clearance between blanket surface and impression-cylinder surface to accommodate printing on substrates with a wide range of thicknesses. For example, sheet-fed presses can be adjusted to print on onionskin or board.

Impression cylinder. After receiving the ink image from the blanket cylinder, the blanket contacts the paper, or substrate, that is carried on a third, impression cylinder. At the lead edge of the cylinder is a device to receive and firmly hold the front edge of the sheet to be printed during the cycle when the blanket contacts the paper. The device is known as the gripper bar, a shaft across the length of the cylinder on which is mounted a series of fingers known as grippers. As the sheet is drawn through the printing nip (the area where blanket and substrate contact each other), the leading edge of the cylinder reaches a position where the impression cylinder grippers release the sheet, while other sets of grippers mounted on traveling chains (one on each side of the press) carry the printed sheet to a delivery pile and release it there.

Lithographic press configurations. The basic lithographic offset press is sheet-fed. It can be fed from a roll by attaching a paper-sheeting device at the feeder. But one of the most important features of the lithographic offset sheet-fed press is the ease with which a multicolor press can be assembled. Presses with up to 10 units are built to meet special needs. If a plate, a blanket, and an impression cylinder in combination are considered as a printing unit, a two-color sheet-fed press comprises two printing units with feeding and delivery components. A six-color sheet-fed press comprises six such units.

One deviation from this approach to multicolor sheet-fed presses is the common impression cylinder design. In sheet-fed presses, two plate cylinders and two blanket cylinders are assembled as couples. Each couple comprises a plate and blanket cylinder. Two such couples, mounted with an impression cylinder between them, results in a two-color common impression cylinder unit (**Fig. 11**). Two such units assembled in tandem, with provisions for sheet transfer from one impression cylinder to the other, become a four-color press. The rotary sheet-fed press has lent itself to the development of presses that can be run as either perfecting (printing both sides of the sheet) or as multicolor, with a very short time required for the changeover of modes.

In-line (tandem) web press. To meet some specialty needs, such as forms and package printing, printing units are assembled in tandem (**Fig. 12**) and fed from

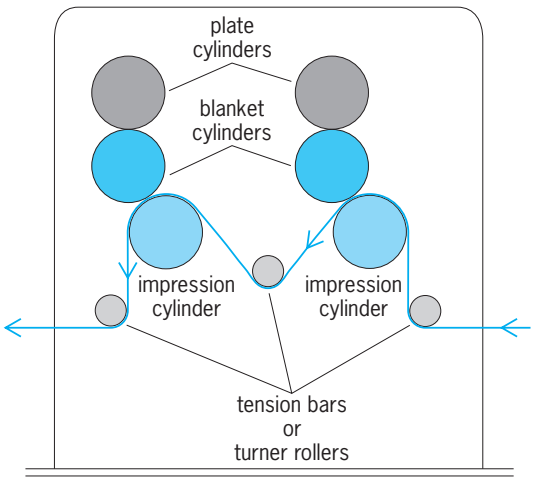

Fig. 12. Unit-type press with printing units assembled in tandem. (*After R. Blair and C. Shapiro, The Lithographers Manual, Graphic Arts Technical Foundation, 1980*)

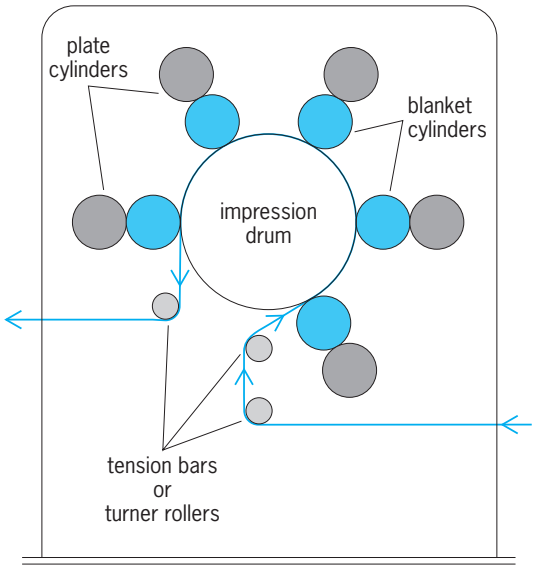

Fig. 13. Common impression drum-type press. (*After R. Blair and C. Shapiro, The Lithographers Manual, Graphic Arts Technical Foundation, 1980*)

a roll of paper (web-fed). Such a press can print lithographic and indirect (offset) letterpress plates on the same run. Presses have been built with three to five printing units in tandem around a common impression cylinder (**Fig. 13**). This configuration has been discontinued for publications presses, but it is used in newspaper presses for process color printing.

Blanket-to-blanket perfecting. The greatest growth in lithography has resulted from a unique capacity of the offset method of printing. That capacity is simultaneous perfecting, or printing, on both sides of a substrate by two opposing blankets (**Fig. 14**). A printing unit in this configuration has two plate/blanket cylinders assembled, such that the blanket cylinders run against each other. Each blanket cylinder transfers the ink for its image and, at the same time, acts as the impression cylinder for the opposing blanket cylinder. These presses are known as blanket-to-blanket presses. The configuration lends itself to a variety of assemblies, depending on the type of product being printed. Up to 12 units have been assembled into a single press. In most commercial printing applications, the web is fed straight through. Many of the presses are equipped so that more than one web may be printed simultaneously. On an eight-unit press, for example, two webs can be run, with each being printed in four colors on both sides. This configuration is used on most magazine and catalog presses.

Ink drying on web lithographic offset presses. After printing, the specially formulated (heat-set) inks are dried in high-temperature ovens. Drying is virtually instantaneous because of the rapid and complete evaporation of the ink solvent. Before the web is passed into a finishing operation (usually folding), the heated web is passed over chilling rollers to set the ink, which is in a molten state when the web leaves the dryer. In newspaper printing by lithography, ink drying results from the absorption of the ink vehicle into the paper.

Relief printing. In relief printing, the image to be transferred to paper (or other substrate) is raised above the surface of the printing plate. Ink is applied to the raised surface and then rolled or stamped onto the substrate (**Fig. 15**). The process is similar to using an inkpad and stamp. Letterpress and flexography are forms of relief printing. Although the image to be printed is raised on the printing plate, it does not necessarily create relief or raised lettering, as in embossing.

Letterpress. Letterpress is printing from a raised surface. It is the oldest of the present printing processes and has been replaced mostly by offset and other printing processes. Letterpress printing is still used for specialized and limited-edition printing, as well as for embossing, die cutting, and foil stamping. It represents less than 5% of all printing. The image is

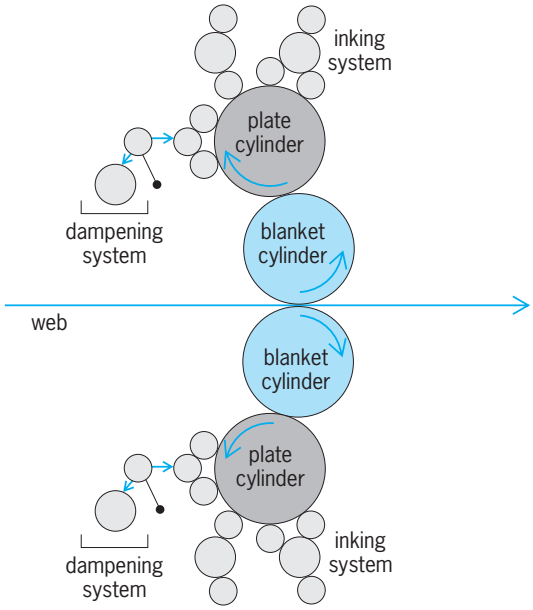

Fig. 14. Typical blanket-to-blanket printing unit of the horizontal type. (*After GATF Staff, Web Offset Press Operating, 4th ed., Graphic Arts Technical Foundation, 1996*)

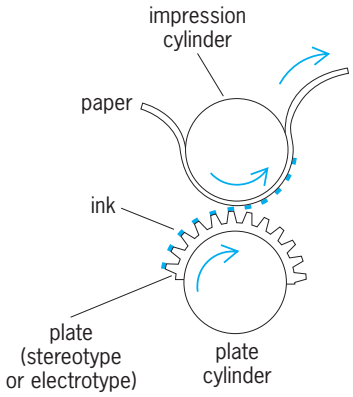

Fig. 15. Relief printing process.

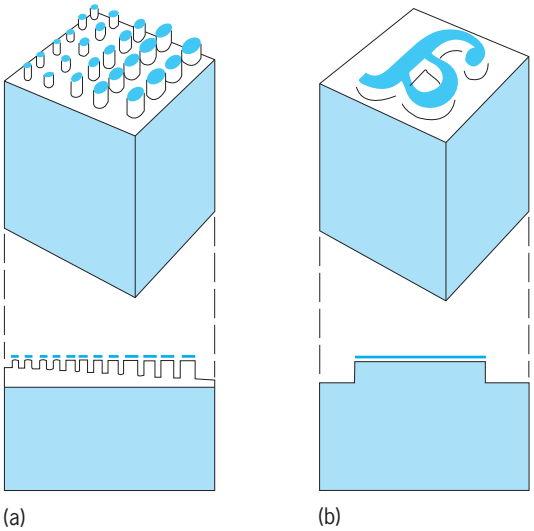

Fig. 16. Printing-image carriers for letterpress. (*a*) Dots of a halftone. (*b*) Type face. (*After V. Strauss, The Printing Industry, Printing Industries of America, 1967*)

wrong reading, or reversed, on the plate and is transferred directly from the plate to the substrate, where the image appears as right-reading (**Fig. 16**). Letterpress plates are typically made of zinc, copper, or a photopolymer.

Three are three major letterpress-printing processes. The platen press (**Fig. 17***a*) has a flat platen and bed. The flatbed cylinder press (Fig. 17*b*) has a flat bed, with an impression cylinder in place of a platen. The rotary press (Fig. 17*c*) has two cylinders, with the flatbed being replaced by a plate cylinder.

The most common size of a platen press is 10 × 15 in. (25 × 38 cm; maximum sheet). The platen press configuration, widely used in die cutting, creasing, and embossing of packaging material, can be considered as an offshoot of letterpress printing. For this purpose, platen presses are available in sizes up to about 64 in. (163 cm).

The flatbed cylinder letterpress was the predominant printing press until the rapid growth of lithography began. This configuration has become virtually extinct except for its use in the packaging field as a cutter/creaser and embosser.

Before 1950, the rotary letterpress was the means that identified printing and set the standard for printing quality. It was responsible for the growth of magazine, catalog, directory, book, and commercial printing in the United States. The great common-impression cylinder presses of the 1930s and 1940s almost completely disappeared from the scene as lithography, gravure, and flexography took over the printing markets.

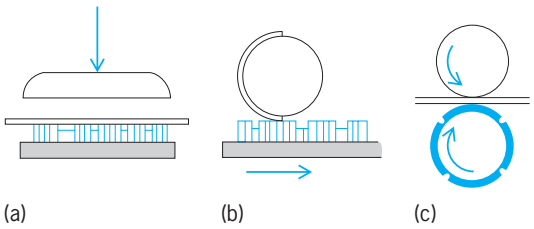

Fig. 17. Letterpress printing presses. (*a*) Platen press. (*b*) Flatbed-cylinder press. (*c*) Rotary press. (*After V. Strauss, The Printing Industries, Printing Industries of America, 1967*)

In 1970, rotary letterpress started a comeback with the development of photopolymer plates, ultraviolet-cured ink systems, and the design of high-speed rotary presses with interstation curing. The presses became popular in label printing, where the narrow-web (up to 25 in. or 63.5 cm) rotary press was used extensively for the production of pressure-sensitive labels and packaging for pharmaceuticals, cosmetics, and toiletries.

The belt press is a type of letterpress (**Fig. 18**). It uses two printing belts to which flexible, shallow-relief photopolymer plates produced from photographic negatives are secured with double-sided tape supplied in wide, accurately sized sheets (sticky-back). The plates are page size and imposed on the belts in the long-grain direction, on a plate mounter, to permit collating of completed books at the delivery end of the press. The use of belts, instead of the fixed-circumference used in conventional presses, permits a virtually limitless assortment of page sizes and page counts to be printed without time-consuming, costly plate changes. The press is ideally suited for the production of mass-market paperback books.

Flexography. Flexography is a growing process and frequently is used for printing on packaging materials. Flexography uses plates that have a slightly raised image made of rubber or plastic, including photopolymers. The inked plates are rotated on a cylinder that transfers the image to the substrate. Flexography is a high-speed printing process for printing on many types of absorbent and nonabsorbent materials, including plastic, foil, acetate film, brown paper, toilet tissue, bags, pressure-sensitive labels, corrugated board, paper and plastic bags, milk cartons, disposable cups, candy bar wrappers, envelopes, labels, and newsprint. It is also used for in-line numbering, in-line folding, dual-sided laminating, booklet placing, pattern gluing, dual-web registration, in-line UV coating on both sides of the web, promotional labels, booklet labels, instant redemption coupons, and encapsulated game pieces. It is also popular for long-run printing requiring one or two flat colors, and where sharpness and the highest print quality are not required. However, there is a growing trend toward high-quality flexography.

Flexography is a form of rotary-web letterpress. The print area or image consists of a raised surface, or relief, which can be inked and pressed onto the substrate. The rubber, plastic, or other flexible plate material is mounted to the printing cylinder with double-faced adhesive or slid on as a sleeve. Nonimage areas are below the printing surface and do not reproduce. The plates are sometimes backed with thin brass or metal sheets and attached to the cylinder with fastening straps for close register. The inked plates are rotated on a cylinder that transfers the image to the substrate.

Fig. 18. Belt press system. (*Cameron*)

Flexography inks are fast-evaporating solvent types, water-based types, or ultraviolet-curing types. The presses have a simple ink-distribution system in which a roller rotates in a fountain or trough of ink. The ink is transferred from the fountain roller to the printing plate by an engraved (anilox) roller with cuplike cells that meter the amount of ink transferred. The addition of a reverse-angle doctor blade to the transfer roller is sometimes used to meter the ink. The anilox method has been used on letterpress and lithographic presses and is known as keyless inking.

There are three types of flexographic presses: stack, common impression cylinder, and in-line (**Fig. 19**). In the stack-type press, the color sections are arranged one above the other or one on both sides of the main press frame. In the common impression cylinder press, a central impression cylinder press, also known as the drum, supports all of the color sections around a single impression cylinder mounted in the main press frame. This prevents register shift from color to color. The in-line unit press has separate color units horizontal one to the other, driven by a common line shaft. The in-line press can print on both sides of a web.

The growth of flexography is largely due to rapid improvement of plates, anilox rollers, and inking systems. Approximately 70% of all packaging is printed by flexography. The simplicity of the flexographic printing process, consistency, and relative low-cost make it increasingly popular among print buyers. Growth in corrugated cartons and flexible packaging as well as a trend to move away from glue-applied labels to pressure-sensitive labels should keep flexography in a growth mode for many years.

Flexographic printing has considerable impact on the environment based on its use of rubber and photopolymer plates, solvent-based inks, and hydrocarbon solvents. As a result, water-based inks are increasingly used by flexographic printers, especially when printing on paper substrates such as corrugated containers and newsprint.

Chrome-plated rolls have been replaced by laser-engraved ceramic surfaces. The doctor blades now wear instead of the roller. Laser-engraved ceramic rollers wear better with the doctor blade in place.

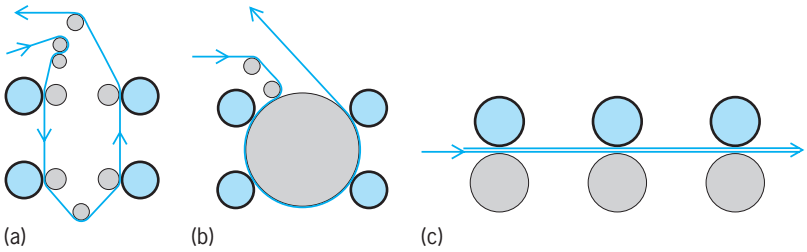

Fig. 19. Flexographic rotary presses. (*a*) Stack type. (*b*) Common impression cylinder type. (*c*) In-line press. (*After V. Strauss, The Printing Industry, Printing Industries of America, 1967*)

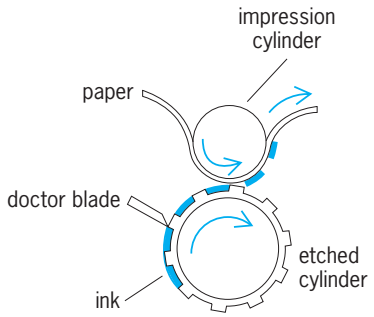

Fig. 20. Gravure (intaglio) printing process.

Several new drying and curing technologies have been developed that reduce energy costs and work well with new ink formulations. Infrared drying systems require moderate capital cost but have high operating costs. They are often used in conjunction with convection air dryers to provide the heating load to the coated web. Radio-frequency systems are an efficient alternative for drying water- and solvent-based coatings. There are high capital costs and moderate operating costs. Supplemental convection air dryers are sometimes used in conjunction with radio frequency.

Intaglio printing. In the various intaglio printing methods, the image to be printed is recessed into the surface of the printing plate. The incised image may be etched with chemicals or engraved with tools. Intaglio printing is used for printing paper currency and for fine-art printmaking. The image areas are filled with ink, and excess ink is wiped from the plates. Heavy pressure is applied to transfer the ink from the plates to the paper. Other forms of printing that use the intaglio process include engraving and gravure.

Gravure. Gravure is the opposite of letterpress. It involves printing from a recessed surface (the image is beneath the surface of the plate or cylinder), and the typical plate is a large copper cylinder (**Fig. 20**). The image is made of microscopic wells that are etched or engraved and filled with ink. The plate is rotated on a cylinder, transferring the image to the paper or other material. Gravure printing is often used for high-volume printing of packaging, wallpaper, and gift-wrap, using fast-drying inks. Although less common, gravure printing may also be used for printing magazines, greeting cards, and high-volume advertising pieces. Gravure is also popular for long-run magazines, packaging, and printing on foils, plastics, and cellophane.

Publication gravure printing is done by relatively few companies, which have large printing plants. Packaging and specialty gravure printing is done by a considerably larger number of companies, ranging from large integrated packaging companies with many gravure press units to small captive operations with only one or two press units.

Gravure is a high-cost and time-consuming platemaking process. It is economical for very long runs and for high-quality printing. The advantages of gravure are the long life of the recessed image and the uniformity of ink, resulting from the fact that the quantity of ink printed is determined by cell volume only. Fast-drying solvent inks, which make possible drying between each impression, and the ability to print well on a wide variety of surfaces, are other advantages.

U.S. Environmental Protection Agency (EPA) emission-control regulations require all printing processes to reduce emissions of volatile organic compounds. Gravure has the advantage that the solvent used in gravure inks can be recovered by carbon adsorption and recycled. In addition, nonpolluting water-based gravure inks have been developed and are widely used in packaging and specialty gravure.

The gravure image carrier is a copper-plated cylinder or copper plate on which the image is in the form of cells or cups engraved in the surface. If chrome-plated, the gravure cylinder can run for millions of impressions. A typical gravure cell is 35 μm deep and 125 μm square, with 22,500 cells per square inch (150-line screen). The walls between cells are known as the bridge or doctor blade support, and the intersections between the walls are called the corner posts. The gravure cylinder is flooded with a liquid gravure ink. The excess ink is removed with a high-quality steel doctor blade 0.006–0.010 in. (0.15–0.25 mm) thick. The ink is transferred by pressing the paper against the cylinder as it turns, using a rubber-covered impression roll.

The basic printing unit consists of an engraved gravure cylinder at the bottom with a rubber-covered impression roll at the top (**Fig. 21**). The gravure cylinder is geared to, and driven by, a motor-driven shaft, which drives the other gravure or converting units in the line. The impression roll is friction-driven through the web by the gravure cylinder. For the wide presses, used in publication printing, the impression roll includes a deflection-compensating

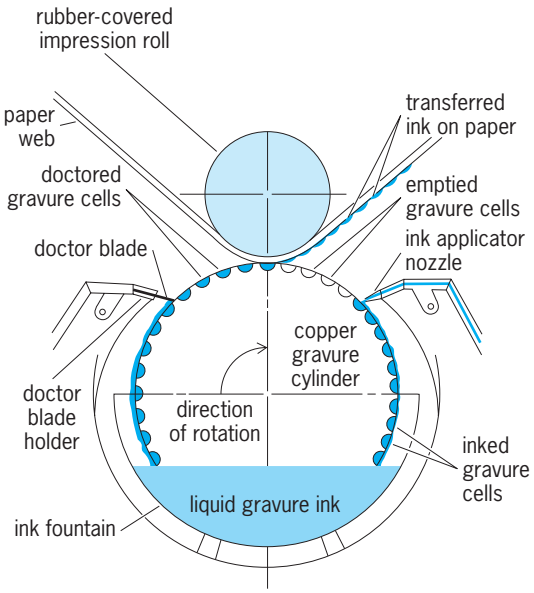

Fig. 21. Gravure printing unit.

mechanism to minimize bending. Mechanical impression screws or hydraulic or pneumatically operated impression pistons apply pressure to the end journals of the impression roll, pressing the rubber against the gravure cylinder.

The gravure cylinder runs in an ink fountain or trough. Spray nozzles may also be used to apply ink at a point close to the exit side of the nip. The ink fountain on high-speed presses using volatile inks is totally enclosed except for a small area between the doctor blade and the nip and on the exit side of the nip. Enclosure of the ink fountain reduces solvent evaporation and permits the use of fast-drying solvents.

The doctor blade holder is provided with an oscillating mechanism to permit adjustment of the blade angle, as well as the horizontal and vertical position of the doctor blade. Doctor blade angles, measuring from a normal to the gravure cylinder, range from 15° on publication presses to 45° on some packaging presses.

The side frames and structural elements of a gravure press unit are usually constructed of heavy cast-iron members to provide the mass and rigidity needed for vibration-free, high-speed running. Gravure press speeds range from 400–800 ft/min (2–4 m/s) in packaging and specialty printing, 1000–2500 ft/min (5–12 m/s) in magazine and supplement publication printing, and up to 4000 ft/min (20 m/s) or more on some printing operations running in line with paper machines.

Once the various press settings are made and the ink viscosity and drying conditions are adjusted for proper ink transfer and printed film characteristics, the gravure printing or coating operations is relatively simple to control. Maintenance of ink viscosity is the main requirement. This is done by periodically adding solvent to the ink to make up for evaporation. The makeup solvent is blended to maintain the primary solvent/resin mixture in the original ink formulation.

Gravure ink transfer is accomplished when the paper is pressed into contact with the liquid ink in the gravure cell. The characteristic hollow dot of the gravure indicates that ink transfer occurs around the edges of the cell where the ink first contacts the paper, and capillary action draws a portion of the remainder of the ink to the paper. In 1967, an electrostatic ink transfer process was introduced in the gravure industry. In this process, the paper web is charged by application of a dc voltage to the top of the paper through a slightly conductive rubber coating on the impression roll. The field lines established from the paper to the grounded gravure cylinder concentrate at the edges of the cell and exert an upward pull on the inductively charged ink, causing the meniscus to be distorted and bulged into contact with the paper. Dramatic improvement in print smoothness resulted.

Digital technology has transformed gravure printing from a craft to a high-technology industry. Virtually all publication and a growing percentage of packaging and specialty gravure cylinders are electronically engraved on electromechanical engraving machines. Predominant is a machine that uses an electromechanically driven diamond stylus (at 4000 cells per second) to engrave pyramid-shaped gravure cells in the copper cylinder. With an eight-engraving-head machine, a large publication gravure cylinder can be completely engraved in about 1 h, plus setup time.

Electronic page makeup systems interfaced to electronic engravers permit filmless production of gravure cylinders. As an interim step before switching to entirely filmless gravure prepress systems, the industry began electronic engraving from halftones instead of the traditional continuous-tone material. This technique, also known as halftone gravure conversion, permits gravure to handle much work formerly intended for offset lithography and reduces gravure prepress cost.

Engraving. Used for fine invitations and sometimes business cards and letterhead, engraving creates a raised surface on the front of the paper. A metal die with an impression cut into the surface is filled with an opaque engraving ink, then pressed against paper, creating raised areas of the paper coated with ink.

Screen printing. Screen printing is the simplest of the major printing processes and is popular for short-run printing. It requires a very thick ink film and is normally a slow manual process, though automated sheet-fed and web presses are available.

Screen printing can accommodate a wide variety of substrates that differ in shape, size, and material, including paper, plastics, textiles, metal, glass, ceramic, leather, and wood. Objects having cylindrical and irregular shapes, such as containers, barrels, glassware, and bottles, are frequently decorated by means of screen printing.

Screen printing is a versatile printing process. Based on simple stenciling techniques, modern screen printing uses precisely woven mesh to support a stencil, where the image is defined by open areas. The open areas allow ink to pass through the top of the screen to the substrate below.

Method. In the basic screen printing method, an industrial woven fabric (**Fig. 22**), generally polyester, nylon, or stainless steel, is stretched to a predetermined tension and affixed to a wood or aluminum frame. Stretching can be accomplished by holding the fabric in a clamping system driven with either mechanical or pneumatic motion to exert the necessary force to pull the fabric to the desired tension level. The fabric is then attached to the frame with a suitable adhesive. Self-tensioning frames can also be used, in which the fabric is clamped into the frame itself and elements in the frame sides are expanded or rolled to create the stretching force.

Fabrics range in mesh count, or fineness, from 25 to over 500 woven threads per inch and have various thread diameters. The woven structure of a screen contains mesh openings that allow ink to pass onto the substrate in the areas not blocked by the stencil during the transfer process. Higher mesh counts are used for fine-detail images, while lower counts are

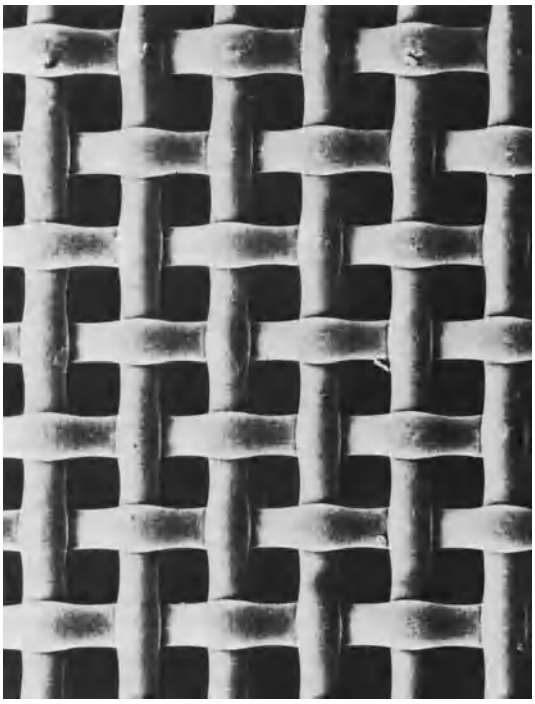

Fig. 22. Enlargement of meshed-screen fabric for screen printing. (*From A. Akermann, Influence of mesh on ink density in halftone printing, Technical Guidebook of the Screen Printing Industry, vol. B, Screen Printing and Graphic Imaging Association, 1987*)

used to deposit larger amounts of ink, as required by textile applications.

There are several types and classifications of stencil material. One of the simplest is a knife-cut stencil, in which the image is hand-cut from the material and the stencil is then adhered to the mesh. The most popular stencil type is a photosensitive emulsion (direct emulsion) that is coated on a screen and allowed to dry. The emulsion is then exposed through a film positive using a high-intensity ultraviolet light source. The unexposed areas of the emulsion are water-soluble and washed out, while the light-exposed areas become insoluble and remain as the stencil. A capillary-film stencil is precoated on a polyester sheet, applied to a wet screen, and then allowed to dry. It is processed the same way as the direct method. An indirect stencil has a gelatin composition and is exposed and processed independently from the screen and then adhered.

The fully prepared printing screen is mounted onto the press, leaving a slight gap between the bottom of the stencil and the substrate being printed. A polyurethane squeegee is positioned to permit a regular sweeping motion, and ink is introduced to the screen (**Fig. 23**). Ink transfer is accomplished when pressure from the squeegee brings the screen and substrate into intimate contact. The forward stroke moving the squeegee then draws the ink across the screen while creating a shearing force that transfers the ink through the open areas of the stencil onto the substrate. Following ink transfer, the screen separates from the substrate because of screen tension, leaving behind the printed image.

In most screen-printing applications, a color must be printed and dried before the next one can be applied. Registration is necessary for multicolor prints, and this is attained by using automated or manual techniques. Textile applications are screen-printed by using direct-contact printing (screen and substrate are completely touching), and colors are printed wet on wet before the ink is cured.

Presses. Various press types are available for screen-printing. The most common is the flatbed press, where the squeegee is drawn back and forth over a stationary screen while the substrate is held firmly in place with a vacuum bed. Textile applications use flatbed presses.

On cylinder presses, the screen is driven back and forth under a stationary squeegee as the substrate is simultaneously moved under the printing point. Similar presses are used to print cylindrical objects, such as cans, by rotating the object under the moving screen.

Most screen-printing presses can handle only piece goods, but web presses are also available for printing roll stock and textiles. Rotary presses are web presses that use a cylindrical screen with a squeegee inside. This press configuration boasts much faster printing speeds than both cylinder and flatbed presses, and is used primarily to print textiles and wall coverings.

Screenless printing. Screenless printing has the potential of coming closer to printing facsimile reproductions of originals than any of the processes using screens or halftones. It has no undesirable moiré patterns caused by overlapping screens, and the colors are purer and more saturated, since they are not degraded by light absorption in the paper between the dots of halftone patterns. Processes involving screenless printing are considered essentially specialty processes because of the difficulty of controlling them. These are collotype and screenless lithography. Digital means also are used to simulate the effects of screenless printing.

Collotype. This type of printing is known in different countries by different names, for example,

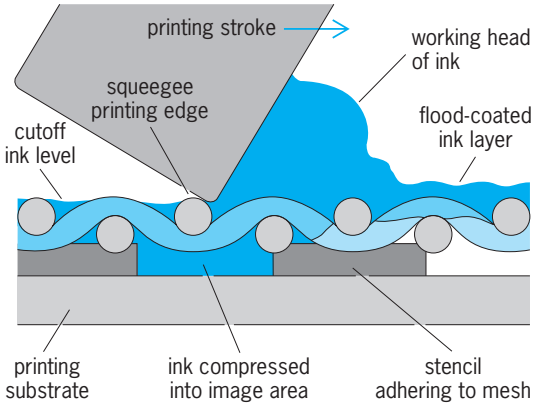

Fig. 23. Basic factors in the screen-printing action. (*After E. J. Kyle, Principles of quality control, Technical Guidebook of the Screen Printing Industry, vol. C, Screen Printing and Graphic Imaging Association, March 1975*)

photogelatin, phototype, heliotype, heliogravure, albertype, and artotype. The process uses no screen, but the prints closely simulate in gradation true photographic prints.

Originally, collotype printing was a hand process in which production might amount to no more than 100 impressions per day. With a good machine, the modern process may average up to 2000 impressions per day. The plate, usually aluminum, is coated with a layer of light-sensitive (bichromated) gelatin, over which is laid an unscreened (continuous-tone) photographic negative of the copy to be reproduced. When exposed to light, the gelatin is hardened in varying degrees in proportion to the amount of light received from the dark, medium, and light portions of the negative. The plate is developed in water to remove unexposed bichromate and immersed in a glycerin or glycol solution. The result is a printing image that varies not only in thickness (in thousands of an inch) but also in moisture content and consequent ability to repel ink or hold it in varying amounts corresponding to exposure of the gelatin. The harder areas accept less moisture and swell less, thus accepting more ink.

The plate is kept moist by maintaining the pressroom relative humidity from 30 to 95%. When the ink roller passes over the plate, which is dampened by the moisture in the air, the greasy ink is picked up in inverse proportion to the amount of moisture in the partially hardened portions of the gelatin.

Technically no two prints are exactly alike in collotype, because the tone reproduction varies with the moisture content of the gelatin. Each impression removes moisture, so the relative humidity must be gradually increased to compensate for the moisture loss. Eventually, the tone reproduction becomes flat, and the plate must then be soaked again in glycerin or glycol solution and the relative humidity reduced to start a new cycle of printing.

Collotype is a short-run process. Runs of only a few hundred prints are common, with 5000 as the practical maximum. The process has been used extensively for art reproductions and for a wide range of advertising work in both black-and-white and color, including counter and theater displays, posters, blowups (enlargements up to 42 × 62 in. or 107 × 157 cm) or advertisement and pictorial copy, and facsimiles of photographs.

Screenless lithography. This is a means of printing lithographic plates made from continuous-tone positives or negatives on an offset press. Best results have been obtained with special photopolymer plates exposed from positives. Unlike collotype, the plates are run conventionally on the press with regular inks and fountain solutions. The random grain pattern of the images, corresponding to the grain of the plate, is almost indistinguishable from the gelatin characteristic of collotype printing, and screenless lithography has been mistaken for collotype.

Plate coating variations, which affect the consistency of printing results, are a disadvantage that has kept the use of screenless lithography to a minimum. The tone reproduction of plate coatings depends on the roughness of the plate surface (grain) and the thickness of the plate coating. Since each is dependent on the other, and neither can be accurately controlled in manufacture, plate results are inconsistent and unpredictable.

Another approach in simulating the effects of continuous-tone reproduction is to use special screens with random grain elements of about 500–600/in. (200–240/cm). An additional method is the production of images on electronic color scanners with very fine elements in random patterns. Inkjet and dye-sublimation color proofing systems also produce screenless images.

Digital printing. First introduced in the 1990s, digital color printing presses have had a major impact on the printing industry sooner than most believed they would. Computer-to-film technology, which bypasses the laborious tasks of graphic arts photography and image assembly, was a major step toward automating the printing processes in the late 1980s. Computer-to-plate technology, an extension of this technology, gained in popularity in the 1990s. However, computer-to-press technology provides the promise of completely integrated printing systems. Direct-to-press color systems were introduced throughout the 1990s, such that by 2000 direct-to-press color-printing systems were competing with commercial printing presses.

On the hardware side, systems that handle a greater variety of substrates, including very lightweight paper, have been developed. The concept of a universal copier/printer also developed. These devices output color and monochrome pages at competitive costs with dedicated color and monochrome printers.

Adding value by integrating all services from front-end to printing to finishing is the key to digital printing, as is flexibility and process improvement. A modern printing operation uses the World Wide Web as a common business tool for receiving information and for executing related business functions. Innovations in raster image processor (RIP), server, and workflow technology are making it easier to integrate digital devices into almost any printer's array of output services.

Within the range of applications that they were designed for, digital presses compare with offset lithography in terms of print quality. While digital systems complement conventional offset lithographic equipment, they do not displace offset as digital presses cannot yet deposit ink or coatings, print complex items such as magazines and die-cut packaging, or turn litho-sized papers into products at thousands of sheets per hour or thousands of feet per minute.

For the most part, digital presses use toners (dry and wet) in place of traditional printing inks. The exceptions are those lithographic presses that use digital technology for imaging plates and then print in the traditional offset manner.

Digital printing press systems have many commonalties. They are all on-demand printing systems, which address the 50- to 20,000-impression market.

In addition, they all provide perfecting (two-sided printing) capabilities, require short or no makeready time, and require no warehousing of finished products because all are produced one at a time and shipped to customers upon completion. Systems using dry toners face some print quality deficiencies related to gloss. And those using wet inks provide images approaching offset lithography in quality. Most of the systems provide variable imaging capabilities.

Color copiers. An offshoot of digital on-demand printing presses is digital high-speed color copiers, which are starting to resemble low-end variable-imaging digital printing presses. This technology addresses the need for up to 50 copies but can easily be cost-effective for much larger quantities.

Wide-format digital printers. Wide-format digital printers provide individuals and companies with access to inexpensive large-sized prints in quantities of one to approximately 100. A growing number systems are available that can produce full-color digital prints in sizes ranging from 36 in. (91 cm) to billboard-size sheets. There are two components to these systems: a large format printer and a raster image processor.

Large-format systems print on a variety of substrates, including paper and Mylar, using inkjet, electrostatic, or thermal wax transfer technology. This market is growing rapidly and is finding its niche in quick printing and on-demand printing. Users of this technology include advertising agencies, screen-printers, service bureaus, and in-plant printing departments of many businesses. The applications are numerous and include items such as art-on-demand, backlit signage, transportation advertising displays, engineering drawings, maps, murals, posters, window graphics, and billboards.

Dye sublimination. Dye-sublimation printers use solid dyes and heat to produce photo-quality images. Cyan, magenta, yellow, and black dyes are coated in page-sized panels on a roll of transparent film. Heating elements in the print head vaporize the dyes onto a specially coated paper. The intensity of the color is controlled by precise variations in temperature.

Dye-sublimation systems print color in continuous tones one color at a time, instead of dots of ink. The color is absorbed into the paper rather than sitting on the surface. Therefore, the output is more photorealistic, more durable, and less vulnerable to fading than other ink technologies. Dye-sublimation printers are favored by some graphic designers for high-end proofing and for applications that require the best possible color.

Inkjet. Typically, an inkjet printer forms images by spraying tiny droplets of liquid ink onto paper. The small size and precision placement of the ink dots produce near-photo-quality images.

Two primary inkjet technologies are thermal bubble, or bubble jet, and piezoelectric. The former uses heat and the latter uses a crystal and an electric charge to apply the ink. Inkjet printers are inexpensive and produce good color but can be slow, especially for short-run color printing. The best results are normally achieved when printing to specially coated inkjet papers. Inkjet printing is popular for variable-data printing, including addressing, coding, computer letters, sweepstakes forms, personalized direct mail advertising, and imprinting of forms and invoices. All these applications involve variable information, which is the basic advantage of digital printing. An important application of inkjet printing is color proofing.

Inkjet printing is also classified as continuous drop (a stream of ink) and drop-on-demand (ink is squirted only when needed). In inkjet printing, the ink is ejected through a nozzle under pressure and transformed into uniform drops by vibration of a piezoelectric crystal; the drops are charged electrostatically and deflected by electrostatic charges.

There are four basic types of inkjet printing systems. One uses a computer-activated single nozzle that oscillates across a substrate much like an electron beam scans a television tube to produce a picture. The second method uses a bank of nozzles, each of which is digitally controlled by a computer program. The third is bubble-jet, which uses heat to produce bubbles that form the ink droplets. The fourth uses phase-change technology, in which color sticks that are solid at room temperature undergo two phase changes during printing: one when the stick melts to form liquid drops, and the other when the ink droplets hit the paper and solidify.

Modifications of inkjet systems are used for short-run, large-format poster or billboard printing, where runs are very short (1–100 impressions) and speed is not an urgent requirement. The large-format market also includes posters, outdoor advertising, transportation advertising, and in-store displays in sizes up to 20 × 60 ft (6 × 18 m) in about 5 h. The substrate is a sheet of vinyl that is wrapped around a huge cylinder 17.5 ft (5.25 m) in diameter (58-ft or 17-m circumference) for imaging. The images are painted with 16 paint jets, called dot shooters, which discharge as many as 31 million dots to create the final image. Single-sheet billboards require derricks for installation, and most are backlit for night viewing. *See* INKJET PRINTING.

Laser. Static electricity is the principle behind laser printers. A revolving drum or cylinder builds up an electrical charge, and a tiny laser beam pointed at the drum discharges the surface in the pattern of the letters and images to be printed, creating a surface with positive and negative areas. The surface is then coated with toner—a fine powder that is positively charged, so it clings only to the negatively charged areas—which is then transferred onto the paper. The paper then passes through heated rollers, which fuse it to the paper. Color lasers make multiple passes, corresponding to the toner colors.

Speed and lower operation costs make laser printers attractive. Laser printers have their own command language that translates the digital information in a file into the bitmap images printed on the page. Adobe PostScript capabilities of many laser printers make them popular with graphic designers and

desktop publishers who often use EPS images and PostScript files.

Electrostatic and electrophotographic. Electrostatic printing is similar to xerographic or electrophotographic printing. It uses toner particles, and the image is formed from an analog original. There is no print drum required. Toner particles are attracted directly to the paper, and no optical system is required. The entire copier glass is exposed at once, and the electrostatic charge is directly deposited on the paper. Liquid or dry toners can be used. Liquid toners are fused to paper by hot air. Dry toners are fused to paper by pressure.

Electrophotographic printing is similar to high-speed copier technology. It operates on the principle of xerography. However, electrophotographic printers operate from digital data and use lasers. Images are created by a moving laser beam across an imaging drum. The image is developed with dry or liquid toners and works well for single-color printing. High-quality color printing is possible, but slow relative to other printing processes.

Electrophotographic systems have drums coated with photocondcutors that are charged overall with a corona charge. The coating is exposed to an image by light that discharges the charge in the nonimage areas. The image is formed using powder or liquid toners in the charged areas, and then transferred to a substrate and fused, or fixed, by heat, solvent vapor, or some other method (**Fig. 24**).

One example is the two-unit web-fed electrophotographic press, which uses cylinders coated with organic photoconductors and liquid toners to produce the image and has many special in-line finishing features and provisions for printing one color on both sides of the paper or two colors on one side. The press prints variable information at speeds up to 300 ft/min (90 m/min), and is capable of changing up to 25% of the information from impression to impression.

Two types of full-color electrophotographic digital presses are used. One uses laser imaging, a special liquid toner, and essentially a single-color offset press with a plate, special heated blanket, and a cold impression cylinder; paper feed and pile delivery system; a means for turning prints over and printing the reverse side (perfecting); and some finishing options. Printing has 800-dpi resolution. Single-color prints can be produced at A3 size (11.5 × 16.5 in.; 297–420 mm) at a press speed of 4000 sheets/h. Since color printing is done by consecutive impressions of the same sheet on the impression cylinder, the output for four-color (one-side prints) is 1000 sheets/h. Duplex printing (two sides) reduces the output to 500 sheets/h, 8 sheets/min, or 32 pages/min (two sides). The system is capable of printing six colors.

The other electrophotographic digital press is a web press with eight in-line units (four on each side of the paper), so it can print both sides of the paper in one pass. Exposure of the photoconductor plates on each cylinder of the printing units is done with light-emitting diodes assembled into arrays of

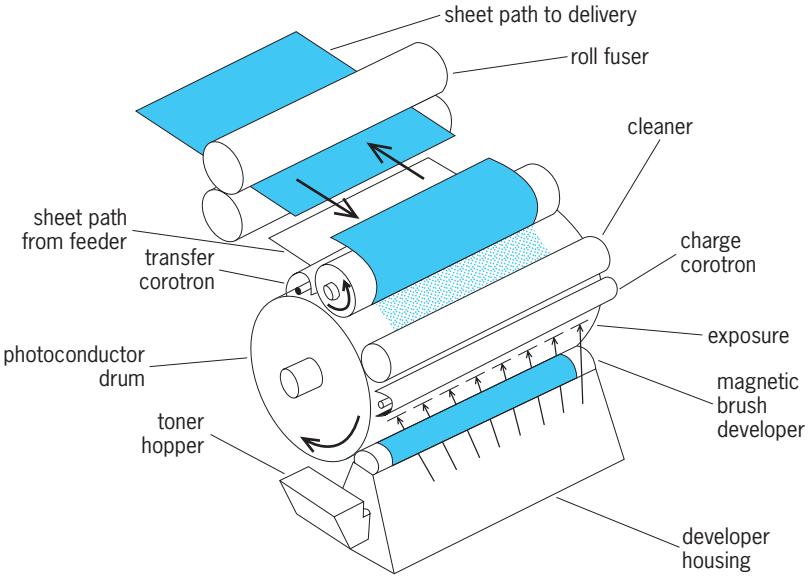

Fig. 24. Typical digital press print unit.

7400 diodes, spaced at a density of 600/in. (236/cm), corresponding to a spatial resolution of 600 dpi. The system has variable spot intensity, with each spot having 64 gray levels, by applying different amounts of a special dry, dual-component toner made up of small particles. The image is composed of single-line screens with continuously variable widths and placed at different angles, according to the amount of color needed. *See* LIGHT-EMITTING DIODE.

There are two digital electrophotographic printing systems for larger-print, poster-type images. One consists of a computer front-end digital printing system that incorporates a wide array of digital color scanners and a 44-in.-wide (340-cm) and a 54-in.-wide (137-cm) electrostatic color plotter modified by a special toner control unit. Imaging is done with a special frequency-modulated (FM) screening algorithm. The system is used for indoor and outdoor posters and paneled billboards. The toners are specially treated to resist fading from exposure to ultraviolet radiation. The other system is a composite electrophotographic system, consisting of existing components with extensive software and hardware modifications. It uses specially treated toners and substrates, as well as absorbing overlaminants to resist ultraviolet radiation. Prints 34 in. (87 cm) wide and up to 26 ft (8 m) long can be produced. Larger images can be panelized. Prints are used for fleet and transit advertising, industry expositions and exhibits, visual merchandising and displays, murals, outdoor exhibits and events, banners, and prototyping. *See* ORGANIC CONDUCTOR; PHOTOCOPYING PROCESSES.

Thermal. Falling somewhere between dye-sublimation and solid ink technologies are thermal wax printers. Thermal wax printers produce vibrant color but require very smooth or specially coated paper or transparencies for best output.

There are two thermal printing processes: direct thermal and wax transfer. In direct thermal printing,

a specially coated paper, when heated by a print head, reacts by turning color. In the wax transfer process the same chemical principle is used; however, a special coated ribbon is used which releases pigmented material. Characters are printed one at a time by a print head that moves back and forth.

Printing Inks

Changes in the technology of printing ink have been driven by environmental regulations, workplace safety, and increasing press speeds. In the United States, federal and state regulations have dominated the technology of all inks since the 1980s. Regulations have spread to include ink formulation, labeling, shipping, solvent type and content, emissions of volatile organic compounds, heavy-metal content of inks, and solid- and liquid-waste effluents from printers and ink manufacturers. All of these factors affect ink technology and development. *See* HAZARDOUS WASTE.

Paste inks. In lithographic printing, a dominant consideration has been the increasing speed of running web and sheet-fed presses. Web offset presses are capable of printing at speeds of 3000 ft/min (900 m/min). For heat-set paste inks, the drying rate is an important consideration, since the print must dry rapidly at high speeds.

Sheet-fed inks. Sheet-fed lithographic presses have also increased in printing speed and run as fast as 18,000 impressions/h. The ink formulations need to be adjusted so that the drying (setting) times are fast enough to avoid the transfer of ink from a printed sheet to the back of the next sheet (set-off) in the stack of output sheets. Most of these formulations have extremely quick setting characteristics and become tack-free within a few seconds after printing.

Ultraviolet curing. Ultraviolet curing is favored, particularly for clear overprint varnishes and in applications that require instant drying and processing of the dried prints. Typical applications include packaging with in-line die cutting, packages for cosmetics and liquors, and compact-disk packaging. The growth of ultraviolet overprint varnishes has been mostly in the area of book and magazine covers printed with either ultraviolet-set or conventional heat-set or sheet-fed inks. The driving forces in these applications are protection, glossy appearance, and enhanced color brightness.

Electron-beam curing. High-energy, linear cathode electron-beam generators are used to cure inks and coatings in Europe and the United States, particularly for the manufacture of aseptic paper/foil packaging for beverages such as citrus drinks, milk, and wine. Food products in such packages have a shelf life of several months without refrigeration. Typically, these inks are printed in a web press and overcoated with a clear electron-beam-curable varnish.

Packaging inks. Air-pollution regulations have shifted the emphasis toward water-based liquid inks. While the use of these inks and varnishes in flexography or gravure packaging is not new, most applications have been for paper or board printing. *See* AIR POLLUTION.

Compared to flexography, packaging gravure is used for a wider variety of substrates, resulting in more adhesion problems. As a result, binders are used that require solvents that are difficult to recover and separate in carbon-bed absorption units. Thus, many packaging gravure printers use solvent incineration or recovery instead of water-based inks. Most of the water-based inks in this market are limited to paper and board substrates.

More than 50% of the flexographic packaging inks are water-based. Some plastic film applications (especially bread bags) have resisted the use of water-based flexography inks because of the nature of the additives in the films, which impede adhesion and wetting. In addition, water-based inks are slower-drying, limiting the printing speeds.

Another packaging area involves radiation curing and a high-speed web offset, wide-web paperboard press. This press uses offset lithography paste inks formulated to be cured with ultraviolet radiation or electron-beam energy to produce finished cartons at speeds of up to 1400 ft/min (420 m/min).

Publication gravure. The major type of ink used throughout the world for this high-speed printing process (up to 3000 ft/min or 900 m/min) is a toluene-based liquid ink using a metallized rosin binder. The toluene vapors emitted during the drying of these inks are captured by carbon absorption beds. The beds are regenerated, and the recovered toluene is reused in ink manufacture. While this system is usually better than 95% effective, there are problems in corrosion and decreased recovery efficiency. In addition, health and safety regulations restrict the amount of toluene vapor allowable in the pressroom atmosphere (around 50 parts per million).

Screen-printing inks. There are four categories: aqueous, evaporative/oxidizing, catalytic, and ultraviolet. In aqueous inks, heat is required to evaporate the primary vehicle, water, leaving the pigments and binders behind to form the dry printed image. Evaporative/oxidizing inks contain solvents or oil vehicles. As with aqueous inks, the vehicle heating is used be evaporate the vehicle or oxidize the oil. Catalytic systems incorporate reactive resins in a chemical process by which the ink is transformed from a liquid to a solid. Ultraviolet-curable inks contain no water, oil, or solvents, and change from a liquid to a solid when exposed to specific wavelengths of ultraviolet light This type of ink can be applied to a wide range of plastics, metals, and paper. *See* DRYING OIL.

Screen-printing inks are available in hundreds of formulations to address different applications. Many specialized screen-printing inks are used to achieve specific effects and functions. Included are thermochromatic, photochromic, fragrant, fluorescent, scratch-off, conductive, ceramic, glitter, and metallic inks.

While other printing processes are primarily restricted to direct and indirect printing on paper,

plastic, and metal sheets, screen-printing is capable of imaging virtually any planar or three-dimensional substrate. Another distinctive characteristic of screen-printing is its ability to lay down a heavier ink deposit than other processes, resulting in bright vibrant color reproduction and exceptional lightfastness for outdoor applications. *See* INK.

Printing Papers

Printing papers are designed primarily for printing one or more colors on each of the two sides. Secondary requirements are that the paper meets the needs of operations subsequent to printing, including folding and binding, and the needs involved in end use. To fulfill the diversified demands of modern printing, printing papers are manufactured in many different finishes (coated and uncoated), degrees of whiteness, colors, basis weights, grades, and types.

Various fibrous and nonfibrous materials, in combination with different degrees of fiber refining and adjustment in paper machine settings, are used to obtain the balance of paper properties required for the desired print quality, press performance, and demands of a specific printing process. Generally, the properties are maximized in accordance with other secondary properties for folding, binding, strength, and cost, and include dimensional stability to maintain printing register, opacity to minimize show-through of the printed image, the degree of smoothness and compressibility demanded by the printing process, and a surface optimized for printability.

Often, a balance must be struck in the fiber and finishing conditions to optimize strength and stiffness to enable the press to feed the sheet mechanically without difficulty and surface properties for printability. Paper coated with minerals such as kaolin and calcium carbonate, and binders such as styrene-butadiene latex and starch can provide a superior printing surface without compromising sheet strength.

Letterpress. For letterpress printing, micro-smoothness (without pits or voids) of the paper surface closely relates to the completeness of image transfer from the printing plate to the paper, and to the extent to which halftone dots are a true reproduction of those on the plate. A level paper surface uniformly absorbent to printing ink is required for consistent density, color, or gloss of the printed image. Compressibility of the paper must be adequate to assist its surface in contacting all areas of a relief printing plate.

Flexography. Paper requirements for flexographic printing are less demanding than for letterpress or planographic printing. Flexography uses comparatively low-tack inks that dry by low-temperature evaporation or absorption into the paper. Consequently, the stresses applied to the paper's surface are less.

Modern photopolymer plates combined with special press designs made flexographic a printing technology capable of process color at the quality level of low-end offset lithography. In order to achieve this quality, flexographic paper must have a smoother and flatter surface than for offset. Flexography is largely water-based, and thus the paper and board used must resist surface deterioration or dimensional change when wetted.

Offset lithography. Offset lithography places more stringent demands on paper than other printing processes.

Sheet-fed. Paper fibers must be well bonded, with sufficient water resistance to prevent them from softening and picking (rupturing at the surface) under the stresses applied to paper as it is repeatedly separated from the inked rubber blanket during multicolor printing. Excessive moisture pickup from repeated contact with water of the lithographic dampening system could lead to intolerable fiber expansion, changes in paper dimensions, and printing misregister. Paper surfaces must be exceptionally clean and have no loosely bonded material that can be lifted off and transferred to the offset rubber blanket, thereby introducing objectionable defects in the prints. Additionally, the paper must not release materials such as fillers or coatings that accumulate on the offset blanket or react with the fountain solution to change its chemistry.

Flatness of the paper sheet with straight undistorted sheet edges is a critical requirement. In addition, it is generally necessary that the paper's moisture content be in balance with the relative humidity of the pressroom to maintain satisfactory register for multicolor printing.

Web. Requirements are basically the same as those for sheet-fed offset lithography, except the moisture requirements of the printing paper is generally less than for sheet-fed offset printing and need not be in balance with pressroom relative humidity. Heat-set web offset papers usually have porous coatings for rapid escape of water vapor in order to prevent blistering during ink drying. A coated web offset paper must be sufficiently strong to fold without cracking after loss of moisture.

To maintain satisfactory printing register and runnability, rolls must unwind with even tension, remain flat across the web, and have minimum paper stretch during printing.

Gravure. Gravure printing requires a flat paper surface that has micro-smoothness for uninterrupted ink transfer from the inked cells to the paper. The compressibility of the paper aids in contacting the surfaces of the small gravure cells. High paper-surface strength is not essential. Paper and board for packaging are often gravure-printed with water-based inks, which must have appropriate characteristics for wetting and ink spread. Dimensional stability is important for color registration, as the inks are dried with hot air between color stations. Paper rolls must feed into the press with even tension and flatness across their wide webs for satisfactory runnability.

Uncoated papers. Uncoated printing papers are manufactured from bleached chemical wood pulps, a mixture of chemical and mechanical wood pulps,

a combination of chemical wood pulps, recycled fibers, or cotton or linen fibers depending upon type, grade, and availability. The hardwood and softwood species that produce short- and long-fiber chemical wood pulps, respectively, are commonly used in various proportions in the papermaking furnish (the mixture of fibrous and nonfibrous materials) to achieve the desired balance of properties for a printing paper. Other components of the papermaking furnish that determine the characteristics of an uncoated printing paper are fillers, sizing, dyes, and starch.

Alkaline papers. The uncoated fine-paper industry largely converted from acid to alkaline papermaking to take advantage of high brightness filler such as precipitated calcium carbonate, and to provide greater permanence. Alkalinity is provided by substituting a synthetic material for the rosin-alum size used in acid papers. Even wood-containing papers are being produced under neutral conditions to avoid strength loss from acid hydrolysis. However, permanence is not considered crucial for daily newspapers, so newsprint remains an acid paper grade.

Newsprint. Newsprint is used chiefly for the printing of newspaper by web offset, some flexography, and letterpress. Its minor uses are advertising supplements, government publications, comic books, and commercial printing. In the United States, the basis weight of newsprint is defined as the weight in pounds per 3000 ft^2 (ream) of paper. In almost all other countries, including Canada, the basis weight is defined in grams per square meter of paper. Newsprint ranges in basis weight from 25–32 lb per ream (40–52 g/m^2). Newsprint can be manufactured by using one or a mixture of the following: thermomechanical pulp, stone-ground wood, chemithermomechanical pulp, recycled newspapers and magazines, and reinforcing chemical pulp. Mechanical fibers that originate from ground wood or thermomechanical pulps have lower brightness and yellow more readily with aging than chemical wood fibers, but they are used for their lower cost and high opacity.

Mechanical pulp-containing papers. These printing papers are made with varying percentages of mechanical and chemical wood pulps, depending upon the grade. They resemble newsprint, but are somewhat brighter and have a smoother finish produced by supercalendering or hot soft calendering. Their basis weights normally range 20–45 lb per ream (30–68 g/m^2). Typical uses are for the printing of periodicals, catalogs, directories, advertising circulars, newspaper supplements and magazine sections, and paperbacks.

Wood-free papers. This group includes offset, book, bond, envelope, and text papers.

Offset lithography provides the versatility of printing papers with all types of finishes. Finishes include antique, eggshell, vellum, wove, smooth, machine, and English. Offset papers may have distinctive patterned finishes obtained by embossing. Their brightness range extends from approximately 75 to 90 or more, depending upon their quality and grade. Offset papers can be manufactured with extra opacity and in various colors to fulfill the varying needs of the printing industry. Standard basis weights range 40–120 lb per ream (59–178 g/m^2).

Uncoated printing papers for book manufacturing are offset papers that are made to bulk a specified number of pages per inch and to assure that the total thickness of the finished book conforms to the predesigned thickness of its cover or casing. Other requirements for book manufacturing are different white shades and finishes that relate to the esthetic qualities of a book and other needs peculiar to book binding.

Text papers are high-quality printing papers made solely from chemical pulp, in many different finishes or textured surfaces, and in white and colors. Some text papers are watermarked and have a deckled edge. Common basis weights are 60, 70, and 80 lb per ream (89, 104, and 118 g/m^2) plus matching cover weights. Text papers are printed mostly by sheet-fed offset lithography.

Lightweight printing papers are characterized by their extraordinarily high opacity for their basis weights, which range approximately 18–40 lb per ream (25–60 g/m^2). Their fiber composition may consist of chemical pulp, recycled cotton, linen fiber, or some combination of these. Lightweight papers are printed mostly by offset lithography. They are used for the bulk reproduction of bibles, encyclopedias, reference books, directories, and manuals; and for lowering first-class mailing costs of printed matter such as financial reports, prospectuses, and proxies.

Coated papers. Mineral coatings such as kaolin, calcium carbonate, or other pigments, and coating binders such as starch or various synthetic binders such as styrene butadiene or acrylic latex are applied to the base paper predominantly by blade coating as well as by roll and air-knife coating either on or off the paper machine. Coating formulations and base papers differ, depending upon the printing process, the quality level, and end-use considerations.

Grade levels of coated papers from the highest to the lowest quality are sometimes designated in six grades: premium and numbers 1, 2, 3, 4, and 5. The premium and number 1, 2, and 3 grades have chemical-pulp base stocks. Grade 4 may or may not contain mechanical fiber. The number 5 grade is made with about 50% mechanical fiber. The higher grades of coated papers are characterized by higher gloss, whiteness, and superior print quality.

Premium and grades 1, 2, 3, and 4 are used for sheet-fed offset lithography. Grades 3, 4, and 5 are used mostly for web offset and gravure printing. Basis weights of coated paper for sheet-fed lithography extend 60–100 lb per ream (90–150 g/m^2) with matching cover weights.

The number 4 and 5 coated grades made with mechanical fiber are used primarily for publication and periodical printing by rotogravure, or web offset, in a basis weight range of 28–50 lb per ream (40–75 g/m^2). Catalog and commercial printing

represents the remaining usage of these lightweight coated papers.

Dull-finish enamel. Dull-finish enamels have a lower gloss than their high-gloss counterparts and are used where the printing quality of an enamel surface with low paper gloss is desired.

Cast-coated paper. Mirrorlike gloss-coated papers having an exceptionally smooth level surface and uniform ink absorptivity are made by a unique method of casting the coating on its base paper against a heated polished metal cylinder. Binders such as casein are used to provide the right adhesion, release, and porosity.

Matte-coated paper. Matte-coated papers provide better lithographic reproduction than uncoated papers while still providing a gloss-free background for printing. Ground calcium carbonate is a common coating pigment.

Coated-one-side papers. Coated-one-side paper is used for the printing of flexible packaging, labels, and book jackets. This type of paper is designed for offset lithography, flexography, or rotogravure and for other postprinting requirements such as lacquering, bronzing, embossing, die cutting, and labeling.

Ink-jet printing. This requires a paper surface that has the capacity to dry the impinging drops of ink by absorption without wicking or spreading of the ink for maximum printing density and sharpness. Uncoated papers for single-color ink jet are generally well sized, and papers for high-quality color ink jet are generally coated with a transparent layer such as silica or gelatin.

Recycled papers. Since paper represents a major part of municipal solid waste, much effort and promotion have been devoted to recycling the paper waste and using recycled papers for printing. In the United States, recycled newsprint has been used for years, but the use of recycled papers for business and commercial printing started in the 1980s. Three cost constraints affect the use of recycled fibers: sorting the waste paper, deinking it, and transporting it to mill and deinking sites far from populated areas. *See* PAPER.

Synthetic paper. Synthetic paper was developed in the 1970s when paper resources were scarce and petroleum was inexpensive. It was also a time when the capacity to produce paper in the United States had decreased because of papermaking machine shutdowns by the Occupational Safety and Health Administration (OSHA) and the EPA, resulting from the discovery of unsafe working conditions and air and water pollution. It was also thought that diminishing wood resources could drive up costs of paper substrates.

Synthetic papers are substrates made of materials other than the natural fibers of trees and the resulting pulp. Most synthetic paper is produced from polypropylene or polystyrene, which may contain minerals and other additives. In addition, any film produced and/or treated to improve its printability, as well as any process that improves the printability of an opaque film substrate, is considered synthetic paper.

Synthetic paper has the look, feel, and processing capabilities of traditional printing paper but is stronger and more durable. Synthetic papers have a higher tensile strength and support higher processing speeds, as well as improved moisture, grease, and chemical resistance.

Ink selection for synthetic paper is important. Most synthetic papers perform best with thin ink films and low mineral oil content, because absorption can be minimal even with auxiliary drying. Oxidizing inks are often recommended for synthetic papers because they dry through evaporation and do not totally rely on absorption. The absorptive nature of the surface of the substrate will determine ink/water balance and ink film thickness when printing on synthetic paper. Clay coatings allow ink absorption, which aids in printing.

Thermal transfer, direct thermal, inkjet, and in some cases laser printing are markets for synthetic paper. Digital printing for short runs is an application for synthetic and tear-resistant papers. Inkjet printing is the most popular, while laser applications are growing.

The dry toner fusion process is popular, but excessive heat causes dimensional instability such as shrinkage, cracking, curling, and sometimes delaminating. Newer copy machines and laser printers have addressed these problems with adjustable-temperature fusion rollers and tear-resistant papers.

In terms of the environmental friendliness, synthetic papers' raw materials are nonrenewable (petrochemicals and minerals), but they can be recycled or burned as an energy source. Since production of synthetic paper is costly, manufacturers waste little in processing.

Synthetic paper is experiencing a high growth in many printing segments. In the past, applications for synthetic paper were for outdoor posters and billboards, maps, and manuals for special purposes. Synthetic papers now are being used for products they are replacing, such as cotton, metal, and foil for tags and labels; laminates of all types; artists' materials; and paper and board. Tag and label applications are growing rapidly, with clothiers and industrial end-users accepting synthetic papers' reliability in wear and print performance. Products often include thickness from 0.003 to 0.01 in. (0.08 to 0.25 mm), or basis weights from 40 to 200 lb (18 to 90 kg). Many products are diversifying into cream shades from high-brightness stocks, and most contain antistatic agents to avoid double feeding.

Another large growth area for synthetic paper is lamination, with identity cards, menus, and tickets being target markets. However, the largest applications are still outdoor banners, maps, and manuals. Traditional paper products compete heavily in these markets because of the large volumes available, somewhat short lifecycles, and lower costs. Synthetic paper improvements in print performance and converting (such as die cutting, embossing, and folding) offer graphic arts producers a high-end alternative and a high profit potential.

Finishing Processes

Finishing processes include folding, stitching, book casing, gluing, perfect binding, scoring, perforating, creasing, sheeting, foil stamping, collating, trimming, die cutting, embossing, laminating, punching, numbering, shrink-wrapping, bundling and banding, in-line finishing, demographic finishing, segmentation, and fulfillment, and mailing and distribution. Finishing processes continue to be enhanced by automation and computerization.

Demographic finishing is a process that personalizes a publication or other print medium for individuals or geographic regions. It typically involves inline inkjet printing from a subscription database and bar codes to address publications or direct-mail advertising, eliminating the need for mailing labels. This technology is widespread in the periodical and catalog industry.

Inline finishing involves binding and finishing operations that are integrated with printing presses and take place as continuous operations, including trimming, slitting, die cutting, embossing, gathering, folding, stitching, numbering, punching, and addressing.

There are significant amounts of automation from production, to bundling and wrapping, to delivering jobs, to loading docks. This involves using bar coding and computer-controlled finishing systems. There is automation for setting up finishing equipment and for removing jobs from press and moving them to the bindery.

The ability to feed precollated sets from a digital press or copier and add a preprinted cover to the set is now possible. The unit then finishes the product by stitching, folding, and trimming. Customer databases allow the operator to retrieve, update, delete, and maintain job requirements. *See* DATABASE MANAGEMENT SYSTEM.

There are robotic palletizers, electronic overhead conveying systems, and automatic, programmable stitching/trimming machines. Some presses have inline UV coating units to enable users to print, UV-coat, slit, and rewind or cut digitally printed pieces in a single pass at full press speed.

Digital presses are driving finish-on-demand developments as markets for digital printing and on-demand finishing accelerates. There is also on-demand print-to-mail. Print-to-mail uses stored or real-time data to create personalized communications with recipients. The Internet is driving this. The most sophisticated use is for data streaming from Web sites to drive personalized communications. Servers residing in the mailing facility capture input information from the site. This drives digital printers, which create customized letters and marketing pieces. They are loaded into an inserting machine, which uses a digital camera-based vision system to verify that the personalized printed pieces match each other. The camera system can also use verified bar code information to instruct an inline inkjet printer to address the outer envelope with the matching address and additional personalization. *See* INTERNET; WORLD WIDE WEB.

Bar code readers enable systems to recall preprogrammed jobs and provide fully automatic slitting, cutting, and creasing. A creasing unit ensures that color-printed documents do not crack when folded.

The major digital print vendors have software codes for describing the attributes of a print job, from file handling and correction, through printing, collating, and finishing. Nearly all vendors of the on-demand printing and finishing equipment are implementing CIP4, the open interface standard for all digital-printing hardware components.

Harvey R. Levenson

Bibliography. *Bridg's CTP Handbook for the Graphic Arts*, 2000; M. H. Bruno, *Pocket Pal*, 19th ed., 2004; M. H. Bruno, *Principles of Color Proofing*, 1986; G. G. Field, *Color and Its Reproduction*, 3d ed., 2004; Foundation of Flexographic Technical Foundation, *Flexography: Principles and Practices*, 6 vols., 5th ed., 2000; Gravure Education Foundation and Gravure Association of America, *Gravure: Process and Technology*, 2d ed., 2003; IDEAlliance SWOP, *SWOP: Specification for Web Offset Publications*, 10th ed., 2005; H. R. Levenson, *Complete Dictionary of Graphic Arts and Desktop Publishing Terminology*, 1995; H. R. Levenson, *Understanding Graphic Communication: Selected Readings*, 2000; F. J. Romano, *Digital Basics 3.0*, 2001; F. J. Romano, *Digital Printing: Mastering On-Demand and Variable Data Printing for Profit*, 2000; F. J. Romano and R. M. Romano, *The GATF Encyclopedia of Graphic Communications*, 1998; R. Sasso, *Publishing Timeline 2000*, 2000; L. Wilson, *What the Printer Should Know about Paper*, 3d ed., 1998.

Prion disease

Transmissible spongiform encephalopathies in both humans and animals. Scrapie is the most common form seen in animals, while in humans the most prevalent form is Creutzfeldt-Jakob disease. This group of disorders is characterized at a neuropathological level by vacuolation of the brain's gray matter (spongiform change). They were initially considered to be examples of slow virus infections. Experimental work has consistently failed to demonstrate detectable nucleic acids—both ribonucleic acid (RNA) and deoxyribonucleic acid (DNA)—as constituting part of the infectious agent. Contemporary understanding suggests that the infectious particles are composed predominantly, or perhaps even solely, of protein, and from this concept was derived the acronym prion (proteinaceous infectious particles). Also of interest is the apparent paradox of how these disorders can be simultaneously infectious and yet inherited in an autosomal dominant fashion (from a gene on a chromosome other than a sex chromosome).

Disorders. Scrapie, which occurs naturally in sheep and goats, was the first of the spongiform encephalopathies to be described. Over the years, an increasing range of animal species have been recognized as occasional natural hosts of this type of disease. Examples of these disorders include

transmissible mink encephalopathy, chronic wasting disease of mule deer and elk, and an epidemic form in cattle in the United Kingdom called bovine spongiform encephalopathy. The last was believed to have been introduced by the accidental contamination of the food chain by meat and bone meal protein supplement derived from sheep presumably harboring scrapie. As a reaction, various measures, including a feed ban, were put in place, with recent surveys indicating the bovine spongiform encephalopathy epidemic is on the decline. The British surveillance unit and its detailed case-control study set up to monitor the incidence of Creutzfeldt-Jakob disease has shown no increase in the number of cases of this disorder, and no geographical clustering. However, the recent occurrence of Creutzfeldt-Jakob disease in persons under the age of 42 has raised the possibility that the bovine disease may have spread to humans. At this stage, the overall risk of transmission appears small. So far, animal models have indicated that only central nervous system tissue has been shown to transmit the disease after oral ingestion—a diverse range of other organs, including udder, skeletal muscle, lymph nodes, liver, and buffy coat of blood (that is, white blood cells) proving noninfectious.

The currently recognized spectrum of human disorders encompasses kuru, Creutzfeldt-Jakob disease, Gerstmann-Sträussler-Scheinker disease, and fatal familial insomnia. All, including familial cases, have been shown to be transmissible to animals and hence potentially infectious; all are invariably fatal with no effective treatments currently available.

Kuru has been almost eradicated since the cessation of ritualistic cannibalism, but occurred endemically among the Fore and some neighboring people of the Eastern Highlands of Papua New Guinea. The disorder was believed to be transmitted by the consumption of infected central nervous system tissues. In its typical form, the symptoms of kuru included a relentlessly progressive loss of muscular coordination (ataxia) associated with an unusual tremor; the duration of the entire illness spanned 1–2 years.

Creutzfeldt-Jakob disease is a rapidly progressive deterioration of intellectual abilities (dementia), which proceeds to death with a median duration of illness of around 4 months. As the disease evolves, it is typically accompanied by a range of additional clinical features which include ataxia, muscular rigidity, and spontaneous irregular limb jerks (myoclonus). For 90% of affected individuals Creutzfeldt-Jakob disease develops sporadically, without any clues as to causation. Approximately 10% of individuals have a genetic basis and occur within families. The final important epidemiological subgroup of Creutzfeldt-Jakob disease comprises those cases arising after accidental person-to-person transmission (iatrogenic).

Gerstmann-Sträussler-Scheinker disease is relatively uncommon, and presents as a chronic progressive familial ataxia with somewhat distinctive neuropathological features. Fatal familial insomnia is the rarest of the disorders, and is characterized by rapidly progressive and untreatable insomnia, ataxia, and myoclonus, with pathological examination revealing severe and selective degeneration of a restricted part of the brain (the thalamus). *See* BRAIN.

Molecular biology. Although not completely resolved, the prevailing view is that the transmissible agent of these disorders is an abnormal conformational form of the constitutively expressed prion protein molecule. The prion protein molecule is normally produced by a wide variety of cells throughout the body, but especially the central nervous system. Although initial investigations minimized its probable biological importance, this host-encoded glycoprotein probably plays a role in receptor-mediated signal transduction, and disordered function may be capable of explaining some of the clinical features observed in Creutzfeldt-Jakob disease.

It has been suggested that in sporadic cases of Creutzfeldt-Jakob disease, the change in the prion protein molecule occurs spontaneously as a random or chance event. However, once in this abnormal shape, the prion protein molecule becomes relatively resistant to enzymes usually capable of degrading proteins (proteases), and also serves as a template or nucleating factor for the conversion of other prion protein molecules, through self-replicating polymerization. In some of the prion disorders, especially Gerstmann-Sträussler-Scheinker disease, the accumulation of the abnormal prion protein molecules can be detected as deposits of amyloid, which may also represent a mechanism of neurotoxicity inducing the spongiform change. However, the precise sequence of events subserving the translation of this subtle molecular change into a devastating fatal neurodegenerative disorder is unknown.

The study of the familial prion diseases has led to the recognition of a number of causal mutations within the prion protein gene located on chromosome 20. Each mutation is generally associated with a particular form of clinical illness, but considerable variability in clinical and neuropathological expression can occur within affected members of the same family. These mutations are thought to greatly increase the chance of the prion protein molecule adopting the abnormal configuration, leading to self-replicating polymerization. Similar to sporadic Creutzfeldt-Jakob disease, the consequent accumulation of abnormal prion protein molecules is thought to be linked to the eventual disruption of cellular function. If the abnormal form is introduced into an individual, it serves as the template for subsequent autocatalytic polymerization. Such a model attempts to reconcile the paradox of how this group of disorders can be transmissible and at the same time inherited in an apparent autosomal dominant fashion.

Although no causative mutation has been demonstrated in the prion protein gene to explain sporadic examples of Creutzfeldt-Jakob disease, a genetic component may still play a part in the susceptibility to developing this disorder, with normal variations in the gene possibly altering the likelihood of developing this and pituitary hormone-related Creutzfeldt-Jakob disease.

Human-to-human transmission. A variety of mechanisms of human-to-human transmission have been described, and include contaminated instruments

such as electroencephalographic depth electrodes, dura mater homografts, corneal transplants, and cadaverically derived human pituitary growth hormone and gonadotrophins. Transmission is due in part to the ineffectiveness of conventional sterilization and disinfection procedures to control the infectivity of transmissible spongiform encephalopathies. Numerically, pituitary hormone-related Creutzfeldt-Jakob disease is the most important form of human-to-human transmission of disease. Although kuru and Creutzfeldt-Jakob disease are experimentally transmissible to nonhuman primates through the oral route, successful human-to-human transmission appears to require more invasive methods of inoculation of the infectious material, usually in the form of tissue transplants or injectable therapies or through the use of contaminated neurosurgical instruments.

Epidemiological evidence suggests that there is no increased risk of contracting Creutzfeldt-Jakob disease from exposure in the form of close personal contact during domestic and occupational activities. Nonetheless, a very small number of physicians, including pathologists and neurosurgeons, and a slightly larger number of allied health personnel (including histologists) have developed Creutzfeldt-Jakob disease. However, there is no clear evidence to suggest that occupational exposure poses a significant risk of contracting Creutzfeldt-Jakob disease, especially if precautionary and preventive guidelines are followed.

Incubation periods in cases involving human-to-human transmission appear to vary enormously, depending upon the mechanism of inoculation. In the case of pituitary hormone therapy and kuru, which presumably utilize peripheral methods of inoculation, incubation periods vary considerably from a few to more than 30 years, with infected individuals usually exhibiting an ataxic syndrome. However, when the infectious material is given more direct access to the central nervous system, as with intracerebral depth electrodes, the incubation periods are usually much shorter, generally of the order of 12–24 months, and cognitive decline similar to sporadic Creutzfeldt-Jakob disease is the usual manifestation.

Current evidence suggests that transmission of Creutzfeld-Jakob disease from mother to child does not occur. This finding accords with the much larger epidemiological experience seen in the related disorder kuru. The evidence encompasses children born to mothers during both the presymptomatic and clinical phase of their illness. The explanation for the apparent transplacental and postpartum barriers to transmission remain to be elucidated.

Transmissibility. The transmissibility of human spongiform encephalopathies was first reported in 1966 after kuru had been successfully transmitted to three chimpanzees. Within 2 years, the successful transmission of Creutzfeldt-Jakob disease using the same animal model was announced. Using nonhuman primates (chimpanzees, squirrel and spider monkeys), transmission rates range from 90% for sporadic Creutzfeldt-Jakob disease, through 95% for kuru, and up to 100% for cases of human-to-human transmission of Creutzfeldt-Jakob disease. Two important factors pertaining to transmissibility are the method of inoculation and the dose of infectious material administered. A high does of infectious material administered by direct intracerebral inoculation is clearly the most effective method of transmissibility and generally provides the shortest incubation time, which in the primate model is approximately 10 months. Peripheral methods of inoculation (for example, subcutaneous, intraperitoneal) are less effective routes for transmissibility which can be partly offset by the administration of higher doses of infectious material.

Smaller nonprimate host animals have been successfully utilized in transmissibility experiments. However, transmissibility rates are generally far lower, despite the use of similar inoculation techniques. Familial forms of disease are less transmissible, averaging around 70%. Within the familial group, the transmission rate also appears dependent on the underlying mutation in the prion gene.

There is some evidence of a species barrier to transmissibility. The prion protein gene appears to modulate both scrapie susceptibility and incubation times, and the more dissimilar the prion protein genes, the greater the species resistance to transmissibility.

Central nervous system tissue has been consistently shown to be the most infectious material in animal transmissibility experiments. However, other tissues have been shown to be infectious, although at lower rates of transmissibility, and include lung, liver, kidney, spleen, and lymph nodes. Although cerebrospinal fluid has been shown to be infectious, vaginal secretions, semen, sputum, saliva, nasal mucus, and tears are not. The testis, prostate, skeletal muscle, peripheral never, adipose tissue, adrenal gland, and thyroid gland have failed to transmit the disease; the status of urine and placental tissue is less clear.

The transmissibility status of blood and blood products is of some concern. White blood cells (buffy coat) and whole blood have been shown to be infectious and transmissible to nonprimate animals when taken from individuals during late stages of Creutzfeldt-Jakob disease. Clearly such findings raise some concern regarding the possible hematogenous transmission of disorders from individuals in the preclinical or clinical phases of their illness, but despite claims to the contrary there is no well-documented and verified example of a Creutzfeldt-Jakob disease being transmitted in this manner in humans.

See MUTATION; NERVOUS SYSTEM DISORDERS; SCRAPIE; VIRUS INFECTION, LATENT, PERSISTENT, SLOW.

Colin L. Masters; Steven J. Collins

Bibliography. P. Brown, M. A. Preece and R. G. Will, "Friendly fire'' in medicine: Hormones, homografts, and Creutzfeldt-Jakob disease, *Lancet*, 340:24–27, 1992; D. M. Knipe et al. (eds.), *Fields' Virology*, 4th ed., 2001; M. S. Palmer and J. Collinge, Human prion diseases, *Curr. Opin. Neurol. Neurosurg.*,

5:895–901, 1992; S. B. Prusiner, Natural and experimental prion diseases of humans and animals, *Cur. Opin. Neurobiol.*, 2:638–647, 1992.

Prism

A polyhedron of which two faces are congruent polygons in parallel planes, and the other faces are parallelograms (see **illus**.). The bases B are the congruent

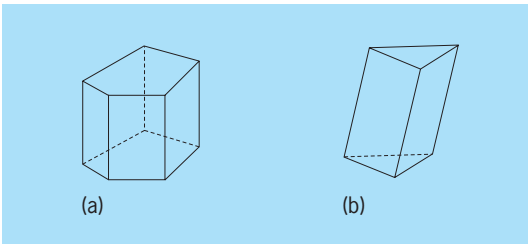

Prisms configurations. (*a*) Right. (*b*) Oblique.

polygons; the lateral faces are the parallelograms; the lateral edges are the edges not lying in the bases; and the perpendicular distance between the bases is the altitude l. Sections parallel to the bases are congruent to the bases. A prism is a right prism if its lateral edges are perpendicular to the bases; an oblique prism otherwise. A prism is called a triangular prism if its bases are triangles; a pentagonal prism if its bases are pentagons; and a parallelepiped if its bases are parallelograms. The volume of any prism is equal to the area of its base times its altitude ($V = Bh$). *See* POLYHEDRON. J. Sutherland Frame

Prismatic astrolabe

A surveying instrument used to make the celestial observations needed in establishing an astronomical position. The instrument (**Fig. 1**) consists of an accurate prism, a small pan of mercury to serve as an

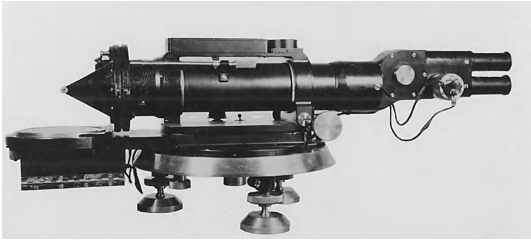

Fig. 1. A prismatic astrolabe, used to make celestial observations. (*U.S. Naval Oceanographic Office*)

artificial horizon, an observing telescope with two eyepieces of different power, level bubbles and leveling screws, a magnetic compass and azimuth circle, adjusting screws, flashlight-battery power source, light, and a rheostat to control the intensity of illumination.

By using a fixed prism, the instrument measures a fixed altitude, usually 45°. As shown in **Fig. 2**, a light ray R from a star enters the upper face of the prism and is refracted so as to pass down the tube of the telescope to the eyepiece. At the same time a parallel ray R' from the same star is reflected from the mercury horizon to the lower face of the prism and is also refracted down the telescope tube.

As a rising star increases altitude past that for which the instrument was constructed, the direct image appears to move upward from the bottom of the field of vision to the top. The image reflected by the mercury horizon appears to move downward from top to bottom. Thus, shortly before the rising star reaches the fixed altitude, the rays (a) and (a') produce images at the bottom and top, respectively, of the field of view as shown in inset (1). As the star continues to rise, the two images approach each other. At the established altitude the rays (b) and (b') both produce images at the center of the field of view, as shown in inset (2). As the star continues to rise, the two images continue their apparent movement, again separating. For a setting body

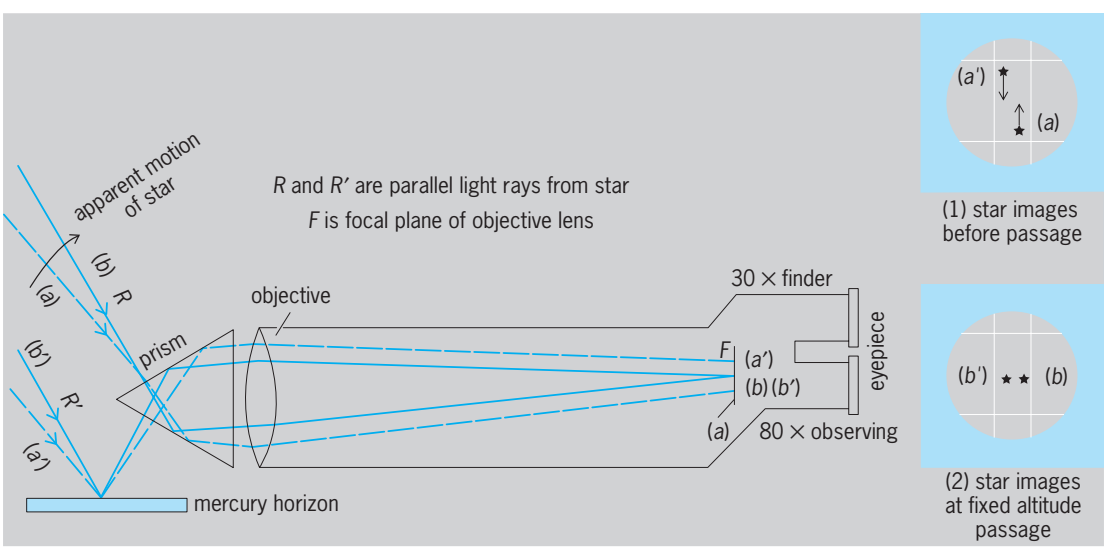

Fig. 2. Optics of a prismatic astrolabe. (*U.S. Naval Oceanographic Office*)

the apparent motion is in the opposite direction. The prism, when installed, is rotated slightly about the telescope axis so that the two images will be close together on a horizontal line when the fixed altitude is reached, as shown, rather than coincident.

A fixed altitude is used to minimize error that is due to variations from standard atmospheric refraction. Each accurately timed observation provides one line of position. The established position is at the center of the inscribed circle of the figure formed by a number of lines of position. An increase in the number of observations decreases the random error of the position.

An "astronomical position" determined by a prismatic astrolabe is subject to a small but sometimes serious error due to deflection of the vertical by gravity anomalies in the Earth. Alton B Moody

Pristiformes

An order of batoid fishes comprising one family (Pristidae), two genera, and seven species and commonly known as sawfishes. Sawfishes are superficially similar to sawsharks (order Pristiophoriformes) in having an elongate sharklike body, sharklike locomotion, two distinct dorsal fins, a distinct caudal fin, and a long bladelike rostrum armed on both sides with large teeth but their fundamental structure is that of other batoids. Each rostral tooth is an enlarged dermal denticle homologous to those of the sawsharks and the dorsal spines of other sharks. *See* BATOIDEA; ELASMOBRANCHII; PRISTIOPHORIFORMES.

Sawfishes are found in tropical and subtropical shallow coastal waters and occasionally enter freshwater. The saw is used to grub in the sand and mud for bottom-dwelling invertebrates. Also, the toothed blade is used to slash to and fro in schools of small fish, stunning or killing them. The maximum length of the smallest species is 1.4 m (4.6 ft); the largest is 7.6 m (25 ft), possibly more, placing them among the largest elasmobranchs. Regardless of their size and their possession of a formidable weapon, unprovoked attacks upon bathers are highly unlikely.

All species of the family are listed as critically endangered on the International Union for the Conservation of Nature and Resources (IUCN) Red List of Threatened Animals. The principal threat to all sawfishes is fisheries, both targeted and bycatch (unintentional capture by fishing gear), because their long tooth-studded saw makes them extraordinarily vulnerable to entanglement in any sort of net gear. *See* ENDANGERED SPECIES. Herbert Boschung

Sawfish (*Pristis microdon*). (*Photo from Aquarium of the Pacific, Long Beach, California*)

Bibliography. L. J. V. Compagno, Checklist of Chondrichthyes, pp. 503–547, in W. C. Hamlett (ed.), *Reproductive Biology and Phylogeny of Chondrichthyes: Sharks, Batoids and Chimaeras*, Science Publishers, Enfield, NH, 2005; L. J. V. Compagno, Checklist of living elasmobranches, pp. 471–498, in W. C. Hamlett (ed.), *Sharks, Skates, and Rays: The Biology of Elasmobranch Fishes*, Johns Hopkins University Press, Baltimore, 1999; L. J. V. Compagno and P. R. Last, Pristidae, Sawfishes, pp. 1410–1417 in K. E. Carpenter and V. Niem (eds.), *FAO Identification Guide for Fishery Purposes: The Living Marine Resources of the Western Central Pacific*, FAO, Rome, 1999; M. R. de Carvalho and J. D. McEachran, Family Pristidae (sawfishes), pp. 17–21, in R. E. Reis, S. O. Kullander, and C. J. Ferraris, Jr. (eds.), *Checklist of the Freshwater Fishes of South and Central America*, EDIPUCRS, Porto Alegre, Brasil, 2003; R. Froese and D. Pauly (eds.), FishBase, World Wide Web electronic publication, <www.fishbase.org>, version (05/2005); J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Pristiophoriformes

An order of sharks comprising one family (Pristiophoridae) and two genera, commonly known as sawsharks. Sawsharks (see **illustration**) are superficially similar to sawfishes, which are rays, but are distinguishable from all other elasmobranchs by the following combination of characters: snout shaped as a long flat blade bearing alternate large and small teeth on each side that are weakly embedded and periodically replaced; one pair of long barbels on the underside of the blade; no visible dorsal fin spines; anal fin absent; large spiracles; and ovoviviparous, with embryos feeding solely on yolk. *See* ELASMOBRANCHII; PRISTIFORMES; SELACHII.

Sawsharks occur on continental and insular shelves and slopes in temperate and tropical waters of the western Atlantic (around the Bahamas, Florida, and Cuba), the southwestern Indian Ocean off South Africa, and the western Pacific from southern Australia to Japan, rarely entering estuaries.

The genus *Pliotrema* (one species) has 6 pairs of gill slits, serrated larger rostral teeth and teeth with prominent transverse ridges on basal ledges, and a maximum length of 170 cm (67 in.). *Pristiophorus* (four described species) has 5 pairs of gill slits, smooth larger rostral teeth and teeth without prominent transverse ridges on basal ledges, and a maximum length of 80 to 137 cm (32 to 54 in.),

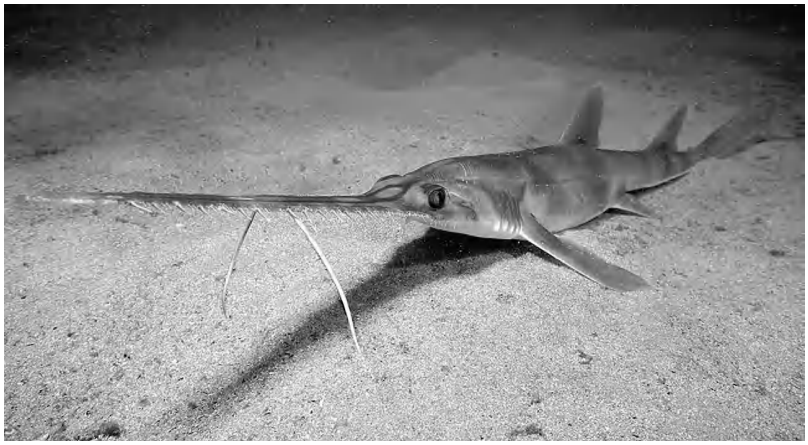

Longnose sawshark (*Pristiophorus cirratus*). (*Photo by Marty Snyderman, www.martysnyderman.com*)

depending on the species. Sawsharks cruise the bottom feeding on bony fishes, shrimps, and squids. Herbert Boschung

Bibliography. L. J. V. Compagno, Checklist of Chondrichthyes, pp. 503–547 in W. C. Hamlett (ed.), *Reproductive Biology and Phylogeny of Chondrichthyes: Sharks, Batoids and Chimaeras*, Science Publishers, Enfield, NH, 2005; L. J. V. Compagno, Checklist of living elasmobranches, pp. 471–498, in W. C. Hamlett (ed.), *Sharks, Skates, and Rays: The Biology of Elasmobranch Fishes*, Johns Hopkins University Press, Baltimore, 1999; L. J. V. Compagno, FAO species catalogue, in *Sharks of the World: An Annotated and Illustrated Catalogue of Shark Species Known to Date–Part 1: Hexanchiformes to Lamniformes*, FAO Fish. Synop. (125) Vol. 4, pt. 1, pp. 212–249, 1984; R. Froese and D. Pauly (eds.), FishBase, World Wide Web electronic publication, www.fishbase.org, version (05/2005); J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Private branch exchange

A customer premises communications system that provides its station subscribers with a wide variety of telephony services and access to public switched telephone network (PSTN) trunk circuits provided by local exchange carrier (LEC) and interexchange carrier (IEC) services. Standard private branch exchange (PBX) system hardware elements include the common equipment (cabinets, shelf carriers, printed circuit boards) and telephone instruments. Peripheral servers may also be included in the system design configuration to support systems management and voice messaging requirements. A PBX system is privately owned and physically located at the user subscriber's premises. Telephony services provided by a PBX range from basic communications operations (provisioning of dial tone and switched connections between peripheral endpoints) to station user-controlled call processing features (such as call transfer, call forwarding, and call waiting) to advanced software applications (such as multisystem networking and call contact centers).

A defining attribute of a PBX system is that trunk circuits are segmented into customer-defined groups for incoming and outgoing call requirements, and allocated to station subscribers as pooled resources. Specific trunk routes and individual circuits are assigned on a call-by-call basis based on the PBX's automatic route selection (ARS) feature as programmed by the system administrator. In contrast, trunk circuits are not pooled in a small-premises communications system designated a key telephone system (KTS). KTS station subscriber telephone instruments have designated key buttons that are used to dial and answer off-premises calls. PBX subscribers have no control over the trunk circuit they use; KTS subscribers can select an available outside-line circuit for placing off-premises calls by depressing a specific key on their telephone instrument. *See* KEY TELEPHONE SYSTEM.

Elements. There are several architecture design elements common to all PBXs, regardless of underlying system technology (analog or digital transmission, circuit-switched or packet-switched): common control elements, the switching network, service circuit boards, and port interface circuit cards.

Common control elements include, at minimum, a call processor and memory storage for generic feature software and the customer database. A PBX common controller may be a proprietary carrier housed in a proprietary equipment cabinet, a proprietary server cabinet or server blade, or a customer-provided server conforming to manufacturer-recommended specifications.

The switching network used to connect PBX peripheral endpoints may be a circuit-switching system (center stage switch, local transmission buses) integrated into proprietary common equipment cabinets or an external Ethernet LAN (local-area network) packet-switched network. *See* DATA COMMUNICATIONS; LOCAL-AREA NETWORKS; PACKET SWITCHING.

Service circuits include tone senders (such as dial and busy), dual-tone multiple-frequency (DTMF)

detectors for push-button analog telephones, registers to temporarily store dialed digits, and input/output (I/O) interfaces for peripheral devices such as cathode-ray-tube (CRT) terminals and printers. *See* COMPUTER PERIPHERAL DEVICES; TELEPHONE.

Port interface circuit cards are required to support station equipment [analog, digital, ISDN BRI (Integrated Services Digital Network Basic Rate Interface), IP (Internet Protocol)] and off-premises trunk circuits [GS/LS (Ground Start/Loop Start), DID (Direct Inward Dialing), E&M (Ear and Mouth), T1-interface, ISDN BRI/PRI (Primary Rate Interface), Auxiliary, FX (Foreign Exchange)]. Distinct port circuit cards are usually required for each type of station or trunk port. *See* INTEGRATED SERVICES DIGITAL NETWORK (ISDN).

System designs. PBX systems are currently available in a wide range of system designs based on several call processing, switching, and hardware equipment platforms, and individual models can be cost-effectively designed to support relatively small port capacity requirements (less than 20 station subscribers) or extremely large configuration needs (tens of thousands of station subscribers). PBX systems compete with medium-line-size KTS-hybrid systems, local telephone company Centrex services, and third-party-hosted service offerings, but are the dominant system for enterprise customers with port size requirements above 80 stations.

Evolution of digital PBX. Since the 1970s, PBX system technology has evolved through several distinct phases as mechanical analog telephone switching systems were transformed by advances in semiconductor technology and computer software programming. The two major PBX technology breakthroughs during the 1970s were computer-stored program control and digital circuit switching using time division multiplex (TDM) transmission buses for internal system transport between port circuit terminations. There were competing techniques used to digitize voice, but pulse code modulation (PCM) using 8-kilohertz sampling and 8-bit encoding emerged as the standard, and remains so for circuit-switched communications systems. Another important PBX system innovation was electronic telephone instruments with programmable line-feature buttons. *See* MULTIPLEXING AND MULTIPLE ACCESS; PULSE MODULATION; SWITCHING SYSTEMS (COMMUNICATIONS).

Important PBX system design advances during the 1980s included stackable modular cabinets, dispersed processing, and digital telephone instruments with data terminals interfaces. This decade also saw the introduction of many evolutionary PBX feature and function options, including intelligent transparent private networking, integrated automatic call distribution (ACD) and reporting, voice messaging system interface, and open application interface (OAI).

Since the mid-1990s, the most important PBX technology innovation has been the integration of Internet Protocol (IP) communications standards into the core system design. The first IP telephony system designs were characterized by an Ethernet LAN-connected call telephony server, IP telephone instruments, and media gateways for analog telephone instruments and PSTN trunk circuit interface requirements. This type of PBX system is now referred to as a client-server IP telephony system. An IP telephony system utilizes Ethernet packet-switched networks for premises transport of call control or voice communications signals, and IP WAN (wide-area network) routers for off-premises transmission

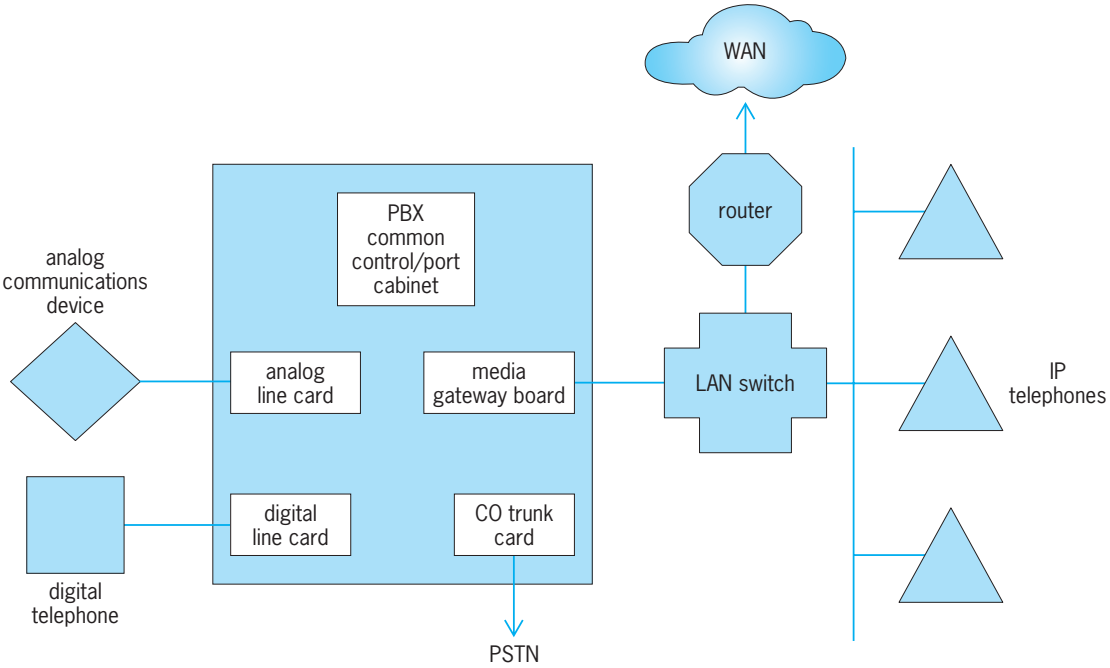

Fig. 1. IP-enabled PBX system design. PBX = private branch exchange, CO = central office, PSTN = public switched telephone network, WAN = wide-area network, LAN = local-area network, IP = Internet Protocol.

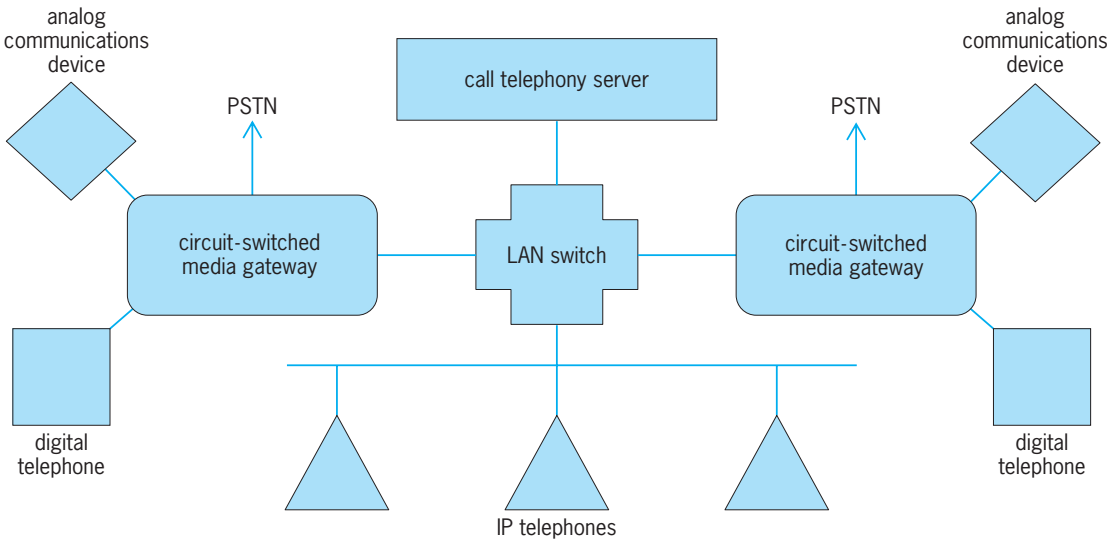

Fig. 2. Converged IP-PBX system design.

requirements. Most of the early IP telephony systems were based on ITU H.323 communications protocol specifications, but the evolving standard that all systems will most likely support is the Internet Engineering Task Force (IETF) Session Initiated Protocol (SIP) specifications, particularly core standards such as RFC (Request for Comments) 3261, RFC 1889, and RFC 2833. *See* INTERNET; WIDE-AREA NETWORKS.

IP telephony PBX system designs. There are three basic IP telephony PBX system designs: IP-enabled, converged, and client-server (sometimes referred to as softswitch).

The following are definitions of each design category:

1. IP-enabled PBXs are communications systems originally based on a circuit-switched network design that are upgraded to support IP telephony requirements using fully integrated media gateway port interface circuit cards to support IP station equipment and/or voice over IP (VoIP) trunk circuits (**Fig. 1**). *See* VOICE OVER IP.

2. Converged IP-PBXs are communications systems based on either a traditional common control cabinet or a LAN-connected call telephony server. The common control cabinet or call telephony server can support distributed circuit-switched common equipment cabinets across a LAN-WAN infrastructure. IP peripheral endpoints can be supported using direct call signaling from the primary call control processor or through integrated media gateway interface boards (**Fig. 2**).

3. Client-server IP-PBXs are communications system based on a LAN-connected call telephony server with direct call control signaling to IP telephone instruments. Media gateway equipment is required to support traditional TDM-PCM endpoints, such as analog telephone instruments or circuit-switched PSTN trunk circuits (**Fig. 3**).

IP-enabled PBXs share the following system design attributes:

1. System architecture based on a traditional circuit-switched system design platform.

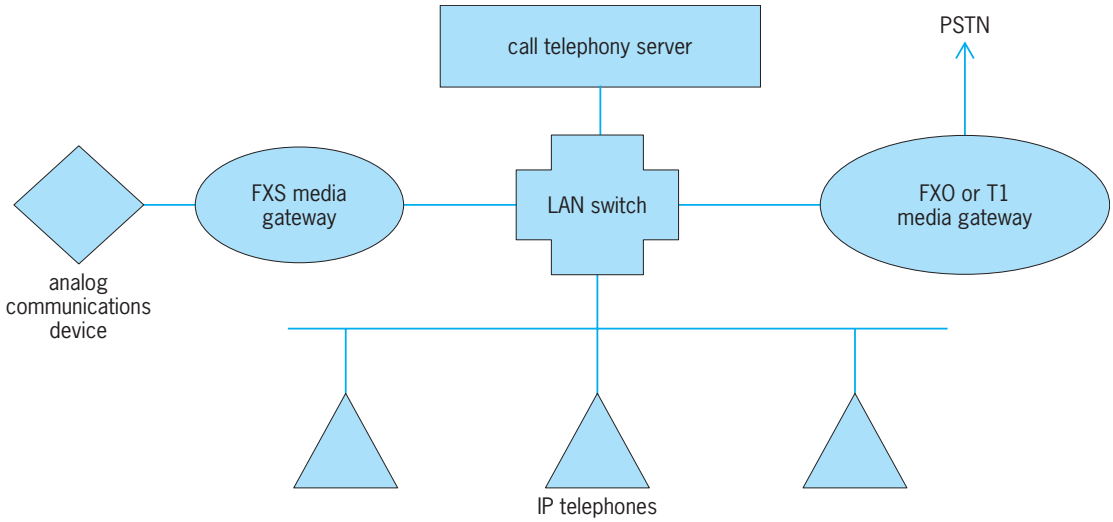

Fig. 3. Client-server IP-PBX system design. FXS = foreign exchange subscriber, FXO = foreign exchange office.

2. A proprietary common control complex housed in a dedicated control cabinet or carrier shelf.

3. Port cabinets or carriers with port interface card slots that physically connect port circuit card terminations to integrated TDM transmission bus backplanes.

4. Integrated media gateway boards that support LAN-connected IP voice terminals.

5. Integrated media gateway boards that support VoIP trunk calls.

Converged IP-PBXs share the following system design attributes:

1. Common control may be a proprietary traditional common control complex housed in a circuit-switched common-equipment cabinet or carrier, a self-standing call telephony server; or a proprietary call processor server blade.

2. The common control supports distributed LAN-WAN-connected TDM-based port cabinets or carriers for traditional circuit-switched analog and digital port interface requirements.

3. LAN-distributed circuit-switched cabinets or carriers require media gateway boards to communicate with each other and provide transmission links between TDM-PCM ports and IP peripheral endpoints.

Call control signaling over the LAN between the system common control and peripheral IP endpoints may require an intermediary media gateway or control signaling board if a traditional common control design is implemented. Server designs usually have direct signaling links to the LAN and all peripheral ports.

Client-server IP-PBXs share the following system design attributes:

1. A call telephony server as a common control complex.

2. Direct call control signaling over LAN-WAN to all IP peripheral endpoints and distributed media gateways.

3. Media gateways used for passing call control signaling and providing transmission channel connections to all non-IP ports.

Media gateways may be embedded in a proprietary call telephony server, or be provisioned as desktop modules, self-standing carriers, or circuit boards housed in Ethernet LAN switches or IP WAN routers.

Generic software features. The PBX system generic software includes hundreds of basic and advanced telephony features and system diagnostic and administration functions. Basic system features include provisioning of dial tone, a uniform numbering and dialing plan, and call routing operations under the control of the customized ARS program. The most widely used and well-known station user features are hold, transfer, forward, conference, call release, and last number redial. Several enhanced applications services can be integrated into the PBX system generic software package, including voice messaging, automatic call distribution (ACD), and networking. IP communications options that can be integrated into the PBX system offering or run on peripheral application servers include presence management, conferencing (audio, video, and Worldwide Web), and collaboration (file sharing and whiteboarding, the placement of shared files on an on-screen "shared notebook" or "whiteboard").

Voice terminals. There are a variety of voice terminal options available to PBX station subscribers: analog, digital, and IP desktop telephone instruments; personal computer (PC) client softphones; and mobile communications devices (cellular handsets, wireless LAN handsets, and personal digital assistants, or PDAs). Attendant operator positions require specialized console instruments or softphone clients to satisfy high call volume and complex call processing operations. *See* TELEPHONE SERVICE.

Allan Sulkin

Bibliography. K. Camp, *IP Telephony Demystified*, McGraw-Hill, 2003; A. Sulkin, IP PBX market chugging along, *Bus. Commun. Rev.*, 36(1):21–29, January 2006; A. Sulkin, IP-telephony enhances business continuity, *Bus. Commun. Rev.*, 36(8):24–26, August 2006; A. Sulkin, *PBX Systems for IP Telephony*, McGraw-Hill, 2002.

Probability

Although probability theory derives its notion and terminology from intuition, a vague statement such as "John will probably come" is as remote from it as the statement "John is forceful and energetic" is remote from mechanics. Probability theory constructs abstract models, mostly of a qualitative nature, and only experience can show whether these reasonably describe laws of nature or life. As always in mathematics, only logical relations and implications enter the theory, and the notion of probability is just as undefinable (and as intuitive) as are the notions of point, line, or mass. An actual assignment of numerical probabilities is frequently unnecessary or impossible. For example, telephone exchanges are based on a theoretical comparison of several possible systems; only the optimal ones are built and the others discarded. Thus a huge industry depends on theoretical models of exchanges which will never exist.

An uncomplicated illustration of the nature of probability models is found in an experiment of Lord Rutherford. To measure radioactive intensity, he proceeded as follows. Observers A_1 and A_2 counted scintillations on a screen and observed, respectively, N_1 and N_2 scintillations; of these, N_{12} were common to both observers. To estimate the unknown true number X, Rutherford assumed that each scintillation has fixed probabilities p_1 and p_2 to be observed by A_1 and A_2, and furthermore that the observations are independent in the sense that a scintillation observed by A_1 has still probability p_2 to be observed by A_2. In reality the likelihood of observing a scintillation varies with growing fatigue and the proximity of the preceding scintillation; also, the observers are affected by common causes and are therefore not independent. Equating probabilities with observed frequencies (another approximation), Rutherford set $N_1 = Xp_1$, $N_2 = Xp_2$, $N_{12} = Xp_1p_2$, whence $X = N_1N_2/N_{12}$.

The three equations may be solved for p_1 and p_2, but these "probabilities" are purely fictitious and, as experience shows, inaccessible to experimental verification. The model is justified by plausibility and success.

Sample space. One speaks of probabilities only in connection with conceptual (not necessarily performable) experiments and must first define the possible outcomes. Thus, by convention, tossing coins results in heads H or tails T; regardless of experimental or philosophical difficulties the age of a person is taken as an exact number and each positive number is taken as a possible age. Throwing two dice results in one of the 36 combinations (1.1), (1.2), . . . , (6.6). An outcome such as "sum 4" is a compound event which can be further decomposed by enumeration: Sum 4 occurs if the outcome is (1.3), (2.2), or (3.1). Thus it is necessary to distinguish between elementary (indivisible) and compound outcomes or events. Each elementary outcome is called sample point; their aggregate is the sample space. The conceptional experiment is defined by the sample space, and it must be introduced and established at the outset.

For example, the experiment "distributing 3 balls in 3 cells" has 27 possible outcomes (sample points) listed in tabulation (1).

$$\begin{array}{lll}
1.\ \{abc|-|-\} & 10.\ \{a|bc|-\} & 19.\ \{-|a|bc\} \\
2.\ \{-|abc|-\} & 11.\ \{b|ac|-\} & 20.\ \{-|b|ac\} \\
3.\ \{-|-|abc\} & 12.\ \{c|ab|-\} & 21.\ \{-|c|ab\} \\
4.\ \{ab|c|-\} & 13.\ \{a|-|bc\} & 22.\ \{a|b|c\} \\
5.\ \{ac|b|-\} & 14.\ \{b|-|ac\} & 23.\ \{a|c|b\} \\
6.\ \{bc|a|-\} & 15.\ \{c|-|ab\} & 24.\ \{b|a|c\} \\
7.\ \{ab|-|c\} & 16.\ \{-|ab|c\} & 25.\ \{b|c|a\} \\
8.\ \{ac|-|b\} & 17.\ \{-|ac|b\} & 26.\ \{c|a|b\} \\
9.\ \{bc|-|a\} & 18.\ \{-|bc|a\} & 26.\ \{c|b|a\}
\end{array} \qquad (1)$$

Note that "n balls in 7 cells" may represent the distribution of n hits among 7 targets, or of n accidents in 7 weekdays, and so on.

Consider next the experiment of placing 3 indistinguishable balls into 3 cells. Whether actual balls are indistinguishable is irrelevant; they are treated as such and, by convention, there is now a space of only 10 sample points. It is listed in tabulation (2).

$$\begin{array}{lll}
1.\ \{***|-|-\} & 4.\ \{**|*|-\} & 8.\ \{-|**|*\} \\
2.\ \{-|***|-\} & 5.\ \{**|-|*\} & 9.\ \{-|*|**\} \\
3.\ \{-|-|***\} & 6.\ \{*|**|-\} & 10.\ \{*|*|*\} \\
 & 7.\ \{*|-|**\} &
\end{array} \qquad (2)$$

In playing roulette, each point on a circle represents a possible outcome and the sample space is the interval $0 \leqq \vartheta < 2\pi$. When one observes the motion of a particle under diffusion, every function $x(t)$ represents a conceivable outcome and the sample space is a complicated function space.

Events. In examining a bridge hand, one may ask whether it contains an ace or satisfies some other condition. In principle each such event may be described by specifying the sample points which do satisfy the stipulated condition. Thus every compound event is represented by an aggregate of sample points, and in probability theory these terms are synonymous. The standard notations of set theory are used to describe relations among events. *See* SET THEORY.

Given an event A one may consider the case that A does not occur. This is the negation or complement of A, denoted by A'; it consists of those sample points that do not belong to A. Given two events A and B, the event C that either A or B or both occur is the union of A and B and denoted by $C = A \cup B$. In particular $A \cup A'$ is the whole sample space $\mathfrak{S}$ which therefore represents certainty. The event D, both A and B occur, is the intersection of A and B and written $D = A \cap B$. It consists of the points common to A and B. If there are no such common points (as in the case of A and A'), A and B cannot occur simultaneously and they are called mutually exclusive, written $A \cap B = 0$. The event "A but not B" is simply $A \cap B'$.

For example, in tabulation (1), the event A "one cell multiply occupied" is the aggregate of the points numbered 1–21. The event B "first cell not empty" is the aggregate of the points 1, 4–15, and 22–27. Because every point belongs either to A or to B (or both). $A \cup B = \mathfrak{S}$ is the certain event. Next, $D = A \cap B$ consists of the points 1, 4–14. Finally, A' may be described as "no cell empty."

Probabilities in finite spaces. If the sample space $\mathfrak{S}$ contains only N points $E_1, \ldots, E_N$ their probabilities may be any numbers such that $P\{E_j\} = 0$ and $P\{E_1\} + \cdots + P\{E_N\} = 1$. The probability $P\{A\}$ of an event A is the sum of the probabilities of all points contained in A; thus $P\{\mathfrak{S}\} = 1$. To find $P\{A \cup B\}$ one considers all points belonging to either A or B, but those belonging to both A and B are counted only once. Therefore $P\{A \cup B\} = P\{A\} + P\{B\} - P\{A \cap B\}$. In particular, for mutually exclusive events, there is the addition rule $P\{A \cup B\} = P\{A\} + P\{B\}$.

Frequently considerations of symmetry lead one to consider all E_j as equally likely; that is, to set $P(E_j) = 1/N$. In this case $P(A) = n/N$ where n is the number of points in A; for a gambler betting on A, these represent the "favorable cases." For example, in throwing a pair of "perfect" dice, one naturally assumes that the 36 possible outcomes are equally likely. This model does not lose its justification or usefulness by the fact that actual dice do not live up to it. The assumption of perfect randomness in games, card shuffling, industrial quality control, or sampling is rarely realized, and the true usefulness of the model stems from the experience that noticeable departures from the ideal scheme lead to the detection of assignable causes and thus to theoretical or experimental improvements.

How the success of probability theory depends on the disregard of preconceived philosophical ideas

and on the readiness to adapt models to unexpected circumstances is illustrated by Bose-Einstein statistics. In the example of tabulation (1), the notion of perfect randomness leads to the assignment of probability 1/27 to each point. In the case of indistinguishable balls, tabulation (2), it has been argued that an experiment is unaffected by failure to distinguish between balls; physically there remain 27 possibilities grouped in 10 distinguishable forms. This argument leads to assigning probability 1/27 to each of the points 1–3, probability 1/9 to each of the points 4–9, and 2/9 to point 10. This reasoning (sound in certain situations) has been accepted as evident in statistical mechanics for the distribution of r particles in n cells (Maxwell-Boltzmann statistics). Surprisingly, it turned out that no physical particles behave this way and it was revolutionary when Bose and Einstein showed that for one type of particle all distinguishable arrangements are equally likely. This model assigns probability 1/10 to each point of tabulation (2). *See* BOLTZMANN STATISTICS; BOSE-EINSTEIN STATISTICS.

A useful, although vague, intuitive description of probability describes $P\{A\}$ as the relative frequency of the event A if the experiment is repeated many times under identical circumstances. The laws of large numbers render this more precise, but the description often lacks operational meaning. Experiments in agriculture and human sampling cannot be repeated under remotely similar conditions, and in the case of telephone exchanges, useful probability models refer to situations which will never materialize.

Probabilities in infinite spaces. Two examples, unending coin tossing and roulette, illustrate the novel features of this topic.

Unending coin tossing. In the study of limit laws, one must consider potentially infinite sequences of coin tossings. The possible outcome of this experiment is an infinite sequence of heads and tails, and every sequence such as *HTHTTH* ... represents a simple point. Finitely many tosses are the beginning of an infinite sequence, and the event "first four trials resulted in *HTTH*" is the aggregate of the infinitely many sequences with the prescribed beginning. Such an event is called an interval of length 4. There are 2^n intervals of length n, and they are mutually exclusive. For reasons of symmetry, one attributes the probability 2^{-n} to each interval of length n. Thus the assignment of basic probabilities refers to intervals rather than to points. A point such as *HTHT* ... is the limit of an infinite sequence of contracting intervals *H*, *HT*, *HTH*, ..., and therefore probability zero must be attributed to each individual point.

The probabilities of other events are similarly defined by limiting procedures. For example, consider the event A that an infinitely prolonged sequence of trials never produces a run of at least two consecutive heads, or two consecutive tails. It is more convenient to enumerate the points of the complementary event A' that two equal symbols do occur in succession. Clearly A' is the union of the infinitely many mutually exclusive intervals *HH*, *TT*; *HTT*, *THH*; *HTHH*, *THTT*, and so on. Here there are 2 intervals of length $n = 2$, and therefore $P\{A'\} = 2(2^{-2} + 2^{-3} + 2^{-4} + \cdots) = 1$, whence $P\{A'\} = 0$. The indicated result of the experiment is thinkable, but probability zero is attributed to it. A similar, although more complicated, limiting procedure leads to the law of large numbers according to which the event "the frequencies of H and T in the first n trials tend to 1/2 as $n \to \infty$" has probability one.

Roulette. Here the sample space consists of the angles $0 \leqq \vartheta < 2\pi$, and the notion of a perfect roulette assumes equal probabilities for intervals of equal length; thus an interval of length a carries probability $a/2\pi$. If the roulette is divided into 32 equal numbered intervals, the event "even number" consists of 16 intervals and has probability 1/2.

The situation encountered here is not peculiar to probability but is common in measure theory. One starts with a collection of basic events, called intervals, and attributes probabilities to them. By simple and natural limiting procedures, probabilities can then be defined for a much wider class $\mathfrak{F}$ of events which are obtainable by applying the operations of set theory to intervals (in finite or infinite numbers). F is the Borel field generated by the intervals. Probability is simply a measure on $\mathfrak{F}$; that is, to each event A in $\mathfrak{F}$, there corresponds a probability $P\{A\} = 0$ which is completely additive. If A is the union of the mutually exclusive events $A_1, A_2, \ldots$, then $P\{A\} = \Sigma P\{A_i\}$. The probability of the whole space is, of course, unity.

The extension of the addition rule from finitely to infinitely many summands may be defended by considerations of continuity, but ultimately it is justified by its simplicity and its success.

Conditional probability—independence. Suppose that a population of N people includes N_A color-blind persons and N_H females. To the event A "a randomly chosen person is color-blind" can be ascribed probability $P\{A\} = N_A/N$, and similarly for the event H that a person be female one has $P\{B\} = N_H/N$. If N_{AH} is the number of color-blind females, the ratio N_{AH}/N_H may be interpreted as probability that a randomly chosen female be color-blind; here the experiment "random choice in the population" is replaced by a selection from the female subpopulation. In the original experiment, N_{AH}/N is the probability of the simultaneous occurrence of both A and H, so that $N_{AH}/N_H = P\{A \cap H\}/P\{H\}$. Similar situations occur so frequently that it is convenient to introduce Eq. (3) and to call this the conditional probability

$$P\{A|H\} = \frac{P\{A \cap H\}}{P\{H\}} \tag{3}$$

of the event A relative to H. This concept is useful whenever it is desired to restrict the consideration to those cases where the event H occurs (or where the hypothesis H is fulfilled). Thus, in betting on an event A the knowledge that H occurred would induce one to replace $P\{A\}$ by $P\{A|H\}$. If all sample points are equally likely, $P\{A|H\}$ still represents the

ratio of favorable cases to the total of cases possible when it is known that H has occurred.

Despite its simplicity the notion of conditional probability is exceedingly important, and frequently the probabilities in sample space are defined only in terms of conditional probabilities. An example of the concept is provided by a bolt factory in which three machines manufacture, respectively, 25, 35, and 40% of the total. Of their output 5, 4, and 2% are defective bolts. Classification of the bolts according to the number of the machine and the quality (d for defective, c for conforming) gives the six categories c_1, c_2, c_3, d_1, d_2, and d_3. A random choice of a bolt results in one of these six outcomes, but their probabilities are not given directly. Instead, the data relating to the first machine are given by

$$P\{c_1 \cup d_1\} = 0.25$$
$$P\{d_1 | c_1 \cup d_1\} = 0.25$$

It follows that $P\{d_1\} = 0.0125$ and similarly for the other points. This example may also serve to illustrate the reasoning following Bayes concerning the probability of causes. Supposing a bolt was found to be defective (hypothesis H), what is the probability that it came from the first machine (cause A)? Here

$$\begin{aligned} P\{H\} &= P\{d_1\} + P\{d_2\} + P\{d_3\} \\ &= 0.0125 + 0.0140 + 0.0080 = 0.0345 \\ P\{A \cap H\} &= P\{D_1\} = 0.0125 \end{aligned}$$

and thus the required answer is given by

$$P\{A|H\} = 0.0125/0.0345 = 25/69$$

In tabulation (1) the probability of H "ball a is in the first cell" equals 1/3, and the probability of A "first cell is multiply occupied" $= 7/27$. Now given that the ball a is in the first cell, the conditional probability that this cell is multiply occupied becomes 5/9. The knowledge that H has occurred should increase one's readiness to bet on A. By contrast, for the event B "ball b is in the second cell," $P\{B|H\} = 1/3 = P\{B\}$, and so the knowledge that H has occurred gives no clue as to B. Therefore, B is said to be independent of H if $P\{B|H\} = P\{B\}$, that is, if

$$P\{B \cap H\} = P\{B\}P\{H\}$$

Clearly, in this case $P\{H|B\} = P\{H\}$ so that H is also independent of B. Accordingly, two events B and H are independent of each other if the probability of their simultaneous occurrence follows the multiplication rule $P\{B \cap H\} = P\{B\}P\{H\}$. This notion carries over to systems of more than two events.

Independent trials. The intuitive frequency interpretation of probability is based on the concept of experiments repeated under identical conditions; a theoretical model for this concept can be developed.

Consider an experiment described by a sample space $\mathfrak{S}$; for simplicity of language it can be assumed that $\mathfrak{S}$ consists of finitely many sample points $E_1, \ldots, E_N$. When the same experiment is performed twice in succession, the thinkable outcomes are the N^2 pairs of sample points (E_1,E_1), (E_1,E_2), ... , (E_N, E_N), and these now constitute the new sample space. It is called the combinatorial product of $\mathfrak{S}$ by itself and denoted by $\mathfrak{S} \times \mathfrak{S}$ with reference to analytic geometry, one speaks of the first and second coordinate of the point (E_i, E_j). These notions apply equally to infinite sample spaces and to products $\mathfrak{S} \times \mathfrak{S} \times \mathfrak{S} \cdots$ of more than two factors. For example, the cartesian plane of points (x, y) is the product of the real line by itself. In tossing a coin once, $\mathfrak{S}$ contains only the points H and T; tossing the coin n times leads to the n-tuple product $\mathfrak{S}: \cdots \times \mathfrak{S}$ whose points have n coordinates and are of the form $(HT \cdots T)$.

Probabilities must be assigned to the events in $\mathfrak{S} \times \mathfrak{S}$. The case of dependent trials will be treated in the next section; if the second trial is independent of the first, the probabilities in $\mathfrak{S} \times \mathfrak{S}$ follow the productive rule $P\{E_i, E_j\} = P\{E_i\}P\{E_j\}$.

In the case of n tossings of a coin, this rule leads to the probability 2^{-2} for each sample point in agreement with the requirement of equally likely cases. The present approach is more flexible and more general as shown by the Bernoulli trials.

Bernoulli trials. Suppose each trial results in success S or failure F, and $P\{S\} = p$, $P\{F\} = q$ where $p + q = 1$. (This may be considered as the model of a skew coin.) A succession of n independent trials of this kind leads to the sample space of n-tuples ($SFFS \cdots FS$), and the probability of such a point is the product ($pqqp \cdots qp$) obtained on replacing each S by p and each F by q.

This model has obvious applications to repeated observations and to gambling. Independence is an assumption to be verified experimentally. Conceivably a coin could be endowed with memory and avoid runs of more than 17 successive heads. That the sex distribution within families resembles Bernoulli trials is purely a matter of experience. Many gamblers fully accept the independence and yet believe that they can influence fate by using "systems," for example, by skipping the game after each failure, or waiting for a run of 3 successes, and so on. The theorem on systems shows this to be a fallacy; a gambler not endowed with foresight may use any system or random choice of the times when he plays or skips the game; he remains confronted with Bernoulli trials and is exactly in the same situation as if he played at each trial. *See* DISTRIBUTION (PROBABILITY).

Geometric probabilities. In the interval $0 < x < 1$ a point is chosen at random. This interval is the sample space $\mathfrak{S}$ and the probability of each subinterval equals its length. The sample space $\mathfrak{S} \times \mathfrak{S}$ is the unit square of the x,y plane, and the probability of any figure equals its area. The event "the two successive choices result in a sum <1" is represented by the triangle below the main diagonal and has probability 1/2. The event "the greater of the two choices is $<t$" is represented by the square $0 < x < t$, $0 < y < t$ and has probability t^2.

Dependent trials; Markov chains. Many phenomena can be analyzed in terms of dependent trials. In their description the convenient and picturesque terminology of urn models shall be adopted, which should

not detract from the general nature of the schemes being presented.

Consider an urn containing N balls, of which r are red R and $b = N - r$ black B. Assuming perfect randomness, the probability that a randomly drawn ball be red equals r/N. If the ball is replaced and the procedure repeated, the result is Bernoulli trials with $p = r/N$. Without replacement, the sample space corresponding to two drawings contains four points RR, RB, BR, and BB, to which probabilities are assigned as follows: If the first ball drawn is red (probability r/N), the conditional probabilities of R and B at the second trial become $(r-1)/(N-1)$ and $b/(N-1)$. By Eq. (3), therefore,

$$P[RR] = r(r-1)/N(N-1)$$
$$P[RB] = P[BR] = rb/N(N-1)$$
$$P[BB] = b(b-1)/N(N-1)$$

A more general urn model is obtained by letting the composition of the urn vary from trial to trial. For definiteness consider the following scheme: Each time a ball is drawn, it is replaced, and c balls of the color drawn and d balls of the opposite color are added to the urn. Here c and d are fixed numbers which may be negative. This scheme contains interesting special cases such as the following:

1. When $c = d = 0$, drawing with replacement occurs, and for $c = -1$, $d = 0$, drawing without replacement occurs. In the latter case, the process terminates after N drawings.

2. The Polya model of contagion is the special case when $c > 0$ is fixed and $d = 0$. Here the drawing of either color increases the probability of the same color at subsequent trials, just as in the contagious disease each occurrence increases the probability of further occurrences. This model represents only a crude first approximation to phenomena of contagion, but it leads to comparatively simple formulas and has been applied with astonishing success to a variety of experiences from sickness insurance to baseball scores.

3. The Ehrenfest model for heat exchange considers two containers, I and II, and N particles distributed in them. A particle is chosen at random and removed from its container into the other. This scheme differs only linguistically from the urn scheme. If the particles in I are called red and those in II black, then each trial changes the color of one ball and gives the special case $c = -1$ and $d = 1$.

The probabilities of the various possible outcomes in the general scheme are obtained as above. For example, $P\{RBR\} = r(b+d)(r+c+d)/N(N+c+d)(N+2c+2d)$, and so on.

Markov chains represent another important scheme for dependent trials. Suppose that at each trial the possible outcomes are $E_1, \ldots, E_N$ and that whenever E_i occurs the conditional probability of E_j at the next trial is p_{ij}, independently of what happened at the preceding trials. Here, of course, $p_{ij} = 0$ and $p_{ij} + p_{i2} + \cdots + p_{iN} = 1$ for each i. The p_{ij} are called transition probabilities. The whole process is now determined if the initial probabilities, π_i, at the first trial are known. For example, $P\{E_aE_bE_c\} = \pi_a p_{ab} p_{bc}$. The probability of the event "E_c at the third trial" is obtained by summation over all a and b, and so on. Markov chains, and their analog with continuous time, represent the simplest type of stochastic process. The Ehrenfest model considered above may be treated as a Markov chain by letting E_i represent the event that container I contains i particles. Then $p_{i,i-1} = i/N$, $p_{i,i+1} = (N-i)/N$, and $p_{ij} = 0$ for all other combinations of ij. Other examples of Markov chains are the gambler's accumulated fortune, the composition of a deck of cards under random shuffling, and random walks. Important applications are to queueing theory where one encounters also processes with more complicated aftereffects. *See* QUEUEING THEORY; STOCHASTIC PROCESS.

Random variables and their distributions. The theory of probability traces its origin to gambling, and the gambler's gain may still serve as the simplest example of a random variable. With every possible outcome (sample point) there is associated a number, namely, the corresponding gain. In other words, the gain is a function on the sample space, and such functions are called random variables. (In infinite spaces the idea is the same, but a somewhat more cautious definition is in order.) With the same experiment, one may associate many random variables. As an example, consider the sample space of tabulation (1) with probability 1/27 for each point. A typical random variable is the number N of occupied cells; it assumes the value 1 at the three points numbered 1–3; the value 2 at the eighteen points 4–21; and the value 3 at the six points 22–27. One says, therefore, that the probability distribution of N is given by $P\{N=1\} = 1/9$, $P\{N=2\} = 2/3$, $P\{N=3\} = 2/9$. Another variable is the number X of balls in the first cell. An inspection of tabulation (1) shows that its probability distribution is given by $P\{X=0\} = 8/27$, $P\{X=1\} = 12/27$, $P\{X=2\} = 6/27$, $P\{X=3\} = 1/27$. One may also consider the two variables simultaneously and find, for example, that the combination $N=1$, $X=0$ occurs at two points, whence $P\{N=1, X=0\} = 2/27$. The probabilities of all pairs are given by the joint probability distribution of N and X exhibited in tabulation (4). Adding the

N \ X	0	1	2	3	Distribution of N
1	2/27	0	0	1/27	3/27
2	6/27	6/27	6/27	0	18/27
3	0	6/27	0	0	6/27
Distribution of X	8/27	12/27	6/27	1/27	

(4)

entries in the rows and columns gives the distribution of N and X, respectively, and they are therefore occasionally called marginal distributions. *See* COMBINATORIAL THEORY.

The "Geometric probabilities" example given earlier may be used to illustrate the case of

continuous random variables. This example considers two consecutive selections of a point in the interval $0 < x < 1$. Let S be the random variable denoting the sum of the two choices, and L the larger of the two. One sees that for $0 < t < 1$ the event $L \leqq t$ has probability t^2; thus, setting $P\{L < t\} = F(t) = t^2$ when $0 \leqq t \leqq 1$; for $t < 0$ and for $t > 1$ one has trivially $F(t) = 0$ and $F(t) = 1$, respectively. This is the distribution function of L. From it can be calculated all probabilities relating to L. Similarly the event $S \leqq u$ is represented by the region in the unit square below the line $x + y = u$; therefore, the distribution function of S, namely, $P\{S \leqq u\} = G(u)$, is given by $G(u) = 0$ for $u \leqq 0$, $G(u) = (1/2)u^2$ for $0 \leqq u \leqq 1$, $G(u) = 1 - (1/2)(2 - u)^2$ for $1 \leqq u \leqq 2$, and $G(u) = 1$ for $u \leqq 2$. In like manner, the joint distribution function $P\{L \leqq t, S \leqq u\} = H(t,u)$ of the pair L,S can be calculated.

Every random variable X has a distribution function $F(t) = P\{X \leqq t\}$. If X assumes only finitely many values, then $F(t)$ is a step function. Thus, in tabulation (4), $F(t)$ assumes the values 0, 8/27, 20/27, 26/27, and 1, respectively, in intervals $t < 0$, $0 \leqq t < 1$, $1 \leqq t < 2$, $2 \leqq t < 3$, and $t \leqq 3$. In such cases the notion of distribution function is used mainly for uniformity of language. The notion is really convenient when $F(t)$ is not only continuous but also has a derivative $f(t) = F'(t)$; then $f(t)$ is called the probability density of X. In the above example the variable L has a density defined by $2t$ for $0 < t < 1$ and 0 elsewhere; the density of S is 0 for $u < 0$ and $u > 2$; it equals u for $0 < u < 1$ and equals $2 - u$ for $1 < u < 2$.

The notion of independence carries over: Two random variables X and Y are independent if $P\{X \leqq x, Y \leqq 2\} = P\{X \leqq s\} \cdot P\{Y \leqq t\}$. It is easily seen that for independent variables with distribution functions $F(t)$ and $G(t)$ the distribution function of the sum $S = X + Y$ is given by the convolution shown in Eq. (5). In terms of densities, Eq. (5) reads as Eq. (6). In the random choice example, the coordinates of the points chosen are independent variables with the rectangular density $f(s) = g(s) = 1$ for $0 < s < 1$. The distribution of their sum S which has been calculated above can be found also using Eq. (6).

$$P\{S \leq u\} = \int_{-\infty}^{+\infty} F(u - s)\, dG(s) = \int_{-\infty}^{+\infty} G(u - s)\, dF(s) \tag{5}$$

$$h(u) = \int_{-\infty}^{+\infty} f(u - s)g(s)\, ds = \int_{-\infty}^{+\infty} g(u - s)f(s)\, ds \tag{6}$$

Expectations. Given a random variable X one may interpret its distribution function $F(t)$ as describing the distribution of a unit mass along the real axis such that the interval $a < x \leqq b$ carries mass $F(b) - F(a)$. In the case of a discrete variable assuming the values $x_1, x_2, \ldots$ with probabilities $p_1, p_2, \ldots$ the entire mass is concentrated at the points x_i; if $F'(x) = f(x)$ exists, it represents the ordinary mass density as defined in mechanics. The center of gravity of this mass distribution is called the expectation of X; the usual symbol for it is $E(X)$, but physicists and engineers use notations such as $\langle X \rangle$, $\langle X \rangle_{\text{Av}}$, or $\overline{X}$. In the cases mentioned,

$$E(X) = \sum p_i x_i$$

$$E(X) = \int_{-\infty}^{+\infty} x f(x)\, dx$$

In all cases $E(X)$ is given by the Stieltjes integral over $x\, dF(x)$. (To be precise, one speaks of expectations only when the integral converges absolutely.)

Before discussing the significance of the new concept, a few frequently used definitions are appropriate. Put $m = E(X)$. Then $(X - m)^2$ is, of course, a random variable. In mechanics, its expectation represents the moment of inertia of the mass distribution. In probability, it is called variance of X; its positive root is the standard deviation. Clearly,

$$\text{Var}(X) = E(X - m)^2 = E(X^2) - m^2$$

The variance is a measure of spread: It is zero only if the entire mass is concentrated at the point m, and it increases as the mass is moved away from m. In the case of two variables X_1 and X_2 with expectations m_1 and m_2 it is necessary to consider not only the two variances $s^2_i = E[(X_i - m_i)^2]$ but also the covariance $\text{Cov}(X_1,X_2) = E[(X - m_1)(X_2 - m_2)] = E(X_1X_2) = m_1m_2$. The covariance divided by s_1s_2 is called the correlation coefficient of X_1 and X_2. If it vanishes, X_1 and X_2 are called uncorrelated. Every pair of independent variables is uncorrelated, but the converse is not true.

If $X_1, X_2, \ldots, X_n$ are random variables with expectations $m_1, \ldots, m_n$ and variances $s^2_1, \ldots, s^2_n$, the expectation of their sum $S_n = X_1 + \cdots + X_n$ is always given by $E(S_n) = m_1 + \cdots + m_n$; if all the covariances of X_I and X_j vanish, then clearly $\text{Var}(S_n) = s^2_1 + \cdots + s^2_n$.

When X represents a physical quantity, then $X^* = (X - m)s^{-1}$ represents the same quantity measured from a different origin and in new units. In the physicist's terminology, X^* is the quantity X referred to dimensionless units. In probability, X^* is called the reduced or standardized variable.

It was once assumed that every reasonable random variable has finite expectation and variance. Modern theory refutes this assumption. Many recurrence times in important physical processes have no finite expectations. Even in the simple coin-tossing game, the number of trials up to the time when the gambler's accumulated gain first reaches a positive level has infinite expectation.

Laws of large numbers. To explain the meaning of the expectation and, at the same time, to justify the intuitive frequency interpretation of probability, consider a gambler who at each trial may gain the amounts $x_1, x_2, \ldots, x_n$ with probabilities $p_1, p_2, \ldots, p_n$. The gains at the first and second trials are independent random variables X_1, X_2 with the indicated distribution and the common expectation $m = + p_ix_i$. The event that an individual gain equals x_i has

probability p_i, and the frequency interpretation of probability leads one to expect that in a large number n of trials this event should happen approximately np_i times. If this is true, the total gain $S_n = X_1 + X_2 + \cdots + X_n$ should be approximately nm; that is, the average gain $(1/n)S_n$, should be close to m. The law of large numbers in its simplest form asserts this to be true. More precisely, for each $\epsilon > 0$ it assures one that

$$P\left\{\left|\frac{1}{n}S_n - m\right| > \epsilon\right\} \to 0 \quad \text{as } n \to \infty$$

As a special case, one can obtain a frequency interpretation of probability. In fact, consider an event A with $P\{A\} = p$ and suppose that in a sequence of independent trials a gambler receives a unit amount each time when A occurs. Then the expectation of the individual gain equals p, and S_n is the number of times the event A has occurred in n trials. It follows that

$$\left\{\left|\frac{1}{n}S_n - p\right| > \epsilon\right\} \to 0$$

That is, the relative frequency of the occurrence of A is likely to be close to p.

Without this theorem, probability theory would lose its intuitive foundation, but its practical value is minimal because it tells one nothing concerning the manner in which the averages $n^{-1}S_n$ are likely to approach their limit m. In the regular case where the X_j have finite variances, the central limit theorem gives much more precise and more useful information; for example, it tells one that for large n the difference $S - np$ is about as likely to be positive as negative and is likely to be of the magnitude $n^{1/2}$. When the X_k have no finite variances, the central limit theorem fails and the sums S_n may behave oddly in spite of the law of large numbers. For example, it is possible that $E(X_k) = 0$ but $P\{S_n < 0\} \to 1$. In gambling language this game is "fair," and yet the gambler is practically certain to sustain an ever-increasing loss.

There exist many generalizations of the law of large numbers, and they cover also the case of variables without finite expectation, which play an increasingly important role in modern theory. *See* GAME THEORY; PROBABILITY (PHYSICS); STATISTICS.

William Feller

Bibliography. W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1, 3d ed., 1968, vol. 2, 2d ed., 1971; H. J. Larson, *Introduction to Probability*, 1995; S. M. Ross, *A First Course in Probability*, 5th ed., 1997.

Probability (physics)

To the physicist the concept of probability is like an iceberg. The part of it which the physicist uses and which is, therefore, in full view for him or her is but a small fraction of what is hidden in other and larger disciplines. The philosophy and mathematics of probability have become increasingly interesting and important. *See* PROBABILITY.

Bernoulli's problem. One of the most basic problems encountered in the application of probability to physics was solved by J. Bernoulli. It concerns the probability of achieving a specified number (x) of successes in n independent trials when the probability of success in a single trial is known. Denote by p the probability of success; let $q = 1 - p$ be the probability of failure. Equation (1) represents

$$C_x^n \equiv \frac{n!}{x!(n - x!)} \tag{1}$$

the well-known binomial coefficient. The probability in question is given by Eq. (2). For example, one

$$w_n(x) = C_x^n p^x q^{n-x} \tag{2}$$

may consider an urn containing a black balls and b white balls; n drawings are to be made from this urn, with replacement of the drawn ball each time. The probability that x white balls shall turn up in these n drawings is required. In this example, Eqs. (3) hold.

$$p = \frac{b}{a+b}$$
$$w_n(x) = C_x^n \left(\frac{b}{a+b}\right)^x \left(\frac{a}{a+b}\right)^{n-x} \tag{3}$$

Equation (2) defines Bernoulli's distribution; because of the newtonian binomial coefficients which it involves, it is sometimes called Newton's formula. It satisfies relations (4)–(6). Hence, the σ^2 is called

$$\sum_{x=0}^{n} w_n(x) = 1 \tag{4}$$

$$\bar{x} = \sum_x x w_n(x) = np \tag{5}$$

$$\sigma^2 = \sum_x (x - \bar{x})^2 w_n(x) = \bar{x^2} - \bar{x}^2 = npq \tag{6}$$

the dispersion of the distribution $w_n(x)$. The so-called standard deviation σ equals $\sqrt{npq}$.

The physicist regards the emission of particles from nuclei, of photons from hot bodies, and of electrons from hot filaments as random phenomena controlled by probability laws. One of the simplest methods for testing this assumption is based upon these formulas. Suppose that a radioactive specimen emits particles for a finite period T, and that their number is x. By repeated observations on the emission of particles in a period of T seconds, one obtains a series of numbers, x_1, x_2, x_3, etc. Now imagine the period T to be subdivided into a very large number, n, of intervals, each of length τ, so that $n\tau = T$. Indeed, τ will be assumed to be so small that, at most, one particle is emitted within the interval τ. A single observation made during that infinitesimal interval would therefore yield the result: either no emission

or emission, say, with probabilities p or q. By hypothesis, n such observations yield x emissions. It is seen, therefore, that the problem involves an application of Bernoulli's distribution. While it is difficult to specify the values of p, q, and τ, the result given by Eq. (6) must nevertheless be true. Since q is extremely small, p is very nearly 1. Hence, in view of Eq. (5), $\sigma^2 = npq = nq = \bar{x}$. This relation has been tested experimentally and found to be true in all instances.

Approximations. In many practical applications, the values of x and n are very large, and a direct use of Newton's formula becomes impossible because the binomial coefficients are difficult to evaluate for large x and n. The example of particle emission just discussed is a case in point. Under these circumstances, two approximations to Newton's formula are available, one derived by C. F. Gauss, the other one by S. D. Poisson and bearing his name. Gauss's law is the limiting form of Eq. (2), when both x and n are large, so large in fact that $1/\bar{x}1/\overline{x} = 1/np$ and $1/npq$ are both negligible. This implies that p and q are numbers not greatly different from unity. In that case, Eq. (7) holds. This is Gauss's law.

$$\lim_{n\to\infty} w_n(x) = \frac{1}{\sqrt{2\pi px}} \exp\left[\frac{-(x-\bar{x})^2}{2px}\right] = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[\frac{-(x-\bar{x})^2}{2\sigma^2}\right] \tag{7}$$

If p is so small that the mean np is of the order of unity in any given application, Bernoulli's distribution is approximated by Poisson's law, Eq. (8). The

$$\lim_{n\to\infty} w_n(x) = \frac{(np)^x e^{-np}}{x!} \tag{8}$$

validity of this formula has also been confirmed in many instances of particle emission.

The density fluctuations which occur in a finite volume of gas present an interesting application of Gauss's law. Suppose that a vessel of volume V contains but a single molecule. One may then select a small element of volume υ and imagine successive observations to be made in order to determine whether the molecule is in this element or not. The probability of its being there, which will be denoted by p, is clearly the ratio of the small volume element to the volume of the total container, $p = \upsilon/V$. Now let n observations be made upon the volume υ. The probability that, in x of them, the one molecule is found in υ is given by Newton's formula. If the vessel contains not one but n molecules, then the probability of finding x of them in υ simultaneously is the same as that of finding the one molecule x times in υ on n successive occasions. Thus, the probability of observing x of the n molecules in υ is given by Eq. (9).

$$w_n(x) = C_x^n p^x q^{n-x} \tag{9}$$

Equation (9) may be approximated by Gauss's law. Thus, on substituting p and q, as shown in Eqs. (10) and (11), Eq. (12) results. The expected mean and

$$p = \frac{\upsilon}{V} = \frac{1}{k} \tag{10}$$

$$q = 1 - p \tag{11}$$

$$w(x)\,dx = \frac{k}{\sqrt{2\pi n(k-1)}} \times \exp\left[-\frac{(kx-n)^2}{2n(k-1)}\right] dx \tag{12}$$

the dispersion of this distribution must be the same as the corresponding quantities for Bernoulli's distribution. Hence Eqs. (13) and (14) hold. If k is sufficiently large, $\sigma^2 = \bar{x}$ approximately.

$$\bar{x} = np = \frac{n}{k} \tag{13}$$

$$\sigma^2 = npq = \frac{n}{k^2}(k-1) = \bar{x}\left(1 - \frac{\bar{x}}{n}\right) \tag{14}$$

It is often convenient to introduce a quantity δ known as the relative fluctuation and defined by Eq. (15).

$$\delta = \frac{x - \bar{x}}{\bar{x}} \tag{15}$$

In view of Eq. (13), Eq. (16) holds.

$$\delta = \frac{kx - n}{n} \tag{16}$$

In terms of δ, the law of fluctuations takes on the simple form given by Eq. (17). Here $\bar{x}$ is the mean

$$w(\delta)\,d\delta = \sqrt{\frac{\bar{x}}{2\pi q}} \exp\left[\frac{-x}{2q}\delta^2\right] d\delta \tag{17}$$

number of molecules within υ, and q is very nearly equal to 1 if $\upsilon \gg V$. Equation (17) is well supported by experiment.

Theory of errors. Perhaps the most important application of probability theory to exact science occurs in the treatment of errors. Measurements are accompanied by errors of two kinds: determinate and random. The former arise from actual mistakes, either on the part of the observer or from faulty instruments; they are not susceptible of mathematical treatment. Random errors, however, because they are numerous, small, and likely to combine in linear fashion, are subject to the laws of probability analysis.

Gauss error law. Let there be n measurements of some physical quantity resulting in the numbers X_1, X_2, $X_3, \ldots, X_n$. If the true value of the quantity (usually unknown) is denoted by X, then the errors are given by Eqs. (18). It can be shown mathematically and

$$x_1 = X_1 - X \qquad x_2 = X_2 - X, \ldots \qquad x_n = X_n - X \tag{18}$$

has been confirmed by numerous observations that the relative frequency of occurrence of an error x is

represented by Bernoulli's distribution which, under the present conditions, takes on the form of Gauss's law, Eq. (19). In this formula the parameter h^2 equals

$$N(x)\,dx = \frac{h}{\sqrt{\pi}} e^{-h^2x^2}\,dx \tag{19}$$

$1/(2\sigma^2)$ and is called the index of precision. Clearly, the greater the value of h the narrower the distribution given by Eq. (19). The latter is often called the Gauss error law or the normal distribution of errors. In writing it, the assumption has of course been made that the numbers x_1, x_2, etc., may be replaced by a continuous distribution.

The probability that a single measurement will contain an error between the limits $-a$ and $+a$ is given by Eq. (20). The integral occurring here is

$$\frac{h}{\sqrt{\pi}} \int_{-a}^{a} e^{-h^2x^2}\,dx = \frac{2h}{\sqrt{\pi}} \int_{0}^{a} e^{-h^2x^2}\,dx \tag{20}$$

called the error function, and is denoted by erf (a); it is tabulated in most textbooks that discuss the theory of errors. The factor in front of the integral has been chosen so that Eq. (21) holds.

$$\int_{-\infty}^{\infty} N(x)\,dx = 1 \tag{21}$$

Returning to the set of measured values $X_1, X_2, \ldots, X_n$, one may ask: Which is the most probable value (the one most likely to be true) consistent with this set of numbers? The answer is given by the principle of least squares, which affirms that the most probable value is the one for which the sum of the squared errors $x_1^2 + x_2^2 + \cdots + x_n^2$ is a minimum. *See* LEAST-SQUARES METHOD.

From this one may prove directly that the most probable value of X is its arithmetical mean, given by Eq. (22).

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} \tag{22}$$

Kinds of errors. The reliability of a set of measurements, such as the sequence $X_1, X_2, \ldots, X_n$, is specified by certain measures called errors, but in a slightly different sense from that previously employed. Three kinds of "error" will be described: the average error a, the root-mean-square error m, and the probable error r. All three refer, not to a single measurement as did the quantity x_i, but to the entire distribution in Eq. (19). The average error a is the arithmetical mean of all individual errors without regard to sign, as shown in Eq. (23).

$$a = \frac{\Sigma|x_i|}{n} \tag{23}$$

When the averaging process is carried out with the use of Eq. (19), Eq. (24) is obtained.

$$a = \int_{-\infty}^{\infty} |x|N\,dx = \frac{2h}{\sqrt{\pi}} \int_{0}^{\infty} xe^{-h^2x^2}\,dx = \frac{1}{h\sqrt{\pi}} \tag{24}$$

The measure of precision h has already been defined. It is an inverse measure of the width of the gaussian curve, but it cannot be established without ambiguity from a finite set of measurements X_i. If the most probable value of h is calculated, it turns out to be the root-mean-square error given by Eq. (25).

$$m = \sqrt{\Sigma x_i^2/n} \tag{25}$$

Comparison with Eq. (24) shows that Eq. (26) is

$$a = m\sqrt{2/\pi} \tag{26}$$

valid. The quantity m is identical with what was previously called the standard deviation.

The probable error r is defined as follows: It marks that value of x which divides the area under the curve $N(x)$ between zero and infinity into two equal parts. Hence, r is an error such that a given error x_i has an equal chance of being greater or smaller than r. Mathematically, the value of r is found from Eq. (27).

$$\text{erf}\,(hr) = 1/2 \tag{27}$$

The numerical relations between r, m, and a are given by Eqs. (28).

$$\begin{aligned} r &= 0.6745m = 0.8453a \\ m &= 1.4826r = 1.2533a \\ a &= 0.7979m = 1.1829r \end{aligned} \tag{28}$$

Probability in statistical mechanics. A complex physical system made up of many constituents, for example, molecules, obeys the empirical laws of thermodynamics. Statistical mechanics is that science which attempts to explain these laws by an appeal to the laws of ordinary mechanics. In doing so, it encounters problems of the following sort.

Suppose one wishes to explain why a gas exerts a pressure upon the walls of a container. Pressure is the momentum lost by the molecules per unit time as they strike a unit area of wall. The molecules are very numerous, and their momentum losses vary erratically from instant to instant of collision, and from point to point on the surface. It is necessary, therefore, that some average be taken. Should this average be over the different collisions which a single molecule experiences in time? Or should one take the average at a given time over the whole area under consideration? Clearly, there are many ways of computing the average, each involving a particular collective in the aforementioned sense and each requiring a specification of elementary probabilities.

The science of mechanics does not clearly dictate which collective is the proper one to be used, and there is a considerable latitude of choice. One of these collectives which has proved most successful, the ensemble of J. W. Gibbs, will be discussed briefly.

Phase space. The motion of a single molecule is described in terms of two variables—its position and its momentum at any given time. In the simplest case of motion along a single axis, for example, the x axis, two numbers suffice to describe the motion; they are

the position x and the momentum p_x. If a plane is constructed with x and p_x laid off on two perpendicular axes within that plane, the motion of the molecule is represented by a curve in the plane, and the plane is called the phase space of the moving molecule. A molecule whose motion is in a plane requires x, y, p_x, and p_y, that is, four numbers, for its complete description, and these four numbers define a point in a phase space of four dimensions. Similarly, a molecule moving in three-dimensional space has a phase space of six dimensions (axes x, y, z, p_x, p_y, and p_z) and its dynamic behavior is depicted by a curve in this six-dimensional space.

This idea can be generalized and applied to a gas containing N molecules. The phase space of the gas will have $6N$ dimensions, and the motion of the entire gas corresponds to the trajectory of a single point in this $6N=$ dimensional space. A single point within it describes the physical condition of the entire gas.

Ensembles. Here Gibbs introduced the notion of an ensemble. He imagined a great number of thermodynamic systems, all similar to the given one. If the latter be a vessel filled with gas, he imagined a very large number of similar vessels, all filled with the same quantity of the same gas. This collection of imaginary vessels is called an ensemble. Each member of the ensemble will have its fate represented by a point moving in $6N$-dimensional phase space, and the whole ensemble, when viewed in that space, will appear like a cloud of dust, with each individual dust particle following its own path. The density of this cloud of dust will differ from place to place and will change in time at any given place. From the laws of mechanics it may be shown that this imaginary cloud of dust behaves like an incompressible fluid.

There is a set of conditions, however, under which the cloud will not change its density in time, even though its individual points are in motion. One such condition amounts to the existence of a special density distribution known as the canonical distribution. It has the simple form given by Eq. (29), where H

$$D(x_1 \cdots p_n) = \text{ constant} \times \exp\left[\frac{-H(x_1 \cdots p_n)}{kT}\right] \tag{29}$$

is the energy, T the temperature of the gas, and k is Boltzmann's constant.

This distribution is relevant to the problem discussed above. All dynamical variables, such as the pressure of the preceding example, which need to be averaged in order to correspond to the observables of thermodynamics, are to be averaged over the probability distribution D. When this is done, the laws of thermodynamics follow; in that sense, Gibbs' probability distribution provides an explanation of thermodynamics.

The success of Gibbs' theory in classical mechanics is remarkable. In order for it to be applicable, however, to systems which follow the laws of quantum mechanics, certain modifications are necessary. *See* BOLTZMANN STATISTICS; QUANTUM STATISTICS; STATISTICAL MECHANICS; STATISTICS.

Henry Margenau

Probability in quantum theory. Probability enters into both classical and quantum physics, but on very different footings. The toss of a coin provides an example of probability in classical physics. It is possible to learn to toss sufficiently gently, with little spin and the right initial conditions, so as to generally guarantee either outcome, heads or tails. But vigorous tosses with lots of spin produce outcomes beyond practical control. Classically, there is no doubt that the outcome of any particular toss, with its particular initial conditions, is fully deterministic. Rather, what the seeming randomness reflects is that for vigorous tosses the outcome is very sensitive to initial conditions, switching from one result to another with tiny changes in those initial conditions (or tiny changes, too, in external influences such as passing breezes).

In contrast, probability notions enter into quantum mechanics intrinsically, at the very foundations. According to quantum theory, the most that can be known about the state of a system is not enough to determine with certainty the outcome of measurements on that system. That is, repeated experiments on systems in exactly the same quantum state will generally yield a spread of outcomes depending on the particular state and on the dynamical quantity being measured. However, although the outcome of any particular measurement is not predictable, the probability distribution is sharply determined for any given state; and the evolution over time of that state, hence of probability distributions, is fully deterministic. If one of the surprises of quantum mechanics is this probabilistic foundation, another—a feature that gave quantum mechanics its name—has to do with the universe of possible outcomes for measurements of various dynamical quantities. In the case of linear momentum, for example, the possible outcomes range continuously over all magnitudes and directions just as they do classically; so too for position. But for angular momentum, in striking contrast with the classical situation, only certain discrete outcomes are possible; and so too in certain cases for energy measurements. The possible outcomes for certain quantities, that is, are quantized. *See* ANGULAR MOMENTUM; ENERGY LEVEL (QUANTUM MECHANICS).

But what is most distinctive about quantum mechanics is the rule for composition of probabilities. This can be illustrated with the example of the scattering of a beam of electrons off a small, weakly scattering target centered at some location. Distant detectors record the number of electrons that emerge at various scattering angles, θ, and the angular probability distribution is labeled $P_1(\theta)$. Next, the first target is removed and replaced with another, centered at a new location. The new probability distribution is labeled $P_2(\theta)$. Finally, both targets are set up, each at its previous location (supposing that there is no overlap), and the resulting probability distribution is labeled $P(\theta)$. Classically, an electron detected at any particular angle would have had to

scatter off one target or the other, so classically it would be expected that $P = P_1 + P_2$. But this is not what happens quantum-mechanically. Instead, quantum mechanics deals with probability amplitudes, complex quantities whose absolute squares determine the probabilities; and it is the amplitudes to which the simple composition law applies: $A = A_1 + A_2$. This interference of amplitudes is what gives quantum phenomena their wavelike character. *See* INTERFERENCE OF WAVES.

In the case of a single particle moving in a potential, the state of the system, alluded to above, is described by a complex wave function of space ($\boldsymbol{r}$) and time (t), $\Psi(\boldsymbol{r},t)$. The wave function incorporates all that can be known about the state. It evolves in time deterministically, as governed by the Schrödinger equation, a partial differential equation. Quantum theory is concerned with the principles for extracting probabilistic predictions from the wave function and with the application of these principles to various physical problems. To give one elemental example: the probability of finding the particle in an infinitesimal volume element d^3x centered around the position $\mathbf{r}$ is proportional to $\}\Psi\}^2 d^3x$, where $\}\Psi\}$ is the absolute value of the complex wave function Ψ. The rules for extracting other kinds of probability distributions, for example, predictions concerning the outcome of energy measurements, are equally definite, although mathematically more complicated. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; SCHRÖDINGER'S WAVE EQUATION.

Sam Treiman

Bibliography. W. T. Grandy and P. W. Milonni (eds.), *Physics and Probability*, 1993; S. P. Gudder, *Quantum Probability*, 1988; R. W. Hamming, *The Art of Probability for Scientists and Engineers*, 1993; A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3d ed., 1991; B. P. Roe, *Probability and Statistics in Experimental Physics*, 2d ed., 2001; D. ter Haar, *Elements of Statistical Mechanics*, 3d ed., 1995.

Problem solving (psychology)

Problems represent gaps between where one is and where one wishes to be, or between what one knows and what one wishes to know. Problem solving is the process of closing these gaps by finding missing information, reevaluating what is already known or, in some cases, redefining the problem.

The complexity of human behavior is frequently overlooked, so that everyday activities and experiences are often viewed as uneventful and typical. The regularity of overt human behavior is taken for granted, and people often miss the demands that are made upon cognition by the events and activities in their daily lives. Psychologists, however, investigate the complexity of human behavior and seek to understand its nature and action. Pioneers such as William James realized that behavior was not an isolated response to objects and forces in the environment, but rather represented a systematic adaptation to specific environmental demands and opportunities. One kind of behavior that is created in response to the environment is problem solving. Research into this behavior has attempted to discover the information that is processed and the techniques for processing it.

For many years the main thrust of psychological research on problem solving has focused upon the identification of specific behaviors that enable individuals to reach the goal in a given problem situation. The solution to a problem was thought to consist of an infrequently occurring response to the problem statement. Much of the early work along this line tended to focus on what has come to be known as well-structured problems. These are typically situations with a known beginning, a known end, and a well-defined set of intermediate states. Solving a well-structured problem consists of finding an infrequently used path connecting the initial state of the problem with its end state. One example of such a problem is the game of tic-tac-toe. Having a well-structured problem does not mean, however, that it will be easy to solve. In a game such as chess the set of possible intermediate states (moves) is enormous, and an exhaustive search through this set in order to find a good move is not feasible. People solve well-structured problems such as chess not by exhaustively searching through the set of possibilities, but rather by heuristically identifying good starting places and productive lines of search.

Two early pioneers in the field of human problem solving were Allen Newell and Herbert Simon. According to their view, people solve problems by first constructing an internal representation, called a problem space. This representation contains the information that must be processed in order for a problem to be solved. The problem space contains knowledge of the initial state and the goal state of the problem as well as possible intermediate states that must be searched in order to link up the beginning and the end of the task. Transitions between one state and another are achieved by selecting and applying operations that process information from memory or the task, in order to reduce the differences between what is given and what is to be achieved.

The structure of a given task helps determine the size and complexity of the problem space that is created to solve it. Characteristics of the human information-processing system, such as limited attention and a fallible memory, constrain the problem space so that the actual set of possibilities considered is usually smaller than the total set available. Simon coined the term satisficing to describe the way in which human problem solvers attempt to limit and redefine a task so that the problem space to be attacked is a simpler version of the one that could be described. While the strategy of simplifying or changing a given problem may seem arbitrary, it is also quite common. The same set of facts and information given to someone from the marketing department in a large corporation will frequently be interpreted as a marketing problem, while when given to a person

in the operations area will appear to be a production problem. The fact that people tend to see new situations in terms of their own past experience enables them to cope with problem complexity. But it also means that behavior is often subject to distortion and bias. Successful consultants and top-level managers recognize this and are constantly on the lookout for evidence of satisficing in their own behavior and in the behavior of others. *See* INFORMATION PROCESSING (PSYCHOLOGY).

Arenas of study. One means of discovering how the human problem-solving process is carried out by human beings, regardless of the type of problem-solving task encountered, is based upon observations made through the use of thinking-aloud protocols. This technique requires an individual to talk about whatever comes to mind as a problem is being solved. Analysis of transcriptions of such protocols reveals that for individuals who are relatively inexperienced at a task (novices), the activity of problem solving often consists of general strategies for linking up one stage with another in the search for a solution. One such general strategy is the means-ends analysis, which consists of identifying some future point in the task, perhaps the end, and attempting to find a means for reducing the difference between one's current status and that future point. A less powerful though more general strategy of a similar sort is referred to by computer scientists as generate-and-test and by psychologists as trial-and-error behavior. It consists of picking a possible answer, trying it out, and if it does not work, trying another. Means-ends analysis and trial-and-error behavior can require large amounts of time to complete if the problem is complex, and may not lead to a solution at all in a practical amount of time. They have been successful primarily in games and relatively simple problem-solving tasks. The analysis of thinking-aloud protocols reveals that for experts the activity of problem solving involves the use of problem-specific and knowledge-intensive methods and techniques (often referred to collectively as heuristics). These heuristics are acquired through experience and represent the basis for expertise in a specific domain of problem-solving tasks such as chess, physics, or medicine.

While in fairly well-defined problem areas such as chess playing, a solution consists of an attempt to convert an initial state of a task (for example, a chess board before the first move is made) into a final or goal state (checkmate) through a series of actions or problem-solving steps, in many other problem areas either the initial state, the final state, or both may be poorly understood (these are referred to as ill-structured problems). If the necessary problem-solving steps are known ahead of time, problem solving consists of finding the right steps to apply at the right time. If some of the problem-solving steps are not known, then the problem-solving process requires creating or inventing new ways to convert one state of a task into another.

Chess. Following early work using protocol techniques with fairly simple, well-structured problems, investigators began to branch out to the study of more complex problem-solving situations. One of the earliest domains attacked was skill in playing chess. Enlarging upon early work by Adrian DeGroot, investigators in the United States began to study the differences between grand masters and average chess players. This work has shown that the important characteristics of those especially skilled in a problem-solving domain such as chess is not in a better set of general problem-solving skills, but rather in a vast amount of knowledge that the expert uses to identify important cues and especially good paths to take to a potential solution. Thus in the case of chess, the difference between a grand master and lesser player is not in the number of moves that can be searched in attempting to find a good move, but rather in the set of possible moves considered.

Physics. As in chess, novices (undergraduates) as well as experts (graduate students or professors) have been observed solving textbook problems through the technique of protocol analysis. Study of these protocols has identified two types of problem-solving steps: representing the problem situation, and applying physics principles to the problem representation in order to generate a solution. Representation involves translating the problem statement into a standard form that is accessible to physics principles. For many physics problems, this translation may entail identifying the problem as an instance of some standard or prototype problem (for example, a harmonic oscillator, free fall at constant gravity, or a block on a frictionless plane) that has known solution methods. More generally, translation consists of equating aspects of the problem with idealized objects such as massless rods, frictionless surfaces, and point masses, whose behavior can be represented fairly clearly by a few basic principles. Of course, a given representation can be inappropriate or flatly wrong, and even a good representation may not be tractable (for example, the classic three-body problem in physics).

Once a representation is created, problem solving proceeds by finding physics principles that will generate a solution. In some instances, single principles (such as uniform linear motion) will be sufficient to derive a solution, while in other cases combinations of principles (rotational as well as linear motion) are necessary. In order to be applied efficiently either singly or in combinations, principles of physics need to be selected in such a way that they fit the assumptions and data of a given problem. When several principles are involved, the coordination of their application becomes important. Prototypical problems in physics are those for which single principles suffice, or ones for which the coordination of principles is taught explicitly in introductory textbooks (for example, the simultaneous use of momentum and kinetic energy conservation in an elastic two-body collision). In more advanced texts, procedures are identified for handling additional complexities, and algorithms are derived for dealing with general cases (such as lagrangian and hamiltonian equations in advanced mechanics).

Problem solving sometimes fails because relevant principles are either not known or have not been learned. In such cases, specific heuristics or shortcuts must be employed. In mechanics problems, for example, the goal of problem solving is typically to infer the behavior of a physical system over long spans of time, given only a description of its initial state. One heuristic for accomplishing this goal is to begin with the initial state of the system (for example, a body at rest) and piece out the time course of system behavior by applying interactively the laws that say what happens in the next instant of time (this process is often referred to as simulation). A second heuristic for specifying the behavior of a physical system is to integrate the differential laws of motion to yield an expression that describes the state of the system at any point in the future as a function of its initial state. Because mathematical complexities often make either of the above heuristics difficult to carry out in practice, additional heuristics are sometimes employed. One such heuristic involves describing steady-state solutions of the system (the places where the laws of motion state that nothing changes). Another heuristic involves describing asymptotic behavior of the system (what happens to it in the long run). Yet another heuristic is to specify a property of the system (such as energy or momentum) that is invariant over time, even though there is variation in other aspects of the system. These heuristics (and others like them) are typically learned and applied by individuals in specific problem-solving tasks. Indeed, the power of heuristics as problem-solving tools derives from their development in the successful solution of particular types of problem situations.

The use of fairly elegant heuristics (such as those described above) is often associated with level of expertise; they are used more often by experts than novices. Such an association does not mean, however, that individuals at more advanced stages of training and experience do only heuristic thinking, nor even that this is their preferred method of solving problems. Like other cognitive activities, problem solving requires the expenditure of mental effort. Because individuals at all levels of expertise seem to prefer to expend the least amount of effort that will do the job, experts as well as novices generally solve problems by relating them to their past experience in the form of problem-solving examples (often referred to as prototypes). It is only when truly unfamiliar or difficult problems are encountered that the expert resorts to the more demanding skills of heuristic thinking, which are for the most part unavailable to novices.

Medicine. Many of the features just described for solving problems in chess and physics can also be found in the behavior of individuals attempting to solve the even more ill-structured problems of medicine. Like problem solving in physics, medical problem solving has been studied intensively by using techniques for analyzing thinking-aloud protocols. Whereas in physics equal amounts of effort are often expended on problem representation and solution generation, in medicine one aspect of problem solving deals almost exclusively with finding a good representation (the problem of medical diagnosis).

Solving a diagnostic problem in medicine seems to consist of applying hypotheticodeductive reasoning, in which the basic units are questions, cues, and diagnostic hypotheses. In order to obtain a good representation, diagnosis typically begins with the elicitation of a small number of cues (the patient's signs and symptoms) elicited via questions (usually through the taking of a patient history). These cues suggest a limited number of diagnostic hypotheses which, in turn, suggest new revised hypotheses. Eventually, each hypothesis is evaluated with respect to its relative ability to account for the available cues. Physicians at all levels of skill appear to generate hypotheses in similar quantities, to use those hypotheses to guide collection and interpretation of additional patient data, and to evaluate the relative ability of the hypotheses to account for those data. Differences between experts and novices (and among members of each group) occur in the detailed content of the medical knowledge they possess and in the way it interacts with the specific features of the case they are working on.

At a detailed level, the activity of medical problem solving is guided by specific representations of the problem situation. Images of "what the problem could be" enable forward-searching strategies for the collection of additional information as well as heuristics for generation and evaluation of possible categories within which to interpret this information. Such strategies propose a large number of recurring patterns or prototypical combinations of data (such as combinations of pieces in chess, or combinations of patient signs and symptoms in medicine) that are familiar and can be recognized. The organization of these prototypes is hierarchical, with the more general patterns at the top and the most specific ones at the bottom. Such patterns are stored in memory in the form of templates to be matched against external data, so that a particular configuration of data can be interpreted as an instance of a given pattern or prototype. For the expert, access to a prototype occurs at multiple levels in the stored hierarchy (top, middle, bottom) and is a function of other patterns (pattern-to-pattern links) and data (data-to-pattern links). Unfamiliar experiences are interpreted by building new prototypes. This occurs either through a form of data-driven processing in which commonalities are abstracted from recurring patterns or cues, or through a more conceptually driven form of processing in which higher-level rules are used to derive patterns that can then be matched with unfamiliar data.

General skills. Although much of the work on problem solving today proceeds on a field-by-field basis, some investigators continue to search for a set of general problem-solving skills. Such work suggests that if a more generalized set of skills for solving problems exists, they have much in common with the basic skills of learning. Part of what it means to learn, is

to be able to solve the problem of linking up a given situation with some specific goal or outcome.

Attention as a resource. Attention is a psychological process that plays a major role in the development of efficient and skilled performance in many complex problem-solving tasks. One definition of attention is simply a mechanism with a limited capacity for processing information. In this sense, attention provides a bottleneck through which information about the environment, and the interpretation of that information, must pass. Attention can also be thought of as an active process for allocating cognitive processing resources to components of a task. According to this view, problem-solving performance is a function of the demand for processing resources by the mental operations that are applied to achieve a given problem solution. Complex problems in which large amounts of information or knowledge must be processed require efficiency in order to overcome basic human processing limitations. Therefore the performance that is observed in less experienced individuals is often a function of their attempt to cope with limited resources and an overabundance of problem-solving information.

An important characteristic of problem-solving skill is the manner in which experts circumvent their processing limitations. Skilled performance results when individuals are able to focus upon a small set of cues and relevant task information. Skilled performance also results from automatic control over the type of processing resources used. An expert who has developed skill in solving a specific class of problems typically reduces the processing demands of the task by changing an information-processing step that requires more of the available resources into one which requires less but is just as effective (for example, rather than figuring out a complex problem-solving step each time it occurs, the expert simply remembers it). Efficient information processing is an important component of expert skill that is developed through extended practice in specific problem-solving activities.

An example of cognitive efficiency in problem solving is provided by the skill of learning to read. The beginning reader must give attention to discovering and differentiating many visual features such as the lines and angles that compose the letters printed on a page. After repeated attention to such features, the reader begins to code patterns of features as a single unit (a letter). With enough exposure and practice, the reader eventually learns to recode letters into higher-order units such as words and phrases. As these higher-order units are developed, the lower-order units no longer require conscious attentional processing—they are automatically coded. Such automatic processing allows the skilled reader to allocate attentional capacity to the more complex task of comprehension and deriving meaning from the printed material. *See* COGNITION.

Garden path problems. The fact that experts achieve their expertise in part through an efficient allocation of problem-solving resources is a two-edged sword. Research has identified classes of problems in which errors regularly occur, even among fairly expert problem solvers, due to the misapplication of attention. The usual desirable strategy of beginning a task with fairly general concepts and then working toward more specific levels of detail can fail when the structure of a problem is such that the initial choice of a representation either does not apply or else fails to emphasize the kind of knowledge required for a successful solution. Some types of problem situations have the additional property that the dominant interpretation they elicit appears to be correct when in fact it is not. Problems of this type are referred to as garden path problems since they induce performance that has the false appearance of success.

One famous example of a garden path problem is from physics and is attributed to the work of Harvey Cohn. In Cohn's original problem, termed Milko, a milk bottle is filled with a mixture of milk and cream and left to stand. After a time, the milk and cream separate and the cream rises to the top without any change in volume in the container. Cohn asked simply whether the pressure at the base of the bottle is the same or different after the separation of the two fluids. Expert scientists and engineers as well as more novice problem solvers often fail to correctly solve the Milko problem.

Difficulties with Milko appear to be due primarily to an improper choice of a problem representation as well as an incorrect level of detail in reasoning. Pressure change is thought to be associated with change in either the height or amount of a fluid above the base of its container. And since neither of these has changed in the Milko problem, individuals infer that there has been no change in pressure (the garden path answer). What problem solvers fail to consider, however, is that the situation deals with two variables—a nonhomogeneous fluid and a container with slanted sides. If one of these variables is present by itself, there is no pressure change, something which is described in many introductory physics books as Pascal's principle. If both variables are present, however, the pressure at the base of the container is less following the separation of the two fluids. It is interesting to note that if individuals working on the Milko problem are given a second problem in which the shape is made explicit by drawing a flask with slanted sides, the more expert individuals will often recognize their error in the first problem and proceed to do both problems correctly. For some experts and most nonexperts, however, the illusion that the pressure does not change is so great that a third problem is necessary, one in which explicit values are given for the variables of shape and fluid density so that an exact calculation can be made. Faced with calculations showing that the pressure in the separated state is indeed less, experts not only recover and adopt the correct problem-solving approach but also frequently generalize this recovery both forward and backward. They immediately infer that answers to previous problems were wrong, and additional problems of the same type are correctly solved. Novices, by contrast,

even when given an opportunity to make calculations, typically do not generalize either to the previous problems or to new ones, and will frequently abandon the correct approach when given another problem.

Garden path problems, due mainly to the allocation of attention to one thing while overlooking another, exist in all fields. In medicine, for example, the possible diseases for a given set of signs and symptoms often have information in common. Because data that trigger the correct disease also trigger others, the wrong hypothesis may be chosen. In fact, in some areas of medicine, research has shown that the early cues in a case inevitably trigger the wrong answer so that the expert must avoid overcommitting processing resources to an early, incorrect response in favor of a later, correct alternative.

Computer applications. Scientists working in the field of artificial intelligence have developed computer programs that are capable of solving problems long thought attackable only by human experts. Computer programs exist that are capable of solving problems in medical diagnosis, oil exploration, and engineering design. Having computer programs that can solve complex, ill-structured problems has made it possible to turn understanding of expert problem solving into a useful resource. Computer programs (usually called expert systems) are currently being developed to give advice to problem solvers working in hospitals, industry, and government.

In addition to potential for providing resources for problem solving and training, programming computers to solve problems has served to highlight some interesting facts about the human problem-solving process. Among the most important of these is the difference that exists between problem solving that proceeds on the basis of recognizing situations as instances of ones that have been dealt with in the past (the use of prototypes), and problem solving that proceeds by reasoning through a solution in terms of basic facts and principles from beginning to end. The former permits great efficiency and economy of thought. Through experience, individuals build up a vocabulary of problem-solving prototypes that represent recurring problem situations. It is these prototypes that form the basis for the ability to recognize new problems as instances of ones that have been solved in the past. On the other hand, research also has shown that these prototypes often serve as a source of misinterpretation, bias, and error.

As mentioned at the beginning of this article and as illustrated with the Milko problem in physics, a housing problem may be interpreted by one individual as landlord exploitation, by another as tenant neglect, and by a third as poverty or racial discrimination. When used to solve problems, prototypes have both descriptive and normative functions. These functions are guided by the metaphors contained in the prototypes. Such metaphors allow one to describe and understand complex problem-solving situations. Many social problems, for instance, can be seen in terms of a health metaphor. Families are sick, housing is decayed, or neighborhoods are unhealthy. A health metaphor may suggest different solutions than a need metaphor. Because prototypes can lead to alternative interpretations of a given problem-solving situation, problem solvers need to be aware of them. By programming computers to solve problems, researchers are learning how to achieve a more balanced and meaningful formulation of a given problem-solving task. *See* ARTIFICIAL INTELLIGENCE.

Instruction. Not surprisingly, teaching problem-solving skill has been a topic of interest to psychologists and educators alike. The biggest payoff for instruction from current research in problem solving has been simply learning more about what task-specific knowledge and strategies need to be learned and where garden paths and other pitfalls may lie as the skills of problem solving are acquired and practiced. *See* LEARNING MECHANISMS; MEMORY.

Paul E. Johnson

Bibliography. A. DeGroot, *Thought and Choice in Chess*, 2d ed., 1978; A. S. Elstein, L. S. Shulman, and S. A. Sprafka, *Medical Problem Solving*, 1978; J. G. Greeno, Trends in the theory of knowledge for problem solving, in D. T. Tuma and F. Reif (eds.), *Problem Solving in Education*, 1980; J. R. Hays, *The Complete Problem Solver*, 1989; P. E. Johnson, What kind of expert should a system be?, *J. Med. Phil.*, 8:77–97, 1983; J. Larkin et al., Expert and novice performance in solving physics problems, *Science*, 208:1335–1342, June 1980; A. Newell and H. Simon, *Human Problem Solving*, 1972.

Proboscidea

An order of placental mammals containing the largest and most powerful living terrestrial mammals—the elephants. Proboscideans once had a wide distribution, reaching every continent except Australia and Antarctica. Although early forms lacked both significant tusks and a proboscis, they are still formally considered proboscideans. *See* ELEPHANT; MAMMALIA.

Classification. The African elephant (*Loxodonta africana*; **Fig. 1**) historically occurred throughout Africa from the Mediterranean Sea to the Cape of Good Hope, except in parts of the Sahara and some

Fig. 1. African bush elephant (*Loxodonta africana*). (*Photo by Gary M. Stolz/U.S. Fish and Wildlife Service*)

other desert regions. Most researchers agree that there are two kinds of African elephant: a smaller, darker forest elephant found in the tropical rainforests of West and Central Africa and a larger, paler bush or savannah elephant in the remainder of the current range. Some feel that these two forms should be regarded as separate subspecies—*L. a. cyclotis* for the forest elephant and *L. a. africana* for the bush elephant; others feel that the forest elephant should be a separate species. In addition, some researchers have identified a pygmy elephant (*L. pumilio*) in the dense lowland jungles from Sierra Leone to Zaire. However, these small forms may just be unusually small adult *L. africana*.

The Asiatic or Indian elephant (*Elephas maximus*; **Fig. 2**) historically occurred from Syria and Iraq east across Asia to Indochina, the Malay Peninsula, China, Sri Lanka, Sumatra, and possibly Java.

A third member of this order is the woolly mammoth (*Mammuthus primigenius*). It once occurred throughout most of Europe and the northern halves of North America and Asia. Mammoths survived up to about 3700 years ago on Wrangel Island in the Arctic Ocean off extreme northeastern Siberia. Intact carcasses have been recovered from permanently frozen ground in Alaska and Siberia. Mammoths had a dense covering of hair and small ears (both adaptations to prevent heat loss), and a trunk that ended in two fingerlike projections at its tip.

Characteristics and description. Proboscideans have a proboscis (trunk), enlarged incisors that form paired tusks, the largest ears of any mammal, and pillarlike legs whose structure prevents the elephant from jumping even a short distance. There are five toes on each forefoot and either three or four toes on each hindfoot. The nose and upper lip are drawn into the long trunk, which is used as a hand. The trunk measures about 6 ft (2 m) in length and weighs about 300 lb (140 kg). It is a boneless mass of flesh and muscles that bends easily. With its trunk the elephant can deftly pick up a peanut or a 600-lb (275-kg) log. The elephant smells, drinks, and feeds itself with its trunk. It depends more on smell than on any other sense. The trunk is used to carry food and water to its mouth. Water is sucked up by the trunk and squirted into the mouth. Elephants feel with the sensitive tips of their trunks. They can tell the shape of an object, and whether it is rough or smooth, hot or cold. They also use their trunk to caress their mate and their young. In a fight, the trunk is used to grasp an enemy. The tusks are modifed incisors that are made of ivory. About two-thirds of each tusk sticks out of the upper jaw; the rest is in the skull. The tusks, which continue to grow throughout the life of the animal, are used to dig for food and to fight. The molars, or grinding teeth, are each at least 1 ft (30 cm) in length and weigh about 8.5 lb (4 kg) each. The elephant has only four of these teeth at any one time. One tooth is on each side of each jaw. New molars form in the back of the mouth and push forward as the old teeth in front wear down and drop out. During its lifetime, an elephant grows six sets of four molars each. The last set comes in when the animal is about 40 years old.

Fig. 2. Asian elephant (*Elephas maximus indicus*) in Kerala region, South India. (*Photo by Nidish K. Viswanathan*)

Elephants are covered with a thick gray hide that is leathery and tough. Because their skin is 2.5 cm (1 in.) thick, elephants are often referred to as pachyderms. In spite of the thick skin, an elephant is very sensitive to cold. Even a slight frost will give it a severe case of cramps. The short stiff hairs on the body may be spaced as much as 1 in. (2.5 cm) apart, and many wear off, resulting in the skin of some animals appearing almost hairless. The enormous ears [3–4 ft (0.9–1.2 m) wide and up to 5 ft (1.5 m) long] serve as fans and are kept constantly in motion during the heat of the day to circulate the air. Elephants have small eyes and poor eyesight, long eyelashes [over 5 in. (13 cm) long], and excellent hearing. They are strictly herbivorous and feed on grasses, leaves, small branches, bark, berries, mangoes, coconuts, corn, and sugarcane. A wild bull may consume 500 to 600 lb (230 to 270 kg) of food per day.

Midway between the elephant's eye and ear lies the temporal gland. It gives off a dark oily material that has a strong smell and stains the elephant's face. The temporal gland becomes active when the elephant is in a condition called musth. At these times, the animal is either dangerously violent or very quiet. Musth occurs in all male elephants and in some females. Some scientists believe it takes place during mating peroids. *See* GLAND.

Differences between African and Asian elephants. How do African and Asian elephants differ? African elephants are taller and heavier. An adult male (bull) is about 11.5 ft (3.5 m) tall at its shoulder, and weighs from 12,000 to 14,000 lb (5440 to 6350 kg). The largest known African bull stood 13 ft 2 in. (4.0 m) tall. An adult female (cow) weighs from 8000 to 10,000 lb (3600 to 4500 kg). An Asian bull stands about 9 ft (2.7 m) tall and weighs from 7000 to

12,000 lb (3200 to 5400 kg). An Asian cow is about 7.5 ft (2.3 m) tall and weighs from 5000 to 10,000 lb (2300 to 4500 kg). The trunks, tusks, and ears of African elephants are much larger than those of Asian elephants. The trunk of an African elephant has two fingerlike knobs on the tip, whereas only one knob is present on the trunk of an Asian elephant. African elephants have slightly hollowed backs; the backs of Asian elephants are arched. African elephants have a more rounded, sloping forehead, whereas the Asiatic elephant has a bulging, protuberant forehead. Most African elephants are dark gray; most Asian elephants are light gray. Both male and female African elephants possess well-developed tusks. Those of a bull grow from 6 to 8 ft (1.8 to 2.4 m) and weigh up to 80 lb (36 kg). The tusks of an Asian bull may grow 4 to 5 ft (1.2 to 1.5 m) long, although the cow's tusks are always small. African elephants are fiercer and more difficult to train.

Behavior. Wild elephants live in herds of 10 to 50 or more. The leader of the herd is usually a female in this matriarchal society. They may travel long distances in search of food or water or to escape pests. They usually travel in single file at a speed of about 6 mi (10 km) per hour. When going down steep slopes, they slide on their bellies, with the back legs stretched out behind and the front legs extended forward. They are good swimmers and often bathe in lakes and rivers and wallow in muddy water.

Reproduction and development. Elephants have the longest gestation period of any animal—20 to 22 months. Pregnancy normally results in one offspring (calf), although occasionally there may be twins. At birth, an elephant weighs about 200 lb (91 kg) and stands about 3 ft (91 cm) tall. The calf nurses by curling its trunk over its head so its mouth can reach the mother's breasts. The calf sheds its milk tusks at 5 or 6 months after birth, but it continues to nurse for another 2 years and remains under its mother's care for 2 years longer. Elephants reach sexual maturity at about 14 to 15 years of age, and attain their adult size when about 20 years old. Most elephants live for about 60 years.

Elephants have few enemies. Tigers in Asia kill calves. Humans are the main enemy through poaching and habitat destruction. Donald W. Linzey

Bibliography. D. Macdonald, *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Procellariiformes

A large order of strictly marine birds found far offshore except when breeding. The procellariiforms, or tube-nosed swimmers, are most closely related to their descendant group, the penguins. *See* SPHENISCIFORMES.

Classification. The Procellariiformes comprise four families: Diomedeidae (albatrosses; 13 species; southern and Pacific oceans); Procellariidae (shearwaters, petrels, and fulmars; 66 species; worldwide); Hydrobatidae (storm petrels; 21 species; worldwide); and Pelecanoididae (diving petrels; 4 species; subantarctic oceans). Some taxonomists have enlarged the Ciconiiformes to include the Procellariiformes, as in the Sibley-Ahlquist taxonomy (a radical new approach to taxonomy based on DNA hybridization studies that differs greatly from the traditional approach). *See* CICONIIFORMES.

Fossil record. The fossil record of procellariiforms is rather incomplete. Most interesting is that a record of fossil albatrosses is known from North America and Europe from the middle Miocene into the lower Pleistocene, even though albatrosses are not now found in the Atlantic Ocean north of the cold waters of the South Atlantic. Their disappearance from most of the Atlantic is, therefore, a recent Quaternary event, and likely related to the lack of suitable nesting islands in the Atlantic Ocean; a fossil albatross colony was recently described from the Quaternary of Bermuda. Few shearwater (Procellariidae) and storm petrel (Hydrobatidae) fossils are known, and little is known of the evolutionary history and past distributions of these groups. One diving petrel fossil dates from the Pliocene of South Africa, where the group is unknown today.

Characteristics. The tube-nosed swimmers are characterized by having their nostrils enclosed in a tube, which is paired in albatrosses; dense plumage; webbed feet, and long wings. They range in size from the sparrow-sized storm petrels to large albatrosses (see **illustration**), which have the greatest wingspan of all living birds, up to 12 ft (3.7 m). The procellariiforms are excellent fliers, as evidenced by the migratory—actually nomadic wandering—flights of many thousands of miles. Albatrosses and shearwaters are dynamic soarers, when their wings are held rigidly away from the body to use the updrafts along wave crests to power their flights. Their flapping flight is slower and more awkward. Their plumage, which varies from white through shades of gray to black, is generally somber but sometimes has a bold pattern. Procellariiforms swim well, and the pelecanoidids and some shearwaters can dive

Black-footed albatross (*Phoebastria nigripes*). (*Photo from Olympic Coast National Marine Sanctuary/National Oceanic & Atmospheric Administration*)

under water, using their wings for propulsion. They are highly pelagic (oceanic) and feed on fish, squids, and crustaceans, which are concentrated into oils and waxy esters in their stomach for transport and feeding to their young, which can digest waxes. Procellariiforms have a well-developed olfactory sense to locate food and apparently to locate their nesting burrows at night.

The larger species mature slowly; some albatrosses begin breeding only after reaching 6 to 8 years of age. Tight pair bonds are formed during courtship, which can take elaborate forms, as is seen in albatrosses. Their nests can be solitary or part of a colony and may be built on the ground as is done by albatrosses, or in burrows as in many shearwaters and storm petrels. The downy young stay in the nest and are cared for by both parents. Development of the young is slow, but the fully grown youngster is very fat, often weighing more than the adults. The breeding season is precarious for procellariiforms, and most species breed on isolated islands. A reduction in numbers has usually resulted from the introduction of mammals such as rats and mongooses which prey on the eggs and young. After the young have learned to fly, they and the adults leave the breeding area and live at sea, usually far from land, until the next breeding season. Some shearwaters and even the small storm petrels follow erratic routes that take them from their subantarctic nesting islands to subarctic oceans. Wandering albatrosses (*Diomedea exulans*) may fly around Antarctica between breeding seasons.

The diving petrels are confined to the colder waters of the southern oceans, and are generally found year-round in the regions of their nesting grounds. They are convergent in structure and habits with the auks (Alcidae) of the Northern Hemisphere. *See* CHARADRIIFORMES.

Economic importance. Aside from being a welcome sight at sea, the procellariiforms have little contact with humans. Several species of tropical albatrosses, however, have been hunted almost to extinction for their feathers and for oil from their fat. A number of species of shearwaters and petrels are seriously endangered or have become extinct because of the introduction of predatory rodents and mongooses onto their breeding islands. *See* AVES; ENDANGERED SPECIES.

Walter J. Bock

Bibliography. M. Brooke, *Albatrosses and Petrels across the World*, Oxford University Press, 2003; J. Warham, *The Behaviour, Population Biology and Physiology of the Petrels*, Academic Press, 1996; J. Warham, *The Petrels: Their Ecology and Breeding Systems*, Academic Press, 1990.

Process control

A field of engineering dealing with ways and means by which conditions of processes are brought to and maintained at desired values, and undesirable conditions are avoided as much as possible. In general, a process is understood to mean any system where material and energy streams are made to interact and to transform each other. Examples are the generation of steam in a boiler; the separation of crude oil by fractional distillation into gas, gasoline, kerosine, gas-oil, and residue; the sintering of iron ore particles into pellets; the polymerization of propylene molecules for the manufacture of polypropylene; the incineration of waste; and the roasting of green coffee beans.

In a wider sense, process control also encompasses determining the desired values. In the narrow sense, it merely refers to keeping process conditions as close as possible to these values, which generally is realized by automation. Modern instrumentation for process control can hardly be thought of without process computers and visual display units.

Process control in the wider sense includes a number of basic functions, which can be arranged as follows:

Scheduling
Mode setting
Mode correction and quality control
Regulatory control Sequence control
Coping with off-normal conditions

Generally, not only do these basic functions determine the performance of conventional process control systems, but they are also predominantly embedded in more advanced integrated intelligent process control systems.

From Conventional to Integrated Intelligent Control

In conventional process control systems, the basic functions are often organized into a hierarchy. At the lower levels, data acquisition is performed directly from the dynamic technical process. Also, basic regulatory control and sequence control are carried out mainly on these levels. Mode correction and quality control as well as coping with off-normal conditions such as fault diagnostics belong to the intermediate levels of the hierarchy. The higher levels include mode setting and scheduling. This bottom-up view of process control is more data-oriented whereas the top-down view emphasizes the goal-oriented perspective.

Advanced process control systems combine the data and the goal orientations by means of a network of data and knowledge flows for achieving integrated intelligent control. The hierarchies of processes, functions, and tasks become less stringent. More overlapping and duplicating functionalities are introduced. The advanced technologies with extended functionalities include expert systems for control and diagnostics, fuzzy process control, and artificial neural networks.

It is important to ensure an integration between conventional control systems and these advanced technologies for the purpose of reaching the overall goals of efficient and safe systems performance as well as sustained and satisfying human operator performance. Thus, the functionalities of well-designed human-machine interaction have gained predominance in advanced process control.

Another extension beyond the conventional control hierarchy is the integration of the advanced process control systems into overall plant control and business management. This organizational extension goes beyond planning and scheduling and relates to the next higher levels of the hierarchy. Thereby, integrated intelligent process control systems are embedded into plantwide control and communication. These developments are changing process control toward human-centered automation. *See* COMPUTER-INTEGRATED MANUFACTURING.

Basic Functions

The basic functions listed above will now be discussed.

Scheduling. At the highest level of the basic control hierarchy, the schedule for process operation is determined, that is, the sequence of the required modes of operation. This basic function indicates, for example, whether product A will be made before B, or vice versa, and when the plant will be shut down for general maintenance. In some cases, scheduling is very important, for example, for continuous distillation units processing a variety of crude oils into different series of intermediate products; in other cases it is practically nonexistent, for example, for the manufacturing of standard solvents. The primary goal is to obtain a feasible schedule, that is, a schedule which does not overflow or empty storage vessels for feedstocks and (intermediate) products. Only if a feasible schedule is possible can one try to find an optimum schedule.

Scheduling can be realized by computer algorithms or by human search-and-try, possibly with the aid of computer simulation. *See* OPERATIONS RESEARCH; SIMULATION.

Mode setting. At the next level in the hierarchy, the desired values are determined for each individual mode of operation. When little is known about the process operation, the mode is characterized by a recipe, obtained from experience. In cases where knowledge is sufficient to set up a mathematical model, optimal desired values can be computed by a suitable optimizing algorithm. Optimization corresponds to maximization of an objective function, which, for continuous processes, might have the form of the following equation:

$$J = \sum_i p_i P_i - \sum_j f_j F_j - \sum_k u_k U_k - \sum_l w_l W_l$$

Here P_i = production rate of product i; F_j = consumption rate of feedstock j; U_k = use rate of utility k (electricity, fuel, cooling water, and so forth); W_l = additional wear rate of process apparatus l; and p_i, f_j, u_k, w_l = corresponding unit values and costs. *See* NONLINEAR PROGRAMMING; OPTIMIZATION.

For integrated production processes, where the product of one process is the feed for another, it is rather difficult to find representative figures for these unit values and costs.

Maximization of the above equation can be interpreted as a compromise between maximum yield of valuable product, minimum consumption of raw material, minimum use of energy, and minimum wear of process apparatus.

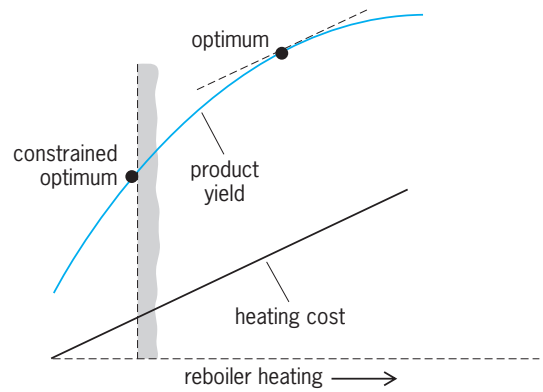

Fig. 1. Optimum operation of distillation column.

Figure 1 shows a simple example. It is the compromise between product yield and reboiler heating for a simple distillation column. More heat costs money but increases the rate of valuable product as well. The optimum is at the point where the marginal value and cost are equal, unless this point is above the maximum capacity of the column. Then, of course, the constrained optimum is at maximum capacity.

The optimal values need not always be the same; they can vary because of changes in feed properties, process load, fouling of process equipment, ambient temperature, and so forth. Here an on-line computer (directly coupled to process sensors) is a suitable device for the computations. Expert systems with their extended knowledge-based functionalities may be applied. The results can be shown to the process operators (advisory control), or they can be automatically transmitted to the automatic control system (supervisory control).

Other examples of constrained optima are controlling the catalyst temperature in a chemical reactor near the sintering point for maximizing product yield, and setting the condenser "wide open" in order to minimize heat requirements in distillation. *See* DISTILLATION.

Special cases are mode switches. For instance, a polyethylene reactor has to be switched from producing a lower molecular weight to producing a higher one. The purpose is to minimize the quantity of intermediate material, that is, to move as rapidly as possible from one to the other mode of operation. Here the optimum desired values depend on the time; they follow a path which can be found by dynamic optimization techniques. *See* CALCULUS OF VARIATIONS.

Analogous results are obtained for batch processes where the sequence of operation steps is done in the same process apparatus. For instance, a batch chemical reactor is started with maximal heating, followed, at a critical moment of time, by a switch to maximum cooling in order to reach the desired reaction temperature without overshoot in minimum time.

Some continuous processes appear to work more efficiently when they are subjected to cyclic variations than when under equilibrium conditions. Evidently, the optimum desired values here are a periodic function of time.

Mode correction and quality control. The next level in the hierarchy is quality control. In some cases, product quality can be completely specified by concentration, for example, less than 1% impurities. Alternatively, some physical property is a good measure of quality, for example, density or average molecular weight. In still other cases, a performance test has to be done, which bears some resemblance to the way in which the customer uses the product. A straightforward example is the octane number for gasoline, which is determined in a standardized internal combustion engine.

Traditionally, quality measurements are done in a control laboratory by skilled analysts. This incurs a rather long delay, as samples have to be brought to and routed inside the laboratory. A faster response time is achieved by local semiautomatic instruments which can be handled by process operators. Still better results can be obtained with completely automatic quality analyzers, which take their own samples. However, these instruments are available only for the more common and straightforward applications. Moreover, they are rather expensive and require much highly skilled maintenance. Consequently, in practice, quality control often has to be done with sporadic and late quality data. In between, process operators try to estimate product quality changes from the values of more easily measured variables. They can be helped by the on-line computer, provided some type of estimation model is available (**Fig. 2**).

In discrete-parts manufacturing, statistical techniques are commonly applied to supervise product quality. Applications in the process industries are being generalized to statistical process control as a tool to assist the operator in correcting mode settings. However, superior results can be expected from estimation techniques (Fig. 2). *See* QUALITY CONTROL.

Regulatory control. Most process variables can be measured with inexpensive and reliable sensors. Common ones are temperature, pressure, level (liquid or solid), flow rate, and rotational speed. It is customary to maintain important process variables of these types near desired values by automatic control.

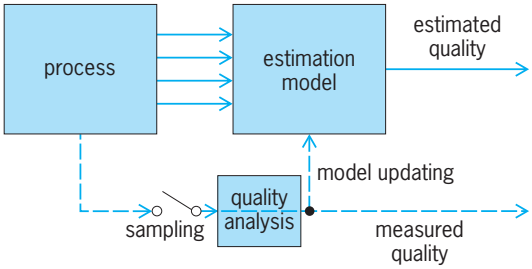

Fig. 2. Estimation of product quality by on-line computer.

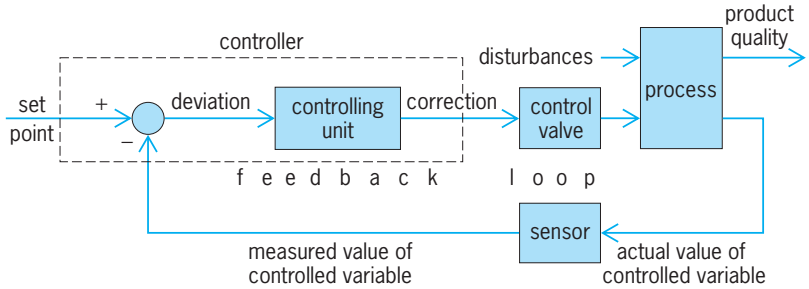

Fig. 3. Block diagram illustrating feedback loop around process.

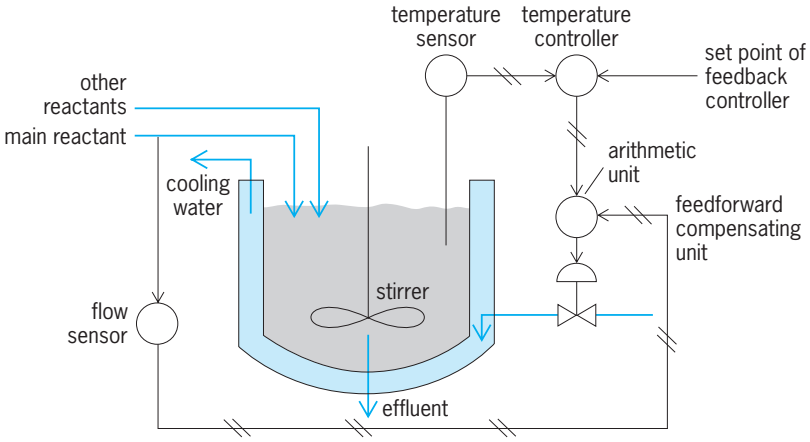

Fig. 4. Feedback and feedforward control on continuous stirred chemical reactor.

Feedback control. The most commonly used method is feedback. In an automatic feedback control system (**Fig. 3**), a process variable such as the temperature in a continuous stirred chemical reactor (**Fig. 4**) is to be maintained at a desired value; hence it is called the controlled variable. It is measured by a sensor, for example, a thermocouple. The measured value is compared with the set point, which can be manually adjusted. The difference between the set point and the measured value, called the deviation, is the input of the controlling unit. Here a correction is generated that should eliminate, or at least reduce, any difference between set point and measured value. The correction influences an input of the process, the manipulating variable (in Fig. 4 the rate of cooling water supplied to the reactor jacket) via a control valve. *See* THERMOCOUPLE.

The process operator still is the master. This worker is relieved from the task of continuously introducing corrections, but still retains the overall supervision. When, for instance, the product quality deviates from the desired value, the process operator will adjust the set point in order to reduce quality deviations.

When a reliable automatic quality analyzer is available, the operator's supervisory action can be partly replaced by a master control loop. The result is a so-called cascade control system, with the accompanying reduction in the operator's work load.

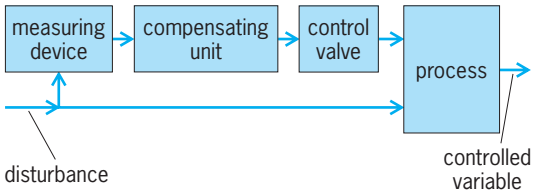

Fig. 5. Block diagram of typical feedforward control.

Feedforward control. Feedforward control is another principle, which is often applied to supplement feedback control (**Fig. 5**).

A process variable that is a source of disturbances (in Fig. 4 the feed rate of the main reactant) is measured, and the measured value is made to influence the control valve via a compensating unit. The compensating unit is adjusted in such a way that the direct influence of the disturbance on the controlled variable is balanced by the indirect influence (through the measuring device, compensating unit, and control valve). In Fig. 4 an increase in main reactant rate will gradually increase the reaction rate in the reactor, hence generate more heat of reaction. A gradual opening of the control valve introduces more cooling, which takes care of removal of the additional heat of reaction; hence the temperature remains constant.

With feedforward control, perfect compensation of disturbance effects is possible in principle. (With feedback control this cannot be achieved because the controlling unit can only generate a correction when there is a deviation.) Perfect compensation, however, requires detailed knowledge of the process dynamics (that is, the way in which processes respond to disturbances and corrections as time progresses). Any significant mismatch between the direct and the indirect influences causes rather severe deviations (100–80 instead of 100–100 makes 20 instead of 0). This does not hold for feedback control, where even large changes in process dynamics have only little effect on performance (a change of 20% does little harm). Another disadvantage of feedforward control is that the designer must provide compensating units for all important sources of disturbances. In view of these difficulties, it is usually more practical to use a combination of feedforward and feedback control, whereby the former helps the latter in reducing the effects of the most important disturbances.

Features of automatic regulatory control. Automatic regulatory control has the following features:

1. It relieves the process operator of tedious time-consuming control tasks, enabling a few people to supervise and control large production plants.

2. Compared to manual regulatory control, performance is more consistent and reliable. This has led to plants with smaller residence times and less buffering, thus decreasing costs and improving safety (smaller amounts of materials under hazardous conditions).

3. Via a judicious choice of controlled variables, product quality variations are reduced, which simplifies quality control.

4. Reduction of the effects of disturbances decreases the chance that the plant runs into off-normal or unsafe conditions.

Sequence control. What regulatory control does for continuous process operation, sequence control does for batch-type operation. In its simplest form, sequence control resembles the actions of the program in a washing machine: at predetermined moments of time, valves are opened or closed, motors are started or stopped, and so on.

In more sophisticated versions, the timing is not fixed, but depends on process conditions: in a chemical reactor, catalyst injection is started after the reactor contents have reached a certain temperature, which may take a shorter or longer time. Further, logical conditions are included too: a centrifugal pump is started only if the suction valve is open and the discharge valve is closed.

Automatic sequence control can be realized with programmable logic controllers (PLCs; a digital device which performs logical operations). It enables one process operator to supervise a series of batch processes, relieving the worker of repetitive tasks. Moreover, the accurate reproduction of mode recipes enhances constancy of product quality. *See* PROGRAMMABLE CONTROLLERS.

Automatic sequence control is also being applied in continuous processes for operations like washing of filters, regeneration of catalysts, and switching of heat exchangers. In some processes, for example, electric power generation, automatic sequence control also takes care of startup and shutdown.

Coping with off-normals and faults. This level is of vital importance for safety and plant availability. It is devoted to the detection, diagnosis (identification, interpretation), and compensation of off-normal process conditions and faults in process equipment and instrumentation. Diagnostic functionalities constitute the most important and most difficult phase of fault management.

Well-protected plants have five barriers against hazardous situations:

1. Automatic regulatory and sequence control systems reduce deviations from desired values, hence decrease the probability of undesired values.

2. The operator observes trends in important process variables in order to detect and identify the start of off-normal conditions.

3. If the previous barriers have been ineffective, an alarm signal is generated, usually in auditive and visual form, to alert the operator. The choice of the alarm limit is a compromise between allowing the process operator time to take action and causing too many false alarms.

4. If the operator does not take action, an automatic protection system shuts down the plant or brings it to a safe condition (standby).

5. If the automatic protection system fails to act, the process operator is the last line of defense. The first three barriers are intended to keep the

process going, while the last two barriers resort to standby or shutdown, with accompanying loss of production.

Advanced Technologies

The extended functions of advanced process control that were mentioned above will now be discussed.

Expert systems. In integrated intelligent process control systems, increasing demands and needs of end users are handled by control-system modules with overlapping and extended functionalities. These functionalities are well suited for integration with the above levels of scheduling, mode setting, and coping with off-normal conditions. The knowledge and experience of technical experts are incorporated into the process control system, enabling it to function at a higher intelligent level. Supervisory control and diagnostic functionalities are particularly suited for implementation with real-time expert systems. The expert knowledge is represented explicitly in the knowledge base, while an inference mechanism is responsible for problem solving under a wide range of circumstances. The inference or reasoning mechanism leads to better problem-solving results when the knowledge representations in the knowledge base are more adequately capturing a multitude of process conditions and, correspondingly, a rich repertoire of well-suited human operator behaviors.

Several forms of deep and shallow knowledge may be available, such as model-based knowledge or embedded simulations, rule-based knowledge, and frame-based knowledge with the inheritance of object attributes and values between different levels of an object hierarchy. Thus, procedural as well as declarative knowledge can be represented in the knowledge bases. Several knowledge bases are generally necessary, covering areas such as process plant components, process states (particularly macrostates), failure cases and failure symptoms, fault diagnosis strategies, and control laws. The knowledge to be incorporated into the knowledge bases has to be elicited from different classes of technical experts, such as operators, engineers, and maintenance personnel. Knowledge elicitation and acquisition is an empirical social interaction process between experts and knowledge engineers. It takes much time, needs much experience to avoid pitfalls, and is therefore sometimes called the bottleneck of knowledge engineering and expert systems design. *See* OBJECT-ORIENTED PROGRAMMING.

The special features of expert systems for industrial applications of dynamic processes include their real-time behavior as well as their online or embedded functionalities within the whole technical system. Process variables are viewed mainly as symbolic expressions in linguistic terms. For example, a temperature is denoted as "high" if the actual numerical value is anywhere between two relevant boundaries. The real numerical values are not of direct interest during the problem-solving process. The coexistence of particular linguistic values of several process variables may define characteristic macrostates of the process control system. Since the variables change their values asynchronically, the real-time aspects of process control require a buffering and regular updating of process variables at the instances of time stamps between successive macrostates. This ensures piecewise-constant conditions for the performance of the reasoning process.

The application of expert systems in supervisory control allows the knowledge-based guidance of the process along desired trajectories between different steady states. Different kinds of regulatory control mechanisms can be influenced by the supervising heuristic expert control. In fault management situations, the diagnostic process can be handled in several steps or phases. After the actual state of conditions is identified, the two steps of hypothesize-and-test of possibly faulty system components are performed. An example is the diagnosis of technical faults in a preheater system of a power plant. Learning mechanisms for improving control through experience are also feasible. *See* EXPERT SYSTEMS.

Fuzzy process control. The methodologies of fuzzy set and systems theory are now applied in process control. Several possibilities for such applications exist. On a higher intelligent level, problem solving and decisions for scheduling and diagnostics are sometimes fuzzy, rather than crisp and precise, in the overall control of complex industrial processes. The process variables themselves may be transformed into fuzzy linguistic variables, based on the degree of membership with which each numerical value of a measured process variable belongs to one or several overlapping linguistic terms. This transformation from precise measurements to fuzzy linguistic variables is called fuzzification.

The inference mechanism of a rule-based expert system may also be described with fuzzy rules and relations. Satisficing rather than optimizing a problem solution is often enough, and sometimes is the only achievable possibility in a complex practical context. Heuristic knowledge for supervisory control and diagnostics can often be formalized only with fuzzy rule-based reasoning. In such cases, evaluation is not restricted to a single path through the rule base, and, indeed, several rules are activated and several inferencing paths contribute to the end result of a problem solution. This end result is expressed in terms of membership functions of linguistic terms. A transformation is then needed into linguistic expressions that are understandable by the human user of the decision outcome in the application context. All definitions of linguistic terms, all membership functions, and all fuzzy rules and relations have to be derived from human expert knowledge by means of empirical studies.

On a lower level nearer to the technical process, regulatory control can be performed with fuzzy logic controllers. As in the case of fuzzy expert systems, precise quantities are changed into linguistic variables by fuzzification (**Fig. 6**). The inference mechanism is now the fuzzy controller, which consists of

linguistic IF-THEN rules such as:

IF temperature is "positive big" or "positive medium" and

IF temperature rate is "negative very small"

THEN coding input change is "positive medium".

A fuzzy decision combines several conclusions of the THEN parts whenever a number of rules are applied in parallel for a fuzzy-logic-based control policy. Most industrial processes cannot interpret fuzzy command inputs. Therefore, the fuzzy controller and decision output need to be transformed into precise control quantities by the defuzzification operation. *See* FUZZY SETS AND SYSTEMS.

Artificial neural networks. Another knowledge-based technology of generally high potential is available with artificial neural networks. Whereas expert systems contain more of the "know-how" type of knowledge such as in heuristics and principles, artificial neural networks are more data-driven and thus represent the "know-what" type of knowledge. Large sets of data such as multi-inputs and corresponding multioutputs of dynamic technical processes can be observed as vectors or patterns of data and can appropriately be matched by the general pattern-matching capabilities of artificial neural network (ANN) models.

In no way do the parameters of these models relate to any type of explicit knowledge, as in expert systems. The knowledge of artificial neural network models is implicitly contained in the combination of the large number of parameters, the so-called weights, in the overall web-type network.

Figure 7 illustrates the most common use of artificial neural networks in process modeling and control. The process input data are the parallel control inputs at certain moments in time, and the process output data are the corresponding parallel process variables, such as temperatures and flow rates. Empirical models based on process input-output data can be developed, particularly for nonlinear and poorly understood processes such as those found in many chemical and biotechnological plants. Artificial neural network models are powerful empirical models with good approximation capabilities. Typically, neural networks with three layers of artificial neurons are applied: the input layer, the hidden layer as a kind of intermediary, and the output layer.

The weights of neurons in the hidden layer and in the output layer can be tuned by means of a learning or training algorithm. The input-output data of the process are used as training data, and the output data of the artificial neural network model are compared with the process outputs (Fig. 7). The back-propagation algorithm is the most commonly employed learning algorithm in process control applications. It is a gradient descent optimization technique developed for suitable use in multilayered iterative updating of the weights. The knowledge about the process behavior is implicitly represented in the combination of all these weights. However, an application-related interpretation of the parameters is not feasible.

Once the artificial neural network model has been fine-tuned, it can be used to predict outputs in regulatory process control. Various control structures with one or two artificial neural network models have been applied in various combinations with control theoretical algorithms to realize a large number of process control applications. Process monitoring and fault diagnosis can also be supported. Several distinct faults in different process components, such as a leakage in a condenser, can be predicted by different artificial neural network models, each with a typical match of its input-ouput patterns as a representation for the corresponding fault. Practical installations of artificial neural network models for process modeling, control, monitoring, and diagnosis can be found in chemical plants, robots, steel plants, and other applications. Further, neural networks have been combined with fuzzy logic to form neuro-fuzzy networks which achieve better interpretations and robustness than artificial neural network models alone. This is accomplished by a clearer modeling structure, for example, by partitioning the process operation into several fuzzy operating regions. *See* NEURAL NETWORK.

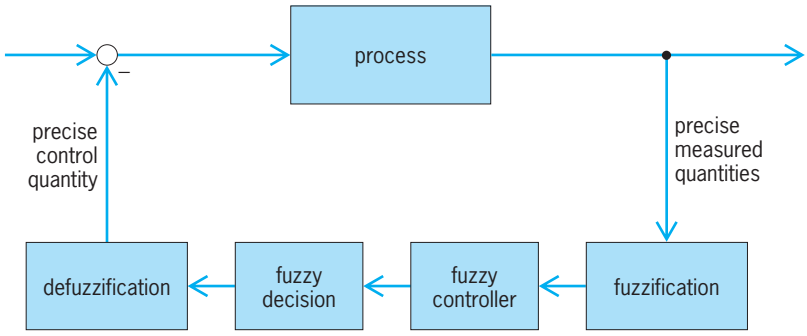

Fig. 6. Closed-loop fuzzy process control.

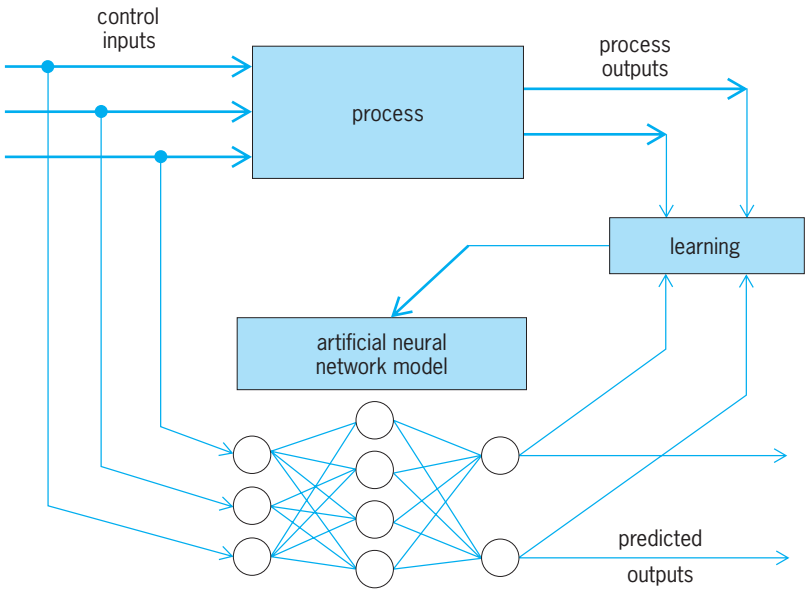

Fig. 7. Neural network for process modeling and control.

Process Control Components

The hardware and software implementations of process control systems are accomplished with the integration of a large number of components.

Sensor and actuator systems. The sensor systems and the actuator systems make up the process-near instrumentation. Pneumatic, mechanical, electronic, and mixed types of instrumentation have been realized for practical process control applications. All relevant outputs of a particular technical system, the process variables, have to be measured by suitable sensors. Product and process properties have to be continually measured in real time. Examples of the large number of possible process and product variables to be measured are pressures, temperatures, flow rates, rotational speeds, mass flows and levels, energy consumptions, viscosities, and concentrations. Different physical principles are applied to realize a variety of sensor technologies, including optical, acoustic, inductive, capacitive, radiometric, hydrostatic, chromatographic, and spectrometric. *See* TRANSDUCER.

The actuator systems transform control signals, mainly electrical outputs of a digital control computer, into the appropriate physical forms of control inputs for the process behavior which has to be influenced. Servo drives such as motors are used to operate the final actuator elements. These can be valves, pumps, heating devices, and all kinds of switches and contacts. *See* SERVOMECHANISM; VALVE.

Distributed control systems. A hierarchical and decentralized organization of all process monitoring and control functions is realized by means of distributed control systems. The required functions include controlling and supervising the process, planning and diagnosing, improving safety and reliability, and displaying information to human users as well as recording and evaluating.

The most common type of information transmission in distributed control systems is the serial data bus. All sensors and control valves are coupled via microcomputers to a bidirectional high-speed transmission line, for example, a coaxial cable or an optical fiber (**Fig. 8**). The microcomputers convert signals from sensors into binary numbers, and put these on the transmission line as a sequence of short (voltage or light) pulses, for example, each of 1 microsecond duration. At the other side the pulses are picked up and stored in a digital register for introduction into a computer. An analogous series of steps takes place between computer and control valves. *See* COAXIAL CABLE; OPTICAL FIBERS.

As all data are transmitted along the same transmission, sender and receiver have to find each other by a communication procedure, coded in pulse trains, the so-called protocol. By way of example, assume that the computer always acts as bus master, both when requesting and when supplying data. Here the computer starts a transmission with an initialization code, which alerts all sensors and control valves, followed by an address code, which activates a particular device. This device responds with a ready code, to which the computer reacts with a request data code. The device then sends the information, if necessary ending with an end data code, and returns to the listening state (to be ready for the next initialization code). More complicated protocols are required when the bus mastership can transfer from one device (or computer) to another. *See* DATA COMMUNICATIONS; LOCAL-AREA NETWORKS.

The main feature of this type of data transmission is the saving in cabling and installation costs. With traditional transmission, thousands of pneumatic tubes or electric wire pairs have to be installed and correctly connected. With serial transmission, only one cable (or optical fiber) runs along all devices in the plant, and insertion of address codes replaces tedious testing of wire (tube) connections.

However, a single serial transmission is too risky: when it fails, the complete instrumentation system breaks down. It is therefore desirable to achieve redundancy by installing at least two independent serial transmissions, which follow individual paths to all devices. If one fails, the other can take over.

The data load on the transmission and communication system can be decreased by incorporation of local data processing (Fig. 8). A number of sensors and control valves, pertaining to a small part of the production processes, are connected to a local computer, which takes care of basic information and control functions. Further decentralization can be achieved by so-called smart sensors, which are provided with microcomputers for linearization, calibration, and fault identification. *See* MICROCOMPUTER; MICROPROCESSOR.

Figure 9 gives a schematic example of a modern distributed control system. It contains many computers for a wide variety of functions, both in local and in central data processing. The latter part is interconnected by a parallel bus, which allows a higher transmission rate than a serial bus. The highest level in this hierarchy of decentralized functionalities is accomplished by the production computer. On the intermediate levels, the medium and lower basic

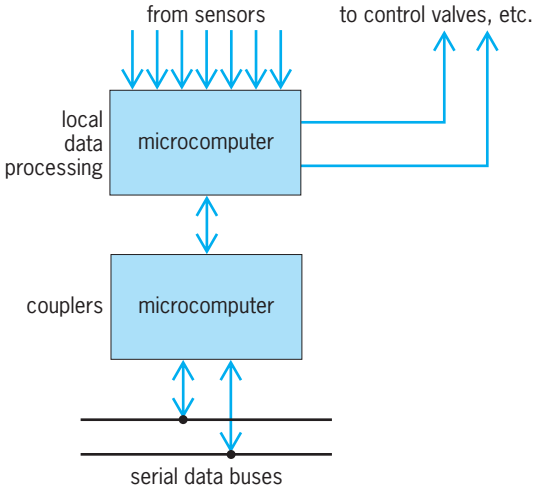

Fig. 8. Local computer system.

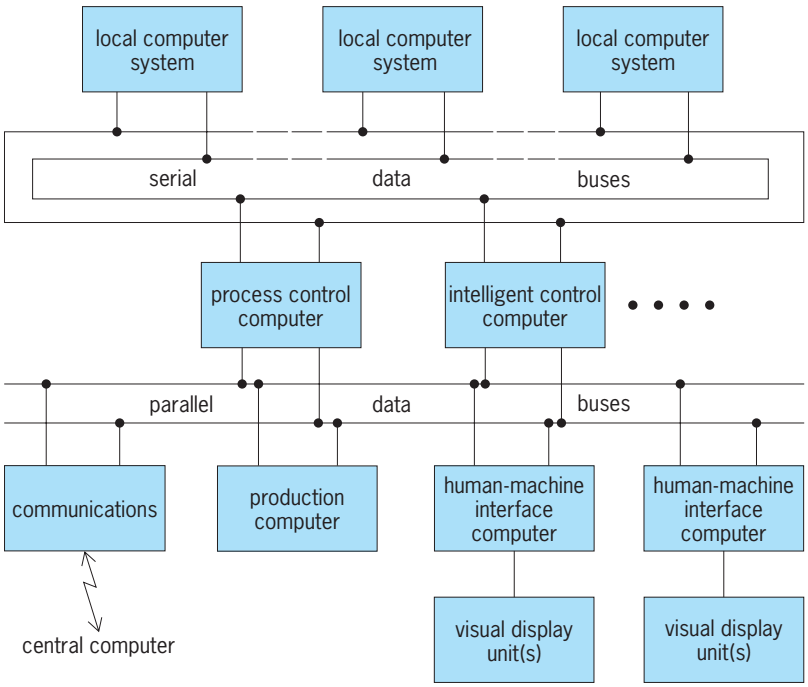

Fig. 9. Schematic example of modern distributed control system.

functions are realized with the process control computer (in the narrower sense), the extended functions by means of the intelligent control computer, and the safety as well as the recording functions with the same or, optionally, additional dedicated computers. The implementation of distributed control systems with a hierarchical decentralized organization and with supervisory control functionalities supports the integration between conventional and integrated intelligent control systems. Further, modern distrbuted control systems have provisions for coupling to office computer networks and even to distant computers.

A very important feature of distributed control systems is system software for configuration of regulatory and sequence control, as well as of higher-level and intelligent control, and for display of process information. In this way, flexibility is combined with efficiency. *See* DISTRIBUTED SYSTEMS (CONTROL SYSTEMS).

Information systems. It is often difficult to make a clear distinction between process control systems and process information systems. As automation never is complete (even space satellites are supervised from Earth), there is interaction between process operators and automation systems, which, among others, requires an operator information system. Moreover, individual computers for data processing in various departments of the factory (or the enterprise) are coupled into an integrated information and control system. In this way, sales, design, internal transport, storage, packaging, expedition, accounting, maintenance, and management send and receive information to and from production. For instance, customer orders are entered into the system in the sales department, and they appear at the production scheduling section, where the sequence of operations is determined for production processes, based on information from storage.

The practical realization of information systems comprises computer networks and management, communication structures and protocols, knowledge and data mining technologies, and quality-assurance architectures and tools. In information systems it is very important to select the type and form of information in accordance with the needs of the user. A plant manager requires quite different reports than a process operator. Systematic user participation during the phases of design and development of process control and information systems becomes a strong requirement.

Human-Machine Interaction

Computerized instrumentation has revolutionized the interaction with plant personnel, in particular the process operators. Traditionally, the central control room was provided with long panels or consoles, on which alarm lights, indicators, and recorders were mounted. Costs were rather high, and surveyability was poor. In computerized instrumentation, visual display units are the main components for information presentation in the human-machine interfaces (Fig. 9). They can provide information in a concise and flexible way, adapted to human needs and capabilities. Many visual display units contain cathode-ray tubes, similar to those used in television sets (**Fig. 10**). *See* CATHODE-RAY TUBE.

A computerized visual display unit offers the designer greater freedom in structuring process information. Process flow schemes can be displayed which indicate the state of process equipment (on/off), alarms, and values of process variables (information for supervisory control).

Historical information can be shown in sectors of the screen. Operators can be offered a choice of display structures and even facilities to generate formats to their individual preferences.

Process flow schemes are visualized for the human operators as topological views on the process in a component-oriented fashion. User-oriented and goal-oriented human-machine interfaces can additionally supply functional knowledge. Much more information processing than previously is required in these human-machine interfaces. Goals to be achieved can be related to functions of the process, which again relate to the components of the process. Multimodel-based approaches use fuzzy logic and qualitative modeling adapted to human needs. Different abstraction levels and different degrees of detail can be visualized. The process visualizations include representations of the fulfillment degrees of goals, the violation degrees of conditions, flow functions of mass and energy balances, and ecological gestalt graphs for the relations between process states and control inputs. Some of these process visualizations are fairly abstract and, therefore, the user needs to be able to navigate between them and the more traditional

Fig. 10. Central control room of the ESSO refinery, Rotterdam, Netherlands. (*RO Studio, Vlaardingen, Netherlands*)

topological views within one human-machine interface.

Intensive participation of prospective end users as well as usability testing is necessary in order to design more advanced human-machine interfaces with supportive knowledge-based functionalities. This participative design approach can be extended further toward the consideration of users' needs early in the development of the whole process control system, ultimately leading to the ideal of human-centered automation.

Other developments in human-machine interfaces for process control include large projection screens and multimedia presentations, in most cases in addition to the process visualizations with visual display units. The main objective of the multimedia approaches for process control is to make the best use of human sensory modalities in addition to vision. This utilization will avoid the visual overload of human operators. Also, it can bring the operator back closer to the environment where audio signals, such as noises from process components, and vibrations can be exploited for transmitting important information. *See* MULTIMEDIA TECHNOLOGY.

Instrumentation system designers can obtain criteria and results about the human-machine interaction from human-factors engineering. *See* HUMAN-COMPUTER INTERACTION; HUMAN-FACTORS ENGINEERING; HUMAN-MACHINE SYSTEMS.

Plant-Wide Control and Communication

Modern distributed control systems offer an infrastructure that meets most needs for process control in the wide sense. New developments toward human-centered automation can be realized by combining basic and intelligent control and communication technologies with human-machine and cognitive systems engineering. Communication with the central office and the outside world can support integration of various production activities in the industrial enterprise. Of particular interest are interaction between processing and packaging (or other discrete production activites); integration with logistics control; integrated plant-wide, factory-wide, or enterprise-wide information systems; and combined systems for process model development and recipe improvement. The main problems encountered in these activities are not of a technical nature, but are related to reorganization and to training and changing the attitudes of all involved people. *See* AUTOMATION; CONTROL SYSTEMS.

Gunnar Johannsen

Bibliography. F. Jovic, *Expert Systems in Process Control*, Kluwer (Chapman & Hall), Dordrecht, 1992; T. E. Marlin, *Process Control*, 2d ed., McGraw-Hill, New York, 2000; M. Polke (ed.), *Process Control Engineering*, VCH Verlagsgesellschaft, Weinheim, 1994; W. F. Ramirez, *Process Control and Identification*, Academic Press, Boston, 1994; M. Rao and H. Qiu, *Process Control Engineering*, Gordon & Breach Science, Newark, NJ, 1993; J. E. Rijnsdorp, *Integrated Process Control and Automation*, Elsevier Science, Amsterdam, 1991; T. J. Ross, *Fuzzy Logic with Engineering Applications*, McGraw-Hill, New York, 1995; F. G. Shinskey, *Process Control Systems*, 4th ed., McGraw-Hill, New York, 1996; S. G. Tzafestas (ed.), *Methods and Applications of Intelligent Control*, Kluwer, Dordrecht, 1997.

Process engineering

A branch of engineering in which a process effects chemical and mechanical transformations of matter, conducted continuously or repeatedly on a substantial scale. Process engineering constitutes the specification, optimization, realization, and adjustment of the process applied to manufacture of bulk products or discrete products. Bulk products are those which are homogeneous throughout and uniform in properties, are in gaseous, liquid, or solid form, and are made in separate batches or continuously. Examples of bulk product processes include petroleum refining, municipal water purification, the manufacture of penicillin by fermentation or synthesis, the forming of paper from wood pulp, the separation and crystallization of various salts from brine, the production of liquid oxygen and nitrogen from air, the electrolytic beneficiation of aluminum, and the manufacture of paint, whiskey, plastic resin, and so on. Discrete products are those which are separate and individual, although they may be identical or very nearly so. Examples of discrete product processes include the casting, molding, forging, shaping, forming, joining, and surface finishing of the component piece parts of end products or of the end products themselves. Processes are chemical when one or more essential steps involve chemical reaction. Almost no chemical process occurs without many accompanying mechanical steps such as pumping and conveying, size reduction of particles, classification of particles and their separation from fluid streams, evaporation and distillation with attendant boiling and condensation, absorption, extraction, membrane separations, and mixing.

Bulk Products

Process engineering begins with a need for a product and some sense of the required scale of production. The end product is always defined; frequently the desired raw materials can also be stipulated at this stage, although sometimes it is very important to consider wholly different sources (for example, motor fuel from coal, petroleum, or renewable resources) and the degree of vertical integration desired—for example, should the process begin with natural products or purchased intermediates. Then one or more process flow sheets are drawn, depicting the series of steps through which the raw materials must pass in the course of transformation into product. On the flow sheet, all lines represent movement of material from one process step to another, and all lines must show the intended direction of flow. Each processing step is represented by an enclosed area (circle, rectangle, and so on) on whose edges the lines terminate (unless only heat exchange is intended, in which case at least one line goes through the area unbroken to indicate thermal contact without mixing). Each process flow sheet envisions a definite sequence of operations, including the selection of separation processes and choice of chemical reactions. The sequence may contain subsequences that are serial or parallel or enclosed by a recycle stream.

Batch versus continuous operation. In designing almost any process, an early choice between batch and continuous operation must be made. Small amounts of material can be taken through a series of process steps over time, often using the same apparatus (for example, a thermally jacketed stirred tank) for several or even all steps. Commercial production in batches is feasible and often preferred when small amounts are to be manufactured, when a sequence of many similar reaction steps is envisioned, or when it is feasible for multiple products to be manufactured in the same equipment. However, batch operations are labor-intensive and more difficult to control, analyze, and design (if done with the same care given to continuous processes). Continuous operation is easier to analyze, design, and control, but the plant is expensive, unique to the product, and difficult to operate effectively at other than design rate.

Material and energy balances. Once the process flow sheet is drawn, it is possible—for either batch or continuous processes—to generate material and energy balances. These are drawn for the overall process and for each step on the basis of conservation and rate laws as well as certain assumptions, some of which are only temporary. For example, in a mechanical step, such as dewatering of a fiber slurry during paper manufacture, a dilute stream containing 1% dry fiber and 99% water may be presumed to enter a separator at the rate of 22,000 lb/min (10,000 kg/min), to be separated into two product streams, one containing 99% of the fibers fed at a concentration of 5% fibers and the other, a recycle water stream. The proportioning of the fibers between the streams and the composition of the one outlet stream are assumptions that may well be adjusted at later stages. In this example, the conservation laws require that water and fiber be separately conserved since no reaction is envisioned, and it is immediately possible to calculate that the concentrated fiber stream emerges at the rate of 4360 lb/min (1980 kg/min), while the balance of the 22,000 lb/min (10,000 kg/min) leaves as a stream containing about 0.0125% fiber. In a complex process, many hundreds of such steps must be assessed and matched to each other. [In the case of the process step just considered, accepting the 22,000 lb/min (10,000 kg/min) stream constitutes matching this step to a predecessor; and two subsequent steps must be matched to the two exit streams just considered.] When the assessment and matching of individual steps are complete, the overall result must be considered. Is the production rate correct? Is the yield of product satisfactory, or has too much material been turned into by-products or lost in waste streams? Has too much water or heat been used? Is the product pure enough? Frequently one or more of these targets are not met. Then assumptions must be revised and the computations repeated. Sometimes the flow sheet itself must be changed: additional steps added, a new recycle stream provided,

one kind of operation replaced by another. Because of the complexity of this process, it is often executed with the aid of computers. Once a flow sheet has been fixed and consistent heat and material balances reckoned, the duty of each process unit is established, and their individual design can be commenced.

Process units and reactors. For many years chemical engineers have recognized that the design of process units was much more dependent upon the kind of operation under consideration (for example, heat transfer, or fluid flow, or distillation) and much less dependent on what was being processed or where in the sequence of process steps a particular unit occurred. Thus evolved the concept of unit operations and a set of principles and rules for design of process units according to their type of unit operation rather than according to the process itself. Notable among unit operations that have been extensively developed are the membrane processes (reverse osmosis, dialysis, electrodialysis, and electroosmosis), techniques for separating solids from fluids, and methods of controlling transport across phase boundaries such as freeze drying.

Reactor design. In completing the engineering of a process, many features peculiar to it must be considered, such as the selection of pipes and tanks strong enough to withstand the temperatures and pressures of operation and the choice of materials that will not be corroded by the materials being processed. Because the chemical reaction is the most critical step in most chemical processes, reactor engineering is a very important component of process engineering. Chemical reactions may be carried out either with a single phase (homogeneously) or near phase boundaries and interfaces (heterogeneously) with or without a catalyst. Process reactors may be stirred tanks containing one fluid, one fluid dispersed in another, or a suspension. Some reactors are fluidized beds: fine particles which may be either reactant or catalyst kept suspended and moving by a flowing gas or liquid. Tubular reactors are used frequently, filled only with fluid, or packed, or equipped with complex so-called internals. Many reactors are jacketed or otherwise equipped for removing heat. Some reactions are conducted in flames or in illuminated chambers. However, in spite of this great geometric variety, general principles of reactor engineering have been developed, based on expressions for the rates of chemical reactions and the transport of molecules and heat as well as the laws of conservation for mass and energy.

Integration and optimization. The widespread availability of computers has led both to higher integration of processes and to much more nearly optimal design. Integration refers to the increased use of recycle streams to recover heat or material that would otherwise be discarded. It is motivated both by possible economies and by its ability to decrease the burdens imposed on the environment by a process. Integration makes design more difficult, usually results in a higher capital cost for the plant, and may make it more difficult to control. Modern petroleum refineries tend to be particularly highly integrated. Optimization is a relative term. In practice it denotes computation of the best split of streams and duties among units and selection of the best operating conditions for each unit to optimize some overall quantity (often plant profitability) subject to constraints (such as environmental restrictions, plant size, usage of certain utilities). *See* OPTIMIZATION.

Extreme conditions. Extremes of high temperature and high and low pressure have been extended, so that reactions can be carried out in plasmas at temperatures of 17,500°F (10,000 K), while pressures may go as high as several hundred thousands of pounds per square inch (several gigapascals) and as low as 10^{-10} torr (10^{-8} pascal). Extension of extremes of feasible operation has been facilitated by development of new materials of construction.

Economic considerations. Process engineering is an activity supported almost entirely by industry, and is thus greatly influenced by economic considerations, much augmented however by regulatory considerations designed to minimize environmental impacts, protect workers, guarantee the safety and efficacy of products, and carry out governmental policies for conservation of resources, especially energy. Still, economics figures heavily in process design: in choosing to make a product, in specifying the degree of vertical integration, in fixing the scale of operations, and in balancing capital costs against operating costs. Other economically significant decisions include estimating the desirable operating lifetime of the process and plant and assessing the degree of risk to be taken in using attractive but not fully proved new technology in the process realization. Thus process engineering utilizes not only principles from the chemical and physical sciences but also detailed assessments of the state of commerce and the economy. *See* DIALYSIS; DISTILLATION; EVAPORATION; EXTRACTION; ION-SELECTIVE MEMBRANES AND ELECTRODES; MECHANICAL CLASSIFICATION; MECHANICAL SEPARATION TECHNIQUES; MIXING.

Edward F. Leonard

Discrete Products

The process engineer must consider all available processes capable of changing the shape of a raw material to the geometrical configuration specified on an engineering drawing. Thus, the process engineer must consider not only those processes within his or her own plant but also those in vendor or subcontractor plants. Usually, several ways exist to produce any part, one being best for a given set of conditions. The process engineer has the responsibility of determining which is best. The principal constraints to be considered in the selection of a basic process in order to bring raw material more closely to the specifications of a functional design are (1) type of raw material; (2) size of the raw material that the equipment involved can handle; (3) geometric configurations that equipment characteristic of the process is capable of imparting to the raw material; (4) tolerance and surface finish capabilities of the equipment; (5) quantity of finished parts needed and their

TABLE 1. Part of guide sheet for geometry and other primary factors

Process	Applicable geometry*	Intricacy of geometry	Applicable materials†	Minimum lot for which process is economical	Secondary operation cost	Decision equation for primary process	Expected rate of production per hour	Expected time in weeks to get into production
Die casting	1,2,3,4,5, 6,7,8,9	*X*	*t*	3000	1	$D_t = 1.623N + 1515$	200–500 injections	12
Investment casting	1,2,3,4,5, 6,7,8,9	*X*	*s*, *t*	100	5	$D_t = 2.370N + 320$ $D_s = 1.000N + 320$	10–20 molds	5
Permanent mold	1,2,3,4,5, 6,7,8	*Y*	*s*, *t*	500	9	$D_t = 0.770N + 465$ $D_s = 0.335N + 465$	20–30 molds	10
Plaster mold	1,2,3,4,5, 6,7,8,9	*X*	*t*	100	5	$D_t = 2.160N + 320$	15–28 molds	5
Shell mold	1,2,3,4,5, 6,7,8,9	*X*	*s*, *t*	300	6	$D_t = 1.965N + 320$ $D_s = 0.595N + 320$	35–45 molds	5

*Numbers refer to the classifications in text.
†*S* = ferrous; *t* = nonferrous.

delivery requirements; and (6) economics of the process. Similar constraints apply in the selection of the most favorable secondary operations. *See* PRODUCTION ENGINEERING.

Choice of process. To identify the most favorable basic processes and secondary processes to be used to produce a given functional design, the process engineer may develop a selector guide to assist in the decision-making process.

Selector guide sheets are usually developed for a given size classification and are shown in part in **Tables 1–3**. In using these guide sheets, an analyst refers to the correct sheets for the size of the part under study. There may be several size classifications to handle all products being made by a given company. Next the analyst determines the geometry that best describes the component. Table 1 indicates nine geometric configurations needed to be considered in connection with the five casting processes listed. These nine classifications may be defined and illustrated to facilitate the classification of the geometry of the product being studied. For example:

Class 1 geometry—Solid or partly hollow rounds involving one or more diameters along one axis. The depth of hollow not to be more than two-thirds the diameter of the end containing the hollow.

Class 2 geometry—Hollow or partly hollow rounds involving one or more outside diameters and one or more inside diameters along one axis. The depth of hollow is more than two-thirds the diameter of the end containing the hollow.

Class 3 geometry—Solid or partly hollow shapes other than rounds such as square, triangular, octagonal or irregular, including one or more cross-sectional areas along one axis. The depth of hollow is not more than two-thirds the distance of the major diameter.

Class 4 geometry—Hollow or partly hollow shapes other than rounds involving one or more cross-sectional areas along one axis. The depth of hollow is more than two-thirds the major diameter of the hollow.

Class 5 geometry—Bowl-shaped concentric.

Class 6 geometry—Dish-shaped nonconcentric.

Class 7 geometry—Flats with or without configuration.

Class 8 geometry—Flanged.

Class 9 geometry—Complex characterized by unsymmetrical shapes.

The next parameter that is considered is the material used. Table 1 illustrates only two material considerations, ferrous (*s*) and nonferrous (*t*). The required

TABLE 2. Part of guide sheet for raw materials and manufacturing costs

Process	Raw material form	Raw material cost	Tooling cost	Equipment cost	Setup cost	Production cost	Scrap cost	Expected tolerance, mm	Expected surface finish, micrometers
Die casting	Pig or liquid	2	10	4	4	1	3	±.08 except .18 across parting line	.191
Investment casting	Pig	2	4	6	6	8	4	±.13	.318
Permanent mold	Pig	2	6	6	3	4	4	±.38	.381
Plaster mold	Pig	2	3	6	6	7	3	±.13	.254
Shell mold	Pig	2	4	6	6	4	3	±.13 except ±.22 across parting line	.318

TABLE 3. Part of guide sheet for quality considerations

Process	Expected reliability	Directional flow lines	Tool marks	Remarks
Die casting	A for Zn B for Al C for brass	No	Yes	Ejection pins may cause tool marks; surface finish depends on surface finish for die; casting of ferrous allows generally not economically feasible
Investment casting	B	No	No	Best for complicated design; no parting line; frozen mercury results in better finish than wax or plastic
Permanent mold	C	No	No	Lightest wall thickness 1/8 in. (3 mm); limited complexity because of rigid molds; not suitable for class 9 geometry
Plaster mold	A	No	No	Little finishing required
Shell mold	B	No	No	

number of parts needed is estimated. The constraint associated with quantity is the minimum number of parts required before it would become economical to provide tooling for the process under study. Decision equations may be developed for typical manufacturing environments of a company. The final step for the analyst is to solve the decision equations for processes still under consideration; most will already have been eliminated by this time because of the constraints of geometry, material, and quantity. The process that provides the smallest value based on the solution of the decision equation will be considered favorable for use. If Table 1 were given in full, it would be seen that two decision equations are provided for most of the processes. The D_s equation is used in conjunction with ferrous materials, and the D_t equation in conjunction with nonferrous materials.

To do a thorough job of process planning, use of the digital computer is recommended. Manual process selection may not give consideration to all possible alternatives. The computer can handle the mathematical manipulation of solving the various decision equations and the final selection problem. The equation used to develop the computer program based on the parameters shown in Table 1 is shown below,

$$D_t = A_{ijk} \times N + B_{ij}$$

where A is the coefficient of N, the number of units to be processed, i is the process to be considered, j is the size classification of the part being process planned, and k is the material specified. B_{ij} is the fixed element. *See* DECISION THEORY.

Sequence of secondary operations. In determining the sequence of operations to be performed, the process engineer considers the logical process order and the geometrical and dimensional control that can be maintained with each processing operation.

In the logical process order, basic operations are performed first and final inspection near the end of the process, just prior to a product's packaging for shipment. Holes that require reaming would necessarily be drilled before being reamed and, similarly, threaded inside diameters would necessarily be drilled prior to tapping.

In general, the final finishing of internal work is done in advance of external finishing. The principal reason for this sequence is that internal surfaces are less likely to be damaged in material handling and subsequent processes so that their surfaces can be completed earlier in the processing. For internal work, the logical sequence of operations is drilling, boring, recessing, reaming, and tapping. The logical sequence of external work is turning, grooving, forming, and threading.

Rough work involving heavy cuts and liberal tolerances should be performed early in a sequence. Heavy cuts will reveal defects in castings or forgings more readily than light cuts, and it is advantageous to identify defective raw material as soon as possible. Because both tolerance and finish of external surfaces can be adversely affected by subsequent material handling and clamping, close-tolerance operations should be performed late in the processing sequence.

In considering the geometrical and dimensional control that can be maintained, the process engineer should keep in mind those operations that establish locating or mastering points in order to schedule them early in the process. For example, it is usually easier to maintain control from a large, plane surface than from a curved, irregular, or small surface; so the process engineer establishes a surface which is best qualified for the location of critical features, and schedules it for finishing at a point early in the sequence.

Specifications concerning tools. It is the responsibility of the process engineer not only to specify how a piece should be made but also to order all special tools to be used in conjunction with the specified manufacturing processes. Special tools refer to those jigs, fixtures, dies, and gages that are required to hold work, guide the cutting tool, and inspect results. The requisition for the special tools should specify (1) those areas or points best suited for locating the workpiece while it is being processed; (2) that portion of the workpiece suited for supporting or holding it while it is being processed; and (3) that portion or area best suited for clamping so that the workpiece is securely held during the processing.

In specifying the location and number of locators to contact the work, the analyst must keep in mind that three locators are needed to locate a plane, two are needed to determine a line, and one will determine a point. A workpiece can move in either of

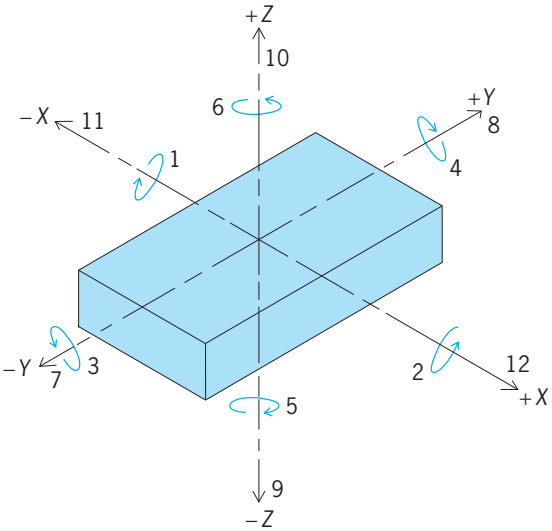

Fig. 1. Twelve degrees of freedom of a workpiece.

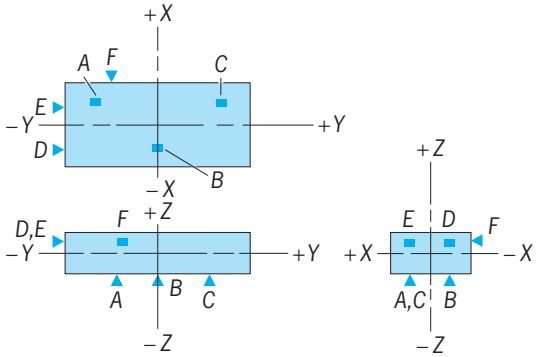

Fig. 2. Three views of a workpiece.

two opposed directions along three perpendicular axes (*X, Y,* and *Z*). Also, the work may rotate either clockwise or counterclockwise around each of these three axes. Each of these possible movements is a degree of freedom; hence 12 degrees of freedom exist. These 12 degrees of freedom are illustrated in **Fig. 1**. *See* DESCRIPTIVE GEOMETRY.

Work can be located positively by six points of contact in the tooling. These six points include three points on one plane. For example, in **Fig. 2** the three locations *A*, *B*, and *C* on the bottom of the block prevent the work from moving downward and from rotating about the *X* and *Y* axes. By adding two locating points *D* and *E* on a plane parallel to the plane containing the *X* and *Z* axes, the work is prevented from rotating about the *Z* axis and also from moving negatively along the *Y* axis. When the sixth and final locating point *F* is added on a plane parallel to the *Y* and *Z* axes, movement upward is prevented. Thus the first three locators prevented movements 1, 2, 3, 4, and 9, as shown in Fig. 1. The next locators prevent movements 5, 6, and 7, and the final locator prevents movement 11. This 3-2-1 locating procedure has prevented movement in 9 of the 12 possibilities. The three remaining degrees of freedom (8, 10, and 12) must not be restricted because they are needed in order to provide clearance to load and unload the tooling.

When requisitioning special tooling, the process engineer should indicate how the work is to be supported and specify the position of the locators. To assure correct position throughout the process, there must be rigid support so that the work does not deflect because of its static weight or the holding and tool forces applied to it, or both. Adding a support to avoid deflection during processing does not mean adding another locator. The support should not provide a point of location, and the workpiece should never contact the support until the tool or tooling forces are applied; the only purpose of the support is to avoid or limit deflection and distortion.

The holding force must be of sufficient magnitude to allow all locators to contact the workpiece during the processing cycle, but it is also important not to have holding forces so large that the work becomes marred. It is usually advantageous to place holding forces directly opposite locators. A nonrigid workpiece may require holding forces at several locations to hold the work against all locators. *See* PROCESS CONTROL.

Benjamin W. Niebel

Bibliography. W. D. Fellers and W. W. Hunt, *Manufacturing Processes for Technology*, 2d ed., 2000; M. P. Groover, *Fundamentals of Modern Manufacturing: Materials, Processes, and Systems*, 2d ed., 2001; B. W. Niebel and A. B. Draper, *Modern Manufacturing Process Engineering*, 1989; R. Robin and D. Wield (eds.), *Product Design and Technological Innovation*, 1986; J. A. Schey, *Introduction to Manufacturing Processes*, 3d ed., 1999; Society for the Advancement of Material and Process Engineering, *Processes: The Intercept Point*, 1988.

Prochirality

The property displayed by a prochiral molecule or a prochiral atom (prostereoisomerism). A molecule or atom is prochiral if it contains, or is bonded to, two constitutionally identical ligands (atoms or groups), replacement of one of which by a different ligand makes the molecule or atom chiral. Structures (**1**)–(**5**) are examples of prochiral molecules.

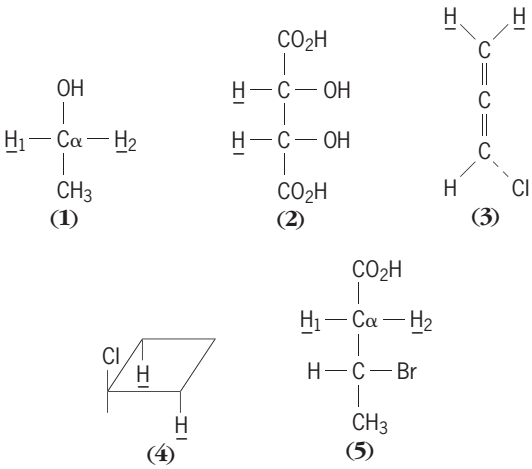

None of molecules (**1**)–(**4**) is chiral, but if one of the underlined pair of hydrogens is replaced, say, by deuterium, chirality results in all four cases. In compound (**1**), ethanol, a prochiral atom or center can be discerned (Cα:CH_2); upon replacement of H by D, a chiral atom or center is generated, whose configuration depends on which of the two pertinent atoms (H_1 or H_2) is replaced. Molecule (**5**) is chiral to begin with, but separate replacement of H_1 and H_2 (say, by bromine) creates a new chiral atom at Cα and thus gives rise to a pair of chiral diastereomers. No specific prochiral atom can be discerned in molecules (**2**)–(**4**), which are nevertheless prochiral [(**3**) has a prochiral axis].

Faces of double bonds may also be prochiral (and give rise to prochiral molecules), namely, when addition to one or other of the two faces of a double bond gives chiral products. Structures (**6**)–(**8**) have prochiral double-bond faces.

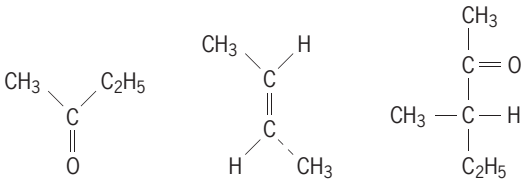

(**6**) (**7**) (**8**)

Hydrogenation of (**6**) gives chiral 2-butanol (addition to one face giving rise to one enantiomer, addition to the other to the opposite one), and HBr addition to one or the other face of (**7**) gives rise to enantiomeric chiral 2-bromobutanes, CH_3CHBr-CH_2CH_3. Hydrogenation from one face or the other of (**8**) also leads to a new chiral center, but in this case diastereomers rather than enantiomers result by addition to the prochiral carbonyl faces.

Prostereoisomerism. Although the term prochirality is widely used, especially by biochemists, a preferred term is prostereoisomerism. This is because replacement of one or other of the two corresponding ligands (called heterotopic ligands; see below) or addition to the two heterotopic faces often gives rise to achiral diastereomers without generation of chirality. Examples are structures (**9**)–(**12**). Replacement

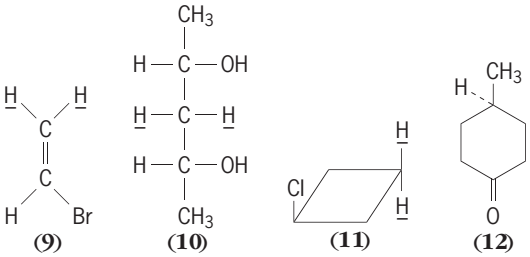

(**9**) (**10**) (**11**) (**12**)

of one or another of the underlined pairs of hydrogens in (**9**) by bromine gives rise to *cis*- or *trans*-1,2-dibromoethene; similar replacement in (**10**) by OH gives two different *meso*-2,3,4-pentanetriols; analogous replacement in (**11**) by Cl gives *cis*- or *trans*-1,3-dichlorobutane; catalytic addition of hydrogen to the carbonyl double bond in (**12**) gives *cis*- or *trans*-4-

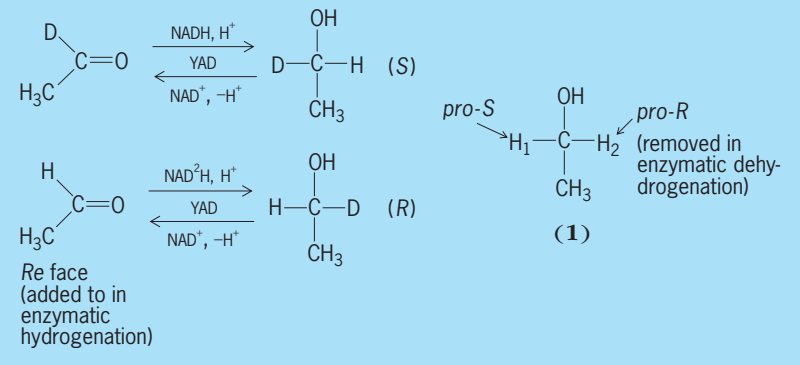

Reactions illustrating the ability of enzymes to distinguish between enantiotopic ligands or faces.

methylcyclohexanol. These are stereoisomeric pairs in all cases, but are devoid of chirality. Thus compounds (**1**)–(**12**) display prostereoisomerism, but only (**1**)–(**7**) display prochirality.

Heterotopic ligands and faces. The pertinent groups (H's) or faces (C═C, C═O) replaced or added to in (**1**)–(**12**) are called heterotopic (or, more precisely, stereoheterotopic) ligands or faces. If their replacement [(**1**)–(**4**)] or an addition to them [(**6**) and (**7**)] gives rise to enantiomers, they are called enantiotopic. If replacement or addition results in diastereomers [(**5**), (**8**)–(**12**)], the ligand faces are called diastereotopic.

The term (stereo)heterotopic comprises both enantiotopic and diastereotopic ligands and faces. Ligands are called homotopic when their separate replacement by different ligands gives rise to identical products, as in $C\underline{H}_3Cl$ or $C\underline{H}_2Cl_2$ (for example, replacing H by Br). Similarly, faces are called homotopic when addition to one or another of them gives rise to identical addition products, as in $(CH_3)_2C{=}O$ or $H_2C{=}CH_2$.

Nomenclature. To name heterotopic ligands, a conceptual replacement is made of one of them by a heavier isotope. If the configuration of the resulting chiral molecule is *R*, the ligand replaced is called *pro-R*; if *S*-configuration results, the ligand is called *pro-S*. Thus replacement of H_1 by D in structure (**1**) gives rise to (*S*)-ethanol-1-*d*; hence H_1 is *pro-S* and, by default, H_2 is *pro-R*, a fact which can be established independently in that replacement of H_2 by D yields (*R*)-ethanol-1-*d*.

To name heterotopic faces, the Cahn-Ingold-Prelog system (sequence rule) is applied to the three ligands in the plane; if the resulting sequence (highest to lowest precedence) is clockwise, the face is called *Re*; if it is counterclockwise, *Si*. Thus the front face of the methyl ethyl ketone (**6**) is *Si* ($O \rightarrow C_2H_5 \rightarrow CH_3$) and the front face of the *trans*-2-butene (**7**) is *Re* ($=C \rightarrow CH_3 \rightarrow H$); the rear faces of these two molecules would have the respective opposite designations [*Re* for compound (**6**), *Si* for compound (**7**)].

Significance. It is well known that enzymes can usually distinguish between enantiomers; thus liver alcohol dehydrogenase catalyzes the dehydrogenation of (*S*)-lactic acid, (*S*)-$CH_3CHOHCO_2H$, but not of

the *R* isomer, to pyruvic acid, CH_3COCO_2H. Similarly enzymes can distinguish between enantiotopic (as well as diastereotopic) ligands or faces, as illustrated by the classical work of F. Westheimer and B. Vennesland (see **illus.**). The fact that yeast alcohol dehydrogenase (YAD) mediates oxidation of (*S*)-ethanol-1-*d* by the coenzyme NAD^+ with total retention of deuterium (that is, clean loss of hydrogen), but of (*R*)-ethanol-1-*d* with integral loss of deuterium (despite the opposing isotope effect), is best explained by saying that the dehydrogenation of ethanol (labeled or not) by YAD-NAD^+ always leads to removal of the *pro-R* hydrogen [H_2 in compound (**1**)]. Correspondingly, the reverse reaction, reduction of acetaldehyde or acetaldehyde-1-*d* by NAD^2H and NADH, respectively, in the presence of YAD, invariably leads to addition of deuterium or hydrogen to the *Re* face. This type of stereochemical preference is due to the specificity of the fit of the substrate with the active site of the enzyme.

Diastereotopic, but not enantiotopic, groups are in principle distinct (anisochronous) in nuclear magnetic resonance (NMR) spectra. Thus ethanol (**1**) displays a single (though spin-split) methylene peak, but the underlined protons in (**5**) and (**9**)–(**11**) are distinct in their signals. This distinction must be kept in mind in the interpretation of NMR spectra.

In citric acid [structure (**13**)] all four marked H's

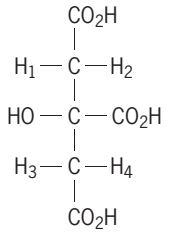

(**13**)

are heterotopic and therefore distinct toward enzymes; only one specific one is implicated in the aconitase-mediated dehydration of citric acid to *cis*-aconitic acid, $HO_2C{-}CH{=}C(CO_2H)CH_2\ CO_2H$. The CH_2CO_2H branch toward which the elimination occurs also can be specifically identified. Moreover, in the biosynthesis of citric acid from oxaloacetate and acetyl-CoA (aldol condensation), the oxaloacetate-derived and acetate-derived branch can be distinguished: the former is *pro-R*, the latter *pro-S*. Since aconitase leads to dehydration toward the *pro-R* branch, it is the *pro-S* branch, that is, the acetyl-CoA derived one, which remains intact. This was demonstrated (by labeling) early in the investigation of the citric acid cycle, and appeared mysterious at the time; the concept of prochirality provides a logical explanation. *See* ENZYME; MOLECULAR ISOMERISM; NUCLEAR MAGNETIC RESONANCE (NMR); STEREOCHEMISTRY.

Ernest L. Eliel

Bibliography. D. Arigoni and E. L. Eliel, *Top. Stereochem.*, 4:127, 1969; E. L. Eliel, Stereochemical nonequivalence of ligands and faces (heterotopicity), *J. Chem. Educ.*, 57:52–55, 1980; K. R. Hanson, Applications of the sequence rule, I: Naming the paired ligands *g,g* at a tetrahedral atom X_{ggij}, II: Naming the two faces of a trigonal atom Y_{gbi}, *J. Amer. Chem. Soc.*, 88:2731, 1966; W. B. Jennings, Chemical shift nonequivalence in prochiral groups, *Chem. Rev.*, 75:307, 1975; K. Mislow and M. Raban, *Top. Stereochem.*, 1:1, 1967.

Prochlorophyceae

A class of prokaryotic organisms coextensive with the division Prochlorophycota in the kingdom Monera. Because prochlorophytes carry out oxygen-evolving photosynthesis, they may be considered algae. They are distinguished from Cyanophyceae, the only other prokaryotic algae, by the presence in their photosynthetic lamellae of chlorophyll *b* in addition to chlorophyll *a* and the absence of phycobilin pigments. Otherwise, they resemble Cyanophyceae biochemically and ultrastructurally. The class comprises a single genus, *Prochloron*, with one species.

Prochloron didemni is a green, coccoid unicell 10–30 micrometers in diameter that lives only as an extracellular symbiont of colonial tunicates (didemnid ascidians) in the tropical Indo-Pacific. It has been intensely studied because advocates of the theory of origin of organelles by serial endosymbiosis believe that the precursor of green chloroplasts may have been an organism like *Prochloron*. Because *Prochloron* has not been grown in bacteria-free culture successfully, certain questions pertinent to its phylogeny have not been answered. *See* ALGAE.

Paul C. Silva; Richard L. Moe

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae: Structure and Reproduction*, 2d ed., 1997; R. A. Lewin, *Prochloron:* A status report, *Phycologia*, 23:203–208, 1984.

Procyon

The brightest star in the constellation Canis Minor, apparent magnitude +0.36. Procyon (α Canis Minoris) is among the stars nearest to the Earth, at a distance of only 3.5 parsecs (1.08×10^{14} km or 6.7×10^{13} mi). Its spectral type is F5, but Procyon is slightly overluminous compared to a main-sequence star of the same spectral type, which indicates that Procyon has already begun to evolve off the main sequence. Its intrinsic luminosity is about seven times that of the Sun, and the effective temperature is estimated to be 6500 K (11,200°F). The radius of the star is about 2.2 solar radii. *See* SPECTRAL TYPE; STELLAR EVOLUTION.

Irregularities in the proper motion of Procyon noticed in 1840 led to the prediction that a relatively massive unseen companion was altering the path of the star, causing it to move from side to side with a period of about 40 years. The faint 11th-magnitude companion was discovered in 1896 about 5″ away, but is very difficult to see in the glare of the bright

primary. The mean physical separation is about 16 astronomical units. Procyon B was subsequently shown to be a white dwarf in the final stages of its evolution, with a luminosity only 1/2000 that of the Sun. From the astrometric orbit the masses of the primary and its companion have been computed as 1.75 and 0.62 solar masses, respectively. The progenitor of the white dwarf was originally the more massive of the two stars and underwent the final stages of its stellar evolution sooner than the original secondary, which is now seen as Procyon A. The age of the system is approximately 1 billion years. Procyon B is only about 1.5 times the size of the Earth, which implies that its mean density is enormous, probably exceeding 300,000 times the density of water, or 6 tons/in.^{3} (300 kg/cm^3). It is the second nearest white dwarf, the nearest being the companion of Sirius. *See* BINARY STAR; STAR; WHITE DWARF STAR.

David W. Latham

Product design

The determination and specification of the parts of a product and their interrelationship so that they become a unified whole. The design must satisfy a broad array of requirements in a condition of balanced effectiveness. A product is designed to perform a particular function or set of functions effectively and reliably, to be economically manufacturable, to be profitably salable, to suit the purposes and the attitudes of the consumer, and to be durable, safe, and economical to operate. For instance, the design must take into consideration the particular manufacturing facilities, available materials, know-how, and economic resources of the manufacturer. The product may need to be packaged; usually it will also need to be shipped so that it should be light in weight and sturdy of construction. The product should appear significant, effective, compatible with the culture, and appear to be worth more than the price. The emphasis may differ with the instance. Durability in a paper napkin is different from durability in a power shovel.

To determine whether a design is well adjusted to the gross array, criteria are needed. Some are objective and measurable, such as clearances and efficiency, whereas others are quite subtle and even subjective. In a way, product design is an industrial art.

Ultimately the purpose of product design is to ascertain that the product will satisfy human wants and wishes either directly as consumer goods, or indirectly as capital equipment or components.

Except in the case of basic inventions, product design is a redesign to suit changed conditions, criteria, or enlightenment. Change may appear capricious as in some fashions and toys, may be the result of technological progress, or may result from a change in attitude toward the product or its function. In some areas trends can be discovered, but future preferences of buyers must be predicted, gambled on, or forced to occur.

There are various steps in product design which are not necessarily in particular order. They are analytical studies, creative synthesis, drawings and models for appearance, function and specifications, plus calculations, experiments, and tests. *See* PROCESS ENGINEERING; PRODUCTION ENGINEERING; PRODUCTION PLANNING.

Richard I. Felver

Product quality

Quality is the collection of features and characteristics of a product that contribute to its ability to meet given requirements. Though quality concepts date back to ancient times, the early work in modern times was on methods and processes to create standards for producing acceptable products. By the mid-1950s, mature methods had evolved for controlling quality, including statistical quality control and statistical process control, utilizing sequential sampling techniques for tracking the mean and variance in process performance. During the 1960s, these methods and techniques were extended to the service industry. During 1960–1980, there was a major shift in world markets, with the position of the United States declining while Japan and Europe experienced substantial growth in international markets. Consumers became more conscious of the cost and quality of products and services. Firms began to focus on total production systems for achieving quality at minimum cost. This trend has continued, and today the goals of quality control are largely driven by consumer concerns and preferences.

Characteristics. There are three views for describing the overall quality of a product. First is the view of the manufacturer, who is primarily concerned with the design, engineering, and manufacturing processes involved in fabricating the product. Quality is measured by the degree of conformance to predetermined specifications and standards, and deviations from these standards can lead to poor quality and low reliability. Efforts for quality improvement are aimed at eliminating defects (components and subsystems that are out of conformance), the need for scrap and rework, and hence overall reductions in production costs. Second is the view of the consumer or user, who can have different needs. To consumers, a high-quality product is one that well satisfies their preferences and expectations. This consideration can include a number of characteristics, some of which contribute little or nothing to the functionality of the product but are significant and influential in providing customer satisfaction. A third view relating to quality is to consider the product itself as a system and to incorporate those characteristics that pertain directly to the operation and functionality of the product. This approach should include overlap of the manufacturer and customer views.

An example of how the product quality of an automobile is examined is given in **Table 1**. From the manufacturer's viewpoint, this product could be rated as high quality, since units leave the final

TABLE 1. Vehicle product quality data

Characteristic	Value
Warranty	3 years/36,000 mi (58,000 km)
Outgoing inspection	98%
Engine and power-train mean time to failure (MTTF)*	42,000 mi (67,000 km)
Acceleration 0–60 mi/h (0–100 km/h)	10 s
Braking distance 55–0 mi/h (88–0 km)	480 ft (140 m)
Fuel consumption	16 mi/gal (6.8 km/liter) in town, 20 mi/gal (8.5 km/liter) on highway

*Usage (total distance traveled) is a common measure for vehicle failures, and is conventionally designated MTTF.

assembly processes with only a 2% observed rate of defects. Assuming that few failures occur during the early stages of use and that the major failures are likely to occur after the warranty has expired, the vehicle seems to be manufactured within the engineering and production specifications for which it was designed. Moreover, the warranty expenses should be relatively low. The customer, however, might view this product differently, particularly in the absence of any highly attractive features through options that might offset his or her satisfaction level. While the early failures might be few in number, the failures that occur after warranty can be quite expensive for the customer. In this example, from the viewpoint of a product system the product has more of a moderate level of quality based on the acceleration, braking distance, and fuel economy. *See* MANUFACTURING ENGINEERING.

Elements. A number of dimensions or factors have evolved that contribute to overall product quality as it is presently defined. Some dimensions are more dominant than others in describing a particular product. The following elements are most common in describing quality.

Operating performance comprises measures of actual operating performance of the product. Examples include automobile gasoline consumption in miles per gallon (kilometers per liter), brake horsepower for an engine, range of a weapon, and hours of life for a battery.

Durability refers to the amount of use before product deterioration, normally measured in the time or hours of usage before a product fails.

Reliability refers to the probability of a product failing within a specified period of time, having survived to that point. Primarily applied to durable goods, reliability is assessed by the average rate of failure, mean time between failures (MTBF), and related failure time measures.

Conformance is the degree to which design and operating characteristics meet preestablished standards, as measured through defect count, scrap rate, and amount of rework.

Serviceability refers to the degree or ease of restoring service, measured in time to restore service, frequency of calls, and mean time to repair (MTTR).

Usability refers to the ease of use that a consumer experiences in developing the necessary preparation, skills, and proficiency for using a product as it was intended. The most common measure of usability is learning time.

Perceived quality concerns the overall image of a product among users and potential users. Though some objective measures are involved through surveys and opinions, perceived quality is generally quite subjective.

Esthetics concerns the way in which the product is sensed through looks, feel, sound, taste, and smell. *See* RELIABILITY, AVAILABILITY, AND MAINTAINABILITY.

Many of the quality elements involve measurable product attributes that are easily quantified, but some are subjective and are influenced by selective group and individual preferences. This is particularly true of perceived quality and esthetics, which characterize the customer view of products.

Quality control. Quality control (QC) is the collection of methods and techniques for ensuring that a product or service is produced and delivered according to given requirements. This includes the development of specifications and standards, performance measures, and tracking procedures, and corrective actions to maintain control. The data collection and analysis functions for quality control involve statistical sampling, estimation of parameters, and construction of various control charts for monitoring the processes in making products. This area of quality control is formally known as statistical process control (SPC) and, along with acceptance sampling, represents the traditional perception of quality management. Statistical process control focuses primarily on the conformance element of quality, and to somewhat less extent on operating performance and durability. *See* PROCESS CONTROL; QUALITY CONTROL.

Performance measures. It is convenient to think of product quality as a vector of attributes that span all important characteristics of the item relative to the overall performance according to the three bases—the manufacturer, the customer, and the overall product viewed as a system. Some of these elements are qualitative, but some are entirely quantitative. The challenge is to incorporate these elements into an overall measure of product performance suitable for monitoring progress in quality improvement and for making management decisions about the product quality and costs.

The overall performance of a product depends on each attribute separately. For example, if a product is considered with regard to two attributes (for example, quality and reliability), and if each of these attributes has three possible values (high, medium,

TABLE 2. Thresholds of overall quality and reliability

Dimension	Threshold	Category
Quality (defects/1000)	<5	High
	5–20	Medium
	>20	Low
Reliability (mean time between failures, hours)	>200	High
	100–200	Medium
	<100	Low

and low), then there are $3 \times 3 = 9$ possible pairs of values of the two attributes, and each pair can give rise to a different value of overall performance. Moreover, the overall performance can vary with time as the values of the attributes vary, and it is often useful to graph this variation.

For example, a new treadmill is brought to market, and the manufacturer has determined overall quality and reliability thresholds as shown in **Table 2**. The average time for customers to learn how to use all of the features on the unit ranges 2–5 h.

Here the most qualitative (subjective) element of interest is usability, because of the relative uncertainty in the training time. Therefore, the first step is to construct the set of combinations of low (L), medium (M), and high (H) attribute values for quality and reliability, giving (L,L), (L,M), (L,H), (M,L), (M,M), (M,H), (H,L), (H,M), (H,H). To incorporate the usability dimension, the learning time is categorized into intervals, such as "easy" for times up to 2.5 h and "normal" for times between 2.5 and 5 h. This forms an expanded range of attribute values. By assigning variables to each of these combinations and employing a process of pairwise comparisons among these attributes, relative weights can be assigned to establish an overall measure of product performance, $\Pi(t)$ [see **illus.**].

Quality, reliability, and warranties. Two types of product failures can generally be distinguished: those that occur relatively soon after they are produced and sold to consumers, and those that occur later after the product has been in use, perhaps even without fault for many cycles or hours of usage. From a manufacturer's point of view, the failures that occur early in the product life are generally due to low quality and are attributed to a lack of conformance to manufacturing and production standards. Here quality of the product can be thought of as a state of acceptance that relates directly to the manufacturing processes, materials, and workmanship. Accordingly, quality is maintained and controlled through developed standards and specifications for characteristics such as dimensions and tolerances, composition and structure of materials, hardness and surface finish, and strength of linkages and mechanism.

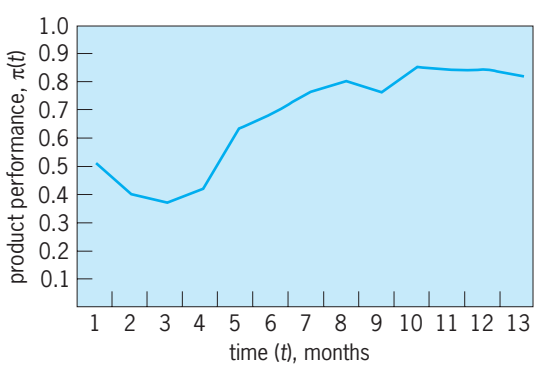

Sample of monthly product performance, $\Pi(t)$. The variable $\Pi(t)$ is an index of the level of performance, with 1 corresponding to perfect and 0 corresponding to worst level.

Those failures that occur later in the life or usage cycle of a product are commonly due to low reliability, caused by problems in the design and engineering of the item. These failures typically occur as stress fractures, fatigue, and shearing, which relate to problems in the materials, basic design, and production processes. Sound planning early in the design process is necessary to overcome these and other problems in order to achieve high product reliability.

Warranties require the producer to pay all or some of the expenses involved in the repair or replacement of products that fail within a prescribed period of time. Whether planned by the manufacturer as part of a market strategy or stipulated by liability legislation, the costs for warranties can represent a significant portion of the cost to manufacture a product. Low product quality caused by poor conformance, low reliability, or both will lead to high warranty costs.

Operational measures of product performance. Warranty costs are expenses that occur because of low product quality. Therefore, measures that relate to the warranty expenditures also relate to product quality. Common operational measures that are used to track quality-related performance in production include the percentage of failures or defects, the number of warranty claims per unit sales, and the average warranty costs per unit sold. These measures, which involve warranty claims, also reflect customer attitude toward the product.

Concurrent engineering, quality function deployment, and total quality management (TQM) are modern management approaches for improving quality through effective planning and integration of design, manufacturing, and materials management functions throughout an organization. Quality improvement programs typically include goals for reducing warranty claims and associated costs because warranty data directly or indirectly impact most of the product quality dimensions. *See* ENGINEERING DESIGN.

Marlin U. Thomas

Bibliography. D. A. Garvin, Competing on the eight dimensions of quality, *Harvard Bus. Rev.*, 65(6):107–109, 1987; D. L. Goetsch and S. B. Davis, *Introduction to Total Quality*, 2d ed., Prentice Hall, 1997; C. L. Karnes, S. V. Sridharan, and J. J. Kanet, Measuring quality from the consumer's perspective: A methodology and its application, *Int. J. Prod. Econ.*, 39:215–225, 1995; N. A. Morgan and N. F. Piercy, Market-led quality, *Ind. Market Manag.*, 21:111–118, 1992.

Product usability

A concept in product design, sometimes referred to as ease of use or user-friendliness, that is related directly to the quality of the product and indirectly to the productivity of the work force. Customer surveys show that product quality is broken down into six components (in descending order of importance): reliability, durability, ease of maintenance, usability, trusted or brand name, and price. Ease of maintenance and usability both relate to product usability. Reliability also has a component of usability to it. If a product is too difficult to use and thus appears not to work properly, the customer may think that it has malfunctioned. Consequently, the customer may return the product to the store not because it is unreliable but because it does not work the way the customer thinks it should.

Since quality is often the main determinant in a customer's decision whether or not to buy the product, products that are highly usable are more likely to be purchased. Usability is especially important for computer products, both hardware and software, because poor products can be almost impossible to operate. As e-commerce grows on the Internet, the difference between a successful Web site and an unsuccessful one is often the usability of the site. If customers cannot easily make a purchase, an alternative site or method will be chosen. *See* HUMAN-COMPUTER INTERACTION; INDUSTRIAL ENGINEERING; PRODUCTIVITY.

Usability criteria. There are criteria by which a product's usability can be measured, including:

1. *Time to perform a task, or the execution time*. The ability to execute tasks quickly is a sign of good usability. Execution time is dependent on the number of steps that must be performed by the user. Intelligent layout of the keys, buttons, and other controls can also have a positive effect on execution time and usability.

2. *Learnability*. Usable products operate simply, and can be learned quickly. The steps are straightforward, intuitive, and easy to follow.

3. *Mental workload, or the mental effort required to perform a task*. A usable product will contain clues to using the product in the physical design instead of relying on people's memory of how to perform a task. Humans can retain about three to nine items in short-term memory at one time. If, due to a poor product design, the user is required to remember too much, then the usability of the product will suffer.

4. *Consistency in the design*. If several tasks must be performed while using the product, the operations should be consistent. Tasks are consistent if they involve similar steps or methods. In addition, product lines from the same company should operate consistently so that the user's learning experience is transferred from one product to the next. Learning time can be decreased because familiarity with one product can be applied to using a new product.

5. *Errors*. Errors can be categorized according to criteria 1–4 above. For instance, errors will occur if a task requires so many steps that the user cannot remember the steps or their proper order; if a task is too difficult to learn; if the effort required to perform the task is too high; or if the steps in the task are so inconsistent with other steps they are difficult to remember.

The usability of a product usually cannot be optimized for all five criteria at the same time. Trade-offs will occur. As an example, a product that is highly usable in terms of fast execution times will often have poor usability in terms of the time needed to learn how to use the product. A product designer must be aware that it may not be possible for a product to be highly usable by all usability criteria, and so design according to the criteria that are most important to potential customers. Casual users of a product will have different demands on a product compared to expert users. *See* CONTROL SYSTEMS; HUMAN-FACTORS ENGINEERING; HUMAN-MACHINE SYSTEMS.

Laboratory methods for improving usability. Many companies, especially computer or consumer electronics companies, have laboratories in which to test the usability of their products. The methods of usability testing are formal experimentation, informal experimentation, and task analyses.

Formal experiments. Formal experiments take place in a laboratory in which different conditions can be controlled and manipulated. Usability is measured most often by counting the number of errors or recording the time needed to execute a task. Often, performance on different products or product features will be compared to determine which can be executed the fastest with the fewest errors. In a typical product design cycle, formal experimentation is often used to test users on early mock-ups, later on the product prototypes, and then on the final product.

The advantage of formal experimentation is that clear cause-and-effect relationships can be established. As an example, the experiment could show that a certain product design causes better usability performance either in faster use speed or reduced errors.

Formal experiments have some disadvantages when compared to other methods for testing usability. First, formal experimentation can be a costly procedure. Typically, each participant in the experiment is tested individually. In addition, setting up the experiment to control for possible confounding factors can be time-consuming. Another disadvantage is that a controlled setting may be good for establishing a cause-and-effect relationship, but it can also produce a sterile experimental environment that may not correspond to how the product will be used in practice. Finally, the data collected may not be very rich. The data will indicate that one product may be better than another, but it cannot necessarily be used to generate new ideas. In most cases, therefore, formal experimentation is used in the latter stages of the product development cycle. Informal experimentation in the early stages can help identify potential product features that may improve usability.

Informal experiments. Surveys, questionnaires, observations in natural settings, and focus groups are ways

to provide product designers with ideas on how a product can be improved. The surveys and questionnaires can be used throughout the design process to assess users' perceptions of the product, including what users like and dislike about the product. At the conception of the product, questionnaires can assess the need for a new product or the need to redesign an existing product. Once the product is placed on the market, product surveys can assess how the users have received the product. Questionnaires can also accompany formal experiments to measure users' satisfaction with the product.

Users can be observed in various settings performing various tasks with the product. Typically, products are tested inside the company designing the product. They can also be brought to consumer product shows or placed in stores. The people using the products in these settings are usually videotaped. The videotapes are often analyzed using a technique called content analysis. For content analysis, definable events, such as obvious displays of frustration, are counted when viewing the videotapes. If too many customers show high levels of frustration, then the problem tasks can be identified and corrected. The videotapes can be especially useful for identifying errors from which the users cannot recover. These kinds of errors are likely to cause product returns.

Focus groups, groups of usually four to eight people, are brought together to use a product and discuss its advantages and disadvantages. A focus group is designed to elicit comments from the participants that normally would not be elicited individually. A focus group moderator can direct the questions and discussions.

Informal experiments are particularly important at the beginning and end of the design process. Although clear cause-and-effect relationships cannot be established, the information received from these techniques can be very valuable in directing the product designers to potential usability problems.

Task analysis. Two kinds of task analysis are typically used: protocol analysis and cognitive analysis. For a protocol analysis, a single user will be brought to the lab to perform various tasks with the product. Generally, the user will be asked to verbalize his or her thought processes (provide verbal protocols) during execution of a task. The verbal protocols then can be analyzed to determine if the participant had the correct goal for accomplishing a task, if the goal was executed correctly, and what kind of confusion occurred. Verbal protocols provide an experimenter with a rich set of data that can be used to improve the product's usability.

In a cognitive model, the cognitive processes of a potential user can be modeled and the model can be executed, similar to a computer program. Some models are derived from the verbal protocols of subjects. Cognitive models can be used to make quantitative predictions of the usability of products. These predictions, which can be used in lieu of the more costly method of formal experimentation, have been shown to be very accurate. *See* COGNITION; MODEL THEORY; SIMULATION; SYSTEMS ANALYSIS; WORK MEASUREMENT.

Summary. Many of the best and most successful companies place great emphasis on usability of their products. Products that are not usable will lose market share, and customers will return these products to the store if they think that they are not working properly. In many cases, they are not working properly because the customer cannot understand how to use them. Usable products also have an effect on worker productivity. Excessive learning times, execution time, or errors can slow the accomplishment of tasks and reduce productivity. Products that are poor in usability force the worker to concentrate on how to use the product instead of concentrating on how to perform the task. Although laboratory methods for improving usability can increase the cost of the product design, the benefits (market share, productivity) will outweigh the costs. *See* METHODS ENGINEERING; OPTIMIZATION.

Ray Eberts

Bibliography. R. W. Bailey, *Human Performance Engineering*, Prentice Hall, Upper Saddle River, NJ, 1996; R. Eberts, *User Interface Design*, Prentice Hall, Englewood Cliffs, NJ, 1994; D. Norman, *The Psychology of Everyday Things*, Basic Books, New York, 1988; M. S. Sanders and E. J. McCormick, *Human Factors in Engineering and Design*, 7th ed., McGraw-Hill, New York 3d ed., 1993.

Production engineering

A branch of engineering that involves the design, control, and continuous improvement of integrated systems in order to provide customers with high-quality goods and services in a timely, cost-effective manner. It is an interdisciplinary area requiring the collaboration of individuals trained in industrial engineering, manufacturing engineering, product design, marketing, finance, and corporate planning. In many organizations, production engineering activities are carried out by teams of individuals with different skills rather than by a formal production engineering department. Although production engineering originated in the manufacturing sector, recently it has found increasing application in the service industries, such as banking, airlines, and entertainment. Thus "product" denotes both a traditional manufactured product and a service product, such as an insurance policy or a bank loan. While traditional production engineering activities have taken place within a single factory, it is becoming more common for the production engineering team to include not only in-house personnel from the different skill sets mentioned above, but also personnel from several geographic locations and even other companies. It is quite common for a firm's design activity to take place in one country, production of components in another, and assembly of the final product in a third. Important components and subsystems of the finished product may be subcontracted to other firms, requiring production engineering teams to include personnel from different companies. Recent developments in information

technology, such as the Internet and computer-supported cooperative work, are important enablers of this process.

Product design. In today's global markets, where companies from all over the world can compete to sell in a given area, rapid response to changing market conditions is essential to a company's survival. The company must rapidly introduce new products to exploit market opportunities, and continuously improve its operations to deliver higher-quality products faster and at lower cost than its competitors. The introduction of a new product is critical for production engineering. Initially, the production engineering team works with the designers, helping them to develop a product that can be manufactured economically while preserving its functionality. Features of the product that will significantly increase its cost are identified, and alternative, cheaper means of obtaining the desired functionality are investigated and suggested to the designers. Decisions as to whether to manufacture a given component of the product internally or to purchase it from outside suppliers must be made, based on cost and quality considerations and the reliability of the production process or supplier. Similar considerations in the service industries involve decisions as to whether or not to subcontract certain services. The process of concurrently developing the product design and the production process is referred to by several names such as design for manufacturability, design for assembly, and concurrent engineering. *See* ACTIVITY-BASED COSTING; DESIGN STANDARDS; PROCESS ENGINEERING; PRODUCT DESIGN; PRODUCTION PLANNING.

Production process. The specification of the production process should proceed concurrently with the development of the product design. This involves selecting the manufacturing processes and technology required to achieve the most economical and effective production. The technologies chosen will depend on many factors, such as the required production volume, the skills of the available work force, market trends, and economic considerations. In manufacturing industries, this requires activities such as the design of tools, dies, and fixtures; the specification of speeds and feeds for machine tools; and the specification of process recipes for chemical processes. Deciding whether individual operations should be performed manually or using automated machinery is another recurring decision. In a service industry, the decisions might be for the specification of the steps required for a loan approval procedure or the processing of an insurance claim, along with the type of computer support to be provided at each step. The production engineering team must interact with other groups, such as research and development, to determine the viability of developing new technologies; machine and tooling design groups, to provide input on the needs of the process; information systems, to develop the necessary computer support; and finance and marketing groups, to evaluate the impact on costs and customer satisfaction of the decisions made. There is also interaction with human resource and industrial engineering groups to design workplaces in which workers can perform efficiently and safely, and to provide workers the necessary training. There should be a constant dialogue between the production engineering team and the product designers, enabling the latter to take advantage of the latest process knowledge to enhance the product design, and the former to adapt to changes in required functionality. *See* AUTOMATION; HUMAN-FACTORS ENGINEERING; MANUFACTURING ENGINEERING.

A critical factor in selecting the appropriate production process is the nature of the market for which the product is designed. If the product will serve a broad consumer market, such as consumer electronics or automobiles, the firm must produce a relatively standard product in very high volumes for a reasonably long time. In this situation, referred to as mass production, the firm can afford to develop highly specialized production equipment that is capable of performing specific operations for a particular product. Manual tasks can be divided into simple, repetitive elements which do not require high worker skill levels. Production engineering is active in the design of the production equipment and its integration with workers into an effective manufacturing process.

Pilot production. Actual production of physical products usually begins with a few prototype units being manufactured in research and development or design laboratories for evaluation by designers, the production engineering team, and sales and marketing personnel. Once a basic product design and production process have been agreed upon by all parties, relatively small volumes of the product will be produced on a pilot production line, which is composed of the same basic equipment and processes that will be used in the actual production facility. The goal of this pilot production phase is to give the production engineering team hands-on experience producing the product, allowing problems to be identified and remedied before investing in additional production equipment or shipping defective products to the customer. The pilot production process involves changes to the product design and fine-tuning of unit manufacturing processes, work methods, production equipment, and materials to achieve an optimal trade-off between cost, functionality, and product quality and reliability. Once a satisfactory degree of efficiency and predictability has been achieved, the production process is installed in a full-scale production facility capable of producing the product in sufficient volume to meet the market demand. *See* PILOT PRODUCTION; PROTOTYPE.

Production facility. The production facility itself can be designed around the sequence of operations required by the product, referred to as a product layout (**Fig. 1**). In another scenario, the firm develops custom products to meet the specialized needs of individual customers. This requires the firm to produce a wide range of products in small volumes, which is referred to as job shop production. General-purpose

production machinery is used, and often must be set up for each individual job, incurring significant changeover times while this takes place. This type of production facility is usually organized in a process layout (**Fig. 2**), where equipment with similar functions is grouped together. The highly job-dependent nature of the production operations also requires a skilled work force capable of evaluating different alternatives and making good choices. *See* HUMAN-MACHINE SYSTEMS; PRODUCTION METHODS.

Many firms fall between these two extremes, where a limited range of parts are produced in medium volumes. This type of environment is known as batch manufacturing, since products are made in batches to amortize changeovers of machines over an economic number of units of the product. Cellular or group technology layouts, where groups of machines are dedicated to the production of families of similar parts, are often used in these environments. An important production technology for the latter environment is flexible manufacturing systems, which are generally composed of a number of computer-controlled machines linked by a centralized computer controller and automated material-handling equipment. The goal of these systems is to be able to produce a variety of products with minimal changeover time, which is achieved by reprogramming the computer control as opposed to adjusting or reconfiguring the actual hardware. *See* FLEXIBLE MANUFACTURING SYSTEM.

The design of the production facility, an important part of production engineering, also overlaps significantly with industrial engineering. The number of each type of machine needed and how the machines are organized in the plant space available determines not only the production capacity of the facility but also to a great extent the degree of difficulty of the activities involved in coordinating the flow of material through the plant to manufacture the product. The day-to-day coordination of production activities (production planning and control), such as how to prioritize the different tasks requiring machine time or attention by personnel, are the domain of industrial engineering and operations management. However, the nature of these tasks and their difficulty is affected heavily by the plant layout and the choice of process technology. The effective performance of these tasks is a major part of any successful production engineering effort. *See* INDUSTRIAL ENGINEERING; INDUSTRIAL FACILITIES.

Process improvement. The production engineering process does not stop once the product has been put into production. A major function of production engineering is continuous improvement—continually striving to eliminate inefficiencies in the system and to incorporate and advance the frontier of the best existing practice. Production processes tend to evolve over time as more cost-effective methods become available. Market and product characteristics also change, requiring modification of the production processes. The task of production engineering is to identify potential areas for improving the performance of the production system as a whole,

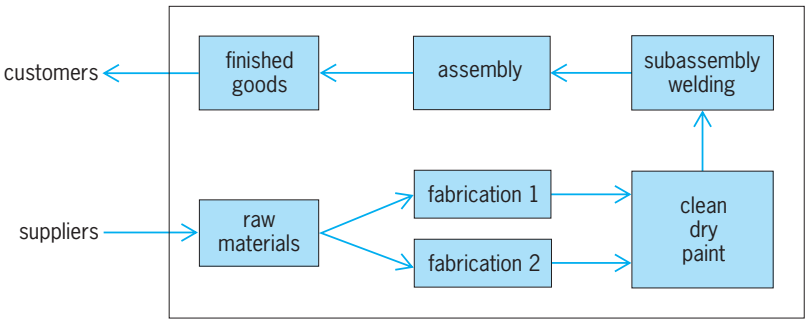

Fig. 1. Example of a process for a mass-production environment.

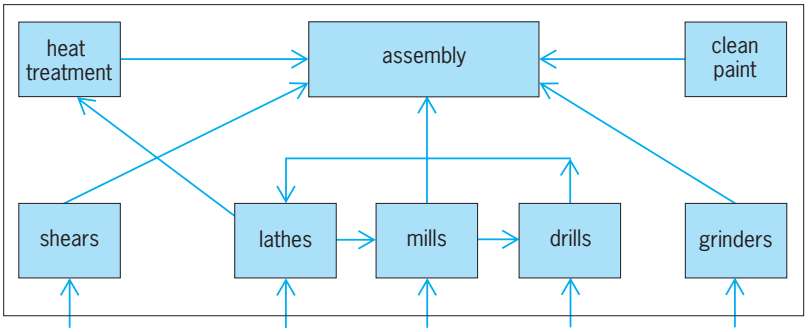

Fig. 2. Process layout for a job shop environment.

and to develop the necessary solutions in these areas. This often requires considerable involvement of the work force, with the production engineering team serving as facilitators and technical consultants. The solutions developed may range from modifications of production equipment and work methods to changes in product design or the revision of a quality specification. A successful continuous improvement program serves to eliminate waste in the company's operations, providing it cost and quality advantages over its competitors. *See* PRODUCT QUALITY.

Reha M. Uzsoy

Bibliography. R. G. Askin and C. R. Standridge, *Modeling and Analysis of Manufacturing Systems*, John Wiley, New York, 1993; J. Browne, J. Harhen, and J. Shivnan, *Production Management Systems*, Addison-Wesley, 2d ed., 1996; J. Heizer and B. Render, *Principles of Operations Management*, Allyn and Bacon, 6th ed., 2005; J. A. Tompkins and J. A. White, *Facilities Planning*, John Wiley, 3d ed., 2002; J. A. White (ed.), *Production Handbook*, John Wiley, 1987.

Production lead time

The total time required to convert raw materials into a finished product. More than the simple sum of machine operation and assembly times, it also includes time spent waiting for machines to be available, moving between machines, and performing quality inspections and other nonoperation functions. Short production lead times are ideal, as long times reduce the responsiveness of a company and result in high inventories.

While customer lead time is the interval of time between placing the order and receiving the order, production lead time typically focuses on the time required to produce the order once all the necessary materials are present. It does not take into consideration the time required to obtain the necessary materials or the time the materials wait in the plant before production begins.

Production lead time is very important in make-to-order industries (such as the aerospace industry). There, lead time can be a deciding factor in obtaining orders. In addition to low cost and superior quality, companies win orders by promising to deliver them quickly. Lead time is not as much of an issue in make-to-stock industries (such as small appliances and other retail products) in that these industries are building inventories against future demand, thus providing a buffer against longer than expected production lead times.

Components. The primary components of production lead time are setup, processing, handling, and nonoperation. Setup time (t_s) is required to prepare the machine for the particular order of parts. Processing time (t_p) includes the time spent loading and unloading the part from the machine as well as actual machining time. Handling time (t_h) is the time an order spends being transported between machines. Nonoperation time (t_n) includes wait time (for both machines and transportation) and time spent on certain other activities such as inspection. Assuming that an order consists of Q identical parts and that there are N operations in the processing sequence (with i indicating the operation in the processing sequence), production lead time (LT) is expressed by Eq. (1).

$$\mathrm{LT} = \sum_{i=1}^{N} [t_{si} + Qt_{pi} + t_{hi} + t_{ni}] \qquad (1)$$

For example, consider an order that consists of 20 identical parts. Five operations must be performed to produce this part. Setup, processing, handling, and nonoperation times are provided in the **table**. The production lead time for this order as computed from Eq. (1) is 44.7 h. Assuming a 40-h work week, the production lead time for the order of 20 parts is 1.1 weeks.

Setting due dates. Knowledge of the lead time for a product is important for setting due dates for orders. When a customer places an order, either the customer states when the order is needed or asks for an estimated time of completion. In either case, a method of estimating lead time is required. If the customer requires a specific delivery date, the company needs to be able to determine if the customer due date is achievable given current production orders. If the customer requested date is not feasible, an estimate of when the order will be done is necessary for negotiating a new due date.

One method for setting due dates requires knowledge of the average production lead time and its standard deviation (σ_{LT}). Given this information, an estimated lead time is calculated such that there is a good probability of completing the order on time. The percentage of orders completed on time is known as the service level (s), where z_s denotes the value on the measurement axis (for which s of the area under the z curve lies to the right of z_s). The z curve is the standard normal curve. The standard normal table lists the values of z_s for values of s ranging from 0 to 100%. To be more precise, the table actually works in reverse in that for values of z_s ranging from −3.49 to +3.39, the value of s is given. Assuming that production lead time is normally distributed, the quoted lead time for an order can be computed by Eq. (2).

$$\text{Estimated lead time} = \mathrm{LT} + z_s\sigma_{\mathrm{LT}} \qquad (2)$$

For example, suppose the average production lead time is 2 weeks with a standard deviation of 0.75 week. If a 95% service level is desired, an estimated lead time for the order is computed to be 3.23 weeks in Eq. (3). The value of 1.645for z_s was obtained by

$$\begin{aligned}\text{Estimated lead time} &= 2 \text{ weeks} + (1.645)(0.75)\\ &= 3.23 \text{ weeks} \qquad (3)\end{aligned}$$

referencing a standard normal table. First, the value 0.95 is found in the body of the table (corresponds to $s = 95\%$). The corresponding row and column label values are added together to obtain the value of z_s. The additional 1.23 weeks (1.645×0.75) is often called the safety lead time. Caution should be exercised when using this method to estimate lead times when the standard deviation is very high. A large standard deviation will lead to a large safety lead time. Since one-half of all orders are expected to complete in less than the average lead time, a large safety lead time can lead to a large finished goods inventory if orders cannot be delivered early.

Managing lead time. Production lead time can make the difference between receiving customer orders and going bankrupt. In all industries, a short production lead time is preferred to a long time. One way to reduce production lead time is to reduce the individual components, such as setup, handling, and nonoperation times. In some industries (such as metalworking factories), these three components constitute 95% of the total production lead time. Setup time can be reduced through the use of common fixtures and by scheduling similar orders to run back to back. Handling time can be reduced by improving the layout of the facility. Machines that follow one another

Data for lead time calculation

Operation (i)	Setup time (t_s), h	Processing time (t_p), h/part	Handling time (t_h), h	Nonoperation time (t_n), h
1	2	0.10	0.25	3
2	3	0.05	0.25	4
3	1	0.07	0.25	2
4	2	0.15	0.25	6
5	4	0.20	0.25	5

in the processing sequence of many products should be located near one another. Finally, nonoperation time can be reduced by efficient scheduling practices and the elimination of unnecessary inspections and other nonoperation activities.

Production lead time is also related to the amount of work in process (orders currently being worked on in the plant) on the production floor and the production rates of the machines. This relationship, shown in Eq. (4), is known as Little's law. For indus-

$$\text{Production lead time} = \frac{\text{Work in process}}{\text{Production rate}} \qquad (4)$$

trial plants operating at less than 100% capacity, production lead time is largely unaffected by the amount of work in process in the plant. However, as the production rate reaches its maximum value (capacity), increasing the work in process merely serves to increase the production lead time. Therefore, controlling the release rate of materials to the plant is an important part of controlling production lead time. Little's law also shows that production lead time can be reduced by increasing the production rate of the plant. This approach typically requires an increase in the plant's capacity. Capacity can be increased by purchasing additional equipment or by working more hours (for example, authorizing overtime or adding another shift).

Production lead time is essential to the success of a company. In the global market place, customers demand low-cost, high-quality products in a short period of time. As a result, effective management of production lead time is as important as striving for low cost and high quality. Short production lead times allow a company to be responsive to its customers' needs. *See* INDUSTRIAL ENGINEERING; INVENTORY CONTROL; MANUFACTURING ENGINEERING; MATERIAL RESOURCE PLANNING; MATERIALS HANDLING; PRODUCTION PLANNING; QUALITY CONTROL in the McGraw-Hill Encyclopedia of Science & Technology.

Elin M. Wicks

Bibliography. M. P. Groover, *Automation, Production Systems, and Computer Integrated Manufacturing*, Prentice Hall, 2d ed., 2000; W. J. Hopp and M. L. Spearman, *Factory Physics*, Richard D. Irwin, 2d ed., 2000.

Production planning

The function of a manufacturing enterprise responsible for the planning, scheduling, and coordination of all production activities. The planning phase involves forecasting demand and translating the demand forecast into a production plan that optimizes the company's objective, which is usually to maximize profit while optimizing customer satisfaction. The twin objectives of maximum profit and maximum customer satisfaction are not always compatible. Sometimes, a highly-valued customer may request a product that must be made under high-cost conditions. During the scheduling phase the production plan is translated into a detailed, sometimes hour-by-hour, schedule of

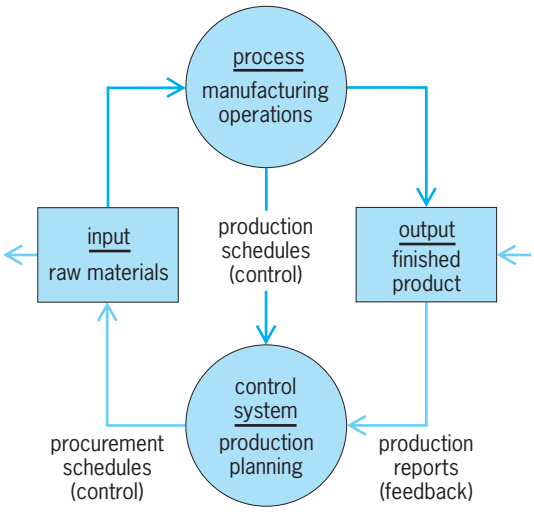

The production process as an input-output process.

products to be made. During the coordination phase, actual product output is compared with scheduled product output, and this information is used to adjust production plans and production schedules. *See* ACTIVITY-BASED COSTING; OPTIMIZATION.

If the production or manufacturing process is viewed as an input-output process, then the production planning function can be viewed as a control process with feedback (see **illus.**). The control is in the form of schedules and plans, while the feedback results from the comparison of the production reports with the production schedules. *See* CONTROL SYSTEMS.

Techniques. In the planning process, material requirements planning (MRP-I) or manufacturing resources planning (MRP-II) may be used. MRP-I is an inventory or materials planning tool, and MRP-II is a resources planning tool that includes planning for labor hours, machine hours, cash, floor space, and so forth. The primary difference is that MRP-II includes capacity resources planning (CRP). Essentially, MRP techniques are based on a master schedule that is created for delivery of product. A bill of materials (BOM) is used to establish the detailed schedule for all components, subassemblies, assemblies, and deliveries. Just-in-time (JIT) is both a technique and a philosophy of manufacturing planning and control. The objective is to have no inventory in the plant that is not being worked on actively. Ideally the work-in-process (WIP) should be zero under a just-in-time system. Usually true just-in-time is not achieved but serves as a goal. *See* MATERIAL RESOURCE PLANNING.

The quantity of material ordered at any given time may be determined either by the MRP system or by use of economic order quantity (EOQ) or economic lot sizing (ELS) techniques, which consider cost of carrying inventory and the fixed costs associated with placing an order. The economic order point (EOP) and the reorder point (ROP) are used to determine when an order or a reorder is placed as a function either of time or of inventory level.

The scheduling function determines the routing of parts through the shop and the loading of

machines. The objective should be to balance the work load on machines, departments, and the entire plant. Computer-aided process planning (CAPP) techniques are helpful in establishing the necessary operations, machines, and the sequences of operations. This information is useful for machine loading and part routing. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

The operations on the shop floor may be controlled by visible records such as cards (in Japanese, *kanbans*) when the factory is using a just-in-time system. In a just-in-time system, every part is expected to arrive at its point of use just as the operation is ready to use it. Moving a *kanban* upstream is equivalent to placing an order, and moving a *kanban* downstream corresponds to the delivery of goods. Modern computer technology has put terminals on the shop floor. These terminals are used to download schedules and work assignments and to upload completion of product, hours used, and so forth; they tend to remove the paperwork from the shop floor. Many different types of charting techniques are available to accomplish these tasks. The Gantt chart, the basis of many charts, is the oldest such technique. *See* COMPUTER; DIGITAL COMPUTER; GANTT CHART.

Constraints. Production planning generally has several constraints: demands must be met within reasonable limits; inventories of raw materials, in-process materials, and finished goods must not be excessively large or small; variations in personnel are usually limited; and the available production facilities are limited. A customer will not wait indefinitely for a product to be delivered. The length of time that is acceptable is dependent on the industry and the product. Inventories represent an investment in material that is not producing immediate revenue and therefore must be limited. Some inventories are necessary for efficient manufacturing in lot sizes and for balancing personnel and equipment requirements with customer demand. *See* INVENTORY CONTROL; PRODUCTION ENGINEERING; PRODUCTION SYSTEMS.

Applications. The functions of production planning are dependent upon the type of industry and type of product. If the manufacturing process is of an assembly-line nature, the planning is primarily of outputs and inputs and the avoidance of bottlenecks. If the manufacturing process is of the job-shop type, that is, if the product is made to the customer's order, the planning must involve a detailed schedule of individual parts through the manufacturing process. In both cases, the planning starts with a forecast of outputs or demands and proceeds to determine personnel and material inputs consistent with demands and company policy. Some of the readily identifiable tasks within the production planning function are planning, scheduling, dispatching, machine loading, expediting, inventory record keeping, and material requisitioning. All of these tasks involve the collection, analysis, interpretation, and dissemination of data. The sophistication of the data system is highly variable from company to company and also within departments or branches of a company.

The production control function does not exist independently of other functions within the company. Its activities must be coordinated with purchasing, sales, finance, production, tool engineering, design engineering, maintenance, and so on. Its primary responsibility is to coordinate the manufacture of the product to best satisfy the customer and the constraints imposed by the company management. *See* OPERATIONS RESEARCH. John E. Biegel

Bibliography. A. Ahmadian, R. Afifi, and W. D. Chandler, *Readings in Production and Operations Management: A Productivity Approach*, 1990; E. S. Buffa and R. K. Sarin, *Modern Production Operations Management*, 8th ed., 1987; J. L. Riggs, *Production Systems: Planning, Analysis, and Control*, 1987, reprint 1992; T. E. Vollmann, W. L. Berry, and D. C. Whybark, *Manufacturing, Planning and Control Systems*, 4th ed., 1997.

Production systems

A collection of workstations that perform operations such as manufacturing, assembly, inspection, finishing, and testing to create products. A workstation is a collection of resources that performs the same or similar operations for the same set of products. The resources at a workstation may be automated machines, machines operated by humans, or human operators performing manual operations. The storage and delivery of both raw materials and finished products are also part of the production system, but the focus here will be on the system that transforms materials and/or components into products.

The design or configuration of workstations in a production system is defined by the flow of products in the system. The flow of products is determined by how products move among the workstations during production, and is dictated by the volume and variety of products that the system will produce. In general, a simpler flow corresponds to lower product variety at higher volumes, and a more complex flow corresponds to higher product variety at lower volumes.

The specific production operations performed at workstations will vary greatly depending on the nature of the product and its requirements. For example, the production of automobile body components is performed through a series of metal-forming steps. The assembly of automobile bodies is generally accomplished in a set of automated welding process steps. The final assembly of the finished automobile is typically completed in a separate series of manual operations. Not all production systems are large and require the use of expensive machinery—for example, fast-food restaurants, office work, construction operations, and many others are production systems.

The design and operation of a production system can be quite complex and requires that many interrelated decisions be made. Examples of design decisions are the assignment of work to workstations (line balancing), the sizing of storage space for work-in-process between workstations, the methods used to control production, and the methods used for moving product from workstation to workstation. Examples of operational decisions are setting staff

		Product variety					
		Very low	Low	Moderate	High		
Product volume	Low				Job shop	High	Product flow complexity
	Moderate			Group technology/ batch production		Moderate	
	High		Production line			Low	
	Very High	Continuous flow				Low	

The relationship between product variety and volume, production system type, and product flow complexity.

levels and shift lengths, production scheduling, and production system bottleneck isolation and improvement.

Production system configurations. The flow of products among the workstations in a production system will define its configuration or design. The type and complexity of product flow designed into a production system are determined by the volume and variety of products produced. At the two ends of the flow complexity spectrum are production lines with simple flow and job shops with complex flow. In between, there are hybrid systems called batch production systems and group technology systems (see **illustration**). In reality, production facilities may be a combination of several types of production system configurations.

A special case is that of a fixed-position production system, where the product stays fixed in position and various workstations move to the product to complete their operations. Examples of such products are ships, airliners, and buildings. These products are often unique or are produced this way because of the product size. Conceptually, these systems can viewed as very large job-shop and assembly systems.

Production lines and continuous-flow systems. Production lines and continuous-flow systems have very simple product flow. Product flow in these systems is conceptually a straight line, with all products moving through the same sequence of workstations. Examples are production lines such as automobile final assembly systems and fast-food preparation systems. In these systems, parts are distinct and countable (such as automobiles and sandwiches). These are referred to as discrete product production lines. Production systems with straight-line flow that produce fluids and powders are referred to as continuous-flow production systems, and are common in the chemical, food, and cosmetics industries.

Production lines and continuous-flow systems are normally high-volume systems. In early production lines, it was necessary that the operations performed at each workstation on each product were almost always the same to achieve high volume. Modern production lines can incorporate more product variety than in the past, but are still primarily limited to variations of a basic product architecture or "platform."

Early production lines consisted of workstations of manually operated machines or of human operators performing manual operations. The advent of computer-controlled machinery has permitted many workstations previously requiring manual operation to become fully automated. A modern automobile body assembly system is an example of where products are automatically moved from workstation to workstation, consisting of machinery and robots performing welding operations. As impressive as fully automated production lines are, many products are still produced most efficiently using manual operations. The prior examples of automobile final assembly and fast-food preparation come to mind.

Job-shop production systems. Job-shop production systems are normally intended for applications where the products are made either in single units or in limited production runs. The intent of job-shop production systems is to produce a high variety of products (in some cases an unlimited variety of custom products) at low volumes. Job shops are typically organized by process. For example, a manufacturing job shop may consist of separate areas for milling, drilling, and turning. The product flow in a job shop is very complex since each product will usually require a separate sequence of operations (called a product routing), and there is essentially no set limit on the number of times a product may have to visit the various processes. From a product flow viewpoint, a hospital is organized as a job shop. Patients (analogous to products in manufacturing) may need to visit a number of different processes, such as admissions, radiology, and orthopedics, but not necessarily in the same order or the same number of times.

Similar to what has been seen in modern production lines, computer-controlled machinery has also had an impact on job shops. Operations that previously required manually operated machinery and highly skilled labor have been replaced in many situations by computer-controlled machinery. This is more common in job shops that produce a number of products in larger volume, since there is relatively high overhead invested in programming of machines for different products. This requires higher volumes to be economical. Job shops strictly producing custom and prototype products are still commonly staffed with highly skilled workers, such as machinists, toolmakers, and diemakers, who perform operations using of a variety of general-purpose machine tools.

Hybrid systems. Because many products or sets of related products are neither high-volume nor low-volume, as would be found in production lines and job shops, hybrid systems have evolved. These systems strike a compromise between volume and variety, falling between that of production lines and job shops. Both batch and group technology production systems can be considered hybrid systems. The fundamental concept implemented in these types of systems is to group similar products and maintain a separate production system for each group. Differentiation of group technology from batch production is found in how product groups are formed, although there is no clear distinguishing boundary. In batch production, grouping of products is relatively simple. A criterion, such as size, is used to group products. An example of a batch production system is automotive sheet metal stamping (forming). Separate stamping lines, which accommodate tooling of approximately the same size and weight, are used to produce a group of similar-sized products. In group technology, the criteria for grouping products is generally more complex and is based on similarities in production operations and routings, tooling, and components. In this sense, group technology may be considered a special case of batch production. Grouping is done to minimize changeover time and cost when switching production to different products. In electronic circuit board assembly, a group of products may use components with similar circuit board insertion methods and similar size, thus permitting the use of the same automated machinery without the need of long setup times between product batches.

The production systems for a specific group of products within a hybrid system may fall anywhere between a job shop and production line, depending on the requirements of products in the group. In the sheet metal stamping example, the production system for a group is a production line and is relatively high-volume. In the electronic circuit board assembly example, the production system may contain automated machinery, which also can accommodate several different product routings. In this sense, the production system has some of the flexibility characteristic of a job shop.

Production control and scheduling. An important feature of production systems, regardless of type, is the method or methods used to control product flow into the production system, and how waiting products are processed at workstations. This is determined by the methods for initiating new production or production control, and how work at individual workstations is scheduled.

The two major categories of production control are "push" and "pull" production control systems. Production plants may use both types of production control at different areas or subsystems within the plant. In a push production control system, initiation of production is based on future orders and forecasted future demand. In a pull production control system, new production is initiated to replenish inventory that has been consumed. The decision on what method to use depends on many criteria, but demand variability, inventory and shortage costs, and production lead time are critical factors. As a simplified example, consider two different types of fast food: hamburgers and fried chicken. In anticipation of mealtime demand, a burger store may cook several burgers in advance (similar to a push system) and then cook more only to replace burgers purchased. This pull system works because the time to cook a burger is short. In contrast, the fried chicken store must cook a much larger quantity of chicken in anticipation of mealtime demand, as well as any large catering and phone orders received. More chicken is cooked in anticipation of demand and not necessarily because of chicken purchased. This push system is used since the time to cook chicken is much longer than the time to cook a burger. The shortage cost in terms of lost business and customer good will is much greater for the chicken store. *See* INVENTORY CONTROL.

The scheduling of work at workstations is a complicated problem that affects the performance of the production system. In many production systems, scheduling equates to sequencing or ordering of the products waiting for work at a workstation. In a production line, scheduling may only be necessary before the first workstation since all products follow the same routing. Depending on how products are physically moved from workstation to workstation, changing the sequence in a production line may be very difficult. Since volume is a critical performance measure for production lines, scheduling of jobs is often executed to maximize production volume. Software packages are often used to help sequence jobs in production lines.

In job shops, scheduling is much more complicated. Because of high product variety and the number of different product routings, the products available for scheduling are different at different workstations and are constantly changing. The criteria used for scheduling products are also more complicated. Some common objectives of scheduling are to minimize the average time that products are in the production system, and to minimize the average lateness of products. Some software exists that can help schedule work in jobs shops. In many cases, scheduling is performed by employees based on years of experience and intuition. *See* SCHEDULING.

Performance measures. There are many ways to measure the performance of a production system. In today's competitive business environment, it is a given that a production system must produce the correct product at the required quality level. The efficiency at which this is achieved will vary. From a production systems perspective where the concept of product flow is applicable to many different products and situations, there are three main interrelated performance measures often used that answer the basic questions of how fast, how long, and how much.

1. Throughput (for example, products/hour), measures how much product a system can produce in a given time.

2. Time-in-system, flow time, or cycle time is a measure of how long it takes to produce a product

from start to finish, and reflects the responsiveness of the system to changes in customer demand.

3. Work-in-progress is a measure of how much partially completed product is typically found in the production system, and is related to the cost of production and the responsiveness of the system.

Total systems approach. The resources, processes, and methods used to produce, store, and deliver products often constitute a large and complex production system. This system exists to meet the goals and objectives of some organization. As such, the design and operation of the system and its components and subsystems should be evaluated in the context of the total production system.

Other critical components of the production system include the layout or arrangement of workstations in a facility, and the equipment and methods used to move and store material and products in the facility.

At the workstation level, many important decisions must be made about the operations and methods for completing the operations. In more typical or traditional production systems, this includes detailed decisions about manufacturing processes such as casting and molding, shearing and forming, machining/material removal, heat treating, finishing, assembly, and inspection. *See* INDUSTRIAL ENGINEERING; MANUFACTURING ENGINEERING. David S. Kim

Bibliography. R. G. Askin and C. R. Standridge, *Modeling and Analysis of Manufacturing Systems*, 1993; W. J. Hopp and M. L. Spearman, *Factory Physics: Foundations of Manufacturing Management*, 2d ed., 2001; S. Nahmias, *Production and Operations Analysis*, 5th ed., 2005.

Productivity

In a business or industrial context, the ratio of output production to input effort. The productivity ratio is an indicator of the efficiency with which an enterprise converts its resources (inputs) into finished goods or services (outputs). If the goal is to increase productivity, this can be done by producing more output with the same level of input. Productivity can also be increased by producing the same output with fewer inputs. One problem with trying to measure productivity is that a decision must be made in terms of identifying the inputs and outputs and how they will be measured. This is relatively easy when productivity of an individual is considered, but it becomes difficult when productivity involves a whole company or a nation.

Calculating productivity. A simple example of productivity is a paper carrier on a bicycle. If the paper carrier delivers 100 papers in 50 min, his productivity is 2 papers per minute. The output is the number of papers delivered. The input is the time required to deliver the papers (labor minutes). To calculate the productivity, the output is divided by the input (100 papers by 50 min). Later, the paper carrier analyzes the route taken to deliver the papers. Changes in the route are made to eliminate some backtracking and to take some additional shortcuts. Now when the carrier delivers the same 100 papers, only 40 min are required. Thus, the new value for productivity is 100 papers divided by 40 min, or 2.5 papers per minute. The carrier has increased productivity by using the available resources (time) more effectively. Note that if productivity is calculated based on time, different units could be selected. Instead of minutes, the time units could be hours. The value 2.5 papers per minute could be converted to 150 papers per hour. Many industry and government productivity figures use hours as the time base.

For large organizations, it is difficult to determine every input that goes into making a product or providing a service. If the subject of a study is an automobile manufacturer, the many parts and services required to produce each automobile must be taken into consideration. Some of them are easier to quantify, such as the amount of steel, vinyl, and other materials that go into each automobile; the number of labor hours required to assemble each automobile (direct labor); and the cost of storing the parts in the factory while awaiting assembly. However, there are many factors that are not directly involved with the production of the automobile without which the company could not produce automobiles. These factors are more difficult to determine, but they are still inputs to the production process. Some of these factors include the cost of the required secretarial services, the cost of engineers that designed the automobile and set up the production line, and the cost to maintain and pay taxes on the factory.

Productivity measures. Industry and government officials have adopted three common types of productivity measures. However these measures are not exclusive. Many companies develop their own measures based on their unique situations. Therefore, it is often difficult to compare the productivity figures of different companies or nations. A comparison of productivity figures can be made only for those values that are calculated by using the same type of measure.

Partial productivity. In this simplest type of productivity measure, a single type of input is selected for the productivity ratio.

The previous example of paper carrier productivity uses a partial productivity measure where direct labor minutes is the input. Typically the company or organization selects an input factor that it monitors in daily activity. Direct labor hours is a factor that most companies monitor because they pay their employees based on hours worked. Therefore, the U.S. Department of Labor uses output per employee hour as a productivity measure in many of its published indices. Other factors that are commonly used in partial productivity ratios are capital (money invested in facilities and equipment), materials, or energy.

The major problem with a partial productivity measure is that it does not include all of the inputs required to make a product or provide a service. In the paper delivery example, if the carrier buys a car to deliver the papers, the delivery time can be reduced to 20 min. Based on the same partial productivity measure, the new productivity would be 5 papers per minute, or double what could be achieved by

using a bicycle. However, if a partial, productivity measure based on capital is used, the carriers productivity would actually decline, because a car is much more expensive than a bicycle. It would now require a great deal more capital to deliver the papers. This example illustrates the problem with choosing a single factor upon which to base a productivity measure.

Total factor productivity. This measure combines labor and capital, two of the most common input factors used in the partial productivity measure. This measure is often used at the national level, because many governments collect statistics on both labor and capital. In calculating at the national level, the gross national product (GNP) is used as the output. A problem with this measure is that it does not include some types of inputs required to yield products.

Total productivity. This measure incorporates all the inputs required to make a product or provide a service. The inputs could be grouped in various categories as long as they determine the total inputs required to produce an output.

Because labor, capital, material, and energy all have different units, they cannot be directly added. Normally, they are all converted to constant standard-currency amounts before they are added. An advantage of using total productivity is that it includes all the inputs required to achieve the product. A disadvantage is that companies or governments often do not have the data required to determine all the input factors.

Factors affecting productivity. Many factors affect productivity. Some general categories for these factors are product, process, labor force, capacity, external influences, and quality.

Product. The type of product and the design affects productivity. Ease of manufacture, assembly, and maintenance have come to be emphasized. Materials that are easier and cheaper to process have been developed. These factors have had a positive effect on productivity.

Process. The process that is used to manufacture a product or perform a service can affect the productivity. A simple example is a worker using a manual drill compared to worker using an electric drill. The capital difference between the two drills is small, but the worker with the electric drill will drill holes faster.

Labor force. Many factors in this category affect productivity. If a good working environment, with proper training and incentives, is provided, the employees will be more productive. The type of supervision and relationship with management can also affect productivity. Employees must also have the proper tools and supplies to perform their task.

Capacity. The capacity of the company will affect the productivity. If a company buys expensive equipment and builds a large factory but uses only a small part of it, productivity will be low. To raise productivity, the company must make more product with the same capital investment. The facility and equipment will be idle a smaller portion of the time, and productivity will increase.

External influences. There are factors that a company cannot control that affect productivity. Consumer demand is one factor. Generally, a company will not produce more units than consumers will buy. If demand declines, the company will reduce the output. In this case the company will either have idle capacity or will switch capacity to different products. If idle capacity increases, productivity will decline.

Government regulation can affect productivity. In many countries, regulations require companies to provide safety equipment and safe conditions for employees. They also require companies to dispose of hazardous chemicals in environmentally safe ways. Some of these regulations increase costs significantly. Without these regulations, some companies would be irresponsible to both their employees and the environment. In the short term, total cost would be reduced and productivity would go up. However, in the long term, negative effects could be experienced by many.

Quality. Increasing quality can both help and hurt productivity. Quality awareness has caused many companies to adopt more productive processes and to design products that can be produced more efficiently. In turn, the cost has been lowered and customer demand raised. All of these factors help to raise productivity. However, sometimes more expensive processes or materials must be used to increase quality. This can increase total cost and reduce total productivity.

Productivity improvement plans. There are many different plans that companies develop in an attempt to improve productivity. Wage incentive plans and changes in management structure are two ways that companies focus on the labor force. Investment in research and development allows companies to develop new products and processes that are more productive. Quality improvement programs can reduce waste and provide more competitive products at a lower cost. *See* METHODS ENGINEERING; OPERATIONS RESEARCH; PRODUCTION PLANNING.

Gregory L. Tonkay

Bibliography. D. I. Cleland and B. Bidanda (eds.), *The Automated Factory Handbook: Technology and Management*, 1990; W. K. Hodson (ed.), *Maynard's Industrial Engineering Handbook*, 1992; M. E. Mundel, Measures of productivity, *Ind. Eng.*, 8(5):32–36, 1976; R. G. Schroeder, *Operations Management: Decision Making in the Operations Function*, 4th ed., 1993; D. S. Sink, *Productivity Measurement Improvement, Evaluation, and Control*, 1985.

Progesterone

A steroid hormone produced in the corpus luteum and placenta. The hormone has an important physiological role in the luteal phase of the menstrual cycle and in the maintenance of pregnancy. In addition, progesterone produced in the testis and adrenals has a key role as an intermediate in the biosynthesis of

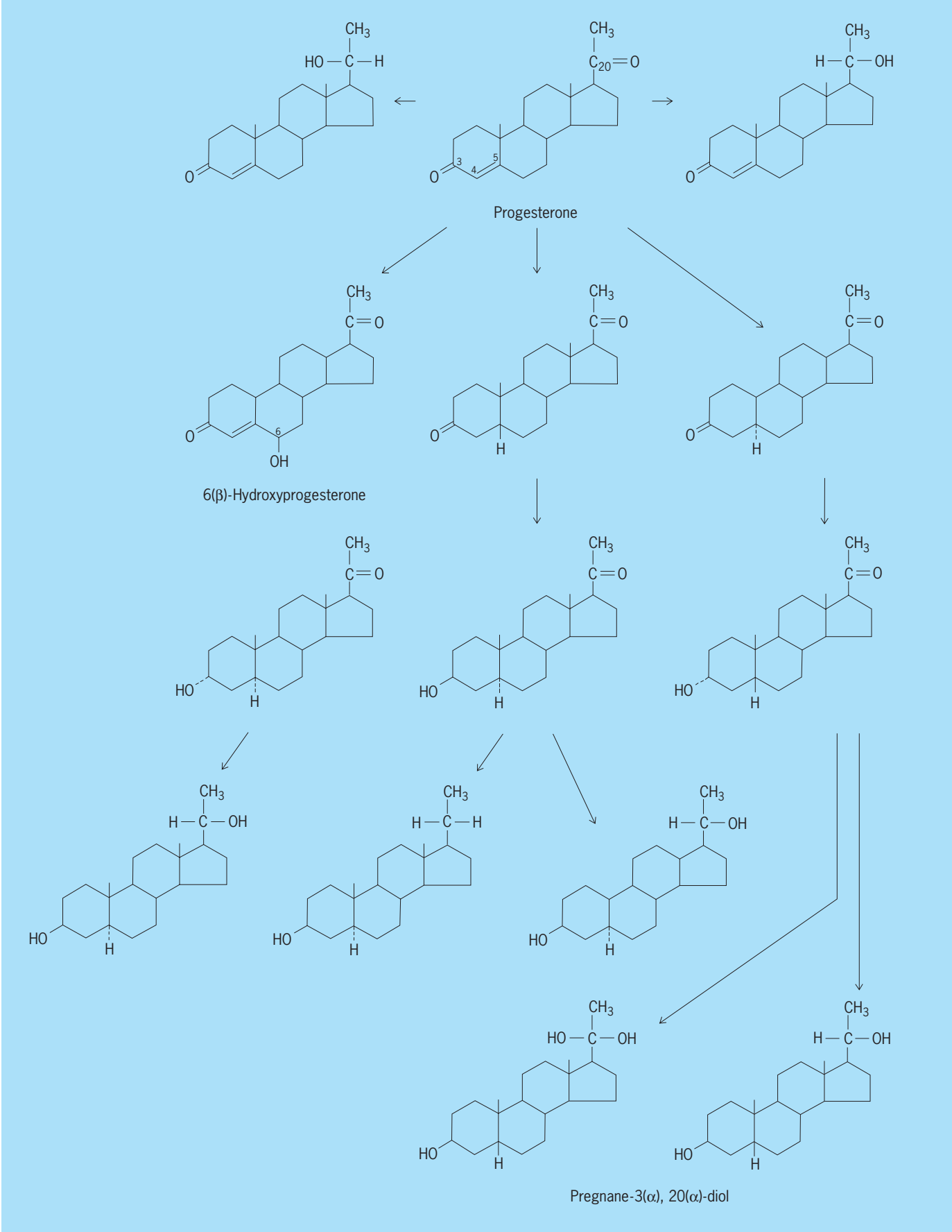

Fig. 1. Catabolic pathways of progesterone.

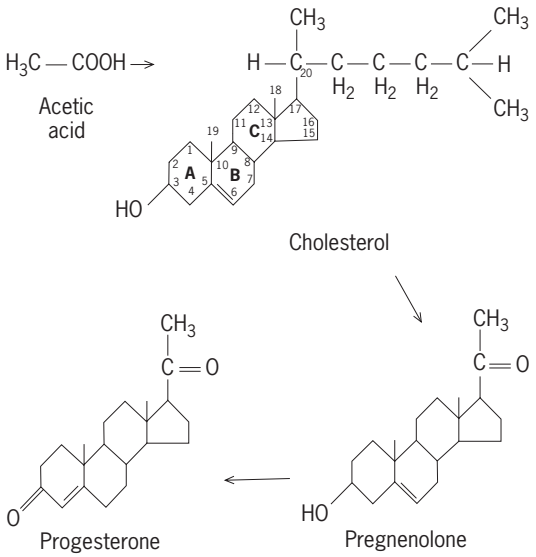

Fig. 2. Biosynthetic pathway for progesterone.

androgens, estrogens, and the corticoids (adrenal cortex steriod).

Catabolism of progesterone involves reductive reactions at carbon atoms 3, 4, 5, and 20, as well as hydroxylation carbon atom 6 (**Fig. 1**). The biosynthetic pathway to progesterone (**Fig. 2**) is common to all steroid-producing tissues and involves acetate→cholesterol→pregnenolone→progesterone.

The blood contains, in addition to progesterone, the 20(α)- and 20(β)-reduced metabolites, which have reduced biological activity.

Pregnane-3(α),20(α)-diol is, quantitatively, the most important metabolite in humans, and its determination in the urine is of clinical importance in conditions involving menstrual irregularities and also in abnormal pregnancies. *See* ANDROGENS; CHOLESTEROL; ESTROGEN; PREGNANCY; STEROID; STEROL.

Ralph I. Dorfman

Bibliography. C. W. Bardin et al. (eds.), *Progesterone and Progestins*, 1982; E. Gurpide (ed.), *Biochemical Actions of Progesterone and Progestins*, 1977.

Programmable controllers

Electronic computers, designed for durability in industrial environments, that are used for the control of machines and manufacturing processes through the implementation of specific functions such as logic, sequencing, timing, counting, and arithmetic. They are also known as programmable logic controllers (PLCs). Historically, process control of a single or a few related devices has been implemented through the use of banks of relays and relay logic for both the control of actuators and their sequencing. The advent of small, inexpensive microprocessors and single-chip computers, or microcontroller units, brought process control from the age of simple relay control to one of electronic digital control while neither losing traditional design methods such as relay ladder diagrams nor restricting their programming to that single paradigm. The computational power of programmable controllers and their integration into networks has led to their having capabilities approaching those of distributed control systems, and plantwide control is now a mixture of distributed control systems and programmable controllers. Applications for programmable controllers range from small-scale, local process applications in which as few as 10 simple feedback control loops are implemented, up to large-scale, remote supervisory process applications in which 50 or more process control loops spread across the facility are implemented. Typical applications include batch process control and materials handling in the chemical industry, machining and test-stand control and data acquisition in the manufacturing industry, wood cutting and chip handling in the lumber industry, filling and packaging in food industries, and furnace and rolling-mill controls in the metal industry. *See* DIGITAL COMPUTER; DISTRIBUTED SYSTEMS (CONTROL SYSTEMS); MICROPROCESSOR.

Although programmable controllers have been available since the mid-1970s, developments—such as the ready availability of local area networks (LANs) in the industrial environment, standardized hardware interfaces for manufacturer interchangability, and computer software to allow specification of the control process in both traditional (ladder logic) and more modern notations such as that of finite-state machines—have made them even more desirable for industrial process control. *See* LOCAL-AREA NETWORKS.

Implementation. Programmable logic controllers are typically implemented by using commonly available microprocessors combined with standard and custom interface boards which provide level conversion, isolation, and signal conditioning and amplification. Microprocessors used in programmable controllers are similar or the same as those used in personal computers. A significant difference between personal computers and programmable controllers is the type of software operating system. The software of a programmable controller must respond to interrupts and be a real-time operating system, and these are characteristics which the typical operating system of a personal computer does not possess. *See* MICROCOMPUTER; OPERATING SYSTEM; REAL-TIME SYSTEMS; SOFTWARE.

Connectivity. Perhaps the biggest benefit of programmable controllers is their small size, which allows computational power to be placed immediately adjacent to the machinery to be controlled, as well as their durability, which allows them to operate in harsh environments. This proximity of programmable controllers to the equipment that they control allows them to effect the sensing of the process and control of the machinery through a reduced number of wires, which reduces installation and maintenance costs. The proximity of programmable controllers to processes also improves the quality of the sensor data since it reduces line lengths, which can introduce noise and affect sensor

calibration. Communication devices in programmable controllers, such as synchronous and asynchronous serial communications ports, allow these distributed programmable controllers to be connected through a local-area network to an operator interface for monitoring and system control, or connected to a workstation for program development and downloading of operational software into the programmable controllers. The versatility of a local-area network also allows for remote location of a central supervisory-monitor computer, as well as for field personnel to connect locally to the local-area network for local control while still monitoring signals that are passing through the local-area network-based system. The extensive computational capability of programmable controllers also allows for local monitoring of such conditions as low battery voltages, lack of sufficient cooling, or other failures of its own control capability. Depending on the extent of the failure, the programmable controller may trigger a local alarm or pass the information back to a central supervisory computer. This monitoring can also be extended to the monitoring of the process under control itself, with local analysis significantly reducing the amount of data that must be transferred over a network for centralized control and monitoring.

The availability of a local-area network also allows for the remote programming of local microcontroller units on the fly. That is, one control algorithm running in microcontroller unit can be replaced with another control algorithm from the remote site without local operator intervention. Hardware options can also be reconfigured and their status maintained in nonvolatile read-write memory, which is readily available in some programmable controllers, and as electrically erasable programmable read-only memory (EEPROM). Troubleshooting can be controlled from a remote site by downloading diagnostic software into the local programmable controller. In the most general sense, every programmable controller can be identical to every other one, with the remote supervisory computer downloading to each local programmable controller the program that is appropriate to the device or process that is being controlled. *See* COMPUTER STORAGE TECHNOLOGY; SEMICONDUCTOR MEMORIES.

Open architecture. Open architecture allows standardization of local-area-network interfaces, host software, operating systems, field buses, and programmable controller backplanes. This open architecture allows process-control engineers to integrate the programmable controllers from various manufacturers that are best suited to individual tasks into a comprehensive customized control system. This approach also allows for an easy upgrade path in which one manufacturer may have made a complete installation and, at a later date, a more advanced system from another manufacturer which has become available can be integrated with components of the first installation.

Beyond logic control. Because of the increased computational power associated with programmable controllers, control algorithms more complex than simple ladder logic can be implemented, and libraries of standard control algorithms are made readily available by manufacturers. This availability allows the process-control engineer to spend more time on the development of the control strategy and the associated parameters than on the debugging and verification of in-house-written software. This increased level of sophistication and partitioning of the design problem has led to software packages for cross-platform development (that is, development that can be carried out with various operating systems) of operating software for programmable controllers. In addition to libraries of control algorithms, signal-processing routines are available for the conditioning and compensation of sensor data from simple smoothing to spectral component analysis.

Software development. Cross-platform development of programmable controller software allows the use of more powerful computers such as personal computers or workstations in the designing, debugging, development, and downloading of this software. Additionally, the quality and user-oriented interface of the development software enables the user to develop the control program either by using traditional ladder-logic diagrams, which are shown graphically on the development computer's display, or through other methods such as state diagrams, sequential function charts, or function blocks. This improved user-interface capability makes programmable controllers more accessible to a broader range of process-control personnel who have different backgrounds and different working languages. Software for individual components of programmable controllers, such as a proportional-integral-derivative (PID) control algorithm for a robotic arm, are not programmed by the process-control engineer, but are indirectly programmed by the engineer inserting the relevant parameters into blanks in on-screen templates (**Fig. 1**). These parameters are then automatically transformed into fully functional software modules, which are interfaced to other components of the

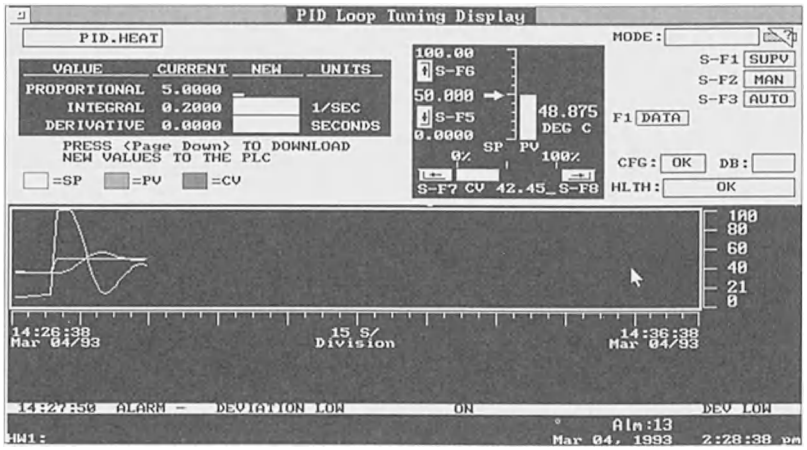

Fig. 1. Windows-based operator interface for the specification of parameters to be used for the tuning of a proportional-integral-derivative (PID) controller. (*After Allen-Bradley, Take Control of Your Application with Process Configuration and Operation (PCO) Software, Publ. 6190-2.1, 1993*)

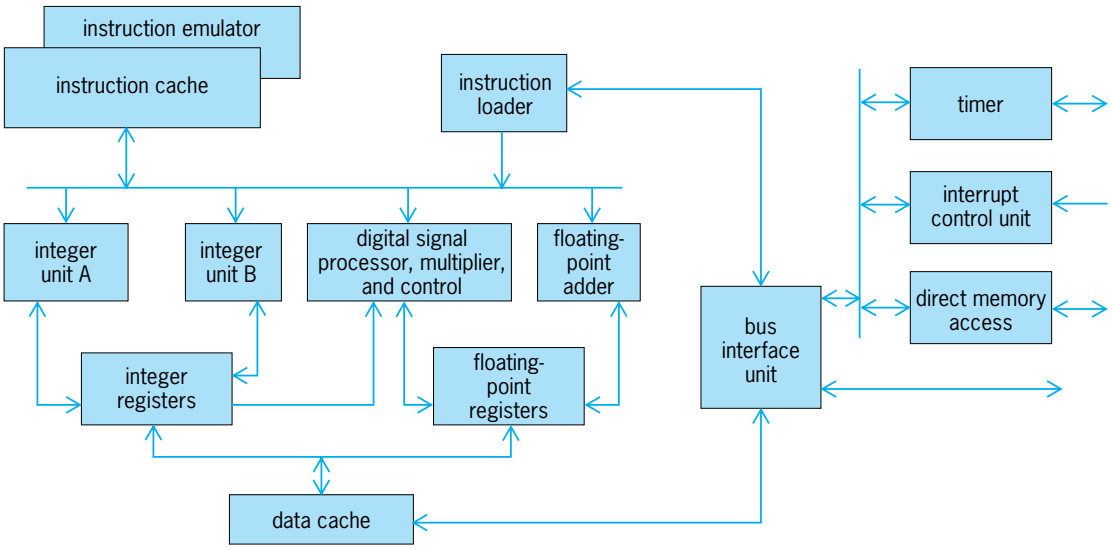

Fig. 2. Functional architecture of a typical high-performance single-chip microcontroller used for control and signal processing. (*National Semiconductor Corp.*)

control process. Another capability of cross-platform development software is the ability to simulate the control process before it is programmed into the programmable controller and connected to operating machinery.

Microcontrollers. Because of their self-contained input/output capability and the incorporation of customized on-chip coprocessors, single-chip microcontrollers are used extensively for the embedded control of processes in stand-alone equipment that is not associated with industrial control. These microcontroller units must be considered programmable controllers from a functional point of view, even though they are not used directly in an industrial setting. Microcontroller units are single-chip versions of general-purpose computers which have all five components that define a computer resident in a single integrated circuit. That is, a single integrated circuit contains the arithmetic logic unit (ALU), the control section, a local memory which may or may not be expandable off-chip, an input section, and an output section. An unusual feature of microcontroller units is their on-chip input/output section and signal coprocessors, which make them particularly suitable for small embedded control systems. For example, 32-bit-word-length microcontrollers are readily available which contain on-chip digital signal processing capability of sufficient power to implement a stand-alone facsimile send/receive system while simultaneously controlling the associated printer motors and user-interface requirements. These embedded processors are used to control devices from laser printers to home security systems, as well as automobile engine ignition timing and fuel injection for the reduction of pollution (**Fig. 2**). *See* EMBEDDED SYSTEMS; INTEGRATED CIRCUITS. Kenneth J. Hintz

Bibliography. Allen-Bradley, *Take Control of your Application with Process Configuration and Operation (PCO) Software*, Publ. 6190-2.1, 1993; G. L. Batten, *Programmable Controllers: Hardware, Software, and Applications*, 2d ed., 1994; K. J. Hintz and D. Tabak, *Microcontrollers: Architecture, Implementation, and Programming*, 1992; C. T. Jones and L. A. Bryan, *Programmable Controllers*, 1983.

Programming languages

The different notations used to communicate algorithms to a computer. A computer executes a sequence of instructions (a program) in order to perform some task. In spite of much written about computers being electronic brains or having artificial intelligence, it is still necessary for humans to convey this sequence of instructions to the computer before the computer can perform the task. The set of instructions and the order in which they have to be performed is known as an algorithm. The result of expressing the algorithm in a programming language is called a program. The process of writing the algorithm using a programming language is called programming, and the person doing this is the programmer. *See* ALGORITHM.

In order for a computer to execute the instructions indicated by a program, the program needs to be stored in the primary memory of the computer. Each instruction of the program may occupy one or more memory locations. Instructions are stored as a sequence of binary numbers (sequences of zeros and ones), where each number may indicate the instruction to be executed (the operator) or the pieces of data (operands) on which the instruction is carried out. Instructions that the computer can understand directly are said to be written in machine language. Programmers who design computer algorithms have difficulty in expressing the individual instructions of the algorithm as a sequence of binary numbers. To alleviate this problem, people who develop algorithms may choose a programming language. Since the language used by the programmer and the language understood by the computer are different, another computer program called a compiler translates the program written in a programming language into an equivalent sequence of instructions that the

computer is able to understand and carry out. *See* COMPUTER STORAGE TECHNOLOGY.

Computers were first developed in the 1940s and have been used in industry since the early 1950s. As more is learned about the programming process and about what makes a good algorithm, the notation used by the programmer, the programming language, has evolved. The requirements of a good program today are vastly different from those in the 1950s.

Evolution of Programming Languages

Programming languages have progressed from machine languages, to assembly languages, to high-level languages.

Machine language. For the first machines in the 1940s, programmers had no choice but to write in the sequences of digits that the computer executed. For example, assume that a computer has 100,000 memory locations (numbered from 00,000 to 99,999) and has 10 registers (numbered from 0 to 9) for performing arithmetic calculations. The instructions for this machine are in an 8-digit format *WWXYYYYY*, where *WW* stands for an arithmetic operation, *X* signifies a register (called an operand register) to contain one argument for an operation, and *YYYYY* signifies the memory location. List (1)

10	Add contents of location YYYYY to X
11	Subtract contents of location YYYYY from X
12	Negate contents of register X. Ignore YYYYY
21	Move contents of location YYYYY to X
22	Move contents of location X to YYYYY
31	Next instruction to execute is at YYYYY if X is positive

(1)

defines a few simple instructions. Assume we want to compute the absolute value of $A + B - C$, where A is the value at machine address 3012, B is the value at address 3013, and C is the value at address 3014, and then store this value at address 3015. Sequence (2) of instructions, beginning at location 02345, would be:

Address	*Instruction*	*Interpretation*
02345	21103012	Get contents of 03012 (A) and put into register 1
02346	10103013	Add contents of 03013 (B) to register 1
02347	11103014	Subtract contents of 03014 (C) from register 1
02348	31102350	Execute at address 02350 if register 1 is positive; Otherwise go to next instruction
02349	12100000	Make negative contents of register 1 positive
02350	22103015	Move $\lvert A + B - C\rvert$ to location 03015

(2)

It should be clear that programming in this manner is difficult and fraught with errors. Explicit memory locations must be written, and it is not always obvious if simple errors are present. For example, at location 02347, writing 101... instead of 111... would compute $|A + B + C|$ rather than what was desired. This is not easy to detect.

Assembly language. Since each component of a program stands for an object that the programmer understands, using its name rather than numbers should make it easier to program. By naming all locations with easy-to-remember names, and by using symbolic names for machine instructions, some of the difficulties of machine programming can be eliminated. A relatively simple program called an assembler converts this symbolic notation into an equivalent machine language program. Example (2) could be programmed in assembly language as sequence (3):

Label	*Operation*	*Operand*	*Comments*
	Load	R1, A	Place operand A into register 1
	Add	R1, B	Compute A + B by adding B to register 1
	Subtract	R1, C	Compute A + B − C by subtracting C from register 1
	Ifplus	R1, SKIP	Goto location SKIP if register 1 is positive
	Negate	R1, 0	Compute $\lvert A + B - C\rvert$ by negating register 1
SKIP	Store	R1, D	Save answer in D by moving register 1 to location D
A	Define	03012	Define A as location 03012
B	Define	03013	Define B as location 03013
C	Define	03014	Define C as location 03014
D	Define	03015	Define D as location 03015

(3)

The symbolic nature of assembly language greatly eased the programmer's burden, but programs were still very hard to write. Mistakes were still common. Programmers were forced to think in terms of the computer's architecture rather than in the domain of the problem being solved.

High-level language. The first programming languages were developed in the late 1950s. The concept was that if we want to compute $|A + B - C|$, and store the result in a memory location called D, all we had to do was write $D = |A + B - C|$ and let a computer program, the compiler, convert that into the sequences of numbers that the computer could execute. FORTRAN (an acronym for Formula Translation) was the first major language in this period; in

it, the above example would be written as (4)

$$D = \text{abs}\ (A + B - C) \tag{4}$$

where *abs* represents a function that computes the absolute value of its argument. If a language does not have the function *abs* to invoke, we could have written the program as (5)

```
D = A + B − C;
IF D < 0 THEN
        D = − D;
WRITE (D);
```
(5)

which consists of three statements: an assignment statement, an IF-THEN statement to take care of the case where the initial calculation was negative and the absolute value requires a positive result, and a WRITE statement to write out the answer.

Although FORTRAN statements were patterned after mathematical notation, the last statement in (5), $D = -D$, shows the basic difference between mathematics and a program. In mathematics the = symbol implies that both sides of the equation have the same value. However, in FORTRAN and some other languages, the equal sign is known as the assignment operator. The action carried out by the computer when it encounters this operator is, "Make the variable named on the left of the equal sign have the same value as the expression on the right." Because of this, in some early languages the statement would have been written as $-D \rightarrow D$ to imply movement or change, but the use of $\rightarrow$ as an assignment operator has all but disappeared.

The compiler for FORTRAN converts that arithmetic statement into an equivalent machine language sequence, such as in (2). In this case, we did not care what addresses the compiler used for the instructions or data, as long as we could associate the names *A*, *B*, *C*, and *D* with the data values we were interested in.

At first there was some resistance to these higher-level languages (called higher-level because they were deemed more complex or higher than the assembly languages of that day). Computers were extremely expensive, and it was important to keep them busy doing useful work. Early compilers often generated inefficient sequences of instructions in compiling a program, and it was believed that even though programmers made mistakes in assembly code, it was still cheaper and more efficient to fix those errors in order to produce programs that were more efficient on these expensive machines.

We take a different view today of high-level languages. Computers have become orders-of-magnitude cheaper and faster; even if a program is not as efficient as assembly code, the extra running time is not important. Moreover, compiler technology has become so good that compiled programs often run as fast or faster than hand-written assembly code. Finally, we have learned, to our dismay, that programmers still make errors. Significant research has gone into developing programming languages that minimize the occurrence of those errors. We have learned that it is more important to get the algorithm right rather than to try to execute every machine cycle on a modern computer. Most new language design features today are concerned with their effects on the programming process rather than on their ultimate efficiency on the underlying machine hardware.

Early high-level languages. The period from 1957 through 1962 was the most significant in the development of programming languages. Almost all modern concepts first appeared in the four important languages developed in that period. Those languages and the major features they introduced are as follows.

FORTRAN. FORTRAN was developed by IBM under the leadership of John Backus in 1957. It became the model for most current languages, including BASIC, C, C++, and Java. A program executes a sequence of statements in which each statement looks like a mathematical expression. Example (4), where $D = abs(A + B - C)$, enables mathematicians to easily translate mathematical relationships into programming language algorithms.

In FORTRAN, the basic program unit is the subroutine or function. Each subroutine is a self-contained unit that computes some value. The goal is to design and develop each subroutine independently such that each can be used repeatedly at multiple places in a program. For example, if a function to compute the standard deviation of a list of numbers was needed, a single *StandDev* function could be written and used wherever the standard deviation was required.

ALGOL-60. The name of this language, developed in 1960 by an international committee led by Peter Naur and John Backus, is an acronym for ALGOrithmic Language, 1960. This international effort extended the ideas present in FORTRAN. A notation called BNF (Backus-Naur Form), which is a grammatical form called a context-free grammar, defined the language. Because of advances in compiler technology, defining languages using BNF grammars allows for efficient compilers to be developed almost automatically. ALGOL-60 further developed the concept of the subroutine (called a procedure in ALGOL-60) to permit the easy communication of data from the calling procedure to the called procedure.

ALGOL-60 was the forerunner of programming language control structures. It also included a primitive form of inheritance, although it was not until the early 1970s that the work of David Parnas on information hiding and Alan Kay on Smalltalk led to the development model we now call object-oriented design.

COBOL. The U.S. Defense Department under the leadership of Grace Hopper organized the development of COBOL in 1960. The name is an acronym for Common Business Oriented Language. This language has been a mainstay of the business community since its conception. COBOL introduced the concept of record to provide access to data for business applications. During the 1970s, Niklaus Wirth further refined this concept with the type declaration and the record type in the Pascal programming language.

LISP. About the same time that FORTRAN was being designed, John McCarthy at the Massachusetts Institute of Technology was investigating software for problem solving. He developed LISP (for LISt Processing). This form of programming, called applicative programming, is the basis for almost all artificial intelligence and expert system developments since then. *See* ARTIFICIAL INTELLIGENCE; EXPERT SYSTEMS.

Later high-level languages. Figure 1 presents an overview of the development of the major programming languages. Well over 500 additional languages have been identified, but the languages in Fig. 1 generally have made the most lasting contribution to language design. *See* COMPUTER PROGRAMMING; DIGITAL COMPUTER.

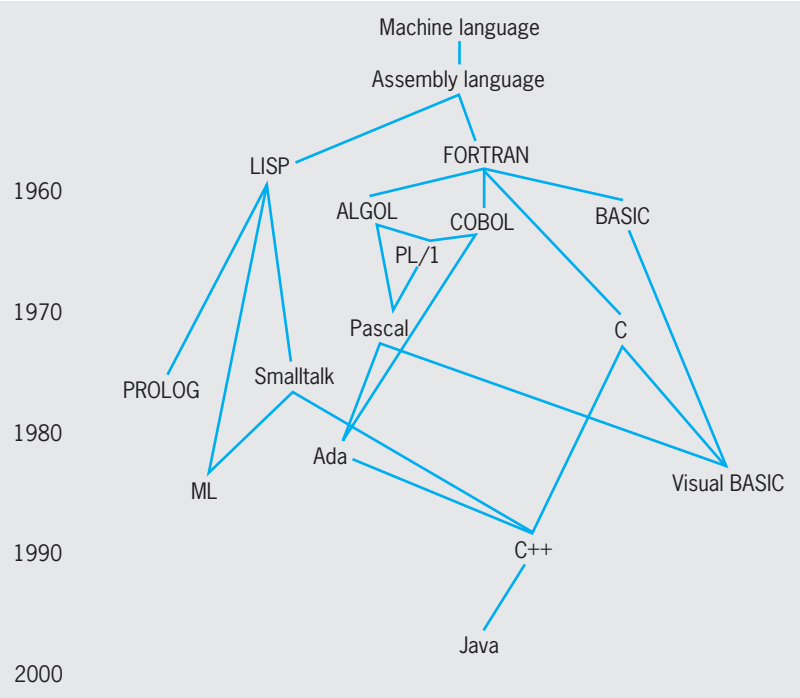

Fig. 1. Evolution of languages.

Structure of Programming Languages

Programs written in a programming language contain three basic components: (1) a mechanism for declaring data objects to contain the information used by the program; (2) data operations that provide for transforming one data object into another; (3) an execution sequence that determines how execution proceeds from start to finish.

Data declarations. Data objects can be constants or variables. A constant always has a specific value. Thus the constant 42 always has the integer value of forty-two and can never have another value. On the other hand, we can define variables with symbolic names. The declaration of variable *A* as an integer informs the compiler that *A* should be given a memory location much like the way the variable *A* in example (2) was given the machine address 03012. The program is given the option of changing the value stored at this memory location as the program executes.

Each data object is defined to be of a specific type. The type of a data object is the set of values the object may have. Types can generally be scalar or aggregate. An object declared to be a scalar object is not divisible into smaller components, and generally it represents the basic data types executable on the physical computer. In a data declaration, each data object is given a name and a type. The compiler will choose what machine location to assign for the declared name. For example, in C, declaration (6) does

$$Int\ A,\ Myvalue; \qquad (6)$$

the following: (1) *A* and *Myvalue* are both declared as scalar variables. (2) *A* and *Myvalue* are both declared with an integer data type. (3) The compiler allocates separate machine memory locations for each variable.

Scalar types. Programming languages usually include the following scalar types:

Integer. As in mathematics, integers represent the whole numbers, 0, 1, 2, 3, and so forth, and the negative numbers, −1, −2, −3, and so forth. However, since words in a computer memory are of a fixed size [for example, the 8-digit machine language from (1)], integers in a programming language will have a minimum and a maximum value.

Float. Float or *real* is the computer representation of rational numbers written in scientific notation. Thus $1.2E^2$ represents 1.2×10^2 or 120. As with integers, because of the fixed size of a machine's memory, not all rational numbers can be represented. Typically a real data object can have from 8 to 15 significant digits, and can represent exponents with values extending from approximately 10^{-300} to 10^{+300}. It must be emphasized that even though these data objects are called real variables, they are mathematically rational numbers. The expression (1.0/3.0) * 3.0 is not equal to 1 because 1.0/3.0 is equal to .333333. . . for as many places as the machine is capable of storing. Multiplying this value by 3.0 gives a result of only .99999. . . for approximately 8 to 15 digits and is not equal to 1.0.

Boolean. This scalar type represents the boolean values *true* and *false*. Operations on boolean data types follow the rules of boolean algebra. *See* BOOLEAN ALGEBRA.

Character. Each individual memory location in most modern computers represents a single character. Most computers today use the 8-bit byte as a memory location; thus there are 2^8 or 256 possible byte values. Character codes such as ASCII (American Standard Code for Information Interchange) or IBM's EBCDIC (Extended Binary Coded Decimal Interchange Code) define the externally printable character for each internal bit representation. For example, the ASCII bit pattern 01000001 (the number 65 in decimal) represents the character A; 01000010 (or the number 66) represents the character B. Defining data objects as characters allows for manipulating data that represent written text.

Enumerated. Languages today also include a programmer-defined scalar type called an

enumerated type. For example, to define the seasons of the year, declaration (7)

$$Enum\ \{spring,\ summer,\ fall,\ winter\}\ A; \quad (7)$$

may be used in C. Data object A can take on one of the values *spring*, *summer*, *fall*, or *winter* and no other value. This allows the programmer to define types relevant to the problem being solved, and allows the compiler to find some common errors, such as if the programmer then tried to write (8).

$$A = 7; \quad (8)$$

The value 7 would not be of the appropriate type.

References. A reference is a scalar data object that is the address of another object. These are also called pointer variables. If A is an integer data object and *Ptr1* and *Ptr2* are both reference variables whose values are the address of A, then the value of A may be changed by referencing either *Ptr1* or *Ptr2* in conjunction with a dereferencing operator as done in the C-like family of languages.

Aggregate types. Aggregate or composite data types are data types whose objects can be divided into smaller components, usually scalar objects or other aggregate types. The following are the usual aggregate types:

Arrays. An array is an ordered sequence of data objects. The array has a lower and an upper bound; thus we can refer to the first member of the array, the second member, and so forth. Each array member is itself a data object of some type. Example (9)

$$Int\ A[10]; \quad (9)$$

would define an array of 10 elements (from $A[0]$ through $A[9]$) in C, with each array element of the array being an integer. We can also defined multiple dimensions. A two-dimensional array called B with 10 rows and 10 columns can be defined as $B[10,20]$. This array consists of 10×20 different elements. Each element is uniquely identified by the value of its row and column. The first element of row 1 is $B[1,1]$, the second element of row 1 is $B[1,2]$.... The last element of the last two elements of the last row of the array are $B[10,19]$ and $B[10,20]$.

An array element can usually be of any valid type. Therefore we can define an array of integers, an array of characters, or an array of an array of reals.

Strings. A string is an ordered sequence of characters. In many programming languages a string is simply an array of characters. For example, in Pascal a string of length 20 could be defined as example (10).

$$Type\ string20 = array[1..20]\ of\ char; \quad (10)$$

Records. A record is a named collection of other data objects. Whereas in an array each of the objects is of the same type, in a record each object may be of a different type and is individually named. In Pascal a record that defines an address could be example (11).

```
Type address = record
        Name: string20;
        Street: string20;
        City: string20;                    (11)
        State: string20;
        PostalCode: integer
        end;
```

$$Var\ MyAddress: address; \quad (12)$$

Pascal declaration (12) declares *MyAddress* to be of type address, and allows the program to access the data objects *MyAddress.Name*, *MyAddress.Street*, *MyAddress.City*, *MyAddress.State*, and *MyAddress.PostalCode*.

Data operations. Data operations provide for setting the values into the locations allocated for each declared data variable. In general this is accomplished by a three-step process: a set of operators is defined for transforming the value of each data object, an expression is written for performing several such operations, and an assignment is made to change the value of some data object.

Operations. For each data type, languages define a set of operations on objects of that type. For the arithmetic types, there are the usual operations of addition, subtraction, multiplication, and division. Other operations may include exponentiation (raising to a power), as well as various simple functions such as modula or remainder (when dividing one integer by another). There may be other binary operations involving the internal format of the data, such as binary *and*, *or*, *exclusive or*, and *not* functions. Usually there are relational operations (for example, equal, not equal, greater than, less than) whose result is a boolean value of *true* or *false*. There is no limit to the number of operations allowed, except that the programming language designer has to decide between the simplicity and smallness of the language definition versus the ease of using the language.

Other data types include other relevant operations. Boolean operations would include the *and*, *or*, *exclusive or*, and *not* operations. For strings we usually have only the relational operations for comparing character data.

Expressions. The expression is the primary mechanism for transforming one data value into another. Most languages use an *infix* notation, similar to common mathematics, where the operation symbol is between the operands; however, there are three common notations for presenting expressions. $(A * B) - (C * D)$ is expressed in examples (13), (14), and (15).

Expression notation	*Example*	*Meaning*	
Infix	A ∗ B − C ∗ D	Operator between operands	(13)
Prefix	− ∗ A B ∗ C D	Operator precedes operands	(14)

Postfix A B ∗ C D ∗ − Operator follows operands (15)

Infix notation depends upon an implied ordering of operations. Thus (13) is assumed to mean $(A * B) - (C^* D)$ rather than an interpretation of $A * (B - C) * D$. This is based the concept of precedence. Precedence defines the order of evaluation. We normally give multiplication and division highest precedence, meaning do these operations first, and then do addition and subtraction. With additional operators, such as character operations, logical operations such as *and* and *or*, binary operations on bits, and so forth, the C language defines 17 levels of precedence.

Since $A * B - C * D$ means $(A * B) - (C * D)$, the expression $A * (B - C) * D$ requires the use of parentheses. The Polish mathematician Jan Lukasiewicz discovered in the 1920s that one could unambiguously write expressions without using parentheses, as is required with the usual infix notation. This notation, originally called Polish notation, now goes by the name of prefix notation. By placing the operator immediately before its operands, the order of evaluation of the operations becomes unambiguous. Rather than the expression in (13), the prefix for the expression $A * (B - C) * D$ is given by $* * A - BCD$. Both expressions have different and unambiguous prefix formats.

Polish notation remained a mathematical curiosity until the 1960s, when it was employed in the ALGOL compiler. By converting all expressions into postfix notation, a variation on Lukasiewicz's prefix notation with the operator following its operands, extremely fast compilers could be built. The evaluation of an expression in postfix uses algorithm (16).

If the next object is an operand, then save it.
If the next object is an operator, then apply operator to saved objects and save result (16)

For the expression in (15), $(A\ B * C\ D * -)$, its evaluation proceeds as in sequence (17).

Input	*Action*	*Saved values*
A	Save A	A
B	Save B	A, B
∗	Apply ∗ to A and B and save (A ∗ B)	(A*B)
C	Save C	(A ∗ B), C
D	Save D	(A ∗ B), C, D
∗	Apply ∗ to last 2 saved values, C and D, and save (C − D)	(A ∗ B), (C − D)
—	Apply − to last 2 saved values, and save result	(A ∗ B) − (C ∗ D)

(17)

The use of postfix notation radically changed the design of programming language compilers. It allowed compilers to generate machine language that was comparable or better than the best human-designed machine language, and it opened the door for future languages to be more concerned with improving the program design process than with just generating efficient machine language instructions.

Assignment statement. The assignment statement is usually the primary mechanism for an expression to have an impact on the data variables in a program. It is usually given by the syntax *variable* = *expression*, and means to evaluate *expression* and assign its value to the storage location defined by the symbolic name *variable*.

Execution sequence. The purpose of a program is to manipulate some data in order to produce an answer. While the data operations provide for this manipulation, there must be a mechanism for deciding which expressions to execute in order to generate the desired answer. That is, an algorithm must trace a path through a series of expressions in order to arrive at an answer. Programming languages have developed three forms of execution sequencing: (1) control structures for determining execution sequencing within a procedure; (2) interprocedural communication between procedures; and (3) inheritance, or the automatic passing of information between two procedures.

Control structures. Corrado Böhm and Giuseppi Jacopini showed in 1966 that a programming language needs only three basic statements: an assignment statement (discussed above), an IF statement, and a looping construct. Anything else can simplify programming a solution, but is not necessary. If we add an input and an output statement, we have all that we need for a programming language. Languages execute statements sequentially with the following variations to this rule.

IF statement. Most languages include both common variations, (18) and (19),

IF expression
THEN Statement_1; (18)

IF expression
THEN Statement_1
ELSE Statement_2; (19)

of the IF statement. In the IF-THEN statement (18), the expression is evaluated, and if the value is *true*, then Statement_1 is executed next. If the value is *false*, then the statement after the IF statement is the next one to execute. The IF-THEN-ELSE statement (19) is similar, except that specific true and false options are given to execute next. After executing either the THEN or ELSE part, the statement following the IF statement is the next one to execute.

Statement_1 and Statement_2 may actually be several statements. Most languages include a bracketing construct to group sequences of statements. Thus, many languages use BEGIN Statement_1; Statement_2;... END to indicate a single grouping of statements.

Looping constructs. The usual looping constructs are the WHILE statement and the REPEAT statement. While only one is necessary, languages usually have both.

The WHILE statement is given by syntax (20).

```
WHILE expression DO
  Statement;                    (20)
```

This has the interpretation of evaluating *expression*. If *true*, execute *Statement*, and then reexecute the WHILE. In order for this process to terminate, something in *Statement* has to eventually cause the expression to become false when it is next evaluated.

The REPEAT statement is given by syntax (21).

```
REPEAT Statement
  UNTIL expression;             (21)
```

It is similar to the WHILE statement except that *Statement* is executed first and then *expression* is next evaluated. The process repeats if the expression is false, and exits to the following statement if true. Although the continuation process between the WHILE and REPEAT statements is different (the WHILE continues if the expression is true and the REPEAT continues if the expression is false), the major difference between the two statements is that the WHILE statement may not execute *Statement* at all (that is, if the initial value for expression was *false*), whereas the REPEAT statement will execute *Statement* at least once.

Other control structures. Most languages include other control structures to ease the programming process. Among the most common are the following.

Controlled repetition. The WHILE and REPEAT statements both iterate for an undetermined number of times. If the number of times is already known, a FOR statement can be used. The FOR statement specifies the explicit number of times a statement is executed, such as FOR I := 1 to 20. The single statement following this construct will be repeated 20 times. The control variable I is incremented after each execution of the statement until it reaches the limit value of 20.

Case statement. The CASE or SWITCH statement is like a multiway IF statement. Control passes to one of many possible statements, depending upon the value of the initial expression.

Go to statement. The GOTO had the simple syntax of GOTO *label*, where *label* is the symbolic name attached to a statement. It provides for a specific branch point to be specified after testing for an expression value. Example (5) using GOTO statements could be written example (22).

```
            D = A + B − C;
            IF D < 0 THEN GOTO NEXTSTMT;
            D = −D;
NEXTSTMT:   WRITE (D);
                                (22)
```

This construct has been found to be very error-prone and totally unnecessary to the writing of good programs. Until the work of Böhm and Jacopini, it was assumed that the GOTO statement was a requirement for programming languages. They showed that a GOTO was unnecessary, since any program function could be written using only the IF and the WHILE constructs.

Various software products are sold with the stated purpose of "structuring bad programs into good," with the Böhm and Jacopini theorem as the justification. But this is using their result backwards. A poorly designed, unstructured program will just become a poorly designed "structured" one. What the Böhm and Jacopini theorem does say is that you will be able to solve your problem using only IF and WHILE. It is an existence proof that a solution does exist. It is up to the programmer to find a good solution.

Interprocedural communication. The various control structures provide a mechanism for executing within a given procedure. However, the essence of good program design is to build a large system consisting of multiple procedures and functions, design each one independently, and then have them communicate necessary information among themselves. The procedure call is the primary communication mechanism, but there are others as well.

Procedure call. The procedure call is typically given by the syntax *FunctionName*(*arguments*), such as the example *abs*($A + B - C$) in (23).

In this case, the expression $A + B - C$, called an argument, is evaluated, and that value is passed to the function *abs*. The function *abs* uses that value to compute an answer, which is passed back to the procedure that invoked the procedure call.

In designing the called procedure, the passed argument is referred to as an actual parameter and given some symbolic name, as in example (23). Within the

Calling procedure	*Argument*		
abs(A + B − C)	A + B − C	↔	*Called procedure*
	Parameter		
	X		abs(X)

(23)

procedure, X behaves as if it is a new data declaration.

The relationship between the formal parameter [X in (23)] and its actual argument [$A + B - C$ in (23)] can differ. Several mechanisms for communication between the argument and the parameter have been developed. The following four are the most common:

1. In call by reference, the parameter is actually the address of the argument. If one writes *abs*(A) and the formal parameter in *abs* is named X, then changes to the value of X are actually reflected as references to A. X is an alias to the variable A, and both names refer to the same actual data object. This method is the most common and efficient form of argument transmission in programming languages.

2. In call by value, the parameter is actually the value of the argument. If one writes *abs*(A) and the formal parameter in *abs* is named X, then upon entry to the procedure, X is set to the value of A. Changes to the value of X do not affect the value of A. The programming language creates a new data object for the parameter. This method is the most efficient since all references to the parameter are data variables local to the called procedure. It is limited, however, in

that the purpose of a called procedure is to compute several values and return them to the calling procedure. Call by value does not permit those values to be returned.

3. Call by result is similar to call by value except that upon returning to the calling procedure, the final value of the parameter is passed back as the value of the argument. This provides the flexibility of returning values to arguments, as is the case with call by reference, with some of the efficiency of the call-by-value implementation. It is a useful model, but one that is not implemented often.

4. Call by name was first described in the report defining ALGOL-60. It has the simplest theoretical definition. In example (23), the function *abs* is simply a replacement for the definition of the actual absolute value function. In practical terms, it means that if X is a call-by-name parameter for the argument $A + B - C$, then each time X is referenced in *abs*, the actual argument $A + B - C$ is reevaluated. This can be extremely inefficient, since in most programs the reevaluation of $A + B - C$ will not change each time.

More seriously, call by name can lead to anomalous situations. Consider example (24)

```
Calling procedure     Called procedure
A = 1;                PROCEDURE F(NAME X,
                        REFERENCE Y);
B = 2;                INTEGER Z;                (24)
CALL F (A + B, B);    Z = X;
                      Y = 1;
                      IF Z = X THEN WRITE
```

where parameter X is a call-by-name parameter and parameter Y is a call-by-reference parameter. Procedure F is passed the arguments $A + B$ and B. In F the assignment to Z evaluates to 3 (1 + 2). The following assignment to Y changes the value of B to 1 since it is a call-by-reference parameter. The IF statement evaluates to false, since X now is 2 (1 + 1) and Z is the old value of X or 3. Because of these issues, call by name is rarely used as a parameter mechanism today.

The above four classical parameter mechanisms are defined by their implementation. In each case, the definition describes how the variables will be implemented. The Ada language uses a difference approach, involving *in-out* parameters. It describes the behavior of the parameter. A value, which is passed into a procedure and used within the procedure, will be an *in* parameter. Upon exit from the procedure, the value is not passed back. This has the same behavior as the call-by-value parameter and is often implemented that way. However, the language definition does not specify the implementation, only the behavior.

A parameter that has a value passed back to the calling argument is an *out* parameter, and one that is passed in and then the final value passed back to the calling argument is an *in-out* parameter. *In-out* parameters are often implemented as call by result, or even call by reference.

Functional operations. For statement-oriented languages such as C, ALGOL-60, or Java, the discussion parameters above describes the transmission of information into procedures. But languages such as LISP, and more recently ML, have a totally different execution model. These languages are based upon a functional model of execution that does not explicitly declare data objects. One creates the function that is the result of the computation that one wants.

Example (22) is the standard statement-oriented version of the simple *abs* example given in (5). We can write these solutions functionally in LISP or ML as in example (25).

```
Language    Expression
  LISP      (write (abs (−(+ A B) C)))      (25)
  ML        write (abs (A + B − C))
```

LISP uses a prefix notation as given in (14) and ML uses the more common infix notation (13). What is important to notice, however, is that neither language uses intermediate data objects. The argument $A + B - C$ is evaluated and passed to the *abs* function. The value returned from the *abs* function is then passed to the *write* function. In no case does this value need to be stored as an intermediate value with its own name, as is the case with other variable transmission mechanisms.

The actual code for computing example (25) [using the algorithm in (5)] could be written in ML as example (26).

```
write (LetD = A + B − C
        in
      if D < 0 then D                    (26)
               else −D);
```

This states that *write* will write the value of its argument, which is the value of the IF expression. The expression $A + B - C$ is assigned the name D and, in the IF statement, either D or its negation is returned. In no case is there an explicit variable storage for any data object.

Messages. In 1974 the Smalltalk language introduced a new model for describing information between procedures. This is the concept of the message. The mental image being portrayed is of a stationary object being passed messages to procedures called methods. The object responds to the message and waits for the next message. Rather than a program being a collection of procedures communicating with one another, a program is now a collection of objects, each defining its own environment, and each responding to queries about its environment from other objects. By adding the concept of inheritance, we have the principal idea behind object-oriented languages.

In Smalltalk we would have integer objects. Example (25) would be implemented with the *print* method and *abs* method for integers. Operations in Smalltalk follow their operands, and execution proceeds from left to right with no operation hierarchy as is the case with FORTRAN or C. Since the value $A + B - C$ is an instance of an integer object, the program equivalent for (25) would be (27).

```
(A + B − C) abs print                      (27)
```

Execution proceeds as in sequence (28).

Smalltalk expression	*Execution*
A	A is an integer object
+B	The + (addition) method for integer A is passed the value B. The integer object (A + B) is returned by the addition method of A.
−C	The − (subtraction) method for integer object (A + B) is passed the value C. The integer object (A + B − C) is returned.
abs	The abs method for integer object (A + B − C) is called and it returns the integer value object that is the value of abs.
print	The print function for the integer value result of the abs function is called. The value of the expression is printed.

(28)

Although we still have functions, procedures, and data variables, by describing programming objects with this model we get a design process that looks very much like the problem we are trying to solve. This model, built in Smalltalk in 1974, is the one major feature that was not present in the early days of programming language design.

Inheritance. Inheritance is the third major form of execution sequencing. In this case, information is passed automatically between program segments. The concept of inheritance began with ALGOL-60 and its static scope rules, and was extended with the information hiding concepts of Parnas and implemented in Smalltalk. This became the basis for the current models used in today's object-oriented languages C++ and Java.

Scope. The scope of a variable declaration is the set of procedures where it can be used. In general, a variable is declared; the variable name can be used throughout that procedure. However, in some languages (ALGOL-60, C, Pascal), a segment can be declared within another segment. For example, in C we can declare a block using the symbols { and }. Consider example (29).

C program	*Comment*
{ int I, J;	I and J declared integer variables
I = J+1;	
{ int J, K;	A new J and K declared integer variables. I is also available in this block.
}	
}	

(29)

In this case, the scope for *I* is the entire example, both the outer block and the inner block. The scope for the *J* declared on the first line is only the outer block since there is a declaration for a new variable *J* in the inner block. The use of scope was one of the first examples of passing information between program segments.

Encapsulation. In 1972, Parnas introduced the concept of encapsulation. The idea was that good program design would consist of a data object definition, and only those functions that needed to know the data object definition would have access to it. For example, the Pascal *MyAddress* record was defined in example (11). Only functions using the type *MyAddress* could have access to the actual structure of the record. A module for an address list could have the functions *CreateAddress*, *PrintAddress*, *RetrieveSpecificAddress*, or any other function required to manage an address list. All other parts of the program would access an address list by calling one of these specific address list functions.

Encapsulation was a clear description of how to divide a program into manageable parts for effective construction. The problem was that there was no current language that effectively implemented that structure. Simula, a variation of ALGOL-60 used for simulation problems, did contain the concept of a data class for declaring data objects. The development of Smalltalk solved the encapsulation problem.

Smalltalk developed the concept of a class object. There were integer class objects, string class objects, file class objects, and so forth. Data objects would then be instances of these class objects. Therefore, we could have an instance of an integer object, an instance of a string, an instance of a file. However, the real power in Smalltalk was that objects would inherit the properties of the objects from which they were created. Thus, if an integer object were designed with the methods (that is, functions) of addition and subtraction, each instance of an integer object would inherit those same functions. One would only need to develop these operations once and then the functionality would pass on to the derived object.

The real power of Smalltalk derives from the fact that all objects are derived from one master object called an Object. An Object is the parent class of objects such as magnitude, collection, and stream. Magnitude now is the parent of objects that have values, such as numbers, characters, and dates. Collections can be ordered collections such as an array or an unordered collection such as a set. Streams are the parent objects of files. From this structure an entire class hierarchy can be developed.

If we develop a method for one object (for example, *print* method for *object*), then this method gets inherited to all objects derived from that object. Therefore, there is not the necessity to always define new functionality. If we create a new class of integer that, for example, represents the number of days in a year (from 1 to 366), then this new integerlike object will inherit all of the properties of integers, including the methods to add, subtract, and print values. It is this concept that has been built into C++, Java, and current object-oriented languages.

Given a function call (or message in Smalltalk), because of inheritance, we often cannot tell where the method is actually defined. This allows us to develop polymorphic functions. A polymorphic function is one which each different set of arguments can call

an entirely different function to execute. Thus the Smalltalk message *print* in each of examples (30)

Smalltalk message	*Interpretation*	*Value printed*
2 print	Integer variable print method	2
'ABC' print	String variable print method	'ABC'
Object print	Object class print method	An Object
Integer print	Integer class print method. This is the same method inherited from Object print	An Integer

(30)

calls an entirely different method to execute.

Once we build concepts around a class definition, we have a separate package of functions that are self-contained. We are able to sell that package as a new functionality that users may be willing to pay for rather than develop themselves. This leads to an economic model where companies can build add-ons for existing software, each add-on consisting of a set of class definitions that becomes inherited by the parent class. *See* OBJECT-ORIENTED PROGRAMMING.

Computing Environments

Since the 1950s, the use of computers has undergone several radical shifts. In response to changing computer hardware and computing environments, programming language development has tried to keep pace with the needs of those environments.

Mainframe era. The period from the 1940s to 1970 can be regarded as the mainframe era. Computers were large, expensive machines used to solve corporate problems. Data entry for these machines utilized punched cards (the 80-column Hollerith card). Languages were needed so that programmers could submit their program one day and get their results the next. This was the era of FORTRAN, ALGOL-60, COBOL, and similar languages.

Personal computer era. The 1970s saw the introduction of terminals to mainframe computers. There was still a single large mainframe, but users could access computers from their desks. Programming started to become more interactive, and quick turnaround from 1 day to several minutes became feasible. Languages that encouraged quick turnaround, such as BASIC, became popular.

The situation was to change in 1976 with the introduction of the personal computer (PC). Users no longer needed to be connected to a mainframe, and standalone computing on an inexpensive machine was now possible. Interactive languages, which provided immediate feedback to the user, such as BASIC, became the standard on such machines.

The personal computer entered the mainstream of business computing in 1981 with the introduction of the IBM PC. The role of the mainframe computer became that of a corporate data server, and computing moved to the desktop, especially as desktop computers became faster and cheaper.

During the 1970s, many of the modern principles of the graphical user interface (GUI) [for example, windows and the use of a mouse for point and click data entry] were developed at the Xerox Palo Alto Research Center. These features became widely available with the introduction of the Apple Macintosh (1984) and subsequently with Microsoft Windows for IBM-compatible PCs. *See* HUMAN-COMPUTER INTERACTION.

With the graphical user interface, building window applications became important, and object-oriented program design became an economic necessity. Class libraries could be sold, allowing users to easily build applications that include windows and features such as menus, scroll bars, and buttons. Visual Basic evolved with predefined methods for including these window-class objects.

Internet era. Beginning in 1970, the U.S. Advanced Research Projects Agency (ARPA) began research in linking machines together. Initially a project to allow machines to continue to communicate and survive in the event of an attack on the United States, the initial Defense Department ARPANET evolved into the worldwide Internet by the mid-1980s. Today the Internet has millions of computers connected to it, mostly PCs. *See* INTERNET.

The Internet was a researcher's domain until the late 1980s. CERN, the European nuclear research facility in Geneva, Switzerland, was developing a method whereby researchers could share technical documents. In 1989, Tim Berners-Lee at CERN came up with the concept of the HTML (hypertext markup language) hyperlink as a way to navigate around the Internet and retrieve documents in a system known as the World Wide Web. With the development of the MOSAIC web browser in 1993, the general populace discovered the Internet, and by the end of the 1990s millions of users were connecting to it each day to search for information. *See* WORLD WIDE WEB.

Use of the World Wide Web led yet again to a change in the role for programming languages. Early web pages were static: A report could be displayed, a picture shown, or control could navigate to another web site. However, as an industry started to develop around the Web (e-commerce), a need developed for active web pages. Web developers wanted the ability to obtain data from users (personal information, ordering information) and then process that information in some meaningful way. This led to the need for web languages, such as Java.

Programming Paradigms

Today, a great deal is understood about the programming process. There are many languages that satisfy different needs. The following describes the basic programming models in use today and the various languages that use those programming models.

Imperative programs. An imperative program is a sequence of commands or statements telling the computer how to execute a particular algorithm.

This was the model for FORTRAN, ALGOL, and COBOL in the 1960s, and is still the major model in use for languages such as C, C++, BASIC, and Java. Execution of a program proceeds from statement to statement, each statement accessing and modifying specific data objects in a machine's memory. We can think of an imperative program as developing the steps that transform a set of data items from initial values into their final values; each statement performs one step in this complex series of transformations. It reflects the basic architecture of a computer and seems most natural for describing the programming process. However, it is not always the best model for solving a particular problem.

Applicative programs. Whereas imperative programming is concerned with the series of steps needed to transform data from some initial set of values into their final set of values, in an applicative program one constructs the function that transforms all the data from initial value to final value. Because of its close association with functions, it is also called functional programming. While this sounds similar to the imperative model, it leads to a very different model of programming. The goal is to transform all the data, function by function, until the final answer is obtained.

Examples (31) and (32) show the difference between the two models. Here, given a sequence of numbers, create a sequence of the same numbers in the reverse order.

An imperative solution, such as written in FORTRAN or C, would proceed as in (31).

1. Set *Last* to be end of the first sequence
2. Set *New* to be 1 (Location of the new sequence)
3. Move the object at *Last* to the object at *New* (Move one member of sequence)
4. Subtract 1 from *Last* and add 1 to *New*
5. Repeat steps 3 and 4 for all members of the sequence

(31)

As shown in **Fig. 2***a*, the imperative solution proceeds by moving each item in the sequence, one at a time, from last to first.

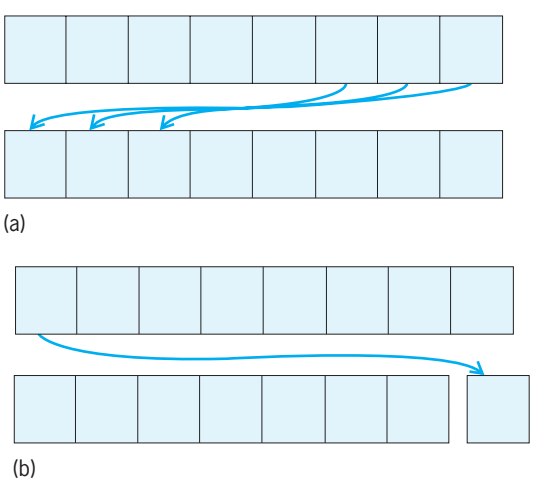

Fig. 2. Diagrams of (*a*) imperative solution and (*b*) applicative solution.

The applicative solution is given in Fig. 2*b* and is specified as (32).

1. Remove the first member from the list.
2. Reverse the remaining members of the list.
3. Add the first member to the reverse remainder of the list.

(32)

In the applicative case, we are unconcerned with specific data items; we consider the switching operation on the entire sequence at one point. In this case, the applicative solution applies to the entire sequence. The first member of the sequence is moved to the end and then the process is repeated for the remaining members, with the second members now being the beginning of the list.

This is an example of a recursive solution, since the solution, in step 2, required using the solution we were still developing. In general, recursive solutions are shorter but harder to write than imperative solutions. However, once written, they can be shown to be correct more easily than the other method.

Applicative programming is the basis for LISP. Problems requiring searching large databases often can be expressed this way, and applicative programs have been useful in developing artificial intelligence applications. Each function in an applicative solution to a problem represents some fact about the solution. The structure of a functional program then becomes (33).

$$\begin{aligned} Fact1 &= statements\ describing\ fact\ 1 \\ Fact2 &= statements\ describing\ fact\ 2 \\ Fact3 &= statements\ describing\ fact\ 3 \end{aligned} \tag{33}$$

The solution to the problem then becomes (34).

$$Solution = fact1\ and\ fact2\ and\ fact3\ and\ fact4 \tag{34}$$

Executing the program is simply the statement: *answer* = *solution*(*input data*).

Declarative programs. In both applicative and imperative programs, the programmer presents the algorithm for computing a solution. In the imperative case, the program is a step-by-step sequence of statements of computing the answer from the initial set of data items. In the applicative case, the program is a sequence of functions that are executed to transform the input data into the answer.

Prolog represents a third method. In this case, the structure of the solution is given, and the programming system determines how to find a solution. The above example could be modified to state: Given a sequence of numbers, create a sequence of the same numbers in ascending order.

Then, we could represent a solution according to Prolog-like algorithm (35).

1. Let I represent the Ith member of the sequence.
2. A (I,I + 1) :- greaterthan (I, I + 1), switch (I, I + 1).
3. A (I, I + 1) :- write (A).

(35)

Although not true Prolog, this represents how Prolog

searches for a solution. In this case, line 2 of (35), we have a relation A. If I and $I + 1$ have the property that the *greaterthan* function is true (that is, member at I is greater than the member at $I + 1$), then the *switch* function is called (I and $I + 1$ reverse order). However, if *greaterthan* always fails, then no variable is out of order, and statement 3 is next executed. The only way that 2 fails is for A to now be in order, so the *write* function prints out the value of the sorted array.

In this case we simply specified constraints on the output: For any member of the output sequence, the following member must have a greater value. We did not specify how the computer would provide for this condition, only that when the program executed the *write* function in line 3 of (34), the array A would be sorted.

Applications requiring searching large databases are amenable to this technique. It has also been found to be useful is solving logic problems; thus the method is also called logic programming. Much of the work in a compiler that automatically translates a BNF description of a programming language into some internally useful form is another example use of this declarative model. A compiler writer simply gives the BNF description of a language, and the compiling system automatically generates a program that will recognize programs written in that language.

Current Programming Language Models

Of the languages given in Fig. 1, and of the hundreds not presented here, a few stand out as currently significant notations for which to solve computer programs.

C. C was developed by AT&T Bell Laboratories during the early 1970s. At the time, Ken Thompson was developing the UNIX operating system. Rather than using machine or assembly language as in (2) or (3) to write the system, he wanted a high-level language. Starting with a language named BCPL, he created a small subset named B. By adding data types and additional operations, he renamed B as C. *See* OPERATING SYSTEM.

C has a structure like FORTRAN. A C program consists of several procedures, each consisting of several statements, that include the IF, WHILE, and FOR statements. However, since the goal was to develop operating systems, a primary focus of C was to include operations that allow the programmer access to the underlying hardware of the computer. C includes a large number of operators to manipulate machine language data in the computer, and includes a strong dependence on reference variables so that C programs are able to manipulate the addressing hardware of the machine.

The relationship between C arrays and string variables illustrates the close relationship of C programs to the underlying hardware. Declaration (36)

$$\textit{char } A[10]; \tag{36}$$

creates an array of 10 characters, $A[0]$ through $A[9]$. However, the C compiler actually creates a reference variable for A pointing to the first element of the array. Therefore, A is just a pointer to $A[0]$. The value at this location (called dereferencing A) can be accessed by the array reference $A[0]$ or by the pointer dereference operation (that is, *) as in $^*A$. Accessing the Ith element of the array can be denoted as either $A[I]$ or as $^*(A + I)$. This duality between C types, such as strings and the underlying integer representation for those types, makes C both an extremely efficient language in which to write systems programs and a very difficult language in which to write correct programs.

C++. C++ was developed in the early 1980s as an extension to C by Bjarne Stroustrup at AT&T Bell Labs. The idea was to extend it with Smalltalk-like classes. Each C++ class would include a record declaration as well as a set of associated functions. In addition, an inheritance mechanism similar to Smalltalk was included in order to provide for a class hierarchy for any program. The Pascal record declaration from (11) could be extended with class functions via declaration (37).

```
Class   address {
        String20 Name;
        String20 Street;
        String20 City;
        String20 State;                          (37)
        int PostalCode;
        void CreateAddress( );
        void PrintAddress( );
        void RetrieveSpecificAddress( ); }
```

In order to access *address* records, one of the functions *CreateAddress*, *PrintAddress*, or *RetrieveSpecficAddress* would have to be called. This provides for the encapsulation so important for good program design.

C++ also included inheritance. A new class, *MyAddress*, could be derived from *address* via declaration (38).

$$\textit{Class MyAddress}: \textit{address}\{\ldots \textit{additional functions}\} \tag{38}$$

Here *MyAddress* inherits all of the properties (that is, record description and class functions) of *address* as well as creation of any new functions specific for that class. C++, then, includes the efficiency of a C program with the encapsulation, inheritance, and flexibility of object-oriented programming in Smalltalk.

Java. By the early 1990s, the World Wide Web was becoming a significant force in the computing community, and web browsers were becoming ubiquitous. However, for security reasons, the browser was designed with the limitation that it could not affect the disk storage of the machine it was running on. All computations that a web page performed were carried out on the web server accessed by web address (its Uniform Resource Locator, or URL). That was to prevent web pages from installing viruses on user machines or inadvertently

(or intentionally) destroying the disk storage of the user.

The effect of this was that all information had to be transferred from the user to the web server. James Gosling at Sun Microsystem led development of a product called Oak, which evolved into Java in 1995, that could execute on the user's machine. Thus, instead of transferring all of the user's data to the server machine, the server would transfer the application program to the user.

In order to make this efficient, the transferred program had to be small, and it had to be capable of running on every browser on every computer connected to the Internet. Sun decided to transfer a version of the source program (the bytecode), rather than the usually longer machine language program. A program built into the browser would execute the bytecode as if it were a machine language program for that particular machine. Such a program is typically called an interpreter.

The constraint on this bytecode language was that it could not affect the disk storage of the user's machine. Initially, C++ was the programming language of choice due to its class hierarchy and efficient execution model. But because of pointer references and binary operations on machine addresses in the underlying C base language, it would be impossible to assure adherence to the web security model. C++ had to be dropped as the base language, and it has been replaced in this context by Java. Java bears a strong similarity to C++, but has eliminated many of the problems of C++. The three major features addressed by Java are:

1. There are no reference variables, thus no way to explicitly reference specific memory locations. Storage is still allocated by creating new class objects, but this is implicit in the language, not explicit.

2. There is no procedure call statement; however, one can invoke a procedure using the member of class operation. A call to *CreateAddress* for class *address* would be encoded as *address.CreateAddress*().

3. A large class library exists for creating web-based objects.

The Java bytecodes (called applets) are transmitted from the web server to the client web site and then execute. This saves transmission time as the executing applet is on the user's machine once it is downloaded, and it frees machine time on the server so it can process more web "hits" effectively. *See* CLIENT-SERVER SYSTEM.

Visual Basic. Visual Basic, first released in 1991, grew out of Microsoft's GW Basic product of the 1980s. The language was organized around a series of events. Each time an event happened (for example, mouse click, pulling down a menu), the program would respond with a procedure associated with that event. Execution happens in an asynchronous manner.

Each displayed screen is called a form. Forms are the visual components, and they are event controllers. Controls (for example, push buttons, radio buttons, check boxes, labels, text boxes, and scroll bars) are added to the form. Each control creates events, the set of actions that occur as a result of an action in the user interface, for example, pressing a key on the keyboard or clicking a mouse button.

Visual Basic provides a different program design model. Rather than looking at a program as a series of statements that form an algorithm, in Visual Basic the approach is to determine what actions occur in response to a series of events.

Prolog. Although Prolog development began in 1970, its use did not spread until the 1980s. Prolog represents a very different model of program execution, and depends on the resolution principle and satisfaction of Horn clauses of Robert A. Kowalski at the University of Edinburgh. That is, a Prolog statement is of the form p:- $q.r$, which means p is true if both q is true and r is true [as in example (35)].

A Prolog program consists of a series of Horn clauses, each being a sequence of relations concerning data in a database. Execution proceeds sequentially through these clauses. Each relation can invoke another Horn clause to be satisfied. Evaluation of a relation is similar to returning a procedure value in imperative languages such as C or C++.

As an example, assume there is a relation *parent_of*. *Parent_of* (X,Y) means that X is the parent of Y. Assume that there is a database of relationships (39).

parent_of (John,Mary)	parent_of (Susan,George)	
parent_of (Susan,Mary)	parent_of (Henry,John)	(39)
parent_of (John,George)	parent_of (George,Alice)	

In Prolog, one queries the database. If all the facts are specified, then Prolog responds either *true* (40) or *false* (41), depending upon whether the data are in the database. However, some facts are omitted, Prolog will search the database to see if a match can be found (42).

Prolog query	*Prolog response*	
parent_of (John,George)	true	(40)
parent_of (Alice,Mary)	false	(41)
parent_of (X,John)	X = Henry	(42)

Prolog programs are built by creating new relationships. A *grandparent_of* (X,Y) relationship can be defined by saying there is some person Z who is the parent of Y, and X is the parent of Z. This is specified in Prolog as (43).

grandparent_of (X,Y) :- parent_of (X,Z), parent_of (Z,Y) (43)

Prolog executes this by first trying to satisfy the first *parent_of* relationship. That defines a value for X and Y. It then sees if that association meets the second relation. If so, a true match is found. As with the single relation, if any operand is missing, Prolog tries

to find a match, as in examples (44), (45), and (46).

Prolog query	*Prolog response*	
grandparent_of (John,George)	false	(44)
grandparent_of (John,Alice)	true	(45)
grandparent_of (X,Alice)	X = John	(46)

Unlike the other languages mentioned, Prolog is not a complete language. That means there are algorithms that cannot be programmed in Prolog. However, for problems that are amenable for searching large databases, Prolog is an efficient mechanism for describing those algorithms. *See* SOFTWARE; SOFTWARE ENGINEERING. Marvin V. Zelkowitz

Bibliography. T. Pratt and M. Zelkowitz, *Programming Languages: Design and Implementation*, Prentice Hall, 4th ed., 2001; R. Sebesta, *Concepts of Programming Languages*, Benjamin Cummings, 5th ed., 2001; R. Sethi, *Programming Languages: Concepts and Constructs*, Addison Wesley, 2d ed., 1996.

Progression (mathematics)

Ordered, countable sets of numbers, $x_1, x_2, x_3, \ldots$, not necessarily all different. In general such sets are called sequences, whereas the term progression is usually confined to the special types: the arithmetic, in which the difference $x_k - x_{k-1}$ between successive terms is constant; the geometric, in which the ratio x_k/x_{k-1} is constant; and the harmonic, in which the reciprocals of the terms are in arithmetic progression.

Arithmetic progressions. If the first term is a and the common difference b,

$$x_1 = a, x_2 = a + b, x_3 = a + 2b, \ldots,$$

$$x_n = a + (n+1)\,b, \ldots \quad (1)$$

In the sum of n terms S_n, two terms equidistant from the ends always have the same sum $x_1 + x_n$; hence $2S_n = n(x_1 + x_n)$ and

$$S_n = n\frac{x_1 + x_n}{2} = n\left(a + \frac{n-1}{2}b\right)$$

If x_1, x_2, x_3 are in arithmetic progression, $x_2 = (x_1 + x_3)/2$ is called the arithmetic mean of x_1 and x_3.

Geometric progressions. If the first term is a and the common ratio r,

$$x_1 = a, x_2 = ar, x_3 = ar^2, \ldots,$$

$$x_n = ar^{n-1}, \ldots \quad (2)$$

Excluding the case $r = 1$ (when all terms are the same), the sum S_n of n terms satisfies $S_n - rS_n = a - ar^n$; hence,

$$S_n = a\frac{1 - r^n}{1 - r}$$

If $|r| < 1$, $r^n \to 0$ as $n \to \infty$; hence the sum of the infinite geometric series

$$\sum_{n=1}^{\infty} ar^{n-1} = \frac{a}{1-r} \qquad |r| < 1$$

If x_1, x_2, x_3 are in geometric progression, $x_2 = \sqrt{x_1x_3}$ is called the geometric mean of x_1 and x_3. Since

$$(\sqrt{x_1} - \sqrt{x_3})^2 = x_1 + x_3 - 2\sqrt{x_1x_3} \geq 0$$

and

$$\frac{x_1 + x_3}{2} \geq \sqrt{x_1x_3}$$

the arithmetic mean of two unequal positive numbers exceeds their geometric mean.

The arithmetic mean A and the geometric mean G of n positive numbers are defined as

$$A = \frac{x_1 + x_2 + \cdots + x_n}{n} \text{ and } G = \sqrt[n]{x_1x_2\cdots x_n}$$

Also, $A \geq G$.

Harmonic progression. The reciprocals of sequence (1) form a harmonic progression. There is no compact expression for the sum of n terms. If x_1, x_2, x_3 are in harmonic progression,

$$x_2 = \frac{2x_1x_3}{x_1 + x_3}$$

is called their harmonic mean.

Sum sequence. A general method of summing a sequence of n terms depends upon a theorem in the difference calculus which is the analog of the fundamental theorem of the differential calculus.

If $x_n = f(n)$ is defined for $n = 0, 1, 2, \ldots$, the difference $f(n)$ is defined as

$$\Delta f(n) = f(n+1) - f(n)$$

Δ is a linear operator; that is,

$$\Delta[af(n) + bg(n)] = a\Delta f(n) + b\Delta g(n)$$

In the difference calculus the factorial powers $n^{(k)}$ for integral k are defined by

$$n^{(k)} = \begin{cases} n!/(n-k)! & k < n \\ n! & k = n \\ 0 & k > n \end{cases}$$

Thus

$$n^{(k)} = n(n-1)(n-2)\cdots(n-k+1)$$

$$0 < k \leq n$$

$$n^{(-k)} = \frac{1}{(n+1)(n+2)\cdots(n+k)} \qquad k > 0$$

In particular

$$n^{(0)} = 1, n^{(1)} = n, n^{(-1)} = \frac{1}{n+1}$$

TABLE 1. Differences

$f(n)$	$\Delta f(n)$
constant	0
r^n	$(r-1)r^n$
$n^{(k)}$	$kn^{(k-1)}$
$\cos n\alpha$	$-2\sin \frac{1}{2}\alpha \sin (n+\frac{1}{2})\alpha$
$\sin n\alpha$	$2\sin \frac{1}{2}\alpha \cos (n+\frac{1}{2})\alpha$

TABLE 2. Antidifferences

$f(n)$	$\Delta^{-1}f(n)$	
0	constant	
r^n	$\dfrac{r^n}{r-1}$	$(r \neq 1)$
$n^{(k)}$	$\dfrac{n^{(k+1)}}{k+1}$	$(k \neq -1)$
1	n	$(k = 0)$
$\cos n\alpha$	$\dfrac{\sin (n-\frac{1}{2})\alpha}{2\sin \alpha/2}$	
$\sin n\alpha$	$\dfrac{\cos (n-\frac{1}{2})\alpha}{2\sin \alpha/2}$	

Moreover, $n^{(k)}$ satisfies the functional equation

$$n^{(k)} \cdot (n-k)^{(b)} = n^{(b)} \cdot (n-b)^{(k)} = n^{(b+k)}$$

This leads to the general definition of the factorial power

$$x^{(k)} = \frac{\Gamma(x+1)}{\Gamma(x+1-k)}$$

for all real values of x and k for which the gamma functions exist. When both of the gammas are infinite (for arguments $0, -1, -2, \ldots$), define $x^{(k)} = \lim (x+\epsilon)^{(k)}$ as $\epsilon \to 0$; thus

$$\begin{aligned}(-1)^{(2)} &= \lim(\varepsilon-1)^{(2)} = \lim \Gamma(\varepsilon)/\Gamma(\varepsilon-2)\\ &= \lim(\varepsilon-1)(\varepsilon-2) = 2!\end{aligned}$$

The last two entries of **Table 1** follow from

$$\Delta e^{in\alpha} = (e^{i\alpha}-1)e^{in\alpha} = 2i\ \sin \frac{\alpha}{2e^{i[n+(1/2)]\alpha}}$$

on taking real and imaginary parts.

If $\Delta F(n) = f(n)$, $F(n)$ is called the antidifference of $f(n)$ and written $\Delta^{-1}f(n)$. Two antidifferences of $f(n)$ differ at most by a constant (or by a periodic function of period 1 for functions $f(x)$ of a continuous variable). Table 1 implies the table of antidifferences (**Table 2**).

TABLE 3. Sequence of cubes

n	0	1	2	3	
$P(n)$	0	1	8	27	$P(0) = 0$
$\Delta P(n)$	1	7	19		$\Delta P(0) = 1$
$\Delta^2 P(n)$	6	12			$\Delta^2 P(0)/2! = 3$
$\Delta^3 P(n)$	6				$\Delta^3 P(0)/3! = 1$

Just as antiderivatives are used to compute definite integrals, antidifferences are used to compute definite sums:

$$\sum_{n=p}^{q} f(n) = \Delta^{-1}f(n)\ \Big|_{p}^{q+1} = F(q+1) - F(p)$$

The proof is immediate on replacing $f(n)$ by $\Delta F(n) = F(n+1) - F(n)$ and performing the indicated summation.

Example 1. The arithmetic progression, as in notation (1):

$$\begin{aligned}\sum_{n=1}^{N} a + b(n-1) &= (a-b)n + {}^1/_2 bn^{(2)}\ \Big|_{1}^{N+1}\\ &= (a-b)N + {}^1/_2 b(N+1)N\end{aligned}$$

Example 2. The geometric progression, as in notation (2):

$$\sum_{n=1}^{N} ar^{n-1} = \frac{ar^{n-1}}{r-1}\ \Big|_{1}^{N+1} = a\frac{r^N-1}{r-1}$$

Example 3. The cosine sequence:

$$\begin{aligned}\sum_{n=1}^{N} \cos n\alpha &= \frac{\sin(n-{}^1/_2)\alpha}{2\sin {}^1/_2\alpha}\ \Big|_{1}^{N+1}\\ &= \frac{\sin(N+{}^1/_2)\alpha - \sin {}^1/_2\alpha}{2\sin {}^1/_2\alpha}\end{aligned}$$

In order to sum a polynomial $P(n)$ of degree k, $P(n)$ may be expressed in terms of factorial powers $n^{(1)}$, $n^{(2)}, \ldots, n^{(k)}$ by Newton's theorem:

$$P(n) = P(0) + \frac{\Delta P(0)}{1!}n^{(1)} + \frac{\Delta^2 P(0)}{2!}n^{(2)} + \cdots + \frac{\Delta^k P(0)}{k!}n^{(k)}$$

The differences of $P(n)$ when $n = 0$ are computed as in **Table 3**.

In Table 3, $P(n) = n^3$ and hence

$$n^3 = n^{(1)} + 3n^{(2)} + n^{(3)}$$

$$\begin{aligned}\sum_{n=1}^{N} n^3 &= {}^1/_2 n^{(2)} + n^{(3)} + {}^1/_4 n^{(4)}\ \Big|_{1}^{N+1}\\ &= {}^1/_4 n^2(n-1)^2\ \Big|_{1}^{N+1} = {}^1/_4(N+1)^2N^2\end{aligned}$$

Summation by parts,

$$\begin{aligned}&\Delta^{-1}\left[f(n)\ \Delta g(n)\right]\\ &\qquad = f(n)\,g(n) - \Delta^{-1}\left[g(n+1)\ \Delta f(n)\right]\end{aligned}$$

is frequently useful in finding antidifferences. With $g(n) = (-1)^n$, $\Delta g(n) = 2(-1)^{n-1}$, this gives

$$\begin{aligned}&2\Delta^{-1}\left[(-1)^{(n-1)}f(n)\right]\\ &\qquad = (-1)^n f(n) - \Delta^{-1}\left[(-1)^{n-1}\Delta f(n)\right]\end{aligned}$$

When $P(n)$ is a polynomial of degree k, $\Delta^{k+1}P(n) = 0$, and this formula repeatedly applied gives

$$\Delta^{-1}[(-1)^{n-1}P(n)]$$

and hence the sum

$$\sum_{n=1}^{N}(-1)^{n-1}P(n)$$

The following tabulation permits the summation of alternating powers $(-1)^{n-1}n^k$:

k	$\Delta^{-1}[(-1)^{n-1}n^k]$
1	$\frac{1}{2}(-1)^n(n-\frac{1}{2})$
2	$\frac{1}{2}(-1)^n n(n-1)$
3	$\frac{1}{2}(-1)^n(n-\frac{1}{2})(n^2-n-\frac{1}{2})$
4	$\frac{1}{2}(-1)^n n(n-1)(n^2-n-1)$
5	$\frac{1}{2}(-1)^n(n-\frac{1}{2})(n^2-n-1)^2$

Summation by parts also gives antidifferences such as $\Delta^{-1}[r^nP(n)]$. Thus $\Delta^{-1}(nr^n) = r^{n-1}p(n-p)$, where $p = r/(r-1)$. *See* SERIES. Louis Brand

Bibliography. W. G. Kelley and A. C. Peterson, *Difference Equations*, 2d ed., 2001; L. M. Milne-Thomson, *The Calculus of Finite Differences*, updated edition, 2004; P. K. Rees, F. W. Sparks, and C. S. Rees, *College Algebra*, 10th ed., 1990.

Projective geometry

A geometry that investigates those properties of figures that are unchanged (invariant) when the figures are projected from a point to a line or plane. Although isolated theorems appeared earlier, the first systematic treatment of the subject was given by the French army officer. J. V. Poncelet (1788–1867), who began his *Traité des propriétés projectives des figures* in 1812 while a prisoner of war in Russia.

Two features of plane projective geometry are (1) introduction of an ideal line that each ordinary line g intersects (the intersection being common to all lines parallel to g), and (2) the principle of duality, according to which any statement that is obtained from a valid one (theorem) by substituting for each concept involved, its dual, is also valid. ("Line" and "point" are dual, "connecting two points by a line" is dual to "intersecting two lines," and so on.) The subject has been developed both synthetically (as a logical consequence of a set of postulates) and analytically (by the introduction of coordinates and the application of algebraic processes). The characteristic properties of geometrical projection are that it defines a one-to-one correspondence that preserves cross ratio between the two sets of entities involved (for example, the points of two lines, the lines of two pencils, where a pencil of lines consists of all the lines on a point). Thus if P_i, P'_i $(i = 1, 2, 3, 4)$ are corresponding points of two lines, and x_i, x'_i $(i = 1, 2, 3, 4)$ are their respective coordinate, then relation (1) holds. All such correspondences be-

$$\frac{(x_3-x_1)(x_4-x_2)}{(x_3-x_2)(x_4-x_1)} = \frac{(x'_3-x'_1)(x'_4-x'_2)}{(x'_3-x'_2)(x'_4-x'_1)} \qquad (1)$$

tween one-dimensional forms are given by the linear transformation $x' = (ax + b)/(cx + d)$, $ad - bc \neq 0$; between two-dimensional forms (for example, two planes) by relations (2), with nonvanishing determi-

$$\begin{aligned} x' &= (a_{11}x + a_{12}y + a_{13})/(a_{31}x + a_{32}y + a_{33}) \\ y' &= (a_{21}x + a_{22}y + a_{23})/(a_{31}x + a_{32}y + a_{33}) \end{aligned} \qquad (2)$$

nant $|a_{ij}|$ $(i, j = 1, 2, 3)$. The latter transformations form an eight-parameter group, the projective group of the plane. It contains, as a proper subgroup, the three-parameter group (rotations and translations) that defines euclidean geometry of the plane; and so euclidean geometry is a subgeometry of projective geometry.

Conics are defined as the class of points of intersection of corresponding lines of two projective pencils of lines (on distinct points). If these intersections are collinear, the pencils are perspective and the conic is degenerate. Otherwise, the conic is nondegenerate. Any two nondegenerate conics are equivalent; that is, a projective transformation exists that carries one into the other.

One of the well-known theorems of projective geometry was proved by the French architect G. Desargues (1593–1662): If the lines joining corresponding vertices of two triangles are concurrent, the corresponding sides intersect in collinear points. The converse of the theorem is its dual, and hence is also valid. *See* CONFORMAL MAPPING; EUCLIDEAN GEOMETRY. Leonard M. Blumenthal

Prokaryotae

Prokaryotic cells are usually microscopic, but some recently discovered prokaryotes are visible to the naked eye. A taxonomic kingdom comprising microorganisms that lack a true nucleus or any other membrane-bound organelle. Instead of a nucleus, prokaryotic cells contain a nucleoid, which is a DNA-dense region of the cell where DNA replication and transcription occur. The chromosome of these cells is circular and, in contrast to eukaryotic cells, continuously codes for cellular proteins and other molecules. Some cells also contain a plasmid, which is a smaller separate piece of circular DNA. The ribosomes of prokaryotes are found floating loose in the cytoplasm, whereas in eukaryotic cells the ribosomes are found in or on the rough endoplasmic reticulum. *See* CLASSIFICATION (BIOLOGICAL); EUKARYOTAE.

Bacteria and Archaea. Prokaryotae include the Bacteria and the Archaea. Although the Archaea and the Bacteria look identical under the microscope, they differ in their chemical composition and are unrelated. Bacteria typically have a peptidoglycan cell wall and a cytoplasmic membrane, while Archaea do not have peptidoglycan in their cell wall. Archaea can often be found in extreme environments, such as hot springs, in which most Bacteria cannot survive. *See* BACTERIA.

Characteristics. Endospores are survival structures made by some prokaryotes; some have been found to be viable after thousands of years. Prokaryotic cells undergo binary fission, where one cell

divides into two equal daughter cells. DNA may also be passed from one cell to another by conjugation, transduction, or transformation. Conjugation involves the formation of a long hollow tube, known as the sex pilus, down which DNA passes from the donor cell to the recipient cell. Transduction occurs after one cell becomes infected with a bacterial virus or bacteriophage, which can sometimes carry pieces of bacterial DNA to a new, uninfected cell. Transformation is the uptake of naked DNA released by a cell that has lysed. This DNA may contain genes useful to the prokaryote, such as antibiotic-resistance genes. *See* BACTERIOPHAGE; DEOXYRIBONUCLEIC ACID (DNA); VIRUS.

Motility. Motile prokaryotic cells move by gliding over surfaces or by flagella. Most flagella are external, found at one end of the cell (polar), at both ends of the cell (bipolar), or over the entire surface of the cell (peritrichous). The spirochetes, a group of prokaryotes that are spirally shaped, contain internal flagella known as axial filaments. Most prokaryotes are single-cell microorganisms; however, a few species are filamentous, mycelial, or colonial. *See* CELL MOTILITY.

Respiration. Different species of prokaryotes have different oxygen requirements. Some may be aerobic, requiring atmospheric oxygen; others are anaerobic, functioning in the absence of oxygen. A few prokaryotes are microaerophilic, requiring lowered oxygen and increased carbon dioxide concentrations to grow and reproduce. Prokaryotes respire to produce energy. Respiration is associated with the cytoplasmic membrane, where the electron transport chain is located. Energy is released from the reactions that occur as electrons are passed along carriers of the electron transport chain. *See* RESPIRATION.

Fermentation. Some prokaryote species use fermentative pathways for metabolism. Humans take advantage of fermentation by using these prokaryotes to produce cheese, yogurt, and pickled vegetables. Eukaryotic yeast cells also ferment and are used to produce wine and beer. *See* FERMENTATION.

Marcia Pierce

Bibliography. M. Hoppert and F. Mayer, Prokaryotes, *Amer. Scientist*, 87:518–525, 1999; J. Lederberg, *Encyclopedia of Microbiology*, 2d ed., Academic Press, San Diego, 2000; E. Nester et al., *Microbiology: A Human Perspective*, 4th ed., McGraw-Hill, New York, 2004; L. M. Prescott, J. P. Harley, and D. A. Klein, *Microbiology*, 6th ed., McGraw-Hill, New York, 2005; H. N. Schulz and B. B. Jørgensen, Big bacteria, *Annu. Rev. Microbiol.*, 55:105–137, 2001.

Prolacertiformes

A group of diapsid reptiles classified within the Archosauromorpha and recorded from Upper Permian to Lower Jurassic deposits around the world (Australia, Antarctica, India, China, Algeria, South Africa, Israel, Europe including Russia, and North America). Prolacertiforms (also known as protorosaurs) share a suite of distinctive characters, including elongate neck vertebrae with long low neural spines and long slender ribs, reduction of the lower temporal bar (a bone forming a portion of the lateral aspect of the skull and part of the base of the cranium), and emargination of the quadrate (part of the upper jaw joint) to support a tympanum (**Fig. 1**). Despite the similarities, however, there is now some doubt that prolacertiforms form a natural (monophyletic) group since some (such as the archetypal *Prolacerta*) seem to lie closer to archosaurs than others (such as *Protorosaurus* and *Tanystropheus*). *See* ARCHOSAURIA; DIAPSIDA; REPTILIA.

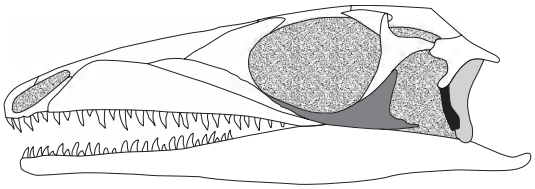

Fig. 1. Skull of *Prolacerta* in lateral view. The lower temporal bar is incomplete, with the jugal shown in dark gray, the reduced quadratojugal in black, and the emarginated quadrate in light gray. (*Redrawn and simplified from C. E. Gow, The morphology and relationships of Youngina capensis Broom and Prolacerta broomi Parrington, Palaeontol. Afr., 18:89–131, 1975*)

Characteristics. Before the adoption of modern methods of phylogenetic analysis (cladistics), prolacertiforms were one of the problematic groups for reptile systematists. Many prolacertiforms were classified with lizards or protolizards on the basis of a small set of skull characters (loss of the lower temporal bar, emargination of the quadrate to support an eardrum), despite major differences in the rest of the skeleton. Postcranial characters provide strong support for a relationship with archosauromorphs. The presence of strong rib processes on the vertebrae is an archosauromorph character, as are features of the hindlimb. In the ankle, the astragalus and calcaneum were movably articulated with one another so that the foot could face forward throughout the stride. In addition, the calcaneum had a posterior process, like the heel in mammals, which served as a lever in plantar flexion of the foot.

Diversity. Although many Late Permian and Triassic reptiles have been attributed to the prolacertiforms, many are fragmentary and/or incompletely described. Among the best-known are the primitive *Protorosaurus* (Upper Permian, Europe), *Prolacerta* (Lower Triassic, South Africa and Antarctica), *Tanystropheus* (Middle to Upper Triassic, Europe and the Middle East), and *Tanytrachelos* (Upper Triassic, North America).

Protorosaurus. *Protorosaurus* from the Upper Permian of Britain and Germany was up to 6 ft (2 m) in length, but most specimens are smaller than this. It was one of the first fossil reptiles to be formally described (in 1710), being interpreted first as a crocodile and then as a lizard. In life, *Protorosaurus* may have resembled a large monitor lizard, but there is no relationship with true lizards.

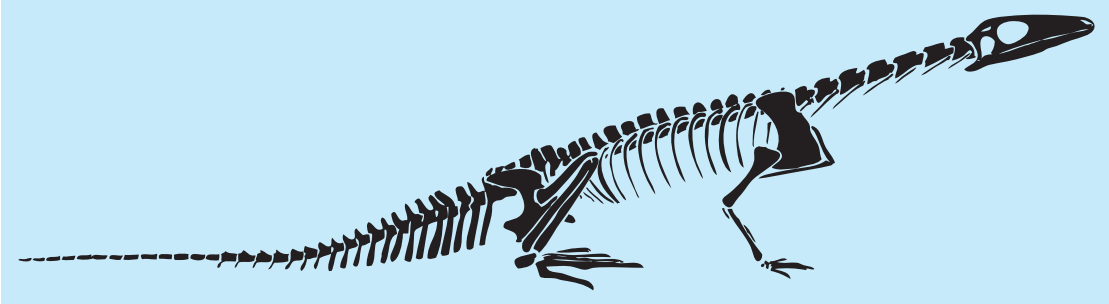

Fig. 2. Skeleton of *Prolacerta*. (*Simplified from C. E. Gow, The morphology and relationships of Youngina capensis Broom and Prolacerta broomi Parrington, Palaeontol. Afr., 18:89–131, 1975*)

Prolacerta. *Prolacerta* is smaller and more derived than *Protorosaurus*. The first specimen to be recovered (Lower Triassic, South Africa) was an isolated skull. It was originally misidentified as an ancestor of lizards. With the discovery of the postcranial skeleton, it became clear that *Prolacerta* was more closely related to archosaurs. *Prolacerta* is now known from numerous specimens (**Fig. 2**).

Tanystropheus. The most distinctive prolacertiform was *Tanystropheus* from the Middle to Upper Triassic of Europe and the Middle East. Individuals of the largest species of *Tanystropheus* could reach 9 ft (3 m) in length, but much of this consisted of an extraordinary neck that was at least three times the length of the trunk region. Not only were there more neck vertebrae than in other prolacertiforms (12 versus 8–9), but each vertebra was exceptionally elongated. Adults of *Tanystropheus* were aquatic, but juveniles may have been at least semiterrestrial, as they have sharply pointed tricuspid teeth more suited to insectivory.

Tanytrachelos. *Tanytrachelos* was closely similar to *Tanystropheus*, but smaller, from the Upper Triassic/Lower Jurassic of North America. On current evidence, *Tanytrachelos* was the last surviving prolacertiform. Susan E. Evans

Bibliography. M. J. Benton, *Vertebrate Palaeontology*, 3d ed., Blackwell, Oxford, 2004; M. J. Benton (ed.), *The Phylogeny and Classification of the Tetrapods*, vol.1: *Amphibians, Reptiles and Birds*, Clarendon Press, Oxford, 1988; R. L. Carroll, *Vertebrate Paleontology and Evolution*, W. H. Freeman, New York, 1997; A. S. Romer, *Osteology of the Reptiles*, University of Chicago Press, 1956.

Promethium

A chemical element, Pm, atomic number 61. Promethium is the "missing" element of the lanthanide rare-earth series. The atomic weight of the most abundant separated radioisotope is 147. *See* PERIODIC TABLE.

Although a number of scientists have claimed to have discovered this element in nature as a result of observing certain spectral lines, no one has succeeded in isolating element 61 from naturally occurring materials. It is produced artificially in nuclear reactors, since it is one of the products that results

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

from the fission of uranium, thorium, and plutonium.

All the known isotopes are radioactive. Its principal uses are for research involving tracers. Its main application is in the phosphor industry. It has also been used to manufacture thickness gages and as a nuclear-powered battery in space applications. *See* RARE-EARTH ELEMENTS. Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

Pronghorn

An antelopelike animal, *Antilocapra americana*, the sole representative of the family Antilocapridae, and of uncertain taxonomic affinities (see **illustration**). This animal is reputed to be the fastest ungulate in North America, sprinting as fast as 60 mi/h (96 km/h) and sustaining speeds of 30 mi/h (48 km/h) for about 15 mi (24 km). It is the only hollow-horned ungulate with branched horns present in both sexes. Like the deer, the pronghorn sheds its horns each fall; the new growth is complete by midsummer. The adult is about 3 ft (0.9 m) high, weighs about 100 lb (45 kg), is covered with coarse brittle hair, and has 32 narrow teeth with the dental formula I 0/3 C 0/1 Pm 3/3 M 3/3. *See* ANTELOPE; DENTITION.

The pronghorn lives in small herds in rather wild, rocky desert country. It blends well into the environment because of its cream-colored fur. Cartilaginous pads on the feet allow the animal to travel quickly

Pronghorn (*Antilocapra americana*), found in western North America. (*Photo by James C. Leupold/U.S. Fish and Wildlife Service*)

and quietly on the rocky terrain. Pronghorns feed on cactus, sagebrush, and other vegetation. In late summer, a buck begins to accumulate a harem of about 10–15 does. Two young are usually born in the spring after a gestation period of 35 weeks. Average lifespan for a pronghorn is about 8 years. *See* ARTIODACTYLA.

Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999; D. E. Wilson and S. Ruff, *The Smithsonian Book of North American Mammals*, 1999.

Propagator (field theory)

The probability amplitude for a particle to move or propagate to some new point of space and time when its amplitude at some point of origination is known. The propagator occurs as an important part of the probability in reactions and interactions in all branches of modern physics. Its properties are best described in the framework of quantum field theory for relativistic particles, where it is written in terms of energy and momentum. Concrete examples for electron-proton and proton-proton scattering are provided by the Feynman diagrams in the **illustration**. The amplitude for these processes contains the propagators for the exchanged photon and meson, which actually specify the dominant part of the probability of each process when the scattering occurs at small angles. In similar fashion, for any electromagnetic process, a propagator for each internal line of the Feynman diagram (each line not connected directly to the outside world) enters the probability amplitude. *See* FEYNMAN DIAGRAM; PROBABILITY (PHYSICS); QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; QUANTUM MECHANICS.

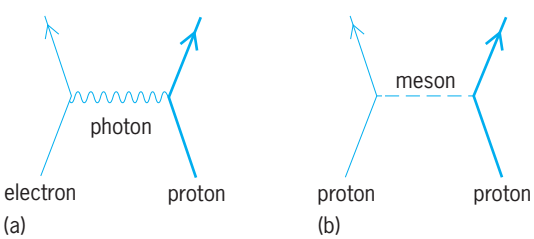

Feynman diagrams for scattering processes.
(*a*) Electron-proton scattering via photon exchange.
(*b*) Proton-proton scattering via meson exchange.

Given the wave equation for a particle or field, the corresponding Green's function is obtained by solving the wave equation when a point source for the particle or field is turned on for an instant. This Green's function solution is the probability amplitude or propagator for the particle or wave to propagate from the source to the point of observation. This procedure can readily be applied to the Schrödinger equation in nonrelativistic quantum mechanics, to Maxwell's equations for the electromagnetic field, to the meson field described by the Klein-Gordon equation in Yukawa theory, to the relativistic electron as described by the Dirac equation, and so on. The point source acting for an instant is described by the Dirac delta function $\delta(x'-x)\,\delta(y'-y)\,\delta(z'-z)\,\delta(t'-t)$, abbreviated as $\delta^3(\mathbf{x}'-\mathbf{x})\,\delta(t'-t)$, which says that a source of unit strength at (x, y, z) in three-dimensional space is turned on at the instant t. *See* GREEN'S FUNCTION; MAXWELL'S EQUATIONS; NONRELATIVISTIC QUANTUM THEORY; RELATIVISTIC QUANTUM THEORY; WAVE EQUATION.

An illustration of this calculational procedure can be given by use of the relativistic version of the Schrödinger equation for a free particle, Eq. (1), and its Fourier transformation, Eq. (2).

$$-\hbar^2\frac{\partial^2}{\partial t^2}\phi(\mathbf{x},t) = (-\hbar^2c^2\nabla^2 + m^2c^4)\phi(\mathbf{x},t) \qquad (1)$$

$$-\hbar^2\frac{\partial^2}{\partial t^2}\tilde{\phi}(\mathbf{p},t) = (\mathbf{p}^2c^2 + m^2c^4)\tilde{\phi}(\mathbf{p},t) \qquad (2)$$

The point source delta function can be inserted in Eq. (1) and $\phi(\mathbf{x},t)$ relabeled as $G(\mathbf{x}'-\mathbf{x}, t'-t)$, yielding Eq. (3). The result for each space Fourier component, from Eq. (2), is given by Eq. (4). When the corresponding time-energy Fourier components are inserted, this equation yields Eq. (5). The propagator in momentum space is formally obtained by division of Eq. (5), to yield Eq. (6), where, for illus. (*b*), E

$$\left(-\hbar^2\frac{\partial}{\partial t^2} + \hbar^2c^2\nabla^2 - m^2c^4\right)G(\mathbf{x}'-\mathbf{x}, t'-t) = \delta^3(\mathbf{x}'-\mathbf{x})\delta(t'-t) \qquad (3)$$

$$\left(-\hbar^2\frac{\partial^2}{\partial t^2} - \mathbf{p}^2c^2 - m^2c^4\right)\tilde{G}(\mathbf{p}, t'-t) = \delta(t'-t) \qquad (4)$$

$$(E^2 - \mathbf{p}^2c^2 - m^2c^4)\tilde{G}(\mathbf{p}, E) = 1 \qquad (5)$$

$$\tilde{G}(\mathbf{p}, E) = \frac{1}{(E^2 - \mathbf{p}^2c^2 - m^2c^4)} \qquad (6)$$

and $\mathbf{p}$ are the energy and momentum carried by the exchanged meson. The quantity $G\dot{w}$ is defined when this denominator goes to zero by specifying boundary conditions which G must satisfy; for example, for the retarded propagator, $G = 0$ if t' is less than t.

Kenneth E. Lassila

Bibliography. J. D. Bjorken and S. Drell, *Relativistic Quantum Fields*, 1965; J. D. Bjorken and S. Drell, *Relativistic Quantum Mechanics*, 1964; C. Itzykson and J. B. Zuber, *Quantum Field Theory*, 1980; M. Kaku, *Quantum Field Theory: A Modern Introduction*, 1993.

Propellant

Usually, a combustible substance that produces heat and supplies ejection particles, as in a rocket engine. A propellant is both a source of energy and a working substance; a fuel is chiefly a source of energy, and a working substance is chiefly a means for expending energy. Because the distinction is more decisive in rocket engines, the term propellant is used primarily to describe chemicals carried by rockets for propulsive purposes. *See* AIRCRAFT FUEL; THERMODYNAMIC CYCLE.

Propellants are classified as liquid or as solid. Even if a propellant is burned as a gas, it may be carried under pressure as a cryogenic liquid to save space. For example, liquid oxygen and liquid hydrogen are important high-energy liquid bipropellants. Liquid propellants are carried in compartments separate from the combustion chamber; solid propellants are carried in the combustion chamber. The two types of propellants lead to significant differences in engine structure and thrust control. For comparison, the effectiveness of either type of propellant is stated in terms of specific impulse. *See* ROCKET PROPULSION; SPECIFIC IMPULSE; THRUST.

Liquid Propellants

A liquid propellant releases energy by chemical action to supply motive power for jet propulsion. The three principal types of propellants are monopropellant, bipropellant, and hybrid propellant. Monopropellants are single liquids, either compounds or solutions. Bipropellants consist of fuel and oxidizer carried separately in the vehicle and brought together in the engine. Air-breathing engines carry only fuel and use atmospheric oxygen for combustion. Hybrid propellants use a combination of liquid and solid materials to provide propulsion energy and working substance. Typical liquid propellants are listed in the **table**. Physical properties at temperatures from storage to combustion are important. These properties include melting point, boiling point, density, and viscosity. *See* METAL-BASE FUEL.

The availability of large quantities and their high performance led to selection of liquefied gases such as oxygen for early liquid-propellant rocket vehicles. Liquids of higher density with low vapor pressure (see table) are advantageous for the practical requirements of rocket operation under ordinary handling conditions. Such liquids can be retained in rockets for long periods ready for use and are convenient for vehicles that are to be used several times. The high impulse of the cryogenic systems is desirable for rocket flights demanding maximum capabilities, however, such as the exploration of space or the transportation of great weights for long distances.

Performance. Jet propulsion by a reaction engine, using the momentum of the propellant combustion products ejected from the engine, is not limited to atmospheric operation if the fuel reacts with an oxidizer carried with the engine. Performance of the propellant in such an engine depends upon both the heat liberated and the propellant reaction products. Combustion with air of effective fuels for air-breathing engines gives approximately 18,000 Btu/lb (42 megajoules/kg), whereas fuels which are more effective in rocket engines may give only 15,000 Btu/lb (35 MJ/kg). A high heat of reaction is most effective with gaseous products which are of low molecular weight.

Performance is rated in terms of specific impulse (occasionally specific thrust), the thrust obtained per pound of propellant used in 1 s. An alternate measure of performance is the characteristic exhaust velocity. The theoretical characteristic exhaust velocity is determined by the thermodynamic properties of the propellant reaction and its products. Unlike the specific impulse, the characteristic exhaust velocity is independent of pressure, except for second-order

Physical properties of liquid propellants

Propellant	Boiling point, °F (°C)	Freezing point, °F (°C)	Density, g/ml	Specific impulse,* s
Monopropellants				
Acetylene	−119 (−84)	−115 (−82)	0.62	265
Hydrazine	236 (113)	35 (2)	1.01	194
Ethylene oxide	52 (11)	−168 (−111)	0.88	192
Hydrogen peroxide	288 (142)	13 (−11)	1.39	170
Bipropellants				
Hydrogen	−423 (−253)	−433 (−259)	0.07	
Hydrogen-fluorine	−306 (−188)	−360 (−218)	1.54	410
Hydrogen-oxygen	−297 (−183)	−362 (−219)	1.14	390
Nitrogen tetroxide	70 (21)	12 (−11)	1.49	
Nitrogen-tetroxide-hydrazine	236 (113)	35 (2)	1.01	290
Red nitric acid	104 (40)	−80 (−62)	1.58	
Red fuming nitric acid-*uns*-dimethyl hydrazine	146 (63)	−71 (−57)	0.78	275

*Maximum theoretical specific impulse at 1000 psi (6.895 megapascals) chamber pressure expanded to atmospheric pressure.

effects, such as reactions modifying the heat capacity ratio of the combustion gases.

The relationship between these parameters is given by Eq. (1), in which I_s is specific impulse in

$$I_s = \frac{F}{\dot{w}} = \frac{c^* C_F}{g} \tag{1}$$

seconds, F is thrust, and $\dot{w}$ is flow rate of propellant. The characteristic exhaust velocity c^* is given in feet per second. C_F is the thrust coefficient, and g is the gravitational constant.

The actual exhaust velocity of the combustion gases is given by the product of the characteristic exhaust velocity and the thrust coefficient, $c^* C_F$. The thrust coefficient is a function of the heat capacity ratio of the combustion gases (the ratio of the heat capacity at constant pressure to the heat capacity at constant volume) and of the ratio of the chamber pressure to the exhaust pressure. The heat-capacity ratio of common propulsion gases varies from 1.1 to 1.4.

Increase in the combustion-chamber pressure increases the specific impulse (**Fig. 1**). Variation in the stoichiometry of the propellant reaction (the oxidizer-fuel ratio) also affects performance. A slightly fuel-rich reaction gives higher performance with common liquid propellants despite the lower heat of reaction because of more favorable working-gas composition. Increase in chamber pressure usually moves the optimum performance point toward the stoichiometric reaction ratio. A nonstoichiometric ratio may be used to give low combustion temperatures if required by the structural materials.

A properly designed engine can give 95–100% of the theoretical performance shown in the table.

Combustion. The energy of liquid propellants is released in combustion reactions which also produce the working fluid for reaction propulsion. The liquids in a bipropellant system may ignite spontaneously on contact, or they may require an ignition device to raise them to ignition temperature. In the first case they are called hypergolic liquids; in the second case, anergolic liquids. Combustion can be initiated with a spark, a hot wire, or an auxiliary hypergolic liquid. Monopropellant combustion, or more properly decomposition, can also be ignited by catalysis with an active surface or by a chemical compound in solution. Ignition of common hypergolic bipropellants occurs in a period of 1–100 ms following initial contact of the liquids. Catalytic quantities of detergents or of certain compounds of metals with several oxidation states, such as vanadium pentoxide, decrease the ignition delay period of specific bipropellants.

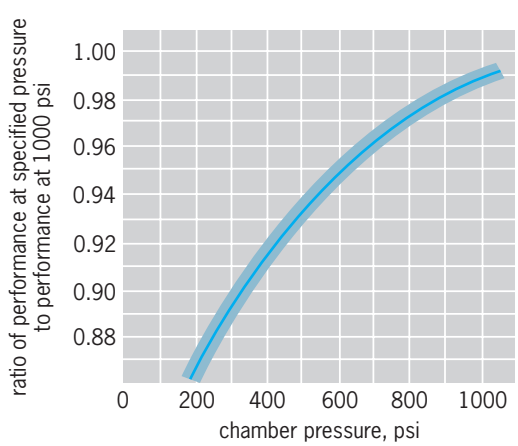

Fig. 1. Approximate effect of chamber pressure on specific impulse. 1 psi = 6.9 kPa.

The combustion chamber in operation contains a turbulent, heterogeneous, high-temperature-reaction mixture. The liquids burn with droplets of various sizes in close proximity and traveling at high velocity. Larger masses of liquid may be present, particularly at the chamber walls. Very high rates of heat release, of the order of 10^5–10^6 Btu/(min)(ft^3)(atm) [10^3–10^4 joules/(s)(m^3)(pascal)] are encountered.

Oscillations with frequencies of 25–10,000 Hz or more may accompany combustion of liquids in jet-propulsion engines. Low-frequency instability (chugging) can result from oscillations coupling the liquid flowing into the combustion chamber with pressure pulses in the chamber. Higher frequencies (screaming) can result from gas oscillations of the acoustic type in the chamber itself.

Engine performance, in contrast to theoretical propellant performance, depends upon effective combustor design. Mixing and atomization are essential factors in injection of the propellants into the combustion chamber. Injector and chamber design influence the flow pattern of both liquid and gases in the chamber. The characteristic chamber length L^* is given by Eq. (2), in which V_c is the chamber volume and A_T is the area of the nozzle throat. In general, monopropellants require larger L^* than bipropellants to provide an equal fraction of theoretical performance in a rocket engine, as expected from the slower combustion exhibited by monopropellants.

$$L^* = \frac{V_c}{A_T} \tag{2}$$

Stanley Singer

Solid Propellants

A solid propellant is a mixture of oxidizing and reducing materials that can coexist in the solid state at ordinary temperatures. When ignited, a propellant burns and generates hot gas. Although gun powders are sometimes called propellants, the term solid propellant ordinarily refers to materials used to furnish energy for rocket propulsion.

Composition. A solid propellant normally contains three essential components: oxidizer, fuel, and additives. Oxidizers commonly used in solid propellants are ammonium and potassium perchlorates, ammonium and potassium nitrates, and various organic nitrates, such as glyceryl trinitrate (nitroglycerin). Common fuels are hydrocarbons or hydrocarbon derivatives, such as synthetic rubbers, synthetic resins, and cellulose or cellulose derivatives. The additives, usually present in small amounts, are chosen

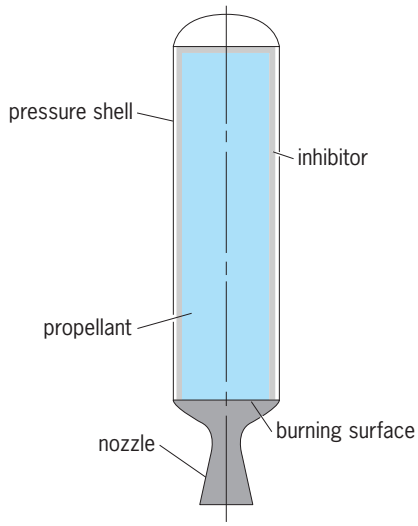

Fig. 2. End-burning grain loaded in rocket.

from a wide variety of materials and serve a variety of purposes. Catalysts or suppressors are used to increase or decrease the rate of burning; ballistic modifiers may be used for a variety of reasons, as to provide less change in burning rate with pressure (platinizing agent); stabilizers may be used to slow down undesirable changes that may occur in long-term storage.

Solid propellants are classified as composite or double base. The composite types consist of an oxidizer of inorganic salt in a matrix of organic fuels, such as ammonium perchlorate suspended in a synthetic rubber. The double-base types are usually high-strength, high-modulus gels of cellulose nitrate (guncotton) in glyceryl trinitrate or a similar solvent.

Propellants are processed by extrusion or casting techniques into what are often intricate shapes that are commonly called grains, even though they may weigh many tons. The double-base types and certain high-modulus composites are processed into grains by casting or extrusion, and are then loaded by insertion of the cartridgelike grain or grains into the combustion chamber of the rocket. This technique requires some type of mechanical support to hold the propellant in place in the chamber. Certain types of composite propellants, bonded by elastomeric fuels, can be cast directly into the chamber, where the binder cures to a rubber and the grain is then supported by adhesion to the walls. Most high-performance rockets are made by this case-bonding technique, permitting more efficient use of weight and combustion-chamber volume.

Burning rate. The thrust-time characteristic of a solid-propellant rocket is controlled by the geometric shape of the grain. Often it is desired that burning not take place on certain portions of the grain. Such surfaces are then covered with an inert material called an inhibitor or restrictor. Neutral-burning grains maintain a constant surface during burning and produce a constant thrust. Progressive burning grains increase in surface and give an increasing thrust with time. Degressive or regressive grains burn with decreasing surface and give a decreasing thrust.

An end-burning grain is shown in **Fig. 2**. This type of configuration is neutral, because the surface stays constant while the grain burns forward. For most applications, radial-burning charges which burn outward from the inside perforation are superior because most of the wall area of the chamber can be protected from hot gas generated by combustion. Such protection is a built-in feature of the case-bonded grain; with the cartridge-loaded, inhibited charge, protection is provided by the addition of obturators to prevent gas flow around the outside of the grain.

Figure 3 shows a progressive design called an internal-burning cylinder. Various star-shaped perforations can be used to give neutral or degressive characteristics. By ingenious use of geometry, the thrust-time characteristic can be designed to meet almost any need. Another important neutral-grain design, the uninhibited, internal-external-burning cylinder, is used widely in short-duration applications such as the bazooka rocket weapon, which contains many such grains.

Propellant burns at a rate r proportional to a fractional power n of the pressure P as expressed by Eq. (3) in which K_1 is the coefficient

$$r = K_1 P^n \tag{3}$$

of proportionality. This rate may be determined at various pressures by measurements of the burning rate of propellant strands in a strand burner (**Fig. 4**). If the propellant is to operate properly, the exponent n in the burning-rate equation must be less than 1. As illustrated by **Fig. 5***a*, if $n < 1$, there is a stable operating pressure at which the lines of fluid generation and fluid discharge intersect. Pressure cannot rise above this value because gas would then be discharged at a faster rate than

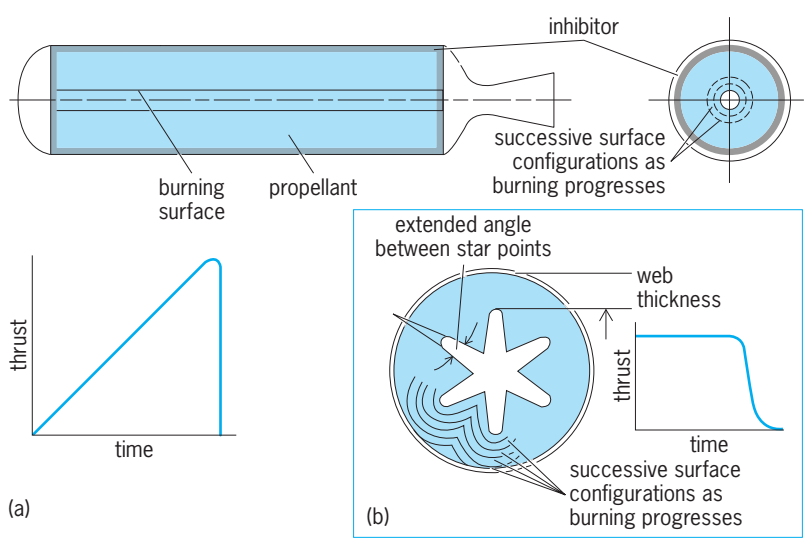

Fig. 3. Internal-burning solid-propellant charge configurations with typical thrust-time characteristics. (*a*) Cylindrical cavity. (*b*) Star-shaped cavity for level, or neutral, thrust-time characteristic.

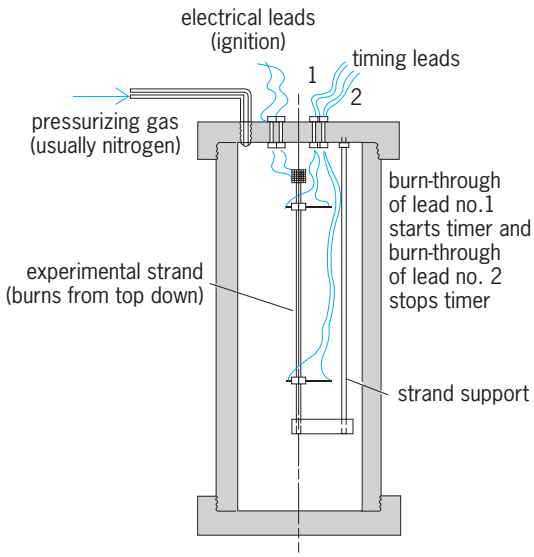

Fig. 4. Strand burner apparatus.

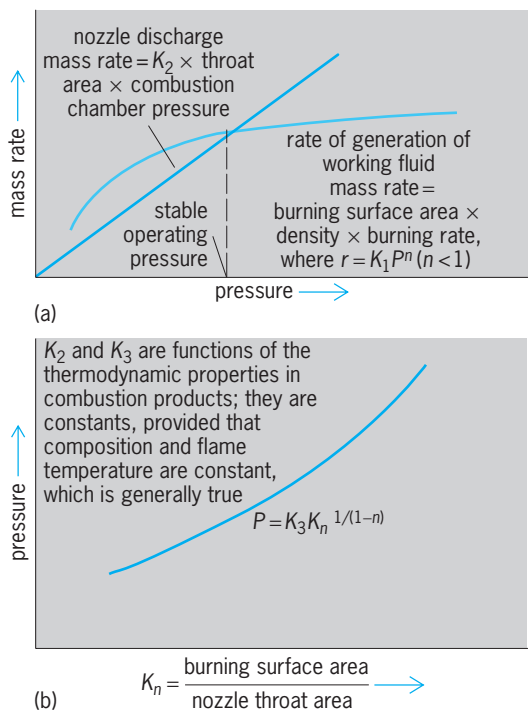

Fig. 5. Solid propellant. (*a*) Stable burning condition. (*b*) Typical pressure characteristic.

it is generated by burning of propellant. When the propellant meets the requirement $n < 1$, a relationship exists between operating chamber pressure and a design parameter known as K_n, which is the ratio of the propellant burning area to the nozzle throat area. This ratio is illustrated in Fig. 5*b*.

Specific impulse of solid propellants is normally rated with the rocket operating at a chamber pressure of 1000 lb/in.^{2} (6.9 megapascals) and exhausting through an optimum nozzle into sea-level atmosphere. Under these conditions, solid propellants in use today can give an impulse of about 250, which is near the ceiling imposed by compositions based on ammonium perchlorate and hydrocarbons, and is 5–10% lower than impulses obtainable from liquid oxygen and gasoline.

The lower specific impulse of solid propellants is partly overcome by their densities, which are higher than those of most liquid propellants. In addition, solid-propellant rockets are easy to launch, are instantly ready, and have demonstrated a high degree of reliability. Because they can be produced by a process much like the casting of concrete, there seems to be no practical limit to the size of a solid-propellant rocket. H. W. Ritchey

Bibliography. A. Davenas (ed.), *Solid Rocket Propulsion Technology*, 1992; L. DeLuca and M. Summerfield (eds.), *Nonsteady Burning and Combustion Stability of Solid Propellants*, 1992; P. G. Hill and C. R. Peterson, *Mechanics and Thermodynamics of Propulsion*, 2d ed., 1992; G. A. Olah and D. R. Squire (eds.), *Chemistry of Energetic Materials*, 1997; G. P. Sutton and D. M. Ross, *Rocket Propulsion Elements: An Introduction to the Engineering of Rockets*, 6th ed., 1992; M. Timnat (ed.), *Advanced Chemical Rocket Propulsion*, 1987.

Propeller (aircraft)

A hub-and-multiblade device for changing rotational power of an aircraft engine into thrust power for the purpose of propelling an aircraft through the air (**Fig. 1**). An air propeller operates in a relatively thin medium compared to a marine propeller, and is therefore characterized by a relatively large diameter and a fairly high rotational speed. It is usually mounted directly on the engine drive shaft in front of or behind the engine housing. *See* PROPELLER (MARINE CRAFT).

Propellers range in size from diameters of 15 cm (6 in.) for model airplanes up to 5.8 m (19 ft) for

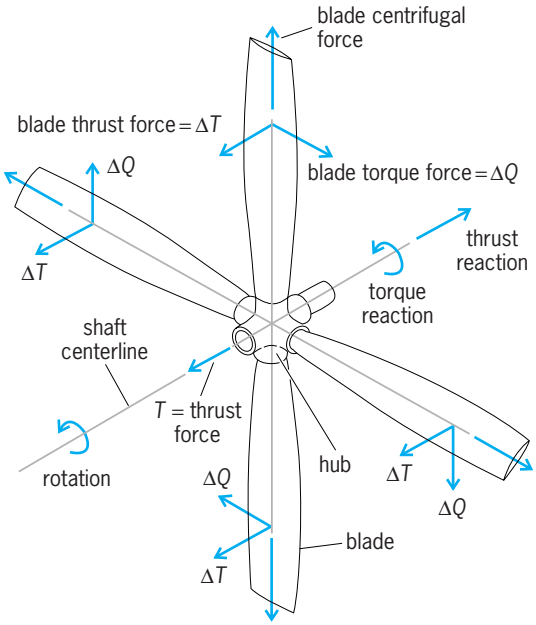

Fig. 1. Typical four-bladed propeller system.

some post–World War II bomber-type airplanes. Usually propellers have two, three, or four blades. For high-speed or high-powered airplanes, six or more blades are used. In some cases these propellers have an equal number of opposite rotating blades on the same shaft, and are known as dual-rotation propellers. The engine power converted to thrust power by propellers ranges from a fraction of a horsepower to values in excess of 10,000 hp (1 hp = 746 W). The usual operating forward speed of propellers ranges from 0 to 450 knots (230 m/s). The rotational tip speed is in the range of 198 to 290 m/s (650 to 950 ft/s), corresponding to a Mach number range of 0.58 to 0.85. For low propeller noise output, it is important to operate at low values of rotational tip speed.

Until the development of turbine-powered straight-jet or turbofan engines, the propeller was the only efficient device available for propelling an airplane through the air. Except for a few special-purpose military and cargo airplanes, the propeller is used mainly on general aviation airplanes having reciprocating and turbo shaft engines with power ratings of 75 to 1500 hp (55 to 1100 W). *See* GENERAL AVIATION.

The performance of a propeller is measured in terms of its efficiency η, which is the ratio of the power output over the power input. Since the power output is equal to the thrust force of the propeller times the speed V of the airplane, the efficiency is given by Eq. (1). The efficiency of well-designed propellers is 85–90% at cruise flight conditions.

$$\eta = \frac{\text{power out}}{\text{power in}} = \frac{\text{thrust} \cdot \text{velocity}}{\text{power}} = \frac{TV}{P} \qquad (1)$$

Principles of operation. A propeller blade advances through the air along an approximate helical path which is the result of its forward and rotational velocity components (**Fig. 2**). This reaction is similar to a screw being turned in a solid surface, except that in the case of the propeller a slippage occurs because air is a fluid. Because of the similarity to the action of a screw, a propeller is also known as an airscrew. To rotate the propeller blade, the engine exerts a torque force. This force is reacted by the blade in terms of lift and drag force components produced by the blade sections in the opposite direction (Fig. 2). As a result of the rotational forces reacting on the air, a rotational velocity remains in the propeller wake with the same rotational direction as the propeller. This rotational velocity times the mass of the air is proportional to the power input. The sum of all the lift and drag components of the blade sections in the direction of flight is equal to the thrust produced. These forces react on the air, giving an axial velocity component opposite to the direction of flight. By the momentum theory, this velocity times the mass of the air going through the propeller is equal to the thrust.

The axial momentum efficiency of a propeller η_m is inversely proportional to the axial velocity component v in the final wake. At an aircraft velocity V

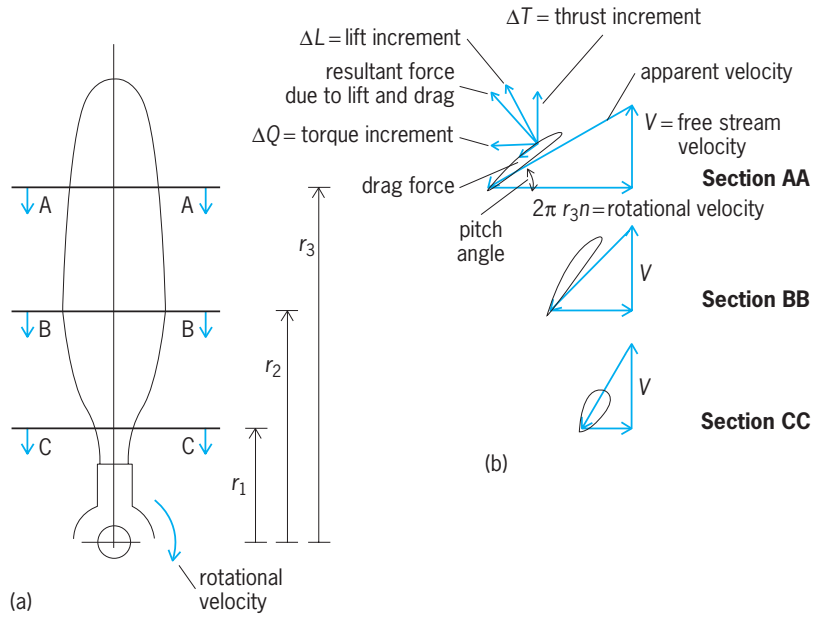

Fig. 2. Propeller blades with velocity and force components. (*a*) Front view. (*b*) Cross sections. Here *n* = angular rotational velocity, in rpm.

the axial momentum efficiency is given by Eq. (2).

$$\eta_m = \frac{2V}{2V + v} \qquad (2)$$

Thus, for high efficiency, the slip velocity in the final propeller wake should be small. Since the thrust is equal to the change in axial momentum, mv, it is apparent that for high efficiency the mass of air, m, should be as high as possible. This implies that, for a given power input, high efficiency is obtained with the highest-diameter propeller possible and is a typical propeller design characteristic. In a jet engine the thrust is produced with a low mass flow and a high wake or slip velocity, thus giving a lower efficiency.

To develop lift, and therefore thrust, on a blade, airfoil sections similar to those used on a wing are placed along the blade span from the hub to the tip. These sections have a blade angle distribution that decreases from hub to tip approximately as the pitch of the helix of the screw (Fig. 2). The size of the airfoil and the number of blades on the propeller are chosen so that the angle of attack of the section is low, and so that the sections are operating with the lift required to absorb the given power input. When the airfoils are operating at low angles of attack relative to the apparent velocity vector (Fig. 2), the high lift drag ratios needed for high efficiency are obtained. As with a wing, the drag of the airfoil depends on the operating lift coefficient, section type, and speed relative to the speed of sound (the Mach number). When the Mach number is high, large losses can be encountered that result in low values of efficiency. Propellers with swept-back blades, like wings, to reduce the compressibility losses may be used to obtain high efficiency at high flight speeds. *See* AIRFOIL; WING.

Propeller blades. A propeller blade is essentially a twisted wing with a radius six to seven times its

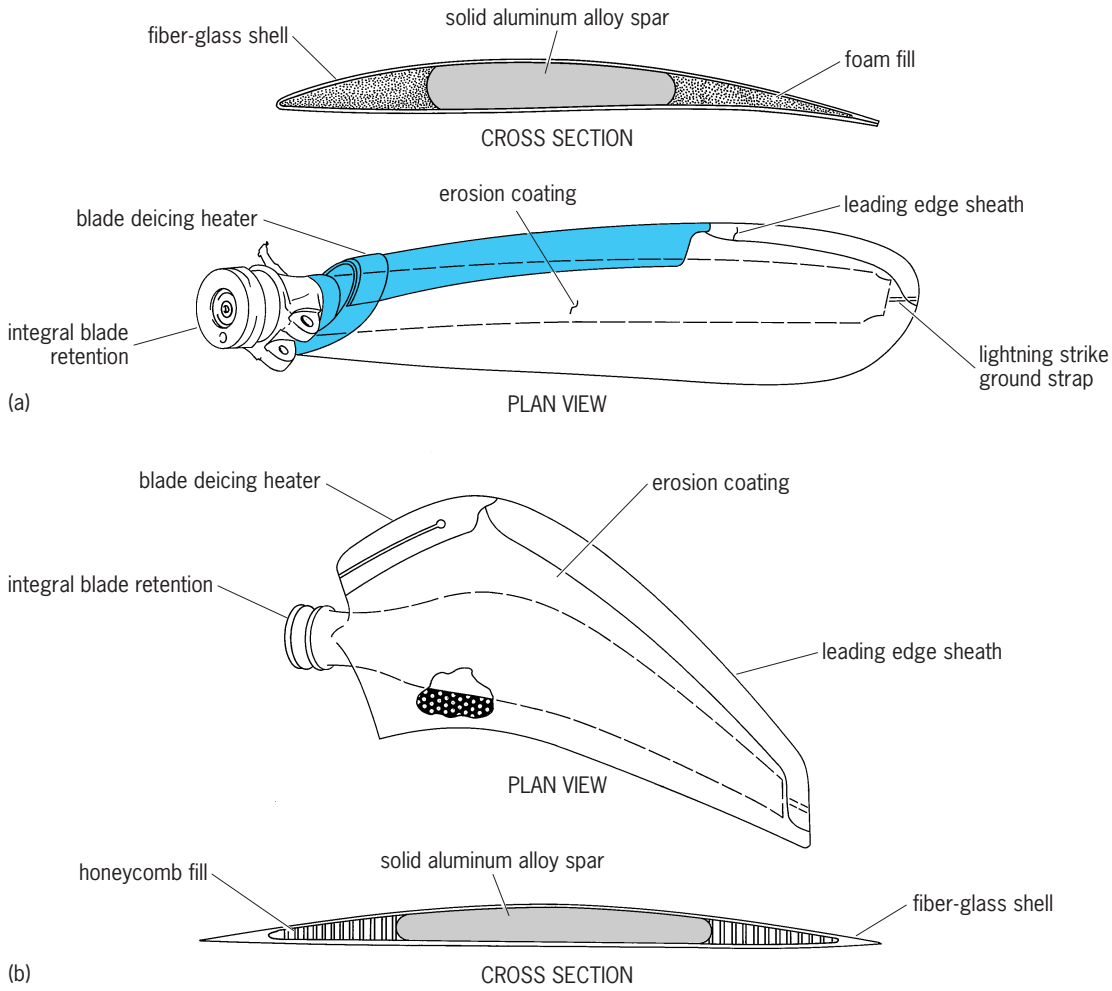

Fig. 3. Composite propeller blade construction. (*a*) Fiber-glass shell with solid aluminum alloy spar and foam fill. (*b*) Fiber-glass shell with solid aluminum alloy spar and honeycomb fill.

chord. The blade (Fig. 2 and **Fig. 3**) uses airfoil sections that have a thickness-to-chord ratio of about 4% at the tip and 25% at its root. For best efficiency, the airfoil sections operate at a small angle of attack, producing lift and drag forces.

A propeller blade must be designed to withstand very high centrifugal forces, which increase from tip to root and can be of the order of 25 tons (222 kilonewtons) where the blade attaches to the hub. The blade also must withstand the thrust force produced plus any vibratory forces generated, such as those due to uneven flow fields. Such flow fields exist when the propeller shaft axis is at an angle of attack relative to the flight direction. Vibratory forces are also transmitted to a propeller by the gas pressure forces generated by reciprocating engines and by the flow field produced by the various components of the airplane. The vibratory forces can be especially high for pusher propeller installations.

To withstand the high stresses due to rotation, propeller blades such as those in Fig. 3 have been made from a number of materials, including wood, aluminum, hollow steel, plastic composites, and the construction illustrated. The most common material used has been solid aluminum. However, the composite blade construction illustrated, plus blades having a fiber-glass shell with a foam filler, are being used for new turboprop installations because of their very light weight and high strength characteristics.

Propeller types. For a small, low-power airplane, very simple, fixed-pitch, single-piece, two-blade propellers are used. The rotational speed of these propellers depends directly on the power input and forward speed of the airplane. Because of the fixed-blade angle of this type of propeller, it operates near peak efficiency only at one condition and will limit the engine rotational speed at the low forward speeds of the aircraft. Usually such a propeller is designed with a blade set for good climb or cruise performance.

To overcome the limitations of the simple fixed-pitch propeller, configurations that provide for variable blade angles are used. The blades of these propellers are retained in their hub so that they can be rotated about their centerline while the propeller rotates. For the normal range of operation, the blade angle varies from the low blade angle needed for takeoff to the high blade angle needed for the maximum speed of the airplane. The lowest blade angle is that angle which allows the engine to develop

full power at the rated engine rpm (revolutions per minute). As the speed increases, the blade angle increases to maintain the desired rpm for good engine performance and to operate at conditions for high efficiency.

The operating speed of the propeller is controlled by the propeller governor, which is set to the desired rpm by the pilot. With an increase in power, the propeller rpm increases; if it exceeds the desired value, the governor causes the blade angle to increase, which causes the propeller to absorb more power and thus decrease the rpm. A decrease in rpm results in a decrease in blade angle, with the opposite effect.

Propellers operating at blade angles between those needed for takeoff and high-speed flight are used on single-engine aircraft. On multiengine airplanes the blade angle range is increased so that the propeller can develop negative thrust for braking during landing. As a safety feature, a full-feathering blade angle is also provided. Here the blade angle is increased to almost 90° so that the propeller will not windmill and its drag will be a minimum when an engine fails.

There have been many different means used for changing the blade angle of a propeller. These include the use of mechanical, electrical, and hydraulic actuators. The hydraulic actuator controlled by an rpm governor is the system most commonly used today. The relatively simple nonfeathering propeller uses hydraulic pressure to actuate a piston that is attached to the blade and increases the blade angle with an increase of pressure. When the pressure on the piston is reduced, the centrifugal twisting moment of the blade is in the direction to reduce blade angle. With a full-feathering and reversing propeller blade, counterweights and a spring are used instead of hydraulic pressure to cause the angle to increase. With this type of propeller (**Fig. 4**), the hydraulic pressure increase causes the blade angle to be reduced by overcoming the twisting moment of the counterweights and the spring. The governor controls the hydraulic pressure to the actuator, and thus

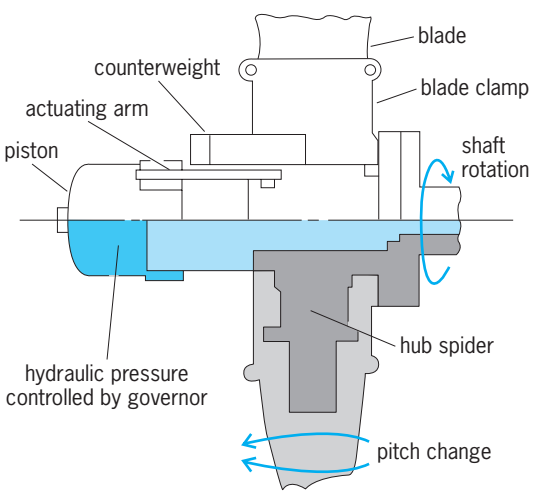

Fig. 4. Actuating mechanism to change propeller pitch.

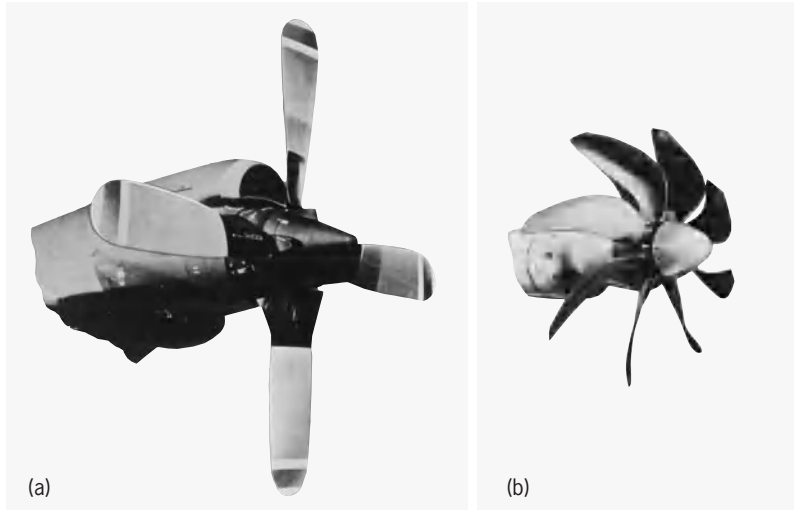

Fig. 5. Size and configuration comparison of (*a*) conventional turboprop installation and (*b*) advanced turboprop fan. (*Hamilton Standard Division of United Technologies*)

the blade angle. This system has important inherent safety features since it increases the blade angle in the event of a loss of hydraulic pressure, thus preventing a dangerous overspeeding of the engine and propeller.

With large propellers, the addition of counterweights and a spring results in a large weight increase. To reduce weight and eliminate the counterweights, hydraulic pressure is used to increase and decrease the blade angle of the propeller. With such a system, safety devices must be used to control the system and prevent unwanted changes of blade angle in the event of failures.

A typical four-blade conventional propeller for a high-powered turbo propeller is illustrated in **Fig. 5***a*. This propeller has a power loading of 80 to 120 kW/m^2 or 10 to 15 SHP/D^2, where SHP is the shaft horsepower input to the propeller with a diameter of Dfeet. Compared to the conventional propeller is a propeller fan scaled to reflect the higher power loading of 300 kW/m^2 (37.5 SHP/D^2) used for this configuration (Fig. 5*b*). The propeller fan, with light, swept-back blades, is designed for use with 0.8-Mach cruise high-speed airplanes. *See* AIRCRAFT PROPULSION; AIRPLANE; HELICOPTER. Henry V. Borst

Bibliography. F. Delp, *Aircraft Propellers and Controls*, 1979; H. Glauert, *The Elements of Aerofoil and Airscrew Theory*, 1944, reprint 1983; M. Kroes and T. W. Wild, *Aircraft Powerplants*, 7th ed., 1994.

Propeller (marine craft)

The component of a ship propulsion power plant that converts engine torque into propulsive force or thrust, thus overcoming the ship's resistance to forward motion by creating a sternward accelerated column of water. Early steamship propulsion forms, such as the jet propeller (1782) and the paddle wheel (1801), were gradually replaced by the screw propeller (1804), which since 1860 has been the only propeller type used in ocean transport, mainly

because of the evolution of the marine engine toward higher rotative speed.

The advantages of the screw propeller include light weight, flexibility of application, good efficiency at high rotative speed, and relative insensitivity to ship motion. The fundamental theory of screw propellers is applicable to all forms of marine propellers. In its current form, a screw propeller consists of a streamlined hub, attached outboard to a rotating engine shaft, on which are mounted two to seven blades. The blades can be integral with the hub, detachable, or movable. The screw propeller indeed has the characteristic motion of a screw: It rotates about the shaft axis as the ship advances through the water. The screw blades are approximately elliptical in outline. Propellers are being developed in which the major axis of the ellipse is an arc, which can give the propeller blades the appearance of a scimitar. One or more propellers are usually fitted as low as possible at the ship's stern to act as thrust-producing devices. The low position of the propellers affords good protection and sufficient immersion during the pitching motions of the ship. The diameter of the propeller usually is not larger than 0.80 of the loaded draft in seagoing single-screw vessels. *See* SHIP DESIGN.

Geometry. The side of a propeller blade away from the hull is called the face or pressure side. The opposite side is called the back or suction side. In its simplest form the blade face is part of a helicoid surface, and its axial advance per revolution defines the propeller's pitch which, in the case of a constant-pitch propeller, is constant over the radius, R. Many modern propeller designs which are adapted to the complex flow pattern behind the vessel's stern, called the wake, employ a radially varying pitch (variable-pitch propeller). The pitch occurring at $0.7R$ is then taken to be representative for the nominal pitch. The dimensionless ratio of the pitch, P, to the propeller diameter, D, is the pitch ratio. It is an important characteristic of the screw propeller.

The difference between the pitch velocity, nP (where n is revolutions per second), and the mean axial speed of advance, V_A, of the screw through the surrounding medium is called the slip velocity. Expressed as a fraction of nP, this yields a dimensionless quantity called the real slip ratio, S_R, as given in Eq. (1),which is one of the fundamental quantities

$$S_R = \frac{nP - V_A}{nP} = \frac{1 - V_A}{nP} \tag{1}$$

determining propeller action and efficiency. In propeller technology this relationship is expressed as propeller advance ratio, Eq. (2). The relationship be-

$$J_A = \frac{V_A}{nD} \tag{2}$$

tween real slip ratio and advance ratio is expressed in Eq. (3).

$$J_A = \frac{P}{D}(1 - S_R) \tag{3}$$

Blade shape. The propeller blade outlines can take a number of different forms depending upon the design objective. The majority are essentially elliptical with a small amount of skew. Skew is the lateral shift of the centerline of the blade from a straight reference line radiating from the centerline of the blade at the hub. This skew line can take any continuous shape from the root to the tip of the blade. On a right-hand-turning propeller, most of the skew is to the left of the reference line. The magnitude of the skew at the tip is usually defined as the angle in degrees that the tip makes with the reference line at the centerline axis of the propeller. Conventional propellers usually have less than 10° of skew. Propellers with more than 25° of skew are usually considered "highly" skewed and require special care in design to ensure adequate strength in the blades. **Figure 1** shows an example of a highly skewed propeller. Propellers with this blade shape produce significantly less propeller-induced blade-frequency vibration. Some propeller blades, such as those used in controllable-pitch propellers with high skew, skew the blade to the right of the reference line in the root region and then heavily skew the blade in the tip region to the left of the reference line. The purpose is to produce a blade that has low spindle torque about the blade axis in order to minimize the blade turning torque and obtain the advantages of the high-skew propeller in reducing the blade-frequency vibratory forces.

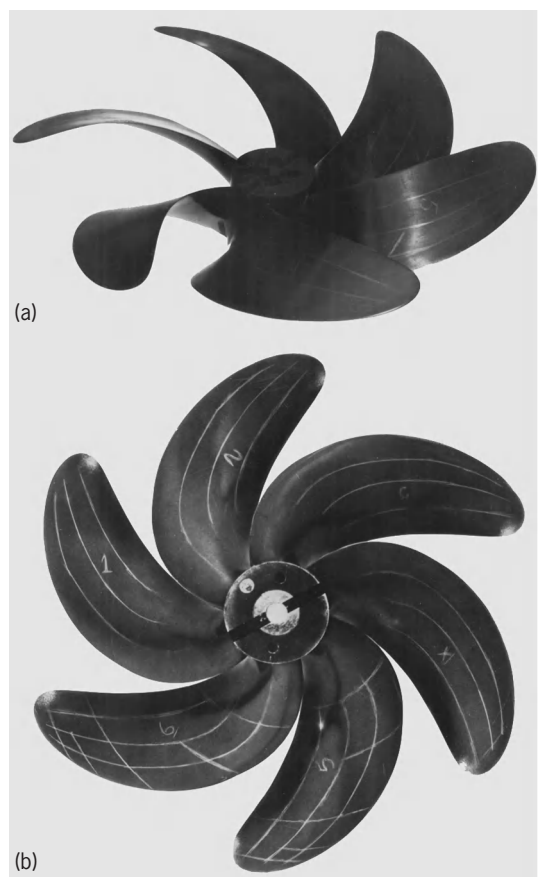

Fig. 1. Highly skewed model propeller. (*a*) Side view. (*b*) Front view.

Theory. Propeller theory developed along two main lines: the momentum theory and the circulation theory. In the former, the origin of the thrust is explained by the change in momentum that takes place in the fluid. This theory views the propeller as an actuator disc (infinite number of blades with no thickness) in an ideal fluid, thus offering no indication as to the proper shape of the blades. It does provide useful information on the ideal efficiency, contraction of the slipstream (the bounding cylindrical streamline which contains the flow accelerated by the propeller when creating thrust), and the velocities inside and outside the propeller slipstream.

In order to design the detailed shape of the blades for the desired thrust, it became necessary to utilize a theory based on the lift-drag characteristics of propeller blade sections. Initially, various blade element theories were developed, but these required empirically derived corrections to provide the desired results. These older theories have been superseded by circulation theory, which treats the additional axial and rotational velocities in the slipstream as being induced by vortex systems distributed radially in the lifting line theory and over the whole blade in the lifting surface theory.

The lifting line representation uses a radial lifting line to represent the radial distribution of the propeller blade loading. It provides the overall propeller performance at the design condition. The theory has been extended to represent the loading on the whole surface of the blade, thus providing the detailed shape of the propeller blade sections to create the desired thrust. The theory has been augmented by the panel methods, which use appropriate combinations of sources, sinks, vortices, and dipoles distributed over the blade surface. The panel methods are particularly useful in predicting off-design and cavitation performance. The use of these theories, particularly the panel methods, requires the aid of digital computers but does permit the designer to achieve the desired performance at both the design and off-design conditions.

Model tests. To aid in propeller design, open-water experiments are done in towing tanks using small-scale model propellers. Self-propulsion tests with a corresponding model of the ship establish the interaction of the propeller with the hull. These experiments are used to verify design calculations, particularly the thrust, torque, and revolutions per minute (rpm). In some instances, to assist in vibration analysis, a specialized dynamometer is used to measure the blade-frequency forces generated by the propeller. *See* TOWING TANK.

Number of propellers. The choice of the number of propellers to incorporate into a vessel design is based on several factors. In general, a single-screw arrangement yields a higher propulsive efficiency than multiple screws, particularly when most of the propeller is operating in the boundary layer of the ship and can recover some of the energy loss. In addition, single-screw propulsion systems generally result in savings in machinery cost and weight in comparison to multiple-screw arrangements.

Shaft powers in excess of 90,000 horsepower (67,100 kW) can be successfully transmitted by a single screw propeller in large, fast oceangoing ships. Greater power output necessitates the use of multiple (mostly twin or quadruple) screw arrangements. However, operational considerations may also result in a decision to opt for a multiple-screw system. In some cases, as in powerful shallow-draft towboats, the maximum available propeller diameter limits the efficient conversion of engine power per shaft. In others, preference for a certain number of diesel engine units to produce the propulsive power needed determines the number of propellers. Other considerations, such as safety or operability of the vessel in the event of engine breakdown (especially in the case of large passenger vessels), may favor multiple-screw arrangements.

Number of blades. Both the maximum open-water efficiency and the optimum blade diameter for the highest propulsive efficiency may decrease slightly with a large number of blades. In practice the number of blades varies from two to seven, with four, five, and six the most common. Two-bladed propellers are used in fast racing craft and in sailing vessels with auxiliary engine power where the propeller can be locked in an upright position behind the keel to reduce drag when sailing. In oceangoing ships, the natural frequencies of vibration of the propeller shaft, the hull, and the propulsive plant and their major harmonics in the operating speed range are the governing factors in determining the number of blades. Excitation of hull or propeller shafting resonance at blade frequency should be avoided.

Propeller vibrations. The propeller exciting forces may be of the once-per-revolution frequency or of blade frequency or its multiples. Once-per-revolution forces are produced by mechanical imbalance or inequalities of blade spacing, pitch, or contour, or by blade damage. Blade frequency forces are produced by the interaction of the propeller with the irregular flow pattern in which it operates and with the adjacent hull surface. Vertical and horizontal transverse forces acting at the stern excite vibration in the hull. Torque and thrust variations excite vibration in the propeller shaft and the machinery. A hull stern shape conducive to more uniform flow into the propeller and generous clearance between the propeller and the hull are desirable to reduce vibratory effects.

Under special circumstances, propeller blade vibration can occur that produces an audible noise ("singing propeller"). In most instances, this can be prevented by the application of special chisel-shaped edges to provide fixed flow separation points along the trailing edge of the propeller blades. *See* SHIP POWERING, MANEUVERING, AND SEAKEEPING.

Propeller cavitation. The formation and collapse of vapor-filled bubbles, or cavities, causes noise, vibration, and often rapid erosion of the propeller blade material, especially in fast, high-powered vessels. This phenomenon is known as cavitation. As long as the rotational and translational speeds of the

propeller are not too high, the onset of cavitation can be delayed or limited to an acceptable amount by utilization of special blade sections. The application of circulation theory in design is of great help, but it is highly advisable to have a model of the propeller tested in a variable-pressure water tunnel, water channel, or vacuum tank under simulated operating conditions and at the proper cavitation index to assure good performance. *See* WATER TUNNEL (RESEARCH AND TESTING).

To minimize the effects of cavitation, propellers are usually made of a corrosion-resistant zinc-free nickel aluminum bronze. Certain stainless steel alloys are also finding increasing application for heavy-duty propellers. *See* CAVITATION.

Supercavitating and superventilated propellers. Supercavitating and superventilated propellers are designed to have fully developed blade cavities which spring from the leading edge of the blade, cover the entire back of the blade, and collapse well downstream of the blade trailing edge. The blade of such propellers has unique sections which usually are wedge-shaped with a sharp leading edge, blunt trailing edge, and concave face. Supercavitating and superventilating propellers are distinguished by the nature of the gases in the cavity. Supercavitating propellers have cavities filled with water vapor and small amounts of gases dissolved in the fluid medium. Superventilated propellers have cavities filled primarily with air or gases other than water vapor. They may be fully submerged propellers with a gas supply system through the propeller shaft and propeller blades into the cavity, or partially submerged propellers drawing air from the water surface as the blade enters the water.

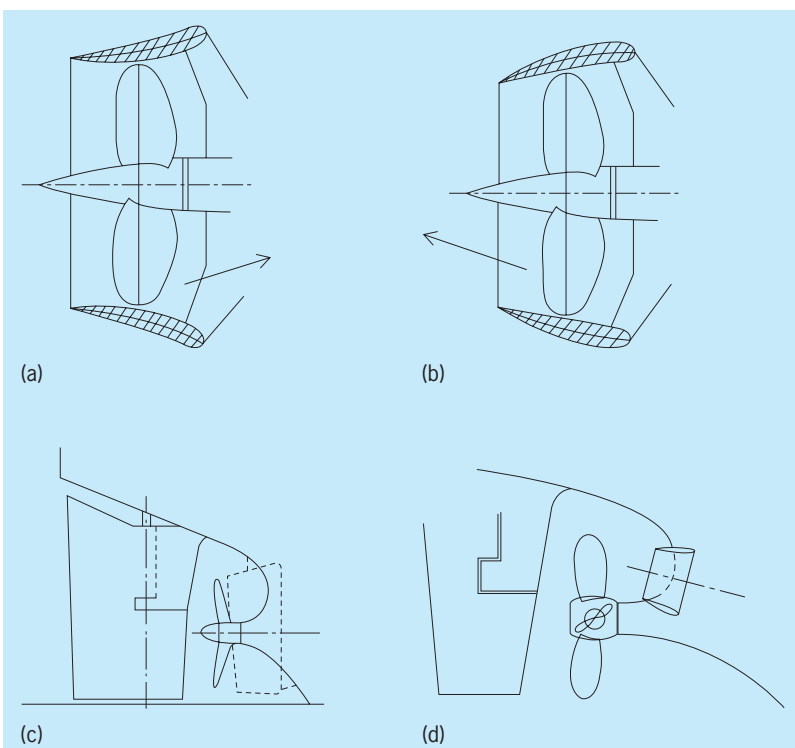

Fig. 2. Ducted propeller arrangements. (*a*) Accelerating or Kort nozzle. (*b*) Decelerating nozzle or pumpjet. (*c*) Mitsui duct ahead of propeller. (*d*) Wake-equalizing duct.

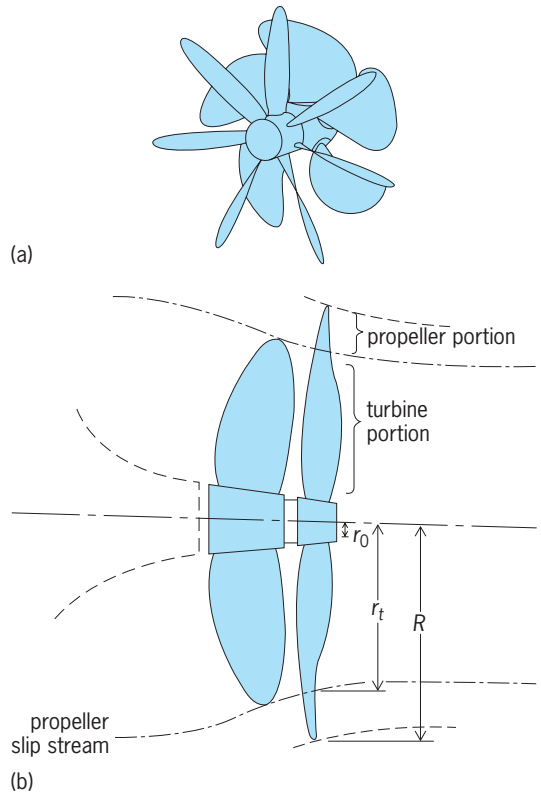

Fig. 3. Typical propeller with vane wheel. (*a*) Oblique view. (*b*) Side view.

Both of these propeller types are intended for use in high-speed craft (speeds usually greater than 45 knots or 83 km/h) such as hydrofoil boats, surface-effect ships, and the higher-speed planing craft with high propeller-shaft rotational speeds. The application of supercavitating propellers has been very limited. They were applied to a number of high-speed hydrofoil patrol vessels. *See* AIR-CUSHION VEHICLE; HYDROFOIL CRAFT.

The partially submerged propeller is widely used in racing hydroplanes (prop-riders). It has evolved by trial and error, and the blade sections are shaped similar to those developed for the supercavitating propeller. They are very efficient two-bladed propellers with thicker sections than would be needed on a superventilating propeller to sustain the high fatigue loads. Speeds up to 175 knots (320 km/h) have been attained. Interest in the partially submerged propeller is being extended to other vessels, such as surface-effect ships, using more blades and blade sections of the type that were developed for the superventilating propeller.

Controllable-pitch propellers. For ships which normally operate at widely varying speeds and propeller loadings (for example, towboats, rescue vessels, trawlers, and ferryboats), the application of controllable-pitch (rotatable-blade) propellers permits the use of full engine power at rated rpm under all operational conditions, ensuring maximum thrust production and utmost maneuverability. Since these

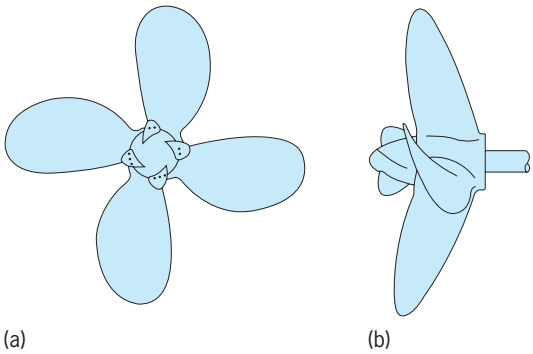

Fig. 4. Propeller with boss cap fins. (*a*) Front view. (*b*) Side view.

propellers are also reversible, they permit the use of nonreversible machinery (such as gas turbines). The hydraulic or electric servomotor used to adjust the pitch of the blades requires a hollow tailshaft for its operation. The propeller pitch can be controlled directly from the ship's bridge. In each case the operational advantages of the controllable-pitch propeller must be weighed against the disadvantages of more complex construction, higher manufacturing cost, and the possibility of a small decrease in efficiency due to the larger diameter of the propeller hub and the exposed shafting.

Ducted propellers. It was found in the early 1930s on river towboats that the tow force could be increased up to 40% by enclosing the propeller in a nonrotating, coaxially converging streamlined nozzle which accelerated the flow into the propeller. With time this has spawned a number of applications which use a duct in conjunction with the propeller to improve the propulsive performance.

Accelerating duct (Kort nozzle). This coaxial nozzle around the propeller, named after its inventor, is shaped such that it accelerates the flow into the propeller (**Fig. 2***a*). At low speeds and heavy propeller loadings, the accelerated flow increases the propeller thrust, with the duct itself providing an additional component of thrust. Research has shown that the Kort nozzle has wide application whenever the propeller is heavily loaded. It has been installed on all types and sizes of ships, from small coastal fishing trawlers, ice breakers, and tugboats, to very large crude-oil tankers. The Kort nozzle in a tunneled stern is widely used on powerful river towboats of modern design.

Decelerating duct. This coaxial nozzle around the propeller is shaped such that it decelerates the flow into the propeller (Fig. 2*b*). Decelerating the flow delays the onset of cavitation, thus reducing the radiated noise generated by the propeller. It is intended for use on vessels that need to minimize radiated noise at as high a speed as possible in order to accomplish their mission. These are primarily military vessels engaged in stealth operations and antisubmarine warfare. The propulsive efficiency of this type of duct is lower than that of a conventional propeller.

Mitsui duct. This coaxial nozzle is placed a short distance forward of the propeller so that the propeller no longer operates within the nozzle, thus substantially reducing the complexity of installation (Fig. 2*c*). The duct is shaped to accelerate the flow into the propeller. However, it does not increase the propeller efficiency to the same extent as the Kort nozzle, nor does the nozzle produce as much thrust. In modifying the flow around the stern of full-formed vessels such as tankers and bulk carriers, it tends to reduce the hull resistance. Although the gains in propulsive efficiency are smaller than from an accelerating nozzle, the benefits are such that many large tankers have been converted to this form of propeller propulsion.

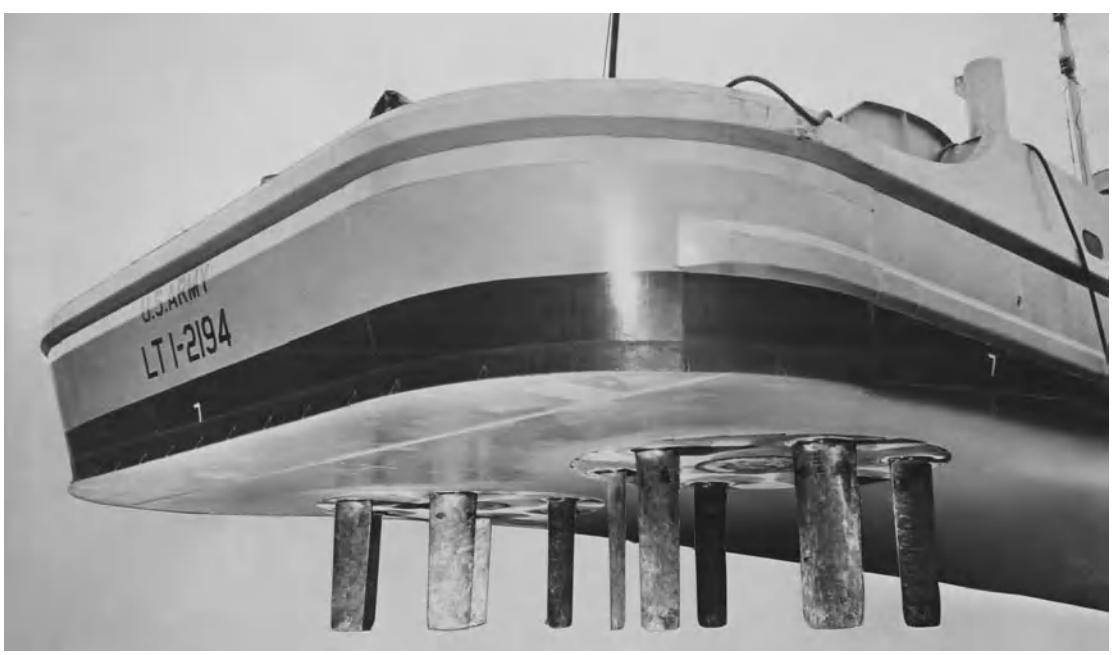

Fig. 5. Twin cycloidal propellers which are rated at 1100 hp (800 kW) each on a U.S. Army towboat. The length of the spadelike vertical blades is 4.5 ft (1.4 m). (*Pacific Car and Foundry Co.*)

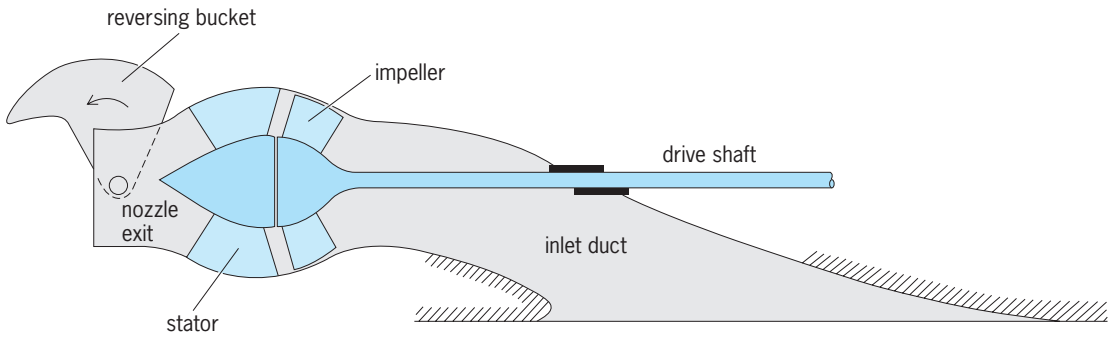

Fig. 6. Typical waterjet arrangement.

Wake equalizing duct (WED). This nozzle consists of two semicircular ducts located on the port and starboard sides of the hull forward of the propeller in the upper part of the propeller disc (Fig. 2*d*). The nozzle is shaped to accelerate the flow into the upper part of the propeller disc where the velocity behind the hull is lowest. This tends to make the total flow into the propeller more uniform, thereby reducing the magnitude of the propeller-induced vibratory forces and increasing propeller efficiency. Like the Mitsui duct, the WED can modify the flow around the stern of full-formed ships in the region of the duct and reduce the hull resistance a small amount. The WED is a relatively inexpensive device which can be applied during or after ship construction with prospects of a short amortization period. It is being applied to a wide variety of single-screw commercial ships.

Contrarotating propellers. Since the rotational energy in the propeller race represents an energy loss, the idea of recovering the lost energy is attractive and has led to the development of contrarotating propellers working on coaxial, contraturning shafts. If the propellers are properly designed, there are minimum rotational losses in the race behind the second propeller. Such propellers have been used for many years on torpedoes, achieving high efficiency, balancing the reaction torque, and preventing the body from rotating. The need to conserve energy has revived interest in this form of propulsion for commercial vessels. Installations on Japanese vessels have shown reductions in shaft horsepower as great as 15%. Although the gains in propulsive efficiency are great, the increased complexity and capital cost of the propulsion system have inhibited its general application.

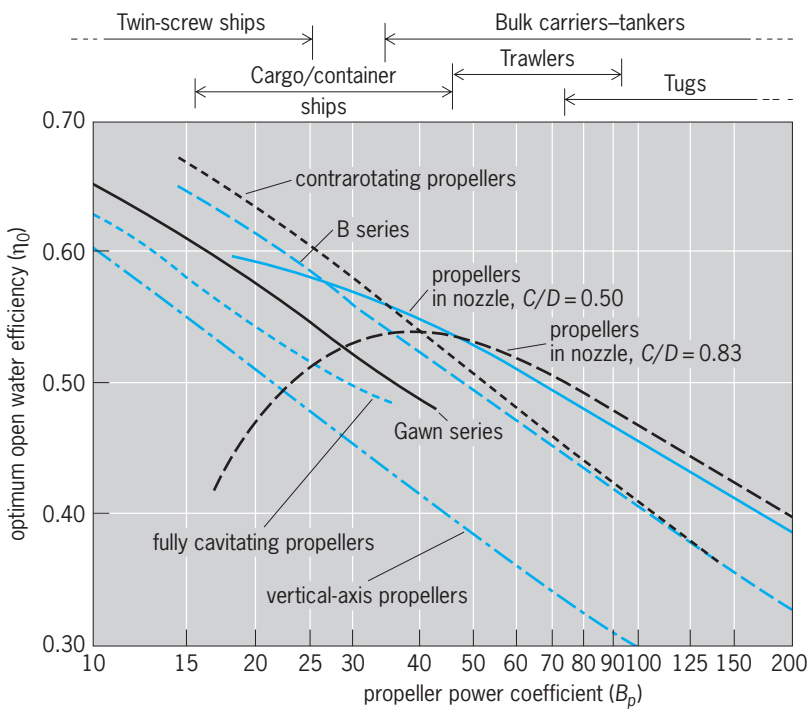

Fig. 7. Typical optimum open-water efficiencies for different propeller types. C = length of duct; D = diameter of propeller. (*After J. D. van Manen, The choice of the propeller, Mar. Technol., April 1966*)

Propeller and vane wheel. Another approach to recovering some of the rotational energy loss in the propeller race utilizes a freely rotating vane wheel mounted on the same shaft but not mechanically connected to the propeller (**Fig. 3**). The vane wheel is somewhat larger in diameter than the propeller; thus the part of the vane wheel inside the slipstream acts as a turbine and the part outside as a propeller producing additional thrust. The vane wheel usually rotates at about 40% of the speed of rotation of the propeller and generally consists of seven or more relatively narrow blades. Application has been primarily to single-screw commercial vessels with gains in thrust of 5–10% reported.

Propeller boss cap fins (PBCFs). This device is designed to recover some of the rotational energy loss generated at the root of the propeller blades (**Fig. 4**). It consists of fins located on the propeller hub cap equal to the number of propeller blades. In addition to a small increase in propulsive efficiency, the PBCF reduces the strength of the hub vortex, thereby reducing the likelihood of cavitation-induced erosion on the leading edge of the rudder when installed behind the propeller shaft. Gains in propulsive efficiency of 2–8% have been reported. PBCFs have been installed primarily on commercial ships.

Pod propulsion. Pod propulsion is a specialized application of propeller propulsion in which the propeller is mounted on a pod which, in turn, is attached to the hull of the vessel by means of a strut. Marine propellers are usually located at the stern of the vessel and are driven by line shafting directly connected to the main engine or its reduction gears. In the pod

configuration, the propeller is mounted on the trailing or leading end of the pod; in the former case it is usually called a pusher and in the latter a tractor. If the vessel has electric propulsion, the electric motor can be mounted in the pod and connected directly to the propeller. Units with electric motors up to 15 kW are now being manufactured. The pod may be either fixed or trainable in azimuth, thus eliminating the need for a rudder. In the latter configuration the pod propeller provides a very high degree of maneuverability to the vessel. In applications in which the propeller is heavily loaded, a Kort nozzle is added to the configuration.

Cycloidal propeller. This type of propulsor rotates about a vertical axis with the blades rotating in a cycloidal path. The cycloidal propeller was introduced in the 1920s by Kirsten-Boeing and has evolved into the Voith-Schneider design. It consists of a circular disk set flush in the vessel's bottom, carrying near its periphery a number (usually six or eight) of spadelike rotating vertical blades which undergo a cycloidal movement about the vertical axes (**Fig. 5**). Depending on the phasing of the cycloidal motion, the propeller's thrust can be directed in any direction, thus providing unusual flexibility and maneuverability at low ship speed without need for a rudder. However, the vulnerability of the blades to damage and the complicated mechanism limit application to special inland-waterways craft.

Waterjet propulsion. This type of propulsor is rapidly finding use in special types of craft (**Fig. 6**). It is a reaction propulsor in contrast to the screw propeller and its derivatives, which derive their thrust as lifting devices. In operation, water is drawn from under the craft through an inlet duct by an internal pump, usually an axial-flow pump, where energy is added to the water, which is then expelled at higher velocity through an outlet nozzle in the stern. The thrust is generated by the increase in momentum supplied to the water.

Comparative performance. The performance of some of the propulsors discussed above in comparison to conventional commercial propellers is dependent upon the particular application, how the propulsor interacts with the hull, and how it responds to the flow field created by the hull. Nominal comparisons can be made if considreation is limited to the propulsor efficiency as a singular entity, ignoring the interaction with the hull. Typical comparative propeller efficiencies of the identified propulsors as a function of the propeller power coefficient, B_P, are shown in **Fig. 7**. This coefficient is defined in Eq. (4),

$$B_P = P^{1/2}N/V_A^{2.5} \tag{4}$$

where B_P is the power coefficient, P is the power (ft-lb/s), N is the speed of the propeller (rmp), and V_A is the speed of advance of the propeller (ft/s). The line identified as "B series" in the figure is the conventional merchant ship propeller, in this case having four blades. The line identified as "Gawn series" is a special three-bladed series intended for high-speed vessels.

The supercavitating, superventilating, partially submerged propellers, and the waterjet are omitted from this figure, as their comparitive performance is more a function of speed than of propeller loading. They can outperform the conventional propeller at speeds exceeding 30–35 knots (55–65 km/h). *See* MARINE ENGINE; MARINE MACHINERY.

Jacques B. Hadler

Bibliography. R. A. Cumming, W. B. Morgan, and R. J. Boswell, Highly skewed propellers, *Soc. Nav. Arch. Mar. Eng. Trans.*, 80:98–135, 1972; E. V. Lewis (ed.), *Principles of Naval Architecture*, Society of Naval Architects and Marine Engineers, vol. II, pp. 164–240, 1988; E. R. May, *The Theory and Practice of Controllable Pitch Propellers*, 1979; F. R. Young, *Cavitation*, 1989.

Proprioception

The sense of position and movement of the limbs and the sense of muscular tension. The awareness of the orientation of the body in space and the direction, extent, and rate of movement of the limbs depend in part upon information derived from sensory receptors in the joints, tendons, and muscles. Information from these receptors, called proprioceptors, is normally integrated with that arising from vestibular receptors (which signal gravitational acceleration and changes in velocity of movements of the head), as well as from visual, auditory, and tactile receptors. Sensory information from certain proprioceptors, particularly those in muscles and tendons, need not reach consciousness, but can be used by the motor system as feedback to guide postural adjustments and control of well-practiced or semiautomatic movements such as those involved in walking. *See* MOTOR SYSTEMS.

Sensory discriminations. Position and movement of limbs can be sensed either when the muscles are relaxed and movement is passive or during active muscular contraction (the latter sense called kinesthesis). In the former case, the threshold for sensing small changes in the angle of the joints, as measured experimentally, depends upon the extent and velocity of the change in angle and is considerably better at proximal joints such as hip or shoulder joints than at more distal joints such as those in the fingers. For example, thresholds vary roughly between 0.2° for the hip to 0.7° for the main joint of the big toe when the speed of displacement is 10°/min. If passive movements of a joint are superimposed, by the experimenter, on active movements of the same joint, discriminatory capacity is improved. Human observers can also identify precisely the absolute position of a limb (placed either actively by the subject or passively by an experimenter) by actively matching the position with the corresponding limb on the opposite side of the body. Humans are also accurate in judging the amount of force exerted by their muscles in making a movement or in maintaining the position of a joint against an opposing force. The latter capacity can be assessed from discriminations between the

active forces required to compress springs of different stiffness held between the thumb and finger. In fact, discriminations analagous to these can be made to a lesser degree without any information from proprioceptors, indicating that there is a "sense of effort" by which the motor system sends a copy of information sent to the muscles to higher sensory processing centers. In normal conditions such central feedback enables the nervous system to determine which proprioceptive information resulted from intended movements and which came from stimuli externally applied.

Proprioceptors. Receptors for proprioception are the endings of peripheral nerve fibers within the capsule or ligaments of the joints or within muscle. These endings are associated with specialized end organs such as Pacinian corpuscles, Ruffini's cylinders, and Golgi organs (the latter resembling histologic Golgi structures in the skin), and muscle spindles. *See* CUTANEOUS SENSATION.

Pacinian corpuscles are found mainly in the connective tissue surrounding the joint and are sensitive to vibrations or abrupt changes in position of the joint. Joint receptors, such as the Ruffini and Golgi types, within the capsule are compressed or stretched when the joint moves, and signal the direction and velocity of movement as well as the steady position of the joint. There are also Golgi organs located at the junction between a muscle and its tendon. When the muscle is lengthened and the ends of a muscle are fixed, then active muscle contraction will shorten the contractile part of the muscle, thereby stretching the inelastic tendon fascicles and exciting the Golgi receptors. Thus, these receptors are excited by muscle tension or force.

Two other stretch receptors are located within the muscle spindle, the latter consisting of small (intrafusal) muscle fibers attached at each end to the main (extrafusal) fibers of skeletal muscle. Each spindle is supplied with two types of sensory nerve endings. One type is excited mainly by changes in the length of the muscle, while the other is sensitive not only to changes in length but also to the velocity of muscle elongation. In addition, the muscle spindle is supplied with axons from neurons called gamma motoneurons, excitation of which causes the intrafusal fibers to contract. Whenever the intrafusal fibers are contracted more than the extrafusal fibers, the sensory endings of the spindle are excited. Stretch receptors are included in many reflexes involving the spinal cord and higher centers in the central nervous system. These reflexes have roles in the control of movement and posture. For example, when a static muscular contraction (such as that occurring in the legs during standing) is subjected to an added load, muscle spindles are stretched, resulting in a reflexive increase in contraction of extrafusal fibers (and relaxation of intrafusal fibers) to maintain the load. Activity in the gamma motoneurons then serves to readjust the length of the muscle spindle (via contraction of the intrafusal fibers) to its previous value in order to maintain the muscle spindle's sensitivity to subsequent varying conditions.

One or more types of muscle receptors may also contribute to the sense of movement and position of the joints since this sense is still possible, although impaired, following local anesthetization of joint receptors or after replacement of the joint with a prosthesis. Conversely, the sense of movement and position also remains, albeit impaired, after elimination of sensory signals from muscle receptors.

Cutaneous receptors may provide some relevant information (particularly in signaling that limb movement has occurred), although the results of experiments in which the skin is anesthetized indicate that this information is not essential. Thus, the sense of movement and position of the limbs probably depends upon the central integration of information from several different types of proprioceptors. *See* NERVOUS SYSTEM (VERTEBRATE); SENSATION; SOMESTHESIS.

Robert Lamotte

Bibliography. E. Autrum et al. (eds.), *Progress in Sensory Physiology*, vol. 9, 1989; S. I. Fox, *Human Physiology*, 7th ed., 2001; R. F. Schmidt (ed.), *Fundamentals of Sensory Physiology*, 3d ed., 1986.

Propulsion

The process of causing a body to move by exerting a force against it. Propulsion is based on the reaction principle, stated qualitatively in Newton's third law, that for every action there is an equal and opposite reaction. A quantitative description of the propulsive force exerted on a body is given by Newton's second law, which states that the force applied to any body is equal to the rate of change of momentum of that body, and is exerted in the same direction as the momentum change. *See* NEWTON'S LAWS OF MOTION.

In the case of a vehicle moving in a fluid medium, such as an airplane or a ship, the required change in momentum is generally produced by changing the velocity of the fluid (air or water) passing through the propulsion device or engine. In other cases, such as that of a rocket-propelled vehicle, the propulsion system must be capable of operating without the presence of a fluid medium; that is, it must be able to operate in the vacuum of space. The required momentum change is then produced by using up some of the propulsive device's own mass, called the propellant. *See* AERODYNAMIC FORCE; AIRFOIL; FLIGHT CHARACTERISTICS; FLUID MECHANICS; PROPELLANT.

Propulsion principle. Any change in momentum, according to Newton's second law, must be exactly equal to the propulsive force exerted on the body. For example, one may examine the simple propulsive means consisting of a hollow sphere with a small hole of area A (**Fig. 1**). If the sphere is filled with gas whose pressure is higher than that of the surrounding medium, the gas will exert pressure on the inside surface of the hollow sphere, with the exception of the hole, through which gas will rush out into the surroundings. Directly opposite the hole is an area A of the sphere on which the compressed gas does

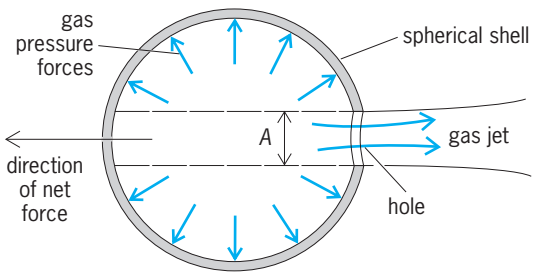

Fig. 1. A shell containing an aperture and saturated with gas at higher pressure than any outside fluid illustrates the principle of propulsion.

exert a force. In all other radial directions inside the sphere, pressure forces balance each other. Consequently, the net force on the sphere is in the direction shown. To compute the size of the unbalanced force, one can either measure and add or subtract all the pressures at the shell surface (including the gas pressure in the hole), or, more simply, in accordance with Newton's second law, calculate the momentum of the escaping gas jet.

This simple propulsive device shows that a rocket gives more thrust in space than in the atmosphere, because atmospheric pressure tends to increase the pressure in the hole, reducing the net force in the forward direction and, of course, reducing by an equal amount the effluent gas momentum. Also, it shows that although the amount of net thrust can be calculated by using the momentum change of the fluid medium, the thrust itself is produced, as it must be, by something pushing against the body— this case, the pressure exerted by a gas. *See* ROCKET PROPULSION.

Most propulsion engines are more complicated than the simple spherical rocket discussed here. Consider, for example, an airplane in level flight at constant speed (**Fig. 2**). Within a rectangular region (the control volume) around the airplane, net forces must be in equilibrium. The sum of all momentum changes occurring across the surfaces of the control volume is balanced by the total of all the forces on the airplane, in accordance with Newton's law.

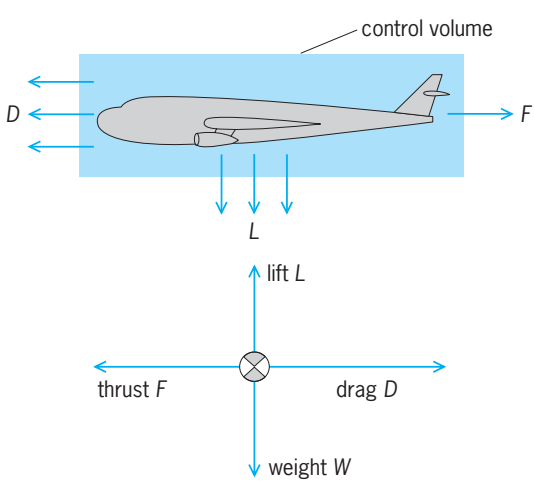

Fig. 2. All forces, including momentum forces, within a control volume about a vehicle are in equilibrium.

Some air passing over the wings is given downward momentum, resulting in upward reaction forces, called lift, on the airplane. This air is also given some forward momentum as it is accelerated downward, resulting in a rearward force on the airplane called induced drag. Viscous drag is also produced (in the rearward direction) because the airplane carries some air forward with it through viscous adherence, giving the air forward momentum. Finally, the airplane propulsive means (such as an engine-driven propeller or a turbojet engine) takes in quiescent air, which is thereby accelerated to the same speed as the airplane, providing a rearward force called ram drag, and then the propulsive means discharges the air at a much higher rearward velocity (with high rearward momentum), resulting in a forward force called gross thrust. If all these forces are balanced, the airplane flies at constant speed in a straight line; if not, the airplane is accelerated in whichever direction the forces are unbalanced, thereby providing another momentum change to again bring net forces within the control volume into equilibrium.

Thrust. Propulsion capability is measured in terms of the thrust delivered to the vehicle. In general, the net thrust delivered by any propulsive means, neglecting the effect of ambient pressure, is given by Eq. (1), where F = net thrust in lb; V_e = velocity of

$$F = (W_a + W_f)V_e - W_a V \tag{1}$$

gas leaving the propulsive means relative to the vehicle in ft/s; V = flight velocity of the vehicle through still air in ft/s; W_a = mass flow rate of the ambient medium (air or water) through the propulsive means in (lb-s^2/s)/ft; and W_f = mass flow rate of fuel (or on-board propellant) through the propulsive means in (lb-s^2/s)/ft.

The first term is the gross thrust; the second is the ram drag. In a rocket engine, no ambient fluid (air or water) is taken into the propulsive means; thus there is no ram drag. The net thrust of a rocket in vacuum is thus given by Eq. (2), where W_f represents the total

$$F = W_f V_e \tag{2}$$

mass flow rate of on-board propellant. *See* TURBINE ENGINE SUBSYSTEMS.

Efficiency. Propulsion efficiency was formerly defined as the useful thrust power divided by the propulsion power developed. Unfortunately, because useful thrust power can only be determined as the thrust multiplied by the flight velocity (FV), and because flight velocity is basically a relative term rather than an absolute term, propulsion efficiency has little real meaning as a measure of propulsion effectiveness.

The two terms most generally used to describe propulsion efficiency are thrust specific fuel consumption (SFC) for engines using the ambient fluid (air or water), and specific impulse (I_{sp}) for engines which carry all propulsive media on board. Thrust

specific fuel consumption is givenby Eq. (3), where

$$\mathrm{SFC} = \frac{W_f g_0}{F} = g_0 \frac{W_f}{W_a} \frac{1}{\left(1 + \dfrac{W_f}{W_a}\right) V_e - V}$$

$$\approx g_0 \frac{W_f}{W_a} \frac{1}{V_e - V} \tag{3}$$

SFC is expressed in pounds of fuel per second per pound of thrust, g_0 is acceleration due to gravity in ft/s^2, and, usually $W_f/W_a < 1$.

Specific impulse is determined similarly (although inversely) from Eq. (4). Obviously, effective propul-

$$I_{\mathrm{sp}} = \frac{F}{W_f g_0} = \frac{V_e}{g_0} \tag{4}$$

sion performance is indicated by low values of SFC or high values of I_{sp}.

The energy source for most propulsion devices is the heat generated by the combustion of exothermic chemical mixtures composed of a fuel and an oxidizer. An air-breathing chemical propulsion system generally uses a hydrocarbon such as coal, oil, gasoline, or kerosine as the fuel, and atmospheric air as the oxidizer. A non-air-breathing engine, such as a rocket, almost always utilizes propellants that also provide the energy source by their own combustion. Here the choice is wider: Fuels may be hydrocarbons again, but may also be more efficient low-molecular-weight chemicals such as hydrazine (N_2H_4), ammonia (NH_3), or even liquid hydrogen itself. Oxidizers are usually liquid oxygen, nitrogen tetroxide (N_2O_5), or liquid fluorine.

Where nuclear energy is the source of propulsive power, the heat developed by nuclear fission in a reactor is transferred to a working fluid, which either passes through a turbine to drive the propulsive element such as a propeller, or serves as the propellant itself. Nuclear-powered ships and submarines are accepted forms of transportation. In the case of nuclear-powered vehicles, the concept of specific fuel consumption is no longer valid, because the loss in weight of the nuclear fuel is virtually zero. However, the specific impulse is still a meaningful measure of performance, inasmuch as it is based on thrust and propellant flow rate, not on mass consumption of fuel. *See* AIRPLANE; SHIP NUCLEAR PROPULSION; SHIP POWERING, MANEUVERING, AND SEAKEEPING; TURBINE PROPULSION. Jerry Grey

Prosobranchia

The largest and most diverse subclass of the molluscan class Gastropoda. The group includes mostly marine snails but with a few fresh-water and land genera, all retaining an anterior mantle cavity and internal evidence of torsion. Adult prosobranchs always retain the streptoneurous (twisted-commissure) condition of the central nervous system, with the commissures to the visceral ganglia in the characteristic figure-of-eight pattern. This pattern reflects the torsion through 180° during larval (or embryonic) development which has brought the mantle cavity to a position above the head and facing forward. This contrasts with conditions in the other two gastropod subclasses, Opisthobranchia and Pulmonata, in which the effects of torsion are reduced or obscured in adults by secondary processes of development and growth. *See* OPISTHOBRANCHIA; PULMONATA.

The diversity of functional morphology exhibited by the prosobranchs is not equaled by any comparable subclass in the entire animal kingdom. From two-gilled forms with symmetrical cardiac and renal structures (the "Diotocardia") which can be numbered among the most primitive of all living mollusks, evolution within prosobranch stocks has involved increasing asymmetry of pallial, cardiac, and renal systems and greater hydrodynamic efficiencies. Torsion and the anterior mantle cavity create locomotory, circulatory, sanitary, and hydraulic problems which have been solved in a variety of ways in different prosobranchs. Four orders are commonly recognized in subclass Prosobranchia: Archaeogastropoda, Neritacea, Mesogastropoda, and Neogastropoda.

Archaeogastropoda. All primitive two-gilled (or zygobranch) snails belong to the order Archaeogastropoda, as do all forms retaining a right functional kidney or other symmetric renal and cardiac structures. To deal with the problems of sanitation created by the mantle cavity opening anteriorly over the head, various shell slits (*Pleurotomaria*) or serial shell openings (*Haliotis*, abalones, and ormers) are developed. Other archaeogastropods, which have a single aspidobranch gill but still some risk of pallial clogging, include the acmaeid limpets and the elegant top-shells (Trochidae). All of these forms are ecologically limited to clean waters over hard substrata.

Neritacea. The Neritacea also possess a single aspidobranch gill, but have the functional kidney on the left side, thus allowing the genital system (on the right) to develop more elaborate secondary structures. The most primitive neritaceans are littoral snails, largely tropical, living on coral and similar rock surfaces. However, this group has evolved completely independently of (though occasionally parallel to) the rest of the gastropods. The morphological separation of renal function allows the development of a penis and of a uterus capable of retaining large shelled eggs, so that neritaceans have successfully invaded land and fresh waters (the family Helicinidae shows the independent evolution of a lung for air breathing). Neritacea show a patchy zoogeographic distribution; over one-third (about 120 species) of terrestrial neritaceans live in Jamaica, and about 70% of these are endemic to that island.

Mesogastropoda and Neogastropoda. In the two other prosobranch orders, Mesogastropoda and Neogastropoda, there is usually a one-sided comb-shaped or pectinibranch ctenidium or "half-gill" with its axis fused to the mantle wall, and always a simplified mantle cavity which is hydraulically efficient and

from which both feces and nitrogenous excretory wastes are discharged to the right-hand side of the snail. Such snails can live on, and even in, the softer substrata of the sea bottom, and their gill systems can operate in silty waters. There is also a functional separation of genital from renal ducts, allowing the development of internal fertilization, of large eggs, and even of viviparity.

Mesogastropoda. Among the mesogastropods, a wide variety of marine life-styles are found, from that of the slipper limpet, *Crepidula*, which filter-feeds by using the pectinibranch gill (with considerable elongation of individual ctenidial leaflets into filaments), to those of the naticid moon snails like *Polinices* which are predatory carnivores on clams. Another group, *Littorina* and its allies, are among the most successful of seashore animals, with some high-level species becoming physiologically almost terrestrial. Fresh-water mesogastropod genera include *Valvata*, which is unusual in retaining an aspidobranch gill, and *Viviparus*, in which females are ovoviviparous, producing "near-adult" young after a 10-month "gestation."

Neogastropoda. The most highly specialized prosobranch order is the Neogastropoda, comprising marine carnivores and carrion feeders. In addition to the efficient simplified gill and mantle cavity (now with a characteristic inhalant siphon with a chemoreceptive osphradium for detecting prey), there is a simplified gut and a flesh-tearing stenoglossan radula borne on a long extensible proboscis. The order is anatomically rather uniform despite a wide diversity of habitats and a large number of species, including many carnivores which specialize on a single prey species (such as some species of *Conus* which eat flatfish after paralyzing them with a neurotoxin). *See* NEOGASTROPODA.

Evolution. As a group, the prosobranchs constitute the most diverse and largest subclass in the whole phylum Mollusca. The most primitive stem stock of prosobranchs is represented by the extant genus *Pleurotomaria* as a living "archetype" with gills, heart auricles, kidneys, gonads, and pallial sense organs all as paired structures. This neontological statement based on functional anatomy is reinforced by the fossil record: the Pleurotomariacea date from the mid-Cambrian, while fossil siphonate genera (corresponding to the Neogastropoda) appear later and increase markedly in the Mesozoic and Cenozoic. *See* GASTROPODA; LIMPET; MOLLUSCA.

W. D. Russell-Hunter

Bibliography. V. Fretter and A. Graham, *British Prosobranch Molluscs*, 1962; W. D. Russell-Hunter, *A Life of Invertebrates*, 1979; K. M. Wilbur (ed.), *The Mollusca*, vols. 1–8, 1983, 1984.

Prospecting

Exploration for mineral deposits; also referred to as ore search. The result of prospecting is the discovery of potentially economic mineralization, that is, the prospect. Mineral exploration also continues beyond prospecting to include the delineation and evaluation of the prospect to determine its minability as an orebody or economic mineral deposit. A successful prospect is developed into a mine. *See* MINING.

Prospecting generally pertains to the search for deposits of metallic ore minerals, but it also includes the search for nonmetallic or industrial minerals and rocks such as sulfur, potash, and limestone, and mineral fuels such as petroleum, coal, and oil shale.

Prospecting may be carried out by individuals, but is generally done by well-supported multidisciplinary teams or groups of geologists and specialists working for mining and exploration companies or for government mining organizations.

In a successful program of regional prospecting, a dozen or more target areas may be chosen for the final stages of investigation, including drilling. A complete regional prospecting program is expected to take several years, and it requires an expenditure of several million dollars.

With much of the Earth's readily accessible surface having been investigated for minerals, prospecting is increasingly directed toward the discovery of deeper mineralization in recognized mining districts; mineralization hidden beneath overlying rocks, sediments, and soils; and mineralization in the less-known jungle, arctic, and offshore parts of the world. *See* MARINE MINING; OIL AND GAS, OFFSHORE.

Characteristics

The traditional approach to prospecting involves an examination of surface exposures for direct geologic evidence of ore mineralization and a subsequent examination of the subsurface by pitting, drilling, or underground prospect workings. The traditional approach commonly has become integrated with less direct methods that involve projections and interpretations from geologic mapping as well as geochemical, geobotanical, remote-sensing, and geophysical surveys for finding anomalous or extraordinary conditions that can be evaluated as evidence of underlying mineralization.

Much of modern prospecting is done in a regional exploration sequence of programmed stages. The early or strategic stages in regional exploration constitute broad geologic studies and field reconnaissance to identify specifically favorable subregions or districts. Additional reconnaissance takes place in the more favorable terrain, and follow-up or tactical stage prospecting is centered in the target areas recognized as having a high mineral potential. Specific exploration targets are identified; more detailed surface and subsurface prospecting work then provides for the actual discovery of ore mineralization and the designation of a prospect.

Some prospecting programs may take place entirely within a recognized target area, a mining district where current or past mining operations have already demonstrated a geologic potential for additional deep or hidden ore mineralization. This kind of program calls for intensive subsurface investigation through geologic mapping, geochemical sampling,

and geophysical surveys in existing mines, accompanied by deep drilling from surface or underground sites, and the driving or sinking of new underground workings.

A central element in all prospecting programs and activities is the acquisition of mineral rights, some form of concession or ownership that will provide for the eventual mining of a discovered orebody. Most governments consider the stagewise sequence in regional exploration and thus provide reconnaissance permits or hunting permits, exploration licenses that convey preference rights within a defined area, and ultimate mining concessions. In the United States, many minerals on federal lands are subject to "location" by the staking of mining claims; other minerals such as coal, petroleum, and certain leasable industrial minerals on federal lands, offshore, and on most state lands are subject to prospecting permits and to mining lease arrangement by competitive bidding. Mining claims that have been patented—that is, purchased from the government—and privately owned mineral rights on older private land in the United States are obtained by direct lease or purchase.

Geologic Guides

Prospecting is carried out on the basis of the guides to ore associated with a conceptual image of the anticipated orebody. The image is referred to as an exploration model, and it is drawn from the characteristics of known orebodies in similar terrain. The exploration model and its guides to ore are expressed in terms of the regional and local geologic pattern; it has a certain diagnostic mineralogical character, it will commonly have a halo or envelope of associated guide minerals, and it will be expected to have a recognizable geochemical and geophysical expression. *See* ORE AND MINERAL DEPOSITS.

An exploration model actually has two appearances, one being the fundamental subsurface appearance of primary mineralization at depths below the influence of weathering, and the other being the near-surface weathered derivative. Except in recently glaciated or actively eroding terrain with a shallow zone of oxidation, it is the weathered derivative that will appear in geologic evidence. Where the targeted mineralization relates to a concealed or blind orebody, interpreted geophysical and geochemical anomalies provide signatures representing the geological condition at depth.

Metallogenic associations. The broadest geologic guidelines are applied to an exploration model during the reconnaissance stage in prospecting; they deal with the probable location of an undiscovered mineral deposit relative to regional trends of known and related deposits. Where a belt or an area of lead-zinc or copper-molybdenum deposits, for example, is of a certain geologic age or is associated with a major geologic feature such as the edge of an ancient depositional basin or a zone of tectonic disturbance, a metallogenic trend is indicated. Spatial projections of the trend or other comparable geologic trends in linear zones may be selected as favorable terrain for prospecting.

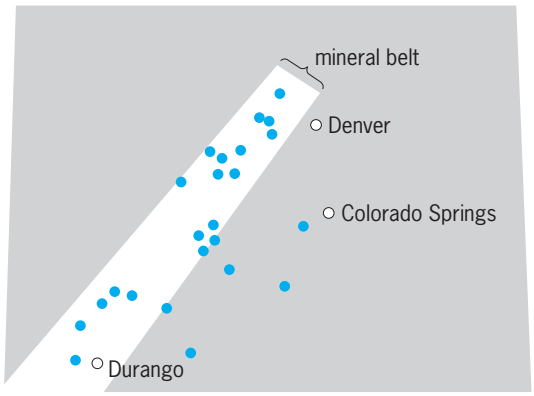

Fig. 1. Major metal mines (color dots) in the Colorado Mineral Belt, constituting a metallogenic guide in regional prospecting.

The largest of metallogenic trends are along the edges of lithospheric plates where interactions between continental and oceanic crustal segments have provided conditions for the genesis of ore deposits; the copper, iron, and tin belts of the Central Andes are of this kind. More localized metallogenic trends involved in prospecting are represented by the Colorado Mineral Belt (**Fig. 1**), a prominent zone of associated mines and prophyry intrusions; by the Mother Lode of California, a fault zone extending along the west side of the Sierra Nevada; and by the Carlin Trend, an alignment of gold deposits in northeastern Nevada. *See* PLATE TECTONICS.

In further defining favorable areas during prospecting, characteristic associations between certain rock types, geologic ages, and mineral deposits are used as guides. The association may relate to syngenetic ore deposition and processes of volcanism, sedimentation, or igneous intrusion, as with Precambrian sequences of banded iron formation in many parts of the world, the copper-bearing Precambrian ore shale of Zambia, and the layers of ultrabasic rock and chromium ore in the Precambrian Bushveld Complex of South Africa. An association may also pertain to the occurrence of epigenetic orebodies in certain preexisting host rocks, as with the Mississippi Valley–type lead-zinc ore in recognized lithologic sequences of limestones and dolomites; or it may pertain to spatially associated rocks such as the granitic intrusions that accompany many tin, copper, tungsten, and molybdenum deposits. Associations between rocks and mineral deposits have been specified in artificial intelligence models such as the Prospector expert system, and used in assessing the probability of new ore deposits in chosen areas. *See* EXPERT SYSTEMS.

Structural associations. The relation of ore deposits to structural features such as faults, brecciated areas, unconformities, and the borders of intrusions is used as a guide in prospecting. Mineralizing solutions will have commonly migrated along fractures and will have deposited ore in veins, breccia pipes (**Fig. 2**), stockworks, and contact zones in older

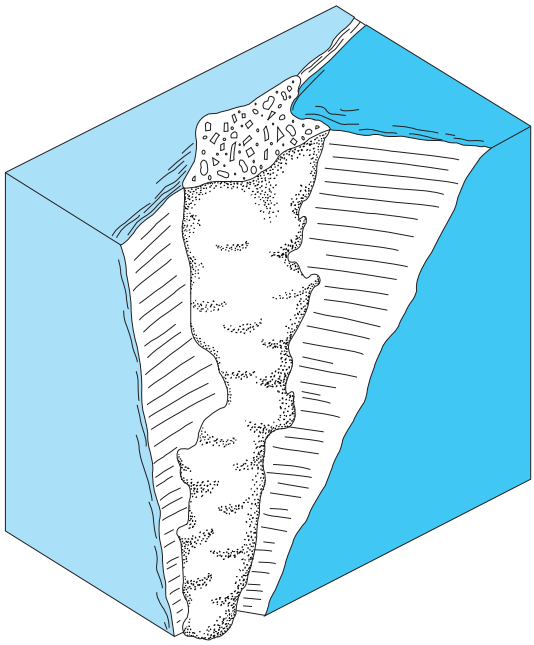

Fig. 2. Breccia pipe orebody at a fault intersection, representing a structural association and a guide in prospecting.

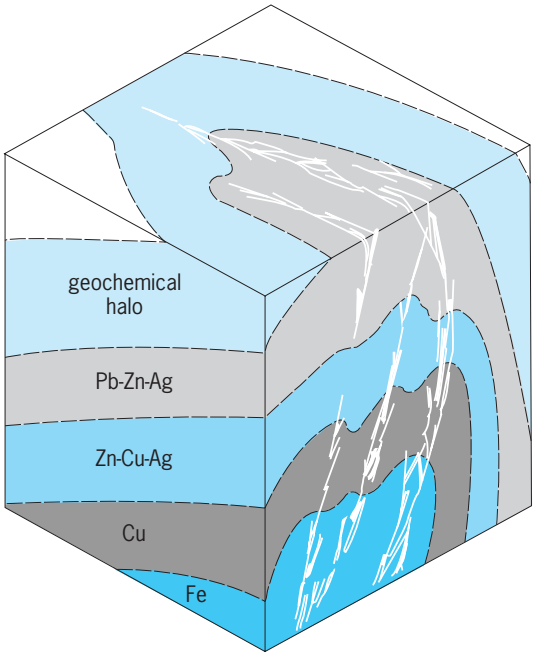

Fig. 3. District metal zonation pattern in a vein complex, constituting a geologic guide to prospecting for deeper ore.

rocks or at exhalative centers on an ancient sea floor. Thus, the overall structural pattern in a region is studied during reconnaissance, and significant centers of volcanism and structural deformation are considered as target areas for follow-up prospecting. During the more detailed stage of prospecting, weakly mineralized or geochemically anomalous structural features and their geophysical patterns are traced toward potentially favorable intersections and lithologic contacts. *See* ORE AND MINERAL DEPOSITS; STRUCTURAL GEOLOGY.

Zoning. One closely followed guide to ore is the expected pattern of mineral zoning. Ore and gangue minerals change in character with distance from the center of mineralization, and there is commonly an additional outer halo or envelope of alteration minerals and trace amounts of index elements in the enclosing rock. The entire zonal sequence may occur within a distance of a few meters, or it may involve groups of orebodies in a system that extends for several kilometers. Regional zoning may extend even farther, and it may incorporate a series of orebodies that are in themselves zoned.

A single mineral deposit or a group of closely associated orebodies with a central iron (Fe)-rich zone of pyrite, magnetite, and quartz may be surrounded by a more or less complete copper-iron (Cu-Fe) zone of chalcopyrite, pyrite, and quartz and by an outer zinc-copper-silver (Zn-Cu-Ag) zone characterized by sphalerite, tetrahedrite, and ankerite (**Fig. 3**). In addition, there may be a still more outlying zone of scattered lead-zinc-silver (Pb-Zn-Ag) mineralization with galena and sphalerite as the principal ore minerals and with calcite as the principal gangue mineral. The host rock beyond the outermost lead-silver (Pb-Ag) zone may have a geochemical signature; and it may have a bleached or discolored appearance due to the presence of alteration minerals such as mica and chlorite that can be identified by microscopic study or laboratory analysis. The outlying ore, gangue, and alteration minerals may also have a recognizable geophysical signature. In prospecting, the expected zonal pattern of target rings in an exploration model is used in relating information from widely dispersed outcrops to a more complete system of mineral deposits that may still be hidden.

Outcrops. Gossan, the residue of red, brown, and yellow iron oxides and silica that remains from the weathering and near-surface leaching of sulfide ore minerals, is used in prospecting as a guide to hidden primary mineralization (**Fig. 4**). The color and textural characteristics of gossan are related to the minerals in the orebody from which they were derived; outcrops of gossan are therefore examined for indications of the nature of the underlying deposit.

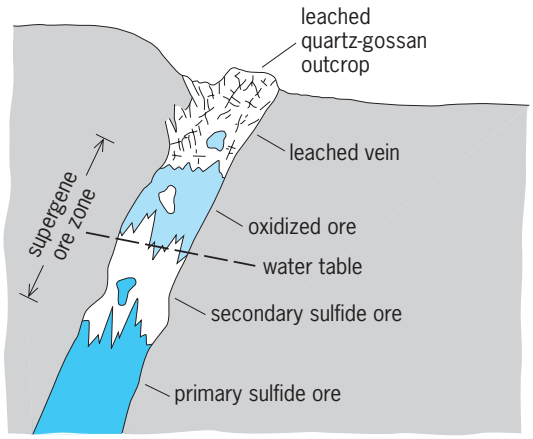

Fig. 4. Leached, supergene, and primary ore in an outcropping copper sulfide–quartz vein.

Due to the weathering of other iron-bearing minerals and due to the general complexity of weathering processes, it is often hard to discriminate between true gossan and other forms of limonitic iron, and it may be difficult to make a precise evaluation of gossan. Even so, the appearance of a cellular and heavily iron-stained outcrop, sometimes with additional staining in green or blue from oxidized copper minerals, black from manganese, or pink from cobalt, is an important factor in selecting targets for details prospecting. *See* WEATHERING PROCESSES.

In some gossans, secondary accumulations of gold and silver minerals may actually be higher-grade than the underlying primary deposit. As a result, some exposures of gossan that were mined for precious metals in earlier times only down to the primary sulfide mineralization can be evaluated again as guides to currently economic ore deposits. Oxidized and secondarily enriched (supergene) ores of copper, lead, nickel, and zinc also accumulate at shallow depth in the zone of weathering; a characteristic gossan remains at the surface.

Prospecting Methods

Both direct and indirect methods of prospecting make use of an appropriate exploration model and its guides to ore. Conventional or direct methods involve the recognition of the guides to ore in their surface expression. Indirect methods make use of geochemical, geobotanical, remote-sensing, and geophysical surveys for the recognition of anomalies, geologic incongruities that can be related to a hidden deposit or to a favorable structural or lithologic condition in the terrain (**Fig. 5**). In a prospecting program, the sets of data acquired by direct and indirect methods are integrated in the development of successively defined target areas. The guiding exploration model is now a complex association in many levels of geomorphology, geology, geochemistry, geophysics, and economics. The spatial data are collected, studied, and interpreted using geographic information systems (GIS), computer-based technologies for managing and visualizing exploration programs. *See* GEOGRAPHIC INFORMATION SYSTEMS.

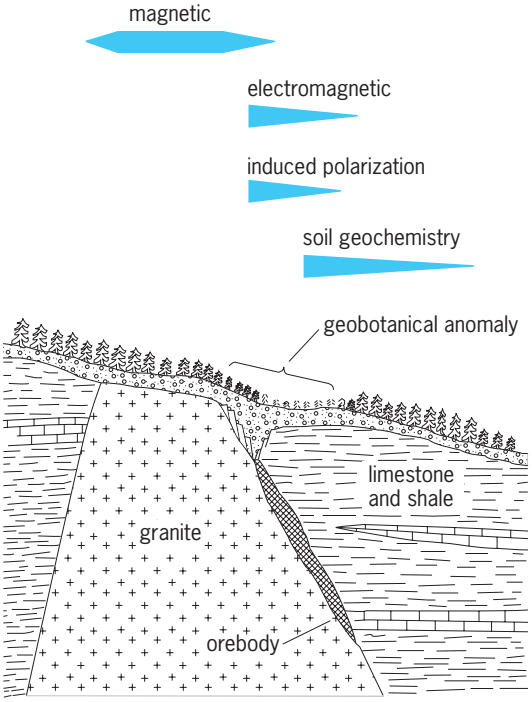

Fig. 5. Geophysical and geochemical expression of a granitic intrusion and a sulfide orebody beneath a cover of glacial sediments and soil, with topographic and geobotanical indications of mineralization. The geobotanical anomaly is revealed in stunted trees, smaller bushes, and grass. Each wedge-shaped symbol indicates one of the other specified anomalies, with the height corresponding to its intensity at each point, and the length, therefore, marking the extent of the anomaly.

Direct methods. The topography itself may give evidence of abrupt depressions related to the leaching and collapse of sulfide ore minerals, or it may show boldly exposed silicified zones associated with ore. Some of the latter expressions of ore mineralization represent outcrops of siliceous iron formation host rocks, quartz-filled breccia pipes, and the prospector's classic quartz reefs that indicate vein deposits. Aerial photography and satellite imagery are valuable in searching for the topographic expression of potential ore mineralization; they are also used in the reconnaissance study and mapping of potentially ore-associated structural patterns and in the search for color and texture manifestations of altered areas and gossan.

Outcrops of gossan are examined in the field for evidence of underlying ore mineralization, and trains of float (fragments of ore and gossan) are traced toward their apparent topographic origin. In glaciated terrain, trains of ore boulders are mapped and traced systematically toward their apparent sources; the copper deposit at Outokumpo, Finland, was located in this way from a boulder train 30 mi (50 km) long.

Placer gold and placer accumulations of other minerals such as platinum, cassiterite, rutile, and diamonds are sought as economic deposits in themselves and are used as guides to upstream deposits of associated minerals. In addition, resistant and relatively dense minerals in stream gravels and residual heavy minerals in soil are collected by the long-established prospector's method of panning the loose material, and these are traced to a source area.

Geochemical and geobotanical methods. Geochemical prospecting is based on two characteristics of orebodies: an association with anomalous concentrations of chemical elements within primary halos in the surrounding rock, and an association with secondary dispersal patterns of chemical elements in the surficial products of their weathering and erosion. Geochemical methods involve the field and laboratory analysis of sampled rock, soil, vegetation, and other natural materials for trace amounts (in parts per million or billion) of the principal indicator elements of an orebody and of the related pathfinder elements that provide more recognizable or farther-reaching anomalies. *See* GEOCHEMICAL PROSPECTING.

Geochemical sampling. In rock sampling, primary alteration halos and leakage mineralization may be

identified for several meters away from the principal mineralization and may be detected for a considerable distance from large ore deposits. In the lead-zinc deposits of Ireland, manganese pathfinder halos extend for as much as 4 mi (7 km), and in the western United States leakage anomalies have been used to locate blind silver-lead-zinc orebodies beneath some 200 ft (60 m) of overlying rock.

The sampling of soil, glacial debris, and stream sediments for the identification of secondary geochemical patterns is a modern correlative to the tracking of visible minerals in gossan, float, ore boulders, and alluvium. Halos or downslope fans of indicator and pathfinder elements occur in residual soil above weathered ore deposits, and longer dispersion trains of trace elements occur in stream and glacial sediments. The discovery of La Caridad, a major copper deposit in Mexico, was aided by the recognition of a diagnostic geochemical train that extended for 12 mi (20 km) in stream sediments.

Lake sediments and water, especially in such terrain as that of the Canadian Shield, reflect the geochemical nature of a local catchment area, and they are sampled in geochemical reconnaissance. Dissolved metal salts and dissolved gases occur in the waters of streams, wells, and springs; these sites are sampled as guides to mineralization in surface- and ground-water drainage systems. Snow and meltwater have also been sampled for entrained mercury vapor and for gaseous complexes given off during the weathering of orebodies, and gases have been sampled directly from the soil and in the atmosphere.

Vegetation sampling. Also known as biogeochemical prospecting, this method is an indirect means of collecting information from deeper soil, rock, and ground water. Roots of trees and bushes accumulate certain elements from the subsurface and transmit them to the foliage; the root systems of phreatophytes in desert areas can reach to ground-water levels at depths of 160 ft (50 m) or more. Samples, commonly twigs or leaves, are taken for analysis from a relatively well-distributed type of vegetation such as pine trees or juniper bushes.

Geobotanical prospecting. In this method the presence of morphological or mutational changes in vegetation is taken as a guide to local geochemical conditions. The Central African copper flower, the European zinc violet, and the selenium-associated poison vetch of the western United States are examples of indicator plants used in prospecting. Areas with significant numbers of indicator plants and areas with dwarfed, discolored, or characteristically metal-stressed plant growth can be detected in surface reconnaissance, but their detection is aided considerably by the study of color aerial photographs and aircraft or satellite imagery. *See* GEOCHEMICAL PROSPECTING.

Remote sensing. Imagery provided by remote sensing from aircraft and orbiting satellites is of fundamental importance in prospecting and in the patterns of exploration data associated with GIS. Photogeology, the interpretation and mapping of terrain from aerial photographs, is in wide use, especially in reconnaissance as an aid to the direct identification of guides to ore and as a means of broader geologic investigation. Aerial photography in spectral bands of the near-infrared and near-ultraviolet frequencies is also used in photogeology for discriminating between types of exposed rock and soil and for emphasizing the appearance of bleached and stained areas as well as geobotanical anomalies. Airborne remote-sensing systems have provided radar imagery of terrain in the prospecting of cloud-covered jungle regions, and they have furnished thermal-band infrared surveys for recognizing anomalously warm areas that may be associated with mineralization. Airborne multispectral sensors with the capability of identifying some of the specific kinds of minerals in altered zones have been tested for use in prospecting. *See* AERIAL PHOTOGRAPH; REMOTE SENSING.

The most widely used satellite imagery in prospecting is from the American *Landsat* series and the French *SPOT* series of spacecraft. The resolution of terrain features is less than that of airborne remote sensing, but the global coverage of major geologic features in multiple passes and seasonal conditions affords an especially suitable base for reconnaissance work in prospecting. Higher-resolution color photography, multispectral imagery, and radar data from newer spacecraft, including Canada's *Radarsat*, Japan's *JERS*, and the European Space Agency's *ERS*, have been applied to reconnaissance prospecting in those regions where coverage is available.

Geophysical methods. Geophysical exploration is based on the measurement of physical properties associated with geologic features. As a means of both airborne and ground prospecting for mineral deposits, it involves the recognition of contrasts in properties between the deposit and the adjacent rock, generally to depths on the order of 330–660 ft (100–200 m), and the definition of deeper structural and lithologic features to be used as guides to ore mineralization. The signals gathered in geophysical exploration are separated as far as possible from irrelevant noise and resolved into patterns associated with known and assumed geologic conditions; departures from the regional and local patterns are geophysical anomalies, and these are interpreted in respect to appropriate exploration models.

Magnetic, electrical, electromagnetic, and radioactive methods are the most widely used in prospecting for ore and industrial minerals deposits. Seismic and gravity methods are employed to some extent but not nearly so much as in petroleum prospecting. Ground-penetrating radar can provide images of shallow features for environmental investigations. Geophysical surveys are often made by several methods, so that more than one physical property can be taken into account. *See* GEOPHYSICAL EXPLORATION.

Magnetic surveys. Surveys are made with magnetometers—sensitive instruments placed at successive ground stations along traverse lines, carried in a helicopter, or carried in a fixed-wing aircraft in a pattern of flight lines (**Fig. 6**).

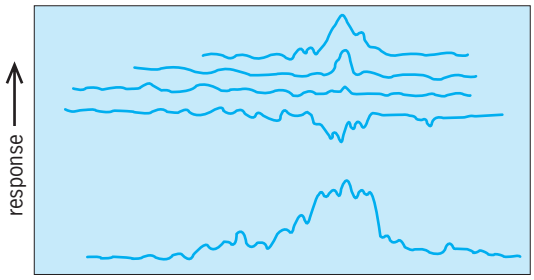

Fig. 6. Response curves for airborne electromagnetic (upper) and magnetometer (lower) prospecting system over a magnetic (pyrrhotite) and conductive stockwork deposit. The peak represents the location of the orebody.

The measurements in magnetic surveys are related to the degree of induced magnetization imposed on a rock or a mineral deposit by the Earth's magnetic field. The most common magnetic minerals are magnetite, ilmenite, pyrrhotite, and specular hematite. Magnetite, the mineral with the highest magnetic susceptibility, is a major component of iron orebodies, an abundant mineral in lithologic zones associated with contact metamorphic deposits, a fundamental component of mafic intrusive and volcanic rocks, and a minor mineral in many other rocks. A magnetic anomaly could therefore have several geologic explanations. A distinct and sharp anomaly, for example, may be associated with a body of iron ore, a contact metamorphic copper deposit, or a mafic dike; alternatively, it may represent a nickel orebody with pyrrhotite as the principal mineral, or a titanium orebody composed mainly of ilmenite. A subtle and broad magnetic anomaly, taken into account with an expected geologic pattern, may be the expression of a granitic intrusion or of a contact between basaltic volcanic rocks and limestone. *See* GEOMAGNETISM; MAGNETOMETER.

Electrical surveys. In electrical methods of prospecting, pairs or series of electrodes are placed in the Earth, and measurements are made of natural or induced currents associated with geologic features. The self-potential or spontaneous potential method involves traversing an area to find the patterns of natural currents set up by electrochemical action in conductive orebodies, especially sulfide mineral bodies. With resistivity methods, an electrical field supplied to the current electrodes during traversing is distorted in the Earth by variations in resistivity between geologic features, and the resulting pattern is measured at receiving or potential electrodes. A conductor, a zone of anomalously low resistivity, is interpreted in relation to an appropriate exploration model and alternative geologic conditions. A conductor may represent a mineral deposit, but it may also be related to such things as a conductive body of graphitic schist or clay. *See* ROCK, ELECTRICAL PROPERTIES OF.

Induced polarization, the most popular of electrical methods, makes use of resistivity measurement and also an electrochemical polarization effect that occurs between two modes of electrical conduction, ionic (in pore fluids) and electronic (in metallic minerals). Mineralized rocks, where both kinds of conductors are present, can be detected by time domain surveys in which a chargeability or characteristic flow of current is measured between pulses of applied current; mineralized rocks can also be detected by frequency domain surveys in which the resistivity differs with changes in the frequency of the applied current. In unmineralized rock, where the current is carried only by ionic conduction, there is no flow between pulses, and the resistivity is independent of the applied frequency. The apparent resistivity and the chargeability, the frequency effect, or an inherent phase shift are measured at successive stations along a traverse; these values and a derived parameter, a metal factor, are determined for a series of depth points below the surface.

Electromagnetic surveys. Electromagnetic methods provide a means of determining differences in the electrical conductivity of rocks without having to make direct contact with the Earth. An electric current in a transmitting coil or at a distant source such as an atmospheric disturbance or a very low frequency communication station generates a primary electromagnetic field in the Earth's crust; eddy currents are induced in conductive bodies, and the resulting secondary field is measured with a detector coil (**Fig. 7**). As with electrical methods, conductors must be evaluated in relation to the appropriate exploration model and to other conductive geological conditions.

Airborne electromagnetic methods have been quite successful in locating anomalies associated with massive sulfide orebodies, especially in areas such as the Canadian Shield where conductive mineralization is hidden at shallow depth beneath glacial debris. Airborne electromagnetic surveys (Fig. 6), generally flown in combination with magnetic and radioactive systems, are also used as an aid to geologic mapping in regions where outcrops are scarce.

Ground electromagnetic traverses are commonly made in selected target areas with various configurations of transmitting and detecting coils. In some

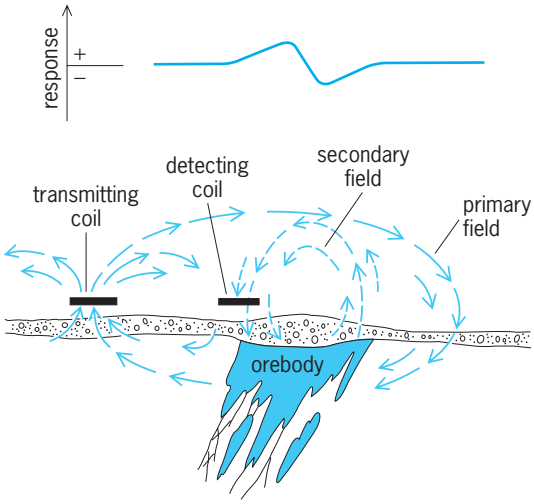

Fig. 7. Effect of a conductor on a ground electromagnetic system and the associated response amplitude curve.

electromagnetic methods, the transmitter is a ground cable or loop several miles long, and the small detector coil is carried across the area or is flown above the area in a helicopter.

Radiometric surveys. The principal methods in airborne and ground radioactivity surveying are gamma-ray spectrometry and total-radiation radiometrics. In both methods, the detector is a crystal that emits a flash of light, a scintillation, when struck by a gamma ray. The scintillation is converted to a voltage pulse that is read or recorded as a count rate. Gamma rays are emitted by uranium-, thorium-, and potassium-bearing minerals; the gamma-ray spectrometer is a specifying detector, equipped with separate channels for measuring the diagnostic radiation from each of three sources. *See* GAMMA-RAY DETECTORS.

Radioactivity surveys are the principal approach to prospecting for uranium deposits; the deposit or its uranium-bearing weathered equivalent must be near the surface, inasmuch as gamma rays are completely absorbed by a few inches of water or a few feet of barren soil and rock. Radioactivity surveys are also used in prospecting for deposits of phosphate rock, titanium-bearing sand, and rare-earth minerals, many of which contain minor amounts of associated uranium and thorium.

Airborne radioactivity surveys are made during regional exploration as an aid to geologic mapping; this application is based on the presence of potassium minerals and minor amounts of uranium and thorium in granitic intrusions and in several types of sedimentary rock.

Subsurface Prospecting

Direct subsurface investigations are a part of prospecting within target areas and in the detailed stages of target investigation. Subsurface prospecting is done by pits, trenches, and drill holes, and by underground workings such as shafts and tunnels.

Drilling is the principal method of subsurface prospecting where evidence of ore mineralization and geophysical or geochemical anomalies indicates a target for prospecting at a depth of more than a few feet. The common methods are diamond drilling, rotary drilling, and percussion drilling. In diamond drilling (**Fig. 8**), a solid core is taken as a sample, and holes may be drilled at various orientations to depths of several thousand feet. In rotary and percussion drilling, rock chips or cuttings come to the surface in a stream of air, foam, or mud, and these are used as samples of the material at depth. *See* DRILLING AND BORING, GEOTECHNICAL; ENGINEERING GEOLOGY.

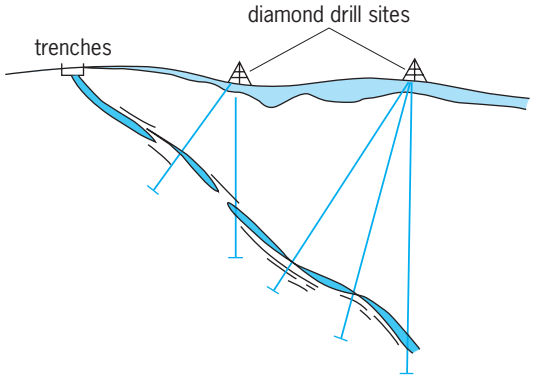

Fig. 8. Subsurface exploration pattern with vertical and angle diamond-drill holes.

Geophysical information is obtained by the probing or logging of drill holes. Electrical and electromagnetic logging is done in holes drilled in search of metallic orebodies; with these methods, the radius of search is extended considerably beyond that of the small-diameter cylinder of sampled rock. Gamma-ray methods of geophysical drill-hole logging have become standard practice in prospecting for uranium ore, and there are several methods of nuclear activation in which other rocks and ore minerals in the walls of drill holes are subjected to neutron bombardment to provide a secondary radioactivity signature. *See* WELL LOGGING.

The investigation of exploration target areas by sinking shafts and driving tunnels is a subsurface prospecting approach with a long history. However, underground prospect workings are more often used in the delineation and evaluation of orebodies after a discovery has been made in drill holes. Even so, there are situations in which an irregularly mineralized orebody can still be identified best in larger-scale underground sampling and in drilling from mine workings, in comparison with drill holes from the surface.

William C. Peters

Bibliography. J. Chaussier and J. Morer, *Mineral Prospecting Manual*, 1987; A. M. Evans (ed.), *Introduction to Mineral Exploration*, 1995; R. J. Goldfarb and R. L. Nielson (eds.), *Integrated Methods for Discovery: Global Exploration in the 21st Century*, Society of Economic Geologists, 2002; M. Kuzvart and M. Böhmer, *Prospecting and Exploration of Mineral Deposits*, 2d rev. ed., 1986; W. C. Lacy (ed.), *Mineral Exploration*, 1983; R. Marjoribanks, *Geological Methods in Exploration and Mining*, 1997; W. C. Peters, *Exploration and Mining Geology*, 2d ed., 1987; J. Tilton, R. G. Eggert, and H. Landsberg (eds.), *World Mineral Exploration*, 1988.

Prostate gland

A triangular body in men, the size and shape of a chestnut, that lies immediately in front of the bladder with its apex directed down and forward. It is found only in the male, having no female counterpart. The prostatic portion of the urethra extends through it, passing from the bladder to the penis. This organ contains 15–20 branched, tubular glands which form lobules. The gland ducts open into the urethra. Between the gland clusters, or alveoli, there is a dense, fibrous connecting tissue, the stroma, which also forms a tough capsule around the gland, continuous with the bladder wall. Penetrating the prostate to empty into the urethra are the ejaculatory ducts from the seminal vesicles which are

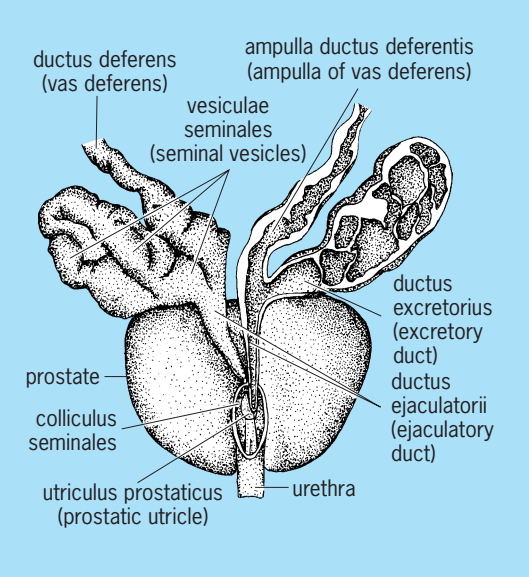

Prostate gland and seminal vesicles. (*After W. A. N. Darland, ed., American Illustrated Medical Dictionary, 19th ed., Saunders, 1942*)

located above and behind the organ (see **illus.**). The prostatic gland secretes a viscid, alkaline fluid which aids in sperm motility and in neutralizing the acidity of the vagina, thus enhancing fertilization.

A midline groove, palpable on rectal examination, indicates the two lateral lobes of glandular tissue; a medial lobe lies posteriorly. Small at birth, the prostate enlarges to adult size during puberty. After middle age, the prostate is sometimes subject to new tissue growth, usually benign, that may result in interference with urine flow through the compressed urethra. *See* PROSTATE GLAND DISORDERS.

Walter Bock

Prostate gland disorders

Disorders of the prostate gland that can cause voiding symptoms and sexual dysfunction. The three most common diseases are prostatitis, benign prostatic hyperplasia, and prostate cancer. Prostatitis is an inflammatory condition of the gland, causing a variety of voiding symptoms and pain in the area around the prostate. Benign prostatic hyperplasia is a benign growth of the inner portion of the gland that causes progressive obstruction of the urinary channel (urethra) at the base of the bladder. Prostate cancer is the most common nonskin malignancy in men and is a major health concern for the aging population. *See* CANCER (MEDICINE).

Benign prostatic hyperplasia. Benign prostatic hyperplasia is a common disease of the aging male. Approximately 30–50% of all men develop urinary symptoms because of an enlarged prostate. While the exact cause of this disease is not known, it is believed that the male hormones testosterone and dihydrotestosterone play a significant role in the development and maintenance of benign prostatic hyperplasia.

The disease process results in the development of outlet obstruction indicated by urinary hesitancy, weak stream, dribbling, and the sensation of incomplete bladder emptying. In addition, for reasons that are not completely understood, a number of individuals also develop irritative voiding symptoms: frequency, nocturia (excessive urination at night), and urgency.

Not all individuals have outlet obstruction. As many as 25% may have a coexisting neurogenic bladder and, if treated solely for prostatism, usually will not improve.

Treatments. Transurethral resection is the most common treatment offered, but there are other forms of therapy. The transurethral resection is associated with a 70–80% success rate and a 3–10% complication rate. All other forms of therapy have a lower complication rate but also a lower success rate. The alternate approaches include watchful waiting, medication, and minimally invasive procedures. Watchful waiting is an appropriate form of management in individuals with only mild symptoms. The majority of these individuals will not progress in the short term and will suffer no damage to the lower urinary tract if observation is elected. Those who have moderate symptoms can be offered a number of options ranging from medical therapy to surgery.

Medication. Two types of medication are used to treat the symptoms associated with prostatism (the symptom complex caused by urethral obstruction). Because the enlarged prostate is dependent on androgens, pharmacologic manipulation that lowers levels of dihydrotestosterone results in prostate gland shrinkage and improvement in the urinary symptoms.

Another form of medical therapy reduces the tone of the prostatic urethra by relaxing the prostate. This effect is achieved by an agent that blocks the alpha receptors which, when stimulated, cause contractions of the prostate gland.

The decision of which drug to use is based upon the clinical condition of the individual.

For individuals with severe symptoms or obstruction, surgery remains the best alternative. Individuals with very large prostates are best treated with an open prostatectomy. Those with moderately enlarged glands can have a transurethral resection. The smaller prostates can be treated by a transurethral incision. These operations will effectively cure 70–80% of the patients. Approximately 10% of the individuals will require a reoperation within 10 years. Complications are greater with this form of treatment in comparison to medical therapy or the minimally invasive techniques.

Prostate cancer. In men, prostate cancer is the most common cancer and the second leading cause of cancer death. In the United States the number of new cases has dramatically increased owing in part to the use of prostate specific antigen for the early detection of prostate cancer. When the antigen is used in screening or for early diagnosis, the detection rate

increases significantly. The American Cancer Society recommends an annual rectal examination starting at age 40 and an annual screening study beginning at age 50.

Stages of prostate cancer. Once a diagnosis of prostate cancer is made, the treatment options depend on the stage of the disease. Several factors influence therapeutic choices prior to making recommendations for treatment. These include the individual's age and health, and the stage and grade of the lesion.

Individuals whose disease is confined to the prostate (stage A or B) may be managed by watchful waiting, radical prostatectomy, or radiation therapy. In those with locally advanced disease (stage C), radiation or hormonal therapy may be indicated. In individuals with metastatic disease (stage D), hormonal therapy is used.

With the use of prostate specific antigen testing, a greater number of individuals show localized disease. Thus the number of individuals treated by radical prostatectomy has significantly increased. Radical prostatectomy involves removing the pelvic lymph nodes. If they are free of metastatic disease, the entire prostate is removed, from the bladder to the urinary sphincter.

Major modifications of this operation have greatly decreased the complications. The nerve-sparing or anatomic radical prostatectomy has resulted in a decrease in the risk of urinary incontinence from 10% to less than 3% of individuals, and an increase in preservation of potency to as high as 70%.

Noninvasive techniques. Conformal therapy utilizes three-dimensional computerized images to reconstruct the prostate in planning the external beam therapy. This technique sculpts the beam to the prostate and eliminates irradiation to noncancerous contiguous tissues (small bowel, rectum, and bladder). Conformal therapy may also allow for increased dose and possibly improved results. Real-time ultrasound guided brachytherapy allows for the placement of radioactive sources (either iodine-125 or palladium-103) directly into the substance of the prostate gland without spreading the radiation to the contiguous tissues.

The optimal treatment for individuals with stage C prostate cancer is not known. Younger men might best benefit from hormonal therapy given for 3–4 months prior to radical prostatectomy.

Individuals with metastatic prostate cancer are best managed with hormonal therapy, which deprives the prostate cancer cells of androgens, causing the cancer to regress. Complete hormonal therapy is accomplished by combining castration with medication that blocks the adrenal androgens. This medication, called an antiandrogen, works at the level of the androgen receptor within the prostate cancer cell. The combination of the antiandrogen with castration effectively removes all available androgens from the prostate cancer cell. The result is a longer time in remission and a prolonged survival. *See* CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; PROSTATE GLAND; RADIOLOGY.

Nelson N. Stone

Prosthesis

An artificial replacement of a body part. It may be an internal replacement such as an artificial joint or an external replacement such as an artificial limb. A prosthesis may be individually manufactured for each client, such as a limb prosthesis, or may be manufactured in large numbers but in different sizes, such as artificial hip joints. Prostheses of all types are lighter and more functional than their predecessors; the broad field of prosthetics has benefited from advances in materials, miniaturization, and computer-generated fabrication.

Limb prosthetics. A standard nomenclature is used to refer to level of amputation and related prostheses. The term trans is used when an amputation goes across the axis of a long bone, such as transtibial (across the tibia of the leg) or transhumeral (across the humerus of the arm). When there are two bones together such as the tibia and fibula, the primary bone is identified. Amputations between long bones or through a joint are referred to as disarticulations and identified by the major body part, such as knee disarticulation. The term partial is used to refer to a part of the foot or hand distal to the ankle or wrist that may be amputated. Artificial limbs are designed to enable the individual to participate in all aspects of life, including sports and leisure activities.

The majority of lower limb amputations are performed as a result of vascular disease; more than 50% are performed on individuals over the age of 60, who have peripheral vascular disease secondary to diabetes. The next most frequent cause of amputation is trauma, with cancer and congenital deformities following. The major causes of juvenile amputations are trauma, congenital abnormalities, osteomyelitis, and cancer. In the upper extremity, the major cause is trauma followed by disease. *See* VASCULAR DISORDERS.

Upper and lower extremity prostheses need to be comfortable, functional, and cosmetic. If the prosthesis is not comfortable, the individual will not wear it; pain or discomfort can be the greatest impediment to successful prosthetic rehabilitation. The prosthesis must also allow the individual to perform desired activities that could not be done without a prosthesis, and to do them with the lowest possible expenditure of energy.

For most individuals, a well-fitting prosthesis will allow a greater range of mobility activities with less expenditure of energy than crutches or a wheelchair. This may not always be true for elderly individuals with transfemoral amputations or for individuals with bilateral amputations. While many individuals of all ages with bilateral transtibial amputations regain functional mobility with prostheses, adults with bilateral transfemoral amputations usually find ambulation too fatiguing. When the energy demands for prosthetic mobility are greater than for mobility without a prosthesis, the prosthesis is rarely worn.

The upper extremity has always posed a major challenge since the great complexity of hand

function is difficult to duplicate mechanically. The loss of sensation limits the function of the hand or hook, and replacement for sensory function has not been developed.

Sockets. Socket design varies with the level of amputation and the configurations of the individual residual limb. The prosthetic socket must support body weight and hold the residual limb firmly and comfortably during all activities. Each area of the residual limb tolerates pressure differently; therefore the tissues are selectively loaded so that the most weight is borne by pressure-tolerant tissue (such as wide and flat bony areas or tendons) and the least by pressure-sensitive tissue (such as nerves that are close to the skin and sharp bony prominence). This is accomplished by relief (socket concavity) over pressure-sensitive areas and socket convexity over the pressure-tolerant areas. Additionally, the socket needs to grip the residual limb firmly to reduce movement between the socket and the skin. The more movement between the residual limb and socket, the less secure the client will be during activities and the more risk for skin abrasions. Total contact between the distal end of the residual limb and the socket is required to aid in sensory feedback and to prevent swelling and skin problems. Sockets are individually constructed for each client from a cast made of the residual limb. The socket may be constructed with computer-aided design and computer-aided manufacture techniques. An electronic scanner is used to determine the shape of the amputation limb, and then a test socket is fitted to ensure comfort and to provide for appropriate weight-bearing and stabilizing pressures. The socket may be constructed by using a machine that molds plastic directly over the positive model of the residual limb. Sockets may be hard and rigid, or flexible and supported by a rigid frame.

Suspension. There are several methods of suspending each type of prosthesis. Suction sockets allow suspension without belts, sleeves, or cuffs. Negative pressure is created between the skin and socket by putting the residual limb in the socket in a manner to eliminate all air and create an intimate fit. In the transtibial prosthesis a rubber sleeve or a cuff that fits on the thigh may be used. In the transfemoral prosthesis a belt around the pelvis provides suspension.

Feet. Foot and ankle function is complex, and a variety of prosthetic feet are designed to respond dynamically to the pressure of walking and running. They store energy at the moment of heel contact, then return it at toe-off. The degree of dynamic response varies with the materials used in the foot, with carbon graphite providing the greatest degree of response.

Knee mechanisms. Replacement of knee function requires, at a minimum, control of both stance and swing. There are mechanical knee joints that provide some degree of each but not both. Knee mechanisms using hydraulic control can provide some degree of control of both stance and swing phases. Complex knee mechanisms allow an athlete to run by using a regular step-over-step pace. (**Fig. 1**).

Fig. 1. Lower-limb prosthesis. Flexible plastic socket, combined with hydraulically controlled knee, plus ankle-foot mechanism permits controlled movement and sufficient speed to allow this above-knee amputee to continue many sports activities. Cosmetic outer covering has been removed to display the internal components. (*Mauch Laboratories, Inc.*)

Terminal device. Of paramount importance is the prosthetic component that will substitute for the missing hand. There is no device that completely replaces the appearance or function of the anatomic hand. Nevertheless, the prosthetic counterpart, called a terminal device, can contribute to the wearer's function and self-esteem. The two types of terminal device are the hand and the hook. Either is secured to a plastic socket encasing the forearm.

The hand may be active or passive: passive hands have no moving parts; active hands have a mechanism that permits the client to control finger position by appropriate action in the proximal part of the amputated limb. Both active and passive hands are made in a range of sizes to suit most children and adults, and are covered with a flexible plastic glove that matches the individual's skin tone.

The most popular active hand is operated myoelectrically (**Fig. 2**). The individual wears a socket with one or more skin electrodes that contact appropriate muscle groups. When the muscle contracts, the electrode transmits the microvoltage generated by the muscle to a mechanism that causes the electrical signal to operate a small motor that, in turn, enables the hand mechanism to open or close the fingers. The motor is powered by a battery worn inside the prosthetic socket or, in the case of wrist disarticulation, elsewhere on the body. Myoelectric control can provide excellent grasp force. The alternative type of prosthetic hand has a steel cable that is attached to a trunk harness. By flexing the shoulder, the individual puts tension on the cable that pulls on the hand mechanism, causing the fingers either to open (voluntary opening) or to close (voluntary closing). Relaxing tension on the cable allows the fingers to revert to the opposite position.

A hook is the other basic type of terminal device. Made either of aluminum or steel, hooks have two fingers that the client can open and close. Myoelectrically controlled hooks are available; however, most individuals who wear hooks have cable-operated ones that are either voluntary-opening or voluntary-closing. Cable-operated hooks are much lighter in weight than any type of hand and are less expensive. Hooks are also more durable.

Other components. Sockets are custom-made of plastic and are intended to fit snugly to distribute pressure over the largest area. The proximal portion of the transradial socket terminates in the vicinity of the humeral epicondyles. A self-suspending socket supports the weight of the terminal device and wrist unit by a secure hold at the epicondyles. A harness is necessary to suspend the socket that is not supported by the epicondyles, to transmit shoulder motions to a cable, or to accomplish both functions. The customary design, figure-of-8, has a loop around each shoulder that connects in back.

The prosthesis for the individual whose amputation is at the elbow disarticulation level or higher includes a terminal device, wrist unit, forearm shell to replace the length of the absent forearm, an elbow unit, socket, and suspension. All elbow units have a hinge for elbow flexion and a locking mechanism to enable the wearer to retain the desired flexion angle. The elbow unit may incorporate a turntable that allows the individual to rotate the forearm shell medially and laterally. The figure-of-8 harness suspends the prosthesis and transmits shoulder and shoulder girdle motions to the cable system. Ordinarily, one cable operates the terminal device and the elbow hinge, and a second cable operates the elbow lock.

Joint replacements. Artificial replacements of joints, such as hip, knee, or shoulder, are another type of prosthesis. The knee and hip joints are the most frequently replaced, occasionally for trauma but usually for degenerative disease. Lower limb joint replacements must have load-bearing capabilities and are fabricated in different sizes of metal. Joint replacements may sometimes be secured by polymer adhesive. Hip and knee joint replacements are common, and a high degree of restoration of function is usually obtained. Most individuals experience major reduction or elimination of joint pain and 90–100% return of function.

In the upper extremity, shoulder and metacarpal phalangeal joints are most frequently replaced. The joint replacement in the upper limb needs to be light and to allow a great range of movement.

Other prostheses. There are a great variety of other types of replacements of body parts. Women who lose a breast to cancer are fitted with a prosthesis that can be fabricated of a number of lightweight materials and molded to resemble in shape and texture the remaining breast. In some instances an artificial breast made of biocompatible gel may be permanently inserted subcutaneously.

Prosthodontics are devices used to replace teeth. Lost teeth may be replaced by the insertion of a bridge containing several teeth resembling in shape and size the missing ones. Dentists can also implant a single artificial tooth permanently into the jaw. *See* DENTISTRY.

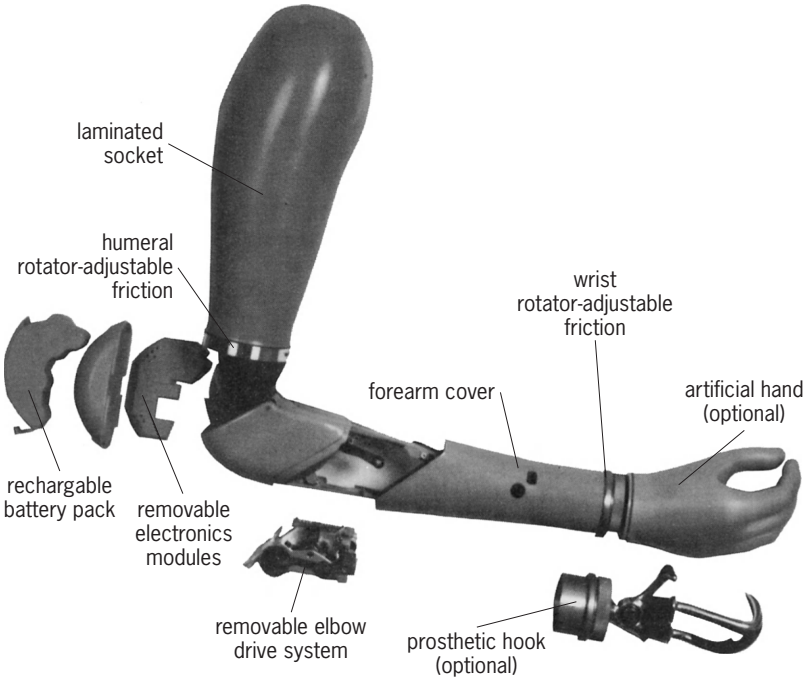

Fig. 2. Upper-limb prosthesis. Miniature electronic circuitry uses muscle activity from the upper-arm remnant to control the electric elbow. When the elbow is locked in position, control is automatically transferred to the electronic wrist and hand. (*Motion Control, Inc.*)

The cardiac pacemaker is a form of a prosthesis replacing the natural electrical stimulation of the heart with a battery-operated device inserted within the body. Damaged heart valves are replaced with artificial valves attached directly to the heart muscle and to the major blood vessels. Artificial valves may be of disk or ball-and-seat design and allow normal flow of the blood from one chamber to the other. The artificial valve also prevents backflow. Prosthetic eyes, fabricated to resemble the remaining eye in color and configuration, serve a cosmetic function only.

Bella J. May

Biologic-prosthetic systems. Although mechanical or electromechanical prosthetic devices represent the main option in many situations, regeneration of destroyed or resected tissue is the preferred goal. The use of inert materials to provide a basis for regenerated tissue growth offers intermediate options for reconstruction.

A porous material—metal, ceramic, or polymer—of good biocompatibility will host the ingrowth of specialized tissue and its associated vascularity. But this ingrowth may require many months for full investment of the available void spaces, and so the prosthetic system must provide adequate support and protection during this period. The ingrowth phenomenon may be utilized, for example, to develop attachment of a prosthesis to a bone site by investment with normal bony cells. Clinical trials have been conducted for several designs of "cementless'' total hip and total knee prostheses. On a more modest scale, porous metal systems are used for the

reconstruction of cancer resections of the long bones and for the fusion of vertebrae.

Specialized tissue growth can utilize porosity, in the form of a fabriclike implant, as a substrate which stimulates and guides the growth. A development of this nature is an artificial skin system, which involves an underlying mesh of a biodegradable polymer into which fibroblasts and keratinocytes, two types of human skin cells, can be made to grow. *See* BIOMEDICAL ENGINEERING; SURGERY. W. Rostoker

Bibliography. J. H. Bowker and J. W. Michael (ed.), *Atlas of Limb Prosthetics: Surgical and Prosthetic Principles*, 1992; J. P. Paul, J. D. Baylor, and J. M. Courtney (eds.), *Biomaterials in Artificial Organs*, 1985; R. B. Salter, *Textbook of Disorders and Injuries of the Musculoskeletal System*, 3d ed., 1999.

Protactinium

A chemical element, Pa, atomic number 91. Isotopes of mass numbers 216, 217, and 222–238 are known, all of them radioactive. Only ^{231}Pa, the parent of actinium, ^{234}Pa, and ^{233}Pa occur in nature. The most important of these is ^{231}Pa, an α-emitter with a half-life of 32,500 years. The artificial isotope, ^{233}Pa, is important as an intermediary in the production of fissile ^{233}U. Both ^{231}Pa and ^{233}Pa can be synthesized by neutron irradiation of thorium. *See* ACTINIUM; PERIODIC TABLE; RADIOACTIVITY; URANIUM.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Protactinium is, formally, the third member of the actinide series of elements and the first in which a 5*f* electron appears, but its chemical behavior in aqueous solution resembles that of tantalum and niobium more closely than that of the other actinides. *See* NIOBIUM; TANTALUM.

Metallic protactinium is silver in color, malleable, and ductile. The crystal structure is body-centered tetragonal. Samples exposed to air at room temperature show little or no tarnishing over a period of several months. The numerous compounds of protactinium that have been prepared and characterized include binary and polynary oxides, halides, oxyhalides, sulfates, oxysulfates, double sulfates, oxynitrates, selenates, carbides, organometallic compounds, and noble metal alloys. *See* ACTINIDE ELEMENTS. H. W. Kirby

Bibliography. P. W. Atkins et al., *Inorganic Chemistry*, 4th ed., 2006; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; S. Cotton, *Lanthanide and Actinide Chemistry*, 2d ed., 2006.

Protandry

That condition in which an animal is first a male and then becomes a female. It occurs in many groups, in addition to oysters and cyclostomes. The reverse condition is protogyny. *See* PROTOGYNY.

Tracy I. Storer

Proteales

An order of flowering plants, division Magnoliophyta, in the eudicots. Consisting of three families, the order is one of the most controversial in current classifications, with a lack of obvious morphological characters linking the family Nelumbonaceae (two species) to the other two families, Platanaceae (seven species) and Proteaceae (about 1350 species). However, DNA sequences indicate that the three families form a natural group, and detailed studies have revealed that Platanaceae and Proteaceae share similar wood anatomy. *See* MAGNOLIOPHYTA; MAGNOLIOPSIDA; WOOD ANATOMY.

Nelumbonaceae are aquatic, rhizomatous herbs with peltate leaves held above the water on long petioles, and have often been thought to be related to the true waterlilies (Nymphaeales). The fruit is highly characteristic with loose seeds embedded in the hard receptacle. *Nelumbo* (sacred lotus and American lotus) is used as a source of food (seeds, receptacle, and rhizome), as an ornamental, and as a sacred plant in several Asian countries. Platanaceae, from the Northern Hemisphere, are deciduous, monoecious trees with simple, palmately lobed leaves, flaking bark, and branched hairs. The fruits are hairy achenes in globose clusters which hang down. Plane trees (*Platanus*) are common street trees, especially in London, due to the regular loss of bark (which gives some resistance to pollution), and they also provide timber. *See* BARK; DECIDUOUS PLANTS; FRUIT; LEAF; NYMPHAEALES.

Proteaceae (predominantly tropical and subtropical, in the Southern Hemisphere) are evergreen shrubs and trees that often accumulate aluminum. The majority of species are from the Cape of South Africa and Australia. The roots often have short lateral branches (proteoid roots). Leaves are normally spirally arranged, simple (sometimes pinnate), and with xerophytic adaptations. Flowers often have specialized modes of pollen presentation. Fruits are of various types, often releasing seeds only after exposure to fire. Several genera (including *Banksia*, *Protea*, and *Leucospermum*) are widely cultivated for cut flowers; *Grevillea* and other genera are used for timber; and *Macadamia* yields edible nuts. *See* EVERGREEN PLANTS; ROOT (BOTANY). Michael F. Fay

Proteasome

A large proteolytic complex that degrades intracellular proteins. Intracellular proteins are continuously synthesized and degraded, and their levels in cells reflect the fine balance between these two processes. The rate of breakdown of individual proteins inside the cell varies widely and can be altered according to changes in the cellular environment. In eukaryotic cells, the site for degradation of most intracellular proteins is a large proteolytic particle termed the proteasome. Proteasomes are a major cell constituent, constituting up to 2% of cellular protein, and are essential for viability. They are found in the cytoplasm and nucleus of all eukaryotic cells. Simpler but homologous forms of the proteasome are also present in archaea and bacteria. Much of the knowledge about the structure and function of eukaryotic proteasomes has been gained from studies of these simpler systems.

The form responsible for most intracellular protein breakdown is termed the 26S proteasome, which functions as a component of the ubiquitin-proteasome pathway (**Fig. 1**). This system catalyzes the rapid turnover of many critical regulatory proteins (for example, transcription factors, cell-cycle regulators, oncogenes); thus the proteasome plays a pivotal role in controlling a wide variety of cellular processes, ranging from cell division to circadian rhythms, gene transcription, and immune responses. Most normal long-lived proteins (which make up the bulk of proteins in cells) and damaged, misfolded, or mutated proteins (which could cause disease if they accumulated in cells) are also degraded by this pathway.

Proteasome versus protease. The proteasome differs from a typical proteolytic enzyme in many important respects. The typical protease is a single subunit enzyme of 20,000 to 40,000 daltons (Da). By contrast, the proteasome is up to 100-fold greater in size (2.5 million Da) and contains about 50 proteins with multiple enzymatic functions, some of which require adenosine triphosphate (ATP). Unlike typical protease pathways, the ubiquitin-proteasome pathway is ATP-dependent and uses the energy stored in ATP to mark, unfold, and transfer the protein substrate into the degradative chamber of the proteasome. Traditional proteases simply cleave a protein and release the partially digested fragments, whereas the proteasome binds and cuts the protein substrate into small peptides ranging from 3 to 25 residues in length, ensuring that partially digested proteins do not accumulate within cells. Peptides released by the proteasome are rapidly hydrolyzed to amino acids by peptidases in the cytosol, which are then reutilized for new protein synthesis.

Ubiquitin conjugation. Ubiquitin is a small (76 amino acids) protein found in all eukaryotic cells but not in bacteria and archaea. Most proteins degraded by the 26S proteasome are first marked by the covalent linkage to multiple ubiquitin molecules (Fig. 1). Ubiquitin conjugation represents a means of providing selectivity and specificity to the degradation process. Long chains of ubiquitin molecules are attached to proteins through the action of three enzymes called E1, E2, and E3. The E1 enzyme activates ubiquitin in an ATP-dependent manner and transfers it to one of the cell's approximately 15 E2's, which function as ubiquitin carrier proteins. The cell has hundreds of distinct E3 enzymes, each of which binds a specific group of proteins destined for degradation and catalyzes the transfer of the activated ubiquitin molecules from the E2 to a lysine side chain in the substrate protein. Proteins tagged

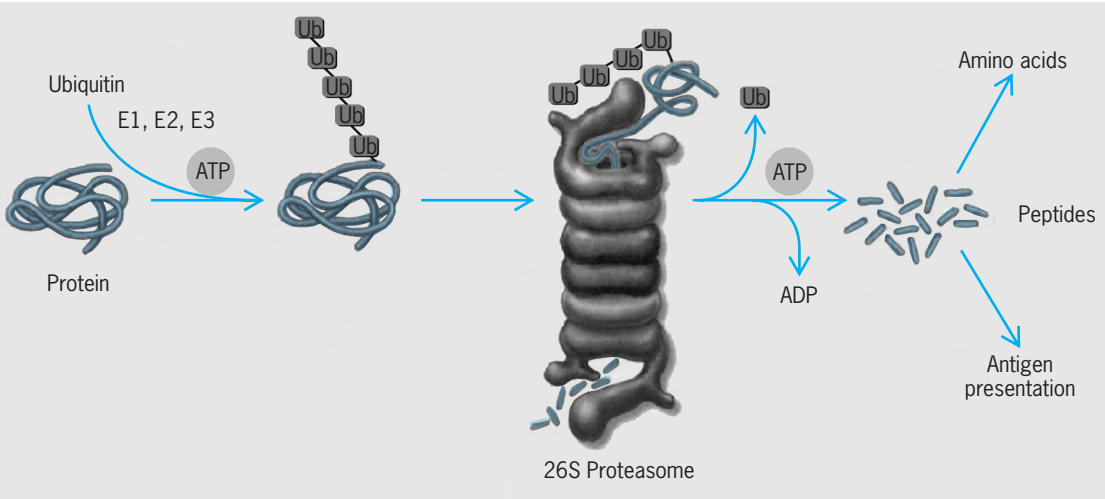

Fig. 1. Most intracellular proteins destined for degradation by the 26S proteasome are bound to multiple ubiquitin molecules through the action of three enzymes: a ubiquitin-activating enzyme (E1), a ubiquitin-conjugating enzyme (E2), and a ubiquitin-protein ligase (E3). Degradation of ubiquitinated proteins occurs within the central chamber of the proteasome. Ubiquitin conjugation, unfolding of a substrate, and its translocation into the inner cavities of the proteasome require energy obtained by ATP hydrolysis to ADP. The great majority of peptides released by the proteasome are further digested into amino acids by other peptidases in the cytosol, but a small fraction of peptides escape destruction and serve in MHC class I antigen presentation.

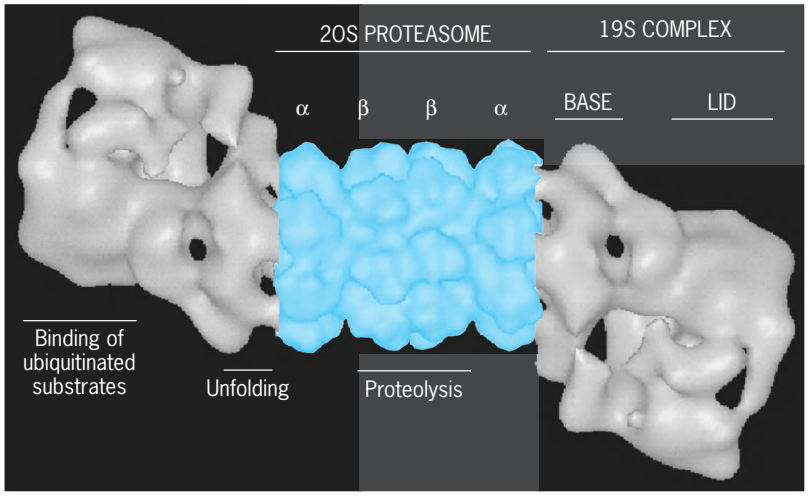

Fig. 2. Three-dimensional computerized reconstruction from electron microscopic images of the 26S proteasome. This large proteolytic complex is composed of about 50 different subunits that form subparticles, the 20S proteasome and the 19S regulatory particle. The lid of the 19S particle binds polyubiquitinated protein substrates and removes ubiquitin chains from the substrate. The substrate is then unfolded by the six ATPases located in the base of the 19S particle and translocated into the inner cavities of the 20S proteasome. Degradation occurs inside the central chamber of the 20S proteasome, formed by the two adjacent β-rings which contain six proteolytic sites. (*After W. Baumeister et al., The proteasome: Paradigm of self-compartmentalizing protease, Cell, 92:367–380, 1998*)

with ubiquitin chains are then rapidly bound and destroyed by the 26S proteasome.

Structure and function. The 26S proteasome is a complex composed of two very different components—the core 20S particle and one or two 19S regulatory particles (**Fig. 2**). Proteins are degraded within the 20S (720-kDa) particle, a 28-subunit barrel-shaped structure composed of four stacked rings, each containing seven distinct but homologous subunits surrounding a central cavity. The two inner rings (formed by the β-subunits) contain the proteolytic active sites and together form a central chamber in which proteins are degraded. The two outer rings (formed by the α-subunits) together with the β-rings form two antechambers through which protein substrates must pass. The proteolytic active sites of the β-subunits face the central cavity so that cleavage of proteins can occur only within the cavity. Only three of the seven β-type subunits possess catalytic activity. Each is specific for a different type of amino acid sequence: two cleave preferentially after basic amino acids, two after large hydrophobic amino acids, and two after acidic residues. Thus, the 20S particle can cleave nearly all the peptide bonds in proteins.

Substrates can enter the 20S particle only at the center of the α-ring. The N-terminal sequences of the α-subunits form a gate that is normally maintained in a closed position. Proteins or peptides cannot, therefore, freely enter the inner cavity of the 20S proteasome from the cytosol. This gate must be opened by a 19S regulatory particle for protein substrates to gain access to the catalytic chamber of the proteasome. The small diameter of the entry channel—even in the opened state—requires polypeptides to be unfolded in order to enter the chamber and be degraded. The 19S particle not only binds ubiquitinated proteins but also unfolds them and promotes their translocation into the 20S particle. The architecture of the particle, in which proteolysis is isolated to a separate chamber and access is tightly regulated, prevents inappropriate destruction of cell proteins.

The 19S (890-kDa) regulatory particle associates with one or both ends of the 20S proteasome and is composed of at least 17 different subunits. It consists of two different functional entities, the lid and the base (Fig. 2). The lid provides specificity to proteolysis and contains the binding site for ubiquitin chains and enzymes that break down the polyubiquitin chains so that the ubiquitin can be recycled for use in subsequent rounds of proteolysis. The base, which touches the α-rings of the 20S core particle, contains six different ATPases (a type of enzyme that hydrolyzes ATP and uses the energy released to drive biological processes). These ATPases form a ring that uses ATP to unfold protein substrates. They also have the capacity to open the gate in the α-ring of the 20S proteasome. In doing so, they promote translocation of substrates into the inner chambers of the 20S particle, where they are degraded. These structural features of the proteasome, the need for ubiquitination prior to degradation, and the energy requirement for proteolysis clearly evolved to provide a remarkable degree of selectivity and regulation to the degradative process and to safeguard against the nonspecific digestion of essential cellular proteins.

This complex system evolved from earlier proteolytic complexes found in archaea and bacteria, which lack ubiquitin and 26S proteasomes. Archaea, for example, contain 20S proteasomes that are arranged as in eukaryotes but appear to function in protein breakdown in association with an ATPase ring complex called PAN (proteasome-activating nucleotidase). PAN stimulates the ATP-dependent degradation of proteins by archaeal proteasomes, but it does not stimulate hydrolysis of small peptides that appear to freely enter into these particles. In prokaryotes, proteins bind directly to PAN. This hexameric-ring structure is homologous to the six ATPases found in the 19S base of the eukaryotic proteasome and has the capacity to unfold globular proteins and promote their entry into the 20S particle. Many features suggest that this PAN complex was the evolutionary precursor of the 19S regulator base of the eukaryotic 26S proteasome. The lid of the 19S regulator evolved from a distinct enzyme complex that allows degradation of ubiquitinated proteins. Thus, with evolution, the ATP-dependent function of the eukaryotic proteasome became linked to ubiquitin conjugation to provide greater selectivity and regulation, and lid subunits became necessary to bind the ubiquitinated protein substrates.

Catalytic mechanism. Proteolytic enzymes are normally classified according to which catalytic amino acids are in their active sites; proteins undergo a different nucleophilic attack depending on whether a serine, cysteine, aspartic, or metal group is in the active site of the enzyme. The sequences of the proteasome's catalytic β-subunits are not homologous to those of known proteases, and the pattern of

sensitivity to various inhibitors differs from that of any known protease family. X-ray diffraction studies and mutagenesis of different amino acids in the proteasome have uncovered a new type of proteolytic mechanism. The active-site nucleophile of the proteasome is the hydroxyl group of a threonine at the amino terminus of the β-subunit. Because no other protease of this type is known, it has been possible to synthesize or isolate inhibitors that specifically block proteasomes, without affecting other cellular enzymes. These inhibitors have been valuable tools in clarifying the proteasome's mechanism and determining its role in cells.

Antigen presentation. In mammals, a small portion of peptides generated by proteasomes escapes destruction in the cytosol and is transported into the endoplasmic reticulum (ER). In the ER, peptides of defined lengths (between 8 and 11 amino acids) are bound by specialized membrane-bound proteins, known as major histocompatibility complex (MHC) class I molecules, and are subsequently transported to the cell surface. The immune system is continually surveying the surface of all cells for nonself peptides originating from viral or mutated proteins. If cells carrying such nonnative peptides are encountered, they are rapidly destroyed by cytotoxic T lymphocytes. Specialized proteasomes, called immunoproteasomes, play an important role in the generation of most MHC class I peptides. They are normally found in lymphoid tissues (for example, the spleen and thymus) but are induced in other cells of the body during infection or in response to the cytokine interferon-γ. Immunoproteasomes contain three alternative proteolytic β-type subunits that replace constitutive β-subunits. These alternative subunits change the way that peptide bonds are cleaved within proteins in order to favor the production of peptides that can preferentially bind MHC class I molecules.

20S proteasomes can also associate with an activator called PA28, or the 11S regulator. PA28 is a heptameric ring composed of two types of subunits that can be attached to one or both ends of the 20S proteasome. PA28 greatly enhances the proteasome's degradation of small peptides (but not of intact or ubiquitinated proteins) by opening up the channel in the base of the particle. Like the alternative β-subunits of the immunoproteasome, PA28 is induced by the cytokine interferon-γ. This cytokine also enhances the formation of a hybrid proteasome particle that consists of a 20S core particle capped with the 19S regulator on one side and with PA28 on the other. It is thought that within these hybrid particles ubiquitinated substrates initially bind to the 19S regulator and peptide products are released through PA28. These hybrid particles appear to play a major role in the generation of MHC class I antigens.

Role in disease. The novel catalytic mechanism of the proteasome has permitted the synthesis of selective inhibitors that can be used to study the function of proteasomes in intact cells. These studies resulted in many of the present insights about the role of the ubiquitin-proteasome pathway in various biological processes. Besides being involved in the rapid clearance of mutated and abnormal proteins which accumulate in various inherited diseases, this pathway is implicated in regulated degradation of transcription factors, oncogenes, tumor-suppressor genes, and cell-cycle regulators. Therefore it is not surprising that several prevalent diseases are caused by the gain or loss of function of this proteolytic system. For example, carcinomas arising from cells lining the uterine cervix are caused by certain strains of human papilloma virus. Upon infecting cells, these viruses produce a protein (E6) that in concert with an E3 (called E6-AP) causes rapid ubiquitination and proteasome degradation of the major tumor suppressor p53. The lack of p53 dramatically increases the tendency of cells to become malignant. Mutation of another E3, the von Hippel-Lindau tumor suppressor, is responsible for some types of kidney tumors. Among its many functions, the suppressor regulates the levels of a specific transcription factor that controls the growth of blood vessels into tissues. This protein is normally maintained at a low level by rapid ubiquitination and degradation by proteasomes; however, when the suppressor is mutated, the transcription factor accumulates in cells, and tumors are able to form new blood vessels and grow rapidly.

The ubiquitin-proteasome system is also very important in the onset of inflammation. The transcription factor known as nuclear factor-kappa B (NF-κB) is the key regulator of expression of many inflammatory mediators. Normally, it is localized to the cytoplasm through binding of an inhibitory protein, termed IκB, which prevents its translocation to the nucleus. However, bacterial surfaces, oxidative stress, cytokines, and a cytokine tumor necrosis factor-α cause rapid phosphorylation, ubiquitination, and degradation of IκB, allowing NF-κB to enter the nucleus, bind to DNA, and activate inflammatory gene expression.

A general acceleration of the ubiquitin-proteasome pathway is responsible for the excessive breakdown of cell proteins in muscle during various disease states characterized by muscle wasting, such as cancer cachexia, diabetes mellitus, sepsis, inflammatory diseases (for example, AIDS), chronic renal failure, and nerve injury or disuse atrophy. There is appreciable interest in using proteasome inhibitors to treat various pathological states because of the key role of the proteasome in human disease. Promising results have been reported in animal models of various diseases, including cancer, rheumatoid arthritis, and stroke. Human trials with these agents are in progress. *See* ADENOSINE TRIPHOSPHATE (ATP); AMINO ACIDS; CYTOLYSIS; ENZYME; HISTOCOMPATIBILITY; NUCLEOPROTEIN; PROTEIN; TUMOR VIRUSES.

Tomo Saric; Stewart Lecker; Alfred L. Goldberg

Bibliography. A. Ciechanover, A. Orian, and A. L. Schwartz, Ubiquitin-mediated proteolysis: Biological regulation via destruction, *Bioessays*, 22(5):442–451, 2000; D. H. Lee and A. L. Goldberg, Proteasome inhibitors: Valuable new tools for cell biologists, *Trends Cell Biol.*, 8:397–403, 1998; J. M. Peters,

J. R. Harris, and D. Finley (eds.), *Ubiquitin and the Biology of the Cell*, Plenum Press, New York, 1998; K. I. Rock and A. L. Goldberg, Degradation of cell proteins and the generation of MHC class I-presented peptides, *Annu. Rev. Immunol.*, 17:739–779, 1999; A. L. Schwartz and A. Ciechanover, The ubiquitin-proteasome pathway and pathogenesis of human disease, *Annu. Rev. Med.*, 50:57–74, 1999; D. Voges, P. Zwickl, and W. Baumeister, The 26S proteasome: A molecular machine designed for controlled proteolysis, *Annu. Rev. Biochem.*, 68:1015–1068, 1999.

Protective coloration

A strategy that organisms use to avoid or deflect the attacks of predators by misleading the latter's visual senses. Protective devices directed at the hearing or the sense of smell of the predators are harder for researchers to detect, and are known in only a small number of species.

Types. Protective coloration can be classified according to whether the functioning or malfunctioning of the vertebrate visual system is exploited. Exploiting the malfunction of the system means simply "not being seen'': the prey fails to attract the attention of the predator, usually because it is the same color as the general background (the green of leaves, the mottling of lichen, or the brown of soil) or because it fails to cast a shadow. The organism avoids producing shadow by flattening itself against the substrate, or by countershading, in which the lower parts of a cylindrical prey such as a caterpillar are more lightly colored than the upper parts. As shadows normally form on the underside of cylinders, the shading cancels the shadow and makes the caterpillar optically flat. Animals that match their background often have an ability to select the appropriate background to rest on, or much less frequently (although widely known in the case of the chameleon) can change their own color to match.

Exploiting the functioning of the vertebrate perceptual system takes many forms: the vertebrate visual cortex decodes the image on the retina in a hierarchical process starting with the detection of edges, and then fills the shapes so defined with colors. A predator may thus detect the shape of a moth which it will fill with the color of the moth; this process enables detection. A moth may counter this by possessing strikingly contrasted patches of color on its wings, arranged in a random way that resembles no interesting shape. What the predator then sees is a random arrangement of, for example, black and white, which it cannot decode as a significant shape; the outline of the moth is broken up.

The prey may alternatively exploit the learning capacity of the predator. Many objects in the world are probably accurately perceived but ignored as of no interest. For example, insectivorous birds assuredly see leaves but do not attack them because they have learned (or perhaps know innately) that these are not edible. Resemblances to leaves, twigs, thorns, flowers, parts of flowers, and more bizarre objects like fresh turds (usually bird droppings) or discarded head capsules of ants are very widespread. This type of camouflage is termed mimetic camouflage. Camouflage in general is often termed cryptic coloration.

Mimetic coloration and warning color. Coloration may be considered mimetic if protection is achieved by a resemblance to some other existing object, which is recognized by the predator but not associated in its mind with feeding. Usually this negative or neutral association is learned, but in a minority of instances it is almost certainly innate. Small birds have an innate flight response to large eyes in close-up (which normally indicate that a cat or a predatory bird is dangerously close). This reaction is exploited by many moths and other insects, which have eye-like markings, sometimes very convincing in their shading and highlighting, on concealed parts of the wings. Attack by a predator causes such a moth not to flee but to change its posture rapidly to reveal the fake eyes, thus frightening away the attacker. Motmots (birds which habitually prey on snakes) have a similar innate fear of the red, black, and yellow striping patterns of the deadly coral snakes. These patterns are mimicked by various nonvenomous snakes. The fear that most birds have of snakes in general results in effective mimicry of even cryptically colored snakes by caterpillars with snakelike coloring and defensive behavior that usually combines the display of large eyelike markings with the posture of a snake about to strike. The ancestor of the domestic silkworm had such coloring, and the now colorless silkworm continues to use the display.

Flash coloration. Flash coloration seems to be an intermediate between a warning color and large eyespots. The prey is cryptic when at rest, but reveals brilliantly colored parts while escaping. This behavior seems to function simply by startling the predator. Very small eye marks at the tips of the wings, or a false head at the wrong end of the body (shown by some coral reef fish, for example) may cause the predator to misdirect its attack.

Aposematic defense. Protection through the possession of a chemical or physical defense that is dangerous to one's potential predator, accompanied by a strikingly conspicuous pattern known as warning coloration (often black, red, yellow, and white), is widespread—the ensemble of defense and color is termed aposematic. Examples occur in Hymenoptera, Lepidoptera, and sea slugs. The actual defense ranges from toxic venoms through stings (in wasps, for example), to the oozing of noxious foams or hemolypmph (as in ladybirds), to the possession of toxic chemicals (cyanides, cardiac glycosides, alkaloids) that will poison the predator or simply produce a revolting taste. The defense may be synthesized de novo or may be obtained from a host plant (as with many butterflies and moths) or from the animal's own prey (as with the use of nematocysts by some sea slugs). The function of the warning color is to remind the predator of its previous unpleasant experience, and is, for a variety of reasons connected

with the neurology and cognitive aspects of learning, easier to remember than a pattern which resembles the background.

Bright coloration must have evolved in circumstances in which the disadvantages of being greatly conspicuous were outweighed by the advantages conferred by more rapid learning on the part of the predators. Related factors involve the prey's being so tough that it can often survive an attack (when the predator will readily learn not to attack it again), or the prey's being gregarious, so that individuals, although killed when attacked, confer benefits on their immediate relatives through their death (gregarious groupings usually being siblings from the same batch of eggs).

The warningly colored life-style involves the development of some distinctive behavior, such as remaining motionless and oozing liquid (or otherwise threatening the predator) when attacked; and in butterflies, habitually flying more slowly and in more shady habitats (where the body temperature remains lower on account of the absence of direct solar heating). Camouflaged butterflies in comparison fly rapidly in open sunny areas and have very high body temperatures. These flight differences lead to the aposematic butterflies having wings of greater aspect ratio, and thus affect the evolution even of the butterflies' morphology.

Batesian and Müllerian mimicry. Warningly colored species can be exploited as models for mimicry, just as can inedible objects such as thorns and leaves. Sometimes the term mimicry is restricted to such resemblances between edible species and actively defended and warningly colored models. Much is known about the evolution of this kind of mimicry in butterflies (and to a lesser extent, in bees and flies). If the mimic is entirely edible, the relationship is parasitic; the mimic benefits from the resemblance, but as every encounter with a mimic reduces the predator's aversion, the model suffers some increase in the rate of attack; such mimicry has traditionally been termed Batesian mimicry. Alternatively, the mimic may be almost or fully as defended as the model, leading to a mutualistic relationship known as Müllerian mimicry, in which both the model and the mimic species suffer a decreased rate of predation (compared with what would happen if each had a different color pattern).

The evolution of Müllerian mimicry has been a powerful force in shaping the appearance of tropical insects. In most areas of the tropics (at least within the forest habitats), the numerous species of warningly colored butterflies tend to share only five or six patterns (known as mimicry rings); this number is also known from tropical wasps, and from bumblebees in the temperate zone (where butterfly mimicry is much less extensive, probably because of a reduced number of species).

Müllerian rings in any one area are parasitized by the local Batesian mimics. They can frequently be distinguished from the Müllerian mimics by retaining their camouflage on the underside of their wings (the part seen when the butterfly is resting). Sometimes mimicry is confined to the female (so that the male retains an ancestral camouflaged or sexual display pattern). Occasionally, Batesian mimics can be distinguished from the Müllerian mimics by being polymorphic, that is, by exhibiting several different patterns within the one species, each pattern belonging to a different mimicry ring.

Evolution of mimicry. In general, the evolution of mimicry occurs in two distinct phases. In the first phase, a mutation of rather large effect produces a rough resemblance between the future mimetic species and its model. This occurs partly by happenstance (out of the great range of mutations which a species may produce, one will eventually resemble something else), and partly because similar processes of development occur in the embryogenesis of the pattern: all butterflies probably potentially share certain physicochemical processes during the laying down of the pattern on their wings. If the rough resemblance is close enough, the new mutation spreads.

During the second phase, the pattern of the mimic gradually changes to resemble the model ever more closely, as the fine discriminatory powers of the predators eliminate poor resemblances at the expense of better ones. This modification phase is thought to involve the establishment of genes of individually rather small effect on the pattern, which cumulatively refine the resemblance. During this phase, the Batesian mimic places the model at a disadvantage; the model therefore evolves away from the pattern of the mimic. It is, however, inevitably built into the dynamics of the system that the mimic can evolve faster than the model; this is the only case in which scientists know in evolutionary terms why the host is unable to shake off its parasite.

Evolution in these two phases occurs in both Batesian and Müllerian mimicry; Müllerian mimics join, or even switch between, mimicry rings in this fashion. It is possible that evolution in general is not always entirely gradual, but may involve small discontinuous changes. *See* ANIMAL EVOLUTION; MACROEVOLUTION; POPULATION ECOLOGY.

John R. G. Turner

Bibliography. W. B. Cott, *Adaptive Coloration in Animals*, 1940; M. Edmunds, *Defence in Animals*, 1974; E. B. Ford, *Ecological Genetics*, 4th ed., 1975; P. M. Sheppard, *Natural Selection and Heredity*, 4th ed., 1975; J. R. G. Turner, Butterfly mimicry: The genetical evolution of an adaptation, *Evol. Biol.*, 10:163–206, 1977; R. L. Vane-Wright and P. R. Ackery (eds.), *The Biology of Butterflies*, 1984.

Protein

Polymeric compounds made up of various monomeric units called amino acids. Amino acids are joined together in a chain by peptide (amide) bonds between the α-carboxyl groups and the α-amino groups of adjacent amino acids. The first amino acid in a protein usually contains a free

α-amino group, as shown below. Proteins generally

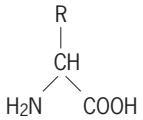

Amino acid

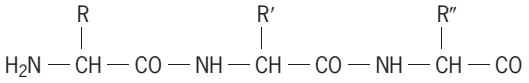

Peptide

contain from 50 to 1000 amino acids per chain. Small chains of up to 50 amino acids are usually referred to as peptides or polypeptides. *See* AMINO ACIDS; PEPTIDE.

Occurrence. Proteins are central to the processes of life. They are fundamental components of all biological systems, performing a wide variety of structural and functional roles. For example, proteins are primary constituents of structures such as hair, tendons, muscle, skin, and cartilage. Several hormones, such as insulin and growth hormone, are proteins. The substances responsible for oxygen and electron transport (hemoglobin and cytochromes, respectively) are conjugated proteins that contain a metalloporphyrin as the prosthetic group. Chromosomes are highly complex nucleoproteins, that is, proteins conjugated with nucleic acids. Viruses are also nucleoprotein in nature.

Protein enzymes are the catalysts of nearly all biochemical transformations. Pepsin and rennin are examples of digestive enzymes involved in breaking down food. DNA polymerases are enzymes that duplicate DNA for cell division, other enzymes are needed to repair damaged DNA, and gene expression is carried out by RNA polymerases. The chemical reactions used in metabolic pathways for carbohydrates (citric acid cycle), lipids, amino acids, and energy production (oxidative phosphorylation) are catalyzed by protein enzymes. All living things contain proteins because they serve as the molecular tools and machines of life. *See* ENZYME.

Of the more than 200 amino acids that have been discovered either in the free state or in small peptides, only 20 amino acids, or derivatives of these 20, are present in mammalian proteins. They are alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine. Each has a unique side chain (represented by R, R′, R″ in the peptide structure above) that defines the chemical properties of the amino acid. Except for glycine, amino acids have an inherent asymmetry referred to as chirality. In all natural proteins, they are of the L-configuration. Amino acids with the mirror-image configuration are D-amino acids. Both D- and L-amino acids occur in the free state and in small peptides. The amino acids can be grouped according to similarities in size, shape, charge, aromaticity, and polarity. Amino acids with nonpolar, hydrophobic side chains are alanine, isoleucine, leucine, methionine, tryptophan, phenylalanine, proline, and valine. Another group with polar, uncharged side chains is asparagine, cysteine, glutamine, serine, threonine, and tyrosine. Finally, five amino acids have side chains that can be charged at physiological pH values. Arginine, lysine, and histidine can be protonated to bear a positive charge, whereas aspartic acid and glutamic acid can be deprotonated to bear a negative charge. Commonly occurring derivatives of these 20 amino acids are cystine, hydroxyproline, and hydroxylysine. Cystine results from the oxidation of two cysteines to form a disulfide bridge. Hydroxyproline and hydroxylysine are formed by the enzymatic hydroxylation of specific proline and lysine, respectively. Modification of these two amino acids takes place while the polypeptide chain of collagen is being synthesized on the ribosome. Other examples of biological modification of parent amino acids are *O*-phosphoserine in phosvitin, ^{3}*N*-methylhistidine in actin, and ε-*N*-methyllysine in histones.

Biosynthesis. The linear arrangement of amino acids in a protein is termed its sequence (primary structure). The sequence is highly specific and characteristic for each particular protein. It is determined by the DNA sequence of each protein's gene that is expressed in the form of messenger RNA. Elucidation of the mechanism by which proteins are built up from free amino acids has been one of the key problems of molecular biology. *See* DEOXYRIBONUCLEIC ACID (DNA); MOLECULAR BIOLOGY; RIBONUCLEIC ACID (RNA).

Although a few proteins such as collagen are stable indefinitely in adulthood, most body proteins are in a continual process of degradation and synthesis (turnover). For example, the half-life of serum proteins in humans is about 10 days. Each amino acid is activated by adenosine triphosphate (ATP) by a specific enzyme called an aminoacyl-tRNA synthetase. These enzymes establish the relationship between specific nucleotides and amino acids, which is the basis for the genetic code. This activated amino acid is covalently attached to a transfer RNA that contains a triplet of nucleotides (anticodon) at one end. Each triplet of nucleotides is unique for one amino acid. Protein synthesis, the joining of amino acids to form polypeptides, takes place on the ribosome, a large ribonucleoprotein complex that translates mRNA into proteins. On the ribosome, aminoacyl-tRNA binds specifically to a segment on mRNA through bonding between the anticodon on the tRNA and a complementary codon of three nucleotides on the mRNA. Two activated aminoacyl-tRNA complexes bind adjacent mRNA codons and then react together to form a peptide bond. The process continues until the synthesis of the entire protein is complete. *See* RIBOSOMES.

Structure. The task of determining the sequences of proteins was a major preoccupation of many scientists in the twentieth century. Frederick Sanger was awarded a Nobel Prize in 1958 for determining the first protein sequence—insulin, a 51-amino acid protein. Sanger's work proved for the first time

that a protein was a pure chemical entity comprising a single amino acid sequence and not a collection of closely related amino acid sequences (sometimes referred to as statistical proteins). Later developments in DNA sequencing led to a Nobel Prize for Walter Gilbert and Frederick Sanger (his second). The relative ease of DNA sequencing combined with many additional technical innovations has permitted the whole-scale sequencing of genomes. Thus, most protein sequences are now deduced from the corresponding gene sequence. With this wealth of sequence information, researchers in the fields of protein chemistry and molecular and cellular biology are focused on understanding the functional roles and properties of proteins.

Proteins of similar function from different species have a common genetic origin and thus have related sequences. When differences exist in the sequence, changes are usually conservative, such as the replacement of one small hydrophobic amino acid by a different small hydrophobic amino acid. The degree of similarity between one protein in two different animals has been used to evaluate the relative genetic distance of the two animals. This type of comparison was made for the β-globin chain of hemoglobin, the blood protein that carries oxygen and carbon dioxide, for chimpanzees, monkeys, orangutans, gorillas, and humans. The relatively close relationship for all these animals is reflected by their nearly identical amino acid sequences for β-globin. However, some amino acid differences exist between β-globin chains. For example, chimpanzee β-globin is identical to gorilla β-globin at 146 of the 147 amino acids. The single difference is an arginine in one sequence and a lysine in the other. In this example, arginine and lysine are both positively charged amino acids and β-globin functions with either. Thus, this change is conservative. *See* HEMOGLOBIN; PROTEINS, EVOLUTION OF.

Nonconservative changes can alter the ability of a protein to function, as is evident by the single amino acid difference between normal and sickle human β-globin. Sickle β-globin contains valine at a position normally occupied by glutamic acid. This single amino acid change alters the overall structure of hemoglobin which, in turn, alters the shape of the red blood cell. When a person inherits the gene for sickle β-globin from both parents, it results in the disease sickle cell anemia. *See* SICKLE CELL DISEASE.

Proteins are not stretched-out polymers; rather, each adopts a specific extended or compact and organized structure called its native structure. It is still not completely understood how proteins "fold" into their structures, nor can we accurately predict the complete structure from its amino acid sequence. Pioneering studies that won Christian Anfinsen the Nobel Prize in 1972 showed that proteins fold themselves—that all the information needed to reach the correct three-dimensional structure is contained in the linear sequence of amino acids. Thus, after synthesis on the ribosome, a polypeptide chain folds to its native structure. In order to prevent the misfolding of certain proteins, chaperone proteins

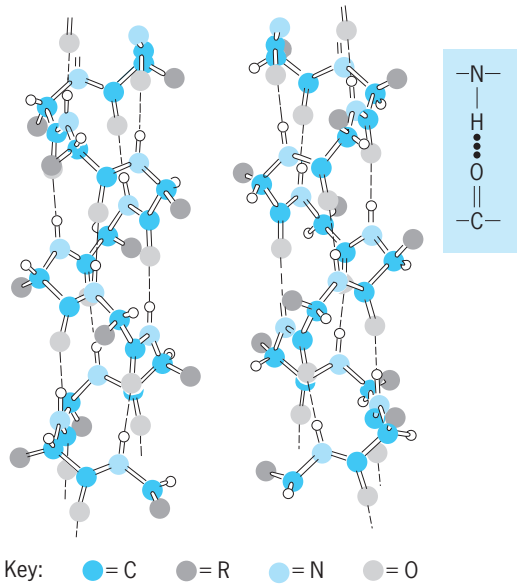

Fig. 1. The α-helix proposed by L. Pauling and R. B. Corey. The repeating —NH—CO—C units form the backbone which spirals up in a left-handed or a right-handed fashion. (The right-handed configuration is adopted in natural proteins with L-amino acids.) Hydrogen bonds are indicated by the broken lines. Note that the side chains (R) are all directed out from the helix. The pitch of the helix (axial displacement per complete turn of the helix) is about 0.54 nm, and there are 3.7 amino acids contained in one complete turn. The structure at the right shows the hydrogen bond. (*After J. T. Edsall and J. Wyman, Biophysical Chemistry, vol. 1, Academic Press, 1958*)

intervene to facilitate proper folding. Proteins that become misfolded or that have lost their native structure pose a major problem in cells. Diseases such as cystic fibrosis and Alzheimer's disease are characterized by misfolded proteins. In cystic fibrosis, a mutation in the chloride ion channel protein prevents the final stages of protein folding and leads to deficiencies in ion transport across cell membranes. In Alzheimer's, a protein known as amyloid precursor protein takes on a misfolded shape that causes the protein to clump together, ultimately forming plaques in the brain. *See* MOLECULAR CHAPERONE.

The polypeptide backbone of protein can fold in several ways by means of hydrogen bonds between the carbonyl oxygen and the amide nitrogen (**Fig. 1**, inset). Structural elements created by backbone hydrogen bonding interactions in the polypeptide are called secondary structures and include features like α-helices, β-sheets, and turns. In helices, the backbone is coiled in a regular fashion that brings peptide bonds separated by several amino acids into close spatial approximation. The stability of a helix is attributed to hydrogen bonds between these peptide bonds. The configuration of an α-helix, which is the most abundant secondary structural element in globular proteins, is described by a backbone amide hydrogen of one peptide unit bonded to a carbonyl oxygen (four amino acids away) that is part of another peptide unit (Fig. 1). Helical structures are often made up of peptide segments of around 10 amino acids, which corresponds to about three turns of helix. In keratin, the fibrous protein found

in hair and nails, much longer helices are formed and pack with helices from other keratin polypeptides to give tensile strength. *See* HYDROGEN BOND.

In addition to α-helices, polypeptides form β-sheet structures that are made of two or more segments that run parallel or antiparallel to each other and connect through backbone hydrogen bonds. The amino acid side chains stick out on alternating sides of the β-sheet surface. The human porin protein is made almost entirely of β-strands arranged to form a central pore that allows sugars, amino acids, ions, and other small metabolites in and out of cells.

The third level of folding in a protein (tertiary structure) comes through interactions between different parts of the molecule. At this level of structure, various secondary structure elements are brought together and interact through many types of associations. Hydrogen bonds between different amino acids and peptide bonds, hydrophobic interactions between nonpolar side chains of amino acids such as phenylalanine and leucine, and salt bridges such as those between positively charged lysyl side chains and negatively charged aspartyl side chains, all contribute to the tertiary structure specific to a given protein. Disulfide bridges formed between two cysteines at different linear locations in the molecule can stabilize parts of a three-dimensional structure by introducing a covalent bond as a cross-link. The result is a unique architecture that is predetermined by the particular sequence of amino acids in the protein.

Finally, some proteins contain more than one polypeptide chain per molecule. This feature is referred to as the quarternary structure. There is usually a high degree of interaction between each subunit, for example, between the α- and β-globin chains of hemoglobin. The tetrameric structure of hemoglobin exemplifies how four individual polypeptide chains (two α- and two β-chains) function cooperatively. The binding of oxygen to one subunit causes a conformational change in the protein that facilitates the binding of oxygen to the other subunits. Thus, the cooperative binding of oxygen to hemoglobin is coordinated to facilitate oxygen binding in the lungs and its release in the tissues where oxygen pressure is slightly lower.

Among proteins with a known three-dimensional structure, there are a few recurring structures that predominate. These structural superfolds include closely related proteins, which have a high level of identical amino acids and perform similar functions, as well as unrelated proteins, which have independent amino acid sequences and function despite overall similarity in architecture. Nine superfolds dominate the known protein structures. These are the globin, trefoil, up-down, immunoglobulin, $\alpha\beta$ sandwich, jelly roll, doubly wound, UB $\alpha\beta$ roll, and TIM barrel folds (**Fig. 2**). Hemoglobin and myoglobin are related and are representative of members of the globin fold family, which is characterized by eight α-helices arranged in a Greek key motif (the topological signature of many β-barreis and a majority of β-sandwich structures). Antibodies contain

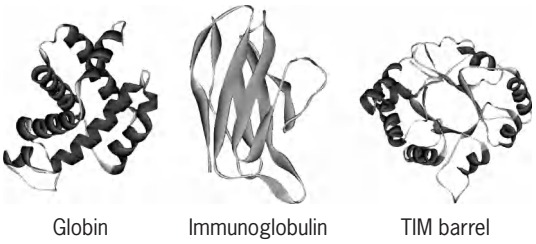

Fig. 2. Three commonly occurring protein folds.

several repeating segments that have eight-stranded β-sheet structures belonging to the immunoglobulin fold family. Several enzymes that carry out the chemical reactions of metabolism are members of the TIM barrel fold family. This fold is made of eight β-strand/α-helix structures wrapped into a barrel. The namesake enzyme TIM (triose-phosphate isomerase) is used in the breakdown of sugars. Another member of the TIM barrel fold family is ribulose 1,5-bisphosphate carboxylase oxygenase (Rubisco), a plant protein that fixes carbon dioxide from the air into organic carbon needed to sustain life and is the most abundant protein on Earth. While many proteins are found to adopt one of the superfolds, many others have a unique architecture.

Preparation. Many of the structural and functional studies of proteins are performed with purified preparations of proteins. Historically, proteins were obtained from their tissue of origin, but modern DNA technology allows proteins to be produced in high yield in bacteria, yeast, or isolated cells from insects and mammals. The isolation of a protein from any of these sources is accomplished by a variety of techniques. The primary concern is that the protein be isolated in its native form—in other words, not denatured or degraded during preparation.

All purification methods seek to exploit differences in solubility, charge, size, and resin-binding specificity in order to enrich the solution for the desired protein. The first step in preparation is usually to mechanically disrupt the tissue or cells containing the protein. Next, any of several methods are applied, often sequentially, to purify the protein. A few of the commonly applied methods are described.

Salting out. High concentrations of neutral salt tend to precipitate proteins; the salt concentration necessary for such precipitation varies from protein to protein. The most commonly used salt for this purpose is ammonium sulfate, which has the advantages of high solubility at the low temperatures (around freezing temperature to room temperature) used to isolate proteins. By adding salt stepwise to a protein solution and separating the precipitate by centrifugation at each step, fractions are obtained in which the desired protein is concentrated in one of them.

Ion-exchange chromatography. Ion-exchange resins and modified cellulose derivatives are powerful tools for protein purification. A solid resin is chemically modified to introduce charged groups for anion (negative charge) or cation (positive charge) binding. Proteins, which typically contain many positive and negative

charges on their surfaces, will interact with a specific resin with different binding affinities. The strength of the protein–resin interaction can be modulated by altering the pH of the solution and thus the protonation state of the protein, as well as the ionic strength of the buffer. Buffers, mixed to produce a gradient of pH and ionic strength, are used to differentially bind and elute proteins from the resin. Ion-exchange chromatography has the capacity to handle gram quantities of proteins. Low-molecular-weight proteins, especially those with markedly acidic (positively charged) or basic (negatively charged) properties, such as ribonuclease, histones, and lysozyme, can be separated from most other proteins in a mixture in a single experiment on an ion-exchange column. *See* CHROMATOGRAPHY.

Affinity chromatography. The resolving power of ion-exchange chromatography can be remarkably enhanced by the covalent attachment of a small molecule or of a protein that has an affinity for the target. For example, streptavidin is a protein from *Streptomyces* bacteria that tightly binds biotin. It is readily purified by binding to biotin-conjugated resins, which has thus led to its use as a handle for purification of other proteins. Therefore, through protein engineering, streptavidin is covalently linked to a second protein of interest. The two linked proteins can be purified by biotin affinity chromatography. Immunoglobulins, or antibodies, are tightly bound by protein A, made by *Staphylococcaus* bacteria. Purification of immunoglobulins by affinity chromatography on protein A–coupled resin is an important method of isolating commercially produced immunoglobulins for medical therapies in diseases ranging from arthritis to cancer. In some cases, total purification in one step can be achieved with affinity chromatography. *See* PROTEIN ENGINEERING.

Countercurrent distribution. Countercurrent distribution is an effective system for purification that can separate peptides based on their differential solubilities in two immiscible liquids. However, the range of application has been narrow. The power of the method was shown by the successful separation of insulin into two components differing by only a single amino acid. This method has been much more widely utilized in fractionation of peptides than in fractionation of proteins. *See* EXTRACTION.

Preparative ultracentrifugation. Proteins of high molecular weight can be concentrated by applying a high gravitational field to the protein, contained in a tube, for a prolonged time. An ultracentrifuge produces a high gravitational field in which the protein sediments through the field, concentrating protein at the bottom of the tube. However, it is possible to achieve significant purification only for the heaviest or the lightest components in a mixture, and the degree of purification will depend on the range of sedimentation velocities represented. Solutions with densities less than that of the solvent allow proteins to float to the surface under the influence of a strong gravitational field. This principle has been successfully applied to the study of serum lipoproteins which, by virtue of their lipid content, have solution densities less than those of other, non-lipid-containing serum proteins. Sufficient salt is added to the serum to raise the solution densities of non-lipid-containing proteins above those of the lipoproteins. Centrifugation to equilibrium brings the lipoproteins to the surface, where they can be collected. This technique has been refined by the use of repeated centrifugation steps at progressively raised salt densities, allowing subfractions of lipoproteins of different solution density classes to be obtained on a preparative scale. *See* ULTRACENTRIFUGE.

Gel filtration. Gel filtration is a technique for the purification of proteins that separates mainly on the basis of molecular weight. Dextran beads, just micrometers in diameter, are used to exclude proteins of high molecular weight but to permit smaller proteins to diffuse in and out. With a column of such beads, a continuous redistribution and separation occurs as the protein mixture moves through the column. Beads of different porosities are available to separate proteins of different molecular weights. For example, with one type of dextran bead, proteins with molecular weights of about 10,000 can be easily separated from those whose molecular weights are greater than 50,000. *See* FILTRATION.

Analysis. Methods for protein analysis include amino acid analysis, protein sequencing, mass spectrometry, isoelectric focusing, gel electrophoresis, and specific activity.

Amino acid analysis. Because of the relatively high molecular weights of proteins, the classical methods of organic chemistry are not adequate to establish composition or structure. A more meaningful empirical formula for a protein is given in terms of the total number of amino acids and the relative proportion of each amino acid. Thus, a protein can be hydrolyzed in hydrochloric acid to liberate the individual amino acids, which can be quantitatively estimated by amino acid analysis. These data, together with an independent measurement of molecular weight (that can be converted into the total number of amino acids), yield the amino acid composition, that is, the number and kind of each amino acid. In those few cases in which a protein is known to lack a particular amino acid (for example, tryptophan is absent from ribonuclease, and isoleucine is absent from human adult hemoglobin), the amino acid composition can be used to estimate the purity of a sample.

Protein sequencing. Developed by Pehr Edman in the 1950s, sequential degradation (one amino acid at a time from one end of the chain) of proteins and peptides was used for decades to determine the sequence. The procedure was done by coupling the free N-terminal α-amino acid with a chemical (particularly phenylisothiocyanate) that removes the amino acid (under appropriate conditions) and leaves the remainder of the polypeptide chain intact. It is then possible to identify the cleaved derivative of the first amino acid and repeat the process on each truncated peptide that is successively generated.

Protein sequencing is rarely used anymore to determine the complete sequence of amino acids

in a protein, because deduced protein sequences are available from the DNA sequences of many organisms. However, it remains a frequently used method to determine protein identity. Identity can be uniquely established from the sequence of the first six to ten amino acids and comparison of that sequence to those predicted from the sequences of the genes in the organism from which the protein was isolated.

Not all proteins generate useful sequence information from the amino terminus, based on reaction with phenylisothiocyanate. Sometimes the amino terminus of some proteins is blocked, or covalently modified, so the Edman method cannot work. To circumvent this problem, proteins can be randomly broken down into smaller polypeptides or peptides. The sequences of the individual peptides are then determined. However, the sequences of these pieces have to be combined in the correct order, and this can be challenging. Alternatively, stepwise removal of amino acids from the carboxyl end by enzymes called carboxypeptidases is used to determine the amino acid sequence at the opposite end of the polypeptide. *See* PROTEIN DEGRADATION.

Mass spectrometry. Powerful new applications of mass spectrometry have catalyzed its use in biology. Mass spectrometry has become an important tool for protein sequencing and identification, and for localization of modifications that occur after the protein has been synthesized. Mass-spectrometric analysis yields the molecular weight of the protein. The masses of the peptides in a proteolytic digest of a protein can also be determined. Then by comparison to protein sequence databases, the protein can be identified.

Isoelectric focusing. The net charge of a protein is largely a function of the relative number of basic (lysine, histidine, and arginine) and dicarboxylic amino acids (aspartic and glutamic acids). This net charge strongly influences the solubility of a protein at different pH values because the solubility depends in part on the proportion of polar groups. When the hydrogen-ion concentration is high (low pH), the net charge is positive; when the hydrogen-ion concentration is low (high pH), the net charge is negative. The pH at which the net charge of the protein is zero is defined as the isoelectric point (pI). Isoelectric focusing—the focusing by electrophoresis of individual proteins according to their respective pI's—is usually carried out in a vertical column that contains a sucrose concentration gradient or in a semisolid phase of dextran. Synthetic polymers of amino acids (polyamino acids) of varying isoelectric points are added to the protein mixture that is to be resolved. In an electric field, each synthetic polymer migrates to the point where it has zero net charge. Thus, a stable gradient of these polyamino acids is established that, in effect, is a pH gradient. The added proteins then are separated according to the intrinsic pI (isoelectric point) of each one. *See* ISOELECTRIC POINT; PH.

Gel electrophoresis. Protein samples are analyzed for purity and size by various methods of polyacrylamide gel electrophoresis (PAGE). Under native conditions, that is, in the absence of denaturing agents, proteins migrate according to their charge density and size and shape. At slightly basic pH, most proteins carry a net negative charge. The lower the charge density (less charge per molecular mass), the slower a protein migrates. The frictional force of the gel matrix creates a sieving effect, retarding the movement of proteins according to size.

Under denaturing conditions, mixtures of proteins are separated by PAGE based on size. Sodium dodecyl sulfate (SDS) is added to denature the protein. (Regardless of the shape of the native, undenatured protein, all proteins assume a rod shape when denatured. As a consequence, the shape differences between proteins are eliminated.) SDS is negatively charged and has an inherent tendency to stick to the protein backbone in an amount proportional to a protein's molecular weight. Thus, each SDS–protein complex migrates through the gel according to molecular weight. The size of a protein is calculated from the distance it migrates in the gel, when compared to proteins of known molecular mass.

Two-dimensional electrophoresis combines the resolving power of isoelectric focusing and denaturing PAGE to separate complex mixtures of proteins. To separate proteins that migrate at a single pI, SDS-PAGE is run at a right angle to the direction of the first (isoelectric focusing) electrophoresis. By these methods, thousands of proteins can be separated into a two-dimensional array of "spots" composed of single pure proteins. This technique is widely used in the field of research known as proteomics.

Specific activity. Function is the ultimate outcome of structure. Therefore, the function or activity of a protein should be measured whenever possible. The activity of a protein refers to the specific function of the native protein. In the case of an enzyme, the activity is the catalysis of a chemical transformation of substrate into product. Enzyme activities are usually reported in activity units such as micromoles of product formed per minute. The specific activity of an enzyme is the total activity divided by the amount of protein. During the purification of an enzyme, specific activity increases because contaminating proteins are removed.

For example, luciferase is an enzyme found in the light-producing organ of fireflies, which create light to signal other fireflies. The enzyme catalyzes a chemical reaction that converts luciferin, adenosine triphosphate, and oxygen (O_2) into oxyluciferin, adenosine monophosphate, pyrophosphate, and light. Thus, one way to measure the activity of firefly luciferase is to measure the amount of light produced.

A pure, completely native enzyme will have the highest specific activity. As the protein becomes degraded or denatured, the specific activity will be reduced. Mutations that affect (negatively or positively) the functioning of an enzyme can be assessed by measuring the specific activity and comparing it to that of the wild-type enzyme. Those mutations that disrupt activity are important because their locations

in the structure can identify the residues needed for catalysis.

Karla L. Ewalt; Paul Schimmel; Gertrude E. Perlmann; James M. Manning

Bibliography. M. C. Anson et al. (eds.), *Advances in Protein Chemistry*, vol. 40, 1990; C. F. Barbas III et al., Semisynthetic combinatorial antibody libraries: A chemical solution to the diversity problem, *Proc. Nat. Acad. Sci. USA*, 89:4457–4461, 1992; S. J. Benkovic, Catalytic antibodies, *Annu. Rev. Biochem.*, 61:29–54, 1992; R. A. Bradshaw et al. (eds.), *Molecular Biology of Intracellular Protein Sorting and Organelle Assembly*, 1988; J. R. Jacobsen et al., An efficient antibody-catalyzed aminoacylation reaction, *Science*, 256:365–367, 1992; P. Schimmel, Alanine transfer RNA synthetase: Structure-function relationships and molecular recognition of transfer RNA, *Adv. Enzymol. Related Areas Mol. Biol.*, 63:233–270, 1990; G. E. Schultz and R. H. Schirmer, *Principles of Protein Structure*, 1988; W. Voelter et al. (eds.), *Chemistry of Peptides and Proteins*, vol. 3, 1985.

Protein degradation

The enzymatic destruction of proteins. Like other macromolecular (large) components of an organism, proteins are in a dynamic state of synthesis (generation) and degradation (destruction, proteolysis). During proteolysis, the peptide bond that links the amino acids to one another is hydrolyzed, and free amino acids are released. The process is carried out by a group of enzymes called proteases. *See* ENZYME; PROTEIN.

Mechanisms. Distinct proteolytic mechanisms serve different physiological requirements. It is necessary to distinguish between destruction of "foreign" and "self" proteins. Foreign proteins contained in the diet are degraded "outside" the body, in the digestive tract. This is necessary in order to prevent a response of the immune system to the intact proteins. The released amino acids are then absorbed and serve to build the organism's self proteins. Self proteins can be divided into extracellular and intracellular. The two groups are degraded by different mechanisms. Extracellular proteins that in solution in the plasma (the blood fluid phase), such as blood coagulation factors, immunoglobulins, cargo carrier proteins (such as the low-density lipoprotein, LDL, that carries cholesterol; or transferrin that carries iron), and peptide hormones (such as insulin) are taken up by pinocytosis (engulfment of small droplets) or following binding to specific receptors. They are carried via a series of vesicles (endosomes) that fuse with lysosomes, where they are degraded by the proteases contained in these organelles. During this process, the extracellular proteins are never exposed to the intracellular environment (cytosol). Degradation of proteins in lysosomes is not specific, and all engulfed proteins are degraded at similar rates.

Several observations lead to the prediction that degradation of many intracellular proteins must be carried out by a completely different and nonlysosomal mechanism. In contrast to lysosomal degradation, the destruction of many cellular proteins is specific, and different proteins have half-life times that range from a few minutes (such as the tumor suppressor protein p53) to several days (such as the muscle proteins actin and myosin). Furthermore, the stability of certain proteins changes under different conditions. For example, cell-cycle regulators such as mitotic cyclins are stable throughout the cell cycle and are degraded only during the short mitotic time window, an event essential for the cell in order to exit mitosis. Other proteins are degraded only following a specific signal that is transmitted from the extracellular environment, and their degradation serves to turn off the signaling event. Also, inhibitors of lysosomal degradation do not affect these processes, a finding that suggests it occurs in the cytosol. Taken together, it has become clear that the proteolytic system involved in intracellular protein degradation must be distinct from lysosomal proteolysis. The discovery of the ubiquitin–proteasome proteolytic pathway has resolved this enigma.

Degradation of a protein via the ubiquitin pathway involves two steps: tagging of the substrate by covalent attachment of multiple ubiquitin molecules, and degradation of the tagged protein by the 26S proteasome with release of free and reusable ubiquitin. In addition to serving as a proteolytic signal, ubiquitination can regulate other processes. Depending on the character of the internal linkage between the ubiquitin moieties, polyubiquitination can also lead to activation of transcriptional regulators. Modification by a single moiety of ubiquitin can target proteins for degradation in the lysosome. In addition to ubiquitin, other modifiers that belong to the ubiquitin family of proteins have been discovered, and designated ubiquitin-like proteins, UbLs. Modification by these proteins can serve a variety of nonproteolytic functions, such as activation of enzymes or routing of the tagged proteins to their subcellular destination. Thus, ubiquitination has emerged as a critically important posttranslational modification that plays major roles in regulating a broad array of basic cellular processes. Aberrations in the system have been implicated in the pathogenesis of many diseases, converting the system into a platform for drug targeting (for a scheme of the ubiquitin system and its multiple functions, see the **illustration**).

Ubiquitination. Conjugation of ubiquitin, a 76-residue polypeptide, to the protein substrate proceeds via a three-step cascade mechanism catalyzed by these enzymes: E1, E2, and E3. E1 activates ubiquitin, and an E2 transfers it from E1 to the substrate that is bound to an E3. E3 enzymes catalyze the last critical step—covalent attachment of ubiquitin to the substrate. Following conjugation of the first ubiquitin moiety, other ubiquitin moieties are added to one another to generate a polyubiquitin chain. The chain is recognized by the downstream 26S proteasome. E3's play a key role in the ubiquitin system, as they serve as the specific substrate recognition elements

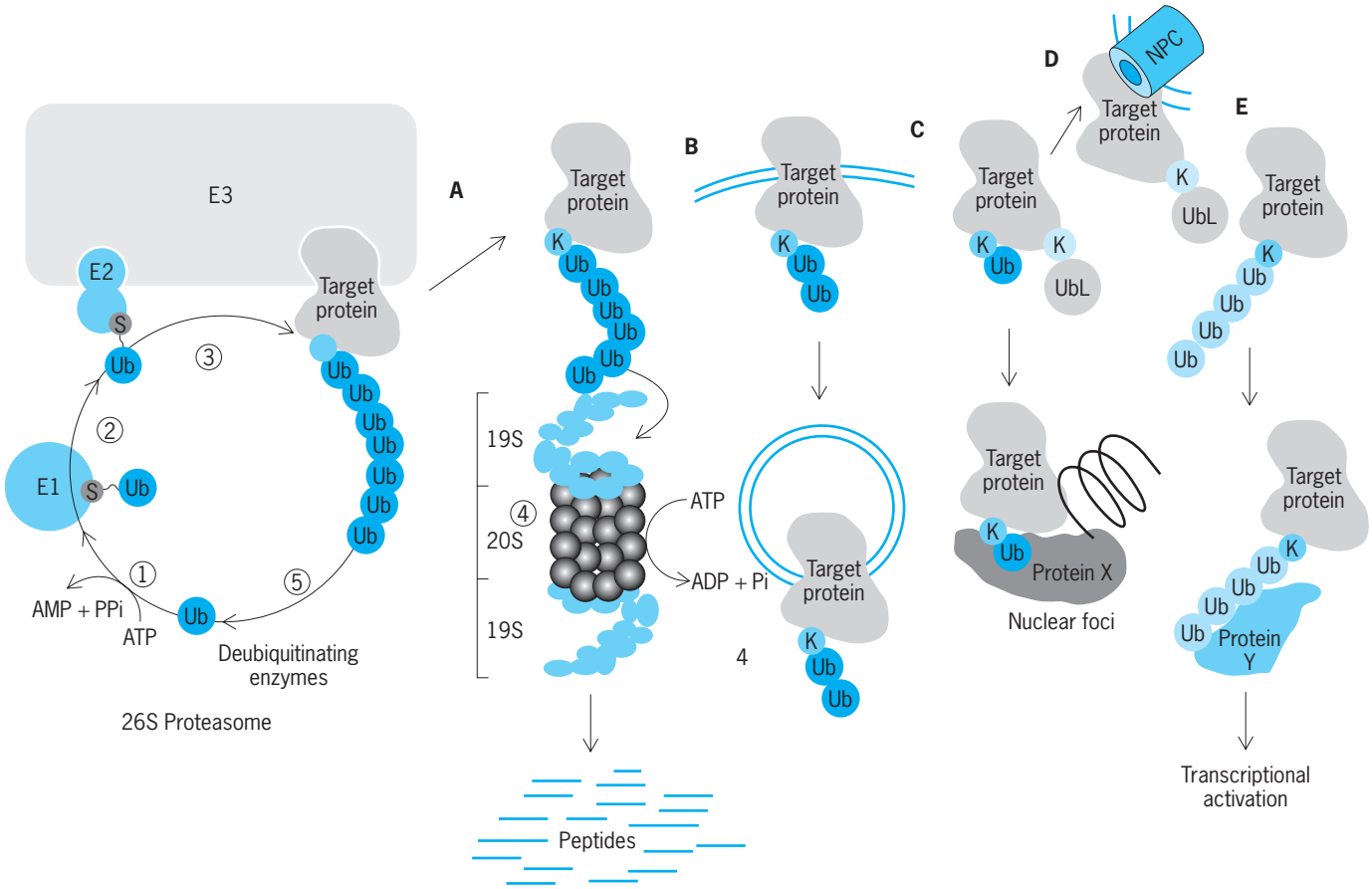

The ubiquitin system and its functions. (A) Proteasomal-dependent degradation. Ubiquitin is activated by the enzyme E1 (1), followed by its transfer to a ubiquitin-carrier protein E2 (2). E2 transfers the activated ubiquitin to the protein substrate that is bound specifically to a unique ubiquitin ligase E3 (3). Successive conjugation of ubiquitin moieties to one another (via Lys^{48} of the previous ubiquitin moiety) generates the polyubiquitin chain that serves as the degradation signal for the downstream 26S proteasome (4). The substrate is degraded to short peptides (4), and free and reusable ubiquitin is released by deubiquitinating enzymes (5). (B) Mono- or oligoubiquitination targets membrane proteins to degradation in the lysosome/vacuole. (C) Monoubiquitination or (D) a single modification by a UbL protein, SUMO for example, can target proteins to different subcellular destinations such as nuclear foci or the nuclear pore complex. (E) Generation of a polyubiquitin chain in which the moieties are bound to one another in a different manner than in the chain recognized by the proteasome (via Lys^{63} rather than Lys^{48}) can activate transcriptional regulators. Ub denotes ubiquitin, K denotes Lys, and S denotes Cys (to which ubiquitin binds on E1, E2, and certain E3's, prior to its transfer to the substrate).

of the system. Approximately 1000 different E3's have been identified in the human genome. UbLs are activated in a similar, albeit not an identical, mechanism.

Degradation. The proteasome is a multicatalytic protease that degrades polyubiquitinated, but not untagged, proteins. It is composed of two subcomplexes: a 20S core particle (CP) that carries the catalytic activity, and a 19S regulatory particle (RP). The 20S core particle is a barrel-shaped structure composed of four stacked rings: two identical outer seven-subunits-each α rings and two identical inner seven-subunits-each β rings. The catalytic sites are localized to some of the β subunits. Each extremity of the 20S can be capped by a 19S regulatory particle. One important function of the 19S regulatory particle is to recognize ubiquitinated proteins. Other functions of the 19S regulatory particle are to unfold the substrate and to open an orifice in the α ring to allow entry of the substrate into the proteolytic chamber. Following degradation, short peptides are generated, as well as reusable ubiquitin that is released by deubiquitinating enzymes; one of them is a subunit of the 19S regulatory particle.

Substrate recognition. A major problem is how the ubiquitin system recognizes its innumerable substrates in a specific manner. Why are certain proteins stable while others are extremely short-lived? Why are certain proteins degraded only at a particular time point along the cell cycle or following specific stimuli while they are stable under other conditions? Specificity of the ubiquitin system is determined by two distinct groups of proteins: E3's, and modifying enzymes and ancillary proteins. Within the ubiquitin system, substrates must be recognized and must bind to a specific E3 as a prerequisite to their ubiquitination. In a few cases, recognition is constitutive (the targeting motif can be a single amino acid residue such as the N-terminal residue). In most cases, however, the substrates must undergo a posttranslational modification such as phosphorylation (for example, the two neighboring Ser residues in the case of

IκBα, an inhibitor of nuclear factor κB) or oxidation (hydroxyproline in the hypoxia-inducible factor-1α subunit, HIF-1α), or must associate with a molecular chaperone in order to be recognized. In some instances, it is the E3 that must be "switched on" by undergoing posttranslational modification (phosphorylation is required, for example, to activate the proto-oncogene c-Cbl).

Functions of the ubiquitin system. The ubiquitin system is involved in regulating numerous basic cellular processes. Among them are cell division, differentiation and development, the cellular stress response, DNA repair, transcriptional regulation, and modulation of the immune and inflammatory responses. The list of cellular proteins that are targeted by ubiquitin is growing rapidly. The proteins include cell cycle regulators, tumor suppressors, transcriptional activators and their inhibitors, cell surface receptors, and endoplasmic reticulum proteins. Mutated/denatured/misfolded proteins are recognized specifically and removed efficiently. Thus, the system plays a key role in the cellular quality control and defense mechanisms.

Regulation of the ubiquitin system. The ubiquitin system can be regulated at the level of ubiquitination or of proteasome activity. Since conjugation and proteasomal degradation are required for a multitude of cellular functions, regulation must be specifically tuned. In a few cases, general rather then specific components of the pathway are modulated. For example, upregulation of the pathway is observed during massive degradation of skeletal muscle proteins that occurs following long fasting, and also under pathological conditions such as malignancy-induced cachexia, severe sepsis, metabolic acidosis, or following denervation. In most cases, however, regulation is specific, affecting defined substrate(s) and a distinct ligase.

Human diseases. Inactivation of a major enzyme such as E1 is obviously lethal. Yet, mutations or acquired changes in enzymes or recognition motifs in substrates that do not affect vital pathways or that affect the involved process only partially, may result in a broad array of diseases. The pathological states associated with aberrations in the system can be classified into two groups: those that result from loss of function—mutation in a ubiquitin system enzyme or target substrate that results in stabilization of the latter; and those that result from gain of function—abnormal or accelerated degradation of the substrate. A few examples can highlight the two states—accelerated and inhibited degradation.

Malignancies. Alterations in ubiquitination and deubiquitination reactions have been directly implicated in the etiology of several malignancies. In general, cancers can result from stabilization of oncoproteins or destabilization of tumor suppressors. *See* CANCER (MEDICINE).

In the case of uterine cervical carcinoma, the level of the tumor suppressor protein p53 is extremely low. Most of these malignancies are caused by high-risk strains of the human papillomavirus. It was shown that p53 is targeted for destruction by the virally encoded oncoprotein E6. Degradation is mediated by the naturally occurring E3, E6-associated protein (E6-AP), where E6 serves as an ancillary protein that allows recognition of p53. E6-AP will not recognize p53 in the absence of E6. In the absence of p53, the virus can integrate its genome into the host genome, an event that would have otherwise caused p53 to induce apoptosis. While the native substrates of E6-AP are still obscure, a mutation in the enzyme has been implicated as the cause of Angelman syndrome, which is characterized by mental retardation, seizures, out-of-context laughter, and abnormal gait.

β-Catenin plays an important role in differentiation of the colorectal epithelium, and aberrations in ubiquitin-mediated regulation of its levels play an important role in the development of colorectal tumors. Under resting conditions, β-catenin is phosphorylated, which leads to recruitment of the β-TrCP ubiquitin ligase ε3, followed by its ubiquitination and degradation. Signaling promotes dephosphorylation, stabilization, and subsequent activation of β-catenin via complex formation with other inactive subunits. β-Catenin interacts with the tumor suppressor APC (adenomatous polyposis coli) to generate a complex that appears to regulate its level. Aberrations in degradation of β-catenin that lead to its stabilization, accumulation, and subsequent oncogenic activation can result from two different mechanisms: mutations in the phosphorylation motif of the protein and mutations in the APC/axin machinery (APC and axin are interacting proteins that regulate the cytoplasmic levels of β-catenin by binding to it).

Aberrations in components of the ubiquitination machinery can also cause malignancies. Mutations in the von Hippel-Lindau (*VHL*) gene predisposes individuals to a wide range of highly vascularized malignancies such as renal cell carcinoma and cerebellar hemangioblastomas. The tumors arise from constitutive expression of the hypoxia inducible factor subunit HIF-1α, which results in overexpression of hypoxia-inducible gene products such as the vascular endothelial growth factor, the erythropoietin receptor, and glycolytic enzymes. pVHL is a subunit in a ubiquitin ligase that targets HIF-1α for ubiquitination and degradation. Under normoxic conditions, HIF-1α is hydroxylated specifically on Pro residue 564 to generate hydroxyproline. The hydroxylated Pro is recognized by the pVHL E3 complex that targets the molecule for ubiquitination and degradation. Under hypoxic conditions, HIF-1α is stable, as the efficiency of the hydroxylation reaction under these conditions is low. Loss of VHL function due to a mutation stabilizes HIF-α, which can explain the stimulation of vascular growth in these tumors.

Neurodegenerative disorders. Accumulation of ubiquitin conjugates of certain proteins in inclusion bodies have been reported in a broad array of chronic neurodegenerative diseases, such as Alzheimer's disease and Parkinson's disease. However, in all these cases a direct pathogenetic linkage to aberrations in

the ubiquitin system has not been established. *See* ALZHEIMER'S DISEASE; PARKINSON'S DISEASE.

The case of Parkinson's disease highlights the complexity of the involvement of the ubiquitin system in the pathogenesis of neurodegeneration. Mutations in several proteins have been described that link the system to the pathogenesis of the disease. One important player is Parkin, which is an E3 ubiquitin ligase. Mutations in the gene appear to be responsible for the pathogenesis of autosomal recessive juvenile parkinsonism (AR-JP), a rare disease by itself, but one of the most common familial forms of the disease. Parkin ubiquitinates and promotes the degradation of several substrates, such as CDCrel-1, a synaptic vesicle-enriched septin GTPase, or the Pael receptor, a putative G protein–coupled endoplasmic reticulum transmembrane polypeptide. It is possible that aberration in the degradation of one or more of these substrates that leads to its accumulation is neurotoxic and underlies the pathogenesis of AR-JP.

Cystic fibrosis. The cystic fibrosis gene encodes the cystic fibrosis transmembrane regulator (CFTR), which is a chloride channel. One frequent mutation in the channel is ΔF508. The mutation leads to a multisystem disorder characterized by chronic obstruction of airways and severe maldigestion due to exocrine pancreatic dysfunction. $CFTR^{\Delta F508}$ does not reach the cell surface and is retained misfolded in the endoplasmic reticulum, where it is subsequently degraded.

Immune and inflammatory disorders. As described above, the human papillomavirus evolved a mechanism for proteolytic removal of p53. Another interesting example of a virus that evolved a mechanism that exploits the ubiquitin system is the human cytomegalovirus. The virus encodes two endoplasmic reticulum proteins, US2 and US11, that downregulate the expression of MHC (major histocompatibility complex) class I heavy-chain molecules. The MHC molecules are synthesized and transported to the endoplasmic reticulum, but shortly thereafter, in cells expressing US2 or US11, they are transported back to the cytosol, deglycosylated, and degraded by the proteasome following ubiquitination. Thus, the virus inactivates the cell's ability to present antigens to cytotoxic T cells, which enables it to replicate and propagate.

Drug development. Because of the central role that the ubiquitin system plays in a broad array of basic cellular processes, development of drugs that modulate the activity of the system may be difficult. Inhibition of enzymes common to the entire pathway, such as the proteasome, may affect many processes nonspecifically, although a narrow window between beneficial effects and toxicity can be identified for a short-term treatment. Indeed, such a drug, Velcade, has been shown to be beneficial in multiple myeloma, a malignancy that affects bone marrow.

A completely different approach to drug targeting can be the development of molecules that interfere with the activity of specific E3's. Thus, a drug that will bind specifically to the p53 pocket in Mdm2, the E3 of p53 that is upregulated in many tumors and acts as an oncogene, will prevent association of the enzyme with p53, which results in upregulation of the tumor suppressor. Aaron Ciechanover; Kazuhiro Iwai

Bibliography. A. Ciechanover, From the lysosome to ubiquitin and the proteasome, *Nat. Rev. Mol. Cell Biol.*, 6:79–86, 2005; C. de Duve and R. Wattiaux, Functions of lysosomes, *Annu. Rev. Physiol.*, 28:435–492, 1966; M. H. Glickman and A. Ciechanover, The ubiquitin-proteasome pathway: Destruction for the sake of construction, *Physiol. Rev.*, 82:373–428, 2002; A. L. Goldberg and A. C. St. John, Intracellular protein degradation in mammalian and bacterial cells: Part 2, *Annu. Rev. Biochem.*, 45:747–803, 1976; G. E. Mortimore and A. R. Poso, Intracellular protein catabolism and its control during nutrient deprivation and supply, *Annu. Rev. Nutr.*, 7:539–564, 1987; R. T. Schimke and D. Doyle, Control of enzyme levels in animal tissues, *Annu. Rev. Biochem.*, 39:929–976, 1970; R. Schoenheimer, The Dynamic State of Body Constituents, Harvard University Press, Cambridge, MA, 1942.

Protein engineering

The design and construction of new proteins or enzymes with novel or desired functions, through the modification of amino acid sequences using recombinant DNA (deoxyribonucleic acid) technology. The sizes and three-dimensional conformations of protein molecules are also manipulated by protein engineering. The basic techniques of genetic engineering are used to alter the genes that encode proteins, generating proteins with novel activities or properties. Such manipulations are frequently used to discover structure-function relationships, as well as to alter the activity, stability, localization, and structure of proteins. *See* GENETIC ENGINEERING; PROTEIN.

Point mutants. Many subtle variations in a particular protein can be generated by making amino acid replacements at specific positions in the polypeptide sequence. Each protein is unique by virtue of the sequence of its amino acids. At any position in the sequence, an amino acid can be replaced by another to generate a mutant protein that may have different characteristics by virtue of the single replaced amino acid. For example, pancreatic ribonuclease A is an enzyme comprising 124 amino acids that cleaves the covalent bonds that join ribonucleic acids (RNA). If at position 119 in the sequence the naturally occurring histidine is replaced with an alanine, the mutant protein is referred to as a histidine 119 → alanine (H119A) mutant of ribonuclease A. This mutant protein is expected to have little or no biological activity, because histidine 119 is important for that activity. Other mutations have very little effect on their proteins. This is particularly true when the amino acid being changed is substituted with other closely related amino acids and when the amino acid is not conserved in the same protein found in other organisms. Typically, site-directed protein engineering targets amino acids that are involved in a particular biological activity. *See* AMINO ACID; MUTATION.

Because there are 20 different amino acids, 20

different variants of ribonuclease A can be created just by having a set that differs only by the amino acid at position 119. If changes are also made at another position, for example, position 41 (where normally there is a lysine), then in principle there are 400 different variants that can be created by making all possible combinations of substitutions at just these two positions in the sequence. As more positions are varied, the number of combinations becomes enormous so that, with just six different positions subjected to all possible variations, 64 million different proteins can be generated. These proteins will share the identical sequence, except at the six varied positions.

Deletion mutants. In addition to the substitution of amino acids at specific positions, amino acids can be deleted from the sequence, either individually or in groups. These proteins are referred to as deletion mutants. Deletion mutants may or may not be missing one or more functions or properties of the full, naturally occurring protein. Proteins from eukaryotes, such as mammals, plants, and fungi, can have very long amino acid sequences and as a consequence tend to be organized as several modular protein domains linked together. Deletion mutants have been a useful way to create smaller proteins that contain only one or a few of these domains so that the individual properties can be studied.

Hybrid/fusion proteins. Protein sequences can be joined or fused to that of another protein. The resulting protein is called a hybrid, fusion, or chimeric protein, which generally has characteristics that combine those of each of the joined partners. Protein fusions have been extensively used to study interactions between two or more proteins. For example, an application of protein fusion methods called the yeast two-hybrid screen was developed to identify proteins that interact with each other (see **illustration**). The system was developed by separating a yeast transcription-activating protein into two functional domains. The first domain of the transcription-activating protein is fused to a protein that is being studied for interactions. The second domain of the transcription-activating protein is fused to a collection of proteins encoded in a protein library. The first fusion protein and one from the protein library are expressed together in yeast. When the two fusion proteins interact, they can bring together the two halves of the transcription activator and turn on the expression of a reporter gene (see illustration). Another example is the fusion of a protein to an inherently fluorescent protein, such as the green fluorescent protein (GFP) from the jellyfish *Aequorea victoria*. Fluorescent fusion proteins have been used to study the location, movement, appearance, and degradation of proteins in living cells. *See* FLUORESCENCE.

Another common application for fusion proteins is to facilitate purification by affinity chromatography. A short sequence of amino acids or a protein domain is fused to one end of the polypeptide. Six histidines can be used to purify a protein by affinity to nickel. The Fc region of an antibody (region of an antibody molecule that binds to cell-surface antibody receptors) can be fused to a protein to purify it by affinity to protein A, a protein from bacteria that specifically binds antibodies. *See* LIQUID CHROMATOGRAPHY.

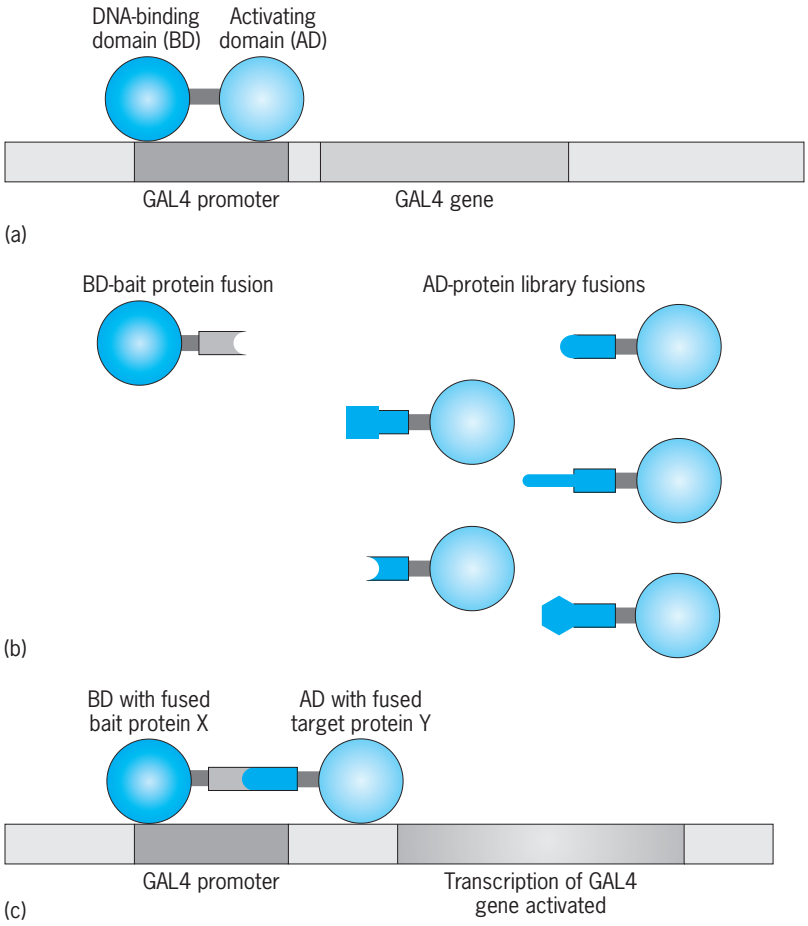

Yeast-two-hybrid screen. (*a*) The yeast transcription-activating protein GAL4 has two separate domains: a DNA-binding domain (BD), which binds the promoter region of the GAL4 gene, and a transcription-activating domain (AD), which stimulates production of the GAL4 RNA transcript. (*b*) These two domains are modular and can be fused onto other proteins while retaining their function. This forms the basis of the yeast two-hybrid assay (screen). The first hybrid, or fusion, protein consists of a protein of interest (bait protein) fused to the GAL4 BD. The second hybrid is a collection of fusion proteins made between the GAL4 AD and a library of many proteins. (*c*) If a protein from the library is able to interact with the first fusion protein, visualized in the figure as a protein with a complementary shape, the GAL4 transcription factor is reconstituted, and transcription of the gene is activated. Proteins that do not interact, those with noncomplementary shapes, fail to activate the system.

Novel activities or properties. Point and deletion mutants and hybrid proteins are constructed to obtain polypeptides with new properties. These proteins are either created individually, by site-directed mutagenesis, or they are generated as a large pool or library of millions of variants. The library is then screened or subjected to a special selection procedure to obtain the protein or proteins with the desired characteristics.

Proteins generally have stable and unique three-dimensional structures and are active at physiological temperatures of 37°C (98.6°F). They usually lose their three-dimensional shapes as the temperature is raised more than 5 or 10° above 37°C and, as a consequence, lose their biological activities. In some instances, a protein may be cold-sensitive; that is, it may lose its conformation and activity at lower temperatures. The effect of temperature on protein

stability can be modified by protein engineering, in particular, by introducing amino acid replacements that enhance or destabilize the molecular packing interactions in the core of the protein structure, such as has been done with the enzymes T4 lysozyme and staphylococcal nuclease.

Protein engineering has been used to produce therapeutic proteins with improved properties such as increased solubility and stability. For example, insulin was engineered through mutagenesis to create monomeric forms that are fast acting (insulin lispro and insulin aspart). Conversely, another form (insulin glargine) was created by mutagenesis to precipitate upon injection and give a sustained release of insulin. Mutation of a free cysteine in aldesleukin (a synthetic version of interleukin-2 used to treat some forms of cancer) or interferons beta-1b (used to treat relapsing-remitting multiple sclerosis) was used to produce therapeutics with decreased aggregation. *See* CANCER (MEDICINE); INSULIN; INTERFERON; INTERLEUKIN; MULTIPLE SCLEROSIS.

Proteins can be engineered to acquire new biological activities. For example, a catalytic antibody is a variant of an antibody. Antibodies are proteins that normally bind to a specific molecule but do not alter the bound molecule in any way. A catalytic antibody is one which has been changed by mutations to have a novel sequence that folds into a structure that catalyzes a specific reaction (such as amide bond formation, ester hydrolysis, and decarboxylation). Catalytic antibodies function like enzymes, and are created to catalyze reactions for which there are no naturally occurring enzymes. Fifty or more different reactions have been catalyzed by the action of catalytic antibodies that were obtained individually by methods of protein engineering. *See* ANTIBODY; CATALYTIC ANTIBODY.

Structure-function relationships. With over 180 genomes sequenced, including human and mouse, the amino acid sequence of a particular protein is now generally available from many different organisms. This advance during the past decade has created a situation in which protein sequences are available even when there is no information on the biological activity of the protein. Thus, the focus has shifted from sequencing to understanding the function of all of these proteins. Techniques such as the yeast two-hybrid system have been applied to identify protein interactions, and crystallographic structures are being determined to elucidate biological information about all the proteins in an entire organism. Still, the chemical basis for activity is often not completely understood, even after the biological activity and high-resolution structure have been determined. The construction of mutant proteins can elucidate the role of a particular amino acid at a specific position in the sequence. *See* GENE.

Protein engineering has been applied to understand how enzymes catalyze reactions, from identifying which amino acids are essential for catalysis to analyzing how amino acid changes alter particular aspects of a reaction such as substrate specificity. Antibiotics such as erythromycin are made by large multidomain proteins called polyketide synthases. Site-directed mutagenesis has been used to modify the substrate specificity of one polyketide synthase reaction so that the product contains a malonate unit, whereas the product of the original enzyme contained a methylmalonate unit. In addition to site-directed mutagenesis, the order of the polyketide synthase domains have been shuffled to create proteins that could catalyze the synthesis of new antibiotics. *See* ANTIBIOTIC; ENZYME.

An extension of site-directed mutagenesis allows nonnatural amino acids to be incorporated into proteins. Nonnatural amino acids are not naturally encoded by the genome, but instead include a wide variety of amino acids that are present in cells or produced by synthetic methods. This protein engineering method allows the chemical properties of a particular amino acid to be changed beyond that naturally encoded by a gene. For example, nonnatural amino acids incorporated into proteins may contain sugars, nucleophiles (electron donors), electrophiles (electron acceptors), crosslinkers (a chemical compound that forms covalent bonds between adjacent polymer chains that lock the chains in place), or altered shapes and sizes. Detailed studies on ion channels with over 60 nonnatural amino acids incorporated illustrate the chemical power of this approach.

Minimalist proteins. Proteins are usually large molecules composed of several hundred or even more than a thousand amino acids. Yet, the portion of the protein responsible for a specific biological activity is usually concentrated in a small part of the structure. In these circumstances, there are advantages to creating a minimalist protein having enough of the structure to retain the desired activity. A smaller protein is often superior for analyzing and manipulating structure-function relationships, and the reduced cost of materials for manufacturing a small protein can also be an important motivation for reducing the size.

To obtain minimalist proteins with novel properties, libraries with large numbers of different amino acid sequences have been created. These sequences are often so short (less than 50 amino acids) that they are referred to as peptides, oligopeptides, or polypeptides, and not as proteins. (A protein is regarded as a long polypeptide, generally at least over 80 amino acids in length and usually much longer.) Those with novel properties, such as an ability to bind to a specific ligand (a unique molecule that is usually small) are then selected. In these instances, the sequences of the peptides are created de novo and are not based on that of a naturally occurring protein. *See* GENE; GENE ACTION; GENETIC CODE; PEPTIDE.

Karla L. Ewalt; Paul Schimmel

Bibliography. J. A. Branningan and A. J. Wilkinson, Protein engineering 20 years on, *Nature Rev. Mol. Cell Biol.*, 3:964–970, 2002; J. Lippincott-Schwartz and G. H. Patterson, Development and use of fluorescent protein markers in living cells, *Science*, 300:87–91, 2003; S. A. Marshall et al., Rational design and engineering of therapeutic proteins, *Drug Discovery Today*, 8:212–221, 2003; P. G. Schultz et al., The

chemistry of the antibody molecule, *Angew. Chem. Int. Ed. Engl.*, 41:4427–4437, 2002; L. Wang and P. G. Schultz, Expanding the genetic code, *Chem. Commun.*, 1:1–11, 2002.

Protein folding

The process by which a long, flexible, unbranched chain of amino acids (the unfolded state of a protein) spontaneously changes shape to form a stable three-dimensional conformation (the native or folded state). This specific folding process is required for the protein's biological function. The sequential arrangement of amino acids is encoded by the genetic material, deoxyribonucleic acid (DNA) [**Fig. 1**]. The 20 most frequently used amino acids have a region in common that forms the covalent peptide bonds in the chain, termed the backbone, and a region that is chemically distinct, identified as the side chain. Amino acid side chains vary in size and shape, as well as the propensity to interact with water, oil, and ions. As a result, the side chains are differentially able to participate in noncovalent chemical bonds (hydrophobic interactions, electrostatic interactions, and hydrogen bonds) with each other and with the solvent. In addition, the amino acid cysteine can form covalent disulfide bonds with other cysteines. Together, this chemical interaction code determines the network of bonds that drives protein folding, stabilizes the native structure, and mediates protein function for any given sequence of amino acids.

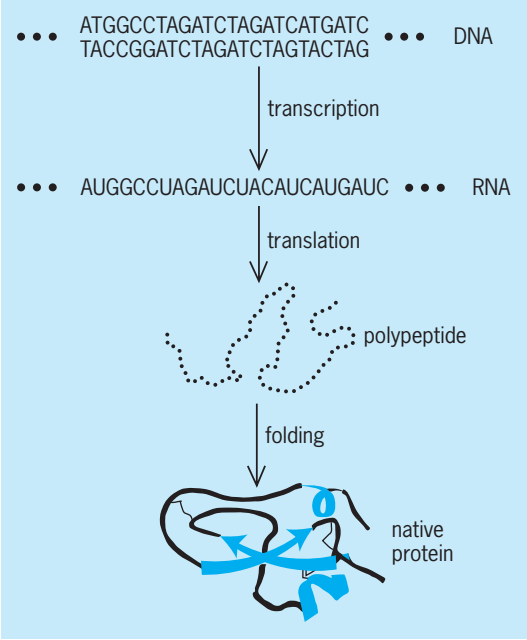

Fig. 1. An integrated view of protein production with folding as the terminal step. The instructions for building proteins are stored in the DNA sequence, which is transcribed into RNA and translated into an amino acid chain. The varying chemical properties of amino acid side chains drive protein folding. The final structure is characterized on different levels: primary structure is the sequence of amino acids; the local structure depicted below as arrows (β-strands) and coils (α-helix) is termed secondary structure; and the arrangement of secondary structural elements is tertiary structure. Many proteins consist of multiple polypeptide chains that work together as a single unit; the number and arrangement of these chains define the quaternary structure for these proteins. Some proteins must noncovalently bind heterologous chemical groups in order to complete the folding process. These groups (many of which are dietary vitamins and minerals) stabilize the protein's structure and may expand the chemical reactions the protein can perform. In contrast, several amino acids in a previously folded protein can be modified by covalent addition of chemical moieties (for example, sugars, lipids, phosphate or acetyl groups) that may alter the structure and/or function of the protein.

Since a single protein chain can consist of thousands of amino acids, and 20 different amino acids are commonly used as assembly units, the number of possible protein sequences and configurations is immense. The function of a protein is determined by its three-dimensional structure, and hence protein functions are diverse: among many functions, these complex molecules act as efficient and specific biological machines to catalyze biochemical reactions for metabolism, provide structural support, regulate biological processes, and transport other molecules. The process of protein folding places the protein's atoms in the correct three-dimensional arrangement required for that protein's unique function. Loss of the protein's native structure through denaturation involves a simultaneous loss of the protein's function as well. Therefore, a comprehensive understanding of protein folding may explain protein malfunctions in vivo, as well as suggest a means to design artificial proteins with novel functions. *See* AMINO ACIDS; BIOORGANIC CHEMISTRY; PEPTIDE; PROTEIN.

Process. Understanding the folding process and the forces that maintain the native structure is one of the oldest and most challenging problems in biochemistry. The native state of a protein is generally 5–10 kcal/mol more stable than the unfolded state, which is actually an ensemble of loose structures. Levinthal's paradox theorizes that a random search of all possible configurations to find the one with lowest energy (most stable) would take a minimum of 10^{77} years. In reality, local formations can occur in 10–100 nanoseconds, loop formations can fold in 500 ns, and the minimum time for the folding of a small protein is 1 millisecond. Therefore, protein folding must be a directed process. Folding intermediates have been proposed to limit the conformational search by defining one or more pathways to the native state. In addition, intermediates are essential for certain processes such as translocation across or insertion into membranes, release of nonpolar ligands, and association with other cellular components.

Discrete intermediates in the folding process have been isolated for several proteins. Due to the extremely fast time scales of protein folding, partially folded states are often studied at equilibrium where they are stably populated at high or low temperatures, in the presence of chemical denaturants, at pH extremes, or at high pressures. Often these partially constrained intermediates have well-formed local structures but fluctuating or absent long-range

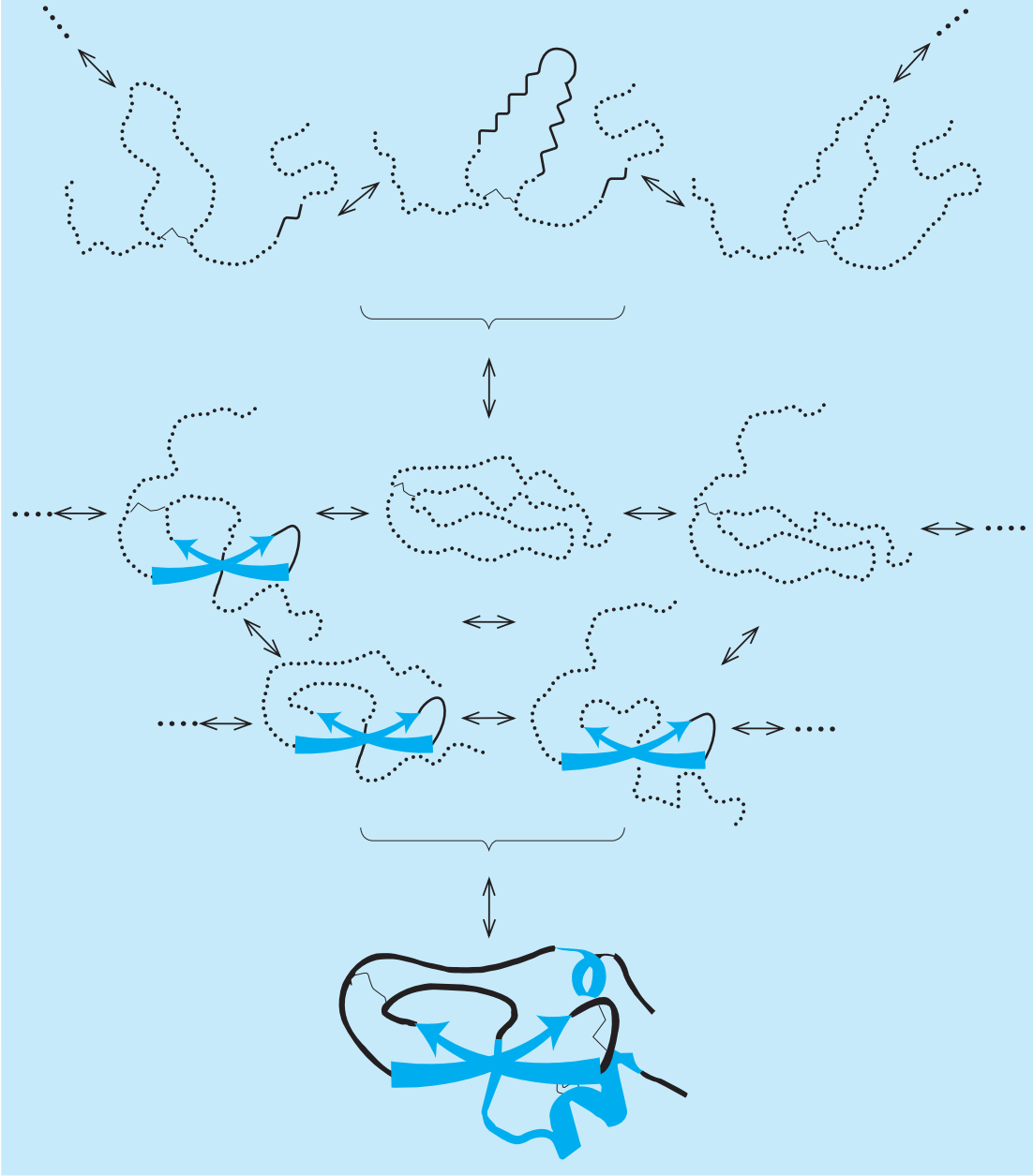

Fig. 2. Schematic of protein bovine pancreatic trypsin inhibitor during the folding process. Arrowhead ribbons are β-strands, and coiled ribbons are α-helices. (*Structures provided by Elisar Barbar and Clare Woodward*)

tertiary interactions (**Fig. 2**). Moreover, local structures may not be evenly dispersed throughout the folding intermediate; some regions may remain unfolded. Equilibrium intermediates are frequently detected by spectroscopic methods that monitor formation of local structural elements, or by the differential fluorescence of amino acids buried within the protein. Structural data (determined by nuclear magnetic resonance or x-ray crystallography) have even been obtained for some folding intermediates. Equilibrium intermediates are thought to be relevant to the kinetic folding process for several reasons. Conditions that favor the formation of the intermediates also speed up overall folding rates. Also, kinetic intermediates (partly folded forms that occur transiently during protein folding) often have structures similar to equilibrium intermediates for those few proteins for which kinetic structural data are available. *See* NUCLEAR MAGNETIC RESONANCE (NMR); X-RAY CRYSTALLOGRAPHY.

Models. Several theoretical models have been proposed to describe the sequence of intermediates assumed by a protein in the course of its folding reaction. In the hydrophobic collapse model, hydrophobic residues are excluded from the aqueous solvent, driving an initial collapse of the polypeptide chain. The native structure is then achieved by reorganizing and forming noncovalent interactions within the collapsed state. In the nucleation-condensation model, the most stable local structures form and propagate, followed by fusion of these various folded regions to form the native protein. Finally, in the

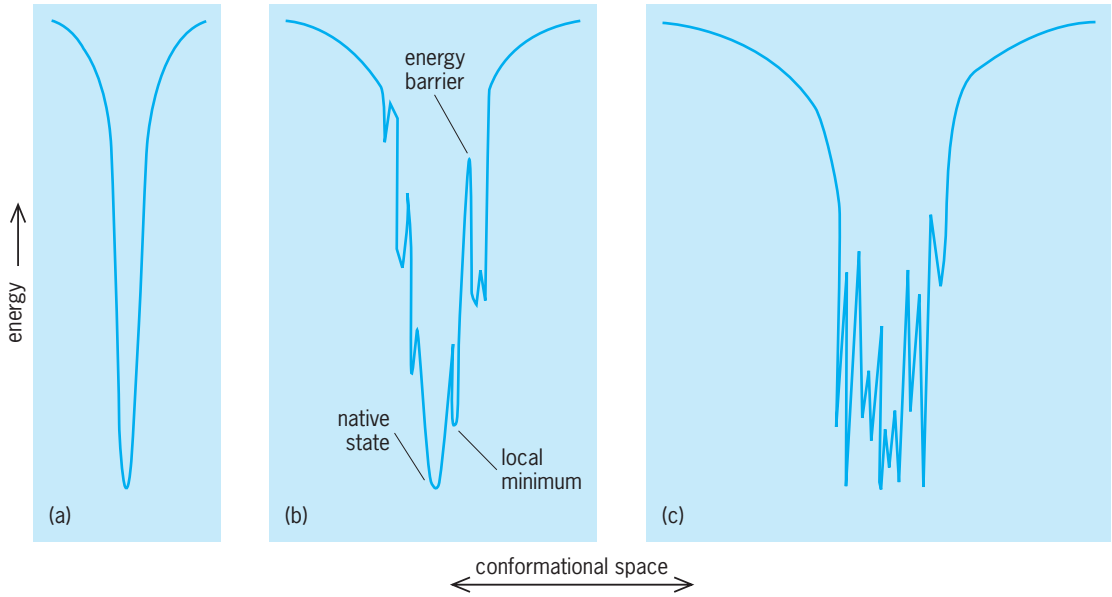

Fig. 3. Two-dimensional cross section through representative folding funnels. (*a*) A protein that folds without intermediates. (*b*) A protein that folds through discrete intermediates, located at the local minima. (*c*) A random heteropolymer with no unique structure.

diffusion-collision model, elements of local structure form and encounter each other by the process of diffusion. Experimental evidence supports each of these models for different proteins, demonstrating that the folding pathway for discrete proteins can be substantially different. Thus, the model that best describes the folding process for any single protein may also vary significantly.

A generalized description of protein folding has been developed that statistically characterizes the energy landscape of folding as a funnel in which the wide top represents the ensemble of high-energy unfolded configurations. The deepest valley represents the native state, the most stable—and therefore lowest energy—conformation (**Fig. 3*b***). Intermediates along the pathway are described as local minima along the sides of the funnel and are separated from the native state by energy barrier peaks. A funnel with smooth sides describes a single, cooperative phase transition (the transition proceeds from starting structure to final structure) with no intermediates (Fig. 3*a*). The opposite extreme is a funnel with very rough sides and a large number of low-energy structures separated by tall barriers. Such a funnel characterizes the folding of a heteropolymer (a compound comprising two or more molecules that are different from one another) that does not adopt a discrete native state (Fig. 3*c*).

In vivo, a protein can be trapped in a structure corresponding to a local minimum in its energy funnel. Therefore, cells have evolved chaperone systems that lower the energy barriers and thereby speed folding and prevent aggregation. A commonly studied bacterial chaperone system is the GroEL/GroES protein complex. This complex sequesters unfolded protein in its hydrophobic center, which is cylindrical in shape, allowing it to fold without exposure to other unfolded proteins which would promote aggregation.

Much progress has been made toward understanding folding phenomena in individual proteins. The computer revolution has enabled modeling of the process, but simulating all atoms in the folding of even a small protein still requires a great deal of time, despite using the largest supercomputers. Modeling is limited to a large degree by incomplete descriptions of the force fields describing the noncovalent interactions that drive folding. With current tools and understanding, structure prediction is possible for peptides and very small proteins, and simple proteins have been designed and created. The current challenge is to extend this knowledge to the large, multidomain and multisubunit proteins that perform the processes crucial to living organisms. The ability to predict folding patterns is critical for interpreting data from the Human Genome Project, which can be used to generate the amino acid sequences, but not the structural and consequently functional information, of all proteins in the human body. *See* HUMAN GENOME.

Exceptions to the folding paradigm. Because protein folding is a balance between a large number of weak intraprotein and protein-solvent interactions, this process is strongly dependent on environmental conditions such as the solvent involved, salinity, temperature, and pressure. In vitro, partially folded or misfolded proteins form under nonoptimal conditions. These incorrect folding states can irreversibly coalesce to form various types of aggregates: amorphous aggregates, films, or amyloid fibrils. Such aggregation is frequently an unwelcome side product during research and for production of protein-based pharmaceuticals.

In vivo, misfolded proteins in a healthy cell are either refolded by chaperones or degraded,

preventing a buildup of aggregates. However, nonoptimal conditions such as inflammation, overproduction of protein, aging/oxidation, and mechanical or chemical stresses may increase the amount of misfolded protein in vivo. These misfolded proteins may then form amyloid fibrils, which are associated with a large number of diseases, among them Alzheimer's disease, Down syndrome, Parkinson's disease, and cataracts. A comprehensive understanding of protein folding and structure will contribute toward the logical design of therapies for these diseases.

Despite the stringent limits imposed upon protein structure by function and the potential effects of aggregation in cells, amazingly regions within many proteins are unfolded or disordered in their native state. Many of these proteins fold upon binding other cellular components—for example, other proteins, nucleic acids, or small molecules—allowing regulation of protein function by a structural oscillation between active and inactive forms. Indeed, many intrinsically disordered proteins participate in inherently transient and tightly controlled processes such as transcription regulation and cell signaling.

Continuing challenges and potential benefits. Protein folding is the least well characterized step in the process by which genetic information is stored, transmitted, and implemented (Fig. 1). Several significant challenges remain, including generating better algorithms or sufficient computational power for modeling large proteins in silico (via computer), monitoring kinetic folding events on very fast time scales, describing the unfolded state, identifying all possible folds (ternary structures), understanding the forces that drive proteins to misfolded states or aggregates and how these forms affect living cells, discovering general principles of protein folding despite varying folding pathways and diverse final structures, and elucidating how and why motion and disorder, like structure, can be encoded by amino acid sequence without eliciting protein destruction. Addressing these issues not only will illuminate a basic and critical aspect of biology but also will expand our ability to dissect uncharacterized biological pathways by predicting the structure and function of novel genes identified by genome sequencing projects. Finally, a thorough description of protein folding would allow exploitation of their functional specificity and efficiency by de novo design of artificial proteins as research reagents, industrial catalysts, and pharmaceutical products.

Sarah Bondos; Kathleen Matthews

Bibliography. Y. Bai, Hidden intermediates and Levinthal paradox in the folding of small proteins, *Biochem. Biophys. Res. Commun.*, 305:785–788, 2003; T. E. Creighton, *Proteins: Structures and Molecular Properties*, 2d ed., W. H. Freeman, 1993; H. J. Dyson and P. E. Wright, Intrinsically unstructured proteins and their functions, *Nat. Rev. Mol. Cell Biol.*, 6:197–208, 2005; K. H. Mayo, Recent advances in the design and construction of synthetic peptides: For the love of basics or just for the technology of it, *Trends Biotechnol.*, 18:212–217, 2000; E. R. McCarney, J. E. Kohn, and K. W. Plaxco, Is there or isn't there? The case for (and against) residual structure in chemically denatured proteins, *Crit. Rev. Biochem. Mol. Biol.*, 40:181–189, 2005; P. G. Wolynes, Energy landscapes and solved protein-folding problems, *Philos. Trans. A Math. Phys. Eng. Sci.*, 363:453–464, 2005.

Protein kinase

One of a family of enzymes that exert regulatory effects on a variety of cellular functions, as well as malignant transformation, by adding a phosphate group to proteins according to the equation

$$\text{ATP} + \text{protein—OH} \xrightarrow{\text{protein kinase}} \text{protein—OPO}_3^{2-} + \text{ADP}$$

(ATP represents adenosine triphosphate, ADP represents adenosine diphosphate, and —OH is a hydroxyl group attached to an amino acid residue.) Based upon the nature of the phosphorylated —OH group, these enzymes are classified as serine/threonine protein kinases and tyrosine protein kinases, where serine, threonine, and tyrosine are amino acid residues found in proteins. Furthermore, there is a small group of dual-specificity kinases, which closely resemble serine/threonine kinases, that catalyze the phosphorylation of both threonine and tyrosine on target proteins. The ratio of phosphoserine/phosphothreonine/phosphotyrosine in proteins from animal cells is about 3000/300/1. Despite the scarcity of tyrosine protein phosphate, it plays a paramount role in cell physiology.

Actions. Protein kinases are enzymes that play a role in nearly every aspect of cell biology. Work in the 1960s showed that protein kinases play a regulatory role in carbohydrate metabolism. Subsequent work indicates that protein kinases participate in nearly all cellular activities including regulation of gene expression, cell division, apoptosis (programmed cell death), differentiation, development, and antibody production. The brain, moreover, is an especially rich site of protein kinase activity. *See* CELL (BIOLOGY); NEUROBIOLOGY.

Protein phosphorylation represents the chief regulatory mechanism in animal cells. In some cases, phosphorylation leads to activation, and in other cases phosphorylation leads to inactivation of the phosphorylated protein. It is estimated that perhaps a quarter of all animal proteins can be phosphorylated by protein kinases. Moreover, many proteins can be phosphorylated at more than one residue by a given protein kinase, and many proteins can be phosphorylated by several protein kinases. *See* ENZYME; PROTEIN.

Protein kinase genes. There are more genes encoding protein kinases than there are genes encoding any other family of human enzymes except for proteases. Workers have identified 518 protein kinase genes in humans (478 typical and 40 atypical), about 2% of the total number of human genes (see **table**). All protein kinases have a similar overall three-dimensional structure and chemical mechanism. *See* GENE; GENE ACTION.

Number of protein kinase genes found in humans

Class	Number		Examples[a]
Protein-serine/threonine	428		
Receptor[b]		12	Transforming growth factor-β
Nonreceptor[c]		369	ERK, cyclin-dependent protein kinases, protein kinase A, protein kinase C, raf protein kinase
Dual specificity		47	MEK
Protein-tyrosine	90		
Receptor[b]		58	Insulin receptor, HER
Nonreceptor[c]		32	Src, Abl
TOTAL	518		

[a]ERK, extracellular regulated kinase; HER, human epidermal growth factor, MEK, mitogen-activated, ERK-activating kinase.
[b]Possess an extracellular domain that binds to a first messenger.
[c]Lacks an extracellular domain and is regulated by second messengers and by phosphorylation-dephosphorylation.

Serine/threonine protein kinases. There are many serine/threonine protein kinases, and three general classes are described here.

Protein kinase A. Protein kinase A, or cyclic adenosine monophosphate (AMP)-dependent protein kinase, was one of the first protein kinases to be described. Many hormones (first messengers) interact with their receptors on the cell membrane, and this interaction leads to the activation of adenylyl cyclase. This enzyme catalyzes the intracellular production of cyclic AMP (a second messenger) from ATP. Cyclic AMP interacts with protein kinase A and activates it. Protein kinase A can lead to the phosphorylation of many different proteins, with many different effects. For example, protein kinase A–mediated phosphorylation of the enzyme that promotes triglyceride degradation leads to its activation, whereas protein kinase A–mediated phosphorylation of the enzyme that promotes glycogen synthesis is inhibitory.

Protein kinase C. Protein kinase C is another serine/threonine kinase that can lead to the phosphorylation of many different proteins. (The C refers to its requirement for ionic calcium.) It also requires phospholipids and diglyceride for full expression of activity. As a result of the interaction of a hormone or growth factor (first messengers) with a receptor on the exterior of a cell membrane, intracellular second messenger molecules such as diglyceride and calcium are liberated, and these activate protein kinase C.

Many protein kinases function in a series, or a cascade, where one enzyme catalyzes the phosphorylation of a second enzyme that in turn mediates the phosphorylation of a third enzyme. A generic form of this cascade is given in the **illustration**. A mitogen-activated protein kinase (MAPK) brings about a physiological response. This enzyme is phosphorylated by a kinase acting on a kinase (MAPKK), and this is then activated by a kinase (MAP kinase kinase kinase, or MAPKKK). A stimulus initiates the process by activating the mitogen-activated protein kinase kinase kinase (MAPKKK). There are several examples of cascades that follow this paradigm, one of which (the Raf-MEK-ERK cascade) is shown in illustration *b*. Protein kinase C–mediated phosphorylation of a serine/threonine protein kinase called Raf leads to its activation, and the activated Raf participates in a cascade that leads to growth and cell division (see illustration). Activation of Raf leads to the phosphorylation and activation of two different protein kinases, MEK1 and MEK2. [MEK stands for mitogen-activated extracellular regulated kinase (ERK)–activating kinase.] MEK1 and MEK2 are examples of dual-specificity protein kinases. These lead to the phosphorylation of both a threonine and a tyrosine residue in each of the protein kinases ERK1 and ERK2, and this dual phosphorylation leads to their activation. ERK1 and ERK2 are serine/threonine protein kinases that mediate the phosphorylation of other protein kinases and regulators of gene transcription. As a result of this cascade of events, cell growth and cell division can result. *See* CELL DIVISION; GROWTH FACTOR; HORMONE.

Cyclin-dependent protein kinases. Cyclin-dependent protein kinases are another family of serine/threonine

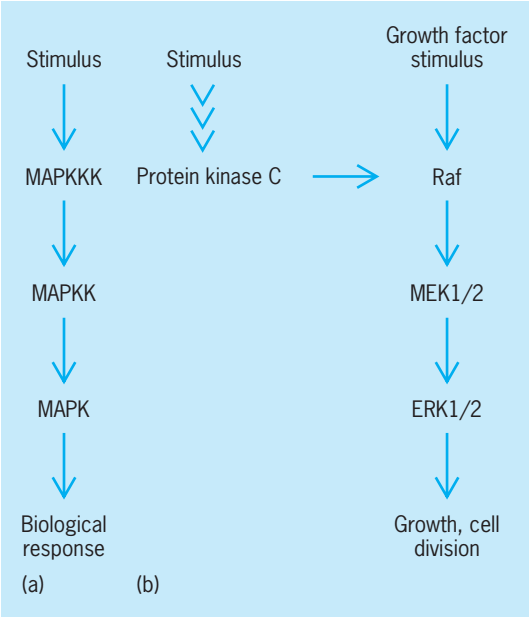

Protein kinase cascades. (*a*) A general protein kinase cascade and (*b*) the Raf-MEK-ERK protein kinase cascade. Protein kinase cascades resulting in an inflammatory response and in apoptosis resemble the general protein kinase cascade. MAPKKK, mitogen-activated protein kinase kinase kinase; MAPKK, mitogen-activated protein kinase kinase; MAPK, mitogen-activated protein kinase.

protein kinases that are essential for cell division. Cyclins are proteins that are synthesized and activated during specific phases of the cell cycle, and they are degraded after they have acted. While the level of cyclin varies during the cell cycle, the level of their cognate protein kinases remains constant throughout the cell cycle. *See* CELL CYCLE.

Tyrosine protein kinases. Receptor tyrosine protein kinases possess an extracellular domain, a transmembrane domain, and an intracellular kinase domain. Binding of a growth factor to the extracellular domain results in receptor dimerization (a chemical reaction in which two identical molecular entities react to form a single molecule, a dimer) and protein kinase activation. The paired kinases phosphorylate each other and other substrate proteins. The phosphorylated tyrosines provide binding, or docking, sites for specific proteins. Binding of a target protein to a docking site results in its activation, and the activated protein initiates a series, or cascade, of downstream phosphorylation reactions. The binding of human epidermal growth factor to members of a family of four receptor tyrosine protein kinases attracts docking proteins that lead to the activation of the Raf-MEK-ERK cascade (see illustration) and to growth and cell division. Variations of this cascade occur in yeast and insects in addition to higher animals. *See* SECOND MESSENGERS.

Protein phosphatases. Protein phosphorylation is a reversible process. Protein kinases add phosphoryl groups to proteins, and protein phosphatases remove these groups during a hydrolysis reaction:

$$\text{Protein—OPO}_3^{2-} + \text{H}_2\text{O} \xrightarrow{\text{protein phosphatase}} \text{protein—OH} + \text{HOPO}_3^{2-}$$

To achieve coordinate regulation of cellular processes, both protein kinase activity and protein phosphatase activity are carefully regulated. Much more is known about the regulation of kinases than phosphatases.

Human cells contain genes corresponding to 120 protein phosphatases: 32 are serine/threonine protein phosphatases, 42 are tyrosine protein phosphatases, and 46 are dual-specificity (threonine/tyrosine) phosphatases. There are three subclasses of tyrosine protein phosphatases. One group contains an extracellular receptor domain, a transmembrane segment, and an intracellular catalytic domain. A second class is tethered to the interior of the cell membrane; this class has the potential to be regulated by transmembrane receptors. A third class of tyrosine protein phosphatases is located within the cell, and this class functions in both the cytosol and the cell nucleus.

Serine/threonine protein phosphatases are generally located in the cell cytoplasm. Phosphoprotein phosphatase-1 is inhibited by a protein called phosphoprotein phosphatase inhibitor-1. Phosphoprotein phosphatases-2A, -2B, and -2C are not inhibited by this protein. Phosphoprotein phosphatase-2B is activated by calcium, a regulatory ion that also activates some protein kinases. There is a balance in the regulation of kinases and phosphatases that leads to physiological responses. Substrates for dual-specificity phosphatases include ERK1, ERK2, and other MAP kinases; these are enzymes that require phosphorylation of both threonine and tyrosine residues for activation.

Drug targets. Owing to the numerous functions of protein kinases and phosphatases, these enzymes are recognized as important drug targets. The HER2 (human epidermal growth factor receptor) protein is overexpressed in about 30% of human breast cancers, and these can be treated effectively with trastuzumab (Herceptin), an antibody directed to the exterior of this receptor tyrosine protein kinase. Imatinib (Gleevec) is an inhibitor of the Bcr-Abl tyrosine protein kinase that occurs in 95% of people with chronic myelogenous leukemia. Both of these medicinal agents are therapeutically effective. A large amount of research on identifying drug targets for kinases and phosphatases is being performed owing to the success of trastuzumab and imatinib. *See* CANCER (MEDICINE).

Robert Roskoski, Jr.

Bibliography. A. Alonso et al., Protein tyrosine phosphatases in the human genome, *Cell*, 117:699–711, 2004; J. M. Berg, J. L. Tymoczko, and L. Stryer, *Biochemistry*, 5th ed., 2001; B. D. Gomperts, I. M. Kramer, and P. E. R. Tatham, *Signal Transduction*, 2002; G. Manning et al., The protein kinase complement of the human genome, *Science*, 298:1912–1934, 2002; D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 4th ed., 2004; R. Roskoski, *Biochemistry*, 1996.

Protein metabolism

The transformation and fate of food proteins from their ingestion and assimilation to the elimination of their excretion products. Proteins are polymers of α-amino acids that are connected by peptide bonds. An average polypeptide chain in a protein contains about 500 amino acid residues. Insulin is a small protein with 51 amino acid residues. Titin, which is the largest known human protein, contains 34,350 residues. A dipeptide contains two amino acid residues, an oligopeptide contains several, and a polypeptide contains many amino acid residues. Proteins are the chief structural and functional components of all living organisms. Their importance was recognized by scientists in the midnineteenth century who coined the name from the Greek *proteios*, meaning first or primary. Proteins are the main building blocks of the cells, tissues, organs, and systems of the body. Proteins of one species differ from those of another species and, within a single animal, proteins of muscle differ from those of the brain, kidney, liver, and other organs. *See* AMINO ACIDS; METABOLISM; PEPTIDE; PROTEIN.

Functions of protein. A 70-kg (154-lb) person contains 10–12 kg (22–26 lb) of protein. After water, proteins are the most abundant component of humans, other animals, and bacteria. Muscle accounts

for about half of the mass of human protein. Collagen is the most abundant protein in animals and is responsible for the strength of the connective tissue found in muscle and tendons. Actin is a protein that occurs in nearly all animal cells and is responsible for cell motility. Myosin is the chief contractile protein of muscle. Hemoglobin is the main protein of red blood cells and is responsible for oxygen transport from the lungs to the tissues. Many hormones, including insulin, are protein in nature. Enzymes that catalyze the myriad reactions in the breakdown of glucose and fatty acids in intermediary metabolism are proteins. Moreover, the adenosine triphosphate synthase that mediates the conversion of adenosine diphosphate and phosphate into ATP, the common currency of energy exchange in all living organisms, is a protein. *See* ADENOSINE TRIPHOSPHATE (ATP); COLLAGEN; ENERGY METABOLISM; ENZYME; HEMOGLOBIN; MUSCLE; MUSCLE PROTEINS.

Protein turnover. Body proteins are continually broken down and resynthesized, a process called turnover. There are two main pathways that participate in intracellular protein degradation: the ubiquitin-proteasome pathway and the lysosomal pathway. The first step in the former pathway involves the joining of ubiquitin molecules with lysine residues in the target protein in a process that requires ATP as an energy source. Ubiquitin is a small protein containing 76 amino acid residues. Following the attachment of several ubiquitin molecules to the target protein, the complex interacts with the proteasome, a large multicomponent complex that degrades the target protein into peptides while liberating ubiquitin, which is then recycled. Proteasomes are located in the soluble portion of the cell, or the cytosol. The proteasomal pathway is responsible for degrading proteins that are damaged because of oxidative stress or are no longer needed owing to changing metabolic needs. The lysosome is a membrane-enclosed vesicle inside the cell that contains a variety of proteolytic enzymes and operates under acidic conditions. It degrades proteins that combine with receptors found on the plasma membrane of the cell and that are subsequently taken up by the cell where they fuse with lysosomes. *See* CELL (BIOLOGY); LYSOSOME.

Normal adults are in nitrogen balance; nitrogen intake equals the amount of excreted nitrogen. Proteins, which are made up of amino acid residues, are the main nitrogen-containing compounds in animals. Growing animals are in positive nitrogen balance; nitrogen intake exceeds nitrogen excretion. During starvation, animals are in negative nitrogen balance; nitrogen excretion exceeds nitrogen intake. People with infections, traumatic injuries (such as from automobile accidents), and major burns also exhibit negative nitrogen balance. Under these conditions, tissue and protein breakdown occurs more rapidly than protein synthesis. Even in the absence of trauma, surgery itself results in accelerated protein breakdown, possibly as a result of increased adrenal steroid (glucocorticoid) production and action. When these conditions are accompanied by reduced food intake, which typically occurs in people with these maladies, this imbalance results in wasting and an impaired immune response. During the healing phase of these illnesses, people exhibit positive nitrogen balance.

Besides the breakdown of endogenous proteins, ingested proteins contribute to the amino acid pool. Protein is digested to amino acids in the stomach and small intestine. These amino acids are absorbed and undergo a degree of metabolic interconversion prior to release into the portal vein, which takes them to the liver. Many of the amino acids are taken up by the liver, and the remainder circulates to the other tissues and organs in the body. The newly absorbed amino acids equilibrate with amino acids in cells that result from protein breakdown.

The amino acid pool serves as the source of precursors for new protein synthesis and for the conversion to other metabolites. Protein synthesis occurs on the ribosome in a messenger ribonucleic acid–dependent process. About 50 grams of muscle protein, 12 g of albumin, 2 g of fibrinogen, 3 g of γ-globulin (antibodies), 8 g of hemoglobin, 30 g of gastrointestinal epithelial protein, and 25 g of enzymes and mucins in the salivary glands, stomach, intestine, and pancreas are synthesized and degraded in the body daily. For normal human adults, about 250 g of protein is degraded and synthesized each day. To the 250 g of amino acids that result from protein degradation, about 100 g per day are obtained from the diet.

Nutritive value of dietary proteins. Plants can synthesize all of the amino acids found in proteins; animals cannot, even though the amino acids are essential for life. Animals ingest protein primarily to supply amino acids for the formation of new and depleted body protein and as a source of various other body constituents derived from amino acids. Twenty amino acids are required for protein synthesis. The amino acids of proteins fall into two nutritional categories: essential or indispensable and nonessential or dispensable. The essential amino acids cannot be synthesized by the body or cannot be synthesized in adequate quantities and therefore must be taken in the diet, while the nonessential amino acids can be synthesized by the body in sufficient quantities to sustain health. The 20 genetically encoded amino acids required for protein synthesis are as follows:

Essential or indispensable	*Nonessential or dispensable*
1. Histidine	1. Alanine
2. Isoleucine	2. Arginine
3. Leucine	3. Asparagine
4. Lysine	4. Aspartate
5. Methionine	5. Cysteine
6. Phenylalanine	6. Glutamate
7. Threonine	7. Glutamine
8. Tryptophan	8. Glycine
9. Valine	9. Proline
	10. Serine
	11. Tyrosine

The relative proportions of the 20 amino acids vary in different foods. Foods of animal origin, including meat, fish, eggs, milk, and dairy products, are sources of high-quality protein. Their essential amino acid content meets the human need for protein synthesis. Gelatin, an animal protein that lacks the essential amino acid tryptophan, is not a high-quality protein. Foods of plant origin can be more difficult to digest and are of lower quality. Plant proteins contain less lysine, methionine, and tryptophan than animal proteins, and these amino acids thus limit the synthesis of body proteins. However, a varied vegetarian diet can readily fulfill human protein requirements if the protein-containing foods are balanced such that their essential amino acids complement each other. For example, legumes such as beans are high in lysine and low in methionine, while rice has a higher percentage of methionine. Thus, if beans and rice are eaten together, their joint amino acid composition will complement each other and provide a higher-quality diet than would either food alone. Traditional food patterns in various cultures have made good use of protein complementarity.

The recommended daily protein intake is 0.8 g/kg body weight for adults, for example, 56 g for a 70-kg (154-lb) person. This amount of protein is easily obtained in a judicious diet: for example, one cup of dry cereal with a cup of low-fat milk for breakfast, a peanut butter sandwich and an 8 oz glass of milk for lunch, and 4 oz of fish, a cup of rice, and cup of split peas amount to 70 g of protein, enough for the average adult. A daily 2000 calorie diet consisting of 6 oz of grains, $2^1/_2$ cups of vegetables, 2 cups of fruit, 3 cups of milk, and $5^1/_2$ oz of meat and beans contains about 90 g of protein. The recommended daily protein intake is greater for infants, for growing children, and for women during pregnancy and during nursing. *See* NUTRITION.

Inadequate protein intake affects health. Kwashiorkor, also called protein malnutrition, is a condition in young children with adequate calorie but inadequate protein intake. Kwashiorkor is most often encountered in developing countries in which the diet is high in starch and low in protein. Signs and symptoms include reddish skin, abdominal swelling, weakness, and nervous irritability. Marasmus results from both insufficient protein and calorie intake in infants under conditions of famine. Marasmus is characterized by growth retardation (in weight more than in height), emaciation, and generalized wasting of muscle and subcutaneous fat. Ribs, joints, and facial bones are prominent owing to the loss of muscle and fat. *See* MALNUTRITION.

Digestion of protein. Protein digestion is initiated in the stomach and completed in the duodenum of the small intestine. The main proteolytic enzyme of the stomach is pepsin, which is secreted by the chief cells in an inactive form called pepsinogen. The hydrogen ion concentration in the stomach is thousands of times greater than that of the blood or in cells; stomach acid is secreted by parietal cells. The conversion of precursor pepsinogen to active pepsin is accelerated by stomach acid and by active pepsin. The activation process involves the hydrolytic removal of an inhibitory 44-amino-acid peptide from pepsinogen. Pepsin preferentially hydrolyzes peptide bonds adjacent to an aromatic amino acid (phenylalanine, tyrosine, tryptophan), and pepsin requires an acid medium in which to function. Stomach acid also unfolds, or denatures, ingested proteins and increases their susceptibility to digestion. *See* DIGESTIVE SYSTEM; PEPSIN.

The acidic stomach contents, which contain partially degraded proteins, are discharged as chyme (semifluid, partially digested food mass) into the slightly alkaline fluid in the duodenum of the small intestine. The pancreatic juice released into the duodenum contains several digestive enzymes that are secreted as zymogens (inactive enzyme precursors) including trypsinogen, chymotrypsinogen, procarboxypeptidase, and proelastase. The activation of trypsinogen involves its proteolytic cleavage by yet another enzyme secreted by the pancreas, namely enteropeptidase, to produce active trypsin. Trypsin, in turn, catalyzes the formation of chymotrypsin, carboxypeptidase, and elastase from their inactive zymogen precursors. In each of these processes, certain peptide bonds are broken to yield the active enzymes. The cells lining the small intestine contain various aminopeptidases that are not liberated into the gut, but act at the cell surface to further proteolytic degradation and to participate in the absorption of the degraded proteins.

Trypsin, chymotrypsin, and elastase are endopeptidases; that is, they cleave peptide bonds that occur in the interior of proteins. Trypsin catalyzes the cleavage of peptide and protein substrates adjacent to positively charged lysine and arginine residues, while chymotrypsin catalyzes peptide bond cleavage next to hydrophobic residues such as methionine, leucine, and the aromatic amino acids. Elastase mediates peptide bond cleavage next to small amino acids such as glycine, alanine, and serine. These three proteases catalyze the cleavage of proteins into small peptides. Aminopeptidases, which are exopeptidases that cleave amino acid residues from the amino-terminal ends of peptides, help to complete protein hydrolysis. Carboxypeptidases, which are also exopeptidases, catalyze the removal of amino acids at the C-terminus of substrate peptides or proteins. Amino acids, dipeptides, and tripeptides are absorbed by the intestinal cells. The hydrolysis of peptides to amino acids is completed intracellularly.

The amino acid digestion products are absorbed as rapidly as they are liberated. The absorption is confined chiefly to the small intestine and is a process that involves the metabolic participation of the cells of the intestinal lining. Amino acids are absorbed from the gut and liberated into the hepatic (liver) portal vein. The outflow of amino acids into the portal vein does not reflect the amino acid composition of dietary protein owing to metabolic transformation by intestinal cells. Most amino acids are taken up and metabolized by the liver except for valine, leucine, and isoleucine—the branched-chain amino acids. These three amino acids may account for 60%

of the amino acids entering the circulation from the gut despite their 20% contribution to the total amino acids of dietary protein. The branched-chain amino acids are taken up chiefly by skeletal muscle. Since they contribute to only 20% of skeletal muscle protein, excess branched-chain keto acid metabolites, derived from these amino acids following transamination (transfer of the amino group of an amino acid to a keto acid to form another amino acid), can be oxidized by muscle or transported to other cells. The liver amino acids are converted into other amino acids (the nonessential amino acids) and into cellular and plasma proteins.

Utilization of absorbed amino acids. Protein homeostasis differs from that of carbohydrate and lipid because protein lacks a storage form. Carbohydrate is stored as glycogen, and lipids are stored as triacylglycerol or triglyceride. The absorbed amino acids become part of the amino acid pool of the body. From these amino acids, new tissue proteins are synthesized to meet body needs. To serve the needs for protein synthesis, all of the 20 genetically encoded amino acids must be present. The absence of any amino acid retards the rate of protein synthesis. *See* CARBOHYDRATE; CARBOHYDRATE METABOLISM; LIPID; LIPID METABOLISM.

Tissue protein synthesis. Turnover of tissue proteins occurs in the adult animal in nitrogen balance, with no net gain of body nitrogen. Body proteins are continually undergoing synthesis and degradation, but the total body protein remains relatively constant in quantity. The rate of replacement varies greatly for different tissues. In humans, it is estimated that the average half-life of total body protein is 80 days; that of lung, brain, bone, and muscle together is about 160 days, while that of liver and serum proteins is only 10 days.

Plasma protein synthesis. The liver is the major organ for plasma protein synthesis. It synthesizes albumin, the predominant blood plasma protein. The liver also synthesizes fibrinogen, one of the main components of blood clots. The plasma proteins offer the most readily available test material in determining the protein nutritional status of an individual. A blood sample is easily drawn, and estimation of the different plasma proteins is now a standard procedure. Plasma proteins exhibit fluctuations in conditions associated with inadequate intake or abnormal protein metabolism. The serum albumin level is the traditional standard for nutritional assessment. Protein deprivation diminishes the albumin content to levels below 30 g/liter, whereas the normal levels are 35–70 g/liter. *See* ALBUMIN; FIBRINOGEN; LIVER.

Blood proteins have numerous physiological functions. Albumin is a major factor in the regulation of blood volume by counteracting fluid expulsion into tissues as a result of hydrostatic blood pressure. Fibrinogen is the chief protein that results in blood clot formation. The globulins include proteins that transport iron, copper, cholesterol, cholesteryl esters, and other lipids. Antibodies, which are synthesized by cells of the immune system, form part of the globulin fraction of plasma proteins. Antibodies defend the body against various pathogenic bacteria and viruses. Protein hormones such as insulin and glucagon, which form a tiny but important fraction of plasma protein, are synthesized in endocrine glands. *See* BLOOD; HORMONE.

Amino acid derivatives. Selected components of the general amino acid pool serve as precursors for numerous nonprotein organic compounds. For example, glycine is converted into heme, an organic component found in hemoglobin. Glycine and aspartate serve as building blocks for the nucleotide bases that occur in deoxyribonucleic acid and ribonucleic acid. Both histamine, which is derived from histidine, and serotonin, which is derived from tryptophan, function as signaling molecules. Tyrosine is converted to norepinephrine, a neurotransmitter, in selected nerve cells and to the melanin pigment in skin cells. *See* DEOXYRIBONUCLEIC ACID (DNA); RIBONUCLEIC ACID (RNA).

Nitrogen excretion. Excess amino acids in cells, tissues, and organs undergo metabolic degradation. The amino group is lost by transamination or deamination (hydrolysis of amino compounds, removing the amino group). About 85% of excess nitrogen is excreted as urea in the urine. Urea $[C(=O)(NH_2)_2]$ is produced in the liver by the Krebs-Henseleit cycle. The overall process is

$$\begin{aligned}&\text{Ammonia} + \text{carbon dioxide} + \text{aspartate} + 3\,\text{ATP} + 2\,H_2O \rightarrow \\ &\quad \text{urea} + \text{fumarate} + 2\,\text{ADP} + 2\,\text{phosphate} \\ &\quad + \text{adenosine monophosphate} + \text{pyrophosphate}\end{aligned}$$

The carbon and oxygen are derived from carbon dioxide, one nitrogen is derived from ammonia, and the second nitrogen is derived from the amino acid aspartate. The process requires the expenditure of three ATP molecules. Ornithine, a nonprotein amino acid, is required for urea biosynthesis; it is continuously regenerated and recycled for additional urea synthesis:

$$\text{Arginine} + H_2O \rightarrow \text{urea} + \text{ornithine}$$

About 2% of nitrogen in the urine is excreted as uric acid, which is derived from DNA, RNA, and nucleotides. A small amount of nitrogen is eliminated in the feces. The remainder is secreted in approximately equal amounts as ammonia, amino acids, and creatinine. Creatinine is a breakdown product of muscle creatine phosphate, an energy-rich storage compound. The residual carbon portions of the amino acids are further metabolized by pathways common to the other major foodstuffs—carbohydrates and fats. Proteins account for about 15% of the calories consumed in an average diet. *See* NITROGEN EXCRETION; UREA.

Robert Roskoski, Jr.; David M. Greenberg

Bibliography. D. H. Baker, Comparative nutrition and metabolism: Explication of open questions with emphasis on protein and amino acids, *Proc. Nat. Acad. Sci. USA*, 102:17897–17902, 2005; J. M. Berg, J. L. Tymoczko, and L. Stryer, *Biochemistry*, 6th ed., W. H. Freeman, New York, 2006; P. C. Champe, R. A. Harvey, and D. R. Ferrier, *Lippincott's*

Illustrated Reviews: Biochemistry, Lippincott Williams & Wilkins, Philadelphia, 2005; P. J. Reeds and P. J. Garlick, Protein and amino acid requirements and the composition of complementary foods, *J. Nutr.*, 133:2953S–2961S, 2003; G. M. Wardlaw, *Contemporary Nutrition*, 6th ed., McGraw-Hill, New York, 2005.

Proteins, evolution of

Proteins are large organic molecules that are involved in all aspects of cell structure and function. Proteins are made up of polypeptide chains, each constructed from a basic set of 20 amino acids, covalently linked in specific sequences. Each amino acid is coded by three successive nucleotide residues in deoxyribonucleic acid (DNA); the sequence of amino acids in a polypeptide chain, which determines the structure and function of the protein molecule, is thus specified by a sequence of nucleotide residues in DNA. *See* GENE; PROTEIN.

Sequence analyses of polypeptides which are shared by diverse taxonomic groups have provided considerable information regarding the genetic events that have accompanied speciation. Interspecies comparison of the amino acid sequences of functionally similar proteins has been used to estimate the amount of genetic similarity between species; species that are genetically more similar to each other are considered to be evolutionarily more closely related than those that are genetically less similar.

The study of functionally related proteins from different animal species has suggested that single amino acid substitutions are the predominant type of change during evolution of such proteins. Insertions or deletions of one or more amino acids have also been reported. The function of a protein is determined by its shape, and the shape is determined by the particular sequence of amino acids; in proteins that serve the same function in dissimilar species small differences in the amino acid sequence will often not affect overall functioning of the protein molecule.

In taxonomic protein sequence analysis, the amino acid sequence of a protein from one species is compared with the amino acid sequence of the protein from another species, and the minimum number of nucleotide replacements (in DNA) required to shift from one amino acid to another is calculated. Peptide "genealogies" can be constructed from many such comparisons in a related group of organisms.

Classical versus protein-derived phylogenies. It has been recognized for a long time that the amino acid sequences of a protein are species-specific. Protein sequencing has been used widely since the mid-1960s to examine taxonomic relationships. Results indicate that, in general, genealogical relationships (phylogenies) based on sequence analyses correspond fairly well with the phylogeny of organisms as deduced from more classical methods involving morphological and paleontological data.

Evolutionary biologists are turning increasingly to the new nucleic acid sequencing technology as an alternative to determining the amino acid sequence of proteins. Knowing the actual nucleotide sequences of genes rather than having to infer them from protein sequence data allows more accurate data to be used in determining genealogical relationships of organisms. For example, silent nucleotide substitutions (that is, base changes in DNA codons that do not result in amino acid changes) can be detected. Nucleotide sequence data has also resulted in increased knowledge of introns (sections of DNA that are not translated into protein), which are usually less conserved than the nucleotide sequences of the coding regions and unexpressed genes, or pseudogenes. Pseudogenes closely resemble a known gene, yet differ from it due to the presence of additions or deletions in their structure, which prevent their normal transcription and translation. Pseudogenes are generally considered to be the diverged products of duplicated genes. *See* INTRON.

Types of proteins studied. The reconstruction of phylogenies from analysis of protein sequences is based on the assumption that the genes coding for the proteins are homologous, that is, descendants of a common ancestor. Those sequences whose evolutionary history reflects that of the species in which they are found are referred to as orthologous. The cytochrome *c* molecules (present in all eukaryotes) are an example of an orthologous gene family. Organisms as diverse as humans and yeast have a large proportion of the amino acids in these molecules in common; they derive from a single ancestral gene present in a species ancestral to both these organisms and to numerous others (see below).

Sequences which are descendants of an ancestral gene that has duplicated are referred to as paralogous. Paralogous genes evolve independently within each species. The genes coding for the human α, β, γ, δ, ϵ, and ζ hemoglobin chains are paralogous. Their evolution reflects the changes that have accumulated since these genes duplicated. Analysis of paralogous genes in a species serves to construct gene phylogenies, that is, the evolutionary history of duplicated genes within a given lineage.

Rate of evolution. Sequence data from numerous proteins have shown that different proteins evolve at different rates. Some proteins show fewer amino acid substitutions, or more conservation, than others (see **table**). Proteins such as immunoglobulins, snake venom toxins, and albumins have changed extensively. Their function apparently requires relatively less specificity of structure and therefore has relatively greater tolerance for variance. By contrast, certain proteins, such as various histones, have changed relatively little over long periods of time. Histone H4 shows extreme conservation; it has essentially the same sequence in all eukaryotes examined. Such extensive sequence conservation is generally interpreted to indicate that the functions of H4 are extremely dependent on its entire structure; thus little or no change is tolerated in its structure. The

Rates of protein evolution*

Protein	Unit evolutionary period†
Histones	
H4	400
H3	330
H2A	60
H2B	60
H1	8
Fibrous proteins	
Collagen (α-1)	36
Crystallin (αA)	22
Intracellular enzymes	
Glutamate dehydrogenase	55
Triosephosphate dehydrogenase	20
Triosephosphate isomerase	19
Lactate dehydrogenase H_4	19
Lactate dehydrogenase M_4	13
Carbonic anhydrase B	4
Carbonic anhydrase C	2.1
Electron carriers	
Cytochrome *c*	15
Cytochrome b_5	11
Plastocyanin	7
Ferredoxin	6
Hormones	
Glucagon	43
Corticotropin	24
Insulin	14
Thyrotropin (β)	9
Lipotropin (β)	8
Lutropin (α)	7
Proparathyrin	7
Prolactin	5
Growth hormone	4
Lutropin (β)	3
Insulin C peptide	1.9
Oxygen-binding proteins	
Myoglobin	6
Hemoglobin (α)	3.7
Hemoglobin (β)	3.3
Secreted enzymes	
Trypsinogen	6
Lysozyme	2.5
Ribonuclease	2.3
Immunoglobulins	
γ chains (C)	1.7
λ chains (C)	1.7
κ chains (V)	1.0
κ chains (C)	0.9
λ chains (V)	0.8
γ chains (V)	0.7
Snake venom toxins	
Long neurotoxins	0.9
Cytotoxins	0.9
Short neurotoxins	0.8
Other proteins	
Parvalbumin	5
Albumin	3
Lactalbumin	2.3
Fibrinopeptide A	1.7
Casein (κ)	1.4
Fibrinopeptide B	1.1

*After A. C. Wilson et al., Biochemical evolution, *Annu. Rev. Biochem.*, 46:573–639, 1977.

†Unit evolutionary period is the average time, in millions of years, required for a 1% difference in amino acid sequence to arise between two lineages; it is thus an inverse measure of the rate of protein evolution.

rates at which different proteins evolve are therefore thought to be due to different functional constraints on the structure of the proteins—the more stringent the conditions that determine the function of a protein molecule, the smaller the chance that a random change will be tolerated in its structure.

Each protein generally has a nearly constant evolutionary rate (the rate of acceptance of mutations) in each line of descent. Exceptions to this rule have been reported, however, and much effort has been spent on determining whether these anomalies are genuine. Some anomalies have been shown to be due to comparison of nonhomologous proteins, and others due to sequencing errors. Other deviations from constant rate of sequence evolution remain to be explained; once uncovered, these may provide useful information about the mechanisms of evolution at the molecular level.

Mitochondrial cytochrome c. Cytochrome *c* is a protein involved in cellular respiration and is found in the mitochondria of plant and animal cells. Although this protein is located in the mitochondria, it is encoded by a nuclear gene. Cytochrome *c* gene sequences have been studied from a wide range of organisms, including protozoa, fungi, various vertebrate and invertebrate groups, and higher plants.

The cytochrome *c* molecules have generally been considered to be an excellent example of a set of orthologous proteins. However, studies have shown that for various mammalian species, cyctochrome *c* pseudogenes may be present, and that there may be cytochrome *c* molecules that are tissue-specific or that occur only during certain developmental stages.

The phylogenetic relationships between various organisms, based on the minimum numbers of nucleotide substitutions necessary to account for the amino acid differences in the cytochrome *c* molecules of these organisms, correspond fairly well with the phylogeny of the organisms, as determined from fossil evidence and comparative morphology (**Fig. 1**).

Cytochrome *c* molecules have evolved slowly. Organisms spanning the range from humans to *Neurospora* have a large proportion of the amino acids in their cytochrome *c* molecules in common. While the evolutionary conservation of such a molecule makes possible the study of genomic differences between remotely related species, this same property prevents determination of evolutionary change among closely related organisms since they may have cytochrome *c* molecules that are completely or nearly identical. For example, sequences from mammals within the same order differ by only 1 or 2 residues out of approximately 100. The primary structure of cytochrome *c* in humans and chimpanzees (which according to paleontological evidence diverged about 10 million years ago) is identical, and in humans and rhesus monkeys (which diverged about 40 to 50 million years ago) it differs by only one amino acid.

Since different proteins evolve at different rates, however, evolutionary relationships among closely related organisms such as humans and chimpanzees can be determined by examining the amino acid sequences of rapidly evolving proteins, such as fibrinopeptides in mammals. Other methods can also be employed to study genetic differences among closely

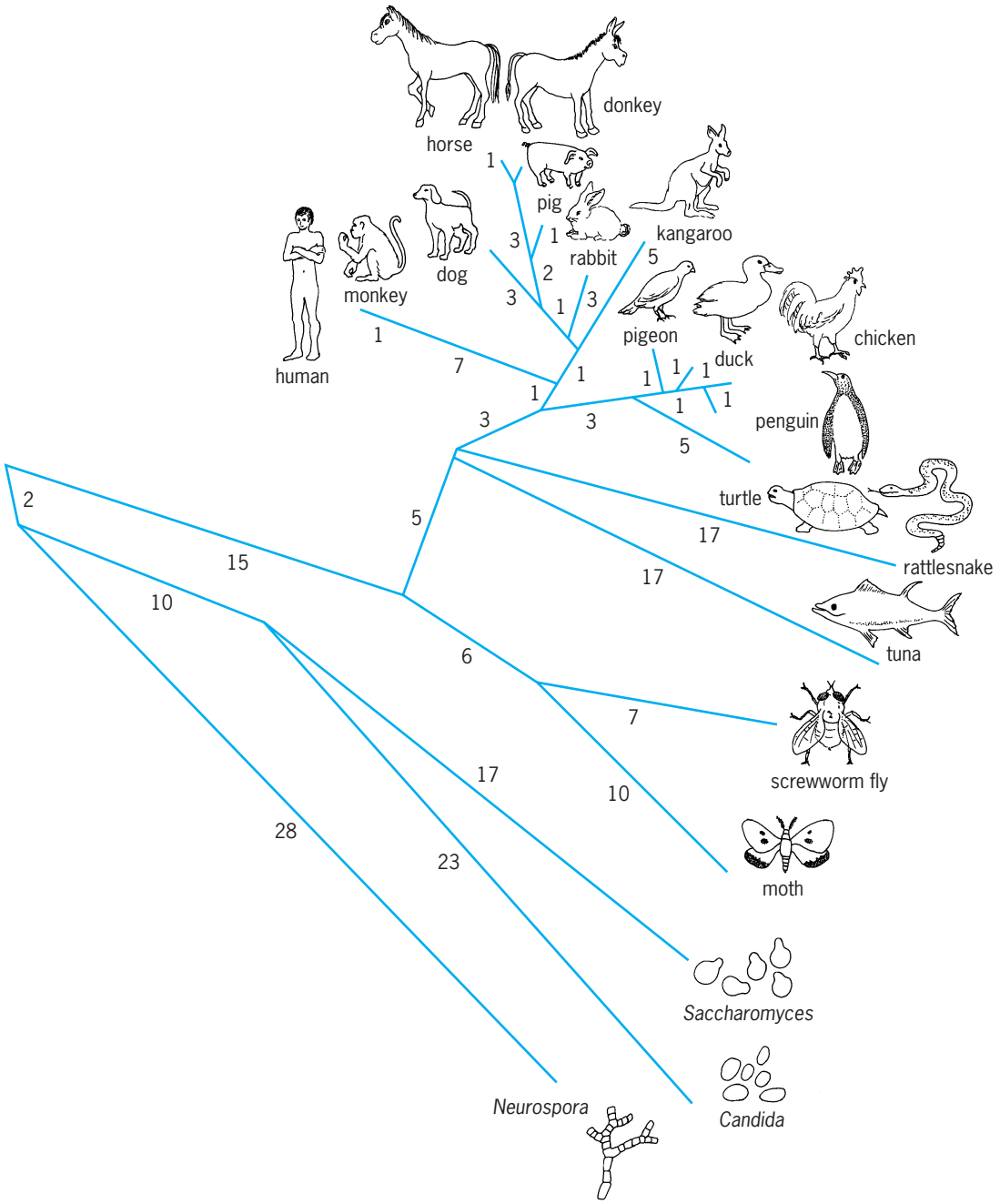

Fig. 1. Phylogeny of 20 organisms, based on differences in the amino acid sequence of cytochrome *c*. The minimum number of nucleotide substitutions required is shown. (*After W. M. Fitch and E. Margoliash, Construction of phylogenetic trees, Science, 155:279–284, 1967*)

related species. These include measurement of immunological cross-reaction between proteins from different species; heteroduplex formation between DNAs from different sources and measurement of their thermostability; examination of protein migration in polyacrylamide gels; use of x-ray crystallography to define the three-dimensional structure of proteins; and direct DNA sequencing of genes.

Globin genes. The genes that code for globin, the protein component of hemoglobin, have been extensively studied. Analysis of the globin gene family has provided detailed information regarding its evolutionary history over the past 500 million years.

Hemoglobin is a tetramer, composed of two α-like and two β-like globin molecules. Humans have five different β-like globins (ϵ, γ^G, γ^A, δ, and β) and three α-like globins (ζ and 2α). These α- and β-globins are synthesized at different times during human development and assemble around a heme group in specific combinations to form functional hemoglobin tetramers.

The human α- and β-globin gene clusters are located on chromosomes 16 and 11, respectively. In both clusters, genes are arranged along the chromosome in the order in which they are expressed during development.

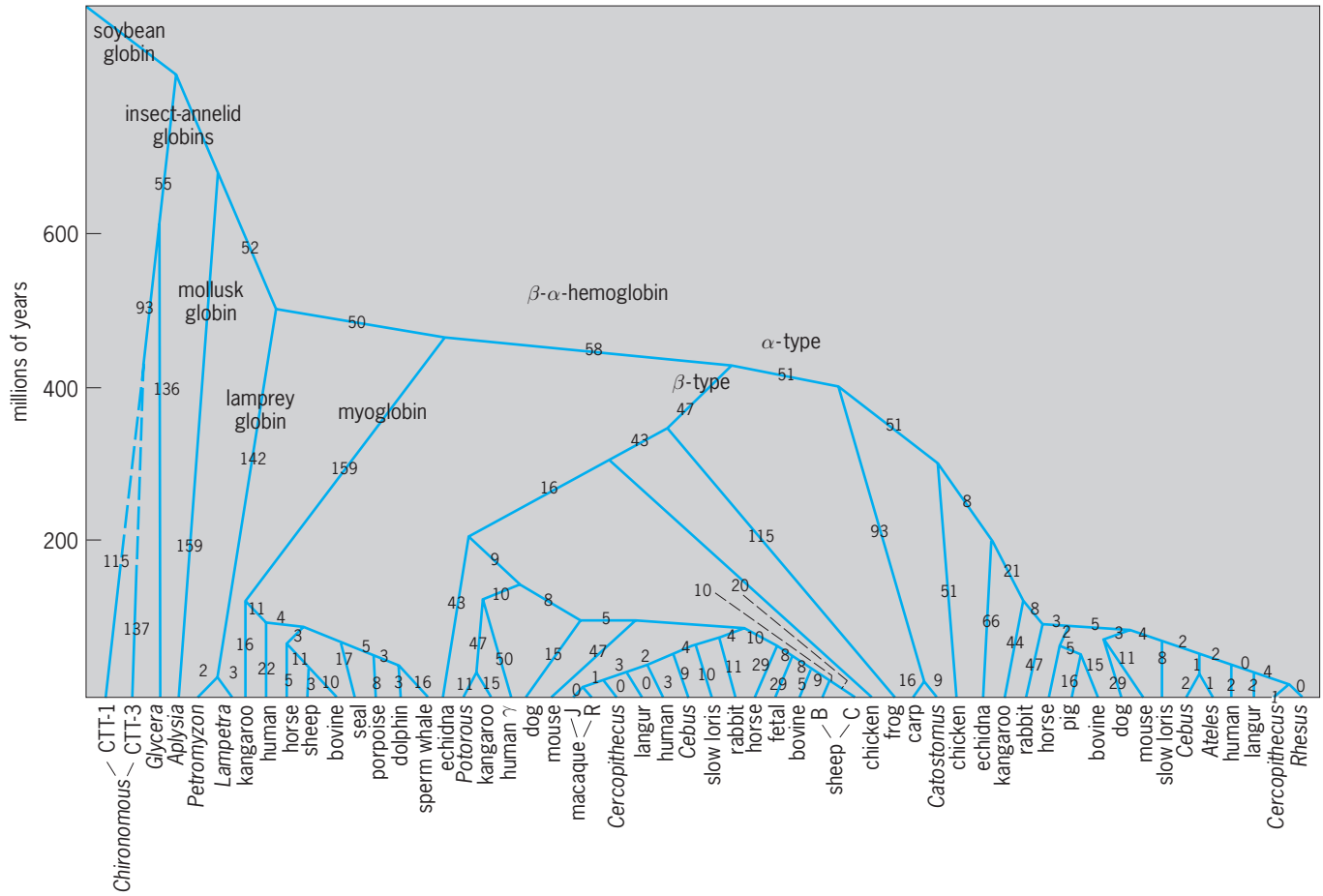

Fig. 2. Globin genealogical tree. Link lengths represent the number of nucleotide replacements between ancestral and descendant sequences. The ordinate scale, in millions of years, is based on paleontological views concerning the ancestral separations of organisms from which the globins came. (*After U. Goodenough, Genetics, 3d ed., Saunders, 1984*)

The similarity in nucleotide sequence of α- and β-globin genes indicates that they arose from a single ancestral gene, presumably by gene duplication. Phylogenetic trees constructed on the basis of amino acid sequence data from vertebrate and invertebrate globin genes suggest that this duplication occurred approximately 425 to 500 million years ago (**Fig. 2**). From such phylogenetic trees, in which the separation of organisms based on paleontological views is fitted against genealogical trees based on nucleotide replacement rates, it is possible to estimate the rate of molecular evolution for proteins. During the early vertebrate period, when monomeric hemoglobin evolved into a tetramer, an accelerated rate of nucleotide substitutions is observed. The sequences directly affected during this period appear to be those directly involved with heme contacts and others that facilitated oxygen binding. Similarly, following the α-β gene duplication, the sequences in the hemoglobin molecule which were directly involved with efficient oxygen delivery evolved much more rapidly than other positions.

Rates of rapid nucleotide substitutions are correlated with periods in which natural selection was presumably acting to produce beneficial adaptations in the α and β chains. Once optimal versions of these chains were produced, stabilizing selection presumably acted to conserve the perfected adaptions at the functionally important sites in the hemoglobin molecules.

How did the α- and β-gene clusters become dispersed on different chromosomes? Recent studies have shown that in *Xenopus laevis* the α- and β-globin genes are closely linked and that the linked genes are found on two different chromosomes ($\alpha^1\beta^1$ and $\alpha^2\beta^2$). In a related species, *X. tropicalis*, however, the α- and β-globin genes are found on only one chromosome. It is speculated that this type of duplication resulted from chromosome duplication by tetraploidization. Duplication of whole chromosome sets in an ancient ancestor followed by switching off of α genes in one location and β genes in another may also have taken place in vertebrate evolution. Since the α and β gene clusters are found on different chromosomes in birds and mammals, the dispersal of these genes presumably occurred 300 to 500 million years ago (before the divergence of birds and mammals from their tetrapod ancestor).

Analysis of α- and β-globin gene structure shows that each gene contains two introns, one of which is generally longer than the other. This is consistent with the idea that these genes had a common origin. Indeed, in all genes examined so far from mammals, birds, and amphibians, these introns, although

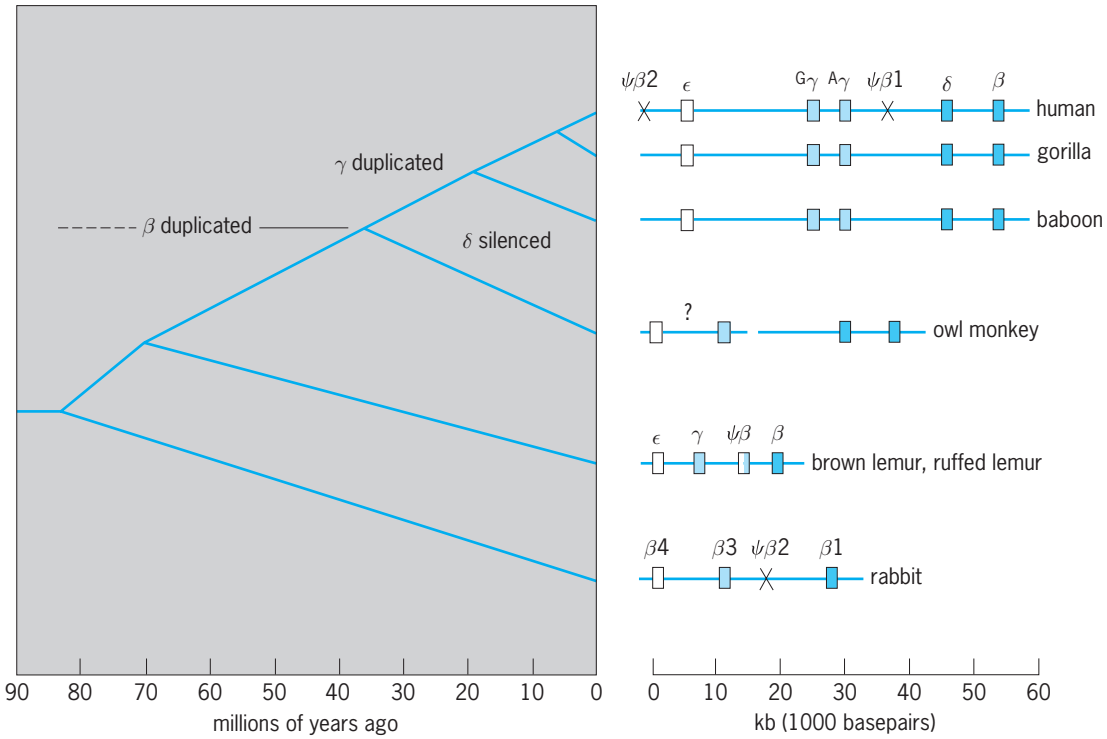

Fig. 3. Phylogeny of the β-globin clusters. The maps of β-globin clusters reveal the sequence in which certain genes changed through evolutionary time. The positions of pseudogenes are shown by crosses. (*After R. Lewin, Evolutionary history written in globin genes, Science, 214:427, 1981*)

varying in length, have remained in the same homologous positions.

The soybean leghemoglobin gene contains three, rather than two, introns. The first and third introns of the leghemoglobin genes are homologous to the first and second introns of the vertebrate genes, and are located at the same homologous positions found in the vertebrate gene. It is possible that the ancestral hemoglobin genes contained three introns, and the central intron was subsequently deleted, before the ancestral gene duplicated to establish the α and β groups. This explanation is supported by the observation that globin's amino acids are assembled in four distinct regions. The first region corresponds with the first exon, and the last with the third exon; the middle two regions are coded by the single middle exon. Leghemoglobin's "extra" intron, however, falls precisely at the division between the vertebrate globin's middle exon and its two heme-binding regions.

In contrast to the coding regions of the genes, the intervening sequences of genes generally accumulate mutations. Nucleotide sequence analysis of globin genes has shown that the coding regions have remained much more conservative than the noncoding regions, which include the 5′ leader and 3′ terminal regions, and introns. Some nucleotide sequences in the noncoding regions, however, such as the sequences involved in the RNA splicing mechanism, and those involved in poly-A initiation, have been preserved. The sequences in the α- and β-globin gene introns have nonetheless diverged considerably, although the sizes and positions of these sequences have remained relatively stable. In contrast to the α- and β-globin genes, most introns examined in other genes do not display such constraint. It is inferred from these properties that the hemoglobin introns may be involved in some unknown function.

The globin gene cluster is thought to have arisen as a gene duplication, followed by changes in gene sequences to produce a family of related although distinct genes that function at different points in development. The human β-globin gene cluster, which contains five closely related functional genes, plus two pseudogenes, has been examined extensively.

A phylogeny of the β-globin gene clusters shows that this gene cluster has remained remarkably stable with respect to the size and organization of the gene cluster for 20 to 40 million years of evolution (**Fig. 3**). Since 95% of the β-globin cluster contains DNA that does not code for proteins, it is inferred that this DNA might perform some as yet undefined function.

The complexity of the β-gene cluster is further illustrated by a number of phenomena. For example, although the γ-gene duplicated 20 to 40 million years ago, the high degree of similarity in the sequences of the duplicated genes suggests a very recent separation. Such similarity can be explained, however, by concerted evolution, whereby repeated genes exchange sequences by certain mechanisms, and maintain similar structures over long periods of evolutionary time.

Another observation in the β-globin cluster concerns the δ-globin gene, which is expressed in humans, apes, and New World monkeys, but not in

Old World monkeys. This gene, however, has been shown to be present in the β gene cluster of Old World monkeys as a pseudogene. It has been speculated that the few defects in this gene might be restored (for example, by back mutation) and the gene expressed in some future Old World monkey lineage. Pseudogenes can therefore be seen not only as evolutionary relics but also as potential new genes.

Pamela K. Mulligan

Bibliography. A. Efstratiadis et al., The structure and evolution of the human beta-globin gene family, *Cell*, 21:653–668, 1980; M. Goodman, *Macromolecular Sequences in Systematic and Evolutionary Biology*, 1982; D. Jones, M. Goodfellow, and F. G. Priest (eds.), *Twenty-Five Years of Numerical Taxonomy*, 1983; R. Lewin, Evolutionary history written in globin genes, *Science*, 214:426–429, 1981; A. C. Wilson, S. S. Carlson, and T. J. White, Biochemical evolution, *Annu. Rev. Biochem.*, 46:573–639, 1977.

Proteocephaloidea

An order of tapeworms of the subclass Cestoda. With one exception, these worms are intestinal parasites of fresh-water fishes, amphibians, and reptiles. The holdfast organ bears four suckers and, frequently, an apical organ which may be suckerlike. The segmental anatomy is very similar to that of the Tetraphyllidea. Most authorities recognize two families, the Proteocephalidae, in which the reproductive organs are within the central mesenchyme of the segment, and the Monticellidae, in which some or all of the organs are in the cortical mesenchyme and which parasitize catfishes. The life histories of several proteocephalids have been studied. After leaving the host's intestine, the embryos must be ingested by an arthropod, typically a copepod of the genus *Cyclops*, in the body cavity of which the young worms grow, developing into procercoids and finally plerocercoids. The vertebrate host becomes infected by eating *Cyclops* containing plerocercoid larvae.

Proteocephalid-like forms in the ancient freshwater fishes probably were ancestral to the other orders of Cestoda. *See* CESTODA; TETRAPHYLLIDEA.

Clark P. Read

Proterozoic

A major division of geologic time spanning from 2500 to 543 million years before present (Ma). The beginning of Proterozoic time is an arbitrary boundary that roughly coincides with the transition from a tectonic style dominated by extensive recycling of the Earth's continental crust to a style characterized by preservation of the crust as stable continental platforms. The end of the Proterozoic coincides with the Precambrian-Cambrian boundary, which is formally defined on the basis of the first appearance of diverse coelomate invertebrate animals. Proterozoic Earth history testifies to several remarkable biogeochemical events, including the formation and dispersal of the first supercontinent, the maturation of life and evolution of animals, the rise of atmospheric oxygen, and the decline of oceanic carbonate saturation (see **table**). Tremendous iron and lead-zinc mineral deposits occur in Proterozoic rocks, as do the first preserved accumulations of oil and gas. *See* CAMBRIAN; PRECAMBRIAN.

Tectonics. Many of the Earth's Archean cratons are blanketed by little-deformed sequences of Proterozoic sedimentary rocks, which indicate that vigorous recycling of the Earth's crust, characteristic of Archean time, had slowed markedly by the beginning of Proterozoic time. This decrease in crustal recycling is attributed to the development of thick continental roots, which stabilized the cratons, and the decrease in heat that was escaping from the Earth's interior, believed to drive thermal convection in the Earth's mantle and recycling of the crust. The sedimentary record that is present atop the major cratons is sufficiently continuous and well enough preserved that a rather detailed accounting of major events in Proterozoic Earth history is possible.

Most of the Earth's Archean cratons appear to have participated in the formation of a supercontinent in Mesoproterozoic time, about 1200 Ma. This supercontinent, called Rodinia, was the first for which there is now substantial evidence. Rodinia seems to have assembled with the North American craton (Laurentia) at its center. Rodinia persisted until the latest part of the Neoproterozoic, about 600 Ma, when it fragmented and ejected Laurentia in the process. Almost immediately, the other cratons reassembled to form the supercontinent of Gondwanaland, whose assembly was complete by about 450 Ma. *See* ARCHEAN; CONTINENTS, EVOLUTION OF; EARTH, CONVECTION IN; EARTH, HEAT FLOW IN; EARTH CRUST; EARTH INTERIOR; PLATE TECTONICS.

Environment. Giant iron oxide deposits were formed by precipitation from seawater about 2000 Ma. These iron formations coincide with the first appearance of red, oxidized sedimentary rocks formed in terrestrial river environments and with the retention of ferric iron in soil horizons. During most of Paleoproterozoic time the oceans and atmosphere were reducing and ferrous iron was abundant in seawater. When photosynthetic microbes produced oxygen in shallow waters, the ferrous iron became oxidized to ferric iron, which is highly insoluble; and the ferric iron precipitated and settled to the sea floor, making the giant iron formations. The iron formations were produced until the iron sink was depleted about 2000 Ma, whereupon oxygen was free to accumulate in the atmosphere and shallow ocean. *See* ATMOSPHERE, EVOLUTION OF.

The partial pressure of carbon dioxide on the early Earth was very high. During Proterozoic time, much of the mass of carbon shifted from the ocean and atmosphere to the solid Earth. Enormous volumes of limestone [$CaCO_3$] and dolostone [$CaMg(CO_3)_2$] were deposited and testify to this shift. These carbonate deposits show significant textural variations, indicating that in the Paleoproterozoic supersaturation with respect to calcium carbonate was very

Proterozoic subdivisions and the major events

Eon	Era	Historical benchmarks
Proterozoic	Neoproterozoic	Evolution of animals Further rise of atmospheric oxygen Dispersal of Rodinia
	1 billion years	
	Mesoproterozoic	Assembly of Rodinia supercontinent Evolution of multicellular eukaryotes
	1.6 billion years	
	Paleoproterozoic	Deposition of iron formations and rise of atmospheric oxygen Precipitation of limestones Evolution of eukaryotes

high but declined markedly during the Mesoproterozoic. Neoproterozoic levels of calcium carbonate saturation were similar to those documented for Phanerozoic time. *See* DOLOMITE ROCK; LIMESTONE; SEDIMENTARY ROCKS.

Glaciers covered significant parts of the Earth during two widely separated times in Proterozoic history. The first episode occurred about 2200 Ma, and glacial deposits of that age cover various parts of North America and Scandinavia. The second episode consisted of at least two different pulses spanning from 750 to 600 Ma during Neoproterozoic time. Glaciers formed at that time were of almost global extent, and this was possibly the most frigid time in Earth history. Glaciers are thought to have extended from the poles to the Equator, according to the snowball Earth hypothesis.

Carbon and strontium isotopes present in limestones and dolostones suggest that several other environmentally significant events occurred toward the end of Neoproterozoic time. The presence of strontium isotopes suggests that high mountain ranges developed, probably in association with the earliest stages of Gondwanaland assembly. Rapid erosion of those mountains would have caused an increase in sedimentation rates in the deep ocean. Carbon isotopes suggest that much organic matter, normally a major oxygen sink, was buried because of the high sedimentation rates. As a consequence, oxygen was free to build up to unprecedented levels, perhaps triggering the evolution of animals.

Life. A number of significant events in the evolution of life occurred during Proterozoic time. The record of biological activity is rich, consisting of actual body fossils, in addition to organism traces and impressions, and complex chemical biomarkers. The geologic record of evolution has been independently tested by comparisons with the phylogenetic trees of living organisms constructed from information sequences contained within macromolecules such as ribonucleic acid (RNA) and various proteins. *See* RIBONUCLEIC ACID (RNA).

Eukaryotic microbes appear to have evolved by about 1900 Ma, when they became major players in ecosystems present at that time. The fossil record is independently confirmed by the presence of steranes, known only from the membranes of eukaryotes, in 1700-Ma-old oil deposits. By the beginning of Neoproterozoic time, about 1000 Ma, multicellular eukaryotic algae are present in numerous sedimentary basins around the world. These eukaryotes, in addition to their reproductive cysts, evolved to fill distinct niches in ecologically complex ancient seas, as shown by onshore-offshore patterns of distribution. The eukaryotic fossil record shows evidence of a major adaptative radiation and subsequent extinction near the end of Neoproterozoic time. *See* EUKARYOTAE.

The evolution of animals did not take place until the close of Neoproterozoic time. The first direct evidence consists of relatively large impressions of soft-bodied organisms whose origin is still broadly debated. The consensus view is that the impressions were formed by unskeletonized animals such as cnidarians and bilaterians. The broader question of why these organisms evolved at this particular time in Earth history remains unanswered. General opinion proposes that it was likely the result of the confluence of a number of environmental factors, such as the rise in oxygen related to increased rates of carbon burial, on top of the previous 3 billion years of molecular and evolutionary preconditioning. Whatever the cause of their origin, these existed until at least 543 Ma, when another major evolutionary adaptive radiation began. This was the most impressive in Earth history, in which most of the extant phyla and classes of marine invertebrates evolved over a remarkably short time interval of about 10 million years. The beginning of this explosive diversification (545 Ma) marks the onset of Cambrian time and the end of the Proterozoic Eon. *See* ANIMAL EVOLUTION; EXTINCTION (BIOLOGY); GEOLOGIC TIME SCALE. John P. Grotzinger

Bibliography. S. Bengtson, *Early Life on Earth*, Nobel Symp. 84, 1994; J. P. Grotzinger et al., Biostratigraphic and geochronologic constraints on early animal evolution, *Science*, 270:598–604, 1995; P. F. Hoffman, Did the breakout of Laurentia turn Gondwanaland inside-out?, *Science*, 252:1409–1412, 1991; A. H. Knoll, End of the Proterozoic Eon, *Sci. Amer.*, 265:64–73, 1991; J. W. Schopf and C. Klein, *The Proterozoic Biosphere*, 1992.

Protista

The kingdom comprising all single-celled forms of living organisms in both the five-kingdom and six-kingdom systems of classification. Kingdom Protista encompasses both Protozoa and Protophyta,

allowing considerable integration in the classification of both these animallike and plantlike organisms, all of whose living functions as individuals are carried out within a single cell membrane. Among the kingdoms of cellular organisms, this definition can be used to distinguish the Protista from the Metazoa (sometimes named Animalia) for many-celled animals, or from the Fungi and from the Metaphyta (or Plantae) for many-celled green plants. This system simplifies the lower-level classification (orders and superfamilies) where closely allied forms of flagellate protistans may possess chlorophyll and, therefore, be plantlike in their metabolism, or lack chlorophyll and feed as animals. However, there are other difficulties, not only of semantics but also of fundamental biology, involved in attempting diagnostic separation of the Protista from the other major groups of living organisms. *See* METAZOA.

The most significant biological distinction is that which separates the bacteria and certain other simply organized organisms, including blue-green algae (collectively, often designated Kingdom Monera), from both Protista and all many-celled organisms. Distinctions between certain plantlike Protista and the lower many-celled plants are very difficult to make. A distinction between the animallike Protista (or Protozoa) and many-celled animals is possible by a moderately complicated definition. Separation of the bacteria (and other Monera) is also easier semantically. The bacteria are described as prokaryotic; both the Protista and the cells of higher plants and animals are eukaryotic. Structurally, a distinguishing feature is the presence of a membrane, closely similar to the bounding cell membrane, surrounding the nuclear material in eukaryotic cells, but not in prokaryotic ones. This diagnostic feature is paralleled by other features of functional morphology; for example, eukaryotic cells can be considerably more compartmentalized with membrane-bound structures (organelles) responsible for particular cell functions. There are also fundamental differences between prokaryotic and eukaryotic organisms in the mechanics of cell multiplication, particularly as concerns division of nuclear material. *See* EUKARYOTAE; PROKARYOTAE; PROTOZOA.

It is important to state that individual protistans are both whole organisms and individual cells. Evidence from studies by electron microscopy has settled earlier controversies. Semantically, protistans are unicellular rather than acellular. In the majority of its ultrastructures, the individual protistan cell appears to be strictly comparable to an individual generalized cell of a metazoan. The definition that can separate the Protista from many-celled animals is that the protistan body never has any specialized parts of the cytoplasm under the sole control of a nucleus. In some protozoa, there can be two, a few, or even many nuclei, rather than one, but no single nucleus ever has separate control over any part of the protistan cytoplasm which is specialized for a particular function. In contrast, in metazoans there are always many cases of nuclei, each in control of cells of specialized function.

Most authorities would agree that the higher plants, the Metazoa, and the Parazoa (or sponges) almost certainly evolved (each independently) from certain flagellate stocks of protistans. This evolutionary significance of the Protista remains valid, but work on the ultrastructure of and physiology in protozoa provides much evidence of considerable levels of adaptive complexity. Thus, the early hope that study of the apparent simplicity of protistans would more readily provide understanding of certain life processes than similar studies in higher forms has proven largely illusory. *See* CLASSIFICATION, BIOLOGICAL.

W. D. Russell-Hunter

Protobranchia

A subclass of bivalve mollusks characterized by a foot with a sole that is divided sagittally and longitudinally and has papillate margins. The foot is without a true byssus gland, although a nonhomologous byssal gland is frequently present in the heel. The byssal gland does not produce threads and has a structure totally different from the byssus gland of the subclass Lamellibranchia. The gill filaments, which are not reflected, are usually platelike with abfrontal cilia. Thus the gills (etenidia) resemble those of aspidobranch gastropods rather than those of the lamellibranch bivalves. Digestion is entirely extracellular.

The Protobranchia are divided into two orders, Solemyoida and Nuculoida. The Solemyoida have an edentulous shell, and if teeth are present, they are not chevron-shaped. The palps are small, triangular, and without palp proboscides. This is in contrast to the Nuculoida, which have well-developed, chevron-shaped hinge teeth and large palps with palp proboscides.

The Nuculoida are subdivided into two superfamilies, Nuculacea and Nuculanacea. The Nuculacea are asiphonate, without fusion or particular specialization of the mantle edge. For example, there is no anteroventral mantle sense organ. The Nuculanacea are siphonate, and ventral to the siphons the mantle is modified to form a feeding aperture. An anterior sense organ and a siphonal tentacle are attached to the siphonal embayment.

The Protobranchia are found throughout the seas of the world, and they are particularly common in the deep sea, where they may form up to 10% of the invertebrate infauna. The subclass has one of the longest geological records within the animal kingdom, dating from the Early Ordovician if not the Late Cambrian. During the course of evolution, they have changed little in their shell form. They probably owe their success in the deep sea to their deposit-feeding habit, which—coupled with extracellular digestion, low metabolic rate, and slow growth—has enabled them to exploit the low organic content of deep-sea soft sediments. Some deep-sea species are extremely long-lived: one species (*Tindaria callistiformis*) is known to live for 120 years, yet grows to little more than 0.4 in. (1 cm) total length. *See* BIVALVIA; MOLLUSCA.

John A. Allen

Bibliography. J. A. Allen and F. J. Hannah, A reclassification of the Recent genera of the subclass Protobranchia (Mollusca: Bivalvia), *J. Conchol.*, 32:225–249, 1986; E. R. Trueman (ed.), *The Mollusca*, vol. 10: *Evolution*, 1985.

Protogyny

A condition in hermaphroditic or dioecious animals in which the female reproductive structures mature before the male structures. It is of rare occurrence. Botanically, protogyny occurs in some plant species in which the stigma develops, withers, and dies before the anthers mature. *See* PROTANDRY.

Tracy I. Storer

Protolepidodendrales

An extinct order of the class Lycopsida (clubmosses) of the Devonian Period. Most of these clubmosses are poorly preserved and not well understood; the notable exception is the widespread *Leclercqia complexa*, which has been largely reconstructed from fragmentary fossils.

Leclercqia (Early to Middle Devonian) was a medium-sized herb; its horizontal rhizomes generated vertical axes up to 0.4 in. (1 cm) in diameter (see **illus.**). The vascular tissue formed a star-shaped actinostele, with inward maturation of the metaxylem tracheids. Additional structural support was provided in the cortical cylinder by thickened hypodermal fibers. The microphyllous leaves were not routinely shed. Most branched into five unequal, needlelike lobes, each supplied by a lateral branch of the median vascular trace, and arranged three-dimensionally. Zones of sterile leaves alternated with zones of fertile sporophylls; both sterile and fertile leaves bore a small ligule on the upper surface. The stomata-bearing sporangia were located on the same surface of the sporophyll, somewhat closer to the stem than the ligule. The sporangia were elliptical and released spores of the same size; thus, *Leclercqia* is primitively homosporous.

Other putative protolepidodendraleans together span the Devonian Period. *Protolepidodendron* (Early Devonian) had bifid leaves and sporophylls, whereas the leaves of *Colpodexylon* (Middle to Late Devonian) were trifid; their stems reached 0.8 in. (2 cm) in diameter. Although traditionally included, *Estinnophyton* (Early Devonian) is a doubtful member of the group, as it bore two pairs of spindle-shaped sporangia per sporophyll rather than the single sporangium characteristic of the Lycopsida. *Haskinsia* and *Archaeosigillaria* (both Middle to Late Devonian) possessed relatively large axes that bore spear- or arrow-shaped microphyllous leaves.

Complex leaf morphologies are important characteristics in assigning these genera to the Protolepidodendrales, even though it may be that they represent multiple independent evolutionary events. The presence of a ligule has been demonstrated only in *Leclercqia*; yet it is the combination of homosporous reproduction and the presence of a ligule that delimits the order and testifies to its evolutionary intermediacy between the eligulate homosporous Lycopodiales and the ligulate heterosporous Selaginellales. *See* LYCOPHYTA; LYCOPODIALES; LYCOPSIDA; SELAGINELLALES.

Richard M. Bateman; William A. DiMichele

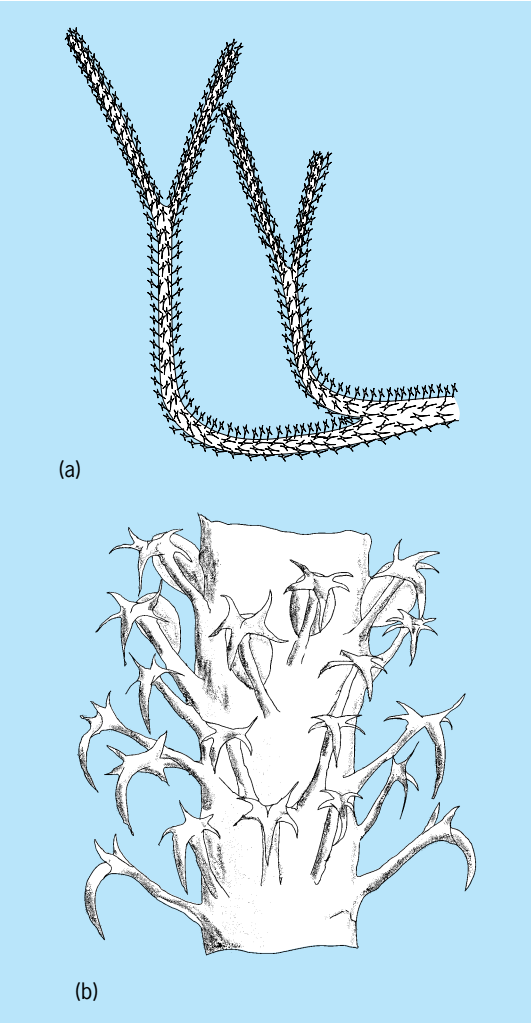

Early to Middle Devonian *Leclercqia complexa*, the best-known protolepidodendralean. (*a*) General growth habit. (*b*) Pentafid sporophylls bearing sporangia in the fertile region of the axis. (*After P. A. Bonamo et al., Leclercqia, Haskinsia, and the role of leaves in delineation of Devonian lycopod genera, Bot. Gaz., 149:222–239, 1988*)

Bibliography. P. G. Gensel, Phylogenetic relationships of the zosterophylls and lycopsids, *Ann. Mo. Bot. Gard.*, 79:450–473, 1992; F. M. Hueber, Thoughts on the early lycopsids and zosterophylls, *Ann. Mo. Bot. Gard.*, 79:474–499, 1992; W. N. Stewart and G. W. Rothwell, *Paleobotany and the Evolution of Plants*, 2d ed., 1993.

Proton

A positively charged particle that is the nucleus of the lightest chemical element, hydrogen. The hydrogen atom consists of a proton as the nucleus, to which a single negatively charged electron is bound by an

attractive electrical force (since opposite charges attract). The proton is about 1836 times heavier than the electron, so that the proton constitutes almost the entire mass of the hydrogen atom. Most of the interior of the atom is empty space, since the sizes of the proton and the electron are very small compared to the size of the atom. *See* ATOMIC STRUCTURE AND SPECTRA; ELECTRIC CHARGE; HYDROGEN.

For chemical elements heavier than hydrogen, the nucleus can be thought of as a tightly bound system of Z protons and N neutrons. An electrically neutral atom will then have Z electrons bound comparatively loosely in orbits outside the nucleus. *See* NEUTRON; NUCLEAR STRUCTURE.

Most of the overall properties of the proton can be experimentally determined in straightforward ways. Free protons can be easily obtained for experimental purposes by ionization (electron removal) of hydrogen atoms in an electrical gas discharge. In some cases, measurements on the proton are best carried out by using neutral hydrogen atoms. In magnetic resonance experiments, the magnetic dipole moment of the proton has been measured with high accuracy. The numerical values of some overall properties of the proton can be summarized as follows: charge, 1.602×10^{-19} coulomb; mass, 1.673×10^{-27} kg; spin, $(^1/_2)\hbar$ (where $\hbar$ is Planck's constant h divided by 2π); magnetic dipole moment, 1.411×10^{-26} joule/tesla; radius, about 10^{-15} m. *See* FUNDAMENTAL CONSTANTS; ION SOURCES; MAGNETIC RESONANCE; NUCLEAR MOMENTS; SPIN (QUANTUM MECHANICS).

Size and structure. In order to understand basic questions concerning the size and structure of the proton, it is instructive to contrast its properties with those of the electron. The behavior of the electron has also been studied in many precision experiments at both high and low energies, and all of the electron's properties have been found to be those expected of a spin-$^1/_2$ particle which is described by the Dirac equation of quantum mechanics. Such a Dirac particle has no internal size or structure. *See* ELECTRON; RELATIVISTIC QUANTUM THEORY.

By contrast, although it also has a spin of $^1/_2$, the proton definitely does not behave as a geometrical point Dirac particle. Its magnetic moment, which is different from that for a Dirac particle, and its binding with neutrons into nuclei strongly suggest that the proton has some kind of internal structure, rather than being a point particle. That this is actually the case has been shown in many high-energy physics experiments. Experiments to explore the structure of the proton need to be carried out by using high-energy particles for two reasons. First, the uncertainty relationship between momentum and position, which is a fundamental principle of quantum mechanics, requires the use of high-momentum particles to observe small regions of space. Second, Einstein's equation of special relativity, $E = mc^2$, where E is energy, m is mass, and c is the speed of light, shows that to produce the additional mass which is needed to form excited states of the proton (described below) requires high-energy collisions. *See* UNCERTAINTY PRINCIPLE.

Correspondingly, two different kinds of high-energy physics experiments have been used to study the internal structure of the proton. An example of the first type of experiment is the scattering of high-energy electrons, above say 1 GeV, from a target of protons. The angular pattern and energy distribution of the scattered electrons give direct information about the size and structure of the proton. The second type of high-energy experiment involves the production and study of excited states of the proton, often called baryonic resonances. It has been found that the spectrum of higher-mass states which are produced in high-energy collisions follows a definite pattern. *See* BARYON.

In 1963, M. Gell-Mann and, independently, G. Zweig pointed out that this pattern is what would be expected if the proton were composed of three spin-$^1/_2$ particles, quarks, with two of the quarks (labeled u) each having a positive electric charge of magnitude equal to $^2/_3$ of the electron's charge (e), and the other quark (labeled d) having a negative charge of magnitude of $^1/_3 e$. Subsequently, the fractionally charged quark concept was developed much further, and has become central to understanding every aspect of the behavior and structure of the proton.

Quark structure. As a result of experimental and theoretical research in high-energy physics since the 1960s, a broad understanding of the internal structure and constituents of the proton has been achieved. The behavior of the proton is essentially that of a system of three quarks, with the charges specified above. The excited, higher-mass states of the proton have been shown to correspond in some detail to the theoretical expectations for such a three-quark system. This correspondence is not yet on a precise basis, because the equations of the quantum-mechanical theory of interactions between quarks, quantum chromodynamics (QCD), have not yet been solved rigorously for a bound system of three quarks.

Quantum chromodynamics has led to other radically new insights into the behavior of the quarks. Among the most fundamental and yet paradoxical properties of the quarks, according to quantum chromodynamics, is the existence of a new quantum number called color, which acts in such a way as to forbid the stable existence of a single, isolated quark. The requirement that isolated physical systems have zero net color in fact forbids the existence of any free particles which have fractional charge. This conclusion has been tested by many experiments which have searched unsuccessfully for such particles. *See* COLOR (QUANTUM MECHANICS).

For the proton, the same principle leads to the conclusion that the proton is the lowest allowed energy state of three quarks. The proton cannot dissociate or decay into separated quarks, for that would lead to the existence of free quarks. In other words, the proton is absolutely stable with respect to the quantum chromodynamics interactions of quarks. *See* QUANTUM CHROMODYNAMICS; QUARKS.

Search for decay. For many years, it has been recognized not only that the stability of the proton is of great practical importance, but also that this stability is probably due to some fundamental conservation law in physics. The proton and the neutron are the building blocks of which all nuclei are formed, and the observed stability of the atoms in the universe is a direct consequence of the stability of the proton itself.

As described above, quantum chromodynamics gives a clear description of why the proton, as a three-quark system, cannot break apart into free quarks, but this leaves open the question as to whether the quarks themselves might decay. Such a question goes beyond the scope of quantum chromodynamics, but it can be answered by the unified electroweak theory of S. Weinberg, A. Salam, and S. Glashow. This theory gives a description of all known electromagnetic and weak (radioactive decay) phenomena, and has so far correctly predicted the results of all experimental measurements of these phenomena. The theory also gives a definite prediction that the proton is stable. *See* WEAK NUCLEAR INTERACTIONS.

However, an important class of fundamental theories, called grand unification theories (GUTs), makes the opposite prediction—that the proton will decay. The predicted lifetime of the proton is very long, about 10^{30} years or more—which is some 10^{20} times longer than the age of the universe—but this predicted rate of proton decay may be detectable in practical experiments. *See* GRAND UNIFICATION THEORIES.

The fundamental physics issue at stake in the search for proton decay is whether quarks, whose behavior is described in quantum chromodynamics, will ultimately be transformed into electrons or muons, which are described by a completely separate (electroweak) theory. If this process does occur, it will be unambiguous evidence for a single mechanism being at the root of all known forces in nature, excepting gravity. It will also permit the unification of all known particle types into one family.

To search for proton decay, experiments are being carried out which can detect and identify one or two proton decay events per year in a sample whose mass is of the order of 1000 tons (900 metric tons). To achieve such unprecedented sensitivity, the entire experiment is carried out in a deep underground laboratory. This shields the apparatus from most of the background events which would otherwise be caused by cosmic rays. The **illustration** shows a proton decay detector being constructed in a laboratory beneath the Alps. The detector is constructed of 900 metric tons (992 short tons) of thin iron plates, each 20 ft × 20 ft × 0.12 in. (6 m × 6 m × 3 mm). Particles resulting from proton decay within an iron nucleus will be detected as they traverse Geiger tubes and flash chambers which are sandwiched between the thin iron plates. Even in an underground laboratory, there is an unavoidable background caused by neutrinos, which are secondary products of cosmic rays. This neutrino background will set an ultimate limit on the sensitivity of proton decay experiments.

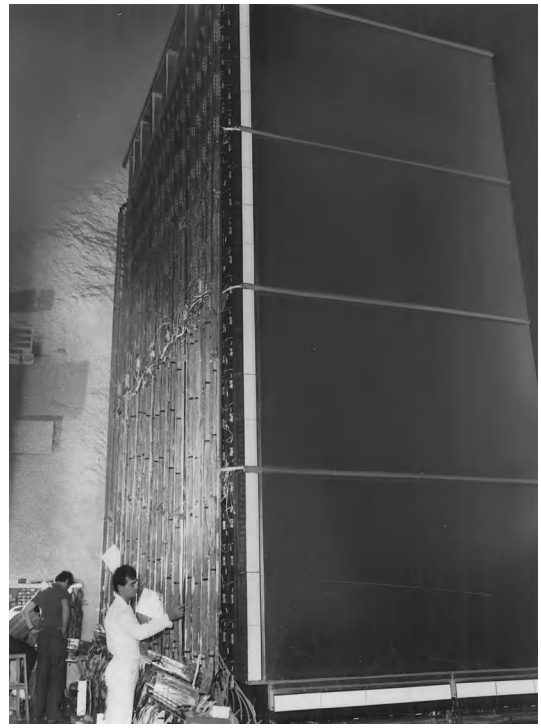

Proton decay experiment being assembled in an underground laboratory in the Frejus tunnel under the French Alps.

See GEIGER-MÜLLER COUNTER; NEUTRINO; PARTICLE DETECTOR.

If the proton is observed to decay, this new interaction will also have profound consequences for understanding of cosmology. The very early times of the big bang (about 10^{-30} s) are characterized by energies so high that the same grand unified interaction which would allow proton decay would also completely determine the subsequent evolution of the universe. This could then explain the remarkable astrophysical observation that the universe appears to contain only matter and not an equal amount of antimatter. The present standard theories of fundamental particle interactions, quantum chromodynamics and the electroweak theory, allow only processes in which an equal number of particles and antiparticles is created. These fundamental theories are therefore inconsistent with the apparent absence of antimatter in the universe. *See* COSMOLOGY; ELEMENTARY PARTICLE; INFLATIONARY UNIVERSE COSMOLOGY.

Thomas H. Fields

Bibliography. M. Gaillard, Toward a unified picture of elementary particle interactions, *Amer. Sci.*, 70:506–514, 1982; H. Georgi, A unified theory of elementary particles and forces, *Sci. Amer.*, 244(4):48–63, 1981; R. G. Roberts, *The Structure of the Proton: Deep Inelastic Scattering*, 1993; W. B. Rolnick, *The Fundamental Particles and Their Interactions*, 1994; L. Sulak, Waiting for the proton to decay, *Amer. Sci.*, 70:616–625, 1982; S. Weinberg, The decay of the proton, *Sci. Amer.*, 244(6):64–75, 1981; V. F. Weisskopf, The origin of the universe, *Amer. Sci.*, 71:473–480, 1983.

Proton-induced x-ray emission (PIXE)

A highly sensitive analytic technique for determining the composition of elements in small samples. Proton-induced x-ray emission (PIXE) is a nondestructive method capable of analyzing many elements simultaneously at concentrations of parts per million in samples as small as nanograms. PIXE has gained acceptance in many disciplines; for example, it is the preferred technique for surveying the environment for trace quantities of such toxic elements as lead and arsenic. There has also been a rapid development in the use of focused proton beams for PIXE studies in order to produce two-dimensional maps of the elements at spatial resolutions of micrometers.

Apparatus. The typical PIXE apparatus uses a small Van de Graaff machine to accelerate the protons which are then guided to the sample. Nominal proton energies are between 1 and 4 MeV; too low an energy gives too little signal while too high an energy produces too high a background. The energetic protons ionize some of the atoms in the sample, and the subsequent filling of empty inner orbits results in the characteristic x-rays. These monoenergetic x-rays emitted by the sample are then efficiently counted in a high-resolution silicon (lithium) detector which is sensitive to the x-rays of all elements heavier than sodium. *See* JUNCTION DETECTOR; X-RAYS.

Advantages. The advantages of PIXE over electron-induced x-ray techniques derive from the heaviness of the proton which permits it to move through matter with little deflection. The absence of scattering results in negligible continuous radiation (bremsstrahlung) so that the backgrounds under the characteristic x-ray lines come mainly from secondary effects. As a result, proton-induced x-ray techniques are two to three orders of magnitude more sensitive to trace elements than are techniques based on electron beams. Another important consequence is that the proton beam can be extracted from the vacuum so that samples, in particular of biological material, can be studied in controlled atmospheric environments. *See* BREMSSTRAHLUNG.

Application. PIXE was developed in European laboratories in the early 1970s. At the University of Lund, Sweden, and at the Niels Bohr Institute in Copenhagen, the emphasis was on using broad beams to obtain great sensitivity to trace elements in bulk samples; at the Harwell National Laboratory in England the emphasis was on focused beams to obtain spatial resolution at some sacrifice in analytic sensitivity. Some 40 to 50 PIXE laboratories now carry out investigations in such diverse fields as materials science, earth and planetary sciences, biomedicine, environment, and archeology. A number of these laboratories have sophisticated facilities in which samples are scanned under computer control by micrometer-size beams. Previously the high cost of the proton accelerator confined PIXE to laboratories which used existing Van de Graaffs, which were constructed mainly in the 1950s and 1960s for nuclear studies. Several companies now market small accelerators which are suited to PIXE, and commercial laboratories offer analytic PIXE services to the technology community. *See* ACTIVATION ANALYSIS.

Lee Grodzins

Protostar

A hydrostatic sphere of gas in the process of becoming a star. A protostar forms by the gravitational collapse of a dense core within a giant cloud of dust and molecular gas (mostly H_2). As the core collapses from the inside out, it surrounds a central protostar with a cocoon of accreting dust and gas that hides it from view at optical wavelengths. Observations at longer wavelengths penetrate this material and reveal that the protostar is radiating due to the impact of infalling gas and dust. Most of the envelope material accretes first onto a circumstellar disk, from which it is then conveyed to the protostellar surface. As the envelope dissipates around a low-mass star such as the Sun, the protostar becomes visible as a pre-main-sequence T Tauri star that continues for a time to add to its mass from the accretion disk. It becomes a full-fledged star when the core temperature reaches the level required for nuclear fusion of hydrogen (temperature $T \sim 10^7$ K). For stars at least 10 times more massive than the Sun, nuclear burning begins before infalling material is dissipated, and the resulting high-intensity radiation quickly clears away the remaining envelope. *See* T TAURI STAR.

Observations. Dark voids in the Milky Way were believed early on to be the result of intervening cold dust clouds that might serve as sites for star formation. Pioneering infrared astronomical surveys in the 1960s and 1970s revealed the first candidate protostars embedded deep within these clouds. Groundwork for a modern understanding of star formation was laid in the following decade. Improved surveys with ground-based telescopes led to an empirical classification of young stars according to their distribution of radiation with wavelength; those with radiation that peaked at the longest wavelengths were interpreted as youngest. Surveys of molecular clouds with submillimeter and millimeter-wave techniques provided the first maps of individual cores in molecular clouds; the Doppler shift of their molecular lines revealed the unexpected presence of bipolar molecular outflows and established their rotation rates. Finally, the *Infrared Astronomical Satellite* (*IRAS*) detected infrared protostars in the center of many of these cores and provided a better measure of their radiation at wavelengths of 12 to 100 micronmeters. Cores without *IRAS* detections were believed to reveal conditions in pre-protostellar cores. However, subsequent observations with the *Spitzer Space Telescope* extended infrared studies to much fainter levels and showed that at least some of these cores have infrared sources that are either proto-brown dwarfs or protostars in their earliest phases of growth. Since the mid-1990s, high-resolution images of the disks and jets from protostars and T Tauri stars have also been obtained with millimeter-wave interferometers, large ground-based telescopes, and the *Hubble Space Telescope*. *See* DOPPLER EFFECT; HUBBLE

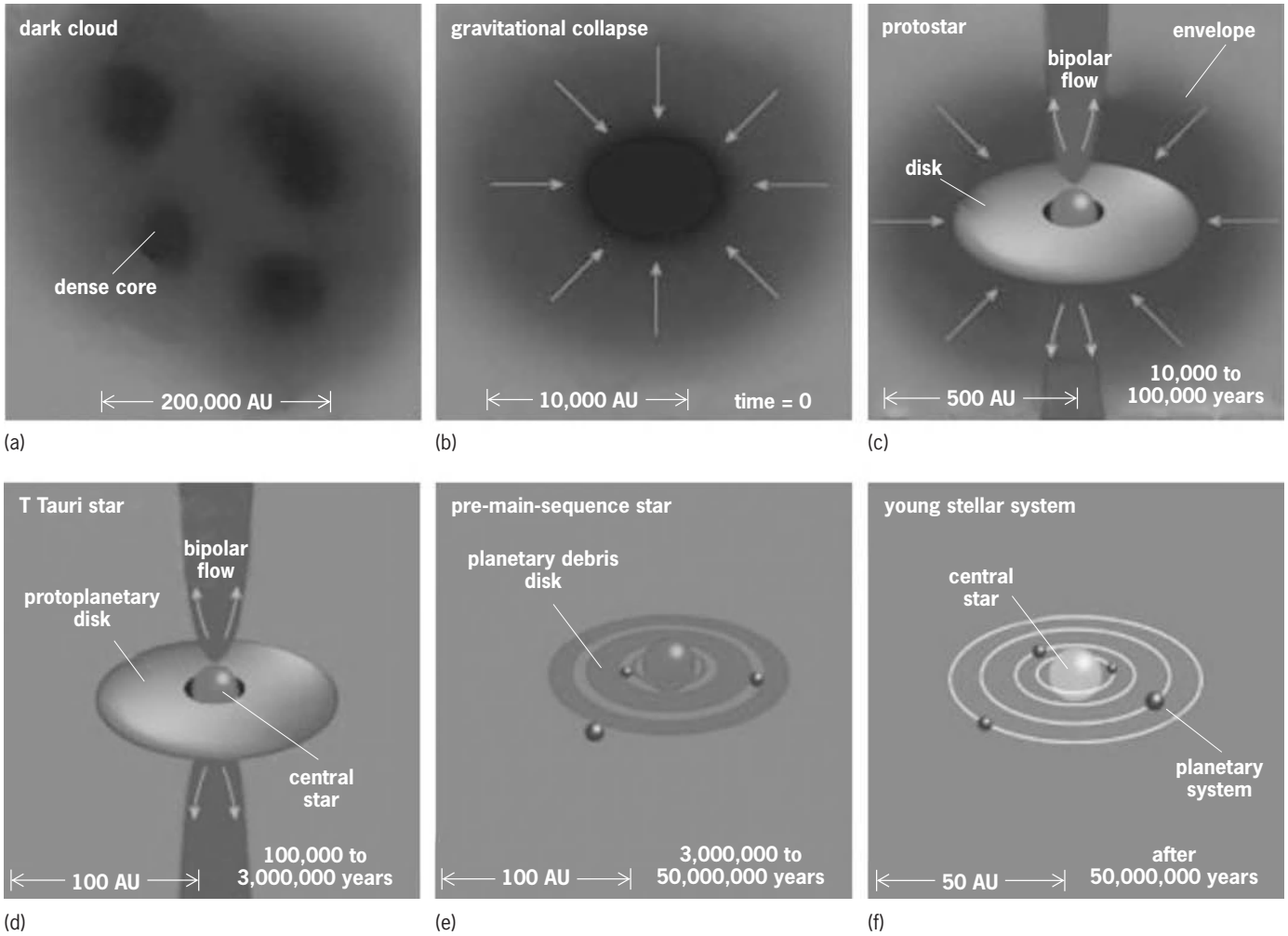

Early evolution of a protostar. (*a*) Dense cores appear within in a larger molecular cloud. AU = astronomical unit. (*b*) Some cores become unstable to gravitational collapse. (*c*) Mass rapidly accumulates on a circumstellar disk accreting onto the star and fueling a bipolar outflow. An opaque dusty envelope hides the protostar from view at optical wavelengths. (*d*) The dusty envelope is dissipated, and the protostar becomes visible as T Tauri star. (*e*) The disk disperses, perhaps through planet formation, while the star is still in a pre-main-sequence phase. (*f*) A true main-sequence star emerges when core temperature is hot enough for nuclear fusion. (*After T. Greene, Protostars, Amer. Scientist, 89:316–325, 2001*)

SPACE TELESCOPE; INFRARED ASTRONOMY; MOLECULAR CLOUD; SPITZER SPACE TELESCOPE; SUBMILLIMETER ASTRONOMY.

Evolution. On the basis of observations and theoretical studies, astronomers now recognize distinct evolutionary phases in the life of a protostar (see **illustration**). Cold condensations in molecular clouds appear initially as the result of turbulent conditions in the molecular gas with sizes of a fraction of a parsec (which is approximately 200,000 astronomical units), temperatures of 10 to 50 K, and densities of 10^4 to 10^5 atoms per cubic centimeter, significantly more than the surrounding medium (illus. *a*). James Jeans showed in the early twentieth century that such a core would collapse under its own gravity if it exceeded a critical mass for its size. During this earliest collapse phase, a protostar may be undetectable even at infrared wavelengths due to the opacity of the surrounding envelope (illus. *b*). Modern theory modifies Jeans' calculations by taking into account core rotation and magnetic fields in order to explain the observed formation of a circumstellar disk with bipolar outflow (illus. *c*). Studies of the number of objects in each phase within a single molecular cloud yield estimates of their evolutionary time scales, from 10^4 years for earliest collapse phase (illus. *b*), to 10^5 years for the optically obscured phase (illus.*c*), to a few times 10^6 years for the T Tauri phase (illus. *d*). Stars such as the Sun reach the main sequence in several times 10^7 years when their cores reach the temperatures of the order of 10^7 K that are needed to begin nuclear hydrogen burning. The disk disperses in the pre-main-sequence phase (illus. *e*). However, upon reaching the main sequence, some stars may still have residual disk material that is no longer accreting onto the star but is probably related to the formation of a young planetary system (illus. *f*).

The above description must be revised for stars much more massive than the Sun. These evolve rapidly to a highly luminous phase in which intense ultraviolet radiation ionizes the surrounding material to form an expanding H II region. The infall envelope dissipates in about 10,000 years, but toroidal

material around one or more high-mass stars may continue to accrete and fuel powerful bipolar outflows. The exact nature of the accretion process in this environment is not yet well understood. Observations with instruments now under development such as the Atacama Large Millimeter Array and both the *Herschel* and the *James Webb* space telescopes will undoubtedly shed more light on this problem. *See* INTERSTELLAR MATTER; NEBULA; STELLAR EVOLUTION.

David Koerner

Bibliography. P. André, D. Ward-Thompson, and M. Barsony, From prestellar cores to protostars: The initial conditions of star formation, in V. Mannings, A. P. Boss, and S. S. Russell (eds.), *Protostars and Planets IV*, pp. 59–96, University of Arizona Press, 2000; T. Greene, Protostars, *Amer. Sci.*, 89:316–325, 2001; C. J. Lada and F. H. Shu, The formation of sunlike stars, *Science*, 248:564–572, 1990; F. H. Shu, F. C. Adams, and S. Lizano, Star formation in molecular clouds: Observations and theory, *Annu. Rev. Astron. Astrophys.*, 25:23–84, 1987; S. W. Stahler, F. Palla, and P. T. P. Ho, The Formation of massive stars, in V. Mannings, A. P. Boss, and S. S. Russell (eds.), *Protostars and Planets IV*, pp. 327–351, University of Arizona Press, 2000.

Prototheria

One of the four subclasses of the class Mammalia. Prototheria contains a single order, the Monotremata. No ancestral genera of fossil monotremes are known, and the structure of the living monotremes is so specialized that the affinities of the Prototheria are largely conjectural. Most mammalogists believe that the prototheres arose from a different stock of therapsid reptiles than the one that gave rise to the Theria.

This origin from a different stock would mean that the history of the monotremes has been separate from the history of other mammals for at least 1.75×10^8 years.

No fossils earlier than the Pleistocene are known, and these come from Australia. The duckbill platypus (*Ornithorhynchus anatinus*) and several species of the echidnas, *Tachyglossus* and *Zaglossus*, are living representatives of this group. They are found in the Australian region. Everything indicates that the Prototheria represent a very small and relatively unsuccessful group that has miraculously survived in an isolated corner of the Earth. *See* MAMMALIA; MONOTREMATA; PLATYPUS; THERAPSIDA.

D. Dwight Davis; Frederick S. Szalay

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999; J. M. Taylor, *Oxford Guide to Mammals of Australia*, 1984.

Prototype

A first or original model of hardware or software. Prototyping involves the production of functionally useful and trustworthy systems through experimentation with evolving systems. Generally, this experimentation is conducted with much user involvement in the evaluation of the prototype.

Primary uses. A primary use for prototyping is the acquisition of information that affects early product development. For example, if requirements for human-computer interfaces are ambiguous or inadequate, prototyping is frequently used to define an acceptable functional solution. It is a method for increasing the utility of user knowledge for purposes of continuing development to a final product. Information obtained through prototyping is important to designers, managers, and users in identifying issues and problems. Prototyping conserves time and resources prior to the commitment of effort to construct a final product. *See* HUMAN-COMPUTER INTERACTION.

Prototyping is an important way to obtain user input related to a product that is being developed. The major ways in which the user participates in development of a product include involvement during requirements identification, review of operational functionality and environmental changes that influence installation and maintenance, and work with designers to gain a better understanding of product application. In large-scale complex systems, such as real-time embedded avionics systems, it is often necessary to construct full-scale models of portions of the system, both hardware and software, to demonstrate the feasibility of the proposed solution. *See* EMBEDDED SYSTEMS.

Prototyping increases worker productivity through modeling of the development process. It provides support for requirements identification and the transformation of requirements into product specifications, as well as for understanding the operational environment. Prototyping assists in understanding the intended and desired system-level functionality in terms of what the product should accomplish. For example, in the development of a database to support a complex system, all the operations involving database structures, database query or report features, and data storage are prototyped to assure performance, size, and access of the database. *See* DATABASE MANAGEMENT SYSTEM.

Classification. Several schemes are used to classify prototypes. Examples include functional, structural, purposeful, operational, rapid, experimental, and exploratory. Functional prototyping helps demonstrate that the behavior of the product is what the user intended. Structural prototyping aids in determining that the particular design approach used is feasible and is evolving toward production. Purposeful prototyping is used to ascertain that user requirements and product performance are consistent. Operational prototyping is done in the field as a temporary add-on to an existing system. Rapid prototyping supports early definition and validation of requirements. Experimental prototyping is used for the investigation of alternative approaches to a solution. Exploratory prototyping is used to assist in the formulation of a problem or issue.

Prototypes are also classified according to the uses intended. Examples include architectural, organizational, and simulation or animation. Architectural prototyping assists in visualization of the allocation of functions to system structure. Organizational prototyping assists in the visualization of functional operations along enterprise lines. Simulation and animation prototyping serve to test the viability of proposed solutions and provide graphical representations of the solutions. *See* SIMULATION.

A prototype is generally a functionally immature model of a proposed product that is built to explore requirements, investigate alternative approaches, or demonstrate model feasibility. The model may or may not evolve into a mature, functionally useful product.

Methods of use. Prototyping generally is applied in stages. Methods used in these stages may be classified into three groups. The first, concept prototyping, is used to encourage dialogue between the user and the developer to reach mutual agreement that a proposed concept will satisfy user requirements. The concept prototype can vary from a totally static demonstration model to one involving system dynamics. A structural, functional, or purposeful prototype can be built as a concept prototype. Rapid prototyping, involving interface designs, animation of the activity, or multimedia representations, is most often used for concept development.

The next stage, laboratory prototyping, involves a working model and begins as soon as the concept prototype is accepted. Advantage is taken of a friendlier development environment and development tools than might be available in the actual operating environment. Structural (incremental or evolutionary) or functional (experimental) prototyping would be utilized at this stage.

Field prototyping begins after the laboratory prototype is accepted and becomes the first unit to be installed in the field. It is possible to test and evaluate a field prototype under actual operating conditions. This is still regarded as a prototype since user feedback may result in modifications. A field prototype would usually be a structural prototype.

Application phases. There are seven phases in the application of prototyping. These move from the initial request, through selection of an approach, to selection of the methods, to construction and validation.

The first phase involves identifying and evaluating the request for software development services. Preliminary information requirements concerning user needs are determined. The most appropriate approach to product development is identified; in particular, whether prototyping is appropriate is determined in this phase. Influential factors include application area and complexity, user experiential familiarity with the application area, and developer experience with prototyping.

The remaining phases involve identifying the type of prototyping that is most appropriate, which will be some form of structural, functional, or purposeful prototyping; developing a set of user and product requirements appropriate for the prototyping approach selected and the resulting prototype; identifying a set of design specifications for the prototype; creating, testing, and refining the prototype; involving the client or user in use of the prototype in an appropriate way relative to accomplishment of prescribed prototyping tasks; and determining whether the prototype developed is acceptable. If the prototype is acceptable, the next phase in the product development life cycle can be undertaken. Otherwise, the designer must iterate back to identifying the type of prototyping that is most appropriate, if it is believed that the initial judgment concerning prototyping type is flawed; or to refining the prototype if the existing prototype is appropriate.

Characteristics and uses. In many hardware and software development projects, the first prototype product built is barely usable. It is usually too slow, too big, too awkward in use. Hence, the term throwaway prototype is generally applied to describe this early use of prototyping. Usually this is due to lack of understanding of user requirements. There is no alternative but to start again and build a redesigned version in which these problems are solved.

A developmental prototyping approach for incremental design of subsystems is often used to reduce the risk involved in building a system-level prototype. In this prototyping environment an incremental approach to rapid prototyping of subsystems development is used. This provides for management oversight of the entire process to assure that resource usage is effective and efficient. Product assurance is implemented throughout the process to make certain that the prototype operation contains the necessary components to satisfy subsystem requirements. Requirements analysis is performed and reviewed, then incremental specifications are developed and reviewed, followed by design of the approved specifications, and completed by implementation of the product.

The decision concerning whether a developmental prototype should be refined into the complete product or be thrown away should be based on the costs and benefits of each alternative. Two key factors to be addressed are the degree of functionality in the prototype and whether or not the prototype design supports a maintainable product. Response to these concerns determines whether the prototype is worth the investment of more effort and money.

The following example illustrates the use of prototypes: A firm is experiencing difficulties in its correspondence operation and is unable to provide accurate and timely responses and reports to customers. Resolution is required immediately, but system requirements are not clearly stated. A description of the flow shows correspondence is received, reviewed, filed, and then manually distributed to executive offices for review and response generation. The process involves the selection of offices that are to generate a response, called action offices, and selection of reviewers to comment on the

proposed response. There is no automated means of maintaining a list of action offices and reviewers appropriate for different subject areas. A packet, containing the correspondence and any background material, is manually circulated throughout the operation, starting and ending with an executive secretariat. Correspondence may come from any number of sources and may need to be responded to in a timely manner. Corporate reports and other materials need to be distributed to grantees, supplier organizations, and other interested parties, both inside and outside the company. Currently, copies of reports are generated on paper and mailed to outside individuals. The listings of groups to which reports are sent are not maintained in the corporate management information system.

Solutions were proposed using rapid prototyping together with animation of current and proposed document flows. Current document flows showed multiple handling of the same document by the same desk and no automatic tracking of the location and status. No reporting capabilities were provided, and no database of reports was available. The rapid prototype animation allowed executive secretariat personnel to view alternative approaches and finally settle on a single approach that satisfied correspondence management needs. The rapid prototype animation was performed on a personal computer using commercially available software.

Advantages. The major potential advantages to the use of prototyping relate to the ability to rapidly construct a product that is representative of the final product at relatively low cost and effort. Nearly all products are developed by using a life cycle, and prototyping enhances the descriptive power of the product life cycle development process. Prototyping allows the usual product evolution to take place in a modeling environment that simulates the natural environment without the expenses associated with replication of that environment. It provides control over the product development process by incorporating aspects of resource management, configuration management, verification, and validation in early phases as well as later phases of the life cycle. Prototyping enables the construction of larger and more complex products with smaller development teams and promotes user feedback early in the development process. *See* MODEL THEORY; SOFTWARE ENGINEERING; SYSTEMS ENGINEERING.

James D. Palmer

Bibliography. W. W. Agresti (ed.), *New Paradigms for Software Development*, 1986; V. E. Church et al., An approach for assessing software prototypes, *ACM Software Eng. Notes*, 11(3):65–76, July 1986; S. Hekmatopour and D. Ince, *Software Prototyping, Formal Methods and VDM*, 1988; H. Hunke (ed.), *Software Engineering Environments*, 1981; J. D. Palmer, J. O. Berkey, and V. Kountouris, Software development environments for reusability and prototyping, *Large Scale Sys.*, 12:201–216, 1987; A. P. Sage and J. D. Palmer, *Software Systems Engineering*, 1990; I. Sommerville, *Software Engineering*, 4th ed., 1992.

Protozoa

A group of eukaryotic microorganisms traditionally classified as one of the five animal kingdoms. Although the name signifies primitive animals, some protozoa (phytoflagellates and slime molds) show enough plantlike characteristics to justify claims that they are plants. This apparent conflict may be reconciled to some extent by considering protozoa as descendants of microorganisms which preceded plants and animals. Such an assumption might account for *Euglena gracilis*, a green flagellate which can fix carbon dioxide (CO_2) in light as algae do but, unlike algae, can also carry on a plantlike fixation in darkness. In addition, *E. gracilis* can fix carbon dioxide by pathways characteristic of animals. Inclusion of such metabolically unspecialized organisms, along with certain photoautotrophic chlorophyll-bearing species, might imply that the phylum is partly an arrangement for taxonomic convenience rather than a phylogenetically homogeneous group.

Protozoa are almost as widely distributed as bacteria. Free-living types occur in soil, wet sand, and in fresh, brackish, and salt waters. Protozoa of the soil and sand live in films of moisture on the particles. Habitats of endoparasites vary. Some are intracellular, such as malarial parasites in vertebrates, which are typical Coccidia in most of the cycle. Other parasites, such as *Entamoeba histolytica*, invade tissues but not individual cells. Most trypanosomes live in the blood plasma of vertebrate hosts. Many other parasites live in the lumen of the digestive tract or sometimes in coelomic cavities of invertebrates, as do certain gregarines. *See* COCCIDIA; GREGARINIA; TRYPANOSOMATIDAE.

Many protozoa are uninucleate, others binucleate or multinucleate, and the number also may vary at different stages in a life cycle. Size ranges from less than 3 micrometers to a number of centimeters, as in various Mycetozoia. Colonies are known in flagellates, ciliates, and Sarcodina. Although marked differentiation of reproductive and somatic zooids characterizes certain colonies, such as *Volvox*, protozoa have not developed tissues and organs.

Taxonomy

Various systems of classification have been proposed, and none is yet accepted by all protozoologists. One of those sponsored by the Society of Protozoologists follows:

Phylum Protozoa
- Subphylum 1: Sarcomastigophora
  - Superclass 1: Mastigophora
    - Class 1: Phytamastigophorea
      - Order 1: Chrysomonadida
      - Order 2: Silicoflagellida
      - Order 3: Coccolithophorida
      - Order 4: Heterochlorida*
      - Order 5: Cryptomonadida
      - Order 6: Dinoflagellida
      - Order 7: Ebriida

Order 8: Euglenida
Order 9: Chloromonadida
Order 10: Volvocida (Phytomonadida)
Class 2: Zoomastigophorea
Order 1: Choanoflagellida
Order 2: Bicosoecida
Order 3: Rhizomastigida
Order 4: Kinetoplastida
Order 5: Retortomonadida (Protomastigida)
Order 5: Retortomonadida
Order 6: Diplomonadida
Order 7: Oxymonadida
Order 8: Trichomonadida
Order 9: Hypermastigida
Superclass 2: Sarcodina
Class 1: Rhizopodea
Subclass 1: Lobosia
Order 1: Amoebida
Order 2: Arcellinida
Subclass 2: Filosia
Order 1: Aconchulinida
Order 2: Gromiida
Subclass 3: Granuloreticulosia
Order 1: Athalamida
Order 2: Foraminiferida
Order 3: Xenophyophorida
Subclass 4: Mycetozoia
Order 1: Acrasida
Order 2: Eumycetozoida
Order 3: Plasmodiophorida
Subclass 5: Labyrinthulia
Class 2: Piroplasmea
Class 3: Actinopodea
Subclass 1: Radiolaria
Order 1: Porulosida
Order 2: Oculosida
Subclass 2: Acantharia
Order 1: Acanthometrida
Order 2: Acanthophractida
Subclass 3: Heliozoia
Order 1: Actinophryida
Order 2: Centrohelida
Order 3: Desmothoracida
Subclass 4: Proteomyxidia
Subphylum 2: Sporozoa
Class 1: Telosporea
Subclass 1: Gregarinia
Order 1: Archigregarinida
Order 2: Eugregarinida
Order 3: Neogregarinida
Subclass 2: Coccidia
Order 1: Protococida
Order 2: Eucoccida
Suborder 1: Adeleina*
Suborder 2: Eimeriina*
Suborder 3: Haemosporina
Class 2: Toxoplasmea
Order 1: Toxoplasmida
Class 3: Haplosporea
Order 1: Haplosporida
Order 2: Sarcosporida
Order 4: Entodiniomorphida
Subphylum 3: Cnidospora
Class 1: Myxosporidea
Order 1: Myxosporida
Order 2: Actinomyxida
Order 3: Helicosporida
Class 2: Microsporidea
Class 1: Ciliatea
Subclass 1: Holotrichia
Order 1: Gymnostomatida
Order 2: Trichostomatida
Order 3: Chonotrichida
Order 4: Apostomatida
Order 5: Astomatida
Order 6: Hymenostomatida
Order 7: Thigmotrichida
Subclass 2: Peritrichia
Subclass 3: Suctoria
Order 1: Suctorida*
Subclass 4: Spirotrichia
Order 1: Heterotrichida
Order 2: Oligotrichida
Order 3: Tintinnida
Order 4: Entodiniomorphida
Order 5: Odontostomatida
Order 6: Hypotrichida

See separate articles on each taxonomic group except those marked with an asterisk.

Morphology

A protozoon may be a plastic organism (ameboid type) but changes in form are often restricted by the pellicle. In *Trypanosoma lewisi*, parallel microtubules beneath the pellicle may support this layer and may also, as probably contractile structures, cause the typical undulations of these flagellates. Although protozoan form varies considerably, there is a tendency toward universal symmetry in floating species and radial symmetry in sessile types. Bilateral symmetry is rare; most swimmers show appreciable spiral torsion.

A protective layer is often secreted outside the pellicle, although the pellicle itself may be strengthened by incorporation of minerals. Secreted coverings may fit closely, for example, the cellulose-containing theca of Phytomonadida and Dinoflagellida, analogous to the cell wall in higher plants. The dinoflagellate theca (**Fig. 1***a*) may be composed of plates arranged in a specific pattern. Tests, as seen in Rhizopodea (Arcellinida, Gromida, Foraminiferida), may be composed mostly of inorganic material, although organic (chitinous) tests occur in certain species. Siliceous skeletons, often elaborate, characterize the Radiolaria (Fig. 1*d* and **Fig. 2***a–e*). A vase-shaped lorica, from which the anterior part of the organism or its appendages may be extended, occurs in certain flagellates (Fig. 1*b*) and ciliates (Fig. 1*c*). Certain marine ciliates (Tintinnida) are actively swimming loricate forms.

Colonies and aggregates. An important feature of colonies is the material which holds the component zooids together. In spheroid colonies (**Fig. 3**) a

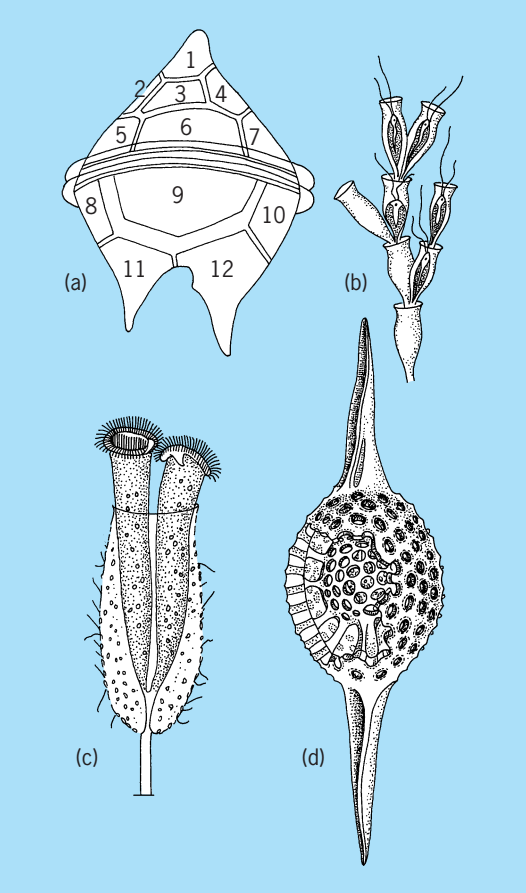

Fig. 1. External coverings of protozoa. (*a*) Theca of dinoflagellate (*Peridinium*), showing separate plates. (*b*) Lorica of a colonial chrysomonad, *Dinobryon*. (*c*) Two zooids within a lorica of a peritrich, *Cothurnia*. (*d*) A radiolarian skeleton, siliceous type. (*After L. H. Hyman, The Invertebrates, vol. 1, McGraw-Hill, 1940*)

matrix is secreted during development of the colony. Spheroid colonies of a few Radiolaria (*Colozoum*), unusually large examples, may measure 4–6 cm across. In arboroid colonies (**Fig. 4***b* and *f*) a branching pattern is produced by a branching stalk (Fig. 4*b*) or by attachment of loricae to one another (Fig. 1*b*). Dispersal of such colonial species sometimes involves migratory stages. In certain spheroid colonies of *Pleodorina* (Fig. 3*b* and *c*) and *Volvox*, reproductive and somatic zooids are differentiated. More often members of a colony are similar. In addition to colonial types, certain protozoa may form aggregates by not separating promptly after fission. Palmella stages of Phytomonadida, analogous to spheroid colonies, are aggregates of nonflagellated organisms. Chains of *Gonyaulax catenella* and certain other Dinoflagellida are another type of aggregate.

Flagella. These organelles occur in active stages of Mastigophora and flagellated stages of certain Sarcodina and Sporozoa. A flagellum consists of a sheath enclosing a matrix in which an axoneme extends from the cytoplasm to the flagellar tip. In certain groups the sheath shows lateral fibrils (mastigonemes) which increase the surface area and also may modify direction of the thrust effective locomotion. In electron micrographs axonemes show two central and nine pairs of peripheral microtubules. The base (blepharoplast) in certain species may contain only nine peripheral triplets. The blepharoplast is considered a self-replicating structure from which new flagella arise in fission. Flagellar number ranges from one or two for most free-living species to many as seen on certain symbionts of termites and wood roaches. Most flagella extend forward in swimming; others extend posteriorly, sometimes driving the body forward, sometimes serving as a ski in gliding, or forming the margin of an undulating membrane. Certain organelles, as the kinetoplast (Kinetoplastida), axostyle, parabasal body (Trichomonadida), or sometimes other structures, may be attached to or adjoin the blepharoplasts. *See* CILIA AND FLAGELLA; KINETOPLASTIDA.

Flagellar movements include planar sinusoidal waves, helical waves, and lateroposterior strokes. When a planar wave travels from base to tip, the force is directed basally, pushing the organism (*Ceratium* and others). However, a base-to-tip beat pulls *Ochromonas*, a chrysomonad, the mastigonemes changing a potential push into a pull. The flagellate is pulled (trypanosomes) when a planar wave travels from tip to base. Helical waves from tip to base (trichomonad undulating membrane) push the flagellate. A similar wave, from tip to base, pulls in *Peranema*. Lateroposterior beating (*Monocercomonas neosepsorum* and certain other trichomonads) more or less resembles a breast stroke in human swimming. An unusual jet propulsion observed for *Medusochloris* sp., a phytomonad, involves contractions of a medusa-shaped body. Another substitute for flagellar drive is seen in *Mixotricha paradoxa*, a parasite of an Australian termite. Most of the body is covered with "brackets," each of which bears a posterior pit into which the tip of a spirochete fits. Coordinated undulations of the spirochetes in metachronal waves, beginning at the anterior end of the flagellate, drive the organism smoothly forward. *See* MASTIGOPHORA.

Cilia. Although typically shorter than flagella, cilia are similar in structure, even to the fibrillar triplets in the base (basal granule and kinetosome). Multiplicity of cilia entails complex fibrillar systems. It has been suggested that several sets of fibrils help support the cortex and ciliary system. A more delicate fibril, or tubule (**Fig. 5**), running along each row of basal granules (*Tetrahymena*) seemingly offers little support and may, as suggested, function in coordination (Fig. 5). Such ciliates as hypotrichs, with compound organelles, may have quite complex fibrillar systems. Compound ciliary organelles represent fused bundles or rows of cilia, with the basal granules detectable in basal plates of the organelles. Membranes (longitudinal rows) occur along the right margin of the buccal cavity, if they are present. Membranelles (transverse rows) often form an adoral zone (Fig. 4*a*) along the left margin of the buccal cavity. Cirri (Fig. 4*a*), typically ventral, are fused tufts of cilia, several to 50 or more in large cirri, important in locomotion.

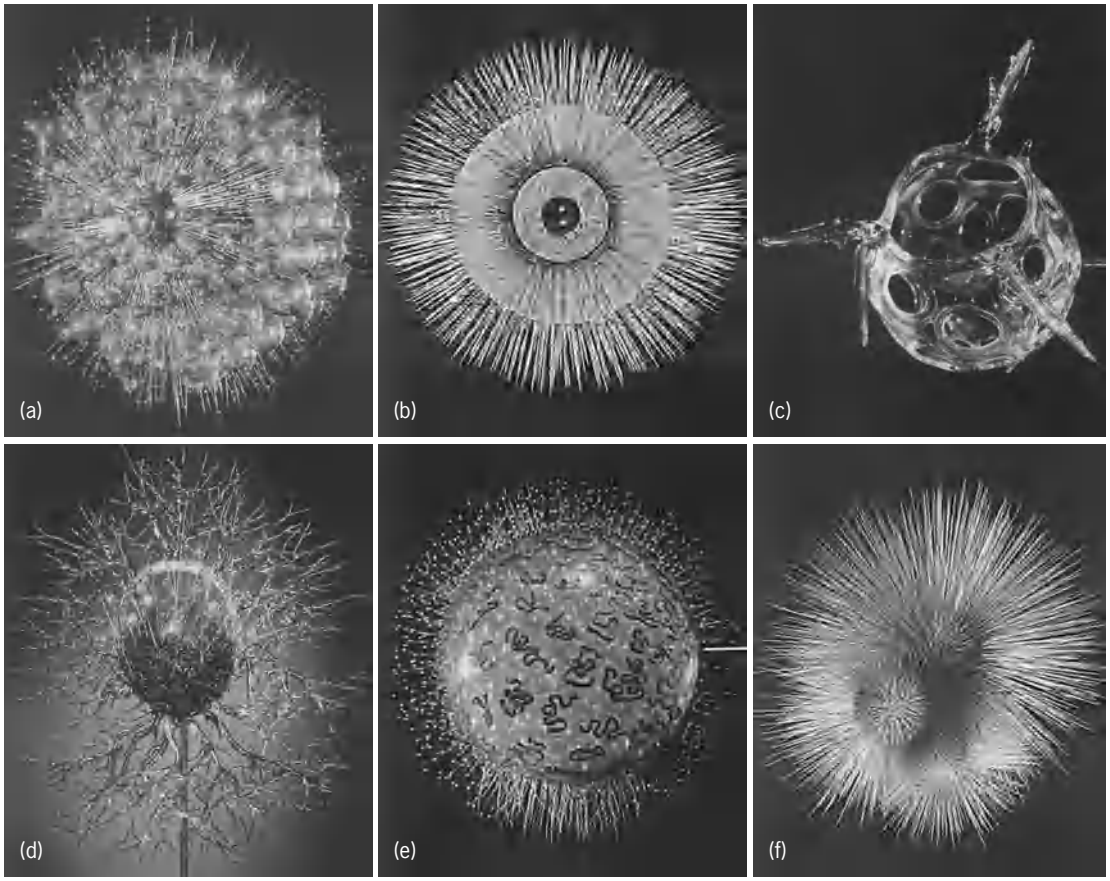

Fig. 2. Glass models of marine protozoa. Radiolarian types: (*a*) *Trypanosphaera transformata* (Indian Ocean); (*b*) *Actissa princeps* (Indian and Pacific oceans); (*c*) *Peridium spinipes* (Pacific Ocean); (*d*) *Lithocircus magnificus* (Atlantic Ocean); (*e*) *Collozoum serpentinum* (Atlantic Ocean). (*f*) Foraminiferan type: the pelagic *Globigerina bulloides* (which is found in all seas). (*American Museum of Natural History*)

Cortical myonemes (**Fig. 6**) are well developed in certain large ciliates (*Spirostomum* and *Stentor*). Similar cortical myonemes, composed of several to a dozen or more fibrils, occur in various gregarines. Ectoplasmic myonemes of certain ciliates also show longitudinal fibrils. Analogous stalk "muscles" occur in the stalks of *Vorticella* and certain colonial ciliates (*Carchesium* and *Zoothamnium*). *See* CILIOPHORA.

Pseudopodia. Two major types have been described, the contraction-hydraulic and the two-way flow types. The first are lobopodia with rounded tips and ectoplasm denser than endoplasm. The larger ones commonly contain granular endoplasm and clear ectoplasm.

According to the contraction-hydraulic theory, pseudopodia of *Amoeba proteus* arise as a result of pressure from the contracting ectoplasm on the endoplasm, forcing the latter forward, away from the posterior contraction. Local weakening of the ectoplasm results in a bulge, or developing pseudopodium. Endoplasm reaching the tip of the growing pseudopodium is diverted to the ectoplasm and converted into gel, thus becoming part of the ectoplasm. In the meantime, ectoplasm at the other end is being converted into endoplasm, which flows forward again under pressure from the ectoplasm.

Two-way flow pseudopodia include reticulopodia of Foraminiferida and related types, filoreticulopodia of Radiolaria, and axopodia of certain Heliozoia. Axopodia of some genera show an axoneme composed of microtubules extending from the centroplast; in others, of a cylinder surrounded by microtubules. In addition, similar filoreticulopodia (no axonemes) occur in Radiolaria and certain other groups. In all these, cytoplasmic flow in opposite directions occurs on opposite sides of a pseudopodium. *See* HELIOZOIA; RADIOLARIA; SARCODINA.

Trichocysts. At maturity these lie in the cortex of certain ciliates and flagellates. Mucocysts (Fig. 5) seem homogeneous in some ciliates in electron micrographs; in others, there is a dense inner zone. The mucocysts of *Tetrahymena pyriformis* show inclusions forming a "houndstooth" pattern. Analogous mucocysts occur in Euglenida and Dinoflagellida. These inclusions probably serve in secretion of cyst walls and similar materials. Filamentous trichocysts (*Paramecium* and *Frontonia*) in the cortex are homogeneous bodies attached to a differentiated tip. During discharge there is formed a cross-striated shaft, ending in a dense tip. Discharged trichocysts are replaced by endoplasmic formation of new ones which, as they mature, are moved to the usual cortical location. Similar trichocysts occur in *Oxyrrhis*, a dinoflagellate. Cross striations of the shaft show a particular periodicity (55 nanometers) in various ciliates. Fibrillar trichocysts of

Paramecium may be extruded partially to anchor a feeding ciliate. Toxicysts of *Dileptus* show, after discharge, a nonstriated tubule with a rodlike tip. Before discharge a capsule contains telescoping tubules and is closed apically by a cap. New toxicysts arise from endoplasmic vesicles which migrate to an adoral position during development. Toxicysts may, upon contact, induce paralysis of lysis of various Protozoa and sometimes paralysis of rotifers. Such organelles aid carnivores in capture of prey.

Inclusions. In addition to nuclei, food vacuoles (gastrioles) in phagotrophs, chromatophores and stigma in many phytoflagellates, water-elimination vesicles in many Protozoa, and sometimes other organelles, the cytoplasm may contain mitochondria, Golgi material, pinocytotic vacuoles, stored food materials, endoplasmic reticulum, and sometimes pigments of various kinds.

Chromatophores. These contain chlorophylls and commonly additional pigments such as carotenes and xanthophylls. Accessory pigments, varying with the group, may be abundant enough to produce yellowish-green, brown, and even blue or red chromatophores. In different species, chromatophores range from a large one (**Fig. 7***e*, *f*, and *i*) to a number of flattened bodies (Fig. 7*g* and *k*). A chromatophore contains, within a double membrane, electron-opaque lamellae separated by transparent material in which granules may be embedded. Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), chemically different from nuclear DNA and cytoplasmic RNA, occur in chromatophores of *Euglena gracilis*. DNA has been reported also from chromatophores of *Chlamydomonas moewusii*. Certain strains of *E. gracilis* have been bleached by exposure to high temperature (about 97°F or 36°C), maintenance in darkness, or treatment with streptomycin, erythromycin, pyribenzamine, and a nitrosoguanidine, among other drugs. Bleaching has been temporary or permanent with different treatments. Certain antibiotics damage chloroplast DNA (preventable by supplementary hydroxylamine). Plastids, containing starch but no chlorophyll, lie below the cortical mitochondria in *Polytoma uvella*. *See* CHROMATOPHORE.

Pyrenoids, ranging from a single body to a cluster, are associated with the chromatophore in certain phytoflagellates but not in others (Fig. 7*b*, *c*, and *j*). Taxonomic distribution may be erratic; for example, certain species of *Euglena* have pyrenoids but others do not. In *Chlamydomonas* the pyrenoid, lying within a chloroplast, is surrounded by starch granules. Electron micrographs show delicate tubules extending from the pyrenoid to the lamellae of the chloroplast.

Stigma. A stigma (Fig. 7*d*) occurs in many green flagellates and a few of their colorless relatives. Observations on behavior have suggested that the stigma functions in orientation with relation to a light source. Also, stigmaless mutants (*Chlamydomonas*) respond to light more erratically than the green parent stock. Whether the stigma is a photoreceptor or merely a shield for a photosensitive organelle is debatable. A typical stigma shows globules containing carotenoid pigment. The organelle lies on the wall of the reservoir, as in *Euglena*, or may be adherent to or lie within a chromatophore, as in the phytomonads. The pigment globules (100 nm or larger) are packed into one or more plates forming the stigma. In experimentally bleached *Euglena*, the fate of the stigma is not closely correlated with that of the chromatophores. A more complex ocellus of certain dinoflagellates (Fig. 7*b* and *m*) shows a lenslike body in a pigmented cup. *See* PHOTORECEPTION.

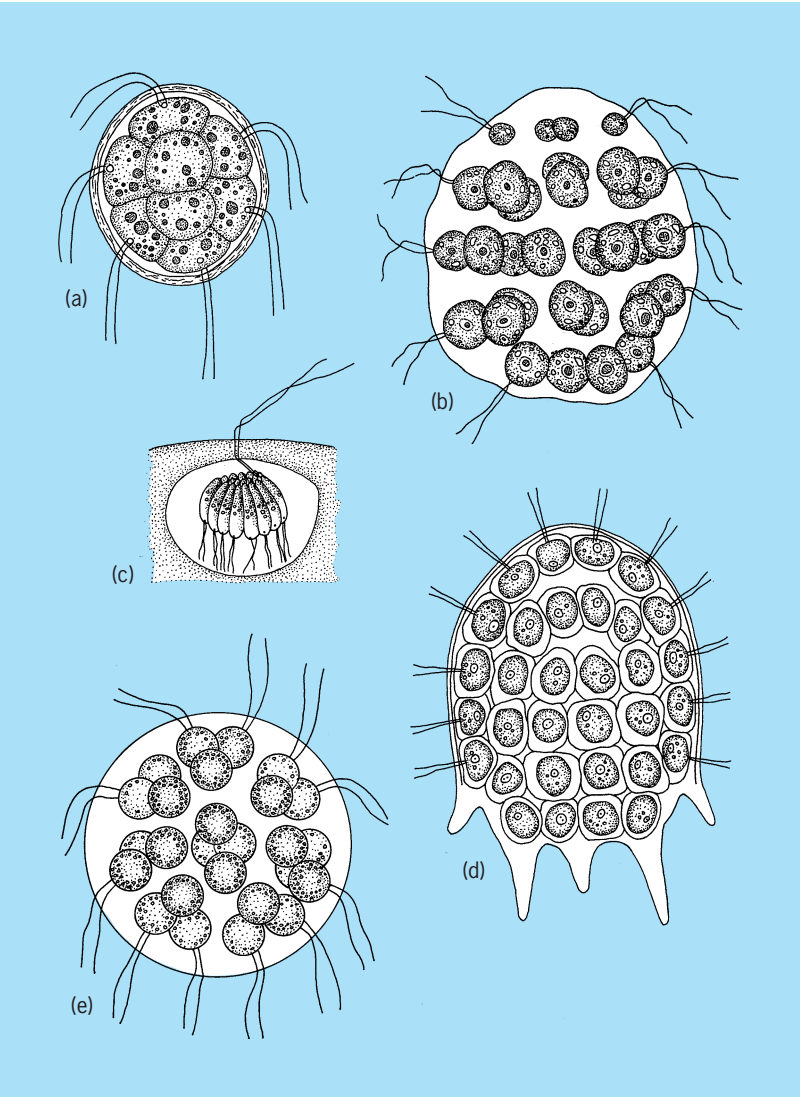

Fig. 3. Colonies of Phytomonadida. (*a*) *Pandorina*. (*b*) *Pleodorina illinoisiensis*. (*c*) Sperm packet, *Pleodorina*. (*d*) *Platydorina caudata*. (*e*) *Eudorina*. (*After L. H. Hyman, The Invertebrates, vol. 1, McGraw-Hill, 1940*)

Cytoplasmic pigments. These occur in many phytoflagellates and certain other protozoa, such as *Blepharisma* and *Stentor*, ranging from violet to red. Reddish pigments are common in Euglenida and Phytomonadida. In *Haematococcus pluvialis* production of such pigments (largely astaxanthin) in abundance results in a "hematocyst" in which the green color is masked. In *Euglena rubra* a similar pigment, normally endoplasmic, is moved to the periphery upon exposure to bright light or

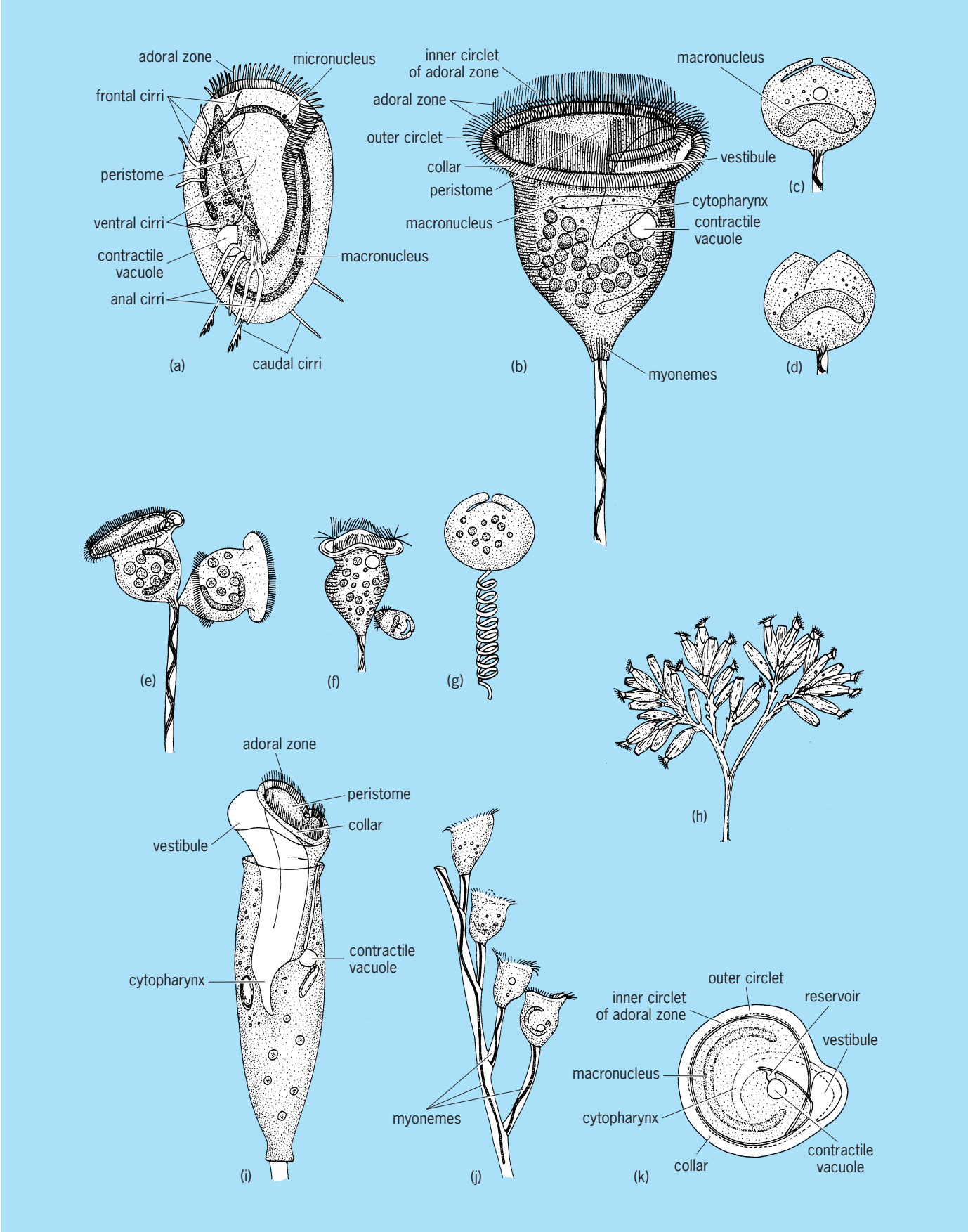

Fig. 4. Hypotricha and Peritricha. (*a*) *Euplotes*. (*b*) *Vorticella*. (*c–e*) Stages of fission in *Vorticella*. (*f*) Macro- and microconjugants in *Vorticella*. (*g*) *Vorticella*, stalk contracted. (*h*) Arboroid colony, *Opercularia*. (*i*) Zooid from *Opercularia* colony. (*j*) Portion of colony (*Carchesium*) showing stalk muscle. (*k*) Peristome of *Vorticella*, seen from above. (*After L. H. Hyman, The Invertebrates, vol. 1, McGraw-Hill, 1940*)

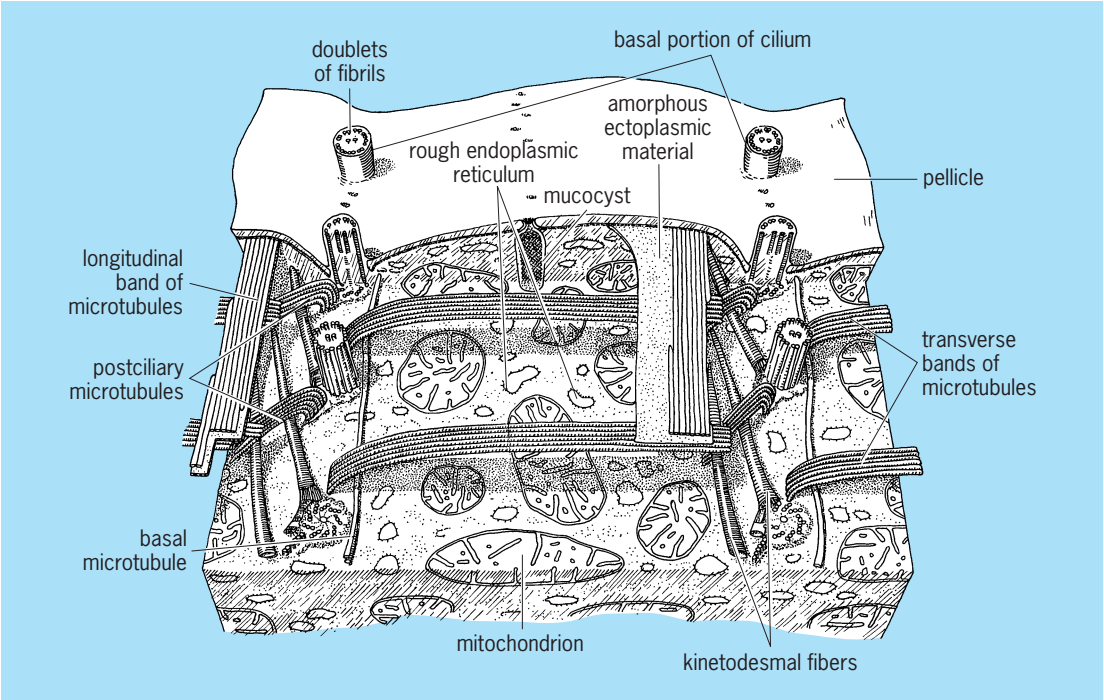

Fig. 5. Schematic of cortex of *Tetrahymena pyriformis*. (*After Allen, J. Protozool., 14:553–565, 1967*)

temperatures near 104°F (40°C). Red blooms may develop in ponds during hot summers.

Certain ciliates also produce cytoplasmic pigments. Of these the pink pigments, similar to hypericin of plants, of *Blepharisma undulans* are converted by bright light into products toxic to microorganisms, including *Blepharisma*. Pigment extracts, tested in light, show similar toxicity. An analogous material, ichthyotoxin, produced by a flagellate (*Prymnesium parvum*), is inactivated in bright light.

Mitochondria. Although apparently absent in certain anaerobes, these normally occur in aerobic protozoa **Fig. 8***a–c*). These inclusions are usually elongated and may fuse into nets in certain species (Fig. 8*a*). Morphology may be influenced by the condition of the organism. When reproduction slows down in a culture of *Tetrahymena pyriformis*, the previous ovoid and peripheral mitochondria round up and become more widely distributed. Also, ribosomes which are normally clumped along mitochondrial villi tend to disappear. DNA has been detected in mitochondria of several species. Oxidative phosphorylation seems to be associated with protozoan mitochondria (*Euglena*), as would be expected. *See* MITOCHONDRIA.

Electron micrographs indicate two general kinds of protozoan mitochondria. In one, folds of the inner membrane form cristae extending into the central lumen; this variety, characteristic of higher animals, occurs in certain phytoflagellates (*Euglena* and chrysomonads) and apparently certain trypanosomes. The kinetoplast may be continuous with such a mitochondrion, and metamorphosis from a leishmanial to a leptomonad stage may involve an increase in the size and complexity of trypanosomid mitochondria. In the other type (amebas, ciliates, and certain parasitic flagellates), cristae are replaced by microvilli.

Lysosomes. Small vacuoles (0.5–1.0 μm) in *Tetrahymena pyriformis* (**Fig. 9**) are considered lysosomes containing acid hydrolases (not yet involved in digestion) and probably arising from the rough endoplasmic reticulum. *See* ENDOPLASMIC RETICULUM; LYSOSOME.

Golgi material. In electron micrographs a typical Golgi body is a stack of flattened vesicles, sometimes up to 30 (*Peranema*). Although the identity of protozoan Golgi material is confused in older literature, some of the described Golgi elements should prove to be such. Although functional significance in Protozoa is not yet certain, some workers favor a secretory activity of these inclusions. *See* GOLGI APPARATUS.

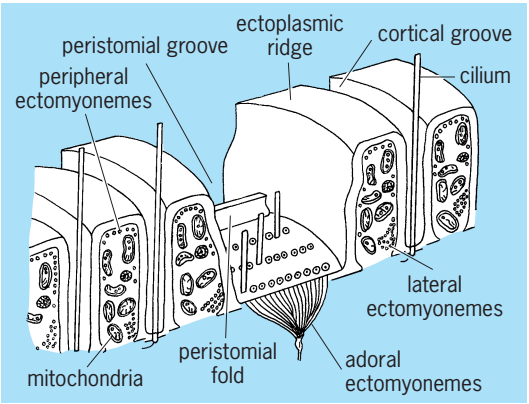

Fig. 6. Diagrammatic reconstruction of the cortex of *Spirostomum ambiguum*. (*After Finley, Brown, and Daniel, J. Protozool., 11:264, 1964*)

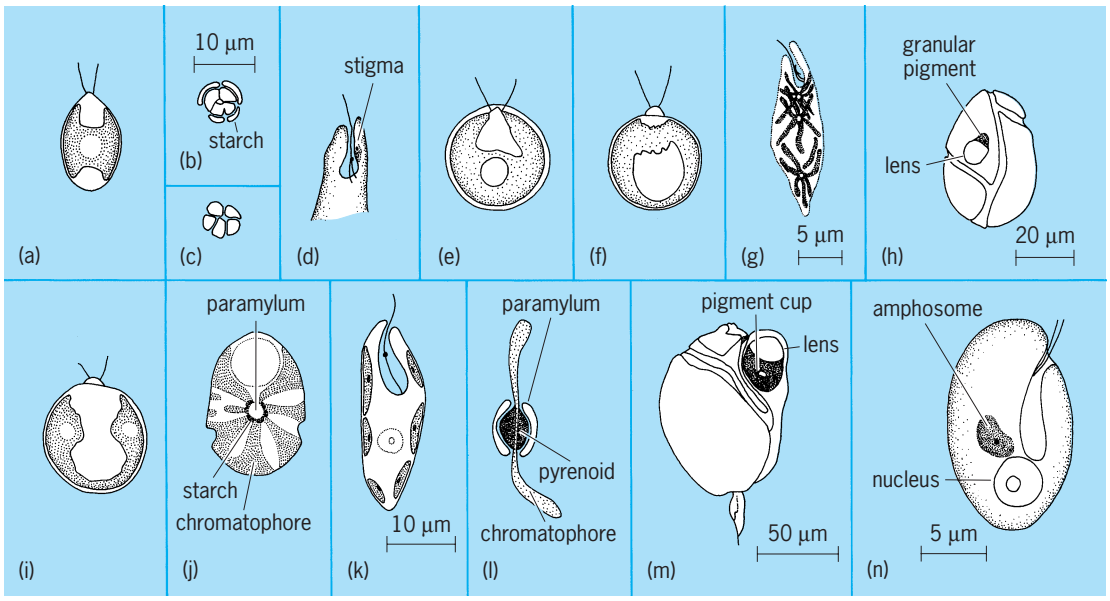

Fig. 7. Chromatophores, ocellus, amphosome, and pyrenoids. (*a*) Chromatophores in *Chlamydomonas agloeformis*. (*b*, *c*) Compound pyrenoids of *Pyramidomonas*, with adherent starch. (*d*) Stigma of *Euglena*. (*e*) Chromatophores in *Chlamydomonas umbonata*. (*f*) Chromatophores in *Chlamydomonas inversa*. (*g*) Chromatophores in *Euglena geniculata*. (*h*) Ocellus in *Protopsis*. (*i*) Chromatophores in *Chlamydomonas bicocca*. (*j*) Chromatophore in *Peridinium umbonatum*. (*k*) Chromatophores in *Colacium*. (*l*) Chromatophore having a pyrenoid and a paramylum. (*m*) Ocellus in *Erythropsis*. (*n*) Amphosome (stained) in *Cryptomonas*. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)

Food reserves. Polysaccharides are often stored as visible inclusions. These include glucose polymers: paramylum of Euglenida, leucosin of Chrysomonadida, starches of Cryptomonadida, Phytomonadida, and Dinoflagellida; and glycogens of Sarcodina, Zoomastigophorea, and ciliates. Protozoan glycogens resemble liver and muscle glycogen but differ in molecular weight and number of glucose residues in unit chains. Lipids may be stored as small to large globules which are characteristic of old rather than young laboratory cultures. A few flagellates, like certain diatoms, accumulate oils responsible for undesirable flavors and odors of drinking water. Nitrogenous reserves include chromatoid bodies, described in axenic cutures of trophozoites of *Entamoeba invadens* and cysts of certain other species. These inclusions arise before encystment by aggregation of small vesicles into bodies of characteristic form. Such inclusions aid in differentiating encysted *E. histolytica* and *E. coli*. Protein and RNA have been reported from chromatoid bodies, each of which contains small (about 20 nm) globules, interpreted by some workers as ribosomes. In electron micrographs (*E. invadens*) these inclusions show fibrils in a closely packed helical arrangement. Chromatoid bodies gradually disappear as the cyst matures. *See* AMOEBIDA.

Water expulsion vesicle. This organelle (**Figs. 10** and **11**), commonly termed a contractile vacuole, is characteristic of fresh-water protozoa and occurs also in certain marine and parasitic phagotrophs. Such vesicles show cyclic increase in volume, followed by collapse during which the contents are discharged. Motion pictures show no contraction of the "contractile" vacuole in *Amoeba proteus* (Fig. 11). Instead, cytoplasmic pressure increases with accumulating fluid until the vesicle is forced to the surface and its proximal surface is thrown into folds as the cortex ruptures to release the contents through a pore. The result is collapse of the vesicle. Electron micrographs of *Tetrahymena pyriformis* suggest a

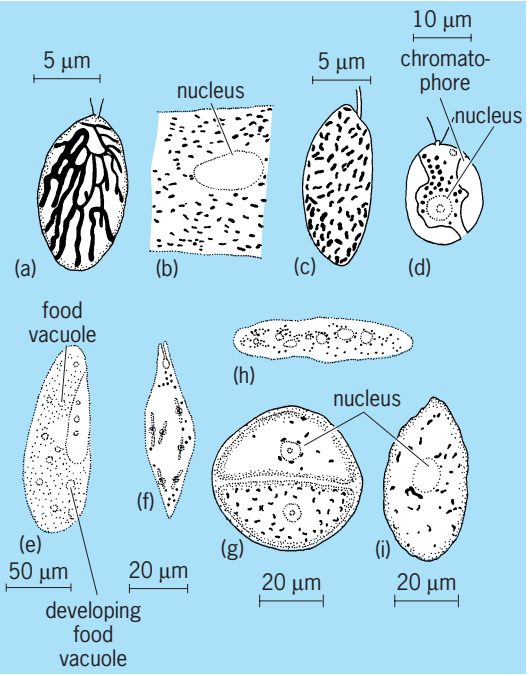

Fig. 8. Mitochondria, granules, and osmophilic inclusions. Mitochondria (*a*) in *Polytoma*; (*b*) in *Protoopalina*; and (*c*) in *Chilomonas*. Granules (*d*) in *Chlamydomonas*; (*e*) in *Paramecium*; and (*f*) in *Euglena*. Osmiophilic inclusions (*g*) in *Gregarina*, gametocytes in cyst; (*h*) in *Paramecium*; (*i*) in *Protoopalina*. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)

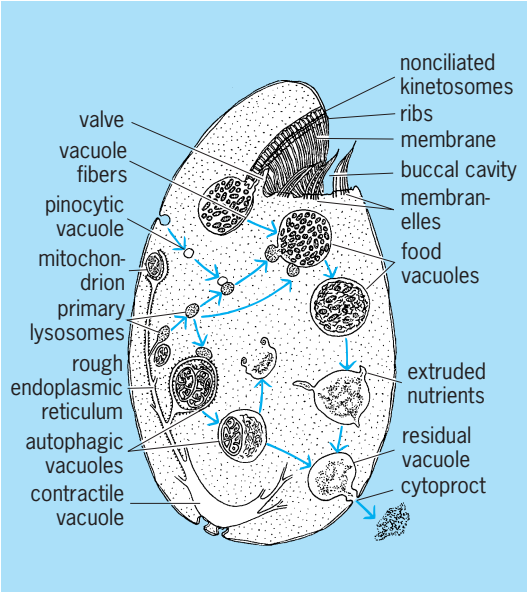

Fig. 9. Digestive pathways in *Tetrahymena pyriformis*. (*After Elliot, J. Protozool., 13:311, 1966*)

similar collapse of the vesicle as its contents are discharged through the pores. Accumulation of fluid in the enlarging vesicle is aided by collecting tubules (Fig. 10*f* and *g*) which may, as in *T. pyriformis*, extend through much of the cytoplasm as a "nephridial" system apparently continuous with the endoplasmic reticulum. The major function of the vesicle is elimination of excess water, which may be traced to ingestion (phagotrophs), endosmosis, or metabolic origin. Cyclic activity of the vesicle may be accelerated by adding ATP to the medium. Since there is no evidence that only pure water is secreted into the collecting tubules, it is possible that soluble materials also are eliminated by the vesicle.

Sensory vacuoles. These occur in certain parasitic and free-living ciliates (Fig. 10*c–e*). The similarity to statocysts of certain invertebrates inspired the assumption implied by the name.

Nuclei. Dimorphic nuclear equipment (micronucleus and macronucleus) is typical of ciliates. Analogous dimorphism is known in a few Foraminiferida but has not been observed in various others. Except for possibly similar cases yet to be investigated carefully, protozoa outside the ciliates contain one kind of nucleus, although the number ranges from one to many. The polyploid macronucleus of ciliates develops from a diploid derivative of the synkaryon (conjugation and autogamy), presumably by repeated replication of chromosomes in a macronuclear anlage. Filamentous "chromosomes" containing DNA have been described in macronuclei of *Blepharisma*. Later development involves increase in volume and sometimes elongation (*Euplotes* and *Vorticella*) or constriction into nodes (*Stentor* and some strains of *Blepharisma*). Micronuclear differentiation may involve a slight decrease in size. Both kinds of nuclei may contain about twice as much RNA as DNA, while the macronuclear content may be about 40 times that of the micronucleus (*Paramecium*).

In protozoa other than ciliates, nuclei vary in size and content of chromatin and nucleoli (or endosomes). Nuclear division is mitotic except in the ciliate macronucleus. It is not certain that the polyploid "somatic" nuclei of various Foraminiferida ever divide normally.

Nutrition

In protozoan feeding, either phagotrophic (holozoic) or saprozoic (osmotrophic) methods predominate in particular species. In addition, chlorophyll-bearing flagellates profit from photosynthesis; in fact, certain species have not been grown in darkness and may be obligate phototrophs.

Phagotrophic. Ingestion of food, followed by digestion in vacuoles, is characteristic of Sarcodina, ciliates, and many flagellates. Digestion follows synthesis of appropriate enzymes and their transportation to the food vacuole. Ingestion of nonnutritive latex particles also leads to acid-phosphatase activity of the "food" vacuoles. Primary lysosomes, which can fuse with the food vacuole, may transport hydrolases from the rough endoplasmic reticulum, as suggested for *Tetrahymena* (Fig. 9). Even a few

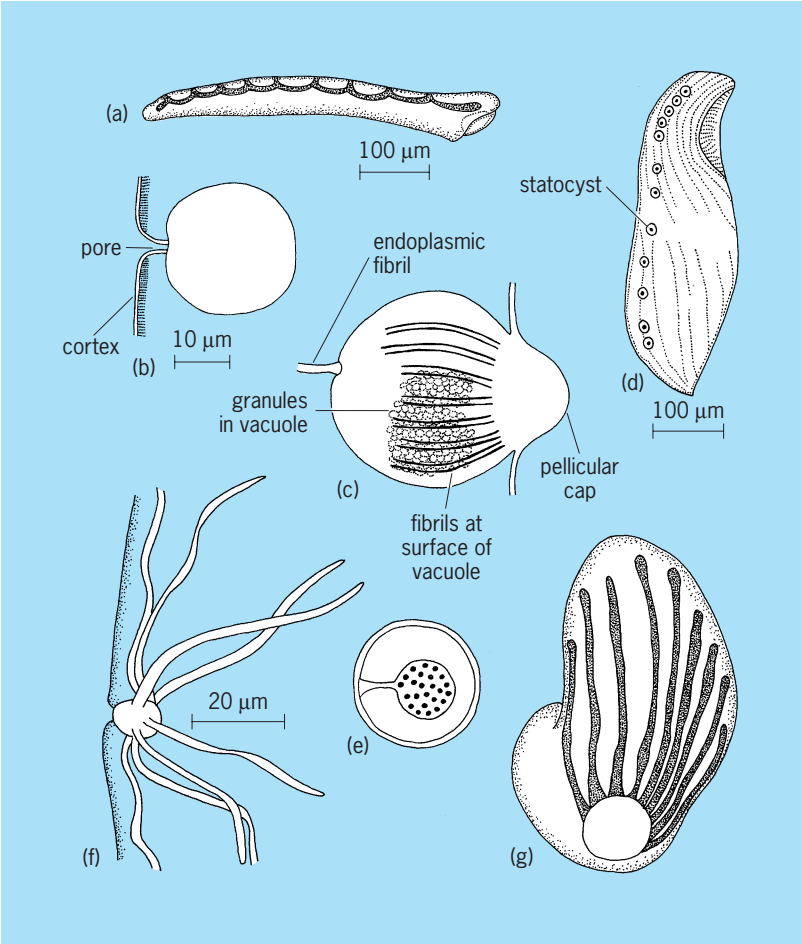

Fig. 10. Water expulsion vesicles and sensory vacuoles of protozoa. (*a*) Contractile tube in *Haptophrya*, an intestinal ciliate from a salamander. (*b*) Water expulsion vesicle and pore in *Eudiplodinium*, a parasitic ciliate from ruminants. (*c–e*) Sensory vacuoles of *Blepharoprosthium* and *Loxodes*. (*f*, *g*) Water expulsion vesicle and contributory canals in *Paramecium* and *Tillina*. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)

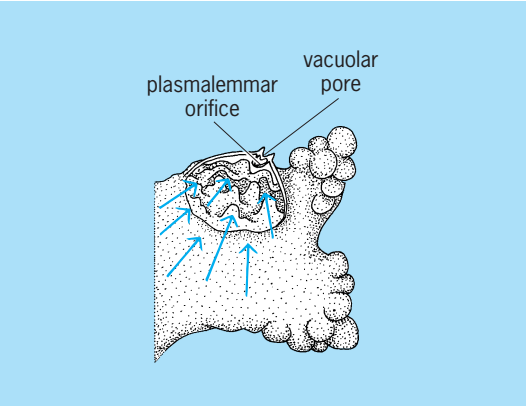

Fig. 11. Water elimination vesicle in *Amoeba proteus*, based on tracings of photomicrographs. Arrows indicate movement of granules. Endoplasm gets pushed against vesicle membrane, producing wrinkles. (*After Wigg, Bovee, and Jahn, J. Protozool., 14:104, 1967*)

Sporozoa, the malarial parasites and certain gregarines, are known to ingest host proteins or sometimes tissue cells or portions of them. Details of ingestion vary. Formation of food cups, or gulletlike invaginations, to enclose prey is common in more or less ameboid organisms, such as various Sarcodina, many flagellates, and at least a few Sporozoa. Entrapment in a sticky reticulopodial net occurs in Foraminiferida and certain other Sarcodina. A persistent cytostome and gullet are involved in phagotrophic ciliates and a few flagellates. A small cytostome and gullet in *Trypanosoma mega* suggest ability to ingest food, although importance of this activity in trypanosomes is still uncertain. Many ciliates (Fig. 9) have buccal organelles (membranes, membranelles, and closely set rows of cilia) arranged to drive particles to the cytostome. Particles pass through the cytostome into the cytopharynx (gullet), at the base of which food vacuoles (gastrioles) are formed. Digestion occurs in such vacuoles. An additional feature, reported in *Paramecium* and *Tetrahymena*, is entanglement of particles in mucus secreted by the peristomial cortex or buccal cavity. The effect is a clumping of particles into larger aggregates. In many less specialized ciliates the cytostome opens at the surface of the body. Such organisms (*Didinium* and *Perispira*) usually ingest large prey, such as other Protozoa. In some cases rodlike trichites surrounding the cytopharynx appear analogous to toxicysts of certain other ciliates.

Saprozoic. By definition this involves passage of dissolved foods through the cortex. It is uncertain to what extent diffusion is responsible, but enzymatic activities presumably are involved in uptake of various simple sugars, acetate and butyrate in *Polytoma* and even potassium and phosphate ions. In addition, external factors, for example, the pH of the medium, may strongly influence uptake of fatty acids and phosphate.

Pinocytosis. This process resembles phagocytosis in miniature but with ingestion apparently limited to materials in solution. First recognized in *Amoeba proteus,* the process has since been described in other amebas, certain ciliates, and *Eimeria* (Sporozoa). Pinocytosis is stimulated by "inducers," such as proteins, certain salts, and toluidine blue. In the process small vacuoles are formed at the base of a tubular invagination from a surface, either cortex or the wall of cytopharynx. The vacuoles drift into the endoplasm and mingle with other inclusions. In certain ciliates, a similar process yields secondary vacuoles from a food vacuole. In certain malarial parasites, but not in others, such pinocytosis precedes completion of digestion. *See* ENDOCYTOSIS.

Nutritional requirements. Establishment of a strain in axenic cultures in defined media make possible a determination of specific food requirements. In nature, these requirements are met by the available diet which includes other microorganisms and materials dissolved in the external medium. *See* AXENIC CULTURE.

Mineral requirements. Certain materials, including calcium, silicon, and strontium salts, are needed for tests and other skeletal structures, and metal ions are needed in traces for enzyme systems. Trace requirements of one species or another include calcium, cobalt, copper, iron, potassium, magnesium, manganese, phosphorus, sulfur, and zinc. Although indispensable status is uncertain, aluminum, barium, boron, iodine, silicon, sodium, and vanadium have each stimulated growth of one or more species.

Nitrogenous foods. Nitrogen requirements of many phytoflagellates are met by an ammonium salt or by either a nitrate or an ammonium salt. Other protozoa need specific amino acids which they cannot synthesize, the number ranging from 1 (*Perenema*) to 7 for an ameba (*Hartmannella rhysodes*), 13 for *Paramecium multimicronucleatum,* and at least 15 for *Leishmania tarentolae*. In a few cases, radioactive amino acids have been used in tracing their fate in metabolism. In analogous fashion, labeled food organisms have been used to identify species ingested by Foraminiferida in laboratory aquariums. In addition to amino acids, certain protozoa need a purine, or sometimes both a purine and a pyrimidine. *See* AMINO ACIDS; PURINE; PYRIMIDINE.

Carbon requirements. The only carbon source needed by certain chlorophyll-bearing flagellates is CO_2. In fact, *Gymnodinium breve*, a dinoflagellate occurring in "red tides," apparently cannot use any organic compound which has been tested and seems to be an obligate phototroph. *Euglena gracilis*, in contrast, can grow in darkness in a good medium. Requirements of heterotrophs, as a group, may be met by a polysaccharide, by a fatty acid, by an acid of the tricarboxylic acid (TCA) cycle, or by an alcohol. Suitability varies with the species.

Growth factors. Organic substances, such as vitamins, are needed in traces. A few phytoflagellates, namely, certain species of *Brachiomonas, Chlamydomonas,* and *Chlorogonium*, are known to synthesize all required vitamins. Other protozoa need one or more vitamins and sometimes other growth factors. Loss of an ability to synthesize a particular growth factor (niacin or thioctic acid) has occurred in *Tetrahymena pyriformis* after prolonged

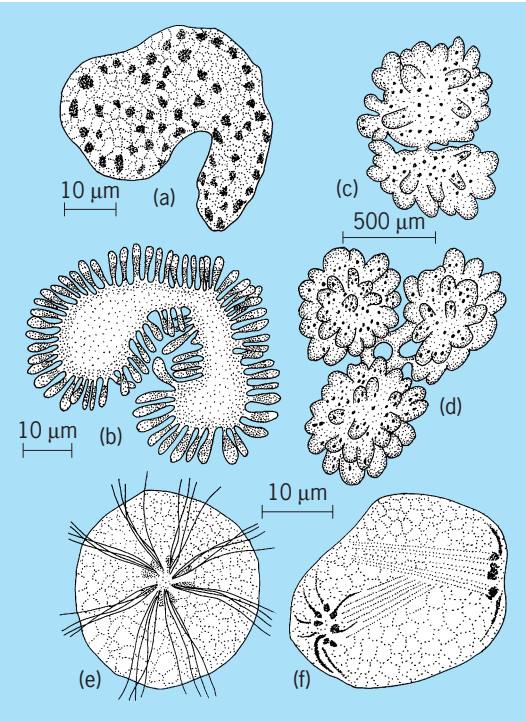

Fig. 12. Schizogony and plasmotomy. (*a*) Schizogony in *Ovivora thalassemae*. (*b*) Schizont of *O. thalassemae*. (*c*, *d*) Plasmotomy in *Pelomyxa carolinensis*. (*e*) *Coronympha octonaria*, vegetative stage showing nuclei and flagellar groups. (*f*) Nuclear groups at end of telophase, as they appear just before plasmotomy in *C. octonaria*. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)

maintenance of strains in laboratory cultures. Certain amebas (*Acanthamoeba* and *Hartmannella*) need only three vitamins (B_{12}, biotin, and thiamine); similar requirements are known for certain phytoflagellates. A ciliate, *Glaucoma chattoni,* needs biotin, folic acid, nicotinic acid, pantothenic acid, pyridoxine, riboflavin, thiamine, thioctic acid, plus certain lipids. Less common requirements include ethanol or methanol for *Tetrahymena setifera,* chlorine for certain European strains of *T. pyriformis*, and a pteridine for *Crithidia fasciculata*. Identification of vitamin requirements was followed by use of certain protozoa in biological assays—vitamin B_{12} (*Ochromonas malhamensis*, *Euglena gracilis*), thiamine (*O. malhamensis*), pyridoxine (*Tetrahymena pyriformis*), biotin (*Ochromonas danica*), nicotinic acid (*T. pyriformis*), and pantothenic acid (*T. pyriformis*). *See* VITAMIN.

Reproduction

Reproduction occurs after a period of growth which ranges, in different species, from less than half a day to several months (certain Foraminiferida). General methods include binary fission, budding, plasmotomy, and schizogony. Fission, involving nuclear division and replication of organelles, yields two organisms similar in size. Budding produces two organisms, one smaller than the other. In internal budding (certain Suctoria) a bud is formed within a brood pouch, an almost closed invagination in the "parental" body. The bud, after it develops cilia, is expelled through a "birth pore" as a sort of larva which, after a few hours of activity, undergoes metamorphosis (losing cilia, developing tentacles) into a typical suctorian. In external budding, the bud is formed at the surface of the "parental" body. In plasmotomy (**Fig. 12***c* and *d*), a multinucleate organism divides into several, each containing a number of nuclei. Schizogony, characteristic of Sporozoa, follows repeated nuclear division, yielding many uninucleate buds (Fig. 12*a* and *b*). In some cases, as in *Eimeria bovis,* the schizont first divides into several multinucleate forms, in each of which the nuclei approach the surface to take part in schizogony (merogony).

Reproduction includes formation of new organelles, either by division of old structures or by neoformation, sometimes preceded by resorption of an old organelle. Blepharoplasts and basal granules, in particular, seem to be self-replicating. Formation of a new mouth (stomatogenesis) in ciliates often involves origin of an anlage from basal granules of one or more ciliary rows and its differentiation into new buccal organelles. Even the formation of a new pellicle in Euglenida may involve formation of new pellicular strips between adjacent strips of the original pellicle.

Life cycles. Simple life cycles include a cyst and an active (trophic) stage undergoing growth and reproduction. In certain free-living and parasitic species, no cyst is developed. Dimorphic cycles show two active stages; polymorphic show several. The former include adult and larva (Suctoria); flagellates and ameba (certain Mastigophora and Sarcodina); flagellate and palmella (nonflagellated; certain Phytomonadida); and ameba and plasmodium (Mycetozoia especially).

Encystment. This process involves resorption of locomotor and feeding organelles, loss of water, often rounding up of the body, and secretion of one or more cyst membranes. The wall usually is composed

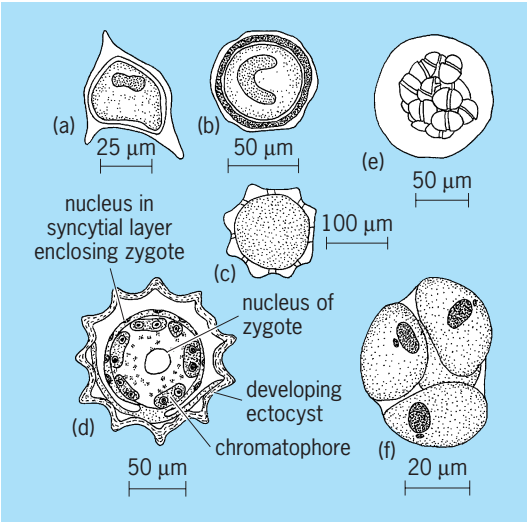

Fig. 13. Protective and reproductive cysts. (*a*) Protective cyst of *Ceratium*; (*b*) of *Didinium*; (*c*) of *Bursaria*. (*d*) Encysted zygote of *Volvox*. (*e*) Reproductive cyst of *Gyrodinium*, a dinoflagellate and (*f*) of *Colpoda*, a ciliate. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)

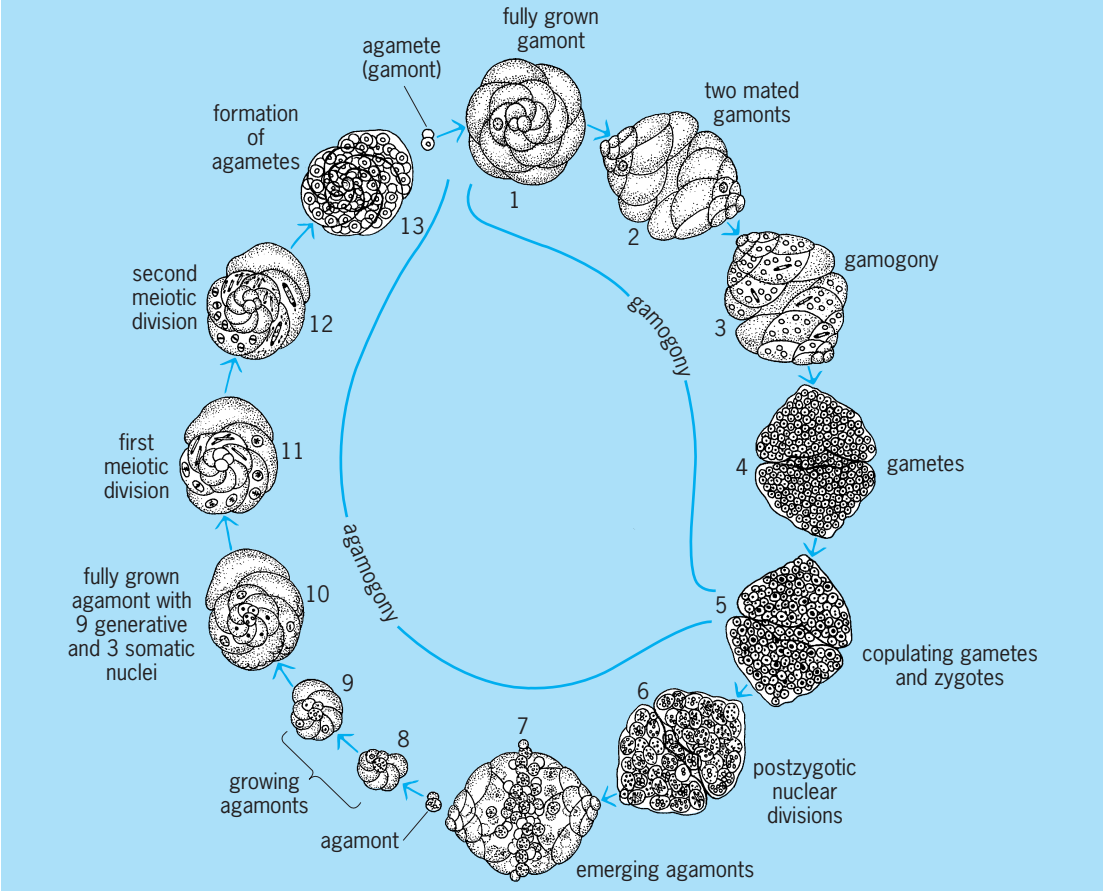

Fig. 14. Life cycle of *Glabratella sulcata*, showing alternation of haploid phase (gamogony) and diploid phase (agamogony). Two meiotic divisions (11 and 12) precede the production of young haploid organisms (13). (*After Grell*, *Research in Protozoology*, *vol. 2*, *Pergamon Press*, *1968*)

of secreted materials (proteins, carbohydrates, keratin, and others), although foreign particles may be cemented into a wall. Precystic activities may include storage of reserves. Protective cysts (**Fig. 13***a–d*), with fairly thick walls, often resist desiccation. Dried cysts of *Colpoda cucullus* have remained viable for 5 years; those of certain amebas and flagellates, for 49 years in dry soil samples. Within reproductive cysts (Fig. 13*e* and *f*) fission, budding or schizogony, sometimes gametogenesis, and syngamy occur in various species. The usually thin-walled reproductive cysts afford relatively little protection.

Excystment involves absorption of water leading to rupture of cyst membranes and, when necessary, development of organelles for feeding and locomotion.

Sexual activities. These involve production of haploid (gametic) nuclei and their fusion into a synkaryon (zygotic nucleus). Meiosis occurs at different stages in life cycles. In haploid organisms (Phytomonadida, certain flagellates of termites and wood roaches, Coccidia, malarial parasites, and gregarines) the zygote is the only diploid stage. Meiosis occurs as the zygotic nucleus divides. In these haploid cycles a heterozygous condition can exist only in the zygote. Diploid cycles, in which meiosis occurs during gametogenesis, are found in ciliates, Heliozoia (*Actinophryx* and *Actinosphaerium*), and in certain Hypermastigida. In several flagellates of the wood roach, sexual activity is stimulated by ecdysone (molting hormone) of the host. A few Foraminiferida show an alternation of generations involving a haploid gamete-producing phase and a diploid phase with ordinary reproduction (**Fig. 14**). *See* ECDYSONE.

Sexual activities include syngamy, autogamy, and conjugation. Syngamy involves fusion of two gametes which may be morphologically similar (isogamy) or different (anisogamy). Autogamy occurs within a single organism; haploid nuclei are formed and subsequently fuse into a synkaryon. Conjugation involves pairing of ciliates (**Fig. 15**), apparently following appearance of specific protein-containing substances at the surface, including cilia. Pairing in *Tetrahymena* includes fusion of the two pellicles and formation of channels joining the conjugants. In tests with radioactive foods, cytoplasmic materials (RNA and proteins) passed from one conjugant to the other. Pairing stimulates nuclear activities. Old macronuclei gradually degenerate as a rule, although fragments may be exchanged occasionally in *Stylonychia*. The micronucleus undergoes three divisions, the second being reductional. The third produces gametic nuclei which, in *Tetrahymena,* arise from one haploid nucleus. A migratory gametic nucleus, whose status is apparently determined by

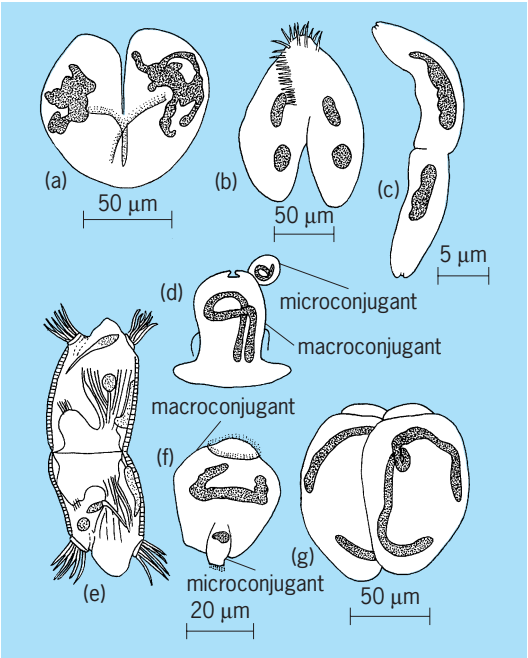

Fig. 15. Early conjugation, just after fusion of conjugants. (*a*) *Nyctotherus*. (*b*) *Pleurotricha*. (*c*) *Ancistrocoma*. (*d*) *Cycloposthium*. (*e*) *Scyphidia*. (*f*) *Vorticella*. (*g*) *Euplotes*. (*After R. P. Hall, Protozoology, Prentice-Hall, 1953*)

position, passes from each conjugant to its mate, where fusion with the stationary nucleus occurs. The resulting synkaryon, in different species, divides one or more times and the products become macro- and micronuclei. Reorganization may involve one or more fissions which restore nuclear numbers to normal. In this sense, a pair of conjugants produces two or more ciliates. In the peritrichs, with micro- and macroconjugants, only the latter survives.

Autogamy occurs in certain ciliates and a few parasitic flagellates, *Barbulanympha,* for example. Since the gametic nuclei arise from one nucleus, the resulting synkaryon is homozygous.

Normal conjugation involves ciliates of complementary mating types, two or more belonging to a syngen, originally termed "variety." Each species contains a number of syngens. Typically, mating types within one syngen can conjugate normally. Abortive conjugation, sometimes fatal, follows pairing of ciliates from different syngens.

Syngamy in Phytomonadida typically requires two kinds of gametes (+ and −). Syngens analogous to those of ciliates have been reported in certain Phytomonadida. In *Astrephomene* at least, different syngens may have different chromosome counts. In dioecious species (*Chlorogonium* and some species of *Chlamydomonas*) the two gametes come from different clones; each zygote produces two clones of each "sex." In monoecious types (several species of *Chlamydomonas* and *Haematococcus pluvialis*) a single clone produces both types of gametes. In colonial phytomonads gametes may be isogamous (*Gonium* and *Pandorina*) or anisogamous (*Eudorina* and *Volvox*). Differentiation of male colonies in *Volvox aureus* is induced by a rather thermostable protein-containing substance. In less specialized genera (*Gonium*) all zooids become gametes; at the other extreme (*Volvox* and *Eudorina*) gametes arise in the posterior half of the colony.

According to the concept of a "physiological life cycle," laboratory strains of ciliates pass through phases of youth; maturity, when conjugation is possible; senesence, when conjugation is impossible; and death. Senescence supposedly is averted by opportune occurrence of conjugation. Although certain species apparently can live indefinitely without conjugation, it is possible that autogamy replaces conjugation in rejuvenating such ciliates.

Parasitic Protozoa

Parasites occur in all major groups. Sporozoa are exclusively parasitic, as are some flagellate orders (Trichomonadida, Hyperamastigida, and Oxymonadida), the Opalinata, Piroplasmea, and several ciliate orders (Apostomatida, Astomatida, and Entodiniomorphida). Various other groups contain both parasitic and free-living types. Protozoa also serve as hosts of other Protozoa, certain bacteria, fungi, and algae. An unusual addition is *Leptomonas karyophilus,* a trypanosomid flagellate which invades the macronucleus of *Paramecium trichium*. Such symbionts as *Chlorella paramecii* in *Paramecium bursaria* are considered symbiotes. A similar example is a bacterial symbiont of *Crithidia oncopelti,* which apparently contributes products to nitrogen metabolism of the host.

Many protozoan symbionts, called commensals, are more or less harmless. Cellulose-digesting flagellates of termites and ciliates of ruminants are symbiotic in the sense that they benefit their hosts. In addition to helpful digestive activities, rumen ciliates are potential food for the herbivores.

Pathogens. Relatively few parasites are distinctly pathogenic, causing amebiasis, visceral leishmaniasis (kala azar), African sleeping sickness, Chagas' disease, malaria, tick fever of cattle, dourine of horses, and other diseases. *See* MALARIA.

Harmful effects may be produced in different ways. Tissue cells are invaded and destroyed in Chagas' disease, leishmaniasis, coccidiosis, and malaria. Tissues may be destroyed, as in abscesses and ulcers involving *Entamoeba histolytica*. Although a potent "toxotoxin" is reported for *Toxoplasma gondii* and a Chagas toxin from *Trypanosoma cruzi,* little is yet known about possible toxins of protozoan parasites.

Epidemiology. Transfer may involve direct migration of parasites living in aquatic hosts; direct contact (osculatory, veneral); ingestion with contaminated food or drink; and vector transfer, as in malaria or Chagas' disease. Variations include placental transfer from mother to fetus, known in malaria and Chagas' disease; ovarian transfer, to eggs in ovary, as in ticks infected with *Babesia bigemina;* and lacteal transfer, from female to suckling young, reported in laboratory trypanosomiasis. *See* MEDICAL PARASITOLOGY.

Richard P. Hall

Bibliography. A. C. Chandler and C. P. Read, *Introduction to Parasitology: With Special References to*

the Parasites of Man, 10th ed., 1961; J. O. Corliss, *The Ciliated Protozoa: Characterization, Classification, and Guide to the Literature*, 2d ed., 1979; J. Kreier, *Parasitic Protozoa*, 2d ed., 1997; R. R. Kudo, *Protozoology*, 5th ed., 1977; *Progress in Protistology*, 2 vols., 1986–1987; M. Sleigh, *Protozoa and Other Protists*, 3d ed., 2002; J. J. Wolken, *Euglena: An Experimental Organism for Biochemical and Biophysical Studies*, 1967.

Protura

An order of ancestrally wingless insects (subclass Apterygota). The order is usually classified into four families in two suborders (Eosentomoidea and Acerentomoidea), and about 500 species are recognized.

The insects are under 0.08 in (2 mm) in length, having an elongated body with a small head, relatively large abdomen, and three pairs of functional legs. The first pair of legs is normally held out in front of the head, antennalike. They serve to replace the sensory role of true antennae, which, like eyes, are absent. Paired pseudoculi (probably chemosensory) occupy eyelike positions on the head. The mouthparts include an elongate and styliform mandible, a maxilla, and a two-part labium. All parts are enclosed within a pocket in the head, so that only their tips are exposed. The thorax bears five-segmented legs, each with an apical claw. In the adult, the abdomen has 12 segments; the anus is terminal, and cerci (segmented sensory appendages on the last abdominal segment) are lacking. The sterna of the first through third segments bear short styli, probably representing vestigial limbs. A large gland opens at the rear of the eighth dorsal plate, which sometimes has a comblike lid. The genitalia are enclosed within a ventral pouch between the eleventh and twelfth segments, and have paired protrusible stylets (slender, elongated appendages).

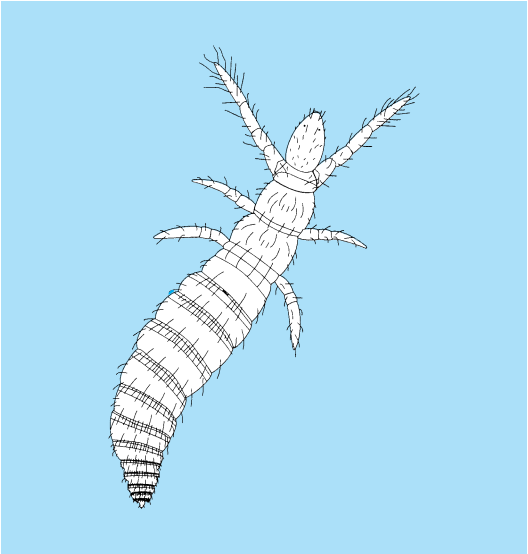

A proturan, *Acerentulus barberi* (dorsal view).

Protura live in forest soil, leaf litter, and similar places, and are thought to feed on mycorrhizal fungi. They are distributed globally but unevenly in temperate and tropical climates, and they sometimes occur in enormous numbers, as in some Oregon fir forests in the United States. *See* INSECTA.

William L. Brown, Jr.

Bibliography. G. Imadate, *The Insects of Australia*, 2d ed., 1991; S. L. Tuxen, *Protura: A Revision of the Species of the World,* 1964.

Proustite

A mineral having composition Ag_3AsS_3 and crystallizing in the hexagonal system. It occurs in prismatic crystals terminated by steep ditrigonal pyramids, but is more commonly massive or in disseminated grains. There is good rhombohedral cleavage. Hardness is 2–2.5 (Mohs scale) and specific gravity is 5.55. The luster is adamantine and the color ruby red. It is called light ruby silver in contrast to pyrargyrite, dark ruby silver. Proustite is less common than pyrargyrite but the two minerals are found together in silver veins. Noted localities are at Chañarcillo, Chile; Freiberg, Germany; Guanajuato, Mexico; and Cobalt, Ontario, Canada. *See* PYRARGYRITE; SILVER.

Cornelius S. Hurlbut, Jr

Provenance (geology)

In sedimentary geology, all characteristics of the source area from which clastic sediments and sedimentary rocks are derived, including relief, weathering, and source rocks. *See* WEATHERING PROCESSES.

The goal of most provenance studies of sedimentary rocks is the determination of source characteristics of the mountains or hills from which the constituent sediment was derived. Such determinations are difficult to make because sediment composition and texture are continually modified during erosion, transport, deposition, and diagenesis (postdepositional modification). These modifications must be accounted for in order to determine source characteristics (provenance). It is most straightforward to determine provenance in situations in which these modifying effects are minimal; provenance may be indeterminate or ambiguous in situations involving extensive modification of sediment composition and texture. The former situation is most common in tectonically active areas, resulting in rapid uplift and erosion of mountains, rapid transport and deposition, and slight diagenetic modification after deposition. In contrast, stable continental areas (for example, cratons) provide ample opportunity for intense weathering so that chemical, mineralogical, and textural characteristics of sediment are intensely modified, thus rendering determination of source characteristics difficult. *See* DIAGENESIS; EROSION; SEDIMENTARY ROCKS.

Clastic (or detrital) sediment is commonly recycled during multiple episodes of mountain building,

erosion, sedimentation, lithification, and renewed mountain building. This process constitutes the rock cycle, within which igneous, sedimentary, and metamorphic rocks are created and modified. During this process, it is common for older sedimentary rocks to be uplifted and eroded, so that individual sedimentary particles (clasts) are recycled to form new sediment, which may be lithified to form new sedimentary rock. Provenance studies must determine the proportion of a sedimentary rock derived directly from indicated source rocks versus the proportion derived directly from another sedimentary rock (that is, rocks exposed during previous cycles of sedimentation). This determination is essential, but commonly difficult to accomplish because the multicyclic nature of sediment may be difficult to recognize. *See* METAMORPHIC ROCKS.

Fundamentally different methods of study are utilized, depending on what aspect of provenance is emphasized, what type of sediment is studied, and what scale of sampling and study is attempted. Grain size of detrital sediment is a dominant control over what methods may be employed.

Coarse-grained sediment (gravel and conglomerate) provides the most direct determination of source rocks because individual clasts may be studied by any method available for the study of source rocks themselves. Statistical counts of clasts may be compared directly to proportions of rock types in potential source areas. However, gravel and conglomerate is unavailable for study in many locations, and many rock types do not survive weathering or transport in this grain size. *See* CONGLOMERATE; GRAVEL.

Intermediate grain sizes (sand and sandstone) are the most commonly studied because they are abundant and preserve many source-rock characteristics. Statistical point-counting of thin sections of sand and sandstone mounted on glass slides using polarizing microscopes is the most common method of study. Details of point-counting methods must be carefully chosen in order to maximize the usefulness of the data for different types of studies. *See* SAND; SANDSTONE.

Fine-grained sediment (mud and mudrock) owes much of its history to weathering and alteration of sediment during transport, deposition, and burial. Therefore, composition of fine clastic particles (primarily clay minerals) reflects these factors more than it reflects provenance. For this reason, provenance is more difficult to decipher based on fine-grained rocks, although whole-rock chemical characteristics may persist through various stages of sediment modification.

The first-order control on rock composition at Earth's surface is plate tectonics. Mountain building (orogeny) and rifting (taphrogeny) are the direct results of plate tectonic processes; most studies of provenance, therefore, account for modern or ancient plate tectonic processes. Provenance studies must also account for scale of study; scale of sampling directly reflects the scale of variation of rock composition in source areas. For example, local drainages produce sediment derived from local rock types, with possible great variability, whereas continental-scale sedimentary systems (e.g., big rivers and turbidite fans) produce homogeneous sediment reflecting plate-tectonic setting. *See* OROGENY; PLATE TECTONICS.

Methods of determining provenance include direct determination of rock types (primarily used for coarse to medium grains); direct determination of mineralogy (used for all grain sizes); whole-rock geochemistry (used for medium to fine grains); geochemistry of individual mineral species (used for all grain sizes, but especially for medium grains); and radiometric dating of individual mineral species (primarily for medium grains). *See* GEOCHEMISTRY; MINERALOGY; ROCK AGE DETERMINATION.

A guiding principle in provenance studies is actualism (the concept that the present is the key to the past). Studies of modern environments of uplift, erosion, transport, deposition, and diagenesis provide key insights into how sedimentary systems work. Fundamental controls on composition of gravel, sand, and mud can be determined from the study of modern systems, and compositional models may be developed based on these controls. The models in turn may be applied to the study of ancient conglomerate, sandstone, and mudrock; significant progress in the reconstruction of paleogeography and paleotectonics inevitably results. *See* PALEOGEOGRAPHY.

Raymond V. Ingersoll

Bibliography. W. R. Dickinson, Provenance and sediment dispersal in relation to paleotectonics and paleogeography of sedimentary basins, in K. L. Kleinspehn and C. Paola (eds.), *New Perspectives in Basin Analysis*, pp. 3–25, Springer-Verlag, New York, 1988; H. Ibbeken and R. Schleyer, *Source and Sediment: A Case Study of Provenance and Mass Balance at an Active Plate Margin* (*Calabria, Southern Italy*), Springer-Verlag, New York, 1991; R. V. Ingersoll, Actualistic sandstone petrofacies: Discriminating modern and ancient source rocks, *Geology*, 18:733–736, 1990; R. V. Ingersoll et al., The effect of grain size on detrital modes: A test of the Gazzi-Dickinson point-counting method, *J. Sed. Petrol.*, 54:103–116, 1984 (see also Discussions and Replies, 55:616–621); R. V. Ingersoll, A. G. Kretchmer, and P. K. Valles, The effect of sampling scale on actualistic sandstone petrofacies, *Sedimentology*, 40:937–953, 1993; M. J. Johnsson and A. Basu (eds.), Processes controlling the composition of clastic sediments, *Geol. Soc. Amer. Spec. Pap.*, no. 284, 1993; L. J. Suttner, Sedimentary petrographic provinces: An evaluation, *Soc. Econ. Paleontol. Mineralog. Spec. Publ.*, no. 21, pp. 75–84, 1974.

Prussian blue

The common name for a material prepared by either the addition of trivalent iron (Fe^{III}) to ferrocyanide $[Fe(CN)_6]^{4-}$ or Fe^{II} to ferricyanide $[Fe(CN)_6]^{3-}$ with the composition of $Fe^{III}{}_4[Fe^{II}(CN)_6]_3 \cdot xH_2O$, where $x \sim 15$. Prussian blue, also known as Turnbull's blue or ferric hexacyanoferrate, was discovered

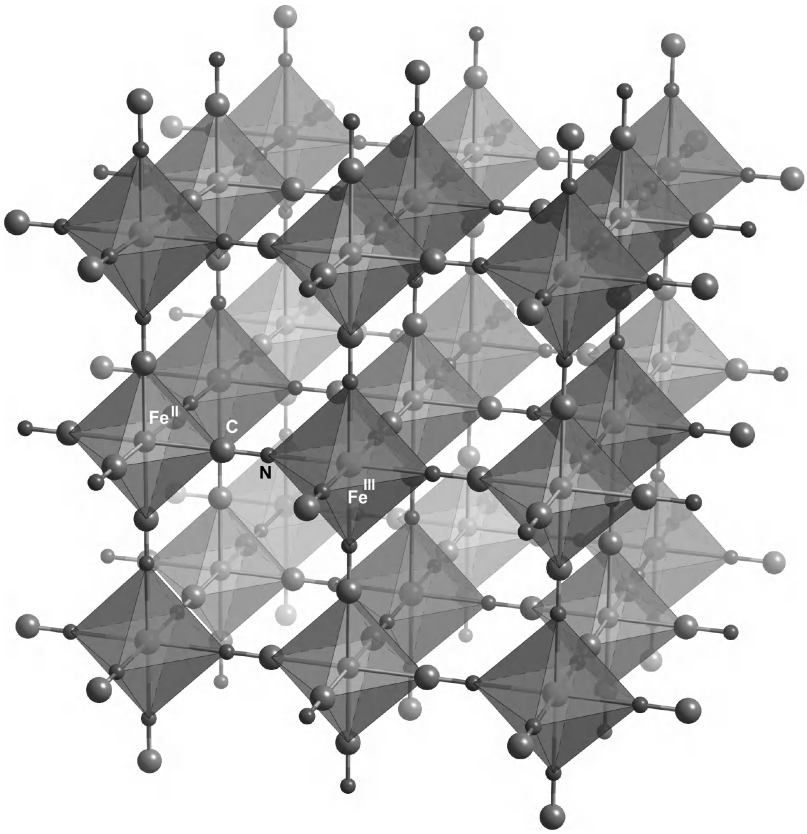

Diagram of cubic structure of Prussian blue.

accidently 300 years ago by Diesbach of Berlin, and its structure has been determined by several diffraction methods. Prussian blue has a structure with each Fe^{II} bound to six C-bounded cyanides (CN^-), while the nitrogens (N) are bonded to Fe^{III}; thus, it has a cubic structure with $Fe^{II}CNFe^{III}$ linkages along each of the edges of the cube (see **illustration**). Owing to the presence of more Fe^{III} than Fe^{II}, some Fe^{II} sites are vacant and filled with water. The presence of iron in two different oxidation or valence states in Prussian blue makes it a mixed-valence compound. This gives Prussian blue its intense blue color and its application as a pigment. The reduction of Prussian blue results in the formation of Everitt's salt, $K_2Fe^{II}[Fe^{II}(CN_6)]$.

Uses of Prussian blue include artists' colors, carbon paper, typewriter ribbons, blueprinting, and printing inks. It is prescribed for the treatment of radioactive $^{137}Cs^+$ and thallium poisoning. Substitution of the Fe^{II} and Fe^{III} ions with other metal ions leads to Prussian blue–structured materials with different metal ions. These materials typically are magnets. Substitution of the Fe^{II} and Fe^{III} ions with Cr^{III} and V^{II} ions, respectively, results in a magnet with an ordering temperature around 100°C (212°F). *See* COORDINATION CHEMISTRY; COORDINATION COMPLEXES; CURIE TEMPERATURE; CYANIDE; FERRICYANIDE AND FERROCYANIDE; IRON. Joel S. Miller

Bibliography. K. R. Dunbar and R. A. Heintz, Chemistry of transition metal cyanide compounds: Modern perspectives, *Prog. Inorg. Chem.*, 45:283–392, 1997; J. S. Miller, 3-D-network structured cyanide-based magnets, *MRS Bull.*, 25(11):60–64, 2000; A. G. Sharp, *The Chemistry of Cyano Complexes of the Transition Metals*, Academic Press, New York, 1976.

Prymnesiophyceae

A class of algae (also known as Haptophyceae) in the chlorophyll *a-c* phyletic line (Chromophycota). In protozoological classification these organisms constitute an order, Prymnesiida or Haptomonadida, in the class Phytamastigophora. Most of the approximately 300 species of prymnesiophytes are biflagellate monads.

This class has been segregated from the Chrysophyceae, with which it shares many biochemical and ultrastructural characters. Prymnesiophytes differ from chrysophytes, however, in several significant characters: (1) the typical monad bears a filiform organelle, the haptonema, between the two flagella; (2) except in the order Pavlovales, the flagella are of equal length and smooth; (3) organic scales, which may be calcified, cover most motile cells. (By contrast, motile cells of Chrysophyceae lack haptonemata, have one smooth and one pleuronematic flagellum, and have siliceous scales, if any.) Depending on the species, the haptonema is longer than the flagella and slender, short and bulbous, or vestigial. It consists of a bundle of longitudinal microtubules surrounded by a fenestrated sheath of endoplasmic reticulum, the whole being enclosed in the plasmalemma. At its base, which may be interconnected with the basal bodies of the flagella, the microtubular bundle consists of nine microtubules in a closely packed hexagonal arrangement. The emergent structure has the form of a shallow trough (an arc in cross section), enclosing seven microtubules. No function for haptonemata is known. Although they are capable of bending or flickering and, if long, can coil rapidly and uncoil slowly in response to shocks, haptonemata apparently play no role in either locomotion or capture of prey.

Calcified scales (cellulosic scales impregnated with calcite) are characteristic of many prymnesiophytes. These scales, which are given the general term coccoliths, were discovered in marine sediment before they were observed on living cells. Coccoliths are classified into several morphological types (such as rhabdoliths, discoliths, zygoliths, and ceratoliths), which are of great diagnostic value in the taxonomy of the prymnesiophytes that bear them (coccolithophorids).

Cells of prymnesiophytes contain one or two chloroplasts, with ultrastructure and pigmentation similar to that in Chrysophyceae. A pyrenoid is usually present, and an eyespot usually absent. The single Golgi apparatus with characteristically distended central cisternae is the site of scale production. The metabolic storage product is chrysolaminaran. All species are photosynthetic, although some are facultatively osmotrophic or phagotrophic.

Members of the order Pavlovales, which resemble other prymnesiophytes in having a haptonema, differ

by lacking platelike scales, by having eyespots, and by having unequal flagella that are covered with fine hairs.

In most prymnesiophytes a nonmotile phase alternates with a motile phase. The nonmotile phase is a free-living unicell or a palmelloid or pseudofilamentous colony. In some cases the alternation is mediated by sexual reproduction. Usually, however, reproduction is effected by binary fission or the production of zoospores.

Prymnesiophytes are primarily marine, with coccolithophorids constituting one of the three major components of phytoplankton (the others being diatoms and dinoflagellates). Their importance in past geologic periods is documented by the vast amounts of coccolith deposits (for example, the white cliffs of Dover) dating back at least to the Jurassic (140 million years ago). More than 350 genera of fossil coccolithophorids have been described, so that the living representatives must be considered relict. Some prymnesiophytes are restricted to brackish or fresh water or are euryhaline. A few species are grown as food sources for invertebrates in mariculture projects, while other (*Prymnesium parvum*) develop explosively (bloom) as weeds in brackish-water fish ponds and poison the fishes. *Phaeocystis pouchetii*, which is characteristic of North Atlantic and Antarctic waters, can bloom so densely in the open sea that fishing nets are clogged by the gelatinous colonies and patterns of herring migration are affected. *See* ALGAE; COCCOLITHOPHORIDA; PHYTOPLANKTON; PROTOZOA. Paul C. Silva; Richard L. Moe

Bibliography. E. R. Cox (ed.), *Phytoflagellates*, pp. 273–317, 1980; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Pseudoborniales

An order of fossil plants found in Middle and Upper Devonian rocks. The group is related to Sphenophyllales and includes a single family and two monotypic genera. *Pseudobornia ursina* is known from Bear Island (north of Norway), Alaska, and Germany. *Prosseria grandis* is found in New York State. Sphenopsid characters are more firmly established in this order than in Hyeniales.

Pseudobornia (see **illus.**), the better known of the two, has rhizomes and stems up to 2.4 in. (6 cm) wide and over 3 ft (1 m) long. The axes are jointed. Larger axes bear two branches at a node and smaller ones have whorls of four leaves. The leaves are 1.4–2.4 in. (3.5–6 cm) long, short-stalked, and palmately divided, and they have laciniate margins. Fertile shoots bear reduced leaves with sporangia on their lower sides. *See* HYENIALES; PALEOBOTANY; SPHENOPHYLLALES. Harlan P. Banks

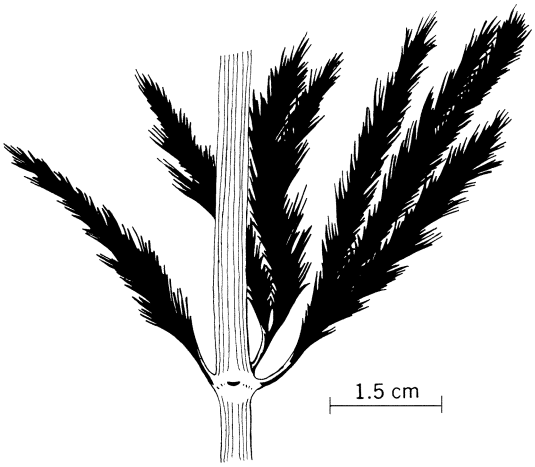

Pseudobornia **node with a whorl of three leaves.**

Pseudomonas

A genus of gram-negative, non-spore-forming, rod-shaped bacteria. *Pseudomonas* is a heterogeneous taxon; division at the genus level may be revised with some species being removed to *Xanthomonas*, *Comamonas*, *Alteromonas*, and *Acetobacter*. The type species is *P. aeruginosa*.

Motile species possess polar flagella. They are strictly aerobic, but some members do respire anaerobically in the presence of nitrate. Some species produce acids oxidatively from carbohydrates; none is fermentative and none photosynthetic. Pseudomonads are catalase-positive and usually indophenol oxidase–positive. Yellow-green fluorescent pigments are produced by some species, and blue, yellow, red, or brown nonfluorescent pigments are also characteristic of certain species.

Members of the genus *Pseudomonas* cause a variety of infective diseases; some species cause disease of plants. One species, *P. mallei*, is a mammalian parasite, and is the causative agent of glanders, an infectious disease of horses that occasionally is transmitted to humans by direct contact. *Pseudomonas pseudomallei* is the causative agent of melioidosis, an endemic glanderslike disease of humans and animals, in Southeast Asia. Pseudomonads such as *P. maltophilia*, *P. cepacia*, and *P. stutzeri* have been implicated in nosocomial human infections. *Pseudomonas aeruginosa* is the most significant cause of hospital-acquired infections, particularly in predisposed patients with metabolic, hematologic, and malignant diseases. The spectrum of clinical disease ranges from urinary tract infections to septicemia, pneumonia, meningitis, and infections of postsurgical and posttraumatic wounds.

Such toxic factors as lipase, esterase, lecithinase, elastase, endotoxin, and exotoxin are produced by *P. aeruginosa*, some of which may contribute to its pathogenesis. Serological, bacteriophage, and pyocin typing systems are used as markers to trace sources of infections to the epidemic foci. Serious *P. aeruginosa* infections require combination drug therapy with an aminoglycoside and a beta-lactam penicillin. *See* GLANDERS; HOSPITAL INFECTIONS; MENINGITIS; PASTEURELLOSIS; PNEUMONIA. Gerald L. Gilardi

Bibliography. M. Bendinelli, M. Campu, and H. Friedman (eds.), *Pseudomonas Aeruginosa as an Opportunist Pathogen*, 1993; R. Hugh and G. L. Gilardi, *Pseudomonas*: *Manual of Clinical Microbiology*, 6th ed., 1995; R. Y. Stanier, N. J. Palleroni, and M. Doudoroff, The aerobic pseudomonads: A taxonomic study, *J. Gen. Microbiol.*, 43:159–271, 1966.

Pseudophyllidea

An order of tapeworms of the subclass Cestoda, parasitic in the intestine of all classes of vertebrates. Typically, the head is simple in structure with two groovelike attachment organs (**Fig. 1***a*), the bothria.

Most pseudophyllideans are segmented and polyzoic with replication of the reproductive systems, although there are a number which do not show such replication and are monozoic. The genital openings are typically in the midline rather than lateral, and there is usually a uterine pore in the midline from which embryos are discharged (Fig. 1*b*). *Dibothriocephalus latus*, the broad or fish tapeworm of humans and certain piscivorous mammals, is a pseudophyllidean. In humans, this worm sometimes precipitates a pernicious anemia by competing with the host for vitamin B_{12}.

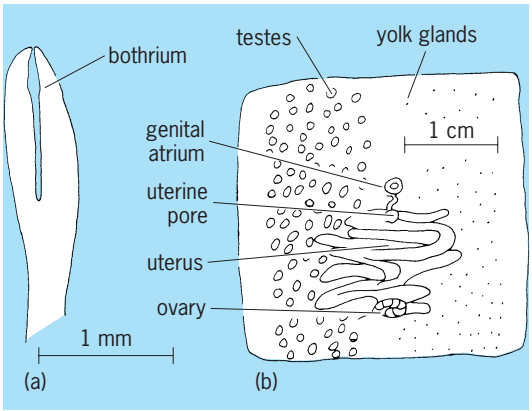

Fig. 1. Pseudophyllidea morphology. (*a*) Scolex of *Dibothriocephalus*. (*b*) A body segment.

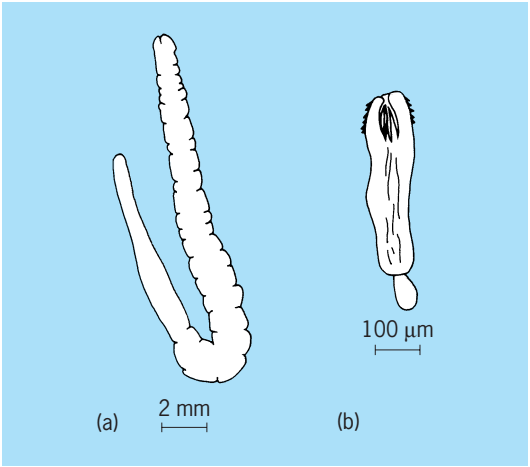

Fig. 2. Larval forms of *Dibothriocephalus*. (*a*) Procercoid larva. (*b*) Plerocercoid larva.

After leaving the intestine of the mammalian host, the ciliated embryo, or coracidium, of *D. latus* escapes from its shell and must be eaten by an arthropod, a copepod of the genus *Cyclops* or *Diaptomus*. In the copepod, the embryo develops into a procercoid (**Fig. 2***a*). For development to continue, the copepod must be eaten by a fish in which the worm develops into a plerocercoid larva (Fig. 2*b*). If this plerocercoid is eaten by the mammalian host, the worm rapidly grows to sexual maturity, sometimes attaining a length of more than 30 ft (9 m). Larval pseudophyllideans are occasionally found as parasites in the extraintestinal tissues of humans, producing a condition known as sparganosis. *See* EUCESTODA; SPARGANOSIS. Clark P. Read

Bibliography. L. H. Hyman, *The Invertebrates*, vol. 2, 1951; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; R. A. Wardle et al., *Advances in the Zoology of Tapeworms*, 1950–1970, 1974.

Pseudoscorpionida

An order of terrestrial Arachnida having the general appearance of miniature scorpions without the postabdomen and sting (**illus.** *a*). These animals are also known as the Chelonethida. The body length is seldom greater than 0.2 in. (5.0 mm). Typically, each finger of the anterior appendages, or chelicerae, has a serrula composed of a row of ligulate plates (illus. *b*). Ducts of silk glands open near the end of

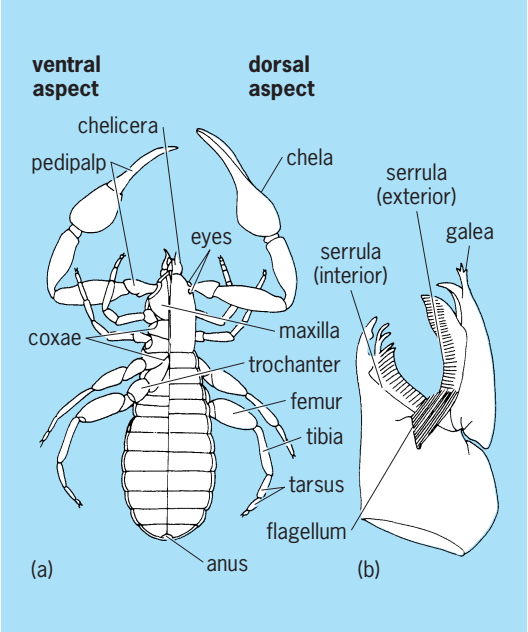

Pseudoscorpionida. (*a*) Adult, in two views. (*b*) Chelicera of male pseudoscorpion. (*After H. S. Pratt, A Manual of the Common Invertebrate Animals, rev. ed., McGraw-Hill, 1951*)

the movable finger, often in connection with a simple or branched spinneret, or galea. The second pair of appendages, or palpi, are large and conspicuous, usually with glands that discharge venom through a terminal tooth on one or both of the chelal fingers. The four pairs of legs are ambulatory. Eggs and larval young are frequently carried on the underside of the abdomen of the female. Pseudoscorpions feed chiefly on small arthropods and, although frequently found on birds, mammals, and insects, are considered nonparasitic. Pseudoscorpions are common in the nests of mammals, birds, and social insects, in woody debris and forest litter, under stones, and in crevices in the bark of trees. A few species are found in barns and chicken houses and the virtually cosmopolitan species *Chelifer cancroides* is frequently found in libraries and dwellings. Tertiary fossil pseudoscorpions from Baltic amber are very similar to modern forms. About 2000 species of pseudoscorpions have been described. *See* ARACHNIDA.

C. Clayton Hoff

Bibliography. M. Beier, *Pseudoscorpionidea*, in F. E. Schulze and W. Kükenthal (eds.), *Das Tierreich*, vols. 57–58, 1932; J. C. Chamberlain, The arachnid order Chelonethida, *Stanford Univ. Publ., Biol. Sci.*, 7(1):1–284, 1931; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; P. Weygoldt, *The Biology of Pseudoscorpions*, 1969.

Pseudosphaeriales (lichenized)

An order of the class Ascolichenes, shared by the class Ascomycetes. The order is also called the Pleosporales. The genera now assigned to this order were formerly classified in the Pyrenulales. They resemble the typical pyrenomycetous lichens except for the structure of the ascocarp, which is not a true perithecium. It is flask-shaped and lined with a layer of interwoven, branched pseudoparaphyses. The asci, with bitunicate walls, are located in scattered locules. Little is known about the structure and development of these ascocarps.

Except for the presence of symbiotic algae, these lichens are very close to the nonlichenized Pseudosphaeriales, but none are pathogenic. The usual habitat is tree bark, and the species are common in temperate and tropical regions. There are two major families. The larger one, Arthopyreniaceae, is a widespread family with at least 5 genera, the largest of which, *Arthopyrenia*, has more than 50 species. The Mycoporaceae is a small family with two well-known genera, *Dermatina* and *Mycoporellum*. All of the species in this order are crustose and many lack a well-defined thallus.

Mason E. Hale

Psilophytales

A group long recognized as an order of fossil plants (subdivision Psilopsida) collected in rocks of Late Silurian and Devonian age. It has been subdivided, on the basis of many data, into three categories whose descriptions include the chief kinds of plants formerly included in Psilophytales. Some authors will treat the three groups as classes under the division Rhyniophyta; others may prefer to consider them as three separate divisions of the subkingdom Embryobionta. They are given in this Encyclopedia as three classes of the division Rhyniophyta: Rhyniopsida, Zosterophyllopsida, and Trimerophytopsida. *See* EMBRYOBIONTA; PALEOBOTANY; RHYNIOPHYTA; RHYNIOPSIDA; TRIMEROPHYTOPSIDA; ZOSTEROPHYLLOPSIDA.

Harlan P. Banks

Psilotophyta

A division of the plant kingdom consisting of only two genera with three living species, *Psilotum nudum*, *P. complanatum*, and *Tmesipteris tannensis*. *Psilotum* is widespread in tropical and subtropical regions of both hemispheres, but *Tmesipteris* is confined to Australia and some of the Pacific islands.

The Psilotophyta have the typical life cycle of vascular cryptogams, with an alternation of sporophyte and gametophyte generations, the sporophyte being much the larger and more complex. The sporophyte consists of a slender, evergreen, more or less dichotomously branched stem with small, alternate leaves, which in *Psilotum* are tiny, simple or bifurcate scales without vascular tissue. The sporangia are compound, trilocular, and terminal on very short branches axillary to, but sometimes grown together with (adnate), some leaves. The underground part of the stem is slightly differentiated from the aerial part, and the plant has no true roots. The gametophyte is tiny, subterranean, and mycorhizal. *See* REPRODUCTION (PLANT).

The Psilotophyta have no economic importance, but they are interesting as possible remnants of an ancient (Silurian and Devonian) group of plants, the Rhyniophyta, which is regarded as ancestral to all other vascular plants. The dichotomously branched stems, morphologically terminal sporangia, and absence of roots are considered particularly significant in linking the two groups, and indeed some systems of classification put them into a single division under the name Psilophyta. Other botanists have suggested a possible relationship of the Psilotophyta to the Polypodiophyta or the Lycopodiophyta. *See* LYCOPHYTA; PLANT KINGDOM; POLYPODIOPHYTA; PSILOPHYTALES; RHYNIOPHYTA.

Arthur Cronquist

Psittaciformes

The parrots, a large order of land birds found worldwide but with most species concentrated on the landmasses of the Southern Hemisphere, particularly Australasia, the Neotropics, Africa, and southern Asia. Only a few species are found in the northern temperate regions. The parrots likely are closely allied to pigeons, from which they have probably evolved. *See* COLUMBIFORMES.

Classification. Although the parrots are generally placed in a single family, the Psittacidae, they are divided into a number of distinct subfamilies, which are listed below:

Order Psittaciformes
- Family Psittacidae
  - Subfamily: Psittacinae (parrots; 247 species; worldwide, but few in the Palearctic)
    - Cacatuinae (cockatoos; 18 species; Australasia to Indonesia)
    - Micropsittinae (pygmy parrots; 6 species; New Guinea)
    - Strigopinae (owl parrot; 1 species; New Zealand)
    - Nestorinae (keas; 3 species; New Zealand)
    - Loriinae (lories; 54 species; Australia to Indonesia)
    - Psittrichadinae (Pesquet's parrot; 1 species; New Guinea)
    - Loriculinae (hanging parrots; 10 species; southeast Asia)

Some workers consider the cockatoos (see **illustration**) more distinctive and have placed them in a separate family, the Cacatuidae. Among other changes, their tongue is considerably reduced in size; although it is still used to manipulate food, cockatoos can no longer drink by lapping water with their tongues. Rather, they have to scoop up water with their broad bill, which restricts them to areas with standing water. Since occupation of Australia by Europeans and the spread of ranching with water troughs for livestock, several species of cockatoos have expanded their range greatly on that continent.

Fossil record. Parrots have a very poor fossil record, as would be expected of a group of forest and grassland birds. The earliest form is from the lower Eocene of England, followed by a late Eocene fossil from France. Only one fossil parrot is known from the North American Tertiary and was found in the late Miocene deposits of Nebraska. Most of the few fossil parrots that have been found are from the Holarctic, outside the present geographic range of the order.

Characteristics. Parrots have a characteristic strong hooked bill. The tongue is large and fleshy; it is used for manipulating food in the bill, for obtaining nectar (many papillae exist on the tongue tip in some groups such as the lories), and for drinking. A short neck connects the large head to the stocky body. The wings are of medium length, varying from pointed to rounded. The legs are short and stout with strong, clawed toes arranged in a zygodactyl pattern, two pointing forward and two backward. The tail varies from short to long. Pigmy parrots climb on the vertical trunks of trees and are said to eat lichens on the bark. Parrots fly well and can attain high speeds, and the larger cockatoos can soar, but even those few parrots that are migratory journey only short distances. One species, the owl parrot, is flightless. Parrots can walk well but not rapidly, and the feet and the bill are used for climbing about in trees as well as for eating. Most parrots are arboreal, but some are terrestrial. A few are nocturnal. Plumage of parrots is variable; most often it is green, but red, orange, yellow, blue, black, and white in bold bright patterns are also common. Parrots are mainly vegetarians, eating seeds, nuts, fruit, nectar, or pollen; a few consume animal food.

Sulphur-crested cockatoo (*Cacatua galerita*). (*Photo by Lloyd Glenn Ingles; © California Academy of Sciences*)

Parrots are social birds and usually live in flocks. Almost all species have a strong pair bond, remaining together year-round for life. The nest is built in a cavity, generally in a tree, but sometimes in the ground, except for the monk parakeet (*Myiopsitta monachus*), which builds a large globular nest of sticks. Usually both sexes incubate the clutch of three or four eggs. The young are unfeathered and remain in the nest to be cared for by both parents. An exception is the owl parrot (Kakapo, *Strigops habroptilus*) of New Zealand. The males of this species establish courting stations in a loose lek, or courtship area, to which the females come. No pair bond is formed, and only the female incubates and cares for the young.

Parrots are found mainly in the tropics and in the southern continents, and the center of parrot radiation is the Australasian–New Zealand region. A large monophyletic radiation of psittacine parrots is found in the Neotropics, and a few species inhabit areas of Africa and tropical Asia.

Many parrots in captivity have an excellent ability to mimic words, and some can be taught a large vocabulary. However, parrots are not known to mimic

the calls of other birds in the wild. Their vocalizing ability is probably important for the constant communication between mates.

Economic significance. Parrots are the most important birds in the avicultural and pet trade, with many species kept in captivity. A number of species have escaped, and feral populations exist in the United States mainly in California and Florida. The monk parakeet is established in more northerly parts of the United States, such as the New York City area. Because the demand for parrots is so great, many countries have imposed exportation bans on native birds or importation restrictions on wild-trapped parrots. Two species, the cockatiel (*Nymphicus hollandicus*) and the budgerigar (*Melopsittacus undulatus*; the parakeet of the pet trade), have been domesticated, and several distinct breeds of each have been developed. Some species are seriously endangered because of habitat destruction, introduction of mammalian predators (such as weasels into New Zealand causing great reduction of the flightless owl parrot), and excessive capture of the pet trade. The Carolina parakeet (*Conuropsis carolinensis*) of North America is now extinct, as are several other species. Some psittaciforms are agricultural pests, consuming tree fruits. Parrots can transmit an especially virulent form of ornithosis to humans. *See* AVES; ENDANGERED SPECIES.

Walter J. Bock

Bibliography. J. M. Forshaw, *Parrots of the World*, 3d ed., Landsdowne Editions, 1989; J. M. Forshaw, *Parrots of the World: An Identification Guide*, Princeton University Press, 2005; T. Juniper and M. Parr, *Parrots: A Guide to the Parrots of the World*, Yale University Press, 1997.

Psocoptera

An order of insects frequently referred to as the Corrodentia, or Copeognatha. Common names for members of this order are book lice, bark lice, and psocids, the latter a general term for all members of the order. They are usually less than 0.25 in. (0.6 cm) long, though rarely some may reach about 0.5 in. (1.2 cm). Wings may be absent, and when present are of differing distinctive venational types. Tarsi are two- or three-segmented, cerci are absent (see **illus.**), and metamorphosis is gradual. Chewing mouthparts usually have a much enlarged clypeus; the lacinia of the maxilla is usually elongate and chisellike, and the antennae have 13 or more segments.

Book lice are most common among old papers on dusty shelves, in cereals, or other domestic situations. They are usually pale, wingless types of insects. Many bark lice, the majority winged, occur on the bark or foliage of trees, and some are found under dead bark or beneath stones. Nymphs of a few species occur on tree trunks as clusters of gregarious individuals, but disperse when mature. Food consists of microscopic molds, cereals, fragments of dead insects, and other organic debris. Sometimes psocids infest granaries, houses, straw packing, or museum collections of insects to the extent that injury, or a nuisance, is involved, but control measures are seldom necessary.

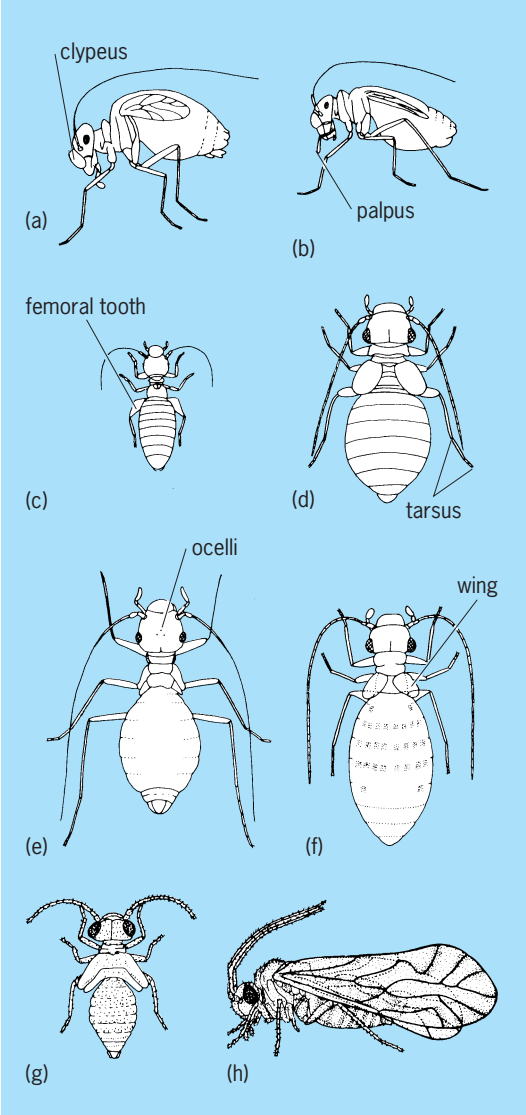

Some genera of Psocoptera. (*a*) *Psocathropos*; (*b*) *Dorypteryx*; (*c*) *Liposcelis*; (*d*) *Lepinotus*; (*e*) *Psyllipsocus*; (*f*) *Trogium* (*after A. B. Gurney, Pest Control Technology, National Pest Control Association*); (*g*) *Lachesilla*, nymph; (*h*) *Lachesilla*, adult.

Psocoptera are worldwide, especially in warm coun-tries, and some 1300 species are known. Current classification now lists about 27 families for this group. About 150 species, in 11 families, have been found in the United States. *See* INSECTA.

Ashley B. Gurney

Bibliography. D. J. Borror and D. M. DeLong, *An Introduction to the Study of Insects*, rev. ed., 1964; P. J. Chapman, Corrodentia of the United States of America, I: Suborder Isotecnomera, *J. N. Y. Entomol. Soc.*, 38:219–290, 319–403, 1930; E. L. Mockford and A. B. Gurney, A review of the psocids or booklice and barklice of Texas, *J. Wash. Acad. Sci.*, 46:353–368, 1956; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; C. N. Smithers, A bibliography of the Psocoptera, *Aust. Zool.*, 13:137–209, 1965.

Psychoacoustics

All of the psychological interactions between humans (and animals) and the world of sound. It encompasses all studies of the perception of sound, as well as the production of speech.

Hearing. The simplest sounds—that is, pure sine waves—can be described in terms of two physical characteristics, frequency and intensity. The primary perceptual attributes of sound are loudness (perceived intensity from soft to loud) and pitch (perceived quality from low to high). However, there is no one-to-one correspondence between the physical characteristics and the perceptual attributes. Although loudness depends largely on a sound's physical intensity, it also depends on frequency. A basic measurement is the audibility function, the absolute threshold for detecting sound as a function of sound frequency. Threshold is far from uniform across the audible spectrum. The physical intensity needed to detect a sound with a frequency of 100 Hz must be a million times greater than the intensity needed to detect a sound with a frequency of 1000 Hz. *See* LOUDNESS.

Similarly, though the pitch of a pure tone depends primarily on its sound frequency, it also depends to some extent on intensity. As intensity increases, low-frequency tones decrease in pitch, and high-frequency tones increase in pitch. A complex sound, composed of many frequencies, has a pitch that usually corresponds to the pitch of the fundamental frequency. *See* PITCH.

Musical sounds also vary in the attribute known as timbre, which characterizes the sounds of individual instruments. The same note played on an oboe and a piano may have the same fundamental frequency and therefore the same basic pitch; however, the notes differ in timbre (quality) because of differences in their patterns of higher-frequency harmonics and because of differences in the temporal characteristics of onset of the notes. *See* MUSICAL ACOUSTICS.

The localization of a source of sound is made possible by the spatial separation of the two ears. A sound coming from the left arrives at the left ear before the right ear, and this difference in time of arrival makes it possible for the auditory system to determine the location of the sounds. Intensity differences at the two ears provide additional information about location. With high-frequency sound waves, the head acts as a barrier, reducing the sound intensity at the opposite ear. Low-frequency sound waves tend to bend around the head, so the intensity difference between the ears is much smaller. Movements of the head are particularly important in enabling humans and animals to locate the source of a sound.

The capacity to detect sounds in the environment and to discriminate among sounds depends on the frequency spectra, intensities, and temporal properties of the sounds. This capacity also depends on the presence of other acoustic energy in the environment, usually noises, which may mask the sounds. In general, masking agents exert their most deleterious effects on sounds with the same or similar frequencies. In order for a person to detect a target sound above the level of background masking noise, the masking noise cannot be more than 10 dB greater in intensity than the target. Even though a sound can be detected in noise, it may not be possible to discern what the sound is. For speech to be fully intelligible, it must have an average intensity level 20–30 dB above the background noise. *See* ACOUSTIC NOISE; HEARING (HUMAN).

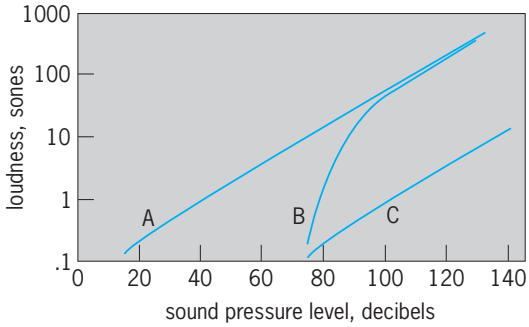

Fig. 1. Relationship between perceived loudness in sones and sound pressure level in decibels for a normal ear (curve A), an ear with a sensory-neural hearing loss (B), and an ear with a conductive hearing loss (C).

Hearing loss and deafness. Hearing deficits may be congenital or acquired. It is especially important to detect cases of congenital hearing loss early in life—before the age of 2 or 3 years—because the acquisition of normal speech and language can be permanently impaired by profound deafness. Modern audiometric techniques permit the evaluation of hearing in infancy, and infants and young children can be fitted with hearing aids tailored to their individual needs in the same way adults can. *See* AUDIOMETRY; HEARING AID.

There are basically two kinds of hearing loss, conductive and sensory-neural. In conductive hearing loss, the deficit stems from the failure of sound to be transmitted properly to the receptors in the inner ear. Often, this results from blockage in the ear canal or from fusion of the bones in the middle ear, the latter condition known as otosclerosis. Because conductive loss consists of a deficiency in sound transmission, its effect appears as a constant relative loss in loudness, as shown in **Fig. 1**. A conductive loss of 60 decibels means that every sound is reduced in physical intensity by a factor of a million. In terms of perceived loudness in sones, an ear with a conductive loss of 60 dB hears sounds 1/64 as loud as a normal ear hears them.

Sensory-neural hearing loss typically results from damage to the receptive elements in the ear or to the auditory nerve itself. This is the kind of hearing loss that in many individuals encroaches with advancing age. Typically, such losses are most evident at high sound frequencies (above about 1000 Hz). Unlike conductive disorders, which affect intense and weak sounds equally, sensory-neural losses show up mainly at low sound levels, which are inaudible or much attenuated perceptually. High sound levels, however, may be perceived to be just as loud by a

person with a sensory-neural loss as by a person with normal hearing. This is shown in Fig. 1. The rapid rise in loudness perception just above threshold is called recruitment. *See* HEARING IMPAIRMENT.

Speech production. Speech sounds are acoustic patterns imposed on air expelled by the lungs. The specific properties of speech—sound intensity, frequency spectrum, and time course—depend on the force with which the air is expelled, by the vibrations of the vocal cords in the larynx, by the shapes of the cavities in the mouth and throat, and by the movements of the tongue, lips, and jaw. It is the position of the tongue that is most important in determining individual vowel sounds, while individual consonants are determined by place and manner of articulation, and by whether they are voiced or unvoiced. *See* PHONETICS; SPEECH.

Speech perception. Most of the energy in the sound waves of speech falls in the frequency region from 50 to 10,000 Hz. Of this, the region from about 200 to 7000 Hz carries most of the information for speech perception. Vowel sounds are characterized by bands of energy concentrated at certain frequencies; these energy bands are known as formants. The spectral locations of the formants correspond to resonances in the vocal tract. The formants are numbered in order or increasing sound frequency: first, second, third, and so on. In general, the first three formants—especially the first and second—are most important in identifying vowels.

Bursts of noise, pauses, and transitions to the formants mark locations in the speech signal where consonants are heard. It is not possible, however, to identify every consonant by means of a single cue in the acoustic spectrum: The very same burst of noise may be heard as *k* when it precedes the vowel *ah* but as *p* when it precedes *oo* or *ee*. Similarly, very different acoustic events can be perceived as the same consonant when they appear in the context of different vowels. The perception of consonants, therefore, depends not only on the acoustic signal at that moment but also on the pattern of acoustic information surrounding it.

The intelligibility of speech signals transmitted over a communication channel depends on such factors as the intensity of the sound, the fidelity of the communication system, and the presence of masking agents such as noise. It is common for communication systems to truncate the sound waves at high amplitudes; this type of distortion is known as peak clipping. Fortunately, intelligibility is affected little by removing high-amplitude peaks in a sound wave, for it is the central region of the speech wave, where information about consonants is represented, that is most important.

Although most of the energy in speech comes from vowels, most of the information comes from consonants. For this reason, as much information is lost if all frequencies above 1900 Hz are removed as if all frequencies below 1900 Hz are removed, despite the fact that the frequencies above 1900 Hz contain only about 2% of the average energy in a speech signal: sound frequencies above 1900 Hz are largely devoted to consonants, frequencies below 1900 Hz to both consonants and vowels.

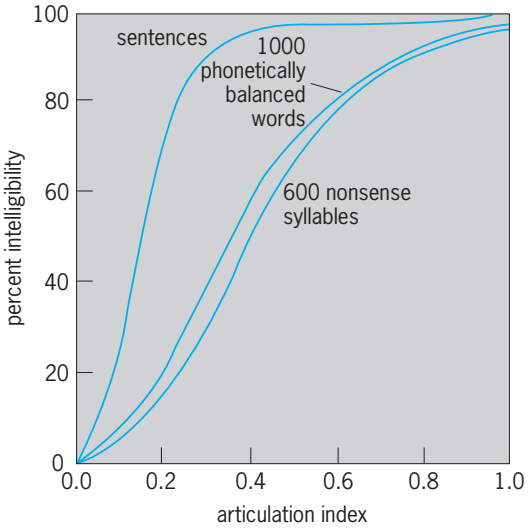

Fig. 2. Relationship between intelligibility and the articulation index for various types of speech.

It is possible to predict the intelligibility of speech by the following scheme: The frequency range from 200 to 6100 Hz is divided into 20 bands, each of which has the potential to contribute equally to intelligibility. Each band's actual contribution is determined by the intensity of the speech signal relative to any masking noise in that band. The actual contributions of all 20 bands sum to yield an articulation index, a variable that ranges from 0.0 to 1.0. From the articulation index one can determine the intelligibility of various kinds of speech, whether simple syllables, single words, or connected discourse (**Fig. 2**). *See* PERCEPTION; PSYCHOLINGUISTICS.

Lawrence E. Marks

Bibliography. R. B. Batkin, *Hearing and Hearing Disorders*, 1988; R. J. Dooling and S. H. Hulse (eds.), *The Comparative Psychology of Audition: Perceiving Complex Sounds*, 1990; N. J. Lass, *Speech, Language and Hearing*, 1989; R. D. Luce, *Sound and Hearing*, 1992.

Psychoanalysis

Psychoanalysis is a psychological theory; a form of psychotherapy, especially for the treatment of neurotic and character or personality disorders; and a method for investigating psychological phenomena. Psychoanalysis was created and developed by Sigmund Freud, who presented his method, clinical observations, and theory in *Interpretation of Dreams* and other major works, including *The Psychopathology of Everyday Life* and *Three Essays on the Theory of Sexuality*, as well as in many of his case studies.

Psychoanalytic theory. Generally, psychoanalysis is concerned with the causal role of wishes and beliefs in human life. More specifically, it attempts to explain mental or behavioral phenomena that do not appear

to make sense as the effects of unconscious wishes and beliefs. Such phenomena include dreams, disturbances in functioning such as slips of the tongue or pen and transient forgetting, and neurotic symptoms.

Typically, unconscious wishes and beliefs are constituents of conflicts. Most simply, the conflict consists of a wish, a belief that there is danger in gratifying the wish, and a solution of sorts. The solution can take the form of a sequence of imagined events, a fantasy, in which the wish is fulfilled in such a way that the danger is evaded or mitigated. Such wishes, beliefs, and fantasies constitute the person's psychic reality, which influences conscious perceptions, feeling states, memories, beliefs, thought processes, and actions.

The term unconscious in psychoanalysis does not mean simply that mental contents are out of awareness. Its psychodynamic meaning is that the person does not want to be aware of these contents, and takes active steps to avoid being aware of them. A fundamental hypothesis of psychoanalysis is that because a mental entity is dynamically unconscious it has the causal power to produce the phenomena that are of interest to psychoanalysis.

At first, the dynamic unconscious was thought to consist of traumatic memories. Later, it was believed to consist of impulses or wishes—especially sexual (and aggressive) impulses or wishes—having some causal link or, at least, some similarity to impulses or wishes in childhood. The significance of memories of actual events is believed to lie in their connection to impulses or wishes. Psychoanalysis now emphasizes that the dynamic unconscious consists of fantasies, which have a history reaching back to childhood. These fantasies are internal scenarios in which sexual (and aggressive) wishes are imagined as fulfilled. Wish fulfillment occurs in association with imagined threats. Therefore, anxiety signals danger when an unconscious fantasy is activated and the person is then motivated to modify the internal scenario in such a way as to mitigate the threat and thereby avoid any further anxiety. A person elaborates the fantasy in attempting to discover a way to fulfill the wish while more or less successfully evading the imagined threat.

The construction of an internal scenario throughout a lifetime, then, involves layers of revisions resulting from different responses to the same threat, or to different or multiple threats, imagined during different developmental epochs. These revisions are essentially symbolic operations (referred to as defense mechanisms, dream work, or primary processes) on the contents of mental states. Defense mechanisms, such as repression, expel or delete contents from consciousness. In addition, they may serve to prevent these contents from returning to consciousness—as reaction formation does by substituting for an expelled wish or impulse its opposite (for example, love for hate, kindness for cruelty). Other defense mechanisms dispose of rejected wishes or impulses by altering the contents of a self-representation or object-representation, or by shifting attention from one to the other. The dream work or primary processes, which include condensation and displacement, transform latent mental contents into manifest contents that may not make sense. Condensation is the use of one manifest content to stand for or allude to a whole set of latent ideas. Displacement is a shift of attention or emphasis from a significant latent content to an indifferent manifest content. Images, which are referentially vague or ambiguous, may be used instead of words to stand for or allude to latent ideas. Iconic (metaphoric) symbols, also known as Freudian symbols, may be used to stand for or allude to latent ideas that go unrecognized in this guise.

Psychoanalysis is distinct in attributing causal powers to unconscious sexual wishes. Such attribution depends on extending the meaning of sexual to encompass the quest for sensual pleasure in childhood (so-called infantile sexuality) and choices of objects and aims other than those exemplified by two adults of the opposite sex engaging in coitus. Unconscious fantasies have themes that correspond to postulated psychosexual stages of development. These stages are characterized and dominated by experiences, thoughts, feelings, kinds of relations with others, and kinds of sensual pleasure associated with a particular zone of the body, including the skin, mouth, teeth, anus, and genitals. One theme that is thought to have particular importance is the Oedipus complex, in which the child rivals one parent in seeking sensual gratifications of various kinds from the other parent.

When an unconscious fantasy is activated, it manifests itself in conscious mental states or in actions—importantly, in emotions; in interpretations of the significance of events or states of affairs; in attributions of motives to others; and in daydreams, dreams, and neurotic symptoms. An unconscious fantasy may be activated periodically and spontaneously, as it is part of its nature to so manifest itself, independently of situational vicissitudes. It may be activated by something in an external situation that reminds the person of the fantasy. For example, a fantasy may come to be manifested by a mood that seems either out of place or not to make sense in the situation that evokes it but that belongs to and makes sense in the unconscious internal scenario. An unconscious fantasy may also be activated by the willful mental activity of a person who wants to experience one of its manifestations, such as a conscious daydream or fantasy, a particular mood or emotion, or pleasure or relief from anxiety.

Unconscious fantasies, as distinct from both conscious reality-oriented imagining and conscious daydreaming, are constructed when imagination functions under very special conditions. The person, immersed in imagining, ignores the fact that he or she is merely imagining: the truth of beliefs is not permitted to matter. Connections are made that have nothing to do with what is true by means of such nonlogical mental operations as condensation, displacement, translation of verbal symbols into imagery, and iconic, or metaphoric, symbolization. The mind is imagined to be both spatial and corporeal: spatial

because it can be inhabited by actors in internal scenarios; and corporeal since it is imagined to be a body or a part of a body and is therefore imagined to have the causal powers of a body or part of a body. Because people can identify with an actor in the internal scenario, what is imagined as giving pleasure can be actually experienced as pleasurable.

This emphasis on fantasy underscores the fact that psychoanalysis gives priority to the relation between wishes (including wishes a person knows could not conceivably be gratified in reality) and imagination (functioning under very special conditions). In particular, psychoanalysis focuses on how sexual wishes and imagination are implicated in the vicissitudes of unconscious wish fulfillments, their construction, and their effects. This is in contrast to other branches of the psychology that, for example, might give priority instead to relations among conscious mental states, and in particular to the causal relation between desire and belief acting together, on the one hand, and rational or irrational action, on the other, where irrational action arises because beliefs are erroneous or relevant beliefs are missing.

Variants of psychoanalytic theory have emphasized the role of destructive object-relations in determining inner reality; how attempts to avoid anxiety or blows to self-esteem lead to distortions in interpersonal relations; the effects of actual early relations, especially with the mother, on the development and maintenance of a sense of identity and self-esteem; and the effects of deficits in early life on psychological capacities and personality. All four of these variants share two characteristics: They tend to assign a major (if not exclusive) causal role to experiences in external reality, especially during the preverbal period of infancy and in relations with the mother. They also tend to deemphasize the causal role of sexual wishes and to reject the primacy of the quest for sensual pleasure.

Psychotherapy. In psychoanalytic psychotherapy, 45- or 50-min sessions are scheduled three to five times a week over a number of years. The patient lies on a couch rather than sitting facing the psychoanalyst, so that the maximum amount of attention is directed to what is within the person rather than to external cues.

Free association is the method of psychoanalysis, both as a form of psychotherapy and as a method of investigation. Patients are encouraged not to talk about some particular problem or aspect of their lives but rather to suspend any conscious purposive organization of what they say, speaking freely whatever comes to mind, and avoiding as much as possible censoring of what may seem trivial, embarrassing, shameful, bad, offensive, or senseless. Both psychoanalyst and patient follow the patient's productions: conscious purposes are replaced by unconscious purposes, which, under these conditions, can determine the direction of the patient's mental processes with less interference. Information that is not easily if ever acquired in other situations can be obtained in the psychoanalytic situation because the participation of the psychoanalyst is disciplined, and the influence of external cues and internal inhibitions on patients and what they say is minimized.

Psychoanalysts do not take what their patients talk about at face value. Interventions are predominantly interpretative; psychoanalysts do not seek primarily to tell their patients what to do, to educate them about the world, to influence their values, or to reassure them in one way or another that everything is or will be all right. Psychoanalysts look for patterns in what each patient says and for signs of feelings of which the patient is more or less unaware. They then engage their patients (who are increasingly aware of these patterns and able to experience and articulate these feelings) in an inquiry about the reasons for them or motives behind them. The focus is on what the patients do not know—and do not want to know—about themselves and their inner life, including strategies for avoiding such knowledge and the consequences of these strategies. Much of the work of psychoanalysis involves coping with the ways in which patients, more or less wittingly, oppose what they regard as efforts to interfere with these strategies. Interpretation of resistance is a distinctive feature of psychoanalytic psychotherapy.

The goal of psychoanalytic psychotherapy is to extend the realm of what patients permit themselves to experience. That does not mean simply to know intellectually about a mental state, but to know firsthand what it feels like to be in such a state. This change occurs especially as patients begin to attribute to their psychoanalysts what are actually their own wishes, impulses, and attitudes or to identify their psychoanalysts with early significant figures who were the object of these wishes, impulses, feelings, or attitudes, or who were seen or imagined as responding to these wishes, impulses, feelings, or attitudes in certain ways. Interpretation of transference is another distinctive feature of psychoanalytic psychotherapy.

Any person can have multiple wishes. Since some of these conflict with each other and some compete for inevitably limited resources for their gratification, they cannot all be gratified. The inevitable normal misery that this as well as unfavorable conditions in external reality causes patients is not what psychoanalytic psychotherapy seeks to alleviate. Rather, it tries to mitigate the misery that patients with a neurotic, character, or personality disorder inflict on themselves by their own more or less unwitting attitudes toward their own wishes and feelings, or by the more or less unwitting strategies that they use to avoid becoming aware of these wishes or feelings or anything associated with them, including their own attitudes toward them.

Method of investigation. The case-study method is characteristic of psychoanalytic research. The arguments that can be used in case studies are not much different from those used by Charles Darwin in his arguments for the credibility of the theory of evolution: namely, analogy (the use of familiar or homely models in which postulated causes and mechanisms can be shown to exist); consilience (the

convergence of inferences from different kinds of information on a common cause); and abduction (inference to the best explanation). *See* PSYCHOTHERAPY.

Marshall Edelson

Bibliography. M. Edelson, *Psychoanalysis: A Theory in Crisis*, 1990; S. Freud, *Interpretation of Dreams*, 1900; S. Freud, *Introductory Lectures on Psychoanalysis,*, 1966, reprint 1989; S. Freud, *Question of Lay Analysis*, 1964, reprint 1969; E. Jones, *The Life and Work of Sigmund Freud*, 3 vols., 1953–1957; L. A. Salome, *Freud*, 1994; R. Waelder, *Basic Theory of Psychoanalysis*, 1960.

Psycholinguistics

An area of study which draws from linguistics and psychology and focuses upon the comprehension and production of language in its spoken, written, and signed forms. Although psychologists have long been interested in language, and the field of linguistics is an older science than psychology, historically scientists in the two fields have had little contact. The two fields were thrown together, however, by the publications of Noam Chomsky in the late 1950s. Chomsky's writing had the effect of making psychologists acutely aware of their lack of knowledge about the structure of language, and the futility of focusing theoretical and research attention upon the surface structure of language while arguing that linguistics is as much concerned with the mind as psychologists should be. Psychologists, in turn, made linguists more aware of the larger cognitive context into which language fits as well as a broader set of methodological approaches available to scientifically test hypotheses about language. As a result, psycholinguists, who have a background of training in both linguistics and psychology, have been attempting since the early 1960s to gain a better understanding of the abstract rules which underlie human language and how they are acquired, used to communicate meaningful messages, and influenced by the biological and psychological context in which they occur. Research has been directed to the evolutionary development of language, the biological bases of language, the nature of the sound system, the rules of syntax, the nature of meaning, and the process of language acquisition. *See* LINGUISTICS.

Evolution and biological bases of language. While animals do convey messages to each other, it is clear that there are species-specific, finite limitations to those messages. Lower animals have limited capacities for generating new and appropriate messages when the occasion arises. Some animals can be taught to expand their communicative capabilities. The achievements of nonhuman primates which have been taught to communicate with humans and other primates via sign language, symbolic objects, and computer terminals are well beyond initial expectations. While it has been hard work for both the animal learners and the human teachers in these experiments, the achievements of the primates have been greater than expected, although hardly comparable to the human use of the symbolic language system so easily and naturally acquired by the human child. *See* ANIMAL COMMUNICATION.

Efforts to reconstruct the evolutionary development of language suggest that early hominid forms were structurally incapable of language. Their supralaryngeal vocal tracts, for example, were similar to those of present-day nonhuman primates. The modification in the supralaryngeal tract, along with other physiological and neurological changes which allowed language to develop, probably occurred long after the evolutionary branching which separated humans from other primates. Thus, although the other primates are the closest phylogenetic relatives of humans, there is little likelihood that they and humans are closely related with regard to symbolic capability.

The neuroanatomical bases of language have been studied by examining the language functioning of individuals with known brain injuries and those whose corpus callosum has had to be severed, by electrical evoked potentials taken from the scalps of normal persons, and by dichotic listening studies. Research in this area supports the early Greeks' hypothesis that there is a lateralization of brain function such that the left hemisphere appears to be much more intimately involved in most language functions than the right hemisphere. There is now evidence, however, that both hemispheres are involved to some extent in most language functioning, especially those aspects relevant to meaning. Differentiation of brain functioning seems to be present from birth, although developmental changes in function, plasticity of function, and cerebral hemispheric support for the functioning continue well beyond birth. The theoretical accounts of the relation between brain function and brain organization remain, however, a matter of considerable controversy. *See* BRAIN; HEMISPHERIC LATERALITY.

Sound system of language. An examination of the acoustic properties of speech reveals that the physical properties of speech sounds are unusually complicated insofar as their relations to the phonological components of the language system are concerned. The fundamental frequency set in motion by the air forced by the lungs through the vocal folds is modulated by the shape and length of the supralaryngeal vocal tract. The constantly changing configuration of the vocal tract during speech produces noise, defining some consonants and transitions to and from relatively steady-state formants, or bands of acoustic energy, defining vowels, to create syllables composed of consonants and vowels. While the vowels have fairly constant acoustic properties modified by the contexts in which they occur, the consonants seem to have no apparent one-to-one relation to the acoustic signal. *See* PHONETICS.

The phonological system is, however, abstracted in terms of discrete segments, phonemes, or syllables from the continuous sound wave. It appears that humans have a phonetic mode, special mechanisms that are engaged to process speech sounds differently than other auditory signals. These

special mechanisms involve an integration of the motor gestural system, usually associated only with production, and the perceptual system, usually associated only with reception, such that neither system operates without engaging the other. This system accounts for the fact that speakers of a language will perceive radically different acoustic signals produced in different contexts as the same phoneme or syllable.

Before acquiring speech, the very young infant appears to be specifically attuned to linguistically relevant acoustic characteristics of the signal and capable of making distinctions among phonemes and syllables of speech as well as clausal units in natural speech. When speech production begins at about 1 year of age, the child begins to acquire the phonology of the language, which includes the distinctive features that define the phonemes of the language, the suprasegmental features such as pitch and stress that extend over syllables or words, and the phonological rules that apply to phonemic strings to alter them in various ways to derive their phonetic pronunciation. *See* SPEECH.

Morphology. Early research ignored the nature of words, or morphology, because in Chomsky's linguistics emphasis had been focused upon the sentence, but some studies have been exploring the manner in which words are stored in the lexicon, or mental dictionary, of the individual, the acquisition of rules that apply to morphemes derived from, for example, Latin as opposed to English, and the influences of context upon access to lexical items. Results suggest that morphemes may be stored independently of the various inflectional forms that may modify them; differential rules for inflecting morphemes derived from different sources are known to the adult speaker and are acquired at a relatively early age; and content effects appear to be important to the lexical processing of words.

Syntax of language. The early research on comprehension and production of syntax began with the assumption that linguistic theories of syntax were psychologically based. That assumption was quickly challenged as it became clear that, in addition to the rules of syntax, processing of syntax involves memory, meaning, context, conversational rules, and the general knowledge of the persons involved in the communication. Much of the early research focused upon the processing of single sentences in isolation. It was demonstrated that the underlying structure of sentences which relates parts of sentences to each other affects, for example, speech production and perception; that ambiguous sentences sometimes take longer to process; that syntactically related sentences are more easily confused than unrelated sentences; and that negative sentences take longer to process than affirmative sentences. Some of these results are attributable to the finding that ambiguous words that have multiple meanings are processed more slowly because all of the meanings are called to mind in the sentence context prior to selecting the appropriate meaning for the particular context. Processing negative sentences appears to involve establishing the affirmative version of the sentence and then negating it. It was also discovered that when sentences are presented in memory tasks, people tend to forget the syntactic form of the sentences but retain the meaning unless they are specifically required to remember the syntax. In addition, it was found that sentences cannot be analyzed in isolation, because each sentence involves presuppositions and implications not explicitly stated but taken into account more or less automatically by a person processing those sentences. Furthermore, it has become clear that many of the influences of syntax on processing of sentences in isolation do not seem to be important when sentences are a part of a larger meaningful conversation or story. It is clear that syntax is required for the comprehension and production of meaningful sentences, but it has become progressively clearer that, from a psychological point of view, syntax cannot be treated in isolation. However, the question of whether the processing of syntactic as well as other sentence components occurs in isolated modules or as an interactive system is a controversial issue. The search for a psychological theory of syntax has led to a host of hypotheses and a number of techniques to test them. The simplest hypotheses have been rejected, and the complexity of sentence production, perception, and memory has become increasingly apparent. *See* INFORMATION PROCESSING (PSYCHOLOGY); MEMORY.

Meaning of language. Obviously language is used to convey meaningful messages from one person to another, and yet meaning, as such, has been the least studied component of language. The reason for this lack of attention has been the lack of a comprehensive theory to stimulate research. Three theoretical approaches, two of which have not proved totally satisfactory and a third too new to evaluate, have dominated thinking in this area. The first, which comes directly from linguistics, takes the word as the unit of analysis and focuses upon a componential analysis of words into feature elements of which meaning is composed. Synonymy, antonymy, and other relations among words are accounted for in terms of the presence or absence of the feature elements making up the meaning of the words involved. Analysis of a whole word into its component elements requires a knowledge of the whole of which the elements are the parts; recognition of this has led other theorists to propose alternative theories which are more holistic in nature. The latter conceptualizations of meaning take the abstract conceptual idea as a unit, and attempt to infer the structure of the concept based upon the relations among exemplars of that concept as used by speakers. The structure of the concept of fruit, for example, may be inferred from the relations among the exemplars apple, pear, guava. The apple exemplar is closer to the core meaning, or is more prototypic, of the concept fruit than is, for example, guava. Based upon this conceptualization of meaning, borrowed from L. Wittgenstein, it has been found that prototypic exemplars can be classified more rapidly, learned more quickly, and remembered more easily than peripheral exemplars.

Neither of these theoretical approaches takes into consideration linguistic and extralinguistic contextual effects upon meaning, nor the fact that the broader knowledge system of the communicator influences the meaning of language.

A theoretical approach called connectionism is based upon the transmission of activation levels in large networks of densely interconnected simple units rather than the manipulation of structured symbolic elements. The tie to the computer and the analogy to neural networks make this an appealing but as yet untested theoretical approach. An important test may come with efforts to account for figurative language, where other theories have had difficulties. In any case, some inroads are being made into the understanding of meaning, but the problem remains relatively intractable. *See* INTELLIGENCE; NEURAL NETWORK.

Language acquisition. Actual production of language is not evidenced until about the first birthday (although the signing child may start sooner), when repeated syllables such as "mama" and "papa" are likely to be the earliest utterances. These early utterances appear to be articulations of the mental and emotional states of the child. When a single word is the productive limitation of the child, extralinguistic context is used effectively to convey the message intended. In effect, the context is used by the child (and the listener) to expand the syntax of the message which the child is unable to produce linguistically. Gradually, the child frees himself or herself from the immediate context by adding to the linguistic portion of the message as syntactic and morphophonemic rules are acquired. Syntactic and morphophonemic complexities are usually expressed initially in forms simpler than those used by adults. The child seems to formulate increasingly complex hypotheses or strategies for coping with the problems of encoding and decoding messages.

As for the complexity of the messages themselves, research has led investigators away from language as such and into the broader realm of cognitive development. The nature of what children discuss is directly related to the level of their cognitive development. Children express their affective and cognitive mental states. However, there seems to be little relation between the acquisition process and what is ordinarily thought of as efforts to teach language by reward and imitation. Clearly children take advantage of exemplary input of various forms when they are ready to acquire those forms, but they are neither imitating nor learning by reward as adults decide to teach. It is apparent that adults and children do adapt their language input to the level of the child acquiring the language, but there is little evidence that the child attempts to imitate their language. Rather, the child formulates broad rules, and the processes of acquisition appear to be a matter of refining those rules to better express his or her mental contents in terms of the specifics of individual structures in the language. While the strategies have been shown to vary among individual children, and the structures for some messages are more complicated in some languages than in others, the construction of utterances that follow the rules for comprehending and producing specific constructions within a language are remarkably regular between children, both within a language and across languages.

In trying to understand the acquisition process, however, developmental psycholinguists have been driven at every turn from the study of phonology and syntax as independent modularly separate units to the interrelations among them and, in particular, to the study of semantics as well. Each component affects the others in the process of acquisition. While sound and syntax provide the medium for the message, it is meaning that is the message. The acquisition of that meaning has been shown to relate to biological givens, the broader cognitive system, and the relation of those biological and cognitive factors to the context in which the child develops. Analytic feature theories, synthetic prototype theories, and connectionist theories have been related to the general cognitive development of the child in an effort to understand the child's semantic development. In addition, learnability theory has been introduced in an effort to specify a class of grammars that include the grammars of all languages and to identify the language-learning procedures that a child could use to learn those languages, given the data available to the child. The child is certainly constrained in some ways in the acquisition process. For example, children overextend some words and underextend others, but they seldom form bizarre relations between words, and concepts. It seems clear that children have some sort of theoretical framework for conceptualizing the world with which they interact, and they search for the proper language tools for expressing their concepts and the relations among those concepts. *See* COGNITION; PSYCHOACOUSTICS.

David S. Palermo

Bibliography. D. W. Carroll, *Psychology of Language*, 2d ed., 1994; V. A. Fromkin and R. Rodman, *An Introduction to Language*, 6th ed., 1998; J. B. Gleason and N. B. Ratner (eds.), *Psycholinguistics*, 2d ed., 1997; R. Hasan and J. R. Martin, *Language Development: Learning Language*, *Learning Culture*, 1989; C. K. Ogden and I. A. Richards, *The Meaning of Meaning*, 1989; M. Singer (ed.), *Psychology of Language: An Introduction to Sentence and Discourse Processing*, 1990.

Psychology

The study of human behavior and mental processes. Psychology is distinguished from similar fields of research and applied science by the unit of analysis, which is the individual. However, some psychological studies stretch this definition. For example, a study of group problem-solving may compare groups of different types, in which case the group is the unit of analysis. However, a psychological theory will still explain the effects by processes that apply at the level of individuals in a group.

Applied and experimental aspects. Psychology is sharply divided into applied and experimental areas. Although the field of psychology is devoted to and determined by clinical treatment of psychological problems, basic research in the field, which includes analysis of many problems, such as thought and perception, has little (if any) relation to clinical applications. However, many fields are represented in both research and applied psychology. For example, clinical psychology exists as an area of application in which individuals are treated for many types of disorders; however, on the research side, there is extensive work on the causes, diagnosis, and treatment outcome of psychological disorders. In many cases, a practitioner or researcher contributes to both aspects of the area. Despite such crossovers, research and application often remain distinct in methodology, purpose, institutional venue, training, professional affiliation, and publication outlets.

Researchers in psychology study a wide range of areas. Cognition includes such areas as human memory, use of language, perception, and attention. Cognitive research is often included as part of subdiscipline called cognitive science. This area examines central issues such as how mental process work, the relation between mind and brain, and the way in which biological transducing systems can convert physical regularities into perceptions of the world. Cognitive science is carved from the common ground shared by computer science [especially simulations of perception and thought (artificial intelligence and machine perception)], cognitive psychology, philosophy of mind, linguistics, neuropsychology, and cognitive anthropology. The study of human attention is a cognitive area that is central in the field. *See* COGNITION.

The study of consciousness involves such basic questions as the physiological basis of mental activity, the freedom of will, and the conscious and unconscious uses of memory. The latter topic can be classified under the rubric of implicit memory, in which people (sometimes also neurological cases with various forms of amnesia) can be shown to dissociate memory use from awareness of memory. That is, they might be unable to retrieve material learned, or may even be unaware that they learned it, while demonstrating in their behavior that they are using the information. *See* INSTINCTIVE BEHAVIOR; MEMORY; PSYCHOLINGUISTICS; SENSATION.

Social psychology. Social psychology includes the study of interactions between individuals and groups, as well as the effects of groups (whether present or imagined) on the attitudes, opinions, and behavior of individuals. The field covers such topics as persuasion, conformity, obedience to authority, stereotyping, prejudice, and decision making in social contexts. Although much interest in the area has been stimulated by questions that arise from pressing social issues, the overall motivation has been to develop general theoretical accounts of human social behavior. *See* MOTIVATION; PERSONALITY THEORY.

Developmental psychology. This discipline has three subfields: life-span development, child development, and aging. Most research in the area concentrates on child development, which examines the development of abilities, personality, social relations, and, essentially, every attribute and ability seen in adults. Life-span development and aging are growing fields of interest. *See* AGING; INTELLIGENCE.

Practitioners. The public awareness of psychology is dominated by the field of clinical psychology, which represents a fraction of theory and research in the field but which is the only point of contact for most of the population. A clinical psychologist is usually known by the term psychologist, which in some states is a term that can be used only by a registered practitioner. A psychiatrist is a physician with a specialty in psychiatric treatment and, in most states, with certification as a psychiatrist by a board of medical examiners. A psychoanalyst is typically trained by a psychoanalytic institute in a version of the Freudian method of psychoanalysis. To be accepted by a psychoanalytic institute for training, an applicant normally has an advanced degree in some form of psychological therapy. A large number of practitioners qualify both as psychoanalysts and psychiatrists, but these terms represent distinct areas of training. *See* PSYCHOANALYSIS.

Neuropsychologists are usually psychologists, who may come from an experimental or a clinical background but who must go through certification as psychologists if they are to render therapy. They treat individuals who have psychological disorders with a clear neurological etiology, such as stroke, brain trauma, and Tourette's syndrome.

Clinical practice includes individual consultation with clients, group therapy, and work in clinics or with teams of health professionals. Psychological therapists work in many settings and on problems ranging from short-term crises and substance abuse, to psychosis and major disorders such as schizophrenia and clinical depression.

Also cutting across the categories of practitioners is the theoretical orientation and therapeutic methodology of choice. While there are definite biases within each field, it is possible for a practitioner with any background to prefer behavior therapy, a humanistic approach, a Freudian (dynamic) approach, or an eclectic approach derived from these and other areas.

Nonclinical professional work in psychology includes the human-factors element, which traditionally is applied to the design of the interface between a machine and its human operator. Cognitive engineering is a branch of applied psychology that deals mainly with software and hardware computer design and other types of human-machine interfaces that are cognitively intensive. Industrial psychology also includes personnel selection and management and organizational planning and consulting.

The use of psychology in forensic matters is a natural result of the fact that much of law is based on psychology. Psychologists have been involved in jury selection, organization of evidence, evaluation of eyewitness testimony, and presentation of material in court cases. Psychiatrists and psychologists

are also called on to diagnose potential defendants for mental disorders and the ability to stand trial.

William P. Banks

Bibliography. H. Gardner, *The Mind's New Science: A History of a Cognitive Revolution*, 1987; H. Gleitman, *Psychology*, 5th ed., 2000; M. I. Posner and M. Raichle, *Images of Mind*, 1997; D. L. Schacter, *Memory: The Brain, the Mind, and the Past*, 1996.

Psychoneuroimmunology

The study of the interactions among behavioral, neural and endocrine, and immune functions. It is perhaps the most recent convergence of disciplines that has evolved to achieve a more complete understanding of adaptive processes. At one time, the immune system was considered an independent agency of defense that protected the organism against foreign material (that is, proteins that were not part of one's "self"). Indeed, the immune system is capable of considerable self-regulation. However, converging data from the behavioral and brain sciences indicate that the brain plays a critical role in the regulation or modulation of immunity. This research indicates that the nervous and immune systems, the two most complex systems that have evolved for the maintenance of homeostasis, represent an integrated mechanism for the adaptation of the individual and the species. Thus, psychoneuroimmunology emphasizes the study of the functional significance of the relationship between these systems—not in place of, but in addition to, the more traditional analysis of the mechanisms governing the functions within a single system—and the significance of these interactions for health and disease. *See* NEUROIMMUNOLOGY.

Brain–immune system interactions. Evidence for nervous system–immune system interactions exists at several different biological levels. Primary (thymus, bone marrow) and secondary (spleen, lymph nodes, gut-associated lymphoid tissues) lymphoid organs are innervated by the sympathetic nervous system, and lymphoid cells bear receptors for many hormones and neurotransmitters. These substances, secreted by the pituitary gland, are thus able to influence lymphocyte function. Moreover, lymphocytes themselves can produce neuropeptide substances. Cytokines produced by macrophages and activated lymphocytes (and by cells of the central nervous system) are critical elements in the cascade of immune responses to antigenic stimulation and also serve to energize the hypothalmic-pituitary-adrenal axis.

Thus, there are anatomical and neurochemical channels of communication that provide a structural foundation for the several observations of functional relationships between the nervous and immune systems. Lesions or electrical stimulation of the hypothalamus, for example, can alter antibody- and cell-mediated immune responses, and the eliciting of an immune response results in an increase in the firing rate of of neurons within the ventromedial hypothalamus at the time of peak antibody production. Changes in hormonal states can influence immunologic reactivity; and conversely, the immune response to antigenic challenges includes the release of cytokines which influence the neural regulation of psychophysiological processes, and is also associated with changes in circulating levels of hormones and neurotransmitter substances.

Stress and immunity. The link between behavior and immune function is suggested by experimental and clinical observations of a relationship between psychosocial factors, including stress, and susceptibility to or progression of disease processes that involve immunologic mechanisms. There are abundant data documenting an association between stressful life experiences and changes in immunologic reactivity. The death of a family member, for example, is rated highly on scales of stressful life events and, depending on gender and age, is associated with depression and an increased morbidity and mortality. Bereavement and depression are also associated with changes in some features of immunologic reactivity such as reduced lymphoproliferative responses [a general measure of the physiological status of T and B (bone marrow-derived) lymphocytes] and impaired activity of natural killer cells, which are lymphocytes capable of destroying cancer and virally infected cells without having had prior contact with the foreign material. Changes in immunity are also associated with the affective responses to other traumas such as marital separation and divorce. Other, less severe, naturally occurring stressful experiences (such as taking examinations) result in transient impairments in several parameters of immune function. *See* DISEASE; STRESS (PSYCHOLOGY).

It should be emphasized, however, that the association between stressful life experiences and disease and the association between stressful life experiences and changes in immune function do not in themselves establish a causal link between stress, immune function, and disease. In animals, a variety of stressors can, under appropriate experimental circumstances, influence a variety of immune responses in a variety of species in a variety of ways. Stressful circumstances can also alter the host's defense mechanisms, allowing an otherwise inconsequential exposure to a pathogen to develop into clinical disease. Understanding of the interactions between neuroendocrine and immune function under normal and stressful conditions, however, is incomplete. Glucocorticoids secreted by the adrenal cortex, a common endocrine feature of the stress response, are, in general, immunosuppressive, and there are numerous examples of stress-induced, adrenocortically mediated changes in immunity. However, there are numerous other observations of stress-induced changes in immunity that are independent of adrenocortical activation. It is evident that the immunologic consequences of stressful experiences involve complex neural, endocrine, and immune response interactions. Since immune responses are themselves capable of altering levels of circulating hormones and neurotransmitters, these interactions probably include complex feedback and feedforward

mechanisms as well. *See* ENDOCRINE MECHANISMS.

The direction, magnitude, and duration of stress-induced alterations of immunity are influenced by a number of factors: (1) the quality and quantity of stressful stimulation; (2) the capacity of the individual to cope effectively with stressful events; (3) the quality and quantity of immunogenic stimulation; (4) the temporal relationship between stressful stimulation and immunogenic stimulation; (5) sampling times and the particular aspect of immune function chosen for measurement; (6) the experiential history of the individual and the existing social and environmental conditions upon which stressful and immunogenic stimulation are superimposed; (7) a variety of host factors such as species, strain, age, sex, and nutritional state; and (8) interactions among these several variables.

Conditioning. Central nervous system involvement in the modulation of immunity is dramatically illustrated by the classical (Pavlovian) conditioning of the acquisition and extinction of suppressed and enhanced antibody- and cell-mediated immune responses. Using a one-trial taste-aversion conditioning situation, a distinctively flavored drinking solution (the conditioned stimulus) was paired with an injection of the immunosuppressive drug cyclophosphamide (the unconditioned stimulus). When subsequently immunized with sheep red blood cells, conditioned animals reexposed to the conditioned stimulus showed a reduced antibody response compared to nonconditioned animals and conditioned animals that were not reexposed to the conditioned stimulus. *See* CONDITIONED REFLEX.

The acquisition and the extinction (elimination of the conditioned response by exposures to the conditioned stimulus without the unconditioned stimulus) of the conditioned enhancement and suppression of both antibody- and cell-mediated immune responses—and nonimmunologically specific host defense responses as well—have been demonstrated under a variety of experimental conditions. For example, the immunological effects of stress have been conditioned, and conditioning effects using antigen itself as the unconditioned stimulus have been demonstrated. The hypothesis that conditioned alterations of immunity are merely a reflection of stress responses, notably, adrenocortical secretions, is not supported by research results. In keeping with the bidirectional nature of nervous and immune system interactions, it is also possible to condition the physiological effects elicited by the products of an activated immune system.

The biological impact of conditioned alterations in immunity is illustrated by experiments in which conditioning operations were applied in the pharmacotherapy of spontaneously developing systemic lupus erythematosus in New Zealand mice. In conditioned animals, substituting conditioned stimuli for active drug on some of the scheduled treatment days delays the onset of autoimmune disease using a cumulative amount of immunosuppressive drug that is ineffective by itself in altering the progression of disease. Similarly, reexposure to a conditioned stimulus previously paired with immunosuppressive drug treatment prolongs the survival of foreign tissue grafted onto mice. These dramatic results address the clinical implications of the behavioral component of research in psychoneuroimmunology, but they have yet to be experimentally verified in humans. *See* IMMUNOSUPPRESSION.

Again, in keeping with the reciprocal nature of the relationship between neural and endocrine and immune responses, there are indications that immune status influences behavior. For example, emotional and cognitive changes are associated with lupus in humans, and there are changes in the behavior of animals that accompany the progression of their autoimmune disease. *See* AUTOIMMUNITY.

Prospects. The mechanisms underlying the functional relationships between the nervous system and the immune system illustrated by conditioned and stressor-induced modulations of immune functions cannot yet be fully described. It is assumed that different conditioning and stressful experiences induce different patterns of neuroendocrine changes that define the milieu within which immunologic reactions occur. This milieu is influenced by neural and endocrine signals to the immune system and signals from the immune system that initiate further neural and endocrine changes—and by regulatory feedback loops between as well as within these "systems." An elaboration of the integrative nature of neural, endocrine, and immune processes and the mechanisms underlying behaviorally induced alterations of immune function is likely to have important clinical and therapeutic implications that will not be fully appreciated until more is known about the extent of these interrelationships in normal and pathophysiological states. *See* ENDOCRINE SYSTEM (VERTEBRATE); IMMUNOLOGY; NERVOUS SYSTEM (VERTEBRATE).

Robert Ader

Bibliography. R. Ader and N. Cohen, Psychoneuroimmunology: Conditioning and stress, *Annu. Rev. Psychol.*, 44:42–85, 1993; R. Ader, D. L. Felten, and N. Cohen, Interactions between the brain and the immune system. *Annu. Rev. Pharmacol. Toxicol.*, 30:561–602, 1990; R. Ader, D. L. Felten, and N. Cohen (eds.), *Psychoneuroimmunology*, 3d ed., 2000.

Psychopharmacology

A discipline that merges the subject matter of psychology, which studies cognition, emotion, and behavior, and pharmacology, which characterizes different drugs. Thus, psychopharmacology focuses on characterizing drugs that affect thinking, feeling, and action. In addition to this general concern, psychopharmacology places particular emphasis on those drugs that affect abnormalities in thought, affect, and behavior, and thus has a relationship to psychiatry.

Psychopharmacology is predominantly, but not exclusively, concerned with four major classes of drugs that are of clinical significance in controlling four

major categories of psychiatric disorder: anxiety, depression, mania, and schizophrenia.

Anxiety. Anxiety is an emotional state that can range in intensity from mild apprehension and nervousness to intense fear and even terror. It has been estimated that 2–4% of the general population suffer from an anxiety disorder at some time. Although anxiety in some form is a common experience, it can become so intense and pervasive as to be debilitating; it may therefore require psychiatric attention and treatment with an anxiolytic drug. There are three major groups of anxiolytics. Members of the first group are called propanediols; meprobamate is the most widely used. The second group is the barbiturates, of which phenobarbital is the most generally prescribed. The third group, most frequently prescribed, is the benzodiazepines, the best known of which is diazepam.

All anxiolytics have, in varying degrees, anticonvulsant and muscle relaxant properties. However, the most clinically relevant features of these drugs are two side effects characteristic of them all—sedation and toxicity. Sedation refers to the drowsiness and minor motor incoordination that all anxiolytics can induce. The propanediols and barbiturates are relatively more sedating than the benzodiazepines. The benzodiazepines are also less toxic in that fatal overdoses are much less likely to occur with them alone. However, it is important to recognize that although suicide attempts with a benzodiazepine alone do occur, such attempts much more often involve a benzodiazepine and another drug such as alcohol or a barbiturate. Benzodiazepines can significantly increase the suicide potential because they potentiate these other drugs.

A major advance in understanding the benzodiazepines was the identification of the cellular sites at which these drugs act (so-called benzodiazepine receptors). The distribution of these receptors in the brain has also been found to have a striking parallel to the distribution of the receptors for a naturally occurring substance called gamma-amino butyric acid (GABA). Furthermore, it is known that GABA has a ubiquitous inhibitory role in modulating brain function. Most importantly, it is now clear that benzodiazepines share a biochemical property in that all augment the activity of GABA.

These findings imply that the augmentation of GABA activity could be the common denominator for all forms of anxiety reduction. Some drugs, such as buspirone, however, do not augment GABA activity but do reduce anxiety, while other drugs, including valproic acid, augment GABA activity but do not reduce anxiety. Thus, although it is likely that augmented GABA activity is the mechanism that mediates the muscle relaxant and anticonvulsant actions of the benzodiazepines themselves, it is not clear that such augmentation is critical in the reduction of anxiety that benzodiazepines also produce. Indeed, other lines of evidence favor a critical role for another endogenous brain substance (5-hydroxytryptophan or its decarboxylated derivative, serotonin) in mediating the effects of the benzodiazepines on anxiety. Attempts at elucidating the mechanisms underlying anxiety reduction have also been complicated by the finding that the extreme forms of anxiety that manifest themselves as panic and phobic states can be controlled by members of another, completely separate group of drugs called antidepressants. *See* ANXIETY DISORDERS; SEROTONIN; TRANQUILIZER.

Depression. The symptoms of depression can include a sense of sadness, hopelessness, despair, and irritability, as well as suicidal thoughts and attempts, which are sometimes successful. In addition, physical symptoms such as loss of appetite, sleep disturbances, and psychomotor agitation (restlessness, hand wringing, pacing, and so on) are often associated with depression. When depression becomes so pervasive and intense that normal functioning is impaired, antidepressant medication may be indicated. It has been estimated that as much as 6% of the population will require antidepressant medication at some time in their lives.

There are two major groups of antidepressant drugs. Members of the first group are called heterocyclics because of their characteristic chemical structures. Members of the second group, which are less often prescribed, are called monoamine oxidase inhibitors because one of their biochemical effects is to inhibit an enzyme called monoamine oxidase. This action does not, however, directly account for the clinical efficacy of these drugs.

The antidepressants typically require at least several weeks of chronic administration before they become effective in alleviating depression. This contrasts with the anxiolytics, which are effective in reducing anxiety in hours and even minutes. Another difference between these two classes of drugs is that the anxiolytics are more likely to be efficacious: anxiolytics are effective in the vast majority of nonphobic, anxious patients, whereas the antidepressants are effective in only about 65–70% of depressed patients. *See* AFFECTIVE DISORDERS.

Historically, the activity of the antidepressants has been linked to the activity of norepinephrine, among other naturally occurring substances. Detailed analyses of this hypothetical linkage were made possible because of two developments: first, the cellular sites at which norepinephrine acts have been isolated; and second, because these receptors can be identified, it has been possible to map out an anatomically defined brain system dependent on norepinephrine for its activity. Furthermore, this system can be supposed to be involved in mediating the signs of depression. Chronic administration of most antidepressants decreases the sensitivity of some norepinephrine-dependent cells because of a reduction in the number of one type of norepinephrine receptors. However, not all efficacious antidepressants reduce the number of these receptors, and other drugs reduce norepinephrine activity but are not antidepressants. *See* NORADRENERGIC SYSTEM.

Mania. Manic episodes are characterized by hyperactivity, grandiosity, flight of ideas, and belligerence; affected patients appear to be euphoric, have racing thoughts, delusions of grandeur, and poor if not

self-destructive judgment. Periods of depression follow these episodes of mania in the majority of patients. The cycles of this bipolar disorder are typically interspersed among periods of normality that are, in most cases, relatively protracted.

Mania can usually be managed by chronic treatment with lithium and can be expected to be effective in 70–80% of the individuals treated. Furthermore, the period of depression that typically follows the manic episode can usually be prevented, or at least attenuated, if lithium treatment is maintained after the manic phase has subsided. Thus, lithium not only controls the period of mania itself but has a prophylactic effect on the subsequent depression. Any periods of depression that do occur can be managed by antidepressant drugs. Lithium is no longer the only drug used in the management of mania. Carbamazepine, an anticonvulsant that is used in the treatment of epilepsy, is also useful in the treatment of periods of mania.

Schizophrenia. Schizophrenia is a form of psychosis; it incorporates a broad range of symptoms that can include bizarre delusions, hallucinations, incoherence of thought processes, inappropriate affect, and grossly disorganized movements. It affects 1–2% of the population.

The symptoms of schizophrenia can be controlled, in varying degrees, by a large group of drugs called antipsychotics. Symptom management requires chronic medication and can be expected in about 80% or more of the schizophrenics treated. However, management is only partially successful in that normal functioning is not completely restored in most patients.

The antipsychotics have a broad range of side effects among which are disturbances of movement that fall into two general classes. The first class includes a loss of muscle tone, motor restlessness, and an array of symptoms very like those characteristic of Parkinson's disease. These disturbances can generally be managed by adjunctive medication with other drugs. Such control is not yet possible for the second class of movement disorder called tardive dyskinesia. Signs of this disturbance typically include involuntary movements that most often affect the tongue and facial and neck muscles but can also include the digits and trunk.

Tardive dyskinesia can develop in the course of chronic treatment with antipsychotics; as a rough rule of thumb, about 25% of those treated will show signs of the disorder. Tardive dyskinesia is relatively permanent once developed. This feature, as well as its refractoriness to treatment, places serious limits on the use of the antipsychotic drugs.

Although different antipsychotic drugs have different kinds and degrees of side effects, all share a single biochemical action: they all attenuate the activity of dopamine, a naturally occurring substance in the brain. The reduction in dopamine activity produced by the antipsychotics directly accounts for their effects on motor behavior. That is, dopamine is known to be present in high concentrations in the nigrostriatal system, an anatomically defined system of the brain that is intimately involved in the regulation of movement. It is to be expected, therefore, that disrupted dopamine activity in this system would produce disturbances of movement as, in fact, the antipsychotic drugs do. It is less clear, however, whether reduced dopamine function is also a factor in the process by which these drugs control psychotic (including schizophrenic) symptoms. *See* SCHIZOPHRENIA.

Laboratory analysis of physiological mechanisms. The mechanisms by which different drugs exert their effects on clinically significant abnormalities in behavior remain to be determined. Four major factors have impeded the kinds of analyses required: (1) The multiplicity of variables that control behavior typically cannot be simply measured. (2) Each of these variables can interact with others so that the behavioral consequences of one interaction are not the same as the consequences of another interaction. (3) It may be that different mechanisms control similar behavior so that patients, all labeled as schizophrenic, for example, may share many behaviors that are actually mediated by different underlying processes. (4) Much of the research cannot be undertaken in humans, for ethical reasons, and therefore laboratory models of various clinical entities must be developed in infrahuman species.

The development of laboratory models is itself a formidable tasks, because the disorders that must be mimicked are, in contrast to most other disorders, uniquely human. Unlike a disorder such as diabetes, there are no direct analogs of mania, depression, or schizophrenia in laboratory animals. While anxiety can be produced in the laboratory, the extent to which such anxiety truly mirrors anxiety in humans is uncertain, because the measurement of anxiety in humans is based on their verbal report.

These problems can be partially circumvented by showing that drug-induced behavioral change in the laboratory has properties parallel to those seen in the clinic. For instance, if chronic, but not acute, administration of a drug produces clinical change, then chronic, but not acute, administration of that drug should produce behavioral change in the laboratory. In addition, if clinical change is selectively induced by one group of drugs, then the laboratory behavior should also be selectively sensitive to that one group but not to other groups of drugs. For example, behavior in a laboratory animal that changes only after chronic, not acute, treatment with both lithium and carbamazepine—but not other drugs—could have promise as a model of mania. As more and more of the characteristics of clinical change are paralleled in laboratory behavior, the greater are the chances that such laboratory behavior will serve as a useful baseline for the elucidation of underlying physiological mechanism. *See* NEUROBIOLOGY. Peter L. Carlton

Bibliography. R. J. Baldessarini, *Chemotherapy in Psychiatry: Principles and Practice*, rev. ed., 1990; P. L. Carlton, *A Primer of Behavioral Pharmacology: Concepts and Principles in the Behavioral Analysis of Drug Action*, 1983; S. G. Dahl and L. F. Gram (eds.), *Clinical Pharmacology in Psychiatry*, 1989; L. F. Hollister and J. Csernansky, *Clinical*

Pharmacology of Psychotherapeutic Drugs, 3d ed., 1990; P. G. Janicak, *Principles and Practices of Clinical Psychopharmacotherapy*, 1993; R. Spiegel, *Psychopharmacology: An Introduction*, 3d ed., 1996.

Psychophysical methods

Methods for the quantitative study of the relations between physical stimulus magnitudes and the corresponding magnitudes of sensation, for example, between the physical intensity of a light and its perceived brightness or the concentration of a sugar solution and its observed sweetness. To establish these relations measurement scales are needed, not only for physical magnitudes but also for subjective magnitudes. Subjective scales are, of course, not obtained directly from observation. They are theoretical models which summarize observed relations between stimuli and responses. *See* SENSATION.

The original three psychophysical methods, first systematically worked out and published by G. Fechner in 1860, were designed to yield the data from which to construct subjective scales. Other techniques, more direct than Fechner's, are superseding his approach to psychophysical problems; his methods and variants of them are still in use, however.

The term psychophysical methods is sometimes extended to include certain scaling techniques which are most often used with, although not limited to, subjective dimensions to which there correspond no simple physical dimensions, for example, food preferences. These techniques have in common the fact that the rationale for scaling involves the amount of dispersion of repeated or multiple judgments. *See* MULTIDIMENSIONAL SCALING.

This article surveys the principal types of methods employed in relating subjective magnitudes to physical stimulus magnitudes. The first group of methods are for measuring the least stimulus, or least stimulus change, that is perceptible; the second group of methods are for establishing quantitative relationships between readily perceptible stimuli and their corresponding sensations.

Threshold methods. Two of the three psychophysical methods developed by Fechner were designed to measure the absolute threshold, defined as the minimum stimulus energy that an organism can detect, and the differential threshold, defined as the minimum detectable change in a stimulus. Both quantities had to be defined as statistical averages, because their repeated measurement under externally constant conditions yielded moment-to-moment fluctuations in the values obtained. To obtain reliable measurements for these averages, Fechner devised the method of limits (also called the method of minimal changes) and the method of constant stimuli. It is necessary in scaling to establish the units of the scale and the zero point. Fechner used the difference threshold as the unit and the absolute threshold as the zero point in scaling the magnitude of subjective sensations.

Method of limits. To measure the absolute threshold by this method, the experimenter begins with a stimulus which is too weak for the subject to detect. In successive presentations, the stimulus intensity is increased in small, equal steps, the subject reporting after each presentation whether the stimulus was perceived. The end of this ascending series is reached when the subject reports that the stimulus has been detected. The descending series is then begun, the stimulus intensity beginning at an above-threshold value and decreasing in steps until the subject signals the disappearance of the stimulus. Many such series are given. To help the subject avoid basing judgments on the number of steps occurring in a series, the initial stimulus intensity of each series is varied on successive repetitions. Frequently, the average threshold value that is obtained from ascending series differs from that obtained from descending series; the usual practice is to average the two.

In measuring the difference threshold, essentially the same procedure is involved, except that the subject now signals the relation of a comparison stimulus to a standard stimulus. The comparison stimulus begins at a value clearly less than that of the standard, and increases in steps until the subject reports the apparent equality of the two. This gives a lower difference threshold. The comparison stimulus is further increased until the subject first reports it as greater than the standard. That is the upper difference threshold. The comparable procedure is followed in a descending series, which yields a descending upper and lower threshold. After a large number of such trials, the average of each of these four threshold values is computed. The differences of each of these four averages from the value of the standard stimulus gives four values for the difference threshold. Usually, these four values are averaged.

The most common modifications of the method of limits involve the omission of descending series in order to minimize sensory adaptation, and the use of continuous rather than discrete stimulus variation. Another procedural modification is the up-and-down method. The stimulus intensity (or intensity difference) is decreased on the following trial if the subject has detected it, increased on the following trial if the subject has not. It is a more efficient method because it concentrates the testing in the region of the threshold, although there is the disadvantage that the level of the next following stimulation is predictable by the subject. It has proved notably useful in its application with the Békésy audiometer for the testing of human hearing and in the study of the visual sensitivity of animals.

Method of constant stimuli. To measure the absolute threshold, the experimenter selects a small number of stimulus values in the neighborhood of the absolute threshold (previously roughly located by informal use of the method of limits) and presents them to the subject a large number of times each, in an irregular order unknown to the subject. Each time a stimulus is presented, the subject reports the presence or absence of sensation.

At the end of the experiment the data provide the proportion of times that each stimulus resulted in a report of sensation by the subject. By interpolation one can estimate the stimulus value that has a probability of .50 of producing sensation, this value being defined as the absolute threshold.

An analogous procedure is followed in obtaining difference thresholds. One finds, for each of a set of comparison stimulus values, the proportion of stimulus presentations in which each of the values is reported to be greater than that of a standard stimulus. By interpolation one can estimate the stimulus value that has a .75 probability of being reported to be greater than the standard stimulus value. The difference between these two stimulus values has conventionally been defined as the difference threshold.

The method of constant stimuli, although a much lengthier procedure than the method of limits, has the advantage that the sequence of stimulus levels is not predictable by the subject. The lengthiness of the procedures is roughly comparable to that required by methods based on the theory of signal detection.

Theory of signal detection. Both the method of limits and the method of constant stimuli have a serious limitation. The threshold values obtained by their use are unavoidably affected by the criterion that a subject adopts in deciding when to report that a stimulus or a stimulus change has been perceived. The discriminations the subject is asked to make are difficult, and a subject who decides not to report that a stimulus or a stimulus change has been detected, until he or she is quite sure of it, will appear to have larger thresholds than would be obtained by being less conservative. Furthermore, as long as the subject is fairly consistent in criterion, there is no way to measure the extent to which it has affected the threshold values obtained. This difficulty is surmounted by the theory of signal detection, an approach to methods of measuring sensory discrimination that dispenses entirely with the concept of threshold.

From the following example of this theory it can be seen how it is possible to assess separately the detectability of a stimulus and the subject's criterion for reporting that it is detected.

Imagine a series of trials, on each of which a subject is to listen for a tone so weak that he or she can never be sure that it has been heard. The subject knows that on some trials the tone will not be sounded and that on others it will be, and after each trial he or she must say whether or not it may have been heard. On some trials the subject will be correct when saying that the tone was or was not present, but there will be some trials on which he or she says the tone was present when it was not, and some trials on which he or she says that the tone was not present when it was. Where there have been enough trials to yield reliable proportions of these kinds of trial outcome, it is possible to infer independently the detectability of the tone and the strictness of the subject's criterion in deciding whether or not the tone has been sounded.

For example, if the tone is of equal detectability for two subjects, but one employs a stricter criterion than the other, the subject with the stricter criterion will less often say that the tone was present when it was not, but will more often say that the tone was not present when it was. Thus, differences in criterion produce an inverse relationship between these two kinds of error. Differences in detectability, on the other hand, produce differences in the total proportion of correct responses. Because detectability and criterion have different effects, it is possible to use the proportions of the subject's responses on trials when a stimulus was present and on trials when a stimulus was absent to assess independently the detectability of the stimulus and the strictness of the subject's criterion. For a more detailed explanation *See* SIGNAL DETECTION THEORY.

It should be noted that the concept of a sensory threshold for a stimulus is not involved, only the detectability of a stimulus as measured by the proportion of correct reports of the presence or absence of the stimulus. This has led a number of investigators to deny that there exists such a thing as a sensory threshold, since this theory has no need of such a concept. Because both detectability and sensory threshold are statistical definitions, the question is not which definition is more valid, but which is more useful. The less laborious method of limits, with its associated concept of the sensory threshold, is often used when a high degree of precision is not required; the methods of the theory of signal detection have been coming into more and more use for accurate work when lengthier procedures are not a handicap.

Subjective metric methods. Fechner proposed to use the results of threshold measurement in developing a subjective metric or scale. He defined the difference threshold, or just noticeable difference (jnd), as the subjective unit and the absolute threshold as the zero point of the subjective scale. Thus the subjective intensity of a particular brightness of light, for example, would be specified when it was given as 100 jnds above threshold. The subjective scale so defined is not a linear function of the physical stimulus scale since jnds, though defined as subjectively equal units, are not of physically equal magnitude throughout the intensity scale. The size of the jnd is approximately proportional to physical stimulus intensity. To the extent that this relation holds, Fechner deduced that subjective intensity should be proportional to the logarithm of the stimulus intensity.

There has been considerable debate over the merits of this formulation and the assumptions involved. The chief modern objection is an empirical one, that the results obtained by Fechner's methods are different from those resulting from other, more direct techniques. A tone 40 jnds above the threshold does not sound twice as loud as a tone 20 jnds above threshold. Furthermore, the employment of the direct ratio discrimination methods that is outlined below suggests that subjective magnitude is a power function of stimulus intensity, rather than the logarithmic function that is required by Fechner's postulates. However, these same methods indicate

that on at least some nonintensive dimensions, such as length, jnds may be subjectively equal.

Direct-ratio discrimination methods. Rather than requiring of the subject merely either yes-no or ordinal judgments, these methods require the subject to make direct-ratio discriminations. For instance, he or she may be presented with a moderately loud tone, and then required, by turning a knob, to adjust the loudness of a comparison tone until it is half as loud, or twice as loud, as the first. The first case illustrates the method of fractionation, the second the method of multiplication. In the method of magnitude estimation, the subject is given a stimulus, such as the brightness of a light, to serve as a modulus with a value assigned to it, for example, 10. The task, as other lights of different intensities are presented, is to assign them numbers which shall stand in the same ratio to 10 as their brightness stands to that of the modulus. One twice as bright is given the designation 20; one half as bright is 5. In these and other similar methods, whether the subject's task is to estimate or to produce the prescribed ratio or the prescribed fraction, there are certain common characteristics. Direct-ratio assessments are obtained from the subject; there can be experimental checks on internal consistency of the results, and since the individual judgments are not of high precision, repetition is required if stable averages are to be obtained.

The empirical results obtained by the various methods are in fairly good agreement. They agree in that, to at least a first approximation, subjective magnitudes on a variety of dimensions are found to be power functions of suprathreshold stimulus intensity; that is to say, subjective magnitude is proportional to the suprathreshold stimulus magnitude raised to a power. The powers have a range from 0.3 for auditory loudness to 3.5 for subjective intensity of alternating current that is applied to the skin.

There is not yet general agreement on whether these power functions indicate a fundamental relationship between physical stimulus intensities and corresponding magnitudes of sensation, or whether they also reflect people's response characteristics when they are asked to gage their sensations. Although recordings of electrical activity in sensory nerve fibers sometimes appear to support the former interpretation, until more is learned about how the brain processes these incoming signals, this line of evidence is inconclusive.

Ratio methods are applicable to nonintensitive, or qualitative, attributes, such as auditory pitch and hue of colors, and their application is not necessarily limited to stimuli from simple physical dimensions. For example, the methods have been used to develop a subjective scale of the seriousness of various crimes.

Direct-matching methods. The subject is not required to produce or assess the ratio of one subjective magnitude to another, but only to adjust a comparison stimulus until some attribute appears to match that of a standard stimulus. For example, the subject might be asked to adjust the physical intensities of tones of various frequencies until their loudness matched that of a 1000-Hz tone of fixed intensity. The result would be an equal-loudness contour, showing the intensities to which tones of various frequencies must be set to produce sensations of equal loudness. These data are of use in acoustics. *See* LOUDNESS.

Method of average error. This method, one of the three methods devised by Fechner, is a special application of direct-matching methods to cases in which the point of interest is in discrepancies between perception and stimulation. The subject adjusts a comparison stimulus to match a standard stimulus; the average of a number of such settings gives the point of subjective equality, and the difference between this point and the standard stimulus is the average error. Two illustrative uses of the method are the measurement of accuracy of distance perception and the measurement of the magnitude of so-called optical illusions. *See* HEARING (HUMAN); PSYCHOLOGY.

John F. Hahn

Bibliography. L. R. Aiken, *Psychological Testing and Assessment*, 10th ed., 1999; N. Arthur and R. R. Fay (eds.), *Human Psychophysics*, 1993; J. R. Barclay, *Psychological Assessment: A Theory Systems Approach*, 1991.

Psychosis

Any disorder of higher mental processes of such severity that judgments pertaining to the reality of external events are significantly impaired. A wide range of psychiatric, medical, and neurological conditions can bring about a psychotic state. They include schizophrenia, mania, depression, ingestion of drugs, withdrawal from drugs, liver or kidney failure, endocrine disorders, metabolic disorders, and Alzheimer's disease, epilepsy, and other neurologic dysfunctions. Not all psychotic states are caused by disease or drugs. The dreams of normal sleep are a form of psychosis; while dreaming, a reality can be seen and heard which, at the moment, is not known to be false but which is roused from upon awakening.

Psychotic alterations of beliefs are called delusions. For manic patients, a typical delusion is that they possess omnipotent powers; for schizophrenics, it is often a belief that their minds are controlled by external forces or that they can project their thoughts into other people's minds. Patients with schizophrenia, Alzheimer's disease, or drug-induced psychosis may have paranoid delusions, such as a belief that other people or evil forces are attempting to do them harm. Psychotic alterations of perception are referred to as hallucinations. Addicts who are undergoing withdrawal from drugs may suffer from tactile hallucinations, for example, a sensation that bugs are crawling on them. Schizophrenics often hear voices of persons who are not actually present. A depressed person may "see" the silhouette of a lost loved one. A person with epilepsy may hallucinate certain smells. Psychotic states that are due to alcoholism, metabolic diseases, or other medical conditions are frequently accompanied by general mental confusion. On the other hand, psychiatric

illnesses and drugs, such as lysergic acid diethylamide (LSD) and phencyclidine (PCP), can produce hallucinations and delusions in the absence of general confusion. *See* PSYCHOTOMIMETIC DRUGS.

Patients are often very distressed by their psychotic symptoms. Unfortunately, very few of those symptoms are unique to a particular illness, which can make proper diagnosis difficult and challenging. Correct diagnosis, however, is critical so that appropriate treatment can be provided. *See* ADDICTIVE DISORDERS; AFFECTIVE DISORDERS; ALZHEIMER'S DISEASE; PARANOIA; SCHIZOPHRENIA. Ralph E. Hoffman

Bibliography. P. Buckley (ed.), *Essential Papers on Psychosis*, 1988; R. Marcus, *Psychosis and Near Psychosis: Ego Function, Symbol, Structure, and Treatment*, 1993; R. E. Hoffman and T. H. McGlashan, Speech, thought, and perceptual disorders, in A. Tasman, J. Kay, and J. A. Lieberman (eds.), *Psychiatry*, vol. II, Saunders, 1996.

Psychosomatic disorders

Disorders characterized by physical symptoms that originate, at least in part, from emotional factors. The *Diagnostic and Statistical Manual of the American Psychiatric Association* (DSM) *IV* now assigns such conditions to two diagnostic categories: somatoform disorders and psychological factors affecting physical condition. Somatoform disorders are characterized by physical symptoms that are suggestive but not fully explained by a diagnosable medical disease. The second category, psychological factors affecting physical condition, refers to the presence of a general medical condition that is negatively affected (for example, prolonged or exacerbated) by psychological factors.

The fact that emotional factors play a role in the pathogenesis of physical disease was known as early as the time of Hippocrates. Sigmund Freud demonstrated that physical symptoms such as paralysis and paresthesia could be caused by emotional conflicts (conversion disorder or hysterical conversion, now known as a type of somatoform disorder). F. Alexander distinguished between conversion disorder and psychophysiologic disorders, now considered to be psychological factors affecting physical condition. His specificity theory also proposed that specific psychological conflicts caused specific medical diseases, such as peptic ulcer and hypertension, that he considered to be "psychosomatic."

Evolution of concept. The explosive developments in the fields of molecular biology, genetics, and computer-assisted large-scale epidemiologic studies fueled an expansion of the concept of psychosomatic medicine during the past several decades. Stress research, in humans and animals, has shown that general stress (for example, common life changes such as marriage and divorce) as well as specific stresses (such as bereavement and combat) can have a profound impact on a number of organ systems (including the immune system, the cardiovascular system, and the brain) and the behavior of the organism, including its coping styles. It has been shown that emotional factors, including stress, may play an important role in diseases such as cancer and various infections which previously had not been considered psychosomatic. On the other hand, the role of genetic factors and nonspecific stress is now considered to be as great as that of specific psychological conflicts. The development of a disease state (pathogenesis) must be considered in terms of genetics, physical and psychological development, and the interaction of the individual with the physical and social environment. Thus, the model of psychosomatic medicine is now an interactional systems model that encompasses all disease (see **illus.**).

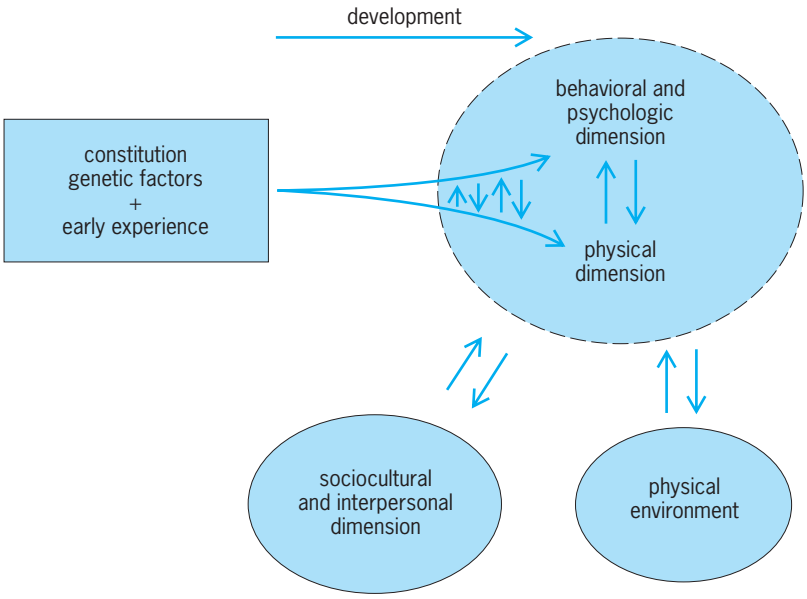

Comprehensive model for all diseases. (*After H. Leigh and M. F. Reiser, Major trends in psychosomatic medicine, Ann. Intern. Med., 87:233–239, 1977*)

Developments in medical sociology and psychology have also elucidated the role of psychological and psychosocial factors not only in pathogenesis but also in help-seeking behavior, the course of illness, compliance to a medical regimen, and convalescence from practically all diseases. It has been determined, for example, that stress and anxiety tend to increase help-seeking behavior and that depression tends to worsen the prognosis of any procedure and delay recovery from disease. In modern medicine, the psychosomatic approach involves investigating the interaction of biological, psychological, and social factors in each person, regardless of the diagnosis. This integrative approach is referred to as the biopsychosocial model for the patient and disease.

Psychological factors affecting physical condition. This category encompasses disorders characterized by the existence of a general medical condition that is negatively affected by psychological factors, such as stress. For example, stress can delay recovery from or exacerbate the symptoms of a general medical condition or even cause additional health problems. There is nothing mysterious about the idea that the mind changes body states: voluntary movements are

certainly psychosomatic. For example, when one thinks about sour apples, salivation occurs as a classically conditioned response. Psychological states influence body organs through a combination of three interrelated mechanisms: neural, hormonal, and immunologic. Voluntary movements (for example, clenching the teeth) are mediated through the motor neurons by the conscious command of the brain. In stress, clenching of the teeth, mediated by the same motor neurons, may also occur, but the act may not be voluntary and conscious. Stress usually causes an activation of the sympathetic nervous system and the hypothalamo-pituitary-adrenal axis followed by a decrease in immunocompetence. Immune mechanisms may be suppressed in part through corticosteroid activation, but a decrease in T-lymphocyte activity in stress (for example, following bereavement) may not be mediated by hormones.

The voluntary nervous system as well as the autonomic nervous system are capable of learning through operant conditioning (reward and punishment), and the immune system can be classically conditioned. Individual specific, but inadvertent, conditioning of specific conflict or stress to specific bodily malfunction may be an important psychosomatic mechanism. *See* CONDITIONED REFLEX; NEUROIMMUNOLOGY; STRESS (PSYCHOLOGY).

Somatoform disorders. Included in this category are disorders whose symptoms are suggestive of but not fully explained by a physical disease. Somatoform disorders differ from factitious disorders and malingering in that the symptoms are not under intentional (voluntary) control. There are six types of somatoform disorders: conversion disorder, somatization disorder, pain disorder, hypochondriasis, body dysmorphic disorder, and undifferentiated somatoform disorder.

Conversion disorder. Conversion disorders are characterized by physical symptoms referrable to the somatosensory nervous system or special sensory organs that cannot be explained on the basis of a medical or neurologic disease and that are caused by psychological factors. Common symptoms include paralysis, blindness, ataxia (lack of muscular coordination, especially in walking), aphonia (loss of voice and power of speech), and numbness of the feet (stocking anesthesia). The symptoms may represent a psychological conflict or may be a form of body-language communication.

Somatization disorder. In somatization disorder (also known as Briquet's syndrome), the patient recurrently complains of multiple somatic symptoms that are referrable to practically every organ system in the body. Upon medical investigation, however, the symptoms turn out not to be a diagnosable physical disease. This chronic, fluctuating disorder begins before age 30 and has a familial tendency; both genetic and environmental factors seem to play a role. The chronicity and multiplicity of symptoms in somatization disorder distinguish it from conversion disorder. The symptoms do not usually symbolize psychological conflicts but may represent general dysphoria and distorted illness behavior, such as a psychological addiction to the sick role. There is no definitive treatment; patients should be managed by one physician who coordinates all diagnostic and treatment plans and provides ongoing support and follow-up without unnecessary invasive procedures.

Pain disorder. When pain is the prominent clinical symptom and psychological factors are considered to play an important role in the onset, severity, exacerbation, or maintenance of the pain, pain disorder may be diagnosed. If pain is part of somatization disorder, depression, or another major psychiatric disorder, a separate diagnosis is not appropriate. *See* PAIN.

Hypochondriasis. Hypochondriasis is preoccupation with the idea of having a serious medical disease based on a misinterpretation of one or more bodily symptoms or signs. Such preoccupation continues in spite of a negative reasonable medical workup, and causes significant distress or impairment in functioning. There may be an amplification of somatic sensations in hypochondriasis. Hypochondriasis is not diagnosed if another major psychiatric disorder, such as anxiety disorder or major depression, accounts for the preoccupation.

Body dysmorphic disorder. Otherwise known as dysmorphophobia, this condition denotes a preoccupation with a defect in appearance that may be either imagined or greatly exaggerated, and that causes significant distress or impairment in function. If such preoccupation occurs in the presence of another major psychiatric disorder, such as schizophrenia or anorexia nervosa, it is subsumed under that diagnosis.

Therapeutic approaches. The management of psychological factors affecting physical condition should exemplify the biopsychosocial model, that is, effective treatment of the physical condition as well as psychotherapeutic and psychopharmacologic interventions as needed to ameliorate the psychological factors, including stress management and relaxation techniques.

The treatment of choice for conversion disorder is psychotherapy designed to provide understanding and resolution of the psychological issues underlying the condition. Physical or speech therapy may be indicated to facilitate recovery of function. Supportive psychotherapy and judicious psychopharmacologic treatment may be indicated for pain disorder, hypochondriasis, and body dysmorphic disorder. Antidepressants may be effective in treating pain disorder, and supportive psychotherapy, relaxation techniques, and antianxiety medications may be useful for patients with various somatoform disorders. *See* PSYCHOPHARMACOLOGY; PSYCHOTHERAPY. Hoyle Leigh

Bibliography. R. Ader, D. Felton, and N. Cohen (eds.), *Psychoneuroimmunology*, Academic Press, San Diego, 2001; F. Alexander, *Psychosomatic Medicine: Its Principles and Applications*, W. W. Norton, New York, 1987; *Diagnostic and Statistical Manual Of Mental Disorders*, 4th ed. (DSM IV), American Psychiatric Association, 1994;

H. Leigh (ed.), *Biopsychosocial Approaches in Primary Care: State of the Art and Challenges for the 21st Century*, Plenum Press, New York, 1997; H. Leigh and M. F. Reiser, Major trends in psychosomatic medicine, *Ann. Intern. Med.*, 87:233–239, 1977; H. Leigh and M. F. Reiser, *The Patient: Biological, Psychological, and Social Dimensions of Medical Practice*, 3d ed., Plenum Medical Book, New York, 1992.

Psychotherapy

Any treatment or therapy that is primarily psychological in nature. In recent years, counseling also has been included in this categorization. The number of different forms of psychotherapy has increased tremendously, particularly in the past three or four decades. Once the main forms of psychotherapy were Freud's psychoanalysis and those of his former disciples. Now there are at least 400 different forms of psychotherapy.

Types. In addition to designating psychotherapies in terms of their theoretical or procedural differences, they can be classified in other ways—for example, individual or group therapy, marital therapy, adult or child therapy, crisis intervention, time-limited or brief therapy, long-term or time-unlimited therapy, and manualized therapy. Thus, one must be cautious when discussing psychotherapy generally.

Psychodynamic therapies. Historically, psychoanalysis has played an important role in the growth and development of psychotherapy. It was created by Sigmund Freud, who developed his views over a period of years. Central to his theories was the importance of unconscious conflicts in producing the symptoms and defenses of the patient. Thus, the goal of therapy is to gradually overcome the patient's defenses and to help the patient attain insight into the repressed conflicts which are the source of difficulty. Since patients resist these attempts to uncover and bring to consciousness the painful repressed material, therapy must proceed slowly. Consequently, psychoanalysis is a long-term therapy requiring several years for completion and almost daily visits.

Since Freud's time, there have been important modifications associated with former disciples such as Alfred Adler and Carl Jung, as well as by a number of others in more recent decades. Self psychology and ego psychology are among more recent emphases. However, the popularity of psychoanalysis has waned for a number of reasons, including cost, length of treatment, selectivity of patients, and evidence of effectiveness. Still, several briefer forms of psychodynamic therapy derived from psychoanalysis have emerged and are clearly in tune with recent emphases on brevity, cost, and health maintenance organizations. *See* PSYCHOANALYSIS.

Experiential therapies. A number of related therapies are included in this group. Probably best known was the patient-centered therapy of Carl Rogers appearing in the 1940s. Like many of the therapies that have appeared since then, it was developed because of dissatisfaction with analytic therapy. In Rogers' therapy, a major emphasis is placed on the ability of the patient to change when the therapist is empathic and genuine and conveys nonpossessive warmth. The therapist is nondirective in the interaction with the patient and attempts to facilitate the growth potential of the patient. Rogers also was a pioneer in using recorded therapy sessions for training and research.

Other therapeutic approaches considered as experiential include the Gestalt therapy of Perls, existential approaches, and transpersonal approaches. The facilitation of experiencing is emphasized as the basic therapeutic task, and the therapeutic relationship is viewed as a significant potentially curative factor. The uniqueness of the patient and his or her potential for personality growth is also emphasized.

Cognitive, behavioral, and interpersonal therapies. Another group of therapies that have received increased attention in recent years are the behavioral, cognitive-behavioral, and interpersonal therapies. Behavioral therapies, which developed from behaviorism in the field of psychology, emphasized theories and procedures that differ markedly from psychodynamic and experiential approaches. The therapists play a more directive role in therapy. The emphasis is on changing the patient's behavior, using positive reinforcement, and increasing self-efficacy. They also emphasized research. More recently, cognitive therapies such as those of A. T. Beck have tended to be combined with behavioral emphases. The cognitive-behavioral therapies have focused on changing dysfunctional attitudes into more realistic and positive ones and providing new information-processing skills. Cognitive-behavioral therapies have become quite popular in the last two decades or so for several reasons. They tend to be brief forms of therapy, rather detailed manuals for conducting therapies for depression and anxiety disorders have been produced, and an impressive amount of research has been published to support the claims made for the effectiveness of these therapies. *See* COGNITION.

Interpersonal therapy is somewhat different from the behavioral and cognitive-behavioral therapies, and most of its developments have occurred in work with depressed patients. The goal of interpersonal therapy (a brief form of therapy) is centered on increasing the quality of the patient's interpersonal interactions. Emphasis is placed on enhancing the patient's ability to cope with stresses, improving interpersonal communications, increasing morale, and helping the patient deal with the effects of the depressive disorder. Like the cognitive-behavioral approaches, a detailed manual was prepared by Klerman, Weissman, and colleagues, and interpersonal therapy was compared with Beck's cognitive-behavioral therapy in a large-scale multisite study of depression. At the end of treatment, the results were roughly comparable and positive. *See* PERSONALITY THEORY.

Eclectic and integrative therapies. Although there are a large number of psychotherapies, several

surveys of different groups of psychotherapists have indicated that the largest number consider themselves to be eclectics. They do not adhere strictly to one theoretical orientation or school but use any procedures that they believe will be helpful for the individual patient. Rather than having the patient adapt to a specific form of therapy, the therapeutic procedures are adapted to the patient. More recently, eclecticism has been linked with the development of a movement for integration in psychotherapy. The emphasis in this new development is on openness to the views of other approaches, a less doctrinaire approach to psychotherapy, and an attempt to integrate two or more different theoretical views or systems of psychotherapy. There is thus a greater emphasis on theoretical integration than is generally true for eclecticism. This also provides more opportunity for research aimed at evaluating therapeutic factors or techniques of different approaches.

Group, family, and marital therapy. Most psychotherapy is conducted on a one-to-one basis—one therapist for one patient—and the confidentiality of these sessions is extremely important. However, there are other instances where more than one patient is involved because of particular goals. Marital therapy, for example, involves both marital partners, as the most efficacious approach to such problems. Therapeutic work with families has also developed from the recognition of the importance of the family when working with individuals who have psychological difficulties. This emphasis has been noted particularly in the case of children and severely disturbed individuals. A variety of group therapies have also been developed and used in different ways. Outpatient groups have been used for smoking cessation, weight loss, binge eating, and similar problems as well as for what were traditionally viewed as psychoneurotic problems. Inpatient group therapy was frequently employed in mental hospital settings along with physical, recreational, and pharmacological therapies.

Therapeutic factors. With so many different forms of psychotherapy available, one may wonder what factors or variables make for positive therapeutic change, particularly when the theoretical emphases appear so different. For example, in psychodynamic therapies the relationship to the therapist and the interpretation of it by the therapist is stressed, as well as the recovery of repressed negative thoughts and feelings. The ability of the patient to secure insight into the causes of the patient's feelings and behavior is also very important. In patient-centered therapy, the therapist's empathy, warmth, and genuineness are the therapeutic variables stressed. In the behavioral therapies, other factors are considered important, including learning to relax, facing the situations that produce anxiety, receiving positive reinforcement and feedback for desired behaviors, and learning self-control. Cognitive therapies emphasize replacing negative cognitions with positive ones, and interpersonal therapists pay particular attention to the patient's interpersonal problems and behaviors in the patient's current life.

Are the therapies really so different? Most of them have not been systematically evaluated, and in a number of the comparative studies the differences in outcome have generally been small. It is important to note that the participants in psychotherapy, the client and the therapist, play a significant role in the types of outcome. There has been an increasing belief that factors common to most psychotherapies may be of some importance. The expectations and motivations of the patient, the type of patient, and the type of therapeutic relationship that develops have been considered possible common factors. Patients who are highly motivated to cooperate with the therapist, whose problems are not severe, and with whom a positive relationship can be developed are most likely to secure positive results regardless of the form of therapy. Furthermore, in most forms of therapy the therapist is likely to provide explanations, information, suggestions, encouragement, and opportunities for emotional expression. *See* PSYCHOSIS; PSYCHOSIS.

Effectiveness. As the popularity of psychotherapy has increased, a number of events have led to increased research evaluating its effectiveness. Better means of appraisal and refined research designs have contributed. Most recently, a design has been used in which carefully evaluated research subjects are assigned randomly to the groups being compared. One form of therapy can be compared to a no-treatment control group, to a group receiving medication, or to another form of therapy. Patients with specific diagnoses are compared during therapy, at termination, and at posttherapy intervals. In general, the results have favored psychotherapy over no-treatment controls. The therapies have tended to be brief and time-limited, and few differences have been secured between pharmacotherapy and psychotherapy, although these latter findings are more complex. In some instances, such as depression, an antidepressant works more quickly and appears to act somewhat better with the most severe cases. However, there may be more "drop-outs" from treatment and possible side effects with medication. There has also been research on the combined use of medication and psychotherapy. In general, where two highly successful treatments are combined in cases with depressive or anxiety disorders, there appears to be little gain in effectiveness. However, in several studies of hospitalized patients with schizophrenia where individual psychotherapy has been ineffective, a combination of psychotherapy and medication has produced better results than medication alone. The combination of certain types of family therapy and psychotropic medication for patients diagnosed with schizophrenia has proven more effective than the use of either therapy alone. Thus, increased knowledge of all treatment modalities as well as types of psychopathology may gradually lead to improved outcomes. Sol L. Garfield

Bibliography. D. H. Baucom and N. Epstein, *Cognitive-Behavioral Marital Therapy*, 1990; B. D. Beitman and G. L. Klerman (eds.), *Integrating*

Pharmacotherapy and Psychotherapy, 1991; A. E. Bergin and S. L. Garfield (eds.), *Handbook of Psychotherapy and Behavior Change*, 4th ed., 1994; G. L. Klerman and M. M. Weissman (eds.), *New Applications of Interpersonal Psychotherapy*, 1993; S. L. Garfield, *The Practice of Brief Psychotherapy*, 2d ed., 1998; S. L. Garfield, *Psychotherapy, An Eclectic-Integrative Approach*, 2d ed., 1995.

Psychotomimetic drugs

A class of drugs, also called psychedelic or hallucinogenic drugs, that induce transient states of altered perception resembling or mimicking the symptoms of psychosis and are characterized by profound alterations in mood. The term psychotomimetic was originally coined to reflect the similarities between the mind-altering effects of certain psychoactive drugs and some of the symptoms of schizophrenia. *See* PSYCHOSIS; SCHIZOPHRENIA.

Effects. The net effect of any psychoactive drug is a complex interaction between the personality and expectations of the individual, the pharmacological effects of the drug, and the setting in which the drug is used. Psychotomimetic drugs produce a range of perceptual alterations and distortions. Sensory inputs are heightened or diminished (colors, textures, sounds, tastes, and touch become more or less intense). Sensations may assume compelling significance, to the point that an individual becomes completely immersed in an experience and disregards more salient stimuli. Individuals may experience illusions (distortions in the perceived nature or meaning of real objects), such as seeing a grotesque distortion of a familiar face, or frank hallucinations (perceptions in the absence of any obvious stimuli). Typically the hallucinations produced by psychotomimetic drugs are in the visual domain, and individuals realize that the perceptions are false. Synesthesias may occur, wherein individuals perceive stimuli from one sensory modality in another modality—for example, "seeing" sounds, or "hearing" colors. Often the body or its parts may appear larger, smaller, heavier, or lighter than normal. Time perception may be distorted, with time seeming to pass slower or faster. Concentration and attention can also be impaired, and subjects may sometimes focus on seemingly trivial details of an experience. Inability to "filter" out trivial external or internal stimuli may result in the feeling of being overwhelmed by sensations. Individuals may feel detached from themselves or the immediate environment (depersonalization). This may be described as watching oneself as a spectator or observer. Similarly, individuals may experience their surroundings as unreal or dreamlike (derealization). At times, individuals may describe having "greater understanding" or enlightenment of events, their own existence, or the motivation of others while under the influence of psychotomimetic drugs. Rarely,} individuals may develop false fixed beliefs or inferences that are not consistent with reality (delusions), such as the paranoid belief that people in their environment are trying to harm them. *See* HALLUCINATION; PARANOIA.

These effects on perceptions and thinking are invariably accompanied by changes in the individual's mood. Thus, individuals may experience intense and often rapidly shifting mood states ranging from mild apprehension to panic or severe depression to elation, or they may experience concurrent emotions (such as sadness and joy) which are not ordinarily experienced simultaneously.

Classes of drugs. Psychotomimetic drugs can be classified loosely into several groups based on chemical structure and their profile of effects.

Classical serotonergic hallucinogens. These drugs include mescaline (**1**), lysergic acid diethylamide (LSD) (**2**),

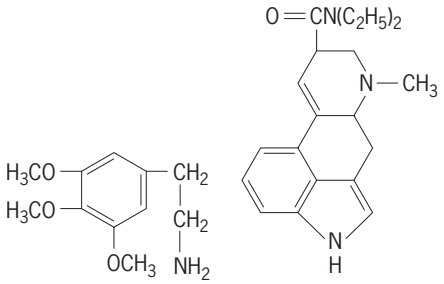

(**1**) (**2**)

and dimethyltryptamine (DMT), as well as the "magic mushroom" (*Psilocybe mexicana*) derivatives psilocybin and psilobin. Mescaline, used in highly structured Native American tribal religious rituals, is derived from peyote buttons, which are the dried tops of a cactus found in the southwestern United States.

Albert Hoffman synthesized LSD-25 (the twenty-fifth in a series of synthesized derivatives) in 1938 with compounds derived from ergot (a fungus, *Claviceps purpurea*, which infests such grasses as rye). He discovered its psychotomimetic properties in 1943. *See* ERGOT AND ERGOTISM.

Serotonergic hallucinogens produce their effects by partial stimulation of the 2_A subtype of serotonin receptors in the brain. While these drugs differ in the onset, duration, and intensity of effects, their acute behavioral effects are quite similar. LSD, the most potent drug in this class, initially produces somatic symptoms including paraesthesias (tingling of the extremities), dizziness, weakness, and tremor. These are followed by perceptual alterations including blurring of vision, illusions, visual hallucinations, less discriminant hearing, and altered time perception. Dreamlike imagery may develop when eyes are closed. Synesthesia sometimes occurs. Abstract geometric patterns are frequently seen. These effects are accompanied by changes in emotions: several emotions may occur at once, or emotions may become intensified and change rapidly. Depersonalization (the feeling of being separated from oneself) and difficulty separating self from environmental stimuli may occur. Cognitive disturbances are generally mild and rare, and include impaired memory and

difficulty with concentration. Individuals may report a capacity for greater introspection and the revelation of "new insights" about themselves, leading to the reputation of these drugs as "mind revealing" or "psychedelic." A "bad trip" is a panic reaction occurring during the drug state (usually manageable with sensible reassurance, or sedatives). "Flashbacks" can occur in a minority of individuals, and these vivid momentary reactivations of the drug state in the absence of drug can be frightening. *See* SEROTONIN.

Tolerance to the physical effects of these drugs develops rapidly, and decreases rapidly with cessation of use. However, the drugs are not associated with dependence or withdrawal and have low reinforcement liability relative to stimulants. Unlike other drugs of abuse, most do not reinforce self-administration in animals.

Amphetamine. Amphetamine and related drugs are stimulants, increasing the release of dopamine, norepinephrine, and serotonin in addition to blocking the reuptake of these neurotransmitters. These drugs can be ingested, smoked, or injected. They produce a "rush" (an orgasmlike state) followed by a state of euphoria, alertness, and a feeling of increased energy and self-confidence. Relative to classical serotonergic hallucinogens, perceptual alterations are not as prominent with amphetamine. Some individuals start to binge (bursts of repeated consumption spanning hours to a few days), and this pattern may be associated with symptoms similar to those of paranoid schizophrenia. Bingeing is followed by a "crash" phase or withdrawal syndrome characterized by depression, agitation, anxiety, and low energy, followed by insomnia and fatigue, and culminating in increased appetite and hypersomnolence (excessive sleepiness).

Amphetamine and other dopaminergic stimulants are associated with tolerance (decreased effect upon repeated use), intense psychological dependence, and withdrawal, and may be associated with sensitization (increased effect upon repeated use). *See* DOPAMINE; NORADRENERGIC SYSTEM.

Substituted amphetamine analogues. These drugs include methylenedioxymethamphetamine (MDMA) or "ecstasy," *N*-ethyl-3,4-methylenedioxyamphetamine (MDEA), 2,5-dimethoxy-4-methylamphetamine (DOM or "STP"), methylenedioxyethylamphetamine (MDE or "EVE"), and methylenedioxyamphetamine (MDA). They were designed from amphetamine, but share properties of both amphetamine (stimulant) and mescaline (hallucinogen). MDMA was developed in 1914 as a parent compound for the synthesis of other drugs. Though intranasal use has been reported, these drugs are typically ingested. The psychoactive properties of these drugs appear to be related to some combination of effects on both serotonin and dopamine release, and partial agonist effects at 5-HT_{2a} receptors. The psychoactive effects begin with a very mild rush followed by other stimulantlike effects including euphoria, increased self-confidence, and decreased appetite. Like serotonergic hallucinogens, these drugs produce perceptual alterations (heightened sensory perception, illusions, hallucinations), difficulty in concentrating, and changes in mood (calmness, or anxiety and panic). Other effects include ataxia (instability of gait), insomnia, jaw clenching (bruxism), and muscle rigidity. MDMA reportedly makes people feel more "open" or more connected to others. Its use became illegal in the United States in 1985 after reports of neurotoxicity, and suggestions that MDMA and MDA may cause irreversible damage to brain serotonin neurons.

Cannabinoids. One of the oldest and most widely used psychotomimetic drugs is derived from the flowers and leaves of the *Cannabis sativa* plant. Delta-9-tetrahydrocannabinol (THC) is the principal active ingredient of cannabis, but there are more than 80 other cannabinoids present in cannabis which may also contribute to its effects. While usually smoked, cannabis is sometimes consumed in brownies, cakes, and sweetened milk. Cannabinoids produce their psychotomimetic effects by actions at brain cannabinoid receptors. The psychotomimetic effects of cannabis appear rapidly (5–10 min) after smoking but slowly (up to 1–2 h) when eaten. The effects of cannabis include changes in perception, cognition, and mood. The cognitive effects include impairment of short-term memory and free recall, difficulty with attention and concentration, difficulty keeping track of thoughts, rapid flow of ideas, and poor judgement. Perceptual alterations may include a heightened sensitivity to all sensory modalities (colors and sounds seem brighter and richer), distortion of time, body image, and space (objects seem or the body feels heavier or longer), depersonalization, derealization, illusions, hallucinations, and synesthesias. Individuals may report a sense of great insight or understanding of life events. Paranoia (irrational suspiciousness or distrustfulness) is not infrequent and frank psychosis, while rare, is possible. The initial effect on mood is euphoria, often accompanied by uncontrollable laughter and a feeling of hilarity. This is followed by feeling calm, relaxed, mellow, passive, apathetic, and drowsy. However, anxiety and panic attacklike symptoms are also possible. "Bad trips" and flashbacks can also occur.

There is some evidence suggesting tolerance to some of the effects of cannabis, and increasing evidence suggesting the presence of a mild and subtle (though definite) cannabis withdrawal syndrome.

NMDA antagonists. These agents were originally developed as "dissociative" anesthetics. Rather than complete loss of consciousness, their usual dosage produces an inattentiveness to pain, an amnesia for ongoing events, and a kind of stupor in which the eyes are open but "not seeing." The class includes 1-[1-phenylcyclohexyl]piperidine (phencyclidine, PCP or "angel dust"), ketamine ("special K"), and 1-(1-2-thienyl-cyclohexyl)piperidine (TCP). They produce their effects by noncompetitively blocking the NMDA (N-methyl-D-aspartate) subtype of glutamate receptor. PCP can be snorted, smoked, consumed orally, or injected intravenously. The

effects occur within minutes of smoking the drug and include initial euphoria, followed by perceptual alterations (feelings of detachment, distortion of body image, heightened sensitivity to sounds and images, illusions, and hallucinations) and cognitive deficits (disruption of memory, difficulty in planning, impairment of judgement, difficulty with abstract thinking, and difficulty concentrating and maintaining a train of thought). Other effects include numbness, paresthesias, unsteady gait, and slurred speech. Some of the effects of PCP and ketamine resemble some of the symptoms of schizophrenia, and in rare instances PCP use may be associated with a psychotic syndrome that outlasts intoxication.

There is little evidence to suggest that the abuse of these drugs by humans is associated with tolerance, dependence, or withdrawal. *See* PSYCHOPHARMACOLOGY. Cyril D'Souza

Bibliography. D. J. Balfour, *Psychotropic Drugs of Abuse*, 1990; D. S. Charney, E. Nestler, and B. S. Bunney (eds.), *Neurobiology of Mental Illness*, 1999; J. R. Cooper, F. E. Bloom, and R. H. Roth, *Biochemical Basis of Neuropharmacology*, 2002; K. L. Davis, et al. (eds.), *Neuropsychopharmacology: The Fifth Generation of Progress*, 2002; B. L. Jacobs (ed.), *Hallucinogens: Neurochemical, Behavioral and Clinical Perspectives*, 1984; B. L. Jacobs, How hallucinogenic drugs work, *Amer. Scientist*, 75:386–392, 1987; M. Lader (ed.), *Psychopharmacology of Addiction*, 1994.

Psychrometer

An instrument consisting of two thermometers which is used in the measurement of the moisture content of air or other gases. The bulb or sensing area of one of the thermometers either is covered by a thin piece of clean muslin cloth wetted uniformly with distilled water or is otherwise coated with a film of distilled water. The temperatures of both the bulb and the air contacting the bulb are lowered by the evaporation which takes place when unsaturated air moves past the wetted bulb. An equilibrium temperature, termed the wet-bulb temperature (T_W), will be reached; it closely approaches the lowest temperature to which air can be cooled by the evaporation of water into that air. The water-vapor content of the air surrounding the wet bulb can be determined from this wet-bulb temperature and from the air temperature measured by the thermometer with the dry bulb (T_D) by using an expression of the form $e = e_{SW} - aP\,(T_D - T_W)$. Here e is the water-vapor pressure of the air, e_{SW} is the saturation water-vapor pressure at the wet-bulb temperature, P is atmospheric pressure, and a is the psychrometric constant, which depends upon properties of air and water, as well as on speed of ventilation of air passing the wet bulb.

Other moisture parameters, such as relative humidity and dew-point temperature, can be conveniently evaluated from the wet- and dry-bulb measurements by means of psychrometric tables.

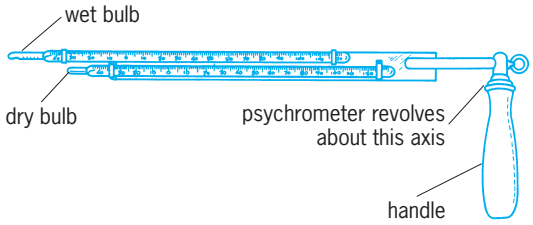

Sling psychrometer. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

These tables are accurate for a ventilation speed of 4.6 m/s (15 ft/s) when applied to measurements made using standard U.S. Weather Bureau thermometers. Other sensors may require different ventilation rates.

The common sling psychrometer consists of two mercury-in-glass thermometers which are mounted side by side on a metal or plastic frame (see **illus.**). The frame is attached to a handle by means of a chain or a suitable bearing in such a manner that the frame can be freely rotated about the handle. A clean muslin patch is tightly mounted over the bulb of one of the thermometers and is thoroughly wetted with distilled water prior to an observation. To make a measurement, the frame is rotated about the handle at a rate of approximately four revolutions per second, and this rotation is continued until the wet-bulb temperature is decreased to a constant value.

Another form of the psychrometer consists of two temperature-sensing elements, one of which is wetted, mounted in an air duct. A suction fan is used to aspirate air past the sensing elements at a rate sufficient to assure that the wet-bulb temperature will be reached for the particular thermal sensor employed in the instrument. This form of the instrument is used frequently at permanent weather stations.

The error associated with the measurement of atmospheric moisture by means of a psychrometer originates from a number of sources. These can be classified as errors due to (1) inherent deficiencies in temperature sensing, (2) improper ventilation, (3) improper muslin preparation or bulb coverage, (4) impure water, (5) extraneous heat flow to the bulb from radiative or conductive sources, and (6) an uncertainty in the phase of the water when the wet-bulb temperature is below the freezing point.

In general, the accuracy of the instrument is greatest at temperatures above the freezing point of water and decreases substantially at low temperatures. For example, normal calibration errors in the standard sling psychrometer can lead to a measurement error of 25% in air of nominal 50% relative humidity when the dry-bulb temperature is 5°F (−15°C). The corresponding measurement error at a dry-bulb temperature of 59°F (15°C) is only 5%. *See* MOISTURE-CONTENT MEASUREMENT. Richard M. Schotland

Bibliography. C. F. Marvin, *Psychrometric Tables*, U.S. Department of Commerce, W.B. 235, 1941; A. Wexler (ed.), *Humidity and Moisture*, 1965.

Psychrometrics

A study of the physical and thermodynamic properties of the atmosphere. The properties of primary concern in air conditioning are (1) dry-bulb temperature, (2) wet-bulb temperature, (3) dew-point temperature, (4) absolute humidity, (5) percent humidity, (6) sensible heat, (7) latent heat, (8) total heat, (9) density, and (10) pressure.

The atmosphere in a clean pure state is a mechanical mixture of dry air and water vapor. Each is independent of the other and follows the laws of physics in accordance with its respective physical properties. Normal atmosphere contains impurities such as carbon dioxide, ozone, dust, and dirt in varying quantities.

Graphic representation. Even though dry air and water vapor are independent entities, there are certain relationships which permit the inclusion of these factors on a common chart, called a psychrometric chart (see **illus.**).

The psychrometric chart is a valuable tool for understanding the relationship of the many variables encountered in the atmosphere and for solving problems in air conditioning.

Dry-bulb temperature. The dry-bulb temperature is the ambient temperature of the air and water vapor as measured by a thermometer or other temperature-measuring device in which the thermal element is dry and shielded from radiation. Temperature scales are usually in degrees Fahrenheit or degrees Celsius. The vertical lines on the chart are dry-bulb temperature lines.

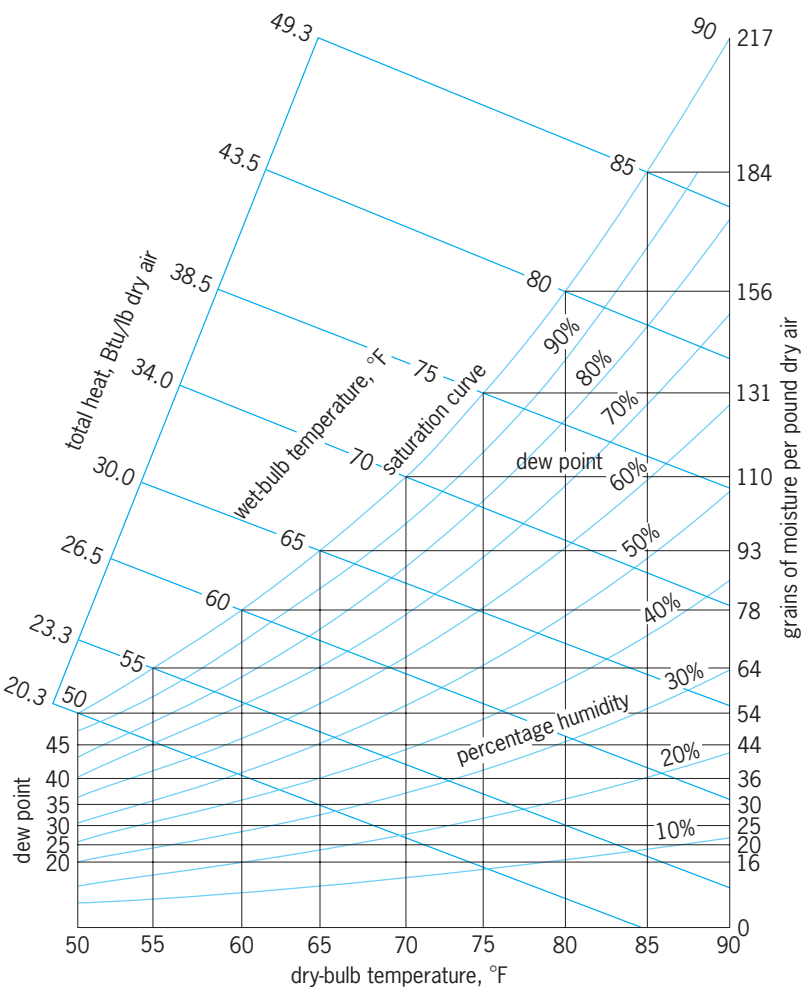

Psychrometric chart for air at a pressure of 29.92 in. of mercury (10^2 kPa). 1 grain/pound = 1/7 g/kg; 1 Btu/lb = 2326 J/kg; °C = (°F−32)/1.8.

Wet-bulb temperature. If the bulb of a dry-bulb thermometer is covered with a silk or cotton wick saturated with distilled water and the air is drawn over it at a velocity not less than 1000 ft/min (5 m/s), the resultant temperature will be the wet-bulb temperature.

If the atmosphere is saturated with water vapor, the water on the wick cannot evaporate and the wet bulb will read the same as the dry bulb. If the atmosphere is not saturated, water will evaporate from the wick at a rate dependent upon the percentage of saturation of the atmosphere. The cooling produced by the evaporation will result in a lowering of the temperature of the bulb and the consequent reading is the wet-bulb tempera- ture.

The lines on the psychrometric chart extending from upper left to lower right are wet-bulb temperature lines. Where the dry-bulb and wet-bulb temperatures are the same, the atmosphere is saturated and a line drawn through these points is called the saturation or 100% relative humidity curve.

Dew-point temperature. The dew-point temperature is the temperature at which the water vapor in the atmosphere begins to condense. This is also the temperature of saturation at which the dry-bulb, wet-bulb, and dew-point temperatures are all the same. *See* DEW POINT.

The dew-point temperature is a measure of the actual water-vapor content in the atmosphere. The water-vapor content is constant for any dew-point temperature regardless of the dry-bulb or wet-bulb temperature. The dew-point temperatures are the horizontal lines on the psychrometric chart.

Absolute humidity. The actual quantity of water vapor in the atmosphere is designated as the absolute humidity. It is measured as pounds of water vapor per pound of dry air, grains of water vapor per pound of dry air, grains of water vapor per cubic foot of dry air, or grams of water per kilogram of dry air. One pound of water vapor is equal to 7000 grains. Because the absolute water-vapor content is directly related to dew-point temperature, the vapor content in grains per pound of dry air is shown on the vertical scale on the right-hand side of the psychrometric chart for different dew-point temperatures. *See* HUMIDITY.

Percentage or relative humidity. Percentage or relative humidity is the ratio of the actual water vapor in the atmosphere to the quantity of water vapor the atmosphere could hold if it were saturated at the same temperature. For example, if the atmosphere at 50°F (10°C) is saturated with water vapor, there would be 53 grains of water vapor per pound of dry air (7.57 grams per kilogram of dry air). If the dry-bulb temperature of this atmosphere were then raised to 69°F (21°C), it could hold 106 grams of water vapor per pound or 15.14 grams per kilogram (corresponding to a saturation temperature of 69°F or 21°C). As no moisture was added in changing the temperature

from 53 to 69°F (12 to 21°C), the actual vapor content is still 53 grains. The percentage humidity would then be 53/106 or 50% at 69°F.

Sensible heat. Sensible heat, or enthalpy of dry air, is heat which manifests itself as a change in temperature. It is expressed in British thermal units (Btu) or joules (1 Btu = 1055 J). One Btu will raise the temperature of 1 lb of water 1°F (or 1 kg of water by 0.252°C). One Btu will also raise the temperature of 1 lb of dry air approximately 4.25°F (or 1 kg by 1.07°C) at atmospheric temperatures and sea-level pressure. *See* ENTHALPY.

Latent heat. Latent heat, or enthalpy of vaporization, is the heat required to change a liquid into a vapor without change in temperature. For example, it would require 1054 Btu to change 1 lb (or 2.45 megajoules to change 1 kg) of water at 70°F (21°C) from a liquid to a dry, saturated vapor at 70°F.

Latent heat is sometimes referred to as the latent heat of vaporization and varies inversely as the pressure. The higher the pressure (or saturation temperature) the lower the Btu required to evaporate water from a liquid to a vapor. **Table 1** shows the thermodynamic properties of water for various temperatures.

Total heat. The total heat, or enthalpy, of the atmosphere is the sum of the sensible heat, latent heat, and superheat of the vapor above the saturation or dew-point temperature. At saturation, the total heat measured in Btu/lb (1 Btu/lb = 2326 J/kg) of dry air is measured at the wet-bulb temperature and includes both the sensible heat of the dry air and the latent heat of the water vapor at the temperature measured. For example, at 60°F (15°C) on the saturation curve the total heat is 26.46 Btu/lb (61.55 kilojoules/kg) of dry air with water vapor to saturate the air at that temperature.

Total heat is relatively constant for a constant wet-bulb temperature, deviating only about 1.5–2% low at relative humidities below 30%. **Table 2** shows enthalpy values for saturated air.

Density. The density of the atmosphere varies with both altitude and percentage humidity. The higher the altitude the lower the density, and the higher the moisture content the lower the density. *See* DENSITY.

TABLE 1. Properties of water*

Temperature, °F (°C)	Pressure, lb/in.^{2} absolute	Enthalpy of evaporation, Btu/lb
40 (4.44)	0.122	1070.46
45 (7.22)	0.147	1067.81
50 (10.00)	0.178	1064.99
55 (12.78)	0.214	1062.16
60 (15.56)	0.256	1059.34
65 (18.33)	0.306	1056.52
70 (21.11)	0.363	1053.71
75 (23.89)	0.430	1050.89
80 (26.67)	0.507	1048.07
85 (29.44)	0.596	1045.23
90 (32.22)	0.698	1042.40
95 (35.00)	0.816	1039.56
100 (37.77)	0.950	1036.72

*1 lb/in.^{2} = 6.895 kPa; 1 Btu/lb = 2326 J/kg.

TABLE 2. Temperature-volume-enthalpy relation of saturated air*

Temperature, F (°C)	Volume, ft^3/lb of dry air	Enthalpy, Btu/lb of dry air
40 (4.44)	12.70	15.23
45 (7.22)	12.85	17.65
50 (10.00)	13.00	20.30
55 (12.78)	13.16	23.22
60 (15.56)	13.33	27.15
65 (18.33)	13.50	30.06
70 (21.11)	13.72	34.09
75 (23.89)	13.88	38.61
80 (26.67)	14.09	43.69
85 (29.44)	14.31	49.43
90 (32.22)	14.55	55.93
95 (35.00)	14.80	63.32
100 (37.77)	15.08	71.73

*1 ft^3/lb = 0.0624 m^3/kg; 1 Btu/lb = 2326 J/kg.

TABLE 3. Altitude and pressure (standard atmosphere)*

Altitude, ft	Pressure, in. Hg
−1,000	31.02
−500	30.47
0 (sea level)	29.92
500	29.38
1,000	28.86
5,000	24.89
10,000	20.58

*1 in. Hg = 3.386 kPa; 1 ft = 0.3 m.

At sea level (29.92 in. Hg or 101.325 kilopascals absolute pressure) and 59°F (15°C) the density is 0.0765 lb/ft^3 (1.225 kg/m^3). At 5000 ft (1524 m) elevation and 50°F (10°C), the density would be 0.0637 lb/ft^3 (1.020 kg/m^3).

At sea level and 65°F (18°C) saturated, the density is 0.074 lb/ft^3 (1.185 kg/m^3). At 65°F (18°C) dry bulb and 30% saturation, the density is 0.0752 lb/ft^3 (1.205 kg/m^3). The reciprocal of density (cubic feet per pound of dry air) is usually used rather than the density. Table 2 shows temperature-volume relations for saturated air.

Pressure. Atmospheric pressure, usually referred to as barometric pressure, is measured either in inches of mercury (29.92 in. Hg at sea level), in pounds per square inch absolute [14.7 lb/in.^{2} absolute (psia) at sea level], in millibars (1013.25 mb at sea level), or in kilopascals (101.325 kPa at sea level).

Pressure varies inversely as elevation, as temperature, and as percentage saturation. Pressure decreases with elevation as shown in **Table 3**. *See* MOISTURE-CONTENT MEASUREMENT; PSYCHROMETER.

John Everetts, Jr.

Bibliography. D. J. Croome and B. M. Roberts, *Air Conditioning and Ventilation of Buildings*, vol. 1, 2d ed., 1981; D. J. Croome and A. F. Sherratt (eds.), *Condensation in Buildings*, 1972; A. Pande, *Handbook of Moisture Determination and Control: Principles, Techniques, Applications*, vol. 1, 1974, vol. 3, 1975; B. Stamper and R. L. Koral (eds.), *Handbook of Air Conditioning, Heating, and Ventilating*, 3d ed.,

1979; A. Wexler (ed.), *Humidity and Moisture Measurement and Control in Science and Industry*, vols. 1, 3, and 4, 1964–1965.

Pteridosperms

A group of extinct seed plants characterized by fernlike leaves that produced naked seeds. The discovery of the seed ferns was a major contribution to the study of plant evolution because it demonstrated the existence of a group of vascular plants that is today extinct. Although the seed ferns probably have their ultimate origin within the progymnosperm order Aneurophytales, the best evidence of the group comes from lower Carboniferous (Mississippian) and younger sediments. Some seed ferns are reconstructed as trees with stout stems, while perhaps the majority were vines or lianas that supported massive fernlike fronds (**Fig. 1**).

The seed ferns consist of six Paleozoic orders (Calamopityales, Buteoxylonales, Lyginopteridales, Medullosales, Callistophytales, Glossopteridales) and three orders (Peltaspermales, Corystospermales, Caytoniales) found in Mesozoic rocks.

Calamopityales. This order is included with the seed ferns because of the nature of the wood and features of the cortex (**Fig. 2**). The stems may be protostelic to eustelic, but all are characterized by a conspicuous pith and petiole bases with numerous vascular bundles.

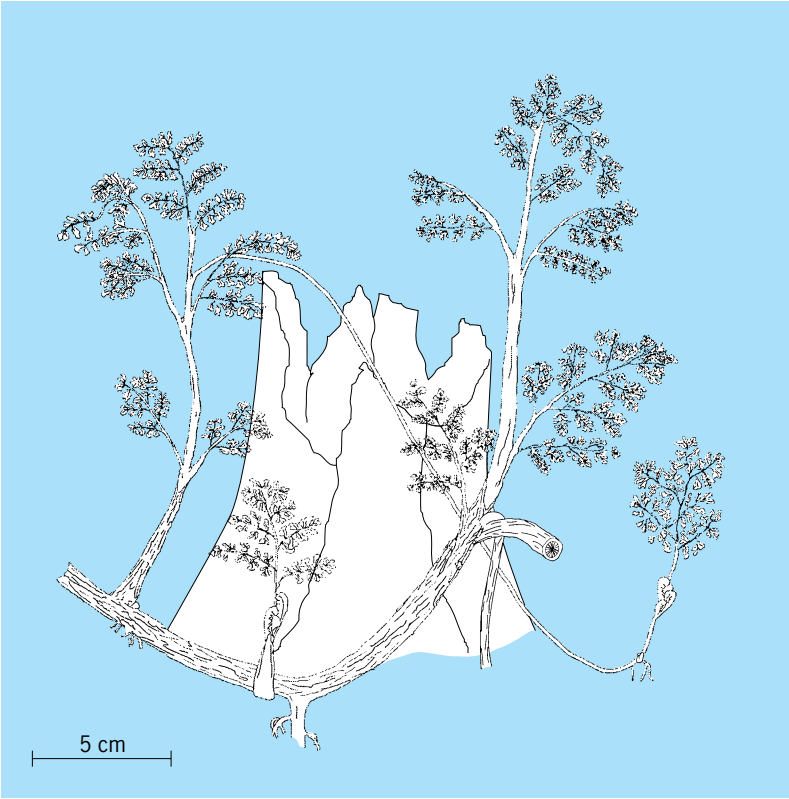

Fig. 1. Suggested reconstruction of *Callistophyton*. (*After G. W. Rothwell, The Callistophytaceae (Pteridospermopsida): I. Vegetative structures, Palaeontographica, 151B:171–196, 1975*)

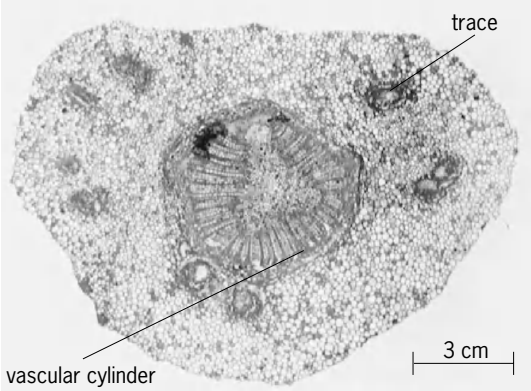

Fig. 2. Transverse section of *Stenomyelon*. (*From C. B. Beck, The appearance of gymnosperm structure, Biol. Rev., 45:379–400, 1970*)

Buteoxylonales. This group of lower Carboniferous seed ferns is also based on the anatomy of stems and petioles, but the primary features of the order include protostelic petioles that are trilobed in cross section. The cortex has sclerenchyma bands around the periphery. As with the Calamopityales, nothing is known about the foliage or reproductive organs of this group.

Lyginopteridales. This group of Carboniferous seed ferns has eustelic stems with a limited amount of secondary xylem. The leaves are large and planated, with numerous small pinnules. The seeds are small and characterized by thick integument and distinct micropyle. *Lagenostoma* is a seed that was approximately 0.2 in. (6 mm) long and was produced in a cup-shaped structure termed a cupule. Extending from the tip of the seed was a small conical projection that trapped wind-borne pollen. One feature that made it possible to unite some of the dispersed organs of the Lyginopteridales was small multicellular glands that extended from the epidermis of the stems, petioles, and seed cupules. Pollen organs consist of elongate pollen sacs that are fused laterally to form synangiate organs that produced spherical, trilete pollen grains. The middle of the pollen organ was hollow, suggesting that the sacs released pollen toward the center of the organ.

Medullosales. These seed ferns extend from the lower Carboniferous into the Permian and are characterized by stems that contain several segments of vascular tissue, each surrounded by secondary xylem. *Medullosa* is the best-known genus of stems, with some Permian forms being up to 20 in. (50 cm) in diameter and containing more than 40 vascular segments. The leaves were massive, with some specimens being several meters long; the petioles contained numerous vascular bundles in the base. Two of the common medullosan leaf types are *Alethopteris* and *Neuropteris*. The seeds of this group were noncupulate and radial in transverse section. Many were large (5 in. or 13 cm long) and ornamented by longitudinally oriented ribs. Pollen organs of the medullosans were large and synangiate, and ranged from simple forms with a few pollen sacs to compound types with more than 80 elongate sacs. The

pollen grains of this group are some of the largest produced by either living or fossil plants, with many up to 0.02 in. (0.5 mm) long.

Callistophytales. Although the seed ferns in this order form only a small group of upper Carboniferous pteridosperms, they are by far the best understood. Stems of *Callistophyton* were eustelic, and they produced pinnately compound leaves with axillary buds or branches at each node. The seeds were small and platyspermic. Pollen organs were synangiate and attached to the lower surface of laminar pinnules. Pollen grains possessed two distally inclined air bladders like those of modern pines. Pollen tubes were produced, and the microgametophyte consisted of an axial row of prothallial cells.

Glossopteridales. These plants are recognized as specialized, highly diverse seed ferns that dominated Southern Hemisphere continents (Gondwana) during the Permian. Leaves of *Glossopteris* are characterized by a reticulate pattern of veins, with some species more than 12 in. (30 cm) long. In transverse section, the underground parts consist of radiating wedges of wood with hollow areas between, a type of anatomy like that found in certain swamp-dwelling plants today. Seeds were produced in cupules or on exposed leaves. Pollen in this group is bisaccate.

Caytoniales. These Mesozoic seed ferns produced palmately compound leaves with a reticulate pattern of venation. Bisaccate pollen was produced in elongate synangia attached to shortened lateral branches. The seed-bearing units (*Caytonia*) consisted of helmet-shaped cupules approximately 0.2 in. (5 mm) in diameter. *See* CAYTONIALES.

Corystospermales. This small group of seed ferns occurs in Triassic sediments. One common foliage type was *Dicroidium*, a leaf characterized by a basal dichotomy of the rachis and small laminar pinnules. Some stems were more than 32 in. (80 cm) in diameter, suggesting that certain corystosperms were woody plants. Pollen sacs were borne on the lower surface of modified pinnule segments and produced bisaccate pollen. Seeds were partially covered by cupules, and the tip of the integument was curved and projected from the cupule surface.

Peltaspermales. These seed plants range from the upper Permian into the Triassic. The genus *Callipteris* is used for large bi- or tripinnate leaves bearing alethopterid pinnules. Seeds were produced on the lower surface of umbrella-shaped modified leaves, and pollen grains in the group were believed to have been boat shaped. *See* PALEOBOTANY.

Thomas N. Taylor

Bibliography. W. N. Stewart and G. W. Rothwell, *Paleobotany and the Evolution of Plants*, 2d ed., 1993; T. N. Taylor and E. L. Taylor, *The Biology and Evolution of Fossil Plants*, 1993.

Pteriomorphia

An order containing seven superfamilies of marine and brackish bivalves with byssal attachment at some time in their geological history (the Ostreacea are exceptional, with the lower valve attached by byssal cement at settlement, and with no record of ancestral attachment by byssal threads). Byssal attachment is associated with departure from the ancestral isomyarian form; Mytilacea and Pinnacea are heteromyarian, with a small anterior adductor muscle; others are monomyarian, with one central posterior adductor. Byssal attachment is lost in some: *Placuna* is free-living; *Pecten maximus* can right itself if inverted; *Chlamys* can swim to escape from starfish; some *Lima* can swim, while the tropical *Amusium* swims strongly. Pearls are obtained from *Pinctada* (Pteriacea).

The Pteriomorphia possess some primitive features and some secondarily simple features: the hinge is edentulous or with inconspicous teeth (except *Spondylus*); there are no siphons; minimal points of fusion occur between the left and right mantle lobes; the Mytilacea, Pteriacea, Pectinacea, and Anomiacea have filibranch ctenidia. Specialized features include simultaneous hermaphroditism in some Pectinacea, and sex reversal and incubation of larvae in some Ostreacea, and protandrous hermaphroditism in others. Few features are common to all Pteriomorphia, and the systematic status is debatable. Studies suggest addition of the isomyarian Arcacea and Limopsacea. *See* BIVALVIA; MOLLUSCA.

Richard D. Purchon

Pterobranchia

A group of small sessile hemichordates that may be colonial (*Rhabdopleura*), pseudocolonial (*Cephalodiscus*), or solitary (*Atubaria*). Each individual, or zooid, lives inside a nonchitinous tube secreted by the protosome, except for *Atubaria*; an aggregation of zooids is called a coenecium and can vary in shape. The protosome, or oral shield, is disciform, closes the mouth ventrally, and secretes the tube. The protocoel has symmetrical pores at the base of the first pair of arms. The mesosome, or collar, has an anteroventral mouth and one to nine pairs of dorsally ciliated tentaculated arms, which are used to collect small organisms for food. The mesocoels extend into both arms and tentacles.

The metasome is divided into a sacciform trunk and a slender ventral stalk that may be free at its end. The trunk contains the U-shaped digestive tract, with the pharynx having a single pair of gills, except in *Rhabdopleura*. The stomach is a sacciform expansion of the gut, and the tubular intestine curves dorsally and opens behind the arms by a middorsal anus, which in *Rhabdopleura* is on the right side.

The nervous system is very simple and lacks a neurochord. The buccal diverticulum is hollow, with an anteroventral heart vesicle and central sinus. The glomerulus is poorly developed, and the circulatory system is simpler that in enteropneusts.

Sexes are separate; the zooids show two anterodorsal sacciform gonads, except for *Rhabdopleura*, which has only one on the right side. The gonads open near the base of the last pair of

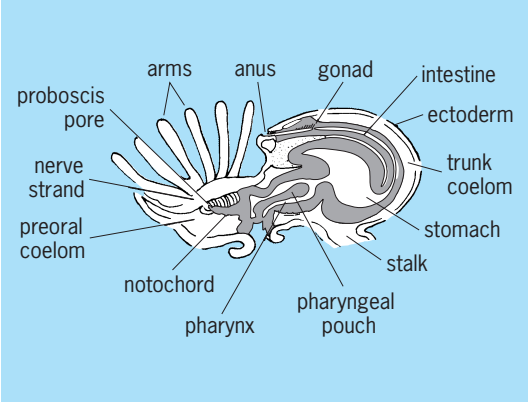

***Cephalodiscus*, of the Pterobranchia. (*After H. W. Rand, Comparative Anatomy, Blakiston-McGraw-Hill, 1936*)**

arms. Development is not completely understood in *Cephalodiscus*; in *Rhabdopleura*, the eggs undergo radial holoblastic cleavage and develop into lecithotropic larvae that feed on stored yolk and have a ventral glandular depression that secretes the larval cocoon. After metamorphosis, which takes place in the cocoon, a hole is formed in the cocoon's wall, and the ancestrula, the first zooid in the colony, builds the coenecium by budding. In *Rhabdopleura*, the ancestrula produces a branching tube that creeps over a substratum and gives off small erect tubes. In the inner side of the creeping tube a syncytial structure, the black stolon, forms and connects all the zooids of the colony by budding. In *Cephalodiscus*, the coenecium is composed of many unconnected individuals that originated from an ancestrula, and budding activity takes place in the distal zone of the stalk (see **illus.**). *See* HEMICHORDATA. Jesús Benito

Pteropsida

A large group of vascular plants characterized by having parenchymatous leaf gaps in the stele (vascular system of the stem) and by having leaves which are thought to have originated in the distant past as branched stem systems. Some botanists regard the Pteropsida as a natural group which they recognize as a class, subdivision, or division. Others regard it as an artificial assemblage of plants that have undergone certain similar changes from a rhyniophyte ancestry. The various components of the Pteropsida are here treated as three separate divisions under the names Magnoliophyta, Pinophyta, and Polypodiophyta. *See* MAGNOLIOPHYTA; PINOPHYTA; PLANT KINGDOM; POLYPODIOPHYTA. Arthur Cronquist

Pterosauria

The "winged reptiles" of the Mesozoic Era, constituting the closest major group to Dinosauria and sharing many features with them. Their common ancestor was a small, bipedal, agile archosaur reptile probably similar to the small *Scleromochlus* of the Late Triassic. This ancestor probably possessed a large, lightly built skull, a short body, long hindlimbs, and digitigrade feet with four long metatarsals. Pterosaurs inherited all these features, and they further evolved the power of flight. Bats have a wing of skin stretched among four fingers and extending from the arms to the legs (and often the tail), and birds have a wing of feathers supported by the arm and the fused wrist–hand bones. Pterosaurs had a wing of skin that was internally supported by long, fine, possibly keratinous stiffening fibers and braced by the forelimb, including a greatly elongated fourth finger (the first three fingers remained small). Their brains were relatively large and somewhat birdlike; the canals of the ear region show that they were well suited for life in the air. *See* DINOSAUR; REPTILIA.

Noteworthy finds. The first remains of pterosaurs were found in Germany in the 1780s, and it was not clear what sort of animals they were. Regarded as aberrant birds by some and aberrant bats by others, they were established as reptiles by the French comparative anatomist Georges Cuvier in the early 1800s. He noted that they were not at all like typical reptiles: they stood erect like birds, they had hollow bones and probably an insulatory covering, and they flew.

Since that time, hundreds of pterosaur skeletons and isolated bones have been discovered on most continents. The best records originally came from Germany, where the Late Jurassic Solnhofen Limestones, the lithified remains of ancient lagoon sediments, preserved many complete and beautiful specimens of *Rhamphorhynchus*, *Pterodactylus*, and other forms. Equally striking were pterosaurs such as *Dorygnathus* and *Campylognathoides* from the Early Jurassic of the Holzmaden region of southern Bavaria, and remains of *Dimorphodon* from southern England (**Fig. 1**). In the late 1800s, many worn and broken pterosaur bones were recovered from the newly explored badlands of Kansas and Nebraska; they included *Pteranodon* and *Nyctosaurus*. The earliest pterosaurs are known from the Late Triassic, about the time that the first dinosaurs appeared: they include *Eudimorphodon*, *Peteinosaurus*, and *Preondactylus*, all from marine rocks of Italy and Greenland. In recent decades, some of the most interesting finds have come from China and from a series of limestone nodules in the Santana region of Brazil, where fossil remains have included three-dimensional skeletons, some preserved soft parts, and bizarre skull crests. Specimens discovered in the lake district of Kazakhstan and Kirgizia preserve a kind of epidermal covering that appears to include both skin and a furlike covering, not identical to mammalian fur. New discoveries from the Early Cretaceous of China have included eggs of pterosaurs with the embryos intact.

Classification. Pterosaurs were traditionally divided into two subgroups. The "Rhamphorhynchoidea" comprised the early pterosaurs with long tails, a long fifth toe, and short metacarpal (palm) bones in the hand; the Pterodactyloidea reduced the tail and fifth toe to stumps and lengthened the metacarpals. The first category has been

abandoned because some rhamphorhynchoids are more closely related to pterodactyloids than to other rhamphorhynchoids. (The term pterodactyl, first used to describe these animals, now informally denotes any pterodactyloid; the term pterosaur describes any member of the Pterosauria.) Pterodactyloids ("wing fingers") were a subgroup of Late Jurassic and Cretaceous pterosaurs that replaced the early pterosaurs. The best-known members of this group include the small *Pterodactylus* and the much larger *Pteranodon* with 7-m (23-ft) wing span (**Fig. 2**). An even larger (11–12 m or 35–40 ft), partially known form, *Quetzalcoatlus*, was discovered in Texas, and similar large remains are known from the Middle East and elsewhere. These larger pterosaurs we were primarily soaring animals, like large birds.

At present there are over 100 known genera of pterosaurs, and most genera have only one known species. Many genera have historically been based on incomplete material, so it is difficult to compare them to others. Several conflicting schemes of the interrelationships of pterosaurs have been proposed; none seems clearly preferable. *See* ARCHOSAURIA; ORNITHISCHIA; SAURISCHIA.

Most analyses of the relationships of pterosaurs conclude that they shared a common ancestor with dinosaurs. These groups share long legs, a thigh bone with an offset head, a reduced fibula (the splint bone on the side of the shin in humans), a hingelike ankle, long foot bones and toes, and other features seen in no other animals. However, it has also been observed that all these features are related to locomotion and could be independently derived (though there is no evidence for this). Unfortunately, pterosaurs are so modified through the rest of their skeletons that they leave little trace of features that might link them to other reptiles. In order to advance this question, further analyses must include more kinds of pterosaurs and other reptiles, and more features held in common.

The hindlimb features that pterosaurs share with birds and other dinosaurs have suggested that pterosaurs also walked upright, on their toes, with a narrow gait, using the hindlimbs exclusively or nearly so. An alternative view suggests that the hindlimbs were too fragile, that the foot was connected to the wing and so impeded terrestrial progression, or that the joints of the feet did not allow walking on the toes. Recent discoveries in France of tracks with forefoot impressions far outside the body midline were clearly made by small pterosaurs slowly foraging for seaside invertebrates while walking on all fours. The feet seem to have been placed beneath the hip joints, and the heels are impressed. Thus, at least some pterodactyloids, with their elongate metacarpals, were habitual quadrupeds on the ground. However, we still know nothing of the stances and gaits of the first pterosaurs. Other indistinct tracks found elsewhere continue to be attributed to pterosaurs, but these tracks do not have clear features of pterosaur anatomy, and the forefoot prints seem too close to the body, so their trackmakers remain unresolved.

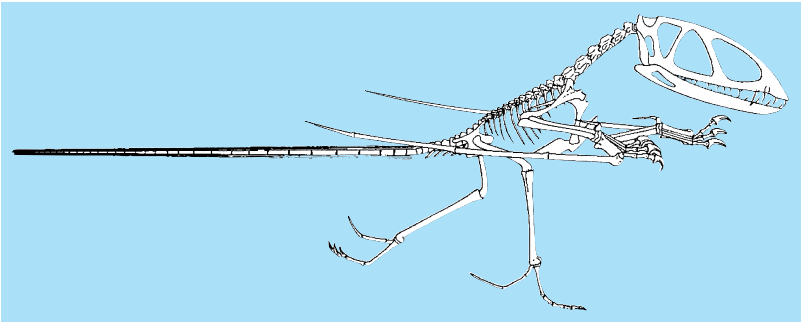

Fig. 1. Skeleton of *Dimorphodon*, an Early Jurassic pterosaur, restored as a biped. Wingspan is about 1.6 m (5.3 ft).

The fifth toe of early pterosaurs, much like the wing finger, was elongated and of variable length. Some reconstructions of pterosaurs have represented this toe as an anchor for the wing, but this idea is controversial because the size of the toe is so variable and its jointed morphology is apparently ill suited for the purpose (compared to the short, stiff, and potentially analogous calcar bone of some bats). Pterodactyloids lost the tail and the fifth toe, which poses difficulties for the explanation of this toe as a wing anchor. *See* SKELETAL SYSTEM.

Flight. Experts agree that pterosaurs were true flying reptiles, powered by wings that flapped vigorously. The flight muscles had extensive attachments to the wing bones and the expanded breastbone, and the shoulder girdle was modified to brace the force of the flight stroke against the skeleton. The wing itself was made of thin skin attached to the bones of the forearm and wing finger. It was reinforced by thin, perhaps keratinous fibers only a few micrometers in diameter but several centimeters long. These radiated over the undersurface of the wing in a pattern similar to the feather shafts of birds, and provided the stiffness and flexibility that made the wing membrane airworthy. All but the largest pterosaurs were active flappers; the largest relied on soaring because muscle power cannot sustain flapping flight for extended periods in such large forms.

Pterosaurs have other modifications for flight. The bones are extremely thin-walled and hollow, and many have openings on the surfaces that are remnants of channels through which the respiratory tissues ran. These tissues opened inside the bones, apparently providing further respiratory

Fig. 2. Late Cretaceous pterodactyloid *Pteranodon*, from western North America (reconstruction).

surface and helping to cool the blood. Birds have similar structures and functions, and their dinosaurian ancestors have at least the beginnings of these structures, which may reflect relatively high metabolic levels. The shoulder girdles of pterosaurs are braced against their large, expanded breastbones that anchored the flight muscles. Many large pterodactyloids have a series of fused vertebrae in the upper back, often showing evidence of sockets in which the ends of the scapulae (shoulder blades) inserted as a further brace. Specialized structures such as internal struts, plywoodlike layers of bone, and torsion-resistant tubular design helped the skeleton resist the forces of flight.

Pterosaurs, like birds and bats, evolved a down-and-forward flight stroke that created a donut-shaped vortex wake that provided the forward component of thrust. Anatomical studies demonstrate that pterosaurs produced various flight strokes. The most basic is roughly the shape of a lazy figure-8, at the shoulder joint, and an up-and-down motion was also possible. The elbow is a hinge, the wrist is blocky but slightly flexible, and most of the length of the wing is formed by a single elongated finger. The wing finger dwarfs the first three, which were free and could still manipulate objects. The base of the wing finger has a pronounced pulley joint, modified so as to flex the finger to the side rather than up and down like the other fingers. The rotation of this joint is very important in the flight stroke.

The shape of the wings of pterosaurs is an open question. Some specimens of *Rhamphorhynchus* suggest a narrow, gull-like wing that tapered toward the waist. Soft tissues preserved in some specimens of *Pterodactylus* suggest an attachment of the wing to the upper leg. However, without evidence of the supporting fibers, it is difficult to tell which soft-part remains represent the true wing and which are merely body remains squashed into two dimensions. Claims have been made that the Kirgiz specimens of *Sordes*, which preserve both wing fibers and other epidermal structures, demonstrate that the wings were attached to the ankles. This would have given the wing a much broader, more batlike planform (the shape and layout of a wing) than suggested by the *Rhamphorhynchus* specimens. In aerodynamic terms, the aspect of the wing (roughly the relationship of the length to the mean breadth) would be lower, and less like the aspect of many vigorously flying birds. This might suggest different styles of flight and different ecological roles. But the *Sordes* specimens have the same problems of flattening and ambiguous interpretation as the *Pterodactylus* specimens. At least one study has suggested that the specimens are systematically distorted and merely create an illusion of wing membranes that are nothing but skin.

Much more needs to be learned about the soft tissues of pterosaurs. Some filamentous epidermal structures appear not only on and near the wing areas but all over the body. These are apparently different from the structural fibers of the wing membranes, but it is not clear whether they grade into each other or are entirely different. The best-preserved specimens of *Rhamphorhynchus* indicate that the wing's structural fibers occur only on the surface of the underside of the wing, and they are embedded superficially in the epidermis, not free. The ability to differentiate the various kinds of fibers would be important in discerning whether a given preserved tissue in a pterosaur fossil is from the wing or another part of the body. Such information could have broad implications for flight style, terrestrial locomotion, and physiology. *See* FLIGHT.

Kevin Padian

Bibliography. S. C. Bennett, Terrestrial locomotion of pterosaurs: A reconstruction based on *Pteraichnus* trackways, *J. Vert. Paleontol.*, 17:104–113, 1997; E. Buffetaut and J.-M. Mazin, *Evolution and Palaeobiology of Pterosaurs*, 2003; K. Padian, A functional analysis of flying and walking in pterosaurs, *Paleobiology*, 9:218–239, 1983; K. Padian, Pterosaur stance and gait, and the interpretation of trackways, *Ichnos*, 10:115–126, 2003; K. Padian and J. M. V. Rayner, The wings of pterosaurs, *Amer. J. Sci.*, 293A:91–166, 1993; D. Unwin, *The Pterosaurs: From Deep Time*, 2005; P. Wellnhofer, *The Illustrated Encyclopedia of Pterosaurs*, 1991; L. M. Witmer et al., Neuroanatomy of flying reptiles and implications for flight, posture and behaviour, *Nature*, 425:950–953, 2003.

Pterygota

A subclass of the class Insecta comprising the winged insects and their secondarily wingless descendants. This subclass, including the great majority of insects, is usually divided into two infraclasses: Paleoptera, or ancient-winged insects (mayflies, dragonflies, and several extinct orders, most of which could not fold the wings over the body), and Neoptera, or folding-winged insects (the rest of the winged insects and some secondarily wingless lineages), which ancestrally folded the wings over the body by a particular articulatory mechanism.

The acquisition of wings occurred once in the Insecta and was the major general adaptation in their evolution. It is possible that wings arose as the muscled gills of aquatic nymphal Apterygota, which gained propulsive capabilities in the water. On the return of the adults to land for mating and reproduction, the movable gills acquired a new function in sexual and agonistic signaling, which soon became supportive for mating flight (as in many extant mayflies). With the basic mechanism already in place, modifications led to directed flight.

The winged insects, both Paleoptera and Neoptera, first appear in the fossil record in the Early Pennsylvania (upper Carboniferous), suggesting their origin in the Mississippian. *See* ENDOPTERYGOTA; EXOPTERYGOTA; INSECTA.

William L. Brown, Jr.

Bibliography. Commonwealth Scientific and Industrial Research Organization, *The Insects of Australia: A Textbook for Students and Research Workers*, 2d ed., 1991; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, vol. 2, 1982.

Public health

An effort organized by society to protect, promote, and restore the people's health. It is the combination of sciences, skills, and beliefs that is directed to the maintenance and improvement of health through collective or social actions. The programs, services, and institutions of public health emphasize the prevention of disease and the health needs of the population as a whole. Additional goals include the reduction of the amount of disease, premature death, disability, and discomfort in the population.

The sciences and techniques utilized in public health aim to protect, preserve, and restore good health. The basic sciences of public health include epidemiology and vital statistics, which measure health status and assess health trends in the population. Epidemiology is also a powerful research method, used to identify causes and calculate risks of acquiring or dying of many conditions. For example, it has been the principal technique for clarifying the nature of infection by human immunodeficiency virus (HIV), the causative agent of acquired immunodeficiency syndrome (AIDS). Many sciences, including toxicology and microbiology, are applied to detect, monitor, and correct physical, chemical, and biological hazards in the environment. Such applications are being used to address concerns about a deteriorating global environment. The social and behavioral sciences have become more prominent in public health since the recognition that such factors as indolence, loneliness, personality type, and addiction to tobacco contribute to the risk of premature death and chronic disabling diseases. *See* EPIDEMIOLOGY.

History and evolution. Some aspects of public health practice have ancient roots in empirical observations dating back to biblical times; for example, lepers carried bells to warn others of potential contagion, and simple sanitary rules were indicated. The Romans also practiced sanitary precautions on military campaigns 2000 years ago, but the nature of their plumbing and water supply systems showed a lack of understanding of the role of polluted water in the transmission of disease. Quarantine began in Europe in the fourteenth century and led to public health laws aimed at protecting populations against infection. Many basic principles of public health law were codified by the Austro-Hungarian physician J. P. Frank in 1779. In 1798, the English physician Edward Jenner discovered a vaccine to prevent smallpox. This find led to the eventual elimination of the disease, declared eradicated by the World Health Organization in 1980—the greatest achievement of coordinated worldwide public health action. *See* SMALLPOX.

Effective sanitary systems, supported by laws and regulations, did not become widespread until the late nineteenth century. Food hygiene followed in the early twentieth century. Medical discoveries in the late nineteenth and early twentieth centuries, beginning with the work of the great microbiologists Robert Koch and Louis Pasteur, led to the development of sera and vaccines to treat and prevent many infectious diseases. Biochemists and physiologists worked out the essential components of a healthy diet. Geneticists identified inherited diseases and conditions due to disorders of genes and chromosomes, laying the foundations for genetic screening and counseling programs that would eventually lead to prevention of some of these conditions. The importance of life-styles in the cause of heart disease and cancer was recognized in the 1960s. All these developments have helped to improve public health. *See* HUMAN GENETICS; NUTRITION; VACCINATION.

Services. It is possible to protect, preserve, and restore good health by making the environment safe, fostering sensible behavior, immunizing against infections, maintaining good nutrition, and providing prudent health care, including prenatal care. All these are part of the work of public health services.

In most industrial nations, public health services are organized nationally, regionally, and locally. National public health services exist primarily to protect the nation's health. They are usually responsible for setting, monitoring, and maintaining health standards, for promoting good health (for example, through national fitness and health education programs), for collecting and compiling national health statistics, and for supporting and performing research on diseases important to public health. Regional (for example, state) public health services deal mainly with major health protection activities such as ensuring safe water and food supplies; they may also operate screening programs for early detection of disease and are responsible for health care of certain groups such as chronic mentally ill persons. Local public health services (in cities, large towns, and some rural communities) conduct a variety of personal public health services, such as immunization programs, health education, health surveillance and advice for mothers and newborn babies, and personal care of vulnerable groups such as the elderly and housebound long-term sick. Local health services also investigate and control epidemics and other communicable conditions such as sexually transmitted diseases.

National public health services communicate with each other in efforts to control diseases of international importance, and they collaborate worldwide under the auspices of the World Health Organization (WHO). While much of the work of WHO has been concentrated in the developing nations, it has also been involved in global efforts to control major epidemic diseases and to set standards for hazardous environmental and occupational exposures. In tropical and developing nations, WHO conducts many programs aimed at controlling and, if possible, eliminating such diseases as malaria, schistosomiasis, infant diarrhea, tetanus, and measles. Premature death and long-term disabling diseases are still serious problems worldwide and in particular in the developing world.

The ultimate purpose of public health is to improve the human condition. Much has been accomplished, but much remains to be done,

especially to improve the health of the disadvantaged, minority groups, the poor, and, above all, people in the developing nations. *See* MEDICAL BACTERIOLOGY; MEDICAL MYCOLOGY; MEDICAL PARASITOLOGY. John M. Last

Bibliography. W. W. Holland, R. Detels, and G. Knox (eds.), *Oxford Textbook of Public Health*, 3d ed., 1997; J. M. Last (ed.), *Maxcy-Rosenau Public Health and Preventive Medicine*, 14th ed., 1998; J. M. Last, *Public Health and Human Ecology*, 2d ed., 1998; World Health Organization, *Global Strategy for Health for All by the Year 2000*, 1981.

Pulley

A wheel with a flat, crowned, or grooved rim (**Fig. 1**) used with a flat belt, V-belt, or a rope to transmit motion and energy. Pulleys for use with V-belt and rope drives have grooved surfaces and are usually called sheaves. A combination of ropes, pulleys, and pulley blocks arranged to gain a mechanical advantage, as for hoisting a load, is referred to as block and tackle. *See* BELT DRIVE; BLOCK AND TACKLE.

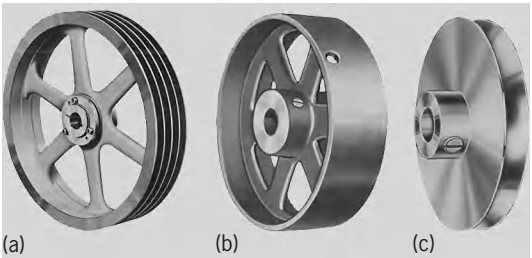

Fig. 1. Some typical examples of pulleys. (*a*) Multigroove V-belt sheave. (*b*) Solid-hub crown face pulley. (*c*) Brass pulley for round belt (*Boston Gear, INCOM International, Inc.*).

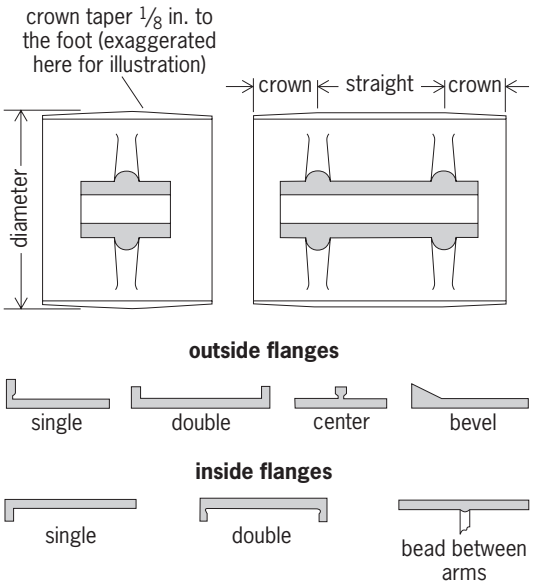

Fig. 2. Diagrams of crowning and flanging on flat-belt pulleys. $^1/_8$ in. to the foot ≅ 1 cm to the meter. (*After D. C. Greenwood, Mechanical Power Transmission, McGraw-Hill, 1962*)

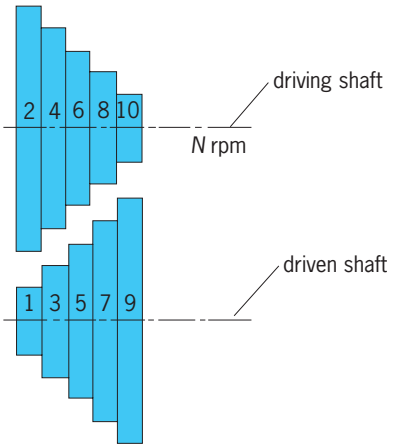

Fig. 3. Equal-stepped pulleys.

Flat-belt pulleys. Pulleys for flat belts are made of cast iron, fabricated steel, wood, and paper. A particular pulley design must be based on such considerations as the ability to resist shock, to conduct heat, and to resist corrosive environments. The face must be smooth enough to minimize belt wear; yet there must be adequate friction between belt and pulley face to carry the load.

Cast-iron pulleys are made in sizes from 3 to 72 in. (7.5 to 183 cm) and are particularly well adapted for service where dampness or acid fumes are present, in contrast to wood or steel pulleys. Because of their greater weight they are useful wherever a flywheel effect is wanted. Available in a wide variety of hub and spoke designs, they can also be furnished in either flat, crowned, or flanged rims. Pulleys are crowned (**Fig. 2**) to keep the belt on the pulley. The amount of crowning seldom exceeds a taper of $^1/_8$ in./ft (1 cm/m); excessive crown is harmful to the belt and may cause the belt to be thrown off the pulley as quickly as if there were no crown at all. Outside flanging (Fig. 2) helps hold the belt on the pulley; inside flanging stiffens the rim.

The arms on cast-iron pulleys are typically elliptical in cross section to achieve both a pleasing appearance and low windage loss, even though an H-section arm can provide equal strength for less material. Commercial cast-iron pulleys are balanced for rim speeds of 3500–4000 ft/min (18–20 m/s). For higher speeds the pulleys have to be specially balanced. Stock pulleys are made to withstand belt tensions of up to 65 lb/in. (115 newtons/cm) of width, though they can be furnished for tensions up to 210 lb/in. (370 newtons/cm).

The majority of pulleys used in flat-belt drives are steel split pulleys. They are low in cost, light in weight, and easily available, and are made with interchangeable bushings which permit their use on any shaft diameter up to the bore size of the pulley itself. Standard sizes range from 3-in. (7.5-cm) diameter by 2-in. (5-cm) width up to 144-in. (366-cm) diameter by 36-in. (91-cm) width.

Where lightness is important and moisture is not present, wood pulleys can be used. The resilience of wood makes these pulleys able to withstand shock

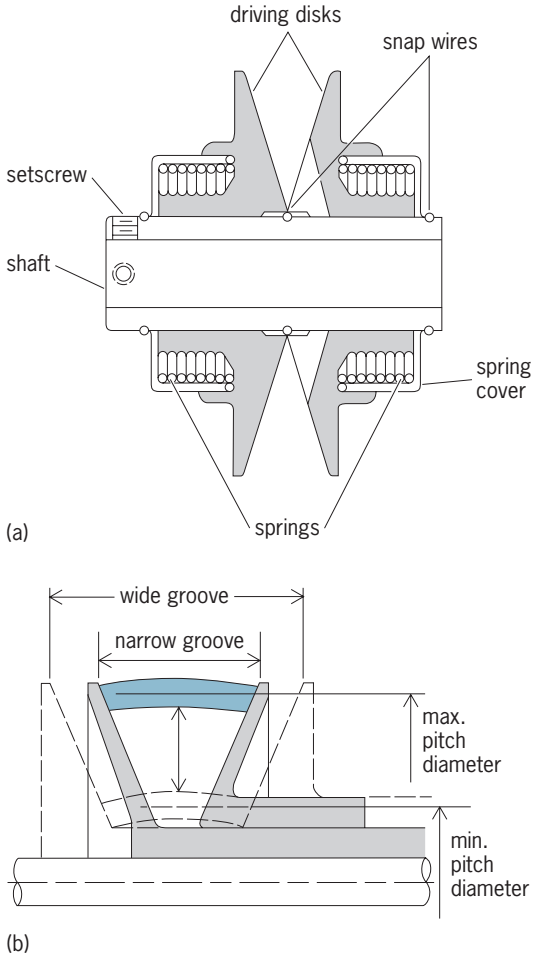

Fig. 4. Typical variable-pitch V-belt pulley (*a*) for standard-section V-belts and (*b*) for wide V-belts. Pulley can accommodate up to six belts. (*Lovejoy, Inc.*)

loads very well. Layers of soft wood are glued together and then compressed under high pressure to increase the density. A ribbed cast-iron hub is forced into an undersize hole in the wood block to complete the construction. Belt speeds of up to 6000 ft/min (30 m/s) are common.

Paper pulleys are made of impregnated paper board and are used particularly for motor pulleys. The recommended rim speed is 5000 ft/min (25 m/s).

If a belt drive must be able to produce several output speeds from a single input speed, stepped pulleys (**Fig. 3**) can be used. It is good practice to select the several output speeds so that they form a geometric progression. For a crossed-belt drive, the sum of the pulley diameters must be the same for each step. In practice a pair of stepped pulleys is often designed so that they have the same dimensions (Fig. 3). As shown in the equation below,

$$N = (n_3 n_7)^{1/2}$$

the driving shaft speed N is then the geometric mean of the output speeds of the two steps adjacent to the middle one.

V-belt pulleys. The two common types of V-belt pulley are pressed-steel and cast-iron. The pressed-steel pulleys are suitable for single-belt drives. For multiple-belt drives, or in single-belt drives where pulley mass should be high to get a flywheel effect, cast-iron pulleys are used (Fig. 1*a*). Generally the pressed-steel pulleys have noninterchangeable hubs, but the cast-iron pulleys have interchangeable hubs. Interchangeability of hubs reduces the number of pulleys that need be kept in stock: A single pulley can be made to fit a range of shaft diameters. Small pulleys have a groove angle of 30° to increase the wedging action of the belt in the groove. For pulley diameters of 5–8 in. (12.5–20 cm) a 34° angle is recommended. Over 8 in. (20 cm) a 38° angle is used.

Variable-pitch pulleys (**Fig. 4**) are made for drives where it is necessary to vary the speed ratio of the drive continuously over a range. Made for capacities up to 5 hp (3.7 kW), they can accommodate one to six belts.

Rope sheaves. Rope drives are much less efficient than V-belt drives, and so V-belt drives have virtually replaced rope drives. But in cranes and power shovels wire-rope drives are still in use. If a relatively soft material such as gray cast iron is used for the sheave, the wire rope will leave its imprint in the groove. A cast-steel sheave will stand up satisfactorily under the rope pressures experienced in typical installations, and heat-treated alloy steel, hardened to about 350 Brinell, will stand very high loads.

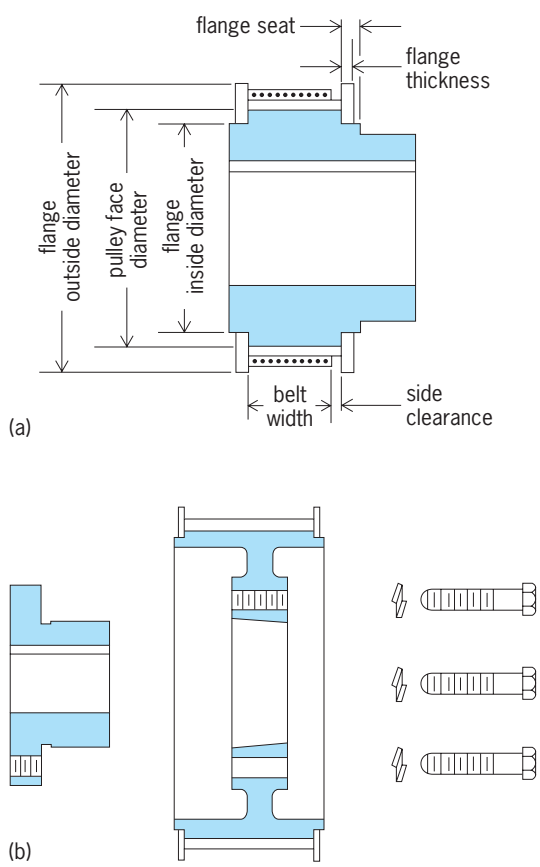

Fig. 5. Timing-belt pulley. (*a*) External plate-type flanges are used to retain belt. (*b*) Standard mounting arrangement permits various pulleys to be used on same shaft diameter. (*T. B. Wood's Sons Co.*)

Timing-belt pulleys. A timing-belt pulley (**Fig. 5**) is rather like an uncrowned flat-belt pulley, except that the grooves for the belt's teeth are cut in the pulley's face parallel to the axis, using a milling cutter, a gear shaper, or a hob. Most of the materials used in the manufacture of gears can be used as well in making timing-belt pulleys. Steel and cast iron are the most common materials, particularly for heavy loads. For light loads aluminum alloys, brass, plastics, and zinc or aluminum die castings are used. External flanges are used to keep the belt on the pulley.

John R. Zimmerman

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; R. T. Kent, *Mechanical Engineer's Handbook*, pt. 1: *Design and Production*, 12th ed., 1950; W. Patton, *Mechanical Power Transmission*, 1979; Trade and Technical Press Editors, *Handbook of Mechanical Power Drives*, 1977.

Pulmonata

A subclass of the molluscan class Gastropoda, containing about 23,500 species of snails that are grouped into three orders (or superorders): Systellommatophora, Basommatophora, and Stylommatophora. Pulmonates include most of the common snails found on land or in fresh waters, although a few prosobranchs have invaded both habitats. Certain pulmonates, the Onchidiidae, Amphibolidae, Otinidae, Siphonariidae, Trimusculidae, and some members of the Ellobiidae are intertidal to subtidal on rocky shores of tropical and temperate regions. *See* BASOMMATOPHORA; STYLOMMATOPHORA; SYSTELLOMMATOPHORA.

Comparatively few characteristics are common to most species, and there are exceptions to almost every statement. Pulmonates usually are hermaphroditic with cross-fertilization, but many inhabitants of temporary ponds, for example, can self-fertilize as an aid to colonization of new areas. All pulmonates, except the Amphibolidae, lack an operculum. Instead of an open mantle cavity with a gill, fusion of the mantle to the neck region has resulted in the formation of a pulmonary sac with a heavily vascularized surface and a small pneumostomal opening; the mantle cavity has become a lung. However, pseudobranchs (false gills not homologous with the ctenidia of other mollusks) are secondarily developed in some fresh-water taxa. Normally there is a spiral shell into which the animal can withdraw completely, but reduction of the shell to an external or internal remnant plate, or even complete loss of the shell, is a common evolutionary pattern in certain land mollusks. Such shell-less snails are called slugs. The animal is detorted, but in a different way from that in opisthobranchs. The single auricle lies (except in some slugs) in front of the ventricle. The size of pulmonates ranges from 0.04 to 9 in. (1 to 225 mm) in maximum adult shell size.

Most pulmonates are secretive dwellers in forests or nonsilted waters. They reach their highest levels of diversity in well-watered limestone hill regions and on relatively isolated islands. A few fresh-water species are of medical importance because they act as intermediate hosts for parasitic worms whose final hosts are humans or domestic animals. A very few land snails and slugs occasionally are pests in gardens or on crops. The vast majority of pulmonate species are beneficial in that they mechanically break up the dead parts of plants into pieces small enough to be quickly reduced by bacteria and fungi into chemicals that can be reused by plants to build new organic matter.

The earliest known pulmonates are stylommatophorans from the Late Carboniferous of eastern North America. All of these species belong to extant families of land snails. The first basommatophorans do not appear in the fossil record until the Jurassic-Cretaceous boundary, almost 1.5×10^8 years after the appearance of the stylommatophorans.

Fresh-water pulmonates belong to the superorder Basommatophora and are adapted to make use of accidental transport on birds or insects from one isolated body of water to another. Mostly they have short life-spans and their numbers vary dramatically with the seasons.

Land pulmonates belong to the superorder Stylommatophora and are marginally terrestrial in that they can be active only when the humidity in their microhabitat is 90% or above. Thus many are cryptic or nocturnal in both feeding and reproduction. They can survive in deserts by estivating for up to 6 years. *See* GASTROPODA.

G. Alan Solem

Bibliography. R. C. Brusca and G. J. Brusca, *Invertebrates*, 2nd ed., Sinauer Associates, Sunderland, MA, 2002; R. Chase, *Behavior and Its Neural Control in Gastropod Molluscs*, Oxford Univ. Press, 2002; V. Fretter and J. Peake, *Pulmonates*, vol. 2A, 1978; C. P. Hickman and L. S. Roberts, *Animal Diversity*, Wm. C. Brown, Dubuque, IA, 1994; L. Hyman, *The Invertebrates*, vol. 6: *Mollusca I*, 1967; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; A. Solem, *The Shell Makers*, 1974.

Pulsar

A celestial radio source producing intense short bursts of radio emission. Since their discovery in 1968, over 1750 pulsars have been found (as of October 2006), and it has become clear that 100,000 pulsars must exist in the Milky Way Galaxy—most of them too distant to be detected with existing radio telescopes. *See* RADIO ASTRONOMY.

Pulsars are distinguished from most other types of celestial radio sources in that their emission, instead of being constant over time scales of years or longer, consists of periodic sequences of brief pulses. The interval between pulses, or pulse period, is nearly constant for a given pulsar, but for different sources ranges from 0.0014 to 8.5 s. The bursts of emission are generally confined to a window whose width is a few percent of the period (**Fig. 1**). Individual pulses

can vary widely in intensity; however, their periodic spacing is accurately maintained.

Neutron stars. The association of pulsars with neutron stars, the collapsed cores left behind when moderate- to high-mass stars become unstable and collapse, is supported by many arguments. Prior to the discovery of pulsars, neutron stars were thought to be unobservable owing to their extremely small size (radius less than 15 km or 10 mi). The standard model for pulsars is a spinning neutron star with an intense dipole magnetic field (surface field of 10^8 teslas or 10^{12} gauss) misaligned with the rotation axis. The off-axis rotating dipole field develops a huge voltage difference between the neutron star surface and the surrounding matter. Charges accelerate in this voltage and generate an avalanche of electrons and positrons, a relativistic current leaving the polar zones of the star. Highly directive radio emission is formed in this current, which is observed as pulses, one per rotation, just like a rotating searchlight. *See* ELECTRON; POSITRON.

A fundamental observation that supports the rotating neutron star model is the remarkable stability of the basic pulsation periods, which typically remain constant to a few tens of nanoseconds over a year. This stability is natural to the free rotation of a compact, rigid object like a neutron star, but is extremely difficult to produce by any other known physical process. A small emitting body is also required by the observed rapid variations within pulses. A spinning, magnetized body will gradually slow down because it emits low-frequency electromagnetic radiation at harmonics of its rotation period. This slowing down is observed in all objects where observations of sufficient precision are available.

Further support of the neutron star model comes from the discovery of a 33-millisecond pulsar in the center of the Crab Nebula, the remnant of the supernova observed in A.D 1054. The lengthening of this pulsar's short period on a time scale of 1000 years matches the age of the remnant. The rotating, magnetized neutron star model is further supported by observations of diffuse radiation near the pulsar in the Crab Nebula by the *Hubble*, *XMM-Newton*, and *Chandra* space observatories. Currently, about 20 associations have been found between pulsars and supernova remnants which lend further support to these ideas. A 66-ms pulsar has been found in the supernova remnant 3C58. This is likely associated with the historical supernova of 1181, making this the youngest radio pulsar currently known. *See* CRAB NEBULA; NEUTRON STAR; SUPERNOVA.

Pulsed emission. The waveform of a pulsar's periodic emission, averaged over several hundred pulses or more, has a distinctive shape characteristic of each individual source (**Fig. 2**). Variations in pulse shape from one pulsar to another are probably the result of differences in detailed structure of the neutron star's magnetic field, or of a different orientation of the rotating beam relative to the Earth, or both. The emission is usually highly polarized, and the angle of the polarization vector (projected onto the plane of the sky) rotates smoothly across the pulse profile.

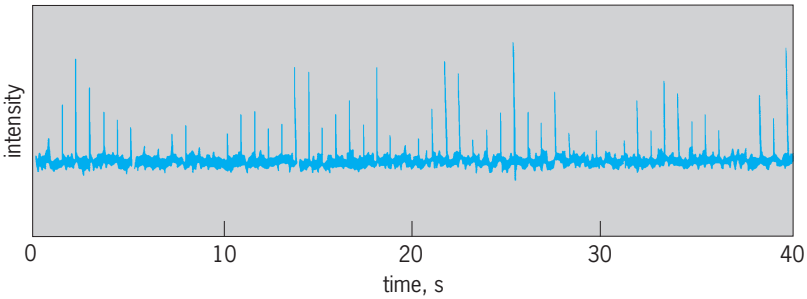

Fig. 1. Chart recording of radio emission from the pulsar PSR B0329+54, observed at a frequency of 400 MHz. This pulsar lies at a distance of about 7500 light-years (7×10^{16} km or 4.5×10^{16} mi) from Earth.

This behavior is taken to be the straightforward result of the changing projection of the pulsar's magnetic field lines as the star spins relative to the line of sight.

Individual pulses are found to vary rapidly in intensity, shape, and polarization (**Fig. 3**). Individual pulses normally consist of one or more subpulses, and the subpulses may in turn contain microstructure of even shorter duration. In some pulsars, subpulses vary in intensity in an obviously nonrandom way, and sometimes exhibit secondary periodicities incommensurate with the basic pulsar period. Individual pulses are more highly polarized than the average pulse profiles, and frequently undergo abrupt changes of polarization angle of almost exactly 90°. *See* POLARIZATION OF WAVES.

Spectra and links to other wavelengths. With a few notable exceptions, pulsars are observable only at radio frequencies, typically between 50 and 5000 MHz. Over this range the average flux density decreases rapidly, usually in approximate proportion to f^1 or f^2, where f is the observing frequency. At frequencies below a few hundred megahertz, a combination of galactic background radiation and terrestrial interference makes observations very difficult; at frequencies above a few thousand megahertz, most pulsars become too weak to be observed with existing instruments.

When extrapolated to optical or higher frequencies, the steep radio-frequency spectra of pulsars predict flux densities far below detectable limits. However, a few of the youngest pulsars, including the pulsar in the Crab Nebula, have been detected at optical, x-ray, and gamma-ray frequencies. Evidently, another emission mechanism is responsible for the high-frequency radiation from these pulsars. A mysterious gamma-ray source called Geminga has pulsed optical, x-ray, and gamma-ray emission but no convincing radio counterpart. This may be a case where the narrow radio beacon does not intercept the Earth. *See* GEMINGA.

Equally mysterious is the discovery of magnetars, pulsars that emit x-rays and gamma rays but may be radio-quiet owing to extreme values of magnetic field that suppress the electron-positron avalanche. However, radio emission has been detected during the outburst of a transient 5.5-s magnetar J1819-197. This remarkable discovery has provided a firm link between radio pulsars and magnetars, and challenges

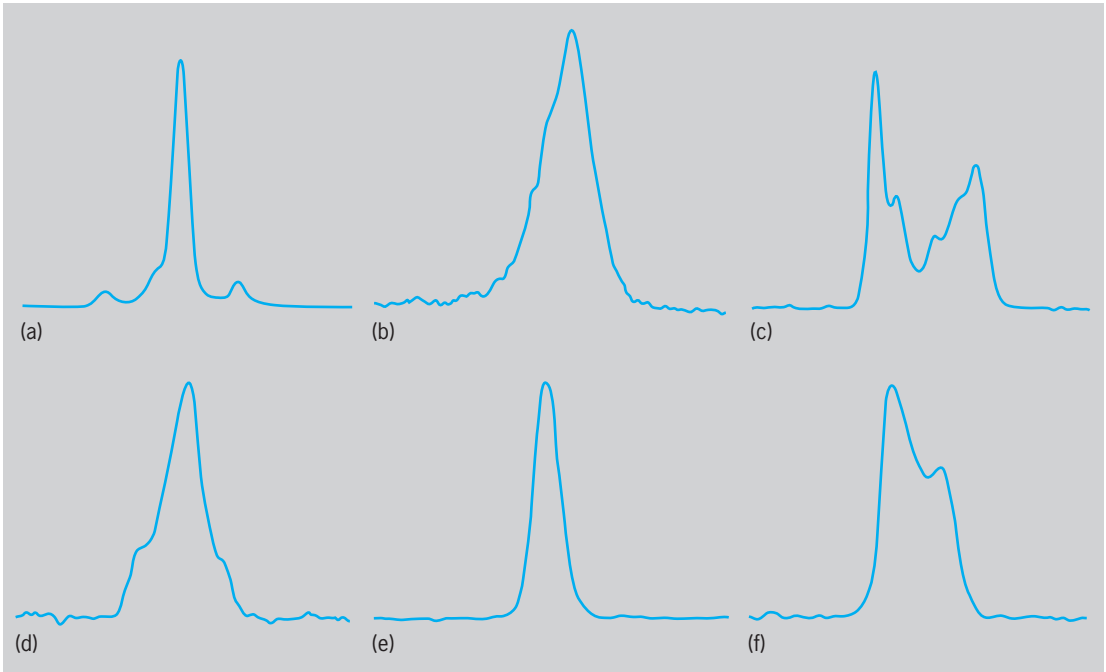

Fig. 2. Average pulse shapes of six pulsars. The data plotted cover a small, roughly 10% fraction of one rotation period. (*a*) PSR B0329+54. (*b*) PSR B0950+08. (*c*) PSR B1237+25. (*d*) PSR B1508+55. (*e*) PSR B1642-03. (*f*) PSR B1919+21.

the details of current theories of radio pulsar emission.

The emissions of most neutron stars are powered by the star's rotation. In contrast, there are accretion-powered neutron stars where matter from the atmosphere of a gravitationally bound companion star heats up to x-ray emission temperatures as it falls into the deep gravitational well of the neutron star. The infalling matter interrupts the process responsible for radio emission. *See* ASTROPHYSICS, HIGH-ENERGY; BINARY STAR; X-RAY ASTRONOMY.

Interstellar medium. Pulsars have provided a unique set of probes for the investigation of the diffuse gas and magnetic fields in interstellar space. Measurement of absorption at 1420 MHz, the frequency of the hyperfine transition in ground-state neutral hydrogen atoms, gives information on the structure of gas clouds, and in many cases provides an estimate of the pulsar distance. The index of refraction for radio waves in the ionized interstellar gas is strongly frequency-dependent, and low-frequency signals propagate more slowly than those at high frequencies. The broadband, pulsed nature of pulsar signals makes them ideal for measurements of this dispersion. Measurements of dispersion when coupled with independent distance estimates lead to both a model of the electron density in the nearby interstellar medium and the pulsar distribution. The average electron density is 0.03 cm^{-3}, and most pulsars are within 30,000 light-years of the Earth (1 light-year equals 9.46×10^{12} km or 5.88×10^{12} mi). The interstellar electrons are not distributed uniformly on small length scales, and their turbulent structures disrupt the observed signals with a variety of diffractive and refractive effects that are related to the twinkling of optical stars. *See* TWINKLING STARS.

Because of the large-scale Galactic magnetic field, the index of refraction in the interstellar medium is different for the two circular components of a linearly polarized wave. For this reason, the plane of polarization rotates along the propagation path, an effect known as Faraday rotation. Measurements of this effect for over 400 pulsars have shown that the magnetic field strength in the solar neighborhood of the Milky Way Galaxy is approximately 2×10^{-10} T (2×10^{-6} G) in a direction along the local spiral arm. The direction of the magnetic field along other spiral arms has also been measured using this technique, which provides strong evidence for field direction reversals in the inner Galaxy. *See* FARADAY EFFECT; INTERSTELLAR MATTER; MILKY WAY GALAXY.

Galactic distribution. The whole sky has been searched for pulsars using a variety of radio telescopes to carry out surveys at different observing frequencies. **Figure 4** shows the distribution of all currently known pulsars in galactic coordinates. The concentration of pulsars within $\pm 10^\circ$ of the galactic equator and in the hemisphere toward the galactic center (longitudes 0–90° and 270–360°) demonstrates that pulsars are members of the Milky Way Galaxy and that their distances are large compared to the thickness of the galactic disk. When distance estimates are combined with the two-dimensional information, a more elaborate analysis shows that pulsars occupy a disk-shaped region of space approximately 2000 light-years thick and 40,000 light-years in radius. The Sun, at a distance of about 30,000 light-years from the galactic center, lies near the outer edge of the pulsar distribution. Twenty pulsars are found in the Magellanic Clouds, while a further 129 pulsars have been associated with galactic globular clusters. *See* MAGELLANIC CLOUDS.

The massive, luminous stars that are believed to be the progenitors of pulsars have a galactic distribution similar to that of the pulsars, except that the thickness of their disk is only about 500 light-years. It appears that pulsars are found at much larger distances from the galactic plane because most pulsars acquire a substantial velocity (300 km s^{-1} or 200 mi s^{-1} or more) at birth, as a result of the supernova explosion. During a typical pulsar's active lifetime, roughly 10^7 years, the pulsar has time to move as much as 10,000 light-years away from its birthplace.

If pulsars typically remain active for 10^7 years, and if the present inferred galactic population of 100,000 pulsars is to be maintained, then a new pulsar must be formed somewhere in the Galaxy approximately every 100 years. This birth rate is reasonably close to the estimated rate of occurrence of galactic supernovas, if allowance is made for those that are optically obscured from view by interstellar dust.

Millisecond pulsars. The view that few if any pulsars spin faster than the 30-ms pulsar in the Crab Nebula was shattered with the discovery in 1982 of PSR B1937+21 with the remarkably short period of 1.6 ms, the first millisecond pulsar. This period is within a factor of 2 or so of the stability limit of neutron stars of 1 solar mass. The period is lengthening, but on time scales of 10^8 years, not 10^6 years or less. Evidently, this object is old and has a low magnetic field. About 150 millisecond-period pulsars are now known. Despite intensive searches over the past 20 years, PSR B1937+21 remains the most rapidly rotating neutron star known in the galactic disk. In 2005, a pulsar was found in the globular cluster Terzan 5 with a spin period of 1.4 ms, that is, a spin rate of 716 Hz. It is an open question as to whether a significant number of pulsars spinning even more rapidly exist.

Millisecond pulsars are extremely accurate clocks which exhibit few or none of the rotation irregularities seen in most pulsars. The observed stabilities rival the best terrestrial atomic time scale, which is derived from cesium clocks maintained around the world. Timing an array of millisecond pulsars will provide the best time scale for durations exceeding 1 year. The data may also improve knowledge of the orbit of the Earth in the solar system, and be used to set limits on, or even detect, a background of gravitational radiation left over from chaotic events in the early universe. *See* ATOMIC CLOCK; ATOMIC TIME; GRAVITATIONAL RADIATION.

Binary pulsars. Over 120 pulsars with orbiting companions are now known. Orbital periods range from hours to years. Timing observations of the binary pulsars, which provide indications of the companion masses, together with optical observations which often identify or constrain the nature of the companion star, show that there are several different types of orbiting companion: planets, white dwarfs, other neutron stars, and main-sequence stars.

The origin of most binary pulsars can be understood by following the evolution of a pair of main-sequence stars. After the supernova explosion of the initially more massive (primary) star, the resulting neutron star initially behaves normally—spinning down to longer periods. The spin evolution changes, however, when the initially less massive (secondary) star in the system reaches the end of its main-sequence lifetime and enters a red giant phase. For favorable orbital configurations, the strong gravitational field of the neutron star can draw material from the evolved secondary. Accretion of mass and angular momentum from the secondary can spin-up the neutron star to millisecond periods. This process is known as recycling, and the resulting rapidly spinning neutron stars are often termed recycled pulsars. *See* GIANT STAR; STELLAR EVOLUTION.

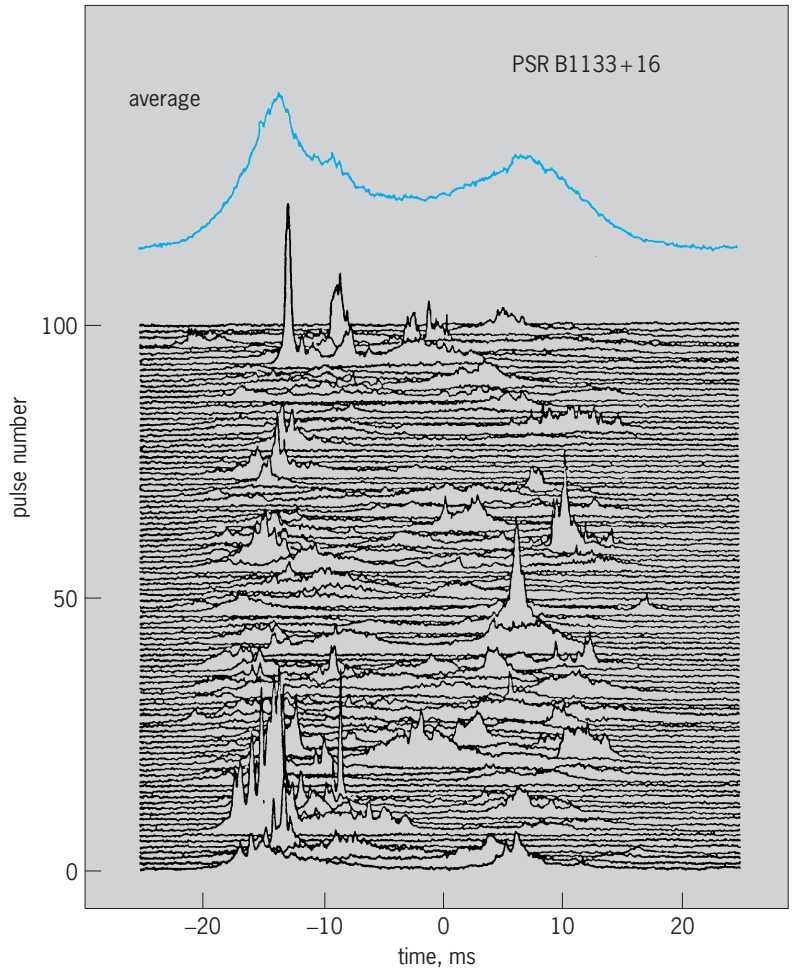

Fig. 3. Individual pulses from pulsar PSR B1133+16, illustrating the variations which take place within pulses and from one pulse to the next.

If the secondary star is massive enough to also explode as a supernova and the binary system survives this explosion, a double neutron star system may be formed. For less massive secondary stars, a white dwarf is formed at the end of the red giant phase. Tidal circularization of the orbit during the recycling phase means that the resulting system is observed as a millisecond pulsar in a circular orbit around a white dwarf. *See* WHITE DWARF STAR.

While the spin-up scenario explains the currently observed spin parameters of most binary and millisecond pulsars, PSR B1937+21 and a few others

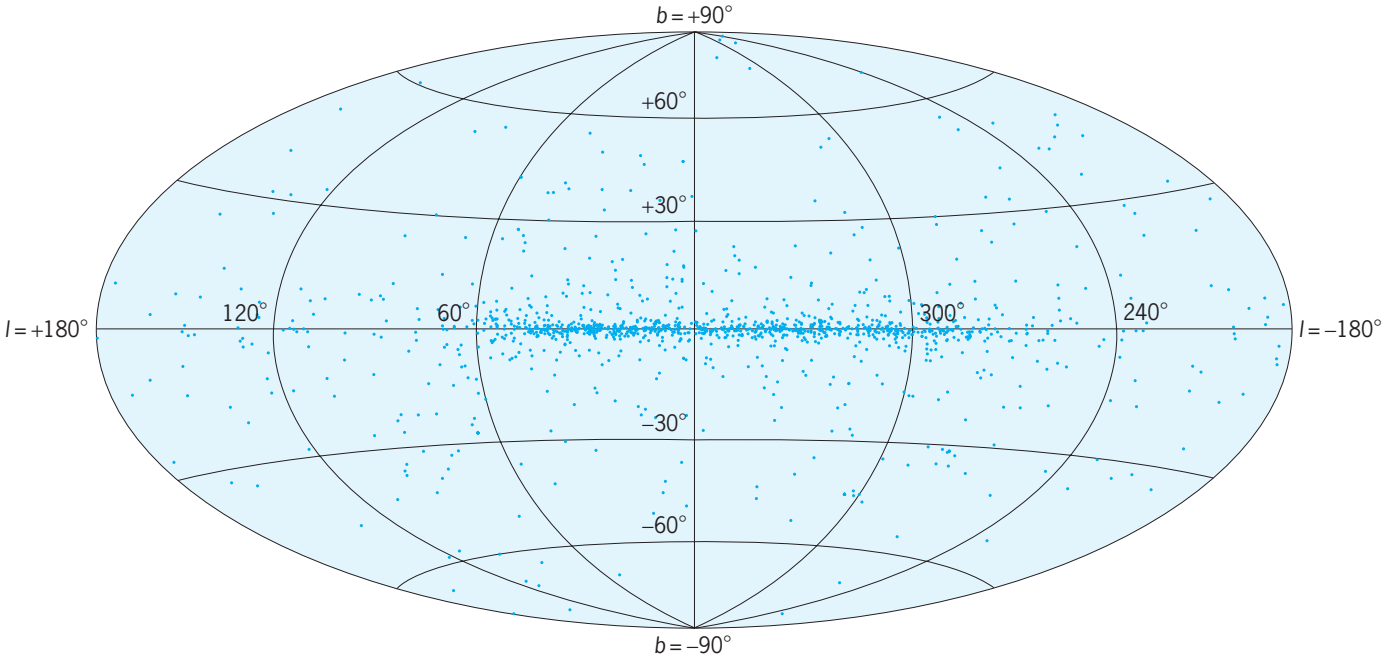

Fig. 4. Distribution of pulsars in Galactic coordinates, showing all pulsars known in late 2006.

are still a puzzle since they have no companion. Several binary millisecond pulsars show eclipses by matter surrounding the companion, which suggests that a pulsar can ablate its companion out of existence. These systems may be the missing links between binaries and solitary millisecond pulsars.

Double pulsar system. Currently (2006), eight examples of double neutron star binary systems are known. In all but one of these systems, only one of the neutron stars has its emission beam passing over the Earth. In 2003, the first system where both pulsar beams are visible was found. This double pulsar binary, PSR J0737-3039, consists of a 22-ms pulsar in a highly relativistic 2.4-h orbit with a 2.7-s pulsar companion. This compact orbital system is viewed almost edge-on to the line of sight and has allowed observers to study the interaction of two radio pulsar beams for the first time. Two effects have been seen so far: the eclipse of the 22-ms, pulsar by the 2.7-s pulsar once per orbit and the modulation of the 2.7-s pulsar by the radiation beam of the 22-ms pulsar at specific orbital phases. Further studies of these interactions may help provide a better understanding about the pulsar radiation mechanism in general.

The pulsars in this unique binary have orbital speeds about one-thousandth the speed of light. The large speeds and intense gravitational fields of the two neutron stars, combined with the accurate clock mechanism provided by both components, make this system an ideal testing ground for relativistic gravity. In addition to the five classical (Keplerian) orbital parameters determined for this system and the mass ratio of the two stars, a further five relativistic (so-called post-Keplerian) parameters have been measured so far. Among the many effects seen in this system is the precession of periastron which occurs at a rate of almost 17° per year—a factor of 10^5 times larger than the general relativistic perihelion advance observed for Mercury.

Within 3 years of its discovery, the double pulsar system surpassed the previous two best double neutron star systems as a precision test of general relativity. Currently (2006) the comparison between observed and predicted post-Keplerian parameters shows that general relativity is correct at the 99.95% confidence level. Ongoing timing measurements should provide even more stringent constraints and new tests of Einstein's theory.

One especially important prediction of the general theory of relativity is that a close binary star system should gradually lose energy by the radiation of gravitational waves, and consequently the two stars should slowly spiral closer together. Observations of double neutron star binary systems show just such an effect, and provide the most stringent experimental evidence available in support of the existence of gravitational waves.

The fate of close double neutron star binaries is a furious coalescence event that will produce a burst of gravitational radiation detectable by Earth-based gravity-wave telescopes now in operation. Although double neutron star mergers in the local Universe are not thought to be frequent enough to be readily observable by current detectors, planned upgrades in sensitivity are expected to allow the direct detection of gravitational waves within the next decade. *See* GRAVITATION; RELATIVITY.

Globular-cluster pulsars. Stars in the centers of globular clusters are only a few light-weeks apart so that frequent stellar encounters occur. Neutron stars in clusters have a 10% chance of capturing a companion during the life of the cluster. This constant production of binaries can lead to an abundance of x-ray objects as well as millisecond pulsars.

Currently, there are 129 pulsars known in 44 globular clusters. Studies of this population exploit the clocklike stability of pulsars to probe the physical properties of the pulsars and the cluster itself. For example, in the cluster 47 Tucanae, in which there are currently 22 pulsars known, timing measurements which map out the pulsar positions in three dimensions have provided the first evidence for the presence of ionized gas in any globular cluster.

The richest cluster, Terzan 5, has 33 pulsars with rotation periods between 1.4 and 80.3 ms. Studies of these pulsars are being used to probe the pulsar properties and overall mass distribution of Terzan 5. Of particular interest is a so-far unidentified radio source in the cluster. This enigmatic source may be a bright millisecond pulsar that has evaded detection by virtue of being in an extremely compact orbit, perhaps around a black hole. The identification of such a system is a high-priority objective in pulsar astronomy, and Terzan 5 may be the place where the system can be found. *See* STAR CLUSTERS.

Duncan R. Lorimer; Donald C. Backer

Bibliography. M. Bailes, D. J. Nice and S. E. Thorsett (eds.), *Radio Pulsars*, Astron. Soc. Pacific Conf. Ser. 302, 2003; D. R. Lorimer, Binary and millisecond pulsars, *Living Rev. Relativity*, vol. 8, no. 7, 2005; D. R. Lorimer and M. Kramer, *Handbook of Pulsar Astronomy*, 2005; A. G. Lyne and M. Kramer, Gravitational labs in the sky, *Phys. World*, 18(3):29–32, March 2005; A. G. Lyne and F. G. Smith, *Pulsar Astronomy*, 2006; T. Piran, Binary neutron stars, *Sci. Amer.*, Spec. Ed., pp. 84–91, 2004; F. A. Rasio and I. H. Stairs (eds.), *Binary Radio Pulsars*, Astron. Soc. Pacific Conf. Ser. 328, 2004; F. G. Smith and M. A. McLaughlin, A magnetopause in the double-pulsar binary system, *Astron. Geophys.*, 46:23–26, 2005; E. P. J. van den Heuvel and J. van Paradijs, X-ray binaries, *Sci. Amer.*, Spec. Ed., pp. 58–67, 2004.

Pulse demodulator

A device that extracts the analog information from a received pulse or from a set of pulses.

Constant-amplitude pulse-width-modulated (PWM) and pulse-position-modulated (PPM) signals carry their information in the time interval that the pulse is "on" in the first case, and in the time interval between the pulse of interest and the framing or timing pulse in the second case. PPM signals can be converted to binary PWM signals (amplitudes of "V" volts when on, amplitude zero when off) by using the signal pulses to latch a gated V-volt signal "on" and the timing signal to turn the gate "off." The source information is in the average value of the resulting PWM signal; therefore, the following source-information-recovery circuit can be a simple low-pass filter whose cutoff frequency must lie above the bandwidth of the source signal to preserve the information content, but below $^1/_2$ the frequency of the framing or timing signal to reject the timing-pulse noise. The source analog information behind ternary PWM signals of amplitudes $+V$, 0, and $-V$ is recovered in the same manner. *See* ELECTRIC FILTER; LOGIC CIRCUITS; PULSE MODULATOR.

Pulse-amplitude-modulated (PAM) signals are pulses of constant width, and usually the product of the source signal and a sampling-pulse train of frequency f_s. The sampling pulses may be regarded as windows through which the source signal is observed. The information is carried in the average amplitude of the pulses. Recovery again is achieved by simple low-pass filtering. The filter cutoff frequency must be above the information bandwidth of the source signal and less than $^1/_2$ the pulse-repetition frequency f_s of the sampling-pulse train in order to avoid foldover distortion.

Pulse-density-modulated (PDM) signals consist of trains of constant-amplitude and constant-duration pulses with variable spacing; therefore, those pulse rates are proportional to the amplitudes of the source signals. The analog source information is again recovered by a simple low-pass filter.

Pulse-code-modulated (PCM) signals are digital representations of their source signals. A conventional digital-to-analog converter (DAC) recovers the analog source signal. Variations on PCM such as delta Modulation (DM), adaptive delta modulation (ADM), and differential PCM (DPCM), all of which have been introduced to improve channel efficiency, require specialized and often complicated recovery circuits. These specialized circuits may be considered as performing a two-step process: recovery of the basic PCM signal followed by a conventional digital-to-analog converter. *See* DIGITAL-TO-ANALOG CONVERTER.

Phase-modulated (PM) signals, sometimes called PSK, or phase-shift keyed, are constant-amplitude sinusoidal segments shifted in phase from a reference sinusoid according to the amplitude of the source signal. The source signal may be recovered by synchronous detection (multiplication) with a reference sinusoid. *See* PHASE MODULATION.

Frequency-modulated (FM) signals are constant-amplitude sinusoids whose frequencies are varied in accordance with the amplitude of a reference sinusoid. A pulse containing a burst of FM data at known discrete frequencies can be demodulated by applying that pulse to a bank of filters tuned to the known discrete frequencies. The power appearing at the filter outputs signifies which frequencies are present in the pulse. FM signals can also be recovered by a phase-locked loop (PLL), but not as quickly. Integration of the recovered FM signal yields PM recovery, an alternative to the method described above. *See* FREQUENCY MODULATION; MODULATION; PHASE-LOCKED LOOPS.

Quadrature-amplitude-modulated (QAM) signals appearing as a burst within a pulse must be synchronously demodulated in two parallel channels by reference signals in quadrature with one another. Subsequent processing of each of the two channels involves integration, digitizing, bit formatting, and data-stream interleaving in order to recover the PCM-formatted data. Finally, a pair of digital-to-analog converters is required to recover the

analog source signals. *See* DEMODULATOR; MODULATOR.

Stanley A. White

Bibliography. G. A. Breed (ed.), *Wireless Communications Handbook*, Cardiff Publishing, Englewood, CO, 1992; S. Gibilisco, *Handbook of Radio and Wireless Technology*, McGraw-Hill, New York, 1999; J. G. Proakis, *Digital Communications*, 4th ed., McGraw-Hill, New York, 2001; T. S. Rappaport, *Wireless Communications*, 2d ed., Prentice Hall, Upper Saddle River, NJ, 2002; M. S. Roden, *Analog and Digital Communications Systems*, 4th ed., Prentice Hall, Englewood Cliffs, NJ, 1995; M. E. Van Valkenburg and W. M. Middleton (eds.), *Reference Data for Engineers: Radio, Electronics, Computer and Communications*, 9th ed., Butterworth-Heinemann, Woburn, MA, 2002.

Pulse generator

An electronic circuit capable of producing a waveform that rises abruptly, maintains a relatively flat top for an extremely short interval, and then rapidly falls to zero. A relaxation oscillator, such as a multivibrator, may be adjusted to generate a rectangular waveform having an extremely short duration, and as such it is referred to as a pulse generator. However, there is a class of circuits whose exclusive function is generating short-duration, rectangular waveforms. These circuits are usually specifically identified as pulse generators. An example of such a pulse generator is the triggered blocking oscillator, which is a single relaxation oscillator having transformer-coupled feedback from output to input. *See* MULTIVIBRATOR.

Pulse generators sometimes include, but are usually distinguished from, trigger circuits. Trigger circuits, by means of *RC* (resistance-capacitance) differentiating, gated *RLC* (resistance-inductance-capacitance) peaking circuits, or blocking oscillators, generate a short-duration, fast-rising waveform for initiating or triggering an event or a series of events in other circuits, such as monostable or bistable multivibrators. In the pulse generator, the pulse duration and shape are of equal importance to the rise and fall times. In this sense the blocking oscillator is a circuit which can be made to perform well in both respects. *See* TRIGGER CIRCUIT.

Digitally controlled pulse generator. The term pulse generator is often applied not only to an electronic circuit generating prescribed pulse sequences but to an electronic instrument designed to generate sequences of pulses with variable delays, pulse widths, and pulse train combinations, programmable in a predetermined manner, often microprocessor-controlled.

For example, in the circuit shown in **Fig. 1**, the basic pulse generator is a bistable *R-S* or *J-K* flip-flop under control of a sequence of clock pulses with a digital logic block to determine the appropriate combinations of set (*S*), reset (*R*), and clock pulses. An output pulse can be programmed to start with the next clock pulse following $R = (1)$ and stop with the clock pulse following $S = (1)$. Thus, the pulse width which can be programmed is nT_c, where T_c is the clock pulse interval and n is any number equal to or greater than unity.

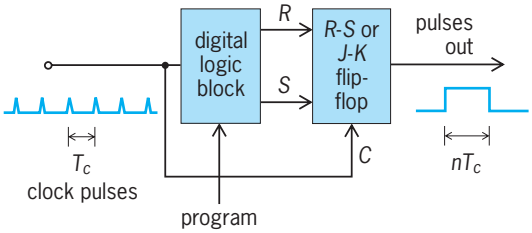

Fig. 1. Digitally controlled pulse generator.

Pulse-forming networks. A network, formed in such a way as to simulate the delay characteristics of a lossless transmission line, and appropriate switching elements to control the duration of a pulse form the basis for a variety of types of pulse generators. *See* DELAY LINE.

A lossless transmission line has a characteristic impedance, given by Eq. (1), where L is the series

$$R_0 = \sqrt{\frac{L}{C}} \tag{1}$$

inductance and C the shunt capacitance per unit length. Such a line may be approximated by a network consisting of a number n of cascaded LC elements. A pulse applied to the input of such a line reaching the output offers a time delay T_d, given by Eq. (2).

$$T_d = n\sqrt{LC} \tag{2}$$

An idealized circuit showing how such a network is used in a pulse generator is shown in **Fig. 2**. A source is connected by a switch to a simulated, unterminated transmission line in series with a resistance R_0 equal to the characteristic impedance of the line. At the time voltage is applied, and until all the capacitors in the line become fully charged, the input impedance of the line is equal to R_0, and the current I is given by Eq. (3). The step

$$I = \frac{V_i}{2R_0} \tag{3}$$

function progresses along the line, charging each C in succession. When it reaches the end, it is reflected back with no change in phase and

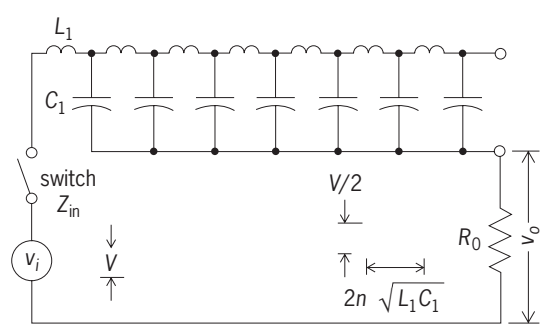

Fig. 2. Principle of line-controlled pulse generator.

returns to the source in a time T, in Eq. (4).

$$T = 2n\sqrt{LC} \tag{4}$$

At this time the line is fully charged, the impedance becomes infinite, and current ceases to flow. The pulse appearing across R_0 is suddenly terminated as shown. For discussion of switching circuits suitable for supplying the line-charging current *see* GATE CIRCUIT

Similar results can be obtained from the use of a current generator as a source and a short-circuited line as the controlling circuit element.

Various forms of delay-line-controlled pulse generators can be found, and some are capable of generating pulses containing considerable amounts of power for such applications as modulators in radar transmitters. Glenn M. Glasford

Bibliography. D. A. Bell, *Solid State Pulse Circuits*, 4th ed., 1997; R. L. Castellucis, *Pulse and Logic Circuits*, 1976; R. J. Tocci, *Fundamentals of Pulse and Digital Circuits*, 3d ed., 1983.

Pulse jet

A type of jet engine characterized by periodic surges of thrust. The pulse jet engine was widely known for its use during World War II on the German V-1 missile (**Fig. 1**). The basic engine cycle was invented in 1908. The inlet end of the engine is provided with a grid to which are attached flap valves. These valves are normally held by spring tension against the grid face and block the flow of air back out of the front of the engine. They can be sucked inward by a negative differential pressure to allow air to flow into the engine. Downstream from the flap valves is the combustion chamber. A fuel injection system is located at the entrance to the combustion chamber. The chamber is also fitted with a spark plug. Following the combustion chamber is a long exhaust duct which provides an inertial gas column.

When the combustion chamber is filled with a mixture of fresh air and fuel, a spark is discharged; it ignites the fuel-air mixture, producing a pressure surge that advances upstream to slam shut the inlet valves and to block off the entrance. Simultaneously, a pressure pulse goes downstream to produce a surge of combustion products out the exhaust duct. Thrust results from the rearward discharge of this gas at high velocity. With the discharge of gas from the combustion chamber, its pressure tends to drop. Inertia causes the column of gas in the exhaust duct to continue to flow rearward even after the explosion

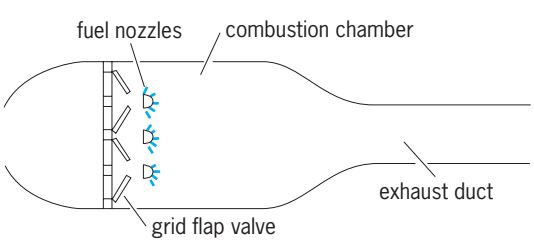

Fig. 1. Diagram of a pulse jet.

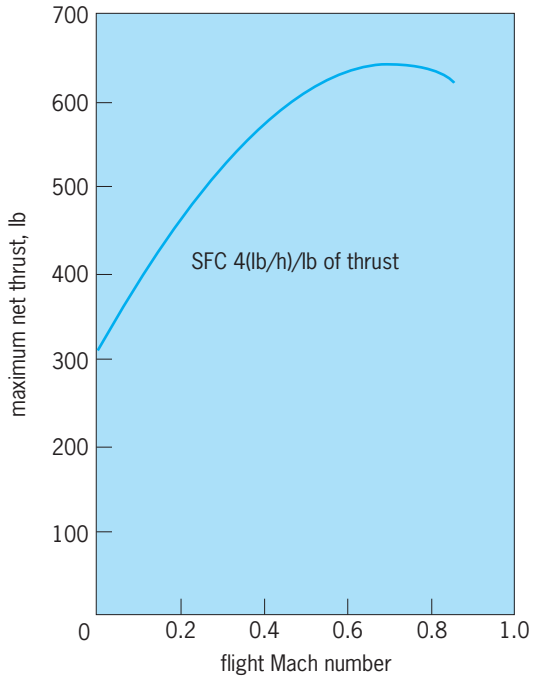

Fig. 2. Effect of flight Mach number on thrust of a German pulse jet; length 137.2 in. (3.49 m), diameter 21.6 in. (0.55 m). 1 lb = 4.448 N. SFC = specific fuel consumption. 4 (lb/h)/lb of thrust = 0.41 (kg/h)/N of thrust.

pressure in the combustion chamber has been dissipated, and this drops the combustion chamber pressure below atmospheric. As a result, the flat valves open and a fresh charge of air enters the combustion chamber. As this air flows past the fuel nozzles, it receives an injection of fuel and the mixture is then ignited by contact with the hot gas residue from the previous cycle. This causes the mixture to explode and the cycle repeats. Thrust increases with engine speed up to a maximum dependent on design (**Fig. 2**). *See* TURBINE ENGINE SUBSYSTEMS.

Unlike the ramjet, the pulse jet has an appreciable thrust at zero flight speed. However, as the flight speed is increased, the resistance to the flow of air imposed by the flap valves eventually causes substantial loss in performance and the pulse jet becomes less efficient than the ramjet. Failure of flap valves and valve seats by fatigue was found to be a problem.

In addition to their use on the German V-1 buzzbomb, pulse jets have been used to propel radio-controlled target drones and experimental helicopters. In the latter case, they were mounted on the blade tips for directly driving the rotor. The high fuel consumption, noise, and vibrations generated by the pulse jet limit its scope of applications. *See* PROPULSION. Benjamin Pinkel

Pulse modulation

A set of techniques whereby a sequence of information-carrying quantities occurring at discrete instances of time is encoded into a corresponding regular sequence of electromagnetic carrier pulses. Varying the amplitude, polarity, presence or absence,

duration, or occurrence in time of the pulses gives rise to the four basic forms of pulse modulation: pulse-amplitude modulation (PAM), pulse-code modulation (PCM), pulse-width modulation (PWM, also known as pulse-duration modulation, PDM), and pulse-position modulation (PPM).

Analog-to-digital conversion. An important concept in pulse modulation is analog-to-digital (A/D) conversion, in which an original analog (time- and amplitude-continuous) information signal $s(t)$ is changed at the transmitter into a series of regularly occurring discrete pulses whose amplitudes are restricted to a fixed and finite number of values. An inverse digital-to-analog (D/A) process is used at the receiver to reconstruct an approximation of the original form of $s(t)$. Conceptually, analog-to-digital conversion involves two steps. First, the range of amplitudes of $s(t)$ is divided or quantized into a finite number of predetermined levels, and each such level is represented by a pulse of fixed amplitude. Second, the amplitude of $s(t)$ is periodically measured or sampled and replaced by the pulse representing the level that corresponds to the measurement. *See* ANALOG-TO-DIGITAL CONVERTER; DIGITAL-TO-ANALOG CONVERTER.

Figure 1 shows an example of an analog signal and its corresponding pulse sequence resulting from the sampling and quantizing process. Here the dynamic range of the amplitude of the signal is assumed to extend from a minimum of 0 volt to a maximum of 7 volts, and the seven quantization levels are uniformly spaced and of magnitude 1 volt. Sampling occurs at the rate of one sample per second, and each sample value is represented by a pulse corresponding to the nearest quantization level and lasting the entire sampling interval of 1 second.

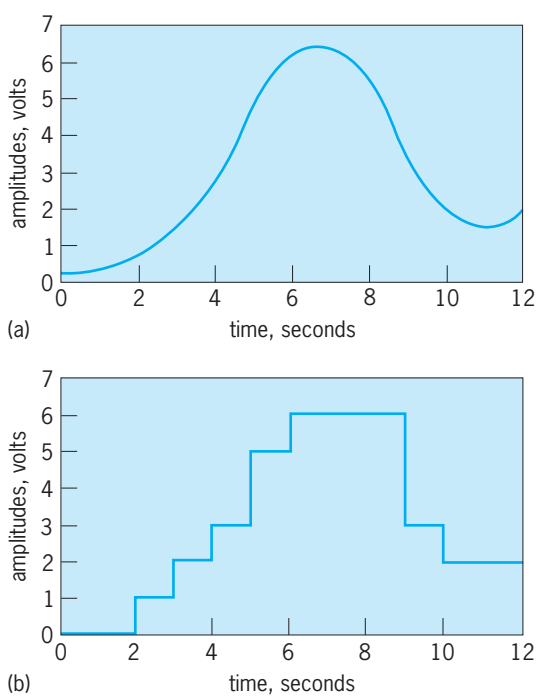

Fig. 1. Signal sampling and quantization. (*a*) Analog signal. (*b*) Corresponding pulse sequence.

According to the Nyquist sampling theorem, if sampling occurs at a rate at least twice that of the bandwidth of $s(t)$, the latter can be unambiguously reconstructed from its amplitude values at the sampling instants by applying them to an ideal low-pass filter whose bandwidth matches that of $s(t)$.

Quantization, however, introduces an irreversible error, the so-called quantization error, since the pulse representing a sample measurement determines only the quantization level in which the measurement falls and not its exact value. Consequently, the process of reconstructing $s(t)$ from the sequence of pulses yields only an approximate version of $s(t)$. Generally speaking, the size of the quantization error, and hence the distortion in the reconstructed version of $s(t)$, is related to the coarseness of the quantization scheme. It can thus be controlled by appropriately selecting the size of the quantization levels or the number of levels covering the dynamic range of the signal. In linear quantization the levels are chosen to be of constant size, whereas nonlinear quantization uses levels nonuniform in size, with the use of smaller levels for signal amplitudes that are particularly sensitive to distortion. But any increase in the number of levels covering a given dynamic range implies that the digital-to-analog conversion process must be able to distinguish between levels whose amplitudes differ by smaller amounts. This may result in an increase in the probability of error in determining the correct quantization level when the signals representing the levels have been corrupted by noise and distortion. *See* ELECTRIC FILTER; INFORMATION THEORY.

Pulse-amplitude modulation. In PAM the successive sample values of the analog signal $s(t)$ are used to effect the amplitudes of a corresponding sequence of pulses of constant duration occurring at the sampling rate. No quantization of the samples normally occurs (**Fig. 2***a*, *b*). In principle the pulses may occupy the entire time between samples, but in most practical systems the pulse duration, known as the duty cycle, is limited to a fraction of the sampling interval. Such a restriction creates the possibility of interleaving during one sample interval one or more pulses derived from other PAM systems in a process known as time-division multiplexing (TDM). *See* MULTIPLEXING AND MULTIPLE ACCESS.

Pulse-width modulation. In PWM the pulses representing successive sample values of $s(t)$ have constant amplitudes but vary in time duration in direct proportion to the sample value. The pulse duration can be changed relative to fixed leading or trailing time edges or a fixed pulse center. To allow for time-division multiplexing, the maximum pulse duration may be limited to a fraction of the time between samples (Fig. 2*c*).

Pulse-position modulation. PPM encodes the sample values of $s(t)$ by varying the position of a pulse of constant duration relative to its nominal time of occurrence. As in PAM and PWM, the duration of the pulses is typically a fraction of the sampling interval. In addition, the maximum time excursion of the pulses may be limited (Fig. 2*d*).

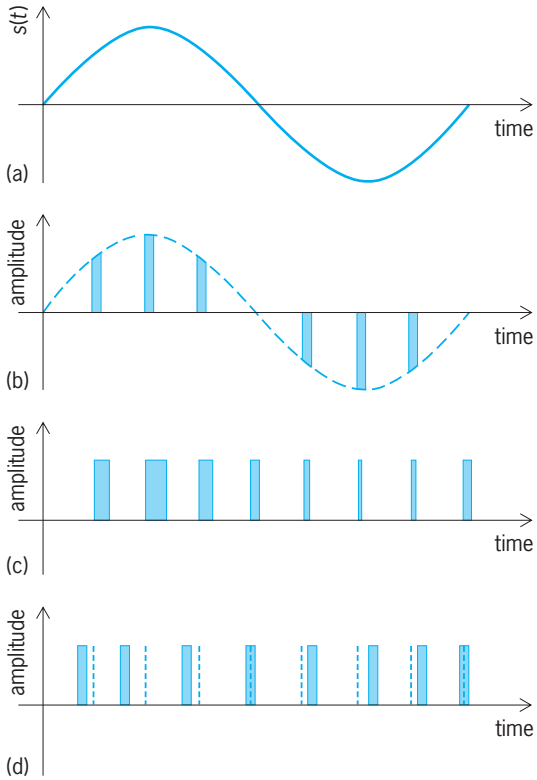

Fig. 2. Forms of pulse modulation for the case where the analog signal, *s*(*t*), is a sine wave. (*a*) Analog signal, *s*(*t*). (*b*) Pulse-amplitude modulation. (*c*) Pulse-width modulation. (*c*) Pulse-position modulation.

Pulse-code modulation. Many modern communication systems are designed to transmit and receive only pulses of two distinct amplitudes. In these so-called binary digital systems, the analog-to-digital conversion process is extended by the additional step of coding, in which the amplitude of each pulse representing a quantized sample of $s(t)$ is converted into a unique sequence of one or more pulses with just two possible amplitudes. The complete conversion process is known as pulse-code modulation.

Figure 3*a* shows the example of three successive quantized samples of an analog signal $s(t)$, in which sampling occurs every T seconds and the pulse representing the sample is limited to $T/2$ seconds. Assuming that the number of quantization levels is limited to 8, each level can be represented by a unique sequence of three two-valued pulses. In Fig. 3*b* these pulses are of amplitude V or 0, whereas in Fig. 3*c* the amplitudes are V and $-V$.

In general, if the original analog signal $s(t)$ has a bandwidth of B hertz, the number of samples per second must be at least $2B$ and the time between successive samples can be no greater than $T = 1/(2B)$ seconds. If the amplitudes of the samples are quantized into L levels, each level can be uniquely represented by a sequence of $\log_2 L$ binary pulses. Consequently the PCM process generates at least $2B \log_2 L$ binary pulses per second. Increasing the number of levels generally reduces the quantization error and therefore improves the fidelity of the reproduced analog signal, but also increases the number of pulses per second required to represent the signal.

PCM enjoys many important advantages over other forms of pulse modulation due to the fact that information is represented by a two-state variable. First, the design parameters of a PCM transmission system depend critically on the bandwidth of the original signal $s(t)$ and the degree of fidelity required at the point of reconstruction, but are otherwise largely independent of the information content of $s(t)$. This fact creates the possibility of deploying generic transmission systems suitable for many types of information. Second, the detection of the state of a two-state variable in a noisy environment is inherently simpler than the precise measurement of the amplitude, duration, or position of a pulse in which these quantities are not constrained. Third, the binary pulses propagating along a medium can be intercepted and decoded at a point where the accumulated distortion and attenuation are sufficiently low to assure high detection accuracy. New pulses can then be generated and transmitted to the next such decoding point. This so-called process of repeatering significantly reduces the propagation of distortion and leads to a quality of transmission that is largely independent of distance. In contrast, repeated amplification of analog signals along a transmission path results in the accumulation of the signal's distortion.

Baseband transmission. The modulated carrier pulses derived from any of the pulse modulation schemes may be transmitted directly, provided that the passband of the transmission medium includes the part of the electromagnetic spectrum that contains the energy of the pulses. In such a case the transmission system is referred to as baseband.

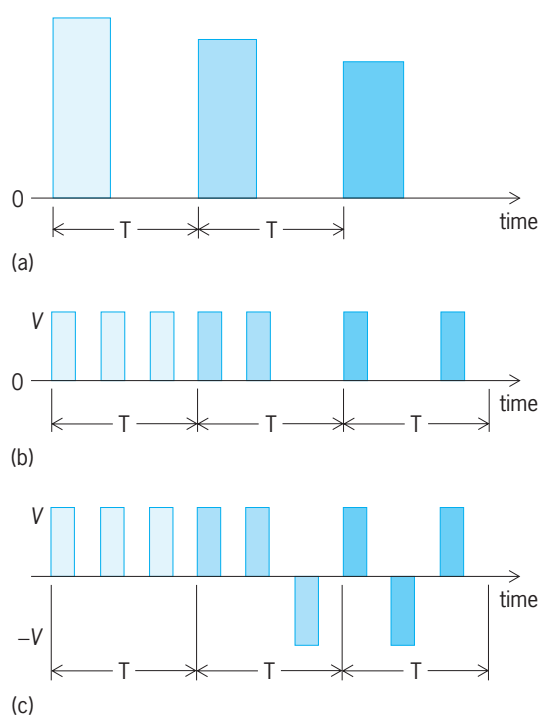

Fig. 3. Pulse-code modulation. (*a*) Three successive quantized samples of an analog signal. (*b*) With pulses of amplitude *V* or 0. (*c*) With pulses of amplitude *V* or −*V*.

Otherwise the modulated carrier pulses may in turn be used to modulate the amplitude, frequency, or phase of an analog carrier signal such as a sine wave in order to shift its energy to a higher frequency range appropriate for the medium.

PCM baseband transmission systems employ several different pulse patterns to represent the binary pulses obtained from the coding process. They are generally distinguished by their ability to provide timing content and error detection capability at the receiver, and by the amount of channel bandwidth required for their reliable transmission. Among the most important of these baseband codes are the nonreturn-to-zero (NRZ), biphase, and bipolar pulse codes. NRZ employs two pulses, typically of equal amplitude but opposite polarity, $+V$ and $-V$, each pulse lasting for the entire time between samples. Biphase, also known as Manchester coding, uses a two-pulse sequence for each sampling interval, with the polarity of the pulse changing from $+V$ to $-V$ or from $-V$ to $+V$ at the midpoint of the sampling interval. Bipolar, also known as alternate mark invert (AMI), employs two pulse amplitudes of equal and opposite polarity $+V$ and $-V$, and a third pulse of zero amplitude. Here $+V$ and $-V$ are used in an alternating manner to represent successive occurrences of one of the binary pulses, with the zero pulse always representing the other binary pulse.

The primary objective in any of these coding schemes is the provision of timing at the receiver by the observation of changes in the polarity of the pulses. In this respect, NRZ is inferior to the other forms due to the fact that polarity changes from one sampling interval to the other occur only when the underlying code pulse changes. The alternating polarity of the nonzero pulses of a bipolar code also provides a certain error detection capability, since the reception of two successive pulses of the same polarity must be the result of an error occurring during transmission. The occurrence of two successive pulses of the same polarity can also be used to signal the presence of a special condition in the information stream.

Time-division multiplexing. An advantage inherent in all pulse modulation systems is their ability to transmit signals from multiple sources over a common transmission system through the process of time-division multiplexing. By restricting the time duration of a pulse representing a sample value from a particular analog signal to a fraction of the time between successive samples, pulses derived from other sampled analog signals can be accommodated on the transmission system. For example, if five analog signals, each of bandwidth 100 hertz, are sampled at the Nyquist rate of 200 samples per second, then the time between successive samples from each signal is 5 milliseconds. By restricting the duration of a pulse to less than 1 millisecond, and by appropriate coordination of the timing of the samples, the pulses from all five signals can be transmitted during one sampling interval. Assigning the pulses from each signal to fixed positions in successive sampling intervals allows the demultiplexing of the composite pulse stream into its individual components by a simple counting process.

One important application of this principle occurs in the transmission of PCM telephone voice signals over a digital transmission system known as a T1 carrier. In standard T1 coding, an original analog voice signal is band-limited to 4000 hertz by passing it through a low-pass filter, and is then sampled at the Nyquist rate of 8000 samples per second, so that the time between successive samples is 125 microseconds. The samples are quantized to 256 levels, with each of them being represented by a sequence of 8 binary pulses. By limiting the duration of a single pulse to 0.65 microsecond, a total of 193 pulses can be accommodated in the time span of 125 microseconds between samples. One of these serves as a synchronization marker that indicates the beginning of such a sequence of 193 pulses, while the other 192 pulses are the composite of 8 pulses from each of 24 voice signals, with each 8-pulse sequence occupying a specified position (**Fig. 4**). T1 carriers and similar types of digital carrier systems are in widespread use in the world's telephone networks.

Bandwidth and distortion. An important consideration in any of the pulse modulation systems is the level of distortion experienced by a pulse in its transmission over a communications channel. In this regard a major effect is due to the band-limited nature of the channel. Theoretically, any pulse of finite duration, no matter what its shape, requires a channel of infinite bandwidth for undistorted reception. Since signals of infinite duration can in practice not be generated, and since the bandwidth of real channels is always finite, some distortion is unavoidable. The amount, however, depends on the duration T of a pulse and its amplitude variation, relative to the channel bandwidth B. In general, it can be said that for a given amount of distortion, pulses in which the amplitude changes abruptly or which are of shorter duration require larger channel bandwidths than pulses with continuously changing amplitudes or longer

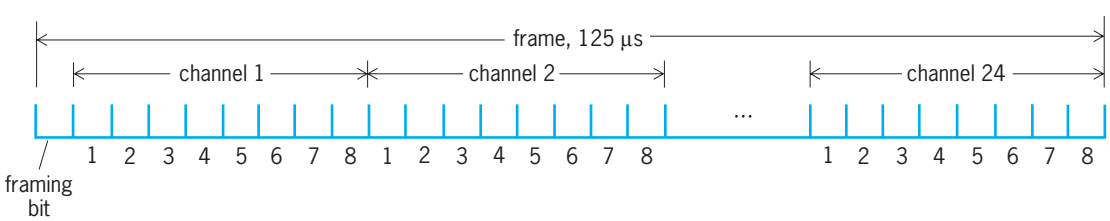

Fig. 4. Frame format for a T1 carrier, showing the arrangement of the pulses for one frame, corresponding to one sampling interval.

duration. As an example, the required bandwidth for a square pulse lasting T seconds is usually taken to equal $1/T$ hertz, whereas for a so-called raised cosine pulse, whose shape approximates that of a cosine wave, the value can be made as low as $1/(2T)$ hertz.

The required channel bandwidth for the transmission of pulses can be reduced by extending them in time over more than one sampling interval. However, such pulses incur intersymbol interference, in which the trailing part of one pulse overlaps with one or more subsequent pulses. Various pulse-shaping techniques that produce controlled amounts of this type of distortion, while at the same time limiting the amount of required channel bandwidth, have been developed. These are referred to as partial response codes. An important consequence of the distortion incurred by a pulse in a band-limited channel is the loss of timing content, since the spreading of the pulse in time obliterates sharp transitions in the voltage levels.

Pulses may also be corrupted by atmospheric noise produced by rain, sandstorms, galactic emissions, ionospheric fading, and other natural phenomena. Other major sources of distortion and interference external to the transmission channel are artificial, and include electric motors, and transmissions from radio and television stations in the same or adjacent spectral ranges. Both atmospheric noise and artificial interference manifest themselves largely in the form of random variations of the pulse amplitude and therefore affect the accuracy with which the quantization levels or sample values of the original signal can be reproduced at the receiver. *See* DISTORTION (ELECTRONIC CIRCUITS); ELECTRICAL INTERFERENCE; ELECTRICAL NOISE.

Bandwidth requirements. Pulse modulation systems may incur a significant bandwidth penalty compared to the transmission of an analog signal in its analog form. An example is the standard PCM transmission of an analog voice signal band-limited to 4000 hertz over a T1 carrier. Since the sampling, quantizing, and coding process produces 8 binary pulses 8000 times per second for a total of 64,000 binary pulses per second, the pulses occur every 15.625 microseconds. Depending on the shape of the pulses and the amount of intersymbol interference, the required transmission bandwidth will fall in the range of 32,000 to 64,000 hertz. This compares to a bandwidth of only 4000 hertz for the transmission of the signal in analog mode. *See* BANDWIDTH REQUIREMENTS (COMMUNICATIONS).

Pulse timing. An important consideration in the design of transmission systems based on any of the pulse modulation methods is the provision of pulse timing information at the receiver. Specifically, the receiver must have accurate knowledge of the points in time when pulses were generated at the sending side in order to determine properly their amplitude, position, or duration. In PAM, PWM, or PCM systems, the timing information usually identifies the beginning, middle, or end of a pulse, whereas in PPM the center of the unmodulated pulse may be indicated.

Pulse timing can be provided in three ways, referred to as autonomous, explicit, and implicit timing. In the autonomous method, timing is determined at the sender and receiver by means of local clocks that are periodically aligned from master reference clocks. These master clocks derive highly stable timing signals from the natural frequency of oscillation or resonance of certain materials such as quartz, cesium, and rubidium. In the explicit timing method, timing pulses are transmitted in parallel with the sequence of information-bearing pulses over a separate channel, and identify the moments in time when the information pulses were generated. Implicit timing relies on a local clock and the presence of timing information in the received pulse sequences themselves from which the local clock can be realigned. For example, in a 100% duty cycle NRZ-coded or AMI-coded pulse sequence, the occurrence of a transition in the voltage level indicates the end of one pulse and the beginning of another and provides an opportunity for the adjustment of the receiver clock. *See* ATOMIC CLOCK; QUARTZ CLOCK.

NRZ or AMI baseband codes suffer from the deficiency that certain information signals may generate long sequences of pulses without any level change, which may cause the receiver clock to drift. Several methods for dealing with this problem are found in operational systems. In so-called pulse stuffing, an extra pulse chosen to interrupt this pattern is inserted at the sending side and extracted at the receiving side after timing extraction. In the technique known as bit sequence transparency coding, a pulse sequence of specified length with no level changes is replaced with a predetermined pattern of pulses that normally does not occur as a sequence of information pulses but contains a high percentage of level changes. In scrambling, the original pattern of information pulses is deterministically altered prior to transmission in a manner that guarantees a sufficiently high incidence of level changes, and is restored at the receiver through an inverse scrambling process. The bipolar code avoids the difficulty altogether in that it guarantees a change in level at the midpoint of every pulse.

Frame synchronization. In TDM systems, besides pulse timing, the additional problem of frame synchronization arises. Here the receiver, in order to demultiplex the pulses in a frame (that is, the pulses received during a sampling interval) must have information relating to the beginning of each such sampling interval. Demultiplexing then can proceed by a simple counting process.

In one method of frame synchronization in widespread use, the sender transmits at the beginning of each frame of information pulses a specially selected pattern of pulses. Synchronization is indicated when a copy of this pattern stored at the receiver matches pulse-by-pulse with a received sequence of pulses. So-called pseudonoise (PN) sequences, whose pulse patterns simulate the random occurrence of binary pulses, and a set of pulse patterns known as Barker codes, find frequent application in frame synchronization.

A somewhat different approach is used in T1

carrier systems, which rely on a single pulse transmitted at the beginning of each frame whose polarity alternates from one frame to the next, that is, every 193 pulses. The receiver derives frame synchronization from the observation of this alternating pattern over several consecutive frames.

In any of the frame synchronization methods, special provision must be made to prevent the appearance of the synchronization pattern in the information portion of the frame, for in such cases false frame synchronization may result. Various techniques have been developed for altering the information sequence when such patterns are detected.

Applications. PAM, PWM, and PPM found significant application early in the development of digital communications, largely in the domain of radio telemetry for remote monitoring and sensing. They have since fallen into disuse in favor of PCM.

Since the early 1960s, many of the world's telephone network providers have gradually, and by now almost completely, converted their transmission facilities to PCM technology. The bulk of these transmission systems use some form of time-division multiplexing, as exemplified by the 24-voice channel T1 carrier structure, and are intended for short-haul transmission. Higher-capacity TDM facilities such as the T1-C, T2, T3, and T4 carriers, which combine 48, 96, 672, and 4032 digital voice channels, respectively, have also been deployed in large quantities to provide connections between large central office switches and toll office switches, and between subscribers and central offices over subscriber lines. Internationally, carrier systems referred to as E1, E2, E3, and E4, which combine 30, 120, 480, and 1920 voice channels, respectively, are in widespread use. These carrier systems are implemented over many types of transmission media, including twisted pairs of telephone wiring, coaxial cables, fiber-optic cables, and microwave. *See* COAXIAL CABLE; COMMUNICATIONS CABLE; MICROWAVE; OPTICAL COMMUNICATIONS; OPTICAL FIBERS; SWITCHING SYSTEMS (COMMUNICATIONS); TELEPHONE SERVICE.

The deployment of high-speed networks such as the Integrated Service Digital Network (ISDN) in many parts of the world has also relied heavily on PCM technology. PCM and various modified forms such as delta modulation (DM) and adaptive differential pulse-code modulation (ADPCM) have also found significant application in satellite transmission systems. *See* COMMUNICATIONS SATELLITE; DATA COMMUNICATIONS; ELECTRICAL COMMUNICATIONS; INTEGRATED SERVICES DIGITAL NETWORK (ISDN); MODULATION.

Hermann J. Helgert

Bibliography. B. P. Lathi, *Modern Digital and Analog Communication Systems*, 3d ed., Oxford University Press, 1998; W. D. Reeve, *Subscriber Loop Signaling and Transmission Handbook: Digital*, IEEE Press, New York, 1995; M. Schwartz, *Information Transmission, Modulation, and Noise*, 4th ed., McGraw-Hill, New York, 1990; D. R. Smith, *Digital Transmission Systems*, 2d ed., Van Nostrand Reinhold, New York, 1993.

Pulse modulator

A device that converts an analog information signal to a pulse format. A pulse modulator encodes the amplitude information at a particular time by a set of short- or long-duration pulses using such methods as pulse-amplitude, pulse-width, pulse-position, pulse-density, or pulse-code modulation. A pulse modulator, therefore, accepts a continuous analog input signal and provides a pulse output that conveys the amplitude of the continuous analog signal as seen at an observation, or sampling, time t_s, often optimistically described as a sampling instant.

Pulse-amplitude modulation. A continuous analog signal, when viewed through equally spaced time windows, provides a pulse-amplitude modulation (PAM) signal. This is analogous to viewing the continuous signal through the vertical slats in a picket fence. The signal is sampled at equally spaced time intervals of T_s, the sampling period. The reciprocal of T_s is the sampling frequency f_s. If the highest frequency in the continuous analog signal is $f_{\max}$, then in order to avoid loss of information through sampling, the sampling theorem dictates that $f_s > 2 \cdot f_{\max}$. The duration of each sampling window, during which time we look at the analog signal, is the aperture time T_a (the spacing between the slats). The duty cycle, T_a/T_s, represents the fraction of the time that the PAM signal is present. The time that the PAM signal is absent (idle time) can be profitably used to multiplex $N = T_s/T_a$ different PAM signals together. Channel efficiency is improved by time-division-multiplexed (TDM) transmission, or the interleaving of N analog signals over a single channel. Multiple analog instrumentation signals are frequently sent to a remote recorder or monitor via a PAM signaling system. *See* INFORMATION THEORY; MULTIPLEXING AND MULTIPLE ACCESS; TELEMETERING.

A pulse-amplitude modulator is simply a transmission gate, or switch, that opens every T_s seconds for duration of T_a seconds to produce a PAM signal. On the other hand, a sample-and-hold circuit may be used. The value of the signal detected at the sampling instant is clamped or "held" until the next sample instant by a sample-and-hold (S&H) device.

The preceding discussion concerns baseband PAM. If the transmited signal is then multiplied by a sinusoidal carrier, it becomes passband PAM.

Phase modulator. A close relative of the pulse-amplitude modulator is the phase-shift-keyed (PSK) modulator. The amplitude of the signal is held constant, but the information is provided in the phase of a carrier. PSK and PAM can be combined to form QAM (quadrature-amplitude-modulated) signals, discussed below. *See* PHASE MODULATION.

Pulse-width modulation. Unlike a PAM signal, the amplitude of each pulse-width modulation (PWM) pulse is constant. The width τ of each output pulse carries the amplitude information of the continuous analog signal during the observation interval. The input signal $x(t)$ may be sampled and "held" or clamped to its value at the sampling instants, thereby generating a stair-step waveform; or it may simply

be used in its continuous form. Most PWM signals are generated by combining the stair-step or continuous analog signal with a triangular or sawtooth waveform in a thresholding device that provides an output of +1 when the sum is positive, and an output of zero otherwise if the input signal is unipolar (binary PWM), that is, if $0 \le x(t) \le 1$. If, for example, the sawtooth function is $-t/T_s$ for $0 \le t \le T_s$, the summed signal $x(t) - t/T_s$ reaches zero at $\tau = T_s \cdot x(\tau)$. Since the threshold circuit output is +1 until $t = \tau$ and then drops to zero, the average value of the PWM output is $\tau/T_s = x(\tau)$. If $-1 \le x(t) \le -0$, we have simply flipped the polarity of the input, and we can similarly flip the polarity of the ramp to make a PWM modulator for negative in put signals. If a −1 is produced in response to a negative input, the two systems can be tied in parallel to provide a ternary PWM system.

An alternative but less power-efficient form of binary PWM outputs a +1 or a −1 (never a 0), maintaining a correct average value over the pulse period. PWM is attractive because the signal can be clamped to amplitudes of +1, −1, and zero, making it very insensitive to additive amplitude noise.

PWM is especially useful in applications where power, as opposed to information or data, is being controlled because electrical power-switching devices are more power-efficient than proportional-control devices. Switched-device resistance is either very low or near-infinite. In dc-motor speed-control operation, PWM motor-drive signals virtually eliminate the static friction (stiction) of motor bearings and mechanical linkages because the torque is applied abruptly. PWM is also used for voltage regulation in switched power supplies, light dimmers, and class D amplifiers for audio applications. *See* ELECTRONIC POWER SUPPLY.

Pulse-position modulation. This is now a seldom-used format. The amplitude information of the analog input signal is carried by the location of a constant-amplitude, constant-duration pulse whose location along the time axis carries the input information. Pulse-position modulation (PPM) involves inexpensive electronics, but is quite susceptible to additive noise. A PPM signal can be generated by differentiating, or high-pass-filtering, a PWM signal.

Pulse-density modulation. The pulse-density modulation (PDM) pulse rate during the sampling period is proportional to the amplitude of the analog input signal, as is its average value. A PDM signal may be generated by driving a voltage-controlled oscillator with the analog input signal, clipping the output to generate a square wave, then differentiating to produce a pulse train whose rate, or density, varies proportionally with the analog input voltage. *See* OSCILLATOR.

Pulse-code modulation. A pulse-code modulator produces numbers, rather than the analog values that are produced by all of the pulse-modulation methods discussed so far, which quantize the signal in time only. In those methods, the signal amplitude is specified at sampling instants t_s (quantized in time), but the signal amplitude can take on any values between the signal maximum and minimum values. For PCM, the signal is quantized also in amplitude steps of size $q = 2^{-B}$, where B is the number of bits used to represent the sample of the amplitude of the continuous analog signal $x(t)$, and it is assumed that $0 \le x(t) \le 1$. The parameter B determines the size of the granularity of the quantization step as well as the error. In binary notation, the signal quantization is given by Eq. (1), where the b_n are the bits 0 or 1 used

$$x(t_s) \approx \sum_{1}^{B} b_n \cdot 2^{-n} \tag{1}$$

to approximate $x(t_s)$. The peak quantization error ε does not exceed $\varepsilon_{\max} = {}^1/_2 q$. Since the quantization error is equally likely to fall anywhere between $-{}^1/_2 q$ and ${}^1/_2 q$, the mean or average squared quantization error (MSE) is given by Eq. (2). The RMS (root-

$$\frac{1}{q}\int_{-q/2}^{q/2} \varepsilon^2 d\varepsilon = q^2/12 = 2^{-2B}/12 \tag{2}$$

mean-squared) error is, of course, the square root of the MSE. This is an additive quantization noise. As B grows, the peak and RMS errors fall. for example, in the quantization of an analog signal sample to 4 bits, that is, $B = 4$, the quantization step size is $q = 2^{-4} = 1/16$, the peak quantization error is $\varepsilon_{\max} = {}^1/_2 \cdot 2^{-4} = 1/32$, and the RMS error is $\varepsilon_{\max}/\sqrt{3} = 0.018$. Now, suppose $x(t_s) = 0.6875 = 1 \cdot 2^{-1} + 0 \cdot 2^{-2} + 1 \cdot 2^{-2} + 1 \cdot 2^{-4}$. The bits are $b_1 = 1$, $b_2 = 0$, $b_3 = 1$, and $b_4 = 1$. That string of bits (or data word), 1011, forms the PCM-coded representation of $x(t_s)$.

In the representation of signed data, positive numbers have a zero as a sign bit appended to the left of the most significant data bit. Conversion to a negative number in two's complement format (the most common) is accomplished easily by copying the data, from right to left, through the first "one," then complementing (making ones into zeros and vice versa) all bits to the left. For the example above, with its sign bit, the PCM form of $x(t_s)$ is 01011; its negative is 10101. *See* NUMERICAL REPRESENTATION (COMPUTERS).

For additional robustness against noise, the digital format is often not straight binary, but reflected binary or Gray code, in which any two adjacent numerical values can differ by only one bit.

Corresponding to each sampling instant, the PCM produces one data word. A PCM signal is said to be digitized because its quantized amplitude values are in digital format. The PCM device itself is usually known as an analog-to-digital converter. A digitized signal is reasonably noise-resistant and can be modulated, demodulated, filtered, and otherwise processed more accurately than an analog signal. Many PCM data streams may be time-division-multiplexed together for transmission. *See* ANALOG-TO-DIGITAL CONVERTER.

Delta modulation (DM) or delta PCM (DPCM). DPCM conversion, in which the difference, rather than the actual values, in amplitudes of successive samples is encoded and transmitted, is often more efficient than PCM. This is particularly so if the signal is oversampled, because large sample-to-sample changes in

amplitude are unlikely to occur. DPCM attempts to approximate the input signal by a sequence of positive or negative steps of fixed size. This method is simple but is degraded by "slope overload," analogous to saturation, if the DPCM is unable to keep up with the changes of the input signal. Because only the error is encoded, fewer bits are necessary, permitting a lower transmission data rate, or improved signal-to-noise ratio. For example, a reduction in word-length requirement can be translated to a faster transmission rate or a 6-dB reduction in quantization noise.

Adaptive delta modulation. This method avoids the slope overload of DM by adjusting the step size according to the demands of the input signal. One adaptive delta modulation (ADM) method, the Song algorithm, begins with some predetermined minimum step size to follow the input signal. At each sample time the step size is either increased or decreased by one unit to reduce the error. The Space Shuttle algorithm is a similar ADM algorithm; however, if the error switches polarity, the step size drops back to the minimum.

Differential PCM. Differential PCM requires memory to predict the next sample value based on fixed prediction-coefficient values and the present and past input samples, and then transmits the prediction error.

Adaptive differential PCM (ADPCM). This is an even more sophisticated method of encoding the signal, in which the coefficients of the predictor are recomputed every sample time. *See* DATA COMPRESSION.

Quadrature amplitude modulation (QAM). QAM (or quadrature PAM) is a unique format because it simultaneously modulates two carriers of the same frequency, but 90° out of phase with one another. This can be viewed as simultaneous modulation of both amplitude (PAM) and phase (PSK), affording an economical and robust means for transmission of a pair of PCM signals.

In two-dimensional signal space, the allowable amplitudes of the values (numbers) modulating the cosine carrier and the sine carrier may be plotted along the I (in-phase or real) and Q (quadrature or imaginary) axes, respectively. This array of M points is known as a constellation. Constellation shapes are usually square [$M = 2 \times 2 = 4$, $M = 3 \times 3 - 1$ (a square with the origin removed), $M = 4 \times 4 = 16$, $M = 6 \times 6 - 4 = 32$ (a square with the 4 corners lopped off); or $M = 16 \times 16 = 256$]; sometimes rectangular, such as $M = 2 \times 4 = 8$; or concentric circles with the points on radials at multiples of 45°. However, any configuration may be used.

In the simplest form, a stream of usually Gray-coded PCM data words are concatenated (laid end-to-end) to form a flowing jumble of zeros and ones. The odd-numbered bits are directed to processor A, the even-numbered bits to processor B. The resulting two data streams are chopped up into two sets of C-bit synthetic data words. To the tail end of each synthetic word an extra one is appended as a least significant bit (LSB). The leading, or most significant, bit (MSB) of each word is complemented; that is, each one become a zero and each zero becomes a one. The MSBs will be interpreted as sign bits. Now, for each pair of $(C + 1)$-bit newly created synthetic words, one from A (the "real" or in-phase component) and one from B (the "imaginary" or quadrature component), we can plot a point in the complex signal space described above. An evenly spaced array of points appears in a $C \times C$ square array in signal space centered on the origin. These new synthetic words are converted to analog values by a pair of digital-to-analog converters. The analog real words are amplitude-modulated onto (multiplied by) a carrier $\cos(2\pi f_c t)$, where f_c is the carrier frequency; similarly, the analog imaginary words are modulated into a second but out-of-phase carrier, $\sin(2\pi f_c t)$. The two modulated curriers are in quadrature, that is, 90° out of phase with one another, to minimize interference between the real and imaginary components. Each of the complex modulated carrier segments that contain $2C$ of the original data bits is called a symbol, and the rate at which they are transmitted is called the symbol rate. *See* DIGITAL-TO-ANALOG CONVERTER.

Stanley A. White

Bibliography. L. W. Couch II, *Digital and Analog Communication Systems*, Prentice Hall, Upper Saddle River, NJ, 2001; S. Gibilisco, *Handbook of Radio and Wireless Technology*, McGraw-Hill, New York, 1999; J. G. Proakis, *Digital Communications*, 4th ed., McGraw-Hill, New York, 2001; T. S. Rappaport, *Wireless Communications*, 2d ed., Prentice Hall, Upper Saddle River, NJ, 2002; M. S. Roden, *Analog and Digital Communications Systems*, 4th ed., Prentice Hall, Englewood Cliffs, NJ, 1995; B. Sklar, *Digital Communications: Fundamentals and Applications*, 2d ed., Prentice Hall, Upper Saddle River, NJ, 2001; M. E. Van Valkenburg and W. M. Middleton (eds.), *Reference Data for Engineers: Radio, Electronics, Computer and Communications*, 9th ed., Butterworth-Heinemann, Woburn, MA, 2002.

Pumice

A rock froth, formed by the extreme puffing up (vesiculation) of liquid lava by expanding gases liberated from solution in the lava prior to and during solidification. Some varieties will float in water for many weeks before becoming waterlogged. Typical pumice is siliceous (rhyolite or dacite) in composition, but the lightest and most vesicular pumice (known also as reticulite and thread-lace scoria) is of basaltic composition. *See* LAVA; VOLCANIC GLASS.

Gordon A. MacDonald

Pump

A machine that draws a fluid into itself through an entrance port and forces the fluid out through an exhaust port (see **illus.**). A pump may serve to move liquid, as in a cross-country pipeline; to lift liquid, as from a well or to the top of a tall building; or

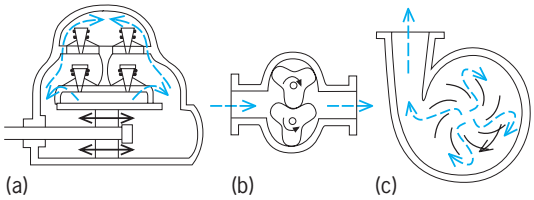

Pumps. (*a*) Reciprocating. (*b*) Rotary. (*c*) Centrifugal.

to put fluid under pressure, as in a hydraulic brake. These applications depend predominantly upon the discharge characteristic of the pump. A pump may also serve to empty a container, as in a vacuum pump or a sump pump, in which case the application depends primarily on its intake characteristic. *See* CENTRIFUGAL PUMP; COMPRESSOR; DISPLACEMENT PUMP; FAN; FUEL PUMP; PUMPING MACHINERY; VACUUM PUMP. Elliott F. Wright

Bibliography. B. Glanfield (ed.), *Pump Technology New Challenges, Where Next*, 1989; T. G. Hicks, *Pump Operation and Maintenance*, 1958, reprint 1983: Institute for Power System, *Pumping Manual*, 9th ed., 1995; J. Karassik et al., *Pump Handbook*, 3d ed., 2000; F. Pollak, *Pump User's Handbook*, 3d ed, 1988; A. J. Stepanoff, *Centrifugal and Axial Flow Pumps: Theory, Design and Application*, 2d ed., 1992; H. L. Stewart, *Pumps*, 3d ed., 1978; R. H. Warring, *Pumps Selection, Systems and Applications*, 1982.

Pumped storage

A process, also known as hydroelectric storage, for converting large quantities of electrical energy to potential energy by pumping water to a reservoir at a higher elevation, where it can be stored indefinitely and then released to pass through hydraulic turbines and generate electrical energy. An indirect process is necessary because electrical energy cannot be stored effectively in large quantities. Storage is desirable, as the consumption of electricity varies throughout the day, between weekday and weekend, as well as among seasons. Consequently, much of the generating equipment needed to meet the largest daytime load is unused or lightly loaded at night or on weekends. During those times the excess generating capacity can be used to generate energy for pumping, hence the necessity for storage. Normally, pumping energy can be obtained from economical sources, and its value will be upgraded when used for peak loads.

Operation. In a typical operation, electrical energy generated during the night or weekend is used to pump water from a lower to a higher elevation, where it is stored as potential energy in the upper reservoir. The water can be retained indefinitely without deterioration or significant loss. During the daylight hours when the loads are greatest, the stored water is released to flow from the higher to the lower reservoir through hydraulic turbines that drive electric generators and is converted to electrical energy. No water is consumed in either the pumping or generating phase. To provide storage or generation merely requires the transfer of water from one reservoir to the other. Pumped storage installations have attained an overall operating efficiency of about 70%. Projected improvements in equipment design promise an efficiency of 75% or more. Postulating one cycle each of generation and pumping per day plus an allowance to change from one mode to the other, the maximum annual generation attainable is 3500 h. *See* HYDRAULIC TURBINE; HYDROELECTRIC GENERATOR; PUMPING MACHINERY.

Description. A typical pumped-storage development consists of two reservoirs of essentially equal volume situated to maximize the difference between their levels. These reservoirs are connected by a system of waterways along which a pumping-generating station is located (**Fig. 1**). Under favorable geological conditions, the station will be located underground; otherwise it will be situated at the lower reservoir. The principal equipment of the station is the pumping-generating unit. In United States practice, the machinery is reversible and is used for both pumping and generating; it is designed to function as a motor and pump in one direction of rotation and as a turbine and generator in opposite rotation. Transformers, a substation, switchyard, and transmission line are required to transfer the electrical power to and from the station. *See* ELECTRIC POWER GENERATION.

The lower reservoir may be formed by impounding a stream or by using an existing body of water. Similarly, an upper reservoir may be created by an embankment across a valley or by a circumferential dike. Frequently, these features produce the most significant environmental impact, which is largely land use and visual. The reservoirs are comparatively small, thus affording some latitude in location to minimize unavoidable effects, such as displacement of developed areas, existing roads, and railways. Problems of emission of particulate matter and gases and of water-temperature rise associated with other generating stations do not exist. On the other hand,

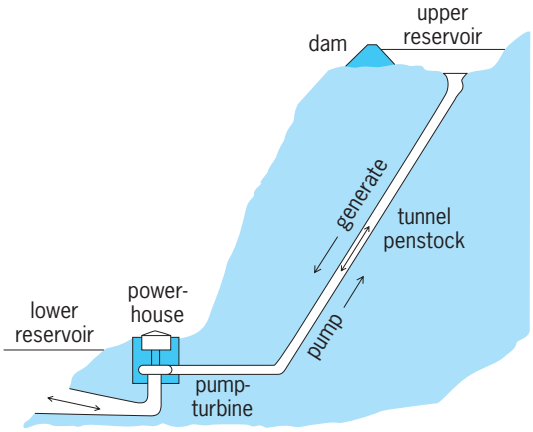

Fig. 1. Diagram of a conventional pumped-storage development.

the reservoirs and surrounding area afford the opportunity to develop recreational facilities such as camp grounds, marinas, boat ramps, picnic areas, and wildlife preserves. In the United States, recreation commands a high priority, and many existing developments have included recreational facilities.

Economics. The economics of pumped storage depend on the system in which it operates. The development must be economically competitive with other types of generation available, namely, nuclear, coal- or oil-fired, gas turbines, hydro, and other storage systems such as compressed air. It has been a generally accepted practice to evaluate the economy of any plant on the basis of its annual cost to the system, but life-cycle costs are now also considered. The annual cost is defined as the sum of fixed costs, operation, maintenance, and fuel. When construction is complete, the fixed costs are established and continue at that level irrespective of the extent of operation. Variations in fuel prices have a major effect. *See* ENERGY SOURCES.

European plants. In European practice, not only system economies receive attention, but also operating requirements and advantages; the latter have materially influenced the design and selection of equipment. European pumped-storage plants have been used not only to provide peak-load power as in the United States, but also to ensure system stability or frequency control. To accomplish this, the units must have very short response times which will enable them to change mode of operation quickly, to follow changes in load, or to provide emergency generation or load. Such operation is the dominant reason for European preference for three-machine sets, that is, separate pump and separate turbine coupled to a motor-generator, as opposed to the two-machine reversible pump-turbine and motor-generator used in the United States. The extent to which European practice incorporates system stability and frequency regulation is shown by the number of pumping and generating starts per day. In one plant, the daily average starts per unit for pumping was 2.5 and for generating 3.5, compared to 1 start for pumping and 1 or possibly 2 per day for generation in typical United States operation.

Deep-underground storage. The deep-underground concept (**Fig. 2**) is important because of high fuel and construction costs for alternate generation, and the flexibility it affords in selecting a site near load centers. Much research has been directed toward the deep-underground scheme owing to its reduced environmental effect and freedom from topographic restraints. The deep-underground concept is similar to conventional pumped storage, having all of the essential features and utilizing similar equipment. There are two notable variations: the upper reservoir is at ground level, while the lower reservoir is a deep underground cavern. Two stages are developed in series to minimize the size and, hence, the construction cost of the lower reservoir. Inherent advantages of the concept are: material reduction in environmental impact since only the upper reservoir is visible, development of heads which utilize the maximum capability of the machines, and elimination of the need for favorable topography. The disadvantages are: the need for large areas to dispose of the excavated material, uncertainties of construction and mining at 3000 ft (900 m) or more below the surface, and substantially increased construction time and costs compared with conventional pumped storage. *See* ENERGY STORAGE; WATERPOWER.

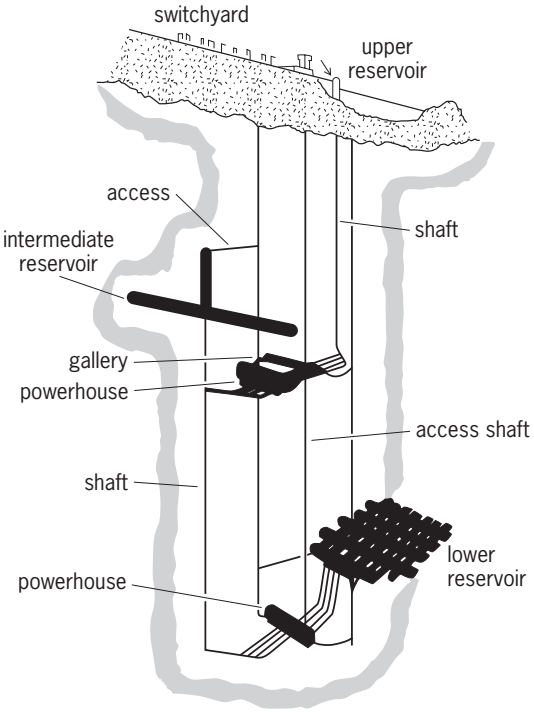

Fig. 2. Diagram of two-stage underground pumped-storage development.

Dwight L. Glasscock

Bibliography. American Society of Civil Engineers, *Civil Engineering Guidelines for Planning and Designing Hydroelectric Developments*, vol. 5: *Pumped Storage and Tidal Power*, 1989; American Society of Civil Engineers, *Compendium of Pumped Storage Plants in the United States*, 1993; M. G. Jog, *Hydroelectric and Pumped Storage Plants*, 1989.

Pumping machinery

Devices that convey fluids, chiefly liquids, from a lower to a higher elevation or from a region of lower pressure to one of higher pressure. Pumping machinery may be broadly classified as mechanical or as electromagnetic.

Mechanical pumps. In mechanical pumps the fluid is conveyed by direct contact with a moving part of the pumping machinery. The two basic types are (1) velocity machines, centrifugal or turbine pumps, which impart energy to the fluid primarily by increasing its velocity, then converting part of this energy into pressure or head, and (2) displacement machines with plungers, pistons, cams, or other

confining forms which act directly on the fluid, forcing it to flow against a higher pressure. A pump deep in a well may raise water or oil to the surface. At a ground level location a pump may deliver fluid to a nearby elevated reservoir or, through long pipelines, to a location at similar or different elevation. In a power plant, pumps circulate cooling water or oil at low pressures and transfer water from heaters at moderate pressure to steam generators at pressures of several thousand pounds per square inch. In chemical plants and refineries pumps transfer a great variety of fluids or charge them into reactors at higher pressure. In hydraulic systems, pumps supply energy to a moving stream of oil or water, which is readily controlled, to move a piston or press platen or to rotate a shaft as required by the specific process. *See* CENTRIFUGAL PUMP; DISPLACEMENT PUMP.

Electromagnetic pumps. Where direct contact between the fluid and the pumping machinery is undesirable, as in atomic energy power plants for circulating liquid metals used as reactor coolants or as solvents for reactor fuels, electromagnetic pumps are used. There are no moving parts in these pumps; no shaft seals are required. The liquid metal passing through the pump becomes, in effect, the rotor circuit of an electric motor. The two basic types are conduction and induction.

The conduction type, used with either direct or alternating current, confines the liquid metal in a narrow passage between the field magnets. Electrodes on each side of this channel apply current through the liquid metal at right angles to the magnetic field and the direction of flow. This current path is like the flow of current in the armature winding of a motor.

The induction or traveling field type operates only on polyphase alternating current. The liquid metal is confined in a thin rectangular or annular passage thermally insulated from the slotted stator. This stator with its windings is similar to the stator of a squirrel-cage motor cut through on one side and rolled out flat. The traveling field induces currents in the liquid metal similar to the currents in a motor armature. *See* ELECTROMAGNETIC PUMP.

Elliott F. Wright

Bibliography. N. P. Chopay (ed.), *Fluid Movers*, 1994; T. G. Hicks, *Pump Operation and Maintenance*, 1958, reprint 1983; Institute for Power System, *Pumping Manual*, 9th ed., 1995; I. J. Karassik et al., *Pump Handbook*, 3d ed., 2000; F. Pollak, *Pump User's Handbook*, 3d ed., 1988; A. J. Stepanoff, *Centrifugal and Axial Flow Pumps: Theory, Design and Application*, 2d ed., 1992; H. L. Stewart, *Pumps*, 3d ed., 1978; R. H. Warring, *Pumps Selection, Systems and Applications*, 1982.

Pumpkin

The term commonly applied to the larger, orange-colored fruit of the *Cucurbita* species, used when ripe as a table vegetable, in pies, or for autumn decoration. The flesh is somewhat coarse and strongly flavored, hence generally not served as a baked vegetable. Although some taxonomists would restrict the term pumpkin to the species *Cucurbita pepper* and *C. moschata*, it is also used in referring to *C. mixta*. Popular varieties (cultivars) are Connecticut Field and Small Sugar. Canned pumpkin is usually made from a blend of pumpkins and winter squashes. Cultural practices are similar to those used for squash. Harvesting generally begins when the fruits are mature, usually 4 months after planting. New Jersey, Illinois, and California are important producing states. *See* SQUASH; VIOLALES. H. John Carew

Puna

An alpine biological community in the central portion of the Andes Mountains of South America. Sparsely vegetated, treeless stretches cover high plateau country (altiplano) and slopes of central and southern Peru, Bolivia, northern Chile, and northwestern Argentina. The poor vegetative cover and the puna animals are limited by short seasonal precipitation as well as by the low temperatures of high altitudes. In spite of severe climate and limited use, some portions are inhabited by native Indians.

Altitudinal zonation. Like the paramos of the Northern Andes, punas occur above timberline, and extend upward, in modified form, to perpetual snow. Due to greater heights of the Central Andean peaks, aeolian regions, that is, regions supplied with airborne nutrients above the upper limit of vascular plants, generally the snowline, are more extensive here than above the paramos. Vascular plants reach higher levels in these low-latitude regions, around 30° south of the Equator, than in the paramos, which are equatorial. Thus the snowline, at about 15,700 ft (4800 m) in southern Peru, is above that of the paramos and is still higher farther south. Similar difference in upper limits of plants is found north of the Equator in the Himalaya Mountains. *See* PARAMO.

The altiplano is above intermountain basins, such as the Peruvian Cuzco area at 11,000 ft (3400 m), these basins being well vegetated and intensely cultivated. Some authorities place the beginning of the puna at 13,000 ft (4000 m), approximately the limit of cultivated crops, and call the uppermost slopes with very few plants "cold deserts."

Plant life-forms and species. The low-growing widely spaced plants with considerable bare ground between them, and the lack of much green color most of the year give the vast puna areas a bleak appearance (**Fig. 1**). The appearance of the puna is due not to severe temperatures, but to a short wet season; almost all precipitation, including snow and hail, occurs from January through March and causes these areas to become green. The average temperature of the Bolivian altiplano is 44°F (7°C). Temperatures below zero are rare. A striking contrast with the paramo is the alkaline nature of the puna soil, the paramo being consistently acid. Some shallow

intermittent lakes are used for salt collection by the natives.

Plant adaptations are indicative of the predominantly dry conditions. Bunch grasses, many with stiff needle leaves, are small; xerophilous herbs and resinous shrubs grow as rosettes, cushions, or mats; and succulent cacti occur in dense cushions. A good indicator plant is *Pycnophyllum* of the pink family, Caryophyllaceae, which forms pale-green mounds of densely packed stems covered with small, imbricated leaves (**Fig. 2**). These masses are locally called yareta, although it would seem better to restrict this term to large *Azorella* cushions of the celery family, Umbelliferae, said to resemble the "vegetable sheep" of New Zealand and to attain larger size in the eastern, somewhat wetter portion of the puna. Many small prostrate shrubs, chiefly of composite family, such as *Baccharis caespitosa* and *Senecio adenophyllus*, add to the scanty ground cover and are used, as is the yareta, for altiplano fuels. *Puya raimondii* of the pineapple family, Bromeliaceae (**Fig. 3**), exemplifies the rosette habit so prominent in *Espeletia* of the paramos. This unusual endemic, with heavily cutinized leaves, occurs in a few isolated sites in Peru and Bolivia. On stream banks and occasional seepages, *Distichia muscoides* of the rush family, Juncaceae, carpets the ground. Low shrub thickets called the tola belt occur along the western margin of the puna. This belt borders upon a desert in characteristics but is floristically related to the puna.

Fig. 1. Vegetation of the puna, with low bunch grass, prostrate shrubs, and cushion growths of cacti appearing white because of dense hairs. This view is from the Lake Titicaca area, Bolivia.

Fig. 2. Mat of *Pycnophyllum*, typical plant of the puna, on rocky soil; Sierra Negra, central Peru.

Fig. 3. *Puya raimondii*, a rare endemic of Peru and Bolivia; Sierra Negra, central Peru.

Fauna. The puna is interesting as the habitation of the American species of the camel family: llamas (*Lama glama*), beasts of burden since ancient Inca times; alpacas (*L. pacos*); and vicuñas (*Vicugna vicugna*), prized for hides and wool. The vicuña and another species, the guanaco (*L. guanicoe*), are wild. All live chiefly upon grasses, such as *Calamagrostis brevifolia, C. vicunarum*, and *Stipa ichu*, and herbs. Their heavy fleece adapts them to the rapid radiation changes which occur with lack of cloud cover at high altitudes. The dung of these animals is source of fuel for native residents. *See* ALPACA; LLAMA.

Agricultural practice. Unlike the potato *Solanum tuberosum*, so widely grown in cold climates, certain tuber producers are little known, including oca (*Oxalis tuberosa*), ullucu (*Ullucus tuberosus*), and anu (*Tropaeolum tuberosum*), all typically Andean and able to grow at high altitudes. Preservation of tubers by alternate dehydration and freezing, with subsequent mashing and forming into chuñu, is a puna practice dating from humans' early occupation of these high areas. *See* ALTITUDINAL VEGETATION ZONES.

Harriet G. Barclay

Bibliography. F. D. Ayres, Giants of the puna, *Nat. Hist.*, 75:54–57, March 1966; H. E. Wright, Jr., and W. H. Osburn (eds.), *Arctic and Alpine Environments*, 1968.

Purine

A heterocyclic organic compound (**1**) containing

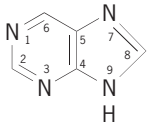

(**1**)

fused pyrimidine and imidazole rings. A number of substituted purine derivatives occur in nature; some, as components of nucleic acids and coenzymes, play vital roles in the genetic and metabolic processes of all living organisms. *See* COENZYME; NUCLEIC ACID.

Chemistry. Purines are generally white solids of amphoteric character. Thus, they can form salts with both acids and bases. The presence of conjugated double bonds in purines results in aromatic chemical properties, confers considerable stability, and accounts for their strong ultraviolet absorption spectra. With the exception of the parent compound, most substituted purines have low solubilities in water and organic solvents. While a variety of methods have been used for the synthesis of purine compounds, the Traube method, shown in the reaction below, may be singled out because of its general usefulness.

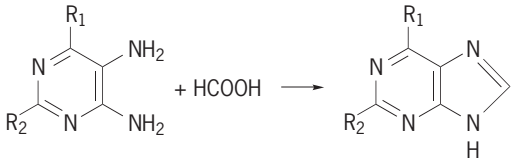

Occurrence. The purine bases, adenine (**2**) and guanine (**3**), together with pyrimidines, are fun-

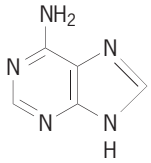

(**2**)

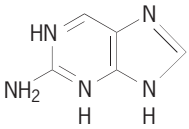

(**3**)

damental components of all nucleic acids. Certain methylated derivatives of adenine and guanine are also present in some nucleic acids in low amounts. In biological systems, hypoxanthine (**4**), adenine, and

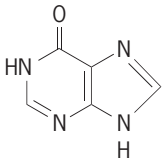

(**4**)

guanine occur mainly as their 9-glycosides, the sugar being either ribose or 2-deoxyribose, although other pentoses are found in rare cases. Such compounds are termed nucleosides generically, and inosine (hypoxanthine nucleoside), adenosine, or guanosine

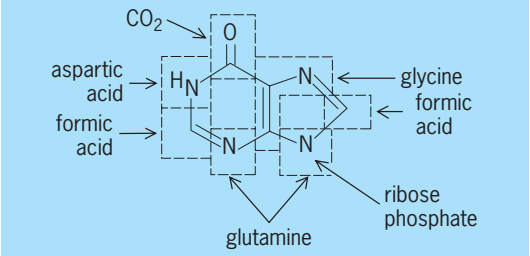

Formation of a purine following enzymatic reactions of its component compounds.

specifically. The name nucleotide refers to phosphorylated derivatives of nucleosides, the principal ones containing 5′-phosphate groups, as in guanosine 5′-phosphate and adenosine 5′-triphosphate (ATP). *See* ADENOSINE TRIPHOSPHATE (ATP).

Nucleic acids are polymers in which nucleotide units are linked by phosphodiester bonds between the 3′-hydroxyl group of one residue and the 5′-hydroxyl group of the next. The methylated xanthines, caffeine (1,3,7-trimethylxanthine), and theobromine (3,7-dimethylxanthine) present in coffee, tea, and cola nuts are responsible for the stimulatory action of these plant products.

Biosynthesis and metabolic functions. Most living organisms are capable of synthesizing purine compounds; in higher animals, this function is largely carried out in the liver and rapidly proliferating tissues such as spleen, bone marrow, and intestinal mucosa. The sequence of enzymatic reactions by which the initial purine product, inosine 5′-phosphate, is formed utilizes glycine, carbon dioxide, formic acid, and amino groups derived from glutamine and aspartic acid (see **illus.**). All of the intermediates in the pathway are derivatives of ribose 5-phosphate. Adenosine 5′-phosphate and guanosine 5′-phosphate are formed from inosine 5′-phosphate by amination reactions requiring aspartic acid or glutamine.

The energy made available from oxidative and nonoxidative metabolism of most foodstuffs in living cells is in part used to condense phosphoric acid and adenosine 5′-phosphate to form ATP. ATP is one of the central compounds in metabolism, acting to transfer, in the form of its energy-rich phosphoanhydride bonds, chemical energy derived from the degradation of foodstuffs to energy-requiring processes of biosynthesis, transport, muscular action, and nerve conduction. As an example, ATP combines with amino acids to form amino acyl adenylates (**5**), the activated intermediates of protein synthesis.

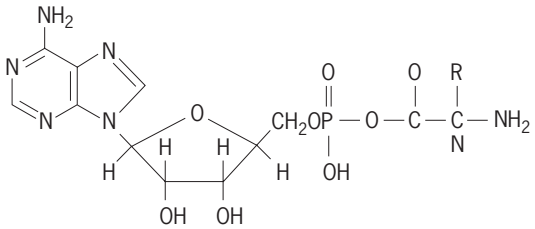

(**5**)

Further, the active forms of several vitamins,

including niacin, pantothenic acid, riboflavin, and cobalamin (B_{12}), are nucleotides containing adenosine phosphate.

Metabolic degradation of purine derivatives may also occur by hydrolysis of nucleotides and nucleosides to the related free bases. Deamination of adenine and guanine produces hypoxanthine and xanthine (**6**), both of which may be oxidized to uric acid (**7**). In most mammals, uric acid is further oxidized to

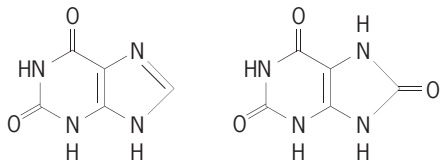

(**6**) (**7**)

the excretory product, allantoin. Primates, including humans, are unable to oxidize uric acid and therefore excrete this substance as the end product of purine metabolism. The arthritic condition known as gout results from precipitation of uric acid or its salts in the joints of individuals with abnormally high levels of these materials in serum and tissues. Birds, certain terrestrial reptiles, and arthropods deposit uric acid or guanine in solid form as the major nitrogenous end product of metabolism. Bird guano, composed largely of uric acid, is harvested in enormous quantities for use as a fertilizer. *See* FERTILIZER; URIC ACID.

Medical uses. A vast number of purine-related compounds have been investigated as potential chemotherapeutic agents. In particular, 6-mercaptopurine (**8**), in the form of its nucleoside phosphate, inhibits several enzymes required for synthesis of adenosine and guanosine nucleotides, and thus proves useful in selectively arresting the growth of tumors. The pyrazolopyrimidine (**9**) has

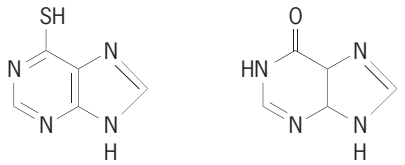

(**8**) (**9**)

been used in gout therapy. As a purine analog, this agent serves to block the biosynthesis of inosine phosphate, as well as the oxidation of hypoxanthine and xanthine to uric acid. As a result of its use, overproduction of uric acid is prevented and the primary cause of gout is removed. *See* CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; GOUT; PYRIMIDINE.

Standish C. Hartman

Bibliography. L. Stryer, *Biochemistry*, 1995; T. W. Trevor and H. A. Simmonds, *Purines: Basic and Clinical Aspects*, 1991.

Push-pull amplifier

An electronic circuit in which two transistors (or vacuum tubes) are used, one as a source of current and one as a sink, to amplify a signal. One device "pushes"

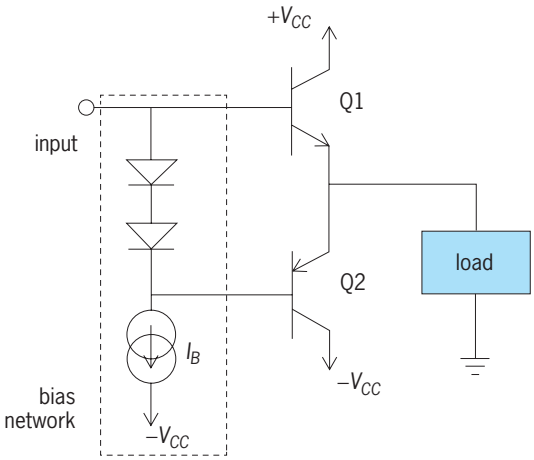

Fig. 1. Complementary-symmetry push-pull output stage. Q1 is an *npn* transistor and Q2 is a *pnp* transistor; I_B is a bias current; positive ($+V_{CC}$) and negative ($-V_{CC}$) power supplies are shown.

current out into the load, while the other "pulls" current from it when necessary. A common example is the complementary-symmetry push-pull output stage widely used to drive loudspeakers (**Fig. 1**), where an *npn* transistor can source (push) current from a positive power supply into the load, or a *pnp* transistor can sink (pull) it into the negative power supply. The circuit functions as an amplifier in that the current levels at the output are larger than those at the input. *See* TRANSISTOR.

Biasing for classes A and B. A so-called bias network in a complementary-symmetry push-pull output stage (Fig. 1) functions to maintain a constant voltage difference between the bases of the two transistors. It can be designed either by setting a bias current, and diode sizes or by replacing it with a different network for class B, class A, or the common compromise, class AB mode of operation. In class B operation, where the bases of the transistors might simply be shorted together, only one transistor is "on" at a time and each is on average "on" for only 50% of the time; when the output current is zero, no current at all flows in the circuit. In class A operation a large voltage is maintained between the bases so that both devices stay "on" at all times, although their currents vary so that the difference flows into the load; and even when the output is zero, a large quiescent current flows from the power supplies. Class B operation is much more efficient than class A, which wastes a large amount of power when the signal is small. However, class B suffers from zero-crossing distortion as the output current passes through zero, because there is generally a delay involved as the input swings far enough to turn one transistor entirely off and then turn the other on. In class AB operation, some intermediate quiescent current is chosen to compromise between power and distortion.

Class AB amplifiers are conventionally used as loudspeaker drivers in audio systems because they are efficient enough to be able to drive the required

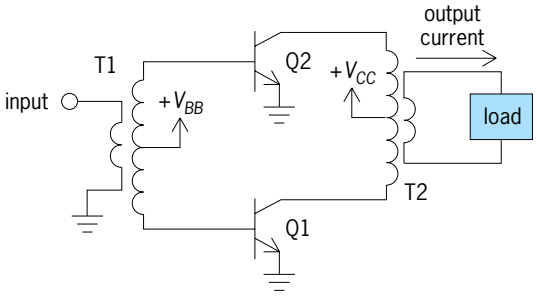

Fig. 2. Transformer-coupled push-pull stage. Both Q1 and Q2 are *npn* transistors, T1 is a phase-splitter transformer, T2 is a center-tapped transformer, and V_{BB} is a bias voltage source. Other symbols are explained for Fig. 1.

maximum output power, often on the order of 100 W, without dissipating excessive heat, but can be biased to have acceptable distortion. Audio signals tend to be near zero most of the time, so good performance near zero output current is critical, and that is where class A amplifiers waste power and class B amplifiers suffer zero-crossing distortion. A class AB push-pull amplifier is also conventionally used as the output stage of a commercial operational amplifier. *See* AUDIO AMPLIFIER; OPERATIONAL AMPLIFIER; POWER AMPLIFIER.

Transformer-coupled amplifier. A second important variant of the push-pull amplifier is the transformer-coupled push-pull stage (**Fig. 2**), where a center-tapped transformer is used to avoid the need for complementary device types (the *npn* and *pnp* transistors in the complementary-symmetry push-pull output stage). Here, both transistors are type *npn*, and when one of them turns on, positive current flows in the load (push), whereas when the other transistor turns on negative, load current flows (pull). Both transistors can only sink current, but their effects on the load are opposite because they are connected to opposite ends of the transformer. This configuration is particularly important when vacuum tubes are used instead of transistors, because there is no complementary tube; tubes operate only by electron flow. A phase-splitter transformer is responsible for turning on one transistor at a time, and can be replaced by an active-circuit implementation if appropriate. A bias voltage is also supplied to the two transistors through the phase-splitter transformer.

Push-pull circuits have an odd symmetry in their operation, meaning that their input-output characteristic satisfies $f(x) = -f(-x)$, because of the symmetry of their structure. This property suppresses even-harmonic distortion, but the circuits may still show substantial distortion at odd harmonics of their input frequency, especially if they are biased in class B where there is zero-crossing distortion. *See* AMPLIFIER; DISTORTION (ELECTRONIC CIRCUITS).

Martin Snelgrove

Bibliography. C. A. Holt, *Electronic Circuits, Digital and Analog*, 1978; J. Millman and A. Grabel, *Microelectronics*, 2d ed., 1987; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Pycnodontiformes

An order of very specialized deep-bodied fishes near the holostean level of organization that are known only from the fossil record. They were a widespread group that first appeared in the Upper Triassic of Europe with the genus *Eomesodon*, flourished during the Jurassic and Cretaceous (*Gyrodus, Coelodus, Gyronchus, Proscinetes, Mesturus*, and others), and persisted to the Upper Eocene (*Pycnodus* and *Palaeobalistum*). Pycnodontiforms are a closely interrelated group most commonly found preserved in marine limestone and associated with coraliferous facies.

They are characterized (see **illus.**) by a laterally compressed, disk-shaped body; long dorsal and anal fins, with each fin ray supported by its own endoskeletal element; an externally symmetrical tail; an axial skeleton having greatly extended neural and hemal spines that are expanded in some forms, and neural and hemal arches which are well ossified—sometimes interdigitating around the notochord—but with no vertebral centra. The ganoid scales, if present, are usually articulated in vertical rows by a modified peg and socket mechanism; often the scalation covers only the anterior half of the body. In the head the subopercle is absent as an individual element and there is no interopercle. The mouth is small but with a powerful mandible. There are a few protruding chisel-like teeth at the front of the jaws. In the upper jaw, the maxillary bone is reduced or absent. Seated on an inner (splenial) expansion of each mandible are two or five rows of blunt, hemispherical teeth that oppose a similar battery of vomerine teeth in the roof of the mouth. No replacement teeth are known. The head is greatly modified in support of this dentition, so much so that the visual clues to relationships with other fishes which usually can be found in skull structure are almost completely obscured in this fossil group.

The origin of the pycnodonts remains one of the puzzling problems in paleoichthyology. These fish exhibit a combination of advanced and archaic characters. For example, the skull roof has a supraoccipital, sometimes with a crest developed, while the notochord is unrestricted. The scales are very like

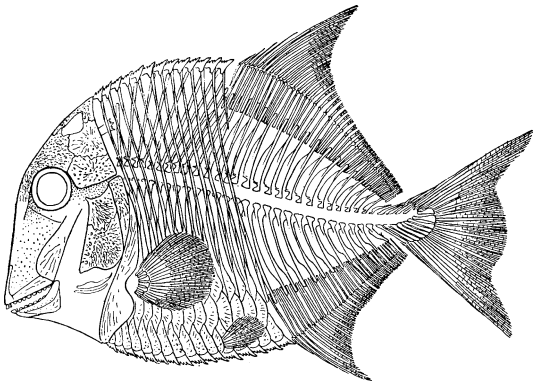

***Coelodus costae*, a pycnodont from Lower Cretaceous of Italy; length to 4 in. (10 cm).**

those of the chondrostean Platysomoidei in gross features but the fins (particularly the caudal) are holostean in nature. The maxilla, when present, is free from the cheek but the preopercle still covers much of the cheek area. The mandible has a high coronoid process, an advanced character; yet the parasphenoid is short and expanded posteriorly as in palaeoniscoids. An attractive hypothesis, of some weight, is the derivation of the pycnodontiforms from a deep-bodied chondrostean stock possibly within the Platysomoidei.

The teeth of pycnodontiforms are adapted for crushing shelled invertebrates, and shell fragments have been found inside the abdominal cavity of some specimens. The body type is an adaptation for life in a confined, quiet water habitat where great maneuverability is necessary, such as the interstices of a coral reef or in an area with much vegetation. A close comparison can be drawn between the modern, foraging type of reef fishes belonging to the teleost level and the pycnodonts. It is probable that the teleost forms as they evolved and spread came into direct contact with some of the pycnodontiforms and competed with them. In like manner, the appearance of the latter may have filled the habitat vacated by the earlier Platysomoidei. *See* PALAEONISCIFORMES.

Ted. M. Cavender

Bibliography. W. N. McFarland et al., *Vertebrate Life*, 1979; A. S. Romer, *Vertebrate Paleontology*, 3d ed., 1966.

Pycnogonida

A class of marine arthropods, consisting of about 600 Recent and 3 Devonian species. The Pycnogonida, or Pantopoda, are commonly called sea spiders.

Morphology. The pycnogonids are characterized by reduction of the body to a series of cylindrical trunk somites supporting the appendages, a large specialized feeding apparatus called the proboscis, gonopores opening on the second joints of the legs, and a reduced abdomen. In many genera, such as *Nymphon* (**Fig. 1**), there are seven pairs of appendages, of which the first four, namely the chelifores, palpi, ovigers, and first walking legs, are on the first or cephalic segment. This segment also bears a dorsal tubercle containing four simple eyes. Each of the remaining three trunk segments bears a single pair of legs. In some genera the chelifores, palpi, or both are absent in the adult, and the ovigers are lacking in the females. Ovigers are always present in males and hold the ball-like egg masses. The extreme of reduction, in which the adults lack anterior appendages, except for ovigers in the male, is represented by *Pycnogonum* (**Fig. 2**), the common intertidal genus from which the group is named. In some species there is an additional trunk somite and pair of legs, and one species has 2 extra somites and 12 walking legs. These polymerous forms occur in paucigeneric families with a large number of species to a genus, and are not known in families with a large number of paucispecific genera.

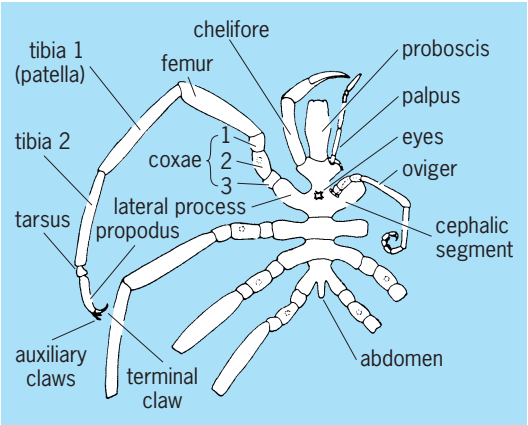

Fig. 1. ***Nymphon*, external features in dorsal view.**

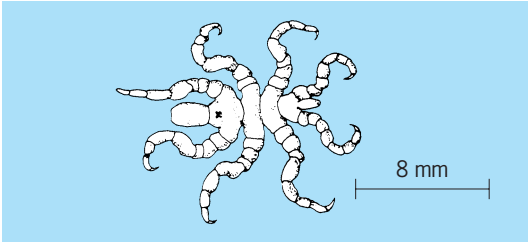

Fig. 2. ***Pycnogonum stearnsi*, a common littoral pycnogonid of the Pacific Coast.**

Reduction of the body is accompanied by increased length of the legs, so that a typical pycnogonid appears to be a bundle of long legs. These contain diverticula of the digestive tract and gonads, and the sexual products ripen in the femora. The proboscis is a rigid tripartite structure, often longer than the combined trunk segments, with a three-cornered mouth at the apex, and containing a straining apparatus which enables the animal to feed on the juices of soft-bodied animals. The nervous system follows the usual arthropod pattern of supraesophageal ganglion, or brain, circumesophageal ring, paired ventral ganglia, and ventral nerve cords. The open circulatory system consists of a dorsal tubular heart with two or three pairs of lateral ostia. Respiratory and excretory systems are lacking. The femora of the males contain specialized cement glands which produce an adhesive substance for forming the egg ball.

Affinities. The chelicerate condition of the first pair of appendages, the patellar joint of the legs, and suppressed deuterocentrum are characters suggesting affinity with the Chelicerata, but the multiple genital openings on the legs and the occurrence of polymerous forms are unique among arthropods. The terminal joints of the oviger have intratarsal muscles, another unique feature, while the pretarsus, or propodus, of the legs has both retractor and extensor muscles, in common with both Chelicerata and Crustacea. The protonymphon larva (**Fig. 3**) of the pycnogonida invites comparison with the crustacean nauplius larva, but the three pairs of legs are not biramous, and the resemblances are superficial. The

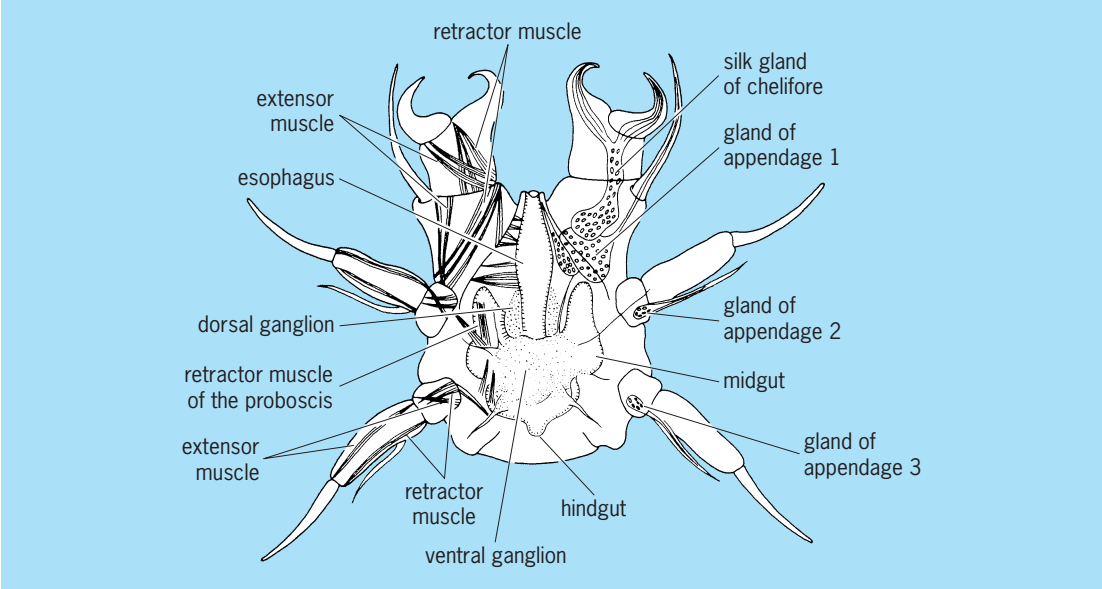

Fig. 3. Protonymphon larva, ventral view.

protonymphon has a number of specialized glands, producing materials for attaching to, or invading, the host, and the structure of the proboscis is well advanced. The common name, sea spider, refers to the appearance of the adults and does not indicate affinity with terrestrial spiders. *See* CHELICERATA.

Mode of life. Pycnogonids are found in all seas except the inner Baltic and Caspian, from intertidal regions to depths of 21,000 ft (6500 m), and one species is bathypelagic at about 3300 ft (1000 m). They are especially common in polar seas. Most of the intertidal species spend their lives in association with coelenterates as encysted, parasitic larval and juvenile stages, or are ectoparasitic as adults, being attached to anemones and hydroids by their claws and proboscides. A few have been found riding in hydromedusae, and some occur in the mantle cavity of bivalves or on nudibranchs and holothurians. Most of the deep-sea species are known only as adults, and their mode of life is a mystery.

Classification. The Pycnogonida are classified primarily on the presence or absence of various anterior appendages; about 60 genera are recognized, grouped in 8 families, which are listed in the **table**. No ordinal distinctions can be recognized for Recent families, and many genera are artificially defined. In order to include Devonian *Palaeopantopus* in the Pycnogonida, the Recent forms are assigned to Pantopoda, and the fossil to Palaeopantopoda. *See* ARTHROPODA; PALAEOISOPUS. Joel W. Hedgpeth

Bibliography. C. P. Arango, Morphological phylogenetics of the sea spiders (Arthropoda: Pycnogonida), *Organisms Diversity Evolution*, 2:107–125, 2002; F. Arnaud and R. N. Bamber, The biology of Pycnogonida, *Adv. Mar. Biol.*, 24:1–96, 1987; J. A. Dunlop and C. P. Arango, Pycnogonid affinities: A review, *J. Zool.*

Recent families of Pycnogonida

Family	Chelifores	Palpi	Ovigers	Genera
Nymphonidae	Present	5-jointed	10-jointed in both sexes	Includes largest genus, *Nymphon* (over 150 species) and polymerous *Pentanymphon* and *Sexanymphon*
Callipallenidae	Present	Lacking or reduced	10-jointed in both sexes	
Phoxichilidiidae	Present	Lacking	5–9-jointed in male only	
Endeidae	Absent	Absent	7-jointed in male only	Monogeneric, paucispecific, *Endeis*
Ammotheidae	Present, usually small and achelate	Well developed, 4–10-jointed	9–10-jointed in both sexes	Many genera
Austrodecidae	Absent	Present, 5–6-jointed	In both sexes reduced, 4–7-jointed	Monogeneric, *Austrodecus*
Colossendeidae	Lacking except in polymerous forms	Long, 9–10-jointed	10-jointed in both sexes	Mostly deep-water forms with large proboscides; genera are *Colossendeis*, polymerous *Pentacolossendeis*, *Decolopoda*, and *Dodecolopoda*
Pycnogonidae	Lacking	Lacking	6–9-jointed in male only	Shallow water to shore forms; includes *Pycnogonum* and polymerous *Pentapycnon*

Syst. Evol. Res., 43(1):8–21, 2005; J. W. Hedgpeth, On the phylogeny of the Pycnogonida, *Acta Zool.* (*Stockholm*), 35:193–213, 1954; P. E. King, *Pycnogonids*, 1974; S. M. Manton, *The Arthropoda: Habits, Functional Morphology, and Evolution*, 1978; D. A. Staples, Sea spiders or pycnogonids (phylum Arthropoda), pp. 1040–1072 in *Marine Invertebrates of Southern Australia, Part IIIM*, S. A. Shepherd and M. Davies, eds., South Australian Research and Development Institute, Adelaide, 1997.

Pygasteroida

An order of Eognasthostomata which exhibits various stages in the backward migration of the anus out of the apical system. They have four genital pores, noncrenulate tubercles, and simple ambulacral plates. All members are referred to a single family, the Pygasteridae. They apparently arose from Triassic Pedinidae and occur in the Jurassic and Cretaceous of the Northern Hemisphere. They were formerly classified with other bilaterally symmetrical echinoids in the artificial assemblage Irregularia. *See* ECHINODERMATA; IRREGULARIA. Howard B. Fell

Pyrenulales

An order of the class Ascolichenes, also known as the Pyrenolichenes. As now circumscribed, the Pyrenulales includes only those lichens with perithecia that contain true paraphyses and unitunicate asci. Other pyrenolichens with pseudoparaphyses and bitunicate asci have been transferred to the Pseudosphaeriales. The flask-shaped perithecia are uniformly immersed in the medulla of the thalli with a small ostiole opening at the surface. The asci and paraphyses arise from a blackened hypothecium and line the walls of the perithecium. The spores eventually burst the ascal walls and ooze out through the ostiole in a jelly matrix. Details of ascal development have been described rather fully for *Dermatocarpon aquaticum*. The Pyrenulales are almost all crustose in growth form, with very simple internal structure. The exception is Dermatocarpaceae, large umbilicate species often confused with the typical rock tripe, *Umbilicaria*, in the Lecanorales.

There are about 10 families, 50 genera, and more than 1500 species in the Pyrenulales. The major taxonomic criteria for separating genera and species are the septation and color spores, since vegetative characters are so poorly developed. The larger families are as follows.

The Dermatocarpaceae contain four genera with umbilicate or squamulose growth form. Most of the species grow on limestone or calcareous soils.

In the Pyrenulaceae are about 10 genera, all crustose species and most common on tree bark in the tropics. They are closely related to the nonlichenized Sphaeriales, but the perithecia are never produced in true stoma.

The Strigulaceae comprise about six genera of crustose species confined chiefly to leaves of evergreen trees in the tropics. These peculiar lichens form extensive crusts on or under the cuticle of leaves without seeming to damage the host plant.

The Verrucariaceae comprise about 8 genera of crustose species typically found on rocks, especially in intertidal or salt-spray zones along rocky coastlines. These are the only truly amphibious lichens. Mason E. Hale

Pyridine

An organic heterocyclic compound containing a triunsaturated six-membered ring of five carbon atoms and one nitrogen atom. Pyridine (**1**) and

(**1**)

pyridine homologs are obtained by extraction of coal tar or by synthesis. The following are available in commercial quantities: pyridine, 2-, 3-, and 4-methylpyridine (also known respectively as α-, β- and γ-picoline), 2,4-dimethyl-, 2,6-dimethyl-, and 3,5-dimethylpyridine (also known respectively as 2,4-, 2,6-, and 3,5-lutidine), 5-ethyl-2-methylpyridine (also called aldehydecollidine), and 2,4,6-trimethypyridine (also called 2,4,6-collidine). Other pyridine derivatives produced on a large scale include nicotinic acid (pyridine-3-carboxylic acid) for preparation of nicotinamide, nicotine for its insecticidal properties, 2-aminopyridine for synthesis of medicinals, piperidine (hexahydropyridine) as a solvent, and 2-vinylpyridine as a polymerizable monomer. The pyridine system is found in natural products, for example, in nicotine (**2**) from tobacco, in ricinine (**3**) from castor bean, in pyridoxine or vitamin B_6 (**4**), in nicotinamide or niacinamide or vitamin P (**5**),

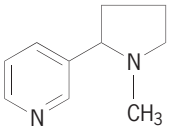

(**2**)

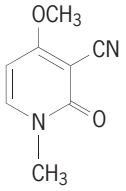

(**3**)

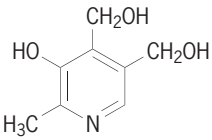

(**4**)

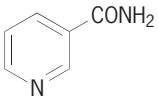

(**5**)

and in several groups of alkaloids. *See* HETEROCYCLIC COMPOUNDS.

Properties. Pyridine (**1**) is a colorless, hygroscopic liquid with a pungent, unpleasant odor. When anhydrous it boils at 115.2–115.3°C (239.4–239.5°F), its

density (20/4) is 0. 98272, and n_D^{20} is 1.50920. Pyridine is miscible with organic solvents as well as with water. A constant-boiling mixture, by 92°C (198°F) forms with three molecules of water. Dry pyridine is obtained by treatment with barium oxide followed by calcium hydride or phosphorus pentoxide. Pyridine is a tertiary amine (pK_a 5.17 at 25°C or 78°F) that combines readily with Brönsted and Lewis acids. The pyridine system is aromatic. It is stable to heat, to acid, and to alkali. It undergoes electrophilic substitution with difficulty, with the 3 position favored in sulfonation and nitration. On the other hand, it undergoes nucleophilic substitution with ease at the 2 position, and lightly less readily at the 4 position. Its resonance energy is 35 kcal/mole. Pyridine is used as a solvent for organic and inorganic compounds, as an acid binder, as a basic catalyst, and as a reaction intermediate.

Pyridine is an irritant to skin (eczema) and other tissues (conjunctivitis), and chronic exposure has been known to cause liver and kidney damage. Repeated exposure to atmospheric levels greater than 5 parts per million is considered hazardous.

Oxidation of pyridine homologs by nitric acid or by permanganate converts the substituent group in a preparative manner to a carboxylic acid. Reaction at the methyl group of 2- and 4-methypyridine tends to occur more readily than at the methyl group of 3-methylpyridine. Thus, 2- and 4-methylpyridine condense with benzaldehyde to give styryl derivatives [**6**; reaction (1)], whereas 3-methylpyridine

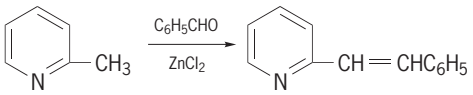

(1)

(**6**)

does not react. With butyllithium or sodium amide, 2- and 4-methypyridines metalate to give pyridylmethyl metals (**7**), which react normally with carbonyl compounds [RCHO; reaction (2*a*)] and with alkyl halides [RX; reaction (2*b*)].

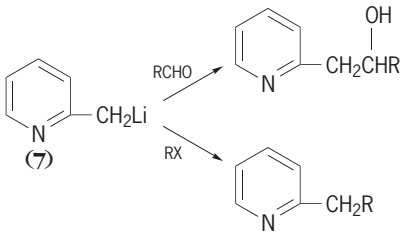

(2*a*)

(2*b*)

(**7**)

Many halogenated derivatives of pyridine are known. Direct bromination at 300–400°C (570–750°F) places bromine at the 3 position and 5 position. 2-Bromopyridine can be prepared from 2-pyridone and phosphorus oxybromide. 3-Bromopyridine is prepared by application of the Sandmeyer process to 3-aminopyridine or by mercurating and then brominating pyridine. 2-Chloropyridine is formed when *N*-methyl-2-pyridone is treated with phosphorus pentachloride. 4-Chloropyridine (**10**) can be prepared by nitrating pyridine-*N*-oxide (**8**), exposing the 4-nitropyridine-*N*-oxide (**9**) to the action of concentrated hydrochloric acid, and removing the oxide oxygen by iron-acetic acid reduction [reaction (3)], but is usually

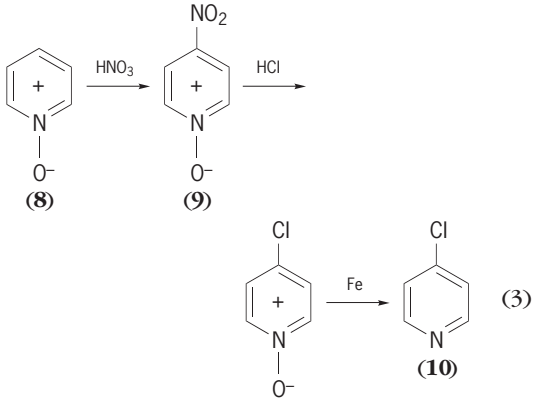

(3)

made by heating *N*-(4-pyridyl) pyridinium chloride (from pyridine and thionyl chloride). The 2- and 4-halo substituents are more readily replaced by hydroxy, alkoxy, and amino than the 3-halo substituent. Bromopyridines form Grignard and lithium derivatives that react normally.

3-Nitropyridine, from nitration of pyridine, can be converted to 3-aminopyridine. The latter is usually prepared by Hofmann rearrangement of nicotinamide (**5**). 4-Nitropyridine-*N*-oxide (**9**) with iron and acetic acid gives 4-aminopyridine. 2-Aminopyridine, prepared on an industrial scale by direct amination of pyridine with sodamide, is utilized in the manufacture of the antihistaminic, pyribenzamine (**11**), and the bacteriostatic agent, sulfapyridine (**12**).

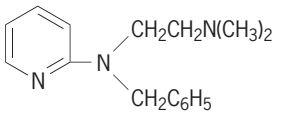

(**11**)

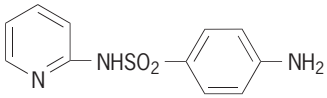

(**12**)

In most of its reactions, 3-hydroxypyridine behaves as a normal phenol, just as 3-aminopyridine behaves as a normal aromatic amine. However, hydroxy or amino groups at the pyridine 2 or 4 position show some reactions that are not characteristic of phenols or aromatic amines. Hydroxyl or amino groups on any pyridine position make electrophilic substitution easier and, as ortho-para directing groups, take control of the orientation. 2-Pyridone is prepared from 2-aminopyridine by a diazotization procedure. 3-Hydroxypyridine is produced by sulfonation of pyridine followed by alkali fusion of the pyridine-3-sulfonic acid.

Preparation. Laboratory synthetic methods can lead to pyridines with no oxygen at positions 2

or 6, or to 2-hydroxypyridines (also known as 2-pyridone), or to 2,6-dihydroxypyridines (also known as 2,6-pyridinone). These two pyridine derivatives exist almost entirely in their tautomeric forms. For example, 2-hydroxypyridine **13***a* can tautomerize to **13***b* as shown in reaction (4) . Glutaconic acid cy-

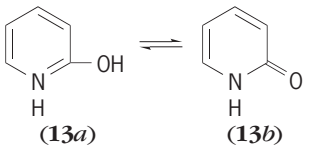

(4)

clizes with ammonia to give 6-hydroxy-2-pyridones. The oxygen at the 2 position and 6 position can be removed by standard conversions with phosphorus oxychloride to the 2,6-dichloro derivative, followed by reductive dechlorination. In this way, for example, β-methyl-glutaconic acid (**14**) can be converted to 4-methylpyridine [**15**; reaction (5)]. 2-Pyridones (**17**)

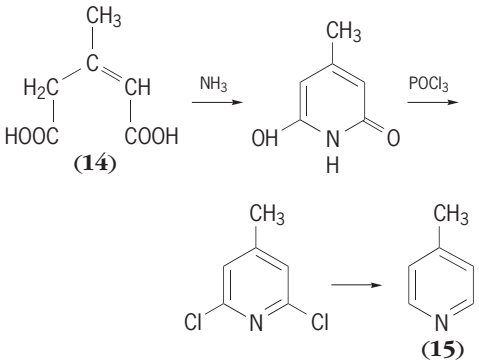

(5)

are formed when 1,3-dicarbonyl compounds (**16**), or their equivalents, react with cyanoacetamide. Subsequent steps remove the cyano groups as well as the oxygen from (**17**) to furnish substituted pyridines [**18**; reaction (6)].

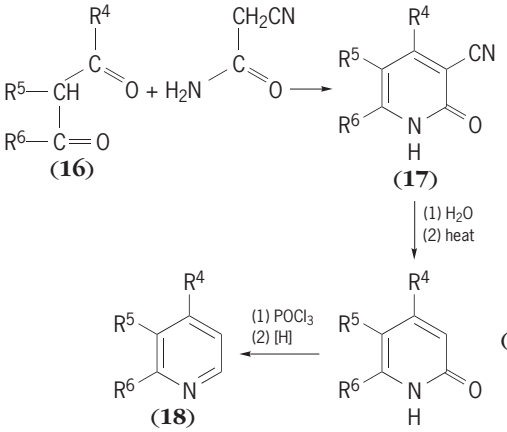

(6)

1,5-Dicarbonyl compounds or their equivalents cyclize with ammonia to give pyridines. Thus, the diketone (**19**) from acetoacetic ester and ethyl orthoformate gives the pyridine (**20**); in a closely related process, the reaction of ethyl β-aminocrotonate (**21**) with ethoxymethyleneacetoacetic ester (**22**) gives the same product [reactions (7)]. The Hantzsch synthesis combines four molecules to form an intermediate dihydropyridine derivative (**23**), which can be readily oxidized to the corresponding completely aromatic derivative [**24**; reaction (8)]. The Chichibabin aldehyde-ammonia synthesis involves the condensation of ammonia with aldehydes and ketones. Generally, mixtures of pyridines are obtained, with the course of the reaction depending on such factors as nature and proportion of reactants, reaction time, temperature, and catalyst. In a commercial process, 5-ethyl-2-methylpyridine (**25**) is obtained in unusually high yields (60–70%) from acetaldehyde

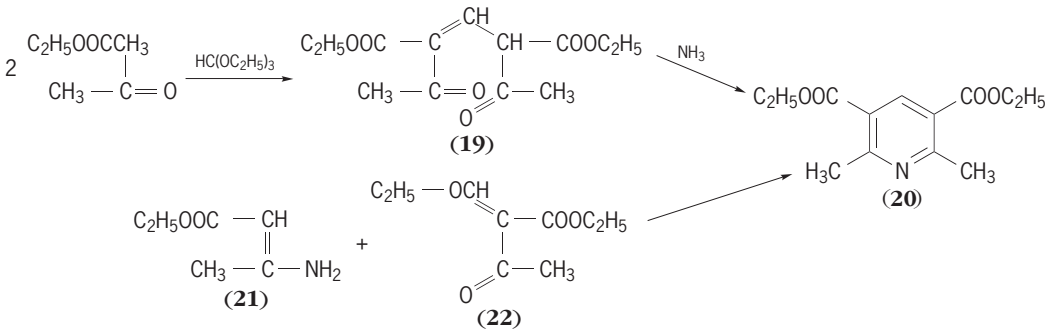

(7)

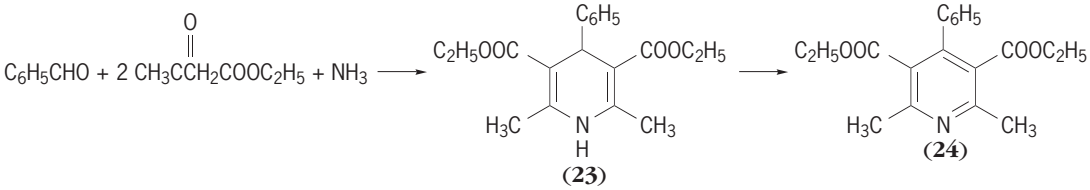

(8)

and ammonia [reaction (9)]. Synthetic industrial pyridine and its homologs are prepared from aldehydes and ammonia over a suitable catalyst.

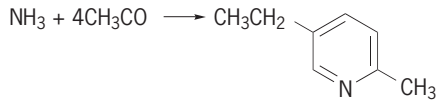

(9)

(25)

Derivatives. Pyridine compounds containing positively charged nitrogen include simple and quaternary pyridinium salts, acylpyridinium salts, and pyridine-*N*-oxides. The reaction of pyridine with alkylating agents xs quaternary salts, crystalline solids whose aqueous solutions conduct electricity. The quaternary salts are readily oxidized by alkaline ferricyanide to the *N*-substituted-2-pyridones. With acyl halides, pyridine forms *N*-oxide (**26**), prepared by oxidation of pyridine with organic peracids, which, in turn, reacts with phosphorus pentachloride to give 2- and 4-chloropyridine [reaction (10)]

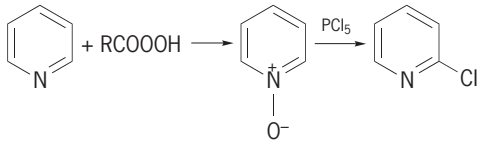

(10)

(26)

and with acetic anhydride to give 2-acetoxypyridine. Pyridine-*N*-oxides nitrate at the 4 position, and by so doing provide a route to 4-substituted derivates.

Pyridine aldehydes are prepared by oxidation of groups already on the ring. Acetylpyridines can be synthesized from pyridine carboxylic esters and ethyl acetate by the Claisen condensation. The reactions of pyridine aldehydes and ketones are normal.

Many pyridine carboxylic acids are known. Their reactions as dipolar materials are not exceptional. Thermal decarboxylation is a standard process, with loss of carboxyl from position 2 easier than from 3 or 4, and loss of carboxyl from position 4 easier than from 3. Nicotinic acid (**27**) is manufactured by oxidation of 3-methylpyridine (β-picoline) or nicotine; it is also obtained by oxidation of either 2-methyl-5-ethylpyridine or quinoline, followed by decarboxylation of the resulting pyridine dicarboxylic acids [reactions (11)]. Pyridine-2-carboxylic acid (α-picolinic acid) can be prepared by oxidation of 2-methylpyridine or by carbonation of 2-pyridyllithium. 4-Pyridine-carboxylic acid (isonicotinic acid) is obtained by oxidation of 4-methylpyridine or by synthesis from citric acid. See (**28**) to (**29**), as in reaction (12). The acid hydrazide of isonicotinic acid (Isoniazid) is a tuberculostatic agent.

Although dihydro- and tetrahydropyridines are known, the hexahydropyridines, or piperidines, are the most common reduced forms of pyridine. The piperidines may be prepared by reduction of pyridines or by cyclization of bifunctional com-

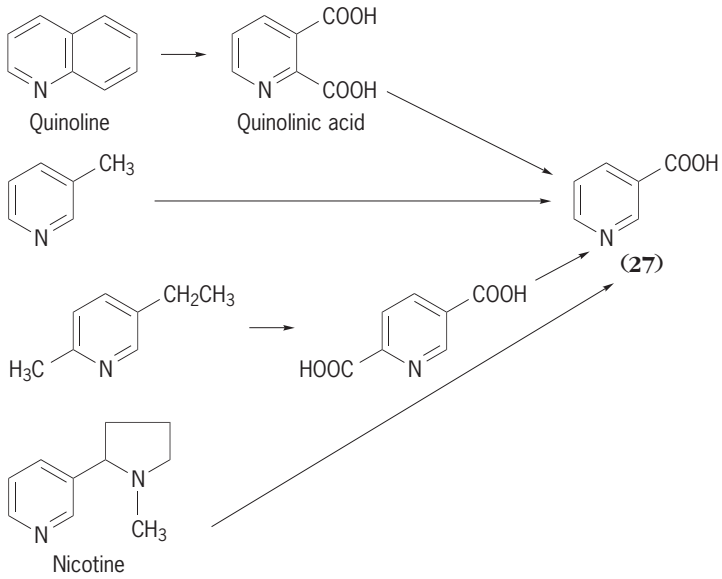

(11)

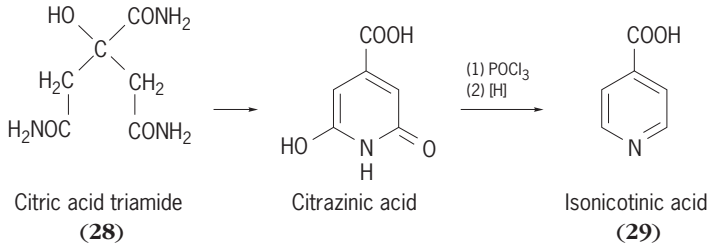

(12)

(28) (29)

pounds, for example, the conversion of 5-bromo-1-aminopentane (**30**) to piperidine [**30**; reaction (13)].

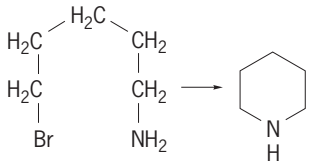

(13)

(30) (31)

Piperidine, the parent compound, is a colorless, unpleasant-smelling liquid (bp 105.6°C or 222.1°F),

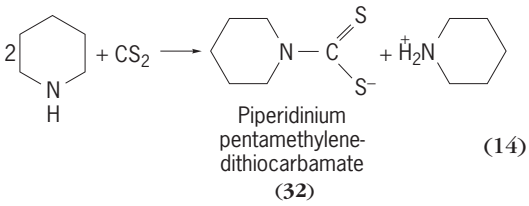

Piperidinium pentamethylene-dithiocarbamate

(14)

(32)

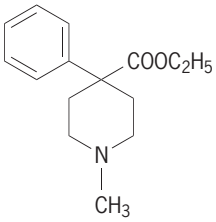

(33)

completely miscible with water. The general properties of piperidine are those of a normal secondary aliphatic amine and, as such, piperidine (pK_a 11.1) is a much stronger base than pyridine (pK_a 5.17). Piperidine carboxylic acids have been investigated in connection with naturally occurring amino acids (for example, pipecolinic acid is piperidine-2-carboxylic acid) as well as with the degradative and synthetic chemistry of quinine. The reaction product (**32**) from piperidine and carbon bisulfide is a rubber accelerator [reaction (14)]. 4,4-Di-substituted piperidines such as Demerol (**33**) are analgesics.

Walter J. Gensler

Bibliography. R. A. Abramovitch (ed.), *Pyridine and Its Derivatives*, suppl., vols. 1–4, 1974–1975; D. H. R. Barton et al. (eds.), *Comprehensive Organic Chemistry*, 1978; G. R. Newkome (ed.), *Pyridine and Its Derivatives*, vol. 14, pt. 5, 1984; P. Tomasik and Z. Ratajewicz, *Pyridine-Metal Compounds*, vol. 14, pt. 6, 1985.

Pyrimidine

A heterocyclic organic compound (**1**) containing ni-

(**1**)

trogen atoms at positions 1 and 3. Naturally occurring derivatives of the parent compound are of considerable biological importance as components of nucleic acids and coenzymes and, in addition, synthetic members of this group have found use as pharmaceuticals. *See* COENZYME.

Chemistry. Pyrimidine (1,3-diazine) and its derivatives are weakly basic substances with characteristic ultraviolet absorption spectra. The diazine structure, like that of benzene, may be regarded as a resonance-stabilized system. As a result it is unreactive to all but very vigorous reagents. A classical mode of synthesis of pyrimidine compounds involves condensation of an ethyl malonate with guanidine, urea, or amidine, as in reaction (1). The substituent groups can be

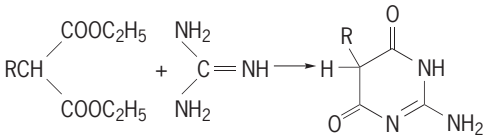

(1)

varied either by choice of different starting reagents or by replacement after formation of the ring system. Certain substituted pyrimidines can exist in tautomeric forms; an example is given by reaction (2).

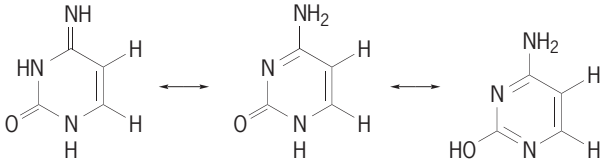

(2)

The forms of the substituent groups shown in the middle structure are the most stable.

Occurrence. Pyrimidine compounds which are found universally in living organisms include uracil (**2**), cytosine (**3**), and thymine (**4**). Together with

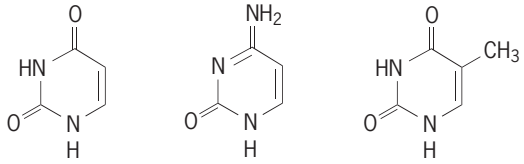

(**2**) (**3**) (**4**)

purines these substances make up the "bases" of nucleic acids, uracil and cytosine being found characteristically in ribonucleic acids, with thymine replacing uracil in deoxyribonucleic acids. A number of related pyrimidines, including 5-methyl-cytosine and 5-hydroxymethyl cytosine, also occur in lesser amounts in certain nucleic acids. Other pyrimidines of general natural occurrence are orotic acid (**5**) and thiamine (vitamin B_1, **6**).

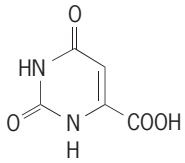

(**5**)

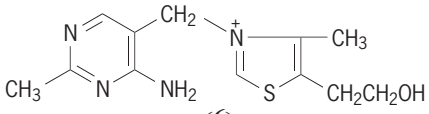

(**6**)

See DEOXYRIBONUCLEIC ACID (DNA); RIBONUCLEIC ACID (RNA).

The pyrimidine bases are normally found in combination with one of two sugars, ribose or 2-deoxyribose, in the form of nucleoside derivatives, for example, uridine (**7**), cytikine, and thymidine (**8**).

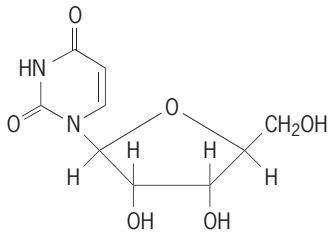

(**7**)

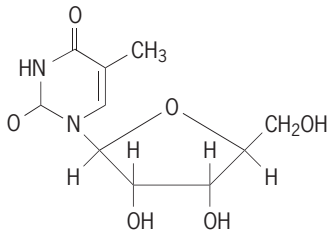

(**8**)

Further, the 5′-hydroxyl group of the sugar residue may exist in ester linkage with phosphoric acid,

forming a nucleoside phosphate, such as uridine 5′-phosphate (uridylic acid, **9**).

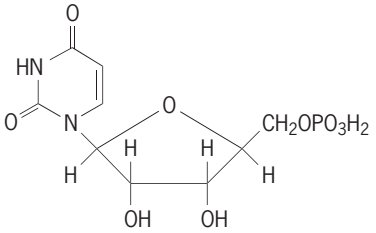

(9)

The class of compounds composed of bases, sugars, and phosphate is referred to as nucleotides, a term including nucleoside di- and triphosphates and a variety of metabolically important coenzymes. Nucleic acids are polymers of nucleotide units which are bridged by phosphodiester bonds between the 3′-hydroxyl group of one unit and the 5′-hydroxyl group of the next. The specificity of nucleic acids in genetic processes is attributed to the formation of complementary pyrimidine-purine pairs (uracil or thymine with adenine, and cytosine with guanine) which are stabilized by intramolecular hydrogen bonds between the ring nitrogen atoms and a substituent amino and carbonyl groups.

Biosynthesis and metabolic function. The synthesis of pyrimidines takes place in virtually all living organisms, and proceeds by way of the intermediate formation of orotic acid. This material arises from ammonia, carbon dioxide, and aspartic acid in a sequence of four enzyme-catalyzed reactions (3). Uridine 5′-

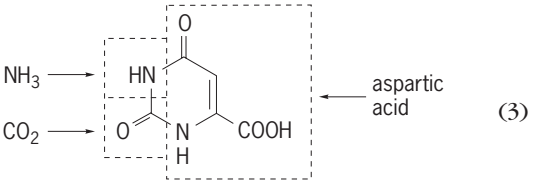

(3)

phosphate is formed after addition of the ribose 5-phosphate moiety to orotic acid, followed by loss of the carboxyl group. Cytidine nucleotides (such as the 5′-triphosphate) result from amination of uridine 5′-phosphate. The biological phosphorylating agent, adenosine triphosphate, acts to convert the nucleoside 5′-phosphates to the corresponding di- and triphosphates. In addition to their direct utilization in nucleic acid synthesis, pyrimidine nucleotides have key metabolic functions in the transfer of activated subunits in the synthesis of many cellular components. As examples, glycogen is built up by incorporation of glucosyl units derived from uridine diphosphoglucose. *See* PURINE.

Cytidine-containing nucleotides play comparable roles in the synthesis of complex lipids found in membranes and nervous tissue. Presumably in such reactions pyrimidine bases are not of primary chemical importance but serve as sites of recognition and specificity in the enzymic processes.

Synthetic pyrimidines of interest. Among the sulfa drugs, the pyrimidine derivatives, sulfadiazine, sulfamerazine, and sulfamethazine, have general formula (**10**). They inhibit folic acid biosynthesis in microorganisms. The barbiturates (**11**) possess potent

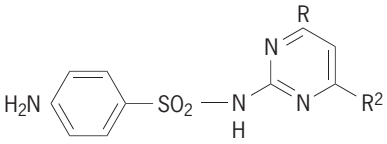

(10)

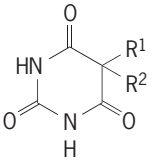

(11)

depressant action on the central nervous system.

This group includes Veronal (R_1 and R_2 = ethyl), phenobarbital (R_1 = ethyl, R_2 = phenyl), Amytal (R_1 = ethyl, R_2 = isoamyl), Seconal (R_1 = allyl, R_2 = 2-pentyl), and pentobarbital (Nembutal, R_1 = ethyl, R_2 = 2-pentyl). The effect of many pyrimidine-related compounds upon nucleotide and nucleic acid metabolism has been examined with the hope of finding useful chemotherapeutic agents. The medical value of such agents has been disappointing even though certain of these have definite biological action. Thus, derivatives of 5-fluorouracil block the methylation step in thymidine phosphate synthesis, and 5-bromouracil and 5-iodouracil are incorporated into deoxyribonucleic acid in place of thymine to produce a polymer of increased susceptibility to mutation. *See* BARBITURATES; CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; MUTATION; NUCLEIC ACID; SULFONAMIDE.

Standish C. Hartman

Bibliography. G. M. Blackburn and M. J. Gait, *Nucleic Acids in Chemistry and Biology*, 2d ed., 1996; M. Florkin and E. H. Stotz (eds.), *Comprehensive Biochemistry,* vol. 19: *Metabolism of Amino Acids, Proteins, Purines, and Pyrimidines,* 1981; D. B. Marks, *Biochemistry*, 3d ed., 1998; K. Mikanagi, *Purine and Pyrimidine Metabolism*, 1989; E. H. Rodd (ed.), *Rodd's Chemistry of Carbon Compounds*, vol. 2, pts. A and B, 2d rev. ed., 1968; E. Smith et al., *Principles of Biochemistry,* 7th ed., 1983.

Pyroclastic rocks

Rocks formed by the accumulation of fragmental volcanic materials (also called pyroclasts) ejected during explosive eruptions. Pyroclasts can vary widely in origin, fragment size and shape, composition, texture, structure (such as their arrangement in the resulting deposit), and other properties (see **table**). Despite such differences, all pyroclastic materials and the great variety of pyroclastic rocks they form share a common, unifying attribute—all are produced by explosive volcanic activity. Pyroclastic deposits are included under a broader category of volcaniclastic deposits, comprising all fragmented volcano-derived material. In some cases, the explosive fragmentation is caused by the interaction of

Classification of pyroclasts and the fall or surface-flow deposits formed by them

Average size of pyroclast	Name and shape of pyroclast	Physical state when erupted or deposited	Deposit resulting from pyroclast accumulation: If mostly unconsolidated	Deposit resulting from pyroclast accumulation: If mostly consolidated
	Block (angular)	Solid	Agglomerate, or deposit of blocks or bombs, or pyroclastic breccia	Agglomerate, or pyroclastic breccia
	Bomb (rounded or aerodynamically shaped)	Semisolid (still fluid)		
	Cinder, scoria (round to angular)	Semisolid or liquid		
	Pumice, spatter (round to angular)	Largely liquid		
64 mm*				
	Lapillus† (round to angular)	Liquid or solid	Lapilli deposit	Lapillistone
2 mm				
	Coarse ash grain (generally angular)	Liquid or solid	Coarse ash deposit	Coarse ash tuff
1/16 mm				
	Fine ash grain, or volcanic dust (generally angular)	Liquid or solid	Fine ash deposit, or volcanic dust deposit	Fine ash tuff, or dust tuff

* Pyroclasts >256 mm are sometimes termed coarse, and those between 64 and 256 mm are fine.
† Lapilli are the smaller-sized equivalents of the types of pyroclasts described in the category >64 mm.

water with hot volcanic ejecta, producing hydrovolcanic materials, which are commonly regarded as pyroclastic *sensu lato* and will be discussed as such. *See* CALDERA; VOLCANO.

Pyroclasts are classified into three principal types according to their origin, which can combine in any proportion in any given pyroclastic rock: juvenile, cognate (or accessory), and accidental. Juvenile pyroclasts are produced by the violent shredding and disintegration of the erupting magma (in liquid or plastic state). Typically, juvenile pyroclasts have a glassy texture and contain small cavities (vesicles) that formerly were bubbles of volcanic gas in the magma. The cognate varieties are formed from the fragmentation of solidified volcanic rocks from previous eruptions from the same volcanic system. Accidental pyroclasts are fragments of geologically older rocks unrelated to the volcano and can be of any composition or rock type (igneous, sedimentary, or metamorphic). Cognate and accidental pyroclasts represent fragments of preexisting solid rock that were incorporated into the rising magma before the explosive eruption. Apart from their origin, pyroclasts are subdivided according to their size and shape, texture (glassy or crystalline), mineral or chemical composition, and other characteristics (see table). Of these distinguishing criteria, fragment size (average diameter) is perhaps the most important, resulting in three main subcategories of pyroclasts: bombs or blocks (>64 mm), lapilli (64–2 mm), and ash (<2 mm). *See* VOLCANIC GLASS.

From studies of pyroclastic rocks formed in the geologic past, combined with direct observations and investigations of pyroclastic deposits formed during historical explosive eruptions, volcanologists have recognized two principal modes of deposition of pyroclastic rocks: fall processes and flowage processes. However, these two processes represent end members of a continuum, and some deposits of pyroclastic rock exhibit features that indicate both kinds of processes were operative to some degree, and sometimes in a complex manner, in their formation. *See* IGNEOUS ROCKS; MAGMA; METAMORPHIC ROCKS; PETROLOGY; SEDIMENTARY ROCKS.

Fall processes and deposits. These processes involve the explosive ejection of pyroclasts high into the air and the fall of these airborne materials, because of gravity, back to the ground, forming pyroclastic fall deposits (sometimes also called air-fall deposits). Tephra refers to any airborne pyroclasts regardless of size, shape, and other physical properties, but it tends mostly to be used for lapilli- and ash-size fragments. Nearly all fall deposits are produced by one or a combination of two processes: (1) fallout of tephra from an eruption column (or eruption plume) jetting upward from the eruptive vent; and (2) fallout of material carried aloft in ash clouds convecting above ground-hugging pyroclastic flows.

Compared with highly destructive pyroclastic flows, the potential hazards to life and property posed by tephra fall processes generally are limited and less severe. The primary concern relates to collapse of roofs because of the weight of the accumulating tephra (especially if wet)—an effect similar to adding the weight of a rapid and intense fall and accumulation of snow.

Fallout from eruption column. An eruption column is the mushroom-shaped plume of volcanic debris and gas that rises above the vent during an explosive eruption. The narrow bottom part of the plume (the mushroom stem) is where the decompressing, expanding volcanic gases are thrusting upward at high velocity. The upper parts of the plume are sustained by thermal energy and buoyancy, before topping out at some height where the bulk density of the plume equals that of the ambient atmosphere. For powerful and sustained explosive eruptions, their eruption columns can rise tens of kilometers high and can

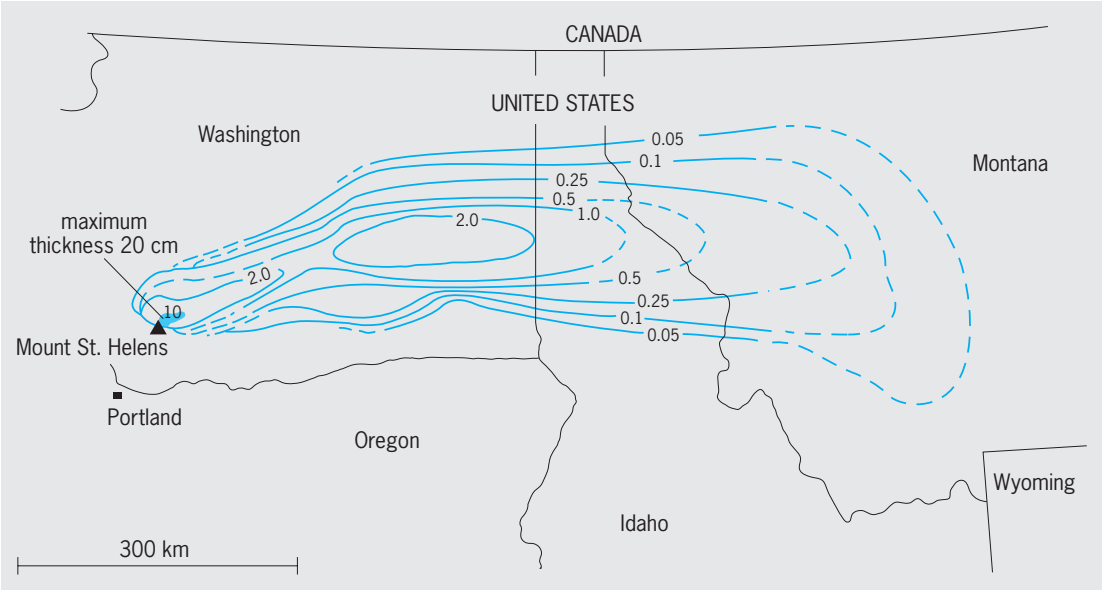

Fig. 1. Distribution downwind of the pyroclastic fall deposit from the climactic eruption of Mount St. Helens, Washington, on May 18, 1980. The contours (isopachs) are drawn through points of measured equal thickness (in centimeters) of this ash layer. (*Modified from USGS Prof. Pap. 1250*)

remain active for hours to days. The largest and heaviest pyroclasts in the plume fall out closest to the vent, and the size of the tephra and the thickness of the resulting fall deposits typically decrease with increasing distance downwind from the vent. *See* VOLCANOLOGY.

Tephra fallout from the climactic eruption of Mount St. Helens (Washington) on May 18, 1980, produced by an eruption column more than 20 km high that was active for about 9 h, formed measurable pyroclastic fall deposits hundreds of kilometers downwind, and covered an area of more than 57,000 km^2 (**Fig. 1**). While impressive, the fall deposits from this eruption of Mount St. Helens pale by comparison to some of the prehistoric explosive eruptions from this volcano and to those produced by more voluminous historical and prehistoric explosive eruptions at other volcanoes: Mount Pinatubo, Philippines, 1991; Novarupta (Katmai), Alaska, 1912; Mount Mazama (Crater Lake), Oregon, 6800 years ago; Yellowstone caldera, Wyoming, which has produced three huge explosive eruptions during the past 2 million years. Because of their widespread distribution, the pyroclastic fall deposits from such powerfully explosive eruptions provide easily identifiable ash layers in the geologic record, and serve as important time markers in geologic mapping studies to reconstruct the geological histories of regions, even those far distant from the eruptive sources. A diagnostic characteristic of such regional ash marker beds is that they tend to drape the surrounding topography, such that their observed thicknesses—in any given direction and distance from the eruptive vent—are similar, whether measured on topographically high elevations or in valleys and other depressions. Tephra from relatively weak explosive eruptions (such as cinder, scoria, and spatter) fall close to the vent, piling up to form conical structures with slopes near or at the angle of repose. *See* STRATIGRAPHY.

Fallout from ash clouds associated with pyroclastic flows. Pyroclastic fall deposits can also be formed from finer-grained pyroclasts settling out from ash clouds that typically roil above pyroclastic flows moving along the ground. Much of the material in such clouds is separated out of the pyroclastic flows by volcanic gases flushing upward through the flows and, once airborne, the material can rise high into the air in turbulent, convecting plumes before settling out. Fall deposits formed in this manner are called ash-cloud-derived fall deposits or co-ignimbrite ash-fall deposits, as opposed to eruption-plume-derived fall deposits. It is not always possible to distinguish unambiguously between these two types of pyroclastic fall deposits. It also should be emphasized that the fall deposits associated with pyroclastic flows can be as voluminous and widely dispersed as those generated from eruption columns. Many of the volcanic ash layers found in deep-sea sediments are believed to be co-ignimbrite ash fall produced by explosive eruptions on land. For example, a prominent ash layer in deep-sea cores taken from the Indian Ocean was derived from fallout from a gigantic explosive eruption at Lake Toba, Sumatra (Indonesia), about 75,000 years ago. *See* IGNIMBRITE.

Flowage processes and deposits. Many but not all pyroclastic rocks are formed by the deposition of hot pyroclasts transported in fast-moving surface flows rather than by deposition of fragments, which are largely cooled to ambient temperature by falling through the air. Even though these flowing masses are heavier than air, they are highly mobile because of their high temperature and gas content. They form pyroclastic flows and surges that sweep down steep slopes and valleys at high speeds (from 10 to several hundred meters per second) and can travel great distances (>100 km) from the eruptive centers. Such flows, which vary considerably in bulk density and flow characteristics, are typically composed of two parts: (1) the pyroclastic flow proper—a

ground-hugging basal flow with a high concentration of solids and a density close to that of the resulting deposit; and (2) the associated pyroclastic surge—an overriding less dense, turbulent cloud of ash with a low concentration of solids and density much less than that of the resulting deposit. Pyroclastic surges flow mostly above rather than hug the ground. Pyroclastic flows and surges are also variously termed glowing avalanches or clouds (*nuées ardentes* in French) or simply ash flows. A less common pyroclastic flow phenomenon is the laterally directed blast (or lateral blast), which appears to share features of both pyroclastic flows and surges. Lateral blasts are generally triggered by the sudden depressurization (uncorking) of a volcanic system by catastrophic collapse at a volcano and associated landsliding.

Because the paths of denser pyroclastic flows are controlled largely by topography, the thicknesses of pyroclastic flow units can vary widely, ranging from less than 0.1 m to more than 100 m where ponded in depressions. If the emplacement temperature and gas content are sufficiently high, the pyroclasts in thick deposits stick together to form welded tuffs or welded ash flows. In steep terrains, pyroclastic flows may drain completely from the upper slopes, so that their deposits are preserved only in the valley floors flanking the volcano. In marked contrast to pyroclastic fall deposits, which invariably become thinner away from the eruptive source, pyroclastic flow deposits can become thicker away from the source. Lower-density pyroclastic flows and surges, because they are little influenced by topography, are able to advance great distances over gently sloping ground, and can surmount ridges hundreds of meters high. The more dense pyroclastic flows tend to fill valleys and other topographic depressions, and the deposits of pyroclastic rock they form can attain great thicknesses. In contrast, the less dense pyroclastic surges, even though they can be equally or more greatly traveled, generally leave much thinner deposits, some of which may not be preserved in the geologic record because of postdeposition erosion and reworking. *See* TUFF.

Pyroclastic flows, with their high abundance of hot solid fragments, are characterized by high temperatures (>800°C), whereas pyroclastic surges and laterally directed blasts can be either hot or cold. Hot pyroclastic surges are almost always associated with pyroclastic flows but, because of their lower ratios of hot solids to gas, rarely exceed 300°C. Cold pyroclastic surges (also called base surges) typically are formed by hydrovolcanic explosions caused by the interaction of rising magma with shallow ground water or surface water, or by the sudden heating of water trapped beneath hot lava flows or pyroclastic flows, such as along streams, lakes, oceans, and ice and snow fields. "Hydrovolcanic" is an inclusive term applied to pyroclastic products formed by explosive interaction between hot volcanic material and water. Hydrovolcanic activity, however, has characteristics quite different from pyroclastic activity *sensu stricto* and commonly produces relatively small volcanic craters (such as maars, tuff rings, and tuff cones) with diameters that are wide relative to the height of their rims.

As with the fall of pyroclastic material into the sea from on-land explosive eruptions, pyroclastic flows and surges from explosive activity of island volcanoes have been observed to enter the sea. Some well-documented examples include the 1883 Krakatau (Java) eruption, the 1902 Mont Pelée eruption, and the 1996–1997 explosive eruptive activity at Soufriére Hills Volcano, Montserrat (Lesser Antilles, Caribbean Sea). The resulting submarine deposits, called subaqueous pyroclastic flows, share many of the features and attributes of their on-land counterparts but differ in details as regards grain-size distribution, texture, and structure. Occurrences of similar-looking pyroclastic rocks, interpreted to be subaqueous in origin, have been recognized throughout the geologic record, dating back more than 2500 million years. Some specialists have proposed that, under certain conditions, pyroclastic flow deposits can also be produced by explosive eruptions at shallow depths underwater. The submarine deposition of pyroclastic rocks can be intermingled with deposition of reworked volcanic debris as well as nonvolcanic sediments, resulting in volcaniclastic aprons that are commonly found offshore of near-ocean continental volcanoes, and around volcanic islands and seamounts. *See* TURBIDITY CURRENT.

Volcanologists have observed or inferred several common flowage processes to generate pyroclastic

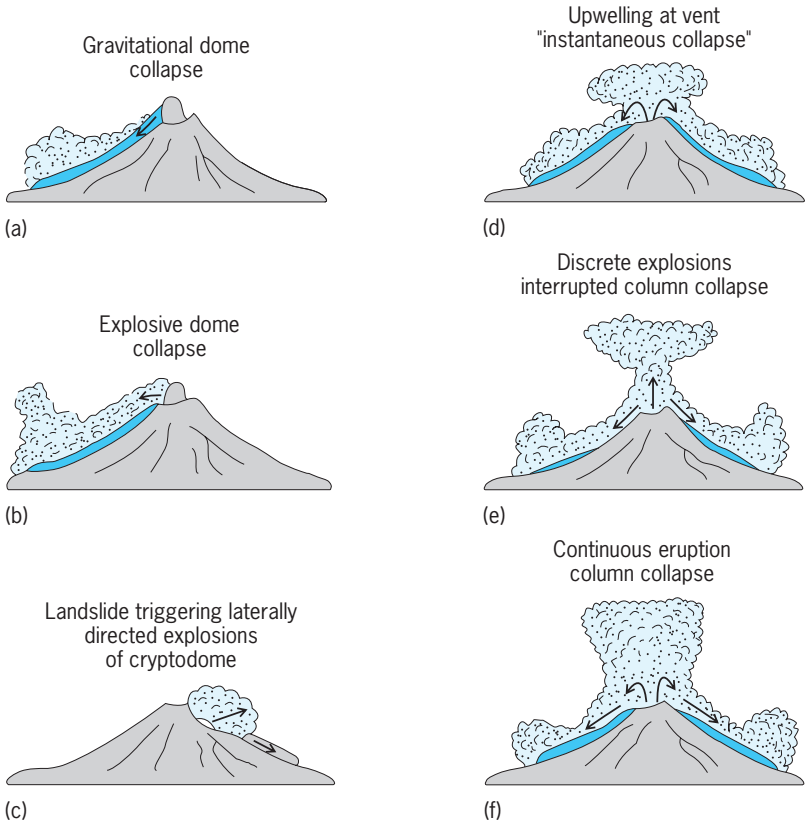

Fig. 2. Observed or postulated mechanisms for the generation of pyroclastic flow and surge deposits (*from W. E. Scott, 1989*). (*a*, *b*) Mechanisms that commonly produce block-and-ash flow deposits. (*b*, *c*) Mechanisms that trigger laterally directed blasts. (*d–f*) Mechanisms that represent the most common of the many proposed origins of pyroclastic flow deposits based on collapse of eruption columns.

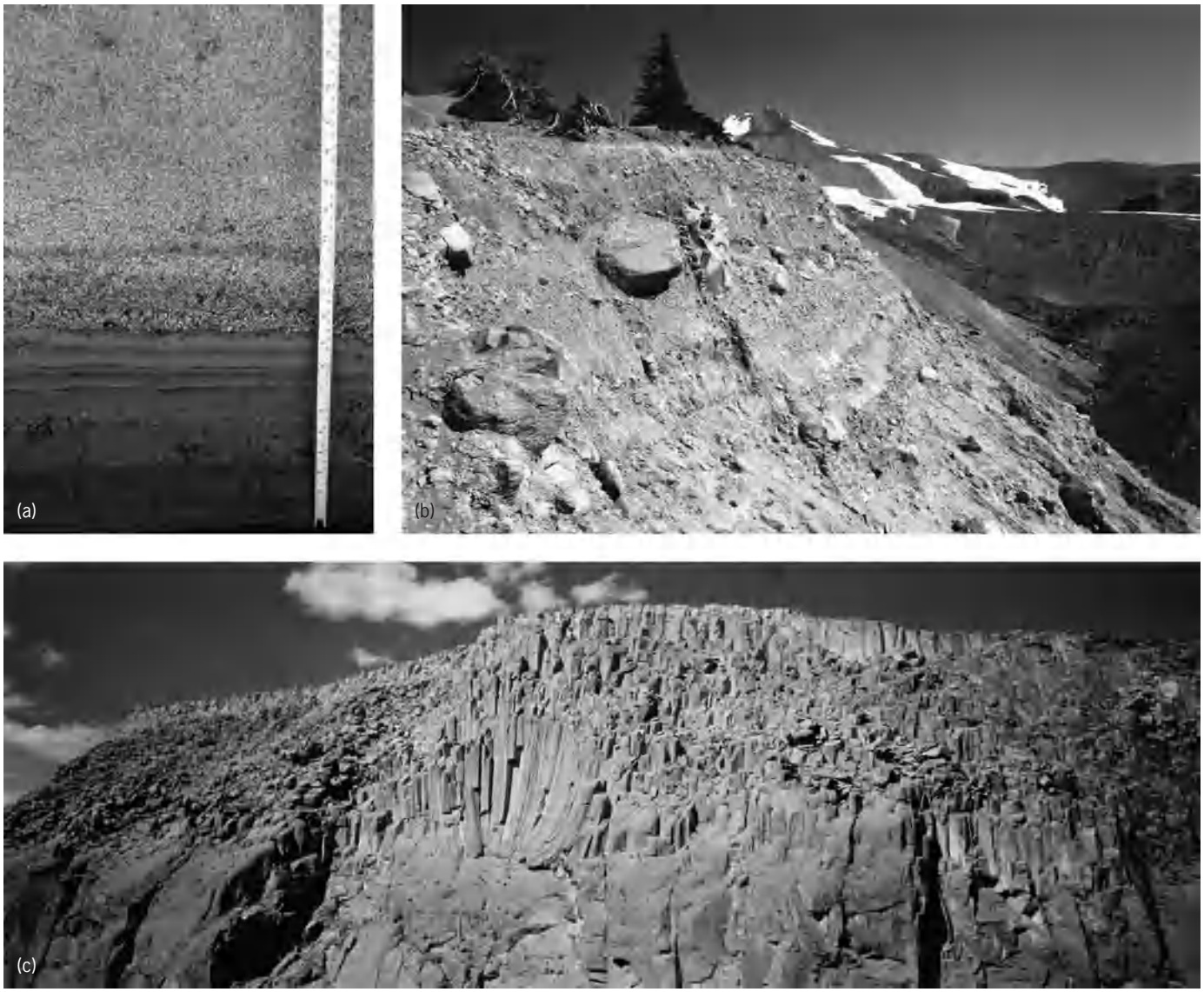

Fig. 3. Some variations in pyroclastic rock formations. (*a*) Layers of pyroclastic fall deposits from the eruption plume of the June 15, 1991, explosive eruption of Mount Pinatubo (Luzon, Philippines). This exposure, located about 9 km (5.5 mi) southeast of the volcano, has a total thickness of about 0.6 m (1.8 ft) (*photograph by W. E. Scott, USGS*). (*b*) Block-and-ash-flow deposit formed by a collapse of a lava dome at Mount Hood Volcano (Oregon) more than 10,000 years ago; similar deposits were formed as recently as about 200 years ago. Some of the blocks are 1 m (3.3 ft) or greater in diameter; the present summit of the volcano can be seen top center (*photograph by D. R. Crandell, USGS*). (*c*) Cliff in the Owens River Gorge showing a massive pyroclastic flow deposit (the Bishop Tuff), which was produced about 760,000 years ago by a huge explosion from Long Valley Caldera (east-central California). At this locality, the deposit is about 150 m (500 ft) thick, is strongly welded in places, and shows spectacular columnar jointing formed from posteruption cooling (*photograph by R. I. Tilling, USGS*).

flows, surges, and directed blasts (**Fig. 2**). Pyroclastic flows formed by collapse of lava domes (Fig. 2*a*,*b*), or of thick lava flows on steep slopes, are called block-and-ash-flow deposits, characterized by large hot blocks (sometimes exceeding 5 m in size) in an ash matrix. Depending on the size and the explosive energy release of the eruption, the triggering mechanism, the texture, structure, and other features of the resulting deposits of pyroclastic rocks can exhibit considerable variation (**Fig. 3**).

Because of their mass, high temperature, high gas content, high velocity, and great mobility, the flowage processes that produce pyroclastic rocks are among the most dangerous and deadly of volcano hazards. These processes can have some or all of the following effects on people and property in their path: asphyxiation, burial, incineration, sand-blasting abrasion, and crushing physical impact. Even worse, hot pyroclastic flows and surges can mix with surface water, or water melted from snow and ice, to produce destructive lahars (volcanic mudflows) and floods that can devastate valleys downstream. Volcanic disasters involving pyroclastic flowage processes that have killed 2000 or more people include Mount Vesuvius, Italy, A.D. 79; Mont Pelée, Martinique, 1902; Mount Lamington, Papua New Guinea, 1951; and El Chichón Volcano, Mexico, 1982.

Robert I. Tilling

Bibliography. R. A. F. Cas and J. V. Wright, *Volcanic Successions: Modern and Ancient*, Chapman & Hall, London, 1988; R. V. Fisher and H.-U. Schmincke, *Pyroclastic Rocks*, Springer-Verlag, Berlin, 1984;

P. Francis, *Volcanoes: A Planetary Perspective*, Oxford University Press, 1993; G. Heiken, *An Atlas of Volcanic Ash*, Smithsonian Contributions to the Earth Sciences, Smithsonian Institution, Washington, DC, 1974; W. E. Scott, Volcanic and related hazards, in R. I. Tilling (ed.), *Short Course in Geology*, vol. 1: *Volcanic Hazards*, American Geophysical Union, Washington, DC, 1989; H. Sigurdsson et al. (eds.), *Encyclopedia of Volcanoes*, Academic Press, San Diego, 2000.

Pyroelectricity

The property of certain crystals to produce a state of electric polarity by a change of temperature. Certain dielectric (electrically nonconducting) crystals develop an electric polarization (dipole moment per unit volume) ΔP when they are subjected to a uniform temperature change ΔT. For a small change ΔT, the components ΔP_i of the polarization vector are given by Eq. (1), where the p_i are constant coefficients. This pyroelectric effect occurs only in crystals which lack a center of symmetry and also have polar directions (that is, a polar axis). These conditions are fulfilled for 10 of the 32 crystal classes. Examples of pyroelectric crystals are tourmaline, lithium sulfate monohydrate, cane sugar, and ferroelectric barium titanate. *See* CRYSTAL STRUCTURE; DIELECTRIC MATERIALS; POLARIZATION OF DIELECTRICS.

$$\Delta P_i = p_i \, \Delta T \qquad i = x, y, z \tag{1}$$

Pyroelectric crystals can be regarded as having a built-in or permanent electric polarization. When the crystal is held at constant temperature, this polarization does not manifest itself because it is compensated by free charge carriers that have reached the surface of the crystal by conduction through the crystal and from the surroundings. However, when the temperature of the crystal is raised or lowered, the permanent polarization changes, and this change manifests itself as pyroelectricity. There is no easy way to determine the magnitude of the total permanent polarization, except for those special pyroelectric crystals, called ferroelectrics, in which the polarization can be reversed by an electric field. *See* FERROELECTRICS.

Magnitude. The magnitude of the pyroelectric effect depends upon whether the thermal expansion of the crystal is prevented by clamping or whether the crystal is mechanically unconstrained. In the clamped crystal, the primary pyroelectric effect is observed, whereas in the free crystal, a secondary pyroelectric effect is superposed upon the primary effect. The secondary effect may be regarded as the piezoelectric polarization arising from thermal expansion. The secondary effect is generally much larger than the primary effect. When heated nonuniformly, even piezoelectric crystals produce a polarization due to the temperature gradients and corresponding nonuniform stresses and strains. This phenomenon is called tertiary, or false, pyroelectricity, because it may be found in nonpyroelectric crystals, that is, crystals which have no polar direction. *See* PIEZOELECTRICITY.

In a typical pyroelectric crystal such as tourmaline, a temperature change of 1.8°F (1°C) produces at room temperature a polarization of about 10^{-5} coulomb/m^2. The same amount of polarization can be produced in a crystal at constant temperature by applying an external electric field of about 70 kV/m. The pyroelectric effect is much larger in ferroelectrics, in particular at temperatures close to the Curie point. The figure of merit, defined as the pyroelectric coefficient divided by the permittivity, is high for ferroelectrics such as triglycine sulfate. Thermal response times of 2 microseconds and less have been reported. *See* PERMITTIVITY.

Inverse effect. As can be shown in thermodynamics, the pyroelectric effect has an inverse, the linear electrocaloric effect. A temperature change ΔT results when the permanent polarization is altered by an externally applied electric field ΔE, as in Eq. (2).

$$\Delta T = \frac{T}{\rho C_p} \sum_i p_i \, \Delta E_i \qquad i = x, y, z \tag{2}$$

In this expression, ρ is the density, C_P the specific heat at constant pressure, p_i are the pyroelectric coefficients, and T is the absolute temperature. The temperature changes that can be realized in typical pyroelectrics are of the order of magnitude 0.02°F (0.01°C).

H. Granicher

Applications. Pyroelectrics have a broad spectrum of potential scientific and technical applications. The most developed is the detection of infrared radiation.

Pyroelectric thermometer. In this instrument a thin pyroelectric plate cut perpendicular to the spontaneous polarization direction is covered with electrodes on both faces. Because of dielectric losses and any bias resistance, there is a resistance R in parallel to this special capacitor, and the potential difference ΔV between electrodes is nil at equilibrium. Let $\Delta T(t)$ be the increase of temperature at time t; a corresponding increase in spontaneous polarization given by Eq. (3) will appear.

$$\Delta P_s = p \, \Delta T \tag{3}$$

If the electric time constant $\tau' = RC$ (where C is the capacitance between the plates) is sufficiently long, the flow of electric charges through R may be neglected and voltage given by Eq. (4) will appear,

$$\Delta V = \frac{\Delta P_s}{C_0} \tag{4}$$

proportional to ΔP_s, and hence to ΔT; here C_0 is the capacitance per unit electrode area. Temperature changes as small as 10^{-5} °F (6×10^{-6} °C) can be detected with a vibrating electrometer in a 1-s time interval.

Pyroelectric manometers. A thermal detector in a gas enclosure is sensitive to small variations of pressure, and the enclosure must be evacuated to avoid significant difficulties. This is simply explained by assuming that any rapid compression is adiabatic and thus leads to an increase of temperature. This effect may be used to measure pressure changes with a pyroelectric thermometer. The effect is especially

large in such detectors, where it may be enhanced by their piezoelectric sensitivity to hydrostatic pressure changes.

Pyroelectric radiation detectors. In most pyroelectric radiation detectors, the radiation is absorbed by a thin electrode whose thickness can be adjusted to make the absorption higher than 50% for the whole electromagnetic spectrum. Some dielectric coatings have absorptions of nearly 100% over a wide spectral interval. There is an increase of temperature and thus an electric signal ΔV. These devices generally detect infrared radiation, but have been used from the far ultraviolet to microwaves. Their normalized detectivity D^* at low chopping frequencies (2–15 Hz) is within an order of magnitude of the limit imposed by background radiation at room temperature (80°F or 300 K), given by Eq. (5). The pyroelectric detector is thus comparable to the best thermal detectors (such as the pneumatic detector), and is superior to thermopiles and bolometers. *See* BOLOMETER; THERMOCOUPLE.

$$D^*_{\text{lim}}\,(300\text{ K}) \approx 2 \times 10^{10}\ \text{W}^{-1} \cdot \text{cm} \cdot \text{Hz}^{1/2} \qquad (5)$$

At higher chopping frequencies the detectivity is significantly lower than D^*_{lim} (300 K). The main reason is that the internal resistance of the crystal is not constant but is purely dielectric, that is, inversely proportional to the chopping frequency, ω. Thus, Johnson noise decreases as $\omega^{-1/2}$, whereas it is constant for a bolometer or a thermocouple, and so the detectivity of pyroelectric detector can be expected to be proportional to $\omega^{-1/2}$ as well.

However, the pyroelectric radiation detector is still considerably better than all other thermal detectors, and is, moreover, mass produced and inexpensive. It is used in fast repetitive Fourier-transform spectroscopy, burglar alarms, and so forth. It is also possible to detect fast particles (for example, H^+ or He^+) from the temperature increase they produce in the pyroelectric.

In some pyroelectric materials a short pulse of radiation (of the order of 10^{-12} s) can produce an instantaneous electric dipole either in an impurity (such as $LiTaO_3 : Cu^{2+}$) or in the crystalline unit cell (because the average value of the dipole moment associated with an anharmonic oscillator can be significant). The result is a quasi-instantaneous pulse of spontaneous polarization followed by an electronic or phonon relaxation, respectively, with a very short time constant (of the order of 10^{-11} s). Special detectors based on this effect can be as fast as quantum detectors. *See* OPTICAL DETECTORS.

Radiometry and pyrometry. Pyroelectric detectors can be used to measure the power generated by a radiation source (in radiometry), or the temperature of a remote hot body (in pyrometry, with corrections due to deviations from the blackbody emission). The temperature change ΔT in Eq. (3) is proportional to the flux of radiant energy impinging on the pyroelectric detector; hence the electric signal ΔV in Eq. (4) is also proportional to this flux. Pyroelectric detectors were first used in the *Nimbus* meteorological satellites [deuterated triglycine sulfate (DTGS) detectors], in the *Pioneer* Venus orbiter, launched on May 20, 1978 (eight DTGS detectors), and in its sounder probe, which descended into the atmosphere of Venus [six lithium tantalate ($LiTaO_3$) detectors, which gave information on the carbon dioxide emission spectrum of Venus and hence on the temperature at different levels]. *See* PYROMETER; RADIOMETRY.

Imaging. An infrared image can be projected on a pyroelectric plate and transformed into a relief of polarization on the surface. The electric charges are read with a beam of electrons as in an ordinary television vidicon. The performances are such that it is actually possible to see in the dark. For instance, temperature differences of 0.9°F (0.5°C) in the landscape can be detected with 15 frames per second. However, a part of the detectivity is lost by the electron beam reading, and still higher performances are expected from linear or two-dimensional arrays of single pyroelectric detectors. *See* INFRARED IMAGING DEVICES.

Pyroelectric probe technique. This technique is one of the best ways to observe ferroelectric domains. A ferroelectric plate is cut perpendicular to the polar axis and coated with electrodes on both sides. A laser spot is swept over one electrode and gives a pyroelectric signal at each inspected point with a plus or minus sign according to the direction of the spontaneous polarization at this point. This technique is capable of inspecting a 100 × 100 micrometer area in a few seconds with a 2-μm limit of resolution. With a 200-kHz chopping frequency only a thin superficial layer of the crystal is inspected. (In the case of triglycine sulfate, the thermal diffusion length is 0.64 μm for that frequency.)

Other applications. Other potential applications of pyroelectricity include solar energy conversion, refrigeration, information storage, geology, and solid-state science.

The pyroelectric radiation detector can be considered as a high-impedance thermoelectric converter, which could be used for solar energy conversion. However, the efficiency achieved so far is low (1%). *See* SOLAR ENERGY.

The use of pyroelectricity for cooling is closely related to adiabatic demagnetization. The temperature has been lowered by 0.5°F (0.3 K) at −442°F (10 K) with strontium titanate ($SrTiO_3$). Larger effects would require new materials. *See* ADIABATIC DEMAGNETIZATION.

When a pyroelectric is also a ferroelectric, the local polarization can be reversed by a laser spot to record information. The reading can also be done with a laser spot. The process is fast (10^6 bits s^{-1}). The signal is much larger than expected, resulting in a high signal-to-noise ratio. *See* COMPUTER STORAGE TECHNOLOGY.

The interior of the Earth has been suggested to be ferroelectric due to the effect of high temperature and pressure on such compounds as magnesium silicate ($MgSiO_3$). This could explain the change of acoustic impedance which has been observed at a depth of 400 mi (700 km).

Pyroelectricity may be related to anharmonic

phonons, as discussed above. In other materials it is connected with atomic displacements. The order of magnitude is 10^{-14} m/°C, while the root-mean-square amplitude of vibrations at room temperature is around 10^{-11} m. The Ti^{4+} displacement at the barium titanate ($BaTiO_3$) transition is up to 2×10^{-11} m, which explains the large spontaneous polarization. Pyroelectricity is thus a very sensitive method of measuring small atomic displacements and can help specify the structure of a pyroelectric crystal. A. Hadni

Bibliography. A. Hadni, Applications of the pyroelectric effect, *J. Phys.,* E14:1233–1240, 1981; A. Hadni and X. Gerbaux, Infrared and millimeter wave absorber structures for thermal detectors, *Infrared Phys.*, 30:465–478, 1990; S. B. Lang, *Pyroelectricity*, 1981; S. B. Lang, Review of recent work on pyroelectric applications, *Ferroelectrics*, 53:189–195, 1984; M. E. Lines and A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials*, 1977, reprinted 2000; J. F. Nye, *Physical Properties of Crystals: Their Representation by Tensors and Matrices*, 1960, reprinted 1987.

Pyroelectricity (biology)

Electrical polarity in a biological material produced by a change in temperature. Pyroelectricity is probably a basic physical property of all living organisms. First discovered in 1966 in tendon and bone, it has since been shown to exist in most animal and plant tissues and in individual cells. Pyroelectricity appears to play a fundamental part in the growth processes (morphogenesis) and in physiological functions (such as sensory perception) of organisms. *See* PYROELECTRICITY.

The elementary components (for example, molecules) of biological (as well as of nonbiological) pyroelectric structures have a permanent electric dipole moment, and are arranged so that all positive dipole ends point in one direction and all negative dipole ends in the opposite direction. This parallel alignment of elementary dipoles is termed spontaneous polarization (P_s, electric dipole moment per unit volume) because it occurs spontaneously without the action of external fields or forces. In this state of molecular order, the structure concerned has a permanent electric dipole moment on a microscopic and macroscopic level (**Fig. 1***a*). *See* DIPOLE MOMENT.

Spontaneous polarization is temperature-dependent; thus any change in temperature causes a change of the dipole moments, measurable as a change of electric charges at both ends of the polar axis. This is the pyroelectric effect (Fig. 1*b*). All pyroelectric structures are also piezoelectric, but the reverse is not true. *See* FERROELECTRICS; PIEZOELECTRICITY.

The existence of an inherent spontaneous polarization (an inherent permanent dipole moment) in biological structures cannot be demonstrated at a constant temperature because the fixed charges on both dipole ends are compensated by free charge carriers within the polar material or from its environment. But the presence of a permanent dipole moment can be established with the aid of the pyroelectric effect by sudden changes of temperature, for example, those produced by the absorption of light flashes (radiant heating, Fig. 1*b*).

Molecular conditions. Prerequisites for the development of spontaneous polarization and pyroelectric activity in biological structures are (1) the presence of a permanent dipole moment in the molecules or molecular aggregates and (2) a molecular shape that favors a parallel alignment as much as possible (or at least does not impede it). Both these conditions are ideally fulfilled in rod-, bar- or board-shaped molecules with a permanent dipole moment along the longitudinal molecular axis. Such molecules readily take on parallel alignment and thus satisfy the physical preconditions for pyroelectric behavior in a defined axial direction.

Several important organic substances have these molecular properties, and therefore behave pyroelectrically in biological structures. The fibrous

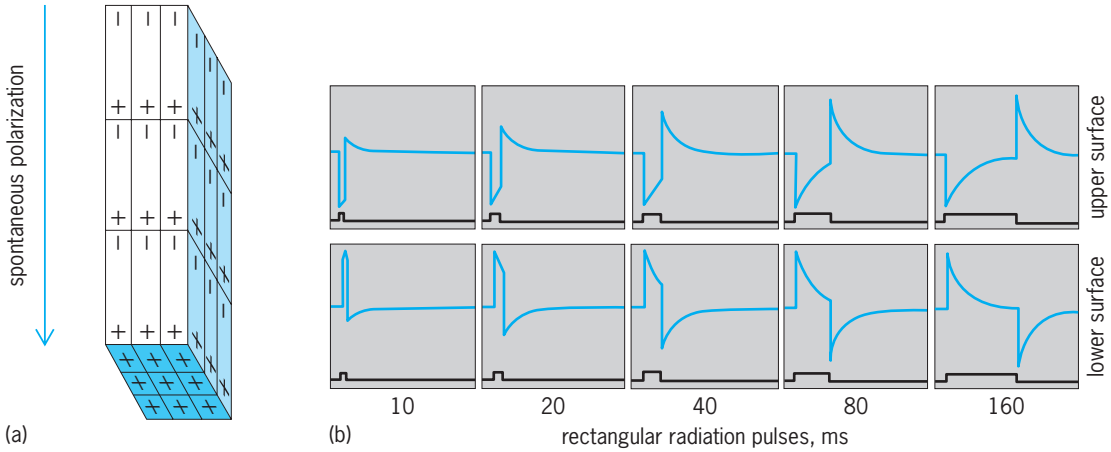

Fig. 1. Schematic diagram of spontaneous polarization and pyroelectric behavior. (*a*) Spontaneous polarization of a polar (pyroelectric) body. (*b*) Pyroelectric responses of upper and lower surfaces of a Z-cut tourmaline disk (a polar body) to rectangular radiation pulses. The responses of both surfaces have opposite electric signs (polar behavior). These are traces of a transient recorder; calibration, vertical 50 mV. The voltage–time course of the pyroelectric responses depends on the rate of change of temperature.

protein collagen is one example. Collagenous tissues (such as bone, cartilage, tendon, and ligaments) are thus strongly pyroelectric. *See* COLLAGEN.

Morphological polarity and inherent spontaneous polarization. Spontaneous polarization and pyroelectric behavior occur in the direction of distinct polar axes. A coincidence of these physical axes and axes of morphological polarity in biological structures is demonstrable with pyroelectric methods.

Individual cells. The polar properties of organisms, their organs, and systems originate in the polarity of cells. To establish which processes cause cell polarity and how to account for its stability during the lifetime of the cell is a principal task of biology. The giant cells of some green algae are favorite test objects for the investigation of cell polarity; both longitudinal and radial cell polarity have been demonstrated.

A permanent electric dipole moment coinciding with the (longitudinal) axis of cell polarity has been found in live individual cells of the green algae *Acetabularia mediterranea* and *Valonia utricularis*. The cells show a polar orientation in a high-voltage dc field if the cell temperature is rapidly changed. The apical cell pole then points to the cathode, the basal pole to the anode (**Fig. 2***a*). The polar orientation is maintained if the direction of the dc field is reversed; that means the apical cell pole turns again to the cathode. Under identical experimental conditions a tourmaline crystal rod shows similar polar orientation.

The cell walls of live *Acetabularia* and *Valonia* cells react pyro- or piezoelectrically on their entire surface to thermal or mechanical stimulations (Fig. 2*b*). Pyroelectric and piezoelectric signals of the outer cell wall surface have opposite electric signs to those of the inner surface; the cell wall is thus polar. The direction of the radial polarization is perpendicular to the cell wall surface.

Epidermis of animals and plants. The epidermis, the outermost cellular layer of animals and plants, represents a highly important interface between the living organism and its environment. The epidermis of invertebrates and of plants is usually a monocellular layer. Pyroelectric measurements on numerous living organisms (invertebrates and vertebrates) as well as on a great many living higher plant species confirmed the existence of a spontaneous polarization perpendicular to the epidermal surface and coinciding with the axis of cell polarity in these layers (**Fig. 3***a*). Pyroelectric and piezoelectric responses of the animal and plant epidermis are analogous. Pyroelectric responses of outer and inner epidermal surfaces have opposite electric signs (Fig. 3*b*), thus proving polar behavior.

The animal and plant epidermis functions on its entire surface as a pyroelectric detector and transducer, converting distinct kinds of physical energy from the environment (for example, rapid changes in temperature, uniaxial and hydrostatic pressure) into electrical energy. Thus the live epidermis has physical properties analogous to nonbiological polar materials which function in pyroelectric detectors and transducers designed for technical applications.

Sensory receptors. Cutaneous receptors are mostly transformed epidermal structures. Since all structures of the animal and plant epidermis that have been investigated showed pyroelectric activities, it may be assumed a priori that cutaneous sensory receptors will behave pyroelectrically (and thus also piezoelectrically). Two examples are the auditory membrane of the locust ear and the infrared-sensitive organ of rattlesnakes.

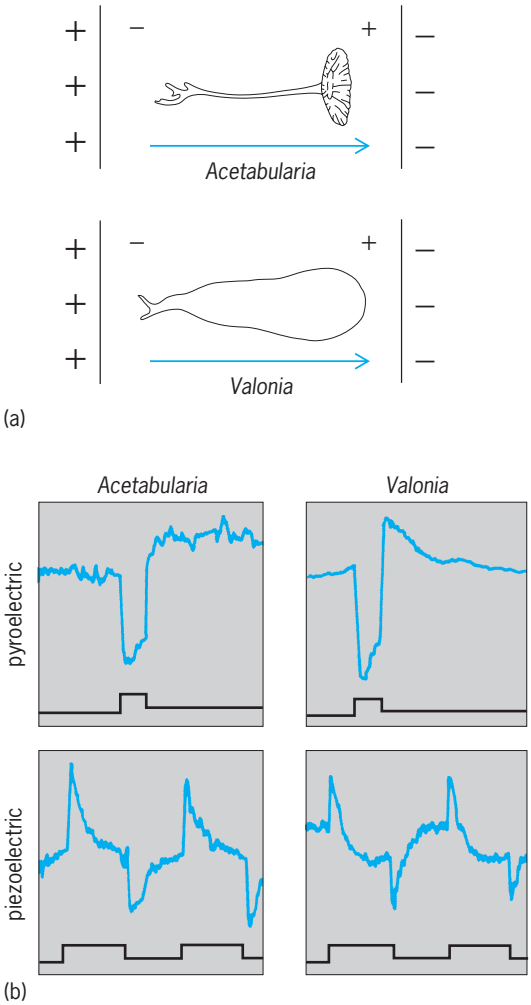

Fig. 2. Longitudinal and radial directions of spontaneous polarization in individual cells. (*a*) Longitudinal polarization; live cells show polar orientation (apical cell pole toward the cathode) in a dc field if the cells' temperature is rapidly changed. (*b*) Radial polarization; the cell walls of live *Acetabularia* and *Valonia* cells show pyroelectric and piezoelectric behavior perpendicular to their surfaces. Pyroelectric voltage responses to rectangular radiation pulses of 20 ms; piezoelectric voltage responses to rectangular mechanical pulses of 100 ms. Calibration: pyroelectric, vertical 5 mV; piezoelectric, 20 mV.

The auditory membrane (tympanum) of the locust ear (**Fig. 4***a–c*) is a part of the insect exoskeleton, the epicuticle. The epicuticle of the arthropod integument has pronounced pyroelectric properties; its pyroelectric coefficient can reach or exceed that of tourmaline. The tympanum of the locust ear acts as a piezoelectric mechanoreceptor (Fig. 4*d*), but it has pyroelectric properties as well. The polar axis of the tympanum is perpendicular to its surface, as it is in the surrounding epicuticle.

The infrared-sensitive organ of rattlesnakes has a

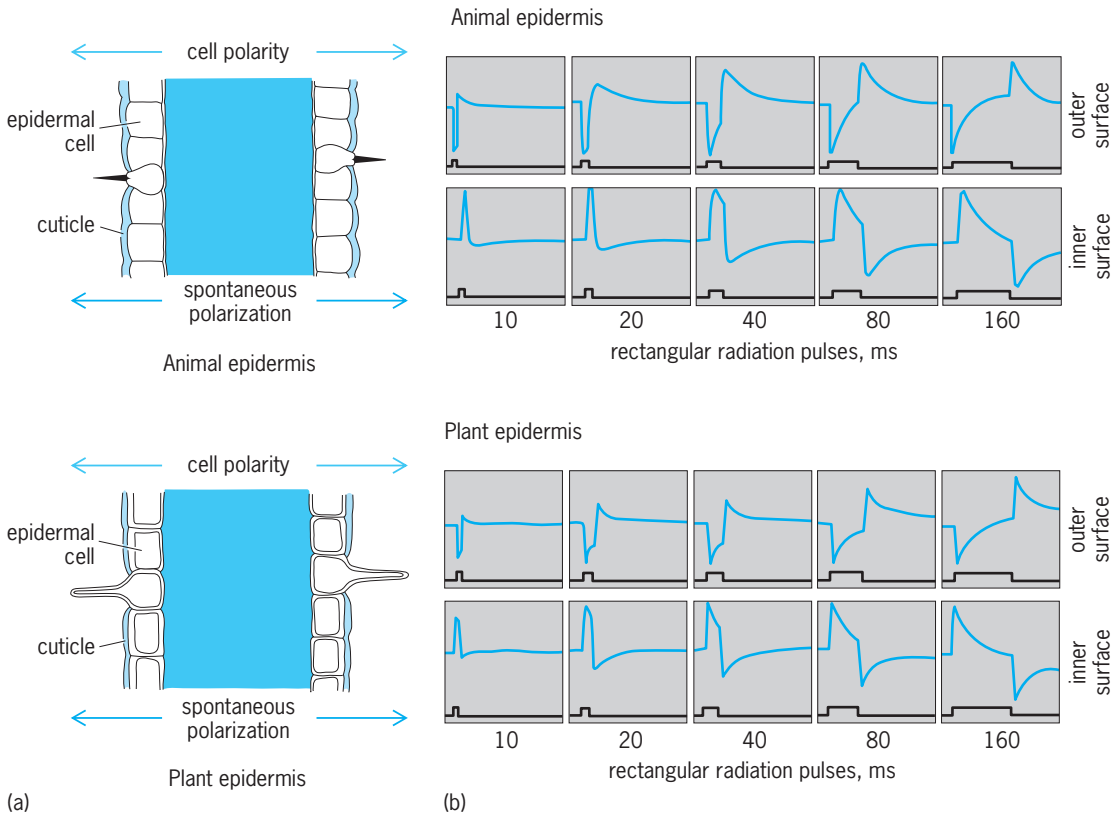

Fig. 3. Spontaneous polarization and pyroelectric behavior of animal and plant epidermis. (*a*) Uniform direction of spontaneous polarization in the animal and plant epidermis. (*b*) Opposite electric signs of pyroelectric voltage responses on the outer and the inner epidermal surface (polar behavior); compare Fig. 1*b*. Animal epidermis: *Periplaneta americana*, cockroach; plant epidermis: *Yucca elephantipes*, yucca. Calibration, vertical 20 mV (*Periplaneta*) and 5 mV (*Yucca*).

similar topography to the locust ear. Its sensitive membrane is suspended between two air-filled cavities and contains numerous nerve endings oriented perpendicular to the membrane's surface. This membrane has strong pyroelectric properties and reacts to minute changes of temperature with electrical signals corresponding to those of technically used pyroelectric infrared detectors, but also has piezoelectric properties.

A physical reaction mechanism common to all sensory receptors is suggested by the following facts. All receptor cells of animals are very similar in their morphological structure and cytochemical organization, in spite of certain differences related to their specific functions. All thermoreceptors, photoreceptors, mechanoreceptors, as well as other receptors, react to the various stimuli (such as variation in temperature, illumination, or mechanical stress) with analogous electric responses (receptor potentials or generator potentials). However, the reactions of one receptor are in most cases not limited to only one external effect. For example, thermoreceptors in the skin react to changes of temperature and mechanical stress with analogous responses; the lateral-line organ of fishes and amphibians can act as a thermal, electrical, chemical, or mechanical receptor; and, in the case of the ampullae of Lorenzini (sense organs found in organisms such as the skate), a function as thermal, electrical, and mechanical receptor has been established. *See* SENSE ORGAN.

Nervous system. The spinal cord of vertebrates constitutes a permanent electric dipole in a longitudinal direction, the negative pole being oriented cranially and the positive pole caudally (**Fig. 5**). Each segment of the spinal cord cut exactly perpendicular to its longitudinal axis shows pyroelectric voltage responses of opposite signs on its cranial and caudal surfaces. Fresh spinal cord preparations thus show a polar orientation in a high-voltage dc field, if the temperature of the specimens is rapidly changed.

A relationship between the direction of longitudinal spontaneous polarization and the physiological function of the central nervous system is evident. Specific reactions of the Chordata (from *Amphioxus* up to humans) that are produced by lengthwise application of dc fields (direct currents) in the spinal cord are dependent on the direction of the direct current (electric narcosis in one direction and electric excitation in the opposite direction; Fig. 5*c* and *d*). This functional polarity of the spinal cord can be explained by the existence of a longitudinal permanent dipole moment.

Experiments on the spinal cord have been continued on the optic nerve (Fig. 5*e*), a brain structure in which the nerve fibers chiefly run in one direction (from the retina of the eye into the brain). In the optic nerve a longitudinal permanent dipole moment of the same direction as in the spinal cord, and pyroelectric and piezoelectric behavior have been established. Thus it is probable that the individual neuron (Fig. 5*f*) has a longitudinal permanent dipole moment and pyroelectric and

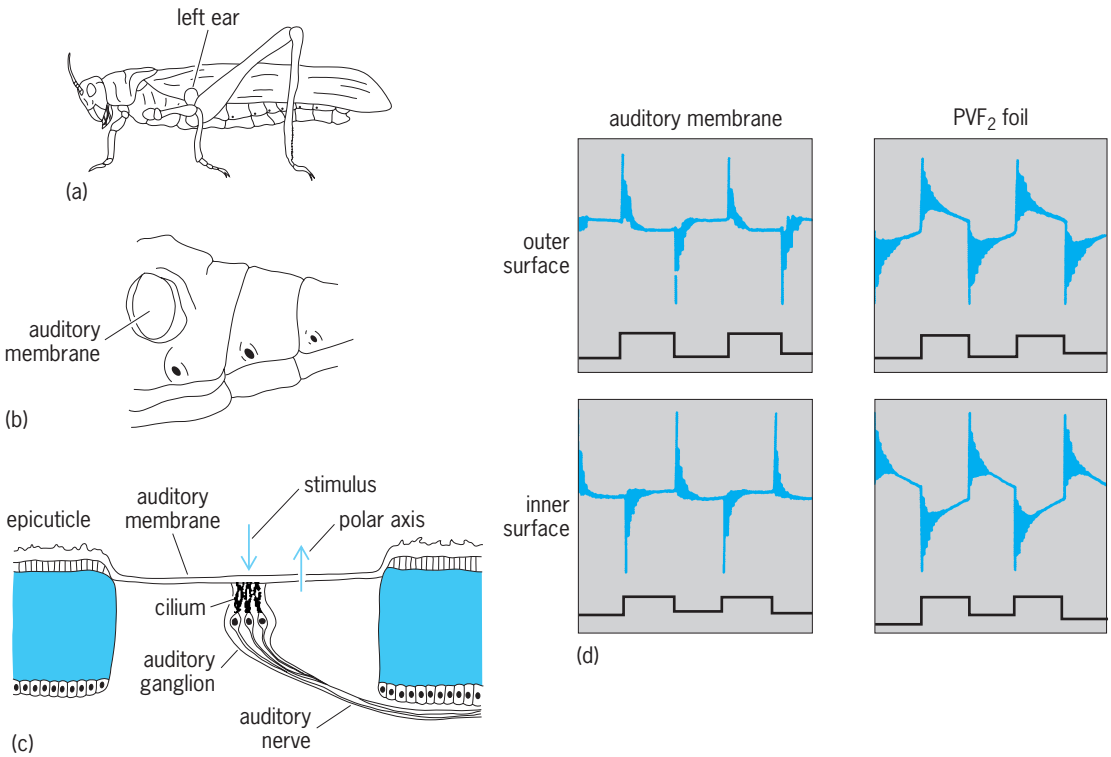

Fig. 4. Piezoelectric behavior of an animal mechanoreceptor (ear of locust). (*a*) *Locusta migratoria*, locust, and its left ear (the two ears of locust are situated on both sides of the first abdominal segment). (*b*) Auditory membrane of the ear. (*c*) Cross section of the ear. (*d*) Piezoelectric voltage responses of outer and inner surface of the auditory membrane and (for comparison) of both surfaces of a piezoelectric polyvinylidene (PVF_2) foil to rectangular pressure pulses of 30 ms. Calibration, vertical 10 mV (auditory membrane) and 100 mV (PVF_2).

piezoelectric properties, too.

Skeletal system. In collagenous tissues (such as bone, cartilage, tendon, and ligaments) of the vertebrate skeleton from the Cyclostomata (such as the *Petromyzon* lamprey) up to humans, a spontaneous polarization pattern has been established that obeys a natural law and is uniform in homologous structures. This polarization pattern is closely related to the morphological polarity of the tissues concerned and to their embryonic development.

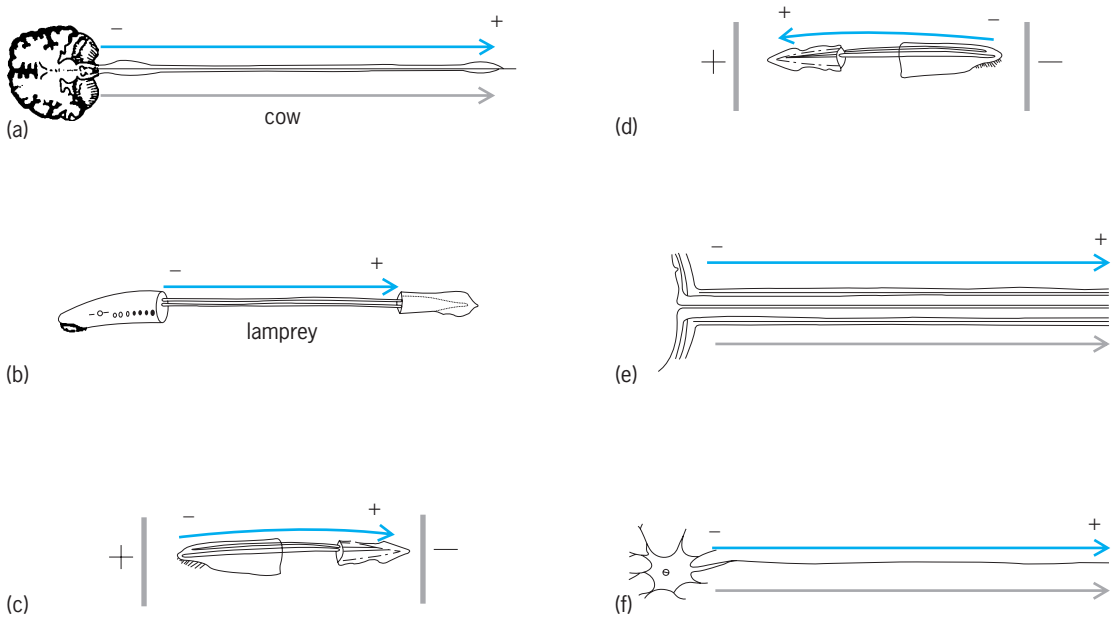

Fig. 5. Longitudinal permanent dipole moment (color arrows) in nerve structures. (*a, b*) Spinal cords of vertebrates: cow and lamprey. (*c*) Cranial end of spinal cord oriented in a dc field to anode causes narcosis and (*d*) cranial end of spinal cord in a dc field oriented to cathode causes excitation (shown for one example, *Branchiostoma lanceolatum*). (*e*) Optic nerve; this brain structure is chiefly composed of nerve fibers having one direction. (*f*) Hypothetical direction of longitudinal permanent dipole moment of an individual neurite. Gray arrows indicate direction of growth.

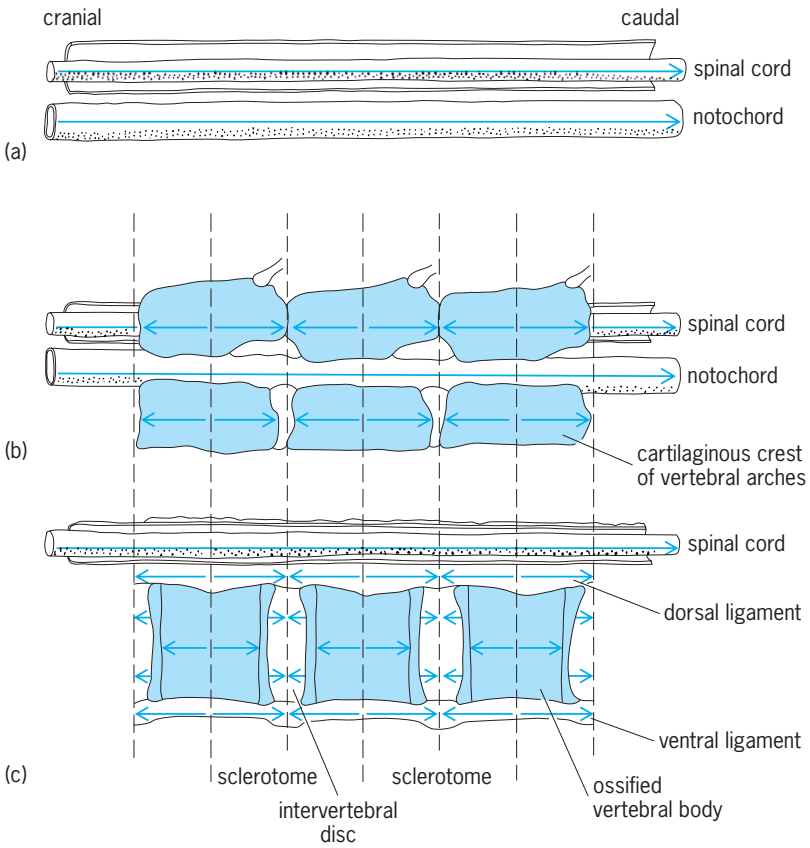

Fig. 6. Directions of spontaneous polarization (arrows) in the axial skeleton of vertebrates. (*a*) *Petromyzon marinus*, lamprey. (*b*) *Acipenser sturio*, sturgeon. (*c*) Human. Boundaries between sclerotomes and the cranial and caudal scleromites are as indicated.

Figure 6 shows experimental results which have been obtained by cutting a continuous series of thin disks at right angles to the longitudinal axis of the tissues under investigation, and by taking individual pyroelectric measurements on each one.

Axial system of Chordata. The chorda dorsalis (notochord), central nervous system, and bilateral row of somites together form the axial system of the chordate germ. The longitudinal polarization of these structures, examined in numerous specimens of juvenile and adult chordates, has corresponding directions throughout the whole range of Chordata, in all stages of development investigated. The direction of the longitudinal dipole moment is the same in notochord and spinal cord; the negative pole lies cranially, the positive pole caudally (Fig. 6*a* and *b*). In the somites the direction of the longitudinal dipole moments changes on the boundaries of the sclerotomes (Fig. 6*c*).

The direction of permanent dipole moments (spontaneous polarization) in all investigated structures of the axial and skeletal systems of vertebrates corresponds with the directions of growth in the structures concerned.

Biological systems as pyroelectric detectors and transducers. Living organisms are able to detect and discriminate between different stimuli in the environment, such as rapid changes of temperature, of illumination, and of hydrostatic and uniaxial pressure. These stimuli represent different forms of energy and are transduced, or converted, into the nearly uniform type of electrical signals whose voltage–time course frequently depends on dX/dt (X = external stimulus, t = time). Such electrical signals have been recorded on cutaneous sensory receptors, on external nerve endings, on epidermal structures, and even on the cell wall of single-cell organisms.

The mechanisms of detection and transduction in these biological systems, still little understood, may lie in the pyroelectric behavior of the structures. Pyroelectric (and thus piezoelectric) behavior has been proved to exist in most biological systems, which means that these systems should in principle be able to function as pyroelectric detectors and transducers. This function is of exceptional significance in epidermal structures, in sensory receptors, and in the nervous system. Herbert Athenstaedt

Bibliography. H. Athenstaedt, Biological systems as pyroelectric detectors and transducers, *Jap. J. Appl. Phys.*, pp. 177–180, 1985; H. Athenstaedt, Epidermis of human skin: Pyroelectric and piezoelectric sensor layer, *Science*, 216:1018–1020, 1982; H. Athenstaedt, "Functional polarity" of the spinal cord caused by its longitudinal electric dipole moment, *Amer. J. Physiol.*, 247:482–487, 1984; H. Athenstaedt, Pyroelectric and piezoelectric properties of vertebrates, *Ann. N.Y. Acad. Sci.*, 238:68–94, 1974.

Pyrolysis

A chemical process in which a compound is converted to one or more products by heat. By this definition, reactions that occur by heating in the presence of a catalyst, or in the presence of air when oxidation is usually a simultaneous reaction, are excluded. The terms thermolysis or thermal reaction have been used in essentially the same sense as pyrolysis.

A simple example of pyrolysis is the classic experiment in which oxygen was first prepared by heating mercuric oxide [reaction (1)]. Similar reactions

$$HgO \xrightarrow[(570°F)]{300°C} Hg + {}^{1}/_{2}O_2 \qquad (1)$$

occur with numerous other metallic oxides and salts. Thermal decomposition or calcining of limestone (calcium carbonate) is the basic step in the manufacture of lime [reaction (2)].

$$CaCo_3 \xrightarrow[(2200°F]{1200°C} CaO + CO_2 \qquad (2)$$

See LIME (INDUSTRY).

Role in organic chemistry. The term pyrolysis is most commonly associated with thermal reactions of organic compounds. In very early chemical investigations, pyrolysis of material from plant and animal sources provided some of the first clues about constitution, as in the formation of isoprene from the thermal breakdown of rubber. Destructive distillation of wood in the absence of air gave "pyroligneous acid," from which methanol and acetic acid were isolated. Similar treatment of bone yielded pyrrole. A range of substances, including benzene, naphthalene, pyridine, and many other aromatic compounds,

was obtained from coal tar, a pyrolysis product of coal. All of these pyrolysis processes lead to formation of volatile products characteristic of the source and also residues of char with high carbon content. *See* DESTRUCTIVE DISTILLATION.

Pyrolysis reactions in contemporary work have been used as preparative methods and as means of generating transient intermediates that can be trapped or observed spectroscopically, or quenched by a further reaction. For preparative purposes, pyrolysis can generally be carried out by a flow process in which the reactant is vaporized with a stream of inert gas through a heated tube, sometimes at reduced pressure. Temperatures range from 400 to 600°C (750 to 1100°F). In flash vacuum pyrolysis (FVP), the apparatus is placed under very low pressure, and the material to be pyrolyzed is vaporized by molecular distillation, usually through a very short zone in a quartz tube at 600–1000°C (1100–1800°F). Under these conditions, contact times are on the order of 0.01 s, and initial unimolecular decompositions can be observed. *See* CHEMICAL DYNAMICS.

Types of reactions. Thermally activated molecules undergo several types of reactions that can be broadly grouped as dissociation into free radicals, elimination of a small molecule, or carbon-carbon bond fission. In many of these pyrolyses, subsequent reactions occur after the initial event.

Dissociation into radicals. At temperatures of 600–800°C (1100–1500°F), most organic compounds acquire sufficient vibrational energy to cause breaking of bonds with formation of free radicals. Early evidence for the transient existence of free radicals was the formation of a metallic mirror when tetraethyllead [$(CH_3)_4Pb$] was vaporized through a heated tube, and removal of the mirror when the tube was heated at a point upstream from the metal deposit. These observations indicate the occurrence of overall reaction (3).

$$(CH_3)_4Pb \rightleftharpoons \underset{\text{Methyl radical}}{4CH_3\cdot} + Pb \qquad (3)$$

See FREE RADICAL.

Alkanes undergo rupture of carbon-hydrogen (C-H) and carbon-carbon (C-C) bonds to two radicals that then react to give lower alkanes, alkenes, hydrogen, and also higher-molecular-weight compounds resulting from their recombination [reactions (4)].

$$CH_3CH_2CH_2CH_3 \xrightarrow[(1300°F)]{700°C} 2CH_3CH_2\cdot \rightarrow CH_3CH_3 + H_2C{=}CH_2 \qquad (4a)$$

$$CH_3CH_2CH_2CH_3 \xrightarrow[(1300°F)]{700°C} CH_3CH_2\cdot CHCH_3 + H\cdot \rightarrow CH_3CH_2HC{=}CH_2 + H_2 \qquad (4b)$$

These reactions are generally nonselective, since the bond dissociation energies ($D°$) do not differ greatly; for simple alkanes, $D°$:C—H = 90–98 kcal, $D°$:C—C = 82–90 kcal. These reactions are the basis of the thermal cracking processes used in petroleum refining. *See* CRACKING.

Pyrolysis of simple aromatic hydrocarbons such as benzene or naphthalene produces aryl radicals, which can attack other hydrocarbon molecules to give bi- and polyaryls, as shown in reaction (5) for

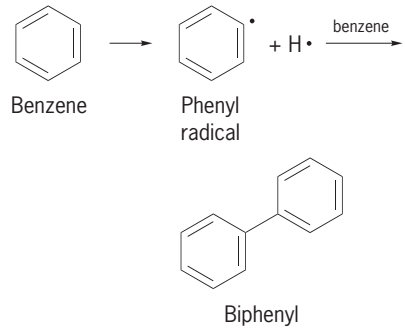

(5)

the formation of biphenyl. With alkyl-substituted and polycyclic hydrocarbons, very complex mixtures of pyrolysis products are obtained by a combination of radical substitutions and disproportionation. Among the end products of these reaction sequences are highly condensed ring systems such as fluoroanthene (**1**), chrysene (**2**), cyclopentapyrene (**3**), and benzopyrene (**4**). These and other polycyclic aromatic

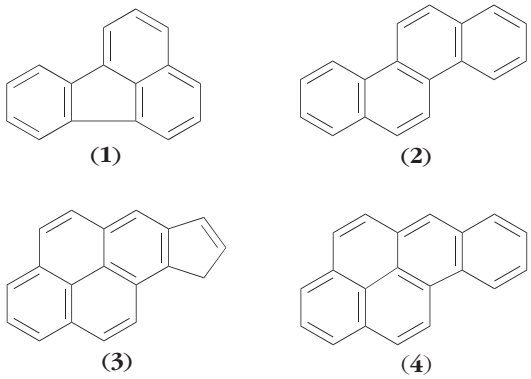

hydrocarbons are present in soot, engine exhausts, and tobacco smoke; they are potent environmental carcinogens. *See* ONCOLOGY.

Elimination. Pyrolytic eliminations can result in formation of a multiple bond by loss of HX from a compound H—C—C—X, where X = any leaving group. A typical example is the pyrolysis of an ester, which is one of the general methods for preparing alkenes. These reactions are commonly carried out with acetate esters at 400–500°C (750–930°F) in a flow reactor and occur by a cyclic transition state [reaction (6)]. Similar elimination of a dithiocarbonate

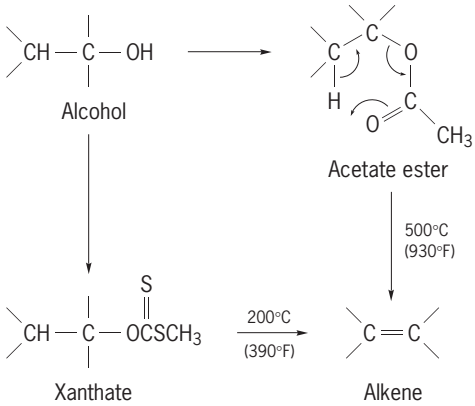

(6)

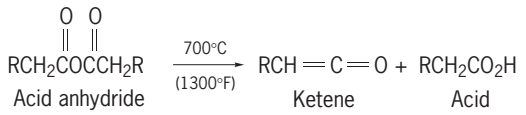

(7)

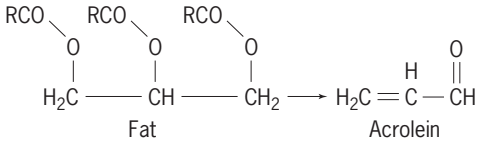

(8)

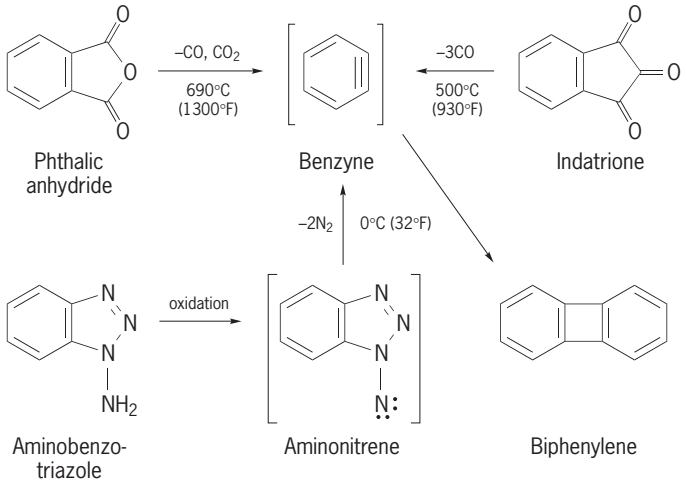

(9)

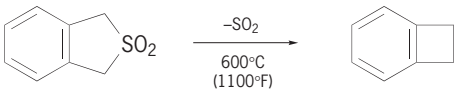

(10)

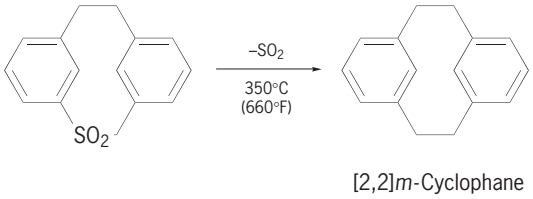

(11)

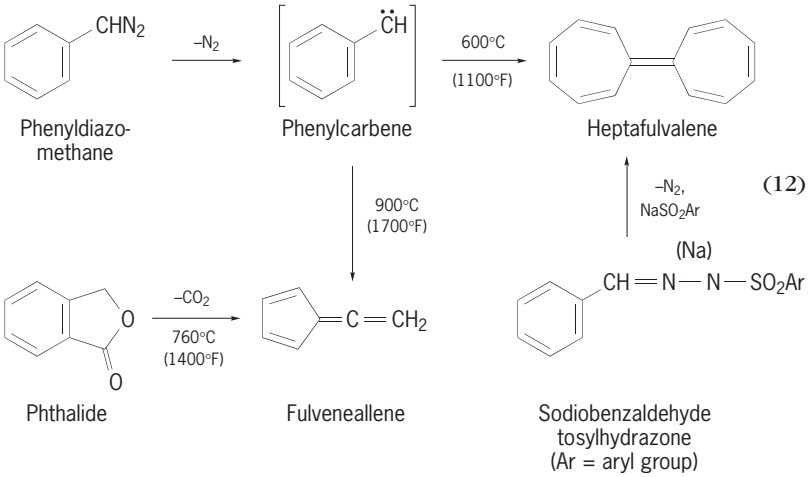

(12)

(xanthate) ester proceeds at 200–300°C (390–570°F). Pyrolytic elimination is particularly useful when acid-catalyzed dehydration of the parent alcohol leads to cationic rearrangement. Another useful application of this process is the production of ketenes from acid anhydrides [reaction (7)]. *See* ACID ANHYDRIDE; ALKENE; ESTER.

The pungent, lacrimatory odor produced by roasting meat at high temperature is due to the formation of acrolein by pyrolysis of fats, which are triesters of glycerol [reaction (8)].

Another type of thermal elimination occurs by loss of a small molecule such as nitrogen (N_2), carbon monoxide (CO), carbon dioxide (CO_2), or sulfur dioxide (SO_2, leading to reactive intermediates such as arynes, diradicals, carbenes, or nitrenes. A few of the pyrolytic routes to benzyne are shown in reactions (9). The conditions vary considerably; thus the nitrene generated from aminobenzotriazole breaks down to benzyne at 0°C (32°F). Benzyne can be trapped by addition reaction or can dimerize to biphenylene. *See* REACTIVE INTERMEDIATES.

Extrusion of SO_2 by pyrolysis of a cyclic sulfone gives a diradical that can then form a new C-C bond. This reaction is an effective preparative method for the synthesis of four-membered rings [reaction (10)] and also cyclophanes [reaction (11)].

When these thermodynamically stable molecules are lost in high-temperature pyrolysis reactions, the energetic and highly reactive fragments are unencumbered by solvent and encounter few intermolecular collisions. Therefore, they can give rise to unusual and otherwise inaccessible products, which may be thought of as a shallow minimum on a potential energy surface.

An example is phenylcarbene, which can be generated by pyrolysis of phenyldiazomethane or other compounds. The carbene can rearrange by either ring contraction or expansion, depending on the source and conditions, to give fulvene or fulvalene products [reactions (12)]. Similarly, nitrenes are formed by pyrolysis of azides [reaction (13)].

Unstable and elusive compounds from flash vacuum pyrolysis can be collected at 20 to 77 K (−253 to −196°C or −423 to −321°F) in an inert matrix for direct spectroscopic observation. By these techniques, cyclobutadiene, triaminoazete, and cyclopentadienone have been characterized [reactions (14)–(16)].

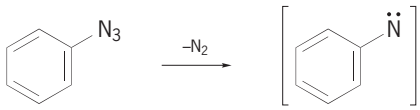

(13)

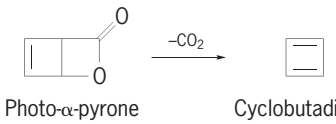

(14)

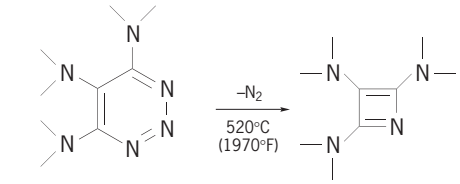

Triamino-1,2,3-triazine Triaminoazete (15)

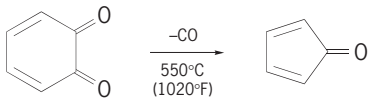

o-Quinone Cyclopenta-dienone (16)

See MATRIX ISOLATION.

Fission of carbon-carbon bonds. A number of pyrolytic reactions involve cleavage of specific C-C bonds in a carbon chain or ring. Fragmentation accompanied by transfer of hydrogen is a general reaction that occurs by a cyclic process. An example is decarboxylation of acids that contain a carbonyl group in the β position, which lose CO_2 on relatively mild heating [reaction (17)]. Acids with a double or triple carbon-to-carbon

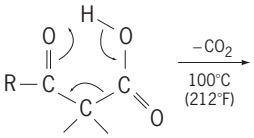

β-Ketoacid

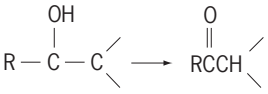

Enol Ketone (17)

bond undergo decarboxylation at 300–400°C (570°–750°F), as in reaction (18). This type of reaction also

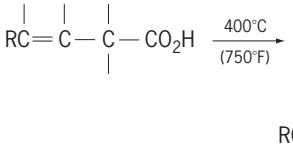

(18)

occurs at higher temperatures with unsaturated alcohols [reaction (19)] and, by transfer of hydrogen

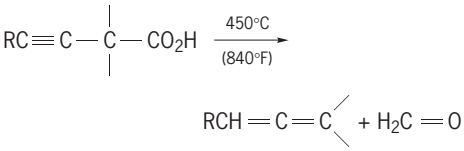

(19)

from a C-H bond, with unsaturated ethers [reaction (20)].

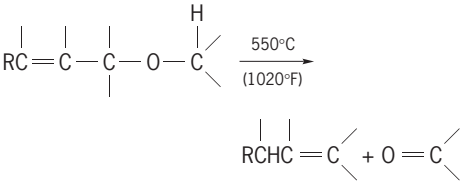

(20)

Cleavage of a ring frequently occurs on pyrolysis. With alicyclic or heterocyclic four-membered rings, cleavage into two fragments is the reverse of 2 + 2 cycloaddition, as illustrated by the cracking of diketene [reaction (21)]. Pyrolysis is an important reaction in

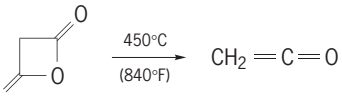

Diketene Ketene (21)

the chemistry of the pine terpenes, as in the conversion of β-pinene to myrcene [reaction (22)].

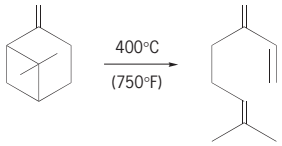

β-Pinene Myrcene (22)

See PINE TERPENE.

Benzocyclobutenes undergo ring opening to *o*-quinone dimethides. By combining this reaction in tandem with formation of the benzocyclobutene and a final Diels-Alder reaction, a versatile one-step synthetic method for the steroid ring system has been developed [reaction (23)].

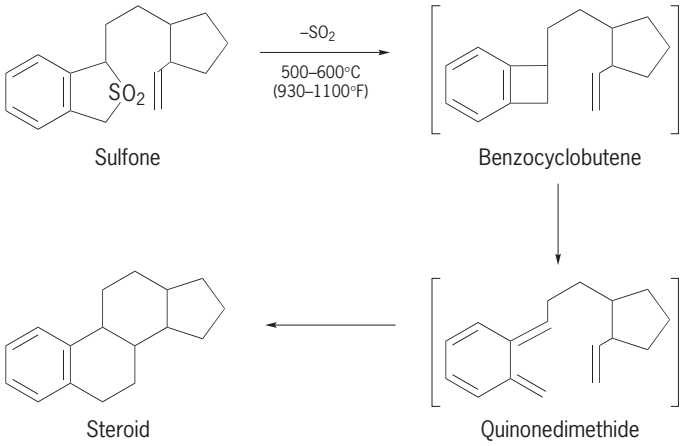

Sulfone Benzocyclobutene

Steroid Quinonedimethide (23)

See DIELS-ALDER REACTION.

Cleavage of a cyclohexene by reverse 4 + 2 cycloaddition is a very general pyrolytic process. Cyclopentadiene dimerizes at room temperature, and this versatile diene is prepared as needed by pyrolysis of the dimer. This ring scission has been employed in the synthesis of the extremely unstable pentalene by flash vacuum pyrolysis of a tetracyclic precursor, in which a cyclopentadiene unit acts as a protective group until the last step [reaction (24).

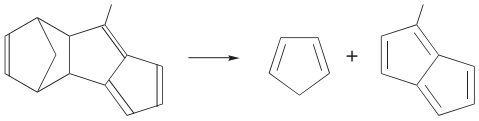

Cyclopenta-diene 1-Methyl-pentalene (24)

Another application of pyrolytic-ring cracking is the reaction of [2,2]-*p*-cyclophane at 600°C (1100°F). The resulting *p*-xylylene is passed in the vapor phase

over surfaces on which a film of the resulting polymer is deposited, forming a tough protective coating [reaction (25)].

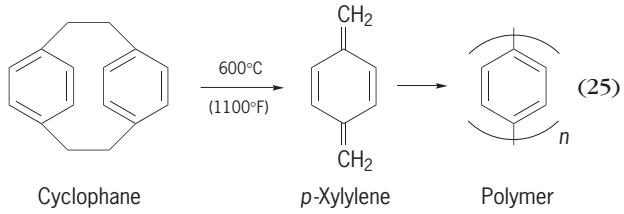

(25)

Many thermal reactions involve isomerization without elimination or fragmentation. These processes can occur by way of intermediates such as diradicals, as in the pyrolysis of pinene [reaction (22)], or they may be concerted pericyclic reactions. An example of the latter is the Claisen-Cope rearrangement of phenyl or vinyl ethers [reaction (26)] and other

(26)

1,5-diene systems. These reactions can be carried out by relatively mild heating, and they are very useful in synthesis. *See* ORGANIC SYNTHESIS; PERICYCLIC REACTION.

Analytical applications. Thermal breakdown of complex structures leads to very complex mixtures of products arising from concurrent dissociation, elimination, and bond fission. Separation of these mixtures provides a characteristic pyrogram that is valuable as an analytical method, particularly for polymeric materials of both biological and synthetic origin. In this application, a small sample is heated on a hot filament or by laser. The pyrolysis products are then analyzed by gas chromatography (Py-Gc), mass spectrometry (Py-Ms), or a combination of both techniques. *See* GAS CHROMATOGRAPHY; MASS SPECTROMETRY; THERMAL ANALYSIS.

Instrumentation with appropriate interfaces and data-handling systems has been developed to permit rapid and sensitive detection of pyrolysis products for a number of applications. One example is the optimization of conditions in petroleum cracking to produce a desired product from varied crude oils. The profile of pyrolysis fragments from a polymer can also be used to detect impurities.

Analytical pyrolysis has also been used in the identification of microorganisms. In this application, a sample of whole cells is pyrolyzed under rigidly controlled conditions, and the entire mixture of volatile products arising from proteins, polynucleotides, cell-wall carbohydrates, and other constituents is analyzed. The Py-Gc or Py-Ms pyrograms obtained from gas chromatography or mass spectrometry contain over 100 identifiable degradation products, ranging from simple alkanes and aromatic hydrocarbons to furans and pyrroles. Many of these compounds are common to all organisms, but computer analysis of the pyrogram can provide a distinctive fingerprint.

James A. Moore

Bibliography. R. F. C. Brown, *Pyrolytic Methods in Organic Chemistry* 1980; W. J. Irwin, *Analytical Pyrolysis*, 1982; V. E. Wiersum, Preparative flash vacuum pyrolysis, *Aldrich Chimica Acta*, 17:31–40, 1984.

Pyrometallurgy

The branch of extractive metallurgy in which processes employing chemical reactions at elevated temperatures are used to extract metals from raw materials, such as ores and concentrates, and to treat recycled scrap metal. The use of heat for pottery making dates as far back as 3500–4000 B.C., and is believed to have gradually led to the first high-temperature furnaces used for smelting copper. Small bowllike furnaces were used around 3500 B.C. for reducing oxidic copper ores mixed with charcoal at high temperatures (1100–1200°C; 2000–2200°F) to metallic copper. The historical metal ages show the development of metallurgy and the beginning of the practice and science of pyrometallurgy. The Chalcolithic Age (roughly 5000–2200 B.C.) led to the Bronze Age (approximately 3000–1300 B.C.), which gave way to the Iron Age (about 1000 B.C. to A.D. 400). The techniques of pyrometallurgy were gradually perfected, especially during the late nineteenth and the twentieth centuries as knowledge, mechanical equipment, and materials of construction improved. Pyrometallurgy continues to be the principal means of metal production.

The simplest reaction in pyrometallurgy for the production of metal M from the compound MX (such as an oxide or sulfide in a concentrate) using a reducing or oxidizing agent R can be written as $MX + R \rightarrow M + RX$. In order for this reaction to proceed, the free-energy term (ΔG°) must be negative. For a practical process, the reaction should proceed rapidly, and M should easily separate from the compound MX. The use of high temperatures prompts higher reaction rates and allows for a sufficiently negative value of ΔG°. During smelting, a liquid slag phase of a lower-density collecting RX forms, which easily separates from the denser liquid metal M. In the case of treating sulfides, R is normally oxygen, so RX would be SO_2, and any accompanying unwanted metals, such as iron, zinc, or nickel, in the feed copper concentrate, for example form oxides such as FeO. In high-temperature metallurgical processing, chemical reactions are rapid, allowing for the processing of large tonnages of metal in a given furnace. Reaction equilibria change as temperature increases so that processes impossible at low temperature become almost spontaneous at higher temperature. In addition, production of the metal as a liquid or a gas facilitates the physical separation of the metal from other materials, which are collected in a liquid slag for disposal. The temperatures used in pyrometallurgy range from 150–200°C (300–400°F) to over 1600°C (2900°F), depending on the specific operation. *See* FREE ENERGY.

For metal production, the pyrometallurgical

operation commences with either a raw material obtained by mining and subsequent mineral and ore processing steps to produce a concentrate, or a recycled material such as separated materials from scrapped automobiles, machinery, or computers. Recycled materials are increasingly important.

Pyrometallurgical preparation processes convert raw materials to forms suitable for future processing, such as roasting of zinc sulfide concentrates to convert the sulfides to oxides for further processing, sintering to oxidize and agglomerate sulfide concentrates to coarser oxides for the blast furnace smelting of lead, or chlorination reactions for certain metals such as titanium. Reduction processes reduce metallic oxides and compounds to metal, as in the iron blast furnace which reduces iron oxide to pig iron, or the lead blast furnace which reduces lead oxide sinter to lead metal. Oxidizing processes oxidize the feed material to an intermediate or a semifinished metal product, such as smelting processes which produce a high-grade copper matte (a copper sulfide rich–iron sulfide material), or converting copper(I) sulfide (Cu_2S) to blister copper (an impure form of copper). Refining processes remove the last of the impurities from a crude metal, as in fractional distillation to remove lead or cadmium from impure zinc, the gaseous reduction (removal of oxygen) of anode copper prior to casting, or copper fire-refining (the refining of molten copper using fluxes such as silica or soda-ash).

The complete production scheme, from raw material (concentrate or recycled scrap) to refined metal, may use pyrometallurgical processes for all steps (steel, lead, tin, zinc), or only the primary extraction processes may be pyrometallurgical with other methods, such as electrorefining, used for final refining (copper or nickel). Alternatively, the primary step may be nonpyrometallurgical (in many cases a leaching step), followed by pyrometallurgical processing (aluminum, magnesium, titanium), or vice versa, such as the production of zinc by first roasting sulfide concentrates, followed by leaching and electrowinning (electrochemical metal extraction from solution). In some cases, certain pyrometallurgical processes are used in predominantly nonpyrometallurgical treatment schemes. *See* ELECTROMETALLURGY; IRON METALLURGY; METALLURGY; PYROMETALLURGY, NONFERROUS.

Phillip J. Mackey

Pyrometer

A temperature-measuring device, originally an instrument that measures temperatures beyond the range of thermometers, but now in addition a device that measures thermal radiation in any temperature range. This article discusses radiation pyrometers. For other temperature-measuring devices *see* BOLOMETER; THERMISTOR; THERMOCOUPLE.

Technique. Figure 1 shows a very simple type of radiation pyrometer. Part of the thermal radiation that is emitted by a hot object is intercepted by a lens and focused onto a thermopile. The resultant heating of the thermopile causes it to generate an electrical signal (proportional to the thermal radiation) which can be displayed on a recorder.

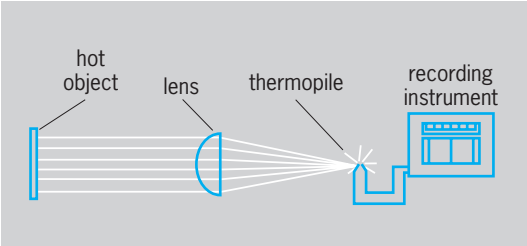

Fig. 1. Elementary radiation pyrometer. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

Unfortunately, the thermal radiation emitted by the object depends not only on its temperature but also on its surface characteristics. The radiation existing inside hot, opaque objects is so-called blackbody radiation, which is a unique function of temperature and wavelength and is the same for all opaque materials. However, such radiation, when it attempts to escape from the object, is partly reflected at the surface. The fraction reflected back into the interior (and hence also the fraction emitted) depends on the type of material, surface roughness, surface films, and microstructure.

The pyrometer readings, then, are ambiguous, since a measured rate of thermal radiation may be due to a temperature of 752°F (400°C) and an internal reflectance of 0.80 (emittance of 0.20) on the one hand, or to a temperature of 932°F (500°C) and an internal reflectance of 0.90 (emittance of 0.10) on the other hand. In order to use the output of the pyrometer as a measure of target temperature, the effect of the surface characteristics must be eliminated from the measurements. *See* BLACKBODY.

This elimination is accomplished as follows. A cavity is formed in an opaque material and the pyrometer is sighted on a small opening extending from the cavity to the surface. The opening has no surface reflection, since the surface has been eliminated. Such a source is called a blackbody source, and is said to have an emittance of 1.00. Any other object, when its surface temperature is measured, has some actual surface reflectance and therefore has an emittance less than 1.00.

By attaching thermocouples to the blackbody source, a curve of pyrometer output voltage versus blackbody temperature can be constructed. Similarly, for a given material (without the cavity) thermocouples can be attached to the surface and the radiation output to the pyrometer output voltage determined at known material temperatures. The emittance factor is taken to be this voltage divided by the voltage output observed for the blackbody at the same temperature. Then, in making measurements on this material remotely and without contact (as on a moving process line), the voltage readings are divided by the proper emittance factor to obtain the true temperature from the blackbody curve. *See* THERMOCOUPLE.

The effects of target size and target distance on the pyrometer readings must also be eliminated. One solution is to design the field of view to increase as the square of the target distance. Thus, as long as the field of view is filled by the target, the increasing target size just compensates for the decreased signal per unit target area as the target distance increases.

Types. Pyrometers can be classified generally into types requiring that the field of view be filled, such as narrow-band and total-radiation pyrometers; and types not requiring that the field of view be filled, such as optical and ratio pyrometers. The latter depend upon making some sort of comparison between two or more signals.

Narrow-band (spectral) pyrometers. These pyrometers (**Fig. 2**) respond to a limited band of wavelengths. (For example, a typical lead sulfide pyrometer will have a filter that limits its spectral response to the 1.8–2.7-micrometer region.) The output may vary as the 10th to 15th power of the temperature. Time constants are of the order of milliseconds (as in lead sulfide) or microseconds (as in silicon cells). Fields of view are of the order of 1-cm diameter at 1 m. The detectors are usually of the photoconductive or photovoltaic type. *See* RADIOMETRY.

Total-radiation pyrometers. The lens, or window and mirror, because of limited spectral transmission, allows only broad-band response, rather than "total" radiation response. Hence the output usually varies as about the 4.5 or 5th power. Detectors are usually of the thermal type (that is, the detector generates an output because of a rise in its temperature), such as a thermopile. Time constants are usually of the order of 1 s or longer. Fields of view of about 2 in. (5 cm) diameter at 40 in. (1 m) are typical (**Fig. 3**).

Optical pyrometer. This instrument should more strictly be called the disappearing-filament pyrometer. In operation, an image of the target is focused in the plane of a wire that can be heated electrically. A rheostat is used to adjust the current through the wire until the wire blends into the image of the target (equal brightness condition), and the temperature is then read from a calibrated dial on the rheostat (**Fig. 4**).

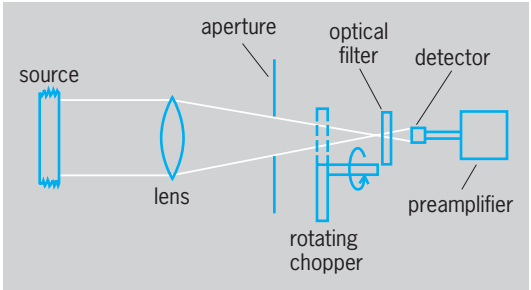

Fig. 2. Narrow-band pyrometer.

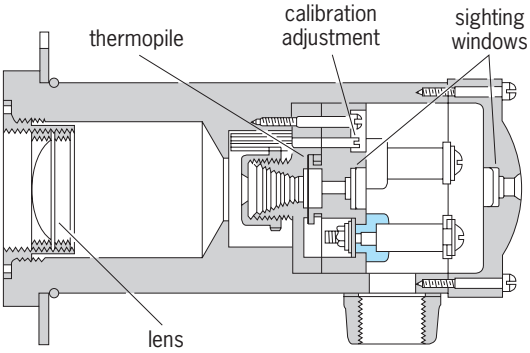

Fig. 3. Radiation pyrometer. (*Honeywell Inc.*)

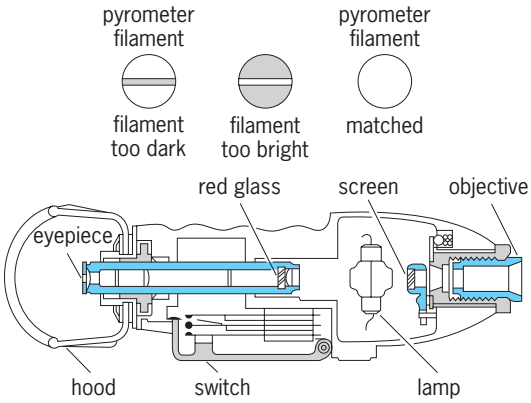

Fig. 4. Optical pyrometer. (*Leeds and Northrup Co.*)

Ratio, or "two-color," pyrometer. This pyrometer makes measurements in two wavelength regions and electronically takes the ratio of these measurements. If the emittance is the same for both wavelengths, the emittance cancels out of the result, and the true temperature of the target is obtained. This so-called gray-body assumption is sufficiently valid in some cases so that the "color temperature" measured by a ratio pyrometer is close to the true temperature.

A given target may be said to have a certain "brightness temperature," "radiation temperature," "color temperature" and, of course, true temperature, depending on how the temperature is measured. These may all differ before the readings are corrected for emittance. The terms luminance temperature and distribution temperature have been adopted to replace brightness temperature and color temperature, respectively.

Radiation pyrometers are particularly useful for (1) temperatures above the practical operating ranges of thermocouples; (2) environments that contaminate and limit the useful life of thermocouples; (3) moving targets; (4) visible targets not easily accessible; (5) targets that would be damaged by contact or insertion; (6) average temperatures of large areas; and (7) rapidly changing temperatures (using fast response detectors).

Errors in temperatures measured by radiation pyrometers can be caused by (1) uncertainty in the emittance value of the target; (2) dust, steam, or smoke in the radiation path; (3) dirt on lens or window; (4) water or dirt on the target surface; and (5) reflection of spurious radiation from the target. *See* TEMPERATURE MEASUREMENT; THERMOMETER.

Thomas P. Murray

Bibliography. American Institute of Physics, *Temperature: Its Measurement and Control in Science and Industry*, vol. 5, 1982; D. P. Dewitt, *Theory and Practice of Radiation Thermometry*, 1988;

T. D. McGee, *Principles and Methods of Temperature Measuremen*, 1988; E. C. Magison, *Temperature Measurement in Industry*, 1990; L. Michalski, K. Eckersdorf, and J. McGhee, *Temperature Measurement*, 2d ed., 2001.

Pyromorphite

A mineral series in the apatite group, or in the larger grouping of phosphate, arsenate, and vanadate-type minerals. In this series lead (Pb) substitutes for calcium (Ca) of the apatite formula $Ca_5(PO_4)_3(F,OH,Cl)$, and little fluorine (F) or hydroxide (OH) is present. Substitution between phosphorus (P) and arsenic (As) gives rise to a complete mineral series with pyromorphite, $Pb_5(PO_4)_3Cl$, and mimetite, $Pb_5(AsO_4)Cl$, as pure end members. Substitution between As and vanadium (V) produces another series between mimetite and vanadinite, $Pb_5(VO_4)_3Cl$.

Pyromorphite generally is a chloride-phosphate-arsenate of lead, $Pb_5(PO_4,AsO_4)_3Cl$. The half of the series in which P > As is called pyromorphite; the remainder of the series with As > P is termed mimetite. Vanadinite is a chloride-vanadate of lead. P and As may replace V in amounts up to P:V = 1:47 and As:V = 1:1.

The pyromorphite series crystallizes in the hexagonal system (see **illus.**). Crystals are prismatic;

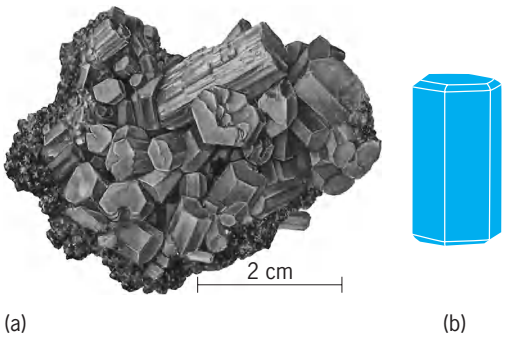

Pyromorphite. (*a*) Hexagonal crystals, Nassau (*specimen from Department of Geology, Bryn Mawr College*). (*b*) Crystal habit (*after C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*)

pyramidal faces may be present. Other forms are granular, globular, and botryoidal. Pyromorphite colors range through green, yellow, and brown; vanadinite occurs in shades of yellow, brown, and red.

Pyromorphites are widely distributed as secondary minerals in oxidized lead deposits. Pyromorphite is a minor ore of lead; vanadinite is a source of vanadium and minor ore of lead. *See* APATITE; LEAD; VANADIUM. Wayne R. Lowell

Pyrophyllite

A hydrated aluminum silicate with composition $Al_2Si_4O_{10}(OH)_2$. The mineral is commonly white, grayish, greenish, or brownish, with a pearly to waxy appearance and greasy feel. It occurs as compact masses, as radiating aggregates (see **illus.**),

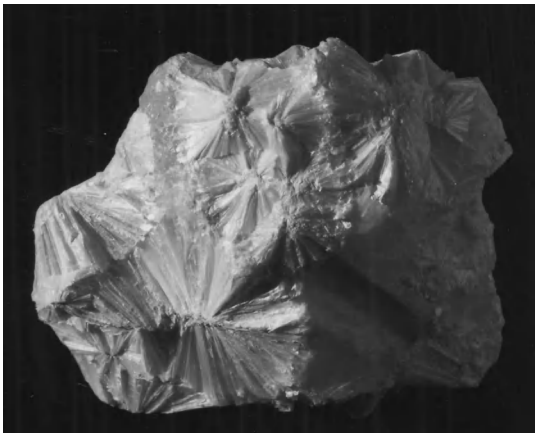

Specimen of pyrophyllite. (*Pennsylvania State University*)

and as foliated masses. Pyrophyllite belongs to the layer silicate (phyllosilicate) group of minerals. The atomic (crystal) structure is similar to that of talc, $Mg_3Si_4O_{10}(OH)_2$, with two of the Mg positions occupied by Al and the third position vacant. The layers are electrically neutral and are held together by weak secondary bonds and long-range ionic (electrostatic) forces. The mineral is soft (hardness 1–$1^1/_2$ on Mohs scale) and has easy cleavage parallel to the structural layers. The mineral is highly stable to acids. Thermal decomposition begins around 750–800°C (1400–1500°F), and recrystallization to mullite and cristobalite occurs at 1200°C (2200°F) and higher temperatures.

Pyrophyllite is used principally for refractory materials, that is, those for use at high temperatures, and in other ceramic applications. In the United States it has been used mainly in making refractories, ceramic materials, and as a carrier for insecticides. The main sources for pyrophyllite in the United States are in North Carolina. An unusual form from the Transvaal, called African wonderstone, is fine-grained and dark gray to black and has the property that it can be shaped mechanically and drilled into complex shapes; after firing it becomes a useful refractory material for special applications. *See* CERAMICS; REFRACTORY; SILICATE MINERALS. George W. Brindley

Bibliography. Bureau of Mines, U.S. Department of the Interior, *Mineral Commodity Summaries*, annually; Bureau of Mines, U.S. Department of the Interior, *Minerals Yearbook*; W. A. Deer, R. A. Howie, and J. Zussman, *Rock-Forming Minerals*, vol. 3, 1962.

Pyrotechnics

Mixtures of substances that produce noise, light, heat, smoke, or motion when ignited. They are used in matches, incendiaries, and other igniters; in fireworks and flares; in fuses and other initiators for primary explosives; in delay trains; for powering mechanical devices; and for dispersing materials such as insecticides.

Greek fire, a pyrotechnic whose composition was a military secret and was lost, was used by the Greeks

in naval warfare in A.D. 673. But black powder, an intimate mixture of potassium nitrate, charcoal, and sulfur, was perhaps the earliest pyrotechnic and remains the most important one. It was discovered prior to A.D. 1000 in China during the Sung dynasty, where it was used in rockets and fireworks. Its introduction to Europe prior to 1242, when Roger Bacon recorded its formula, eventually revolutionized warfare. Its first use in mines was reported in Hungary in 1677. Its use in blasting rock was also revolutionary, enabling rapid fragmentation of ore and other rock with a great reduction in manual labor.

The making of fireworks developed over the centuries as an art whose secrets were passed down from generation to generation. Pyrotechnics still contain much art even in such highly technical applications as the *Apollo* spacecraft, which contained over 200 pyrotechnic devices, including lightweight, remotely operated electrical switches and mechanical releases required to function only once.

A pyrotechnic mixture contains a fuel and an oxidizer, usually another ingredient to give a special effect, and often a binder. The oxidizer is usually a nitrate, perchlorate, chlorate, or peroxide of potassium, barium, or strontium. Fuels may be sulfur, charcoal, boron, magnesium, aluminum, titanium, or antimony sulfide. Examples of binders that are also fuels are dextrin and natural polysaccharides such as red gum. Salts of strontium, calcium, barium, copper, and sodium and powdered magnesium metal when combined with a fuel and oxidizer can give the special effects of scarlet, brick red, green, blue-green, yellow and white flames, respectively; iron and aluminum granules can give gold and white sparks, respectively; organic dyes and a coolant such as sodium bicarbonate can give colored smokes; potassium benzoate can give unstable burning which results in a whistling sound, and powdered titanium metal can give very rapid burning which results in a booming sound. Mixtures such as boron with red lead have no gaseous reaction products and are useful for confined delay trains where gasesous products would be disruptive.

Many pyrotechnic mixtures are easily ignited by impact, friction, flame, sparks, or static electricity. Even those that burn quietly in small quantities can explode violently when ignited under confinement or in larger quantities. Their manufacture is done with elaborate safety precautions owing to the hazards involved. D. Linn Coursen

Bibliography. R. T. Barbour, *Pyrotechnics in Industry*, 1981; J. A. Conkling, Chemistry of fireworks, *Chem. Eng. News*, 59(26):24-32, 1981; R. Lancaster, *Fireworks Principles and Practice*, 3d ed., 1998; C. F. Ruggieri, *Principles of Pyrotechnics*, 1994.

Pyrotheria

An extinct order of primitive, mastodonlike, herbivorous, hoofed mammals restricted to the Eocene and Oligocene deposits of South America. There is only one family (Pyrotheriidae) in the order and four genera (*Pyrotherium*, *Propyrotherium*, *Carlozittelia*, and *Griphodon*) in the family.

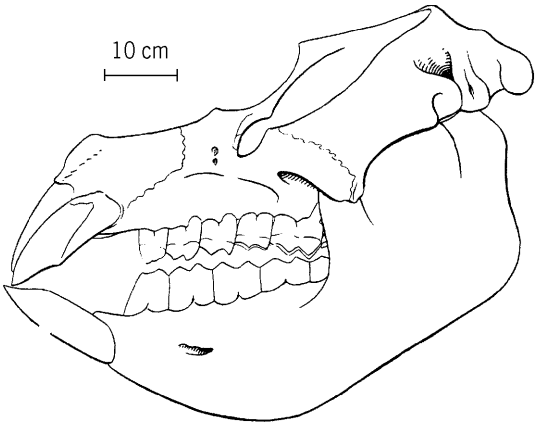

Skull and jaw of *Pyrotherium sorondi*, an early Oligocene pyrothere from South America. 10 cm = 25.4 in.

The characters of this group superficially resembling those in early proboscideans are nasal openings over orbits indicating the presence of a trunk, strong neck musculature, and six upper and four lower bilophodont cheek teeth. They also had four upper and two lower chisellike incisor tusks (see **illus**.). They are distantly related to members of the superorder Paenungulata, including Proboscidea, Xenungulata, and others. *See* PROBOSCIDEA; XENUNGULATA. Gideon T. James

Bibliography. M. J. Benton, *Vertebrate Paleontology*, 1991; R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988.

Pyroxene

A large, geologically significant group of dark, rock-forming silicate minerals. Pyroxene is found in abundance in a wide variety of igneous and metamorphic rocks. Because of their structural complexity and their diversity of chemical composition and geologic occurrence, these minerals have been intensively studied by using a wide variety of modern analytical techniques. Knowledge of pyroxene compositions, crystal structures, phase relations, and detailed microstructures provides important information about the origin and thermal history of rocks in which they occur.

Crystal structures. All silicate minerals in the Earth's crust contain silicon cations (Si^{4+}) bonded to four nearest-neighbor oxygen anions (O^{2-}) arranged at the corners of a tetrahedron centered on the Si^{4+} cation. Silicate tetrahedra may polymerize by means of one or more oxygen anions bonding to two Si^{4+} cations. If each tetrahedron shares two of its four anions with adjacent tetrahedra, single silicate chains are formed. The crystallographic repeats of these single chains may range from two to nine tetrahedra; the pyroxene group includes only those with a repeat of two. The two unshared oxygen anions of each silicate tetrahedron are bonded to divalent (for example, Mg^{2+}, Fe^{2+}) or trivalent (for example, Al^{3+}, Fe^{3+})

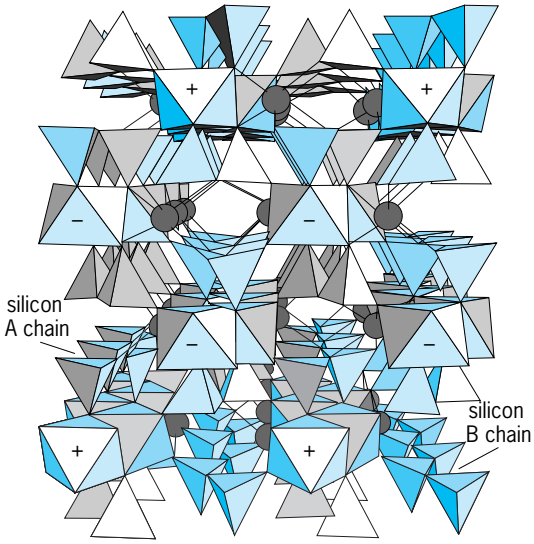

Fig. 1. Perspective view of the orthopyroxene structure. Crystallographic orientation: *a* axis vertical, *b* axis horizontal, *c* axis parallel to the direction of tetrahedral silicate chains. Pyroxene I beam consists of a strip of MI octahedra sandwiched between a nearly straight silicon A chain and a kinked silicon B chain. Relative I-beam orientation is indicated with +1 or −1 on M1 octahedral faces. Shaded circles are M2 sites, showing bonds to nearest-neighbor oxygens.

cations occupying M1 sites, forming octahedral coordination polyhedra sharing edges in chains parallel to the tetrahedral chains (M1 octahedra; **Fig. 1**). The oxygen anions also bond to divalent cations (Ca^{2+}, Mg^{2+}, Fe^{2+}) or sodium (Na^{+}) cations that occupy M2 sites. The composite building block of all pyroxene structures, called an I beam, consists of a chain of edge-sharing M1 octahedra and two attached silicate tetrahedral chains, one above and one below the octahedral chain (Fig. 1). The M2 cations located along the edges of the I beams cross-link them to form the three-dimensional structure. The I beams can be stacked in several ways, leading to pyroxene structures with different space groups and unit-cell sizes. *See* COORDINATION CHEMISTRY.

Crystal chemistry. The general chemical formula for pyroxenes is $M2M1T_2O_6$, where T represents the tetrahedrally coordinated sites, occupied primarily by silicon cations (Si^{4+}). Names of specific end-member pyroxenes are assigned based on composition and structure type. Extensive solid solution exists between many end-member compositions (see **table** and **Fig. 2**). *See* SOLID SOLUTION.

Those pyroxenes containing primarily calcium (Ca^{2+}) or sodium (Na^{+}) cations in the M2 site are monoclinic. Some omphacites (Fig. 2*b*) exhibit ordering of Ca^{2+} and Na^{+} on M2 sites, along with concomitant ordering of magnesium (Mg^{2+}) and aluminum (Al^{3+}) cations on M1 sites. Pyroxenes containing primarily Mg^{2+} or iron(II) (Fe^{2+}) cations in the M2 site are orthorhombic at low temperatures, but they may transform to monoclinic at higher temperature. These clinopyroxene phases exhibit the usual monoclinic I-beam stacking sequence but have two crystallographically distinct silicate chains (A and B; Fig. 1) that differ in their degree of straightness, similar to the orthopyroxenes. They will transform rapidly and reversibly at higher temperatures to a symmetry where all silicate chains are equivalent. The extent of polymorphism a particular pyroxene may undergo thus depends on the content of the M2 sites: those containing significant amounts of Ca^{2+} or Na^{+} cations are always monoclinic, whereas those containing little or no Ca^{2+} or Na^{+} may be orthorhombic or monoclinic; and the particular monoclinic symmetry depends on temperature. *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Important end-member pyroxene minerals

Name	Space group	Formula
Clinoentstatite	$P2_1/c$	$Mg_2Si_2O_6$
High clinoenstatite	$C2/c$	
Enstatite	$Pbca$	
Protoenstatite	$Pbcn$	
Clinoferrosilite	$P2_1/c$	$Fe_2^{2+}Si_2O_6$
Ferrosilite	$Pbca$	
Diopside	$C2/c$	$CaMgSi_2O_6$
Headenbergite	$C2/c$	$CaFe^{2+}Si_2O_6$
Jadeite	$C2/c$	$NaAlSi_2O_6$
Aegirine	$C2/c$	$NaFe^{3+}Si_2O_6$
Ca-Tschermak's pyroxene	$C2/c$	$CaAl(SiAl)O_6$
Johannsenite	$C2/c$	$CaMnSi_2O_6$
Spodumene	$C2/c$	$LiAlSi_2O_6$

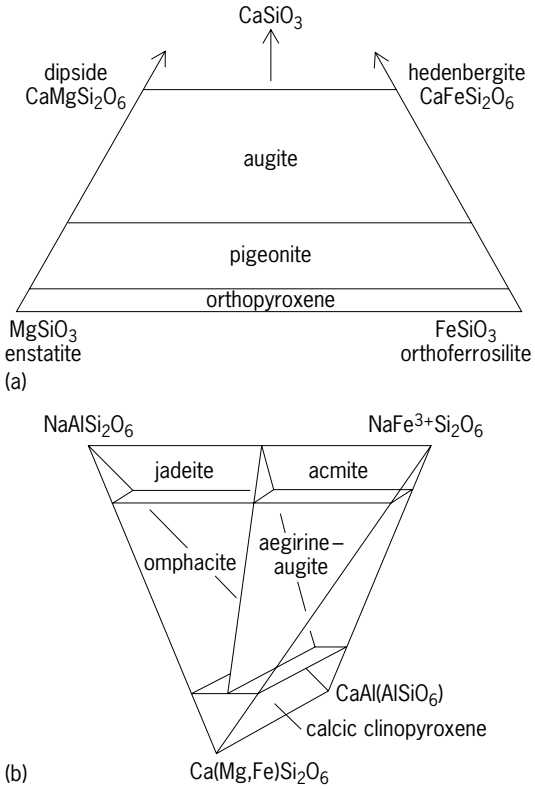

Fig. 2. Nomenclature for common pyroxene solid solutions. (*a*) Rock-forming pyroxene quadrilateral with end members enstatite-ferrosilite-diopside-hedenbergite. (*b*) A clinopyroxene composition tetrahedron, with end members diopside + hedenbergite, jadeite, aegirine, and Ca-Tschermak's pyroxene. (*After W. D. Nesse, Introduction to Optical Mineralogy, 2d ed., Oxford University Press, 1991*)

Within the rock-forming pyroxene quadrilateral (Fig. 2*a*), the $Mg^{2+} \leftrightarrow Fe^{2+}$ solid solution is complete, except that pure ferrosilite ($Fe_2Si_2O_6$) does not occur naturally but can be synthesized at high pressure. However, there is a strong tendency for the Ca-poor clino- and orthopyroxenes and the Ca-rich clinopyroxenes to be immiscible in each other. This immiscibility decreases with increasing temperature, and is clearly related to the differing M2 coordination requirements of Ca^{2+} versus Mg^{2+} and Fe^{2+}. As a consequence, most igneous pyroxenes crystallizing at high temperatures exhibit layers precipitated from the solid crystalline phase (exsolution lamellae) of augite in a pigeonite or orthopyroxene matrix, or vice versa. Pigeonite may transform to orthopyroxene during cooling, further complicating analysis of textural relationships. The orientation of exsolution lamellae of orthopyroxene in clinopyroxene (or the reverse) is parallel to their common (100) planes, whereas exsolution between two clinopyroxenes (pigeonite-augite) is oriented along (001). Detailed study of these exsolution phenomena can provide substantial insight into rock crystallization history. *See* PIGEONITE; SOLID SOLUTION.

The distribution of Mg^{2+} and Fe^{2+} cations between M1 and M2 in the enstatite-ferrosilite series varies with temperature. At low temperatures Fe^{2+} preferentially occupies M2 sites, but as the temperature increases the preference is weakened and disorder prevails. Examination of the distribution of Fe^{2+} over M1 and M2 in these phases using x-ray diffraction techniques or Mössbauer spectroscopy can provide further information about cooling history. *See* MÖSSBAUER EFFECT; SPECTROSCOPY; X-RAY DIFFRACTION.

In addition to the more important pyroxene compositions (see table and Fig. 2), pyroxenes with a significant amount of manganese cations (Mn^{2+}) in M1 or M2 or both are known. Zinc cations (Zn^{2+}) and some titanium cations (either Ti^{3+} or Ti^{4+}) can be accommodated in M1, as can chromium (Cr^{3+}) or scandium (Sc^{3+}) substituting for iron(III) cations (Fe^{3+}) in sodic pyroxenes.

Physical properties. Common pyroxenes have specific gravity ranging from about 3.2 (enstatite, diopside) to 4.0 (ferrosilite). Hardnesses on the Mohs scale range from 5 to 6. Iron-free pyroxenes may be colorless (enstatite, diopside, jadeite); as iron content increases, colors range from light green or yellow through dark green or greenish brown, to brown, greenish black, or black (orthopyroxene, pigeonite, augite, hedenbergite, aegirine). Spodumene may be colorless, yellowish emerald green (hiddenite), or lilac pink (kunzite). *See* SPODUMENE.

Occurrence. Pyroxenes in the rock-forming quadrilateral (Fig. 2*a*) are essential constituents of ferromagnesian igneous rocks such as gabbros and their extrusive equivalents, basalts, as well as most peridotites. These rocks generally contain either augite or pigeonite or orthopyroxene, or mixtures of augite with pigeonite or orthopyroxene, or occasionally all three. Complex exsolution relationships may cause all three phases to be present in the same mineral grain. Pyroxenes may also be present as the dark constituents of more silicic diorites and andesites. Aegirine or aegirine-augite is frequently found in alkaline igneous rocks such as nepheline syenites and some alkali granites. *See* ANDESITE; BASALT; DIORITE; GABBRO; GRANITE; IGNEOUS ROCKS; PERIDOTITE.

Pyroxenes, especially those of the diopside-hedenbergite series, are found in medium- to high-grade metamorphic rocks of the amphibolite and granulite facies. In calc-silicate rocks formed by regional metamorphism of silty carbonate rocks and in skarns formed by contact metamorphism of carbonate rocks, these pyroxenes can form through reaction (1).

$$\underset{\text{Dolomite}}{CaMg(CO_3)_2} + \underset{\text{Quartz}}{2SiO_2} \rightarrow \underset{\text{Diopside}}{CaMgSi_2O_6} + \underset{\text{Carbon dioxide}}{2CO_2} \quad (1)$$

See METAMORPHISM.

Jadeite is found in glaucophane schists, where under moderate pressure it forms from the breakdown of albite feldspar according to reaction (2).

$$\underset{\text{Albite}}{NaAlSi_3O_8} \rightarrow \underset{\text{Jadeite}}{NaAlSi_2O_6} + \underset{\text{Quartz}}{SiO_2} \quad (2)$$

Gem-quality jadeite is one of two materials used as jade in jewelry and artifacts (the other is nephrite, an amphibole). The closely related omphacite is an essential constituent of eclogites (high-grade metamorphosed basalts and gabbros) where it occurs with pyrope garnet. Manganese (Mn)-pyroxenes are restricted to skarn-type metamorphosed manganese deposits. Spodumene is found exclusively in lithium-bearing granite pegmatites, where it can occur in very large crystals, often mined as a source of lithium. Clinoenstatite is rare in terrestrial environments but is a common constituent of stony meteorites. *See* AMPHIBOLITE; ECLOGITE; FACIES (GEOLOGY); FELDSPAR; GLAUCOPHANE; GRANULITE; JADE; METAMORPHIC ROCKS; OMPHACITE; PEGMATITE; SCHIST; SKARN.

The peridotites found in the Earth's upper mantle contain Mg-rich, Ca-poor pyroxenes, in addition to olivine and other minor minerals. At successively greater depths in the mantle, these Mg-rich pyroxenes will transform sequentially to spinel (Mg_2SiO_4) plus stishovite (SiO_2), an ilmenite structure, or a garnet structure, depending on temperature; and finally at depths of around 360–420 mi (600–700 km) to an $MgSiO_3$ perovskite structure thought to be the major stable constituent of the lower mantle to depths of around 1700 mi (2900 km) at the core-mantle boundary. *See* ASTHENOSPHERE; EARTH INTERIOR; GARNET; ILMENITE; OLIVINE; PEROVSKITE; SILICATE MINERALS; SPINEL; STISHOVITE.

Charles W. Burnham

Bibliography. W. A. Deer, R. A. Howie, and J. Zussman, *An Introduction to the Rock-Forming Minerals*, 2d ed., 1996; N. Morimoto et al., Nomenclature of pyroxenes, *Amer. Mineralog.*, 73:1123–1133, 1988; C. T. Prewitt (ed.), Pyroxenes, *Reviews in Mineralogy*, vol. 7, 1980.

Pyroxenite

A heavy, dark-colored, phaneritic (visibly crystalline) igneous rock composed largely of pyroxene with smaller amounts of olivine or hornblende. Pyroxenite composed largely of orthopyroxene occurs with anorthosite and peridotite in large, banded gabbro bodies. These rocks are formed by crystallization of gabbroic magma (rock melt). Some of these pyroxenite masses are rich sources of chromium. Certain pyroxenites composed largely of clinopyroxene are also of magmatic origin, but many probably represent products of reaction between magma and limestone. Other pyroxene-rich rocks have formed through the processes of metamorphism and metasomatism. *See* GABBRO; IGNEOUS ROCKS; PERIDOTITE; PYROXENE.

Carleton A. Chapman

Pyroxenoid

A group of silicate minerals whose physical properties resemble those of pyroxenes. In contrast with the two-tetrahedra periodicity of pyroxene single silicate chains, the pyroxenoid crystal structures contain single chains of $(SiO_4)^{4-}$ silicate tetrahedra having repeat periodicities ranging from three to nine (see **illus.**). The tetrahedron is a widely used geometric representation for the basic building block of most silicate minerals, in which all silicon cations (Si^{4+}) are bonded to four oxygen anions arranged as if they were at the corners of a tetrahedron. In pyroxenoids, as in other single-chain silicates, two of the four oxygen anions in each tetrahedron are shared between two Si^{4+} cations to form the single chains, and the other two oxygen anions of each tetrahedron are bonded to divalent cations, such as calcium (Ca^{2+}), iron (Fe^{2+}), or manganese (Mn^{2+}). These divalent cations bond to six (or sometimes seven or eight) oxygen anions, forming octahedral (or irregular seven- or eight-cornered) coordination polyhedra. *See* PYROXENE; SILICATE MINERALS.

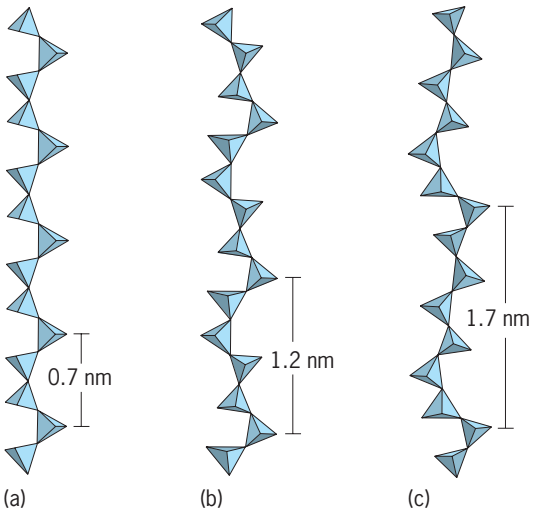

Tetrahedral silicate chains in three pyroxenoid structures: (*a*) wollastonite with three-tetrahedra periodicity, (*b*) rhodonite with five-tetrahedra periodicity, and (*c*) pyroxmangite-pyroxferroite with seven-tetrahedra periodicity.

Pyroxenoid minerals

Name	Formula
Anhydrous	
Wollastonite	$CaSiO_3$
Bustamite-ferrobustamite	~ $Ca(Mn,Fe^{2+})Si_2O_6$
Rhodonite	~ $CaMn_4Si_5O_{15}$
Pyroxmangite-pyroxferroite	~ $Ca(Mn,Fe^{2+})_6Si_7O_{21}$
Ferrosilite III	$FeSiO_3$
Hydrous	
Pectolite-serandite	$(Ca,Mn)_2NaHSi_3O_9$
Babingtonite	$Ca_2(Fe,Mn,Mg)Fe^{3+}HSi_5O_{15}$
Nambulite-marsturite	$(Li,Na)(Mn,Ca)_5HSi_5O_{15}$
Santaclaraite	$CaMn_4HSi_5O_{15}(OH)\cdot H_2O$

The pyroxenoid structures all have composite structural units consisting of strips of octahedra (or larger polyhedra) two or more units wide formed by sharing of polyhedral edges, to which the silicate tetrahedral chains are attached on both top and bottom. The repeat periodicity of the octahedral strips is the same as that of the silicate chains to which they are attached. These composite units are cross-linked to form the three-dimensional crystal structures. Although various crystallographic settings have been employed in the past, there has evolved general agreement that the most useful orientation selects the unit-cell *c* axis to be parallel to the silicate chain and octahedral strip direction, and sets the (100) plane parallel to the plane of the octahedral strips, which also is a plane of approximately close-packed oxygen atoms; this setting is analogous to that used for pyroxenes. The pyroxenoid minerals are all triclinic, with either $C\bar{1}$ or $I\bar{1}$ space-group symmetry depending on the stacking of the composite units, and *c*-axis lengths ranging from about 0.71 nanometer for three-repeat silicate tetrahedral chains to about 2.3 nm for nine-repeat chains. *See* CRYSTAL STRUCTURE.

Types. There are two series of pyroxenoid minerals, one anhydrous and one hydrous (see **table**). The anhydrous pyroxenoids are significantly more abundant. A general formula for anhydrous pyroxenoids is $(Ca, Mn, Fe^{2+})SiO_3$. Silicate-chain repeat length is inversely proportional to mean divalent cation size, for example, wollastonite with three-repeat chains, and ferrosilite III with nine-repeat chains. Wollastonite ($CaSiO_3$) has two other polymorphs: parawollastonite is a rare pyroxenoid having a different stacking sequence of the composite tetrahedral-octahedral units; and pseudowollastonite is a ring silicate and not, therefore, a pyroxenoid. Both bustamite ($CaMnSi_2O_6$) and ferrobustamite ($CaFeSi_2O_6$) are polymorphous with pyroxenes: johannsenite and hedenbergite respectively. The $MnSiO_3$ composition can exist as rhodonite, pyroxmangite, pyroxene, or garnet, depending on temperature and pressure. Ferrosilite III has been synthesized in the laboratory as a pyroxenoid polymorph of the pyroxenes clinoferrosilite and orthoferrosilite; it has not been found occurring naturally. The hydrogen in hydrous

pyroxenoids is hydrogen-bonded between two oxygen atoms; additional hydrogen in santaclaraite is bound as hydroxyl (OH) and as a water molecule (H_2O).

Occurrence. Wollastonite is the most abundant pyroxenoid, occurring primarily as a consequence of high-grade metamorphism of siliceous limestone [$CaCO_3$ with silicon dioxide (SiO_2)]. Under high-temperature conditions, it can form from the decarbonation reaction below. Wollastonite is mined

$$CaCO_3 + SiO_2 \longrightarrow CaSiO_3 + \underset{\text{Carbon dioxide}}{CO_2}$$

in New York and California and, to a lesser extent, in Mexico, Finland, Kenya, South-West Africa, and central Asia. It is used commercially in the ceramic industry, as a filler in high-quality light-colored exterior paints, as a strengthening additive in some plastics, and in ceramic-bonded abrasives. Rhodonite and other manganese-rich pyroxenoids are found chiefly in association with some metamorphosed manganese deposits; particularly well-formed rhodonite crystals are sold as semiprecious gems, with value related primarily to their pink color. Pyroxferroite was first found as a minor phase in lunar basalts; it has also been found in pegmatites in Japan. *See* RHODONITE; WOLLASTONITE.

Charles W. Burnham

Bibliography. W. A. Deer et al., *Rock-Forming Minerals,* vol. 2A, 2d ed., 1978; W. A. Deer, R. A. Howie, and J. Zussman, *An Introduction to Rock-Forming Minerals,* 2d ed., 1992.

Pyrrole

One of a group of organic compounds containing a doubly unsaturated five-membered ring in which nitrogen occupies one of the ring positions. Pyrrole (**1**) is a representative compound. The pyrrole system is found in the green leaf pigment, chlorophyll, in the red blood pigment, hemoglobin, and in the blue dye, indigo. Interest in these colored bodies has been largely responsible for the intensive study of pyrroles. Tetrahydropyrrole, or pyrrolidine (**2**), is part of the structures of two protein amino acids, proline (**3**) and hydroxyproline, and of hygrine (**4**),

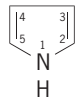

(**1**)

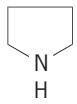

(**2**)

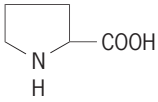

(**3**)

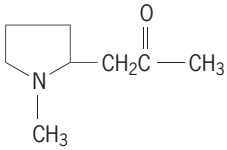

(**4**)

which is an alkaloid from Peruvian coca. *See* HETEROCYCLIC COMPOUNDS; INDOLE; PORPHYRIN.

Properties. Pyrrole (**1**) is a liquid, bp 130°C (266°F), n^{20}_D 1.5085 (1.5098), and density (20/4) 0.948 (0.969), that darkens and resinifies on standing in air, and that polymerizes quickly when treated with mineral acid. Polyalkyl pyrroles are not so sensitive and negatively substituted pyrroles are even less so. Pyrrole is a planar, aromatic compound, with an experimental resonance energy of 22–27 kcal/mole. Familiar substitution processes, such as halogenation, nitration, sulfonation, and acylation, can be realized. Substitution generally occurs more readily than in the corresponding benzene analog. The entering group favors the 2 or 5 position. Pyrrole, by virtue of its heterocyclic nitrogen, is very weakly basic and is comparable in this respect to urea or to semicarbazide. The hydrogen at the 1 position is removable as a proton, and accordingly, pyrrole is also an acid, although a weak one.

Pyrroles are not as resistant to oxidation as the analogous benzene compounds. Controlled chromic acid oxidation converts pyrroles with or without groups in the α positions to the corresponding maleimides. In this way for example, ethyl-2,5-dimethylpyrrole-3-carboxylate (**5**) is oxidized to a maleimide derivative [**6**; reaction (1)]. Conditions are

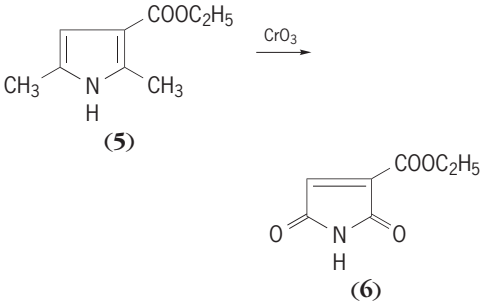

(**5**) (**6**) (1)

available for useful reductions of groups attached to the nucleus. Zinc and hydrochloric acid can reduce the pyrrole ring to dihydropyrrole (pyrroline). Raney nickel under rigorous conditions, or platinum or palladium under milder conditions, serve as catalysts in the hydrogenation of pyrroles to pyrrolidines.

Preparation. The Knorr synthesis [reaction (2)],

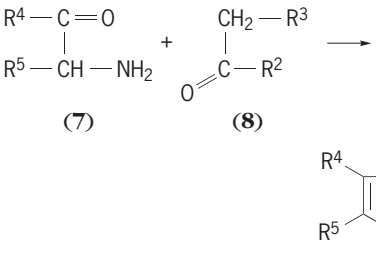

(**7**) (**8**) (2)

probably the most versatile pyrrole synthesis, combines an α-aminoketone (**7**) with an α-methylene carbonyl compound (**8**). The condensation is carried out either in glacial acetic acid or in aqueous alkali. The R groups may be varied widely; however, best results are obtained when R^3 of the carbonyl compounds (**8**) is an activating group, for example, carbethoxy. Another useful general synthesis,

the Paal-Knorr synthesis [reaction (3)], converts a

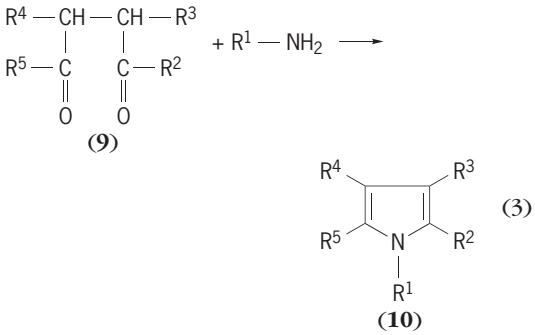

(3)

1,4-dicarbonyl compound (9) by cyclization with ammonia (R^1 = H) or with a primary amine to a pyrrole (10). Pyrrole itself is obtained by pyrolysis of ammonium mucate (saccharate). Pyrolysis of primary amine salts of mucic acid gives 1-substituted pyrroles.

Derivatives. Chloro-, bromo-, and iodopyrroles have been prepared, with bromopyrroles comprising the largest group. Halopyrroles carrying an electronegative group, such as carbethoxy, are more stable than the same compounds without the group. 2,3,4,5-Tetraiodopyrrole, formed by iodination of pyrrole or of tetrachloromercuripyrrole, is useful as an antiseptic in the same areas as iodoform. Pyrroles with ethylmagnesium bromide react as active hydrogen compounds to give Grignard derivatives. Ethyl formate reacts with such pyrryl Grignard reagents to form pyrrole aldehydes; chloroformic ester reacts to give pyrrolecarboxylic esters. Pyrrole aldehydes can also be formed by formylation of a vacant pyrrole position with hydrocyanic acid and hydrogen chloride. Acylpyrroles, that is, pyrryl ketones, are prepared satisfactorily and under mild conditions by Friedel-Crafts acylation of the pyrrole nucleus.

Pyrrolidine (**2**), bp 87–88°C (189–190°F), can be prepared by catalytic hydrogenation of pyrrole or by ring-closure reactions [reaction (4)]. Either

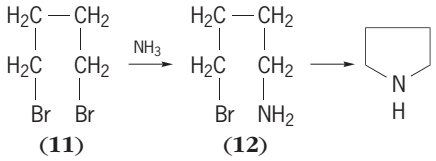

(4)

1,4-dibromobutane (**11**) or 4-bromobutylamine (**12**) can serve as starting material. 2-Ketopyrrolidine, or pyrrolidone (**14**), is of considerable interest in connection with the preparation of polyvinylpyrrolidone. Pyrrolidone, which can be formed from tetrahydrofuran by autoxidation in the presence of a cobalt catalyst and treatment of the resulting γ-butyrolactone (**13**) with ammonia, is combined with acetylene to form vinylpyrrolidone (**15**) [reaction (5)]. Polymerization of this material furnishes polyvinylpyrrolidone, a material of relatively high

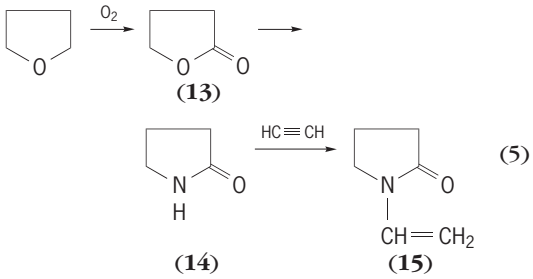

(5)

molecular weight, which is suitable for maintaining osmotic pressure in blood and so acting as an extender for plasma or whole blood. *See* PYRIDINE.

Walter J. Gensler

Bibliography. D. H. R. Barton et al. (eds.), *Comprehensive Organic Chemistry*, 1978; F. A. Carey and R. J. Sundberg, *Advanced Organic Chemistry*, pt. A: *Structure and Mechanisms*, 4th ed., 2000.

Pythagorean theorem

This theorem states that in any right triangle the square on the hypotenuse is equal to the sum of the squares on the other two sides: $r^2 = x^2 + y^2$. More than 100 different proofs have been given for this extremely important theorem of euclidean plane geometry, which is named for the Greek philosopher and mathematician Pythagoras (sixth century B.C.). One of these is based on a comparison of areas in the squares in **Fig. 1**.

The three-dimensional Pythagorean theorem (**Fig. 2**) may be phrased "the square of the diagonal of a rectangular box is equal to the sum of the squares

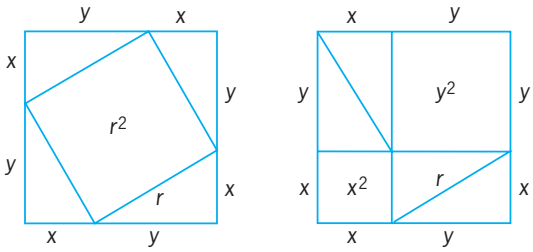

Fig. 1. Proof of Pythagorean theorem based on comparison of areas of squares.

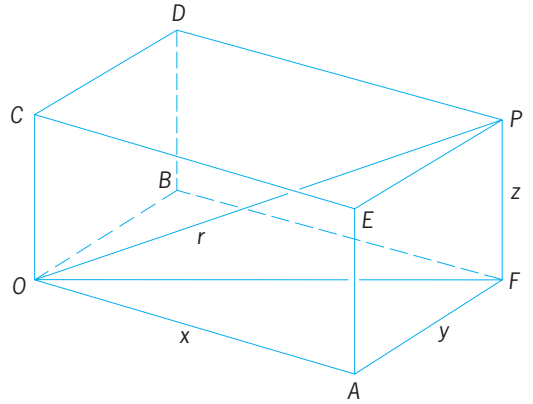

Fig. 2. Three-dimensional Pythagorean theorem.

of three adjacent edges that meet at a vertex: $r^2 = x^2 + y^2 + z^2$."

Squares that are equal to the sum of two or three squares were known to Pythagoras. For example, $3^2 + 4^2 = 5^2$ and $1^2 + 4^2 + 8^2 = 9^2$. The name pythagorean triple is given to any triple (x, y, r) of whole numbers such that $x^2 + y^2 = r^2$, and the name Pythagorean quadruple to such integral quadruples (x, y, z, r) for which $x^2 + y^2 + z^2 = r^2$. All pythagorean triples in which x, y, r are without common factor and x is odd are obtained by replacing the letters a and b in the triple $(a^2 - b^2, 2ab, a^2 + b^2)$ by whole numbers that have odd sum and no common factor; and similarly all such pythagorean quadruples are found by replacing a, b, c, d in the quadruple $(a^2 - b^2 + c^2 - d^2, 2ab + 2cd, 2bc - 2ad, a^2 + b^2 + c^2 + d^2)$ by whole numbers that have odd sum and no common factor.

J. Sutherland Frame

Q (electricity) — Quinone

Q (electricity)

Often called the quality factor of a circuit, Q is defined in various ways, depending upon the particular application. In the simple RL and RC series circuits, Q is the ratio of reactance to resistance, as in Eqs. (1),

$$Q = \frac{X_L}{R} \quad Q = \frac{X_C}{R} \quad \text{(a numerical value)} \tag{1}$$

where X_L is the inductive reactance, X_C is the capacitive reactance, and R is the resistance. An important application lies in the dissipation factor or loss angle when the constants of a coil or capacitor are measured by means of the alternating-current bridge.

Q has greater practical significance with respect to the resonant circuit, and a basic definition is given by Eq. (2), where Q_0 means evaluation at resonance.

$$Q_0 = 2\pi \frac{\text{max. stored energy per cycle}}{\text{energy lost per cycle}} \tag{2}$$

For certain circuits, such as cavity resonators, this is the only meaning Q can have.

For the RLC series resonant circuit with resonant frequency f_0, Eq. (3) holds, where R is the total

$$Q_0 = \frac{2\pi f_0 L}{R} = \frac{1}{2\pi f_0 CR} \tag{3}$$

circuit resistance, L is the inductance, and C is the capacitance. Q_0 is the Q of the coil if it contains practically the total resistance R. The greater the value of Q_0, the sharper will be the resonance peak.

The practical case of a coil of high Q_0 in parallel with a capacitor also leads to $Q_0 = 2\pi f_0 L/R$. R is the total series resistance of the loop, although the capacitor branch usually has negligible resistance.

In terms of the resonance curve, Eq. (4) holds,

$$Q_0 = \frac{f_0}{f_2 - f_1} \tag{4}$$

where f_0 is the frequency at resonance, and f_1 and f_2 are the frequencies at the half-power points. *See* RESONANCE (ALTERNATING-CURRENT CIRCUITS).

Burtis L. Robertson

Q meter

A direct-reading instrument formerly widely used to measure the Q of an electrical component or passive network of components at radio frequencies. The instrument measures the voltage applied to a component or circuit together with the voltage across a key component and calculates the Q of the circuit from their ratio. An impedance-measuring instrument (bridge or analyzer) is more likely to be employed now. *See* ELECTRICAL MEASUREMENTS; Q (ELECTRICITY); RADIO-FREQUENCY IMPEDANCE MEASUREMENTS.

Bryan P. Kibble

Bibliography. D. A. Bell, *Electronic Instrumentation and Measurements*, 2d ed., 1994; A. J. Diefenderfer and B. E. Holton, *Principles of Electronic Instrumentation*, 3d ed., 1994; L. Schnell (ed.), *Technology of Electrical Measurements*, 1993.

Quadrature

The condition in which the phase angle between two alternating quantities is 90°, corresponding to one-quarter of an electrical cycle. The electric and magnetic fields of electromagnetic radiation are in space quadrature, which means that they are at right angles in space. *See* ELECTROMAGNETIC RADIATION.

The current and voltage of an ideal coil are in quadrature because the current through the coil lags behind the voltage across the coil by exactly 90°. The current and voltage of an ideal capacitor are also in quadrature, but here the current through the capacitor leads the voltage across the capacitor by 90°.

In these last two cases the current and voltage are in time quadrature. *See* ALTERNATING-CURRENT CIRCUIT THEORY. John Markus

Quadric surface

A surface defined analytically by an equation of the second degree in three variables. If these variables are x, y, z, such an equation has the form of Eq. (1).

$$ax^2 + by^2 + cz^2 + 2exy + 2fxz + 2gyz + 2px + 2qy + 2rz + d = 0 \quad (1)$$

Every plane section of such a surface is a conic. To simplify the notation, the corresponding capital letters are used to denote the cofactors of the coefficients a, b, c, d, e, f, g, p, q, r in the determinant shown in Eq. (2). Thus $D = abc + 2efg - ag^2 - bf^2 - ce^2$. By λ_1, λ_2, λ_3 are denoted the three roots of the polynomial whose product is D, shown in Eq. (3).

$$\Delta = \begin{vmatrix} a, & e & f & p \\ e & b & g & q \\ f & g & c & r \\ p & q & r & d \end{vmatrix} \quad (2)$$

$$\phi(\lambda) = \begin{vmatrix} a-\lambda & e & f \\ e & b-\lambda & g \\ f & g & c-\lambda \end{vmatrix} \quad (3)$$

Quadrics are classified as central if $D \neq 0$, and noncentral if $D = 0$; nondegenerate if $\Delta \neq 0$, and degenerate if $\Delta = 0$. For further subclassifications it can be assumed that the coefficients in Eq. (1) are real numbers. The roots of $\phi(\lambda)$ will then be real also.

Central quadrics. The central quadric has a center of symmetry at the point $x = P/D$, $y = Q/D$, $z = R/D$. Its equation is freed of linear terms by the translation of axes $x' = x - P/D$, $y' = y - Q/D$, $z' = z - R/D$ and becomes Eq. (4).

$$ax'^2 + by'^2 + cz'^2 + 2ex'y' + 2fx'z' + 2gy'z' = -\frac{\Delta}{D} \quad (4)$$

By an appropriate rotation of axes this equation can be reduced further to the form of Eq. (5). If $\Delta = 0$,

$$\lambda_1 X^2 + \lambda_2 Y^2 + \lambda_3 Z^2 = -\frac{\Delta}{D} \quad (5)$$
$$\phi(\lambda_i) = 0$$

the degenerate central quadric of Eq. (1) is a quadric cone. Whenever the real roots λ_i of $\phi(\lambda) = 0$ are not all of the same sign, the quadric cone is real and is a surface formed of lines connecting the points of a real conic to a point outside its plane, called the vertex. If the roots λ_i are all of the same sign, the quadric cone is imaginary except for a single real point, the vertex.

If $\Delta \neq 0$, the nondegenerate central quadric of (1) is a real ellipsoid, a hyperboloid of one sheet, a hyperboloid of two sheets, or an imaginary ellipsoid according to whether three, two, one, or none of the λ_i are equal, the central quadric is a spheroid or a hyperboloid of revolution. If the three roots λ_i are equal, it is a sphere.

Noncentral quadrics. If $D = 0$, the quadric of (1) is noncentral, and $\phi(\lambda) = 0$ has a root $\lambda_3 = 0$. A suitable rotation of axes reduces Eq. (1) to Eq. (6), where $\Delta =$

$$\lambda_1 X^2 + \lambda_2 Y^2 + 2\mu_1 X + 2\mu_2 Y + 2\mu_3 Z + d = 0 \quad (6)$$

$-\lambda_1\lambda_2\mu_3{}^2$. A translation further reduces μ_1 to 0 if $\lambda_1 \neq 0$, μ_2 to 0 if $\lambda_2 \neq 0$, and d to 0 if $\mu_3 \neq 0$. A nondegenerate noncentral quadric ($D = 0$, $\Delta \neq 0$) is an elliptic paraboloid if $\lambda_1\lambda_2 > 0$, or a hyperbolic paraboloid if $\lambda_1\lambda_2 < 0$. A degenerate noncentral quadric ($D = \Delta = 0$) is either a quadric cylinder (if at least one of the cofactors in Δ is not 0) or a pair of planes, a single plane or a line, or it has no real locus. If $\lambda_1\lambda_2 > 0$ and $\mu_3 = 0$, the surface may be a real elliptic cylinder, a single line, or an imaginary elliptic cylinder, depending on the value of d.

If $\lambda_1\lambda_2 < 0$ and $\mu_3 = 0$, the surface is a hyperbolic cylinder or a pair of planes. If either $\lambda_1\mu_3$ or $\lambda_2\mu_3$ is 0, but not both, the surface is a parabolic cylinder. *See* CYLINDER. J. Sutherland Frame

Qualitative chemical analysis

The branch of chemistry concerned with identifying the elements and compounds present in a sample of matter. Inorganic qualitative analysis traditionally used classical "wet" methods to detect elements or groups of chemically similar elements, but instrumental methods have largely superseded the test-tube methods. Methods for the detection of organic compounds or classes of compounds have become increasingly available and important in organic, forensic, and clinical chemistry. Once it is known which elements and compounds are present, the role of quantitative analysis is to determine the composition of the sample. *See* ANALYTICAL CHEMISTRY; QUANTITATIVE CHEMICAL ANALYSIS.

A considerable part is played by the human senses as analytical tools. There are surprisingly sensitive standard procedures which employ the human sense of smell for the qualitative detection and semiquantitative estimation of extremely small concentrations of objectionable odor-causing substances in water supplies. Perfumers develop a skill based on the sense of smell that is remarkable in identifying and making use of odoriferous substances, both to blend perfumes and to disguise objectionable odors. Tea tasters and wine experts depend on a developed skill that is a combination of taste and smell. The trained eye can serve as a relatively good spectrophotometer, both qualitatively in regard to absorption maxima and to some extent quantitatively to measure color intensity. Prospectors and geologists greatly rely on visual skills for the recognition of ores and rock types.

Inorganic Analysis

The operating principles of all systematic inorganic qualitative analysis schemes for the elements are

similar: separation into groups by reagents producing a phase change; isolation of individual elements within a group by selective reactions; and confirmation of the presence of individual elements by specific tests. In the classical cation scheme, the wide range of solubilities of sulfide salts of the metals was exploited through control of sulfide ion concentration by pH control. Outlines for short qualitative schemes for several metals and three common anions of interest in the environment are presented in **Figs. 1** and **2**. The cation scheme has been adapted from a more extensive non–hydrogen sulfide scheme. The outlines show the reagents and steps used, and incomplete chemical equations are presented to illustrate the reactions and conditions employed. Although they avoid the use of hydrogen sulfide or thioacetamide (important reagents in earlier methods), the schemes include most of the basic qualitative analysis techniques of selective and pH-controlled precipitation, oxidation and reduction, complex ion formation, gas evolution, and a flame test.

Sample size. Through usage and tradition, descriptive terms for sample sizes have been:

macro	0.1 gram or more
semimicro	0.01 to 0.1 gram
micro	1 milligram (10^{-3} g)
ultramicro	1 microgram (10^{-6} g)
submicrogram	less than 1 microgram

For defining the smallest amount of a substance that

Cations included: Hg_2^{2+}, Cr^{3+}, Sb^{3+}, Pb^{2+}, Sr^{2+}, Cd^{2+}, Cu^{2+}, Hg^{2+}, As^{3+}
Step 1: Add HCl

Precipitate: $Hg_2Cl_2 \downarrow$ (white)
$Hg_2Cl_2 + NH_4OH \rightarrow HgO \downarrow$ (black mixture with $HgNH_2Cl$ confirms Hg_2^{2+})

Solution: Cr^{3+}, Sb^{3+}, Pb^{2+}, Sr^{2+}, Cd^{2+}, Cu^{2+}, Hg^{2+}, As^{3+}

Step 2: Adjust pH to 3.5–4.0, add NH_4 benzoate and Na^+ benzoate

Precipitate: $Cr(OH)(benzoate)_2 \downarrow$, $SbOCl \downarrow$
$Cr(OH)(benzoate)_2 + OH^- + H_2O_2 \rightarrow CrO_4^{2-}$
$CrO_4^{2-} + Pb^{2+} \rightarrow PbCrO_4 \downarrow$ (yellow, confirms Cr^{3+})
$SbOCl + Na^+$ tartrate $\rightarrow$ Sb-tartrate complex, acidify with HNO_3
Sb-tartrate complex $+ H_2S \rightarrow Sb_2S_3 \downarrow$ (orange, confirms Sb^{3+})

Solution: Pb^{2+}, Sr^{2+}, Cd^{2+}, Cu^{2+}, Hg^{2+}, As^{3+}

Step 3: Add NaF, adjust pH to 3.4

Precipitate: $PbF_2 \downarrow$, $SrF_2 \downarrow$
$PbF_2 + 2NH_4^+$ acetate $\rightarrow Pb(acetate)_2$
$Pb(acetate)_2 + S^{2-} \rightarrow PbS \downarrow$ (black, confirms Pb^{2+})
$SrF_2 + 2HCl \rightarrow SrCl_2$
$SrCl_2$, flame test $\rightarrow$ crimson flame confirms Sr^{2+}

Solution: Cd^{2+}, Cu^{2+}, Hg^{2+}, As^{3+}

Step 4: Make alkaline with NaOH, boil free of NH_3

Precipitate: $Cd(OH)_2 \downarrow$, $Cu(OH)_2 \downarrow$, $HgO \downarrow$
$Cd(OH)_2 + 2HCl \rightarrow Cd^{2+}$, treat with NH_4OH, acidify with HCl
$Cd^{2+} + S^{2-} \rightarrow CdS \downarrow$ (yellow, confirms Cd^{2+})
$Cu(OH)_2 + 2HCl \rightarrow Cu^{2+}$, treat with NH_4OH, acidify with HCl
$2Cu^{2+} + 4SCN^- \rightarrow (SCN)_2 + 2CuSCN \downarrow$ (white, indicative of Cu^{2+})
$2CuSCN + HNO_3$, oxidize $\rightarrow 2Cu^{2+}$
$Cu^{2+} + 4NH_3 \rightarrow [Cu(NH_3)_4]^{2+}$ (deep blue, confirms Cu^{2+})
$HgO + 2HCl \rightarrow HgCl_2$
$HgCl_2 + NH_4OH \rightarrow HgNH_2Cl \downarrow$
$HgNH_2Cl + HNO_3 \rightarrow HgCl_2$; two tests follow:
a. $2HgCl_2 + Sn^{2+} \rightarrow Hg_2Cl_2 \downarrow$ (white)
$Hg_2Cl_2 + Sn^{2+} \rightarrow 2HgO$ (white turning gray confirms Hg^{2+})
b. $HgCl_2 + 2I^- \rightarrow HgI_2 \downarrow$ (red-orange, confirms Hg^{2+})

Solution: As^{3+}

Step 5: Acidify with 12 *M* HCl, boil to reduce volume, add 15*M* NH_4OH, 6*M* NaOH, and Al granules

Solution: $As^{3+} + 4OH^- \rightarrow AsO_2^- + 2H_2O$
$2AsO_2^- + 2Al + 3H_2O \rightarrow 2AsH_3 \uparrow$ (gas)
$2AsH_3 + 6Ag^+ \rightarrow 6AgO \downarrow$ (black precipitate on $AgNO_3$-treated paper strip confirms As^{3+})

Fig. 1. Outline of a scheme of analysis for cations of interest in the environment. (*After J. L. Lambert and C. E. Meloan, A simple qualitative analysis scheme for several environmentally important elements, J. Chem. Educ., 54:249, 1977*)

Anions included: NO_3^-, SO_4^{2-}, Cl^-
Procedure: Add Na_2CO_3, boil

Precipitate: waste (carbonates, hydroxides, and oxides of metal cations in Fig. 1, except As^{3+})	Solution: NO_3^-, SO_4^{2-}, Cl^- (as Na^+ salts) Divide into two portions, a and b: a. Acidify with HCl to expel CO_2 add 18*M* H_2SO_4 as separate layer without mixing $HNO_3 + Fe^{2+} \rightarrow [Fe(NO)]^{2+}$ (brown ring at interface of solution and 18*M* H_2SO_4 confirms NO_3^-) b. Acidify with HNO_3 to expel CO_2 1. $SO_4^{2-} + Ba^{2+} \rightarrow BaSO_4 \downarrow$ (white, confirms SO_4^{2-}) 2. $Cl^- + Ag^+ \rightarrow AgCl \downarrow$ (white to pale violet, confirms Cl^-)

Fig. 2. Outline of a scheme of analysis for three common anions. (*After J. L. Lambert and C. E. Meloan, A simple qualitative analysis scheme for several environmentally important elements, Chem. Educ., 54:249, 1977*)

can be detected by a given method, the term "limit of identification" is used. Under favorable conditions, an extremely sensitive method can detect as little as 10^{-15} (the ability to detect a single cesium atom in gaseous methane by a laser-induced photoionization procedure has been reported, although it hardly qualifies as an analytical method). If the smallest amount is present in extremely dilute solution, "limiting concentration" defines the lowest concentration of the substance in solution that will give the test. If another substance is present that gives similar results or interferes with the test, "limiting proportion" defines the ratio of the mass of the substance sought to the largest mass of interfering substance that can be present without producing spurious results.

Visual tests. Spot tests are selective or specific single qualitative chemical tests carried out on a spot plate (a glass or porcelain plate with small depressions in which drop-size reactions can be carried out), on paper, or on a microscope slide. On paper or specially prepared adsorbent surfaces, spot tests become one- or two-dimensional through the use of solvent migration and differential adsorption (thin-layer chromatography). Containing selected indicator dyes, pH indicator paper strips are widely used for pH estimation. Solid reagent monitoring devices and indicator tubes are used for the detection and estimation of pollutant gases in air. *See* CHROMATOGRAPHY; PH.

By use of the microscope, crystal size and habit can be used for qualitative identification. With the addition of polarized light, chemical microscopy becomes a versatile method widely employed in forensic analysis, geology, ceramics, and for the identification of particulate matter in air. *See* CHEMICAL MICROSCOPY; FORENSIC CHEMISTRY.

Instrumental methods. Generally speaking, instrumental methods of analysis have been developed primarily for quantitative analysis, but any quantitative method can be adapted to qualitative analysis. Some, such as electrochemical methods, are not often used in qualitative analysis (an exception would be the ubiquitous pH meter). Others, such as column chromatography (gas and liquid) and mass spectrometry, are costly and complicated but capable of providing unique results.

Spectroscopy. The first instrumental method—emission spectroscopy—began as a qualitative tool. Emission spectroscopy still is important in establishing the presence or absence of a suspected element in forensic analysis. The simple qualitative flame test developed into flame photometry and subsequently into atomic absorption spectrometry, which is one of the most widely used analytical methods today. Absorption spectrometry (particularly in the infrared region), mass spectrometry, and nuclear magnetic resonance are extensively used for identification and structural analysis or organic compounds. *See* ATOMIC SPECTROMETRY; EMISSION SPECTROCHEMICAL ANALYSIS.

X-rays, electrons, and ions. Bombardment of surfaces by x-rays, electrons, and positive ions has given rise to a number of methods useful in analytical chemistry. X-ray diffraction is used to determine crystal structure and to identify crystalline substances by means of their diffraction patterns. X-ray fluorescence analysis employs x-rays to excite emission of characteristic x-rays by elements. The electron microprobe uses electron bombardment in a similar manner to excite x-ray emission. Bombarding electrons can also cause emission of electrons having energies characteristic of the target atoms. Bombardment by H^+ or He^+ ions permits identification and determination of elements by measurement of backscattering of the bombarding ions, emission of secondary ions for analysis by a mass spectrometer, or emission of characteristic photons for spectrophotometric analysis. *See* SECONDARY ION MASS SPECTROMETRY (SIMS); X-RAY DIFFRACTION; X-RAY FLUORESCENCE ANALYSIS.

Activation analysis. Neutron activation analysis, wherein neutron bombardment induces radioactivity in isotopes of elements not naturally radioactive, involves measurement of the characteristic gamma radiation and modes of decay produced in the target atoms. This bombardment method is not restricted to surface analysis, and is independent of the chemical combinations of the target elements. Some elements can be detected in solids in concentrations as low as 1 in 10^{12} by this method. Mössbauer spectroscopy, by contrast, involves selective absorption of gamma radiation from an excited isotope by

the stable isotope of the same atomic and mass number, and is exceedingly sensitive to the chemical environment of the target atoms. *See* ACTIVATION ANALYSIS.

Jack L. Lambert

Organic Analysis

Qualitative analysis of an organic compound is the process by which the characterization of its class and structure is determined. Due to the numerous classes of organic compounds (alcohols, aldehydes, ketones, and so forth) and the complexity of their molecular structures, a systematic analytical procedure is often required. A typical procedure entails an initial assignment of compound classification, followed by a complete identification of the molecular structure.

Physical methods. The initial step is an examination of the physical characteristics. Color, odor, and physical state can be valuable clues. For example, nitro compounds are often highly colored. The color of a sample is often misleading. Air oxidation and sample decomposition can result in color formation. Some compound classes have characteristic odors. While many esters have a sweet odor, thiols yield a very disagreeable odor. Caution should be exercised when investigating the odor of an unknown compound. Direct inhalation should be avoided.

The physical constants of an unknown compound provide pertinent data for the analyst. Constants such as melting point, boiling point, specific gravity, and refractive index are commonly measured. The melting point is usually determined by packing a small quantity of the solid in a capillary tube. The tube is then heated in either an oil bath (Thiel tube method) or an aluminum block (electrically heated) to the melting-point temperature. For a more precise measurement, a microscope equipped with a hot stage can be employed. The temperature range over which a solid melts is a measure of compound purity. Samples with wide ranges should be recrystallized. Boiling-point measurements are conducted by utilizing a small-scale distillation apparatus. A "center cut" from the distillation is taken for the boiling point. Wide boiling-point ranges also indicate sample impurity. The measurements of specific gravity and refractive index are less common due to their limited utility. Specific gravity is determined with a gravitometer or a pycnometer. Refractive index can be measured with a refractometer. For solid samples, chemical microscopy (Becke test) may be employed.

Chemical methods. The preliminary chemical tests which are applied are elemental analysis procedures. The elements generally associated with carbon, hydrogen, and oxygen are sulfur, nitrogen, and the halogens. The analyst is usually interested in the latter group. These elements are converted to water-soluble ionic compounds via sodium fusion. The resulting products are then detected with wet chemical tests.

A sample is fused with sodium metal by adding it to the sodium and heating to the melting temperature. After a short heating period, the solution is cooled, the excess sodium is dissolved with ethanol, and the fusion products are dissolved in water. Sulfur can be detected by the formation of a lead sulfide precipitate upon the addition of lead acetate. A variety of procedures for the detection of nitrogen are known. The resulting products are usually precipitates or colored solutions. The results of nitrogen tests are sometimes inconclusive due to the difficulty in detecting nitrogen in nitro compounds. The halogens can be detected by the evidence of silver halide precipitates upon the addition of silver nitrate. The identity of the specific halogen can be determined by more selective reactions or by x-ray diffraction.

The solubility of a compound in various liquids provides information concerning the molecular weight and functional groups present in the compound. A systematic solubility study is outlined in **Fig. 3**. On the basis of solubility, assumptions can often be made that reduce the number of tests necessary to classify a compound. Amines are basic compounds, whereas phenols are weakly acidic.

In order to indicate the presence or absence of a functional group, specific classification reactions are tested. The products of these tests are usually precipitates or colored solutions. The reactions are simple, are rapid, and require a small quantity of sample. No single test is conclusive evidence. A judicious choice of reactions can confirm or negate the presence of a functional group. A list of the popular classification tests is presented in **Table 1**.

Once thee identity of a functional group has been established, the structure must be determined. Derivatization of the unknown is one procedure employed for complete identification. Derivatives prepared from the unknown are compared with those of known compounds. The agreement of melting points is the basis of an identification. The majority of derivatives are solids and are produced in high yield. This allows for a facile isolation and purification. Extensive tabulations of derivative melting points have been compiled. A list of derivatization reagents are presented in **Table 2**.

Chemical structure. Instrumental methods are commonly applied for functional group determination and structure identification. Revolutionary developments in instrumentation and data interpretation have greatly simplified the analytical process. Many of the time-consuming procedures have been replaced by instrumental techniques.

Absorption spectroscopy methods are among the most important techniques. Whenever a molecule is exposed to electromagnetic radiation, certain wavelengths cause vibrational, rotational, or electronic effects within the molecule. The radiation required to cause these effects is absorbed. The nature and configuration of the atoms determine which specific wavelengths are absorbed. Radiation in the ultraviolet and visible regions is associated with electronic effects, and infrared radiation is associated with rotational and vibrational effects.

Infrared absorption spectroscopy is perhaps the most valuable instrument for functional group

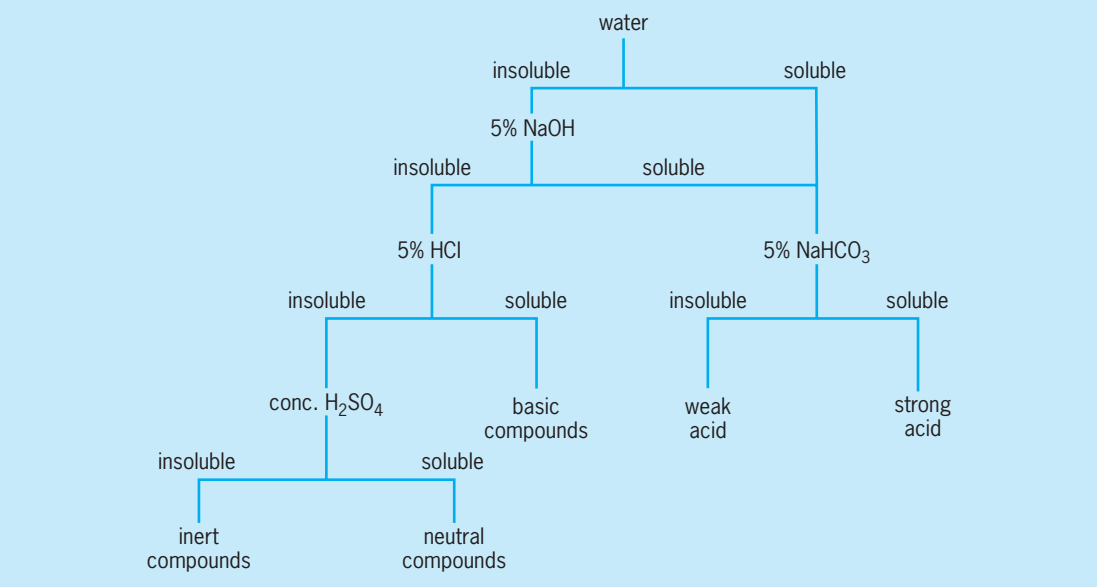

Fig. 3. Diagram of a typical solubility scheme.

determination. A beam of infrared radiation is passed through the sample. The transparent radiation is dispersed by means of a prism or diffraction grating. With the aid of a slit, the intensity of a narrow wavelength region is measured. The entire infrared region is scanned, and an absorbance-versus-wavelength scan is obtained. Based on the wavelengths at which radiation is absorbed, a functional group can be identified. *See* INFRARED SPECTROSCOPY.

The principle of ultraviolet absorption spectroscopy is similar to that of infrared spectroscopy, except that ultraviolet radiation is monitored. Unlike infrared scans, where numerous sharp absorption bands are obtained, ultraviolet scans normally exhibit one or two broad bands of limited utility for qualitative purposes. Visible absorption spectroscopy suffers the same limitations.

Raman spectroscopy is slightly different from the other spectroscopic techniques. A sample is irradiated by a monochromatic source. Depending upon the vibrational and rotational energies of the functional groups in the sample, light is scattered from the sample in a way which is characteristic of the functional groups. Raman spectroscopy is a fine complement to infrared spectroscopy because the two are sensitive for different groups. This technique has not enjoyed widespread application as a qualitative tool due to the high cost of the instrumentation and the difficulty in obtaining good spectra. *See* RAMAN EFFECT.

Instrumental techniques are also valuable for structure identification. The most popular are infrared spectroscopy, nuclear magnetic resonance, and mass spectrometry. Other, less common techniques are electron spin resonance, x-ray diffraction, and nuclear quadrupole spectroscopy.

In addition to the value of infrared for functional group analysis, an infrared spectrum can identify

TABLE 1. Common classification reactions

Compound	Reagent	Comments
Alcohols	Hydrochloric acid/zinc chloride	Forms alkyl halide
	Ceric nitrate	Color change yellow to red
	Chromic anhydride	Distinguishes 1° and 2° from 3° alcohols
Aldehydes	Benedict's solution	Colored precipitate
	Tollens' reagent	Forms silver mirror
	Fuchsin-aldehyde	Color change
Aldehydes and ketones	2,4-Dinitrophenyl hydrazine	Hydrazone precipitate
	Hydroxylamine hydrochloride	HCl evolved
Amines	Benzenesulfonyl chloride	Hinsberg method: distinguishes 1°, 2°, and 3° amines
	Acid chloride	Amide formation
Esters	Hydroxamic acid test	Magenta color formation
	Sodium hydroxide	Saponification
Ethers	Hydriodic acid	Zeisel method: light orange color produced
Nitro compounds	Zinc and ammonium chloride/Tollens' reagent	Silver mirror
	Ferrous hydroxide	Red-brown precipitate
Unsaturation	Potassium permanganate	Loss of color
	Bromine in carbon tetrachloride	Loss of color

TABLE 2. Common derivatization reaction

Compound	Reagent	Derivative
Alcohols	Phenyl isocyanate	Phenyl urethane
	3,5-Dinitrobenzyl chloride	3,5-Dinitrobenzoate
Aldehydes and ketones	2,4-Dinitrophenyl hydrazine	2,4-Dinitrophenyl hydrazone
	Semicarbazide hydrochloride	Semicarbazone
Alkyl and aryl halides	Magnesium/mercuric chloride	Alkyl mercuric halide
Amines 1° and 2°	Benzenesulfonyl chloride	Benzenesulfonamide
	Phenyl isothiocyanate	Phenyl urea
	Acetic anhydride	Acetamide
	Benzoyl chloride	Benzamide
Amines 3°	Methyl iodide	Quaternary ammonium salt
	Benzyl chloride	
Aromatic hydrocarbons	H_2SO_4/HNO_3	Nitro compound
Esters	Benzylamine/ammonium chloride	*N*-Benzylamide
	Hydrazine	Acid hydrazide
Ethers, aromatic	Bromine	Bromo compound
	Picric acid	Picrate
Ethers, aliphatic	3,5-Dinitrobenzyol chloride/zinc chloride	3,5-Dinitrobenzoate
Phenols	Bromine	Bromo compounds
	α-Naphthyl isocyanate	α-Naphthyl urethane
	Acetic anhydride	Acetate

a compound. The spectral region 6.0–14.0 μm is known as the fingerprint region. Comparison to known spectra can identify an unknown compound. Extensive files of infrared spectra are available for this purpose.

Mass spectroscopy is a powerful tool for structure identification. A spectrum is obtained by bombarding a sample with an electron beam. The resulting fragments are accelerated through a magnetic field and are deflected from their original path. The deflection is a function of both mass and charge. The separated fragments are then detected. Based on the fragmentation pattern and the mass-to-charge ratio of the fragments, structural information is obtained. A fingerprint identification method can also be employed. *See* MASS SPECTROMETRY.

Nuclear magnetic resonance is another powerful structural tool. As in other spectroscopic methods, a sample is irradiated and the absorbed radiation is monitored. Similar to electrons, certain nuclei (^{1}H, ^{13}C, and so on) can be considered to spin. The rotation of the charged particle creates a magnetic moment about the axis of spin. When a sample is placed in a magnetic field, the nuclei are aligned in a given direction. In order to reverse the direction, energy is required. The changing of spin states results in absorption of the applied radio-frequency radiation. Either the radio frequency or the magnetic field can be held constant while the other is changed and the radiation absorption is monitored. In practice it is more convenient to hold the radio frequency constant and vary the magnetic-field strength. The energy required to change spin states is also dependent upon neighboring groups in the molecule. Based on the number and intensity of absorptions and the field strength required, structural information can be obtained. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

Physical structure. In addition to chemical structure, the physical structure of a sample can be important. Certain materials behave differently depending on characteristics such as crystallinity, cross-linking, and their association as dimers, trimers, and so on. Thermal analysis is a useful procedure for examining these structural characteristics. Among the most popular methods are thermogravimetry, differential thermal analysis, differential scanning calorimetry, and thermal mechanical analysis. *See* CALORIMETRY.

Although the implementation of instrumental methods have greatly simplified qualitative analytical procedures, no individual instrument is capable of complete identification for all samples. A complement of instrumental and wet chemical techniques is generally required for adequate proof of identification. *See* SPECTROSCOPY.

Sidney Siggia; Kenneth Longmoore

Bibliography. F. Feigl, *Spot Tests in Inorganic Analysis*, 6th ed., 1972; W. Kemp, *Qualitative Organic Analysis: Spectrochemical Techniques*, 2d ed., 1988; J. Lagowski and H. Sorum, *Introduction to Semimicro Qualitative Analysis*, 8th ed., 2004; R. L. Shriner et al., *The Systematic Identification of Organic Compounds*, 8th ed., 2003; S. Siggia and J. G. Hanna, *Quantitative Organic Analysis via Functional Groups*, 4th ed., 1979, reprint 1988; G. I. Svehla, *Vogel's Qualitative Inorganic Analysis*, 7th ed., 1996.

Quality control

The operational techniques and the activities that sustain the quality of a product or service in order to satisfy given requirements. Quality control is a major component of total quality management and is applicable to all phases of the product life cycle: design, development, manufacturing, delivery and installation, and operation and maintenance.

The quality-control cycle consists of four steps: quality planning, data collection, data analysis, and implementation. Quality planning consists of

defining measurable quality objectives. Quality objectives are specific to the product or service and to the phase in their life cycle, and they should reflect the customer's requirements.

The collection of data about product characteristics that are relevant to the quality objectives is a key element of quality control. These data include quantitative measurements (measurement by variables), as well as determination of compliance with given standards, specifications, and required product features (measurement by attributes). Measurements may be objective, that is, of physical characteristics, which are often used in the control of the quality of services. Since quality control was originally developed for mass manufacturing, which relied on division of labor, measurements were often done by a separate department. However, in the culture of Total Quality Management, inspection is often done by the same individual or team producing the item.

The data are analyzed in order to identify situations that may have an adverse effect on quality and may require corrective or preventive action.

The implementation of those actions as indicated by the analysis of the data is undertaken, including modifications of the product design or the production process, to achieve continuous and sustainable improvement in the product and in customer satisfaction.

The methods and techniques for data analysis in quality control are generic and can be applied to a variety of situations. The techniques are divided into three main categories: diagnostic techniques; process control, which includes process capability assessment and control charts; and acceptance sampling.

Diagnostic techniques. Diagnostic techniques serve to identify and pinpoint problems or potential problems that affect the quality of processes and products, and include the use of flowcharts, cause-and-effect diagrams, histograms, Pareto diagrams, location diagrams, scatter plots, and boxplots.

Flowcharts. A flowchart is a graphical description of a process in terms of its constituent steps and the sequence among them (**Fig. 1**). Various symbols are used to represent the steps and the types of relationships among them. The most common types of steps are processing, waiting, transportation, inspection, and storage. Flowcharts are useful in evaluating the way that a process is actually executed relative to its original design. Preparing a flowchart for the process is often a good first step toward improving it. Often, when the process is clearly defined and documented, bottlenecks and redundant steps become apparent, and opportunities for improvement are obvious. It is also helpful to classify the various steps of the process into three groups: (1) Value-adding steps create or modify the product features that are required to meet customer expectations, and typically include manufacturing, assembly, and delivery activities. These steps are the essence of the process and should be carried out as efficiently and effectively as possible. (2) Process-sustaining steps are required to manage the process, such as record-

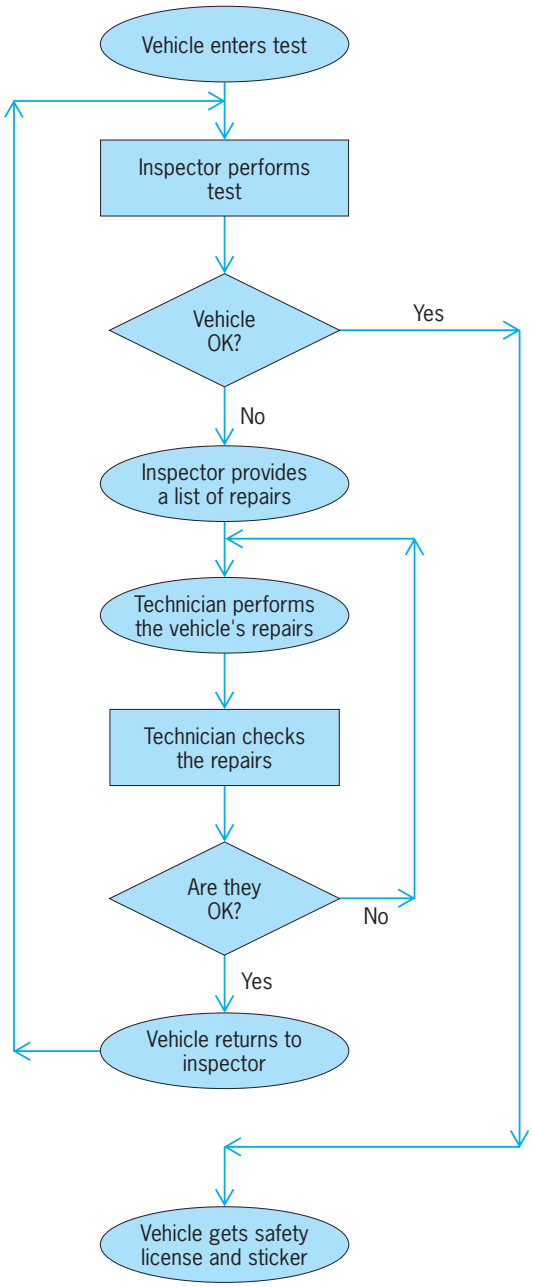

Fig. 1. Flowchart used to monitor a vehicle safety test.

keeping, production control, and invoicing, and they should be minimized to the extent possible. (3) Non-value-adding steps, such as waiting, in-process storage, rework, and certain types of inspection, should be eliminated from the process.

Cause-and-effect diagrams. These are also called fishbone diagrams because of their shape, or Ishikawa diagrams in honor of the Japanese engineer who made their use popular. The cause-and-effect diagram is used to describe the relationship among the factors that may contribute to a quality problem. Construction of the cause-and-effect diagram begins by identifying the variable or phenomenon of interest. This might be the yield of a process, the response rate from a customer survey, the performance of a particular assembly, the variation of a process, or any other

variable that bears on the quality of the product or service. An arrow is drawn from left to right, with the tip pointing to the variable. Then, main contributory factors are drawn as arrows leading into the main arrow, and secondary and lower-order contributory factors are shown as feeding into their respective arrows (**Fig. 2**).

Histogram. A histogram is a frequency distribution plot. The horizontal axis represents different values or categories, and the vertical axis represents the number of occurrences observed for each value or category (**Fig. 3**). Histograms are helpful for organizing the data collected from measurements taken on a product or process. The histogram can then be analyzed for any unusual pattern, such as skewed distribution, multimodal distribution, and long tail. These patterns are analyzed for indications of problems and are interpreted based on the specific quality objectives of the product.

Pareto diagram. This is a histogram of defects or quality problems classified by type and sorted in descending order of frequency (**Fig. 4**). The Pareto diagram, named after the Italian economist Vilfredo Pareto, serves to focus attention on the main sources of problems. A variation of this type of technique is called the Pareto pyramid, where two Pareto diagrams, representing the situation before and after an intervention, are shown side by side with the axes transposed. *See* PARETO'S LAW.

Location diagram. This type of chart shows the distribution of defects by their location on the process or the product. For example, it can show the layout of a printed circuit board, with the number of soldering defects shown next to each soldering point (**Fig. 5**).

Scatter plot. A scatter plot is a two-dimensional plot showing the relationship between two variables, each plotted along one axis (**Fig. 6**). The pattern of points suggests the type of correlation between the two variables (positive, negative, nonlinear) and its strength. Correlation between two variables can be exploited to improve the performance of the dependent variable. This is achieved by manipulating the independent variable in a direction such that it will cause the dependent variable to change in the desired direction.

Box plot. This plot summarizes the results of successive sample observations from the process taken over time. The horizontal axis represents the sample number (or time). On the vertical axis, each box summarizes a sample with five numbers: the upper and lower quartiles, which form the ends of the box; the median, which is the line dividing the box; and the minimum and maximum values, which appear at the end of the whiskers. Box plots analyze the process behavior over time and the relationship between its mean and its variability.

Process control. Process-control methods are applicable to systems that produce a stream of product units, either goods or services. They serve to control the processes that affect those product characteristics that are relevant to quality as defined in the quality objectives. For example, in a system that produces metal parts, some of the processes that might need to be controlled are cutting, machining, deburring, bending, and coating. The relevant product characteristics are typically spelled out in the specifications in terms of physical dimensions, position of features,

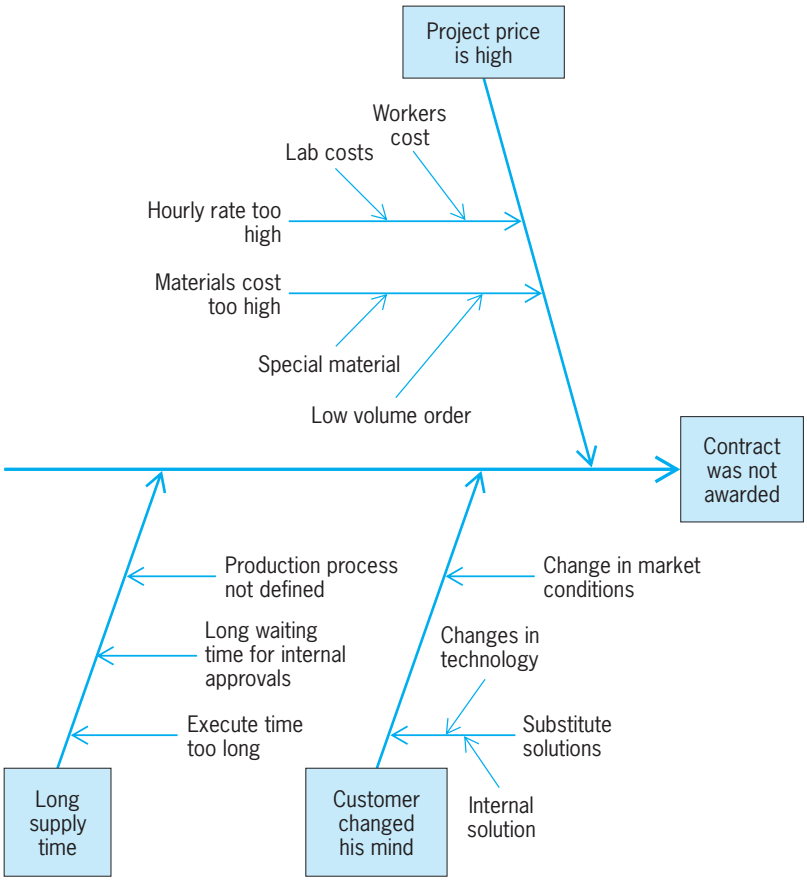

Fig. 2. Cause-and-effect diagram.

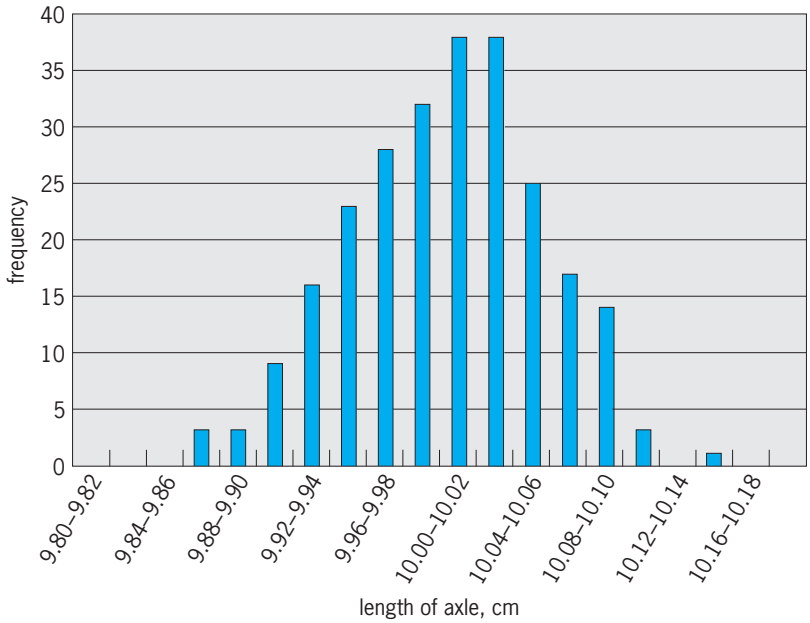

Fig. 3. Histogram showing the length distribution of a manufactured axle.

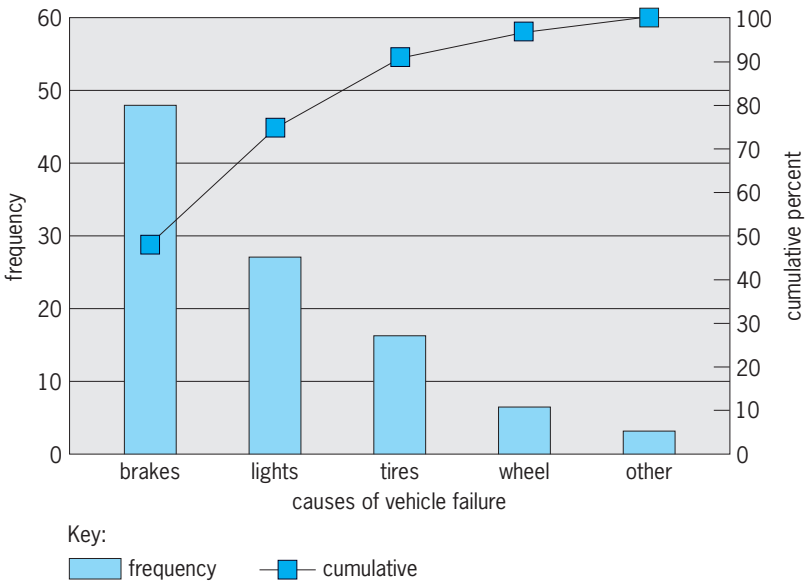

Fig. 4. Pareto diagram showing causes and failure rates in a vehicle test.

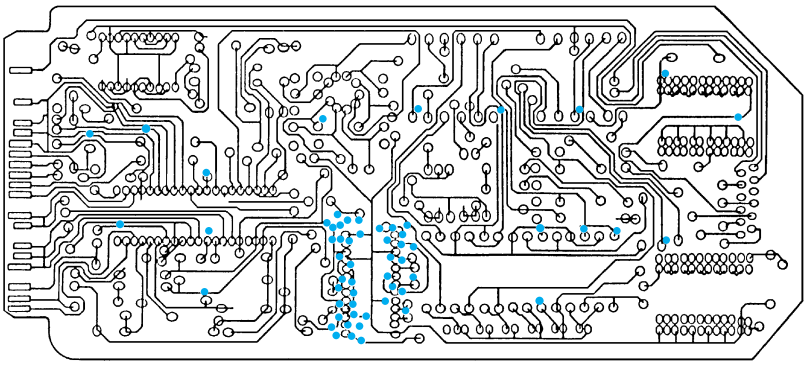

Fig. 5. Location diagram showing the places that are not functioning as expected (white circles) in a jet engine.

surface smoothness, material hardness, paint thickness, and so on. In a system that produces a service, such as a telephone help line, the relevant processes could be answering the call, identifying the problem, and solving the problem. The characteristics that are relevant to quality as perceived by the customer might include response time, number of referrals, frequency of repeat calls for the same problem, and elapsed time to closure.

Process control focuses on keeping the process operating at a level that can meet quality objectives, while accounting for random variations over which there is no control. There are two main aspects to process control: control charts and capability analysis. Control charts are designed to ascertain the statistical stability of the process and to detect changes in its level or variability that are due to assignable causes and can be corrected. Capability analysis considers the ability of the process to meet quality objectives as implied by the product specifications.

Process-control techniques were originally developed for manufactured goods, but they can be applied to a variety of situations as long as the statistical distribution of the characteristics of interest can be approximated by the normal distribution. In other cases, the principles still apply, but the formula may need to be modified to reflect the specific mathematical expression of the probability distribution functions. *See* PROCESS CONTROL.

Control charts. Every process has some inherent variability due to random factors over which there is no control and which cannot economically be eliminated from the process. This inherent variability is the aggregate result of many individual causes, each having a small impact. Control charts are designed to detect excessive variability due to specific assignable causes that can be corrected; they serve to determine whether a process is or is not in a state of statistical control. Being in statistical control means that the extent of variation of the output of the process does not exceed that which is expected based on the natural statistical variability of the process.

Samples are drawn from the process output at regular intervals, and a statistic is calculated from the observations in the sample and is charted over time. A nonrandom pattern of points, such as points outside the control limits, indicates lack of statistical control. If a nonrandom pattern is detected, the process should be stopped and an investigation should be done to determine the causes for the change. Once the causes of the change have been ascertained and any required corrective action has been taken, the process can be resumed. *See* CONTROL CHART.

Process capability analysis. The capability of a process is defined as the range of values that the process generates while operating under statistical control. Typically, the process capability is ± three standard deviations from the process mean, for a total of six standard deviations. The standard deviation of the process should be estimated from historical data collected while the process was operating under normal, stable circumstances. Control charts can be used to establish the statistical stability of the process and the absence of assignable causes.

The process capability is compared to the tolerance spread, that is, the range of values that meet the specifications for the product and are deemed acceptable by the customer. This range is determined

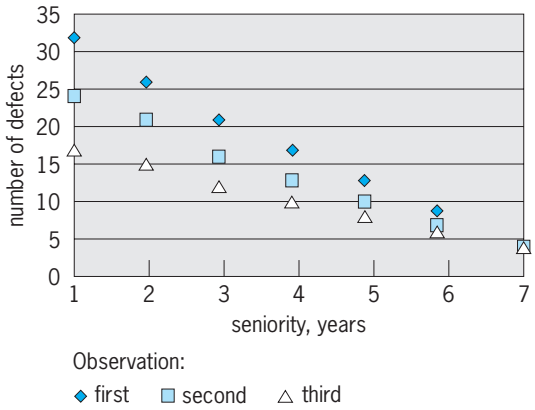

Fig. 6. Scatter plot showing a negative correlation between number of defects and worker seniority.

by the upper specification limit (USL) and the lower specification limit (LSL), which together define the range of values that will meet the requirements from the product. The process capability index (C_p) is defined as the ratio between the tolerance spread (USL–LSL) and the process capability [six standard deviations (6σ)] as in the equation below. If the capability

$$C_p = \frac{\text{USL–LSL}}{6\sigma}$$

index is smaller than 1, a certain fraction of the process output will fall outside the tolerance spread and will not be acceptable, entailing lower quality and higher costs through rework and scrap. When the capability index is equal to 1, if the process is perfectly centered in the tolerance interval, and assuming the normal distribution, virtually all the process output will fall within specifications and the fraction of units that fail to conform to the specifications will be very small, about 0.27%. However, if the process mean shifts to either side, the fraction nonconforming will increase rapidly. For example, a shift of $^1/_2$ standard deviation results in 0.62% nonconforming, and a shift of 1 standard deviation results in 2.28% nonconforming. In general, the capability index should be greater than 1, with some recommended values being at least 1.33 (tolerance spread of $\pm 4\sigma$) and possibly 2.0 (tolerance spread of $\pm 6\sigma$).

There are two avenues for improving the capability index: reducing the natural variability of the process, and increasing the tolerance spread. Reduction of the process variability can be achieved by operating the process under more consistent conditions (such as uniform inputs, better-trained operators, or more frequent adjustments). A focus on continuous process improvement can help in that respect by eliminating sources of variation and promoting more consistent execution. Some of the diagnostic techniques, such as the Pareto diagram and the cause-and-effect diagram, can help identify the aspects in the process that can yield effective variability reductions.

The second avenue, widening the tolerance spread, requires the involvement of the customers to validate the specifications. Often it is possible to modify the design of the product and to loosen its specifications without adversely affecting its users.

Acceptance sampling. Acceptance sampling refers to the procedures used to decide whether or not to accept product lots or batches based on the results of the inspection of samples drawn from the lots. Acceptance sampling techniques were originally developed for use by customers of manufactured products while inspecting lots delivered by their suppliers. These techniques are particularly well suited to situations where a decision on the quality level of product lots and their subsequent disposition needs to be made, but it is not economic or feasible to inspect the entire production output.

Sampling by attributes. The most commonly used acceptance sampling plans are based on inspection by attributes. In inspection by attributes, each product unit is classified as either possessing the required attributes and therefore conforming to specifications, or lacking at least one of the required attributes and therefore nonconforming or defective. The recipient of the product (that is, the consumer) specifies what constitutes the acceptable quality level (AQL) in terms of the fraction defective in a lot. Lots that do not exceed the AQL are accepted in their entirety. Those that do exceed it are considered not acceptable and are either returned to the producer or screened to remove all defective units and only then accepted.

The sampling plan is defined in terms of the sample size (n) and the acceptance number (c). Randomly selected samples of n units are drawn from each lot and inspected. Only if the number of defective units found in the sample does not exceed c is the lot accepted. Because of the random variability inherent in the sampling process, two types of inference errors are possible: rejection of a lot of acceptable quality, and acceptance of a lot of unacceptable quality. The second type of error requires the customer to define what fraction defective constitutes unacceptable quality. This fraction is called the limiting quality level (LQL); and if expressed as a percentage, it may also be referred to as the lot tolerance percent defective (LTPD). The rejection of a lot with fraction defective not exceeding the AQL is called a type 1 error, and its probability of occurrence is called the producer risk (α). The acceptance of a lot with fraction defective greater than the LQL is called a type 2 error, and its associated probability is called the consumer risk (β).

The entire range of performance of a given sampling plan for various incoming quality levels is summarized by the operating characteristic curve (**Fig. 7**). The operating characteristic curve shows on the vertical axis the probability of accepting a lot as a function of its fraction defective. The probabilities are calculated with the appropriate probability distribution function.

Rectifying inspection. Rectifying inspection is the procedure of screening lots that were not accepted by the sampling inspection plan. Screening means 100% inspection of the lot and removal of all the nonconforming units, followed by acceptance of the remaining units in the lot. If it is necessary to maintain a

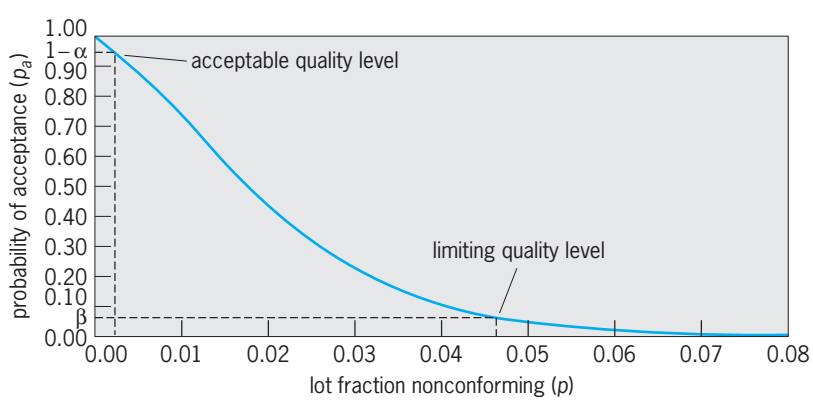

Fig. 7. Sampling by attributes. Operating characteristic curve for the sampling plan: sample size (n) = 100 and acceptance number (c) = 1.

constant lot size, the nonconforming units are replaced by conforming ones.

Under rectifying inspection, the average quality of the accepted lots varies, depending on the incoming fraction nonconforming. If the incoming fraction nonconforming is less than the AQL, most lots will be accepted according to the sampling plan, and the average fraction nonconforming will be low. If the incoming fraction nonconforming is significantly higher than the AQL, most lots will be rejected and screened, which, by removing all nonconforming units, lowers the average outgoing fraction nonconforming. Between these two extremes, the average outgoing fraction nonconforming will be higher. This fraction, which is called the average outgoing quality (AOQ), is a function of the incoming fraction nonconforming and of the given sampling plan.

Variables sampling plans. Variables sampling plans use the actual measurements taken on the units in the sample rather than the number of units that conform to specifications. Variables sampling plans are more complex to design and to administer, and the measurements are typically more costly than inspection by attributes. They also require the assumption that the observations are drawn from a normal distribution. On the positive side, variables sampling plans require a much smaller sample size than attributes sampling plans that provide the same protection in terms of AQL, producer risk, LQL, and consumer risk; and the actual measurements provide more information regarding how close the lot is to meeting specifications. However, if the lots come from a process that runs continuously, control charts are a preferable method for controlling quality.

Double and multiple sampling plans. Double sampling plans consist of two samples and two sets of acceptance and rejection numbers. The sample sizes need not be equal, but in many instances either they are equal or the second sample is twice as large as the first one. A double sampling plan is carried out as follows: (1) A sample of n_1 units is inspected. (2) If the number of nonconforming units does not exceed the first acceptance number (c_1), the lot is accepted. (3) If the number of nonconforming units exceeds the first rejection number (r_1), the lot is rejected. (4) Otherwise, a second sample of n_2 units is inspected. (5) If the total number of nonconforming units does not exceed the second acceptance number (c_2), the lot is accepted. (6) Otherwise, the lot is rejected. For the same protection in terms of AQL, producer risk, LQL, and consumer risk, double sampling plans result in a smaller average sampling number than single sampling plans. The reduction in inspection effort is even greater if the plan incorporates curtailing. Curtailing is the practice of stopping the inspection as soon as the number of nonconforming units found reaches the rejection number for the second sample (r_2), which is equal to $c_2 + 1$.

Multiple sampling plans are an extension of the same concept, involving a series of samples of equal size and their associated acceptance and rejection numbers. To prevent multiple plans from continuing indefinitely, they are designed such that a decision is made by the time the total number of units inspected is twice the sample size of a single sampling plan that offers the same protection to consumer and producer. Multiple sampling plans result in an even lower average sampling number, but they require more decisions and are more complex to design and administer.

Sequential sampling plans. In sequential sampling, units are inspected one at a time, and a decision to accept or reject the lot is made based on the cumulative number of nonconforming units found. The decision numbers are calculated by using the sequential probability ratio (SPR). If the sequential probability ratio is greater than or equal to $(1 - \beta)/\alpha$, the lot is rejected. If sequential probability ratio is smaller than or equal to $\beta/(1 - \alpha)$, the lot is accepted. If neither is true, an additional unit is inspected, and the sequential probability ratio is recalculated. Here again, provision should be made to truncate the test when the total number inspected reaches a certain value, typically a multiple of the sample size of an equivalent single sampling plan or of the average sampling number of an equivalent double or multiple sampling plan.

Chain sampling plans. Chain sampling is a procedure in which the decision to accept or reject a lot depends in part on the results of inspecting the preceding lots. When these procedures were originally developed, it was suggested that two acceptance numbers, 0 and 1, be used. If a number of previous lots (i) had no nonconforming units, the current lot will be accepted if the sample contains no more than one nonconforming unit. Otherwise, the lot will be accepted only if the sample has no nonconforming units. The design parameters for this type of plan are the sample size n and the number of immediately preceding lots. This type of plan increases the probability of acceptance when the fraction nonconforming is in general low, and reduces inspection costs by taking into account the immediate history of the product.

Continuous sampling plans. Continuous sampling plans are designed for processes that generate a continuous stream of product units that does not lend itself to grouping into lots or even meaningful samples. Continuous sampling plans entail periods of screening and periods of sampling. The procedure begins with 100% inspection and, after i consecutive conforming units are found, switches to inspection of a randomly selected fraction (f) of the units. When a nonconforming unit is found, it reverts to 100% inspection.

Standards. Following the recommendation from the American Society for Quality Control, the American National Standards Institute adopted a standard called Sampling Procedures and Tables for Inspection by Attributes. This system of tables is based on a standard used by military procurement, and it covers single, double, and multiple sampling for defectives (nonconforming units) and for defects (nonconformities).

The tables are indexed by AQL, lot size, and inspection level. The inspection level determines the

relationship between the sample size and the lot size, which affects the probability of rejecting lots with quality worse than the AQL. There are three general inspection levels: normal, tightened, and reduced. Once the AQL has been agreed upon by the consumer and the supplier, and the sampling plans have been selected based on the number of units in the lots submitted for acceptance, inspection begins at the normal level. If two or more lots out of any five consecutive lots are rejected, inspection switches to the tightened level, which uses a larger sample size and has a higher probability of rejection. If five consecutive lots are accepted while on tightened inspection, the system allows return to normal inspection. If, while under normal inspection, ten consecutive lots are accepted, inspection can switch to the reduced inspection level, which uses a smaller sample size. If a lot is rejected while on reduced inspection, the system prescribes return to the normal inspection level.

This system of tables is intended for use to protect the consumer in situations where there is a continuing series of lots. A similar system of tables is available for acceptance sampling by variables. *See* CONTROL CHART; INSPECTION AND TESTING; STATISTICS.

Tzvi Raz

Bibliography. L. S. Aft, *Fundamentals of Industrial Quality Control*, CRC Press, Boca Raton, FL, 1998; J. Bentley, *Introduction to Reliability and Quality Engineering*, Addison-Wesley, Harlow, 1999; F. W. Breyfogle, *Statistical Methods for Testing, Development and Manufacturing*, 1992; J. W. Cortada and J. A. Woods, *The McGraw-Hill Encyclopedia of Quality Terms and Concepts*, McGraw-Hill, New York, 1995; J. R. Evans and W. M. Lindsay, *The Management and Control of Quality*, South-Western College Publishers, Cincinnati, 1999; V. E. Sower, M. J. Savoie, and S. Renick, *An Introduction to Quality Management and Engineering*, Prentice Hall, Upper Saddle River, NJ, 1999; D. Straker, *A Toolbook for Quality Improvement and Problem Solving*, Prentice Hall, London, 1995.

Quantitative chemical analysis

The branch of analytical chemistry that deals with the determination of the relative or absolute amounts of the various constituents of materials.

Methods

Methods may be broadly grouped as chemical or physical, classical or instrumental, general or selective, and simple or compound. Any given method is classifiable within each of these groups. All techniques require calibration with respect to some standard of known composition. Caution is necessary to prevent substances other than those being determined from falsely giving signals attributable to the sought-for substance, which is called the analyte. *See* CALIBRATION.

Scale of determination. Selection of a technique is based in part on the size of sample available, the quantity of analyte expected to be in the sample, the precision and accuracy of the technique, and the speed of analysis required. The size of the sample dictates whether a determination is macroscopic (>0.1 g), semimicroscopic (0.01–0.1 g), microscopic (0.001–0.01 g), or ultramicroscopic (<0.001 g). Expected analyte levels may be major constituent (>10%), minor constituent (0.1–10%), or trace constituent (<0.1%).

Chemical and physical methods. Chemical methods employ reactions between the analyte and some other species, followed by observation of some quantitative result. Physical methods employ the interaction of the species of interest with electromagnetic radiation, sound, heat, or some other entity with which chemical reaction does not occur. A chemical method for determining the acidity of a solution would be the use of a series of indicator compounds which change color at specific pH or acidity values. A physical method for the same determination would involve the use of a glass indicator electrode to measure the adsorption of hydrogen ion to the glass by sensing the electrical potential of the glass. *See* ACID-BASE INDICATOR; ION-SELECTIVE MEMBRANES AND ELECTRODES; PH.

Classical and instrumental methods. In a classical method, human senses are used to detect a signal related to the quantity of the analyte. In an instrumental method, mechanical (usually electronic) means are used to detect the signal. Iron may be determined classically by titration of Fe(II) with Ce(IV) by using Ferroin as a visual indicator of the equivalence point. Iron may be determined instrumentally by producing a cloud of iron atoms in an electrical plasma and observing the light emitted by the iron at specific wavelengths. The light intensity at these wavelengths is proportional to the amount of iron introduced into the plasma.

General and selective methods. General methods respond to a wide range of elements or compounds, whereas selective methods respond only to a limited range of species. Measurement of the thermal conductivity of a gas stream can indicate the presence of any analyte species in the stream, so long as the thermal conductivity of the analyte is different from that of the pure stream. For carrier streams of inert gases, this means that most molecular analytes will give rise to signals. Measurement of the ability of the gas to attach electrons (an electron capture detector) will indicate only the presence of highly electronegative species such as the halogens.

Simple and compound techniques. A simple technique is one in which a single measurement operation is performed. A compound technique requires several operations to complete the analysis. Measurement of pH is a simple technique. Chromatographic separation of a mixture followed by thermal conductivity detection of the presence of a species, followed in turn by mass spectrometric detection of the molecular weight and fragmentation of the species, is an example of a compound technique.

Quantitation. Direct measurement of a signal related to concentration or activity of the chemical

species of interest is the most intuitive approach. Generally, a linear relationship between signal (or its logarithm) and concentration (or its logarithm) is sought. The relationship between signal and concentration is called a working curve. The slope of the line describing the relationship is known as sensitivity. The smallest quantity which is measurably different from the absence of analyte is the detection limit. Although a linear working curve is the simplest form to use, nonlinear curves may still be employed (either graphically or with the use of computers).

Titration. This is the process by which an unknown quantity of analyte (generally in solution) is determined by adding to it a standard reagent with which it reacts in a definite and known proportion. A classical or instrumental means is provided to indicate when the standard reagent has consumed exactly the amount of analyte initially present. Each determination is performed with a reagent whose concentration is directly traceable to a primary standard, so that accuracy is frequently superior to that of other methods. *See* STOICHIOMETRY; TITRATION.

Standard additions. Sensitivity of a method may be related not only to the method and analyte of interest, but also to the presence of other species in the sample. The materials other than the analyte constitute a sample matrix. For example, seawater is a matrix quite different from distilled water because of the large amount of dissolved electrolyte present. If a signal can be derived from the analyte which is proportional to the amount of the analyte, but the sensitivity of the signal to concentration varies as a function of the sample matrix, the method of standard additions may be useful (see **illus.**). The signal from the analyte is measured, after which small, known quantities of analyte are successively added to the sample, and the signal remeasured. The sensitivity of the method can thus be obtained, and the initial, preaddition signal interpreted to give the amount of analyte (C_1) initially present by using the relationship shown in the equation below, where S_1 is the initial signal; C_1 the

$$C_1 = \frac{\Delta C}{\Delta S} \cdot S_1$$

initial quantity of analyte; ΔC the standard addition of analyte; and ΔS the change in the signal caused by the addition. The sensitivity is $\Delta C/\Delta S$.

This approach is applicable only in the absence of reagent blanks, that is, signals caused by the presence of analyte in the reagent used for the determination.

Internal standard. In many methods, aliquots of samples are introduced into the measurement instrument. The signal from the analyte may vary with sample uptake rate or volume. To compensate for such effects, an internal standard, or species other than the analyte, may be added to the analyte in a known concentration prior to determination. The signal due to the internal standard is measured simultaneously with the analyte signal. Variation of the signal of the internal standard is interpreted to indicate the variation in sample uptake, which should be the same for both analyte and internal standard. The ratio of analyte signal to internal standard signal is independent of sample uptake. Thus the ratio of analyte to internal standard signal is used to establish the working curve, rather than using analyte signal alone.

Null comparison. If the relationship between signal and analyte concentration is nonlinear, quantitation may require the use of null comparison. A signal is observed from the analyte and from a standard whose concentration can be adjusted in a known way. When the signal from the adjustable standard equals the signal from the analyte, the two have identical concentrations. This condition is called a null, as there is no detectable difference between the sample and reference signals.

Interferences. It is good laboratory practice to check every quantitative measurement for the influence of species other than the one being sought. For example, a glass electrode designed to sense hydrogen ion will also respond to high concentrations of sodium ion. The degree to which a given sensor responds to one species in preference to another is called the selectivity coefficient. If a general detector is desired, this coefficient ideally should be 1.0. If a species specific detector is desired, the coefficient should be infinite. A signal of unknown or general origin which appears to underlie the analyte signal is known as background. This quantity, together with that for the reagent blank, may be subtracted from the raw signal and thus be compensated. However, variations in their level may prevent reliable compensation, particularly when small quantities of analyte are to be determined. Alexander Scheeline

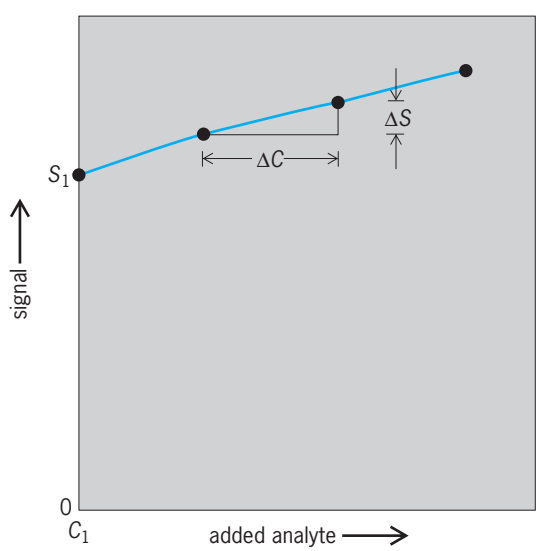

Working curve for the method of standard additions.

Analysis of Organic Materials

For determination of elements, functional groups, or molecules in organic materials, the type of analysis used is determined by the information required. If the total nitrogen content is desired, elemental analysis is used; if only amino nitrogen is desired, then only the amino group is determined. Organic analyses are made on a wide range of materials, from

pure compounds to mixtures such as blood and fertilizer. The identification of structures in conjunction with purity criteria, analyses of isotopic mixtures of organic molecules, and other examples of structure-related problems are increasingly important analytically. Proton nuclear magnetic resonance and other classes of spectrometric measurement are prominent for such purposes.

Determination of metals. This may be done in two ways. In the first method the sample is moistened with sulfuric acid and heated in oxygen to obtain metal oxides or sulfates, which are weighed. Some metals such as gold and silver are weighed as metals. The second method is based on destruction of the organic portion of the sample by heating with nitric and sulfuric acids, followed by determination of the metals by regular procedures. This method is required for metals which are easily volatilized, such as arsenic.

Carbon and hydrogen. These elements are usually determined simultaneously by burning a sample in a stream of oxygen to form carbon dioxide and water. This is done at 600°C (1100°F) in the presence of platinum as a catalyst. If nitrogen is present in the sample, the gas stream is passed through lead peroxide to remove the oxides of nitrogen. Silver wool adsorbs halogens and oxides of sulfur. The water formed is absorbed on a drying agent, such as magnesium perchlorate and from the increase in weight the amount of hydrogen in the sample can be calculated. Carbon dioxide is absorbed by an alkaline solid such as sodium hydroxide on asbestos fibers, and from the increase in weight the amount of carbon can be calculated.

Carbon alone is determined by wet combustion. The sample is heated in a mixture of sulfuric acid and potassium or silver dichromate. The carbon dioxide formed is determined from the amount of sodium hydroxide with which it combined, or by measuring directly the volume of the carbon dioxide. Reaction chromatography provides a newer basis for rapid carbon and hydrogen analyses on small samples of most classes or organic compounds. The weighed sample is reacted or combusted, and the reaction products are passed through suitable chromatographic columns. The products are measured from the areas beneath their chromatographic peaks.

Oxygen. Direct determination of oxygen in organic materials is accomplished by heating the sample in a stream of nitrogen to form water, oxides of carbon, and hydrocarbons. On passage of this mixture through graphite at 1150°C (2100°F), all the oxygen is converted to carbon monoxide. The carbon monoxide is oxidized with iodine pentoxide to form carbon dioxide and iodine. Either the iodine or the carbon dioxide may then be measured. This procedure is difficult to perform, so that oxygen in pure compounds is usually obtained by difference rather than by direct measurement.

Nitrogen. Two methods are used to determine nitrogen. In the Dumas method the sample is mixed with copper oxide and heated in a stream of pure carbon dioxide. The elemental nitrogen formed is collected over 50% potassium hydroxide solution, and its volume is measured. This procedure determines total nitrogen in most samples. In the Kjeldahl method the sample is heated in concentrated sulfuric acid containing a catalyst; this procedure converts the nitrogen to ammonia. On the addition of sodium hydroxide followed by boiling, ammonia is distilled into dilute boric acid solution, which is then titrated with acid. The choice of the catalyst is critical because different nitrogen compounds require a variety of catalysts. Esters of nitric and nitrous acids cannot be analyzed by this procedure.

Other elements. Sulfur, halogens, and phosphorus are determined by conversion to sulfuric acid, halogen acids, and phosphoric acids, which are measured by standard inorganic procedures. In the Carius method the sample is heated with fuming nitric acid in a sealed tube. In the combustion method the sample is burned in oxygen, oxides of sulfur are absorbed in neutral hydrogen peroxide, and halogens are absorbed in sodium bisulfite solution. In the bomb method for sulfur the sample is ignited with sodium peroxide and sugar.

Functional groups. Groups such as carboxyl, hydroxyl, nitro, and amide are determined by chemical reactions which are characteristic for each group. These reactions include neutralization, oxidationreduction, precipitation, condensation, and gas evolution. These methods are based on consumption of a reagent measured by direct titration, or in some cases by determination of excess reagent. For others, a product formed in the reaction is measured. Since most organic compounds are insoluble in water, organic solvents are used. These solvents enhance the acidity or basicity of some groups and also permit some reactions which are not possible in water. Many common organic functionalities can be used quantitatively to undergo reactions in which water is stoichiometrically consumed or evolved. The titrimetric measurement of such water with Karl Fischer reagent provides a useful basis for determining such functional groups.

Other methods. Many organic compounds absorb energy in the ultraviolet or infrared regions. Since this absorption is characteristic for each type of molecule, a direct analysis based on energy absorption is frequently possible. The variation in mass among hydrocarbons is the basis for the mass-spectrometric analysis of petroleum samples. These and other instrumental methods are based on the properties of molecules.

Physical methods of separation are commonly required to obtain molecules free from interferences. Normal distillation, vapor distillation, azeotropic distillation, ion exchange, chromatography, extraction of solids and liquids, and diffusion have a place in quantitative analysis.

In an analytical perspective, the concurrent separation, characterization, and quantitative measurement of the component molecules from very small samples results from such techniques as mass spectrometry or vapor chromatography. With preconcentration or larger samples, similar techniques detect

and measure trace components at ppm (parts per million) to ppb (parts per billion) levels. For complex samples, the preseparated species from a chromatographic column may be introduced at the inlet of a mass spectrometer.

Many enzyme-substrate combinations can be combined with selective-ion electrode responses for potentiometric measurement. For example, the urease-catalyzed hydrolysis of urea produces ammonium ion, which is sensed by an ammonium ion-selective electrode. If an enzyme-gel layer surrounds the electrode, its potential responds to external concentrations of urea. With other techniques, often spectrophotometric, the progress of enzymatic or other organic reactions may be followed with time, and provide a kinetic-based method of analysis. *See* ANALYTICAL CHEMISTRY.

For further information on common quantitative techniques *see* ACTIVATION ANALYSIS; CALORIMETRY; CHEMICAL MICROSCOPY; CHROMATOGRAPHY; COMBUSTION; ELECTRON SPECTROSCOPY; ELECTROPHORESIS; GAS CHROMATOGRAPHY; GEL PERMEATION CHROMATOGRAPHY; IMMUNOASSAY; IMMUNOELECTROPHORESIS; ION EXCHANGE; ISOTOPE DILUTION TECHNIQUES; KINETIC METHODS OF ANALYSIS; MASS SPECTROMETRY; NUCLEAR MAGNETIC RESONANCE (NMR); POLARIMETRIC ANALYSIS; SPECTROSCOPY; X-RAY SPECTROMETRY. Charles L. Rulfs

Bibliography. American Chemical Society, *Analytical Chemistry*, monthly; T. R. Crompton, *Comprehensive Organometallic Analysis*, 1987; J. S. Fritz and G. H. Schenk, Jr., *Quantitative Analytical Chemistry*, 5th ed., 1987; D. C. Harris, *Quantitative Chemical Analysis*, 7th ed., 2006; R. W. Ramette, *Chemical Equilibrium and Analysis*, 1981; S. Siggia and J. B. Hanna, *Quantitative Organic Analysis via Functional Groups*, 4th ed., 1979, reprint 1988; R. M. Silverstein et al., *Spectrometric Identification of Organic Compounds*, 7th ed., 2005; H. H. Willard et al., *Instrumental Methods of Analysis*, 7th ed., 1988.

Quantized electronic structure (QUEST)

A material that confines electrons in such a small space that their wavelike behavior becomes important and their properties are strongly modified by quantum-mechanical effects. Such structures occur in nature, as in the case of atoms, but can be synthesized artificially with great flexibility of design and applications. They have been fabricated most frequently with layered semiconductor materials. Generally, the confinement regions for electrons in these structures are 1–100 nanometers in size. The allowable energy levels, motion, and optical properties of the electrons are strongly affected by the quantum-mechanical effects. The structures are referred to as quantum wells, wires, and dots, depending on whether electrons are confined with respect to motion in one, two, or three dimensions (**Fig. 1**). Multiple closely spaced wells between which electrons can move by quantum-mechanical tunneling through intervening thin barrier-material layers are referred to as superlattices. *See* QUANTUM MECHANICS.

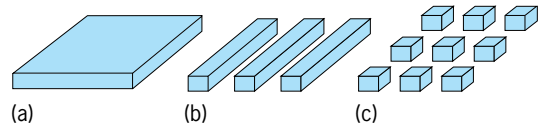

Fig. 1. Geometric configurations of (*a*) quantum well, (*b*) quantum wires, and (*c*) quantum dots.

The most frequently used fabrication technique for quantized electronic structures is epitaxial growth of thin single-crystal semiconductor layers by molecular-beam epitaxy or by chemical vapor growth techniques. These artificially synthesized quantum structures find major application in high-performance transistors such as the microwave high-electron-mobility transistor (HEMT), and in high-performance solid-state lasers such as the semiconductor quantum-well laser. They also have important scientific applications for the study of fundamental two-dimensional, one-dimensional, and zero-dimensional physics problems in which particles are confined so that they have free motion in only two, one, or zero directions. Chemically formed nanocrystals, carbon nanotubes, zeolite cage compounds, and carbon buckyball C_{60} molecules are also important quantized electronic structures.

Properties. Electrons that are confined in quantum structures can have only certain energies instead of the continuous distributions of energies that are available to particles in free space. For particles strongly confined to a single layer of thickness L and not moving parallel to the layer, the allowed energy levels relative to the bottom of the well are $E_n = h^2/(8m^*L^2)$, $4h^2/(8m^*L^2)$, $9h^2/(8m^*L^2)$, . . . , $n^2h^2/(8m^*L^2)$, where $n = 1, 2, 3, \ldots$ are integers, h is Planck's constant, and m^* is the effective mass with which electrons move in the layer. These energies correspond to electron states in which the electrons form standing waves of electron density with $n = 1, 2, 3, \ldots$ peaks in electron density (**Fig. 2**). These are the allowed states of the confined electrons. *See* NONRELATIVISTIC QUANTUM THEORY.

For electron effective mass $m^* = 0.1$ free electron masses (which is typical in semiconductors,

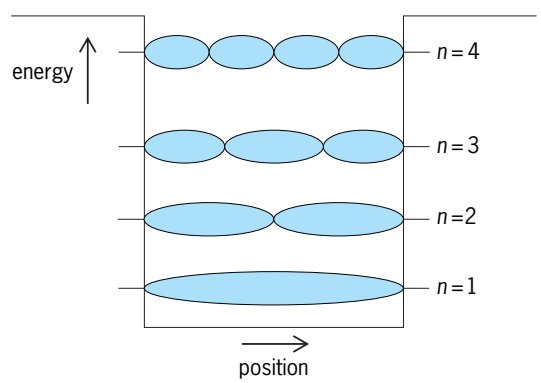

Fig. 2. Electron energy levels and electron standing waves for $n = 1$, 2, 3, and 4 quantum states in a quantum well.

where the electron inside the crystal behaves as if it were lighter than a free electron) and $L = 10$ nm, the allowed electron energies in electronvolts are $E_n = 0.0376n^2$ eV. Such quantized energy levels have clear signatures in the discrete lines of the absorption spectrum of light in high-quality gallium arsenide and other semiconductor quantum wells.

The potential energy for an electron in a quantum well and the potential energy for electrons in the barrier (Fig. 2) are basically determined by the electron affinities of the materials from which the well and barrier layers are composed. The electron affinity is the difference between the energy of an electron at rest in vacuum and the energy of an electron at the bottom of the conduction band of a material. The quantum wells also provide confining potentials for electrons in the valence bands of quantized electronic structures. When electrons are removed from the valence band of a semiconductor quantum well, the resulting hole in the valence-band quantum well leads to quantized hole phenomena that are directly analogous to all of the effects discussed above. *See* BAND THEORY OF SOLIDS; ELECTRON AFFINITY; HOLE STATES IN SOLIDS.

Potential well-shape profiles other than square or rectangular can be produced by growth of mixed alloys of the constituent semiconductors. A particularly basic example is the parabolic quantum well in which the confining potential for electrons varies as the square of the distance from the center of the well and thus closely mimics a simple harmonic oscillator potential for electron confinement. Well shape can also be changed by the application of an electric field perpendicular to the quantum well. Potentials in quantized electronic structures can also be shaped with the electric fields of charged donor and acceptor ions, and form the basis for so-called nipi doping superlattices. *See* ACCEPTOR ATOM; ANHARMONIC OSCILLATOR; DONOR ATOM.

Quantum wires and quantum dots provide additional confinement of electrons in quantized electronic structures and further sharpening of the spectrum of allowed energy levels. Wires and dots can be formed in a thin quantum well layer by lithographic etching of the layer or by placement of metallic gate electrodes on top of the layer. Electrons passing through short quantum wires known as quantum point contacts have shown sharp quantization of the electrical conductance in units of e^2/h, where e is the electron charge. Closely spaced pairs of constrictions produce quantum-dot-like structures where electrons have shown unique Coulomb blockade behavior, in which the electrostatic energy needed to place an additional electron on a dot is clearly observable.

Fabrication. Quantized electronic structures are most frequently formed by epitaxial (that is, crystalline film) growth of layers of the silicon-germanium (group IV) system of semiconductors; in the phosphide, arsenide, and antimonide families of group III–V semiconductors; and in the telluride families of II–VI semiconductors. *See* SEMICONDUCTOR.

In the molecular-beam-epitaxy growth of quantized electronic structures, beams of the constituent atoms and molecules of the semiconductor structures are formed by thermal evaporation of elemental sources of the atoms and molecules in ultrahigh vacuum. In the chemical-vapor-based growth techniques, by contrast, gases of compounds containing the desired elements react on hot substrate surfaces to produce the quantum structures. In each case, films can be grown at rates of the order of one monolayer of deposition per second. Film surfaces can be produced that are smooth at the scale of plus or minus one monolayer. Smoothness is important in quantized electronic structures because smooth films have sharper energy levels and stronger quantum behavior. *See* CRYSTAL GROWTH; SEMICONDUCTOR HETEROSTRUCTURES.

Film thickness is controllable with monolayer precision by molecular-beam shutter operation in molecular-beam epitaxy and by valve openings and closures in the chemical-vapor techniques. Thickness control is also of importance because it determines the magnitude of quantized energies and the operating frequencies or wavelengths for many devices. Fortunately, it is possible to grow strained layers with different lattice constants on top of each other, and strain is frequently usable to fine-tune the properties of the quantum structures.

Quantum dot and wire formation pose a special challenge because lateral dimensional control during epitaxy is less precise than thickness control. Ingenious ways that have been demonstrated to achieve small lateral quantum confinement structures are overgrowth of cleaved edges of quantum wells, growth of repeated alternate submonolayer coverages on regularly stepped surfaces, and nucleation of small islands during growth of thin epitaxial layers of different lattice spacing than the substrate material. Chemical means of producing quantum-size particles have also rapidly developed, particularly the growth of sulfide and selenide nanocrystals and the growth of cage compounds that form templates for particle or wire formation. The buckyball family of closed cage molecules of symmetrically arranged carbon atoms has been extended to carbon nanotubes. These crystalline nanotubes can be millions of times longer than their diameters. They can be formed by vaporization of graphite with lasers or arcs and also by chemical vapor deposition. *See* FULLERENE; ZEOLITE.

Applications. Both the optical and the electronic applications of quantized electronic structures are important.

Optical applications. The optical applications are based on the interactions between light and electrons in the quantum structures. The absorption of a photon by an electron in a quantum well raises the electron from occupied quantum states to unoccupied quantum states. Electrons and holes in quantum wells may also recombine, with the resultant emission of photons from the quantized electronic structure as the electron drops from a higher state to a lower state. *See* ELECTRON-HOLE RECOMBINATION.

The photon emission is the basis for quantum-well semiconductor lasers, which have widespread applications in optical fiber communications and compact disk and laser disk optical recording. Quantum-well lasers operate by electrically injecting or pumping electrons into the lowest-conduction-band ($n = 1$) quantum-well state, where they recombine with holes in the highest-valence-band ($n = 1$) quantum-well state (that is, the electrons drop to an empty $n = 1$ valence-band state; **Fig. 3*a***), producing the emission of photons. These photons stimulate further photon emission and produce high-efficiency lasing. *See* COMPACT DISK; LASER; OPTICAL COMMUNICATIONS; OPTICAL RECORDING.

The photon absorption is the basis for quantum-well photodetectors and light modulators. In the quantum-well infrared photodetector an electron is promoted from lower (say, $n = 1$) to higher (say, $n = 2$) conduction band quantum-well states (Fig. 3*b*) by absorption of an infrared photon. An electron in the higher state can travel more freely across the barriers, enabling it to escape from the well and be collected in a detector circuit. Changes in quantum-well shapes produced by externally applied electric fields can change the absorption wavelengths for light in a quantized electronic structure. The shift in optical absorption wavelength with electric field is known as the quantum-confined Stark effect. It forms the basis for semiconductor light modulators and semiconductor optical logic devices. *See* OPTICAL DETECTORS; OPTICAL MODULATORS; STARK EFFECT.

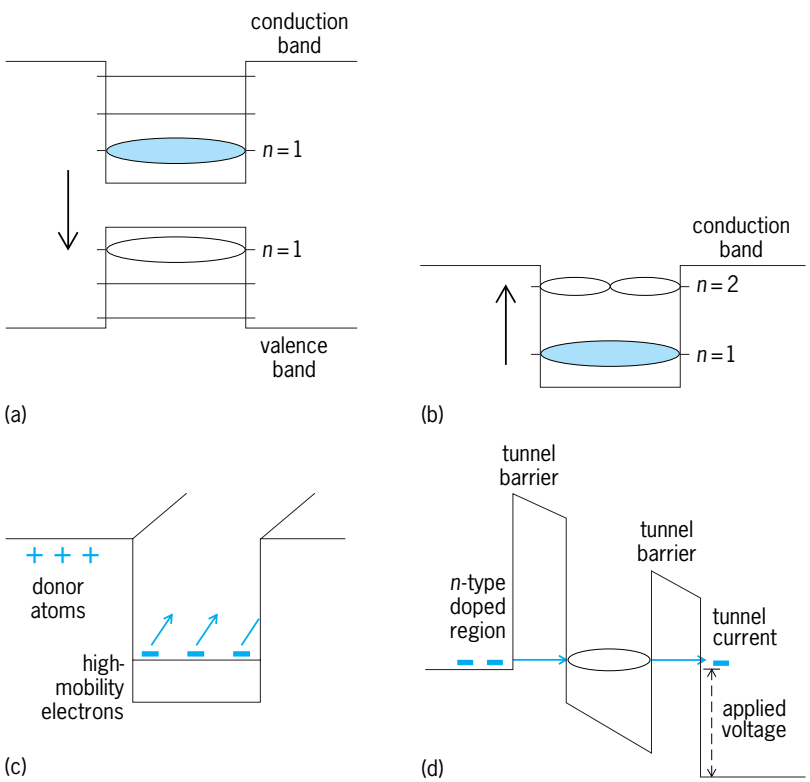

Fig. 3. Principles of operation of quantum-well devices. (*a*) Quantum-well laser. (*b*) Quantum-well infrared detector. (*c*) High-electron-mobility transistor (HEMT or MODFET). Electrons in the quantum well that came from donor atoms in the barrier are free to move with high mobility in the direction perpendicular to the page. (*d*) Resonant tunneling device.

Electrical applications. Modulation doping is a special way of introducing electrons into quantum wells for electrical applications. The electrons come from donor atoms lying in adjacent barrier layers (Fig. 3*c*). Modulation doping is distinguished from conventional uniform doping in that it produces carriers in the quantum well without introducing impurity dopant atoms into the well. Since there are no impurity atoms to collide with in the well, electrons there are free to move with high mobility along the quantum-well layer. Resistance to electric current flow is thus much reduced relative to electrical resistance in conventional semiconductors. This enhances the low-noise and high-speed applications of quantum wells and is the basis of the high-electron-mobility transistor (HEMT), which is also known as the modulation-doped field-effect transistor (MODFET). HEMTs are widely used in microwave receivers for direct reception of satellite television broadcasts. *See* TRANSISTOR.

Modulation-doped structures have also shown a remarkable quantization of the Hall resistance, known as the quantum Hall effect. In this quantum effect, the lateral voltage induced in a flowing electron gas by a magnetic field, divided by the longitudinal current along the two-dimensional electron gas, is precisely quantized with values $R_H = h/(ne^2) = 25{,}812.81/n$ ohms, where $n = 1, 2, 3, \ldots$. This ratio is so precisely determined and measurable that it has been adopted as the basis for the standard of resistance and the ohm. It was discovered in the two-dimensional electron gas formed at the metal-oxide interface of a metal-oxide-semiconductor field-effect transistor (MOSFET). At very low temperatures (less than 1K) and high magnetic fields (greater than 1 tesla), the electrons in the two-dimensional electron gas in high-mobility materials can form ordered states that give rise to fractional quantization of the Hall resistance. The electrons then behave as if they had fractional charges. *See* ELECTRICAL UNITS AND STANDARDS; HALL EFFECT; RESISTANCE MEASUREMENT.

Electrical conductivity in carbon nanotubes occurs without doping and results from the absence of any energy gap in the electronic energy band structure of the nanotubes and the presence of allowed states at the Fermi energy. Individual nanotubes can be electrically contacted. Simple quantum wire transistors displaying quantized electron motion have been formed from single nanotubes.

Quantum-mechanical tunneling is another important property of quantized electronic structures. Tunneling of electrons through thin barrier layers between quantum wells is a purely quantum-mechanical effect without any real analog in classical physics or classical mechanics. It results from the fact that electrons have wavelike properties and that the particle waves can penetrate into the barrier layers. This produces a substantial probability that the particle wave can penetrate entirely through a barrier layer and emerge as a propagating particle on the opposite side of the barrier. The penetration probability has an exponential drop-off with barrier

thickness. The tunneling is greatest for low barriers and thin barriers.

This effect finds application in resonant tunnel devices, which can show strong negative resistance in their electrical properties. In such a device (Fig. 3*d*), electrons from an *n*-type doped region penetrate the barrier layers of a quantum well by tunneling. The tunneling current is greatest when the tunneling electrons are at the same energy as the quantum-well energy. The tunneling current actually drops at higher applied voltages, where the incident electrons are no longer at the same energy as the quantum-state energy, thus producing the negative resistance characteristic of the resonant tunneling diode. *See* ARTIFICIALLY LAYERED STRUCTURES; NANOSTRUCTURE; NEGATIVE-RESISTANCE CIRCUITS; RESONANCE (QUANTUM MECHANICS); TUNNELING IN SOLIDS.

Arthur C. Gossard

Bibliography. J. H. Davies, *The Physics of Low-Dimensional Semiconductors: An Introduction*, Cambridge University Press, 1998; A. C. Gossard (ed.), *Epitaxial Microstructures*, Academic Press, 1994; M. J. Kelly, *Low-Dimensional Semiconductors: Materials, Physics, Technology, Devices*, Clarendon Press, 1995; A. Shik, *Quantum Wells: Physics and Electronics of Two-Dimensional Systems*, World Scientific, 1997.

Quantized vortices

A type of flow pattern exhibited by superfluids, such as liquid ^{4}He below 2.17 K. The term vortex designates the familiar whirlpool pattern where the fluid moves circularly around a central line and the velocity diminishes inversely proportionally to the distance from the center. The strength of a vortex is determined by the circulation k, which is the line integral of the velocity around any path enclosing the central line. For an ordinary vortex, k can possess any value; for a superfluid vortex, k is restricted to a quantized multiple of Planck's constant h divided by m, the mass of the helium atom. Hence the expression quantized vortex line. Although h and m are microscopic (that is, atomic) quantities, their ratio h/m is rather large, being equal to 10^{-3} cm$^2 \cdot$ s^{-1}. *See* VORTEX.

Theory. The possible existence of quantized vortices was suggested on theoretical grounds as early as 1949. A superfluid is believed to be characterized by a macroscopic (that is, large-scale) quantum-mechanical wave function ψ. This wave function locks the superfluid into a coherent state. Considerations on how a uniform flow might affect ψ lead to realization that the superfluid velocity v_s is intimately connected to the manner in which ψ might vary in space. From elementary considerations, one then deduces that quantized vortices should exist.

Since the velocity around the vortex increases without limit as the center is approached, the superfluid density and thus ψ must vanish at the center in order to avoid an infinite energy. Thus the central core of the vortex marks the zeros, or nodal lines, in the macroscopic wave function. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Production. Quantized vortex lines are usually produced by rotating a vessel containing superfluid helium. At very low rotation speeds, no vortices exist: the superfluid remains at rest while the vessel rotates. At a certain speed the first vortex appears and corresponds to the first excited rotational state of the system. If the container continues to accelerate, additional quantized vortices will appear. At any given speed the vortices form a regular array which rotates with the vessel.

Experimental investigation. Quantized vortex lines were first detected in the mid-1950s by their influences on superfluid thermal waves traveling across the lines. Shortly thereafter the quantum of circulation was directly measured by studying the precessional motion of a fine vibrating wire immersed in rotating superfluid helium. If no circulation is present, the plane of vibration of a circular-cross-section wire would remain stationary. In the presence of circulation, the vibration plane precesses at a rate determined by the circulation. It has been found that the stable values of circulation are close to h/m.

Mechanical properties. The mechanical properties of vortices have been studied by detecting various types of bending waves in the vortices. In one early experiment, well-defined normal modes were observed, which were consistent with predictions. Since the effective tension of a single vortex is very small, these mechanical measurements have usually been concerned with systems containing many lines. However, some progress has been made in measuring the very small angular momentum of a single line.

Use of ions to probe vortices. In the late 1950s it was discovered that electrons in liquid helium form tiny charged bubbles which can become trapped on the vortex core but can move quite freely along the line. These electron bubbles (often referred to as ions) have been one of the most useful probes of quantized vortices. The most accurate quantitative proof of the quantized nature of the vortices emerged from the discovery that at low temperatures (below 0.5 K) a rapidly moving ion can create a vortex ring. These rings were found to move precisely like classical smoke rings except that the circulation κ was equal to h/m to within 1%.

Detection of single vortex lines. Researchers have been able to use ions to detect single quantized vortex lines. In one experiment the vortices in a rotating vessel were charged with ions, and subsequently the amount of trapped charge was measured. As the vessel slowly accelerates from rest, the trapped charge is seen to increase as a stepwise function of rotation speed, thus proving that the superfluid comes into rotation in a sequence of quantum steps.

In an extension of the ion technique, the trapped ions are pulled out at the top of the vortex lines, accelerated, and focused onto a phosphor screen. The pattern of light thus produced on the phosphor

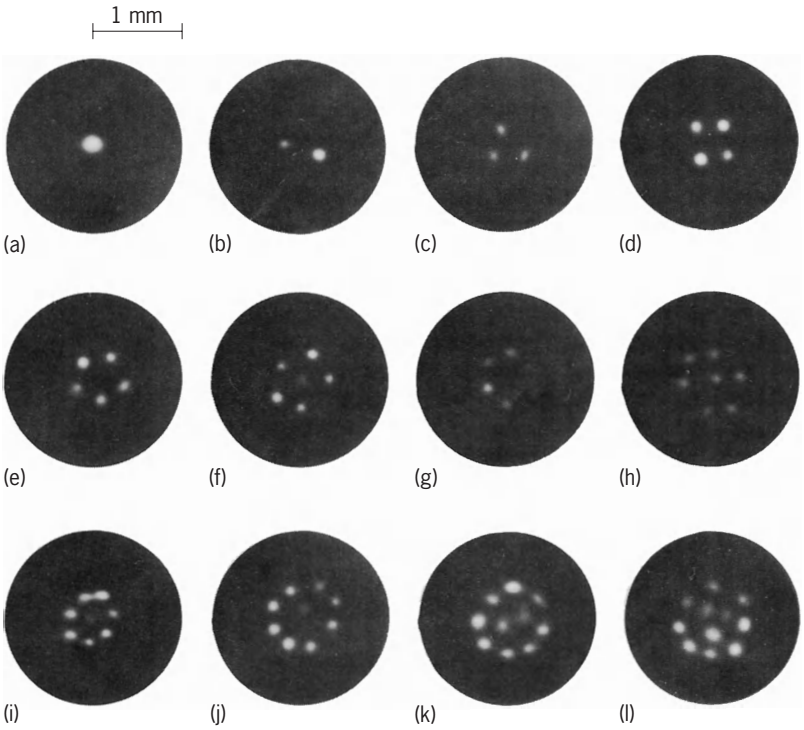

Fig. 1. Stationary configurations of vortices which appear when a cylindrical container of superfluid ^{4}He is rotated about its axis. As the rotation speed increases from *a* to *l*, more vortices appear and the patterns become more complex. (*From E. J. Yarmchuk, M. J. V. Gordon, and R. E. Packard, Observation of stationary vortex arrays in rotating superfluid helium, Phys. Rev. Lett., 43:214–217, 1979*)

is a map of the position of the vortices where they contact the liquid meniscus (**Fig. 1**). The arrays of vortices agree well with theoretical predictions.

Vortices in helium-3. A second isotope of helium, ^{3}He, also exists in superfluid states but only at temperatures below 2×10^{-3} K, 150,000 times colder than room temperature. Superfluid ^{3}He is associated with a pairing of two atoms, a mechanism similar to that responsible for the phenomenon of superconductivity. There have been several varieties of quantized vortices predicted for the ^{3}He superfluids. The simplest type of vortex is similar to that in ^{4}He except that the circulation quantum is expected to equal $h/(2m_3)$, where m_3 is the mass of the ^{3}He atom and the factor 2 arises from the pairing of the atoms. In a direct measurement of the superfluid flow around a fine wire, the circulation was found to be quantized, as expected, in units of $h/(2m_3)$ [**Fig. 2**]. *See* SUPERCONDUCTIVITY.

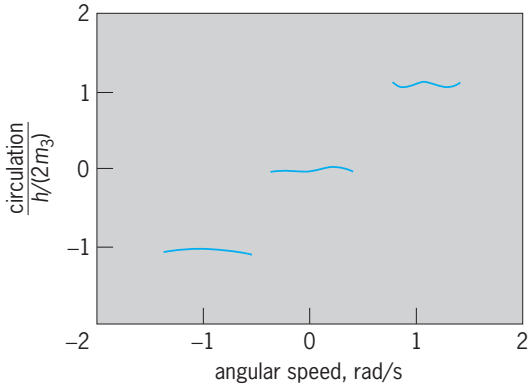

Fig. 2. Measured fluid circulation around a thin wire immersed in superfluid ^{3}He. The horizontal axis of the graph is the angular speed of the rotating refrigerator. There are three stables values of fluid circulation: $h/(2m_3)$, 0, and $-h/(2m_3)$. (*After J. C. Davis et al., Observation of quantized circulation in superfluid ^{3}He-B, Phys. Rev. Lett., 66:329–332, 1991*)

Experiments using magnetic resonance techniques have confirmed the prediction that several varieties of ^{3}He vortices exist. Which particular vortex type exists depends on the temperature and pressure within the superfluid. Using these nuclear magnetism techniques, the appearance of individual ^{3}He vortices has been detected in a small vessel whose rotational speed is steadily changing. *See* MAGNETIC RESONANCE.

Vortices in neutron stars. Numerous sources of very regular, pulsed electromagnetic radiation have been detected. These so-called pulsars are believed to be rotating neutron stars. In addition to the two helium isotopes, it is believed that the interiors of these neutron stars are also superfluid. Since these pulsars are rotating objects, the neutron star's interior is expected to be filled with an array of quantized vortices, and some features of the neutron star's motion have been explained in terms of the related motion in the vortices. A laboratory-sized rotating superfluid ^{4}He analog of a neutron star has been developed, and exhibits features similar to those seen in pulsar signals, supporting the idea that pulsars are indeed rotating superfluid neutron stars. *See* NEUTRON STAR; PULSAR.

Richard E. Packard

Bibliography. M. Krusius, Vortices in rotating ^{3}He, the vortices of superfluid ^{3}He, *J. Low Temp. Phys.*, 91:233–273, 1993; D. R. Tilley and J. Tilley, *Superfluidity and Superconductivity*, 3d ed., 1990; E. J. Yarmchuk and R. E. Packard, Photographic studies of quantized vortex lines, *J. Low Temp. Phys.*, 46:479–515, 1982.

Quantum (physics)

A term characterizing an excitation in a wave or field, connoting fundamental particlelike properties such as energy or mass, momentum, and angular momentum for this excitation. In general, any field or wave equation that is quantized, including systems already treated in quantum mechanics that are second-quantized, leads to a particle interpretation for the excitations which are called quanta of the field. This term historically was first applied to indivisible amounts of electromagnetic, or light, energy usually referred to as photons.

The photon, or quantum of the electromagnetic field, is a massless particle, best interpreted as such by quantizing Maxwell's equations. Analogously, the electron can be said to be the quantum of the Dirac field through second quantization of the Dirac equation, which also leads to the prediction of the existence of the positron as another quantum of this field

with the same mass but with a charge opposite to that of the electron. In similar fashion, quantization of the gravitational field equations suggests the existence of the graviton, a particle which is the quantum of gravitational field. The pi meson or pion was theoretically predicted as the quantum of the nuclear force field by H. Yukawa; other particles associated with nuclear forces have since been found. Another quantum is the quantized lattice vibration, or phonon, which can be interpreted as a quantized sound wave since it travels through a quantum solid or fluid, or through nuclear matter, in the same manner as sound through air. The energy of a phonon is $h\nu$, h being Planck's constant (called the quantum of action in "old" quantum mechanics) and ν the frequency of oscillation of the matter through which the wave propagates. The energy of a photon is also $h\nu$, with h being Planck's constant and ν the frequency of the light wave.

The use of the word quantum as an adjective (quantum mechanics, quantum electrodynamics) implies that the particular subject is to be treated according to the modern rules that have evolved for quantized systems. *See* ELEMENTARY PARTICLE; GRAVITATION; GRAVITON; MAXWELL'S EQUATIONS; MESON; NONRELATIVISTIC QUANTUM THEORY; PHONON; PHOTON; QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; QUANTUM MECHANICS. Kenneth E. Lassila

Quantum acoustics

The investigation of the effects of the laws of quantum mechanics on the propagation and absorption of sound. At the present stage of development of physical science, quantum mechanics is the most fundamental theory of physical phenomena. However, for many applications in the everyday world, a sufficiently accurate description of nature is provided by classical mechanics. Quantum acoustics refers to acoustic experiments that are carried out under conditions such that the results can be understood only in terms of quantum theory. As a general tendency, quantum effects become more important in acoustic experiments that are performed with higher-frequency sound waves or that are carried out at lower temperatures. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Sound and phonons. To understand the quantum nature of sound, it is valuable to consider the origins of the quantum theory of light. According to the classical theory of electromagnetism, the energy of a light wave can have any value. However, experiments in the latter part of the nineteenth century showed that the classical theory of electromagnetism combined with the laws of statistical mechanics could not explain the spectrum of light emitted by a heated surface. To reconcile theory and experiment, Max Planck in 1900 proposed that the energy of a light wave is quantized. This means that only certain values of the energy of the wave are allowed. These allowed values are given by Eq. (1),

$$E = n\hbar\omega \tag{1}$$

where $\hbar$ is Planck's constant divided by 2π, ω is the angular frequency of the light wave, and n is an integer. The modern interpretation of this formula describes this energy in terms of elementary quanta called photons. Each photon has an energy $\hbar\omega$, and the wave is made up of n photons. It was later realized that this same quantization of energy should apply to sound waves. These fundamental units of sound energy were given the name phonons. *See* HEAT RADIATION; LIGHT; PHONON; PHOTON.

Complementarity principle. According to quantum theory, a sound wave should be considered to be a beam of phonon quanta. However, for many practical applications in acoustics, such as the design of musical instruments or the construction of concert halls, the classical theory of sound holds extremely well and it is unnecessary to describe the sound wave in terms of phonons. This is an example of the complementarity principle, which states that under suitable conditions the predictions of classical and quantum theory agree, even though the nature of the two theories may first seem to be quite different.

Size of quanta. For a sound wave of given frequency, the energy $\hbar\omega$ of a single phonon can be calculated. The numerical value of Planck's constant is 6.6×10^{-34} J · s. Thus, for a sound wave of frequency 1000 Hz, the energy of a sound quantum is only 6.6×10^{-31} J. Hence, a wave of this frequency and of total energy as small as 10^{-12} J still contains more than 10^{18} quanta. Consequently, for practical purposes the restriction that the energy of the wave be quantized in units of the phonon energy is not important. However, because the amount of energy in each quantum increases with increasing frequency, the effects of the quantization are much more likely to be significant when the acoustic wave has a very high frequency. With modern techniques it is possible to generate and detect sound waves whose frequency extends far above the audible range, and it is in experiments conducted with these methods that the quantum nature of sound has been convincingly demonstrated. *See* ULTRASONICS.

Phonoatomic effect. One of the earliest experiments that confirmed the idea that light consists of photons was the photoelectric effect. In this effect the energy of light quanta is used to eject electrons from the surface of a metal. Only if the frequency of the light is sufficiently high do the photons have enough energy to knock the electrons out of the metal. The detection of this threshold frequency is thus strong evidence for quantization of energy. An analogous experiment has been performed with sound (**Fig. 1**). In the experiment, sound is generated by a source S in a vessel of liquid helium maintained at very low temperature, less than 0.1 K above absolute zero. At this temperature, even very high-frequency sound waves will travel a long distance through the liquid without appreciable loss of energy. In liquid helium, each atom is bound to the other atoms in

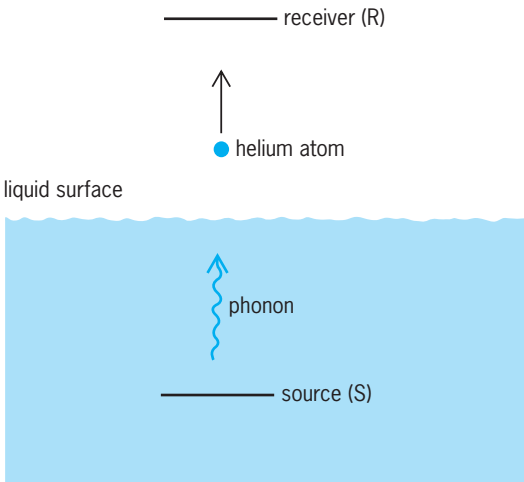

Fig. 1. Experimental demonstration of the phono-atomic effect in liquid helium.

the liquid by an energy that is unusually small compared to that of other liquids. This binding energy E_{bind} is such that $E_{\text{bind}}/k_B = 7.1$ K (where k_B is Boltzmann's constant). For sound of sufficiently high frequency, the sound quanta when they arrive at the surface of the liquid have sufficient energy to knock helium atoms out of the liquid. These atoms can be detected by a suitable receiver R placed above the surface of the liquid. The energy E_{kinetic} of an ejected atom is equal to the energy of a phonon minus the energy E_{bind}, and so the velocity of the atom is given by Eq. (2), where m is the mass of a helium atom.

$$v_{\text{atom}} = \sqrt{\frac{2(\hbar\omega - E_{\text{bind}})}{m}} \tag{2}$$

The time of arrival of the ejected atoms at the receiver R is determined by this velocity and hence is dependent on the frequency of the sound wave. Experiments have confirmed that the helium atoms arrive at the time expected based on this theory. *See* PHOTOEMISSION.

Phonon-phonon interactions. Sound waves in solids are attenuated during their propagation by a wide variety of physical processes. In many materials the most important mechanisms are related to impurities or defects in the solid, such as cracks, grain boundaries, or dislocations. Even when sound travels through a perfect crystal containing no defects, it is found that a measurable attenuation still occurs. In insulating crystals, where there are no free electrons, this attenuation is due to an interaction between the sound wave and the random thermal vibrations of the atoms in the solid. These random vibrations, which constitute the heat energy of the solid, are also quantized and are called thermal phonons. The attenuation of the sound wave can be attributed to collisions with the thermal phonons in which some of the sound quanta are scattered out of the sound beam. This mechanism is referred to as phonon-phonon scattering. *See* CRYSTAL DEFECTS; LATTICE VIBRATIONS; SOUND ABSORPTION.

In a linear elastic solid the elastic stress is exactly proportional to the strain, a phenomenon that is called Hooke's law. If Hooke's law holds, the presence of one wave does not affect the propagation of another, and so there are no interactions between phonons. Phonon-phonon scattering occurs because real solids always exhibit some deviations from linear elastic behavior. This nonlinearity is called anharmonicity. *See* HOOKE'S LAW; NONLINEAR ACOUSTICS.

The most important type of collision (**Fig. 2**) between phonons is the three-phonon process. In this interaction a sound phonon collides with a thermal phonon, and one new phonon is produced. The collision has to conserve energy, and so Eq. (3) must

$$E_{\text{sound}} + E_1 = E_2 \tag{3}$$

be satisfied, where E_{sound}, E_1, and E_2 are the energies of the sound phonon and the two other phonons. Since it is also necessary to conserve momentum, Eq. (4) must be satisfied, where $\vec{p}_{\text{sound}}$, $\vec{p}_1$, and $\vec{p}_2$ are

$$\vec{p}_{\text{sound}} + \vec{p}_1 = \vec{p}_2 \tag{4}$$

the momentum vectors of the three phonons. The momentum of a phonon is related to its wave vector $\vec{k}$ by Eq. (5).

$$\vec{p} = \hbar\vec{k} \tag{5}$$

See CONSERVATION OF ENERGY; CONSERVATION OF MOMENTUM.

The attenuation of a sound wave that occurs as a result of these collisions increases with increasing frequency of the sound wave. It also increases rapidly with temperature because the higher the temperature the more thermal phonons are present in a solid.

Even when the temperature is very low and there are very few thermal phonons, anharmonicity can still give rise to an attenuation of a sound wave. This is because the sound phonons can spontaneously decay into phonons of lower frequencies (Fig. 2*b*). In

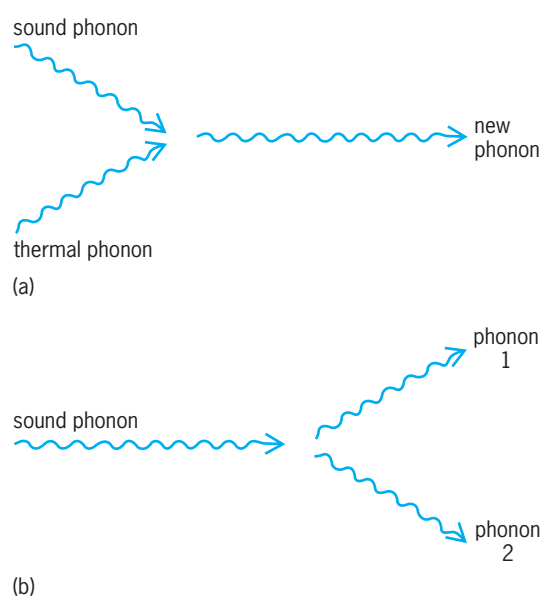

Fig. 2. Phonon-phonon processes. (*a*) A sound phonon collides with a thermal phonon to produce another phonon. (*b*) A sound phonon decays into phonons 1 and 2.

this process the conditions of conservation of energy and momentum are given by Eqs. (6) and (7), with

$$E_{\text{sound}} = E_1 + E_2 \tag{6}$$

$$\vec{p}_{\text{sound}} = \vec{p}_1 + \vec{p}_2 \tag{7}$$

E_1 and E_2 the energies, and $\vec{p}_1$ and $\vec{p}_2$ the momenta of the phonons produced in the decay. The rate at which this decay occurs is proportional to the frequency of the sound wave to the fifth power, and so the attenuation is important only for sound waves of very high frequency. If the number of phonons is sufficiently large, it is possible under some circumstances for there to be a large buildup in the population of some of the decay phonons. This process is called parametric amplification. *See* PARAMETRIC AMPLIFIER.

Sound attenuation by electrons. In a metal a certain number of electrons from each atom are able to move freely through the crystal. The sound quanta can be absorbed by these electrons (**Fig. 3**). This process also leads to an attenuation of the sound wave. Just as for the phonon-phonon interaction, energy and momentum must be conserved in the process. It can be shown that because of these conditions the sound phonons are absorbed primarily by electrons that are moving at right angles to the direction of propagation of the sound wave.

Resonant absorption of phonons. Many solids contain defects or impurities that can absorb energy from a sound wave. The simplest example is a paramagnetic impurity in a crystal. When an atom with an unpaired atomic electron is placed in a magnetic field, the atom can exist in one of two different energy states; these states correspond to different orientations of the spin of the electron with respect to the direction of the magnetic field. If the energy of the phonons in a sound wave is the same as the difference in energy between these two states, the phonons can be absorbed by the impurity atom. In this process the atom makes a transition from the lower energy state to the upper. This process is called resonance absorption because there must be a precise matching of the sound energy and the spacing between the two states.

This same general type of phenomenon can occur in several other situations. Some crystals are composed of atoms that have nuclei that possess a magnetic moment. In a magnetic field these nuclei have several equally spaced energy states. The sound phonons can be absorbed by these nuclei if the phonon energy is the same as the difference in energy of the different nuclear states. For magnetic fields that can be readily produced in the laboratory, this type of absorption occurs for phonons of frequency in the range of roughly 5 to 100 MHz. The attenuation is quite small and can be detected only in specially designed experiments. In semiconducting crystals that have been doped with donor or acceptor impurities, the electrons associated with these impurities often have several different energy states even in the absence of a magnetic field. Sound phonons can also be resonantly absorbed by these electrons.

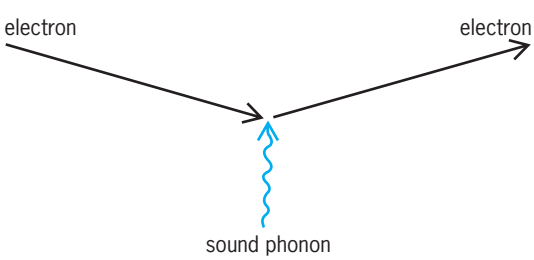

Fig. 3. Absorption of a sound phonon by an electron in a metal.

In the above examples of resonance absorption, the attenuation of the sound wave occurs only if the sound wave has one particular frequency. An especially interesting example of resonance absorption occurs in glasses and other amorphous materials. In these solids there is no regular crystal structure, and so every atom is in a slightly different environment. It is believed that in all glasses there are a small number of atoms that interact with their neighbors in such a way that each one of the atoms can exist in two states that have nearly the same energy. Such an atom can absorb energy from a sound wave if the energy of the phonon quanta matches the difference in the energies of the two states of the atom. However, because of the random nature of the glass this energy difference is different for different atoms. As a result, regardless of the frequency of the sound wave, there will always be some atoms in the glass that are able to absorb the sound phonons. Thus, resonance absorption in glasses occurs over a broad band of frequencies. *See* AMORPHOUS SOLID; GLASS.

Phonon maser. Resonance absorption occurs when a paramagnetic impurity makes a transition from a low energy state to a higher state and a sound phonon is absorbed. If by some means it is possible to prepare a crystal in which the majority of the impurities are in the higher of their two states, the sound phonons will induce transitions from the higher state to the lower state. New phonons will be emitted, and the sound wave will be amplified. This process is analogous to the amplification of microwave electromagnetic radiation that occurs in a maser. *See* MASER.

Humphrey J. Maris

Bibliography. R. T. Beyer and S. V. Letcher, *Physical Ultrasonics*, 1969; M. J. Crocker (ed.), *Handbook of Acoustics*, Wiley, 1998; S. Hunklinger et al. (eds.), *Phonons '89*, 1990; H. J. Maris, Phonon interactions in liquid helium, *Rev. Mod. Phys.*, 49:341–359, 1977; M. Meissner et al. (eds.), *Phonon Scattering in Condensed Matter VII*, 1993; R. Peierls, *Quantum Theory of Solids*, 1955, paperback 2001; W. A. Phillips (ed.), *Amorphous Solids: Low-Temperature Properties*, Springer-Verlag, Berlin, 1981; M. N. Wybourne et al. (eds.), Phonons 2001, *Physica B*, vols. 316–317, 2002.

Quantum anomalies

Phenomena that arise when a quantity that vanishes according to the dynamical rules of classical physics acquires a finite value when quantum rules are used.

For example, the classical Poisson bracket for some entities may vanish; yet the corresponding quantum commutator may be nonzero—this is a commutator anomaly. Alternatively, the flow of some material current may satisfy a continuity equation by virtue of the classical equations of motion, indicating conservative flow; but upon quantization the continuity equation may fail and the flow may no longer be conservative in the quantum theory—this is an anomalous divergence (of the current in question). Since the forms of Poisson brackets and quantum commutators as well as the occurrence of continuity equations for currents are related to symmetries and conservation laws of the theory, quantum anomalies serve to break some symmetries and destroy some conservation laws of classical models. This violation of symmetry is not driven by explicit symmetry-breaking terms in the dynamical equations—rather the quantization procedure itself violates the classical symmetry. The mathematical reason for this phenomenon is that classical dynamics, involving a finite number of degrees of freedom, usually leads to a quantum theory on an infinite-dimensional vector space (Hilbert space), and this "infinity" gives rise to novel effects. *See* CANONICAL COORDINATES AND TRANSFORMATIONS; CONSERVATION LAWS (PHYSICS); EQUATION OF CONTINUITY; HILBERT SPACE; NONRELATIVISTIC QUANTUM THEORY; SYMMETRY LAWS (PHYSICS).

Example from quantum mechanics. While most instances of anomalies arise in quantum field theory, the simplest example is provided by the quantum mechanics of a point particle moving on a plane, described by a two-vector $\mathbf{r}(t)$, and governed by a potential proportional to a two-dimensional delta function, $\delta^2(\mathbf{r})$. Within classical mechanics, the two-dimensional delta-function potential does not deflect the particle's motion. But solving the quantum-mechanical Schrödinger equation yields a bound state and an isotropic scattering cross section $\sigma(\theta)$, independent of the scattering angle θ, that is given by Eq. (1). Here, $-B$ is the energy of the bound state,

$$\sigma(\theta) = \frac{\lambda}{[\pi^2 + \ln^2(B/E)]} \tag{1}$$

E is the energy of the scattered particle, and λ is its quantum-mechanical wavelength. This example also shows that a symmetry of the classical system is broken: When the spatial variable $\mathbf{r}$ is rescaled to $s\mathbf{r}$, and the time variable t is rescaled as s^2t, where s is a number that can have any value, classical equations of motion are unchanged. This invariance of dynamics, called scale symmetry, holds because the model with a delta-function potential has no intrinsic definition of spatial and temporal scales. Consequences of scale symmetry are that σ/λ must have no energy variation, and that bound states cannot exist. Yet the quantum theory for the planar delta-function potential leads to a bound state and to a cross section where σ/λ varies with energy. The model exhibits what is known as a scale anomaly, and its scale symmetry is broken quantum-mechanically. The mathematical reason for the anomaly is the singular nature of the hamiltonian quantum operator. Owing to its $\mathbf{r} = 0$ short-distance singularity in the delta-function potential, the hamiltonian operator is not well defined on the quantum Hilbert space. Moreover, the mathematical procedure for converting this ill-defined operator into a well-defined one (called self-adjoint extension) necessarily breaks scale invariance. *See* OPERATOR THEORY; SCHRÖDINGER'S WAVE EQUATION.

Role in quantum field theory. While this example exhibits all the features of quantum-mechanical or anomalous symmetry breaking, the physically interesting setting for these phenomena is in quantum field theory, especially as applied to elementary particle physics, where the mechanism serves as an important source for symmetry breaking. Physicists are mostly satisfied that the ultimate laws of nature enjoy a high degree of symmetry. This ancient esthetic prejudice is built into field theoretic models of fundamental particle physics by endowing these models with a high degree of symmetry, whose presence also allows the control and renormalization of the ultraviolet infinities that afflict quantum field theory. The symmetries of the "standard" particle physics model include a scale symmetry (analogous to that described previously), a non-abelian gauge symmetry (matrix-valued generalization of the electromagnetic gauge invariance), and a chiral symmetry associated with massless spin-$^1/_2$ fermions—ingredients of the standard model. This symmetry ensures that a chiral current exists, seemingly obeying a continuity equation, since a classical calculation finds the current to be divergence-free. *See* ELEMENTARY PARTICLE; GAUGE THEORY; QUANTUM FIELD THEORY; STANDARD MODEL.

In spite of physicists' preference that descriptions of nature be enhanced by a large amount of symmetry, actual physical phenomena rarely exhibit overwhelming symmetry. Therefore, at the very same time that a theory is constructed with intrinsic symmetry, a way must be found to break the symmetry in the physical consequences of the model. One mechanism for achieving this is spontaneous symmetry breaking. Another is quantum-mechanical symmetry breaking through anomalies. Two instances of anomalous symmetry breaking in the standard model of particle physics are described below.

Scale anomalies. While the standard model appears to possess scale symmetry, nature clearly exhibits various scales, so it is not scale-invariant. The potential conflict between theory and experiment is avoided because scale invariance is anomalously broken. The true quantum behavior (as opposed to the classical behavior) of quantum-mechanically broken scale symmetry is described by the renormalization group. One aspect is that strengths of interactions depend on the energy at which they are measured, and the dependence is inverse-logarithmic, as in the quantum-mechanical example. This variation has been observed experimentally. *See* RENORMALIZATION.

Chiral anomalies. The previously mentioned chiral symmetry is absent from some channels in the fermion sector. The relevant current j^{μ}_{chiral}, which

would satisfy a continuity equation if the symmetry were present, in fact possesses a nonvanishing, anomalous divergence, which obeys Eq. (2). Here **E**

$$\frac{\partial}{\partial x^\mu} j^\mu_{\text{chiral}}(x) \propto \mathbf{E}\cdot\mathbf{B} \neq 0 \qquad (2)$$

and **B** are the electric and magnetic (abelian and nonabelian) gauge fields present in the model. The proportionality constant is determined by the charges of the fermions (quarks and leptons) in the relevant channel. This result is gratifying, for if the chiral current were conserved, the amplitude for neutral pion decay (to two photons) would be much suppressed relative to the observed value. Instead, with the parameters of the standard model determining the proportionality constant for this anomalous divergence of j^μ_{chiral} in the channel containing the neutral pion, the decay amplitude can be calculated and excellent agreement with experiment is obtained. In fact, the chiral current in all fermionic channels potentially can acquire an anomalous divergence. This would spell disaster for the theory, because consistency and renormalizablility require that the current be divergence-free in those channels to which the gauge fields of the theory couple. But each fermion contributes independently the value of its charge q_i to the proportionality constant in the anomalous divergence relevant to gauge-field-coupled channels. So there is a possibility of cancellation if $\sum_i q_i = 0$. With the known charges for each family of quarks and leptons, the sum indeed vanishes. This is striking evidence that anomaly phenomena exist in nature, and are not mere constructions of mathematical physicists.

The symmetry structure of a theory is encoded in the symmetry group's Lie algebra, which is realized by Poisson brackets in a classical theory and by commutators in the quantum theory. The Lie algebra of the apparent chiral symmetry in the standard model is indeed realized by the Poisson brackets, which are calculated according to the rules of classical physics. On the other hand, calculation of quantum commutators exposes a discrepancy: there are commutator anomalies, reflecting the anomalous breaking of chiral symmetry. Thus anomalous current divergences and anomalous commutators are two sides of the same coin. *See* LIE GROUP.

Relation to infinities. The technical reason for the emergence of scale and chiral anomalies in quantum field theory is the presence of ultraviolet, short-distance infinities in the formalism. These infinities must be controlled and renormalized, but that procedure cannot be carried out in a manner consistent with the symmetries. Thus anomalous symmetry breaking is a consequence of the infinities in quantum field theory. Yet the anomalies themselves are finite quantities. Indeed the chiral anomalies, both in current divergences and in commutators, have a mathematical significance related to topological properties of the gauge fields that occur in the anomaly equations; for example, the $\mathbf{E}\cdot\mathbf{B}$ term in the anomalous divergence of the chiral current is the Chern-Pontryagin density, whose space-time integral is an integer-valued measure of the amount of "twist" in **E** and **B**.

Role in other branches of physics. Quantum anomalies play a role in various other branches of physics, in which quantum field theory finds application, including condensed matter, supersymmetry, string theory, and motion in curved space-time. For example, the requirement that chiral anomalies cancel in string theory or that scale anomalies cancel in supersymmetric theories (analogous to the cancellation in gauge field channels of the standard model) severely constrains the structure and determines the model almost uniquely. This fact suggests that anomalies reflect a subtle and deep feature in the mathematical description of nature, which has been only partially uncovered and understood. *See* SPACE-TIME; SUPERSTRING THEORY; SUPERSYMMETRY. Roman Jackiw

Bibliography. R. Bertlmann, *Anomalies in Quantum Field Theory*, Oxford University Press, 2000; B. R. Holstein, Anomalies for pedestrians, *Amer. J. Phys.*, 61:142–147, 1993; S. Treiman et al., *Current Algebra and Anomalies*, Princeton University Press/World Scientific, 1985; S. Weinberg, *The Quantum Theory of Fields*, vol. 1: *Foundations*, 1995, vol. 2: *Modern Applications*, 1996, Cambridge University Press.

Quantum chemistry

A branch of chemistry concerned with the application of quantum mechanics to chemical problems. More specifically, it is concerned with the electronic structure of molecules. Since 1960 the ease with which the quantum chemist may obtain reliable approximate solutions to the nonrelativistic Schrödinger equation has improved by at least eight orders of magnitude. The following is a brief review of developments in ab initio molecular electronic structure theory since 1960. The term "ab initio" implies that no approximations have been made in the one- and two-electron integrals, shown in Eqs. (1) and (2), arising from the ordinary nonrelativistic hamiltonian operator, Eq. (3). In contrast,

$$I(i \mid j) = \int \phi_i^*(1)\left\{\frac{-\nabla_1^2}{2} - \sum_A \frac{Z_A}{r_{1A}}\right\} \cdot \phi_j(1)\, dv(1) \qquad (1)$$

$$(ij \mid kl) = \int \phi_i^*(1)\phi_j^*(2)\frac{1}{r_{12}} \cdot \phi_k(1)\phi_l(2)\, dv(1)\, dv(2) \qquad (2)$$

$$H = \sum_i \left\{\frac{-\nabla_i^2}{2} - \sum_A \frac{Z_A}{r_{iA}}\right\} + \sum_i \sum_{j>i} \frac{1}{r_{ij}} \qquad (3)$$

semiempirical methods resort to various approximate schemes, especially in evaluating the two-electron integrals [$(ij \mid kl)$; also known as electron repulsion integrals]. The present discussion is primarily restricted to the method that dominates the field of quantum chemistry, namely, the Hartree-Fock

or self-consistent-field approximation and the family of methods grouped under the category of density functional theory. *See* HAMILTON'S EQUATIONS OF MOTION; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Elementary theory. For closed-shell molecules, the form of the Hartree-Fock wave function is given by Eq. (4), in which $A(n)$, the antisymmetrizer for n

$$\psi_{\mathrm{HF}} = A(n)\phi_1(1)\phi_2(2)\ldots\phi_n(n) \tag{4}$$

electrons, has the effect of making a Slater determinant out of the orbital product on which it operates. The ϕ's are spin orbitals, products of a spatial orbital χ and a one-electron spin function α or β. For any given molecular system there are an infinite number of wave functions of form (4), but the Hartree-Fock wave function is the one for which the orbitals ϕ have been varied to yield the lowest possible energy [Eq. (5)].

$$E = \int \psi_{\mathrm{HF}}^{*} H \psi_{\mathrm{HF}}\, d\tau \tag{5}$$

The resulting Hartree-Fock equations are relatively tractable due to the simple form of the energy E for single determinant wave functions [Eq. (6)].

$$E_{\mathrm{HF}} = \sum_i I(i \mid i) + \sum_i \sum_{j>i} [(ij \mid ij) - (ij \mid ji)] \tag{6}$$

To make this discussion more concrete, it should be noted that for singlet methylene (the CH_2 molecule), the Hartree-Fock wave function is of the form given in Eq. (7).

$$\psi_{\mathrm{HF}} = A(8)\ 1a_1\alpha(1)\ 1a_1\beta(2)\ 2a_1\alpha(3)\ 2a_1\beta(4)\ 1b_2\alpha(5)\ 1b_2\beta(6)\ 3a_1\alpha(7)\ 3a_1\beta(8) \tag{7}$$

The same energy expression, Eq. (6), is also applicable to any open-shell system for which the open-shell electrons all have parallel spins. This follows from the fact that such Hartree-Fock wave functions can always be expressed as a single Slater determinant. A simple example is triplet methylene, shown in Eq. (8), for which the outer two $3a_1$ and $1b_1$ or-

$$\psi_{\mathrm{HF}} = A(8)\ 1a_1\alpha(1)\ 1a_1\beta(2)\ 2a_1\alpha(3)\ 2a_1\beta(4)\ 1b_2\alpha(5)\ 1b_2\beta(6)\ 3a_1\alpha(7)\ 1b_1\alpha(8) \tag{8}$$

bitals have parallel spins. For clarity it is often helpful to abbreviate Eq. (8) as Eq. (9). Although solution of

$$\psi = 1a_1^2\ 2a_1^2\ 1b_2^2\ 3a_1\alpha\ 1b_1\alpha \tag{9}$$

the Hartree-Fock equations for an open-shell system such as triplet methylene is more difficult than for the analogous closed-shell system, Eq. (7), the procedures are well established.

In fact, methods are readily available for the solution of the Hartree-Fock equations for any system for which the energy expression involves only Coulomb and exchange integrals, Eqs. (10) and (11).

$$J_{ij} = (ij \mid ij) \tag{10}$$

$$K_{ij} = (ij \mid ji) \tag{11}$$

Open-shell singlets are a class of systems that can be treated in this way, and one such example is the first excited singlet state (of $^1B_1$ symmetry) of methylene, Eq. (12). In addition, these same generalized

$$\psi_{\mathrm{HF}} = \frac{1}{\sqrt{2}}\ 1a_1^2\ 2a_1^2\ 1b_2^2\ 3a_1\alpha\ 1b_1\beta - \frac{1}{\sqrt{2}}\ 1a_1^2\ 2a_1^2\ 1b_2^2\ 3a_1\beta\ 1b_1\alpha \tag{12}$$

Hartree-Fock procedures can be used for certain classes of multiconfiguration Hartree-Fock wave functions.

In recent years, density functional methods have come to play an important role in molecular electronic structure theory. Density functional theory can be traced to a critical 1951 paper by John Slater. In order to avoid the evaluation of the exchange integrals in Eq. (11), Slater proposed an approximation to Hartree-Fock theory in which exchange was treated in a statistical manner. For two decades, this statistical exchange approximation was viewed as a theoretically inferior but computationally advantageous method, primarily used in applications to the band structure of solids, for which ab initio methods were not feasible.

In 1965, W. Kohn and L. J. Sham wedded Slater's statistical exchange to a potentially more powerful theoretical methodology, again with many similarities to ordinary Hartree-Fock theory. However, in principle, the Kohn-Sham method could go beyond Hartree-Fock theory to include electron correlation. Today, the most popular methods for molecular electronic structure theory are hybrid methods involving a mixture of Hartree-Fock theory and density functional theory. For example, the popular B3LYP method combines A. D. Becke's three-parameter treatment of exchange with C. Lee, W. Yang, and R. G. Parr's description of electron correlation. The B3LYP method is semiempirical in the sense that the three parameters in the treatment of exchange were chosen to reproduce certain experimental thermochemical data. Becke's treatment of exchange is a hybrid in that it includes 20% exact exchange; that is, 20% of the exchange functional is the Hartree-Fock exchange. Although the relationship of methods such as B3LYP to an exact quantum-mechanical theory is unclear, it is undeniable that B3LYP and related methods are very efficient and very useful in solving chemistry problems.

Basis sets. To solve the Hartree-Fock or Kohn-Sham equations exactly, one must either expand the spatial orbitals χ in a complete set of analytic basis functions or obtain strictly numerical (that is, tabulated) orbitals. The former approach is impossible from a practical point of view for systems with more than two electrons, and the latter has been accomplished only for atoms and, more recently, for some

TABLE 1. Minimum basis set of Slater functions for the carbon atom

Label	Analytic form	Exponent ζ^*
$1s$	$(\zeta^3/\pi)^{1/2} \exp(-\zeta r)$	5.673
$2s$	$(\zeta^5/3\pi)^{1/2}\, r \exp(-\zeta r)$	1.608
$2p_x$	$(\zeta^5/\pi)^{1/2}\, x \exp(-\zeta r)$	1.568
$2p_y$	$(\zeta^5/\pi)^{1/2}\, y \exp(-\zeta r)$	1.568
$2p_z$	$(\zeta^5/\pi)^{1/2}\, z \exp(-\zeta r)$	1.568

* The orbital exponents ζ are optimum for the ^3P ground state of the carbon atom.

diatomic molecules. Therefore the exact solution of the Hartree-Fock (or Kohn-Sham) equations is abandoned for polyatomic molecules. Instead, one adopts an incomplete (but reasonable) set of analytic basis functions and solves for the best variational [that is, lowest-energy; Eq. (5)] wave function of form (4). C. C. J. Roothaan and coworkers at the University of Chicago referred to such a wave function as being of self-consistent-field (SCF) quality. At the limit of very large basis sets, then, it is reasonable to refer to the resulting SCF wave function as near Hartree-Fock.

For large chemical systems, only minimum basis sets (MBS) are typically used in ab initio theoretical studies. The term "large" includes molecular systems with 1000 or more electrons. A large molecule treated by MBS-SCF methods as early as 1991 is the six-base-pair DNA fragment $C_{116}H_{138}N_{46}O_{68}P_{10}$. Significantly larger systems can be treated today with state-of-the-art quantum-chemical methods. A minimum basis set includes one function for each orbital occupied in the ground state of each atom included in the molecule. For the first row atoms B, C, N, O, and F, this means that a minimum basis set includes $1s$, $2s$, $2p_x$, $2p_y$, and $2p_z$ functions.

Historically, minimum basis sets have been composed of Slater functions, such as those seen in **Table 1** for the carbon atom. However, experience has shown that the evaluation of the molecular integrals, Eq. (2), arising when Slater functions are employed is extremely time-consuming. Therefore each Slater function in a minimum basis set is typically replaced by a linear combination of three or four gaussian functions. The resulting chemical predictions obtained with such STO-3G (Slater-type orbital–three gaussian functions) or STO-4G basis sets are usually indistinguishable from the corresponding Slater function results.

Minimum basis sets are inadequate for certain types of chemical predictions. Therefore a basis twice as large, and appropriately designated double-zeta (DZ), is often used in theoretical studies. Here, however, it is not as fruitful to expand each Slater function as a linear combination of gaussians. Instead gaussian functions $x^p y^q z^r e^{-\alpha r^2}$ are used directly in atomic self-consistent-field calculations and then contracted according to the atomic results. Among the most widely used contracted gaussian double-zeta basis sets are those of T. H. Dunning. His double-zeta basis has $9s$ and $5p$ original (or primitive) gaussian functions, and is contracted to $4s$ and $2p$. Thus the basis may be designated C($9s5p/4s2p$). A substantial molecular system studied as early as 1995 with the double-zeta basis set is the giant fullerene C_{240}, a molecule with 1440 electrons. By 1999 G. E. Scuseria and coworkers were able to report a DZ basis set study of a 1026-atom (6767 basis functions) RNA fragment.

Just as the double-zeta basis logically follows the minimum set, the logical extension of the double-zeta set involves the addition of polarization functions. Polarization functions are of higher orbital angular momentum than the functions occupied in the atomic self-consistent-field wave function. That is, for carbon, $d, f, g, \ldots$ functions will be polarization functions. Fortunately, d functions are far more important than f, f functions are far more important than g, and so on. For most chemical applications a double-zeta plus polarization (DZ + P) basis including a single set of five d functions ($d_{x^2-y^2}$, d_{z^2}, d_{xy}, d_{xz}, d_{yz}) will be quite adequate for first-row atoms.

Structural predictions. Ab initio theoretical methods have had the greatest impact on chemistry in the area of structural predictions. A good illustration of this is the methylene radical for which S. F. Boys reported the first ab initio study in 1960. Boys predicted the structure of triplet methylene to be r_e(C—H) = 0.112 nanometer, θ_e(HCH) = 129°. Unfortunately, however, the work of Boys was largely denigrated due to the 1961 spectroscopic conclusion of G. Herzberg that the lowest triplet state of CH_2 was linear. Herzberg's conclusion was greatly strengthened by a very influential 1962 semiempirical study of H. C. Longuet-Higgens, who concurred that the ground state of CH_2 was linear.

It was not until 1970 that a definitive theoretical prediction of the nonlinearity of $^3B_1$ CH_2 appeared. The prediction was swiftly verified by independent electron spin resonance experiments. For many chemists, the structure of triplet methylene was the first genuine example of the usefulness of ab initio theoretical chemistry.

Turning from the specific to the more general, the most encouraging aspect of ab initio geometry predictions is their, perhaps, surprising reliability. Essentially all molecular structures appear to be reliably predicted at the Hartree-Fock level of theory. Even more encouraging, many structures are accurately reproduced by using only minimum-basis-set self-consistent-field methods. This is especially true for hydrocarbons. A fairly typical example is methylenecyclopropane (**Fig. 1**), with

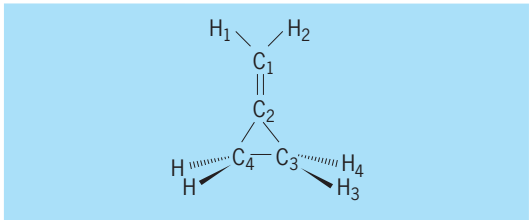

Fig. 1. Methylenecyclopropane structure.

TABLE 2. Minimum-basis-set self-consistent-field geometry prediction compared with experiment for methylenecyclopropane

Parameter*	STO-3G	Experiment
$r(C_1{=}C_2)$	0.1298 nm	0.1332 nm
$r(C_2{-}C_3)$	0.1474 nm	0.1457 nm
$r(C_3{-}C_4)$	0.1522 nm	0.1542 nm
$r(C_1{-}H_1)$	0.1083 nm	0.1088 nm
$r(C_3{-}H_3)$	0.1083 nm	0.109 nm
$\theta(H_1C_1H_2)$	116.0°	114.3°
$\theta(H_3C_3H_4)$	113.6°	113.5°
$\theta(H_{34}C_3C_4)$	149.4°	150.8°

*Here r represents the carbon-carbon or carbon-hydrogen bond distance; θ represents the bond angle in degrees of H-C-H bonds; the numbers on C and H correspond to the numbered atoms in Fig. 1.

its minimum-basis-set self-consistent-field structure compared with experiment in **Table 2**. Carbon-carbon bond distances differ typically by 0.002 nm from experiment, and angles are rarely in error by more than a few degrees. Even for severely strained molecules such as bicyclo[1.1.0] butane, very reasonable agreement with the experimental structure is obtained. It is noteworthy that experimental geometries are available for less than half of the C_4 hydrocarbons studied theoretically to date. Thus, for many purposes, theory may be considered complementary to experiment in the area of structure prediction.

For molecules including atoms in addition to C and H, minimum-basis-set self-consistent-field results are sometimes less reliable. For example, the F_2N_2 molecule has minimum-basis-set self-consistent-field bond distances r_e(N—F) = 0.1384 nm, r_e(N═N) = 0.1214 nm, which are respectively 0.0107 nm longer and 0.0169 nm shorter than the experimental values. Fortunately, vastly improved agreement with experiment results is obtained when a larger basis set is adopted for F_2N_2.

In general, double-zeta self-consistent-field structure predictions are considerably more reliable than those based on minimum basis sets. A noteworthy exception is the water molecule, for which minimum-basis-set self-consistent-field yields a bond angle of 100.0° and double-zeta self-consistent-field predicts 112.6°, compared to the well-known experimental value of 104.5°. More typical are the HF and F_2 molecules, for which the minimum-basis-set, double-zeta, and experimental bond distances are 0.0956, 0.0922, and 0.0917 nm (HF); and 0.1315, 0.1400, and 0.1417 nm (F_2). In fact it can be argued that if the theory will not be pursued beyond the Hartree-Fock (single-configuration) approximation, double-zeta self-consistent-field is often a reasonable stopping point.

Transition states are typically more sensitive to basis set than equilibrium geometries. This is true because potential energy surfaces are often rather flat in the vicinity of a saddle point (transition state). An example is the carboxime–cyanic acid rearrangement given in reaction (13).

$$\text{HONC} \rightarrow \text{HOCN} \qquad (13)$$

Minimum-basis-set self-consistent-field and double-zeta (technically 4-31G) self-consistent-field transition state geometries are compared in **Fig. 2**. There it is seen that the minimum-basis-set and double-zeta predictions are very different. Neither transition-state structure is reliable, as is seen in Fig. 2 by comparison with results from more complete basis sets. It should be noted that for many other transition states, better agreement is found between the minimum and double-zeta basis sets. More typical structural variations are ~0.005 nm in internuclear separations and 5° in angles.

As larger basis sets within the Hartree-Fock formalism are considered, better agreement with experiment is frequently obtained. As implied above, the water molecule bond angle is much improved at the double-zeta plus polarization level, to 106.1°. However, it is often the case that adding polarization functions has only a marginal effect on predicted geometries. A reasonably typical comparison is given by the NH_2F and PH_2F molecules. In this example, the only polarization functions added were sets of d functions on the central N or P atom. The only pronounced improvement with respect to experiment is for the P—F separation in PH_2F, and this is improved by 0.0034 nm when d functions are added to phosphorus. The good agreement with experiment for NH_2F and PH_2F suggests a high degree of reliability for the comparable NHF_2 and PHF_2 predictions, where no experimental structures have been determined.

An interesting comparison of the three frequently used basis sets is given in **Table 3** for the linear HCNO molecule. The most sensitive geometrical parameter is the N—O bond distance, for which the

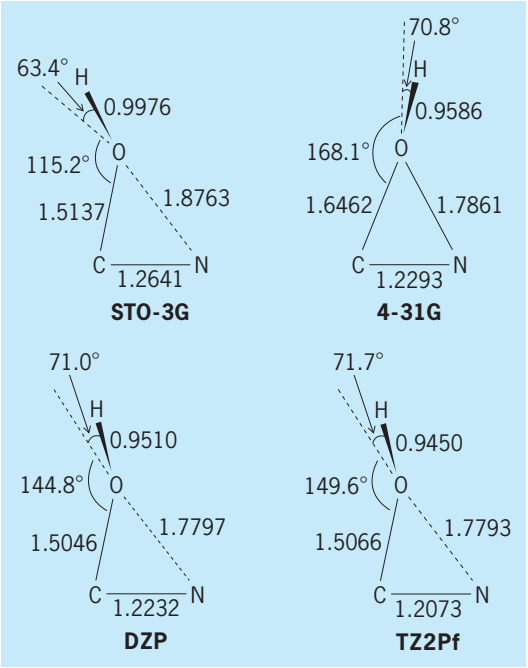

Fig. 2. Comparison of self-consistent-field transition states for the HONC → HOCN reaction using four different basis sets.

TABLE 3. Equilibrium geometry of formonitrile oxide, HCNO, from self-consistent-field theory and experiment*

	MBS	4-31G	DZP	Experiment
r_e(H—C)	0.1065	0.1049	0.1059	0.1027
r_e(C—N)	0.1155	0.1133	0.1129	0.1168
r_e(N—O)	0.1294	0.1255	0.1201	0.1199

*Values are in nanometers.

minimum basis set is 0.0095 nm too long, the double-zeta still 0.0056 nm too long, but the double-zeta plus polarization result is in nearly perfect agreement with experiment. For the CN distance, the minimum-basis-set treatment actually gives the best agreement with experiment. The experimental microwave spectrum is difficult to unravel for a quasi-linear molecule such as HCNO, and it has been suggested that the double-zeta plus polarization prediction for the C—H distance may be more reliable than experiment.

An example of a sizable molecule for which complete geometry optimization became routine in 1998, even using a substantial triple-zeta plus polarization basis set, is the digallium compound $Na_2\{Ga_2[C_6H_3(C_6H_5)_2]_2\}$ [**Fig. 3**]. The complete structural optimization (50 independent degrees of freedom) was carried out using the B3LYP method in conjunction with a basis set of 836 contracted gaussian functions.

Energetic predictions. Among the most chemically important energetic quantities are conformational energy changes, exothermicities or heats of reaction, dissociation energies, and activation energies or barrier heights. In general, only the first of these, and sometimes the second, is reliably predicted at the Hartree-Fock level of theory. In other words, energetic quantities are often sensitive to the effects of electron correlation.

Conformational energy changes are, almost without exception, properly reproduced within the Hartree-Fock formalism. In fact, certain types of barriers, typified by the ethane rotational barrier, are quite satisfactorily predicted at the minimum-basis-set self-consistent-field level of theory. More sensitive problems, such as the ammonia inversion barrier and the rotational barrier of hydrogen peroxide, demand the inclusion of polarization basis functions.

Although Hartree-Fock exothermicities are often unreliable, there is at least one fairly large class of reactions for which consistently good agreement with experiment has been found. Generally speaking, heats of reaction for systems having closed-shell reactants and products are often predicted successfully. More specifically, even better agreement with experiment is found for isodesmic reactions, where the number of bonds of each type is conserved. In fact, reasonable predictions are often made at the minimum-basis-set self-consistent-field level for isodesmic reactions. Further, such information can sometimes be used indirectly (or in conjunction with other thermochemical information) to predict quantities that might be very difficult to evaluate by more straightforward ab initio methods.

The dissociation energies of covalent molecules are generally predicted poorly by single-configuration self-consistent-field methods. Certainly the best-known example is the F_2 molecule, for which the molecular Hartree-Fock energy lies about 1 eV above the Hartree-Fock energy of two fluorine atoms. This problem is often mistakenly attributed to the "perverse" nature of the fluorine atom. In fact, the near-Hartree-Fock dissociation energy of N_2 is 5.27 eV, only about half of the experimental value, 9.91 eV. In addition, the near-Hartree-Fock dissociation energy of O_2 is 1.43 eV, only one-third of the experimental value, 5.21 eV. Thus the Hartree-Fock dissociation energies of covalent molecules are consistently much less than experiment.

Another frequent failing of the Hartree-Fock method is in the prediction of the barrier heights or activation energies of chemical reactions. However, it must be noted that there are many classes of reactions for which Hartree-Fock theory does yield meaningful barrier heights. Two well-studied examples are the isomerizations shown in reactions (14) and (15). For the HNC rearrangement, comparison

$$\mathrm{HNC} \rightarrow \mathrm{HCN} \tag{14}$$

$$\mathrm{CH_3NC} \rightarrow \mathrm{CH_3CN} \tag{15}$$

between self-consistent-field (40 kilocalories or 167 kilojoules) and configuration interaction (CI; 36 kcal or 151 kJ) barriers reveals good qualitative agreement. However, the inclusion of *d* functions in

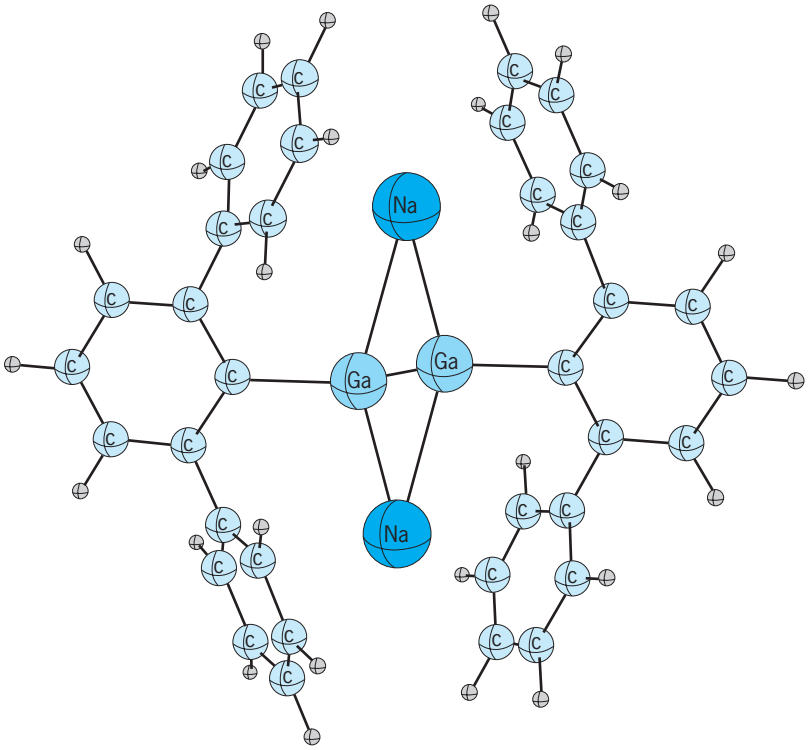

Fig. 3. Example of a sizable molecule for which complete DFT geometrical optimization with a substantial basis set became routine in 1998.

the C and N basis sets appears to be very important. For example, for the methyl isocyanide rearrangement, the self-consistent-field barrier decreases from 60 to 45 kcal (251 to 188 kJ) when polarization functions are added to a double-zeta basis. The remaining discrepancy with the experimental activation energy of 38 kcal (159 kJ) is about equally due to correlation effects and the fact that the zero-point vibrational energy at the transition state is ~3 kcal (13 kJ) less than for CH_3NC. Hydrogen isocyanide is one of the first three simple interstellar molecules (HNC, HCO^+, and HN_2^+) to be identified by ab initio theory prior to their laboratory detection.

Generally speaking, unimolecular reactions seem to be treated more reliably by Hartree-Fock methods than bimolecular systems. A second example is the geometrical isomerization of cyclopropane. This system has been studied in considerable detail by L. Salem and coworkers, who were the first to totally resolve the structure of the transition state within the full 21-dimensional hypersurface. Although their work involved only a minimum basis set, it goes slightly beyond the self-consistent-field model in that 3×3 configuration interaction was included. The predicted barrier height is 53 kcal (222 kJ), in reasonable agreement with the experimental value, 64 kcal (268 kJ). Subsequent studies with much larger basis sets show quantitative agreement with experiment.

For many attractive potential energy surfaces, that is, those having no barrier or activation energy at all, Hartree-Fock methods are frequently reliable. An example that has been carefully documented is reaction (16). One of the most important features of

$$H + Li_2 \rightarrow LiH + Li \qquad (16)$$

the $H + Li_2$ surface is the fact that the C_{2v} HLi_2 structure is a chemically bound entity. Self-consistent-field theory suggests that the dissociation energy relative to LiH + Li is 20 kcal (84 kJ), in excellent agreement with the large-scale configuration interaction result of 22 kcal (93 kJ).

Hybrid Hartree-Fock/density functional methods such as B3LYP usually do a far better job than Hartree-Fock theory for thermochemistry. Often, diatomic molecule dissociation energies have errors an order of magnitude smaller with B3LYP than with Hartree-Fock. Moreover, the density functional methods are typically less sensitive to basis set than is Hartree-Fock theory. Thus it seems likely that the hybrid Hartree-Fock/density function methods will dominate quantum chemistry for the foreseeable future.

Beyond Hartree-Fock and density functional theory. The most important post-Hartree-Fock methods for quantum chemistry are perturbation theory and the configuration interaction (CI) and coupled cluster (CC) methods. These three rigorous approaches may be labeled "convergent" quantum-mechanical methods, as each is ultimately capable of yielding exact solutions to Schrödinger's equation. The present discussion focuses on coupled cluster theory. The coupled cluster method treats excitations based on the number of electrons by which they differ from the Hartree-Fock reference function. Thus the CCSD method incorporates amplitudes differing by single (S) and double (D) excitations from Hartree-Fock. The CCSDT method adds all triple excitations to the CCSD treatment. As one goes to higher and higher excitations (for example, CCSDTQ includes all configurations differing by one, two, three, or four electrons from the Hartree-Fock reference configuration), one approaches the exact quantum-mechanical result.

In fact, coupled cluster theory beyond CCSD becomes impractical for large molecular systems. Thus, although the coupled cluster path to exact results is clear, it becomes a difficult road to follow. Triple excitations are sufficiently important that effective coupled cluster methods have been developed in which the effects of triples are approximated. The best of these methods, CCSD(T), is the closest thing to a panacea that exists today in quantum chemistry for very difficult problems involving smaller molecules. However, the range of applicability of the theoretically superior CCSD(T) method is much narrower than that of the popular hybrid Hartree-Fock/density functional methods. For example, the DZP basis-set coupled cluster treatments of ferrocene, $Fe(C_5H_5)_2$, in 1995 and the small fullerene C_{20} in 1996 were considered state-of-the-art and required the use of supercomputers. In contrast, these systems were readily (though not as reliably) treated in the year 2000 with personal computers costing less than $500 when methods such as B3LYP were employed. Thus, for most chemists the Hartree-Fock and density functional methods are likely to play central roles in molecular electronic structure theory for many years to come. *See* CHEMICAL BONDING; COMPUTATIONAL CHEMISTRY; MOLECULAR ORBITAL THEORY; MOLECULAR STRUCTURE AND SPECTRA; RESONANCE (MOLECULAR STRUCTURE); STRUCTURAL CHEMISTRY.

Henry F. Schaefer III

Bibliography. R. F. Bader, *Atoms in Molecules: A Quantum Theory*, 1994; C. J. Barden and H. F. Schaefer, Quantum chemistry in the twenty-first century, *Pure Appl. Chem.*, 72:1405–1423, 2000; I. N. Levine, *Quantum Chemistry*, 5th ed., 2000; J. P. Lowe, *Quantum Chemistry*, 3d ed., 2005; H. F. Schaefer, The reachable dream, *J. Mol. Struc.*, 398:199–209, 1997; T. Veszpremi and M. Feher, *Quantum Chemistry: Fundamentals to Applications*, 1999; Y. Yamaguchi et al., *A New Dimension to Quantum Chemistry*, 1994.

Quantum chromodynamics

A theory of the strong (nuclear) interactions among quarks, which are regarded as fundamental constituents of matter, structureless and indivisible at the current resolution of about 10^{-18} m. Quantum chromodynamics (QCD) explains why quarks combine in certain configurations to form the observed patterns of subnuclear particles, such as the proton and pi meson. According to this picture, the strong interactions among quarks are mediated by a set of force

particles known as gluons. Quantum chromodynamics successfully accounts for many features of high-energy hard scattering of the strongly interacting particles. Interactions among gluons may lead to new structures that correspond to as-yet-undiscovered particles. The long-studied nuclear force that binds protons and neutrons together in atomic nuclei is regarded as a collective effect of the elementary interactions among constituents of the composite protons and neutrons. *See* NUCLEAR STRUCTURE.

By construction, quantum chromodynamics embodies many features abstracted from empirical observations, but the strength of the theory lies in the predictions it makes and the new understanding it brings. The property of asymptotic freedom implies a realm in which reliable calculations can be extracted using traditional Feynman-diagram techniques. These have been subjected to numerous quantitative tests with impressive success. The development of lattice gauge theory has made possible a growing range of computations of particle properties that emerge from the confinement of quarks. Part of the esthetic appeal of the theory is due to the fact that quantum chromodynamics is nearly identical in mathematical structure to quantum electrodynamics (QED) and to the unified theory of weak and electromagnetic interactions. This resemblance encourages the hope that a unified description of the strong, weak, and electromagnetic interactions may be at hand. *See* ELECTROWEAK INTERACTION; FEYNMAN DIAGRAM; QUANTUM ELECTRODYNAMICS; WEAK NUCLEAR INTERACTIONS.

Gauge theories. At the heart of current theories of the fundamental interactions is the idea of gauge invariance. Gauge theories constructed to embody various symmetry principles have emerged as the correct quantum-mechanical descriptions of the strong, weak, and electromagnetic interactions, at energies up to at least 1 TeV (10^{12} eV). *See* SYMMETRY LAWS (PHYSICS).

Electromagnetism. The simplest example of a gauge theory is electromagnetism. It may be derived from a symmetry principle as follows. Quantum-mechanical observables do not depend upon the phase of the complex wave function that describes the state of a system. Therefore, the phase of a wave function can be rotated by an amount that is the same at all times and all places without affecting the physical consequences of the theory. The choice of phase is thus conventional, as opposed to observable. This is known as a global symmetry principle. It is natural to ask whether it should not be possible to choose this arbitrary convention independently at each point of space-time, again without affecting the physical consequences of the theory. It turns out to be possible to construct a quantum theory that is invariant under local (that is, position- and time-dependent) phase rotations that are proportional to the electric charge of the particles, but only if the theory contains an electromagnetic field with precisely the observed properties as summarized by Maxwell's equations. In the quantum theory, a massless spin-1 particle identified as the photon mediates the electromagnetic interaction. The interactions of matter with electromagnetism thus are essentially prescribed by the requirement of local phase invariance.

Local gauge invariance. Local phase rotations of the kind described above are the simplest examples of local gauge transformations. For a continuous symmetry, global gauge invariance implies the existence of a set of conserved currents. In the case of electromagnetism, it is the electromagnetic current that is conserved. A local gauge invariance requires, in addition, the existence of a massless gauge field corresponding to the set of phase transformations that forms the one-parameter unitary group U(1). The question arises as to whether the notion of local gauge invariance can be used to deduce the theory of nuclear forces.

Yang-Mills theory. This question was addressed in 1954 by C. N. Yang and R. L. Mills and, independently, by R. Shaw. Early in the study of nuclear forces it was established that the nuclear interaction is charge-independent; it acts with the same strength between proton and proton, or proton and neutron, or neutron and neutron. This may be understood by thinking of the proton and neutron as representing two states of the same particle, called the nucleon. Just as an electron can be in a state with spin up or spin down, a nucleon can be in a state with the internal quantum number isospin up (defined as the proton) or isospin down (defined as the neutron). Charge independence then would reflect the invariance of the strong interactions under isospin rotations, characterized by the group SU(2). If isospin is regarded as a gauge group, local gauge invariance requires the existence of three massless spin-1 gauge particles, corresponding to the three generators of SU(2). The interactions of the gauge particles with nucleons are prescribed by the gauge principle. All of this is entirely parallel to the theory of electromagnetism. What distinguishes this SU(2) gauge theory from its U(1) counterpart is that the SU(2) gauge fields carry isospin and thus couple among themselves, whereas the photon is electrically neutral and does not interact with itself. Interacting gauge fields are an attribute of any theory based upon a nonabelian gauge group. *See* I-SPIN.

Spontaneous symmetry breaking. Its mathematical properties notwithstanding, the Yang-Mills theory was unacceptable as a description of nuclear forces because they are mediated by massive particles, whereas the gauge particles are required to be massless. Beginning in the late 1950s, a succession of gauge theories of the weak interactions appeared. At first these too foundered on the prediction of massless gauge bosons, but it was ultimately learned that spontaneous breakdown of the gauge symmetry would endow the gauge bosons with masses. All the elements were successfully combined in 1967 in the theory of weak and electromagnetic interactions proposed separately by S. Weinberg and A. Salam. When in 1971 it was demonstrated that spontaneously broken gauge theories were renormalizable, and hence calculable in the same sense as quantum electrodynamics, experimental interest in the predictions of

the Weinberg-Salam model was stimulated, and theoretical enthusiasm for gauge theories in general was renewed. Spontaneous symmetry breaking was not a cure for the shortcomings of gauge theories of the strong interactions. Instead, thanks to parallel developments described below, a new candidate emerged for the strong gauge group, and with it arose the idea of quantum chromodynamics. *See* GAUGE THEORY; RENORMALIZATION; SYMMETRY BREAKING.

Color. It was shown in 1963 that the observed pattern of strongly interacting particles, or hadrons, could be explained if the hadrons were composed of fundamental constituents called quarks. According to the quark model, a baryon such as the proton, which has half-integral spin in units of $\hbar$ (Planck's constant h divided by 2π), is made up of three quarks, each with spin $^1/_2$. An integral-spin meson, such as the pi meson, is made up of one quark and one antiquark. Three varieties (or flavors) of quarks, denoted up, down, and strange, could be combined to make all of the known hadrons, in precisely the families identified according to the eightfold way.

Although the idea that the strongly interacting particles are built up of quarks brought new order to hadron spectroscopy and suggested new relations among mesons and baryons, the constituent description brought with it a number of puzzles. An obvious question concerns the rules by which the hadrons are built up out of quarks. Mesons are composed of a quark and an antiquark, while baryons are made of three quarks. What prevents two-quark or four-quark combinations? Within this innocent question lurks a serious problem of principle. The Pauli exclusion principle of quantum mechanics is the basis for understanding the periodic table of the elements. It restricts the configurations of electrons within atoms and of protons and neutrons within nuclei. It should be a reliable guide to the spectrum of hadrons as well. According to the Pauli principle, identical spin-$^1/_2$ particles cannot occupy the same quantum state. As a consequence, the observed baryons such as Δ^{++} (uuu) and Ω^- (sss), which would be composed of three identical quarks in the same state, would seem to be forbidden configurations.

To comply with the Pauli principle, it is necessary to make the three otherwise identical quarks distinguishable by supposing that every flavor of quark exists in three varieties, fancifully labeled by the colors red, green, and blue. Then each baryon can be constructed as a "colorless" (or "white") state of a red quark, a green quark, and a blue quark. Similarly, a meson will be a colorless quark-antiquark combination. The rule for constructing hadrons may then be rephrased as the statement that only colorless states can be isolated. *See* COLOR (QUANTUM MECHANICS).

A second issue is raised by the fact that free quarks have not been observed. This suggested that the interaction between quarks must be extraordinarily strong, and perhaps permanently confining. That free quarks are not seen is of course consistent with the idea that colored states cannot exist in isolation. On the other hand, the quark-parton model description of violent collisions rests on the assumption that quarks within hadrons may be regarded as essentially free, as explained below.

This paradoxical state of affairs may be visualized as follows. A hadron may be thought of as a bubble within which the constituent quarks are imprisoned. The quarks move freely within the bubble but cannot escape from it. This picture yields an operational understanding of many aspects of hadron structure and interactions, but it falls far short of a dynamical explanation for the puzzling behavior of quarks.

The quarks are the constituents of strongly interacting particles. The leptons, of which the electron and neutrino are the most common examples, are the fundamental particles that do not interact strongly. Each lepton flavor appears in only a single species. In other words, the leptons are colorless. In other respects, leptons resemble quarks: they are spin-$^1/_2$ particles that have no internal structure at the current limits of resolution. Color may therefore be regarded as the strong-interaction analog of electric charge. Color cannot be created or destroyed by any of the known interactions. Like electric charge, it is said to be conserved. *See* LEPTON.

In light of evidence that color could be regarded as the conserved charge of the strong interactions, it was natural to seek a gauge symmetry that would have color conservation as its consequence. An obvious candidate for the gauge symmetry group is the unitary group SU(3), now to be applied to color rather than flavor. The theory of strong interactions among quarks that is prescribed by local color gauge symmetry is known as quantum chromodynamics. The mediators of the strong interactions are eight massless spin-1 bosons, one for each generator of the symmetry group. These strong-force particles are named gluons because they make up the "glue" that binds quarks together into hadrons. Gluons also carry color (in eight color-anticolor combinations) and hence have strong interactions among themselves.

Asymptotic freedom. The theoretical description of the strong interactions has historically been inhibited by the very strength of the interaction, which renders low-order perturbative calculations untrustworthy. However, in 1973 H. David Politzer, David J. Gross, and Frank Wilczek found that in many circumstances the effective strength of the interaction in Yang-Mills theories becomes increasingly feeble at short distances. For quantum chromodynamics, this remarkable observation implies that the interaction between quarks becomes weak at small separations. This discovery raises the hope that some aspects of the strong interactions might be treated by using familiar computational techniques that are predicated upon the smallness of the interaction strength.

The physical basis for the change in the strength of the strong interaction as a function of the distance may be understood by examining the corresponding question in electrodynamics. If an electric charge is placed in surroundings in which other

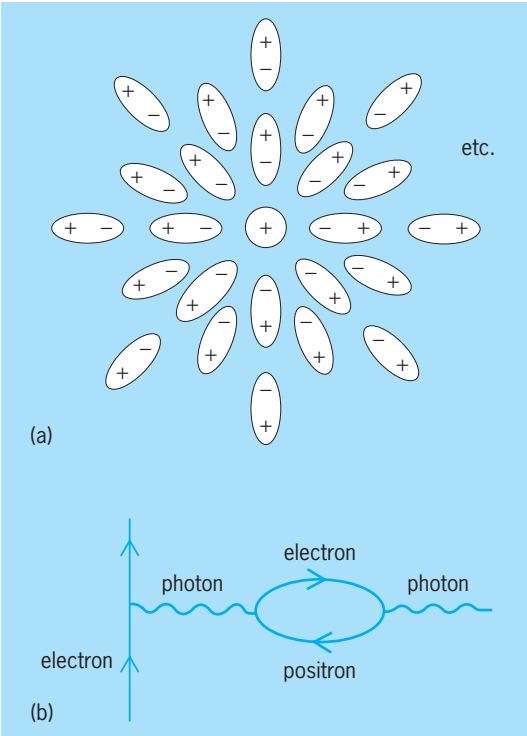

Fig. 1. Screening in electrodynamics. (*a*) Polarization of a dielectric medium by a test charge. (*b*) Feynman diagram contributing to vacuum polarization.

charges are free to move about, the effects of the charge may be modified. An example is a medium composed of many molecules, each of which has a positively charged end and a negatively charged end. Introducing a test charge polarizes the medium (**Fig. 1***a*). Charges of opposite sign are oriented toward the test charge, while those with like charge are repelled. This screening effect means that the influence of the charge is diminished by the surrounding medium. Viewed from afar (but within the medium), the charge appears smaller in magnitude than its true or unscreened value.

A closely related phenomenon in quantum electrodynamics is known as vacuum polarization (Fig. 1*b*). The vacuum, or empty space, is normally thought of as the essence of nothingness. However, in quantum theory the vacuum is a complicated and seething medium in which "virtual" pairs of charged particles, predominantly electrons and positrons, have a fleeting existence. These ephemeral vacuum fluctuations are polarizable in the same way as the molecules of the above example, and they serve to screen the charge at large distances. Consequently, in quantum electrodynamics it is also expected that the effective electric charge should increase at short distances. The behavior of the effective charge in quantum electrodynamics is opposite to that required to understand the strong interactions, where the interaction between quarks must diminish in strength at short distance.

The situation in quantum chromodynamics is complicated by the fact that gluons carry the strong (color) charge, whereas photons are electrically neutral. This means that, in addition to a color polarization phenomenon like the related charge screening of quantum electrodynamics (**Fig. 2***a* and *b*), it is possible for the color charge of a quark to be shared with the gluon cloud (Fig. 2*c* and *d*). Because the color charge is spread out rather than localized, the effective color charge will tend to appear larger at long distances and smaller at short distances. The outcome of the competition between these opposing tendencies depends on the number of gluon species that can share the color charge and on the number of quark types that can screen the color charge. If the color gauge group is SU(3), the net effect is one of antiscreening, that is, of a smaller effective charge at short distances, provided the number of quark flavors is less than 17. Six quark flavors are known. Extremely

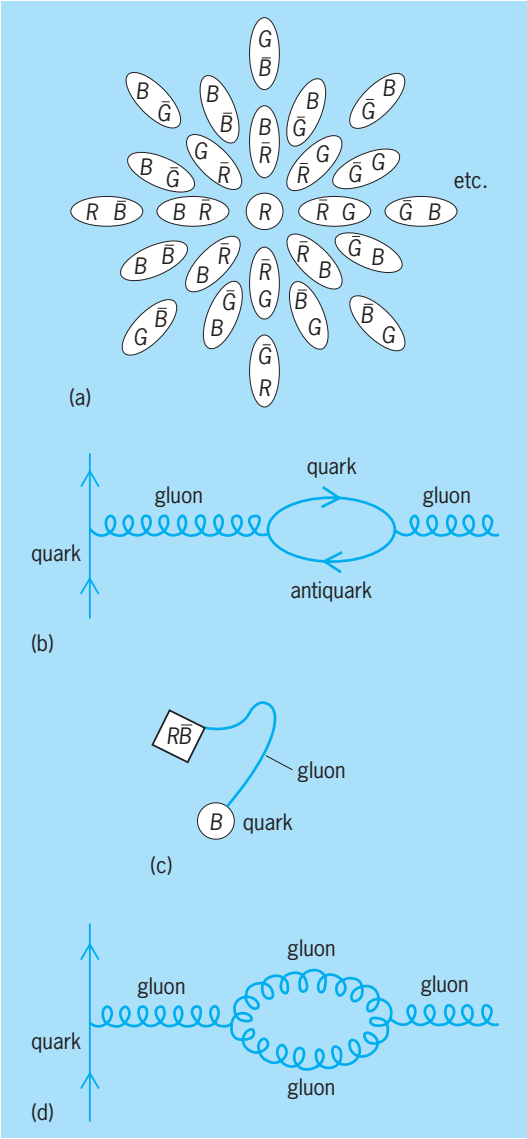

Fig. 2. Screening and antiscreening in quantum chromodynamics. (*a*) Screening of a colored quark by vacuum polarization. Color charges are denoted by R = red, $\bar{R}$ = anti-red, B = blue, $\bar{B}$ = anti-blue, G = green, $\bar{G}$ = anti-green. (*b*) Feynman diagram contributing to color polarization of the vacuum. (*c*) Delocalization of the color charge by gluon radiation. (*d*) Feynman diagram contributing to color antiscreening.

close to a quark, the effective color charge becomes vanishingly small, so that nearby quarks behave as if they were noninteracting free particles. This is the origin of the term asymptotic freedom.

Experimental consequences. Asymptotic freedom has the immediate consequence, within the context of quantum chromodynamics, of providing a partial justification of the parton model description of violent scattering processes. The parton model was invented to explain features of electron-proton collisions in which many particles are created. Experiments first carried out in 1968 indicated that in these deeply inelastic collisions the electron "sees" the proton, not as an amorphous whole, but as a collection of structureless entities that have been identified as quarks. According to this picture, an electron scatters from a single parton, free from the influence of any other parton. This is reminiscent of electron scattering from nuclei, in which an electron may scatter from an individual proton or neutron as if it were a free particle. There is an important difference. Protons and neutrons are lightly bound in nuclei and may be liberated in violent collisions. In contrast, free quarks have not been observed, so they must be regarded as very deeply bound within hadrons. Asymptotic freedom offers a resolution to the paradox of quasifree quarks that are permanently confined. At the short distances probed in deep inelastic scattering, the effective color charge is weak, so the strong interactions between quarks can largely be neglected. As quarks are separated, the effective color charge grows, so infinite energy would be required to extract a quark from an isolated hadron. This characteristic offers a mechanism for quark confinement, but an analytic proof has not yet been constructed. A critical insight of quantum chromodynamics is that most of a proton's mass arises from the energy stored up in confining its component quarks in a small volume, not from the masses of the quarks.

In the regime of short distances probed in violent high-energy collisions, the strong interactions are sufficiently feeble that reaction rates may be calculated by using the diagrammatic methods developed for quantum electrodynamics. These calculations reproduce the simple quark-parton model results as first approximations. For example, in electron-positron annihilations into hadrons, the quark-antiquark production rate correctly anticipates both the structure of the dominant two-jet events and the approximate rate of hadron production. The strong-interaction corrections to this process include the process in which a gluon is radiated by one of the outgoing quarks. Like the quarks, the gluon materializes as a jet of hadrons. The resulting three-jet events are commonplace in electron-positron annihilations.

In high-energy proton-antiproton collisions, the hard scatterings of the partons (**Fig. 3***a*) lead to striking jets of hadrons at large angles to the direction defined by the incident proton and antiproton beams (Fig. 3*b*). Quantum chromodynamics, supplemented by measurements of the parton distributions within protons, accounts for the features of jet production in great detail.

Although quantum chromodynamics provides support for the spirit of the parton model, it also exposes the incompleteness of the parton model description. Because quantum chromodynamics is an interacting field theory of quarks and gluons, probes of different wavelengths, analogous to microscopes of different resolving power, may map out different structures within the proton. Under increasing magnification (**Fig. 4**), quarks undergo virtual dissociation into quarks and gluons. The structure of the proton resolved in high-energy scattering experiments changes in the way predicted by quantum chromodynamics.

Quarkonium. The light hadrons are sufficiently large that the forces between the quarks within them

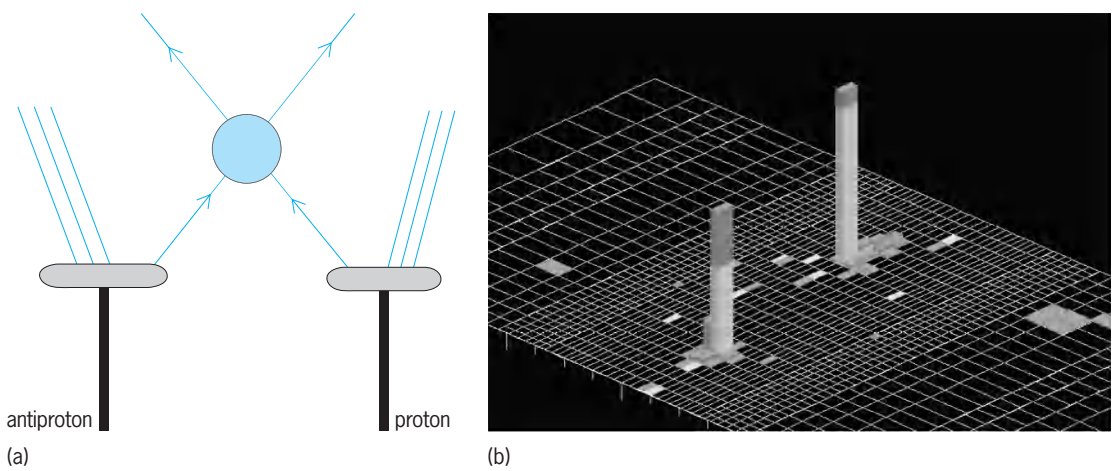

Fig. 3. High-energy proton-antiproton collisions. (*a*) Parton-model description of a violent collision between constituents of a proton and an antiproton. The active particles may be quarks, antiquarks, or gluons. (*b*) The most energetic two-jet event yet observed, as recorded in the Collider Detector at Fermilab. A proton and antiproton, each with energy 980 GeV, collide head on. In this event, a quark from the proton scatters at large angle off an antiquark from the antiproton. Two jets of hadrons emerge, carrying 1364 GeV at right angles to the initial beam directions. The event display depicts the energy deposited in individual cells of the cylindrical detector, unrolled. (*Courtesy CDF Collaboration*, http://www-cdf.fnal.gov)

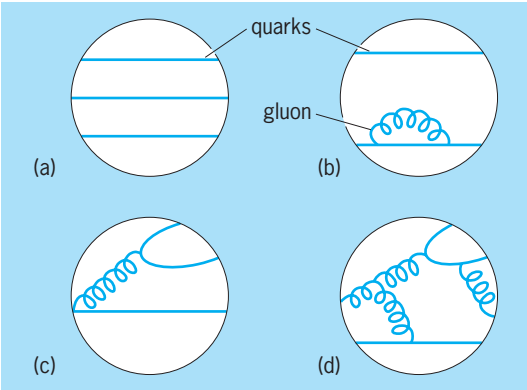

Fig. 4. Structure of the proton as observed with probes of increasing resolution. (*a*) Quarks appear as simple and noninteracting objects under "low magnification." (*b–d*) Quarks reveal increasing complexity because of their interactions with gluons as resolving power is improved.

are quite formidable. A fundamental description of the spectroscopy of light hadrons therefore relies on large-scale computer simulations. It is, however, possible to imagine special situations in which hadron spectroscopy is completely tractable by using available theoretical techniques. It was suggested in 1974 that the bound system of an extremely massive quark with its antiquark would be so small that the strong force would be extremely feeble. In this case, the binding between quark and antiquark is mediated by the exchange of a single massless gluon, and the spectrum of bound states resembles that of an exotic atom composed of an electron and an antielectron (positron) bound electromagnetically in a Coulomb potential generated by the exchange of a massless photon (**Fig. 5***a*). Since the electron-positron atom is known as positronium, the heavy quark-antiquark atom has been called quarkonium. Two families of heavy quark-antiquark bound states, the ψ/J system composed of charmed quarks (Fig. 5*b*) and the Υ system made up of *b* quarks, have been discovered. Both have level schemes characteristic of atomic spectra, which have been analyzed by using tools of nonrelativistic quantum mechanics developed for ordinary atoms. The atomic analogy has proved extremely fruitful for studying the strong interaction. *See* CHARM; J/PSI PARTICLE; POSITRONIUM.

Lattice models. To deal with the existence and properties of the hadrons themselves, it is necessary to devise a new computational approach that does not break down when the interaction becomes strong. The most promising method has been the crystal lattice formulation of the theory. By localizing quarks on individual lattice sites, it is possible to use many of the Monte Carlo simulation techniques developed in statistical physics for the study of spin systems such as magnetic substances. Modern calculations include the influence of dynamical quark-antiquark pairs. Lattice quantum chromodynamics analyses support the color-confinement paradigm. *See* ISING MODEL; STATISTICAL MECHANICS.

Quark-gluon plasma. It seems likely that when hadronic matter is compressed to very great densities and heated to extremely high temperatures hadrons will lose their individual identities. When the hadronic bubbles overlap and merge, quarks and gluons may be free to migrate over great distances.

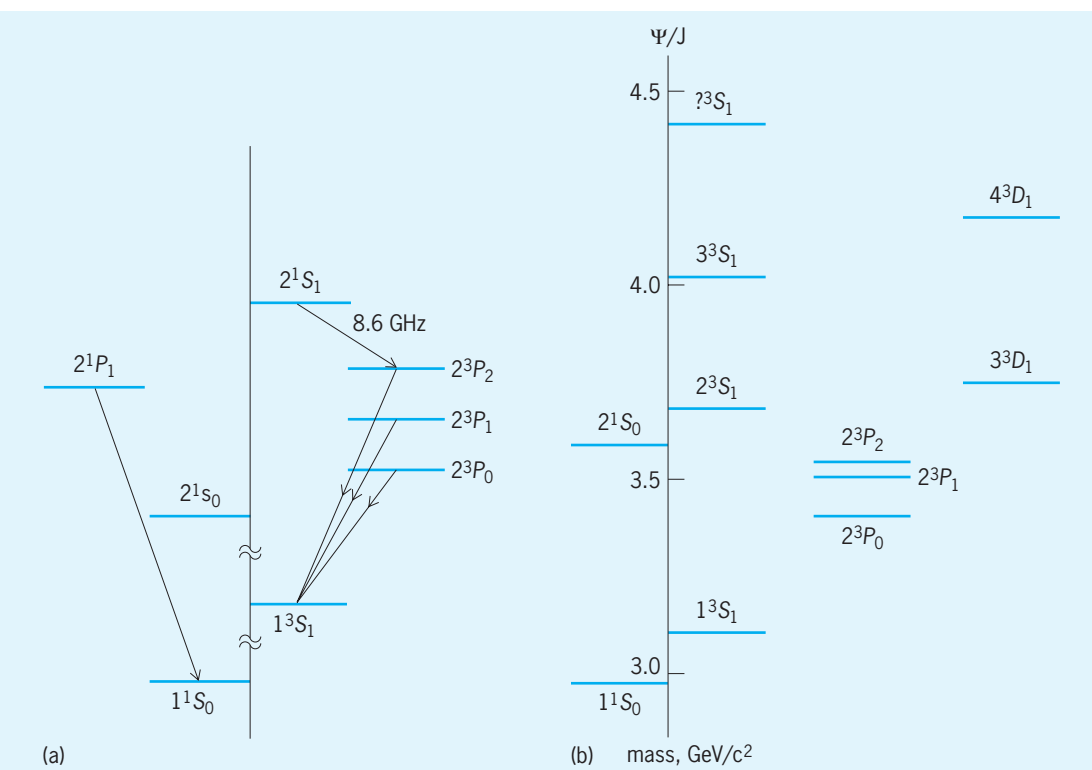

Fig. 5. Spectra of bound particle-antiparticle systems. (*a*) Positronium. (*b*) ψ/J family, known as charmonium.

The resulting new states of matter may exist in the cores of collapsing supernovae and neutron stars. The possibility of creating new quantum chromodynamic states of matter in the laboratory is under active experimental study in collisions of relativistic heavy ions. *See* NEUTRON STAR; QUARK-GLUON PLASMA; SUPERNOVA.

Grand unification. Because quantum chromodynamics and the unified theory of weak and electromagnetic interactions have the same gauge theory structure, it is appealing to speculate that all three interactions have a common origin in a single gauge symmetry. The construction of such grand unification theories is a very active area of theoretical research. *See* ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS; GLUONS; GRAND UNIFICATION THEORIES; QUANTUM FIELD THEORY; QUARKS; STANDARD MODEL; STRONG NUCLEAR INTERACTIONS. C. Quigg

Bibliography. I. J. R. Aitchison and A. J. G. Hey, *Gauge Theories in Particle Physics*, vol. 2, 3d ed., Taylor & Francis, London, 2002; R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics*, Cambridge University Press, 1996, paper 2003; I. Montvay and G. Münster, *Quantum Fields on a Lattice*, Cambridge University Press, 1994, paper 1997; T. Muta, *Foundations of Quantum Chromodynamics*, 2d ed., World Scientific, Singapore, 1998; Y. Nambu, The confinement of quarks, *Sci. Amer.*, 235(5):48–60, November 1976; C. Quigg, *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, Perseus/Westview, New York and Boulder, 1983, reprint 1997.

Quantum computation

Whereas classical computers perform operations on classical bits, which can be in one of two discrete states, 0 or 1, quantum computers perform operations on quantum bits, or qubits, which can be put into any superposition of two quantum states, $|0\rangle$ and $|1\rangle$. Quantum computers hold the promise to perform classically intractable tasks—computations that require astronomical time or hardware resources on the fastest classical computer—in minuscule time with minuscule resources. The building blocks of a quantum computer are atomic particles which, following the laws of quantum physics, behave like waves and exhibit interference phenomena as if they were in several locations at the same time. By equating different locations—for example, an electron in the lowest orbit or in an excited orbit of an atom—to binary digits 0 or 1, one may interpret the time-evolving state of the particles as executing several computations at the same time. One set of locations at a given time describes the result of one computation. Thus one atom can do two computations at once; two atoms can do four; three atoms can do eight. The challenge is to coerce the atoms to follow trajectories that amount to meaningful computations and to read out a definite result from the multitude of computations occurring in parallel. The control of trajectories is the hardware part of the challenge; the design of useful trajectories—algorithms that are superior to classical algorithms—is the software part of the challenge. *See* DIGITAL COMPUTER; SUPERCOMPUTER.

Principles. A classical computer manipulates strings of N classical bits, $(n_1, \ldots, n_N)$ with $n_j = 0$ or 1 ($j = 1, \ldots, N$), in such a way that intermediate states of the computation are also strings of classical bits. A quantum computer manipulates states of N two-level atoms, nuclear spins, or other entities, $|n_1, \ldots, n_N\rangle$, with $n_j = 0$ if the jth atom is in the ground state and $n_j = 1$ if it is in the excited state, in such a way that intermediate states are superpositions of the states $|n_1, \ldots, n_N\rangle$. The 2^N states $|n_1, \ldots, n_N\rangle$ (the computational basis) are product states in which each atom is in either the ground state or excited state, and n_j is called the value of the jth qubit; they represent the strings of classical bits. The superpositions include states in which an atom no longer has a sharp value of n_j (it has indefinite bit value), and states in which an atom no longer exists in a state separate from the other atoms (entangled states); both have no classical counterpart. A quantum computation starts with a product state $|n_1, \ldots, n_N\rangle$; allows the state to evolve according to the Schrödinger equation (1),

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle \tag{1}$$

with initial condition $|\psi(0)\rangle = |n_1, \ldots, n_N\rangle$ and time-dependent Hamiltonian $H(t)$ [energy operator] driving the coupled atoms; and ends with the measurement of the values of the qubits of the state $|\psi(t)\rangle$. The Hamiltonian generates the unitary time evolution operator $U(t)$ which takes the initial state into the final state according to $|\psi(t)\rangle = U(t)\ |\psi(0)\rangle$. The measurement transforms $|\psi(t)\rangle$ into the output state $|n'_1, \ldots, n'_N\rangle$ with probability $|\langle n'_1, \ldots, n'_N|\psi(t)\rangle|^2$. The output is probabilistic because quantum measurements are so. The computation $|n_1, \ldots, n_N\rangle \rightarrow |\psi(t)\rangle$ is a unitary transformation, hence reversible; the readout $|\psi(t)\rangle \rightarrow |n'_1, \ldots, n'_N\rangle$ is a projection (called the collapse of the wave function), hence irreversible. Thus, to perform a specific computation, one must drive the atoms with a specific Hamiltonian; to read out the result, one must send the atoms through a series of state detectors. *See* BIT; SCHRÖDINGER'S WAVE EQUATION.

Power of quantum computation. A quantum computer is more powerful than a classical computer for two reasons:

1. The quantum state space is much larger than the classical state space: N qubits can be in an infinite number of different states (any point on the unit sphere of the complex Hilbert space spanned by the 2^N basis vectors $|n_1, \ldots, n_N\rangle$); N classical bits can be in only 2^N different states (the points where the 2^N coordinate axes intersect the unit sphere; for $N = 1$: **Fig. 1**). If states on the unit sphere can be resolved to within accuracy ε ($\varepsilon \ll 1$), the sphere hosts a number of the order of $\varepsilon^{-(2N-1)}$ [written $O(\varepsilon^{-(2N-1)})$] distinct states. Thus a quantum computer can store and access an exponentially large number of states compared to a classical computer.

2. The quantum computer operates in a massively parallel way: If the initial state is the uniform superposition of all basis states, Eq. (2), the time evolution

$$|\psi(0)\rangle = 2^{-N/2} \sum_{n_1,\ldots,n_N=0,1} |n_1,\ldots,n_N\rangle \qquad (2)$$

computes simultaneously $U(t)|n_1,\ldots,n_N\rangle$ for all 2^N possible inputs $|n_1,\ldots,n_N\rangle$ by linearity of $U(t)$. The matrix element $\langle n'_1,\ldots,n'_N|U(t)|n_1,\ldots,n_N\rangle$ is the probability amplitude that the computation converts the input $|n_1,\ldots,n_N\rangle$ into the output $|n'_1,\ldots,n'_N\rangle$, along all possible classical computational paths in parallel (as given by Feynman's path integral). A classical computation can follow only a single path. The aim is to choose $U(t)$ so that computational paths of no interest cancel each other by destructive interference, and paths of interest add constructively, that is, to select relevant computations by quantum interference (**Fig. 2**). *See* FEYNMAN INTEGRAL; INTERFERENCE OF WAVES; SUPERPOSITION PRINCIPLE.

Logic gates. Any computation $U(t)$, also called a quantum circuit, can be approximated by sequential application of a finite set of unitary transformations that operate on only one or two qubits. Such transformations are called logic gates. A logic gate assigns to any input, consisting of a one-qubit or two-qubit state, a unique output, in terms of a one-qubit or two-qubit state. An example for such a universal set of quantum logic gates are the unitary transformations given in Eqs. (3) [Hadamard gate], (4) [T gate], and (5) [controlled-not gate], where $|n_j\rangle\langle n'_j|$ acts only on

$$H_j = (|0_j\rangle\langle 0_j| + |0_j\rangle\langle 1_j| + |1_j\rangle\langle 0_j| - |1_j\rangle\langle 1_j|)/\sqrt{2} \qquad (3)$$

$$T_j = |0_j\rangle\langle 0_j| + e^{i\pi/4}|1_j\rangle\langle 1_j| \qquad (4)$$

$$C_{jk} = |0_j 0_k\rangle\langle 0_j 0_k| + |0_j 1_k\rangle\langle 0_j 1_k| + |1_j 0_k\rangle\langle 1_j 1_k| + |1_j 1_k\rangle\langle 1_j 0_k| \qquad (5)$$

qubit j, and $|n_j n_k\rangle\langle n'_j n'_k|$ only on qubits j and k. They correspond to the logic gates in a classical computer, but are reversible, unlike the classical "and" and "exclusive or" gates, which are irreversible. A logic gate is called reversible if one can reconstruct the input when one knows the output. The Hadamard gate transforms the states $|0_j\rangle$ and $|1_j\rangle$ into the superpositions $(|0_j\rangle \pm |1_j\rangle)/\sqrt{2}$; the T gate shifts the phase of the excited state relative to the ground state by $n/4$; and the controlled-not gate flips the target qubit k if and only if the control qubit j is in the excited state. The three gates are the analog of an optical beam splitter, which splits an incoming wave into a transmitted wave and a diffracted outgoing wave; a phase shifter (refractive medium); and a conditional half-wave plate, respectively. Only the Hadamard gate creates multiple computational paths; the other two transform a single basis state into a single basis state. To approximate a general $U(t)$ to within accuracy ε requires $O(4^N N^2[\ln(4^N N^2/\varepsilon)]^s)$ gates, where $s \approx 2$ (this is the Solovay-Kitaev theorem); the leading factor 4^N is set by the number of matrix elements of $U(t)$. For special computations, however, often a much smaller number of gates, independent of ε, suffices. *See* LOGIC CIRCUITS.

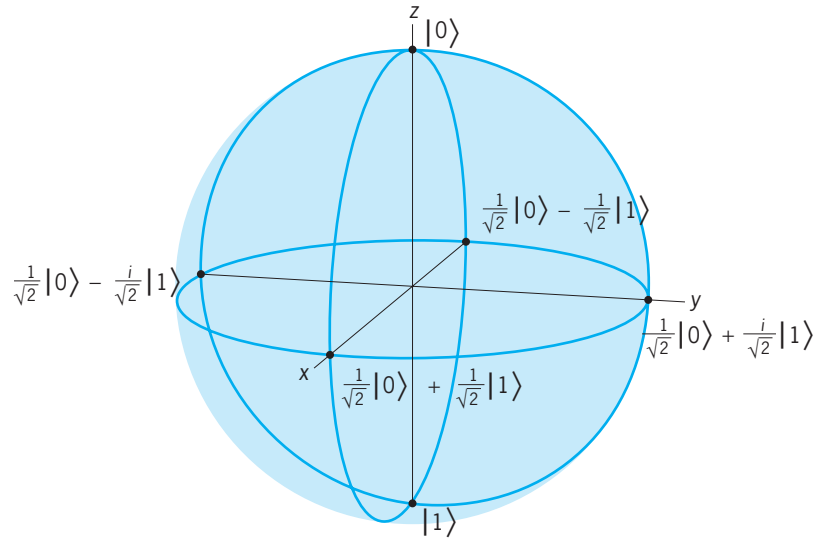

Fig. 1. A classical bit can be in only state $|0\rangle$ or $|1\rangle$ (the "north" and "south" poles of the sphere). A quantum bit can be in any superposition $a|0\rangle + b|1\rangle$, where $|a|^2 + |b|^2 = 1$ (an arbitrary point on the sphere). The sphere is called the Bloch sphere, and the polar and azimuthal angles, θ and φ, of a point on the sphere are given by $a = \cos(\theta/2)$ and $b = \exp(i\varphi)\sin(\theta/2)$.

Fast quantum algorithms. A widely used class of encryption systems for public transmission of sensitive data derive their security from the difficulty of factoring a large, publicly transmitted integer. The fastest known classical algorithm factors an N-bit number in time $O[\exp(\text{const} \times N^{1/3}(\ln/\ln N)^{2/3})]$. Peter Shor's quantum algorithm for the same task requires only time $O[N^2(\ln N)(\ln N)]$. The key task in Shor's algorithm is to find the period of a function related to the number to be factored, and the algorithm does this using the discrete Fourier transform: The values at which the transform does not vanish yield the period—similarly to the way in which Bragg peaks in x-ray diffraction, at which the x-ray intensity does not vanish, yield the crystal periodicity. The quantum fast Fourier transform, with no approximation, requires $O(N^2)$ operations, which is much less than the $O(2^N N)$ operations in the classical fast Fourier transform and much less than the Solovay-Kitaev bound. The period is obtained from $O(\ln N)$ repeated measurements of the qubits of the state representing the

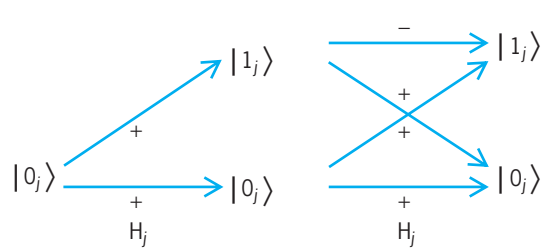

Fig. 2. Application of the Hadamard gate, Eq. (3), to state $|n_j\rangle$ creates two computational paths, one leading to state $|0_j\rangle$ with probability amplitude $1/\sqrt{2}$, and the other to $|1_j\rangle$ with amplitude $1/\sqrt{2}$ if $n_j = 0$, and amplitude $-1/\sqrt{2}$ if $n_j = 1$. Application of the Hadamard gate twice to $|0_j\rangle$ creates four paths, leading to $|0_j\rangle$ with amplitude $(+1)(+1)/2 + (+1)(+1)/2 = 1$, and to $|1_j\rangle$ with amplitude $(+1)(+1)/2 + (+1)(-1)/2 = 0$ (cancellation of trajectories leading to $|1_j\rangle$).

Fourier transform. *See* CRYPTOGRAPHY; FOURIER SERIES AND TRANSFORMS; X-RAY DIFFRACTION.

Other quantum algorithms that outperform classical algorithms by orders of magnitude are Lov K. Grover's algorithm for "finding a needle in a haystack" (search of an item in a database of 2^N items) in $O(2^{N/2})$ instead of $O(2^N)$ steps; estimation of the median and mean of 2^N items to precision ε in $O(N/\varepsilon)$ instead of $O(N/\varepsilon^2)$ steps; search of the minimum of a function sampled at 2^N points in $O(2^{N/2})$ instead of $O(2^N)$ steps; search of two distinct pre-images giving the same image of a two-to-one function sampled at 2^N points, in $O(2^{N/3})$ instead of $O(2^{N/2})$ steps; the Deutsch-Jozsa algorithm to determine if either [1] 2^N numbers are either all 0 or all 1 (a constant function), or [2] half of these numbers are 0 and half are 1 (a balanced function), in one instead of up to $2^{N-1} + 1$ steps; and various allocation tasks and game-theoretic strategies. For $N = 1$, the Deutsch-Jozsa algorithm can be visualized as putting two slabs of material, which may be thin ("0") or thick ("1"), into the apertures of a double-slit apparatus and measuring the interference pattern of a light wave passing through the two slits. If the pattern is unchanged compared to the empty apparatus, the slabs have the same thickness (constant function); or else they have different thickness (balanced function).

Quantum error correction. Noise from imperfect computer operation poses no fundamental barrier to large-scale computations. A quantum error-correction code encodes the N logical qubits into N' carrier qubits ($N' > N$); runs the carrier qubits through a group of accordingly encoded logic gates; transforms the noisy state by appropriate projection operators (which perform syndrome measurements to detect errors) and unitary operators (which recover the original carrier qubits) into an error-corrected state; and feeds the state into the next group of gates. Such periodic error correction prevents accumulation of errors in the state. At the end of the computation, the carrier qubits are decoded. The encoding spreads the state of the N logical qubits over all N' carrier qubits so that, when the syndromes (which provide information about errors) are measured, no information about the state of the logical qubits is revealed: The projections preserve superpositions of the logical qubits, and the state of the carrier qubits is highly entangled even if the logical-qubit state is not. Remarkably, a discrete set of corrections can correct a continuum of errors. For a code to correct any error on any M carrier qubits, a necessary condition is $N' \geq 4M + N$ (the Knill-Laflamme bound), and a sufficient condition for large N is given by Eq. (6) [the Gilbert-Varshamov bound]. For example, a code exists that encodes 1 logical qubit ($N = 1$) into 5 carrier qubits ($N' = 5$) and corrects any error on any 1 carrier qubit ($M = 1$) with 16 pairs of syndrome measurements and recovery operations (thus achieving equality in the Knill-Laflamme bound).

$$N/N' < 1 - 2[-x\log_2 x - (1-x)\log_2(1-x)]_{x=2M/N'} \tag{6}$$

Different codes require different encodings of gates, and of interest are encodings for which an error in an input carrier qubit or in the gate operation propagates only to a small number of output carrier qubits. Specifically, an encoded gate is called fault-tolerant if a failure with probability p of any single component (for example, one of the 10 "wires" feeding two logical qubits, each encoded by 5 carrier qubits, into a controlled-not gate) introduces an error in 2 or more carrier qubits in any logical output qubit with probability cp^2 at most, with c a constant and p small. Such a gate, when followed by error correction with $M = 1$, yields an error-free output with probability $1 - cp^2$; that is, it reduces the error probability from p to cp^2 if $p < 1/c$. Fault-tolerant Hadamard, T, and controlled-not gates exist with $c \approx 10^4$ and $d \approx 10^2$, where d is the number of operations on carrier qubits needed to encode and error-correct the gate. By hierarchical fault-tolerant encoding of all gates, a computation involving L gates can be carried out to within accuracy ε using only $O(L[\ln(L/\varepsilon)]^{\log_2 d})$ operations on carrier qubits if $p < 1/c$. This is the threshold theorem for quantum computation. Thus, if the noise in individual carrier qubits is low enough, so that $p < 1/c \approx 10^{-4}$, arbitrarily large computations can be performed because the overhead for error correction grows only polynomial-logarithmically with the size of the computation, L. *See* FAULT-TOLERANT SYSTEMS; INFORMATION THEORY.

Experimental state of the art. A remarkable array of experimental realizations of quantum computing devices and implementations of algorithms have been achieved. As of 2006, records with respect to the number of qubits that can be controlled and prepared in well-defined states are a 12-qubit nuclear magnetic resonance device (NMR, four ^{1}H, six ^{13}C, and two ^{14}N nuclei in *l*-histidine, each in its spin-up or spin-down state, driven by radio-frequency pulses), an 8-qubit ion-trap device (IT, eight $^{40}Ca^+$ ions in a linear electromagnetic trap, each in the electronic ground state or an excited state, driven by optical pumping), a 3-qubit device based on cavity quantum electrodynamics [CQED, three rubidium (Rb) atoms, each in one of two Rydberg states and coupled to a cavity mode with zero or one photon, driven by microwave pulses], various 2-qubit devices based on superconducting circuits (SC), and a 1-qubit quantum-dot device [QD, two gallium arsenide (GaAs) quantum dots in a spin singlet or triplet state, driven by nanosecond voltage pulses]. IT, SC, and QD devices have the potential for scalable or solid-state architectures (less-than-exponential growth of the physical hardware necessary to manipulate large numbers of qubits). Successful computations carried out include Shor's algorithm on a 7-qubit NMR device, factoring the number 15 with a sequence of 300 radio-frequency pulses; the Deutsch-Jozsa algorithm on a 5-qubit NMR device; finding the order of a permutation on a 5-qubit NMR device; correction of any one-qubit error on a 5-qubit NMR device; spin-flip error correction on a 3-qubit IT device; and Grover's algorithm on a 2-qubit IT device. *See* CAVITY RESONATOR; NUCLEAR MAGNETIC RESONANCE (NMR); PARTICLE TRAP; QUANTIZED ELECTRONIC STRUCTURE

(QUEST); QUANTUM ELECTRODYNAMICS; RYDBERG ATOM; SUPERCONDUCTIVITY.

General hardware requirements. Any experimental realization faces three challenges:

1. The device must be able to control the state of each qubit separately, while allowing neighboring qubits to interact with each other during the operation of 2-qubit gates. Control of individual qubits is achieved by different chemical shifts in NMR, by laser beams driving ions spatially separated by tens of micrometers in IT, and by pulses addressing successive atoms traveling at spatial separation of several centimeters in CQED. Qubits interact via spin-spin coupling in NMR, via a shared phonon state of the ions in IT, and via electric dipole coupling between each atom and the cavity mode in CQED.

2. The system must be switchable, so that interactions executing a prescribed sequence of gate operations [the Hamiltonian $H(t)$] can be turned on and off by external control. Switching is done by magnetic field pulses in NMR, and by laser pulses in IT and CQED.

3. The computer must be well isolated from the environment so that the decoherence time t_d, the time at which the computer and the environment depart significantly from a product state (entanglement of the computer with the environment), is long compared to t_g, the time it takes to operate a single gate. At time t_d, a generic qubit state $a|0_j\rangle + b|1_j\rangle$ will have degraded into the mixture with the density matrix operator $|a|^2|0_j\rangle\langle 0_j| + |b|^2|1_j\rangle\langle 1_j|$, which no longer contains the interference terms necessary for quantum computation. Good isolation is provided by long spin-spin and spin-lattice relaxation times in NMR ($t_d = 10^{-2}$ to 10^8 s, $t_g = 10^{-6}$ to 10^{-3} s), long-lived states and stable trap and laser operation in IT ($t_d = 10^{-1}$ to 10^0 s, $t_g = 10^{-7}$ to 10^{-5} s), low spontaneous emission rate of Rydberg states and low photon escape rate from the cavity in CQED ($t_d = 10^{-3}$ to 10^0 s, $t_g = 10^{-5}$ to 10^{-4} s), and spin-echo pulses in QD ($t_d = 10^{-6}$ to 10^{-5} s, $t_g = 10^{-8}$ to 10^{-7} s). *See* DENSITY MATRIX; QUANTUM THEORY OF MEASUREMENT.

But the condition $p < 1/c$ in the threshold theorem is more demanding than $t_g < t_d$: Let $|\psi_t(t)\rangle$ be the state of the computer and environment interacting with each other, starting from a product state at $t = 0$; let $|\psi_n(t)\rangle$ be the same state, but noninteracting (product state at all times); and assume the overlap $F(t) = |\langle\psi_t(t)|\psi_n(t)\rangle|$ (called the fidelity) decays exponentially with time. The decay constant is the decoherence time, $F(t) = \exp(-t/t_d)$, and the threshold condition requires $1 - \exp(-2t_g/t_d) < 1/c$ [because $p = 1 - F^2(t_g)$], giving $t_g < t_d/(2c)$ for large c. Thus the computer should remain isolated from the environment for about 10^4 gate operations for sustained computations. Present devices can execute about 300 operations and thus are far from this goal.

Outlook. Numerous other approaches have been proposed and are under investigation, for quantum computers with much longer decoherence times and control over many more qubits (a long-term goal is 200 qubits). These include solid-state NMR; solid-state CQED; electron spin resonance of donor atoms in semiconductors; electronic excitation in quantum dots; ion traps with hot ions; electrons trapped on superfluid helium films; neutral atoms, including Bose-Einstein condensates, trapped in an optical lattice; linear optics (beam splitters, phase shifters, and single-photon sources); and topological quantum computing, using braided world lines of quasi-particles in the fractional quantum Hall effect as decoherence-free qubits. With so many lines of developments, quantum computing holds a rich future. *See* BOSE-EINSTEIN CONDENSATION; ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; HALL EFFECT; LIQUID HELIUM; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS. Peter Pfeifer

Bibliography. G. P. Berman et al., *Introduction to Quantum Computers*, World Scientific, Singapore, 1998; H. O. Everitt (ed.), *Experimental Aspects of Quantum Computing*, Springer, New York, 2005; G. Greenstein and A. G. Zajonc, *The Quantum Challenge*, 2d ed., Jones and Bartlett, Boston, 2006; A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation*, American Mathematical Society, Providence, 2002; M. Nakahara et al. (eds.), *Physical Realizations of Quantum Computing: Are the DiVicenzo Criteria Fulfilled in 2004?*, World Scientific, Hackensack, NJ, 2006; M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000; W.-H. Steeb and Y. Hardy, *Problems and Solutions in Quantum Computing and Quantum Information*, 2d ed., World Scientific, Hackensack, NJ, 2006.

Quantum electrodynamics

The field of physics that studies the interaction of electromagnetic radiation with electrically charged matter within the framework of relativity and quantum mechanics. It is the fundamental theory underlying all disciplines of science concerned with electromagnetism, such as atomic physics, chemistry, biology, the theory of bulk matter, and electromagnetic radiation.

Quantum mechanics and relativity. Efforts to formulate quantum electrodynamics (QED) were initiated by P. A. M. Dirac, W. Heisenberg, and W. Pauli soon after quantum mechanics was established. The first step was to remedy the obvious shortcoming of quantum mechanics: that it applies only to the case where particle speeds are small compared with that of light, c. This led to Dirac's discovery of a relativistic wave equation, in which the wave function has four components and is multiplied by certain 4×4 matrices. His equation incorporates in a natural manner the observed electron-spin angular momentum, which implies that the electron is a tiny magnet. The strength of this magnet (magnetic moment) was predicted by Dirac and agreed with observations then available. A detailed prediction of the hydrogen spectrum was also in good agreement with experiment. *See* ATOMIC STRUCTURE AND SPECTRA; ELECTRON SPIN; MATRIX THEORY.

In order to go beyond this initial success and calculate higher-order effects, however, the interaction of charge and electromagnetic field had to be treated

dynamically. To begin with, a good theoretical framework had to be found for describing the wave-particle duality of light, that is, the experimentally well-established fact that light behaves like a particle (photon) in some cases but like a wave in others. Similarly, the electron manifests wave-particle duality, another observed fact. Once this problem was settled, the next question was how to deal with the interaction of charge and electromagnetic field. It is here that the theory ran into severe difficulties. Its predictions often diverged when attempts were made to calculate beyond lowest-order approximations. This inhibited the further development of the theory for nearly 20 years. Stimulated by spectroscopic experiments vastly refined by microwave technology developed during World War II, however, S. Tomonaga, R. P. Feynman, and J. Schwinger discovered that the difficulties disappear if all observable quantities are expressed in terms of the experimentally measured charge and mass of the electron. With the discovery of this procedure, called renormalization, quantum electrodynamics became a theory in which all higher-order corrections are finite and well defined. *See* NONRELATIVISTIC QUANTUM THEORY; PHOTON; QUANTUM MECHANICS; RELATIVISTIC QUANTUM THEORY; RELATIVITY; WAVE MECHANICS.

Scope of QED. Quantum electrodynamics turned out to be much more than a marriage of quantum mechanics and relativity. The wave-particle duality is now fully incorporated. Both electrons and photons are treated as quantized fields. The theory makes it possible to handle not only the dynamics of electrons but also the creation and destruction of electron-positron pairs, a phenomenon unheard of in classical physics. What makes quantum electrodynamics particularly intriguing is that it is the first physical theory ever developed that has no obvious intrinsic limitation and describes physical quantities from first principle. The predictive power of quantum electrodynamics may be appreciated by considering what is needed to describe the hydrogen atom by ordinary quantum mechanics to the same precision as that of quantum electrodynamics. For this purpose, a large number of new terms must be added to the potential-energy operator of quantum mechanics. Unfortunately, quantum mechanics itself is powerless in selecting these terms. Meanwhile, quantum electrodynamics accounts for everything with just one term. *See* ELECTRON-POSITRON PAIR PRODUCTION.

However, nature accommodates forces other than the electromagnetic force, such as those responsible for radioactive disintegration of heavy nuclei (called the weak force) and the force that binds the nucleus together (called the strong force). A theory called the standard model, has been developed which unifies the three forces and accounts for all experimental data from very low to extremely high energies. This does not mean, however, that quantum electrodynamics fails at high energies. It simply means that the real world has forces other than electromagnetism. Predictions of quantum electrodynamics itself have thus far been confirmed by experiments without exception. Breakdown of quantum electrodynamics in any form would imply that there is something in nature besides relativity and quantum mechanics, the two pillars of modern physics. Since the standard model is an adaptation of quantum electrodynamics to include other forces, it would also fail if quantum electrodynamics fails. This is why it is important to keep examining the validity of quantum electrodynamics. *See* FUNDAMENTAL INTERACTIONS; STANDARD MODEL; STRONG NUCLEAR INTERACTIONS; WEAK NUCLEAR INTERACTIONS.

That it has been possible to examine quantum electrodynamics so closely is due to the fortuitous circumstance that the interaction of electromagnetic field and charge is relatively weak and its effect can be evaluated systematically as a small deviation from the noninteracting case. Therefore, the properties of noninteracting (or free) fields will be examined before interactions are considered. *See* PERTURBATION (QUANTUM MECHANICS).

Free electromagnetic field. The simplest electrodynamic system is that of radiation in an empty space, free of any charge, described classically by the Maxwell equations (which are consistent with relativity). In classical physics, the radiation is regarded as confined in a box and expressed as a superposition of normal modes, analogous to the dynamics of a string on a violin, each mode having a well-defined wavelength and oscillating with a well-defined frequency. The complete field can be constructed by specifying the amplitude of each normal mode and adding up the contributions of all modes. In other words, the dynamics of the electromagnetic field is equivalent to that of normal-mode amplitudes. *See* CAVITY RESONATOR; ELECTROMAGNETIC RADIATION; MAXWELL'S EQUATIONS; VIBRATION.

In quantum theory, any normal mode of the radiation field can be found only in a discrete set of quantum states with energy given by Eq. (1), where

$$E_n = h\nu\left(n + \frac{1}{2}\right) \tag{1}$$

h is Planck's constant, ν is the oscillator frequency, and n is restricted to integers 0, 1, 2, ... This state is said to consist of n photons each of energy $h\nu$. The total energy in the field is the sum of energies of all photons. The state with no photons present, called the vacuum, is characterized by assigning all oscillators to the lowest energy state $n = 0$. The total energy of such a state is the sum over all the normal modes of the quantity $^1/_2 h\nu$. This sum is not zero; in fact, it is infinite. This infinity is interpreted as a consequence of quantum fluctuations (the ephemeral presence of photons allowed by the uncertainty principle). It causes no problem since what is observable is the difference of energies between states, in which the infinities cancel out. *See* HARMONIC OSCILLATOR; UNCERTAINTY PRINCIPLE.

Free electron field. As mentioned above, Dirac's equation successfully explains many aspects of electrons and atoms. On closer scrutiny, however, it has some strange features. As noted above, his wave function has four components. While the first

two components correspond to the usual positive energy states, the last two are associated with the states of negative energy, which are unknown in classical physics and seem to have pathological properties. Dirac found, however, that they can be reinterpreted as positive energy states of the electron's antiparticle, the positron, which behaves exactly like an electron except that its charge has opposite sign. This bold prediction was strikingly successful with the subsequent discovery of positron. Dirac's successful reinterpretation required, however, that his "wave function" be treated not as a probability amplitude analogous to the Schrödinger wave function but as a quantized field with infinitely many degrees of freedom. Just like the electromagnetic field, the electron field can be written as a superposition of quantized normal modes, except that each mode can accommodate at most one electron or one positron in accord with the Pauli exclusion principle familiar in chemistry. The state with the lowest energy is one in which neither electron nor positron is present, analogous to the state of the electromagnetic field with no photons present, discussed above. Naturally, such a state is also called a vacuum. Like the photon vacuum, the electron vacuum is filled with fluctuating electrons and positrons. This means that quantum electrodynamics must always deal with (infinitely) many particles, even if the states of just a few particles are under consideration. This is where quantum electrodynamics differs strongly from the ordinary quantum mechanics. *See* ANTIMATTER; EXCLUSION PRINCIPLE.

Interaction of light and electron. The classical Maxwell equations contain in a well-defined way the current and charge of particles as the source of radiation. The quantum-electrodynamics version of the charge-current density is constructed from the electron field and its hermitian conjugate, and is proportional to a quantity, e_0, referred to as the bare charge of an electron. The photon field is represented by a four-vector potential. The interaction energy can be written as the integral over space of a product of the components of these two fields.

The equation of motion that follows from this expression for the interaction energy is nonlinear, and has thus resisted exact solutions. Fortunately, the dynamics of the coupled electron-photon system can often be described to a good approximation by the first few terms of a power series expansion in a small dimensionless parameter. The calculation of expansion terms is facilitated greatly by a pictorial representation (**Fig. 1**) of the interaction energy, where two straight lines (representing electrons or positrons) and one wavy line (representing a photon) meet at a space-time point, called a vertex. One such diagram (Fig. 1*a*) represents a situation in which an incident electron (entering from the bottom of the picture) absorbs an incident photon (also entering from the bottom) to produce a final electron (leaving from the top of the picture). Here a convention is adopted in which time proceeds upward from the bottom toward the top of page. Also, electrons and positrons are distinguished by upward

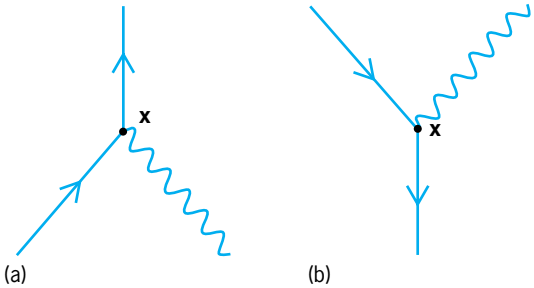

Fig. 1. Diagrams of photon absorption and emission. (*a*) Absorption of a photon (represented by a wavy line) by an electron (represented by a solid line with an upward arrow). (*b*) Emission of a photon by a positron (represented by a solid line with a downward arrow). The photon meets the electron or positron at a space-time point, *x*, called a vertex.

and downward arrows, respectively, on both initial and final solid lines. Another such diagram (Fig. 1*b*) corresponds to emission of a photon by a positron, and six more diagrams can be drawn corresponding to other situations such as creation of an electron-positron pair by a photon.

Feynman diagrams. Actually, the processes discussed above (Fig. 1) cannot occur by themselves since they violate the law of energy-momentum conservation. Nevertheless, they do serve as building blocks out of which allowed processes are constructed. All these processes can be represented by a set of diagrams obtained by connecting the vertices to each other according to a certain rule. In terms of these pictures, referred to as Feynman diagrams, a given process can be readily visualized. Feynman introduced a set of rules (called Feynman rules) which make it possible to write expressions called transition amplitudes. The absolute square of the sum of transition amplitudes gives the probability with which this process will take place. In many problems, it suffices to use only a few of these diagrams to obtain results with sufficient precision.

In general, Feynman diagrams contain several solid lines or wavy lines which originate in the distant past (the bottom of the figure) and represent incident particles. They will subsequently interact (at several vertices). After that, other solid or wavy lines emerge and move toward the distant future (the top of the figure). For example, (**Fig. 2**) an electron and a photon can enter the diagram, interact at two space-time points, **x** and **y**, and then leave. The line originating in the past or emerging to the future is referred

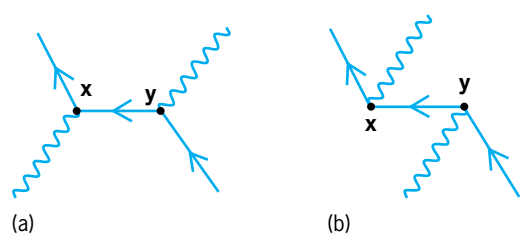

Fig. 2. Diagrams representing scattering of a photon on an electron caused by an exchange of a virtual electron. The electron and photon interact at space-time points *x* and *y*.

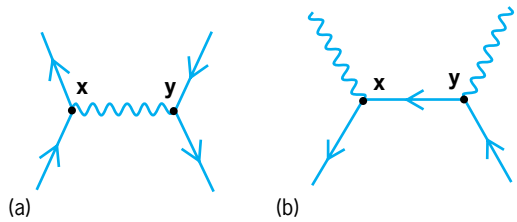

Fig. 3. Diagrams representing encounters of an electron and a positron. (*a*) Scattering of a positron and an electron caused by an exchange of a virtual photon. (*b*) Mutual annihilation of an incident positron and electron into two photons.

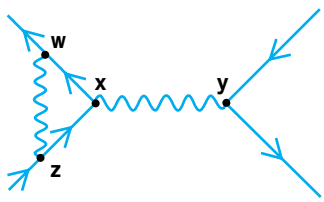

Fig. 4. Example of a diagram of order e_0^4, involving two virtual photons and two virtual electrons. The diagram has a closed path of internal lines connecting the vertices *x*, *z*, and *w*.

to as an external line. The line connecting **x** and **y** is referred to as internal or virtual electron line. To each internal line corresponds a function of **x**−**y**, called an electron propagator. It provides a measure of the likelihood that an electron at point **x** will traverse freely to point **y**. Internal wavy lines can also occur, which correspond to virtual photons or photon propagators. An electron and a positron may undergo scattering (**Fig. 3***a*), or they may interact and subsequently disappear (Fig. 3*b*), producing two photons in the final state. Since the corresponding pictures have two space-time points at which interaction occurs, the amplitude has a strength proportional to e_0^2 (where e_0 is the bare charge that appears in the charge-current density). There are also diagrams proportional to e_0^4, including some that give the so-called radiative corrections to electron-positron scattering (**Fig. 4**). Although such radiative corrections may be less important than diagrams proportional to e_0^2 (Fig. 3*a*) because of the extra two powers of e_0, their contribution must be included if a comparison with precise experimental results is desired. Another feature of Feynman diagrams is that they may have a closed path of internal lines connecting vertices (Fig. 4). Whenever a closed path occurs, the amplitude contains a four-dimensional integral. The situation becomes very complicated when there are multiple closed paths. *See* FEYNMAN DIAGRAM; PROPAGATOR (FIELD THEORY).

Renormalization. It turns out that some of the integrals associated with closed paths diverge; that is, they are infinite. The simplest examples of divergent integrals are given by the self-energy and vertex correction diagrams (**Fig. 5**). Feynman diagrams containing any self-energy or vertex correction diagrams as subdiagrams are divergent. These are the divergences that initially hampered the formulation of quantum electrodynamics until the concept of renormalization was introduced. The basic idea of renormalization is as follows. For any theory, say quantum electrodynamics, there are certain parameters such as electron mass m_0 (called bare mass) and electron charge e_0 (called bare charge) which are needed to describe the theory. In the absence of any interaction, only m_0 is present. When an interaction is included a new parameter, e_0, appears. Now m_0 and e_0 are presumed to be the values that these parameters would have if the interaction was reduced to zero. They are not the values of the electron mass or charge which an experimentalist would determine in some measurement. When Feynman diagrams are evaluated, they are expressed in terms of the parameter m_0 and e_0. However, this is clearly unsatisfactory since they are not measured or "physical" values of the parameters. Thus m and e, the measured values, must be introduced instead. Now, clearly m and e must include the effect of interaction since the experimentalist cannot turn them off. In order to express answers in terms of the measured quantities, it is necessary to replace m_0 and e_0 by writing them as m and e plus correction terms. This leads to some new terms to be computed. These new terms, presumed to be divergent in perturbation theory, now bring about a cancellation of the infinite results when they are expressed as functions of physical mass m and physical charge e. The process of renormalization of mass and charge, although not entirely satisfactory from a mathematical point of view, leads to predictions that agree very well with experiment. *See* RENORMALIZATION.

High-precision tests. High-precision tests have provided excellent confirmation for the validity of the renormalization theory of quantum electrodynamics. In the high-energy regime, the most extensive tests of quantum electrodynamics have been carried out by using electron-positron colliding-beam facilities at various high-energy physics laboratories. These measurements have confirmed the predictions of quantum electrodynamics at center-of-mass energies up to 1.8×10^{11} electronvolts (180 GeV). The uncertainty principle implies that this is equivalent to saying that quantum electrodynamics is valid down to about 10^{-17} meter, a distance 100 times shorter than the radius of the proton. This is remarkable in view of the fact that quantum electrodynamics was tested to only 10^{-10} m when it was first formulated. At energies around 90 GeV, the contribution

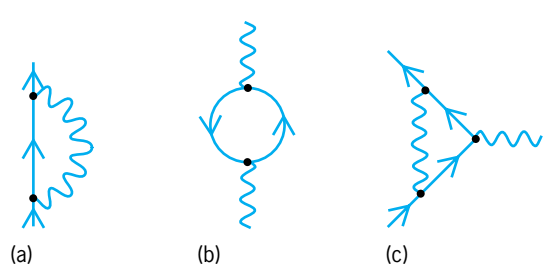

Fig. 5. Examples of divergent diagrams. (*a*) Electron self-energy diagram. (*b*) Photon self-energy diagram. (*c*) Vertex correction diagram.

of the weak boson Z^0 dominates those of quantum electrodynamics. As was mentioned above, however, this does not imply the failure of quantum electrodynamics at short distances. *See* INTERMEDIATE VECTOR BOSON; PARTICLE ACCELERATOR.

High-precision tests of quantum electrodynamics have also been carried out at low energies by using various simple atomic systems. The most accurate is that of the measurement of the magnetic moment of the electron, or the gyromagnetic ratio g, the ratio of spin and rotation frequencies. According to the Dirac equation, g must be exactly equal to 2. However, it was found in a 1947 experiment that g is larger than 2 by about 0.1%. Schwinger showed that this is due to the radiative correction of the electron in a magnetic field. The agreement of his calculation with experiment was one of the crucial successes of the then newly developed theory of quantum electrodynamics. By 1987, measurements of $a_e \equiv (g - 2)/2$, utilizing radio-frequency resonance transitions between quantized levels of an electron in a magnetic field, reached a precision of 3.7 parts in 10^9, as given in Eqs. (2). These measurements show that the

$$\begin{aligned} a_{e^-}^{\text{exp}} &= 0.001\,159\,652\,188\,4(43) \\ a_{e^+}^{\text{exp}} &= 0.001\,159\,652\,187\,9(43) \end{aligned} \tag{2}$$

electron (e^-) and its antiparticle, the positron (e^+), have identical magnetic moments, providing an example of the validity of a general theorem of quantum field theory, including quantum electrodynamics, called the CPT theorem. (The numeral 43 in the parentheses represents measurement uncertainty of $\pm 43 \times 10^{-13}$.) *See* CPT THEOREM.

In 2006, the precision of the measurement of a_e was improved significantly by a substantial refinement of the experiment, yielding the value given in Eq. (3), which is about six times more accurate than

$$a_{e^-}^{\text{exp}} = 0.001\,159\,652\,180\,85(76) \tag{3}$$

the 1987 measurement quoted in Eqs. (2).

Quantum electrodynamics calculations involving many Feynman diagrams provide the coefficients for the expansion of a_e in powers of the ratio α/π, where α is the fine-structure constant, as given in Eq. (4).

$$a_e = A^{(2)}(\alpha/\pi) + A^{(4)}(\alpha/\pi)^2 + A^{(6)}(\alpha/\pi)^3 + \cdots \tag{4}$$

The coefficients $A^{(2)}$, $A^{(4)}$, and $A^{(6)}$ are known analytically; their numerical values are given in Eqs. (5).

$$\begin{aligned} A^{(2)} &= 0.5 \\ A^{(4)} &= -0.328\,478\,965\,579\ldots \\ A^{(6)} &= 1.181\,241\,456\,587\ldots. \end{aligned} \tag{5}$$

Thus far the value of $A^{(8)}$ has been obtained only by numerical integration, yielding Eq. (6), where

$$A^{(8)} = -1.7283(35) \tag{6}$$

35 is the uncertainty in the last two digits. Including contributions of small quantum electrodynamics terms and non-quantum electrodynamics effects, and using the most precise value of the fine-structure constant available at present (2006), given in Eq. (7),

$$\alpha^{-1} = 137.035\,998\,78(91) \tag{7}$$

one obtains the theoretical prediction of a_e given in Eq. (8), in good agreement with experiment, where

$$a_e^{\text{theory}} = 0.001\,159\,652\,188\,70\,(40)\,(771) \tag{8}$$

40 is the remaining theoretical uncertainty and 771 comes from the uncertainty in the measurement of the fine-structure constant α quoted in Eq. (7). The gyromagnetic ratio g is thus correctly predicted by quantum electrodynamics to 12 significant digits. The comparison of a_e^{exp} with a_e^{theory} presented here constitutes the most precise confrontation of any experiment with a theoretical prediction in the history of science. Furthermore, it provides a new value of the fine-structure constant, given by Eq. (9),

$$\alpha^{-1}(g-2) = 137.035\,999\,710\,(96) \tag{9}$$

with a relative uncertainty of 0.7 part per billion (10^9), which is an order of magnitude more precise than any other measurement, including Eq. (7). Further improvement in the non-quantum electrodynamics measurement of α is crucial in pushing the test of the validity of quantum electrodynamics further. At present, other low-energy tests of quantum electrodynamics are less accurate, but their precision is improving with the development of the frequency comb technique and atom interferometry in laser spectroscopy. *See* ELECTRON SPIN; FREQUENCY COMB; LASER SPECTROSCOPY; QUANTUM FIELD THEORY.

Toichiro Kinoshita

Bibliography. J. Bjorken and S. Drell, *Relativistic Quantum Fields*, 1965; J. Bjorken and S. Drell, *Relativistic Quantum Mechanics*, 1965; R. P. Feynman, *QED: The Strange Theory of Light and Matter*, rev. ed., 2006; C. Itzykson and J.-B. Zuber, *Quantum Field Theory*, 1980, paperback, 2006; T. Kinoshita, *Quantum Electrodynamics*, 1990; J. Schwinger (ed.), *Selected Papers on Quantum Electrodynamics*, 1958; G. Sterman, *An Introduction to Quantum Field Theory*, 1993; S. Weinberg, *The Quantum Theory of Fields*, vol. 1: *Foundations*, 1995.

Quantum electronics

Quantum electronics is concerned with the interaction of electromagnetic radiation and matter and with the use of radiation-matter interactions for applications in devices such as lasers and masers. The term "quantum electronics" for this field was coined in 1958 or 1959 by Charles H. Townes, who invented the maser, as this field in its early days brought together ideas from electrical engineering and quantum physics in the development of the maser and the laser. The field has grown immensely, and today quantum electronics is concerned not just with masers and lasers but with a wide range of topics related to

linear as well as nonlinear radiation-matter interactions, quantum-mechanical properties of radiation, noise and photon statistics, physical properties of solid-state materials relevant to radiation interaction, and devices and systems that utilize these basic principles for practical applications. *See* LASER; MASER.

Many electronic devices, such as semiconductor resonant tunneling diodes or the superconductor Josephson junctions, may be considered quantum electronic devices, since their operation is dictated by the principles of quantum mechanics. However, quantum electronics is understood to be concerned with only those devices whose operation involves radiation-matter interactions, such as lasers, parametric oscillators, and optoelectronic devices, such as light modulators, photodetectors, and optical switches. *See* ELECTROOPTICS; OPTICAL BISTABILITY; OPTICAL DETECTORS; OPTICAL MODULATORS.

Radiation absorption and emission. Some of the most fundamental ideas in quantum electronics are related to the absorption and emission of radiation by matter. The early quantum theory of radiation developed by Max Planck in 1901 postulated that radiation can be absorbed or emitted by matter in only discrete bundles or packets that were later called photons, with each photon having energy hf, where h is Planck's constant and f is the frequency of the radiation. Soon after in 1917, Albert Einstein discovered that photon emission by matter can be a spontaneous process or a process stimulated by other photons. Stimulated absorption of photons and spontaneous and stimulated emission of photons by atoms are illustrated in **Fig. 1**. The atoms are assumed to have two energy levels, and an electron in an atom can be in either the lower or the upper energy level. In each process, energy conservation requires that the energy difference between the upper and the lower energy level is equal to the photon energy ($E_2 - E_1 = hf$). Figure 1 shows that the number of photons can be increased substantially through stimulated emission, as each generated photon can in turn stimulate the emission of additional photons. This photon multiplication through stimulated emission is the basic principle behind the operation of lasers and masers. The acronym "laser" (or "maser") stands for light (or microwave) amplification through stimulated emission of radiation. *See* PHOTOEMISSION; PHOTON.

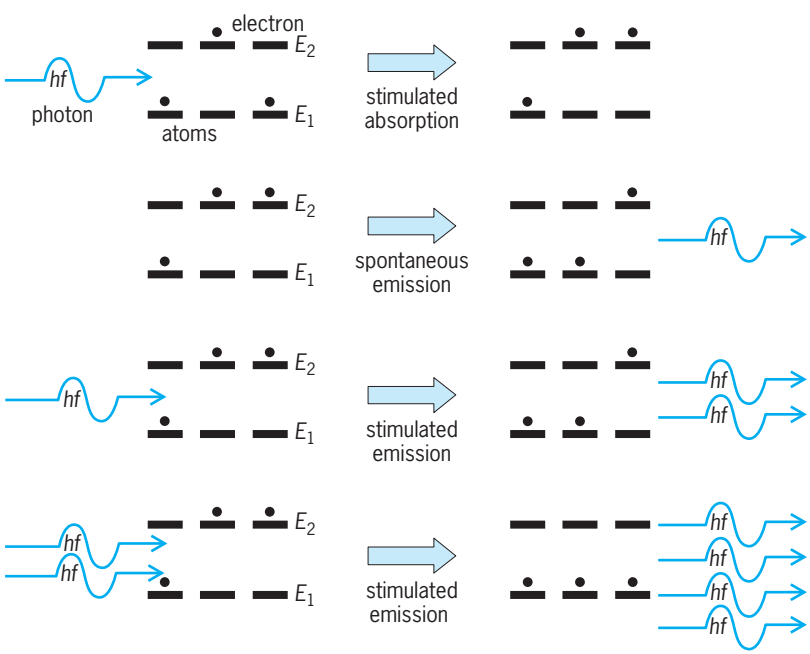

Fig. 1. Stimulated and spontaneous emission of photons by atoms.

Laser operation. In a laser, photons are confined inside a cavity along with atoms that can emit photons. These atoms can be in the form of molecules (as in dye lasers), solid-state crystals (as in semiconductor lasers), separate atoms (as in gas lasers), or impurities in a host material (such as chromium atoms in aluminum oxide in ruby lasers). Photons can be generated inside the cavity via spontaneous and stimulated emission by the atoms. Photons can be lost from the cavity via stimulated absorption by the atoms. In addition, photons can escape from the cavity. When the rate at which photons are generated inside the cavity is larger than the rate at which photons are lost from the cavity, the number of photons inside the cavity increases dramatically and this is called lasing. Lasing requires that the rate of stimulated emission of photons by the atoms exceed the rate of stimulated absorption, and this is achieved by ensuring that electrons in a majority of the atoms are in the upper energy level. This condition is called population inversion. Energy must be expended from outside in order to maintain population inversion during lasing.

Coherent radiation. A feature of lasers that makes radiation from a laser source different from radiation from a blackbody source is the coherence of the laser radiation. Coherent radiation does not have any significant phase jumps in time (or phase noise). In semiclassical terms, the coherence of laser radiation can be related to the fact that radiation emitted via a stimulated process has the same phase as that of the radiation that caused the stimulation. Therefore no phase noise is introduced during radiation amplification through stimulated emission. The small amount of phase noise in laser radiation is due to the spontaneous emission of radiation. Coherent electromagnetic radiation at low frequencies (less than 10^{11} Hz) can also be generated with oscillating electrical currents, as in antennas. Lasers and masers can provide high-power coherent radiation all the way from microwave frequencies (10^9 Hz) to x-ray frequencies (10^{17} Hz).

Nonlinear optical effects. Soon after the invention of the laser, it was discovered that a wide assortment of interesting and useful nonlinear optical phenomena could be produced with appropriate atoms or materials using light from lasers because of the unprecedented high power and coherence of laser radiation. Nonlinear optical effects can be understood in terms of the atomic or material polarization caused by the strong electric fields of laser radiation. It is well known from the classical theory of radiation that oscillating charge dipoles radiate electromagnetic energy just like oscillating currents. In fact,

stimulated emission by atoms in a laser can also be understood in terms of radiation emitted by oscillating charge dipoles induced in the atoms in the presence of the stimulating radiation. A volume density of charge dipoles induced in a material medium by an electric field $E(t)$ is described by the time-dependent material polarization $P(t)$ which can be expanded in a Taylor series in powers of the electric field strength $E(t)$ as

$$p(t) = \varepsilon_0[\chi^1 E(t) + \chi^2 E^2(t) + \chi^3 E^3(t) + \cdots]$$

where ε_0 is the permittivity of free-space, and χ^1, χ^2, and χ^3 are the first-, second-, and third-order optical susceptibilities of the material, respectively. The first-order susceptibility describes the linear optical response of the material, and the second- and third-order susceptibilities describe the nonlinear optical response of the material. Stimulated emission of radiation in lasers can be described in terms of the first-order linear optical susceptibility. Atoms and materials with large values of the second- and third-order susceptibilities have been used to produce a wide variety of nonlinear optical phenomena such as harmonic generation, sum and difference frequency generation, parametric amplification and oscillation, four-wave mixing, self-phase modulation, and Raman and Brillouin scattering (**Fig. 2**). Nonlinear optics has become a major part of the field of quantum electronics. *See* DIFFERENTIATION; DIPOLE; NONLINEAR OPTICS; POLARIZATION OF DIELECTRICS; RAMAN EFFECT; SCATTERING OF ELECTROMAGNETIC RADIATION; SERIES.

Ultrafast optics. Ultrafast optics has been another notable development in the field of quantum electronics since the 1980s. High-intensity single-cycle optical pulses have been produced by mode-locking lasers. Most lasers have many different discrete frequencies all lasing at the same time. These frequencies are almost equally spaced and correspond to the Fabry-Perot modes of the laser cavity. The phases of the different lasing modes inside a laser cavity are usually unrelated. Using a variety of linear and nonlinear optical techniques, the phases of the different lasing modes can be locked with respect to each other. When this happens, light output from the laser is not in the form of a continuous beam but in the form of short periodic pulses. Ultrashort pulses from mode-locked lasers have enabled the study of dynamical processes in physics, chemistry, and biology with femtosecond time resolution. In addition, the high temporal stability of the periodic pulse train from mode-locked lasers has enabled their use in optical frequency metrology applications with performances comparable to atomic clocks. *See* FREQUENCY COMB; INTERFEROMETRY; OPTICAL PULSES; ULTRAFAST MOLECULAR PROCESSES.

Optics on microchips. With the advance in micro- and nanofabrication techniques since the 1980s, quantum electronics has developed a promising new frontier. It has become possible to generate, amplify, and control coherent radiation over a wide range of frequencies on microchips. Micrometer-scale lasers

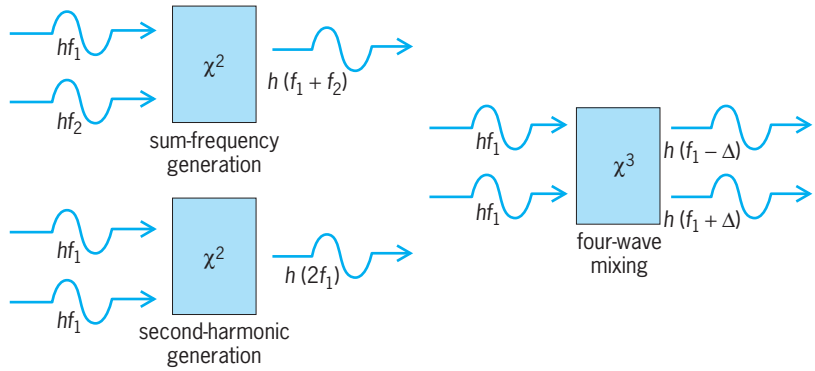

Fig. 2. Some nonlinear optical processes.

and optoelectronic devices are now possible. The development of inexpensive semiconductor lasers and integrated optoelectronic devices has been partly responsible for the wide deployment of fiber-optic communication links. The development of photonic very large scale integrated (VLSI) circuits is an active area of research. Such photonic circuits could provide ultra-high-speed information processing capabilities on microchips. *See* INTEGRATED CIRCUITS; INTEGRATED OPTICS; MICROLITHOGRAPHY; NANOTECHNOLOGY; OPTICAL COMMUNICATIONS. Farhan Rana

Bibliography. R. Boyd, *Nonlinear Optics*, 2d ed., Academic Press, New York, 2003; L. A. Coldren and S. W. Corzine, *Diode Lasers and Photonic Integrated Circuits*, Wiley, New York, 1995; *IEEE Journal of Quantum Electronics*, IEEE Press, New York, published monthly; R. Loudon, *The Quantum Theory of Light*, 3d ed., Oxford University Press, New York, 2000; A. E. Siegman, *Lasers*, University Science Books, Sausalito, CA, 1986; A. Yariv, *Quantum Electronics*, 3d ed., Wiley, NY, 1989.

Quantum field theory

The quantum-mechanical theory of physical systems whose dynamical variables are local functions of space and time. As distinguished from the quantum mechanics of atoms, quantum field theories describe systems with an infinite number of degrees of freedom. Such theories provide the natural language for describing the interactions and properties of elementary particles, and have proved to be successful in providing the basis for the fundamental theories of the interactions of matter. The present understanding of the basic forces of nature is based on quantum field theories of the strong, weak, electromagnetic, and gravitational interactions. Quantum field theory is also useful in the study of many-body systems, especially in situations where the characteristic length of a system is large compared to its microscopic scale. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Quantum electrodynamics. Quantum field theory originated in the attempt, in the late 1920s, to unify P. A. M. Dirac's relativistic electron theory and J. C. Maxwell's classical electrodynamics in a quantum

theory of interacting photon and electron fields. This effort was completed in the 1950s and was extremely successful. At present the quantitative predictions of the theory are largely limited to perturbative expansions (believed to be asymptotic) in powers of the fine-structure constant. However, because of the extremely small value of this parameter, $\alpha = e^2/\hbar c \approx 1/137$ (where e is the electron charge, $\hbar$ is Planck's constant divided by 2π, and c is the speed of light), such an expansion is quite adequate for most purposes. The remarkable agreement of the predictions of quantum electrodynamics with high-precision experiments (sometimes to an accuracy of 1 part in 10^{12}) provides strong evidence for the validity of the basic tenets of relativistic quantum field theory. *See* CLASSICAL FIELD THEORY; ELECTROMAGNETIC RADIATION; MAXWELL'S EQUATIONS; PERTURBATION (QUANTUM MECHANICS); PHOTON; QUANTUM ELECTRODYNAMICS; RELATIVISTIC ELECTRODYNAMICS; RELATIVISTIC QUANTUM THEORY.

Theories of other interactions. Quantum field theory also provides the natural framework for the treatment of the weak, strong, and gravitational interactions.

Weak interactions. The first of such applications was Fermi's theory of the weak interactions, responsible for radioactivity, in which a hamiltonian was constructed to describe beta decay as a product of four fermion fields, one for each lepton or nucleon. This theory has been superseded by the modern electroweak theory that unifies the weak and the electromagnetic interactions into a common framework. This theory is a generalization of Maxwell's electrodynamics which was the first example of a gauge theory, based on a continuous local symmetry. In the case of electromagnetism the local gauge symmetry is the space-time-dependent change of the phase of a charged field. The existence of massless spin-1 particles, photons, is one of the consequences of the gauge symmetry. The electroweak theory is based on generalizing this symmetry to space-time-dependent transformations of the labels of the fields, based on the group SU(2) × U(1). However, unlike electromagnetism, part of this extended symmetry is not shared by the ground state of the system. This phenomenon of spontaneous symmetry breaking produces masses for all the elementary fermions and for the gauge bosons that are the carriers of the weak interactions, the $W^{\pm}$ and Z bosons. (This is known as the Higgs mechanism.) The electroweak theory has been confirmed by many precision tests, and almost all of its essential ingredients have been verified. *See* ELECTROWEAK INTERACTION; GAUGE THEORY; INTERMEDIATE VECTOR BOSON; SYMMETRY BREAKING; SYMMETRY LAWS (PHYSICS); WEAK NUCLEAR INTERACTIONS.

Strong interactions. The application of quantum field theory to the strong or nuclear interactions dates from H. Yukawa's hypothesis that the short-range nuclear forces arise from the exchange of massive particles that are the quanta of local fields coupled to the nucleons, much as the electromagnetic interactions arise from the exchange of massless photons that are the quanta of the electromagnetic field. The modern theory of the strong interactions, quantum chromodynamics, completed in the early 1970s, is also based on a local gauge theory. This is a theory of spin-$^1/_2$ quarks invariant under an internal local SU(3) (color) gauge group. The observed hadrons (such as the proton and neutron) are SU(3) color-neutral bound states of the quarks whose interactions are dictated by the gauge fields (gluons). This theory exhibits almost-free-field behavior of quarks and gluons over distances and times short compared to the size of a hadron (asymptotic freedom), and a strong binding of quarks at large separations that results in the absence of colored states (confinement). *See* ELEMENTARY PARTICLE; GLUONS; MESON; QUANTUM CHROMODYNAMICS; QUARKS.

Gravitation. Quantum field theory has been tested down to distances of 10^{-20} m. There appears to be no reason why it should not continue to work down to Planck's length, $(G\hbar/c^3)^{1/2} \approx 10^{-35}$ m (where G is the gravitational constant), where the quantum effects of gravity become important. In the case of gravity, A. Einstein's theory of general relativity already provides a very successful classical field theory. However, the union of quantum mechanics and general relativity raises conceptual problems that seem to call for a radical reexamination of the foundations of quantum field theory. Among these problems are the lack of renormalizability of standard perturbative quantization schemes, and the quantum-mechanical implications of the ubiquitous singularities that occur in classical general relativity. *See* FUNDAMENTAL INTERACTIONS; GRAVITATION; QUANTUM GRAVITATION; RELATIVITY.

General properties. A field theory can be specified by either lagrangian or hamiltonian methods, the former having the advantage of making manifest the symmetries of the theory and the latter being the basis of canonical quantization, the generalization of canonical quantization for systems with a finite number of degrees of freedom. An interacting relativistic field theory necessarily involves creation and annihilation processes and is therefore a many-body theory.

The problem of constructing exact solutions of the coupled nonlinear equations for interacting four-dimensional field theories has so far proved too formidable for solution. However, there do exist both soluble field-theoretic models and existence proofs for others in three or two space-time dimensions. In four dimensions, however, certain assumptions must be made (the axioms of quantum field theory) in order to obtain physical predictions. These include assuming that a nondegenerate, Lorentz-invariant vacuum state exists; that the spectra of the energies and momenta lie in the forward light cone; and that a discrete spectrum of single-particle states exists. In addition, it is assumed that, for large positive (or negative) times, the matrix elements of the fields under consideration (appropriately smeared out) between normalizable physical states approach constant multiples of the matrix elements of free fields. By using this condition, which expresses the intuitive idea

that any system with only short-range interactions will consist of freely moving particles at asymptotic times, all matrix elements of physical interest can be expressed in terms of matrix elements of products of field operators between vacuum states. In particular, the S matrix, which yields the probability amplitude for the transition from an initial state of free particles to a final state of free particles, can be related to these matrix elements. *See* SCATTERING MATRIX.

Many properties of the S matrix can be established without recourse to the specific dynamics of the fields under consideration, as long as they are local fields satisfying the microcausality condition. This condition expresses the simultaneous observability of field operators at spacelike separations, and is a necessary condition in a Lorentz-invariant theory in which signals do not propagate with a velocity faster than the velocity of light. It can be shown that the axioms of quantum field theory lead to many general properties that are independent of specific dynamical models. One of the most general results is the CPT theorem. This theorem states that the laws of physics must be invariant under a simultaneous reversal of time, charge conjugation (which reverses the sign of all conserved charges and interchanges particles with antiparticles), and parity, even though none of these individual transformations are exact symmetries of nature. A consequence of this theorem is that the masses of particles and antiparticles must be identical; this has been tested in one case to an accuracy of 1 part in 10^{-7}. Another general result is the correlation between spin and statistics, which states that integer-spin and half-integer-spin particles must obey Bose-Einstein and Fermi-Dirac statistics respectively; that is, collections of such particles must be described by totally symmetric and totally antisymmetric wave functions respectively. *See* BOSE-EINSTEIN STATISTICS; CPT THEOREM; FERMI-DIRAC STATISTICS; QUANTUM STATISTICS.

Another class of general results involves the analytic properties of scattering amplitudes and Green's functions as functions of complex energy and momentum, which together with derived or assumed bounds on the asymptotic behavior of the scattering amplitude provide the basis for the derivation of dispersion relations. These dispersion relations are essentially consequences of microcausality, and their verification is strong evidence for the basic structure of quantum field theory. *See* DISPERSION RELATIONS; GREEN'S FUNCTION.

Other general features follow from symmetry properties. In quantum-mechanical systems with a finite number of degrees of freedom, a continuous symmetry implies the existence of a conserved charge that commutes with the hamiltonian. The consequences of such a symmetry are that the particles of the theory fall into degenerate multiplets with the same mass and their scattering amplitudes are linearly related. In quantum field theory, another realization of the symmetry is possible and often occurs, one in which the vacuum or ground state does not respect the symmetry. An example of symmetry breaking is provided by the phenomenon of spontaneous magnetization, wherein a magnetic material exhibits at low temperature a magnetic field pointing in a specific direction, even though the microscopic laws are rotationally invariant. In such a case the physical manifestations of the symmetry are quite different. Spontaneous symmetry breaking implies the existence of massless particles (called Goldstone bosons) with the quantum numbers of the conserved charge. The particles of the theory need not form degenerate multiplets, and the symmetry gives rise to low-energy theorems for the scattering of the Goldstone bosons. If, however, the symmetry is a local gauge symmetry, the Goldstone bosons are absent but the corresponding gauge bosons acquire a mass. This is the Higgs mechanism, and it underlies the unified electroweak theory, in which the exact $SU(2) \times U(1)$ symmetry is spontaneously broken down to the $U(1)$ gauge symmetry of electrodynamics.

Feynman diagrams. In a given lagrangian quantum field theory, Green's functions and the S matrix can be calculated perturbatively in the nonlinear coupling constants by constructing a perturbation series expansion for the unitary transformation that relates the fields under consideration to free fields. Individual terms in the expansion of Green's functions and scattering amplitudes can be represented by Feynman diagrams. These diagrams provide an instructive physical picture of the process under consideration and, with the aid of Feynman rules that associate a specific integral with a given diagram, are of enormous aid in performing calculations. *See* FEYNMAN DIAGRAM.

Path-integral formalism. An alternate formulation of quantum field theory is based on Feynman's representation of transition amplitudes as sums over path histories weighted by the classical action. Although it is difficult to give a rigorous definition of the functional integrals that appear in this formulation, it has proved extremely useful for deriving consistent perturbative expansions for gauge field theories and as a starting point for semiclassical approximations. The path integral is also the starting point for numerical approaches to quantum field theory. Continuous space and time is replaced by a finite-volume lattice, in which case the path integral reduces to a finite product of ordinary integrals. Various numerical techniques (such as Monte Carlo integration) are then used to evaluate physical observables, by extrapolating the lattice calculations to the limit of an infinite volume and vanishing lattice spacing. Such methods have been used in quantum chromodynamics to calculate the hadronic mass spectrum, with some success. *See* FEYNMAN INTEGRAL; MONTE CARLO METHOD.

Renormalization. Quantum field theory is plagued by many divergences such as those arising from the existence of an infinite number of degrees of freedom (which can be dealt with by careful control of the infinite-volume limit), infrared divergences arising from the existence of massless particles (which can be dealt with by careful treatment of asymptotic states), and ultraviolet divergences arising from the pointlike structure of local field theory. The last are

dealt with by the process of renormalization, which eliminates the original (or bare) parameters in the lagrangian (coupling constants and masses) in favor of those with physical significance. The meaning of renormalization is that the parameters in lagrangian refer to the properties of the theory at arbitrarily small distances, whereas the physical parameters measured at normal distances are modified by the interactions. Thus, in quantum electrodynamics charge renormalization is due to the screening of the electric bare charge by virtual electron-positron pairs (vacuum polarization), resulting in a diminished physical electric charge. In a renormalizable field theory, all ultraviolet divergences can be systematically eliminated by perturbatively adjusting the bare parameters to yield finite physical parameters. Nonrenormalized field theories, however require the specification of an infinite number of parameters in perturbation theory or complete nonperturbative control over their short-distance behavior. *See* RENORMALIZATION.

Renormalization group. The renormalization group describes how the dynamics of a system evolves as the scale of the phenomena being observed is changed. A scale-dependent hamiltonian is constructed to describe the system at varying distances. A change of the scale of length induces a transformation of this effective hamiltonian. These transformations often have the property of approaching a fixed point after many iterations. In this case, the large- or small-distance structure of the theory will be relatively independent of its structure at intermediate distances and, if the fixed-point hamiltonian is simple enough, can be explicitly calculated. These techniques have been enormously successful in applications of quantum field theory to the study of critical behavior in statistical mechanics. In the case of the strong interactions the asymptotic freedom of quantum chromodynamics means that the effective dynamics of quarks and gluons at very high energies is that of an almost-free theory, and thus can be calculated. Many of these high-energy predictions have been verified by experiment. *See* CRITICAL PHENOMENA.

The renormalization group also provides a perspective on models of low-energy physics as effective field theories. It is assumed that the final theory of physics (a unified theory of all interactions) can be described at some very high energy scale, Λ, by a local quantum field theory. In the absence of further dynamical information, this theory will be described by the most general possible local quantum field theory, constrained only by the assumed symmetries of the theory. It will also be assumed that all coupling constants that parametrize the strengths of the interactions at the scale Λ are numbers of order one, times appropriate powers of Λ, $\hbar$, and c. Starting with this infinite-parameter theory at high energies, it is possible to integrate out the high-momentum modes of all the fields, thus deriving an effective theory that describes low-energy physics at energies E much less than Λ. The effective theory will depend on the scale of energy E. The renormalization group describes the change in the effective theory as E is changed. The characteristic behavior of the solutions of the renormalization group equations, which describe the flow in the space of theories, is that the solutions approach a finite-dimensional submanifold in this infinite-dimensional space, which is parametrized by the renormalizable couplings. A nonrenormalizable interaction is one whose coupling constant has dimensions of powers of the inverse energy, and thus has a coupling of order g/Λ^d, where d is the dimension of the coupling, Λ is the cutoff, and g is a pure number of order one. A renormalizable coupling, however, is characterized by a dimensionless parameter, g', of order one. To compare the two at low energies of order E, dimensionless ratio must be formed, given by the equation below.

$$\frac{g(E/\Lambda)^d}{g'} \sim (E/\Lambda)^d$$

The effect of the nonrenormalizable interactions on low-energy processes will therefore be of order $(E/\Lambda)^d$. Thus the nonrenormalizable interactions (of which there might be an infinite number) can be totally neglected at low energies, and the low-energy theory, up to small corrections, will then be described by the most general renormalizable quantum field theory consistent with the symmetries, with a finite number of parameters. David Gross

Bibliography. J. Bjorken and S. Drell, *Relativistic Quantum Fields*, 1965; C. Itzkykson and J. Zuber, *Quantum Field Theory*, 1980; M. Kaku, *Quantum Field Theory*, 1993; F. Mandl and G. Shaw, *Quantum Field Theory*, rev. ed., 1993; P. Ramond, *Field Theory: A Modern Primer*, 2d rev. ed., 1990; J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, 4th ed., 2002.

Quantum gravitation

The quantum theory of the gravitational field; also, the study of quantum fields in a curved space-time. In classical general relativity, the gravitational field is represented by the metric tensor $g_{\mu\nu}$ of space-time. This tensor satisfies Einstein's field equation, with the energy-momentum tensor of matter and radiation as a source. However, the equations of motion for the matter and radiation fields also depend on the metric.

Classical field theories such as Maxwell's electromagnetism or the classical description of particle dynamics are approximations valid only at the level of large-scale macroscopic observations. At a fundamental level, elementary interactions of particles and fields must be described by relativistic quantum mechanics, in terms of quantum fields. Because the geometry of space-time in general relativity is inextricably connected to the dynamics of matter and radiation, a consistent theory of the metric in interaction with quantum fields is possible only if the metric itself is quantized. *See* MAXWELL'S EQUATIONS; QUANTUM FIELD THEORY; RELATIVISTIC QUANTUM THEORY; RELATIVITY.

Under ordinary laboratory conditions the curvature of space-time is so extremely small that in most quantum experiments gravitational effects are completely negligible. Quantization in Minkowski space is then justified. Gravity is expected to play a significant role in quantum physics only at rather extreme conditions of strongly time-dependent fields, near or inside very dense matter. The scale of energies at which quantization of the metric itself becomes essential is given by $(\hbar c^5/G)^{1/2} \approx 10^{19}$ GeV, where G is the gravitational constant, $\hbar$ is Planck's constant divided by 2π, and c is the velocity of light. Energies that can be reached in the laboratory or found in cosmic radiation are far below this order of magnitude. Only in the very early stages of the universe, within a proper time of the order of $(G\hbar/c^5)^{1/2} \approx 10^{-43}$ s after the big bang, would such energies have been produced. *See* BIG BANG THEORY.

Quantum fields in curved space-time. In most physical systems the metric is quasistationary over macroscopic distances so that its fluctuations can be ignored. A quantum description of fields in a curved space-time can then be given by treating the metric as a classical external field in interaction with the quantum fields. If the metric is asymptotically flat, asymptotic particle states can be defined associated with positive frequency-solutions of the free-field wave equation. "In" and "out" Hilbert spaces can then be constructed which will be related by a unitary transformation S. In particular, the vacuum states, $|0_{\text{in}}\rangle$ and $|0_{\text{out}}\rangle$, are related by $S|0_{\text{out}}\rangle = |0_{\text{in}}\rangle$, and in general, in a nonstationary background, $|0_{\text{out}}\rangle \neq |0_{\text{in}}\rangle$. This means that particles are created by the nonstationary metric. In the special case of free quantum fields, an explicit form for the S-matrix can be obtained. It is found that particles are created in pairs of which one member is a particle and the other its antiparticle. One exception to particle creation is the case of a massless free field in interaction with a background Robertson-Walker metric which represents a spatially homogeneous and isotropic geometry. This property of space is assumed in the Friedmann models of an expanding universe, discussed below. The conformal invariance of these systems ensures that $|0_{\text{out}}\rangle = |0_{\text{in}}\rangle$. *See* SCATTERING MATRIX.

Quantum effects of black holes. The most striking quantum effect in curved space-time is the emission of radiation by black holes. A black hole is an object that has undergone gravitational collapse. Classically this means that it becomes confined to a space-time region in which the metric has a singularity (the curvature becomes infinite). This region is bounded by a surface, called the horizon, such that any matter or radiation falling inside becomes trapped. Therefore, classically the mass of a black hole can only increase. However, this is no longer the case if quantum effects are taken into account. When, because of fluctuations of the quantum field, particle-antiparticle or photon pairs are created near the horizon of a black hole, one of the particles carrying negative energy may move toward the hole, being absorbed by it, while the other moves out with positive energy. Only these emitted particles can be observed; their number, in a given mode j, is the expectation value in the incoming vacuum $|0_{\text{in}}\rangle$ of the number operator $n_{j\text{out}}$ associated with the outgoing Hilbert space. The index j stands for the frequency $\omega/2\pi$ and other quantum numbers. A careful analysis of the propagation of wave packets shows that the average number per unit frequency of particles emitted per unit time is given by Eq. (1), where $\Gamma(\omega)$ is the absorption

$$\langle n_\omega \rangle = \frac{\Gamma(\omega)}{\exp\,(2\pi\omega/\kappa - 1)} \tag{1}$$

coefficient and κ is the surface gravity of the black hole. For a Schwarzschild black hole of mass M, $\kappa = 1/(4M)$, in geometric units chosen so that $G = \hbar = c = 1$. This result shows that a black hole emits radiation as a body in thermal equilibrium at temperature $T = \kappa/(2\pi k)$, where k is the Boltzmann constant. It lends support to the thermodynamic interpretation of the classical laws of black hole mechanics. *See* BLACK HOLE; GRAVITATIONAL COLLAPSE; HEAT RADIATION.

Back reaction on the metric. As energy is taken out of the black hole by the emission of radiation, its mass must decrease. This is an effect of the back reaction of the quantum field on the metric. As long as the mass of the black hole is very large compared to Planck's mass $(\hbar c/G)^{1/2} \approx 10^{-8}$ kg, a semiclassical quasistationary approach is valid and the rate of change of mass is given by the total energy emitted per unit time. If it is assumed that $\Gamma(\omega)$ is equal to the absorption coefficient of a blackbody of radius equal to the Schwarzschild radius, and is thus proportional to the area of the horizon, then from Eq. (1) it is found that the total rate of emission is inversely proportional to the square of the mass. As the mass decreases to very small values, the black hole evaporates with a burst of radiation. In this process, conservation of baryon number is violated. These conclusions rely on approximations that break down in the last stages of the black hole's evolution; hence they may not be entirely correct. For stellar black holes whose masses are of the order of a solar mass, the emission rate is negligibly small and unobservable. Only primordial black holes, of mass less than 10^{13} kg, formed very early in the quantum era of the universe, would have been small enough to produce quantum effects that could play any significant role in astrophysics or in cosmology.

Quantization of the metric. There are basically two approaches to the quantization of the metric, the canonical and the covariant quantization. Each method for quantizing gravity will be briefly discussed for the case in which other fields are absent. Einstein's field equations then reduce to Eq. (2), where the indices μ and ν take on the values 0, 1, 2, 3.

$$G_{\mu\nu} \equiv R_{\mu\nu} - \tfrac{1}{2}Rg_{\mu\nu} = 0 \tag{2}$$

Canonical quantization. In the canonical quantization, a parametrization of the metric is used. There are

four nonpropagating fields, analogous to the timelike component A_0 of the vector potential A_μ in electrodynamics. The dynamical variables are the components g_{ij} $(i, j = 1, 2, 3)$ of the three-geometry of a spacelike hypersurface and their conjugate momenta π^{ij}. Variation of the Einstein action with respect to the nonpropagating fields gives a set of constraint equations. At this stage it is possible either to use Dirac's method of quantization for constrained systems, or choose a coordinate system by imposing four gauge conditions $F_\mu(g,\pi) = 0$. Taking the time derivative of these gauge conditions and using the equations of motion yields further constraints $\dot{F}_\mu = 0$. This whole set of constraints is then used to eliminate the nonpropagation fields and four of the pairs of variables g_{ij}, π^{ij}, leaving only two pairs of dynamical variables, representing the two degrees of freedom of the system. These are then quantized by means of the usual canonical commutation rules. Solving the constraint is a major technical difficulty in this procedure. *See* HAMILTON'S EQUATIONS OF MOTION; NONRELATIVISTIC QUANTUM THEORY; POTENTIALS; VARIATIONAL METHODS (PHYSICS).

An alternative is to follow J. A. Wheeler's superspace formulation of general relativity using the Hamilton-Jacobi formalism. The Einstein action is taken between two spacelike hypersurfaces where the metric of the three-geometry is given. The minimal of the action with the metric held fixed at the two hypersurfaces is the Hamilton principal function, which is a function of the two given three-geometries. If the second hypersurface is allowed to change, the Hamilton function gives the history of the geometries connecting the two hypersurfaces. The variation of the Hamilton action corresponding to a variation of the metric on the second hypersurface gives an expression for the conjugate momenta, π^{ij}. Replacement of the conjugate momenta with this expression in the constraint equations yields the full geometrodynamics of supersace at the classical level. *See* HAMILTON-JACOBI THEORY.

The theory can be quantized by replacing the dynamical variables g_{ij} and π^{ij} by operators satisfying canonical comutation relations, and imposing the constraint equations on the state vectors of the system. A remarkable feature in this procedure is that the "time'' parameter no longer appears in the evolution equation of the system. Somehow it has to be defined as an observable in terms of the extra components of the metric g_{ij}. The question of defining a complete set of observable in the theory is still open. Attempts to address this problem have also been made using a new set of variables instead of g_{ij}. In these variables the constraints take on a simpler form.

Covariant quantization. In the covariant method the metric is split into a classical background field and a quantum operator. Since not all components of the quantum operator are independent, constraints must be imposed on the physical states, analogous to the constraints in the Gupta-Bleuler formalism in quantum electrodynamics. If the classical metric is chosen to be the Minkowski metric, it is possible to formally write down Feynman rules in momentum space. One objection to this method is that spacelike surfaces are defined with respect to the classical metric, not the total metric. *See* QUANTUM ELECTRODYNAMICS.

Application of path integral representation. A third method, which can be derived from the first and is now most widely used, is based on the Feynman path integral representation for the vacuum-to-vacuum amplitude, which is the generator of Green's functions for the quantum theory. One important feature of this method is that since the topology of the manifold is not specified at the outset it is possible to include a sum over paths in different topologies. The outcome of this idea is that the vacuum would, at the level of the Planck length, $(G\hbar/c^3)^{1/2} \approx 10^{-35}$ m, acquire a foamlike structure. *See* FEYNMAN INTEGRAL; GREEN'S FUNCTION.

Problems. At present a complete, consistent theory of quantum gravity is still lacking. The formal theory fails to satisfy the power-counting criterion for renormalizability. In every order of the perturbation expansion, new divergences appear which could only be canceled by counterterms that do not exist in the original lagrangian. This may not be just a technical problem but the reflection of a conceptual difficulty stemming from the dual role, geometric and dynamic, played by the metric. *See* RENORMALIZATION.

Quantum cosmology. The existence of a singularity in generic solutions of Einstein's equation of general relativity was proven in 1970 under very general and realistic assumptions. Thus the big bang theory of the universe, based on classical general relativity, was given a solid theoretical foundation. However, near the singularity at the origin of the universe the classical description breaks down and the concepts of quantum mechanics are required. The universe should be described in terms of a wave function whose determination depends on what is assumed as initial boundary conditions. If the wave function is constructed by using R. P. Feynman's formalism of sum over histories, it is convenient to rotate the time in the complex plane to the imginary axis, turning "real time" into "imaginary time." The Minkowski metric of real space-time becomes a euclidean metric in the new manifold. Compact geometries of this four-dimensional euclidean manifold can then be considered.

Among the boundary conditions that can be imposed in such geometries is the condition of no boundary. This hypothesis has been advocated as appropriate for the construction of the wave function of the universe. Here not only would the singularity of the classical theory disappear in the quantum theory, but even the need to specify boundary conditions would be circumvented by the no-boundary hypothesis; in the euclidean construction of the wave function, all possible compact four-geometries should be included in the sum over histories without imposing any initial conditions at all. *See* BIG BANG THEORY; COSMOLOGY.

Problem of the cosmological constant. The cosmological constant is a parameter that can be introduced

in general relativity with the effect of giving a nonvanishing curvature to space-time even in the absence of matter. There are very stringent limits derived from astronomical observations on the value of the cosmological constant, which is either zero or extremely small. In a quantum theory, a cosmological constant will result from quantum fluctuations, even if it is absent in the classical action. It appears that a precise fine tuning of the boundary conditions at the Planck scale is required to obtain a vanishing cosmological constant. This is a major problem in quantum gravity whose resolution requires some mechanism to drive the cosmological constant to zero almost independently of the physics below the Planck scale, that is, for most boundary conditions on the scale.

Supergravity and superstrings. Supergravity is a geometric extension of general relativity which incorporates the principle of supersymmetry. Supersymmetry is a kind of symmetry, discovered in the 1970s, that allows for the transformation of fermions and bosons into each other. (Fermions carry half-integer spin while bosons carry integer spin; they also obey different statistics.) Supergravity can be formulated in space-time manifolds with a total of $D = d + 1$ dimensions, where d, the number of space dimensions, can be as large as 10. They constitute truly unified theories of all interactions including gravity.

One motivation for studying such theories was the hope that, because of the extra symmetry, they were better behaved when quantized, that is, that supersymmetry would make the divergencies of the quantum theory of gravity cancel out. This was expected, above all, of the $D = 11$ theory after reduction to four dimensions by some dynamical mechanism of compactification of the extra seven dimensions. All the efforts to achieve this goal proved unsuccessful. Moreover, it became apparent that the theories obtained in four dimensions by this process of dimensional reduction did not have some of the basic features of the physics of the real world at low energies. A possible exception might have been some theory derived from $D = 10$ supergravity with an additional gauge symmetry. *See* GAUGE THEORY.

In the early 1980s, some encouraging results were found with a theory based on the idea that the basic objects of nature are not pointlike but actually one-dimensional objects like strings, which can be open or closed. Consistency of the quantum theory of bosonic strings requires that they exist in a $D = 26$ space-time manifold, but the action principle for their motion is formulated in two dimensions. Conformal invariance of the action is a fundamental symmetry of the theory. It can be shown that the bosonic string theory is not realistic because it contains an unphysical tachyon. Incorporating supersymmetry into the theory leads to a critical dimension $D = 10$ and the removal of the tachyon. *See* TACHYON.

The relevance of superstring theories to quantum gravity is that they contain a zero-mass graviton (spin-2 state), and the equation of motion for a background metric in which the string is quantized reproduces (in some approximation) Einstein's equation.

In the approximation of neglecting string excitations, certain superstring models may be described in terms of local fields as a $D = 10$ supergravity theory. However, at least in every order of a certain perturbation expansion (a loop expansion) the quantized superstring theory is finite. At present, these are the only theories that both include gravity and can be consistently quantized. Although a superstring theory may eventually become the ultimate theory of all the interactions, there is still a very long way to go in making the connection between its fundamental fields and the fields representing the particles and their interactions as observed at low energies. *See* FUNDAMENTAL INTERACTIONS; GRAVITATION; SUPERGRAVITY; SUPERSTRING THEORY; SUPERSYMMETRY.

Samuel W. MacDowell

Bibliography. A. Ashketar and J. Stachel (eds.), *Conceptual Problems of Quantum Gravity*, 1991; S. W. Hawking, *A Brief History of Time*, 1988; S. W. Hawking and W. Israel (eds.), *General Relativity: An Einstein Centenary Survey*, 1979; C. J. Isham, R. Penrose, and D. W. Sciama (eds.), *Quantum Gravity: An Oxford Symposium*, 1975, and *Quantum Gravity 2: A Second Oxford Symposium*, 1981.

Quantum mechanics

The modern theory of matter, of electromagnetic radiation, and of the interaction between matter and radiation; also, the mechanics of phenomena to which this theory may be applied. Quantum mechanics, also termed wave mechanics, generalizes and supersedes the older classical mechanics and Maxwell's electromagnetic theory. Atomic and subatomic phenomena provide the most striking evidence for the correctness of quantum mechanics and best illustrate the differences between quantum mechanics and the older classical physical theories. Quantum mechanics is needed to explain many properties of bulk matter, for instance, the temperature dependence of the specific heats of solids.

The formalism of quantum mechanics is not the same in all domains of applicability. In approximate order of increasing conceptual difficulty, mathematical complexity, and likelihood of future fundamental revision, these domains are the following: (i) Nonrelativistic quantum mechanics, applicable to systems in which particles are neither created nor destroyed, and in which the particles are moving slowly compared to the velocity of light, as shown by relation (1). Here a particle is defined as a material

$$c \cong 3 \times 10^{10} \text{ cm/s} \tag{1}$$

entity having mass, whose internal structure either does not change or is irrelevant to the description of the system. (ii) Relativistic quantum mechanics, applicable in practice to a single relativistic particle (one whose speed equals or nearly equals c); here the particle may have zero rest mass, in which event, its speed must equal c. (iii) Quantum field theory, applicable to systems in which particle creation and

destruction can occur; the particles may have zero or nonzero rest mass.

This article is concerned mainly with nonrelativistic quantum mechanics, which apparently applies to all atomic and molecular phenomena, with the exception of the finer details of atomic spectra. Nonrelativistic quantum mechanics also is well established in the realm of low-energy nuclear physics, meaning nuclear phenomena wherein the particles have kinetic energies less than about 10^8 electronvolts (1 eV = 1.6×10^{-12} erg, and is the energy gained by an electron in traversing a potential difference of 1 volt). Many quantum-mechanical predictions are not as quantitatively accurate for nuclei as for atomic and molecular systems, however, because nuclear forces are not accurately known. *See* ATOMIC STRUCTURE AND SPECTRA.

The novel (from the standpoint of classical physics) features of nonrelativistic quantum mechanics are described in this article. Some of these features are retained, others modified, in the more complicated domains of relativistic quantum mechanics and quantum field theory. *See* QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; RELATIVISTIC QUANTUM THEORY; SYMMETRY LAWS (PHYSICS).

For the formal mathematical structure of nonrelativistic quantum mechanics and justification for many of the assertions made in this article. *See* NONRELATIVISTIC QUANTUM THEORY.

Planck's constant. The quantity 6.626×10^{-34} joule-second, first introduced into physical theory by Max Planck in 1901, is a basic ingredient of the formalism of quantum mechanics. Most of the fundamental quantum-mechanical relations, for example, Schrödinger's equation and Heisenberg's uncertainty principle, explicitly involve Planck's constant, as do many of the well-verified consequences of quantum mechanics, for example, the formula for the energy levels of atomic hydrogen. Planck's constant plays no role in the classical theories. Planck's constant commonly is denoted by the letter h; the notation $\hbar = h/2\pi$ also is standard.

Uncertainty principle. In classical physics the observables characterizing a given system are assumed to be simultaneously measurable (in principle) with arbitrarily small error. For instance, it is thought possible to observe the initial position and velocity of a particle and therewith, using Newton's laws, to predict exactly its future path in any assigned force field. According to the uncertainty principle (W. Heisenberg, 1927), accurate measurement of an observable quantity necessarily produces uncertainties in one's knowledge of the values of other observables. In particular, for a single particle relation (2*a*) holds,

$$\Delta x \Delta p_x > \hbar \tag{2a}$$

$$\Delta \phi \Delta l_z \gtrsim \hbar \tag{2b}$$

$$\Delta x \Delta E \gtrsim \frac{\hbar}{m} p_x \tag{2c}$$

$$\Delta t \Delta E \gtrsim \hbar \tag{2d}$$

where Δx represents the uncertainty (error) in the location of the x coordinate of the particle at any instant, and Δp_x is the simultaneous uncertainty in the x component of the particle momentum. Relation (2*a*) asserts that under the best circumstances, the product $\Delta x \Delta p_x$ of the uncertainties cannot be less than about 10^{-27} erg-s; of course, with poor measurements, the product can be much greater than $\hbar$. On the other hand, there is no restriction on the simultaneous determination of position along x and momentum along y; that is, the product $\Delta x \Delta p_y$ may equal zero. Other typical uncertainty inequalities for a single particle are given by relations (2*b*) and (2*c*).

In relation (2*b*) the particle location is specified in spherical coordinates, with polar axis along z; Δ^{ϕ} is the uncertainty in azimuth angle; and Δl_z is the uncertainty in the z component of the orbital angular momentum. In relation (2*c*) ΔE is the uncertainty in energy and m is the particle mass. The uncertainty relation (2*d*) is derived and interpreted somewhat differently than relations (2*a*)–(2*c*); it asserts that for any system, an energy measurement with error ΔE must be performed in a time not less than $\Delta t \sim \hbar/\Delta E$. If a system endures for only Δt s, any measurement of its energy must be uncertain by at least $\Delta E \sim \hbar/\Delta t$ ergs.

Because the numerical value of $\hbar$ is so small, and since $\Delta p_x = m \Delta \upsilon_x$, υ denoting velocity, the restrictions implied by relations (2*a*)–(2*c*) are utterly inconsequential for macroscopic systems, wherein masses are of the order of grams. For an electron, however, whose mass is 9.1×10^{-28} g, $\hbar/m \sim 1$ cm^2/s, and the uncertainty principle cannot be ignored. Similarly, relation (2*d*) is unimportant for macroscopic systems, wherein energies are of the order of ergs, but is significant for atomic systems where $\Delta E = \hbar/\Delta t$ need not be negligible compared to the actual energy E. *See* UNCERTAINTY PRINCIPLE.

Wave-particle duality. It is natural to identify such fundamental constituents of matter as protons and electrons with the mass points or particles of classical mechanics. According to quantum mechanics, however, these particles, in fact all material systems, necessarily have wavelike properties. Conversely, the propagation of light, which, by Maxwell's electromagnetic theory, is understood to be a wave phenomenon, is associated in quantum mechanics with massless energetic and momentum-transporting particles called photons. The quantum-mechanical synthesis of wave and particle concepts is embodied in the de Broglie relations, given by Eqs. (3*a*) and (3*b*). These give the wavelength λ and

$$\lambda = \frac{h}{p} \tag{3b}$$

$$f = \frac{E}{h} \tag{3b}$$

wave frequency f associated with a free particle (a particle moving freely under no forces) whose momentum is p and energy is E; the same relations give the photon momentum p and energy E associated with an electromagnetic wave in free space (that is,

in a vacuum) whose wavelength is λ and frequency is f. *See* PHOTON.

The wave properties of matter have been demonstrated conclusively for beams of electrons, neutrons, atoms (hydrogen, H, and helium, He), and molecules (H_2). When incident upon crystals, these beams are reflected into certain directions, forming diffraction patterns. Diffraction patterns are difficult to explain on a particle picture; they are readily understood on a wave picture, in which wavelets scattered from regularly spaced atoms in the crystal lattice interfere constructively along certain directions only. Moreover, the wavelengths of these "matter waves," as inferred from the diffraction patterns, agree with the values computed from Eq. (3*a*), as first demonstrated by C. J. Davisson and L. H. Germer in 1927. *See* ELECTRON DIFFRACTION; NEUTRON DIFFRACTION.

Photoelectric effect. The particle properties of light waves are observed in the photoelectric effect. When light of frequency f causes electrons to be emitted from a surface, all the electrons have very nearly the same maximum kinetic energy; the maximum kinetic energy is independent of the light intensity; the number of electrons emitted in unit time is proportional to the light intensity; as f is varied, the maximum electron kinetic energy W varies linearly with f, in fact, $W = hf - C$, $W = hf - C$, C being a constant characteristic of the emitting material. These observations are difficult to understand on the wave picture, wherein the magnitude of the electric field vector (which presumably exerts the force which ejects the electron) is proportional to the square root of the incident light intensity, and is not directly related to nor limited by, the incident light frequency f. The photoelectric effect is interpreted readily on the assumptions that the energy in the light beam is carried in quanta of energy $E = hf$; that emission of an electron results from absorption of a single quantum (a single photon); and that absorption of half or any fraction of a quantum is not possible, because the photons act as discrete indivisible entities. For additional information on the photoelectric effect *see* PHOTOEMISSION.

Compton effect. The particle properties of electromagnetic waves also are demonstrated in the Compton effect, wherein the wavelengths of x-rays are lengthened by scattering from free electrons. The change in wavelength is predicted quantitatively, assuming the scattering results from elastic collisions between photons and electrons, and using Eqs. (3*a*) and (3*b*) for the photon momentum and energy. The diffraction of x-rays by crystals was regarded, in the prequantum era, as conclusive proof that x-rays are waves and not "corpuscles." *See* COMPTON EFFECT; ELECTROMAGNETIC RADIATION; MAXWELL'S EQUATIONS; X-RAY DIFFRACTION.

Interference and diffraction. Wave propagation is distinguished from particle propagation by the phenomena of interference and diffraction. It is a general result of wave theories that interference and diffraction effects largely are confined to an angle (relative to the incident beam) which in radians equals about λ/d, where d is a characteristic dimension of the system causing the diffraction or interference, for example, the width of the slit diffracting the wave, or the distance between two interfering scattering centers. This fact and the magnitudes of λ inferred from Eq. (3*a*) are sufficient to explain why wave effects are not observed in the propagation of ordinary macroscopic bodies, but can be observed in the propagation of electrons, neutrons, and light atoms or molecules. For example, for a mass of 1 g moving at 1 cm/s, $\lambda = 6.6 \times 10^{-27}$ cm. But for a neutron or hydrogen atom (using $p = M\upsilon = 2ME$, where $M = 1.66 \times 10^{-24}$ g) moving at velocity corresponding to room temperature (300 K), $E = 3/2kT$ (k = Boltzmann's constant $= 1.38 \times 10^{-16}$ erg deg^{-1}), and λ turns out to equal 1.45×10^{-8} cm. For an electron with an energy of 100 eV, $\lambda = 1.22 \times 10^{-8}$ cm. For a proton with an energy 10^6 eV = 1 MeV, $\lambda = 2.88 \times 10^{-12}$ cm. These numerical results and the discussion of this paragraph also explain the ability of crystals, wherein interatomic spacings are about 10^{-8} cm, to give a good demonstration of electron and molecular diffraction; suggest the need for quantum mechanics to "understand" atomic systems, wherein atomic dimensions are $\sim 10^{-8}$ cm and electron energies are $\sim$10 eV; suggest the need for quantum mechanics to "understand" atomic nuclei, wherein nuclear dimensions are $\sim 10^{-13}$ cm and neutron or proton energies are $\sim$10 MeV; and explain why quantum effects are more readily observed in H_2 and He than in heavier gases, and at low temperatures rather than high. *See* DIFFRACTION; INTERFERENCE OF WAVES.

Relationship to uncertainty principle. Wave-particle duality is intimately connected with the uncertainty principle in that the uncertainty inequalities can be derived from analyses of specific experiments. For a nonrelativistic particle the connection can be seen from the following argument, which contains the basic elements of the rigorous formal treatment. As explained later, the probability that the x coordinate of the particle will lie in the interval x to $x + dx$ is $|\psi(x)|^2\, dx$, where $\psi(x)$ is called the wave function. Suppose measurement has ascertained that the particle lies in an interval of width Δx centered at $x = 0$; that is, measurement has determined that the dependence of $\psi(x)$ on x is approximately as shown in **Fig. 1**. Because of wave-particle duality, the wave

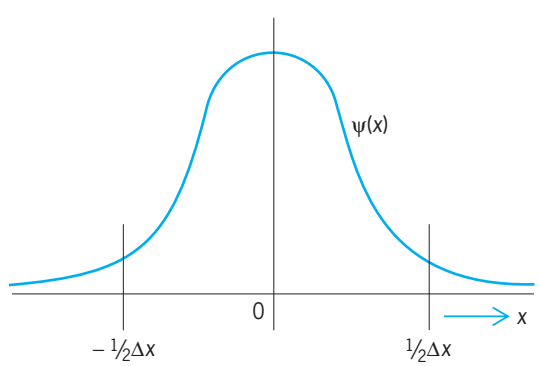

Fig. 1. Plot of $\psi(x)$ versus x.

packet of Fig. 1 can be looked upon as a superposition of waves. Since $\psi(x)$ rises from and falls to a very small value in an interval Δx, the packet must contain waves whose half wavelengths are as small as Δx; furthermore, it can be proved that because $\psi(x)$ does not change sign, the packet must contain waves of very long wavelength. Thus, in the packet, the wavelengths λ run from about $2\Delta x$ to ∞; the reciprocal wavelengths λ^{-1} run from about 0 to $1/(2\Delta x)$. But from Eq. (3*a*), $\Delta p_x = h\Delta\lambda^{-1}$. Hence $\Delta p_x \Delta x$ cannot be less than about $h/2$. Considering the simplicity of the argument, this result is close enough to relation (2*a*). Similarly, relation (2*d*) can be understood from Eq. (3*b*) and from the fact that to transmit information in a time Δt, that is, to turn a measuring instrument on and off in an interval Δt, it is necessary to use frequencies higher than about $(2\Delta t)^{-1}$. Using relation (2*a*), the discussion of this paragraph suggests that in atoms, whose dimensions Δx are about 10^{-8} cm, one must expect to find electrons with energies as given by Eq. (4), which is of the order

$$E = \frac{p^2}{2m} = \frac{\hbar}{2m(\Delta x)^2} = 3.8 \text{ eV} \qquad (4)$$

of magnitude observed. Similarly, in atomic nuclei, whose dimensions Δx are about 10^{-13} cm, one must expect to find neutrons or protons with energies about 20 MeV, again of the order of magnitude observed. On the other hand, if atomic nuclei contained electrons, the electron energies would be as high as 200 MeV. In computing this value, it is necessary to use the relativistic relation $E = m^2c^4 + c^2p^2$ connecting E and p. These energies are much too high to be explained by electrostatic forces between electrons and protons at separations of about 10^{-13} cm. Thus, the uncertainty principle leads to the inference that electrons are not contained in atomic nuclei, which inference is in accord with current theories of nuclear structure and beta decay. *See* NUCLEAR STRUCTURE; RADIOACTIVITY; RELATIVISTIC MECHANICS.

Complementarity. Wave-particle duality and the uncertainty principle are thought to be examples of the more profound principle of complementarity, first enunciated by Niels Bohr (1928). According to the principle of complementarity, nature has "complementary" aspects; an experiment which illuminates one of these aspects necessarily simultaneously obscures the complementary aspect. To put it differently, each experiment or sequence of experiments yields only a limited amount of information about the system under investigation; as this information is gained, other equally interesting information (which could have been obtained from another sequence of experiments) is lost. Of course, the experimenter does not forget the results of previous experiments, but at any instant, only a limited amount of information is usable for predicting the future course of the system.

The well-known double-slit experiment provides a good illustration of the principle of complementarity. Light from a monochromatic point source P (**Fig. 2**) is diffracted by the two slits S_1 and S_2 in the screen Q_1. On the screen Q_2, an interference pattern of alternating bright and dark bands is formed in the region D_1D_2, where the two diffraction patterns (from the slits S_1 and S_2) overlap. Assuming that P is equidistant from the slits, and also that the slits are very narrow compared to the distances PS_1 or S_1O, it follows that interference maxima (bright bands) are observed whenever $S_2O - S_1O$ equals $n\lambda$, λ being the wavelength of the light and n an integer; when $S_2O - S_1O = (n + {}^1/_2)\lambda$, interference minima (dark bands) are observed. In moving from any maxima to an adjacent maximum, the path difference $S_2O - S_1O$ changes by precisely one wavelength. Consequently, measurement of the distance Y between successive maxima, and knowledge of S_1S_2 and of the distance X between Q_1 and Q_2, yields λ via the formula $\lambda = Yd/X$ (valid when $S_1S_2 = d$ is much smaller than X). Evidently, the double-slit experiment is understandable in terms of, and provides information concerning, the wave properties of light.

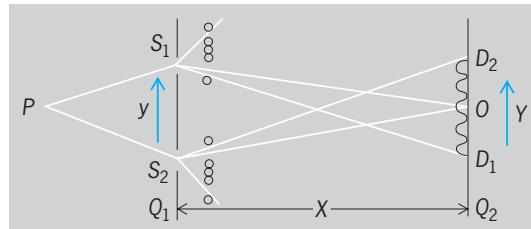

Fig. 2. Double-slit experiment.

The double-slit experiment yields no information concerning the particle properties of light; in fact, introducing the particle picture leads only to conceptual difficulties. These difficulties appear with the recognition that reducing the source intensity does not modify the interference pattern; after a sufficiently long exposure, a photographic film at Q_2 will show exactly the same interference pattern D_1D_2 as is observed by the eye using a more intense source. Since it is possible to make the source intensity so low that two photons almost never will be emitted during the very small time required for light to travel from P to Q_2 via either of the slits, it is necessary to conclude that the interference pattern is produced by independent individual photons, and not by interference between two or more different photons. On the other hand, the interference pattern is destroyed when either of the slits is closed. Thus, the question arises: How can a stream of independent photons, each of which presumably passes through only one of the slits, and half of which on the average pass through S_1, produce an interference pattern that is destroyed by closing one of the slits? Or to put it differently, how can closing or opening a slit through which a photon does not pass affect the likelihood of that photon reaching any particular point on Q_2?

The principle of complementarity meets these difficulties with the assertion that the possibility of

demonstrating that the photons have well-defined trajectories through one or the other slit (a particle-like property) is complementary to the possibility of demonstrating the wavelike property of interference. In the double-slit experimental setup which has been described, until the photon is localized at Q_2 (by the visible evidence that a chemical effect has occurred in a photographic film), it is not possible to locate the photon at any particular point in space, nor is it legitimate to insist that the photon must have passed through only one of the slits. Moreover, according to the principle of complementarity, modifying the experimental setup so as to localize the photon at one of the slits, and thereby to determine through which slit the photon passes, necessarily destroys the interference pattern. This last assertion is supported by analysis of various photon-detection schemes, recognizing that the proposed experiments perforce are entirely *Gedanken* (in the mind); an actual measurement of the slit through which the photon passes demands extreme precision, and this has not been attempted. It is concluded that quantum mechanics involves no inconsistencies of paradoxes. From the standpoint of the complementarity principle, the questions presented in the preceding paragraph, and other similar difficulties, rest always on the specious assumption of more information than actually is obtainable. *See* SUPERPOSITION PRINCIPLE.

The limitations on the experiments are imposed by the requirements of the uncertainty principle, relations (2), and wave-particle duality, Eqs. (3). To illustrate the analysis, suppose indicators (symbolized by circles), free to move in the vertical y direction, are placed immediately behind the slits S_1 and S_2 in Fig. 2; the recoil of an indicator signifies a collision with a photon and places the photon at the indicator. In order that the recoil establish the slit through which the photon has passed, the uncertainty in the vertical location of each indicator must be much less than $d/2$. According to relation (2*a*), this means that the vertical momentum p_y of each indicator will be uncertain by amount $\Delta p_y \gg 2\hbar/d$. There can be no assurance that the indicator has recoiled unless δp_y, the momentum transferred from photon to indicator, equals or exceeds Δp_y. When the slits are narrow and $X \gg d$, a momentum transfer δp_y deflects the photon through an angle $\theta \sim \delta p_y/p$, and changes the point Y where the photon strikes the screen by an amount $\Delta Y = \theta X \cong X\delta p_y/p$, with $p = h/\lambda$, Eq. (3*a*). It follows that when there are observable recoils localizing the photon at one of the slits, then $\Delta Y \gg \lambda X/\pi d$; that is, $\Delta Y \gg Y/\pi$, where $Y/2 = \lambda X/2d$ is the distance between adjacent maxima and minima of the interference pattern. Thus, determining the slit through which the photon passes necessarily gives the photon an uncontrollable random vertical deflection (on the screen Q_2), which is very much larger than the distance between adjacent maxima and minima. The photons now spread with uniform intensity over vertical distances ΔY, which are large compared to the original widths of the light and dark interference bands; in other words, the interference pattern is destroyed.

Quantization. In classical physics the possible numerical values of each observable, meaning the possible results of exact measurement of the observable, generally form a continuous set. For example, the x coordinate of the position of a particle may have any value between $-\infty$ and $+\infty$; similarly p_x, the x component of momentum, may have any value between $-\infty$ and $+\infty$; the kinetic energy T of a particle may have any value between 0 and $+\infty$; the total energy (kinetic plus potential) of an electron in the field of a proton may have any value between $-\infty$ and $+\infty$; the orbital angular momentum vector $\mathbf{l}$ of a particle moving in a central field, for example, an electron in a hydrogen atom, may have any magnitude between 0 and ∞; if the magnitude of $\mathbf{l}$ is known to be l, then, since the plane of the orbit may be arbitrarily oriented, the z component of $\mathbf{l}$ may have any value between $-l$ and $+l$. In quantum mechanics the possible numerical values of an observable need not form a continuous set, however. For some observables, the possible results of exact measurement form a discrete set; for other observables, the possible numerical values are partly discrete, partly continuous; for example, the total energy of an electron in the field of a proton may have any positive value between 0 and $+\infty$, but may have only a discrete set of negative values, namely, -13.6, $-13.6/4$, $-13.6/9$, $-13.6/16$ eV, ... Such observables are said to be quantized; often there are simple quantization rules determining the quantum numbers which specify the allowable discrete values. For example $\mathbf{L}^2$, the square of the orbital angular momentum of a particle, must equal $l(l+1)\hbar^2$, where l is zero or a positive integer; the only allowed values of $\mathbf{J}^2$, the square of the total angular momentum (orbital angular momentum plus intrinsic angular momentum or spin), are $\mathbf{J}^2 = j(j+1)\hbar^2$ where $j = 0, 1/2, 1, 3/2, 2, \ldots$; that is, $j\hbar$ is an integral or half-integral multiple of $h/2\pi$. If the magnitude of the total angular momentum is given by some one of these values, $j\hbar = 3\hbar/2$, for example, then the only allowed values of the z component J_z of $\mathbf{J}$ are $-3\hbar/2, -\hbar/2, \hbar/2, 3\hbar/2$; that is, for a given j, the allowed values of $J_z = -j, -j+1, -j+2, \ldots, j-1, j$, in units of $\hbar$. On the other hand, the observables x, p_x, and T for a relativistic particle are not quantized, and these observables have precisely the same allowed values in quantum mechanics as they do classically. In the formal theory each observable is a linear operator, whose eigenvalues (characteristic values) are the allowed values of that observable. The set of eigenvalues is termed the spectrum of the operator, which spectrum may be discrete, continuous, or mixed. *See* QUANTUM NUMBERS.

The fact that nature is quantized has been amply verified experimentally. For instance, the quantization of energy and momentum in light waves is demonstrated in the photoelectric and Compton effects, described earlier.

Stern-Gerlach experiment. The quantization of angular momentum is strikingly exhibited in the

Stern-Gerlach experiment, wherein a beam of, for example, hydrogen atoms moving through a region of space containing an inhomogeneous magnetic field breaks up into two separate beams. Classically, the force on a hydrogen atom in an inhomogeneous magnetic field depends on the angle between the magnetic field and the plane of the electron's orbit. Thus, classically, the original beam, in which all orientations of the plane of the orbit are possible, is expected to spread out or defocus in the inhomogeneous magnetic field, but is not expected to form two distinct focused beams. Formation of two beams agrees, however, with the quantum mechanical prediction that the square of the total angular momentum of atomic hydrogen is $\mathbf{J}^2 = 3\hbar^2/4 = 1/2(1/2 + 1)\hbar^2$, and therefore that the vector $\mathbf{J}$ (which classically always is perpendicular to the orbit plane) must be either parallel or antiparallel to the applied magnetic field (along z); for these two directions of $\mathbf{J}$, and only for these two, the value of J_z equals one of the permitted values $J_z = -\hbar/2$ or $J_z = \hbar/2$. The Stern-Gerlach experiment also can be performed with beams of other atoms and molecules, as well as with neutron beams. In all cases the observations agree with quantum-mechanical expectation. Refinements of the experiment yield accurate measurements of the parameters characterizing the magnetic interactions of atoms, molecules, and atomic nuclei. *See* MAGNETIC RESONANCE; MOLECULAR BEAMS; NUCLEAR MOMENTS.

Atomic spectra. Spectroscopy, especially the study of atomic spectra, probably provides the most detailed quantitative confirmation of quantization. Granting that the energy in a light wave is quantized, it follows from conservation of energy and Eq. (3*b*) that when an atom emits a photon of frequency f, its initial energy E_i and final energy E_f after emission are related by Eq. (5*a*). Similarly, when a photon of frequency f is absorbed, the initial energy E'_i, and final energy E'_f after absorption, satisfy Eq. (5*b*).

$$E_i - E_f = hf \tag{5a}$$

$$E'_f - E'_i = hf \tag{5b}$$

In quantum mechanics, as in classical mechanics, an electron remains in the vicinity of a proton, that is, it is "bound" to the proton, when and only when the magnitude of its kinetic energy T is less than the magnitude of its negative potential energy V (V is set equal to zero at infinite separation). Thus, the total energy of a stable hydrogen atom necessarily is negative; in other words, the only allowed energies of atomic hydrogen are $-R/n^2$, where n is an integer and $R = -13.6$ eV. Consequently, the radiation emitted by atomic hydrogen must consist of a discrete set of frequencies, or lines, obeying the relation, from Eq. (5*a*), given by Eq. (6),

$$f = \frac{R}{h}\left(\frac{1}{n^2} - \frac{1}{m^2}\right) \tag{6}$$

where m and n are integers, $m > n$. This simple argument provides a convincing explanation of the observation that atomic hydrogen has a line spectrum and not a continuous spectrum; in other words, it radiates discrete frequencies rather than a continuous band of frequencies. The observed lines very accurately satisfy Eq. (6); moreover, except for small relativistic and quantum field theory effects, the nonrelativistic Schrödinger equation accurately predicts not merely the frequencies of the lines, but also their relative intensities and widths. Because on the average the levels E_i and E_f of Eq. (5*a*) endure only for a limited time, their energies are uncertain by an amount ΔE, as explained in the earlier discussion of the uncertainty principle; therefore, the observed lines have a width not less than $\Delta f \cong h^{-1}(\Delta E_i + \Delta E_f)$, according to Eq. (5*a*). Heavier atoms have more complicated line spectra than hydrogen, and the frequencies they emit cannot be described by formulas as simple as Eq. (6). For these atoms, the agreement between experiment and the predictions of the nonrelativistic Schrödinger equation, though always very good, is not as precise as for hydrogen. All the evidence indicates, however, that the discrepancies arise because the Schrödinger equation cannot be solved exactly in many-electron atoms, so that the theoretical predictions necessarily are approximate. If approximations were not necessary (for example, if perfect computing machines were available), there is every reason to think that the predictions of the Schrödinger equation would be exactly correct, except for the aforementioned small relativistic and field theory effects.

Probability considerations. The uncertainty and complementarity principles, which limit the experimenter's ability to describe a physical system, must limit equally the experimenter's ability to predict the results of measurement on that system. Suppose, for instance, that a very careful measurement determines that the x coordinate of a particle is precisely $x = x_0$. This is permissible in nonrelativistic quantum mechanics. Then formally, the particle is known to be in the eigenstate corresponding to the eigenvalue $x = x_0$ of the x operator. Under these circumstances, an immediate repetition of the position measurement again will indicate that the particle lies at $x = x_0$; if the particle is moving in a one-dimensional force field, described by the potential $V(x)$, the particle's potential energy will be exactly $V(x_0)$. Knowing that the particle lies at $x = x_0$ makes the momentum p_x of the particle completely uncertain, however, according to relation (2*a*). A measurement of p_x immediately after the particle is located at $x = x_0$ could yield any value of p_x from $-\infty$ to $+\infty$ a measurement of $T = p_x^2/2m$ could yield any value from 0 to $+\infty$, and in fact the average or expectation value of T in these circumstances would be infinite.

More generally, suppose the system is known to be in the eigenstate corresponding to the eigenvalue α of the observable A. Then for any observable B, which is to some extent complementary to A, that

is, for which an uncertainty relation of the form of relations (2) limits the accuracy with which A and B can simultaneously be measured, it is not possible to predict which of the many possible values $B = \beta$ will be observed. However, it is possible to predict the relative probabilities $P_\alpha(\beta)$ of immediately thereafter finding the observable B equal to β, that is, of finding the system in the eigenstate corresponding to the eigenvalue $B = \beta$. If the system is prepared in the eigenstate α of A a great many times, and each time the observable B is measured immediately thereafter, the average of these observed values of B will equal the expectation value of B, defined by Eq. (7), summed over all eigenvalues

$$\langle B \rangle = \sum \beta P_\alpha(\beta) \tag{7}$$

of B; when the spectrum is continuous, the summation sign is replaced by an integral. To the eigenvalues correspond eigenfunctions, in terms of which $P_\alpha(\beta)$ can be computed. In particular, when α is a discrete eigenvalue of A, and the operators depend only on x and p_x, the probability $P_\alpha(\beta)$ is postulated as in Eq. (8), where $u(x,\alpha)$ is the eigenfunction corre-

$$P_\alpha(\beta) = \left| \int_{-\infty}^{\infty} dx\, v^*(x,\beta) u(x,\alpha) \right|^2 \tag{8}$$

sponding to $A = \alpha$; $\upsilon(x,\beta)$ is the eigenfunction corresponding to $B = \beta$; and the asterisk denotes the complex conjugate. Since measurement of A in the state $A = \alpha$ must yield the result $A = \alpha$, it is necessary that the states $u(x,\alpha)$ satisfy the normalizing and orthogonalizing relation in Eq. (9), where $\delta_{\alpha\alpha'} = 0$

$$\left| \int_{-\infty}^{\infty} dx\, u^*(x,\alpha') u(x,\alpha) \right|^2 = \delta_{\alpha\alpha'} \tag{9}$$

when $\alpha \neq \alpha'$ (eigenfunctions normalized); the eigenfunctions $\upsilon(x,\beta)$ are similarly orthonormal. The integral in Eq. (8) is called the projection of $u(x,\alpha)$ on $\upsilon(x,\beta)$. The projection of the eigenfunction corresponding to $A = \alpha$ on the eigenfunctions of the x operator is $u(x,\alpha)$, and $|u(x,\alpha)|^2\, dx$ is the probability that the system, known to be in the eigenstate $A = \alpha$, will be found in the interval x to $x + dx$. *See* EIGENFUNCTION; EIGENVALUE (QUANTUM MECHANICS).

The formalism just described embodies the essential feature that each measurement on an individual system, as it develops new information, necessarily loses or makes untrue some information gained in the past; in fact, this formalism leads to a rigorous derivation of uncertainty relations (2). For example, suppose B is measured with the system in the state $A = \alpha$, and it is found that $B = \beta$ exactly. The act of measurement necessarily and unavoidably disturbs the system, with the result that after the measurement, the system is in the eigenstate $\upsilon(\beta,x)$. After the measurement, therefore, it no longer is certain that $A = \alpha$; rather the probability of finding $A = \alpha$ is $P_\beta(\alpha)$, which by Eq. (8) equals $P_\alpha(\beta)$. Thus, after starting with $A = \alpha$, and then measuring $B = \beta$, it is possible to find that $A = \alpha' \neq \alpha$. Of course, as stressed previously, these considerations are unimportant for macroscopic systems, where the limitations imposed by the uncertainty principle are inconsequential. Even if all observations were extremely accurate by usual standards, when the momentum p_x of a 1-g mass at x is measured, the position x' immediately thereafter should be indistinguishable from x. Were x' and x distinguishable, the particle position seemingly would have changed discontinuously from x to x', contrary to all classical (macroscopic) experience.

This formalism yields predictions in excellent agreement with observation; furthermore, it can be seen that the formalism is internally consistent. Consequently, the following doctrine, embodied in a formalism, appears well established: Although it is possible to predict the average of a large number of observations on identical systems, the result of a measurement on a single (microscopic) system generally is unpredictable and largely a matter of chance. Nonetheless, some physicists have refused to accept this inherent indeterminancy of nature and believe that this doctrine is a serious deficiency of present physical theory. To put the problem in simplest terms, consider a gram of radium, containing approximately 10^{21} atoms. According to generally accepted theory, it is not possible to predict when any one atom will decay, but it is possible to predict very accurately the average number of atoms decaying every second. The objectors to this doctrine feel that it must be possible to predict the subsequent history of every individual atom; failure to do so represents, not an inherent indeterminism in nature, but rather a lack of obtainable information—and therefore a lack of understanding—concerning the mechanism of radioactive decay. To mention but one possible alternative, nonrelativistic quantum theory can be reinterpreted in terms of hidden variables, which in principle determine the precise behavior of an individual system but whose values are not ascertained in measurements of the type which now can be carried out. This alternative has not led to new predictions, however, and contains some unappealingly ad hoc features. *See* CAUSALITY; PROBABILITY (PHYSICS).

Wave function. When the system is known to be in the eigenstate corresponding to $A = \alpha$, the eigenfunction $u(x,\alpha)$ is the wave function; that is, it is the function whose projection on an eigenfunction $\upsilon(x,\beta)$ of any observable B gives the probability of measuring $B = \beta$. The wave function $\psi(x)$ may be known exactly; in other words, the state of the system may be known as exactly as possible (within the limitations of uncertainty and complementarity), even though $\psi(x)$ is not the eigenfunction of a known operator. This circumstance arises because the wave function obeys Schrödinger's wave equation. Knowing the value of $\psi(x)$ at time $t = 0$, the wave equation completely determines $\psi(x)$ at all future times. In general, however, if $\psi(x,0) = u(x,\alpha)$, that is, if $\psi(x,t)$ is an eigenfunction of A at $t = 0$, then $\psi(x,t)$ will not be an eigenfunction of A at later times $t > 0$. For example, suppose at $t = 0$ a free particle (a particle moving under no forces) is known to

be in an eigenstate for which the uncertainty in x is $(\Delta x)_0$; $(\Delta x)_0$ is approximately the x interval within which $|\psi(x,0)|^2$ is not negligibly small. Suppose further that, at $t = 0$, the product of the uncertainties in position and momentum is as small as possible: $(\Delta x)_0(\Delta p_x)_0 \cong \hbar$, compare relation (2*a*). Then it can be proved that $(\Delta x)_t$, the uncertainty in x at time t, satisfies Eq. (10), where m is the particle mass.

$$(\Delta x)_t = \left[(\Delta x)_0^2 + \frac{t^2}{m^2}(\Delta p_x)_0^2\right]^{1/2} \qquad (10)$$

Equation (10) is readily interpreted. The root-mean-square spread at time t results from $(\Delta x)_0$ and from an uncertainty in the distance the particle has traveled; the latter uncertainty is $t(\Delta \upsilon_x)_0 = t(\Delta p_x)_0/m$. If $(\Delta x)_0 = 0$, meaning $\psi(x,0)$ is an eigenfunction of the x operator, $(\Delta x)_t$ is infinite, showing that $\psi(x,t)$ cannot be an eigenfunction of the x operator. When the particle is free, the projections of the wave function on the eigenfunctions of the momentum operator do not change with time, corresponding to the classical result that the momentum of a free particle does not change. Thus, the probability of measuring any value of the momentum does not change, $(\Delta p_x)_t = (\Delta p_x)_0$, and Eq. (10) shows that $(\Delta x)_t\ (\Delta p_x)_t$ grows with time for a free particle, whatever the value of $(\Delta x)_0$, provided $(\Delta x)_0\ (\Delta p_x)_0 \cong \hbar$. Nonetheless, $\psi(x,t)$ is known no less exactly than the initial wave function $\psi(x,0)$. The magnitude of $\Delta x\ \Delta p_x$ at any instant is no measure of the exactness with which the state of the system is known; increased uncertainties in position or momentum may be the price for increased certainty in the value of some other observable.

A system described by a wave function is said to be in a pure state. Not all systems are described by wave functions, however. Consider, for example, a beam of hydrogen atoms streaming in the x direction out of a small hole in a hydrogen discharge tube. According to the formal theory, if the beam were described by a wave function $\psi(x)$, then Eq. (11)

$$\psi(x) = C_+(x)\, u^+(z) + C_-(x)\, u^-(z) \qquad (11)$$

holds, where $u^+(z)$ is the eigenfunction corresponding to finding a hydrogen atom with its z component J_z of total angular momentum equal to $\hbar/2$; $u^-(z)$ is the corresponding eigenfunction for finding $J_z = -\hbar/2$; $|C_+(x)|^2$ is the probability of finding $J_z = \hbar/2$ at any point x along the beam; $|C_-(x)|^2$ is the corresponding probability of finding $J_z = -\hbar/2$. Since there are only two possibilities, $J_z = \pm\hbar/2$, $|C_+(x)|^2 + |C_-(x)|^2 = 1$, and since there seems no reason to favor either of these possibilities, it is reasonable to suppose that $|C_+(x)|^2 = |C_-(x)|^2 = {}^1/_2$. As Eq. (11) shows, however, to specify $\psi(x)$ it is necessary to know not merely the relative magnitudes of the complex numbers $C_+(x)$ and $C_-(x)$ but also their relative phase. It can be shown that each choice of relative magnitude and phase of C_+ and C_- corresponds to a direction γ for which there is probability 1 of finding $J_\gamma = \hbar/2$, and probability zero of finding $J_\gamma = -\hbar/2$. Thus each choice of relative magnitude and phase of C_+ and C_- puts the system in an eigenfunction $u^+(\gamma)$ or $u^-(\gamma)$, that is, in a pure state. On the other hand, the discharge tube singles out no particular direction in space, so that in a Stern-Gerlach experiment the original beam must break up into two beams of equal intensity, whatever the direction of the external magnetic field. Consequently, the original beam is not in a pure state, but can be regarded as a statistical ensemble or mixture of pure states oriented with equal probability in all directions. Equivalently, Eq. (11) can be used for the original beam, provided calculations are averaged over all relative phases of C_+ and C_-. The distinction between mixtures and pure states is strongly analogous to the distinction between polarized and unpolarized light beams; consequently, beams of particles in pure spin states are termed polarized.

Schrödinger equation. Equation (12) describes a

$$\psi(x, t) = A(\lambda) \exp\left[2\pi i\left(\frac{x}{\lambda} - ft\right)\right] \qquad (12)$$

plane wave of frequency f, wavelength λ, and amplitude $A(\lambda)$, propagating in the positive x direction. The previous discussion concerning wave-particle duality suggests that this is the form of the wave function for a beam of free particles moving in the x direction with momentum $p = p_x$, with Eq. (3) specifying the connections between f, λ and E, p. Differentiating Eq. (12), it is seen that Eqs. (13) hold. Since for a

$$p_x\psi = \frac{h}{\lambda}\psi = \frac{\hbar}{i}\frac{\partial\psi}{\partial x} \qquad (13a)$$

$$E\psi = hf\psi = -\frac{\hbar}{i}\frac{\partial\psi}{\partial t} \qquad (13b)$$

free particle $E = p^2/2m$, it follows also that Eq. (14) is valid.

$$\frac{-\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} = -\frac{\hbar}{i}\frac{\partial\psi}{\partial t} \qquad (14)$$

See WAVE MOTION.

Equation (14) holds for a plane wave of arbitrary λ, and therefore for any superposition of waves of arbitrary λ, that is, arbitrary p_x. Consequently, Eq. (14) should be the wave equation obeyed by the wave function of any particle moving under no forces, whatever the projections of the wave function on the eigenfunctions of p_x. Equations (13) and (14) further suggest that for a particle whose potential energy $V(x)$ changes, in other words, for a particle in a conservative force field, $\psi(x,t)$ obeys Eq. (15).

$$\frac{-\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V(x)\psi = -\frac{\hbar}{i}\frac{\partial\psi}{\partial t} \qquad (15)$$

Equation (15) is the time-dependent Schrödinger equation for a one-dimensional (along x), spinless particle. Noting Eq. (13*b*), and observing that Eq. (15) has a solution of the form of Eq. (16), it is

$$\psi(x, t) = \psi(x)\ \exp(-iEt/\hbar) \qquad (16)$$

inferred that $\psi(x)$ of Eq. (16) obeys the time-independent Schrödinger equation (17).

$$\frac{-\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V(x)\psi = E\psi \tag{17}$$

See FORCE.

Equation (17) is solved subject to reasonable boundary conditions, for example, that ψ must be continuous and must not become infinite as x approaches $\pm\infty$. These boundary conditions restrict the values of E for which there exist acceptable solutions $\psi(x)$ to Eq. (17), the allowed values of E depending on $V(x)$. In this manner, the allowed energies of atomic hydrogen listed in the earlier discussion of quantization are obtained. When a $\psi(x)$ solving Eq. (17) exists, all probabilities inferred from the corresponding $\psi(x,t)$ are independent of time; see Eq. (8). Thus the allowed energies E of Eq. (17) are the energies of the stationary states (states not changing with time) of the system.

The forms of Eqs. (13*a*), (15), and (17) suggest that the classical observable p_x must be replaced by the operator $(\hbar/i)(\partial/\partial x)$. With this replacement, Eq. (18)

$$(xp_x - p_x x)\psi = i\hbar\psi \tag{18}$$

holds. In other words, whereas the classical canonically conjugate variables x and p_x are numbers, obeying the commutative law in Eq. (19*a*), the quantum-mechanical quantities x and p_x are noncommuting operators, obeying Eq. (19*b*).

$$xp_x - p_x x = 0 \tag{19a}$$

$$xp_x - p_x x = i\hbar \tag{19b}$$

In the formal theory the noncommutativity of x, p_x, leads directly to uncertainty relation (2*a*). For a derivation of relation (2*a*) from Eq. (19*b*), as well as for generalizations and more sophisticated derivations of Eqs. (13), (15), and (19*b*), *See* MATRIX MECHANICS; NONRELATIVISTIC QUANTUM THEORY; HAMILTON'S EQUATIONS OF MOTION.

Correspondence principle. Since classical mechanics and Maxwell's electromagnetic theory accurately describe macroscopic phenomena, quantum mechanics must have a classical limit in which it is equivalent to the older classical theories. Although there is no rigorous proof of this principle for arbitrarily complicated quantum-mechanical systems, its validity is well established by numerous illustrations, such as those mentioned in the preceding discussions of the uncertainty principle and wave-particle duality. In general, the classical limit is approached when (i) $h \to 0$; (ii) the mass becomes large; (iii) wavelengths become small; (iv) dimensions become large; (v) quantum numbers become large. (The notation $h \to 0$ refers to a mathematical operation in which one adheres to a fixed set of values of the other quantities involved and considers the effect of making h smaller and smaller.) These simple criteria must be employed cautiously, since they have not been stated in terms of dimensionless parameters; obviously the classical limit, in (ii) for instance, cannot depend on the unit of mass. Nonetheless, these criteria are useful guides; for example, Eq. (19*a*) is the limit of Eq. (19*b*) as $h \to 0$.

Before the introduction of the Schrödinger equation (1926) made possible exact determination of energy levels and related quantum numbers, the correspondence principle was very effectively employed as a heuristic means of arriving at the quantization rules. In particular, for periodic orbits there was evolved the rule in Eq. (20), when n is an integer,

$$\oint p\,dq = nh \tag{20}$$

q and p are canonically conjugate position and momentum variables, and the integration is performed along the orbit for one complete cycle. *See* HAMILTON-JACOBI THEORY; PERTURBATION (QUANTUM MECHANICS).

Equation (20) is not always exact, but nonetheless is useful, especially in the limit of high quantum numbers. In the case of a one-dimensional harmonic oscillator, for example, whose classical frequency is f and whose potential energy is $V(x) = 2\pi^2 mf^2x^2$, Eq. (20) implies that the allowed energies are as defined by Eq. (21*a*), whereas the correct result, deduced from the Schrödinger equation, is given by Eq. (21*b*).

$$E = nhf \tag{21a}$$

$$E = (n + 1/2)hf \tag{21b}$$

See HARMONIC OSCILLATOR.

The existence of the zero-point energy $^1/_2hf$ in the ground state $n = 0$ has been confirmed in analyses of molecular spectra and has significance for many phenomena; for example, *see* IONIC CRYSTALS.

Quantum statistics; indistinguishability. Equation (20) means that the allowed orbits, plotted as functions of q and p, have area nh the area between two allowed orbits equals h. Since in classical statistical mechanics (i) all orbits are allowed, and (ii) the statistical weight of a volume $dq\,dp$ of q, p (phase) space is proportional to $dq\,dp$, the correspondence principle suggests that in quantum statistics each allowed orbit replaces a set of classical orbits which cover an area h when plotted in the q, p plane. It follows that in quantum statistics: (i) each allowed quantized orbit should have the same statistical weight, which may be set equal to unity; and (ii) with unit weight for a quantized orbit, the weight of an unquantized volume $dp\,dq$ of phase space must be $(dp\,dq)/h$. Because the average number of atoms in a state of energy E is proportional to $\exp(-E/kT)$, where k is Boltzmann's constant and T the absolute temperature, it is reasonable that quantum statistics yields different predictions than classical (Boltzmann) statistics when the energy-level spacing is large compared to kT. However, in the limit that the level spacing becomes small compared to kT, which corresponds to

the limit $h \to 0$, quantum statistics must be equivalent to Boltzmann statistics. The quantum theory of specific heat and the theory of blackbody radiation illustrate and confirm the considerations of this paragraph. *See* BOLTZMANN STATISTICS; HEAT RADIATION; QUANTUM STATISTICS; SPECIFIC HEAT OF SOLIDS; STATISTICAL MECHANICS.

Quantum statistics is further complicated by the fact that identical particles are indistinguishable. Classically, if two He atoms in their ground states are placed in a box, the statistical weights of states are computed as if the atoms can be distinguished, that is, as if each atom had an identifying mark. Quantum mechanics rejects the possibility of this identification and simultaneously modifies the statistical weights which otherwise would be used in computing statistical averages. Formally, this modification is accomplished by insisting that for a system of identical particles, the wave function must be symmetric in the coordinates of all the particles, including spin coordinates specifying spin orientations, if the particles have integral spin, namely, 0, $\hbar$, $2\hbar$, . . . ; and antisymmetric in the coordinates of all the particles, including spin coordinates, if the particles have half-integral spin, namely, $\hbar/2$, $3\hbar/2$, $5\hbar/2$, ... *See* BOSE-EINSTEIN STATISTICS; FERMI-DIRAC STATISTICS.

A symmetric wave function is unchanged under interchange of coordinates, for example, for two particles $\psi(x_1, x_2) = \cos(x_1 - x_2)$ is symmetric; an antisymmetric wave function changes sign under interchange of coordinates, for example, $\psi(x_1, x_2) = \sin(x_1 - x_2)$ is antisymmetric. It is postulated that wave functions having other symmetry properties than these specified here never occur, and therefore must not be included in any enumeration of available states, even though without these symmetry restrictions, they might be acceptable solutions of the Schrödinger equation (for the many-particle system). These symmetry restrictions have profound consequences for macroscopic as well as microscopic systems, and are extremely well established for nuclei as well as atoms. The reason for the connection between spin and statistics is beyond the scope of this article, but appears to be a consequence of requirements imposed by the special theory of relativity. *See* EXCLUSION PRINCIPLE.

Edward Gerjuoy

Bibliography. L. E. Ballentine, *Quantum Mechanics*, 2d ed., 1998; A. Bohm, *Quantum Mechanics: Foundations and Applications*, 3d ed., 1994; R. Eisberg and R. Resnick, *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles*, 2d ed., 1985; A. Messiah, *Quantum Mechanics*, vol. 1, 1961, vol. 2, 1962; H. Ohanian, *Principles of Quantum Mechanics*, 1989; R. Shankar, *Principles of Quantum Mechanics*, 2d ed., 1994.

Quantum numbers

The quantities, usually discrete with integer or half-integer values, which are needed to characterize a physical system of one or more atomic or subatomic particles. Specification of the set of quantum numbers serves to define such a system or, in other words, to label the possible states the system may have. In general, quantum numbers are obtained from conserved quantities determinable by performing symmetry transformations consisting of arbitrary variations of the system which leave the system unchanged. For example, since the behavior of a set of particles should be independent of the location of the origin in space and time (that is, the symmetry operation is translation in space-time), it follows that momentum and energy are rigorously conserved. In similar fashion, the invariance of the system to arbitrary rotations of the three axes defining three-dimensional space leads to conservation of total angular momentum. Inversion of these three axes through their origin is likewise a symmetry operation and leads to conservation of parity. (Parity is not conserved in weak interactions and cannot be used as a quantum number to specify the state of a set of particles undergoing such interactions. For such a case, parity is not a "good" quantum number.) For a list of invariance or symmetry operations and the corresponding conserved quantities *see* SYMMETRY LAWS (PHYSICS)

Of these conserved quantities, the common ones used to label physical states are momentum, energy, angular momentum, charge, baryon number, lepton number, parity, charge parity, product of parity and charge parity, isotopic spin and third component thereof, hypercharge, isotopic parity (or G parity), and unitary spin. In general, each physical system must be studied individually to find the symmetry transformations, and thus the conserved quantities and possible quantum numbers. The quantum numbers themselves, that is, the actual state labels, are usually the eigenvalues of the physical operators corresponding to the conserved quantities for the system in question. *See* EIGENVALUE (QUANTUM MECHANICS); ELEMENTARY PARTICLE; NONRELATIVISTIC QUANTUM THEORY; PARITY (QUANTUM MECHANICS).

It is not necessary that the conserved quantity be quantized in order to be regarded as a quantum number; for example, a free particle possesses energy and momentum, both of which can have values from a continuum but which are used to specify the state of the particle. For a particle at rest the momentum is zero, and the energy is given by the rest mass; the additional specification of values for the remaining (discrete) quantities in the list mentioned above completely identifies the particle.

The specification of the quantum numbers for a set of particles constitutes a spectroscopy. In the 1920s and 1930s the systematic study of the energy levels of atoms and molecules became known as atomic and molecular spectroscopy. In the 1960s and 1970s similar intensive studies of the quantum numbers of strongly interacting particles, mesons and baryons, was characterized as hadron spectroscopy. In the 1980s another spectroscopy

involving heavy quarks developed. The quarks, though still remaining unobserved particles postulated as having some 13 fractional quantum numbers, have yielded dramatically fruitful predictions for new experiments and new particles which have been observed. The special quantum numbers, known as flavors, which the quarks carry are labeled *u, d, s, c, b*, and *t*, where the first three are constituent labels of the older familiar particles of hadron spectroscopy, for example, protons, neutrons, and pi mesons. The (heavy) quark carrying the *c* (charm) quantum number was indirectly discovered in 1974 and that carrying the *b* (beauty) quantum number in 1977. The study of the levels of these quark-antiquark systems $c\bar{c}$ and $b\bar{b}$ is often characterized as quarkonium spectroscopy. The *t* quark is much heavier yet, and appears to have produced some experimental indications of its existence in a 1984 study at CERN, Geneva, Switzerland. A special quantum number given the name color is believed to be carried by the massless field exchanged between quarks which is responsible for their binding in particles. *See* COLOR (QUANTUM MECHANICS); QUANTUM CHROMODYNAMICS; QUARKS.

Some important numbers. Probably the most familiar quantum numbers are those associated with the restricted (or quantized) energy values which appear naturally in quantum mechanics whenever one or more particles are confined to a small region of space. The wave equation (Schrödinger equation in the nonrelativistic case) for the confined particle can be solved to express these allowed energy values in terms of one or more quantum numbers. The number of nodes, or zeros, that the solution of the wave equation possesses always enters such an expression and is usually called the principal quantum number *n*. In the case of a charged particle confined to the vicinity of an oppositely charged particle according to Coulomb's law (for example, the electron to the proton in a hydrogen atom), the Schrödinger equation yields possible energies, $-E_B/n^2$, where $n = 1, 2, 3, \ldots$ and for hydrogen $E_B = 13.6$ eV. The magnitude of *n* is also directly related to the most probable distance from the center of the confining region where the electron may be found. *See* ATOMIC STRUCTURE AND SPECTRA; COULOMB'S LAW.

The wave equation for the confined particle also leads to quantized angular momentum, with possible values or quantum numbers $L = 0, 1, 2, 3, \ldots$ in units of $\hbar$ (Planck's constant divided by 2π). In addition, each particle, whether confined or not, possesses an intrinsic spin angular momentum *S* with possible values $0, {}^1/_2, 1, {}^3/\sqrt{2}, \ldots$, in units of $\hbar$. The electron, for example, has a spin of ${}^1/_2$. *See* SPIN (QUANTUM MECHANICS).

Since angular momentum is a quantized vector quantity (a direction in space is associated with it), it might be expected that the three space components would also be quantized observables. However, only one, the *Z*-axis projection, is observable and the associated quantum number m_L has possible values $-L$, $-L + 1, \ldots, L - 1, L$; correspondingly, the spin angular momentum M_S has possible values $-S$, $-S + 1, \ldots, S - 1, S$. These are often called magnetic quantum numbers because they determine how the energy changes when a system with given *n, L, S* is placed in a magnetic field.

Determination of numbers. Thus the properties or state of a confined electron (for example, in the hydrogen atom) are known when the quantum numbers *n*, L, M_L, and M_S ($S = {}^1/_2\hbar$) are specified. Actually, when two or more angular momenta enter, as for the confined electron with an intrinsic $S = {}^1/_2\hbar$ in a state with orbital angular momentum *L* or especially for a system with more than one particle, the overall rotational symmetry of the system is best expressed by a total angular momentum quantum number *J* and a total magnetic quantum number *M*. The resultant *J* is obtained by vectorially adding all individual angular momenta, and *M* is obtained by summoning all the separate M_S's and M_L's. For a many-electron atom or many-nucleon nucleus, the preferred way to perform the vector addition is often indicated by the dynamics, that is, how the individual particle *L* and *S* values enter into interparticle interactions. In such a case an intermediate result in the determination of *J* may be a good quantum number; for example, the seniority quantum number is related to the number of particles whose spins are paired. The objective of atomic and nuclear spectroscopy is the determination of the quantum numbers for the states of many-electron and many-nucleon systems, respectively. Nucleon systems have the additional quantum numbers of isotopic spin and third-axis projection. *See* ELECTRON CONFIGURATION; NUCLEAR STRUCTURE.

Kenneth E. Lassila

Bibliography. J. P. Elliot and P. G. Dawber, *Symmetry in Physics*, 2 vols., 1979, paper 1985; D. Flamm and F. Schoberl, *Quantum Numbers, Gauge Theories and Hadron Spectroscopy*, 1982; W. Ludwig and C. Falter, *Symmetries in Physics*, 2d ed., 1996; U. Mosel, *Fields, Symmetries, and Quarks*, 2d ed., 1999; L. H. Ryder, *Elementary Particles and Symmetries*, 2d ed., 1986.

Quantum solids

Solids in which the quantum-mechanical wave functions of individual atoms overlap the wave functions of neighboring atoms in the quantum ground state of the system. The spatial extension of the wave functions is called zero-point motion since the location of the atoms cannot be determined within the width of the wave function even at the absolute zero of temperature. The zero-point motion becomes large when the constituent atoms have small mass and the attractive force between them is small. The isotopes of helium have the weakest interaction of any atom or molecule (and the smallest mass except for the much more strongly interacting

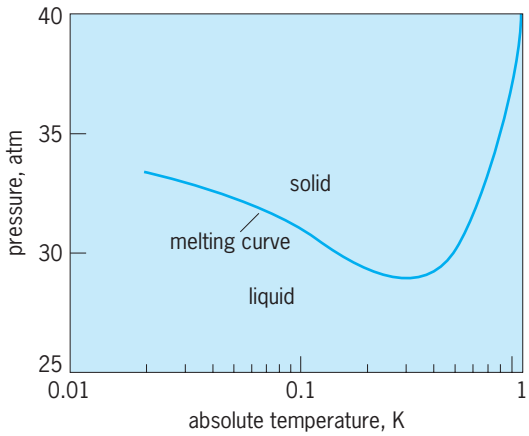

Fig. 1. Melting curve for solid ^{3}He. (*From data of E. R.Grilly, J. Low Temp. Phys., 11:33–52, 1973*)

hydrogen molecules). The root-mean-square zero-point motion in those solids is approximately 25% of the mean distance between atoms. This zero-point motion results in some very unusual properties that are manifestations of the quantum statistical mechanics of many-particle systems. The properties of solid helium are currently an active subject of research with some fundamental questions remaining unanswered. *See* HELIUM; INTERMOLECULAR FORCES; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; QUANTUM STATISTICS.

Melting pressure of ^{3}He and ^{4}He. Unlike all other solids, helium cannot be solidified by simply cooling liquid helium to low temperatures. The liquid must be compressed to at least 25 bar for ^{4}He and 29 bar for ^{3}He in order to freeze (1 bar = 10^5 Pa ≈ 1 atm). In further contrast with other solids, the pressure required to freeze the isotopes of helium increases with decreasing temperature below a certain temperature (**Fig. 1**). In that temperature range, a mixture of liquid and solid must be heated in order to freeze it. A general result of thermodynamics states that the pressure P and temperature T of any melting curve are related by the equation

$$dP/dT_{\text{melting}} = \frac{\Delta s}{\Delta v}$$

Here Δs is the liquid entropy (per mole) minus the solid entropy, and Δv is the liquid volume (per mole) minus the solid volume. When the slope of the melting curve, dP/dT, is negative, the entropy of the liquid is less than that of the solid, since the phase that exists at lower pressure must always have larger molar volume (Δv must always be positive). For all other materials, the solid state has lower entropy than the liquid due to the spatial order of the atoms in the crystal lattice. For helium, the entropy of the liquid is lower than the solid due to quantum statistics. *See* ENTROPY.

Quantum statistics becomes important when the wave functions of individual atoms or molecules overlap. In liquid helium at low temperatures, where thermal motion of the atoms becomes small, the overlap is so great that essentially all of the wave functions overlap each other. However, quantum statistics takes two different forms for the two fundamental types of particles so that a many-particle system consisting entirely of one type or the other will have very different properties at low temperatures. ^{3}He atoms are "Fermi" particles (fermions) for which only one atom can occupy a given quantum state. The quantum states are specified by a momentum quantum number and a spin quantum number. The latter has only two possible values for ^{3}He and represents the orientation of the magnetic moment of the atom with respect to some reference direction such as a magnetic field. Consequently, there are two atoms with each momentum quantum number, and their magnetic moments must be oriented antiparallel so that the bulk magnetization of liquid ^{3}He decreases as temperature decreases. In solid ^{3}He, the wave functions are much more localized so that the magnetization does not decrease until much lower temperatures. At temperatures higher than the ordering of the moments in the solid, the entropy of the liquid is lower than the solid. *See* FERMI-DIRAC STATISTICS.

^{4}He atoms are "Bose" particles (bosons) in which any number of atoms can occupy the same quantum state. At low temperatures, this leads to Bose-Einstein condensation to a highly ordered, low-entropy state in which a large fraction of the atoms are in the same quantum state below a critical temperature. Bose-Einstein condensation is responsible for the remarkable states of superfluidity and superconductivity. In the solid phase, the greater localization suppresses Bose-Einstein condensation. *See* BOSE-EINSTEIN CONDENSATION; BOSE-EINSTEIN STATISTICS; LIQUID HELIUM; SUPERCONDUCTIVITY; SUPERFLUIDITY.

Quantum statistics in solid phases. Although the wave functions overlap much less in the solid than in the liquid, there are still consequences of the quantum statistics.

As ^{3}He atoms are forced closer together, the effect of the statistics is to tend toward some type of

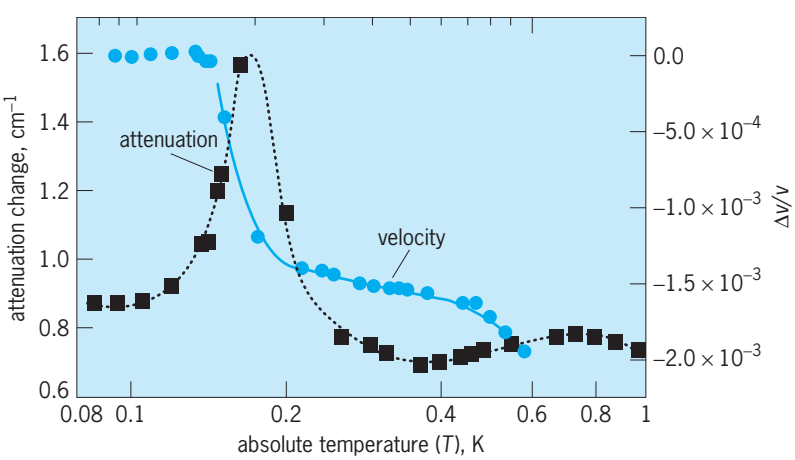

Fig. 2. Changes in the attenuation and speed of sound near 0.2 K for a sample of solid ^{4}He with 27 ppm of ^{3}He impurities. The attenuation is relative to an arbitrary reference value, and the velocity shift, $\Delta v/v$, is the fractional change relative to the velocity at the lowest temperature measured.

antiparallel alignment of neighboring pairs of atoms. This effect of the statistics is called the quantum exchange interaction. It has no analog in Newtonian physics. Although there is an interaction between the magnetic dipole moments of the nuclei, it is smaller than the exchange interaction by a factor of 10^{-4}. The exchange interaction in ^{3}He produces an unusual antiferromagnetic state below about 10^{-3} K. The magnetic moments in two adjacent planes of atoms are oriented in the same direction, and the moments in the next two planes are oriented in the opposite direction. *See* ANTIFERROMAGNETISM; EXCHANGE INTERACTION.

An analogous phenomenon in solid ^{4}He would be some amount of Bose-Einstein condensation at much lower temperatures than the condensation in liquid ^{4}He. If true, this would represent a peculiar state of matter which could display some of the properties of a superfluid as well as the spatial order of a solid. This possibility was proposed theoretically, and some experiments have found behavior suggestive of some type of a phase transition. Proof of its existence is an active topic of current research. The addition of as little as about 20 parts per million of ^{3}He to solid ^{4}He produced a sharp peak in the attenuation of sound and a shift in the speed of sound at about 0.2 K (**Fig. 2**). Some of the observed properties below 0.2 K were also produced by straining crystals without impurities. In a completely different experiment, the rotational moment of inertia of solid helium has been measured by placing the helium solid in a cylindrical container suspended on a torsion fiber. The period of the torsion pendulum was measured as a function of temperature and found to decrease below about 0.2 K (**Fig. 3**). *See* LOW-TEMPERATURE ACOUSTICS; MOMENT OF INERTIA; PHASE TRANSITIONS.

A decrease in the moment of inertia means that some of the mass of the solid is not moving with the container. This could be direct evidence that some of the solid becomes superfluid. However, the phenomenon appears to depend on the ^{3}He impuritiy concentration and the strain of the solid since it becomes smaller and moves to lower temperatures for the pure solid and when the solid is annealed. A test for the flow of solid through small openings showed that, if there is supersolid, it does not flow in the same manner as superfluid. No flow was observed down to flow rates seven orders of magnitude smaller than would have corresponded to the observed velocities in the torsion pendulum. A measurement of the melting pressure in 2006 revealed unexpected behavior below 0.08 K, but it is not consistent with a supersolid phase transition.

Molecular quantum solids. The molecules H_2, HD, D_2, N_2, and a few others also have sufficient zero-point motion that some quantum solid properties might be expected. In the various isotopic combinations of solid hydrogen, these are manifested through the influence of quantum tunneling on nuclear magnetic resonance (NMR) relaxation times and on the orientational order of the molecules. *See* HYDROGEN; NUCLEAR MAGNETIC RESONANCE (NMR); TUNNELING IN SOLIDS.

John M. Goodkind

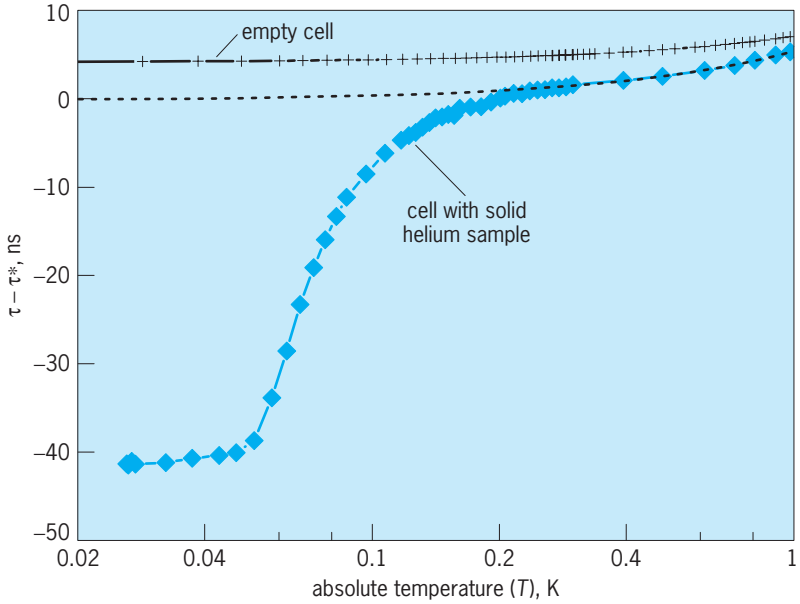

Fig. 3. Variation of the period of a torsion pendulum containing solid ^{4}He as a function of temperature, indicating the decrease in the moment of inertia of the solid below about 0.2 K. The pendulum oscillation amplitude, for these data, was set so that the maximum velocity was 4 μm/s. Temperature dependence of an empty cell was also measured in order to be certain that there was no effect intrinsic to the cell rather than to solid helium. Vertical scale is $\tau - \tau^*$, where τ is the period of the pendulum and τ^* is the value of the period at the temperature where the period starts to decrease.

Bibliography. J. Day and J. Beamish, Pressure-driven flow of solid helium, *Phys. Rev. Lett.*, 96:105304, 2006; P.-C. Ho, I. P. Bindloss, and J. M. Goodkind, A new anomaly in solid ^{4}He, *J. Low Temp. Phys.*, 109:409421, 1997; E. Kim and M. H. W. Chan, Observation of superflow in solid helium, *Science*, 305:1941–1944, 2004; D. D. Osheroff, Nuclear magnetic order in solid ^{3}He, *J. Low Temp. Phys.*, 87:297–342, 1992; I. A. Todoshchenko et al., Melting curve of ^{4}He: No sign of a supersolid transition down to 10 mK, *Phys. Rev. Lett.*, 97:165302, 2006; J. S. Xia, W. Ni, and E. D. Adams, Melting pressure of solid ^{3}He through the magnetic ordering transition, *Phys. Rev. Lett.*, 70:1481–1484, 1993.

Quantum statistics

The statistical description of particles or systems of particles whose behavior must be described by quantum mechanics rather than by classical mechanics. As in classical, that is, Boltzmann statistics, the interest centers on the construction of appropriate distribution functions. However, whereas these distribution functions in classical statistical mechanics describe the number of particles in given (in fact, finite) momentum and positional ranges, in quantum statistics the distribution functions give the number of particles in a group of discrete energy levels. In an individual energy level there may be, according to quantum mechanics, either a single particle or any number of particles. This is determined by the symmetry character of the wave functions. For

antisymmetric wave functions only one particle (without spin) may occupy a state; for symmetric wave functions, any number is possible. Based on this distinction, there are two separate distributions, the Fermi-Dirac distribution for systems described by antisymmetric wave functions and the Bose-Einstein distribution for systems described by symmetric wave functions.

In relativistic quantum theory it is shown that particles having integer spin necessarily obey Bose-Einstein statistics, while those having half-integer spin necessarily obey Fermi-Dirac statistics. (Particles obeying Bose-Einstein statistics are often called bosons; particles obeying Fermi-Dirac statistics, fermions.) For sufficiently high temperatures, both forms of distribution functions go over into the familiar Boltzmann distribution, although strictly speaking no system is correctly described by this distribution. In practice, of course, the Boltzmann distribution gives an exceedingly good description of the experiments, but there are situations, such as those involving the behavior of electrons in metals and liquid helium, where the quantum description is essential. *See* BOLTZMANN STATISTICS; BOSE-EINSTEIN STATISTICS; EXCLUSION PRINCIPLE; FERMI-DIRAC STATISTICS; KINETIC THEORY OF MATTER; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; RELATIVISTIC QUANTUM THEORY; SPIN (QUANTUM MECHANICS); STATISTICAL MECHANICS.

Max Dresden

Quantum teleportation

A way to transfer the state of a quantum system over large distances by employing entanglement. Entanglement is a nonclassical connection between objects that Albert Einstein called "spooky."

To be able to travel from one place to another instantly and over arbitrary distances, or at least to move objects in this way, is an ancient dream. The concept of teleportation is frequently utilized in the literature of science fiction to overcome limitations imposed on space travel by the laws of physics.

In the standard science fiction approach, the sender, Alice, scans the object to be teleported in order to read out all the information needed to describe it. She then sends that information to the receiver, Bob, who uses this information to reconstitute the object, not necessarily from the same material as that of the original. However, according to quantum mechanics, it is impossible to succeed in this way. If only one individual object is at hand, it is impossible to determine its quantum state by measurement. The quantum state represents all that can be known about the object, that is, all possible (in general, probabilistic) predictions that can be made about future observations of the object.

In fact, it is quantum mechanics that comes to the rescue and makes quantum teleportation possible using a very deep feature of the theory, quantum entanglement. It is important to realize that there are significant differences between teleportation as portrayed in science fiction and quantum teleportation as realized in the laboratory. In the experiments, what is teleported is not the substance an object is made of but the information it represents.

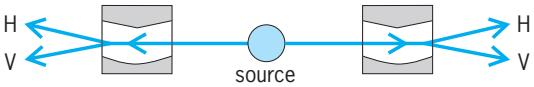

Fig. 1. Principle of quantum entanglement for two photons emitted by a source. Each photon travels to its own two-channel polarizer, each of which can be rotated around the respective beam direction. Independent of the orientation of the polarizer, each detector (H or V) has the same probablility of registering a photon. If the two polarizers are oriented parallel, the two photons will always be registered in different detectors; that is, if one photon is registered in its H detector, the other is registered in its V detector, and vice versa. Yet neither photon carries any polarization before it is measured.

Quantum entanglement. Entangled quantum states as used in teleportation were introduced into the discussion of the foundations of quantum mechanics by Einstein, Boris Podolsky, and Nathan Rosen in 1935. In the same year, Erwin Schrödinger introduced the notion of entanglement, which he called the essence of quantum mechanics.

In order to discuss entanglement, one specific case will be considered, and the possible experimental results will be examined (**Fig. 1**). There are many possible sources that can create many different sorts of entangled states. The source under consideration will be assumed to be the one used in the first teleportation experiments, which produced photons in a singlet polarization state. This means that neither photon enjoys a state of well-defined polarization; each one of the photons on its own is maximally unpolarized. Yet, when one of the two photons is subject to a polarization measurement, it assumes one specific polarization. That specific experimental outcome is completely random. As a consequence of the two photons being in the entangled singlet state, the other photon is instantly projected into a state orthogonal to that of the first photon. The fact that the measurement result on the second photon can be perfectly predicted on the basis of the measurement result of the first photon, even as neither one carries a well-defined quantum state, is known as the Einstein-Podolsky-Rosen paradox. In 1964 John Bell showed that these perfect correlations cannot be understood on the basis of properties that the entangled photons carry individually before the measurement. The resulting conflict between the philosophical position of local realism and the predictions of quantum mechanics, which have been confirmed beyond reasonable doubt in experiment, is known as Bell's theorem. *See* HIDDEN VARIABLES; PHOTON; POLARIZATION OF WAVES.

From an information-theoretic point of view, the interesting feature of entanglement is that neither of the two photons carries any information on its own. All information is stored in joint properties.

Concept of quantum teleportation. It was first realized by Charles H. Bennett and his colleagues that entanglement can be utilized to make teleportation possible (**Fig. 2**). Alice, who is in possession of the

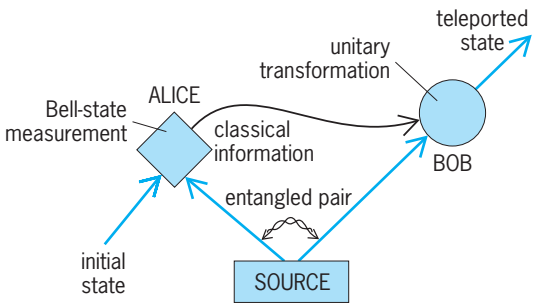

Fig. 2. Quantum teleportation procedure. Alice has an original particle in the initial state, and Alice and Bob also share an ancillary entangled pair. Alice then performs a Bell-state measurement and transmits the random result of that measurement in the form of two classical bits to Bob who, by a simple unitary transformation, can turn his ancillary photon into an exact replica of the original. (*After D. Bouwmeester et al., Experimental quantum teleportation, Nature, 390:575-579, 1997*)

original teleportee photon in a quantum state not known to her, and Bob initially share an ancillary pair of entangled photons, say in the singlet state described above. Alice then subjects her teleportee photon and her member of the ancillary pair to a Bell-state measurement. A Bell-state measurement is designed in such a way that it projects the two photons into an entangled state even if they were previously unentangled. This is a very tricky procedure both conceptually and experimentally. Conceptually, it means that the measurement must be performed in such a way that it is not possible, even in principle, to determine from the measurement result which photon was the teleportee and which was Alice's ancillary. They both have to lose their individuality. The result of the measurement must reveal only how the two photons relate to each other, and ignore individual properties. A Bell measurement has four possible results if the objects considered are defined in a two-dimensional Hilbert space just as is done to describe the photon's polarization. One of the four states is the singlet state discussed above. The other three states also define specific relations between the two photons, though different ones than those for the singlet state.

By the Bell-state measurement, Alice now knows how the unknown state of the teleportee photon relates to her ancillary one. She also knows in which entangled state the two ancillaries were produced, that is, how these two relate to each other. Thus she finally knows precisely how the teleportee relates to Bob's photon. More formally speaking, as a result of Alice's measurement Bob's photon is projected into a state which is uniquely related to the original state; the specific relationship is expressed by which of the four Bell states Alice obtained. Alice therefore informs Bob of her measurement result via a classical communication channel, and he, by applying a simple unitary transformation on his photon, changes it into the original state.

In one of the four cases, Alice obtains the information that her two photons have been projected into the singlet state, the same state in which the ancillaries were produced. Then, she knows that Bob's photon is instantly projected into the original state; the transformation that Bob has to apply is an identity transformation, that is, one that makes no change to his photon. That Bob's photon then instantly becomes an exact replica of the original seems to violate relativity. Yet, while Alice knows instantly that Bob's photon, no matter how far away, is already an exact replica, she has to inform Bob of the Bell measurement result such that he knows that his photon is already in the correct state. That classical information can arrive only at the speed of light. This requirement is also true for the other possible Bell-state measurement results. Bob has to know them in order to apply the correct transformation to his photon.

The result of the Bell measurement is not related at all to any properties that the original photon carries. Thus, that measurement does not reveal any information about its state. Therefore, the operation that Bob has to apply is also completely independent of any properties of the original photon. The reason that quantum measurement succeeds is that entanglement makes it possible to completely transfer the information that an object carries without measuring this information.

Experimental realization. An experiment therefore faces a number of challenges. They include (1) how to produce the entangled photon pairs and (2) how to perform a Bell measurement for independent photons. In the experimental realization by D. Bouwmeester and his colleagues in 1997, the entangled photons were produced in the process of spontaneous parametric downconversion. This is a second-order nonlinear process where a suitable crystal, in the experiment beta barium borate (BBO), is pumped with a beam of ultraviolet radiation. A photon from that beam has a very small probability to decay spontaneously into two photons, which then are polarization-entangled in just the way necessary for the experiment. The more tricky part is the Bell-state measurement because, in essence, it requires that the two photons are registered such that all information about which was the teleportee photon and which the ancillary is irrevocably erased. This is a nontrivial requirement since the two photons are coming from different directions, they might arrive at different times, and so forth. *See* NONLINEAR OPTICS; OPTICAL MATERIALS.

In the experiment, the Bell-state measurement was performed using a semireflecting mirror, which acted as a 50/50 beam splitter. Two photons were incident on the beam splitter, one from its front side and one from its back, and each one had the same probability of 50% to be either reflected or transmitted. If each of the two detectors in the two outgoing beams, again one in the front and one in the back, registered a photon simultaneously, then no information existed as to which incoming photon was registered in which detector, and the two were projected into the entangled singlet state. Narrow-bandwidth filters in front of the detectors further served to erase any time information which could also serve to identify the photons.

In this experiment, only one of the four possible Bell states could be identified, the singlet state. This certainly reduced the efficiency of the procedure, though in those cases in which the two detectors at the Bell-state analyzer registered, teleportation worked with a fidelity escaping all possible classical explanation.

In another experiment, also called entanglement swapping, it was even possible to teleport a photon that was still entangled to another one. That experiment started with two entangled pairs. A joint Bell-state measurement on one photon from each pair projected the other two photons onto an entangled state. In that way, two photons that neither came from the same source nor ever interacted with one another became entangled.

What all these experiments reveal is that the quantum state is really just a representation of the information that has been acquired. In the case of entanglement, it is only information on how objects relate to each other without any information on their individual properties. And in the case of teleportation, Alice's observation changes the quantum state that Bob observes. In other words, what can be said about the situation changes due to an observation by Alice. This gives very strong support to the Copenhagen interpretation of quantum mechanics. The first experiments were done with polarization-entangled photon pairs. Since then a number of experiments teleporting other properties such as continuous variables carried by the electromagnetic field of light, instead of the discrete ones discussed above, have been performed.

Prospects. While the teleportation distance in the first experiments was of the order of 1 m (3 ft), experiments in 2004 extended the distance to the order of 600 m (2000 ft), and there are plans to perform such experiments over much larger distances and even from a satellite down to laboratories on the ground. Other important experimental steps include the teleportation of quantum states of atoms (2004) and the teleportation of the quantum state of a photon onto that of an atomic cloud (2006).

Today quantum teleportation and entanglement swapping—the teleportation of an entangled state—are considered to be key building blocks of future quantum computer networks. At present there is intense research in the development of both quantum communication networks and quantum computers. Future quantum computers would use individual quantum states, for example those of atoms, to represent information in so-called quantum bits. They are expected to allow some algorithms to be performed with significantly higher speed than any existing computers. Quantum teleportation would allow the transfer of the quantum output of one quantum computer to the quantum input of another quantum computer. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM COMPUTATION; QUANTUM MECHANICS; QUANTUM THEORY OF MEASUREMENT.

Anton Zeilinger

Bibliography. M. D. Barrett et al., Deterministic quantum teleportation of atomic qubits, *Nature*, 429:737–739, 2004; C. H. Bennett et al., Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, *Phys. Rev. Lett.*, 70:1895–1899, 1993; D. Bouwmeester et al., Experimental quantum teleportation, *Nature*, 390:575–579, 1997; A. Furasawa et al., Unconditional quantum teleportation, *Science*, 282:706–709, 1998; M. Riebe et al., Deterministic quantum teleportation with atoms, *Nature*, 429:734–737, 2004; J. F. Sherson et al., Quantum teleportation between light and matter, *Nature*, 443:557–560, 2006.

Quantum theory of matter

The microscopic explanation of the properties of condensed matter, that is, solids and liquids, based on the fundamental laws of quantum mechanics. Without the quantum theory, some properties of matter, such as magnetism and superconductivity, have no explanation at all, while for others only a phenomenological description can be obtained. With the theory, it is at least possible to comprehend what is needed to approach a complete understanding.

The theoretical problem of condensed matter, that is, large aggregates of elementary particles with mutual interactions, is that of quantum-statistical mechanics, or the quantum-mechanical many-body problem: an enormous number, of order 10^{23}, of constituent particles in the presence of a heat bath and interacting with each other according to quantum-mechanical laws. What makes the quantum physics of matter different from the traditional quantum theory of elementary particles is that the fundamental constituents (electrons and ions) and their interactions (Coulomb interactions) are known but the solutions of the appropriate quantum-mechanical equations are not. This situation is not due to the lack of a sufficiently large computer, but is caused by the fact that totally new structures, such as crystals, magnets, ferroelectrics, superconductors, liquid crystals, and glasses, appear out of the complexity of the interactions among the many constituents. The consequence is that entirely new conceptual approaches are required to construct predictive theories of matter. *See* CRYSTAL STRUCTURE; FERROELECTRICS; GLASS; LIQUID CRYSTALS; SUPERCONDUCTIVITY.

The usual technique for approaching the quantum many-body problem for a condensed-matter system is to try to reduce the huge number of variables (degrees of freedom) to a number which is more manageable but still can describe the essential physics of the phenomena being studied. In general, the fundamental laws of quantum mechanics give little or no guidance for this task. For example, while it seems natural to describe a solid by a collection of ions and valence electrons, it is not apparent from the elementary equations of motion that this makes sense. Here, the ions consist of the nuclei and the more tightly bound electrons; the valence electrons are the more loosely bound ones which participate in the chemical bonding. The problem remains extraordinarily

complex because of the mutual Coulomb interactions among all these particles.

One-electron approach. In the several decades after the development of quantum theory and the understanding of the Pauli exclusion principle, enormous successes were achieved in developing the band theory of solids. The basic idea was to simplify the effects of all the interparticle interactions by approximating them by average potential fields. In this so-called one-electron approximation, in a crystalline lattice for example, the valence electrons move effectively independently of each other, under the influence of the average potential field arising from their mutual interactions and of the periodic electrostatic potential due to the ions which have arranged themselves in a periodic array. The quantum mechanics of this system then gives conclusions which can explain many observed properties of a large number of materials: the differences between a metal, a semiconductor, and an insulator; optical and thermodynamic properties; heat and electrical conductivity; thermoelectric effects; and some of the effects of impressed magnetic fields. None of these issues found satisfactory explanations before the quantum-mechanical band theory of solids. *See* BAND THEORY OF SOLIDS; ELECTRICAL CONDUCTIVITY OF METALS; FREE-ELECTRON THEORY OF METALS; SEMICONDUCTOR.

Even the motion of the ions must be described quantum-mechanically. In a very good approximation, the deviations of the ionic positions from their regular lattice points can be described in terms of a collection of masses whose near neighbors are coupled by springs. The quantum-mechanical description of the normal modes of these springs is in terms of elementary excitations called phonons whose thermodynamics gives the correct result for the contribution of ionic motion to the specific heat, which was previously unobtainable. *See* LATTICE VIBRATIONS; PHONON; SPECIFIC HEAT OF SOLIDS.

While the band theory, based on the one-electron approximation described above, has been dramatically successful as a detailed microscopic explanation of a variety of properties of matter, it has failed for a number of others. It was realized that a more detailed consideration of the interelectron Coulomb interactions is essential to understand some of the deeply important properties. For example, the band theory cannot distinguish between a metal and a type of magnetic insulator, nor can it lead to a theory of superconductivity.

Many-body approach. The quantum theory of metals has been extended beyond the one-electron band theory to include some effects of the interelectron Coulomb interactions. It seems at first sight that this would be impossible since the interaction energies are not small compared to the characteristic one-electron energy. However, the Pauli principle and the phenomenon of screening reduce the effectiveness of the interaction, and in many cases this allows a systematic treatment. In an elegant formulation of the many-body problem, the many interactions are reduced to a few simple parameters; the result is called a Fermi liquid. The Fermi liquid describes the behavior of the metallic electrons and certain collective excitations which arise as a consequence of the interactions. The same theory accounts for many of the properties of liquid helium-3. *See* LIQUID HELIUM.

Magnetism. One of the most dramatic of the phenomena which elude band theory is magnetism. The pioneers of the quantum theory lost no time in attacking the issue of magnetism, first on the level of atoms. Thus, P. A. M. Dirac's theory of the electron enabled an understanding of the magnetic properties of atoms in terms of the electron's orbital and intrinsic (spin) magnetic moments and the quantization of atomic energy levels. However, the character of the ground state of a given atom depends on the interplay between the Pauli exclusion principle and the Coulomb repulsion between the electrons. This was first accounted for qualitatively by Hund's rules, but the microscopic description requires a detailed treatment of the electron-electron interactions which goes beyond the Bohr one-electron theory for atoms. *See* ATOMIC STRUCTURE AND SPECTRA; ELECTRON SPIN; EXCLUSION PRINCIPLE.

It was quickly recognized by W. Heisenberg and Dirac that the interplay of the Pauli exclusion principle and the Coulomb interactions makes the behavior of a small group of electrons dependent on the relative orientations of their spins. This is called the exchange effect, and is expressed by the Heisenberg-Dirac exchange interaction. The latter is an example of the reduction of many degrees of freedom to a few: The complicated description of the behavior of all the electrons on, for example, two atoms is replaced by a simple expression involving only two variables, the scalar product of the net spin vectors on each atom. They are coupled with a strength which, in principle, is determined by the motions and interactions of all the underlying electrons but in practice is often taken as a phenomenological parameter. In solids, it is conceivable that the coupling parameter has either sign; that is, either spins parallel (ferromagnetic coupling) or spins antiparallel (antiferromagnetic coupling) could be the lowest energy configuration. This fundamental result, when applied to collections of atoms, is the basis for many theories of the magnetism in insulators. The foundation of these theories is the Heisenberg-Dirac exchange interaction with a phenomenologically determined coupling; the microscopic determination of the sign and magnitude of this coupling remains difficult. *See* ANTIFERROMAGNETISM; EXCHANGE INTERACTION; FERROMAGNETISM.

Modern many-body problem. Magnetism in metals such as iron and chromium poses even more complicated questions because, in contrast to insulators where the valence electrons are essentially localized, there are itinerant conduction electrons whose orbital motion is important for the magnetism. In such cases, the electron-electron interaction leads to correlations which play a central role in determining magnetic properties but for which there is no complete theoretical analysis. The same situation

occurs even more dramatically in those magnetic insulators which within the one-electron band theory are predicted to be metals. Striking examples are found among the transition-metal oxides, for example, and the insulating magnetic parent compounds of the family of copper oxide high-temperature superconductors. In these cases, the electron-electron repulsions are so strong that not only is the one-electron approach almost irrelevant but the systematics of the Fermi-liquid theory are no longer valid. This state of affairs also occurs frequently in systems of low dimensionality, which often have behavior dominated by strong interactions. These include the family of quasi-one-dimensional conductors formed by arrays of long chains of molecules, the planar copper oxide high-temperature superconductors, and the state of matter observed in two-dimensional electron systems in a magnetic field when in the regime of the fractional quantum Hall effect. *See* HALL EFFECT; ORGANIC CONDUCTOR.

This situation defines the modern many-body problem in the quantum theory of matter. The properties of many materials are, in a fundamental way, primarily determined by the presence of strong interparticle interactions, and these are also responsible for the presence of magnetism in those materials which exhibit it. The modern challenge for the quantum theory of matter is to give a thorough analysis of the interparticle correlations which arise from the strong interactions and to describe how these correlations lead to the many effects which are observed in a wide class of materials. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; SOLID-STATE PHYSICS.

Elihu Abrahams

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; P. M. Chaikin and T. C. Lubersky, *Principles of Condensed Matter Physics*, 1994; A. Ishihara, *Condensed Matter Physics*, 1991; G. D. Mahan, *Many-Particle Physics*, 3d ed., 2000; R. M. White and T. H. Geballe, *Long Range Order in Solids*, 1979.

Quantum theory of measurement

The attempt to reconcile the counterintuitive features of quantum mechanics with the hypothesis that it is in principle a complete description of the physical world, even at the level of everyday objects. A paradox arises because, at the atomic level where the quantum formalism has been directly tested, the most natural interpretation implies that where two or more different outcomes are possible it is not necessarily true that one or the other is actually realized, whereas at the everyday level such a state of affairs seems to conflict with direct experience.

Interference. It is useful to consider a group (ensemble) of identically prepared microscopic objects such as electrons or photons which start, one at a time, from some state A, and from there can proceed to one of two alternative states, B or C, and from there to E or some other state. (The various states might correspond to spatially separated positions, as in Young's double-slit experiment, or to something more abstract, such as the different strangeness states of the neutral K meson.) Quantum mechanics describes the process of transition of the particles between the various states in terms of a set of probability amplitudes (wave functions) such that the probability of a given process occurring is proportional to the squared modulus of the corresponding amplitude. When a given result can be achieved by more than one process, and no measurement has been made to decide which occurred, the total amplitude for the result to occur is the algebraic sum of the amplitudes for the various contributing processes.

Thus, in the setup described, the probability (P_B) that a particle arrives at E via B can be determined; this is done by blocking alternative C (for example, in the Young's double-slit experiment, by simply obstructing the slit C). This probability is just the squared modulus ($|\psi_B|^2$) of the amplitude ψ_B for the process A→B→E. Similarly, by blocking alternative B, $P_C = |\psi_C|^2$ can be determined.

If both alternatives are now left open, the total amplitude ($\psi_{B\text{ or }C}$) to arrive at E is now the algebraic sum of ψ_B and ψ_C, and the probability of arriving at E is given by the expression in Eq. (1). The first two

$$\begin{aligned} P_{B\,\mathrm{or}\,C} &= |\psi_{B\,\mathrm{or}\,C}|^2 = |\psi_B + \psi_C|^2 \\ &= |\psi_B|^2 + |\psi_C|^2 + 2\,\mathrm{Re}\,(\psi_B^*\psi_C) \qquad (1) \\ &= P_B + P_C + 2\,\mathrm{Re}\,(\psi_B^*\psi_C) \end{aligned}$$

terms in this expression (which is very well verified at the atomic level) correspond to what would be obtained by assuming that each individual particle of the ensemble proceeded either via B or via C. However, the last term cannot be interpreted in this way; rather, it shows the effects of interference between the two possibilities, a characteristically quantum-mechanical phenomenon. In particular, if ψ_C happens to be equal to ψ_B in magnitude and opposite in sign, then P_B or P_C vanishes even though P_B and P_C are both individually nonzero. *See* COHERENCE; INTERFERENCE OF WAVES.

Thus, the properties of the complete ensemble of particles reaching E are not the same as the properties of a combination of the subensembles which are obtained by allowing only alternative B and only alternative C respectively. This strongly suggests that in this situation it is not the case that each individual particle of the ensemble realizes one or other of the two alternatives. Rather, each particle, as it were, keeps its options open. Although strenuous attempts have been made to avoid this conclusion, they all encounter severe difficulties. However, if a measuring device is set up to determine whether one or both paths are followed, each individual particle appears to choose either alternative B or alternative C, at random with a probability $|\psi_B|^2$ and $|\psi_C|^2$ respectively. But under these conditions no interference at E occurs. Thus, it seems that when no measurement is performed, a definite alternative is not realized, but

when a measurement is performed, it is. *See* HIDDEN VARIABLES.

Quantum measurement paradox. In the early days of quantum mechanics, the resolution of the apparent paradox was often located in the act of measurement itself, which was regarded as involving a supposedly unanalyzable interaction of the microscopic system with a macroscopic instrument which had to be described in classical terms. However, if one believes, as most physicists do, that quantum mechanics gives in principle a complete description of the physical world, it must be possible to describe the process of measurement itself, as well as the working of the measuring instrument, in quantum-mechanical language.

A collection of macroscopic counters can be set up to determine which of the paths, B or C, is followed by each individual particle of the ensemble. It may be supposed temporarily (and unrealistically, as discussed below) that whenever the particle in question is drawn from a subensemble described by ψ_B (alone), the final state of the counters, which registers the fact that path B was indeed followed, is some pure quantum state Ψ_B; similarly, when the particle is drawn from a subensemble described by ψ_C, and is therefore guaranteed to follow path C, the final state of the counters is a pure state Ψ_C. Since it is necessary to read off which path was followed directly from an inspection of the counters, the states described by Ψ_B and Ψ_C must be macroscopically distinguishable.

If the particle in question is now drawn from an ensemble described by the so-called linear-superposition wave function ψ_B or $\psi_C \equiv \psi_B + \psi_C$, then because quantum mechanics is by its nature a linear theory, it immediately follows that in this case the final state of the counters is given by Eq. (2).

$$\Psi_B \text{ or } \Psi_C \equiv \Psi_B + \Psi_C \tag{2}$$

That is, it is a linear superposition of macroscopically distinct states. If now it is supposed that no measurement has been made on the measuring instruments themselves, and the same interpretation of the quantum formalism is adopted as seemed to be required at the atomic level, it would seem to follow that no definite outcome of the experiment has been realized even at the everyday level of counters, and so forth—a conclusion which apparently conflicts violently with direct experience at this level. This is called the quantum measurement paradox.

Alleged resolution. The resolution of this paradox that is probably most favored by practicing physicists proceeds in two stages. At stage 1, it is pointed out that the above argument (and in particular the description of the final state of the counters by pure states Ψ_B and Ψ_C) is grossly oversimplified, and that a correct quantum-mechanical analysis shows that, quite generically, whenever the formalism appears to generate a superposition of macroscopically distinct states it is impossible to demonstrate the effects of interference between them. The reasons for this claim include the facts that the initial state of a macroscopic system is likely to be unknown in detail; its extreme sensitivity to random external noise; and most important merely by virtue of its macroscopic nature any such system will rapidly have its quantum-mechanical state correlated (entangled) with that of its environment in such a way that no measurement on the system alone (without a simultaneous measurement of the complete state of the environment) can demonstrate any interference between the two states in question—a result often known as decoherence. Thus, it is argued, the outcome of any possible experiment on the ensemble of macroscopic systems prepared in this way will be indistinguishable from that expected if each system had actually realized one or other of the two macroscopically distinct states in question. Stage 2 of the argument (often not stated explicitly) is to conclude that if this is indeed true, then it may be legitimately asserted that such realization of a definite macroscopic outcome has indeed taken place by this stage.

Most physicists agree with stage 1 of the argument. However, not all agree that the radical reinterpretation of the meaning of the quantum formalism which is implicit at stage 2 is legitimate; that is, that an interpretation in terms of realization, by each individual system, of one alternative or the other, forbidden at the atomic level by the observed phenomenon of interference, is allowed once, on going to the macroscopic level, the phenomenon disappears.

Alternative approaches. Consequently, various alternative interpretations have been developed. One, sometimes known as the statistical interpretation, holds that the symbols of the quantum-mechanical formalism correspond to nothing in the physical world at all, even at the macroscopic level, but are merely a recipe for calculating the probabilities of various directly observable outcomes of the experiment in question. In another direction, the Everett-Wheeler (or relative-state or many-worlds) interpretation holds (at least according to some expositions) that the subjective impression of having observed one and only one outcome of a given experiment is an illusion and that there exist other parallel universes, said to be equally real, in which the unknown outcomes are realized. Several other interpretations have also been proposed.

A quite different approach to the quantum measurement paradox is the conjecture that quantum mechanics is not, after all, a complete description of the physical world but must be supplemented by other laws of physics. Since the predictions of quantum mechanics appear to be very well verified at the atomic level, the corrections due to such new laws must be zero or very small in this regime; however, they must be effective enough at the macroscopic level to effect the realization of a definitive alternative in, at most, a time that is small compared to the minimum interval detectable by human observation. A concrete scenario along these lines was developed by G. C. Ghirardi, A. Rimini, T. Weber, and P. Pearle (the GRWP theory).

Experimental test. The existence, or not, of quantum superpositions of macroscopically distinct states

may be susceptible, to an extent, to experimental resolution. Obviously, to establish the existence of such superpositions it is necessary to observe the effects of interference between the two components (and to show that such effects could not follow from any theory in which realization had taken place by this stage), and thus a system is required where decoherence does not take place on the relevant time scale. Candidates for such systems include magnetic grains, various quantum-optical systems, and superconducting devices incorporating the Josephson effect. *See* JOSEPHSON EFFECT; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; SQUEEZED QUANTUM STATES.

Anthony J. Leggett

Bibliography. D. Z. Albert, *Quantum Mechanics and Experience*, 1992; L. E. Ballentine, The statistical interpretation of quantum mechanics, *Rev. Mod. Phys.*, 42:358–381, 1980; B. d'Espagnat, *Conceptual Foundations of Quantum Mechanics*, 2d ed., 1976, reprint 1989; A. I. M. Rae, *Quantum Physics: Illusion or Reality?*, 1987, paper 1994.

Quark-gluon plasma

A predicted state of nuclear matter containing deconfined quarks and gluons. According to the theory of strong interactions, called quantum chromodynamics, hadrons such as mesons and nucleons (the generic name for protons and neutrons) are bound states of more fundamental objects called quarks. The quarks are confined within the individual hadrons by the exchange of particles called gluons. However, calculations indicate that, at sufficiently high temperatures or densities, hadronic matter should evolve into a new phase of matter containing deconfined quarks and gluons, called a quark-gluon plasma when hot, or quark matter when cold and dense. Such a state of matter is thought to have existed briefly in the very early universe until about 10 microseconds after the big bang, and its cold form might exist today inside the dense cores of neutron stars. One important prediction is that the mechanism, called spontaneous chiral symmetry breaking, that is thought to be responsible for 98% of the mass of the nucleon under normal conditions is not operative in the quark-gluon plasma. *See* BIG BANG THEORY; HADRON; NEUTRON STAR; QUANTUM CHROMODYNAMICS.

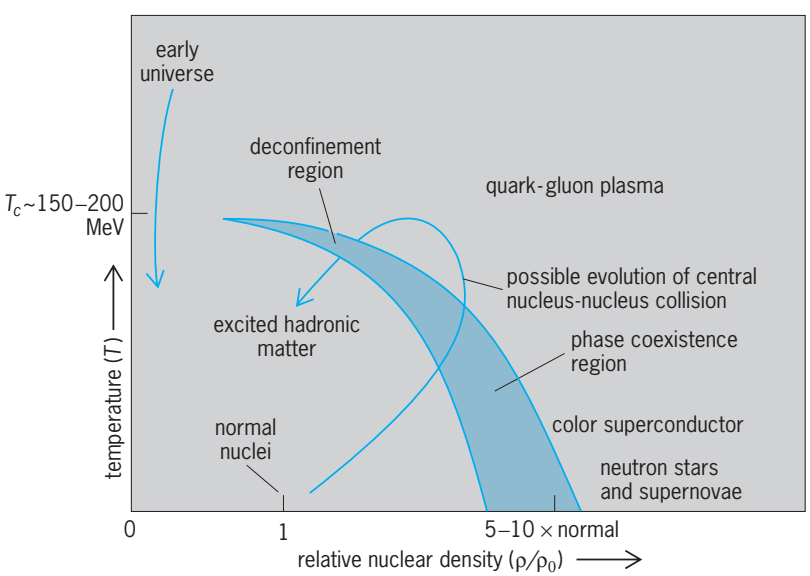

Fig. 1. Conjectured phase diagram of nuclear matter, temperature T versus relative nuclear density ρ/ρ_0, where the normal nuclear density $\rho_0 = 0.16$ nucleon/fm^3. Various regions of matter are shown with normal nuclei being at $T = 0$ and $\rho/\rho_0 = 1$. A possible path of the early universe and a possible trajectory followed by matter produced in a nucleus-nucleus collision are shown.

Thermodynamics of plasma. Since the quark-gluon plasma is made up of strongly interacting quarks and gluons at thermal equilibrium, an appropriate theoretical framework for describing its properties is statistical quantum chromodynamics. In this framework, calculations of quantities such as the energy density, entropy, pressure, and the equation of state of the matter, as a function of the temperature and chemical potential of the system, are carried out. (Temperature refers to the average energy of a particle in the system; the chemical potential denotes the energy required to add a nucleon to the system.) Present calculations, using computer simulations of statistical quantum chromodynamics on a discrete spatial grid (lattice gauge theory), estimate that the critical temperature T_c needed to heat matter to the point that the quarks and gluons become deconfined and form a quark-gluon plasma is about 150–200 MeV (**Fig. 1**), where 1 MeV is equivalent to about 10^{10} K. Such calculations assume that thermal and chemical equilibrium is achieved, and they involve extrapolations of volume and quark masses to realistic values. These assumptions need to be verified experimentally in attempts made to create a quark-gluon plasma in laboratory conditions, for example, in high-energy nuclear collisions. *See* STATISTICAL MECHANICS.

Dense but cold quark matter is predicted to exhibit the property of color superconductivity. This phenomenon occurs when quarks near the Fermi surface bind in pairs that act as bosons and form a Bose condensate, similar to the Cooper pairs of electrons in an electronic superconductor. Depending on the quantum numbers of the pairing quarks and the spatial structure of the Cooper-pair condensate, several different phases of color-superconducting quark matter can be distinguished. Which one of these phases would form in the core of a neutron star when the nuclear density is sufficiently high is still under investigation. *See* SUPERCONDUCTIVITY.

The nature of the phase transition, whether first-order or higher, is an important issue. First-order transitions (such as the transformation of water to ice or steam) generally involve abrupt changes in experimental observables, such as the latent heat, whereas higher-order transitions are continuous. Present calculations indicate that the character of the hadronic matter to quark-gluon plasma transition is very sensitive to the masses of the quarks involved. For realistic masses of the up (u) and down (d) quarks, the constituents of the nucleon, and the strange (s) quark, the transition under the conditions of the early universe (high temperature and nearly zero nuclear

density) appears to be very rapid, but not discontinuous. For cold but dense nuclear matter, the transition is conjectured to be of first order. The order of the transition may affect the ability of experiments to detect the plasma. *See* PHASE TRANSITIONS.

Production. The discovery and study of a new state of matter requires a means for producing it under controlled laboratory conditions. The extreme conditions of heat and compression required for the transition from the hadronic matter to the quark-gluon plasma phase can be achieved on Earth only by collisions of beams of heavy ions such as nuclei of gold or lead (although lighter nuclei can be used) with other heavy nuclei at extremely high energies (Fig. 1). To be most effective, these collisions must be as central as possible to bring to bear the full energy of the incident nuclei. Calculations using the lattice gauge model of quantum chromodynamics indicate that energy densities of at least 1–2 GeV/fm^3 (1 fm = 1 femtometer = 10^{-15} m, 1 GeV = 1000 MeV), about 10 times that found in ordinary nuclear matter, must be produced in the collision for plasma formation to occur (Fig. 1). *See* NUCLEAR REACTION; RELATIVISTIC HEAVY-ION COLLISIONS.

Experimental signatures and results. At the earliest stage of a central collision, multiple nucleon-nucleon collisions occur as the two nuclei interpenetrate, leading to a sudden formation of a highly excited region densely filled with gluons and quark-antiquark pairs. The interaction among these particles causes their fast equilibration and leads to a rapid increase of temperature and pressure (Fig. 1). On the assumption that sufficient energy density is deposited, the transition to a quark-gluon plasma may be initiated over a volume comparable to the overlap region of the colliding nuclei, which have radii of several femtometers, and on a time scale of the order of 10^{-23} s. This hot compressed system then starts to expand under the action of its internal pressure and cools. As the temperature of the expanding matter drops below the critical value T_c needed to deconfine the quarks and gluons, they recombine to form hadrons and eventually restore the pure hadronic matter phase. Thus, no free quarks or gluons are expected at the end of the reaction.

Although no single unambiguous signature for the quark-gluon plasma has been identified, several characteristic probes or observables have been suggested. Signatures are sought coming from the primary hot, dense plasma phase or from transition back into the hadronic phase. Photons and lepton pairs (e^+e^- or $\mu^+\mu^-$) are considered to be among the best probes of the quark-gluon phase owing to their relatively weak interaction with matter, which allows them to leave the production volume without modification of the information they carry about their primary production. *See* LEPTON.

Other sources of information about the plasma phase are hadronic fragments of very energetic quarks or gluons, which are scattered sideways during the original impact between the colliding nuclei. These quarks interact strongly with the medium and thereby can lose a significant part of their energy. The amount of energy lost is a measure of the transparency of the medium to quarks and gluons, which is expected to be low. (This property is analogous to the optical opacity of electromagnetic plasmas.) *See* PLASMA (PHYSICS).

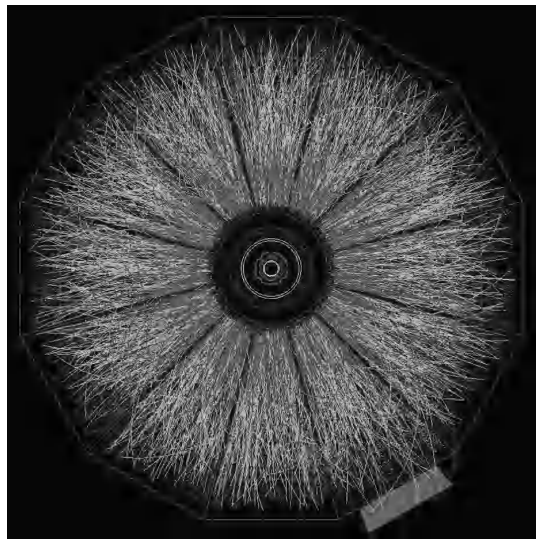

Fig. 2. Time projection chamber picture of a central gold-gold collision in which several thousand charged particles were produced. Each incident gold nucleus had a kinetic energy of 100 GeV per nucleon. The visible tracks of the particles can be analyzed to obtain information about the properties of the state from which the particles are emitted. This experiment was performed by the STAR collaboration at Brookhaven National Laboratory with the Relativistic Heavy Ion Collider (RHIC). (*Brookhaven National Laboratory*)

The momentum spectrum of emitted hadrons is influenced by the collective flow pattern of the matter before its breakup into freely streaming particles. This flow is generated by the internal pressure of the expanding matter and reflects its equation of state as well as the initial conditions prevailing at the moment of formation of an equilibrated phase. The angular anisotropy of the flow pattern in not quite central collisions is predicted to be especially sensitive to the fluid properties of the matter.

Other probes of the properties of the produced matter involve heavier species of quarks, such as the strange (s) quarks and charm (c) or bottom (b) quarks. The reduction of the s quark mass and the ample availability of gluons in the quark-gluon plasma phase allow enough strange quarks to be produced to establish chemical equilibrium. As the system cools, the strange quarks later recombine with other, lighter (u and d) quarks to form strange hadrons, which can thus serve as vestiges of quark and gluon deconfinement during the high-density stage. Bound states of c or b quarks are predicted to dissolve at even higher temperatures due to the screening of the strong interaction in the plasma phase. A large suppression of mesons made exclusively of such heavy quarks would thus serve as evidence for the screening of forces that is characteristic of the plasma phase. *See* CHARM; MESON; STRANGE PARTICLES.

A systematic search for the quark-gluon plasma has been conducted at the Relativistic Heavy Ion Collider (RHIC) since 2000. Experiments at this accelerator facility colliding beams of gold nuclei with energies of up to 100 GeV per nucleon (**Fig. 2**) have found compelling evidence for the formation of a novel state of matter which exhibits many of the predicted properties of the quark-gluon plasma. The final state is found to be compatible with the assumption that particles are emitted from a thermally and chemically equilibrated source. The yield of energetic hadrons is strongly suppressed, indicating a high opacity of the medium to quarks and gluons. The intensity of the measured collective flow pattern among the emitted hadrons indicates that the produced matter is a nearly ideal fluid with an extremely low viscosity near the theoretical limit. Differences observed in the flow patterns of baryons and mesons suggest that the flow was generated in a phase in which quarks were not confined. Detailed studies of weakly interacting or rarely produced probes, such as lepton pairs and hadrons containing *c* quarks, are in progress. Experiments with nuclei of even higher energy are in preparation at the Large Hadron Collider at the European Center for Particle Physics (CERN). These experiments should provide critical tests of the theory of the strong interaction and illuminate the earliest moments of the universe. *See* ELEMENTARY PARTICLE; GLUONS; PARTICLE ACCELERATOR; QUARKS.

Berndt Müller

Bibliography. M. G. Alford, Color superconducting quark matter, *Annu. Rev. Nucl. Part. Sci.*, 51:131–160, 2001; J. H. Harris and B. Müller, The quest for the quark-gluon plasma, *Annu. Rev. Nucl. Part. Sci.*, 46:71–107, 1996; T. Ludlam and L. McLerran, What have we learned from the relativistic heavy ion collider?, *Phys. Today*, 56(10):48–54, October 2003; B. Müller and J. L. Nagle, Results from the Relativistic Heavy Ion Collider, *Annu. Rev. Nucl. Part. Sci.*, 56:93–135, 2006; K. Yagi, T. Hatsuda, and Y. Miake, *Quark-Gluon Plasma: From Big Bang to Little Bang*, Cambridge University Press, 2005.

Quarks

The basic constituent particles of which elementary particles are understood to be composed. Theoretical models built on the quark concept have been very successful in understanding and predicting many phenomena in the physics of elementary particles.

Search for fundamental constituents. Physics research over the past two centuries has been probing progressively deeper into the structure of matter in order to seek, at every stage, the constituents of each previously "fundamental" entity. Thus, a hierarchy proceeding from crystals of bulk matter through molecules, atoms, and nuclei to nucleons (protons and neutrons) and mesons has been revealed. The energies required to dissociate each entity increase from thermal energies to gigaelectronvolts in proceeding from crystals to mesons. It is therefore only a logical extrapolation of past patterns to expect that hadrons—mesons and baryons (nucleons)—might be dissociated into more fundamental constituents if subjected to a sufficiently high energy. *See* ATOMIC PHYSICS; HADRON; NUCLEAR PHYSICS; NUCLEON.

Evidence supporting the quark model. The study of the elastic scattering of electrons on protons demonstrated that the proton has a finite form factor, that is, a finite radial extent of its electric charge and magnetic moment distributions. It was plausible that the charge cloud which constitutes the proton is a probability distribution of some smaller, perhaps pointlike constituents, just as the charge cloud of an atom was learned to be the probability distribution of electrons. Subsequent high-energy, deep inelastic scattering experiments of electrons on protons, leading to meson production, revealed form factors corresponding to pointlike constituents of the proton. These proton constituents, first referred to as partons, are now understood to include the constituent quarks of the proton. The quark structure of the proton was also invoked to explain the nature of secondary π-meson (pion) distributions produced in high-energy proton-proton inelastic collisions.

These high-energy collisions also produced an abundance of resonance states, equivalent to short-lived particles. The spectroscopy of these hadronic states revealed an order and symmetry among the observed hadrons that could be interpreted in terms of representations of the SU(3) symmetry group. This in turn is interpreted as a consequence of the grouping of elementary constituents of fractional electric charge in pairs and triplets to form the observed particles. The general features of the quark model of hadrons have withstood the tests of time, and the static properties of hadrons are consistent with predictions of this model. *See* SYMMETRY LAWS (PHYSICS); UNITARY SYMMETRY.

Thus, the proton and neutron are not fundamental constituents of matter, but each is composed of three quarks, very much as the nuclei of ^{3}H and ^{3}He are made of protons and neutrons, and the molecules of NO_2 and N_2O are made of oxygen and nitrogen atoms.

Kinds of quarks. There are two kinds (or "flavors") of quarks of very low mass of which the proton, neutron, and pions are composed, and a third, more massive quark which is a constituent of "strange" particles, including the *K* mesons and hyperons such as the Λ^0. These are known as the up quark (u), the down quark (d), and the strange quark (s). Baryons are composed of three quarks, for example the proton (uud), neutron (udd), Λ^0 (uds), and Ξ^- (dss). Antiparticles such as antiprotons are formed by the antiquarks of those forming the particle, for example, the antiproton $\bar{p}(\bar{u}\,\bar{u}\bar{d})$. Mesons are composed of a quark-antiquark pair, such as the $\pi^+(u\bar{d})$, $\pi^-(\bar{u}d)$, $K^+(u\bar{s})$, and $K^-(\bar{u}s)$. *See* BARYON; HYPERON; MESON; STRANGE PARTICLES.

The quantum numbers of quarks are added to give the quantum numbers of the elementary particle which they form on combination. The unit of electrical charge of a quark is $+2/3$ or $-1/3$ of the charge

TABLE 1. Properties of quarks

Flavor	Mass[†], GeV/c^2	Electric charge[‡]	Baryon number	Spin[§]	Isotopic spin	Strangeness	Charm
u	0.0015–0.005	+2/3	+1/3	1/2	1/2	0	0
d	0.003–0.009	−1/3	+1/3	1/2	1/2	0	0
c	1.1–1.4	+2/3	+1/3	1/2	0	0	+1
s	0.060–0.170	−1/3	+1/3	1/2	0	−1	0
t	174.3±5.1[¶]	+2/3	+1/3	1/2	0	0	0
b	4.1–4.4	−1/3	+1/3	1/2	0	0	0

[†]As the mass of baryons composed of quarks is strongly influenced by the gluons binding the quarks, and as free quarks are not observed, the masses are theoretical estimates.
[‡]Charge is in units of the magnitude of the charge of an electron; 1.6×10^{-19} Coulomb.
[§]Spin is in units of Planck's constant divided by 2π, written as *h*.
[¶]The top quark mass is deduced from experimental measurements of its decay dynamics.

on a proton (1.6×10^{-19} coulomb), and the baryon number of each quark is +1/3 (**Table 1**). The charge, baryon number, and so forth, of each antiquark are just the negative of that for each quark.

The manner in which the *u*, *d*, and *s* quarks and their antiquarks may combine to form families of mesons may be indicated on two-dimensional plots for multiplets, such as the meson pseudoscalar nonet (**Fig. 1**). On such plots, hypercharge is a quantum number related to the quark strangeness and baryon number, and isospin (here the third component of isotopic spin) is a quantum number related to the *u*-*d* quark difference. *See* HYPERCHARGE; ISOBARIC PROCESS.

In the strong interactions of elementary particles, quark-antiquark pairs of the same flavor may be created if sufficient energy is present, but a quark does not transform into a different quark. The weak interaction, however, does permit quark transformations to occur, so that, for example, a *d* quark becomes a *u* quark in the radioactive decay of a neutron, and particles containing an *s* quark (strange particles) decay to nonstrange particles only through the weak interaction. *See* FUNDAMENTAL INTERACTIONS; STRONG NUCLEAR INTERACTIONS; WEAK NUCLEAR INTERACTIONS.

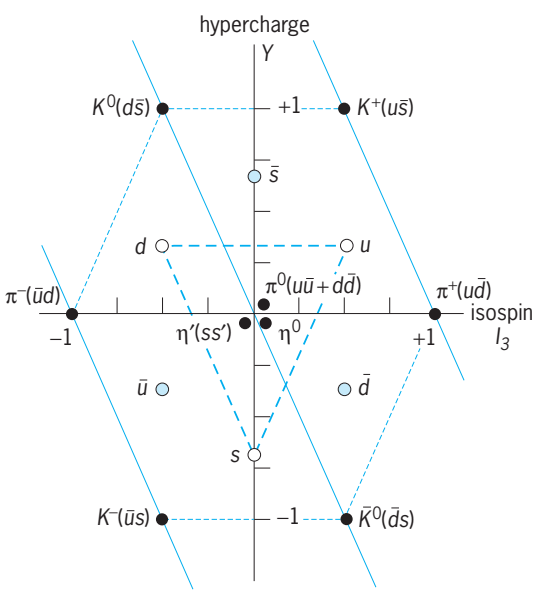

Fig. 1. Pseudoscalar meson nonet and the *u*, *d*, and *s* quarks and antiquarks as represented in isospin (I_3)–hypercharge space. The solid lines sloping downward to the right are lines of constant electric charge.

A particularly interesting set of particles is the group of vector mesons composed of a quark-antiquark pair. An example is the ϕ meson, composed of an $s\bar{s}$ pair. It may decay to a K^+K^- meson pair but is inhibited from decaying to a $\pi^+\pi^-$, as the latter contain no *s* or $\bar{s}$ quarks. The ϕ may also decay electromagnetically to a $\mu^+\mu^-$ or e^+e^- lepton pair; conversely, it may be produced by a photon or by the collision of an electron and positron of sufficient energy.

Such quark-antiquark systems are very analogous to the positronium atom, composed of a positron, e^+, and an electron, e^-. The force binding this atom together is the electromagnetic (Coulomb) interaction, whereas the force binding the quark-antiquark system is the strong interaction. In both cases, the energy levels of excited states of the system are directly related to the details of the force holding the particles together. As with the quark-antiquark particles, positronium decays in a small fraction of a second through the annihilation of its particle-antiparticle constituents. *See* POSITRONIUM.

Charm, bottom, and top quarks. During the 1970s, experiments at electron-positron colliders and proton accelerators detected a relatively long-lived (that is, very narrow, in energy) resonant state of about 3.1 GeV total energy. This was interpreted as evidence for a new quark, the charm (*c*) quark, produced as a $c\bar{c}$ quark-antiquark resonance analogous to the ϕ. The discovery of this J/ψ resonance was followed by the observation and study of meson systems, now labeled *D* mesons, containing a single *c* or $\bar{c}$ quark (paired with an antiquark of another flavor), as well as baryon states containing these quarks. The two-dimensional diagrams (Fig. 1) may be elaborated by adding a third dimension corresponding to the charm quantum number, defining a polyhedron whose vertices correspond to identified charm (*D*), strange, and other mesons (**Fig. 2**). *See* CHARM; J/PSI PARTICLE; PARTICLE ACCELERATOR.

A few years later, experiments with higher-energy proton beams, studying the spectra of muon-antimuon pairs at the Fermi National Accelerator Laboratory, discovered a more massive, narrow resonant

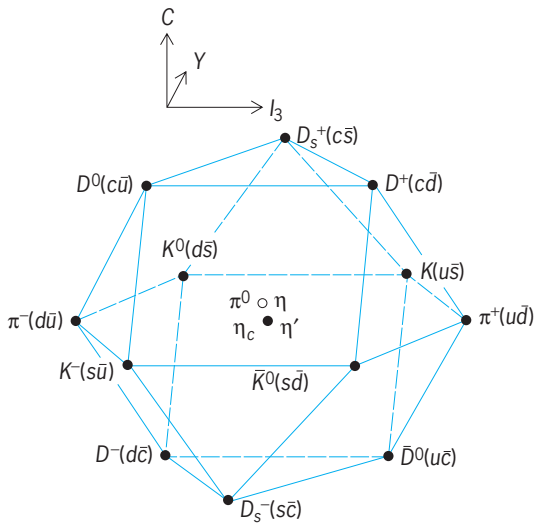

Fig. 2. Pseudoscalar mesons composed of *u*, *d*, *s*, and *c* quarks. The three orthogonal axes are the *z* component of isospin I_3, hypercharge *Y*, and charm *C*. (*After C. Quigg, Lectures on charmed particles, Fermi National Accelerator Laboratory, Fermilab-Conf-78/37-THY, April 1978*)

state at about 9.4 GeV, which was labeled the Υ (upsilon). This was interpreted as evidence for a more massive quark, the *b* (bottom) quark. Subsequent experiments at proton and electron accelerators confirmed the existence of the *b* quark and also observed a corresponding family of meson resonant states, now referred to as *B* mesons. A four-dimensional extension of the three-dimensional diagram of Fig. 2 can be imagined which would include *B* mesons containing the *b* quark, such as $b\bar{u}$, $b\bar{s}$, and $c\bar{b}$. *See* UPSILON PARTICLES.

During the 1990s, experiments observing collisions of protons and antiprotons at an energy of 1.8 TeV in the center of mass established the existence of the *t* (top) quark, primarily through analysis of its decay to a *B* meson and a *W* intermediate vector boson. The *t* mass of 174.3 ± 5.1 GeV/c^2 (about the mass of a tungsten atom) is so great that its weak decay through this channel is very fast, and mesonic states of the *t* and $\bar{t}$ quark (analogous to the Υ, the J/ψ, and the ϕ) are not observed, although the observed *t*'s are from the production of $t\bar{t}$ pairs. *See* INTERMEDIATE VECTOR BOSON.

Color. Quarks are understood to have a spin of 1/2; that is, their intrinsic angular momentum is $\hbar/2$ (where $\hbar$ is Planck's constant *h* divided by 2π), just as for the electron and muon. Thus mesons (composed of a quark-antiquark pair) have integral spins, of 0,1, or greater (in units of $\hbar$). Baryons, made of three quarks, have spins of 1/2, 3/2, and so forth. A problem arose when the structure of observed baryons required two or, in some cases, three quarks of the same flavor in the same quantum state, a situation forbidden for spin-1/2 particles by the Pauli exclusion principle. In order to accommodate this contradiction, a new quantum variable, arbitrarily labeled color, was introduced; the idea is that each quark is red, green, or blue (and the antiquarks, anti-red, and so forth). The color quantum number then breaks the degeneracy and allows up to three quarks of the same flavor to occupy a single quantum state. Confirmation of the color concept has been obtained from experiments with electron-positron storage rings, and the theory of quantum chromodynamics (QCD), based on this concept, has been developed. According to quantum chromodynamics, hadrons must be colorless; for example, baryons must consist of a red, a green, and a blue quark, and mesons of a quark-antiquark pair of the same color (for example, a red quark and an anti-red antiquark). *See* COLOR (QUANTUM MECHANICS); EXCLUSION PRINCIPLE; SPIN (QUANTUM MECHANICS).

The field quanta of quantum chromodynamics are gluons; massless, spin-1 quanta which interact with quarks. This is very analogous to the manner in which photons, the quanta of electromagnetic interaction, interact with particles containing electric charge and are responsible for electromagnetic forces. The QCD theory is part of the now widely accepted standard model of elementary particle interactions, together with the electroweak theory. Experiments have increasingly confirmed details of the standard model to the extent that most physicists are confident that it is fundamentally correct. *See* ELECTROWEAK INTERACTION; GLUONS; QUANTUM CHROMODYNAMICS; STANDARD MODEL.

Generations of particles. There are three sets, or "generations," of quarks and leptons. Each generation contains a charged lepton (electron, muon, or tau lepton); a correspoding neutrino; a charge $-1/3$ quark color triad; and a charge $+2/3$ quark triad. Evidence for a fourth kind of neutrino would signify a new massive lepton and a pair of very massive, new quarks. Detailed measurements of the decays of the Z^0 intermediate vector boson at electron-positron colliders have firmly established the number of neutrinos at three (2.984 ± 0.008 in a 2000 compilation). This confirms the conclusion that the three generations of elementary particles now known are a complete set. *See* LEPTON; NEUTRINO.

Searches for free quarks. The QCD theory argues that free quarks do not exist, because the gluon-mediated force between quarks increases without limit proportional to the separation between single quarks. When two quarks are separated, beyond a certain point the energy required results in the creation of a quark-antiquark pair, resulting in a meson and a baryon, for example. During the 1960s and 1970s there were intensive searches for evidence of free quarks, at particle accelerators, in cosmic rays, and in stable matter.

The signature for a free quark should be clear. A fast-moving particle of charge $\pm 1/3$ or $\pm 2/3$ the charge on a proton or electron would produce ionization in passing through matter 1/9 or 4/9 as great as that of an integrally charged particle (such as an electron or muon). No confirmed evidence for free quarks was found. Upper limits to the possible production cross sections for free quarks from particle accelerator experiments are shown in **Fig. 3**. In spite

of the theory and the body of negative evidence, some physicists still seek evidence for the existence of free quarks. A Stanford University group has an ambitious program to study stable matter using a modernized, automated version of the Millikan oil drop method (first used in 1911 to establish the charge on an electron) to seek free quarks or at least to set much more stringent limits on their existence. *See* ELECTRON.

Unresolved questions. Quarks and the theory of quantum chromodynamics are now firmly established as cornerstones of the standard model of elementary particles (together with the electroweak theory, charged leptons, neutrinos, and so forth). However, unanswered questions remain.

The advanced string and M theories have the property of supersymmetry, which demands that every spin-1/2 quark and lepton must have a partner with integral spin. As of 2000, no experimental evidence for any of these supersymmetric (SUSY) particles had been found; if they exist, they must have rest masses greater than 30–100 GeV/c^2, depending on the candidate particle (or about 30–100 times the mass of a proton). As new accelerators enable searches at higher energies, the search for these particles will continue. *See* SUPERSTRING THEORY; SUPERSYMMETRY.

Contemporary theories also predict that there exists one or more massive particles of integral spin, the Higgs paticles, responsible for the rest masses of the quarks and charged leptons. Experiments have precluded the existence of Higgs particles at masses below 100 GeV/c^2. In late 2000, however, groups working at the CERN electron-positron collider, LEP, reported possible evidence for a Higgs boson of a mass of about 115 GeV/c^2. The statistical evidence for this discovery is weak, and confirmation from future experiments is required before it can be included in tables of "fundamental particles" with confidence. *See* HIGGS BOSON.

Quarks may be permanently stable against decay via the weak interaction; however, it is also possible that quarks spontaneously decay to leptons. Intensive searches for the decay of the proton (into a neutral pion and a positron, for example) have been negative, setting a lower limit of over 10^{32} years for the proton lifetime. However, the apparent asymmetry of the universe between matter and antimatter (there is, at present, no evidence for primordial antimatter) suggests that antiprotons, for example, may spontaneously decay (or transform) more readily. Experiments have shown that the antiproton lifetime is at least 10^6 years; however, the age of the universe is over 10^{10} years. *See* ANTIMATTER; PROTON.

Some theories have postulated that quarks are composed of smaller constituents, just as other objects that were originally believed to be fundamental subsequently were found to have internal structure. Objects called preons by some theorists could make up the different flavors of quarks. Again, there is no evidence for such internal structures. So far, all observations are compatible with the quarks as point objects, like the electron. Lawrence W. Jones

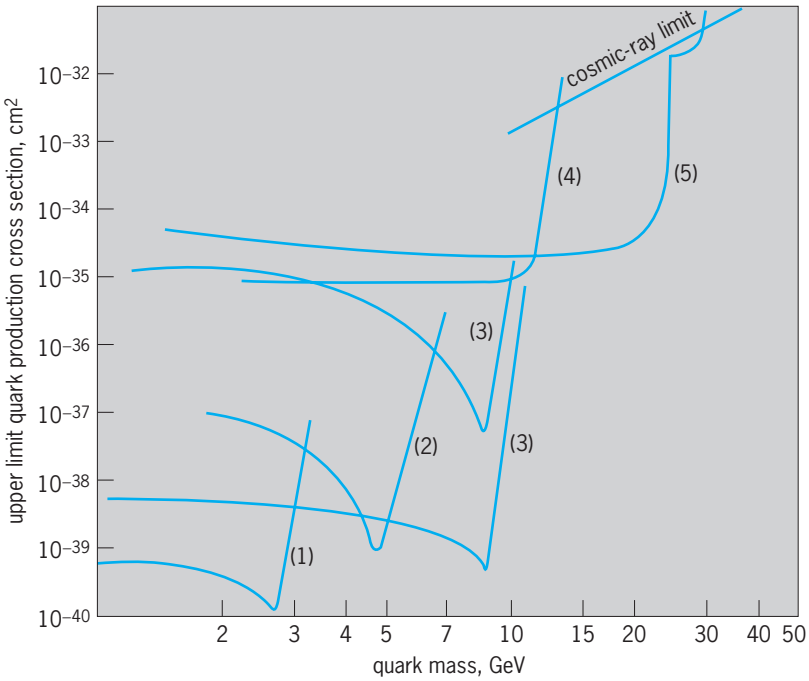

Fig. 3. Upper-limit cross section for production of quarks of charge $-{}^1/_3e$ as a function of quark mass from experiments at various particle accelerators and cosmic rays. The different experiments are indicated by numbers: (1) CERN 28-GeV proton synchrotron; (2) Serpukov 70-GeV proton synchrotron; (3, 4) Fermi National Accelerator Laboratory synchrotron; (5) CERN Intersecting Storage Rings. (*After L. W. Jones, A review of quark search experiments, Rev. Mod. Phys., 494:717–752, 1977*)

Theory. The theory of the properties and interactions of quarks will now be discussed in detail.

Local gauge theories. The theory of the interactions of quarks is expressed, at a fundamental level, in terms of local gauge theories: theories which are invariant under independent symmetry transformations at each space-time point. Such a theory is characterized by its symmetry group, coupling constants, matter fields, and whether or not the symmetry is broken spontaneously. (Spontaneously broken symmetry is described below.) Local gauge theories first entered elementary particle physics in quantum electrodynamics (QED), the theory of the interactions of the electron-positron field with the Maxwell field of electromagnetism. For quantum electrodynamics, the symmetry group is U(1), the group of complex numbers of modulus one which can be identified with rotations of the complex plane; the coupling constant is the electric charge, the matter field is the electron-positron field, and the group is unbroken. *See* QUANTUM ELECTRODYNAMICS.

Local gauge theories were generalized to "non-abelian" theories by C. N. Yang and R. L. Mills. Non-abelian theories have a non-abelian group, that is, a group in which, in general, $g_2g_1 \neq g_1g_2$, where the g's are group elements. Typical non-abelian groups are the SU(N) and SO(N) groups. SU(N) is the group of N by N unitary (U) matrices which have determinant one (or are "special," S). Unitary matrices

preserve the length in a complex space: they have the property $UU\dagger = I$, where I is the unit matrix, and $U\dagger$ is the complex conjugate transpose of U. SO(N) is the group of N by N orthogonal matrices. Orthogonal matrices preserve the length in a real space: they have the property $OO^t = I$, where O^t is the transpose of the matrix O. The group of rotations in three-dimensional space is the group SO(3). *See* GROUP THEORY; MATRIX THEORY.

The groups which are relevant for gauge theories are Lie groups, named after S. Lie. These groups are continuous, like the group of rotations in three-dimensional space, and have further smoothness properties, which are also shared by SO(3). An arbitrary element of a Lie group near the unit or identity can be written as the identity plus a sum of a certain number of "generators" multiplied by real numbers. Each generator corresponds to a small or "infinitesimal" transformation away from the identity. For SO(3), the generators can be chosen to be the deviations from the identity associated with small rotations around the x, y, and z axes. *See* LIE GROUP.

In a local gauge theory, there must be a gauge field, whose associated gauge particle is a spin-1 boson, corresponding to each generator. For U(1), which has one generator, there is one gauge field. In quantum electrodynamics, this is the electromagnetic vector potential. For SU(N) and SO(N), there are $N^2 - 1$ and $^1/_2N(N - 1)$ generators and gauge fields, respectively. Gauge fields are the mediators of the interactions between the matter fields. When the local gauge symmetry is unbroken, the particles associated with the gauge fields are massless, corresponding to a long-range interaction.

For groups such as U(1), SU(N), or SO(N), the interactions are determined by a single constant. This fact, and the fact that the form of the interaction is completely fixed by the group and the matter fields, makes such gauge theories tightly constrained. There is a charge associated with each generator of a gauge theory. For quantum electrodynamics the charge is the electric charge of the electron. A crucial difference between non-abelian and abelian gauge theories is that the gauge fields themselves carry the charges in non-abelian theories, while in abelian theories the gauge fields, for example, the electromagnetic potential in quantum electrodynamics, are neutral (that is, do not carry the charges). Thus the gauge fields in a non-abelian theory can interact by exchanging other gauge fields, while the gauge field (or fields) in an abelian theory only interacts with each other indirectly via the matter fields.

Spontaneous symmetry breaking is the phenomenon in which the symmetry is broken by the vacuum state, although the lagrangian and equations of motion of the theory obey the symmetry. Spontaneous symmetry breaking in a local gauge theory gives rise to a mass for the gauge particles associated with the generators for which the symmetry is broken. The interactions mediated by massive gauge particles are short-ranged. *See* GAUGE THEORY; SYMMETRY BREAKING.

TABLE 2. Three generations of quarks and leptons

Generation:	1	2	3
Quarks*	u_α	c_α	t_α
	d_α	s_α	b_α
Leptons	ν_e	ν_μ	ν_τ
	e^-	μ	τ^-

*The index α runs over the three quark colors. The three quarks with charge $^2/_3e$ (u, c, and t) are called the three up quarks u_i; and the three quarks with charge $-^1/_3e$ (d, s, and b) are called the three down quarks d_i.

Electroweak interactions of quarks. The electromagnetic interactions of quarks can be described by using the electromagnetic current given by Eq. (1), where the

$$j_\mu = \sum_i \left({}^2/_3 \bar{u}_i \gamma_\mu u_i - {}^1/_3 \bar{d}_i \gamma_\mu d_i \right) \tag{1}$$

γ_μ are the Lorentz covariant forms of the Dirac matrices, and the sum runs over the three presently known generations of quarks (**Table 2**). The u_i are the so-called up quarks with charge $^2/_3e$ (u, c, and t), and the d_i are the so-called down quarks with charge $-^1/_3e$ (d, s, and b). The six quark species are called flavors. This form of the electromagnetic current says that the electromagnetic interaction does not change the flavor of the quark, and that quarks interact electromagnetically with a strength proportional to their electric charge. Electromagnetism is mediated by the photon, a neutral massless gauge particle.

The weak interactions of quarks can be described by using three weak currents: charge raising and lowering ones, and a neutral current. The prototype of charged-current weak interactions is beta decay, exemplified by the decay of the neutron, $n \to pe^- \bar{\nu}_e$. The charge raising current has the form of Eq. (2), where d'_i stands for a superposition or

$$j_\mu^+ = \sum_i \bar{u}_i \gamma_\mu d'_i \tag{2}$$

mixture of the three down quarks. N. Cabibbo introduced this mixing between the d and s quarks to take account of the fact that strangeness-conserving decays, such as neutron decay, go at a faster rate (when phase space factors are removed) than the analogous strangeness-violating decays, such as lambda decay, $\Lambda \to pe^-\nu_e$.

The theory with only three flavors of quarks leads to "strangeness-changing neutral-current weak decays," such as $K \to \mu\bar{\mu}$, which experimental data show occur at an extraordinarily low rate. To avoid these unseen decays, it was suggested that there should be a fourth flavor of quark, the charm quark c, whose presence in intermediate states would cancel the unobserved decays, provided the mass of the c quark was not too different from the mass of the u quark. The charm theory predicted new hadronic states containing the c quark. As discussed

above, the discovery of the narrow J/ψ resonance at mass 3.1 GeV/c^2 gave dramatic confirmation of the theory. The further narrow Υ resonance discovered at mass 9.4 GeV/c^2 gave evidence for the b, the fifth flavor of quark; and the sixth flavor of quark, the t, was discovered in proton-antiproton collisions.

Until the work of S. Glashow, S. Weinberg, and A. Salam, who contributed theoretical developments which gave a partial unification of electromagnetism and weak interactions, the theory of weak interactions suffered from divergences in calculations beyond lowest order in the Fermi constant, $G_F \sim 10^{-5} M_P^{-2}$, where M_P is the proton mass. This partial unification used a local gauge theory based on the group SU(2) × U(1), which has four generators, three from the group SU(2) and one from U(1). The gauge fields W^+ and W^- associated with two of the SU(2) generators are identified as the mediators of the charge-changing weak interactions mentioned above. The electromagnetic vector potential is associated with a linear combination (or "mixture") of two of the remaining generators. A fourth gauge field Z^0 associated with the orthogonal mixture of the last two generators is required by the theory. Spontaneous symmetry breaking simultaneously produces the mixing of the generators, and gives masses to the $W^\pm$ and Z^0 fields, as required by the short range of the weak interactions. Since the Z^0 is neutral, this partially unified theory predicts the existence of neutral-current weak processes, such as $\nu_\mu p \to \nu_\mu p$. Such processes were found in neutrino experiments, giving striking confirmation of the electroweak theory. *See* NEUTRAL CURRENTS.

The *CP* violation observed in the $K^0 - \bar{K}^0$ system is a puzzling phenomenon. M. Kobayashi and T. Maskawa extended the mixing of the d and s quarks due to Cabibbo to a three-dimensional mixing of the d, s, and b quarks. The *CP* violation is associated with complex elements of the Cabibbo-Kobayashi-Maskawa (CKM or KM) matrix. For fewer than six flavors of quarks, redefining the arbitrary phases of the quarks can always make the matrix real, so at least six flavors of quarks are needed for quark mixing to lead to *CP* violation. Experiments on weak decays of hadrons have given strong constraints on the elements of the CKM matrix. In general, the diagonal elements are close to one, which means that the mass eigenstates of quarks, which are relevant in strong interactions, are close to the weak eigenstates, which are relevant in weak processes. The elements displaced by one position from the diagonal are small, and the elements doubly displaced from the diagonal are very small or zero, which means that the main mixing is between neighboring generations.

Quantum chromodynamics. The apparently contradictory low-energy (massive, strongly bound, indeed permanently confined) and high-energy (almost massless, almost free) behaviors of quarks can be reconciled by using quantum chromodynamics, which is a local gauge theory based on the group SU(3), as the theory of the strong interactions. The eight gauge fields of quantum chromodynamics are called gluons.

The high-energy properties of quarks follow from the asymptotic freedom of quantum chromodynamics. The interquark potential has the form of Eq. (3).

$$V(r) = \frac{g_{\text{eff}}^2(r)}{r} \tag{3}$$

Because of quantum corrections, the effective or running coupling constant g_{eff} of a quantum field theory depends on the distance scale r at which the coupling constant is measured. A theory is asymptotically free if g_{eff} goes to zero as r goes to zero so that perturbation theory is valid for small r. For large r, the potential grows linearly without bound, as given in Eq. (4), where σ is a constant, so that single quarks

$$V(r) \to \sigma r \tag{4}$$

can never be free. This is the phenomenon of permanent quark confinement. Since confinement occurs when g_{eff} is large, nonperturbative methods are necessary to demonstrate confinement.

Lattice gauge theory. The best evidence for confinement comes from the Monte Carlo numerical approximation to lattice gauge theory on finite lattices. The properties of quantum chromodynamics and of quarks can be calculated by using this discrete formulation in which the dynamical variables are placed on the vertices and links of a lattice. This approach allows the treatment of nonperturbative features of quantum chromodynamics. Two types of problems have been studied: demonstration of confinement and finding its properties, and computation of observables such as hadron mass ratios and weak hadronic matrix elements. The results obtained are promising but not definitive. *See* MONTE CARLO METHOD.

Standard model. The local gauge theory based on the group SU(3) × SU(2) × U(1), in which all fundamental interactions have a gauge theory origin, and SU(3) is unbroken while SU(2) × U(1) is spontaneously broken to the U(1) of electromagnetism, serves as a standard model for all elementary particle physics at energies below about 100 GeV.

Beyond the standard model. A major direction of theoretical research beyond the standard model is the search for a grand unified theory in which quantum chromodynamics and the electroweak interactions, and possibly gravity, are part of a larger scheme with a single coupling constant; however, the striking prediction of grand unified theories that the proton should decay has not been confirmed. Two other areas of such research are the exploration of supersymmetric theories in which all fermions have bosonic partners, and vice versa; and the study of models in which quarks, leptons, and other "elementary" particles are composites of still more basic particles. *See* GRAND UNIFICATION THEORIES.

The most ambitious departure concerns superstring theories, which combine the idea of supersymmetry with the use of a one-dimensional "string" as the object over which the quantum theory is defined

to replace the zero-dimensional point over which conventional quantum field theories are defined. Superstring theories lead naturally to a quantum theory of gravity and, for the first time, seem likely to provide a well-defined theory, free of uncontrollable (technically, unrenormalizable) infinities. The hope that the entire content of elementary particle physics can be derived from a superstring theory has not yet been realized; nonetheless, it seems likely that these theories will play an important role in future developments. O. W. Greenberg

Bibliography. P. C. Bhat, H. B. Prosper, and S. S. Snyder, Top quark physics at the Tevatron, *Int. J. Mod. Phys. A*, 13:5113–5218, 1998; S. L. Glashow, Quarks with color and flavor, *Sci. Amer.*, 233(4):38–50, October 1975; D. E. Groom et al. (Particle Data Group), Review of particle physics, *Eur. Phys. J. C*, 15(1–4):1–878, 2000; N. M. Mar et al., Improved search for elementary particles with fractional electric charge, *Phys. Rev. D*, 53:6017–6032, 1996; Y. Nambu, *Quarks*, 1985; A. Pickering, *Constructing Quarks*, 1984; C. Rebbi, The lattice theory of quark confinement, *Sci. Amer.*, 248(2):54–65, February 1983; R. F. Schwitters, Fundamental particles with charm, *Sci. Amer.*, 237(4):56–70, October 1977; P. F. Smith, Searches for fractional electric charge in terrestrial materials, *Annu. Rev. Nucl. Part. Sci.*, 39:73–111, 1989.

Quarrying

The process of extracting rock material from the outer portions of the Earth's crust. The shape of the excavated space is dependent on the geologic condition and orientation of the rock mass being quarried. This excavated space is either an open pit or an underground cavity that is accessed by tunnel or shaft. An underground quarry is usually called a mine, even though the process for making the excavation is usually the same as on the surface. *See* OPEN PIT MINING; ROCK; UNDERGROUND MINING.

The actual quarrying technique depends on the specific use of the excavated rock. Rock that is excavated for use as dimension stone—relatively flawless blocks or slabs of stone—is carefully extracted as large, 20-ton blocks (**Fig. 1**). These blocks are sawed into thin (3-cm or 1.2-in.) slabs, and then cut to size and polished as finished pieces (**Fig. 2**). Rock that is quarried for crushed stone is used either for construction aggregate or as an industrial mineral source. Crushed stone is broken in place by drilling and blasting into manageable-sized pieces for transport by truck or conveyor to a crushing plant, where it is crushed and screened into segregated, sized particles (**Fig. 3**). *See* STONE AND STONE PRODUCTS.

Fig. 1. Pulling a 100-ton marble block in an underground quarry. This block was cut with a wire saw and will be cut into smaller (20-ton) blocks for transporting to the mill. (*Vermont Quarries, Inc.*)

Fig. 2. Marble slab sawed from a block of marble and polished, ready to be sawed to size. (*Geomapping/Technographics*)

Dimension stone. An expanding construction market during the last decade of the twentieth century created an increased demand for dimension stone. This increased demand for stone over other building materials was due to its beauty and durability, as well as reduced costs for stone products. The reduction in cost was due to improved quarry and finishing techniques, as well as automated machinery. Dimension stone consumption is expected to increase, as more stone is used for countertops and floors in both new and renovated buildings.

The methods of quarrying large (20-ton) blocks, measuring approximately 3 × 2 × 1.5 m (10 × 7 × 5 ft), have drastically changed. With the exception of some deep quarries, there are very few derricks used to lift dimension-stone blocks to the surface. Diesel-driven and hydraulically operated rubber-tired

loaders are the predominant means for transporting blocks from the bottom of the quarry to the surface. Small, transportable electrically driven saws with segmented diamond wire or fixed bars with carbide-impregnated chains are used for cutting manageable-sized blocks from the quarry. Hydraulic jacks and pneumatic wedges have replaced mechanical wedges for splitting blocks.

In practice, the specific system that is selected for quarrying depends on the stone type. Slate is easily split along a major cleavage plane. Most slate, after being sized by circular diamond saws, is mechanically split into roofing tiles. Commercial marble deposits are more massive and do not split as readily as slate, but are easily sawed into blocks. Commercial granite deposits are also massive, but they tend to split in a preferred direction called the "easy way." A light explosive charge placed in equally spaced, drilled holes along this plane will separate individual blocks from the main stone mass. Blocks of soapstone, limestone, sandstone, and travertine are all quarried by using a combination of drill, saw, and various splitting techniques. *See* GRANITE; LIMESTONE; MARBLE; ROCK MECHANICS; SANDSTONE; SLATE; SOAPSTONE; TRAVERTINE.

Crushed stone. Construction aggregate is the stone that is used for making asphalt and concrete roadways. It also is the stone, sand, and gravel used for building foundations and for fill in landscaping. Industrial minerals are the minerals and rock quarried for agricultural and manufactured products. Toothpaste, paint, plastics, paper, and glass products all contain some minerals, including clay, titanium, wollastonite, talc, gypsum, as well as phosphate, limestone, and sandstone rock. Each person living in North America will consume or use in her or his lifetime more then 61 tons of these minerals, all of which have to be quarried from the Earth's crust. The method used for quarrying crushed stone has remained constant, while improvements in the technology and the size of the equipment have allowed a greater tonnage to be quarried, along with minimizing the environmental impact. *See* CONCRETE; CRUSHING AND PULVERIZING; GRAVEL; MINERALS; SAND.

The main task in quarrying crushed stone is to break large rock masses into smaller pieces. Depending on the geology and uniformity of the rock being quarried, large areas are drilled for blasting into piles of smaller rock. The blast holes are 8–25 cm (3–10 in.) in diameter and usually 6–20 m (20–66 ft) deep. The drilled holes are filled with a blasting agent and detonated in a specific sequence, which breaks the rock mass into predetermined-sized pieces. These pieces are transported to a crushing and screening plant for further size reduction and segregation. Additional processing separates specific minerals for industrial use. Softer rock types, such as salt, gypsum, clay, and phosphate rock, are quarried without explosives. Various rock-grinding machines are used to cut, scrape, and remove the soft rock from the quarry floor. These machines produce small fragments, which are further processed into finished products.

Lance Mead

Fig. 3. A 20-ton load of crushed marble being readied for transport to a plant facility to separate the pure calcium carbonate mineral from the rock. The calcium carbonate will be used as filler or extender in industrial products. (*Geomapping/Technographics*)

Bibliography. D. D. Carr (ed.), *Industrial Minerals and Rocks*, 6th ed., Society for Mining Metallurgy and Exploration, 1994; H. L. Hartman (ed.), *SME Mining Engineering Handbook*, Society for Mining Metallurgy and Exploration, 2d ed., vol. 2, chaps. 14.1 and 14.2, 1992; *Mineral Yearbook*, U.S. Geological Survey, annually.

Quartz

The most common oxide on the Earth's surface, constituting 12% of the crust by volume. Quartz is a crystalline form of silicon dioxide (SiO_2). When cooling magmas solidify into igneous rocks and when preexisting rocks metamorphose into new mineral assemblages, the crystallization of quartz is favored by the sheer abundance of silica in the crustal environment. Among the igneous rocks, quartz is especially common within granites, granodiorites, pegmatites, and rhyolites. In addition, quartz can be observed in low- to high-grade metamorphic rocks, including phyllites, quartzites, schists, granulites, and eclogites. Because hydrothermal fluids are enriched in dissolved silica, the passage of fluids through rock fractures results in the emplacement of quartz veins. *See* GRANITE; GRANODIORITE; IGNEOUS ROCKS; METAMORPHIC ROCKS; PEGMATITE; RHYOLITE.

Once quartz has formed, it persists through erosional reworking because of its low solubility in water (parts per million) and its high mechanical hardness (7 on Mohs scale). Consequently, quartz becomes increasingly concentrated in beach sands as they mature, and it is a major component of sandstone. In sedimentary environments, quartz also forms as the final crystallization product during

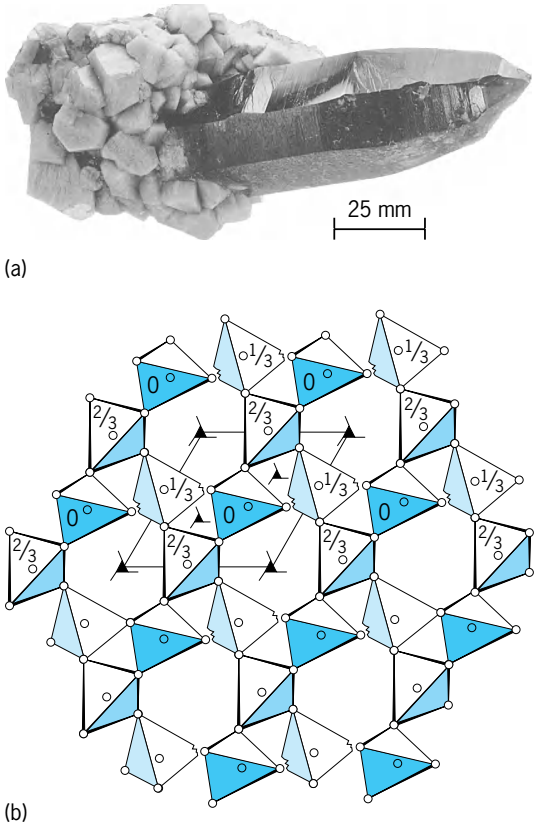

Fig. 1. Quartz. (*a*) Smoky quartz crystal with microcline, found in Florissant, Colorado (*American Museum of Natural History specimen*). (*b*) Projection of the atomic structure of α-quartz along the *c* axis. Each Si^{4+} cation is surrounded by an O^{2-} coordination tetrahedron (outlined). Numbers next to the Si atoms represent fractional coordinates along the *c* axis. Unit cell edge measures 0.491 nanometer.

silica diagenesis; amorphous silica on the sea floor that derives from the skeletons of diatoms, radiolarians, and sponges will transform to quartz upon prolonged exposure to increased temperatures ($\leq$300°C or 572°F) and pressures ($\leq$2 kilobars or 200 pascals) after burial. *See* DIAGENESIS; HARDNESS SCALES; SANDSTONE.

Crystal structure and chemistry. As with virtually all silicates, the atomic framework of the quartz structure consists of Si^{4+} cations that are tetrahedrally coordinated by oxygen anions (O^{2-}) [**Fig. 1**]. Every oxygen anion is bonded to two silicon cations, so that the tetrahedral units are corner-linked to form continuous chains. In low-temperature quartz (or α-quartz), two distinct tetrahedral chains spiral about the crystallographic c axis, creating ditrigonal tunnels along this direction. These chains wrap around the *c* axis in accordance with threefold screw symmetry, and α-quartz belongs to the trigonal crystal system. Prismatic crystals of quartz (Fig. 1*a*) are elongate parallel to these intertwining tetrahedral chains.

Although the silica tetrahedra can be depicted as spirals about the *c* axis in a left-handed sense, right-handed quartz crystals are found in nature as abundantly as are left-handed crystals. These enantiomorphic varieties are known as the Brazil twins of quartz, and they may be distinguished by crystal shape (corresponding crystal faces occur in different orientations) and by opposite optical activities. When polarized light passes through left-handed quartz, the plane of vibration is rotated counterclockwise as the light comes from the crystal toward the observer (levorotation); by contrast, right-handed quartz is dextrorotatory. *See* CRYSTAL OPTICS; POLARIZED LIGHT.

In addition to the Brazil twins, natural quartz crystals frequently contain structural domains called Dauphiné twins. These are regions within an α-quartz crystal that are rotated with respect to each other by 180° about the *c* axis. Usually Dauphiné twin domains are not evident to the unaided eye, but they are visible in reflected light after crystals are etched with hydrofluoric acid. Whereas the opposite handedness that describes the Brazil twins of quartz is a happenstance of crystal growth, Dauphiné twins generally form during the cooling of high-temperature β-quartz to α-quartz, and they may be induced when α-quartz crystals are stressed.

Impurity concentrations in natural α-quartz crystals usually fall below 1000 parts per million. Because the void space within the ditrigonal tunnels is fairly restricted, the most common interstitial cations are the monovalent species with small radii: sodium (Na^+), lithium (Li^+), and especially hydrogen (H^+). Inclusion of even small amounts of hydrogen dramatically weakens quartz, and so-called wet quartz crystals respond to geologic pressures in a plastic rather than brittle fashion. Charge balance is achieved by the substitution of such trivalent cations as aluminum (Al^{3+}) and iron(III) [Fe^{3+}] for silicon cations. Upon irradiation, crystals containing these impurities display characteristic color centers. The violet and yellow hues observed in amethyst and citrine are associated with Fe, and black smoky quartz contains Al. The white coloration of milky quartz reflects light scattering off minute fluid inclusions, and the pink tint in rose quartz is believed to arise from fine-scale intergrowths of a pegmatitic mineral called dumortierite [$Al_{27}B_4Si_{12}O_{69}(OH)_3$]. *See* AMETHYST; DUMORTIERITE.

Silica system. When α-quartz is heated above 573°C (1063.4°F) at atmospheric pressure (1 bar or 100 kilopascals), the structure transforms to β-quartz (**Fig. 2**). This transition requires the bending but not the breaking of primary Si-O bonds, and thus it is rapid and reversible. High-temperature β-quartz has a more open framework with hexagonal symmetry, and the density of β-quartz (2.53 g/cm^3) is lower than that of α-quartz (2.65 g/cm^3). With increasing temperature, β-quartz undergoes a radical restructuring that involves the rupture and reassembly of Si-O bonds to form the mineral tridymite. At even higher temperatures, tridymite in turn experiences a reconstructive transformation to the mineral cristobalite. At 1727°C (3141°F), cristobalite melts to silica liquid. Silica glass, cristobalite, and tridymite may be found naturally in volcanic rocks that have quenched upon eruption. *See* SILICA MINERALS; SILICATE MINERALS; SILICATE SYSTEMS.

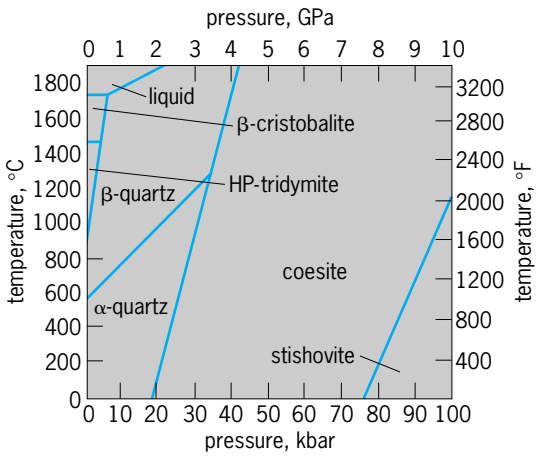

Fig. 2. Phase diagram for the silica system representing fields of thermodynamic stability for the various silica minerals as a function of pressure and temperature. HP indicates hexagonal-symmetry, primitive unit cell. (*After C. Klein, and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*)

When quartz is exposed to high pressure, it converts to the mineral coesite, which is structurally similar to feldspar. Above ~76 kbar (7.6 gigapascals), coesite transforms to stishovite, which is unique among the silica minerals in that each Si atom is coordinated by six rather than four oxygen atoms. Stishovite has the same structure as rutile (TiO_2). Coesite and stishovite occur at the sites of meteorite impacts on the surface of the Earth and may be major constituents of the Earth's mantle. *See* COESITE; COORDINATION CHEMISTRY; FELDSPAR; STISHOVITE.

Microcrystalline quartz commonly is intergrown with a metastable silica mineral called moganite. Microcrystalline varieties usually are classified by their textures: Fine-grained fibrous silica includes chalcedony, agate, and carnelian. By contrast, chert, flint, jasper, and onyx are equigranular varieties. These silica varieties have an extensive and sometimes ambiguous nomenclature arising from their historical role as gemstones. *See* AGATE; CHALCEDONY; CHERT; GEM; JASPER; ONYX.

Uses. Quartz is used predominantly by the construction industry as gravel and as aggregate in concrete. In addition, quartz is important in advanced technologies. Quartz is piezoelectric and has an extremely high quality factor. The high quality factor means that a bell made of quartz would resonate (ring) for a very long time. This property, combined with its piezoelectric behavior, makes quartz the perfect crystal for oscillators in watches.

Compression of α-quartz perpendicular to the c axis creates an electrostatic charge, and this property is exploited in oscillator plates in electronic components. Large flawless crystals of quartz are routinely synthesized for oscillators and for prisms in laser optic systems. Quartz also is employed in abrasives, fluxes, porcelains, and paints. *See* CONCRETE; OSCILLATOR; PIEZOELECTRICITY. Peter J. Heaney

Bibliography. C. Frondel, *The System of Mineralogy*, 7th ed. vol. 3, 1962; P. J. Heaney, C. T. Prewitt, and G. V. Gibbs (eds.), *Silica: Physical Behavior, Geochemistry and Materials Applications*, 1994; M. O'Donoghue, *Quartz*, 1987; R. B. Sosman, *The Phases of Silica*, 1965.

Quartz clock

A clock that makes use of the piezoelectric property of a quartz crystal. When a quartz crystal vibrates, a difference of electric potential is produced between two of its faces. The actual displacement of the surface is generally only a few atomic diameters. The crystal has a natural frequency of vibration that depends on its size and shape. If it is placed in an oscillating electric circuit having nearly the same frequency as the crystal, it is caused to vibrate at its natural frequency, and the frequency of the entire circuit becomes the same as the natural frequency of the crystal. The first quartz clock was invented by Warren Marrison in 1927. *See* OSCILLATOR; PIEZOELECTRICITY.

The actual oscillation frequency depends not only on the quartz crystal but also on the oscillator circuit. In the quartz oscillator, this natural frequency may be used to produce other frequencies such as 1 or 5 MHz. A clock displaying the time of day can also be driven by using one of these frequencies.

The natural frequency of a quartz crystal is nearly constant if precautions are taken when it is cut and polished and it is maintained at nearly constant temperature and pressure. After a crystal has been placed in operation, its frequency usually varies slowly as a result of physical changes. If allowance is made for changes, laboratory quartz-crystal clocks may run for a year with accumulated errors of less than a few thousandths of a second. However, quartz crystals typically used in watches may accumulate errors of several tens of seconds in one year. Modern watches that utilize quartz crystals often use crystals that are chemically etched to shape from thin sheets of quartz. Accuracy depends on the precision of the cut and the degree of contamination of the quartz. *See* WATCH.

For comparison, clocks using rubidium as a frequency standard might be expected to have accumulated errors of less than a few ten-thousandths of a second, while those using cesium might be expected to have accumulated errors better than a few millionths of a second. *See* ATOMIC CLOCK.

The advantage of quartz clocks is that they are reliable, relatively accurate and inexpensive, and easy to use in various applications such as computers and microprocessors. Thus, despite their inaccuracy relative to some other types of clocks, they enjoy wide popularity, particularly in applications requiring accurate timekeeping over a relatively short time span. Natural and cultured quartz crystals are used in modern devices. In these applications, the rates and epochs of the quartz clocks may be readjusted periodically to account for possible accumulated errors. *See* CLOCK; COMPUTER; MICROPROCESSOR; TIME. Dennis D. McCarthy

Quartzite

A metamorphic rock consisting largely or entirely of quartz. Most quartzites are formed by metamorphism of sandstone; but some have developed by metasomatic introduction of quartz, SiO_2, often accompanied by other chemical elements, for example, metals and sulfur (ore quartzites). The geological relations and the shape of quartzite bodies serve to distinguish between them (see **illus.**). The metasomatic quartzites are often found as contact products of intrusive bodies. *See* METASOMATISM; SANDSTONE.

The transition from sandstone to quartzite is gradational. All states of relic clastic structures are encountered. Some sandstones are soon completely metamorphosed. Others are very resistant, and in many highly metamorphic quartzites of the Precambrian, there are relic structures still to be observed.

Pure sandstones yield pure quartzites. Impure sandstones yield a variety of quartzite types. The cement of the original sandstone is in quartzite recrystallized into characteristic silicate minerals, whose composition often reflects the mode of development. Even the Precambrian quartzites correspond to types that are parallel to present-day deposits.

Carbonate cement reacts with silica to produce silicates during high-temperature metamorphism. However, ferric silicate has never been observed. The Fe_2O_3 pigment in deposits of desert sand resists any degree of metamorphism. Therefore, old Precambrian quartzites exhibit the same red color as the present-day sand of Sahara.

Under the conditions of regional metamorphism, cement composed of clay gives rise to sillimanite or kyanite, potash-rich cement yields potash feldspar or mica, lime and alumina yield plagioclase or epidote, dolomitic cement yields diopside or tremolite, and siderite cement yields gruenerite. Wollastonite will crystallize from pure calcite cement. Such quartzites occur as folded layers alternating with layers of other sedimentary rocks.

Waterloo quartzite, with schistosity developed by shearing on bedding planes. Dodge County, Wisconsin. (*USGS*)

In some feldspathic quartzites, thin dark micaceous layers parallel to the foliation superficially suggest relic bedding. Many of the Moine schists of Scotland illustrate this. True bedding is marked, however, by thin strings of zircon, iron ore, or other inherited concentrations of the heavy minerals.

Under the condition of contract metamorphism, the cement of the original sandstone will recrystallize and minerals of the hornfels facies will develop. Quartz itself is usually stable. However, in very hot contacts, against basic intrusions, quartz may invert to tridymite, which again has reverted to quartz; or it may even melt. Such vitrified sandstones resulting from partial fusion of inclusions in volcanic rocks are called buchites. They often contain fritted feldspar fragments, corroded grains of quartz, and a matrix of slightly colored glass corresponding to the fused part of the sandstone. *See* HORNFELS; METAMORPHIC ROCKS; QUARTZ; SILICATE MINERALS.

T. F. W. Barth

Quasar

An astronomical object that appears starlike on a photographic plate but possesses many other characteristics, such as a large redshift, that prove that it is not a star. The name quasar is a contraction of the term quasistellar object (QSO), which was originally applied to these objects for their photographic appearance. The objects appear starlike because their angular diameters are less than about 1 second of arc, which is the resolution limit of ground-based optical telescopes imposed by atmospheric effects. Stars also have angular diameters much less than this, and so they too appear unresolved or pointlike on a photograph.

Discovery. Quasars were discovered in 1961 when it was noticed that very strong radio emission was coming from a localized direction in the sky that coincided with the position of a starlike object. Prior to this time, it was believed that strong radio emission originating beyond the solar system came only from the direction of certain exploding galaxies (radio galaxies) or from previously exploded stars (supernova remnants). Two techniques were used by radio astronomers to measure within a few arc-seconds the direction of the radio emission. The more common technique, called interferometry, makes use of two (or more) antennas pointing in the direction of the source. After the signals are amplified at each antenna, they are combined via cables or a microwave radio link. Another method, very long baseline interferometry (VLBI), records the signals on magnetic tapes which are then synchronously played back so that the signals can be combined. The combined signals yield sinusoidally varying output voltages called fringes. Fringes represent the beating in and out of phase of the incoming radio waves arriving at different times at the two antennas. The phase of the fringes yields an accurate determination of the

direction of arrival of the radio waves. *See* INTERFEROMETRY; SUPERNOVA.

The second technique makes use of the motion of the Moon through the sky in its orbit around the Earth. Occasionally the Moon will pass in front of a radio source and block its emission from reaching the Earth. The exact time of such an occultation is observed by using a radio telescope and is compared with the known position of the edge of the Moon at that instant. When the source reappears from behind the Moon several minutes later, the time of reappearance is also measured and compared with the location of the Moon's edge. These two fixes on the radio source are usually sufficient to permit its position to be measured accurately. *See* OCCULTATION.

When these techniques were used to make accurate position measurements of small-angular-diameter radio sources, the coincidence with starlike objects on optical photographs led to the discovery of a new, hitherto-unsuspected class of objects in the universe, the quasars. The full significance of the discovery was not appreciated until 1963, when it was noted that the hydrogen emission lines seen in the optical spectrum of the quasar 3C 273 were shifted by about 16% to the red from their normal laboratory wavelength. This redshift of the spectral lines is characteristic of galaxies whose spectra are redshifted because of the expansion of the universe and is not characteristic of stars in the Milky Way Galaxy, whose spectra are only slightly shifted to the red or blue from the Doppler effect of their small velocities relative to the Earth and Sun.

Optical characteristics. The color of quasars is generally much bluer than that of most stars with the exception of white dwarf stars. The optical continuum spectrum of stars has a blackbody distribution, while the energy distribution of quasars is generally a power law of the form f^{α}, where f is the frequency and α is the spectral index and is usually negative. The blueness of quasars as an identifying characteristic led to the discovery that many blue starlike objects have a large redshift and are therefore quasars. The quasistellar objects discovered this way turned out to emit little or no radio radiation and to be about 20 times more numerous than the radio-emitting quasistellar radio sources (QSSs). Why some should be strong radio emitters and most others not is unknown at the present time. Several orbiting x-ray satellites have found that most quasars also emit strongly at x-ray frequencies. Gamma rays have also been observed in many quasars. *See* GAMMA-RAY ASTRONOMY; X-RAY ASTRONOMY.

Visible and x-ray emission from some quasars is absorbed by nearby dust grains and is reradiated in the far-infrared near 100-micrometer wavelengths. Unfortunately, this emission is difficult to observe since it is absorbed by the Earth's atmosphere. Several quasars have been observed in the far-infrared from space by the *Infrared Astronomical Satellite (IRAS)*. *See* INFRARED ASTRONOMY.

The emission from quasars also varies with time. The shortest time scale of variability ranges from years to months at short radio wavelengths, to days at optical wavelengths, to hours at x-ray wavelengths. These different time scales suggest that the emissions from the different bands originate from different regions in the quasar. Very highly active quasars are sometimes referred to as optically violent variables (OVVs), blazars, or BL Lac's, after the prototype BL Lacertae, a well-known variable "star" that turned out to be a quasar. The optically violent variables have no or very weak emission lines in their optical spectrum.

The rapid fluctuations indicate that there are some components in quasars that have diameters less than a light-hour or of the order of 10^9 km (10^9 mi), the size of the solar system. The radiation mechanism responsible for the optical and x-ray continuum emission is a combination of thermal blackbody radiation, synchrotron emission, and inverse Compton radiation. *See* COMPTON EFFECT; SYNCHROTRON RADIATION.

Emission lines characteristic of a hot ionized gas are superimposed on the continuum. The emission lines are highly Doppler-broadened, indicating internal velocities of hundreds to thousands of kilometers per second. The most common permitted lines seen are those of H, He, C IV, Si IV, and Mg II, while forbidden lines originating from less dense gaseous regions in the quasar are emitted by the ions [C III], [O II], [O III], [Ne V], and others. *See* DOPPLER EFFECT.

In addition to the emission lines, many quasars, especially those with large redshifts, show very narrow hydrogen absorption lines. In contrast to the emission-line regions, the absorbing regions have internal velocities of about 10 km/s (6 mi/s). Frequently, many absorption-line systems are found with differing redshifts. The absorption-line redshifts are typically less than the emission-line redshifts by amounts corresponding to relative velocities of several thousand kilometers per second. In some quasars, there are absorption-line redshifts substantially less than the emission-line redshift. The most likely interpretation of the latter is that the light from the quasar is passing through and being absorbed by a distant galaxy that just happens to lie along the line of sight. The more numerous other lines could be absorptions by clouds in the vicinity of the quasar. Multiple images due to gravitational lensing by the mass of an intervening galaxy have been observed for a few quasars. *See* GRAVITATIONAL LENS.

For the few thousand quasars measured, the emission-line redshift z ranges from near 0 to 6.4. For quasars having a z greater than about 0.5, many normally ultraviolet spectral lines are seen that are redshifted into the visible region of the spectrum. These lines, such as the Lyman α line at 121.6 nanometers—seen when $z > 1.6$—cannot otherwise be seen in astronomical objects observed from the surface of the Earth because of the atmospheric absorption of all radiation less than about 310 nm.

Not all quasar images are entirely point sources of light. Quasar 3C 273 (**Fig. 1**) has an elongated jet of optical emission 20 arc-seconds from the quasar but pointing radially away from it as if the gas in the jet had been ejected from the quasar. The image of

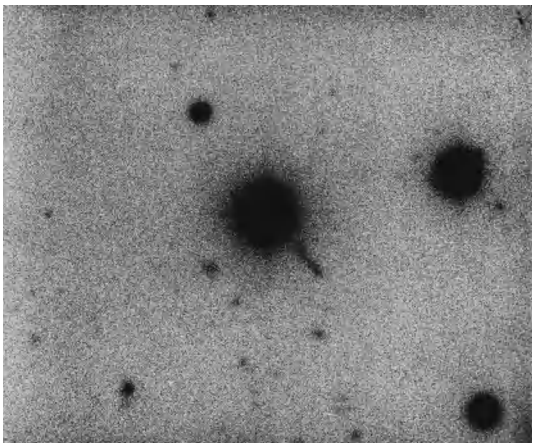

Fig. 1. Photographic negative of the brightest quasar, 3C 273. The cross-shape pattern is due to diffraction within the telescope. The jet extends toward the lower right. (*Palomar Observatory, California Institute of Technology*)

the quasar has been intentionally overexposed to reveal the ejected jet of material. Observations with the Hubble Space Telescope have shown that most quasars have faint nebulous emission surrounding the starlike image. The spectrum of this emission has been found in some cases to be an emission-line spectrum indicating the presence of heated gas extending thousands of light-years from the quasar. However, in most quasars the nebulous emission has a blackbody spectrum with superimposed absorption lines indicating the existence of a galaxylike distribution of stars surrounding the quasar. Furthermore, objects having quasar characteristics but with small redshifts are sometimes found surrounded by a faint distribution of stars. These are called N galaxies and may be similar to the Seyfert galaxies, which are spiral galaxies having a very bright quasarlike nucleus. The existence of these galaxy types and the starlike spectrum of the faint emission seen around some quasars strongly suggest that quasars are exceedingly luminous nuclei of galaxies. Thus, they are often referred to as active galactic nuclei (AGNs). *See* GALAXY, EXTERNAL.

Distances of quasars. When the quasar redshift is plotted against the quasar's brightness or magnitude, as is done for galaxies, there is not a strong correlation between large z and faint magnitudes. Such a correlation is observed for galaxies and indicates that the fainter, more distant galaxies have progressively larger redshifts. This correlation, called Hubble's law, is considered to be proof that the universe is expanding. In the case of the quasars the scatter in the plot is so great that it does not prove that the redshift can be used as a distance indicator. This does not prove that the large-z quasars are not the most distant objects known, since the scatter in the diagram could be due to large differences in the intrinsic luminosities of the quasars. *See* HUBBLE CONSTANT; REDSHIFT.

If the quasars exhibited the Hubble law relationship between their magnitude and redshift, then their redshifts would unambiguously show them to be the most distant known objects in the universe, ranging up to about 10^{10} light-years away. The statistics of the redshift number distribution could then be interpreted as showing a greater abundance of quasars at earlier epochs in the universe. However, the paucity of known quasars with z greater than 6 implies that they may have been absent at even earlier epochs when the density of the universe was hundreds of times greater than its present density.

A few astronomers have argued that the redshifts are not cosmological in origin and hence do not indicate the distances of quasars. One of these arguments rests on the observations that there are examples where two quasars of very different redshift are seen closer together than would be expected by chance if the two quasars were not physically related. On the other hand, the redshifts of quasistellar objects seen in nearby clusters of galaxies have been found to agree with the measured redshifts of the associated galaxies, implying that the redshifts are cosmological. Models postulating noncosmological redshifts, such as gravitational or Doppler redshifts, have met with difficulties. If the redshifts are not cosmological, then they are probably a result of some new unknown physical effect. *See* COSMOLOGY.

Radio characteristics. Some quasars emit a significant fraction of their radiated energy at radio frequencies ranging from about 30 MHz (wavelength, 10 m) to 300 GHz (wavelength, 1 mm). There appear to be two general categories of radio emitters. Those that emit predominantly at the lower radio frequencies (less than 1 GHz) have spectral indices α of about -0.8, showing that the flux of radio emission is rapidly decreasing with increasing frequency. Interferometers have shown that the radio emission from these types of quasistellar radio sources does not originate from the optical quasar but comes from two considerably distant regions symmetrically displaced on either side of the optical object. Such radio-frequency characteristics are nearly identical to those of radio galaxies, which are massive elliptical galaxies, suggesting that the two may be generically related.

The other, more interesting type of quasistellar radio source emits primarily at centimeter and millimeter wavelengths (1 GHz to 300 GHz) and originates in the same region as the optical quasar. The flux of emission of these sources is nearly independent of frequency ($\alpha = 0$) except that there is an abrupt drop-off in the emission spectrum at the low-frequency end. There are a few examples of quasistellar radio sources that exhibit both sets of characteristics, suggesting that one type may evolve into another. In 1965 large variations in the flux of radio emission from the second type of quasistellar radio source were discovered, indicating that, since these types are active, they are young objects while those of the first type are probably decaying remnants. Sudden outbursts of emission are observed (**Fig. 2**), with enormous luminosities sometimes approaching 10^{38} joules/s, nearly 10^{12} times the luminosity of the Sun.

The nature of the origin of this energy is not known and represents one of the most important

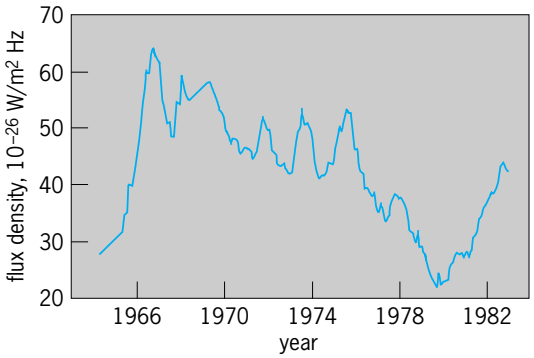

Fig. 2. Outbursts in radio emission which were observed in the nearby quasar 3C 273. Measurements were made at a frequency of 15.5 GHz (1.9 cm) with the 120-ft (37-m) radio telescope of the Haystack Observatory in Massachusetts. Sudden outbursts are evident, particularly the ones in 1966 and 1982.

unanswered questions concerning quasars. The polarization of the radio emission suggests that it is synchrotron radiation produced by electrons with velocities very near the speed of light that are spiraling in magnetic fields. The electrons can absorb their own photons at lower frequencies, producing the drop-off in their emission spectrum. From the observed frequency of this self-absorption, the magnetic field strengths can be estimated as 0.01 to 1 gauss (1 to 100 microtesla).

There have been VLBI observations of some of the active quasars. These observations show that very compact radio components with angular diameters of about 0.001 arc-second exist in quasars of the second type. Furthermore, successive observations seem to show that these compact components are expanding rapidly and separating at speeds greater than the speed of light, when cosmological distances are assumed. A number of theories have been advanced to explain the phenomena without requiring that the quasars be relatively nearby. The most likely explanation is that one of the separating radio components is a jet moving almost directly toward the Earth with a velocity close to the speed of light. At high velocities of approach, effects predicted by the special theory of relativity become important. Intervals of time in a rapidly approaching object would appear to be much shorter, giving the appearance that the object is moving much faster than it really is. This interpretation would also explain the high apparent radio luminosities. *See* RELATIVITY.

Theories of quasars. The many similarities of the observed characteristics of quasars with radio galaxies, Seyfert galaxies, and BL Lacertae objects strongly suggest that quasars are active nuclei of galaxies. There is good statistical evidence which shows that quasars with large redshifts are spatially much more numerous than those with small redshifts. Because high-redshift objects are very distant and emitted their radiation at an earlier epoch, quasars must have been much more common in the universe about 10^{10} years ago. Observations with the *Hubble Space Telescope* have shown that this is the same epoch when galaxies are observed to be forming. Thus it is likely that quasars are associated with the birth of some galaxies.

More than 10^{53} J of energy are released in quasars over their approximately 10^6-year lifetime. Of the known energy sources, only gravitational potential energy associated with a mass about 10^9 times the mass of the Sun can provide this energy, but it is unknown how this gravitational energy produces jets of particles that are accelerated to very near the speed of light.

Several theories have been proposed for quasars. However, the most favored interpretation is that quasars are massive black holes surrounded by rapidly spinning disks of gas in the nuclei of some galaxies. The hot gas in the disk emits the x-ray and optical continuum, a heated halo around the disk produces the emission lines, and the relativistic radio jets are ejected along the rotation axis of the spinning disk. *See* ASTRONOMICAL SPECTROSCOPY; ASTROPHYSICS, HIGH-ENERGY; BLACK HOLE; INFRARED ASTRONOMY; NEUTRON STAR; RADIO ASTRONOMY.

William Dent

Bibliography. G. Burbidge and M. Burbidge, *Quasi-Stellar Objects*, 1967; C. Hazard and S. Mitton (eds.), *Active Galactic Nuclei*, 1979; A. K. Kembhavi and J. V. Narlikar, *Quasars and Active Galactic Nuclei: An Introduction*, 1999; I. Robson, *Active Galactic Nuclei*, 1995; H. L. Shipman, *Black Holes, Quasars, and the Universe*, 1980; D. Weedman, *Quasar Astronomy*, 1986, paperback 1988.

Quasiatom

A structure in which the nuclei of two atoms approach each other closely and their electrons are arranged in atomic orbitals characteristic of a single atom of atomic number equal to the sum of the nuclear charges. Quasiatoms can be formed for short times in atom-atom and ion-atom collisions when the nuclei are much closer than the mean orbital radius of the innermost K-shell electrons. The electrons are then bound in the electric field of both nuclear charges Z_1 and Z_2, which resembles the spherically symmetric $1/r^2$ Coulomb field of a single united atom having charge $Z_{ua} = Z_1 + Z_2$. *See* ATOMIC STRUCTURE AND SPECTRA.

Formation. The formation of transient quasiatoms requires that the electrons of the target atom and projectile ion smoothly rearrange their orbits in response to the changing electric field as the nuclei approach one another. This condition is fulfilled when the orbital velocity of the atomic electrons greatly exceeds the relative velocity of the colliding nuclei. Experimentally, this ratio of electron velocity to nuclear velocity has been achieved in collisions of heavy ions at energies near the nuclear Coulomb barrier. For example, collisions have been studied of 1.4-GeV uranium ions ($Z = 92$) with uranium atoms, in which the K-shell electron velocity v_e is approximately $0.6c$, and the nuclear velocity v_{nuc} is approximately $0.1c$, where c is the velocity of light. The orbits of at least the inner-shell electrons adjust nearly adiabatically

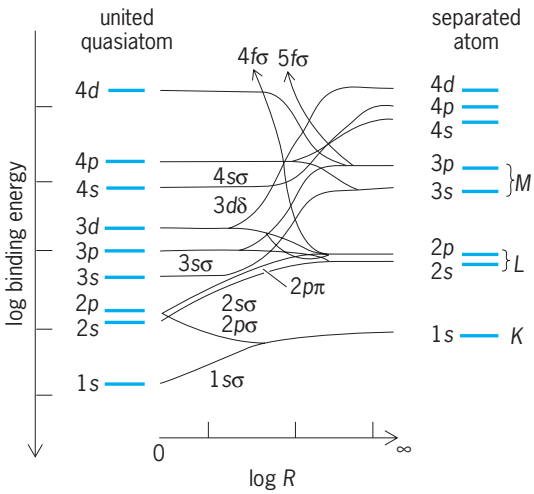

Fig. 1. Binding energies of quasimolecular states evolving from the separated-atom limit to the united-atom limit, as a function of the internuclear separation *R*.

to the changing nuclear charge configuration during the collision.

Initially, at very large internuclear distances, the electrons are bound to the incident projectile nucleus or the target nucleus, with wave functions characteristic of the separated atoms. As the approaching nuclei reach distances comparable to the size of the atomic electron clouds, the electrons respond to the two-centered Coulomb potential of both nuclei. At each moment the electron wave functions resemble those of a diatomic molecule with an interatomic spacing equal to the nuclear separation R. As a function of time, the electrons evolve through a continuous series of quasimolecular states, such as $1s\sigma$ or $2p\pi$ (**Fig. 1**), characterized by $R(t)$. At the smallest internuclear distances of $R = 20$ femtometers (1 fm $= 10^{-15}$ m) for 1.4-GeV uranium-uranium collisions, the nuclei are well within the orbit of the innermost K-shell, and for a short time (10^{-20} s) the electrons are governed not by the individual nuclei but by a single charge distribution with an effective radius of $R/2$. As the nuclei recede, the electrons pass again through the quasimolecular states to the separated-atom limit. *See* MOLECULAR STRUCTURE AND SPECTRA.

Electron excitations. The quasimolecular picture presented above is somewhat idealized, because not all electrons can adjust adiabatically during the collision, especially those that are less tightly bound and have lower orbital velocities. These can be excited to higher quasimolecular orbitals or ejected from the collision system by the time variation of the Coulomb field. Certain electron excitations are enhanced at particular internuclear separations by a process called electron promotion, in which the principal quantum number of an electron is changed during the collision from its value in the separated atom. This occurs when quasimolecular orbitals associated with different principal quantum numbers in the separated-atom limit have nearly equal binding energies at a particular internuclear separation, increasing the probability of induced excitations between these orbitals. In this process, an electron does not always return to its original state in the separated atom after the collision.

The feature of the collision that is responsible for the electron promotion, excitation, and ionization processes is the rapid variation of the two-centered Coulomb field during the approach. Nonadiabatic adjustment of the electrons to the radial component of the electric field as the internuclear distance changes produces a change in their principal quantum number. The rapid rotation of the axis between the nuclei in their plane of scattering as they pass one another can also couple angular momentum to the electrons, changing their orbital-angular-momentum quantum numbers. These excitation mechanisms, unique to the quasimolecular description of ion-atom collisions, explain several experimentally observed phenomena, such as very large excitation and ionization cross sections and the resonant production of K, L, and M x-rays in the separated atoms, which cannot be understood in terms of pure Coulomb excitation during the collisions. *See* X-RAYS.

The first direct experimental evidence for quasimolecule formation came from x-ray observations of medium-heavy collision systems, for example, nickel on nickel. In addition to the characteristic line radiation associated with transitions in the collisionally excited separated atoms, a continuum of x-rays with energies up to the united-atom limit (K x-rays characteristic of nuclear charge $Z_1 + Z_2$) was observed, consistent with the changing quasimolecular binding energies during the collision. A Doppler-shift analysis of this radiation showed that the x-rays were emitted from the center-of-mass system (that is, from the quasimolecule), and detected angular anisotropy was in agreement with the predicted effect of the rotation of the internuclear axis during the collision.

Superheavy quasiatoms. In addition to describing conventional atomic collision phenomena, quasimolecular collisions offer the prospect of forming short-lived quasiatoms at the smallest internuclear separations. This creates a new realm of atomic physics involving very high Z_{ua} atoms and making tests of quantum electrodynamics possible in extremely strong electromagnetic fields, far beyond those available in the heaviest artificial elements. Fundamentally new phenomena are expected to occur in atoms with $Z\alpha \geq 1$ ($\alpha = 1/137$, the fine-structure constant). At $Z = 137$, the Dirac equation for a pointlike nucleus predicts a binding energy of the $1s$ electron equal to its rest-mass energy, $mc^2 =$ 511 keV (m = electron rest mass); and for Z above 137, the Dirac equation breaks down unless the finite size of the nucleus is taken into account. For $Z >$ 173, the $1s$ electron-binding energy is predicted to exceed $2mc^2 = 1.02$ MeV, which opens the interesting possibility of observing a phase change of quantum electrodynamics from a neutral vacuum state to a charged vacuum. In comparison, the heaviest stable atoms that have been studied (such as fermium with $Z = 100$) exhibit binding energies of only 140 keV. *See* RELATIVISTIC QUANTUM THEORY.

Wave functions. Theoretical studies of superheavy quasiatoms indicate that a very different pattern of atomic energy levels is expected in very high-Z systems (**Fig. 2**), in which the electron binding energies and wave functions are dominated by relativistic effects, unlike those encountered in atoms of lower Z. In light, nonrelativistic atoms, the atomic $1s$ electron binding energy scales as Z^2, and the mean radius of the $1s$ orbit scales as $1/Z$. As the atomic number increases past $Z = 137$, the binding energy increases more quickly than Z^2, and the electron wave function becomes more tightly localized about the nucleus as relativistic effects become important. At the highest attainable atomic charges, such as $Z_{ua} = 188$ in collisions of uranium with curium, the $1s$ electron binding energy is almost three times larger than the nonrelativistic prediction, and the wave function is almost ten times more compressed about the nucleus than expected nonrelativistically.

The dominance of relativistic effects in the superheavy quasiatom drastically alters the character of the most tightly bound electron orbitals. For example, the $2p_{1/2}$ orbital, which is nearly degenerate with the $2s_{1/2}$ level in low-Z atoms, is more tightly bound than the $2s_{1/2}$ state by several hundred kiloelectronvolts. Furthermore, the fine-structure splitting between the $2p_{1/2}$ and $2p_{3/2}$ levels is on the order of mc^2 (= 511 keV), comparable to their total binding energies. These effects and the large density of the $2p_{1/2}$ state at the origin are understood in terms of the relativistic spin-orbit coupling, which mixes $l = 0$ and $l = 1$ angular momentum states in the $2p_{1/2}$ wave function and gives it a close resemblance to the $1s_{1/2}$ state. On the other hand, the $2s_{1/2}$ binding is smaller because of the radial node in its wave function, and the $2p_{3/2}$ level because it is an admixture of $l = 1$ and $l = 2$ states, which have low probability density at small radii where the Coulomb field is largest. *See* FINE STRUCTURE (SPECTRAL LINES); RELATIVITY.

Spontaneous decay of the vacuum. An interesting effect is associated with quasiatoms with $Z > 173$, in which the $1s$ binding energy is more than twice the electron rest mass, $E_{1s} > 2mc^2$. If a vacancy exists in this orbital, it is energetically favorable to create an electron-positron pair with the electron bound in this state. The positron would be repelled from the nucleus with kinetic energy equal to $E_{e^+} = |E_{1s}| - 2mc^2$. In the Dirac hole picture, in which the vacuum consists of a negative energy continuum ($E < -mc^2$) filled with electrons, the $1s$ level is said to fall into the negative-energy Dirac sea as Z increases above the critical value, $Z_{cr} = 173$ (Fig. 2). A $1s$ hole (vacancy) becomes embedded in the negative continuum as an unstable resonance state that decays in a time of $\sim 10^{-19}$ s to a bound electron and a spontaneously emitted monoenergetic positron.

The quantum electrodynamic vacuum in the presence of a bare supercritical nuclear charge is therefore unstable and decays to a fundamentally new charged vacuum, which consists of the nucleus with two $1s$ electrons (from the two spin orientations). At higher values of Z_{ua}, as additional quasiatomic levels enter the negative continuum, the charge of the quantum electrodynamic vacuum increases accordingly. If detected, spontaneous positron emission would represent the first observation of a phase transition in a gauge field theory. *See* ANTIMATTER; ELECTRON-POSITRON PAIR PRODUCTION; GAUGE THEORY; PHASE TRANSITIONS; POSITRON; QUANTUM ELECTRODYNAMICS; SUPERCRITICAL FIELDS.

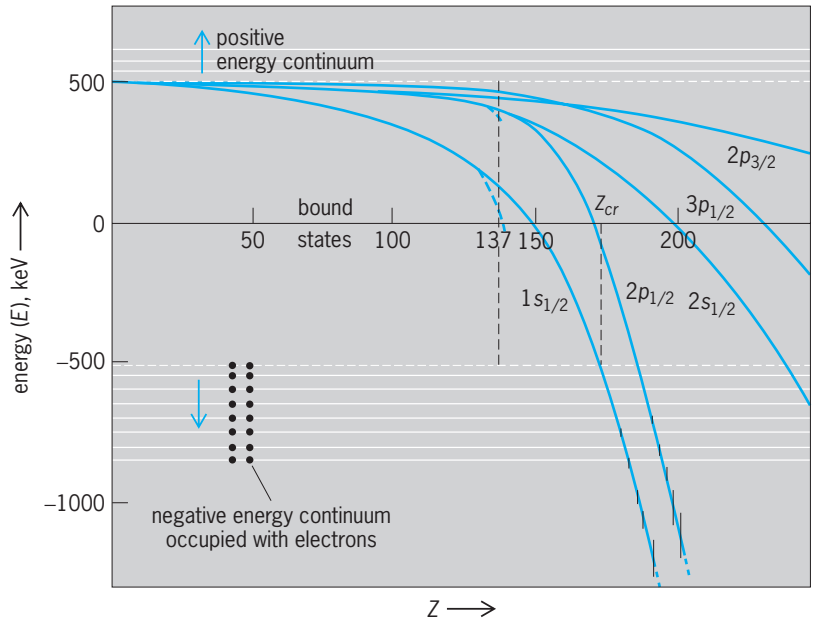

Fig. 2. Calculated quasiatomic binding energies as a function of nuclear charge *Z*. Broken curves represent levels in the field of a point charge; solid curves take into account finite size of nucleus. (*After B. Mueller, Positron creation in superheavy quasi-molecules, Annu. Rev. Nucl. Sci., 26:351–383, 1976*)

Supercritical quasiatomic collisions. It was suggested in 1969 that quasiatomic collisions with $Z_{ua} > 178$ could provide the strong fields necessary to observe the spontaneous decay of the quantum electrodynamic vacuum experimentally. **Figure 3** traces the time development of the internuclear distance and the inner-shell binding energies in a supercritical uranium-uranium collision at 1.4 GeV ($Z_{ua} = 184$). As the ions approach the united-atom limit (at time $t = 0$), the total energy of the $1s_{1/2}$ orbital falls rapidly until, for internuclear separations less than $R_{min} = 27$ fm, the $1s$ state drops into the continuum. On the incoming trajectory, electrons are excited to higher states or ejected to the positive energy continuum as delta electrons, as a result of the interplay between the increased binding and relativistic collapse of the electron wave functions mentioned above. Electron excitation or emission creates vacancies in the inner shells, which, if present in the $1s_{1/2}$ orbital, may decay spontaneously during the short time of supercritical binding ($t_{cr} \sim 2 \times 10^{-21}$ s in elastic collisions). Electron excitation and quasimolecular x-ray emission occur throughout the collision. *See* DELTA ELECTRONS.

In addition to their creation by spontaneous decay of vacancies in the $1s_{1/2}$ state, positrons are produced by the so-called direct ionization of the negative-energy-continuum electrons to positive-continuum states, or by the so-called induced excitation of

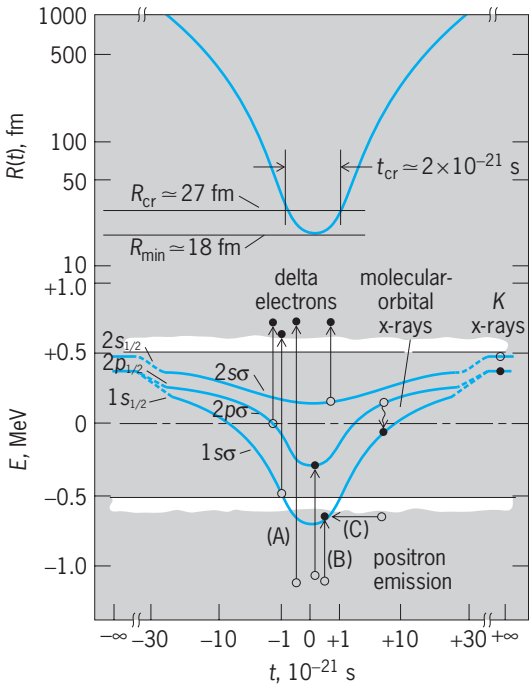

Fig. 3. Internuclear separation *R* and quasimolecular orbital binding energies *E* as a function of time *t* for 1.4-GeV uranium-uranium (^{238}U-^{238}U, $Z_{ua} = 184$) collisions. Delta-electron, x-ray, and (A) direct, (B) induced, and (C) spontaneous positron-emission processes are indicated. (*After J. Eichler, I. V. Hertels, and N. Stalterfoht, eds., Electronic and Atomic Collisions, Elsevier Science, 1984*)

negative-continuum electrons to quasiatomic bound states. This leaves holes in the Dirac sea, and this absence of negative-energy electrons is observed as positive-energy positrons. As opposed to spontaneous positron emission, an inherently static phenomenon occurring as a result of the static field strength, the production of these so-called dynamic positrons is a direct consequence of the rapidly time-varying Coulomb fields present during the quasiatomic collision.

Experimental verification. This picture of superheavy quasiatomic collisions has, to a large extent, been verified experimentally by measurements performed at the heavy-ion UNILAC accelerator in Darmstadt, Germany. In general, the salient feature of these collisions is that the sharply increased electron binding energies and the relativistic collapse of the electron wave functions about the origin with increasing effective nuclear charge combine to dominate the atomic excitation mechanisms. For example, a thousandfold enhancement was found in the K-shell ionization probability in uranium-lead collisions, as against nonrelativistic calculations.

The emission of very high-energy delta electrons, as shown in **Fig. 4*a*** for lead-lead collisions, also provides evidence for relativistic collapse of wave functions and increased binding energy. This follows from kinematic arguments that the maximum energy transferable to a lead K-shell electron from the incident nucleus is only ~200 keV, unless the electron becomes much more tightly bound. The measured energies of up to ~1 MeV, however, imply that the ejected electrons must have been bound with energies of up to $E_{1s} \sim 900$ keV, and localized in the quasiatom with mean radii of less than 100 fm, in agreement with the predictions of quasiatomic theory.

The production of dynamic positrons has also been studied over a large range of values of Z_{ua}. The total probability for emitting positrons in excess of those from nuclear pair-production backgrounds

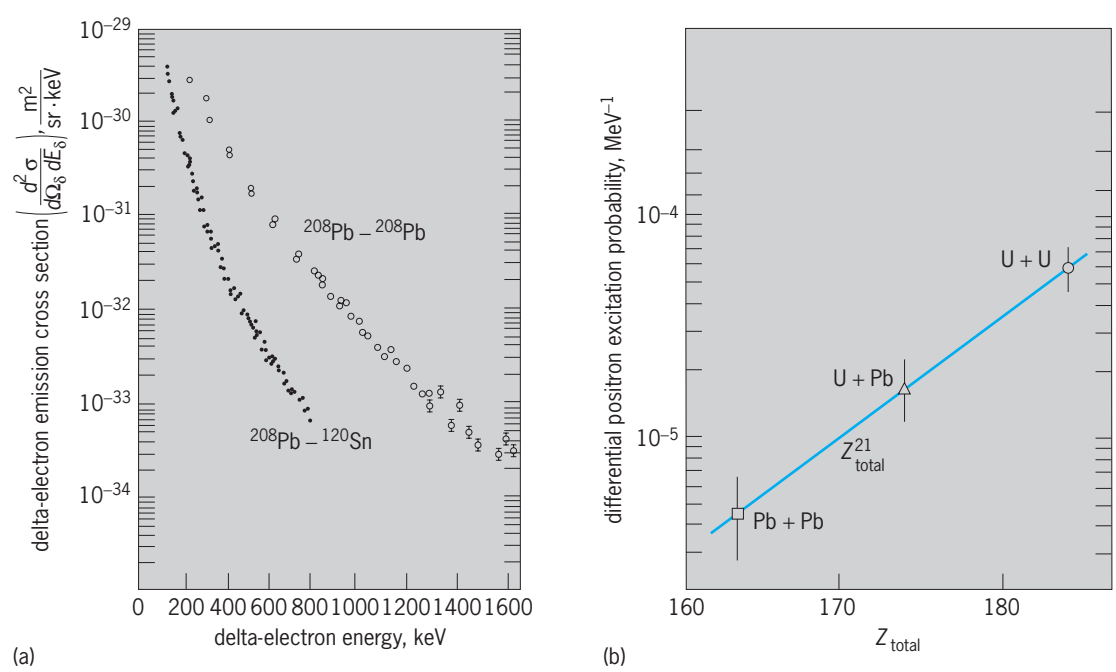

Fig. 4. Electron and positron emission in superheavy quasiatomic collisions. (*a*) Delta-electron energy distributions from lead-lead (^{208}Pb-^{208}Pb, $Z_{ua} = 164$) and lead-tin (^{208}Pb-^{120}Sn, $Z_{ua} = 132$) collisions at bombarding energy of 4.7 MeV/u. Presence of high-energy delta electrons reveals onset of relativistic effects in the higher-Z_{ua} quasiatom (*after S. Datz, ed., Electronic and Atomic Collisions, Elsevier Science, 1982*). (*b*) Total positron creation probability in quasiatomic collisions as a function of combined nuclear charge, $Z_{total} = Z_1 + Z_2$.

rises steeply with an unprecedented $(Z_1 + Z_2)^{20}$ power-law dependence on the combined nuclear charge (Fig. 4*b*), reflecting relativistic effects. The measured dependence of dynamic positron production on internuclear distance and on the incident beam energy, as well as its Z_{ua} dependence, are well reproduced by the theoretical calculations.

An important consequence of the rapid variation of the Coulomb field in quasiatomic collisions is that the distributions of emitted positron, electron, and x-ray energies are all dominated by dynamic effects and are therefore continuous and rather broad (typically ~1 MeV wide). This limits the amount of detailed spectroscopic information that can be experimentally extracted concerning specific quasiatomic energy levels.

Search for spontaneous positron emission. The short time scale of the collision also leads to several consequences that compound the difficulty of detecting spontaneously emitted positrons from quantum electrodynamic vacuum decay: (1) The short time during which the 1*s* state is supercritically bound (2×10^{-21} s) compared to the longer decay time ($\sim 10^{-19}$ s) implies that very few of the vacancies decay spontaneously. (2) By the Heisenberg uncertainty principle ($\Delta E \cdot \Delta t \geq \hbar/2$, where ΔE represents the energy spread, Δt is the time during which emission takes place, and $\hbar$ is Planck's constant divided by 2π), the emitted positron energies would be broadened by 600 keV. (3) The amplitude for spontaneous emission adds coherently with dynamic positron production. For these reasons, no observable signature of spontaneous positron emission is anticipated, and only continuous positron distributions are expected to be found. *See* UNCERTAINTY PRINCIPLE.

It came as a surprise, therefore, when a narrow peak was first discovered in the energy distribution of positrons emitted from uranium-curium collisions ($Z_{ua} = 188$). Perhaps coincidentally, the energy of the positron peak ($E_{e^+} = 316$ keV) corresponds with the binding energy expected for the 1*s* state in the $Z_{ua} = 188$ quasiatom, and its width (~80 keV) implies a long-lived source moving with the center-of-mass velocity of the collision system. These observations encouraged speculation that the peak might be the signature of spontaneous positron emission. Indeed, it was demonstrated that the peak energy, intensity, width, and dependence of the production cross section on the incident beam energy could all be incorporated into the theory of spontaneous positron emission if, in a small fraction (10^{-3}) of the collisions, a dinuclear complex were formed with a lifetime of more than 10^{-19} s, an order of magnitude above the typical collision time. *See* NUCLEAR MOLECULE.

This attractive explanation was brought into question by measurements that probed the theoretically predicted $\sim Z_{ua}^{20}$ dependence of the emitted spontaneous positron energy on combined nuclear charge. Several supercritical collision systems were systematically studied, ranging from uranium-curium ($Z_{ua} = 188$) where the predicted spontaneous

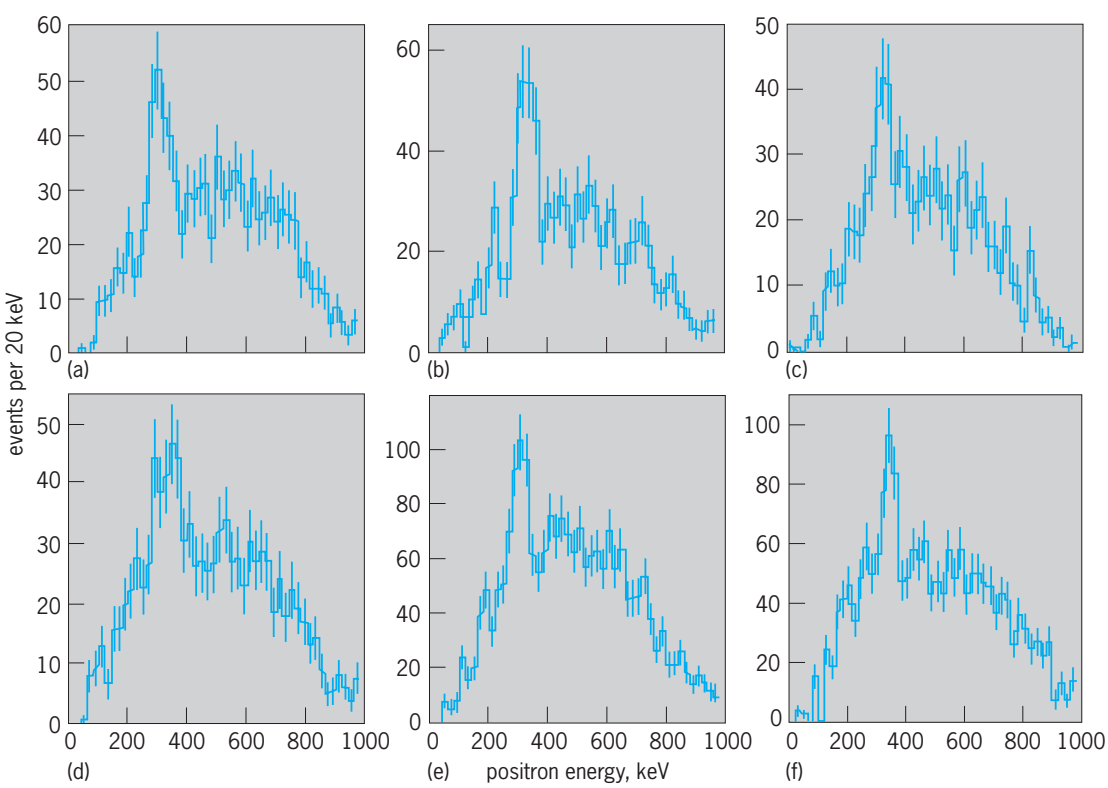

Fig. 5. Energy distributions of positrons emitted from collision systems. The kinematic constraints chosen prominently exhibit the narrow peak structures between 310 and 380 keV. (*a*) Uranium-curium, at bombarding energy of 6.07 MeV/u, $Z_{ua} = 188$. (*b*) Thorium-thorium at 6.02 MeV/u, $Z_{ua} = 186$. (*c*) Uranium-uranium at 5.8 MeV/u, $Z_{ua} = 184$. (*d*) Thorium-uranium at 5.83 MeV/u, $Z_{ua} = 182$. (*e*) Thorium-thorium at 5.75 MeV/u, $Z_{ua} = 180$. (*f*) Thorium-tantalum at 5.78 MeV/u, $Z_{ua} = 163$. (*After M. B. Gilbody et al., eds., Electronic and Atomic Collisions, Elsevier Science, 1988*)

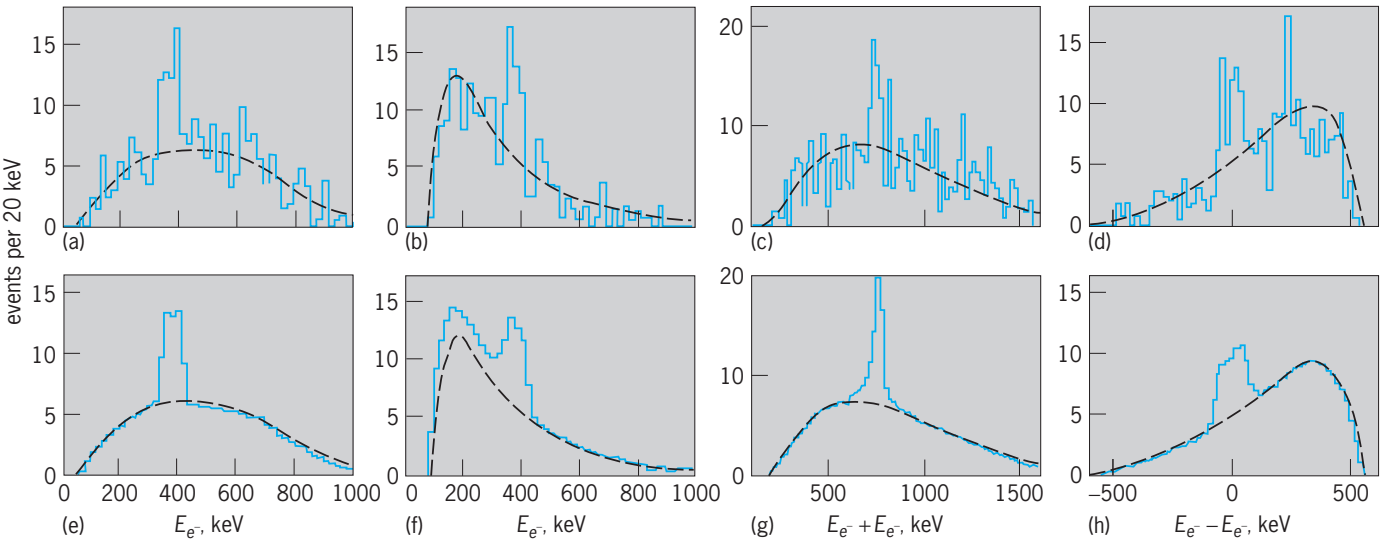

Fig. 6. Coincident positrons and electrons from 1.4-GeV uranium-thorium collisions. Spectra are shown of (*a*) energy E_{e^+}, (*b*) coincident electron energy E_{e^-}, (*c*) sum energy, $E_{e^+} + E_{e^-}$, and (*d*) difference energy, $E_{e^+} - E_{e^-}$ in coincident positron and electron events. Corresponding parts (*e–h*) present Monte Carlo simulation of expected yields, assuming production and decay of a 1.8 MeV/c^2 neutral object. (*After W. Greiner, ed., Physics of Strong Fields, Plenum Publishing, 1987*)

positron energy is $E_{sp} = 300$ keV, and thorium-thorium ($Z_{ua} = 180$) where $E_{sp} = 80$ keV, to the subcritical thorium–tantalum system ($Z_{ua} = 163$), where the nuclear charge is clearly below the supercritical charge threshold and spontaneous positron emission is not expected. Each system studied gave narrow positron peaks between 310 and 380 keV with similar widths and cross sections (**Fig. 5**). No competing gamma-ray or internal conversion electron lines were observed that could explain the data in terms of nuclear-conversion processes, and no instrumental or background processes were found that could produce these structures. All attempts to explain the near independence of the peak energies on Z_{ua} or the presence of a line in the subcritical thorium–tantalum system within the context of spontaneous positron emission have failed.

Correlated electron and positron lines. The similar energies of the positron peaks imply a common origin not involving conventional nuclear or quantum electrodynamic processes. The narrowness of the lines suggest a two-body final state. An obvious possibility pursued experimentally was the two-body decay of a previously undetected neutral object into a positron-electron pair. The signature for this process would be the coincident detection of nearly monoenergetic positrons and electrons, correlated in such a way that the Doppler shifts of their energies cancel, indicating back-to-back emission in the rest frame of their source. *See* DOPPLER EFFECT.

Figure 6 shows the results of a first measurement of coincident positrons and electrons from 1.4-GeV uranium–thorium collisions. An electron peak at $E_{e^-} = 375$ keV (Fig. 6*b*) was detected in coincidence with the positron peak at $E_{e^+} = 380$ keV (Fig. 6*a*). The peak in the distribution of the sum of the positron and electron energies at 750 keV (Fig. 6*c*) is narrower than the individual lines. Each of these spectra, as well as the distribution of the difference of positron and electron energies (Fig. 6*d*), is consistent with that expected from Monte Carlo simulations (Figs. 6*e–b*) of the production and decay of a neutral object of mass 1.8 MeV/c^2 in a small fraction of the collisions. Additional measurements with improved electron energy resolution revealed two additional sets of correlated positron-electron structures with energy peaks at $(E_{e^+} + E_{e^-}) = 817$ keV and 612 keV. Explanations for these structures involving instrumental backgrounds, external or internal pair production, the uncorrelated emission of separate positron and electron lines, and multibody decay are all unable to describe several features of the experimental data, including the narrow sum-energy peak width and the absence of peak intensity in adjacent kinematic regions. Moreover, a kinematic analysis of the widths of the various correlated peaks suggests that the positron and electron are emitted with a 180° back-to-back angular correlation in the rest frame of their source.

These results represent a new source of positron-electron pair production in heavy-ion collisions, and appear to reflect the decay of at least three previously undetected neutral objects with masses of ~1.63, 1.77, and 1.84 MeV/c^2, which are created with very low momenta in superheavy nuclear collisions. Theoretical attempts to explain these data in terms of new elementary particles have met with difficulties due to the negative results of experimental searches for new particles in high-precision quantum electrodynamic measurements, from decays of J/psi and upsilon mesons, in competition with nuclear transitions; from rare pion, kaon, and muon decays; and from electron brehmsstrahlung in beam dumps. Further experiments with superheavy quasiatoms, and attempts to form these objects resonantly in positron-electron scattering, should shed light on the nature of these anomalous correlated positron-electron pairs.

Thomas E. Cowan

Bibliography. D. A. Bromley (ed.), *Treatise on Heavy-Ion Science*, vol. 5: *High Energy Atomic Physics*, 1985; W. Greiner (ed.), *Physics of Strong Fields*, 1987; W. Greiner, B. Muller, and J. Rafelski, *Quantum Electrodynamics of Strong Fields*, 1985.

Quasicrystal

A solid with conventional crystalline properties but exhibiting a point-group symmetry inconsistent with translational periodicity. Like crystals, quasicrystals display discrete diffraction patterns, crystallize into polyhedral forms, and have long-range orientational order, all of which indicate that their structure is not random. But the unusual symmetry and the finding that the discrete diffraction pattern does not fall on a reciprocal periodic lattice suggest a solid that is quasiperiodic. Their discovery in 1982 contradicted a long-held belief that all crystals would be periodic arrangements of atoms or molecules.

Since the eighteenth century, the idea that crystals are periodic arrangements of atoms or molecules gradually moved from a postulate to rationalize the limited number of different rotational symmetries seen in crystals; to a concept experimentally confirmed by the diffraction by crystals of x-rays, neutrons, and electrons; to a possible definition of a crystal. The conjecture that the lowest-energy arrangements of atoms or molecules would always be periodic was widely believed, although it was unprovable and is now known to be false. Early definitions of crystals invoked their polyhedral external form rather than their regular internal arrangements. It is easily shown that in two and three dimensions the possible rotations that superimpose an infinitely repeating periodic structure on itself are limited to angles that are $360^\circ/n$, where n can be only 1, 2, 3, 4, or 6. Various combinations of these rotations lead to only 32 point groups in three dimensions, and 230 space groups which are combinations of the 14 Bravais lattices that describe the periodic translations with the allowed rotations. Until the 1980s, all known crystals could be classified according to this limited set of symmetries allowed by periodicity. Periodic structures diffract only at discrete angles (Bragg's law) that can be described by a reciprocal lattice, in which the diffraction intensities fall on lattice points that, like all lattices, are by definition periodic, and which has a symmetry closely related to that of the structure. *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY; X-RAY CRYSTALLOGRAPHY; X-RAY DIFFRACTION.

Icosahedral quasicrystals were discovered in 1982 during a study of rapid solidification of molten alloys of aluminum with one or more transition elements, such as manganese, iron, and chromium. Since then, many different alloys of two or more metallic elements have led to quasicrystals with a variety of symmetries and structures. Quasicrystals are now made in a number of ways, including several methods of rapid solidification, such as powder and melt-spinning, solid-state transformations using thin-film techniques, and annealing of foils after heavy radiation damage and ion implantation. *See* ALLOY; ALLOY STRUCTURES; CRYSTAL GROWTH; ION IMPLANTATION; RADIATION DAMAGE TO MATERIALS.

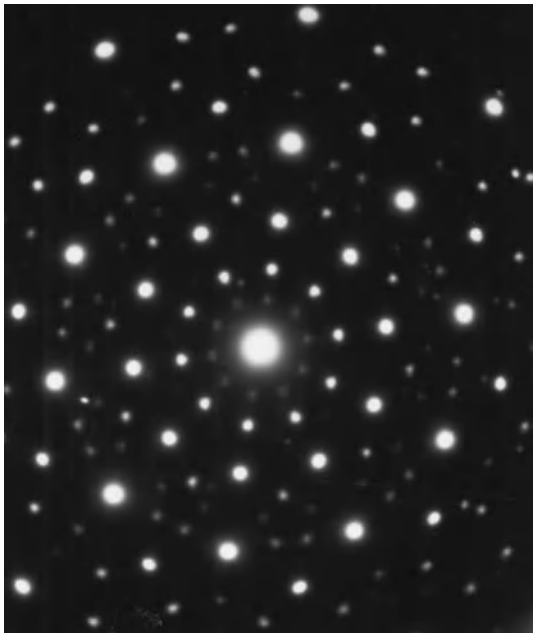

Fig. 1. Diffraction pattern, taken in an electron microscope along the fivefold axis of a rapidly solidified aluminum-manganese alloy, that led to the discovery of icosahedral quasicrystals.

The first quasicrystals were identified in the electron microscope, and their symmetry was established by electron diffraction (**Fig. 1**). There has been a lively debate about whether these early imperfect aperiodic structures would ever evolve toward a truly quasiperiodic structure, or whether the randomness was an intrinsic part of such structures. Because the perfection of the stable aluminum-copper-iron quasicrystals approaches that of crystals, it is now believed that these structures can come as close to the ideal of quasiperiodicity as the skill of the crystal grower permits. **Figure 2** shows the external

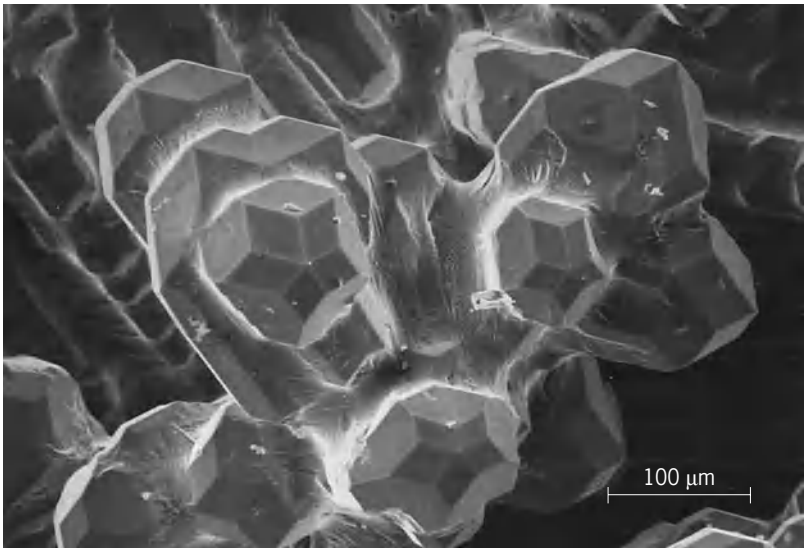

Fig. 2. Quasicrystals of an alloy of aluminum, copper, and iron, displaying an external form consistent with their icosahedral symmetry.

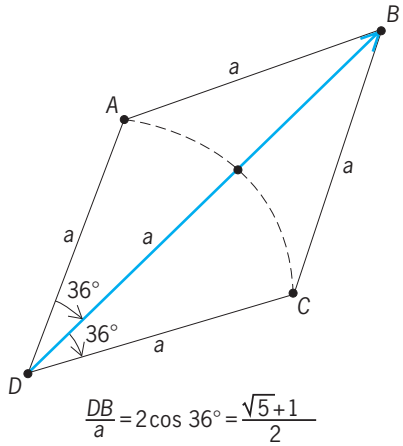

Fig. 3. Simple trigonometric construction showing that distances to successive spots in the diffraction pattern are not in simple rational ratios and therefore cannot be part of a periodic reciprocal lattice.

polyhedral form of an icosahedral aluminum-copper-iron alloy. *See* ELECTRON DIFFRACTION; ELECTRON MICROSCOPE.

Crystallography. The diffraction patterns of quasicrystals violate several predictions resulting from periodicity. Quasicrystals have been found in which the quantity n is 5, 8, 10, and 12. In addition, most quasicrystals exhibit icosahedral symmetry in which there are six intersecting fivefold rotation axes. Furthermore, in the electron diffraction pattern (Fig. 1) the diffraction spots do not fall on a (periodic) lattice but on what has been called a quasilattice. The aperiodicity can be seen in Fig. 1 and in the construction in **Fig. 3** by noting that distances from the center to successive spots along a row are in the ratio 2 cos 36°, which happens to be the algebraic number known as the golden mean, given in Eq. (1).

$$\frac{1+\sqrt{5}}{2} = 1.618034\ldots \tag{1}$$

Because this is not a rational number, the origin and two successive spots on a line cannot belong to a lattice. To locate all the spots in this two-dimensional section, a minimum of four basis vectors is required, rather than the two needed for a two-dimensional lattice. Instead of the usual three for a reciprocal lattice, six basis vectors are needed for indexing the diffraction pattern of this icosahedral alloy. If the diffraction patterns of a structure can be indexed with a finite basis, that is, each spot can be located as the sum of a finite number of vectors, then these patterns fall on a quasilattice and the structure is quasiperiodic. A quasilattice is not periodic, but has the property that any local arrangement is repeated infinitely many times in the infinite domain. A Fibonacci sequence of short and long lengths and corners of Penrose's aperiodic tiling of the plane with tenfold symmetry (**Fig. 4**) are one- and two-dimensional examples of quasilattices in which the golden mean plays a role. Quasilattices are quasiperiodic and lead to diffraction patterns whose peaks do not fall on the points of a lattice. *See* COMBINATORIAL THEORY; POLYGON.

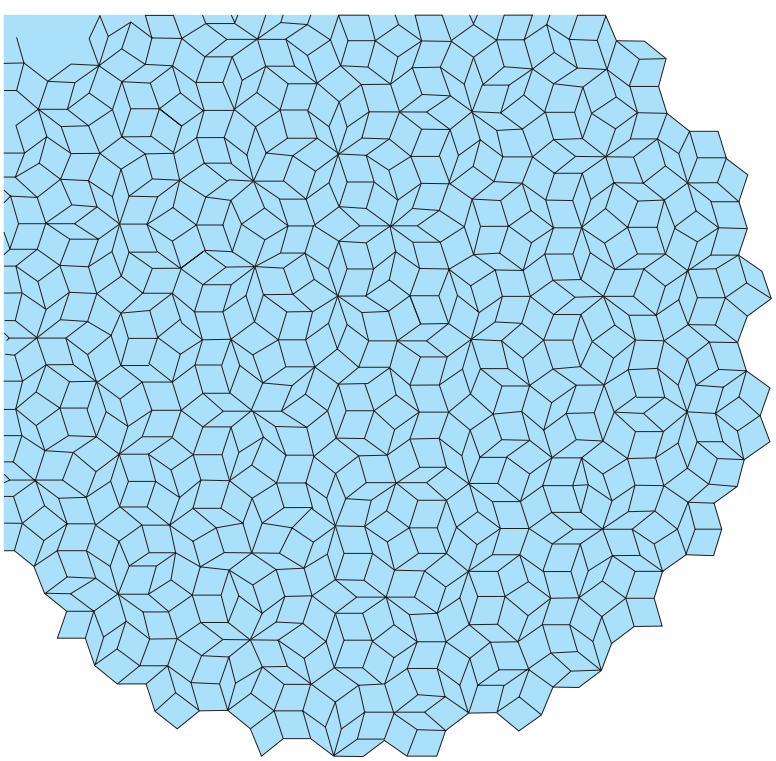

Fig. 4. Penrose tiling generated by projection of the facets of five-dimensional hypercubes in the neighborhood of a two-dimensional cut on an irrational plane. The corners of this quasiperiodic tiling display a tenfold diffraction pattern.

Quasiperiodicity. Strictly periodic structures give sharp diffraction. Deviations from strict periodicity commonly give rise to diffuse scattering. It has long been known and is easily shown theoretically, although not widely appreciated, that there are aperiodic structures that give rise to discrete diffraction patterns with no diffuse scattering. These are the structures whose densities can be described by almost periodic functions. Quasiperiodic functions are special cases of almost periodic functions. *See* FOURIER SERIES AND TRANSFORMS.

Among the interesting properties of quasiperiodic functions that are close to those of periodic functions is the existence, for any value of $\epsilon > 0$, of an infinite set of translation numbers T, shown in Eq. (2), such

$$|f(t+T) - f(t)| < \epsilon \tag{2}$$

that inequality (2) is satisfied for all t. For periodic functions, ϵ can be set equal to zero, and every T is then a multiple of the periodicity, that is, a one-dimensional lattice. For the quasiperiodic functions, the translation numbers T do not form a lattice.

An important concept in the mathematics of quasiperiodic functions is that such a function can be described as a planar cut on an irrational plane of a higher-dimensional periodic function. A periodic crystal cut on an irrational plane gives a two-dimensional quasiperiodic structure, which preserves some of the symmetries in the higher-dimensional crystal. A rational cut of a cubic

structure yields a periodic structure, which, depending on the orientation, could preserve the two-, three-, or fourfold axes. Since an N-dimensional isometric periodic structure always contains an $(N + 1)$-fold axis, quasiperiodic structures can have any integral value of N. The fivefold axes first appear in four-dimensional periodic structures, and the six intersecting fivefold axes needed for icosahedral symmetry first appear in six dimensions. Only irrational cuts of the higher-dimensional space can preserve the symmetries that fall outside those permitted for periodic crystals.

The relationship between a quasiperiodic function and a higher-dimensional periodic function can be easily demonstrated by considering the quasiperiodic function with two incommensurate lengths, 1 and $\sqrt{2}$, given in Eq. (3). The related two-

$$f(x) = \cos x + \cos \sqrt{2}\, x \tag{3}$$

dimensional function is given in Eq. (4). Along the

$$F(x, y) = \cos x + \cos y \tag{4}$$

line of Eq. (5) in the xy plane, the quasiperiodic function is recovered as shown by Eq. (6).

$$y = \sqrt{2}\, x \tag{5}$$

$$f(x) = F(x, \sqrt{2}\, x) \tag{6}$$

Structure determination. Although many methods of crystal structure determination depend heavily on periodicity, those that do not have been adapted for the structure determination of quasicrystals. The methods for analyzing liquid structures have also been used but have not been adapted to depict the order that is present in quasicrystals.

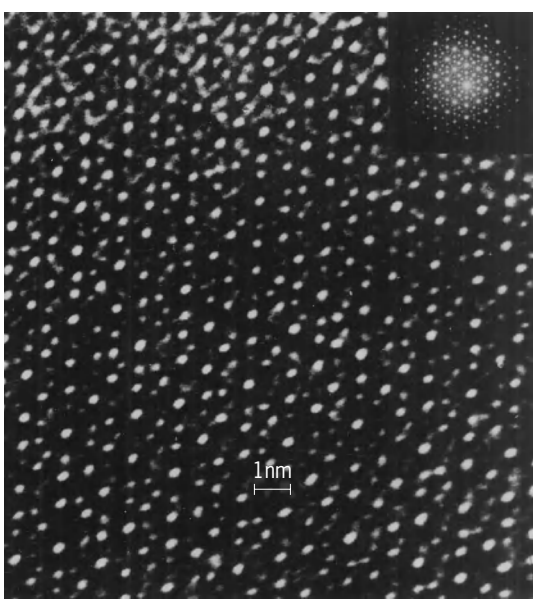

Fig. 5. High-resolution electron micrograph of a thin foil of an icosahedral (aluminum,silicon)-manganese quasicrystal seen in transmission along its fivefold axis. Inset shows electron diffraction pattern. (*From R. Portier et al., High resolution electron microscopy of the icosahedral quasiperiodic structure in Al-Mn system, J. Microsc. Spectrosc. Electron., 10(2):107–116, 1985*)

These aperiodic structures are best described as irrational cuts of periodic higher-dimensional structures that can be determined by standard crystallographic means from diffraction data. The structures examined so far all bear striking similarities to the complicated periodic structures found in intermetallic compounds, often containing hundreds of atoms per unit cell. These unit cells often contain atoms that pack into large icosahedral groupings common for the close packing of spheres of different sizes, but when these groupings pack periodically the fivefold axes are necessarily lost for the crystal as a whole. The quasicrystalline packings preserve these rotational symmetries in the diffraction at the expense of periodicity.

With modern high-resolution electron microscopy, quasicrystals can be seen in transmission (and thus in projection of specimens hundreds of atoms thick) with a resolution that is close to that of the size of atoms. **Figure 5** is an icosahedral example seen along one of the fivefold axes. Such pictures give information about the structure and clearly show the lack of periodicity.

Defects in quasicrystals. Defects can be defined only when the ideal defect-free structure is agreed upon. Not unexpectedly, the study of defects and their role in property determination has given rise to considerable debate, but the array of defects that are emerging is far more varied than those in ordinary periodic crystals. *See* CRYSTAL DEFECTS.

John W. Cahn; Dan Shechtman

Bibliography. C. H. Janot, *Quasicrystals: A Primer*, 2d ed., 1997; M. V. Jaric (ed.), *Introduction to Quasicrystals*, 1988; K. H. Kuo and S. Takeuchi (eds.), *Quasicrystals and Imperfectly Ordered Crystals*, 1994; P. J. Steinhardt and S. Ostlund (eds.), *The Physics of Quasicrystals*, 1987; P. J. Steinhardt and D. Di Vincenzo (eds.), *Quasicrystals: The State of the Art*, 1991.

Quasielastic light scattering

Small frequency shifts or broadening from the frequency of the incident radiation in the light scattered from a liquid, gas, or solid. The term quasielastic arises from the fact that the frequency changes are usually so small that, without instrumentation specifically designed for their detection, they would not be observed and the scattering process would appear to occur with no frequency changes at all, that is, elastically. The technique is used by chemists, biologists, and physicists to study the dynamics of molecules in fluids, mainly liquids and liquid solutions. It is often identified by a variety of other names, the most common of which is dynamic light scattering (DLS).

Several distinct experimental techniques are grouped under the heading of quasielastic light scattering (QELS). Photon correlation spectroscopy (PCS) is the technique most often used to study such systems as macromolecules in solution, colloids, and

critical phenomena where the molecular motions to be studied are rather slow. This technique, also known as intensity fluctuation spectroscopy and, less frequently, optical mixing spectroscopy, is used to measure the dynamical constants of processes with relaxation time scales slower than about 10^{-6} s. For faster processes, dynamical constants are obtained by utilizing techniques known as filter methods, which obtain direct measurements of the frequency changes of the scattered light by utilizing a monochromator or filter much as in Raman spectroscopy. *See* RAMAN EFFECT; SCATTERING OF ELECTROMAGNETIC RADIATION.

Static light scattering. If light is scattered by a collection of scatterers, the scattered intensity at a point far from the scattering volume is the result of interference between the wavelets scattered from each of the scatterers and, consequently, will depend on the relative positions and orientations of the scatterers, the scattering angle θ, the refractive index n of the medium, and the wavelength λ of the light used. The structure of scatterers in solution whose size is comparable to the reciprocal of the scattering vector length, q^{-1}, where $q \equiv [(4\pi n/\lambda) \sin(\theta/2)]$, may be studied by this technique. It is variously known as static light scattering, total-intensity scattering, integrated-intensity light scattering, Rayleigh scattering, or, in the older literature, simply light scattering. It was, in fact, developed in the 1940s and 1950s to measure equilibrium properties of polymers both in solution and in bulk. Molecular weights, radii of gyration, solution virial coefficients, molecular optical anisotropies, and sizes and structure of heterogeneities in bulk polymers are routinely obtained from this type of experiment. Static light scattering is a relatively mature field, although continued improvements in instrumentation (mainly by the use of lasers and associated techniques) are steadily increasing its reliability and range of application.

Both static and quasielastic light-scattering experiments may be performed with the use of polarizers to select the polarizations of both the incident and the scattered beams. The plane containing the incident and scattered beams is called the scattering plane. If an experiment is performed with polarizers selecting both the incident and final polarizations perpendicular to the scattering plane, the scattering is called polarized scattering. If the incident polarization is perpendicular to the scattering plane and the scattered polarization lies in that plane, the scattering is called depolarized scattering. Usually the intensity associated with the polarized scattering is much larger than that associated with the depolarized scattering. The depolarized scattering from relatively small objects is zero unless the scatterer is optically nonspherical.

Photon correlation spectroscopy. The average intensity of light scattered from a system at a given scattering angle depends, as stated above, on the relative positions and orientations of the scatterers. However, molecules are constantly in motion due to thermal forces, and are constantly translating, rotating and, for some molecules, undergoing internal rearrangements. Because of these thermal fluctuations, the scattered light intensity will also fluctuate. The intensity will fluctuate on the same time scale as the molecular motion since they are proportional to each other.

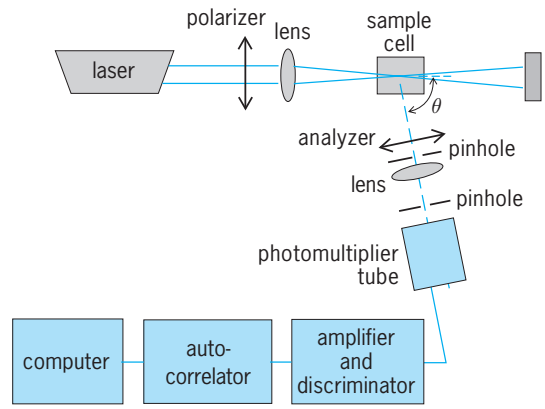

Fig. 1. Schematic diagram of a photon correlation spectroscopy apparatus.

Figure 1 shows a schematic diagram of a typical photon correlation apparatus. Light from a laser source traverses a polarizer to ensure a given polarization. It is then focused on a small volume of the sample cell. Light from the scattering volume at scattering angle θ is passed through an analyzer to select the polarization of the scattered light, and then through pinholes and lenses to the photomultiplier tube. The output of the photomultiplier is amplified, discriminated, and sent to a photon counter and then to a hard-wired computer called an autocorrelator, which computes the time autocorrelation function of the photocounts. The autocorrelator output is then sent to a computer for further data analysis.

The scattered light intensity as a function of time will resemble a noise signal. In order to facilitate interpretation of experimental data in terms of molecular motions, the time correlation function of the scattered intensity is usually computed by the autocorrelator. The autocorrelation function obtained in one of these experiments is often a single exponential decay, $C(t) = \exp(-t/\tau_r)$, where τ_r is the relaxation time.

The lower limit on decay times τ_r that can be measured by photon correlation spectroscopy is about a microsecond, although with special variations of the technique somewhat faster decay times may be measured. For times faster than this, filter experiments are usually performed by using a Fabry-Perot interferometer.

Fabry-Perot interferometry. Light scattered from scatterers which are moving exhibits Doppler shifts or broadening due to the motion. Thus, an initially monochromatic beam of light from a laser will be frequency-broadened by scattering from a liquid, gas, or solid, and the broadening will be a measure of the speed of the motion. For a dilute gas, the spectrum will usually be a gaussian distribution. For a liquid, however, the most common experiment of this type yields a single lorentzian line with its maximum

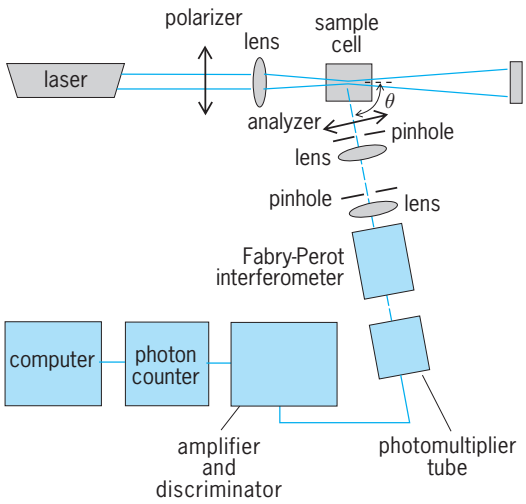

Fig. 2. Typical filter apparatus.

at the laser frequency $I(\omega) = A/\pi[(1/\tau_r)/(\omega^2 + 1/\tau_r^2)]I(\omega) = A/\pi[(1/\tau_r)/(\omega^2 + 1/\tau^2{}_r)]$. **Figure 2** shows a schematic of a typical Fabry-Perot interferometry apparatus. The Fabry-Perot interferometer acts as the monochromator and is placed between the scattering sample and the photomultiplier. Fabry-Perot interferometry measures the (average) scattered intensity as a function of frequency change from the laser frequency. This intensity is the frequency Fourier transform of the time correlation function of the scattered electric field. Photon correlation spectroscopy experiments utilizing an autocorrelator measure the time correlation function of the intensity (which equals the square of the scattered electric field). For scattered fields with gaussian amplitude distributions, the results of these two types of experiment are easily related. Sometimes photon correlation spectroscopy experiments are performed in what is called a heterodyne mode. In this case, some unscattered laser light is mixed with the scattered light on the surface of the photodetector. Photon correlation spectroscopy experiments in the heterodyne mode measure the frequency Fourier transform of the time correlation function of the scattered electric field. *See* INTERFEROMETRY.

Translational diffusion coefficients. The most widespread application of quasielastic light scattering is the measurement of translational diffusion coefficients of macromolecules and particles in solution. For particles in solution whose characteristic dimension R is small compared to q^{-1}, that is, $qR < 1$, it may be shown that the time correlation function measured in a polarized photon correlation spectroscopy experiment is a single exponential with relaxation time $1/\tau_r = 2q^2D$, where D is the particle translational diffusion coefficient. For rigid, spherical particles of any size a photon correlation spectroscopy experiment also provides a measure of the translational diffusion coefficient.

Translational diffusion coefficients of spherical particles in dilute solution may be used to obtain the particle radius R through use of the Stokes-Einstein relation (1), where k_B is Boltzmann's constant, T the

$$D = \frac{k_B T}{6\pi\eta R} \tag{1}$$

absolute temperature, and η the solvent viscosity. If the particles are shaped like ellipsoids of revolution or long rods, relations known, respectively, as the Perrin and Broersma equations may be used to relate the translational diffusion coefficient to particle dimensions. For flexible macromolecules in solution and also for irregularly shaped rigid particles, the Stokes-Einstein relation is often used to define a hydrodynamic radius (R_H).

This technique is routinely used to study such systems as flexible coil macromolecules, proteins, micelles, vesicles, viruses, and latexes. Size changes such as occur, for instance, in protein denaturation may be followed by photon correlation spectroscopy studies of translational diffusion. In addition, the concentration and, in some cases, the ionic strength dependence of D are monitored to yield information on particle interactions and solution structure.

Photon correlation spectroscopy experiments are also used to obtain mutual diffusion coefficients of mixtures of small molecules (for example, benzene–carbon disulfide mixtures) and to measure the behavior of the mutual diffusion coefficient near the critical (consolute) point of binary liquid mixtures. Experiments of this type have proved to be very important in formulating theories of phase transitions.

Rotational diffusion coefficients. Rotational diffusion coefficients are most easily measured by depolarized quasielastic light scattering. The instantaneous depolarized intensity for a nonspherical scatterer depends upon the orientation of the scatterer. Rotation of the scatterer will then modulate the depolarized intensity. In a similar way, the frequency distribution of the depolarized scattered light will be broadened by the rotational motion of the molecules. Thus, for example, for dilute solutions of diffusing cylindrically symmetric scatterers, a depolarized photon correlation spectroscopy experiment will give an exponential intensity time correlation function with the decay constant containing a term dependent on the scatterer rotational diffusion coefficient D_R [Eq. (2)]. A depolarized filter experiment on a similar system will give a single lorentzian with $1/\tau_r$ equal to one-half that given in Eq. (2).

$$\frac{1}{\tau_r} = 2(q^2D + 6D_R) \tag{2}$$

For small molecules (for example, benzene) and relatively small macromolecules (for example, proteins with molecular weight less than 30,000) in solution, filter experiments are used to determine rotational diffusion coefficients. In these cases, the contribution of the translational diffusion to τ_r is negligible. For larger, more slowly rotating macromolecules, depolarized photon correlation spectroscopy experiments are used to determine D_R.

Quasielastic light scattering is the major method of studying the rotation of small molecules in

solution. Studies of the concentration dependence, viscosity dependence, and anisotropy of the molecular rotational diffusion times have been performed on a wide variety of molecules in liquids as well as liquid crystals.

Rotational diffusion coefficients of large ($\geq$100 nanometers) nonspherical particles may also be measured from polarized photon correlation spectroscopy experiments at high values of q. *See* DIFFUSION.

Other applications. There are many variations on quasielastic light scattering experiments. For instance, polarized filter experiments on liquids also give a doublet symmetrically placed about the laser frequency. Known as the Brillouin doublet, it is separated from the incident laser frequency by $\pm C_s q$, where C_s is the hypersonic sound velocity in the scattering medium. Measurement of the doublet spacing then yields sound velocities. This technique is being extensively utilized in the study of bulk polymer systems as well as of simple liquids.

In a variation of the photon correlation spectroscopy technique, a static electric field is imposed upon the sample. If the sample contains charged particles, the molecule will acquire a drift velocity proportional to the electric field strength $v = \mu E$, where μ is known as the electrophoretic mobility. Light scattered from this system will experience a Doppler shift proportional to v. Thus, in addition to particle diffusion coefficients, quasielastic light scattering can be used to measure electrophoretic mobilites. *See* ELECTROPHORESIS.

Quasielastic light scattering may also be used to study fluid flow and motile systems. Photon correlation spectroscopy, for instance, is a widely used technique to study the motility of microorganisms (such as sperm cells); it is also used to study blood flow.

Robert Pecora

Bibliography. L. P. Bayvel and A. R. Jones, *Electromagnetic Scattering and Its Applications*, 1981; B. J. Berne and R. Pecora, *Dynamic Light Scattering with Applications to Chemistry, Biology, and Physics*, 1976, reprint 1990; W. Brown (ed.), *Dynamic Light Scattering: The Method and Some Applications*, 1993; R. Pecora (ed.), *Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy*, 1985.

Quaternary

A period that encompasses at least the last 3×10^6 years of the Cenozoic Era, and is concerned with major worldwide glaciations and their effect on land and sea, on worldwide climate, and on the plants and animals that lived then. The Quaternary is divided into the Pleistocene and Holocene. The term Pleistocene is gradually replacing Quaternary; Holocene involves the last 7000 years since the Pleistocene.

The Quaternary includes four principal glacial stages, each with subdivisions, in the western Rocky Mountains and Sierra Nevada and in the north-central United States, matched by four corresponding stages throughout northern Europe and the Alps. Interglacial stages, periods of mild climate during which ice sheets and alpine glaciers disappeared, intervened between the four main glacial stages. *See* CENOZOIC; HOLOCENE; PLEISTOCENE.

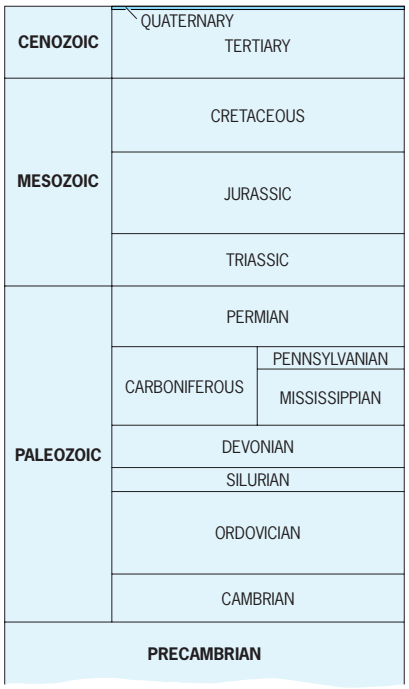

Sidney E. White

Quaternary ammonium salts

Analogs of ammonium salts in which organic radicals have been substituted for all four hydrogens of the original ammonium cation. Substituents may be alkyl, aryl, or aralkyl, or the nitrogen may be part of a ring system. Such compounds are usually prepared by treatment of an amine with an alkylating reagent under suitable conditions. They are typically crystalline solids which are soluble in water and are strong electrolytes. Treatment of the salts with silver oxide, potassium hydroxide, or an ion-exchange resin converts them to quaternary ammonium hydroxides, which are very strong bases, as shown in the reaction below.

Quaternary ammonium salt → Quaternary ammonium hydroxide

A quaternary salt, methyllybenzulphenylammonium iodide, containing four different groups on nitrogen, proved to be resolvable (1899), clearly establishing that nitrogen is tetrahedral in these compounds. In general, optically active quaternary ammonium salts racemize readily, suggesting the

possibility of an equlibrium between the tertiary amine and the alkyl compounds from which it is generated.

Following are some specific examples of useful quaternary ammonium salts and bases: Triton B (**1**)

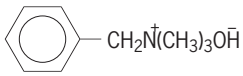

(1)

$CH_3(CH_2)_{15}\,N^+(CH_3)_3Cl^-$

(2)

$HOCH_2CH_2\,N^+(CH_3)_3Cl^-$

(3)

is used in organic synthesis as a basic catalyst for aldo land Michael-type reactions. Benzyltriethylammonium chloride is used as a phase-transfer catalyst in organic synthesis to promote the reaction of water-insoluble reagents in an aqueous medium. Quaternary ammonium salts in which one of the alkyl groups is a long carbon chain, such as hexadecyltrimethylammonium chloride (**2**), are called invert soaps and have useful germicidal properties.

Choline (**3**) is a vitamin, since the human body makes it too slowly to meet all of its needs and a dietary source is required. Choline is a component of complex lipids such as lecithin, and acetylcholine serves in transmitting nerve signals. *d*-Tubocurarine chloride is a complex quaternary ammonium salt isolated from a tropical plant which is used in medicine as a skeletal muscle relaxant. *See* ACETYLCHOLINE.

Other quaternary ammonium salts have found use as water repellants, fungicides, emulsifiers, paper softeners, antistatic agents, and corrosion inhibitors. *See* AMINE; AMMONIUM SALT; SURFACTANT.

Paul E. Fanta

Bibliography. T. S. Eckert, *Organic Chemistry*, 1995; R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; J. C. Stowell, *Intermediate Organic Chemistry*, 2d ed., 1993.

Quaternions

An associative, noncommutative algebra based on four linearly independent units or basal elements. Quaternions were originated in Dublin, Ireland, on October 16, 1843, by W. R. Hamilton (1805–1865), who is famous because of his canonical functions and equations of motion which are important in both classical and quantum dynamics.

The four linearly independent units in quaternion algebra are commonly denoted by 1, i, j, k, where 1 commutes with i, j, k and is called the principal unit or modulus. These four units are assumed to have the multiplication table of Eqs. (1). The i, j, k do not

$$\begin{aligned} 1^2 &= 1 \\ i^2 &= j^2 = k^2 = ijk = -1 \\ i(jk) &= (ij)k = ijk \\ 1i &= i1 \\ 1j &= j1 \\ 1k &= k1 \end{aligned} \tag{1}$$

commute with each other in multiplication, that is, $ij \neq ji$, $jk \neq kj$, $ik \neq ki$, etc. But all real and complex numbers do commute with i, j, k; thus if c is a real number, then $ic = ci$, $jc = cj$, and $kc = ck$. On multiplying $ijk = -1$ on the left by i, so that $iijk = i(-1) = -i$, it is found, since $i^2 = -1$, that $jk = i$. Similarly $jjk = ji = -k$; when exhausted, this process leads to all the simple noncommutative relations for i, j, k, namely, Eqs. (2). More complicated products, for

$$\begin{aligned} ij &= -ji = k \\ jk &= -kj = i \\ ki &= -ik = j \end{aligned} \tag{2}$$

example, $jikjk = -kki = i$, are evaluated by substituting for any adjoined pair the value given in the preceding series of relations and then proceeding similarly to any other adjoined pair in the new product, and so on until the product is reduced to ± 1, $\pm i$, $\pm j$, or $\pm k$. Multiplication on the right is also permissible; thus from $ij = k$, one has $ijj = kj$, or $-i = kj$. Products such as jj and jjj may be written j^2 and j^3.

All the laws and operations of ordinary algebra are assumed to be valid in the definition of quaternion algebra, except the commutative law of multiplication for the units i, j, k. Thus the associative and distributive laws of addition and multiplication apply without restriction throughout. Addition is also commutative, for example, $i + j = j + i$.

Now if s, a, b, c are real numbers, rational or irrational, then a real quaternion q and its conjugate q' are defined by Eqs. (3). In this case $qq' = q'q = s^2 +$

$$\begin{aligned} q &= s + ia + jb + kc \\ q' &= s - (ia + jb + kc) \end{aligned} \tag{3}$$

$a^2 + b^2 + c^2 = N$, and N is called the norm of q; the real quantity $T = \sqrt{N}$ is the tensor of q, and s, a, b, c are components (or coordinates) of q. The part $\gamma = ia + jb + kc$ is the vector of q, and it may be represented by a stroke or vector in a frame of cartesian coordinates, a, b, c being its components. Let now $p = w + ix + jy + kz$ be another real quaternion; if $pq = 0$, either p or q or both are zero, which is called the product law. If, for example, $p = 0$, then all of its components w, x, y, z are zero. When $p = q$, that is, $p - q = 0$, then one must have $w = s$, $x = a$, $y = b$, $z = c$; otherwise Eq. (4) would constitute

$$(w - s) + i(x - a) + j(y - b) + k(z - c) = 0 \tag{4}$$

a linear relation between 1, i, j, k, which would be in conflict with their original definitions as linearly independent.

Multiplication. The product of two quaternions may be found by a straightforward process, and in full is given by Eqs. (5) and (6); and hence

Eq. (7) holds, so that $pq - qp$ is zero only when

$$\begin{aligned} pq = ws &- (ax + by + cz) \\ &+ i(aw + sx + cy - bz) \\ &+ j(bw + sy + az - cx) \\ &+ k(cw + sz + bx - ay) \end{aligned} \quad (5)$$

$$\begin{aligned} qp = ws &- (ax + by + cz) \\ &+ i(aw + sx - cy + bz) \\ &+ j(bw + sy - az + cx) \\ &+ k(cw + sz - bx + ay) \end{aligned} \quad (6)$$

$$\begin{aligned} pq - qp = 2\,[i(cy - bz) &+ j(az - cx) \\ &+ k(bx - ay)] \end{aligned} \quad (7)$$

$cy - bz = az - cx = bx - ay = 0$. This shows that $pq \neq qp$ except under special conditions; quaternion multiplication is not, in general, commutative.

In q, if any two of a, b, c are zero, one has, in effect, an ordinary complex number; if all of a, b, c are zero, then $q = s$ and is an ordinary real number or scalar. Hence real quaternions include the real and ordinary complex numbers as special cases. It will be evident that real quaternions are a kind of extension of the ordinary complex numbers $z = x + \sqrt{-1}y$. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

So far the case in which s, a, b, c are complex quantities has not been included, thus making q a complex quaternion, and for present purposes complex quaternions will be put aside.

It may be noted that the invention of vector analysis was inspired by Hamilton's quaternions; as early as 1846–1852 the Reverend M. O'Brien published papers in which he assumed $i^2 = j^2 = k^2 = 1$, and thus paved the way to the dot or inner product of vector analysis. Fundamentally, quaternion algebra provides much deeper concepts and consequences, and in some practical problems it presents clear advantages over vector analysis. *See* CALCULUS OF VECTORS.

Division. In the division of quaternions, reciprocals of the quaternion units i, j, k are easily found; thus Eqs. (8) hold. More complicated quotients may

$$\begin{aligned} i^{-1} = \frac{1}{i} = \frac{i}{ii} = -i \qquad j^{-1} = -j \\ k^{-1} = -k \qquad kk^{-1} = -k^2 = 1 \end{aligned} \quad (8)$$

be evaluated if care is taken to observe the defined conventions. Thus Eq. (9) holds, but so does Eq. (10).

$$\frac{ij}{j} = ijj^{-1} = i \quad (9)$$

$$\frac{ji}{j} = jij^{-1} = -kj^{-1} = kj = -i \quad (10)$$

It is best to write denominators with negative exponents and place them properly in the numerator to avoid errors. Similarly, the reciprocals and quotients of real quaternions yield unique results. If in $q = s + ia + jb + kc$, all of s, a, b, c are not zero, then Eqs. (11) hold. Further, if $p = w + ix + jy + kz$

$$\begin{aligned} q^{-1} = \frac{1}{q} = \frac{q'}{qq'} = \frac{q'}{N} \\ N = s^2 + a^2 + b^2 + c^2 \end{aligned} \quad (11)$$

is a second real quaternion, then Eq. (12) holds.

$$\frac{p}{q} = \frac{pq'}{qq'} = \frac{pq'}{N} \quad (12)$$

Accordingly, real quaternions admit of division. If $rq = p$, then $r = p/q = pq'/N$; and if $qr = p$, $r = q^{-1}p = q'p/N$. Hence both right division and left division yield unique quotients. If, for example, $s = a = 1$, $b = \sqrt{-2}$, $c = 0$, then $N = 1 + 1 - 2 + 0 = 0$; this result is one reason why the discussion has been limited to real quaternions.

Hamilton adopted the name vectors for directed lines in space. If vectors are denoted with Greek letters, then $\alpha = ix + jy + kz$ is a vector whose components along a conventional right-handed rectangular cartesian coordinate system are x, y, z, respectively. If $\beta = ix' + jy' + kz'$ is another vector in the same coordinate system, their products are given by Eqs. (13) and (14). Both products yield a scalar u

$$\begin{aligned} \alpha\beta = -(xx' + yy' + zz') &+ i(yz' - y'z) \\ &+ j(x'z - xz') + k(xy' - x'y) = -u + \gamma \end{aligned} \quad (13)$$

$$\begin{aligned} \beta\alpha = -(xx' + yy' + zz') &- [i(yz' - y'z) \\ &+ j(x'z - xz') + k(xy' - x'y)] = -u - \gamma \end{aligned} \quad (14)$$

and a vector γ as a sum, and such a sum is by definition a quaternion. Further, Eq. (15) holds. That is, in

$$\alpha\beta + \beta\alpha = -2(xx' + yy' + zz') \quad (15)$$

general, multiplication of nonparallel vectors is not commutative, which is a special case of the noncommutative multiplication of quaternions. If one sets $x = x'$, $y = y'$, $z = z'$, then Eq. (16) results, which,

$$\alpha^2 = -(x^2 + y^2 + z^2) \quad (16)$$

with negative sign, is the square of the length of the vector α whose components are x, y, z. The quotient α/β emerges easily since Eq. (17) is valid. Hence α/β

$$\frac{\alpha}{\beta} = \frac{-\alpha\beta}{(x'^2 + y'^2 + z'^2)} \quad (17)$$

has a unique value if x', y', z' are all real and not all zero. The quotient of two vectors is a quaternion, and previously it was stated that a quaternion can be defined as the ratio of two vectors.

By multiplication a real q and its conjugate q' are found to commute, as in Eq. (18),to give a real posi-

$$qq' = q'q = s^2 + a^2 + b^2 + c^2 = T^2 \quad (18)$$

tive scalar. Further, $q + q' = 2s$, and hence q and q' are the roots of a quadratic equation (19) with real

$$t^2 - 2st + T^2 = 0 \quad (19)$$

coefficients. When this equation is solved in the field

of ordinary complex numbers, Eq. (20) results. This

$$t = s \pm \sqrt{-(a^2 + b^2 + c^2)} \tag{20}$$

simple but important result emphasizes the fact that, in asking for a solution of a given algebraic equation, the field of the quantities for which a solution is desired must be specified.

Applications. Next, let $r = s' + ia' + jb' + kc'$ be another real quaternion, then Eq. (21) holds, where $S = ss' - (aa' + bb' + cc')$, $A = as' + a's - bc' - b'c$, etc. for B and C. Also, Eqs. (22) are valid. But

$$\begin{aligned} qr &= ss' - (aa' + bb' + cc') \\ &\quad + i(as' + a's + bc' - b'c) \\ &\quad + j(bs' + b's + a'c - ac') \\ &\quad + k(cs' + c's + ab' - a'b) \\ &= S + iA + jB + kC \end{aligned} \tag{21}$$

$$\begin{aligned} r'q' &= S - (iA + jB + kC) = (qr)' \\ qr(qr)' &= S^2 + A^2 + B^2 + C^2 \end{aligned} \tag{22}$$

$qr(qr)' = qrr'q' = qq'rr'$, since rr' is a scalar and commutes with q and q'. Therefore, one has Euler's famous result, Eq. (23), which is important in number

$$\begin{aligned} S^2 + A^2 + B^2 + C^2 &= (s^2 + a^2 + b^2 + c^2) \\ &\quad \times (s'^2 + a'^2 + b'^2 + c'^2) \end{aligned} \tag{23}$$

theory since it is used in the proof that every positive rational integer can be represented as a sum of four squares. *See* NUMBER THEORY.

The quotient of real vectors α/β is the sum of a scalar and vector and hence is a quaternion, for example, q. If l_1, m_1, n_1 are the direction cosines of α in a rectangular cartesian coordinate frame, and l_2, m_2, n_2 are the direction cosines of β, then α and β can be expressed by Eqs. (24) and (25), and q is given by Eq. (26). Then $\alpha = q\beta$; that is, q is an op-

$$\begin{aligned} \alpha &= a_0(il_1 + jm_1 + kn_1) \\ a_0 &= \sqrt{x^2 + y^2 + z^2} \end{aligned} \tag{24}$$

$$\begin{aligned} \beta &= b_0(il_2 + jm_2 + kn_2) \\ b_0 &= \sqrt{x'^2 + y'^2 + z'^2} \end{aligned} \tag{25}$$

$$\begin{aligned} q = \frac{\alpha}{\beta} = \frac{\alpha_0}{b_0} &= [l_1l_2 + m_1m_2 + n_1n_2 \\ &\quad + i(m_2n_1 - m_1n_2) + j(l_1n_2 - l_2n_1) \\ &\quad + k(l_2m_1 - l_1m_2)] \\ &= s + ia + jb + kc = s + \gamma \end{aligned} \tag{26}$$

erator which turns and stretches β to coincide with α. From analytical geometry $l_1l_2 + m_1m_2 + n_1n_2 = \cos\omega$, where ω is the angle ($<\pi$) between vectors α and β. Note that Eq. (27) holds, hence the vector

$$\begin{aligned} l_1(m_2n_1 - m_1n_2) + m_1(l_1n_2 - l_2n_1) \\ + n_1(l_2m_1 - l_1m_2) = 0 \end{aligned} \tag{27}$$

part γ of q is perpendicular to α, and the same is true for γ and β. This suggests the relationship of Eq. (28), where $a_0^2/b_0^2 = s^2 + a^2 + b^2 + c^2 = T^2$,

$$q = \frac{a_0}{b_0}(\cos\omega + \epsilon\sin\omega) = TUq \tag{28}$$

$\cos\omega = s/T$, and ϵ is the unit vector $\epsilon = (ia + jb + kc)/\sqrt{a^2 + b^2 + c^2}$, which is perpendicular to both α and β. Because $\epsilon^2 = -1$, the square of any real unit vector is -1. The factor $Uq = \cos\omega + \epsilon\sin\omega$ turns β through the angle ω to coincide with the direction of α and is called the versor of q. U is one of the symbols encountered in the grammar of earlier quaternion theory. If $U'q = \cos\omega - \epsilon\sin\omega$, then $UqU'q = 1$.

There is an important result in quaternion algebra which seems to be due to A. Cayley. With q as before, then its reciprocal is $q^{-1} = q'/T^2$, as before. If some other real quaternion is $p = w + ix + jy + kz$, then the expression qpq^{-1} is called the conical rotation (or transform) of p, and it finds important application in the specification of motions of rigid bodies. One has Eq. (29), where $T^2 = s^2 + a^2 + b^2 + c^2$. If $T^2 = 1$, $w = w$ and X, Y, and Z are given by Eqs. (30), which is a linear transformation of X, Y, Z to x, y, z.

$$\begin{aligned} qpq^{-1} &= \{wT^2 + i[(s^2 + a^2 - b^2 - c^2)x \\ &\quad + 2(ab - cs)y + 2(ac + bs)z] + j[2(ab + cs)x \\ &\quad + (s^2 - a^2 + b^2 - c^2)y + 2(bc - as)z] \\ &\quad + k[2(ac - bs)x + 2(as + bc)y \\ &\quad + (s^2 - a^2 - b^2 + c^2)z]\}/T^2 \\ &= w + iX + jY + kZ = w + \tau \end{aligned} \tag{29}$$

$$\begin{aligned} X &= (s^2 + a^2 - b^2 - c^2)x + 2(ab - cs)y \\ &\qquad + 2(ac + bs)z \\ Y &= 2(ab + cs)x + (s^2 - a^2 + b^2 - c^2)y \\ &\qquad + 2(bc - as)z \\ Z &= 2(ac - bs)x + 2(as + bc)y \\ &\qquad + (s^2 - a^2 - b^2 + c^2)z \end{aligned} \tag{30}$$

It is easy to show that, for example, Eqs. (31)

$$\begin{aligned} (s^2 + a^2 - b^2 - c^2)^2 + 4(ab - cs)^2 + 4(ac + bx)^2 \\ = T^4 = 1 \\ 2(ab + cs)(s^2 + a^2 - b^2 - c^2) \\ + 2(s^2 - a^2 + b^2 - c^2)(ab - cs) \\ + 4(bc - as)(ac + bs) = 0 \end{aligned} \tag{31}$$

hold. Hence the coefficients of x, y, z are direction cosines of the frame X, Y, Z in the x, y, z frame of coordinates, and the transformation is orthogonal (unitary). The a, b, c, s are Euler's parameters (ξ, η, ζ, χ). Clearly the angle ψ between $\alpha = ix + jy + kz$ and $\gamma = ia + jb + kc$ is given by $g(ax + by + kz) = \cos\psi$, with g a known quantity, and some easy algebra shows that the angle between the vector part τ of qpq^{-1} and γ of $q = s + \gamma$ is given by the same expression. Hence qpq^{-1} has rotated α conically about γ, and the magnitude of the rotation comes out, remarkably enough, to be just 2ω, where

again $\cos\omega = s/T = s$. For example, if Eqs. (32) hold,

$$a = b = w = 0 \quad s = \cos\omega \quad c = \sin\omega \tag{32}$$

so that $q = \cos\omega + k\sin\omega$; then the formulas above yield easily Eqs. (33), which shows that p has been

$$\begin{aligned} X &= x\cos 2\omega - y\sin 2\omega \\ Y &= x\sin 2\omega + y\cos 2\omega \\ Z &= z \end{aligned} \tag{33}$$

rotated conically about z and through the angle 2ω.

If λ, μ, ν are the angles which the vector $\gamma = ia + jb + kc$ makes with x, y, z, respectively, and if $T^2 = 1$, then the relations $a = \cos\lambda\sin\omega$, $b = \cos\mu\sin\omega$, $c = \cos\nu\sin\omega$, $s = \cos\omega$, and $a^2 + b^2 + c^2 = \sin^2\omega$ are consistent with the above analysis. Another formulation for a, b, c, s is given by Eqs. (34). The angles ω, ϕ, ψ are the familiar eule-

$$\begin{aligned} a &= \sin\frac{\theta}{2}\sin\frac{1}{2}(\psi - \phi) \\ b &= \sin\frac{\theta}{2}\cos\frac{1}{2}(\psi - \phi) \\ c &= \cos\frac{\theta}{2}\sin\frac{1}{2}(\psi + \phi) \\ s &= \cos\frac{\theta}{2}\cos\frac{1}{2}(\psi + \phi) \end{aligned} \tag{34}$$

rian angles relating a fixed cartesian frame to another with the same origin, the second being assumed, in kinematics, to be fixed in a rigid body moving about a fixed point, namely, the common origin of the two frames.

If α, β, γ, δ are defined by $\alpha = s + ic$, $\alpha = a + ib$, $\gamma = -a + ib$, $\delta = s - ic$, $i = \sqrt{-1}$, then in terms of the eulerian angles, they are given by Eqs. (35), and Eq. (36) holds. These are the Cayley-

$$\begin{aligned} \alpha &= \cos\frac{\theta}{2}e^{i(\phi-\psi)/2} \qquad \beta = i\sin\frac{\theta}{2}e^{i(\phi-\psi)/2} \\ \gamma &= i\sin\frac{\theta}{2}e^{i(\psi-\phi)/2} \qquad \delta = \cos\frac{\theta}{2}e^{i(-\psi-\phi)/2} \\ i &= \sqrt{-1} \end{aligned} \tag{35}$$

$$\alpha\delta - \beta\gamma = s^2 + a^2 + b^2 + c^2 = 1 \tag{36}$$

Klein parameters which lend themselves to homographic (bilinear) transformation, $z' = \alpha z + \beta/\gamma z + \delta$, in the complex z plane so that the motion of a solid body in space can be represented on a plane. The quantities $u = -i\gamma$, $v = i\delta$ are also components of a unit spinor, a function occurring in higher algebra and in the quantum theory (similarly for α, β). If $a_x = (1/2)(u^2 - v^2)$, $a_y = (-i/2)(u^2 + v^2)$, $a_z = uv$ are the components of a complex vector, then $a_x^2 + a_y^2 + a_z^2 = 0$ and the vector is of length zero. Such a vector is the starting point for some treatments of spinor theory. *See* CAYLEY-KLEIN PARAMETERS; EULER ANGLES.

Returning now to the quaternion units 1, i, j, k, they may, as appears to have been first discovered by Sylvester, be represented by various matrices, one 2×2 set being given by Eqs. (37), and the last three,

$$\begin{aligned} 1 &\rightleftharpoons \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad i \rightleftharpoons \begin{bmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{bmatrix} \\ j &\rightleftharpoons \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad k \rightleftharpoons \begin{bmatrix} 0 & \sqrt{-1} \\ \sqrt{-1} & 0 \end{bmatrix} \end{aligned} \tag{37}$$

when multiplied by $\sqrt{-1}$, are the Pauli spin matrices occurring in the quantum theories of electron spin (spin $= {}^1/_2$). It has been shown by S. Bochner that no set of 3×3 matrices have the multiplication table corresponding to 1, i, j, k. A set of 4×4 matrices, due to A. S. Eddington, which represents 1, i, j, k is given by Eqs. (38). These are four out of a group of

$$\begin{aligned} 1 &\rightleftharpoons \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad i \rightleftharpoons \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} \\ j &\rightleftharpoons \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad k \rightleftharpoons \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \end{aligned} \tag{38}$$

sixteen 4×4 matrices used by Eddington in his fundamental theory; five of the sixteen, not including the above, are, when multiplied by $\sqrt{-1}$, the matrices occurring in Dirac's theory of the relativistic wave equation for the electron. Another set of four matrices for 1, i, j, k is given by Eqs. (39), and if, by

$$\begin{aligned} 1 &\rightleftharpoons \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad i \rightleftharpoons \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \\ I \\ j &\rightleftharpoons \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} \quad k \rightleftharpoons \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \end{aligned} \tag{39}$$

ordinary matrix addition rule, $x + ix, + jx_2 + kx_3 = q$ are formed, then Eq. (40) holds, and the determinant of q, $|q|$, has the form of Eq. (41),which is an

$$q = \begin{bmatrix} x & x_1 & x_2 & x_3 \\ -x_1 & x & -x_3 & x_2 \\ -x_2 & x_3 & x & -x_1 \\ -x_3 & -x_2 & x_1 & x \end{bmatrix} \tag{40}$$

$$|q| = (x^2 + x_1{}^2 + x_2{}^2 + x_3{}^2)^2 \tag{41}$$

intriguing result. *See* MATRIX THEORY; RELATIVISTIC QUANTUM THEORY.

The basal quaternion elements 1, i, j, k, and their negatives, are the elements of a non-Abelian group of order eight, which is called the quaternion group. The group contains proper, self-conjugate (invariant) cyclical subgroups, one of the order 1 and, significantly, three of order 4; and there are five conjugate classes, three of which are of order 2 and the others each of order 1. *See* GROUP THEORY.

Hamilton's interest in analytical geometry led to many applications of quaternions in this field,

especially to conics and quadric surfaces. In his hands and especially in those of P. G. Tait, P. Kelland, J. McAulay, C. J. Joly, and A. S. Hardy there have been many applications of quaternions to classical geometry and mathematical physics. These include the use of Hamilton's partial differential operator, Eq. (42).

$$\nabla = i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z} \tag{42}$$

The great advances made in quantum theory, relativity, number theory, algebra, and group theory are associated with scholars who had an easy familiarity with quaternions. Don M. Yost

Bibliography. S. L. Adler, *Quaternionic Quantum Mechanics and Quantum Fields*, 1995; S. L. Altmann, *Rotations, Quaternions, and Double Groups*, 1986; E. Cartan, *The Theory of Spinors*, 1937, reprint 1980; H. Goldstein, C. E. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., 2002; W. R. Hamilton, *Elements of Quaternions*, 2 vols., 3d ed., ed. by C. J. Joly, 1901, reprint 1969; G. H. Hardy and E. M. Wright, *Introduction to the Theory of Numbers*, 5th ed., 1980; H. Jeffreys and B. Jeffreys, *Methods of Mathematical Physics*, 3d ed., 1972; L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory*, 3d ed., 1997; J. C. Slater, *Quantum Theory of Molecules and Solids*, 4 vols., 1963, 1965, 1967, 1974.

Quebracho

Any of a number of trees (see **illus.**) belonging to different genera but having similar qualities, all indigenous to South America and valuable for both wood and bark. The heartwood of one South American tree, *Schinopsis lorentzii* (family Anacardiaceae), is called quebracho (meaning axbreaker) in reference to the exceedingly hard wood, one of the hardest known. Quebracho is the world's most important source of tannin. The logs are chipped and the chips are cooked with steam in copper extractors. The liquor becomes highly concentrated with 40–60% of tannin. It is used either alone or in combination with other chemicals for tanning all kinds of leather. *See* SAPINDALES. Perry D. Strausbaugh; Earl L. Core

Quebracho, branch with fruit.

Queueing theory

The mathematical theory of the formation and behavior of queues or waiting lines. The name is also applied loosely to the mathematical study of a wide variety of problems connected with traffic congestion and storage systems. Uneven flow through a service point, with fluctuating arrivals and service times, constitutes a major topic of operations research, and some current work in queueing theory goes under that title. For the mathematician, queueing theory is particularly interesting because it is concerned with relatively simple stochastic processes, which are in general nonmarkovian and possibly stationary. *See* OPERATIONS RESEARCH; STOCHASTIC PROCESS.

Origin. The principal pioneer of queueing theory was A. K. Erlang, who began in 1908 to study problems of telephone congestion for the Copenhagen Telephone Company. He was concerned with problems such as the following: A manually operated telephone exchange has a limited number (one or more) of operators. When a subscriber attempts to make a call, the subscriber must wait if all the operators are already busy making connections for other subscribers. It is of interest to study the waiting times of subscribers—for example, the average waiting time and the chance that a subscriber will obtain service immediately without waiting—and to examine how much the waiting times will be affected if the number of operators is altered, or conditions are changed in any other way. If there are more operators or if service can be speeded up, subscribers will be pleased because waiting will be reduced, but the improved facility will be more expensive to maintain; therefore, a reasonable balance must be struck.

Related problems in the use of automatic telephone exchanges and of long-distance lines able to carry only a limited number of messages simultaneously have resulted in much mathematical study of telephone traffic problems since Erlang's time. His name is commemorated by telephone engineers in the unit of traffic intensity of a channel, the erlang, which may be defined as the number of requests for service (during some period of time) that are actually made, divided by the number of requests that could have been satisfied if the channel had been used to full capacity the whole time.

Similar problems arise in other contexts. In a factory a number of machines, such as looms, may be under the care of one or more repair person. If a machine breaks down, it must stand idle until a repair person is free from repairing other machines. Machines here correspond to telephone subscribers, breakdown corresponds to attempts to make a call, and repair corresponds to connection. Other examples of congestion situations, mathematically similar or identical to the foregoing, are aircraft flying around in circles waiting to use an airport landing strip, automobiles lining up at a turnpike toll booth,

and (perhaps the most familiar of all) customers lining up at the counter of a retail shop, waiting for service. Much of the literature of queueing theory uses terminology appropriate to the last example.

Theory. One of the simplest portions of queueing theory is the following presentation. Consider customers arriving at a shop counter. Suppose that they arrive singly in a purely random or haphazard fashion, specified by the following conditions: (1) The average number of customers arriving per unit time is constant, signified by λ; (2) the numbers of customers arriving during any two nonoverlapping time intervals are independent; or (3) the chance that a customer will arrive during any specified time interval of infinitesimal duration dt is $\lambda\ dt$. Expressed more exactly, condition (3) means that the chance of one new arrival during any short time interval of length δt is $\lambda\ \delta t + o(\delta t)$, and the chance of more than one new arrival is $o(\delta t)$, as $\delta t \to 0$. Events (arrivals) having this character are said to constitute a Poisson process (stochastic process). Suppose that there is just one server, and if that server is occupied in helping a customer and more customers arrive, the latter will line up to await their turn for service in the order of their arrival. Suppose that the time scale has been so chosen that the average duration of service of a customer is just 1 unit. It will be assumed that customers do not all require the same time for service but that their service times have what is known as an exponential distribution, such that the chance that a customer's service time will exceed any particular duration t is equal to e^{-t}. This service time distribution has the remarkable property that whether or not a customer now being served will have his or her service completed during the next infinitesimal time interval of duration dt is independent of how long the customer has already been served; the chance of completion is just dt. It is assumed that the service times of successive customers are independent and that the number of persons waiting has no effect on the speed of service. It may be noted that, because the mean service time of a customer is equal to 1 unit, λ is equal to the traffic intensity, measured in erlangs.

The queue length L can be defined as the total number of customers at the counter, so that $L = 0$ if there are no customers, and someone is waiting for service if $L = 2$; and T denotes the length of time that a customer has to wait before his service begins. Let $f_r(t)$ denote the chance that $L = r$ at time t. During the infinitesimal time interval $(t, t + dt)$ there is the chance $\lambda\ dt$ that a new customer will arrive and, provided that $L = 1$ at time t, there is the chance dt that the customer at the head of the queue will be served and depart. Hence, by comparing the states of the queue at times t and $t + dt$, it is easy to show that $f_r(t)$ satisfies the system of differential-difference equations (1).

$$\begin{aligned} f_0'(t) &= f_1(t) - \lambda f_0(t) \\ f_r'(t) &= f_{r+1}(t) - (1+\lambda) f_r(t) + \lambda f_{r-1}(t) \\ &\qquad (r \geq 1) \end{aligned} \tag{1}$$

It can be proved that if $\lambda < 1$, the process approaches an equilibrium or steady state, in which the chances $f_r(t)$ are independent of t, and the derivatives on the left-hand sides of Eqs. (1) vanish. The equilibrium values are easily found as shown by Eq. (2).

$$f_r = (1 - \lambda)\lambda^r \qquad (r \geq 0) \tag{2}$$

In particular, the chance that at any arbitrary instant the server is idle is shown by Eq. (3). For the

$$f_0 = 1 - \lambda \tag{3}$$

mean queue length one finds that Eq. (4) holds.

$$\varepsilon(L) = \sum_{r=0}^{\infty} r f_r = \frac{\lambda}{1 - \lambda} \tag{4}$$

Now consider a newcomer to the queue. If there are L persons in the queue the instant before the newcomer arrives, the expected service time for each of these (including the one at the head) is 1, and so $\epsilon(T) = \epsilon(L)$. Thus, the mean waiting time is shown by Eq. (5). This will be large if λ is close to, but less

$$\varepsilon(T) = \frac{\lambda}{1 - \lambda} \tag{5}$$

than, 1. If $\lambda > 1$, it is obvious that the mean queue length increases indefinitely and no equilibrium is reached, and it can be shown that the same is true if $\lambda = 1$.

It has been assumed that the service times follow an exponential distribution. If instead it is assumed that every customer requires exactly the same time for service, namely 1 unit, a somewhat similar argument would lead to the result that, for equilibrium with $\lambda < 1$, Eq. (6) holds, which is just one-half the

$$\varepsilon(T) = \frac{\lambda}{2(1 - \lambda)} \tag{6}$$

previous result. Both results are included in a formula derived by F. Polaczek (1930), as follows. If the service time distribution has mean equal to 1 and variance v and if all the other assumptions previously made are satisfied, then the mean waiting time of a customer, in the equilibrium condition for $\lambda < 1$, is shown by Eq. (7). This will tend to be large if either

$$\varepsilon(T) = \frac{(1 + v)\lambda}{2(1 - \lambda)} \tag{7}$$

λ is close to 1 or v is large (service times are very variable).

Applications. Many variants of the above problem have been studied. Arrivals need not constitute a Poisson process. For example, the arrival times of patients at a doctor's office can show a more even spacing than a Poisson process if they have been given equally spaced appointments, but their arrival times will not be exactly equally spaced. The rule of queue discipline, "first come, first served," will not always apply. Some customers, for example, may enjoy priority and move ahead of nonpriority customers. There may be more than one server, and if

so, separate queues may be formed for each. There are then various possible rules to be considered for assigning a newcomer to a queue. It may happen that the frequency of arrivals or speed of service varies with the queue length; for example, if the queue is too long, new customers may be turned away. Compound queueing systems are formed when successive services have to be rendered to the same customers.

What is most interesting to investigate varies with the circumstances. Sometimes it is the mean waiting time of customers, sometimes the frequency with which the queue length exceeds a given limit, sometimes the proportion of the servers' time that is idle, and sometimes the average duration of a period during which a server is continuously occupied. In the study of stocking a warehouse or retail shop, known generally as the theory of inventories, the frequency with which the stock will be exhausted is considered under various reordering policies. Similar considerations apply in the theory of dams and water storage. *See* LINEAR PROGRAMMING; SYSTEMS ENGINEERING. Francis J. Anscombe

Bibliography. R. B. Cooper, *Introduction to Queueing Theory*, 2d ed., 1981, reprint 1990; J. W. Daigle, *Queueing Theory for Telecommunications*, 1992; B. Gnedenko and I. N. Kovalenko, *Introduction to Queueing Theory*, 2d ed., 1989; D. Gross and C. M. Harris, *Fundamentals of Queueing Theory*, 3d ed., 1997; L. Kleinrock, *Queueing Systems*, 2 vols., 1975, 1976.

Quince

The tree *Cydonia oblonga*, originally from Asia, grown for its edible fruit. The deciduous, crooked-branched tree attains a height of about 20 ft (6 m). It is cultivated in either bush or tree form. The undersides of the leaves are densely tomentose (hairy); the solitary flowers (**illus.** *a*), up to 2 in. (5 cm) across, are snowy white or pale pink; the fruit is a pear-shaped or apple-shaped pome (illus. *b*), characteristically tomentose, up to 3 in. (7.5 cm) in diameter, aromatic, sour, astringent, and green, turning clear yellow at maturity. Used mostly for jam and jelly or as a stewed fruit, the fruit of the quince develops a pink color in cooking. *See* DECIDUOUS PLANTS; ROSALES.

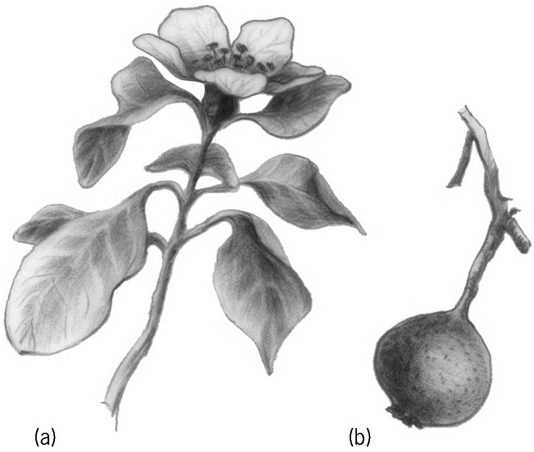

Quince. (*a*) Flower is borne on wood of the same season, not from an autumn fruit bud. (*b*) Short shoots start from the winter buds, and the flowers and fruits are produced on the ends of these shoots.

The quince is propagated on hardwood cuttings or by budding on quince seedlings. It may be used as stock for dwarfing the pear. The tree is only slightly hardier than the peach, the wood being severely injured at −15 to −20°F (−26 to −29°C), and is subject to fire blight (*Bacillus amylovorus*), which is controlled by removal of diseased twigs and avoidance of overly vigorous growth. Trees are planted 15–20 ft (4.5–6 m) apart. Principal varieties are Orange, Champion, and Van Deman. Fruits keep for only a few weeks. There is but little commercial production of this fruit in North America. The dwarf Japanese quince (*Chaenomeles japonica*), with orange-scarlet flowers, is grown as an ornamental shrub. *See* BACTERIA; FRUIT, TREE; PEACH; PEAR; PLANT PATHOLOGY; PLANT PROPAGATION. Harold B. Tukey

Quinine

The chief alkaloid of the bark of the cinchona tree, which is indigenous to certain regions of South America. The structure of quinine is shown below.

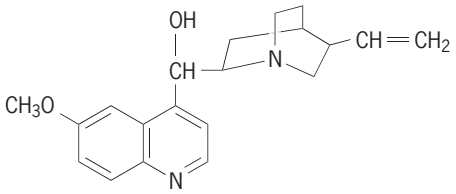

The most important use of quinine has been in the treatment of malaria.

For almost two centuries cinchona bark was employed in medicine as a powder or extract. In 1820 P. Pelletier and J. Caventou isolated quinine and related alkaloids from cinchona bark, and the use of the alkaloids as such rapidly gained favor. Major credit is due to P. Rabe for the postulation of the correct structure of quinine. The difficult laboratory synthesis of quinine by R. Woodward and W. Doering in 1944, although economically unfeasible, corroborated Rabe's structure.

Until the 1920s quinine was the best chemotherapeutic agent for the treatment of malaria. Clinical studies definitely established the superiority of the synthetic antimalarials such as primaquine, chloroquine, and chloroguanide. *See* ALKALOID; CINCHONA; MALARIA. S. Morris Kupchan

Quinoa

An annual herb, *Chenopodium quinoa* (family Chenopodiaceae), 4–6 ft (1.2–1.8 m) tall, a native of Peru, and the staple food of many people in South

America. These plants, grown at high altitudes, produce large quantities of highly nutritious seeds used whole in soups or ground into flour, which is made into bread or cakes. The seeds are also used as poultry feed, in medicine, and in making beer. The ash is mixed with coca leaves to flavor them as a masticatory. In the United States the leaves are sometimes used as a substitute for spinach. *See* CARYOPHYLLALES; SPICE AND FLAVORING.

Perry D. Strausbaugh; Earl L. Core

Quinone

One of a class of aromatic diketones in which the carbon atoms of the carbonyl groups are part of the ring structure. The name quinone is applied to the whole group, but it is often used specifically to refer to *p*-benzoquinone (**1**). *o*-Benzoquinone (**2**) is

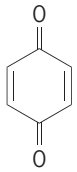

(**1**)

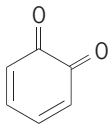

(**2**)

also known but the meta isomer does not exist.

Preparation. Quinones are prepared by oxidation of the corresponding aromatic ring systems containing amino (—NH_2) or hydroxyl (—OH) groups on one or both of the carbon atoms being converted to the carbonyl group. *p*-Benzoquinone is prepared by the oxidation of aniline with manganese dioxide, MnO_2, in the presence of sulfuric acid, H_2SO_4. Oxidation of phenol, *p*-aminophenol, hydroquinone, or *p*-phenylenediamine will also produce *p*-bezoquinone. *o*-Benzoquinone is prepared by oxidation of catechol with silver oxide, Ag_2O, in the absence of water. This quinone is much less stable and more reactive than the para isomer.

Three of the several quinones derived from naphthalene are known: 1,4-napthoquinone (**3**), 1,2-naphthoquinone, and 2,6-naphthoquinone (**4**).

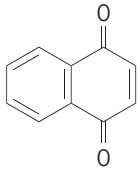

(**3**)

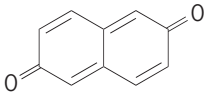

(**4**)

The naphthoquinones are prepared by oxidation of the corresponding aminoaphthols. 9,10-Anthraquinone (**5**) is best prepared, as in reaction (1),

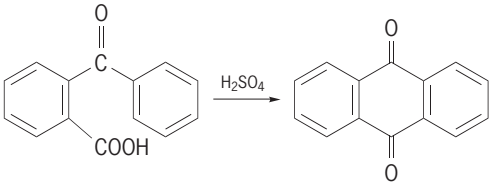

(1)

(**5**)

by dehydration of *o*-benzoylbenzoic acid which is prepared from Friedel-Crafts reaction of benzene and phthalic anhydride. Direct oxidation of phenanthrene with chromic acid yields 9,10-phenanthraquinone (**6**), the further oxidation of which

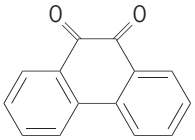

(**6**)

gives diphenyl-2,2′-dicarboxylic acid (diphenic acid). *See* FRIEDEL-CRAFTS REACTION.

Reactions. *p*-Benzoquinone is easily reduced, as in reaction (2), to hydroquinone (**7**) by a variety of

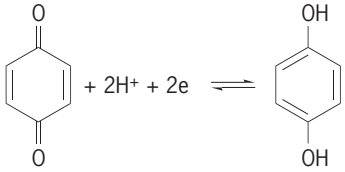

(2)

(**7**)

reagents. Reaction (2) is reversible, and the position of equilibrium can be made to depend on hydrogen-ion concentration and applied electrical potential.

This system (E_0 = 0.699 volt) has been useful for the measurement of hydrogen-ion concentration. The E_0 values for many other quinone-hydroquinone systems have been measured. An intermediate in the reduction of *p*-benzoquinone or in the oxidation of hydroquinone is quinhydrone, a 1:1 molecular complex of these two substances. *See* HYDROGEN ION.

The most characteristic reactions of para quinones are those of the carbon to carbon double bonds and of the conjugated system

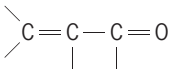

Reaction as a dienophile in the Diels-Alder process, reaction (3) [where *p*-benzoquinonel (**8**) and 1,3-butadiene (**9**) react to give 5,8,9,10-tetrahydro-1,4-napthoquinone (**10**)], is quite general and oc-

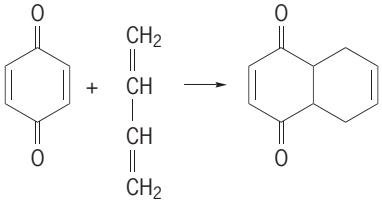

(3)

(**8**) (**9**) (**10**)

curs under mild conditions. The remaining \css\ bond of the quinone ring may also react in the same way. Halogen adds normally to the \css\ bond as in alkenes. Hydrogen halide, however, adds to the conjugated system

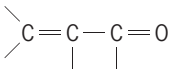

by 1,4 addition, and this is followed by enolization to a hydroquinone derivate, reaction (4). Malonic es-

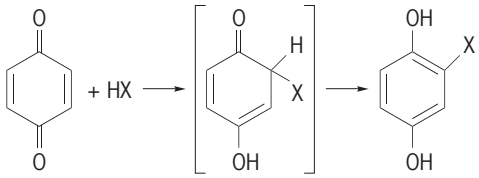

(4)

ter and acetoacetic ester react similarly with utilization of their active hydrogen atoms. Mercaptans and Grignard reagents give mixtures of normal adducts to the C═O group and 1,4 additions to the system. Simple quinones do not often undergo substitution reactions with the electrophilic reagents commonly used for aromatic systems. Free radicals from decomposition of acyl peroxides or lead tetraacetate substitute as in reaction (5).

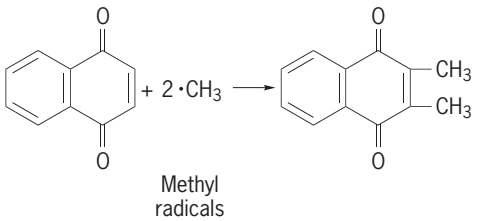

(5)

See FREE RADICAL; GRIGNARD REACTION.

Important naturally occurring naphthoquinones are vitamins K_1 (**11**) and K_2 (**12**) which are found in

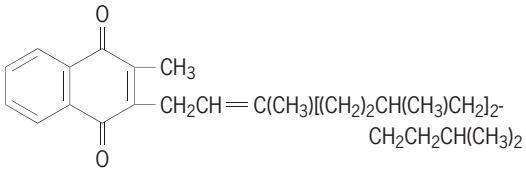

(**11**)

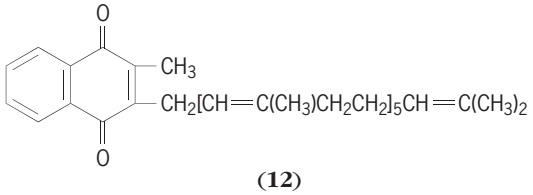

(**12**)

blood and are responsible for proper blood clotting reaction. The long aliphatic chain has been found unnecessary for the clotting reaction; its replacement by a hydrogen atom gives Menadione or 2-methyl-1,4-naphthoquinone, which is manufactured synthetically for medicinal use. *See* VITAMIN K.

A number of quinone pigments have been isolated from plants and animals. Illustrative of these are juglone (**13**) found in unripe walnut shells and spinulosin (**14**) from the mold *Penicillium spinu-*

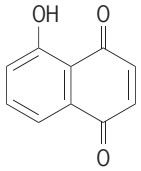

(**13**)

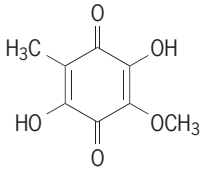

(**14**)

losum. 9,10-Anthraquinone derivatives form an important class of dyes of which alizarin is the parent type. *p*-Benzoquinone is manufactured for use as a photographic developer. *See* AROMATIC HYDROCARBON; DYE; KETONE; OXIDATION-REDUCTION.

David A. Shirley

Bibliography. R. J. Fessenden and J. S. Fessenden, *Organic Chemistry*, 6th ed., 1998; S. Patai, *The Chemistry of Quinonoid Compounds*, 1974; Z. Rappoport (ed.), *The Chemistry of Quinones*, 1988; J. C. Stowell, *Intermediate Organic Chemistry*, 2d ed., 1993; R. H. Thompson, *Naturally Occurring Quinones*, 3d ed., 1987.

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

15 R - RYE

www.MHEST.com

Rabies

An acute, encephalitic viral infection. In humans it is almost invariably fatal. Human beings are infected from the bite of a rabid animal, usually a dog. *See* ANIMAL VIRUS.

Infectious agent. The virus is bullet-shaped with a cylinder diameter of about 10 nanometers and a length of about 210 nm. The nucleic acid is ribonucleic acid (RNA). The virion is enveloped, covered with projections, and ether-sensitive. A grouping of rabies virus, together with vesicular stomatitis virus of cattle and certain viruses of fish, insects, and plants, has been included under the name rhabdovirus group. Canine rabies can infect all warm-blooded animals, and death usually results. Some animals show chiefly paralytic signs, whereas others manifest encephalitic hyperexcitability and viciousness. However, the vampire bat may transmit virus for months while apparently not infected, and in other wild animals infection may take other courses than the fatal encephalitis. The virus will grow in chick embryos or in chick or mouse embryo tissue cultures. Strains freshly isolated from dogs and humans are called street virus; strains of altered pathogenicity and a stable, shortened incubation period, produced by serial passage in rabbit brains, are called fixed. Except in bats, street virus invariably produces cytoplasmic inclusion bodies (Negri bodies) in infected nerve cells. *See* EMBRYONATED EGG CULTURE; TISSUE CULTURE; VIRAL INCLUSION BODIES.

Pathogenesis. The virus is believed to move from the saliva-infected wound through sensory nerves to the central nervous system, multiply there with destruction of brain cells, and thus produce encephalitis, with severe excitement, throat spasm upon swallowing (hence hydrophobia, or fear of water), convulsions, and death—with paralysis sometimes intervening before death.

Diagnosis. Diagnosis in the human is made by observation of Negri bodies in brains of animals inoculated with the person's saliva, or in the person's brain after death. Diagnosis in dogs is essential for guidance concerning human vaccination. A dog which has bitten a person is isolated and watched for 10 days for signs of rabies; if none occur, rabies was absent. If signs do appear, the animal is killed and the brain examined for Negri bodies, or for rabies antigen by testing with fluorescent antibodies. *See* FLUORESCENCE MICROSCOPE.

Epidemiology. Rabies occurs throughout the world. About 1000 cases of human rabies are reported annually to the World Health Organization, almost all from developing countries. Only rarely are human cases reported in the United States, but a reservoir is constantly present, as shown by outbreaks in domestic and wild animals. In South and Central America, vampire bats and fruit- and insect-eating bats transmit the disease to cattle, and occasionally to humans. Since bats may carry the virus without becoming ill, they are important as a constant source of infection for wildlife predators, domestic animals, and human beings. In addition, aerosol transmission of rabies by bat has been proved, adding to the problem of rabies control, especially among persons who explore caves which are inhabited by bats.

Immunity and prevention. Individuals at high risk must receive preventive immunization. If exposure is believed to have been dangerous, postexposure prophylaxis should be undertaken. The rabies virus may remain latent in tissues for some time after being introduced from a bite, but if antibody or immunogenic vaccine is administered promptly, the virus can be prevented from invading the central nervous system. Passively administered antibody provides additional time for a vaccine to stimulate active antibody production before the central nervous system is invaded.

An inactivated rabies virus vaccine is available in the United States. It is made from virus grown in human or monkey cell cultures and is free from brain proteins that were present in earlier Pasteur-type

vaccines. This material is sufficiently antigenic that only four to six doses of virus need be given to obtain a substantial antibody response. An experimental recombinant virus oral vaccine, which consists of vaccinia virus carrying the rabies surface glycoprotein gene, has successfully immunized animals, and it should prove valuable in the immunization of both wildlife reservoir species and domestic animals. *See* BIOLOGICALS; VACCINATION.

Since 1960, one to five cases of human rabies have occurred annually in the United States, but 20,000 to 30,000 persons require some treatment for possible bite-wound exposure each year. All bites should immediately be cleaned thoroughly with soap and water, and a tetanus shot should be considered. The decision to administer rabies antibody, rabies vaccine, or both depends on four factors: the nature of the biting animal; the existence of rabies in the area; the manner of attack (provoked or unprovoked) and the severity of the bite and contamination by saliva of the animal; and recommendations by local public health officials. *See* PUBLIC HEALTH.

Control. Isolated countries, such as Great Britain, that have no indigenous rabies in wild animals can establish quarantine procedures; imported domesticated animals are quarantined for 6–12 months. In countries where dog rabies exists, stray animals should be destroyed, and vaccination of pet dogs and cats should be mandatory. Where wildlife rabies exists—and where contact between domestic animals, pets, and wildlife is inevitable—all domestic animals and pets should be vaccinated, and the incidence of rabies in wild animals should be continually monitored.

Preexposure vaccination is desirable for all persons who are at high risk of contact with rabid animals, particularly veterinarians, animal care personnel, certain laboratory workers, and spelunkers. Persons traveling to countries where rabies control programs for domestic animals are not optimal should receive preexposure prophylaxis if they plan to stay for more than 30 days. Persons who stay in rabies-endemic areas for extended periods and who are at risk of inapparent exposure to rabies or a delay in postexposure prophylaxis should be advised to have a booster vaccination every 2 years or to undergo serum testing for rabies-neutralizing antibody every 2 years. If the antibody titer is inadequate, booster vaccination is advisable. Preexposure prophylaxis does not, however, eliminate the need for postexposure prophylaxis if exposure to rabies occurs.

Joseph L. Melnick

Bibliography. D. M. Knipe et al. (eds.), *Field's Virology*, 4th ed., 2001; S. A. Plotkinet al. (eds.), *Vaccines*, 3d ed., 1999.

Raccoon

A member of the mammalian family Procyonidae, along with ringtails, coatis, kinkajous, olingos, and lesser pandas. These carnivores inhabit areas from southern Canada to northern Argentina and Uruguay and have been introduced into Russia and Germany. Only one species, *Procyon lotor*, inhabits North America.

Raccoon, a member of the family Procyonidae.

Morphology. The raccoon is a medium-sized, heavily built, partially arboreal mammal with long, loose, gray to almost black fur (see **illustration**). A white-bordered black mask extends across the eyes and down onto the cheeks. The foxlike head is broad with a pointed muzzle, and the ears are prominent, erect, and somewhat pointed. The moderately long, well-haired tail is banded with five to seven blackish rings alternating with wider, lighter bands. The plantigrade feet have naked smooth soles and five very long digits, each of which possesses a nonretractile claw. No webbing is present between the toes. The "hands" are well adapted for grasping and manipulating objects. The dental formula is I 3/3, C 1/1, Pm 4/4, M 2/2 for a total of 40 teeth. Raccoons have excellent senses of hearing, vision, smell, and touch. *See* DENTITION.

Adult raccoons are 79.5–135 cm (31–53 in.) long, including a 19–40 cm (7–16 in.) long tail. They stand 22.8–30.4 cm (9–12 in.) at the shoulder and generally weigh 3–10 kg (6–21 lb). Males tend to be 20 to 30% larger than females. Body size is correlated with severity of the climate; raccoons with the smallest overall body size are found in the subtropical Florida Keys, and those with the largest body size are found in areas with 100 or more days of snow cover.

Ecology. Raccoons prefer moist areas and are found mainly in timbered swamps, on river bottoms, along the banks of streams and lakes, and in coastal salt marshes. These primarily nocturnal mammals are excellent climbers, and they swim well, but not far or

often. Dens may be in hollow trees, hollow logs, old stumps, burrows, muskrat houses, or in outbuildings such as barns. Seasonal activity cycles vary with latitude. Raccoons do not hibernate, although in northern areas they may remain in a den for much of the winter. Although normally a solitary species, communal denning is common with up to 23 raccoons having been reported in a single den.

Raccoons are opportunistic omnivores and feed on a wide variety of plant and animal foods including wild fruits, berries, acorns, shellfish, crustaceans, frogs, fish, small mammals, reptiles, and birds. Most male raccoons range over about 600 hectares (1500 acres) annually. Female home ranges are smaller and often overlap.

Mating and development. Mating in the United States extends from December to August, with peaks in February and March. Following an average gestation of 63 days, the annual litter is born in April or May. Litter sizes average 3 or 4, but range 1–7. Newborn animals have blackish skin and are well covered with yellowish-gray fur. Indistinct darker areas show where the facial mask and tail rings will appear prominently later. Weaning is completed at 3–4 months of age. Mother and young remain together into the late fall and often well into the winter. Sexual maturation is attained at about 12–15 months, but most raccoons do not mate until the following year. The average life span is 2–3 years in the wild, although some wild individuals have lived more than 12 years.

Threats. The activities of humans (hunting, trapping, and automobiles) are the principal causes of raccoon mortality. Though the demand for its pelt has fluctuated, the raccoon has long been an important fur species in North America. Foxes, bobcats, hawks, owls, and snakes may prey on young raccoons. Distemper and rabies can temporarily reduce populations.

Other species. Central and South American raccoons include the Tres Marias raccoon (*P. insularis*) which inhabits Maria Madre Island, Mexico; the Barbados raccoon (*P. gloveralleni*) of Barbados; the crab-eating raccoon (*P. cancrivorus*) that ranges from Costa Rica south to northern Argentina; the Cozumel Island raccoon (*P. pygmaeus*) of Cozumel Island, Yucatan, Mexico; the Guadeloupe raccoon (*P. minor*) of Guadeloupe; and *P. maynardi* found on Nassau Island in the Bahamas. Recent researchers feel that the insular forms (*P. insularis*, *P. gloveralleni*, *P. pygmaeus*, *P. minor*, and *P. maynardi*) are conspecific with *P. lotor*. All of the insular forms are classified as Endangered. The Barbados raccoon may be extinct. *See* CARNIVORA; MAMMALIA. Donald W. Linzey

Bibliography. *Grzimek's Encyclopedia of Mammals*, McGraw-Hill, 1990; D. W. Linzey, *The Mammals of Virginia*, McDonald & Woodward, 1998; D. Mcdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford Limited, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; J. O. Whitaker, Jr. and W. J. Hamilton, Jr., *Mammals of the Eastern United States*, 3d ed., Comstock Publishing Associates, 1998; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Racemization

The formation of a racemate from a pure enantiomer. Alternatively stated, racemization is the conversion of one enantiomer into a 50:50 mixture of the two enantiomers (+ and −, or R and S) of a substance. Racemization is normally associated with the loss of optical activity over a period of time since 50:50 mixtures of enantiomers are optically inactive. *See* OPTICAL ACTIVITY.

Racemization is an energetically favored process since it reflects a change from a more ordered to a more random state. But the rate at which enantiomers racemize is typically quite slow unless a suitable mechanistic pathway is available, since racemization usually, but not always, requires that a chemical bond at the chiral center of an enantiomer be broken. Racemization of enantiomers possessing more than one chiral center requires that all chiral centers of half of the molecules invert their configurations. *See* ENTROPY.

Mechanisms. Of the several known racemization mechanisms, those involving the temporary formation of intermediates which possess reflection symmetry are the most common. Depending upon the substance and the conditions employed, the intermediate may be a free radical, carbocation, or carbanion. The racemization of an alkyl halide in the presence of a Lewis acid catalyst [reaction (1)]

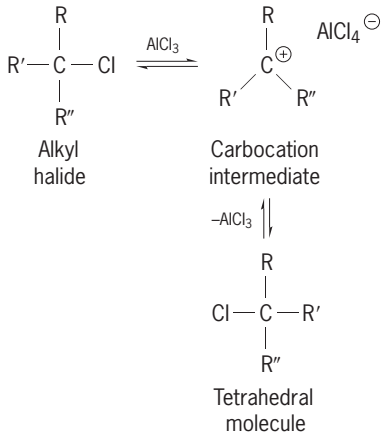

(1)

illustrates the intervention of a planar, hence achiral, carbocation intermediate. Reversal of the ionization or capture of the cation by a solvent molecule can occur with equal probability to either side of the planar intermediate which then reforms a molecule of tetrahedral geometry. When equal numbers of (R)- and (S)-enantiomers are present, racemization is complete.

Chiral compounds with acidic hydrogens at the chiral center are racemized by base with the

intervention of carbanions [reaction (2)]. Delocalization of the unshared electrons makes the enolate intermediate planar. For carbanions for which delocalization is not possible, racemization takes place because a nonplanar carbanion usually inverts its configuration readily.

Racemization via free radicals is exemplified by the chlorination of hydrocarbon which involves an intermediate free radical mechanism [reaction (3)].

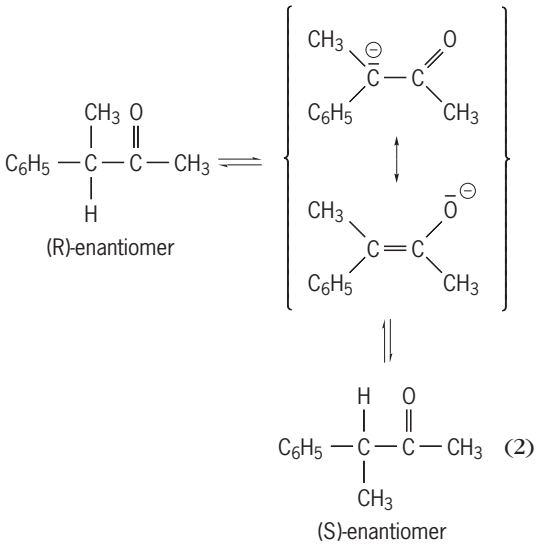

(2)

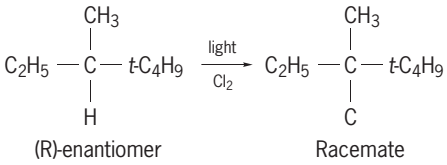

(3)

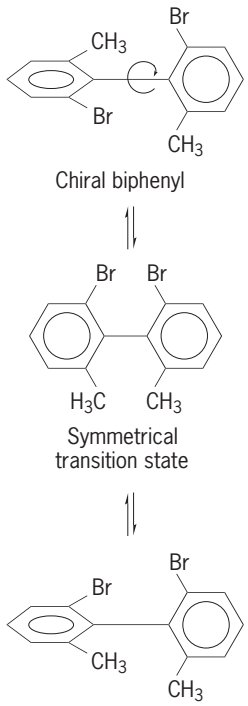

(4)

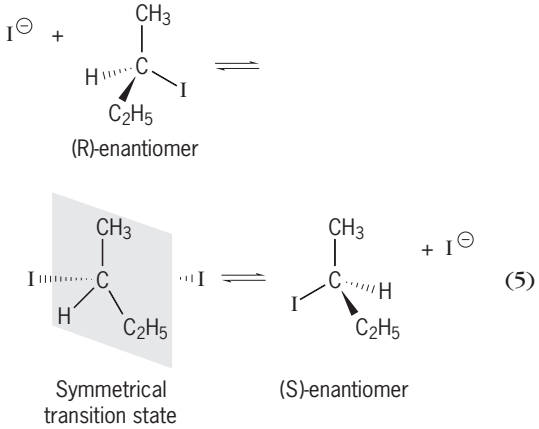

(5)

The product is racemic. The *t* in the structures stands for tertiary (as intertertiary butyl group) and is a descriptor defining one of the isomers of the butyl group (C_4H_9). Heat can also cause racemization without breaking bonds if it is possible for a molecule to pass through a state which possesses reflection symmetry. For example, the chiral biphenyl may be racemized thermally by rotation about the central carbon-carbon bond. In this case the benzene rings become coplanar in the transition state [reaction (4)]. Nucleophilic displacement on chiral alkylhalides may also lead to racemization via symmetrical transition states if the displacing ion is identical to the leaving ion [reaction (5)]. This type of change in configuration is known as Walden inversion. *See* FREE RADICAL; ORGANIC REACTION MECHANISM; REACTIVE INTERMEDIATES.

Significance. The observation and study of racemization have important implications for the understanding of the mechanisms of chemical reactions and for the synthesis and analysis of chiral natural products such as peptides. Moreover, racemization is of economic importance since it provides a way of converting an unwanted enantiomer into a useful one. Synthetic medicinal agents are often produced industrially as racemates. After resolution and isolation of the desired enantiomer, half of the product would have to be discarded were it not for the possibility of racemizing the unwanted isomer and of recycling the resultant racemate.

Inversion of configuration at some but not all chiral centers in diastereomeric peptides and sugars does not lead to total loss of configuration. This process, while analogous to racemization, does not lead to complete loss of optical activity. It is called epimerization. Epimers are diastereomers which differ at only one of several chiral centers. Epimerization of amino acids in proteins takes place very slowly on standing, and is used as a means of dating very old protein samples (aminostratigraphy). *See* STEREOCHEMISTRY.

Samuel H. Wilen

Bibliography. E. L. Eliel and S. H. Wilen, *Stereochemistry of Organic Compounds*, 1994; D. Naspuri, *Stereochemistry of Organic Compounds: Principles and Applications*, 1991; M. Nogradi, *Stereoselective Synthesis*, 2d ed., 1994; F. Vogtle and E. Weber (eds.), *Stereochemistry*, 1984.

Radar

An acronym for radio detection and ranging, the original and basic function of radar. The name is applied to both the technique and the equipment used. Radar is a sensor; its purpose is to provide estimates of certain characteristics of its surroundings of interest to a user, commonly the presence, position, and motion of aircraft, ships, or other vehicles in its vicinity. In other uses, radars provide information about the Earth's surface (or that of other astronomical bodies) or about meteorological conditions. To provide the user with a full range of sensor capability, radars are often used in combinations or with other elements of the complete system.

Radar operates by transmitting electromagnetic energy into the surroundings and detecting energy reflected by objects. If a narrow beam of this energy is transmitted by the directive antenna, the direction from which reflections come and hence the bearing of the object may be estimated. The distance to the reflecting object is estimated by measuring the period between the transmission of the radar pulse and reception of the echo. In most radar applications, this period will be very short since electromagnetic energy travels with the velocity of light.

Kinds of radar. Radar has so many valuable applications that the physical nature of radars varies greatly. Familiar radars that are very small and compact include those used on small boats or in small aircraft for navigation and safety and, even smaller, the handheld radars used for measurement of vehicle speed or even that of a baseball pitch. *See* MARINE NAVIGATION.

The largest radars cover acres of land, long arrays of antennas all operating together to monitor the flight of space vehicles and orbiting debris, an increasing space-flight problem, and also to contribute to radar astronomy. Other very large radars provide coverage at substantial distances well beyond the normal radar horizon; they monitor flight activity and sea-surface conditions, contributing to both air-traffic and meteorological interests. These radars are very large because they must use longer-than-usual radio wavelengths [those in the high-frequency (HF) part of the radio spectrum, and not the much higher microwave frequencies more common in radar] that result in the ionospheric containment of the signal for over-the-horizon operations; these longer wavelengths require very large antennas to form the narrow beams useful in the measurement of target positions or the surface spot being observed. *See* RADAR ASTRONOMY; SPACECRAFT GROUND INSTRUMENTATION.

More common in size are those radars at airports, with rotating antennas 20–40 ft (6–12 m) wide, a room or two of electronic equipment, and operator consoles the size of a desk. Radars intended for mobile use, particularly those in aircraft and in trucks, buses, and automobiles, are quite compact and very modular in construction, well adapted to fit into the vehicle. *See* AIRBORNE RADAR.

Radars intended principally to determine the presence and position of reflecting targets in a region around the radar are called search radars. Other radars examine further the targets detected: examples are height finders with antennas that scan vertically in the direction of an assigned target, and tracking radars that are aimed continuously at an assigned target to obtain great accuracy in estimating target motion. In some modern radars, these search and track functions are combined, usually with some computer control. Surveillance radar connotes operation of this sort, somewhat more than just search alone. *See* SURVEILLANCE RADAR.

The development of quite complex and versatile radars, computer-controlled, generally using

(a)

(b)

Fig. 1. Multifunction phased-array radars. (*a*) One of the four planar faces of the radar of the AEGIS combat system, which is in use on many U.S. cruisers and destroyers (*Lockheed Martin*). (*b*) Pave Paws radar installation. (*Raytheon Company*)

phased-array antennas and called multifunction radars, has been quite successful worldwide. Such antennas (described below) require no physical motion to position the radar beam, so various coverage routines and modes of operation can be scheduled with an efficiency and expediency unachievable in lesser radars. The U.S. Navy's AEGIS combat system features such a radar (**Fig. 1***a*), as does the Army's Patriot missile system. Large land-based multifunction phased-array radars include the Cobra Dane (and variants) system and the Pave Paws radars (Fig. 1*b*), which have contributed to United States defense surveillance for several decades. Pave Paws, as an example, used two array faces, each 72.5 ft (21.8 m) in diameter composed of approximately 14,000 solid-state radiating elements, operating in the UHF frequency band. Many other nations have now deployed such radars of various scale, and developments continue in such systems for land-based, shipboard, airborne, and space-based radars. No longer are such multifunction radars a developmental curiosity.

An important airborne multifunction radar that uses mechanical rotation of its antenna in azimuth and electronic beam positioning in elevation is the Air Force's S-band AWACS (Airborne Warning and Control System) radar, the antenna of which is mounted in a rotodome atop the aircraft (**Fig. 2***a*). The precision-built waveguide array (Fig. 2*b*) features extremely low sidelobes, essential in almost any airborne radar because of the strong ground echo at all angles. New multimission military aircraft will include multifunction radars as an integral part of their design. *See* MILITARY AIRCRAFT.

Very accurate tracking radars intended for use at missile test sites or similar test ranges are called instrumentation radars. Radars designed to detect clouds and precipitation are called meteorological or weather radars. Improved techniques in Doppler-sensitive signal processing have enabled radars to examine more precisely the motion of reflecting objects (windborne dust, insects, and precipitants, for example) indicative of weather disturbances. Such radars are used in critical places such as airports where sudden wind shears (as from downdraft microbursts in the area) have threatened air safety. Many weather radars also use microwave polarimetry to examine the precipitants and characterize the disturbance more accurately. *See* DOPPLER EFFECT; DOPPLER RADAR; METEOROLOGICAL RADAR; STORM DETECTION.

Airborne and spaceborne radars have been developed to perform ground mapping with extraordinary resolution by special Doppler-sensitive processing while the radar is moved over a substantial distance. Such radars are called synthetic-aperture radars (SARs) because of the very large virtual antenna formed by the path covered while the processing is performed. Interferometry can provide topological information (3D SAR), and polarimetry and other signal analysis can provide more information on the nature of the surface (type of vegetation, for example). *See* REMOTE SENSING; SYNTHETIC APERTURE RADAR (SAR).

Some radars have separate transmit and receive antennas sometimes located miles apart. These are called bistatic radars, the more conventional single-antenna radar being monostatic. Some useful systems have no transmitter at all and are equipped to measure, for radarlike purposes, signals from the targets themselves. Such systems are often called passive radars, but the terms "radiometers" or "signal intercept systems" are generally more appropriate. Regular radars may, of course, occasionally be operated in passive modes. Doing so helps locate sources of interference, either deliberate (jamming) or unintentional. Estimating the angle-of-arrival of signal sources can be used to draw other sensors (infrared sensors, for example; other active radars) to those angles for further examination. Angle data also contribute to triangulation processes (with similar data from other sensors) to locate sources more completely. The spectral data, showing perhaps the less occupied frequencies, can be used in governing regular radar operations somewhat adaptively. Passive radars (or modes) can be equipped to analyze the signals being received to help identify targets by association, a common data fusion process in military systems particularly. *See* PASSIVE RADAR.

The terms "primary" and "secondary" are used to describe, respectively, radars in which the signal received is reflected by the target and radars in which

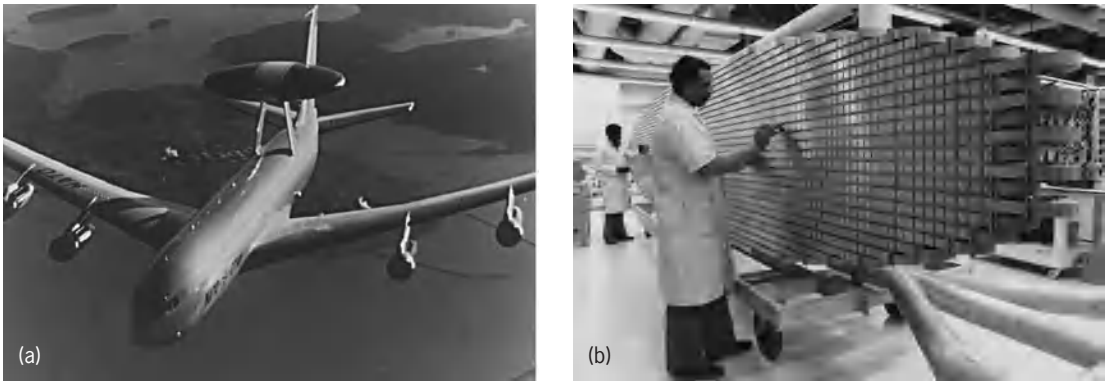

Fig. 2. Air Force AWACS (Airborne Warning and Control System) in use. (*a*) AN/APY-1 rotodome installed. (*b*) Precision-array antenna of the AN/APY-1. (*Northrop/Grumman Corp.*)

the transmission causes a transponder (transmitter-responder) carried aboard the target to transmit a signal back to the radar. The Identification Friend or Foe (IFF) system in both military and civil use is a secondary radar. *See* AIR-TRAFFIC CONTROL; ELECTRONIC NAVIGATION SYSTEMS.

Fundamentals of operation. It is convenient to consider radars as having four principal subsystems: the transmitter, the antenna, the receiver (and signal processor), and the control and interface apparatus.

The transmitter provides the radio-frequency (RF) signal in sufficient strength (power) for the radar sensitivity desired and sends it to the antenna, which causes the signal to be radiated into space in a desired direction. The signal propagates (radiates) in space, and some of it is intercepted by reflecting (or "scattering") bodies. The part of the incident signal that is backscattered returns to the radar antenna, which collects it and routes it to the receiver (along with all incoming signals) for appropriate signal processing and detection. The presence of an echo of the transmitted signal in the received signal reveals the presence of a target. The echo is indicated by a sudden rise in the output of the detector, which produces a voltage (video) proportional to the sum of the RF signals being received and the RF noise inherent in the receiver itself. The range to the target is half the product of the speed of electromagnetic propagation (3×10^8 m/s), and the time between the transmission and the receipt of the echo. The direction or bearing of the target is disclosed by the direction to which the beam is pointed when an echo is received.

In some radars, pulsing may be too rapid to permit range determination by direct echo timing in this unambiguous way. Such radars are said to employ higher pulse-repetition-frequency (prf) modes of operation; the higher sampling rate is necessary for clearly distinguishing moving targets from highly reflective but nearly stationary backgrounds (clutter) through the Doppler effect involved. Range determination then requires resolution of the ambiguities in echo timing by using more than one such pulse repetition frequency. Airborne radars are particularly dependent upon such Doppler-sensitive modes of operation.

Most radars indeed use pulses in their operation. Those that do not are called continuous-wave (cw) radars. Some missile seekers, mostly interested only in the direction to a target, are of this type; police radars interested only in the Doppler shift of the echo for speed measurement are another example. In pulsed radars (**Fig. 3**) with a transmitter of the master-oscillator power-amplifier (MOPA) type, the conversion of electrical power to transmitted signal power takes place in a chain of amplifiers of increasing power level, each controlled by the modulator forming the radar pulse. In some simpler radars, a single high-power oscillator may be used with no further amplification. In newer high-performance radars, the final stage of amplification may be in transmit-receive modules in the antenna structure itself, one such module at each element of a large active array. *See* AMPLIFIER; CONTINUOUS-WAVE RADAR; MODULATOR; OSCILLATOR; POWER AMPLIFIER; RADIO TRANSMITTER.

A duplexer permits the same antenna to be used on both transmit and receive, and is equipped with protective devices to block the very strong transmit signal from going to the sensitive receiver and damaging it. The antenna forms a beam, usually quite directive, and, in the search example, rotates throughout the region to be searched, or otherwise causes the beam to be positioned as desired, perhaps without mechanical motion, as in radars using a computer-controlled phased-array antenna (Fig. 3*b*,*c*). In such phased arrays, the duplexing function is generally built into the feed system and phasing logic. *See* ANTENNA (ELECTROMAGNETISM).

Radars searching with mechanical antennas generally transmit several pulses (perhaps 20–40) as the antenna scans past a target, whereas phased-array radars may search with various "dwells," sometimes with just one transmission per beam position. The use of several pulses allows a buildup of the echo being received, a process called noncoherent or video integration. Most radars are equipped with low-noise RF preamplifiers to improve sensitivity. The signal is then "mixed" with (multiplied by) a local oscillator signal to produce a convenient intermediate-frequency (IF) signal, commonly at 30 or 60 MHz; the same principle is used in all heterodyne radio receivers. The local oscillator signal, kept offset from the transmit frequency by precisely this intermediate frequency, is supplied by the transmitter oscillators during reception. After other significant signal processing in the IF circuitry (of a digital nature in many newer radars), a detector produces a video signal, a voltage proportional to the strength of the processed IF signal. This video can be applied to a cathode-ray-tube (CRT) display so as to form a proportionately bright spot (a blip), which could be judged as resulting from a target echo. However, radars increasingly use artificial computerlike displays based on computer analysis of the video. Automatic detection and automatic tracking (based on a sequence of dwells) are typical of such data processing, reports being displayed for radar operator management and also made instantly available to the user system. *See* CATHODE-RAY TUBE; ELECTRONIC DISPLAY; HETERODYNE PRINCIPLE; MIXER; PREAMPLIFIER; RADIO RECEIVER.

The waveform of an elementary search radar consists of a train of pulses. Duty factors (the ratio of pulse width or transmit "on" time, to total time) in such radars are typically 0.1 to 1.0%. With modern signal processing, much more complicated waveforms are often used, and with the advent of transistors and other solid-state amplifiers, radar systems have been developed with duty factors of 10 to 50%.

Radar carrier-frequency bands. Radar carrier frequencies are broadly identified by a widely accepted nomenclature. The spectrum is divided into bands (see **table**). In addition, some special-purpose

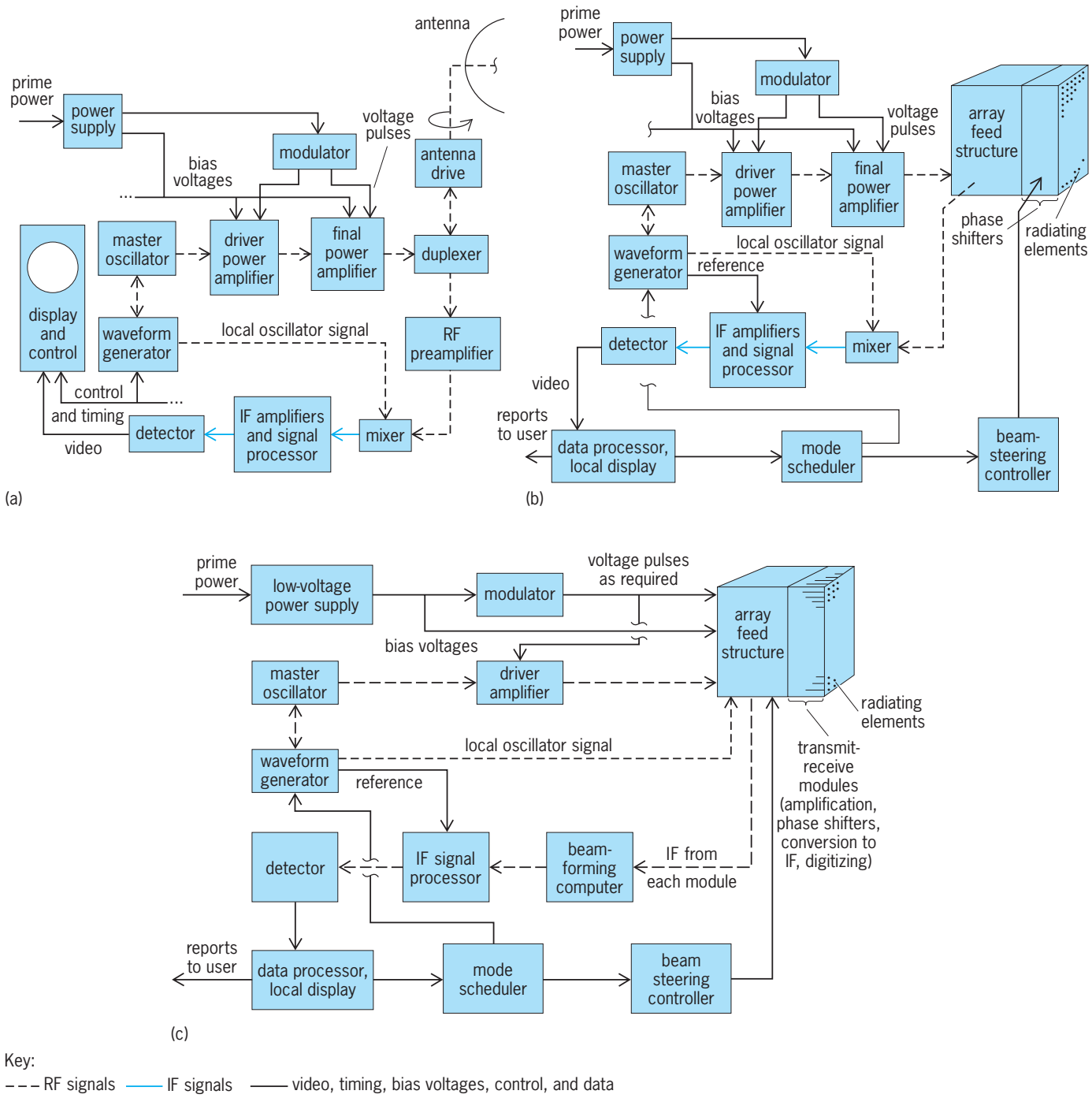

Fig. 3. Block diagram of pulse radar. (*a*) Radar with a mechanically scanned antenna. (*b*) Radar with a phased-array antenna for electronic scanning. Array feed structure includes transmit distribution, receive beam forming, duplexing, and preamplification. (*c*) Radar with active phased array using transmit-receive modules at every element and digital beam forming on receive.

radars are operated with laser transmitters at infrared and near-optical frequencies. *See* LIDAR; RADIO SPECTRUM ALLOCATIONS.

Propagation. There are several noteworthy characteristics of the radar's surroundings in terrestrial operation. The charged layers of the ionosphere present a highly refractive shell at radio frequencies well below the microwave frequencies of most radars. Consequently, over-the-horizon radars operate in the 10-MHz region to exploit this skip path. Further, the Earth's surface makes such an abrupt change of medium that electromagnetic reflections take place. In microwave radars these forward-scattered surface reflections cause a "multipath" condition resulting in the lobed interference structure seen in field strength graphs of radar coverage (**Fig. 4**). These lobes in the plane normal to the reflecting surface result from the alternately constructive and destructive phase interference of the direct and the reflected signals, varying with the elevation angle. When

Radar carrier-frequency bands

Band designation	Nominal frequency range	Representative wavelength
HF	3–20 MHz	30 m at 10 MHz
VHF	30–300 MHz	3 m at 100 MHz
UHF	300–1000 MHz	1 m at 300 MHz
L	1.0–2.0 GHz	30 cm at 1 GHz
S	2.0–4.0 GHz	10 cm at 3 GHz
C	4.0–8.0 GHz	5 cm at 6 GHz
X	8.0–12.0 GHz	3 cm at 10 GHz
K_u	12.0–18 GHz	2 cm at 15 GHz
K	18–27 GHz	1.5 cm at 20 GHz
K_a	27–40 GHz	1 cm at 30 GHz
mm	40–300 GHz	0.3 cm at 100 GHz

terrain is not so reflective, the lobe structure is less pronounced. Because there is always destructive interference (giving minimum signal strength) at 0° elevation, there exists an inherent performance limitation for very low flying targets, often said to fly under the radar beam. Similarly, narrow-beam tracking radars have difficulty measuring the true elevation angle of low-angle targets when the beam grazes the Earth's surface and incurs multipath interference. *See* INTERFERENCE OF WAVES; IONOSPHERE; RADIO-WAVE PROPAGATION; REFLECTION OF ELECTROMAGNETIC RADIATION.

The atmosphere also acts as a lens since its dielectric constant decreases with increasing altitude. Consequently, microwave rays are bent downward slightly, and the radar horizon is somewhat beyond the visual. Under certain conditions defined as standard, the effect is such that if one were to draw the Earth with a radius 4/3 its actual value, the refracted rays would be bent back to straight lines.

Frequently, an atmospheric condition involving a pronounced departure from the smoothly varying dielectric constant with altitude occurs, and a superrefractive or ducting condition results with low-altitude targets detected well beyond 4/3-Earth predictions. The variation in such conditions over time of day, season, geographic location, and local weather makes reliable accounting for the phenomenon difficult. Attenuation of radar signals in the

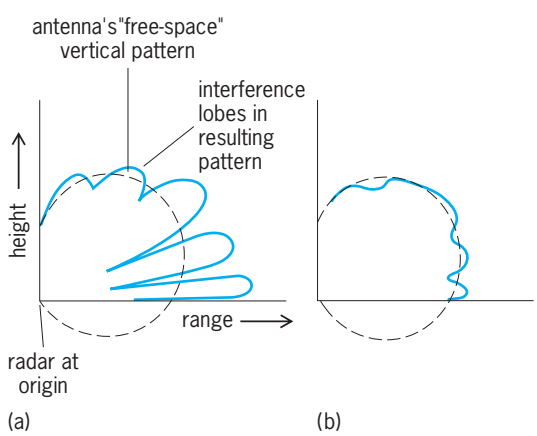

Fig. 4. Effect of a reflecting Earth surface on the radiation pattern of an antenna. (*a*) Reflection nearly perfect. (*b*) Reflection much reduced.

Earth's atmosphere is due to both molecular absorption by resonant excitation of uncondensed gases and scattering by particles such as dust or water-droplets in fog and clouds. The attenuation due to water vapor and to oxygen molecules is negligible below the L-band and tolerable through the X-band. The resonance peaks for these molecules cause much more concern in choosing frequencies in the millimeter region for special-purpose radars. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION.

Target characteristics. This discussion assumes that targets of interest are vehicles such as aircraft, trucks, ships, or spacecraft. Terrain features, buildings, and vegetation are certainly of interest to other (imaging) radars in which signal processing addresses the reflectivity, polarimetric, and other characteristics of such distributed targets. Targets such as vehicles are composed of many scatterers or reflecting surfaces, all of which contribute to the whole reflection. The individual reflections add coherently (that is, with both constructive and destructive phase interference, depending on the many path lengths involved) in composing the total reflection. Consequently, over successive observations, the reflections will differ in magnitude if even the slightest repositioning of the scatterers occurs, or if a different carrier frequency is employed. *See* COHERENCE.

Such targets have been categorized by the statistical nature of these fluctuations in echo strength; the Swerling models (after Peter Swerling) serve as a basis. Advances in more precise modeling of the composition of targets, aided by computer analysis, provide assistance to the designers of both radars and aircraft (or other potential targets). A single average value of a target's radar cross section often suffices for basic radar calculations, provided a statistical allowance is made for the fluctuations. Some targets are designed to have very little reflection back to the radar; stealth aircraft and missiles are examples. Others, however, may be designed to concentrate the reflections back to the radar. For safety reasons, small boats and aircraft are often equipped with small corner reflectors of right-intersecting planes, so that the radar signal is reflected back to the radar regardless of the direction from which it came, making the radar return far larger and more visible to a searching radar than would be the random reflections from the target itself. These and other retrodirective reflectors are available in various physical sizes and aid in instrumentation, navigation, and search-and-rescue situations.

The strength of the echo from complex targets, then, is properly treated as a random variable. An expected value (a mean value) might be used in elementary performance calculations for radar. Radar cross section is defined as a fictitious area which, if intercepting the incident signal and reradiating isotropically the signal intercepted, would return to the radar the same strength echo that the real target returns. Radar cross sections of ships can be in the thousands of square meters, of commercial aircraft several tens of square meters, of smaller

aircraft and missiles a few square meters, and of deliberately stealthy aircraft a small fraction of a square meter (perhaps about the same as a bird or large insect). Retro-reflective aids to navigation can have radar cross sections of tens or hundreds of square meters, even though they are small in physical size.

Noise. The detectability of a radar echo is determined by its strength relative to various competing signals also causing a detector output. Among the sources of such signals is the receiver itself. The receiver produces RF and IF signals as a result of thermal conditions (molecular agitation in the components) and of the mixing process (unintended constituents of the local oscillator signal, for example).

A measure of the receiver's noise production is called its noise factor. In general terms, the noise factor is the ratio of noise power at the output of a network to the noise power that would have resulted from the network simply acting on the noise power at the input. The added noise of the network itself causes a degradation in the signal-to-noise ratio from input to output. It can be seen, then, that a sensitive radar receiver might to good advantage use a first stage of low noise factor but of considerable gain, so that subsequent stages are provided with a relatively strong signal (and noise) at their inputs, their own noise production further degrading the now-amplified signal-to-noise ratio very little. Noise factor when expressed in decibels is more commonly called noise figure. *See* ELECTRICAL NOISE.

Other noise sources can be external to the radar. They include the Sun, other cosmic sources, the Earth's own surface, electromagnetic interference from urban areas, and deliberately energized sources intended to reduce the radar's sensitivity (jamming). *See* ELECTRICAL INTERFERENCE; ELECTRONIC WARFARE; JAMMING; RADIO ASTRONOMY.

Clutter. While some radars are designed to examine land features and others, weather phenomena, generally such radar returns are considered a nuisance, hindering the detection of targets of greater interest (such as aircraft) and producing false alarms. It is possible to state elementary formulas for the power of such returns, but in general the simple expressions fail to account sufficiently for the statistical distribution of clutter backscatter encountered in practice.

Several types of clutter have been carefully investigated. Surface clutter includes scattering from the sea surface (for various sea-state conditions) and from land surface (for various terrain types). Volume clutter includes fog, clouds, rain, snow, and chaff (reflective threads deliberately dispersed in the air as a radar countermeasure), and even birds and insects have been taken into account. Clear-air turbulence and thermal inversions can be reflective; these are often called angels by radar users. *See* CLEAR-AIR TURBULENCE.

In the radar bands, clutter backscatter generally increases with increasing frequencies—in a given amount of rain, the clutter power will be worse for an X-band radar than for an L-band one. Radars using frequencies below L-band are often called all-weather radars. Radars at S-band and higher may use high resolution (narrow beams, very short pulses) in part to reduce the amount of clutter competing with a target at any one time.

In addition to the strength of the backscatter, the spread in frequency of clutter return is of concern in radar design. Such spectral spread in the Doppler-shifted return comes from turbulence among the scatterers as experienced by a surface-based radar; it is particularly severe in airborne radar because the return from the Earth's surface comes from all angles (even through the sidelobes of the directive antenna) while the aircraft moves. In signal processing intended to suppress returns from presumed stationary clutter, this clutter spectrum is a major consideration. Airborne radars generally feature modes of pulse Doppler processing that are unnecessary in most surface-based radars.

Detection process. Since a radar receiver is always exposed to a background of noise (and possibly other disturbance such as clutter, interference), the declaring at some instant that a desired signal (a target echo) is also present is clearly a probabilistic matter. The detection process is, in fact, a direct application of hypothesis testing in which one must decide whether the receiver's output (the video) was produced by the background only or by a target echo in addition to the background (the two hypotheses).

There are, of course, two ways to err in such a process: first, to declare a target to exist when none does (a false alarm), and second, to fail to declare a target when in fact one exists (a miss). Statistical methods are used to quantify the probabilities of false alarm and detection (the complement of miss). A useful relationship, often used in very general radar analysis and with certain simplifying assumptions about the background and the target, is that a ratio of target signal to average noise of about 13 decibels (13 dB, a power ratio of 20:1) permits detection with 0.5 probability while maintaining a false alarm probability of just 10^{-6}. Since a typical radar might be making about one million decisions per second (resolution-dependent—here as if using a pulse width of 1 microsecond), a threshold setting producing 10^{-6} false alarm probability would yield about one false alarm each second, a rate perhaps quite tolerable as an input to successive steps (multipulse integration, confirmation) of the entire detection process. An often misunderstood point in radar design is that this initial threshold should be as low as possible—for as high a rate of false alarms as can be tolerated—so that maximum possible sensitivity to the desired signal is achieved.

Today, the detection process often uses computer analysis of the receiver's output in the radar's control and interface subsystem, described more fully later. In older radars with a human operator viewing a cathode-ray-tube (CRT) display of that output (the video), the process is quite the same, even if unwittingly so. When the operator makes receiver

and display adjustments so that the background is just barely visible (minimally exciting the CRT), he is actually adjusting the "threshold" relative to the experienced background so that he would not likely mistake a noise excursion for a target and yet be reasonably likely to recognize a true echo causing a bright spot, a "blip," on the display. *See* CATHODE-RAY TUBE.

In modern automatic detection, a threshold voltage is established based on a relatively large amount of data by which the background—no matter how composed of clutter, jamming, and noise—is estimated. Detections are declared whenever the video exceeds this threshold and some further criteria are met (as with dwells of multiple transmissions). Such automatic detection may promise no particular improvement in basic sensitivity, but offers very consistent performance and a greater capacity in dense target situations than could be expected of human decision making. Abrupt changes in the background must be recognized in such automation and the threshold adjusted to maintain a constant false alarm rate (CFAR); this is perfectly analogous to human operators tempering their declaring of targets when they sense changes in the background shown on the display.

In automated radar, then, an operator's display will likely show symbols and other computer-generated indications of targets and local conditions useful for his further action and overall management of the system, but generally will not show the receiver's video output directly.

The advent of stealth targets (ones with very low radar cross section) has prompted radar detection methods that involve more than thresholding on signal strength. Processes involving polarimetric signature, greater Doppler dependence, and track behavior estimation have been refined to accommodate detection of target signals no stronger than the background present but differing from it in these other ways.

Tracking. Many processes are used in radar systems to associate successive declared detections (now called contacts or plots) to form a history and a prediction of target motion, that is, to form its track. Historically, such processing was done rather manually, on or near the radar display, showing the tracks on a plotting board. Modern radars use computers to form such tracks by associating the contacts into probable sequences, that is, the tracks. In dense target environments the association process itself is a challenge. Methods of accounting for measurement accuracy, the quality (strength) of each contact, the localized history of successful and failed associations (indicating local confusing clutter conditions or the like), and very probabilistic analysis involving multiple associations as multiple hypotheses (until untrue pairings naturally fail) have all been explored to help in this important step. After association, the track is formed using mathematical algorithms for estimating course and speed (the velocity vector) which are called filters; the process is called smoothing. In automated radars that use regularly scanning antennas, this processing is called track-while-scan; in multifunction phased-array radars, it is called sampled-data-tracking since track updating can be scheduled quite independently of the search mode. Adaptive tracking refers usually to using track filters in which some of the terms can be adjusted to the behavior of either the track or the radar itself being experienced. Sometimes multiple radars (or other sensors) share a common field of view and separately generated contacts are used to form a single set of tracks; this is called integrated automatic tracking, a form of data fusion. *See* ESTIMATION THEORY.

Radars have been built that are dedicated to tracking just one target but to do so very accurately (**Fig. 5**). In these radars a dish antenna is steered to the direction of a known target. Such a designation may have come from a search radar, and the tracking radar may have to perform a small local search in order to acquire the target. After acquisition, a continuous stream of contacts results and is used to keep the antenna steered to the target. The dish antenna produces a beam narrow in both horizontal and vertical angles (a pencil beam). If this beam is scanned in a very small cone around the target's presumed direction, the amplitude and phasing of the resulting modulations of the returns can be detected and used

Fig. 5. AN/FPS-16 tracking radar, a very accurate monopulse radar, operating in the C-band, installed at more than 50 sites around the world. (*RCA Corp.*)

to drive the motors that control the position of the antenna.

In these con-scan trackers, usually only the feedhorn in front of the dish is caused to nutate; the whole dish and feed assembly is repositioned by the director. The final estimate of the target's angle is made by sensing the director's position; the inertia due to its mass acts as a smoothing filter in the track estimate.

Another technique used in such tracking radars is monopulse angle measurement, in which multiple beams are generated by a cluster of feedhorns in a conventional antenna or by special beam-forming networks in phased arrays. A sequence of transmissions is not required, as it is in conical scanning, and fine angle measurement can be made on each transmission. While a tracking radar with a mechanical antenna (such as used in instrumentation radars and in missile guidance and gunfire control systems) dwells continuously on just one target, phased arrays permit scheduling of track dwells only as required to maintain a specified accuracy, and do so on dozens of targets at a time. *See* MONOPULSE RADAR.

Composition of radar systems. It is convenient to discuss radar systems in terms of four principal subsystems, namely the transmitter, the antenna, the receiver and signal processor, and the control and interface apparatus (likely including a display of some sort).

Transmitter. The transmitter converts electrical power (as from a motor-generator set or a utility system) to electromagnetic power at the carrier frequency and in the waveform desired. The three principal components of a transmitter are the power supply, the modulator, and the RF devices. The power supply changes the utility electrical power to the levels and type of electrical power needed by the particular transmitter, the modulator applies that power to the RF devices, and the RF devices convert that applied electrical power to electromagnetic power, the RF signal to be sent to the antenna and radiated.

The main functions of power supplies are transformation, rectification, and filtering. Step-up transformers change the voltage levels of the input (or prime) utility power to the high voltage required, rectifiers change that power from alternating current (ac) to direct current (dc) required by the RF devices, and the filters smooth out the dc high voltage, removing the ripples inherent in rectification. Radar power supplies must also be well regulated, which means that the voltage levels applied to the RF devices must remain constant regardless of the power being drawn by the various waveforms and modes of operation involved. Transmitters using solid-state amplification only (described below) will generally not require the high voltages (often tens of kilovolts) required for high-power vacuum tubes, but good regulation and the handling of higher total current (for similar power levels) must be accommodated. *See* ELECTRIC FILTER; ELECTRONIC POWER SUPPLY; RECTIFIER; RIPPLE VOLTAGE; TRANSFORMER.

Modulators act as switches to apply the power-supply voltage to the RF devices during transmission. Generally, vacuum tubes require bias-voltage modulation whereas solid-state amplifiers do not. The timing (when to switch and for how long) is governed by the waveform generator, often a part of the receiver equipment since the same waveform details are required there during reception. Gas-filled switch tubes (thyratrons, for example) were key in older designs, whereas solid-state circuitry is used increasingly in both modulators and power supplies today. *See* GAS TUBE.

The RF devices in many radar transmitters are vacuum tubes that amplify the signal in multiple stages (predriver, driver, and final stages, for example). Typical amplifiers are traveling-wave tubes, klystrons, and crossed-field amplifiers. The signal itself may be formed by a master oscillator or a set of oscillators constituting a frequency synthesizer, and then fed to the amplifiers. The modulator applies to these tubes the full cathode-anode voltage, or sometimes a lesser voltage to a modulating grid, just as the RF signal passes through, effecting the conversion of electrical power to RF power. *See* KLYSTRON; MICROWAVE TUBE; TRAVELING-WAVE TUBE; VACUUM TUBE.

The development of solid-state devices, generally transistors, as amplifiers at microwave frequencies has provided a design alternative to the use of large high-power vacuum tubes, appealing in several ways. Specifically, the power supply voltages required are far lower (tens of volts) than the biases required for the vacuum tubes, modulators of even that voltage are generally not required (transistors are designed to amplify only when fed an RF signal), and the probability of complete failure of the transmitter in the field is much reduced.

At frequencies from UHF through L-band, the relatively economic doped silicon compounds are commonly used as the semiconductor material, whereas the more costly gallium-arsenide devices provide higher power at frequencies above S-band. In either case, power capability lessens with increasing frequency, from perhaps 100 watts output per transistor to only 10 watts or less over the frequency ranges cited for each above. Therefore, thousands must often be used to achieve useful radar powers. Also, such amplifiers can handle an average power that is a considerable fraction of their peak power capability; hence, radar designers are led to consider modes and waveforms with duty factors of 10% and more (rather than the 0.1 to 1% duty usually associated with vacuum tubes).

In using such a large number of parallel amplifiers in high-power radar (small radars, of course, need far less power transmitted, far fewer transistor modules), two configurations prevail. In one, the power from large assemblies of such amplifiers is combined and the total power routed to the antenna in a rather conventional waveguide (typically) transmission line; in the other, such amplifiers are part of transmit-receive modules, one at each radiating element of an array-type antenna. This "active array" approach is mentioned again in the antenna discussion

below. *See* MICROWAVE SOLID-STATE DEVICES; TRANSISTOR.

Much simpler radar transmitters are also in wide use, in which a single oscillator (such as a magnetron) is used rather than an amplifier chain or large assembly of transistors. The modulator supplies a cathode-to-anode voltage pulse which causes an RF signal to be generated. *See* MAGNETRON.

The transmitter subsystem, particularly a subsystem that generates lethal high voltages, must include many safety features. Transmitters also require very involved cooling systems, since the overall power conversion efficiency is generally quite low and much damaging heat is generated.

Antenna. Of the four radar subsystems, the antenna is the most visible and often the most indicative of the nature of the radar. Antennas form the directive beam being used and position it in space by either physical movement of the antenna or by purely electronic means in an otherwise fixed structure. In the latter, an array of many discrete radiating elements with individual phase control is commonly used and is called a phased array. Some antennas use mechanical positioning in one dimension and electronic steering in another (for example, rotation in azimuth and electronic steering in elevation).

Antennas for mechanical scan are often of the feedhorn and reflector type; transmitter power is carried to the antenna by waveguide (usually rectangular metal piping) through a duplexer (if the antenna is used on both transmit and receive) and a rotary joint to a radiating horn at the focus of a reflector. The reflector (generally a parabolic section) forms the directive beam for both transmit and receive. The nature of the antenna's pattern, containing both the desired main beam and inescapable sidelobes in every direction, is determined by the way the feedhorn distributes power across the reflector's aperture; on receive, the same structure forms a pattern of directional sensitivity, the receive beam (and its sidelobes). Antennas designed for very low sidelobe levels (even 40 dB or more below the main beam) are a valuable part of modern radar. The design and manufacture of such low sidelobe antennas, requiring such care in the illuminating function across the aperture, is greatly assisted by computer-aided methods. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING; MICROWAVE OPTICS; WAVEGUIDE.

Optical properties are such that the beam formed has a width inversely proportional to the dimension of the illuminated aperture—the wider the antenna (in units of wavelengths), the narrower the beam in that dimension. A vertical fan beam is suitable for two-dimensional (2D) search radar making estimates of just the range and bearing of a target; such radars use, then, oblong reflector shapes (**Fig. 6**). A horizontally oriented fan beam might be used in a height-finding radar; its antenna, pointed to the azimuth of a designated target, nods vertically, scanning the narrow beam over the target to estimate its elevation angle.

Three-dimensional (3D) radars measure azimuth, elevation, and range as they rotate in azimuth, using beams that are narrow in both angles (pencil beams). Often, a stack of pencil beams (requiring separate receivers) is used to measure elevation over a small angle for each transmission, then the stack is repositioned vertically for the next transmission. Such multibeam operation permits coverage and timing that is much more limited in single-beam 3D radars. Beams being narrow in both angles, these radars are characterized by rather square or round antenna shapes, equal in size in both dimensions.

Pencil beams are also used in precision tracking radars, those with dish antennas. Both the conical scanning technique and the multibeam monopulse angle measurement technique were described in the "Tracking" discussion earlier in this article.

Many radars use phased-array antennas, in which the transmitter power is divided among many radiating elements (possibly thousands) and in which the phase of each element is controlled by a computer (**Fig. 7**). In such a fixed-aperture antenna, the beam can be steered in only microseconds from one position to another many beam widths away. Consequently a wide variety of dwell routines can be implemented that are not at all constrained by mechanical rotation. Phased-array antennas are quite complex and expensive; they are used, therefore, only where great demand for very versatile operation exists. Phased arrays usually are designed to form pencil beams since 3D operation and even precision tracking are likely among the modes of operation intended.

Some modern radars use active arrays, in which the final stage of amplification of the transmitted signal takes place in transmit-receive modules located at every radiating element. The transmit-receive

Fig. 6. AN/SPS-49 search radar antenna, which forms a vertical fan beam for the two-dimensional L-band radar widely used in the U.S. Navy. (*Raytheon Co.*)

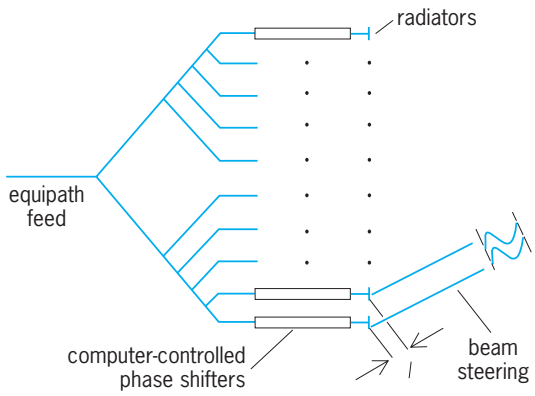

Fig. 7. Basic principle of electronic beam steering with a phased array. Equipath feed is for transmit or receive. All radiators radiate in essentially all forward directions. In beam steering, signals add in phase in the direction shown only if phase shifters are set to establish a phase increment element-to-element to compensate exactly for the path-length difference, *l*, shown.

modules may also contain duplexing circuitry and perform receive preamplification and conversion to intermediate frequency. If receive beam forming is to be done digitally in a computer (Fig. 3*c*) rather than in microwave circuits (Fig. 3*b*), the modules may contain the analog-to-digital converters as well. Still further development of self-adaptive processes, whereby the antenna's pattern can be made inversely proportional to directional sources of interference in its field of view, is well under way. *See* ANALOG-TO-DIGITAL CONVERTER.

Using rather narrow beams provides radars with useful angular (and lateral or cross-range) resolution. The ability to separate closely spaced targets is important in many radar applications, so the maximum practical antenna size (producing the narrowest practical beam width) often limits design. In airborne and spaceborne radars, fine cross-range resolution can be achieved by processing the echoes from the ground in a careful Doppler-sensitive way while the aircraft or satellite moves over a considerable path. The effect is somewhat like using an antenna as long as that segment of the flight path, perhaps a few hundred meters. This synthetic aperture radar (SAR) technique is particularly valuable in imaging the Earth's surface for both military and civil purposes. *See* SYNTHETIC APERTURE RADAR (SAR).

Receiver and signal processor. The signals collected by the antenna are amplified in some radars (those with RF preamplifiers), then converted to the IF for further amplification and processing. Several parallel channels of receivers may be necessary (for the several beams in a stacked-beam radar or for the angle error channels in a monopulse tracker, for example), each requiring all the processing described here.

Many radars require a long pulse for the requisite energy on each transmission, yet require the range resolution of a short pulse. The effect in range resolution of a very short pulse can be achieved by modulating the long pulse with either frequency or phase modulation so that it has the same bandwidth as the very short pulse desired. Then, on receive, the waveform generator supplies a replica of the modulation on the transmitted pulse to the receiver so that a matched-filter action takes place in circuitry called the pulse compressor to produce a narrow output pulse.

In many radar applications it is desired to determine if returns are Doppler-shifted, that is, if certain returns are coming from radially moving targets rather than from stationary clutter. At the microwave frequencies, the Doppler shift is so slight, even for rather high-speed targets, that it cannot be detected on just one short pulse transmission. Instead, multiple pulses must be used over an interval of time that compares to a cycle (at least) of the Doppler shift. A radar equipped to compare the phase of each echo with the phase of the transmission, and then to see if a consistent phase progression exists from pulse to pulse (indicating that the range to the target is changing) is called a "coherent" radar. A simple way of processing such echoes is to compare the phase of each pulse to that of the previous one; this is what a two-pulse canceller circuit does in a moving target indication (MTI) radar. More elaborate cancellers may compare pulses in sets of three or more, with a better shaping of the Doppler response function of the process. More valuable (often) Doppler filtering is achieved by processing the phase data from all the pulses of a single coherent dwell, perhaps scores of pulses. The process involves spectral analysis of that set, wherein as many filters can be formed as there are samples (pulses) and the resolution of the filters (their width and separation in Doppler frequency, roughly speaking) is equal to the reciprocal of the time spanned by the coherent dwell (perhaps several milliseconds). *See* MOVING-TARGET INDICATION.

The fine quantization and processing speeds of modern computers permit these coherent processes to be performed in computerlike circuitry after an appropriate digitizing of the IF signal (as in paired synchronous detection giving two scalar values of the quadrature components of the signal). Most valuable is the process of digital Doppler filtering using discrete and fast Fourier transform methods. Such computer use has also prompted development of adaptive processing wherein heavy clutter that is not uniformly distributed in Doppler (as experienced in airborne radar flying over certain terrain features) can be more effectively countered in the Doppler filtering. *See* FOURIER SERIES AND TRANSFORMS.

Pulse compression and Doppler processing are both coherent processes since complex (amplitude and phase) information is processed. Further steps in the signal processing might now be noncoherent; for example, angle estimation in monopulse radar often uses just the amplitudes of signals in the parallel receiver channels involved. Following Doppler filtering, the amplitude of the signal accumulated in each filter becomes the observable upon which a detection decision is to be made. In older radars, this output might be the video voltage from a radio detector circuit had the filtering been in analog circuitry at the intermediate frequency. In modern

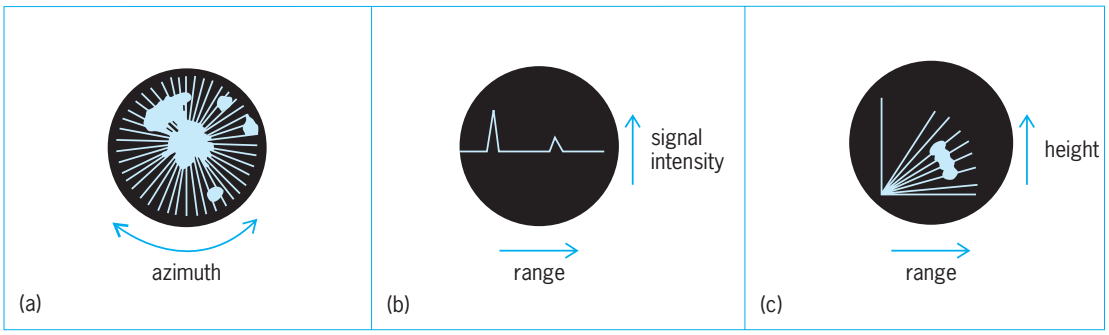

Fig. 8. Cathode-ray-tube radar display formats. (*a*) Type P (plan position indication, or PPI). Range is measured radially from center. (*b*) Type A. (*c*) Range-height indication (RHI).

radar, the digital processing produces numeric values for that output; it is upon these values that the interpretation of the output (to generate contacts, in the search-radar context) is effected in the control and interface apparatus. *See* SIGNAL PROCESSING.

Control and interface apparatus. In simple radars, the operator may observe the output of an elementary video-producing detector on one of several formats on a cathode-ray-tube display. The most common format is the PPI (plan position indication), in which a radial trace (the path of the cathode-ray-tube electron beam exciting the phosphor) is synchronized to the pulse repetition frequency, resulting in radar range being displayed radially from the center (**Fig. 8***a*), and in which this trace is moved around the display, linked electronically to the antenna rotation. The detector output is applied to the intensity grid of the cathode-ray tube so that a bright spot results when a strong video voltage is produced by the detector. The display circuitry may also superimpose upon the radar video certain calibration pulses (range marks and bearing lines, for example), maps of the local terrain, or other symbols that are of importance in the entire system of which the radar is a part.

Other analog displays are illustrated (Fig. 8*b*,*c*) and, together with the PPI format, remain useful display formats for more automated radars. However, such display is generally of computer-generated symbols and data for operator awareness to manage the system, rather than of receiver video for operator performance of the detection process itself.

In the earlier discussion of the "Detection process," the fundamentals of constant false alarm rate (CFAR) automatic detection were described. In the "Tracking" discussion, the association and track filtering processes were described. The computers and attendant displays of this subsystem perform those functions.

Another important part of the control and interface apparatus, particularly in modern multifunction radars, is the mode scheduler. The outputs of the detection and tracking functions combined with external data provided to the radar (mission priorities, dimensions of function-critical regions in the radar surveillance space, information from other sensors of the network, and so on) must feed into a control logic to produce a list of what next should be done. That list must be prioritized and fed to a control apparatus to operate the radar itself. Further, this scheduling must be done within the radar's time and energy constraints. Often such control is planned in short time "modules," planning periods of perhaps a few tens of milliseconds, enough time for a dozen or so radar dwells. The response time from a radar result being interpreted and a new action taken might need to be a small fraction (a tenth) of a second—the most critical loop-closure time in the various modes of the radar will bear heavily upon the dimensioning of the time modules and the architecture of the scheduler.

Finally, the monitoring of the state of the radar itself is an important function of this subsystem. Modern radars have, of necessity, much built-in test equipment (BITE). There may be a resource console where data are collected from various monitoring points throughout the radar equipment. In military radars, this function is further burdened with battle damage assessment and the direction of compensating measures if further operation is possible. This function may include conducting readiness tests from time to time. Internal test target stimuli are affected and results analyzed for assuring acceptable performance. Physical maintenance procedures can be based on these results.

Less sophisticated testing of radars (and certainly of those with less BITE than described above) is still carried out by trained technicians in the field. A rather fundamental measure of a radar's state is called the performance factor, the ratio of the power output of a radar's transmitter and the minimum discernable signal (MDS) of its receiver. For example, a radar with transmitter power (pulse power) of 1 megawatt and an MDS of 10^{-14} watt (an MDS of -140 dBw or -110 dBm -110 decibels below a milliwatt) has a performance factor of 200 dB (that is, of 10^{20}). A convenient way to test for this value is built into many radars; separate test equipment can be used in others.

The "echo box" is familiar to all technicians who have worked in the field with radar. It consists of a resonant cavity into which can be fed a set fraction of the transmitter's power (by a directional coupler built into the radar for this purpose). The well-calibrated and decaying ringing of the resonating pulse is fed into the receiver, where the receiver's response to it (at an appropriate test point) can be

viewed on an oscilloscope. The length of time after the pulse that the ringing remains visible above the receiver noise background is a direct measure of the radar's performance factor. Some echo boxes are equipped with precise tuning and a power meter inside the cavity so that a more careful spectral analysis of the transmitter's performance can be made across the radar's operating band. Knowing the power level also permits the receiver's MDS to be determined. These performance values differ day-to-day in actual radar operation, and this relatively simple diagnostic tool is helpful in keeping the radar working well.

Robert T. Hill

Bibliography. D. K., Barton, *Modern Radar System Analysis*, Artech House, Norwood, MA, 1988; L. V. Blake, *Radar Range Performance Analysis*, Artech House, Norwood, MA, 1986, reprint 1991; R. Buderi, *The Invention that Changed the World*, Simon and Schuster, New York, NY, 1996; F. E. Nathanson, *Radar Design Principles*, SciTech Publishing, Mendham, NJ, 1999; S. Sabatini and M. Tarantino, *Multifunction Array Radar*, Artech House, Norwood, MA, 1994; M. I. Skolnik, *Introduction to Radar Systems*, 3d ed., McGraw-Hill, New York, NY, 2001; M. I. Skolnik (ed.), *Radar Handbook*, 2d ed., McGraw-Hill, New York, NY, 1990; W. D. Wirth, *Radar Techniques Using Array Antennas*, Institution of Electrical Engineers, London, 2001.

Radar-absorbing materials

Materials that are designed to reduce the reflection of electromagnetic radiation from a conducting surface in the frequency range from approximately 100 MHz to 100 GHz. The level of reduction that is achieved varies from a few decibels to greater than 50 dB, reducing the reflected energy by as much as 99.999%.

Theory. The reflection of electromagnetic radiation from the surface of any material can be calculated from Maxwell's equations if the electromagnetic properties are known. The relevant properties are the magnetic permeability μ_m and the dielectric permitivitty ε_m. Both μ_m and ε_m may be expressed as complex numbers, in which case the imaginary part represents the magnetic loss or dielectric loss, respectively. These parameters in turn define the characteristic impedance Z_m, given by Eq. (1), and the propagation coefficient γ_m, given by Eq. (2) [where

$$Z_m = \sqrt{\frac{\mu_m}{\varepsilon_m}} \tag{1}$$

$$\gamma_m = \frac{2\pi_c}{\lambda}\sqrt{\mu_m \varepsilon_m} \tag{2}$$

λ is the wavelength of the radiation and c is the speed of light], which determine the reflection at the material interface and the attenuation of the radiation in the material, respectively. In the case of radar-absorbing materials, the objective is to minimize the reflection by matching the input impedance Z_i of the material layer to the characteristic impedance Z_0 of free space. *See* ELECTRICAL IMPEDANCE; ELECTROMAGNETIC RADIATION; MAGNETISM; MAXWELL'S EQUATIONS; PERMITTIVITY.

Equation (3) gives the input impedance of a layer

$$Z_i = Z_m \tanh(\gamma_m d) \tag{3}$$

of material of thickness d directly adjacent to a conducting surface. The energy reflection coefficient is given by expression (4). Zero reflection will occur

$$\left(\frac{1 - Z_i/Z_0}{1 + Z_i/Z_0}\right)^2 \tag{4}$$

in the case where Z_i is equal to Z_0.

In the simple case where the attenuation in the material is such that there is no reflected energy from the conducting surface, the hyperbolic tangent (tanh) function in Eq. (3) tends to unity, Z_i equals Z_m, and the only reflection is at the free-space/material interface. Zero reflection is obtained when the characteristic impedance of the material and free space are identical, which also implies that Eq. (5) is satisfied.

$$\frac{\mu_m}{\varepsilon_m} = \frac{\mu_0}{\varepsilon_0} \tag{5}$$

This condition must be met for μ_m and ε_m as complex numbers. Since either μ_m or ε_m must contain an imaginary component in order to attenuate the radiation, it necessarily follows that in this case both μ_m and ε_m must be complex numbers with both the real parts and imaginary parts in the same ratio as the ratio of μ_0 and ε_0. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION.

Design methods. In theory, this simple solution offers the possibility of a perfect absorber. In practice, materials have not been found which will give such a good impedance match over an appreciable frequency range. It is therefore necessary to adopt specific design methods to make practical absorbing materials.

Two methods have been widely adopted in order to produce such absorbers. The first is to avoid a discrete change of impedance at the material surface by gradually varying the impedance. The removal of the discontinuity at the surface allows the microwave energy to be transmitted into the absorbing medium without reflection. This transition from the impedance of free space to that of the bulk material is commonly achieved by a geometric profile. The carbon-loaded foam pyramids used as the lining of anechoic chambers are typical of this type of absorber. To produce such absorbers, it is necessary in practice to taper the material over distances which are large compared with the wavelength of the frequencies to be absorbed. Therefore, practical absorbers of this type giving greater than 20 dB absorption vary in thickness from about 0.8 in. (2 cm) at 10 GHz and above to 6 ft (2 m) at 100 MHz and above. **Figure 1** shows the approximate level of absorption at normal incidence as a function of the thickness of the material in wavelengths of the incident radiation. The absorber performance improves with increasing thickness until the point is reached

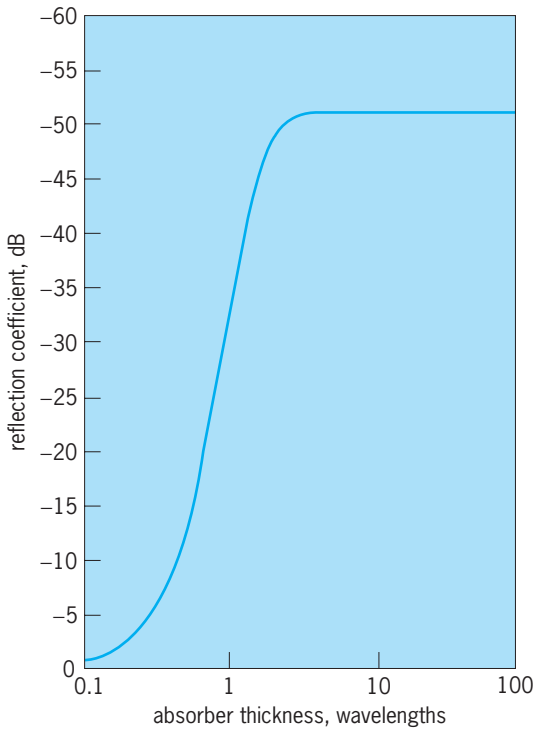

Fig. 1. Absorption of a thick, profiled lossy layer at normal incidence as a function of the thickness of the material in wavelengths of the incident radiation.

where all of the energy that enters the material is absorbed and only the front-face reflection is left. In the example given, the front face reflection is −50 dB relative to a perfect reflector, which is a typical figure for anechoic chamber material. While this type of absorber is capable of producing a very high degree of absorption over a broad bandwidth, it is at the same time a relatively thick material. *See* ANECHOIC CHAMBER.

The second method of absorber design produces much thinner absorbing layers which are capable of producing good absorption (≥25 dB) with restricted bandwidths. These materials achieve the absorption by a combination of attenuation within the material and destructive interference at the interface. The electromagnetic properties and the thickness of the layer are such that the initial reflected wave and the sum of the emergent rays (e_1, e_2, e_3, and so forth) resulting from the multiple reflections within the material are equal in magnitude and opposite in phase. The thickness of the layer is close to a quarter-wavelength at the frequency of operation, giving a 180° phase difference between the interface reflection and the emergent waves (**Fig. 2**). *See* INTERFERENCE OF WAVES.

The characteristic impedance and the propagation coefficient of the materials change with the angle of incidence and polarization of the radiation on the material surface. As a result, the input impedance of the layer and hence the level of absorption are angle- and polarization-dependent. The performance of a typical quarter-wave absorber designed for normal incidence is shown in **Fig. 3**. The absorption against frequency is shown when the electric field (E) vector is parallel and perpendicular to the plane of incidence for a range of angles. *See* POLARIZED LIGHT.

It is possible to tune these materials to any desired center frequency over most of the range from 100 MHz to 100 GHz. When the absorber is optimized for an angle other than normal, it will generally be for only one polarization. While this type of material intrinsically provides a relatively narrow bandwidth frequency performance, it is also possible to broaden the effective bandwidth by various techniques such as multiple-layer absorbers. With two layers of material it is possible to tune one absorber to two different frequencies. By placing these two frequencies appropriately, for example, within one

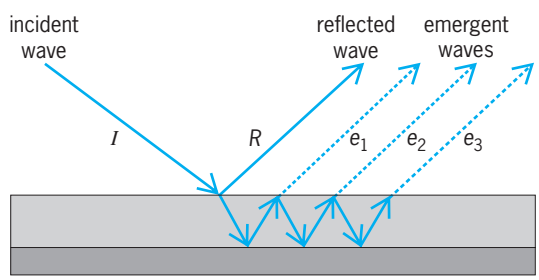

Fig. 2. Mode of operation of a lossy quarter-wave layer. There is a mutual cancellation of the reflected wave and the emergent waves e_1, e_2, e_3, and so forth.

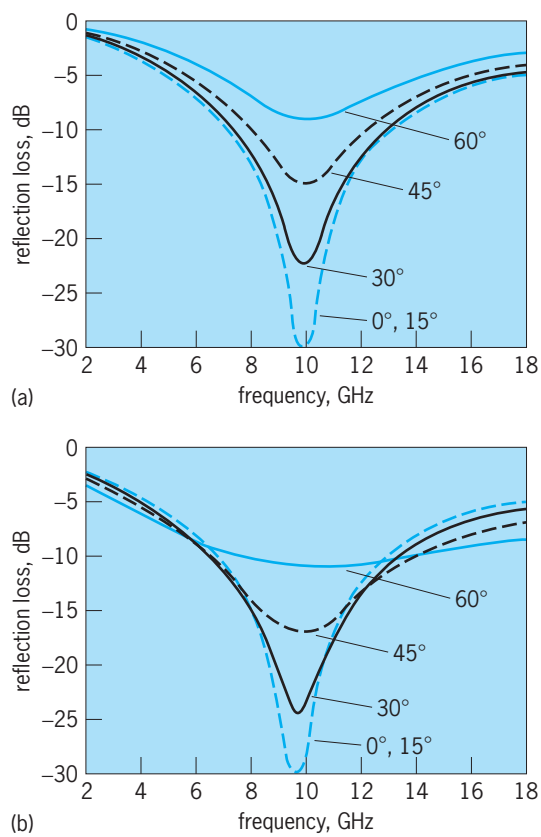

Fig. 3. Performance of a lossy quarter-wave layer designed for optimum performance when the incident radiation is normal to the surface as a function of frequency, polarization, and angle of incidence. (*a*) Perpendicular polarization. (*b*) Parallel polarization. Angles of incidence are indicated.

octave of each other, a broadband absorber is obtained in practice.

Absorbers based on destructive interference can also be made by placing a thin resistive layer one-quarter of a wavelength from a conducting surface. In this case, we consider the system in terms of admittance, which is the reciprocal of impedance. The admittance of free space is 1/377 siemens, and that of a quarter-wavelength space in front of a reflector is zero. Therefore, a thin layer of resistance 377 ohms spaced one-quarter wavelength from a reflector will match the admittance of free space, and there will be no reflection. The absorption characteristics with frequency are similar to those shown in Fig. 3.

Uses. Microwave-absorbing materials are widely used both within the electronics industry and for defense purposes. Their uses can be classified into three major areas: (1) for test purposes so that accurate measurements can be made on microwave equipment unaffected by spurious reflected signals (the most common example is the anechoic chamber); (2) to improve the performance of any practical microwave system by removing unwanted reflections which can occur if there is any conducting material in the radiation path; and (3) to camouflage a military target by reducing the reflected radar signal. *See* ELECTRONIC WARFARE; MICROWAVE; MICROWAVE MEASUREMENTS; REFLECTION OF ELECTROMAGNETIC RADIATION.

S. B. Morris

Bibliography. E. F. Knott, J. F. Shaeffer, and M. T. Tuley, *Radar Cross Section*, 2d ed., 1993; A. R. Von Hippel, *Dielectrics and Waves*, 2 vols., 1955, reprint 1995.

Radar astronomy

A powerful astronomical technique that uses radar echoes to furnish otherwise unavailable information about bodies in the solar system. By comparing a radar echo to the transmitted signal, information can be obtained about the target's size, shape, topography, surface bulk density, spin vector, and orbital elements, as well as the presence of satellites and in certain situations the target's mass and density. While other astronomical techniques rely on passive measurement of reflected sunlight or naturally emitted radiation, the illumination used in radar astronomy is a coherent signal whose polarization and time modulation or frequency modulation are tailored to meet specific scientific objectives. Through measurements of the distribution of echo power in time delay or Doppler frequency, radar achieves spatial resolution of a planetary target despite the fact that the radar beam is typically much larger than the angular extent of the target. This capability is particularly valuable for asteroids and planetary satellites, which appear as unresolved point sources through optical telescopes. Moreover, the centimeter-to-meter wavelengths used in radar astronomy readily penetrate cometary comas and the optically opaque clouds that conceal Venus and Titan, and also permit determination of near-surface roughness (abundance of wavelength-scale rocks), bulk density, and metal concentration in planetary regoliths. *See* ASTEROID; SATELLITE (ASTRONOMY); SATURN; VENUS.

Telescopes. A radar telescope is essentially a radio telescope equipped with a high-power transmitter (a klystron vacuum-tube amplifier) and specialized instrumentation that links the transmitter, low-noise receiver, high-speed data-acquisition computer, and antenna together in an integrated radar system. Planetary radars, which must detect echoes from targets at distances from about 10^6 km (10^6 mi) for closely approaching asteroids and comets to more than 10^9 km (nearly 10^9 mi) for Saturn's rings and satellites, are the largest and most sensitive radars on Earth. *See* KLYSTRON; RADIO TELESCOPE.

The two active planetary radar facilities are the Arecibo (radar wavelengths of 13 and 70 cm) and Goldstone (3.5 and 13 cm) instruments; for each, the shorter wavelength provides the greater sensitivity. The Arecibo radio-radar telescope in Puerto Rico consists of a 305-m-diameter (1000-ft) fixed reflector, the surface of which is a section of a 265-m-radius (870-ft) sphere. Movable feeds suspended from a triangular platform 130 m (427 ft) above the reflector can be aimed toward various positions on the reflector, enabling the telescope, located at 18°N latitude, to point up to 20° from the zenith. The Goldstone radar in California is part of the National Aeronautics and Space Administration's Deep Space Network. The Goldstone main antenna, DSS-14, is a steerable, 70-m (230-ft) parabolic reflector with horn feeds. Arecibo is the more sensitive of the two instruments, with twice the range of Goldstone, but Goldstone can track targets continuously for much longer periods and has access to the whole sky north of −40° declination. Bistatic (two-station) experiments employing transmission from DSS-14 and reception of echoes at the 27-antenna Very Large Array (VLA) in New Mexico have synthesized a beamwidth as small as 0.24 second of arc, versus 2 minutes of arc for single-dish observations with Arecibo or Goldstone. Bistatic experiments using DSS-14 transmissions and reception of echoes at DSS-13, a 34-m (112-ft) antenna 22 km (14 mi) away, have been conducted on several very close targets. Bistatic observations between Arecibo and Goldstone, or using transmission from Arecibo or Goldstone and reception at the 100-m (328-ft) Greenbank Telescope (GBT) in West Virginia, have also proven very powerful.

Techniques. A typical transmit-receive cycle consists of signal transmission for a duration close to the round-trip light time between the radar and the target, that is, until the first echoes are about to return, followed by reception of echoes for a similar duration. The target's apparent radial motion introduces a continuously changing Doppler shift into the echoes. This shift is removed by continuously tuning the instrument according to an ephemeris based on an orbit calculated from optical (and perhaps radar) astrometric observations. In continuous-wave observations, a nearly monochromatic signal is transmitted and the distribution of echo power is measured as a

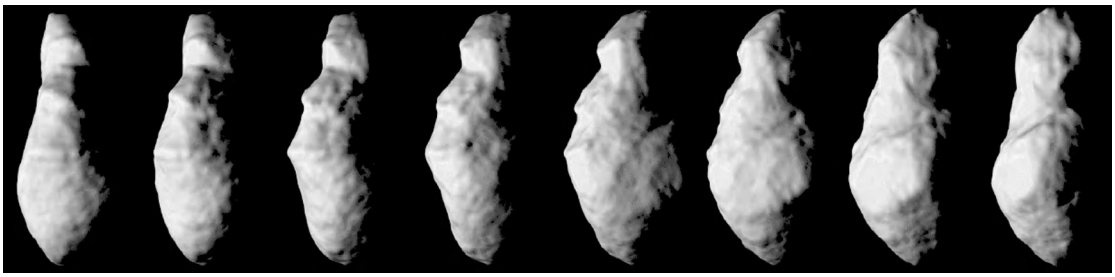

Fig. 1. A model of the shape of near-Earth asteroid 4179 Toutatis based on radar delay-Doppler images. The 4.5-km-long (2.8-mi) object has complex linear features as well as circular craterlike structures down to the ~34-m (112-ft) resolution limit. The noncraterlike features may arise from complex interior configurations involving monolithic fragments with various sizes and shapes, presumably due to collisions in various energy regimes. (*R. S. Hudson, S. J. Ostro, and D. J. Scheeres, High-resolution model of asteroid 4179 Toutatis, Icarus, 161:346–355, 2003*)

function of frequency. The resultant echo spectrum can be thought of as a one-dimensional image, or a brightness scan across the target through a slit parallel to the target's apparent spin vector. *See* DOPPLER EFFECT.

Range resolution can be obtained by using a coherent pulsed continuous waveform, but in practice, because of engineering considerations associated with klystrons, a pulsed waveform is simulated by encoding a continuous-wave signal with a preset sequence of 180° phase reversals. Fourier transformation of time samples taken at the same position within each of many successive range profiles yields the echo power spectrum for the corresponding range cell on the target. Spectral analysis of echoes from multiple range cells yields a delay-Doppler image. If the target's shape is known, it is a straightforward matter to convert that image into a radar brightness map. Heterogeneities in brightness can be introduced by variations in large-scale surface tilts, small-scale roughness, or surface bulk density. By measuring the echo's polarization state, the nature of the scattering process and the severity of wavelength-scale roughness near the surface can be deduced, thereby refining the interpretation of reflectivity measurements. If the shape is unknown, as with asteroids and comets, it can be estimated from delay-Doppler images of the target in a variety of orientations (**Fig. 1**). *See* FOURIER SERIES AND TRANSFORMS; POLARIMETRY.

Observational results. Radar-detected planetary targets include the Moon, Mercury, Mars, Venus, Phobos, the Galilean satellites, Saturn's rings, Titan and six other Saturnian satellites, nine comets, 95 main-belt asteroids, and 169 near-Earth asteroids (including 16 binaries and 95 potentially hazardous objects). The radar signatures of these bodies are extraordinarily diverse. For example, for each of the two most basic radar properties (albedo and circular polarization ratio), estimated values span two orders of magnitude. *See* ALBEDO; BINARY ASTEROID; COMET; JUPITER; MARS; MERCURY (PLANET).

Radar observations have revealed a metallic near-Earth asteroid, the dumbbell shape and metallic composition of a large main-belt asteroid, and the presence of large-particle swarms (with particles at least 1 cm or 0.4 in. in diameter) around many comets (including Halley's Comet). Radar has discovered the existence of contact binaries, binary systems, nonprincipal-axis spin states, and near-Earth-asteroid rotation periods ranging from minutes to weeks. Radar also has achieved the first detection of the nongravitational, thermal-recoil "Yarkovsky" acceleration of an asteroid, using the measurement to estimate the asteroid's mass. Radar's unique capabilities for trajectory refinement and physical characterization give it a natural role in predicting and preventing collisions with small bodies.

Other radar experiments have disclosed the extraordinary diversity of Mars's surface at decimeter-to-meter scales and have produced startling evidence for radar-bright, probably icy, polar caps on Mercury. Thirty years after the radar discovery of Venus's slow retrograde rotation, a radar instrument on the Venus-orbiting *Magellan* spacecraft mapped almost all of that planet's surface with striking geologic clarity. The first four-station radar interferometry assisted selection of the landing sites for the *Mars Exploration Rovers*, and a novel two-station radar speckle displacement technique has produced ultraprecise measurements of Mercury's spin state that indicate a liquid core.

Titan is the primary focus of the *Cassini* spacecraft, which went into orbit about Saturn in 2004. *Cassini*'s radar instrument, a 13.8-GHz (2.2-cm) synthetic aperture radar (SAR) imager, altimeter, scatterometer, and radiometer, is to operate during around half of the 44 Titan flybys, covering about a fifth of the surface with imaging resolution of 2 km (1.2 mi) or finer. Scatterometry observations

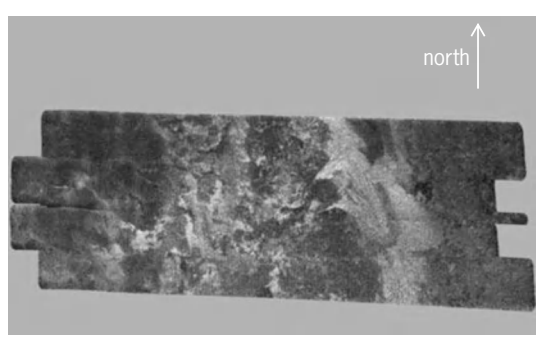

Fig. 2. A 150 × 250-km (90 × 150-mi) portion of *Cassini*'s first synthetic aperture radar (SAR) image strip on Titan, centered at 50°N, 82°W. The smallest details seen on the image are about 300 m (1000 ft) across. (*NASA/JPL*)

are to cover most of the surface, albeit at resolution no finer than tens of kilometers, and a limited number of short altimetry tracks are to give regional topographic information. *See* SYNTHETIC APERTURE RADAR (SAR).

The first Titan radar images (**Fig. 2**) reveal a complex surface with low relief and a wide variety of surface features and terrains, including flows that might be cryovolcanic and parallel linear features that might be eolian or tectonic. Radar-bright sinuous channels and associated fanlike structures suggest the presence of transported material with different radar properties. Dark patches are tentatively identified as low-dielectric materials such as hydrocarbon lakes, organic sludge, water ices, or ammonia hydrates. The small number of features unambiguously identified as craters strongly suggests a young, geologically active surface, possibly modified as well by deposition of atmospheric chemical products. *See* PLANET; RADAR; SPACE PROBE. Steven J. Ostro

Bibliography. G. J. Black, D. B. Campbell, and P. D. Nicholson, Icy Galilean satellites: Modeling radar reflectivities as a coherent backscatter effect, *Icarus*, 151:167–180, 2001; J. K. Harmon, P. J. Perillat, and M. A. Slade, High-resolution radar imaging of Mercury's north pole, *Icarus*, 149:1–15, 2001; J. L. Margot et al., Topography of the lunar poles from radar interferometry: A survey of cold trap locations, *Science*, 284:1658–1660, 1999; S. J. Ostro, Planetary radar astronomy, *Rev. Mod. Phys.*, 65:1235–1279, 1993; S. J. Ostro et al., Asteroid radar astronomy, in *Asteroids III* (W. Bottke et al., eds.), pp. 151–168, University of Arizona Press, Tucson, 2002; G. H. Pettengill et al., *Magellan*: Radar performance and data products, *Science*, 252:260–265, 1991.

Radar meteorology

The application of radar to the study of the atmosphere and to the observation and forecasting of weather. Meteorological radars transmit electromagnetic waves at microwave and radio-wave frequencies. Water and ice particles, inhomogeneities in the radio refractive index associated with atmospheric turbulence and humidity variations, insects, and birds scatter radar waves. The backscattered energy received at the radar constitutes the returned signal. Meteorologists use the amplitude, phase, and polarization state of the backscattered energy to deduce the location and intensity of precipitation, the wind speed along the direction of the radar beam, and precipitation type (for example, rain or hail). *See* METEOROLOGICAL RADAR; RADAR; WEATHER FORECASTING AND PREDICTION.

Echoes from atmosphere. Meteorological radars normally transmit short pulses of energy. The signal received is the combination of electromagnetic waves backscattered toward the radar by each scattering element within the volume illuminated by a pulse. If the scattering elements fill the beam, which is normally true for meteorological applications, this volume is determined from the angular beamwidth θ and the pulse duration τ, and is givenby Eq. (1),

$$V = \frac{\pi}{4}(R\theta)^2\frac{c\tau}{2} \tag{1}$$

where c is the speed of propagation of the waves and R is the distance from the radar to the pulse volume.

If the beam width is 1°, the pulse duration 1 microsecond, and the range 50 km (30 mi), the scattering volume equals about 9×10^7 m^3 (3.2×10^9 ft^3). For precipitation, the volume may contain more than 10^{11} raindrops or ice particles. The contribution of each scattering element in the volume adds in phase to create the returned signal. The returned signal fluctuates from one pulse to the next because the scattering elements move relative to one another. For this reason, the returned signals from many pulses are averaged to determine the average received power, from which the intensity of the precipitation can be estimated. The pulse-to-pulse phase change, due to the average motion of the scattering elements along the radar beam, is also used by Doppler radars to determine air motion toward or away from the radar. *See* DOPPLER RADAR.

The radar range equation for distributed targets such as raindrops may be written as Eq. (2), where

$$P_r = \frac{P_t G^2 \lambda^2}{(4\pi)^3 R^4} V\eta \tag{2}$$

P_r is the average received power, P_t the transmitted power, G the antenna gain, λ the wavelength, R the range, V the scattering volume, and η the reflectivity, which is assumed to be constant over the scattering volume. The radar cross section σ of a spherical water or ice particle whose diameter D is small compared to the wavelength λ is given by the Rayleigh scattering law, Eq. (3), where $|K|^2$, a dimension-

$$\sigma = \frac{\pi^5}{\lambda^4}|K|^2 D^6 \tag{3}$$

less factor that depends on the dielectric properties of the particle, is approximately equal to 0.93 for water and 0.21 for ice. The radar reflectivity of clouds and precipitation is obtained by summing the cross sections of all the particles in the scattering volume and is written as Eq. (4).

$$\eta = \frac{\pi^5}{\lambda^4}|K|^2 Z \tag{4}$$

The radar reflectivity factor Z is defined by Eq. (5), wherethe summation is over the scattering

$$Z = \frac{\sum D^6}{V} \tag{5}$$

volume. It is customary to use cubic meters as the unit for volume and to measure the particle diameters in millimeters, so that Z has conventional units of mm^6/m^3. The radar reflectivity factor varies from about 10^{-5} mm^6/m^3 to 10 mm^6/m^3 in nonprecipitating clouds, upward to about 10^6 mm^6/m^3 in rain, and as high as 10^7 mm^6/m^3 in large hail. In a mixture of rain and clouds, the raindrops dominate the returned

signal because of the sixth-power weighting on diameter. This is true even though the number of raindrops in a cubic meter—typically about 10^3—is perhaps 10^5 fewer than the number of cloud droplets. The radar reflectivity factor Z is of significant meteorological importance because of its relationship to the raindrop size distribution within the scattering volume. In meteorological applications, the averaged returned power is measured and converted to values of Z using Eqs. (1)–(4). Since Z varies across many orders of magnitude, a logarithmic scale, defined by Eq. (6), is used to display the radar reflectivity factor.

$$\mathrm{dBZ} = 10 \log_{10} \left(\frac{Z}{1\ \mathrm{mm}^6/\mathrm{m}^3} \right) \tag{6}$$

Typical radar depictions of weather systems that appear in the media (**Colorplate 1**) display the radar reflectivity factor in logarithmic units. Displays of the radar reflectivity factor overlaid on maps quickly allow the meteorologist to determine the location and intensity of precipitation. The term "radar reflectivity factor" is often shortened to "reflectivity" by meteorologists, although radar specialists reserve this terminology for η.

Raindrops, ice particles, and other objects are not the only source of radar echoes. Reflections from the optically transparent atmosphere can be observed by sensitive radars. These echoes are associated with spatial fluctuations in the atmospheric refractive index created by turbulence in the presence of temperature and humidity gradients. The clear-air radar reflectivity produced by turbulent mixing is given by Eq. (7), where C_n^2 is a quantity that characterizes the strength of the refractive index irregularities.

$$\eta = \frac{5}{6}\pi^{4/3}4^{-5/3}\lambda^{-1/3}C_n^2 = 0.38\lambda^{-1/3}C_n^2 \tag{7}$$

The different dependence on wavelength in Eq. (4) and Eq. (7) shows that the radar sensitivity to turbulent scattering decreases at a much slower rate with increasing wavelength than does the sensitivity to scattering by small particles. The different wavelength dependence is the reason why radars with rather long wavelengths are favored for clear-air studies. The clear-air signal is important for retrieving air motions in some Doppler radars.

The Doppler shift refers to a small change in frequency of the returned signal associated with motion of scattering elements toward or away from the radar. Scattering elements move with the ambient wind. In meteorological applications, measurement of this small frequency shift provides an estimate of the along-beam, or radial, wind component, from which information about winds and other motions of interest can be extracted. To obtain Doppler information from a radar echo, the transmitted signal must have suitable characteristics and special signal processing must occur in the receiver. In particular, the pulse-to-pulse change in the phase of the returned signal must be retained and extracted. Ambiguity problems exist in Doppler radar systems because the same series of sampled phase values can be produced by more than one target radial speed. The maximum unambiguous radial speed measurable by a Doppler radar, termed the Nyquist velocity v_n, is given by Eq. (8), where F is the pulse repetition frequency.

$$v_n = \frac{\lambda F}{4} \tag{8}$$

Radial wind speeds that exceed the Nyquist velocity are folded back into the observable range. Algorithms exist to "unfold" the radial wind measurements to determine the true radial winds for most meteorological situations.

Conventional radars designed for meteorological measurements radiate and receive electromagnetic waves with a fixed polarization, a single fixed orientation (horizontal) of the electric field. Special polarization diversity radars allow variation of the polarization state (horizontal and vertical) of the transmitted or received signal. Polarization measurements are sensitive to particle orientation. As large raindrops fall, they have a tendency to flatten into an oval shape. Hail tumbles as it falls and generally has no preferred orientation. Ice crystals may or may not have preferred orientation, depending on their shape. The radar cross sections of these precipitation particles depend on the orientation of the impinging electric field. Polarization measurements therefore have the potential to discriminate hail from rain, and to identify predominant types of ice particles in clouds. Radars have been developed that switch polarization between horizontal and vertical on successive pulses, or transmit and receive circular and elliptically polarized waves. Researchers are also exploring methods for using polarization techniques to estimate rainfall. *See* PRECIPITATION (METEOROLOGY).

Applications. Radar meteorologists use analysis and animation techniques to track and predict the movement of storms, estimate winds and precipitation, and identify dangerous weather phenomena, such as tornadoes and hail.

Storm development and motion. Thunderstorms appear on radar as cores of high reflectivity, while widespread precipitation appears as larger areas of weaker reflectivity. Widespread precipitation areas, viewed on a large scale, often are organized in bands with widths ranging from tens to hundreds of kilometers. Meteorologists use displays of reflectivity to observe storms developing and to track storms through their life cycles. Meteorologists use individual radars to determine the location, intensity, and dimensions of individual storms and storm complexes. Data composites from radar networks show how precipitation regions are organized around their parent weather systems, which may extend over thousands of kilometers. From animations of radar images, meteorologists can determine whether storms are intensifying or weakening, and the speed and direction of their motion, enabling the prediction of storm passage over critical areas up to several hours in advance. *See* STORM; STORM DETECTION.

Severe thunderstorm detection. Severe thunderstorms may produce large hail, damaging straight-line winds,

or tornadoes. Radars, particularly Doppler, allow meteorologists to issue timely warnings as destructive thunderstorms approach populated regions. A severe thunderstorm will often develop a hooklike appendage in the reflectivity field. If a tornado develops, it will often appear near the center of the hook. The radial velocity field from a Doppler radar can provide clear indications of storm rotation. Rotation appears as a tight couplet of adjacent strong inbound and outbound radial motions (**Colorplate 2**). Animations of the reflectivity and radial velocity fields are used to track hook positions and radial velocity couplets, permitting meteorologists to issue specific warnings concerning the arrival of dangerous conditions. *See* THUNDERSTORM.

The reflectivity within a storm is another indicator of its severity and potential destructiveness. Values of reflectivity exceeding 55 dBZ are often associated with hail. Large hail, which can reach diameters of 1–6 cm (0.4–2.5 in.), appears on radar with reflectivity approaching 70 dBZ. Radars with polarization switching capability can detect hail by comparing the reflectivity measured at horizontal (Z_H) and vertical (Z_V) polarization. When Z_H is large, and the ratio of $Z_H/Z_V \cong 1$, the precipitation particles are approximately circular and are likely to be hail. For large values of the ratio, the particles are oval shape and are likely to be rain.

Precipitation measurement and hydrology. The value of the radar reflectivity factor Z is a general indicator of precipitation intensity; however, no exact relationship between Z and precipitation rate P exists. Many studies have found that P and Z are approximately related by Eq. (9). The coefficient a and the

$$Z = aP^b \tag{9}$$

exponent b take on different values depending on the precipitation type. For moderate, widespread rain, for example, a is typically 200 and b is 1.6 if P is measured in mm/h and Z is in mm^6/m^3. Two methods have been used for establishing relationships of this kind. In one method, a radar is used to take frequent measurements of the reflectivity just above one or more rain gages located on the ground. Direct comparisons can then be made between Z and P to determine the most suitable values of a and b in Eq. (9). The second method is indirect, based on measurements of the sizes and concentrations of raindrops in many rainfall samples. For a given sample, both Z and P can be calculated. Many samples corresponding to rainfalls of different intensities yield a number of combinations of Z and P from which a best-fit relation of the form of Eq. (9) can be derived. Owing to uncertainties in the radar calibration, variability in the empirical quantities a and b, and other sources of error, radar estimates of the short-term precipitation rate at a point can deviate by more than a factor of 2 from gage measurements. Some of these errors are random, so when averaged over larger areas and longer times, radar estimates of the total accumulated rainfall are more accurate. Hydrologists use precipitation estimates from radar to estimate the total precipitation over a watershed and stream runoff. Because radar-based precipitation estimates are almost instantaneously available to hydrologists, radar has become an effective tool for issuing flash flood warnings. *See* HYDROLOGY; HYDROMETEOROLOGY; METEOROLOGICAL INSTRUMENTATION.

Aviation meteorology. Thunderstorms and associated phenomena such as hail and lightning pose a direct threat to aviation. Radar is used routinely to direct air traffic around and away from thunderstorms. A particular hazard for aircraft approaching or departing an airport is wind shear, a sudden sharp change in wind speed along the flight path. Wind shear occurs along gust fronts, the leading edge of fast-moving rain-cooled air rushing outward from precipitating areas of thunderstorms. Wind shear also occurs in downbursts, localized intense downdrafts that induce an outburst of damaging winds on or near the ground. Downbursts can develop when rain falling from a storm passes though a dry layer and evaporates, cooling air in the dry layer and causing it to descend.

Aircraft encountering wind shear can rapidly lose air speed, which reduces lift and may cause loss of control. Many disastrous aircraft accidents have been attributed to low-level wind shear encountered on approach or departure. Doppler radars are now used at many airports to detect wind shear. Radial velocity measurements from the radar are processed by computer algorithms that rapidly deliver information to air-traffic controllers about the expected air speed loss on approach or departure. Relayed to pilots, this information allows an informed decision about continuing or aborting a takeoff or landing. *See* AERONAUTICAL METEOROLOGY; LIGHTNING.

Cloud and precipitation physics. Current understanding of the physics of precipitation processes derives from direct measurement of cloud properties made by aircraft, remote measurements of clouds made by radar and other remote sensors, laboratory measurements, and theoretical studies. Precipitation forms through two distinct processes, the first involving the collision and coalescence of cloud droplets, and the second involving the formation of ice particles. Radar provides one of the best methods for detecting precipitation particles in clouds because of the sixth-power relationship between particle diameter and reflectivity [Eq. (5)]. Radar measurements, combined with measurements of the temperature and height at which early radar echoes form, have been used to study the relative importance of these two processes in different cloud types. The subsequent growth of precipitation particles can also be inferred from changes in the reflectivity as precipitation falls to the Earth's surface (see **illus.**). Radar can also be used to identify areas of new precipitation formation in widespread clouds. Early studies of radar echoes showed that precipitation in cold stratiform clouds often formed in localized regions near cloud top. These regions, called generating cells by radar meteorologists, produce streams of ice particles which

fall into the lower part of the cloud while growing larger (**Colorplate 3**). *See* CLOUD; CLOUD PHYSICS.

Radar can provide information concerning the type and fall speed of precipitation particles in clouds. The fall speed of precipitation particles is related to their size and shape, and to the motion of the air. In clouds with relatively weak vertical motions of a few cm/s or less, the distribution of fall speeds measured by a vertical-pointing Doppler radar can be related to the particle size distribution. Measurements of the change in the spectrum of fall speeds as particles descend from the cloud top to the ground provide information about precipitation growth.

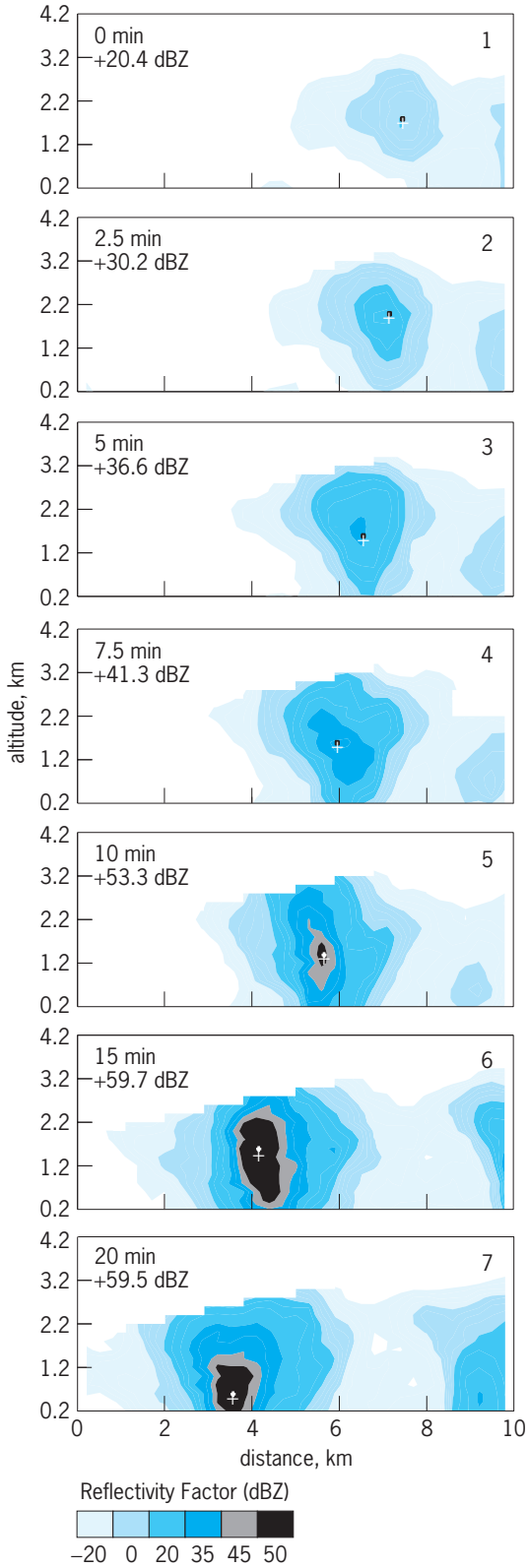

Vertical cross section showing the evolution of the radar reflectivity factor (dBZ) during a 20-min period in a shower located near the coast of the island of Hawaii.

The most prominent feature of radar echoes in widespread clouds is a "bright band" of high reflectivity at the altitude of the melting level. Several effects explain the intensification of the reflectivity within the bright band. As snowflakes begin to melt, their reflectivity rapidly increases because the dielectric factor $|K|^2$ for water exceeds that for ice [Eq. (3)]. Snowflakes become more compact as they continue to melt and descend, finally collapsing into raindrops. The drops have a smaller radar cross section and fall faster than snowflakes, resulting in a more dilute concentration of smaller particles below the melting layer than within it. These effects reduce the reflectivity below the melting layer.

Polarization diversity radars provide a means of discriminating ice and water in clouds and, in some cases, permit identification of the predominant type of ice particles in different cloud regions. The most important application of polarization techniques is hail detection. A second potential application is the detection of liquid water droplets in regions of widespread clouds where the temperature is below 0°C (32°F). Detection of these supercooled water droplets is important because they can cause airframe icing.

Wind measurements. Radars operating at longer wavelengths, called wind profilers, are sensitive to turbulent irregularities in the radio refractive index associated with temperature and humidity variations on a scale of half the radar's wavelength. Profilers typically operate with three to five fixed beams, one pointing upward and the others in orthogonal directions. Doppler frequency measurements enable the estimation of drift velocities of the scattering elements, from which the wind velocity can be obtained. Wind profilers operate best in precipitation-free conditions. These radars can measure a vertical profile of the wind speed and direction over the profiler site at time intervals as short as 5 min (standard wind measurements made worldwide from balloon-borne instruments are available only twice each day).

Shorter-wavelength radars used for storm detection also can be used to generate vertical wind profiles. The technique involves analysis of the change in the radial (along-beam) component of the wind on a circle at a fixed distance from the radar. Distant circles correspond to higher elevations. To make these measurements, echoes must cover a sufficient area around the radar. For this reason, the technique works best when precipitation is present. Together, wind profilers and shorter-wavelength radars can provide wind measurements aloft in most atmospheric conditions.

Storm structure studies. Much of the understanding of the structure of storms derives from measurements made with networks of Doppler radars. Two or more Doppler radars simultaneously observing the same volume of space allow the estimation of more than one component of the wind. With appropriate scanning procedures, Doppler radars can scan the entire volume of thunderstorms or other phenomena in a few minutes, so that the entire wind field within the storm can be determined. Techniques have also been developed to estimate pressure and temperature fields within storms from the wind fields once they are deduced. Many research projects since the mid-1970s have employed networks of Doppler radars to investigate the complete three-dimensional wind fields associated with storms, fronts, and other meteorological phenomena. Robert M. Rauber

Bibliography. D. Atlas (ed.), *Radar in Meteorology*, American Meteorological Society, 1990; L. J. Battan, *Radar Observation of the Atmosphere*, University of Chicago Press, 1973; R. J. Doviak and D. S. Zrnić, *Doppler Radar and Weather Observations*, 2d. ed., Academic Press, 1993; E. E. Gossard and R. G. Strauch, *Radar Observations of Clear Air and Clouds*, Elsevier, 1983.

Radian measure

A radian is the angle subtended at the center of a circle by an arc of the circle equal in length to its radius. It is proved in geometry that equal central angles of two circles subtend arcs proportional to their radii; and the converse is true. Hence the radian is independent of the length of the radius. The **illustration** represents two circles of radius r. Arc AB of length r subtends 1 radian (rad) at the center O of the circle, and arc $A'B'$ of length s subtends θ rad at its center. Since arcs on equal circles are proportional to their subtended central angles, $s/r = \theta/1$, or formula (1) holds. If $\theta = 2\pi$, $s = 2\pi r$,

$$S = r\theta \qquad (1)$$

the circumference of the circle. Therefore 2π rad is the complete angle about a point or 360°, and

$$\begin{aligned} 2\pi \text{ rad} &= 360^\circ \\ 1 \text{ rad} &= 360^\circ/2\pi = 57.2958^\circ = 57^\circ 17' 45'' \\ 1^\circ &= 2\pi/360 \text{ rad} = 0.0174533 \text{ rad} \end{aligned}$$

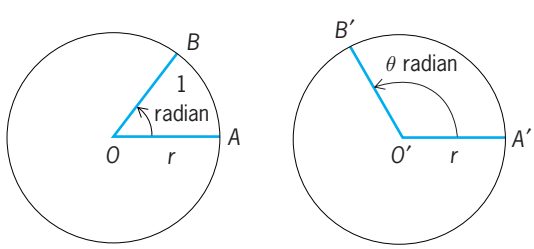

Diagrams showing radian measurement. Symbols are explained in the text.

Observe that

$$\begin{aligned} 30^\circ &= 30\pi/180 \text{ rad} = \pi/6 \text{ rad} \\ 45^\circ &= \pi/4 \text{ rad} \\ 60^\circ &= \pi/3 \text{ rad} \\ 90^\circ &= \pi/2 \text{ rad} \\ 135^\circ &= 3\pi/4 \text{ rad} \end{aligned}$$

The degree as a unit of angle has come down from antiquity. However, its use in various theories involves clumsy constants. The use of the radian avoids these constants. The radian is employed generally as a measure of angle in theoretical discussions; when no unit of angle is mentioned, the radian is understood.

The following examples illustrate the simplicity and convenience of radian measure.

The simple formula (1) would be $s = (\pi/180)r\theta$ if degrees were used. A mariner on a ship observes that a lighthouse, known to be 400 ft (120 m) high, subtends 0.064 rad at the eye, and writes from Eq. (1) that

$$400 = r(0.064) \qquad r = \frac{400}{0.064} \text{ ft} = 6250 \text{ ft}$$

The mariner then concludes that it is approximately 6250 ft (1875 m) to the lighthouse. The important ratios $(\sin\ \theta)/\theta$ and $(1 - \cos\ \theta)/\theta^2$ approach, when θ approaches zero, respective limits of 1 and $^1/_2$ when radians are used, but $\pi/180$ and $\pi^2/64{,}800$ when degrees are used. The use of radians avoids these constants throughout the many fields of application of the trigonometric functions. If a particle A moving on a circle with center O and radius r has velocity v, tangential acceleration a_t, normal acceleration a_n, and if OA has angular velocity ω and angular acceleration α, then relations (2) hold when radians are used. Each of these

$$v = r\omega \qquad a_t = r\alpha \qquad a_n = r\omega^2 \qquad (2)$$

basic formulas would involve an unwieldy constant if degrees were employed for angle measure. The simplicity arising from use of the radian indicates the importance of a wise choice of basic units in mathematics and its applications. *See* UNITS OF MEASUREMENT. Lyman M. Kells

Radiance

The physical quantity that corresponds closely to the visual brightness of a surface. A simple radiometer for measuring the (average) radiance of an incident beam of optical radiation (light, including invisible infrared and ultraviolet radiation) consists of a cylindrical tube, with a hole in each end cap to define the beam cross section there, and with a photocell against one end to measure the total radiated power in the beam of all rays that reach it through both holes (see **illus.**). If A_1 and A_2 are the respective areas of the two holes, D is the length of the tube (distance between holes), and Φ is the radiant flux or power

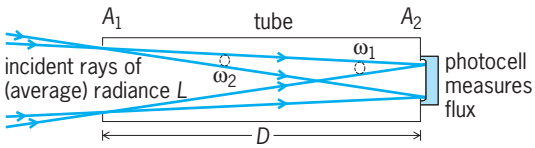

A simple radiometer.

measured by the photocell, then the (average) radiance is approximately given by the equation below.

$$L = \frac{\Phi}{A_1 \cdot A_2/D^2} \mathrm{W \cdot m^{-2} \cdot sr^{-1}}$$
$$= \frac{\Phi}{A_1 \cdot \omega_2} = \frac{\Phi}{A_2 \cdot \omega_1}$$

In the alternate forms on the second line, $\omega_2 = A_2/D^2$ is the solid angle subtended at A_1 by A_2 (at a distance D), and vice versa for ω_1. These approximate relations are good to about 1% or better when D is at least 10 times the widest part of either hole, but the accuracy degenerates rapidly for larger holes or shorter distances between them. A more exact formula, in terms of the ray radiance in each direction at every point across an area through which a beam flows, requires calculus. Then the radiance of a ray in a given direction through a given point is defined as the radiant flux per unit projected area perpendicular to the ray at the point and unit solid angle in the direction of the ray at the point. *See* LIGHT.

Power (flux) is given in SI units of watts (W), areas in square meters (m^2), and solid angles in steradians (sr), and so the units of radiance are, as shown, watts per square meter and steradian ($W \cdot m^{-2} \cdot sr^{-1}$). If the flux of visible radiation is given in units weighted for standardized eye response, called lumens (lm), instead of watts, with everything else exactly the same, the corresponding photometric quantity of luminance is obtained in $lm \cdot m^{-2} \cdot sr^{-1}$. When concerned with interactions between radiation and matter, flux can also be measured in numbers of photons or quanta per second ($q \cdot s^{-1}$) rather than watts.

A complete specification of a beam of optical radiation requires the distribution of radiance, not only as a function of ray position and direction, but also as a function of wavelength, time, and polarization. A distribution in space and wavelength λ is spectral radiance L_λ in watts per square meter, steradian, and nanometer of wavelength ($W \cdot m^{-2} \cdot sr^{-1} \cdot nm^{-1}$). Spectral radiance may also be given in terms of wave number σ in reciprocal centimeters (cm^{-1}) or of frequency ν in terahertz (THz). No special names are given to the distributions with respect to time (or frequency of fluctuation or modulation) and polarization, and they are often overlooked even though their effects may be significant.

All of the foregoing assumes that radiant energy is propagated along the rays of geometrical optics that can form sharp shadows and images. The results may be in error when there are significant diffraction or interference effects, as is often the case with coherent radiation such as that from a laser. *See* DIFFRACTION; GEOMETRICAL OPTICS; INTERFERENCE OF WAVES; LASER.

Fred E. Nicodemus

Bibliography. R. W. Boyd, *Radiometry and the Detection of Optical Radiation*, 1983; R. M. McCluney, *Introduction to Radiometry and Photometry*, 1994; F. E. Nicodemus, Radiance, *Amer. J. Phys.*, 31:368–377, 1963.

Radiant heating

Any system of space heating in which the heat-producing means is a surface that emits heat to the surroundings by radiation rather than by conduction or convection. The surfaces may be radiators such as baseboard radiators or convectors, or they may be the panel surfaces of the space to be heated. *See* PANEL HEATING AND COOLING; RADIATOR.

The heat derived from the Sun is radiant energy. Radiant rays pass through gases without warming them appreciably, but they increase the sensible temperature of liquid or solid objects upon which they impinge. The same principle applies to all forms of radiant-heating systems, except that convection currents are established in enclosed spaces and a portion of the space heating is produced by convection. The radiation component of convectors can be increased by providing a reflective surface on the wall side of the convector and painting the inside of the enclosure a dead black to absorb heat and transmit it through the enclosure, thus increasing the temperature of that side of the convector exposed to the space to be heated.

Any radiant-heating system using a fluid heat conveyor may be employed as a cooling system by substituting cold water or other cold fluid. This cannot be done with electric resistance-type radiant-heating systems. Thermoelectric couples will emit or absorb heat, depending upon the polarity of the direct current applied to them. However, the technique is practical only for very special and small-scale heating and cooling applications, certainly not on the scale required for comfort control of an occupied space. *See* COMFORT HEATING; HEAT RADIATION.

Erwin L. Weber; Richard Koral

Bibliography. M. F. Modest, *Radiative Heat Transfer*, 2d ed., 2003; R. Siegel and J. R. Howell, *Thermal Radiation Heat Transfer*, 4th ed., 2001; R. Watson and K. Chapman, *Radiant Heating and Cooling Handbook*, 2002.

Radiation

The emission and propagation of energy; also, the emitted energy itself. The etymology of the word implies that the energy propagates rectilinearly, and in a limited sense, this holds for the many different types of radiation encountered.

The major types of radiation may be described as electromagnetic, acoustic, and particle, and within these major divisions there are many subdivisions.

For example, electromagnetic radiation, which in the most familiar energy ranges behaves in a manner usually characteristic of waves rather than of

particles, is classified roughly in order of decreasing wavelength as radio, microwave, visible, ultraviolet, x-rays, and gamma rays. In the last three subdivisions, and frequently in the visible, the behavior of the radiation is more particlelike than wavelike.

Since the energy of a photon (light quantum) is inversely proportional to the wavelength, this classification is also on the basis of increasing photon energy. *See* ELECTROMAGNETIC RADIATION.

Acoustic or sound radiation may be classified by frequency as infrasonic, sonic, or ultrasonic in order of increasing frequency, with sonic being between about 16 and 20,000 Hz. Infrasonic sound can result, for example, from explosions or other sources so loud that exceptional waves are set up because the large amplitudes of the source vibrations exceed the elastic limit of the transmitting medium. Ultrasonic sound can be produced by means of crystals which vibrate rapidly in response to alternating electric voltages applied to them. There is a nearly infinite variety of sources in the sonic range. *See* SOUND.

The traditional examples of particle radiation are the alpha particles and beta particles of radioactivity. Cosmic rays also consist largely of particles—protons, neutrons, and heavier nuclei, along with beta particles, mesons, and the so-called strange particles. *See* COSMIC RAYS; ELEMENTARY PARTICLE.

McAllister H. Hull, Jr.

Radiation biology

The study of the action of ionizing and nonionizing radiation on biological systems. Ionizing radiation includes highly energetic electromagnetic radiation (x-rays, gamma rays, or cosmic rays) and particulate radiation (alpha particles, beta particles, neutrons, or heavy charged ions). Nonionizing radiation includes ultraviolet radiation, microwaves, and extralow-frequency (ELF) electromagnetic radiation. These two types of radiation have different modes of action on biological material: ionizing radiation is sufficiently energetic to cause ionizations, whereas nonionizing radiation causes molecular excitations. In both cases, the result is that chemical bonds of molecules may be altered, causing mutations, cell death, or other biological changes. *See* ELECTROMAGNETIC RADIATION; RADIATION.

Ionizing radiation originates from external sources (medical x-ray equipment, cathode-ray tubes in television sets or computer video displays) or from internal sources (ingested or inhaled radioisotopes, such as radon-222, strontium-90, and iodine-131), and is either anthropogenic (medical, industrial, or military) or natural (atmospheric or terrestrial).

Nonionizing radiation originates from natural sources (sunlight, Earth's magnetic field, lightning, static electricity, endogenous body currents) and technological sources (computer video displays and television sets, microwave ovens, communications equipment, electric equipment and appliances, and high-voltage transmission lines).

Ionizing Radiation

The action of ionizing radiation is best described by the three stages (physical, chemical, and biological) that occur as a result of energy release in the biological target material (**Fig. 1**).

Physical stage. All ionizing radiation causes ionizations of atoms in the biological target material. The Compton effect, which predominates at the energies of electromagnetic radiation that are commonly encountered (for example, x-rays or gamma rays), strips orbital electrons from the atoms. These electrons (Compton electrons) travel through the target material, colliding with atoms and thereby releasing packets of energy. About 60 eV of energy is released with each collision, breaking chemical bonds and causing ionizations known as primary ionizations. The typical initial energy of a Compton electron is 200 keV, and so there may be over 3300 primary ionizations from a single Compton electron. For low-energy x-rays, the photoelectric effect predominates, producing photoelectrons that transfer their energy in the same manner as Compton electrons. *See* COMPTON EFFECT; ELECTRON; GAMMA RAYS; X-RAYS.

The absorbed dose of ionizing radiation is measured as the gray (Gy, 1 joule of energy absorbed by 1 kilogram of material). In general, the amount of absorption of ionizing radiation is greater for target material of high density and high atomic number, and it is independent of the chemical bonds of the target material. Because of the very localized absorption of ionizing radiation as compared to heat energy, an amount of ionizing radiation energy equivalent to 1/100 the heat energy in a cup of coffee will result in a 50% chance that the person absorbing the radiation will die in 30 days.

Neutrons with energies between 10 keV and 10 MeV transfer energy mainly by elastic scattering, that is, billiard-ball-type collisions, of atomic nuclei in the target material. In this process the nucleus is torn free of some or all of the orbital electrons because its velocity is greater than that of the orbital electrons. The recoiling atomic nucleus behaves as a positively charged particle. Because the mass of the neutron is nearly the same as that of the hydrogen atom, hydrogenous materials are most effective for energy transfer. A unit known as the kerma (acronym for kinetic energy released in material) is used to measure the amount of energy transfer from neutrons and other indirectly ionizing radiation (for example, x-rays and gamma rays). This quantity is frequently equivalent to the absorbed dose of radiation in the material (the gray) because the ranges of the secondary recoiling charged particles are much shorter than the ranges of neutrons, x-rays, or gamma rays. *See* NEUTRON.

Chemical stage. Chemical changes in biological molecules are caused by the direct transfer of radiation energy (direct radiation action) or by the production of chemically reactive products from radiolysis of water that diffuse to the biological molecule (indirect radiation action). More than half the biological action of low linear-energy-transfer (LET) ionizing

radiation (for example, x-rays and gamma rays) results from indirect radiation action, about 90% of which is due to the action of the hydroxyl radical (OH·). Characteristically, the manifestations of indirect radiation action decrease as the linear energy transfer of the radiation increases. Thus, for high linear-energy-transfer radiation, direct radiation action predominates. Chemicals that react with hydroxyl radicals, rendering them unreactive, provide protection against indirect radiation damage. For example, sulfhydryl compounds (RSH) remove hydroxyl radicals by reaction (1), where R represents

$$2OH\cdot + 2RSH \longrightarrow 2H_2O + RSSR \qquad (1)$$

an organic functional group. *See* LINEAR ENERGY TRANSFER (BIOLOGY).

Indirect radiation action is also responsible for the oxygen effect of radiobiology, which describes the increased biological sensitivity when irradiation occurs in the presence of oxygen compared to the absence of oxygen during irradiation. The oxygen must be present during irradiation because of the limited lifetime (less than 1 millisecond) of the chemical species that react with oxygen. The oxygen effect is observed only in biological systems with membranes; the radiosensitivity of viruses and free deoxyribonucleic acid (DNA) is not enhanced by the presence of oxygen. Thus, two types of radiation damage are probable: oxygen-dependent radiation damage, which involves the membrane as the principal target; and oxygen-independent damage, which involves nonmembrane components, including the genome. *See* RADIATION CHEMISTRY.

The most important biological targets for damage from ionizing radiation are probably the plasma membrane and DNA, because there is only one copy, or a few copies, in the cell; because they serve critical roles for the survival and propagation of cells; and because they are large. The last factor is important because ionizing radiation releases its energy in a random manner; thus the larger the target, the more likely that it will be damaged by radiation.

Membranes are composed of protein (50%) and lipids (50%). Biochemical damage in either of these components can cause membrane damage. Consequences of radiation damage in membranes are changes in ion permeability, with leakage of potassium ions; changes in active transport; and cell lysis. *See* CELL MEMBRANES.

The initial direct radiation action results in hydrogen abstraction [reaction (2)]. In the presence of oxygen (O_2), reactions (3)–(6) occur. Oxidative

$$RH \longrightarrow R\cdot + H\cdot \qquad (2)$$

$$H\cdot + O_2 \longrightarrow HO_2\cdot \qquad (3)$$

$$R\cdot + HO_2\cdot \longrightarrow ROOH \qquad (4)$$

$$R\cdot + O_2 \longrightarrow RO_2\cdot \qquad (5)$$

$$RO_2\cdot + H\cdot \longrightarrow ROOH \qquad (6)$$

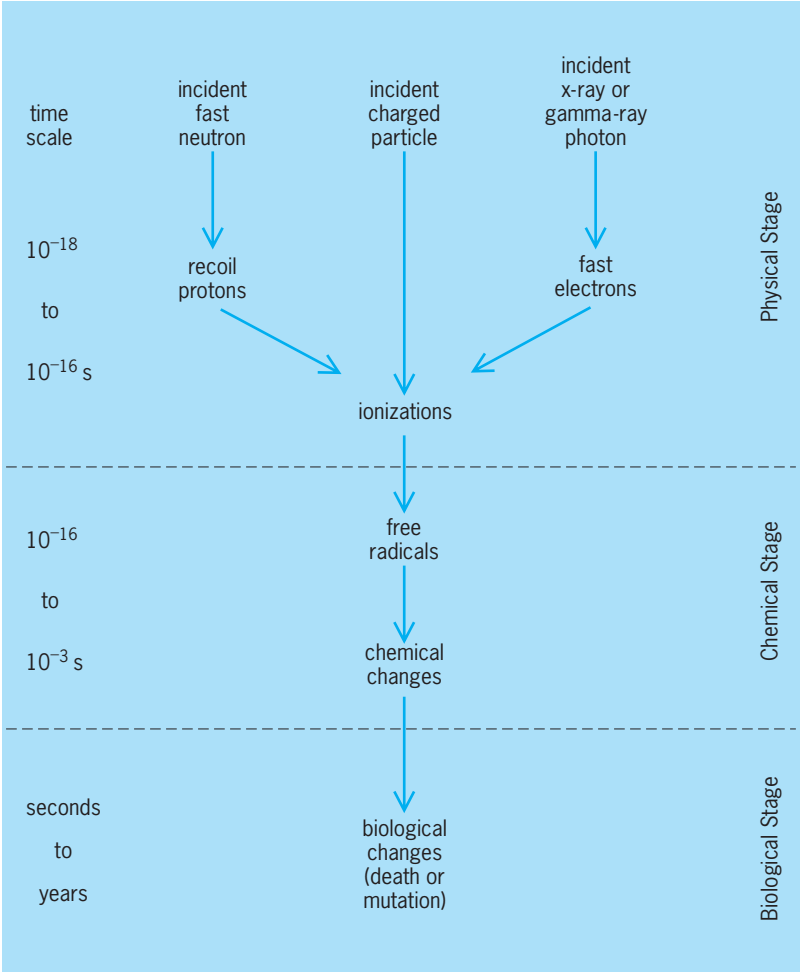

Fig. 1. Diagram of the three stages for the action of ionizing radiation on biological systems. The duration of the physical stage approximates the time required for the ionizing radiation to traverse an atom. During the chemical stage, molecular and thermal equilibrium is achieved. The biological stage involves metabolism of the chemical damage.

decomposition products of the organoperoxides (ROOH) accumulate, and overall lipid damage is increased approximately threefold when irradiation occurs in the presence of oxygen. Two sensitive sites for protein damage are the free amino and sulfhydryl groups, as observed by the release of ammonia and hydrogen sulfide gases during irradiation of proteins. Another sensitive site is the imidazole ring of the amino acid histidine. The peptide bond is also subject to attack, which causes breakage of the polypeptide chain and a reduction of the molecular weight of the protein. *See* LIPID; PROTEIN.

Membrane components can be protected by reactions of the initially formed organic radical with sulfhydryl compounds [reaction (7)]. The sulfhydryl

$$R\cdot + RSH \longrightarrow RH + RS\cdot \qquad (7)$$

compound restores the initially damaged membrane component to its original state. This radioprotective reaction does not appear to be important for protection of DNA.

Lesions in DNA that is irradiated in aqueous solution include single-strand breaks, double-strand

breaks, base damage, interstrand cross-links, and DNA-protein cross-links. Base damage occurs most in thymine and least in guanine, and the yield of single-strand breaks is about 10 times the yield of double-strand breaks and cross-links. The ratio of the radiosensitivities of thymine in DNA is 1:2:6 when the DNA is in the form of free DNA, active chromatin, and condensed chromatin, respectively. This demonstrates protection of DNA by the proteins of the chromatin against lesions from diffusible water (hydroxyl) radicals. Oxidations of sugars in the DNA backbone and loss of whole nucleotides are the main mechanisms for formation of breaks in strands. *See* DEOXYRIBONUCLEIC ACID (DNA).

Biological stage. Various biological effects can result from the biological actions of ionizing radiation (**Fig. 2**). Reproductive death is most pronounced in mammalian cells that are actively dividing and in nondifferentiated tissue. Thus, dividing tissues (bone marrow and the germinal cells of the ovary and testis) are radiosensitive, and nondividing tissues (liver, kidney, brain, muscle, cartilage, and connective tissue) are radioresistant in animals exposed to whole-body radiation. Developing embryos are quite radiosensitive, as predicted by this principle. The radiosensitivity of organisms varies greatly, being related to their intrinsic sensitivity to radiobiological damage and to their ability to repair the damage. Radiation doses resulting in 10% survival range from 3 Gy (mouse and human cells), to 60 Gy (most bacteria and the fruit fly), to 130 Gy (cabbage looper), to 600 Gy (viruses), to greater than 1000 Gy (the bacterium *Deinococcus radiodurans*).

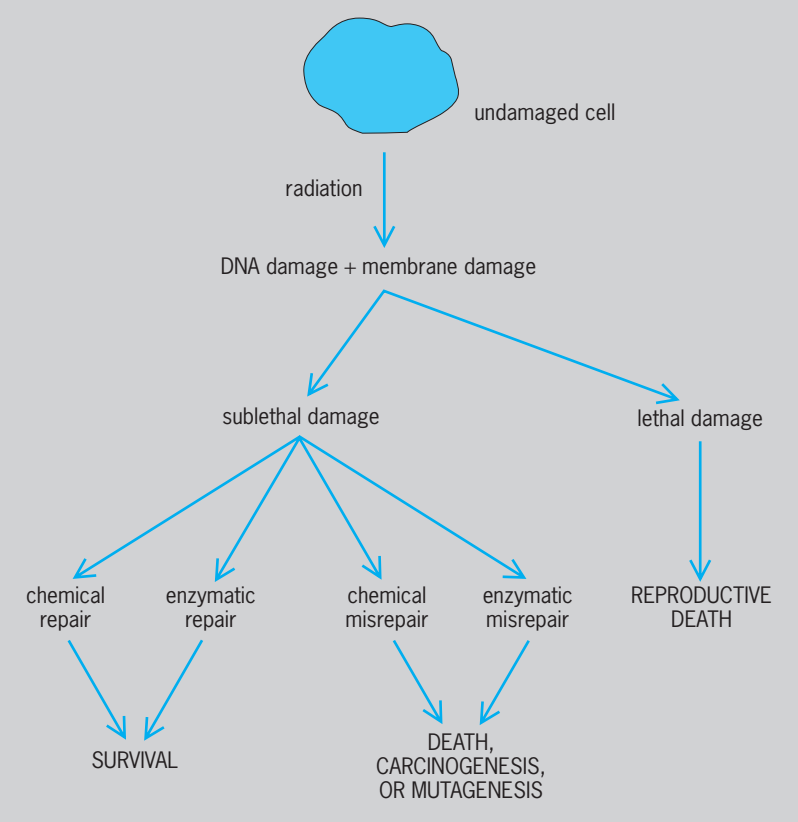

Fig. 2. Possible pathways for cellular response to ionizing radiation.

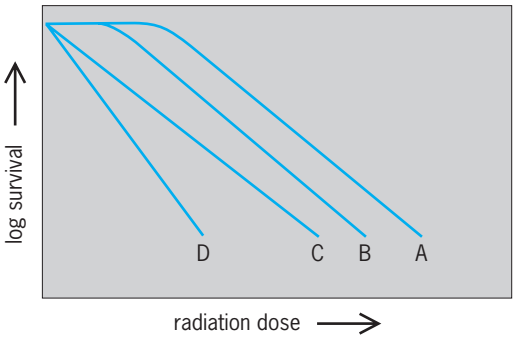

Fig. 3. Biological survival curves. A and B are biphasic with an initial shoulder period followed by exponential inactivation kinetics. C and D are exponential only.

Single-break chromosome damage leads to chromosome aberrations in cells irradiated early in the cell cycle (before DNA synthesis has started) and to aberrations in chromatid aberrations in cells irradiated later in the cell cycle (after DNA synthesis). The yield of chromosome breaks increases twofold when oxygen is present during irradiation. *See* CELL CYCLE; CHROMOSOME ABERRATION.

Cells lose their reproductive capacity when exposed to ionizing radiation. Mammalian cells are most resistant to this action during the late S phase of the cell cycle, and most sensitive during mitosis. *See* MITOSIS.

Repair of biological damage involves repair of damage in the DNA. The kinetics of biological inactivation can be biphasic, with an initial "shoulder" period followed by exponential kinetics (**Fig. 3**, A and B), or exponential only (Fig. 3, C and D). Two types of biological repair are involved in the differences in the magnitude of the shoulder region and the slope of the exponential region: the shoulder region arises from the action of Q repair, which diminishes as dose increases, and the value of the exponential slope is affected by the action of P repair, which is not diminished as dose increases. Curves A and B in Fig. 3 exemplify conditions in which the biological system of curve A has a more active Q-repair system than that of curve B while both have the same amount of P repair. Alternatively, the final slope of a curve can be influenced by the intrinsic radiosensitivity of the biological system. Also, the initial shoulder period can result from the necessity that several critical targets must be inactivated before the biological system is inactivated, or several radiation-damaging events (hits) must occur within the same target in order for biological inactivation to occur (Fig. 3, A and B).

The three organ systems that generally contribute to the death of mammals following a single dose of whole-body irradiation are, in decreasing order of radiosensitivity, the hematopoietic system, the gastrointestinal system, and the cerebrovascular system. At very high doses (above 100 Gy), the survival time of the animal is short (from minutes to several days after exposure, depending on the dose), and death results from damage to the cerebrovascular system.

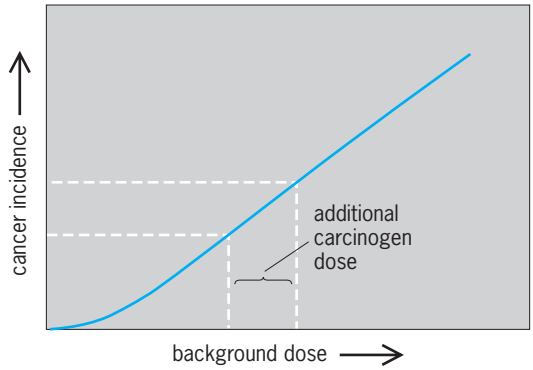

Fig. 4. Additive interaction of carcinogenic agents. (*After A. C. Upton, The question of thresholds for radiation and chemical carcinogenesis, Canc. Invest., 7(3):267–276, 1989*)

At intermediate doses (between 10 and 100 Gy), survival is longer (between 3 and 4 days), and death occurs because of gastrointestinal damage. The longest duration between exposure and death (between 1 week and 1 month) follows radiation doses of 3 to 10 Gy, and arises from failure of the hematopoietic system.

Late somatic effects may take years or decades to appear and include genetic mutations transmitted to subsequent generations, tumor development and carcinogenesis, and shortening of life span. *See* MUTAGENS AND CARCINOGENS; MUTATION; TUMOR.

An important and controversial issue concerning ionizing radiation is the question of the existence of a threshold dose, below which no biological effects occur. The most conservative model uses a linear equation to relate the incidence of biological effects with increasing dose. Use of a quadratic equation to describe the dose response predicts the presence of a threshold dose. Because radiation damage is cumulative, and this damage can interact with other chemical agents in causing biological effects such as carcinogenesis, the linear model is favored (**Fig. 4**). *See* RADIATION INJURY (BIOLOGY).

Nonionizing Radiation

Of all the nonionizing radiations, only ultraviolet radiation, microwaves, and high-voltage electromagnetic radiation are considered in the study of radiation biology.

Ultraviolet radiation. This type of radiation is much less penetrating than ionizing radiation. Since it can penetrate only several layers of cells, the effects of ultraviolet (UV) radiation on humans are restricted to the skin and the eyes. Ultraviolet radiation is divided into UV-C (wavelength of 200–280 nanometers), UV-B (280–320 nm), and UV-A (320–400 nm). The most biologically damaging is UV-C, and the least damaging is UV-A, with UV-B having intermediate efficiency of biological action. The solar spectrum at the Earth's surface contains only the UV-B and UV-A radiations. Stratospheric ozone strongly absorbs UV-C radiation and the shorter wavelength portion of UV-B radiation, thus providing some biological protection. Depletion of the ozone layer caused by stratospheric pollution dramatically reduces this protective action, resulting in a decreased atmospheric absorption of UV-B radiation. *See* AIR POLLUTION; STRATOSPHERIC OZONE.

Biological effects can arise only when absorption of ultraviolet radiation occurs. Absorption is dependent on the chemical bonds of the material, and it is highly specific. Identification of the critical target for biological action has been facilitated by this specific absorption, using a comparison to the action spectrum, that is the relative efficiency for causing biological effects at different wavelengths. The action spectrum for ultraviolet radiation closely matches the absorption spectrum of nucleic acid, at wavelengths between 240 and 300 nm. At wavelengths greater than 320 nm, ultraviolet-mediated damage is oxygen-mediated, because of the interaction of dioxygen (O_2) with photoexcited biological chromophores, whereas below 300 nm, biological damage occurs principally by photochemical damage in DNA. Exposure dose for ultraviolet radiation is measured in units of joules per square meter; it is a measure of the incident energy per unit area impinging on the target material.

Biological responses. Sunburn is a form of erythema produced by overexposure to the UV-B portion of the solar spectrum (which is not transmitted through window glass). A rare but deadly form of skin cancer in humans, malignant melanoma, is induced by exposure to sunlight, with occurrences localized on those regions of the body that are most frequently exposed.

Ultraviolet light can also cause photochemical damage. Cyclobutane pyrimidine dimers are the main photoproduct following exposure to UV-C and UV-B, and they can lead to cell death and precarcinogenic lesions. Other types of dimers are considered to be especially mutagenic. DNA-protein cross-links that are observed after ultraviolet radiation can be lethal.

Repair systems. Survival from ultraviolet irradiation is reduced as the dose of radiation is increased. The shapes of survival curves are similar to those for lethality from ionizing radiation (Fig. 3); they are dependent on the presence or absence of repair systems. The four repair systems that enhance biological survival are discussed below.

Photoreactivation is an enzymatic repair system that enhances biological survival by splitting cyclobutane dimers in the DNA of cells that have been irradiated by ultraviolet light. The process requires light, and the most effective wavelengths are in the blue region of the visible spectrum. The repair system is error-free, and thus it is nonmutagenic. However, the repair must correct the damage before it is copied by DNA replication.

Another system is excision repair. A region of DNA containing bases that have been damaged by ultraviolet radiation is removed enzymatically, followed by synthesis of new DNA to replace the damaged region. There are no errors, and mutant cells that lack this repair system are highly sensitive to ultraviolet radiation. In fact, cells isolated from individuals with the genetic-recessive disease xeroderma

pigmentosum lack excision repair. The clinical symptoms of this disease include a high incidence of skin lesions and early death from malignant melanoma. Excision repair operates in the dark. The damage must be corrected before DNA replication copies it. Not all regions of DNA are equally accessible to the action of this repair system, and it has been observed that actively transcribed regions of DNA in mammalian cells are more efficiently repaired than are nontranscribed regions.

Recombination repair is another repair system. DNA that contains damage is replicated, and gaps in the newly synthesized DNA appear opposite to the sites where there is damage in the parental strand. These gaps are filled in by recombining the portion of the undamaged complementary parental strand with the daughter strand containing the gaps. This repair system is essentially error-free; it is sometimes known as postreplicational repair, because it occurs after replication of the damaged region. *See* RECOMBINATION (GENETICS).

The system known as SOS repair is an inducible repair system. It is activated by a reactive intermediate of DNA metabolism in irradiated bacteria. SOS repair is associated with a complex of responsive genes known as *din* (damage-inducible) genes. SOS repair acts by inducing synthesis of an alternative excision repair system, and the recombination repair system, which increases survival. The alternative excision repair system involves removal of a longer segment of DNA containing the damaged region; it is error-prone, which leads to increased mutations. Another manifestation of the SOS response is the induction of DNA replication that can bypass the damaged template region of DNA. Thus this is not actually a repair system, since the original damage remains in the DNA, and can best be thought of as a system for tolerating DNA damage. Survival is increased, because DNA replication is not blocked by the presence of the damage. *See* ULTRAVIOLET RADIATION; ULTRAVIOLET RADIATION (BIOLOGY).

Microwaves. Microwaves are electromagnetic radiation in the region from 30 MHz to 300 GHz. They originate from devices such as telecommunications equipment and microwave ovens. Metals reflect microwaves; glass transmits them; and aqueous material absorbs them, accompanied by a rise in temperature of the liquid. The 915- and 2450-MHz bands are used for industrial microwave heating and in microwave ovens. Thermal effects of microwaves occur at exposure rates greater than 10 mW/cm^2 (70 mW/in.^{2}), while nonthermal effects are associated with exposure rates less than 10 mW/cm^2.

Microwave radiation is absorbed unevenly in biological tissue because of the heterogeneity of the dielectric properties of the material. Thus, significant temperature gradients can be established, which may enhance the action of thermal heating from microwaves as compared to that from infrared radiation. Certainly, material with a high water content will have a higher absorption coefficient for microwaves, and thus a greater thermal response to microwave action. Microwave absorption is high in skin, muscle, and internal organs, and lower in bone and fat tissue.

Biological effects. Cultured mammalian cells exposed to microwaves at a high power density show chromosome abnormalities after 15 min of exposure. Progression through the cell cycle is also temporarily interrupted, which interrupts DNA synthesis. Chromosome aberrations in peripheral blood lymphocytes are significantly greater for persons who are occupationally exposed to microwaves.

Microwaves can be lethal when the power intensity and exposure time are sufficient to cause a rise in temperature that exceeds an organism's homeostatic capabilities; this occurs when the temperature rise exceeds approximately 5°C (9°F). For example, cataracts can result from exposure to ultraviolet radiation because the eye is not able to dissipate heat very well since there is not blood circulation. The testicles are at a temperature about 2°C (4°F) below that of body temperature, and spermatogenesis is particularly sensitive to temperature rise. Reversible testicular damage, in the form of reduced spermatogenesis and degeneration of the epithelial lining of the seminiferous tubules, therefore can occur from exposure to microwaves.

There are also some nonthermal effects associated with microwaves. There is significant uncertainty in the actual doses among people exposed as a result of their occupations; this reduces the reliability of the observations of the nonthermal actions of microwaves. However, a list of clinical symptoms includes increased fatigue, periodic or constant headaches, extreme irritability, decreased hearing acuity, and drowsiness during work. Laboratory studies involving exposure of animals to microwaves have produced changes in the electroencephalogram, blood–brain barrier, central nervous system, hematology, and behavior. Cell membrane permeability is also altered.

Exposure limits. It would be desirable to establish limits of exposure to microwaves based on the absorbed dose of microwave energy, similar to the gray unit that is used for ionizing radiation. However, the difficulty associated with measuring the absorbed dose makes this impractical. Therefore, radiation protection standards are based on exposure values instead of the previously used specific absorption rate (SAR, measured as W/kg); exposure dose is measured as W/cm^2. Occupational limits are set at a maximum of 1 mW/cm^2 (7 mW/in.^{2}) for frequencies between 30 and 100 MHz, and 10 mW/cm^2 (70 mW/in.^{2}) for frequencies between 1 and 300 GHz, averaged over a 6-min period. These values are selected to limit the specific absorption rate of the average whole body to 0.4 W/kg (0.2 W/lb), and they take into consideration the variable absorption for microwaves of different frequencies. The maximum acceptable limit for leakage from household microwave ovens is set at 1 mW/cm^2 (7 mW/in.^{2}) or less, measured at a distance of 5 cm (2 in.) from a new oven, and never more than 5 mW/cm^2 measured 5 cm (2 in.) from the oven, during its expected lifetime. *See* INDUSTRIAL HEALTH AND SAFETY; MICROWAVE.

Extremely low frequency electromagnetic fields. This type of radiation is generated by the electric and magnetic fields associated with high-voltage current in power transmission lines, and also some household and industrial electrical equipment. Biological effects from ELF radiation are the least understood, and the potential consequences are the most controversial. The issue of potential biological damage from this type of radiation has arisen only since the introduction of very high voltage electric power transmission lines (440 kV and above) and the occurrence of widespread use of various electrical and electronic equipment. *See* ELECTROMAGNETIC PULSE (EMP).

Biological effects. Biological studies on ELF electromagnetic fields have been performed on cells and whole animals; and epidemiological studies have been carried out on populations exposed occupationally. The results share some common features that had not been expected: (1) There is not always a clear dose response; that is, increasing the exposure does not necessarily give rise to an increased biological effect, as is observed commonly for other types of radiation. (2) Some biological effects are seen only at certain frequencies and dose rates. Some of the reported effects are subjective, and may be related to normal physiological adaptation to environmental changes.

In cellular studies, membrane transport of calcium ions by chick brain cells was altered by exposure to ELF electric fields. Chromosome aberrations have not been observed in lymphocytes after long-term exposure to 60-Hz fields of 50 kV/m (15 kV/ft) and 10 gauss (10^{-3} tesla), suggesting that ELF fields whose strengths are less than these are unlikely to cause cancer or mutations; but this does not provide conclusive evidence that such radiation is noncarcinogenic. The major action of ELF electromagnetic radiation is targeted to the cell membrane, and most likely involves changes in membrane activity.

Controlled laboratory studies on developing chick embryos continuously exposed to 60-Hz electric fields up to strengths of 100 kV/m (30 kV/ft) show no effects on the general health, development, mortality, bone growth, malformations, and behavior for embryos that are hatched and followed for 6 weeks of growth. However, rats exposed to 60-Hz electric fields between 2 and 40 kV/m (0.6 and 12 kV/ft) for 21 days had reached nighttime peaks of pineal gland melatonin secretion.

In humans, qualitative biological effects of low-frequency radiation (0 to 300 Hz) include headaches, lethargy, and decreased sex drive. Humans have been noted to perceive the presence of a 60-Hz electric field when the intensity is in the range of 2 to 12 kV/m (0.6 to 3.6 kV/ft), and animals were observed to avoid entering an area where the electric field was greater than 4 kV/m (1.2 kV/ft).

Some epidemiological studies have shown a pattern of increased cancer (leukemias and brain tumors) among individuals occupationally exposed to increased ELF electromagnetic fields. However, the evidence is weak, because no dose dependency has been found; also the evidence is only associative and not of a cause-and-effect relational type. Some studies have reported no link between exposure to ELF fields and cancer.

Exposure limits. Because of the conflicting experimental results on the biological effects of ELF electromagnetic radiation, the guidelines are of an interim nature. Exposure to electric and magnetic fields gives rise to induced current in the human body. Clearly, exposure should be well below values that would lead to acute effects such as heart fibrillation and pain. Measurements show that a current density of 4 mA/m^2 (0.4 mA/ft^2) averaged over the head and trunk regions occurs when an individual is exposed to an electric field of 10 kV/cm (25 kV/in.), or a magnetic field of 500 G (50 milliteslas). The natural current densities in the body are about 10 mA/m^2 (0.9 mA/ft^2). The rationale for setting exposure limits to radiation from ELF electric and magnetic fields is to limit external exposure to no more than the endogenous values.

Studies on humans exposed to a 60-Hz magnetic field of 5 mT combined with a 60-Hz electric field of 20 kV/m (6 kV/ft) for 4 to 6 hours per day over several days revealed no effects on the health or welfare of the volunteer subjects.

The International Radiation Protection Agency (IRPA), in collaboration with the Environmental Health Division of the United Nations World Health Organization (WHO), has established exposure limits to ELF electric and magnetic fields, based upon available knowledge about the endogenous fields, and the biological effects of ELF electric and magnetic fields. Continuous exposure of the general public should not exceed 2 kV/m (0.6 kV/ft) for 50/60-Hz electric fields, and should not exceed 0.1 mT for 50/60-Hz magnetic fields, which includes a safety factor of 5 with regard to the maximum continuous field exposures allowed for industrial safety [10 kV/m (3 kV/ft) and 0.5 mT]. *See* RADIATION INJURY (BIOLOGY).

Phillip M. Achey

Bibliography. Z. M. Bacq and P. Alexander, *Fundamentals of Radiobiology*, 2d ed., 1961; W. Harm, *Biological Effects of Ultraviolet Radiation*, 1980; IRPA/INIRC Committee Report, Interim guidelines on limits of exposure to 50/60 Hz electric and magnetic fields, *Health Phys.*, 58:113–122, 1990; D. E. Lea, *Actions of Radiations on Living Cells*, 2d ed., 1955; A. H. Nias and R. Dimbleby, *An Introduction to Radiobiology*, 2d ed., 1998; U. L. Prenn, *Introduction to Biological Radiation Effects*, 2d ed., 1994; Special issue on radiation effects on man and animals, *Experientia*, 45:1–114, 1989.

Radiation chemistry

The study of chemical changes resulting from the absorption of high-energy, ionizing radiation. Such radiation includes alpha particles, electrons, gamma rays, fission fragments, protons, deuterons, helium nuclei, and heavier charged projectiles. X-rays are distinguished from gamma rays only as being extranuclear in origin. In absorbing materials of low and

intermediate atomic weight such as aqueous systems and most biological systems, these radiations deposit energy in a largely indiscriminate manner, leaving behind a complex mixture of short-lived ions, free radicals, and electronically excited molecules. This contrasts with the absorption of visible and ultraviolet radiation, in which one or a few specific electronically excited species are formed. Radiation chemical change results from the further reaction of these intermediates. *See* PHOTOCHEMISTRY.

Because of the nature of the absorption process, chemical-product yields are expressed in terms of total energy absorption. The extent of chemical change per unit of absorbed energy is known as the G value for the reaction. The SI unit is moles or micromoles of product formed (or reactant decomposed) per joule. The more frequently used unit is molecules changed per 100 electronvolts absorbed: $G(\mu\text{mol/joule}) = 0.103\ G(\text{molecules/100 eV})$.

Despite the initial complexity, a wealth of kinetic information has been obtained from radiation chemical studies.

Sources of high-energy radiation. Sources of high-energy radiations in the laboratory and industry include radioactive nuclides [for example, cobalt-60 (^{60}Co), strontium-90 (^{90}Sr), and hydrogen-3 (^{3}H)] and instruments such as x-ray tubes, Van de Graaff generators, the betatron, the cyclotron, and the synchrotron. An electron accelerator known as the linac (linear electron accelerator) has proved particularly valuable for the study of transient species that have lifetimes as short as 16 picoseconds; and another electron accelerator, known as the Febetron, has been used for the study of the effects of single pulses of electrons with widths of several nanoseconds at very high currents.

Interaction of radiation with matter. The primary absorption processes for high-energy radiation are ionization and molecular excitation. The distribution of the absorbed energy, however, depends significantly upon the nature of the radiation and absorbing medium.

Electrons. If the electron source is a beta-particle-emitting isotope dispersed in the absorbing medium, electron energies (E) will range from essentially zero to an E_{max} characteristic of the decay process. For ^{3}H, E_{max} is 18 keV; for phosphorus-32 (^{32}P), it is 1.71 MeV. Accelerator-produced electron beams are monoenergetic, and for radiation chemical applications may be from several hundred kiloelectronvolts to 10 MeV. *See* BETA PARTICLES.

Fast electrons undergo energy loss through inelastic scattering by atoms and molecules in which ionization and electronic excitation occur. Except for straggling, which involves a statistical fluctuation in the number of encountered molecules, a charged particle has a definite range in a given medium. This range depends upon the particle charge, mass, and energy and upon the density of the medium. The geometrical distribution of energy deposition along the particle path is determined by its linear rate of energy transfer (LET). The LET will vary along the path of the particle, increasing as the particle velocity decreases. An average LET may be obtained as the total particle energy divided by its range.

An electron of 300 keV will travel, on the average, 80 mm in water and produce approximately 9000 ionization events along its path, some by the secondary electrons released. These events are distributed along the path in small volumes known as spurs. Each spur may contain several ion pairs, electronically excited molecules, and free radicals formed from reaction of these species. Spurs may overlap in tracks of dense ionization, or they may be separated by many molecular diameters in tracks of low LET. Chemical change results from reactions within the spurs and through reaction of intermediates diffusing from the spurs into the bulk medium.

An electron slowed to thermal energy in a liquid medium may be trapped by a small group of solvent molecules that have been oriented by its presence. The hydrated electron (e^-_{aq}) is a strong reducing agent, has a mean lifetime of milliseconds in neutral solutions, and has an optical absorption peak at 715 nanometers. It is perhaps the most well characterized intermediate in radiation chemistry. *See* ELECTRON.

Gamma rays. Dependent upon the energy and the absorbing medium, gamma rays transfer energy by three principal interactions.

In photoelectric absorption, the gamma ray imparts all of its energy to an electron that is ejected from the absorbing molecule with an energy equal to that of the gamma ray less the binding energy of the electron. The photoelectron is absorbed, with the production of further ionization and excitation. The photoelectric effect is important for low-energy gamma rays and high-atomic-number absorbers. It is the dominant mode in water for gamma rays or x-rays of less than 20 keV.

In Compton scattering, the gamma ray imparts a portion of its energy to an electron, ejecting it from the molecule. The scattered gamma ray continues with diminished energy, and it may experience further such interactions in the medium. The Compton electron or electrons are responsible for the major energy depositions in the system. Compton scattering is most important for absorbers of low and intermediate atomic numbers. For gamma rays of from 0.1 to several megaelectronvolts, it dominates the absorption process in aqueous and biological media. *See* COMPTON EFFECT.

Pair production may occur with gamma rays of energy greater than 1.02 MeV. The rest-mass energy of the electron is 0.51 MeV; and in an interaction with an atomic nucleus, a gamma ray of twice this energy may be absorbed, with the formation of an electron-positron pair. The pair share the gamma-ray energy in excess of 1.02 MeV and undergo the usual electron absorption behavior. The positron, however, upon reaching thermal energy, will undergo annihilation with an interacting electron, and two 0.51-MeV gamma rays are formed. The probability of pair production increases with gamma-ray energy and with increasing atomic number of the absorber.

It occurs significantly in water and biological material only above 10 MeV. *See* GAMMA RAYS; POSITRON.

Heavy charged particles. These particles interact with matter in qualitatively the same way as electrons, producing ionization and excitation. For a given energy, however, ionization and spur density are much greater for these particles and increase with decreasing particle velocity. Average LET values in water, for 1-MeV particles, are 43, 58, and 190 keV/μm for protons, deuterons, and alpha particles, respectively. *See* ALPHA PARTICLES.

Dosimetry. Evaluation of the yields of radiation-induced reactions requires a knowledge of the energy imparted to the reacting system. The energy deposited in the system is termed the dose, and the measurement process is called dosimetry. This is a distinction from actinometry, referring to the absorption of visible and ultraviolet light. Absorbed energy from ionizing radiation is described in terms of grays (Gy; joule/kg), in rads (100 ergs/g), or in electronvolts per gram or per cubic centimeter.

If the radiation source consists of a uniformly dispersed alpha- or beta-particle-emitting radioactive species, the energy absorption rate may be evaluated from the particle energy and the source intensity. The intensity is determined radiochemically and, for a beta-particle source, the average electron energy must be used. For such as evaluation to be valid, the absorbing sample must be sufficiently large that radiations escaping from the surface can be neglected.

In an aqueous or biological medium containing, for example, 1 millicurie (3.7×10^7 d/s) of iodine-131 (^{131}I; average beta energy, 0.205 MeV) per gram, the dose rate is 1.2×10^{-3} Gy/s or 0.12 rad/s. This will be increased by a small contribution from the accompanying gamma ray.

For systems in which the mechanisms and responses may be determined by the inhomogeneous, microscopic distributions of absorbed energy, for example, biological cells, an average absorbed dose could be inadequate in the interpretation of radiation effects. For the required microdosimetry, specialized techniques and models have been developed.

For a typical radiation-chemical system in which gamma rays such as those from cobalt-60 (^{60}Co) or cesium-137 (^{137}Cs) are the source of ionizing radiation, chemical dosimetry is most usual. The reaction system of interest is replaced in the radiation field with a geometrically similar system having a known chemical response to the absorption of the energy. The latter is a secondary standard that has been calibrated by an absolute method for energy absorption measurement such as calorimetry.

The oxidation of ferrous ion [Fe(II)] to ferric ion [Fe(III)] in aerated solutions containing 0.4 *M* sulfuric acid (H_2SO_4; the Fricke dosimeter) occurs with a *G* value of 15.6 molecules per 100 eV for gamma rays and fast electrons. Dependent upon LET, the *G* values are 5.1 and 3.0 molecules per 100 eV for 5.3-MeV alpha particles and uranium-235 (^{235}U) fission fragments, respectively. The reaction is readily followed by measurement of the Fe(III) absorbance at 305 nm and is used for doses of 50–350 Gy.

Other chemical dosimeters include the cesium [Ce(IV) to Ce(III)] reduction, which occurs in acid solution with a *G* value of 2.5 molecules per 100 eV and the oxalic acid decomposition in an aerated solution having a *G* value of 5 molecules per 100 eV. Both values apply to gamma-ray or fast-electron irradiation.

The radiation chemical system being studied will normally be chemically different from the dosimeter solution. The dose as measured in the dosimeter (D_D) must be corrected to apply to the system of interest. With Compton absorption dominant, the appropriate correction factor is the electron density ratio of the system being studied to that of the dosimeter solution. With doses in grays, rads, or electronvolts per gram, the densities have dimensions of electrons per unit mass; for doses in electronvolts per cubic centimeter, electrons per unit volume are required. These are determined by the chemical composition of the dosimeter and system.

For charged-particle radiation, the incident beam is usually totally absorbed in both dosimeter and sample, and energy absorptions are the same. Geometric distributions may differ, however. For thin samples with partial beam absorption, corrections require the use of atomic stopping powers.

Water. Because of its importance in both chemical and biological systems, the radiation chemistry of water has been extensively studied and serves as an example of radiation-induced chemical change. A primary radiation interaction process may be represented by reaction (1), where H_2O^* represents

$$H_2O \rightarrow H_2O^+, e^-, H_2O^* \tag{1}$$

an electronically excited water molecule. The secondary electron (e^-), if formed with sufficient energy, will form its own trail of ionization and excitation. Within 10^{-10} to 10^{-8} s, reactions within spurs form hydrogen (H) atoms, hydroxyl (OH) radicals, hydrated electrons and molecular products, molecular hydrogen (H_2), and hydrogen peroxide (H_2O_2), as shown in reactions (2)–(7).

$$H_2O^+ + H_2O \rightarrow H_3O^+ + OH \tag{2}$$

$$H_2O^* \rightarrow H + OH \tag{3}$$

$$e^- + H_2O \rightarrow H_2O^-(e^-_{aq}) \tag{4}$$

$$e^-_{aq} + H_3O^+ \rightarrow 2H_2O + H \tag{5}$$

$$2H \rightarrow H_2 \tag{6}$$

$$2OH \rightarrow H_2O_2 \tag{7}$$

In pure water, radicals escaping the spurs undergo further radical-radical reactions and reactions with molecular products. In reactions such as (5) and the reverse of (3), molecular water is formed; and upon continuous irradiation, steady-state concentrations of H_2, H_2O_2, and smaller amounts of dioxygen (O_2) result and no further decomposition occurs.

Initial radical and molecular product yields depend upon the ionization density of the radiation and are a function of LET (**Fig. 1**). The initial *G* value for

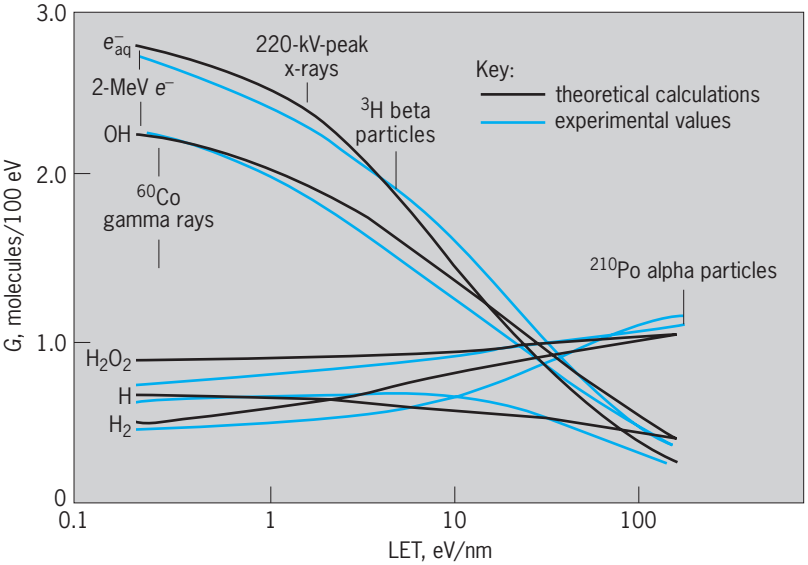

Fig. 1. Variation of *G* with linear energy transfer (LET) for the products solvated electron (e^-_{aq}), OH, H_2O_2, H, and H_2. (*After R. C. Cooper and R. W. Wood, eds., Physical Mechanisms in Radiation Biology, National Technical Information Service, 1974*)

water decomposition, essentially reaction (8), is 4.6

$$2H_2O \rightarrow H_2O_2 + H_2 \tag{8}$$

molecules/100 eV for gamma rays and fast electrons and 2.9 for 12-MeV alpha particles.

In the presence of oxygen, the reducing species, e^-_{aq} and H, form the hydroperoxyl radical and superoxide ion, respectively, and are converted to H_2O_2, as shown in reactions (9)–(13). The hydroperoxyl

$$H + O_2 \rightarrow HO_2 \tag{9}$$

$$HO_2 \rightarrow H^+ + O_2^- \tag{10}$$

$$e^-_{aq} + O_2 \rightarrow O_2^- \tag{11}$$

$$2HO_2 \rightarrow H_2O_2 + O_2 \tag{12}$$

$$O_2^- + HO_2 + H^+ \rightarrow H_2O_2 + O_2 \tag{13}$$

radical behaves as a weak acid [reaction (10)] in aqueous solution. In dilute aqueous systems containing oxygen, the spectrum of radical intermediates is markedly altered, and reactions of the OH radical become most important. This has significant consequences in irradiated biological systems exposed to the atmosphere.

In dilute aqueous solutions, radicals escaping spur recombinations can undergo reaction with the solute. *G* values for these radical yields have been determined from pulsed radiolysis or by a technique in which a specific radical is scavenged by an additive to form a stable molecule. The product is measured, for example, by spectrophotometry. Tetranitromethane [$C(NO_2)_4$] is an effective scavenger for the hydrated electron [reaction (14)]. The product,

$$C(NO_2)_4 + H_2O^- \rightarrow C(NO_2)_3^- + NO_2 + H_2O \tag{14}$$

nitroform [$C(NO_2)_3^-$] is a stable anion with a well-known optical absorption spectrum.

The concept of yield for radicals surviving spur reactions is not a precise one. In addition to a physical LET dependence, it varies slightly with the scavenging system used. For gamma rays and fast electrons and in solutions of pH > 4, generally satisfactory correlations with radiolysis yields are obtained by using 2.8, 2.8, and 0.6 molecules, per 100 eV for e_{aq}, OH, and H radicals, respectively. Corresponding molecular yields for H_2 and H_2O_2 are 0.4 and 0.8 molecule per 100 eV.

In acidic solutions, e^-_{aq} is converted to H through reaction (4), and radical yields are 3.7 and 2.9 molecules per 100 eV for H and OH, respectively. The molecular yields are essentially unchanged.

Radiation chemical mechanisms in aqueous solutions are determined by correlations of chemical change with known radical yields. Oxidation of Fe(II) to Fe(III) in an aerated acid solution (the Fricke dosimeter solution) is an example of such a system. The following mechanism [reactions (15)] has

$$\begin{aligned} H + O_2 &\rightarrow HO_2 \\ HO_2 + Fe^{2+} &\rightarrow HO_2^- + Fe^{3+} \\ HO_2^- + H^+ &\rightarrow H_2O_2 \\ H_2O_2 + Fe^{2+} &\rightarrow Fe^{3+} + OH^- + OH \\ OH + Fe^{2+} &\rightarrow Fe^{3+} + OH^- \end{aligned} \tag{15}$$

been developed for radiolysis. The yield for Fe^{3+}, based upon these radical yields, is 15.6 molecules per 100 eV, consistent with the experimental value.

Applications. In addition to basic kinetics and mechanistic studies, the principles of radiation chemistry find application in any process in which ionizing radiation is used to study, treat, or modify a biological or chemical system.

In radiation therapy, tumors are destroyed by the application of ionizing radiation from external or internally administered sources. Either or both free-radical attack and specific local energy deposition may be involved in cell destruction. Gamma rays are used for treatment of internal tumors; electron or charged-particle beams are applied to external or invasively accessible lesions. Radiation LET is an important variable in treating tumors with particle beams.

The physiological concentration of iodine in the thyroid is the basis for the treatment of hyperthyroidism with ^{131}I. The beta radiation from this isotope is effective in localized tissue destruction.

A goal of any radiation therapy is maximum tumor cell destruction with minimum damage to healthy cells. Chemical sensitization toward, and protection of cells from, radiation damage are areas of continuing research. *See* RADIATION THERAPY.

Disposable medical supplies such as gloves, needles, syringes, and saline solutions are routinely sterilized by radiation. Doses of the order of 10^6 rads (10 kGy) are effective in deactivating most microorganisms.

Radiation processing of sewage sludge enables environmentally safe disposal or conversion to fertilizer. However, while chemically feasible, the process is not economically competitive.

Radiation chemistry and food preservation. Ionizing radiation in sufficient dose is lethal to microorganisms. This was observed shortly after the discoveries of radioactivity and x-rays and was early considered as a means to control the growth of microorganisms in foods.

The inactivation of a microorganism by radiation is a result of reaction by free radicals produced in the organism itself or in the surrounding medium, and by direct action of the radiation on the organism. While all cell constituents are susceptible to radiation damage, it is generally believed that the DNA components in the cell chromosomes are the most critical target. In a direct-action event, nucleic acid molecules are ionized or excited by a primary or secondary electron. A chain of reactions is initiated which may result in damage or death to the organism. In the cytoplasm of cells, which may contain 80% water, free radicals from the water are formed which diffuse to and react with the DNA. Initial chemical events involve radical site formation followed by strand breakage.

The more complex the organism, the greater the sensitivity to radiation damage. The larger DNA units provide a larger target for interaction with radiation and with the radiation-produced intermediates. Viruses, for example, have a greater resistance to radiation than do bacteria. The radiation dose required to kill or inactivate 90% of a specific microbial population is designated as D_{10}. While the effectiveness of a given radiation dose also depends upon the physical state and chemical composition of the surrounding medium, D_{10} values are within the range of 0.1–2 kGy for most bacteria but may be of the order of 10 kGy for some viruses. A dose of 5 kGy has been reported to decrease populations of *Salmonella*, *Staphylococcus aureus*, *Escherichia coli*, *Brucella*, and *Vibrio* by factors of at least 10^6.

The use of ionizing radiation for pathogen control has been approved by most governments for a wide range of foods. In general, limitations on dose have been specified for all products. The Food and Drug Administration (FDA) approved the use of radiation for pathogen control in meat and meat products. A maximum of 4.5 kGy was established for refrigerated meat and meat products and 7.5 kGy for the frozen. The actual dose used for preservation of a commercial product is determined by storage requirements and conditions.

Processing of commercial quantities of food supplies requires a source of stable intensity and a radiation of sufficient penetrating power to deposit energy throughout the product. Current and proposed facilities utilize a conveyor belt on which the product being sterilized is passed for a predetermined exposure time under a cobalt-60 (or cesium-137) gamma-ray source or an electron accelerator. With the accelerator, energetic electrons, for example, 10 MeV, are incident upon a metal film in which the electron energy is converted to the more penetrating x-radiation. With a copper converter, approximately 10% of incident 10-MeV electron energy is converted to x-radiation. If only a surface sterilization is desired, electron beams may be used directly. In the decay of cobalt-60, gammas of 1.17 and 1.33 MeV are emitted; a beta particle emitted in the same decay process is largely absorbed within the source and plays essentially no role in the sterilization process. Cesium-137 emits a single gamma of 0.67 MeV in energy but has a longer half-life of 30 years compared to 5.3 years for cobalt-60.

The radiation processing of food products on a commercial scale requires sources of sufficient intensity to treat large quantities of produce in an economic, brief time period. To achieve this, 10^4–10^6-curie quantities of cobalt-60 or electron accelerators of 10–100-kW beam power are required. While both facility types require extensive shielding for radiation protection, the electron accelerator can be shut off, but the radioisotope source will require more involved and strictly monitored access limitation.

Despite the physical feasibility and federal approval, concern over public acceptance has delayed the implementation of large-scale radiation processing of food supplies. The term "cold pasteurization" has been used in some cases to label food so treated.

Scientific concerns involving possible radiolytic formation of toxic products, destruction of nutritional components of the food, and a detrimental effect on taste or odor have been addressed in over 50 years of extensive research and testing. Radiolysis products in foods are in most cases similar but not identical to those resulting from cooking with heat or by microwave. In addition to water, the main components of foods are carbohydrates, lipids, and proteins. Extensive studies involving analyses for specific products formed in these materials have been carried out. Representative materials, such as starch, corn oil, and alanine, were subjected to radiation doses equal to and greater than that recommended for treatment. In carbohydrates, shorter-chain aldehydes, alcohols, and ketones are the major products of irradiation. From proteins, amino acids and NH_3 and non-nitrogen-containing simpler acids and aldehydes are formed. From lipids, hydrocarbons and aldehydes are major products as well as a variety of esters. Such studies have given no evidence for significant chemical toxins being formed in food subjected to sterilizing doses of radiation.

Most studies of toxicological safety of irradiated foods have involved in vitro experiments with animals. In such studies, an animal population is fed a diet containing a food subjected to a dose of radiation usually many times that proposed for sterilization and increased shelf-life requirements. Comparisons with populations fed a regular diet have demonstrated no adverse effects attributable to the irradiation.

Some vitamin loss does occur in irradiated food samples. Loss normally occurs in the cooking process, and this is sometimes enhanced by irradiation. Thiamine is particularly sensitive. The FDA has concluded that other dietary factors are largely unaffected in nutritional value.

Milk and most dairy products will develop an off-flavor from 10-kGy doses of radiation. Radiation will not be usable for their processing. No

significant tastes or odors have been reported for other foods subjected to sterilization or preservation doses. *See* FOOD PRESERVATION; ISOTOPIC IRRADIATION. Francis J. Johnston

Pulse radiolysis. The chemical effects produced by the absorption of ionizing radiation involve, to a large extent, the production and reaction of free radicals, that is, molecules containing unpaired electrons. In most cases, these free radicals are very reactive and have very short lifetimes, so that they cannot be studied directly by the usual chemical approaches. However, pulse radiolysis methods can be very effectively employed to observe the production and subsequent reaction of free radicals. By using these methods, both chemical and instrumental, the properties of many radicals and other transient intermediates have been described in detail. In many cases, these properties are unavailable from more conventional experiments. A major fraction of pulse radiolysis studies have been carried out on radicals produced in aqueous solutions. Here, the interest is usually in the secondary radicals resulting from reaction of organic or inorganic substrates with the radicals produced in the radiolysis of water, that is hydrogen atoms, hydrated electrons, and hydroxyl radicals. The radiation chemical yields of these initial radicals are known quite well so that quantitative measurements are possible. Since excited states also play an important role in the radiolysis of organic systems, they can be studied by pulse radiolysis techniques. *See* FREE RADICAL.

Pulse radiolysis usually involves observing optical absorption changes following irradiation with nanosecond or microsecond pulses of megaelectronvolt electrons. Once the spectral characteristics of the transient molecules (free radicals) are established, measurements on the time dependence of their optical absorptions allow determination of the rate for their production and reaction. Changes in absorption spectra with time frequently allow details of the mechanisms involved in radical reactions to be elucidated. While optical absorption studies are usually carried out on the nanosecond or microsecond time scale following pulse irradiation, measurements can also be made at times as short as a few picoseconds or at millisecond and longer times. *See* ULTRAFAST MOLECULAR PROCESSES.

Spectroscopic methods. Electron spin resonance and laser resonance Raman spectroscopic methods are also employed in pulse radiolysis. Both approaches provide information on the structure of the transients that otherwise would not be available. Because these approaches are sensitive to the structure of the intermediate, they provide fingerprint identification of the intermediate being examined and allow discrimination against unwanted species. They can be an important adjunct to spectrophotometric experiments. However, very sophisticated equipment is required for such experiments. Conductivity measurements are usually used to provide information on the production of acid or base in oxidation or reduction processes. However, the background conductance of ions effectively restricts these studies to solutions of low ionic strength in the pH range 4–10. Nonpolar media such as hydrocarbons can also be examined by time-resolved conductance methods. In this case, microwave methods are particularly attractive since electrodes are not necessary and polarization problems are absent. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; RAMAN EFFECT.

High-energy electrons. Pulse radiolysis experiments are usually based on pulses of electrons having energies of a few megaelectronvolts from a Van de Graaff or linear accelerator. At these energies the electrons have sufficient penetration that it is possible to irradiate a sample of 1 cm thickness reasonably uniformly. Beam currents during the pulse are usually of the magnitude of amperes, so that a 1-nanosecond pulse of electrons having an energy of 2 MeV produces an initial radical concentration in water of a few micromolar. Many detection methods are sufficiently sensitive that this concentration is adequate for appropriate measurements. It is possible to produce considerably higher initial free-radical concentrations by using longer or more intense pulses, and to produce initial free-radical concentrations in excess of 10 millimolar by using the intense pulses from a commercially available field emission accelerator. However, it is generally desirable to limit the radical concentrations in order to avoid complicating reactions between radicals.

Electron transfer kinetics. Pulse radiolysis is particularly valuable in providing information on electron transfer kinetics and equilibria. In particular, pulse radiolysis allows direct observation of electron transfer reactions. Most data on electron-transfer rates and one-electron reduction potentials of free radicals have been derived from pulse radiolysis. The results obtained represent very significant additions to the information available on a type of reaction which is of considerable importance to many aspects of chemistry. A different approach does not involve free radicals but a determination of the rate for acid dissociation of naphthol (**Fig. 2**). In this case, hydroxide ion, which is a by-product of water radiolysis, attacks the naphthol to produce a naphtholate anion which has a greater absorption than does the naphthol.

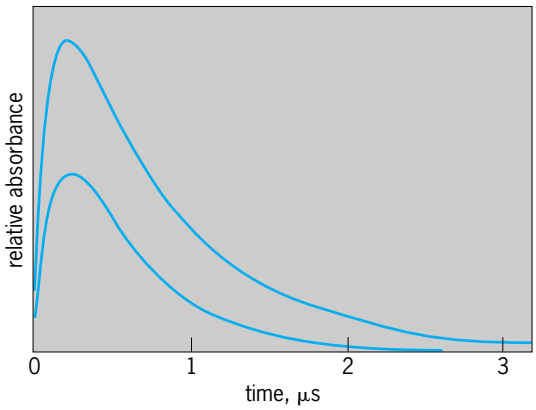

Fig. 2. Growth and decay of the signal of naphtholate anion in the pulse irradiation of a naphthol solution. Determination of the decay rate gives a spontaneous lifetime of 30 milliseconds for loss of the naphthol proton in aqueous solution.

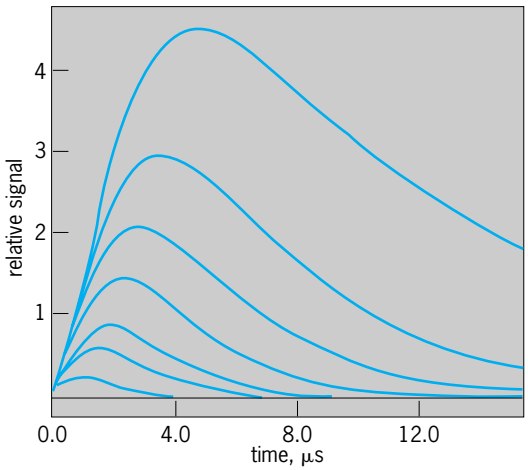

Fig. 3. Time dependence of the electron spin resonance signals of hydrogen atoms in the presence of 50–600-micromolar azide ion. The signal intensities and decay periods decrease with increasing azide ion concentration.

Observations on the recovery in acidic solution provides information on the rate for protonation of the phenolate anion and allows derivation of the spontaneous lifetime for acidic dissociation of the naphthol. Time-resolved electron spin resonance are similarly valuable in providing kinetic information on the reactions of free radicals. In optical studies, hydrogen atom signals are usually masked by other absorptions so that they cannot be observed directly. They are, however, uniquely observed in techniques (**Fig. 3**). Electron spin resonance measurements have provided a significant fraction of the available information on the reaction of hydrogen atoms with organic substrates. Determination of the rates for reaction of radicals with spin traps is another important application of time-resolved electron spin resonance measurements. Time-resolved Raman studies provide information on the vibrational structure of radicals, which is otherwise unavailable. Since the concentrations required for detection can be maintained for only brief periods, pulse techniques are required. Robert H. Schuler

Bibliography. G. V. Buxton et al.,Critical review of rate constants for reactions of hydrated electrons, hydrogen atoms and hydroxyl radicals ($\cdot OH/\cdot O^-$) in aqueous solution, *J. Phys. Chem. Ref. Data*, 17(2):513–886, 1988; W. C. Dewey et al. (eds.), *Radiation Research: A Twentieth Century Perspective*, 1992; J. F. Diehl, *Safety of Irradiated Foods*, 1995; M. Ebert et al., *Pulse Radiolysis*, 1965; *Federal Register*, 62(232):64107–64121, 1997; R. H. Schuler, Three decades of spectroscopic studies of radiation produced intermediates, *Radiat. Phys. Chem.*, 43(5):417–423, 1994; J. W. T. Spinks and R. J. Woods, *An Introduction to Radiation Chemistry*, 3d ed., 1990; Y. Tabata, Y. Ito, and S. Tagawa (eds.), *Handbook of Radiation Chemistry*, 1991; P. Wardman, Reduction potentials of one-electron couples involving free radicals in aqueous solution, *J. Phys. Chem. Ref. Data*, 18(4):1637–1755, 1989; V. M. Wilkinson and G. W. Gould, *Food Irradiation:, A Reference Guide*, 1996.

Radiation damage to materials

Harmful changes in the properties of liquids, gases, and solids, caused by interaction with nuclear radiations. For a discussion of radiation damage in minerals. For a description of damage caused to biological systems by radiation *see* METAMICT STATE; RADIATION INJURY (BIOLOGY)

The interaction of radiation with materials often leads to changes in the properties of the irradiated material. These changes are usually considered harmful. For example, a ductile metal may become brittle. However, sometimes the interaction may result in beneficial effects. For example, cross-linking may be induced in polymers by electron irradiation leading to a higher temperature stability than could be obtained otherwise. *See* RADIOACTIVITY AND RADIATION APPLICATIONS.

Radiation damage is usually associated with materials of construction that must function in an environment of intense high-energy radiation from a nuclear reactor. Materials that are an integral part of the fuel element or cladding and nearby structural components are subject to such intense nuclear radiation that a decrease in the useful lifetime of these components can result. *See* NUCLEAR REACTOR.

Radiation damage will also be a factor in thermonuclear reactors. The deuterium-tritium (D-T) fusion in thermonuclear reactors will lead to the production of intense fluxes of 14-MeV neutrons that will cause damage per neutron of magnitude two to four times greater than damage done by 1–2 MeV neutrons in operating reactors. Charged particles from the plasma will be prevented from reaching the containment vessel by magnetic fields, but uncharged particles and neutrons will bombard the containment wall, leading to damage as well as sputtering of the container material surface which not only will cause degradation of the wall but can contaminate the plasma with consequent quenching. *See* NUCLEAR FUSION; PLASMA (PHYSICS).

Superconductors are also sensitive to neutron irradiation; hence the magnetic confinement of the plasma may be affected adversely. Damage to electrical insulators will be serious. Electronic components are extremely sensitive to even moderate radiation fields. Transistors malfunction because of defect trapping of charge carriers. Ferroelectrics such as $BaTiO_3$ fail because of induced isotropy; quartz oscillators change frequency and ultimately become amorphous. High-permeability magnetic materials deteriorate because of hardening; thermocouples lose calibration because of transmutation effects. In this latter case, innovations in Johnson noise thermometry promise freedom from radiation damage in the area of temperature measurement. Plastics used for electrical insulation rapidly deteriorate. Radiation damage is thus a challenge to reactor designers,

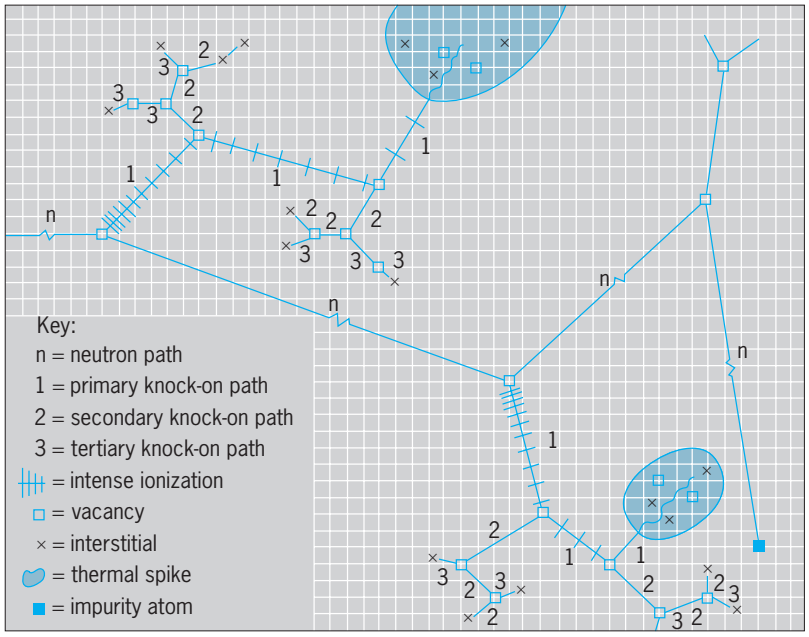

Fig. 1. The five principal mechanisms of radiation damage are ionization, vacancies, interstitials, impurity atoms, and thermal spikes. Diagram shows how a neutron might give rise to each in copper. Grid-line intersections are equilibrium positions for atoms. (*After D. S. Billington, Nucleonics, 14:54–57, 1956*)

materials engineers, and scientists to find the means to alleviate radiation damage or to develop more radiation-resistant materials.

Damage mechanisms. There are several mechanisms that function on an atomic and nuclear scale to produce radiation damage in a material if the radiation is sufficiently energetic, whether it be electrons, protons, neutrons, x-rays, fission fragments, or other charged particles.

Electronic excitation and ionization. This type of damage is most severe in liquids and organic compounds and appears in a variety of forms such as gassing, decomposition, viscosity changes, and polymerization in liquids. Rapid deterioration of the mechanical properties of plastics takes place either by softening or by embrittlement, while rubber suffers severe elasticity changes at low fluxes. Cross-linking, scission, free-radical formation, and polymerization are the most important reactions. *See* RADIATION CHEMISTRY.

The alkali halides are also subject to this type of damage since ionization plays a role in causing displated atoms and darkening of transparent crystals due to the formation of color centers. *See* COLOR CENTERS.

Transmutation. In an environment of neutrons, transmutation effects may be important. An extreme case is illustrated by reaction (1). The ^{6}Li isotope is approximately 7.5% abundant in natural lithium and has a thermal neutron cross section of 950 barns (1 barn = 1×10^{-24} cm^2). Hence, copious quantities of tritium and helium will be formed. (In addition, the kinetic energy of the reaction products creates many defects.) Lithium alloys or compounds are consequently subject to severe radiation damage. On the other hand, reaction (1) is crucial to success of thermonuclear reactors utilizing the D-T reaction since it regenerates the tritium consumed. The lithium or lithium-containing compounds might best be used in the liquid state.

$$^6\text{Li} + n \rightarrow {}^4\text{He} + {}^3\text{H} + 4.8\ \text{MeV} \qquad (1)$$

Even materials that have a low cross section such as aluminum can show an appreciable accumulation of impurity atoms from transmutations. The capture cross section of ^{27}Al (100% abundant) is only 0.25×10^{-24} cm^2. Still reaction (2) will yield several percent of silicon after neutron exposures at fluences of $10^{23} n/\text{cm}^2$.

$$^{27}\text{Al} + n \rightarrow {}^{28}\text{Al}^{\beta-} \xrightarrow{2.3\ \text{min}} {}^{28}\text{Si} \qquad (2)$$

The elements boron and europium have very large cross sections and are used in control rods. Damage to the rods is severe in boron-containing materials because of the ^{10}B(n,α) reaction. Europium decay products do not yield any gaseous elements. At high thermal fluences reaction (3) is most important in nickel-containing materials. The reaction $(n,n') \rightarrow \alpha$ at 14 MeV takes place in most materials under consideration for structural use. Thus, transmutation effects often can be a problem of great importance. *See* NEUTRON SPECTROMETRY; NUCLEAR REACTION; TRANSMUTATION.

$$^{58}\text{Ni} + n \rightarrow {}^{59}\text{Ni} + n \rightarrow {}^{56}\text{Fe} + \alpha \qquad (3)$$

Displaced atoms. This mechanism is the most important source of radiation damage in nuclear reactors outside the fuel element. It is a consequence of the ability of the energetic neutrons born in the fission process to knock atoms from their equilibrium position in their crystal lattice, displacing them many atomic distances away into interstitial positions and leaving behind vacant lattice sites. The interaction is between the neutron and the nucleus of the atom only, since the neutron carries no charge. The maximum kinetic energy ΔE that can be acquired by a displaced atom is given by Eq. (4), where M is mass of the primary knocked-on atom (PKA), m is the mass of the neutron, and E_N is the energy of the neutron.

$$\Delta E = \frac{4Mm}{(M+m)^2} \cdot E_N \qquad (4)$$

The energy acquired by each PKA is often high enough to displace additional atoms from their equilibrium position; thus a cascade of vacancies and interstitial atoms is created in the wake of the PKA transit through the matrix material. Collision of the PKA and a neighbor atom takes place within a few atomic spacings or less because the charge on the PKA results in screened coulombic-type repulsive interactions. The original neutron, on the other hand, may travel centimeters between collisions. Thus regions of high disorder are dispersed along the path of the neutron. These disordered regions are created in the order of 10^{-12} s. The energy deposition is so intense in these regions that it may be visualized as a temporary thermal spike.

Not all of the energy transferred is available for displacing atoms. Inelastic energy losses (electronic

excitation in metals and alloys, and excitation plus ionization in nonmetals) drain an appreciable fraction of the energy of the knocked-on atom even at low energies, particularly at the beginning of its flight through the matrix material. The greater the initial energy of the PKA, the greater is the inelastic energy loss; however, near the end of its range most of the interactions result in displacements. **Figure 1** is a schematic representation of the various mechanisms of radiation damage that take place in a solid.

A minimum energy is required to displace an atom from its equilibrium position. This energy ranges from 25 to 40 eV for a typical metal such as iron; the mass of the atom and its orientation in the crystal influence this value. When appropriate calculations are made to compensate for the excitation energy loss of the PKA and factor in the minimum energy for displacement, it is found that approximately 500 stable vacancy-interstitial pairs are formed, on the average, for a PKA in iron resulting from a 1-MeV neutron collision. By multiplying this value by the flux of neutrons [10^{14-15} $n/(\mathrm{cm}^2)(\mathrm{s})$] times the exposure time [3×10^7 s/yr] one can easily calculate that in a few years each atom in the iron will have been displaced several times.

In the regions of high damage created by the PKA, most of the vacancies and interstitials will recombine. However, many of the interstitials, being more mobile than the vacancies, will escape and then may eventually be trapped at grain boundaries, impurity atom sites, or dislocations. Sometimes they will agglomerate to form platelets or interstitial dislocation loops. The vacancies left behind may also be trapped in a similar fashion, or they may agglomerate into clusters called voids. *See* CRYSTAL DEFECTS.

Effect of fission fragments. The fission reaction in uranium or plutonium yielding the energetic neutrons that subsequently act as a source of radiation damage also creates two fission fragments that carry most of the energy released in the fission process. This energy, approximately 160 MeV, is shared by the two highly charged fragments. In the space of a few micrometers all of this energy is deposited, mostly in the form of heat, but a significant fraction goes into radiation damage of the surrounding fuel. The damage takes the form of swelling and distortion of the fuel. These effects may be so severe that the fuel element must be removed for reprocessing in advance of burn-up expectation, thus affecting the economy of reactor operation. However, fuel elements are meant to be ultimately replaced, so that in many respects the damage is not as serious a problem as damage to structural components of the permanent structure whose replacement would force an extended shutdown or even reconstruction of the reactor. *See* NUCLEAR FISSION; NUCLEAR FUELS.

Damage in cladding. Swelling of the fuel cladding is a potentially severe problem in breeder reactor design. The spacing between fuel elements is minimized to obtain maximum heat transfer and optimum neutron efficiency, so that diminishing the space for heat transfer by swelling would lead to overheating of the fuel element, while increasing the spacing to allow for the swelling would result in lower efficiencies. A possible solution appears to be in the development of low-swelling alloys.

Damage in engineering materials. Most of the engineering properties of materials of interest for reactor design and construction are sensitive to defects in their crystal lattice. The properties of structural materials that are of most significance are yield strength and tensile strength, ductility, creep, hardness, dimensional stability, impact resistance, and thermal conductivity. Metals and alloys are chosen for their fabricability, ductility, reasonable strength at high temperatures, and ability to tolerate static and dynamic stress loads. Refractory oxides are chosen for high-temperature stability and for use as insulators. **Figure 2** shows relative sensitivity of various types of materials to radiation damage. Several factors that enter into susceptibility to radiation damage will be discussed. *See* METAL, MECHANICAL PROPERTIES OF.

Temperature of irradiation. Nuclear irradiations performed at low temperatures (4 K) result in the maximum retention of radiation-produced defects. As the temperature of irradiation is raised, many of the defects are mobile and some annihilation may take place at 0.3 to 0.55 of the absolute melting

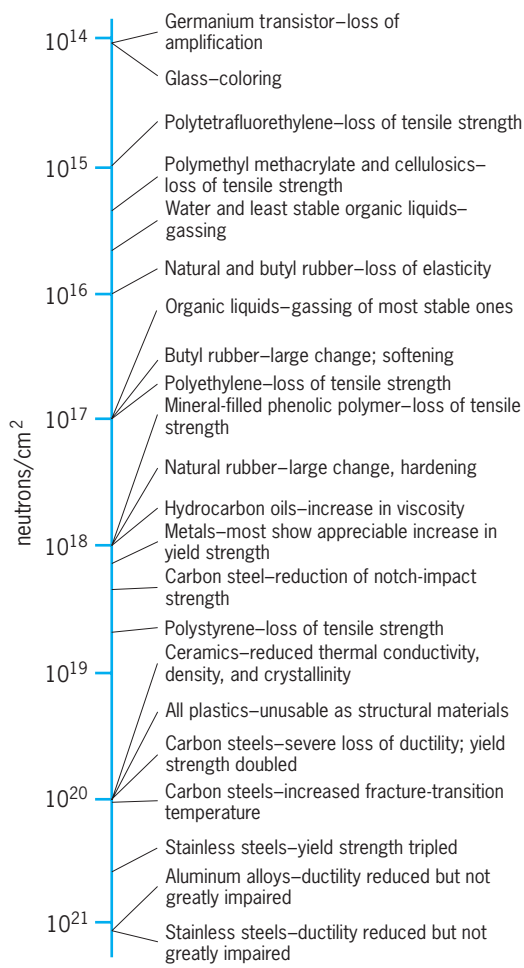

Fig. 2. Sensitivity of engineering materials to radiation. Levels are approximate and subject to variation. Changes are in most cases at least 10%. (*After O. Sisman and J. C. Wilson, Nucleonics, 14:58–65, 1956*)

point T_m. The increased mobility, particularly of vacancies and vacancy agglomerates, may lead to acceleration of solid-state reactions, such as precipitation, short- and long-range ordering, and phase changes. These reactions may lead to undesirable property changes. In the absence of irradiation many alloys are metastable, but the diffusion rates are so low at this temperature that no significant reaction is observed. The excess vacancies above the equilibrium value of vacancies at a given temperature allow the reaction to proceed as though the temperature were higher. In a narrow temperature region vacancy-controlled diffusion reactions become temperature-independent. When the temperature of irradiation is above $0.55T_m$, most of the defects anneal quickly and the temperature-dependent vacancy concentration becomes overwhelmingly larger than the radiation-induced vacancy concentration. However, in this higher temperature region serious problems may arise from transmutation-produced helium. This gas tends to migrate to grain boundaries and leads to enhanced intergranular fracture, limiting the use of many conventional alloys.

Nuclear properties. Materials of construction with high nuclear-capture cross sections are to be avoided because each neutron that is captured in the structural components is lost for purposes of causing additional fissioning and breeding. The exception is in control rods as discussed earlier. Moderator materials, in particular, need to have low capture cross sections but high scattering cross sections. Low atomic weight is an important feature since moderation of fast neutrons to thermal energies is best done by those elements that maximize the slowing-down process. [See Eq. (2).] Beryllium and graphite are excellent moderators and have been used extensively in elemental form. Both elements suffer radiation damage, and their use under high-stress conditions is to be avoided.

Fluence. The total integrated exposure to radiation (flux × time) is called fluence. It is most important in determining radiation damage. Rate effects (flux) do not appear to be significant. The threshold fluence for a specific property change induced by radiation is a function of the composition and microstructure. One of the most important examples is the appearance of voids in metals and alloys. This defect does not show up in the microstructure of irradiated metals or alloys until a fluence of 10^{19} n/cm^2 or greater has been achieved. Consequently, there was no way to anticipate its appearance and the pronounced effect in causing swelling in structural components of a reactor. This and other examples point to the importance of lifetime studies to establish the appearance or absence of any unexpected phenomenon during this time.

Lifetime studies in reactors are time-consuming and are virtually impossible if anticipated fluences far exceed the anticipated lifetime of operating test reactors. A technique to overcome this impasse is to use charged-particle accelerators to simulate reactor irradiation conditions. For example, nickel ions can be used to bombard nickel samples. The bombarding ions at 5 to 10 MeV then simulate primary knocked-on atoms directly and create high-density damage in the thickness of a few micrometers. Accelerators are capable of producing beam currents of several $\mu\text{A/cm}^2$; hence in time periods of a few hours to a few days ion bombardment is equivalent to years of neutron bombardment. Correlation experiments have established that the type of damage is similar to neutron damage. Moreover, helium can be injected to approximate n,α damage when these reactions do not occur in accelerator bombardments. However, careful experimentation is required to obtain correlation between results obtained on thin samples and thicker, more massive samples used in neutron studies. *See* CHARGED PARTICLE BEAMS; PARTICLE ACCELERATOR.

Pretreatment and microstructure. Dislocations play a key role in determining the plastic flow properties of metals and alloys such as ductility, elongation, and creep. The yield, ultimate and impact strength properties, and hardness are also expressions of dislocation behavior. If a radiation-produced defect impedes the motion of a dislocation, strengthening and reduced ductility may result. On the other hand, during irradiation, point defects may enhance mobility by promoting dislocation climb over barriers by creating jogs in the dislocation so that it is free to move in a barrier-free area. Moreover, dislocations may act as trapping sites for interstitials and gas atoms, as well as nucleation sites for precipitate formation. Thus the number and disposition of dislocations in the metal alloy may strongly influence its behavior upon irradiation.

Heat treatment prior to irradiation determines the retention of both major alloying components and impurities in solid solution in metastable alloys. It also affects the number and disposition of dislocations. Thus heat treatment is an important variable in determining subsequent radiation behavior.

Impurities and minor alloying elements. The presence of small amounts of impurities may profoundly affect the behavior of engineering alloys in a radiation field. It has been observed that helium concentrations as low as 10^{-9} seriously reduce the high-temperature ductility of a stainless steel. Concentrations of helium greater than 10^{-3} may conceivably be introduced by the n,α reaction in the nickel component of the stainless steel or by boron contamination introduced inadvertently during alloy preparation. The boron also reacts with neutrons via the n,α reaction to produce helium. The addition of a small amount of Ti (0.2%) raises the temperature at which intergranular fracture takes place so that ductility is maintained at operating temperatures.

Small amounts of copper, phosphorus, and nitrogen have a strong influence on the increase in the ductile-brittle transition temperature of pressure vessel steels under irradiation. Normally these carbon steels exhibit brittle failure below room temperature. Under irradiation, with copper content above 0.08% the temperature at which the material fails in a brittle fashion increases. Therefore it is necessary to control the copper content as well as the phosphorus and nitrogen during the manufacture and heat treatment of these steels to keep the transition temperature

at a suitably low level. A development of a similar nature has been observed in the swelling of type 316 stainless steel. It has been learned that carefully controlling the concentration of silicon and titanium in these alloys drastically reduces the void swelling. This is an important technical and economic contribution to the fast breeder reactor program.

Beneficial effects. Radiation, under carefully controlled conditions, can be used to alter the course of solid-state reactions that take place in a wide variety of solids. For example, it may be used to promote enhanced diffusion and nucleation, it can speed up both short- and long-range order-disorder reactions, initiate phase changes, stabilize high-temperature phases, induce magnetic property changes, retard diffusionless phase changes, cause re-solution of precipitate particles in some systems while speeding precipitation in other systems, cause lattice parameter changes, and speed up thermal decomposition of chemical compounds. The effect of radiation on these reactions and the other property changes caused by radiation are of great interest and value to research in solid-state physics and metallurgy.

Radiation damage is usually viewed as an unfortunate variable that adds a new dimension to the problem of reactor designers since it places severe restraints on the choice of materials that can be employed in design and construction. In addition, it places restraints on the ease of observation and manipulation because of the radioactivity involved. However, radiation damage is also a valuable research technique that permits materials scientists and engineers to introduce impurities and defects into a solid in a well-controlled fashion. *See* ION IMPLANTATION.

Douglas S. Billington

Bibliography. F. L. Bouquet, *Radiation Damage in Materials*, 4th ed., 1994; S.-H. Chen and M. Kotlarchyk, *Interaction of Radiation with Matter and Applications*, 1994; A. Holmes-Siedle and L. Adams, *Handbook of Radiation Effects*, 1993; J. Koutsky and J. Kocik, *Radiation Damage of Structural Materials*, 1994; A. S. Kumar et al., *Effects of Radiation on Materials: 16th International Symposium*, 1993; K. E. Stahlkopf and L. E. Steele (eds.), *Assuring Structural Integrity of Steel Reactor Pressure Boundary Components: Proceedings of the 5th International Seminar*, 1989; L. E. Steele (ed.), *Radiation Embrittlement of Nuclear Reactor Pressure Vessel Steels*, vols. 1–4, 1983–1993.

Radiation hardening

The protection of semiconductor electronic devices and electronic systems from the effects of high-energy radiation. Applications for such devices are in three major areas: (1) satellites, which are exposed to natural space radiation from the Van Allen belts, solar flares, and cosmic rays; (2) electronics, especially sensor and control electronics for commercial nuclear power-generating plants; and (3) perhaps most important, equipment designed to survive the radiation from nuclear explosions.

The expanding technology base has developed with emphasis shifting from one particular type of radiation to another. An early emphasis on neutron fluence and displacement damage effects was followed by a concentration on dose-rate effects from short x-ray and gamma-ray pulses. Next came a large effort to deal with the effects of electromagnetic pulses (EMP) from nuclear weapons on electronic systems and components. The preponderant utilization of complementary metal-oxide-semiconductor (CMOS) devices for satellite electronics increased concern about ionizing-radiation response. This has been incorrectly described as total-dose response, but it usually depends strongly on dose rate as well as total accumulated ionizing dose. The evolution to very large scale integrated-circuit (VLSI) devices has reduced the active volume of devices to levels where a single cosmic ray can produce sufficient ionization to change the logic state of a gate or flip-flop circuit; and this phenomenon, called single-event upset, has aroused strong concern. *See* ELECTROMAGNETIC PULSE (EMP); INTEGRATED CIRCUITS.

Although most radiation effects have been explained and are well understood, much remains to be done. Among the better-understood phenomena are displacement damage from neutrons and photocurrent transients produced by ionizing-radiation pulses. Basic electromagnetic-pulse interactions are also well understood, although their effects on complex electronic systems are extremely difficult to predict because such prediction requires that points of entry and transfer functions be characterized, which is an extremely complex boundary-value problem.

Ionizing dose is discussed with respect to rads as the customary unit of absorbed energy. One rad (material) is the absorption of 100 ergs per gram (10^{-2} joule per kilogram or 10^{-2} gray). Similarly, dose rate is measured in rads (material) per second or ergs per gram per second. Radiation effects depend on both the accumulated ionizing dose (so-called total dose) and the rate at which the dose was accumulated. *See* UNITS OF MEASUREMENT.

Ionizing dose effects. The quasipermanent effects of exposure to ionizing radiation are least understood, and understanding the response of semiconductor devices to this radiation is probably the single most important remaining radiation-hardening problem. Emphasis has been on electrically measurable manifestations such as accumulation of positive charge in silicon dioxide and buildup of negative charge at the oxide–semiconductor interface. The chemical or physical nature of traps has not been adequately explained, although many theories have been set forth. Hardening of metal-oxide-semiconductor (MOS) devices has been accomplished by lower-temperature processing, which probably reduces physical or crystalline defects, and by developing extremely clean processes, which probably reduce chemical defects. The dearth of basic knowledge about damage from ionizing radiation is so serious that presently known electrical, chemical, or physical measurements on semiconductor devices cannot be used to predict behavior in

an ionizing-radiation environment. *See* CRYSTAL DEFECTS; TRAPS IN SOLIDS.

The problem of creating hardened systems for use in ionizing-radiation environments is further exacerbated by the common practice of specifying the environment only in terms of total dose. Semiconductor devices do not have a unique response to total dose; rather they show a complex time-dependent response, which depends on both dose rate and accumulated dose. The damage depends on operating bias and temperature. Thus, it is necessary to specify the environment in terms of dose, dose rate, and temperature.

The most important electrical manifestations of ionizing dose damage in MOS devices are an increase in leakage current; a shift in threshold voltage; and a decrease in speed, transconductance, and channel conductance. The threshold voltage shift is predominantly negative for unhardened devices, and negative at low doses followed by positive at high doses for hardened devices.

Dose-rate effects. Dose-rate effects are well understood. The generation of electron-hole pairs in semiconductors is proportional to the dose; the generation constant for silicon is 4.3×10^{13} electron-hole pairs per rad (4.3×10^{15} pairs per gray). Carriers generated in or near a *pn* junction result in a transient photocurrent proportional to dose rate and the effective volume of the junction. Effective junction volume is the product of junction area and junction width plus a diffusion length on each side of the junction. High dose rates can damage semiconductor devices through logic upset, latch-up, and burnout. In logic upset, logic and memory cells are switched. In latch-up, susceptible devices are triggered into a high current–high voltage state via silicon controlled rectifier action, producing circuit incapacitation and in some cases burnout. Finally, the photocurrents can cause enough energy dissipation, primarily in *pn* junction regions, to cause device burnout. *See* PHOTOVOLTAIC EFFECT.

Displacement damage effects. Displacement damage effects are caused by neutrons, protons, electrons, and other high-energy particles. The production of lattice defects is proportional to the nonionizing energy absorbed by the lattice. The irradiation itself produces Frenkel defects, which are unstable at room temperature and diffuse to produce room-temperature-stable defects, including divacancies, E-centers (donor vacancy defects), and A-centers (oxygen vacancy defects). The defects act as recombination and trapping centers, producing a minority-carrier lifetime decrease proportional to particle fluence. A second-order effect is an increase in resistivity in silicon semiconductors, which is caused by a decrease in mobility and majority-carrier concentration. The dominant effect in bipolar silicon devices is a reduction in common-emitter current gain. This effect is quantitively described by the Messenger-Spratt equation, which shows that the increase in reciprocal common-emitter current gain is directly proportional to particle fluence and inversely proportional to gain-bandwidth product. *See* ELECTRON-HOLE RECOMBINATION.

Single-event phenomena. Single-event upsets are caused at a very low rate in logic and memory circuits by cosmic rays. Rates are low enough (less than 10^{-3} upset per bit per day) that error-detection-and-correction (EDAC) software can be effectively used. This problem is a result of the extremely small effective volumes of large-scale integrated circuits. Integrated circuits with feature sizes greater than 10 micrometers do not show significant single-event upset rates. The upset rate is a function of the maximum linear dimension of the critical device volume, is inversely proportional to the area of the collecting junction, and is proportional to the flux of cosmic-ray particles that exceed the critical linear energy transfer required to produce critical charge in logic devices. This charge is the charge required to change a 0 to a 1 or vice versa, whichever is lower. Single-event upsets may trigger latch-up or second breakdown and hence may also cause catastrophic failure.

Single-event phenomena have led to the development of fault-tolerant architectures to mitigate the effects of random digital upsets, and to the design modification of integrated circuits to prevent an upset even when the active volume is struck by a cosmic ray. Single-event research and development programs are expected to become increasingly important since the problem becomes progressively more serious as integrated circuits utilize smaller active volumes, allowing cosmic rays of lower energy to produce upsets. *See* FAULT-TOLERANT SYSTEMS.

System hardening. Hardening of electronic systems is accomplished by a combination of selecting hardened components, designing circuits more tolerant to radiation-induced degradations, and shielding. Shielding is effective against x-rays and electrons, which have short ranges in materials, and relatively ineffective against gamma radiation, neutrons, and cosmic rays, which have long ranges in materials. Shielding is also effective in mitigating the effect of electromagnetic pulse. *See* RADIATION SHIELDING.

SDI environments. Another major radiation problem results from Strategic Defense Initiative (SDI) environments. Damage effects could result from high-energy-beam weapons, which include lasers, microwave beams, and neutral-particle beams. The frequency range of electromagnetic pulse and the energy range of atomic particles would be extended orders of magnitude beyond the limits thus far considered. Thus, a major research and development program in radiation hardening will be necessary to support the SDI program. *See* RADIATION DAMAGE TO MATERIALS; SEMICONDUCTOR. George C. Messenger

Bibliography. W. R. Dawes, Jr., *Hardening Semiconductor Components Against Radiation and Temperature*, 1990; R. N. Ghose, *EMP Environments and System Hardness Design*, 1984; F. Larin, *Radiation Effects in Semiconductor Devices*, 1968; T. Ma and P. Dressendorfer, *Ionizing Radiation Effects in MOS Devices and Circuits*, 1989; G. Messenger and M. Ash, *The Effects of Radiation on Electronic Systems*, 2d ed., 1992; Summaries of Annual Conferences on Nuclear and Space Radiation Effects, *IEEE Trans. Nucl. Sci.*, December issues.

Radiation injury (biology)

The harmful effects of ionizing radiation on plant and animal cells and tissues. Although the mechanisms of these changes are still poorly known, a wealth of information has been acquired about them. *See* GAMMA RAYS; X-RAYS.

In multicellular organisms, effects on cells are complicated by the interaction of injured and intact cells. Consequently, proper understanding of radiation injury in such organisms calls for appreciation not only of the reaction of individual cells but also of groups of cells in organs and tissues.

Effects of Radiation on Cells

Radiation is known to alter the genetic apparatus of the cell, to interfere with cell division, and to cause cell death.

Effects on genes and chromosomes. The most important action of radiation on the cell is the production of changes in the genetic or chromosomal apparatus. Once established, these changes are largely irreversible and, because of the strategic importance of each gene, may alter the fate of the cell and of its progeny as well. The changes include mutations and chromosome breaks.

Mutation, or gene alteration, consists of a change in the nucleic acid that constitutes the gene. Once established, it may be reversed or otherwise altered only by further mutation. A mutation induced by irradiation is essentially indistinguishable from a natural one. *See* MUTATION.

Chromosome breakage is visible microscopically as a break in the continuity of the chromosomal fiber. The ends of the broken chromosome often reunite within a few minutes, restituting the original structure of the chromosome; but they may unite with fragments of other chromosomes to cause various genetic rearrangements (**Fig. 1**), or they may not rejoin at all.

Radiation may also cause chromosome stickiness and clumping, which presumably affect the chromosomal surface, and may cause chromosomes to adhere to one another and fail to separate normally at cell division. *See* CHROMOSOME ABERRATION.

Effects on cell division. Relatively small amounts of radiation postpone cell division, the delay varying with the dose. When division is eventually resumed in a population of affected cells, the percentage dividing may temporarily exceed, or overshoot, normal; this is often associated with extensive degeneration of the dividing cells. If a large enough dose is administered, the proliferative ability of the cell is permanently abolished. *See* MITOSIS.

Killing of cells. Cells differ markedly in their susceptibility to radiation-induced death, but any cell may be killed if it receives enough radiation. In general, susceptibility varies in proportion to the rate of cell division. Few rapidly dividing mammalian cells are able to survive 5 grays (Gy) of x-rays. (A gray is a unit of absorbed radiation dose, 1 gray equaling 1 joule per kilogram of absorbing tissue.) Although radiation death may occur during or immediately after irradiation, more often it ensues when the cell attempts to divide or after it has divided several times. The relation between the percentage of cells surviving and the radiation dose implies that the dose required to kill all cells in a given population varies with the total number of cells in that population. Hence, although the median lethal dose per cell may be less than 1 Gy, several hundred grays may be required to kill all the billions of cells in a large organ or tumor. This is attributable to the spatial distribution of radiation-induced ion pairs in the absorbing medium, the probability of all cells being appropriately affected decreasing as the total number of cells irradiated is increased.

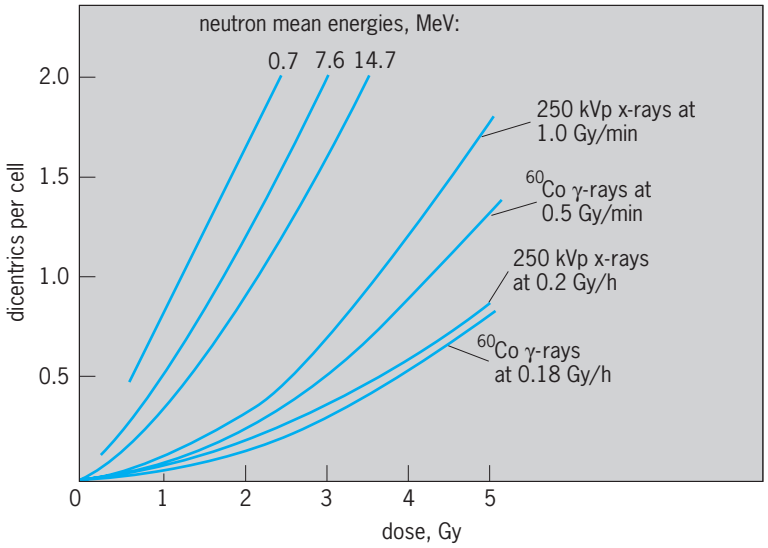

Fig. 1. Frequency of dicentric chromosome aberrations (dicentrics) in human lymphocytes irradiated in the laboratory as a function of the dose and dose rate of neutrons, x-rays, and gamma rays. kVp = kilovolt peak. (*From D. C. Lloyd and R. J. Purrott, Chromosome aberration analysis in radiological protection dosimetry, Radiat. Protect. Dosimetry, 1:19–28, 1982*)

The mechanism by which radiation kills cells is unknown, but the high radiosensitivity of dividing cells, with their tendency to die during cell division, suggests that the most critical type of injury involves the reproductive or chromosomal apparatus of the cell. This is also suggested by the fact that many irradiated cells rendered incapable of division retain their ability to differentiate into more highly specialized forms and to synthesize nucleic acid and protein. Enormous amounts of radiation (many hundred grays) are generally required to stop metabolic activity in such nondividing cells.

Effects of Radiation on Animal Tissues

Tissues differ widely in radiosensitivity, but in general their susceptibility to radiation varies with the rate at which their component cells divide. Accordingly, the most radiosensitive tissues of the body, composed of cells that divide rapidly, are the blood-forming organs, gonads, skin, and intestine.

Radiation injury is not, however, limited to the killing of radiosensitive cells. Invariably, the initial cellular destruction leads to secondary disturbances and reparative processes, often through systemic mechanisms, which modify the primary lesion. In this respect, radiation injury simulates other types of injury and is not unique or specific.

Blood-forming tissues. Since circulating blood cells live only a few days or weeks, they must be replaced continually. The new cells are produced in the bone marrow, spleen, and lymphoid organs through division of blood-cell precursors, which are highly radiosensitive and hence readily destroyed by radiation. Because destruction of these cells leads to shortage of mature blood cells, which may have drastic consequences, damage to the blood-forming organs is one of the most important types of radiation injury. *See* HEMATOPOIESIS.

If only a part of the body is irradiated, blood-forming cells from nonirradiated tissue are carried by the bloodstream to the damaged organs, where they multiply and replace destroyed cells. Consequently, greater quantities of radiation are tolerated if applied to part of the body than if applied to the whole.

The rate of fall in blood count after irradiation varies with the type of blood cell in question, the species, and the radiation dose. In general, the number of cells is depleted more rapidly in small laboratory animals than in humans, the depletion being accelerated in any given species with increasing dose of radiation. In the following remarks, the blood changes are characterized according to their time sequence after a single exposure to whole-body radiation in humans.

The lymphocyte is the first blood cell to be depleted. Diminution in the number of these cells is usually evident within only a few hours after irradiation and becomes maximal within several days. One reason for this may be the relatively short lifespan of such cells; another is the high radiosensitivity of lymphocytes, which is puzzling since they are not rapidly dividing cells. An added factor causing lymphocyte destruction is the nonspecific stress reaction of the body in response to irradiation. Minute amounts of radiation less than 10 millisieverts (mSv); the sievert is a standard unit of radiation dose] cause injured lymphocytes to appear in the blood. The loss of lymphocytes is believed to be responsible to some extent for the depression of immunity that occurs soon after intensive irradiation and that predisposes the affected individual to infection.

The granulocyte, another type of white blood cell, disappears more slowly from the bloodstream. Paradoxically, the number of such cells in the blood may be increased for several days after irradiation, owing to a constitutional reaction of the body to stress. Since mature granulocytes are nondividing cells and are relatively radioresistant, the gradual depletion of their number reflects predominantly the loss of aging cells from the total population, the maximal depression of the granulocyte count not being reached until 3–5 weeks after exposure. Because these cells constitute the body's most important defense against bacterial invasion, infection is a common cause of death in individuals that have been heavily irradiated.

Simultaneously with the fall in granulocyte count, the number of circulating blood platelets also declines, maximal deficiency of these corpuscles also occurring 3–5 weeks after irradiation. Depression of the platelet count below a certain level leads to hemorrhage. The bleeding, which may be either localized or generalized, varies in severity and can be fatal.

The red blood cell count falls gradually in the absence of hemorrhage, severe anemia occurring 5 or more weeks after irradiation. The anemia results primarily from underproduction of new red blood cells owing to destruction of precursors in the bone marrow. Although red-blood-cell-forming elements are usually among the first to regenerate after radiation injury, regeneration may be delayed, and the anemia may therefore be severe and prolonged, even in the absence of bleeding. *See* BLOOD.

Skin. The response of the skin to radiation has been studied extensively, because the skin is the most exposed structure of the body and because its radiosensitivity at one time seriously limited the amount of radiation that could be applied to underlying tumors. The extent of skin injury varies from transient reddening, a reaction formerly used to measure dose in radiotherapy, to ulceration and sloughing. Development of injury is characteristically slow, maximal damage not being evident until weeks after irradiation. Months later, permanent effects consisting of thinning of the skin, scarring of the underlying connective tissue, and dilation of cutaneous blood vessels may gradually appear. This sequence of changes, known collectively as radiation dermatitis, predisposes to subsequent development of skin cancer.

Injury of accessory skin structures is manifested in loss of the hair. The epilation, or loss of hair, appears several weeks after irradiation and may be temporary or permanent, depending on the amount of radiation absorbed.

Gastrointestinal tract. Because the dividing cells lining the small intestine are extremely radiosensitive, relatively small amounts of radiation applied to the abdomen elicit profound effects. These vary from slight disturbances of motility and secretion to ulceration and sloughing of the lining of the bowel.

Nausea and vomiting usually ensue within a few hours after exposure to doses above 2 Gy. After several times this much radiation, sloughing of the intestinal lining may lead to ulceration, intractable diarrhea, dehydration, and invasion of the bloodstream by bacteria that normally inhabit the lumen of the bowel. This sequence is usually fatal and constitutes one of the major causes of death after massive irradiation of the whole body.

Gonads. Since the developing germ cells are highly radiosensitive, their irradiation may result in sterility. Sterility does not usually occur immediately, however, owing to survival of more radioresistant mature eggs or sperm, but only after the supply of these cells is exhausted. Even then, sterility is transitory unless too few precursors survive to resume adequate production of germ cells. In humans, as in most other mammals studied, permanent sterilization requires amounts of radiation that are lethal when absorbed by the entire body; hence sterility is not generally a complication of whole-body irradiation.

Apart from killing of germ cells, mutations may be induced by radiation and be passed on via the

eggs and sperm of successive generations of progeny. The risk of these genetic disturbances is thought to be inceased by even the minutest amounts of radiation, and is therefore considered to be one of the limiting factors in the radiation dose permissible for humans.

Eye. The part of the eye most easily injured by radiation is the lens, opacification of which (cataract formation) has been observed after acute exposure to as little as 2 Gy of x-rays. Smaller doses of neutrons are estimated to have caused cataracts, examples of which have been noted among early neutron physicists. The cornea, conjunctiva, and the retina withstand much more radiation. The retina, however, is highly radiosensitive early in its embryonic development. Minute amounts of ionizing radiation are visible through radiochemical reactions in the retina, which are harmless. *See* EYE (VERTEBRATE).

Nervous system. Although the developing nervous system is highly radiosensitive, only relatively large amounts of radiation will kill nerve cells in the adult. Even in the adult, however, transitory functional disturbances may be elicited by relatively low doses, and after intensive exposure of the brain to a dose in excess of 50 Gy, incapacitating neurological effects may lead to death within minutes or hours. *See* NERVOUS SYSTEM (VERTEBRATE).

Bones and teeth. A dose as low as a few grays applied to bone- and tooth-forming cells in infancy or early childhood cause disturbances of dentition and skeletal growth. In contrast, mature bones and teeth are relatively radioresistant. Large amounts of radiation, however, such as may accumulate from locally deposited radioisotopes or from the treatment of cancer, produce demineralization and necrosis of bone that can lead to fractures, loosening of the teeth, bone cancer, and other complications. *See* BONE.

Vascular system. Transitory dilation of blood vessels, causing erythema or reddening of the skin, is one of the earliest known reactions to ionizing radiation. It occurs after only a few grays and may be accompanied by increased permeability of blood capillaries. Larger amounts of radiation may severely injure or kill cells that line the walls of blood and lymph vessels, giving rise to rupture, occlusion, permanent dilation, or scarring of vessels. Adverse effects on the blood supply may lead to secondary effects, such as metabolic disturbances and atrophy, in involved tissues. *See* CARDIOVASCULAR SYSTEM; LYMPHATIC SYSTEM.

Endocrine glands. The glands of internal secretion have traditionally been regarded as radioresistant because of their ability to withstand relatively large amounts of radiation without developing morphological lesions. There is growing evidence, however, that smaller doses (<10 Gy) may elicit changes in endocrine function and may, in some instances, induce lasting functional impairment. Apart from radiation injury itself, the endocrine system's adaptational response to acute effects of radiation resembles its response to other types of stress. *See* ENDOCRINE SYSTEM (VERTEBRATE).

Urinary system. The kidney and lower urinary tract are relatively radioresistant. Doses in excess of 10–20 Gy may, however, cause gradually progressive scarring and atrophy of the kidney, which can lead to fatal loss of renal function. *See* URINARY SYSTEM.

Lungs. Although relatively radioresistant, the lungs may be injured by intensive irradiation. The result is a chronic, pneumonialike disease, with scarring of lung tissue and blood vessels. Another complication is cancer of the lung, which has been noted in miners of radioactive ore and other irradiated populations.

Effects of Whole-Body Irradiation

When the entire body is irradiated, killing of cells in the various radiosensitive organs causes a complex series of disturbances, the predominant signs and symptoms of which (**Table 1**) are referable to injury of the intestinal tract, blood-forming organs, and skin (acute radiation syndrome). The injury, if severe enough, is fatal within 30 days in most laboratory animals, but in humans death may be delayed until the second month after irradiation.

Susceptibility to death varies from species to species, the average median lethal radiation dose for mammals being about 5 Gy (**Table 2**). As with most other lethal agents, radiation in doses below a certain minimum threshold causes negligible mortality, but radiation in doses above a maximum level is uniformly lethal (**Fig. 2**).

TABLE 1. Symptoms of acute radiation syndrome*

Time after exposure	Supralethal dose, 10 Gy	Median lethal dose, 3–5 Gy	Sublethal dose, 2 Gy
First week	Nausea and vomiting, first day	Nausea and vomiting, first day	
Second week	Continued nausea, vomiting, diarrhea, fever, inflammation of throat, prostration, emaciation, leading to death		
Third week		General malaise, loss of appetite, loss of hair, hemorrhage, pallor, diarrhea, fever, inflammation of throat, emaciation leading to death in 50% of victims	Loss of appetite, loss of hair inflammation of throat, pallor, hemorrhage, diarrhea; recovery begins; no deaths in absence of complications

*After United Nations Scientific Committee on the Effects of Atomic Radiation, *Sources, Effects and Risks of Ionizing Radiation: Report to the General Assembly with Annexes*, 1988.

TABLE 2. Median lethal dose of x-rays for various species of organisms*

Organism	Median lethal dose, Gy
Viruses	
Tobacco mosaic	2000
Rabbit papilloma	1000
Bacteria	
? *coli*	50
? *mesentericus*	1300
Algae	
Mesotenium	85
Pandorina	40
Protozoa	
Colpidium	3300
Paramecium	3000
Vertebrates	
Goldfish	7.5
Mouse	4.5
Rabbit	8.0
Rat	6.0
Monkey	4.5
Human	4.0(?)

*After R. Paterson (ed.), *The Treatment of Malignant Disease by Radium and X Rays,* E. Arnold, 1948.

The cause of death varies with the species, dose, dose rate, and type of radiation. The predominant cause of death in most mammals is injury of blood-forming tissues, which results in infection, hemorrhage, and anemia. Lethal injury of the intestine, usually requiring somewhat higher doses than lethal injury of the marrow, is another major cause of death. In this case, death results from diarrhea, dehydration, loss of body salts, and massive bacterial invasion from the lumen of the bowel. With even larger doses of radiation, death may ensue rapidly from injury of the brain, as mentioned earlier. When the latter mode of death is prevented by shielding the brain, other lethal mechanisms are encountered. Hence, it would appear that sufficient injury of any one of a variety of organs is potentially lethal.

Individuals surviving the acute radiation syndrome usually appear outwardly to have recovered and to be normal by the sixth month after irradiation. That such individuals are not fully recovered, however, is indicated by their subsequent increase in mortality from cancer.

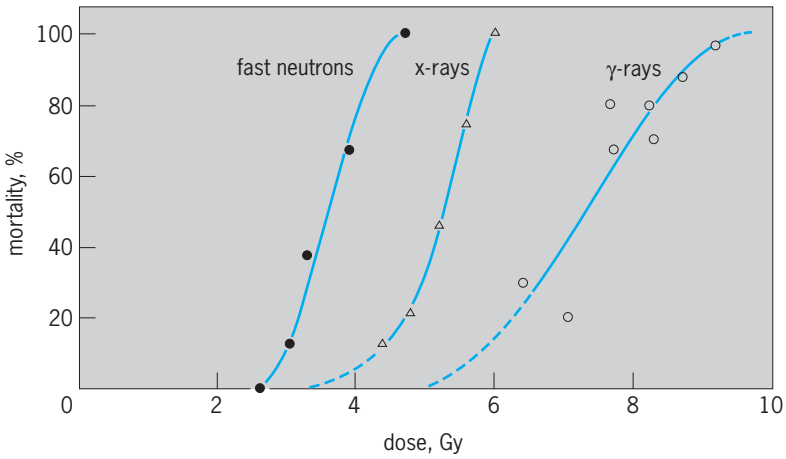

Fig. 2. Thirty-day mortality of mice exposed acutely to x-rays, fast neutrons, and gamma rays. (*After A. C. Upton et al., Radiation Research, vol. 4, Academic Press, 1956*)

Radiation-Induced Cancer

It is paradoxical that ionizing radiation, which is effective in the treatment of cancer, should also be capable of causing cancer. As such, however, it is but one of many agents, including ultraviolet rays, viruses, and a variety of chemicals, that are known to have cancer-inducing potency. The earliest known example of radiation-induced cancer was reported in 1902, less than 10 years after the discovery of x-rays. Since then, numerous instances have been observed in humans, and the process has been studied extensively in experimental animals. *See* ONCOLOGY.

Although susceptibility to cancer induction varies widely among species and organs, virtually all types of cancer have been induced experimentally, and it is clear that all types of ionizing radiation share cancer-forming potency. The radiation-induced cancers that have been noted most often in humans are leukemia and cancers of the skin, breast, lung, thyroid, stomach, colon, and bone.

Cancer of the skin was encountered early in the twentieth century as a complication of radiation dermatitis, developing on the hands and fingers of many of the pioneer radiologists. This effect resulted from prolonged manipulation of radiation equipment, the hands being exposed to large cumulative doses of radiation in the era before the attendant hazards were suspected. Although this disease claimed the lives of more than 100 of the first workers in radiology, it is no longer an occupational threat because of safeguards.

Another malignant disease encountered with unusual frequency in radiologists is leukemia. This disease is also abnormally prevalent in other populations exposed to radiation, such as the Japanese atomic bomb survivors (**Table 3**).

Radiation-induced bone cancer was first noted in painters of luminous watch dials who inadvertently ingested toxic quantities of radium-containing paint by habitually drawing their paint brushes to a point between their lips. Deposition of the radium in the skeleton, where it delivered relatively large doses of radiation to small foci of bone, led to the development of skeletal cancer. Osseous tumors are also abnormally prevalent in those who have consumed radium for medicinal purposes.

Cancer of the lung in miners employed in the pitchblende mines of Saxony, traditionally the predominant cause of death among these workers, has been attributed to their prolonged exposure to radon, which is present in high concentrations in the air of such mines (2–200 becquerels per liter). It is noteworthy, however, that other complicating factors also may have contributed to the effects of radiation in causing lung cancer in this population. *See* RADON.

Cancer incidence versus radiation dose. Although irradiation increases the incidence of many types of cancer, the precise relation between incidence and dose is not known for any type of cancer, especially

TABLE 3. Mortality from leukemia in atomic bomb survivors of Hiroshima and Nagasaki, 1950*

Radiation dose to bone marrow, Gy	Number of survivors	Number of leukemias	Relative risk†
0	35,290	61	1.0
0.01–0.05	19,740	33	1.1
0.06–0.09	4,059	5	1.4
0.10–0.19	5,210	11	1.7
0.20–0.40	6,375	23	2.7
0.50–0.99	3,042	24	4.6
1.0–1.99	1,578	24	8.2
2.0–2.99	412	15	13.4
3.0–3.99	130	2	18.9
4.0 +	155	4	25.0

* After Y. Shimizu, H. Kato, and W. J. Schull, *Life Span Study Report 11,* Part 2: *Cancer Mortality in the Years 1950–1985 Based on the Recently Revised Doses* (DS86), RERF TR 5–88, Radiation Effects Research Foundation, 1988.
† Fitted relative risk, values rounded.

at low radiation dose levels. The yield of tumors per unit dose cannot, therefore, be accurately estimated, nor can it be assumed that any amount of radiation, however small, will increase the cancer rate. At the same time it cannot be assumed that a threshold dose exists below which no cancers are induced. For most experimentally induced cancers the incidence is not linearly proportional to dose, and the evidence suggests that radiation is appreciably less effective at low dose levels than at high dose levels. Paradoxically, moreover, irradiation actually reduces the frequency of certain types of tumors in experimental animals. Accordingly, no generalization about the relation between cancer incidence and radiation dose can be applied to all types of cancer. Estimates of lifetime risks of the major types of radiation-induced cancer in humans are shown in **Table 4**.

Effect of constitutional variables. Susceptibility to cancer induction in experimental animals varies widely with constitutional variables, such as genetic inheritance, age at irradiation, sex, and hormonal activity. Furthermore, the influence of these variables on susceptibility to tumor formation differs depending on the type of tumor and the variables in question. As yet, little is known about the action of these variables on susceptibility in experimental animals and almost nothing about their action in humans.

Mechanisms of cancer induction. The essential change caused by radiation that leads ultimately to the development of cancer is unknown, as is the nature of the cancer change itself. There is growing evidence, however, that the transformation of a normal cell into a cancer cell entails a series of alterations and is not merely a one-step process. This is suggested by the relatively long induction period intervening between irradiation and the onset of malignant growth, a period averaging 5–20 years in humans, depending on the type of cancer induced. It is also suggested by the tendency of naturally occurring cancer to develop late in life and to become progressively more frequent with advancing age.

One theory of cancer formation ascribes the cancer change to one or more mutations in the cell that predispose it to unrestrained growth, perhaps by activating oncogenes or by inactivating tumor-suppressor genes. According to this view, the cancer-forming action of ionizing radiation stems from its mutagenic potency. In some situations, radiation appears to promote the development of cancer through other mechanisms as well, including the activation of latent cancer viruses. *See* CANCER (MEDICINE); ONCOGENES.

TABLE 4. Estimated lifetime risks of different forms of cancer attributable to whole-body irradiation

Type of cancer	Additional fatal cases of cancer per 10,000 persons per gray
Leukemia	90–100
Respiratory tract	50–200
Gastrointestinal tract	130–300
Female breast	40–70
Urinary tract	20–40
Genital organs	20–30
All other	100–300
Total	450–1040
Percentage of normal lifetime cancer risk	2.5–6.0

Effects on the Embryo and Fetus

Radiation injury of human embryos has been amply documented, but most knowledge of the effects of radiation on embryonic and fetal development comes from experiments with laboratory animals. These studies have shown that the individual is more vulnerable to killing by radiation early in development than at any other time of life. The most sensitive period is that preceding implantation of the fertilized egg in the uterus. At this stage, death may be caused by less than one-third the dose required to kill the adult. After implantation and with increasing maturation, susceptibility to killing declines, and death tends to be postponed until birth or thereafter.

The teratogenic effects of radiation depend on the stage of development of the embryo at the time of irradiation, because different parts of the body are formed according to a definite order and time schedule. Malformation of almost any organ may be induced by irradiation at the appropriate moment during or preceding the development of that organ. Since the major period of organ formation occurs early in embryonic life (during the first trimester of

pregnancy in humans), susceptibility to gross malformation decreases in the middle and latter part of gestation. Even late in gestation, however, irradiation may cause abnormalities in the development of slowly maturing organs such as the brain, eye, and gonads. In atomic bomb survivors who were irradiated between the eighth and fifteenth week of embryonic development, the frequency of mental retardation increased with increasing dose. Subtle degrees of abnormality may not be evident at birth.

How radiation causes malformations is not yet fully known, although interference with organ development through the killing of embryonal cells must play an important role in the process. Whatever the mechanism of malformation, the sensitive period for induction of a given abnormality is frequently very short. With increasing radiation dose, however, the sensitive period tends to be prolonged and the incidence and severity of abnormalities increased. Doses as low as 0.25 Gy have been shown to cause developmental disturbances in laboratory animals if applied at the critical time. For this reason, care should be taken to avoid exposing the abdomen of a pregnant woman to radiation early in her pregnancy.

In addition to malformations, other delayed injuries of the type induced by irradiation in adult life may be produced by exposure of the embryo. These, however, tend to manifest themselves long after birth.

Heritable Effects of Radiation

Since radiation-induced mutation may alter characteristics inherited through eggs or sperm and may be passed on through successive generations, such effects are of greater potential significance than those affecting only the exposed individual. Although there is virtually no information about the genetic effects of radiation in human beings, certain facts established by investigations with other organisms appear applicable to all species, including humans.

Natural mutation rate. Through various mechanisms, genes are altered at random during the course of time. These alterations vary from gross chromosomal changes, affecting many genes, to changes affecting only one gene at a time (point mutations). The former are, in general, more deleterious in their effects than the latter, but even the latter are preponderantly detrimental and, in certain instances, lethal. Since genes are present in homologous pairs and one gene of each pair is inherited from each parent, the effects of a mutation depend on whether the mutant gene behaves as dominant or recessive or interacts with other genes. In humans, certain dominant mutations are estimated to appear naturally in 4–40 of every million eggs or sperm. *See* LETHAL GENE.

Radiation-induced mutations. For a given dose of radiation the yield of detectable mutations in germ cells varies from gene to gene, and the relative frequencies of the various mutants differ from those obtained with mutagenic chemicals. Because of this variation and because mutations have been observed thus far in very few of the thousands of genes present in any cell, only crude estimates of the effects of radiation on the overall mutation rate can be made.

Although the frequency of point mutation in some types of irradiated cells varies in direct proportion to the dose, regardless of the dose rate, experiments with mice suggest that the latter may not be true of the complex populations of germ cells in mammals. These experiments disclose that highly fractionated or chronic gamma-irradiation causes fewer detectable mutations in germ cells than the same total dose delivered in a brief single exposure. Because of these complexities, and because of the absence thus far of any evidence of radiation-induced mutations in irradiated human populations, including the children of the atomic bomb survivors, any estimates of the mutation-inducing effectiveness of radiation in humans are of necessity highly speculative. Such estimates have, nonetheless, been made in an effort to assess the genetic hazards of radiation, the mutation rate-doubling dose for humans being considered to be no smaller than 1 Gy.

Practical significance of heritable effects. Until more is known about the extent to which the mind and body are normally handicapped by unfavorable genes, the effects of added deleterious mutants may be only guessed at. Consequently, even if the mutagenic effectiveness of radiation were known accurately, the diverse biological and social consequences of a given radiation-induced increase in the mutation rate could not be estimated at present. It is generally accepted, however, that the genetic burden is regulated by the rates at which mutant genes are produced and subsequently eliminated from the population through natural selection. Hence it is thought that any increase in the mutation rate above natural levels augments the genetic burden.

In the absence of more adequate knowledge of human genetics, attempts have been made to assess the genetic burden in terms of the morbidity from specific genetic traits, such as albinism, achondroplasia, and aniridia. Of such traits, which are detectable in about 4% of all live births, only about one-fourth appear to be attributable to simple genetic mechanisms such as a single mutant gene. Accordingly, if the mutation rate were doubled by radiation, although the latter group would presumably be doubled, the frequency of the other groups would not be increased so greatly. The total incidence then, of all such traits, although elevated from approximately 4% to a level above 5%, would probably still remain below 8%, or twice normal.

In addition to the aforementioned traits, however, the occurrence of which is determined by all-or-none mechanisms, there are characteristics that appear to be quantitatively influenced by heredity, such as general vigor, length of life, fertility, and perhaps, intelligence. Animal experimentation has provided evidence that vigor, length of life, and fertility are reduced in the offspring of irradiated parents, but the data are still fragmentary. Nevertheless, because of the importance of these characteristics, similar radiation effects in humans would be very significant.

Effects of Radiation on Plants

Plant cells have been used extensively for studying the genetic effects of radiation and for investigating the effects of radiation on chromosomes and on cell division. Relatively little work has been done in plants, however, on radiation injury at the tissue or organ level. From available evidence, it appears that radiation effects on plant and animal tissues are not qualitatively different if allowances are made for physiological discrepancies between the two types of organisms.

Irradiation has produced a wide variety of disturbances in plants. These vary from subtle changes resulting from mutations in seeds to marked inhibition of growth and killing. The cause of death is not known, but it is noteworthy that the process of photosynthesis seems relatively insensitive to radiation. Many of the effects of radiation on the growth of higher plants can be attributed to destruction of growth hormone or depression of its synthesis.

Induction of mutations by irradiation of seed has been used to good advantage in breeding new varieties of plants; however, it must be remembered that the mutations so induced occur more or less at random. But since the great majority of mutations are deleterious, one improved plant is gained only at the cost of thousands of others. *See* BREEDING (PLANT).

Factors Affecting Radiation Response

The biological effects of radiation, that is, the response of a cell or tissue to radiation, depend on a variety of direct and indirect factors including radiosensitivity, physical characteristics of the radiation, reaction of surrounding tissues to irradiated cells, and the effectiveness of agents capable of modifying radiosensitivity.

Radiosensitivity. The great variation among cells in susceptibility to radiation injury is largely unexplained, although a number of factors are known to influence radiosensitivity. Of these, perhaps the most important is position on the phylogenetic scale, for example, more than 10,000 Gy are required to kill certain bacteria, whereas less than 10 Gy kill most mammals and less than 1 Gy is lethal for many types of mammalian cells (Table 2).

Another determinant of radiosensitivity is the number of sets of chromosomes present in the cell. This is logical, since each set of chromosomes contains genes duplicating those of another. Hence, in any given family of cells, those with multiple sets of chromosomes are more resistant than those with only a single set. On the other hand, radiosensitivity in certain diploid plants increases in relation to chromosomal content. The radiosensitivity of the cell also varies with the stage in the division cycle at which it is irradiated, as well as with its rate of division, as mentioned earlier. Similarly, the radiosensitivity of the organism varies at different stages in its development.

In addition to the aforementioned intrinsic factors affecting radiosensitivity, extrinsic variables such as temperature, moisture, light, and oxygen tension influence susceptibility. Cells irradiated in the frozen state are, in general, relatively radioresistant, presumably because of the reduced formation and diffusion of radicals at low temperature. The influence of temperature is complex, however; it varies with the time the temperature is altered in relation to irradiation and with the type of cells studied. In most cells, vulnerability to radiation varies with oxygen tension, presumably because oxidizing radicals, such as hydrogen peroxide, are important in the production of radiation injury. Oxygen also exerts other effects, however, since in some types of cells the healing of broken chromosomes requires energy obtained through oxidative pathways.

Physical factors of radiation. The biological effects of radiation depend not merely on the amount of radiant energy absorbed but also on the distribution of the radiation in time (dose rate, or radiation intensity) and space (linear ion density, or linear energy transfer). Certain effects, such as chromosomal rearrangements, generally require more than one ionizing particle or radical to act simultaneously within a small volume of the cell. For multihit effects of this type, densely ionizing radiations such as protons have a high relative biological effectiveness in comparison with sparsely ionizing radiations such as gamma rays.

The dose rate may influence the effects of radiation in other ways too, since some types of injury undergo repair with the passage of time. For such injuries, a given total dose is less effective if fractionated into successive widely separated increments than if administered in a single brief exposure. Paradoxically, however, fractionation in some instances increases the effectiveness of radiation if the successive doses are of appropriate size and periodicity. The influence of fractionation depends, therefore, on many variables, including the total dose, total duration of irradiation, number of exposures, dose per exposure, dose rate per exposure, interval between exposures, and recovery capacity of the system irradiated.

Indirect biological effects of radiation. In addition to affecting the cells irradiated, radiation indirectly alters neighboring or distant cells. These indirect effects may conceivably result from toxic substances liberated by the dying cells or from reactive changes occurring as part of the body's adaptation to injury.

Attempts to demonstrate the liberation of toxic materials from irradiated cells have yielded inconsistent results. Hence it is not established whether such materials are produced in significant quantities. It would appear, however, that they are rarely, if ever, of major importance in the reaction to ionizing radiation.

On the other hand, indirect effects resulting from adaptive or reparative processes are well documented, although their relative importance is not always known. These include alterations caused by local inflammation, scarring, and occlusion of blood vessels, in addition to constitutional changes such as immunological depression, hormonal disturbances, and debilitation.

Modification of radiation injury. It has long been the goal of radiobiologists and radiotherapists to be able to increase or decrease radiosensitivity at will. Radiotherapists have sought ways of increasing destruction of cancer cells without damaging normal tissues. Although their objective is yet to be fully attained, methods for modifying radiosensitivity are now known and are being tested at the experimental and clinical levels.

One of these methods consists in administering drugs that reduce the effectiveness of radiation, principally by blocking or inactivating radiation-induced radicals. The most potent of such drugs yet discovered are cysteine and its derivatives, some of which lower the effectiveness of x-radiation in microorganisms and laboratory animals by a factor of 2 or more. Although most of the chemicals of this type thus far tested have proved too toxic for human use, a few of the newer agents appear promising.

The opposite approach, enhancement of the effectiveness of radiation, is likewise still in the experimental stage. Perhaps the most notable of the radiosensitizing agents studied is oxygen. Irradiation of cancer under increased oxygen tension has been prompted by the fact that many tumors are relatively poorly oxygenated in contrast to normal tissues and to that extent relatively radioresistant.

Among other modifying agents under investigation are chemicals such as nitrogen mustard, urethane, and fluorinated pyrimidines, some of which themselves possess radiomimetic, or radiationlike, activity. These compounds, which are effective by themselves in curbing the growth of cancer cells, exert additive or synergistic effects when administered concomitantly with radiation.

In addition to attempts at preventing or modifying radiation injury by certain processes applied before or during exposure, treatments after irradiation are being explored. Some cellular effects hitherto considered irreversible may be inhibited if appropriate measures are taken promptly enough after irradiation. A particularly notable advance is the discovery that otherwise lethal injury of the bone marrow may be repaired by transplantation of nonirradiated blood-forming cells. Unfortunately, however, since immunological reactions usually occur when the donor of the marrow is genetically different from the recipient, use of this method in humans will be greatly restricted until the immunological complications can be overcome. *See* TRANSPLANTATION BIOLOGY.

Effects of Internally Deposited Radioelements

The radiation injuries caused by radioisotopes within the body are basically no different from those caused by radiation penetrating from without. Because of inhomogeneities of distribution, however, and other variables peculiar to internally deposited emitters, the effects of radioisotopes should be considered separately. *See* RADIOACTIVITY; RADIOISOTOPE.

Anatomic distribution of isotope. In distribution of injury, the area damaged depends on the localization of the radioelement in the organism and the energy of the emitted radiation.

The localization of the isotope is determined by its chemical nature, physical properties, and route of entry into the body. Radioactive iodine-131, for example, behaves chemically like stable iodine, being concentrated from the bloodstream in the thyroid gland. Strontium-90, on the other hand, behaves chemically like calcium, being deposited primarily in the mineral salts of bone. In contrast to these two elements, which become localized in specific organs if dissolved in the bloodstream, tritium, like hydrogen,

TABLE 5. Distribution and excretion of some radioactive fission products and some other radioelements

Radionuclide	Organ or tissue of maximum deposition	f*	Half-life, days: Biological	Half-life, days: Physical
Carbon-14	Total body	1.0	40	2,090,000
Calcium-45	Bone	1.0	(18,000)†	165.2
Cobalt-60	Liver	0.6	6	1,924
		0.2	60	—
		0.2	800	—
Strontium-90	Bone	1.0	(3,900)†	10,512
Iodine-131	Thyroid	1.0	120	8.04
Cesium-137	Muscle	0.1	2	11,012
		0.9	110	—
Barium-140	Bone	1.0	(200)†	12.79
Polonium-210	Liver‡	1.0	50	138.38
	Kidney	1.0	50	—
	Spleen	1.0	50	—
Radium-226	Bone	1.0	(16,000)†	584,000
Uranium-233	Bone	0.9	20	159,000
		0.1	5,000	—
	Kidney	0.996	6	—
		0.004	1,500	—
Plutonium-239	Liver	1.0	14,600	8,800,000
	Bone	1.0	36,500	—

*f = fraction of organ burden retained with designated biological half-life.
† Retention is better represented by a complex exponential plus power function equation.
‡ When more than one organ has been given, the initial deposition has been determined to be equally distributed.

permeates the entire body. If, conversely, these same elements are incorporated into the body as insoluble particulates, their distribution is altered in that they are then concentrated in phagocytic cells, the location of which depends on the portal of entry of the particulates.

The microscopic distribution of a radioelement—whatever its organ distribution—is usually irregular, with the result that its radioactivity tends to be concentrated in foci. For this reason and since the intensity of emitted radiation diminishes with the square of the distance from the radioactive atom, the distribution of radiation damage tends to be patchy, especially when the emission consists of alpha particles or low-energy beta particles. On the other hand, if the emission consists of penetrating gamma rays, distant organs that contain no radioisotope may be injured.

Elimination of isotope from the body. The degree of injury caused by a given radioelement depends on the amount of radiation it delivers to surrounding tissue. This, in turn, depends on the disintegration rate of the isotope (physical decay) and on its rate of elimination from the tissue by metabolism or excretion (biological decay). Since the physical decay of an isotope occurs at an exponential rate and since for practical purposes its elimination from the body also occurs at a more or less exponential rate after its initial uptake and distribution, the amount remaining in the body at any given time (the effective concentration) will vary, depending on the physical and biological half-lives (**Table 5**). There is, of course, no consistent correlation between the two half-lives; both types of values vary enormously from one isotope to another. In general, however, elements that are concentrated in bone tend to be eliminated very slowly and are therefore particularly hazardous. Included in this group are many heavy elements and fission products.

Practical Hazards of Ionizing Radiation

It is evident that, barring occasional massive exposures from nuclear accidents or atomic warfare, the hazards of radiation result from small doses accumulated over a long period of time. The effects of such doses fall into two categories, genetic effects and delayed somatic effects.

Because of concern about the potential danger of fallout from nuclear weapon tests, systematic studies have been made to ascertain the amount of radiation to which humans are now exposed. This radiation comes from (1) naturally occurring radioelements present in the Earth, atmosphere, and human bodies; (2) cosmic rays; and (3) human-made devices, such as x-ray machines, watch dials, nuclear weapons, and television apparatus. It is now apparent that the natural background level has been materially increased by human-made radiation, particularly medical exposure; however, the contribution from weapon fallout is almost negligible to date, despite its notoriety.

Although, as mentioned above, the biological effects of small doses of radiation cannot yet be estimated accurately, certain recommendations have been made in an effort to safeguard humans against exposure to levels of radiation likely to cause detectable injury. The National Council on Radiation Protection and Measurements in the United States has therefore recommended that for the general population the maximum permissible dose of radiation from human-made sources other than medical exposure not exceed 1 mSv per year.

Concerning the hazard of delayed somatic radiation injury such as cancer, there is no conclusive evidence as yet that doses only slightly above the natural background are damaging. It is noteworthy, however, that on the basis of extrapolation from the experimental and clinical data now available, natural background radiation cannot conceivably account for more than a small fraction of the natural incidence of cancer. *See* RADIATION BIOLOGY; RADIOACTIVE FALLOUT.

Arthur C. Upton

Bibliography. C. B. Meinhold (ed.), *Limitation of Exposure to Ionizing Radiation*, 1993; National Academy of Sciences–National Research Council, Committee on Biological Effects of Ionizing Radiation, *Health Effects of Exposure to Low Levels of Ionizing Radiation* (BEIR V), 1990; United Nations Scientific Committee on the Effects of Atomic Radiation, *Sources, Effects and Risks of Ionizing Radiation: Report to the General Assembly with Annexes*, 1988.

Radiation pressure

The force on an object exposed to electromagnetic radiation. It has been known since the days of J. C. Maxwell in the nineteenth century that electromagnetic radiation (which includes visible light) carries both energy and momentum. If radiation impinges on a material body and becomes absorbed, the energy gives rise to heat and is readily detectable. When radiation interacts with an object and is absorbed or scattered, there is also a change in the momentum of the light. By conservation of momentum, this gives rise to a force on the object. This is called radiation pressure. The magnitude of this momentum for visible light is quite small and is difficult to detect. Only near or inside stars, where the intensity is enormous, do light forces have large effects. *See* ELECTROMAGNETIC RADIATION; LIGHT; MAXWELL'S EQUATION; POYNTING'S VECTOR.

This situation is different with lasers. Lasers generate coherent light beams and can be focused to spot sizes of about one wavelength, giving rise to very high intensities and intensity gradients, using modest total powers. Using Planck's law, the energy of a light photon is $E_{\text{photon}} = h\nu$, where h is Planck's constant and ν is the frequency. From Einstein's law, the photon's energy in terms of its effective mass is $E_{\text{photon}} = m_{\text{effective}}\, c^2$, where c is the velocity of light. The photon momentum then is $m_{\text{effective}}\, c = h\nu/c$. If, for example, a laser beam strikes a 100% reflecting mirror, it generates a radiation pressure force $F_{\text{rad}} = (P/h\nu)\ (2h\nu/c) = 2P/c$, where P is the power carried by the beam. For $P = 1$ watt,

$F_{rad} \cong 10^{-8}$ N $= 10^{-3}$ dynes. *See* LASER; PHOTON; PLANCK'S CONSTANT; QUANTUM MECHANICS; RELATIVITY.

For a macroscopic mirror of mass of about 1 gram, the acceleration due to the radiation is $A = F/m \cong 10^{-3}$ cm/s$^2 \cong 10^{-6}\,g$, where g is the acceleration of gravity, about 10^3 cm/s^2. This is quite small. For a 1 micrometer-sized particle, however, with a density of about 1 g/cm^3, the acceleration is $F/m \cong 10^{-3}/10^{-12} = 10^9$ cm/s$^2 \cong 10^6\,g$, which is very large and should be readily observable.

Particle guidance and trapping. Based on these considerations, Arthur Ashkin in 1970 showed experimentally that a fraction of a watt of laser power striking a small transparent polystyrene particle several micrometers in diameter was sufficient to propel it many diameters per second through water in the direction of the light. This force in the direction of the light, called the scattering force, agreed with the above calculation. However, it was discovered that, when the particle was located off the beam axis in the region of a strong gradient of light intensity, there was an additional force component pushing the particle into the high-intensity region at the center of the beam. This component of force was termed the gradient component. It served to guide the particle and keep it on axis as it moved along the beam.

Figure 1 shows that both these force components do indeed originate from radiation pressure. It shows a sphere with high index of refraction, many wavelengths in diameter, placed off axis in a mildly focused beam with a Gaussian intensity profile. A typical pair of rays, a and b, are shown, striking the sphere symmetrically about its center. Neglecting relatively minor surface reflections, most of the rays refract through the particle, giving rise to forces F_a and F_b in the direction of the momentum change. Since the intensity of ray a is higher than ray b,

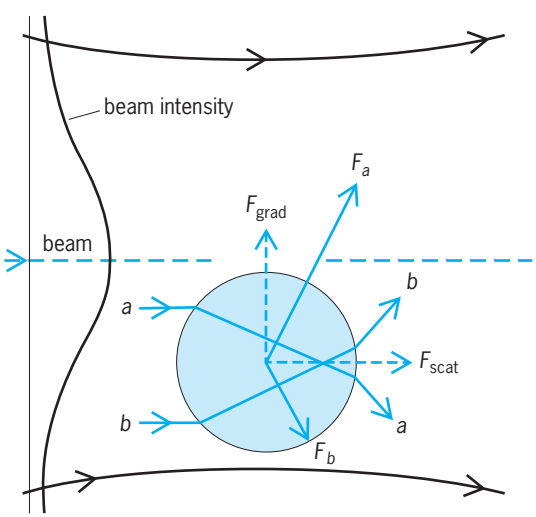

Fig. 1. Origin of scattering force component F_{scat} and gradient force component F_{grad} for a sphere with high index of refraction displaced from the axis of a mildly focused Gaussian beam. Curved lines with arrows indicate rays of the beam.

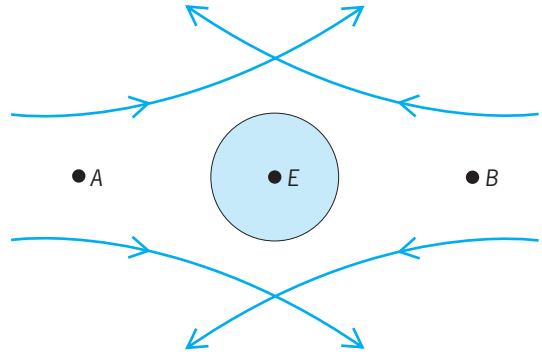

Fig. 2. Geometry of a two-beam trap. Curved lines with arrows indicate rays of the two beams.

the force F_a is greater than F_b. Adding all such symmetrical pairs of rays striking the sphere, one sees that the net force can be resolved into two components: the scattering force component F_{scat}, in the direction of the light; and the gradient component F_{grad} arising from the gradient in light intensity and pointing transversely toward the high-intensity region of the beam. For a particle on axis or in a plane wave, $F_a = F_b$, and there is no net gradient force component.

The understanding of the magnitude and properties of these two basic force components made it possible to devise the first stable three-dimensional (3D) optical trap for single neutral particles. The trap consisted of two opposing moderately diverging Gaussian beams focused at points A and B (**Fig. 2**). The predominant effect in any axial displacement of a particle from the equilibrium point E (that is, in a direction parallel to the beam axes) is a net opposing scattering force. Any radial displacement (that is, in a direction perpendicular to the beam axes) is opposed by the gradient force of both beams. The trap was filled by capture of randomly diffusing small particles which wandered into the trap.

At this time it was proposed that similar stable trapping of atoms was possible using the same two force components. Much progress was made in the 1970s in understanding light forces on macroscopic transparent dielectric particles and atoms. In 1972 Ashkin and colleagues stably levitated transparent particles against gravity using a single, mildly divergent, vertically directed laser beam in air and in vacuum. In vacuum an optical feedback system was developed to lock levitated particles to an external reference and also to provide optical damping to avoid fluctuation in position and possible escape of particles in the absence of viscous air damping. It acted as a sensitive force-measuring device as well, since the application of any type of external force resulted in a change in the laser power needed to maintain a fixed position. High-resolution measurement of the charge of the electron was made using an applied electric field in a modern version of the Millikan oil-drop experiment. With a tunable dye laser with feedback, extremely high Q resonant responses of a transparent

oil droplet were observed for the first time. These resonances corresponded to optical whispering gallery modes of the spherical droplet. They gave a means of determining particle size and the index of refraction to unprecedented accuracy. Highly resonant quartz spheres can be used as filters in optical communications, as resonators in quantum electrodynamics experiments, and as high-Q resonators in nonlinear optics experiments.

Atom traps. With atoms, the scattering force can be understood as the absorption of resonant photons and their subsequent reemission. The gradient force on atoms can be thought of as force on an optically induced dipole in the gradient of the optical electric field. In the early 1970s, experiments were done on the deflection of an atomic beam by a resonant light beam for purposes of isotope separation. This geometry avoids the effects of Doppler shifting.

In 1973 Theo W. Hänsch and Arthur L. Shawlow proposed using the Doppler shift to provide optical damping or cooling of an atomic gas. If a pair of opposing laser beams in one dimension is tuned below an atomic resonance, then an atom moving in either direction feels a net damping because it is Doppler-shifted to a higher frequency, closer to resonance for the opposing beam, and to a lower frequency, away from resonance, for the trailing beam. With three pairs of beams in three dimensions, all the degrees of freedom can be damped. This creates what is known as an optical molasses. Residual fluctuations of the scattering force limit the cooling to a temperature known as the Doppler limit. For sodium atoms this is about 240 microkelvin. Attempts to design traps to capture atoms on the intensity maxima of the standing waves using the gradient force have problems because of excessive saturation, or the ability of the atom to absorb and emit only a finite number of photons. In 1978 Ashkin solved this problem by designing gradient force traps detuned far below resonance to reduce absorption. Two-beam traps and a single-beam, strongly focused gradient trap were proposed for holding the atoms, cooled by optical molasses. The single-beam gradient trap, later referred to as a tweezers, is the simplest of all traps, consisting of a single strongly focused Gaussian beam. *See* DOPPLER EFFECT; LASER COOLING; PARTICLE TRAP.

The proposed technique to increase the gradient force was checked in 1978 in experiments showing focusing and guiding of atoms in detuned light beams. This marked the beginning of what is known as atom optics, and also suggested that stable atom traps were feasible. By 1984 William D. Phillips and Wolfgang Ertmer succeeded in overcoming the strong Doppler shifts involved in slowing thermal atomic beams to a temperature of about 1 K, using the scattering force of an opposing light beam. Such atoms provided an excellent source for further three-dimensional cooling. In 1985 Steven Chu entered the field and, with J. E. Bjorkholm, Ashkin, and L. W. Hollberg, demonstrated the first molasses cooling of sodium atoms to around 240 microkelvin. These atoms were used in 1986 to demonstrate the first stable trapping of neutral atoms using the off-resonance single-beam gradient trap. Later in 1986 another type of large-volume, molasses-cooled trap called a magneto-optic-trap (MOT) was demonstrated. This could trap many more atoms than the first tweezer trap.

Overcoming the Doppler limit. In 1988 Phillips and his colleagues showed that by changing the molasses detuning, it was possible to reach a new cooling regime, where the atom's temperature was around 20 microkelvin, ten times lower than the Doppler limit. The theory of this process was explained principally by Claude Cohen-Tannoudji and his colleagues and was called Sisyphus cooling. Applications of these ultracold atoms soon appeared, such as improved atomic clocks using atomic fountains, atom interferometers to measure the gravitational force with much improved accuracy, and the study of cold-atom collisions in the regime of large de Broglie wavelengths. *See* ATOMIC CLOCK.

Biological applications. In 1986 Ashkin and his colleagues showed that tweezer traps could also trap submicrometer and micrometer-sized dielectric particles. In further experiments in 1987, Ashkin and J. M. Dziedzic serendipitously discovered that infrared tweezers could trap living bacteria, viruses, and other biological organisms without causing optical damage. It was even possible to optically manipulate organelles inside larger cells and observe the mechanical properties of cytoplasm. This initiated the application of laser tweezers to biology.

Study of motor molecules. A major application of tweezers is the study of motor molecules that drive organelle movement, cell locomotion, and muscle action. Steven M. Block and his colleagues improved earlier in-vitro motility assays by optically manipulating single molecules attached to dielectric spheres in what is known as the handles technique. For kinesin they resolved the detailed motion along a microtubule into a sequence of 8-nanometer steps. They measured the complete force-velocity relationship with a maximum force of about 5 to 6 piconewtons. For myosin, Jeffrey T. Finer and his colleagues used a new feedback-enhanced dual optical trap to measure stepwise motion on actin filaments of about 11 nm and forces of 3 to 4 pN. These studies often used mutant forms of actin and kinesin molecules. They measured the ATP hydrolysis cycle and single-enzyme kinetics, and the efficiency of kinesin motors. Another advance was the development of a novel feedback-controlled molecular force clamp that maintains constant force on moving single molecules. Such force clamps were used to study another class of motors, the nucleic acid motor enzymes such as RNA polymerase, which synthesizes a RNA replicate of DNA, and DNA polymerase, which catalyzes DNA replication. These motors are slow and powerful. Resolutions down to about a single DNA base pair were achieved, and novel behavior only previously

postulated was observed, such as the backtracking by RNA polymerase during long pauses in transcription, during which the proofreading and error-correcting mechanisms can cleave and discard flawed sections. Block and his colleagues, using identically prepared strands, showed that RNA polymerase tracks the DNA groove during the search for the promoter sequence that initiates transcription. Carlos Bustamante introduced the study of DNA polymerase. Data on replication in base pairs per second taken at different tensions using constant force feedback gave the kinetic tension or stall force at which replication ceases. *See* CELL MOTILITY; CELL ORGANIZATION; MITOSIS; MUSCLE PROTEINS; NUCLEOPROTEIN.

Other DNA-associated molecules were studied with tweezers, including chromatin, protamines, arginines, MukBEF, and RecBCD. Many structural measurements were made on DNA and other macromolecules, including stretching and applying controlled torques. Using traps, David G. Grier and others have made the surprising observation of confinement-induced attractive forces between like colloidal particles in charge-stabilized suspensions. Despite considerable theoretical attention, this effect remains unexplained.

Optically driven motors. The rotation of particles by light forces has been much studied using laser beams. The rotation of shaped (vanelike) particles caused by the linear momentum of light was observed. Also, rotation induced by the intrinsic angular momentum of the photon and the orbital angular momentum of higher-order rotating light beam modes was investigated. Such optically driven motors are expected to be important for study of physical and biological materials.

Fluidic sorting. Another topic studied was the fluidic sorting of particles and cells using dynamic holographic tweezers. With nematic spatial light modulators it is possible to form arrays with hundreds of adjustable tweezer traps. One can spatially sort mixtures of particles by flowing them through fixed arrays of weak potentials or by capturing them and transporting them through space in what is called optical peristalsis. Or one can drive particles through lithographically formed microchannels using optically driven gear-type fluid pumps and gate valves. Sorters of this type have advantages over large-volume conventional-flow cytometric sorters.

Bose-Einstein condensation. A major advance involves evaporative cooling of atoms from magnetic quadrupole traps to submicrokelvin temperatures (the coldest temperatures in the universe) at which Bose-Einstein condensation (BEC) can occur. In Bose-Einstein condensation, the de Broglie wavelengths of atoms overlap, and they occupy the lowest energy state of the trap, where they behave as a single quantum particle. Experiments have shown the coherence of atoms in such condensates, and crude atom lasers were demonstrated. Applications of Bose-Einstein condensates are possible: in Josephson devices, coherent atom lithography, quantum computers and cryptography, and even more accurate atomic clocks. *See* ATOMLASER; BOSE-EINSTEIN CONDENSATION.

Subsequently, fermionic atoms were cooled well below the Fermi temperature. Using all-optical traps, it is possible to magnetically tune the interaction strength across Feshbach magnetic resonances for all hyperfine levels without affecting the trapping ability. In this way molecular Bose-Einstein condensates were formed for the first time, when tuned above a Feshbach resonance. Strongly interacting fermionic pairing was also observed for tunings below resonance where the gas acts as a superfluid. Weakly interacting Cooper pairs, as in superconducting solids or superfluids, have not as yet been achieved. Using all-optical traps, it was possible to demonstrate five signatures of strongly interacting pairs. It is now clear that optical traps are superior to magnetic traps for making and exploiting fermionic Bose-Einstein condensates. Observations on fermionic pairing may shed light on the phenomenon of high-temperature superconductivity. Radiation pressure forces from lasers can be expected to play an important role in physics, chemistry, and the biological sciences in future work on small particles. *See* RESONANCE (QUANTUM MECHANICS); SUPERCONDUCTIVITY.

Arthur Ashkin

Bibliography. A. Ashkin, Applications of laser radiation pressure, *Science*, 210:1081–1088, 1980; A. Ashkin, History of optical trapping and manipulation of small neutral particle, atoms, and molecules, *IEEE J. Selected Topics Quant. Electr.*, 6(6):841–855, 2000; K. C. Neuman and S. M. Block, Optical tapping, *Rev. Sci. Instrum.*, 75(9):2787–2809, 2004.

Radiation shielding

Physical barriers designed to provide protection from the effects of ionizing radiation; also, the technology of providing such protection. Major sources of radiation are nuclear reactors and associated facilities, medical and industrial x-ray and radioisotope facilities, charged-particle accelerators, and cosmic rays. Types of radiation are directly ionizing (charged particles) and indirectly ionizing (neutrons, gamma rays, and x-rays). In most instances, protection of human life is the goal of radiation shielding. In other instances, protection may be required for structural materials which would otherwise be exposed to high-intensity radiation, or for radiation-sensitive materials such as photographic film and certain electronic components.

For the effects of radiation on living organisms *see* HEALTH PHYSICS; NUCLEAR RADIATION (BIOLOGY); RADIATION BIOLOGY; RADIATION INJURY (BIOLOGY) For the effects on inanimate materials *see* RADIATION DAMAGE TO MATERIALS; NUCLEAR EXPLOSION.

Radiation sources. Nuclear-electric generating stations present a variety of shielding requirements. In the nuclear fission process, ionizing radiation is

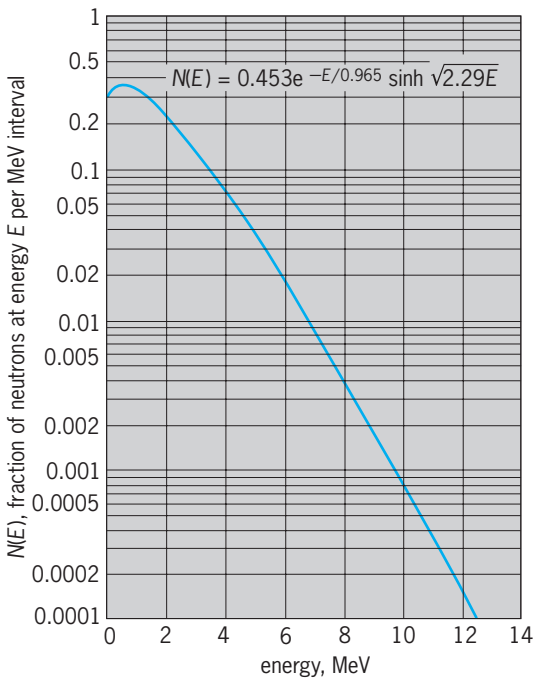

Fig. 1. Neutron spectrum as produced from the fission of ^{235}U by thermal neutrons. (*After L. Canberg et al., Phys. Rev., 103:662–670, 1956*)

released not only at the instant of fission but also through radioactive decay of fission products and products of neutron absorption. Of the energy released in the fission process, about 2.5% is carried by prompt fission neutrons and 3% by delayed gamma rays. About 2.5 neutrons are released per fission with average energy of 2 MeV. About 90% of the prompt neutrons have energies less than 4 MeV, but because of their greater penetrating ability the remaining 10% are of greater concern in radiation shielding. Neutrons are distributed in energy approximately as shown in **Fig. 1**, which gives the distribution for fission of uranium-235. The distribution for other fissionable isotopes is similar. *See* NUCLEAR FISSION; NUCLEAR REACTOR; REACTOR PHYSICS.

Fission of one atom releases about 7 MeV of energy in the form of prompt gamma rays. This energy is distributed over some 10 gamma-ray photons with the energy spectrum as shown in **Fig. 2**. Evaluation of the intensity of delayed gamma-ray sources within a nuclear reactor requires knowledge of the operating history of the reactor. Absorption of neutrons in structural or shielding materials results in the emission of capture gamma rays. Likewise, gamma rays result from inelastic scattering of neutrons. Because these gamma rays (especially capture gamma rays) are generally of higher energy than prompt fission gamma rays and because they may be released deep within a radiation shield, they require careful consideration in shielding design for nuclear reactors. *See* GAMMA RAYS; NEUTRON; NEUTRON SPECTROMETRY; NUCLEAR REACTION.

Controlled thermonuclear reactors, deriving energy from the nuclear fusion of deuterium and tritium, present radiation shielding requirements similar in kind to those of fission reactors. Highly penetrating (14-MeV) neutrons are released in the fusion process along with charged particles and photons. Capture gamma rays again require careful consideration. *See* NUCLEAR FUSION.

X-ray generators vary widely in characteristics. Typical units release x-rays with maximum energies to 250 keV, but high-voltage units are in use with x-ray energies as high as tens of megaelectronvolts. The dominant nature of x-ray energy spectra is that of bremsstrahlung. X-ray generators are but one of many types of charged-particle accelerators and, in most cases, the governing shielding requirement is protection from x-rays. For very high-energy accelerators, protection from neutrons and mesons produced in beam targets may govern the shield design. *See* PARTICLE ACCELERATOR; RADIOLOGY; X-RAYS.

Although many different radionuclides find use in medical diagnosis and therapy as well as in research laboratories and industry, radiation shielding requirements are of special importance for gamma-ray and neutron sources. Alpha and beta particles from radionuclide sources are not highly penetrating, and shielding requirements are minimal. Space vehicles are subjected to bombardment by radiation, chiefly very high-energy charged particles. In design of the space vehicle and in planning of missions, due consideration must be given to radiation shielding for protection of crew and equipment. *See* COSMIC RAYS; RADIOISOTOPE; RADIOISOTOPE (BIOLOGY).

Attenuation processes. Charged particles lose energy and are thus attenuated and stopped primarily as a result of coulombic interactions with electrons

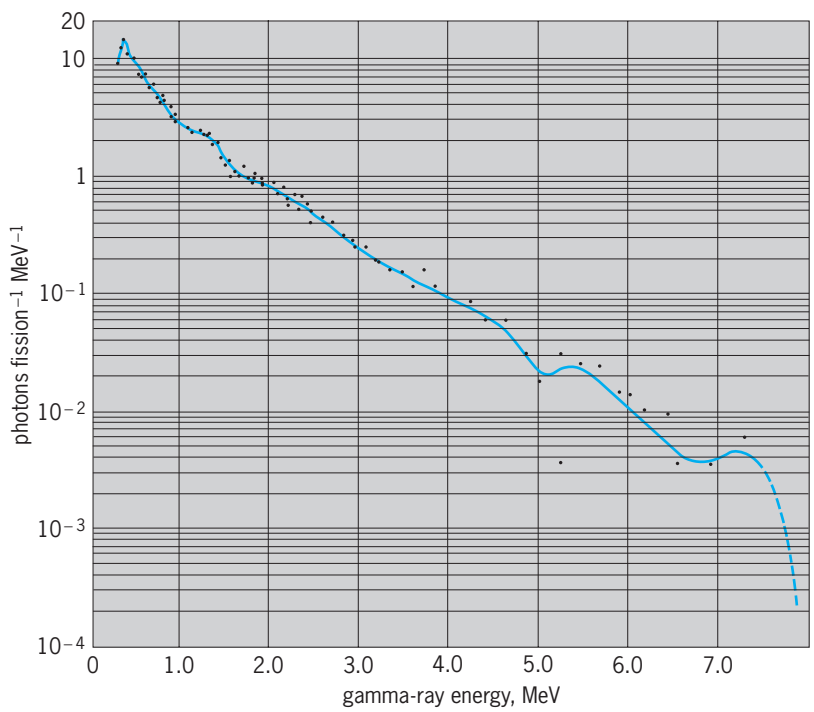

Fig. 2. Energy spectrum of gamma rays observed within 10^{-7} s after fission.

of the stopping medium. For heavy charged particles (protons, alpha particles, and such), paths are nearly straight and ranges well defined. Electrons may suffer appreciable angular deflections on collision and may lose substantial energy radiatively. Very high-energy charged particles may lose energy through nuclear interactions, resulting in fragmentation of the target nuclei and production of a wide variety of secondary radiations.

Gamma-ray and x-ray photons lose energy principally by three types of interactions: photoemission, Compton scattering, and pair production. In photoemission, or the photoelectric effect, the photon transfers all its energy to an atom, and an electron is emitted with kinetic energy equal to the original energy of the photon less the binding energy of the electron in the atom. In Compton scattering, the photon is deflected from its original course by, and transfers a portion of its energy to, an electron. In pair production, the gamma ray is converted to a positron-electron pair. At least 1.02 MeV of gamma-ray energy is required for the rest mass of the pair, and any excess appears as kinetic energy. Ultimately, the positron and an electron recombine and, in annihilation, release two 0.505-MeV gamma rays. Photoemission is especially important for low-energy photons and for stopping media of high atomic number. Compton scattering usually dominates at intermediate photon energies, and pair production at high photon energies. *See* COMPTON EFFECT; ELECTRON-POSITRON PAIR PRODUCTION; PHOTOEMISSION.

Neutrons lose energy in shields by elastic or inelastic scattering. Elastic scattering is more effective with shield materials of low atomic mass, notably hydrogenous materials, but both processes are important, and an efficient neutron shield is made of materials of both high and low atomic mass. The fate of the neutron, after slowing down as a result of scattering interactions, is absorption frequently accompanied by emission of capture gamma rays. Suppression of capture gamma rays may be effected by incorporating elements such as boron or lithium in the shield material. The isotopes ^{10}B and ^{6}Li have large cross sections for neutron capture without gamma-ray emission.

Shielding concepts. The cross section of an atom or electron for interaction with radiation is the effective "target" area presented for the interaction. It is usually given the symbol σ and the units cm^2 or barns (1 barn = 10^{-24} cm^2). The cross section depends on the type of interaction and is a function of the energy of the radiation. When the cross section is multiplied by the number of atoms or electrons per unit volume, the product, identified as the linear attenuation coefficient μ or macroscopic cross section Σ, has the units of reciprocal length and may be interpreted as the probability per unit distance of travel that the radiation experiences in an interaction of a given type. The total attenuation coefficient for radiation of a given energy is the sum of attenuation coefficients for all types of interactions in the shielding medium. The quotient of the linear attenuation coefficient and

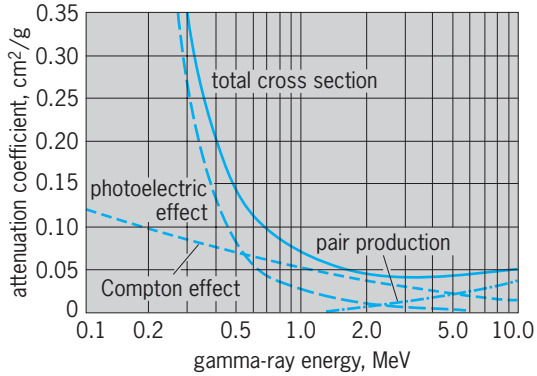

Fig. 3. Gamma-ray attenuation coefficients for lead.

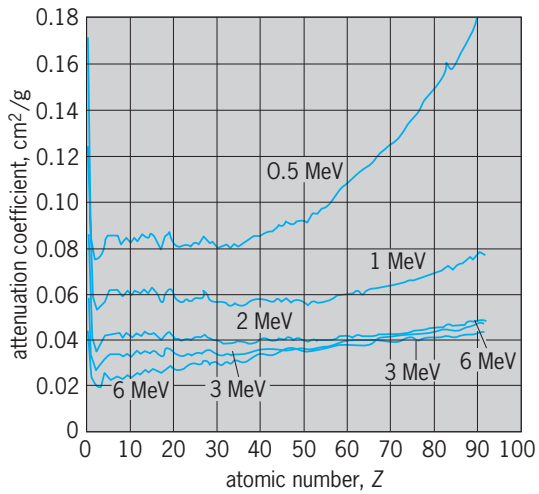

Fig. 4. Gamma-ray attenuation coefficients versus atomic number of absorbing material for various energies.

the density, μ/ρ, is called the mass attenuation coefficient. If μ/ρ for gamma rays is weight-averaged by the fraction of the gamma-ray energy locally dissipated subsequent to an interaction, the average, μ_a/ρ, is called the mass energy absorption coefficient. The total and component parts of the mass attenuation coefficients for lead are illustrated in **Fig. 3**. In **Fig. 4** are shown total mass attenuation coefficients for all elements at several photon energies.

The flux density, or fluence rate, characterizes the intensity of radiation. It may be thought of as the path length traveled by radiation per unit volume per unit time. It is usually given the symbol ϕ and has units $cm^{-2}\ s^{-1}$.

To illustrate these concepts, consider the flux density at a distance r from a point source isotropically emitting S monoenergetic gamma rays of energy E per second in a uniform medium with total attenuation coefficient μ. If μ were zero, the flux density would be determined just by the inverse square of the distance from the source. If μ were not zero, attenuation of the gamma rays would also be exponential with distance. Thus the equation shown below

$$\phi = \frac{S}{4\pi r^2} e^{-\mu r}$$

would hold. This is the flux density of gamma rays which have traveled distance r without having experienced any interactions in the stopping medium. The energy locally dissipated per unit mass per unit time due to these gamma rays as they experience their first interactions is the product $E(\mu_a/\rho)\phi$. However, some gamma rays reach distance r already having experienced scattering interactions. To account for these secondary gamma rays, a buildup factor B is employed. B is a function of the energy of the source gamma rays and the product μr and depends on the attenuating medium. The total energy locally dissipated per unit mass per unit time is thus the product $BE(\mu_a/\rho)\phi$. Similar concepts apply to neutron shielding; however, the treatment of scattered neutrons is considerably more complicated, and the buildup-factor concept is not well established. In the final phases of shielding design, digital computers are usually employed to carry out the required calculations. It is also common practice to base shielding design, at least in part, on measurements made on a prototype.

Shielding materials. The most common criteria for selecting shielding materials are radiation attenuation, ease of heat removal, resistance to radiation damage, economy, and structural strength.

For neutron attenuation, the lightest shields are usually hydrogenous, and the thinnest shields contain a high proportion of iron or other dense material. For gamma-ray attenuation, the high-atomic-number elements are generally the best. For heat removal, particularly from the inner layers of a shield, there may be a requirement for external cooling with the attendant requirement for shielding the coolant to provide protection from induced radioactivity.

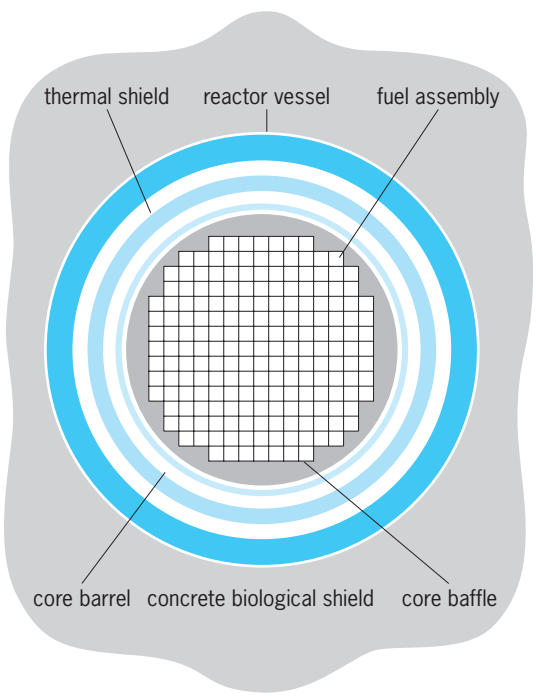

Fig. 5. Typical power reactor shield configuration.

Metals are resistant to radiation damage, although there is some change in their mechanical properties. Concretes, frequently used because of their relatively low cost, hold up well; however, if heated they lose water of crystallization, becoming somewhat weaker and less effective in neutron attenuation.

If shielding cost is important, cost of materials must be balanced against the effect of shield size on other parts of the facility, for example, building size and support structure. If conditions warrant, concrete can be loaded with locally available material such as natural minerals (magnetite or barytes), scrap steel, water, or even earth.

Typical shields. Radiation shields vary with application. The overall thickness of material is chosen to reduce radiation intensities outside the shield to levels well within prescribed limits for occupational exposure or for exposure of the general public. The reactor shield is usually considered to consist of two regions, the biological shield and the thermal shield. The thermal shield, located next to the reactor core, is designed to absorb most of the energy of the escaping radiation and thus to protect the steel reactor vessel from radiation damage. It is often made of steel and is cooled by the primary coolant. The biological shield is added outside to reduce the external dose rate to a tolerable level (**Fig. 5**). Richard E. Faw

Bibliography. Center for Occupational Research and Development Staff, *Course 19: Radiation Shielding*, 1984; A. B. Chilton, K. Shultis, and R. E. Faw, *Principles of Radiation Shielding*, 1984; R. G. Jaeger et al. (eds.), *Engineering Compendium on Radiation Shielding*, vol. 1, 1968, vol. 2, 1975, vol. 3, 1970; International Atomic Energy Agency, *Reactor Shielding*, 1981; M. F. Kaplan, *Concrete Radiation Shielding*, 1989; A. E. Profio, *Radiation Shielding and Dosimetry*, 1979; J. Wood, *Computational Methods in Reactor Shielding*, 1982.

Radiation therapy

The use of ionizing radiation to treat disease; the method is also known as radiotherapy and therapeutic radiology. Radiotherapy was widely used in the past to treat diseases of the skin, lymph nodes, and other organs. However, because radiation can cause cancer and because alternative treatments for these diseases have been discovered, radiation therapy is now mainly limited to treating malignant tumors: the medical specialty is called radiation oncology. *See* ONCOLOGY; RADIATION INJURY (BIOLOGY).

Radiobiologic basis. The exact mechanisms by which radiation kills cells remain uncertain. Most likely, electrons dislodged from water or biological molecules disrupt the bonds between atoms of the nuclear deoxyribonucleic acid (DNA), resulting in double-strand breaks. Although usually not immediately fatal, such damage may cause the death of cells when they attempt to divide. Complex enzymatic mechanisms can repair some of the damage if given sufficient time.

The fraction of cells surviving after irradiation depends on many factors. Radiation affects both normal and cancerous cells in a similar manner qualitatively, but different cell lines vary greatly in their quantitative sensitivity. Rapidly dividing tissues—such as the skin, bone marrow, and gastrointestinal mucosa—usually display the greatest sensitivity experimentally and the most immediate side effects clinically. The cellular environment is important; for example, hypoxic cells that receive less-than-normal amounts of oxygen from the circulatory system are often found in large tumors, and are less sensitive to x-rays than well-oxygenated cells. Particles that produce a dense path of ionizations (such as pions) are more effective at killing cells than is sparsely ionizing radiation (such as photons). Treatment parameters such as the energy of the radiation, the dose given in a single treatment (fraction size), the speed with which it is delivered (dose rate), the interval between treatments (fractionation), the number of days of the treatment course, and the total dose critically influence outcome.

Clinical aspects. The goal of radiation therapy is either to cure the disease permanently (radical treatment) or to reduce or eliminate symptoms (palliative treatment) by destroying tumors without causing unacceptable injuries to normal tissues. For many cancers—such as cancers of the reproductive organs, lymphomas, and small head and neck tumors—a cure usually is possible, and the chance of functionally significant complications is small.

Some tumors, however, contain too many cells to be entirely destroyed by tolerable doses. Also, some neoplasms, such as sarcomas and glioblastomas, are relatively resistant to irradiation and are difficult to eradicate even when only small numbers of cells are present. Such situations are best handled by using surgery to remove all visible tumor. Microscopic deposits of tumor cells often remain in and near the surgical field in tissues that cannot be removed without risk of further injury. Radiotherapy reduces the chance of a local recurrence following surgery. This approach is used successfully in managing the early stages of cancer of the breast (following limited, breast-preserving surgery, or lumpectomy) and in treating tumors in many other body sites, including the brain, head and neck, lung, gastrointestinal tract, and gynecological organs. Radiotherapy is sometimes used before surgery (rather than after surgery) to increase the ease of surgical removal or to lessen the chance of tumor cells being spread by surgical manipulations.

Also, radiation often is used to treat tumors that cannot be removed because of extensive invasion into surrounding structures, that are located in sites where surgery would cause unacceptable complications, or that occur in patients unable to tolerate the stresses of surgery. Except for small or sensitive tumors, success rates with conventional radiotherapy usually have been moderate or poor. Chemotherapy given in combination with radiation (sequentially or simultaneously) may reduce tumor bulk and sensitize cells to irradiation, as well as destroy tumor cells outside the radiation fields. Total body irradiation is used in some diseases (mainly leukemias and lymphomas) to destroy both tumor cells and the bone marrow in preparation for bone marrow transplantation. *See* CHEMOTHERAPY.

Technical aspects. The ionizing radiations most commonly used clinically are photons in the x-ray and gamma-ray energy range (10,000 eV to 45 MeV) and electrons (3 to 20 MeV) produced either by linear or circular particle accelerators or by radioactive isotopes (predominantly radium-226, cobalt-60, iridium-192, cesium-137, and iodine-125). Beams of particles or photons may be directed at the patient from a distance (teletherapy or external-beam therapy). In general, higher energy results in deeper tissue penetration and relatively lower doses to superficial structures such as the skin. Radioisotopes may be placed on the body surface (mold techniques), inside body cavities (intracavitary application), or temporarily or permanently implanted into tissues (interstitial therapy). Some radioisotopes, such as phosphorus-32, gold-198, and iodine-131, have also been administered orally, by intravenous injection, or by instillation into body cavities. Dose is defined as the energy absorbed per unit tissue mass; its unit is the gray (1 Gy = 1 joule/kg = 100 rad). *See* RADIOISOTOPE.

The radiotherapist locates tumor masses by using physical examination, imaging studies, and surgical findings. Knowledge of the patterns of spread and natural history of the disease is used to determine which apparently uninvolved areas might contain tumor cells. These areas are usually irradiated for at least part of the overall course of treatment. Careful planning of the directions of entry and sizes of the radiation beams (fields), the use of multiple converging fixed or rotating fields, and the use of intracavitary and interstitial techniques (brachytherapy) ordinarily permit delivery of much higher doses to the tumor than to surrounding organs. Manipulating the parameters of treatment, particularly dose rate, fraction size, and fractionation, can increase the chance of tumor eradication while lowering the likelihood of complications. The total dose prescribed depends on the exact type of tumor and clinical situation, but it is usually 20–80 Gy delivered over 1–8 weeks.

Technique refinements. Many improvements to existing radiotherapy techniques and practices have been explored. Computed tomography and magnetic resonance imaging, which visualize tumors and organs in three dimensions with great precision, are increasingly used in combination with sophisticated treatment planning and dosimetry programs. Computer control of accelerator motion and dose output, intraoperative radiotherapy, and elaborate localization and immobilization (stereotactic) techniques help to concentrate radiation on tumors and spare normal tissues. Certain drugs (radiosensitizers) and increasing tumor temperatures (hyperthermia) may increase tumor-cell killing while adding few or no complications. Radio-protective drugs for normal tissues are also under investigation. Neutrons, protons, pions, and charged atomic nuclei such as helium,

which interact with matter differently than photons, may be useful in certain clinical situations. *See* CANCER (MEDICINE); RADIATION BIOLOGY.

Abram Recht

Bibliography. G. C. Bentel et al., *Treatment, Planning, and Dose Calculation in Radiation Oncology,* 4th ed., 1989; B. Blackburn (ed.), *Blackburn's Introduction to Clinical Radiation Therapy Physics,* 1989; V. T. DeVita, S. Hellman, and S. A. Rosenberg (eds.), *Cancer: Principles and Practices of Oncology,* 6th ed., 2001; J. M. Vaeth and J. L. Meyer (eds.), *Radiation Tolerance of Normal Tissues,* Frontiers of Radiation and Oncology Series, vol. 23, 1990.

Radiative transfer

The study of the propagation of energy by radiative processes; it is also called radiation transport. Radiation is one of the three mechanisms by which energy moves from one place to another, the other two being conduction and convection. *See* ELECTROMAGNETIC RADIATION; HEAT TRANSFER.

The kinds of problems requiring an understanding of radiative transfer can be characterized by looking at meteorology, astronomy, and nuclear reactor design. In meteorology, the energy budget of the atmosphere is determined in large part by energy gained and lost by radiation. In astronomy, almost all that is known about the abundance of elements in space and the structure of stars comes from modeling radiative transfer processes. Since neutrons moving in a reactor obey the same laws as radiation being scattered by atmospheric particles, radiative transfer plays an important part in nuclear reactor design.

Each of these three fields—meteorology, astronomy, and nuclear engineering—concentrates on a different aspect of radiative transfer. In meteorology, situations are studied in which scattering dominates the interaction between radiation and matter; in astronomy, there is more interest in the ways in which radiation and the distribution of electrons in atoms affect each other; and in nuclear engineering, problems relate to complicated, three-dimensional geometry.

Physical processes. Radiative transfer is a complicated process because matter interacts with the radiation. This interaction occurs when the photons that make up radiation exchange energy with matter. These processes can be understood by considering the transfer of visible light through a gas made up of atoms. Similar processes occur when radiation interacts with solid dust particles or when it is transmitted through solids or liquids. *See* PHOTON.

If a gas is hot, collisions between atoms can convert the kinetic energy of motion to potential energy by raising atoms to an excited state. Emission is the process which releases this energy in the form of photons and cools the gas by converting the kinetic energy of atoms to energy in the form of radiation. The reverse process, absorption, occurs when a photon raises an atom to an excited state, and the energy is converted to kinetic energy in a collision with another atom. Absorption heats the gas by converting energy from radiation to kinetic energy. Occasionally an atom will absorb a photon and reemit another photon of the same energy in a random direction. If the photon is reradiated before the atom undergoes a collision, the photon is said to be scattered. Scattering has no net effect on the temperature of the gas. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; ATOMIC STRUCTURE AND SPECTRA; SCATTERING OF ELECTROMAGNETIC RADIATION.

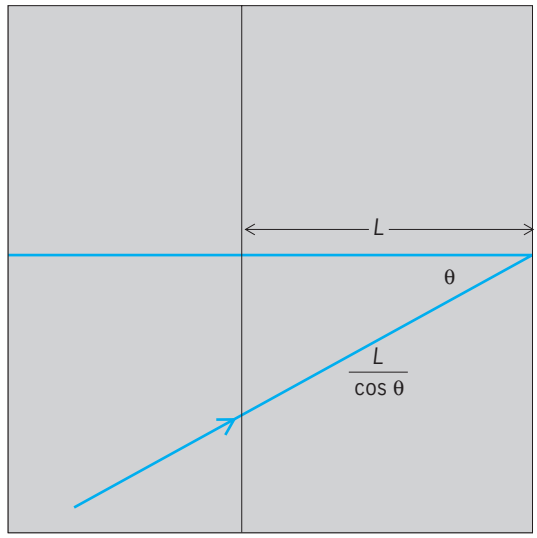

Fig. 1. Beam of light illuminating a slab of gas of thickness *L*, shown in cross section.

Equation of radiative transfer. The equation of radiative transfer is built up by considering the three processes—absorption, emission, and scattering. First, however, it is necessary to define some geometry and, in particular, to consider an infinitely wide slab of gas (**Fig. 1**). Radiation enters the slab at some angle θ. If the slab has a thickness L, the radiation must cross a distance L divided by the cosine of the angle θ; that is, the distance traveled is $L/\cos\theta$. It is common to write the cosine of the incidence angle as μ; that is, $\mu = \cos\theta$. In this notation, the distance crossed is L/μ.

Absorption. First, one may consider a gas that absorbs radiation but does not scatter or emit it. This model is an idealization since such a gas is gaining energy which it must eventually radiate away. However, it is a reasonably good approximation of a gas that emits radiation at a much lower energy than the incident radiation.

Figure 2 shows a slab of gas divided into 5 layers; each layer is assumed to absorb half the radiation impinging on it. If the gas is illuminated from the left with a beam of light of some intensity, for example, I_0, the solid dots show how much radiation reaches each layer. If there are 10 layers each absorbing 25% of the incident energy, then the energy falls off as shown by the circles.

The curve shows the intensity at each point if this process is continued to the limit of extremely

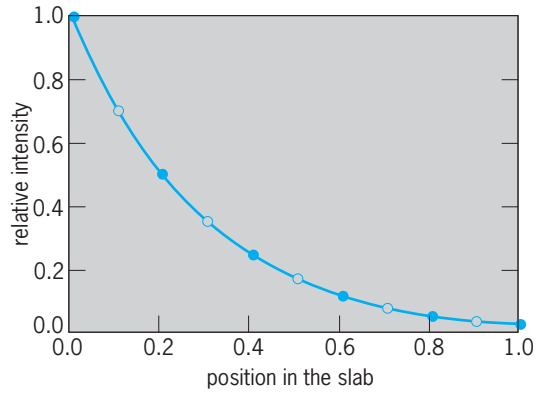

Fig. 2. Absorption of radiation in a slab of gas. The solid circles show the amount of radiation reaching each interface if each layer absorbs 50% of the radiation reaching it. The open circles are for 10 layers, each absorbing 25% of the radiation. The curve shows the result in the limit of a continuous medium.

thin layers absorbing a constant fraction of the incident energy per unit length.

The absorption illustrated in Fig. 2 can be expressed as Eq. (1) where I_n is the intensity leaving

$$I_{n+1} - I_n = -I_n \kappa \frac{\Delta x}{\mu} \tag{1}$$

layer n, $\Delta x/\mu$ is the distance the radiation travels in crossing the layer, and κ is the absorption coefficient or opacity, a quantity that measures the fraction of the energy absorbed per unit length. On the average, a photon will travel a distance $1/\kappa$ before being absorbed, and $1/\kappa$ is called the mean free path. The quantity I is called the specific intensity and measures the amount of energy crossing a unit area per unit time per unit solid angle in a given direction. Typical units are joules per square centimeter per second per steradian.

Taking the limit of Eq. (1) as the layers get thinner yields the differential equation (2). If the opac-

$$\mu \frac{dI}{dx} = -\kappa I \tag{2}$$

ity κ does not depend on position in the slab, the gas is said to be homogeneous, and Eq. (2) has the solution given by Eq. (3). In other words, the inten-

$$I = I_0 e^{-\kappa x/\mu} \tag{3}$$

sity decreases exponentially with distance from the boundary of the slab; the opacity determines how rapidly the intensity falls off.

Since the solution depends only on the product of the opacity and distance, the optical depth of a point a distance x into the gas is defined as $\tau(x) = \kappa x$. The optical thickness of a layer [$\tau(L)$ in the example of Fig. 1] is a measure of the absorbing power of the gas. A thin layer of gas with a large absorption coefficient can absorb as much energy as a thick layer of gas with a small absorption coefficient.

Even if the opacity depends on position, Eq. (2) can be solved by combining κ and dx and calling the product $d\tau$. Equation (2) becomes Eq. (4), which has the solution given by Eq. (5). In order to compute

$$\mu \frac{dI}{d\tau} = -I \tag{4}$$

$$I = I_0 e^{-\tau/\mu} \tag{5}$$

the optical depth from the opacity and the physical depth, it is necessary to add the contributions of each thin piece of the gas; that is, integration is carried out over distance to obtain Eq. (6).

$$\tau(x) = \int_0^x \kappa(z)\, dz \tag{6}$$

All the details of the absorption process are contained in the absorption coefficient κ. For example, the opacity normally depends on wavelength, which means that both the optical depth and specific intensity depend on wavelength.

Emission. It follows from Kirchhoff's law of radiation that if matter absorbs energy it must either radiate away the energy or heat up. In many radiative transfer problems the emission of energy can be ignored because it is in a different wavelength domain than the one under study. For example, in the study of the transfer of visible radiation through the Earth's atmosphere, the emitted radiation, which is primarily in the infrared part of the spectrum, is usually of no concern. However, there are many situations in which both the absorption and emission of radiation must be considered.

In the simplest approximation, the hot gas radiates like a blackbody. The emission is isotropic and has an intensity, denoted $B_\lambda(T)$, given by the Planck radiation law. The subscript λ indicates that the emission is wavelength-dependent, while the argument T indicates the temperature dependence of the emission. The equation of radiative transfer now becomes Eq. (7). In the following discussion, the subscript λ

$$\mu \frac{dI_\lambda}{d\tau_\lambda} = -I_\lambda + B_\lambda(T) \tag{7}$$

will be omitted, although most quantities depend on wavelength. *See* HEAT RADIATION.

If the temperature through the gas is constant, Eq. (7) has the solution given by Eq. (8). This equa-

$$I = I_0 e^{-\tau/\mu} + B(T)(1 - e^{-\tau/\mu}) \tag{8}$$

tion is used in astronomy to determine the properties of interstellar clouds, and in the laboratory to find the atomic properties of gases.

If the temperature is not constant, the situation is more complicated. The properties of the gas can still be analyzed, however, if Eq. (8) is replaced by Eq. (9). The equation indicates that the radiation at

$$I = I_0 e^{-\tau/\mu} + \int_0^\tau B(T) e^{-t/\mu} \frac{dt}{\mu} \tag{9}$$

a point in the slab is made up of two parts. The

first part is the incoming radiation that has not yet been absorbed ($I_0 e^{-\tau/\mu}$). The second part is the radiation emitted by the hot gas. Since radiation emitted far away will be attenuated more than that emitted nearby, Eq. (9) contains a term $e^{-\tau/\mu}$. In addition, the radiation reaching a point comes from all over the gas so it is necessary to add up the contributions, that is, to integrate over all optical depths.

Scattering. When radiation interacts with matter, there is a probability that it will simply change direction. If there is no exchange of energy between matter and radiation, the scattering is called conservative.

In its simplest form, scattering is isotropic; that is, the direction of the radiation after the scattering is independent of its incident direction. If the probability that the radiation is scattered is denoted by ω, then the probability that it is absorbed is $1 - \omega$. The equation of radiative transfer becomes Eq. (10), where the source function S is defined by Eq. (11), and $d\Omega$ denotes that the integration is over all directions.

$$\mu \frac{dI}{d\tau} = -I + S \tag{10}$$

$$S = (1 - \omega)B(T) + \omega \int I \, d\Omega \tag{11}$$

The source function S contains two contributions. The first is the thermal emission from the hot gas, $(1 - \omega)B(T)$. The second term is needed because the specific intensity I is a measure of energy traveling in a specific direction. The integral in Eq. (11) counts the contribution of radiation that was traveling in any direction before it was scattered but is now traveling in the particular direction being considered. Since the source function depends on the intensity, there is no simple analytic solution to Eq. (10).

Nature of solutions. The solutions to the transfer equation are quite different, depending on what processes are significant. If the slab is very thick, the amount of radiation at a given layer decreases exponentially with depth for pure absorption, approaches $B_\lambda(T)$ if both absorption and emission are considered, and decreases linearly with depth if only scattering is important.

Only the simplest form of the equation of radiative transfer has been presented. Since the problems are so formidable, each discipline has concentrated on a particular set of approximations and there is relatively little crossover from one field to another.

Scattering dominated problems. As noted above, the transfer of visible radiation through the Earth's atmosphere is dominated by scattering. In other words, the single-scatter albedo ω is near one; that is, the scattering often is nearly conservative. If the radiation left a particle in a random direction on each scattering, that is, if isotropic scattering prevailed, then the solution of Eq. (10) would be tedious but not difficult. In fact, there is an analytic solution in this case if the gas is an infinitely wide slab.

The probability that a photon will be scattered into a given direction is described by the scattering phase function $P(\Theta)$, where Θ is the angle between the incident and outgoing radiation. The phase function is determined by the size of the particles relative to the wavelength of light being scattered and by the composition of the particles. Metallic particles have very different properties from the nonconducting particles that are found in the atmosphere.

If the scattering is not isotropic, Eq. (11) is written as Eq. (12). The factor $P(\Theta)$ under the integral

$$S = (1 - \omega)B(T) + \omega \int P(\Theta) I \, d\Omega \tag{12}$$

accounts for the fact that photons traveling in certain directions have a greater probability of being scattered into the direction of interest than do others.

If the scattering particles are very small compared to the wavelength of the radiation, the scattering is described by the Rayleigh phase function, Eq. (13),

$$P(\Theta) = {}^3/_4(1 + \cos^2 \Theta) \tag{13}$$

shown in **Fig. 3**. Equal amounts of energy are scattered into the forward ($\Theta = 0^\circ$) and backward ($\Theta = 180^\circ$) directions, and less energy is scattered perpendicular to the incident direction ($\Theta = 90^\circ$).

When the scattering particles are approximately the same size as the wavelength of the light being scattered, the phase function is more complicated. Figure 3 also shows a phase function typical of rain clouds. The value of the phase function at $\Theta = 0^\circ$ is nearly 25, well off the scale used for the figure. In other words, the scattering changes the direction of travel of only a small amount of radiation; most of it is scattered straight ahead. Since absorption in clouds is quite low, even very thick clouds can transmit an appreciable amount of light. *See* METEOROLOGICAL OPTICS.

Radiation and atomic physics. A completely different set of problems arises if the radiation changes the absorption and scattering properties of the gas. The effect is most apparent when atomic or molecular transitions in diffuse gases are considered.

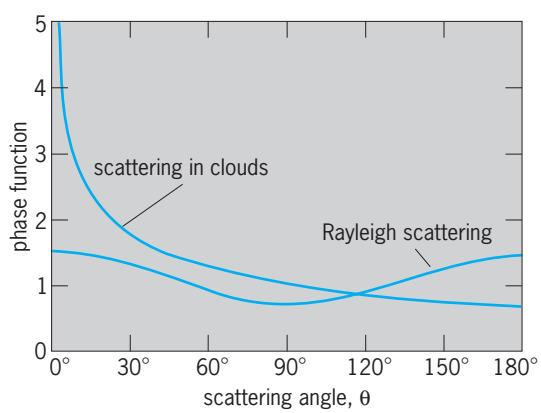

Fig. 3. Representative scattering phase functions, representing the probability that a photon will be scattered through an angle Θ.

If the gas is dense, collisions between atoms are frequent, and the distribution of electron energies depends only on the temperature. Such a gas is said to be in thermodynamic equilibrium. If the conditions of thermodynamic equilibrium are approximately satisfied in some region, the gas is said to be in local thermodynamic equilibrium, and the effect of radiation on the energy distribution of the electrons can be ignored. *See* KINETIC THEORY OF MATTER.

If the gas density is low enough, the distribution of electron energies depends on the radiation as well as the temperature. Since the absorption and scattering properties of the gas depend on the electron distribution, the radiation and electron distribution equations are coupled. Such a gas is said to be in a state of nonlocal thermodynamic equilibrium.

To further complicate matters, atoms can be perturbed while in an excited state. Therefore, electrons in different atoms jumping between the same pair of energy levels emit photons at slightly different wavelengths. The spectral line shape must then be taken into account, and the radiative transfer equation must be solved for a number of frequencies simultaneously. In the simplest situation, Eq. (10) becomes Eq. (14), where ϕ_λ, the spectral line shape, is

$$\mu \frac{dI_\lambda}{d\tau_\lambda} = -I_\lambda + \int_0^\infty \phi_\lambda S_\lambda \, d\lambda \qquad (14)$$

the probability that the electron transition results in a photon of wavelength λ. *See* LINEWIDTH; PERTURBATION (QUANTUM MECHANICS).

Complicated geometries. If the scattering phase function is not simple, or if the radiation and gas properties are coupled, it is impractical to try to solve the radiative transfer equation in geometrically complicated domains. However, in nuclear reactors it can normally be assumed that the scattering is either isotropic or nearly so, and that the absorption and scattering properties of the medium are independent of the neutron distribution. These assumptions make it possible to model reactors with hundreds or thousands of elements. The equation of radiative transfer can be written in the quite general form of Eq. (15), where c is the neutron velocity, $\hat{n}$ is a unit

$$\frac{1}{c}\frac{dI}{dt} + \hat{n} \cdot \nabla I = I - S \qquad (15)$$

vector, and ∇ denotes the gradient operator. Since reactor designers are interested in how the neutron flux changes with time, they include the time derivative as done in Eq. (15). *See* CALCULUS OF VECTORS.

One reason nuclear reactors generate energy is that each collision between an atom and a neutron can result in the emission of additional neutrons. In terms of the radiative transfer equation, this means that the single-scatter albedo ω is greater than unity. A part of the reactor is said to be critical when $\omega \geq 1$. Other regions can be subcritical, that is, $\omega < 1$. The goal of reactor design is to balance the critical and subcritical domains to meet the desired energy production. *See* REACTOR PHYSICS.

Difficulties. The modeling of radiative transfer is difficult. Problems with complex geometries that have complicated scattering phase functions and in which the absorption and scattering properties depend on the radiation are too hard to solve even with the largest computers available.

One simple problem that has not yet been solved accurately is the calculation of the radiation emerging from a single cloud in an otherwise clear sky. Nor has an accurate solution been presented for the transfer of radiation through a clear sky that includes the effects of horizontal variations in the reflectivity of the ground.

Even if such problems could be solved, more difficult problems remain. For example, radiation can be polarized, that is, the waves making up the light are lined up. The intensity of polarized radiation has four distinct components, so that Eq. (10) must be replaced by four equations. Furthermore, calculating the scattering phase function is considerably more difficult than if polarization is ignored. Only in the 1980s was substantial progress made on this problem. *See* POLARIZED LIGHT.

Alan H. Karp

Bibliography. M. Q. Brewster, *Thermal Radiative Transfer and Properties*, 1992; S. Chandrasekhar, *Radiative Transfer*, 1960; J. Lenoble, *Atmospheric Radiative Transfer*, 1993; D. Mihalas, *Stellar Atmospheres*, 2d ed., 1978; M. F. Modest, *Radiative Heat Transfer*, 1993; H. C. van de Hulst, *Light Scattering by Small Particles*, 1957, reprint 1982; H. C. van de Hulst, *Multiple Light Scattering*, 2 vols., 1980.

Radiator

Any of numerous devices, units, or surfaces that emit heat, mainly by radiation, to objects in the space in which they are installed. Because their heating is usually radiant, radiators are of necessity exposed to view. They often also heat by conduction to the adjacent thermally circulated air. *See* HEAT RADIATION.

Radiators are usually classified as cast-iron (or steel) or nonferrous. They may be directly fired by wood, coal, charcoal, oil, or gas (such as stoves, ranges, and unit space heaters). The heating medium may be steam, derived from a steam boiler, or hot water, derived from a water heater, circulated through the heat-emitting units.

Cast-iron radiators are made in sections of varying widths and heights and are assembled in the required number by top and bottom nipples. Some are made in a flat panel. They are set on legs or similar supports or affixed to walls by adjustable hangers. The preferred location is under windows.

When windows are low, cast-iron baseboard radiators, less than a foot high, are available for use under the windows. Because of the limited heat from such a radiator, it is frequently extended along the nonwindow baseboard. This radiator is assembled with valve and end sections, extension blocks, and inverted and

projecting corner covers. Cast-iron convectors may be used also.

Finned-tube radiators consist of a pipe or tube with affixed fins of steel, copper, or aluminum. They are available with or without enclosures, and usually are wall-mounted. They are more compact than cast-iron radiators, having greater heating area per volume.

Mass production was responsible for the introduction of convectors, which, like finned-tube radiators, consist of steel, copper, or aluminum tubes with extended fin surfaces affixed to them. They have almost completely displaced cast-iron radiators, for the fact that they are lightweight causes freight costs to be lower. They also require less labor to install. They cannot, however, be used as replacements in an existing hot-water heating system with cast-iron radiators for heating surfaces, because the heat-dissipating characteristics of convectors are completely different from those of cast-iron radiators with changes in hot-water temperatures.

Convectors are customarily made of nonferrous finned tubes in a wide variety of enclosures for free-standing, recessed, or wall-hung installation. Air is circulated over the elements by natural convection or by fans, in which case they are known as unit ventilators. Convectors and unit ventilators are not true radiators because they emit most of the heat by conduction to the circulated air.

A limited amount of cooling may be produced by passing chilled water through convectors or radiators, but care must be exercised to dispose of moisture condensation in humid weather.

Electric heating elements may be substituted for fluid heating elements in all types of radiators, convectors, and unit ventilators. *See* COMFORT HEATING; ELECTRIC HEATING; HOT-WATER HEATING SYSTEM; RADIANT HEATING; STEAM HEATING.

Erwin L. Weber; Richard Koral

Bibliography. American Society of Heating, Refrigerating and Air Conditioning Engineers, *ASHRAE Handbook*: *Equipment*, 1988.

Radio

Communication between two or more points, employing electromagnetic waves ("carriers") as the transmission medium. *See* ELECTROMAGNETIC RADIATION.

At the start of the twenty-first century, radio transmission services are in the midst of a fundamental transition from traditional (and increasingly antiquated) analog technology to highly sophisticated and advanced digital technology. The basic methods of analog transmission and reception of audio signals were developed principally in the twentieth century (**Fig. 1**). For transmission, an audio pulse and a carrier pulse are blended into a modulated carrier wave, which is then amplified and fed into the antenna. The process of blending audio and carrier waves is called modulation. For reception, the receiving antenna and tuner catch the weak signal, amplify it, sort the audio pulse from the carrier, and play a now reamplified audio pulse through the speaker at home. Digital radio techniques differ from analog in the nature of the information being transmitted (audio represented digitally, as a sequence of 1's and 0's) and the methods of modulation employed. Digital techniques are significantly more flexible, robust (that is, less susceptible to interference and other deleterious effects), and more efficient than the analog techniques they replace. *See* ANTENNA (ELECTROMAGNETISM); ELECTRICAL COMMUNICATIONS; MODULATION; RADIO RECEIVER; RADIO TRANSMITTER; RADIO-WAVE PROPAGATION.

Methods of information transmission. Radio waves transmitted continuously, with each cycle an exact duplicate of all others, indicate that only a carrier is present (also called an unmodulated carrier). The two principal characteristics of a carrier are its frequency (the number of cycles per second) and amplitude (which determines how much power is contained in the signal). The information desired to be transmitted must cause changes in the carrier which can be detected at a distant receiver. The method used for the transmission of the information is determined by the nature of the information which is to be transmitted as well as by the purpose of the communication system (**Fig. 2**).

Code transmission. This method, the first ever used, dates from the dawn of radio, and was patented by Gugliemo Marconi in 1896. The carrier is keyed on and off to form dots and dashes (Morse code). The technique, traditionally used in ship-to-shore and amateur communications, has been almost entirely superseded by more efficient methods. This primitive method is actually digital (binary) in nature, since only one of two possible symbols is transmitted at any given moment. *See* TELEGRAPHY.

Frequency-shift transmission. The frequency of the carrier is shifted a fixed amount to correspond with telegraphic dots and dashes or with combinations of pulse signals identified with the characters on a typewriter. This technique was widely used in handling the large volume of public message traffic on long circuits, principally by the use of teletypewriters. *See* TELETYPEWRITER.

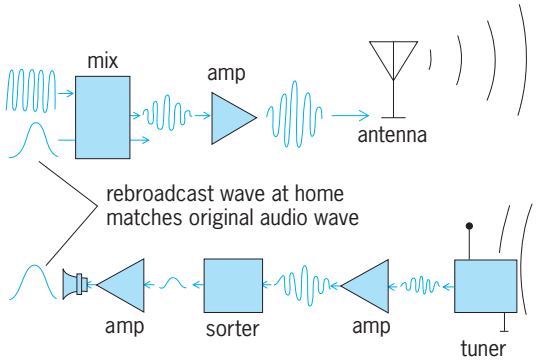

Fig. 1. Transmission of audio information by radio.

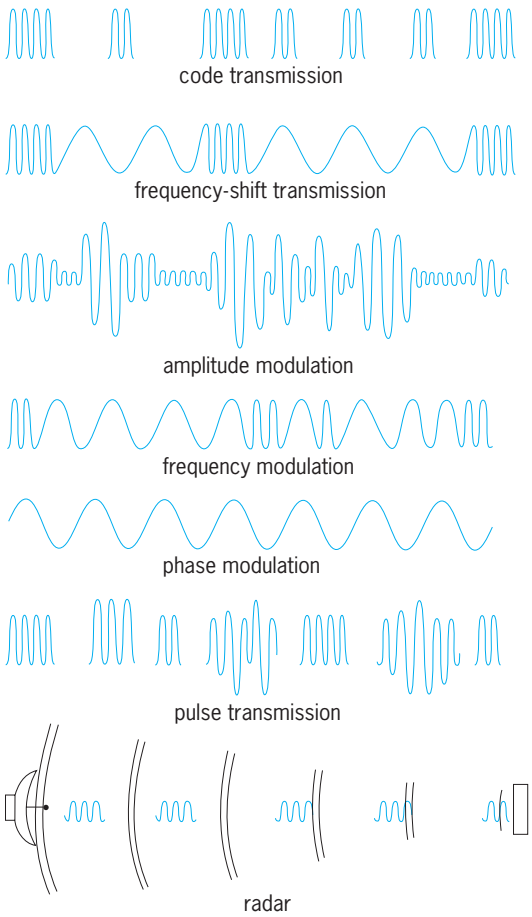

Fig. 2. Methods for transmission of information.

Amplitude modulation. The amplitude of the carrier is varied in a way that conforms to the fluctuations of a sound wave, or, in the case of digital transmission, the specific sequence of 1's and 0's being transmitted. This technique is used in amplitude-modulation (AM) terrestrial radio broadcasting, analog television picture transmission, and many other services. One widely used digital method, called quadrature AM (QAM), is employed in some terrestrial digital radio systems. *See* AMPLITUDE MODULATION; AMPLITUDE-MODULATION DETECTOR; AMPLITUDE-MODULATION RADIO; AMPLITUDE MODULATOR; TELEVISION.

Frequency modulation. The frequency of the carrier is varied in a way that conforms to the fluctuations of the modulating wave. This technique is used in frequency-modulation (FM) broadcasting, analog television sound transmission, and analog microwave relaying. *See* FREQUENCY MODULATION; FREQUENCY-MODULATION DETECTOR; FREQUENCY-MODULATION RADIO; FREQUENCY MODULATOR; MICROWAVE.

Phase modulation. The phase of the carrier is varied in a way that conforms to the fluctuations of the modulating wave. A family of popular digital methods, known as phase-shift keying (PSK), are used in digital satellite broadcasting. Specific forms include binary PSK (BPSK), quadrature PSK (QPSK), and octal PSK (8PSK), differentiated by the number of digital bits used to create each transmitted symbol. *See* PHASE MODULATION; PHASE-MODULATION-DETECTOR; PHASE MODULATOR.

Pulse transmission. The carrier is transmitted in short pulses, which change in repetition rate, width, or amplitude, or in complex groups of pulses which vary from group to succeeding group in accordance with the message information. These forms of pulse transmission are identified as pulse-code, pulse-time, pulse-position, pulse-amplitude, pulse-width, or pulse-frequency modulation. These complex techniques are employed principally in microwave relay and telemetry systems. *See* PULSE DEMODULATOR; PULSE MODULATION; PULSE MODULATOR.

Radar. The carrier is normally transmitted as short pulses in a narrow beam, similar to that of a searchlight. When a wave pulse strikes an object, such as an aircraft, energy is reflected to the radar station, which measures the round-trip time and converts it to distance. A radar receiver can display varying reflections in a maplike presentation on a video display. *See* RADAR.

Uses of radio. The first practical application of radio, in the 1900s between ships and shore stations, was followed quickly by overseas communication and communication between other widely separated points. Subsequently, the applications have become widely diversified. Some of the important uses are listed below.

1. Public safety: marine and aviation communications, police and fire protection, forestry conservation, and highway traffic control.

2. Industrial: power utilities, pipelines, relay services, news systems, financial institutions, agriculture, and petroleum processing.

3. Land transportation: railroads, motor carriers, urban public transit and taxicabs, and automobile emergency needs.

4. Broadcasting: television (terrestrial and satellite), terrestrial radio broadcasting (FM, AM, digital), satellite digital audio radio services (SDARS), and shortwave broadcasting. *See* RADIO BROADCASTING; RADIO BROADCASTING NETWORK; SATELLITE RADIO; SATELLITE TELEVISION BROADCASTING.

5. Military and space: radar, communications, navigation, telemetering, missile tracking, satellite photographic surveys, and missile guidance. *See* ELECTRONIC NAVIGATION SYSTEMS; REMOTE SENSING; SPACE COMMUNICATIONS.

6. Fixed point-to-point: long-distance message, data, and picture transmission.

7. Relaying: television, sound, data, picture, and public message relaying over long distances. Communications satellites handle both domestic and overseas communications. These satellites supply not one but many services, including thousands of narrow channels for voice communications with wider channels for television and data transmission. *See* COMMUNICATIONS SATELLITE.

8. Telemetering: remote indication of water levels in reservoirs and rivers, performance of

experimental aircraft, missiles, and satellites. *See* TELEMETERING.

9. Weather reporting: early warning and location of hurricanes and other storms, trends of weather for industrial and public information. *See* WEATHER FORECASTING AND PREDICTION.

10. Interpersonal: cellular telephones, personal communications services, amateur radio services, radio paging systems, interconnection (for example, Bluetooth®, Wi-Fi®, WiMax). *See* AMATEUR RADIO; MOBILE RADIO; RADIO PAGING SYSTEMS.

Frequency separation. Hundreds of thousands of radio transmitters exist, each requiring a carrier at some radio frequency. To prevent interference, different carrier frequencies are used for stations whose service areas overlap, and receivers are built which can select the carrier signal of the desired station. Resonant electronic circuits in the receiver are adjusted, or tuned, to accept the desired frequency and reject others. Digital technology, in particular the availability of high-speed, low-cost microprocessors and analog-to-digital converters, has made possible the development of software radios which select and demodulate radio signals according to a computer program which runs on a microprocessor. Software radios are more flexible and offer higher performance than do traditional radios.

Each station operates within a specific radio-frequency (RF) channel. All other stations within a geographical area are excluded from using this channel. However, an advanced digital technology called code-division multiple access (CDMA) allows multiple stations to operate on the same frequency without interfering with one another. CDMA technology has found its principal application in some cellular telephone systems. Another digital technology, time-division multiple access (TDMA), allows stations to share a frequency by precisely coordinating the time at which each station transmits, making sure that only a single station is transmitting at any one time. A new technology called ultra-wideband (UWB) spreads the information over many thousands of megahertz, allowing it to operate at such a low power level per megahertz that it does not interfere significantly with other overlapped services. *See* MULTIPLEXING AND MULTIPLE ACCESS; ULTRAWIDEBAND (UWB) SYSTEMS.

Each channel must be wide enough to accommodate the message information, provide tolerance for small carrier frequency drift, perhaps provide a guard band, and allow for imperfect receiver selectivity capabilities. The minimum usable channel widths (or bandwidths) vary from service to service, depending upon the amount of information a channel must accommodate. In analog television it is 6000 kHz because of the large amount of essential picture information. In FM radio broadcasting it is 200 kHz and in AM radio broadcasting 10 kHz. Digital systems can accommodate corresponding services with smaller bandwidths. For example, digital television services can accommodate multiple video programs in the same 6000-kHz bandwidth required for a single video program using analog services. The great demand for new frequency authorizations requires efficient channel utilization and has been a major factor in the drive to develop and deploy digital services. *See* BANDWIDTH REQUIREMENTS (COMMUNICATIONS); RADIO SPECTRUM ALLOCATIONS; TELEVISION.

Regulation of spectrum usage. All nations have a sovereign right to use freely any or all parts of the radio spectrum. But numerous international agreements and treaties divide the spectrum and specify sharing among nations for their mutual benefit and protection.

Each nation designates its own regulatory agency. Functioning within the international agreements, it issues authorizations; assigns frequencies; polices operations; creates technical standards, rules, and practices; and safeguards and protects the public interest.

In the United States all nongovernmental radio communications are regulated by the Federal Communications Commission (FCC), according to the provision of the Communications Act of 1934, as amended. Creation of a radio station or service requires authorization by the FCC. Upon construction of an authorized facility, a license to operate is issued. Radio stations may be inspected by engineers attached to FCC field offices. Stations must comply with the terms of their authorization regarding carrier-frequency tolerance, power limitations, permissible communications, calls signals, and control by properly licensed personnel.

Governmental use of spectrum in the United States, which is extensive, is regulated by the National Telecommunications and Information Administration (NTIA), an agency which also serves as the principal advisor to the President on telecommunications and information policy issues. In managing spectrum the NTIA's stated goals include serving the national security and defense, supporting crime prevention and law enforcement, and promoting scientific research, development, and exploration.

The International Telecommunication Union (ITU) is a specialized agency of the United Nations with headquarters in Geneva, Switzerland. It is responsible for the international agreements necessary to assure provision of radio and telecommunications services among the nations of the world.

The ITU is divided into three sectors. The radio communication sector (ITU-R) is responsible for the allocation of radio spectrum, allocation of orbital slots used by international satellites, and establishment of various technical characteristics of operating radio systems. The telecommunication standardization sector (ITU-T) is responsible for developing internationally agreed technical and operating standards and defining tariff and accounting principles for international telecommunications services. The telecommunication development sector (ITU-D) promotes investment and fostering the expansion of telecommunication infrastructure in developing nations throughout the world. Additional ITU functions include the maintenance of an

international registry of radio transmitters, and the availability to facilitate the resolution of international disputes on uses of the radio spectrum among member countries.

David H. Layer; John D. Singleton; Michael C. Rau

Bibliography. American Radio Relay League, *The ARRL Handbook for Radio Amateurs*, rev. ed., annually; G. Jones, *A Broadcast Engineering Tutorial for Non-Engineers*, NAB and Focal Press, 2005; J. G. Proakis and M. Salehi, *Communication Systems Engineering*, 2d ed., Prentice Hall, 2001; T. S. Rappaport, *Wireless Communications: Principle and Practice*, 2d ed., Prentice Hall PTR, 2001; C. H. Sterling and J. M. Kittross, *Stay Tuned—A History of American Broadcasting*, 3d ed., Lawrence Erlbaum Associates, 2002; M. E. Van Valkenburg and W. M. Middleton (eds.), *Reference Data for Engineers: Radio, Electronics, Computer, and Communications*, Newnes, 9th ed., 2002; J. Whitaker (ed.), *NAB Engineering Handbook*, 9th ed., National Association of Broadcasters, 1999.

Radio astronomy

The study of celestial bodies by examination of the energy they emit at radio frequencies. Celestial radio noise originates by one of several processes. These include both broadband continuum radiation owing to (1) thermal radiation from solid bodies such as the planets, (2) thermal or bremsstrahlung radiation from hot gas in the interstellar medium, (3) synchrotron radiation from ultrarelativistic electrons moving in weak magnetic fields, (4) coherent processes as found in the Sun and on Jupiter, and (5) pulsed radiation from the rapidly rotating neutron stars, as well as (6) narrow "spectral line" radiation from atomic or molecular transitions that occur in the interstellar medium or in the gaseous envelopes around stars.

Radio astronomy observations cover the entire radio spectrum from less than 1 mm, where the cosmic signals become heavily attenuated by the atmosphere, to wavelengths of tens of meters, beyond which the incoming signals are attenuated by the ionosphere. However, radio waves between about 1-cm and 1-m wavelength not only get through the Earth's atmosphere and ionosphere with little distortion, they also penetrate much of the gas and dust in space as well as the clouds of planetary atmospheres. Radio astronomy can therefore sometimes give a much clearer picture of stars and galaxies than is possible by means of conventional observation using visible light. Sophisticated antennas equipped with very sensitive radio receivers that are used to detect cosmic radio emission are referred to as radio telescopes. *See* RADIO TELESCOPE.

Spectral line emission occurs at specific wavelengths characteristic of the atomic and molecular species. However, due to the motion of gas clouds, the wavelength is shifted toward longer wavelengths if the source is moving away from the observer and toward shorter wavelengths if it is moving toward the observer. Precise measurements of this so-called Doppler shift are used to determine the radial velocities due to the random motion of gas clouds, the rotation of the Galaxy, and even the expansion of the universe. *See* DOPPLER EFFECT; REDSHIFT.

Historical background. In 1932, the physicist Karl Jansky, working at the Bell Telephone Laboratories, first detected cosmic radio noise from the center of the Milky Way Galaxy while investigating radio disturbances interfering with transoceanic telephone service. Grote Reber, a young engineer and amateur radio operator, later built the first radio telescope at his home in Wheaton, Illinois, and found that the radio radiation came from all along the plane of the Milky Way and from the Sun. Solar radio emission was also detected during the World War II by military researchers in the United States and the United Kingdom but was not announced until after the conclusion of hostilities due to wartime secrecy.

During the late1940s and 1950s, Australian and British radio scientists were able to use techniques and instrumentation developed for wartime radar to locate a number of discrete sources of celestial radio emission which they associated with old supernovae and active galaxies later known as radio galaxies. The construction of ever larger antenna systems and radio interferometers, improved radio receivers, and data processing methods have allowed radio astronomers to study fainter radio sources with increased angular resolution and improved image quality. Today more than 1 million sources of radio emission are known. The precise measurement of their position in the sky has usually allowed their identification with an optical, infrared, x-ray, or gamma-ray counterpart; however, some cosmic radio sources remain unidentified at other wavelengths even with the most sensitive instrumentation.

Solar system radio astronomy. Solar flares and sunspots are strong sources of radio emission. Their study has led to increased understanding of the complex phenomena near the surface of the Sun and provides advanced warning of dangerous solar flares which can interrupt radio communications on the Earth and endanger sensitive equipment in satellites and the health of astronauts. Even in the absence of activity, the quiet Sun is the strongest radio source in the sky, but only because it is so close to the Earth. *See* SUN.

Radio telescopes are used to measure the surface temperatures of all the planets as well as some of the moons of Jupiter and Saturn. Radio studies of the planets have revealed the existence of a greenhouse effect on Venus which results in a surface temperature of about 600°F (315°C), intense Van Allen belts surrounding Jupiter, powerful radio storms in the Jovian atmosphere, and an internal heating source deep within the interiors of Jupiter, Saturn, Uranus, and Neptune. *See* JUPITER; NEPTUNE; SATURN; URANUS; VAN ALLEN RADIATION; VENUS.

Astronomers also use radar observations to image features on the surface of Venus, which is completely obscured from visual scrutiny by the heavy cloud cover that permanently enshrouds the planet.

Terrestrial-based radar measurements also have revealed the rotation of Mercury, which was previously thought to always keep the same side toward the Sun. Accurate measurements of the travel time of radar signals reflected from Venus near superior conjunction (that is, nearly opposite the Earth on the other side of the Sun) show that radio waves passing close to the Sun slow down owing to gravity and thereby provide a new independent test of Einstein's general theory of relativity. *See* MERCURY; RADAR ASTRONOMY; RELATIVITY; SOLAR SYSTEM.

Radio emission from the Milky Way. Broadband continuum radio emission is observed throughout the radio frequency spectrum from a variety of stars including flare stars, binary stars, x-ray binaries, novae, and supernovae; from supernova remnants; and from magnetic fields and relativistic electrons in the interstellar medium thought to be produced in supernova explosions (for example, **Fig. 1**). Bursts of radio emission are frequently observed from some binary star systems thought to be the result of mass transfer between the two components of the binary system. The remnants of powerful supernovae explosions are often observed as bright sources of radio emission due to synchrotron radiation from relativistic electrons accelerated by supernova events. One of the brightest observed radio sources in the Milky Way is the Crab Nebula, which is the remnant of a supernova observed by Chinese astronomers in the year 1054. Of particular interest to astronomers is the complex radio source at the Galactic Center known as Sagittarius A, which contains both thermal and nonthermal continuum radiation as well as a rich mixture of molecular and ionized hydrogen gas. High-resolution radio observations also show a weak, embedded, very compact radio source called Sgr A*, which is thought to lie at the precise center of the Galaxy (**Fig. 2**). Sgr A* is believed to be associated with a black hole containing the mass equivalent of 2 million suns within a region only 5 solar diameters across. *See* BINARY STAR; CRAB NEBULA; MILKY WAY GALAXY; NOVA; SUPERNOVA; SYNCHROTRON RADIATION; X-RAY ASTRONOMY.

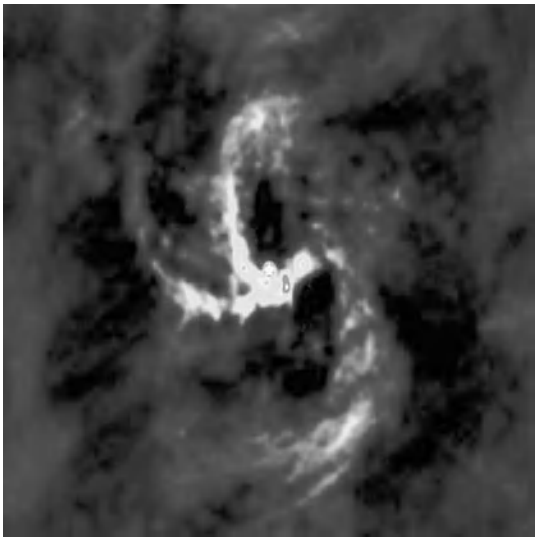

Fig. 2. Radio image of the center of our Milky Way Galaxy known as Sagittarius A. The bright spot is the very compact radio source known as SgrA*, which is associated with the million-mass black hole at the very center of the Galaxy.

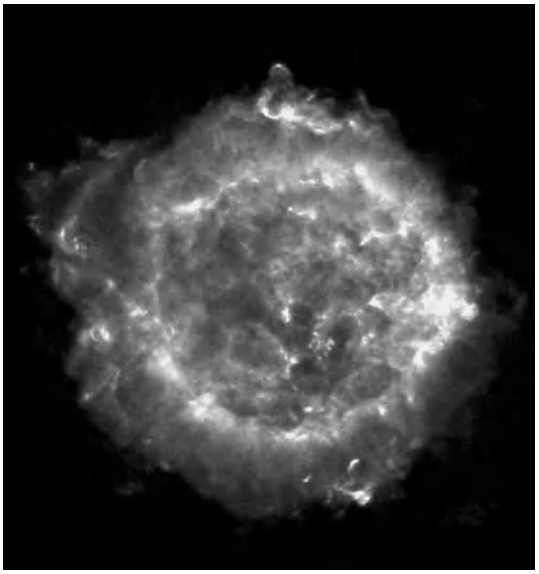

Fig. 1. Radio image, known as Cassiopeia A, of the remnants of a supernova which exploded about 300 years ago. (*NRAO/AUI/NSF*)

Atomic and molecular gas. The important spectral line of neutral atomic hydrogen (H I), which occurs at 1421.405 MHz (21 cm), was first observed in 1951. H I is observed throughout the Milky Way Galaxy as well as in other spiral galaxies, although the galactic hydrogen is heavily concentrated toward the plane of the Milky Way. The radial (along the line of sight) velocity of interstellar H I clouds is determined by measuring the shift in the wavelength of the observed lines arising from the Doppler effect. Measurements of the position and velocity of H I are used to trace out the spiral arms in the Milky Way system to study the rotation of the Galaxy.

H I is also observed in nearby galaxies. The observed dependence of rotation velocity on distance from the center of the galaxy is referred to as a rotation curve and is used to study the distribution of mass in galaxies. Observations of rotation curves in a number of spiral galaxies show the presence of dark matter located far beyond the extent of visible stars and nebula. The nature of this dark matter is unknown. It could be composed of ordinary matter in the form of planetlike objects, brown dwarfs, or neutron stars; or neutrinos; or some unknown form of matter. *See* DARK MATTER.

Molecular gas found in dense interstellar clouds radiates at radio wavelengths characteristic of each molecule. The wavelength of emission corresponds to transitions between different rotational and sometimes vibrational energy levels of the molecular species. More than 100 different molecular species have been observed, only a few of which can be observed at optical wavelengths. These include carbon monoxide, ammonia, water, ethyl and methyl

alcohol, formaldehyde, and hydrogen cyanide as well as some heavy organic molecules. Most of the observed molecules are found only in dense giant molecular clouds, especially near the center of the Galaxy. Carbon monoxide has been detected in some of the most distant known galaxies and quasars.

In the case of hydroxyl (OH), water (H_2O), and silicon oxide (SiO), collisions or ultraviolet radiation from nearby stars can "pump" the molecules to an excited state resulting in maser (microwave amplification by stimulated emission of radiation; a maser is the radio equivalent of a laser) emission. Maser emission is characterized by very narrow line widths and angular dimensions as small as one thousandth of a second of arc. Masers are found around both newly forming and dying stars. Very powerful H_2O masers found in other galaxies are known as megamasers. *See* ATOMIC STRUCTURE AND SPECTRA; GALAXY, EXTERNAL; INTERSTELLAR MATTER; MASER; MOLECULAR CLOUD; MOLECULAR STRUCTURE AND SPECTRA.

Pulsars. Pulsars (a term derived from pulsating radio stars) were discovered by Jocelyn Bell and Anthony Hewish in Cambridge, England, in 1967. Pulsars are neutron stars. Their electrons have all combined with protons in nuclei to form neutrons, and they have shrunk to a diameter of only a few kilometers following the explosion of the parent star in a supernova. Because they have retained the angular momentum of the much larger original star, neutron stars spin very rapidly, up to hundreds of times per second, and contain magnetic fields as strong as 10^8 tesla (10^{12} gauss) or more. The radio emission from pulsars is concentrated along a thin cone which produces a series of pulses corresponding to the rotation of the neutron star, much like a beacon from a rotating lighthouse lamp. Due to dispersion in the ionized interstellar gas, the arrival time of pulses observed at long wavelengths is delayed compared with short wavelengths, and observations of this dispersion measure are used to estimate the interstellar electron density and magnetic field strength. The rotation rates of pulsars are extremely stable. They are the most precise clocks known in nature with pulse repetition rates accurate to better than 1 second in a million years. Pulsars are found mostly close to the plane of the Milky Way, especially toward the direction of the center of the Galaxy.

Studies by Russell Hulse and Joseph Taylor of the slowing down of the pulsar PSR 1913+16 in close orbit about another neutron star led to the first observations of the effect of gravitational radiation. Precise measurements of the arrival times from the PSR 1257+12 in 1992 detected small perturbations due to the motion of planets orbiting the pulsar. This was the first detection of an extrasolar planetary system. *See* GRAVITATIONAL RADIATION; NEUTRON STAR; PULSAR.

Galaxies, quasars, and active galactic nuclei (AGN). Normal galaxies, particularly spiral galaxies, are sources of weak radio emission with a luminosity of about 10^{37} watts/Hz, comparable to the luminosity of our own Milky Way Galaxy. In some galaxies, regions of intense star formation release cosmic rays from supernovae causing intense radio emission which can be observed even at very great distances. Even more powerful are the radio galaxies, which can emit more than 100 million times more radio radiation than normal spiral galaxies do. Radio galaxies are surrounded by huge clouds of relativistic plasma located up to millions of light-years from the parent galaxy (**Fig. 3**) and containing electrons moving with nearly the speed of light in weak magnetic fields to produce the synchrotron radiation observed throughout the radio spectrum. By contrast, in the normal and star-forming galaxies, the radio emission comes from within the optically visible region of the galaxy.

In 1963, the study of radio galaxies led astronomer Maarten Schmidt to discover quasars. Quasars are

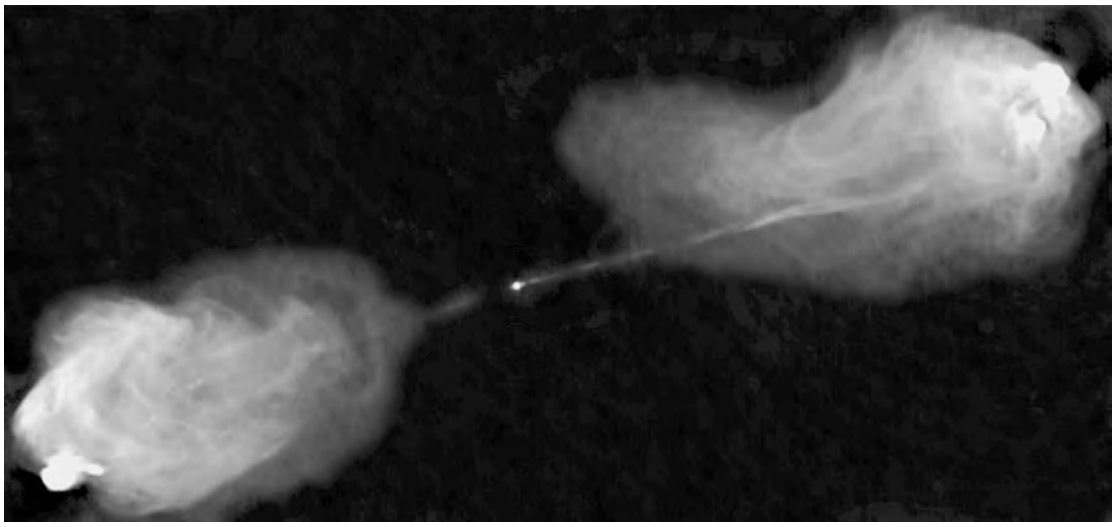

Fig. 3. Radio image of the radio galaxy Cygnus A showing the thin radio jet extending from the bright active galactic nucleus (AGN) toward the distant radio lobes. This 5-GHz (6-cm) image was made using the NRAO Very Large Array located near Socorro, New Mexico. (*NRAO/AUI/NSF*)

found in the central regions of rare galaxies and may shine with the luminosity of up to 100 ordinary galaxies. While some quasars also contain distant radio lobes, more often the radio emission from quasars is confined to a very small region coincident with the optically visible quasar. Very high resolution radio images may show relativistic plasma appearing to flow along a narrow jet away from the quasar with an apparent velocity greater than the speed of light. This so-called superluminal motion is an illusion due to the finite signal travel time from a relativistically moving source oriented close to the line of sight so that the radio source nearly keeps up with its own radiation. Most quasars are only weak radio sources, and the reason why only some 10 to 15% of all quasars become strong radio sources is unclear. Quasars which are not powerful radio sources are referred to as radio-quiet quasars. *See* QUASAR; SUPERLUMINAL MOTION.

The energy contained in the form of relativistic particles and magnetic fields surrounding radio galaxies and quasars is enormous, reaching in some cases as much as 10^{61} ergs. It is widely believed that the source of this energy is massive black holes at the center of the galaxy or quasar containing up to 10^9 times the mass of the Sun (see **colorplate**). Matter is accreted in an accretion disk surrounding the black hole. Electrons are heated to very high energy in the accretion disk and are then propelled out along a thin jet to form the distant radio-emitting plasma clouds (Fig. 3). The best direct evidence for the existence of a massive black hole comes from precise measurements of the angular velocity and acceleration of H_2O masers in the galaxy NGC 4258, leading to the conclusion that the masers are in orbit about a massive black hole containing 100 million times the mass of the Sun. The observed motions appear to accurately follow Kepler's laws of planetary motion, which allowed radio astronomers to more accurately determine the distance to the galaxy than previously possible. *See* BLACK HOLE; KEPLER'S LAWS.

The discovery of relativistic motion in radio jets was the foundation of the so-called unified models which postulate that the appearance of a radio source depends on its geometry as well as on its intrinsic properties. Special relativity shows that the radiation from a rapidly moving object is focused into a narrow cone along the direction of motion. For an object moving at 99% of the speed of light, the radiation can appear to be boosted by up to a factor of 1000 over an object at rest. The strong compact radio sources associated with quasars and active galactic nuclei (AGN) are thought to be the result of relativistic boosting of synchrotron radiation from plasma clouds ejected toward the observer from galactic nuclei. But when the plasma clouds are ejected in other directions, astronomers see only the radiation from the large, stationary, more distant clouds characteristic of radio galaxies.

Sometimes the observed radio emission from a distant AGN, quasar, or galaxy passes very close to a massive galaxy or cluster of galaxies, which bends both the radio and optical waves so that astronomers may see more than one image arriving from different directions. This phenomenon is known as gravitational lensing. Depending on the nature of the lensing galaxy and its angular separation from the remote source, there may be multiple images or just a single ring image, known as an Einstein ring, surrounding the lens. *See* GRAVITATIONAL LENS.

Since the first observations of large numbers of radio galaxies in the early 1960s, it has been clear that the number or luminosity of radio galaxies increases with increasing distance from our Galaxy. Due to the finite signal travel time of radio waves, this means that in the early universe, conditions were very different from the local (current epoch) universe, and such studies provided the earliest evidence that we are living in an evolving or changing universe.

Cosmic microwave background radiation. In 1965, Robert Wilson and Arno Penzias detected the faint comic microwave background (CMB) signal left over from the original big bang, which occurred some 14×10^9 years ago. As the universe expanded, it cooled to its present value of 2.7° above absolute zero. Subsequent observations with both ground- and space-based radio telescopes have detected the fine details in the CMB of less than one thousandth of a percent corresponding to the initial formation of structure in the early universe. These studies give new insight into the formation of galaxies and the rate of expansion of the universe, and show a surprising apparent acceleration of the expansion corresponding to an unknown form of dark energy.

When the faint CBR passes through a cluster of galaxies, some of the radiation may be scattered by relativistic electrons contained in the intergalactic gas, causing a slight diminution of the CMB observed in the direction of the cluster. This is known as the Sunyaev-Zeldovich effect. At the same time, the intergalactic electrons are then boosted in energy and radiate x-rays which are observed by astronomers using orbiting x-ray telescopes. Comparison of the x-ray emission, which depends on distance with the radio signal decrement, which is independent of distance, gives a direct measure of the distance to remote galaxies, independent of the conventional complex and less direct methods used by optical astronomers. *See* BIG BANG THEORY; COSMIC BACKGROUND RADIATION; COSMOLOGY.

Kenneth Kellermann

Bibliography. B. F. Burke and F. Graham Smith, *An Introduction to Radio Astronomy*, Cambridge University Press, 2d ed., 2002; J. D. Kraus, *Radio Astronomy*, 2d ed., Cygnus-Quasar Books, 1986; K. Rohlfs and T. Wilson, *Tools of Radio Astronomy*, 4th ed., Springer, 2003; G. L. Verschuur, *The Invisible Universe Revealed: The Story of Radio Astronomy*, Springer, 1986; G. L. Verschuur and K. I. Kellermann, *Galactic and Extragalactic Radio Astronomy*, Springer, 1991.

Radio broadcasting

The transmission, via radio-frequency electromagnetic waves, of audible program material for direct reception by the general public. Electromagnetic waves can be made to travel or propagate from a transmitting antenna to a receiving antenna. By modifying the amplitude, frequency, or relative phase of the wave in response to some message signal (a process known as modulation), it is possible to convey information from the transmitter to the receiver. In radio broadcasting, this information usually takes the form of voice or music. *See* ELECTRICAL COMMUNICATIONS; ELECTROMAGNETIC WAVE TRANSMISSION; MODULATION; RADIO.

Radio broadcasting occurs in seven frequency bands (**Fig. 1**). So-called longwave broadcasting is permitted by international agreement in a portion of the low-frequency band from 150 to 290 kHz in Europe. The most widely used broadcast band is in the medium-frequency (mf) range between 525 and 1700 kHz. It is commonly known as AM after amplitude modulation, the technique employed. So-called shortwave broadcasting is permitted worldwide in eight frequency bands between 5950 and 26,100 kHz. The very high frequency (VHF) band of 88 to 108 MHz is used for what is commonly called FM broadcasting, after frequency modulation that is used for transmissions. During the 1990s, a digital audio broadcasting (DAB) service in the 1452–1492- and 174–240-MHz frequency bands was put in place in Europe, Canada, and other countries. This band is unavailable in the United States for this service, so an alternate DAB system is being devised for use there. Radio broadcasting from satellites to listeners has been authorized in the 2310–2360-MHz frequency band. *See* AMPLITUDE MODULATION; AMPLITUDE-MODULATION RADIO; FREQUENCY MODULATION; FREQUENCY-MODULATION RADIO.

The frequencies set aside for radio broadcasting have been assigned by the International Telecommunications Union (ITU), an organization similar to the United Nations in which all countries may participate. The ITU is charged with allocating use of the entire radio spectrum (not just radio broadcasting), and developing technical standards through the International Telecommunications Union—Radio communications (ITU-R). The recommendations and technical standards for spectrum use are formally adopted as treaties among the member countries. *See* RADIO SPECTRUM ALLOCATIONS.

AM Medium-Frequency Band

This section discusses the more important technical aspects of AM broadcasting as it is currently used.

AM transmission standards. Broadcast stations in the medium-frequency band use amplitude modulation of a carrier wave to transmit information. The amplitude of the wave is modified in response to the changing amplitude of an audible voice or music signal. The AM receiver detects these amplitude changes and converts them back into audible signals, which can then be amplified and reproduced on acoustical transducers or speakers. *See* RADIO RECEIVER; RADIO TRANSMITTER.

The audible frequency range is generally considered to extend from 20 to 20,000 Hz. As a practical matter, AM broadcasting transmissions are limited to

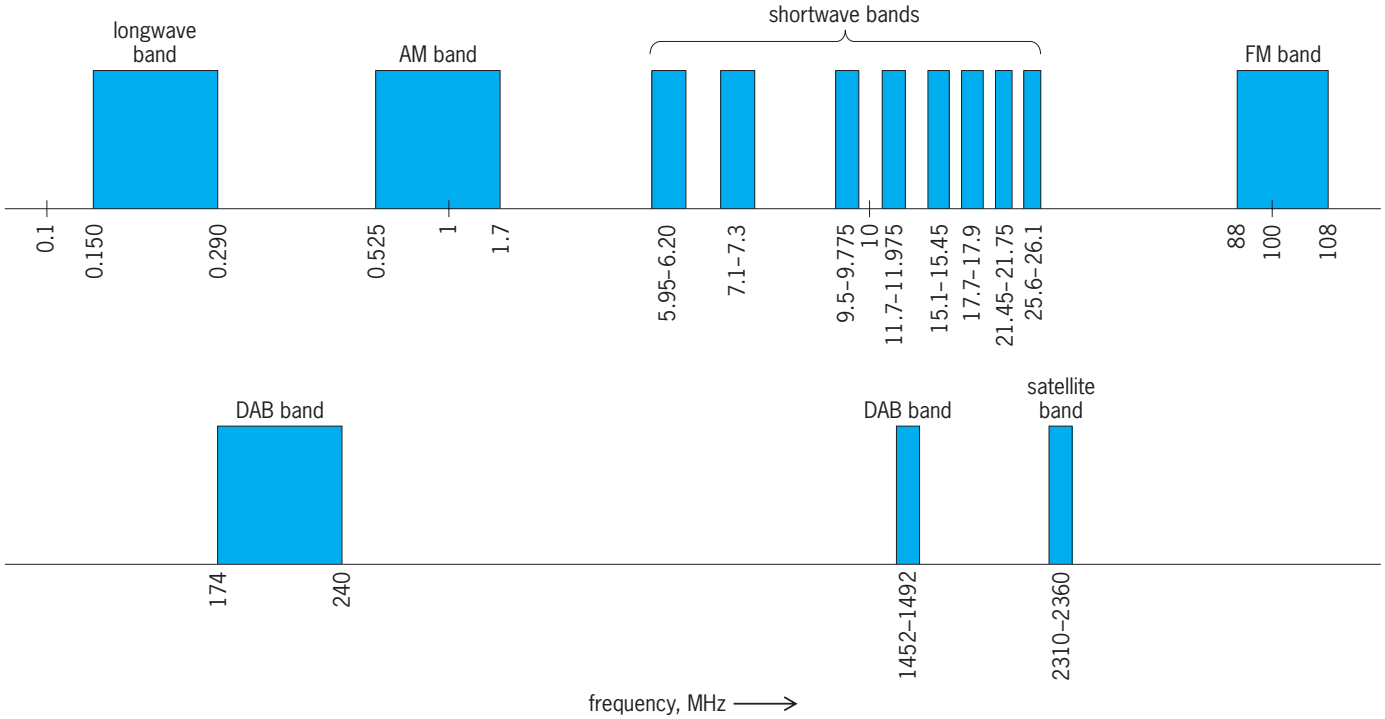

Fig. 1. Seven frequency bands used in radio broadcasting. Bands for digital audio broadcasting and satellite broadcasting are at higher frequencies.

a range of 50 to 10,000 Hz. Because of transmission components and directional antennas, the fidelity of many stations is more severely restricted, resulting in voice transmission that is still acceptable but music transmission that is of relatively low fidelity.

In North and South America and the Caribbean (ITU Region 2), AM stations are assigned to frequencies spaced 10 kHz apart. In most of the rest of the world, stations are spaced at 9-kHz intervals.

In the United States, the maximum permitted power is 50 kW. The vast majority of AM stations in the United States operate with 5 kW or less. Stations are often licensed with higher power levels for daytime than nighttime hours to allow for the increased potential for interference at night. In other parts of the world, AM transmitters operate at much greater power levels, with 1-MW power levels rather common in Europe, the Mediterranean, and the Middle East.

AM station allocation criteria. The fundamental concept in allocating and licensing broadcast stations is to limit interference from one station to the service area of another. For AM stations, the service area is defined as those locations receiving a signal level greater than 0.5 millivolt per meter. Because of atmospheric noise and unregulated artificial noise sources such as appliances, automobiles, and electrical systems, as a practical matter the 0.5-mV/m signal level may provide useful service only in electrically quiet rural areas. In urban areas, a signal level of 2 mV/m or higher may be necessary to overcome noise and provide useful service. *See* ELECTRICAL INTERFERENCE; ELECTRICAL NOISE.

To protect this service area, other stations on the same frequency (cochannel stations) are required to locate far enough away, or limit their transmitting power, so that their interfering signal is less than 5% of the desired station's signal strength at the boundary of the desired station's 0.5-mV/m service area, a 20:1 ratio of desired signal to undesired signal. For stations on adjacent channels ($\pm$10 kHz difference in carrier frequency), the required ratio is 1:1. This latter ratio is used primarily in the Western Hemisphere as an allocation standard; different, more restrictive, adjacent-channel ratios are used in other countries. Additionally, less restrictive desired-to-undesired signal ratios, or prohibitions on overlapping field-strength contours, are also employed when the frequency separation between stations is $\pm$20 and $\pm$30 kHz.

During nighttime hours a cochannel interference protection ratio of 20:1 is also employed. In addition, during nighttime hours the 0.5-mV/m, 50%-time sky-wave signal service area of stations in the highest class (class A) may also be protected from interference.

Transmitting antennas. An AM station may use a single tower for an antenna, resulting in an omnidirectional radiation pattern; or two or more transmitting towers to augment the radiation in certain directions while suppressing it in others, in order to comply with station allocation criteria. Such a directional antenna is developed by making use of the geometric relationship of the towers, and the relative amplitudes and phases of the currents in each tower, to create controlled constructive and destructive phase interference relationships in the desired directions. Directional antennas that use three to six towers are common, and stations with up to 12 towers have been built.

Since the allocation restrictions may be different during the daytime and nighttime hours, many stations employ two different directional antennas. **Figure 2** shows the daytime and nighttime radiation patterns for a three-tower directional antenna used by an AM station. The radiation from the antenna is expressed in millivolts per meter at 1 km (0.62 mi) from the antenna. *See* ANTENNA (ELECTROMAGNETISM).

Medium-frequency signal propagation. AM broadcast signals propagate from the transmitter by three mechanisms: ground-wave, space-wave, and sky-wave.

Ground waves travel along the ground surface (the boundary between the Earth and the atmosphere). Because they are surface waves, they penetrate into the ground, resulting in the energy being diminished because of losses in the ground. The degree of energy lost or signal attenuation is a function of the conductivity and permittivity of the near-surface ground, the frequency of operation, and the presence of any major surface discontinuities such as mountain ranges. The conductivity along the transmission path is the quantity primarily used to determine the extent of ground-wave propagation. The conductivity along salt-water paths can be as high as 5000 millisiemens per meter, while rocky terrain can have conductivities as low as 0.1 mS/m. Fertile farmland, alluvial plains, and other flat open areas can have conductivities ranging from 4 to 30 mS/m. Depending on the frequency, power level, and conductivity, useful ground-wave signals can propagate from a few tens of miles out to several hundred miles from the station.

The current flowing in the antenna also produces space waves, which travel through the atmosphere from transmitter to receiver. Space-wave propagation is usually limited by intervening terrain obstacles or the curvature of the Earth. Consequently, space-wave propagation is not as important for AM broadcasting as ground-wave and sky-wave propagation.

Sky-wave propagation occurs when space waves directed toward the ionosphere are reflected toward the Earth. This phenomenon can result in substantial signal strengths at distances of several hundred miles from the antenna. AM sky-wave propagation occurs primarily during nighttime hours by reflections from the E and F layers of the ionosphere at about 60 and 130 mi (100 and 220 km) altitude above the Earth's surface, respectively. During daytime hours the Sun's radiation ionizes the D layer of the ionosphere at about 40 mi (60 km) altitude above the Earth's surface. When the D layer is ionized, it tends to absorb and scatter the skyward-directed space waves from the AM antenna, thus preventing them from reaching the reflective E and F layers. As nighttime

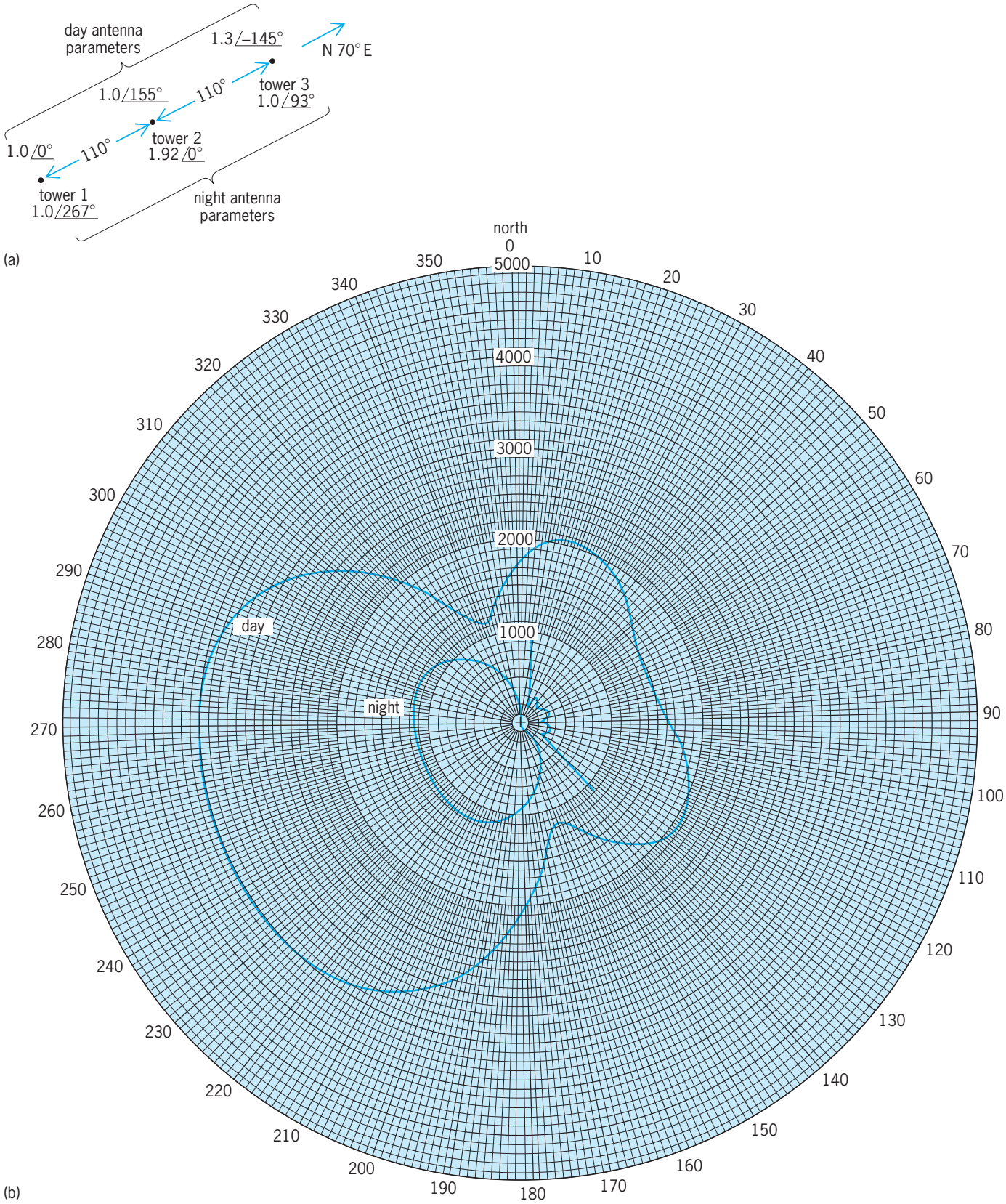

Fig. 2. Directional antenna array of the AM station KLOK in San Jose, California, broadcasting at 1170 kHz. (*a*) Configuration of array. Spacing between towers is given in electrical degrees. For the frequency of 1170 kHz, 1 electrical degree = 2.3 ft = 0.71 m and 110° = 257 ft = 78 m. The antenna parameters shown are field ratios and relative phases for each tower. (*b*) Resulting daytime (power = 50 kW) and nighttime (power = 5 kW) radiation patterns. The radiation from the antenna is expressed in millivolts per meter at 1 km (0.62 mi) from the antenna.

approaches, the ionization of the D layer diminishes, and so the AM signals pass through it to the E and F layers. The reflected sky waves increase in amplitude to levels at which they can provide useful service or substantial interference to other stations. *See* IONOSPHERE; RADIO-WAVE PROPAGATION.

Broadcast service area. The daytime ground-wave signal level protected by allocation criteria is usually 0.5 mV/m, although much higher signal strengths may be necessary to overcome noise from atmospheric and artificial sources, especially in highly urbanized areas. The extent of ground-wave service may be determined by using theoretical calculations or by measurements of the field strengths. The accuracy of the theoretical prediction method is limited primarily by the limited accuracy of the available conductivity data taken from maps, and secondarily by the approximate nature of the simple mathematical formulas used to calculate radiation from the antenna. Field-strength measurements provide an inefficient but definitive method of assessing station coverage.

During nighttime hours, the service area for most AM stations (other than class A stations) is usually limited by interference from other cochannel stations. In such cases, the service area is confined to the locations where the desired ground-wave field strength is 20 times greater than the RSS (root of the sum of the squares) of all the strongest interference signals on the channel at a given location. Because of interference during the night, the useful service may be limited to areas inside the 2.5 to 25 mV/m ground-wave contours. For low-power stations, this may be only 10 mi (16 km) or less from the transmitter.

For the highest class of stations (class A), nighttime service is protected by the allocation criteria to the 0.5-mV/m, 50%-time sky-wave contour. Because such service is subject to the time variations and fading of sky-wave propagation, the 0.5-mV/m, 50%-time contour is considered to be a secondary service area.

Figure 3 shows a plot of the theoretically predicted daytime and nighttime service areas for the station whose radiation patterns were shown in Fig. 2. The daytime service area is taken to be the 0.5-mV/m ground-wave contour, while the nighttime service area is limited by cochannel interference as described above to 8.37 mV/m. The distances to the contours vary because of radiation variations in the directional antenna and because of the different ground conductivity values in the region around the station, including that of salt water in San Francisco Bay.

AM stereo. In the United States, the Federal Communications Commission (FCC) has declined to designate a standard AM stereo transmission system. By 1990, approximately 600 stations, about 12% of the AM stations in the United States, operated in stereo. Of these, approximately 500 stations used a phase-modulation-type system, while another 100 used an independent-sideband system with the left and right channels transmitted on the lower and upper sidebands to form a complete composite AM stereo signal. *See* PHASE MODULATION; STEREOPHONIC RADIO TRANSMISSION.

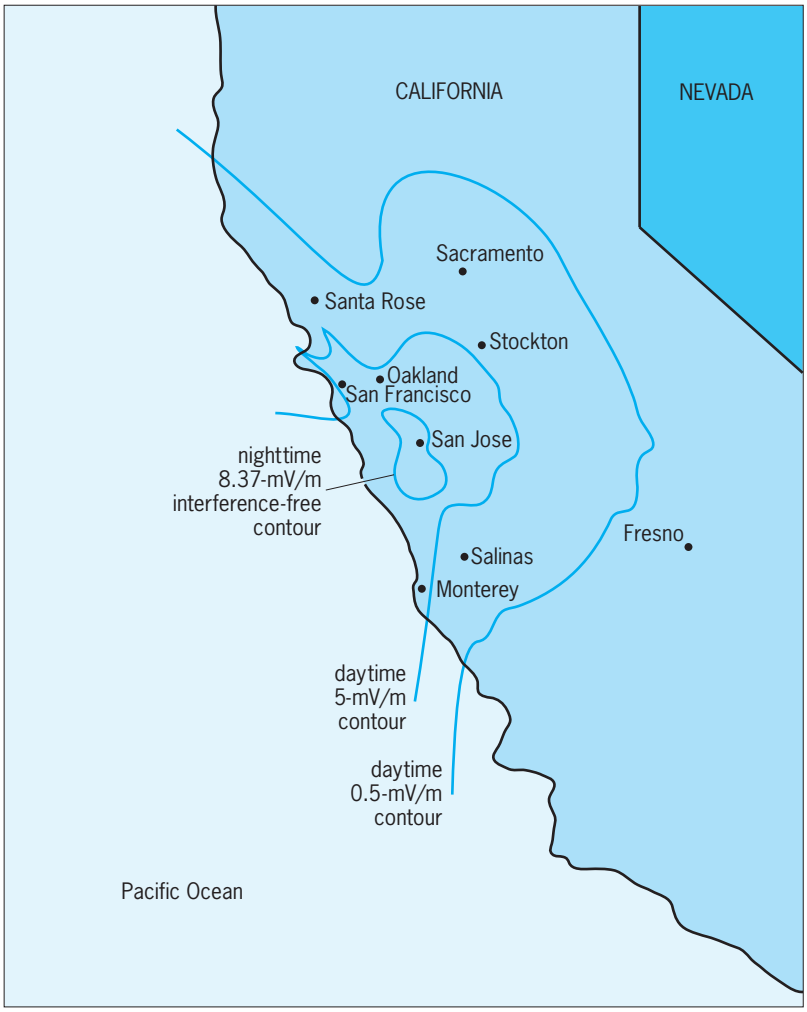

Fig. 3. Predicted daytime and nighttime interference-free field strength coverage contours of the AM station KLOK in San Jose, California.

FM VHF Band

FM broadcasting has become the dominant broadcast service in the United States primarily because of its better fidelity and its superior reception, which is less subject to noise and interference than that of AM.

FM transmission standards. Information is conveyed by frequency modulation or deviation of a carrier wave. In the United States, the carrier frequency may be deviated ±75 kHz around the assigned carrier frequency. The carrier frequencies, or channels, are spaced at 200-kHz intervals in the United States; a few other countries use slightly different channel spacings.

Nearly all FM stations transmit in stereo. A stereo audio signal consists of left and right channels. For broadcasting, the stereo signal is constructed at audio or baseband frequencies before being sent to the FM transmitter. The baseband frequency portion of the spectrum from 0 to 15 kHz (**Fig. 4**) is used for a signal that is the sum of the left and right channels ($L + R$). The frequency band from 23 to 53 kHz is used for an amplitude-modulated signal that is the difference between the left and right channels ($L - R$). A pilot tone at 19 kHz is also transmitted so that the receiver can recognize when a stereo signal is pres-

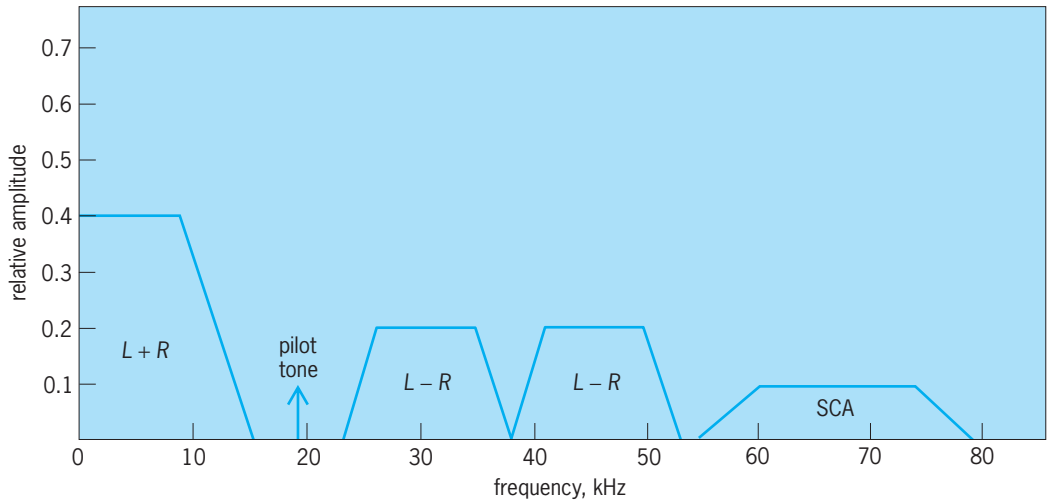

Fig. 4. FM broadcast baseband frequency spectrum.

ent. The pilot tone is also used to synchronize the detection of the $L - R$ signal. In the receiver, the $(L + R)$ signal is added to and subtracted from the $(L - R)$ signal to yield separate $2L$ and $2R$ audio signals at the receiver output. In this way the two separate left and right channels are reconstructed from the composite signal. This stereo transmission method was chosen so that it would be compatible with monaural FM broadcasts, which use only the 0–15-kHz $(L + R)$ signal.

The power of FM stations is limited by regulatory agencies so that the degree of coverage and potential interference to other stations is limited. A wide range of station classes has evolved using power levels from 3 to 100 kW. For FM, the allowed power rating is the effective radiated power (ERP) from the antenna, not the transmitter power. The effective radiated power is approximately equal to the transmitter power multiplied by the antenna gain minus any system losses.

FM station allocation criteria. FM stations are assigned to channels in the United States by using a fixed-distance separation table between transmitter sites. The amount of required separation depends on the station class, which in turn depends on the allowed effective radiating power and antenna height above average terrain (HAAT). For the highest class of station (class C) with 100 kW effective radiating power and 1969 ft (600 m) antenna height above average terrain, the required separation for stations on the same channel is 180 mi (290 km). Required separations are less for stations of lower class and for stations that are separated in frequency by ±200, ±400, or ±600 kHz.

Rather than use a fixed-distance separation scheme, some other countries use desired-to-undesired signal ratios to determine station locations to avoid interference. In still other countries, FM frequencies are freely assigned, with no plan for protecting service areas or limiting interference.

VHF signal propagation. The service area of an FM station depends on the propagation of space waves from the transmitter to the receiver. Ground waves and sky waves, which are the dominant propagation mechanisms for AM broadcasting, are unimportant in the VHF band. Space waves propagate through the atmosphere and are diffracted around, and reflected off, mountains, buildings, and other objects. Propagation within areas that have an unobstructed line-of-sight from transmitter to receiver is most reliable and predictable. When obstructions lie along the path, the FM signal strength is attenuated below the level that would exist for a line-of-sight path of the same length. The amount of attenuation can be calculated by using simple physical models based on diffraction theory. The accuracy of the calculation is reduced by the inability of simple models to account for real hills and mountains. *See* DIFFRACTION.

Because of reflections that occur with VHF FM transmission, a condition known as multipath can exist. Multipath occurs when signals from more than one direct or reflected signal source arrive at the receiver. The combination of these signals can be such that they destructively interfere with each other, resulting in diminished signal strength and distorted program reproduction.

The degree of signal diffraction and reflection that occurs depends on the reception location, and can vary with time because of changing atmospheric or other propagation path conditions. For this reason, VHF FM broadcast signal strengths are usually described statistically. For example, the field strength level exceeded 50% of the time at 50% of the locations is denoted F(50,50). A higher-reliability signal level is one exceeded 90% of the time at 90% of the locations, denoted F(90,90).

FM VHF service area. The primary service area of FM stations is generally taken as that area where the F(50,50) signal strength exceeds 1.0 mV/m. In the United States, the FCC requires each station to have a city of primary service, where the F(50,50) signal level must exceed 3.16 mV/m.

For FCC purposes, the service area of an FM station can be found by using a set of curves showing field strength as a function of distance, effective radiated power, and antenna height above average

Fig. 5. Predicted 3.16-mV/m coverage contour for the 100-kW station KUGN in Eugene, Oregon. The broken line represents coverage based on FCC field strength curves. The solid lines and "islands" represent a more realistic depiction of coverage that explicitly takes into account terrain obstructions.

terrain. For the highest class of station (class C), the 1.0 mV/m service extends to about 60 mi (96 km) according to FCC curves. In flat terrain, with a high antenna, the effective service may be rendered with field strengths much less than 1.0 mV/m, resulting in service areas that extend to 130 mi (210 km) or more from the transmitter.

Because of mountains and other obstructions, FM signal strength does not decrease monotonically with distance, but may fluctuate with distance (as well as with time). Consequently, the FCC propagation curves provide only an approximation of FM service. By explicitly taking into account signal attenuation over terrain obstacles (mountains), a more realistic estimation of FM service may be obtained. **Figure 5** shows the 3.16-mV/m service area for a 100-kW FM station. The broken line is the predicted FCC service area, while the solid lines show the service area prediction with terrain obstructions considered. The "islands" within the broken line are locations where the signal strength falls below 3.16 mV/m because of terrain obstacles.

Subcarrier transmission. To achieve greater utility from their transmitting facilities, FM broadcast stations have taken advantage of baseband spectrum space above 53 kHz for so-called subsidiary communication authorization (SCA) transmissions. For stereo stations, any frequency between 53 and 99 kHz may be used for an SCA, although the frequencies of 67 and 75 kHz are typically used. An SCA centered at 67 kHz is shown in Fig. 4. The amount of power that can be devoted to the SCA is limited

to 10% of the total station power, but within these loose technical constraints an FM broadcast station may transmit a wide variety of audio and digital signals on the SCA. Because the SCA signal is not detected by a normal FM stereo receiver, but requires a special receiver that is distributed to a limited audience (who usually pay for the service), SCA transmission services are sometimes referred to as narrowcasting rather than broadcasting. SCAs carry a wide range of services including background music, stock market data, and sophisticated nationwide paging services with watchband receivers that can automatically scan the FM frequency band and lock in on the FM broadcast station transmitting the appropriate digital coded message on the SCA. *See* RADIO PAGING SYSTEMS.

Shortwave Broadcasting

For reaching audiences in foreign countries or other distant places, shortwave broadcasting is most often used. Nearly 600 million shortwave radio receivers are in use worldwide.

Technical standards. Shortwave broadcasting is permitted worldwide in eight frequency bands from 5950 to 26,100 kHz. The assigned transmitting frequencies are spaced at 5-kHz intervals, resulting in a limited usable audio bandwidth. Voice transmissions are most effective, while music transmissions have limited fidelity.

Double-sideband (DSB) amplitude modulation is usually employed, although to make better use of the spectrum space some single-sideband (SSB) transmission is used. In single-sideband transmission, one sideband of a normal double-sideband signal is suppressed, thus reducing the occupied frequency bandwidth by about one-half without sacrificing audio bandwidth. The penalty with single-sideband transmission is that a more complex and stable receiver is required for suitable reception.

Transmitter powers for shortwave stations range from 1 to 500 kW, with the higher powers commonly used. Transmitters are normally coupled to high-gain directional transmitting antennas, which can result in effective radiated powers of several megawatts.

Frequency assignments. Four times each year, a shortwave station desiring to transmit to particular areas conducts a theoretical computer study of the propagation conditions from its transmitter site to the reception area for a coming season. The result of the study is a list of proposed frequency assignments and times when those frequencies are put to best use. This list is then submitted through the national government to the ITU. The ITU analyzes the frequency and time requests from all countries and then publishes a list of so-called incompatibilities or potential interference problems 3 months before the season begins. During this 3-month period, stations with conflicts have the opportunity to resolve them in bilateral agreements. In most cases, the interference conflicts can be resolved so that 85–90% of the transmissions will not suffer interference when operation for the season actually begins (based on the theoretical predictions).

To take best advantage of the varying propagation conditions throughout the day, a shortwave station may shift frequencies every 4 h to different bands, or redirect particular frequencies to different reception areas. The operating frequencies and schedules of a station may become quite complex.

Shortwave signal propagation. Shortwave signals propagate via sky waves that are reflected one or more times from the E and F layers of the ionosphere. Multiple reflections are possible because a signal can also bounce off the Earth's surface after reflecting off the ionosphere in a "Ping-Pong" effect. With multiple reflections, signals can propagate around the Earth, reaching countries distant from the broadcasting station. Because the degree of ionization of the E and F layers varies with the Sun's radiation, shortwave propagation is highly dependent on the solar illumination along the transmission path between the transmitter and the receiver.

Shortwave service areas. Several factors limit the reception of a shortwave signal, chief among them being atmospheric noise, interference, and signal fading. Through the use of high-gain antennas, shortwave stations can target a desired reception area by concentrating the energy in that direction. Usually a signal strength of 0.25 to 1.0 mV/m, 50% of the time, is needed in the targeted service area to overcome atmospheric noise and provide good reception. It may be impossible to overcome a strong cochannel interference source, and sometimes such interference has been used to deliberately jam reception of another signal. Signal fading is a result of the varying conditions of the ionosphere; however, usable signal strengths can sometimes exist for several hours even with changing atmospheric conditions.

Digital Audio Broadcasting

To enhance audio quality, a digital audio broadcasting service has been put in place in Europe and elsewhere, operating primarily in the 1452–1492-MHz band. As in VHF FM broadcasting, DAB uses transmitters located at elevated locations (mountaintops, building roofs) that provide the best line-of-sight paths to the intended service area. Digital broadcasting differs from VHF FM broadcasting in that the audio signal (voice or music) is first converted to a stream of binary digits (data bits) that represent the audio signal. These data bits are then used to modulate the radio-frequency carrier signal using one of several techniques. After transmission via the radio waves, the radio-frequency carrier is demodulated at the receiver to recover the stream of data bits, and the bits are then converted back to the audio signal. DAB offers improved reception quality and fidelity because error-correcting codes in the digital signal can be used to eliminate many flaws that may occur during transmission. It is not possible to use such codes to correct similar flaws in AM and FM analog transmissions. *See* INFORMATION THEORY; MODULATION.

The DAB system in use in Europe and elsewhere is known as Eureka 147, and employs a digital modulation technique known as orthogonal

frequency-division multiplexing (OFDM). This technique provides improved resistance to transmission problems resulting from fading and multipath that limit the coverage and quality of traditional VHF FM broadcasting. *See* MULTIPLEXING AND MULTIPLE ACCESS.

Due to the shortage of radio spectrum in the United States, a different DAB system that shares the same radio channel as medium-wave AM or VHF FM is being developed. Using sophisticated digital techniques, the aim of these systems is to allow simultaneous transmission of traditional analog AM and FM signals in the same channel as the digital broadcasting service.

Satellite Broadcasting

Satellite broadcasting also uses digital modulation techniques. However, in this case the transmitters are located on satellites high above the Earth. From this position, satellite broadcasting can achieve essentially universal coverage of an entire nation or most of a continent from one or two transmitters. Both geostationary and low-Earth-orbit (LEO) satellites are used for these systems. Specialized receiving antennas placed in locations visible to the sky are usually needed, along with a special radio receiver specifically designed for satellite service. To extend coverage into areas that are not visible to the sky, some satellite broadcasting networks employ a network of ground-based transmitters (repeaters) to supplement the coverage of the satellite signal. Satellite broadcasting is capable of providing 100 or more channels of audio programming, usually on a subscription or fee basis. *See* DIRECT BROADCASTING SATELLITE SYSTEMS.

Harry R. Anderson

Bibliography. D. G. Bobbett (ed.), *World Radio-TV Handbook 2001*, Watson-Guptill Publications, 2001; Federal Communications Commission, *Rules and Regulations, Part 73*, 1999; Institute of Electrical and Electronic Engineers, *IEEE Transactions on Broadcasting* (journal), quarterly; International Telecommunications Union, *Recommendations and Reports of the ITU-R*, vol. 10: *Broadcasting Service* (*Sound*), 1998; National Bureau of Standards, Transmission Loss Prediction for Tropospheric Communication Circuits, *NBS Tech. Note*, no. 101, 1967.

Radio broadcasting network

A group of radio broadcast stations interconnected by leased channels on wire, fiber, microwave, or satellite to one or more central feed points for the purpose of receiving and rebroadcasting program material of a timely nature to a geographically diverse audience; or alternatively, an organization which provides programming for, but does not actually own or operate any, broadcast stations. The terms "radio broadcasting network" and "radio network" are often used interchangeably.

The need for radio networks arises from the local nature of terrestrial radio broadcasting, which is due both to the physical nature of how radio waves propagate and to the way in which terrestrial radio services are regulated. At best, each AM and FM radio station will be receivable only on the order of 100 mi (160 km) from the transmitter site, with the exception of high-powered AM stations at night (so-called clear channel stations) which, because of a phenomenon known as skywave propagation, can be received over much greater distances (albeit with substantial interference). This distance is often further reduced by interference from what otherwise might be possible due to the manner in which broadcasting licenses are allocated. *See* RADIO BROADCASTING; RADIO-WAVE PROPAGATION.

In the United States, nongovernmental use of radio-frequency (RF) transmissions are regulated by the Federal Communications Commission (FCC), including the allocation of "channels" in the AM and FM broadcast bands which utilize RF frequencies 550–1705 kHz and 88–108 MHz, respectively. These channels, which listeners typically refer to by their broadcast frequency or call sign (a four-letter combination beginning with either W or K), are allocated by the FCC on a community-by-community basis. Before channels are allocated to a particular community, engineering analyses are performed to determine the impact on other nearby radio signals already allocated on either the same frequency or on the frequencies adjacent to the desired allocation channel. Based on these studies, the FCC determines the allowable distance a station can transmit without causing interference to its neighbors, which dictates the power level at which the broadcast station can operate. *See* RADIO SPECTRUM ALLOCATIONS.

The use of radio networks makes it possible to overcome these distance limitations and broadcast live and prerecorded programs to many communities, or even a national or global audience, through affiliated terrestrial radio stations (affiliates); the stations make national and regional markets available to advertisers and offer stations entertainment and public service programs.

Types of radio networks. Radio networks exist in many forms and are distinguished by whether they are government-owned (or sponsored) or are operated as a commercial enterprise (being either publicly or privately owned), whether they are free or subscription services, and the transmission medium used for reception. Many countries around the world have broadcasting systems which consist solely of government-run or -affiliated networks, although over time the trend is that many of these countries are developing commercial broadcast segments as well.

Commercial networks. Radio networks in the United States began as commercial enterprises and today remain a model for non-government-centric broadcasting systems which other countries (having a history of government-only networks) emulate. Commercial radio networks obtain revenue primarily

through the sale of advertisements which are broadcast along with network-produced or -purchased programming (such as news, entertainment, and sports).

Public networks. Public networks are nonprofit in nature and can be owned by either the government or not-for-profit corporations. They are funded either exclusively by the government or through a combination of government funding and private donations. In the United States, National Public Radio (NPR), a not-for-profit corporation, is the principal public radio network with over 780 affiliate radio stations in the AM and FM bands. NPR's government funding is provided principally by the Corporation for Public Broadcasting (CPB).

International networks. Many countries operate radio networks to deliver programming outside of their own, often to provide free and uncensored information to citizens of countries where government censorship is practiced. An example is the Voice of America (VOA), which operates a global network of transmitting stations that transmit over 1000 hours of news, informational, educational, and cultural programs every week to an audience of more than 100 million people worldwide in 44 languages. VOA broadcasts primarily in the AM and shortwave (approximately 6–18-MHz) bands.

Satellite Digital Audio Radio Service (SDARS). Radio services broadcast directly from a satellite to the listener are a new form of radio system made possible by advances in perceptual audio coding, mobile satellite antenna, and digital signal processing technologies. An SDARS system utilizes one or more satellites (in most cases complemented by terrestrial repeaters) to provide multiple channels of programming (on the order of 100) to listeners over a broad area, resulting in a coverage area commensurate with that of a radio network but with far fewer transmitters. So far, all SDARS systems are subscription services and have been built and operated by individual companies, which distinguishes them from terrestrial radio networks in that there are no affiliate stations. SDARS systems make use of RF signals which are at much higher frequencies than those of terrestrial radio stations, typically in the L-band (1452–1492 MHz) and S-band (2310–2360 MHz). The first SDARS system, WorldSpace, began broadcasting to Africa in 1999 and to Asia in 2000. The United States has two SDARS providers, XM Radio and Sirius, which began operation in November 2001 and July 2002, respectively. *See* SATELLITE RADIO.

Internet radio. The proliferation of the Internet and audio streaming technology has led to the development of so-called Internet radio stations, which have the ability to reach a global audience since their transmission medium is the Internet. As wireless Internet installations increase in number, utilizing wireless fidelity (Wi-Fi) technology, Internet radio will be available to mobile listeners and may over time offer competition to terrestrial and satellite radio services. Currently the principal limitations of Internet radio are that listening is primarily tied to the use of a personal computer, and the limited capacity of the computer servers from which each Internet radio station originates, since the amount of bandwidth required to broadcast the station increases incrementally with each listener. As multicast server technology is implemented throughout the Internet, this bandwidth problem will diminish. *See* INTERNET; WIRELESS FIDELITY TECHNOLOGY.

Early history. The development of radio broadcasting networks (originally called chain broadcasting) in the first part of the twentieth century is analogous to the development of the first printing press in that it revolutionized the manner in which information was disseminated. Radio networks made possible, for the first time, the simultaneous, instantaneous delivery of news and information on a continental (later global) scale. It is generally agreed that the first example of a radio network was demonstrated in 1923 involving two stations—WEAF(AM) in New York and WNAC(AM) in Boston—connected using a telephone line. Other methods of connection were experimented with, including telegraph lines and shortwave radio links (able to be received over greater distances than AM broadcast signals), but these proved to be unreliable and did not provide for audio of sufficient quality.

The first bona-fide radio network was the National Broadcasting Company (NBC), created in 1926 as a joint venture of the Radio Corporation of America (RCA), Westinghouse, and General Electric. This early network (and others too) owned the broadcast facilities and was responsible for the content broadcast over them. As networks proliferated, the government became concerned about the level of control being exercised by the networks on their affiliates, and in 1940 the FCC issued its "Report on Chain Broadcasting" which, in addition to reducing this level of control, placed limits on the amount of time the networks may broadcast through their affiliates.

Distribution technologies. Distribution of network programming to affiliates was accomplished using wired telephone circuits or wireless terrestrial microwave circuits until the 1980s, when the availability of affordable satellite uplink and downlink equipment and the ability for radio networks to lease satellite capacity made possible widespread use of geosynchronous satellite circuits as well. This distribution technology is similar to terrestrial microwave, using uplink and downlink frequencies in both the C-band (4–6 GHz) and Ku-band (12–14 GHz). Each satellite contains banks of transponders, each of which receives the uplink signal for amplification and retransmission by the downlink transmitter, which returns the signal to Earth station receivers. *See* COMMUNICATIONS SATELLITE.

These geosynchronous satellites, operated and maintained in their orbiting positions by the satellite service providers, repeat the signal on assigned transponder frequencies, beaming a broad pattern of signal which can cover an entire continent, or alternatively, beaming a narrower pattern, using a "spot" beam, for more localized coverage. The signal may be received in any location within the footprint of the

satellite beam by the affiliated radio station utilizing low-noise receiving amplifiers centered in the focus of microwave receiving antennas aimed toward the satellite.

NPR is generally acknowledged to have first successfully demonstrated stereo frequency-modulation (FM) distribution via satellite facilities, and today operates the Public Radio Satellite System (PRSS), a satellite-based distribution network which includes more than 400 downlinks, and is the principal means by which content is delivered to NPR affiliate stations. NPR continues to be a leader in radio network distribution technology. Its Content Depot system is based on Internet Protocol (IP) technology and will support distribution over a variety of media including satellite, terrestrial (point-to-point), and the Internet. *See* RADIO. David H. Layer

Bibliography. *Broadcasting and Cable Yearbook*, annually; W. B. Emery, *National and International Systems of Broadcasting: Their History, Operation, and Control*, Michigan State University Press, East Lansing, 1969; Federal Communications Commission, *Report on Chain Broadcasting*, Commission Order no. 37, Docket no. 5060,1941; Federal Communications Commission, *Rules and Regulations*, Parts 73 and 74, and *Amendments*, 2004; W. L. Hetrich, *An Improved Audio Pipeline: NPR's Satellite System*, National Public Radio, 1982; H. Sterling and J. M. Kittross, *Stay Tuned—A History of American Broadcasting*, 3d ed., Lawrence Erlbaum Associates, 2002.

Radio compass

A popular term for an automatic radio direction finder used for navigation purposes on ships and aircraft. It is not strictly a compass, because it indicates direction with respect to the radio station to which it is tuned, rather than to the north magnetic pole. The modern radio compass uses a nondirectional antenna in combination with a bidirectional loop antenna to provide a unidirectional bearing indication. The navigator thus knows at all times whether the craft is traveling toward or away from the radio station used as a reference. The antennas are used with a special radio receiver that provides a visual indication of direction on a meter or cathode-ray indicator. *See* DIRECTION-FINDING EQUIPMENT. John Markus

Radio-frequency amplifier

A tuned amplifier that amplifies the signals commonly used in radio communications. Amplifier designs in the radio-frequency (RF) range differ significantly from conventional low-frequency circuit approaches; they consequently require special, distributed circuit considerations. The need for a distributed treatment is warranted whenever the size of the physical circuit is greater than about one-tenth of the signal wavelength.

The wavelength of a radio-frequency signal in free space is obtained from Eq. (1), where c is the speed

$$\lambda = c/f \tag{1}$$

of light ($3 \cdot 10^8$ m/s) and f is the frequency of the signal in hertz. The wavelength of a 3-gigahertz signal in free space, for example, is $\lambda = 3 \cdot 10^8/3 \cdot 10^9 = 0.1$ m, or 10 cm (4 in.). Thus, lumped circuit theory will not accurately model the behavior of 3-GHz circuits whose dimensions are in the 1-cm range, and Kirchhoff's voltage and current laws must be abandoned in favor of voltage and current wave propagation phenomena. An immediate implication is that appropriate distributed matching networks are needed to reduce standing-wave effects and avoid undesirable oscillations. The designs of suitable matching networks in conjunction with stability and noise figure analyses are usually the first steps in the amplifier design process. Additional considerations for a radio-frequency amplifier include gain, gain flatness, gain compression, output power, bandwidth, harmonic generation, and bias conditions.

Transition from lumped to distributed design. The starting point for a conventional amplifier circuit is the selection of an active device, either a bipolar junction transistor (BJT) or a field-effect transistor (FET). For optimal power transfer and amplifier stabilization, the device typically requires that input and output impedance matching networks be interposed between its input source and its load. In their simplest forms, each of these networks consists of two reactive elements that transform the impedances of the radio-frequency source and load, typically 50 ohms, to new values at the active device plane. The level of the input source must be matched to the input impedance of the transistor, which is typically very low. For the transistor amplifier to deliver maximum output power to its load, its output impedance—typically higher than 50 ohms—must be decreased to the 50-ohm level of its load. *See* ELECTRICAL IMPEDANCE; IMPEDANCE MATCHING; TRANSISTOR.

Figure 1 depicts an example of a lumped circuit model for a narrow-band amplifier where a capacitor C_1 followed by an inductor L_1 is used to create the input network, while an inductor L_2 followed by a capacitor C_2 is used for the output network. The inductors, RFC (for radio-frequency choke), and capacitors, C_B, isolate the radio-frequency signal paths from the dc bias sources. *See* CAPACITOR; INDUCTOR.

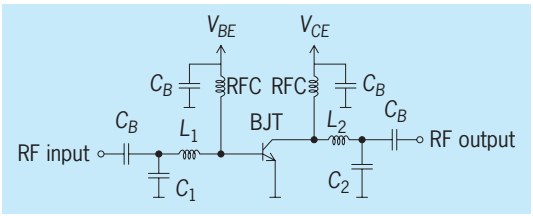

Fig. 1. Generic lumped element amplifier with input (L_1, C_1) and output (L_2, C_2) matching networks. Capacitors C_B and inductors RFC separate the radio-frequency paths from the required bipolar junction transistor (BJT) bias conditions.

Unfortunately, this design implementation suffers serious drawbacks at high operational frequencies. Specifically, the lumped elements in the radio-frequency paths no longer act as ideal inductors and capacitors; resistive losses and parasitics cause undesired departures from the ideal electric circuit behaviors at high frequency. Fortunately, transmission lines—pairs or conductors of carefully chosen dimensions—can be made to function as capacitors or inductors at frequencies where discrete inductors and capacitors perform poorly. Whether a length of transmission line acts as a capacitor or an inductor depends on its length at its operating frequency and on whether its far end is open or shorted. *See* TRANSMISSION LINES.

The signal propagation speed in transmission lines is usually considerably less than that in free space, so the electrical wavelength of a signal on a transmission line is noticeably shorter than its free-space wavelength. The electrical wavelength is given by Eq. (2), where c is the speed of light, f is the circuit frequency of operation, and ε_r is the relative dielectric constant of the circuit-board material. For a relative dielectric constant of 4, the electrical wavelength of a signal on a circuit board will be one-half its wavelength in free space. Therefore the wavelength of the 3-GHz signal on the circuit board would be 5 cm (2 in.). Most lines used in circuit design are about 1/4 wavelength long, around 1.25 cm (0.5 in.) in the 3-GHz case.

$$\lambda = c/(f\sqrt{\varepsilon_r}) \tag{2}$$

Figure 2 depicts an amplifier circuit using transmission-line segments whose response is equivalent to the circuit of Fig. 1. Because of its parasitic properties, the circuit of Fig. 1 will not perform properly at high frequencies, while the one of Fig. 2 will. Open-circuited transmission lines of less than 1/4 wavelength at the frequency of operation of the amplifier appear as capacitors, while those greater than 1/4 wavelength act as inductors. The value of capacitive or inductive reactance of an open-circuited line depends on the length of the line and on its characteristic impedance. The special-case line segments denoted $\lambda/4$ are exactly 1/4 wavelength long. When a transmission line is exactly 1/4 wavelength long, its input terminals will look like a short circuit if its far end is open-circuited, and an open circuit if its far end is short-circuited. This quarter-wave property is used to emulate the action of a radio-frequency choke. In Fig. 2, for example, the far end of each quarter-wave line effectively is shorted by a capacitor that has a very low reactance at the desired operating frequency. At the transistor terminals, 1/4 wavelength away, this short appears as an open circuit to radio-frequency energy, and these connections do not disrupt the operation of the matching networks. At dc the lines are direct low-resistance connections between the power sources and the amplifier collector and base bias circuits. *See* CHOKE (ELECTRICITY).

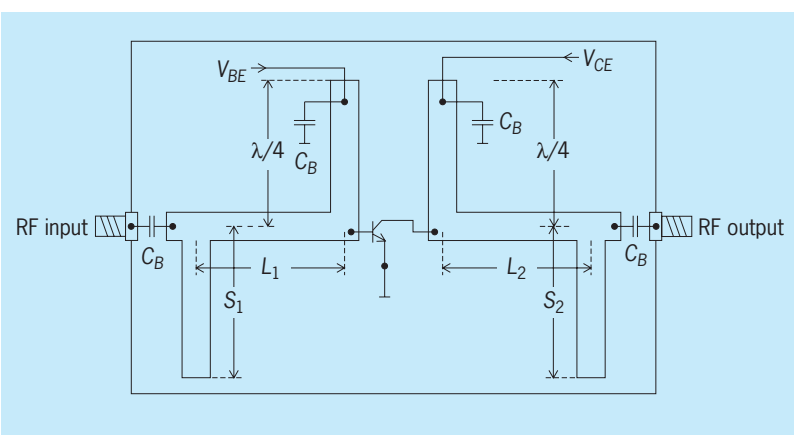

Fig. 2. Basic layout of input and output matching networks with microstrip lines.

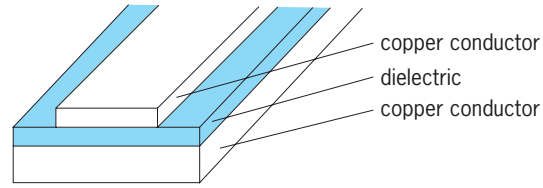

Fig. 3. Microstrip line.

Microstrip, a particularly useful form of transmission line for these applications, is implemented by laying down copper strips on the insulating surface of a circuit board whose opposite side is completely covered by a thin layer of copper (**Fig. 3**). The critical parameters of the line are its copper thickness and width, the thickness and relative permittivity of the insulating (dielectric) layer between the strip and the copper substrate, and, to a lesser extent, the length of the line.

Microstrip design approach. A microstrip line implementation of the input and output matching networks of Fig. 2 for a narrow-band amplifier at 3 GHz takes the following form: For a 0.76-mm-thick (30-mil) circuit board material with relative dielectric constant $\varepsilon_r = 4.4$, a trace width of 1.45 mm will achieve a characteristic impedance of $Z_0 =$ 50 ohms. The effective dielectric constant, however, is only about 3.33 since the microstrip line resides at an air-dielectric interface. Therefore the electrical wavelength of a 3-GHz signal on this line is given by Eq. (3): about 55 mm. The two-element input matching network, L_1 and S_1, needs to transform the transistor input impedance to 50 ohms. A transmission line segment, L_1, of length 9.6 mm followed by a capacitive stub, S_1, of length 5.5 mm does the job.

$$\begin{aligned}\lambda &= c/(f\sqrt{\varepsilon_{eff}})\\ &= 3\times10^8/(3\times10^9\sqrt{3.33}) = 0.0548\ \text{m}\end{aligned} \tag{3}$$

Similarly, the output matching network, L_2 and S_2, transforms the conjugate complex impedance of the transistor to the external load impedance of 50 ohms. The network capacitive stub, S_2, is of length 9.1 mm and its transmission line segment, L_2, is 12.7 mm

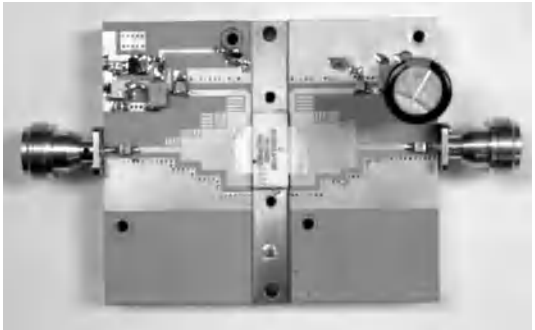

Fig. 4. Board layout for field-effect transistor (FET) power amplifier with microstrip networks for input and output matching. Total board dimensions are 76 mm (3 in.) in length by 40 mm (1.575 in.) in width. (*From J. E. Davis et al., The Characterization of a Power Amplifier, Senior Project Report, Worcester Polytechnic Institute, Worcester, MA, 2005*)

long. The amplifier gain with this matching arrangement is about 16.57 dB.

Finally, the $\lambda/4$ line segments, each 13.7 mm in length, transform radio-frequency ground (at 3 GHz, capacitor C_B is a short circuit) to very large impedances at the transistor terminals.

Power amplifier matching. A practical matching network for a power amplifier is shown in **Fig. 4**. A high-performance metal-oxide-semiconductor (MOS) field-effect transistor is employed in this case. The amplifier is designed for cellular base station applications in the frequency range 1.9–2.4 GHz. The MOS device is rated for 24 W average output power and 150 W peak power. It operates at a nominal drain-source voltage of 28 V and delivers up to 15 dB of gain.

Figure 4 depicts the device installed on a circuit board with a relative dielectric constant of 3.5, a height of 0.76 mm (30 mil), and a copper conducting surface of thickness 35 μm (1.4 mil). The conductor routing forms the biasing interconnections and the transmission line matching circuits. The bias voltages are supplied through 17.866-mm-long (0.7034-in.) $\lambda/4$ microstrip lines terminated by bypass capacitors on the dc side. The matching networks are constructed from microstrip segments of varying width and length ratios. The widest section of microstrip adjacent to the device has very low characteristic impedance; this is necessary to transform the very low device input impedance to a reasonable level. Several free-floating copper patches are supplied around the matching circuits. These can be connected selectively to the microstrip segments to adjust the matching. Although not visible in Fig. 4, the copper base plate is attached to a water cooling system. *See* AMPLIFIER.

Reinhold Ludwig

Bibliography. C. Coleman, *An Introduction to Radio Frequency Engineering*, Cambridge, 2004; R. Ludwig and P. Bretchko, *RF Circuit Design: Theory and Applications*, Prentice Hall, 2000; F. H. Raab et al., RF and microwave power amplifier and transmitter technologies—Part 1, *High Freq. Electr.*, 2(3):22–36, May 2003.

Radio-frequency impedance measurements

Measurements of electrical impedance at frequencies ranging from a few tens of kilohertz to about 1 gigahertz. This broad radio-frequency (rf) range encompasses frequencies from the extralow-frequency band, through the long-, medium-, and short-wave ranges, and then through the very high frequency (VHF) and the ultrahigh-frequency (UHF) bands. In the electrical context, impedance is defined as the ratio of voltage to current (or electrical field strength to magnetic field strength), and it is measured in units of ohms (Ω). *See* ELECTRICAL IMPEDANCE; RADIO SPECTRUM ALLOCATIONS.

At zero frequency, that is, when the current involved is a direct current, both voltage and current are expressible as real numbers. Their ratio, the resistance, is a scalar (real) number. However, at nonzero frequencies, the voltage is not necessarily in phase with the current, and both are represented by vectors, and therefore are conveniently described by using complex numbers. To distinguish between the scalar quantity of resistance at zero frequency and the vectorial quantity at nonzero frequencies, the word impedance is used for the complex ratio of voltage to current. The real part R of the complex impedance, $Z = R + jX$, is the (series-equivalent) resistance, and the imaginary part, the reactance X, is the reactive impedance of the (series-equivalent) inductance or capacitance (for positive or negative signs, respectively). Admittance Y is defined as the inverse of impedance, $Y = 1/Z$, and some instruments measure admittance directly, in units of siemens (S). Network theory states that at a given frequency the impedance of any two-terminal linear (that is, level independent) circuit may be represented either as a series or a parallel connection of a resistance and a reactance. *See* ALTERNATING-CURRENT CIRCUIT THEORY; DIRECT CURRENT; ELECTRICAL RESISTANCE; NETWORK THEORY.

The measurement of impedance at radio frequencies cannot always be performed directly by measuring an rf voltage and dividing it by the corresponding rf current, for the following reasons: (1) it may be difficult to measure rf voltages and currents without loading the circuit by the sensing probes; (2) the distributed parasitic reactances (stray capacitances to neighboring objects, and lead inductances) may be altered by the sensing probes; and (3) the spatial voltage and current distributions may prevent unambiguous measurements (in waveguides, for instance).

Impedance measurements may be classified according to the principle used, such as resonance methods, bridge methods, or reflectometers, or may be classified according to the physical nature of the conductors, such as two-terminal, three-terminal, coaxial line, stripline, or waveguide.

Bridge methods. At low frequencies, impedance measurements are often carried out by measuring separately the resistive and reactive parts, using either Q-meter instruments (for resonance methods), or reconfigurable bridges, which are sometimes

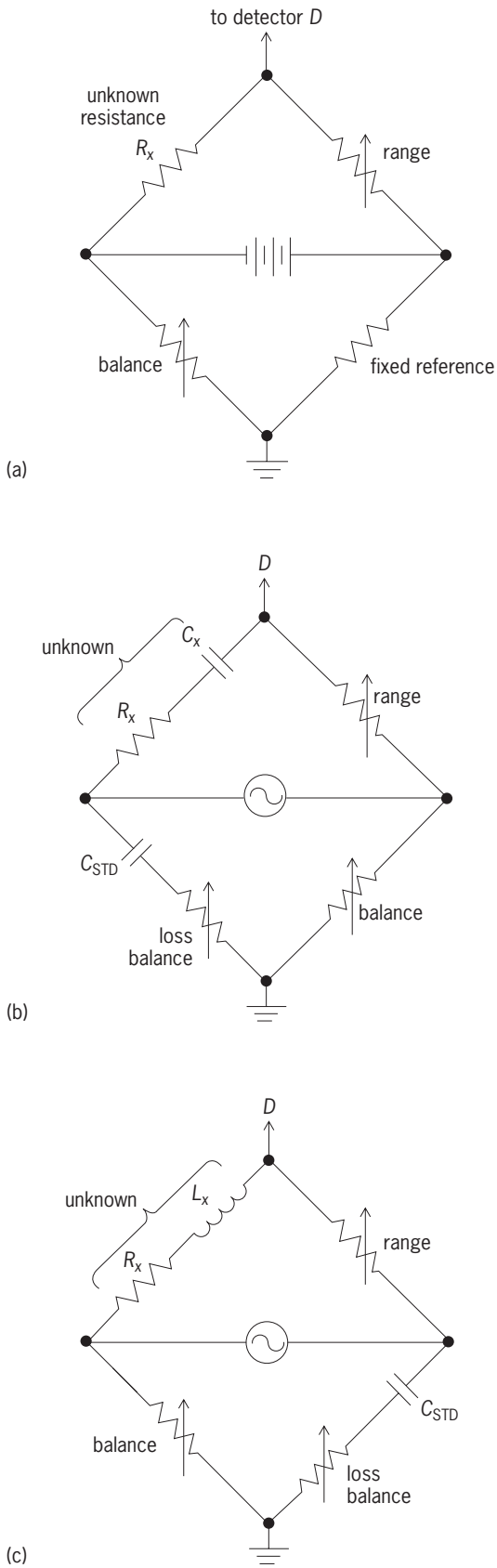

Fig. 1. Three internal bridge circuits used in an *LCR* bridge. (*a*) Wheatstone bridge. (*b*) Series-resistance-capacitance bridge for capacitance measurements; C_{STD} is standard capacitance. (*c*) Maxwell bridge for the measurement of inductances.

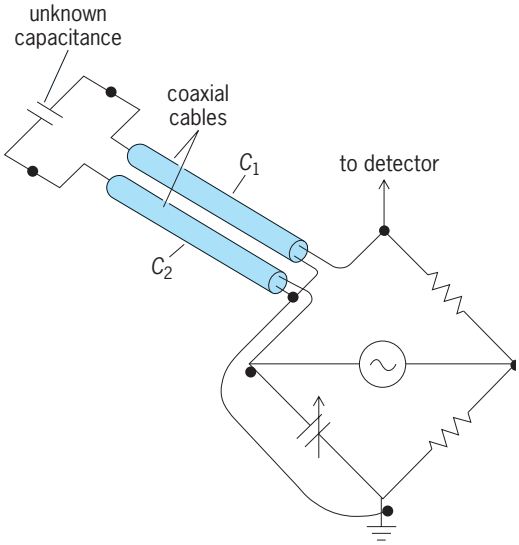

Fig. 2. Three-terminal measurement of impedance. This connection permits the remote measurement of impedance, since the measurement is between the inner conductors of the connecting coaxial cables, and lead capacitances are shielded from each other. Cable capacitance C_1 is across the detector. For accurate measurements, C_2 may be subtracted from the standard capacitance.

called universal *LCR* (inductance-capacitance-resistance) bridges. In one such bridge the resistive part of the impedance is measured at dc with a Wheatstone bridge (**Fig. 1***a*). Capacitive reactance is measured with a series-resistance-capacitance bridge (Fig. 1*b*), and inductive reactance is measured with a Maxwell bridge (Fig. 1*c*), using alternating-current (ac) excitation and a standard capacitance. If the impedances of open structures or remote components have to be measured, lead capacitances may significantly alter the impedance sought. In such cases, a quasi-three-terminal measurement may be performed by making connections to the measurand via shielded (coaxial) cables, in which case the bridge senses primarily the impedance between the inner conductors. The outer conductors of the cables are grounded (**Fig. 2**). The capacitance of one cable is across the detector, and therefore has no effect, and for accurate measurements the capacitance of the other cable may be measured and subtracted from the value of the standard capacitance. A modern, microprocessor-controlled *LCR* bridge operates in the frequency range of 100 Hz–40 MHz and has an accuracy of 0.17%. *See* WHEATSTONE BRIDGE.

Transformer bridge. Transformer bridges (**Fig. 3***a*) are capable of operating up to 100 MHz. The use of transformers (Fig. 3*b*) offers the following advantages: (1) Only two bridge arms are needed: the standard and the unknown arms. Phase opposition between them (necessary for a bridge null) is provided by one of the transformers, driven at a tap on the winding. (2) Both the detector and the source may be grounded at one of their terminals, minimizing ground-loop problems and leakage. The neutral point (Fig. 3) may be regarded as an rf virtual ground

point, and for three-terminal measurements is used for the connection of the coaxial cable outer conductors, because capacitances effectively shunt the transformers and have negligible effect on the balance. *See* TRANSFORMER.

Coaxial line admittance bridge. A coaxial line admittance bridge (**Fig. 4**) is usable from 20 MHz to 1.5 GHz. The currents flowing in three coaxial branch lines are driven from a common junction, and are sampled by three independently rotatable, electrostatically shielded loops, whose outputs are connected in parallel. Two of the branch lines are terminated by conductance and susceptance (inverse of reactance) standards, and the third by the unknown admittance. The currents in the conductance and susceptance arms are in quadrature, and by rotating the loops the correct proportions of these two currents are found to balance the current in the unknown arm. Scales beneath the arms which rotate the loops can be read directly in units of admittance. (The third rotatable loop in the unknown arm provides a common scaling factor to the real and imaginary parts of the admittance.) The attainable accuracy of this instrument is between 3 and 5%, depending on frequency.

Reflection coefficient. A quantity related to impedance is the complex (voltage) reflection coefficient, defined as the ratio of the reflected voltage to the incident voltage, when waves propagate along a uniform transmission line in both directions. Usually, uppercase gamma (Γ) or lowercase rho (ρ) is

$$\Gamma = \frac{Z_T - Z_0}{Z_T + Z_0} \tag{1}$$

$$\text{VSWR} = \frac{1 + |\Gamma|}{1 - |\Gamma|} \tag{2}$$

used to represent the reflection coefficient. When a

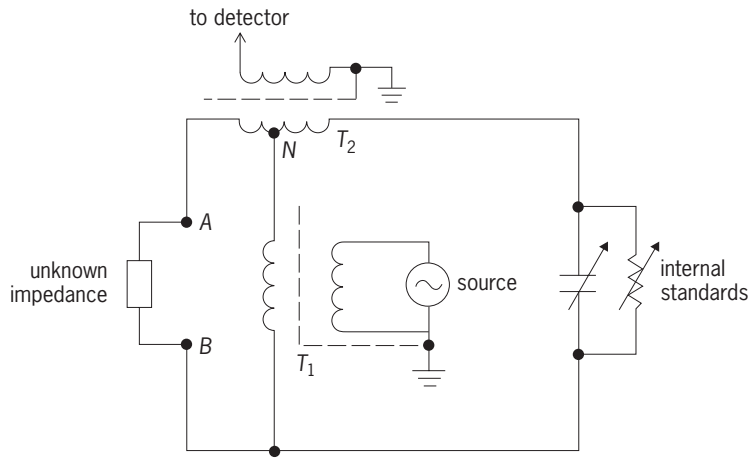

Fig. 3. Basic circuit diagram of transformer bridge usable up to 100 MHz, with transformers T_1 and T_2. Transformer T_2 provides phase opposition between standard and unknown arms. Point *N* (neutral) is for the connection of the ground terminal of three-terminal measurements. Capacitance between *A-N* and *B-N* effectively shunt the transformers.

transmission line of characteristic impedance Z_0 is terminated in impedance Z_T, the reflection coefficient at the load is given by Eq. (1), and the voltage standing-wave ratio (VSWR) is related to the magnitude of Γ by Eq. (2). *See* REFLECTION AND TRANSMISSION COEFFICIENTS; TRANSMISSION LINES.

Resistive bridge. Resistive bridges employed as reflectometers use a matched source and detector, and therefore differ from the Wheatstone bridge (Fig. 1*a*), which aims to use a zero-impedance voltage source and an infinite-impedance detector. It can be shown that these kinds of bridges, from a network-analysis point of view, are equivalent to a directional coupler having a 6-dB coupling coefficient and having a 6-dB attenuator between the source and its input port. When the bridge is driven by a matched generator and uses a matched rf detector (that is,

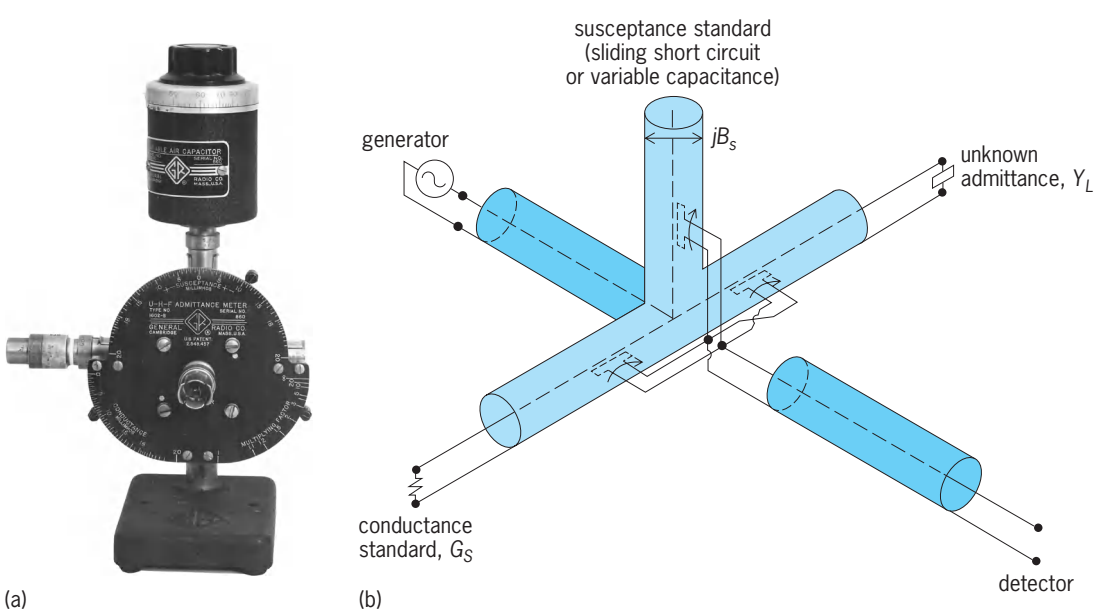

Fig. 4. Admittance bridge. (*a*) Instrument (*GenRad, Inc.*). (*b*) Circuit.

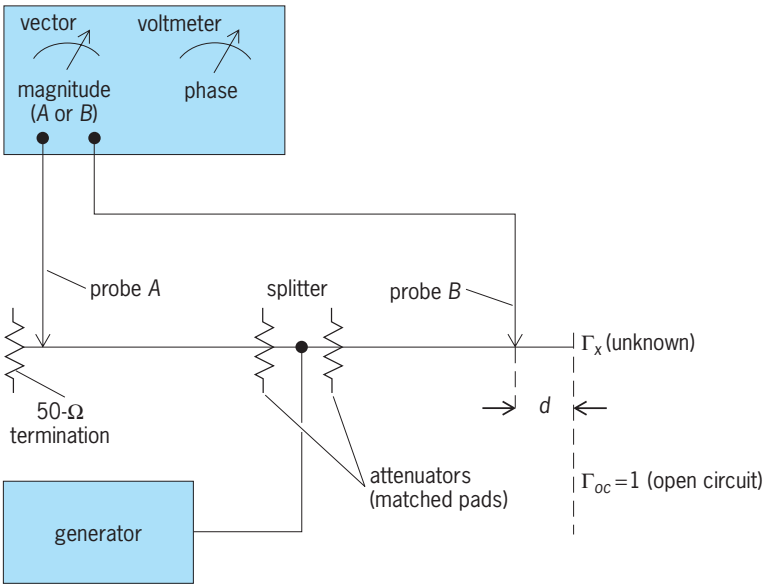

Fig. 5. Possible connection of a vector voltmeter for the measurement of impedance. (*Hewlett-Packard*)

with the impedances of the generator and detector equal to the characteristic impedance Z_0), the output is proportional to Γ. If this output is measured with a voltmeter and with a suitably referenced phase meter (or a vector voltmeter, described below), the complex reflection coefficient, and from it the complex impedance Z_T, of Eq. (1) is found. When it is sufficient to measure only the voltage standing-wave ratio (VSWR), only an amplitude detector is needed. If a square-law detector (power meter or diode detector at low levels) is used, the output is proportional to $|\Gamma|^2$, from which the VSWR may be calculated by using Eq. (2). Such a bridge is first calibrated by connecting a known termination, such as an open or a short circuit (with reflection coefficients $+1$ or -1, respectively). The advantages of the VSWR bridge are low cost and potentially wideband operation. Without adopting special techniques to reduce measurement uncertainty, $|\Gamma|$ may be measured with an uncertainty of ± 0.01. *See* PHASE-ANGLE MEASUREMENT; VOLTMETER.

Voltage-current method. Some specialized electronic instruments make use of the basic definition of impedance, and effectively measure voltage and current. One such instrument is called an rf vector impedance meter. Instead of measuring both the voltage and the current, it drives a constant current into the unknown impedance, and the resultant voltage is measured. With a constant driving current, the rf voltage across the unknown is proportional to the magnitude of its impedance, $|Z|$. The phase difference between the current and the voltage is measured with a discrete phase-measuring circuit. The unknown impedance is thus obtained in magnitude and phase. The instrument linearly downconverts the measuring frequency (0.5 to 108 MHz) to a 5-kHz intermediate frequency, where the amplitude and phase measurements are performed.

Vector voltmeter method. Vector voltmeters (VVM) are instruments with two (high-impedance) voltmeter probes, which display the voltages at either probe (relative to ground) as well as the phase difference between them. One type operates from 1 MHz to 1 GHz, and linearly converts to a 20-kHz intermediate frequency by sampling. This instrument can be connected to measure impedance (**Fig. 5**). The signal is split and sent toward a matched termination (say, 50 Ω) and toward the unknown impedance. The reference probe A senses the voltage in the matched transmission line while, in the unknown arm, the other probe B measures the voltage at a position equidistant from the junction. Attenuators may be inserted in both arms to minimize interaction. The measurement is normalized by connecting a (shielded) open circuit in the plane of the unknown, and the complex voltage at probe B is noted as V_{oc}. The unknown is connected and the indication of probe B is noted as V_x. Then the unknown reflection coefficient is calculated from Eq. (3), where ϕ

$$\Gamma_x = e^{j2\phi}\left[\frac{V_x}{V_{oc}}(1 + e^{-j2\phi}) - 1\right] \tag{3}$$

is the electrical angle between probe B and the unknown (equal to $2\pi d/\lambda$, where d is the distance between probe B and the unknown and λ is the wavelength). Because the substitution is on the same arm, the lack of symmetry of the splitter, the lack of equality of the attenuators, and the lack of equality of the two voltmeter channels A and B do not cause errors.

Since the probe responds to the sum of the incident and reflected waves, the resolution becomes poorer as the reflection coefficient magnitude approaches zero. This method is therefore not suitable for measuring impedances near the characteristic impedance of the transmission line (50 Ω in this case). For the accurate measurement of such impedances, directional couplers are used. *See* DIRECTIONAL COUPLER.

At radio frequencies the realization of directional couplers is quite different from microwave realizations because the component is much shorter than the wavelengths encountered. For this reason, lumped-element directional couplers may be constructed. Equal electric and magnetic coupling results in directional behavior so that the coupled output responds (ideally) only to signals traveling in one direction in the main line. The electric and magnetic couplings may be provided by voltage and current transformers, whose signals are summed. Distributed-circuit directional couplers may also be constructed by twisting a pair of insulated copper wires together to provide mainly capacitive coupling between them. Then the twisted pair is formed into a coil to provide additional inductive coupling. By stretching the coil, the magnetic coupling may be varied, and adjusted to equal the capacitive coupling. Such directional couplers have been constructed from 1 to 300 MHz and, if equipped with small trimmer capacitors near the terminals, can achieve directivities in excess of 50 dB at selected frequencies.

A coaxial-line, rf impedance-measuring circuit can be built using a vector voltmeter and directional couplers. One probe serves as a phase reference only. The other probe samples the reflected signal, and therefore the vector voltmeter indication is proportional to the reflection coefficient. The readings are normalized by first connecting a reflection standard, which may be a short or an open circuit, or a mismatched resistor. In measuring small reflection coefficients (impedances near 50 Ω), the sensitivity of the vector voltmeter may be increased in 10-dB steps, offering high resolution. If the analog magnitude and phase outputs are converted into their rectangular form, these voltages may be connected to the X and Y deflection inputs of an oscilloscope to give a display of the impedance on the circle diagram (Smith chart). By controlling the frequency and the calibration sequence with a computer, the VVM-coupler method may be automated, and this forms the basis of automatic network analyzers (ANA). *See* MICROWAVE IMPEDANCE MEASUREMENT.

Resonance method. When the magnitude of the reactive part of the impedance is much greater than the resistive part at a given frequency f, resonance methods may be employed to measure impedance. The most commonly used instrument for this purpose is the Q meter. The Q (quality-factor) of an impedance is defined as 2π times the ratio of energy stored to energy lost per cycle. Numerically, Q is equal to $\omega L/R$ or $1/R\omega C$ for the series combination of a resistor R and inductor L or a resistor and capacitor C respectively, where $\omega = 2\pi f$.

A Q-meter-type measurement requires a radio-frequency source of frequency f, which drives a current through a current-limiting injection resistor R_i into the Q-meter circuit (**Fig. 6**). The ratio of a magnitude of the voltage E developed across the calibrated capacitor C to a voltage e across the combination of capacitor, inductor, and resistor is indicated with a high-impedance rf voltmeter, and has a maximum value of Q_{ind} when C is set to maximize this reading.

To measure low-loss capacitors, the circuit is first resonated (with a suitably chosen inductor of high Q) to find Q_{ind}. The capacitor setting at resonance is noted. Then the unknown capacitor is paralleled with C, and the reduction in C needed to restore resonance is the value of C_x.

A low-loss inductor is measured by resonating it with C and substituting into Eq. (4). Since the magnitude of E/e is given by Eq. (5), the Q factor of the inductor is given by Eq. (6), which, for high-Q circuits, approximates to $Q = Q_{\text{ind}}$. The series resistance is given by Eq. (7), and the inductance is given by Eq. (8), which, for $Q_{\text{ind}} \gg 1$ reduces to Eq. (4).

$$L_x \simeq \frac{1}{\omega^2 C} \tag{4}$$

$$\left|\frac{E}{e}\right| = [(\omega^2 L_x) + (\omega R_x C)^2]^{-1/2} \tag{5}$$

$$Q = \left(Q_{\text{ind}}^2 - 1\right)^{1/2} \tag{6}$$

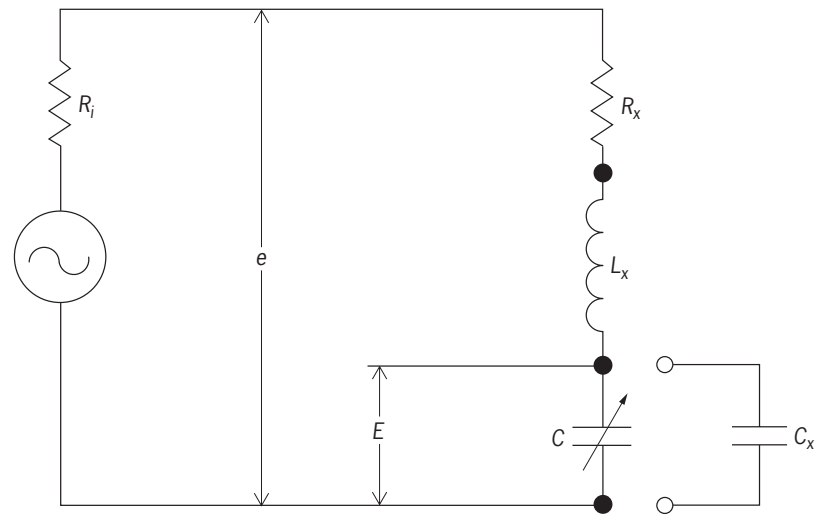

Fig. 6. Circuit of a Q meter.

$$R_x = \frac{\left(Q_{\text{ind}}^2 - 1\right)^{1/2}}{Q_{\text{ind}}^2 \omega C} \tag{7}$$

$$L_x = \frac{Q_{\text{ind}}^2 - 1}{Q_{\text{ind}}^2 \omega^2 C} \tag{8}$$

See Q METER.

Slotted lines and six-port methods. At the upper end of the rf range, microwave methods of impedance measurement may also be used, employing slotted lines and six-port junctions. For completeness, these methods will be briefly discussed.

A slotted line is a transmission line, with an axial slot cut in the outer conductor, permitting a traveling probe to sample the electric field. When an unknown impedance is connected to one end and a generator to the other, the ratio of the maximum to minimum signals as the probe is sliding is the voltage standing-wave ratio, and is related to the reflection coefficient by Eq. (2). The phase of Γ is found from Eq. (9), relating the positions of the minima when the

$$\phi_x = \pi\left[1 - \frac{4(d_{sc} - d_x)}{\lambda}\right] \tag{9}$$

unknown is connected and when a short circuit is connected, where d_{sc} and d_x are the probe positions for the minima for the short circuit and the unknown connected, and λ is the wavelength.

According to the theory of six-port junctions, an unknown impedance may be uniquely determined from four power-meter readings (if the six-port junction is such that the power meters are sufficiently independent of each other). If diode detectors sample the power along a coaxial line at intervals of λ/8, then the differences between the outputs of alternate pairs of detectors yield a Smith chart display directly, and the outputs of three of the four detectors together with a sample of the incident power can be used for six-port operations. Six-port junctions

are calibrated by connecting a number of impedance standards in place of the unknown.

Automatic network analyzers. The most modern impedance measuring units are classed as automatic network analyzers, which may be scalar automatic network analyzers (SANA), measuring insertion loss, gain, return loss, voltage standing-wave ratio, and power; and vector automatic network analyzers (VANA), which, in addition, measure vectorial quantities, like the complex *S*-parameters of networks.

The *S*-parameters (scattering parameters) are defined by Eqs. (10), where a_1 and a_2 are the incident

$$b_1 = S_{11}a_1 + S_{12}a_2 \qquad (10a)$$

$$b_2 = S_{21}a_1 + S_{22}a_2 \qquad (10b)$$

signals on ports 1 and 2 of a two-port network, and b_1 and b_2 are the emerging signals of these ports, respectively. The parameters S_{ij} represent signal reflection coefficients, and S_{ij} represent signal transmission coefficients. These quantities are measured by using internal directional couplers which sample the incident and reflected signals in the two measuring ports (between which the network to be analyzed is connected) and, after internal linear frequency down-conversion using mixers, the voltage amplitudes, their ratios, and phase differences are measured with precision low-frequency voltmeters and phase meters. The resulting quantities can be displayed (in the frequency domain, that is, as functions of frequency) on a linear or logarithmic scale versus frequency, or on the Smith chart.

Vector automatic network analyzers are available in numerous frequency bands, from 5 Hz to above 100 GHz. Because of the broad frequency coverage of these instruments, the internal directional couplers used cannot be of sufficiently high directivity, and their use would result in unacceptably large errors in the measurements. This difficulty is overcome by digital error correction. The automatic network analyzer is first calibrated by connecting to it the necessary number of impedance standards, and then the instrument performs the necessary measurements, stores the results, and computes the error-correcting terms at all the measured frequencies. After calibration, the automatic network analyzer is ready for use, and the displayed results are digitally corrected at all the measured frequencies. As an option, voltage automatic network analyzers can be equipped with a time-domain facility which, after digital Fourier transformation of the time-domain results, plots the results as a function of time (that is, distance along the component tested). This permits finding the locations of troublesome reflections along transmission-line components, coaxial lines, or waveguides. *See* FOURIER SERIES AND TRANSFORMS. Peter I. Somlo

Bibliography. A. C. Lynch and A. E. Bailey (eds.), *Radio Frequency Bridges*, 1988; P. I. Somlo and J. D. Hunter, *Microwave Impedance Measurement*, 1985; M. Sucher and J. Fox, *Handbook of Microwave Measurements*, 1963; M. Valkenburg (ed.), *Reference Data for Engineers: Radio, Electronics, Computers, and Communications*, 8th ed., 1993.

Radio paging systems

Systems consisting of three basic elements—a personal paging receiver, radio transmitter, and an encoding device—whose primary purpose is to alert an individual, or group of individuals, and deliver a short message, typically of a temporary or perishable nature. Characteristics that are used to define a specific paging system include distance covered, radio frequency, modulation type, paging code format, and message type.

Every country has a radio regulatory agency, such as the Federal Communications Commission (FCC) in the United States, which regulates radio spectrum usage and issues licenses for the operation of radio transmitters. Most licenses for radio paging systems limit the number of transmitters and power level that a system is allowed to use. These restrictions limit the distance over which a paging signal may be transmitted. On-site systems cover a single building or a small complex of buildings typically utilizing one low-power transmitter. Wide-area systems can cover an entire city or country and usually use multiple transmitters which simulcast the paging signals.

Nationwide or international paging can be accomplished in several ways. Many countries have allocated specific radio spectrum to provide multicity nationwide paging on a single radio channel. Systems using one of these channels either will simulcast all their transmitters throughout the nation or, at the paging user's direction, will route a paging signal to a specifically requested region or city. This allows users of these systems to travel throughout the country and still receive their messages. Paging in Europe is standardized on a system named ERMES. Users traveling throughout Europe can have their paging messages sent to the country or region to which they are traveling. This protocol allows for the use of up to 16 designated radio channels, and an ERMES paging receiver scans all of these channels, looking for its unique pager address code.

Independent nationwide paging system operators who use the same radio channel in their respective countries have coordinated the assignment of their codes and have connected their paging systems to one another. This has allowed the paging users of these systems to travel to other countries and carry their "hometown" personal pager with them to receive their messages.

Most paging systems now utilize the very high-frequency (vhf) or ultrahigh-frequency (uhf) radio spectrum using frequency modulation (FM). There are systems which use a special subcarrier on an FM broadcast transmitter, and others which use the citizens-band (CB) channels. The latter is often used for "auto-theft" pagers and small on-site systems using amplitude modulation (AM). *See* AMPLITUDE MODULATION; FREQUENCY MODULATION; RADIO SPECTRUM ALLOCATIONS.

Paging formats. Each individual paging receiver is alerted by transmitting its unique address code or codes. Many paging formats are available, and they fall into two categories—analog and binary digital.

Analog codes use audible, or sometimes subaudible, tones. The most common analog formats are two-tone sequential and five/six-tone decimal digital. Two-tone typically provides for several thousand individual codes, while five/six-tone allows 1 million unique codes and is used in many wide-area systems. Binary digital paging formats are composed of binary patterns and can support tens of millions of codes. Moreover, binary digital codes are particularly well suited for encoding numeric and alphanumeric messages.

An additional characteristic of paging formats is the speed at which they transmit the paging code and message. The most common digital format of the 1980s initially was transmitted at 600 bits per second. It was modified to 1200 bps, and there are a few implementations at 2400 bps. The increased number of paging users and the longer messages that they receive have given rise to a family of high-speed, flexible paging codes. One such code enables paging systems to operate at 1600, 3200, or 6400 bps, thus allowing system operators to adapt their system as needs arise. The system uses receivers that can dynamically adjust to any of the speeds.

To complement this code, and to allow a paging user to respond to a message, a two-way paging format has been implemented. This enables the operation of a class of advanced paging services which the FCC has designated narrow-band personal communication services (NBPCS).

Paging receivers. Paging receivers fall into four basic categories: tone alert, tone and voice, numeric, and alphanumeric. Tone pagers emit a beep when they are signaled. Some models silently alert the user with a vibration in place of a beep; other models use differing staccato beeps to provide the user with several alert messages. Tone and voice pagers allow the initiator of the page to transmit a simple voice message which will follow a pager's beep alert. Numeric pagers, sometimes called digital pagers, allow the initiator to convey numerical information. These messages are typically composed by using a tone telephone key pad. Alphanumeric pagers (**Fig. 1**) allow the initiator to send a complete textual message to the pager user. These messages are composed on word processors, personal computers, or dedicated terminals which can connect to a paging terminal. *See* MICROCOMPUTER; WORD PROCESSING.

Fig. 1. Alphanumeric personal message pager. (*Comfort Net*)

Fig. 2. Card pager. (*Motorola Inc.*)

In addition to the basic pagers, there are advanced devices. The pager card (**Fig. 2**) is a pager designed in a standard computer card format. This allows the card to be inserted into a computer or personal digital assistant standard accessory slot, enabling these devices to wirelessly receive updated information or messages. The two-way pager includes an integral transmitter, allowing the user to respond to a received message. This helps eliminate the need for a paging user to find a telephone to respond to a page. A class of alphanumeric pagers allows the receipt and management of selective broadcast information. Services such a news headlines, financial information, and sports scores can be sent to a large number of interested users. These information pagers can be combined with standard paging functions, or can be manufactured as stand-alone information appliances.

Encoders. Every paging system contains an encoding device which provides the interface between the page initiator and the pager user. The encoder's basic function is to receive a pager number and message from the initiator and convert this information into the correct format to be transmitted to the paging receiver. The output from the encoder is connected directly to the system's radio transmitter. Encoders range from simple manual desktop models to large telephone interconnected central offices. Manual encoders are used to initiate pages from a single location such as a dispatcher or switchboard operator. In many large on-site or area-wide systems, it is desirable to initiate pages from the standard telephone network. These systems use an interconnected automatic paging terminal. In addition to generating the

correct pager code and message format, terminals develop page/message queues, store system usage statistics, and even store short voice messages in solid-state recorders. *See* MOBILE RADIO; TELEPHONE SERVICE.

James A. Wright

Bibliography. D. Baker, *Comprehensive Guide to Paging*, 1992; C. A. Bean, *Overview of the U.S. Radio Paging Industry*, 1988; N. J. Boucher, *The Paging Technology Handbook*, 1998; W. C. Y. Lee, *Mobile Communications Design Fundamentals*, 2d ed., 1993; Telocator, *The Future of Paging II*, 1993.

Radio receiver

The part of the radio communications system that extracts information from radio-frequency (rf) energy intercepted by the antenna. Radio receivers are the most common electronic equipment worldwide and a vital part of all radio, television, and radar systems. Since the 1960s, radio receiver performance has improved greatly, while size, weight, and cost have fallen dramatically. In the past, radio receivers were built from analog circuits, but increasingly they are realized by digital signal processing. *See* RADIO; SIGNAL PROCESSING.

The antenna intercepts a band of energy in the radio frequencies containing many transmissions. These may have different modes of modulation; the two most common are amplitude modulation (AM) and frequency modulation (FM). Signals have a large size range, from a large fraction of a volt down to a small fraction of a microvolt. The receiver must be selective, responding to only one signal, must demodulate the signal, extracting the impressed information from the radio-frequency wave, and must raise it to an acceptable power level by amplification.

Amplitude-modulation and frequency-modulation receivers will be described because they are the most common, but other modes of modulation are also used. Single-sideband (SSB) transmissions are similar to amplitude modulation, but with one of the symmetrical pair of sidebands eliminated and the carrier suppressed. Single-sideband is significant because it conserves electromagnetic spectrum, introducing less spectrum pollution than any other modulation. Because of the growing prevalence of digital signals, digital modulation systems are increasingly important. In the most common of these, known as binary frequency shift keying (FSK), the 0 and 1 states of the binary sequence transmitted are represented by two frequencies. Binary frequency shift keying can thus be regarded as a digital variant of frequency modulation. Systems of digital modulation that combine a form of amplitude modulation (called envelope shaping) with frequency modulation are becoming more common, the purpose again being to reduce spectrum pollution. *See* AMPLITUDE MODULATION; FREQUENCY MODULATION; MODULATION; SINGLE SIDEBAND.

Amplitude-modulation architectures. A simple receiver for amplitude-modulated signals is shown in **Fig. 1**. A band-pass filter selects the required signal, which, after optional amplification, is passed to the demodulator (obsolete term: detector), which in this version consists of a limiting amplifier and a multiplier. The high-gain limiting amplifier has an output of the same sign as the input but constant magnitude (a square wave). When multiplied by the amplitude-modulated waveform, this inverts negative half-cycles. After low-pass filtering to remove residual radio-frequency ripple, the modulating waveform is obtained, shown as an audio waveform, which passes to the loudspeaker. There are other amplitude demodulator circuits, all of which either suppress or invert alternate half-cycles of the amplitude-modulated waveform. *See* AMPLIFIER; AMPLITUDE-MODULATION DETECTOR; ELECTRIC FILTER.

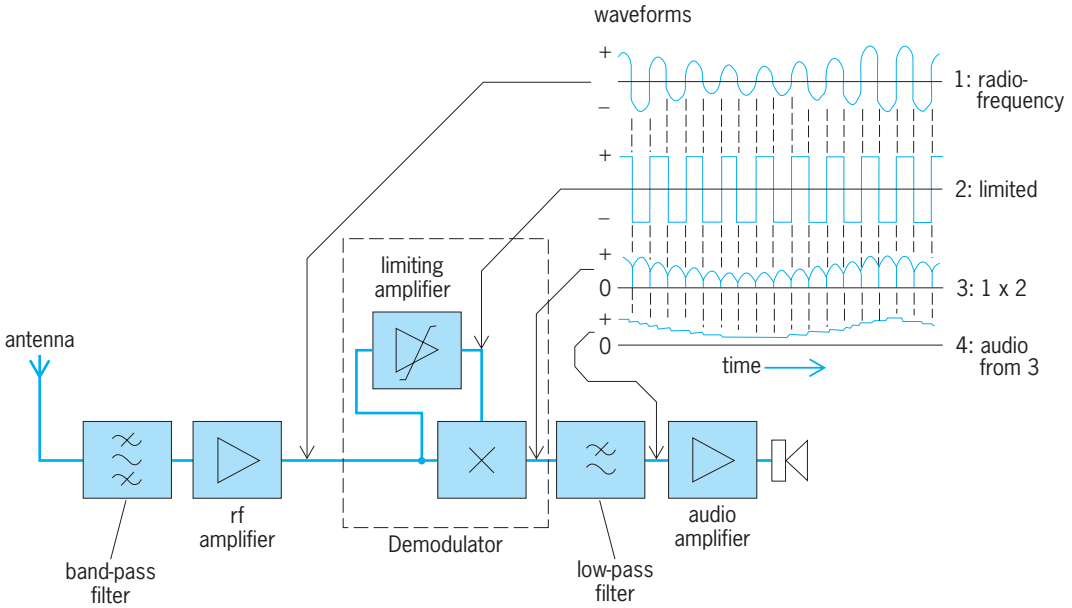

Fig. 1. Simple tuned-radio-frequency (TRF) receiver for amplitude-modulated (AM) signals.

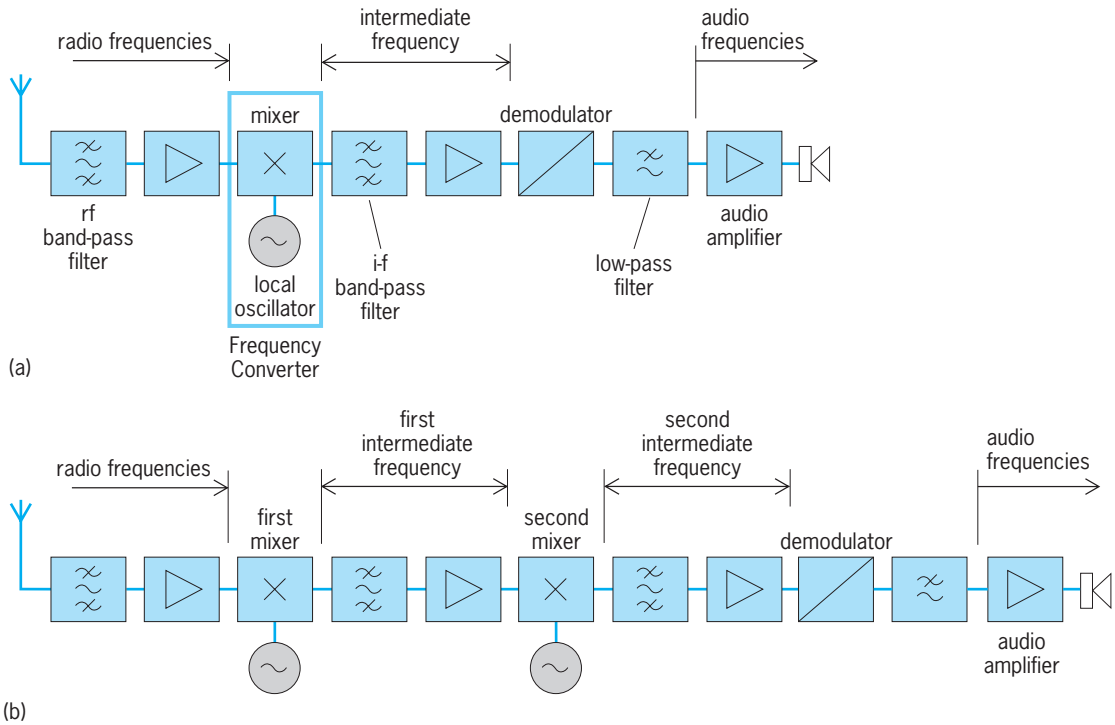

Fig. 2. Superheterodyne (superhet) receivers. (*a*) Basic architecture. (*b*) Double superheterodyne, used to combat image response problems.

Since the input signal may be on the order of microvolts whereas volts are required to operate a loudspeaker, the total receiver gain may range up to the order of a million or more. In this simple tuned-radio-frequency (TRF) receiver it is difficult to adjust the center frequency of a high-performance band-pass radio-frequency filter in order to receive other stations; also, the radio frequency may be too high for effective amplification. For this reason the superheterodyne (superhet) receiver (**Fig. 2***a*) superseded it.

Superheterodyne receivers. After a radio-frequency filter and amplifier, the signal passes to a mixer, from which the output is formed by the product of the signal and a locally generated wave. If the signal carrier is at f_c and the local oscillator is at f_o, the result is a wave at a frequency $\pm(f_o - f_c) = f_{\text{i-f}}$, known as the intermediate frequency (i-f), and this difference frequency can be kept constant, whatever the value of f_c, by a suitable choice of f_o. Subsequent to the mixer, the superheterodyne receiver becomes similar to the tuned-radio-frequency receiver except that it is now operating at a fixed frequency, so that, provided a low intermediate frequency is chosen, amplification is easy and high-performance filters can be used. Fixed-frequency intermediate-frequency filters, based on mechanical resonance in piezoelectric ceramics and crystals or on surface-acoustic waves, give extreme attenuation outside the passband. The superheterodyne architecture may be used with amplitude modulation, frequency modulation, and other modulation types. *See* HETERODYNE PRINCIPLE; MIXER; OSCILLATOR; PIEZOELECTRICITY; SURFACE-ACOUSTIC-WAVE DEVICES.

Assuming that a low value of $f_{\text{i-f}}$ is chosen, a problem arises because of image response. If the carrier is below the oscillator frequency at $f_c = (f_o - f_{\text{i-f}})$, it follows that a frequency $f_s = (f_o + f_{\text{i-f}}) = (f_c + 2f_{\text{i-f}})$ will give the same result, since the sign of the difference is not significant, and the receiver will therefore have a spurious response at f_s, known as the image frequency, separated from the wanted carrier by $2f_{\text{i-f}}$. It is this image response that the first radio-frequency filter must suppress. This is easier the larger the value of $f_{\text{i-f}}$, in conflict with the requirement for a low value explained above. Usually a compromise value of $f_{\text{i-f}}$ is chosen, but in exacting cases the receiver architecture may be changed to a double superheterodyne (Fig. 2*b*), using a relatively high first intermediate frequency which is followed by a second conversion to a lower intermediate frequency, thus gaining the advantages of both at the cost of greater complexity.

Automatic gain control. In amplitude-modulation systems, the amplitude of the signal carries the information, and so to keep the audio output from wide level changes it is desirable to adjust the strength of the signal at the demodulator to be roughly constant, despite variations at the antenna due to transmitter power or range. Thus, automatic gain control (AGC) is used in most amplitude-modulation receivers. The output from the demodulator splits, and, as well as going forward to the audio stages, passes to a low-pass filter which cuts off at a few hertz, well below the audio frequencies. The output from this filter is a voltage proportional to the mean amplitude of the amplitude-modulated wave. After amplification, it controls the gain of the amplifiers. If the amplitude-modulated signal rises, and hence the

control voltage also rises, the gain is cut back, tending to restore the original demodulator output level. *See* AUTOMATIC GAIN CONTROL (AGC).

Single-sideband receivers. Receivers for single-sideband signals are similar to amplitude-modulation receivers, except for the demodulator, a frequency converter in which the single-sideband signal at a nominal carrier frequency converted to f_{i-f} is mixed with a local oscillator also at f_{i-f}, thus translating the carrier frequency to zero and the sidebands to their correct audio frequencies. Because there is no constant carrier component that can be used as an amplitude reference, automatic gain control presents problems in single-sideband receivers.

Frequency-modulation architectures. For frequency modulation, the signal is of constant amplitude but varies in frequency, and so it is usual for the intermediate-frequency amplifier to be a limiter, giving a constant output. In other respects the receiver may be identical with one of those in Fig. 2. Automatic gain control is often omitted, but may be combined with limiting for the highest performance. The frequency demodulator is quite different, often utilizing a phase-sensitive detector (PSD). *See* LIMITER CIRCUIT.

The phase-sensitive detector (**Fig. 3*a***) consists essentially of a multiplier and a low-pass filter, shown in Fig. 3*b* working with square-wave inputs, from limiters. When the two waves are in phase the output from the filter is positive (Fig. 3*c*), but as they shift in phase it declines, reaching zero when the waves are at 90°, then going to a negative peak when they are at 180°, returning to zero at 270°, a positive peak again at 360°, and so on.

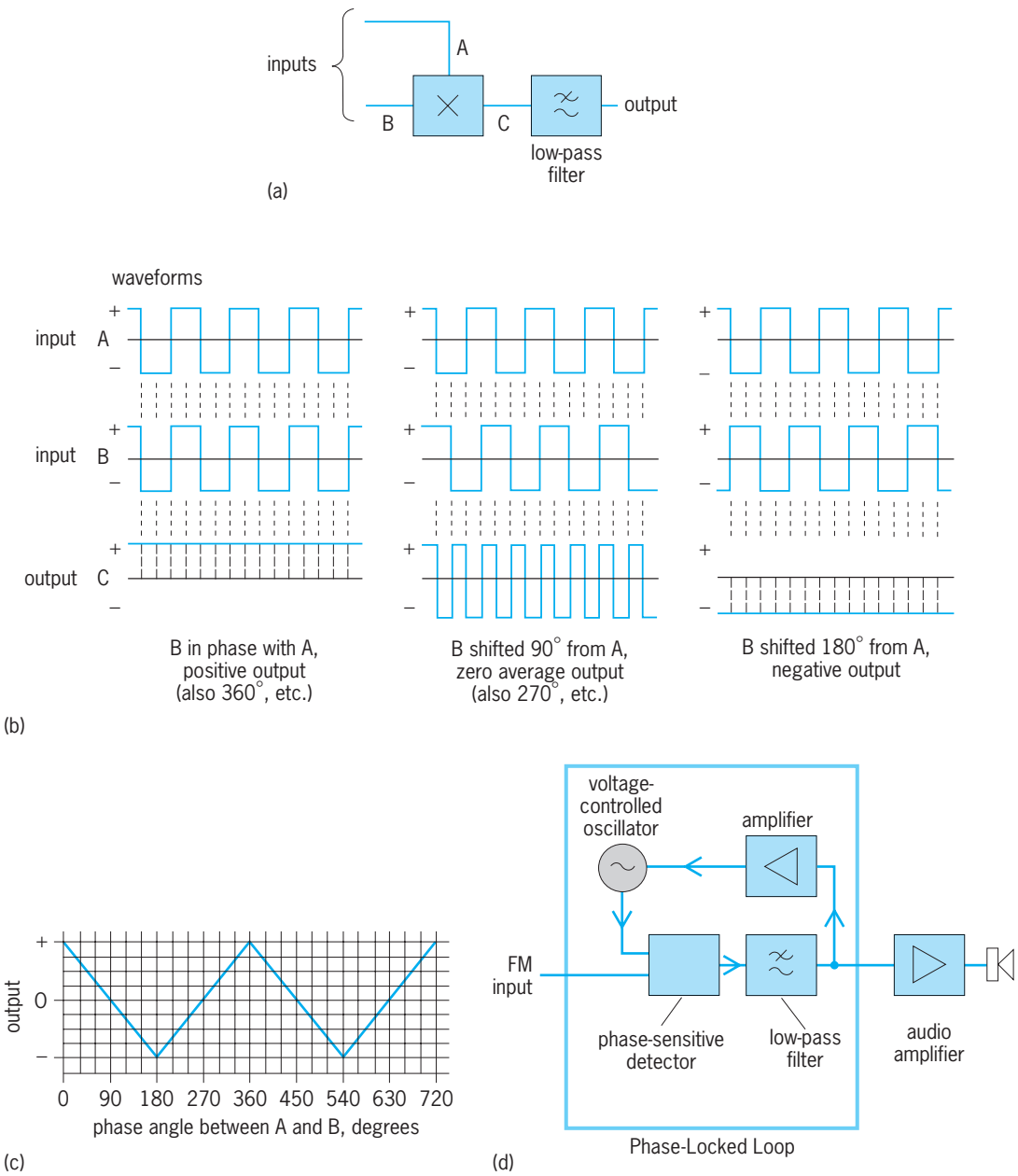

Fig. 3. Phase-sensitive detector (PSD). (*a*) Diagram. (*b*) Waveforms of inputs and multiplier output. (*c*) Output from low-pass filter versus phase angle from inputs A and B. (*d*) Frequency demodulator using a phase-locked loop (PLL) with a phase-sensitive detector.

A sophisticated frequency demodulator uses the phase-sensitive detector to create a phase-locked loop (PLL; Fig. 3*d*). An important component is a voltage-controlled oscillator (VCO). Its frequency can be controlled over a range larger than the excursion of the frequency-modulated wave by an externally applied voltage. This can be achieved on a silicon chip in many ways, such as by charging a frequency-determining capacitor from a voltage-controlled current source.

Since the frequency of a wave is the rate of change of its phase, a small frequency error between the voltage-controlled oscillator and the frequency-modulated waveform may be expressed as a phase error growing larger with time, and thus produces an increasing output from the phase-sensitive detector. This output is amplified and applied to the voltage-controlled oscillator, where it pushes the frequency toward the correct value if it is of the right polarity. (If the polarity is at first wrong, the phase simply slips into the next half-cycle, where the sense of variation of output with phase is reversed.) The amplifier gain is large, and so for an output in the working range its input must be very small. The phase of the voltage-controlled oscillator must lock to a value very close to 90° (or a multiple) from that of the frequency-modulated waveform, since only then is the output of the phase-sensitive detector minimized.

Once the loop has phase-locked, the frequency of the voltage-controlled oscillator faithfully follows that of the frequency-modulated wave, and so its frequency-control voltage represents the frequency deviation of the frequency modulation. This is the basis of the phase-locked-loop frequency demodulator. *See* FREQUENCY-MODULATION DETECTOR; PHASE-LOCKED LOOPS.

A radio for receiving frequency-modulation broadcasts may generate a stereo audio signal pair, which consists of left and right audio signals. In such receivers, the frequency demodulator produces a multiplex signal which contains both sum and difference signals (**Fig. 4**). The sum and difference signals are further processed to create left and right audio signals, which are then supplied to a pair of stereo speakers. To facilitate extraction of the difference signal from the multiplex signal, the receiver uses a low-level pilot signal also present in the multiplex signal. The difference signal is located within the multiplex signal as a double-sideband signal.

The process of extracting the sum and difference signals and the creation of the left and right audio signals is shown in **Fig. 5**. The 19-kHz pilot signal is doubled to create a 38-kHz signal, which is used to frequency-shift the difference signal to its normal audio position at 0 Hz. This operation uses a mixer similar to the mixer shown in Fig. 1. The sum and difference signals are isolated in filters and then added and subtracted to produce left and right audio signals. *See* STEREOPHONIC RADIO BROADCASTING.

The multiplex signal may contain a data signal. In the United States, this signal allows reception of information from the source of the radio broadcast, in accordance with the Radio Data Broadcasting System (RDBS) standard (Fig. 4). The information may include identification of the broadcaster. In Europe, a similar system operating according to the Radio Data System (RDS) standard allows reception of information on a network of broadcasters providing the same audio material, allowing the receiver to automatically select the broadcaster providing the signal with the least reception impairments.

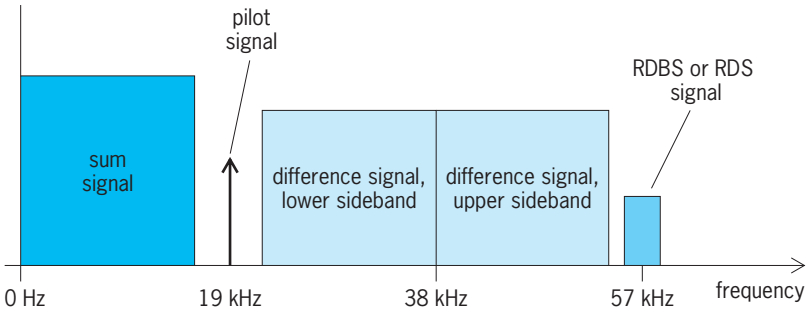

Fig. 4. Frequency spectrum of frequency demodulator output in a radio receiver that generates stereo audio signals, showing content of the multiplex signal.

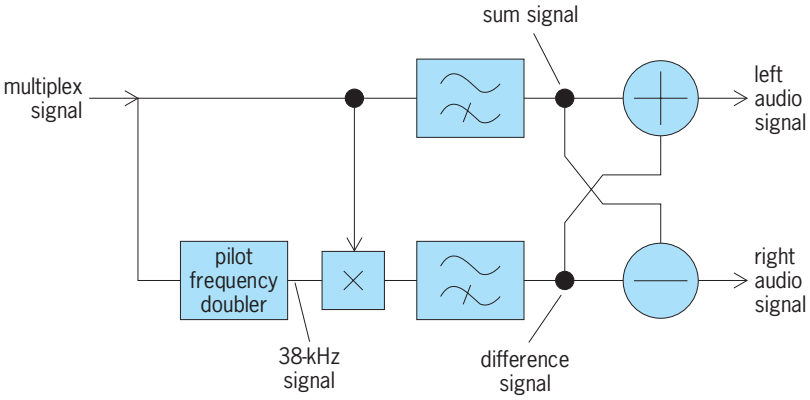

Fig. 5. Process to extract left and right audio signals from a multiplex signal.

Typical consumer radio receivers provide reception of both amplitude- and frequency-modulated signals. Such receivers employ common amplifier, filter, and mixer circuits that operate for receiving either type of signal, depending on which is selected by the listener. The use of common circuits allows a lower cost to manufacture the radio.

Homodyne receivers. The image frequency response in a superheterodyne receiver occurs at a frequency $2f_{i\text{–}f}$ from the wanted signal. If $f_{i\text{–}f}$ is reduced to nearly zero, the image and wanted responses coalesce. This is the basis of homodyne, or zero-intermediate-frequency, receivers (**Fig. 6**). The intermediate frequency (corresponding to the signal carrier wave) is made zero by setting the local oscillator to the carrier frequency. Sidebands, whether higher or lower, are converted to low frequencies, and all adjacent channels come out higher. The desired signal may be selected by low-pass filters, which are easier to realize than band-pass filters.

Frequencies both above and below the carrier are converted into the same range (since the sign of the frequency difference is immaterial). With a symmetrical spectrum signal, such as in amplitude modulation, this would not matter if the converted lower sidebands fell exactly on the converted upper

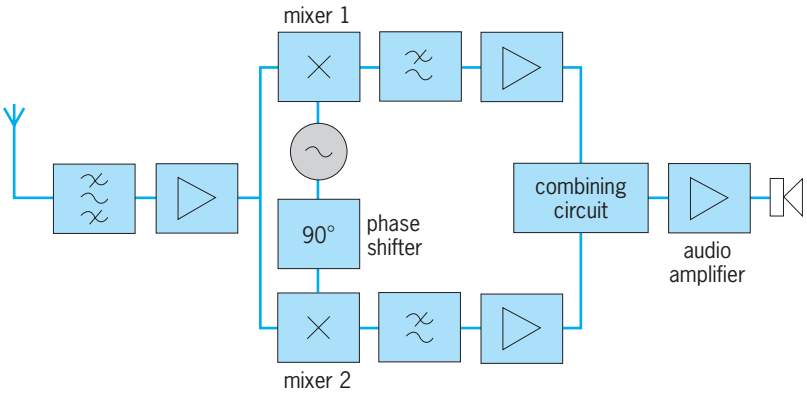

Fig. 6. Homodyne (zero-intermediate-frequency) receiver.

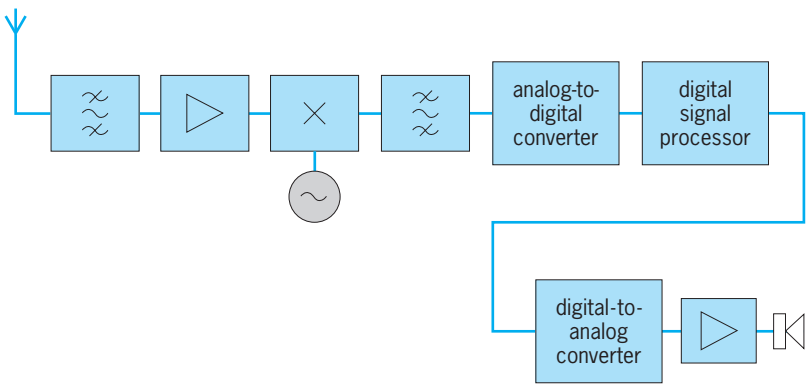

Fig. 7. Superheterodyne receiver with digital signal processing.

sidebands, but this could happen only if the frequency of the local oscillator were exactly locked to that of the incoming carrier, which is difficult because the incoming carrier is so weak. In any event, the method would not work for asymmetrical spectrum signals such as single sideband.

A better alternative is to use two frequency converters with oscillator inputs of the same frequency $f_o = f_c$ but at 90° phase difference (Fig. 6), which makes it possible to separate the upper and lower sideband products. If the oscillator that feeds one of the mixers is in phase with the signal carrier, then, whether the signal frequency is above or below the carrier, the output will be the difference frequency, without phase shift. However, if the second mixer is 90° ($\pi/2$ radians) phase-advanced relative to the carrier, its output will be phase-advanced by 90° ($\pi/2$) for frequencies above the carrier and phase-retarded by 90° ($-\pi/2$) for frequencies below, since in the latter case the sign of the frequency difference $f - f_c$ is negative, and hence opposite to that of the phase angle. The output phases of the two carriers are summarized in the **table**.

This difference in phase relationships between the two outputs makes it possible to distinguish positive and negative frequency excursions, and demodulation is facilitated. For example, in the case of a frequency-modulated signal, the output from mixer 1 is given by Eq. (1) and its derivative by Eq. (2), where E is the maximum output. The output from mixer 2 is, similarly, given by Eqs. (3). Multiplication of the

$$e_1 = E\cos\left[2\pi(f - f_c)t\right] \tag{1}$$

$$\frac{de_1}{dt} = -2\pi(f - f_c)E\sin\left[2\pi(f - f_c)t\right] \tag{2}$$

$$e_2 = E\cos\left[2\pi(f - f_c)t + \frac{\pi}{2}\right] \tag{3a}$$

$$\text{for } (f - f_c) > 0$$

$$e_2 = E\cos\left[2\pi(f - f_c)t - \frac{\pi}{2}\right] \tag{3b}$$

$$\text{for } (f - f_c) < 0$$

differentiated output of mixer 1 and the output of mixer 2, after low-pass filtering to remove radio-frequency components, results in a voltage proportional in sign and magnitude to $f - f_c$, forming a frequency demodulator. Different combining circuits can demodulate amplitude, single-sideband, and other modulations.

The homodyne configuration is well suited to digital signal processing in place of analog circuits. By inserting analog-to-digital converters in the two intermediate-frequency channels, the subsequent filtering, processing, and combining can be performed by means of digital calculations, using a digital processing chip. After combining in this way the digital signals are reconverted to analog form for the audio circuits. Receivers of this kind are inexpensive to fabricate and are increasingly commonplace, as in cellular telephones and consumer broadcast receivers. Superheterodyne receivers may also employ digital signal processing, where the output of the analog superhet mixer and filter intermediate-frequency circuitry is converted to digital form and mixed, filtered, and demodulated by digital calculations. *See* ANALOG-TO-DIGITAL CONVERTER; DIGITAL-TO-ANALOG CONVERTER.

A typical arrangement of a superheterodyne receiver with digital signal processing is shown in **Fig. 7**. A significant number of the functions of the receiver is processed in the digital domain, including filtering, amplifier, automatic gain control, mixing, and demodulation, and this processing may be performed in one or more digital processing chips. A key characteristic of most digital signal processing circuits is programmability. For a receiver, the digital signal processor can be programmed to process either amplitude- or frequency-modulated signals. Within the digital signal processor, a homodyne structure may additionally be employed. Processing may also be included to extract the RDBS or RDS signals.

Output phases of mixers in a zero-intermediate-frequency receiver

	Output phase	
	Mixer 1	Mixer 2
$f > f_c$	0	+90° or $\pi/2$ radians
$f < f_c$	0	−90° or $-\pi/2$ radians

Phase-locked-loop applications. The phase-locked loop has important applications beside the frequency demodulator described above. One is for tracking receivers, which are able to follow a signal drifting in carrier frequency. These are used, for example, in space vehicles, where, because of the high velocity, there is a considerable Doppler shift of received frequencies. *See* DOPPLER EFFECT.

To overcome this problem, the receiver (frequency modulation or amplitude modulation) uses a phase-locked loop feeding back to the frequency converter. In this so-called long-loop receiver, this feedback pulls the frequency of the first oscillator so as to center the wanted signal at $f_{i\text{-}f}$. The intermediate-frequency filter can then be made just wide enough to accommodate the signal without adding extra width to accommodate frequency drift, so that interfering signals or noise is reduced.

A more important application of phase-locked loops is in frequency synthesizers. Since the intermediate frequency $f_{i\text{-}f}$ is fixed in a superheterodyne receiver, the value of f_o determines the signal f_c that is received, so the frequency stability of the local oscillator is critical. Sometimes receivers use quartz crystals as piezoelectric mechanical resonators to determine oscillator frequency directly [they can give better than 20 parts per million (ppm) repeatability without difficulty], but since each crystal is ground for a single frequency, many crystals are needed to receive different transmissions. A better solution is to use one crystal as a high-accuracy clock (as in a quartz watch), deriving all required frequencies from it by using a phase-locked loop. *See* QUARTZ CLOCK.

In such a frequency synthesizer (**Fig. 8**), a crystal oscillator generates a stable frequency f_Q, which passes to a counter circuit (easily realized in digital electronics), where it is counted down to another frequency, say f_Q/n. Similarly, a voltage-controlled oscillator (part of a phase-locked loop) is counted down by m to f_o/m, in this case by a programmable counter, so that m can be set by external command. The phase-locked loop brings the two counted-down waveforms into phase lock by pulling the frequency of the voltage-controlled oscillator. Hence $f_Q/n = f_o/m$, or $f_o = m \cdot (f_Q/n)$. *See* DIGITAL COUNTER; FREQUENCY DIVIDER.

Usually radio channels are equally spaced in the bands, and thus f_Q/n is chosen to give the correct channel spacing; for example, with a 5-MHz crystal, if $n = 200$, the channel spacing will be 25 kHz, a commonly used value in the ultrahigh frequencies. The value of m is then selected to correspond to the channel required and changed whenever the receiver is retuned. With a synthesizer of this kind and a crystal maintained at a constant temperature, it is not difficult to attain a frequency setting accuracy of under 1 ppm.

Other receiver types. Radio receivers are used in many types of broadcast and communication systems, each of which uses some combination of amplifiers, filters, mixers, and demodulators to fulfill the purpose of that system. More sophisticated systems may combine the basic principles of superheterodyne or homodyne architectures with digital modulation techniques to allow reception of complex signals, such as those containing large amounts of data. Many receiver systems now employ digital signal processing, which in many cases is required for their practical realization.

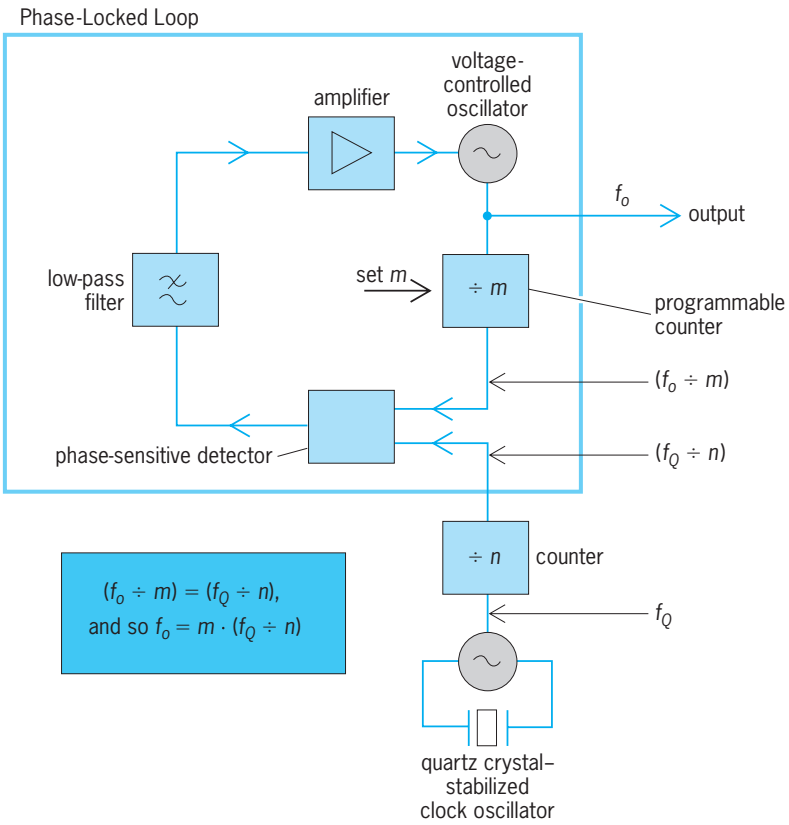

Fig. 8. Frequency synthesizer used to generate accurate local oscillator frequencies.

Television. A television receives radio frequencies modulated with a video signal representing the picture to be displayed. Television receivers typically employ the superheterodyne architecture, with the picture and picture synchronization signals obtained from the demodulator. Analog television receivers use vestigial sideband modulation, a form of amplitude modulation. A digital television receives the picture in digitally modulated form, which provides higher picture quality than that of analog modulation. *See* TELEVISION RECEIVER.

Rake receiver for cellular telephones. In code division multiple access (CDMA) cellular telephones, the rake receiver is used to take advantage of multipath appearing in the received signal. A signal traveling from a transmission source to a receiver typically takes multiple paths due to reflections from buildings and other objects in the environment. Each path will have a different reduction in signal power and a different time delay. The signals traveling on each path combine at the point in space of the receiver antenna input. A rake receiver locks to a reference signal inserted at the transmission source to separate the components in the received signal that have taken the different paths to reach the receiver. Once separated, the signal from each path

is demodulated independently, each in a finger of the rake, and the demodulator outputs are time synchronized and then combined to create a single received signal. The use of the multiple paths in this manner results in a higher-quality received signal than that which could be obtained by only demodulating the incoming combined signal. *See* MOBILE RADIO; SPREAD SPECTRUM COMMUNICATION.

Multiple-carrier systems. Multiple-carrier systems have been developed for transmitting and receiving signals consisting of high-speed digital data streams. The multiple carriers are each digitally modulated to represent data bits, using for example phase-shift keying. The bits from all the carriers taken together are structured to represent some type of digital data signal or signals, such as multiple digital audio data streams. Using a method called orthogonal frequency division multiplexing (OFDM), the carriers are arranged to allow the receiver to easily separate the carriers and extract the bits from each carrier without interference from the other carriers. This approach is used for digital broadcast receivers, in which the multiple-carrier signal is combined with a traditional amplitude- or frequency-modulated carrier.

Impairments. All receivers suffer certain impairments in their performance: the problem of image responses has already been described, together with a solution, the homodyne.

However, the most important impairment is noise. Electrical noise is both received by the antenna and generated within components of the receiver, the first radio-frequency amplifier and mixer being critical, because the signal is weak at this stage. Amplifiers generate less noise than mixers in most radio bands, and so high-performance receivers usually have a radio-frequency amplifier first. Noise can also come from the local oscillator, often run at a high power to improve the ratio of wanted output to noise. Noise is reduced by keeping the receiver bandwidth as narrow as possible, thus reducing the total noise power passed; this is one reason for using tracking receivers, as described above. *See* ELECTRICAL INTERFERENCE; ELECTRICAL NOISE.

A second important impairment is due to nonlinearity, particularly at the front end. The radio-frequency amplifier and converter should give output linearly proportional to their input, but in practice the relationship is better represented by a polynomial. This results in spurious receiver responses—to harmonics and subharmonics—but, more significantly, also in spurious sensitivity very close to the wanted response, due to complex third- and fifth-order interactions. If there are strong unwanted signals near the wanted transmission frequency, these spurious responses cause problems, and are not easily suppressed by the radio-frequency filter. Although better filters can help, there is no solution except to make the early stages of the receiver as linear as possible, and much attention is given to this goal in the design stage. *See* DISTORTION (ELECTRONIC CIRCUITS); LINEARITY.

Impairments to reception performance may also originate from sources outside the receiver. Amplitude-modulation receivers are susceptible to noise from electrical power lines. Noise pulses from electrical motors and switches can enter receivers and lower performance. To combat degradation due to these pulses, receivers may include special noise blanking circuits to detect and reduce the pulses. Noise may also come from electronic equipment near the receiver.

Multipath can degrade receivers that do not have a means of separating the signal components from the multiple paths, such as in the rake receiver already described. A common type of receiver that suffers from multipath is the frequency-modulated broadcast receiver. The result of the multipath is a drop in the level, or fading, of all or part of the signal entering the receiver, causing bursts of noise in the receiver's audio output. Sophisticated mobile receiver systems may select from two or more antennas coupled into a single receiver or may use multiple antennas coupled to multiple receivers to take advantage of the fact that the multipath condition will usually be less at one of the antennas at different times.

More complex receivers may contain means that allow the receiver to modify its operating characteristics automatically depending on the condition of the received signal. For example, the receiver may detect that the incoming signal is very weak, which will result in a noisy audio signal. To reduce the noise in the audio signal the receiver may respond by automatically adding audio filtering to remove some of the noise. William Gosling; J. William Whikehart

Bibliography. *The ARRL Handbook for Radio Amateurs*, American Radio Relay League, annually; H. L. Davidson, *Radio Receiver Projects You Can Build*, Tab Books, 1993; S. J. Erst, *Receiving System Design*, Artech House, 1984; K. Feher, *Wireless Digital Communications*, Prentice Hall PTR, 1995; F. M. Gardner, *Phaselock Techniques*, 2d ed., Wiley, 1979; W. Gosling (ed.), *Radio Receivers*, Peter Peregrinus, 1986; H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, Wiley, 1980.

Radio spectrum allocation

The process of identifying specific frequency bands or blocks of the radio spectrum and determining the broad category of radio communication service or services that will be permitted there, and under what specified conditions. The radio spectrum is the part of the natural spectrum of electromagnetic radiation, lying between the frequency limits of 9 kilohertz and 300 gigahertz, that has been extensively used to provide communications services by means of electromagnetic waves. To make effective use of this resource, the radio spectrum is divided, in the frequency dimension, into a series of discrete bands or blocks, arranged in a Table of Frequency Allocations. Because different portions of the radio spectrum are marked by different propagation features and the nature and scope of radio communication services vary greatly, it has become vital that allocation decisions be made in a way that permits radio communication

services to operate in an efficient and effective manner and that minimizes, to the extent feasible, the potential for interference between services. This process, generally, is referred to as frequency management and involves the work of governmental and international bodies. *See* ELECTROMAGNETIC RADIATION.

The radio portion of the electromagnetic spectrum is not subject to a fixed definition, but has gradually expanded as technology has permitted and demand for radio communication services has grown. For example, in 1959 the "upper end" of the usable radio spectrum band was considered to be 40 GHz, which was lifted to 275 GHz in 1971, and is now considered to be 300 GHz. More significantly, there is nearly constant activity within the radio band as allocations are added, deleted, and modified. This is done in order to facilitate the deployment of new radio communication services, to accommodate the expanded needs of existing services, to recognize efficiencies that can be realized through technological improvements or because of diminished use of a particular radio service, and to react to international decisions and treaties between individual countries regarding use of the radio spectrum.

Need for Spectrum Allocation

Two broad principles underlie spectrum allocation. First, the radio spectrum is a finite resource that multiple users can easily render unusable if they cause harmful interference (that is, interference which seriously degrades, obstructs, or repeatedly interrupts a radio communication service) to each other by, for example, trying to simultaneously broadcast in the same band at the same power in the same geographic area. One means of avoiding this potential problem while simultaneously promoting robust radio use is for the government licensing bodies to adopt specific service and technical rules for the use of radio frequencies within a specific band, and to limit entry to specified parties (licensees). Other times, technical protocols developed by standard-setting bodies, such as those associated with Wi-Fi equipment use, can allow multiple users to effectively operate within a frequency band, even though that is lightly regulated by the government and there are minimal barriers to entry. *See* ELECTRICAL INFERFERENCE; ELECTROMAGNETIC COMPATIBILITY.

Regardless of whether a government decides to designate a spectrum band for a narrowly defined radio service or it permits entities to make more flexible use of the band (and regardless of whether it uses licensing as a means of promoting efficient use of a band), such decisions are still made within the framework of the allocation process, and the Table of Frequency Allocations remains the primary means of organizing the radio spectrum. Allocations are sufficiently broad to permit a wide variety of uses, but are designed to avoid conflicts between incompatible uses. For example, a jurisdiction may choose to allow for cellular telephone service by adopting very specific licensing and service rules within a band broadly allocated for mobile use. However, such a mobile allocation would not typically be considered to coexist within a band allocated for broadcasting, due to the fact that it would be difficult to reconcile potential interference between mobile services and broadcasting services operating in the same band.

A second characteristic of radio waves is that they do not respect political boundaries, and radio communication use in one country can negatively affect the ability of citizens of another country to use the same radio spectrum. This is particularly significant in that the propagation characteristics in certain radio bands, such as those associated with shortwave radio and many of the amateur bands, can be heard on the other side of the world. Thus, radio spectrum allocations are necessarily a matter of worldwide concern. The International Telecommunication Union (ITU), an international organization within the United Nations System where governments and the private sector coordinate global telecommunications networks and services, is the primary entity in matters of international radio allocation. The ITU publishes the Radio Regulations, which is a binding international treaty that governs the allocation and use of radio frequencies worldwide. The Radio Regulations includes an International Table of Frequency Allocations (International Table). For purposes of the International Table, the ITU divides the world into three regions and provides specific allocations for different classes of radio services. Within the allocations, there are some differences between regions, and some frequency bands are subject to greater conditions on use than others. At periodic world radio communication conferences, the ITU evaluates spectrum sharing and allocation matters and provides necessary updates and modifications to the International Table.

It is not uncommon to have regional and national variations of the International Table. First, regional bodies exist to develop consensus for and promote common interests in telecommunications-related matters, and may only be able to foster support for changing the International Table in a single region. One such body is CITEL, the Inter-American Telecommunication Commission, that is an entity of the Organization of American States. Also, individual countries retain sovereign control over spectrum use within their borders and may decide to permit a spectrum use that is different from that contained in the International Table. In these cases, countries that are signatories of the Radio Regulations must ensure that such services are provided in a way that does not cause interference with services that are operating in accordance with the International Table.

Within the United States, spectrum management responsibilities are divided by the Federal Communications Commission (FCC) and the National Telecommunications and Information Administration (NTIA), which is a part of the Department of Commerce. The NTIA governs all federal government use of spectrum, while the FCC governs all non–federal government use of spectrum and is responsible for other telecommunications-related issues. Frequency allocations within the United States

are governed by the United States Table of Frequency Allocations (U.S. Table), which consists of the federal government Table of Frequency Allocations and the non–federal government Table of Frequency Allocations. Within the United States, frequency bands are allocated for exclusive federal government use, exclusive non–federal government use, and shared federal government/non-government use. The Federal Communications Commission publishes a Table of Frequency Allocations that includes, for reference, both the U.S Table and the International Table.

Both the ITU and the United States government have incorporated the concept of primary and secondary use into their tables of frequency allocations. Stations of a secondary service are not permitted to cause harmful interference to stations of primary services to which frequencies are already assigned or to which frequencies may be later assigned, and cannot claim protection from harmful interference caused by stations of a primary service but are entitled to protection from interference caused by other secondary stations. This primary–secondary distinction serves to promote more robust use of the radio spectrum and can permit new services to be deployed in already-occupied spectrum bands, while maintaining protections for those services that have been designated as primary.

Making Allocation Decisions

The characteristics of radio waves vary greatly throughout the radio spectrum, and the propagation characteristics of a particular band, along with the type of services expected to be deployed, will greatly influence allocation decisions. Efficient spectrum allocations take into account these characteristics. Because the wavelength of a broadcast signal can range from hundreds of meters to mere millimeters depending on where in the electromagnetic spectrum the operations are located, the distance a radio signal can travel, its ability to penetrate buildings, and its susceptibility to atmospheric conditions can vary greatly. For example, it is advantageous to allocate broadcasting services in bands that are less susceptible to rain fade (the weakening of a radio signal as it passes through raindrops). Other technical considerations include the ability to control interference, both within a given service and with other, adjacent services, as well as apparatus limitations such as antenna size and power requirements. Allocation decisions also involve a number of policy and economic considerations, including the public need and benefits for the service (both standing alone, and in competition with other services that could be provided in the spectrum), as well as the amount of spectrum required to promote efficient use and ensure that the service can be made economically viable.

Allocations to particular services can often be found throughout different frequency bands. Moreover, as suggested above, allocated spectrum is usually further refined by designating it for a particular service, which involves the adoption of specific service and technical rules, and—if the designated service is a licensed service—by assigning it to entities who then hold license rights to develop and deploy systems within the spectrum. As part of this process, policymakers typically subdivide a frequency band into a series of smaller channels, which are sometimes further divided geographically. For example, within the broadband personal communications service (PCS) in the United States, the "A" block consists of two paired 15-MHz blocks licensed in each of 51 Major Trading Areas. The choice of a specific band plan will also take into account a number of technical factors, such as the separation between frequencies that is necessary to avoid harmful interference, the characteristics of the transmitters and receivers that can be expected to be deployed in the band, the likely distance between units being used in the service and receivers operating on an adjacent band, and any constraints on transmitter power and location. Accordingly, the wide variation in the necessary separation between assigned frequencies that are adjacent to each other within the radio spectrum is due to both the unique characteristics associated with each service and the overall propagation characteristics of that portion of the electromagnetic spectrum. *See* BANDWIDTH REQUIREMENTS (COMMUNICATIONS); RADIO-WAVE PROPAGATION.

Allocated Services

Specific allocated services include the broadcasting service, fixed and mobile services, unlicensed services, and amateur and personal radio services.

Broadcasting service. A broadcast allocation permits transmissions that are intended for direct reception by the general public, and includes sound, visual, and other types of transmissions. The commercial AM radio band consists of the frequency range 535 kHz–1.7 MHz, FM radio operates in the band 88–108 MHz, and television stations operate on several frequency bands—54–88 MHz for channels 2–6, 174–220 MHz for channels 7–13 (collectively, the very high frequency or VHF channels), and 470–794 MHz for channels 14–69 (the ultrahigh-frequency or UHF channels). Channels above 70 (794 MHz) are no longer assigned to television. *See* AMPLITUDE-MODULATION RADIO; FREQUENCY-MODULATION RADIO; RADIO BROADCASTING; TELEVISION.

High-frequency (HF) broadcasting, also known as shortwave broadcasting, is an international service that operates between 5900 and 26,100 kHz. Its broadcast signals propagate over very long distances and thus can be received by the general public in foreign countries. However, due to these propagation characteristics, broadcasts in these bands are subject to interference from HF stations around the world, and the bands are extremely congested. Use of these bands is subject to international planning and coordination, and HF broadcasters are authorized frequencies on a seasonal basis that allows broadcasters to account for changes in propagation conditions, altered programming needs, and objectionable interference situations. The 2003 World Administrative Radio Conference in Geneva adopted a transition plan for the HF bands that, among other

things, increases the amount of spectrum allocated to exclusive HF broadcasting use throughout the world to 3720 kHz in ten frequency bands after April 1, 2007, and permits broadcasts to use new digital modulation techniques that result in clearer reception and more efficient use of the spectrum.

For domestic commercial radio bands in the United States, the FCC began exploring rules for digital broadcasting in 2004, and currently permits in-band on-channel (IBOC) broadcasts, which consist of the transmission of a digital radio broadcast signal centered on the same frequency as the AM or FM station's present frequency. Digital radio is expected to provide near–compact disk (CD) quality reception for stations operating in the FM broadcast band. For AM stations, IBOC digital radio is expected to provide reception approximately equal to present-day analog FM reception.

Within the commercial television bands, jurisdictions have been working to establish a transition plan to digital broadcasting. Because the public reception of digital broadcasts will require new or upgraded tuners, the United States established a complex mechanism by which broadcasters were required to commence digital transmissions between 1999 and 2003 (based on the size and classification of the station), but in which they were also permitted to continue analog operations on their existing 6-MHz-wide channels. Analog transmissions will continue until an official cut-off, which is scheduled to take effect at the end of 2006, but which may be extended under particular circumstances set forth by law, including for those markets where 15% or more households do not have access to either digital television equipped receivers or multichannel video. Ultimately, due to the efficiency of the digital standard, television broadcasting will be able to operate in a "core" broadcast spectrum consisting of channels 2–51, but will also be able to greatly enhance the quality of the video, audio, and supplemental data that are being broadcast. The reclamation of television spectrum has been addressed in two proceedings: the "Upper 700 MHz" band (channels 60–69) which comprises 60 MHz, and the "Lower 700 MHz" band (channels 52–59) which comprises 48 MHz. Within the Upper 700 MHz band, public safety entities will receive 24 MHz of the reclaimed spectrum, with the remaining 36 MHz slated for commercial use including fixed, mobile, and broadcasting services. The Lower 700 MHz band, slated to be auctioned in 2005, is designated for flexible use, and can accommodate fixed, mobile, and broadcast services. In both of these bands, new licensees will be unable to deploy fully operations until television operations cease.

Fixed and mobile services. The fixed service is a radio communication service between specified fixed points, whereas a mobile station is intended to be used while in motion or during halts at unspecified points. From an allocation standpoint, it is most common to create a shared fixed and mobile band in order to promote the most flexible use of the radio spectrum. The types of fixed and mobile services are extensive, and these applications serve many vital commercial, private, and public safety radio needs. One example of a fixed service is microwave communications, which are used mostly for point-to-point communication and serve as backhaul and backbone links for commercial wired and wireless networks in remote or rugged terrain. Mobile services include popular commercial radio telephone services—cellular and broadband PCS—as well as private land mobile radio systems that are used by companies, local governments, and other organizations to meet a wide range of communication requirements, including coordination of people and materials, important safety and security needs, and quick response in times of emergency. Because public safety services are provided under fixed and mobile allocations, a consistent theme in making these allocations has been to provide sufficient spectrum for public safety purposes. In 2002, for example, the United States allocated 50 MHz of spectrum in the 4940–4990 MHz band (4.9-GHz band) for fixed and mobile services (except aeronautical mobile service) and designated the band for use in support of public safety.

One development has been the identification of significant blocks of spectrum to support advanced commercial wireless needs. These voice and data services are commonly known as 3G (third-generation mobile) and IMT-2000, for the study that set forth the characteristics of the advanced services that were expected, and concluded that at least 160 MHz of spectrum were needed to be made available in order to support such applications. These services are expected to build on the success of existing cellular mobile phones. The 2000 World Administrative Radio Conference identified three bands that could be used for such purposes: 806–960 MHz, 1710–1885 MHz, and 2500–2690 MHz. While the bulk of this spectrum was intended for terrestrial-based mobile services, a small portion was also identified for a satellite-based component. One difficulty in making use of these bands on a worldwide basis is the fact that many are already used by jurisdictions for fixed, mobile, and broadcasting operations. For example, the 1920–1980 MHz band—identified for mobile unit transmissions to base stations and allocated as such in European jurisdictions—overlaps with the 1930–1990 MHz broadband PCS base station transmit band in the United States. The United States has identified two 45-MHz blocks of spectrum, consisting of the 1710–1755 MHz and 2110–2155 MHz bands, for the deployment of advanced wireless services, and has begun the process to set service rules for and award licenses in the band. Up to 40 additional megahertz within the 1910–2180 MHz range might also be made available for such use.

Spectrum in the 70-, 80- and 90-GHz bands, the so-called millimeter wave bands, has been made available for new fixed applications, including those that could provide increased broadband capacity. These bands had previously been used primarily for radio astronomy and military applications, and until the 2003 decision authorizing operations in the band, the United States had only authorized operation of

radio technology above 50.2 GHz on an experimental basis and prohibited the marketing of devices that could be used for either licensed or unlicensed operations above 77 GHz. *See* MOBILE RADIO.

Unlicensed services. An increasing number of services are provided in bands whose allocations permit unlicensed uses, where there are few regulatory barriers to entry but in which an individual user has no rights to prevent interference caused by other users. Traditionally, such industrial, scientific, and medical equipment (ISM) bands permitted the development and deployment of equipment or appliances that were designed to generate and use energy in the radio-frequency bands but not for communications purposes. Common consumer ISM appliances are microwave ovens and ultrasonic jewelry cleaners. Other ISM categories include industrial heating equipment and medical diathermy equipment. Another type of unlicensed operation involves the use of equipment that is specifically designed to radiate energy for purposes of communication. While this category has typically been dominated by such devices as cordless phones, garage door openers, television remote controls, and toys, recent growth in the use of this spectrum has been in the area of wireless communication devices that use a common platform established by standards-setting bodies but which can be deployed using off-the-shelf consumer-marketed equipment. Examples of these standards include Wi-Fi®, Bluetooth®, and WiMAX.

Wi-Fi. Wi-Fi is short for "wireless fidelity," and is the term used to describe wireless local-area networks operating on a common set of standards. Wi-Fi applications generally are used to connect personal computers and laptops to a network, and can be used both in private settings (such as a home or business) and via public wireless "hotspots." Wi-Fi is a relatively short-range service, with the best reception taking place within several hundred feet (or less) of a base station. In addition, the speed of data transmission decreases the farther one moves from a base station. Wi-Fi was initially deployed in the 2.4-GHz band, which is shared with such other unlicensed devices as cordless phones, and its increasing popularity led to fears that spectrum congestion could stifle its growth. In the United States, Wi-Fi use has expanded to an additional 300 MHz of spectrum in the 5-GHz band. Internationally, a major focus of the 2003 World Administrative Radio Conference was to expand similarly the available bands for Wi-Fi use, both in order to meet anticipated demand and to simplify the manufacturing of equipment though the establishment of worldwide allocations. The Conference provided a 5-GHz allocation that can be used for Wi-Fi, but with the encouragement that regulators restrict its use to the indoors out of concern for possible interference with existing services in the 5-GHz range. *See* WIRELESS FIDELITY TECHNOLOGY.

Bluetooth. Bluetooth is similar to Wi-Fi in that it is based on a wireless standard and is designed to operate in the 2.4-GHz band. However, the Bluetooth standard operates over a much shorter distance, typically in the range of 30 ft (10 m), and was designed primarily to use short-range radio links to replace cables connecting fixed and portable electronic devices. While Bluetooth has been deployed in such devices as the wireless computer mouse and wireless telephone headset, it can be used across technology platforms, and has been incorporated in cellular phones, personal digital assistants (PDAs), laptop computers, and even automotive electronics. Additional wireless standards have been developed and are in the process of finding both commercial support and adequate spectrum for deployment.

WiMAX. WiMAX is a relatively new wireless networking standard that offers greater range and bandwidth than the Wi-Fi standard. For example, WiMAX can transfer around 70 megabits per second over a distance of 30 mi (50 km), and a single base station can serve thousands of users. Although WiMAX is based on standards that allow for the use of any frequency band between 2 and 11 GHz, it has yet to find a spectrum "home" that would allow for widespread deployment and the necessary economies of scale to produce reasonably priced equipment. Moreover, based on the prospective bands identified by proponents of the service, WiMAX could operate on either licensed or unlicensed spectrum, or possibly even both.

Amateur and personal radio services. The Amateur Radio Service is a voluntary noncommercial communication service used by qualified persons of any age who are interested in radio technique with a personal aim and without pecuniary interest. The amateur service dates back to the 1900s, and its users have made numerous contributions to the development of radio. The amateur service is also noteworthy in its allocations and the wide scope of communications employed by its users. Twenty-seven small frequency bands throughout the spectrum are allocated to this service internationally. Some 1300 digital, analog, pulse, and spread-spectrum emission types may be transmitted. Amateurs throughout the world communicate with each other directly or through ad hoc relay systems and amateur satellites, and use voice, teleprinting, telegraphy, facsimile, and television to communicate. Although amateur radio spectrum is often looked upon as a source of additional spectrum by those seeking to deploy or expand commercial and government services, the amateur community has been quite effective in minimizing any encroachment on these bands. The 2003 World Administrative Radio Conference realigned the allocations near 7 MHz in order to expand the worldwide 40-m amateur band by 100 kHz. Once international broadcasters operating in this band relocate to other frequencies in 2009, the band will be allocated exclusively to the amateur service in all three regions. Also, in 2003 the FCC provided access to five channels in or near the 5250–5400 kHz band on a secondary basis for the amateur service, and upgraded the existing secondary amateur service allocation to primary status in the 2400–2402 MHz band. *See* AMATEUR RADIO.

The personal radio services provide short-range, low-power radio for personal communications, radio signaling, and business communications not

provided for in other wireless services. The range of applications is wide, spanning from varied one- and two-way voice communications systems to non–voice data transmission devices used for monitoring patients or operating equipment by radio control. The Citizens Band (CB) Radio is included in the personal radio services, as is the Family Radio Service—a popular Citizen Band service with a 1-mi (1.6-km) range that is well suited for family use in a neighborhood or during group outings. Additions to the personal radio services have supported wireless medical needs. For example, the Wireless Medical Telemetry Service in the 608–614 MHz, 1395–1400 MHz, and 1427–1432 MHz bands is designed for remote monitoring of patients' health through radio technology and transporting the data via a radio link to a remote location.

Other allocations. Many other important radio services have distinct allocations; among the most notable are space- and satellite-based services, the maritime services, and aeronautical services. In addition, important science research takes place within the radio spectrum, and specific bands have been allocated for the radio astronomy and space research services. Radionavigation-satellite services, for example, include the United States' Global Positioning System (GPS) and Russia's *Glonass* satellite system. The scope of allocated services is reflected in the 30 distinct categories of radio service allocations listed in the U.S Table. International radio telecommunication conferences have considered such issues as providing allocations for the mobile satellite service in the 2-GHz band as part of the larger IMT-2000 allocation for advanced services, finding additional spectrum for high-altitude platform stations, and allowing an aeronautical mobile satellite service in the 14-GHz band to permit in-flight access to the Internet. *See* RADIO ASTRONOMY; SATELLITE NAVIGATION SYSTEMS.

Trends and Developments

A consistent theme in spectrum allocation is the quest to meet the demand for more spectrum. Because technological developments are only able to expand the range of usable radio spectrum to a small degree, and the applications that reasonably can be accommodated in the upper frequency bands are limited due to the different propagation characteristics of different radio frequencies, those parties involved in spectrum allocation decisions have had to employ other means to accommodate increased demand for radio spectrum. The primary tool is the promotion of increased sharing. Different services can use the same spectrum if the risk of interference is minimal and the uses are compatible or can be coordinated. However, increased sharing usually comes at the expense of increased complexity in system design and, thus, greater cost.

Perhaps the easiest means of sharing is by way of geographic separation. Ships and railroads are able to share maritime frequencies in the VHF band due to the distance between land-based rail systems and ocean-going vessels. Geographic separation also permits the reuse of frequencies within a particular service, such as a cellular radio telephone network in which the same frequencies are reused across multiple cell sites within the license area.

A second way to make more use of spectrum is through efficient system design. Directional antennas and antenna discrimination permit fixed microwave links to share the same spectrum with satellite Earth station uplinks, for example. The switch from analog to digital modulation techniques, along with improved coding, has served to increase the amount of information that can be transmitted over a fixed quantity of bandwidth, which in turn has permitted "refarming" in some bands. Refarming maintains the same number of existing channels but reduces the bandwidth occupied by each channel, resulting in additional spectrum that can be put to new uses while permitting existing services to continue to be provided on the narrower channels, by way of more efficient broadcasting techniques. In addition, by employing system protocols such as "listen before talk" and moving to trunked systems versus inflexible dedicated frequencies, users can derive more robust use from a limited amount of spectrum. *See* ANTENNA (ELECTROMAGNETISM); DATA COMPRESSION; DIRECTIVITY; MODULATION.

A third mechanism is band clearing, where lightly used spectrum is made available for new services by moving existing operations to other bands, or even other modes of communication. The deployment of broadband PCS involved the relocation of existing fixed microwave users in order to clear sufficient spectrum to permit PCS licensees to deploy high-value mobile phone services in the band. Some of the legacy microwave users were relocated to other frequency bands where they deployed more modern, spectrally efficient equipment. Others were able to meet their communications needs in different ways, such as by use of fiber-optic cables or subscription to commercial communications services. However, growing demand and band congestion makes successful band clearing an increasingly difficult and contentious task.

A notable development in the quest to make more spectrum available is the deployment of complex systems that have been made possible by advances in computer processing. Cellular phone systems—in which phone calls must be seamlessly handed off from one cell site to another and the base and mobile units must constantly monitor and adjust their signals to maintain a robust connection—would not be possible without advanced computer applications. Many of the unlicensed devices are able to make efficient use of spectrum through such techniques as low-power-sharing protocols and adaptive technologies. For example, cognitive radio technologies can enable a radio device and its antenna to adapt its spectrum use in response to its operating environment, allowing the radio to identify available spectrum that is unusable under current conditions and adjust as conditions change. By contrast, older radios are often programmed to transmit and receive a limited number of frequencies, which can be changed only by physically modifying each radio unit. Mesh networks allow individual radios to connect with other radios

to transmit and relay information instead of connecting to a defined base station, offering the potential for a more robust network and reducing the need for expensive fixed base sites. Spectrum underlays permit the beneficial use of frequencies already being used for traditional high-powered services by adding noninterfering low-powered operations onto the same band. Finally, ultrawideband and spread spectrum applications break away from the traditional view of fixed and discrete band use and instead permit the development of applications that use numerous frequency bands. *See* SPREAD SPECTRUM COMMUNICATION; ULTRAWIDEBAND (UWB) SYSTEMS.

The technology that permits these new developments in spectrum use is likely to outpace the political considerations inherent in spectrum use decision making. Incumbent stakeholders have traditionally been reluctant to share bands that they perceive as having an exclusive right to use, and many of the new technologies do not neatly fit the traditional radio spectrum allocation and licensing models. Nevertheless, because the future of radio spectrum allocation is likely to be driven by these types of technological innovations, the traditional view that radio spectrum allocation requires a central policymaker who must identify specific frequency blocks for the exclusive use of discrete and narrowly defined radio services is quickly becoming outmoded. However, because the radio spectrum continues to be in high demand and increasing quantities of information are being transmitted via radio waves, the core purposes of radio spectrum allocation—the efficient use of the radio resource and the prevention of harmful interference—remain vital to the successful and productive use of this resource. Jamison Prime

Bibliography. Federal Communications Commission, *Rules and Regulations* (annually within vol. 47 of *Code of Federal Regulations*, National Archives and Records Administration, Washington, D.C.); International Telecommunications Union, *Radio Regulations*, 2004 (ITU, Geneva); B. Z. Kobb, *Wireless Spectrum Finder: Telecommunications, Government and Scientific Radio Frequency Allocations in the US, 30 MHz–300 GHz*, McGraw-Hill, 2001; A. K. Maitra, *Wireless Spectrum Management: Making Effective Use of RF Spectrum*, McGraw-Hill, 2004; J. A. Manner, *Spectrum Wars: The Policy and Technology Debate*, Artech House, 2003; P. C. Roosa Jr., *Federal Spectrum Management: A Guide to the NTIA Process*, NTIA Spec. publ. 91–25, 1992.

Radio telescope

An instrument used in astronomical research to study naturally occurring radio emission from stars, galaxies, quasars, pulsars, interstellar clouds, and other astronomical bodies between wavelengths of about 1 mm (300 GHz) and 10 m (30 MHz). At the short wavelength end, the performance is limited by the opacity of the terrestrial atmosphere, and at the long wavelength end by the opacity of the ionosphere.

The performance of a radio telescope is measured by its sensitivity or ability to measure weak sources of radio emission, angular resolution or sharpness of the radio image, and the image quality or ability to image over a wide range of radio brightness. Radio telescopes vary widely in design and appearance, depending on the wavelength of operation and astronomical goals, but they are all composed of one or more antenna elements, each equipped with a radio receiver.

Because cosmic radio sources are extremely weak, radio telescopes are usually very large and equipped with the most sensitive radio receivers available. Moreover, the weak cosmic signals can be easily masked by human-made radio interference, and great effort is taken to protect radio telescopes from locally generated interference by locating them in remote environments, often surrounded by mountains to provide shielding from terrestrial radio transmissions. But with the increasing use of satellite radio transmissions to provide communications, entertainment, and navigation throughout the world, it is increasingly difficult to avoid the effects of radio interference.

The radio signals that arrive at the Earth from cosmic radio sources are as weak as 10^{-19} watt. In order to detect the faint broadband signals generated in interstellar or intergalactic space, radio telescopes are designed to receive over a very large bandwidth of hundreds of megahertz, and the data may be averaged for tens of hours or more to reduce the effect of random noise generated by the radio receiver itself. Also, many atoms and molecules found in interstellar space and around some stars radiate only at discrete frequencies, mostly at short, centimeter and millimeter wavelengths. These spectral lines are studied with radio telescopes able to simultaneously observe up to thousands of different narrow-frequency channels. The most straightforward type of radio spectrometer employs a large number of filters, each tuned to a separate frequency and followed by a separate detector to produce a multichannel, or multifrequency, receiver. However, it is difficult to build a receiver with more than a few hundred frequency channels in this way. Very large multichannel receivers convert a single broad-bandwidth signal into digital form which is then analyzed by the mathematical process of autocorrelation and Fourier transformation to produce the radio spectrum. *See* ANTENNA (ELECTROMAGNETISM); FOURIER SERIES AND TRANSFORMS; RADIO RECEIVER.

Filled-aperture radio telescopes. The simplest type of radio telescope uses a parabolic antenna—the so-called single-dish or filled-aperture radio telescope—which operates in the same manner as a television satellite receiving antenna, to focus the incoming radiation onto a small antenna element called the feed, which is connected to the sensitive radio receiver. The larger the parabolic surface, the more energy is focused on the feed and the more sensitive the telescope. Cryogenically cooled solid-state amplifiers with very low internal noise are used to obtain the best possible receiver sensitivity.

The performance of a radio telescope is limited by various factors: the accuracy of the reflecting surface that may depart from the ideal shape because of manufacturing irregularities; the effect of winds; thermal deformations that cause differential expansion and contraction; and deflections due to changes in gravitational forces as the antenna is pointed to different parts of the sky. Departures from a perfect parabolic surface become important when they are more than a few percent of the wavelength of operation. Since small structures can be built with greater precision than larger ones, radio telescopes designed for operation at millimeter wavelength are typically only a few tens of meters across, whereas those designed for operation at centimeter wavelengths range up to 100 m (328 ft) in diameter.

The largest fully steerable filled-aperture radio telescope currently in operation is the 110 by 100 m (361 ft by 328 ft) antenna (**Fig. 1**) operated by the National Radio Astronomy Observatory in Green Bank, West Virginia. The moving structure of the Green Bank telescope weighs 16 million pounds (7.3 million kilograms) and is able to point to any direction in the sky with an accuracy of a few arcseconds. Each of the 2004 surface panels making up the parabolic surface is held in place by computer-controlled motors which keep the surface accurate to better than half a millimeter, allowing performance at wavelengths less than 1 cm. Green Bank is located in a valley far from high population densities in the National Radio Quiet Zone, which offers protection from local sources of human-made interference. Somewhat less precise 100-m, 76-m, and 64-m diameter (328, 249, 210 ft) antennas are operated by the Max-Planck-Institut für Radioastronomie near Effelsberg, Germany, the Jodrell Bank Observatory near Manchester, UK, and the CSIRO Australia Telescope National Facility near Parkes, NSW, respectively.

The largest filled-aperture radio telescope in the world is the 305-m (1000-ft) fixed spherical reflector operated by Cornell University near Arecibo, Puerto Rico (**Fig. 2**). Although the Arecibo antenna has an enormous collecting area, it can be used only over a limited angle of about 20° from the zenith. The Russian RATAN-600 telescope, located near Zelenchukskaya in the Caucasus Mountains, has 895 reflecting panels, each 7.4 m (24 ft) high, arranged in a ring 576 m (1890 ft) in diameter. Using long parabolic cylinders, standing reflectors, or simple dipole elements, researchers in Australia, France, India, Italy, Russia, and Ukraine have also built antennas with very large collecting areas, but operating at relatively long meter wavelengths where the required precision is less demanding. Several smaller, more precise radio telescopes used at millimeter wavelength are located on mountains, where clear skies and high altitudes minimize absorption and distortion of the incoming signals by the terrestrial atmosphere. A 45-m-diameter (148 ft) radio telescope near Nobeyama, Japan, is used for observations at wavelengths as short as a few millimeters. The French-German Institut de Radio Astronomie Millimétrique (IRAM) in Grenoble, France, operates a 30-m (98-ft) antenna at an altitude of 2850 m (9350 ft) on Pico Veleta in the Spanish Sierra Nevada for observations at wavelengths as short as 1 mm. The 15-m (50-ft) James Clerk Maxwell Telescope located near the summit of Mauna Kea, Hawaii, at elevations above 4000 m (13,100 ft), operates at wavelengths below 1 mm. *See* SUBMILLIMETER ASTRONOMY.

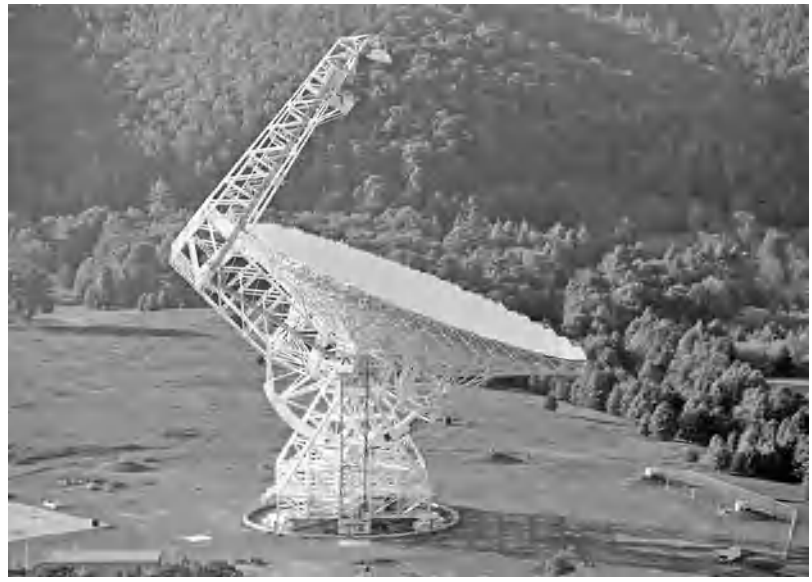

Fig. 1. The 110 by 100 m (361 by 328 ft) fully steerable radio telescope in Green Bank, West Virginia. (*NRAO/AUI/NSF*)

Interferometry and aperture synthesis. The angular resolution, or ability of a radio telescope to distinguish fine detail in the sky, depends on the wavelength of observations divided by the size of the instrument. But even the largest antennas when used at their shortest operating wavelength have an angular

Fig. 2. The 1000-ft (305-m) NAIC fixed spherical dish near Arecibo, Puerto Rico. (*NAIC/NSF-Arecibo Observatory*)

Fig. 3. NRAO 27-element Very Large Array near Socorro, New Mexico, in its most compact configuration. (*NRAO/AUI/NSF*)

resolution only a little better than 1 arcminute, which is comparable to that of the unaided human eye at optical wavelengths. Because radio telescopes operate at much longer wavelengths than optical telescopes, they must be much larger than optical telescopes to achieve the same angular resolution. Fortunately at radio wavelengths, the distortions introduced by the atmosphere are less important than at optical wavelengths, and so the theoretical angular resolution of a radio telescope can be achieved in practice, even for the largest dimensions.

The high angular resolution of radio telescopes is achieved by using the principles of interferometry to synthesize a very large effective aperture from a number of small elements. In a simple two-element radio interferometer, the signals from an unresolved, or point, source alternately arrive in phase and out of phase as the Earth rotates and cause a change in the difference in path from the radio source to the two elements of the interferometer. This produces interference fringes in a manner similar to that in an optical interferometer. If the radio source has finite angular size, the difference in path length to the elements of the interferometer varies across the source. The measured interference fringes from each interferometer pair thus depend on the detailed nature of the radio "brightness" distribution in the sky.

Each interferometer pair measures one Fourier component of the brightness distribution of the radio source. By using multiple antenna elements with a variety of separations and taking advantage of the rotation of the earth, radio arrays can sample a sufficient number of Fourier components to synthesize the effect of a large aperture and thereby reconstruct high-resolution images of the radio sky. Because radio signals can be divided and distributed over large distances without distortion, it is possible to build radio telescope systems of essentially unlimited dimensions. Although the laborious computational task of doing Fourier transforms to obtain images from the large amount of interferometer data was once formidable, these tasks are now easily accomplished by powerful PCs.

The most powerful radio array of this type is the Very Large Array (VLA) located on the Plains of San Agustin near Socorro, in central New Mexico (**Fig. 3**). The VLA consists of 27 parabolic antennas, each measuring 25 m (82 ft) in diameter. The total collecting area is equivalent to that of a single 130-m (426-ft) antenna. However the angular resolution is equivalent to that of a single antenna 35 km (22 mi) in diameter. Each element of the VLA can be moved along a Y-shaped railroad track; thus, it is possible to change the length of the arms between 600 m (2000 ft) and 21 km (13 mi) to vary the field-of-view and the resolution much like a photographic zoom lens. When used at short, centimeter wavelengths in the largest antenna configuration, the angular resolution of the VLA is better than one-tenth of an arcsecond, or about the same as that of the *Hubble Space Telescope* at optical wavelengths. The VLA is operated by the U.S. National Radio Astronomy Observatory and is used by nearly 1000 astronomers each year for a wide variety of research programs.

In conventional interferometers and arrays, coaxial cable, waveguide, or fiber-optic links are used to distribute a common reference signal to each antenna and also to return the received signal from an individual antenna to a central laboratory where it is correlated with the signals from other antennas. In cases in which antennas are spaced more than a few tens of kilometers apart, however, it becomes prohibitively expensive to employ real physical links to distribute the signals. Very high frequency (VHF) or ultrahigh frequency (UHF) radio links can be used, but the need for a large number of repeater stations makes this impractical for spacings greater than a few hundred kilometers. The Multi-Element Radio-Linked Interferometer Network (MERLIN), operated by the Nuffield Radio Astronomy Laboratories at Jodrell Bank in the UK, uses microwave radio links to connect seven antennas separated by up to 200 km

Fig. 4. Artist's conception of the 10-element Very Long Baseline Array. Signals from each of the 10 antennas are independently recorded on magnetic tape or disks which are then sent to the array headquarters in Socorro, for processing to form radio images with an angular resolution better than 0.001 arcseconds HST. (*NRAO/AUI/NSF*)

(124 mi) in the southern part of England. It is used primarily to study compact radio sources associated with quasars, active galactic nuclei (AGN), and cosmic masers with a resolution of a few hundredths of an arcsecond.

Interferometer systems of essentially unlimited element separation can be formed by using the technique of very long baseline interferometry, or VLBI. In a VLBI system the signals received at each element are recorded by broad-bandwidth tape recorders or large-capacity computer disks. The recorded tapes or disks are then transported to a common location where they are replayed and the signals combined to form interference fringes. A single magnetic tape capable of recording for several hours can contain 10^{12} bits of information, which is roughly equivalent to storing the entire contents of a modest-sized library.

The Very Long Baseline Array (VLBA) consists of ten 25-m (82-ft) dishes spread across the United States from the Virgin Islands to Hawaii **Fig. 4**. The VLBA operates at wavelengths from 3 mm to 1 m and is used to study quasars, galactic nuclei, cosmic masers, pulsars, and radio stars with a resolution as good as 0.0001 arcsecond, or more than 100 times better than that of the *Hubble Space Telescope*. Data tapes or computer disks from the ten individual antenna elements of the VLBA are shipped to a special processing center in New Mexico where they are replayed and the signals analyzed to form images. Precise timing between the elements is maintained by a hydrogen maser atomic clock located at each antenna site.

In 1997, Japanese radio astronomers working at the Institute for Space Science near Tokyo launched an 8-m (26-ft) dish, known as VSOP, in Earth orbit. Working with the VLBA and other ground-based radio telescopes, VSOP gave interferometer baselines up to 20,000 km (12,400 mi). Russian radio astronomers plan a similar space VLBI mission with baselines extending more than 100,000 km (62,000 mi) from the Earth.

Interferometers and arrays are also used at millimeter and submillimeter wavelengths to study the formation of stars and galaxies with resolution better than can be obtained with simple filled-aperture antennas. The operation of arrays at millimeter and submillimeter wavelengths is very difficult and requires siting the instrument at very high and dry locations to minimize the phase distortions of signals as they propagate through the terrestrial atmosphere. Millimeter interferometers and arrays are found at the IRAM Plateau de Bure facility in France, the Japanese Nobeyama Observatory, and near the Owens Valley in California. The Harvard-Smithsonian Center for Astrophysics in collaboration with the Academia Sinica of Taiwan operates the eight-element Submillimeter Array (SMA) near the summit of Mauna Kea at an elevation of 4080 m (13,385 ft). The SMA is designed to work at wavelengths as short as 0.3 mm. A major new international facility under construction in the Atacama Desert in northern Chile at an elevation of more than 5000 m (16,400 ft) is expected to be completed by 2012 and will consist of more than fifty 12-m (39-ft) dishes also operating at wavelengths as short as 0.3 mm. *See* ASTRONOMICAL OBSERVATORY; RADIO ASTRONOMY. Kenneth Kellermann

Bibliography. B. F. Burke and F. Graham Smith, *An Introduction to Radio Astronomy*, Cambridge University Press, 2d ed., 2002; J. D. Kraus, *Radio Astronomy*, 2d ed., Cygnus-Quasar Books, 1986; K. Rohlfs and T. Wilson, *Tools of Radio Astronomy*, 4th ed., Springer, 2003; G. L. Verschuur, *The Invisible Universe Revealed: The Story of Radio Astronomy*, Springer, 1986; G. L. Verschuur and K. I. Kellermann, *Galactic and Extragalactic Radio Astronomy*, Springer, 1991.

Radio transmitter

A generator of radio-frequency (rf) signals for wireless communication over some distance, which can vary from the short ranges within a building to intercontinental distances. Most applications utilize signals from very low frequencies (VLF) to extremely high frequencies (EHF); some applications require frequencies as low as 45 Hz or as high as 100 GHz. The radio-frequency output power varies from a fraction of a watt in emergency beacons and portable equipment to several megawatts in long-range, low-frequency transmitters. *See* RADIO SPECTRUM ALLOCATIONS.

The architecture (organization) of a radio transmitter is determined by the type of signal it is intended to produce. The four basic architectures are those used for continuous-wave, frequency-modulation, amplitude-modulation, and single-sideband signals. Transmitters for some applications (for example, television) use a combination of these architectures

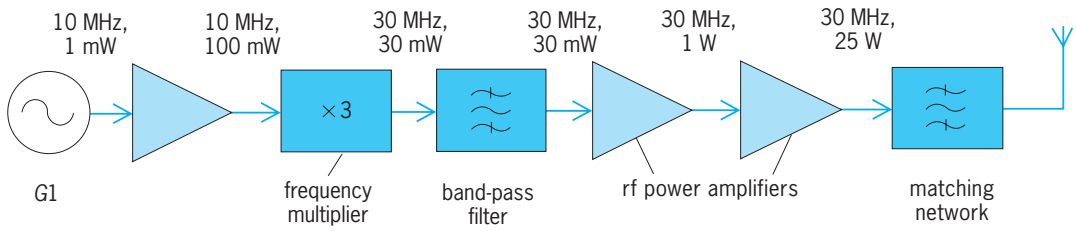

Fig. 1. Basic continuous-wave (CW) transmitter.

(for example, frequency modulation for sound and single sideband for video), while others (for example, Loran C) use unique architectures. Alternative architectures such as envelope elimination and restoration or outphasing can be used to improve efficiency. *See* ELECTRICAL COMMUNICATIONS; MODULATION.

Continuous-wave (CW) transmitter. The most basic type of radio transmitter produces only a continuous-wave signal. Such transmitters are often switched on and off (keyed) to produce telegraph signals. *See* TELEGRAPHY.

CW architecture. The block diagram of a simple continuous-wave transmitter is shown in **Fig. 1**. The oscillator *G*1 produces a low-power signal, which is boosted to the final output power by a series of progressively larger power amplifiers. The optional inclusion of a frequency multiplier improves stability by allowing the frequencies of the oscillator and high-power amplifiers to be different. *See* FREQUENCY MULTIPLIER; OSCILLATOR; POWER AMPLIFIER.

The architecture shown in **Fig. 2** includes both frequency translation and power splitting, which makes it more suitable for generating high-power signals at various frequencies. While at a relatively low level, the signal is translated by a mixer to the desired output frequency. After amplification by a chain of power amplifiers, it is split into two parts to drive two final power amplifiers whose outputs are combined to produce the transmitter output. *See* MIXER.

Typically, each stage of amplification increases the signal amplitude by 10–20 decibels. Since the signal has a constant amplitude, class C power amplifiers are generally used to provide good efficiency in dc-to-radio-frequency power conversion. For higher efficiency, power amplifiers of classes D, E, or F can be used. *See* AMPLIFIER.

Since the applications for continuous-wave transmitters generally require operation over only a small bandwidth, tuned matching networks are generally employed. These networks provide the final power amplifier with the load impedance required to produce the desired power output, given the available power supply voltage. *See* IMPEDANCE MATCHING.

Keying. Keying for telegraphy is generally produced by an analog switch inserted between two low-level stages. Alternatively, the bias points of one or more power amplifiers can be shifted to put their transistors into a cut-off condition. A shaping filter is often employed to produce the desired pulse shape.

Frequency-modulation (FM) transmitter. Analog frequency modulation is widely used for voice communication, high-quality audio broadcasting, and television audio. Frequency-shift keying (FSK) and phase-shift keying (PSK) are widely used for transmission of digital data via radio-frequency signals. *See* FREQUENCY MODULATION; FREQUENCY-MODULATION RADIO; MOBILE RADIO; RADIO BROADCASTING; TELEVISION.

FM architecture. Frequency-modulated and phase-modulated (PM) signals have constant amplitudes and are therefore produced by transmitters with architectures similar to those of the continuous-wave transmitter (Figs. 1 and fig. 2). The principal change is the replacement of oscillator *G*1 by a frequency or phase modulator. In frequency-modulation transmitters, the frequency multiplier increases the frequency deviation as well as the carrier frequency of the frequency-modulated signal. *See* PHASE MODULATION.

Frequency modulators. In communication applications, the frequency modulator is typically a voltage-controlled crystal oscillator (VCXO) in which the

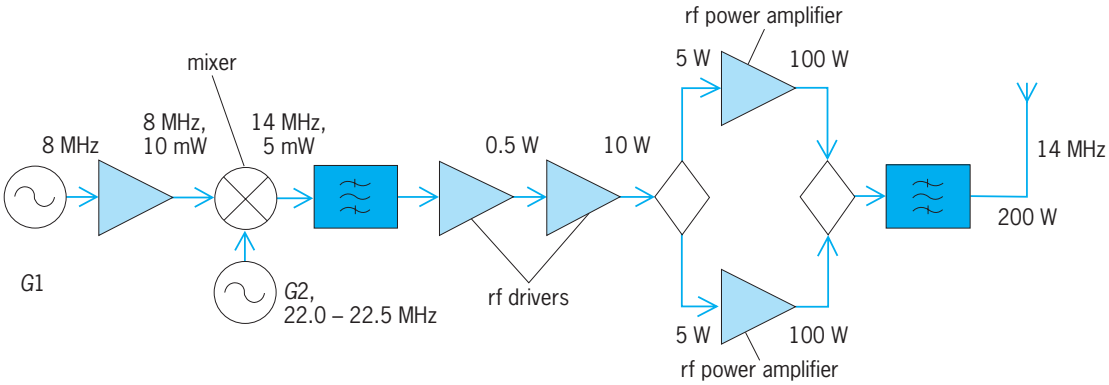

Fig. 2. Continuous-wave (CW) transmitter with frequency conversion and power splitting.

capacitance of a varactor diode is used to vary slightly the frequency of a crystal oscillator. Other applications employ various types of modulators, including phase-shift, phase-locked-loop, comparator, and Armstrong. *See* FREQUENCY MODULATOR.

Frequency-shift-keying signals are generally produced by selecting the output of one of two oscillators according to the current data bit. Binary phase-shift keying is typically produced by using a double-balanced mixer to invert the polarity of the carrier. Other forms of phase-shift keying generally employ digital generation (and selection) of the several phases.

Amplitude-modulation transmitter. Full-carrier amplitude modulation is used in medium-frequency (MF) broadcasting, high-frequency (HF) international broadcasting, citizen-band communication, aircraft communication, and nondirectional navigation beacons. *See* AMPLITUDE MODULATION; ELECTRONIC NAVIGATION SYSTEMS.

AM architecture. Most modern full-carrier amplitude-modulation transmitters produce the output signal by amplitude modulation of the final radio-frequency power amplifier. Generally, the modulation is accomplished by varying the supply voltage of the radio-frequency power amplifier with a high-power radio-frequency amplifier. Since the radio-frequency carrier has constant amplitude until the final power amplifier, the architecture of the radio-frequency chain (**Fig. 3**) is similar to that of a continuous-wave or frequency-modulation transmitter. *See* AUDIO AMPLIFIER.

Amplitude modulators. Class B audio-frequency power amplifiers are sometimes used in low-power transmitters. However, modern high-power transmitters use either a class S audio-frequency power amplifier or pulse-step modulation for high efficiency. The input to the audio-frequency power amplifier is a sum of audio-frequency components (sound) and dc (carrier level).

Grid-bias modulation is sometimes employed in inexpensive, low-power amplitude-modulation transmitters. In this technique, the audio-frequency signal and a bias voltage are fed to the gate of a final class C radio-frequency power amplifier that uses field-effect transistors (FETs). The variation of bias causes variation of the conduction angle, hence variation of the amplitude of the output. While this method is not as linear as series modulation, it eliminates the need for a high-power audio amplifier.

Transformer coupling of high-power audio signals into the supply voltage of the final radio-frequency power amplifier was often employed in older transmitters. However, this technique is now little used because of the size and weight of the transformer. It is, of course, possible to produce full-carrier amplitude-modulated signals with linear radio-frequency power amplifiers, as in the single-sideband transmitter, discussed below. *See* AMPLITUDE MODULATOR.

Single-sideband (SSB) transmitter. Single-sideband amplitude modulation is widely used for high-frequency voice communications, including military, marine, aeronautical, diplomatic, and amateur. It also finds use (as amplitude-compandored single-sideband, or ACSB) at very high frequencies (VHF) and ultrahigh frequencies (UHF).

Although single sideband is technically a form of amplitude modulation, the single-sideband signal itself has variations of both amplitude and phase. Signals such as multitone, independent sideband (ISB), and vestigial sideband (VSB), used for video, also possess such characteristics. Consequently, these signals are traditionally amplified by a chain of linear radio-frequency power amplifiers operating in class B. *See* SINGLE SIDEBAND.

SSB architecture. The architecture of a modern, broad-band single-sideband transmitter is shown in **Fig. 4**. The low-level output of the single-sideband modulator is first shifted by the local oscillator *G*2 to an intermediate frequency (i-f) that is at least twice the highest output frequency. The intermediate-frequency signal is then shifted downward to the desired output frequency by the variable-frequency oscillator (VFO) *G*3. The mixer output is low-pass-filtered and then amplified to the desired power. *See* ELECTRIC FILTER.

This high-side-mixing architecture creates only

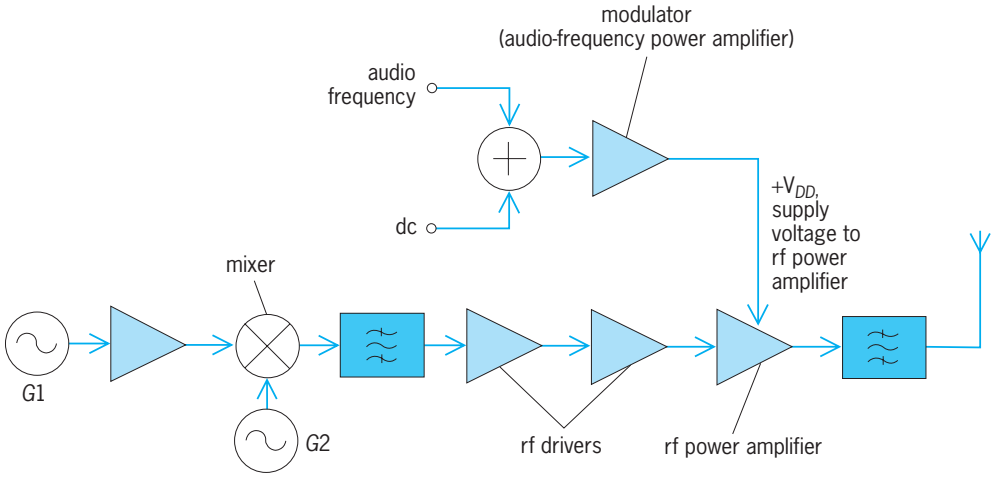

Fig. 3. Amplitude-modulation (AM) transmitter.

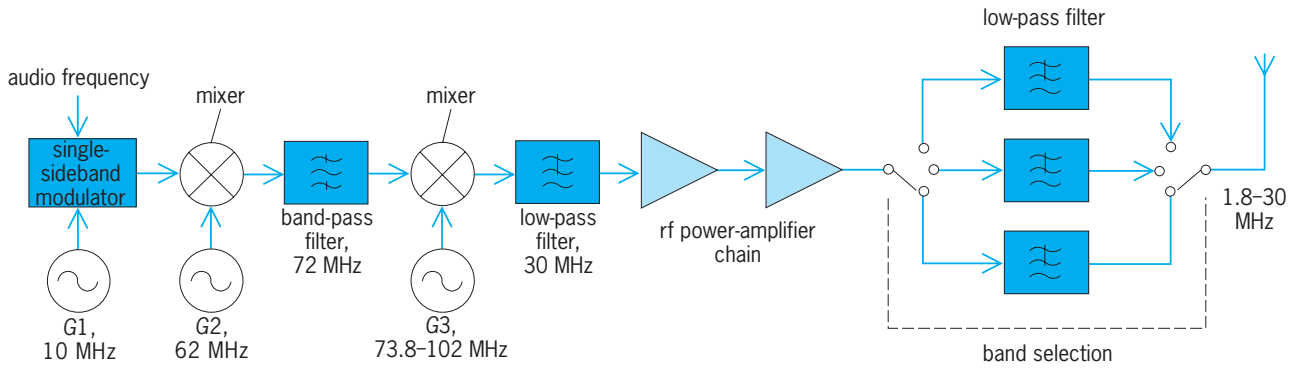

Fig. 4. Single-sideband (SSB) transmitter.

modest demands on the synthesizer used to implement the variable-frequency oscillator *G*3. Further, no first-order products in the final mixer output fall within the output band. Consequently, the entire chain of radio-frequency power amplifiers is untuned and can amplify signals without adjustment for any frequency in the intended range of operation. It is necessary only to switch-select an appropriate output filter to ensure adequate suppression of harmonics. Such filters are generally based upon elliptic designs, and each covers approximately 1.5 octaves of frequency.

Older single-sideband transmitters designed for operation on only a limited number of frequency bands (for example, amateur) typically vary the frequency of oscillator *G*1 for tuning within a given band and select one of several different frequencies for oscillator *G*2 to produce an output in the desired band. In such transmitters, each stage of amplification is generally tuned to the current band of operation.

Direct-conversion architecture is used in some single-sideband transmitters. Such transmitters are analogous to high-power phasing-type single-sideband modulators. While this approach offers some simplicity, a very high degree of amplitude and phase matching is required to achieve good rejection of the unwanted sideband.

SSB modulators. The traditional method for generating high-quality single-sideband signals first provides a double-sideband/suppressed-carrier amplitude-modulation (DSB/SC-AM) signal in a double-balanced mixer. The unwanted sideband is then removed by a mechanical or crystal filter.

Single-sideband signals can also be generated by the phasing technique. The audio signal is applied to an all-pass filter, which produces two outputs that differ in phase by 90°. The two audio-frequency signals then modulate two radio-frequency carriers that differ in phase by 90°. The addition or subtraction of the two double-sideband/suppressed-carrier outputs cancels the unwanted sideband.

The Weaver (third-method) single-sideband modulator is sometimes used, especially when the filtering and modulation processes are accomplished digitally. This technique eliminates the need for audio-frequency all-pass filters by modulating the input audio signal onto two midband carriers (for example, 1.6 kHz) that are in phase quadrature. The outputs of these mixers are then low-pass-filtered (for example, at 1.5 kHz) to remove the carrier and unwanted sideband. The two low-pass-filtered signals modulate in-phase and phase-quadrature carriers. The outputs of these two mixers are then summed to produce the described sideband, as in the phasing method.

Techniques for high efficiency. High-efficiency transmitters are based upon a combination of high-efficiency radio-frequency power amplifiers, modulation techniques, and high-efficiency modulators. These techniques not only increase the efficiency when the output is near the peak output power but also keep it relatively constant at all signal levels. Consequently, the average efficiency can be several times that of a conventional power amplifier for signals with large peak-to-average ratios, such as speech and multiple tones. The increased efficiency reduces the power costs for broadcast transmitters and reduces size, weight, and power supply requirements for portable, remote-relay, and satellite-borne transmitters.

Power amplifiers. Conventional power amplifiers (classes A, B, and C) use their active devices (tubes, bipolar junction transistors, or field-effect transistors) as variable resistors and therefore inherently dissipate a portion of the dc input power as heat. High-efficiency power amplifiers (classes D, E, F, and S) use their active devices as switches or employ wave-shaping circuits, allowing more of the dc input power to be converted into radio-frequency output. *See* WAVE-SHAPING CIRCUITS.

A class D power amplifier employs a pair of transistors driven to act as a switch. An output filter allows only the fundamental-frequency component of the resultant square wave to reach the load. The efficiency is ideally 100% and does not inherently decrease with mismatched loads. Class D power amplifiers are now practical through high frequencies and the lower range of very high frequencies at power levels of 1 kW or more.

A class E power amplifier employs a single transistor driven to act as a switch. The output filter is selected so that it brings the drain voltage to zero at

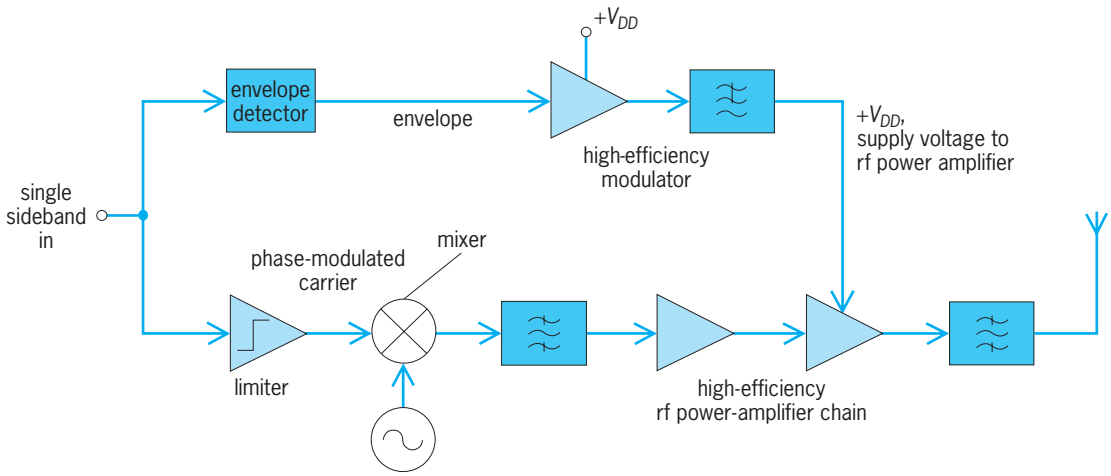

Fig. 5. Envelope-elimination-and-restoration (EER) transmitter.

the instant the transistor is switched on. The power losses due to charging the drain capacitance are eliminated, and the efficiency is ideally 100%. As a result, class E power amplifiers are attractive at very high frequencies and ultrahigh frequencies.

A class F power amplifier employs a single transistor and a multiple-resonator output circuit, preferably implemented from transmission lines. The multiple resonances facilitate the formation of approximate square waves, thus reducing the dissipation of the amplifier. Class F power amplifiers can be implemented at ultrahigh frequencies and microwave frequencies where switching-mode operation of the transistors is not possible.

High-efficiency modulators. Both class S amplification and pulse-step modulation (PSM) are currently in use for high-efficiency modulation. Both ideally achieve 100% efficiency.

Class S modulators are based upon pulse-width modulation (PWM) with a switching frequency several times the highest output frequency. The pulse-width-modulated signal is boosted to the desired power level by switching amplifiers, after which the desired audio output is obtained by a low-pass filter. *See* PULSE MODULATION.

A pulse-step modulator is basically a high-power digital-to-analog converter that selects the one power supply voltage of several (8–12) that is closest to the desired modulating voltage at any given time. *See* DIGITAL-TO-ANALOG CONVERTER.

Envelope elimination and restoration. Any narrow-band signal can be regarded as having simultaneous amplitude (envelope) and phase modulation. The envelope-elimination-and-restoration (EER) technique uses this principle to facilitate high-efficiency linear radio-frequency amplification.

An EER system (**Fig. 5**) first detects the envelope and hard-limits the carrier. The hard-limited carrier is a phase-modulated signal suitable for amplification by efficient but nonlinear radio-frequency power amplifiers (for example, classes C, D, E, and F). The envelope consists of audio-frequency and dc components and is therefore amplified efficiently by a class S modulator or power supply modulator. Amplitude modulation of the final radio-frequency power amplifier restores the envelope, producing the desired high-power single-sideband signal.

In some applications (for example, AM stereo), the envelope and phase can be produced easily by digital signal processing, thus simplifying the EER system. *See* STEREOPHONIC RADIO TRANSMISSION.

Other techniques. Digital power addition is used to produce amplitude-modulated signals at low frequencies (lf) and medium frequencies. Power amplifiers of various outputs are connected in series and toggled on or off to produce the desired instantaneous output voltage.

In outphasing, the outputs of two power amplifiers are combined. However, the power amplifiers are driven with signals of different phases. The phase difference is varied dynamically so that the vector sum of the two outputs has the desired amplitude.

The Doherty technique combines the outputs of two power amplifiers via a quarter-wavelength line. At intermediate signal levels, one power amplifier operates near its peak output power and is therefore at its maximum efficiency. The second power amplifier is used to produce peak system output power.

Transmitters for the pulsed-continuous-wave signal used in Loran C radio navigation use a half-cycle technique. Sequential discharge of several inductance-capacitance (*L-C*) networks into the antenna efficiently produces the desired high-power signal. *See* LORAN; RADIO. Frederick H. Raab

Bibliography. *The ARRL Handbook for the Radio Amateur*, annually; H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid State Radio Engineering*, 2d ed., 2002; *Proceedings of RF Expo East '87*, Boston, November 11–13, 1987; *Proceedings of RF Technology Expo '88*, Anaheim, California, February 10–12, 1988; W. E. Sabin and E. O.Schoenike (eds.), *Single-Sideband Systems and Circuits*, 2d ed., 1995; V. Shakhgildyan, *Radio Transmitter Design*, 1987; J. R. Smith, *Modern Communication Circuits*, 2d ed., 1997.

Radio-wave propagation

The means by which radio signals are transported through space from a transmitting antenna to a receiving antenna. Radio signals are electromagnetic waves which travel with the velocity of light and can be reflected, refracted, diffracted, scattered, and absorbed. Unlike visible light, radio frequencies cover many octaves, from about 10 kHz to 60,000 MHz (wavelengths from 30,000 m to 0.5 cm). Since frequency is an important parameter, radio propagation characteristics vary over a wide range. At the higher radio frequencies the similarity with visible light is evident. At the lower frequencies the radio waves follow the surface of the Earth by a mechanism that in geometrical optics is unimportant and relatively unknown.

The power radiated from a transmitting antenna is ordinarily spread over a relatively wide area. As a result, the power available at most receiving antennas is only a very small fraction (10^{-8} to 10^{-16} or less) of the radiated power. The principal function of radio-wave propagation analysis is to make possible estimation of the expected received power in order to predict the usefulness of a radio signal at any location remote from the transmitter. For a discussion of antenna radiation *see* ANTENNA (ELECTROMAGNETISM); ELECTROMAGNETIC RADIATION.

Typical distances that can be achieved with usual types of equipment are shown in **Fig. 1**. The frequencies around 20 kHz can be received reliably at distances of thousands of miles but are limited to telegraph-type signals and require very large transmitting antennas. Higher frequencies are needed for voice, and still higher frequencies for television transmission. As the frequency increases, the transmission range tends to decrease. Frequencies above 100 MHz can transmit wideband signals, but they are limited to approximately line-of-sight distances with the usual type of equipment. However, distances of 200 mi (320 km) or more are possible by the use of high power and large antennas to provide narrow "searchlight" beams.

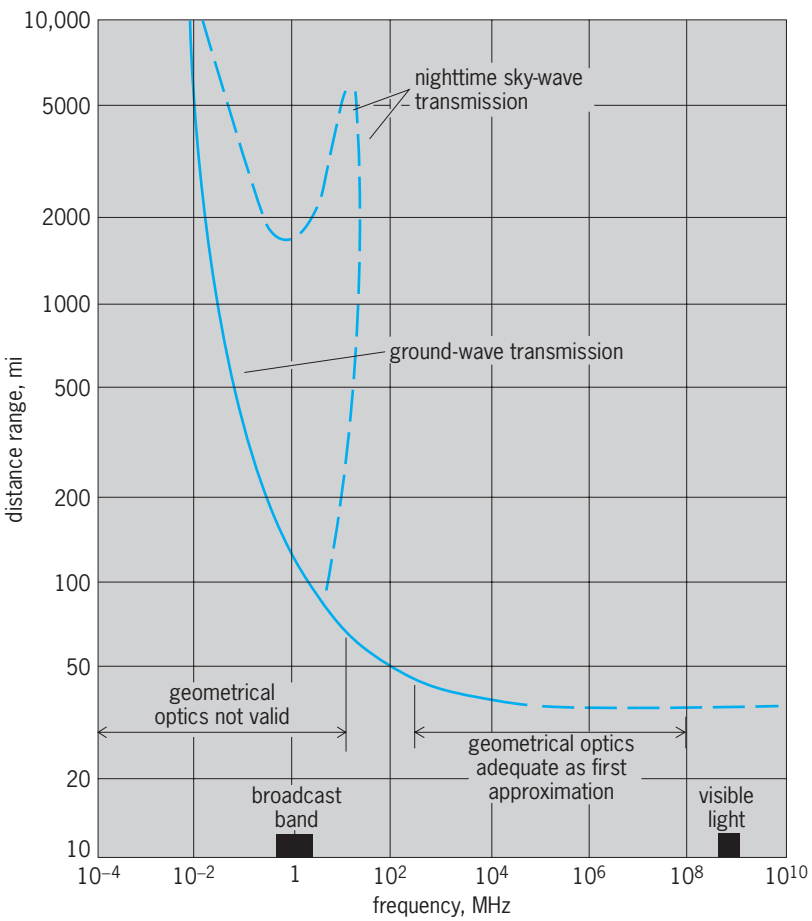

Fig. 1. Transmission range of electromagnetic waves. 1 mi = 1.6 km.

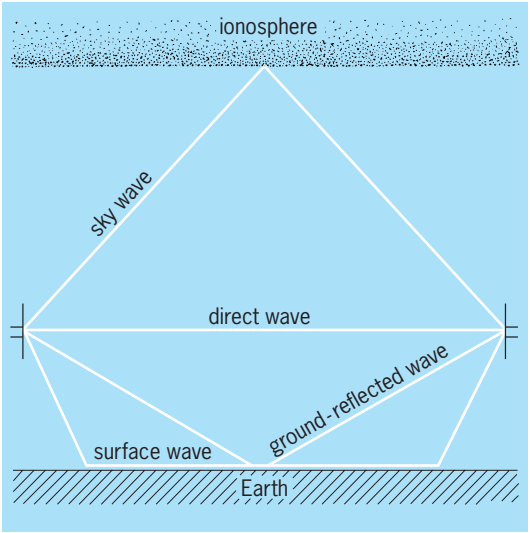

Fig. 2. Possible transmission paths between antennas.

Reflections from the ionosphere (ionized layers 50–250 mi or 80–400 km above the Earth's surface) provide a useful but variable long-distance service at frequencies less than about 30 MHz. These reflections account for the long-range broadcast coverage at night and for the shortwave intercontinental communication. *See* IONOSPHERE.

The principal components of the received radio signal are shown symbolically in **Fig. 2**. The vector sum of the direct, reflected, and surface waves has been called the space wave, ground wave, or tropospheric transmission to differentiate it from the ionospheric reflections. The ionospheric and surface waves are the principal components at frequencies below 10–30 MHz. The direct and reflected rays are the principal factors at frequencies above 30–50 MHz. Although the ionospheric, direct, and ground-reflected waves can be easily visualized as rays, the surface wave is more difficult to understand; it originates at the air-Earth boundary because the Earth is not a perfect reflector.

Line-of-sight transmission. This is the first and simplest concept in radio-wave propagation. Radio transmission in free space results in a decrease in energy per unit area in accordance with the inverse-square law. The ratio of the transmitter power to the received power is called the radio transmission loss;

for nondirectional antennas in free space it is $P_T/P_O = (4\pi d/\lambda)^2$, where the distance d and wavelength λ are expressed in the same units. Here P_T is the transmitted power, and P_O is the power received under free-space conditions. The wavelength $\lambda = c/f$, where c is the velocity of light, and f is the frequency in hertz (cycles per second). *See* INVERSE-SQUARE LAW.

Ground effects. The presence of the ground may affect the antenna impedance and efficiency, and reflections from the ground tend to interfere with the free-space wave. The received power near the ground is ordinarily less than would be expected in free space. At frequencies in the amplitude-modulation (AM) broadcast band and lower, transmission to moderate distances is primarily by the surface wave because the direct and reflected waves tend to cancel. For the surface wave the received power P_R is less than the free-space value by the fraction shown as relation (1). The factor K is greater for

$$\frac{P_R}{P_0} \approx \left| K \frac{\lambda}{d} \right|^2 \tag{1}$$

vertical polarization of the waves than for horizontal polarization and is greater for vertically polarized waves over seawater than for transmission over dry land.

At frequencies in the television band and higher, both the transmitting antenna and the receiving antenna are usually several wavelengths above the ground. In this case the surface wave and its dependence on polarization becomes less important than the direct and reflected waves. Reflections from the ground may either cancel or reinforce the direct wave, and the power P_R received over flat ground is given approximately by Eq. (2). Here h_1 and h_2 are

$$\frac{P_R}{P_0} = \left| 2 \sin \frac{2\pi h_1 h_2}{\lambda d} \right|^2 \tag{2}$$

the antenna heights in the same units as the wavelength and distance.

At still higher frequencies, that is, in the microwave region, irregularities in the surface of the Earth are frequently large compared to wavelength, and the magnitude of the reflection coefficient becomes substantially less than unity. *See* REFLECTION AND TRANSMISSION COEFFICIENTS.

As the reflected wave is weakened, radio propagation approaches free-space transmission, as long as there is adequate clearance over all intervening obstacles. In free-space transmission between two directive antennas the received power is related to the transmitted power by Eq. (3). Here A_T and A_R are

$$\frac{P_R}{P_T} = \frac{A_T A_R}{(\lambda d)^2} \tag{3}$$

the effective aperture areas of the transmitting and receiving antennas, respectively, measured in the same units as the wavelength and distance.

Fading. Variations in signal level with time are caused by changing atmospheric conditions. The severity of the fading usually increases as either the frequency or path length increases. The path of a radio wave is not a straight line, except for the ideal case of a uniform atmosphere. The transmission path may be refracted down or up, depending on atmospheric conditions. This bending may either increase or decrease the effective path clearance, and unfavorable (upward) bending may have the effect of transforming a line-of-sight path into an obstructed one. This type of fading may last for several hours at a time. The frequency of its occurrence and its depth can be reduced by increasing the path clearance, particularly in the middle of the path.

Most of the fading that occurs on "rough" paths with adequate clearance is the result of interference between two or more rays traveling slightly different routes in the atmosphere. This fast multipath type of fading is relatively independent of path clearance.

Most fading is a temporary diversion of energy to some direction other than the intended location, but absorption effects are important in the microwave region. At frequencies above 5000–10,000 MHz the presence of rain, snow, or fog introduces an absorption in the atmosphere that depends upon the amount of moisture, the particle size, and the frequency. At frequencies higher than 15,000–20,000 MHz the additional attenuation caused by heavy rain tends to limit the path length to only a few miles, if high reliability is required. In addition to the effect of rain, some selective absorption may result from the oxygen and water vapor in the atmosphere. The first absorption peak due to water vapor occurs at about 23,000 MHz, and the first absorption peak for oxygen occurs at about 60,000 MHz. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION.

Radar transmission. The received radar echo is very weak because only a small part of the transmitted energy reaches the target, and only a small fraction of the energy reflected from the target is returned to the radar receiver. The transmission loss under free-space conditions is given by Eq. (4), where A'_R and A'_T pertain to the reflecting properties

$$\frac{P_R}{P_T} = \frac{A_T A'_R}{(\lambda d)^2} \frac{A'_T A_R}{(\lambda d)^2} \tag{4}$$

of the target. When accurate information on the target characteristics is unavailable, it is usually assumed that A'_R is the geometrical cross section of the target and that $A'_T = \lambda^2/4\pi$ (uniform radiation in all directions).

Outer-space vehicles. Radio transmission to Earth satellites and other outer-space vehicles is an extension of line-of-sight transmission to long-distance, free-space conditions. Each tenfold increase in distance requires a hundredfold increase in transmitter power, when the antenna size and other parameters remain constant. The assigned frequency needs to be high enough to penetrate the ionosphere and to provide adequate directivity and bandwidth; an upper limit is set by the limitations of rain attenuation and atmospheric absorption. *See* SPACE COMMUNICATIONS.

Passive satellites, such as *Echo 1* or *Echo 2*, are like radar targets or mirrors in that the transmitted signal is reflected but not amplified. The resulting weak echo limits their information capacity to much less than is possible with active satellites.

Active satellites, such as the *Intelsat* and *Molniya* communication satellites or the lunar and interplanetary probes, contain radio receivers and radio transmitters with the necessary power supply on board. Since the transmission loss before amplification pertains to a one-way rather than a round-trip path, the resulting signal is strong enough either to carry more information or to go farther than is possible with a passive satellite. *See* COMMUNICATIONS SATELLITE.

Propagation beyond optical horizon. Radio transmission beyond the line of sight can be achieved by three principal methods: refraction, diffraction, and reflection. When these effects are intermixed and cannot be separated easily, the energy is said to be scattered.

Refraction. The dielectric constant of the atmosphere normally decreases gradually with increasing altitude. The result is that on the average the radio ray is bent or refracted toward the Earth so that the distance to the radio horizon is slightly greater than to the optical horizon. The amount of refraction is variable, and exceptionally long-range transmission may occur occasionally. The corresponding phenomenon in optics makes it possible to see lights or other objects that are normally below the horizon. Conversely, when the radio energy is bent away from the Earth (upward bending), the transmission loss is increased. *See* REFRACTION OF WAVES.

Diffraction. Radio waves are also transmitted around the Earth by the phenomenon of diffraction. When light waves are partially blocked by an obstacle, the transition from light to dark at the edge of a shadow is gradual rather than infinitely sharp. The amount of energy diffracted around an obstruction decreases as the frequency is increased. Typical obstructions include hills, trees, and buildings, as well as the curvature of the Earth. *See* DIFFRACTION.

Reflection and scatter. Most of the experimental data at points far beyond the horizon are intermediate between the values expected for diffraction over a smooth sphere and for diffraction over a ridgelike obstruction. Various theories have been advanced to explain these effects. The explanation most commonly accepted is that energy is reflected or scattered from atmospheric layers and turbulent air masses. At points far beyond the horizon the median received power, relative to free-space transmission, decreases as the first power of the frequency and as the fifth or sixth power of distance. Beyond 300–400 mi (480–640 km) the decrease in signal changes gradually from a power law to an exponential decrease of about $e^{-0.02d}$, where d is measured in miles. Although useful signals can be obtained at distances of 200 mi (320 km) or more at all frequencies up to the 5000–10,000-MHz limit set by rain attenuation, the optimum frequency range is below 1000 MHz for most applications. Beyond-horizon or tropospheric-scatter circuits require very high power and large antennas; they are, however, economically feasible in isolated areas and for overwater paths. Rapid fading occurs nearly all the time, but with ample fading margin and the use of diversity reception, high quality and high reliability can be obtained. It is possible to transmit 100 or more voice channels on a single radio carrier. It is also possible to provide an acceptable grade of monochrome television on a single link of approximately 150–200 mi (240–320 km). *See* TROPOSPHERIC SCATTER.

Ionospheric propagation. Ionospheric reflections return to Earth useful radio energy at frequencies up to 25 MHz, and possibly up to 60 MHz. Information about the nature of the ionosphere has been obtained mainly by transmitting radar-type signals directly overhead and recording the intensity and time delay of the echoes returned from the ionized layers. At night all frequencies below the critical frequency of 2–4 MHz (for vertical incidence) are returned to Earth with a received power that is close to the value that would be expected in free space for the round-trip distance. During the day the critical frequency is two to three times greater than the corresponding night value. This apparent increase in useful frequency range is largely offset by the strong daytime ionospheric absorption, which reaches a maximum in the 1–2-MHz range at about noon. The difference between day and night transmission means that most sky-wave circuits require two or more frequencies for reliable service covering all hours.

For oblique incidence the maximum usable frequency (MUF) is greater; however, even at the longer distances it does not exceed 3–3.5 times the critical frequency for vertical incidence at that time, season, and latitude.

The frequencies most suitable for transmission over distances of 1000 mi (1600 km) or more will ordinarily not be reflected at the high angles needed for much shorter distances. As a result, the range of sky-wave transmission usually does not overlap the range of ground-wave transmission. The intermediate region of very weak or undetectable signals is called the skip zone. At frequencies around 1–2 MHz the ground wave and sky wave may overlap, resulting in severe fading in the region where the two signals are comparable in amplitude.

In addition to the diurnal variations, there are systematic changes with season, latitude, and the 11-year sunspot cycle. For example, on a summer night in middle latitudes during a year of low sunspot activity, the MUF for the longest distance may be limited to less than 7–8 MHz. On the other hand, during a year of high sunspot activity, the corresponding MUF for a winter afternoon may be 40 MHz or higher.

In addition to the normal daytime absorption, there is a second type of absorption which can occur either day or night and is particularly troublesome on transmission paths that travel near or through the polar regions. During periods of magnetic storms the auroral zones expand over an area much larger than normal, thereby disrupting ionospheric communications by introducing unexpected absorption. These conditions of poor transmission can last for hours and sometimes for days.

In addition to the regular ionospheric reflections, strong signals sometimes occur at frequencies as high as 60–70 MHz from the E layer, which is a region of relatively high ionization located about 50–70 mi (80–110 km) above the Earth's surface. These reflections are known as sporadic E signals because they are erratic in both time and space.

All ionospheric circuits are subject to rapid multipath fading with echo delays up to several milliseconds. These delays are about 10,000 times as long for tropospheric transmission. As a result of these relatively long delays, uncorrelated selective fading can occur within an interval of a few hundred hertz, and this effect produces the well-known distortion on voice circuits characteristic of shortwave transmission.

Scatter. As the frequency is increased above that normally reflected from the ionosphere, the signal intensity decreases rapidly, but it does not drop out completely. Although the signal intensity is low, reliable transmission can be obtained at frequencies up to 50 MHz or higher and to distances up to at least 1200–1500 mi (1900–2400 km). Such ionospheric-scatter circuits require much higher power and larger antennas than are ordinarily used in ionospheric transmission. Ionospheric-scatter transmission is suitable for continuous transmission of a few telegraph channels or for one-voice circuit, but the useful bandwidth is limited by the severe selective fading that is characteristic of all ionospheric transmission. Ionospheric scatter is apparently the result of reflections from many patches of ionization which occur in the E region.

Momentary peaks of the ionospheric-scatter signal intensity occur every few seconds and seem to be the result of the relatively strong ionization produced by the passage and disintegration of small meteors. Meteor-burst communication systems transmit information only during the peaks and are inoperative when the signal intensity is significantly less than its maximum value. The required transmitter power is much less than for ionospheric-scatter systems, but the intermittent meteor-burst method is limited to discontinuous telegraph-type service of relatively low information capacity.

Radiation belts. Beyond the E and F ionospheric layers are the Van Allen radiation belts and other regions of ionization and magnetic activity, which are not yet clearly understood. These regions extend to distances of several earth radii and, although they have a negligible effect on terrestrial communications, they may affect radio communication to outer space. The associated ionization and magnetic activity are weaker and more variable with time and space than is the case for the better-understood ionospheric layers.

All these effects are influenced strongly by particles and radiation from the Sun. A continuing exploration and measurement of these variations should improve understanding of the solar system. *See* VAN ALLEN RADIATION.

Underground, underwater, and jungles. Radio waves can be transmitted hundreds of feet underground or through jungles and a few tens of feet to submarines under the sea. When the usual radiation component traveling along a straight line between transmitter and receiver is absorbed in the lossy medium, an alternate "up, over, and down" path becomes dominant. This transmission path consists of three parts: (1) from the transmitter up through the absorbing medium to the open air, (2) along and above the Earth's surface to a point over the receiver, and (3) from the surface down through the absorbing medium to the receiving antenna. Under these conditions unconventional antennas and antenna orientations are needed to match the impedance and the direction of the path that provides the strongest signal.

Noise. The usefulness of a radio signal is limited by noise. The noise may be either unwanted external interference or noise originating in the receiver itself. Atmospheric static is caused by lightning or other natural electrical disturbances and is propagated over the Earth by ionospheric transmission. It is generally higher at night than in the daytime and is higher in the warm tropical areas, where storms are frequent, than in the colder northern regions. Atmospheric static is ordinarily predominant at frequencies below a few megahertz, whereas set noise is the primary limitation at frequencies above 200–500 MHz. In the intermediate region the controlling noise depends on location and time of day and may come from human sources, such as the operation of electric switches or automobile-ignition devices.

The very low frequency components of lightning discharges are propagated along lines of magnetic force to the antipodes, giving rise to a phenomenon known as whistlers. Cosmic and solar noise are of considerable interest in the fields of radio astronomy and satellite transmission but ordinarily are not the controlling factors in terrestrial communication. *See* ELECTRICAL NOISE; MICROWAVE; RADIO ASTRONOMY; RADIO BROADCASTING; SCATTERING OF ELECTROMAGNETIC RADIATION; SFERICS; WAVEGUIDE.

Kenneth Bullington

Bibliography. L. Boithias, *Radiowave Propagation*, 1988; K. G. Budden, *The Propagation of Radio Waves: The Theory of Radio Waves of Low Power in the Ionosphere and Magnetosphere*, 1985, paper 1988; J. Griffiths, *Radiowave Propagation and Antennas: An Introduction*, 1987; T. S. Maclean and Z. Wu, *Radiowave Propagation over Ground*, 1993; K. Rawer, *Wave Propagation in the Ionosphere*, 1993; S. Shibuya, *A Basic Atlas of Radio-Wave Propagation*, 1987.

Radioactive beams

Beams of radioactive (unstable) nuclei. In several nuclear physics laboratories, a capability exists to produce such beams and, before these nuclei spontaneously decay, use them to gain insight into the reactions on and structure of nuclei never before accessible. Radioactive beams are particularly useful to study stellar explosions such as novae, supernovae,

and x-ray bursts. These explosions are some of the most catastrophic events in the universe, generating enormous amounts of energy while synthesizing the elements that make up lifeforms and the world. These spectacular explosions involve, and in some cases are driven by, reactions where the atomic nuclei of hydrogen (protons) and helium (alpha particles) fuse with (are captured by) radioactive isotopes of heavier elements to form new elements. The capability to produce beams of radioactive nuclei allows direct measurements of these reactions, providing crucial information needed to theoretically model cataclysmic stellar events and to understand the origin of many chemical elements.

Nuclear reactions in stars. Stars are born when gas and dust in space condenses under its own gravity to a large body with a central density and temperature sufficiently high to ignite fusion reactions between atomic nuclei. The carbon-nitrogen-oxygen (CNO) cycle is a catalytic sequence of nuclear reactions that generates energy, making massive stars shine and balancing the crushing inward pressure of gravity. Carbon, nitrogen, and oxygen "seed" nuclei capture four protons (the fuel) and eject an alpha particle, liberating energy in the process (**Fig. 1***a*). Typical temperatures and densities at the core of a massive star where these nuclear reactions occur may be approximately 10^7 kelvins and 100 grams per cubic centimeter, respectively. Under these conditions, the rates of such nuclear fusion reactions (which have been measured in the laboratory) are extremely slow. Some CNO cycle reactions transform catalytic seed nuclei into radioactive isotopes of different elements, which then spontaneously decay to a stable nucleus in typically a few to a few hundred seconds, well before they will undergo any additional fusion reactions. *See* CARBON-NITROGEN-OXYGEN CYCLES; NUCLEOSYNTHESIS; STELLAR EVOLUTION.

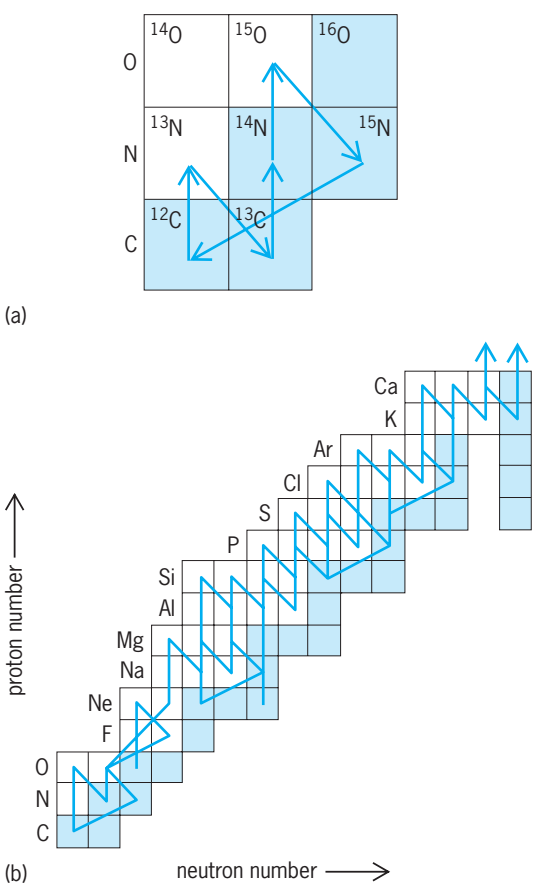

Fig. 1. Sequences of nuclear reactions occurring in stars. The shaded boxes represent individual stable nuclei, the white boxes radioactive nuclei. The arrows represent nuclear reactions leading to transforming one nucleus to another. (*a*) Carbon-nitrogen-oxygen (CNO) cycle, occurring in nonexploding stars. (*b*) Reactions on radioactive isotopes occurring in stellar explosions.

Stellar explosions. During stellar explosions such as novae or x-ray bursts, a different sequence of nuclear reactions occurs. Nova explosions are thought to occur in binary star systems (two stars orbiting each other) when one star of the pair—a cold compact white dwarf star—is able to gravitationally pull material (mainly hydrogen) off the surface of its close binary companion in a phenomenon known as accretion. This situation is not uncommon: there are approximately 40 novae in the Milky-Way Galaxy each year. After approximately 10,000 years, enough hydrogen has collected, or accreted, on the white dwarf surface to raise the temperatures and densities to 10^8 kelvins and 100–10,000 grams per cubic centimeter, respectively. These conditions are sufficient to ignite nuclear fusion reactions between the accreted hydrogen and the carbon and oxygen (or in some cases oxygen, neon, and magnesium) heavy nuclei which make up the white dwarf. The fusion reactions generate energy, which raises the temperature of the accreted layer, which exponentially increases the rates of all the fusion reactions. This becomes a runaway thermonuclear explosion where the temperatures may reach $3\text{–}4 \times 10^8$ kelvins and the energy generated by the nuclear reactions causes ejection of a sizable fraction of the accreted layer off the surface of the white dwarf. The energy released in these enormous thermonuclear explosions can be up to $10^{31}\text{–}10^{38}$ joules, in a thousand seconds—10^{24} times the energy released in powerful thermonuclear explosions on Earth.

Since nuclear fusion reaction rates increase exponentially with increasing temperature, they occur approximately 10^{16} times faster in stellar explosions than in a normal star. This is so fast that protons can be captured by radioactive isotopes before they have a chance to spontaneously decay. Sequences of nuclear reactions occurring in novae are therefore different from those in ordinary stars (Fig. 1*b*). Heavier nuclei, perhaps as high as mass 40 (calcium) and beyond, can be assembled in novae, ejected into space, and eventually incorporated into new stars and planets; even human bodies are made up of this "stardust." The high rate of nuclear reactions in novae also means energy is generated faster, influencing many aspects of the explosion. A similar situation is thought to occur in x-ray bursts, where the accretion from a companion star onto the surface of a neutron star occurs. Here, the temperatures can peak at more than 10^9 kelvins, the densities may reach as high as

10^6 grams per cubic centimeter, and isotopes up to mass 100 (tin) and beyond may be synthesized. *See* CATACLYSMIC VARIABLE; NOVA.

Measurements. Until recently, direct measurement of the fusion reactions occurring in stellar explosions was impossible. This is because beams of radioactive nuclei were not available, and targets of these nuclei would spontaneously decay before a measurement could be made. Computer models of nuclear burning in stellar explosions therefore had to rely on theoretical estimates of many crucial reaction rates. In some cases, these estimates have been shown to be incorrect by many orders of magnitude. Because of these uncertainties, critical comparisons of theory to astrophysical observations to determine the temperatures, densities, and duration of these events, as well as the origin of the elements in the world, were difficult to make.

The availability of beams of radioactive nuclei in nuclear physics laboratories has enabled experimental determinations of the nuclear reactions that drive stellar explosions. One approach to radioactive beam production, pioneered at CERN (Switzerland), is the isotope separator on-line (ISOL) technique, used at Louvain-la-Neuve (Belgium) and at Oak Ridge National Laboratory (United States) [**Fig. 2***a*]. One accelerator bombards a target with a beam of stable nuclei, and a small number of the radioactive atoms of interest are produced through nuclear reactions. These atoms are transported, by various techniques, including thermal diffusion, to an ion source where they are ionized (removing or adding electrons to give atoms an electrical charge) and extracted. The radioactive ions are then mass-separated from other ions and accelerated to energies needed for nuclear physics experiments by a second accelerator. The ISOL technique can produce very high beam qualities, purities, and intensities; the disadvantages are that only a few radioactive beam species can be generated from each combination of production target and primary beam, and that beams with short lifetimes (less than 1 s) are difficult to produce. *See* ION SOURCES; MASS SPECTROSCOPE.

To measure the fusion rate between radioactive heavy ions and hydrogen nuclei, a number of techniques can be used once a radioactive beam is produced. First, a beam of radioactive nuclei at a specific energy is directed at a thin (gas or foil) target containing hydrogen or helium. A direct measurement technique then involves counting the products of the fusion reaction (nuclei and protons) in sophisticated detection systems (**Fig. 3**). One effective system consists of a mass separator followed by charged particle detectors. By varying the energy of the radioactive beam, the probability of fusion as a function of energy is determined. Combined with a calculation of the relative energy distribution of the particles at a given temperature in a star, the fusion rate is determined as a function of temperature. Indirect techniques are also used to determine reaction rates, such as measuring the scattering of the heavy-ion radioactive beam off protons or alphas to discover resonances that may enhance the fusion reaction. Reactions that have been studied with ISOL beams include the fusion of the unstable ^{13}N, ^{17}F, ^{18}F, and ^{19}Ne nuclei with protons and ^{18}Ne with alphas. These reactions are all important for understanding element production in nova explosions.

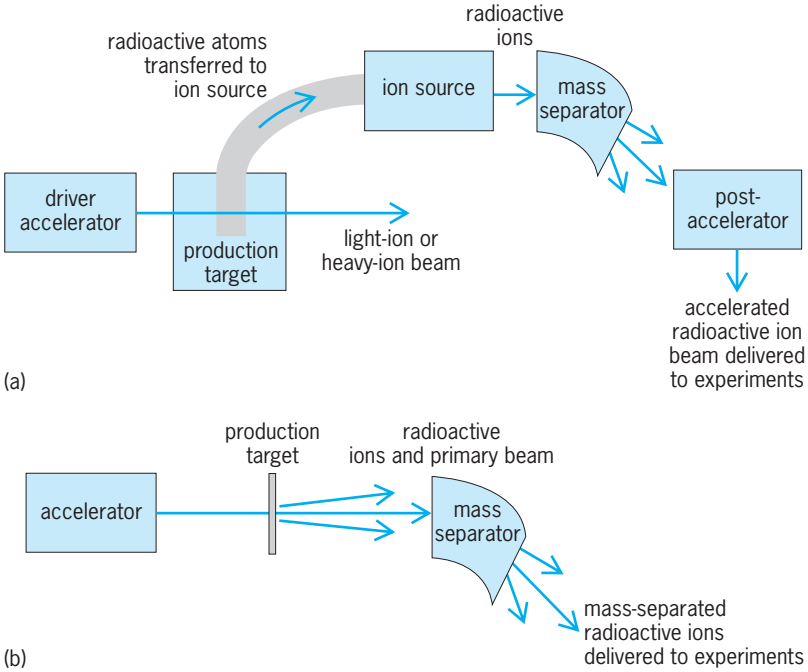

Fig. 2. Techniques for producing beams of radioactive nuclei. (*a*) Isotope separator on-line (ISOL) technique. (*b*) Projectile fragmentation technique.

A complementary radioactive beam production technique is projectile fragmentation, used at Michigan State University (United States), RIKEN (Japan), GANIL (France), and GSI (Germany) [Fig. 2*b*]. When a high-energy beam of stable heavy ions passes through a thin target, the beam particles (projectiles) can break up into fragments—some of which are the radioactive isotope of interest. The desired fragments are then mass-separated from other ions and steered toward a target to undergo the reaction of interest. The projectile fragmentation technique can produce beams of very short lifetimes (10^{-6} s or less), and the same setup can be used to produce many different

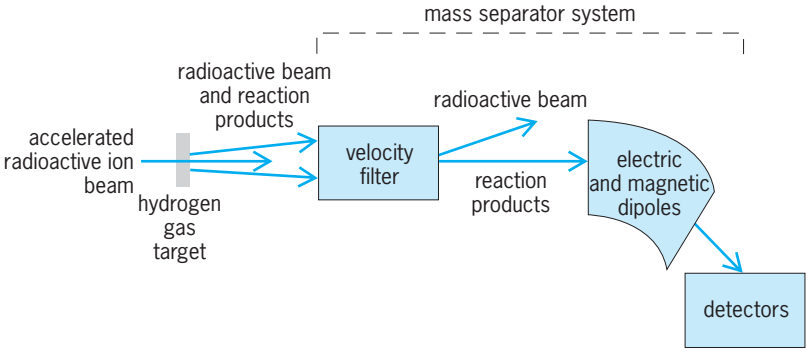

Fig. 3. Technique to measure a thermonuclear fusion reaction occurring in stellar explosions. A radioactive beam bombards a target with hydrogen (protons), and the products of a fusion reaction (as well as unreacted beam particles) enter a mass separator. The products of the thermonuclear reaction are steered onto a detector for counting, while the unreacted beam particles are deflected away.

beam species; the disadvantages are that high beam quality, purity, and intensity are difficult to obtain. Reactions studied with projectile fragmentation beams include the fusion of ^{7}Be and ^{13}N with protons, both by measuring the time-reversed reactions (for example, ^{8}Be breaking up to a ^{7}Be plus a proton). Both ISOL and projectile fragmentation techniques have also been used to measure the properties (masses, spins, parities, energy levels, and decay properties) of nuclei that are important in astrophysics calculations. Other techniques for producing radioactive beams have been used at Notre Dame University and Argonne National Laboratory in the United States. All together, these measurements enable nuclear physicists to directly measure some of the nuclear reactions that power cataclysmic stellar explosions. *See* NUCLEAR FUSION; NUCLEAR REACTION; PARTICLE ACCELERATOR.

Michael S. Smith

Bibliography. D. Arnett, *Supernovae and Nucleosynthesis*, Princeton University Press, 1996; D. D. Clayton, *Principles of Stellar Evolution and Nucleosynthesis*, University of Chicago Press, 1983; C. E. Rolfs and W. S. Rodney, *Cauldrons in the Cosmos: Nuclear Astrophysics*, University of Chicago Press, 1988.

Radioactive fallout

Whenever radioactive materials become airborne, either from a nuclear device detonation or from a nuclear release accident, the resultant contaminated atmospheric plume will ultimately return radioactivity to the Earth's surface. Material settling from the radioactive plume and its subsequent surface deposition is known as radioactive fallout.

Radioactivity. Radioactive materials consist of unstable atoms which emit gamma rays, beta particles, or alpha particles. These emissions—rays and particles—are unique in that they cause ionizations in neighboring atoms. The energy of the emitted radiation and subsequent ionizations can be a cause of concern if absorbed by living systems. *See* ALPHA PARTICLES; BETA PARTICLES; GAMMA RAYS; RADIOACTIVITY.

Nuclear bombs. Nuclear explosions (that is, atomic bombs) depend on the very rapid uncontrolled fissioning of a critical mass of uranium-235 (^{235}U) or plutonium-239 (^{239}Pu). Atomic fission occurs when a ^{239}Pu or ^{235}U atom absorbs an incoming neutron, causing the atom's nucleus to become unstable and to split into two or more parts, called daughters or fission products, each usually radioactive. In addition, the fission reaction releases heat and neutrons. The latter interacting with adjacent plutonium or uranium can induce further fissions in a very short time span (microseconds). When uncontrolled, this chain reaction causes an explosion. The process releases massive amounts of energy, usually measured in kilotons (thousands of tons) of TNT explosive equivalent. Fission bombs can release thousands of kilotons of energy, and the larger thermonuclear or hydrogen bombs can release millions of tons of TNT equivalent (that is, megaton explosions). *See* NUCLEAR EXPLOSION; NUCLEAR FISSION.

Atomic power. At nuclear power plants, the neutron concentration in the fuel is carefully regulated, and the rate of nuclear fission is slowed in a controlled chain reaction. Control is achieved by inserting rods of neutron-absorbing material, which reduces the flux of free neutrons available for further interaction with the fuel. The fissionable uranium in fuel is more dilute than that in a bomb, and the fissioning rate is carefully controlled. Heat from fission is extracted by the nuclear reactor's coolant to make the steam which is converted to electric power, or the coolant is used to generate electricity through direct conversion turbines. *See* CHAIN REACTION (PHYSICS); NEUTRON; NUCLEAR POWER.

Fission products. Fissioning of uranium and plutonium produces isotopes of about 70 different atoms; each atom may have several different isotopic forms. Examples of daughter products are the isotopes of elemental strontium, which are efficiently produced in nuclear fission. The isotope strontium-90 (^{90}Sr) has a 28-year half-life, while ^{89}Sr is produced in slightly higher concentrations but has a half-life of only 50 days. In addition to the differing half-lives, each isotope emits a unique radiation spectrum. The two strontium isotopes emit beta particles of different energies. Other isotopes such as cesium-137 (^{137}Cs, half-life 30 years) and iodine-131 (^{131}I, half-life 8 days) emit both beta particles and gamma rays. *See* CESIUM; IODINE; ISOTOPE; STRONTIUM.

Other radioactive materials. Spent nuclear fuel contains mainly its original uranium, an inventory of fission products, and small amounts of activation products and actinide elements. [Activation products are stable atoms made radioactive (activated) by neutron bombardment in and near the nuclear fuel.] Atom bomb debris also contains fission products, activation products, and unfissioned uranium or plutonium, all of which are radioactive. *See* ATOMIC BOMB; NUCLEAR FUELS.

Atmospheric fallout. Following a major destructive disassembly of a nuclear reactor, such as at Chernobyl, Ukraine, in April 1986, some of the fuel can be released in the form of particles and gases. If the accident is large enough, a fraction of the release becomes airborne, as was the case with Chernobyl. In this disaster, radioactivity injected high into the atmosphere spread over and covered most of the Northern Hemisphere. A surface or air burst of an atomic weapon produces the infamous "mushroom cloud" containing radioactive materials which also can be transported for long distances. *See* NUCLEAR REACTOR.

Global fallout. Soon after the end of World War II, the United States and the Soviet Union began an atomic weapons arms race. France, the United Kingdom, and China joined in, and many aboveground bomb tests were executed. Within 10 years, the amount of radioactivity lofted into the atmosphere began to reach alarming levels, and political pressure increased to halt aboveground atomic-bomb testing. In 1962, a limited test ban treaty forbidding

atmospheric nuclear weapons testing was signed by most nations possessing atomic weapons. Fallout levels began to drop, and today are only a few percent of their peak concentrations in 1962–1964. Most of the atmospheric testing was halted after 1962, but because some of the smallest particles were lofted to the highest altitudes and settle back to Earth at very slow rates, atmospheric fallout is still present. Most of the nuclear detonations occurred in the Northern Hemisphere; in Nevada and the Pacific for the United States, in Kazakhstan for the Soviet Union, in Lop Nor for China, and in Algeria for France. In the Southern Hemisphere, a few tests were conducted by the United Kingdom in Australia's desert and by France on some of its South Pacific islands. Tracking the radioactive clouds from atmospheric tests yielded valuable information on the dispersion patterns and particle size distributions of fallout. Tropospheric and stratospheric concentrations generally followed a westerly trajectory. Little Northern Hemisphere fallout migrated to the Southern Hemisphere; the traces generally stayed in the latitudes in which they were first injected. *See* STRATOSPHERE; TROPOSPHERE.

Local fallout. Most of the fallout from early atomic weapons testing consisted of larger, heavier particles and those smaller particles which agglomerated onto dust particles. These particles settled out close to their source and were at dangerous concentrations adjacent to detonation sites. During the early tests in Nevada, plumes of radioactivity containing mainly fine particles and volatile materials rather than the heavier particulate fractions, blanketed much of the United States. Dose reconstructions of the fallout plumes have shown that many areas of the United States received significant radioactive iodine exposures. Some exposures were high enough to have caused an increase in thyroid cancer risk. Tests at Semipalatinsk, Kazakhstan, produced similar fallout clouds.

A chemical explosion in 1957 (the "Kyshtym accident" at a nuclear materials processing facility in Cheliabinsk in the southern Ural Mountains of Russia) ejected a large, highly radioactive plume of long-lived radionuclides containing about 20 million curies, mainly ^{90}Sr and ^{137}Cs. The resultant "East Urals Radioactive Trace" (EURT) contaminated 25,000 km^2 (10,000 mi^2) of land, exposed about a quarter million people, and caused thousands to be relocated. In 1993, a similar but much smaller explosion in the Tomsk-7 nuclear materials processing facility near Tomsk in Siberia contaminated about 250 km^2 (100 mi^2).

Fallout factors. Atmospheric fallout can be scavenged by rainfall. Wet deposition, involving washing out of atmospheric fallout, can increase local deposition patterns. This was the case following the Chernobyl accident, where local rainfall in Belarus, Ukraine, and Russia washed high concentrations of radioactive iodine and cesium out of the plume and onto the spring pasture. Radioactive iodine and cesium are relatively volatile and were more easily "boiled" out of Chernobyl's burning core. Cesium is a congener of potassium and therefore is fairly uniformly distributed throughout the body once inhaled or ingested. The result is a whole-body radiation dose. Furthermore, the energetic gamma-ray emission from ^{137}Cs adds a source of external radiation from surface deposits on the ground. These factors, in addition to its long half-life and relatively high concentration, make ^{137}Cs the major long-term contamination concern from fallout. For example, although the Chernobyl accident occurred in 1986, precautions must still be taken against potential intake doses of ^{137}Cs: inhalation doses can occur when burning wood from contaminated trees, and consuming mushrooms grown in contaminated forests delivers an ingestion dose.

Resuspension of surface radioactivity concentrations can cause secondary local radioactive fallout. Simple vehicle traffic over or next to a contaminated surface causing particle resuspension is one example. A more extreme example was produced following a prolonged drought in the Ural Mountains of Russia, which lowered the water level in a radioactivity storage pond. A sudden strong wind storm in 1967 resuspended fallout from the dry lake bed, which contained over a half million curies (1.8×10^{16} becquerels), adding further contamination to the 1957 EURT.

Radiation risks. Levels of ^{131}I in the plumes of radioactive fallout are of particular concern. With an 8 day half-life and a strong beta- and gamma-ray emission, this radionuclide concentrates almost exclusively in the thyroid gland. Recent dose reconstructions in the United States show that the 1950s and 1960s fallout radiation doses from ^{131}I were large enough to have increased the risk for thyroid cancer, especially in children.

In the 1950s, large-scale Pacific tests deposited significant concentrations of weapons fallout on some of the Marshall Islands. Seventeen of the nineteen children under 10 years of age received thyroid doses of about 1000 rads (10 grays) and were later found to have developed thyroid lesions (one became malignant, the others were treated surgically).

Following its release into the environment during the Chernobyl accident, radioactive iodine was inhaled and, more importantly, incorporated into milk following fallout onto dairy herd pasture. Subsequent consumption of this milk caused significant thyroid doses. Over a thousand children living downwind of Chernobyl, most of whom were infants at the time of the accident, have developed thyroid cancer, when only a dozen cases might normally have been expected. This was probably caused by the consumption of large amounts of contaminated milk by infants, and the small size of the thyroid at that stage of development. *See* THYROID GLAND.

In 1955, the United Nations established the Scientific Committee on the Effects of Atomic Radiation, due to concern over possible risks from fallout. It issues comprehensive reports about every 5 years and has collected and documented the world's literature on radioactive fallout and, more recently, on

possible radiation consequences. *See* RADIATION INJURY (BIOLOGY). Marvin Goldman

Bibliography. L. R. Anspaugh et al., The global impact of the Chernobyl reactor accident, *Science*, 242:1513–19, 1988; M. Eisenbud, *Environmental Radioactivity*, 3d ed., 1987; M. Goldman, The Russian radiation legacy: Its integrated impact and lessons, *Environ. Health Perspect.*, 105:1385–1391, 1997; J. N. Stannard, *Radioactivity and Health*, 1988; United Nations, *Sources, Effects and Risks of Ionizing Radiation*, 1988.

Radioactive minerals

Minerals that contain uranium (U) or thorium (Th) as an essential component of their chemical composition. Examples are uraninite (UO_2) or thorite ($ThSiO_4$). There are radioactive minerals in which uranium and thorium substitute for ions of similar size and charge. There are approximately 200 minerals in which uranium or thorium are essential elements, although many of these phases are rare and poorly described. These minerals are important, as they are found in ores mined for uranium and thorium, most commonly uraninite and its fine-grained variety, pitchblende, for uranium. Thorite and thorogummite are the principal ore minerals of thorium. Minerals in which uranium and thorium occur in trace amounts, such as zircon ($ZrSiO_4$), are important because of their use in geologic age dating. The isotope uranium-238 (^{238}U) decays to lead-206 (^{206}Pb); ^{235}U decays to ^{207}Pb; ^{232}Th decays to ^{208}Pb; thus, the ratios of the isotopes of uranium, thorium, and lead can be used to determine the ages of minerals that contain these elements. *See* DATING METHODS; GEOCHRONOMETRY; LEAD ISOTOPES (GEOCHEMISTRY); THORITE; URANINITE.

Classification. Typically, the uranium and thorium minerals are classified according to their principal anions: oxides, carbonates, sulfates, molybdates, phosphates, vanadates, and silicates. Thorium exists in a single valence state (Th^{4+}); uranium, in two valences (U^{4+} and U^{6+}). The U^{6+} ion exists as the uranyl ion (UO_2^{2+}), which is soluble; the U^{4+} ion is much less soluble. The double oxidation state of uranium results in the dissolution (on oxidation), transport (under oxidizing conditions), and precipitation (under reducing conditions), which leads to the formation of uranium deposits. The roll-front deposits in the sandstones of the Colorado Plateau are an example. *See* ORE AND MINERAL DEPOSITS.

Properties. The multiple oxidation states of uranium lead to great complexity in the structures of the radioactive minerals. Those in which the principal oxidation state is IV (U^{4+} and Th^{4+}) are dark in color or opaque and do not fluoresce in ultraviolet light. The lower oxidation states are characteristic of minerals formed under reducing conditions, such as pegmatites or hydrothermal deposits. The U^{6+} minerals comprise the uranates and the uranyls. The uranates ($M_2^{+}U_x^{6+}O_{3x+1}$) are red, orange, or brown, with a rare orange or red fluorescence; the uranyls contain the $(UO_2)^{2+}$ ion. These compounds are bright lemon yellow, yellow, orange, or green and commonly fluoresce in ultraviolet radiation.

The U^{6+} phases form as low-temperature alteration products in near-surface deposits. Weathering may lead to multicolored sequences of these minerals up to tens of meters in thickness. The classic uranium deposits of the Shaba District of Zaire are typical. In addition to the uranyl ion in these highly oxidized phases, water of hydration is common. The great number of possible phases, the variable uranium oxidation states, and the water of hydration make these phases extremely difficult to identity. Some exist in a mixed oxidation state. An important phase in this regard is ianthinite [$(UO_2)(5UO_3)(10H_2O)$], but even apparently simple phases such as uraninite actually have complex formulas.

Radioactive decay of ^{238}U, ^{235}U, and ^{232}Th results in the emission of alpha particles (helium nuclei), alpha-recoil nuclei, beta particles (electrons), and gamma radiation (x-rays). Although the beta particles and gamma radiation can cause discoloration due to ionization, the alpha-decay event leads to extensive atomic displacements (approximately 1000 displaced atoms per decay event), until in some minerals the structure becomes amorphous, the metamict state. Some phases such as zircon are susceptible to radiation damage and may become completely amorphous. Other phases, such as uraninite, retain their crystallinity by annealing, despite doses that may be as high as 50 displacements per atom. Minerals with U^{6+} are seldom metamict, as they are generally of secondary origin. Metamict minerals are glassy in appearance, are optically isotropic, and have a conchoidal fracture, reduced hardness (25%), lower density (17%), and an increased susceptibility to chemical alteration. The changes in physical and chemical properties that are due to alpha-decay events are of interest because of potential radiation effects on nuclear waste forms used in radioactive waste disposal. *See* ALPHA PARTICLES; AMORPHOUS SOLID; BETA PARTICLES; GAMMA RAYS; METAMICT STATE; PLEOCHROIC HALOS; RADIOACTIVE WASTE MANAGEMENT; URANIUM; THORIUM. Rod Ewing

Bibliography. B. C. Chakoumakos et al., Alpha-decay—Induced fracturing in zircon: The transition from the crystalline to the metamict state. *Science*, 236:1556–1559, 1987; R. J. Finch and R. C. Ewing, The corrosion of uraninite under oxidizing conditions, *J. Nucl. Mat.*, 190:133–156, 1992; C. Frondel, *Systematic Mineralogy of Uranium and Thorium*, Geol. Surv. Bull. 1064, 1958; E. W. Heinrich, *Mineralogy and Geology of Radioactive Raw Materials*, 1958.

Radioactive tracer

A radioactive isotope which, when injected into a chemically similar substance or artificially attached to a biological or physical system, can be traced by radiation detection devices. Many problems in biology, medicine, and industrial engineering not amenable

to other approaches can be solved by the use of these tracers. *See* RADIOACTIVITY; RADIOACTIVITY AND RADIATION APPLICATIONS; RADIOISOTOPE; RADIOISOTOPE (BIOLOGY).

Uses in industry. Radioactive tracers are useful in industrial applications, such as the testing of wear and corrosion in mechanical components. Piston ring wear within an engine can be determined accurately by transforming some of the carbon within the ring into carbon-14. This is done by placing the ring in an intense neutron flux. The piston ring is then installed in the engine and, after a period of use, samples of the engine oil are removed and tested for carbon-14 content. In this way amounts of wear much smaller than could be found by weight measurements and other techniques can be determined.

On a larger scale, the entire operation of an industrial plant can be tested and controlled by tracer techniques. In one industrial survey, sulfur-35, prepared in an appropriate chemical form, was introduced into a batch of coke entering an iron-smelting plant. The fate of the sulfur within the plant and its presence in the iron were studied by taking samples and measuring the iron's sulfur-35 content. It is entirely possible that similar techniques could be used in the routine operation and monitoring of many industrial processes. Such applications have been restricted in the past chiefly because of the difficulty of hazard control, but with more sensitive detectors such procedures may become commonplace.

Uses in biology. The simplest radioactive tracer studies consist of the tagging of a biological entity with a radioactive isotope (radioisotope). The entity is then tracked by following the radiation from the isotope. Such a method has been used to study the nocturnal behavior of bats. Small sources of cobalt-60 were attached to the legs of the bats. A radiation detector connected to an automatic recorder registered the presence or absence of the bats from their nests at night. In a similar manner, cockroach migration in an urban sewer system has been studied, as well as behavior of various insects and small animals.

The operation becomes more complex when a large number of biological particles are labeled, for example, in the tagging of red blood cells or bacteria. When the labeled substance is injected into an animal, it is impossible to follow the individual labeled particles, but their average movement can be tracked by observations of the radiation. In this way it is possible to measure the average lifetime of a red blood cell or the diffusion rate of bacteria.

Finally, a radioisotope of a particular element can be used to tag that element. Phosphorus-32 can be introduced into the soil where a plant is growing, and the amount of phosphorus absorbed and its distribution throughout the plant can be studied.

In one of the earliest tracer studies, G. von Hevesy used radioactive lead to trace the pathways of lead metabolism in plants. It was essential in this study, as in any other tracer experiments of this type, that the chemical behavior of the radioisotope be identical to that of the stable isotope in order for it to circulate normally within the biological entity being studied. For most isotopes this is essentially true.

In most biological tracer experiments, the radio-isotope is introduced into the system and its radiation subsequently measured with Geiger-Müller counters or scintillation detectors. Extremely soft (low-intensity) radiations can be detected by the use of photographic film. In a typical experiment using this technique, part of a plant containing a radioactive isotope is sectioned and placed next to the film. The radiation from the plant darkens the film, and the pattern on the film shows the distribution of the radioisotopes in the plant. Such pictures are called autoradiographs. Autoradiographs can be taken of cellular and even subcellular particles. Studies have been made of the reproduction of chromosomes using such techniques. Similar studies can be carried out in animals. *See* AGRICULTURAL SCIENCE (ANIMAL); AUTORADIOGRAPHY; GEIGER-MÜLLER COUNTER; PHOTOSYNTHESIS; SCINTILLATION COUNTER.

Gordon L. Brownell

Uses in medicine. A radioactive atom can be attached to a molecule or more complex substance, which can then be used to examine a chemical reaction in a test tube, or it can be administered to a patient by ingestion or injection and subsequently be incorporated into a biochemical process. The radioactive emissions from the radioactive atom can be used to track (trace) the behavior of the labeled molecule or substance in biological processes by means of medical imaging, utilizing techniques such as positron emission tomography (PET) or single-photon-emission computed tomography (SPECT). *See* MEDICAL IMAGING.

The tracer principle makes it possible to characterize the dynamic state of biological systems, including the living human body. Radioactive tracers revolutionized biochemistry and are widely used in the practice of medicine as well as in biomedical research.

The branch of medicine that uses radioactive tracers in the care of patients is called nuclear medicine. Radiotracers of practically every element can be produced in nuclear reactors or cyclotrons. *See* NUCLEAR MEDICINE.

An example of an early application of the tracer principle in medicine is the use of radioactive iodine (iodine-131), discovered in the late 1930s, to detect abnormalities in thyroid gland metabolism. The rate of accumulation of radioiodine by the thyroid can be measured with a radiation detector, making it possible to detect abnormalities in iodine metabolism, such as reduced thyroid function (hypothyroidism) or increased function of the thyroid (hyperthyroidism). Another early use of biological tracers involved measuring the life-span of red blood cells, platelets, and white blood cells. *See* THYROID GLAND DISORDERS.

In nuclear medicine, radioactive tracers are measured within the body by using radiation detectors directed at the body from the outside. Nuclear medicine also uses measurements of radioactive tracers in body fluids (such as blood and urine) in test

tubes. The emission of gamma rays from radioactive tracers within the human body makes possible the measurement of regional function and biochemistry in practically every organ.

In a nuclear medicine procedure, small amounts of radioactive chemicals are injected into the patient's bloodstream, typically via an arm vein. Then the amount of radioactivity in different regions of the body or within an organ is measured by means of externally placed radiation detectors. The selection of a specific radioactive tracer depends on the particular bodily function and regional biochemical reactions that are to be studied.

Radioactive tracers provide an entirely new way of looking at disease. They can reveal characteristic patterns of blood flow, biochemical abnormalities, and the location, severity, and extent of disease in the brains of patients.

Radioactive tracers play a major role in the study not only of the brain but of practically every other organ, including the heart. Radioactive tracers are used as part of the diagnostic process. Three radionuclides—carbon-14, tritium (hydrogen-3), and phosphorus-32—remain the backbone of modern biomedical sciences, especially in the fields of biochemistry, molecular biology, and genetics.

The cell remains the basic unit of biology, but with radioactive tracers one can examine intracellular chemistry as well as the means of communication among cells and organs. The exquisite sensitivity with which radioactivity can be measured (there are instruments sensitive enough to detect the decay of a single radioactive atom) makes it possible to quantify exceedingly low concentrations of chemical substances in the body, not only in body fluids such as blood but also in the organs themselves.

To carry out such studies, persons are administered radioactive tracers that take part in specific chemical reactions in different organs of the body. The choice of the radioactive substance is determined by the chemical or physiologic process to be examined. The movement of the tracer is displayed in a series of scans or images that reveal regional biochemistry. Positron-emitting radiotracers are usually made in cyclotrons, chiefly by using carbon-11 or fluorine-18 and a PET scanner. Other radioactive tracers are imaged with a SPECT scanner.

Images from PET or SPECT instruments are cross-sectional pictures of the distribution of radiotracers within the body at various times after a radioactive tracer is injected into a vein. These cross sections, or tomographic slices, provide information relating to slices of the person's body. Mathematical models are used to calculate regional concentrations of chemical concentrations and reaction rates, which are then compared from one region of the body to another.

Radioactive tracers reveal abnormalities in the chemical workings of the body, whereas computer tomography (CT) detects abnormalities in anatomy. PET can often reveal disease before anatomical changes show up on a CT scan. Since biochemical problems often occur before anatomical changes, PET can be used to diagnose certain diseases at an earlier stage than was possible in the past.

Henry N. Wagner, Jr.

Radioactive waste management

The treatment and containment of radioactive wastes. These wastes originate almost exclusively in the nuclear fuel cycle and in the nuclear weapons program. Their toxicity requires careful isolation from the biosphere. Their radioactivity is commonly measured in curies (Ci). The curies, chosen to approximate the activity of 1 gram of radium-226 (^{226}Ra), is equal to 3.7×10^{10} becquerels. The becquerel (Bq), the SI unit of activity (radioactive disintegration rate), is the activity of a radionuclide decaying at the rate of one spontaneous nuclear transition per second. Considering its toxicity, the curie is a rather large unit of activity. A more appropriate unit is the microcurie ($1\ \mu\text{Ci} = 10^{-6}$ Ci), but the nanocurie ($1\ \text{nCi} = 10^{-9}$ Ci) and picocurie ($1\ \text{pCi} = 10^{-12}$ Ci) are also frequently used. A common unit of nuclear generating capacity is the gigawatt (electric) [GW(e)], equal to 10^9 watts of electric power, as opposed to thermal power. *See* UNITS OF MEASUREMENT.

Radioactive wastes are classified in four major categories: spent fuel elements and high-level waste (HLW), transuranic (TRU) waste, low-level waste (LLW), and uranium mill tailings. Minor waste categories, such as radioactive gases produced during reactor operation, radioactive emissions resulting from the burning of uranium-containing coal, or contaminated uranium mine water, will not be discussed.

Spent fuel elements arise when uranium is fissioned in a reactor to generate energy. The fuel elements needed for the production of 1 GW(e)-year of electrical energy contain 40 metric tons (44 short tons) of uranium; the spent fuel contains 1 metric ton (2200 lb) of fission-product nuclides, and also transuranic nuclides such as plutonium and americium produced by neutron capture in uranium nuclei. Spent fuel elements arising in the civilian energy program will not be chemically reprocessed. In the United States, weapons-grade plutonium for nuclear explosives is produced in special reactors. In order to extract this plutonium, the spent fuel must be chemically reprocessed. The resulting high-level waste contains most of the fission products and transuranic elements, including residual plutonium. Transuranic waste, arising mainly during this reprocessing, is now defined as solid material contaminated to greater than 100 nCi/g (3.7×10^6 Bq/kg) with certain alpha-emitting radionuclides. (Prior to 1984, this limit had been set at 10 nCi/g.) Uranium mill tailings are the residues of the chemical extraction of uranium from the ore. Finally, low-level waste is a very broad category of wastes, covering almost every form of radioactive waste not falling into the other categories. Projections of the radioactivity and the volume of the different waste categories (**Figs. 1** and **2**) assume no new nuclear plants will be ordered, leading to an installed civilian nuclear

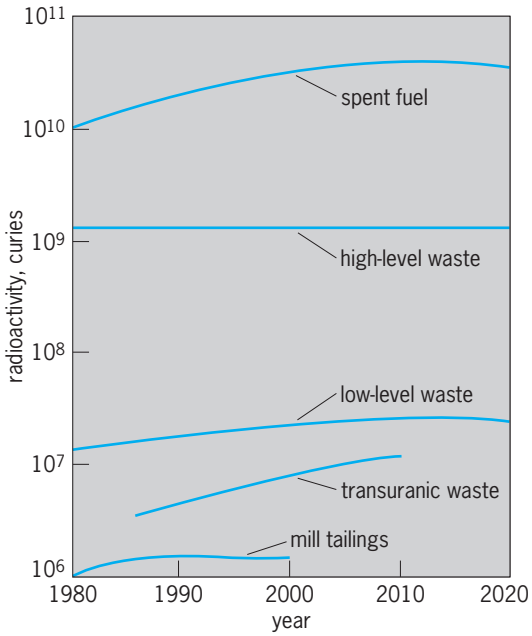

Fig. 1. Past and expected future accumulations of the four major waste categories: high-level waste and spent fuel, transuranic waste from the weapons program, low-level waste, and uranium mill tailings. The radioactivity in the mill tailings is the sum of the radioactivity of all uranium daughters. (*After U.S. Department of Energy, Spent Fuel and Radioactive Waste Inventories, Projections, and Characteristics, Rep. DOE/RW-0006, Rev. 5, Nov. 1989*)

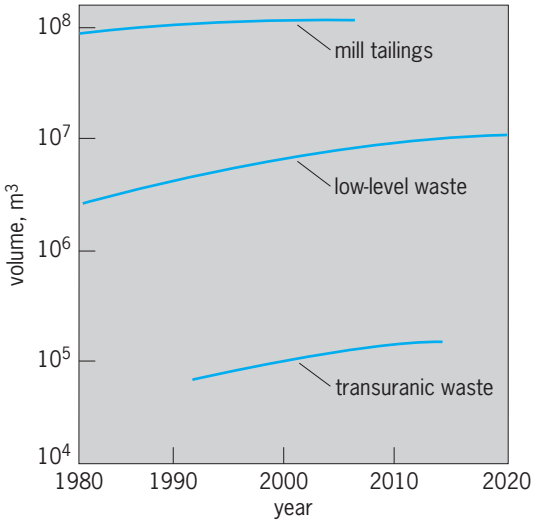

Fig. 2. Volume of the accumulated wastes shown in Fig. 1, except for that of the high-level waste and spent fuel, which is very small (less than 10,000 m^3). 1 m^3 = 35 ft^3.

generating capacity of 102.5 GW(e) in the year 2000, and 51.6 GW(e) in 2020, as the plants reach the end of their lifetimes [1989: 97.5 GW(e)]. Considerably more radioactivity has been produced in the civilian nuclear energy program than in the weapons program. By the year 2020, the spent fuel is expected to consist of 75,000 metric tons (83,000 short tons) of uranium containing transuranic and fission-product nuclides. *See* ATOMIC BOMB; NUCLEAR FISSION; NUCLEAR FUEL CYCLE; NUCLEAR FUELS REPROCESSING; TRANSURANIUM ELEMENTS.

Since no practical methods exist to detoxify radioactive nuclides, protection against their harmful radiation must rely on their isolation from the biosphere until their radioactivity has decayed. Because each of the waste categories poses different problems, they will be discussed separately. The management of the radioactive wastes that arise during the dismantling of nuclear facilities will also be discussed.

Spent fuel and high-level waste. Most of the existing radioactivity is contained in this waste (Fig. 1). To identify and compare the major contributors to the radiotoxicity of the different nuclides, the water dilution volume is defined as the volume of water required to dilute these nuclides to acceptable concentrations, according to guidelines for occupational exposure, as specified by the International Commission of Radiological Protection (1979). The water dilution volumes for the most important isotopes contained in spent nuclear fuel have been calculated (**Fig. 3**; for high-level wastes, the results would be

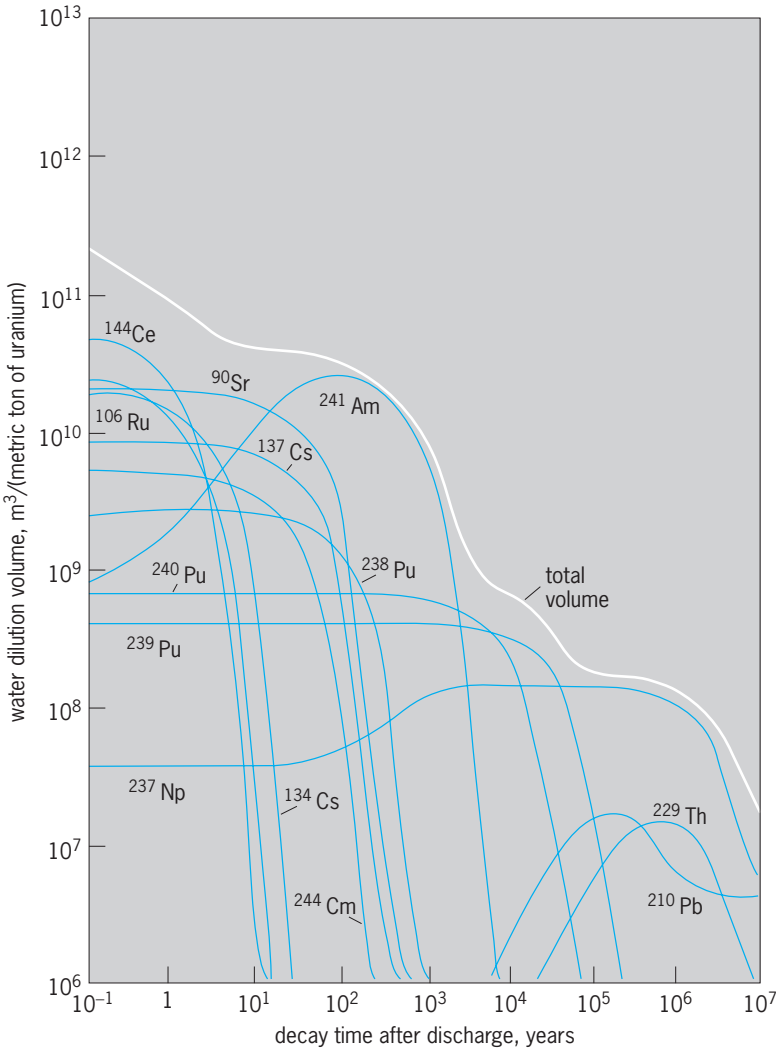

Fig. 3. Water dilution volume, as explained in the text, of spent fuel containing 1 metric ton (2200 lb) of uranium. The total amount of fuel needed to generate 1 GW(e)-year of electric energy is approximately 40 metric tons. The concentration of some of the transuranics increases for some time, as their parent nuclides decay. 1 m^3 = 35 ft^3. (*After Board on Radioactive Waste Management, National Research Council, A Study of the Isolation System for Geologic Disposal of Radioactive Wastes, National Academy Press, 1983*)

qualitatively similar, except for a reduced contribution by transuranics.) For the first 100 years, the toxicity is dominated by the beta- and gamma-emitting fission products [such as strontium-90 (^{90}Sr) and cesium-137 (^{137}137), with half-lives of approximately 30 years]; thereafter, the long-lived, alpha-emitting transuranium elements [for example, plutonium-239 (^{239}Pu), with a half-life of 24,000 years] and their radioactive decay daughters [for example, americium-241 (^{241}Am), with a half-life of 432 years, a daughter of plutonium-241 (^{241}Pu), with a half-life of 13 years] are important. Burial in geologic formations at a depth of 500–1000 m (1600–3200 ft) appears at present the most practical and attractive disposal method, although as of 1994 none of this waste had been disposed of in this way anywhere in the world. Most of the commercial spent fuel is stored in water-cooled basins at the reactor sites. Construction of some interim or emergency storage facilities aboveground for spent fuel has been considered. Such storage would further increase the risks of transportation accidents. The military reprocessed high-level waste is stored in tanks in liquid or solid form. Incorporation of high-level waste into glass blocks suitable for burial was expected to start in the United States in 1996. It is expected that by 2020 only 36% of the existing high-level waste, as measured by its radioactivity, will have been prepared in this form.

Geology as a predictive science is still in its infancy, and many of the parameters entering into model calculations of the long-term retention of the waste in geologic media are questionable. The major single problem is the heating of the waste and its surrounding rock by the radioactive decay heat. This heating can accelerate the penetration of groundwater into the repository, the dissolution of the waste, and its transport to the biosphere. Much effort has been devoted to the development of canisters to encapsulate the spent fuel elements or the glass blocks containing high-level waste, and of improved waste forms and overpacks that promise better resistance to attack by groundwater. The suitability of rock salt, which for 30 years had been widely considered the preferred host rock because of its superior heat dissipative properties and because of its plasticity (which would help to seal cracks), is now questioned because of several shortcomings. The major drawback is that rock salt is a natural resource with a large number of uses. It is impossible to assure that drilling into or through a salt formation housing a nuclear waste repository will never happen. The consequences of such drilling activities add so much to the complexity of predicting the safe containment of the nuclear waste that they seem to outweigh the geophysical advantages that salt might have over other rocks that are not natural resources and not water-soluble.

In the United States, the suitability of a disposal site in tuff, adjacent to the Nevada Nuclear Test Site (Yucca Mountain), is being studied. As planned, this repository could accept all the spent fuel elements, 70,000 metric tons (77,000 short tons), produced up to the year 2020 [used to generate approximately 2000 GW(e)-year of electrical energy]. If approved, the site would not be operational until 2010 at the earliest. *See* TUFF.

Transuranic waste. Although the radioactivity of the transuranic wastes is considerably smaller than that of high-level waste or spent fuel (Fig. 1), the high radiotoxicity and long lifetime of these wastes also require disposal in a geologic repository. Present plans call for disposal in bedded salt at the Waste Isolation Pilot Project (WIPP) in Carlsbad, New Mexico. Suitability for permanent disposal is to be tested over a period of approximately 5 years. During its expected subsequent period of operation, until 2018, the WIPP would accept all transuranic wastes generated up to that time. Waste with less than 100 nCi/g (3.78 Bq/kg) of transuranic elements will be treated as low-level waste.

Uranium mill tailings. Uranium is naturally radioactive, decaying in a series of steps to stable lead. It is currently a rare element, averaging between 0.1 and 0.2% in the mined ore. At the mill, the rock is crushed to fine sand, and the uranium is chemically extracted. The residues, several hundred thousand cubic meters for the annual fuel requirements of a 1-GW(e) reactor, are discharged to the tailings pile (**Fig. 4**). The tailings contain the radioactive daughters of the uranium. The long-lived isotope thorium-230 (^{230}Th, half-life 80,000 years) decays into radium-226 (^{226}Ra, half-life 1600 years), which in turn decays to radon-222 (^{222}Rn, half-life 3.8 days). Radium and radon are known to cause cancer, the former by ingestion, the latter by inhalation. Radon is an inert gas and thus can diffuse out of the mill tailings pile and into the air. It has been estimated by the Environmental Protection Agency that a person living 500 m (1600 ft) away from an unprotected tailings pile containing 280 pCi/g (1×10^4 Bq/kg) of ^{226}R would have a 30% higher chance of lung cancer from the radon gas it produces than the average person. (The average radium concentration in the existing tailings piles is twice as high, and hence, for these piles, the estimate should be doubled.) Ground-water pollution by radium that has leached from the pile has also been

Fig. 4. Partial view of an operating mill tailings pile (Homestake Mining Company, Milan, New Mexico). The pile measures approximately 1000 m × 1000 m (3000 ft × 3000 ft), is 13 m (42 ft) high on average, and is entirely unprotected. In its 1.7×10^7 m^3 (6×10^8 ft^3), it contains 8000 Ci (2.9×10^{14} Bq) of ^{226}Ra. The extracted uranium was enough to generate 100 GW(e)-year of electric energy in a reactor.

observed around tailings piles, but its health effects are more difficult to estimate, since the migration in the ground water is difficult to assess and also highly site-specific. *See* RADIOACTIVITY; RADIUM; RADON; URANIUM.

Although the radioactivity contained in the mill tailings is very small relative to that of the high-level waste and spent fuel, it is comparable to that of the transuranic waste (Fig. 1). It is mainly the dilution of the thorium and its daughters in the large volume of the mill tailings (Fig. 2) that reduces the health risks to individuals relative to those posed by the transuranium elements in the transuranic wastes. However, this advantage is offset by the great mobility of the chemically inert radon gas, which emanates into the atmosphere from the unprotected tailings. New mill tailings piles will be built with liners to protect the ground water, and will be covered with earth and rock to reduce atmospheric release of the radon gas. None of these measures provides protection on the time scales required for the ^{230}Th to decay. Permanent disposal methods, like chemical extraction of the thorium for disposal in a geologic repository, or burial of the tailings in deep mines, are not planned at this time. Because of the enormous volume of the tailings, a permanent disposal—should this be considered necessary—would present serious technical as well as economic problems.

Low-level wastes. By definition, practically everything that does not belong to one of the three categories discussed above is considered low-level waste. This name is misleading because some wastes, though low in transuranic content, may contain very high beta and gamma activity. For example, ion-exchange resins or activated components from reactors, which have radioactivities of hundreds of curies per cubic meter, may even require biological shielding during handling and transport. Of the 130,000 m^3 (4.6×10^6 ft^3) of low-level waste generated in 1988, about 70% originated in the weapons program, 20% in the commercial nuclear fuel cycle, 7% in industry, and 3% for research and medical purposes. The radioactivity contained in the last category was very small, less than 1% of the total radioactivity contained in the low-level waste, and is also mostly short-lived. The radioactivity of the wastes generated for medical purposes alone was less than 0.1%.

The current method of low-level waste disposal is shallow-land burial, which is relatively inexpensive but provides less protection than a geologic repository. [Prior to 1970, a total of 10^5 Ci (3.7×10^{15} Bq) of low-level waste generated in the United States weapons program was also disposed of by ocean dumping and prior to 1983, 1.3×10^6 Ci (4.8×10^{16} Bq) was also injected with grout into hydrofractured shale formations underlying the Oak Ridge National Laboratory. These practices have since been stopped.] At present, there are six major Department of Energy shallow-land burial sites and six commercial sites (only two of which are now operating). While sites in arid environments have generally performed in an acceptable manner, those located in humid environments have commonly not performed as hoped, and the fact that three of the four commercial sites in the eastern United States are no longer operating illustrates the problems they have encountered. It also points to the difficulties to be expected in developing new regional low-level waste disposal sites, which are to be located in the humid northeastern United States, as required by law. Besides the very large volume, the nonuniformity of the waste presents formidable problems for low-level waste disposal. Current programs focus on improved waste forms, better site selection criteria, and engineering improvements to the site in order to restrict releases.

Decommissioning of nuclear facilities. At the end of their lifetime, nuclear facilities have to be dismantled (decommissioned) and the accumulated radioactivity disposed of. Nuclear power plants represent the most important category of nuclear facilities, containing the largest amounts of radioactive wastes, and will therefore be discussed here. These wastes can be grouped in three classes: neutron-activated wastes, surface-contaminated wastes, and miscellaneous wastes. *See* NUCLEAR POWER; NUCLEAR REACTOR.

The neutron-activated wastes are mainly confined to the reactor pressure vessel and its internal components, which have been exposed to large neutron fluences during reactor operation. These components contain significant amounts of long-lived nontransuranic radioactive isotopes such as niobium-94 (^{94}Nb, an impurity in the stainless steel), which emits highly penetrating gamma rays and has a half-life of 20,000 years. These wastes are unacceptable for shallow-land disposal as low-level wastes. Disposal in a geologic repository is envisioned. Surface-contaminated and miscellaneous radioactive wastes derive mainly from faulty fuel pin claddings, which allow radioactive material to escape from the fuel. While the surface contamination can be effectively removed from smooth surfaces, such decontamination will inevitably increase the amount of the miscellaneous wastes (solvents, filters, and so forth). *See* DECONTAMINATION OF RADIOACTIVE MATERIALS.

Delaying the dismantling of the facility, in a procedure called safe storage or entombment, allows much of the radioactivity to decay. While this reduces the dismantling cost, it will shift the burden to future generations. For a decommissioning scenario assuming a 2-year wait after shutdown before dismantling and decontaminating the facility, the cumulative volume and radioactivity for the 68 reactors expected to be shut down between 1989 and 2020 has been estimated to be 8.4×10^5 m^3 (3.0×10^7 ft^3) and 3.3×10^7 Ci (1.2×10^{18} Bq), respectively. Of this volume, 0.6% will contain 97% of the radioactivity (mainly the neutron-activated reactor internals). The remainder, containing 1×10^6 Ci (3.7×10^{16} Bq), will be buried in shallow-land disposal sites, where it will constitute by the year 2020 approximately 10% of the accumulated volume and 3% of the accumulated radioactivity (Figs. 1 and 2). Robert O. Pohl

Bibliography. R. E. Berlin, *Radioactive Waste Management*, 1989; Board on Radioactive Waste Management, National Research Council, *A Study of the*

Isolation System for Geologic Disposal of Radioactive Wastes, 1983; B. W. Burton et al., *Overview Assessment of Nuclear Waste Management*, Los Alamos Sci. Lab. Rep. LA-9395-MS, August 1982; R. D. Lipschutz, *Radioactive Waste: Politics, Technology, and Risk*, 1980; A. G. Milnes, *Geology and Radwaste*, 1985; U.S. Department of Energy, *Spent Fuel and Radioactive Waste Inventories, Projections, and Characteristics*, Rep. DOE/RW-0006, Rev. 9, March 1994; U.S. Environmental Protection Agency, *Draft Environmental Impact Statement for Standards for the Control of By-product Materials from Uranium Ore Processing* (40 CFR 192), Rep. EPA 520/1-82-022, March 1983.

Radioactivity

A phenomenon resulting from an instability of the atomic nucleus in certain atoms whereby the nucleus experiences a spontaneous but measurably delayed nuclear transition or transformation with the resulting emission of radiation. The discovery of radioactivity by Henri Becquerel in 1896 was an indirect consequence of the discovery of x-rays a few months earlier by Wilhelm Roentgen, and marked the birth of nuclear physics. *See* X-RAYS.

On the other hand, nuclear physics can also be said to begin with the proposal by Ernest Rutherford in 1911 that atoms have a nucleus. On the basis of the scattering of alpha particles (emitted in radioactive decay) by gold foils, Rutherford proposed a solar model of atoms, where negatively charged electrons orbit the tiny nucleus, which contains all the positive charge and essentially all the mass of the atom, as planets orbit around the Sun. The attractive Coulomb electrical force holds the electrons in orbit about the nucleus. Atoms have radii of about 10^{-10} m and the nuclei of atoms have radii about 2×10^{-15} m, so atoms are mostly empty space, like the solar system. Niels Bohr proposed a theoretical model for the atom that removed certain difficulties of the Rutherford model. *See* ATOMIC STRUCTURE AND SPECTRA.

However, it was only after the discovery of the neutron in 1932 that a proper understanding was achieved of the particles that compose the nucleus of the atom. The nucleus contains protons that carry positive charge and neutrons with slightly higher mass and zero net electrical charge. These protons and neutrons (called nucleons) are held inside the nucleus by the nuclear force between these particles. This force gives rise to the binding energy of the nucleus, which is the energy required to pull all the protons and neutrons apart. The binding energy makes the mass of a nucleus less than the masses of the Z protons and N neutrons that make up the nucleus. In all radioactive decays the total number of nucleons, $A = Z + N$, before and after the decay is a constant; that is, the number of nucleons is conserved. Differences in binding energies produce differences in masses which in turn determine what type of radioactive decay can occur. Radioactive decay occurs when the masses of all the particles after the decay are less than the mass of the original radioactive nucleus. *See* NEUTRON; NUCLEAR BINDING ENERGY; NUCLEAR STRUCTURE; NUCLEON; PROTON.

In 1934, Irère Curie and Frédéric Joliot demonstrated that radioactive nuclei can be made in the laboratory. All chemical elements may be rendered radioactive by adding or by subtracting (except for hydrogen and helium) neutrons from the nucleus of the stable ones. Studies of the radioactive decays of new isotopes far from the stable ones in nature continue as a major frontier in nuclear research. The availability of this wide variety of radioactive isotopes has stimulated their use in many different fields, including chemistry, biology, medicine, industry, artifact dating, agriculture, and space exploration. *See* ALPHA PARTICLES; BETA PARTICLES; GAMMA RAYS; ISOTOPE; RADIOACTIVITY AND RADIATION APPLICATIONS.

A particular radioactive transition may be delayed by less than a microsecond or by more than a billion years, but the existence of a measurable delay or lifetime distinguishes a radioactive nuclear transition from a so-called prompt nuclear transition, such as is involved in the emission of most gamma rays. The delay is expressed quantitatively by the radioactive decay constant, or by the mean life, or by the half-period for each type of radioactive atom.

The most commonly found types of radioactivity are alpha, beta negatron, beta positron, electron capture, and isomeric transition (**Table 1**). Each is characterized by the particular type of nuclear radiation which is emitted by the transforming parent nucleus. In addition, there are several other decay modes that are observed more rarely in specific regions of the periodic table (Table 1). Several of these rarer processes are in fact two-step processes (**Figs. 1** and **2**). In addition, there are several other processes predicted theoretically that remain to be verified.

In alpha radioactivity (Table 1) the parent nucleus spontaneously emits an alpha particle. Since the

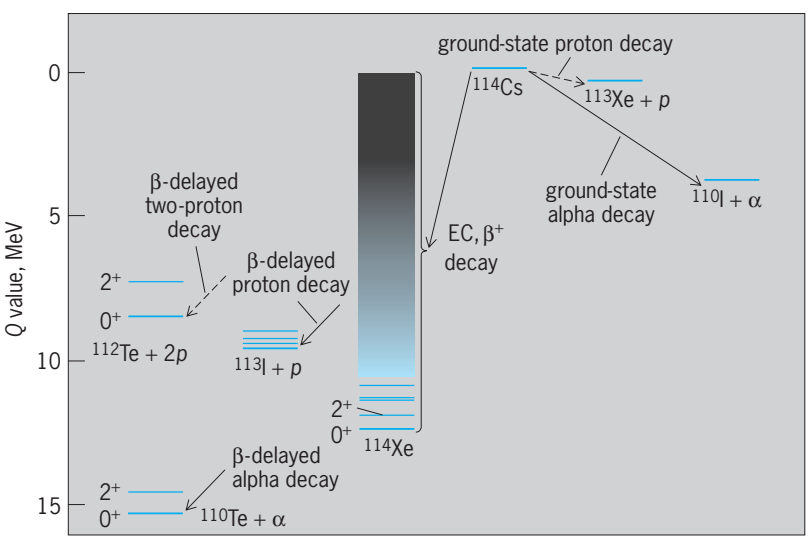

Fig. 1. Decay modes of ^{114}Cs based on Q values from the droplet-model formula for nuclear masses. (*After J. H. Hamilton et al., eds., Future Directions in Studies of Nuclei Far from Stability, 1980*)

TABLE 1. Types of radioactivity

Type	Symbol	Particles emitted	Change in atomic number, ΔZ	Change in atomic mass number, ΔA	Example
Alpha	α	Helium nucleus	−2	−4	${}^{226}_{86}\mathrm{Ra} \rightarrow {}^{222}_{84}\mathrm{Rn} + \alpha$
Beta negatron	β^-	Negative electron and antineutrino[a]	+1	0	${}^{24}_{11}\mathrm{Na} \rightarrow {}^{24}_{12}\mathrm{Mg} + e^- + \overline{\nu}$
Beta positron	β^+	Positive electron and neutrino[a]	−1	0	${}^{22}_{11}\mathrm{Na} \rightarrow {}^{22}_{10}\mathrm{Ne} + e^+ + \nu$
Electron capture	EC	Neutrino[a]	−1	0	${}^{7}_{4}\mathrm{Be} + e^- \rightarrow {}^{7}_{3}\mathrm{Li} + \nu$
Isomeric transition[b]	IT	Gamma rays or conversion electrons or both (and positive-negative electron pair)[c]	0	0	${}^{137m}_{56}\mathrm{Ba} \rightarrow {}^{137}_{56}\mathrm{Ba} + \gamma$ or conversion electrons
Proton	p	Proton	−1	−1	${}^{151}_{71}\mathrm{La} \rightarrow {}^{150}_{70}\mathrm{Yb} + p$
Spontaneous fission (hot)	SF	Two intermediate-mass nuclei and 1–10 neutrons	Various	Various	${}^{238}_{92}\mathrm{U} \rightarrow {}^{133}_{50}\mathrm{Sn} + {}^{103}_{42}\mathrm{Mo} + 2n$
Spontaneous fission (cold)	SF	Two intermediate-mass nuclei (zero neutrons)	Various	Various	${}^{252}_{98}\mathrm{Cf} \rightarrow {}^{106}_{42}\mathrm{Mo} + {}^{146}_{56}\mathrm{Ba}$
Ternary spontaneous fission (hot)	TSF	Two intermediate-mass nuclei, a light particle (${}^{2}\mathrm{H}$, α, up to ${}^{12}\mathrm{C}$), and neutrons	Various	Various	${}^{252}_{98}\mathrm{Cf} \rightarrow {}^{100}_{40}\mathrm{Zr} + {}^{146}_{56}\mathrm{Ba} + \alpha + 2n$
Ternary spontaneous fission (cold)	TSF	Two intermediate-mass nuclei and a light particle	Various	Various	${}^{252}_{98}\mathrm{Cf} \rightarrow {}^{96}_{38}\mathrm{Sr} + {}^{146}_{56}\mathrm{Ba} + {}^{10}_{4}\mathrm{Be}$
Isomeric spontaneous fission	ISF	Heavy fragments and neutrons	Various	Various	${}^{244f}_{95}\mathrm{Am} \rightarrow {}^{134}_{53}\mathrm{I} + {}^{107}_{42}\mathrm{Mo} + 3n$
Beta-delayed spontaneous fission	(EC + β^+)SF	Positive electron, neutrino, heavy fragments, and neutrons	Various	Various	${}^{246}_{99}\mathrm{Es} \rightarrow \beta^+ + \nu + {}^{246f}_{98}\mathrm{Cf} \rightarrow {}^{138}_{54}\mathrm{Xe} + {}^{107}_{44}\mathrm{Ru} + n$
	β^-SF	Negative electron, antineutrino, heavy fragments, and neutrons	Various	Various	${}^{236}_{91}\mathrm{Pa} \rightarrow \beta^- + \nu + {}^{236f}_{92}\mathrm{U} \rightarrow {}^{139}_{53}\mathrm{I} + {}^{94}_{39}\mathrm{Y} + 3n$
Beta-delayed neutron	$\beta^- n$	Negative electron, and antineutrino, neutron	+1	−1	${}^{11}_{3}\mathrm{Li} \rightarrow \beta^- + \nu + {}^{11}_{4}\mathrm{Be}^* \rightarrow {}^{10}_{4}\mathrm{Be} + n$
Beta-delayed two-neutron (three-, four-neutron)	$\beta^- 2n(3n, 4n)$	Negative electron, antineutrino, and two (three, four) neutrons	+1	−2 (−3, −4)	${}^{11}_{3}\mathrm{Li} \rightarrow \beta^- + \nu + {}^{11}_{4}\mathrm{Be}^* \rightarrow {}^{9(8)}_{4}\mathrm{Be} + 2n(3n)$
Beta-delayed proton	$\beta^+ p$ or (β^+ + EC)p	Positive electron, neutrino, and proton	−2	−1	${}^{114}_{55}\mathrm{Cs} \rightarrow \beta^+ + \nu + {}^{114}_{54}\mathrm{Xe}^* \rightarrow {}^{113}_{53}\mathrm{I} + p$
Beta-delayed two-proton	$\beta^+ 2p$	Positive electron, neutrino, and two protons	−3	−2	${}^{22}_{13}\mathrm{Al} \rightarrow \beta^+ + \nu + {}^{22}_{12}\mathrm{Mg}^* \rightarrow {}^{20}_{10}\mathrm{Ne} + 2p$
Beta-delayed deuteron	$\beta^- {}^{2}_{1}\mathrm{H}$	Negative electron, antineutrino, and deuteron	0	−2	${}^{6}_{2}\mathrm{He} \rightarrow \beta^- + {}^{6}_{3}\mathrm{Li} \rightarrow {}^{2}_{1}\mathrm{H} + {}^{4}_{2}\mathrm{He}$
Beta-delayed triton	$\beta^- {}^{3}_{1}\mathrm{H}$	Negative electron, antineutrino, and triton	0	−3	${}^{11}_{3}\mathrm{Li} \rightarrow \beta^- + \nu + {}^{11}_{4}\mathrm{B}^* \rightarrow {}^{8}_{3}\mathrm{Li} + {}^{3}_{1}\mathrm{H}$
Beta-delayed alpha	$\beta^+ \alpha$	Positive electron, neutrino, and alpha	−3	−4	${}^{114}_{55}\mathrm{Cs} \rightarrow \beta^+ + \nu + {}^{114}_{54}\mathrm{Xe}^* \rightarrow {}^{110}_{52}\mathrm{Te} + \alpha$
	$\beta^- \alpha$	Negative electron, antineutrino, and alpha	−1	−4	${}^{214}_{83}\mathrm{Bi} \rightarrow \beta^- + \nu + {}^{114}_{84}\mathrm{Po}^* \rightarrow {}^{210}_{82}\mathrm{Pb} + \alpha$
Beta-delayed alpha-neutron	$\beta^- \alpha, n$	Negative electron, antineutrino, alpha, and neutron	−1	−5	${}^{11}_{3}\mathrm{Li} \rightarrow \beta^- + \nu + {}^{11}_{4}\mathrm{B}^* \rightarrow {}^{6}_{2}\mathrm{He} + \alpha + n$
Double beta decay	$\beta^- \beta^-$	Two negative electrons and two antineutrinos	+2	0	${}^{82}_{34}\mathrm{Se} \rightarrow {}^{82}_{36}\mathrm{Kr} + 2\beta^- + 2\nu$
	$\beta^+ \beta^+$	Two positive electrons and two neutrinos	−2	0	${}^{d130}_{56}\mathrm{Ba} \rightarrow {}^{130}_{54}\mathrm{Xe} + 2\beta^+ + 2\nu$
Double electron capture[d]	EC EC	Two neutrinos	−2	0	${}^{d130}_{56}\mathrm{Ba} + 2e^- \rightarrow {}^{130}_{54}\mathrm{Xe} + 2\nu$
Neutrinoless double beta decay[e]	$\beta^- \beta^-$	Two negative electrons	+2	0	${}^{e82}_{34}\mathrm{Se} \rightarrow {}^{82}_{36}\mathrm{Kr} + 2\beta^-$
Two-proton	$2p$	Two protons	−2	−2	${}^{45}_{26}\mathrm{Fe} \rightarrow {}^{43}_{24}\mathrm{Cr} + 2p$
Neutron[d]	n	Neutron	0	−1	

(*cont.*)

TABLE 1. Types of radioactivity (*cont.*)

Type	Symbol	Particles emitted	Change in atomic number, ΔZ	Change in atomic mass number, ΔA	Example
Two-neutron[d]	$2n$	Two neutrons	0	−2	
Heavy clusters[f]	${}^{14}_{6}C$	${}^{14}_{6}C$ nucleus	−6	−14	${}^{223}_{88}Ra \rightarrow {}^{209}_{82}Pb + {}^{14}_{6}C$
	${}^{20}_{8}O$	${}^{20}_{8}O$ nucleus	−8	−20	${}^{227}_{89}Ac \rightarrow {}^{207}_{81}Tl + {}^{20}_{8}O$
	${}^{24}_{10}Ne$	${}^{24}_{10}Ne$ nucleus	−10	−24	${}^{232}_{92}U \rightarrow {}^{208}_{82}Pb + {}^{24}_{10}Ne$

[a] The neutrinos and antineutrinos emitted in beta decay are electron neutrinos and antineutrinos.
[b] Excited states with relatively long measured half-lives are called isomeric and are identified by placing the symbol *m* for metastable after the mass number, as in ${}^{137m}Ba$. Excited states with essentially prompt decay are identified by asterisks, as in ${}^{11}Be^{*}$.
[c] Pair emission occurs as an additional competing decay mode when the decay energy exceeds 1.022 MeV.
[d] Theoretically predicted but not established experimentally.
[e] Theoretically possible and predicted in some grand unified theories.
[f] There are other possible clusters in addition to those shown. Cold SF is also a type of cluster radioactivity.

alpha particle is the nucleus of the helium-4 atom, it contains two protons and two neutrons. Thus, the atomic number, or nuclear charge Z, of the decay product is 2 units less than that of the parent, and the nuclear mass A of the product is 4 atomic mass units less than that of the parent, because the emitted alpha particle carries away this amount of nuclear charge and mass. This decrease of 2 units of atomic number or nuclear charge between parent and product means that the decay product will be a different chemical element, displaced by 2 units to the left in a periodic table of the elements. For example, radium has atomic number 88 and is found in column 2 of the periodic table. Its decay product after the emission of an alpha particle is a different chemical element, radon, whose atomic number is 86 and whose position is in column 0 of the periodic table.

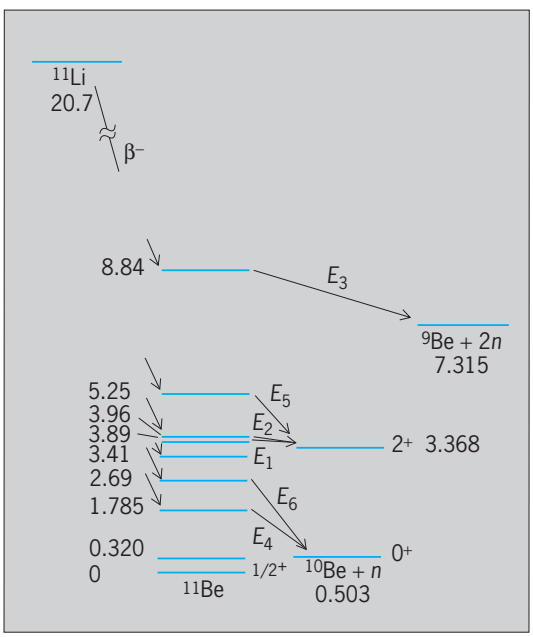

Fig. 2. Observed decay modes of ${}^{11}Li$. Energies in megaelectronvolts. (*After J. H. Hamilton et al., eds., Future Directions in Studies of Nuclei Far from Stability, 1980*)

Transition Rates and Decay Laws

This section covers radioactive decay constant, dual decay, exponential decay law, mean life, and half-period.

Radioactive decay constant. The rate of radioactive transformation, or the activity, of a source equals the number A of identical radioactive atoms present in the source, multiplied by their characteristic radioactive decay constant λ. Thus Eq. (1) holds, where the

$$\text{Activity} = A\lambda \text{ disintegrations per second} \qquad (1)$$

decay constant λ has dimensions of s^{-1}. The numerical value of λ expresses the statistical probability of decay of each radioactive atom in a group of identical atoms, per time. For example, if $\lambda = 0.01\ s^{-1}$ for a particular radioactive species, then each atom has a chance of 0.01 (1%) of decaying in 1 s, and a chance of 0.99 (99%) of not decaying in any given 1-s interval. The constant λ is one of the most important characteristics of each radioactive nuclide: λ is essentially independent of all physical and chemical conditions such as temperature, pressure, concentration, chemical combination, or age of the radioactive atoms. There are a few cases where measurable effects are observed for different chemical combinations. One of the largest observed is a 3.2% change in λ for the 24-s isomer in ${}^{90}Nb$. The half-period is inversely proportional to λ.

The identification of some radioactive samples can be made simply by measuring λ, which then serves as an equivalent of qualitative chemical analysis. For the most common radioactive nuclides, the range of λ extends from $3 \times 10^{6}\ s^{-1}$ (for thorium C′) to $1.6 \times 10^{-18}\ s^{-1}$ (for thorium).

Dual decay. Many radioactive nuclides have two or more independent and alternative modes of decay. For example, ${}^{238}U$ can decay either by alpha-particle emission or by spontaneous fission. A single atom of ${}^{64}Cu$ can decay in any of three competing independent ways: negatron beta-particle emission, positron beta-particle emission, or electron capture. When two or more independent modes of decay are possible, the nuclide is said to exhibit dual decay.

The competing modes of decay of any nuclide have independent partial decay constants given by the probabilities $\lambda_1, \lambda_2, \lambda_3, \ldots$ per second, and the total probability of decay is represented by the total decay constant λ, defined by Eq. (2). If there are A

$$\lambda = \lambda_1 + \lambda_2 + \lambda_3 + \cdots \tag{2}$$

identical atoms present, the partial activities, as measured by the different modes of decay, are $A\lambda_1$, $A\lambda_2$, $A\lambda_3, \ldots$, and the total activity $A\lambda$ is given by Eq. (3).

$$A\lambda = A\lambda_1 + A\lambda_2 + A\lambda_3 + \cdots \tag{3}$$

The partial activities, $A\lambda_1, \ldots$, such as positron beta particles from ^{64}Cu, are proportional to the total activity, $A\lambda$, at all times.

The branching ratio is the fraction of the decaying atoms which follow a particular mode of decay, and equals $A\lambda_1/A\lambda$ or λ_1/λ. For example, in the case of ^{64}Cu the measured branching ratios are $\lambda_1/\lambda = 0.40$ for negatron beta decay, $\lambda_2/\lambda = 0.20$ for positron beta decay, and $\lambda_3/\lambda = 0.40$ for electron capture. The sum of all the branching ratios for a particular nuclide is unity.

Exponential decay law. The total activity, $A\lambda$, equals the rate of decrease $-dA/dt$ in the number of radioactive atoms A present. Because λ is independent of the age t of an atom, integration of the differential equation of radioactive decay, $-dA/dt = A\lambda$, gives Eq. (4), where ln represents the natural log-

$$\ln \frac{A}{A_0} = -\lambda(t - t_0) \tag{4}$$

arithm to the base e, and A atoms remain at time t if there were A_0 atoms initially present at time t_0. If $t_0 = 0$, then Eq. (4) can be rewritten as the exponential law of radioactive decay in its most common form, Eq. (5). The initial activity at $t = 0$ was $A_0\lambda$, and

$$A = A_0 e^{-\lambda t} \tag{5}$$

the activity at t, when only A atoms remain untransformed, is $A\lambda$. Because λ is a constant, the fractional activity $A\lambda/A_0\lambda$ at time t and the fractional amount of radioactive atoms A/A_0 are given by Eq. (6). In cases

$$\frac{A\lambda}{A_0\lambda} = \frac{A}{A_0} = e^{-\lambda t} \tag{6}$$

of dual decay, the partial activities $A\lambda_1, A\lambda_2, \ldots$ also decrease with time as $e^{-\lambda t}$, not as $e^{-\lambda_1 t} \ldots$, because $A\lambda_1/A_0\lambda_1 = A/A_0 = e^{-\lambda t}$ where λ is the total decay constant. This is because the decrease of each partial activity with time is due to the depletion of the total stock of atoms A, and this depletion is accomplished by the combined action of all the competing modes of decay.

Mean life. The actual life of any particular atom can have any value between zero and infinity. The average or mean life of a large number of identical radioactive atoms is, however, a definite and important quantity.

If there are A_0 atoms present initially at $t = 0$, then the number remaining undecayed at a later time t is $A = A_0 e^{-\lambda t}$, by Eq. (5). Each of these A atoms has a life longer than t. In an additional infinitesimally short time interval dt, between time t and $t + dt$, the absolute number of atoms which will decay on the average is $A\lambda dt$, and these atoms had a life-span t. The total L of the life-spans of all the A_0 atoms is the sum or integral of $tA\lambda\, dt$ from $t = 0$ to $t = \infty$, which is given by Eq. (7). Then the average lifetime L/A_0, which is called the mean life τ, is given by Eq. (8),

$$L = \int_0^\infty tA\lambda\, dt = \int_0^\infty tA_0\lambda e^{-\lambda t} dt = \frac{A_0}{\lambda} \tag{7}$$

$$\tau = \frac{1}{\lambda} \tag{8}$$

where λ is the total radioactive decay constant of Eq. (2). Substitution of $t = \tau = 1/\lambda$ into Eq. (6) shows that the mean life is the time required for the number of atoms, or their activity, to fall to $e^{-1} = 0.368$ of any initial value.

Half-period (half-life). The time interval over which the chance of survival of a particular radioactive atom is exactly one-half is called the half-period T (also called half-life, written $T_{1/2}$). From Eq. (4), Eq. (9) is

$$-\ln \frac{A}{A_0} = \ln \frac{A_0}{A} = \ln 2 = 0.693 = \lambda T \tag{9}$$

obtained. Then the half-period T is related to the total radioactive decay constant λ, and to the mean life τ, by Eq. (10). For mnemonic reasons, the half-period

$$T = \frac{0.693}{\lambda} = 0.693\tau \tag{10}$$

T is much more frequently employed than the total decay constant λ or the mean life τ. For example, it is more common to speak of ^{232}Th as having a half-period of 1.4×10^{10} years than to speak of its mean life of 2.0×10^{10} years or its total decay constant of 1.6×10^{-18} s^{-1}, although all three are equivalent statements of the average longevity of ^{232}Th atoms.

Any initial activity $A_0\lambda$ is reduced to $^1/_2$ in 1 half-period T, to $1/e$ in 1 mean life τ, to $^1/_4$ in 2 half-periods $2T$, and so on (**Fig. 3**). The slope of the

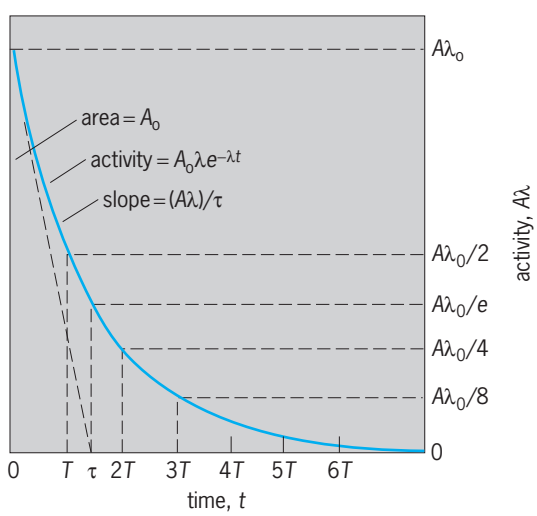

Fig. 3. Graphical representation of relationships in decay of a single radioactive nuclide.

activity curve, or rate of decrease of activity, is $d(A\lambda)/dt = -\lambda dA/dt = -\lambda(A\lambda)$. Thus the initial slope is $-\lambda(A_0\lambda) = -(A_0\lambda)/\tau$. The area under the activity curve, if integrated to $t = \infty$, is simply A_0, the total initial number of radioactive atoms. Also, the initial activity $A_0\lambda$, if it could continue at a constant value for one mean life τ, would exactly destroy all the atoms because $(A_0\lambda)\tau = A_0$.

Radioactive Series Decay

In a number of cases a radioactive nuclide A decays into a nuclide B which is also radioactive; the nuclide B decays into C which is also radioactive, and so on. For example, ${}^{232}_{90}\mathrm{Th}$ decays into a series of 10 successive radioactive nuclides. Substantially all the primary products of nuclear fission are negatron beta-particle emitters which decay through a chain or series of two to six successive beta-particle emitters before a stable nuclide is reached as an end product. *See* NUCLEAR FISSION.

Let the initial part of such a series be represented by reaction (11), where radioactive atoms of types A,

$$A \xrightarrow{\lambda_A} B \xrightarrow{\lambda_B} C \xrightarrow{\lambda_C} D \xrightarrow{\lambda_D} \cdots \tag{11}$$

$B, C, D, \ldots$ have radioactive decay constants given by $\lambda_A, \lambda_B, \lambda_C, \lambda_D, \ldots$. Then if there are initially present, at time $t = 0$, A_0 atoms of type A, the numbers A, B, C, ... of atoms of types A, B, C, ..., which will be present at a later time t, are given by Eqs. (12)–(14), and the activities of A, B, C, ... are $A\lambda_A$, $B\lambda_B$,

$$A = A_0 e^{-\lambda_A t} \tag{12}$$

$$B = A_0 \frac{\lambda_A}{\lambda_B - \lambda_A}(e^{-\lambda_A t} - e^{-\lambda_B t}) \tag{13}$$

$$C = A_0 \left(\frac{\lambda_A}{\lambda_C - \lambda_A} \frac{\lambda_B}{\lambda_B - \lambda_A} e^{-\lambda_A t} + \frac{\lambda_A}{\lambda_A - \lambda_B} \frac{\lambda_B}{\lambda_C - \lambda_B} e^{-\lambda_B t} + \frac{\lambda_A}{\lambda_A - \lambda_C} \frac{\lambda_B}{\lambda_B - \lambda_C} e^{-\lambda_C t} \right) \tag{14}$$

$C\lambda_C$, ... (**Fig. 4**). General equations describing the amounts and activities of any number of radioactive decay products are more complicated and are given in standard texts.

Radioactive equilibrium. In Fig. 4 the ratio $B\lambda_B/A\lambda_A$ of the activities of the parent A and the daughter product B change with time. The activity $B\lambda_B$ is zero initially and also after a very long time, when all the atoms have decayed. Thus $B\lambda_B$ passes through a maximum value, and it can be shown that this occurs at a time t_m given by Eq. (15). The situation in which

$$t_m = \frac{\ln(\lambda_B/\lambda_A)}{\lambda_B - \lambda_A} \tag{15}$$

the activities $A\lambda_A$ and $B\lambda_B$ are exactly equal to each other is called ideal equilibrium, and exists only at the moment t_m.

If the parent A is longer-lived than the daughter B, as occurs in many cases, then at a time which is long compared with the mean life τ_B of B, the activity ratio approaches a constant value given by Eq. (16), where T_A and T_B are the half-periods of A

$$\frac{B\lambda_B}{A\lambda_A} = \frac{\lambda_B}{\lambda_B - \lambda_A} = \frac{T_A}{T_A - T_B} \tag{16}$$

and of B. When the activity ratio $B\lambda_B/A\lambda_A$ is constant, a particular type of radioactive equilibrium exists. This is spoken of as secular equilibrium if the activity ratio is experimentally indistinguishable from unity, as occurs when T_A is very much greater than T_B.

Equilibrium concepts are applied also between a long-lived parent and any of its decay products in a long series. For example, in a sufficiently old uranium ore, radium ($T = 1620$ years) is in secular equilibrium with its ultimate parent uranium ($T = 4.5 \times 10^9$ years) although there are four intermediate radioactive substances intervening in the series between uranium and radium. Here, secular equilibrium shows that activities of radium and uranium continue to be equal to each other even though the activity of the parent uranium is decreasing with time.

When T_B is comparable with T_A, Eq. (16) shows that the equilibrium ratio will clearly exceed unity; this situation is spoken of as transient equilibrium. For example, in fission-product decay series (17) the

$${}^{140}\mathrm{Ba} \rightarrow {}^{140}\mathrm{La} \rightarrow {}^{140}\mathrm{Ce} \tag{17}$$

half-period of ^{140}Ba is 307 h and that of ^{140}La is 40 h. In an initially pure source of ^{140}Ba the activity of ^{140}La starts at zero, rises to a maximum at $t_m = 135$ h [Eq. (15)], then decreases, and after a few hundred hours is in transient equilibrium with its parent, when the ^{140}La activity [by Eq. (16)] is 307/(307–40) = 1.15 times the activity of its parent ^{140}Ba.

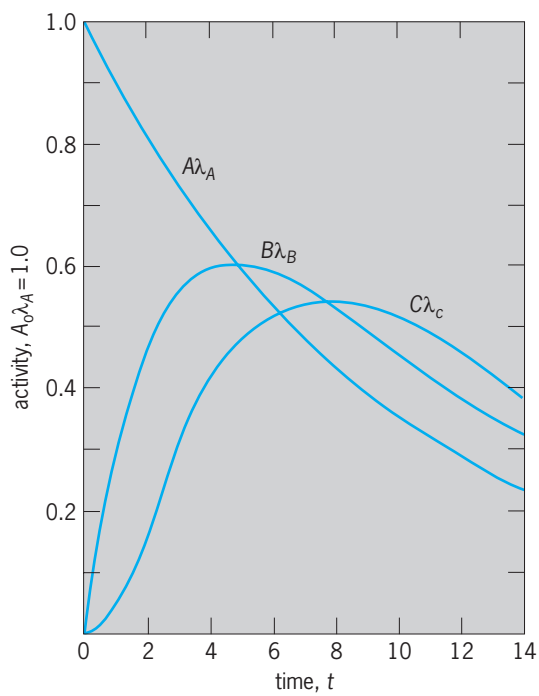

Fig. 4. Growth and decay of the activity $B\lambda_B$ of the daughter product, and $C\lambda_C$ of the granddaughter product, in an initially pure source of a radioactive parent whose activity at $t = 0$ is $A_0\lambda_A$.

TABLE 2. Parent radioactive nuclides found in nature

Nuclide					
Atomic number, Z	Mass number, A	Percent abundance in nature	Half-period, years	Radioactive transitions observed*	Disintegration energy, MeV*
19 K	40	0.0117	1.3×10^9	β^-, β^+, EC	β^- 1.3 – EC 1.5
34 Se	82	9.19	1.1×10^{20}	$\beta^-\beta^-$	3.0
37 Rb	87	27.83	4.8×10^{10}	β^-	0.3
48 Cd	113	12.2	9×10^{15}	β^-	0.3
49 In	115	95.77	5.1×10^{14}	β^-	0.5
52 Te	130	34.49	2×10^{21}	Growth of $^{130}_{54}Xe$†	1.6
57 La	138	0.089	1.1×10^{11}	β^-, EC	β^- 1.0 – EC 1.75
60 Nd	144	23.8	2.1×10^{15}	α	1.9
62 Sm	147	15.07	1.1×10^{11}	α	2.3
62 Sm	148	11.3	8×10^{15}	α	1.99
64 Gd	152	0.20	1.1×10^{14}	α	2.2
71 Lu	176	2.6	3.6×10^{10}	β^-, γ	0.6
72 Hf	174	0.16	2×10^{15}	α	2.5
75 Re	187	62.6	4×10^{10}	β^-	0.003
78 Pt	190	0.013	6×10^{11}	α	3.24
90 Th	232	100	1.4×10^{10}	α	4.08
92 U	235	0.715	7.0×10^8	α, SF	α 4.68
92 U	238	99.28	4.5×10^9	α, SF	α 4.27

*EC = electron capture. The EC energy is between ground states, but, in ^{138}La decay, 1.44 MeV of the EC energy goes to a gamma ray that feeds the ground state.
†Indirect evidence for $\beta^-\beta^-$ decay.

Radioactivity in the Earth. A number of isotopes of elements found in the Earth are radioactive (**Table 2**). All known or theoretically predicted isotopes of elements above bismuth are radioactive. Because the Earth is composed of atoms which were believed to have been created more than 3×10^9 years ago, the naturally occurring parent radioactive isotopes are those which have such long half-periods that detectable residual activity is still observable today. As a general rule, one can detect the presence of a radioactive substance for about 10 half-lives. Therefore activities with $T \lesssim 0.3 \times 10^9$ years should not be found in the Earth. For example, present-day uranium is an isotopic mixture containing 99.3% ^{238}U, whose half-period is 4.5×10^9 years, and only 0.7% of the shorter-lived uranium isotope ^{235}U, whose half-period is 0.7×10^9 years, whereas these isotopes presumably were produced in roughly equal amounts in the Earth a few billion years ago. Geophysical evidence indicates that originally some ^{236}U was present also, but none is found in nature now as expected with its half-period of 0.02×10^9 years. The elements technetium ($Z = 43$) and promethium ($Z = 61$) are not found in the Earth's crust because all their isotopes are radioactive with much shorter half-periods (their longest-lived are $T = 2.6 \times 10^6$ years for ^{97}Tc and $T = 17.7$ years for ^{145}Pm).

Uranium-238 decays through a long series of 14 radioactive decay products before ending as a stable isotope of lead, ^{206}Pb. Some of these members of the ^{238}U decay chain have very short half-periods, so their existence in nature is entirely dependent on the presence of their long-lived parent, and thus is a genealogical accident. For example, radium occurs in nature only in the minerals of its parent, uranium. The decay series of ^{235}U supports 14, and the decay series of ^{232}Th supports 10, short-lived radioactive substances found in nature.

A few of the common elements contain long-lived, naturally radioactive isotopes. For example, all terrestrial potassium contains 0.012% of the radioactive isotope ^{40}K, which has a half-period of 1.3×10^9 years, and emits negatron or positron beta particles (plus decay via electron capture) and gamma rays in a dual decay to stable ^{40}Ca and ^{40}Ar. This isotope is the principal source of radioactivity in a normal human being; each human contains about 0.1 microcurie (3.7×10^3 becquerels) of the radioactive potassium isotope ^{40}K.

Geological age measurements are based on the accumulation of decay products of long-lived isotopes, especially in the cases of ^{40}K, ^{87}Rb, ^{232}Th, ^{235}U, and ^{238}U. *See* GEOCHRONOMETRY.

Laboratory-produced radioactive nuclei. With particle accelerators and nuclear reactors, the order of 2000 radioactive isotopes not found in detectable quantities in the Earth's crust have been produced in the laboratory since 1935, including those of 26 new chemical elements up to element 118 (as of 2004), with element 117 not known. Earlier titles of induced or artificial radioactivities for these isotopes are misnomers. Many of these now have been identified in meteorites and in stars, and others are produced in the atmosphere by cosmic rays. There are over 5000 isotopes theoretically predicted to exist. As one approaches the place where a proton or neutron is no longer bound in a nucleus of an element (the limits of the existence of that element), the half-periods become extremely short. *See* TRANSURANIUM ELEMENTS.

For example, carbon-14 is a negatron beta-particle emitter, with a half-period of about 5600 years, which can be produced in the laboratory as the

product of a variety of different nuclear transmutation experiments. Nuclear bombardment of ^{11}B nuclei by alpha particles (helium nuclei) can produce excited compound nuclei of ^{15}N which promptly emit a proton (hydrogen nucleus), leaving ^{14}C as the end product of the transmutation. The same end-product ^{14}C can be produced by bombarding ^{14}N with neutrons, resulting in nuclear reaction (18).

$$^{14}N + \text{neutron} \rightarrow {}^{15}N^{*} \rightarrow {}^{14}C + \text{proton} \qquad (18)$$

This reaction is easily carried out by using neutrons from nuclear accelerators or a nuclear reactor. This particular transmutation reaction is one which occurs in nature also, because the nitrogen in the Earth's atmosphere is continually bombarded by neutrons which are produced by cosmic rays, thus producing radioactive ^{14}C. Mixing of ^{14}C with stable carbon provides the basis for radiocarbon dating of systems that absorb carbon for times up to about 50,000 years ago (10 half-lives). *See* COSMIC RAYS; NUCLEAR REACTION; NUCLEAR REACTOR; PARTICLE ACCELERATOR; RADIOCARBON DATING.

Radioactive hydrogen, ^{3}H, is also formed in the atmosphere from the ^{14}N + neutron → ^{12}C + ^{3}H reaction. Also, ^{3}H is produced in the Sun, and the Earth's water as well as satelleites show an additional concentration of ^{3}H from the Sun. Over two dozen radioactive products, ranging in half-life from a few days to millions of years, have been identified in meteorites that have fallen to Earth. The carbon and hydrogen burning cycles that produce energy for stars produce radioactive ^{13}N, ^{15}O, ^{3}H. At higher temperatures the radioactivities ^{7}Be and even ^{8}Be ($T \approx 10^{-16}$ s) help burn hydrogen and helium. In addition to the production of radioactive as well as stable isotopes prior to the formation of the solar system, nucleosynthesis continues to go on in stars with the production of many short-lived radioactive atoms by different processes. *See* CARBON-NITROGEN-OXYGEN CYCLES; NUCLEOSYNTHESIS; PROTON-PROTON CHAIN.

The yield of any radioactivity produced in the laboratory is the initial rate of the activity under the particular conditions of nuclear bombardment. When a target material A is bombarded to produce a radioactive product B whose radioactive decay constant is λ_B, the number of atoms B which are present after a bombardment of duration t, and their activity $B\lambda_B$, are given by Eq. (19), where the yield Y has dimen-

$$B\lambda_B = \frac{Y}{\lambda_B}(1 - e^{-\lambda_B t}) \qquad (19)$$

sions equivalent to curies of activity produced per second of bombardment. The yield Y depends on the number of atoms A present in the target, the intensity of the beam of bombarding particles, and the cross section, or probability of the reaction per bombarding particle under the conditions of bombardment.

Radioactive transformation series. As noted in Eqs. (12)–(14), many radioactive substances have decay products which are also radioactive. Thus many long chains or series of radioactive transformations are known. The three naturally occurring transformation series are headed by ^{232}Th, ^{235}U, and ^{238}U (**Fig. 5** and **Table 3**).

Each of the naturally occurring radioactive isotopes in these transformation series has two synonymous names. For example, the commercially important radioisotope whose classical name is mesothorium-1 is now known to be an isotope of radium with mass number of 228 and is designated as radium-228 (^{228}Ra). Table 3 summarizes the names, symbols, and some radioactive properties of these three transformation series. However, these chains are not complete, and their uniqueness or importance as chains is an accident of the very long half-lives of ^{232}Th, ^{235}U, and ^{238}U. For example, element 105 of mass 260 has a succession of seven alpha decays and one electron capture and positron decay to ^{232}Th. The special importance of the chains in Table 3 is related to the fact that they were essentially the only early sources of radioactive materials, and they also play a role in nuclear power. *See* NUCLEAR POWER.

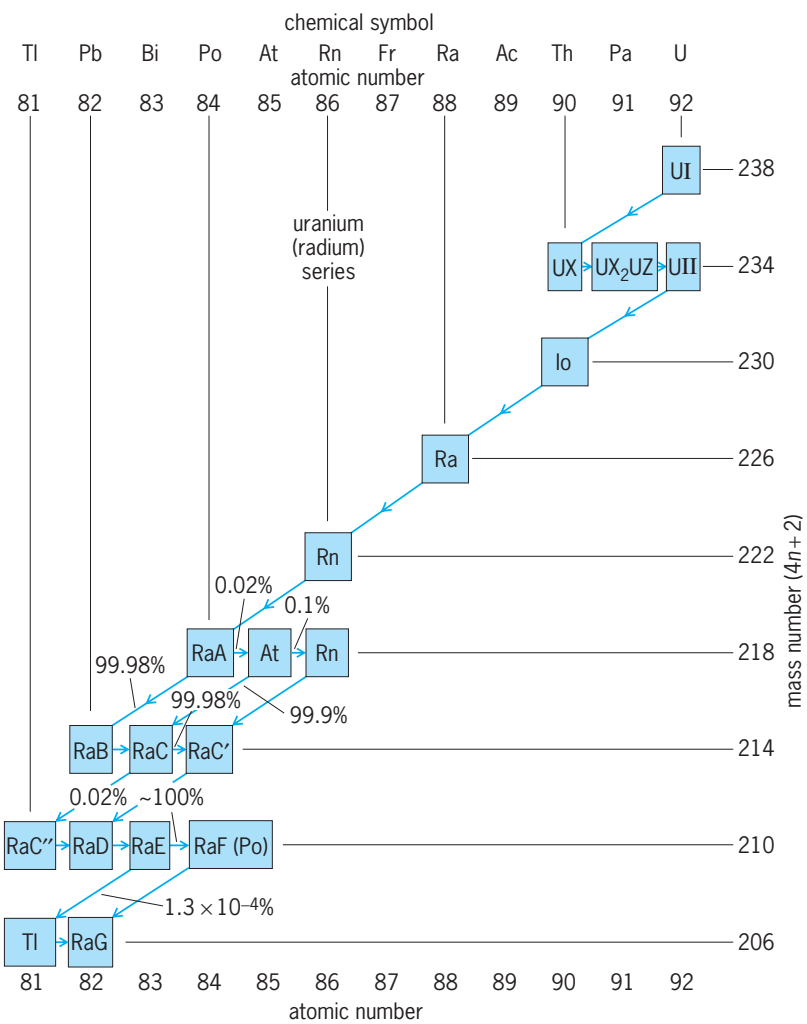

Fig. 5. Main line of decay of uranium series, or $4n + 2$ series, of heavy radioactive nuclides, headed in nature by uranium-238. Each member has a mass number given by $4n + 2$, where n is an integer.

TABLE 3. Names, symbols, and radioactive properties of members of the three naturally occurring radioactive transformation series*

Early name†	Early symbol†	Atomic number	Mass number	Isotopic symbol	Half-period	Type of decay
			Uranium (4n + 2) series			
Uranium I	UI	92	238	^{238}U	4.5×10^9 y	α, SF
Uranium X_1	UX_1	90	234	^{234}Th	24 d	β^-
Uranium X_2	UX_2	91	234	^{234m}Pa	1.2 m	IT, β^-
Uranium Z	UZ	91	234	^{234}Pa	6.7 h	β^-
Uranium II	UII	92	234	^{234}U	2.5×10^5 y	α
Ionium	Io	90	230	^{230}Th	8×10^4 y	α
Radium	Ra	88	226	^{226}Ra	1600 y	α
Radon	Rn	86	222	^{222}Rn	3.8 d	α
Radium A	RaA	84	218	^{218}Po	3.0 m	α, β^-
Astatine	At	85	218	^{218}At	1.5 s	β^-, α
Radon	Rn	86	218	^{218}Rn	3.5 ms	α
Radium B	RaB	82	214	^{214}Pb	27 m	β^-
Radium C	RaC	83	214	^{214}Bi	20 m	β^-, α
Radium C′	RaC′	84	214	^{214}Po	1.6×10^{-4} s	α
Radium C″	RaC″	81	210	^{210}Tl	1.3 m	β^-
Radium D	RaD	82	210	^{210}Pb	22 y	β^-
Radium E	RaE	83	210	^{210}Bi	5.0 d	β^-, α
Radium F	RaF	84	210	^{210}Po	138 d	α
Thallium	Tl	81	206	^{206}Tl	4.2 m	β^-
Radium G	RaG	82	206	^{206}Pb	Stable	Stable
			Thorium (4n) series			
Thorium	Th	90	232	^{232}Th	1.4×10^{10} y	α
$Mesothorium_1$	$MsTh_1$	88	228	^{228}Ra	5.8 y	β^-
$Mesothorium_2$	$MsTh_2$	89	228	^{228}Ac	6.1 h	β^-
Radiothorium	RdTh	90	228	^{228}Th	1.9 y	α
Thorium X	ThX	88	224	^{224}Ra	3.7 d	α
Thoron	Tn	86	220	^{220}Rn	56 s	α
Thorium A	ThA	84	216	^{216}Po	0.15 s	α
Thorium B	ThB	82	212	^{212}Pb	10.6 h	β^-
Thorium C	ThC	83	212	^{212}Bi	1.0 h	β^-, α
Thorium C′	ThC′	84	212	^{212}Po	3×10^{-7} s	α
Thorium C″	ThC″	81	208	^{208}Tl	3.1 m	β^-
Thorium D	ThD	82	208	^{208}Pb	Stable	Stable
			Actinium (4n + 3) series			
Actinouranium	AcU	92	235	^{235}U	7.0×10^8 y	α, SF
Uranium Y	UY	90	231	^{231}Th	26 h	β^-
Protactinium	Pa	91	231	^{231}Pa	3.3×10^4 y	α
Actinium	Ac	89	227	^{227}Ac	22 y	β^-, α
Radioactinium	RdAc	90	227	^{227}Th	19 d	α
Actinium K	AcK	87	223	^{223}Fr	22 m	β^-, α
Actinium X	AcX	88	223	^{223}Ra	11 d	α
Astatine	At	85	219	^{219}At	0.9 m	α, β^-
Actinon	An	86	219	^{219}Rn	4.0 s	α
Bismuth	Bi	83	215	^{215}Bi	7.6 m	β^-
Actinium A	AcA	84	215	^{215}Po	1.8×10^{-3} s	α
Actinium B	AcB	82	211	^{211}Pb	36 m	β^-
Actinium C	AcC	83	211	^{211}Bi	2.2 m	α, β^-
Actinium C′	AcC′	84	211	^{211}Po	0.5 s	α
Actinium C″	AcC″	81	207	^{207}Tl	4.8 m	β
Actinium D	AcD	82	207	^{207}Pb	Stable	Stable

*Radon-223 has been shown to have a very weak radioactive ^{14}C decay branch to ^{209}Pb. Several of the isotopes in these chains, such as $^{234,235,238}U$, ^{226}Ra, and others are predicted to have such very weak, heavy cluster decay branches.

†These were the names and symbols used before the different isotopes of these elements were known.

Transformation series are now known for every element in the periodic table except hydrogen. Chains of neutron-rich isotopes have been produced and studied among the products of nuclear fission. Heavy-ion-induced reactions and high-flux reactors have been used to extend knowledge of the elements beyond uranium. The elements from number 93 (neptunium) to 118 (as yet unnamed except for element 117), which have so far not been found on Earth, were made in the laboratory. Both proton- and heavy-ion-induced reactions have extended knowledge of chains and neutron-deficient isotopes of the stable elements.

Alpha-Particle Decay

Alpha-particle decay is that type of radioactivity in which the parent nucleus expels an alpha particle (a helium nucleus). The alpha particle is emitted with a speed of the order of 1 to 2×10^7 m/s (10^4 mi/s), that is, about 1/20 of the velocity of light.

In the simplest case of alpha decay, every alpha particle would be emitted with exactly the same

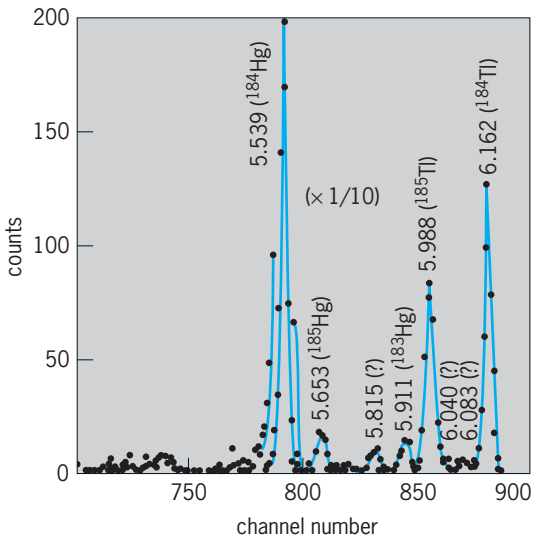

Fig. 6. Alpha groups in the decay of ^{184}Tl ($T = 11$ s) and ^{184}Hg and weak groups from ^{183}Hg and ^{185}Hg, very far off stability (17 neutrons less than the lightest stable thallium isotope). Energies in megaelectronvolts. (*After K. S. Toth et al., Observation of α-decay in thallium nuclei, including the new isotopes ^{184}Tl and ^{185}Tl, Phys. Lett., 63B:150–153, 1976*)

velocity and hence the same kinetic energy. However, in most cases there are two or more discrete energy groups called lines (**Fig. 6**). For example, in the alpha decay of a large group of ^{238}U atoms, 77% of the alpha decays will be by emission of alpha particles whose kinetic energy is 4.20 MeV, while 23% will be by emission of 4.15-MeV alpha particles. When the 4.20-MeV alpha particle is emitted, the decay product nucleus is formed in its ground (lowest energy) level. When a 4.15-MeV alpha particle is emitted, the decay product is produced in an excited level, 0.05 MeV above the ground level. This nucleus promptly transforms to its ground level by the emission of a 0.05-MeV gamma ray or alternatively by the emission of the same amount of energy in the form of a conversion electron and the associated spectrum of characteristic x-rays. Thus in all alpha-particle spectra, the alpha particles are emitted in one or more discrete and homogeneous energy groups, and alpha-particle spectra are accompanied by gamma-ray and conversion electron spectra whenever there are two or more alpha-particle groups in the spectrum.

Geiger-Nuttall rule. Among all the known alpha-particle emitters, most alpha-particle energy spectra lie in the domain of 4–6 MeV, although a few extend as low as 2 MeV ($^{147}_{62}$Sm) and as high as 10 MeV ($^{212}_{84}$Po or ThC′). There is a systematic relationship between the kinetic energy of the emitted alpha particles and the half-period of the alpha emitter. The highest-energy alpha particles are emitted by short-lived nuclides, and the lowest-energy alpha particles are emitted by the very long lived alpha-particle emitters. H. Geiger and J. M. Nuttall showed that there is a linear relationship between log λ and the energy of the alpha particle.

The Geiger-Nuttall rule is inexplicable by classical physics but emerges clearly from quantum, or wave, mechanics. In 1928 the hypothesis of transmission through nuclear potential barriers, as introduced by G. Gamow and independently by R. W. Gurney and E. U. Condon, was shown to give a satisfactory account of the alpha-decay data, and it has been altered subsequently only in details. The form of the barrier-penetration equations is such that correlation plots of log λ against $1/E$ give nearly straight lines.

Nuclear potential barrier. At distances r which are large compared with the nuclear radius, the potential energy of an alpha particle, whose charge is $2e$, in the field of a residual nucleus, whose charge is $(Z - 2)e$, is $2(Z - 2)e^2/r$. At very close distances this electrostatic repulsion is opposed and overcome by short-range, nuclear, attractive forces. The net potential energy U as a function of the separation r between the alpha particle and its residual nucleus is the nuclear potential barrier.

One of several operating definitions of the nuclear radius R is the distance $r = R$ at which the attractive nuclear forces just balance the repulsive electrostatic forces. At this distance, called the top of the nuclear barrier, the potential energy is about 25–30 MeV for typical cases of heavy, alpha-emitting nuclei (**Fig. 7**).

Inside the nucleus the alpha particle is represented as a de Broglie matter wave. According to wave mechanics, this wave has a very small but finite probability of being transmitted through the nuclear potential energy barrier and thus of emerging as an alpha particle emitted from the nucleus. The transmission of a particle through such an energy barrier is completely forbidden in classical electrodynamics but is possible according to wave mechanics. This transmission of a matter wave through an energy barrier is analogous to the familiar case of the transmission of ordinary visible light through an opaque metal such as gold: if the gold is thin enough, some light does get through, as in the case of the thin gold leaf which is sometimes used for lettering signs on store windows. *See* QUANTUM MECHANICS.

The wave-mechanical probability of the transmission of an alpha particle through the nuclear

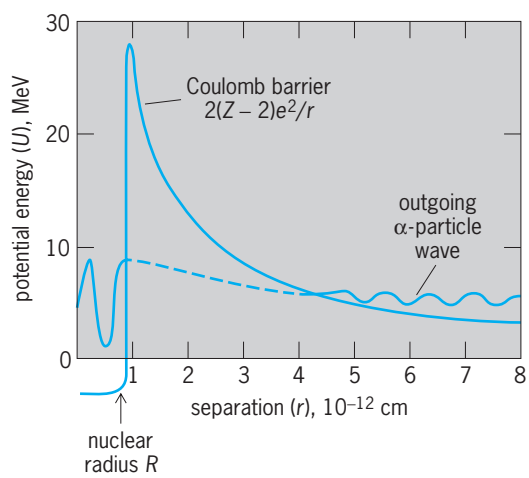

Fig. 7. Schematic of nuclear potential barrier, illustrating emission of an alpha particle as a wave which can be transmitted through the barrier.

potential barrier is very strongly dependent upon the energy of the emitted alpha particle. Analytically the probability of transmission T depends exponentially upon a barrier transmission exponent γ according to Eq. (20). To a good approximation, Eq. (21) holds,

$$T = e^{-\gamma} \tag{20}$$

$$\gamma = \left(\frac{4\pi^2}{h}\right)\frac{(Z-2)2e^2}{V} - \left(\frac{8\pi}{h}\right)[2(Z-2)2e^2MR]^{1/2} \tag{21}$$

where $h = 6.626 \times 10^{-34}$ joule-second is Planck's constant, and M is the so-called reduced mass of the alpha particle. For the alpha decay of ^{226}Ra, the numerical value of γ is about 71: hence $T = e^{-71} = 10^{-31}$. The first term on the right side of Eq. (21) is about 154 and is therefore the dominant term. When this term is taken alone, $e^{-(4\pi^2/h)(Z-2)2e^2/V}$ is called the Gamow factor for barrier penetration.

Inspection of Eq. (21) shows that the barrier transmission decreases with increasing nuclear charge $(Z-2)e$, increases with increasing velocity V of emission of the alpha particle, and increases with increasing radius R of the nucleus. When the experimentally known values of alpha-decay energy are substituted into Eq. (21), with R about 10^{-12} cm and Z about 90, the transmission coefficient $T = e^{-\lambda}$ is found to extend over a domain of about 10^{-20} to 10^{-40}. This range of about 10^{20} is just what is needed to relate the alpha-disintegration energy to the broad domain of known alpha-decay half-periods. Equation (21) thus explains the Geiger-Nuttall rule very successfully (**Fig. 8**). From Eq. (20), one can derive the relationship between the mean life τ in seconds and the alpha-particle energy E_α in MeV, $\ln \tau = AE_\alpha^{-1/2} + B$, where A and B are constants for different parent nuclei.

Since 1970, knowledge of alpha-emitting isotopes has been greatly enlarged through the identification of many isotopes far off stability in the region just above tin and in the broad region from neodymium all the way to uranium. For example, fusion reactions between 290-MeV ^{58}Ni ions and ^{58}Ni and ^{63}Cu targets have been used to produce and study very neutron deficient radioactive isotopes, including 12 alpha emitters between tin and cesium. These results provide important data on the atomic masses of nuclei far from the stable ones in nature. These data test understanding of nuclear mass formulas and their validity in new regions of the periodic table.

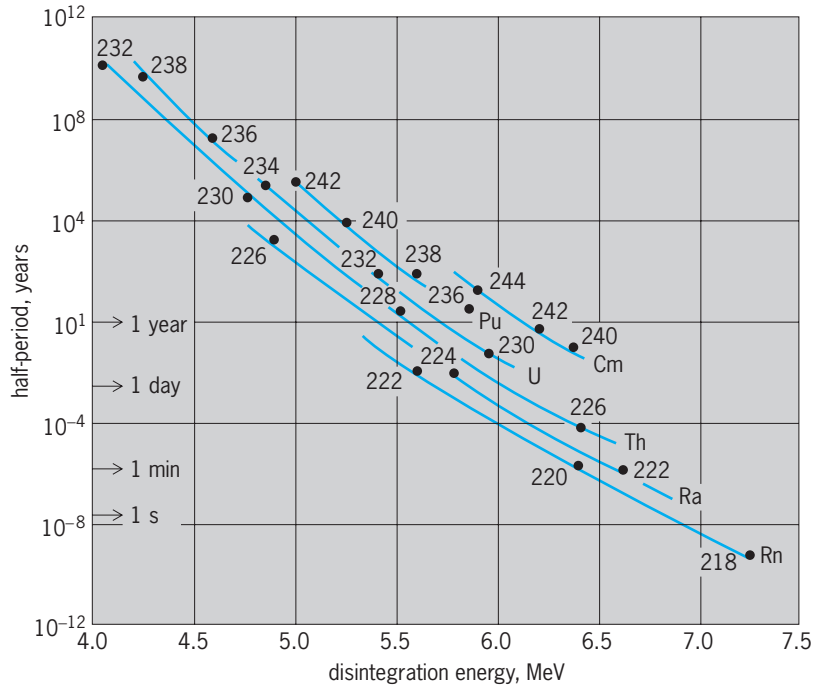

Fig. 8. Systematics of the broad range of half-periods for alpha-particle decay and their strong dependence on alpha-decay energy and weaker dependence on nuclear charge. Numbers beside experimental points are mass numbers of parent alpha-particle emitters. Lines connect parent isotopes and are drawn using wave-mechanical theory of alpha-particle transmission through nuclear potential barriers.

Beta-Particle Decay

Beta-particle decay is a type of radioactivity in which the parent nucleus emits a beta particle. There are two types of beta decay established: in negatron beta decay (β^-) the emitted beta particle is a negatively charged electron (negatron); in positron beta decay (β^+) the emitted beta particle is a positively charged electron (positron). In beta decay the atomic number shifts by one unit of charge, while the mass number remains unchanged (Table 1). In contrast to alpha decay, when beta decay takes place between two nuclei which have a definite energy difference, the beta particles from a large number of atoms will have a continuous distribution of energy (**Fig. 9**). *See* POSITRON.

For each beta-particle emitter, there is a definite maximum or upper limit to the energy spectrum of beta particles. This maximum energy, E_{max}, corresponds to the change in nuclear energy in the beta decay. Thus $E_{max} = 0.57$ MeV for β^- decay of ^{64}Cu, and $E_{max} = 0.66$ MeV for β^+ decay of ^{64}Cu. For positron decay to occur, the total decay energy must exceed 1.022 MeV (twice the rest energy of the electron). The total decay energy for β^+ decay is then $E_{max}(\beta^+)$ plus 1.022 MeV. As in the case of alpha decay, most beta-particle spectra are not this simple, but include additional continuous spectra which have less maximum energy and which leave the product nucleus in an excited level from which gamma rays are then emitted.

For nuclei very far from stability, the energies of these excited states populated in beta decay are so large that the excited states may decay by proton, two-proton, neutron, two-neutron, three-neutron or alpha emission, or spontaneous fission. In some cases, the energies are so great that the number of excited states to which beta decay can occur is so large that only the gross strength of the beta decays to many states can be studied.

Neutrinos. The continuous spectrum of beta-particle energies (Fig. 9) implies the simultaneous emission of a second particle besides the beta particle, in order to conserve energy and angular momentum for each decaying nucleus. This particle is

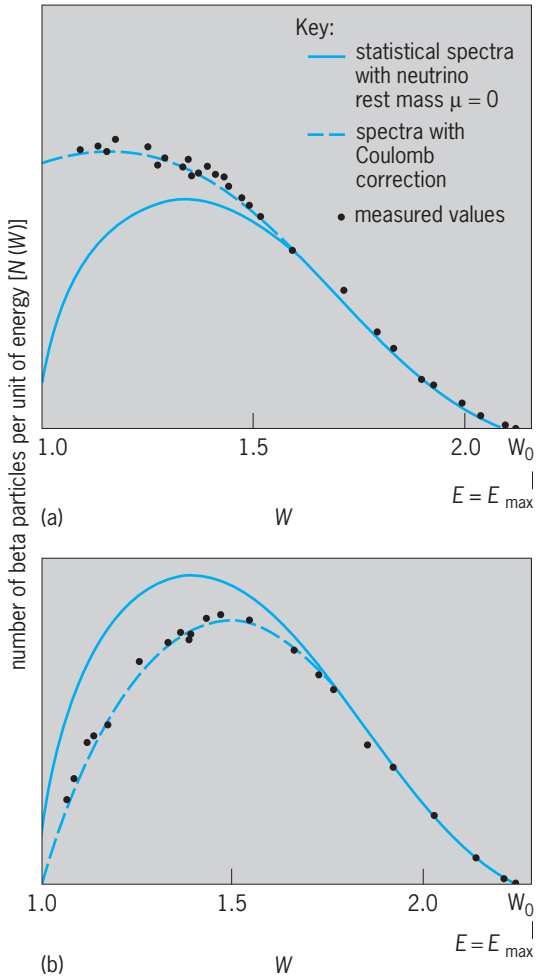

Fig. 9. Spectra of beta particles. The quantity *W* is the total beta-particle energy in units of the electron rest energy. (*a*) β^- decay of ^{64}Cu. (*b*) β^+ decay of ^{64}Cu. (*After L. M. Langer, R. D. Moffat, and H. C. Price, The beta-spectrum of ^{64}Cu, Phys. Rev., 76:1725–1726, 1949*)

the neutrino. The sum of the kinetic energy of the neutrino and the beta particle equals E_{max} for the particular transition involved except in the rare cases where internal bremsstrahlung or shake-off electrons are emitted along with the beta particle and neutrino. The neutrino has zero charge and nearly zero rest mass, travels at essentially the same speed as light (3×10^8 m/s or 1.86×10^5 mi/s), and is emitted as a companion particle with each beta particle.

Earlier careful measurements of the beta spectra of ^{3}H established an upper limit for the neutrino rest mass as less than 0.0005 times the rest energy of the electron (**Fig. 10**). In 1980, however, some indirect evidence and a new ^{3}H beta spectra measurement yielded evidence for a rest mass much smaller than this limit, but finite. If the neutrino does have a nonzero rest mass, this will have many consequences, such as the size of the total mass of the universe but will not radically change the general features of the beta decay as presented here.

Two forms of neutrinos are distinguished in beta decay. In positron beta decay, a proton p in the nucleus transforms into a neutron n in the nucleus, thus reducing the nuclear charge by 1 unit. At the time of this transition, two particles, the positron β^+ and the neutrino ν, are created and emitted. The emitted β^+ and ν together carry away the energy E_{max} of the transition and provide for conservation of energy, momentum, angular momentum, charge, and statistics. Thus positron beta decay is represented by decay (22). Negatron beta decay is a closely

$$p \rightarrow n + \beta^+ + \nu \qquad (22)$$

related process, except that a neutron n changes to a proton p in the nucleus, and a negatron beta particle β^- and its characteristic companion particle, the antineutrino ν, are emitted, as in decay (23). The antineutrino is the antiparticle of the neu-

$$p \rightarrow n + \beta^- + \bar{\nu} \qquad (23)$$

trino as the β^+ is the antiparticle of the β^-. The ν and $\bar{\nu}$ have the same properties of zero charge and essentially zero rest mass, and differ only with respect to the direction of alignment of their intrinsic spin along their direction of motion. In most beta-decay contexts, the term "neutrino" includes both its forms, neutrino and antineutrino. *See* ANTIMATTER.

There are, in fact, three classes of neutrinos. The neutrinos emitted in the two types of beta decay [decays (22) and (23)] are called electron neutrinos. In addition, there is a neutrino and antineutrino associated with the mu meson ($\mu\pm$) and neutrinos and

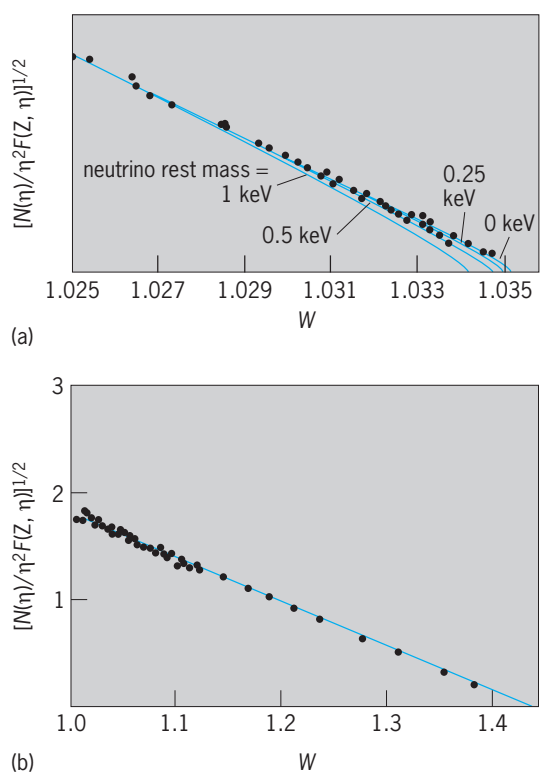

Fig. 10. Kurie plots. The quantity *W* is the total beta-particle energy in units of the electron rest energy; other quantities are defined in text. (*a*) Allowed decay of ^{3}H (*after L. M. Langer and R. J. D. Moffat, The beta-spectrum of tritium and the mass of the neutrino, Phys. Rev., 88:689–694, 1952*). (*b*) Once-forbidden decay of ^{147}Pm (*after J. H. Hamilton, L. M. Langer, and W. G. Smith, The shape of the ^{143}Pr spectrum, Phys. Rev., 112: 2010–2019, 1958*).

antineutrinos associated with the tau ($\tau\pm$). The mu neutrinos are well established. The first evidence for a tau neutrino was reported in 1998. These three particles, electron, mu, and tau, together with their neutrinos, ν_e, ν_μ, and ν_τ, and their respective antiparticles, make up a class of particles called leptons. The number of leptons in a decay or reaction is considered to be conserved; this rule is called lepton number conservation. There are separate conservation numbers for e, μ, and τ. There must be the same net number of each type of lepton on each side of a decay or reaction. For example, since there are no leptons on the left sides of decay (22) and (23), there must be no net leptons on the right sides. Thus decay (22) has on the right side an antielectron (positron) with lepton number $L = -1$ and a neutrino with lepton number $L = 1$. Decay (23) has an electron with $L = 1$ and an antineutrino with $L = -1$, and so the lepton number is $L = 1 - 1 = 0$ leptons on the right side as well. Decays (22) and (23) also conserve nucleon number. *See* LEPTON.

Because of the fact that the neutron rest mass is greater than the proton rest mass, free neutrons can undergo beta decay [decay (23)], but protons must use part of the nuclear binding energy available inside a nucleus to make up the rest mass difference in decay (22).

The interaction of neutrinos with matter is exceedingly feeble. A neutrino can pass all the way through the Sun with little chance of collision. The thickness of lead required to attenuate neutrinos by the factor $^1/_2$ is about 10^{18} m (10^{15} mi), or 100 light-years of lead!

Neutrinos play critical roles in understanding other fundamental questions. One question concerns the rest mass of the neutrino. The current upper limit on the rest-mass energy of the electron neutrino is 2 eV (the rest energy of an electron is 511,000 eV). However, since there are so many neutrinos in the universe (every second every square centimeter of a person's body is bombarded by 10^{10} neutrinos from the Sun), even this small mass can influence the total mass in the universe and whether there is sufficient mass to cause the expansion of the universe to stop and contract. The nature of the neutrino is also important in grand unified theories as outlined below. *See* COSMOLOGY; NEUTRINO; UNIVERSE.

Average beta energy. Charged particles, such as beta particles or alpha particles, are easily absorbed in matter, and their kinetic energy is thereby converted into heat. In beta decay the average energy E_{av} of the beta particles is far less than the maximum energy E_{max} of the particular beta-particle spectrum. The detailed shape of beta-particle spectra and hence the exact value of the ratio E_{av}/E_{max} varies somewhat with Z, E_{max}, the degree of forbiddenness of the transition, and the sign of charge of the emitted beta particle. A rough rule of thumb which covers many practical cases is $E_{av} = (0.40 \pm 0.05)E_{max}$, with slightly higher values for positron beta-particle spectra than for negatron beta-particle spectra. The remaining disintegration energy is emitted as kinetic energy of neutrinos and is not recoverable in finite absorbers.

There are other processes that carry off part of the energy of beta decay, including internal bremsstrahlung (gamma rays) and shake-off electrons (atomic electrons). The total probabilities for these additional two processes are on the order of 1% or much less per beta decay, and the probability of their emission decreases rapidly with increasing energy so they are mainly low-energy (less than about 50 keV) radiations. In internal bremsstrahlung, through an interaction of the beta particle and the emitting nucleus, part of the decay energy is emitted as a gamma ray. In the shake-off process, part of the beta-decay energy is given to one of the atomic electrons. The gamma rays are not absorbed in matter as easily as the beta particles. In addition, if one tries to absorb the beta particles in matter, the beta particles can interact with the atoms and give off external bremsstrahlung (gamma rays). The number of these gamma rays again is a strongly decreasing function of energy, but their emission extends up to the maximum energies of the beta particles. *See* BREMSSTRAHLUNG.

Fermi theory. By postulating the simultaneous emission of a beta particle and a neutrino, as in reaction (22), E. Fermi developed in 1934 a quantum-mechanical theory which satisfactorily gives the shape of beta-particle spectra (Fig. 9), and the relative half-periods of beta-particle emitters for allowed beta decays. The energy distribution of beta particles in allowed transitions is then given by Eq. (24). Here

$$N(W)\,dW = \frac{|P|^2}{\tau_0} F(Z, W)(W^2 - 1)^{1/2}(W_0 - W)^2 W\,dW \quad (24)$$

$N(W)\,dW$ = number of beta particles in energy range W to $W + dW$; $W = 1 + E/(m_0c^2)$ = total energy of beta particle in units of rest energy $m_0c^2 = 0.51$ MeV for an electron (m_0 = electron mass, c = velocity of light); $W_0 = 1 + E_{max}/(m_0c^2)$ = maximum energy of the beta-particle spectrum; $|P|^2$ = squared matrix element for the transition, and is of the order of unity for allowed transitions; τ_0 = time constant $\cong$ 7000 s; $F(Z,W)$ = complex, dimensionless function involving the nuclear radius, nuclear charge, beta-particle energy, and whether the decay is β^- or β^+.

Physically this distribution function involves the product of the energy W and momentum $(W^2 - 1)^{1/2}$ of the beta particle times the energy $(W_0 - W)$ and the momentum $(W_0 - W)/c$ of the neutrino. The product of these factors gives a "statistical" distribution for the number of beta particles as a function of energy (Fig. 9). The observed spectra show an excess of low-energy β^- and a deficiency of low-energy β^+ particles. This arises because of the Coulomb attraction and repulsion of the nucleus for β^- and β^+. The statistical spectrum is corrected by the Fermi function, $F(Z,W)$, and the new distribution agrees with experiments (Fig. 9).

Equation (24) essentially matches the energy spectra of allowed beta-particle transitions and therefore

furnishes one type of experimental verification of the properties of neutrinos. Its counterpart in terms of the beta-particle momentum spectrum is often used for the analysis of spectra, and is given by Eq. (25).

$$N(\eta)d\eta = \frac{|P|^2}{\tau_0}F(Z,\eta)(W_0 - W)^2\eta^2 d\eta \qquad (25)$$

The momentum distribution is much more nearly symmetric than its corresponding energy spectrum. Here $N(\eta)d\eta$ = number of beta particles in the momentum interval from η to $\eta + d\eta$; $\eta = (W^2 - 1)^{1/2}$ = momentum of the beta particle in units of m_0c; and $F(Z,\eta) = F(Z,W)$ of Eq. (24).

Konopinski-Uhlenbeck theory. After the work of Fermi which explained allowed decay, E. J. Konopinski and G. E. Uhlenbeck in 1941 developed the theory of forbidden beta decay. Allowed decays occur between nuclear states which differ in spin by 0 or 1 unit and which have the same parity. Konopinski and Uhlenbeck developed a theory to describe beta decays where energy is available for decay but the allowed selection rules on spin or parity or both are violated. These beta transitions occur at a slower rate and are called forbidden transitions. In 1949 the theory of forbidden beta decay was confirmed by L. M. Langer and H. C. Price. The orders of forbiddenness, which retard the rate of decay, are the following: once-forbidden decay when the change in nuclear spin ΔJ is again 0 or 1 as in allowed decay, but a parity change $\Delta\pi$ occurs; once-forbidden unique decay when $\Delta\pi$ changes and $\Delta J = 2$; n-times forbidden decay when $\Delta J = n$, $\Delta\pi = (-)^n$, where $\Delta\pi = -$ indicates a parity change; and n-times forbidden unique decay when $\Delta J = n + 1$, $\Delta\pi = (-)^n$ [**Table 4**]. In forbidden decays the first-order allowed matrix elements of the Fermi theory in Eq. (24) vanish because of the selection rules on angular momentum and spin. Then the much smaller higher-order matrix elements that can be neglected compared to the large allowed matrix elements come into play. *See* PARITY (QUANTUM MECHANICS); SELECTION RULES (PHYSICS).

Comparative half-lives, f*T*. The half-period T of beta decay can be derived from Eq. (24) because the radioactive decay constant $\lambda = 0.693/T$ is simply the total probability of decay, or $N(W)\,dW$ integrated over all possible values of the beta-particle energy from $W = 1$ to $W = W_0$.

For allowed decays, the matrix elements are not functions of the beta energy and can be factored out of Eq. (24), so Eq. (26) is valid, where f is given by Eq. (27), and the constants include $|P|^2$ of Eq. (24). Equation (26) can be rearranged as Eq. (28).

$$\lambda = \frac{0.693}{T} = \text{constants} \times \text{f} \qquad (26)$$

$$\text{f} = \int_1^W F(Z,W)(W^2 - 1)^{1/2}(W_0 - W)^2 W\,dW \qquad (27)$$

$$\text{f}T = \frac{0.693}{\text{constants}} = \text{comparative half-life} \qquad (28)$$

For different beta decays, T varies over a range greater than 10^{18} and inversely depends on the beta-decay energy in analogy to the Geiger-Nuttall rule for alpha decay. However, Eq. (28) says that the comparative half-life should be a constant. Indeed it is found experimentally that different classes of beta decay do have very similar fT values. It is generally easier to give the $\log_{10}$ fT for comparison. The groups (Table 4) include, in addition to the forbidden decays, three classes of allowed decays: the favored or superallowed decays of nuclei whose structures are very similar so that the matrix element in the denominator of Eq. (28) is large and log fT is small; normal allowed; and allowed *I*-forbidden where the total angular momentum selection rule holds, but the individual particle that is undergoing beta decay has a change of 2 units of orbital angular momentum. The matrix elements for each degree of forbiddenness get progressively smaller and so log fT values increase sharply with each degree of forbiddenness. The ranges of these fT values for each degree of forbiddenness are in general so well established that measurements of fT values can be used to establish changes in spins and parities between nuclear states in beta decay.

TABLE 4. Selection rules for beta decay and log f*T* values

Type	ΔJ	$\Delta\pi$	Log fT	Examples
Allowed (favored)	0 or 1	No	3	n, ^{3}H
Allowed (normal)	0 or 1	No	4 to 7	^{35}S, ^{30}P
Allowed (*I*-forbidden)	1	No	6 to 9	^{32}P, ^{65}Ni
Once-forbidden	0 or 1	Yes	6 to 8	^{111}Ag, ^{143}Pr
Once-forbidden (unique)	2	Yes	8 to 9	^{42}K, ^{91}Y
Twice-forbidden	2	No	11 to 14	^{36}Cl, ^{59}Fe
Twice-forbidden (unique)	3	No	12 to 14	^{22}Na, ^{60}Co
Third-forbidden	3	Yes	17 to 19	^{87}Rb, ^{138}La
Third-forbidden (unique)	4	Yes	(~18)	^{40}K
Fourth-forbidden	4	No	~ 24	^{115}In
Fourth-forbidden (unique)	5	No		

Kurie plots. For allowed transitions, the transition matrix element $|P|^2$ is independent of the momentum η. Then Eq. (25) can be put in the form of Eq. (29).

$$\left(\frac{N(\eta)}{\eta^2 F(Z,\eta)}\right)^{1/2} = \text{const}\,(W_0 - W) \qquad (29)$$

Therefore a straight line results when the quantity $(N/\eta^2F)^{1/2}$ is plotted against beta-particle energy, either as W or as E, on a linear scale. Such graphs are called Kurie plots, Fermi plots, or Fermi-Kurie plots. These are especially useful for revealing deviations from the allowed theory and for obtaining the upper energy limit E_{max} as the extrapolated intercept of N/η^2F on the energy axis. Practically all of the results on the shapes of beta-particle spectra are published as Kurie plots, rather than as actual momentum or energy spectra.

When spectral data give a straight line (Fig. 10), then $N(\eta)$ is in agreement with the Fermi momentum

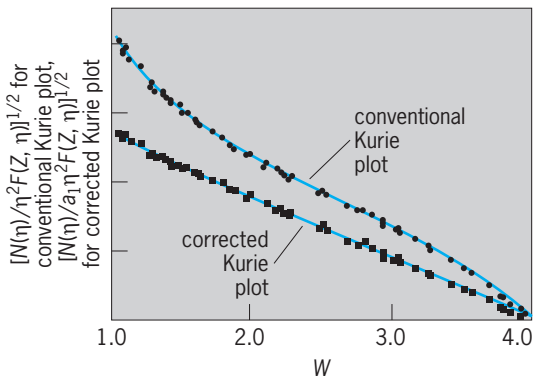

Fig. 11. Once-forbidden spectrum of ^{91}Y: conventional Kurie plot and Kurie plot corrected by the unique shape factor, $a_1 = W^2 - 1 + (W_0 - W)^2$, given by Konopinski and Uhlenbeck, which linearizes the data. The quantity W is the total beta-particle energy in units of the electron rest energy; other quantities are defined in text. (*After L. M. Langer and H. C. Price, Shape of the beta spectrum of the forbidden transition of yttrium 91, Phys. Rev., 75:1109, 76:641, 1949*)

distribution, Eq. (25); and the intercept of this straight line, on the energy axis, gives the disintegration energy E_{max}. For β^+ decay the total decay energy is E_{max} (β^+) plus 1.022 MeV. In Fig. 10*a*, theoretical curves are given for various values of the neutrino rest mass, and the data points, which are experimental values, lie on the curve corresponding to zero mass.

In addition to allowed decays, all but one known once-forbidden decays have Kurie plots that are essentially linear in energy (Fig. 10*b*). The once-forbidden unique decays have a pronounced characteristic energy dependence for their matrix elements, and thus the conventional Kurie plot has a characteristic shape that differs from a straight line (**Fig. 11**). When the data are corrected by the unique shape factor, given by Konopinski and Uhlenbeck, a linear Kurie plot is again obtained. This unique shape was the key to the discovery of forbidden beta decay by Langer and Price (Fig. 11). The higher-order forbidden spectra each show different strong energy dependences in their Kurie plots, each characteristic of their degree of forbiddenness.

Double beta decay. When the ground state of a nucleus differing by two units of charge from nucleus A has lower energy than A, then it is theoretically possible for A to emit two beta particles, either $\beta^+\beta^+$ or $\beta^-\beta^-$ as the case may be, and two neutrinos or antineutrinos, and go from Z to $Z \pm 2$. Here two protons decay into two neutrons, or vice versa. This is a second-order process and so should go much slower than beta decay. There are a number of cases where such decays should occur, but their half-lives are of the order of 10^{20} years or greater. Such decay processes are obviously very difficult to detect.

The first direct evidence for two-neutrino double-beta-minus decay of selenium, reaction (30), was

$$^{82}_{34}\text{Se} \rightarrow {}^{82}_{36}\text{Kr} + 2e^- + 2\bar{\nu} \qquad (30)$$

found only in 1987, some 40 years after the first attempt to observe this rare process. The trajectories of the electrons from this decay have been detected in a time-projection chamber. A half-period of $1.08^{+0.26}_{-0.06} \times 10^{20}$ y has been obtained. Previous geochemical and cosmochemical evidence of double beta decay had been only indirect because it consisted of observations of buildup of the noble gases $^{82}_{32}$Kr and $^{130}_{54}$Xe in the $^{84}_{32}$Se and $^{130}_{52}$Te samples, respectively. *See* TIME-PROJECTION CHAMBER.

Some unifications of the electroweak and strong forces suggest that the electron neutrino and antineutrino are identical ($\nu_e \equiv \bar{\nu}_e$), that is, that the neutrino is a Majorana particle. There is no experimental evidence at present for this. If true, this allows double beta decay to proceed without the emission of neutrinos. In this neutrinoless double beta decay, the first neutrino absorbs the second one so the decay is decay (31).

$$^{82}_{24}\text{Se} \longrightarrow {}^{82}_{36}\text{Kr} + 2e^- \qquad (31)$$

Searches have shown that the half-period of this decay is greater than 2.7×10^{22} y, which is over 200 times larger than the half-period of the ^{82}Se double beta decay with two-neutrino emission. Lepton number is not conserved in decay (31). This process is predicted to occur in some grand unified theories. The experimentally established lower limits of half-lives for neutrinoless double beta decays have been used to derive upper limits for the mass of the Majorana neutrino ($m_{\nu e}c^2 < 2$ eV) and for the right-handedness of the weak interaction (less than 2×10^{-8} of the left-handed strength). *See* GRAND UNIFICATION THEORIES.

Electron-capture transitions. Whenever it is energetically allowed by the mass difference between neighboring isobars, a nucleus Z may capture one of its own atomic electrons and transform to the isobar of atomic number $Z - 1$ (Table 1). Usually the electron-capture (EC) transition involves an electron from the K shell of atomic electrons, because these innermost electrons have the greatest probability density of being in or near the nucleus. *See* ELECTRON CAPTURE.

In EC transitions, a proton p bound in the parent nucleus absorbs an electron e^- and changes to a bound neutron n. The disintegration energy is carried away by an emitted neutrino ν as in transition (32). The residual nucleus may be left either

$$p + e^- \rightarrow n + \nu \qquad (32)$$

in its ground level or in an excited level from which gamma-ray emission follows.

EC transitions compete with all cases of positron beta-particle decay. EC has an energetic advantage over β^+ decay equivalent to the mass of two electrons, or 1.02 MeV, because in transition (32) one electron mass e^- enters on the left and is available, whereas in decay (22) one electron mass β^+ must be produced as a product of the positron beta-particle decay. For example, in the radioactive decay of $^{64}_{29}$Cu, twice as many transitions go by EC to $^{64}_{28}$Ni as go by positron beta decay to the same decay

product. In the heavy, high-Z elements, EC is greatly favored over the competing β^+ decay, and examples of measurable β^+ decay are practically unknown for Z greater than 80, although there are a large number of examples of electron capture. As the energy for decay increases beyond 1.02 MeV, the probability of β^+ decay increases relative to EC and dominates at several megaelectronvolts of energy.

Several examples are known of completely pure EC radioactivity in which there is insufficient nuclear energy to allow any positron beta-particle decay (total decay energy is less than 1.022 MeV). For example, $^{55}_{26}$Fe emits no positron beta particles, but transforms with a half-period of 2.6 years entirely by EC to the ground level of $^{55}_{25}$Mn. This radioactivity is detectable through the K-series x-rays which are emitted from ^{55}Mn when the atomic electron vacancy, produced by nuclear capture of a K electron, refills from the L shell of atomic electrons. Also, double electron capture, analogous to double beta decay, is theoretically predicted to exist. Here two atomic electrons are captured and two neutrinos emitted.

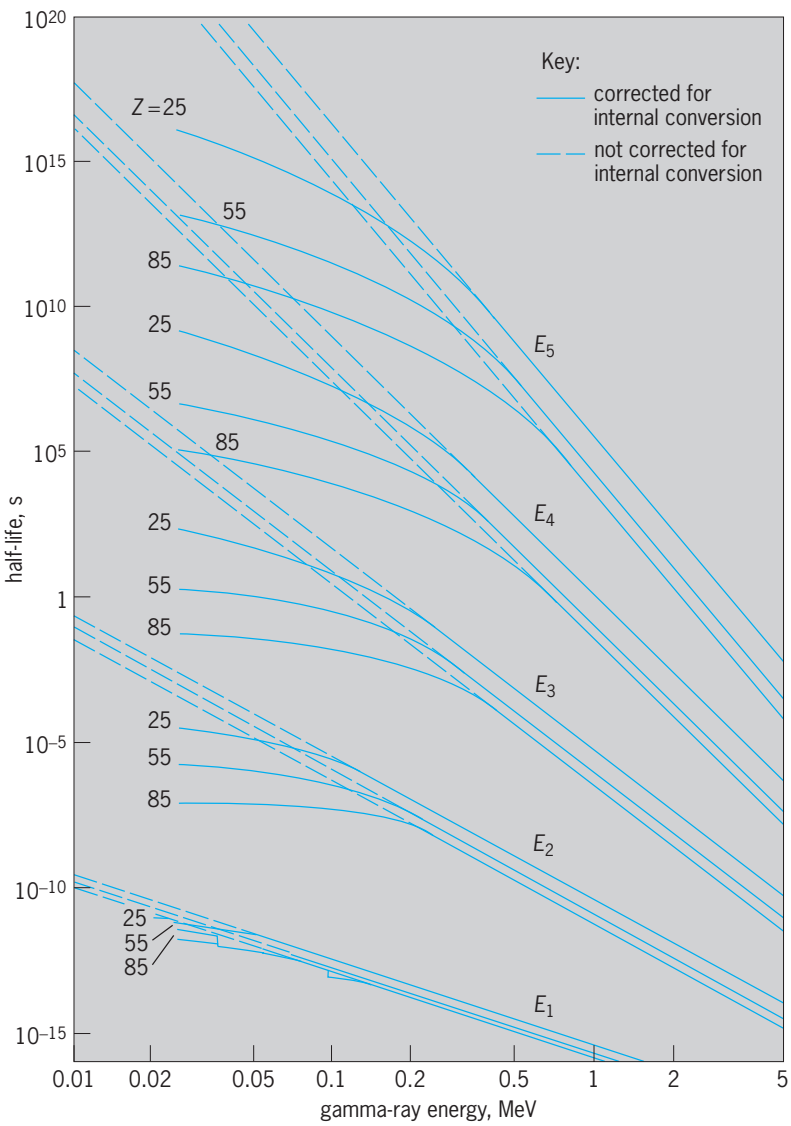

Fig. 12. Theoretical values of half-lives for decay of nuclear levels by emission of gamma rays and conversion electrons, for electric multipoles. (*After A. H. Wapstra, G. J. Nijgh, and R. van Lieshout, Nuclear Spectroscopy Tables, North-Holland, 1959*)

Gamma-Ray Decay

Gamma-ray decay involves a transition between two excited levels of a nucleus, or between an excited level and the ground level. A nucleus in its ground level cannot emit any gamma radiation. Therefore gamma-ray decay occurs only as a sequel of one of the processes in Table 1 or of some other process whereby the product nucleus is left in an excited state. Such additional processes include gamma rays observed following the fusion of two nuclei, as occurs in bombarding ^{58}Ni with ^{16}O to form an excited compound nucleus of ^{74}Kr. This compound nucleus first promptly gives off a few particles like two neutrons to leave ^{72}Kr* or two protons to leave ^{72}Se*, both of which will be in excited states which will emit gamma rays. Or one may excite states in a nucleus by the Coulomb force between two nuclei when they pass close to each other but do not touch (their separation is greater than the sum of the radii of the two nuclei). There are also other nuclear reactions such as induced nuclear fission that leave nuclei in excited states to undergo gamma decay. *See* COULOMB EXCITATION.

A gamma ray is high-frequency electromagnetic radiation (a photon) in the same family with radio waves, visible light, and x-rays. The energy of a gamma ray is given by $h\nu$, where h is Planck's constant and ν is the frequency of oscillation of the wave in hertz. The gamma-ray or photon energy $h\nu$ lies between 0.05 and 3 MeV for the majority of known nuclear transitions. Higher-energy gamma rays are seen in neutron capture and some reactions. *See* ELECTROMAGNETIC RADIATION.

Gamma rays carry away energy, linear momentum, and angular momentum, and account for changes of angular momentum, parity, and energy between excited levels in a given nucleus. This leads to a set of gamma-ray selection rules for nuclear decay and a classification of gamma-ray transitions as "electric" or as "magnetic" multipole radiation of multipole order 2^l where $l = 1$ is called dipole radiation, $l = 2$ is quadrupole radiation, and $l = 3$ is octupole, l being the vector change in nuclear angular momentum. The most common type of gamma-ray transition in nuclei is the electric quadrupole (E2). There are cases where several hundred gamma rays with different energies are emitted in the decays of atoms of only one isotope. *See* MULTIPOLE RADIATION.

Mean life for transitions. A reasonably successful approximate theory of the mean life for gamma-ray decay was developed by V. F. Weisskopf in 1951, using the single-particle shell model of nuclei (**Fig. 12**). An E2 transition of about 1 MeV is expected to take place with a mean life, τ_{el}, or mean delay in the upper level, of about 10^{-11} s. Thus most gamma-ray transitions are prompt transitions, in which the mean life of the excited level is too short to be measured easily. The mean life τ_{mag} for magnetic multipoles is of the order of 30 (for $A = 20$) to 150 (for

$A = 200$) times longer than τ_{el} for electric multipole transitions (Fig. 12).

At low energies or high Z, or both, the internal conversion process becomes a very important additional mode of decay that markedly shortens the mean lives of the nuclear levels. In addition, in many cases the structure of the nucleus comes into play and alters the observed mean lives considerably compared to those predicted by the Weisskopf theory (Fig. 12). Electric dipole (E1) transitions are generally retarded (longer mean lives) by factors of 10^6 over the Weisskopf estimates (Fig. 12). On the other hand, A. Bohr and B. Mottelson developed a model of collective nuclear motions where E2 transitions are enhanced by factors of 100 or more (shorter τ) over the Weisskopf single-particle estimates, and these predictions are confirmed by experiments. The magnetic dipole (M1) transitions are also often hindered by factors of 100 or more. Measurements of the mean lives for gamma-ray decay provide important tests of nuclear models.

Internal conversion. An alternative type of deexcitation which always competes with gamma-ray emission is known as internal conversion. Instead of the emission of a gamma ray, the nuclear excitation energy can be transferred directly to a bound electron of the same atom. Then the nuclear energy difference is converted to energy of an atomic electron, which is ejected from the atom with a kinetic energy E_i given by Eq. (33). Here B_i is the original atomic

$$E_i = W - B_i \tag{33}$$

binding energy of the particular electron, which is ejected, and W is the nuclear transition energy which would otherwise have been emitted as a gamma-ray photon having energy $h\nu = W$.

The spectrum of internal conversion electrons is then a series of discrete energies, or "lines," each corresponding to an individual value of B_i, for the K, L (L_1, L_2, L_3), $M, \ldots$ electrons in each shell and subshell of the atom. Thus conversion electron spectra are much more complex than gamma spectra (**Fig. 13**). From the spacing of the E_i values in this conversion electron spectrum, it is possible to assign definitely the atomic number Z of the atom in which the nuclear transition W took place. In this way it is known that the conversion electron and the competing gamma-ray emission are sequels and not antecedents of alpha decay, beta decay, and electron-capture transitions.

The total internal conversion coefficient α_T is the ratio of the number of transitions proceeding by internal conversion, N_{eT}, to the number going by gamma-ray emission, N_γ, for any particular nuclear transformation from an excited level to a lower-lying level, as in Eq. (34). The total internal conversion co-

$$\alpha_T = \frac{N_{eT}}{N_\gamma} \tag{34}$$

efficient is a sum of the conversion coefficients for each shell [K, $L(L_1 + L_2 + L_3)$, $M(M_1 + \cdots)$, and so forth], and is given by Eq. (35), where $\alpha_K = N_{eK}/N_\gamma$,

$$\alpha_T = \alpha_K + \alpha_L + \alpha_M + \cdots \tag{35}$$

$\alpha_L = N_{eL}/N_\gamma = \alpha_{L1} + \alpha_{L2} + \alpha_{L3}$ ($\alpha_{L1} = N_{eL1}/N_\gamma, \ldots$), and N_{eK}, N_{eL} ($N_{eL1}, \ldots$), and so forth are the numbers of electrons ejected from the K, L (L_1, L_2, L_3), $\ldots$ shells, respectively. In general, this probability of internal conversion relative to gamma-ray emission increases with increasing atomic number Z, with increasing multipole order 2^l, and with decreasing nuclear deexcitation energy W. In middle-weight elements, for $W = 1$ MeV, α is of the order of 10^{-2} to 10^{-4}; while for $W = 0.2$ MeV, α is of the order of 0.1 for electric $l = 2$ transitions, and 10 or larger for electric $l = 5$ transitions. When internal conversion electron decay occurs, this process is always followed by the emission of characteristic x-rays of the element and Auger electrons from outer shells when the inner shell vacancy is filled. This emission can include K x-rays, L x-rays and x-rays from higher shells, and K Auger electrons, L Auger electrons, and so forth. *See* AUGER EFFECT.

Radiationless transitions. There are cases where gamma-ray emission is strictly forbidden and conversion electron emission allowed (Fig. 13). This occurs when both nuclear states have zero spin and the same parity. The conversion electrons are called electric monopole radiations, EO. These transitions occur because of the penetration of the atomic electrons into the nuclear volume where they interact directly with the nucleus. EO radiation can occur in principle whenever two states have the same spin and same parity, but in practice, EO decays are found to be very, very small in these cases. There are some exceptions in well-deformed nuclei and in nuclei which have states with quite different shapes. In these cases, EO decays can totally dominate the electron emission for transitions that have no change in spin and that involve decays between states with large differences in their nuclear shapes (Fig. 13).

The EO decays which arise because of the penetration of the atomic electrons into the nuclear volume are thus sensitive measures of changes in shape between two nuclear states, and have played important roles in establishing vibrations of the nuclear shape and the coexistence of states with quite different deformation in the same nucleus. There also are other circumstances where the penetration of the atomic electron into the nuclear volume gives rise to additional contributions to the conversion-electron decay. Again these penetration effects probe details of the structure of the nucleus.

Internal pair formation. When the energy between two states in the same nucleus exceeds 1.022 MeV, twice the rest mass energy of an electron, it is possible for the nucleus to give up its excess energy to an electron-positron pair—a pair creation process. This is a third alternative mode to gamma decay and conversion electron decay. This process becomes more important as the gamma-ray energy increases. It is relatively unimportant below 2–3 MeV of decay energy. *See* ELECTRON-POSITRON PAIR PRODUCTION.

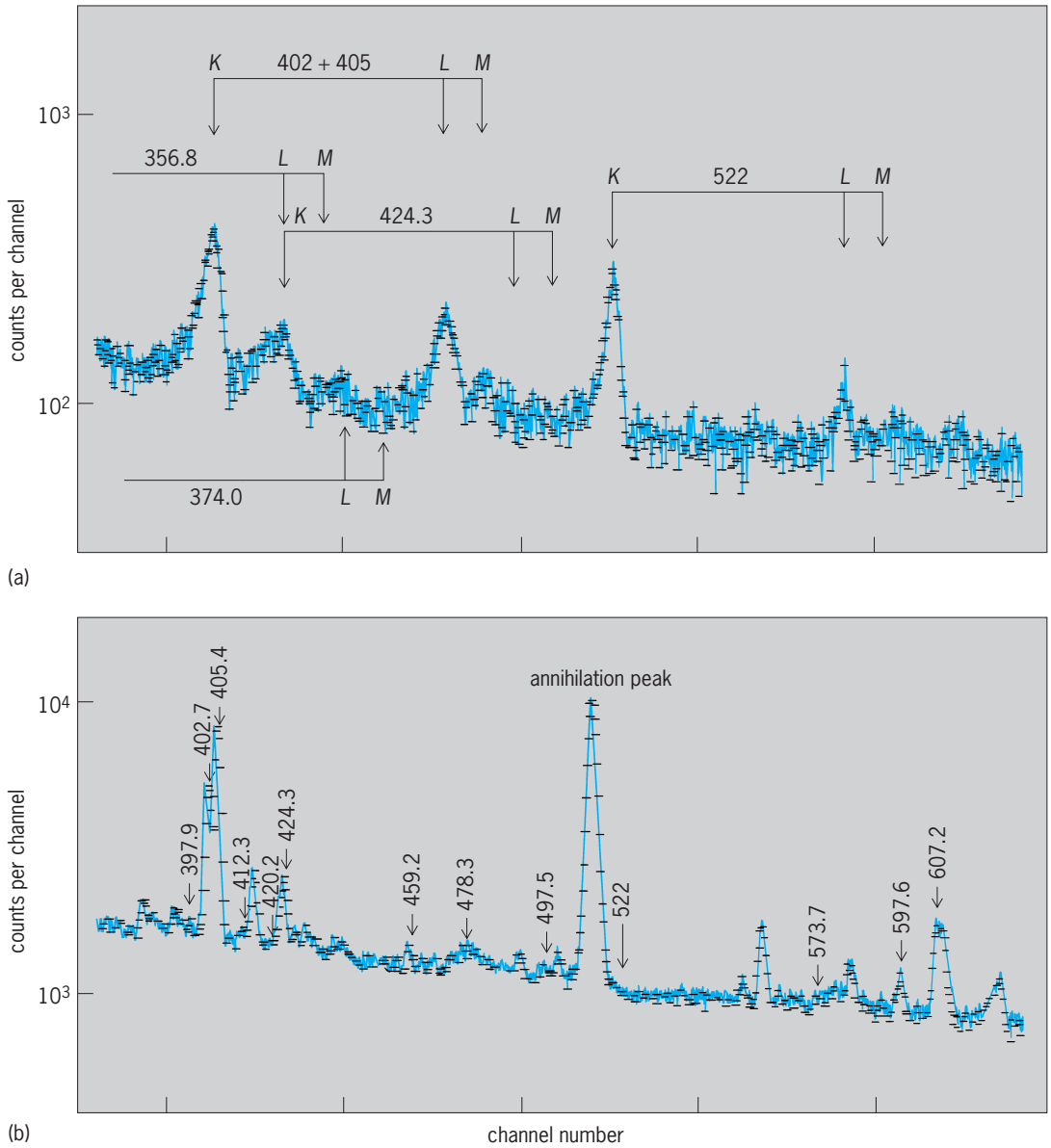

Fig. 13. Spectra from the decay of 30-s half-life ^{186}Tl far off stability (15 neutrons less than the lightest stable thallium isotope). Nuclear transition energies are given in kiloelectronvolts. (*a*) Internal conversion electrons. (*b*) Gamma rays. The strong 522-keV electron transition has no gamma ray associated with it. The strong 511-keV gamma ray is from annihilation of positrons and is not a nuclear transition. (*After J. H. Hamilton et al., Shape coexistence in ^{186}Hg and the decay of ^{186}Tl, Phys. Rev., C16: 2010–2018, 1977*)

Isomeric transitions. Measurably delayed radioactive transitions from an excited level of a nucleus are known as isomeric transitions. The measurably long-lived excited level is called an isomeric or metastable level or an isomer of the ground level. What constitutes an isomer is not well defined. The terminology arose when it was difficult to measure mean lives shorter than 10^{-7} s. States with longer mean lives were isomers. Now mean lives down to 10^{-13} s can be measured for many transitions in different nuclei, but these are not generally called isomers. The break point is simply not defined.

Figure 12 shows that if the excitation energy is small (say, 0.5 MeV or less) and the angular momentum difference l is large (say, $l = 3$ or more), then the mean life of an excited level for gamma-ray or conversion-electron emission can be of the order of 1 s up to several years.

Most of the long-lived isomers occur in nuclei which have odd mass number A. Then either the number of protons Z in the nucleus is odd, or the number of neutrons N in the nucleus is odd. The frequency distribution of odd-A isomeric pairs, excited level and ground level, displays so-called islands of isomerism in which the odd-proton or odd-neutron number is less than 50 or less than 82. The distribution is one of several lines of evidence for closed shells of identical nucleons at N or $Z = 50$ or 82 in nuclei, and it plays an important role in the so-called shell model of nuclei. *See* MAGIC NUMBERS; NUCLEAR ISOMERISM.

Spontaneous Fission

This involves the spontaneous breakup of a nucleus into two heavy fragments (two intermediate atomic number elements, for example, with $Z = 42$ and 50) and neutrons (Table 1). Spontaneous fission can occur when the sum of the masses of the two heavy fragments and the neutrons is less than the mass of the parent undergoing decay. After the discovery of fission in 1939, it was discovered that isotopes like ^{238}U had very weak decay branches for spontaneous fission, with branching ratios on the order of 10^{-6}. Some isotopes with relatively long half-lives such as ^{252}Cf have large (3.1%) spontaneous fission branching. In these cases, the nucleus can go to a lower energy state by spontaneously splitting apart into two heavy fragments of rather similar mass plus a few neutrons. This process liberates a large amount of energy compared to any other decay mode. Thus, ^{252}Cf has become important in many applications in medicine and industry as a compact energy source, a source of neutrons or as a source of nuclear radiation, since the fragments themselves are left in excited states and so emit gamma rays.

An important isomeric decay mode was discovered in the early 1960s in the very heavy elements, spontaneous fission isomers. Here the nucleus in an excited state, rather than emit a gamma ray or conversion electron, spontaneously breaks apart into two heavy fragments plus neutrons exactly as in spontaneous fission. To identify these isomers, the symbol f is often placed after their atomic mass, for example, $^{244f}_{95}$Am. Their half-lives are generally short, 10^{-3} to 10^{-9} s. It is now understood that these fission isomers are states with much larger deformation than the ground states of these isotopes. The Coulomb barrier against fission is in fact a double-hump barrier with the fission isomers in the valley at large deformation (**Fig. 14**). The study of these fission isomers has provided important tests of understanding of the behavior and structure of nuclei with very large deformation.

Fission with the emission of neutrons is called hot fission because the fragments have high excitation energy that is carried away by the neutron emission. In 1994 cold spontaneous fission was observed for ^{252}Cf and, subsequently, ^{242}Pu, in which no neutrons are emitted, in contrast to hot spontaneous fission, where one to ten neutrons are emitted. Cold ternary spontaneous fission and hot ternary spontaneous fission have also been observed. In ternary spontaneous fission, in addition to the two intermediate-mass nuclei, one light nucleus such as an alpha particle, ^{3}H, or ^{6}Li up to ^{14}C, has been observed. In cold ternary spontaneous fission no neutrons are emitted, while in hot ternary spontaneous fission one or more neutrons are emitted. These rare decay modes provide new insight into the fission process, clustering in nuclei, and the theory of multifragmentation, as discussed below. *See* NUCLEAR FISSION.

Joseph H. Hamilton

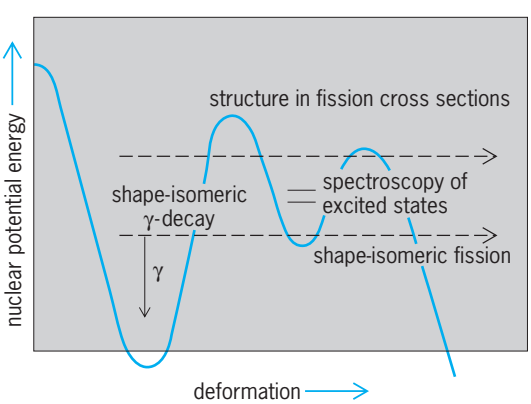

Fig. 14. Observable consequences of a double-humped nuclear potential barrier against fission. The potential well at the larger deformation gives rise to isomeric spontaneous fission.

Heavy Cluster Decays

Alpha-particle decay and spontaneous fission are two natural phenomena in which an atomic nucleus spontaneously breaks into two fragments, but the fragments are of very different size in one case and almost equal size in the other. On the basis of fragmentation theory and the two-center shell model, new kinds of radioactivities that are intermediate between alpha particle decay and fission were predicted in 1980. Subsequently, it was shown theoretically that the new processes should occur throughout a very broad region of the nuclear chart, including elements with atomic numbers higher than 40. However, experimentally observable emission rates could be expected only for nuclei heavier than lead, in a breakup leading to a very stable heavy fragment with proton and neutron numbers equal or very close to $Z = 82$, $N = 126$ ($^{208}_{82}$Pb or its neighborhood).

For more than 150 kinds of cluster emission, the predicted half-periods of the parent nuclei are shorter than 10^{23} years. The main competitor is always alpha-particle decay. The predicted branching ratios relative to this alpha-particle decay are smaller than 10^{-9}, with a maximum value for $^{14}_{6}$C radioactivity of $^{223}_{88}$Ra, the first experimentally observed case. Clusters are emitted through fission processes in which the fragments retain compact shape configurations, with a relatively high kinetic energy of about 2 MeV per nucleon.

In 1984, a series of experimental confirmations began with the discovery of $^{14}_{6}$C radioactivity of $^{223}_{88}$Ra. Initially a semiconductor telescope identification technique was used. In subsequent experiments on $^{14}_{6}$C emission from other radium isotopes, magnetic spectrometers (a superconducting solenoid and an Enge split pole) allowed suppression of the strong background of alpha particles.

A very promising technique uses solid-state track-recording detectors with special plastic films and glasses that are sensitive to heavier clusters but not to alpha particles. This technique has been applied to the entire range of cluster emissions measured previously and has yielded results in good agreement with theoretically predicted half-lives in the emission of $^{14}_{6}$C from radium isotopes with mass numbers

222–224 and 226; ${}^{20}_{8}$O from ${}^{288}_{90}$Th; ${}^{23}_{9}$F from ${}^{231}_{91}$Pa; ${}^{24}_{10}$Ne from ${}^{230}_{92}$Th, ${}^{231}_{91}$Pa, and ${}^{232-234}_{92}$U; ${}^{28}_{12}$Mg from ${}^{234}_{92}$U; and many others. *See* PARTICLE TRACK ETCHING.

Cluster radioactive decay to excited states of the daughter nucleus or involving excited clusters has been predicted, and experimentally verified.

In cold fission, the fragments are not very deformed or excited, just as in cluster radioactivity. Cold fission has been interpreted, according to the two-center shell model, as cluster emission. As noted above, cold fission has been clearly established in the spontaneous fission of ^{252}Cf and ^{242}Pu, and cold ternary spontaneous fission has also been observed. These decay modes provide tests of a unified theory that includes cold fission, light cluster radioactivity, and alpha decay.

While quite rare, these modes also provide interesting tests of the understanding of the structure of heavy nuclei, including shell and pairing effects, deformation, large-amplitude collective motion, and clustering, with particular emphasis on very heavy clusters like ${}^{132}_{50}$Sn and ${}^{208}_{82}$Pb. Because of its particularly strong shell effects, the ${}^{208}_{82}$Pb cluster is involved not only in cluster decays but also in the asymmetric cold fusion process, which has been used in the synthesis of the heaviest elements. *See* ELEMENT 112.

Walter Greiner; Joseph H. Hamilton

Proton Radioactivity

Proton radioactivity is a mode of radioactive decay that is generally expected to arise in proton-rich nuclei far from the stable isotopes, in which the parent nucleus changes its chemical identity by emission of a proton in a single-step process. Its physical interpretation parallels almost exactly the quantum-mechanical treatment of alpha-particle decay. It is also theoretically predicted that one can have the simultaneous emission of two protons—two-proton radioactivity. Although proton radioactivity has been of considerable theoretical interest since 1951 and is expected to be a general phenomenon, for many years only a few examples of this decay mode were observed, because of the narrow range of half-lives and decay energies where this mode can compete with other modes. However, in the late 1990s, experimental techniques using new recoil mass spectrometers, which can separate rare reaction products, and new double-sided silicon strip detectors became available and opened up the discovery of many new proton radioactivities.

The first nuclide found, in 1970, to decay by proton radioactivity was ${}^{53m}_{27}$Co (**Fig. 15**), where the *m* (metastate) denotes a (relatively) long-lived isomeric state. Because of its very high angular momentum of 19/2 and odd parity, gamma decay is highly forbidden. This mode of decay is essentially the same as that of β-delayed proton emission discussed below, except that the energy of the excited nuclear level is low, and angular momentum selection rules highly forbid gamma-ray decay so the state lives a relatively long time in comparison to those states populated in beta decay. It was produced in the laboratory by the compound nucleus reactions ${}^{16}_{8}\mathrm{O} + {}^{40}_{20}\mathrm{Ca} \rightarrow {}^{53m}_{27}\mathrm{Co} + p + 2n$ and ${}^{54}_{24}\mathrm{F} + p \rightarrow {}^{53m}_{27}\mathrm{Co} + 2n$. This 247-ms isomer exhibits two different decay modes: though it predominantly decays by positron (β^+) emission to a similar $19/2^-$ level in ${}^{53}_{26}$Fe, a 1.5% branch in its decay occurs via direct emission of a 1.59-MeV proton to the ${}^{52}_{26}$Fe ground state. The calculated half-life that ${}^{53m}_{27}$Co would possess if proton radioactivity were the only decay mode (its partial half-life for this decay branch) is the surprisingly long time of 17 s.

The first two cases of ground-state proton radioactivities were reported in 1982. These were ${}^{151}\mathrm{Lu} \rightarrow {}^{150}_{70}\mathrm{Yb} + p$ ($T_{1/2} = 85$ ms) and ${}^{147}_{69}\mathrm{Tm} \rightarrow {}^{146}_{68}\mathrm{Er} + p$ ($T_{1/2} = 0.56$ s) produced in the reactions ${}^{58}\mathrm{Ni} + {}^{96}\mathrm{Ru}$ and ${}^{58}\mathrm{Ni} + {}^{92}\mathrm{Mo}$. Others have now been found as heavy-ion experiments have reached isotopes still further off stability, for example, ${}^{113}_{55}$Cs. Nevertheless, the windows for the observation of direct proton decays are small and, therefore, such decays are very difficult to identify.

As one continues to remove neutrons from the nucleus of a given element with atomic number *Z*, one reaches a point, called the proton drip line, where for that *Z* and *N* a single proton becomes unbound and can drip off the nucleus. The Coulomb barrier (discussed earlier for alpha decay) and the centrifugal (angular momentum) barrier can hold the proton in for a limited time, so the nucleus undergoes radioactive decay with a certain half-life for the emission of a proton. Studies of proton radioactivities probe the limits of stability of proton-rich nuclei. Beyond the proton drip line, proton radioactivity is 100% of the decay. As one moves inside the line to more stable nuclei, proton radioactivity competes with positron decay, and, with the addition of more neutrons, positron beta decay becomes 100% of the radioactivity.

There are numerous examples of proton radioactivities with half-lives from hundreds of milliseconds

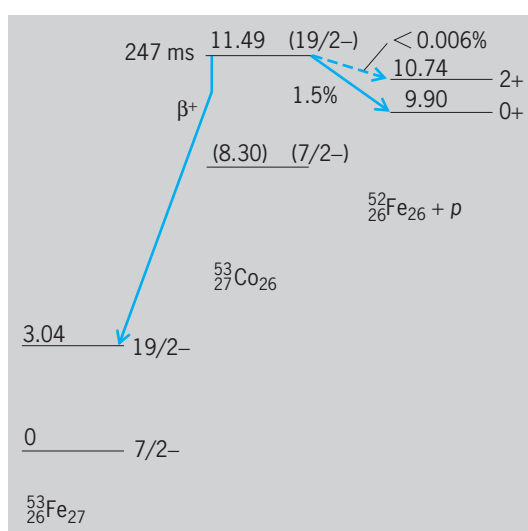

Fig. 15. Decay scheme of ${}^{53m}_{27}$Co. Numbers to left of levels represent energies in megaelectronvolts, relative to ground state of ${}^{53}_{26}\mathrm{Fe}_{27}$. Symbols to right of levels are spin and parity. (*After J. Cerny et al., Further results on the proton radioactivity of ${}^{53m}_{27}$Co, Nucl. Phys., A188:666–672, 1972*)

down to microseconds. The shortest-lived case is ^{145}Tm ($T_{1/2} = 3.5\ \mu$s), discovered in 1998. Most of the newer proton radioactivities have Z between 63 and 82, and most of these proton-radioactive nuclei are spherical. However, an example of proton radioactivity has been found for well-deformed ^{131}Eu, which decays by proton emission to ^{130}Sm, which is likewise well deformed. Fine structure in the proton spectrum is observed with decays to the ground and first excited states of ^{130}Sm to establish the large deformation of this nucleus. Proton fine-structure radioactivities have been observed both in the ^{146}Tm ground state and from an isomeric state ^{146m}Tm. (**Fig. 16**).

Proton radioactivities make it possible to probe the structure of nuclei at the limits and beyond the limits of stability. The half-lives for proton decays are strongly dependent on the energies of the proton and on the angular momentum carried away by the proton. Proton emitters are odd-Z nuclei because in such nuclei the energy to break apart a proton pair in the nucleus is not needed for proton emission to occur. The angular momentum carried away by the proton gives insight into which orbit the proton occupied prior to emission. Moreover, studies in which the proton emission has been used to tag the recoils of a heavy-ion reaction after mass separation have allowed the observation of the gamma rays emitted by the recoil nucleus prior to proton emission (Fig. 16). Thus, it is possible to study excited states in nuclei beyond the limits of proton stability. Both such data test theories of nuclear structure under new extreme conditions. In addition, the energies of the protons provide information about the mass differences of nuclei at the drip lines and so probe mass formulas out to new limits as well.

Two-proton radioactivities from ground states of nuclei are now observed. The first example was the two-proton decay of ^{45}Fe, $^{45}_{26}\mathrm{Fe} \rightarrow {}^{43}_{24}\mathrm{Cr} + 2p$. Two-proton decay of an excited state in ^{18}Ne to ^{16}O was observed earlier.

Neutron Radioactivity

In very neutron rich nuclei near the boundary line of nucleus stability, one may find nuclei with ground or excited states which are unstable to the emission of one or two neutrons. Here there is no Coulomb barrier to hold the neutron in the nucleus, but one can have a centrifugal barrier that may give rise to one- or even two-neutron radioactivity. These processes for ground states would be very near the limits where nuclei become totally unstable to the addition of a neutron, the neutron drip line, and very difficult to even make much less measure. However, there may be neutron-rich nuclei with high spin isomeric states where the high spin analogous to the one in ^{53m}Co gives rise to a large centrifugal barrier. Such isomeric states may undergo one- or two-neutron radioactivity.

Delayed Particle Emissions

Thirteen types of beta-delayed particle emissions have been observed (Table 1). Over 100 beta-delayed particle radioactivities are now known. Theoretically, the number of isotopes which can undergo beta-delayed particle emission could exceed 1000. Thus, this mode, which was observed in only a few cases prior to 1965, is among the important ones in nuclei very far from the stable ones in nature. Studies of these decays can provide insights into the nucleus which can be gained in no other way.

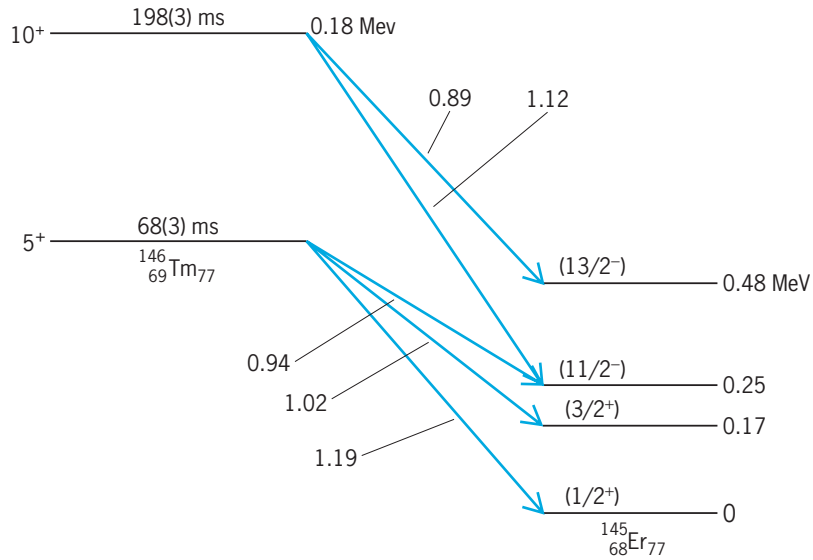

Fig. 16. Proposed decay scheme of the $^{146}_{69}$Tm ground state and of the isomeric state $^{146m}_{69}$Tm. Symbols at left of levels are spin and parity. Numbers above $^{146}_{69}$Tm levels are half-lives. Numbers to right of levels are energies relative to ground state in megaelectonvolts. Decays are labeled with their energies in magelectronvolts. (*After K. P. Rykaczewski et al., Proceedings, Conference on Nuclei at the Limits, Argonne, IL, July 2004*)

Beta-delayed alpha radioactivity. The β^- decay of ^{214}Bi to ^{214}Po leaves the nucleus in such a high-energy excited state that it can emit an alpha particle and go to ^{210}Pb as an alternative to gamma-ray decay to lower levels in ^{214}Po. This is a two-step process with beta decay the first step. After beta decay the nucleus is in such a highly excited state that it can emit either an alpha particle or gamma ray.

The β^- delayed alpha emission has been found relatively rarely, but in many cases beta (β^+, EC) delayed alpha emission has been discovered. In proton-rich nuclei far from stability, the conditions are more favorable for beta (β^+, EC) delayed alpha emission because of the excess of nuclear charge, and a number of such beta-delayed alpha emitters are now known.

Beta-delayed neutron radioactivity. In 1939, shortly after the discovery of nuclear fission, it was proposed that the delayed neutrons observed following fission were in fact beta-delayed neutrons. That is, after the nucleus fissioned, the beta decay of the neutron-rich fission fragments populated high-energy excited states that could promptly undergo dual decay, emitting either a gamma ray or neutron (Fig. 2). The processes of beta-delayed two- (Fig. 2) and three-neutron emission were discovered in 1979 and 1980 in the decay of ^{11}Li, and $\beta^- 2n$ to $4n$ decays were subsequently observed in other nuclei.

The process of beta-delayed neutron emission is essential for the control of nuclear fission reactors.

When neutrons absorbed by ^{235}U cause the ^{236}U nucleus formed to fission, many of the fission products undergo beta-delayed neutron emission. These neutrons are important in producing more fission events. In a nuclear reactor, the rate of neutron-induced fission must be controlled to prevent the fission reactions from running away and destroying the reactor. The rate of fission depends on the number of neutrons available. The numbers of neutrons can be controlled by moving in and out of the reactor control rods, which contain material with very high neutron absorption rates. Since many of the neutrons emitted in fission are delayed by beta-decay half-lives, these half-lives allow time for the control rods to be mechanically inserted and removed to control the rate of fission. *See* DELAYED NEUTRON.

Beta-delayed proton radioactivities. In addition to proton radioactivity, one can have beta-delayed proton and beta-delayed two-proton radioactivities which again ultimately result in emission of protons from the nucleus. These latter processes also occur in quite proton-rich nuclei with very high decay energies; however, they are complex two-step decay modes whose fundamental first step is beta decay.

Over 40 nuclei ranging from $^{9}_{6}C$ to $^{183}_{80}Hg$ have been identified to decay by the two-step mode of beta-delayed proton radioactivity. Typical is the decay of $^{33}_{18}Ar$ (**Fig. 17**), with a half-life of 173 ms; it was produced by the $^{32}_{16}S + ^{3}_{2}He \rightarrow ^{33}_{18}Ar + 2n$ reaction. This isotope decays by superallowed and allowed β^+ decay to a number of levels in its daughter nucleus $^{33}_{17}Cl$, which immediately (in less than 10^{-17} s) breaks up into $^{32}_{16}S$ and a proton. More than 30 proton groups arising from the decay of $^{33}_{18}Ar$ are observed, ranging in energy from 1 to approximately 6 MeV and varying in intensity over four orders of magnitude. Although it is normally very difficult to study many β-decay branches in the decay of a particular nuclide—because of the continuous nature of the energy spectrum of the emitted beta particles—it is possible to do so when investigating beta-delayed proton emitters. The observed proton group energies and intensities can be correlated with the levels fed in the preceding beta decay and their transition rates, thereby permitting sensitive tests via beta decay of nuclear wave functions arising from different models of the nucleus. β^+-delayed two-proton decay has also been discovered, as well as β^--delayed deuteron and triton decay.

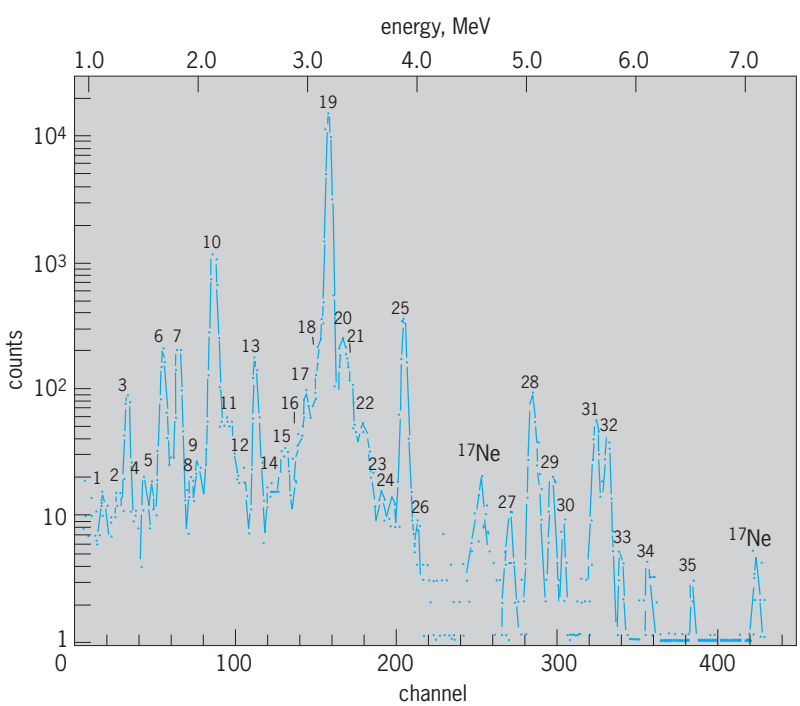

Fig. 17. Spectrum of β^+-delayed protons from the decay of $^{33}_{18}Ar$ as observed in a counter telescope; the proton laboratory energy is indicated at the top. Proton groups are numbered 1 through 35. (*After J. C. Hardy et al., Isospin purity and delayed-proton decay: ^{17}Ne and ^{33}Ar, Phys. Rev., C3:700–718, 1971*)

Beta-delayed spontaneous fission. There are also observed beta-decay processes where the excited nucleus following beta decay has a probability of undergoing spontaneous fission rather than gamma-ray decay. This is the same process as in spontaneous or isomeric spontaneous fission. The excitation energy of the nuclear level provides the extra energy to make fission possible. The nucleus splits into two nearly equal fragments plus some neutrons. This process is like isomeric spontaneous fission except that the lifetime of the nuclear level is so short that the level would not normally be called an isomer.

Joseph H. Hamilton

Bibliography. D. A. Bromley (ed.), *Nuclei Far from Stability*, vol. 8 of *Treatise on Heavy-Ion Science*, Plenum, 1989; J. M. Eisenberg and W. Greiner, *Nuclear Theory*, 3 vols., 3d ed., Elsevier, 1987–1988; R. B. Firestone and V. S. Shirley (eds.), *Table of Isotopes*, 8th ed., Wiley, 1996; E. J. Konopinski, *The Theory of Beta Radioactivity*, Clarendon Press, 1966; D. N. Poenaru (ed.), *Nuclear Decay Modes*, Institute of Physics, 1996; A. Sandulescu and W. Greiner, New radioactivities, *Sci. Amer.*, 262(3):58, March 1990; F. Yang and J. H. Hamilton, *Modern Atomic and Nuclear Physics*, McGraw-Hill, 1996.

Radioactivity and radiation applications

The field in which the subatomic fragments emitted in radioactive decay (alpha particles, beta particles, gamma rays) or produced by high-voltage accelerators (electrons, protons, x-rays) are applied to the problems of science, engineering, industry, and medicine. The techniques are extraordinarily versatile and sensitive and are basically inexpensive. A disadvantage that limits the range and extent of these applications is the health hazard that may be involved. *See* RADIOACTIVITY; RADIOISOTOPE.

Tracer applications. These techniques are based on two principles. First is the chemical similarity of radioactive atoms and other atoms of the same element. Periodically a few of the radioactive atoms decay, emitting some penetrating subatomic fragments that can be detected one by one, usually through their ability to cause ionization. Thus the movement of a particular element can be followed through various chemical, physical, and biological steps. The second principle involves the characteristic half-life and nature of the emitted fragments. This

makes a radioactive species unique and thereby detectable above a background of radioactive emitters associated with elements.

The use of tracers in pure and applied research, process control, and medical diagnosis is one of the triumphs of modern science. Examples include the following: (1) A small amount of radioactive additive injected into an opening in a piping system can be used to measure the flow rate throughout the system and to detect the leaks in buried portions of the pipe. (2) In the metals industry, wear studies have been successfully performed using radioactively tagged tool bits and piston rings; radioactivity buildup in the lubricant, on the metal being shaped, or on the wall of the cylinder makes possible a simple and rapid measure of the durability of the tagged component. (3) Tracers added to the bloodstream have been used to monitor the movement of biochemical components throughout the body. (4) Addition of short-lived tracers to industrial waste streams have been used to determine their flow patterns. (5) A good way to determine when mixing is complete in a 50,000-bbl (8000 m^3) tank at an oil refinery is to add a small amount of radioactivity and find out when the activity level in random samples is constant. (6) Important contributions in understanding the mechanism of combustion have been obtained by means of tracer studies. (7) Diffusion in a high-viscosity medium or in a crystal is a slow process whose measurement is tedious and difficult by any technique except the tracer method. For discussions of radioisotope techniques relating to tracer methodology *see* ACTIVATION ANALYSIS; ISOTOPE DILUTION TECHNIQUES; NONDESTRUCTIVE EVALUATION; RADIOACTIVE TRACER; RADIOECOLOGY; RADIOISOTOPE (BIOLOGY).

Penetration and scattering applications. These applications arise from the fact that subatomic fragments can penetrate a thick section of a material, and yet a small fraction of the incident particles can be backscattered by a relatively thin section.

The oldest application of the penetrating properties of energetic ionizing photons is radiography. A more recent extension of this technique is autoradiography. *See* AUTORADIOGRAPHY; RADIOGRAPHY.

Since World War II the penetration and scattering properties of beta particles and gamma rays have been applied in industry in the form of thickness gages, particularly for process control in the continuous production of paper and metal films. In the usual application of a penetration-thickness gage, a beta source is placed on one side of the continuous sheet, and a beta detector is mounted on the other side. The basic absorption law is given by the notation below. Because the number of beta particles

$$N \propto e^{-x/\bar{x}}$$

penetrating the sheet N is an exponential function of the thickness x, small changes in thickness can be detected when a beta source with optimum penetration properties is used. The optimum condition is one in which the thickness of the sheet equals the apparent mean free path of the beta particle, $\bar{x}$, which varies directly with the energy of the beta particle.

A related application is the use of the amount of beta backscatter to measure the thickness of a thin coating on a thick backing of different composition. A backscatter gage using thallium-204 beta particles can easily monitor changes of 4×10^{-5} in. (10^{-4} cm) in the thickness of a tin coating on a steel backing.

An important application of the scattering and penetration of neutrons and gammas is in the field of well logging. *See* WELL LOGGING.

Initiation of chemical reactions. The absorption of small amounts of energy from ionizing particles and ionizing photons have chemical effects that have been the basis of several practical applications. A dose of 1 rad (0.01 gray) is equivalent to the deposition in each gram of the absorber of 100 erg (1 erg = 10^{-7} joule = 7.4×10^{-8} ft-lbf) from a source of radiation. An absorbed whole body dose of 500 rads (5 gray) may be sufficient to cause death in a human, but the same amount of energy in the form of heat would do no more than raise the body temperature by 0.002°F (0.001°C).

The oldest application of this principle is radiation therapy. For example, in cancer therapy the local affected areas are irradiated by external beams of gammas from cobalt-60 or of radiation from accelerators. Radioactive sources have also been administered internally to induce beneficial biochemical reactions in patients with various ailments. Injection of gold-198 in colloidal form into tumor tissue is a common therapeutic procedure. An iodine-131 "cocktail" is useful in the treatment of an overactive thyroid gland. The intravenous injection of phosphorus-32 is a standard method of treating leukemia and other blood diseases. *See* ISOTOPIC IRRADIATION; RADIOLOGY.

A related area is the radiation sterilization of biomedical supplies. There are about 50 installations throughout the world employing ionizing radiation (usually in the form of cobalt-60 gamma rays) to give sterilizing doses (2.5–4 megarads or 25–40 kilograys) to disposable medical supplies. The advantages to this method of biochemical destruction of microscopic life are that (1) unlike steam sterilization, it can be performed at low temperatures on plastics and other thermally unstable materials, and (2) unlike germicidal gases, ionizing radiation can reach every point in the treated product. Radiation-sterilized objects are not radioactive.

The radiation preservation of food is an area of considerable promise. Small doses can inhibit sprouting in potatoes (10,000 rads or 100 Gy), kill insects in wheat (50,000 rads or 500 Gy), and sterilize pork products (3×10^6 rads or 30,000 Gy) but practical applications have been sharply limited due to a cautious role by regulatory authorities in approving such procedures. There is also industrial concern over the possibility of consumer rejection of foods preserved by treatment with ionizing radiation. Despite the short-term problems, there are grounds for optimism in the long run. The energy requirements for radiation preservation of food are far less than those of a preservation technology based on steam sterilization and refrigeration. Rising energy costs as well as increasing laboratory and marketing experience

with irradiated foods should lead to a more favorable attitude.

Although ionizing radiation has not yet proved to be an attractive alternative to nonradiation methods of chemical synthesis, it has assumed a significant role in the modification of plastics. Electron-beam irradiation of wire insulated by polyethylene or poly(vinyl chloride) formulations is an important means of upgrading the properties of the insulation at higher temperatures. The ionizing radiation induces chemical reactions that lead to the cross-linking of polymer chains to form a network called a gel, which, in effect, does not melt and does not dissolve in solvents. Electron-beam irradiations of polyethylene packaging is used to manufacture heat-shrinkable form-fitting pouches for turkeys, hams, electronic parts, and so on. The principle behind the application is as follows. Polyethylene is highly cross-linked by irradiation; when heated above its normal melting point, it behaves like a rubber. Thus a pouch made of cross-linked polyethylene, which is then heated and stretched, retains the stretched configuration when cooled. When a ham is placed in this pouch and the entire package is heated, the now-rubbery pouch shrinks to form a skin-tight covering that retains its shape after cooling. The same principle has been successfully applied to produce heat-shrinkable electrical-insulation components that have widespread industrial use.

Another important application of electron irradiation is the curing of paint applied to plastic or metal backings. Some automobile dashboards are coated with paint that is "cold-cured" by exposure to accelerated electrons. A variety of automobile parts are treated in a similar manner. The process replaces one involving long-heating ovens with high thermal energy losses and with the liberation to the atmosphere of volatile paint components.

Electron machines are used for a radiation process that produces an estimated 2×10^9 ft^2 (2×10^8 m^2) per year of a permanent-press polyester-cotton blend with excellent soil-release properties. The process involves irradiation of the fabric, which has been soaked with a reagent that ultimately provides the permanent-press characteristics. The irradiation provides reactive sites on the fabric to which the reagent molecules attach themselves. The fabric is then exposed to the soil-release agent, which then hooks on to some long-lived reactive sites that persist after the irradiation.

The role of ionizing radiation in sterilization and chemically oriented applications is significant, but it is not a major factor in modern technology. A major motivating force behind the research leading to radiation processes is the potential availability of large inventories of fission products. Nevertheless, the only successful industrial irradiators involve either cobalt-60 or electron accelerators. The fission product, cesium-137, has some potential future value as a gamma source. *See* RADIATION CHEMISTRY.

Radioactive power sources. Kinetic energy of emissions in radioactive decay can be converted to useful forms of light, heat, and electricity.

A mixture of a phosphor and radionuclide produces a luminous paint that has been used for watch and instrument dials. Such a light source in combination with a photovoltaic cell can provide electricity at a power level up to a microwatt. *See* LUMINOUS PAINT.

The heat produced by betas from strontium-90 on their absorption in a thermoelectric converter has been used to power buoys and remote arctic radiotransmitters. Similar converters of heat to electricity using promethium-147 and plutonium-238 are the microwatt power sources for pacemakers that regulate heart action. Larger versions of the plutonium-238 battery have been proposed for use in an artificial heart.

Although direct conversion to electricity of the energy of alphas and betas has been accomplished, the most efficient and practical converters are based on the conversion to heat in an intermediate step. *See* NUCLEAR BATTERY. Joseph Silverman

Bibliography. G. Foldiak (ed.), *Industrial Applications of Radioisotopes*, 1986; M. F. L'Annunziata, *Radionuclide Tracers*, 1987; E. Reichmanis, C. W. Frank, and J. H. O'Donnell (eds.), *Irradiation of Polymeric Materials*, 1993; J. W. Spinks and R. J. Woods, *An Introduction to Radiation Chemistry*, 3d ed., 1990; W. M. Urbain, *Food Irradiation*, 1986; P. Von der Hardt and H. Rottger (eds.), *Irradiation Technology*, 1983; R. R. Wolfe, *Radioactive and Stable Isotope Tracers in Biomedicine*, 1992.

Radioactivity standards

Calibrated standard sources of radioactive substances used to determine, by comparison, the strength or activity of samples of the same substances in terms of the number of radioactive atoms they contain or in terms of some figure proportional to this number. The calibration of the standard source in terms of the number of radioactive atoms is usually an elaborate procedure but need only be carried out once, and the calibration may be made at a standardizing laboratory, such as the National Institute of Standards and Technology, having special equipment for the work. Comparison between a sample and the standard is usually made by finding the ratio of the response of an ionization chamber, or other detector of radiation, to the radiation from a sample and from the standard. In each case the intensity of the radiation, and therefore the response of the detector under identical conditions, is proportional to the number of radioactive atoms in the source, because this number is also proportional to the activity or disintegration rate of a source. *See* HALF-LIFE; RADIOACTIVITY.

The situations in which standards of radioactivity can be used are limited by a number of factors. For example, the radioactive species involved must have a half-life long enough that the standards will continue to have sufficient activity for the period of time in which they are to be used. Thus standards of radium, in which the half-life is about 1620 years, have an almost permanent value. Standards of radioisotopes

with very much shorter half-lives, of the order of a day, require preparation and calibration immediately prior to their use. Leon F. Curtiss; Karl Z. Morgan

Bibliography. *American Institute of Physics Handbook*, 3d ed., 1972; H. Etherington (ed.), *Nuclear Engineering Handbook*, 1958; W. B. Mann, The preparation and maintenance of standards of radioactivity, *Int. J. Appl. Radiat. Isotop.*, 1:3–23, 1956; National Council on Radiation Protection and Measurement, *Handbook of Radioactivity Measurements Procedures*, NCRP Rep. 58, 2d ed., 1985; J. Shapiro, *Radiation Protection: A Guide for Scientists and Physicians*, 3d ed., 1990.

Radiocarbon dating

A method of obtaining age estimates on organic materials which has been used to date samples as old as 75,000 years. The method was developed immediately following World War II by Willard F. Libby and coworkers, and has provided age determinations in archeology, geology, geophysics, and other branches of science.

Radiocarbon (^{14}C) determinations can be obtained on wood; charcoal; marine and fresh-water shell; bone and antler; peat and organic-bearing sediments; carbonate deposits such as tufa, caliche, and marl; and dissolved carbon dioxide (CO_2) and carbonates in ocean, lake, and ground-water sources. Each sample type has specific problems associated with its use for dating purposes, including contamination and special environmental effects. While the impact of ^{14}C dating has been most profound in archeological research and particularly in prehistoric studies, extremely significant contributions have also been made in hydrology and oceanography. In addition, beginning in the 1950s the testing of thermonuclear weapons injected large amounts of artificial ^{14}C ("bomb ^{14}C") into the atmosphere, permitting it to be used as a geochemical tracer.

Basis of the Method

Carbon (C) has three naturally occurring isotopes. Both ^{12}C and ^{13}C are stable, but ^{14}C decays by very weak beta decay (electron emission) to nitrogen-14 (^{14}N) with a half-life of approximately 5700 years. Naturally occurring ^{14}C is produced as a secondary effect of cosmic-ray bombardment of the upper atmosphere (**Fig. 1**). As $^{14}CO_2$, it is distributed on a worldwide basis into various atmospheric, biospheric, and hydrospheric reservoirs on a time scale much shorter than its half-life. Metabolic processes in living organisms and relatively rapid turnover of carbonates in surface ocean waters maintain ^{14}C levels at approximately constant levels in most of the biosphere. The natural ^{14}C activity in the geologically recent contemporary "prebomb" biosphere was approximately 13.5 disintegrations per minute per gram of carbon. *See* COSMOGENIC NUCLIDE; ISOTOPE.

To the degree that ^{14}C production has proceeded long enough without significant variation to produce an equilibrium or steady-state condition, ^{14}C levels

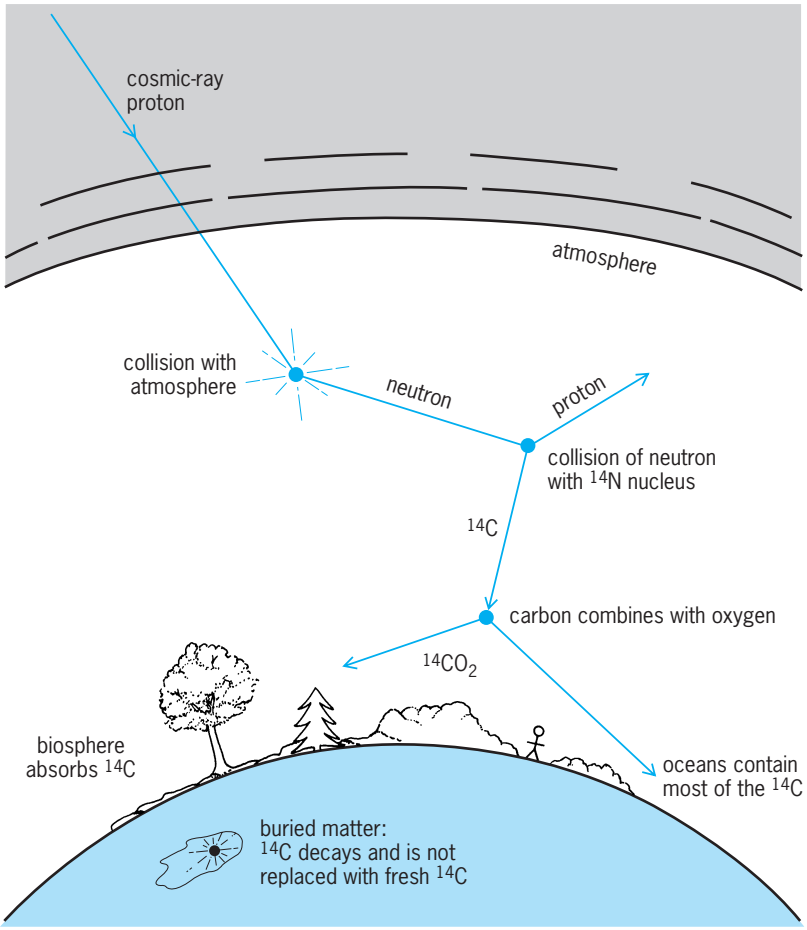

Fig. 1. Generation, distribution, and decay of ^{14}C.

observed in contemporary materials may be used to characterize the original ^{14}C activity in the corresponding carbon reservoirs. Once a sample has been removed from exchange with its reservoir, as at the death of an organism, the amount of ^{14}C begins to decrease as a function of its half-life. A ^{14}C age determination is based on a measurement of the residual ^{14}C activity in a sample compared to the activity of a sample of assumed zero age (a contemporary standard) from the same reservoir. The relationship between the ^{14}C age and the ^{14}C activity of a sample is given by the equation below, where t is radiocarbon years B.P. (before the present), λ is the decay constant of ^{14}C (related to the half-life $t_{1/2}$ by the expression $t_{1/2} = 0.693/\lambda$), A_o is the activity of the contemporary standards, and A_s is the activity of the unknown age samples. Conventional radiocarbon dates are calculated by using this formula, an internationally agreed half-life value of 5568 ± 30 years, and a specific contemporary standard. Most laboratories define the contemporary standard value by using one of the standards prepared by the U.S. National Bureau of Standards [NBS; now known as the U.S. National Institute of Standards and Technology (NIST)], or a standard with a known relationship to the NBS/NIST oxalic acid preparations.

$$t = \frac{1}{\lambda} \ln \frac{A_o}{A_s}$$

Measurement of Radiocarbon

The naturally occurring isotopes of carbon occur in the proportion of approximately 98.9% ^{12}C, 1.1% ^{13}C, and 10^{-10}% ^{14}C. The extremely small amount of radiocarbon in natural materials was one reason why ^{14}C was one of the isotopes which had been produced artificially in the laboratory before being detected in natural concentrations. The routine development of the radiocarbon method was made possible by the development by Libby of a practical method of low-level counting. To detect the very weak beta-decay characteristic of ^{14}C, a means had to be devised to introduce the sample directly into the sensitive volume of a detector. In all of Libby's early work, the sample was converted to solid carbon (amorphous elemental carbon) and deposited on a sleeve which fitted inside a screen-wall type of Geiger counter. The counting rate of an unshielded screen-wall counter was on the order of 500 counts per minute.

Since the activity from ^{14}C decay of a modern sample was expected to be about six or seven counts per minute, the total background counting rate had to be radically reduced. This was accomplished initially by placing the instrument in an iron shield with 20-cm (8-in.) walls. This reduced the activity in the detector to 120 counts per minute, still unacceptably high. The final reduction was made possible by enclosing the sample counter in a ring of smaller Geiger counters. The sample counter and the outer guard ring were connected together electronically so that any pulse from any of the outer Geiger tubes would inactivate the sample counter for about 10^{-3} s. This anticoincidence system reduced the background in the center detector to about five counts per minute.

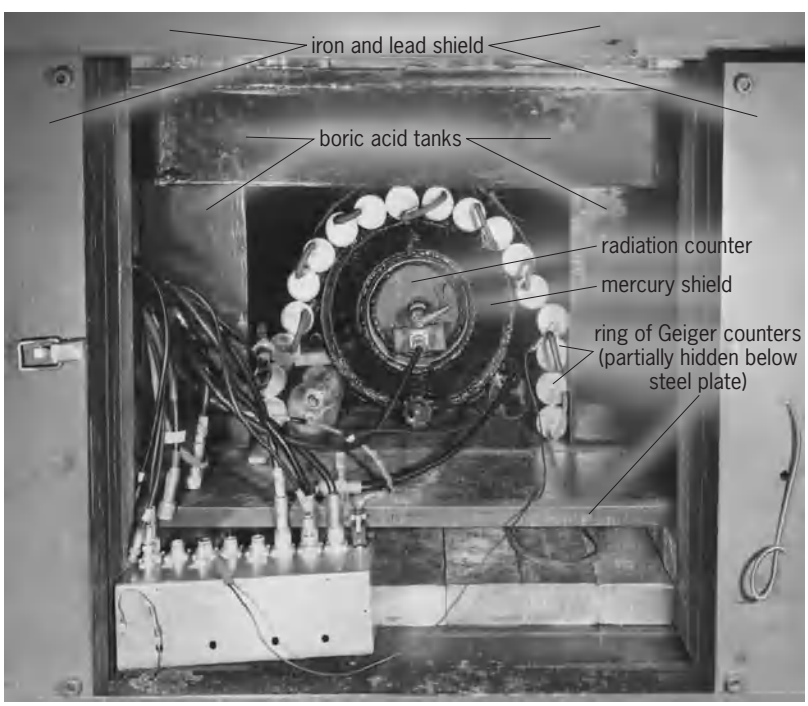

Fig. 2. Large-volume proportional counter, which is used for carbon-14 measurements. The outer shield is closed by rolling doors. The sample is introduced into the radiation counter in the form of carbon dioxide gas. (*Geochemical Laboratory*, *Lamont-Doherty Geological Observatory*, *Columbia University*)

With this system, the maximum age that could be measured was about 23,000 years and required the use of 10–12 g (0.35–0.42 oz) of carbon from sample materials. Because of self-absorption of the weak betas in the sample, the efficiency of the detector was only about 5%. Because of this and the susceptibility of the carbon black to contamination from airborne radioactive fallout, the solid carbon technique was replaced by either gas counters or liquid scintillation systems. *See* LOW-LEVEL COUNTING; RADIOACTIVITY.

Gas counters. In the early 1950s, both proportional (**Fig. 2**) and Geiger gas counters were employed in ^{14}C work, using carbon dioxide, carbon disulfide, acetylene, methane, or ethane as counting gases. As in the case of the solid carbon system, the center counter containing the sample was surrounded by individual Geiger tubes or an annular or continuous ring guard, all housed within an iron or lead shield assembly. Efforts to reduce the background values in gas detectors have resulted in various types of experimental arrangements, including the location of counters in underground vaults. In such underground facilities, the contribution of the meson flux, the major contributor to the background rate, can be significantly reduced. Because of the 90–95% efficiency in most gas detector systems, the typical maximum age limits were extended to 40,000–60,000 years, depending on the experimental configuration, including the volume of the detectors and the level of the background count rates in specific detectors. Isotopic enrichment of sample gases permits the maximum age attainable to be extended several additional half-lives. In general, sample-size requirements with gas detectors were reduced from that required with the solid carbon method—special systems being designed to permit the measurement of a sample with as little as 0.1 g (3.5×10^{-3} oz) of carbon. *See* GEIGER-MÜLLER COUNTER; IONIZATION CHAMBER; MESON.

Liquid scintillation systems. Current liquid scintillation systems involve the conversion of samples to benzene. The addition of a scintillator chemical allows beta-decay events to be monitored by photomultiplier tubes. Earlier liquid scintillation systems used for ^{14}C measurements generally required larger samples. However, currently the amounts required are comparable to gas counting systems. Continuing developments in liquid scintillation technology for low-level measurement have provided the ability to monitor counter performance in much greater detail than typically is possible in gas counting, and have also resulted in reduced background values. In these systems, the maximum ages that can be measured can be extended beyond that possible with typical gas systems.

Direct detection. Both of the conventional decay counting methods share a common problem in that they employ an inherently inefficient means of monitoring ^{14}C concentrations in samples. In 1 g of modern carbon, for every decay per minute there are about 4×10^9 atoms of ^{14}C. Counting methods have

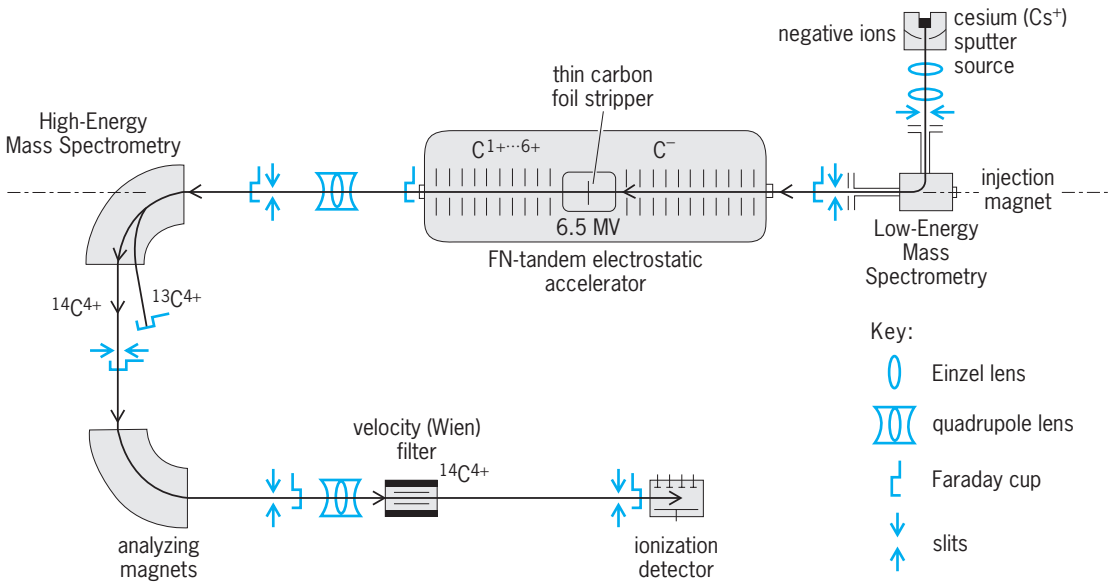

Fig. 3. Accelerator mass spectrometry system for the direct detection of ^{14}C atoms. (*Center for Mass Spectrometry, Lawrence Livermore National Laboratory*)

been developed which employ particle accelerators as very sensitive mass spectrometers, thus counting the ^{14}C atoms directly.

It has long been recognized that if the ^{14}C atoms could be detected directly, rather than by waiting for their decay, smaller samples could be used for dating and older dates could be measured. A simple hypothetical example to illustrate this point is a sample containing only one atom of ^{14}C. To measure the age (that is, the abundance of ^{14}C), the sample can be placed into a mass spectrometer and that atom counted, or the sample can be placed into a Geiger counter and counted, requiring a wait on the average of 8000 years (the mean life of ^{14}C) for the decay. In practice, neither the atoms nor the decays can be counted with 100% efficiency, but the huge advantage for atom counting remains.

Until 1977, attempts at direct atom counting for ^{14}C and other natural radioisotopes failed, because of the extremely low concentration of ^{14}C. Ordinary mass spectrometers could not see the tiny ^{14}C signal in the background of other atoms and molecules in the sample. Even trace amounts of nitrogen would swamp the ^{14}C signal, since ^{14}N forms ions with nearly identical mass and identical charge to that of the ^{14}C atoms.

The technique for detecting ^{14}C atoms [and other natural radioisotopes such as tritium, beryllium-10 (^{10}Be), and chlorine-36 (^{36}Cl)] is a combination of mass spectrometry and accelerator technology, called accelerator mass spectrometry (AMS). The approach was first demonstrated using a cyclotron. This accelerator, which sent particles along a spiral trajectory, was used as an ultrasensitive mass spectrometer to distinguish ionized carbon isotopes by their charge-to-mass ratio. Detecting ^{14}C by this means was possible, but consistent results proved difficult to achieve despite years of effort. *See* MASS SPECTROMETRY.

Another type of AMS technology uses a tandem electrostatic accelerator (**Fig. 3**). The device employs two stages. First, a negative ion beam is accelerated and passed through a stripper, which removes the electrons, converting the beam to positive ions. Then the particles are further accelerated. The stripping process breaks up molecules of mass 14, which would otherwise interfere with the detection of ^{14}C, and the negative ion beam eliminates ^{14}N since there are no known stable negative nitrogen ions. Almost all AMS ^{14}C applications currently use tandem accelerators to accomplish routine measurements. *See* PARTICLE ACCELERATOR.

The advent of AMS technology in the late 1970s brought about an enormous boost in detection efficiency that promised three important advantages for ^{14}C dating. First the amount of carbon required was reduced from grams to milligrams. Second, counting times were reduced from days, weeks, or even months to minutes. Finally, it was initially thought that the detection sensitivity would increase so that the maximum age datable with ^{14}C might be extended to 100,000 years. However, sensitivity is limited by very small amounts of contamination introduced during sample preparation. Much of this contamination stems from the requirement in most laboratories that samples be converted to graphitic carbon for measurement. In routine operation, current AMS technology can measure between 40,000 and 50,000 years, and rarely, 60,000 years.

The ability to use milligram (rather than gram) samples is very important for dating. Certain irreplaceable objects (for example, parchments, cloth, and chips of wood) would have had to be destroyed in order to extract the gram of carbon required for a date; for the accelerator method, only a small piece of the artifact is required. In addition, the ability to date by using only milligrams of carbon allows careful selection of the sample used. Any part of the object

which may have been contaminated by modern carbon can be ignored; small seeds trapped in the object, or even specific amino acid compounds which are less likely to come from modern carbon contamination, can be selected.

Despite the sensitivity of the accelerator technique, decay dating will probably continue to be used for ^{14}C dating when gram amounts of carbon are available. However, the technology of atomic mass spectrometry continues to be developed and will increasingly be used for routine ^{14}C analysis. The accuracy of ^{14}C values based on atomic mass spectrometry has become essentially comparable to that obtained with decay counting.

Accuracy of Radiocarbon Determinations

A measurement of the ^{14}C content of an organic sample will provide an accurate determination of the sample's age if it is assumed that (1) the production of ^{14}C by cosmic rays has remained essentially constant long enough to establish a steady state in the $^{14}C/^{12}C$ ratio in the atmosphere, (2) there has been a complete and rapid mixing of ^{14}C throughout the various carbon reservoirs, (3) the carbon isotope ratio in the sample has not been altered except by ^{14}C decay, and (4) the total amount of carbon in any reservoir has not been altered. In addition, the half-life of ^{14}C must be known with sufficient accuracy, and it must be possible to measure natural levels of ^{14}C to appropriate levels of accuracy and precision. Studies have shown that the primary assumptions on which the method rests have been violated both systematically and to varying degrees for particular sample types. Several approaches have been developed to provide calibration and corrections of conventional ^{14}C values. The basis of the calibration and correction procedures will be discussed in the context of a brief review of the assumptions of the method.

Constancy in radiocarbon production rates. Carbon-14 determinations on known-age samples have revealed systematic discrepancies in the ^{14}C time scale. The first hint of such anomalies came from early ^{14}C measurements on Egyptian archeological materials. Samples which, on historical grounds, should have dated to the early part of the third millennium B.C. yielded ^{14}C values some 700–800 years too young. Carbon-14 determinations carried out on dendrochronologically dated wood, periglacial varves, and lake sediments confirmed the fact that there have been systematic variations in ^{14}C values over time. The data which first contributed most directly to the study of these anomalies was the dendrochronological time scale provided by the bristlecone pine (*Pinus longaeva*), from the White Mountains of east-central California, developed by C. W. Fergusson. His data provide an unbroken tree-ring series back to almost 6700 B.C. An independently developed bristlecone tree-ring chronology from a different locality in the southern portion of the White Mountains, developed by V. C. La Marche and T. P. Harlan, supports the accuracy of the Fergusson chronology at least as far back as about 3500 B.C. Carbon-14 determinations on bristlecone pine as well as tree-ring-dated samples from the sequoia (*Sequoia gigantea*) and European oaks (*Quercus* spp.) have been undertaken by a number of laboratories, and ^{14}C determinations on these samples provide data over the last 10,000 years. *See* DENDROCHRONOLOGY; VARVE.

Main trend. Upon examination of the data in Fig. 4 and other similar plots, it becomes apparent that radiocarbon years and calendar years are not necessarily equivalent. If such had been the case, all of the data points plotted on **Fig. 4** would lie along the horizontal 0 line. In fact, some of the points lie above the line, indicating that ^{14}C values in these periods are too old. Conversely, those below the 0 line are too young when compared to the tree-ring data. This plot indicates that there are two major components to the deviations. The first is a general main-trend secular variation phenomenon (the curved line) exhibiting during the Holocene, a sine-wave function with an apparent period of about 8500–9000 years, with a maximum deviation of about 800 years, approximately 8000 years ago.

The characteristic of the secular variation anomalies in the period before about 10,000 years ago cannot, at present, be documented by tree-ring/^{14}C data. It has been argued, however, that the major part of the effect may be estimated by examining the record of the Earth's dipole geomagnetic field over time. Variations in the intensity of the dipole field modulate the cosmic-ray flux in the vicinity of the Earth. An increase in the field strength, for example, diverts more of the cosmic-ray particles away from the Earth, resulting in a decrease in the production of ^{14}C.

Geophysicists have collected data which document changes in the intensity of the Earth's dipole field for the last few hundred thousand years. Because of the apparent inverse relationship between the intensity of the field and the ^{14}C production rate, it would, in theory, be possible to extrapolate the maximum and minimum secular variation deviations back to the limit of the ^{14}C method. Unfortunately, such data are not yet as precise as might be wished.

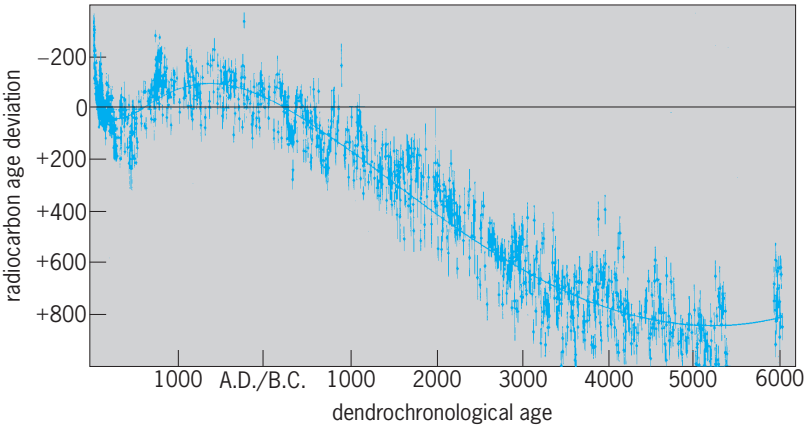

Fig. 4. Secular variation/major trend; relationship between radiocarbon and dendrochronological age of wood samples. (*After J. Klein et al., Calibration of radiocarbon dates, Radiocarbon, 24(2):103–150, 1982*)

However, comparisons of ^{14}C ages with uranium-thorium (U-Th) ages obtained on cores from coral deposits support conclusions based on the ^{14}C/tree-ring data up to the limit of the current dendrochronological data. Uranium-thorium values can be used to continue to examine the ^{14}C deviations over the last 30,000 years. Such data indicate that radiocarbon ages earlier than 10,000 years B.P. continue to be systematically younger than U-Th ages, with a maximum difference of about 3500 years approximately 20,000 years ago. *See* PALEOMAGNETISM; ROCK MAGNETISM.

De Vries effect. In addition to the long-term secular variation phenomenon, the bristlecone pine data have revealed the presence of high-frequency components to the variation in ^{14}C activity. These short-term oscillations or wiggles have sometimes been called the De Vries effect after the pioneering Dutch researcher, Hessel de Vries, who was one of the first to call attention to the existence of systematic anomalies in ^{14}C values. Although almost all investigators concerned with the issue have agreed that the tree-ring/^{14}C data definitely reveal the presence of a number of short-term perturbations, the frequency and magnitude of earlier episodes during the Pleistocene, (that is, > 10,000 years B.P.) have not been resolved. Likewise, there are uncertainties as to the causes of the De Vries effect, although variation in solar activity (heliomagnetic effects) has been seen as an important factor.

Calibration of radiocarbon dates. The existence of main trend and De Vries deviations has important implications in the interpretation of ^{14}C determinations. The long-term variations result in the necessity to calibrate conventional ^{14}C dates in terms of the known variation between radiocarbon time and real or calendar time as documented by the dendrochronological/^{14}C values. The magnitude of the calibration varies depending on from what time period a sample is derived. For the period back to about 1000 B.C., corrections required by virtue of the secular-variation deviations do not exceed about 150 years. Prior to 1000 B.C., the magnitude of the correction steadily increases. By using data such as those presented in Fig. 4, various approaches have been developed to "calibrate" radiocarbon values. In this context, calibration involves taking a ^{14}C age value expressed as a conventional radiocarbon date and adding or subtracting the number of years required to bring the conventional age into conformity with the ^{14}C determinations on known-age tree-ring-dated samples.

The documentation of the presence of the De Vries or short-term anomalies has introduced a second problem in the calibration of ^{14}C values. Periods of rapid change in the ^{14}C content of the atmosphere result in situations where a single ^{14}C value may reflect two or more points in real time. The characteristics of the short-term anomalies are illustrated in **Fig. 5**. During periods of particularly rapid change in ^{14}C activity, it is usually not possible to use ^{14}C data to document temporal intervals in units of less than a few hundred years. Thus the dendrochronologically based calibration data can be used to identify the general degree of deviation of ^{14}C values from real time and also the degree of maximum precision which is possible for specific temporal intervals.

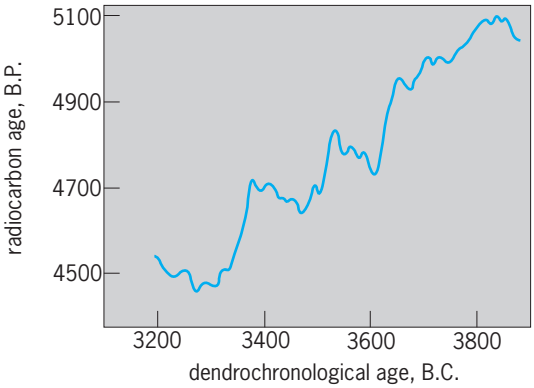

Fig. 5. Secular variation/De Vries effect. Example of short-term variations in ^{14}C activity; detail of De Vries effects using high-precision measurements for period of approximately 4500–5100 ^{14}C years B.P. (*After A. F. M. de Jong and W. G. Mook, Medium-term atmospheric ^{14}C variations, Radiocarbon, 22(2):267–272, 1980*)

The impact of the De Vries effects on the precision of ^{14}C values, as applied to archeological and historical problems, must be considered. The need to take into account these shorter-term variations have, for example, been demonstrated in the evaluation of ^{14}C values on twelfth- and fourteenth-century European medieval archeological materials. *See* ARCHEOLOGICAL CHRONOLOGY.

Lack of geographical and altitude variations. Concern has been expressed as to whether variation in the ^{14}C content of wood samples taken from a small number of localities in the Northern Hemisphere can be used to document worldwide secular variation effects. This question has been specifically answered as a result of studies of ^{14}C concentrations in tree rings from Patagonia, Canada, and Europe. The maximum deviations noted in contemporaneous woods were between those grown in the Southern and Northern hemispheres. However, even in this case, the variation did not exceed 0.5% or about the equivalent of 40 years.

Studies have also shown the lack of any significant altitude effect on ^{14}C concentrations in wood. A concern had been expressed that solar protons of appropriate energies would interact with nitrogen to form ^{14}C directly in wood samples growing at high elevations. The projected effect would be to inflate the ^{14}C content of a sample so that the ^{14}C age would appear to be significantly younger than its true age. However, calculations indicate that the maximum effect that could be obtained, assuming the most advantageous parameters, would not exceed 40 years. Other evidence suggests the actual effect is much less. The bristlecone pine wood of the White Mountains grows at about the 3350-m (11,000-ft) level. That there is no measurable ^{14}C produced as a function of altitude is indicated by the essential agreement in age between the high-altitude

bristlecone and low-altitude European and American sequoia samples of the same dendrochronological age.

Variability in radiocarbon distribution. An important feature of the ^{14}C method is its potential to provide directly comparable age determinations on a worldwide basis for a wide variety of organic samples. For this potential to be realized, ^{14}C, following its production, has to be mixed rapidly and completely throughout all of the carbon-containing reservoirs on a time scale not exceeding a few tens of years. To the degree that such conditions prevail, the contemporary ^{14}C content of all organic samples will be essentially identical. It was quickly determined that such is not the case. The initial ^{14}C content of samples could be significantly affected as a result of environmental conditions. A classic illustration of the problem was the discovery of living organisms from a fresh-water lake exhibiting ^{14}C ages of approximately 2000 years. In this case a large percentage of the carbon used by the organisms was derived from dissolved CO_2 from the limestone bed of the lake. The fictional age of the modern samples had been produced as a result of the dilution of contemporary ^{14}C activity by "dead" carbon (that is, containing no ^{14}C) from the limestone. Another example is provided from trees growing in active volcanic areas. The CO_2 emitted during volcanic discharges is characteristically depleted of its ^{14}C. Living trees exhibiting apparent ages as much as 1000 years from such environments have been reported.

One effect of the recognition that the geochemical environment of a sample can affect its initial ^{14}C concentration has been to cast doubt on the reliability of particular types of samples. The use of shells in ^{14}C studies has been affected, since a tradition arose that their use should be discouraged. Terrestrial shells (gastropods) from most fresh-water environments generally merit this negative evaluation, since they typically take up carbonate which is not in equilibrium with atmospheric ^{14}C. The reputation of marine shells was adversely affected primarily as a result of early experiences with shells taken from several archeological sites along the Peruvian coast. Marine shell samples were found to have ^{14}C values that exhibited an apparent age as much as 900 years greater than that of charcoal samples assumed to have been deposited contemporaneously. It was therefore assumed that marine shells would consistently yield anomalous values.

Subsequent studies showed that marine shells can yield generally acceptable values if the conventional ^{14}C values can be corrected for upwelling effects. By examining ^{14}C concentrations in shells collected alive in the period before nuclear testing contributed bomb ^{14}C, it was determined that many marine shells exhibited apparent ages ranging as high as 1200 years. Part of the reason has to do with the fact that ocean water depleted of ^{14}C by long residence times in the deeper parts of the ocean is periodically upwelled or brought to the surface and mixed with surface ocean water. The effect is to dilute the contemporary ^{14}C activity of the surface ocean near the westward-facing continental margins, resulting in a spurious apparent age for the organisms utilizing surface-ocean-water carbonates. Shells growing in locations adjacent to the outlets of major river systems whose water is depleted of ^{14}C as a result of exchange with limestone or other carbonate-bearing rocks can also give spurious apparent ages. This is probably the explanation for the false ages exhibited by shells growing in the Gulf of California (Colorado River discharge) and the northern part of the Gulf of Mexico (Mississippi River discharge). *See* CONTINENTAL MARGIN; UPWELLING.

Upwelling and other reservoir effects are highly variable depending on location and specific environmental conditions. For the western coasts of North and South America, for example, the magnitude of the upwelling effects can range from about 80 to 1000 years. It is most severe along the Peruvian coast, contributing to an explanation for the problematical ^{14}C values on marine shells from that region. Unfortunately, it is possible for shells from highly localized regions to exhibit a sizable range in apparent ages. Samples from the Galápagos Islands show a variation of about 350 years. Such a fluctuation in such a relatively small area emphasizes the fact that the magnitude of upwelling effects for any region must be carefully established by multiple sampling of closely spaced areas.

Problems similar to those associated with marine shells arise for a number of sample types and geochemical environments where contemporary samples may not be in equilibrium with the atmosphere. In each case, empirically derived values for the contemporary standard must be obtained for each sample type or locality, or both. In practice, this is accomplished by determining the degree of deviation from whatever contemporary standard is used to define modern or "zero ^{14}C age" samples. For example, specific values are required for marine shells from different oceanic regions, for fresh-water shells in specific terrestrial environments, and for Arctic and Antarctic specimens.

Variability in carbon isotope ratios. For the ^{14}C method, the basic physical measurement used to index time is the $^{14}C/^{12}C$ ratio. However, carbon has three naturally occurring isotopes. Variation in this ratio can be effected by influences other than the decay of ^{14}C. The most common problem occurs when carbon-containing compounds not indigenous to the original samples are physically or chemically introduced into the sample matrix resulting in the contamination of the sample. Usually less difficult to deal with are fractionation effects in which a variation in the stable carbon ratio translates into a change in the $^{14}C/^{12}C$ ratio.

Contamination. The sources and effects of the introduction of foreign organics into samples are complex; they depend on the nature and condition of the sample materials, the characteristics of the environment to which the samples were exposed, and the period of time over which the exposure occurred. Precautions exercised to avoid contamination effects are unique to each sample type and source locality.

A series of procedures to remove potential contaminants in samples has been established by research laboratories. Most sample preparation techniques are concerned with completely removing what is assumed to have not been present when the original sample died or was removed from exchange with its carbon reservoir. Samples such as wood and charcoal, which can be subjected to treatment with strong acids and bases to facilitate the removal of absorbed carbonates and soil humic and fulvic acids and other soluble soil organic matter, are preferred. Less desirable are cases where it is difficult to distinguish between contamination and the original sample as with various types of carbonate samples, such as tufa and caliche.

It is usually possible to infer the effect of known contamination effects on a given sample in terms of the direction that the age change will take for a given type of contamination, but the magnitude of the errors can be calculated only if the true age of the original sample, the age of the contaminant, and the percentage contribution of the contaminant are all known. Usually this is difficult to determine. However, with few exceptions, problems of contamination for samples with ages of less than about 10,000 years can be solved, usually by applying standard pretreatment approaches developed by ^{14}C laboratories. For materials with expected ages in excess of 10,000 years, sample contamination problems typically become more serious, and laboratories must exercise even more rigorous care in the pretreatment processes.

Fractionation effects. While all the isotopes of carbon follow the same chemical or physical pathway, the rate at which this occurs varies as a function of their difference in mass. The pioneering studies of Harmon Craig pointed to the need to consider variations in the stable isotope ratio ($^{13}C/^{12}C$) of samples to obtain precise ^{14}C values. Variations equivalent to up to several hundred years can result if ^{14}C values are not standardized in light of $^{13}C/^{12}C$ ratios. Fortunately, no significant fractionation effects are usually observed in standard sample materials such as charcoal or wood. Problems arise, however, when it is necessary to compare ^{14}C values from a variety of sample types such as grasses, grains, seeds, succulents, and marine carbonates, as well as standard terrestrial organics. In such cases, it is necessary to use the stable isotope ratios to correct the ^{14}C values onto a common scale.

Variability in amount of carbon. In addition to the variation in production and distribution of ^{14}C over time and within portions of various carbon reservoirs, variations may result from situations where carbon not in equilibrium with the contemporary standard values is added or removed from any reservoir. Two instances are well documented since they occurred within the last century as a result of human intervention in the carbon cycle. The first, beginning in the middle of the nineteenth century, is known as the industrial or Suess effect. The combustion of fossil fuels added enough "dead" ^{14}C to the atmosphere to result in the reduction by about 3% in biospheric ^{14}C activity. In the more recent atomic bomb or Libby effect, relating to the detonation of thermonuclear devices in the atmosphere in the early 1950s, large amounts of artificial ^{14}C were produced, almost doubling the amount of ^{14}C in the terrestrial biosphere. When combined with the late-eighteenth-century De Vries excurses, the Suess effect makes it difficult to distinguish ^{14}C concentrations within the last two centuries. This is one of the reasons why laboratories generally use 100 or 150 years as the minimum age which can be cited. It also explains why laboratories cannot use modern wood as a contemporary reference.

Half-life of radiocarbon. The fundamental constant which permits the conversion of a $^{14}C/^{12}C$ ratio into an age value is the half-life or decay constant of ^{14}C. Initially in the development of the method, Libby and collaborators used the value 5720 ± 47 as the half-life figure, but soon adopted the weighted average of three independently obtained measurements. The average value was 5568 ± 30 and this became identified as the Libby half-life. In 1962, at the 5th Radiocarbon Dating Conference at Cambridge, it was decided that 5730 ± 40 probably represented a more accurate approximation of the actual half-life. It was agreed, however, that the Libby half-life would be used in the calculation of conventional ^{14}C determinations. The stated reason was that any changes in the value would introduce unneeded confusion in the radiocarbon literature.

The issue of the correct half-life for ^{14}C has lost a considerable amount of its significance because of the discovery and documentation of similiar variation and De Vries effects. The existence of dendrochronologically documented relationships between ^{14}C age and calendar age, for samples up to about 10,000 years old, enables researchers to circumvent the problem of the actual ^{14}C half-life and proceed to calibrate these ^{14}C age values directly.

Statistical and contextual uncertainties. Most ^{14}C determinations are expressed in the form: age value (in ^{14}C years B.P.) $\pm$ statistical uncertainty. The age value is calculated by using the equation previously presented. The measurement uncertainty results from statistical considerations inherent in the random decay process characteristic of all radioactive isotopes. A date, for example, of 5600 ± 80 ^{14}C years B.P. reflects the fact that the count rate of the sample is about 50% of the modern reference standard (that is, it has decayed for a period of about one Libby half-life) and the age value is known to about 1% or 80 years. Statistical uncertainties in ^{14}C work are usually cited in terms of one standard deviation errors. The expression 5600 ± 80 is a shorthand manner of stating that there are two chances out of three that the age equivalent of the counting rate for this sample will be contained within the range 5520 to 5680. An accurate statement of the results of a ^{14}C determination must include a listing of the measurement uncertainty. In addition, some laboratories increase this value to take into consideration changes that affect counting conditions, such as drift of electronic

equipment and changes in barometric pressure. Statistical errors are not cited only when the counting rate of a sample is statistically indistinguishable from the background counting rate for the counter being used. The result is expressed as a minimum or infinite value by stating that the age is "greater than" a limit imposed by the characteristics of the counting system being employed (for example, >40,000).

The processing and counting of a sample to determine its ^{14}C age is a challenging analytical procedure. However, a technically correct value which has been carelessly collected may be scientifically worthless. Often the significance and importance of a ^{14}C determination are only as good as the attention to detail which went into documenting the geological, historical, or archeological context of the sample. It is important to be aware of what a ^{14}C date does and does not indicate. A ^{14}C value provides a temporal index of when the sample was removed from its reservoir. For example, a ^{14}C determination on a piece of charcoal or wood provides an age value for the tree rings which make up the sample. A ^{14}C date taken on a piece of wood taken from a beam excavated from a ruined structure may indicate the time when the building was constructed if the sample was taken from the outside rings of the tree used as the source of the timber and if the timber itself did not happen to be reused from an earlier structure. Problems such as these often confront geologists and archeologists as they attempt to critically interpret the dating evidence provided by the ^{14}C method.

R. E. Taylor

Bibliography. E. Bard et al., Radiocarbon calibration by means of mass spectrometric $^{230}Th/^{234}U$ and ^{14}C ages of corals: An updated base including samples from Barbados, Mururoa and Tahiti, *Radiocarbon*, 40:1085–1092, 1998; D. Polach, *Radiocarbon Dating Literature: The First 21 Years*, *1947–1968*, 1988; M. Stuiver, J. von der Plicht, and A. Long (eds.), Calibration Issue INTCAL 98, *Radiocarbon*, 40:1–1160, 1998; R. E. Taylor, Fifty years of radiocarbon dating, *Amer. Scientist*, 88:60–67, 2000; R. E. Taylor, Radiocarbon dating, in R. E. Taylor and M. J. Aitken (eds.), *Chronometric Dating in Archaeology*, 1997; R. E. Taylor, *Radiocarbon Dating: An Archaeological Perspective*, 1987; R. E. Taylor, R. Kra, and A. Long (eds.), *Radiocarbon After Four Decades: An Interdisciplinary Perspective*, 1992.

Radiochemical laboratory

A laboratory or facility for investigation and handling of radioactive chemicals which provides a safe environment for the worker and the public. It's features can vary depending on the type of radioactive emissions to be handled, the quantity, the half-life, and the physical form (solid, liquid, gas, or powder). Special measures to minimize spread of contaminated material and to dispose of radioactive waste are required. Working surfaces should be smooth and easily washable to permit effective decontamination if necessary. Good ventilation and detectors for monitoring radiation and contamination on surfaces or people are also typical features. Space for clothing change, lockers, and showers are commonly provided.

Investigations utilizing only very small amounts (a few microcuries) of beta or gamma emitters which are not readily dispersed (no powders or volatile liquids) may sometimes be performed without special facilities on the bench top. In this case, precautions such as working on plastic-backed absorbent paper (lightweight disposable diapers are useful) and wearing protective gloves and lab coat may be sufficient. A special bag or can for disposing of the paper and gloves as radioactive waste is required. If the radioactive isotopes are solely alpha-particle emitters, containment and isolation from direct contact are more serious concerns. Due to the limited penetration but high biological toxicity of alpha particles, it is essential to avoid ingestion or inhalation. For very small quantities, work may take place with double rubber gloves in a fume hood with appropriate filter. An air velocity greater than 125 ft/min (38 m/min) into the hood at the opening is necessary for suitable isolation. Generally, an enclosed glove box (**Fig. 1**) is used, situated inside a hood and maintained at negative pressure with respect to the face of the hood and the room. Sensors to monitor proper differential pressure and adequate airflow are usually used to assure containment and to generate an alarm if conditions degrade. Many alpha emitters decay into other radioactive isotopes which subsequently decay into still other radioactive species. Some members of these long decay chains are gamma-emitting or radon isotopes. If the quantity of radioactivity is high, shielding from the penetrating gamma rays and special traps for radon gas emanating from the materials are required.

For work with pure beta emitting isotopes, for example phosphorus-32, confinement within a hood and the use of indirect means of handling the radioactivity is recommended (**Fig. 2**). Long-handled tongs or other tools, sometimes with an attached piece of Plexiglass, approximately 3/16 in. (0.5 cm) thick, are used for higher levels of radioactivity in order to shield the hands. Generally, since most beta emission is also accompanied by penetrating gamma emission, the entire work area must be enclosed in heavily shielded enclosures. *See* ALPHA PARTICLES; BETA PARTICLES; RADIATION CHEMISTRY; RADIOACTIVITY; RADIOCHEMISTRY; RADIOISOTOPE.

Hot laboratories. These special radiochemistry laboratories contain walled enclosures for remotely handling larger quantities of gamma-emitting isotopes. A small enclosure is usually referred to as a cave, while large ones are called hot cells. Hot cells are usually equipped with remote manipulators and thick windows made from high-density lead glass (**Fig. 3**).

Construction. Many different shielding materials have been used for hot cells. High-density concrete is useful because of its relatively low cost. This shielding can be improved with the addition of iron powder to the concrete. When floor space is limited or more effective shielding is required, steel or lead can be

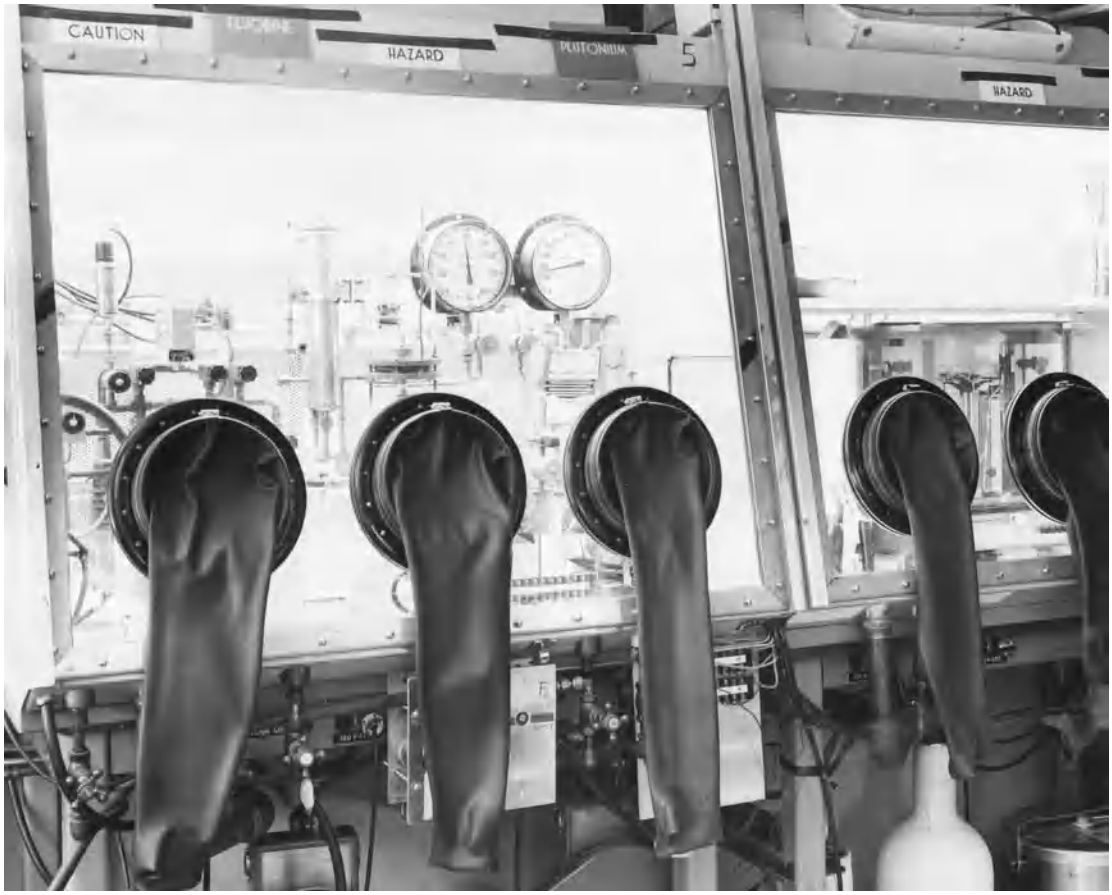

Fig. 1. Radiochemical laboratory, hood-equipped, with rubber-gloved openings. (*Argonne National Laboratory*)

Fig. 2. Fume hood set up with portable shielding and small tongs for working with low-level beta- and gamma-emitting isotopes. (*Brookhaven National Laboratory*)

used for walls and ceilings. Most commercially available hot cells utilize poured lead or interlocking lead brick encased in steel plate. S-shaped penetrations in the walls or ceiling are usually provided for electrical wires, and for tubing to allow water connections and reagent addition. This prevents any straight-line path for radiation to shine through. For disposing of aqueous radioactive waste, small cup sinks with shielded drain lines must be installed. Shielded storage tanks for this waste are usually installed under the floor or in a basement room. Some radiochemistry laboratories specialize in investigations with the short-lived positron-emitting isotopes fluorine-18 (half-life 128 min), carbon-11 (20 min), nitrogen-13 (10 min), and oxygen-15 (2 min). In this case, radioactive waste disposal is significantly simplified by just putting it aside and waiting for decay to turn it into normal waste. Thus, extensive storage areas for either solid or liquid radioactive waste are not necessary.

Hot-cell viewing windows are made from lead glass with a thickness providing equivalent shielding to the walls. Dopants containing cerium are often added to the innermost glass panes to help resist darkening and discoloration caused by high levels of radiation exposure. Since high-density lead glass is yellow in color, bright interior lighting is required for good visibility. It is good design to allow a light bulb to be changed from outside the hot cell. Sometimes closed-circuit television or magnifying

Fig. 3. Exterior view of two hot cells with manipulators. The walls are 7-in.-thick (17.75-cm) lead, encased in 1/4-in.-thick (0.6-cm) steel plate. The lead glass windows are 14 in. (35.5 cm) thick and provide shielding equivalent to the walls. (*Brookhaven National Laboratory*)

periscopes are provided for special inspection needs. Before first use, the shielding integrity of the entire assembly should be checked with a small size, high-level radioactive source. Special attention needs to be devoted to surveying radiation levels around the outside of wall-to-wall joints, conduit or reagent penetrations, and ventilation ducts. *See* RADIATION SHIELDING.

Operations. Hot cells are usually grouped together in order to provide separate processing areas for different isotopes. This minimizes the chances of cross contamination. However, it is efficient to have a single common hot cell to receive and open irradiated targets and to dispose of waste, as well as a single common hot cell with a porthole for dispensing of final product. The materials to be processed can be moved with manipulators through internal connecting portholes between the hot cells, or by conveyor belts or equivalent. In several installations, large-gauge toy electric trains serve this purpose. Most hot cells have small portholes or drawers for sample insertion and removal. In addition, there must be a larger opening, such as a door or a removable wall panel, for major maintenance, cleanup work, and equipment installation. The type of work carried out in hot cells is similar to that in conventional laboratories. This includes materials inspection, wet chemistry, and analytical measurements. The specific equipment must be inserted into the hot cell prior to use with hot sources. Since space is at a premium, selection and location of the equipment (such as balances, remotely operated pipettes, automatic bottle capping device and crimp sealers, and specialized glassware) are important.

A significant difference in operation exists between hot laboratories specializing in preparing radiopharmaceuticals with the short-lived positron emitters compared to those handling longer-lived radioisotopes. The very short lifetimes require very rapid processing and compound labeling. Also, as many such agents are used clinically in human diagnostic nuclear medicine procedures, they must be prepared several times a day, five or six days a week. This preparation burden is therefore very labor intensive and expensive. In response, the preparation of many of the standard, frequently used radiopharmaceuticals, such as F-18 fluorodeoxyglucose, has been automated into "black box" synthesis modules for hot cell use. Solutions are moved internally with small liquid pumps or by vacuum through the synthesis steps. These commercially available systems provide rapid, sterile preparations ready for human use. Alternatively, computer programmable robot arms can control synthesis by operating syringes, moving flasks, and other motions. This approach requires more setup effort than a black box assembly, but also provides more flexibility since the robot arm can be reprogrammed for different motions. *See* NUCLEAR MEDICINE; RADIOACTIVE WASTE MANAGEMENT; ROBOTICS.

Remote manipulators. Many hot-cell operations are performed using remote manipulators. Many varieties are now available in different sizes and capacities. The most generally useful is the mechanical master-slave manipulator. This device features a mechanical arm with a wrist and jaws at the end located inside the hot cell. The movements of this slave device precisely follow the movements of the operator's arm, wrist, and hand. Generally the force applied by the operator is directly transmitted without magnification up to the capacity limit of the device. Manipulators with lift capacity ranging 10–100 lb (4.5–45 kg) are the most common. Many manipulators are mounted such that the slave arm penetrates the ceiling of the hot cell through a rotating steel or lead shield ball. Since a wide range of motion is desirable, it is difficult to completely shield the area at the manipulator penetration. Small leaks through a ceiling area may be tolerable since it is not generally an occupied space. Another style uses a horizontal barrel mounted through the front face of a hot cell near the top, with offset shield plugs at both the inside and outside of the penetration. This style allows the inner slave arm to electrically swing forward and back, or swing left and right without moving the master arm. This extended reach capability is critical in large hot cells because it allows the operator to stay in a good viewing location while still reaching the back or corners. Both styles usually cover the slave arm with a plastic sleeve called a boot to minimize contamination of the gears and pulleys inside the arm. This can be important if repair of the manipulator outside the hot cell becomes necessary. Many manipulators allow the jaws or tongs to be remotely removed and changed. Although different styles can thus be accommodated, it is still usually necessary to design special tools or tool handles in order to

improve the reliability of gripping objects (especially small ones). *See* REMOTE-CONTROL SYSTEM; REMOTE MANIPULATORS.

Ventilation. Good ventilation is critical for radiochemical laboratories, particularly with hot cells, to sweep any airborne radioactivity away from the worker. The radioactivity may be a gas, dust, or fume. Filters are then necessary to trap this material and prevent it from entering the local atmosphere. Because dust is a good vehicle for spread of contamination, filtered inlet air can reduce spread of contamination and prolong the lifetime of outlet filters. The primary outlet filter required in such laboratories is the HEPA (high-efficiency particulate air) filter. HEPA filters must be 99.97% efficient, and are required to be tested annually. Typically a roughing filter is also used upstream to reduce the dust loading on the HEPA. In many hot-cell installations, an on-line differential pressure gauge continuously monitors the performance of the HEPA filter. A low-differential pressure value may indicate a rip in the filter, while too high a value would imply that the filter is clogged.

In a multi-hot-cell installation, proper air balance between the hot cells must be maintained to minimize cross contamination. Adequate airflow and static pressure must also be maintained when portholes or hot-cell doors are opened. This is ordinarily accomplished with either an oversized system, or with variable-speed ventilation fans which speed up in response to a drop in hot-cell airflow rate. In laboratories where use of acids is common, for example for dissolving solid irradiated targets, acid fumes can be sucked off into a separate system with an inverted funnel attached to a flexible hose. These fumes are then passed through a neutralizing buffer before joining the main ventilation duct. Charcoal filters are also required if the emission of gaseous radioactivity, for example iodine-131, is anticipated. Where performance of the ventilation system is crucial for radiation safety or environmental protection, electronic sensors with alarms are used to monitor total system flow rate, individual hood or hot-cell flow rate, static pressure, and filter performance. Backup electric power may be necessary to guarantee negative pressure in these urgent situations. *See* AIR FILTER; VENTILATION.

Safety. In addition to containment, radiation shielding, and good ventilation, several other safety measures are commonly employed. To limit spread of contamination outside the laboratory, radiation monitors are often placed at the doorway. These may be simple hand-held Geiger counters with a large-area detecting probe, called pancake detectors. Or in some radiochemistry laboratories, large step-in detector arrays, called portal monitors, can check the entire body for radioactivity. When high levels of radioactivity are present in hot cells, radiation sensors mounted inside can be interlocked with doors and portholes to prevent access. Area radiation monitors are often installed in the room as well as airborne radiation monitors. Both can initiate audible and visual alarms if trip points are reached. *See* INDUSTRIAL HEALTH AND SAFETY; RADIATION BIOLOGY.

Leonard F. Mausner

Bibliography. K. J. Connor and I. S. McLintock, *Radiation Protection Handbook for Laboratory Workers*, 1994; *Manual on Safety Aspects of the Design and Equipment pf Hot Laboratories: 1981*, 1981; *Safe and Secure: Guide to Working Safely with Radiolabelled Compounds*, Amersham International, 1992.

Radiochemistry

A subject which embraces all applications of radioactive isotopes to chemistry. It is not precisely defined and is closely linked to nuclear chemistry. The widespread use of isotopes in chemistry is based on two fundamental properties exhibited by all radioactive substances. The first property is that the disintegration rate of an isotopic sample is directly proportional to the number of radioactive atoms in the sample. Thus, measurement of its disintegration rate (with a Geiger counter, for example) serves to analyze a radioactive compound. With nearly all chemical elements (the most notable exceptions being nitrogen and oxygen, which have no suitable radioactive isotopes), an isotope may be incorporated in a chemical compound, and thereafter, masses of this compound as small as 10^{-6} to 10^{-10} g may be measured with a high precision. Because experimental chemistry depends largely upon analysis, isotopes may be employed in most chemical problems, especially those requiring high analytical sensitivity. The second fundamental property is that the disintegration rate is completely unaffected by the chemical form of the isotope, and conversely, the property of radioactivity does not affect the chemical properties of the isotope. By substituting or labeling a particular atom within a molecule, isotopes can be used to trace the fate of that atom during a chemical reaction. In contrast to physical migration tracer studies, the compounds arising in a reaction must first be isolated in separated pure forms before radioactive assays can be performed. *See* RADIOACTIVE TRACER.

In general, radiochemical studies can be classified according to whether the use of isotopes represents a convenient or a unique solution to a problem.

Convenient applications. These applications usually exploit the high sensitivity of tracer techniques because alternative analytical procedures are slower and often less accurate. The efficiencies of chemical separations, such as those based on selective precipitation, solvent extraction, ion exchange, and electrodeposition reactions, are studied by labeling the desired compound and following the radioactivity during the separations. The rate and extent of adsorption on solid surfaces of either labeled solutions or labeled gases are rapidly determined by assaying periodically the mobile phase, or better, the solid phase. New chemical phenomena, such as the coprecipitation of trace elements and radiocolloid formation, occur at submicro concentrations (10^{-10} g/liter

of solution) and may be studied most conveniently with isotopes. The solubility of an "insoluble" precipitate is measured by saturating a solvent medium with a radioactive solid. Similarly, the vapor pressure of a solid is measured by saturating an evacuated volume with vapor or, for pressures below 10^{-4} mm, by effusing vapor into a cooled target, which is later assayed. Qualitative and quantitative analysis for most trace elements present in parts per million or less in a sample is possible by radioactivation analysis. The sample is irradiated in a flux of neutrons or other suitable particles, and the trace element is identified and determined by its induced activity. Depending upon the element, quantitative determination of masses of 10^{-8} to 10^{-12} g is usually possible. *See* NUCLEAR REACTION.

Unique applications. An understanding of diffusion processes is of considerable importance because the rates of many chemical reactions are governed by the rate at which chemical species can diffuse through a medium to the point of reaction. For example, the rate of many electrode processes depends upon the rate of diffusion of electrolyte to the electrode, and the rate of oxidation of copper is determined by the rate of diffusion of copper ions up to the metal surface. If a layer of radioactive copper is sandwiched between two ordinary copper samples, it is found that at elevated temperatures copper ions will diffuse considerable distances within the metal. The rate of diffusion of copper in copper (that is, the self-diffusion rate of copper) can be observed only by the transfer of radioactivity from the labeled region into unlabeled regions, thin slices being removed from the solid at known distances and then being assayed. The **illustration** shows some experimental points obtained in such an experiment. The distribution curve is that expected from the integrated form of Fick's law of diffusion, Eq. (1), where c is

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2} \tag{1}$$

the concentration (activity) of the diffusing tracer, t is the diffusion time, x the diffusion distance, and D the self-diffusion coefficient. In addition to solid-state studies with elements, alloys, metallic oxides, and inorganic salts (all of which have important metallurgical implications), self-diffusion experiments are performed with liquids and gases. In all cases, they provide valuable information on the nature of the intermolecular forces which determine the magnitude of D.

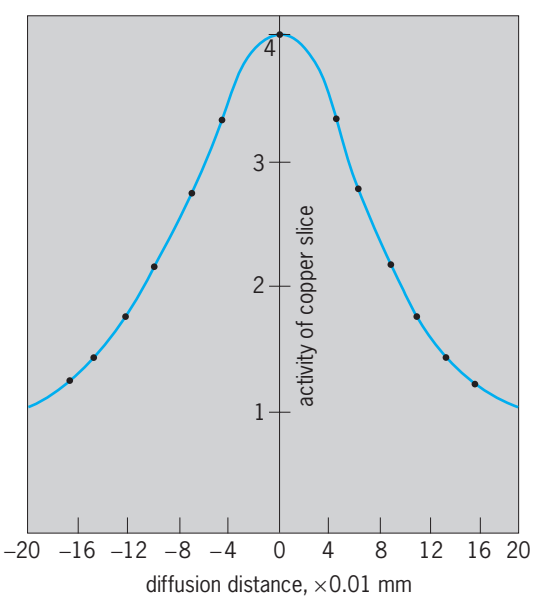

Distribution of radioactive copper after 605 min diffusion at 1742°F (950°C). The full curve corresponds to $D = 8.92 \times 10^{-10}$ cm²/s.

Isotopic exchange reactions. When slightly soluble lead chloride crystals are mixed with an aqueous solution of sodium chloride, labeled with chlorine-36, radioactivity rapidly appears in the lead chloride as a result of exchange of chloride ions between the two compounds. Both compounds produce chloride ions on dissociation, and some chloride ions, originating from the sodium chloride, become associated with the lead ions, thereby leading to radioactivity in the lead chloride, as in reaction (2), where s

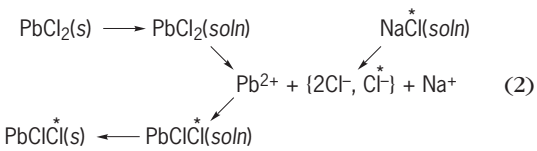

(2)

indicates solid and *soln* indicates solution. Exchange processes are proceeding continuously, but they can be detected only with isotopes, hence the term isotopic exchange reactions. Exchange reactions may occur between any two species of molecules having a common atom or group. They may be due to a dissociation process (as above) or to a collision between the two species in which chemical bonds are formed and broken (a bimolecular process), as in the exchange of radioactive sulfur (denoted by an asterisk) between sulfite and thiosulfate ions, reaction (3).

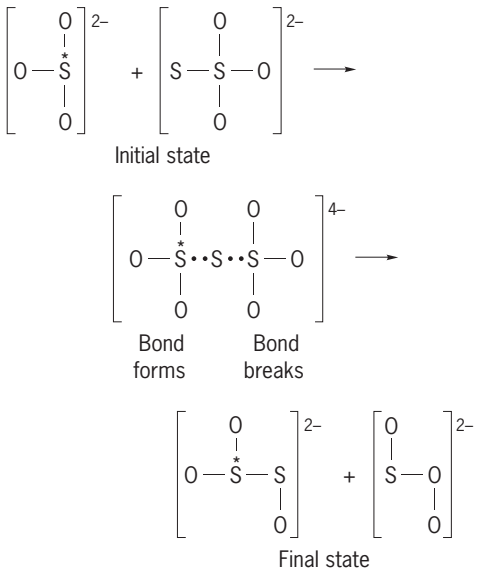

(3)

With two oxidation states of an element, exchange occurs by the transfer of an electron (an electron exchange reaction), for example, reaction (4), although

$$Fe^{II} + \overset{*}{Fe}^{III} \rightarrow [Fe^{II} \xrightarrow{e} \overset{*}{Fe}^{III}] \rightarrow Fe^{II} + \overset{*}{Fe}^{II} \qquad (4)$$

in many cases the electron may actually be transferred by an atom or a group bridging between the reactants. The rates of exchange reactions are measured by separating the two reactants at various times and determining the fraction F of the total radioactivity in each species. For an initially unlabeled species, F increases exponentially with time to an equilibrium value corresponding to equal concentrations (specific activities) of isotope in each species; for an initially labeled species, F will decrease exponentially with time. As in ordinary chemical kinetics, the exchange mechanism is deduced from the dependence of the exchange rate upon reactant concentrations. Rate studies of exchanges by dissociation provide information concerning ionization (including acid-base equilibria) of both solutes and solvents, the reaction of solvents with molecules (solvolysis), the thermal dissociation of gases, and the dissociation of gases on catalyst surfaces. Bimolecular exchanges are exceptionally important in studies of oxidation-reduction reactions and of substitution reactions of coordination complexes of the transitional elements. The rates of exchange by substitution provide a direct measure of the lability of these coordination complexes.

Isotopic tracer studies. Details of reaction mechanisms are provided by labeling a specific atom within a molecule. Thus, when carboxyl-labeled propionic acid is oxidized in acid dichromate, the product, as anticipated, is only radioactive carbon dioxide, reaction (5). However, the oxidation mechanism is

$$CH_3CH_2\overset{*}{C}OOH \rightarrow HOOC{-}COOH + \overset{*}{C}O_2 \qquad (5)$$

more complex in the case of alkaline permanganate because the isotopic distribution is 75% labeled oxalate and only 25% carbonate. When the masses of the labeling and the normal isotopes differ markedly (notably with hydrogen, carbon, and oxygen), the isotopic molecule will react more slowly than the normal molecule if the reaction mechanism involves significant stretching of the chemical bond to the isotopically substituted atom. Such isotope-effect studies provide mechanistic information.

New elements. The artificially produced transuranium elements and technetium, astatine, and francium are all radioactive, and their chemistry is being elucidated with radiochemical techniques. However, this topic is normally accepted as nuclear chemistry.

Recoil studies. For a discussion of recoil studies *see* NUCLEAR CHEMISTRY; RADIATION CHEMISTRY; RADIOACTIVITY; RADIOCHEMICAL LABORATORY.

Donald R. Stranks

Bibliography. R. A. Faires and G. G. Boswell, *Radioisotope Laboratory, Techniques*, 1980; C. Keller, *Radiochemistry*, 1988; R. Lambrecht, *Nuclear and Radiochemistry*, 1982.

Radioecology

The study of the fate and effects of radioactive materials in the environment. As a hybrid field of scientific endeavor, it is founded upon, and derives its basic principles from, both of its parent disciplines, that is, basic ecology and radiation biology. Following the discovery of ionizing radiation and radioactive particles in biological studies in the 1940s and 1950s, these phenomena were studied under what were usually controlled laboratory conditions. Soon after, however, a need was demonstrated for a better understanding of the fate and effects of radioactive materials that were being released into the environment following the use or testing of nuclear weapons. This need became an important factor in the emergence of radioecology as a scientific discipline in its own right.

Radiation effects. Because of its early association with nuclear weaponry, radioecology was initially focused on radiation effects, usually under the umbrella of concern for survival under postnuclear attack conditions. Mortality induced by radiation exposure has usually been measured by using some form of lethal dose 50. However, the amount of radiation exposure required to produce mortality under free-living conditions in the field (ecological lethality) may often be substantially less than that required to cause physiological lethality when the organism is held under protected conditions in the laboratory. In the case of young bluebirds (*Sialia sialis*), for example, exposure to radiation levels as low as 800–900 roentgens (R) causes the stunting of wing feather growth, which would reduce the bird's ability to fly and escape predators. Therefore mortality in the field would result from these levels of exposure. When birds are hand-raised and protected from predators in the laboratory, however, exposure to as much as 2500 R would be required to produce 50% mortality. *See* LETHAL DOSE 50.

Similar considerations also apply to sublethal radiation effects, such as perturbations of growth rate, reproduction, and behavior. These responses to radiation stress have consequences for both the individual organism and for the population, community, or ecosystem of which it is a part. When populations or individuals of different species differ in their sensitivities to radiation stress, for example, the species composition of the entire biotic community may be altered as the more radiation-sensitive species are removed or reduced in abundance and are replaced in turn by more resistant species. Such changes have been documented by studies in which natural ecological systems, including grasslands, deserts, and forests, were exposed to varying levels of controlled gamma radiation stress (**Fig. 1**). *See* POPULATION ECOLOGY.

Techniques of laboratory toxicology are also available for assessing the responses of free-living animals to exposure to low levels of radioactive contamination in natural environments. This approach uses sentinel animals, which are either tamed, imprinted on the investigator, or equipped with miniature

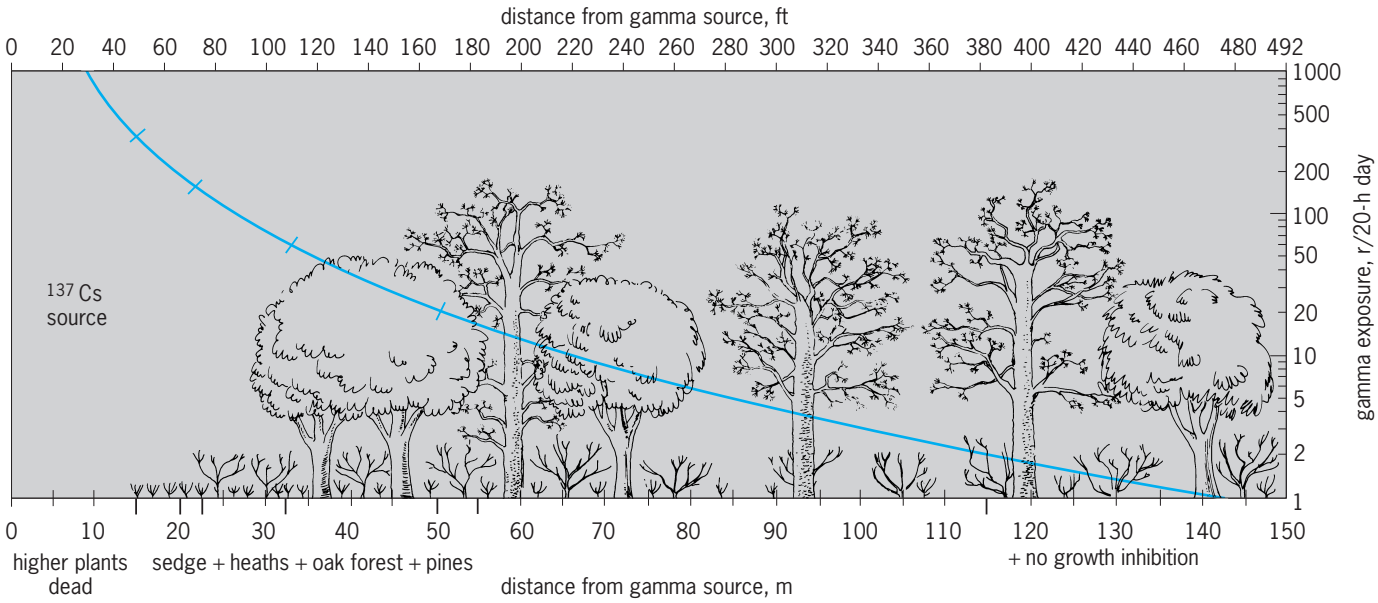

Fig. 1. Gradations in response of the vegetation of an oak-pine forest to continuous exposure to gamma radiation stress. The radiation source was located at the left side of the figure and, as indicated by the curve, exposure levels decreased from left to right with increasing distance from the source. (*After G. M. Woodwell, Radiation and the patterns of nature, Science, 156:461–470, 1967*)

radio transmitters, to permit their periodic relocation and recapture as they forage freely in the food chains of contaminated habitats. When the animals are brought back to the laboratory, their level of radioisotope uptake can be determined and blood or tissue samples taken for analysis. In this way, even subtle changes in deoxyribonucleic acid (DNA) structure can be evaluated over time. These changes may be suggestive of genetic damage by radiation exposure. In some cases, damage caused by a radioactive contaminant may be worsened by the synergestic effects of other forms of environmental contaminants such as heavy metals.

In other cases, assessment of responses to radiation exposures has documented the phenomenon of hormesis, whereby an organism responds to very low-level exposures in the opposite way from which it would respond to higher levels of radiation stress. For example, while growth rates of plants and animals normally decline after high levels of radiation stress, exposure to low levels of radiation may actually result in a stimulation of growth. Moreover, the documentation of such effects may suggest that some physiological processes may be adapted to perform optimally under very low levels of background radiation exposure, such as are provided naturally from cosmic rays and natural sources of radioisotopes in certain rock and soil substrates. The nuclear accidents and other releases of radioactive contaminants at various sites in the former Soviet Union have provided natural laboratories for ecological studies of the effects of radiation contamination.

Radionuclide tracers. Because of the ease with which they can be detected and quantified in living organisms and their tissues, radioactive materials are often used as tracers to study the rates and patterns by which biological processes take place under natural conditions. The use of such radioactive tracers in medicine is well known. In the same way, radioactive tracers can also be used to trace food chain pathways or determine the rates at which various processes take place in natural ecological systems (**Fig. 2**). Gamma-emitting isotopes are particularly useful in this regard, since their emissions from deep within the body of a plant or animal allow them to be detected and quantified by sensitive counting

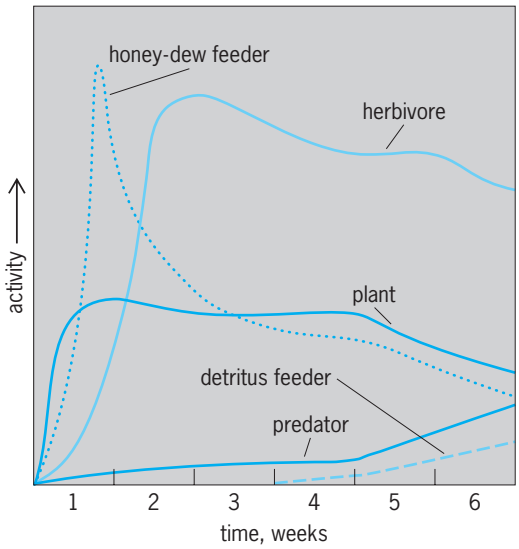

Fig. 2. Varying patterns of uptake of radioactivity by organisms occupying different trophic levels in an old-field ecosystem, following the introduction of radioactive tracers into the plants upon which their food web was based. Animals that suck plant juices and directly consume the vegetation acquired maximum activity levels earlier than did predators and detritus feeders. (*After R. G. Wiegert, E. P. Odum, and J. H. Schnell, Forb-arthropod food chains in a one-year experimental field, Ecology, 48:75–82, 1967*)

equipment without the need for sacrifice and dissection. Other uses of radioactive tracers include tagging adult animals in order to later identify the particular eggs or young they produce, and using radioactive markers to later relocate and recapture animals that are too small to carry even the tiniest of radio transmitters.

Although most of these tracer experiments were performed in the past by deliberately introducing a small amount of radioactive tracer into the organism or ecological system to be studied, they now take advantage of naturally tagged environments where trace amounts of various radioactive contaminants were inadvertently released from operating nuclear facilities such as power or production reactors or waste burial grounds. As long as the amounts of such releases are less than the amounts that could be considered hazardous to health, such low-level contaminated habitats have actually become valuable research sites for conducting studies of functional processes associated with the native flora or fauna of the area. *See* RADIOACTIVE TRACER.

Radioactive contamination. In some cases, the inadvertent release of radioactive materials into the environment creates concern for the health and well-being of humans and other organisms living in the area. An important component of radioecology, and one that is closely related to the study of radioactive tracers, is concerned with the assessment and prediction of the movement and concentration of these radioactive contaminants in the environment in general, and particularly in food chains that may lead to humans. Primary concern is, of course, focused on agricultural pathways through which radionuclides may enter the human food chain through crop plants or meat, milk, or other products of domestic livestock. Often overlooked, however, is the transfer of radioactive contaminants to humans who may consume fish or wild game as food. In many parts of the world, wild game animals such as deer have been found to have consistently higher concentrations of environmental radioactive contaminants than domestic livestock grazing fertilized pastures in the same region. This phenomenon is most noticeable in regions where natural vegetation has difficulty obtaining sufficient nutrients from the soil and must rely more heavily on "fallout" nutrients distributed by atmospheric processes. In high alpine regions and arctic tundra, for example, the concentration of radioactive contaminants in species such as reindeer may be particularly severe, as was observed following the Chernobyl nuclear accident in the Soviet Union. *See* FOOD WEB.

The Chernobyl accident was also important because it demonstrated the potential for global transport of radioactive contaminants by both physical forces, such as atmospheric circulation and meteorological factors, and biological vectors such as waterfowl, which might accumulate contaminants at the site of accidental release and then distribute them elsewhere along their migratory journeys. Studies of the migratory habits of these birds suggest that such biological transport of contaminants may not always follow the same patterns and rates as the distribution of contaminants by physical forces. This, in turn, emphasizes the importance of thoroughly understanding all aspects of the basic ecology and natural history of organisms involved in important issues of radioecological study and concern.

Studies in radioecology, when conducted under field conditions, often produce results that are either unexpected or may reverse the conventional wisdom derived from studies conducted under controlled laboratory conditions alone. An example of the latter is the failure of data from some radioecological field studies to support the popular notion of biomagnification, the systematic increase in levels of contaminant concentration in higher levels of food chains. Although some other environmental contaminants such as pesticides or heavy metals may behave in this fashion in natural food webs, such is certainly not always the case in situations documented for radioactive materials (**Fig. 3**).

Finally, situations where the environmental behavior of radioactive contaminants fails to agree with the predictions of laboratory studies demonstrate how radioecological research can contribute new information about some of the ecological mechanisms responsible for contaminant uptake, cycling, and transport in a more general sense. This information can help to provide a better understanding of the environmental behavior of other forms of contaminants such as pesticides or heavy metals, which, in contrast to radionuclides, are more difficult to detect and measure under natural

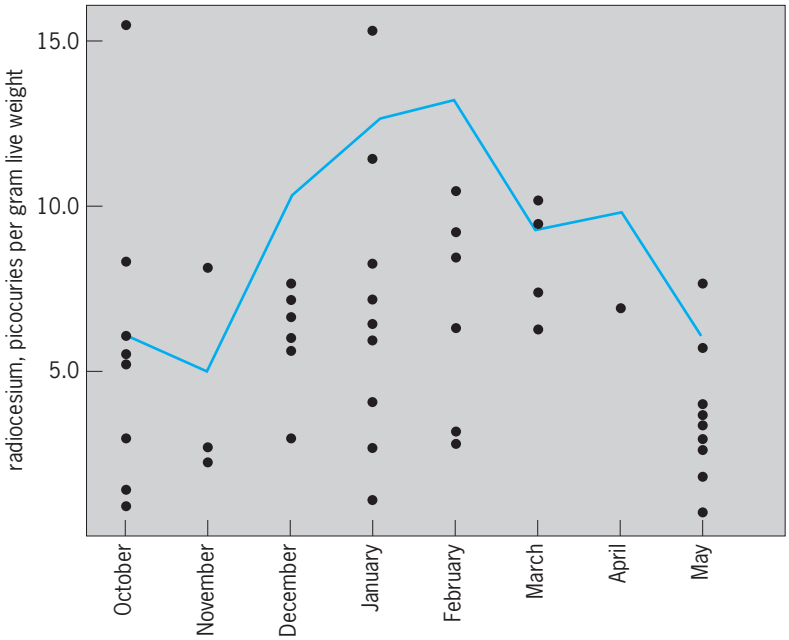

Fig. 3. Monthly changes in levels of radioactive contaminants accumulated by migratory waterfowl wintering on a nuclear-production-reactor cooling reservoir. Contamination levels shown by the strictly herbivorous American coot (*Fulica americana*; curve) nearly always exceeded those shown by other more omnivorous and carnivorous species (solid circles). (*After I. L. Brisbin, Jr., R. A. Geiger, and M. H. Smith, Accumulation and redistribution of radiocaesium by migratory waterfowl inhabiting a reactor cooling reservoir, in Environmental Behavior of Radionuclides Released in the Nuclear Industry, International Atomic Energy Agency, 1973*)

field conditions. *See* ECOLOGY; ENVIRONMENTAL RADIOACTIVITY; ENVIRONMENTAL TOXICOLOGY; RADIATION BIOLOGY. I. Lehr Brisbin, Jr.

Bibliography. M. Eisenbud, *Environmental Radioactivity: From Natural, Industrial and Military Sources*, 4th ed., 1997; M. Garcia-Leon and G. Madurga (eds.), *Low Level Measurements and Their Application to Environmental Radioactivity*, 1988; E. Holm, *Radioecology*, 1994; T. D. Luckey, *Hormesis with Ionizing Radiation*, 1980; E. P. Odum, Feedback between radiation ecology and general ecology, *Health Physics*, vol. 11, 1985; V. Schultz and F. W. Whicker, *Radiological Techniques*, 1982; F. S. Sterrett (ed.), *Environmental Sciences*, 1987; F. W. Whicker and V. Schultz, *Radioecology: Nuclear Energy and the Environment*, 2 vols., 1982.

Radiography

The technique of producing a photographic image of an opaque specimen by the penetration of radiation such as gamma rays, x-rays, neutrons, or charged particles. When a beam of radiation is transmitted through any heterogeneous object, it is differentially absorbed, depending upon the varying thickness, density, and chemical composition of this object. The image registered by the emergent rays on a photographic film adjacent to the specimen under examination constitutes a shadowgraph or radiograph of its interior. Radiography is the general term applied to this nondestructive film technique of testing the gross internal structure of any object, whether it be of the chest of a patient for evidence of tuberculosis, silicosis, heart pathology, or embedded foreign objects; of bones in case of fractures or of arthritis or other bone diseases; or of a weld in a pipe to observe cracks, inclusions, or voids. Radiography with x-rays is commonly used in both medical and industrial applications. Industrial work also involves gamma and neutron radiography. Radiography with charged particles is under development. Most of this discussion will be concerned with radiography with x-rays and gamma rays. *See* CHARGED PARTICLE BEAMS; GAMMA RAYS; NEUTRON; X-RAYS.

Industrial radiography enables detection of internal physical imperfections such as voids, cracks, flaws, segregations, porosities, and inclusions. It is frequently used for visualization of inaccessible internal parts in order to check their location or condition. It is extensively applied wherever internally sound metallic components are required such as (1) in the foundry industry to guarantee the soundness of castings; (2) in the welding of pressure vessels, pipelines, ships, and reactor components to guarantee the soundness of welds; (3) in the manufacture of fuel elements for reactors to guarantee their size and soundness; (4) in the solid-propellant and high-explosives industry to guarantee the soundness and physical purity of the material; and (5) in the automotive, aircraft, nuclear, space, oceanic, and guided-missile industries, whenever internal soundness is required.

The general term applied to radiation imaging and inspection is radiology. This includes film and similar photographic image methods, such as radiographic paper, under the term radiography. In medical circles, the term roentgenography, derived from the name of the discoverer of x-rays, W. C. Roentgen, is used. The older technique of registering an image on a fluorescent screen is called fluoroscopy. The fluoroscopic prompt-response imaging of radiation has largely been replaced by electronic detection with image intensifiers or sensitive television cameras. This technique, called radioscopy, is now widely used in both medical and industrial applications. *See* LIGHT AMPLIFIER; RADIOLOGY; TELEVISION CAMERA TUBE.

Other variations of radiation imaging include xeroradiography, microradiography, flash radiography, and computerized tomography. Xeroradiography is a dry-plate, electrostatic image method similar to that used in photocopy machines. Microradiography involves a magnified image to improve spatial resolution and the detection of small detail. Flash radiography is the production of an x-ray image in a very short time, of the order of nanoseconds, in order to stop fast motion. The computerized tomography (CT) method recreates an image that is essentially a slice through the object. Computerized tomography has made a strong impact on medical diagnosis and industrial inspection. *See* COMPUTERIZED TOMOGRAPHY; MICRORADIOGRAPHY; PHOTOCOPYING PROCESSES.

In many cases, radiographic inspection is a satisfactory method, but there is often a need for complementary, nondestructive inspection methods, such as ultrasonics (also in medicine), magnetic-particle inspection, and electromagnetic infrared and microwave test methods. *See* NONDESTRUCTIVE EVALUATION.

Absorption laws. X-radiography is a straightforward application of the well-known exponential absorption law given below, where I_0 is the initial

$$I_x = I_0 e^{-\mu x} = I_0 e^{-(\mu/\rho)\rho x}$$

intensity of the x-ray beam; I_x the intensity after passage through the object of thickness x, ρ the density, μ the linear absorption coefficient, and μ/ρ the mass absorption coefficient. The ratio μ/ρ is a function of the atomic number of the absorbing chemical element and of the wavelength of the x-ray beam. The value of μ/ρ for a compound is equal to the sum of the weight-proportional values for each of the elements present. The penetrating ability of x-rays increases as the wavelength of the rays decreases (hard x-rays). The x-rays of larger wavelength (soft x-rays) are relatively less energetic (less penetration power). The coefficient μ/ρ is proportional to the cube of the wavelength. The product ρx is mass per unit area. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION.

This exponential law and related expressions quantitatively predict why bones absorb more radiation than soft tissues and thus are clearly delineated, or why cracks or blowholes in castings absorb to a

TABLE 1. Types of x-ray tubes and power sources used in radiography

Energy of radiation	X-ray tube	Power source
1–50 keV	Conventional	High-tension transformer
50–250 keV	Conventional	Regular or resonance transformer; gas-filled, portable with tube integral
250–500 keV	Conventional	Cascade or resonance transformer
1–3 MeV	Sealed tube with high potential applied over gradient in center of high potential generator	Resonance transformer or Van de Graaff generator, with tube integral
1–20 MeV	Tuned resonant microwave waveguide	Linear accelerator

lesser degree than the solid metal and thus are disclosed on the radiograph.

Equipment. Because of the wide range of materials subjected to radiographic inspection, a wide array of radiation-emitting sources is required.

X-ray equipment. Somewhat arbitrarily the x-ray equipment and operating conditions may be classified according to energy of radiation (**Table 1**). X-ray equipment with an energy up to about 300 keV can be considered portable. However, linear accelerators have been mounted on trailers. *See* PARTICLE ACCELERATOR; X-RAY TUBE.

Gamma-ray sources. Natural radioactive isotopes, such as radium, and many artificial radioactive isotopes are used in industrial radiography (**Table 2**). The artificial isotopes are obtained by bombarding the respective inert isotopes with an intense neutron flux in a nuclear reactor. The most commonly used radioactive sources are iridium-192 and cobalt-60. Medical therapy is also accomplished with radioisotope sources. *See* NUCLEAR MEDICINE; RADIOACTIVITY; RADIOACTIVITY AND RADIATION APPLICATIONS; RADIOISOTOPE (BIOLOGY); RADIOISOTOPE (BIOLOGY).

General principles and techniques. Some of the more important general principles and techniques common to radiography follow.

Exposure techniques. Thickness, physical density, and absorption coefficient of the materials under investigation determine the most suitable radiation energy (**Fig. 1**).

Film radiography. Since radiographs are usually interpreted from film negatives, each object under examination must be subjected to a controlled technique, and properly selected film in terms of contrast, latitude, and sensitivity. Optimum density, D, or blackening is 1.5–2.5, where $D = \log L_0/L$. The photometrically measured light intensities are L_0 before and L after the passage through the film. The normal eye can detect with certainty a minimum blackening difference between adjacent areas of 0.02. Experienced radiographers can detect density differences as small as 0.005. *See* PHOTOGRAPHY.

Table 2. Radioactive sources for industrial radiography

Element or isotope	Half-life	Steel thickness normally inspected
Radium	1580 years	3–8 in. (8–20 cm)
Cesium-137	33 years	1–4 in. (2.5–10 cm)
Cobalt-60	5.3 years	1.5–5 in. (4–13 cm)
Iridium-192	75 days	0.5–2.5 in. (1.3–6 cm)

Contrast is the difference in density produced by a change in object thickness. Latitude is the extent of object thickness that can be reproduced in the working range of the film density. Sensitivity is the smallest fractional increase in thickness detectable; in industrial radiography such changes are usually in the range of 0.5–3%. Definition is the fidelity (or sharpness) with which a radiograph delineates a sharply defined inhomogeneity.

For better interpretation of the size, shape, and location of discontinuities, two or more radiographic views of the object are obtained under predetermined angles to each other. Stereoradiography is also used for this purpose.

If the thickness of the specimen varies, complete coverage can be obtained by one of three methods: (1) various radiographs may be taken under exposure conditions as determined by the average thickness of the respective section of the object; (2) the multiple film method may be applied by placing a combination of fast and slow films in the same film holder; and (3) compensators may be used by building up the thinnest parts of the object to an average uniform thickness with similar radiographically opaque material.

Target-to-specimen distance. For conventional x-ray tubes, this should be as great as possible, consistent with the fact that radiation intensity decreases inversely as the square of the distance, in order to decrease the unsharpness of images caused by penumbra (**Fig. 2**). If F represents the size of the focal spot of the x-ray tube; d, the distance from the focal spot to the specimen; and t, the distance from defect to film; then $U = Ft/d$ is the additional width of the penumbral shadow (or unsharpness). As discussed below, however, when microfocus x-ray sources are used, the specimen may be placed closer to the x-ray source in order to magnify the image.

Focal spot or source size. This should be as small as possible, to gain emission from (as nearly as possible) a point source, since otherwise there is unsharpness from penumbra (Fig. 2). However, radiation intensity tends to decrease as the focal spot decreases.

Scattered and secondary radiation. Most of this type of radiation is produced by the impact of primary radiation on objects, such as stands, walls, or test specimens. It travels in directions other than the

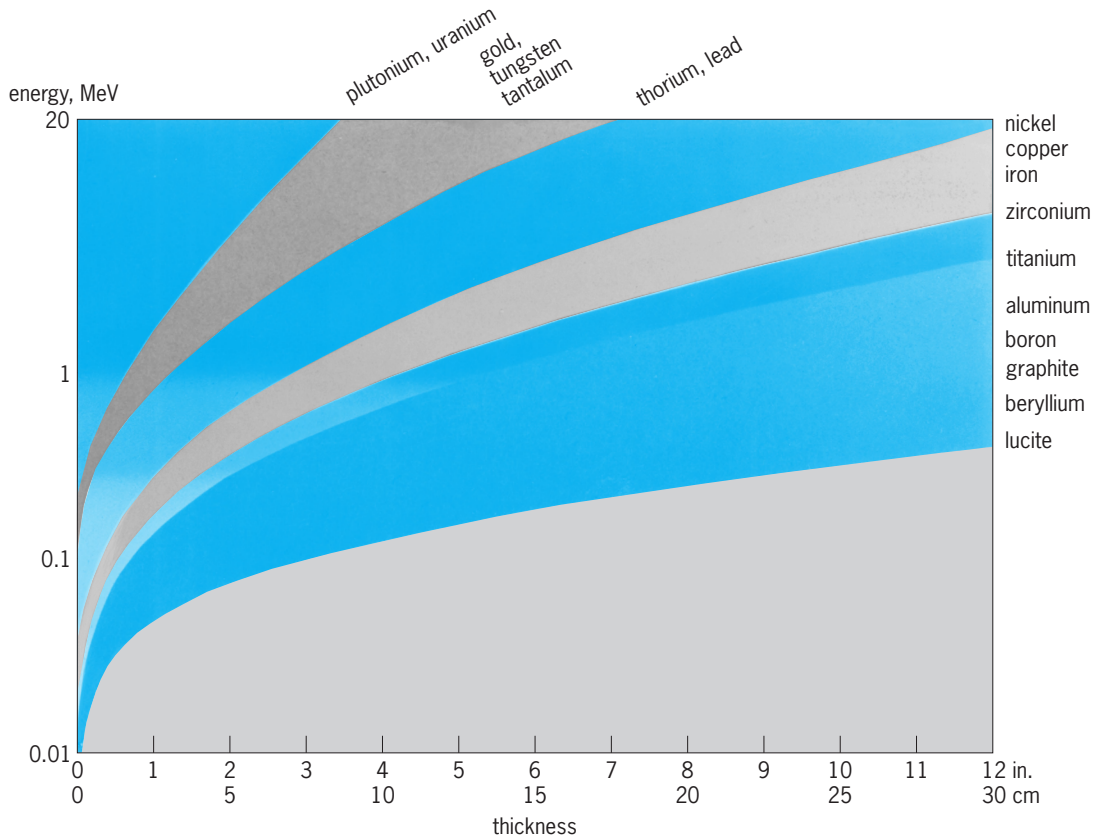

Fig. 1. Relation of the penetration characteristics of x-rays to the type and thickness of various materials.

primary radiation which originates in the x-ray tube or isotope source. It contributes to film fogging and poor radiographic definition; it can be controlled by lead screens (also simultaneously used as intensifying screens), diaphragms, lead sheets, and increased object-to-film distance. An advantage of microradiography and geometric enlargement, discussed below, is that the detection of scatter from the specimen is reduced because of the increased distance of the specimen from the detector.

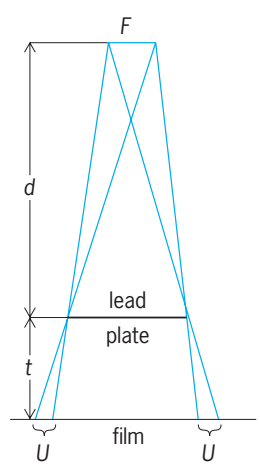

Fig. 2. Geometry of radiographic image formation. *F* = size of focal spot of x-ray tube; *d* = distance from focal spot to specimen (lead plate); *t* = distance from plate to film; *U* = unsharpness.

Radiographic inspection. In the field of medical radiography, many special techniques are used to visualize the contours of organs and soft tissues. Staining with heavily absorbing chemicals (iodine, barium, and thorium compounds) is a common procedure.

In industrial radiography the adequacy of the radiographic technique is determined by the use of penetrameters or image quality indicators. They usually are plates of the same material as the object under investigation, and their thickness is in definite proportion to the object thickness (normally 2%). They have centrally located holes with diameters of one, two, and four times their thickness and are normally placed on the source side of the object. The visualization of the smallest hole determines the radiographic sensitivity. Radiographic inspection procedures, as accepted by private and governmental industry or technical organizations, normally require the use of such penetrameters (**Fig. 3**).

Stereoscopy. Two exposures are made at tube positions roughly corresponding to interpupillary distances ($^{29}/_{\sqrt{16}}$ in. or 57 mm). The two negatives are observed in a stereoscope which fuses the two images into one with the illusion of a third dimension. This permits location of any detail of gross structure (defect in a casting, bullet in a chest, location of cancer) below the surface. Electronic, real-time stereoscopy has also been accomplished by changing the x-ray image perspective in time intervals compatible with the $^{1}/_{30}$-s television frame time. This can be

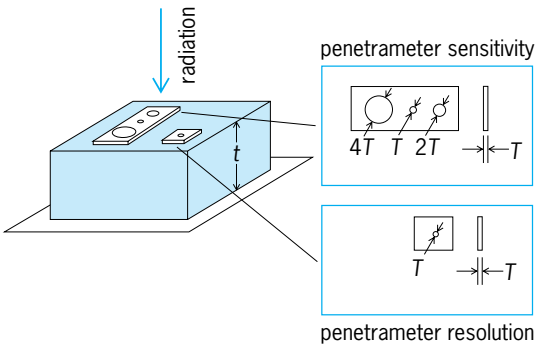

Fig. 3. Penetrameter. The windows show plan views and cross sections of the penetrameter plates. *T* is plate thickness; *t* is object thickness.

done, for example, by rotating the inspection object in the x-ray imaging beam or by electronically moving the x-ray focal spot in a microfocus x-ray tube set up to image with geometric magnification. *See* STEREOSCOPY.

Tomography. Of the three components—tube, subject, and film—two are moved during exposure synchronously so that it is possible to register the radiographic image of one plane in the object while images outside this slice have a continuous relative displacement and are blurred. A series in various planes enables three-dimensional exploration called serioscopy.

Tomographic information can also be obtained by computerized transverse axial scanning. In this technique, use is made of collimated x-ray beams, electronic detectors, and computer reconstruction of images. The radiation beam is made to pass through the object in many directions by rotating the collimated source around the object, or rotating the object. Each ray passing through the object is detected in a detector array. The attenuation in the many radiation paths is analyzed to reconstruct a density image of the inspected object. A computerized tomography image of a tree, for example, would show the growth rings as though the tree had been cut in that location to expose a visual image.

One system, which has been extensively used in medical diagnosis, utilizes 160 linear measurements across a slice of the object, these being taken at 180 one-degree intervals around the object. These 28,800 (180×160) measurements are used to reconstruct computer images. The high sensitivity of this method (0.5% variations in density) has enhanced the diagnosis of brain and head malfunctions.

Industrial applications of computerized tomography include inspections of rocket motors and missiles to detect inhomogeneities, voids, cracks, delaminations, and so forth, and of turbine blades to detect flaws and to measure internal passages and details. This use of computerized tomography as a dimensional measurement approach for internal details shows great promise. A computerized tomography system to inspect automobile engine block castings has been placed in operation for dimensional measurements of internal surfaces.

Kymography. Internal motion in an object (heartbeat, breathing of lungs) can be recorded by interposing between object and film a lead diaphragm with a 0.04-in. (1-mm) slit. The film is moved at a speed which is compatible with the motion, and records a sawtooth image as a time record of such motion.

Flash radiography. With an intense x-ray beam from field-emission pulse discharge, it is possible to make flash exposures in 2×10^{-7} s so that a rapidly moving object such as a bullet or shell appears to stand still. This has been of great significance in studies of ballistics, moving parts in engines, and other objects. A series of high-intensity exposures leads to cineradiography—moving pictures. Exposures of 0.01–0.001 s are conveniently made with rotating-target x-ray tubes, which permit passage of electron currents of 500–1000 milliamperes without damage to the target. Thus, images of heart and lungs or moving machinery may be registered between motions (**Fig. 4**).

X-ray generators (linear accelerators) capable of discharging x-ray bursts have been built. With an intensity of up to 60 amperes, at an energy of up to 20 MeV and a pulse duration of 2×10^{-7} s, such instruments are capable of penetrating very dense objects for the purpose of observing fast-moving shock waves through such specimens. *See* SHOCK WAVE.

Monochromatic radiography. By proper techniques of filtration, or reflection from crystal faces or mirrors, it is possible to produce monoenergetic beams,

Fig. 4. Chest radiograph of foundry worker made with intense beam from rotating-target x-ray tube, showing nodules in lungs which are caused by silicosis, and shadows of skeleton, heart, and stomach.

especially below 50 keV. The radiographs of thin or small objects (especially when magnified) are greatly superior to those from beams with a whole spectrum of rays each with a different absorption in the specimen.

Electronic radiography or radioscopy. The basis for this technique is the application of direct image converter tubes or the use of television pickup or electronic scanning. The resultant signals are amplified and presented for viewing on a television monitor. This inspection method permits remote or daylight viewing using x-ray energies up to 15 MeV to inspect steel or other materials. For example, some large rocket motors (about 8 ft or 2.4 m in diameter) are inspected by high-energy x-rays and television detection systems. The electronic method offers technical advantages because of the increased brightness of the radiographic image and because the electronic image signals can be processed to enhance contrast or sharpen image detail.

Digital radiologic imaging. Electronic system images can be digitized for easier processing, storage, and retrieval. Solid-state cameras, such as charge-coupled-device (CCD) cameras, can provide excellent sensitivity and images with extremely wide dynamic range (10^3–10^4), leading to x-ray image systems with 12-bit (4096 levels of gray) capability or more. *See* CHARGE-COUPLED DEVICES.

Radiographic film images can also be digitized. Large x-ray films (14 × 17 in. or 35.6 × 43.2 cm) can be scanned in less than a minute, in a 12-bit image acquisition system to handle films having a density range of 3.5 or more. Digitizing spatial resolution is typically of the order of 35–100 micrometers.

These digital radiologic methods make it possible to store, retrieve, process, and transmit x-ray images produced by film or filmlike methods and by electronic techniques. Optical disk storage of such images makes it much easier to recall and view x-ray images for both medical and industrial inspection. In medicine, a hardware-software-protocol approach known as the Picture Archiving and Communication System (PACS) has been established. PACS units collect image information from a wide variety of diagnostic techniques, including x-radiography and radioscopy, computed tomography, ultrasonic imaging, and magnetic resonance imaging (MRI), making possible improved medical diagnosis by bringing together many test records. In industry, nondestructive testing workstations are assembled. These again combine many test results, but through a neutral image format, so that new methods can easily be entered into the system. *See* COMPUTER STORAGE TECHNOLOGY; MEDICAL IMAGING; OPTICAL RECORDING.

Teleradiology permits the transmission of x-ray image information over telephone lines, via satellite, or by hardwire within an organization such as a hospital. This has been used mostly in medicine to aid in diagnosis in remote communities or to obtain assistance from specialists.

One-step radiography. As in one-step photography, this technique includes a one-step developing process through which a nearly dry, paper-supported radiographic image is obtained within about 2 min after exposure, without the use of a conventional darkroom or wet processing. This process is best suited to low-energy x-radiography but has also been used for neutron and proton radiography. The short density range of the paper is most applicable to x-radiography of high-contrast specimens, such as metal-plastic assemblies.

Paper radiography. Radiographic paper is an alternative to the more expensive x-ray film. Paper costs are typically about one-third those of film. Like one-step prints, the radiographic paper image is viewed in reflection rather than in transmission. A practical range of reflection densities is 0.4–1.0. Paper images are wet-processed in special solutions. These images are not as permanent as those made on film. After about 3 months' storage, the paper images tend to darken, but extended life can be obtained if the paper is chemically fixed like a film and dried. Paper radiographic images for aluminum in the x-ray energy range of 45–160 kV typically show penetrameter sensitivities of 2% or better.

Photothermographic radiography. This technique makes use of silver halide emulsions which can be processed by dry heat. The x-ray image is recorded with light-emitting rare-earth phosphor screens in contact with the film. The dry thermal developing process produces a silver image and stabilizes the translucent film. The dry film processing takes only 10 s and can easily be done in a small portable dark enclosure wherever there is electric power for the processor. These films can be viewed as either a reflection or transmission image. The light-sensitive emulsions display good spatial resolution, to 200 lines per inch (8 per millimeter), and hold the image in storage for periods up to 7 years.

Photostimulable phosphor radiography. Screens made of phosphors such as barium fluorohalide can store x-ray image information for extended periods before the energy is released by exposure to red light. The incident x-ray modify the europium activation ions in the phosphor crystals so that extra electrons are absorbed by the conduction band in a stable configuration. When the exposed phosphor is stimulated by red light, typically a scanning laser beam, the stored energy is released in the form of light in the ultraviolet-blue part of the spectrum. This light is detected, for example by a photomultiplier tube, with the resulting electronic signal converted to a digital image. The phosphor screen images show a very wide dynamic range, as large as 10^5, and a spatial resolution as good as 125 line pairs per inch (5 per millimeter). Images can be displayed on a television monitor or printed on a transparent base to produce a film. Advantages of this technique include excellent sensitivity, the quantitative results inherent in the digital nature of the signal, and the possibility of adjusting the final image processing to compensate for x-ray exposure variations.

Microradiography. Finer definition for thin specimens can be obtained by using microradiography.

Objects can be radiographed in close contact with fine-grain film, and the film then enlarged. Detail as small as 40 microinches (1 μm) can be brought out by this photographic-enlargement process.

An improved approach makes use of a microfocus x-ray tube, a tube having an extremely small source of x-radiation (typically of the order of 10 μm). With such a small focal spot, the inspection object can be placed close to the x-ray tube (instead of the detector), and the resultant image can be magnified with the diverging x-ray beam. This method can provide definition some 10 times smaller than the photographic method. Geometric magnification of 10 to 20× is common in industrial applications, and geometric magnification of 100× has been demonstrated.

Neutron radiography. Neutron beams, obtained from nuclear reactors, accelerators, or radioactive sources, can penetrate matter with relative ease since they are not electrically charged. Slow thermal neutrons (those in thermal equilibrium with their surroundings at or near room temperature) cover a broad band of wavelengths similar to the general, or "white," radiation from an x-ray tube. *See* THERMAL NEUTRONS.

The attenuation of neutrons by most materials is relatively small because the neutron carries no electric charge and consequently is neither attracted nor repelled by the charged particles in the nucleus, nor by the electron clouds associated with the atoms of the material through which the neutron passes. Such attenuation as does occur arises from the occasional capture of a neutron by a nucleus to form another nucleus of different mass number, or from scattering by the nucleus. Whereas in the case of x-radiation the mass absorption coefficient increases in a regular fashion with increasing atomic number of the absorber, the mass absorption coefficients (cross sections) for neutrons are randomly distributed. There are low attenuations of thermal neutrons in most of the heavy elements; therefore, large thicknesses of heavy materials can be inspected by neutrons in less time than would be required by x-radiography or gamma radiography. On the other hand, the neutron absorption coefficients of some elements with low atomic numbers are high; hydrogen, lithium, and boron are particularly attenuating. This reversal of attenuation properties between neutrons and x-rays leads to complementary properties for the two radiographic methods. With neutrons, it is possible to visualize materials such as liquids, adhesives, rubber, plastic, or explosives even when they are in metal assemblies.

Neutron radiography offers three major advantages. First is the described sensitivity to light material. Second, neutrons are sensitive to particular isotopes rather than elements. Therefore, a neutron radiograph can show differences between isotopes of the same element, for example, between normal uranium (essentially ^{238}U) and ^{235}U, the isotope of uranium that fissions more easily. A third advantage is that highly radioactive material can be radiographed without problems of fogging the radiographic film. In one method for detecting neutron images, the transfer technique, the image is detected by a thin sheet of potentially radioactive material, such as indium or dysprosium. The neutrons are transmitted through the object to the detector. The radioactive image on the detector is made visible by later placing it on film, remote from the neutron beam and object, and allowing the radioactive decay radiation to expose the film. This transfer method is extensively used to inspect highly radioactive nuclear reactor fuel.

Other methods for detecting neutron images are similar to those used for x-radiography. X-ray film, one-step film, and television methods are all used, but the conversion screens are different. For film the common screen is thin gadolinium metal; this emits internal conversion electrons as a result of gamma radiation promptly excited by neutron capture. Screens that emit light are also used for both film and television systems. These employ materials such as gadolinium, lithium, or boron. Gadolinium emits electrons, as noted above; lithium and boron emit alpha particles. In each case the radiation excites the emission of light from a phosphor material.

Neutron radiography is used to inspect nuclear reactor fuel both before and after irradiation in the reactor, explosive and adhesive-bonded components, as well as a wide variety of other materials and assemblies. The inspection of small, explosive devices, to confirm the presence of explosive, and investment-cast turbine blades to detect the residual ceramic core have been major areas of application (**Fig. 5**). (Ceramic remaining in critical turbine blades can

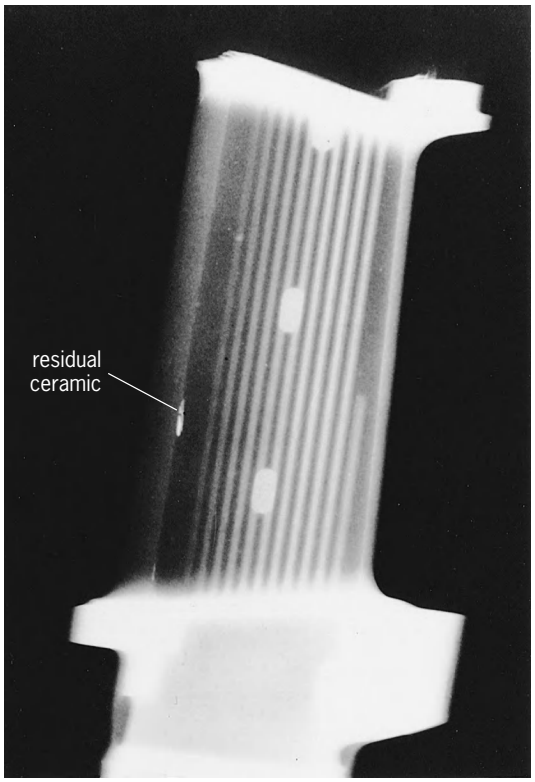

Fig. 5. Thermal neutron radiograph of an investment-cast turbine blade, showing image of residual ceramic. (*Industrial Quality, Inc.*)

lead to overheating and blade failure in hot areas of an engine because of the heat-blocking property of the residual ceramic. This residual ceramic can be removed by additional leaching.)

Proton radiography. This method employs beams of protons. A rapidly moving, high-energy proton or other charged particle moves through material with little attenuation until it slows sufficiently for the charges on the particle and the material to interact. A monoenergetic charged particle travels a well-defined distance, called the range, in a given material before it is stopped. Since most of the attenuation of the charged particles occurs near the end of the range, a very small change in material thickness will result in a large change in radiation transmission. Therefore, the sensitivity of this method to small changes in object thickness is very great, if the total path for the radiation approximates the range. This is a major advantage of proton radiography. Changes in object thickness as small as 0.05% have been imaged with one-step film.

Gerold H. Tenney; George L. Clark; Harold Berger

Bibliography. H. Berger (ed.), *Practical Applications of Neutron Radiography and Gaging*, American Society for Testing and Materials, ASTM STP 586, 1976; L. Bryant and P. McIntire (eds.), *Radiography and Radiation Testing*, vol. 3: *Nondestructive Testing Handbook*, 2d ed., 1985; D. C. De Vos, *Basic Principles of Radiographic Exposure*, 2d ed., 1995; S. J. Dwyer III and R. G. Jost (eds.), *Medical Imaging IV: PACS Systems Design and Evaluation*, SPIE vol. 1234, International Society for Optical Engineering, 1990; R. Halmshaw, *Industrial Radiology: Theory and Practice*, 2d ed., 1995; H. V. Lemke et al. (eds.), *Proceedings of International Symposium on Computer Assisted Radiology*, biannually; V. J. Orphan (ed.), *X-Ray Detector Physics and Applications*, SPIE vol. 2009, International Society for Optical Engineering, 1993; U.S. Nuclear Regulatory Commission, *Review and Evaluation of Technology, Equipment, Codes and Standards for Digitization of Industrial Radiographic Film*, 1992.

Radioimmunoassay

A general method employing the reaction of antigen with specific antibody, permitting measurement of the concentration of virtually any substance of biologic interest, often with unparalleled sensitivity. The basis of the method is summarized in the competing reactions shown in **Fig. 1**. The unknown concentration of the antigenic substance in a sample is obtained by comparing its inhibitory effect on the binding of radioactively labeled antigen to a limited amount of specific antibody with the inhibitory effect of known standards.

Method. A typical radioimmunoassay is performed by the simultaneous preparation of a series of standard and unknown mixtures in test tubes, each containing identical concentrations of labeled antigen and specific antibody, as well as variable amounts of standards or the unknown sample. It is generally advisable to have 10 to 20 standards whose concentrations cover the range to be assayed. With the use of automatic pipetting devices, the number of unknown samples that can be handled conveniently in a single assay may be 1000 or more. After an appropriate reaction time, which may be hours or days depending on the association constant for reaction with the particular antiserum, the antibody-bound (B) and free (F) fractions of the labeled antigen are separated by one of a variety of techniques, such as adsorption of free antigen to solid-phase material, precipitation of antigen-antibody complexes, and adsorption or complexing of antibody to solid-phase material.

Under the usual conditions for an assay, the antigen-antibody complexes do not spontaneously precipitate, even if they would do so at much higher concentrations of reactants. The B/F ratios in the standards are plotted as a function of the concentration of unlabeled antigen (standard curve), and the unknown concentration of antigen is determined by comparing the observed B/F ratio with the standard curve (**Fig. 2**).

Substances measured. The specific antibody is produced in a suitable animal, usually a guinea pig or a rabbit, by multiple immunizations with the substance to be measured or another substance with similar immunologic properties. Since a peptide hormone as small as vasopressin (~1000 mol wt) is antigenic in at least some animal species, and since virtually any substance of biologic interest can be rendered antigenic by coupling it chemically to a large polypeptide or protein, the method has been applied to the measurement of thousands of hormonal and nonhormonal substances.

Competitive radioassay. The radioimmunoassay principle is not limited to immune systems. It has been extended to systems in which, in place of the specific antibody, there is a specific reactor (that is, binding substance), which may be, for instance, a specific binding protein in plasma, an enzyme, or a tissue receptor site. The more general term "competitive radioassay" has been used to describe such applications. However, this term is not completely inclusive, since the principle is applicable to the use

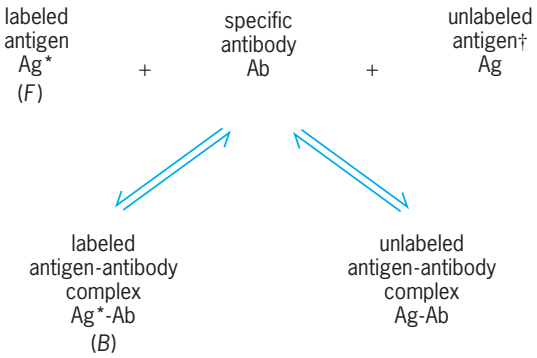

Fig. 1. Competing reactions that form the basis of radioimmunoassay. Dagger indicates antigen in known standard solutions or unknown samples. B indicates bound antigen; F, free antigen.

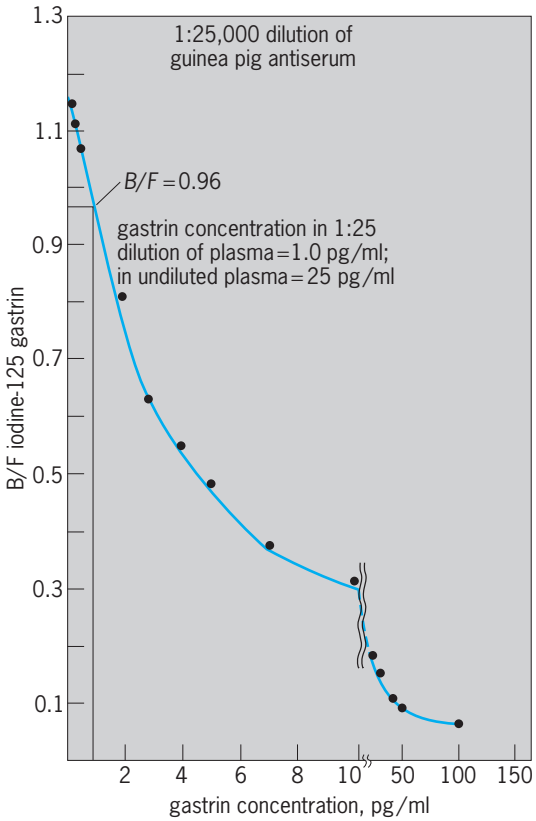

Fig. 2. Standard curve for radioimmunoassay of gastrin. Hormone concentration in unknown plasma sample is determined as shown.

of markers other than radioisotopic tracers (modified phages, enzymes, and so on).

Sensitivity considerations. Radioimmunoassay has found wide application because of its potential for high sensitivity and specificity, and because the methodology is easily adaptable to automation and computerization. Each of the diverse applications does not require all of these attributes. The requirements for sensitivity depend on the minimal quantity that must be measured. For peptide hormones in the nonstimulated state, the concentrations may be as low as 10^{-12} to 10^{-10} *M*: the concentrations of thyroidal and steroidal hormones are higher by a factor of about 10^5; the concentrations of drugs may be even higher. Theoretical analysis of the radioimmunoassay principle demonstrates that the inherent sensitivity of the system is limited by the equilibrium constant that characterizes the reaction of antigen with the predominating antibodies in the antiserum. Since animals, even of the same species, vary considerably in the concentration and potential sensitivity of the antibodies they produce, the probability of obtaining a satisfactory antiserum increases with the number of animals immunized. Since the presence of other immunological reactions does not interfere with the reaction between labeled antigen and its specific antibody, immunization with several unrelated antigens can be performed simultaneously, affording an advantage in reducing the number of animals to be immunized and bled by a factor equal to the number of antigens employed. Monoclonal antibodies, which can be produced in quantity and remain absolutely uniform, offer a great improvement in radioimmunoassay technology.

To achieve the maximal sensitivity obtainable with a particular antiserum, the chemical amount of antigen in the radioactively labeled antigen should be low enough that further reduction in its size does not increase the B/F ratio in the absence of added unlabeled antigen. This requirement determines the choice of radioisotope used for labeling. Because of the low concentrations of peptide hormones in plasma, high-specific-activity tracers are required. Iodine-125 (half-life = 60 days) is the radiochemical of choice and can readily be substituted on a tyrosyl or histidyl residue of the antigen using a variety of chemical procedures. For assays of nonpeptidal hormones and drugs, ^{3}H-labeled (half-life = 12 years) antigens prepared and purified in commercial laboratories are generally employed. Long-lived isotopes such as ^{14}C (half-life = 5800 years) are of limited value except when the concentrations of antigen to be measured are quite high, that is, concentrations of 10^{-6} *M* or higher. *See* RADIOISOTOPE (BIOLOGY).

Nonspecific interference. Radioimmunoassay differs from traditional bioassay in that it is an immunochemical procedure which is not affected by biologic variability of the test system. However, nonspecific factors can interfere with chemical reactions, and the possible presence of cross-reacting related antigens in the samples to be assayed must be considered. Some nonspecific factors that may interfere in the immune reaction include extremes of alkalinity or acidity or high concentrations of salt, anticoagulants, enzyme inhibitors, and so on. It is therefore important that the contents of the incubation tubes be identical for the standards and unknowns except for the differing amounts of antigen. Differences between concentrations of peptide hormones as determined by bioassay and immunoassay may be caused by the presence of precursor hormones (prohormones), molecular fragments, or related but not identical hormones. Requirements for specificity merits consideration in assays for other substances as well. For instance, if the clinical problem relates to the toxicity of a particular drug, the question of whether the assay measures the biologically active form of the drug is relevant. If the question relates simply to whether or not a drug of abuse (morphine, LSD, and so forth) has been used, the presence of metabolites or the exact form of the drug may be irrelevant. *See* BIOASSAY.

Applications. The radioimmunoassay principle has found wide application in the measurement of a large and diverse group of substances in a variety of problems of clinical and biological interest. It is therefore not unexpected that there are differences in the specific methods employed for the assay of a particular substance. It seems that virtually any substance of biologic interest can be measured using radioimmunoassay, modified according to the characteristics of the particular substance. In recent years there has been a tendency for laboratories to try as far as

possible to replace radioimmunoassays with nonisotopic techniques. Radioisotopes are highly regulated substances and are increasingly expensive and difficult to obtain; disposal is cumbersome and represents a potential hazard to the environment. The radioisotope is replaced by a fluorescent marker, for example, in the fluorescence polarization immunoassay that many large clinical laboratories prefer. Enzyme-linked immunosorbent assays (ELISA) have also replaced many radioimmunoassays. Nevertheless, radioimmunoassay remains one of the few techniques available to measure amounts of small (monovalent) molecules in the usual research laboratory. *See* ANTIBODY; ANTIGEN; IMMUNOASSAY; CLINICAL IMMUNOLOGY. Rosalyn S. Yalow; J. John Cohen

Bibliography. S. A. Berson and R. S. Yalow (eds.), *Methods in Investigative and Diagnostic Endocrinology*, vol. 2: *Peptide Hormones*, 1974; D. S. Hage, Immunoassays, *Anal. Chem.*, 71:294R–304R, 1999; F. Laszlo and T. Janaky (eds.), *Clinical Application of Radioimmunoassay*, 1992; R. Yalow, Nobel Prize Lecture, Dec. 8, 1977: *Radioimmunoassay: A Probe for Fine Structure of Biological Systems*, in T. Charc, *An Introduction to Radioimmunoassay and Related Techniques*, 4th ed., 1991.

Radioisotope

A radioactive isotope (as distinguished from a stable isotope) of an element. Atomic nuclei are of two types, unstable and stable. Those in the former category are said to be radioactive and eventually are transformed, by radioactive decay, into the latter. One of the three types of particles or radiation (alpha particles, beta particles, and gamma rays) is emitted during each stage of the decay. *See* ISOTOPE; RADIOACTIVITY.

The term radioisotope is also loosely used to refer to any radioactive atomic species. Whereas approximately a dozen radioisotopes are found in nature in appreciable amounts, hundreds of different radioisotopes have been artificially produced by bombarding stable nuclei with various atomic projectiles.

The production, separation, purification, and shipment of radioisotopes are usually considered primary production; secondary processing of the primary materials into compounds, special shapes, radiation sources, and medical radiopharmaceuticals is usually required before use in hundreds of applications that have been developed. Radioisotopes are produced by irradiation of the elements with high-energy particles (protons, deuterons) in accelerators and in the high-intensity neutron fluxes of relatively small research-type reactors.

A few radioactive elements occur in nature, for example, uranium, radium, and thorium, which were produced when the Earth was formed. Some of these naturally occurring radioisotopes (primary parents of series) have half-lives (time in which one-half of the atoms decay) greater than 10^8 years, and therefore have not had time to disappear. During the early years of work with radioactive materials, only natural radioactivity was available; The first radioactive tracer experiment was in 1934 with radiolead (^{212}Pb) obtained from thorium decay products.

The first artificially induced radioactivity was also produced in 1934 by irradiating aluminum foil with alpha particles, thus initiating the use of radioactivity on a wide scale in scientific work. A very wide variety of radioisotopes are produced in particle accelerators, such as the cyclotron. Charged particles, such as deuterons (D^+) and protons (H^+), are accelerated to great speeds by high-voltage electrical fields and allowed to strike targets in which nuclear reactions take place; for example, proton in, neutron out (p,n), increasing the target-atom atomic number by one without changing the atomic mass; and deuteron in, proton out (d,p), increasing the atomic mass by one without changing the atomic number. The target elements become radioactive because the nuclei of the atoms are unbalanced, having an excess or deficit of neutrons or protons. Although the particle-accelerating machines are most versatile in producing radioisotopes, the amount of radioactive material that can be produced is relatively smaller than that made in a nuclear reactor [less than curie amounts; a curie (abbreviated Ci) is that quantity of a radioisotope required to supply 3.7×10^{10} disintegrations per second or 3.7×10^{10} becquerels (Bq)]. For large-scale production, nuclear reactors with neutron fluxes of 1×10^{10} to 5×10^{15} neutrons per square centimeter per second are required. *See* NUCLEAR REACTION; NUCLEAR REACTOR; PARTICLE ACCELERATOR; REACTOR PHYSICS; UNITS OF MEASUREMENT.

Production. Radioisotope production proceeds in a series of steps. First, target materials containing the desired element are selected in the most stable heat- and radiation-resistant form (for example, metal oxides) and the highest chemical purity, determined by chemical, spectrographic, or neutron activation analysis. The target material is encapsulated in a low-neutron-absorbing material (such as aluminum or magnesium) and sealed by heli-arc welding or an equivalent process, and the capsule is tested for leakage. Next, the capsule is irradiated for the optimum time period, determined by the half-life of the desired radioisotope, the absorption cross section of the target atom, and the growth of possible undesired impurities. The irradiated capsule is then removed to a remote-control hot cell for chemical processing, and the preparation is analyzed and assayed by the most advanced methods that are amenable to remote-control operations. Aliquots are dispensed as needed for users or secondary processors, and the radioisotope is packaged, inspected for radioactive contamination, and shipped by common carriers. Shipment may be subject to local, state, federal, or international shipping regulations, depending upon the nature, size, and destination of the radioactive material.

Reactor production. Most radioisotopes produced in quantity are made in the nuclear reactor by one of four reactions (**Table 1**): neutron-gamma, neutron-proton, neutron-alpha, and neutron-fission. The neutron-gamma reaction is the most common,

TABLE 1. Reactor radioisotope production reactions

Reaction	Examples
(1) Neutron-gamma (n,γ)	${}^{59}_{27}\mathrm{Co} + {}_0n^1 \rightarrow {}^{60}_{27}\mathrm{Co} + \gamma$
(2) Neutron-proton (n,p)	${}^{32}_{16}\mathrm{S} + {}_0n^1 \rightarrow {}^{32}_{15}\mathrm{P} + {}^{1}_{1}\mathrm{p}$
(3) Neutron-alpha (n,α)	${}^{35}_{17}\mathrm{Cl} + {}_0n^1 \rightarrow {}^{32}_{15}\mathrm{P} + {}^{4}_{2}\gamma$
(4) Neutron-fission (n,f)	${}^{235}_{92}\mathrm{U} + {}_0n^1 \rightarrow {}^{131}_{52}\mathrm{Te} + {}^{102}_{42}\mathrm{Mo} + \sim 2{}_0n^1$

because many elements have a good neutron-capture cross section (relative ability for capturing neutrons). By simple neutron capture, important radioisotopes, such as ^{24}Na, ^{59}Fe, ^{60}Co, and ^{198}Au, are made. The procedure is quite simple in this case. Highly purified materials (to prevent contamination by neutron-capturing impurities) in amounts ranging from a few milligrams to several hundred grams are sealed in pure aluminum cans and placed in the reactor. In the higher-flux reactors (flux greater than 10^{14} neutrons per square centimeter per second) the highly purified target material is sealed in quartz ampules which are in turn weld-sealed in small aluminum tubes, well designed for heat removal (**Fig. 1**). Pneumatic or hydraulic tubes are used for passing samples in and out of the reactor. Also, reactors may have special magnesium, beryllium, or aluminum holders for insertion into the reactor lattice. Care must be taken to avoid putting maerials that decompose easily, such as organic compounds, into the reactor because gas pressure may be produced in the container. In general, thermally stable materials are also fairly stable under neutron irradiation. Metals or stable oxides of elements are usually used as target materials.

The production rate for a radioisotope in a reactor depends upon the neutron flux, amount of target atom, the half-life of the radioisotope, and the neutron-activation cross section of the target atom. The formula written below can be used for calcu-

$$A = N\phi\sigma(1 - e^{-\lambda t})$$

lating the amount of radioisotope produced, where A is the activity in disintegrations per second, N is the number of target atoms, ϕ is the neutron flux in neutrons per square centimeter per second, σ is the activation cross section for the reaction in square centimeters per atom, and the expression $(1 - e^{-\lambda t})$ is the saturation factor. The irradiation time t and the decay constant λ are in compatible units (λ = 0.693/half-life). The activity A, in disintegrations per second, can be converted to millicuries by dividing by 3.70×10^7 disintegrations per second per millicurie.

Fig. 1. Aluminum capsule target for high-neutron-flux reactor.

Target materials of some elements can be obtained quite pure; when such elements have only one isotope that has a high activation cross section, radioisotopes are easy to produce and little chemical purification is required after irradiation. In other cases there is multiple production in the target material, sometimes four or five different radioactive species in one target, which must be chemically separated from the main product. Virtually every known chemical-separation procedure is used in this kind of work, which usually must be done by remote methods because of the high radiation levels involved. *See* NUCLEAR CHEMISTRY.

When radioisotopes are produced by neutron-proton or neutron-alpha reactions or by a neutron-gamma reaction followed by beta decay, the radioelement is of a different chemical species from the target element and can be separated from it chemically to produce carrier-free radioisotopes.

Very often concentrations of radioelements are too low for them to be precipitated directly, so they are carried on the surface of a flocculent precipitate, such as $Fe(OH)_3$; similarly, others are coprecipitated, where isomorphous compounds are brought down together, for example, $^{140}BaSO_4$ with $PbSO_4$. Many methods are more adaptable to work with low concentrations of material or amenable to remote operation, such as solvent extraction and ion exchange. A preferred purification method is gasification, because if the radioisotope goes through various stages in the gaseous phase, very high purity usually results. Distillation is sometimes used, as in purification of ^{131}I (distilled as the element) and $^{103\text{-}106}$Ru (distilled as the tetroxide). Ion exchange is practically the only method for fractionating the rare earths in the low concentrations usually found in radioisotope work.

Fission-product separation. The fission products, fragments of the fissioned uranium or plutonium atom, are an extremely important group of radioisotopes, ranging from zinc (atomic number 30) to samarium (atomic number 62). The fission-product fragments occur in two groups: light atoms with atomic masses between 72 and 110 units and heavy atoms with masses between 110 and 162 units. Those fission products with high yields occur at one of the peaks, and the companion fission fragment occurs with the same high yield on the other peak; for example, mass 140 (Xe through Ba) has companion fragments of about mass 95 (Y through Mo). Fission-product yields are given as the percent of fissions that result in fragments of a certain mass; there are two fragments for each fission, so the overall yield arithmetically totals up to 200%. *See* NUCLEAR FISSION.

The fission products can be separated and purified for use as radioisotopes and fall into two groups (**Table 2**).

With the exception of ^{131}I, the short-lived fission products were first separated almost entirely for research purposes. The extraordinary growth in the use of ^{99}Mo to produce the daughter ^{99m}Tc for medical use has made ^{99}Mo the most important commercial radioisotope. The parent radioisotope, 67-h ^{99}Mo, is also made by neutron-gamma

TABLE 2. Isotopes ordinarily separated from fission products

Radioisotope	Half-life	Fission yield, %
Principal long-lived fission products*		
^{85}Kr	10.27	1.3
^{90}Sr	28	5.8
^{99}Tc	212,000	6.1
^{106}Ru	1	0.4
^{137}Cs	30	5.9
^{144}Ce	0.8	5.7
^{147}Pm	2.6	2.4
Principal short-lived fission products†		
^{89}Sr	54	4.8
^{91}Y	59.5	5.8
^{95}Zr-^{95}Nb	65–35	6.3
^{99}Mo	2.8	6.2
^{103}Ru	39.8	3.0
^{131}I	8.08	2.9
^{133}Xe	5.3	6.5
^{140}Ba	12.8	6.4
^{141}Ce	32.5	6.0
^{143}Pr	13.9	6.0
^{147}Nd	11.1	2.4

*Half-life in years.
†Half-life in days.

reaction on natural molybdenum, or, in some cases, enriched ^{98}Mo target material is irradiated. The very high specific activity fission product ^{99}Mo does have some advantages for making ^{99}Tc generators. However, because of the high-radiation-level, highly technical separation process, very few commercial radioisotope producers have been interested in taking it over.

The short-lived fission products are made by irradiating a small piece of ^{235}U-Al alloy in the reactor for a few weeks. The target is promptly dissolved in nitric acid upon discharge from the reactor, the radioiodine and xenon separated from the gases are discharged during dissolution, and the uranium-aluminum nitric acid solution is processed by solvent extraction and ion exchange, principally to obtain ^{99}Mo, ^{95}Zr-^{95}Nb, ^{89}Sr, and the rare earths.

The long-lived fission products, such as ^{90}Sr and ^{137}Cs, are separated from long-cooled waste from the processing of spent fuel from power reactors. Krypton-85 is also separated from the waste gases at the fuel reprocessing plants. The long-lived fission products have potential as sources of radiation and heat.

Accelerator production. The production of radioisotopes using high-energy particles is more versatile than reactor production, especially for radioisotopes that are neutron-deficient (most reactor radioisotopes have an excess of neutrons in the nucleus). Many of the neutron-deficient radioisotopes decay by *K*-shell electron capture with the emission of soft x-rays, or by positron emission with accompanying annihilation radiation (the collision of positive and negative electrons, e^- and e^+, with each electron mass converted to 0.51-MeV gamma radiation). Particularly in nuclear medicine, short-lived radioisotopes are needed that have little beta radiation and thus cause minimum tissue damage, or radioisotopes with directional annihilation radiation to aid location in scanning.

Small cyclotrons have been developed that will accelerate protons, deuterons, and alpha particles. A few of these are now devoted to radioisotope production.

Attachments have been put onto a few large research cyclotrons to get protons of extremely high energy to produce radioisotopes by spallation. This occurs when the particle hitting the target atom is so energetic that "pieces fly off." Many of the particles making up the nucleus are lost and a variety of radioisotopes of lower atomic number and weight are produced. *See* SPALLATION REACTION.

Shipping. The packaging and shipment of radioisotopes present some unusual problems—the penetrating radiation must be shielded and leakage must be prevented because of the extreme toxicity of many radioisotopes. Packaging must be done by remote control. Most radioisotopes are dispensed as water solutions by pipetting from glass storage bottles to glass shipping bottles. Tempered-glass bottles are commonly used, with closures lined with polyethylene inserts. For some special preparations, such as carrier-free ^{45}Ca or ^{32}P, polyethylene storage and shipping bottles are used to cut down losses by

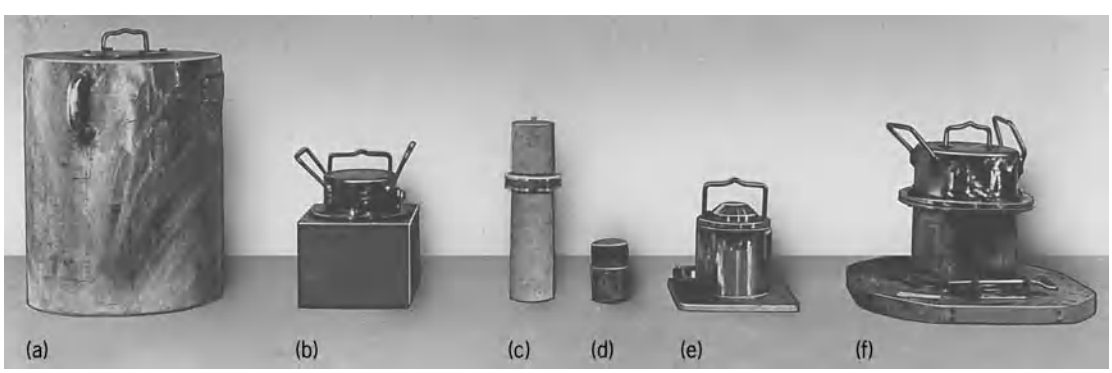

Fig. 2. Radioisotope shipping containers. (*a*) Lead shield, stainless-steel encased for gamma sources, such as ^{60}Co. (*b*) Lead shield, stainless-steel encased, in wooden block for pile units and small sources. (*c*) Gas shipping container with internal sealed cylinder for ^{85}Kr and tritium gas. (*d*) Nonreturnable lead shield for processed radioisotope bottles. (*e*) Solid stainless-steel container for bottles of processed ^{90}Sr. (*f*) Lead shield, stainless-steel encased, on a shipping pallet for distributing weight in aircraft, for pile units and small gamma sources. (*Oak Ridge National Laboratory*)

Fig. 3. Nonreturnable shipping container for radioisotopes. (*Oak Ridge National Laboratory*)

adsorption on the walls of the container. High standards of cleanliness must be maintained, although no attempt is made to ensure sterility of solutions except for the special medical radioisotope preparations shipped by pharmaceutical suppliers.

A few preparations have sufficient mass per unit of radioactivity and weak enough radioactivity to permit handling and packaging as dry solids. An example is ^{14}C, which is dispensed as barium carbonate. Materials that are not chemically processed, such as reactor irradiation units or metallic cobalt, are sent out as solids in metal containers.

Two general kinds of shipping container are used; returnable, on which a deposit is paid; and nonreturnable, which the receiver may keep. Returnable containers are usually more massive and expensive, especially those used for shipping large amounts of ^{60}Co or fission products. Pure beta emitters, such as ^{32}P, require little shielding other than the absorbent packing material. Gases are shipped in special glass or metal containers (**Figs. 2** and **3**).

Shipment of radioactive materials by rail or truck is covered by federal government regulations (Department of Transportation) and international regulations, coordinated through the International Atomic Energy Agency. *See* ISOTOPE SEPARATION; RADIOACTIVE TRACER. Arthur F. Rupp

Bibliography. *Accelerator-Produced Nuclides*, Brookhaven National Laboratory, BNL-50448, 3 vols., 1975, 1978, 1983; F. Helus (ed.), *Radionuclides Production*, 2 vols., 1983; International Atomic Energy Agency, *Radioisotope Production and Quality Control*, Tech. Rep. Ser. 128, 1971; International Atomic Energy Agency, *Radiopharmaceuticals and Labelled Compounds*, vol. 1, IAEA-SM-171/35, 1973; F. F. Knapp and T. A. Butler (eds.), *Radionuclide Generators: New Systems for Nuclear Medicine Applications*, 1984; U.S. Department of Energy, *Isotope Production and Distribution*, 1994.

Radioisotope (biology)

A radioactive isotope used in studying living systems. This article discusses the use of radioactive isotopes in the investigation of metabolic processes.

The usefulness of radioisotopes as tracers arises chiefly from three properties: (1) At the molecular level the physical and chemical behavior of a radioisotope is practically identical with that of the stable isotopes of the same element. (2) Radioisotopes are detectable in extremely minute concentrations. (3) Analysis for radioisotopic content often can be achieved without alteration of the sample or system. In some applications radioisotopes are not essential but are used because of the advantages of sensitivity or convenience. In other applications, however, principally those in which reaction rates and transfer rates are studied, isotopes, particularly radioisotopes, have unique advantages as tracers, and their use in studies of the kinetics of steady-state systems, self-diffusion, and metabolic pathways has opened numerous fields previously thought to be inaccessible. *See* ISOTOPE; RADIOISOTOPE.

Methods. Radioisotope methods include preparing the labeled material, introducing it into the system to be studied, detecting its presence, and interpreting the results.

Preparation of labeled compounds is one of the most important and most difficult steps in tracer methodology. In synthesis with radioactive reagents, the chemical procedures used must often be modified for remote handling, and other special precautions must be taken to prevent contamination of laboratories and personnel with radioactive material and to minimize irradiation of the personnel. Containment, shielding, and distance are the main items in protecting personnel from radiation. Some compounds labeled by biosynthesis, derived from bacteria or plants grown on media containing radioactive substrates, are commercially available. Before a labeled product is used, it should be tested for radiochemical purity, which means that ordinarily there should be only one radioactive species present and that all of it should be in the same chemical state. Labeling of an element present in two valence states has caused confusion. The chemical stability of the label (that is, the radioactive element) should also be established. It is unwise to use material labeled in easily exchangeable positions if the fate of the material itself (rather than the exchange process) is the

point of interest. *See* RADIOCHEMICAL LABORATORY; VALENCE.

The amount of isotope to be used and the path by which it is introduced into the system are governed by many factors. Sufficient tracer to be detectable must be used, but the amounts of material which are introduced must be small enough not to disturb the system by their mass, pharmacological effects, or the effects of radiation. The mass of 1 curie, the unit of disintegration rate, depends inversely upon the half-life and directly upon the atomic weight of the particular radioisotope; it is 1 gram for ^{226}Ra (half-life 1620 years), but only 8 micrograms for ^{131}I (half-life 8.0 days). In tracer experiments with small animals, microcurie quantities are usually adequate. It is sometimes necessary to choose between administering the isotope in a single injection or using a continuous perfusion, for example, to maintain a constant concentration in the arterial blood. Usually the single injection is easier to achieve, and the resulting data are easier to analyze.

There are many methods for detecting the presence of radioactive material. The Geiger counter has largely been displaced by thallium-activated sodium iodide scintillation crystals for counting gamma rays, but Geiger counters and proportional counters are still useful for counting alpha and beta particles. A well-type scintillation counter may have more than 25% efficiency for ^{131}I. In a scintillation crystal the passage of a gamma ray through the crystal produces a flash of light which is detected by a photomultiplier tube connected to the crystal. The signal from the photomultiplier tube may be fed into a scaler or rate meter, as with the signal from Geiger tubes, or after amplification it may be used as the input signal for a pulse-height analyzer in a method called gamma-ray spectroscopy. The pulse-height analyzer makes it possible to record pulses of a single height, corresponding to single gamma-ray energies, so that one isotope may be counted selectively in the presence of others. With multichannel pulse-height analyzers, simultaneous determinations may be made of a number of gamma-emitting isotopes in a mixture. Beta-particle spectroscopy is also possible but much more difficult both to achieve and to interpret, in part because beta particles are emitted with a spectrum of energies, whereas gamma rays are typically monoenergetic. For very low-energy β^- particles (negative electrons), particularly those emitted by ^{14}C and tritium (^{3}H), liquid scintillation counting is useful. In this method a solution of the sample and a phosphor take the place of the scintillation crystal. Low-energy β^- emitters can also be counted by introduction, in the gas phase, directly into proportional counters. In histological and cytological studies the method of autoradiography, in which photographic film is exposed through contact with the specimen, is very useful. The autoradiographic method is also used extensively in conjunction with paper or column chromatography, particularly in studies of metabolic pathways. *See* AUTORADIOGRAPHY; CHROMATOGRAPHY; GEIGER-MÜLLER COUNTER; PARTICLE DETECTOR; PHOTOMULTIPLIER; SPECTROSCOPY.

In many applications the interpretation of the data obtained with radioisotopes is quite simple. The volume of dilution and total exchangeable mass, for example, are inversely related to the final concentration. Following a metabolic pathway or determining the uptake of radioisotopes by various organs or tissues is analogous to following a flock of sheep by belling the leader. In studies of rate processes, however, particularly when material is moving in opposing directions simultaneously, extensive use of mathematics is needed for correct interpretation of the data. The mathematical basis for such interpretation is outlined in the following paragraphs.

Interpretation of tracer kinetics. For mathematical analysis it is convenient to designate the distinct phases or volumes of a system as compartments. In steady-state systems it is assumed that, except for the tracer, the content of a compartment does not change; the rates of flow of material into and out of the compartment are equal. It is customary, but not always realistic, to assume that within a given compartment mixing of the labeled and unlabeled material is instantaneous and complete; the specific activity (the ratio of the amount of the labeled form of a substance to the total substance present) within a compartment is assumed to be uniform in space, although it may be a function of time.

Given a compartmental model regarded as representing the system of interest (often greatly simplified), the expected behavior of the tracer is described with a set of differential equations in terms of the invariants of the system. The solutions of these equations contain constants which would be determined experimentally from observations on an ideal system. Although prediction of tracer behavior from information about the system has many applications, the reverse problem, that of using the experimentally determinable constants describing the behavior of the tracer to deduce the invariants characteristic of the system, is the usual goal of the tracer experiment.

The passage of a tracer through a chain of steady-state compartments with one-way flow, as in the system $\rightarrow$ A $\rightarrow$ B $\rightarrow$ C$\rightarrow$ D $\rightarrow$, is described with equations identical in form to the Bateman equations, which describe radioisotopic decay chains. The solution equation for the nth compartment will include n exponential terms, with each exponential constant being equal to one of the n turnover-rate constants (the ratio of the rate of entrance or exit of the substance of interest to the amount present in a given compartment). For example, the disappearance of a tracer from the simplest model, $\rightarrow$ A $\rightarrow$, assuming that, after time zero, no tracer enters, is described by Eq. (1), where x is the specific activity in the com-

$$x = x_0 e^{-\lambda t} \tag{1}$$

partment; x_0 is the value at time zero; t is time; e is the base of the natural logarithms, 2.718 . . . ; and λ is the fractional disappearance constant. If, instead, there is no tracer in the compartment at time zero but there is a constant specific activity x_∞ in the inflow from time zero on, the buildup of activity in the

compartment is described by Eq. (2).

$$x = x_\infty (1 - e^{-\lambda t}) \tag{2}$$

In both cases λ is equal to the turnover-rate constant k. Because all the λ's appear in the equation for the last compartment of a chain, all the turnover-rate constants may, in principle, be determined by sampling the last compartment only, but this is insufficient for determining the sequence in the chain. *See* NUCLEAR REACTION; RADIOACTIVITY.

When there is flow in both directions, the simple relationship between the k's and λ's is lost. Equations (3) and (4) are differential equations for the

$$\frac{dx_A}{dt} = -k_A(x_A - x_B) \tag{3}$$

$$\frac{dx_B}{dt} = k_B(x_A - x_B) \tag{4}$$

two-compartment, steady-state, closed system $A \rightleftharpoons B$, for which, if at time zero $x_A = x_{A0}$ and $x_B = 0$, the solutions are expressed by Eqs. (5) and (6), where

$$\frac{x_A}{x_{A_0}} = C_E + C_1 e^{-\lambda t} \tag{5}$$

$$\frac{x_B}{x_{A_0}} = C_E - C_E e^{-\lambda t} \tag{6}$$

the C's are constants determined by the amounts of material present. C_E is the equilibrium value, and C_1 is the difference between the initial value and the equilibrium value of the material present. In this example $\lambda = k_A + k_B$ rather than being equal to one or the other turnover-rate constant. Calculation of the individual turnover-rate constants from the experimental data as represented by Eqs. (5) and (6) can, however, be achieved through use of Eqs. (7).

$$k_{AB} = C_1\lambda \qquad k_{BA} = C_E\lambda \tag{7}$$

In the general case of N compartments in the steady state, Eq. (8) describes the behavior of the

$$\frac{dx_J}{dt} = -K_J x_J + \sum_{R=A}^{R=N} k_{JR} x_R \qquad R \neq J \tag{8}$$

tracer in the Jth compartment. Here k_{JR} is the fractional turnover-rate constant resulting from exchange with the Rth compartment, and K_J is the total turnover-rate constant for compartment J. The solutions have the general form of Eq. (9), with $N\lambda$'s

$$x_J = C_{J1}e^{-\lambda_1 t} + C_{J2}e^{-\lambda_2 t}, \ldots, C_{JN}e^{-\lambda_N t} + C_E \tag{9}$$

for open systems and $(N - 1)\lambda$'s for closed systems. Interpretation of experimental data in terms of the size of and rates of transfer among the various compartments involves fitting Eq. (9) to the data and using the constants (the C's and λ's) to calculate the k's and other characteristics of the system. Complete sets of formulas for such calculations for the three-compartment systems are available. Because of the unwieldy equations encountered, explicit formulas for deriving the k's from the C's and λ's have

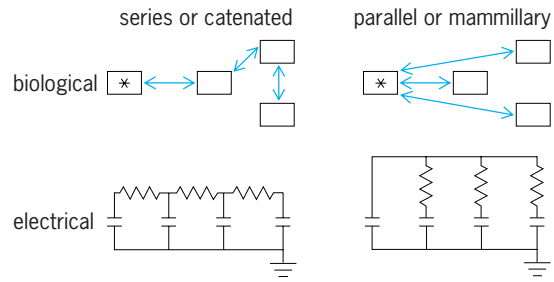

Fig. 1. Simple electrical analog scheme for a four-compartment steady-state system. The asterisk indicates the injected compartment in the biological system, corresponding to an initial charge on the left-hand condenser in the electrical system.

not been published for systems involving more than three compartments, but numerical examples can be worked with matrix algebra using the general relationship $|k| = |C\| - \lambda \, \|C|^{-1}$, where $|k|$ is a matrix of the K's and k's in Eq. (8), $|C|$ is a matrix of the coefficients in Eq. (9), $|-\lambda|$ is a diagonal matrix of the $-\lambda$'s, and $|C|^{-1}$ is the inverse of $|C|$. This relationship is derived by equating the derivatives of Eq. (9) to those of Eq. (8).

Complicated systems can also be analyzed with electronic analog and digital computers, which eliminate most of the intermediate mathematics involved in translating descriptions of the data into descriptions of the systems. Analog computers have the advantage of directness and simplicity of programming, whereas digital computers have the advantage of numerical accuracy.

Figure 1 depicts an electrical analog for a four-compartment steady-state system which uses a direct analog between the electrical and biological components. The equations which describe the voltage changes in the electrical system are identical in form with those which describe the net tracer movement in the biological system. Adjustment of the electrical parameters therefore provides a direct method for determination of the values of the biological system components. For non-steady-state systems, operational amplifiers and an electrical circuit analogous to the differential equation of the biological model may be used to meet the requirement of unequal flow rates in the two opposing directions. *See* ANALOG COMPUTER; DIGITAL COMPUTER.

Applications. In chemistry the tracer method has been applied to isotopic exchange reactions, chemical kinetics, structural chemistry, self-diffusion studies, and analytical chemistry. Self-diffusion (really isotopic interdiffusion) has been studied in solids, as well as in liquids and gases. In activation analysis a sample is subjected to nuclear bombardment in a reactor or accelerator, and the induced activity provides an extremely sensitive method for assay of the amounts of certain trace constituents.

In biology one of the outstanding achievements in which radioisotopes have played a role has been the use of carbon-14 in the elucidation of the metabolic path of carbon in photosynthesis. The products produced in the first few seconds following exposure

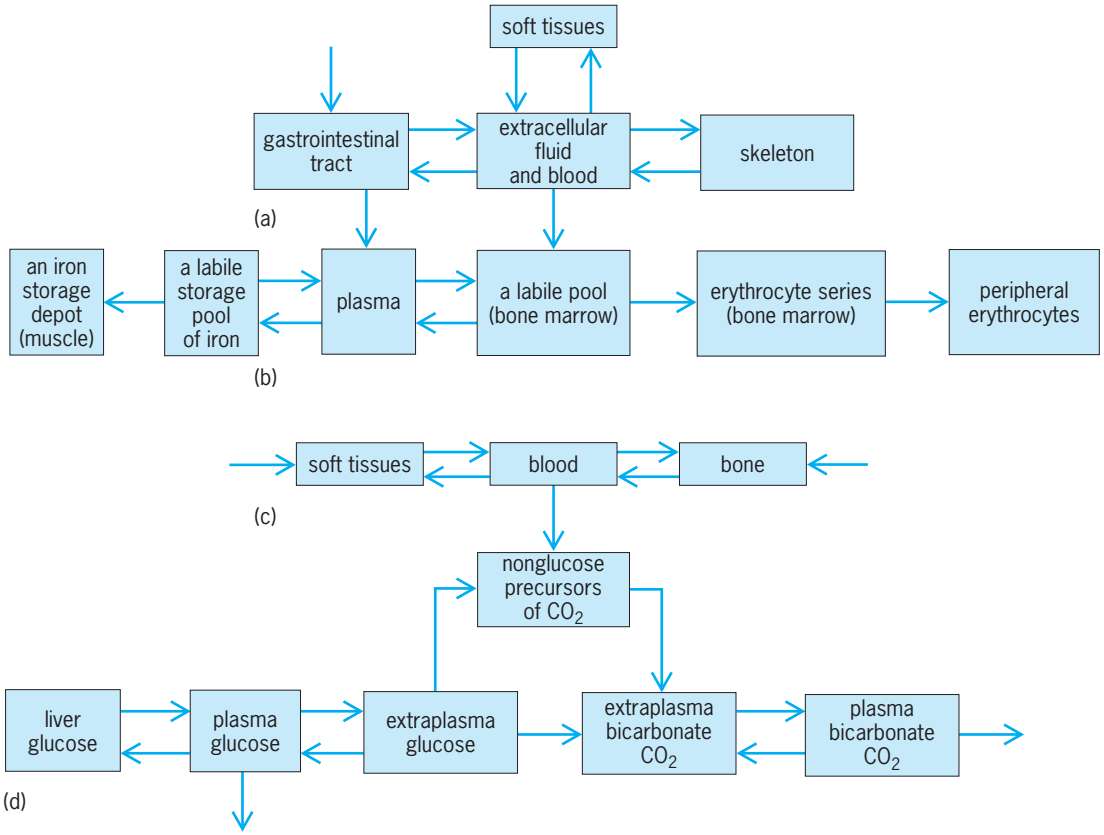

Fig. 2. Models used in analyses of metabolizing systems. (*a*) Calcium. (*b*) Iron. (*c*) Carbonate. (*d*) Glucose. In *b*, the iron returns from the peripheral erythrocytes to the plasma very slowly.

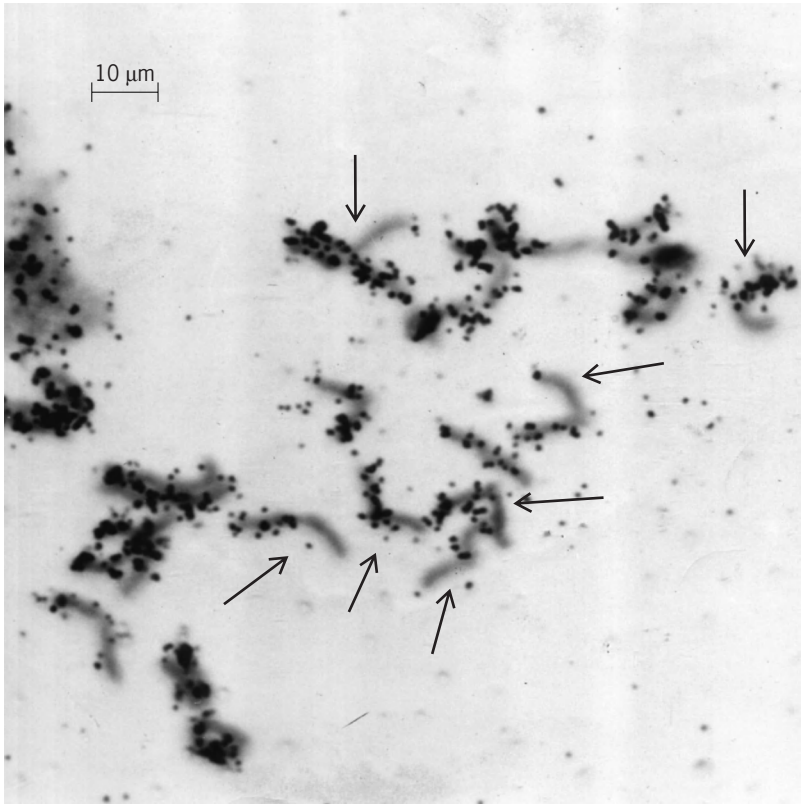

Fig. 3. Photomicrograph of an autoradiograph of *Vicia faba* chromosomes in metaphase labeled with tritiated thymidine. In the chromosome pairs indicated by the arrows, one chromosome is labeled but the other is not. This has been interpreted as evidence that the newly formed deoxyribonucleic acid (DNA) does not exchange with the previously existing DNA. Feulgen stain was used. (*Courtesy of W. L. Hughes*)

to light have been identified by combinations of paper chromatography and autoradiography. Somewhat similarly, the extrathyroidal metabolism of iodine, the path of iodine in the thyroid gland, and other problems of intermediary metabolism have been studied intensively. The concept of the dynamic state of cell constituents is largely attributable to discoveries made with isotopic tracers. At one time it was thought that concentration gradients across cell membranes depended upon their being impermeable, but the use of isotopes has refuted this hypothesis by proving that in many such cases the substances involved are normally transported in both directions across the membrane.

In physiology, radioisotopes have been used in a wide variety of permeability, absorption, and distribution studies. Representative model systems are depicted in **Fig. 2**. As an example, calcium metabolism is described here in some detail. Calcium is absorbed from the intestine into the blood, from which it is distributed to the soft tissues and skeleton or is excreted via the feces, urine, and perspiration. More than 80% of ^{45}Ca injected intravenously leaves the circulation within 5 min, but the serum specific activity is still decreasing 2 months after injection. The initial rapid drop in specific activity is explained by dilution of the ^{45}Ca in the relatively large pool of rapidly exchangeable bone calcium. Although less than 2% of the bone Ca is available for exchange, this is 10–100 times the amount in the serum. Bone also takes up ^{45}Ca by accretion into the

"nonexchangeable" pool. In growing boys the rate of bone-salt formation was found to be two to three times the rate of calcium intake from the intestine, the deficit being made up by resorption. The combination of excretion and slow turnover of nonexchangeable bone accounts for the slow components of the serum-specific activity curve. In adults the processes of accretion and resorption continue at slower rates and are balanced. The distinction between accretion and exchange is possible only with isotopic tracers. Another distinction made possible with ^{45}Ca is the finding, made by comparing fecal contents following ^{45}Ca administration by mouth and by vein, that about 10% of fecal calcium is of endogenous origin, whereas the rest represents a lack of absorption. Comparison of ^{45}Ca absorption and excretion in a boy with that in an adult suggests that absorption depends more on intake than on the need for calcium, and that calcium balance is regulated by excretion in the urine of calcium absorbed in excess of the needs of the body rather than by control of absorption. Depending upon the intake, 15–50% of ingested calcium is excreted in the feces. *See* ABSORPTION (BIOLOGY).

Tritium is especially useful in microautoradiographs because the average beta particle emitted has a range of only 1 micrometer in tissue, giving especially sharp resolution.

As illustrated in **Fig. 3**, chromosomes can be labeled with tritium through the use of tritium-labeled thymidine, a substance which localizes in deoxyribonucleic acid (DNA), which in turn occurs only in the genetic material. After the cells were labeled by being grown in the tritiated thymidine medium, they were transferred to a nonradioactive medium for growth of the second generation. The rate and mechanism of synthesis and degradation of DNA are under intensive study. Several lines of study, such as the labeling of successive generations of bacteria and the absence of label in the second-generation chromosomes marked by the arrows in Fig. 3, indicate that DNA is not normally subject to exchange or turnover processes, but is synthesized only when the cell is preparing to divide. *See* DEOXYRIBONUCLEIC ACID (DNA).

The kinetics of cellular proliferation has provided a rich vein for application of radioisotopic methods. For example, the lifetime of human red blood cells (about 120 days) was established with the use of ^{59}Fe-labeled cells. The life cycles of blood-cell precursors and of white blood cells and platelets have also been studied with a variety of radioisotopes, including tritium.

Some applications, such as the intake of ^{131}I by the thyroid, the measurement of the red-cell mass with ^{51}Cr-labeled red cells, and the absorption of ^{60}Co-labeled vitamin B_{12}, are of practical clinical importance in the diagnosis and treatment of disease. Knowledge of the rates of distribution and disposal of a wide variety of radioactive substances is basic to the problem of evaluation of the hazard from fallout radiation. Determination of induced activities, particularly ^{24}Na, provides a basis for calculating the neutron exposure in reactor accidents and in therapeutic exposures to neutrons. James S. Robertson

Bibliography. L. G. Colombetti (ed.), *Biological Transport of Radiotracers,* 1982; J. S. Robertson, Theory and use of tracer in determining transfer rates in biological systems, *Physiol. Rev.*, 37:133–154, 1957; R. J. Slater (ed.), *Radioisotopes in Biology: A Practical Approach*, 1990; F. Stohlman, Jr. (ed.), *Hemopoietic Cellular Proliferation,* 1970.

Radioisotope (geochemistry)

A branch of environmental geochemistry and isotope geology concerned with the occurrence of radioactive nuclides in sediment, water, air, biological tissues, and rocks. The nuclides have relatively short half-lives ranging from a few days to about 10^6 years, and occur only because they are being produced by natural or anthropogenic nuclear reactions or because they are the intermediate unstable daughters of long-lived naturally occurring radioactive isotopes of uranium and thorium. The nuclear radiation, consisting of alpha particles, beta particles, and gamma rays, emitted by these nuclides constitutes a potential health hazard to humans. However, their presence also provides opportunities for measurements of the rates of natural processes in the atmosphere and on the surface of the Earth. *See* ALPHA PARTICLES; BETA PARTICLES; GAMMA RAYS.

The unstable daughters of uranium and thorium consist of a group of 43 radioactive isotopes of 13 chemical elements, including all of the naturally occurring isotopes of the chemical elements radium, radon, polonium, and several others. Several nuclides in this group are used for measuring the rates of deposition of sediment in lakes and in the oceans and for dating calcium carbonate precipitated inorganically or in the skeletons of corals, mollusks, and other organisms. In addition, the radioactive daughters of long-lived uranium-238 (uranium-234, thorium-230, radium-226, and lead-210) have been used to study the origin of lava flows on active volcanoes, such as Mount Vesuvius near Naples, Italy, and Kilauea on the island of Hawaii.

A second group of radionuclides is produced by the interaction of cosmic rays with the chemical elements of the Earth's surface and atmosphere. This group includes hydrogen-3 (tritium), beryllium-10, carbon-14, aluminum-26, silicon-32, chlorine-36, iron-55, and others. The half-lives of these nuclides range from 2.8 years for iron-55 to 1.5×10^6 years for beryllium-10, which makes them useful for measurements of the rates of a wide range of geological, geochemical, and biochemical processes. The carbon-14 method of dating is especially important in archeology and the geology of the Quaternary Period. This group also includes short-lived radionuclides such as beryllium-7, phosphorus-32, phosphorus-33, and sulfur-35, whose half-lives are measured in days and which find application in the study of atmospheric processes. *See* COSMOGENIC NUCLIDE; DATING METHODS.

A third group of radionuclides is produced artificially by the explosion of nuclear devices, by the operation of nuclear reactors, and by various particle accelerators used for research in nuclear physics. The radioactive fission products and transuranium elements (neptunium, plutonium, and americium) that accumulate in the fuel rods of nuclear reactors must be isolated in underground repositories because of their intense radiation which is harmful to humans and all other forms of life. The dispersion of these radionuclides as a result of accidental explosions of nuclear reactors or during the testing of nuclear weapons creates a health hazard, especially when the radionuclides occur in food, drinking water, and air. The release of these radionuclides into the atmosphere has resulted in their dispersion around the Earth. Some of the radionuclides produced in nuclear reactors decay sufficiently slowly to be useful for geochemical research, including strontium-90, cesium-137, iodine-129, and isotopes of plutonium. The explosion of nuclear devices in the atmosphere has also contributed to the abundances of certain radionuclides that are produced by cosmic rays such as tritium and carbon-14. *See* RADIOACTIVE FALLOUT; TRANSURANIUM ELEMENTS.

Ionium. Thorium-230 (ionium) is a daughter of long-lived uranium-238. It is radioactive and decays to radium-226 with a half-life of 75,200 years. Ionium is used to measure the rate of deposition of sediment in the oceans. The method is based on the fact that atoms of thorium are concentrated with the rock and mineral particles deposited in the ocean, whereas atoms of uranium can remain in seawater for long periods of time. Ionium is thereby separated from its parent, and decays with its characteristic half-life as time passes and sediment continues to accumulate. For practical reasons, the rate of decay of unsupported ionium in the sediment is expressed as the ratio of the activity of ionium divided by the activity of thorium-232, which is the long-lived naturally occurring isotope of this element. The natural logarithm (ln) of the activity ratio (R) of ionium to thorium-232 in the sediment at some depth (h) below the bottom of the ocean is related to the rate (a) of sediment deposition by the equation $\ln R = \ln R_0 - \lambda h/a$, where R_0 is the activity ratio of recently deposited sediment and λ is the decay constant of ionium. The sedimentation rate is determined by analyzing a series of samples taken at intervals along a piston core collected from an oceanographic ship at sea. The activity ratios expressed as $\ln R$ are then plotted versus depth in the core (h) and a straight line is fitted to the data points. The slope (m) of this line yields the sedimentation rate (a) from the equation $m = -\lambda/a$. The results of such measurements indicate that the rate of accumulation of sediment in the ocean basins is of the order of a few millimeters per thousand years or a few meters per million years. The thickness of sediment in the deep basins of the oceans is generally less than 500 m (1600 ft), which indicates that they are quite young when viewed in the perspective of the history of the Earth. This conclusion is compatible with the theory of plate tectonics and sea-floor spreading. *See* MARINE SEDIMENTS; PLATE TECTONICS; THORIUM.

Thorium-230/uranium-234. The calcium carbonate that is precipitated by organisms such as corals and mollusks excludes thorium-230 but admits uranium-234, which is the parent of thorium-230 in the decay chain of uranium-238. Therefore, calcium carbonate, precipitated in the oceans or in saline lakes either by organisms or by inorganic chemical reactions, does not contain thorium-230 at the time of deposition. However, as the calcium carbonate deposit (composed of either calcite or aragonite) ages, thorium-230 is produced by decay of uranium-234. Therefore, the radioactivity due to the presence of thorium-230 increases with time until radioactive equilibrium between uranium-234 and thorium-230 is reestablished. When this condition of secular radioactive equilibrium is achieved, thorium-230 decays at exactly the same rate at which it is produced by decay of uranium-234. However, before that happens the activity ratio of thorium-230 to uranium-234 increases with time. After the thorium-230/uranium-234 activity ratio of a calcium carbonate samples has been measured, its age is determined. This method can be used to date calcium carbonate samples that are less than about 250,000 years old. It was used successfully to determine the ages of coastal terraces on the island of Barbados in the Lesser Antilles. *See* URANIUM; THORIUM.

Lead-210. Radon-222, a radioactive daughter of uranium-238, tends to escape from uranium-bearing minerals into the atmosphere, where it decays rapidly and eventually produces lead-210, which is removed from the atmosphere by rain and snow and is deposited on the surface of the Earth. Lead-210 decays to bismuth-210 with a half-life of 22.3 years, and is therefore useful for dating sediment deposited in lakes and in the coastal ocean during the twentieth century. Sediment deposited in the recent past reflects the effects of pollution of lakes, rivers, and the atmosphere by discharge of industrial and domestic waste. Lead-210 is also used to measure the rate of deposition of snow on the ice sheets of Antarctica and Greenland. The rate of deposition is obtained from the slope of a straight line fitted to measurements of the disintegration rate of lead-210 in snow collected at increasing depths below the surface. An interesting result obtained by this method is that the average annual rate of precipitation at the South Pole is only 2.4 in. (6 cm) per year. Evidently the interior of Antarctica is a cold desert where an ice sheet more than 1.8 mi (3 km) thick has formed by accumulation of snow and ice. The glaciation of Antarctica probably began more than 2.5×10^7 years ago, and has exerted a strong influence on the climate of the Earth for a long time. *See* LEAD ISOTOPES (GEOCHEMISTRY); RADON.

Beryllium-10. Cosmic rays cause nuclear reactions in the atmosphere that result in the formation of beryllium-10, which is rapidly brought down by meteoric precipitation and accumulates in sediment at the bottom of the oceans and in soil on the continents. Beryllium-10 is also produced by nuclear

spallation reactions of high-energy cosmic-ray neutrons in quartz (SiO_2) when this mineral is exposed at the surface. The amount of "cosmogenic" beryllium-10 (and aluminum-26) in quartz exposed to cosmic rays increases with time at a rate that depends on the intensity of the cosmic-ray flux, which increases with elevation above sea level and the geomagnetic latitude of the site. Therefore, the concentrations of beryllium-10 (and aluminum-26) in quartz that has been exposed to cosmic rays can be used to determine the duration of the exposure or its so-called exposure age. This method of dating has been used to determine exposure ages of quartz sandstone on the summits of the Transantarctic Mountains, which are emerging from the East Antarctic ice sheet either because the ice sheet is getting thinner or because the mountains are being uplifted, or both. Cosmic-ray exposure ages can also be measured by means of the concentration of neon-21 in quartz and chlorine-36 in ordinary rocks. These kinds of measurements are possible because of the availability of tandem accelerator mass spectrometers whose detection limit for beryllium-10 and aluminum-26 is about 1 million atoms. *See* BERYLLIUM; COSMIC RAYS; MASS SPECTROSCOPE.

Cesium-137. The testing of nuclear weapons in the atmosphere from about 1945 resulted in the deposition of cesium-137 and other fission products on the surface of the Earth. The injection of this and other fission products into the atmosphere varied from year to year, depending on the frequency and character of the weapon tests. High rates of input of cesium-137 occurred in 1963 and also in 1954. The high radioactivity of sediment deposited in these years is used to correlate sediment deposited in different places and to estimate the rate of deposition. A large quantity of cesium-137 and other radionuclides was discharged into the atmosphere on April 26, 1986, when a nuclear reactor at Chernobyl in the northern Ukraine exploded catastrophically. Fallout from this event was detected in Europe 2 days later, and subsequently spread around the world. *See* CESIUM; ENVIRONMENTAL RADIOACTIVITY; RADIOISOTOPE.

Gunter Faure

Bibliography. N. A. Chapman and G. M. McKinley, *The Geological Disposal of Nuclear Waste*, 1987; G. Faure, *Principles and Applications of Inorganic Geochemistry*, 1998; G. Faure, *Principles of Isotope Geology*, 2d ed., 1986; T. -L. Ku, The uranium-series methods of age determination, *Rev. Earth Planetary Sci.*, vol. 4, 1976; W. W. Nazaroff and A. V. Nero, Jr. (eds.), *Radon and Its Decay Products in Indoor Air*, 1988.

Radiolaria

A group of marine protists, regarded as a subclass of Actinopodea in older classifications, but not recognized as a natural group in some modern systems owing to its heterogeneity. Radiolarians occur almost exclusively in the open ocean as part of the plankton community, and are widely recognized for their ornate siliceous skeletons produced by most of the groups (**illus.** *a–c*). Their skeletons occur abundantly in ocean sediments and are used in analyzing the layers of the sedimentary record (biostratigraphy). Although most radiolarians occur only in open ocean locations, a unique skeleton-bearing nasselarian radiolarian (*Lophophaena rioplatensis*) has been identified in a coastal location in the Rio de la Plata estuary in South America.

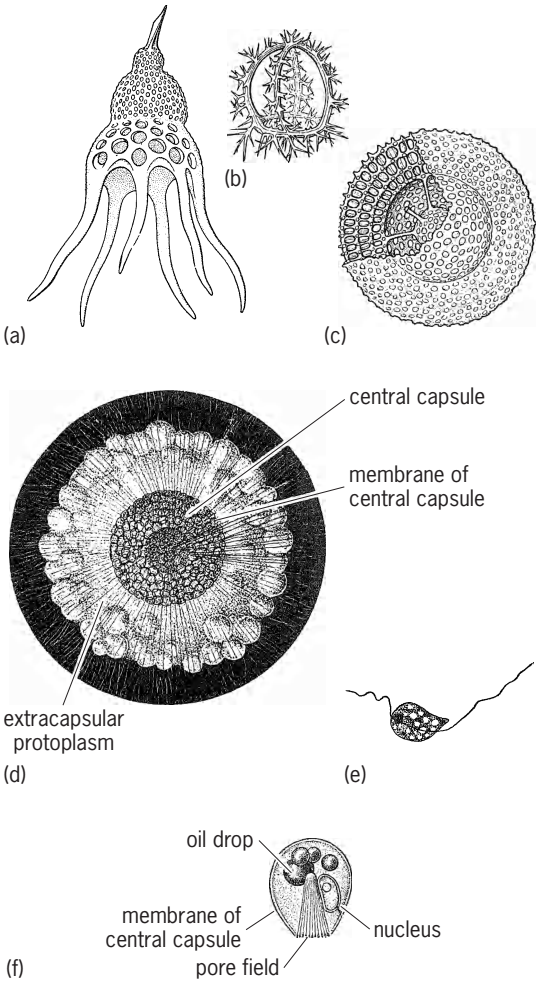

Radiolaria. Skeletons representing (*a*, *b*) certain Nasselarida and (*c*) certain Spumellarida. (*d*) *Thalassicolla*, one of the Spumellarida without skeletal elements. (*e*) Biflagellate gametes. (*f*) Central capsule, one of the Nasselarida showing one group of pores. (*After L. H. Hyman, The Invertebrates, vol. 1, McGraw-Hill, 1940*)

Phylogeny. In some classification systems, the Radiolaria are subdivided into two classes, Polycystinea and Phaeodarea. However, clarification of the evolutionary relationships (phylogeny) of the Radiolaria using modern molecular genetic analyses indicates that Polycystinea and the Phaeodarea are undoubtedly polyphyletic: that is, they are not derived from closely related phylogenetic ancestors. Therefore, current taxonomic systems separate them.

Molecular genetic analyses have begun to clarify the phylogenetic relationships among some of the major groups of Polycystinea. The colonial and skeletonless species (for example, *Thalassicolla*) form

a lineage that is distinct from the skeleton-bearing species. Also, there is evidence that the Polycystinea and Acantharia (with skeletons composed of strontium sulfate) are phylogenetically closely related. These data may indicate that Acantharia and Polycystinea should be grouped together taxonomically. This is of interest historically since Ernst Haeckel, a prominent nineteenth-century zoologist, originally grouped them together as members of the class Radiolaria. However, we now recognize that much of his system of classification is artificial; that is, it is based on superficial similarities of form that are not indicative of close evolutionary relationships. Much additional work needs to be done using modern techniques of electron microscopy and molecular genetics to more fully elucidate the natural systematic relationships of species included in the Radiolaria.

Morphology. A characteristic feature of the group is the capsule, a central mass of cytoplasm bearing one or more nuclei, food reserves, and metabolic organelles. This is surrounded by a perforated wall and a frothy layer of cytoplasm known as the extracapsulum, where food digestion generally occurs and numerous axopodia (stiffened strands of cytoplasm) and rhizopodia radiate toward the surrounding environment. Radiolarians feed on microplankton captured by the sticky axopodia. Their prey include bacteria, algae, some protozoa, and small arthropod zooplankton such as copepods. Algal symbionts, including dinoflagellates, green algae, and golden-brown pigmented algae, occur profusely in the extracapsulum. The algal symbionts living within the protection of the extracapsulum provide photosynthetically derived food for the radiolarian host.

Polycystinea. In the order Spumellarida (class Polycystinea) the central capsule is perforated by numerous pores distributed evenly on the surface of the capsular wall. These pores, containing strands of cytoplasm, provide continuity between the cytoplasm in the central capsule and the surrounding extracapsulum. The skeletons of the Spumellarida are characteristically developed on a spherical organizational plan, but some are spiral-shaped (resembling snail shells) or produce elongate skeletons composed of numerous chambers built one upon another. In some genera, such as *Thalassicolla* (illus. *d*), there is no skeleton; in others there are rods or spicules, or often a single or multiple concentric latticework skeleton (illus. *c*). The shape of the skeleton is determined by the dynamic action of the radiolarian cytoplasm that secretes the silica within a specialized part of the rhizopodial system called a cytokalymma.

Multicellular aggregates (colonies of spumellaridans), measuring several centimeters in diameter (or even several meters in some rare elongate forms), consist of numerous radiolarian central capsules enclosed within a gelatinous envelope and interconnected by a web of rhizopodia that bears abundant algal symbionts. A thin halo of feeding rhizopodia protrudes from the surface of the colony and is used to capture prey. The kinds of prey observed in solitary and colonial forms of Spumellarida include copepods, small crustacean larvae, and a variety of algae (diatoms, dinoflagellates, and others). Reproduction is poorly understood. In some colonial forms, daughter colonies are produced by asexual reproduction (fission). Flagellated swarmers (illus. *e*) released from mature central capsules of some species are possibly gametes and have a large crystal inclusion of strontium sulfate.

In the Nasselarida (Polycystinea), the central capsule is often ovate and the pores are localized at one pole (illus. *f*). The axopodia and rhizopodia emerge from this pore field and are supplied by a cone-like array of microtubules within the central capsule. The skeleton, when present, is often shaped like a dome or helmet, but some forms may be ornamented with lateral lattices, spines, or other embellishments. Only dinoflagellate algae have been observed in association with the Nasselarida.

Phaeodarea. Radiolarians in the class Phaeodarea possess a central capsule with two types of pore areas, a larger one (astropyle) that serves as a kind of cytopharynx where food is carried into the central capsule, and two accessory openings (parapylae) where smaller strands of cytoplasm emerge. The phaeodarian skeleton exhibits a wide range of shapes, including geodesic-like lattice spheres and small porous clam-shaped shells. No symbiotic algae have been reported in the Phaeodarea.

Given the peculiarities of the two kinds of pores in the central capsule, and their distinct differences from the fusules (cytoplasmic strands) of the Polycystinea, it is very probable (as indicated by some modern molecular genetic analyses) that the Phaeodarea are not closely related to Polycystinea. Similarities in the capsular organization may occur through adaptive convergence; that is, over evolutionary time, these two distinct lineages have adapted to their environment by developing similar morphological features, including siliceous skeletons.

Fossil forms. Radiolarians have a fossil record that extends back to early Paleozoic time, about a half billion years ago. Compared with other groups of skeleton-bearing marine microplankton, they are highly diverse, several hundred species having inhabited the oceans at any given time. Their skeletons of opaline silica are more readily dissolved than are calcareous and organic-walled microfossils; therefore, radiolarians occur in rocks much less commonly than do foraminifera, calcareous nannofossils, and palynomorphs (pollen, spores, and the like). However, the microfossil record of radiolarians in deep-sea sediments and in some rock formations, including cherts and limestones, is very good and has been used by micropaleontologists to elucidate the geological history of the Earth.

In Cenozoic sediments radiolarians usually occur in clays or shales, but older occurrences are often in calcareous nodules or cherts, where they have been protected from the destructive effects of pressure, chemical change, and dissolution. Some notable

land-based occurrences are in late Cenozoic diatomites of California, Japan, and the Mediterranean region; Mesozoic cherts of California and the Alps; and Paleozoic calcareous and phosphatic nodules in England, Australia, and the Canadian Arctic.

Radiolarians have undergone such profound evolutionary changes through time that it is difficult to determine the relationships between Paleozoic, Mesozoic, and Cenozoic forms. Because they are planktonic and have undergone continuous evolutionary change, radiolarians are particularly useful for determining time equivalence (and geological ages) of marine sedimentary deposits at widely separated localities. As with many other groups of organisms, a particularly abrupt change occurred at the Mesozoic-Cenozoic boundary. The Cenozoic record of radiolarians in sediments, particularly on the deep-sea floor, is sufficiently complete to show the course of evolutionary change in considerable detail.

Assemblages of fossil radiolarians also provide clues to oceanic conditions during the geological past. Each of the major oceanic water masses has its characteristic radiolarian fauna, and so changes in the distribution or composition of these assemblages can be interpreted in terms of changes in the pattern of water masses, or in their oceanographic properties. Since some species of Radiolarians inhabit surface water of the oceans in specific climatic regions with narrow temperature requirements for survival, the skeletons of radiolarians in sediments can be used to clarify the ancient climates of the Earth. They also have been used to analyze ancient terranes, since the species of radiolarian fossils isolated from surface rocks can be used to date the origin of the rocks and, in some cases, their original geographic locale.

Some species of radiolarians dwell only in the surface water of the ocean, requiring sufficient biological productivity to feed and survive, while others occur only at great depths. The deep-dwelling species are less dependent on surface productivity. Hence, the proportion of surface-dwelling species to deep-dwelling species deposited as microfossils within a stratum of the sediments can be used as evidence of the amount of productivity of the surface water in the ancient seas. *See* ACANTHARIA; ACTINOPODEA; OCULOSIDA; PROTOZOA; SARCODINA; SARCOMASTIGOPHORA.

O. Roger Anderson; William R. Riedl

Bibliography. O. R. Anderson, The physiological ecology of planktonic sarcodines with applications to paleoecology: Patterns in space and time, *J. Eukaryotic Microbiol.* 43:261–274, 1996; O. R. Anderson, *Radiolaria*, 1983; O. R. Anderson, The trophic role of planktonic foraminifera and radiolaria, *Mar. Microb. Food Webs*, 7:31–51, 1994; D. Boltovskoy et al., First record of a brackish radiolarian (Polycystina): *Lophophaena rioplatensis* n. sp. in the Rio de la Plata estuary, *J. Plankton Res.*, 25:1551–1559, 2003; B. U. Haq and A. Boersma (eds.), *Introduction to Marine Micropaleontology*, 1978; D. Lazarus, A brief review of radiolarian research, *Paläontologische zeitschrifft*, 79:183–200, 2005.

Radiology

The medical science concerned with x-rays, radioactive materials, and other ionizing radiations, and the application of the principles of this science to diagnosis and treatment of disease. Nonionizing radiations of infrared and ultrasound are also used for diagnosis.

Diagnostic radiology. This technology uses radiation, usually x-rays, to study the configuration of anatomical structures or the function of body organ systems. The most common techniques used in diagnostic radiology are discussed below.

Radiography is the formation of an image, usually on a photographic emulsion, by the action of ionizing radiation. The image is the result of the differential absorption of the radiation in its passage through a part of the body. The ionizing radiation is usually x-rays, but protons and fast neutrons have also been used. Contrast media have been developed which help to visualize body structures such as blood vessels, kidney, and gallbladder. Most of these media contain firmly bound iodine, which has a relatively high atomic number. There is a probability, usually quite low, that the patient may have a reaction to the contrast media. Physicians must always be concerned about benefit-cost considerations when obtaining radiographic studies, so as to balance the benefit of the information gained against the monetary, radiation, and possible drug reaction cost. *See* RADIOGRAPHY.

Fluoroscopy is the examination of the structure of a part of the body by passing a beam of x-rays through it and viewing the image formed on a fluorescent screen by the transmitted radiation. This permits visualization of internal organs in motion. In some cases, such as a gastrointestinal examination, contrast is increased by having the patient drink a dense fluid such as a suspension of barium sulfate.

Image intensification of the fluoroscopic image by an electronic device (orthicon, plumbicon, vidicon) permits visualization of the internal organs with x-ray beams of lower intensity than those used in direct fluoroscopy. This permits a decrease in radiation exposure received by the patient. The intensified image may be viewed on a television screen, a practical advantage for the physician and a source of considerable interest to some patients.

There are studies in diagnostic radiology with special names which usually indicate the organs to be examined, such as cholestography (gallbladder), angiocardiography (heart and large blood vessels), and neuroradiology (brain and spinal cord).

Nuclear medicine. Radioactive isotopes are used to obtain images of organ systems and functions. A compound such as radioactive technetium-pertechnetate (^{99m}Tc-pertechnetate) after intravenous injection is

used to locate brain tumors, and thyroid function is studied by the uptake of radioactive iodine (usually 131I) by the thyroid gland after oral administration. The accumulation of isotope in a tumor or an organ such as the thyroid is recorded by a suitable gamma-ray detector such as a gamma-ray-sensitive crystal attached to an electronic amplifier and recording equipment. The image of the radioactivity concentrated in an organ is viewed on a television-type screen and recorded on a photographic print. A large number of body organs and functions can be studied with a variety of commercially available radiopharmaceutical compounds.

Ultrasound. Sound waves of 1–10 MHz/s in pulses of 1.5 μs duration are transmitted from a crystal transducer, and after amplification are displayed as A scan or B scan on an oscilloscope and recorded on a photographic print. The ultrasound pulses demonstrate organ structures such as the heart, liver, and spleen. Although the resolution is less fine than that obtained with x-ray there is an advantage in that the ultrasound is nonionizing radiation. Ultrasound is particularly useful, therefore, in determining the size and degree of development of the human fetus.

Thermography. Infrared radiation from the human body is used to detect tumors such as breast tumors, which are near the body surface. The technique, which is still experimental, is based on the idea that tumors are warmer than the surrounding normal tissue. This increase in temperature is detected by an infrared device, and the "hot spot" scan is displayed on a television-type screen, with permanent records kept on photographic prints.

Ultrasound and thermography are techniques which are used to supplement the information obtained by diagnostic radiology and nuclear medicine studies.

Therapeutic radiology. Radiation therapy deals with the treatment of disease with ionizing radiation. The diseases most commonly treated are cancer and allied diseases. Radiation therapy has been found useful in the management of some diseases such as ringworm of the scalp and bursitis, but because of possible serious complications occurring many years later, the use of ionizing radiation is generally avoided if alternative methods of treatment are available.

In cancer therapy the objective is to destroy a tumor without causing irreparable radiation damage in normal body tissues that must of necessity be irradiated in the process of delivering a lethal dose to the tumor. This applies particularly to important normal structures in the vicinity of the tumor. The relative radiosensitivity of the tumor with respect to these normal structures is the chief factor determining the success of the treatment. The effect of the radiation on the tumor and other tissues becomes apparent gradually, in a matter of weeks. During this latent period a certain amount of recovery or repair may also take place.

The optimal differential between the effect on the tumor and the effect on normal tissues of the patient as a whole results from the proper adjustment of many treatment factors, which in general requires great clinical experience. The time factor is very important. This involves the administration of the therapeutic dose in one of three ways: by one short treatment, by protraction as continuous irradiation over a long time, or by fractionation in small repeated doses. The importance of the distribution of radiation in the patient's body is obvious. Ideally, only the tumor should be irradiated, but this is impossible, because in general there is infiltration of adjacent normal tissues of unknown extent. If the radiosensitivity of the tumor is much higher than that of the surrounding normal tissue, the problem is relatively simple, but in general this is not the case.

Various techniques have been devised to bring about a reasonable distribution of radiation in the region of the tumor. When x-rays are used for the treatment of deep-seated tumors, the beam must traverse a considerable thickness of normal tissues. To minimize absorption in these tissues, higher and higher voltage x-rays have been used. Supervoltage x-ray therapy makes use of x-rays produced at several hundred kilovolts to 2000 kV. Megavoltage x-rays range from a few to 70 MV. Cross-fire treatment involves the use of more than one beam aimed at the tumor, but passing through different skin areas. Rotation therapy involves the use of one beam constantly aimed at the tumor, while either the patient or the source of radiation is rotated about an axis passing through the center of the tumor. When the source is rotated, 1000–2000 Ci of cobalt-60 (^{60}Co) are generally used to provide a well-collimated beam of gamma rays.

Fast neutrons, high-energy electrons, and deuterons, as external beams, have been used for therapeutic purposes. By using radioactive substances, other techniques are possible. In the case of intracavitary therapy, the source, in a suitable container, is placed within a body cavity in which cancer is present. In the case of interstitial therapy, many needlelike sources are distributed throughout the tumor and removed at the completion of the treatment. By using radon, which has a short half-life, the sources may be made into radon seeds, which are very small and may be left in place permanently. Sources of tantalum-182 (^{182}Ta) are sometimes used in a similar way. Attempts to infiltrate a tumor with radioactive material in fluid form have not been very successful, but some good results have been obtained by using colloidal gold-198 in the treatment of cancer of the prostate. The same material has been used for the palliative treatment of ascites tumors in the pleural or abdominal cavity. Phosphorus-32 has been found useful in the treatment of polycythemia vera and leukemia by injection into the bloodstream. Iodine-131, which, when introduced into the body by any route, tends to concentrate almost entirely in the thyroid, is very effective in reducing the size and function of the thyroid in patients suffering from hyperthyroidism. Certain types of cancer of the thyroid and metastases can be treated similarly with radioactive iodine, but the doses must be much larger and the arrest of the tumor process is generally temporary. *See* NUCLEAR RADIATION (BIOLOGY); RADIATION BIOLOGY.

Lee B. Lusted

Radiometry

A branch of science that deals with the measurement or detection of radiant electromagnetic energy. Radiometry is divided according to regions of the spectrum in which the same experimental techniques can be used. Thus, vacuum ultraviolet radiometry, intermediate-infrared radiometry, far-infrared radiometry, and microwave radiometry are considered separate fields, and all of these are to be distinguished from radiometry in the visible spectral region. Curiously, radiometry in the visible is called radiometry, optical radiation measurement science, or photometry, but it is not called visible radiometry. *See* ELECTROMAGNETIC RADIATION; INFRARED RADIATION; LIGHT; MICROWAVE; ULTRAVIOLET RADIATION.

The use of the word photometry to mean radiometry in the visible portion of the spectrum is misleading. Strictly speaking, photometry is the measurement of electromagnetic radiation according to its ability to produce visual sensation. This, of course, is a subjective attribute of the radiation. Radiometry, on the other hand, is concerned with physical attributes of the radiation such as its energy content and spectral distribution. The reason for the confusion between photometry and radiometry is that the meanings of these words have changed considerably over the last 200 years. *See* PHOTOMETRY.

Development and nomenclature. During the 1700s, techniques were developed for measuring light using the human eye as a null detector to compare the relative intensity of light coming from different sources. This was referred to as photometry. During the same period, radiant heat was studied with sensitive liquid-in-glass thermometers, and this activity was referred to as radiometry. Also during this period, the photochemical activity of solar radiation was studied by observing the photoinduced decomposition of silver compounds into metallic silver, and this activity was called actinometry. At that time it was not known that radiant heat and light were the same phenomenon; the demonstrations that light was a wave phenomenon and that mechanical energy and heat were different forms of a more general conception of energy were to come; and the electromagnetic and quantum theories of radiation had not been proposed. Therefore, the emphasis of these studies was not to quantify the radiation, but rather to learn the nature of radiation and its sources.

Following the discovery of infrared radiation in 1800 and of ultraviolet radiation in 1801, a great deal of effort was devoted to studies of their properties. The only practical detectors of ultraviolet radiation at that time were photochemical reactions, and the only detectors of infrared radiation were thermometers, while the eye remained the most sensitive detector of visible radiation. Thus actinometry, photometry, and radiometry became synonymous with studies of radiation in the ultraviolet, visible, and infrared spectral regions, respectively.

It is modern practice to refer to any radiation detector (such as a thermometer) that responds to an increase in temperature caused by the absorption of radiant energy as a thermal detector. Similarly, any detector (such as a photochemical reaction) that responds to the excitation of a bound electron is called a photon or quantum detector. Almost all of the many different types of modern radiation detectors fall into one of these two classes.

By the end of the 1800s, a number of different effects that could be used as the basis for thermal detectors had been discovered, and some of these were sensitive enough for use in the ultraviolet. Similarly, a number of effects that could be used as the basis for photon detectors had been discovered, and some of these responded to infrared radiation. Furthermore, both thermal and quantum detectors could be used quite satisfactorily with visible radiation for certain applications. As a result of these developments, the distinctions between radiometry, actinometry, and photometry in terms of spectral regions became artificial, and the words began to assume new meanings. Radiometry became the more general term, while photometry and actinometry came to denote the visual and photochemical effects of electromagnetic radiation, respectively.

Nevertheless, the older meanings continued to coexist with the newer. Even now, astronomers refer to radiometry as photometry, and their reference to infrared photometry would be a contradiction in terms according to the modern definition of photometry in terms of the ability to produce visual sensation.

Thermal detectors. Liquid-in-glass thermometers are sluggish and relatively insensitive. The key to developing thermal detectors with better performance than liquid-in-glass thermometers has been to secure a large and rapid rise in temperature associated with a high sensitivity to temperature changes.

The temperature of any thermal detector will increase until the rate of loss of heat to its surroundings is equal to the rate at which heat is absorbed. To secure the largest rise for a given radiation, the detector should absorb the radiation as completely as possible, while the loss of heat, by reradiation, conduction, and convection, must be as small as possible. To secure a rapid rise in temperature, the heat capacity of the detector should be as small as possible. To obtain the least interference from temperature fluctuations, the absorbing surface should be as small as possible.

These criteria lead to detectors which are very small and thin, with thin black coatings housed in vacuum containers with a window to admit the radiation. Special applications often require that this optimized design be sacrificed for improved accuracy, convenience, or robustness. Thus large-area detectors without windows have been used as standards against which the most sensitive detectors are calibrated.

Thermal detectors have been based upon a number of different principles. Radiation thermocouples produce a voltage, bolometers undergo a change in resistance, pyroelectric detectors undergo a change in spontaneous electric polarization, and the gas in pneumatic detectors (Golay cells) and photoacoustic detectors expands in response to incident radiation.

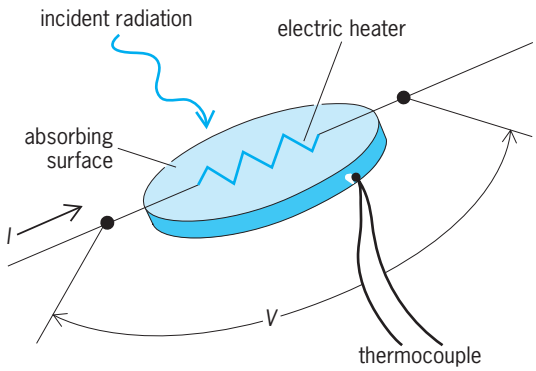

Fig. 1. Absolute radiometer. The electrical power dissipated in the heater is given by the product of the voltage drop *V* across the heater times the current *I* passing through the heater.

The periodic expansion and contraction of the gas in response to high-frequency modulated radiation is detected by a sensitive microphone in the case of the photoacoustic detector. The Golay cell, on the other hand, uses a sensitive photomultiplier and a reference beam of light to detect distortion of a flexible membrane mirror caused by the expansion and contraction of the gas. *See* BOLOMETER; PYROELECTRICITY; THERMOCOUPLE.

Thermal detectors with electrical heaters built into the part of the detector that absorbs the radiation can be used as radiation standards. The detector output caused by the unknown radiation is compared with the output caused by a measured quantity of electrical power dissipated in the heater to determine the radiant power content of the radiation. Thermal detectors of this nature are called absolute radiometers (**Fig. 1**).

The main problem with thermal detectors is that they respond not only to electromagnetic radiation but to any source of heat. This makes their design, construction, and use rather difficult, because they must be made sensitive to the radiation of interest while remaining insensitive to all other sources of heat, such as conduction, convection, and background radiation, that are of no interest in the particular measurement.

Photon detectors. These detectors, on the other hand, respond only to photons of electromagnetic radiation that have energies greater than some minimum value determined by the quantum-mechanical properties of the detector material. Since heat radiation from the environment at room temperature consists of infrared photons, photon detectors for use in the visible can be built so that they do not respond to any source of heat except the radiation of interest.

Except for certain measurements of interest in photochemistry and photobiology, chemical reactions that respond to radiation are not very convenient detectors. The key to developing photon detectors that are more convenient than photochemical reactions has been to base them on some electrical principle, so that the most advanced electronic techniques could be used to detect and process the detector signal. So many different approaches have been used that only the most important will be mentioned here.

After vacuum-tube technology was introduced, its techniques were applied to the external photoelectric effect to produce vacuum photodiodes and photomultipliers. For ultraviolet and short-wavelength visible radiation, photomultipliers are the most sensitive detectors available. By using photomultipliers in properly designed experiments, it is possible to record the absorption of a single photon. This mode of operation is called photon counting. *See* PHOTOMULTIPLIER.

Following the introduction of planar silicon technology for microelectronics, the same technology was quickly exploited to make planar photodiodes based on the internal photoelectric effect in silicon. In these devices (**Fig. 2**) the separation of a photogenerated electron-hole pair by the built-in field surrounding the p^+n junction induces the flow of one electron in an external short circuit (such as the inputs to an operational amplifier) across the electrodes. The number of electrons flowing in an external short circuit per absorbed photon is called the quantum efficiency. The use of these diodes has grown to the point where they are the most widely used detector for the visible and nearby spectral regions. Their behavior as a radiation detector in the visible is so nearly ideal that they can be used as a standard, their cost is so low that they can be used for the most mundane of applications, and their sensitivity is so high that they can be used to measure all but the weakest radiation (which requires the most sensitive photomultipliers). *See* JUNCTION DIODE; PHOTODIODE; SEMICONDUCTOR DIODE.

Research efforts have been directed at producing photon detectors based on more exotic semiconductors, and more complicated structures to extend the sensitivity, time response, and spectral coverage of the radiometric techniques that have been

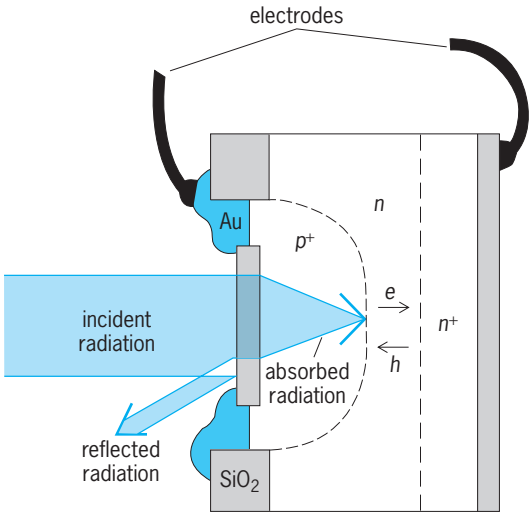

Fig. 2. Silicon photodiode. The separation of a photogenerated electron-hole pair (*e*-*h*) by the built-in field surrounding the p^+n junction induces the flow of one electron in an external short circuit across the electrodes.

developed around the planar silicon photodiode. *See* OPTICAL DETECTORS.

Radiometric sources. Any source of optical radiation can be used as a radiometric source. Lasers have become important radiometric sources. *See* LASER.

Some sources of radiation have properties that have made them useful as standards at various times. Early standard sources include candles of specified composition and construction, gas mantles, and lamps. The best-known standard source is the blackbody, which emits a unique spectral distribution of radiation dependent only upon its temperature. This distribution is given by Planck's law. No real body can behave as a true blackbody, but cavities surrounded by freezing metals can be used as blackbodies to a very high level of approximation. **Figure 3** illustrates

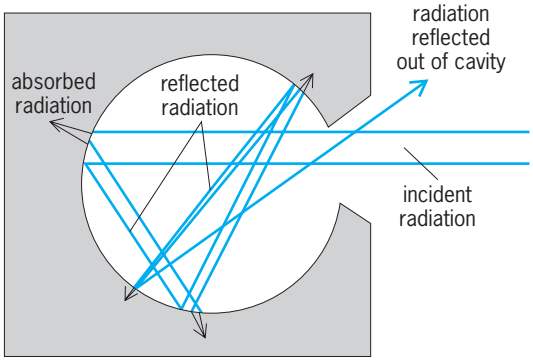

Fig. 3. Cavity approximating a blackbody, that is, an object that absorbs all of the radiation incident on it.

how the cavity geometry reduces the reflectance of the cavity opening below that of the surface of the cavity by multiple reflection and absorption of the radiation before it leaves the interior of the cavity. Thus the cavity appears blacker (more absorptive) than a flat surface made from the same material. Other standard sources in use include the radiation from plasma, and synchrotron radiation. These are used as standards primarily in the vacuum ultraviolet. *See* HEAT RADIATION; SYNCHROTRON RADIATION.

Jon Geist

Bibliography. R. W. Boyd, *Radiometry and the Detection Optical Radiation* 1983; W. Budde, *Physical Detectors of Optical Radiation*, vol. 4 of F. Grum and C. J. Bartleson (eds.), *Optical Radiation Measurements*, 1983; F. Grum and R. J. Becherer, *Radiometry*, vol. 1 of F. Grum (ed.), *Optical Radiation Measurements*, 1979; R. McCluney, *Introduction to Radiometry and Photometry*, 1994; C. L. Wyatt, *Radiometric Calibration: Theory and Methods*, 1978; C. L. Wyatt, *Radiometric Systems Design*, 1987.

Radish

A cool-season annual or biennial crucifer, *Raphanus sativus*, of Chinese origin belonging to the plant order Capparales. The radish is grown for its thickened hypocotyl, which is eaten uncooked as a salad vegetable (see **illus.**). Propagation is by seed.

Radish (*Raphanus sativus*), cultivar Red Boy. (*Joseph Harris Co., Rochester, New York*)

Varieties (cultivars) are classified according to root shape and season or time of maturity. Colors include red, yellow, white, black, pink, and red-white combinations. Popular varieties are short-season (21–25 days), Early Scarlet Globe and Comet; medium-season (30–50 days), Crimson Giant; and long-season or winter varieties (50–70 days), Black Spanish. Commercial production, largely of the round red short-season varieties, is primarily in the field but radishes are also produced commercially in greenhouses. Harvesting by hand or machine begins when roots are approximately $^1/_2$ to 1 in. (13 to 25 mm) in diameter, often only 21–23 days after planting. *See* CAPPARALES.

H. John Carew

Radium

A chemical element, Ra, with atomic number 88. The atomic weight of the most abundant naturally occurring isotope is 226. Radium is a rare radioactive element found in uranium minerals to the extent of 1 part for about every 3×10^6 parts of uranium. Chemically, radium is an alkaline-earth metal having properties quite similar to those of barium. Radium is important because of its radioactive properties and is used primarily in medicine for the treatment of cancer, in atomic energy technology for the preparation of standard sources of radiation, as a source for actinium and protactinium by neutron bombardment,

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb

actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Physical properties of radium

Property	Value
Atomic number	88
Atomic weight	226.05
Valence states	0, 2+
Specific gravity	6.0 at 20°C
Melting point	700°C (1290°F)
Boiling point	~1140°C (2080°F)
Ionic radius, Ra^{2+}	0.245 nm (estimated)
Atomic parachor	~140
Decomposition potential	1.718 volt
Heat of formation of oxide	130 kcal/mole
Magnetic susceptibility	Feebly paramagnetic

and in certain metallurgical and mining industries for preparing gamma-ray radiographs. *See* PERIODIC TABLE.

Thirteen isotopes of radium are known; all are radioactive; four occur naturally; the rest are produced synthetically. Only ^{226}Ra is technologically important. It is distributed widely in nature, usually in exceedingly small quantities. The most concentrated source is pitchblende, a uranium mineral containing about 0.014 oz (0.4 g) of radium per ton of uranium.

Biologically, radium behaves as a typical alkaline-earth element, concentrates in bones by replacing calcium, and, as a result of prolonged irradiation, causes anemia and cancerous growths. The tolerance dose for the average human being has been estimated at a total of 1 μg of radium fixed within the body. However, because radiations from radium and its decay products preferentially destroy malignant tissue, radium and radon, the gaseous decay product of radium, have been used to check the growth of cancer.

When first prepared, nearly all radium compounds are white, but they discolor on standing because of intense radiation. Radiation causes a purple or brown coloration in glass on long contact with radium compounds. Eventually the glass crystallizes and becomes crazed. Radium salts ionize the surrounding atmosphere, thereby appearing to emit a blue glow, the spectrum of which consists of the band spectrum of nitrogen. Radium compounds will discharge an electroscope, fog a light-shielded photographic plate, and produce phosphorescence and fluorescence in certain inorganic compounds such as zinc sulfide. The emission spectrum of radium compounds is similar to those of the other alkaline earths; radium halide imparts a carmine color to a flame.

When freshly prepared, radium metal has a brilliant white metallic luster. Some of its physical properties are shown in the **table**. Chemically, the metal is highly reactive. It blackens rapidly on exposure to air because of the formation of a nitride. Radium reacts readily with water, evolving hydrogen and forming a soluble hydroxide. *See* ALKALINE-EARTH METALS; NUCLEAR REACTION; RADIOACTIVE MINERALS; RADIOACTIVITY; RADON. Murrell L. Satutsky

Bibliography. D. R. Lide, *CRC Handbook of Chemistry and Physics*, 85th ed., CRC Press, 2004; G. Van Kaick et al., *Health Effects of Internally Deposited Radionuclides: Emphasis on Radium and Thorium: Proceedings*, World Scientific, 1995; V. M. Vdovenko and Yu. V. Dubasov, *Analytical Chemistry of Radium*, 1975.

Radius of gyration

The radial distance from a rotation axis at which the mass of an object could be concentrated without altering the moment of inertia of the body about that axis. If the mass M of the body were actually concentrated at a distance k from the axis, the moment of inertia about that axis would be Mk^2. If this quantity is equal to the actual moment of inertia I about the rotation axis, then k is the radius of gyration given by $k = \sqrt{I/M}$. The quantity k has dimensions of length and is measured in appropriate units of length such as meters. *See* MOMENT OF INERTIA.

For example, consider a rigid object of arbitrary shape with mass M rotating about the z axis with a moment of inertia I (see **illustration**). The equivalent moment of inertia would be attained if the entire mass of the object could be concentrated at one point a distance k from the rotation axis. That distance k is the radius of gyration. Edwin R. Jones

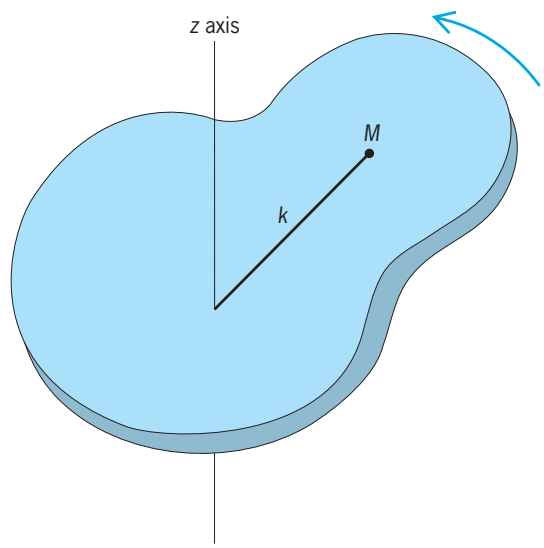

A rigid object of mass *M* rotates about the *z* axis with a moment of inertia *I*. A point mass *M* a distance *k* from the *z* axis rotates about the axis with the same moment of inertia.

Bibliography. H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3d ed., Addison-Wesley, Boston, 2002; J. B. Marion and S. T. Thornton, *Classical Dynamics of Particles and Systems*, 5th ed., Brooks Cole, Belmont, CA, 2004; K. R. Symon, *Mechanics*, 3d ed., Addison-Wesley, Boston, 1971.

Radome

A protective housing for radar, satellite communications, or similar antennas, often spherical but sometimes shaped to satisfy aerodynamic or other requirements of the installation. Radomes are designed to

satisfy both electromagnetic and mechanical requirements.

Radomes protect antennas from harmful environmental conditions. Shielding mechanically scanning antennas from winds, for example, reduces the mechanical power needed to ensure proper motion. Properly designed radomes can lessen the detrimental effects of ice, snow, and rain accumulation. *See* ANTENNA (ELECTROMAGNETISM).

Radomes on aircraft must be shaped for proper flight characteristics. Top-mounted larger antennas (as in airborne early warning radars) are housed in rotating radomes or "rotodomes," turning with the antenna inside; smaller antennas (as in airborne intercept radars or those in commercial aircraft) operate behind shaped stationary radomes at the nose of the aircraft or beneath it.

Spherical radomes for surface installations, surmounting a tower or on a perimeter foundation on the ground, may be tens of meters in diameter, large enough to allow the motion of the antenna inside. They may be constructed of thin dielectric sheets using scores of polygonal pieces, joined at the turned-in edges of the pieces or by attaching them to a skeletal framework of short metal members. Some domes are made of thin flexible material supported by positive relative air pressure maintained on the inside.

Some antennas have protective covers mounted on the antenna itself; the dish-shaped antennas of precision tracking radars and microwave terrestrial relays are often protected this way. Phased-array antennas (usually planar fixed apertures) may have flat dielectric sheets as the outer surface, contributing to both the antenna's impedance matching and its protection from the physical environment.

In general terms, the radome must be simultaneously transparent to the radar or communication signal and structurally sound. The designer has available many materials of varying characteristics to help in meeting these joint requirements such as thin rubberlike skins for air-supported domes and rigid sheets of various resins with reinforcing glass fibers when more strength is required. The dielectric nature of all such materials, causing both absorptive loss in the material itself and surface reflections, is of prime interest to the designer. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; RADAR-ABSORBING MATERIALS; REFLECTION OF ELECTROMAGNETIC RADIATION.

Most materials used are nonmagnetic (relative permeability near unity) but have widely varying permittivity (dielectric constant). Permittivity is a complex value, and the ratio of its imaginary and real values, the "loss tangent," relates to how lossy the material is. Loss tangents in the range of 0.01 to 0.001 are common; absorptive loss in thin sheets can be as little as 0.1 dB and in thicker rigid material still considerably less than 1 dB. *See* PERMITTIVITY.

When strength requires thicker material, materials of low loss tangent will achieve low absorptive loss, leaving surface reflections the major concern. Thickness of the material can be chosen to minimize the combined reflection from the two surfaces, inside and out; thicknesses of integral multiples of half-wavelengths (in the dielectric) give this favorable resonant condition. Higher values of permittivity sustain good performance over a large range of incidence angles required for large antennas in shaped radomes. Material of dielectric constant 4.2 and loss tangent 0.014 has, for about half-wavelength (in the material) thickness, a transmission efficiency of 94% at 0° incidence, but much less for incidence greater than 30° or so. Values of dielectric constant 9.0 and loss tangent 0.0004 give greater than 93% transmission efficiency for incidence up to 80°. The higher-permittivity material, however, requires far tighter manufacturing tolerances and introduces some other detrimental effects (such as polarization distortions). When still thicker material is required for strength, various "sandwich" designs with multiple layers are used. *See* INTERFERENCE OF WAVES; REFRACTION OF WAVES.

Special radomes can be made transparent in just one frequency band and reflective (opaque) at all other frequencies, generally by using conductor-clad dielectric material with slots of frequency-sensitive dimensions etched in the conducting surfaces. Radomes have also been built with artificial dielectric materials (with embedded conductors or such) that can be electronically excited for control of the characteristics. *See* DIELECTRIC MATERIALS.

Not all radar antennas require radomes in their installation. When required, however, the radome becomes a significant component of the radar demanding careful attention to its design. *See* RADAR.

Robert T. Hill

Bibliography. H. E. Schrank et al., Chap. 6, *Reflector antennas*, in M. E. Skolnik (ed.), *Radar Handbook*, 2d ed., McGraw-Hill, 1990.

Radon

A chemical element, Rn, atomic number 86. Radon is produced as a gaseous emanation from the radioactive decay of radium. The element is highly radioactive and decays by the emission of energetic alpha particles. Radon is the heaviest of the noble, or inert, gas group and thus is characterized by chemical inertness. More than 25 isotopes of radon have been identified. All isotopes are radioactive with short half-lives. *See* PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Radon is found in natural sources only because of its continuous replenishment from the radioactive decay of longer-lived precursors in minerals containing uranium, thorium, or actinium. ^{222}Rn (half-life 3.82 days), ^{220}Rn (thoron; half-life 55 s), and ^{219}Rn (actinon; half-life 4.0 s) occur in nature as members of the uranium (U), thorium (Th), and actinium (Ac) series, respectively. All three decay by the emission of energetic alpha particles. *See* ACTINIUM; ALPHA PARTICLES; RADIOACTIVITY; RADIUM; THORIUM; URANIUM.

Any surface exposed to ^{222}Rn becomes coated with an active deposit which consists of a group of short-lived daughter products. The radiations of this active deposit include energetic alpha particles, beta particles, and gamma rays. The ultimate decay products of radon following the rapid decay of the active deposit to lead-210 include bismuth-210, polonium-210, and, finally, stable lead-206. Radon possesses a particularly stable electronic configuration, which gives it the chemical properties characteristic of noble gas elements. It has a boiling point of −62°C (−80°F) and a melting point of −71°C (−96°F). The spectrum of radon has been extensively studied, and resembles that of the other inert gases. Radon is readily adsorbed on charcoal, silica gel, and other adsorbents, and this property can be used to separate the element from gaseous impurities.

Earl K. Hyde

The rocks and soils of the Earth's crust contain approximately 3 parts per million of ^{238}U, the long-lived head of the uranium series; 11 ppm of ^{232}Th, the head of the thorium series; but only about 0.02 ppm of ^{235}U, the long-lived member of the actinium group. The radon isotopes ^{222}Rn and ^{220}Rn are produced in proportion to the amount of the parent present. Some of the newly formed radon atoms which originate in or on the surface of mineral grains escape into the soil gas, where they are free to diffuse within the soil capillaries. Some of the radon atoms eventually find their way to the surface, where they become a part of the atmosphere. Even though thorium (^{232}Th) is generally more abundant than uranium in the Earth's crust, the probability for decay is smaller; hence, the production rate of ^{222}Rn and ^{220}Rn in the soil is roughly the same. Much of the ^{220}Rn decays before reaching the Earth's surface due to its short half-life.

When radon (^{222}Rn or ^{220}Rn) passes from soil to air, it is mixed throughout the lower atmosphere by eddy diffusion and the prevailing winds. Mean radon levels are found to be higher during those times of year when atmospheric stability is the greatest such as may occur during the fall months. Radon and its daughters play an important role in atmospheric electricity. Near the Earth's surface almost half of the ionization of the air is due to ^{220}Rn and ^{222}Rn and their daughter products. The alpha emitters from these chains typically produce about 10^7 ion pairs per second per cubic meter. *See* ATMOSPHERIC CHEMISTRY; ATMOSPHERIC ELECTRICITY.

Radon is readily soluble in water. Since ground and surface waters are in close contact with soil and rocks containing small quantities of radium, it is not surprising to find radon in public water supplies.

The radon isotopes ^{220}Rn and ^{222}Rn are used widely in the study of gaseous transport processes both in the underground environment and in the atmosphere. Radon accumulates to high levels of the order of 4000 becquerels/m^3 or more in caves unless natural or artificial ventilation occurs. Changes in ^{222}Rn concentrations in spring and well water and in soil and rocks have been suggested as a means of predicting earthquakes.

The tendency of the decay products of radon to attach to aerosols means that these nuclides will be inhaled and deposited in the bronchial epithelium and lungs. The daughter products, therefore, make up the major part of the internal radiation dose from radon. Ways of reducing radon levels within homes or workplaces include increased ventilation and sealing of major sources of entry from soil and building materials. Workers in uranium mines may encounter radon and decay product levels of the order of 50,000 Bq/m^3 or more. Ventilation procedures and special filters for the miners must be used. *See* BIOSPHERE; RADIATION INJURY (BIOLOGY); RADIOECOLOGY; RADIOISOTOPE (GEOCHEMISTRY).

M. Wilkening

Bibliography. W. Burkart et al. (eds.), *High Levels of Natural Radiation and Radon Areas: Radiation Dose and Health Effects by International Conference on High Levels of Natural Radiation and Radon*, Elsevier Science, 2002; D. R. Lide, *CRC Handbook of Chemistry and Physics*, 85th ed., CRC Press, 2004; W. W. Nazaroff and A. V. Nero (eds.), *Radon and Its Decay Products in Indoor Air*, Wiley-Interscience, 1988.

Rafflesiales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Rosidae of the class Magnoliopsida (dicotyledons). The order consists of three families and fewer than a hundred species, nearly all tropical or subtropical. The plants are highly specialized, nongreen, rootless parasites which grow from the roots of the host. They have few or solitary, rather large to very large flowers with numerous ovules and a single set of tepals that are commonly united into a conspicuous, corolloid calyx. *Rafflesia* is famous for its gigantic flowers, which in *R. arnoldii* of Sumatra are about 3 ft (1 m) in diameter. No other family has individual flowers even approaching this size. *See* MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM; ROSIDAE.

Arthur Cronquist

Railroad control systems

Those devices and systems used to direct or restrain the movement of trains, cars, or locomotives on railroads, rapid-transit lines, and similar guided ground-transportation networks. Such control varies from

the use of simple solenoid valves to fully automatic electronic-electromechanical systems.

A primary function of railroad control systems is to ensure the safe movement of trains. This is generally accomplished by providing train operators and track-side operators with visual indications of equipment status. The simplest form of control consists of track-switch-position indicators combined with track-side manually operated "stop" or "proceed" signals, which the train operator follows. Advanced systems incorporate fully automated train control, subject to human supervisory control and potential intervention when faults occur in automated systems.

Automatic block signaling. Block signaling significantly improves the safety of railroad operations. Automatic block signaling is accomplished by sectionalizing the track into electrical circuits to detect the presence of other trains, engines, or cars. Logic circuits in the control system detect the locations of the trains and the positions of switches, and then set the necessary signals to inform the train operators when to stop, run slowly, or proceed at posted speeds. The control system automatically detects the presence of a leading train, selects the signal to be given, and then sets the signal indications for the following train operators to read so that they may perform accordingly. If a train is within a block, its wheels and axles form a circuit with the rails. This produces a path of very low electrical resistance, and therefore the block signal receives negligible current. This low-current condition sets the signal at "stop" for any following train. In a block that is unoccupied, all track current passes through the signal circuit and sets the signal at "clear," allowing an approaching train to proceed.

Automatic train stop. In conjunction with automatic block signals, many subway rapid-transit lines incorporate automatic trip stops along the tracks to ensure that train operators obey the stop signals. If a train passes certain restrictive signals, such as a red light indication stop, the air brakes on the train are automatically actuated by a trip arm located alongside the tracks. The train cannot be restarted until the mechanism aboard the car is reset and the train operator goes through a safety startup procedure.

Automatic cab signaling. Automatic cab signaling systems display signaling information (traditionally, permitted speeds) on board the train. Coded information is transmitted to the train, generally via the running rails (**Fig. 1**). Antennas and receivers aboard the train pick up, amplify, decode, and distribute the intelligence, which then causes the proper signal aspects to be displayed in the cab. The display is mounted near the normal line of sight of the operator. Automatic cab signaling reduces or eliminates the need for wayside signals and improves the all-weather capability of trains and the train-handling capacity of the track.

Automatic train control. Automatic train control (ATC) subsystems, located wholly on board the train, sense whether or not the train is operating within safe speed limits. If it is not, automatic train control sets the brake to bring the train to a stop or to a speed below the allowed speed.

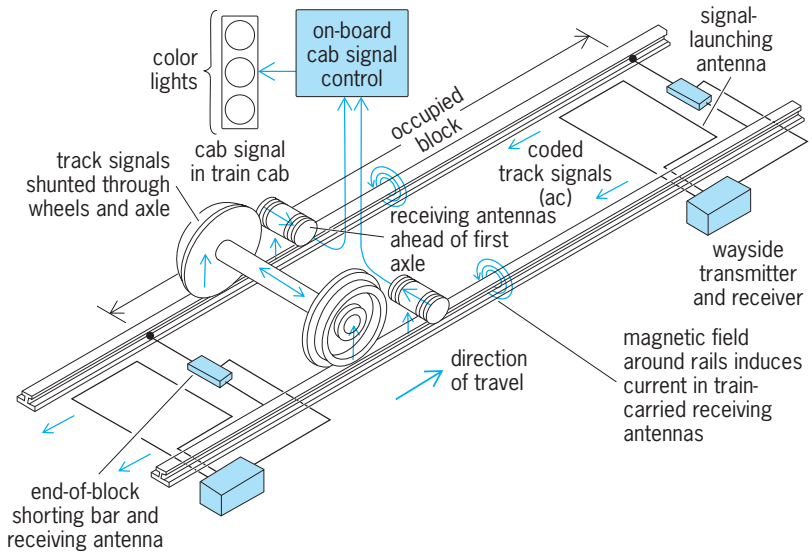

Fig. 1. Track circuit and cab signal using alternating-current track circuits.

Automatic cab signaling with automatic train control is used on many railroads and several rapid-transit lines in the United Sates and on systems in Europe and Japan.

Automatic train operation. Automatic train operation (ATO) subsystems perform nonvital operating functions such as starting, running at the prescribed speeds, slowing down, and stopping, and on some rapid-transit installations include passenger-door controls (**Fig. 2**). Automatic train operation

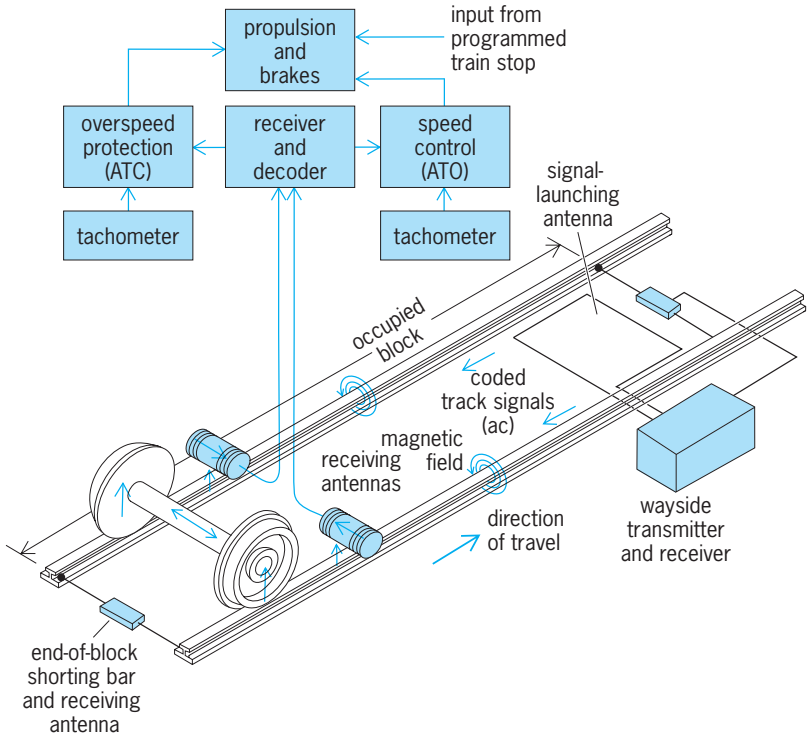

Fig. 2. Automatic train operation with fail-safe or vital overspeed protection and provision for programmed train stopping.

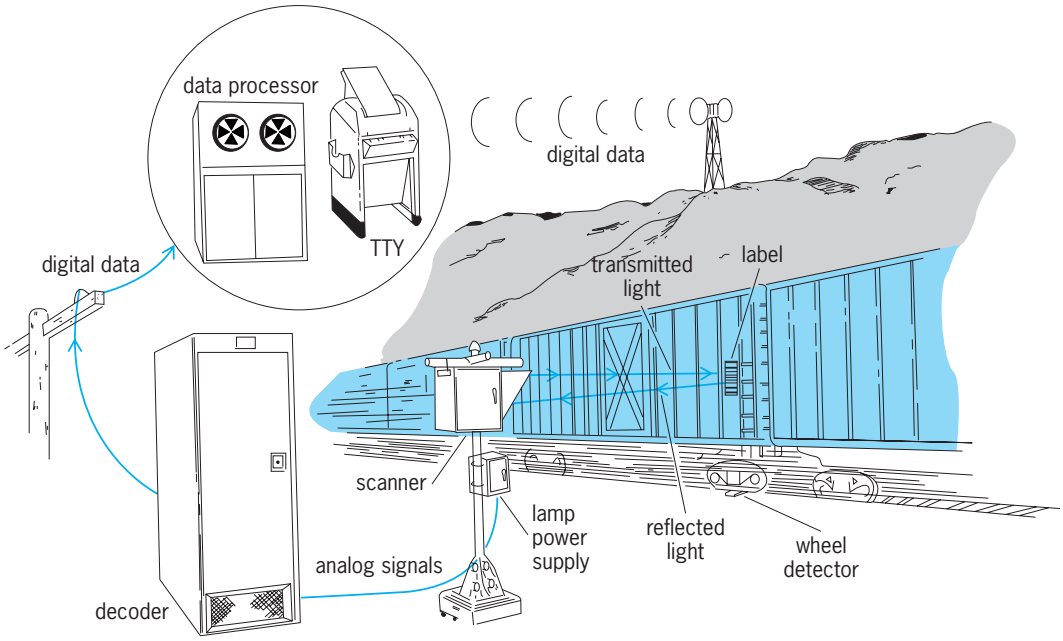

Fig. 3. Automatic car identification system.

builds upon the information transmitted to the train as part of automatic cab signaling, and is a logical next step in automating train operations.

The civil speed limits of the railroad are those imposed by the physical constraints of the railroad right-of-way, such as curves, grades, and the location of stations. Traffic speed restrictions supersede the civil speed limit when another train in the vicinity requires a lower, more restrictive speed than that allowed by the civil limit. Thus, a train approaching a slower train from behind will receive progressively more restrictive speed signals as it closes in. A speed of zero, the most restrictive speed signal, is used to stop the approaching train at a safe distance from the one ahead. In the case of cab signaling with automatic train control, the operator is responsible for obeying speed signals. With automatic train operation, the train automatically slows down, and comes to a complete stop when necessary. When automatic train operation systems are used, automatic train control (with or without cab signaling) is installed as a separate subsystem to act as a fail-safe backup for the automatic train operation.

Programmed train stop. Station stopping presents a special set of requirements for rapid transit, commuter railroads, and mainline railroad passenger operations. Accurate positioning of car doors at the station platform and smooth deceleration at relatively high rates are desirable for passenger comfort and efficient operating performance. A special subsystem of control referred to as programmed train-stop systems (or station-stop systems) are a combination of on-board and wayside electronic and electromechanical equipment that can bring a train to rest within inches of its stopping-point target.

Central line supervision. Efforts by railroads and transit operations to provide more efficient service is increasing demands on train dispatchers and schedulers. These demands are being met by a number of control systems designed to provide information about the trains and cars and to perform some of the less critical functions previously assigned to dispatchers and tower or central operators.

Car identification systems are one example of central line supervision. This system scans and decodes a series of colored and patterned lines placed on the side of each car to identify an individual car (**Fig. 3**). This information is transmitted to the operations area, where a computer system is used to establish routing, determine maintenance schedules, and so on. Dispatchers and central operators can also use computer workstations to obtain information on system status.

Highly automated railroad. The use of advanced technology and automation in railroad operations is growing. A number of systems employ the control subsystems described above. The San Francisco Bay Area Rapid Transit (BART) regional system has a fully automatic train-control system as well as computerized central line supervision. BART trains can operate at 80 mi/h (128 km/h) at 90-s headways. Similar systems have been installed worldwide, including those in São Paulo, Brazil, and Washington, D.C. Japan's Tokaido Bullet Trains, running at 125 mi/h (200 km/h), are too fast for wayside signals to be read. Cab signals are used, and automatic train control prevents overspeeding.

Higher performance standards are written into specifications for each new system and system upgrade, making it clear that the trend is toward complete systemwide automatic train control. Safety, reliability of higher-speed performance, availability of equipment, return on investment, and other factors build a good case for automation.

Terminal control systems. Railroad terminals (points of origin and destination of trains) are

critical to the efficient, cost-effective operation of railroads, so they represent a major focus for automation.

A terminal generally contains three types of yards: a receiving yard, where incoming trains from the main line are temporarily stored; the hump yard, where cars are classified and resorted into new trains; and a departure yard, where trains are assembled and stored for dispatch onto the main line.

Gravity or hump yards. These have evolved as an efficient method of sorting and reforming large trains. Typical trains may contain 100–200 units.

As the train moves into the hump yard to be classified, cars or coupled assemblies of cars (called cuts) are moved to the crest of a small hill, or hump, beyond which they roll by gravity onto the appropriate classification track. The steady, one-direction movement allows a single locomotive to move three to five cars per minute.

The hump yard presents interesting control problems. The car or cut is given an initial velocity by the locomotive. Ideally, the rolling cut should couple with the next car standing on the same classification track at a speed of 2–4 mi/h (3–6 km/h). Too low a speed may prevent the couplers from engaging properly. Too high a speed can damage the lading or rolling stock.

The initial speed given to each cut must reflect the physical characteristics of the cut, the track and geometric conditions, and the distance that the cut must travel. Rolling performance of the cut is affected by a number of physical characteristics, including weight, type, and condition of the bearings, the center of gravity, and the cross-sectional area. The track and geometric conditions that impact rolling performance include curves, track condition, and the grade (slope) of the track. The distance that the car must travel on the selected classification track depends not only on the length of the track but also on the number of cars already stored there.

The speed of cuts within the hump yard is controlled primarily by the use of remotely controlled retarders which act as external breaks on the car wheels to slow the cut to a desired speed. The cuts are directed into the proper classification track by manipulation of remotely controlled track switches.

Hump yard control. Automatic control of the hump operation has many advantages. Operators can maintain a high output with moderate effort and training. More importantly, automation reduces damage claims, derailments, and misclassifications.

Advanced automatic controls for hump yards involve digital process control computers combined with automatic car-identification systems. The control system accepts input information about the cars, such as their destination and characteristics, and inputs from the yard, such as the speeds of the cars, the number of vehicles already on the classification tracks, and the positions of the switches.

The digital process control computer provides major advantages over older systems that employ automatic relay and analog computer systems. The use of digital process control techniques in hump yards opens up the entire terminal to automation, including control of switches and signals in the receiving and departure yards, as well as the handling of terminal operating paperwork and accounting functions. *See* CONTROL SYSTEMS; RAILROAD ENGINEERING; TRAFFIC-CONTROL SYSTEMS.

James Costantino; Donna C. Nelson

Bibliography. American Railway Engineering Association, *AREA Manual of Railway Engineering*, annually; American Railway Engineering Association, *The Signal Manual of American Railroads*, 4 vols., 1984; F. George, How a digital computer controls a class yard, *Railway Sig. Commun.*, 60(12):24–31, 1968: *Railway Age Mag*. pp. 50–51, September 29, 1975.

Railroad engineering

A branch of engineering concerned with the design, construction, maintenance, and operation of railways. Railway engineering includes elements of civil, mechanical, industrial, and electrical engineering. It is unique in being concerned with the interaction between moving vehicles (mechanical engineering) and infrastructure (civil engineering). The employment of both a load-supporting guideway and groups or strings of connected vehicles on flanged wheels for the transport of goods and people sets railways apart from other modes of transport.

Track geometry. Components of track geometry are alignment, curvature, and track gage.

Horizontal alignment and curvature. The plan view of a railroad track is known as the horizontal alignment. It is made up of a series of curves (arcs of simple circles), tangents (straight tracks), and spirals joining the curves and tangents. Deviations from any of the three are flaws. These imperfections are corrected periodically by a technique known as lining the track.

The radii of simple railroad curves are usually very large, so that the centers of the circles of which the curves are an arc are frequently inaccessible from the track. For convenience, measurements of curve severity and regularity are made in terms of midchord offsets (**Fig. 1**). In North America, curves are described in terms of degrees. The degree of curve is the number of degrees of central angle subtended by a 100-ft chord. It is also equal to the midchord offset, in inches, of a 62-ft chord. Outside of North America and on some transit properties, curves are described in terms of their radii.

On mainline tracks the outer rail of curves is superelevated above the inner rail to compensate for centrifugal force. The amount of superelevation depends on train speeds and the degree of curvature.

Vertical alignment. The side or elevation view of track, composed of a series of straight portions and the vertical curves joining them, is known as the vertical alignment. The vertical change in elevation, in feet, over a horizontal distance of 100 ft is the percent grade. Because the friction coefficient of steel wheels on steel rails is low, railroad grades must also be low, with values from zero to 1.5% fairly common.

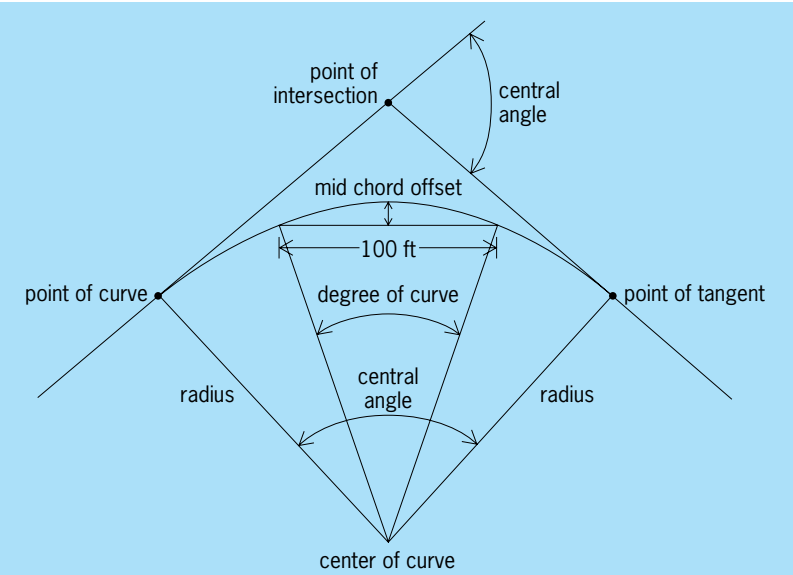

Fig. 1. Simple railroad curve without spirals.

Two-percent grades are severe, usually requiring helper locomotives. Grades that are more severe, up to about 4%, can be surmounted only with considerable extra operating care and at significant additional expense. The macro view of elevation is the profile of the railroad, while the micro view is the surface. Irregularities on the surface are corrected by raising the low spots and packing ballast under the raised track.

Gage. Track gage is the distance between rails. The standard gage throughout the world is 4 ft 8.5 in. Narrow gages of 3 ft and broad gages of 5 and 6 ft have all been tried at various times and in different places. The reasons are diverse and include attempts to reduce construction cost, to intentionally make rolling stock of other countries incompatible, and to hinder attempts at invasion, as in Russia where broad gage prevails.

Track components. These include rails, welded rails, crossties, and ballast.

Rails. The function of rail is to guide wheels and distribute their vertical and lateral loads over a wider area. Neither cast nor wrought iron was ideally suited to this task. The development of the Bessemer furnace to produce steel that was three times harder than wrought iron at reasonable cost made it possible for the weight of vehicles and therefore the productivity of railways to increase. Unfortunately Bessemer steel hardness was only half that of cast iron, causing more frequent renewal, especially on heavy curves. Nonetheless, steel was so vastly superior overall that it quickly supplanted its predecessors. *See* CAST IRON; STEEL; STEEL MANUFACTURE; WROUGHT IRON.

For a given shape the moment of inertia of rail (and its resistance to bending) is proportional to its cross-sectional area. Rail size was rapidly increased from 30 or 50 lb per yard of length (15 or 50 kg/m), to a maximum of 155 lb/yd (77 kg/m) as engineers attempted to reduce the cost of maintaining proper vertical and horizontal alignment of track subject to heavier loads and higher speeds. *See* MOMENT OF INERTIA.

In attempting to find the optimum distribution of material to resist deflection, wear and service failures, and to facilitate economic manufacture, the shape and contour of rail sections have undergone many adjustments.

As competitive pressure heightened, freight-hauling railroads in the United States tended to increase the proportion of 100-ton (90-metric-ton) carrying capacity cars. Some special-purpose noninterchangeable cars, with even higher axle loads, have become common as well. Speeds are also increasing in both passenger and freight traffic. This has led North American railroads to search for stronger and harder steel. Fine-grain pearlitic steel (high-strength steel) with tensile strengths of approximately 150,000 lb/in.^{2} (10 gigapascals) and hardness in the 340–388 Brinell range is being produced to meet this need. *See* HARDNESS SCALES.

Welded rails. Rails are joined end to end by butt welding, whereby continuous rails of over 1000 ft (300 m) in length can be produced. When laid in track, the rails are heavily anchored to restrain movement due to temperature changes. The stresses produced by temperature changes are well below the strength of the steel.

Crossties. Crossties play important roles in the distribution of wheel loads vertically, longitudinally, and laterally. Each tie must withstand loads up to one-half that imposed on the rail by a wheel. The crosstie must then distribute that load to the ballast surrounding it.

Timber crossties, varying in section from 6 in. × 6 in. (15 cm × 15 cm) for the lightest applications to 7 in. × 9 in. (18 cm × 23 cm) for heavy-duty track and in length from 8 ft 6 in. (2.6 m) to 9 ft (2.7 m) in length, are sawed in highly automated saw mills. Ties may be prebored for fastenings; and to prevent splitting, irons, dowels, or nailplates are frequently applied. Air drying or artificial seasoning is necessary before pressure treatment with preservatives, most often creosote–coal tar or creosote–petroleum mixtures. To protect timber crossties from crushing, it is standard practice to employ steel tie plates between rails and ties to further spread train loads beyond that accomplished by the wide rail base.

Well-treated hardwood ties such as red oak, locust, beech, hickory, and fir in well-maintained track may be expected to last 30 years or more. Timber crossties become unserviceable after time because of splitting, decay, insect attack, center cracking, mechanical wear, and crushing. Tie life varies with species, weather, traffic, track condition, and geometry.

In North America, crossties are replaced selectively on a periodic basis. At certain intervals, all the crossties that are considered badly deteriorated, but only those, are renewed. In other parts of the world, all the crossties in a given length of track are replaced along with the rail and ballast, and those components considered to have substantial remaining life are

recycled by installing them in lighter traffic lines, yard tracks, or sidings.

Prestressed concrete monoblock crossties are standard in Great Britain and parts of continental Europe. In North America, they are less extensively used than timber. Concrete crossties are used extensively in the high-speed tracks of Amtrak's Northeast Corridor. Several western railroads employ them, especially in mountainous terrain where heavy curvature requires frequent rail renewal or transposition, the practice of interchanging the inner and outside rail of curves to equalize wear. Rail fasteners used with concrete ties may be removed and reinstalled easily without damage to the crosstie. Tests conducted by the Association of American Railroads (AAR) have shown a significant increase in track stiffness in concrete tie track over timber tie track of otherwise similar construction and material.

Ballast. The granular material that supports crossties vertically and restrains them laterally is known as ballast. Ideal ballast is made up of hard, sharp, angular interlocking pieces that drain well and yet permit adjustments to vertical and horizontal alignment. Materials that crush and abrade, creating fines that block drainage or that cement, should not be used. Soft limestones and gravel, including rounded stones, are examples of poor ballast, while crushed granite, trap rock, and hard slags are superior.

The ballast must distribute crosstie load over an area sufficiently wide so that the underlying subsoil will not be overstressed. Tests conducted by the AAR during 1988–1991 showed that a 31,500-lb (14,200-kg) wheel load generates an average pressure of 37 lb/in.^{2} (260,000 Pa) at the bottom of a tie and that 18 in. (45 cm) of ballast spreads this over sufficient area to reduce its magnitude to about 10 lb/in.^{2} (70,000 Pa). Depending on the nature of the subsoil a depth of 12 to 18 in. (30 to 48 cm) beneath the bottom of the tie is usually adequate. A subballast of a fine-grain, well-draining material such as cinder or sand may be used beneath the stone or slag top ballast.

Track stiffness. A loaded wheel depresses the track it rolls on, and lifts the track ahead and behind. The depression is greatest directly below the wheel. Thus, as the wheel moves it creates a wave action in the track. A characteristic of track representing its stiffness is the modulus of elastic track support, frequently called the track modulus. This quantity is determined mathematically and is used to calculate track depression. Track modulus is more significant than rail stiffness in resistance to track depression under load. Because a rolling wheel is always climbing out of its own depression, track stiffness is an important contributor to train resistance and therefore energy conservation. There is a limit to the amount of stiffness desirable, for as stiffness increases, the attenuation of dynamic load decreases, as well as the spreading of wheel load. A value of track modulus approaching 10,000 lb/(in.)(in.) [70,000 newtons/(m)(m)] is considered close to ideal, given current suspension designs and track geometries.

Train resistance. Railroads were developed to reduce the effort required to move people and goods. However, a train of freight or passenger cars does resist movement. This resistance is the result of a combination of factors.

The most critical factor is gravity. Locating engineers seek the lowest possible grade consistent with the intended route. Practices such as following river valleys and stream beds, and adding length to reduce the percent grade in mountainous territory are common. Grade resistance, estimated at about 20 lb/ton (10 kg/metric ton) of train weight for each percent of grade, means a 10,000-ton (9,000-metric ton) train requires 200,000 lb (90,000 kg) of force just to overcome gravity on a 2% grade.

The second factor is the track deformation. Other factors may be grouped into the category of friction. There is the friction of axles turning in their journals, and of wheel flanges rubbing the sides of rails. The conical shape of wheels is intended to make them roll straight with respect to the track and to be self-correcting when they veer away from that path. However, irregularities of track alignment and wear of wheels and curves make following a true course without flange contact impossible. At certain speeds the wheels follow a zigzag path, striking first one rail and then the other. This is called truck hunting, and it consumes energy.

As a train negotiates curved track, the flanges bear against the outer rail, causing wear of both rails and flanges and consuming additional energy estimated at about 8 lb/ton (4 kg/metric ton) per degree. This form of resistance is sufficiently severe that when sharp curves occur on grades the grade is frequently reduced throughout the curve to compensate for the added resistance. Such grades are known as compensated grades.

Many of the friction factors vary directly with speed. Aerodynamic resistance, the last factor, varies with the square of speed. Up to speeds of 40–50 mi/h (24–30 km/h) it is not critical, but at speeds above these values it assumes considerable importance.

Locomotives. Force must be applied to the leading coupler of a train to overcome the combination of resistances. To supply this force, locomotives convert some form of potential energy to produce useful work.

A locomotive's performance, measured in a number of ways, begins with its tractive effort, and in particular, its starting tractive effort (the force it can apply to a standing train). The locomotive's starting tractive effort is usually equated to its weight on the driving wheels times the coefficient of friction of those wheels on steel rail. This coefficient can vary from about 0.15 to about 0.30, depending on the rail surface condition. The most modern locomotives employ electronic slip detection and control to enhance their starting tractive effort, and almost all locomotives use sand when necessary to increase the coefficient of friction.

Once a locomotive is moving, the tractive force that it can apply to its train once the train is moving is a function of speed, usually decreasing as speed

increases. Tractive force varies with speed for several types of locomotives. The locomotive's power is the product of its tractive force at a given speed and that speed. The difference between the locomotive's power and the train's resistance is the power available for acceleration or for overcoming additional resistance. Generally, train resistance increases with speed, especially when aerodynamic resistance is a factor. Therefore, a speed known as the balancing speed is attained where the locomotive power equals train resistance, and the entire power capability of the locomotive is consumed just to maintain speed. Any increase in resistance will slow the train until a new balance is reached.

Other significant factors in rating locomotive performance are thermal efficiency, reliability, and availability (time available for duty versus time in shops for repair or service). For all of these, and especially availability, the diesel-electric locomotives, which first appeared in the late 1930s, surpassed steam. In addition, since diesel-electric power does not require the proliferation of expensive servicing facilities such as coaling and watering stations, by the early 1960s the transition to diesel was complete.

But even diesel-electric locomotives have limitations. The exhaust gases containing oxides of nitrogen and sulfur have become environmental concerns. A longstanding difficulty has been the expense of maintaining diesel-electric traction motors and especially their commutators. *See* COMMUTATOR; DIESEL ENGINE.

Electric traction motors exert their maximum torque at starting and slow speeds. This characteristic is desirable for railroad applications, since it is in starting that train resistance is greatest (until aerodynamic resistance takes over). Tractive effort curves showing tractive force versus speed are useful in rating locomotives for various types of service and for developing computer-generated simulations of train performance. By matching tractive effort curves against those for train resistance, these programs predict elapsed time, braking distance, and fuel consumption.

The earliest diesel-electric switching locomotives developed about 600 horsepower and road freight units 1350 hp. Single locomotive units of 4000 hp are common. Common practice since dieselization began has been to employ a number of locomotive units coupled together to form a single, more powerful power source.

Locomotive development includes designs using liquefied natural gas as a fuel to reduce environmental pollutants. Locomotives equipped with the necessary power conditioning equipment and squirrel-cage asynchronous motors, made possible by the advent of high-capacity solid-state electronics, exhibit superior adhesion and have no troublesome commutators. They are adept at hauling heavy-tonnage mineral freight, fast passenger trains, or high-speed merchandise trains (freight trains that haul primarily high-value merchandise, as opposed to low-cost raw materials such as coal or grain) [**Fig. 2**]. *See* LOCOMOTIVE.

Freight services. Until World War II, several general-purpose cars such as open hoppers, open gondolas, closed box, and flat cars, which any shipper could load with any commodity, met most needs. However, since then a number of special-purpose cars were designed which are frequently assigned to or even owned by individual shippers or third parties, such as open and covered hopper cars for coal or grain, insulated and refrigerator cars for fresh food and beverages, and bi- and trilevel cars for automotive products.

The high level of productivity of railroads in America largely has been made possible by the semiautomatic coupler, invented to replace the dangerous link-and-pin system, and the Westinghouse air brake. The formation of the Master Carbuilders' Association, a forerunner of the AAR, set standards for railway cars and permitted them to move freely on all lines, thereby obviating the need to transfer lading from one car to another at railroad junctions. The AAR administers the interchange agreement under which cars and car components are tested and certified for free interchange.

Air brakes permitted much longer trains to move safely and eliminated the need for brake personnel to leap from car to car to set manual brakes. Janney couplers facilitated the making up and breaking of long trains. They also introduced slack of about 6 in. (15 cm) per car so that long trains could be started by steam locomotives by overcoming starting resistance one car at a time. *See* AIR BRAKE.

Both of these attributes have come to be viewed as mixed blessings. Free slack together with the delay in propagating pneumatic signals to actuate brakes in very long trains can lead to longitudinal train action and dynamic shock loads. The use of semipermanent solid drawbars, which eliminate longitudinal slack, is preferred in solid trains where all cars have a single origin and destination. Such trains have increased, but there is a tendency for other trains to bypass as many switching yards as is feasible.

The fastest-growing segment of freight business in North America is intermodal traffic: highway trailers mounted on flat cars and containers (both domestic and overseas) mounted either on flat cars or specially

Fig. 2. Locomotive with ac asynchronous traction motors. (*General Motors*)

Fig. 3. Articulated double-stack container train, showing single four-wheel truck under one end each of two adjacent cars, and containers stacked two high. (*Gunderson, Inc.*)

designed container cars. Many of these container cars are actually a series of cars permanently connected by articulated joints in which a single truck supports one end each of two adjacent cars. The platforms of these cars have depressed centers or wells, so that two containers can be stacked on top of one another without violating certain height restrictions (**Fig. 3**).

Passenger services. Rail passenger systems such as the Shinkansen in Japan (1964); TGV (Très Grande Vitesse; 1981) in France; ICE (Inter City Eisenbahn; 1991) in Germany; X-2000 in Sweden (1990); or Britain's several High-Speed Intercity Trains are notable for speed and convenience, as well as magnetic levitation. These systems have developed in context of an awareness of deteriorating highway infrastructures, serious concern with the air pollution generated by automobiles and trucks, increasing traffic congestion in urban areas, and worries over petroleum usage and supply. *See* AIR POLLUTION; MAGNETIC LEVITATION.

Therefore questions have arisen concerning the possibility of high-speed rail passenger service in the United States. The answerers are many, complex, and relate only indirectly to technology. Using the accepted criteria of 125 mi/h (200 km/h) to qualify as high speed, there are no less than 18 such weekday Metroliners operated between New York and Washington. These trains are powered by 5300-continuous-horsepower class AEM-7 rectifier electric locomotives capable of 7600 hp for short periods. The 4000-hp Genesis diesel-electric locomotives (**Fig. 4**) are also in use. In addition, Amtrak successfully tested in revenue service both the Swedish X-2000 tilting train and the German ICE at speeds up to 140 mi/h (220 km/h).

Fig. 4. The genesis, 4000-hp passenger locomotive capable of speeds up to 103 mi/h (165 km/h). (*Amtrak*)

The lack of extensive services in the United States outside the Northeast Corridor can be partly explained by demographic factors. Population centers in both Japan and Europe are typically closer to each other than in North America, making construction costs for dedicated high-speed trackage more feasible and travel times far more competitive with air travel. In addition, high-speed passenger trains require significantly different track configurations (for example, curve superelevation and turnout designs) than do slower freight trains, even those high-valued merchandise or intermodal highway trailers and containers, which frequently travel at speeds of 70 mi/h (110 km/h). These engineering differences are impractical on lines primarily moving heavy mineral freight, where axle loadings and speeds differ even more radically from high-speed passenger trains. The United States freight lines are also frequently burdened with highway grade crossings that are incompatible with passenger trains traveling at speeds well over 100 mi/h (160 km/h). In spite of all these difficulties, advances in the caliber of passenger service continue at a rapid pace in and outside the Northeast Corridor, particularly in the area of suburban commuter rail.

Environmental issues. Railroads are considered to be three to four times more energy efficient than other forms of land transport, with lower air contaminates per ton-mile for diesel-electric powered trains. The goal is to reduce the level of the oxides of nitrogen and sulfur even further. Railroads are the one form of heavy-freight transport that can be electrified by using available technology.

One of the more serious environmental concerns of railroads is posed by potential spills and leaks of hazardous materials. As a result of engineering studies, many steps such as head shields and special nonoverriding couplers on tank cars have already been taken, and further studies are under way to improve the integrity of these cars. *See* RAILROAD CONTROL SYSTEMS; TRANSPORTATION ENGINEERING.

George H. Way

Bibliography. American Railway Engineering Association, *Manual of Recommended Practice*, 1993, supplemented annually; American Railway Engineering Association, *Stresses in Railroad Track—The Talbot Reports*, 1980; H. I. Andrews, *Railway Traction*, 1986; J. H. Armstrong, *The Railroad, What It Is, What It Does*, 3d ed., 1990; W. W. Hay, *Railroad Engineering*, 2d ed., 1982; V. A. Profillidis, *Railway Engineering*, 2d ed., 2000; *Railway Age Magazine*, pp. 48–53, 55–77, January 1994.

Rain shadow

An area of diminished precipitation on the lee side of mountains. There are marked rain shadows, for example, east of the coastal ranges of Washington, Oregon, and California, and over a larger region, much of it arid, east of the Cascade Range and Sierra Nevadas. Precipitation on the northern Oregon coast is around 100 in. (2.5 m) per year; east of the coastal ranges in the Willamette Valley it is about 40 in. (1.0 m); and east of the Cascades it drops to around 10 in. (0.25 m). All mountains decrease precipitation on their lee; but rain shadows, as is shown by annual totals, are sometimes not marked if moist air often comes from different directions, as in the Appalachian region.

The causes of rain shadow are (1) precipitation of much of the moisture when air is forced upward on the windward side of the mountains, (2) deflection or damming of moist air flow, and (3) downward flow on the lee slopes, which warms the air and lowers its relative humidity. *See* CHINOOK; PRECIPITATION (METEOROLOGY). J. R. Fulks

Rainbow

An optical effect of the sky formed by sunlight falling on the spherical droplets of water associated with a rain shower. The circular arc of colors in the rainbow is seen on the side of the sky away from the Sun. The bright, primary rainbow shows the spectrum of colors running from red, on the outside of the bow, to blue on the inside. Sometimes a fainter, secondary bow is seen outside the primary bow with the colors reversed from their order in the primary bow. The shape of each bow is that of a circle, centered on the antisolar point, a point in the direction exactly opposite to that of the Sun, which is marked by the shadow of the observer's head (see **illus.**).

As a light ray from the Sun strikes the surface of a water drop, some light is reflected and some passes through the surface into the drop. The primary bow results from light that enters the drop, reflects once inside the drop, and then leaves the drop headed toward the observer's eye. Light that is reflected twice inside the drop produces the secondary bow. The change of direction that occurs when a light ray enters or leaves the waterdrop (refraction) is different for the different colors that make up white sunlight. As a result, the size of the circle is different for each color, thereby separating the colors into the rainbow sequence. *See* METEOROLOGICAL OPTICS. Robert Greenler

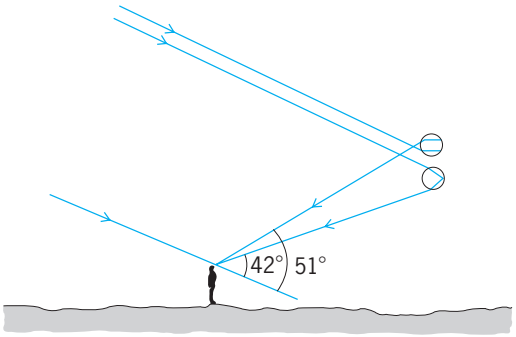

Raindrops contribute to the primary rainbow (with one internal reflection) and to the secondary rainbow (with two internal reflections). (*After R. Greenler, Rainbows, Halos and Glories, Cambridge University Press, 1980*)

Rainforest

Forests that occur in continually wet climates with no dry season. There are relatively small areas of temperate rainforests in the Americas and Australasia, but most occur in the tropics and subtropics.

Occurrence. The most extensive tropical rainforests are in the Americas. These were originally 1.54×10^6 mi^2 (4×10^6 km^2) in extent, about half the global total, and mainly in the Amazon basin, extending north to the Orinoco basin and the Guyanas. A narrow belt also occurs along the Atlantic coast of Brazil, and a third block lies on the Pacific coast of South America, extending from northern Peru to southern Mexico.

The eastern tropical rainforests (970,000 mi^2 or 2.5×10^6 km^2) represent the second most extensive block. These are centered on the Malay archipelago, extending toward Queensland (Australia), the western Pacific islands, and continental Asia; west to the western Ghats of India; and north to Xishuangbanna in southernmost China (see **illus.**).

Aerial view of a typical rainforest in Windsor tableland, Queensland, Australia.

The African tropical rainforests, originally 700,000 mi^2 (1.8×10^6 km^2), have the least extent. They are situated mainly in the Zaire basin, with outliers in east Africa, and extending along the coast of west Africa, where they have now largely been cleared for farms.

Characteristics. Tropical rainforests have a continuous canopy (commonly 100–120 ft or 30–36 m tall) above which stand huge emergent trees, reaching 200 ft (60 m) or even taller. Within the rainforest canopy are trees of many different sizes, including pygmies, that reach only a few feet. Trees are the main life form and are often, for purposes of description and analysis, divided into strata or layers. Trees form the framework of the forest and support

an abundance of climbers, orchids, and other epiphytes, adapted to the microclimatic conditions of the different zones of the canopy, from shade lovers in the gloomy, humid lower levels, to sun lovers in the brightly lit, hotter, and drier upper levels. Many trees have buttresses, and some bear their flowers directly on the trunk (cauliflory). Many leaves, especially in the lower part, have long extended drip tips. Most trees have evergreen leaves, many of which are pinnate or palmate. These features of forest structure and appearance are found throughout the world's lowland tropical rainforests.

There are other equally distinctive kinds of rainforest in the lower and upper parts of perhumid tropical mountains, and additional types on wetlands that occur either as permanent or periodic swamp forests. Peat swamp rainforests can be found in Asia and in smaller areas elsewhere. These forests and others that exist on limestone and ultramafic rocks form a total of about a dozen distinct types, commonly called formations. Each type is associated with a particular site and is usually, therefore, sharply bounded.

Climate. Rainforests occur where the monthly rainfall exceeds 4 in. (100 mm) for 9–12 months. They merge into other seasonal or monsoon forests where there is a stronger dry season (3 months or more with 2.5 in. or 60 mm of rainfall). The annual mean temperature in the lowlands is approximately 64°F (18°C). There is no season unfavorable for growth.

Primary and secondary communities. Rainforests, like other vegetation, are dynamic; the canopy is in a continual state of flux. As trees die, singly or in groups, gaps develop and new trees grow from seedlings to fill them. Thus, the forest consists of a mosaic of patches at different stages of a growth cycle, and any layering (stratification) of tree crowns is continually changing as trees mature and eventually die. Most species have seedlings that can grow in the shade on the forest floor, filling in the smaller gaps of the canopy. Thus, the species maintain themselves from generation to generation, and although different species may fill any given gap, replacing what was there before, overall the canopy composition remains the same. Such species are often called climax species.

Where a section of rainforest is cleared, or if a natural catastrophe (such as a cyclone, landslide, or volcanic mudflow) occurs, much larger gaps in the canopy can result. However, these gaps are colonized by a different set of tree species that grow up as a single canopy layer from seed after the creation of the gap. These species are adapted to the hotter, drier, brighter conditions of large gaps, and are commonly called pioneers, and they form secondary forest. Their seedlings cannot establish in shade, and pioneers are replaced as they die by climax species, thereby restoring the original or primary forest. The complete succession may take several hundred years.

Primary rainforests are exceedingly rich in species of both plants and animals. There are usually over 100 species of trees 2.5 in. (10 cm) in diameter or bigger per 2.4 acres (1 ha). The richest forests yet known (in Kalimantan, Peru and Ecuador) have over 250 species. There are also numerous species of climbers and epiphytes.

In addition, flowering and fruiting occur year-round, but commonly there is a peak season; animal breeding may be linked to this. Secondary rainforests are much simpler. There are fewer tree species, less variety from location to location, and fewer epiphytes and climbers; the animals are also somewhat different.

Total clearance of rainforest for pasture or for growing food crops is followed by species-poor, structurally simple secondary forest when habitation ceases. However, extraction of timber causes less severe damage by creating gaps in the canopy. Damage can be minimized, and the forest can recover, remaining a rich mixture of climax species. Where damage is heavier and large canopy gaps are formed, secondary forest fills them. Thus, when logging-induced disturbances mimic natural gap creation, less alteration of forest species composition and structure occurs. *See* ECOLOGICAL SUCCESSION.

Useful products. Tropical rainforests are a source of resins, dyes, drugs, latex, wild meat, honey, rattan canes, and innumerable other products essential to rural life and trade. Modern technology for extraction and for processing has given timber of numerous species monetary value, and timber has come to eclipse other forest products in importance. The industrial nations use much tropical hardwood for furniture, construction, and plywood. Rainforest timbers, however, represent only 11% (1×10^{10} ft^3 or 284×10^6 m^3) of world annual industrial wood usage, a proportion that has doubled since 1950. West Africa was the first main modern source, but by the 1960s was eclipsed by Asia, where Indonesia and Malaysia are the main producers of internationally traded tropical hardwoods. Substantial logging has also developed in the neotropics. For example, since the 1980s parts of the Brazilian Amazon have been heavily exploited for the south Brazil market; and there is also an export trade in mahogany. In Asia, second to timber in value are the rattans, climbing palms largely confined to these eastern rainforests. *See* FOREST TIMBER RESOURCES.

Tropical deforestation and climatic change. Tropical deforestation is the source of about one-quarter of the annual increase of atmospheric carbon dioxide, or about 7300 million tons. This net flux to the atmosphere is the sum of carbon dioxide from burning and decay of residues, oxidation of soil humus, and slow oxidation of harvested products, offset by carbon fixation in vegetation that regrows on the site. (Three-quarters of annual carbon dioxide increase, however, comes from burning fossil fuel.) In addition to an increase in the production of carbon dioxide, deforestation adds methane and nitrous oxide to the atmosphere. These three gases together contribute 75% of the annual radiative forcing, the greenhouse effect [the rest comes from synthetic chloro (fluoro) carbons], and 20–25% of the greenhouse effect is attributable to deforestation. *See* GREENHOUSE EFFECT.

Conservation. Tropical forests as a whole are estimated to have disappeared at a rate of 0.8% or 37 × 10^6 acres (15.4 × 10^6 ha) per year during the 1980s, and much of this was rainforests. Perhaps half as much rainforest has been logged but left to regenerate. For this reason, there is worldwide concern regarding the destruction of rainforests. Apart from esthetic and moral issues, numerous potentially valuable species are being lost or threatened. Countries with rainforests have set aside 5–10% of this forest for national parks; however, not all are ideally sited to be adequately representative. Given the high species richness, and the very sparse occurrence of some species (for example, carnivorous animals at the top of the food chain), the national park network is inadequate. Another approach has been to skillfully conduct logging and extraction procedures of other useful products in a manner that maintains the forest ecosystems and plant and animal habitats. *See* CONSERVATION OF RESOURCES; FOREST ECOSYSTEM.

T. C. Whitmore

Bibliography. L. A. Bruijnzeel, *Hydrology of Moist Tropical Forests and Effects of Conversion: A State of Knowledge Review*, 1990; A. Gomez-Pompa, T. C. Whitmore, and M. Hadley (eds.), *Rain Forest Regeneration and Management*, 1990; M. D. Swaine and T. C. Whitmore, On the definition of ecological species groups in tropical rain forests, *Vegetatio*, 75:81–86, 1988; J. Terborgh, *Diversity and the Tropical Rain Forest*, 1992; T. C. Whitmore, *An Introduction to Tropical Rain Forests*, 2d ed., 1998; T. C. Whitmore, *Tropical Rain Forests of the Far East*, 2d ed., 1984.

Rajiformes

One of four orders making up the Batoidei. Aside from the few exceptions noted below (mostly in family Rajidae), the combination of characters that distinguish the order from other batoids are a single unbranched rostral process on front of the cranium, extending to the tip of the snout; antorbital (in front of the eye socket) cartilage that does not extend forward to help support the anterior part of the disc; two dorsal fins on the tail; lack of spines on the tail; lack of electric organs; a snout that is not a long flat blade; a head lacking anterior finlike expansions of pectoral fins; and ovoviviparous reproduction. Their food consists of various crustaceans, mollusks, and small fishes.

Order Rajiformes
- Rhinidae (genus *Rhina*)
- Rhynchobatidae (*Rhynchobatus*)
- Rhinobatidae (*Aptychotrema, Rhinobatos, Trygonorrhina, Zapteryx*)
- Rajidae (18 genera)

Rhynchobatidae. The general shape of these batoids is intermediate to typical sharks and typical rays. They are distinguished by a moderately flattened body, and narrow, pointed snout with the length in front of the eyes about equal to the width of the head at the eyes. The caudal fin is bilobed, with both lobes sharp-pointed. The dorsal fins are relatively large; the origin of the first is about over midpoint of the pelvic fins. The rear margin of pectoral fins do not extend posteriorly to the origin of the pelvic fins. The back is sometimes variously marked by small white spots. Their habitat is shallow inshore waters and estuaries.

The family consists of one genus, *Rhynchobatus*, and four species, all of which occur in the western Indian Ocean from the Red Sea to South Africa and east to Malaysia and the Philippines and north to Japan.

Rhinidae. This family, comprised of only one species, *Rhina ancylostoma*, occurs in the tropical Indian Ocean to Malaysia, Southeast Asia, and Australia. It is similar to species of Rhynchobatidae, but differs in having a short, broadly rounded snout, its length in front of the eyes being less than the width of the head at the eyes; and large spines on the head and shoulders. It inhabits coastal areas over substrates of mud and sand, as well as coral reefs.

Rhinobatidae (guitarfishes). In body form, these batoids are between sawfishes and skates, and they vary from a relative flat wedge-shaped body sector (for example, *Aptychotrema* and *Rhinobatos*, see **illustration**) to a broadly rounded body sector and stout tail sector (for example *Zapteryx*) to an almost circular body sector and long narrow tail sector (that is, banjo shaped, such as *Trygonorrhina*). The origin

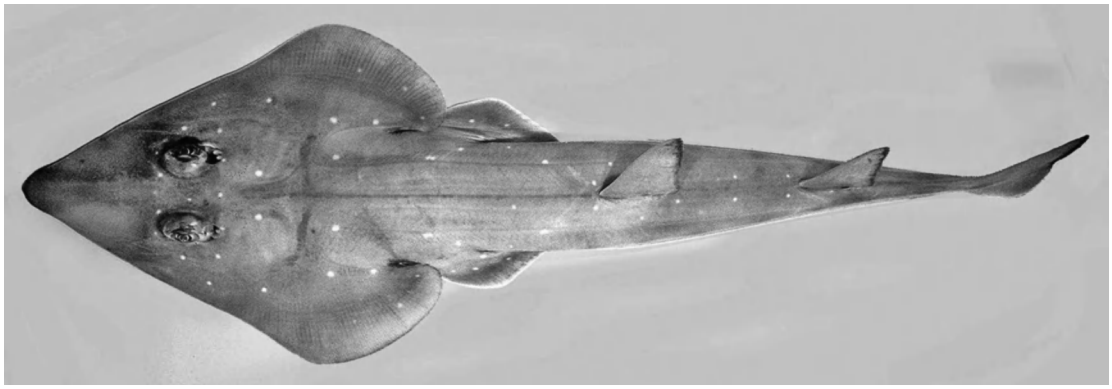

Spotted guitarfish (*Rhinobatos punctifer*). (*Photo © John E. Randall*)

of the first dorsal fin is behind the pelvic fins; the caudal fin is without a definite lower lobe; and the rear margin of the pectoral fins extends posteriorly to the origin of the pelvic fins, or beyond. The numerically largest genus, *Rhinobatos*, contains about 36 species, accounting for about 75% of the family.

Guitarfishes swim near the bottom, where they use their robust tail sector for propulsion and their pectoral fins as elevator and rudder. They often lie half buried in the sand or mud with their eyes and spiracles exposed. Occurring primarily in tropical and warm temperate coastal waters of all oceans, they rarely enter estuaries or freshwater.

Rajidae (skates). In body form, skates are more raylike than sharklike. The head and body are strongly depressed; the snout is moderately slender; the pectoral rays are united with the sides of the head, to very near the tip of the snout; the outer margin of the pelvic fins is concave or notched (so deeply notched in *Cruriraja* and *Anacanthobatis* that the upper lobes appear to be little legs); the tail sector is sharply distinct from the body sector; there is a single unbranched rostral process reaching near the tip of the snout (falling short in *Breviraja*, and absent in *Psammobatis*, *Sympterygia*, *Malacoraja*, and *Anacanthobatis*); there are two small dorsal fins near the tip of the tail (one in *Arhynchobatis*, none in *Anacanthobatis*); the caudal fin is reduced to a small membranous fold or is completely absent in some adults; and the upper surface of the disc and tail is variously rough with prickles and thorns (though smooth in *Anacanthobatis*). *Raja* is the typical genus and has the largest number of species; however, there is one genus in addition to those noted above that depart radically from the typical. *Dactylobatus armatus* is unique among batoids in having six or eight of its pectoral rays, at midlength of its pancake-shaped disc, extending beyond the edge as a tongue-shaped lobe.

The family consists of 18 genera and about 200 species. It is represented in all seas of the world, from the tropics to subpolar waters, from shallow depths to at least 1000 m (3300 ft). Like other rajiforms, skates are bottom feeders, relying primarily on their pectoral fins for locomotion. *See* BATOIDEA.

Herbert Bosehung

Bibliography. H. B. Bigelow and W. C. Schroeder, Sawfishes, guitarfishes, skates and rays, in Fishes of the Western North Atlantic, *Mem. Sears Found. Mar. Res. Mem.*, 1(pt. 2):1–514, 1953; J. D. McEachran and T. Miyake, Phylogenetic interrelationships of skates: A working hypothesis (Chondrichthyes, Rahjoidei), pp. 285–304, in H. L. Pratt, Jr., et.al. (eds.), Elasmobranchs as living resources: Advances in the biology, ecology, systematics, and the status of the fisheries, *NOAA Tech. Rep. 90*, 1990; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994.

Raman effect

A phenomenon observed in the scattering of light as it interacts with a material medium in which the incident light suffers a change in frequency due to internal energy change of the molecular scatterers. Raman scattering differs in this respect from Rayleigh scattering in which the incident and scattered light have the same frequency. Shifts in frequency are determined by the type of molecules in the scattering medium, and spectral analysis of the scattered light can provide a "fingerprint" of the chemical structure of the scatterers. Both incoherent and coherent forms of the Raman effect exist. Spontaneous Raman scattering, the usual (incoherent) form, is very weak, with Raman signals 4–5 orders of magnitude smaller than Rayleigh scattering and about 14 orders of magnitude smaller than fluorescence. In stimulated Raman scattering, a coherent form, the signals may be quite large. In addition, there are nonlinear forms of Raman scattering, including hyper-Raman scattering and coherent anti-Stokes Raman scattering (CARS). *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

Discovery. Raman scattering was experimentally observed in 1928 by the Indian physicists C. V. Raman and K. S. Krishnan, who examined sunlight scattered by a number of liquids. With the help of complementary filters, they found frequencies in the scattered light that were lower than the frequencies in the filtered sunlight. By using light of a single frequency from a mercury arc, they showed that the new frequencies were characteristic of the scattering medium. Within a few months of Raman and Krishnan's announcement of their discovery, the Soviet physicists G. Landsberg and L. Mandelstam communicated their independent discovery of the effect by studying scattering of light in crystals, and in the Russian literature the phenomenon is still often referred to as combination scattering.

Compton scattering is another example of inelastic scattering. In this case, x-ray or gamma-ray photons cause a scattering electron to recoil, transferring energy to it and leading to a reduction in photon frequency. *See* COMPTON EFFECT.

Physical principles. The mechanism of the Raman effect can be understood either by the particle picture of light or by wave theory. Both features emerge in the quantum theory of radiation.

Photon picture. The particle model envisages light quanta or photons traversing a molecular medium undergoing scattering. If the interaction is elastic, the photons scatter with unchanged energy E and momentum, and hence with unchanged frequency ν, using Planck's relationship, $E = h\nu$, where h is Planck's constant. Such a process gives rise to Rayleigh scattering. If, however, the collision is inelastic, the photons can gain or lose energy from the molecules. A change ΔE in the internal energy of the scatterer causes a change $-\Delta E$ in photon energy, hence a corresponding frequency change $\Delta\nu = -\Delta E/h$. Frequency-shifted scattered photons can occur at lower or higher energy relative to the incident photons, depending on whether the scattering molecule is in the ground or excited vibrational state. In the former case, photons lose energy by exciting a vibration and the scattered light appears at a lower frequency, $\nu_s = \nu - \Delta E/h$, called the Stokes

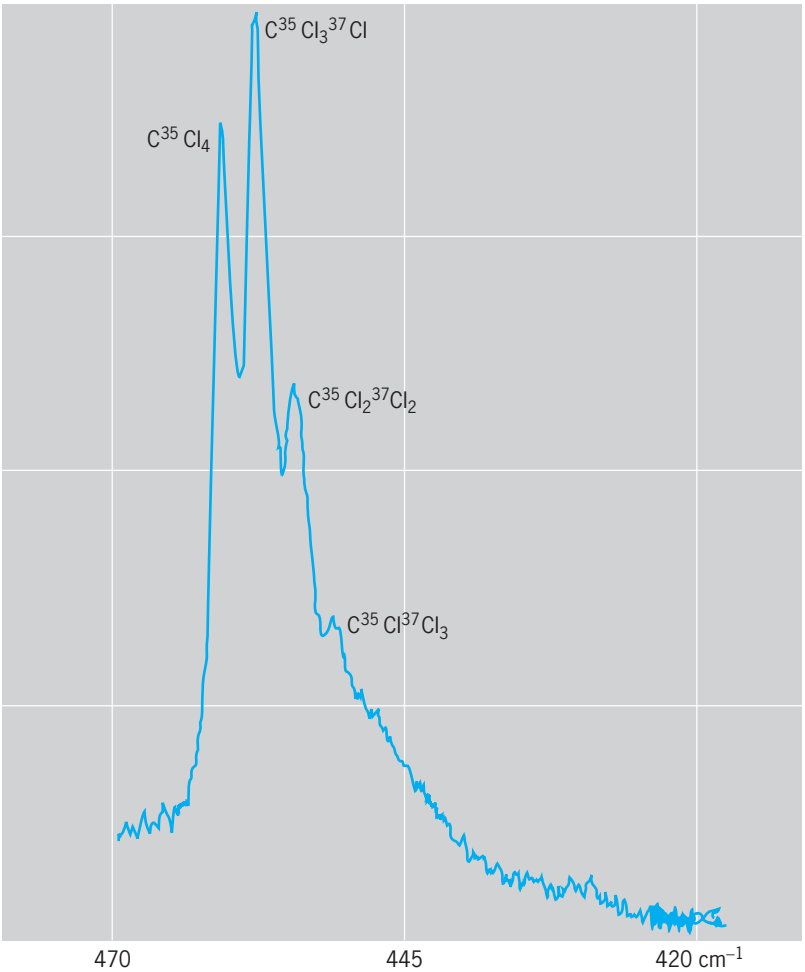

Fig. 1. Raman spectrum of carbon tetrachloride excited by He-Ne laser light at 632.8 nm. The scattered signal is recorded as a function of the frequency shift from the exciting spectral line. These shifts are usually expressed in units of inverse wavelength, $1/\lambda$, with λ measured in centimeters, called wave numbers and denoted by cm^{-1}, rather than actual frequency. (The frequency is $\nu = c/\lambda$, with c the speed of light in vacuum, so the two are proportional.) The Rayleigh-scattered spectral line (peak at 0 cm^{-1}) is 3 orders of magnitude more intense than the Raman lines, and its maximum is therefore far off scale. The Stokes lines appear at negative frequency shifts and the (weaker) anti-Stokes lines at positive frequency shifts. The lower scale gives the wavelengths in nanometers.

frequency. In the latter case, by interacting with a molecule in an excited vibrational state, the photons gain energy from the molecular vibrations and the scattered signal appears at higher frequency $\nu_{aS} = \nu + \Delta E/h$, the anti-Stokes frequency. In general, anti-Stokes Raman scattering results in much smaller scattering signals than Stokes scattering because only a small fraction of molecules is in an excited vibrational state and can contribute to anti-Stokes Raman scattering. *See* LIGHT; QUANTUM MECHANICS; SCATTERING EXPERIMENTS (ATOMS AND MOLECULES).

Wave picture. In the wave picture, light as an electromagnetic wave interacts with the molecules in the medium, inducing an oscillating dipole moment by separating the positively charged nuclei and negatively charged electrons. In Rayleigh scattering (elastic), the oscillation frequency is ν, the same as that of the incident light. However, in the case of Raman scattering (inelastic), scattering from the molecules, vibrating at frequency $\Delta E/h$, gives rise to two frequency components at $\nu \pm \Delta E/h$, with ΔE the internal energy of the scatterer. This induced dipole moment interacts with the light wave, producing emission not only at the frequency ν of the incoming field (elastic component), but also at sidebands at the sum and difference frequencies between the incoming light and the molecular vibrations (Stokes and anti-Stokes components ν_S and ν_{aS}). *See* DIPOLE MOMENT.

In general, any scatterer with internal energy structure can undergo the Raman effect by exchanging internal energy with the radiation field. In molecules, rotations, as well as vibrations, can produce Raman effects. Similarly, atoms can give rise to Raman scattering. However, in the context of revealing chemical structure, the Raman effect is of the greatest interest in the spectroscopy of molecules and crystals. *See* MOLECULAR STRUCTURE AND SPECTRA.

Raman spectroscopy. Raman scattering is analyzed by spectroscopic means. When a medium is excited by monochromatic incident light, the collection of new frequencies in the spectrum of scattered light is called its Raman spectrum. As the Raman effect probes vibrations of the molecule, which depend on the kinds of constituent atoms and their bond strengths and arrangements, a Raman spectrum provides a high degree of chemical structural information. **Figure 1** shows a typical Raman spectrum. The kind of information provided by this spectrum is similar to that provided by infrared spectroscopy; in fact, Raman and infrared spectra often provide complementary data about molecular structure. *See* INFRARED SPECTROSCOPY.

The advent of lasers as intense sources of monochromatic light was a milestone in the history of Raman spectroscopy and resulted in dramatically improved scattering signals. Today, intense sources of laser light over a wide range of frequencies from the near-ultraviolet to the near-infrared region are used in Raman scattering studies, enabling selection of optimum excitation conditions for each sample. Tunable dye lasers and, most recently, high-intensity diode lasers have become useful sources of near-infrared light and have made this spectral region attractive for practical applications. In particular, diode lasers provide compact, inexpensive Raman instrumentation. Development of low-noise, high-quantum-efficiency multichannel detectors [charge-coupled device (CCD) arrays], combined with high-throughput single-stage spectrographs used in combination with holographic laser rejection filters (such as Raman "notch" filters), has led to high-sensitivity Raman spectrometers. In addition, by using a microscope, particularly in combination with diaphragms, for focusing the excitation light and collecting the scattered light, Raman signals from femtoliter volumes (about 1 μm^3) can be observed, enabling spatially resolved measurements from very small structures (micro-Raman spectroscopy). Also, using pulsed laser sources, Raman spectra can be measured on the picosecond time scale, providing information about short-lived species. *See* CHARGE-COUPLED DEVICES; LASER; SPECTROGRAPH.

Figure 2 shows the layout of an advanced Raman facility. Tunable excitation provides access to all

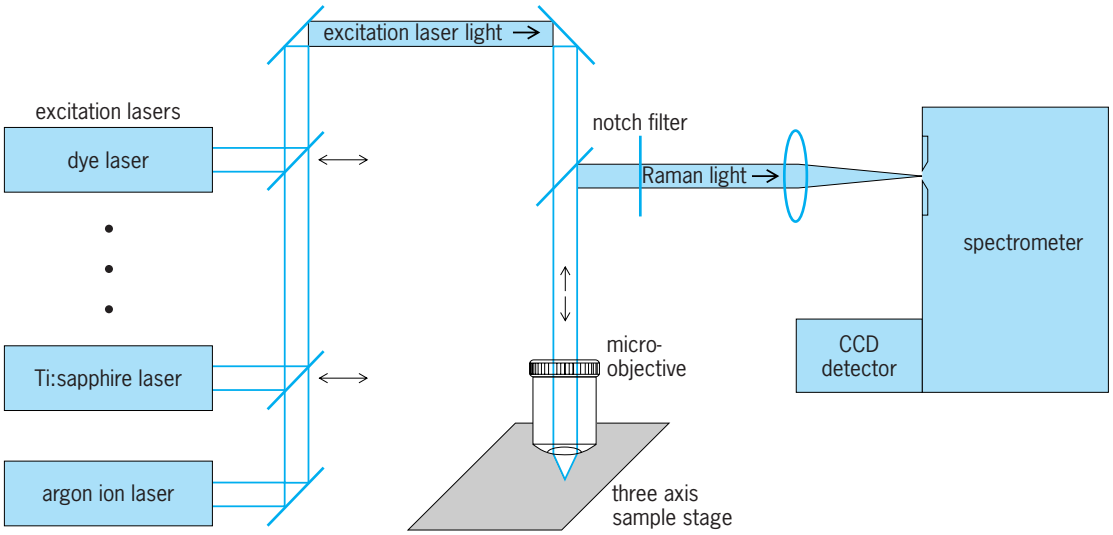

Fig. 2. Layout of a modern Raman facility.

visible and near-infrared wavelengths for resonance Raman studies. In addition, a microscope, with a three-axis control stage and multiple objectives, enables excitation regions as small as 0.5 μm in diameter to be studied. *See* CONFOCAL MICROSCOPY; OPTICAL MICROSCOPE.

The main advantage of Raman spectroscopy is its capability to provide rich information about the molecular structure of the sample. Sophisticated data analysis techniques based on multivariate analysis, in which spectral information from many Raman spectral lines is used simultaneously, have made it possible to exploit the full information content of Raman spectra, and to elucidate the chemical structure and composition of very complex systems such as biological materials. *See* SPECTROSCOPY.

Incoherent Raman scattering. The above discussion on Raman spectroscopy illustrates the most common form of incoherent Raman spectroscopy, also known as the spontaneous Raman effect. Several other variants of incoherent Raman scattering are also important.

Resonance Raman scattering (RRS). In usual Raman scattering studies, the frequency of the excitation light is not resonant with any actual molecular transition. In resonance Raman scattering, however, the excitation laser frequency is tuned to match a specific absorption peak of the scattering medium. Under these conditions the Raman scattering can become up to 10,000 times stronger, but in most molecules it is a different and much weaker process compared to fluorescence, which is excited in the same way and, in many cases, produces overlapping emission. In addition to the increased size of the Raman signals, resonance Raman scattering has the advantage of higher specificity, since the resonance Raman spectrum is dominated by vibrations related to the part of the molecule responsible for the appropriate resonant electronic transition. This is useful in understanding and assigning Raman spectra of large and complex molecules.

Raman scattering and fluorescence can be distinguished by the characteristics of their emission spectra. The fluorescence emission spectrum, determined by the electronic transition frequencies of the molecule, always occurs at the same frequency, regardless of the frequency of the excitation light. In contrast, the Raman emission spectrum is always shifted from the frequency of the excitation light by the same amount, regardless of its frequency. In many cases, resonance Raman scattering and fluorescence can also be differentiated in time-resolved experiments: Whereas resonance Raman scattering emission occurs at the same time as the excitation laser pulse, fluorescence occurs with some time delay, on the order of picoseconds to nanoseconds. *See* FLUORESCENCE.

Surface-enhanced Raman scattering (SERS). As discussed above, Raman scattering signals are usually exceedingly weak, rendering detection difficult and limiting practical applications. Therefore, about 50 years after the discovery of the Raman effect, a development which showed unexpectedly strong Raman signals from pyridine on a rough silver electrode surface attracted considerable attention. The effect was called surface-enhanced Raman scattering. Later experiments showed that it is not so much a surface effect as it is a nanostructure effect. (A nanostructure is a collection of nanoparticles, regular arrangements or clusters.) The strongly enhanced Raman signals arise from molecules attached to silver and gold nanostructures, such as small colloidal particles in solution, or silver or gold nano-island films. The enhancement of Raman signals in SERS can be as large as 14 orders of magnitude. Such a tremendous enhancement effect is primarily due to the optical properties of metal nanoparticles, based on the collective oscillations of their conduction electrons, known as surface plasmons. For silver and gold nanoparticles, these oscillations can span a relatively broad spectral range, and fall in the frequency range of Raman excitation and scattered signals. Resonances between the excitation light, the Raman scattered light, and the surface plasmons create strongly enhanced and spatially confined local optical fields in the vicinity of the nanoparticles, where Raman scattering takes place,

resulting in strongly increased Raman signals. SERS enhancement factors associated with enhanced electromagnetic fields have been found to be up to 12 orders of magnitude in size. Chemical interactions between the molecule and the metal can further enhance the Raman signal size in ways that are still not completely understood. *See* NANOSTRUCTURE; PLASMON.

In addition to its general scientific interest, since the mid-1990s SERS has gained considerable interest as a spectroscopic tool. Enhancement factors of 14 orders of magnitude increase the sensitivity of SERS to the single-molecule level. At present, SERS is the only way to detect and simultaneously identify the chemical structure of a single molecule based on its strongly enhanced Raman signal. Another important feature of SERS is its ability to precisely locate Raman scattering. Exploiting local optical fields of special metallic nanostructures, SERS can result in probed volumes as narrow as 5 nm or less in lateral dimension.

Hyper-Raman scattering (HRS). At extremely high excitation intensities, scattering signals Raman-shifted relative to twice the frequency of the excitation laser can be observed. This effect, called hyper-Raman scattering, is most easily described in the photon picture: with an intense laser source there is a small probability that two photons can simultaneously interact with the same molecule in a manner similar to that of a single photon of approximately twice the frequency. Due to its two-photon nature, hyper-Raman scattering can probe vibrations (called silent modes) that cannot be studied in either Raman scattering or infrared absorption. Hyper-Raman scattering is a very weak effect and requires extremely strong pulsed laser excitation. However, it can be enhanced in an analogous fashion to normal Raman scattering in SERS, when the two-photon excited Raman process takes place in the vicinity of silver and gold nanoparticles. Hyper-Raman scattering is an example of a nonlinear incoherent Raman effect.

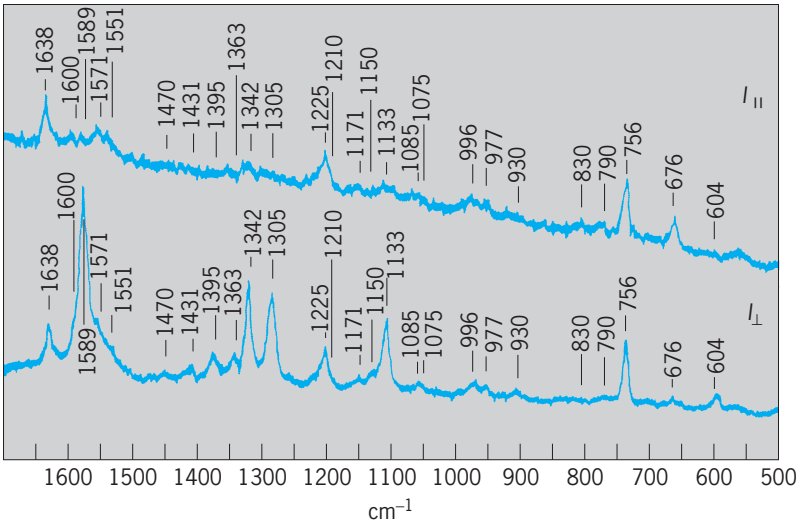

Fig. 3. Resonance Raman spectrum of oxyhemoglobin. (*After T. G. Spiro and T. C. Strekas, Resonance Raman spectra of hemoglobin and cytochrome c, Proc. Nat. Acad. Sci. USA, 69:2622–2626, 1972*)

Coherent Raman scattering. In addition to the incoherent Raman processes discussed above, the availability of lasers has also opened the field of study of coherent or stimulated Raman processes.

Stimulated Raman scattering (SRS). Stimulated Raman emission (or scattering) in a Raman transition bears the same relationship to spontaneous Raman emission as does stimulated emission to spontaneous emission in a conventional molecular transition. In stimulated Raman scattering, a sample is pumped by an intense laser source, generating gain at the Stokes frequency. Spontaneous Raman light emitted along the pump direction can then be amplified, producing strong coherent Raman emission along that direction, comprising in most cases one strong Raman line. This intense Stokes light can, in turn, act as pump light for generating a second Stokes field. In this way a comb of intense frequencies can be generated.

Raman lasers and inverse Raman scattering. This Raman gain can also be used to convert intense light from pump to Raman frequencies. In this case, two incident laser fields offset in frequency by a molecular vibration frequency are incident on a Raman-active medium. As the waves propagate through the medium, pump field photons are converted into Stokes photons, so that the Raman field grows. In a closely related configuration, the higher-frequency field is broadband. Absorption can then be observed at the anti-Stokes frequencies, a process known as inverse Raman scattering or Stoicheff absorption.

Coherent anti-Stokes Raman scattering (CARS). Using two excitation lasers offset in frequency by the Stokes shift, stimulated scattering at the anti-Stokes frequency can also be obtained. This method of generating anti-Stokes light is called coherent anti-Stokes Raman scattering (CARS). It is a coherent four-wave mixing interaction, that is, three coherent waves (two pump fields and one Stokes field) interact nonlinearly to produce a fourth. Unlike the other stimulated scattering processes described above, the anti-Stokes light is phase-coherent with the excitation light waves. CARS has been studied for several decades, but the effect has recently become of interest as a spectroscopic method for imaging a specific chemical compound by subjecting the sample to two laser fields offset by the molecular vibrational frequency of the molecule to be imaged. *See* NONLINEAR OPTICS.

Applications. Raman spectroscopy is of considerable value in determining molecular structure and in chemical analysis. It can be used to characterize the chemical and structural composition of nearly all kinds of molecules in liquid, gas phase, and solid samples for both scientific study and industrial monitoring, such as control of products and processes. As with many optical spectroscopy effects, the Raman effect can be used to interrogate a substance without altering it under ambient conditions in almost every environment. Measuring a Raman spectrum does not require special sample preparation techniques, in contrast with infrared absorption. Materials can be

studied remotely, either with line-of-sight techniques or via optical fiber probes (with appropriate filters). *See* FIBER-OPTIC SENSOR.

Biomedical applications. Raman spectroscopy can be used to characterize the chemical and structural composition of complex systems, such as biological materials. It has the potential to become a central tool for medical diagnosis, based on quantitative extraction of molecular information. The technique can monitor small changes in the chemical composition of a biological system such as human tissue, which can provide very early indications of development of disease. It is also used for noninvasive (that is, noncontact) measurement of glucose and other blood analytes in the body. As an example, (**Fig. 4**) shows Raman spectra of four tissue samples—one from normal tissue, two from benign lesions, and one from a breast cancer—along with stained microscopy sections of biopsy samples used in making the pathology diagnoses. The underlying composition, obtained from multivariate analysis, is also shown. The spectral differences and concomitant changes in composition observed indicate the power of the technique.

Trace monitoring. With developments in SERS, Raman spectroscopy not only is a tool for structural analysis, but also allows ultrasensitive and trace detection of substances down to the single-molecule level. Moreover, SERS is an analytical technique that can provide information about surface and interface processes. SERS opens up exciting opportunities in laboratory medicine and for basic research in biophysics and biochemistry, where it provides ultrasensitive detection and characterization of biophysically and biomedically relevant molecules and processes, as well as a vibrational spectroscopy with extremely high spatial resolution.

Nanostructure materials. Raman spectroscopy has become an important tool for characterizing nanostructures such as carbon nanotubes and nanowires.

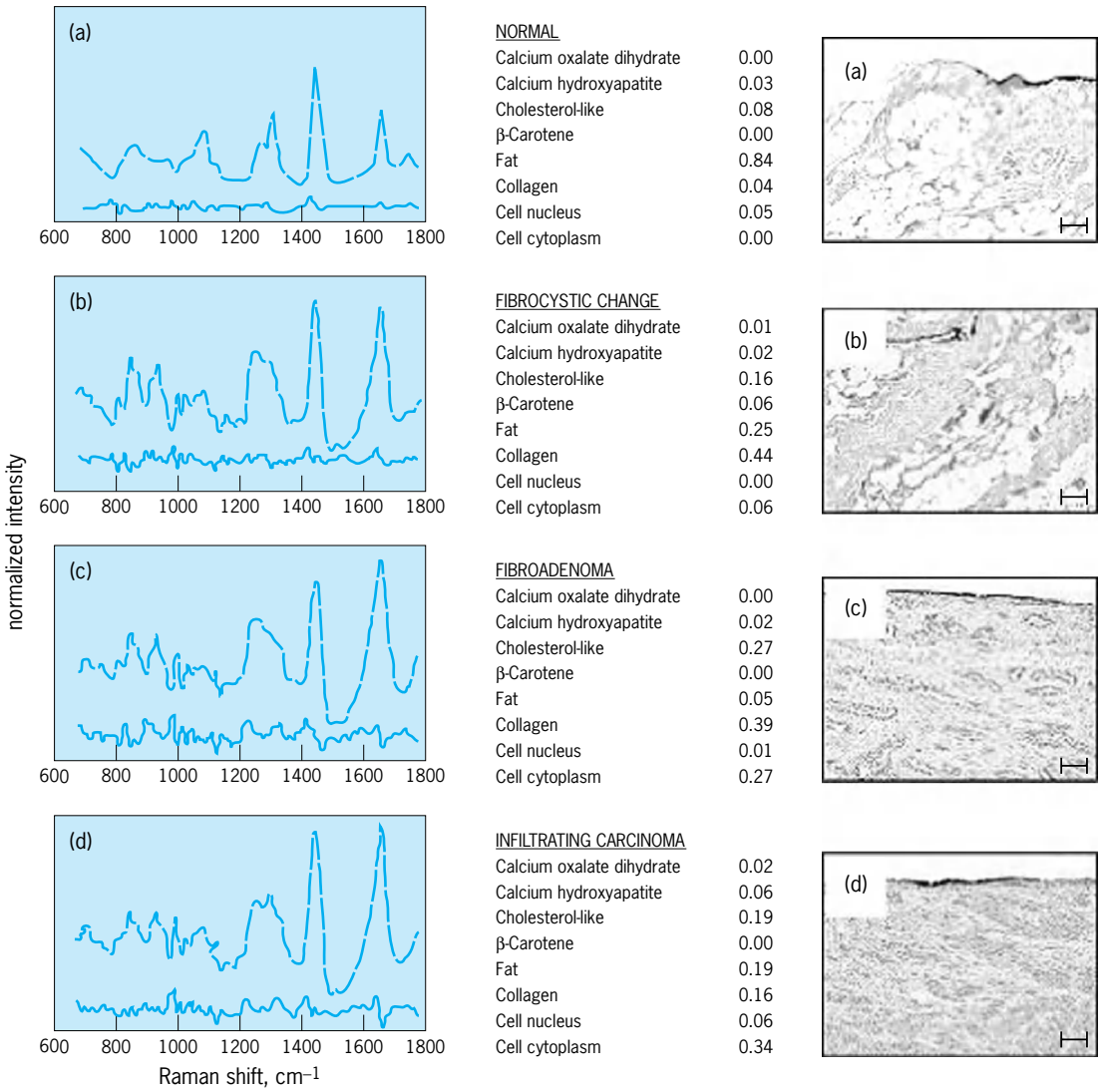

Fig. 4. Raman spectra of four human breast samples. The spectra at the left are from (*a*) normal tissue, (*b*,*c*), benign lesions, and (*d*) a breast cancer. At right, the corresponding pathology sections are shown. The chemical/structural compositions, based on multivariate spectral analysis, are shown in the center. In each panel on the left, the upper traces show the observed Raman spectrum and multivariate fit, and the spectrum below shows the difference between them.

Due to their extremely small size, in the nanometer range, the electronic transitions of nanostructures exhibit very specific and narrow novel features, which lead to interesting new and strong effects in resonance Raman spectroscopy, providing information about both the vibrational and electronic structure. Katrin Kneipp; Michael S. Feld

Bibliography. N. Bloembergen, *Nonlinear Optics*, 4th ed., World Scientific, 1996; D. B. Chase, *Fourier Transform Raman Spectroscopy*, Academic Press, 1994; R. J. Clark and R. E. Hester (eds.), *Advances in Infrared and Raman Spectroscopy*, vols. 1–18, 1975–1989, *Advances in Spectroscopy*, vols. 19–26, Wiley, 1991–1998; N. B. Colthup, L. H. Daly, and S. E. Wiberly, *Introduction to Infrared and Raman Spectroscopy*, 3d ed., Academic Press, 1990; A. M. K. Enejder et al., Raman spectroscopy for non-invasive glucose measurements, *J. Biomed. Opt.*, 10(3):031114, May/June, 2005; J. R. Ferraro, K. Nakamoto, and C. W. Brown, *Introductory Raman Spectroscopy*, 2d ed., Academic Press, 2003; E. B. Hanlon et al., Prospects for in vivo Raman spectroscopy, *Phys. Med. Biol.*, 45(2):R1–R59, 2000; K. Kneipp et al., Ultrasensitive chemical analysis by Raman spectroscopy, *Chem. Rev.*, 10:2957–2975, 1999.

Ramie

The plant *Boehmeria nivea*, a stingless member of the nettle family; the only member of the family used commercially for fiber. This herbaceous perennial is erect, usually nonbranching, and 3 to 6 ft (1 to 2 m) tall at maturity (**Fig. 1**). The fiber, which comes from the inner bark, is exceptionally strong and has uses similar to those for fiber flax.

Fig. 1. Ramie (*Boehmeria nivea*), height 3–6 ft (1–2 m).

Cultivation. Ramie is grown most in the tropics and subtropics (to about 30–35° latitude) of the Far East and Brazil. It is planted from vegetative parts, usually root (rhizome) cuttings, but sometimes from stem cuttings. Because of the extreme variability of seedlings, seed is used only for variety development.

The rhizomes are placed in rows at a depth of about 5 cm (2 in.) on fertile, well-drained soil that has been worked two or more times to kill weeds that have sprouted. During the first year, while the plants spread and establish solid stand, weeds must be kept out, and the ramie must be kept cut often enough to prevent production of seed. Fertilizers expediate the establishment of the stand and are required in considerable amounts in subsequent years to replace the large quantities of nutrients removed by the crop.

Fiber extraction. Harvesting is usually done about every 60 days during the growing season. Three to as many as six crops are harvested annually. Plants are usually defoliated, either chemically or mechanically, before they are cut and brought to a central location for decortication (separation of the inner bark containing the fiber from the remainer of the stem).

On very small plantings, the fiber is usually extracted by hand-stripping and scraping methods. Elsewhere, hand methods have generally given way to mechanized decortication, most frequently a small hand-fed raspidor, driven by electricity or more often by a small internal combustion engine. Automatic decorticators similar to the ones used for sisal are used frequently on plantations. Ramie leaves are highly nutritious, and the tops or an entire young crop is sometimes used as livestock feed.

Diseases. Few diseases affect ramie and are usually limited to minor outbreaks. Probably the most serious is the white fungus, *Rosellina necatris*, reported in the Far East. The fungus attacks the roots and, by the time it is detected, may kill the plant. If infected planting stock is suspected, rhizome cuttings should be soaked in a mercuric chloride solution. *Rhizoctonia solani* (**Fig. 2**), a fungus that destroys young plantings and causes broken stems in older plants,

Fig. 2. Young plants attacked by *Rhizoctonia solani*.

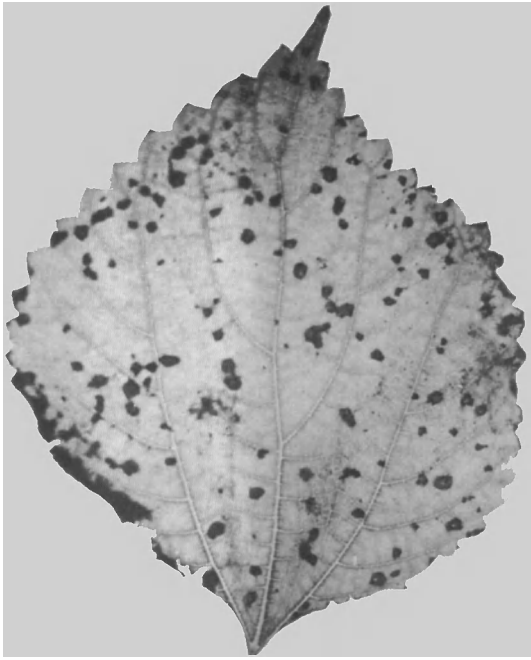

Fig. 3. Leaf spot caused by *Cercospora boehmeria*.

may be controlled by treating planting material with captan. The leaf-spot diseases, caused by *Cercospora boehmeria* and *C. krugliana*, retard the growth of young plants and may defoliate plants of poor vigor (**Fig. 3**). These diseases can be controlled by spraying or dusting infected areas with Dithane.

Many insects feed on ramie, but few are a serious threat. Without control, which is not difficult, several may become harmful. The leaf rollers, *Pilcrosis ramentalis* and *Sylepta silicalis*, are considered the most serious threat and are the most widespread. The larvae of *Pilcrosis* feed on the leaves, then roll up the leaf and pupate within it. Heavy infestations result in defoliation, and the plant ceases to grow. This pest may be controlled by BHC. The black caterpillar, *Cocytodes coerulea*, can also completely defoliate ramie. Dieldrin is an effective control. *See* FLAX; LINEN; NATURAL FIBER; PHLOEM; TEXTILE.

Elton G. Nelson

Bibliography. H. F. Linskens and J. F. Jackson (eds.), *Plant Fibers*, 1989; H. R. Mauersberger, *Mathew's Textile Fibers*, 6th ed., 1954.

Random matrices

Collections of large matrices, chosen at random from some ensemble. Random-matrix theory is a branch of mathematics which emerged from the study of complex physical problems, for which a statistical analysis is often more enlightening than a hopeless attempt to control every degree of freedom, or every detail of the dynamics. Although the connections to various parts of mathematics are very rich, the relevance of this approach to physics is also significant.

Principles. Random matrices were introduced by Eugene Wigner in nuclear physics in 1950. In quantum mechanics the discrete energy levels of a system of particles, bound together, are given by the eigenvalues of a hamiltonian operator, which embodies the interactions between the constituents. This leads to the Schrödinger equation which, in most cases of interest in the physics of nuclei, cannot be solved exactly, even with the most advanced computers. For a complex nucleus, instead of finding the location of the nuclear energy levels through untrustworthy approximate solutions, Wigner proposed to study the statistics of eigenvalues of large matrices, drawn at random from some ensemble. The only constraint is to choose an ensemble which respects the symmetries that are present in the forces between the nucleons in the original problem, and to select a sequence of levels corresponding to the quantum numbers that are conserved as a consequence of these symmetries, such as angular momentum and parity. The statistical theory does not attempt to predict the detailed sequence of energy levels of a given nucleus, but only the general properties of those sequences and, for instance, the presence of hidden symmetries. In many cases this is more important than knowing the exact location of a particular energy level. This program became the starting point of a new field, which is now widely used in mathematics and physics for the understanding of quantum chaos, disordered systems, fluctuations in mesoscopic systems, random surfaces, zeros of analytic functions, and so forth. *See* CONSERVATION LAWS (PHYSICS); EIGENVALUE (QUANTUM MECHANICS); MATRIX THEORY; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Initially Wigner introduced simple gaussian probability distributions for the matrix elements, with the requirement that the ensemble be invariant; that is, it should not favor any preferred direction in the space on which the matrices act. He could then prove that, for large matrices, the density of eigenvalues, that is, their number per unit interval of energy, followed a "semicircle" law. This means that there are essentially no eigenvalues outside an interval, for instance $[-1,1]$ in the appropriate normalization, and within this interval the density ρ varies with eigenvalue λ according to a law of the type given by Eq. (1). [The

$$\rho(\lambda) = \frac{2}{\pi}\sqrt{1-\lambda^2} \qquad (1)$$

graph of this function is a semicircle.] The statistics of the eigenvalues, of the correlations between their positions, and of the distribution of their spacings led to many intricate new mathematical problems, even for the simple gaussian ensembles, but most of them are now well understood. The distribution of the spacings of successive eigenvalues is one of the characteristics of random matrices that is commonly used for practical applications. It exhibits level repulsion, which reflects the symmetries of the ensemble; indeed, eigenvalues of random matrices

behave quite differently from independent random numbers (whose spacings would follow the Poisson distribution, that is, exhibit no level repulsion). In a comparison with experiment, a spacing distribution must be selected that is appropriate to the basic properties of the system. For instance, when the dynamics is invariant under time reversal, which is true when there is no magnetic field acting on the system, real random matrices can be chosen. In the presence of a magnetic field, the matrices are necessarily complex and the spacing distribution is quite different. *See* DISTRIBUTION (PROBABILITY).

However, some doubts were soon raised concerning the applicability of the theory to real physical situations, since the density of nuclear levels, for instance, increases rapidly with growing excitation energies, and is certainly not given by a semicircular law. It was then conjectured, and proved much later, that the local statistics of levels, that is, the correlations and the spacings at the scale of the mean level separation, were universal, that is, largely independent of the probability distribution of the matrices. This universality is undoubtedly the main reason for the vast range of applicability of the theory, provided the research is limited to the study of the local statistics of levels.

Applications. A few examples outside the initial range of nuclear physics will illustrate some of the modern uses of random matrix theory.

Quantum chaos. Classically the time evolution of a system with several degrees of freedom is generically chaotic; that is, it shows a sensitive dependence upon its initial data (the butterfly effect of meteorologists). It has been conjectured that the same dynamical system, treated now by quantum mechanics instead of newtonian dynamics, exhibits random-matrix-level statistics in its energy levels. Numerical studies of nonintegrable systems, such as the hydrogen atom in a magnetic field, strongly support this conjecture, but it is still a matter of active research to understand the basic reason for the replacement of a deterministic nonrandom problem by a probabilistic treatment. *See* CHAOS; NONLINEAR PHYSICS.

Mesoscopic systems. In semiconductor physics, numerous interesting phenomena and applications have been found for quantum dots, which are small chaotic cavities, coupled via narrow leads to the rest of the system. The theory of quantum transport relies in such cases upon random matrices. Similarly the diffusive motion of electrons in small metallic grains, when there is enough time for the electrons to diffuse over the whole size of the grain, leads to a statistical distribution of energy levels in agreement with random matrix theory. *See* QUANTIZED ELECTRONIC STRUCTURE (QUEST).

The previous examples belong to a long list of phenomena concerning mesoscopic systems, that is, small semiconductor or metallic devices at very low temperatures. In these systems the electrons, scattered by the crystalline ions and by the impurities of the structure, retain the full quantum-mechanical phase coherence which allows for interference between different electronic paths. As a consequence of the small size of these systems, fluctuations of their conductance are observed from sample to sample, or in a single sample as a function of an applied magnetic field. The universality of these fluctuations was soon discovered theoretically and checked by experiment. This striking phenomenon is related to the universality present in the statistics of levels of random matrices. *See* AHARONOV-BOHM EFFECT; MESOSCOPIC PHYSICS.

String theory. Some other domains of application are more surprising. For instance, large random matrices have been found to encode, in some limit, the combinatorics of randomly triangulated surfaces of arbitrary genus. Those surfaces appear in string theory, an approach which attempts to embody Albert Einstein's gravity at the quantum level. The elementary objects of the physical world are no longer points in space, but bits of strings, which sweep out a two-dimensional surface during their motion. Random-matrix theory has brought some partial resolution to such problems, still very incomplete, but in a domain in which there are still very few explicit results. *See* SUPERSTRING THEORY.

Riemann zeta function. In mathematics, this function, defined by Eq. (2), is one of the most studied an-

$$\varsigma(s) = \sum_{1}^{\infty} \frac{1}{n^s} \tag{2}$$

alytic functions of a single complex variable. A famous conjecture of Bernhard Riemann has resisted the efforts of many waves of mathematicians to prove it. It states that, in the band $0 < \mathrm{Re}\ s < 1$ of the complex s-plane, this function vanishes exclusively on the line $\mathrm{Re}\ s = 1/2$. A huge number of zeros (hundreds of millions) have been computed numerically, and they have all been found to agree with Riemann's hypothesis. Furthermore, the statistics of those zeros, for instance the distribution of their spacings, measured in units of their local density, is in striking agreement with the spacing distribution of the eigenvalues of random matrices, drawn from an ensemble invariant under unitary transformations. The explanation of these remarkable observations remains a challenge to mathematicians. *See* NUMBER THEORY.

Relation to other fields. The mathematical theory underlying the properties of random matrices overlaps with several active fields of contemporary mathematics, such as the asymptotic behavior of orthogonal polynomials at large-order, integrable hierachies, tau functions, semiclassical expansions, combinatorics, and group theory; and it is the subject of active research and collaboration between physics and mathematics.

Edouard Brézin

Bibliography. C. E. Beenakker, Random-matrix theory of quantum transport, *Rev. Mod. Phys.*, 69:731–808, 1997; D. J. Gross, T. Piran, and S. Weinberg (eds.), *Two Dimensional Quantum Gravity and Random Surfaces*, Jerusalem Winter School, vol. 8, World Scientific, 1992; M. L. Mehta, *Random Matrices*, 2d ed., Academic Press, 1991.

Rangefinder (optics)

An optical instrument for measuring distance, usually from its position to a target point. Light from the target enters the optical system through two windows spaced apart, the distance between the windows being termed the base length of the rangefinder. The rangefinder operates as an angle-measuring device for solving the triangle comprising the rangefinder base length and the line from each window to the target point. Rangefinders can be classified in general as being of the coincidence or the stereoscopic type.

Coincidence rangefinders. In these types, one-eyed viewing through a single eyepiece provides the basis for manipulation of the rangefinder adjustment to cause two images or parts of each to match or coincide. This type of device is used, in its simpler forms, in photographic cameras. The range triangle for such a rangefinder is shown in **Fig. 1**, where B is the base, angle A a right angle, D the range, and L the convergence angle at the target T. The relationship is given by Eq. (1). The basic optical arrangement is

$$D = B \cot L \tag{1}$$

shown in **Fig. 2**, where M_1 and M_2 are a semitransparent mirror and a reflecting mirror, respectively. When coincidence is obtained, that is, when the target T is seen in the same apparent position along either path, the rangefinder equation is satisfied. In many small rangefinders coincidence is achieved by rotating mirror M_2 through a small angle while viewing the images in mirror M_1; other deviating means which do not require rotating of mirror M_2 are also employed. **Figure 3***a* and *b* shows the appearance of the images in a superposed field rangefinder. If mirror M_1 is made fully reflecting and arranged to fill only the lower half of the field of view, allowing the eye to view target T directly through the upper half, a split-field type of rangefinder is produced; the images appear as shown in **Fig. 4***a* and *b*.

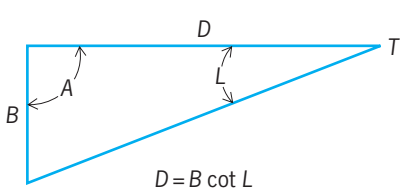

Fig. 1. Range triangle; symbols explained in the text.

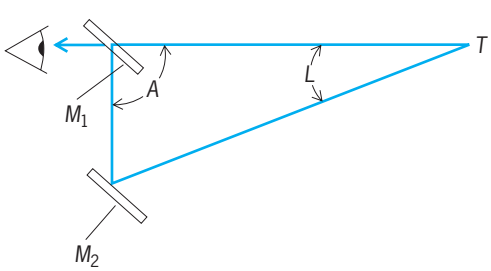

Fig. 2. Simple coincidence rangefinder; symbols explained in the text.

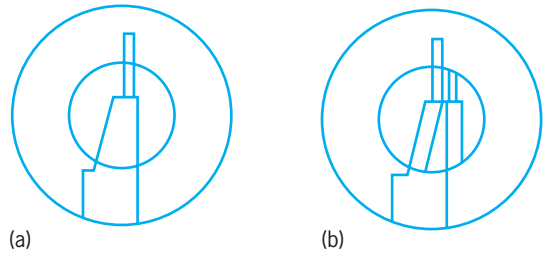

Fig. 3. View in superposed field rangefinder. (*a*) Images in coincidence. (*b*) Images unmatched.

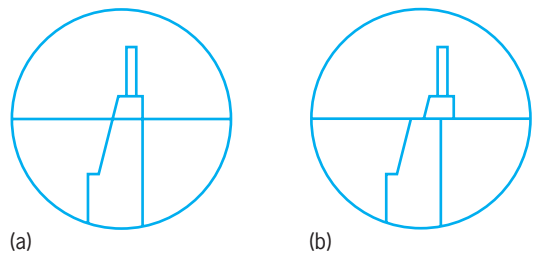

Fig. 4. View in split-field rangefinder. (*a*) Images matched. (*b*) Images unmatched.

Camera rangefinders. These usually have base lengths varying from less than 1 in. (25 mm) to more than 3 in. (75 mm), with magnifications between 0.5 and 2.5. The lower powers are found in arrangements where the rangefinder and viewfinder are combined for viewing through a single eyepiece; here a magnification of unity or less is necessary to cover the usual 42–52° angle of view.

Military types. Coincidence rangefinders of the military type ordinarily employ a pentaprism, or its two-mirror equivalent, at each window. These prisms, permanently fixed in place, reflect the light rays inward at an angle exactly 90° to the angle of incidence. The light from each prism passes through a telescope objective, and an image of the target formed by each objective is combined and directed by a prism assembly in a single eyepiece for simultaneous viewing. **Figure 5** shows the geometry to such a rangefinder. Here b is the base length of the rangefinder, and f' is the focal length of its objective lenses. Light from an infinitely distant target would arrive along lines TG and $T'H$, while light from a target distance R would arrive along lines TG and TK. The objectives, located at G and H and with their axes on lines TG and TH, respectively, form images of the target T at J and at K, respectively. With an infinitely distant target, objective H would form its image at L rather than at K. This displacement d, equal to LK, is called the parallactic displacement and, since triangles TGH and HLK are similar, is the measure of range. The relationship is given by Eq. (2).

$$R = \frac{bf'}{d} \tag{2}$$

The optical arrangement of a typical military coincidence rangefinder is shown in **Fig. 6**. The rangefinder is adjusted so that if the object being

viewed is at infinity (dotted line path through objective H), both images of the object are seen exactly superimposed. If the object is closer than infinity, the axis of the image which is formed by objective H is displaced by a deviation means, such as prism D, in the manner and amount described by Eq. (2), so that superimposition is achieved. The amount of such displacement is shown on the scale as the distance reading. Other means for obtaining deviation include oppositely rotating wedges (diasporometer), two equal coaxial oppositely arranged prisms with variable spacing, and slidable positive and negative lenses (swing wedge).

Stereoscopic rangefinders. These are entirely different, although externally they resemble coincidence rangefinders except for the fact that they possess two eyepieces. Both eyes are used with this instrument, rangefinding being accomplished by stereoscopic vision. It is essentially a large stereobinocular fitted with special reticles which allow a skilled user to superimpose the stereo image formed by the pair of reticles over the images of the target seen in the eyepieces, so that the reticle marks appear to be suspended over the target and at the same apparent distance. In the typical stereoscopic rangefinder (**Fig. 7**) the objectives form their images in the plane of their respective reticles. A diasporometer in one half of the instrument is used to vary the angle between the beams reaching the eyes from the target until it equals that between the beams reaching the eyes from the two reticles, at which point the reticle image appears to be at the same distance as the target. *See* STEREOSCOPY.

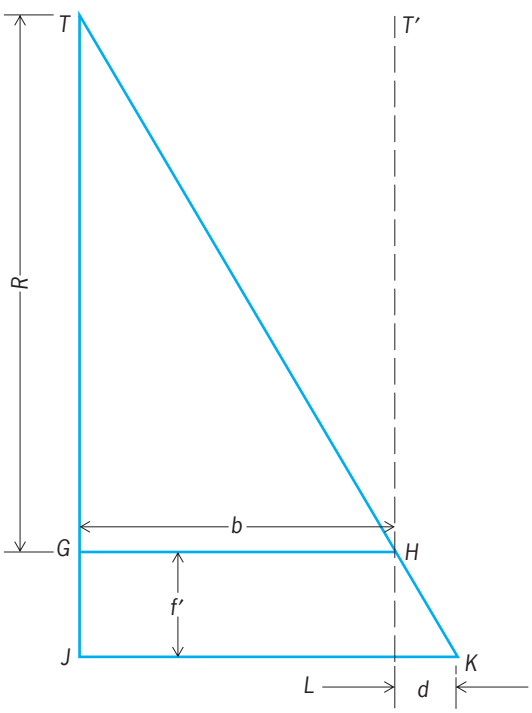

Fig. 5. Range triangle for military coincidence rangefinder; symbols explained in text.

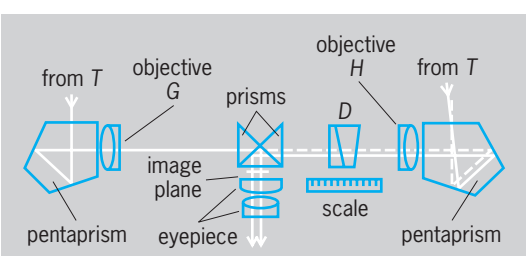

Fig. 6. Military coincidence rangefinder.

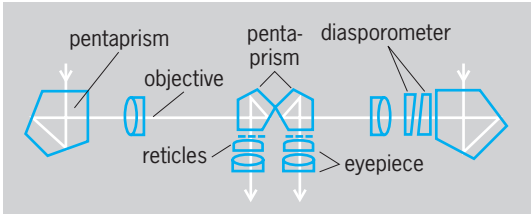

Fig. 7. Stereoscopic rangefinder.

Rangefinder errors. The accuracy of a rangefinder depends upon the accuracy of the eye in judging the coincidence of two lines, as well as upon the characteristics of the instrument and the range distance. Rangefinder errors are proportional to the square of the range and inversely proportional to the base length and to the magnification. A 1.1-yd (1-m) base and a 3.3-yd (3-m) base rangefinder of 15 power would have errors of 7.3 and 2.7 yd (6.7 and 2.5 m) at 1000-yd (914-m) range and 470 yd (430 m) and 170 yd (155 m) at 8000-yd (7315-m) range, respectively.

Heightfinders. These are modified forms of rangefinders originally used in the directing of antiaircraft fire. A heightfinder depends for its operation upon the knowledge of the range R to the aircraft and its angle of elevation θ. The height is given by Eq. (3).

$$H = R \sin \theta \tag{3}$$

In practice, one way that the height may be measured is by multiplying range and elevation angle in a gear train computer. Another method employs an additional pair of deviation prisms within the rangefinder, geared so as to produce a deviation which varies as $\sin \theta$.

Depression rangefinders. These are used in coast defense positions, where such instruments can be installed on a cliff top, and this elevation is used as the rangefinder base. The device measures the depression angle of the target, which it converts directly to range after including necessary corrections for tide, curvature of the Earth, and refraction of the atmosphere.

Stadia methods of rangefinding. These are suitable for some applications. If the size of the object is known or can be guessed accurately, its distance can be determined by measuring its size in the eyepiece reticle of a telescopic system. *See* LENS (OPTICS); OPTICAL PRISM.

Edward K. Kaprelian

Bibliography. Optical Society of America, *Handbook of Optics*, 2 vols., 2d ed., 1995; W. J. Smith, *Modern Optical Engineering*, 3d ed., 2000.

Rankine cycle

A thermodynamic cycle used as an ideal standard for the comparative performance of heat-engine and heat-pump installations operating with a condensable vapor as the working fluid. Applied typically to a steam power plant, as shown in the **illustration**, the cycle has four phases: (1) heat addition *bcde* in a boiler at constant pressure p_1 changing water at *b* to superheated steam at *e*, (2) isentropic expansion *ef* in a prime mover from initial pressure p_1 to back pressure p_2, (3) heat rejection *fa* in a condenser at constant pressure p_2 with wet steam at *f* converted to saturated liquid at *a*, and (4) isentropic compression *ab* of water in a feed pump from pressure p_2 to pressure p_1.

This cycle more closely approximates the operations in a real steam power plant than does the Carnot cycle. Between given temperature limits it offers a lower ideal thermal efficiency for the conversion of heat into work than does the Carnot standard. Losses from irreversibility, in turn, make the conversion efficiency in an actual plant less than the Rankine cycle standard. *See* CARNOT CYCLE; REFRIGERATION CYCLE; THERMODYNAMIC CYCLE; VAPOR CYCLE.

Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; G. N. Hatsopoulos and J. H. Keenan, *Principles of General Thermodynamics*, 1965, reprint 1981; J. B. Jones and G. A. Hawkins, *Engineering Thermodynamics: An Introductory Textbook*, 1960.

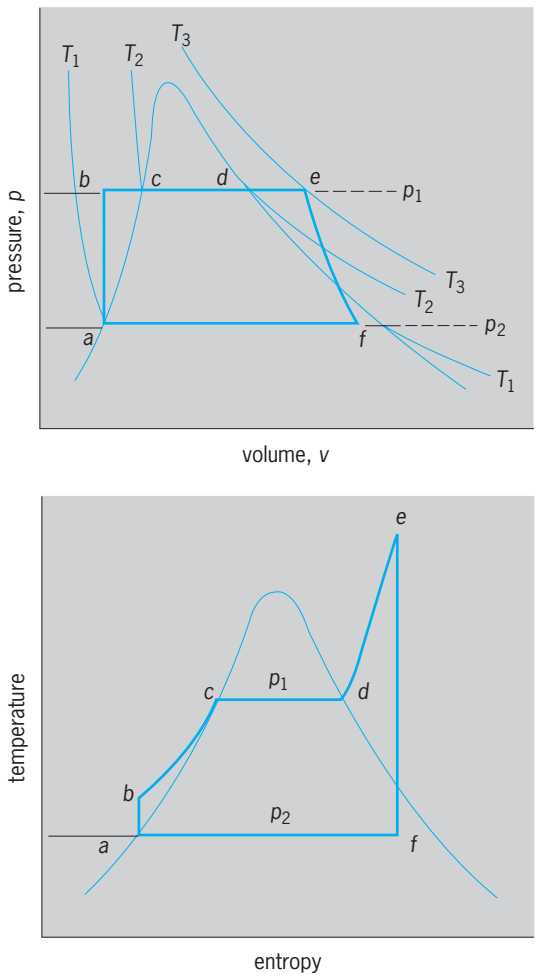

Rankine-cycle diagrams (pressure-volume and temperature-entropy) for a steam power plant using superheated steam. Typically, cycle has four phases.

Ranunculales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Magnoliidae of the class Magnoliopsida (dicotyledons). The order consists of 8 families and about 3200 species. The vast majority of the species belong to only 3 families, the Ranunculaceae (2000), Berberidaceae (650), and Menispermaceae (425). Within its subclass, the order is characterized by its mostly separate carpels, triaperturate pollen, herbaceous or only secondarily woody habit, frequently numerous stamens, generally more than two sepals, and lack of ethereal oil cells. Many members of the order produce benzyl-isoquinoline alkaloids or aporphine alkaloids, or both. The Ranunculales are generally considered to be derived from the Magnoliales, a woody order in the same subclass. The barberry (*Berberis*), in the family Berberidaceae, and the buttercup (*Ranunculus*), columbine (*Aquilegia*; see **illus.**), larkspur (*Delphinium*), and windflower (*Anemone*), in the family Ranunculaceae, are familiar genera.

Colorado columbine (*Aquilegia coerulea*) of family Ranunculaceae. (*U.S. Forest Service photograph by C. A. Kutzleb*)

See MAGNOLIALES; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM. Arthur Cronquist; T. M. Barkley

Rape

Rape (*Brassica napus*) and turnip rape (*B. campestris*) plants are members of the Cruciferae family. The name is derived from the Latin *rapum*, meaning "turnip," to which these plants are closely related. The aerial portions of rape plants have been bred to produce oilseeds, fodder, and vegetable crops. Rape seed is small (0.07 to 0.18 oz or 2 to 5 g per 1000 seeds), round, and usually black, although varieties with yellow seed coats are also grown. The seeds, borne in long slender pods or siliques, contain over 40% oil. The plant germinates rapidly and forms a rosette of bluish-green leaves from which bolts an indeterminate racemose inflorescence (see **illus.**). Both annual and biennial forms of the crop are grown; the biennial form will not flower without extended exposure to freezing temperatures. *See* CAPPARALES.

Brassica oilseed crops (including *B. juncea*) are the world's third most important source of edible vegetable oil, with production centered in Canada, northern Europe, China, and the Indian subcontinent. Seed, oil, and meal from rapeseed plants that contain superior nutritional qualities are known worldwide as canola. Canola plants were developed by genetically blocking the biosynthesis of the nutritionally undesirable fatty acid, erucic, and sulfur compounds called glucosinolates. Canola oil is a highly nutritious salad and cooking oil and is used in the manufacture of margarine and shortening. It is also used in Europe as a diesel fuel, while rape oil is used industrially in the manufacture of plastics and as a lubricant. Canola meal, the by-product of oilseed extraction, is a high-quality protein feed supplement for livestock and poultry. In some Asian and Indian countries, where conversion to canola quality varieties has not yet occurred, rape meal is used as a fertilizer.

Well-podded plants of annual rape (*Brassica napus*) with inset showing the characteristic inflorescence and basal leaves.

Rape is widely used for forage. In some countries the whole plant is cut and fed to cattle, but more frequently biennial rape, such as the Dwarf Essex variety, is used as summer or late fall and winter pasture for sheep and swine. It is in the latter form that rape is most widely used throughout the United States and New Zealand. Some forms of turnip rape are used as leaf vegetables in Asia. *See* TURNIP. R. K. Downey

Bibliography. R. K. Downey and R. Rimmer, Agronomic improvement in oilseed brassicas, *Adv. Agron.*, 50:1–66, 1993; G. Röbbelen, R. K. Downey, and A. Ashri (eds.), *Oil Crops of the World*, 1989; F. Shahidi (ed.), *Canola and Rapeseed*, 1990.

Raphidophyceae

A small class of poorly known biflagellate unicellular algae (raphidomonads) in the chlorophyll *a–c* phyletic line (Chromophycota). It is sometimes called Chloromonadophyceae, and its members called chloromonads, but this nomenclature is confusing because it calls to mind *Chloromonas*, a totally unrelated genus of green algae (Chlorophyceae). The alternative nomenclature used here is derived from *Raphidomonas*, a generic name within the class.

The cell of a typical raphidomonad is flattened, dorsiventral or bilaterally symmetrical, and naked. Two flagella are borne at the base of an anterior groove. One flagellum has hairs and is directed anteriorly, while the other is smooth and trails close to the cell body. The cytoplasm is segregated into an endoplasm, which contains the nucleus and mitochondria, and an ectoplasm, which contains numerous discoid chloroplasts (in photosynthetic forms), an anterior assemblage of small contractile vacuoles (in fresh-water forms), and commonly ejectile structures (trichocysts and mucocysts). Chloroplasts have four membranes and three thylakoids per lamella. Photosynthetic pigments include chlorophyll *a* and *c*, β-carotene, and various xanthophylls, of which fucoxanthin has been reported only in marine forms. Depending upon pigment composition, the chloroplasts are bright green, yellow-green, or yellow-brown. Pyrenoids are present only in marine forms. Reserve food particles are unknown, and the storage product apparently is oil. Eyespots are absent. The nucleus is relatively large, located in the center or anterior portion of the cell, and capped by an array of Golgi bodies. *See* CELL PLASTIDS; CILIA AND FLAGELLA.

Gelatinous palmelloid colonies composed of nonflagellate spheroidal cells have been observed.

Reproduction apparently is restricted to longitudinal binary fission, which can take place while the cell is motile or palmelloid.

Two families are recognized. All photosynthetic raphidomonads are placed in the Vacuolariaceae. The Thaumatomastigaceae comprises a few colorless forms that bear pseudopodia. They are osmotrophic or phagotrophic or both.

Most genera in both families occur in fresh water (acidic to neutral), but there are brackish-water and marine forms that produce conspicuous blooms. *Heterosigma*, for example, is a frequent cause of red tide in the inland seas of Japan. *See* ALGAE. Paul C. Silva; Richard L. Moe

Bibliography. E. R. Cox (ed.), *Phytoflagellates*, 1980.

Rare-earth elements

The group of 17 chemical elements with atomic numbers 21, 39, and 57–71; the name lanthanides is reserved for the elements 58–71. The name rare earths is a misnomer, because they are neither rare nor earths. The early Greeks believed that everything in the world was made of four elements: air, earth, fire, and water. The earths were substances which could not be changed with heat by the temperatures then available to the scientist; and in the early part of the nineteenth century, when the first rare earths were discovered, they resembled the common earths, which were really oxides of magnesium, calcium, and aluminum. Since the rare earths were found in very rare minerals, they were thus called rare earths. They are not rare, however, since cerium is reported to be more abundant in the Earth's crust than tin; yttrium more abundant than lead; and even the scarce rare earths, except promethium, more abundant than the platinum-group elements. All these elements form trivalent bonds, and when their salts are dissolved in water, they ionize to form trivalent ions and the solutions exhibit very similar chemical properties. The elements scandium, yttrium, lanthanum, and actinium in the II column of the extended periodic table show similar properties in aqueous solution. Yttrium and lanthanum are always found associated with the rare earths in nature.

The similarity of properties among the lanthanide rare earths (atomic numbers 58 through 71) originates from the fact that as the atomic number increases in this part of the periodic table, the increased charge on the nucleus is compensated for by electrons which start filling an inner incomplete subshell (4*f*). This subshell can hold up to 14 electrons; there are accordingly 14 elements in the lanthanide series. These extra electrons, however, play almost no role in the valency forces between atoms. The elements with atomic numbers 90 through 103 also occur in the periodic table at a place where a similar inner subshell (5*f*) is being completed. In many ways, these elements resemble the lanthanides. They are frequently referred to as the actinide rare earths. Both groups are usually displayed at the bottom of the periodic table as an appendage of two rows. *See* ACTINIDE ELEMENTS; LANTHANIDE CONTRACTION; PERIODIC TABLE.

Properties. The rare-earth elements are metals possessing distinct individual properties which make them potentially valuable as alloying agents. They are usually reduced thermally by treating the anhydrous halide with calcium, lithium, or other alkali metals and then remelting under vacuum to volatilize the last traces of the reductant. They can also be reduced eletrolytically from fused-salt baths, as is done commercially for cerium and misch metal (a mixed rare-earth metal, mainly cerium, with small amounts of iron present).

Samarium, europium, and ytterbium cannot be reduced by the above methods. These elements form divalent halides when so treated. Fortunately, these metals have very low boiling points so they can be distilled from a mixture of lanthanum or cerium and the oxides of these elements. The pure samarium, europium, and ytterbium are collected on the condenser, and lanthanum or cerium oxides are left behind.

Table 1 gives the properties of the metals. The anhydrous salts show greater differences in properties among the elements than do the hydrated salts.

Many of the properties of the metals and alloys are quite sensitive to temperature and pressure. They are also different when measured along different crystal axes of the metal; for example, electrical conductivity, elastic constants, and so on. If the properties are plotted against temperature and pressure, they show anomalies whenever a crystal or magnetic transition takes place.

The rare earths form organic salts with certain organic chelate compounds. These chelates, which have replaced some of the water around the ions, enhance the differences in properties among the individual rare earths. Advantage is taken of this technique in the modern ion-exchange methods of separation. *See* CHELATION; CHROMATOGRAPHY; ION EXCHANGE; MAGNETOCHEMISTRY; TRANSITION ELEMENTS.

Occurrence. Although the rare earths are widely distributed in nature, they generally occur in low concentrations. They also are found in high concentrations as mixtures in a number of minerals (**Table 2**). The relative abundance of the different rare earths in various rocks, geological formations, and the stars is of great interest to the geophysicist, astrophysicist, and cosmologist. The development of precision methods for determining the abundance of the various elements present in trace amounts by means of mass spectroscopy, optical and infrared spectroscopy, and radio and nuclear chemistry permits the relative abundance of the rare earths in any sample to be determined with considerable precision. It is found that the deep basic rocks, such as basalt, contain a few parts per million of mixed rare earths, and that the more acid silicate rocks contain a slightly higher concentration. When these two types of rocks come in contact in the molten state, the more acid rocks extract some of the mixed rare

TABLE 1. Properties of the rare-earth metals

Symbol	Transitions	Transition temperature, °C (°F)	ΔH transition, cal	Lattice parameters, Å		c/a	Ionic radius RE^{3+}, nm
				a	c		
Sc	hcp (ABAB)	—	—	3.3088	5.2680	1.592	0.0732
	bcc	1335 (2435)	958	—	—	—	—
Y	hcp (ABAB)	—	—	3.6482	5.7318	1.571	0.0893
	bcc	1478 (2692)	1189	4.08	—	—	—
La	dhcp (ABAC)	—		3.7740	12.171	1.6125*	0.1061
	fcc (ABC)	310 (590)	67	—	—	—	—
	bcc	865 (1589)	753	4.26	—	—	—
Ce	fcc (ABC)	—	—	4.85	—	—	—
	dhcp (ABAC)	–178 (–288)	—	3.68	11.92	1.619*	0.1034
	fcc (ABC)	~100 (heating) (212)	—	5.160	—	—	Ce^{4+}
		–10 (cooling) (14)	—	—	—	—	0.0092
	bcc	726 (1339)	700	4.12	—	—	
Pr	dhcp (ABAC)		—	3.6721	11.832	1.611*	0.1013
	bcc	795 (1463)	760	4.13	—	—	Pr^{4+}
							0.090
Nd	dhcp (ABAC)	—	—	3.6583	11.7966	1.612*	0.0995
	bcc	863 (1586)	713	4.13	—	—	
Pm							0.0979
Sm	hcp (nonprimitive)	—	—	3.6290	26.207	1.605	0.0964
	bcc	926 (1699)	744	—	—	—	Sm^{2+}
							0.111
Eu	bcc	—	—	4.5827	—	—	0.0950
							Eu^{2+}
							0.109
Gd	hcp (ABAB)	—	—	3.6336	5.7810	1.59	0.0938
	bcc	1238 (2255)	935	4.05	—	—	—
Tb	hcp (ABAB)	—	—	3.6055	5.6966	1.580	—
	bcc	1289 (2352)	1203	—	—	—	0.0923
							TB^{4+}
		—	—	—	—	—	0.084
Dy	hcp (ABAB)			3.5915	5.6501	1.573	0.0908
	bcc	1381 (2518)	955	—	—	—	—
Ho	hcp (ABAB)	—	—	3.5778	5.6178	1.570	0.0894
Er	hcp (ABAB)	—	—	3.5592	5.5850	1.569	0.0881
Tm	hcp (ABAB)	—	—	3.5375	5.5540	1.570	0.0869
Yb	fcc (ABC)	—	—	5.4848	—	—	0.0858
	bcc	795 (1463)	418	—	—	—	Yb^{2+}
							0.093
Lu	hcp (ABAB)			3.5052	5.5495	1.583	0.0848

*The values are the author's opinion of the best average values. The actual values vary slightly in the original literature, depending on the investigator and the purity of the metal at hand. Most of the values were obtained on 99.8–99.9% pure metal, the principal impurities being hydrogen, oxygen, nitrogen, carbon, and tantalum. Some of these values will change slightly as purer metal are produced.

earths from the basalt, and the light rare earths are extracted to a larger extent than the heavy ones. Thus, by determining a profile of the relative abundance of the rare earths in given rocks, it can be determined whether they have been molten. Such tests were performed on the various rocks brought back from the Moon. Also, if deposits of rich rare-earth minerals are found in a formation, one can determine the physical condition with regard to temperature and pressure to which the formation must have been subjected in order to form such minerals. The relative abundance of the various rare earths present in stars as determined from their spectra gives clues as to what nuclear processes are taking place in those stars. The numerous scientific papers published in this field make interesting reading, since theories as to the formation of the solar system and the universe are in part based on such data. Monazite, xenotime, and bastnasite are among the more important sources. These minerals are usually concentrated from other rock and minerals by mechanical means, such as flotation or magnetic cross-belt separation methods. The rare earths are leached from the minerals with acid in the case of the phosphate or silicate minerals. Some minerals, such as the columbotantalates, have to be heated with carbon or treated with strong caustic before undergoing the leaching process. *See* ORE DRESSING.

Separation. The mixed rare earths can be separated from the acid solutions by means of oxalate precipitation or by other insoluble precipitates; ignition of the oxalate gives a mixed rare-earth oxide. They are frequently concentrated directly by ion-exchange methods, using the acid leach from the minerals.

The rare earths occur in solution as hydrated trivalent ions whose properties are very much alike; therefore they tend to form mixed crystal precipitates or solid solutions. A single chemical operation

Melting point, °C (°F)	Boiling point, °C (°F)	Heat of vaporization ($\Delta H_{v,0}$), kcal/g-atm	Density 25°C, g/cm^3	Molal volume 25°C, cm^3/g-atm	Radius metal atom, nm	Electrical resistivity (4.2 K), ohm-cm 10^{-6}	Residual resistivity (4.2 K), ohm-cm 10^{-6}	Compressibility, cm^2/kg 10^{-6}
1541	2831 (5128)	89.9	2.9890	15.041	0.1640	52	3	2.26
1522	3338 (6040)	101.3	4.4689	19.894	0.1801	59	2	2.68
921 (1690)	3457 (6255)	103.1	6.1453	22.603	0.1879	61–80	Superconductor	4.04
799 (1470)	3426 (6199)	101.1	6.672	21.001	0.1820	70–80	10	4.10
931 (1708)	3512 (6353)	85.3	6.773	20.805	0.1828	68	1	3.21
1021 (1870)	3068 (5554)	78.5	7.007	20.585	0.1821	65	7	3.0
1168 (2134)	2700 (est.) (4892)	—	—	—	—	—	—	(2.8)
1077 (1971)	1791 (3256)	49.2	7.520	20.001	0.1804	91	7	3.34
822 (1531)	1597 (2907)	(41.9) $\Delta H^\circ = 29$	5.2434	28.981	0.1984	91	1	8.29
1313 (2395)	3266 (5911)	95.3	7.9004	19.9041	0.1801	127	1	2.56
1356 (2473)	3123 (5653)	93.4	8.2294	19.3119	0.1783	114	4	2.45
1412 (2574)	2562 (4646)	70.0	8.5500	19.0058	0.1774	100	5	2.55
1474 (2685)	2695 (4883)	72.3	8.7947	18.7533	0.1766	88	3	2.47
1529 (2874)	2863 (5145)	76.1	9.066	18.4499	0.1757	71	3	2.39
1545 (2813)	1947 (3537)	55.8	9.3208	18.1244	0.1746	74	3	2.47
819 (1506)	1194 (2181)	36.5	6.9654	24.8428	0.1939	28	2	7.39
1663 (3025)	3395 (6143)	102.2	9.8404	17.7808	0.1735	60	2	2.38

only slightly enriches one rare earth over another; thus, to isolate pure compounds of the individual elements by these methods, the processes have to be repeated many times.

Historically, the elements were purified by fractional processes such as fractional crystallization or fractional decomposition. The enormous amount of work that was involved permitted the separation of only very small amounts; consequently, the pure rare earths were very costly and gained the reputation of being rare. Fractionation methods are still used commercially to separate crude rare earths, particularly for lanthanum and cerium, since cerium can then be separated from lanthanum by taking advantage of the quadrivalent state of cerium. The other members of the rare-earth series, if desired in high purity, are separated by means of ion-exchange processes; if great purity is not desired, liquid-liquid extraction processes can be employed. Some of the rare earths which show anomalous valence are separated from the others by taking advantage of this property.

Yttrium and europium, used in the television industry, are usually separated by means of liquid-liquid extraction. Europium shows the divalent characteristic, and advantage can be taken of this. Yttrium, which occurs in another row in the periodic table, can be shifted either in liquid-liquid extraction or ion exchange along the rare-earth series by means of different complexing agents. Advantage can also be taken of the fact that some minerals contain very little of the heavy rare earths, but this fraction contains the yttrium, and liquid-liquid processes can be quite effective. Furthermore, when ton quantities are desired, the liquid-liquid extraction processes are usually more economical.

Uses. Most of the early uses of the rare earths took advantage of their common properties and were centered principally in the glass, ceramic, lighting,

TABLE 2. Some common rare-earth minerals

Mineral	Crystal form	Formula composition*
Monazite†	Monoclinic	$CePO_4$ with $Th_3(PO_4)_4$
Xenotime	Tetragonal	YPO_4
Gadolinite	Monoclinic	$2BeO \cdot FeO \cdot Y_2O_3 \cdot 2SiO_2$
Bastnasite†	Hexagonal	$CeFCO_3$
Samarskite	Orthorhombic	$3(Fe,Ca,UO_2)_3O \cdot Y_2O_3 \cdot 3(Nb,Ta)_2O_5$
Fergusonite	Tetragonal	$Y_2O_3 \cdot (Nb,Ta)_2O_5$
Euxenite	Orthorhombic	$Y_2(NbO_3)_3 \cdot Y_2(TiO_3)_3 \cdot 1^1/_2H_2O$
Yttrofluorite	Cubic	$2YF_3 \cdot 3CaF_2$

*Only the most abundant rare earth is listed. Ce minerals are rich in light members of the rare-earth group. Y minerals are rich in heavy members of the rare-earth group.
†One of the two most important minerals used by the rare-earth industry in the United States.

and metallurgical industries. Today these applications use a very substantial amount of the mixed rare earths just as they are obtained from the minerals, although sometimes these mixtures are supplemented by the addition of extra cerium or have some of their lanthanum and cerium fractions removed.

Cerium oxide or the mixed rare-earth oxides make excellent abrasives for glass polishing, and several million pounds of these materials are sold each year for polishing mirrors, television face plates, lenses, and plate glass. The rare earths have long been used in the manufacture of specialty glasses and in the ceramic industry, particularly as glazes.

The elements exhibit very complex spectra, and the mixed oxides, when heated, give off an intense white light which resembles sunlight. Consequently mixtures of the oxides are used in cored carbon arcs, such as those employed in the movie industry.

The rare-earth metals have a great affinity for the nonmetallic elements, as, for example, hydrogen, carbon, nitrogen, oxygen, sulfur, phosphorus, and the halides. Considerable amounts of the mixed rare earths are reduced to metals, such as misch metal, and these alloys are used as "getters" in the metallurgical industry. The rare-earth metals combine with the nonmetals in other metals and alloys, and these impurities are partially removed along with the slag or, as in nodular cast iron, the rare earths convert the carbon inclusions into nodules, and this procedure imparts considerable strength to the iron. Alloys made of cerium and the mixed rare earths are used in the manufacture of lighter flints.

Rare earths are also used in the petroleum industry as catalysts. The rare-earth ions are introduced into zeolite-type oxide catalyst, say in the form of the chloride. The resulting catalyst, which contains about 5% mixed rare-earth oxides, produces petroleum with a higher yield of the desired petroleum fraction.

Since the individual rare earths have been available commercially at reasonable prices, a sizable market has developed for a number of them. This is not surprising, since they represent, as a group, about one-sixth of the naturally occurring elements and about one-fourth of the elemental metals. The fact that these elements form compounds and alloys, many of whose properties change in a predictable manner as one progresses across parts of the series, makes them of particular interest to scientists working in various fields of research. These scientists develop theories about numerous experimental phenomena, and they can use the rare-earth metals, alloys, compounds, and solutions as test materials to see whether the theories still apply when tested on a series of neighboring elements whose properties shift in a known manner.

The television industry has been using highly purified yttrium and europium oxides in large quantities. A europium-activated yttrium phosphor emits an intense red light when used in a television screen. Use of this type of phosphor greatly improves the color reproduction, and pictures are much more brilliant than when a non-rare-earth phosphor was used. The mercury-activated electric street lights, which are economical to operate, give a purplish-blue color; for about the same energy consumption they can be made to give a brilliant white light by the use of rare-earth phosphors deposited on the glass.

Another very important use of individual rare earths is in the manufacture of solid-state microwave devices widely used in radar and communications systems. Yttrium-iron garnets are especially good, since they transmit short-wave energy with low energy losses. The devices, however, are very small, so the total use of rare earths is not large.

Still another important use of individual rare earths is in the construction of lasers. A high percentage of the patent applications for new lasers involves the use of a rare earth as the active constituent of the laser.

The rare earths show interesting magnetic properties. Alloys of cobalt with the rare earths, such as cobalt-samarium, produce permanent magnets that are far superior to most of the varieties on the market, and many uses for these magnets are developing.

Many phosphors contain rare earths, and barium phosphate-europium phosphor finds applications in x-ray films that form satisfactory images with only half the exposure time of conventional x-rays.

Yttrium aluminum garnets (YAG) are used in the jewelry trade as artificial diamonds. Single crystals of YAG have a very high refractive index, similar to diamond, and when these crystals are cut in the form of diamonds, they sparkle in the same manner as a diamond. Also, they are very hard, so that they scratch

glass, and only an expert can tell the difference between a YAG diamond and a real one.

Other small but important uses are mentioned in the articles on the individual rare earths. (For the list of rare-earth elements see Table 1.)

Frank H. Spedding

Bibliography. R. J. Elliott (ed.), *Magnetic Properties of Rare Earth Metals*, 1972; K. Gschneidner and L. Eyring, *Handbook on the Physics and Chemistry of Rare Earths*, vol. 1: *Metals*, vol. 2: *Alloys and Intermetallics*, 1978; C. M. Lederer et al., *Table of Isotopes*, 6th ed., 1967; J. E. Powell et al., The separation of rare earths, *J. Chem. Educ.*, 37:629–633, 1960; F. H. Spedding and A. H. Daane (eds.), *Rare Earths*, 1961.

Rare-earth minerals

Naturally occurring solids, formed by geological processes, that contain the rare-earth elements—the lanthanides (atomic numbers 57–71) and yttrium (atomic number 39)—as essential constituents. In a rare-earth mineral, at least one crystallographic site contains a total atomic ratio of lanthanides and yttrium that is greater than that of any other element. The mineral name generally has a suffix, called a Levinson modifier, indicating the dominant rare-earth element; for example, monazite-(La) [$LaPO_4$] contains predominantly lanthanum, and monazite-(Ce) [$CePO_4$] contains predominantly cerium. *See* MINERAL; MONAZITE; PERIODIC TABLE; RARE-EARTH ELEMENTS.

So far, about 170 distinct species of rare-earth minerals have been described. A large number of carbonates, phosphates, silicates, niobates, and fluorides are known as rare-earth minerals. However, no sulfide of the rare earths has been reported as a mineral species. It is necessary to obtain structural as well as chemical information about a mineral to judge the essentiality of its rare-earth elements (that is, whether the rare-earth element is part of the mineral's ideal formula or is basically an impurity). Sometimes, minerals with significant rare-earth content are treated as rare-earth minerals, even if the rare-earth element content appears unessential to the mineral. More than 60 mineral species, including the apatite group minerals, garnet group minerals, and fluorite, are in this category. *See* APATITE; FLUORITE; GARNET.

Occurrence. Despite their name, rare-earth elements are not extremely rare in the Earth's crust. Their abundance is not as low as the rarest elements such as gold and mercury. Rare-earth elements are contained in many common rock-forming minerals as trace impurities, and are concentrated by geological activities to form rare-earth minerals. The concentration depends on many factors such as the relative content of the rare-earth elements in source melts or fluids, oxidation states, temperature, pressure, and fit of ionic radius to developing crystal structure. Rare-earth minerals can be observed \pagebreak as accessory minerals in igneous rocks, such as monazite-(Ce) in granite. Carbonatite is the typical host rock of rare-earth minerals such as bastnäsite-(Ce) [$CeCO_3F$] and monazite-(Ce). Rare-earth minerals also often occur in pegmatite. In both carbonatite and pegmatite, rare-earth elements are concentrated by primary crystallization from melt and by hydrothermal reactions. Carbonatite deposits containing rare-earth elements are found throughout the world in such places as Mountain Pass, California, and Bayan Obo, China. Chemically stable, rare-earth minerals are not weathered easily. As a result, they have been deposited as heavy minerals in beach sand. Such deposits are found in Southeast Asia and Western Australia. *See* CARBONATITE; PEGMATITE.

Among the rare-earth minerals, bastnäsite-(Ce) is the most important source of rare-earth elements. Monazite-(Ce), synchysite-(Ce) [$CaCe(CO_3)_2F$], xenotime-(Y), britholite-(Ce) [$(Ce_3Ca_2)(SiO_4)_3(OH)$], and allanite-(Ce) [$CaCeAl_2Fe(Si_2O_7)(SiO_4)O(OH)$] are also sources. Rare-earth elements have been leached with acid from the surface of clay minerals. Rare-earth minerals containing radioactive nuclear species, such as thorium and uranium, are not used as source materials. *See* CLAY MINERALS.

Chemistry. The rare-earth minerals always contain a variety of rare-earth elements. The study of the rare-earth elemental distribution in minerals and rocks has been an active subject in geochemistry and mineral chemistry for many years. Elemental distribution follows Oddo-Harkins's rule: in natural minerals, a lanthanide with an even atomic number is more abundant than the neighboring lanthanides with odd atomic numbers. *See* GEOCHEMISTRY; MINERALOGY.

Rare-earth minerals can be classified into the following three groups based on their distribution of rae-earth elements: those with a predominance of the Y-group rare-earth elements, such as yttrium, having relatively small ionic radii; those rich in the Ce-group rare-earth elements, such as cerium, having larger ionic radii; and those in which either Y-group or Ce-group rare-earths elements may predominate (for example, gadolinite-(Y) [$Y_2FeBe_2Si_2O_{10}$] and gadolinite-(Ce) [$Ce_2FeBe_2Si_2O_{10}$]).

Rare-earth elements are usually trivalent in minerals. Notable exceptions are the divalent europium (Eu^{2+}) and tetravalent cerium (Ce^{4+}) ions. Crystal structures have been reported for about half of the known rare-earth minerals. The structure analyses have revealed that many of them are solid solutions of rare-earth ions and other heterovalent ions, such as sodium (Na^+), calcium (Ca^{2+}), and thorium (Th^{4+}), with ionic radii comparable to those of rare-earth ions. It is known that the similarity of ionic radii and electronegativities are important factors for isomorphous substitutions (substitution of similar-sized ions within crystal structures), whereas the similarity of chemical properties is not as significant. Studies of isomorphous substitution in rare-earth minerals are important for understanding the behavior of rare-earth elements in solids. *See* ELECTRONEGATIVITY.

Crystal structure. Crystallographic studies are impeded by the fact that some rare-earth minerals always occur in the metamict state, an amorphous (glasslike) state mainly caused by radiation damage to crystal structures due to radioactive decay of elements such as uranium and thorium. Crystallographic data and crystal structures of metamict minerals are therefore determined using samples recrystallized by annealing. In such cases, there is some doubt about the correlation between the crystal structures of recrystallized phases and the original structures of pre-metamict minerals. *See* METAMICT STATE; RADIOACTIVE MINERALS.

Anionic groups. Rare-earth minerals fall into six classes based on the anionic groups in their crystal structure:

1. Minerals with crystal structures containing isolated triangular anionic groups—carbonates such as lanthanite-(Nd) [$Nd_2(CO_3)_3 \cdot 8H_2O$] and bastnäsite-(Ce) [$Ce(CO_3)F$].

2. Minerals with crystal structures containing only isolated tetrahedral anionic groups—phosphates such as xenotime-(Y) [YPO_4] and monazite-(Ce) [$CePO_4$].

3. Minerals with crystal structures containing linked anionic groups of tetrahedral ions—silicates such as thalenite-(Y) [$Y_3Si_3O_{10}(OH)$] and yttrialite-(Y) [$Y_2Si_2O_7$].

4. Minerals with crystal structures containing tetrahedral and octahedral ions—aluminum silicate and titanium silicate such as allanite-(Ce) [$CaCeAl_2Fe(Si_2O_7)(SiO_4)O(OH)$] and mosandrite [$Na(Na,Ca)_2(Ca,Ce)_4Ti(Si_2O_7)_2(O,F)_2F_2$].

5. Minerals with crystal structures containing anionic groups of octahedral ions—titanates and niobates such as aeschynite-(Ce) [$CeTi_2(O_5OH)$] and fergusonite-(Y) [$YNbO_4$].

6. Minerals with crystal structures without anionic groups—fluorides and simple oxides such as fluocerite-(Ce) [CeF_3] and cerianite-(Ce) [CeO_2].

Isomorphous substitution. The coordination numbers of the rare-earth atoms are generally high, and their coordination polyhedra are as large as those of sodium ions, calcium ions, thorium ions, and so on. Y-group rare-earth elements and Ce-group rare-earth elements show a difference in their coordination numbers. The coordination numbers of the Y-group range from 6 to 11. Those of the Ce-group range from 7 to 12. The coordination numbers of calcium resemble those of the Ce-group; hence isomorphous substitution between calcium and Ce-group rare-earth elements has been observed often.

The charge compensation mechanisms for mutual substitution of heterovalent ions found in the structures of rare-earth minerals have been classified into the following four types (where RE represents a generic rare-earth element):

1. Substitution accompanying vacancies (denoted by □).

$$3Ca^{2+} \rightleftharpoons 2RE^{3+} + \square$$
$$RE^{3+} + \square \rightleftharpoons Ca^{2+} + Na^{+}$$

2. Coupled substitution within a rare-earth site.

$$2RE^{3+} \rightleftharpoons Ca^{2+} + Th^{4+}$$
$$2Ca^{2+} \rightleftharpoons RE^{3+} + Na^{+}$$

3. Coupled substitution at two independent sites.

$$RE^{3+} + Na^{+} \rightleftharpoons Ca^{2+} + Ca^{2+}$$
$$RE^{3+} + Si^{4+} \rightleftharpoons Ca^{2+} + (P,As)^{5+}$$
$$RE^{3+} + Ti^{4+} \rightleftharpoons Ca^{2+} + (Nb, Ta)^{5+}$$
$$RE^{3+} + Mg^{2+} \rightleftharpoons Ca^{2+} + Al^{3+}$$
$$RE^{3+} + (OH)^{-} \rightleftharpoons Ca^{2+} + H_2O$$
$$RE^{3+} + O^{2-} \rightleftharpoons Ca^{2+} + (OH, F)^{-}$$

4. Substitution accompanying valence variation.

$$RE^{3+} + Fe^{2+} \rightleftharpoons Ca^{2+} + Fe^{3+}$$

In almost all rare-earth minerals, complex solid solutions result from these isomorphous substitutions. Type 1, accompanying a vacancy site, may occur only in crystal structures having tight frameworks formed by covalent bonding, because the number of atoms changes before and after the substitution. Type 2 is a simple one occurring at an independent site. These are the most common substitutions observed in the crystal structures of rare-earth minerals because they do not require any other concomitant substitution. The complex solid solutions formed by types 3 and 4 make it especially difficult to study their chemical compositions and crystal structures.

Polyhedra shape. The shapes of the coordination polyhedra of rare-earth atoms are, in almost all cases, not regular polyhedra, but distorted polyhedra because the bonds between the rare-earth atoms and anions are ionic rather than covalent and their coordination numbers are high. Considering the distortion of polyhedra, the coordination polyhedra composed of rare earths and surrounding anions can be expressed as an octahedron, a monocapped trigonal prism, a cube, a dodecahedron, a square antiprism, an icosahedron, and so on.

The differences in coordination numbers and sizes of coordination polyhedra between the Y- and Ce-groups are reflected in the crystal structures of carbonates and phosphates. These minerals do not have infinite framework structures, and the isolated anionic groups in the structures are connected by rare-earth atoms to form three-dimensional structures. Therefore, the size of the rare earth–polyhedron affects the whole structure. On the other hand, structures having a tight framework, as seen in the structures of gadolinite-(Y), allanite-(Ce), and aeschynite-(Y), are not affected by the sizes of the rare earth–polyhedra, because they can accept rare earths of various sizes without rearrangement of the structure. *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Ritsuro Miyawaki

Bibliography. K. A. Gschneidner, Jr., and L. Eyring (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, vol. 16, Elsevier Science, Amsterdam, 1993; A. P. Jones, F. Wall, and C. T.

Williams (eds.), Rare Earth Minerals: Chemistry, Origin and Ore Deposits, *Mineralogical Society Series*, no. 7, Chapman & Hall, London, 1996; B. R. Lipin and G. A. McKay (eds.), Geochemistry and Mineralogy of Rare Earth Elements, *Reviews in Mineralogy*, vol. 21., Mineralogical Society of America, Washington DC, 1989.

Rarefied gas flow

Flow of gases below standard atmospheric pressure, sometimes called low-pressure gas flow. The flow may be confined to pipes between a chamber or vessel to be evacuated and a pump, or it may be the beam of molecules issuing from an orifice into a large evacuated chamber or the plume of exhaust gases from a rocket launched into the upper atmosphere, for example. The flow velocity is measured with respect to a fixed boundary such as the wall of a pipe, the surface of a rocket or jet plane, or a model in a wind tunnel. *See* FLUID FLOW; GAS; MOLECULAR BEAMS; PIPE FLOW; VACUUM PUMP; WIND TUNNEL.

For flow through ducts, the gases concerned are initially those of the original atmosphere inside a chamber that must be evacuated. However, even after the bulk of the original gas has been removed, the pumps must continue to remove gas evolved from surfaces and leaking in through imperfections in the walls. In some cases, gas is introduced through valves at a controlled rate as part of the process being carried out at a low pressure.

Since the flow through pipes involves an interaction or drag at the walls, a pressure drop is generated across the entrance and exit of the pipe. Also, gaseous impurities from the pump may flow toward the chamber when the pressure is very low. Proper design of the duct system therefore involves selecting pipes and valves of adequate internal diameter to ensure a minimal pressure drop and the insertion of baffles or traps to prevent impurities from the pumps from entering the process chamber.

Flow regimes. The resistance due to the walls depends on the mass flow velocity υ, and may depend on the gas viscosity η and the pressure P or density ρ of the gas. The mean free path λ of molecules between collisions with other molecules in the gas can be calculated from Eq. (1), where υ_m is the

$$\lambda = \frac{2\eta}{\rho \upsilon_m} \tag{1}$$

mean molecular speed in a gas obeying the Maxwell-Boltzmann distribution law in a reference frame moving with the stream velocity υ. The speed υ_m is given by Eq. (2), and the density ρ is given by the perfect gas law, Eq. (3). Here, M is the molecular weight

$$\upsilon_m = \left(\frac{8R_0T}{\pi M}\right)^{1/2} \tag{2}$$

$$\rho = \frac{MP}{R_0T} \tag{3}$$

(in kg/mol in SI units, and in g/mol in cgs units); P is the pressure; R_0, given by Eq. (4), is the molar

$$R_0 = 8.314\frac{\text{J}}{\text{mol}\cdot\text{K}} \qquad \text{(SI units)}$$
$$= (8.314\times 10^7)\frac{\text{erg}}{\text{mol}\cdot\text{K}} \qquad \text{(cgs units)} \tag{4}$$

gas constant; and T is the absolute temperature. *See* KINETIC THEORY OF MATTER; MEAN FREE PATH; VISCOSITY.

The analysis of low-pressure flow is divided into three or four flow regimes depending on the value of the Knudsen number Kn defined as the ratio of the mean free path λ to a characteristic length d [Eq. (5)],

$$\text{Kn} = \frac{\lambda}{d} \tag{5}$$

and the dimensionless Reynolds number, defined by Eq. (6). The characteristic length d may be chosen

$$\text{Re} = \frac{\upsilon\rho d}{\eta} \tag{6}$$

as the mean pipe diameter in the case of confined flow or as some length associated with a test model suspended in a wind tunnel, for example. *See* GAS DYNAMICS; KNUDSEN NUMBER; REYNOLDS NUMBER.

Another dimensionless number used in gas flow dynamics is the Mach number, defined as the ratio of the mass flow velocity υ to the local velocity of sound υ_a in the gas. It is given by Eq. (7), in which γ

$$\text{Ma} = \frac{\upsilon}{\upsilon_a} = \frac{2\upsilon}{\upsilon_m}\left(\frac{2}{\pi\gamma}\right)^{1/2} \tag{7}$$

is the ratio of the specific heat at constant pressure to the specific heat at constant volume. When the gas molecules are diffusely scattered from the walls, these numbers are related by Eq. (8).

$$\text{Kn} = \left(\frac{\pi\gamma}{2}\right)^{1/2}\frac{\text{Ma}}{\text{Re}} = \frac{\lambda}{d} \tag{8}$$

See MACH NUMBER.

When the mean free path is much smaller than the pipe diameter (Kn < 0.01), the gas flows as a continuous viscous fluid with velocity near the axis of the pipe at locations well beyond the pipe entrance much higher than the velocity in gas layers near the wall. The velocity profile as a function of radial distance from the axis depends on the distance from the entrance and the viscosity of the gas. When the Reynolds number is less than 2000, the profile is a simple curved surface so that the flow is laminar (laminar flow regime). When the mean free path becomes greater than about 0.01 times the diameter, the profile is distorted by boundary-layer effects, and the velocity near the wall does not approach zero (sometimes referred to as slip flow). *See* LAMINAR FLOW.

For Reynolds numbers above the critical value (approximately 2100), the flow is subject to instabilities depending on the geometry of the boundary and at high Reynolds numbers becomes turbulent (turbulent flow regime). *See* TURBULENT FLOW.

When the mean free path is about equal to or greater than the pipe diameter (Kn $\geq$ 1), the gas molecules seldom collide with each other, but can either pass through the pipe without striking the wall or scatter randomly back and forth between various points on the wall and eventually escape through the exit or pass back through the entrance. This type of gas flow is known as free-molecule flow (molecular flow regime). The transition region (0.01 < Kn < 1) between the laminar flow regime and the molecular flow regime is referred to as the Knudsen or transition flow regime.

The flow may also be classified by the boundary conditions or by the Mach number. For example, Couette flow involves the flow of rarefied gas between two surfaces that are moving with respect to each other with different parallel tangential velocities. For hypersonic flow, Ma $\geq$ 5.

Vacuum system pressure. Pressure may be expressed by the SI unit, the pascal (Pa; one newton of force per square meter), or the commonly used unit in vacuum technology, the torr (approximately 133 Pa). Standard atmospheric pressure is 101,325 Pa. Pressure measurement usually involves inserting the tubulation of a vacuum gage through the wall of a vacuum system with the plane of the opening parallel to the axis of flow so that so-called static pressure is measured, but ionization-type vacuum gages are sometimes installed directly inside the walls with no envelope or tubulation so that gas molecules can arrive from all directions. Such so-called nude gages then measure gas density or molecular concentration rather than pressure. When the temperature varies widely over different parts of a vacuum system, it is usually preferable to analyze the gas flow in terms of mass flow or molecular flux and the average number of molecules in unit volume. *See* PRESSURE MEASUREMENT; VACUUM MEASUREMENT.

Many applications require a high vacuum, corresponding to pressures less than 0.1 Pa. These usually involve a process chamber connected by a short duct of large cross section equal to or greater than the intake area of a high-vacuum pump, such as a diffusion pump, turbomolecular pump, or cryopump. Diffusion pumps and turbomolecular pumps require a forepressure below a limiting value and therefore must be connected by a pipe to a backing pump or forepump, such as a mechanical rotary-piston pump or rotary blower pump (**Fig. 1**). The connecting pipe can be relatively long and of small diameter since the mean pressure in the foreline is generally high enough for the flow to be viscous and the pressure drop is small up to the entrance of the backing pump.

Pumping speed. Since most high-vacuum pumps and mechanical forepumps have a broad range of intake pressures over which the volumetric flow rate of gas is approximately constant, it is customary to express the quantity of gas flowing into the pump in pressure–volume units (at room temperature) rather than mass units. In designing vacuum systems, it is essential to know the maximum volumetric flow rate at room temperature, or pumping speed S_0, for each pump. The speed is measured by the manufacturer by bleeding air (or other gases) through a calibrated control valve into a chamber (or test dome) placed directly over the intake entrance and measuring the pressure near the entrance according to an industry standard. The actual speed S_p varies with the absolute temperature T and molecular weight M of the gas. For high-vacuum pumps it can usually be assumed that Eq. (9) holds, where S_0 is the speed for air at 296 K (73°F).

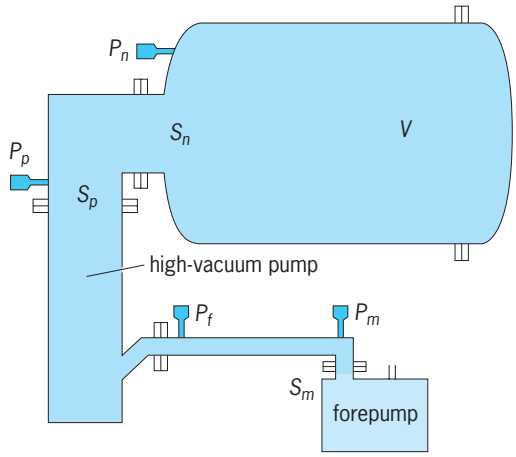

Fig. 1. Vacuum system consisting of a process chamber of volume V at pressure P_n, high-vacuum pump of speed S_p, and forepump of speed S_m with connecting pipes. P_p = pressure near entrance to high-vacuum pump; P_f = discharge pressure (forepressure) at exit of high-vacuum pump; P_m = intake pressure of forepump; S_n = flow rate of gas from the process chamber through the opening to the pipeline.

$$S_0 = \left(\frac{28.98\,T}{296\,M}\right)^{1/2} \tag{9}$$

Common units for pumping speed are liters per second (L/s or l/s); cubic feet per minute (usually abbreviated cfm; sometimes ft^3/min); or the SI unit, cubic meters per second (m^3/s).

Throughput. The quantity of gas in pressure–volume units, at a specified temperature, flowing through a cross section of the pipe in unit time is called the throughput. Unless otherwise specified, it is understood that the prevailing gas temperature is that of the walls near the point at which the pressure is measured; in most cases this will be room temperature. For steady flow through a pipe with walls at a uniform temperature, the throughput is a conserved quantity when no adsorption or desorption occurs on the walls.

When the pressure P_p in the gas near the entrance to the pump is measured under conditions similar to those in the standard test dome used to measure pump speed S_p for the same gas, the throughput is given by Eq. (10).

$$Q = S_p P_p \tag{10}$$

Most gases at low pressures obey the ideal gas law of Eq. (3), so that the mass rate of flow in g/s

is given by Eq. (11), where the numerical value of

$$Q_m = Q\frac{M}{R_v T} \tag{11}$$

the molar gas constant R_v is 62.36 when Q is in torr · L/s, 8.314×10^3 when Q is in Pa · L/s, or 8.314 when Q is in Pa · m^3/s.

Conductance and resistance. The pressure P_n in the process chamber depends on the resistance to flow in the pipeline between the chamber and the pump. This resistance W_{np} is defined by Eq. (12), where P_p

$$Q = \frac{P_n - P_p}{W_{np}} = C_{np}(P_n - P_p) \tag{12}$$

is the pressure at the junction with the pump and C_{np} is the conductance, or reciprocal of the resistance. Similarly, for the pipe between a high-vacuum pump and a mechanical forepump of speed S_m at the intake pressure P_m, the steady-state throughput is given by Eq. (13), where P_f is the discharge pres-

$$Q = C_{fm}(P_f - P_m) = S_m P_m \tag{13}$$

sure (or forepressure) at the exit of the high-vacuum pump and C_{fm} is the conductance of the pipe between the points at which P_f and P_m are measured (Fig. 1).

The conductance C_{np} is independent of internal pressure since the flow occurs in the molecular flow regime, but the flow in the line between high-vacuum pump and forepump may be in the transition regime for certain values of Q and will be in the viscous flow regime during the initial pump-down from atmospheric pressure. The conductance of a pipe in the viscous or transition flow regime increases with the mean pressure.

Net pumping speed. The volumetric flow rate of gas from the chamber of volume V through the opening to the pipeline at the system temperature is known as the net speed, S_n as given by Eq. (14), where P_n is the

$$S_n = \frac{Q_0}{P_n} - \frac{V}{P_n}\frac{dP_n}{dt} \tag{14}$$

pressure in the chamber, and Q_0 is the quasi-steady-state gas load due to outgassing, leaks, and so forth, in throughput units, which can be regarded as constant during the time required for initial evacuation. When dP_n/dt is small, equations of continuity can be written as Eq. (15), and Eq. (16) follows. Equa-

$$\begin{aligned} Q &= S_n P_n = C_{np}(P_n - P_p) = S_p P_p \\ &= C_{fm}(P_f - P_m) = S_m P_m \end{aligned} \tag{15}$$

$$\frac{1}{S_n} = \frac{1}{C_{np}} + \frac{1}{S_p} \tag{16}$$

tions (12)–(16) are valid only for steady-state flow and uniform temperature throughout the system, or when the variables have been referred to standard room temperature by correction factors.

Molecular flow regime. A conductance C_{12} can be defined in terms of the pressure drop, $P_1 - P_2$, across any section of a complex pipe system, but depends on the geometry of the system at the points where P_1 and P_2 are measured. If a section of the pipeline of length L and uniform cross section is located between two relatively large chambers in which the pressures, as measured some distance from the pipe entrance or exit planes, are respectively P_1 and P_2, then the molecular flow conductance of the pipe section will include end corrections.

The molecular flow conductance can be calculated from Eq. (17). Here, f is the effusion law factor

$$C_{12} = \frac{Q}{P_1 - P_2} = fA_1 w_{12} \tag{17}$$

for molecular flow through an orifice in a thin plate, given by Eq. (18) [where Eq. (2) has been used to ex-

$$f = \frac{v_m}{4} = \left(\frac{R_0 T}{2\pi M}\right)^{1/2} \tag{18}$$

press v_m]; A_1 is the entrance area; and w_{12} is defined as the probability that a molecule entering the pipe will cross the exit and not return to the entrance chamber.

The transmission probability w_{12} for a circular pipe of radius R between two large chambers is given approximately by Eq. (19), where L' is an end-

$$w_{12} = \frac{L'}{L + L'} \tag{19}$$

correction term, which varies from $L' = {}^8/_3\ R$ for pipes with $L/R > 100$ to $L' = 2R$ for L/R approaching zero. More exact equations for w_{12} for straight pipes of circular (and some other) cross sections have been derived. The transmission probability factor w for straight circular tubes between large chambers is plotted as a function of L/R in **Fig. 2**.

The overall transmission probability w_{13} for gas from a large chamber entering a pipeline consisting of a first pipe of uniform cross section A_1 connected in series by a short transition piece, or by an annular junction, to a second pipe of uniform cross section A_2

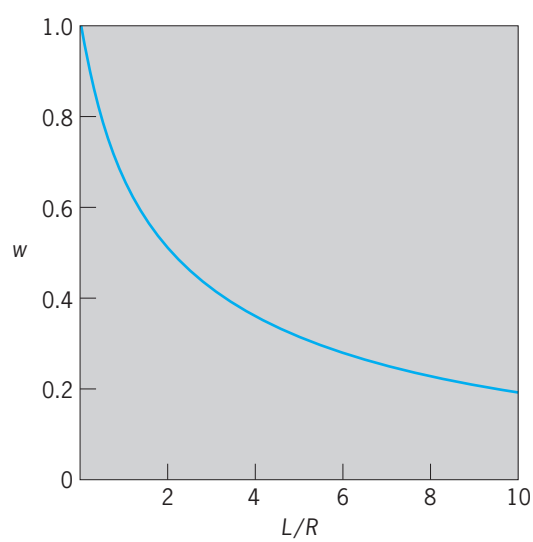

Fig. 2. Molecular flow transmission probability *w* for straight cylindrical tube of length *L* and radius *R*.

exiting into a large chamber is given approximately by Eq. (20). Here, w_{12} is the transmission probabil-

$$\frac{(1/w_{13})-1}{A_1} = \frac{(1/w_{12})-1}{A_1} + \frac{(1/w_{23})-1}{A_2} \quad (20)$$

ity through the first pipe, calculated as though the pipe were located by itself between two large chambers, and w_{23} is similarly the transmission probability for the second pipe alone. This equation can be extended to any number of different pipe sections in series.

The molecular flow transmission probability for pipes with elbows and for ducts with baffles or complex shapes can be obtained from Monte Carlo computer programs for the trajectory of molecules crossing the entrance. *See* MONTE CARLO METHOD.

For an orifice of area A_1 in a thin plate between two chambers of cross section much larger than A_1, the value of w_{12} is approximately equal to 1, corresponding to $L = 0$ in Eq. (19).

In common units, when L is much greater than L', the conductance (in L/s) of a straight circular tube of radius R (in cm), length L (in cm), and effusion law factor f (in cm/s) is given by Knudsen's formula, Eq. (21). For air at 23°C (73°F), the effusion law factor $10^{-3}f = 11.63$ L/(s · cm^2).

$$C_{12} = \frac{{}^{8}/_{3}(10^{-3}f)\pi R^3}{L} \quad (21)$$

Laminar flow regime. For long straight tubes ($L > 20R$) of length L (in cm) and uniform radius R (in cm) in the laminar flow regime, the conductance (in L/s) is given by the Poiseuille equation (22),

$$C_{fm} = \frac{(\pi R^4/8\eta L)P_{fm}}{100} \quad (22)$$

where P_{fm} is the mean pressure in pascals, given by Eq. (23), and η is the viscosity in poise. For short tubes a correction factor must be introduced.

$$P_{fm} = \frac{P_f + P_m}{2} \quad (23)$$

Transition flow regime. The Knudsen semiempirical formula for conductance (in L/s) of a long circular tube of length L (in cm) and radius R (in cm) in the transition regime is Eq. (24), where C_{fm} is the

$$C_t = C_{fm} + F_k C_{12} \quad (24)$$

Poiseuille flow term given by Eqs. (22) and (23), C_{12} is the molecular flow conductance given by Eq. (21), and F_k is given by Eq. (25). Here, f is given by

$$F_k = \frac{1 + (2/\pi)^{1/2}R(10P_{fm}/f\eta)}{1 + 1.24(2/\pi)^{1/2}R(10P_{fm}/f\eta)} \quad (25)$$

Eq. (18) in centimeters per second, and P_f by Eq. (23) in pascals. Benjamin B. Dayton

Bibliography. G. A. Bird, *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*, 1994; W. Fiszdon, *Rarefied Gas Flows*, 1982; E. P. Muntz, D. H. Campbell, and D. P. Weaver (eds.), *Rarefied Gas Dynamics*, 3 vols., 1994; H. Oguchi, *Rarefied Gas Dynamics*, 2 vols., 1985; A. Roth, *Vacuum Technology*, 3d ed., 1990, reprint 1998; G. L. Saksaganskii, *Molecular Flow in Complex Vacuum Systems*, 1988.

Raspberry

The horticultural name for certain species of the genus *Rubus*, plant order Rosales. In these species the fruit, when ripe (unlike the blackberry), separates thimblelike from the receptacle. Raspberry plants are upright shrubs with perennial roots and prickly, biennial canes (stems). There are several species, both American and European, from which the cultivated raspberries have been developed. Varieties are grouped as to color of fruit—black, red, and purple, the last being hybrids between the red and black types. Red raspberries, with upright canes, are propagated by suckers or by root cuttings (**Fig. 1**);

Fig. 1. Red raspberry branch, Loudon variety. (*USDA*)

the black varieties, with long canes which arch over and touch the ground, are propagated by tip layers. The hybrid purple varieties are usually progagated by tip layers. Raspberry breeding has been carried on so extensively that all the important red and purple and most of the black varieties are the result of breeding experiments. Raspberries are grown extensively in home gardens over most of the United States. Leading states in commercial production are Michigan, Oregon, New York, Washington, Ohio, Pennsylvania, New Jersey, and Minnesota. The fruit is sold fresh for dessert purposes, is canned, and is made into jelly or jam, but quick freezing is the most important processing method. *See* BREEDING (PLANT); ROSALES. J. Harold Clarke

Raspberry plants are seriously damaged by diseases caused by viruses and viruslike organisms, bacteria, fungi, and nematodes.

Raspberry mosaic and raspberry leaf curl are spread by aphids and are common in most areas

Fig. 2. Anthracnose disease on canes (stems) of black raspberry.

in the United States where raspberries are grown. Tomato ringspot virus, spread by dagger nematodes (*Xiphinema* sp.), and pollen-spread raspberry bushy dwarf virus weaken plants and cause crumbly fruit. The use of resistant cultivars, certified stocks, isolation from infected plantings, and nematode control (for tomato ringspot virus) are recommended controls.

Crown gall caused by *Agrobacterium tumefaciens* is characterized by tumorlike outgrowths at ground level on infected raspberry plants. Certified stocks should be planted where fruit crops have been grown. *See* CROWN GALL.

Fungi cause diseases of raspberry leaves, canes, roots, fruits, and of the whole plant. Raspberry leaf spot, caused by *Sphaerulina rubi*, causes whitish spots on leaves. Powdery mildew, caused by *Sphaerotheca humuli*, involves entire young leaves in dusty white growth. Anthracnose, caused by *Elsinoë veneta* (**Fig. 2**), spur blight, caused by *Didymella applanata*, and yellow rust, caused by *Phragmidium rubi-idaei*, can produce spots on leaves and cankers on canes. *Leptosphaeria coniothyrium* causes cane blight when it gains entrance into tissues through wounds. Resistant cultivars, proper pruning, and approved fungicides help control these diseases. Verticillium wilt caused by *Verticillium dahliae* is a root-infecting fungus that produces wilting and death of canes. Avoidance of land previously planted with solanaceous crops (like potatoes) helps control this disease. Red raspberry fruits are susceptible to gray mold fruit rot caused by *Botrytis cinerea*, and less commonly to rots by other fungi including *Alternaria, Cladosporium*, and *Rhizopus* sp. Prompt picking and marketing, cool handling, and application of approved fungicides aid in fruit rot control. Orange rust, caused by *Gymnoconia peckiana*, systemically infects black raspberry plants. Growing points are invaded, and the fungus produces abundant orange pustules on leaves at harvest time. Use of certified planting stock controls the disease. Red raspberries are immune. *Pratylenchus* and *Xiphinema* are the most important plant-parasitic nematodes that feed on raspberry roots and weaken the plants. *Xiphinema* also spreads certain viruses. Control is achieved by using certified planting stock and by not planting in soil with high population of these nematodes. *See* PLANT PATHOLOGY.

Richard H. Converse

Bibliography. M. A. Ellis et al., *Compendium of Raspberry and Blackberry Diseases and Insects*, 1991; N. W. Frazier, *Virus Diseases of Small Fruits and Grapevines*, USDA Farmers Bull. 2208, 1978.

Ratchet

A wheel, usually toothed, operating with a catch or a pawl so as to rotate in a single direction (see **illus.**).

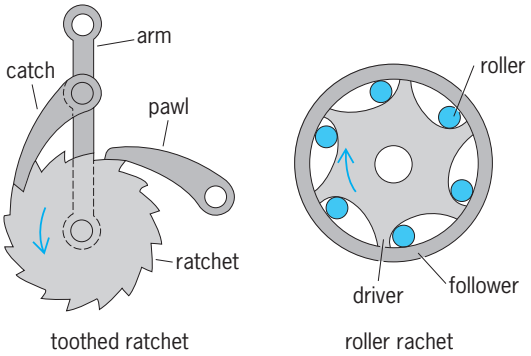

Toothed ratchet is driven by catch when arm moves to left; pawl holds ratchet during return stroke of catch. In roller ratchet, rollers become wedged between driver and follower when driver turns faster than follower in direction of arrow.

A ratchet and pawl mechanism locks a machine such as a hoisting winch so that it does not slip. The locking action may serve to produce rotation in a desired direction and to disengage in the undesired direction as in a drill brace. A further adaptation is to drive the catch in a to-and-fro motion against the ratchet to produce intermittent circular motion. The catch or pawl may be of various shapes such as an eccentrically mounted disk or ball bearing. Gravity, a spring, or centrifugal force (with the catch mounted internal to the ratchet) are commonly used to hold the pawl against the ratchet. A ratchet and pawl provides an arresting action, whereas an escapement provides an arresting action followed by a self-initiated momentary release. In high-speed machines, the abrupt action of a toothed ratchet produces severe shock. In such situations, a continuously variable yet directionally sensitive action is achieved by wedging rollers or specially shaped sprags between the input and output members. *See* BRAKE; ESCAPEMENT; PAWL.

Frank H. Rockett

Bibliography. R. H. Creamer, *Machine Design: Mechanical Engineering Technology*, 3d ed., 1984; A. Esposito and J. R. Thrower, *Machine Design*, 2d ed., 1991; K. Lingaiah, *Machine Design Data Handbook*, 2d ed., 2001.

Rate-of-climb indicator

An aircraft instrument that provides an indication of the vertical change of the aircraft position within the air mass. It is more commonly known as the vertical-velocity or vertical-speed indicator. Contained within a sealed case, it is connected to the aircraft static pressure source, the so-called at-rest air pressure outside the aircraft, through a calibrated leak (see **illus.**). Although the instrument operates from a static pressure source, it is a differential pressure indicator. The differential pressure is established between the static pressure in the diaphragm or pressure capsule and the trapped static pressure within the case. When the aircraft changes vertical position, the static pressure in the diaphragm changes immediately but, because of the metering action of the calibrated leak, the case pressure will remain at its prior value and cause the needle to show a change in vertical speed. The needle is usually calibrated in feet per minute but may be calibrated in any appropriate unit of length over time. Because of the calibrated leak in the system, a finite time (usually 6–9 s) is required for the pressure inside the instrument to equalize and a reliable rate indication to become available.

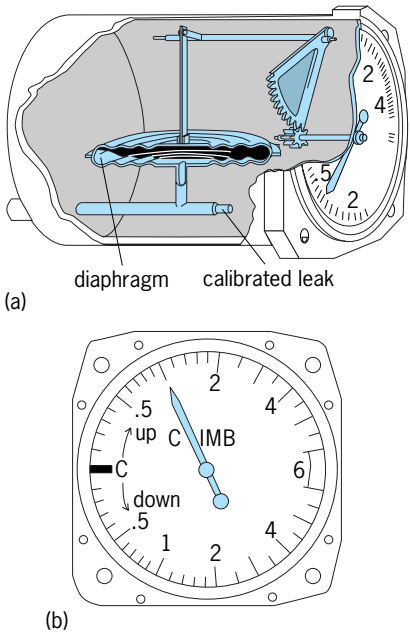

Rate-of-climb indicator or vertical-speed indicator. (*a*) Components. (*b*) Face. (*After Federal Aviation Administration, Airframe and Powerplant Mechanics Airframe Handbook, AC 65-15A, 1976*)

This lag induced by the calibrated leak has led to the introduction of a more reliable instrument known as the instantaneous vertical-speed indicator. It incorporates acceleration pumps to eliminate the limitations associated with the calibrated leak. During climb entry, the vertical acceleration causes the pumps to supply extra air into the diaphragm to stabilize the pressure differential without the usual lag time. The opposite phenomenon occurs during descent. During stabilized climbs and descents or level flight, the instrument operates on the same principles as the vertical-speed indicator. Both the vertical-speed indicator and instantaneous vertical-speed indicator usually incorporate a small screw on the lower left face of the instrument to allow ground adjustment to zero rate of climb.

In many modern aircraft with flight computers, the rate of climb or descent is electronically calculated by differentiating the altitude from the pitot-static source. *See* AIRCRAFT INSTRUMENTATION; PITOT TUBE.

Grady W. Wilson

Ratites

A group of flightless, mostly large, running birds characterized by their flat, keelless sternum, which were formerly segregated as a superorder of birds, the Palaeognathae, but whose interrelationships have been a long-standing controversy. Species include the ostriches (**Fig. 1**), emus, cassowaries, and rheas (**Fig. 2**). Traditionally the flying tinamous, in spite of their also possessing the typical palaeognathous palate, were not included in this group despite their close similarities to the rheas. Ornithologists, stressing the disjunctive distribution of these flightless birds and the many possible convergent features due to their large size, have concluded that the several groups of ratite birds represent end points of unrelated phyletic lineages from several different ancestral avian stocks and hence should be placed in separate orders. Comparative morphologists, on the other hand, stressing the basic similarity of the bony

Fig. 1. Ostrich (*Struthio camelus* spp. *massaicus*). (*Gerald and Buff Corsi © 2002 California Academy of Sciences*)

Fig. 2. Greater rhea (*Rhea americana*). (*Dr. Lloyd Glenn Ingles © 2001 California Academy of Sciences*)

palate and other anatomical features, have concluded that the ratites are a strictly monophyletic group of birds. The latter view is supported by evidence that the structure of the bony palate, type of cranial kinesis, and other cranial features are homologous in all ratites in spite of their other differences. Additional evidence such as general behavior, structure of the rhamphotheca, and some biochemical features also support this conclusion.

The ratites apparently represent two or three phyletic lineages [the African ostriches and elephant birds; the Australasian emus, cassowaries, moas and kiwis; and the Neotropical rheas and tinamous, the Australasian and Neotropical groups being perhaps more closely related to one another than to the African group] which evolved from a common ancestral stock—the Holarctic fossil Lithornithidae of the Paleocene-Eocene periods, which were possibly much like the still volant (capable of flight) tinamous of Central and South America. Each phyletic line probably reached its present center of distribution (African, Australasia, and the Neotropics) before becoming gigantic in size and flightless. The ratites are advanced birds, not a primitive avian group as argued by a number of earlier workers. *See* AVES; NEOGNATHAE; PHYLOGENY; STRUTHIONIFORMES. Walter J. Bock

Bibliography. S. Davis, *Ratites and Tinamous*, Oxford University Press, 1999; A. Feduccia, *The Origin and Evolution of Birds*, 2d ed., Yale University Press, 1999; B. Monroe and C. G. Sibley, *A World Checklist of Birds*, Yale University Press, 1997.

Rauwolfia

A genus of mostly poisonous, tropical trees and shrubs of the dogbane family (Apocynaceae). Certain species are the source of valuable emetics and cathartics. The species *Rauwolfia serpentina* (see **illus.**) has received special attention as the source of

Branch of *Rauwolfia serpentina.*

tranquilizing drugs. For centuries in India the drug has been used in the treatment of hypertension. It came into use in Western countries because of its effect in reducing blood pressure. Although not a sedative, as that term is usually construed, it often has a quieting influence on the patient. Among the purified alkaloids obtained from *R. serpentina*, reserpine (Serpasil) is perhaps the one most used as a tranquilizing agent. *See* GENTIANALES; TRANQUILIZER. Perry D. Strausbaugh; Earl L. Core

Raw water

Water obtained from natural sources such as streams, reservoirs, and wells. Natural water always contains impurities in the form of suspended or dissolved mineral or organic matter and as dissolved gases acquired from contact with earth and atmosphere. Industrial or municipal wastes may also contaminate raw water. *See* WATER POLLUTION.

If admitted to a steam-generating unit, such contaminants may corrode metals or form insulating deposits of sediments or scale on heat-transfer surfaces, with resultant overheating and possible failure of pressure parts.

Principal scale-forming impurities are compounds of calcium and magnesium, or silica. Principal corrosive agents are dissolved oxygen and carbon dioxide. In some localities, raw water has a mineral acidity. Oil and grease impair wetting and heat removal from the steam-generating surfaces and may also form corrosive scale or sludge. Certain organic materials or a high concentration of dissolved solids in the boiler water may cause foaming which contaminates the steam.

Raw water can be treated to remove objectionable impurities or to convert them to forms that can

be tolerated. For steam generation, suspended solids are removed by settling or filtration. Scale-forming hardness is diminished by chemical treatment to produce insoluble precipitates that are removable by filtration, or soluble compounds that do not form scale. Essentially complete purification is achieved by demineralizing treatment or evaporation. Demineralization consists of passing the water through beds of synthetic ion-exchange resin particles. Certain of these exchange hydrogen for metallic cations; others exchange hydroxyl for sulfate, chloride, or other anions in solution. The hydrogen and hydroxyl ions combine to form water. The resins may be used in separate or mixed-bed arrangements. They require periodic regeneration by acid and alkaline solutions, respectively; the mixed resins can be separated for such treatment by virtue of the fact that they differ in specific gravity. *See* WATER SOFTENING; WATER TREATMENT.

Evaporation requires expenditure of heat for complete vaporization of the water. The vapor is subsequently condensed and collected as purified distillate. Low-pressure steam is used as a heat source; multiple-effect heat exchange provides thermal economy. F. G. Ely

Reactance

The imaginary part of the impedance of an alternating-current circuit.

The impedance Z of an alternating current circuit is a complex number given by Eq. (1). The imaginary

$$Z = R + jX \tag{1}$$

part X is the reactance. The units of reactance, like those of impedance, are ohms. Reactance may be positive or negative. For example, the impedance of an inductor L at frequency ω is given by Eq. (2), so

$$Z = jL\omega = jX \tag{2}$$

X is positive, and X is referred to as the inductive reactance. The impedance of a capacitor C is given by Eq. (3), so X is negative, and X is referred to as the capacitive reactance.

$$Z = -\frac{j}{C\omega} = jX \tag{3}$$

The reactance of a circuit may depend on both the resistors and the inductors or capacitors in the circuit. For example, the circuit in the **illustration** has admittance [Eq. (4)] and impedance [Eq. (5)], so that the value of the reactance [Eq. (6)], depends on both the capacitor C and the resistor R.

$$Y = \frac{1}{R} + jC\omega \tag{4}$$

$$Z = \frac{R}{1 + jRC\omega} \tag{5}$$

$$X = -\frac{R^2C\omega}{1 + R^2C^2\omega^2} \tag{6}$$

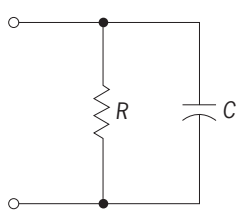

Circuit with a resistor and capacitor in parallel.

See ADMITTANCE; ALTERNATING-CURRENT CIRCUIT THEORY; ELECTRICAL IMPEDANCE. J. O. Scanlan

Reaction turbine

A power-generation prime mover utilizing the steady-flow principle of fluid acceleration, where nozzles are mounted on the moving element. The rotor is turned by the reaction of the issuing fluid jet and is utilized in varying degrees in steam, gas, and hydraulic turbines. All turbines contain nozzles; the distinction between the impulse and reaction principles rests in the fact that impulse turbines use only stationary nozzles, while reaction turbines must incorporate moving nozzles. A nozzle is defined as a fluid dynamic device containing a throat where the fluid pressure drops and potential energy is converted to the kinetic form with consequent acceleration of the fluid. For details of the two basic principles of impulse and reaction as applied to turbine design *see* IMPULSE TURBINE. Theodore Baumeister

Reactive intermediates

Unstable compounds that are formed as necessary intermediate stages during a chemical reaction. Thus, if a reaction in which A is converted to B requires that A first be converted to C, then C is an intermediate in the reaction (A → C → B). The term reactive further implies a certain degree of instability of the intermediate; reactive intermediates are typically isolable only under special conditions, and most of the information regarding the structure and properties of reactive intermediates comes from indirect experimental evidence.

Organic Reactions

In organic reactions, the most common types of reactive intermediates arise from dissociative reactions, in which a carbon atom dissociates from some appropriate leaving group, resulting in a carbon reactive intermediate with a decreased valence (**Table 1**). Associative reactions, in which some reactive species forms a new bond to carbon, can also give rise to some of the same classes of carbon reactive intermediates, as well as other classes in which carbon has increased valence. The most common types of reactive species to initiate associative reactions are nucleophiles which bring in a pair of electrons for the new bond; free radicals which bring in a single electron for the new bond; and electrophiles which contribute no electrons toward the new bond (**Table 2**). *See* VALENCE.

TABLE 1. Most common types of organic reactive intermediates*

Dissociative reaction	Intermediate	Simplest example
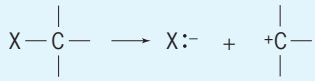	Carbocation	CH_3^+ (methyl cation)
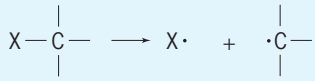	Free radical	$CH_3\cdot$ (methyl radical)
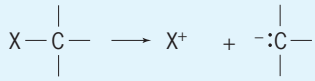	Carbanion	$CH_3{:}^-$ (methyl anion)
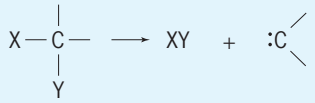	Carbene	$CH_3{:}$ (methylene)

*X^-, $X^\cdot$, X^+, XY represent appropriate leaving groups.

Carbocations. These are compounds in which carbon bears a positive charge. The simplest type of carbocation is a carbon atom that makes only three bonds and has no additional electrons. With a total of only six valence electrons (from the three bonds), such a carbon atom bears a positive charge. The simple trivalent carbocations have a planar structure. The presence of only six valence electrons represents an electron deficiency, and any structural features that offer electron donation stabilize the carbocation. Particularly effective are resonance effects, in which electron density is donated from a pair of electrons from an adjacent atom or from an adjacent pi bond or aromatic ring, as in notation (1).

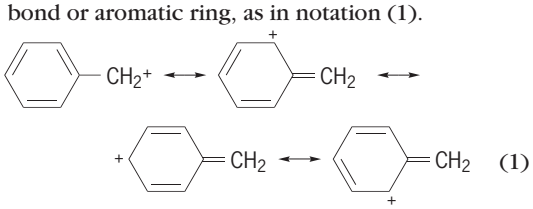

(1)

See RESONANCE (MOLECULAR STRUCTURE).

Reactivity of carbocations is dominated by their electron deficiency. Virtually all reactions of carbocations involve bonding to an electron pair, by covalent bonding to an anion or neutral molecule with an available pair of electrons (a nucleophile), or by bonding to a pair of electrons from a pi bond, or by shifting a pair of electrons from an adjacent bond. Carbocations are intermediates in a large variety of organic reactions (**Table 3**). *See* CHEMICAL BONDING.

Carbocations also exist in which carbon has the normal complement of eight valence electrons but still bears a positive charge. In these cases, sometimes called nonclassical carbocations, the key structural feature is a three-center, two-electron bond. The simplest example is the methonium ion, CH_5^+; and a typical reaction, electrophilic substitution of methane by deuterium ion, is shown in reaction (2).

TABLE 2. Organic reactive intermediates produced by associative reactions

Associative reaction	Intermediate
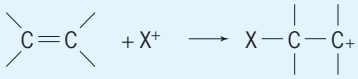	Carbocation (X^+ = electrophile)
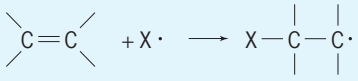	Free radical (X^- = free radical)
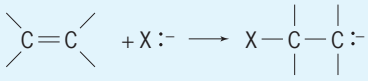	Carbanion (X^- = nucleophile)
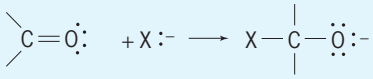	Tetrahedral intermediate (X^- = nucleophile)
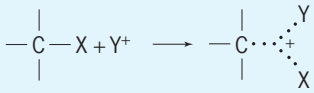	Carbonium ion (Y^+ = electrophile)

TABLE 3. Some reactions involving carbocation intermediates

Reaction	Reaction type
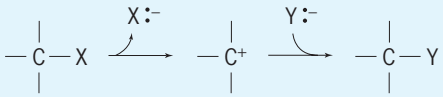	Nucleophilic substitution (X^- = leaving group, Y^- = nucleophile)
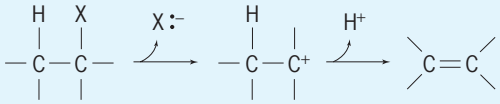	Elimination (X^- = leaving group)
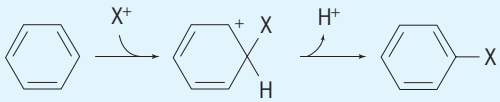	Electrophilic aromatic substitution (X^+ = electrophile)
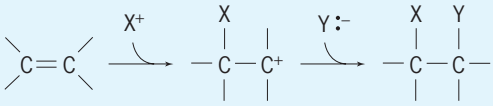	Electrophilic addition (X^+ = electrophile, Y^- = nucleophile)
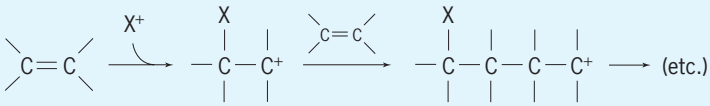 (etc.)	Cationic polymerization (X^+ = electrophile)

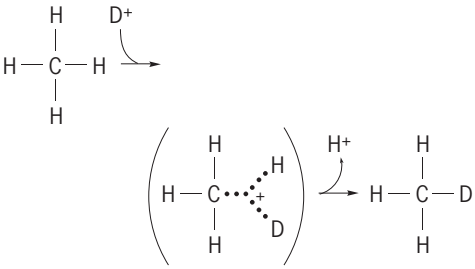

(2)

Carbanions. These are compounds in which carbon bears a negative charge. A carbanion will always have a positive counterion in associated with it; depending upon the particular cation and the stability of the carbanion, the association may be ionic, coavalent, or some intermediate combination of ionic and covalent bonding, as in notation (3), where M = metal.

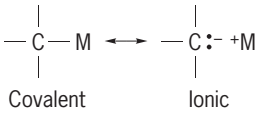

(3)

Compounds involving primarily covalent bonding of carbon to a metal are considered organometallic compounds, which frequently have properties similar to carbanions but substantially modified by the presence of the metal. *See* ORGANOMETALLIC COMPOUND.

Carbanions are trivalent, with eight valence electrons. Simple alkyl carbanions assume a pyramidal structure, analogous to ammonia, and are extremely strong bases. Carbanions are stabilized by structural features that allow electron withdrawal, in particular, inductive effects of halogen substituents or resonance effects of nitro (—NO_2), cyano (—CN), or carbonyl

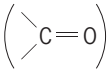

substituents. Such resonance-stabilized carbanions have a planar structure and are relatively easily formed in acid-base reactions. An example is shown in reaction (4).

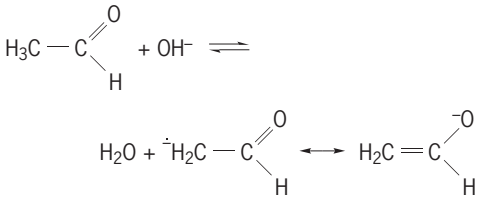

(4)

Resonance-stabilized enolate anion

Carbanions are electron-rich; their reactions are mostly with electron-deficient species (**Table 4**). The use of carbanions (or organometallic analogs) in alkylation and condensation reactions generates new carbon-carbon bonds and provides some of the most useful reactions for organic synthesis. *See* ORGANIC SYNTHESIS.

Free radicals. These are neutral compounds having an odd number of electrons and therefore one unpaired electron. Carbon free radicals are trivalent, with seven valence electrons, and typically assume a planar structure. Free radicals are primarily electron-deficient species and are stabilized by structural features that donate electron density or delocalize the odd electron by resonance. Structural features that prevent the normal reactions of radicals, such as large substituents that surround the free-radical site, are very effective in increasing radical stability. *See* DELOCALIZATION.

TABLE 4. Some reactions involving carbanion intermediates

Reaction	Reaction type
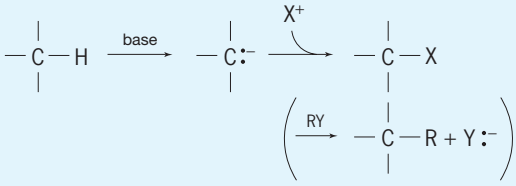	Electrophilic substitution (alkylation) (X^+ = electrophile, Y^- = leaving group)
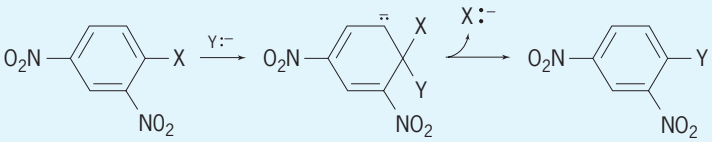	Nucleophilic aromatic substitution (Y^- = nucleophile, X^- = leaving group)
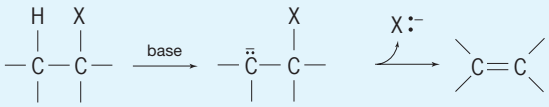	Elimination (X^- = leaving group)
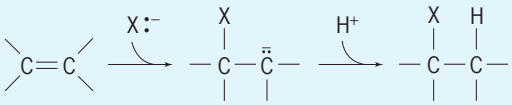	Nucleophilic addition (X^- = nucleophile)
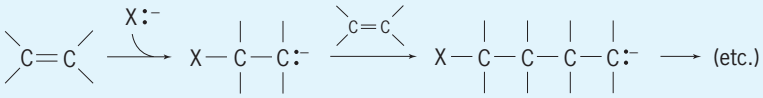	Anionic polymerization (X^- = nucleophile)
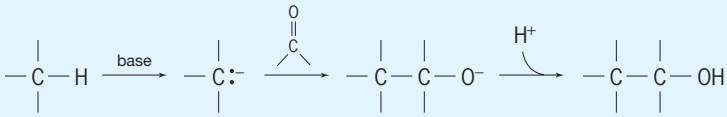	Condensation

An important characteristic of many reactions involving free-radical intermediates is the tendency to follow a chain reaction mechanism. An example is shown.

Initiation:

$$Cl_2 \rightarrow 2Cl\cdot$$

Propagation:

$$Cl\cdot + CH_4 \rightarrow CH_3\cdot + HCl$$
$$CH_3\cdot + Cl_2 \rightarrow CH_3Cl + Cl\cdot$$

Termination:

$$2Cl\cdot \rightarrow Cl_2$$
$$2CH_3 \rightarrow CH_3CH_3$$
$$CH_3\cdot + Cl\cdot \rightarrow CH_3Cl$$

Overall substitution reaction:

$$CH_4 + Cl_2 \rightarrow CH_3Cl + HCl$$

Addition by a free-radical chain reaction:

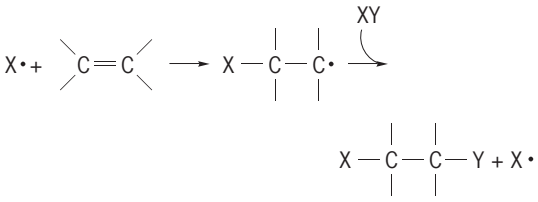

Radical chain polymerization:

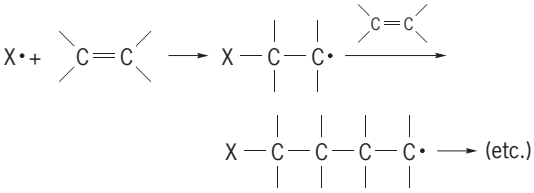

Chain reactions have three fundamental components: initiation, propagation, and termination steps. Initiation steps generate the reactive free radical from an appropriate precursor (or radical initiator). Propagation steps utilize the reactive free radical to generate product, and also regenerate another reactive free radical so that those steps may continue recycling. The overall reaction is the sum of the propagation steps, which may recycle thousands of times. Termination steps involve the combination of free radicals; each combination effectively terminates two chains.

Other common free-radical reactions include addition reactions, including polymerization. *See* CHAIN REACTION (CHEMISTRY); FREE RADICAL; POLYMERIZATION.

Radical ions. These compounds are charged and have an unpaired electron; they are either radical cations (positively charged) or radical anions (negatively charged). In many cases a radical ion is derived from a stable neutral molecule by addition of

one electron (radical anion) or removal of one electron (radical cation); thus, radical ions are frequently encountered as intermediates in electrochemical or other one-electron reduction or oxidation processes, as shown in the pinacol reduction reaction (5).

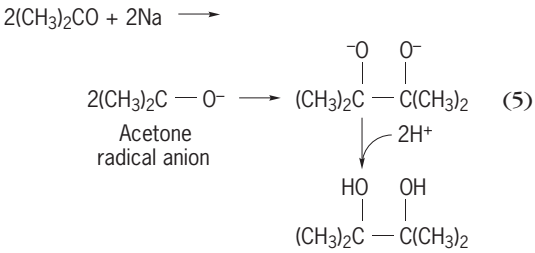

(5)

See ELECTROCHEMISTRY; OXIDATION-REDUCTION.

Carbenes. These are compounds which have a divalent carbon. The divalent carbon also has two nonbonded electrons, for a total of six valence electrons. The two nonbonded electrons may have either the same spin quantum number, which is a triplet state, or opposite spin quantum number, which is a singlet state. The triplet state of carbenes is typically lower in energy than the singlet state. Singlet and triplet carbenes normally undergo the same reactions, but because of the spin differences they must follow different mechanisms, so there are differences in reactivity, selectivity, and stereochemical details of the reactions. Typical singlet carbene reactions are concerted, involving a single step, and usually maintain stereochemistry. Triplet carbene reactions typically proceed through a triplet radical pair or a triplet biradical intermediate, which frequently randomizes stereochemistry. *See* ELECTRON SPIN; TRIPLET STATE.

Generation of carbenes is most commonly by photolysis or thermolysis diazo compounds or ketenes, or by alpha-elimination reactions. Carbenes are extremely reactive because of their electron deficiency. Typical reactions may be classified as reactions with single bonds (insertions) and those with double bonds (cycloadditions). Carbene cycloaddition reactions are among the best synthetic procedures for making cyclopropane derivatives. Internal rearrangements of carbenes are also common, depending on the particular structure of the carbene (**Table 5**). *See* PERICYCLIC REACTION.

TABLE 5. Reactions of carbenes

Reaction	Classification
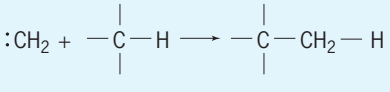	Insertion
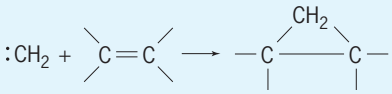	Cyclo-addition
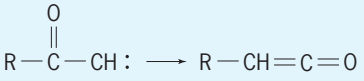	Wolff rearrangement

Carbenoids are carbene complexes which have reactivity similar to carbenes. Often their reactivity is moderated in a way that is very useful synthetically, such as the Simmons-Smith reagent (CH_2I_2-Zn), which undergoes cycloaddition reactions well without any competing insertion reactions.

Nitrenes. These are compounds which contain a monovalent nitrogen atom. The monovalent nitrogen will have a nonbonded pair of electrons and two additional electrons which may also be spin-paired (a singlet nitrene) or unpaired (a triplet nitrene). Nitrenes, like carbenes, have only six valence electrons and are highly electron-deficient. There are strong similarities between nitrenes and carbenes in terms of both their structure and their chemical reactivity. Nitrene insertion reactions form new C—N bonds, and nitrene cycloadditions provide a good synthetic route to aziridines. Nitrenes are conveniently generated by thermolysis or photolysis of azides.

Tetrahedral intermediates. These are compounds derived from addition of a nucleophile to a carbonyl group. Depending upon the substituents on the carbonyl group, the initially formed tetrahedral intermediate may give either an addition or a substitution product. Aldehydes and ketones give addition products, whereas carboxyl derivatives (including carboxylic acids, esters, amides, and acid anhydrides) give substitution products (**Table 6**).

Excited states. These are unstable electronic states of a molecule in which the electron distribution is different from the most stable distribution for the molecule (the ground state). Occasionally, vibrational excited states are encountered, in which a molecule contains an abnormally large amount of vibrational energy. Vibrational excited states lose their excess vibrational energy very rapidly, particularly in condensed phases, and generally the term excited states refers to electronically excited molecules. Typically, electronically excited states are generated by absorption of a photon of light; hence the chemistry of excited states is called photochemistry. Absorption promotes an electron from a normally occupied molecular orbital into a normally unoccupied molecular orbital (see **illus.**). Because of this electron redistribution, excited states are always better oxidizing agents as well as better reducing agents that the molecule in its ground state, as measured by ionization potential (IP; the energy to remove one electron) and electron affinity (EA; the energy gained on adding an electron). Thus many photochemical reactions involve one-electron transfer as their initial step.

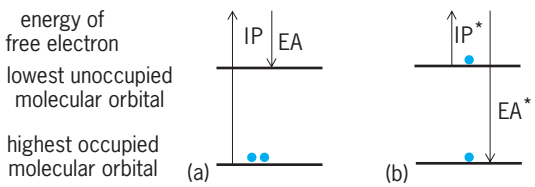

Ionization potentials (IP) and electron affinities (EA) for a molecule in (*a*) its ground state (A) and (*b*) its lowest electronically excited state (A*).

TABLE 6. Addition and substitution reactions via tetrahedral intermediates*

Reaction	Product
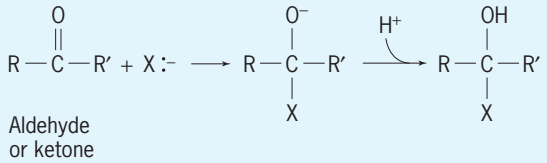 Aldehyde or ketone	Nucleophilic addition (X^- = nucleophile)
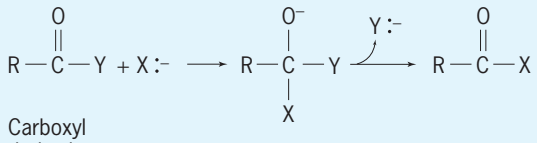 Carboxyl derivative	Nucleophilic substitution (X^- = nucleophile, Y^- = leaving group)

*R, R′ = C or H; Y = O, N, S, halogen.

Like carbenes and nitrenes, excited states have two unpaired electrons whose spin quantum numbers may be opposite (a single excited state) or parallel (a triplet excited state). A triplet excited state will always be lower in energy than the corresponding singlet excited state.

Because an excited state differs from the more stable ground state only in its electron distribution, the predominant unimolecular reactions of excited states involve return to the ground state, by emission of light (fluorescence from a singlet state, phosphorescence from a triplet state) or by conversion of the excess energy to heat. These unimolecular processes are very rapid, and excited state lifetimes are very short (for singlets, typically 10^{-6} to 10^{-12} s; for triplets, 10^{-6} to 10 s). Bimolecular reactions between excited states and other molecules also contribute to the deactivation of excited states. Such reactions include energy transfer, electron transfer, and a variety of specific reaction types depending on the compounds involved. However, to compete effectively with the rapid unimolecular deactivation of excited stated, bimolecular reactions must be extremely rapid, typically with rate constants close to diffusion control. *See* ELECTRON-TRANSFER REACTION.

Photochemistry results from excited states when a new compound is formed, either by unimolecular reaction of the excited state (rearrangement or fragmentation, for example) or by bimolecular reaction with another molecule (photoadditions, photooxidations, and photoreductions, for example). Some of the special features of photochemical reactions which differentiate them from normal (thermal) chemical rections include the following: (1) Compounds in a mixture may be selectively excited, depending upon the absorption spectrum of the compounds and the wavelength of light used. (2) Absorption of a photon provides a molecules with large amounts of energy (ultraviolet photons of 286 nanometers provide a total of 100 kcal or 420 kilojoules per mole) relative to thermal energy (around 1 kcal or 4.2 kJ per mole).

Considering these two points together, low-temperature photochemistry offers the opportunity for both the selective formation and the stabilization of compounds which would be highly unstable at ambient temperatures. Many highly strained reactive intermediates have been isolated and characterized under these conditions, including benzyne [reaction (6)]. Other intermediates which have been

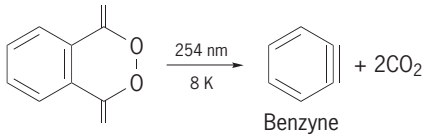

(6)

immobilized at low temperatures include free radicals, carbenes, nitrenes, and ylides. *See* PHOTOCHEMISTRY; YLIDE.

Other types. There are many kinds of organic reactive intermediates which do not fit into the previous classifications. Benzyne is a reactive intermediate which can be derived from an aryl halide by elimination with strong base: benzyne rapidly undergoes nucleophilic addition to give a net result of a substitution, reaction (7*a*), or may undergo cycloaddition, reaction (7*b*).

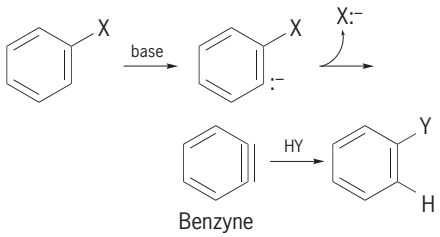

(7*a*)

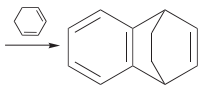
(7*b*)

See SUBSTITUTION REACTION.

Other reactive intermediates are simply compounds which are unstable for a variety of possible reasons, such as structural strain or an unusual oxidation state.

Carl C. Wamser

Inorganic Reactions

Inorganic reactions may involve reactive intermediates as atoms, ionic species, covalently bound radicals or radical ions, and coordination compounds. In the gas phase, intermediates occur mainly as free radicals. The complexity of inorganic intermediates arises mainly from the possible participation of more than 100 elements and their varied oxidation states. However, with the tremendous upsurge of inorganic and organometallic research since the mid-1960s, some of the complex mechanisms leading to these intermediates as well as their reactivity are more clearly understood. It is almost impossible to provide a uniform classification of the common inorganic intermediates in the same manner that is possible for organic reactive intermediates.

On the other hand, some organometallic intermediates resemble organic intermediates. Divalent species of group 14 metals such as germylenes with the structure

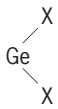

(X = F^-, Cl^-, or phenyl)

exhibit properties analogous to those of carbenes. Other examples are compexes containing transition metal (M) to carbon double bonds, such as cationic alkylidene complexes, $M^+{=}CH_2$. In some cases, these reactive intermediates could be regarded as metallocarbocations.

Inorganic reactive intermediates, such as metal atoms, radicals, radical ions, or metal ions in unusual valency states, are very reactive species. They can be generated either by ordinary chemical processes or by special methods, such as continuous or pulse radiolysis, photolysis, or electron beam vaporization. In particular in this section, intermediates proposed or "detected" in oxidation-reduction and substitution reaction of the metal complexes are discussed in detail. *See* OPTICAL PULSES; PHOTOLYSIS; ULTRAFAST MOLECULAR PROCESSES. For other types of processes or methods which could give rise to reactive inorganic intermediates; *see* ELECTROCHEMICAL PROCESS; HOMOGENEOUS CATALYSIS; INORGANIC PHOTOCHEMISTRY; LASER PHOTOCHEMISTRY; OSCILLATORY REACTION.

Inorganic radicals and radical ions. Three major short-lived intermediates, hydrated electron (e^-_{aq}), hydrogen atom (H·), and hydroxyl radical (OH·), are produced by irradiation of aqueous solutions, as shown by reaction (8). These primary radicals react

$$H_2O \xrightarrow{\text{irradiation}} e^-_{aq} + H\cdot + OH\cdot \qquad (8)$$

with solutes in water to form other short-lived and also very reactive intermediates, such as oxyanion, halide, pseudohalide, and cation-radical ions.

Hydrated electron. This radical is a very strong reducing agent ($E^\circ = -2.77$ V, where E° represents standard electrode potential), and has a broad absorption band with maximum at 715 nm, corresponding to a blue solution. In the presence of oxygen, e^-_{aq} is rapidly converted into other reactive species, superoxide ion O_2^-, and hydroperoxyl radical $HO_2\cdot$, as shown by reaction (9).

$$e^-_{aq} + O_2 \rightarrow O_2^- \longrightarrow H + HO_2\cdot \qquad (9)$$

Hydrated electrons also can be completely converted into hydroxyl radicals when a totally oxidizing system is obtained. This can be achieved when N_2O is present as a solute, as shown by reaction (10).

$$e^-_{aq} + N_2O \rightarrow N_2 + O^-(+H_2O) \rightarrow OH\cdot + OH^- \qquad (10)$$

If other solutes are present, such as peroxodisulfate ions ($S_2O_8^{2-}$), reaction (11) occurs. In this example, a new reactive sulfate radical anion is produced.

$$e^-_{aq} + S_2O_8^{2-} \longrightarrow SO_4^-\cdot + SO_4^{2-} \qquad (11)$$

A radical-cation intermediate can be generated by the very fast reaction of hydrated electrons with cobalt(III)-ammine complexes containing a mono- or dinitrobenzoato ligand, as shown by reaction (12).

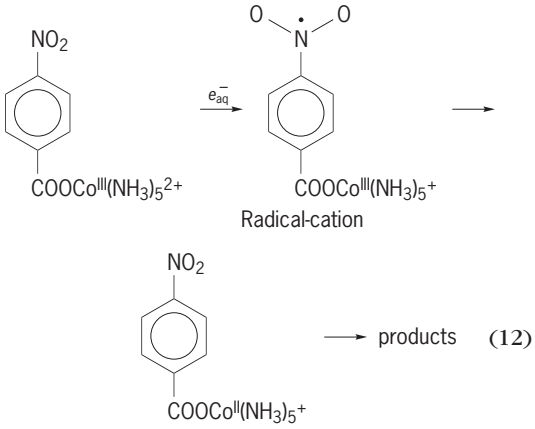

(12)

Hydrogen atom. This is the simplest chemical species, carrying a single electron, which decays very rapidly in aqueous solution. In alkaline solution the hydrogen atom is converted into the hydrated electron [reaction (13)], and in acidic solution the hydrated electron is reconverted to the hydrogen atom [reaction (14)].

$$H\cdot + OH^- \longrightarrow e^-_{aq} + H_2O \qquad (13)$$

$$e^-_{aq} + H^+ \longrightarrow H\cdot \qquad (14)$$

Although the hydrogen atom is a strong reducing agent, there are reactions where it acts as an oxidant: as an example, the formation of the reactive hydrido complex, $Fe^{3+}H^-$ (**Table 7**).

Hydroxyl radical. This radical is a strong oxidant ($E^\circ = 1.9$ V) which reacts rapidly with organic and inorganic compounds, often producing new radicals. A hydroxyl radical is frequently produced as a reactive intermediate in reactions in solutions and gas phase, for example, reactions (15), where hydroxyl

$$Co(CN)_5^{3-} + H_2O_2 \rightleftharpoons Co(CN)_5OH^{3-} + OH\cdot \qquad (15a)$$

$$Co(CN)_5^{3-} + OH\cdot \longrightarrow Co(CN)_5OH^{3-} \qquad (15b)$$

TABLE 7. Some reactions of radicals and radical ions with inorganic and organic solutes

Radical or radical ion		Solute		Products
$OH\cdot$	+	Fe^{2+}	→	$Fe^{3+} + OH^-$
O^-	+	$Fe^{2+}(+H_2O)$	→	$Fe^{3+} + 2OH^-$
$HO_2\cdot$	+	Fe^{2+}	→	$Fe^{3+}\ OH_2^-\ (+H^+) \rightarrow Fe^{3+} + H_2O_2$
O_2^-	+	Fe^{3+}	→	$Fe^{2+} + O_2$
$H\cdot$	+	Fe^{2+}	⇌	$Fe^{3+}\ H^-\ (+H^+) \rightarrow Fe^{3+} + H_2$
$H\cdot$	+	Fe^{3+}	→	$Fe^{2+} + H^+$
$Cl_2\cdot^-$	+	Fe^{2+}	→	$FeCl^{2+} + Cl^-$
$Cl_2\cdot^-$	+	Fe^{2+}	→	$Fe^{3+} + 2Cl^-$
e^-_{aq}	+	C_2H_5OH	→	$C_2H_5O^- + H\cdot$
$H\cdot$	+	C_2H_5OH	→	$C_2H_4OH + H_2$
$SO_4\cdot^-$	+	C_2H_5OH	→	$C_2H_4OH + HSO_4^-$

radicals, OH, are short-lived reactive intermediates generated in the oxidation-reduction step (15*a*) and consumed in the addition step (15*b*).

In the pulse radiolysis of aqueous solution, hydroxyl radical is completely converted into a reactive carbon dioxide radical anion, $CO_2^{\cdot-}$, in the presence of a solute, such as methanol or formate. In this case, a totally new reducing system is achieved, as is shown in reaction (16). An additional example of

$$OH\cdot + HCOO^- \longrightarrow CO_2^{\cdot-} + H_2O \qquad (16)$$

hydroxyl radical reactivity is shown in reaction (17),

$$OH\cdot + Cl^- \rightleftharpoons ClOH^{\cdot-} + (H^+) \rightleftharpoons H_2O + Cl\cdot + (Cl^-) \rightleftharpoons Cl_2^{\cdot-} \qquad (17)$$

where the chloride radical, $Cl_2^{\cdot-}$, is produced. *See* HYDROXYL.

Gas-phase systems. There are many systems, such as in the gas phase, where the radicals are also common intermediates. An example is in the decomposition of ozone, O_3, to oxygen, which involves oxygen atoms as intermediates [reactions (18)].

$$O_3(g) \rightleftharpoons O_2(g) + O(g) \qquad (18a)$$

$$O_3(g) + O(g) \rightarrow 2O_2(g) \qquad (18b)$$

See OZONE.

Radicals are common reactive intermediates in gas-phase synthesis. An example is in the formation of hydrogen bromide, HBr, by the reaction of hydrogen and bromine gases at high temperature [reactions (19)]. This reaction proceeds by a chain mech-

$$Br_2(g) \rightarrow 2Br(g) \qquad (19a)$$

$$Br(g) + H_2(g) \rightleftharpoons HBr(g) + H(g) \qquad (19b)$$

$$H(g) + Br_2(g) \rightarrow HBr(g) + Br(g) \qquad (19c)$$

anism with H and Br radicals as intermediates.

Metal intermediates: oxidation-reduction. Reactive intermediates such as metal atoms and metal ions in unusual valency states and with metal complexes as intermediates are mainly classified according to the types of reactions in which they may arise. A major class involves oxidation-reduction.

Metal ions in unusual valency states. Some metals can exhibit higher reduced or higher oxidized or intermediate states than normally shown. In such states, metal ions are usually very reactive, and they are generally produced by radiation methods. A simple example is provided by reaction (20), where transition metal

$$e^-_{aq} + M^{n+} \rightarrow M^{(n-1)+} \qquad (20)$$

(M) in reduced state can be generated by the hydrated electron, e^-_{aq}. A few examples illustrate the wide variety of reactions of metals in an unusual oxidation state, such as Cu(III), Co(I), Eu(II), Ni(I), and Tl(II). (**Table 8**). It is also believed that the metal ions are reactive intermediates in some oxidation-reduction reactions in solutions. For example, it is assumed that the thallium ion Tl^{2+} is an intermediate in the oxidation of the ferrous ion Fe^{2+} by Tl^{3+} [reactions (21)].

$$Tl^{III} + Fe^{2+}_{aq} \rightleftharpoons Tl^{2+}_{aq} + Fe^{III} \qquad (21a)$$

$$Tl^{2+}_{aq} + Fe^{2+}_{aq} \rightarrow Tl^{+}_{aq} + Fe^{III} \qquad (21b)$$

Metal atoms. Metals in the atomic state, also sometimes known as the high-surface-area state, are the most reactive form and hence are technologically very important in the fields of synthetic and catalytic chemistry. The reactive form of the metal can be obtained, for example, by electron beam or resistive heating vaporization. Atomic metals participate in a variety of chemical reactions (**Table 9**).

Binuclear complexes. The most common intermediates arising in the inner-sphere oxidation-reduction reactions between two metal complexes of transition elements are binuclear complexes. Depending on the state of valencies of the metals in the binuclear centers, these reactive intermediates may be subdivided into three types. The first type is generated when the oxidant and the reductant become mutually attached, usually through a ligand (L) of the oxidant. In this type, the metals in the binuclear site exhibit the initial oxidation states. This is apparent in reactions (22), where type I represents a binuclear

$$Co^{III}(NH_3)_5LH^+ + Fe^{II} \overset{fast}{\rightleftharpoons} \underset{\text{Type I}}{(NH_3)_5Co^{III} - L - Fe^{II}} \qquad (22a)$$

$$\text{Type I} \rightarrow Co^{II} + Fe^{III}L \qquad (22b)$$

intermediate, termed the precursor complex [reaction (22*a*)].

TABLE 8. Some reactions involving metals in unusual valency states

Reaction	Reaction type
$Cu(OH)_2^+ + Cu(OH)_2^+$	Adduct formation
$Tl^{2+} + Tl^{2+}$	Dismutation
$Co^+ + Cu^{2+}$	Electron-transfer
$Eu^{2+} + Co(NH_3)_5Cl^{2+}$	Inner-sphere electron transfer
$Ni(L)^+ + H_3O^+$	Proton transfer

TABLE 9. Some chemical reactions of metal atoms

Atoms		Reaction substrate		Products	Reaction type
2Cu	+	$2BCl_3$	⟶	$2CuCl + B_2Cl_4$	Abstraction
Pd	+	RC(O)Cl	⟶	RC(O) — Pd — Cl	Oxidative addition
K	+	O_2	⟶	$K^+ O_2^-$	Electron transfer
Fe	+		⟶	$Fe + H_2$	Simple orbital mixing

Reactions (23) give an example of the second type of binuclear complex. Usually termed the successor complex, type II could originate in reaction step (23*b*), where k_{is} represents rate constant for the inner-sphere electron transfer process. Both metals in this type exhibit their final oxidation states.

$$Ir^{IV}Br_6^{2-} + Co^{II}(CN)_5^{3-} \overset{\text{fast}}{\rightleftharpoons} Br_5Ir^{IV} — Br — Co^{II}(CN)_5^{5-} \quad (23a)$$

$$Br_5Ir^{IV} — Br — Co^{II}(CN)_5^{5-} \xrightarrow{k_{os}} Br_5Ir^{III} — Br — Co^{III}(CN)_5^{5-} \quad (23b)$$

Type II

$$\text{Type II} + H_2O \xrightarrow{58\%} Ir^{III}Br_6^{3-} + Co^{III}(CN)_5OH_2^{2-} \quad (23c)$$

$$\text{Type II} + H_2O \xrightarrow{42\%} Ir^{III}Br_5OH_2^{2-} + Co^{III}(CN)_5Br^{3-} \quad (23d)$$

The third type of binuclear intermediate contains metals in the mixed valence: the oxidant metal ion is in the initial oxidation state, and the reductant metal ion is already in the final oxidation state. The most widely known example is the radical-cation intermediate, type III, shown in reactions (24). (The dot inside the ring represents electron-rich pyrazine ligand.) *See* LIGAND.

Other types. Inevitably, many intermediates generated in the oxidation-reduction reactions of metal compounds cannot be neatly placed in the above categories. It has been reported that in some systems the "dead-end" intermediate can be formed. The best-known example is illustrated in reactions (25), where EDTA represents ethylenediaminetetraacetic acid and k_{os} represents the outer-sphere rate constant. In the first step [reaction (25*a*)], the dead-end intermediate appears, but the only productive oxidation-reduction process is the outer-sphere step [reaction (25*b*)].

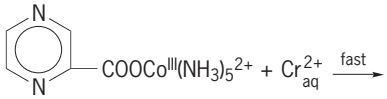

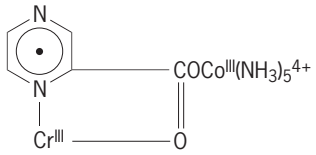

Type III (24*a*)

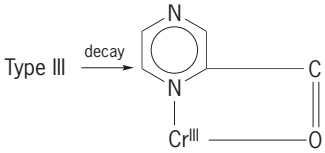

(24*b*)

$$Co^{II}EDTA^{2-} + Fe^{III}(CN)_6^{3-} \overset{\text{fast}}{\rightleftharpoons} EDTACo^{III} — NC — Fe^{II}(CN_5)^{5-} \quad (25a)$$

"Dead-end"

$$Co^{II}EDTA^{2-} + Fe^{III}(CN)_6^{3-} \xrightarrow{k_{os}} Co^{II}EDTA^{-} + Fe^{II}(CN)_6^{3-} \quad (25b)$$

Metal intermediates: substitution and homogeneous catalysis. Reactive intermediates may exist for a very short time in dissociative and associative substitution reactions of metal complexes, or during homogeneous catalysis by transition metal complexes. The most common type is five-coordinate intermediates.

Five-coordinate: dissociative reactions and catalysis. This type of coordinatively unsaturated reactive intermediate is generated by dissociative substitution reactions in octahedral Werner-type complexes (**Table 10**) or in six-coordinate organometallic complexes.

TABLE 10. Some dissociative substitution reactions of Werner-type complexes involving five-coordinate intermediates

Reaction type	Metal complex	Entering ligand	Intermediate
Anation	*trans*-$Coen_2SO_3OH_2^+$	N_3^-	$Coen_2SO_3^+$
Base hydrolysis	$Co(NH_3)_5Cl^+$	OH^-	$Co(NH_3)_4NH_2^{2+}$
Induced aquation	$Co(NH_3)_5Br^+$	$(Hg^{2+})H_2O$	$Co(NH_3)_5^{3+}$
Replacement	$Fe(CN)_5SO_3^{5-}$	CN^-	$Fe(CN)_5^{3-}$
Solvolysis	$Cr(H_2O)_5I^{2+}$	$H_2O — CH_3OH$	$Cr(H_2O)_5^{3+}$

The intermediate of reduced coordination number of Werner-type complexes lives long enough to show some discrimination between potential ligands in its vicinity. For example, in the ligand replacement [reactions (26)], the generation of a five-coordinate

$$MA_5L \underset{k_2}{\overset{k_1}{\rightleftharpoons}} MA_5 + L \qquad (26a)$$

$$MA_5 + Y \xrightarrow{k_3} MA_5Y \qquad (26b)$$

complex, MA_5, is shown in step (26*a*). The rate constants k_1, k_2, and k_3 are given for the individual reactions; k_1 represents the unimolecular dissociative rate constant, k_2 represents the rate constant for the addition of ligand (L) to the five-coordinate intermediate (MA_5), and k_3 represents the formation of a product with a different ligand (Y).

The competition ratio, k_2/k_3, is most often used to test the ability of the intermediate to discriminate between two ligands. It is helpful to recognize that the shape of a five-coordinate intermediate could be trigonal-bipyramidal or square-pyramidal, and it may be either rigid or fluxional. It is most probable that in the base hydrolysis of Co(III)-ammine and -amine complexes, the intermediate has trigonal bipyramidal geometry. On the other hand, in the base hydrolysis of Rh(III), its shape is probably square pyramid. There is apparently very little energy difference between the two geometries. A retention of stereoconfiguration is expected if reaction proceeds through square pryamid. A reaction proceeding through a trigonal bipyramid, on the other hand, would lead to a mixture of stereoisomers.

The nature of intermediates in organometallic reactions and catalysis is more complex. It is believed that dissociative reactions of 18-electron organotransition metal complexes proceed via 16-electron intermediates, and that 16-electron complexes proceed via 14-electron intermediates. Most of the common intermediates are three- or five-coordinated. The ligand-deficient intermediates are highly reactive, and the competition ratio shows that they are nondiscriminatory toward ligands of widely varying nucleophilicity. An example of a reaction which proceeds via five-coordinate 16-electron intermediate involves an acyl intermediate [reactions (27)].

$$RMn(CO)_5 \rightleftharpoons [RC(O)Mn(CO)_4] \qquad (27a)$$

$$[RC(O)Mn(CO)_4] + L \rightarrow RC(O)Mn(CO)_4L \qquad (27b)$$

This intermediate is generated by CO insertion [reaction (27*a*)] in alkyl transition metal carbonyl complexes.

Generally, the cyclic nature of catalysis requires that homogeneous catalytic reactions proceed by a series of ligand dissociation and ligand association steps in which intermediates with 16- and 18-electron configuration occur alternately.

Five-coordinate: associative reactions. It is generally accepted that substitution reactions of square-planar Pt(II), Pd(II), Au(II), Rh(I), and Ni(II) complexes occur by an associative mechanism with the possibility of the formation of a five-coordinate intermediate with increased coordination number. The trigonal-bipyramidal geometry of the intermediate has been assumed to be more favorable than square-pyramidal. The reaction of square-planar $Pt(NH_3)_3Cl^+$ complex with Cl^- ion to give trans-$Pt(NH_3)_2Cl_2$ is a typical example, which probably involves a five-coordinate trigonal bipyramidal intermediate, $Pt(NH_3)_3Cl_2$, as shown in reaction (28). The

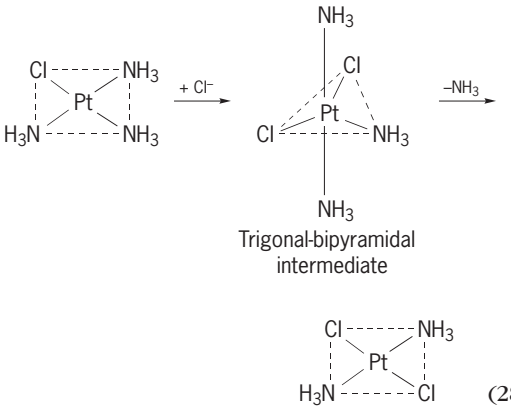

(28)

stability, structure, and reactivity of this type of intermediate is greatly influenced by solvents, ligands, and the nature of the metals involved in the reactions. In some systems, an intermediate exists for a sufficient time for it to undergo pseudorotation. *See* CHEMICAL DYNAMICS; COORDINATION CHEMISTRY; COORDINATION COMPLEXES; ELECTROPHILIC AND NUCLEOPHILIC REAGENTS; MOLECULAR ORBITAL THEORY; TRANSITION ELEMENTS.

Zdravko Bradic

Bibliography. J. D. Atwood, *Inorganic and Organometallic Reaction Mechanisms*, 1985; F. A. Cotton and G. Wilkinson, *Advanced Inorganic Chemistry*, 5th ed., 1988; A. B. Lever (ed.), *Excited States and Reactive Intermediates: Photochemistry, Photophysics, and Electrochemistry*, 1986; T. H. Lowry and K. S. Richardson, *Mechanism and Theory in Organic Chemistry*, 3d ed., 1987; J. March, *Advanced Organic Chemistry*, 5th ed., 2000; C. J. Moody and G. H. Whitham, *Reactive Intermediates*, 1992.

Reactive power

The concept of reactive power arises in electrical circuits operated from alternating-current sources. Electric fields (associated with capacitive effects) and magnetic fields (associated with inductive effects) are alternately charged and discharged according to the transfer of energy that takes place in each cycle. Reactive power is related to the peak of power absorbed by either capacitance or inductance in the circuit. The term "real or active power" is used to describe the average value of the power waveform. All power system equipment generates or absorbs reactive power, including synchronous generators, synchronous condensers, static compensators, capacitive and inductive compensators, overhead lines and underground cables, transformers, and consumer loads. Real power controls the

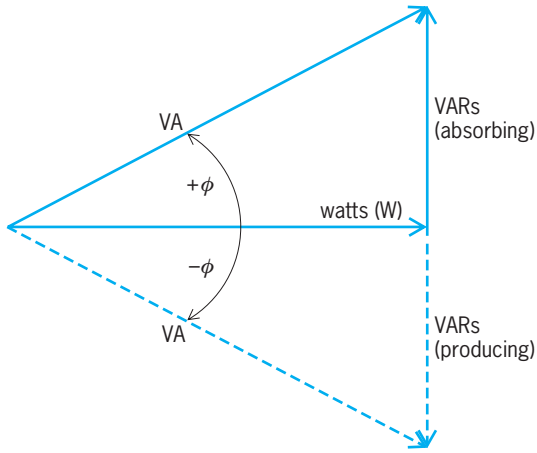

Fig. 1. The power triangle.

frequency of the system, while reactive power controls the voltage. Both are required for successful operation of the power system, and a balance must be respected between the real and reactive power at every point in the network. To violate the real power balance is to invite a condition known as transient instability, while violating the reactive power balance may lead to voltage instability or eventually collapse.

Power triangle. The real power consumed by a single-phase device will be denoted as P, the reactive power as Q, and the complex power as S, where $S = P + jQ$, and where j is the imaginary number $\sqrt{-1}$. The reactive power is measured by a unit called a VAR, for volt-ampere reactive, while the real power is measured in watts (W). As shown in **Fig. 1**, the vector (complex) sum of these two components defines the complex power, S, whose magnitude $|S|$, termed the apparent power, is measured in units of volt-amperes (VA). The graphical relation between S, P, and Q is known as the power triangle. *See* POWER; VOLT-AMPERE.

Denoting the voltage magnitude across the device as V, the current magnitude as I, and the phase angle between the voltage and the current as ϕ, we can express the real power as $P = VI\cos(\phi)$, the reactive power as $Q = VI\sin(\phi)$, and the apparent power as VI. As shown in Fig. 1, the term $\cos(\phi) = P/S$, defined as the power factor, is a dimensionless quantity measured in watts/VA denoting the fraction of the apparent power which can be usefully converted into other forms of energy. As seen in Fig. 1, a device can absorb or produce VARs. By convention, a capacitor produces reactive power, while an inductor absorbs reactive power. *See* CAPACITOR; INDUCTOR.

Power-factor correction. The idea of power-factor correction is to compensate for reactive power consumption by inductive loads by generating reactive power as close as possible to the load which requires it, rather than supplying it from a remote power station. Alternatively, if a load generates reactive power, power-factor correction would be provided by consuming VARs close to the load.

Figure 2*a* shows a single-line diagram of a load with admittance $Y = G + jB$ supplied by a constant voltage source, V. The supply current is, therefore, $I = YV = VG + jVB = I_R + jI_X$. The voltage and current, V and I, are represented by the phasor diagram of Fig. 2*b*, in which V is the reference phasor. The load current has a resistive component, I_R, in phase with V, and a reactive component, I_X, which is in quadrature with V. The current supplied by the power system is larger than necessary to supply the real power alone by the factor $I / I_R = 1/\cos(\phi)$. *See* ADMITTANCE.

Figure 3 shows the load after 100% compensation by placing a device at the load bus with a purely reactive admittance $-jB$. Then, the compensator current is $I_C = -V(jB)$, while the current supplied by the power system is $I_+ = V(G + jB) - V(jB) = VG = I_R$. This then means that the compensated supply current I_+ is in phase with V, making the overall power-factor unity. The supply current I_+ now has the smallest value capable of supplying full power at the voltage V, and all the reactive power required by the load is supplied locally by the compensator. Relieved of the reactive requirements of the load, the supply now has excess capacity which is available for supplying other loads.

The terms "leading" and "lagging" are often written in conjugate with the power factor. They are defined by the current through the load. If the current through a load leads the voltage across it, the load is said to have a leading power factor. If the current through a load lags the voltage across it, the load is said to have a lagging power factor. In other words, capacitive networks have a leading

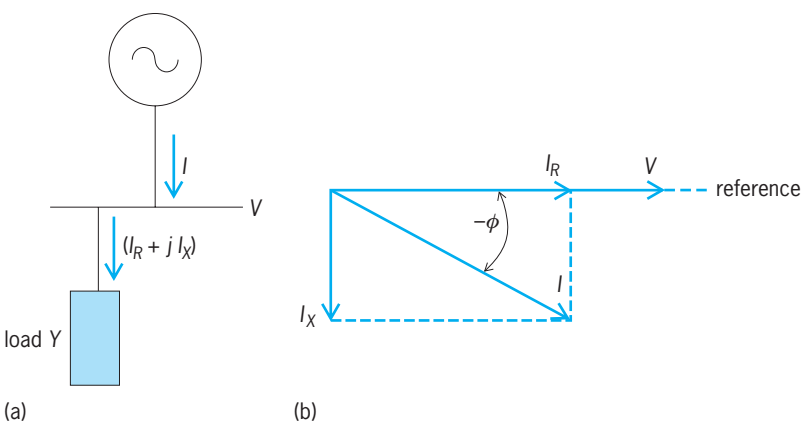

Fig. 2. Diagrams of uncompensated load. (*a*) Single-line diagram. (*b*) Phasor diagram.

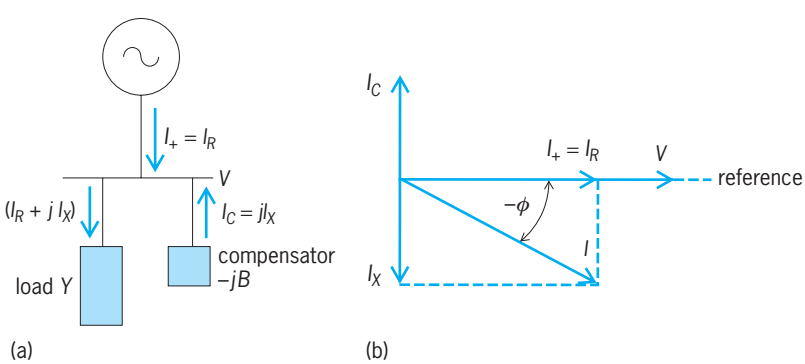

Fig. 3. Diagrams of compensated load. (*a*) Single-line diagram. (*b*) Phasor diagram.

power factor and inductive networks have a lagging power factor. *See* ALTERNATING-CURRENT CIRCUIT THEORY; ELECTRIC POWER TRANSMISSION.

D. McGillis; K. El-Arroudi; Francisco D. Galiana

Bibliography. A. E. Emanuel, On the definition of power factor and apparent power in unbalanced polyphase circuits with sinusoidal voltage and currents, *IEEE Trans. Power Delivery*, 8(3): 841–852, July 1993; A. E. Knowlton, Reactive powers in need of clarification, *AIEE Trans.*, 52:744–805, September 1933; R. M. Mathur, Static *Compensator for Reactive Power Control*, Cantext Publication, Canadian Electricity Association, Winnipeg, 1986; T. J. E. Miller, *Reactive Power Control in Electric Systems*, Wiley, 1982; Reactive Power Reserve Work Group Technical Studies Subcommittee, *Voltage Stability Criteria, Undervoltage Load Shedding Strategy and Reactive Power Reserve Monitoring Methodology*, Western Electricity Coordinating Council (WECC), Final Report, May 1998; C. W. Taylor, *Power System Voltage Stability*, McGraw-Hill, New York, 1994; U.S.-Canada Power System Outage Task Force, *Final Report on the August 14, 2003 Blackout in the United States and Canada: Cases and Recommendations*, April 2004; Western Electricity Coordinating Council (WECC), *WECC Policy Regarding Extreme Contingencies and Unplanned Events*, no. 12, May 2002; J. Willems, J. Ghijselen, and A. Emanuel, The apparent power concept and the IEEE standard 1459–2000, *IEEE Trans. Power Delivery*, 20(2):870–884, April 2005.

Reactor (electricity)

A device for introducing an inductive reactance into a circuit. Inductive reactance x is a function of the product of frequency f and inductance L; thus, $x = 2\pi fL$. For this reason, a reactor is also called an inductor. Since a voltage drop across a reactor increases with frequency of applied currents, a reactor is sometimes called a choke. All three terms describe a coil of insulated wire. *See* CHOKE (ELECTRICITY); INDUCTOR.

According to their construction, reactors can be divided into those that employ iron cores and those where no magnetic material is used within the windings. The first type consists of a coil encircling a circuit of iron which usually contains an air gap or a series of air gaps. The air gaps are used to attenuate the effects of saturation of the iron core. The second type, called an air-core reactor, is a simple circular coil, wound around a cylinder constructed of nonmagnetic material for greater mechanical strength. This strength is necessary for the coil to withstand the electromagnetic forces acting on each conductor. These forces become very large with heavy current flow, and their direction tends to compress the coil into less space: radial forces tend to elongate internal conductors in the coil and to compress the external ones while the axial forces press the end sections toward the center of the coil.

Both iron-core and air-core reactors may be of the air-cooled dry type or immersed in oil or a similar cooling fluid. Both types of reactors are normally wound with stranded wire in order to reduce losses due to eddy currents and skin effect. In addition, it is important to avoid formation of short-circuited metal loops when building supporting structures for air-core reactors since these reactors usually produce large magnetic fields external to the coil. If these fields penetrate through closed-loop metal structures, induced currents will flow, causing both losses and heating of the structures. Which of these two reactor types should be used depends on the particular application, which also provides a reactor designation. *See* EDDY CURRENT; SKIN EFFECT (ELECTRICITY).

Ballast reactor. A ballast reactor consists of a coil wound on an iron core and connected in series with a fluorescent or vapor lamp. The coil compensates for a negative-resistance characteristic of the lamp by providing an increased voltage drop as the current through the lamp is increased. Normally, two coils are magnetically coupled (wound on the same core) to ensure an equal current distribution between two fluorescent lamps operating in parallel. *See* FLUORESCENT LAMP; VAPOR LAMP.

Commutating reactor. The commutating reactor is found primarily in silicon controlled rectifier (SCR) converters, where it is connected in series with a commutation capacitor to form a highly efficient resonant circuit. This circuit is used to cause a current oscillation that turns off (commutates) the conducting SCR. The inductor is normally made of litz wire to reduce losses caused by skin effect and is usually of air-core type. *See* CONTROLLED RECTIFIER; RECTIFIER; SEMICONDUCTOR RECTIFIER.

Current-equalizing reactor. This type is used to achieve desired division of current between several circuits operating in parallel.

Figure 1 shows two interphase reactors connected between two parallel SCR converters. The reactor provides balanced system operation when both converters are conducting by acting as an inductive voltage divider. Since the net direct current through the reactor is zero, the reactor supports only the harmonic voltage generated by SCR switchings.

Figure 2 shows two magnetically coupled reactors used to force equal currents through two parallel conducting paths. The voltage induced across each

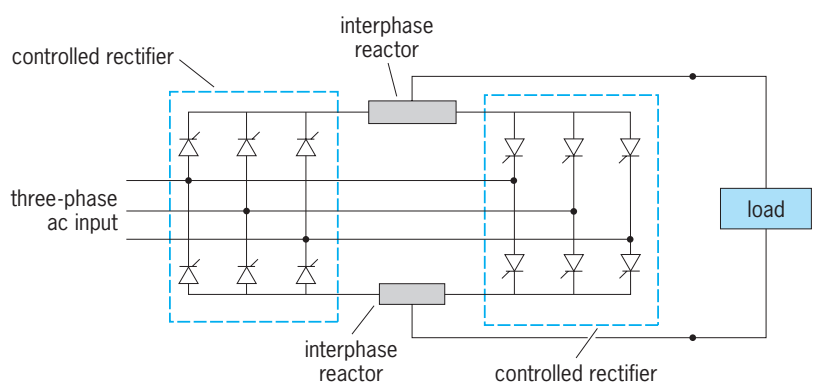

Fig. 1. Two interphase reactors that enable simultaneous (circulating current) operation of two controlled rectifiers.

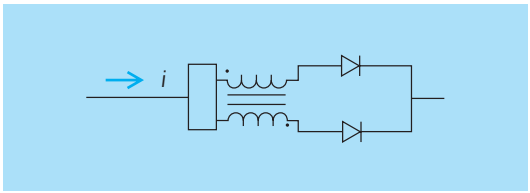

Fig. 2. Two magnetically coupled reactors used to force equal currents through two parallel diodes.

inductor is proportional to the difference between currents through the two diodes and with a polarity which tends to equalize the two currents. Since the reactors should be as linear as possible and have minimum hysteresis, magnetic materials are not used in the cores.

Current-limiting reactor. This type is used for protection against excessively large currents under short-circuit or transient conditions. By increasing the circuit's time constant (defined as L/R, where R is the total resistance), an inductor decreases the rate at which the current can change. In this way, a reactor provides more time for various protection devices (such as circuit breakers) to act before the current reaches a dangerous level. *See* ELECTRIC PROTECTIVE DEVICES.

Current-limiting reactors are always connected in series with a circuit which they are protecting, are of the air-core type, and can be divided into the following groups.

Bus reactor. This is an air-core inductor connected between two buses or two sections of the same bus in order to limit the effects of voltage transients on either bus. During normal operation both buses are at approximately the same voltage, and a free exchange of current takes place. As a fault occurs, the voltage on that bus abruptly changes and a heavy current starts to flow from one to the other bus. At that instant, a significant voltage drop is developed across the reactor, the magnitude of which depends on the rate of change of the current and on the reactor size (inductance). This voltage drop limits the current flow, minimizing the effect of the fault on the rest of the system and allowing a circuit breaker to disconnect the faulted section.

The di/dt reactor. This is a small inductor connected in series with a silicon controlled rectifier in order to limit the rate of rise of the SCR current. In small static converters, the di/dt reactor consists of an air-core inductor. In larger units, a saturable reactor is used, with a saturation point being just below the SCR rated current.

Feeder reactor. This is essentially the same type (but usually of higher power rating) used as a bus reactor. It is connected in series with a feeder circuit in order to limit and localize disturbances due to faults on the feeder. The principle of operation is the same as for bus reactors.

Generator reactor. This is the same as the feeder reactor but is connected between power plant generators and the rest of the network.

Starting reactor. This type is normally used to limit starting current of electric motors. A starting reactor usually consists of iron-core inductors connected in series with the machine stator winding. The starting reactor is short-circuited by a switch once the motor reaches a predetermined speed.

Filter reactor. This type is used to attenuate (filter) high-frequency currents. Since the voltage drop across an inductor is equal to a product between inductance, current, and frequency, the inductor opposes the flow of high-frequency currents, while a direct current is limited only by the resistance of the wire used to make the reactor. Filter reactors are connected in series and are usually wound on an iron core with an air gap to prevent saturation. Filter reactors are used extensively in both alternating-current and direct-current circuits, especially in conjunction with static power converters. *See* ELECTRIC FILTER.

Grounding reactor. This type is used in grounded power systems and connected between the system's neutral point and a ground. The main purpose of a grounding reactor is to limit disturbances due to ground faults and atmospheric discharges on high-voltage transmission lines. A special grounding reactor, tuned to the rest of the network, which helps to extinguish a line-to-ground fault current, is called a Petersen coil. All grounding reactors are made of heavy cable that is wound on an iron core. *See* GROUNDING.

Saturable reactor. This is an iron-core reactor, adjusted to saturate at a predetermined current (**Fig. 3**), and used primarily in alternating-current circuits. Since the slope of the voltage-current curve (Fig. 3) is proportional to the reactor inductance, saturation substantially decreases this inductance, meaning that any further increase in the reactor current will result in very little change in the voltage across the reactor. Saturable reactors are used in a wide variety of circuit, whenever a nonlinear inductance is needed. By placing an additional, direct-current-supplied, control winding on the reactor core, one obtains a magnetic amplifier (amplistat or magamp). The direct current is used to control the point at which the reactor saturates and thus the voltage drop across the reactor. Magnetic amplifiers are often used in conjunction with diode rectifiers to regulate dc power flow, but are now being replaced by semiconductor (thyristor or silicon controlled rectifier) converters. *See* SATURABLE REACTOR.

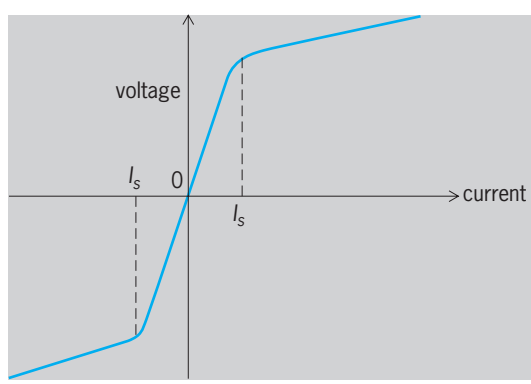

Fig. 3. Voltage-current characteristic of a saturable reactor; I_s is the current level at which the inductor saturates.

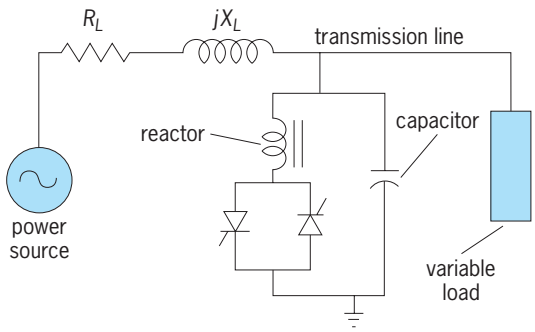

Fig. 4. Simplified single line diagram of a static reactive power system. Silicon controlled rectifiers are used to regulate the current through the reactor. R_L and X_L are the equivalent resistance and reactance of the transmission line.

Shunt reactor. A shunt reactor consists of an iron-core inductor and is used to compensate for capacitance in a transmission line. For this reason, it is also called a compensating reactor. Shunt reactors are normally connected at selected points between a transmission line and a ground. Examples of its application are the following:

1. Compensating reactors, connected at predetermined points along an underground transmission line to compensate the leading (capacitive) characteristic of an underground cable.

2. The Pupin coil, connected at regular intervals on long telephone lines to compensate line capacitance and to provide a constant line characteristic (impedance) at lower frequencies.

3. A parallel reactor in a static reactive power system (**Fig. 4**), used in conjunction with silicon controlled rectifiers to regulate voltages on an extra-high-voltage (EHV) overhead transmission line. The regulation is achieved by controlling the amount of current flow through the reactor and is used to prevent voltage variations caused by switching loads.

Tesla reactor. Also known as a tesla coil, this consists of two magnetically coupled coils wound on an air core. The primary winding is made of a few turns of heavy wire, while the secondary consists of fine wire with a very large number of turns. By circulating a high-frequency current in the primary, very high voltage is induced in the secondary. The tesla coil is used in radio and television receivers as well as for producing high discharge voltages. *See* INDUCTION COIL; TRANSFORMER.

Victor R. Stefanovic

Bibliography. B. J. Baliga, *Power Semiconductor Devices*, PWS, 1995; S. B. Dewan and A. Straughen, *Power Semiconductor Circuits*, Wiley, 1975; A. J. Pansini, *Power Transmission and Distribution*, CRC Press, 2d ed., 2005.

Reactor physics

The science of the interaction of the elementary particles and radiations characteristic of nuclear reactors with matter in bulk. These particles and radiations include neutrons, beta (β) particles, and gamma (λ) rays of energies between zero and about 10^7 eV. *See* BETA PARTICLES; GAMMA RAYS.

The study of the interaction of beta and gamma radiations with matter is, within the field of reactor physics, undertaken primarily to understand the absorption and penetration of energy through reactor structures and shields. For a discussion of problems of reactor shielding *see* RADIATION SHIELDING

With this exception, reactor physics is the study of those processes pertinent to the chain reaction involving neutron-induced nuclear fission with consequent neutron generation. Reactor physics is differentiated from nuclear physics, which is concerned primarily with nuclear structure. Reactor physics makes direct use of the phenomenology of nuclear reactions. Neutron physics is concerned primarily with interactions between neutrons and single nuclei or with the use of neutron beams as analytical devices, whereas reactor physics considers neutrons primarily as fission-producing agents. In the hierarchy of professional classification, neutron physics and reactor physics are ranked as subfields of the more general area of nuclear physics. *See* NEUTRON; NUCLEAR PHYSICS.

Reactor physics borrows most of its basic concepts from other fields. From nuclear physics comes the concept of the nuclear cross section for neutron interaction, defined as the effective target area of a nucleus for interaction with a neutron beam. The total interaction is the sum of interactions by a number of potential processes, and the probability of each of them multiplied by the total cross section is designated as a partial cross section. Thus, a given nucleus is characterized by cross sections for capture, fission, elastic and inelastic scattering, and such reactions as (n,p), (n,α) and $(n,2n)$. An outgrowth of this is the definition of macroscopic cross section, which is the product of cross section (termed microscopic, for specificity) with atomic density of the nuclear species involved. The symbols $N\sigma^j_i$ or Σ^j_i are used for macroscopic cross section, the subscript referring to the nuclear reaction involved, and the superscript to the isotope. The dimensions of Σ are cm^{-1}, and Σ_t (total cross section) is usually of the order of magnitude of unity. *See* NEUTRON SPECTROMETRY.

Cross sections vary with energy according to the laws of nuclear structure. In reactor physics this variation is accepted as input data to be assimilated into a description of neutron behavior. Common aspects of cross section dependence, such as variation of absorption cross section inversely as the square root of neutron energy, or the approximate regularity of resonance structure, form the basis of most simplified descriptions of reactor processes in terms of mathematical or logical models.

The concept of neutron flux is related to that of macroscopic cross section. This may be defined as the product of neutron density and neutron speed, or as the rate at which neutrons will traverse the outer surface of a sphere embedded in the medium, per unit of spherical cross-sectional area. The units of flux are neutrons/(cm^2)(s). The product of flux and

macroscopic cross section yields the reaction rate per unit volume and time. The use of these variables is conventional in reactor physics.

The chain reaction, a concept derived from chemical kinetics, is the basis of a physical description of the reactor process. The source of energy in a nuclear reactor is the fission of certain isotopes of heavy elements (thorium, uranium, and plutonium, in particular) when they absorb neutrons. Fission splits the elements into two highly radioactive fragments, which carry away most of the energy liberated by the fission process (about 160 MeV out of about 200 MeV per fission), the bulk of which is rapidly transformed into heat. Neutrons of high energy are also liberated, so that a chain of events alternating between neutron production in fission and neutron absorption causing fission may be initiated. Because more than one neutron is liberated per fission, this chain reaction may rapidly branch out to produce an increasing reaction rate, or divergent reaction; or if the arrangement of materials is such that only a small fraction of the neutrons will ultimately produce fission, the chain will be broken in a convergent reaction. *See* NUCLEAR FISSION.

Criticality. The critical condition is what occurs when the arrangement of materials in a reactor allows, on the average, exactly one neutron of those liberated in one nuclear fission to cause one additional nuclear fission. If a reactor is critical, it will have fissions occurring in it at a steady rate. This desirable condition is achieved by balancing the probability of occurrence of three competing events: fission, neutron capture which does not cause fission, and leakage of neutrons from the system. If ν is the average number of neutrons liberated per fission, then criticality is the condition under which the probability of a neutron causing fission is $1/\nu$. Generally, the degree of approach to criticality is evaluated by computing k_{eff}, which is the ratio of fissions in successive links of the chain, as a product of probabilities of successive processes.

Reactor constants are the parameters used in determining the probabilities of the various processes which together define k_{eff}. They comprise two sets, those used to characterize nuclear events, and those used to characterize leakage. In the former group are fast effect ϵ, resonance escape probability p, thermal utilization factor f, and neutrons emitted per fuel absorption η. In the latter group are neutron age or slowing-down area τ or L^2_s, migration area M^2, thermal diffusion area L^2, diffusion coefficient D, and buckling B^2.

Fast effect. The fast effect occurs in thermal reactors containing significant quantities of ^{238}U or ^{232}Th. These reactors comprise the bulk of plutonium production and civilian power reactors. The isotopes mentioned can undergo fission only when struck by very energetic neutrons; only a little more than one-half of the neutrons born in fission can cause fission in them. Moreover, these fast neutrons (energy greater than about 1.4 MeV) are subject to energy degradation by the competing reaction of inelastic scattering and by collision with moderator. In consequence, only a small number of fast fissions occur.

The fast effect is characterized by a ratio R of fast to nonfast fissions. From this ratio, a quantity ϵ, the fast effect, is derived, which determines the neutrons available to the chain reaction after the convergent chain of fast fissions has been completed, per neutron born of nonfast fission.

The common magnitude of ϵ varies between 1.02 and 1.05. R varies between 0 and 0.15 commonly. When $R < 0.01$, the fast effect is usually omitted from consideration.

Resonance escape probability. The resonance escape probability p is a significant parameter for the same group of reactors that have a significant fast effect and is defined as the probability that a neutron, in the course of being moderated, will escape capture by ^{238}U (or ^{232}Th) at any of the many energies at which the capture cross section is unusually high (resonance energies). The existence of this resonance absorption prevents most reactors homogeneously fueled with natural uranium from going critical, because resonance absorption, coupled with other losses, causes neutron depletion below the requirement for criticality. Therefore, low-enrichment reactors are heterogeneous, the fuel being deposed in lumps. The lumping of fuel, originally proposed by E. Fermi and E. P. Wigner, increases the probability of resonance escape. Spatial isolation of resonance absorber from moderator increases the number of neutrons which are slowed down without making any collisions with the absorber. Also, because absorption is very probable at resonance energies, the neutrons can travel only very short distances into the fuel lump before being absorbed, so that the interior of the lump is hardly exposed to resonant neutrons. Another effect of lumping is the removal of excess neutron scattering near the absorber so that there is a lesser probability of absorption following multiple collision.

With slight enrichment, homogeneous assemblies can be made critical. The homogeneous problem also provides the formalism by which p can be calculated. Use is made of the resonance integral RI, defined as the absorption probability per absorbing nucleon per neutron slowed down from infinite source energy in a moderator of unit slowing-down power. In a highly dilute system the resonance interval is given by Eq. (1), where σ_a is the neutron absorption cross section of the absorber as a function of energy, and E_0 is an energy taken as the lower limit of the resonance region. The slowing-down power of a moderator nucleus is $\xi\sigma_s$, where ξ is the mean increase in lethargy of the neutron per scattering, and σ_s is the moderator scattering cross section. Thus the resonance absorption probability of a system would appear to be given by Eq. (2), where N_r and N_m are,

$$RI = \int_{E_0}^{\infty} \sigma_a \frac{dE}{E} \tag{1}$$

$$1 - p = \frac{N_r(RI)}{N_m \xi \sigma_s} \tag{2}$$

respectively, atomic concentrations of absorber and moderator in the reactor volume. However, because the probabilities of escaping capture by successive nuclei must be multiplied, Eq. (3) is better.

$$p = \exp\left[\frac{-N_r(RI)}{N_m \xi \sigma_s}\right] \tag{3}$$

The value of E_0 used in the definition of RI follows one of two conventions: it is either taken as some defined thermal cut-off energy between about 0.3 and 1 eV, or as an energy just below the lowest resonance, about 6 eV for U and 20 eV for Th. In the latter case, RI is spoken of as $1/\nu$ corrected.

Thermal utilization factor. Thermal utilization factor f is the fraction of neutrons which, once thermalized, are absorbed in fuel. In a homogeneous array, f may be calculated from the atomic densities (that is, the number of atoms per cubic centimeter) and thermal-absorption cross sections of the various constituents of the reactor. The problem becomes more complex in a highly absorbing system and in the presence of nuclei (such as Pu or Cd) whose absorption cross sections do not vary with energy in the usual $1/\nu$ fashion. In this case, it becomes necessary to evaluate the neutron spectrum and average the absorption cross sections over this spectrum in order to obtain reaction ratios.

For low absorptions, the neutron spectrum is given by Eq. (4), the maxwellian expression, where

$$N(E)dE = \frac{2\pi}{(\pi kT)^{3/2}} E^{1/2} \times \exp\left(\frac{-E}{kT}\right) dE \tag{4}$$

k is the Boltzmann constant and T the absolute temperature of the medium. Deviations from this shape caused by absorption were first formulated by Wigner and J. E. Wilkins. This problem has become very significant with the advent of highly absorbing systems operating with considerable quantities of Pu.

In heterogeneous systems the problem is further complicated by the spatial nonuniformity of the neutron flux. It is therefore necessary to calculate reaction rates in various regions of the lattice by multiplying local absorption cross sections and atomic densities by local neutron fluxes; or, in effect the same thing, by introducing flux weights into the cross sections.

The calculation of neutron flux may be performed by diffusion theory or by more exact and elaborate methods for solving the neutron-transport problem. The problem may be further complicated by spatial effects on spectrum.

The term disadvantage factor, applied in simple systems originally to describe the ratio of mean moderator to mean fuel flux, has fallen into disfavor because of vague and local definitions. F, called the fuel disadvantage factor, is still in use to describe the ratio of surface to mean volume flux in a fuel lump with isoperimetric flux. *See* THERMAL NEUTRONS.

Fission neutrons per fuel absorption. This constant, η, is a characteristic of the fuel and of the neutron spectrum, but not of the spatial configuration of the system. Pure fissionable materials do not always undergo fission when they absorb neutrons; sometimes they lose their energy of excitation by emission of a gamma ray. The number of neutrons per fission ν varies only slightly with incident neutron energy (a few percent per megaelectronvolts), but η can fluctuate considerably even within a fraction of an electronvolt. Consequently, the specification of η is dependent upon a good evaluation of neutron spectrum.

For unirradiated, low-enrichment assemblies, the custom of defining fuel as all uranium, ^{235}U and ^{238}U, persists. Thus, f considers total uranium captures, and η is lowered from the value for pure ^{235}U by the fractional absorption rate of ^{235}U in uranium. This custom leads to excessive complexity as plutonium builds into the fuel and is therefore declining in use.

Infinite multiplication constant. The infinite multiplication constant, k_∞, is the ratio of neutrons in successive generations of the chain in the absence of leakage. In the formalism just described, the chain is taken from thermal neutron through fission and back to thermal neutron and from the definition of terms, $k_\infty = \eta\epsilon pf$. This formula is known as the four-factor equation.

For reactors other than weakly absorbing thermal systems, the simplified description breaks down. Thus in a fast reactor significant fission and capture occur at all neutron energies, and no moderator is present; in a very strongly absorbing thermal system, an appreciable fraction of neutrons react at energies between the thermal region and the lowest ^{238}U resonance. For such systems, the definition of the neutron chain is usually made in terms of a total time-dependent fission-rate expression, and the parametric representation of k_∞ becomes appreciably more complex. Generally, the spectrum is broken up into energy groups, and k_∞ is defined as the sum of the fission neutron production rate over all groups divided by the sum of absorption rates over all groups.

Neutron age. The neutron age τ is a reactor parameter defined in various ways, all related to the probability of leakage of a fast neutron from a reactor system. The basic definition is that 6τ measures the mean square distance of travel between injection of a neutron at one energy or energy spectrum and its absorption at some other energy in an infinite system. The term age is used because in weakly moderating systems the equation describing neutron slowing down in space and energy has the same form as the time-dependent heat-conduction equation, with τ substituted for time. The British and Canadian usage is L^2_s for slowing-down length (squared), which more accurately describes the physical parameter.

The common injection spectrum is a fission spectrum, and the common points of measurement or application are absorption at the indium resonance energy, 1.4 eV, or at some arbitrarily defined thermalization energy. The ages of these energies are denoted as τ_{In} or τ_{Th}. When a source other than the fission spectrum is considered, other subscripts are used.

A. M. Weinberg has pointed out that the shape of the spatial distribution in an infinite system for which τ is the second moment can be closely correlated with the fast leakage of a bare finite system. If the finite system has a source and sink distribution which is a solution of Laplace's equation, $\nabla^2\phi \times B^2\phi = 0$, then the Fourier transform of the τ distribution for given B will yield a quantity $P_\infty(B^2)$, which is almost exactly the nonleakage probability during slowing down. Two particularly significant cases are those for which the τ distribution is a gaussian or an exponential curve. In the former case, $P_\infty(B^2) = e^{-\tau B2}$; in the latter, $P_\infty(B^2) = 1/(1 + \tau B^2)$. The gaussian distribution is experimentally and theoretically verified for moderators as heavy as Be or heavier. The exponential distribution is crudely applicable to H_2O-moderated systems, and D_2O has a definitely mixed distribution.

In some cases, τ is used in a synthetic way to describe the leakage under some simple approximation to the slowing-down distribution. Thus, τ_{2G} is a number which is used in a two-energy group neutron model to give correct fast nonleakage probability as $P(B^2) = 1/(1 + \tau_{2G}B^2)$. When this model is not a good approximation, $6\tau_{2G}$ is not the second moment of the τ distribution.

For a heavy moderator, the age between two energies is given by the simple approximation of Eq. (5),

$$\tau_{E_1 \to E_2} = \int_{E_2}^{E_1} \frac{D(E)}{3(N\xi\sigma_s)} \frac{dE}{E} \tag{5}$$

where D is the diffusion coefficient, N is atomic density, σ_s is scattering cross section, and ξ is mean lethargy (logarithmic energy) gain/collision.

Thermal diffusion area. The thermal diffusion area L^2 is one-sixth the mean square distance of travel between thermalization and absorption. It is thus the analog of τ for thermal neutrons. Because thermal migration is usually well represented by a diffusion equation, which has an exponential absorption distribution, thermal nonleakage probability is given by $1/(1 + L^2B^2)$. L^2 is defined by $L^2 = D_{\text{Th}}/N\sigma_{a,\text{Th}}$, where D and σ_a are spectrum-averaged diffusion coefficient and absorption cross section, respectively.

Migration area. The migration area M^2 is one-sixth the mean square distance of travel from birth to death of a neutron. For large, small-leakage systems, $1/(1 + M^2B^2)$ is an excellent approximation to the total nonleakage probability. When only thermal absorption exists, $M^2 = \tau + L^2$.

Diffusion coefficient. The diffusion coefficient D is essentially a scaling factor applied to validate Fick's law, a relation between neutron flux and current (the latter being defined as a vector describing net rate of flow of neutron density). Fick's law is expressed by Eq. (6), where $\mathbf{j}$ is current, $\nabla\phi$ flux gradient, and D

$$\mathbf{j} = -D\nabla\phi \tag{6}$$

diffusion coefficient. For generalized systems, D is a tensor, but in regions where $|\nabla\phi|/\phi$ is small, D is approximated by a scalar of magnitude $1/(3\Sigma)$.

Buckling, B^2, is mathematically defined as $\nabla^2\phi/\phi$ in any region of a reactor where this quantity is constant over an appreciable volume. The name is derived from the relationship between force and deflection in a mechanical system with constrained boundaries. In a bare reactor, the buckling is constant except within one neutron mean free path ($\lambda \equiv 1/\Sigma$) of the boundary. For a slab of width t, $B^2 = (\pi/t)^2$; for an infinitely high cylinder of radius a, $B^2 = (2.404/a)^2$; for a sphere of radius a, $B^2 = (\pi/a)^2$; and for other geometries, it is again a geometrical constant. Because B^2 is a definite eigenvalue of Laplace's equation, it is the appropriate number to be used in the formulations of $P(B^2)$ previously described.

Effective multiplication k_{eff} is the ratio of neutron production in successive generations of the chain reaction. It is given by the product of k_∞ and $P(B^2)$, where $P(B^2)$ is itself the product of fast and thermal (or equivalent) nonleakage probabilities.

Reactivity. Reactivity is a measure of the deviation of a reactor from the critical state at any frozen instant of time. The term reactivity is qualitative, because three sets of units are in current use to describe it.

Percent k and millikay are absolute units describing the imbalance of the system from criticality per fission generation. Because $k_{\text{eff}} = 1$ describes a critical system, one says that it is 1% super- or subcritical, respectively, if each generation produces 1.01 or 0.99 times as many neutrons as the preceding one. Millikay are units of 0.1% k, and are given plus sign for supercriticality and minus sign for subcriticality. In both cases, k_{eff} is the base, as in Eqs. (7). These units

$$\begin{aligned} \%k &= \frac{100\,|k_{\text{eff}} - 1|}{k_{\text{eff}}} \\ \text{Millikay} &= \frac{1000(k_{\text{eff}} - 1)}{k_{\text{eff}}} \end{aligned} \tag{7}$$

are used primarily in design and analysis of control rods.

Dollars describe reactivity relative to the mean fraction of delayed neutrons per fission. Because the delayed neutrons are the primary agents for permitting control of the reaction, supercriticalities of less than 1 dollar are considered manageable in most cases. Thus, there is 1 dollar to "spend" in maneuvering power level. The dollar is subdivided into 100 cents. Because the delayed neutron fraction is a function of both neutron energy and fissionable material, the conversion rate between dollars and % k varies among reactors.

Inhours are reactivity units based on rate of change of power level in low-power reactors. If a low-power reactor is given enough reactivity so that its level would steadily increase by a factor of e per hour (which is also known as a 1-h period), that much reactivity is 1 inhour (from inverse hour). It is only for very small reactivities, however, that reactivity in inhours may be obtained from reciprocal periods.

Reactivity is measured in inhours primarily by operators of steady-state reactors, in which only small reactivities are normally encountered.

Reflectors. Reflectors are bodies of material placed beyond the chain-reacting zone of a reactor, whose

function is to return to the active zone (or core) neutrons which might otherwise leak. Reflector worth can be crudely measured in terms of the albedo, or probability that a neutron passing from core to reflector will return again to the core.

Good reflectors are materials with high scattering cross sections and low absorption cross sections. The first requirement ensures that neutrons will not easily diffuse through the reflector, and the second, that they will not easily be captured in diffusing back to the core.

Beryllium is the outstanding reflector material in terms of neutronic performance. Water, graphite, D_2O, iron, lead, and ^{238}U are also good reflectors. The use of Be, H_2O, C, and D_2O as reflectors permits conversion of neutrons leaking at high energy into thermal neutrons diffusing back to the core. Because the reverse flow of neutrons is always accompanied by a neutron-flux gradient, these reflectors show characteristic thermal flux peaks outside the core. They are, therefore, desirable materials for research reactors, in which these flux peaks are useful for experimental purposes.

The usual measurement of reflector worth is in terms of reflector savings, defined as the difference in the reflected dimension between the actual core and one which would be critical without reflector. Reflector savings are close to reflector dimensions for thin reflectors and approach an asymptotic value dependent upon core size and reflector constitution as thickness increases.

Reactor dynamics. Reactor dynamics is concerned with the temporal sequence of events when neutron flux, power, or reactivity varies. The inclusive term takes into account sequential events, not necessarily concerned with nuclear processes, which may affect these parameters. There are basically three ways in which a reactor may be affected so as to change reactivity. A control element, absorbing rod, or piece of fuel may be externally actuated to start up, shut down, or change reactivity or power level; depletion of fuel and poison, buildup of neutron-absorbing fission fragments, and production of new fissionable material from the fertile isotopes ^{232}Th, ^{234}U, ^{238}U, and ^{240}Pu make reactivity depend upon the irradiation history of the system; and changes in power level may produce temperature changes in the system, leading to thermal expansion, changes in neutron cross sections, and mechanical changes with consequent change of reactivity.

Reactor control physics. Reactor control physics is the study of the effect of control devices on reactivity and power level. As such, it includes a number of problems in reactor statics, because the primary question is to determine the absorption of the control elements in competition with the other neutronic processes. It is, however, a problem in dynamics, given the above information, to determine what motions of the control devices will lead to stable changes in reactor output.

Particular problems occurring in the statics of reactor control stem from the particular nature of control devices. Many control rods are so heavily absorbing for thermal neutrons that elegant refinements of neutron-transport theory are needed to estimate their absorption. Other types of control rods include isotopes with heavy resonance absorption, and the interaction of such absorbers with ^{238}U resonances must be examined. The motion of control rods changes the material balance of reactor regions, and with water-moderated reactors, peaks in the fission rate occur near empty rod channels. By virtue of high absorption of their constituent isotopes, some absorbing materials (for example, cadmium and boron) burn out in the reactor, and a rod made of these materials loses absorbing strength with time. As a final example, the motion of a control rod may change the shape of the power pattern in the reactor so as to bring secondary pseudostatic effects into play. *See* NUCLEAR REACTOR.

Reactivity changes. Long-term reactivity changes may represent a limiting factor in the burning of nuclear fuel without costly reprocessing and refabrication. As the chain reaction proceeds, the original fissionable material is depleted, and the system would become subcritical if some form of slow addition of reactivity were not available. This is the function of shim rods in a typical reactor. The reactor is originally loaded with enough fuel to be critical with the rods completely inserted. As the fuel burns out, the rods are withdrawn to compensate.

In order to decrease requirements on the shim system, many devices to overcome reactivity loss may be used. A burnable poison may be incorporated in the system. This is an absorbing isotope which will burn out at a rate comparable to or greater than the fuel. Burnable poisons are therefore limited to isotopes with very high effective neutron-absorption cross sections. Combinations of poisons, and the use of self-shielding of poisons can, in principle, make the close compensation of considerable reactivity possible without major control rod motions, but the technological problems in their use are formidable.

A more popular method for compensating reactivity losses is the incorporation of fertile isotopes into the fuel. This is desirable because the neutrons captured in the fertile material are not wasted, but used to manufacture new fissionable material; and also because (as with ^{238}U in ^{235}U reactors or ^{240}Pu in ^{239}Pu systems) the fertile material is normally found mixed with the fissionable, and isotopic separation may be circumvented or minimized. Depending on the conversion ratio (new fissionable atoms formed per old fissionable atom burned) and the fission parameters of the materials, a reactor so fueled loses reactivity relatively slowly, and in some cases, may show a temporary reactivity increase. The various isotopes produced by successive neutron capture in uranium and plutonium must also be considered, the higher isotopes becoming prominent at very long exposures.

A final consideration of long-term reactivity is the extra parasitic absorption of the fission products as formed. At long exposure, this absorption becomes significant because of the relatively high absorption

of many of the fission products. At shorter exposures, isotopes of very high absorption, mainly ^{149}Sm and ^{135}Xe, are more prominent. These materials have such high cross sections that they reach a steady-state concentration relatively quickly, burning out by neutron capture as rapidly as they are formed in fission.

^{135}Xe is particularly interesting because its cross section is abnormally large, its fission yield is high, it is preceded by an isotope of low cross section and approximately 7-h half-life (^{135}I), and it undergoes beta decay to the low absorption ^{135}Cs with a half-life of about 10 h. This combination of properties gives several interesting effects. Chief of these is that, at high flux, a reservoir of ^{135}I is formed which continues to decay to ^{135}Xe even after the reactor is shut down. Because ^{135}Xe is maintained at steady state during operation by a balance of buildup against burnout, the shutdown also removes the chief mode of ^{135}Xe destruction. Hence, the ^{135}Xe concentration increases immediately after shutdown. The reservoir of ^{135}I is so large that reactivity is rapidly lost as ^{135}Xe builds up; and in some cases, it may be impossible to restart the reactor after a short shutdown. The operator must then wait almost 2 days for the ^{135}Xe to disappear by radioactive decay. At very high fluxes, this effect becomes so severe that even a small temporary reduction in power may lead to ultimate subcriticality of the reactor.

Another effect caused by ^{135}Xe in high-flux reactors is that, if the reactor is large enough, the ^{135}Xe may force the power pattern into oscillations of 1- or 2-day periods. Although this is not a serious dynamic problem, it does emphasize the necessity of monitoring not only the total power, but also the power pattern, so that appropriate countermeasures may be taken. *See* NUCLEAR FUELS; NUCLEAR FUELS REPROCESSING.

Reactor kinetics. This is the study of the short-term aspects of reactor dynamics with respect to stability, safety against power excursion, and design of the control system. Control is possible because increases in reactor power often reduce reactivity to zero (the critical value) and also because there is a time lapse between successive fissions in a chain resulting from the finite velocity of the neutrons and the number of scattering and moderating events intervening, and because a fraction of the neutrons is delayed. *See* DELAYED NEUTRON.

Prompt-neutron lifetime. This is the mean time between successive fissions in a chain, and it is the basic quality which determines the time scale within which controlling effects must be operable if a reactivity excursion is touched off. In thermal reactors the controlling feature is the time between thermalization and capture, because fast neutrons spend less time between collisions, and the total time for moderation of a neutron is at most a few microseconds. Prompt-neutron lifetimes vary from 10 to a few hundred microseconds for light-water reactors, the shorter times being found in poorly reflected, highly absorbing systems, and the longer in well-reflected systems, in which time spent in the reflector is the dominating factor. Other thermal reactors have longer lifetimes, with some heavy-water reactors having lifetimes as long as a few milliseconds. Fast reactors have lifetimes of the order of 0.01–0.1 μs, the controlling factor being the amount of scattering material used as diluent.

Delayed neutrons. Delayed neutrons are important because a complete fission generation is not achieved until these neutrons have been emitted by their precursors. A slightly supercritical reactor must wait until the delayed neutrons appear, and this delay allows time for the system to be brought under control. The fraction of delayed neutrons β ranges from about $^1/_3$% for thermal fission of ^{239}Pu and ^{233}U, to about $^3/_4$% for thermal fission of ^{235}U, to several percent for some fast fission events. In these latter cases, however, the extra delayed neutrons have such short lifetimes that they are of only slight extra utility.

In any case, the influence of delayed neutrons is felt only to the extent that they are needed to maintain criticality. When the system is supercritical enough that the delayed neutrons are not needed to complete the critical chain, it is known as prompt critical. Prompt criticality represents in a qualitative sense the threshold between externally controllable and uncontrollable excursions, and it is for this reason that the dollar unit is popular in excursion analysis (a prompt critical system has a reactivity of 1 dollar).

Although it has now been established that a larger number of fission products emit delayed neutrons, the distribution in time after fission of the delayed neutron emission rate is accurately represented for all purposes by a sum of six negative exponentials. For many purposes, however, a three, two, or one group approximation is adequate.

Reactors moderated by D_2O and Be have additional delayed neutrons contributed by photoneutron reactions between the moderator and fission product gamma rays. Although not a large fraction, this effect makes such reactors unresponsive to small reactivity fluctuations and gives them unusual operational smoothness.

Reactor period. This is the asymptotic time required for a reactor at constant reactivity to increase its power by a factor e. When a critical reactor is given extra reactivity, its power will rise. At first, the power production rate has a complex shape on a time plot, but ultimately the power will rise exponentially. The period is the measure of this exponential rate.

The relation between the reactor period and the reactivity is known as the inhour equation. If l is the reactor lifetime in seconds, β_i the fraction of delayed neutrons in group i, λ_i the delay constant of group i delayed neutrons in s^{-1}, S reactor period in seconds and ρ reactivity in thousands of millikay, then Eq. (8) is the inhour equation. The equation

$$\rho = \frac{l/S + \sum_i [\beta_i/(1 + \lambda_i S)]}{1 - \sum_i [\beta_i/(1 + \lambda_i S)]} \tag{8}$$

has $I + 1$ solutions for S for a given ρ, I being the number of groups; and there is always a real value

of S with a higher value than the real part of any other solution. This highest S is the period. For very small values of ρ, that is, for very large periods, this value is approximatelythat given by notation (9). For

$$\frac{l + \sum_i (\beta_i/\lambda_i)}{\rho} \tag{9}$$

very large ρ, and therefore small S, the period is approximately $l/(\rho - \beta)$. This same result would be found if the delayed neutrons were thrown away completely.

Reactivity coefficients. There are several functions relating changes in reactivity to changes in the physical state of the reactor. The power coefficient is the change in reactivity per unit change in reactor power; the temperature coefficient relates reactivity to temperature change, and is often broken down into fuel, moderator, and coolant coefficients; for low-power graphite reactors there exists a barometric coefficient; one may define also coolant circulation rate coefficients and void coefficients.

Because the reactivity is commonly a complicated function of all the pertinent variables, the reactivity coefficient generally is the coefficient of the first term in a series expansion of the reactivity about the operating point. This in turn describes a linear theory of reactor dynamics. The theory may be extended to reactivity effects of arbitrary type by considering reactivity coefficients as functionals.

The basic problem of reactor dynamics is specifying the power coefficient of reactivity. The chain, power affects reactivity which affects power, is thus analyzable, using the power coefficient functional along with the reactor kinetic equations. The power coefficient is, however, predictable only in terms of changes in temperature and flow resulting from power changes in the system. The analysis of the power coefficient implies exhaustive knowledge of system behavior.

Reactivity coefficients may be prompt or delayed, and most delayed effects can be characterized as either of decay or transport type. An example of the decay type of coefficient is the contribution of coolant temperature change to power coefficient. Here, a power pulse gives a thermal effect on the coolant which is instantaneously observable, and which decreases exponentially with a time constant imposed by the heat-transfer equations. An example of a transport type of delay is the delay attributable to coolant circuit times. Here, a finite time lapse exists between the cause and the observable response. Effects due to fuel heating are examples of prompt effects.

Some types of power coefficient yield dangerous or unstable situations. Thus a power coefficient may contain a prompt positive (autocatalytic) term and a larger delayed negative term. Even though such a system may be stable against slow power-level increases, it will undergo a violent excursion whenever power is raised rapidly enough to outstrip delayed effects. Again, a system with prominent delayed effects of the transport type is always unstable beyond some critical power, even if the effect opposes the power shift; here, there is a possibility of phase instability.

The dynamic behavior of a reactor is usually analyzed by techniques common to all feedback systems. *See* CONTROL SYSTEMS; SERVOMECHANISM.

Bernard I. Spinrad

Bibliography. G. I. Bell and S. Glasstone, *Nuclear Reactor Theory*, 1970, reprint 1979; A. R. Foster and R. L. Wright, Jr., *Basic Nuclear Engineering*, 4th ed., 1983; S. Glasstone and A. Sesonske, *Nuclear Reactor Engineering*, 4th ed., 1993; D. L. Hetrick, *Dynamics of Nuclear Reactors*, 1993; J. R. Lamarsh, *Introduction to Nuclear Engineering*, 2d ed., 1983; R. Stammler and M. J. Abate, *Methods of Steady State Reactor Physics in Nuclear Design*, 1983; J. Weisman, *Elements of Nuclear Reactor Design*, 2d ed., 1983.

Reagent chemicals

High-purity chemicals used for analytical reactions, for the testing of new reactions where the effects of impurities are unknown, and in general for chemical work where impurities must either be absent or at known concentrations. If the concentration of impurity in any reagent is critical, an analysis should be made.

Methods of purification. Chemicals are purified by a variety of methods, the most common being recrystallization from solution. For many inorganic chemicals a saturated solution is prepared in water at the boiling point. After filtration to remove insoluble matter, the chemical crystallizes out of solution as it cools. The crystals are removed by filtration on a small scale or by centrifugation on a large scale, washed with water to remove impurity-containing solution on the surface, and dried. If the substance to be purified is not appreciably more soluble in hot solution than in cold solution, then one can use isothermal crystallization, the removal of solvent at constant temperature by reducing the pressure. The recrystallization process is not always successful as a purification method. In some cases precipitation of crystals from a saturated water solution by adding a second solvent, for example, ethyl alcohol, is the simplest method. One difficulty is that the impurity crystals may have the same structure as the desired ones, that is, the two crystals may be isomorphous. If water of hydration is present, storage in an atmosphere of known humidity may be necessary to obtain a definite hydrate. Occasionally it is easier to remove the impurity by dissolving it in a solvent in which the desired material is not very soluble.

If the desired chemical is volatile and the impurities are not volatile, sublimation is an effective method of purification. For example, iodine and arsenious oxide are easily purified by sublimation. For liquid chemicals, distillation is an effective procedure. Finally, the simplest procedure may be to synthesize the desired reagent from pure materials; for example, the addition of pure ammonium carbonate solution

to pure calcium chloride solution precipitates pure calcium carbonate.

Standards of purity. Commercial chemicals are available at several levels of purity. Chemicals labeled "technical" or "commercial" are usually quite impure. The grade "USP" indicates only that the chemical meets the requirements of the *United States Pharmacopeia*. The term "CP" means only that the chemical is purer than "technical." Chemicals designated "reagent grade" or "analyzed reagent" are specially purified materials which usually have been analyzed to establish the levels of impurities. The last two classes are the ones usually used in the laboratory. The American Chemical Society has established specifications and tests for purity for some chemicals. Materials which meet these specifications are labeled "Meets ACS Specifications."

A special group of extremely pure chemicals are called "primary standard" reagents. These reagents are usually readily available, easily purified, and unreactive with components of air such as water and carbon dioxide. The total sum of impurities should be less than 0.02%. These reagents are used to determine the concentrations of solutions used in volumetric analysis or for other purposes in which impurities must be quite low in concentration.

Care is necessary to prevent contamination by dust or by other chemicals. Transfers should be made directly from the container by pouring and not with spatulas or other tools. No material should be returned to the container. The maintenance of purity in an opened container is a problem.

Selective and specific reagents. Chemical reagents are often classified on the basis of utility. Many chemicals are general reagents, that is, they may react with many others. For example, any acid will be neutralized to some extent by a base. However, there are some reagents which react with only a limited number of other chemicals. These reagents are called selective reagents. The silver salts of chloride, bromide, iodide, thiocyanate, and a few other ions are insoluble in water; therefore silver nitrate is a selective precipitating reagent for these ions. A limited number of reagents are known which react appreciably with only one ion under specified conditions. These reagents are called specific reagents. The specific property is determined by the product formed. The ion dimensions, the charge densities, and the electron arrangements of the reagent and the ion which react must fall within certain limits or no reaction will occur. Variations in the acidity of the solution and the presence of other ions can change the condition of the ion so that no reaction can occur. Most inorganic reagents are at best selective. Most specific reagents are organic in nature.

These organic reagents have two types of reactive groups. One group forms electrovalent bonds by charge neutralization, and the second group forms covalent bonds by sharing electrons. The products are cyclic because both bonds are to the same ion, and they are called chelate compounds. These chelates are frequently very insoluble in water and intensely colored, making them very useful in analysis. *See* CHEMICAL SEPARATION TECHNIQUES; CRYSTALLIZATION; DISTILLATION.

Kenneth G. Stone
Charles Rulfs

Bibliography. American Chemical Society, *Reagent Chemicals: ACS Specifications*, 9th ed., 1999; L. Fieser and M. Fieser (eds.), *Fieser and Fieser's Reagents for Organic Synthesis*, vols. 9 and 10, 1981–1982; M. Fieser, *Reagents for Organic Synthesis*, 8 vols., 1980.

Real variable

A variable whose range is a subset of the real numbers. By extension the term is also used to refer to the theory of functions of one or more real variables. This theory has to do with properties of broad classes of functions, such as continuity, types of discontinuities, differentiability of functions, oscillation and variation of functions, and the various kinds of integrals. *See* INTEGRATION.

Real numbers. Real numbers are those commonly used in the geometric theory of measurement. The integers and fractions, also called rational numbers, are included among the real numbers. In practice an irrational number x is specified by telling which rational numbers are less than x and which are greater than x. Such a division of the rational numbers into two classes was used by J. W. R. Dedekind as the formal definition of a real number and is called a Dedekind cut.

The system of real numbers has the familiar algebraic properties and is also ordered. This order is related to the algebraic operations of addition and multiplication by the following properties:

If $a < b$, then $a + c < b + c$ for every number c.

If $a < b$, then $ac < bc$ for every number $c > 0$.

If $a > 0$ and $b > 0$, there exists an integer n such that $a < nb$.

Various subsets of the real numbers may be defined by means of inequalities. The set of numbers x satisfying the inequalities $a < x < b$ is called an open interval and denoted by (a,b). Here a and b are the ends of the interval but are not included in it. When the left end a is included, the notation is changed to $[a,b)$. The interval $[a,b]$, including both ends, is called a closed interval. A more complicated set, called the Cantor discontinuum, which is very frequently of use in the theory of functions, may be constructed as follows. From the interval [0,1] remove the open intervals (1/3, 2/3), (1/9, 2/9), (2/3 + 1/9, 2/3 + 2/9), and in fact all intervals $(c - 1/3^n, c)$, where $c = d_1/3 + d_2/3^2 + \cdots + d_{n-1}/3^{n-1} + 2/3^n$ and each $d_i = 0$ or 2. It may be proved that the remaining set contains just as many numbers as the original interval [0,1]. Another set to which reference will be made consists of all the rational numbers x satisfying the inequality $s^2 < 2$. *See* INFINITY.

An upper bound for a set S of numbers is a number U such that $x \leqq U$ for every x in S, and a lower bound is a number L such that $L \leqq x$ for every x in S. A set need not have an upper bound, for example, the set

of all rational numbers. A least upper bound for S is an upper bound U such that V is not an upper bound when $V < U$. There cannot be more than one least upper bound for a set of real numbers. A distinguishing property of the real number system states that every set having an upper bound also has a least upper bound in the system. This property fails for the system consisting only of the rational numbers, as is shown by the fact that the least upper bound of the set consisting of the rational numbers x such that $x^2 < 2$ is the irrational number $\sqrt{2}$.

Functions. Consider a function $f(x)$ defined on an interval $a \leqq x \leqq b$. By the oscillation of f on a subinterval $[c,d]$ of $[a,b]$ is meant the number $\omega(c,d) =$ least upper bound of $|f(x_1) - f(x_2)|$ for x_1 and x_2 in $[c,d]$. The function f is continuous at a point x of $[a,b]$ in case $\omega(c,d)$ tends to zero when c and d approach x, and f is continuous on $[a,b]$ when it is continuous at every point of $[a,b]$. A function f continuous on a closed interval $[a,b]$ attains a greatest and a least value on $[a,b]$. Moreover, it takes every value between $f(a)$ and $f(b)$. If f has a derivative f' on the interval $[a,b]$, then f' also takes every value between each pair of its values, although f' may fail to be continuous. However, a continuous function f may fail to have a derivative at any point whatever. This suggests a useful classification of functions according to the number of continuous derivatives they possess.

The variation of a function f on an interval $[a,b]$ is defined to be the least upper bound $V(a,b)$ of the sums shown in expression (1) for

$$\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \tag{1}$$

all finite sets of points $x_0, \ldots, x_n$ satisfying $x_0 = a$, $x_{i-1} < x_i$, $x_n = b$. One can write $V(a,b) = +\infty$ when no upper bound exists, and in all other cases one says that f has bounded variation. A continuous function f may fail to have bounded variation—for example, relation (2) on the interval [0,1]. On

$$\begin{aligned} f(x) &= x \sin\left(\frac{1}{x}\right) \qquad 0 < x \leq 1 \\ f(0) &= 0 \end{aligned} \tag{2}$$

the other hand, a function of bounded variation may be discontinuous; however, it always has right-hand and left-hand limits at each point, so all its discontinuities are jump discontinuities. This follows readily from the fact that a function of bounded variation always is equal to the difference of two nondecreasing functions. A very remarkable theorem states that every function f of bounded variation has a derivative $f'(x)$, except at the points of a set which can be enclosed in a sequence of intervals the sum of whose lengths is arbitrarily small.

If S is a set composed of nonoverlapping intervals $[c_k,d_k]$, the variation $V(S)$ of f over S may be defined as the sum of the variations $V(c_k,d_k)$. If $V(S)$ tends to zero with the sum of the lengths of the intervals, the function f is said to be absolutely continuous on $[a,b]$. An absolutely continuous function f admits a generalization of the fundamental theorem of integral calculus, as shown in Eq. (3), where the integral is a Lebesgue integral.

$$f(b) - f(a) = \int_a^b f'(x)\, dx \tag{3}$$

Continuous functions may be characterized in a way different from the definition given above. Thus a function f defined on a finite closed interval $[a,b]$ is continuous on $[a,b]$ if and only if for every positive number ϵ there exists a polynomial function $p(x)$ which approximates $f(x)$ with an error less than ϵ on the entire interval $[a,b]$, that is, $|f(x) - p(x)| < \epsilon$ for $a \leqq x \leqq b$. A variation of this condition (which may seem minor) yields a much more restricted class of functions than the continuous functions—that is, if one requires the approximating polynomials to be of the form shown in Eq. (4), where the coefficients c_k

$$p(x) = \sum_{k=0}^{n} c_k (x - x_0)^k \tag{4}$$

belong to a fixed infinite sequence $c_0, c_1, c_2, \ldots$, and x_0 is a point of $[a,b]$, the function f will have continuous derivatives of all orders and will have many other properties which are best understood when the variable x is allowed to be complex. On the other hand, if it is required only that there exist a sequence of polynomials p_n such that Eq. (5) holds for each

$$\lim_{n \to \infty} p_n(x) = f(x) \tag{5}$$

x on $[a,b]$, then $f(x)$ may be discontinuous. Such a function f is said to belong to Baire's class 1. A sequence f_n of functions in Baire's class 1 may converge to a discontinuous function not in the class 1. This is the beginning of an infinite collection of classes of discontinuous functions. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES. Lawrence Murray Graves

Bibliography. S. K. Berberian, *A First Course in Real Analysis*, 1994; S. A. Lang, *Real Analysis*, 3d ed., 1993; M. H. Protter and C. B. Morrey, *A First Course in Real Analysis*, 2d ed., 1993; H. L. Royden, *Real Analysis*, 3d ed., 1988; W. Rudin, *Real and Complex Analysis*, 3d ed., 1987; A. Torchinsky, *Real Variables*, 1988.

Real-time systems

Computer systems in which the computer is required to perform its tasks within the time restraints of some process or simultaneously with the system it is assisting. Usually the computer must operate faster than the system assisted in order to be ready to intervene appropriately.

Types of systems. Real-time computer systems and applications span a number of different types. Reactive real-time systems have an ongoing interaction with their external environment. Embedded real-time systems control specialized tasks in the system in which they are incorporated. Synchronous real-time systems control events at predictable time

intervals, whereas asynchronous real-time systems react to their environment to control or respond to unpredictable events requiring something from the real-time system.

Examples. At the race track in the last few minutes before a race, most bets are placed; the asynchronous real-time computer system must record the bets and update the betting odds for up to tens of thousands of bets per minute. In storm weather tracking, the real-time system must obtain data from up to hundreds of thousands of sensors per local area and process them to arrive at multiple weather forecasts per minute. Factory just-in-time real-time scheduling of machines and transportation occurs when trucks arrive with materials and subassemblies which need to be assembled into final products, packaged, palletized and loaded onto trucks, without intermediate warehousing, ready for delivery to end customers. Personal computer users require that their keystrokes and mouse actions are handled immediately (in the real time of human thinking and action). Weapons fire control requires microsecond real-time control. At the supermarket, real-time inventory restocking is synchronized by taking into account sales (as registered from bar code readers at the check-out counter), just-in-time ordering and delivery (time-lag scheduling), and shelf stocking timing. New jet engine design testing requires testing a strapped-down prototype engine in the factory by running it to its extreme limits; the real-time system must process thousands of sensor data inputs per microsecond to control the experiment and to shut down the engine just before it is about to fly to pieces. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING; MICROCOMPUTER; STORM DETECTION.

Real-time control and real-time process control. In these applications the computer is required to process systems data (inputs) from sensors for the purpose of monitoring and computing system control parameters (outputs) required for the correct operation of a system or process. The type of monitoring and control functions provided by the computer for subsystem units ranges over a wide variety of tasks, such as turn-on and turn-off signals to switches; feedback signals to controllers (such as motors, servos, and potentiometers) to provide adjustments or corrections; steering signals; alarms; monitoring, evaluation, supervision, and management calculations; error detection, and out-of-tolerance and critical parameter detection operations; and processing of displays and outputs. *See* CONTROL SYSTEMS; PROCESS CONTROL.

Real-time assistance. Here the computer is required to do its work fast enough to keep up with a person interacting with it (usually at a computer terminal device of some sort, for example, a screen and keyboard). These are people-amplifier-type real-time computer systems. The computer supports the person or persons interacting with it and provides access, retrieval, and storage functions, usually through some sort of database management system, as well as data processing and computational power. System access allows the individual to intervene (control, adjust, supply parameters, direct, and so forth) in the system's operation. The real-time computer also often provides monitoring or display information, or both. *See* DATABASE MANAGEMENT SYSTEM; HUMAN-COMPUTER INTERACTION; MULTIACCESS COMPUTER.

Real-time robotics. In this case the computer is a part of a robotic or self-contained machine. Often the computer is embedded in the machine, which then becomes a smart machine. If the smart machine also has access to, or has embedded within it, artificial intelligence functions (for example, a knowledge base and knowledge processing in an expert system fashion), it becomes an intelligent machine. *See* ARTIFICIAL INTELLIGENCE; EMBEDDED SYSTEMS; EXPERT SYSTEMS; ROBOTICS.

Evolution. Real-time computer systems have been evolving constantly since the interrupt function, which allowed the computer to be synchronized with the external world, was invented. There are five primary paths along which real-time systems continue to advance:

First-generation real-time control systems comprise process control (such as an oil refinery); guidance and control (such as antiballistic missile and intercontinental ballistic missile systems); numerical control (such as factory machine operations); dedicated (mini) computer systems; and store-and-forward message switching. *See* DATA COMMUNICATIONS; GUIDANCE SYSTEMS.

Second-generation real-time computer systems comprise time-shared and multiprocessor computing; interactive computing; smart and intelligent terminals; and computer networks (distributed computers, distributed smart machines, and distributed intelligent machines). *See* MULTIPROCESSING.

Third-generation real-time assistance systems comprise operating systems; CAD (computer-aided design), CAM (computer-aided manufacturing), CAI (computer-assisted instruction), MIS (management information systems), and DSS (decision support systems); personal computers; word processing and work stations; and artificial intelligence expert systems. *See* DECISION SUPPORT SYSTEM; INFORMATION SYSTEMS ENGINEERING; OPERATING SYSTEM; WORD PROCESSING.

Fourth-generation real-time machines comprise smart machines with embedded computers; intelligent machines; and robots (dumb, smart, and intelligent).

Fifth-generation real-time integrated systems comprise the factory-of-the-future (totally automated factories); just-in-time (JIT) systems (in the factory and in distribution); computer utilities and knowledge utilities; and knowledge inference processing systems. *See* COMPUTER-INTEGRATED MANUFACTURING.

Artificial intelligence and expert systems. Artificial intelligence (AI) is rapidly advancing beyond the research stage into practical use by scientists, management, and many other areas of business and society. One branch is directed toward the development of real-time expert systems. Thus, widespread use of artificial intelligence, knowledge bases, expert

systems, and people-amplifiers is expected to develop, and decision making will be amplified with real-time and intelligent computer systems. *See* ARTIFICIAL INTELLIGENCE; EXPERT SYSTEMS.

Real-time simulation. The real-time computer can serve as a tool allowing simulation of models of the real world. By coupling this tool with artificial-intelligence knowledge-base expertise, scientific experimentation becomes possible without cumbersome laboratory equipment and procedures. Scientists can play serious real-time experimental and mathematical games with the object of their research without first needing to learn sophisticated laboratory techniques or mathematics. The detailed mathematical and discipline-oriented experimental skills and procedures are embedded within the computerized model. Such robot simulators allow the scientist to concentrate upon the investigation, rather than the scientist becoming buried within the relevant mathematics and discipline crafts. However, the scientist must first learn his or her field plus the computer simulation/modeling language. Thus, robot simulator assistants increasingly do for the scientist what the calculator does for the average person: they remove the need to perform bulky, precise, and rote skill functions, allowing the researcher to get more quickly and easily to the core of scientific investigation.

With robot simulators the researcher can ask questions of a computer modeling system, and have simulated experiments performed that are otherwise nearly impossible or too time-consuming and costly. For example, the researcher can ask "What happens if . . . ?" After performing the simulated experiment, the computer gives an answer, while the experimenter views the progress of the computerized experiment and intervenes when desired. Then the experimenter can ask "What happens if something else is done instead?" to arrive at a different comparable answer from the robot simulator. Through such real-time interactive simulations the researcher becomes directly involved in the experiment as a surrogate participant.

Future robot simulators are expected to take the form of advanced hand-held calculators with voice dialog and artificial intelligence capabilities, and to contain specialized expert knowledge making them capable of general decision making. Such smart robot simulators will be used as people-amplifiers or electronic assistants by managers, programmers, doctors, politicians, voters, and others. *See* COMPUTER; DIGITAL COMPUTER; SIMULATION. Earl C. Joseph

Bibliography. S. T. Allworth and R. N. Zobel, *Introduction to Real-Time Software Design*, 2d ed., 1990; J.-P. Calvez, *Embedded Real-Time Systems: A Specification and Design Methodology*, 1993; S. Goldsmith, *A Practical Guide to Real-Time Systems Development*, 1993; P. A. Laplante, *Real-Time Systems Design and Analysis: An Engineer's Handbook*, 2d ed., 1997; M. Schiebe (ed.) *Real-Time Systems Engineering and Applications*, 1992; C. Vickery, *Real-Time Systems Programming for PCs*, 1993.

Rearrangement reaction

A reaction in which an atom or bond moves or migrates, having been initially located at one site in a reactant molecule and ultimately located at a different site in a product molecule. A rearrangement reaction may involve several steps, but the key feature defining it as a rearrangement is that a bond shifts from one site of attachment to another. The simplest examples of rearrangement reactions are intramolecular, that is, reactions in which the product is simply a structural isomer of the reactant [reaction (1)].

(1)

See MOLECULAR ISOMERISM.

More complex rearrangement reactions occur when the rearrangement is accompanied by another reaction, for example, a substitution reaction (2).

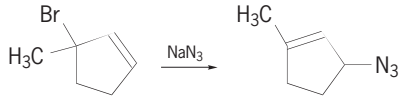

(2)

Classification. Rearrangement reactions are classified and named on the basis of the group that migrates and the initial and final location of the migrating bond. The initial bond location is designated as position 1, and the final location as position i, where the number of atoms is simply counted along the connection from 1 to i. Such a migration is called a [1,i] rearrangement or [1,i] shift. If the migrating group also reattaches itself at a different site from the one to which it had originally been attached, then both shifts are indicated, as in [i,j] shift. Reaction (1) is an example of a [3,3] rearrangement, because the initial carbon-oxygen (C-O) bond that breaks designates the 1 position for each component, and both components then rearrange and form a new C-C bond by reattaching at position 3 for each component. Reaction (2) is an example of a [1,3] rearrangement with substitution, commonly called an S_N2' reaction.

In addition, classification can be based on how many electrons move with the migrating group. Of the two electrons in the initial bond that breaks, the migrating group may bring with it both electrons (nucleophilic or anionotropic), one electron (radical), or no electrons (electrophilic or cationotropic). If the rearrangement is a concerted reaction in which there is a cyclic delocalized transition state that results in shifts of pi bonds as well as sigma bonds, the reaction is called a sigmatropic rearrangement [for example, reaction (1)]. *See* DELOCALIZATION; PERICYCLIC REACTION.

Carbocation rearrangements. The great majority of rearrangement reactions occur when a molecule develops a severely electron-deficient site, such as a carbocation. Shift of a nearby atom or group, with its pair of electrons, can serve to satisfy the electron deficiency at the original site, although it typically leaves behind another site with electron deficiency. As long as the final site can bear the electron deficiency better than the original site, the

rearrangement will be favorable. The stability of carbocations relies on electron donation from attached substituents and nearby groups. For simple alkyl groups and other electron-donating substituents, carbocation stability increases with greater numbers of substituents. Thus a simple hydrogen shift ([1,2]-H) can convert a less stable carbocation into a more stable one [reaction (3)].

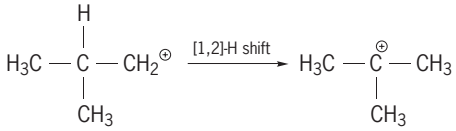

(3)

In such a reaction, the hydrogen migrates with a pair of electrons, so the reaction is sometimes called a hydride shift ($H:^-$ is the hydride anion). Hydrogen is the atom most commonly observed to rearrange. If an adjacent hydrogen is not available for a [1,2] shift, alkyl groups [for example, the methyl group (CH_3)] may also migrate [reaction (4)].

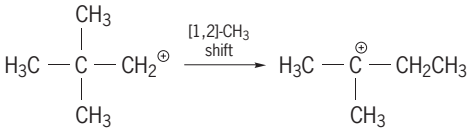

(4)

Rearrangements in carbocations can create subtle changes or major changes in the carbon skeleton of a molecule. The phenomenon of carbon skeletal rearrangements was first recognized by pioneers in the area of terpene chemistry, for example, in the conversion of isoborneol (1) to camphene (2) [reaction (5)].

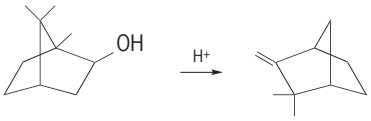

(5)

(1) (2)

See TERPENE.

The acid-catalyzed dehydration of alcohols proceeds through carbocation intermediates, and carbon skeletal rearrangements occurring during the dehydration are known as Wagner-Meerwein rearrangements. The specific rearrangement in the conversion of isoborneol to camphene is simply a [1,2]-carbon shift, well camouflaged. Without the methyl substituents as markers, detection of the rearrangement can be even more subtle, as in the conversion of exo-2-norbornyl derivatives, which are chiral, into their enantiomeric form. Enantiomers are mirror-image isomers, and the overall reaction results in racemization and loss of optical activity. *See* OPTICAL ACTIVITY; RACEMIZATION; REACTIVE INTERMEDIATES.

Migratory aptitude. The pinacol rearrangement [reaction (6) where Ph represents a phenyl group], represents a case in which there are two different groups that may migrate; the major product indicates that the phenyl group (C_6H_5) prefers to shift rather than the methyl group. The ability of a group to shift with a pair of electrons is called migratory aptitude, and it is determined by the ability of the migrating group to

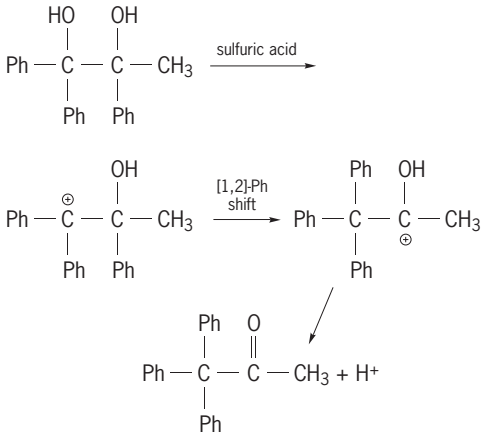

(6)

stabilize the electron deficiency. In general, aromatic groups have a greater migratory aptitude than alkyl groups.

Other rearrangements to electron-deficient carbon. The benzilic acid rearrangement [reaction (7)] represents the migration of a phenyl group to a carbonyl carbon (C═O).

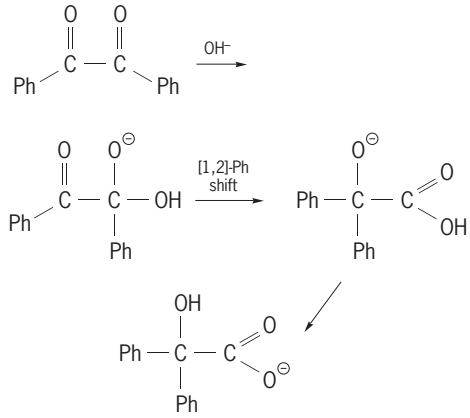

(7)

Rearrangements to electron-deficient nitrogen or oxygen. A number of related rearrangements occur in which an electron deficiency develops at a heteroatom (any atom other than hydrogen or carbon) and a carbon group shifts in response. The Beckmann rearrangement [reaction (8)] converts an

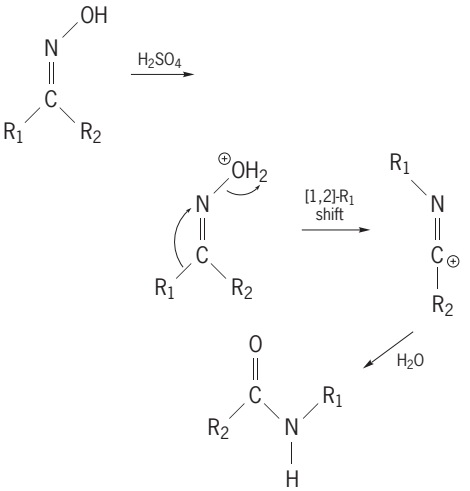

(8)

oxime (which is readily prepared from a ketone) to an amide. Of the two possible groups that may migrate to nitrogen (N), only the group anti (opposite) to the group that leaves from nitrogen migrates. This is taken as evidence that the migration is concerted (simultaneous) with the departure of the leaving group.

The Beckmann rearrangement, used in the synthesis of caprolactam (the monomer in the manufacture of Nylon 6), involves a ring expansion reaction. Rearrangement to electron-deficient oxygen occurs in the Baeyer-Villiger rearrangement, which converts a ketone to an ester. A related reaction involving peroxides is an important industrial reaction for the production of phenol and acetone.

Free-radical rearrangements. Free radicals undergo rearrangements much less readily than carbocations, an effect attributed to the molecular orbital requirements for migration. The transition state for a typical [1,2] migration involves bonding of three atoms simultaneously. One delocalized molecular orbital can accommodate two electrons, but any additional electrons, as in a radical migration, would have to be accommodated in higher-energy orbitals. Radical rearrangements are observed when the odd electron can be stabilized, for example, in migration of aryl groups where the odd electron can be delocalized into the pi system [reaction (9)].

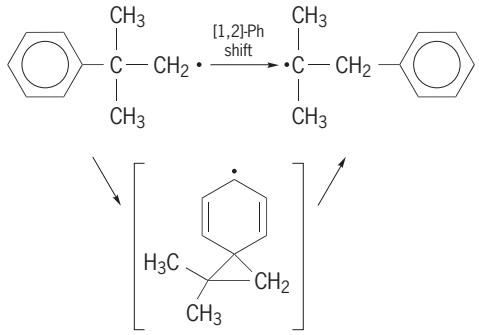

(9)

See FREE RADICAL; MOLECULAR ORBITAL THEORY.

Sigmatropic rearrangements. Concerted rearrangements that fall under the general category of pericylic reactions are called sigmatropic rearrangements and are covered by the Woodward-Hoffmann rules. Such reactions have a cyclic delocalized transition state, and the rearrangement involves shifts of both sigma and pi bonds. reaction (1) is an example of a [3,3] sigmatropic rearrangement, specifically a Cope rearrangement. A [1,5] sigmatropic rearrangement of hydrogen [reaction (10)] is an example of a degen-

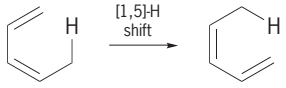

(10)

erate rearrangement, that is, the product is identical to the reactant. Detection of degenerate rearrangements requires subtle techniques, of which the most common and useful is nuclear magnetic resonance (NMR) spectroscopy, which can distinguish the specific environment of all of the hydrogen or carbon atoms in a molecule. Nuclear magnetic resonance spectra indicate that all of the carbon atoms are equivalent in bullvalene (3) [reaction (11)]. This is

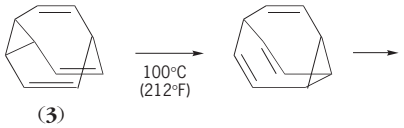

(3) (11)

attributed to multiple Cope rearrangements, which allow every carbon atom to take its turn at any of the various sites in the molecule; in all, there are over 1 million possible forms of rearranging the 10 carbon atoms in bullvalene, and the molecule appears to sample all of them. *See* NUCLEAR MAGNETIC RESONANCE (NMR).

Application of the Woodward-Hoffmann rules to sigmatropic rearrangements allows the prediction of what sizes of rearrangements are favorable (or allowed). Cyclic delocalized transition states containing $(4n + 2)$ electrons are most favorable, since they are predicted to proceed with the optimum stereochemistry; that is, all bonding can occur on the same side of the molecule (suprafacial). Thus the Cope rearrangement, a [3,3] rearrangement, involves a total of six electrons and is favorable. *See* WOODWARD-HOFFMANN RULE.

The knowledge of the favored stereochemistry also allows specific predictions regarding the correlation of reactant and product stereochemistry. A specific example of a [1,5]-H shift [reaction (12)] illus-

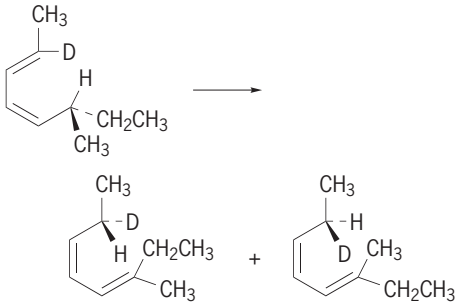

(12)

trates that the stereochemistry at the saturated centers that are the origin and terminus of the migration can be correlated with the stereochemistry of the original and final double bonds. The correlation indicates that the migration of H occurred specifically on the same side of the molecule, that is, suprafacially, part of the time along the top face and part of the time along the bottom face. *See* CHEMICAL BONDING; DEUTERIUM; ORGANIC SYNTHESIS; STEREOCHEMISTRY.

Carl C. Wamser

Bibliography. P. de Mayo, *Molecular Rearrangements*, 2 vols., 1963; P. de Mayo, *Rearrangements in Ground and Excited States*, 3 vols., 1980; J. March, *Advanced Organic Chemistry*, 5th ed., 2001.

Reciprocating aircraft engine

A fuel-burning internal combustion piston engine specially designed and built for minimum fuel consumption and light weight in proportion to developed shaft power. The rotating output shaft of the

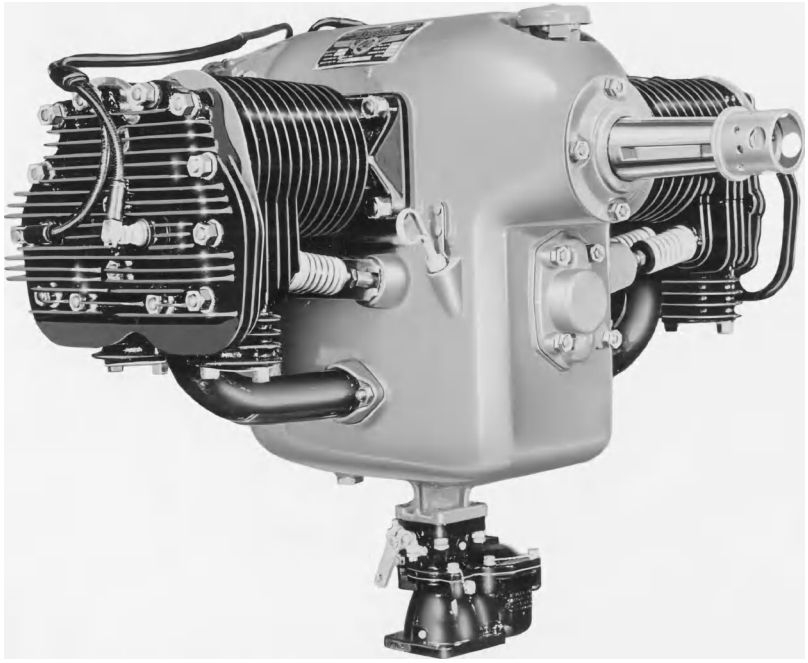

Fig. 1. Continental A-40 engine. (*Teledyne Continental Motors*)

engine may be connected to a propeller, ducted fan, or helicopter rotor that, in turn, accelerates the surrounding air, imparting an equal and opposite thrust force for propulsion of the aircraft.

Reciprocating aircraft engines are used in about 75% of all powered aircraft flying in the United States. Most of the aircraft powered by these engines belong to the general aviation segment of the domestic aviation fleet, which includes all aircraft except those flown by scheduled airlines and the military. There are over 200,000 active aircraft in the United States general aviation fleet. The reciprocating aircraft engine is used to power single-engine and multiengine airplanes, helicopters, and airships. It is the principal engine used in aircraft for air taxi, pilot training, business, personal, and sport flying as well as aerial application of seed, fertilizer, herbicides, and pesticides for farming. *See* AGRICULTURAL AIRCRAFT; GENERAL AVIATION.

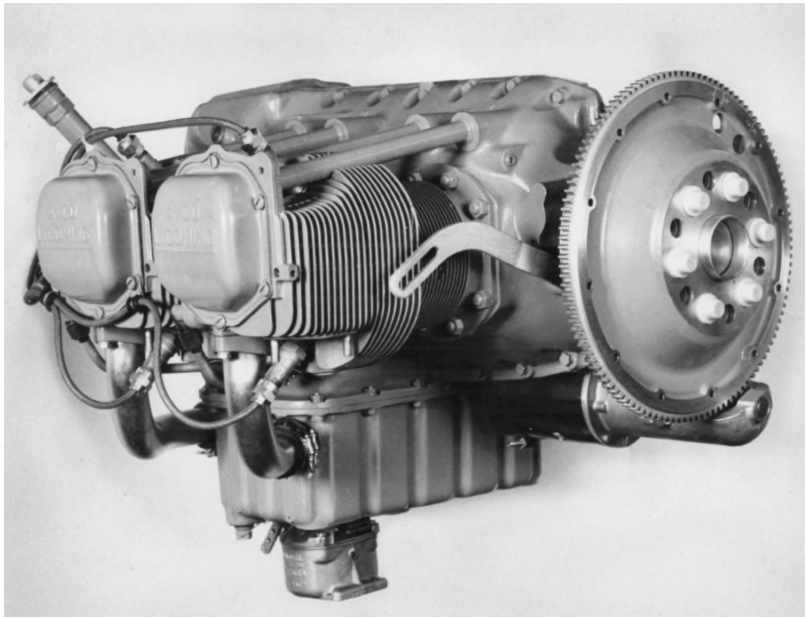

Fig. 2. Lycoming O-235 engine. (*AVCO Lycoming*)

Predominantly, reciprocating aircraft engines operate on a four-stroke cycle, where each piston travels from one end of its stroke to the other four times in two crankshaft revolutions to complete one cycle. The cycle is composed of four distinguishable events called intake, compression, expansion (or power), and exhaust, with ignition taking place late in the compression stroke and combustion of the fuel-air charge occurring early in the expansion stroke. These spark-ignition engines burn specially formulated aviation gasolines and produce shaft power by the force of combustion gas pressure on pistons acting on connecting rods turning a crankshaft. Major parts are the crankcase, crankshaft, connecting rods, pistons, cylinders with intake and exhaust valves, camshafts, and auxiliary operating systems such as ignition, fuel injection or carburetion, and fuel and oil pumps. *See* INTERNAL COMBUSTION ENGINE.

Four-stroke-cycle compression-ignition (diesel) reciprocating engines are currently certified in a limited segment of the general aviation market. The two-stroke-cycle spark-ignition engine has found widespread use in powered ultralight airplanes where horsepower requirements are generally less than 50 hp (37.5 kW). *See* ULTRALIGHT AIRCRAFT.

The reciprocating engine powered all aircraft for the first 40 years of heavier-than-air flight, including all military, commercial, and general aviation types. The advent of the turbojet engine near the end of World War II started a rapid conversion to turbine power for military aircraft. This conversion was accelerated by the Cold War and the Korean War. The emphasis on government financing of turbine engine development resulted in a virtual halt in further development of reciprocating engines after the end of World War II. The major swing to turbine power by commercial airlines began in 1958, and the design of smaller turbojet and turboprop engines started the gradual influx of business jets into general aviation. About 7% of all general aviation aircraft are turbine-powered, a proportion not expected to change drastically in the near future.

Most modern aircraft using engines with up to 450 hp (336 kW) output are powered by air-cooled, horizontally opposed, reciprocating engines. The smaller turbine engines that fall within this power range are used in helicopters (turboshaft engines), and the few turboprop engine designs below 450 hp have failed to meet the low cost and fuel economy necessary to compete with existing reciprocating engines. The gap in power between the largest reciprocating engine and the smallest popular turboprop has been filled by the rapidly dwindling supplies of post-World War II radial engines which are no longer manufactured, but for which spare parts are still available. This gap will probably disappear altogether as lower-cost, more efficient turboprop engines are developed to fill the needs of a growing general

aviation aircraft fleet. The trend in modern reciprocating engine development is toward lower engine weight and improved fuel economy rather than increased power.

Early development. Through World War I, aircraft had relatively low velocity head available for cooling. Velocity head is the term used to indicate the air pressure due to the forward velocity of the airplane in flight. A limited knowledge of air cooling of cylinders also tended to limit use of air-cooled fixed cylinders. Major combat aircraft were therefore mostly powered by water-cooled 6-cylinder in-line, 8-cylinder V, and 12-cylinder V types. There was, however, considerable use by the French of air-cooled rotary radial engines for fighter or pursuit aircraft and some use by the British of air-cooled V engines.

World War I demonstrated the great importance of aircraft as a military weapon. This resulted in the financing and development of a great variety of engine types and forms, particularly those promising higher power and improved aircraft performance.

During the years between the two world wars, with advancements in design and materials, the variety of engine types diminished to a few well-proved designs. These were the single-row 9-cylinder and two-row 14-cylinder air-cooled radials and the liquid-cooled V-12 engines for high-speed military aircraft. During the late 1930s, larger and much higher-powered two-row radials of both 14 and 18 cylinders were developed. Design work also emphasized higher outputs and high temperature cooling on liquid-cooled 12-cylinder V engines.

This period also saw the emergence of small, low-power engines which were to become the power plants of the general aviation industry. Before 1930, there were few reliable small aircraft engines. Manufacturers of small civil aircraft used surplus World War I engines like the Curtiss OX-5 liquid-cooled V-8, or modified automotive power plants. Two former automobile engine manufacturing companies turned to the light aircraft engine business in 1928 when it became apparent that the successful car manufacturers were interested in manufacturing their own engines. The first two engines were the Continental Model A-70, a 160-hp (119-kW) 7-cylinder air-cooled radial, and the Lycoming R-680, a 215-hp (160-kW) 9-cylinder radial. These two engines and their derivatives went on to power many civil and military aircraft, including the Boeing PT-13 Stearman biplane.

There was a demand for even lower-powered engines for the emerging civil aviation market. In 1931 a small 4-cylinder, horizontally opposed engine designated the A-40 (**Fig. 1**) was manufactured. It was first installed in the Taylorcraft E-2 airplane. The engine produced 37 hp (28 kW) at 2500 revolutions per minute and had a recommended major overhaul period of 500 h. The development of the horizontally opposed reciprocating aircraft engine is considered the single most important factor in the growth of the general aviation industry. The A-40 was followed in rapid succession by the A-50, A-65, A-75, and the Lycoming O-145. The O-145 engine powered the first successful helicopter, built and flown by Igor Sikorsky. These engines continued to be produced as power plants for aircraft manufactured by Aeronca, Luscombe, Piper, Taylorcraft and many others, up until 1942 when production was halted due to World War II. *See* HELICOPTER.

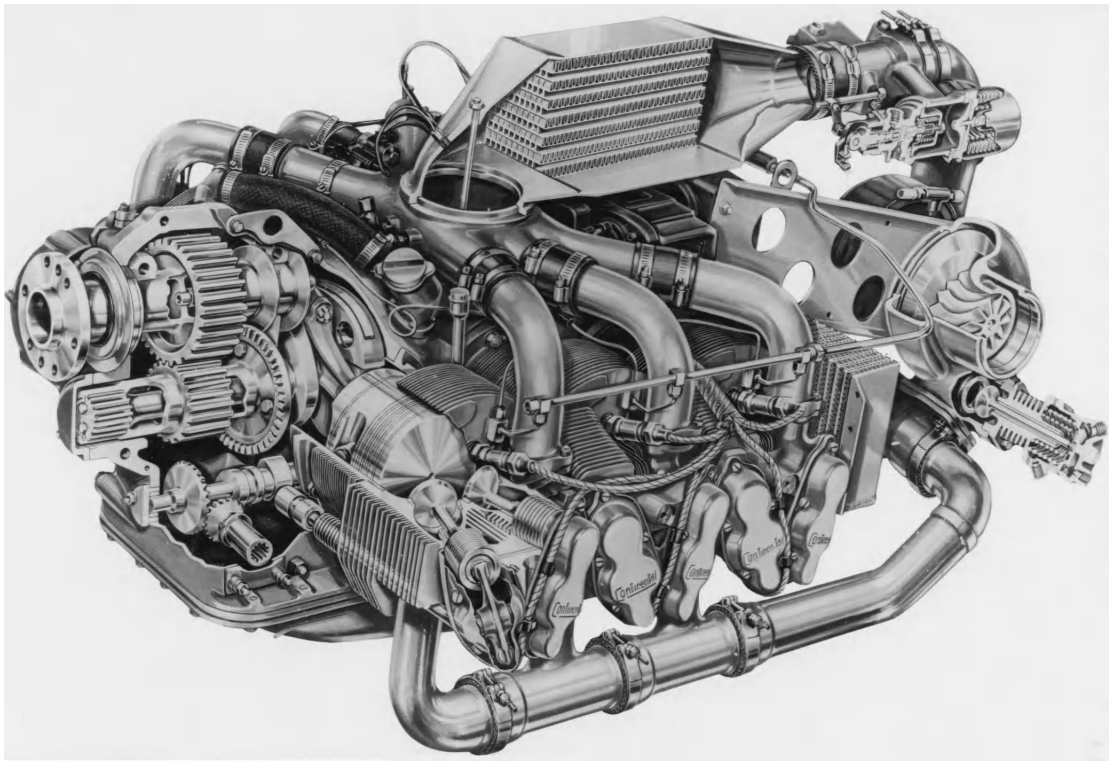

Fig. 3. Continental GTSIO-520 engine. (*Teledyne Continental Motors*)

Modern engines. Modern reciprocating engines are purely commercial products meeting the needs of the general aviation aircraft market. Modern aircraft reflect greatly improved economy, durability, and versatility in service. This is due in no small part to the advances made in small reciprocating engines since their inception in the early 1930s. Following the early 4-cylinder, horizontally opposed engines, production expanded after World War II to include 6-cylinder and 8-cylinder engines of the same format. Gradually these engines adopted mechanically driven superchargers, turbochargers with aftercooling, oil coolers to cope with the increasing power per unit displacement, fuel injection, propeller governors for constant speed props, gear reduction from the crankshaft to the propeller shaft, and sonic venturis for aircraft cabin pressurization.

The engines in production range from the 4-cylinder 115-hp (86-kW) Lycoming O-235 (**Fig. 2**) to the 6-cylinder 435-hp (324-kW) Continental GTSIO-520 (**Fig. 3**). These engines are used in over a hundred different aircraft designs, from small, two-seat training airplanes to sophisticated 18-seat commuter airliners. Some of the airplanes powered by the larger turbocharged engines fly at maximum true airspeeds of up to 260 knots (482 km/h) and altitudes of 30,000 ft (9144 m).

Fig. 4. Bombardier/Rotax Type 447 two-cylinder, in-line engine. (*BRP-Rotax GmbH*)

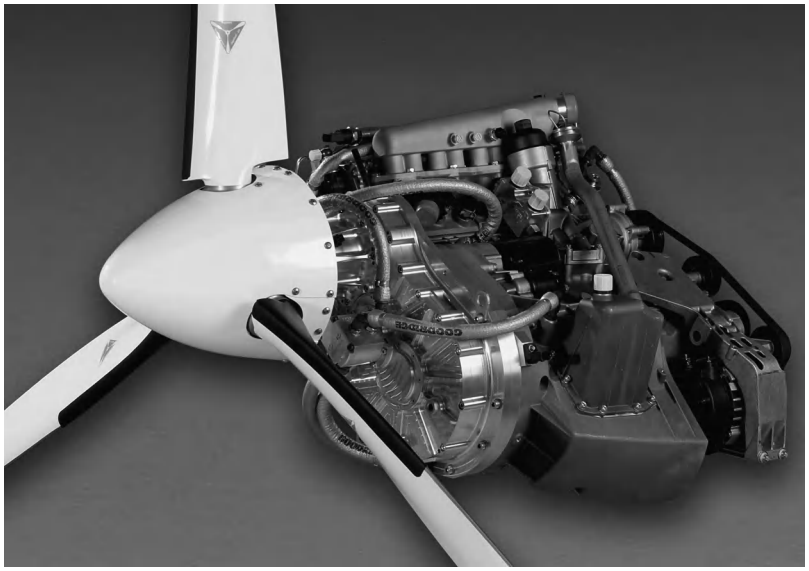

Fig. 5. Thielert Centurion 1.7 diesel aircraft engine. (*Thielert Aircraft Engines GmbH*)

Fig. 6. IOL-200, four-cylinder, liquid-cooled engine used in the *Voyager* airplane during its nonstop, unrefueled circumnavigation of the globe. (*Teledyne Continental Motors*)

On the smaller end of the scale, the demand for low-cost recreational flying has opened another growth area for the reciprocating aircraft engine market. Powered ultralight airplane manufacturers are increasingly taking advantage of new stronger and lighter materials, advanced structural design techniques, and improved aerodynamics to produce a class of airplanes which is not required to be certified by the U.S. Federal Aviation Administration (FAA) and for which no pilot's license is needed. Most of the engines used to power these ultralights are 1-cylinder or 2-cylinder, two-stroke-cycle, and fan-cooled, and operate on automobile gasoline. These engines are adapted from existing uses (for example, lawnmower, snowmobile) and drive pusher propellers through a multiple-sheave-belt reduction drive. The power output of these engines is usually less than 50 hp (37.5 kW). The Bombardier/Rotax Type 447 engine (**Fig. 4**) is typical of the 2-cylinder, in-line engines.

Thielert Aircraft Engines GmbH of Germany has certified a 135-hp (99-kW), four-stroke-cycle, four-cylinder, turbocharged diesel engine (**Fig. 5**), which operates on either automotive diesel or jet fuel. As of 2005, there were over 500 general aviation airplanes flying with these Centurion 1.7 engines. The engine incorporates a full-authority digital engine control (FADEC) system which reduces the pilot's workload by controlling both the fuel flow and the rotational speed of the propeller.

Voyager engines. In December 1986, the experimental aircraft *Voyager*, piloted by D. Rutan and J. Yeager, completed the first nonstop, unrefueled circumnavigation of the globe by an airplane. *Voyager* was powered by two reciprocating aircraft engines and carried over 7000 lb (3175 kg) of fuel.

Voyager's front engine was a conventional, air-cooled, 4-cylinder O-240 model that had been in production for 18 years. Rated at 130 hp (97 kW), it

was operated in conjunction with the rear engine for the initial takeoff and for various periods during the flight to climb to higher altitudes to avoid mountains and bad weather.

The main, liquid-cooled rear engine, an experimental design designated IOL-200 (**Fig. 6**), was a 4-cylinder, horizontally opposed design producing 110 hp (82 kW) at 2750 revolutions per minute. Specially designed for low oil and fuel consumption, it had a high-turbulence combustion chamber and a compression ratio of 11.4:1. The rear engine ran for the entire 9 days of the flight, except for a 5-min period when it shut down because of a fuel-transfer problem. At the end of the flight, with only 106 lb (48 kg) of fuel remaining, the *Voyager* had averaged a fuel consumption of 22 mi/gal (9.35 km/liter).

Kenneth J. Stuckas

Bibliography. B. Gunston, *The Development of Piston Aero Engines*, 2d ed., Patrick Stevens, Ltd., 1999; IAP Staff (ed.), *Aircraft Reciprocating Engines*, 1985; P. Jackson (ed.), *Jane's All the World's Aircraft*, Jane's Information Group, annually J. L. Kerrebrock, *Aircraft Engines and Gas Turbines*, 2d ed., MIT Press, 1992; M. Kroes and T. W. Wild, *Aircraft Powerplants*, 7th ed., McGraw Hill, 1994; H. Smith, *A History of Aircraft Piston Engines*, Sunflower University Press, 1993; P. H. Wilkenson, *Aircraft Engines of the World*, 1966–1967, 1964–1965, 1944–1962, 1941.

Reciprocity principle

In the scientific sense, a theory that expresses various reciprocal relations for the behavior of some physical systems. Reciprocity applies to a physical system whose input and output can be interchanged without altering the response of the system to a given excitation. Optical, acoustical, electrical, and mechanical devices that operate equally well in either direction are reciprocal systems, whereas unidirectional devices violate reciprocity. The theory of reciprocity facilitates the evaluation of the performance of a physical system. If a system must operate equally well in two directions, there is no need to consider any nonreciprocal components when designing it.

Examples of reciprocal systems. Some systems that obey the reciprocity principle are any electrical network composed of resistances, inductances, capacitances, and ideal transformers; systems of antennas, with restrictions according to Eq. (2); mechanical gear systems; and light sources, lenses, and reflectors.

Devices that violate the theory of reciprocity are transistors, vacuum tubes, gyrators, and gyroscopic couplers. Any system that contains the above devices as components must also violate the reciprocity theory. The gyrator differs from the transistor and vacuum tube in that it is linear and passive, as opposed to the active and nonlinear character of the other two devices. *See* FARADAY EFFECT.

Rayleigh's theorem of reciprocity. Reciprocity is concisely expressed by a theorem originally proposed by Lord Rayleigh for acoustic systems and later generalized by J. R. Carson to include electromagnetic systems. Both mathematical expressions of the theory of reciprocity are closely related to the mathematical theorem known as Green's theorem. The acoustical reciprocity theorem of Lord Rayleigh is as follows: In an acoustic system consisting of a fluid medium having boundary surfaces $s_1, s_2, \ldots, s_k$ and subject to no impressed body forces, surface integral (1) holds. Here p_1 and p_2 are the pressure

$$\int_s (p_1 v_{2n} - p_2 v_{1n})\, ds = 0 \qquad (1)$$

fields produced respectively by the components of the fluid velocities v_{1n} and v_{2n} normal to the boundary surfaces $s_1, s_2, \ldots, s_k$. The integral is evaluated over all boundary surfaces. *See* GREEN'S THEOREM.

For a region containing only one simple source H. L. F. Helmholtz has shown that the theorem can be expressed as follows: A simple source at A produces the same sound pressure at B as would have been produced at A had the source been located at B. In other words, the response of a human ear at B due to a vibrating tuning fork at A is the same as the response of the ear at A due to the same tuning fork when located at B. The human ear, the tuning fork, and the intervening acoustical media constitute a physical system that obeys the theory of reciprocity.

Electromagnetic systems. The generalization of Lord Rayleigh's theorem to electromagnetic systems can be mathematically expressed by volume integral (2) where $\mathbf{E}_1$ and $\mathbf{H}_1$ are the electric and mag-

$$\int_v \nabla \cdot (\mathbf{E_1} \times \mathbf{H_2} - \mathbf{E_2} \times \mathbf{H_1})\, dv = 0 \qquad (2)$$

netic field vectors describing a state due to one electromagnetic sound and $\mathbf{E}_2$ and $\mathbf{H}_2$ describe another state due to a second source. The above relation is valid as long as the medium is isotropic and the field vectors are finite and continuous, and vary according to a linear law (thus excluding ferromagnetic materials, electronic space charges, and ionized gas phenomena).

By means of Maxwell's equations, relation (2) can be expressed in another form when restricted to systems of conduction current only where $\mathbf{J}_1$ and $\mathbf{J}_2$ are the conduction current densities in an electromagnetic system due to the action of the external electric fields $\mathbf{E}_1$ and $\mathbf{E}_2$, respectively.

Equation (3) is readily applied to antennas and ra-

$$\int_v (\mathbf{E_1} \cdot \mathbf{J_2} - \mathbf{E_2} \cdot \mathbf{J_1})\, dv = 0 \qquad (3)$$

diation. If, in **Fig. 1**, $\mathbf{J}_1$ is the resulting current density in antenna B due to an electric field $\mathbf{E}_1$ established by antenna A, and $\mathbf{J}_2$ is the current density in antenna A due to electric field $\mathbf{E}_2$ established by antenna B, then $\mathbf{J}_1 = \mathbf{J}_2$, provided $\mathbf{E}_1 = \mathbf{E}_2$. The two emfs need not be applied at the same instant of time. The integral in Eq. (3) over all space reduces to an integral over the two antennas since $\mathbf{J}_1$ and $\mathbf{J}_2$ are zero elsewhere. From this particular application of the

reciprocity theorem it is seen that the transmitting and receiving patterns of an antenna are the same.

Equation (3), when evaluated over an N-mesh electrical network, reduces to Eq. (4), where a and b are

$$\sum_{j=1}^{N} V_{aj} i_{bj} = \sum_{j=1}^{N} V_{bj} i_{aj} \tag{4}$$

two different states of the network and the j subscript denotes in which of the N meshes the voltage and current are measured. For the two-mesh network in **Fig. 2**, Eq. (4) gives Eq. (5). Expressed in

$$V_{a1} i_{b1} = V_{b2} i_{a2} \tag{5}$$

words: If an emf source of magnitude V and zero internal impedance, when applied to terminals 1–1, produces a current I at terminals 2–2, then the same current I will be measured at terminals 1–1 when the emf V is applied to terminals 2–2. This statement, that is, Eq. (5), is probably the most familiar form of the theorem of reciprocity.

Electrostatic systems. The statement of reciprocity for electrostatics is given by Eq. (6), where V_1 and

$$\int_v \rho_1 V_2 dv = \int_v \rho_2 V_1 dv \tag{6}$$

V_2 are the electric potentials produced at some arbitrary point due to the volume charge distributions ρ_1 and ρ_2, respectively. When the integral expression in Eq. (6) is applied to the electrostatic system of two charged conductors in **Fig. 3**, it becomes Eq. (7). Here V_a is the potential on conductor a due

$$V_a q_a = V_b q_b \tag{7}$$

to charge q_b on conductor b; the remaining quantities are similarly defined. In other words, if a charge

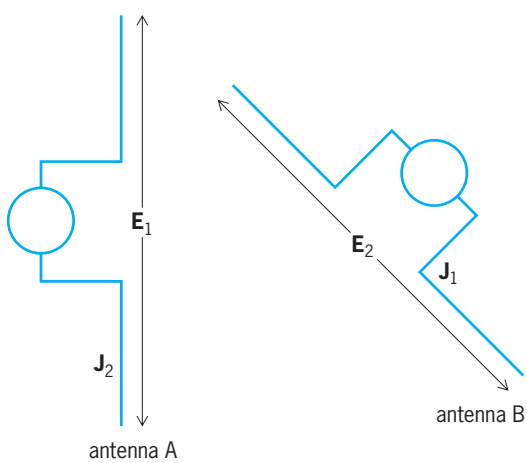

Fig. 1. Antenna system.

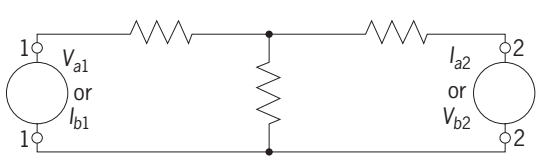

Fig. 2. Two-mesh network.

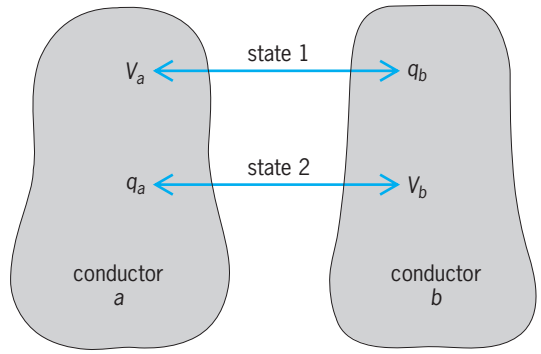

Fig. 3. Charged conducting bodies.

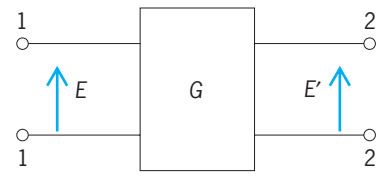

Fig. 4. Four-terminal, black-box network.

q_b on conductor b raises the potential of conductor a to V, then the same charge on conductor a raises the potential of conductor b to V.

Electrical networks. A somewhat different approach to reciprocity is the so-called black box, or two-terminal pair, method illustrated in **Fig. 4**. The box might contain a mechanical, acoustical, optical, or electrical system. The applied excitation or cause is E and the response or effect is E'. The ratio of E/E' (or E'/E) is the transfer function G for the system within the black box. Using the subscript notation of G_{12} when E is impressed at terminals 1–1 and E' is measured at terminals 2–2, then G_{21} represents a response measured at 1–1 for an excitation at 2–2. Mathematically the general behavior of the box to excitations at both sets of terminals can be expressed by Eqs. (8), as long as the response bears a linear

$$\begin{aligned} E_1 &= G_{11} E_1' + G_{12} E_2' \\ E_2 &= G_{21} E_1' + G_{22} E_2' \end{aligned} \tag{8}$$

relation to the excitation.

If, in addition to its linear characteristic, the system satisfies Eq. (9), the principle of reciprocity is

$$G_{12} = G_{21} \tag{9}$$

obeyed, and the device will operate equally in either direction. Whenever $G_{12} \neq G_{21}$, the system violates the theory of reciprocity, with the result that the response in one direction is different from that obtained in the other direction. *See* NETWORK THEORY.

Hugh S. Landes

Bibliography. B. Bleaney and B. I. Bleaney, *Electricity and Magnetism*, 3d ed., 1976, reprint 1989; D. E. Gray (ed.), *American Institute of Physics Handbook*, 3d ed., 1972; P. Hammond and J. K. Sykulski, *Engineering Electromagnetism: Physical Processes and Computation*, 1994; J. D. Kraus, *Antennas*, 2d ed., 1988; M. Valkenburg (ed.), *Reference Data*

for Engineers: Radio, Electronics, Computers, and Communications, 8th ed., 1992; J. Vanderlinde, *Classical Electromagnetic Theory*, 1993.

Recombination (genetics)

The formation of new genetic sequences by piecing together segments of previously existing ones. Classical genetics is based on the transmission of traits, attributable to particular genetic determinants or genes, through sexual generations. The genes are predominantly present in chromosomes, the thread-like structures that fill the cell nucleus and divide as the cells divide. The key component of the chromosomes is deoxyribonucleic acid (DNA); each chromosome has a single coiled and folded double-stranded DNA molecule running along its length, and the genes of the chromosomes occupy segments of this DNA. *See* CHROMOSOME; DEOXYRIBONUCLEIC ACID (DNA).

Recombination often follows DNA transfer in bacteria and, in higher organisms, is a regular feature of sexual reproduction. The basic feature of sexual reproduction is the cyclical alternation between the diploid state, with a double set of chromosomes (two of each kind), and the haploid state, with a single set. Animals and most flowering plants are diploid, and their germ cells (eggs and sperm, or eggs and pollen grains) are haploid. In most fungi, the organism is normally haploid, but sexual fertilization creates a transient diploid cell that immediately divides to form a tetrad of four haploid spores. *See* REPRODUCTION (ANIMAL); REPRODUCTION (PLANT).

In all sexually reproducing organisms, the transition from the diploid to the haploid state is brought about by the process of meiosis—two rounds of cell division accompanied by only one division of the chromosomes. The two chromosomes of each kind present in the diploid meiotic mother cell are necessarily segregated into different haploid meiotic products. *See* MEIOSIS.

Free reassortment and linkage. When the diploid cell undergoing meiosis is heterozygous, that is, carries two alternative forms (alleles) of the same gene inherited from different parents, the meiotic segregation of chromosome pairs will ensure that half of the meiotic products (germ cells) will carry one alternative and half the other. When the diploid is doubly heterozygous (*A/a B/b*), the two pairs of alleles segregate independently of one another at meiosis, provided the two gene differences are associated with different chromosome pairs. Consequently, four possible products—*AB*, *Ab*, *aB*, and *ab*—will occur with statistically equal frequency. *See* ALLELE; MENDELISM.

If, in a double heterozygote *Aa Cc*, the two genes are in the same chromosome and they are not very far apart, they show linkage, with the parental combinations (*AC* and *ac*, for example) tending to appear in the meiotic products more frequently than the recombinant types (*Ac* and *aC*). The fact that recombinants types *Ac* and *aC* occur at all is due to reciprocal exchanges between chromosomes (crossing over) that take place in the first meiotic division. The reciprocal nature of these events is shown most clearly in fungi such as yeast and the filamentous fungus *Neurospora crassa*, where the four products of a single meiosis can be recovered together as a tetrad of haploid cells (or, in the case of *Neurospora*, of spore pairs). In such organisms, one recombinant type (*Ac*) is generally accompanied in the same tetrad by the reciprocal product (*aC*). *See* CROSSING-OVER (GENETICS); LINKAGE (GENETICS).

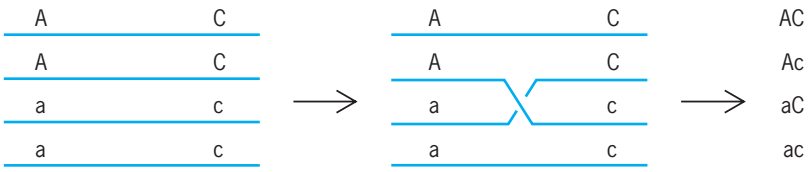

Fig. 1. Crossing-over in meiotic tetrads. Only two of the four products usually show crossing-over of linked markers.

Only two of four tetrad members are generally recombinant. Tetrads of constitution *AC*, *Ac*, *aC*, *ac* are far commoner than those containing four recombinant products, the inference being that crossing-over between chromosome pairs occurs after the chromosomes have replicated and that each crossover involves only half of each divided chromosome (chromatid) [**Fig. 1**]. If more than one crossover occurs within the linkage group, it is a matter of chance which chromatid of each chromosome is involved in each chiasma.

Chromosome maps. All the segregating genetic markers found in a given organism can be assigned to any of a number of linkage groups. Markers in different groups reassort independently, whereas markers in the same group are linked. The number of linkage groups is the same as the haploid number of chromosomes, and through correlations of changed linkage relationships with specific chromosome aberrations, each linkage group can be assigned to a specific chromosome (**Fig. 2**).

The measure of linkage is recombination frequency $[(Ac + aC)/(AC + Ac + aC + AC)]$, which is often expressed as a percentage, and it must be significantly less than 50% to be distinguishable from random reassortment. The map distance, in map units or centimorgans, between two markers is 100 times the mean number of crossovers per chromatid, or 50 times the mean number per meiotic tetrad. If there is never more than one crossover between the markers, recombination frequency equals map distance. However, if there is a significant number of double or multiple crossovers, recombination frequency is an underestimate of map distance because each additional crossover within a marked interval has as much chance of canceling recombination as of causing it. Above a certain distance, recombination frequency approaches a limiting value of 50%. In practice, the linkage map of each chromosome is built up from the summing of a series of intervals, each too short for double crossovers to have a significant effect. A corollary of the theory of chromosome mapping is that the total map length of a

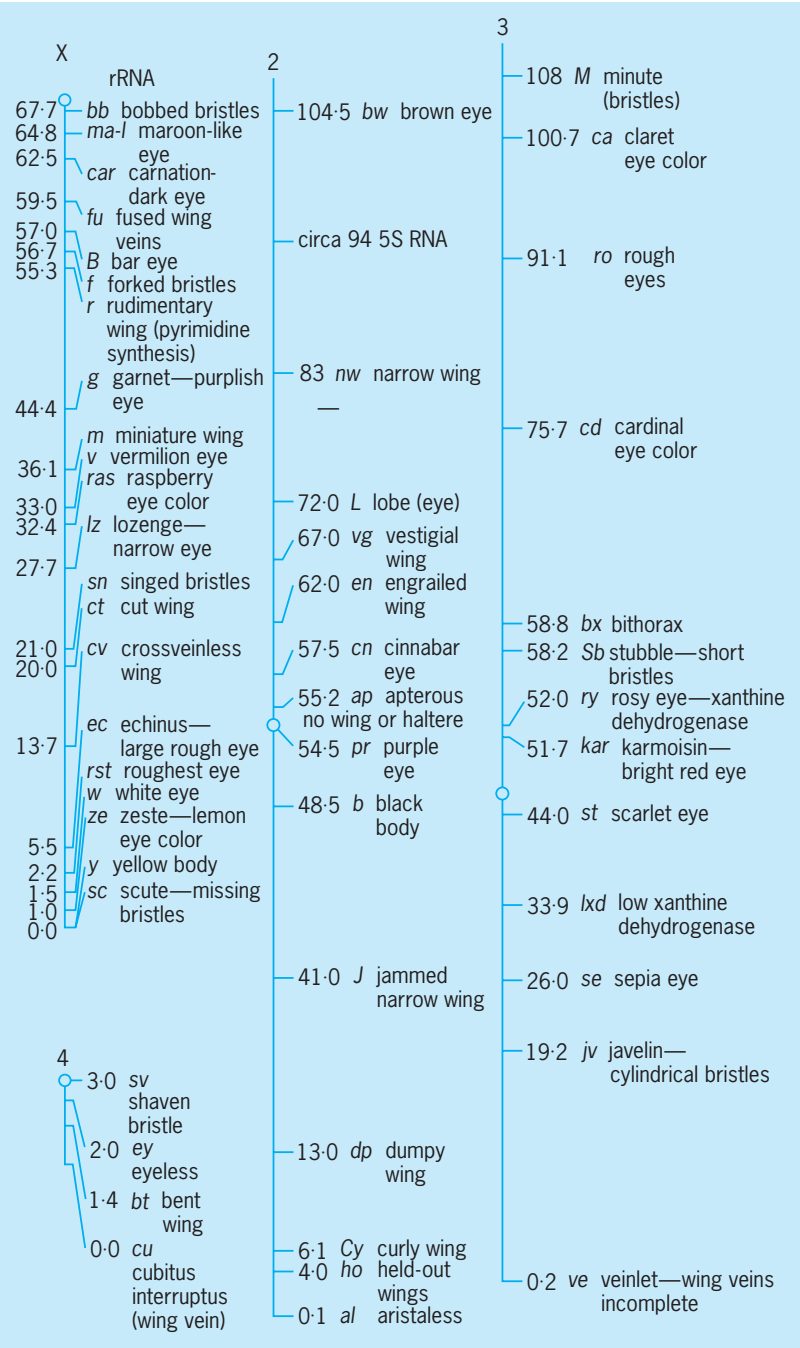

Fig. 2. Linkage maps for chromosomes X, 2, and 3 of *Drosophila melanogaster* at the metaphase state of mitosis. Numbers indicate the distance from the upper end of the chromosome of important genes (in map units). (*From J. R. S. Fincham, Genetics, Jones and Bartlett, 1983*)

linkage group should be 50 times the mean chiasma frequency of the corresponding chromosome pair at meiosis. This has been confirmed in organisms, such as corn (*Zea mays*), where both relatively complete linkage maps and accurate chiasma counts are available. *See* GENETIC MAPPING.

The average amount of DNA corresponding to 1% recombination (one map unit) varies enormously between species, from less than 10,000 DNA base pairs in the yeast *Saccharomyces cerevisiae* to about a million base pairs in humans and even higher values in some amphibia and flowering plants. Furthermore, the value varies from one part of the genome to another within a species, since not all chromosome regions have the same probability of crossing-over.

A standard method of determining the sequence of three linked markers is the three-point test cross. This measures the frequencies of the eight possible kinds of meiotic product formed by a diploid segregating with respect to three linked markers, say *ACD/acd*. If the markers are in the order written, recombinant meiotic products of constitution *AcD* or *aCd* will be relatively infrequent compared with *Acd*, *aCD*, *ACd*, and *acD*, since, unlike the latter classes, which arise from single crossovers, they require relatively rare double crossovers for their formation. The method would be seriously undermined if crossovers tended to occur in clusters such that doubles and multiples were as probable as singles. In fact, the occurrence of a crossover in one chromosome interval tends to reduce the likelihood of a crossover in an adjacent interval, an effect called crossover or chiasma interference. A decreased probability of occurrence of crossovers beyond the first has also been inferred from chiasma counts under the microscope. Whereas the chance of at least one chiasma is nearly 100% (very few chromosome pairs are left unjoined by chiasmata at the first meiotic division), shorter chromosomes commonly have no more than one and even the longest may never have more than three or four.

Conversion. Recombination was once thought to occur only between genes, never within them. Indeed, the supposed indivisibility of the gene was regarded as one of its defining features, the other being that it was a single unit of function. The functional criterion was failure of alleles to complement one another. If two defective mutants of independent origin failed to compensate for each other's defects to determine a more normal phenotype when present together in a heterozygote, then neither was thought capable of supplying the normal function that was defective in the other; hence, both must be defective in the same gene, defined as a unit of function. The two definitions agreed to the extent that mutations placed in the same gene by the complementation criterion were always at least very closely linked. However, examination of very large progenies shows that, in all organisms studied, nearly all functionally allelic mutations of independent origin can recombine with each other to give nonmutant products, generally at frequencies ranging from a few percent (the exceptionally high frequency found in *Saccharomyces*) down to 0.001% or less. The indivisible unit of recombination is no larger than a single DNA base-pair.

Mutations within a gene can be mapped in a linear order either by using the criterion of recombination frequency or, more accurately, by deletion mapping, which uses deletions of extended segments of the gene. Overlapping deletion mutations cannot recombine to give normal recombinants since neither has all of the sequence in which the other is defective; for the same reason, a deletion mutation gives no recombinants with any single-base-pair mutation

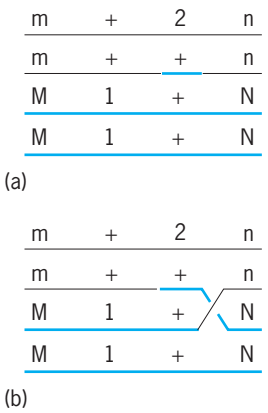

Fig. 3. Typical patterns of intragenic recombination in meiotic tetrads: 1 and 2 are mutant sites within the gene (+ is the wild-type site); *M,m* and *N,n* are flanking markers. (*a*) Conversion of site 1 to wild type, to give a wild-type recombinant gene, is accompanied by adjacent crossing-over in this tetrad but (*b*) not in the other tetrad. Contributions of the two recombining chromosomes are distinguished by black and colored lines.

falling within the region defined by the deletion. Given a series of overlapping deletions, any point mutation can be positioned within one or any other of the segments defined by the deletions and their overlaps.

Tetrad analysis, especially in *S. cerevisiae*, shows that recombination within genes is most frequently nonreciprocal. In a cross between two allelic mutants, both nonmutant and double-mutant products are formed, but seldom in the same tetrad (**Fig. 3**). This phenomenon is termed gene conversion, and must represent the nonreciprocal transfer of DNA sequence between chromatids of paired chromosomes. Gene conversion is also found in crosses involving only a single mutational difference. In addition to the expected 2:2 tetrad segregation pattern, up to several percent 3:1 and 1:3 ratios can be found in yeast; frequency in other experimental organisms is generally very much lower. Both in *Saccharomyces* and other fungi, genes often show a gradient of conversion frequency from one end of a gene to the other, an effect called conversion polarity.

Conversion sometimes affects only half a chromatid, equivalent to a single DNA strand, rather than both strands of the double helix. In this case, formation of a completely converted gene is delayed until after the first replication of the DNA following meiosis. When donor and recipient genes differ in more than one mutational marker, the two markers are often transferred together (coconverted). The frequency of coconversion, compared with single-marker conversion, diminishes as the distance between the markers increases; the relationship between frequency and distance indicates that conversion involves the transfer of segments of DNA commonly of the order of several hundred to a few thousand base-pairs in length. Coconversion does not result in any marker recombination; in order to recombine two markers within a gene, a conversion segment has to include one but not the other. *See* MUTATION.

Unifying theory. Recombination between and within genes may seem to be two distinct processes involving reciprocal crossing-over and nonreciprocal conversion, respectively. It is, however, attractive to postulate that both follow from the same kind of initiating event. Conversion within a gene is often (with a frequency of 40–50%) accompanied by a closely adjacent crossover, detectable through the reciprocal recombination of markers flanking the gene undergoing conversion (Fig. 3). Both conversions and crossovers are probably formed from precursor structures involving local unilateral transfer of DNA segments between chromatids. Such structures, which are called conversion tracts, will lead to observable gene conversion only in the rare event that the transferred segment happens to include a genetic marker. If the structure develops into a crossover (which is presumed to happen in yeast with a probability of 40–50%), it will always lead to reciprocal crossing-over between flanking markers, however distant they may be from the initiating event. In the case of recombination between extremely close markers (particularly those within the same gene), the initiating event is necessarily very close to the markers being recombined, which will thus have a high probability of being included in the conversion tract.

Molecular mechanisms. There are two prevailing hypotheses as to the nature of the event that initiates recombination. In the first model, a chromatid involved in meiotic pairing is cut in one strand of its DNA, probably at one of a large number of special sites distributed along the chromosome. A stretch of single-stranded DNA then unwinds from one side of the cut and invades the DNA of the other chromatid, displacing the corresponding single-strand segment from its duplex. In the simplest sequence of events, the invading single strand becomes sealed into the recipient duplex, creating a mismatch between the complementary strands at any point at which the recipient and donor duplexes differ. Such a mismatch can be left unrepaired, in which case the result will be half-chromatid conversion and postmeiotic segregation; or (more frequently, at least in yeast) it is repaired by excision and replacement of the mismatch in one strand either to restore the original arrangement or to convert the recipient whole chromatid to the type of the donor. To explain the association between conversion and crossing-over, it is supposed that the single-strand bridge can lead to a reciprocal single-strand exchange (a Holliday junction) and then to a reciprocal double-strand exchange, that is, a regular crossover. The chance of a single-strand transfer leading to a crossover is variable but seldom exceeds 50% (**Fig. 4*a***).

A radically different hypothesis proposes that the initiating event is a double-stranded break, which may be subsequently enlarged to a gap, in one chromatid. The gap is healed by copying of both strands of the undamaged partner chromatid. The repair process involves the formation of Holliday junctions on both sides of the gap and, depending on the way these are resolved, crossing-over may take place. The

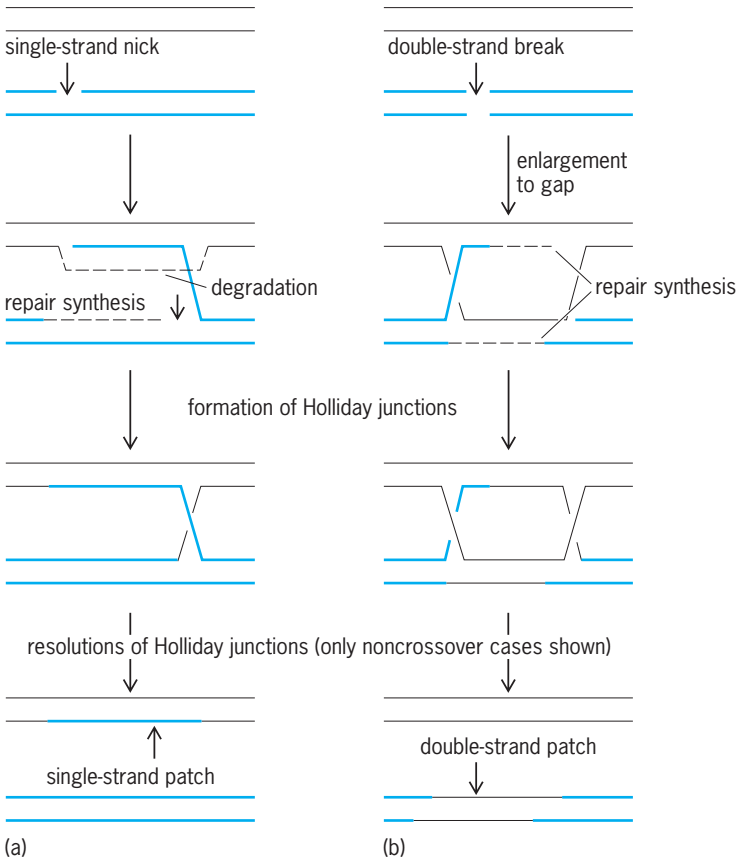

Fig. 4. Typical models of general recombination. (*a*) Single-strand assimilation model. (*b*) Double-strand gap repair model. In these models, conversion occurs without crossing-over, but other, supposedly equally probable, models of resolution of the Holliday junctions lead to crossing-over adjacent to the conversion event. Contributions of the recombining chromatids are differentiated by black and colored lines.

double-strand break-repair hypothesis accounts for conversion either with or without adjacent crossing-over, but it does not readily explain the occurrence of conversion events that affect only half a chromatid, leading to postmeiotic segregation. However, this too can be explained if the regions of heteroduplex that would be formed on each side of the gap (from annealing of single-strand ends into the undamaged duplex) are sufficiently extensive (Fig. 4*b*).

Both hypotheses provide for a scattering of conversion segments in all chromosome pairs, some of which are associated with crossovers. Both explain polarity in gene conversion as due to the initial DNA cuts, whether single- or double-stranded, occurring at preferred sites in the DNA; the probability of heteroduplex formation or gapping, and thus the probability of a marker undergoing conversion, should decrease with increasing distance from such a site. The double-strand break-repair model appears more probable, inasmuch as a site at the high-conversion end of at least one yeast gene appears to undergo double-strand breakage at meiosis. However, elements of both proposed mechanisms may have to be included in a more refined model.

Most speculation on recombination mechanisms has focused on recombining chromosomes as DNA molecules, which is an oversimplification. Little is yet known about how the higher-order DNA protein structure of chromosomes affects their recombination. The role of the synaptonemal complex, across which the exchanged strands of DNA may need to pass, is also obscure. *See* MOLECULAR BIOLOGY.

Mitotic recombination. Crossing-over between homologous chromosome pairs can also occur during the prophase of mitotic nuclear division. The frequency is very much lower than in meiosis, presumably because the mitotic cell does not form the synaptic apparatus for efficient pairing of homologs. Mitotic crossing-over has been studied in the fruit fly *Drosophila melanogaster*, in the filamentous fungus *Aspergillus nidulans*, and in *Saccharomyces* yeast. In these species it is detected through the formation of homozygous clones of cells in an initially heterozygous diploid. There is a 50% chance of homozygosity in daughter cells whenever a crossover occurs between chromatids in the interval between the marker and the centromere, the chromosomal site of attachment to the mitotic spindle (**Fig. 5**). The frequency of mitotic crossing-over is greatly increased by radiation. *See* MITOSIS.

Integration of bacterial DNA fragments. Bacteria have no sexual reproduction in the true sense, but many or most of them are capable of transferring fragments of DNA from cell to cell by one of three mechanisms. (1) Fragments of the bacterial genome can become joined to plasmid DNA and transferred by cell conjugation by the same mechanism that secures the transfer of the DNA of transmissible plasmids. (2) Genomic fragments can be carried from cell to cell in the infective coats of bacterial viruses (phages), a process called transduction. (3) Many bacteria have the capacity to assimilate fragments of DNA from solution and so may acquire genes from disrupted cells.

Fragments of DNA acquired by any of these methods can be integrated into the DNA of the genome in place of homologous sequences previously present. If the incoming DNA has no homology with that of the recipient cell, it usually cannot be integrated and is lost for lack of ability to replicate autonomously. Homologous integration in bacteria is similar, in its nonreciprocal nature and perhaps also in its mechanism, to gene conversion in eukaryotic organisms. *See* BACTERIAL GENETICS; BACTERIOPHAGE; TRANSDUCTION (BACTERIA).

Site-specific recombination. Bacteriophages, plasmids, bacteria, and unicellular eukaryotes provide many examples of differentiation through controlled and site-specific recombination of DNA segments. In vertebrates, a controlled series of deletions leads to the generation of the great diversity of gene sequences encoding the antibodies and T-cell receptors necessary for immune defense against pathogens. All these processes depend upon interaction and recombination between specific DNA sequences, generally but not always with some sequence similarity, catalyzed by site-specific recombinase enzymes. The molecular mechanisms may have some similarities with those responsible for general meiotic recombination, except that the latter does not depend on any specific sequence, only on similarity (homology) of the sequence recombined.

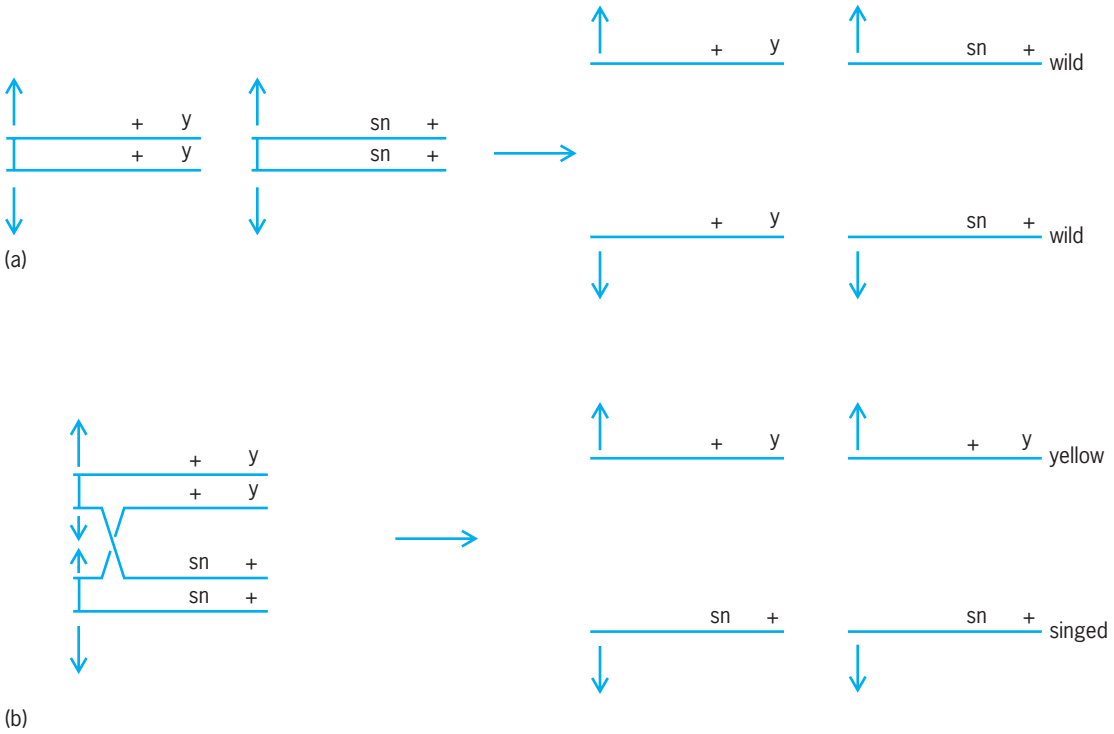

Fig. 5. Mitotic crossing-over, as demonstrated by one *Drosophila* X chromosome in the female carrying the cuticle marker *yellow* (y), and the other the linked marker *singed* (sn), affecting bristle morphology. (*a*) Normal mitosis, with heterozygosity of both markers maintained and a wild-type appearance of the fly, since both mutations are recessive. (*b*) The effect of a crossover in the position indicated during embryonic development, which gives (with 50% probability) adjacent yellow and singed patches on the surface of the adult fly.

Nonhomologous recombination. Techniques have been devised for the artificial transfer of DNA fragments from any source into cells of many different species, thus conferring new properties upon them (transformation). In bacteria and the yeast *S. cerevisiae*, integration of such DNA into the genome (on which the stability of transformation generally depends) requires substantial sequence similarity between incoming DNA and the recipient site. However, cells of other fungi, higher plants, and animals are able to integrate foreign DNA into their chromosomes with little or no sequence similarity. These organisms appear to have some unidentified system that recombines the free ends of DNA fragments into chromosomes regardless of their sequences. It may have something in common with the mechanism, equally obscure, whereby broken ends of chromosomes can heal by nonspecific mutual joining. *See* TRANSFORMATION (BACTERIA).

Recombinant DNA in genetic engineering. The science of genetics has been revolutionized by the development of techniques using isolated cells for specific cleaving and rejoining of DNA segments and the introduction of the reconstructed molecules into living cells. This artificial recombination depends on the use of site-specific endonucleases (restriction enzymes) and DNA ligase. *See* GENE; GENE ACTION; GENETIC ENGINEERING; GENETICS; RESTRICTION ENZYME. J. R. S. Fincham

Bibliography. J. R. S. Fincham and P. Oliver, Initiation of recombination, *Nature*, 338:14–15, 1989; J. Ilan (ed.), *Translational Regulation of Gene Expression*, 2d ed., 1993; M. Karin (ed.), *Gene Expression: General and Cell-Type-Specific*, 1993; K. B. Low (ed.), *The Recombination of Genetic Material*, 1988; M. Meselson and C. M. Radding, A general model for recombination, *Proc. Nat. Acad. Sci. USA,* 72:821–828, 1982; A. Prokop, R. K. Bajpapi, and C. S. Ho, *Recombinant DNA Technology and Applications*, 1991; J. W. Szostak et al., The double-strand break repair model of recombination, *Cell,* 33:26–35, 1983.

Rectifier

A nonlinear circuit component that allows more current to flow in one direction than in the other. An ideal rectifier is one that allows current to flow in one (forward) direction unimpeded but allows no current to flow in the other (reverse) direction. Thus, ideal rectification might be thought of as a switching action, with the switch closed for current in one direction and open for current in the other direction. Rectifiers are used primarily for the conversion of alternating current (ac) to direct current (dc). *See* ELECTRONIC POWER SUPPLY.

A variety of rectifier elements are in use. The vacuum-tube rectifier can efficiently provide moderate power. Its resistance to current flow in the reverse direction is essentially infinite because the tube does not conduct when its plate is negative with respect to its cathode. In the forward direction, its resistance is small and almost constant. Gas tubes, used primarily for higher power requirements,

also have a high resistance in the reverse direction. The semiconductor rectifier has the advantage of not requiring a filament or heater supply. This type of rectifier has approximately constant forward and reverse resistances, with the forward resistance being much smaller. Mechanical rectifiers can also be used. The most common is the vibrator, but other devices are also used. *See* GAS TUBE; MECHANICAL RECTIFIER; SEMICONDUCTOR RECTIFIER.

A rectifying element can be illustrated by assuming a device having a forward resistance R_1 and a reverse resistance R_2, which is much greater than R_1. A sinusoidal alternating voltage $E_m \sin 2\pi ft$ is applied to the rectifier, where E_m is the maximum value of the applied voltage, f is the frequency of the voltage wave, and t is time. The magnitude of the current in the forward direction is $(E_m/R_1) \sin 2\pi ft$. This current flows from t equals 0 to $1/(2f)$, or for one-half the cycle of the alternating voltage wave. The forward current, averaged over one cycle, is $E_m/(\pi R_1)$. The reverse current has the magnitude $(E_m/R_2) \sin 2\pi ft$ and flows from t equals $1/(2f)$ to $1/f$, or for the other half-cycle. The reverse current, averaged over one cycle, is $E_m/(\pi R_2)$. The net forward average current is $E_m(R_2 - R_1)/(\pi R_1 R_2)$.

If reverse resistance R_2 is extremely large compared to R_1, the average current approaches $E_m/(\pi R_1)$. If the average current is subtracted from the current flowing in the rectifier, an alternating current results. This ripple current flowing through a load produces a ripple voltage which is often undesirable. Filter and regulator circuits are used to reduce it to as low a value as is required. *See* ELECTRIC FILTER; VOLTAGE REGULATOR.

Half-wave rectifier circuit. A half-wave rectifier circuit is shown in **Fig. 1**. The rectifier is a diode, which allows current to flow in the forward direction from A to B but allows practically no current to flow in the reverse direction from B to A. The ac input is applied to the primary of the transformer; secondary voltage e supplies the rectifier and load resistor R_L.

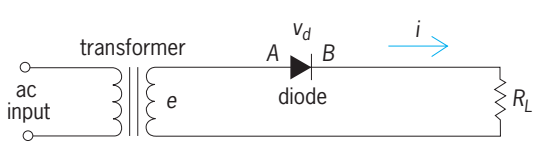

Fig. 1. Half-wave diode rectifier.

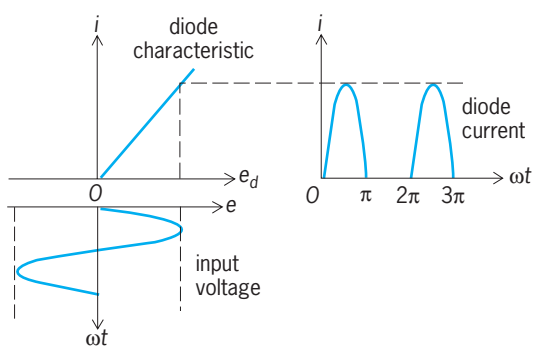

Fig. 2. Rectifying action of half-wave diode rectifier.

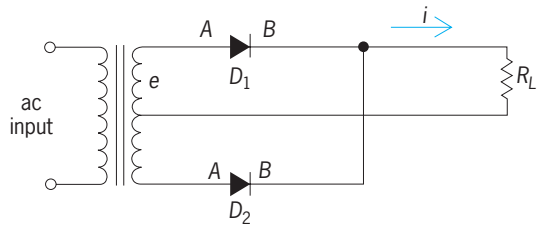

Fig. 3. Full-wave diode rectifier.

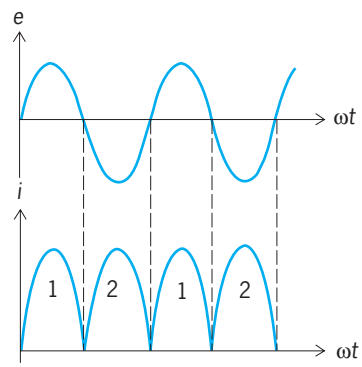

Fig. 4. Applied voltage and output current of full-wave rectifier.

The rectifying action of the diode is shown in **Fig. 2**, in which the current i of the rectifier is plotted against the voltage e_d across the diode. The applied sinusoidal voltage from the transformer secondary is shown under the voltage axis; the resulting current i flowing through the diode is shown at the right to be half-sine loops. Averaging the value of these half-sine loops gives the direct current flowing; the ripple current is the variation of load current about the average value.

Full-wave rectifier circuit. A full-wave rectifier circuit is shown in **Fig. 3**. This circuit uses two separate diodes, D_1 and D_2. The center tap of the transformer is connected through the load resistance R_L to the B sides of both diodes. During one-half cycle of the ac input the A terminal of diode D_1 is positive with respect to the B terminal, and D_1 conducts current from A to B. This current passes through the load resistor R_L in the direction shown by i. During this time diode D_2 is not conducting because terminal A is negative with respect to B. When the ac potential goes through zero, A of diode D_1 becomes negative and the diode stops conducting. The potential on A of D_2 then becomes positive, and D_2 starts to conduct. The resulting current wave shape is shown in **Fig. 4**.

The effect of using two diodes instead of one is to produce a more continuous flow of direct current through load R_L because the first diode conducts for the positive half-cycle and the second diode conducts for the negative half-cycle, as shown in Fig. 4. Comparison of Figs. 2 and 4 indicates that a full-wave circuit will produce a more nearly uniform output than will the half-wave circuit.

Polyphase rectifier circuits. When high dc power is required by an electronic circuit, a polyphase rectifier circuit may be used. It is also desirable

when expensive filters must be used. This is particularly true of power supplies for the final radio-frequency and audio-frequency stages of large radio and television transmitters. The rectifier employed in polyphase circuits generally is a gas tube that has a low voltage drop in the forward direction and thus has a high efficiency. Semiconductor diodes are also used. The number of phases used in these circuits is most often 3, but 2, 4, 6, and 12 phases are used occasionally.

The simplest polyphase circuit is the three-phase half-wave circuit of **Fig. 5**. The primaries of the transformers are connected in delta to the three-phase ac line, and the secondaries are connected in wye with the common connection going to one end of the load resistor. The other end of the load resistor is attached to the B terminals of the three rectifier diodes required in the circuit. The A terminals are connected to the separate ends of the three transformer secondaries.

Operation of the circuit is such that diode D_1, connected to the first secondary, conducts for 120° of the ac cycle. As soon as the voltage on secondary 2 equals that of secondary 1, diode D_2 starts to conduct and diode D_1 stops conducting. Secondary voltages e_{o1}, e_{o2}, and e_{o3} are shown in **Fig. 6**, and diode currents are indicated as i_1, i_2, and i_3. The resulting load current i_L is also shown in Fig. 6. This current is much closer to a true direct current than is the current for the single-phase circuits of Figs. 1 and 3. Ripple voltage is much lower, and less elaborate filter circuits are needed to smooth the output wave. The diodes in Fig. 5 could be replaced by gas rectifier tubes or ignitrons if higher load currents were required.

Another common polyphase rectifier circuit is the three-phase full-wave or six-phase half-wave circuit shown in **Fig. 7**. Here the tubes conduct for 60° instead of 120°, as in the circuit of Fig. 5. The ripple voltage for the full-wave circuit is much smaller.

Many other polyphase rectifier circuits are possible.

Bridge rectifier circuits. Bridge rectifier circuits are useful in both single-phase and polyphase applications in which a transformer must be used whose secondary has no center tap or in which dc voltages approximately equal to the total secondary voltage of the transformer must be obtained. Another use of the bridge circuit is in ac rectifier-type meters. The bridge circuit is shown in **Fig. 8**. Four separate half-wave rectifier diodes are used. When the left-hand side of the transformer secondary is positive, current flows through diode D_1, the load resistor R_L, and diode D_3. When the secondary voltage reverses, the current flows through diodes D_2 and D_4, passing through resistor R_L in the same direction as during the first half cycle.

Parallel rectifiers. If greater current is desired, two or more rectifiers can be connected in parallel. In such an arrangement small resistors or inductors are put in series with the rectifiers before they are connected in parallel.

Controlled rectifiers. Controlling the current delivered by a rectifier can be accomplished by varying

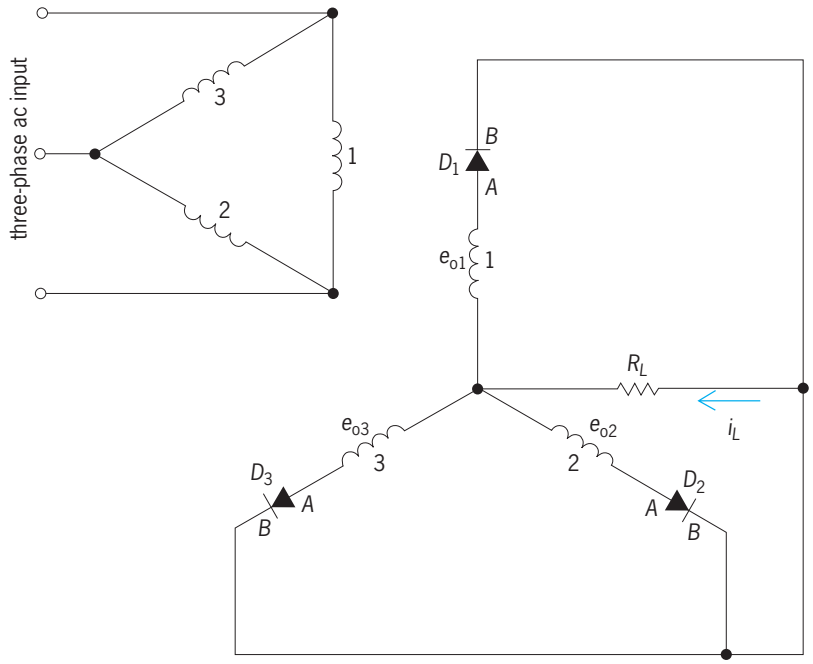

Fig. 5. Three-phase half-wave rectifier.

the primary voltage of the power transformer or by changing a resistance in series with the load resistor. The first technique has the disadvantage of being expensive; the second leads to poor efficiency. A more convenient and less expensive method is to control the angle at which the rectifier tube starts to conduct. Special gas tubes that accomplish this control are thyratrons and ignitrons. Silicon controlled rectifiers (SCR) may also be used for this purpose. Thyratrons are hot-cathode gas tubes with a large grid structure that prevents the arc from being ignited until the correct voltage is applied to the grid. An ignitron is a cold-cathode pool-type tube with an igniter grid actuated by an electrical pulse. The igniter of the ignitron

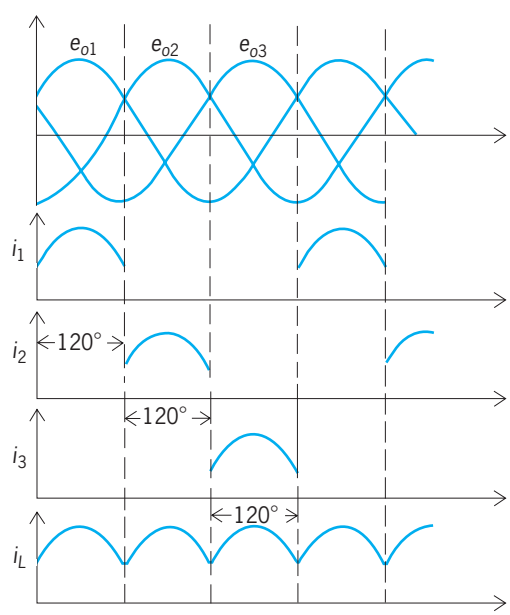

Fig. 6. Transformer voltages, diode currents, and load current in a three-phase half-wave rectifier.

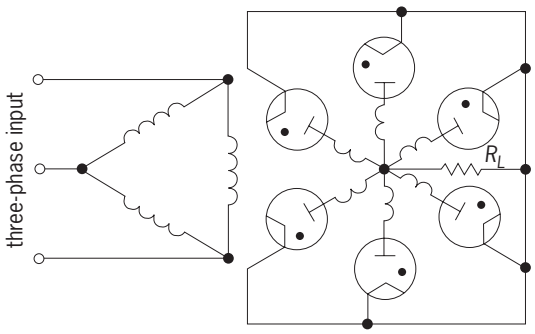

Fig. 7. Three-phase full-wave or six-phase half-wave rectifier, an example of polyphase rectifier circuit.

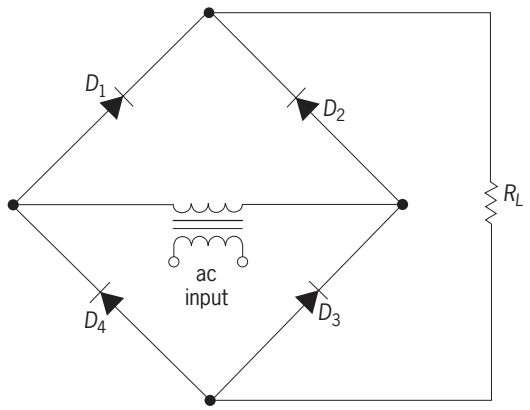

Fig. 8. Single-phase full-wave bridge circuit.

requires a substantial amount of power, which is usually supplied by an auxiliary thyratron in the control circuit. The SCR is a silicon semiconductor device which conducts when the trigger gate electrode is raised to the triggering potential. *See* CONTROLLED RECTIFIER; SEMICONDUCTOR RECTIFIER.

One circuit for an SCR is shown in **Fig. 9**, with control of the triggering point of the SCR possible over a full-half-cycle of 180 electrical degrees. When the upper terminal of the ac input is positive, capacitor C will charge to the triggering voltage of the SCR in a time determined by the RC time constant and the voltage across the SCR. Current then flows through load resistor R_L until the ac input voltage starts to reverse. When the lower terminal of the ac input is positive, capacitor C charges negatively through the diode D_2, and the cycle of conduction repeats. Diode D_1 prevents the negative peak voltage from appearing on the gate of the SCR. By varying the resistor R, the triggering point can be varied from zero degrees when the SCR will conduct the maximum current and to 180° when the SCR conducts zero current. Hence a continuous control from zero to maximum current is achieved. Other control circuits can also be used.

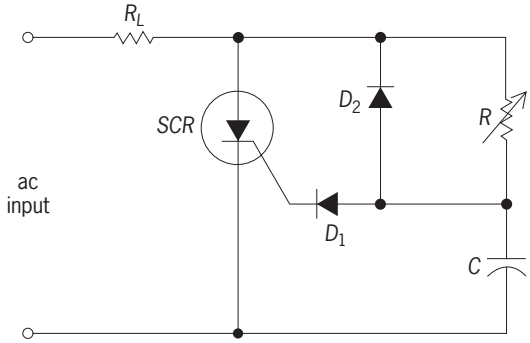

Fig. 9. One circuit for a silicon controlled rectifier with phase control triggering network.

Inverse voltage. The inverse voltage of a rectifier is the voltage that the rectifier must withstand when it is not conducting or when it is conducting slightly in the reverse direction. As an example, in the full-wave rectifier circuit of Fig. 3, when diode D_2 is conducting, diode D_1 has impressed upon it the total secondary voltage of the power transformer minus the voltage drop in D_2. For a well-designed power supply the maximum value of the inverse voltage should not exceed the rated value of the rectifier specified by the manufacturer.

Current ratings. Another important rating for a rectifier is the average current through it. The average rectifier current of a half-wave single-phase rectifier is the same as the average load current. For a full-wave single-phase rectifier the average rectifier current is one-half the average load current. The maximum value of instantaneous current through the rectifier should not exceed the peak current rating of the rectifier. This is particularly true when capacitor-input filters are employed, because these filters generally produce high peak-to-average current ratios in the rectifier. For information on electronic circuits in general *see* CIRCUIT (ELECTRONICS)

Donald L. Waidelich

Bibliography. T. H. Barton, *Rectifiers, Cycloconverters, and AC Controllers*, 1994; A. A. Jaecklin (ed.), *Power Semiconductor Devices and Circuits*, 1993; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Rectilinear motion

Motion is defined as continuous change of position of a body. If the body moves so that every particle of the body follows a straight-line path, then the motion of the body is said to be rectilinear.

When a body moves from one position to another, the effect may be described in terms of motion of the center of mass of the body from a point A to a point B (see **illustration**). If the center of mass of the body moves along a straight line connecting the points A and B, then the motion of the center of mass of the body is rectilinear. If the body as a whole does not rotate while it is moving, then the path of every particle of which the body is composed is a straight line parallel to or coinciding with the path of the center of mass, and the body as a whole executes rectilinear motion. This is shown by the straight line connecting points P_1 and P_2 in the illustration. *See* CENTER OF MASS.

Rectilinear motion is an idealized form of motion which rarely, if ever, occurs in actual experience, but

it is the simplest imaginable type of motion and thus forms the basis for the analysis of more complicated motions. However, many actual motions are approximately rectilinear and may be treated as such without appreciable error. For example, a ball thrown directly upward may follow, for all practical purposes,

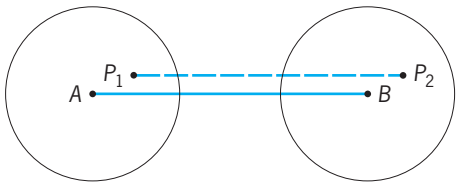

Rectilinear motion. All points move parallel to the center of mass.

a straight-line path. The motion of a high-speed rifle bullet fired horizontally may be essentially rectilinear for a short length of path, even though in its larger aspects the ideal path is a parabola. The motion of an automobile traveling over a straight section of roadway is essentially rectilinear if minor variations of path are neglected. The motion of a single wheel of the car is not rectilinear, although the motion of the center of mass of the wheel may be essentially so. *See* BALLISTICS; MOTION. Rogers D. Rusk

Recursive function

A function that maps natural numbers to natural numbers and is special in that it must be calculable by using a precisely specified algorithm. The mathematical definitions of partial recursive functions and recursive functions were developed to give a precise mathematical characterization of those functions or operations on the natural numbers which are computable by using effective procedures.

Effective procedure. An effective procedure is a procedure or process determined by a finite list of precise instructions. There is no upper bound on the number of instructions in this list. These instructions can be carried out in a discrete one-step-at-a-time fashion. The only equipment needed in executing them are pencil and paper, and an unlimited supply of both is assumed to be available. No creativity is involved in applying these instructions, and no random devices such as flipping a coin are allowed in carrying them out. The instructions are very specific about all matters, in particular, as to when to terminate these instructions and when to give various outputs. Consequently, each time that these instructions are applied under the same circumstances, the calculation or computation proceeds in exactly the same way with exactly the same outcome. The computation may never end and, consequently, may go on forever. Even when the computation does terminate, there is no predetermined bound on the time required to carry it out.

While a computer program can be viewed as an effective procedure, in general, more than an ordinary computer would be needed to implement the computation specified by a program in the sense of effective procedure, since there would be no restrictions on the size of the program or on the available memory and running time. Effective procedures give a theoretical upper bound on the kind of computations that a computer or person could ever actually carry out.

Effectively computable function. An effective procedure may be used to calculate a function as follows: The calculator inputs to the effective procedure a natural number (0, 1, 2, 3, . . .) and allows the procedure to be carried out on this input. If an output is obtained, the function is said to be defined at this input and its value at this input is the output so obtained. Any function which can be computed in this way is called effectively computable. For example, addition and multiplication of natural numbers are effectively computable functions.

Turing machines. The notion of effective procedure is intuitive and vague. In the setting of an idealized computing device called a Turing machine, a more precise definition of certain effectively computable functions can be given. A Turing machine is a very simple machine which operates on a paper tape divided into blocks. Each block can contain either a B (for blank) or a 1. The paper tape is potentially infinitely long in both directions, and runs through the Turing machine, which can view or scan only a single block at a time (see **illus.**). The machine can

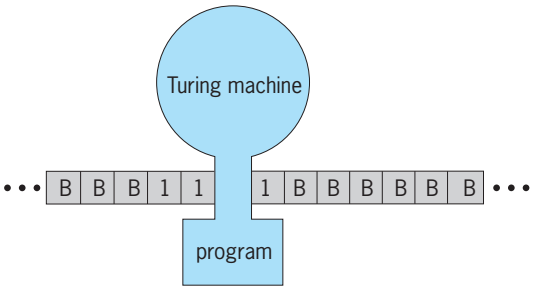

Turing machine scanning a single block on its tape.

perform only four simple actions on the tape: move the tape one block to the left (so that after this action it is scanning the block immediately to the right of the one previously scanned), move the tape one block to the right (with the corresponding effect), print a B (erase a 1), or print a 1 on the block currently being scanned. (These actions are labeled L, R, B, and 1 respectively.)

A program for a Turing machine is specified as follows: A finite set of states is used, denoted by the symbols $s_0, \ldots, s_k$. Any instruction for the machine is a quadruple of the form $(s_i; a; b; s_j)$, where $0 \leq i, j \leq k$; a is B or 1; and b is B, 1, L, or R. This instruction is executed as follows: If the machine is in state s_i and is scanning a block with the symbol a on it, it does the action specified by b and enters the

state s_j. A program P for a Turing machine is any finite set of quadruples of the above form such that there are never two different instructions in P applicable at the same time. *See* AUTOMATA THEORY.

Partial recursive functions. A Turing machine with such a program can be used to compute a unary function on certain natural numbers as follows: A number n is input by printing $n + 1$ consecutive symbols 1 on the tape (B's every place else), the tape is arranged so that the left-most 1 is the block being scanned by the machine, and the machine is started in initial state s_0. The computation proceeds according to the quadruples in its program and terminates when the machine encounters a situation for which no instruction is applicable. If the computation terminates, the output or value of the function is taken to be the total number of 1's appearing on the tape at that time. If the computation never terminates, the function is considered undefined at that input n. A k-ary function can be computed by using a Turing machine with the same kind of program by varying the input; for example, an input for a binary function consists of two strings of 1's separated by a B.

Any function that can be calculated in this way is called a partial recursive function. A recursive function (also called a general recursive function or a total recursive function) is a partial recursive function that is defined at all inputs.

Church's thesis. Besides the Turing-machine approach, many alternatives have been proposed for a more precise mathematical definition for the intuitive notion of effectively computable function, but they have all led to the class of partial recursive functions defined above. Also, all known examples of effectively computable functions on the natural numbers have been shown to be partial recursive functions. This gives evidence for Church's thesis, a widely accepted principle that any effectively computable function on the natural numbers must be partial recursive function. Church's thesis cannot be proven because it asserts that the intuitive notion of effectively computable function is precisely captured by the mathematical definition of partial recursive function; that is, that this is a good definition. Many results in mathematical logic take on a greater significance by assuming Church's thesis. For example, proving that a certain function is not recursive shows that an effective procedure could never be found for computing it. This makes it possible to prove that certain mathematical problems, whose solution would require that certain functions be effectively computable, do not have a solution or are unsolvable. *See* GÖDEL'S THEOREM; LOGIC.

George C. Nelson

Bibliography. J. Barwise (ed.), *Handbook of Mathematical Logic*, 1977; M. Davis, *Computability and Unsolvability*, 1982; H. B. Enderton, *A Mathematical Introduction to Logic*, 2d ed., 2000; S. C. Kleene, *Introduction to Metamathematics*, 1971; H. Rogers Jr., *Theory of Recursive Functions and Effective Computability*, 1987.

Recycling technology

Methods for reducing solid waste by reusing discarded materials. Essentially, recycling is a three-part process that uses waste materials to make new products. The three integral phases involve the collection of recyclable materials, manufacture or reprocessing of recyclable materials into new products, and purchase of recycled-content products. Various techniques have been developed to recycle plastics, glass, metals, paper, and wood.

Plastics

Plastic discards represent an estimated 9–10% by weight and up to 26% by volume of the municipal solid waste in the United States discarded after materials recovery. About 2% by weight of discarded plastics is recovered. Most of the effort to recycle postconsumer plastic involves high-density polyethylene (HDPE) and polyethylene terephthalate (PET) resins. Plastic waste made of these two resins is more easily identified and separated from other plastic waste as a result of industry standardization of materials for soft drink bottles (polyethylene terephthalate), base cups from these soft drink bottles (high-density polyethylene), and milk bottles (high-density polyethylene). Perhaps 26% of all plastic containers are recovered.

Plastic available for recycling. Approximately half of plastic waste consists of single-use convenience packaging and containers. Many manufacturers prefer plastic for packaging because it is lightweight, resists breakage and environmental deterioration, and can be processed to suit specific needs. Once plastics are discarded, these attractive physical properties become detriments.

Six resins are used for most commodity plastic products: low-density polyethylene (LDPE), high-density polyethylene, polyethylene terephthalate, polyvinyl chloride (PVC), polypropylene (PP), and polystyrene (PS). These are single-resin thermoplastics and are considered suitable for recycling because they become pliable when heated and can be remolded. Acrylonitrite butadiene styrene (ABS), also a thermoplastic, is used to make many computer casings. *See* POLYOLEFIN RESINS; POLYSTYRENE RESIN; POLYVINYL RESINS.

Plastics are also used with paper and metal as multicomponent films, mostly for packaging. They are more difficult to recycle than single-resin plastic material discards, and can be converted only to mixed-plastic products.

Recycled plastic products. Secondary materials are products recovered from the waste stream and reformulated into new physical forms serving end uses other than those of the original materials. A variety of secondary materials are produced from plastic waste. The outer plastic layer of disposable diapers has been recycled into garbage bags, flowerpots, and plastic lumber. Polyethylene terephthalate is recycled into scouring pads, fiber fill for jackets, paintbrushes, carpet fibers, and other products. Polyethylene terephthalate is also recycled to produce polyol for use in urethane foam and furniture, and

unsaturated polyester for boat hulls, pools, auto body parts, and appliance parts. High-density polyethylene is recycled into base cups for soft drink bottles, flowerpots, toys, and pallets. Projects to recycle bottles made from polyvinyl chloride resins have focused on products such as bottles for shampoo and vegetable oil.

Recycling postconsumer polystyrene foam has been targeted in certain pilot projects by industry for reuse as coat hangers, building insulation, office accessories, trash receptacles, and flowerpots. As much as 95% of clean preconsumer industrial plastic waste (plastic scrap from an industrial or manufacturing process) is recycled. Companies have also recycled unclean industrial plastic waste for use in inner layers, with virgin resins as outer layers, for products such as multilayer detergent bottles; however, the actual amount of unclean plastic industrial scrap that is recycled is unknown.

Research in recycling of plastics has expanded to include the development of lumber from a mix of all types of postconsumer plastic. Such a mix, commonly called commingled plastics, comprises various plastic resins, pigments, additives used in manufacturing, and nonplastic contaminants. Plastic lumber has a number of advantages over wood, such as resistance to rot, chemicals, water, and insects. Plastic lumber is being used for fence posts, poles, marine pilings and bulkheading, dock surfaces, park benches, landscape timbers, retaining walls, palettes, and parking space bumpers. Engineering tests have shown that it is better to make plastic lumber from a single polymer than from a commingled plastic.

Methods and technologies. Postconsumer plastic may be collected in a curbside recycling program or at designated drop-off centers. There are a number of problems associated with the collection of plastic waste: its high volume can burden an existing collection program; the material does not crush easily; and plastics made from different resins may be mixed together.

To assist the public and laborers working in material recovery facilities with identifying different plastics, the Society of the Plastics Industry sponsors a voluntary coding system. However, in the United States it is required by law in 39 states. The coding system consists of a triangular arrow stamp with a number in the center and letters underneath to identify the resin used in the container.

The collected plastic waste is usually separated manually from the waste stream, and often it is cleaned to remove adhesives or other contaminants. It is sorted further, based on different resins. Mechanical techniques to separate plastics are available, and they can be used to sort plastics based on unique physical or chemical properties. It is anticipated that more sophisticated methods under development will provide better systems for separating plastics made of different resins.

The technologies used to manufacture recycled plastic are virtually the same as those used to manufacture products from virgin plastic resins. The recycled products are melted; for some products, additives or virgin resins are used to improve the properties. The plastic is then extruded into specific products or pellets.

For secondary lumber-type materials made from mixed plastic waste, the technology generally used is known as Extruder Technology 1 (ET/1). ET/1 has three main components: an extruder, a molding unit, and an extraction unit. The extruder is used for melting the plastic at temperatures of 360–400°F (182–204°C). Plastic with higher melt temperatures, such as polyethylene terephthalate, become encapsulated within the melted plastic. The molding unit consists of linear molds that are water cooled. The product shrinks within the mold during cooling. It is possible in some units to make lengths that exceed those of typical wooden lumber. It is also possible to chemically break down the plastic polymers into monomers or oligomers. These are cleaned and can then be made back into the polymer. This process perhaps offers the best opportunity for reusing the plastic for the same product as was originally made from the virgin material.

Engineering and environmental issues. A significant problem in plastic recycling is the presence of contaminants such as dirt, glass, metals, chemicals from previous use, toxicants from metallic-based pigments, and other nonplastic materials that are part of or have adhered to the plastic products. The U.S. Food and Drug Administration (FDA) has expressed concern over potential risks in using recycled plastic products in contact with food. However, a redesigned chemical recycling process has been used to recycle PET into soft-drink and salad-dressing bottles. Other constraints involve inconsistencies in the amount of different plastic resins in commingled plastic wastes used for recycling, and engineering aspects of recycled plastic products, such as lessened chemical and impact resistance, strength, and stiffness, and the need for additional chemicals to counteract other types of degradation for reprocessing. There may be limitations to the number of times a particular plastic product can be effectively recycled as compared to steel, glass, or aluminum, which can be recycled many times with no loss of their properties and virtually no contamination.

The long-term engineering properties of recycled products are still largely unknown. The many products made from recycled plastic include gimmick toys, penholders, and paper trays; these may appear in the waste stream in a short time.

The recycling of plastics creates a waste stream of its own—contaminated wastewater and air emissions. Many additives used in processing and manufacturing plastics, such as colorants, flame retardants, lubricants, and ultraviolet stabilizers, are toxic and may be present in the waste stream.

Economic and social acceptability. Markets for recycled single-resin plastics have had more success than mixed-plastic products. The reason may be that bottle return laws lead to better collection of single-resin polyethylene terephthalate beverage containers and to focusing by the plastics industry on

recycling technologies for single-resin plastics. Marketing problems that arise with products made from commingled plastics include inconsistency in feedstock, lack of suitable engineering specifications, and unpredictable performance.

An important issue related to the marketability of recycled products is whether such products are, in fact, cheaper than the products they will be replacing. Plastic lumber, for example, is expensive compared to wooden lumber, concrete, and other materials for which it can be substituted. However, product lifetime must also be taken into consideration with the initial cost of a material. The average lifetime of plastic lumber and other products made from commingled plastics is undetermined. Other factors, such as plastic lumber's low maintenance costs, are also significant considerations.

For products made from commingled plastics, marketing problems due to a reduced esthetic appeal may arise: there is a lack of black and very dark colors, and impurities such as bits of paper or metal can be seen within the material.

A notable marketing advantage for recycled plastic products over items made from virgin plastics or other material is that they are appealing from an environmental standpoint to many consumers.

Casings for computers, often acrylonitrile butadiene styrene, are a growing part of the waste stream. These casings are made of an engineered thermoplastic and therefore can be recycled. Unfortunately, few of the component parts of computers are recycled, largely because of the cost of disassembling. Recycling plastics is an essential part of a national waste prevention and waste reduction strategy. Recycling plastic that may otherwise be diverted to landfills, incinerators, or roadside litter is a positive step in reducing the waste stream and should be incorporated in all programs for solid-waste management. Research, engineering specifications, government regulations, and continued interest are necessary for future success of products made from recycled postconsumer plastic.

R. L. Swanson; Vincent T. Breslin; Marci L. Bortman

Glass

Recycling of glass is an important aspect of solid-waste management. Glass containers are a usual ingredient in community recycling programs; they are 100% recyclable and can be recycled indefinitely. Moreover, the recycling process creates no additional waste or by-products.

In 1993 in the United States, glass containers, which constitute 6% of the solid-waste stream by weight, were recycled at a rate of 35%. This means that nearly one-third of the glass containers available for consumption in the United States were cycled back into glass containers and other useful items such as glasphalt, or were returned as refillable bottles.

Collection methods. There are three primary methods of collecting glass: drop-off centers, buy-back centers, and curbside collection. Drop-off centers and buy-back centers typically collect less material, but they also have much lower capital and operating costs than curbside programs.

Drop-off centers are one of the simplest forms of glass recovery. These may be mobile collection stations or permanent sites maintained by local municipalities. In some areas, fiberglass domes known as igloos are used as around-the-clock collection points. In other places, a drop-off may be as simple as several 55-gal (208-liter) metal drums, each marked for the color of glass it is to hold.

Buy-back centers purchase glass and other recyclables from people who voluntarily transport the material to the site. The centers serve as a convenient market for recyclers who cannot sell directly to a glass plant.

Comprehensive curbside collection programs are increasingly popular. Curbside collection programs allow residents the opportunity to put their recyclable materials out at the curb, separated from their regular trash, for pick-up. Many programs offer commingled collection, which affords a resident the opportunity to put all recyclable material into one container or bag. Glass is an integral part of these programs. Often aluminum and glass provide the necessary revenue to sustain the community's program.

Processing specifics. The process of recycling glass is relatively straightforward. Used glass bottles and jars are mixed with silica sand, soda ash, and limestone in a melting furnace at temperatures up to 2800°F (1540°C). The molten glass is poured into a forming machine, where it is blown or pressed into shape. The new containers are gradually cooled, inspected, and shipped to the customer. Before glass can be recycled, however, it must be furnace-ready, that is, sorted by color and free of contaminants.

Color sorting. Sorting of green, amber, and flint (clear) glass containers is essential in order to maintain color consistency during the manufacturing process. Darker brown and green containers are used in many cases to help preserve the shelf life of a food product.

Different chemical agents are used to color glass. For example, chromium salts are generally used to create a green color, selenium salts are used to produce clear bottle glass, and iron oxides impart an amber finish. The coloring properties of these chemicals are so strong that a small amount of miscellaneous colors in a glass batch mix will adversely affect the result. Thus it is critical to deliver color-separated glass to manufacturing facilities in order to avoid costly problems involving equipment and production.

A considerable amount of research is focused on developing automated color-sorting equipment that can be used at glass processing and manufacturing plants. In addition, there is experimentation with the staining of flint glass bottles. The coating would be burned off in the furnace, making it unnecessary to separate by color.

Beneficiation plants. Cullet must meet a standard of quality similar to that of the raw material it replaces. Contamination from foreign material will result in the cullet being rejected by the plants, as it poses

a serious threat to the integrity and purity of the glass packaging being produced. Contaminants include metal caps, lids, stones, dirt, and ceramics. Paper labels do not need to be removed for recycling, as they burn off at high furnace temperatures.

Many glass manufacturers have invested in glass-processing equipment. These beneficiation units remove metals and other contaminants. For example, the New York State Energy Research and Development Authority, with funding and technical assistance support from several private entities including the Glass Packaging Institute, introduced a new glass beneficiation system that can optically sort out ceramics and separate nonferrous metals from the cullet. A nonferrous metal separator has been developed. This system removes aluminum, lead, and small contaminants, including the rings of residue in bottlenecks. Following this process, the cullet is moved to the optical ceramic sorter, which removes unwanted ceramics by releasing quick jets of air activated by an infrared light that detects the ceramic's opacity. Several other efforts are going on across the country to develop improved methods of processing furnace-ready cullet.

Energy and environment benefits. From a manufacturing standpoint, cullet can reduce wear and tear on furnaces, and maintenance expenses, since it can be melted at a lower temperature than that required to combine virgin materials. This reduced melting point can save as much as 25% of the energy used to make virgin glass, depending on what type of heating is used in the furnace. The energy saved by recycling a single glass bottle is estimated to be the equivalent of lighting a 100-W bulb for 4 h. Similarly, for every soft-drink bottle recycled, enough energy is saved to run a television set for $1^1/_2$ h.

Glass recycling also has environmental benefits, such as reductions in emissions to the atmosphere. It has been estimated that 27.8 lb of air pollution is produced for every ton of new glass produced (13.9 kg/metric ton), and recycling glass reduces this source off pollution by 14–20%. Recycling a ton of glass saves 1330 lb (599 kg) of sand, 433 lb (195 kg) of soda ash, 433 lb (195 kg) of limestone, and 151 lb (68 kg) of feldspar. Generally costs for cullet are less than for soda ash, which it replaces at a ratio of 3.5 to 1.

End markets. The primary recycling market for container cullet consists of the manufacturers of glass containers. However, over the years difficulties involving color sorting and transportation have spawned new applications for cullet. Thus cullet use in secondary materials has increased. There markets include fiberglass, glasphalt, roadbed, reflective beads, decorative glass, and drainage. One of the most visible secondary markets is manufacture of glasphalt, which is a type of asphalt mix that incorporates crushed glass. Glasphalt was developed as an application of commingled cullet. One useful characteristic of glasphalt is its ability to retain heat. Glass acts as an insulating agent, making glasphalt workable under much colder environments. Manufacture of fiberglass, predominantly used in the form of glass wool for thermal and acoustical insulation, is another common secondary market for cullet.

Nathan Tyler; Natalie U. Roy

Metals

A modern society depends on a tremendous quantity of metals to provide for transportation, housing, communications, and other essentials of a healthy, comfortable existence. To meet these needs with minimal environmental damage, metals must be recycled to alleviate the need to mine more ore, to reduce energy consumption, to limit the dissemination of metals into the environment, and to reduce the cost of metals. In the United States a substantial portion of these needs are met by recycling metals.

The extensive recycling of materials is very important for three reasons: (1) The energy required to recycle a metal is considerably less in comparison to producing it from ore. (2) Extracting the metal from ore produces a tremendous amount of waste material. For instance, to produce 1 ton (0.9 metric ton) of copper metal requires the moving of 350 tons (315 metric tons) of overburden (the waste material above the ore body); 150 tons (135 metric tons) of ore, of which 145 tons (131 metric tons) ends in tailings (the waste material left after the metal is removed from the ore); plus using 120,000 gal (450,000 liters) of water and consuming 6 tons (5 metric tons) of air. In addition, 2.7 tons (2.4 metric tons) of sulfuric acid and 1.8 tons (1.6 metric tons) of slag are produced. The United States produces about 1.5×10^6 tons (1.4×10^6 metric tons) of copper from ore each year, resulting in the generation of large waste streams. (3) Metals that are not recycled eventually begin to be dissipated throughout the environment. Since many metals are toxic, this can result in the pollution of water and soil. Lead and cadmium are examples of metals that are considered toxic. Lead was used in paints and as a fuel additive and thus was widely disseminated throughout the environment. With the recognition of its toxic properties, it was virtually eliminated as a fuel additive and its use greatly reduced in paints. Cadmium is used in rechargeable batteries, and these batteries have become much more widely used. However, methods to recover cadmium from these batteries has not kept pace and, as a consequence, landfills are being contaminated with these batteries. *See* HAZARDOUS WASTE.

Challenges. While it is beneficial to recycle materials, there are at least two important problems that hinder recycling: collection and impurity buildup. When a metal becomes scrap and is a candidate for recycling, it must be collected at a cost that makes it attractive to recyclers. It is useful to divide scrap into three categories—home scrap, new or prompt scrap, and old or obsolete scrap—whose methods for collection differ significantly. Home scrap is waste produced during fabrication, and includes casting waste (for example, risers), shearings and trimmings, and material that has been rejected. This scrap is usually recycled within the plant, and therefore it is not recognized as recycled material in recycling

statistics. New or prompt scrap is waste generated by the user of semifinished material, that is, scrap from machining operations (such as turnings or borings), trimmings, and rejected material. This material is collected and sold to recyclers and, if properly labeled and segregated, it is easy to recycle and is valuable. Old or obsolete scrap is waste derived from products that have completed their life cycle, such as used beverage cans, old automobiles, and defunct batteries. The collection and the impurity buildup problems are most severe when considering old or obsolete scrap. The recycling of automobiles will be used to demonstrates these problems.

Automotive recycling. The recycling of automobiles illustrates the problems associated with recycling old or obsolete scrap. The first step in the recycle process is carried out by the dismantler. The dismantler's role is to remove the high-value parts, plus engine blocks and fuel tanks that have economic value but also must be removed for safety reasons. The auto hulk is then compressed into a flattened body. Next, the shredder takes the flattened bodies and tears them into fist-sized pieces by using hammer mills. By using techniques such as magnetic separation (to remove the ferrous material), air streams, or hand sorting, the ferrous metals, nonferrous metals (primarily aluminum), and automotive shredder residue are separated. The automotive shredder residue, which consists of polymers, glass, fibers, dirt, and other materials, goes primarily to landfills, since it is not economic to recycle. The ferrous scrap goes to minimills to make low-carbon steels used in fencing, reinforcing bars, and other items; and to integrated steel producers to be used in basic oxygen furnaces, where it is mixed with pig iron to make various grades of steel.

The nonferrous metal scrap is then separated into primarily three groups; aluminum, copper, and zinc. The lead has been separated previously by the dismantler by taking out the batteries, of which 97% are recycled. The aluminum generally goes to the secondary aluminum market, where it is used to produce aluminum foundry alloys for use in the auto industry. Copper must be separated from ferrous material, because it severely degrades the properties of steel. Zinc is becoming more of a problem because the steel sheet used in the body of the car is being coated with zinc to reduce rusting.

Major new problems have developed in auto recycling. The desire to produce autos of lighter weight for greater fuel economy has resulted in increased use of high-strength alloy steels, aluminum, and polymers. Little recycling of polymers occurs, and therefore more automotive shredder residue is produced that must be deposited in landfills. The ferrous materials retain some of the alloying elements plus some contamination from copper used in wiring, making the recycled steel lower in value. In the future it will be necessary to make automobiles easier to disassemble if their recycling is to remain economically viable.

Problems have developed because of the undesirable alloying elements in the recycled material. There are hundreds of different alloys of steel and aluminum, each with very desirable properties for certain applications, but disastrous in other applications. For instance, small amounts of lead added to steel (free-machining steel) make it much easier to machine, but this same lead can make steel susceptible to unexpected brittle fracture when stressed at elevated temperatures. With more and different materials being mixed together in the waste streams, it is ever more likely that impurity contamination will be a problem.

Iron/steel. The tonnage of iron and steel used in the United States dwarfs all other metals combined, and therefore it is by far the most important to recycle.

A major change in the production of steel began in the 1960s, with the basic oxygen furnace and the electric-arc furnace replacing open-hearth steel making. About 70% of steel is made by the basic oxygen furnace process, and 30% by electric-arc furnaces. The principal feedstock for electric-arc furnaces is recycled steel, with a significant portion being recycled automobiles. Some scrap metal is also used in the basic oxygen furnace, but it is significantly less in comparison to that used in the old open-hearth method. Therefore, the two processes are complementary in balancing the use of virgin and recycled metal.

Before 1965, the electric-arc melting furnace was used primarily to produce specialty steels (such as stainless or tool), but it is now a major process for producing low-carbon steel for wire and rod for making nails, netting and mesh, tire cord, spring wire, fencing, and re-bar in minimills. *See* ARC HEATING; IRON METALLURGY; STEEL MANUFACTURE.

Aluminum. Recycling of aluminum is done by independent secondary aluminum smelters in addition to the large integrated aluminum producers. The secondary smelters use gas- or oil-fired reverberatory furnaces of 30,000–100,000 lb (13,500–45,000 kg) capacity to remelt and refine scrap aluminum, with much of the product going into die castings for the automotive and appliance industries. Old scrap represents about 60% of the 2.7×10^6 tons (2.4×10^6 metric tons) of aluminum produced from scrap, and of this about 50% is used in beverage cans. *See* ALUMINUM.

Copper. Since copper is a relatively expensive material, it is recycled extensively. However, it is used in products that have a long lifetime, for example, electrical machinery (30-year life). Consequently, less than half of the production of copper comes from scrap, of which slightly more than 50% is new scrap. *See* COPPER METALLURGY.

Lead. The uses of lead have changed dramatically. Before 1972, lead was used extensively in gasoline additives, pigments, ammunition, and chemicals; however, with the concern about lead toxicity these uses were greatly reduced. Now the primary use of lead is in batteries, and very effective means of recycling have been established (primarily the buy-back of used batteries by retailers). Used lead batteries are sent to a secondary lead smelter, where they are

shredded and material is produced that is sold back to the battery makers.

Zinc. About 75% of the zinc produced is used in the metal form, with the remaining 25% used to produce chemical compounds that have many applications in agriculture and manufacturing. The compounds are generally not recycled; about 90% of the elemental metal is used in the galvanizing of steel and its alloys. Consequently, only about 10% of the production of zinc is from old scrap. Many techniques are used to recycle zinc, with the recycling of zinc-coated steel from automobiles growing in importance. *See* ZINC. Dale F. Stein

Paper

Paper and paperboard used for making recycled paper products come from a variety of sources, including offices, retail businesses, converters, printers, and households. The accepted term for such is "recovered paper." "Wastepaper" is applied to paper stock that is not recovered for recycling and is disposed of in landfills or incinerated. Recovered paper products that have been distributed, purchased, and in some way have served their intended primary purpose are classified as postconsumer paper. Examples are used newspapers and magazines, corrugeted containers, writing paper, and copy paper. Recovered paper from paper manufacturing (mill broke) and converting operations (trimmings), which is of excellent quality and is often used to make recycled paper, is not considered postconsumer paper. It is sometimes referred to as preconsumer paper. "Recycled paper" is the paper made, at least in part, from recovered paper. Recycled paper often contains a combination of postconsumer paper fiber, preconsumer paper fiber, and virgin fiber. The percentge of recycled fiber (and often the percentage of postconsumer fiber) is specified for the product.

Recovered paper grades. There are over 50 grades of recovered paper and several broad classifications. Five categories that are often used throughout the industry are (1) old corrugated containers, which account for nearly half of the paper and paperboard recovered in the United States and include used corrugated shipping packaging; (2) old newspapers, the second most recovered grade of paper; (3) mixed papers (a catch-all category), including old mail, magazines, office paper, and computer paper; (4) high-grade deinking, which are mostly printed and unprinted white papers collected from converting operations, printing plants, and offices; and (5) pulp substitutes, which are a general classification of high-quality recovered paper grades that can be used directly in paper manufacture without cleaning, bleaching, or deinking. Over 95% of pulp substitutes have been recycled for many years because they are a clean, high-quality material.

Each category comprises different types of papermaking fibers that are useful for manufacturing different types of recycled paper products. The efficiencies in the collection and transportation to recycling mills, as well as advances in fiber processing technology, also have an impact on the economies of using one grade over another. Although not all grades of paper can be recycled because of their components or prior use (for example, medical papers or tissues), recycling will increase until a practical limit is reached. This limit is considered to be 50–60%. In 2000, the paper recovery rate (total paper recovered for recycling as a percentage of paper consumed) was 48%.

Recovered paper fibers are used in the manufacture of many recycled paper products such as paperboard, corrugated containers, tissue products, newspapers, and printing and writing papers. They can also be used in other products such as insulation, packing materials, egg cartons, and flowerpots. *See* PRINTING.

Processing. Before paper can be recovered and used again, it must be collected, sorted, and cleaned. The first step in recycling, collection occurs in curbside programs, in drop-off centers, in paper drives, and more recently in commercial collection programs run in conjunction with waste (garbage) collection for landfill or incineration. Many factors have an impact on collection programs, including legislation, landfill availability, proximity to recycling mills, and levels of community commitment and participation.

Reprocessing begins by sorting recovered papers by grade and level of cleanliness. This can be done at the source, during collection, or by a broker. At the recycling plant, the recovered paper is processed to remove contaminants and produce a pulp that is sufficiently clean and strong to be used for papermaking. An important difference between processing of virgin pulp and recovered pulp is that recovered paper requires the removal of a much wider variety and often a much higher level of contaminants. Many of these materials are inherent in the production, converting, and end use of paper products. Some common contaminants are inks, plastic, polystyrene foam, metals, glass, dirt or sand, and adhesives. There are also a number of nonfibrous materials used in making paper, such as fillers (clay, calcium carbonate, titanium dioxide), starch, latex, and resin, that may need to be separated from the fibers. Some contaminants have an adverse affect not only on the quality of the paper being made but also on the papermaking process. If they are not removed in the recycling process, contaminants such as adhesives, resins, waxes, hot melts, and ink binders can build up on the papermaking equipment as sticky deposits, which require expensive downtime for cleanup. In general, the amount of contaminants removed depends on the type present, the grade of recovered paper being used, and the requirements of the paper being made. For example, grades such as boxboard and corrugated containers are less demanding and require a simpler cleaning process than printing and writing grades.

The basic operations used at a recycling mill to remove contaminants include (1) pulping to break down the recovered paper to a fibrous state and begin separation of contaminants from the fibers; (2) screening to mechanically separate large

contaminants from the pulp; (3) centrifugal cleaning to separate by specific gravity the contaminants that are heavier and lighter than the fibers; (4) flotation to separate and remove ink particles; and (5) washing to remove very small particles not eliminated in the previous steps.

In the pulping step, the recovered paper (usually in the form of bales) is mixed with a large quantity of water in a slusher or pulper to produce a fiber-and-water slurry called pulp. Chemicals may be added to the pulper to decolorize the paper and separate the contaminants from the fibers for removal in later stages. After the paper has been sufficiently agitated, the pulp is withdrawn from the bottom of the pulper through a screen plate to remove large contaminants (greater than about 5 mm or 0.2 in.).

The pulp is then diluted with water and pumped through a screening operation that separates contaminants that differ in size and shape from the fibers. Both coarse screens with holes (often 1.6 mm or 0.062 in. in diameter) and fine screens with slots (0.25–0.10 mm or 0.01–0.004 in. in width) may be used. Flakes, plastics, hot melts, labels, and films are removed in screening.

Screening is usually followed by centrifugal cleaning which subjects the pulp to a vortex in a tapered cone and separates particles that differ in specific gravity from the fibers. Specially designed cleaners are used to separate high-specific-gravity materials such as sand, metal, and fillers, and low-specific-gravity materials such as Styrofoam™, waxes, polyethylene, and some plastics.

If the pulp requires deinking, a flotation step is included in the recycling process. There are various designs of flotation cells, but the basic process involves injecting air into the pulp. Ink particles (and some small adhesive particles) attach to the air bubbles and float to the surface, where they are removed as an inky foam.

The smallest particles, often 30 micrometers (0.001 in.) or less, are dislodged from the pulp in a subsequent washing stage. Various systems are used that dilute the pulp with water, loosen the particles, and wash them from the pulp by gravity, vacuum, or pressure.

Multiple screening, cleaning, and flotation stages may be used to ensure an adequate level of contaminant removal. If materials are still present that cause objectionable specks or operational problems, some mills add a dispersion step. By using mechanical energy or heat and pressure, dispersion reduces these contaminants to a minute size, minimizing their negative effects. If the recycled pulp is to be used to make white paper, a bleaching or color-stripping step may be added to brighten the pulp and remove color.

Trends and challenges. Continuing issues include the management and control of fiber from recovered paper, the need for expanded recycling programs (supply), and environmental considerations.

As with virgin pulp, the purpose of managing recovered fiber is to ensure that the right mix of fibers, at an economical cost, is available for the paper products being made. A mill making paper from virgin pulp will control the species of trees and the type of market pulp it uses, while a recycled paper mill has to control the grades of recovered paper (and the associated level and type of contaminants) it accepts. The need by a mill for a consistent supply of uniform-quality recovered paper grades requires that paper procurement practices be carefully coordinated. This may involve working closely with community curbside collection programs as well as with the many brokers and dealers who serve as intermediaries for buying and selling recovered paper.

The amount of recovered paper used is steadily increasing. Currently, about 81% of the paper recovered in the United States is recycled into new products in the nation's paper mills. About 16% is exported, and the rest is used domestically in nonpaper products such as cellulose insulation, animal bedding, and mulch. Requirements by the federal government regarding the procurement of paper with a minimum recycled fiber content have helped spur the production and use of recycled paper. With this increased consumption comes the challenge of continued supply. Already certain grades of recovered paper, such as old corrugated containers and old newspapers, are approaching their practical recovery limit. This means that more efficient collection systems have to be developed for these grades, and underutilized grades and sources have to be used. These underutilized grades often have a higher level of contaminants and contain materials that are difficult to handle. A notable example is the pressure-sensitive adhesives on labels and stamps. Unless removed in the recycling process, these materials can build up on papermaking equipment as sticky deposits. Paper printed with certain inks can also present problems. The need to use these less desirable papers will continue to drive the development of more efficient recycling technologies. At the same time, more attention is being given to product life cycle management and the use of "environmentally benign" materials in paper and packaging that are easily removed in the recycling process. Another trend is the increased use of nonchlorine bleaching processes for color removal from recovered paper.

With the use of paper having more contaminants comes the challenge of finding ways to use the residue that the mills have to discard. Often this material goes to landfill, but some has been incorporated into products such as concrete and peat pots. Other research challenges include the need to develop standardized methods to measure the quality of both recovered and recycled fiber, and continued improvement in the properties of fibers used to make recycled papers. *See* PAPER.

Paul W. Resler

Wood

Waste is generated at every stage of the process by which a forest tree is turned into consumer and industrial products. Additional waste is generated in the disposal of those products. Trimmings and limbs are left in the forest. Historically, this remainder was simply piled in the forest and burned as slash. Wastes from the sawmill or from other forest-products

manufacturing facilities, for example, furniture and millwork companies, are called mill residues. Crates and pallets used for packaging are eventually disposed of as waste. Wood waste is also produced by the homeowner and by large and small businesses and is generated from landscaping and agricultural operations such as pruning and tree removal. The United States generates about 1.5 lb (0.7 kg) of wood waste per person per year, and wood waste accounted for 6% of total municipal solid waste (see **illustration**).

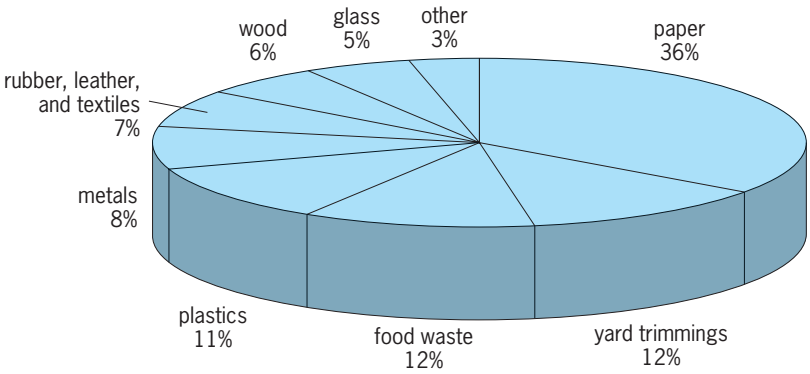

Municipal solid waste. (*U.S. Environmental Protection Agency*)

For each type of wood residue, there exists a recycling technology. Lumber and flooring can be reclaimed during the renovation and demolition of residential and commercial buildings. In addition, much of the wood residue generated is burned for energy recovery or is composted. While these processes do not strictly return wood to the economy in its original form, they have the effect of diverting wood residues from the landfill, and thus they may be included under a broadly interpreted definition of recycling. Several types of waste have been recycled for many years. For example, the particleboard industry originally developed in response to the availability of inexpensive sawdust. Nonetheless, much wood waste ends up in landfills. Diverting this wood from the landfill has become more important as disposal costs have risen. Municipal solid waste contains 5–20% solid wood and solid wood products such as plywood and particleboard, depending upon the season and geographic area.

Mill residues. These wastes consist of bark, sawdust, edgings, clippings, slabs, shavings, or chips produced by primary industries in the production of lumber, timbers, poles, plywood, and other wood products. Substantially all of this waste is recovered by the primary industries. Most of it is used for fuel, but the remainder is recycled into other products. It is considered a clean waste in comparison to that obtained from municipal solid waste. The secondary wood products industry also produces some clean waste wood. These are firms that use the products of the primary industry to produce more complex products, for example manufacturers of furniture, fencing, and windows.

Industrial and municipal wastes. The solid wood in municipal solid waste comes from several sources: wood left over from the construction, remodeling, or demolition of houses and other buildings (known as construction and demolition waste); wood from discarded pallets, crates, and other types of packaging; wood used in railroad crossties, utility poles, and other types of wood that have been treated with chemicals (treated wood).

Construction and demolition waste. This type of wood waste includes items such as siding removed from old barns, which is often reused in new buildings and valued for its weathered look. It also includes wood that has been treated with paints or stains, glues, binders, or resins (found in plywood, particleboard, and laminated veneer lumber), and wood composites (such as vinyl- or melamine-laminated wood for countertops). If construction or demolition waste is segregated while the structure is being built or demolished, the wood and some other materials can be recycled. The solid wood from the demolition waste stream is either removed from the building before it is demolished or separated from the other wastes. If not, it is invariably mixed with other materials such as cement, gypsum board, or glass. Then it must be separated before it can be recycled, and this separation must be done by hand. Therefore the value of the wood recovered must justify the expense of the separation. Only a small proportion of construction and demolition waste wood is directly recycled and used again as lumber. Most of this wood is processed into other products.

Packaging waste. Waste wood from packaging consists primarily of pallets. Since the mid-1970s the pallet industry has redesigned and standardized the pallet for recycling. This has led to a rapid increase in pallet recycling. Over 90% of the wood in recycled pallets is used to make new pallets. The remaining wood is used either as a fuel or in various products. An unknown number of pallets are also diverted by landfill operators and firms other than pallet manufacturers. These pallets are probably ground and used primarily for fuel or raw material for composite-board products.

Treated wood. A unique class of waste wood is treated wood, which has been injected with chemicals to prevent its decay. Disposal of these items presents problems due to the toxicity of the chemicals in the wood. In the United States, laws in most states allows treated wood to be disposed of in municipal landfills. Wood treated with creosote can be burned for fuel in permitted facilities. However, other types of treatment, such as chromated copper arsenate, yield wood containing heavy metals that cannot be safely burned. The preferred method for dealing with treated wood removed from service is recycling. For example, railroad ties that have been removed from service often continue to serve a useful purpose as landscape timbers. However, the disposal of treated wood and the contamination of the solid waste stream by treated wood are problematic.

Processing. Wood recycling begins with wood separation from the waste stream. Recovered

materials can then be processed into various products, including fuel, raw material furnish for the production of particleboard or other wood-composite panel products, compost, landscaping mulch, animal bedding, landfill cover, amendments for municipal solid waste and sludge compost, artificial firewood, wood-plastic composite lumber and other composite products, charcoal, industrial oil absorbents, insulation, and specialty concrete.

Most of these products require that the wood be ground into small particles. A typical grinder used for this application is a hammermill, although a variety of grinders are used. The size of the particles produced is determined by the end use of the wood; sizes smaller than about 20 mesh are called wood flour (the particles passing the 20-mesh screen are usually less than 0.8 mm in size). Though not produced by a saw, wood flour has the appearance of sawdust. Wood may then be passed over an electromagnet which removes items made of ferrous metals such as nails or staples. If additional processing is performed, it is typically to separate wood particles by size. This is accomplished in two ways: the particles can be passed through a series of screens of different mesh size and the various sized particles collected from the screens; or the particles may be separated in a tower with air blown in the bottom and out the top. The particles distribute themselves in the tower with the smaller, lighter particles on the top and the larger, denser ones at the bottom.

Wood that has been ground for use as fuel is known as hog fuel. The type of burner or furnace in which the hog fuel is used determines the applicable particle size. The range of particle sizes used in wood fuel is very large, from fine powders in fluidized-bed burners to fireplace logs. An important consideration in this use is moisture content. As the moisture increases, the amount of energy that can be recovered decreases, because some of the heat of combustion is used to convert the contained moisture into steam. About 30% of waste wood from all sources is used for fuel. However, this proportion varies from perhaps 15 to 50%, depending upon geographic region. *See* WOOD PRODUCTS. John Simonsen

Bibliography. M. J. Coleman (ed.), *Recycling Paper: From Fiber to Finished Product*, vols. 1 and 2, 1990; Earthworks Group, *The Recyclers Handbook*, 1990; F. R. Field III and J. P. Clark, Automobile recycling: Environmental policymaking in a constrained marketplace, *J. Metals*, pp. 17–21, April 1994; Glass Packaging Institute, *Glass Recycling: Why? How?*, 1991; T. Laufenberg, R. Rowell, and S. Sobozinski (eds.), *Materials Interactions Relevant to Recycling of Wood-Based Materials*, 1992; H. F. Lund, *The McGraw-Hill Recycling Handbook*, 2d ed., 2000; R. McKinney, *The Technology of Paper Recycling*, 1993; R. K. Miller and M. E. Rupnow, *Paper Recycling*, 1991; G. A. Smook, *Handbook for Pulp and Paper Technologists*, 2d ed., 1997; R. J. Spagenberg, *Secondary Fiber Recycling*, 1993; C. G. Thompson, *Recycled Papers: The Essential Guide*, 1992; Y. Virtanen and S. Nilsson, *Environmental Impacts of Waste Paper Recycling*, 1993.

Red dwarf star

A low-mass main-sequence star of spectral classes M and L. Red dwarf stars range from about 0.6 solar mass at class M0 down to 0.08 solar mass in cool M and warm L, below which the proton-proton chain cannot run. Lower-mass bodies are termed brown dwarfs. At the transition, within class L, the two are difficult to distinguish. Effective temperatures range from 3800 K (6400°F) at class M0 down to 2000 K (3100°F) at class L0, and absolute visual magnitudes from +9 to +20. Downward along the main sequence, red dwarfs produce progressively more radiation in the infrared. Bolometric corrections (which account for invisible radiation) range from −1.2 magnitudes to around −6, so luminosities range from about 0.05 down to 2×10^{-4} the solar luminosity. Radii range from about 0.5 to down 0.1 solar. Spectra become increasingly complex, with titanium oxide and neutral metals appearing at the warm end of class M, these and vanadium oxide appearing toward the cool end, and neutral metals and metal hydrides appearing in class L. *See* BROWN DWARF; MAGNITUDE (ASTRONOMY); PROTON-PROTON CHAIN.

Red dwarfs constitute over 70% of all stars, and as a result constitute about half the visible mass of the Milky Way Galaxy. The nearest stars are also dominantly red dwarfs, with the closest, Proxima Centauri (class M5), only 1.3 parsecs (4.2 light-years) away. Yet their intrinsic faintness renders them all invisible to the unaided eye; Proxima Centauri is of the 11th magnitude. Their lifetimes are so long that none has ever evolved in the Milky Way Galaxy's lifetime. *See* MILKY WAY GALAXY.

Warmer red dwarfs have (like the Sun) convective envelopes and radiative cores, but below about 0.3 solar mass, around class M4, they become convective throughout. Rotation and convection produce strong magnetic dynamos and complex fields that in turn produce powerful flares, especially common between M0 and M5. In a typical flare, the star increases in brightness by several tenths of a magnitude in less than a minute, the flare producing ultraviolet light and x-rays as well. *See* DWARF STAR; SPECTRAL TYPE; STAR; STELLAR EVOLUTION. James B. Kaler

Red Sea

A body of water that separates northeastern Africa from the Arabian Peninsula. The Red Sea forms part of the African Rift System, which also includes the Gulf of Aden and a complex series of continental rifts in East Africa extending as far south as Malawi. The Red Sea extends for 1920 km (1190 mi) from Ras (Cape) Muhammed at the southern tip of the Sinai Peninsula to the Straits of Bab el Mandab at the entrance to the Gulf of Aden. At Sinai the Red Sea splits into the Gulf of Suez, which extends for an additional 300 km (180 mi) along the northwest trend of the Red Sea and the nearly northward-trending Gulf of Aqaba. The 175-km-long (109-mi) Gulf of Aqaba forms the southern end of the Levant transform, a

primarily strike-slip fault system extending north into southern Turkey. The Levant transform also includes the Dead Sea and Sea of Galilee and forms the northwestern boundary of the Arabian plate. *See* ESCARPMENT; FAULT AND FAULT STRUCTURES.

Physiography. The coastlines of the Red Sea are remarkably straight and parallel for nearly 500 km (310 mi) south of the Gulf of Suez to about 24°N with a separation of 190 km (120 mi). South of there, the coastlines become sinuous, and the sea gradually widens to a maximum width of 350 km (210 mi) between 16°N and 17°N before narrowing to about 40 km (25 mi) at the Straits of Bab el Mandab. Irregular, eroded escarpments face seaward from the uplifted rift shoulders that flank the Red Sea. The rift shoulders average 1000–3000 m (3300–9900 ft) in elevation and expose Late Precambrian granitic, metamorphic, and volcanic rocks. The total width of the rift occupied by the Red Sea increases from about 250 km (150 mi) in the north to 500 km (300 mi) in the south. *See* PRECAMBRIAN.

The Red Sea consists of narrow marginal shelves and coastal plains and a broad main trough with depths ranging from about 400 to 1200 m (1300 to 3900 ft). The main trough is generally bounded by narrow marginal shelves and coastal plains. However, south of 20°N on the eastern side and 17°N on the western side, shallow reefs and carbonate banks nearly fill the entire main trough. Carbonate reefs are found along the entire length of the Red Sea, but are confined to nearshore areas further north. From 15°N to 20°N, the main trough is bisected by a narrow (<60 km or 37 mi wide) axial trough with a very rough bottom morphology and depths of greater than 2000 m (6600 ft). The maximum recorded depth is 2920 m (9580 ft). *See* REEF.

Oceanography. Water circulation in the Red Sea is driven by monsoonal wind patterns and changes in water density due to evaporation. Evaporation in the Red Sea is sufficient to lower the sea level by over 2 m (6.6 ft) per year. No permanent rivers flow into the sea, and there is very little rainfall. As a result, there must be a net inflow of water from the Gulf of Aden to compensate for evaporative losses. During the winter monsoon, prevailing winds in the Red Sea are from the south, and there is a surface current from the Gulf of Aden into the Red Sea. Evaporation increases the salinity of surface water, causing it to become more dense and to sink, particularly in the north where surface temperatures are lower. This dense water spreads southward along the bottom of the Red Sea. Some Red Sea deep water spills over the sill in the Straits of Bab el Mandab and can be traced as a distinct water mass at intermediate depths through the Gulf of Aden and into the Indian Ocean. During the summer monsoon, the wind in the Red Sea blows strongly from the north, causing a surface current out of the Red Sea. Gulf of Aden water flows into the Red Sea under the surface current. The southerly surface flow causes upwelling in the northern Red Sea and a northward bottom current. Warm, high-salinity water then spreads to the south under the northward-flowing Gulf of Aden water, and some of this dense water spills over into the Gulf of Aden. *See* MONSOON METEOROLOGY.

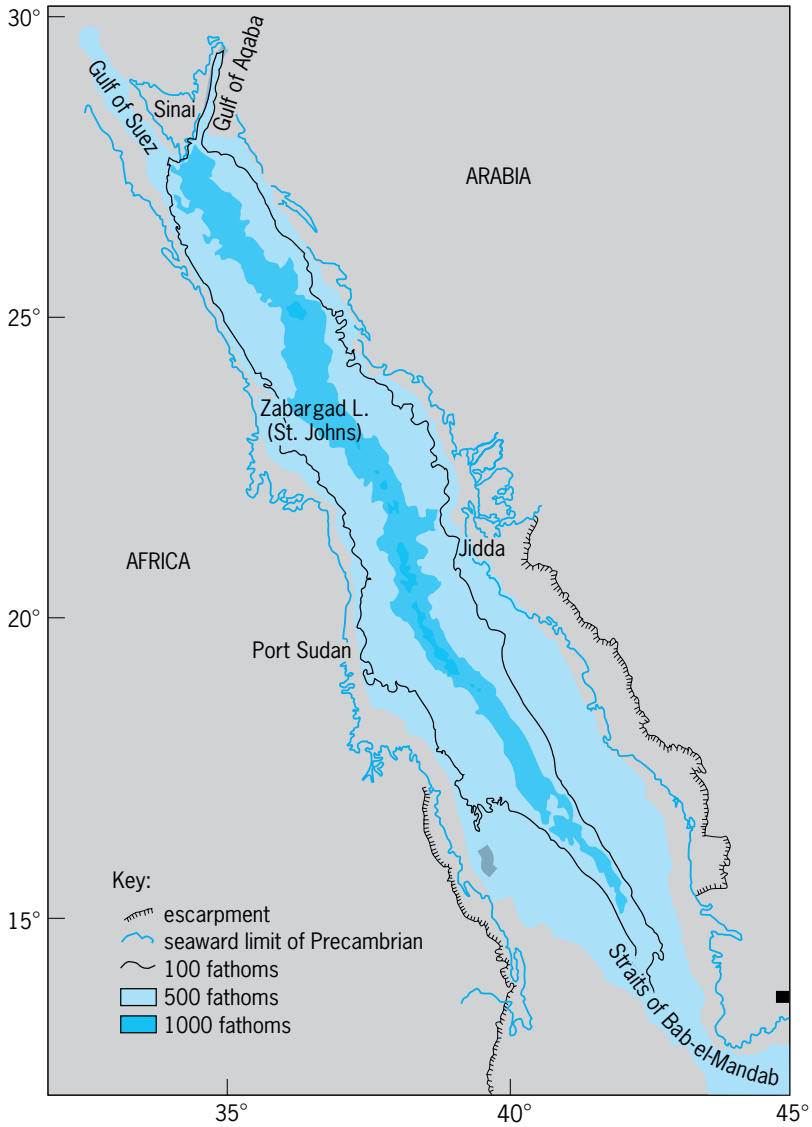

Division of the Red Sea into regions representing different stages in the formation of an ocean basin. (*After F. Martinez and J. R. Cochran, Structure and tectonics of the northern Red Sea: Catching a continental margin between rifting and drifting, Tectonophysics, 153:1–32, 1988*)

Geologic development. The rifting that has created the Red Sea began about 30 million years ago (Ma) [see **illus.**]. During the early development of the sea, the Gulf of Aqaba did not exist, and the Gulf of Suez formed the northern continuation of the Red Sea rift. The Gulf of Suez was cut off by the beginning of motion on the Levant transform in the middle Miocene (about 12–15 Ma). A total of about 140 km (87 mi) of motion has occurred between Africa and Arabia in the northern Red Sea and about 300 km (186 mi) of motion in the southern Red Sea. The present rates of motion between Africa and Arabia range from 0.75 cm/year (0.29 in./year) in the northern Red Sea to 1.5 cm/year (0.59 in./year) in the southern Red Sea.

Faunal evidence from sediments, particularly from the Gulf of Suez where there has been extensive

drilling for oil, indicates that during the early part of its development the Red Sea communicated with the Mediterranean. This connection was cut off or became extremely restricted in the middle Miocene, perhaps by uplift in Sinai related to the beginning of motion on the Levant transform. High evaporation rates caused sea level to fall rapidly, and the Red Sea became a shallow, hypersaline evaporitic basin precipitating primarily gypsum, anhydrite, and halite. The sea apparently did not completely dry out, but received either a small, steady supply or periodic influxes of water through the Gulf of Suez that maintained the salinity at a level where that mixture of evaporites was produced. These conditions persisted for perhaps as long as 10 million years, resulting in the accumulation of very thick (several kilometers) evaporitic deposits. About 5 Ma, tectonic activity in the southern Red Sea and Gulf of Aden created a connection with the Indian Ocean, ending evaporite deposition and reestablishing a more normal marine environment. *See* INDIAN OCEAN; MIOCENE; PLATE TECTONICS; SALINE EVAPORITES.

The axial trough in the southern Red Sea is associated with large-amplitude, lineated "sea-floor-spreading" magnetic anomalies resulting from the creation of new crust at a mid-ocean ridge spreading center. This interpretation is supported by the recovery of fresh oceanic basalt from the axial trough. The spreading center appears to have nucleated near 17°N at 4–5 Ma and to have propagated north and south to 19.5°N and 15°N, where the magnetic anomalies can be correlated to the geomagnetic time scale only to about 2 Ma. *See* GEOMAGNETISM; MID-OCEANIC RIDGE; PALEOMAGNETISM.

The axial trough becomes discontinuous north of 19.5°N and forms a series of isolated deeps spaced at 50–70-km (31–43-mi) intervals. These deeps are 25–40 km (16–25 mi) long, 10–25 km (6–16 mi) wide, and over 2000 m (6562 ft) deep, and are floored by oceanic basalt. They are associated with lineated sea-floor-spreading magnetic anomalies that can be identified to less than 2 Ma. The deeps are separated by intertrough zones that are shallower and broader, and are covered by several kilometers of sediment. These sediments are highly faulted and disturbed. The central Red Sea deeps appear to represent cells of organized sea-floor spreading that have not yet coalesced to form a continuous axis.

The main trough outside the axial trough has low-amplitude magnetic anomalies and appears not to have been formed by sea-floor spreading. The thick evaporitic sediments prevent investigation of the nature of the crust beneath the main trough by using seismic reflection techniques. Rifting in the southern Red Sea area has been accompanied by extensive volcanism. Large quantities of volcanic rocks are exposed onshore in Yemen, Ethiopia, and southwestern Saudi Arabia; and much of the crust along the margins of the Red Sea in these regions is made up of igneous intrusions. This pattern probably continues offshore, and it appears that extension between Africa and the Arabian Peninsula in the southern Red Sea has largely been accommodated by the injection of igneous material. These intrusions did not, however, become focused at a well-defined spreading center until 5 Ma. In contrast to the southern Red Sea, there has been very little volcanic activity accompanying the rifting in the northern half. Gravity measurements show a regular pattern of linear highs and lows that have been interpreted as resulting from a series of large rotated fault blocks. It thus appears that most of the extension in the northern Red Sea has been accommodated mechanically through faulting and thinning of the continental crust. *See* VOLCANOLOGY.

There is no evidence of organized sea-floor spreading in the northern 500 km (310 mi) of the Red Sea. The morphology of the northern Red Sea consists of a series of terraces stepping down to an axial region at a depth of about 1200 m (3937 ft) characterized by disturbed and deformed sediments. The axial region of the northern Red Sea is very similar in depth and appearance to the intertrough zones of the central Red Sea. Small deeps, generally associated with large magnetic anomalies, are spaced at roughly 60-km (36-mi) intervals along the northern Red Sea axis. These deeps are much smaller (a few kilometers across) and shallower (1400–1600 m or 4593–5250 ft deep) than the large central Red Sea deeps; and they are generally floored by thick sediments. The magnetic anomalies are all dipolar, indicating localized igneous intrusions, rather than lineated as are the sea-floor-spreading anomalies to the south. They also show that all of the intrusions are positively magnetized, implying that they were emplaced during the past 700,000 years since the last reversal of the Earth's magnetic field. It appears that the northern Red Sea deeps represent the first stage in the establishment of an oceanic spreading center within a rift.

Hot brines. Some small, enclosed basins within the central Red Sea deeps trap pockets of very dense, salty hot water. The salt content of the brine pools is greater than 25% by weight compared to normal Red Sea salinity of about 4%. Temperatures of the hot brines can be in excess of 50°C or 122°F (compared with average Red Sea bottom temperatures of 22°C or 72°F and normal deep-ocean bottom temperatures of 1–2°C (34–36°F)). The density contrast between the hot brines and the overlying seawater is great enough to cause a reflection of the sound waves used by echo sounders. The hot water results from hydrothermal circulation through the highly fractured, young oceanic crust, while the salt comes from the thick layer of evaporites. The hydrothermal circulation, in addition to heating the water, results in leaching of heavy minerals from the igneous rocks of the upper crust. These minerals become concentrated in the brine pools and precipitate out, resulting in the formation of sediments with very high concentrations of iron, manganese, zinc, lead, copper, and silver. Various schemes have been proposed to exploit these mineral deposits, but none have proved feasible. *See* AFRICA; SEAWATER. James R. Cochran

Bibliography. R. G. Coleman, *Geologic Evolution of the Red Sea*, 1993; E. T. Degens and D. A. Ross (eds.),

Hot Brines and Recent Heavy Metal Deposits in the Red Sea, 1969; A. L. Edwards and S. M. Head (eds.), *Key Environments: Red Sea*, 1987; I. I. Gass, *Crustal and Mantle Processes: Red Sea Case Study*, 1980; X. LePichon and J. Cochran (eds.), The Gulf of Suez and Red Sea rifting, *Tectonophysics*, vol. 153, special issue, 1988; S. Locke and R. C. Thunell, Paleoceanographic record of the last glacial/interglacial cycle in the Red Sea and Gulf of Aden, *Palaeogeog. Palaeoclimatol. Palaeoecol.*, 64:163–187, 1988; F. Martinez and J. R. Cochran, Structure and tectonics of the northern Red Sea: Catching a continental margin between rifting and drifting, *Tectonophysics*, 150:1–32, 1988.

Redbeds

Clastic sediments and sedimentary rocks that are pigmented by red ferric oxide which coats grains, fills pores as cement, or is dispersed as a muddy matrix. These conspicuously colored rocks commonly constitute thick sequences of nonmarine, paralic (marginal marine), and less commonly shallow marine deposits. Clastic redbeds accumulated in many parts of the globe during the past 10^9 years of Earth history. They were among the first sedimentary deposits to be considered climatic indicators, because their color was assumed to reflect a unique, identifiable condition of deposition. On the contrary, development of red pigment in most redbeds was complex and is difficult to decipher in detail. Ferric oxides also pigment marine chert, limestone, and cherty iron formations and ooidal ironstones, but these chemical deposits are not usually included among redbeds.

Some redbeds contain abundant grains of sedimentary and low-grade metamorphic rocks and relatively few grains of iron-bearing minerals. Most of them, however, contain feldspar and relatively abundant grains of opaque black oxides derived from igneous and high-grade metamorphic source rocks. In older redbeds the black grains are predominantly specular hematite replacing magnetite; in younger ones magnetite and ilmenite are more common. Dark grains of iron-bearing silicate minerals are rather rare in red sandstones; but they are more common in early-cemented concretions and in modern analogs, pointing to their postdepositional destruction which released pigmenting ferric oxide. Clay minerals in older redbeds, as in most other ancient clastic deposits, are predominantly illite and chlorite, thus providing no specific clue to the climate in the source area or at the place of deposition. *See* ARKOSE.

In many of the younger redbeds the pigmenting ferric oxide mineral cannot be identified specifically because of its poor crystallinity. In most of the older ones, however, hematite is the pigment. As seen under the scanning electron microscope, the hematite is in the form of hexagonal crystals scattered over the surface of grains and clay mineral platelets. In red mudstones most of the pigment is associated with the clay fraction.

Redbeds do not contain significantly more total iron than nonred sedimentary rocks. Normally, iron increases with decreasing grain size of redbeds. Moreover, the amount of iron in the grain-coating pigments is small compared with that in opaque oxides, dark silicates, and clay minerals. These facts demonstrate that chemical and mineralogical data cannot differentiate redbeds formed in moist climates from those formed in deserts.

Source areas. Actively eroded source areas of most redbeds supplied relatively abundant grains of magnetite and iron-bearing silicate minerals, whether in humid or dry climates. Accordingly, the common grains of specular hematite in redbeds must have been produced by an oxidizing burial regime. Similarly, the sparse crop of dark silicates commonly rimmed with red pigment resulted from postdepositional alteration, emphasizing its role in pigmentation and as a source of some of the clay mineral matrix as well.

Regardless of climate or color of their soils, source areas generally deliver brown sediments. Accordingly, the inherited brown muddy matrix that contained amorphous to very poorly crystalline hydrated ferric oxide (ferrihydrite) must have transformed slowly to red hematite pigment after deposition. This process is favored by exposure during accumulation and by persisting oxygenating conditions of burial in nonmarine to paralic environments. Postdepositional conversion of brown hydrated ferric oxide to red pigments may have also operated in shallow- to deeper-water marine deposits. In all of these red deposits, the final color is no direct clue to the climate in the source area.

Most redbeds were made from rather ordinary alluvium that accumulated in nonmarine to shallow-marine environments, under a pervasive oxygenating condition of burial within a common range of water chemistry. None was red at the time of deposition, and for each the environment during and after accumulation was crucial for postdeposition dehydration (aging) of ferrihydrate to hematite. Criteria such as associated fauna, flora, eolian sands, evaporites, or coal measures provide the most reliable evidence about the climate.

Mode of deposition. Redbeds occur in some alluvial fans, floodplains, deltas, lakes, and deserts. The fan and braided deposits are generally uniformly red; the floodplain and delta plain deposits are red-banded; and the desert alluvial, lake, and eolian deposits are generally uniformly red. In any case, diagenesis is, in a real sense, the control of redbeds. The processes that create red hematite are dehydration of brown ferric oxide, derivation from iron silicates, derivation from ilmenite and magnetite, and replacement of iron-bearing grains by dolomite cement.

Redbeds common in late Precambrian and Phanerozoic nonmarine to paralic deposits comprise several quite different facies produced by different combinations of factors. Some accumulated in basins whose adjacent highlands had essentially the same climate; others had relatively remote source areas where the climate may have differed from that of the

basin. Sediments deposited in a hot, dry climate were derived either from adjacent arid uplands or from more distant sources that may have been moister. The most convincing analogs of these desert redbeds are forming diagenetically today in nonmarine Cenozoic deposits in the arid and semiarid regions of the southwestern United States and northwestern Mexico. Deposits of this sort are composed of sandy alluvium, dune sand, and fanglomerate with relatively little mudstone. They were pigmented by postdepositional destruction (hydrolysis) of iron-bearing grains. Examples include late Paleozoic redbeds of Colorado and eastern Canada, as well as the Permian upper Rotliegendes of central Europe and Penrith eolian sandstone of Great Britain. *See* DIAGENESIS.

Extensive, well-sorted, fine-grained redbeds derived from distant sources accumulated on evaporitic tidal flats and in saline lagoons. They were also pigmented largely by postdepositional destruction of iron-bearing minerals. Thick and extensive red claystone in some of these successions suggests derivation from distant soil-mantled uplands with a moister climate, implying a reddening of the mud by postdepositional dehydration of inherited brown pigment. This facies includes the Permian redbeds of Texas, Oklahoma, New Mexico, and Arizona; Triassic redbeds of Wyoming and South Dakota; and Jurassic redbeds of the Colorado Plateau. *See* GYPSUM; HALITE; SALINE EVAPORITES.

Sediments that accumulated in savannas were derived either from adjacent uplands with seasonally humid climate or from more distant sources that may have been even moister. Savanna redbeds were supplied with detritus that included both fresh grains of iron-bearing minerals and a soil-derived clay fraction pigmented with brown hydrated ferric oxide that dehydrated after deposition to red hematite. Commonly, savanna redbeds comprise interbedded red and drab layers and lenses. Red-banded (mottled) deposits are usually reddish brown, maroon, and lavender as well as greenish gray, yellowish gray, and yellowish orange. Most of these red-banded deposits are fining-upward fluvial sequences with lower drab channel sandstone that accumulated under reducing conditions below the water table, and upper oxidized reddish-brown overbank (floodplain) mudstone that retained its inherited free ferric oxide during deposition. In long interflood periods of exposure, paleosols commonly developed in the upper part of the floodplain and tidal-flat mud. These horizons are marked by extensive burrowing, brown calcareous nodules, and root-tube fillings. Paleosols of this sort point to subhumid to semiarid climate with seasonal rainfall and low water table, yet with sufficient moisture to maintain an immature soil profile in the slowly aggrading floodplain mud. Similarly varicolored and "violet" paleosols in red-banded deposits have been interpreted as evidence of subtropical, humid to semiarid climate. In addition, some savanna deposits bear reliable faunal and floral evidence of a warm, moist climate, as in the redbeds of the late Paleozoic coal measures of Great Britain and eastern Canada, the lower Rotliegendes of central Europe, and the early Cenozoic deposits in the Rocky Mountain region. Other deposits are suggestive of some-what drier climate, for example, the Devonian Catskill redbeds of New York and Pennsylvania, the early Carboniferous Mauch Chunk Formation of Pennsylvania, and the Permian New Red Sandstone of Scotland. *See* PALEOSOL; SAVANNA.

Some of the red mudstones and siltstones that formed before the development of advanced land plants in mid-Paleozoic time accumulated in offshore marine environments, as in the Silurian marine redbeds of Great Britain and the United States, either on quite deep shelves or in basins of rapid sedimentation and burial. Moreover, some deepwater marine redbeds are resedimented nearshore deposits that were swept into deep basins by turbidity currents, as in the early Paleozoic of eastern Canada. *See* NEARSHORE PROCESSES; SEDIMENTARY ROCKS; TURBIDITY CURRENT.

Tectonics and continental drift. A genetic association of major redbeds with tectonic activity has long been recognized. As one tectonic end member, extensive uniform redbed-evaporite sequences accumulated on stable cratons, whereas some thick wedges of desert redbeds with border conglomerates accumulated in rift valleys. Overall, the tectonic background most commonly associated with redbeds is a late to post orogenic framework, as noted in the late Hercynian deposits of Europe and northwestern Africa, the late Alpine deposits of southern Europe, and the late Andean deposits of South America. Repeated episodes of deformation along the Appalachian belt produced redbeds at the close of three Paleozoic orogenies.

On a global scale, paleomagnetic evidence of the distribution of redbeds relative to their pole position corroborates paleogeographic data, suggesting that most redbeds, evaporites, and eolian sandstones accumulated less than 30° north and south of a paleo equator where hot, dry climate generally prevailed. But diagenetic development of red hematite may be acquired long after deposition. Moreover, continental drift reconstructions reveal that the most widespread redbeds in the geologic record developed near the Equator in late Paleozoic and early Mesozoic time when the continents were assembled in a great landmass, Pangaea. *See* CONTINENTS, EVOLUTION OF; PALEOGEOGRAPHY; PALEOMAGNETISM.

Franklyn B. Van Houten

Bibliography. R. H. Blodgett, J. P. Crabaugh, and E. F. McBride, The color of red beds: A geologic perspective, *Soil Sci. Soc. Amer. Spec. Publ.*, 31:127–159, 1993; T. M. Bown and M. J. Kraus, Lower Eocene alluvial paleosols (Willwood Formation, northwest Wyoming, U.S.A.) and their significance for paleoecology, paleoclimatology, and basin analysis, *Paleogeogr. Paleoclimat. Paleoecol.*, 34:1–30, 1986; H. Falke (ed.), *The Continental Permian in Central, West, and South Europe*, pp. 240–282, 1976; K. W. Glennie, Desert sedimentary environments, *Dev. Sedimentol.*, 14:173–193, 1970; J. C. Lorenz, *Triassic-Jurassic Rift-Basin Sedimentology*, 1988; E. F. McBride, Significance of color in red, green,

purple, olive, brown and gray beds of Difunta group, northeastern Mexico, *J. Sediment. Petrol.*, 44:760–773, 1974; J. G. McPherson, Calcrete (caliche) paleosols in fluvial redbeds of the Aztec Siltstone (Upper Devonian), southern Victoria Land, Antarctica, *Sediment. Geol.*, 22:267–285, 1979; P. Turner, *Continental Red Beds*, 1980; F. B. Van Houten, Origin of red beds: A review—1961–1972, *Annu. Rev. Earth Planet. Sci.*, 1:36–61, 1973; D. L. Woodrow, F. W. Fletcher, and W. F. Ahrnsbrah, Paleography and paleoclimate at the deposition sites of the Devonian Catskill and Old Red facies, *Geol. Soc. Amer. Bull.*, 84:3051–3064, 1973; A. M. Ziegler and W. S. McKerrow, Silurian marine red beds, *Amer. J. Sci.*, 275:31–56, 1975.

Redshift

A systematic displacement toward longer wavelengths of lines in the spectra of distant galaxies, and also of the continuous part of the spectrum. First studied systematically by E. Hubble, redshift is central to observational cosmology, in which it provides the basis for the modern picture of an expanding universe. There are two fundamental properties of redshifts.

First, the fractional redshift $\Delta\lambda/\lambda$ is independent of wavelength. ($\Delta\lambda$ is the shift in wavelength of radiation of wavelength λ.) This rule has been verified from 21 cm (radio radiation from neutral hydrogen atoms) to about 6×10^{-5} cm (the visible region of the electromagnetic spectrum) and leads to the interpretation of redshift as resulting from a recession of distant galaxies. Though this interpretation has been questioned, no other mechanism is known that would explain the observed effect.

Second, redshift is correlated with apparent magnitude in such a way that when redshift is translated into recession speed and apparent magnitude into distance, the recession speed is found to be nearly proportional to the distance. This rule was formulated by Hubble in 1929, and the constant of proportionality bears his name. Based on the results of the Hubble Key Project and the *Wilinkson Microwave Anisotropy Probe* (*WMAP*), Hubble's constant is currently estimated to lie between 67 and 75 km per second per megaparsec or 2.2 and 2.4×10^{-18} s^{-1}. *See* HUBBLE CONSTANT; MAGNITUDE (ASTRONOMY); WILKINSON MICROWAVE ANISOTROPY PROBE.

Until 1975 the largest redshifts that could be routinely measured for optical galaxies were around $\Delta\lambda/\lambda = 0.2$. Progressively higher redshifts have been measured as instruments have improved. Modern telescopes such as the Hubble Space Telescope and the Keck telescope have measured galaxy redshifts up to 6.68 and quasar redshifts up to 6.4. These objects were discovered in large surveys using photometric redshifts; this technique compares the predicted brightness of an object in various colors to the brightnesses observed, yielding an approximate redshift that can later be measured precisely using spectroscopy. *See* QUASAR; TELESCOPE.

The recession speed indicated by the redshift in the spectrum of a given galaxy is not the current value for that galaxy but the value appropriate to the epoch when the light now reaching the Earth was emitted. Consequently, the observed relation between redshift and apparent magnitude contains information about past values of Hubble's constant, as well as about the present value. If this information could be extracted from the record, it would enable astronomers to choose among various model universes that have been proposed by cosmologists. Current efforts to do this focus on type Ia supernovae, and the unexpected dimness of distant supernovae has been interpreted as evidence for acceleration; but the dimming may also plausibly be explained by supernova evolution or intergalactic dust. *See* SUPERNOVA.

Since redshifts of galaxies are nearly proportional to their distances, they can be used in conjunction with measured galaxy locations on the sky to construct very large scale three-dimensional maps of the galaxy distribution. When this was first done in 1988 for 1100 galaxies, the maps revealed structures in the galaxy distribution with sizes up to 100 megaparsecs (3×10^8 light-years). Subsequent redshift surveys (with up to 26,000 galaxies) have verified this result, and given statistical information about the large-scale distribution of cosmic matter. Redshift surveys now in progress, such as the Sloan Digital Sky Survey, will measure millions of additional galaxies. *See* COSMOLOGY; GALAXY, EXTERNAL; SLOAN DIGITAL SKY SURVEY; UNIVERSE.

David Layzer; Anthony Aguirre

Bibliography. E. J. Chaisson, *The Hubble Wars: Astrophysics Meets Astropolitics in the Two-Billion-Dollar Struggle over the Hubble Space Telescope*, Harvard University Press, 1998; P. J. E. Peebles, *Principles of Physical Cosmology*, Princeton University Press, 1993; S. Perlmutter et al., Discovery of a supernova explosion at half the age of the universe, *Nature*, 391:51, 1998; B. P. Schmidt et al., The high-Z supernova search: Measuring cosmic deceleration and global curvature of the universe using type Ia supernovae, *Astrophys. J.*, 507:46–63, 1998.

Redtop grass

One of the bent grasses, *Agrostis alba* and its relatives, which occur in cooler, more humid regions of the United States on a wide variety of soils. Redtop tolerates both wet and dry lands and acid and infertile soils, and it is used where other species of grasses do not thrive. Redtop is a perennial, spreads slowly by rootstocks, and makes a coarse, loose turf. Top growth is 2–3 ft (0.6–0.9 m) tall, with moderately leafy, wiry stems. The inflorescence is a reddish open panicle. Redtop is used for pasture and hay and is fairly nutritious if harvested promptly when heading occurs. It is effective in preventing erosion by holding banks of drainage ditches, waterways, and terrace channels. *See* CYPERALES; GRASS CROPS.

Howard B. Sprague

Redwood

A member of the pine family, *Sequoia sempervirens*, is the tallest tree in the Americas, attaining a height of 350 ft (107 m) and a diameter of 27 ft (8.2 m). Its range is limited to a strip along the Pacific Coast about 35 mi (55 km) wide and 500 mi (800 km) long, extending from southwest Oregon to about 100 mi (160 km) south of San Francisco, California. The leaves are evergreen, sharply pointed, small (0.3–1 in. or 8–25 mm long), distichous (disposed in two vertical rows) on short branches, and scalelike on the main stem (see **illus.**). The cones are egg-shaped, about 1 in. (25 mm) long and 0.5 in. (13 mm) broad. The bark is a dull red-brown, on old trees sometimes 1 ft (0.3 m) thick, densely fibrous, and highly resistant to fire. The tree gets its common name from the color of the bark as well as that of the heartwood. The wood is moderately light in weight and strength and low in warp and shrinkage, and is not difficult to work. The heartwood is especially valuable, being highly resistant to decay. The wood holds paint well. It is used for bridge timbers, tanks, flumes, silos, posts, shingles, paneling, doors, caskets, furniture, siding, and many other building purposes. *See* FOREST AND FORESTRY; PINALES; PINE; TREE.

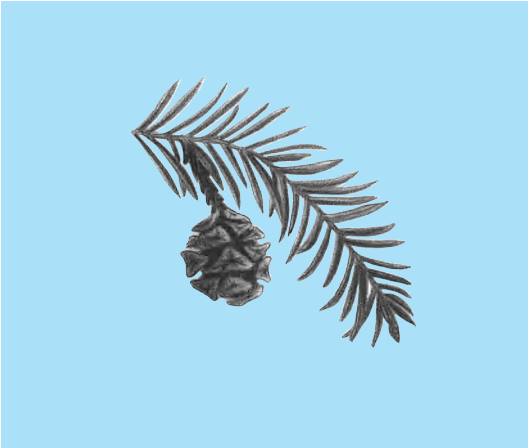

Redwood (*Sequoia sempervirens*) leaves and cones.

Arthur H. Graves; Kenneth P. Davis

Reef

A rigid and wave-resistant marine structure that stands above its surroundings. Biologists and geologists specify that reefs are constructed by organisms that secrete calcium carbonate skeletons. To navigators, a reef is any rocky structure that poses a threat to navigation.

Coral reefs. Modern reef-builders include algae, mollusks, bryozoans, worms, and sponges, but the most important are corals.

Distribution. Coral reefs are most common in warm water (typically $\geq 20^\circ$C) in which calcium carbonate precipitates more easily. Most shallow-water corals are hermatypic, and they host symbiotic zooxanthellae (algal-like dinoflagellates) that provide metabolites through photosynthesis. This provides energy for the coral while removing carbon dioxide that inhibits calcification.

While coral reefs are more common in the shallow tropics, they can be found in both deeper and colder water, where they do not have zooanthellae. According to the National Oceanic and Atmospheric Administration, reefs dominated by the coral *Lophelia pertusa* occur off the coast of Florida at depths of 500–850 m; similar features built by *Oculina varicosa* are found at water depths of 70–100 m in the Gulf of Mexico and the Straits of Florida. Off the coast of Norway, *Lophelia* reefs occur at depths of 75–155 m, at temperatures ranging from 4 to 13°C. Coral diversity is much lower in these deep-water reefs, but they can harbor a variety of organisms. Between 350 and 400 species of mollusks, crustaceans, and other fauna have been reportedly found within the branches of *Oculina* off the eastern coast of Florida. *See* SCLERACTINIA.

Living corals and reefs. Most corals are colonial, with each individual polyp measuring from several millimeters to a few centimeters across (**Fig. 1***a*). Each polyp secretes an external calcium carbonate skeleton and sits within a cuplike calyx (Fig. 1*b*) that is part of the larger colony (Fig. 1*c*). At night, the coral polap extends its tentacles to feed. During the day, the coral retreats into its "skeletal cup" and relies on zooxanthellae for nutrition.

Reef types. Generally, coral reefs can be divided into three categories: fringing, barrier, and atoll. This classification is based on Darwin's original observations of mid-Pacific reefs.

Volcanoes form as a moving oceanic plate passes over a rising plume of hot mantle (**Fig. 2**). As the sea floor moves away from this hot spot, the volcanic lava (basalt) cools and subsides. Coral reefs initially form along the steep slopes of a new volcano in shallow waters close to shore. As the volcano subsides (moves away and sinks), reefs build vertically to stay close to sea level and the light needed for photosynthesis; the lagoon behind the reef progressively widens, and the influence of land decreases. Fringing reefs, which are closest to shore, occur around newly formed and thus high volcanic islands. As the volcanic island slowly subsides, reefs transition to barriers, and eventually to atolls, once the volcano has sunk from sight beneath the surface waters. *See* ATOLL; OCEANIC ISLANDS; VOLCANO.

Coral zonation. The species and shapes of corals on a reef are controlled primarily by light availability, wave energy, and sedimentation, resulting in predictable reef zones (**Fig. 3**). Branching corals, such as *Acropora palmata* in the Caribbean Sea (Fig. 1*c*; background), occur along the reef crest, where abundant light drives the photosynthesis needed to support rapid growth. Vigorous wave energy clears sediment from the branching species that have no

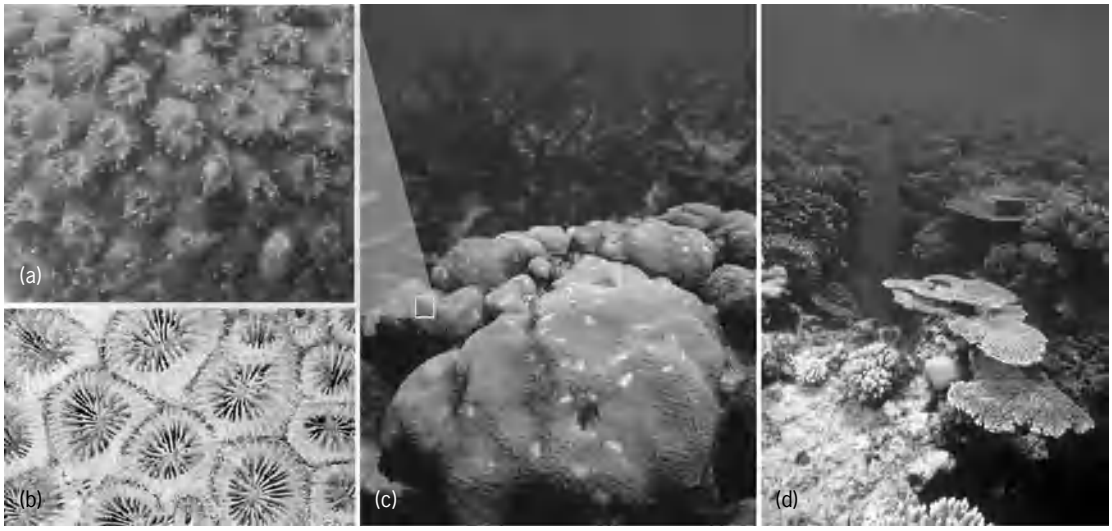

Fig. 1. Photographs illustrating the relationship between (*a*) individual coral polyps, (*b*) the skeletal structures they form, and (*c*) the colony in which they live. The hemispherical colony of *Montastraea faveolata* (*c*; foreground) sits on a sloping Caribbean forereef in 10 m of water. Branching colonies of *Acropora palmata* (*c*; background) occupy shallower water near the reef crest. (*d*) The branching/tablelike corals near the crest are responding to similar conditions and give way to more massive forms in deeper water along the Great Barrier Reef of Australia.

biological means of sediment removal. Although the species of branching and table-like corals vary by ocean, these types of corals usually occur near the reef crest, in response to similar processes (Fig. 1*d*). Further down the reef front, slower-growing hemispherical corals (Fig. 1*c*, foreground; Mixed/Massive Zone in **Fig. 3**) are better suited for removing sediment coming down the slope. In even deeper water, platelike colonies (Fig. 3) respond to continually decreasing light levels. This shape places all the polyps on upward-facing surfaces, optimizing the colony's ability to gather light—much like solar panels.

Reef building. Reefs have traditionally been thought of as structures built by intact and interlocking corals

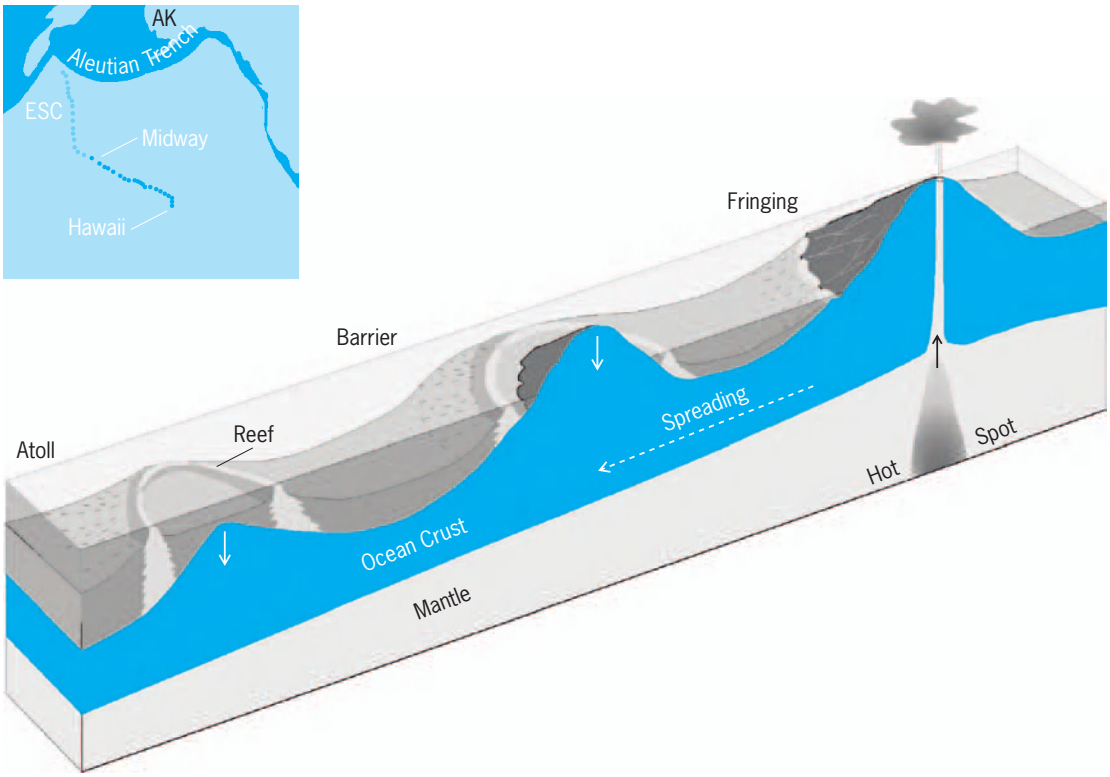

Fig. 2. Diagram of the evolution of fringing reefs into barrier reefs and atolls as subsiding volcanoes move away from a hot spot. The inset shows the trend of volcanic islands (dark color) and seamounts (sunken volcanoes, light color) associated with the hot spot near Hawaii. The volcanoes get older away from Hawaii.

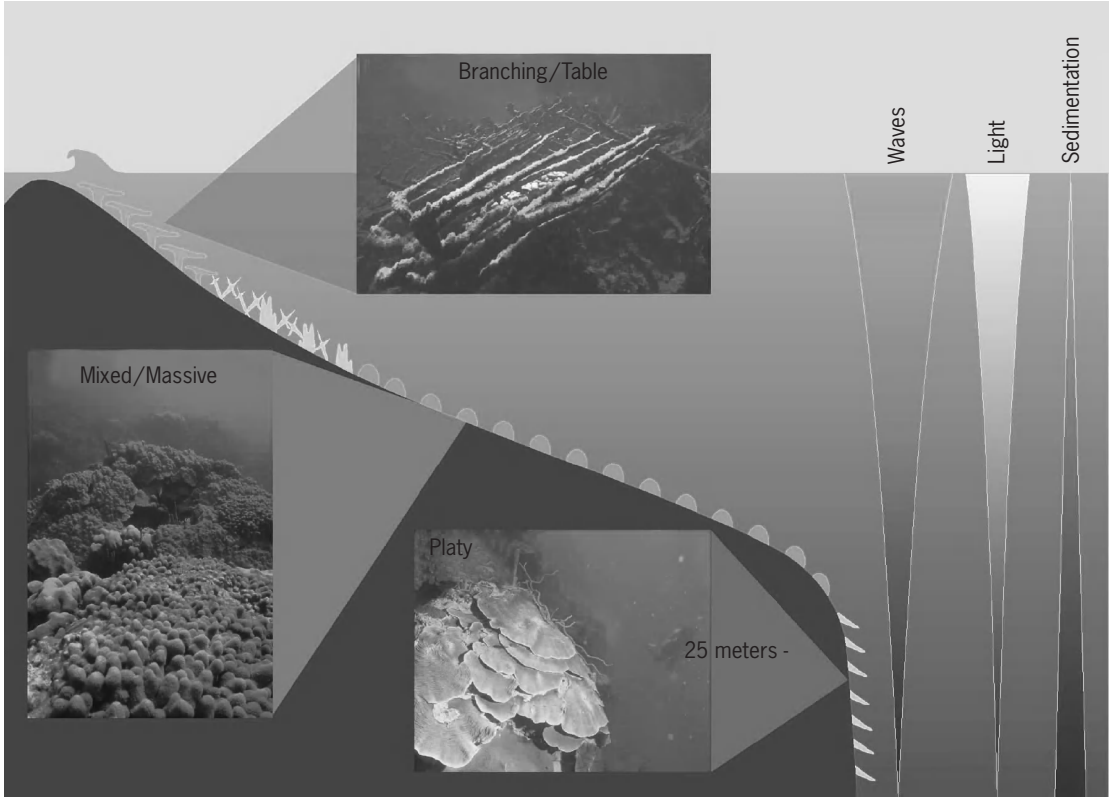

Fig. 3. Zonation on a typical Caribbean reef. The shift from branching corals to mixed/massive and platy corals down the reef front is a response to decreasing light and wave energy, accompanied by increasing sedimentation with depth.

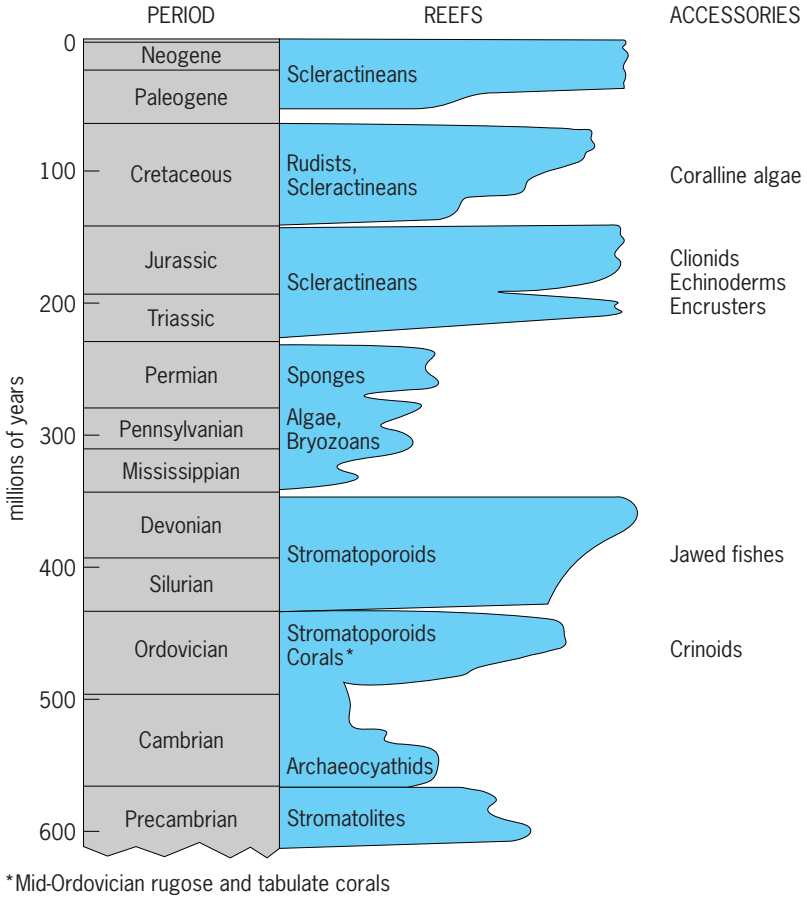

Fig. 4. Major reef-building episodes (Reefs column) through geologic time. The major organisms responsible are listed. The origin of other important organisms that were associated with these reefs are also indicated (ACCESSORIES column).

growing atop one another over the course of hundreds to thousands of years (6000–12000 years for modern reefs). However, recent coring studies have recognized that much of the reef interior consists of toppled and degraded corals, sediment, and open cavities. Most of the sediment is produced by (1) fish and urchins grazing on algae that cover dead coral surfaces and (2) sponges, mollusks and worms that excavate into the carbonate substrate seeking shelter (a process known as bioerosion). This nearly equal mix of intact and broken corals plus loose sediment is subsequently bound together by encrusting organisms (for example, coralline algae) and carbonate cement. Thus, the edifice that is the coral reef owes its final structure as much to bioerosion, cementation, and encrustation as it does to skeletal production by corals. *See* ALGAE; CORALLINALES.

Reefs through time. Reefs have not always been built by corals. Microbial stromatolites formed nonskeletal mounds approximately 3 billion years ago (**Fig. 4**). Photosynthesis by these widespread reef-builders provided the oxygen needed for the evolution of complex organisms. Over the ensuing 550 million years, the evolution of reef-building organisms was controlled by changes in ocean chemistry (which affected calcification), changes in sea level (which increased or decreased space for colonization), increases in habitat variability, and increases in competition and predation. *See* PALEOECOLOGY.

In the mid-Cambrian period (543–490 million years ago), archaeocyathids, followed by early

corals (rugose and tabulate) and stromatoporoids (calcifying sponges), may have evolved due to increased grazing pressure (Fig. 4). Early coral evolution paralleled the rapid evolution of herbivorous fish, particularly those with jaws in the Devonian (417–354 million years ago). The diversification of photosynthesizing dinoflagellates may have coincided with the radiation of both grazing and boring species from Triassic through Cretaceous time (approximately 250–65 million years ago). In this regard, increases in grazing/boring versus calcification resembled an "arms race" between organisms that built reefs and those that bioeroded them for food or shelter. *See* ARCHAEOCYATHA; RUGOSA; STROMATOPOROIDEA; TABULATA.

Reefs under threat. Dramatic changes have occurred in reefs over the past three decades. In the mid-1970s, coral cover commonly exceeded 50–75%. Since then, coral abundance has plummeted on many reefs, especially those closer to population centers. The low nutrient levels characteristic of healthy reefs rose in response to increased septic waste, agricultural fertilizers, and sedimentation. Macroalgae (seaweeds) that were previously limited by low nutrients suddenly took over. Sedimentation from dredging and deforestation not only blocked sunlight needed for photosynthesis, it also eliminated the hard, stable substrate needed for coral recruitment. The ensuing shift from coral-dominated reefs to dense macroalgal communities was exacerbated by overfishing (especially of grazers that kept algae populations balance) in and the loss of other key species, such as the long-spined sea urchin (*Diadema altillarum*) in 1983 from disease. Even the deep reefs described earlier are under threat, mostly by breakage related to trawling for deep-sea fishes.

Recently, global factors have added to the more direct assaults on reef habitat and the organisms that keep the benthos in balance. Rising temperatures have caused corals to expel their zooxanthellae, exposing the bare skeleton through the remaining clear coral tissue (bleaching). If temperatures remain elevated for more than 2–3 weeks, the corals die. Starting with white band disease in the late 1970s, corals have been assaulted by episodes of tissue degeneration. Arguments abound over the specific causes of recent declines, but the scientific community is unanimous in its assessment of likely anthropogenic ties for the global degradation of coral reefs. Dennis Hubbard

Bibliography. D. K. Hubbard, Dynamic processes of coral-reef development, in C. Birkeland (ed.), *Life and Death of Coral Reefs*, Chapman and Hall, 1997; T. P. Hughes, Catastrophes, phase shifts and large-scale degradation of a Caribbean coral reef, *Science*, 265:1547–1551, 1994; J. Pandolfi et al., Are U.S. coral reefs on the slippery slope to slime?, *Science*, 307:1725–1726, 2005; L. L. Richardson, Coral diseases: What is really known?, *Trends Ecol. Evolut.*, 13:438–443, 1998; G. D. Stanley, *The History and Sedimentology of Ancient Reef Systems*, Klewer Academic/Plenum Press, New York, 2001.

Reengineering

The application of technology and management science to the modification of existing systems, organizations, processes, and products in order to make them more effective, efficient, and responsive. Responsiveness is a critical need for organizations in industry and elsewhere. It involves providing products and services of demonstrable value to customers, and thereby to those individuals who have a stake in the success of the organization. Reengineering can be carried out at the level of the organization, at the level of organizational processes, or at the level of the products and services that support an organization's activities. The entity to be reengineered can be systems management, process, product, or some combination. In each case, reengineering involves a basic three-phase systems-engineering life cycle comprising definition, development, and deployment of the entity to be reengineered.

Systems-management reengineering. At the level of systems management, reengineering is directed at potential change in all business or organizational processes, including the systems acquisition process life cycle itself. Systems-management reengineering may be defined as the examination, study, capture, and modification of the internal mechanisms or functionality of existing system-management processes and practices in an organization in order to reconstitute them in a new form and with new features, often to take advantage of newly emerged organizational competitiveness requirements, but without changing the inherent purpose of the organization itself.

Process reengineering. Reengineering can also be considered at the levels of an organizational process. Process reengineering is the examination, study, capture, and modification of the internal mechanisms or functionality of an existing process or systems-engineering life cycle, in order to reconstitute it in a new form and with new functional and nonfunctional features, often to take advantage of newly emerged or desired organizational or technological capabilities, but without changing the inherent purpose of the process that is being reengineered.

At the level of processes alone, the reengineering effort would be almost totally internal. It would consist of modifications to whatever standard life-cycle processes are in use in a given organization in order to better accommodate new and emerging technologies or new customer requirements for a system. For example, an explicit risk-management capability might be incorporated at several different phases of a given life cycle and accommodated by a revised configuration of management processes. Alternatively, the work-flow processes that lead to services could be reengineered. This capability could be implemented into the processes for research, development, test, and evaluation (RDT&E); acquisition or procurement; and systems planning and marketing. Basically, reengineering at the level of processes would consist of the determination, or synthesis, of an efficacious process for ultimately fielding a product on the basis of a knowledge of generic

customer requirements, and with appropriate consideration for the objectives and critical capabilities of the systems-engineering organization. *See* RISK ASSESSMENT AND MANAGEMENT.

However, reengineering of processes alone, without attention to reengineering at a higher level, may, in many instances, represent an incomplete and not fully satisfactory way to improve organizational capabilities. Thus, the processes considered as candidates for reengineering should be high-level managerial processes as well as operational processes. Information technology is a major enabling catalyst for process reengineering. The benefits of using a process reengineering include shorter development cycles, fewer engineering change orders, products that fulfill customer expectations, and reduced program and product development costs throughout the life cycle. Thus, the process improvement results ultimately in an increase in effectiveness of product for the same cost, a reduction in cost for the same effectiveness, or some blend of these two.

Product reengineering. The term "reengineering" could mean some sort of reworking or retrofit of an already engineered product, and could be interpreted as maintenance or refurbishment. Reengineering could also be interpreted as reverse engineering, in which the characteristics of an already engineered product are identified, such that the product can perhaps be modified or reused. Inherent in these notions are two major facets of reengineering: it improves the product or system delivered to the user for enhanced reliability or maintainability, or to meet a newly evolving need of the system users; and it increases understanding of the system or product itself. This interpretation of reengineering is almost totally product-focused.

Thus, product reengineering may be redefined as the examination, study, capture, and modification of the internal mechanisms or functionality of an existing system or product in order to reconstitute it in a new form and with new features, often to take advantage of newly emerged technologies, but without major change to the inherent functionality and purpose of the system. This definition indicates that product reengineering is basically structural reengineering with, at most, minor changes in purpose and functionality of the product. This reengineered product could be integrated with other products having rather different functionality than was the case in the initial deployment. Thus, reengineered products could be used, together with this augmentation, to provide new functionality and serve new purposes. There are a number of synonyms for product reengineering, including renewal, refurbishing, rework, repair, maintenance, modernization, reuse, redevelopment, and retrofit.

A specific example of a product reengineering effort is that of taking a legacy system written in Cobol or Fortran, reverse-engineering it to determine the system definition, and then reengineering it in C^{++}. Depending upon whether any modified user requirements are to be incorporated into the reengineered product, the product is reengineered through a forward engineering effort, either just after reverse engineering has determined the initial development (technical) system specifications, or after reverse engineering has been carried out far enough to determine user requirements and user specifications, and these have then been updated. This reverse engineering concept, in which salient aspects of user requirements or technological specifications are recovered from examination of characteristics of the product, predates product reengineering. *See* PROGRAMMING LANGUAGES.

In the late 1990s, much product reengineering was focused on eliminating the potential Y2K difficulty that might ensue when older software was used that assumed that all years could be expressed by 19*xy*, where *x* and *y* varied between 0 and 9. The year 2000 could not be accommodated in such a representation, and so calculations made using the *xy* number might be in serious error. Often, this software was written many years ago, and full documentation of the software might be lacking.

Taxonomy and life-cycle phases. Much of product reengineering is very closely associated with reverse engineering to recover either design specifications or user requirements. Then follows refinement of these requirements or specifications and forward engineering to achieve an improved product (see **illus.**). Forward engineering is the original process of defining, developing, and deploying a product, or realizing a system concept as a product; whereas reverse engineering, sometimes called inverse engineering, is the process though which a given system or product is examined in order to identify or specify the definition of the product either at the level of technological design specifications or at system- or user-level requirements. Reverse engineering may be broken down into redocumentation and design recovery.

Redocumentation is a subset of reverse engineering in which a representation of the subject system or product is recreated for the purpose of generating functional explanations of original system behavior

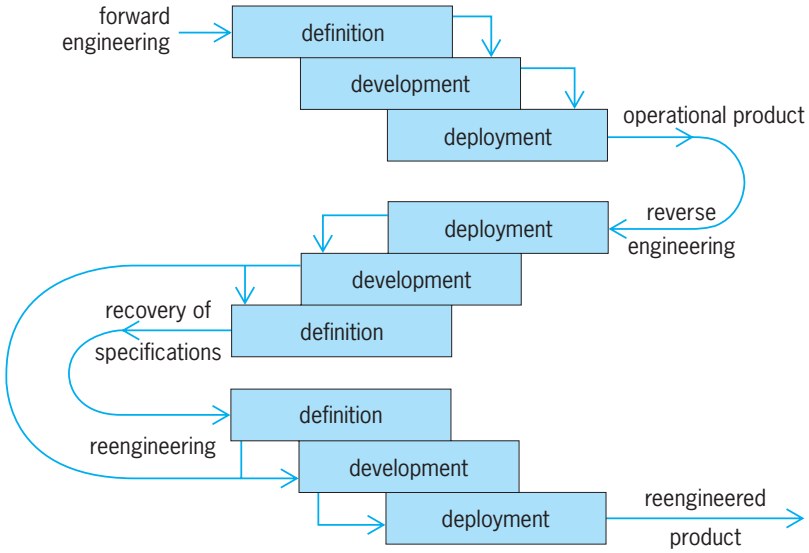

Product reengineering as a sequence of forward, reverse, and forward engineering.

and, perhaps more importantly, to aid the reverse engineering team in better understanding the system at both a functional and a structural level. A number of redocumentation tools are available for software. One major purpose of redocumentation is producing new documentation for an existing product where the existing documentation is faulty, or perhaps virtually absent.

Design recovery is a subset of reverse engineering in which the redocumentation knowledge is combined with other efforts, often involving others' experiences with and knowledge of the system, that lead to functional abstractions and enhanced product or system understanding at the level of function, structure, and even purpose.

Restructuring involves transformation of the reverse engineering information concerning the original system structure into another representation form. Restructuring generally preserves the initial functionality of the original system, or modifies it slightly in a purposeful manner that is in accord with the user requirements for the reengineered system and the way in which they differ from the requirements for the initial system.

Objectives. There are four major objectives for product reengineering.

1. *Improved maintainability*. Often, it is not easy to justify reengineering on the basis of improving maintainability because of the lack of appropriate maintenance data. *See* RELIABILITY, AVAILABILITY, AND MAINTAINABILITY.

2. *Migration to a new platform*. It may be necessary to move the system, especially software, to a better operational environment, often because of improvements in technology, such as the development of a better microprocessor or a new operating system.

3. *Greater reliability*. There may be reliability problems associated with the old system, and a reengineered version may be considerably more reliable. Restructuring reveals most such potential defects.

4. *Functional integration*. Reengineering the system into a collection of subsystems may improve the structure of these systems and thereby make it easier to change or add functionality through the addition of new subsystems without affecting existing ones.

These are not mutually exclusive objectives, and a reengineered product may gain performance in accordance with each of them.

Justification. Often, it is not easy to justify reengineering an existing product. A reengineering judgment is often reached as a last alternative after it turns out not to be possible to engineer a new product from the beginning, to purchase an available commercial product, or to continue to maintain the old one. A program of systematic measurements that can demonstrate improved reliability, maintainability, functionality, and organizational value is needed. Generally, a number of concerns must be addressed in order to demonstrate the cost effectiveness, or lack thereof, of a product reengineering effort. It is necessary to determine the operating costs, maintenance costs, business value, and estimated future life of the current system; and the costs, estimated time to completion, and associated risks of the reengineering process. For the reengineered system, the operating costs, the business value and maintenance costs after reengineering, and the system life must be estimated.

Planning. Planning for product reengineering is essential. Any planning effort must involve the definition phase, the development phase, and the deployment phase. In term, the definition phase involves formulation of the reengineering issue such as to determine the need for, and requirements to be satisfied by, the reengineered product; identification of potential alternative candidates for reengineering; analysis of the alternatives to enable determination of costs and benefits of the various alternatives; and interpretation and selection of a preferred plan for reengineering. In the development phase, the detailed specifications for implementation of the reengineering plan are determined; and in the deployment phase, operational plans, including contracting, are set forth to enable reengineering of the product in a cost-effective and trustworthy manner.

Needs. Several needs must be considered if a product reengineering process is to yield appropriate and useful results:

1. Long-term organizational and technological issues must be examined in developing a product reengineering strategy.

2. Human, leadership, and cultural issues must be considered, including how they will be affected by the development and deployment of a reengineered product, as a part of the definition of the specifications for the product.

3. It must be possible to demonstrate that the reengineering process and product are, or will be, cost-effective and high-quality, and that they support continued evolution of future capabilities.

4. Reengineered products must be considered within an enhanced framework that also considers the potential need for reengineering at the levels of systems management and organizational processes, since technological fixes alone will generally not be sufficient to resolve organizational difficulties at these levels.

5. In particular, product reengineering for improved postdeployment maintainability must consider maintainability at the level of process rather than at the level of product only, such as would result in the case of software through rewriting source-code statements. Use of model-based management systems or code generators should yield much greater productivity, in this connection, than rewriting source code.

6. Product reengineering must consider the need for reintegration of the reengineered product with existing legacy systems that have not been reengineered.

7. Increased conformance to standards must be a result of the product reengineering process.

8. Product reengineering must consider legal issues associated with reverse engineering.

Relation to product reuse. Product reengineering is closely related to product reuse, especially for software. The reengineering of legacy software and the reuse-based production of new software are closely related concepts. The cost of developing software for one or a few applications is approximately the same as the cost of developing domain reuse components and reengineering approaches to legacy software.

Risks. Reengineering is accompanied by a variety of risks that are associated with processes, people, tools, strategies, and the application area for reengineering. (1) Integration risk is associated with having a reengineered product that cannot be satisfactorily integrated with, or interfaced to, existing legacy systems. (2) Maintenance improvement risk is the possibility that the reengineered product will exacerbate, rather than ameliorate, maintenance difficulties. (3) Systems management risk arises when the reengineered product attempts to impose a technological fix on a situation where the major difficulties call not for greater support but for organizational reengineering at the level of systems management. (4) Process risk is associated with having a reengineered product that might well represent an improvement in a situation where the specific organizational process in which the reengineered process is to be used is defective and in need of reengineering. (5) Cost risk is associated with having major cost overruns in order to obtain a deployed reengineered product that meets specifications. (6) Schedule risk is associated with having schedule delays in order to obtain such a product. (7) Human acceptance risk is associated with obtaining a reengineered product that is not suitable for human interaction, or one that is unacceptable to the user organization for other reasons. (8) Application supportability risk is associated with having a reengineered product that does not really support the application or purpose it was intended to support. (9) Tool and method availability risk is associated with proceeding with reengineering a product based upon promises for a method or tool, needed to complete the effort, which does not become available or which is faulty. (10) Leadership, strategy, and culture risk is associated with imposing a technological fix in the form of a reengineered product in an organizational environment that cannot adapt to this product. These risks are not mutually exclusive, the risk attributes are not independent, and there are additional risks, such as legal risks associated with ownership issues. *See* INFORMATION SYSTEMS ENGINEERING; SOFTWARE ENGINEERING; SYSTEMS ENGINEERING. Andrew P. Sage

Bibliography. M. Hammer, *Beyond Reengineering: How the Process-Centered Organization Is Changing Our Work and Our Lives*, Harper Business, New York, 1996; S. Hannaford, *Workflow Reengineering*, Hayden Publishing, 1996; M. A. Ould, *Business Processes: Modeling and Analysis for Re-engineering and Improvement*, John Wiley, New York, 1995; A. P. Sage, *Systems Management for Information Technology and Software Engineering*, Wiley, New York, 1995; A. P. Sage and W. B. Rouse (eds.), *Handbook of Systems Engineering and Management*, Wiley, New York, 1999.

Reference electrode

An electrode with an invariant potential. In electrochemical methods, where it is necessary to observe, measure, or control the potential of another electrode (denoted indicator, test, or working electrode), it is necessary to use a reference electrode, which maintains a potential that is practically unchanged during the course of an electrochemical measurement. Potentials of indicator or working electrodes are measured or expressed relative to reference electrodes. *See* POTENTIALS.

Hydrogen electrode. One such electrode, the normal hydrogen electrode, has been chosen as a reference standard, relative to which potentials of other electrodes and those of oxidation-reduction couples are often expressed. A hydrogen electrode consists of a platinum foil, the surface of which is covered with finely divided platinum (platinum black). This is achieved by electroplating spongelike platinum onto the surface of the foil from a solution of chloroplatinic acid. The surface of the electrode is saturated with hydrogen (H_2) gas that is bubbled into the test solution (**Fig. 1***a*). Alternatively, the hydrogen gas (saturated with water vapor) can be bubbled continuously over the surface of the platinum (Fig. 1*b*). Electrical contact to a potentiometer can be made through the mercury in the glass tube to which the platinum foil is attached, and the metal leads are immersed into mercury.

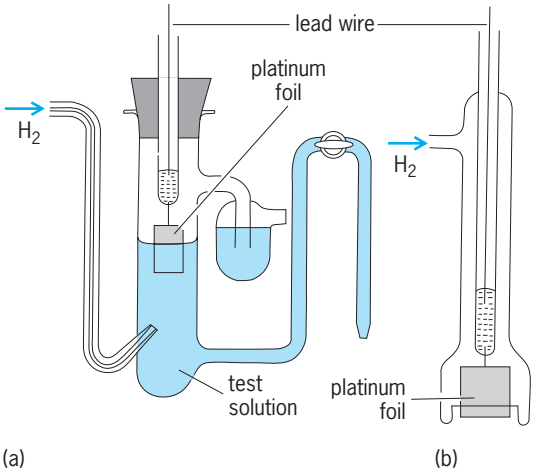

Fig. 1. Hydrogen electrodes. (*a*) Stationary bubbling type. (*b*) Dipping type.

When the hydrogen electrode is immersed in a solution containing hydrogen ions, such ions accept electrons from the electrode and form hydrogen atoms [reaction (1)], which then combine to form hy-

$$H^+ + e \rightleftharpoons H \qquad (1)$$

drogen molecules [reaction (2)]. In the presence of

$$2H \rightleftharpoons H_2 \qquad (2)$$

platinum black as a catalyst, both of these equilibria are rapidly established, forming a reversible system. The potential (E) of a hydrogen electrode depends on the pressure of hydrogen gas (p_{H2}) and on activity of hydrogen ions (a_{H^+}) in the solution. The potential is given by Eq. (3), where R is the gas constant, T

$$E = -\frac{RT}{F}\ln a_{H^+} + \frac{RT}{2F}\ln p_{H_2} \qquad (3)$$

the absolute temperature, and F the Faraday. At 25°C (77°F) this equation has the form of Eq. (4).

$$E = -0.0591 \log a_{H^+} + 0.0296 \log p_{H_2} \qquad (4)$$

Hence, at a constant hydrogen pressure, a tenfold change in hydrogen-ion activity yields a change of 0.0591 V in the potential of the hydrogen electrode. Alternatively, at constant hydrogen-ion activity, a tenfold change in the pressure of the hydrogen gas results in a change of 0.0296 V in the potential of the hydrogen electrode. *See* ACTIVITY (THERMODYNAMICS).

Measurement of the potential of the hydrogen electrode can be used for determination of the pressure of hydrogen gas; however, such applications are rare. On the other hand, by maintaining a constant pressure of hydrogen gas the potential of a hydrogen electrode can be used for determination of the activity of hydrogen ions in the tested solution. Achievement of a well-functioning hydrogen electrode is difficult, and some components of the test solution can affect its function. Thus, in practice the determination of the hydrogen-ion activity (pH) is performed by using a glass electrode. The hydrogen electrode itself is used only in fundamental studies and some nonaqueous solutions. *See* PH.

The hydrogen electrode, however, remains important for providing a reference standard. For this purpose, the normal hydrogen electrode standard is achieved with a hydrogen electrode immersed in a test solution in which the hydrogen-ion activity is unity and the hydrogen gas pressure is 1 atm (100 kilopascals). The potential of this electrode, the normal hydrogen electrode, is conventionally assigned a value of 0 V at all temperatures. All electrodes with positive potential relative to the normal hydrogen electrode have greater reducing power than hydrogen; conversely, those of negative potential have lower reducing power. Such comparisons enable arrangement of the chemical elements into a series according to their reducing power (electrochemical series). *See* ELECTROCHEMICAL SERIES; OXIDATION-REDUCTION.

Secondary reference electrodes. In practice, potentials are measured against reference electrodes that are easier to work with than the normal hydrogen electrode. Such electrodes are known as secondary reference electrodes; the most common are the calomel and silver–silver chloride electrodes.

Calomel electrode. This consists of a small pool of mercury covered with a layer of a paste formed by dispersing a small amount of metallic mercury in mercury(I) chloride. The electrode is placed in an aqueous solution of potassium chloride of known concentration. Most frequently it is a saturated solution, and such an electrode is known as a saturated calomel electrode (SCE). Sometimes calomel electrodes contain either a 1.0 M (molar) or a 0.1 M solution of potassium chloride. The potential of the saturated calomel electrode measured against the normal hydrogen electrode is given by Eq. (5), where t is the

$$E = -0.2440 + 0.00076(t - 25^\circ\text{C}) \quad \text{volts} \qquad (5)$$

temperature in degrees Celsius.

When the calomel electrode is used in measurement of pH and potentials of other electrodes, in polarography, and in voltammetry, a reservoir for the solution of potassium chloride may be provided on the side arm, which can serve as a salt bridge. The reservoir enables flushing and replacing the solution in the side arm with a fresh potassium chloride solution (**Fig. 2**).

Silver–silver chloride electrode. This consists of silver wire or foil covered with a layer of silver chloride, which is usually produced by electrolysis. The silver wire or foil is made an anode (positive electrode) in a solution containing chloride ions. Silver may also

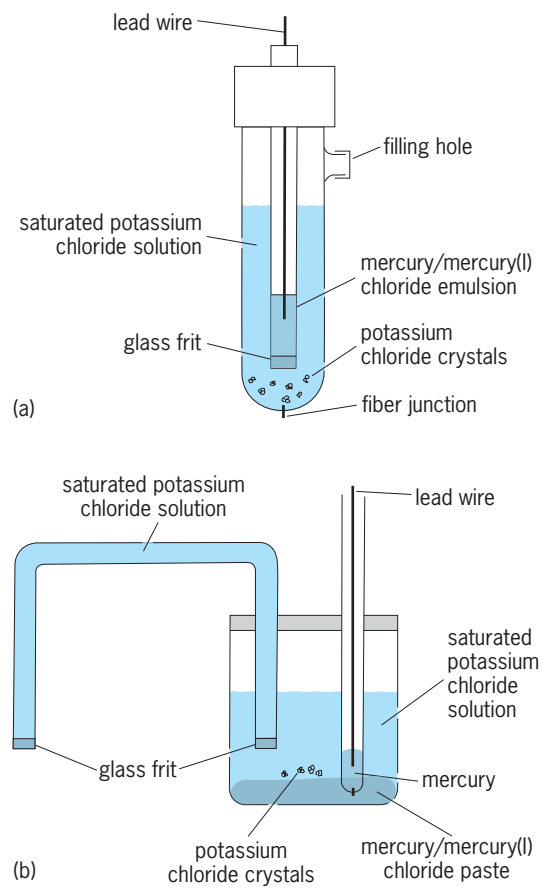

Fig. 2. Calomel electrodes. (*a*) "Upside-down" type, the type most used commercially. (*b*) Two liquid-junction potential types.

be deposited at first on a noble metal by reduction of silver oxide or silver oxalate by heating to 500°C (932°F). This process is then followed by electrodeposition of silver chloride. Alternatively, an intimate mixture of silver oxide and silver chlorate is pasted on platinum and heated to 500°C.

The potential of the silver–silver chloride electrode is defined only in a solution of known concentration of chloride ions. It is a logarithmic function of chloride-ion concentration, and measurement of the potential of such an electrode can be used for determination of chloride-ion concentration. When the silver–silver chloride electrode is used as a reference electrode, the concentration of the chloride ions must be known and stated. In 1 *M* potassium chloride solution at 25°C (77°F), the potential of this electrode is −0.222 V relative to the normal hydrogen electrode. Silver–silver chloride electrodes are used in potentiometric measurements (particularly in connection with ion selective electrodes), polarography, and voltammetry. *See* ION-SELECTIVE MEMBRANES AND ELECTRODES; POLAROGRAPHIC ANALYSIS.

In nonaqueous systems, an electrode consisting of a silver wire or foil in a solution of silver nitrate is often used as a reference. Such an electrode is suitable for comparison of standard, half-wave or peak potentials of various compounds in a given solvent. Nevertheless, the use of such reference electrodes containing one solvent (including water) for comparison of potentials of a given compound in different solvents is difficult. At the interface between the solvent in the reference electrode and the solvent in the studied solution, a liquid junction potential is established. No simple means for measuring this potential are available.

To avoid this complication, a different approach is used, involving adoption of a reference system that is different from the standard hydrogen electrode used in aqueous solutions. These reference systems are oxidation-reduction couples, the potential of which can be assumed to be only slightly sensitive to the composition of the solvent. Ferrocene-ferricinium ion, cobaltocene ion, and bis(diphenyl) chromium are the most frequently used systems. A platinum electrode immersed in the solution containing equal concentrations of the oxidized and reduced forms in the given solvent can serve as a reference electrode.

Alternatively, one of these systems can be added to the investigated solution and the potential of the studied compound measured relative to added system; this is known as an internal standard. Such an approach can be used even in aqueous solutions in voltammetric and polarographic studies. Reduction of thallium(I) ions is independent of composition of the solution in most common electrolytes. Potential of thallium(I) ion is then used as an internal standard, and potentials of studied substances are measured relative to it. *See* ELECTRODE; SOLVENT. Petr Zuman

Bibliography. V. S. Bagotzky, *Fundamentals of Electrochemistry*, 1993; A. J. Bard, R. Parsons, and J. Jordan (eds.), *Standard Potentials in Aqueous Solutions*, 1985; J. Koryta, J. Dvořák, and L. Kavan, *Principles of Electrochemistry*, 1993; P. H. Rieger, *Electrochemistry*, 2d ed., 1993.

Reflecting microscope

A microscope whose objective is composed of two mirrors, one convex and the other concave. The imaging properties are independent of the wavelength of light, and this freedom from chromatic aberration allows the objective to be used even for infrared and ultraviolet radiation. Although the reflecting microscope is simple in appearance, the construction tolerances are so small and so difficult to achieve that the system is used only when refracting objectives are unsuitable. The distance from the objective to the specimen can be made very large; this large working distance is useful in special applications, such as examining objects situated within metallurgical furnaces. Reflecting microscopes have been mainly used for microspectrometry in the infrared and the ultraviolet, and for ultraviolet microphotography.

Optical system. The concave mirror collects light from the specimen under examination. The convex mirror intercepts this light just above the point where the rays would be focused by the concave mirror, and redirects the light up the microscope tube to the plane where the primary image is formed (see **illus.**). The curvatures and spacing of the mirrors can be so chosen that the optical aberrations

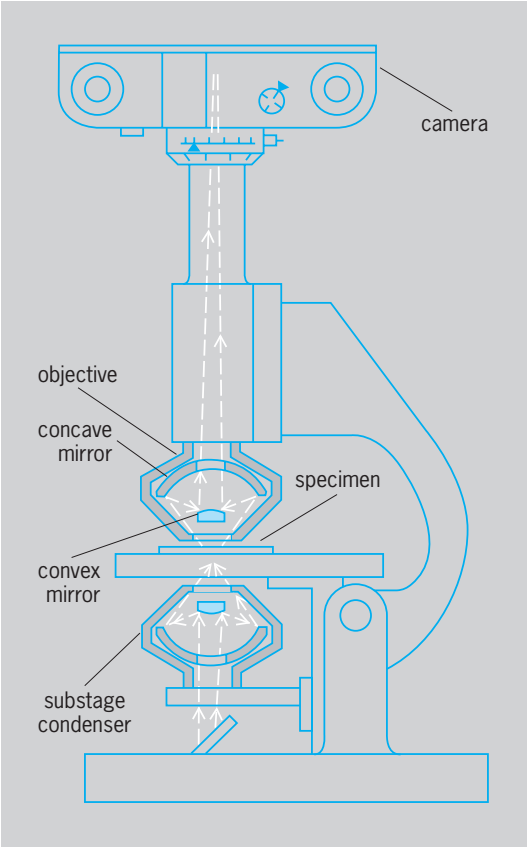

Reflecting microscope arranged for photomicrography.

of the mirrors are mutually compensating, provided that the numerical aperture does not exceed about 0.60 (visible light). The limiting aperture is somewhat smaller than this for ultraviolet radiation and somewhat larger for infrared, and is smaller for large working distances than for small working distances.

The optical designer, in attempting to achieve numerical apertures larger than about 0.60, may resort to mirrors that depart slightly from spherical curvature; alternatively, he may add one or more lens elements. By such means, objectives of numerical aperture greater than unity and free of chromatic aberration can be realized.

The convex mirror unavoidably obscures a small central portion of the aperture. This alters the diffraction of light (upon which the ultimate resolving power of the objective depends) but fortunately this effect is negligible if proper care is used in the design and construction.

Substage illumination of the specimen can sometimes be achieved by a conventional refracting condenser if only visible light is used. Usually, however, the chromatic aberration of the condenser is troublesome and a condenser similar to the objective is used. Chromatic aberration of the ocular is usually not serious; hence the ocular may be of conventional form, although usually constructed of fused silica, fluorite, or other materials transparent to the infrared or ultraviolet. The ocular may be omitted for special applications in which the primary image formed by the objective is examined directly. David S. Grey

Illumination. The reflecting microscope has several unique characteristics that make it highly suitable for applications which require nonvisible radiation to illuminate the specimen. Reflecting objectives are achromatic to wavelengths ranging from ultraviolet to the infrared. As a result, specimens can be illuminated with single- or multiple-wavelength radiation without the problem of correcting for chromatic aberrations. There is also a fairly large depth of focus at higher magnifications, so that there is little or no need to adjust the focus as the wavelength of the light source is changed, making photomicrography less problematic for the microscopist using nonvisible illumination.

The reflecting objective lens has a long working distance at higher magnifications and, therefore, is more useful than the common refracting objectives for performing micromanipulations on a specimen. Since their introduction in the 1700s, reflecting lenses have been dramatically improved and can now be constructed with high numerical apertures (0.90) and fewer requirements for aspherical corrections.

Uses. The reflecting microscope can be used with a photometer system for the nondestructive spectral investigation of the absorption and transmission of visible and nonvisible radiation from various objects such as specific regions of cells and tissues, crystals, minerals, synthetic or natural dyed fibers, hairs, inks on documents, paint chips, cosmetics, and polymers. An exceedingly small amount of specimen, with little or no sample preparation, can reveal important optical and chemical information. The reflecting microscope in conjunction with ultraviolet illumination and an ultraviolet spectrophotometer is useful in characterizing the position and content of deoxyribonucleic acid (DNA) in various normal and abnormal cells. The image resolution with infrared illumination is somewhat poorer than with ultraviolet illumination. However, when the microscope is interfaced with a Fourier transform infrared (FTIR) spectrometer, it reveals molecular microstructural information on infrared-illuminated, opaque or transparent objects as small as 1 micrometer in diameter. As a result, the Fourier transform infrared reflecting microscope has found a wide range of applications in the biochemical, medical, and forensic sciences. *See* SPECTROSCOPY. Lawrence Kobilinsky

Bibliography. G. H. Needham, *The Practical Use of the Microscope Including Photomicrography*, 1958; J. A. Reffner, Molecular microspectral mapping with the FT-IR microscope, *Inst. Phys. Conf. Ser. 98*, 1990.

Reflection and transmission coefficients

When an electromagnetic wave passes from a medium of permeability μ_1 and dielectric constant ϵ_1 to one with values μ_2 and ϵ_2, part of the wave is reflected at the boundary and part transmitted. The ratios of the amplitudes in the reflected wave and the transmitted wave to that in the incident wave are called the reflection and transmission coefficients, respectively. For oblique incidence, the reflection and refraction formulas of optics are most convenient, but for normal incidence of plane waves on plane boundaries, such as occur with transmission lines, waveguides, and some free waves, the concept of wave impedance and characteristic impedance is useful.

For a z-directed wave with electric intensity $\mathbf{E}$ in the x direction and magnetic intensity $\mathbf{H}$ in the y direction, the total phasor fields on the incident side are given by Eqs. (1) and (2), where primes are used

$$\check{E}_x = E_0 e^{-jkz} + \check{E}'_0 e^{jkz} \tag{1}$$

$$\check{H}_y = (\eta)^{-1}(E_0 e^{-jkz} - E'_0 e^{jkz}) \tag{2}$$

for reflected quantities and η is the wave impedance. The sign difference in Eqs. (1) and (2) is due to the fact that Poynting's vector, $^1/_2\mathbf{E} \times \mathbf{H}$, is positive for the incident and negative for the reflected wave. For the transmitted wave, Eqs. (3) hold. Since the tangen-

$$\check{E}''_x = E''_0 e^{-jkz} \qquad \check{H}''_y = (\eta'')^{-1} E''_0 e^{-jkz} \tag{3}$$

tial components of $\mathbf{E}$ and $\mathbf{H}$ are continuous across the boundary at $z = 0$, $\check{E}_x = \check{E}''_x$ and $\check{H}_y = \check{H}''_y$, so that Eqs. (4) hold. The ratios for the reflected and trans-

$$E_0 + E'_0 = E''_0 \qquad \eta''(E_0 - E'_0) = \eta E''_0 \tag{4}$$

mitted fields obtained by solving these equations are given by Eqs. (5), which are the reflection and

$$\frac{E'_0}{E_0} = \frac{\eta'' - \eta}{\eta'' + \eta} \qquad \frac{E''_0}{E_0} = \frac{2\eta''}{\eta'' + \eta} \tag{5}$$

transmission coefficients, respectively. *See* ELECTROMAGNETIC RADIATION; POYNTING'S VECTOR.

Coefficients for optics. Equations (5) hold for normal incidence in optics, if the velocities υ and υ'' are written for η and η''. For a plane wave whose electric vector is normal to the plane of incidence and whose direction makes an acute angle θ with the normal to the interface, the reflection and transmission coefficients are given by Eqs. (6), where υ'' sin θ''. When

$$\frac{E_0'}{E_0} = -\frac{\sin(\theta - \theta'')}{\sin(\theta + \theta'')} \qquad \frac{E_0''}{E_0} = \frac{2\sin\theta''\cos\theta}{\sin(\theta + \theta'')} \tag{6}$$

the electric vector lies in the plane of incidence, the coefficients are given by Eqs. (7). The ratio υ/υ'' is

$$\frac{E_0'}{E_0} = \frac{\tan(\theta - \theta'')}{\tan(\theta + \theta'')} \qquad \frac{E_0''}{E_0} = \frac{2\sin\theta''\cos\theta}{\sin(\theta + \theta'')\sin(\theta - \theta'')} \tag{7}$$

the index of refraction. *See* REFRACTION OF WAVES.

Waveguides. In waveguides, as in free space, the characteristic impedance is defined as the ratio of the transverse electric field E_t to the transverse magnetic field H_t. For waveguides, this radio depends on the frequency and the dimensions of the waveguide, as well as on the permeabilities and dielectric constants. For a transverse interface, the boundary conditions used for Eqs. (4) on the tangential fields still hold. Thus, Eqs. (5) for the reflection and transmission coefficients are valid if η and η'' are replaced by the characteristic impedances on the incident and emergent sides, respectively. *See* WAVEGUIDE.

Transmission lines and networks. Let $\check{Z}$ and $\check{Z}''$ be the characteristic impedances on the incident and emergent sides of a discontinuity in a transmission line or on the two sides of a junction between two networks. Then the relations between the potentials and currents of the incident, reflected, and transmitted waves are given by Eqs. (8*a*)–(8*c*), respectively.

$$\check{V} = \check{Z}\check{I} \tag{8a}$$

$$\check{V}' = -\check{Z}\check{I}' \tag{8b}$$

$$\check{V}'' = \check{Z}''\check{I}'' \tag{8c}$$

At the discontinuity, potential and current must be continuous so that Eqs. (9) hold. Solution for the ratios gives Eqs. (10), which are the reflection and

$$\check{V} + \check{V}' = \check{V}'' \qquad \check{I} + \check{I}' = \check{I}'' \tag{9}$$

$$\frac{\check{V}'}{\check{V}} = \frac{\check{Z}'' - \check{Z}}{\check{Z}'' + \check{Z}} \qquad \frac{\check{V}''}{\check{V}} = \frac{2\check{Z}''}{\check{Z}'' + \check{Z}} \tag{10}$$

transmission coefficients. *See* TRANSMISSION LINES.

Coefficients for acoustics. Equations (10) hold in acoustics, provided acoustic impedance is substituted for electrical impedance. Acoustic impedance is defined as the product of the density of a medium by the speed of sound in it. There are two types of waves in solid mediums, longitudinal waves and shear waves, and thus there are two impedances. *See* ACOUSTIC IMPEDANCE.

William R. Smythe

Bibliography. M. F. Iskander, *Electromagnetic Fields and Waves*, 2000; P. Lorrain and D. R. Corson, *Electromagnetic Fields and Waves*, 3d ed., 1987; S. Ramo, J. R. Whinnery, and T. Van Duser, *Fields and Waves in Communication Electronics*, 3d ed., 1994; W. R. Smythe, *Static and Dynamic Electricity*, 3d ed., 1968, reprint 1984; D. H. Staelin, A. W. Morgenthaler, and J. A. Kong, *Electromagnetic Waves*, 1998.

Reflection of electromagnetic radiation

The returning or throwing back of electromagnetic radiation such as light, ultraviolet rays, radio waves, or microwaves by a surface upon which the radiation is incident. In general, a reflecting surface is the boundary between two materials of different electromagnetic properties, such as the boundary between air and glass, air and water, or air and metal. Devices designed to reflect radiation are called reflectors or mirrors.

Reflection angle. The simplest reflection laws are those that govern plane waves of radiation. The law of reflection concerns the incident and reflected rays (as in the case of a beam from a flashlight striking a mirror) or, more precisely, the wave normals of the incident and reflected rays and the normal to the reflecting surface all lie in one plane, called the plane of incidence, and that the reflection angle θ_{refl} equals the angle of incidence θ_{inc} as in Eq. (1) [**Fig. 1**]. The

$$\theta_{\text{refl}} = \theta_{\text{inc}} \tag{1}$$

angles θ_{inc} and θ_{refl} are measured between the surface normal and the incident and reflected rays, respectively. The surface (in the above example, that of the mirror) is assumed to be smooth, with surface irregularities small compared to the wavelength of the radiation. This results in so-called specular reflection. In contrast, when the surface is rough, the reflection is diffuse. An example of this is the diffuse scattering

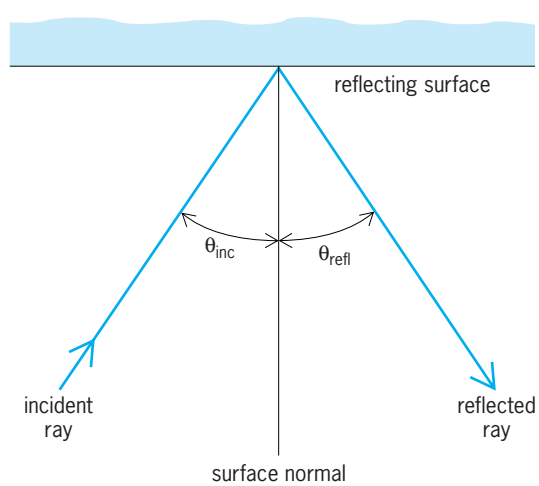

Fig. 1. Reflection of electromagnetic radiation from a smooth surface.

of light from a screen or from a white wall where light is returned through a whole range of different angles.

Reflectivity. The reflectivity of a surface is a measure of the amount of reflected radiation. It is defined as the ratio of the intensities of the reflected and incident radiation. The reflectivity depends on the angle of incidence, the polarization of the radiation, and the electromagnetic properties of the materials forming the boundary surface. These properties usually change with the wavelength of the radiation. Reflecting materials are divided into two groups: transparent materials and opaque conducting materials. Radiation penetrating a transparent material propagates essentially unattenuated, while radiation penetrating a conducting material is heavily attenuated. Transparent materials are also called dielectrics. In the wavelength range of visible light, typical dielectrics are glass, quartz, and water. Conducting materials are usually metals such as gold, silver, or aluminum, which are good reflectors at almost all wavelengths. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; DIELECTRIC MATERIALS.

Reflection from metals. The reflectivity of polished metal surfaces is usually quite high. Silver and aluminum, for example, reflect more than 90% of visible light. The reflectivity of some common metals as a function of wavelength is given in **Fig. 2** for normal incidence ($\theta_{inc} = 0$). The reflectivities vary considerably with wavelength, generally falling off toward the shorter wavelengths, with silver exhibiting a reflection "window" at 320 nanometers. The reflectivity values depend somewhat on the way the metal surface was prepared; for example, whether or not it was polished or was produced by evaporation. The presence of an oxidation layer is also a factor influencing (and usually decreasing) the reflectivity. In ordinary mirrors the reflecting surface is the interface between metal and glass, which is thus protected from oxidation, dirt, and other forms of deterioration. When it is not permissible to use this protection for technical reasons, one uses "front-surface" mirrors, which are usually coated with evaporated aluminum. The aluminum has a high reflectivity and deteriorates relatively little in the atmosphere. Front-surface mirrors are used in scientific applications such as interferometry and in large reflecting telescopes.

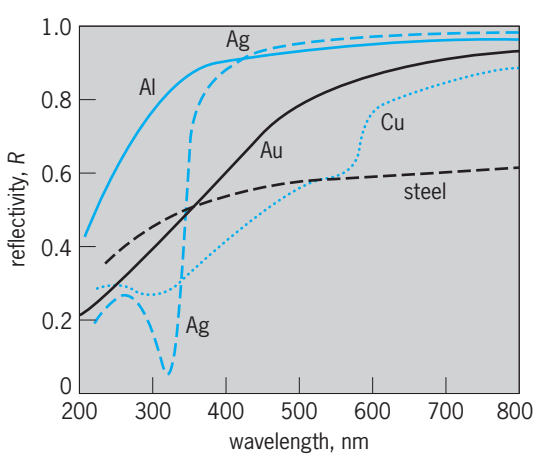

Fig. 2. Reflectivity of some common metals for normal incidence as a function of wavelength. (*After F. A. Jenkins and E. White, Fundamentals of Optics, 4th ed., McGraw-Hill, 1976*)

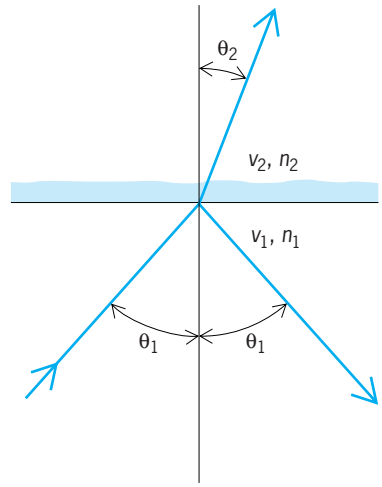

Fig. 3. Reflection and refraction of electromagnetic radiation at an interface between two dielectric media.

Reflection from dielectrics. The material property that determines the amount of radiation reflected from an interface between two dielectric media is the phase velocity v of the electromagnetic radiation in the two materials. In optics one uses as a measure for this velocity the refractive index n of the material, which is defined by Eq. (2) as the ratio

$$n = \frac{c}{v} \tag{2}$$

of the velocity of light c in vacuum and the phase velocity in the material. For visible light, for example, the refractive index of air is about $n = 1$, the index of water is about $n = 1.33$, and the index of glass is about $n = 1.5$. *See* PHASE VELOCITY; REFRACTION OF WAVES.

For normal incidence ($\theta_{inc} = 0$) the reflectivity R of the interface is given by Eq. (3), in which the

$$R = \left(\frac{v_1 - v_2}{v_1 + v_2}\right)^2 = \left(\frac{n_2 - n_1}{n_2 + n_1}\right)^2 \tag{3}$$

material constants are labeled 1 and 2 as shown in **Fig. 3**, where the radiation is incident in material 1. The reflectivity of an air-water interface is about 2% ($R = 0.02$) and that of an air-glass interface about 4% ($R = 0.04$); the other 98% or 96% are transmitted through the water or glass, respectively.

A ray incident upon the interface at an oblique, nonnormal angle θ_1 is deviated as it penetrates material 2 as shown in Fig. 3. This is called refraction, and the refracted angle θ_2 follows from Snell's law of refraction, Eq. (4). Again, there is partial reflection, with the reflectivity depending on the angle of incidence θ_1 and the polarization of the radiation.

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{4}$$

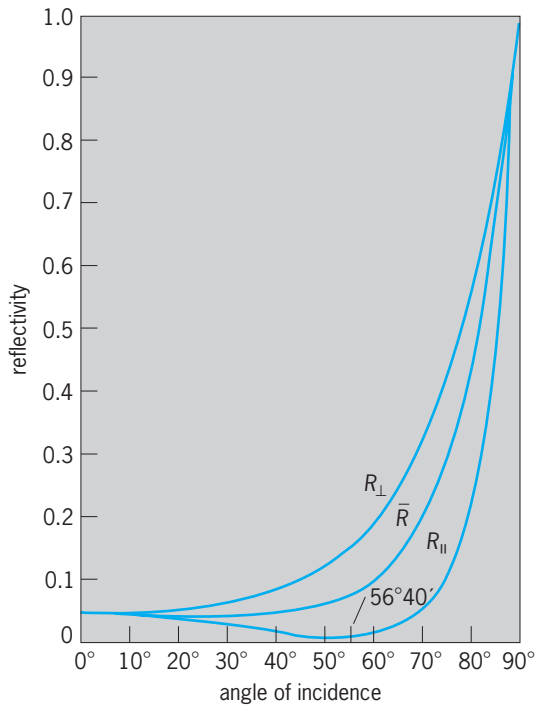

Fig. 4. Reflectivity as a function of angle of incidence for an air-glass interface. (*After M. Born and E. Wolf, Principles of Optics, 5th ed., Pergamon Press, 1975*)

When the electric field is polarized parallel to the plane of incidence, the reflectivity $R_\parallel$ is given by Eq. (5), and when the electric field is polarized per-

$$R_\parallel = \frac{\tan^2(\theta_1 - \theta_1)}{\tan^2(\theta_1 + \theta_2)} \tag{5}$$

pendicular to the plane of incidence, the reflectivity $R_\perp$ is given by Eq. (6). These formulas are known as

$$R_\perp = \frac{\sin^2(\theta_1 - \theta_2)}{\sin^2(\theta_1 + \theta_2)} \tag{6}$$

the Fresnel formulas. For unpolarized radiation, in which the electric field varies rapidly in a random, unpredictable manner, the reflectivity $\overline{R}$ is the average of $R_\parallel$ and $R_\perp$. **Figure 4** shows $R_\parallel$, $R_\perp$, and $\overline{R}$ for an air-glass interface as a function of angle. The reflectivity approaches 100% at grazing incidence ($\theta_1 \approx 90^\circ$) for both polarizations. *See* REFLECTION AND TRANSMISSION COEFFICIENTS.

At an angle of about $\theta_1 = \theta_B = 56^\circ$ the reflectivity $R_\parallel$ assumes zero value. This angle is called Brewster's angle, which is, in general, obtained from formula (7). At this angle, which is also called the

$$\tan\theta_B = \frac{n_2}{n_1} \tag{7}$$

polarizing angle, only radiation polarized perpendicular to the plane of incidence is reflected.

Total internal reflection. Total reflection, that is, reflection of 100% of the incident radiation, occurs at the interface of two dielectrics when the radiation is incident in the denser medium, that is, when $n_1 > n_2$ and when the angle of incidence θ_1 is larger than the critical angle θ_0 given by Eq. (8). Total internal

$$\sin\theta_0 = \frac{n_2}{n_1} \tag{8}$$

reflection can be observed, for example, by a submerged diver looking up at the water-air interface for which the critical angle is about $\theta_0 = 49^\circ$. For a glass-air interface the critical angle is approximately $\theta_0 = 42^\circ$.

Selective reflection from crystals. The discussion of reflection from dielectrics to this point has been concerned with reflectivity of nonabsorbing media far removed from absorption bands. These bands are located in the spectral regions where the frequency of the radiation corresponds to a resonance frequency of the atoms, molecules, or crystal lattice of the medium. Since this radiation is strongly absorbed, it is also strongly reflected. The metallic sheen of dye crystals, which have very strong absorption bands in the visible spectrum, is caused by selective reflection. Crystalline solids such as rocksalt or quartz, the lattices of which are built up of atoms bearing net electric charges, show strong selective reflection in the infrared region at wavelengths near those of the strong absorption bands associated with lattice vibrations in the crystal. By reflecting an infrared beam several times from such a material, highly monochromatic radiation can be obtained at the specific wavelengths. These monochromatic beams are referred to as residual rays or reststrahlen. **Figure 5** shows residual rays for some crystals. *See* IONIC CRYSTALS.

Antireflection coatings. In order to reduce undesired reflections from surfaces of optical components such as photographic lenses, one can coat the surface with a thin film. To minimize reflection, the refractive index n_f of this film should be the geometric mean of the indices n_1 and n_2 of the incident and refracting media, Eq. (9), and the film thickness

$$n_f^2 = n_1 n_2 \tag{9}$$

d_f should be equal to a quarter of the wavelength of light in the film material, that is, it should satisfy Eq. (10), where λ is the vacuum wavelength of the

$$n_f d_f = \frac{\lambda}{4} \tag{10}$$

light for which the reflectivity of the surface is minimized. Such a film is called a quarter-wave layer. Materials such as magnesium fluoride with an index of $n = 1.38$ are used for antireflection coatings have been made which reduce the reflectivity to less than 0.1% on certain materials; on glass surfaces less than $^1/_2$% has been achieved.

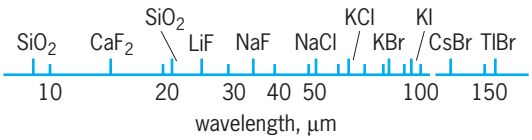

Fig. 5. Residual rays for various crystals.

High-reflectivity multilayers. A series of dielectric quarter-wave layers of alternating high and low refractive index can be used to produce reflectors of high reflectivity at a specified wavelength. This is called a quarter-wave stack. The high reflectivity is due to an interference effect. Stacks with about 20 layers are used to make reflectors with reflectives higher than 99%. Mirrors such as these are used for the construction of laser resonators. Multilayer coatings are also used to enhance the reflectivity of metals such as aluminum. About four quarter-wave layers can enhance the metal reflectivity to values better than 99%. *See* ALBEDO; GEOMETRICAL OPTICS; INTERFERENCE FILTER; MIRROR OPTICS; REFLECTION OF SOUND.

Herwig Kogelnik

Bibliography. *American Institute of Physics Handbook*, 3d ed., 1972; M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; B. D. Guenther, *Modern Optics*, 1990; E. Hecht and A. Zajac, *Optics*, 3d ed., 1997; F. A. Jenkins and E. White, *Fundamentals of Optics*, 4th ed., 1976; J. R. Meyer-Arendt, *Introduction to Classical and Modern Optics*, 1995.

Reflection of sound

The return of sound waves from surfaces on which they are incident. Suppose a sound wave strikes a smooth surface that is large compared to the wavelength of the sound and suppose that the path of this sound wave is represented by a ray, that is, by a line perpendicular to the advancing wavefront. By the law of reflection, the angle of reflection r for this ray equals the angle of incidence i, and the reflected ray lies in the plane of incidence. **Figure 1** shows reflection from a plane surface.

The geometrical laws for reflection of sound waves are the same as those for light waves. The apparent differences involve only questions of scale, because the average wavelength of sound is about 100,000 times that of light. For example, a mirror or lens used to produce a beam of sound waves must be enormously large compared to mirrors and lenses used in optical systems. *See* REFLECTION OF ELECTROMAGNETIC RADIATION.

A concave surface tends to concentrate the reflected sound waves (**Fig. 2**). Such surfaces are sometimes used to advantage as reflectors, but if used indiscriminately they may lead to poor acoustics as a result of undesirable focusing effects. Convex reflectors (**Fig. 3**) tend to spread the reflected waves. Therefore, when placed at the boundaries of a room, they tend to diffuse the sound throughout the room. For this reason, some radio-broadcasting studios employ cylindrical convex panels as part of their wall construction to promote diffusion.

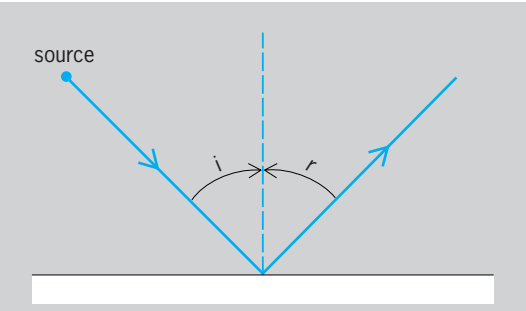

Fig. 1. Reflection from a plane surface.

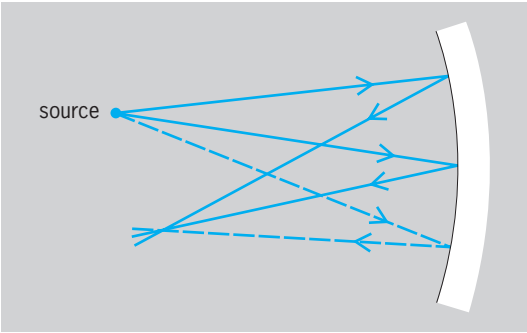

Fig. 2. Reflection from a concave spherical surface.

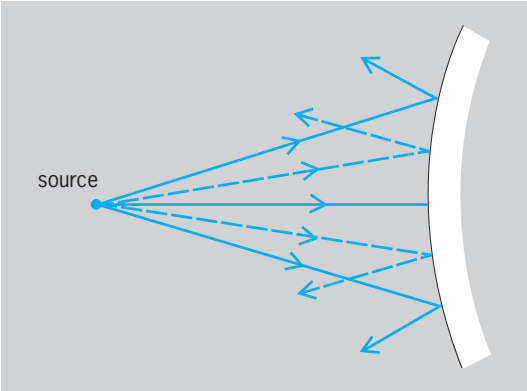

Fig. 3. Reflection from a convex surface.

The law of reflection is often used in investigating the effects of various shapes of a proposed room on the distribution of sound in that room. Such studies can lead to the design of interior surfaces that will give beneficial reflections, or to the elimination or modification of surfaces that otherwise would give rise to echoes. However, caution must be exercised in such applications of the law of reflection, because the wave properties of sound are neglected in this simplification of the behavior of sound. For a discussion of sound reflections in water *see* UNDERWATER SOUND; ARCHITECTURAL ACOUSTICS; ECHO; SOUND.

Cyril M. Harris

Reflex

A simple, unlearned, yet specific behavioral response to a specific stimulus. Reflexes are exhibited by virtually all animals from protozoa to primates. Along with other, more complex stimulus-bound responses such as fixed action patterns, they constitute much of the behavioral repertoire of invertebrates. In higher animals, such as primates, where learned behavior dominates, reflexes nevertheless persist as an important component of total behavior.

Reflexes are characteristic of a species and its relatives; that is, they are genetically determined. In humans, no effort is needed to acquire reflexes; they simply occur automatically, and many take place below the conscious level. For example, although heart rate and ventilation rate are constantly adjusted by reflexes, most of the time people are unaware of these changes. Similarly, pupillary diameter and blood pressure are reflexively regulated without conscious knowledge. Other reflex responses, however, such as perspiring, shivering, blinking, and maintaining posture, are more visible but occur without conscious intervention.

The simplest known reflexes require only one neuron or, in the strictest sense, none. For example, ciliated protozoa, which are single cells and have no neurons, nevertheless exhibit apparently reflexive behaviors. When a paramecium swims into an object, it reverses the power stroke of the cilia, backs away a short distance, turns, and again moves forward. When touched caudally, the animal accelerates in the forward direction. Here, the animal's cell membrane itself serves as the receptor of the stimulus, and the cilia act as effectors for directed movement. In higher organisms, there are also very simple reflexes. For example, when human skin is injured sufficiently to stimulate a single pain neuron, an unknown substance is released which causes small local blood vessels to dilate. In contrast to these simple responses, however, most reflexes require activity in a large circuit of neurons. *See* NEURON.

Anatomy and physiology. The neurons involved in most reflexes are connected by specific synapses to form functional units in the nervous system. Such a circuit begins with sensory neurons and ends with effector cells such as skeletal muscles, smooth muscles, and glands, which are controlled by motor neurons. The central neurons which are often interposed between the sensory and motor neurons are called interneurons. When first described, this sequence of neurons was called a reflex arc. The sensory side of the reflex arc conveys specificity as to which reflex will be activated. That is, the sensory cells themselves determine which environmental change is sensed, either within the body or without. The remainder of the reflex response is governed by the specific synaptic connections that lead to the effector neurons.

A familiar reflex is the knee-jerk or stretch reflex. It involves the patellar (kneecap) tendon and a group of upper leg muscles. Other muscle groups show similar reflexes, but this one is described here because of its widespread clinical use. It is a relatively simple reflex arc in which the terminals of the sensory cells synapse directly on the effector neurons. It is thus an example of a monosynaptic reflex (**Fig. 1**). When the patellar tendon is tapped, a brief stretch is imposed on the attached muscles (quadriceps femoris), including the muscle spindles which are embedded within the quadriceps. This stretch causes the sensory neurons of the spindles to fire impulses which return to the spinal cord. Within the spinal cord, excitatory synapses are activated on the specific motor neurons supplying the same muscles that were stretched, that is, the quadriceps. This results in a volley of impulses in the motor neurons which activate a brief contraction of the quadriceps, which in turn extends the leg. A possible function of stretch reflexes is to oppose sudden changes in length of skeletal muscle, and these reflexes are likely operating even during normal movements.

The stretch reflex as well as other more complex reflexes involve the inhibition of antagonistic muscle groups. This requires the activation of inhibitory interneurons in the spinal cord which synapse upon the motor neurons of antagonistic muscles. For example, although not shown in Fig. 1, the sensory neurons of the quadriceps also produce inhibition of motor neurons that contact antagonistic (flexor) muscles on the opposite side of the leg. In still other complex reflexes, there is a spread of activity to higher and lower levels of the spinal cord (multisegmental) and to the opposite side of the body (contralateral). Examples of reflexes with these characteristics include the flexor reflex and accompanying crossed extensor reflex (**Fig. 2**). These reflexes are initiated by a painful stimulus. The returning sensory signals excite the motor neurons of the leg flexor muscles while inhibiting those of the leg extensor muscles on the same side as the stimulus (ipsilateral side). This assures that the foot is quickly lifted away from the harmful stimulus. Balance, however, must also be maintained, and this is accomplished by exciting the leg extensors of the opposite (contralateral) limb; thus the term crossed extensor reflex. The flexor reflex is a prime example of a protective response (see below). It occurs so rapidly that the pain is felt only after the withdrawal response is complete. Speed of response is particularly important when the severity of the injury is time-dependent, such as in response to a chemical exposure or burn.

Although only a few neurons were drawn into Figs. 1 and 2 for clarity, many thousands are usually

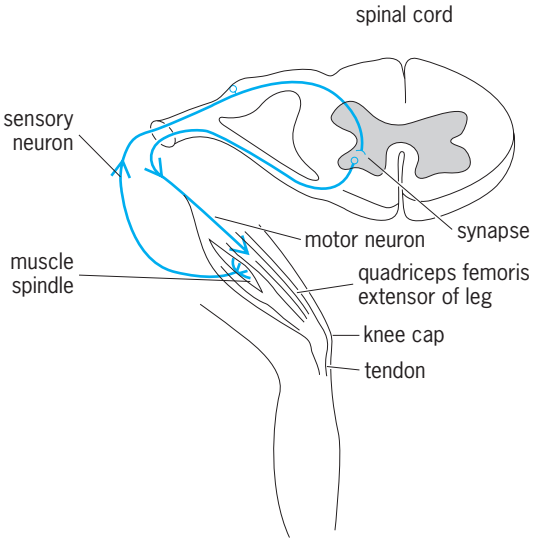

Fig. 1. Simplified diagram of the excitatory component of a stretch reflex. This reflex is said to be monosynaptic because there is a single synapse between the sensory neuron and motor neuron.

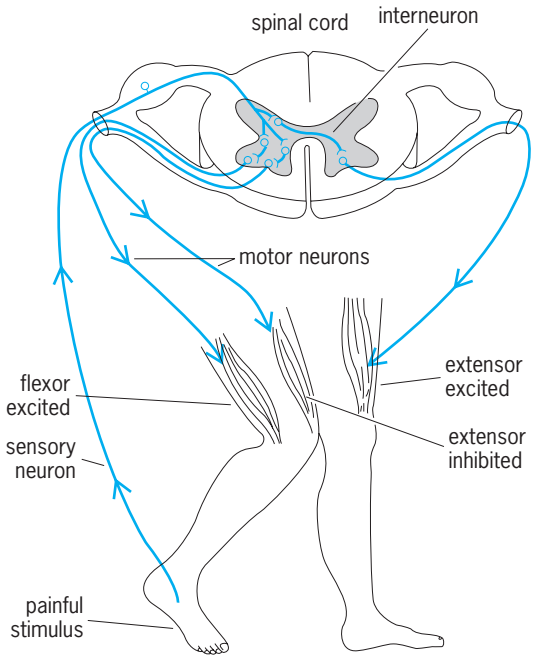

Fig. 2. Diagram of the flexor and crossed extensor reflexes. This reflex is said to be monosynaptic because there is a single synapse between the sensory and motor neuron.

involved. A few to many sensory neurons may activate several hundred to several thousand motor neurons, depending upon the reflex involved. Each motor neuron in turn branches to synapse on as many as several hundred muscle cells. In addition, in multisegmental reflexes and in those reflexes with bilateral effects, another thousand or more interneurons may become involved.

Functions. Reflexes adjust many important biological functions rapidly and efficiently without conscious effort, while other reflexes are largely protective. For example, the most delicate and sensitive sensory systems, the eye and the ear, can be damaged or destroyed by overstimulation or by accident. The amount of light admitted to the retina is controlled by the pupillary reflex, which in humans can effect a change in pupillary diameter from about 8 mm (0.30 in.) in darkness to 2 mm (0.08 in.) in bright light. A sudden flash of intense light can evoke the reflex closure of the eyelids, further protecting the delicate retina. The eyeball itself is protected from drying by the blink reflex and from mechanical injury by the eyelid closure reflex. The latter reflex is triggered when an object approaches the eye or when the lashes or the cornea are touched. The ear is similarly protected from potentially damaging sounds through the reflex contraction of middle-ear muscles in response to loud noise. This reflex serves to lower the efficiency with which sound vibrations are transmitted through the bones of the middle ear, and thus reduces the possibility of damage to the delicate hair cell receptors of the inner ear (cochlea). *See* EAR (VERTEBRATE); EYE (VERTEBRATE); HEARING (HUMAN).

Stereotyped escape responses constitute another category of protective reflexes exhibited by many animals. For example, a startled squid takes evasive action by contracting its mantle muscles, forcing a jet of water through its siphon. Fishes respond to waterborne vibrations by contracting muscle on one side of their bodies. The reflex contraction occurs on the side opposite the source of vibration, and the result is a sudden move away from possible harm.

Equally important for survival are the various feeding reflexes exhibited by animals. For example, flies and many other insects have chemoreceptors on their feet, mouthparts, and antennae. If a hungry fly walks on a surface moistened with nutrients, a set of reflexes is initiated. When the receptors on the feet are stimulated, there is a reflex extension of the proboscis. If the proboscis receptors are favorably stimulated, the animal begins to drink. Drinking continues until the crop is sufficiently distended to stimulate its stretch receptors. Finally, this stimulus initiates the reflex termination of feeding.

Instinctive or innate behaviors such as courtship rituals, nestbuilding, aggressive, and territorial behaviors show considerable similarities to reflexes. Although generally more complex, they are, like reflexes, unlearned, species-specific, genetically determined, and stereotypic in nature. Importantly, fixed action patterns such as these are like reflexes in being initiated by a specific stimulus, called a sign stimulus or a releaser. In the broad sense, both forms of behavior are also similar in that they are thought to be controlled by specific sets of neurons that underlie each behavior. Like the still more complex learned behaviors, the neural basis of the instinctive behaviors remains largely known. *See* INSTINCTIVE BEHAVIOR; REPRODUCTIVE BEHAVIOR; TERRITORIALITY.

Modification (plasticity). In general, the strength of a reflexive response depends upon the magnitude of the stimulus. However, many reflexes tend to gradually weaken if the stimulus is applied repeatedly. This phenomenon, called habituation, allows an animal to ignore a familiar or repeated stimulus. If a strong, novel stimulus is presented along with a reflex-evoking stimulus, however, the strength of the reflex is often enhanced. This process, called sensitization, causes the animal to attend to a potentially threatening situation by responding strongly to its environmental cues. Recent analysis of the changes in reflex circuits during habituation and sensitization have shown changes in synaptic transmission between neurons in the reflex. These changes are now understood in some detail at the cellular and molecular levels.

Similarly, reflexes may be modified by classical (pavlovian) conditioning, a form of associative learning. For example, in dogs, the sight and smell of food cause the reflexive secretion of saliva. A light or the sound of a bell does not. Ivan Pavlov showed that pairing the sight and smell of food (unconditioned stimuli) with a light or bell caused the animal to associate the light or bell with food. Thus, after several training sessions, presentation of the light or bell alone caused salivation. Whether classical conditioning provides a significant portion of learned behavior is still controversial. *See* CONDITIONED REFLEX.

Finally, reflexes are often modified during voluntary movement or locomotion. Motion itself

stimulates many sensory receptors and could therefore elicit reflexes which oppose the intended movement. Such reflexes are thought to be overridden or suppressed by commands from the brain in order for the desired movements to be executed.

Clinical uses. Since reflex responses are essentially invariant in healthy individuals, their examination can often provide valuable information about neurological disorders. Relatively simple tests, such as pupillary responses or auditory perceptions, can give information about the normal operation of brain functions. Scratching the skin of the foot, which leads normally to flexion of the toes, can show the status of the spinal circuits that control them. For example, if the toes extend and spread in response to scratching the sole of the foot (Babinski response), this is indicative of a spinal lesion in the pyramidal tracts descending from the motor centers of the cerebral cortex. When observations on these and many other reflexes are combined with data on paralyses, spasms, or losses of sensation, it is often possible to localize the site of the lesion. *See* NERVOUS SYSTEM (INVERTEBRATE); NERVOUS SYSTEM (VERTEBRATE); NEUROBIOLOGY. J. L. Larimer

Bibliography. J. M. Camhi, *Neuroethology*, 1984; R. Gilles et al. (eds.), *Advances in Comparative and Environmental Physiology*, vol. 2, 1988; J. Nolte, *Principles of Neurobiology*, int. ed., 1991; R. A. Rhodes and G. A. Tanner (eds.), *Medical Physiology*, 1995; G. M. Shepherd, *Neurobiology*, 3d ed., 1997; R. N. Singh and N. J. Strausfeld (eds.), *Neurobiology of Sensory Systems*, 1990.

Reforestation

The reestablishment of forest cover either naturally or artificially. Given enough time, natural regeneration will usually occur in areas where temperatures and rainfall are adequate and when grazing and wildfires are not too frequent. For example, Canada was basically treeless 16,000 years ago (during the last glacial period). During that time, spruce trees probably dominated much of present Illinois and Missouri. Due to global warming, spruce forests gradually spread north and now cover much of Canada. Within 16,000 years, more than 1 billion acres (400 million hectares) of forests in Canada were regenerated naturally.

Human impact. Rapid reforestation was not important when human populations were low in number and food was obtained by gathering wild plants and following game. Mature natural forests usually produced wood at a rate of 14–28 ft^3/(acre)(yr) [1–2 m^3/(ha)(yr)], and this often exceeded the rate of harvesting firewood. However, with the adoption of agriculture, humans settled down and their populations increased, placing pressures on forests. Not only did the demand for fuelwood increase, but in many places forests were cut down to make room for agriculture. For example, in 1600 about 49% of the continental United States was covered with forests, but this had been reduced to 33% by 1900. In just 300 years, population pressures reduced forestlands by about 1 million acres/yr (400,000 hectares/yr). Since natural regeneration was not keeping up with the rate of harvests, the need for rapid reforestation became more apparent.

During the late 1800s, a few foresters and government officials expressed concern over future wood supplies in the United States and the floods caused by deforestation. Two solutions to these problems were proposed. One resulted in the creation of national forests where wood and clean water could be produced in perpetuity. The other promoted artificial regeneration (the establishment of trees by planting or direct seeding). Due, in part, to a combination of improved technology in both agricultural and artificial regeneration, the United States now has about the same amount of forestland as it did 80 years ago. This is not true in many countries where human populations are increasing but artificial regeneration has been limited.

Planting and seeding. Reforestation occurs on land where trees have been recently removed due to harvesting or to natural disasters such as a fire, landslide, flooding, or volcanic eruption. When abandoned cropland, pastureland, or grasslands are converted to tree cover, the practice is termed afforestation (where no forest has existed in recent memory). Afforestation is common in countries such as Australia, South Africa, Brazil, India, and New Zealand. Although natural regeneration can occur on abandoned cropland, planting trees will decrease the length of time required until the first harvest of wood. Planting also has an advantage in that both tree spacing and tree species can be prescribed. The selection of tree species can be very important since it affects both wood quality and growth rates. Direct seeding is also used for both afforestation and reforestation, although it often is less successful and requires more seed than tree planting. Unprotected seed are often eaten by birds and rodents, and weeds can suppress growth of newly germinated seed. For these reasons, direct seeding accounts for only about 5% and 1% of artificial reforestation in Canada and the United States, respectively.

Growth rate improvements. In just a 5-year period (1990–1995), the world population increased by 0.4 billion and forests declined by 139 million acres (56 million hectares) due mostly to an increase in pastureland and cropland. Since the world population is expected to increase by more than 3 billion by the year 2050, some predict wood consumption will increase by 50%. How much this additional demand will deplete natural stands will depend, in part, on how much wood can be produced from tree plantations. The amount of plantation wood harvested in 2050 is predicted to meet half the demand. However, the actual amount will depend on improvements in growth rates and on the amount of afforestation (**Fig. 1**). Growth rates of plantations can be increased by selecting fast-growing species, by genetic selection, and by various cultural treatments.

Species selection. The objectives of the landowner determine which tree species are planted. Some

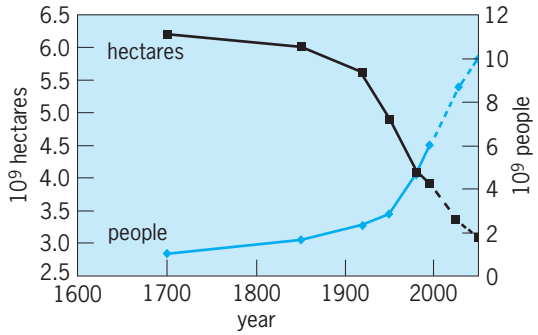

Fig. 1. Forests and people in the world.

species grow fast and straight, while others grow slow and crooked. If a landowner's objective is to produce fuelwood quickly, then an exotic species might be selected. In many cases, exotic species grow faster than native trees and sometimes can produce the same weight of wood from one-fifth the land. Some believe the faster growth rates of exotics is due to a lack of belowground pests.

Genetic improvement. Most developed countries have improvement programs for their most commonly planted tree species. Trees with rapid growth rates and desirable wood quality traits are bred in special areas called seed orchards. The progeny of various crosses are tested, and the best offspring are selected to be included in future seed orchards. In some cases when exceptional crosses are made, the genotypes are replicated using various vegetative propagation techniques. The most common method is to produce rooted cuttings. For example, about 30% of the pines planted in New Zealand are rooted cuttings. A more expensive method of vegetative propagation involves tissue culture. Although tissue culture techniques are used mainly by forest researchers, some companies in Brazil, Canada, and New Zealand are producing a limited amount of trees using either organogenesis or somatic embryogenesis. If the cost of these methods can be greatly reduced, it will be much easier to mass-produce genotypes with high growth rates. *See* FOREST GENETICS.

Seedling size. The best genetic improvements can be wasted if the nursery produces inferior seedlings. The size and condition of the seedling is related to its survival and growth after transplanting. Seedlings with small roots usually exhibit a limited amount of new root growth, undergo transplanting shock, and may die when soil moisture is limited. Small seedlings usually do not compete well with weeds and may require more intensive site preparation than seedlings with large root systems. In some regions of Canada, companies have moved away from applying herbicides to increase the growth of small seedlings. Instead, they are now planting large stock that can outgrow the weeds. It was once believed that money could be saved by purchasing small nursery stock. But it is now generally recognized that in many cases total establishment costs can be reduced by planting large stock instead of applying overly intensive mechanical site preparation to improve the growth of small seedlings.

Mechanical site preparation. On some sites, environmental conditions limit early growth of trees, and large machines are used to ameliorate the soil. The soil is sometimes too compacted, or too wet or too cold to allow rapid root growth. When the water table is close to the soil surface, ditches may need to be made to allow the soil to breathe. On some poorly drained sites, mounds or ridges are used to prevent young seedlings from being inundated during wet weather. Rippers are sometimes used to break up compacted soils. Although mechanical site preparation is sometimes used to control woody weeds, a reduction in site productivity can occur when part of the topsoil is removed during root-raking operations. In some regions the use of herbicides has reduced soil degradation caused when inappropriate mechanical methods are used.

Weed control. Foresters tend to group weeds into either herbaceous or woody. Herbaceous weeds include grasses and annuals, while woody weeds include shrubs and noncrop trees. Herbaceous weeds are usually found on afforestation sites and compete with trees mainly during the early years of the stand. Controlling herbaceous weeds at the time of planting can increase early growth and can shorten rotation lengths by 1 or 2 years. Controlling herbaceous weeds after the plantation is 10 years old will likely have no effect on wood production. In contrast, woody weeds can compete with crop trees throughout the rotation. Therefore, reducing the growth of woody weeds at either time of establishment or at midrotation can increase the growth of crop trees. In many cases, complete failures have occurred when woody sprouts were not controlled. Although it was once believed that a mixture of tree species would be the best economic approach, this is not the case since all woody species do not have the same value at harvest.

Fertilization. In the Southern Hemisphere, afforestation sites are usually fertile and growth can exceed 300 ft^3/(acre)(yr) [21 m^3/(ha)(yr)]. In contrast, in developed countries in the Northern Hemisphere, naturally fertile sites are usually occupied by agricultural crops and forest trees are planted on eroded and nutrient-poor sites. Some plantations on nutrient-poor sites produce only 140 ft^3/(acre)(yr) [10 m^3/(ha)(yr)]. Not only can growth be reduced, but wood quality can be reduced when certain micronutrient deficiencies exist. On phosphorus-deficient sites, fertilization can greatly increase yields of pine plantations. In the southern United States, past agricultural practices have greatly reduced topsoil depth, and although rainfall and temperatures are favorable, a limited amount of available nutrients restricts tree growth. Therefore, fertilization in pine plantations is conducted by forest industry. In total, forest industry uses only about 0.2% of all the nitrogen fertilizer used in the United States. Currently more than 1.2 million acres (485,000 ha) of plantations in this region were fertilized in 1998. It is estimated that more than half of all forest fertilization in the world is used on nutrient-deficient sites in the United States.

Fertilization can be made either at time of planting or at some point in midrotation. In general, phosphorus applications are made at the time of planting. Nitrogen applications are more beneficial after the tree roots have expanded their size and are better able to capture the added nitrogen. In either case, it is often wise to control weeds so they do not benefit from the fertilization more than the trees. *See* SOIL ECOLOGY.

Tree planting rates. Globally, the rate of tree planting gradually increased throughout the twentieth century. For example, only about 139,000 acres (56,000 ha) of land were planted in 1930, but now more than 2.5 million acres (1 million hectares) are planted annually in the United States. Some of this is due to afforestation since farmers are finding it more profitable in the long run to plant trees on cropland than to grow certain annual crops. Although tree planting is currently increasing on land owned by farmers and companies, the acreage of tree planting on federal land has been declining since 1990, when it peaked at 388,000 acres (157,000 ha). The planting rate by the federal government is now about half that amount, due mainly to a reduction in harvesting from National Forests. Governments in several other developed countries have also reduced their planting rate. Currently, a majority of industrial plantations are located in developed countries in Europe, North America, and the former Soviet Union. However, this may change in the future as governments in countries with expanding populations are continuing to encourage the planting of trees.

Fiber farms. Tree plantations can be classified into three groups according to management intensity. Most tree farms are managed with a low level of management intensity, where almost no inputs are made after the trees are 2 years old. These plantations might never receive any fertilization, weed control, insect control, or irrigation. This level of management is often used by a large number of private landowners who usually have limited amounts of capital. Medium-level management may involve herbicides and fertilization at establishment, and may also receive fertilizers one or two times during midrotation. Insect populations may be monitored, and preventative treatments may be applied when economical. Forest companies usually have the capital for this level of management. The increased growth rates allow the company to produce the same amount of wood on fewer acres (**Fig. 2**). This allows some land to be devoted to other objectives such as wildlife and streamside management. A high level of input on fiber farms increases the yields. Fiber farms typically grow hardwoods, have a 10-year or less rotation, and are irrigated and fertilized annually, and each farm is less than 25,000 acres (10,000 ha) in size. The management of fiber farms is closer to that of intensive horticulture than it is to traditional forestry. Currently less than 25 fiber farms exist in North America. But over time, as pressures by a growing population increase, the higher yields from fiber farms [400–800 ft^3/(acre)(yr) or 28–56 m^3/(ha)(yr)] will help to reduce the demand for wood that people place on natural forests [at 14–28 ft^3/(acre)(yr) or 1–2 m^3/(ha)(yr)]. *See* BREEDING (PLANT); FOREST AND FORESTRY; FOREST MANAGEMENT.

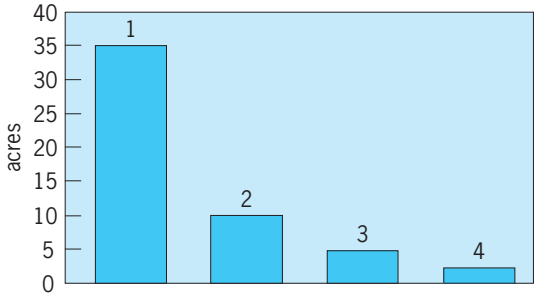

Fig. 2. Relative amounts of land needed to produce 1000 ft^3 of wood per year. (1) Mature, upland, natural hardwood stand. (2) Pine plantation owned by a farmer and managed with minimum inputs. (3) Pine plantation managed by a forest company and managed with medium inputs. (4) Hybrid-poplar fiber farm using trickle irrigation, fertilization, and weed control.

David B. South

Bibliography. J. R. Boyle et al., *Planted Forests: Contributions to the Quest for Sustainable Societies*, Kluwer Academic, Dordrecht, 1999; J. Evans, *Plantation Forestry in the Tropics*, Oxford University Press, Oxford, 1992; Food and Agriculture Organization of the United Nations, *State of the World's Forests*, FAO, Rome, 1999; T. Gardner-Outlaw and R. Engelman, *Forest Futures*, Population Action International, Washington, DC, 1999; R. W. Haynes, *An Analysis of the Timber Situation in the United States:* 1989–2040, USDA Forest Service, Washington, DC, 1990; J. Pöyry, *Global Outlook for Plantations*, Australian Bureau of Agricultural and Resource Economics, Canberra, 1999; M. Williams, *Americans and Their Forests: A Historical Geography*, Cambridge University Press, Cambridge, 1989.

Reforming processes

Those processes used to convert, with limited cracking (thermal decomposition), petroleum liquids into higher-octane gasoline. When the demand for higher-octane gasoline developed during the early 1930s, attention was directed to ways of improving the octane number of fractions within the boiling range of gasoline. Straight-run (distilled without change) gasoline frequently had low octane numbers, and any process that would improve its octane numbers would aid in meeting the demand for higher-octane gasoline. Such a process, called thermal reforming, was developed from thermal cracking processes and is used widely. *See* CRACKING; GASOLINE; OCTANE NUMBER.

The production of liquid product streams by distillation or by cracking processes is the first of a series of steps leading to the production of marketable liquid petroleum products. Several other unit processes are involved in the production of a final product. Such processes are generally called secondary processes or product improvement processes since they

are not used directly on the crude petroleum (oil) but are used on primary product streams that have been produced from the crude petroleum. Product improvement processes, such as reforming processes, are those processes in which the molecular structure of the feedstock is reorganized. Examples are the conversion (reforming, molecular rearrangement) of *n*-hexane to cyclohexane, and cyclohexane to benzene. These processes reform or rearrange one particular molecular type to another, thereby changing the properties of the product relative to the feedstock. *See* PETROLEUM.

Chemistry. Upgrading by reforming may be accomplished, in part, by an increase in volatility (reduction of molecular size) or by the conversion of *n*-paraffins to isoparaffins, olefins, and aromatics, and of naphthenes (cycloalkanes) to aromatics. The nature of the final product is influenced by the structure and composition of the straight-run (virgin) naphtha (hydrocarbon mixture) feedstock. In thermal reforming, the reactions resemble those in the cracking of gas oils. The molecular size is reduced, while olefins and some aromatics are synthesized. For example, an alkane can be converted to another alkane and an olefin [reaction (1), where $n > x + y$], or an alkane can be converted to a cycloalkane, which is in turn converted to an aromatic compound [reaction (2)]. Hydrocracking of high-molecular-weight paraffins to lower-molecular-weight paraffins and an olefin:

$$\underset{\text{Alkane}}{CH_3-(CH_2)_n-CH_3} \longrightarrow \underset{\text{Alkane}}{CH_3(CH_2)_x-CH_3} + \underset{\text{Olefin}}{CH_3-(CH_2)_y-CH=CH_2} \quad (1)$$

Dehydrocyclization of paraffin compounds to aromatic compounds:

$$\underset{\text{Alkane}}{CH_3-(CH_2)_4CH_3} \longrightarrow$$

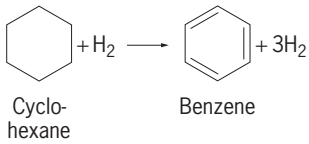

(2)

Hydrocracking of high-molecular-weight paraffins to lower-molecular-weight paraffins:

$$\underset{\text{Dodecane}}{CH_3(CH_2)_{10}CH_3} \xrightarrow{H_2} \underset{\text{Hexane}}{2CH_3(CH_2)_4CH_3}$$

$$\underset{\text{Dodecane}}{CH_3(CH_2)_{10}CH_3} \xrightarrow{H_2} \underset{\text{Octane}}{CH_3(CH_2)_6CH_3} + \underset{\text{Butane}}{CH_3(CH_2)_2CH_3} \quad (3)$$

Isomerization of *n*-paraffins to isoparaffins:

$$\underset{\text{Hexane}}{H_3C-CH_2-CH_2-CH_2-CH_2-CH_2} \longrightarrow$$

$$H_3C-CH_2-CH_2CH(CH_3)-CH_3 \quad (4)$$

IsoHexane

Isomerization of methylcyclopentane to cyclohexane:

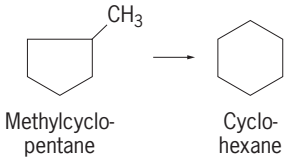

(5)

The liquid product (reformate) has the potential for further reaction by virtue of the presence of the olefinic constituents, but still contains appreciable quantities of these unsaturated materials and, as such, is classed as being unstable. Further hydrotreatment is required for stabilization. *See* PARAFFIN.

In the presence of catalysts and in the presence of the hydrogen available from dehydrogenation reactions, hydrocracking of paraffins to yield two lower-molecular-weight paraffins takes place, and olefins that do not undergo dehydrocyclization (dehydrogenation accompanied by formation of cyclic compounds) are dehydrogenated (with or without isomerization) so that the end product contains only traces of olefins. Examples of typical reactions are dehydrocyclization and dehydrogenation [reaction (2)], hydrocracking [reaction (3)], and isomerization [reactions (4) and (5)].

Other reactions such as the demethylation of branched paraffins to *n*-paraffins and the dealkylation of aromatics also occur. *See* DEHYDROGENATION; HYDROCRACKING; ISOMERIZATION.

Thermal reforming. Thermal reforming was a natural development from thermal cracking, since reforming is also a thermal decomposition reaction. Cracking converts heavier oils into gasoline constituents, whereas reforming converts (reforms) these gasoline constituents into higher-octane molecules. The equipment for thermal reforming is essentially the same as for thermal cracking, but higher temperatures are used.

In carrying out thermal reforming, a feedstock, such as a 205°C (400°F) end-point naphtha or a straight-run gasoline, is heated to 510–595°C (950–1100°F) in a furnace much the same as a cracking furnace, with pressures of 400–1000 lb/in.^{2} (2.8–6.9 megapascals). As the heated naphtha leaves the furnace, it is cooled or quenched by the addition of cold naphtha. The quenched, reformed material then enters a fractional distillation tower where any heavy products are separated. The remainder of the reformed material leaves the top of the tower and is separated into gases and reformate. The higher octane number of the product (reformate) is due primarily to the cracking of longer-chain paraffins into higher-octane olefins. *See* DISTILLATION.

The products of thermal reforming are gases, gasoline, and residual oil, the last formed in very small amounts (about 1%). The amount and quality of the reformate are very dependent on the temperature. As a rule, the higher the reforming temperature, the higher the octane number of the product but the lower the reformate yield. For example, a gasoline with an octane number of 35 when reformed at

515°C (960°F) yields 92.4% of 56-octane reformate; when reformed at 555°C (1030°F) the yield is 68.7% of 83-octane reformate. By using catalysts, as in the catalytic reforming processes described later, higher yields of much higher-octane gasoline can be obtained for a given temperature.

Thermal reforming is less effective and less economical than catalytic processes and has been largely supplanted. As it was practiced, a single-pass operation was used at 540–760°C (1000–1140°F) and 500–1000 lb/in.^{2} (3.4–6.9 MPa). The degree of octane-number improvement depended on the extent of conversion but was not directly proportional to the extent of cracking. The octane number was changed by the severity of the cracking, and the product had increased volatility, compared to the volatility of the feedstock. At high conversion, the production of coke and gas increased. The gases produced, although high in methane, generally contained olefins, and the process was generally accompanied by either a separate gas polymerization operation or one in which the C_3-C_4 gases, autogenous as well as extraneous, were added back to the reforming system.

Modifications of the thermal reforming process due to the inclusion of hydrocarbon gases with the feedstock are known as gas reversion and polyforming. Thus, olefinic gases produced by cracking and reforming can be converted into liquids boiling in the gasoline range by heating them under high pressure. Since the resulting liquids (polymers) have high octane numbers, they increase the overall quantity and quality of gasolines produced in a refinery.

The gases most susceptible to conversion to liquid products are olefins with three and four carbon atoms. These are propylene ($CH_3CH{=}CH_2$), which is associated with propane in the C_3 fraction, and butylene ($CH_3CH_2CH{=}CH_2$ or $CH_3CH{=}CHCH_3$) and isobutylene [$(CH_3)_2C{=}CH_2$], which are associated with butane ($CH_3CH_2CH_2CH_3$) and isobutane [$(CH_3)_2CHCH_3$] in the C_4 fraction. When the C_3 and C_4 fractions are subjected to the temperature and pressure conditions used in thermal reforming, they undergo chemical reactions that result in a small yield of gasoline. When the C_3 and C_4 fractions are passed through a thermal reformer as a mixture with naphtha, the process is called naphtha-gas reversion or naphtha polyforming.

Reversion and polyforming are essentially the same but differ in the manner in which the gases and naphtha are passed through the heating furnace. In gas reversion, the naphtha and gases flow through separate lines in the furnace and are heated independently of one another. Before leaving the furnace, both lines join to form a common soaking section where the reforming, polymerization, and other reactions take place. In naphtha reforming, the C_3 and C_4 gases are premixed with the naphtha and pass together through the furnace. Except for the gaseous components in the feedstock, both processes operate in much the same manner as thermal reforming and produce similar products.

Catalytic reforming. Like thermal reforming, catalytic reforming converts low-octane gasoline into high-octane gasoline (reformate). Although thermal reforming can produce reformate with a research octane number of 65–80 depending on the yield, catalytic reforming produces reformate with octane numbers on the order of 90–95. Catalytic reforming is conducted in the presence of hydrogen over hydrogenation-dehydrogenation catalysts, which may be supported on alumina or silica-alumina. Depending on the catalyst, a definite sequence of reactions takes place, involving structural changes in the feedstock. Furthermore, this process has rendered thermal reforming obsolete. *See* CATALYSIS.

Although subsequent olefin reactions occur in thermal reforming, the product contains appreciable amounts of unstable unsaturated compounds. In the presence of catalysts and hydrogen (available from dehydrogenation reactions), hydrocracking of paraffins to yield two lower paraffins occurs. Olefins that do not undergo dehydrocyclization are also produced. The olefins are hydrogenated with or without isomerization, so that the end product contains only traces of olefins.

The addition of a hydrogenation-dehydrogenation catalyst to the system yields a dual-function catalyst complex. Hydrogen reactions—hydrogenation, dehydrogenation, dehydrocyclization, and hydrocracking—take place on one catalyst, and cracking, isomerization, and olefin polymerization take place on the other (acid) catalyst. *See* HYDROCRACKING.

Under the high-hydrogen partial pressure conditions used in catalytic reforming, sulfur compounds are readily converted into hydrogen sulfide which, unless removed, builds up to a high concentration in the recycle gas. Hydrogen sulfide is a reversible poison for platinum catalysts and decreases the catalysts' dehydrogenation and dehydrocyclization activity. In the first catalytic reformers, hydrogen sulfide was removed from the gas cycle stream by absorption in diethanolamine. Sulfur is generally removed from the feedstock by use of a conventional desulfurization over cobalt-molybdenum catalyst. An additional benefit of desulfurization of the feed to a level of <5 ppm sulfur is the elimination of hydrogen sulfide (H_2S) corrosion problems in the heaters and reactors.

Organic nitrogen compounds are converted into ammonia under reforming conditions, and this neutralizes acid sites on the catalyst, thus repressing its activity for isomerization, hydrocracking, and dehydrocyclization reactions. Straight-run materials do not usually present serious problems with regard to nitrogen, but feeds such as coker naphtha may contain around 50 ppm nitrogen and removal of this quantity may require high-pressure hydrogenation (800–1000 lb/in.^{2}; 5.5–6.9 MPa) over a nickel-cobalt-molybdenum on alumina catalyst.

The commercial catalytic processes can be broadly classified as the moving-bed, fluid-bed, and fixedbed types. The fluid-bed and moving-bed processes use mixed nonprecious metal oxide catalysts in units equipped with separate regeneration facilities. Fixed-bed processes use predominantly

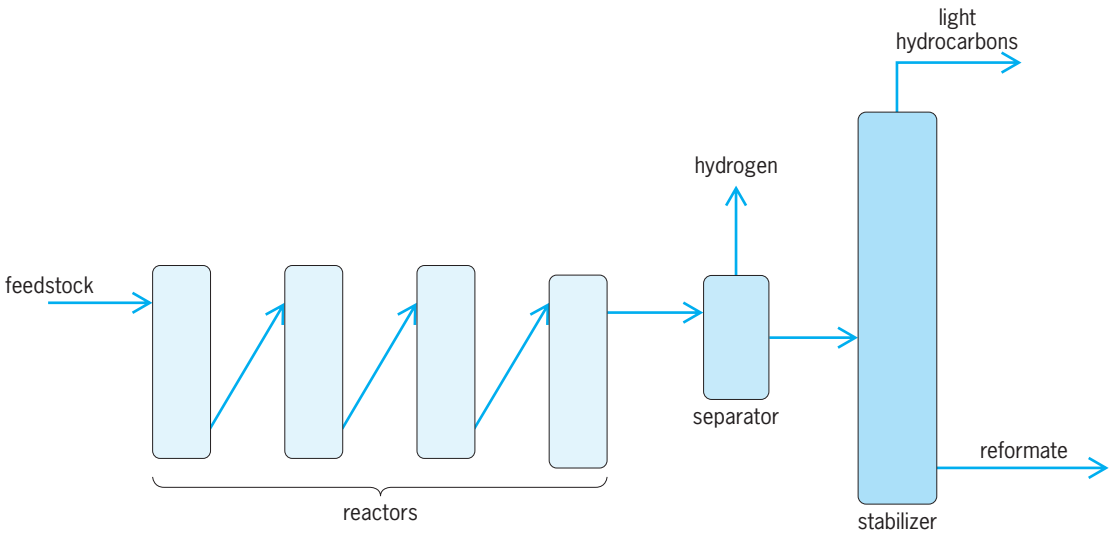

Catalytic reforming process.

platinum-containing catalysts in units equipped for cycle, occasional, or no regeneration.

Catalytic reformer feeds are saturated (not olefinic) materials. In most cases, the feed is a straight-run naphtha; however, other by-product low-octane naphtha (such as coker naphtha) can be processed after treatment to remove olefins and other contaminants. Hydrocarbon naphtha, which contains substantial quantities of naphthenes, is also a suitable feed.

The yield of gasoline of a given octane number and at given operating conditions depends on the hydrocarbon types in the feed. For example, high-naphthene stocks, which readily give aromatic gasoline, are the easiest to reform and give the highest gasoline yields. Paraffinic stocks which depend on the more difficult isomerization, dehydrocyclization, and hydrocracking reactions, require more severe conditions and give lower gasoline yields than the naphthenic stocks. The end points of these feeds are usually limited to about 190°C (375°F), partially because of increased coke deposition on the catalyst at the end point during processing at about 15°C (27°F). Limiting the feed end point avoids redistillation of the product to meet the gasoline end-point specification of 205°C (400°F), maximum.

Dehydrogenation is a main chemical reaction in catalytic reforming, and hydrogen gas is consequently produced in large quantities. The hydrogen is recycled through the reactors where the reforming takes place to provide the atmosphere necessary for the chemical reactions; and the hydrogen also prevents carbon from being deposited on the catalyst, thus extending its operating life. An excess of hydrogen above whatever is consumed in the process is produced, and as a result, catalytic reforming processes are unique in that they are the only petroleum refinery processes to produce hydrogen as a by-product. *See* DEHYDROGENATION; HYDROGENATION.

Catalytic reforming is usually carried out by feeding a naphtha (after pretreating with hydrogen if necessary) and hydrogen mixture to a furnace where the mixture is heated to 450–520°C (840–965°F) and then passed through fixed-bed catalytic reactors at hydrogen pressures of 100–1000 lb/in.^{2} (0.69–6.9 MPa). Normally two (or more) reactors are used in series, and reheaters are located between adjoining reactors to compensate for the endothermic reactions taking place. Sometimes as many as four or five are kept on-stream in series while one or more is being regenerated. The on-stream cycle of any one reactor may vary from several hours to many days, depending on the feedstock and reaction conditions (see **illus.**).

The product from the last catalytic reactor is cooled and sent to a high-pressure separator where the hydrogen-rich gas is split into two streams: one stream goes to recycle, and the other stream is excess hydrogen available for other uses. The excess hydrogen is vented from the unit and used in hydrotreating, as a fuel, or for the manufacture of chemicals (such as ammonia). The liquid product (reformate) is stabilized (by removal of light ends) and used directly in gasoline or extracted for aromatic blending stocks for aviation fuel. *See* AIRCRAFT FUEL.

Fixed-bed processes. Reforming processes that use catalysts arranged in beds are known as fixed-bed processes.

Hydroforming. The hydroforming process uses molybdena-alumina (MoO_2-Al_2O_3) catalyst pellets arranged in mixed beds. The hydroformer has four reaction vessels or catalyst cases, two of which are on the process cycle, while the other two are being regenerated. Naphtha feed is preheated to 400–540°C (900–1000°F) and passed in series through the two catalyst cases under a pressure of 150–300 lb/in.^{2} (1–2 MPa). Gas containing 70% hydrogen produced by the process is passed through the catalyst cases with the naphtha. The material leaving the final catalyst case enters a four-tower system where it is fractionally distilled as hydrogen-rich gas, a product (reformate) suitable for gasoline, and an aromatic polymer boiling above 205°C (400°F).

After 4–16 h on process cycle, the catalyst is regenerated. This is done by burning carbon deposits from the catalyst at 565°C (1050°F) by blowing air diluted with flue gas through the catalyst. The air also reoxidizes the reduced catalyst (9% molybdenum oxide on activated alumina pellets) and removes sulfur from the catalyst.

Platforming. The first step in the platforming process is preparation of the naphtha feed. For gasoline manufacture, the naphtha feed is distilled to separate a fraction boiling in the 120–205°C (250–400°F) range. Since sulfur adversely affects the platinum catalyst, the naphtha fraction may be treated to remove sulfur compounds. Otherwise, the hydrogen-rich gas produced by the process, which is cycled through the catalyst cases, must be scrubbed free of its hydrogen sulfide content.

The prepared naphtha feed is heated to 455–540°C (850–1000°F) and passed into a series of three catalyst cases under a pressure of 200–1000 lb/in.^{2} (1.4–6.9 MPa). Further heat is added to the naphtha between each of the catalyst cases in the series. The material from the final case is fractionated into hydrogen-rich gas and reformate. The catalyst is composed of 1/8-in. (0.3-cm) pellets of alumina containing chlorine and about 0.5% platinum. Each pound of catalyst reforms up to 100 barrels (16 m^3) of naphtha before losing its activity. It is possible to regenerate the catalyst, but it is more usual to replace the spent catalyst with new catalyst.

Other fixed-bed processes include catforming, in which the catalyst is a platinum, alumina, silica-alumina composition, which results in very high hydrogen purity. Regeneration to prolong catalyst life is practiced on a blockout basis (one of several reactions is taken off stream) with a dilute air in-stream mixture. In addition, Houdriforming is a process in which the catalyst may be regenerated, if necessary, on a blockout basis. A guard case catalytic hydrogenation pretreating stage using the same Houdry catalyst as the Houdriformer reactors is available for high-sulfur feedstocks. Lead and copper salts are also removed under the mild conditions of the guard case operation.

Powerforming. The cyclic Powerforming process is based on frequent catalyst regeneration (carbon burn-off) and permits continuous operation. Reforming takes place in four or five reactors, and regeneration is carried out in the last, or swing, reactor. Thus the plant need not be shut down to regenerate a catalyst reactor. The cyclic process assures a continuous supply of hydrogen gas for hydrorefining operations and tends to produce a greater yield of higher-octane reformate. The choice between the semiregenerative process and the cyclic process depends on the size of plant required, the type of feedstocks available, and the octane number needed in the product.

Iso-Plus Houdriforming. This is a combination process using a conventional Houdriformer operated at moderate severity, in conjunction with one of three alternatives: (1) conventional catalytic reforming plus aromatic extraction and separate catalytic reforming of the aromatic product; (2) conventional catalytic reforming plus aromatic extraction and recycling of the aromatic product aligned to the reforming state; and (3) conventional catalytic reforming followed by thermal reforming of the Houdriformer product and catalytic polymerization of the C_3 and C_4 olefins from thermal reforming. A typical feedstock for this type of unit is naphtha, and the use of a Houdry guard case permits charging stocks of relatively high sulfur content.

Rexforming. Rexforming is a combination process using platforming and aromatic extraction processes in which low-octane raffinate is recycled to the platformer. Operating temperatures may be as much as 27°C (50°F) lower than conventional platforming, and higher space velocities (rates of feedstock introduction into reactor) are used. A balance is struck between hydrocyclization and hydrocracking, excessive coke and gas formation thus being avoided. The glycol solvent in the aromatic extraction section is designed to extract low-boiling high-octane isoparaffins as well as aromatics.

Fixed-bed swing reactor. Fixed-bed, continuous catalytic reforming can be classified by catalyst type: (1) cyclical regenerative with nonprecious metal oxide catalysts and (2) cyclic regenerative with platinum-alumina catalysts. Both types use swing reactors to regenerate a portion of the catalyst while the remainder stays on-stream.

The cyclic regenerative fixed-bed, operation, using a platinum catalyst, is basically a low-pressure process (250–350 lb/in.^{2}; 1.7–2.4 MPa) which gives higher gasoline yields because of fewer hydrocracking reactions, as well as higher-octane products from a given naphtha charge and better hydrogen yields because of more dehydrogenation and fewer hydrocracking reactions. The coke yield, with attendant catalyst deactivation, increases rapidly at low pressures.

Selectoforming. The Selectoforming process uses a fixed-bed reactor operating under a hydrogen partial pressure. Typical operating conditions depend on the process configuration but are in the ranges 200–600 lb/in.^{2} (1.4–4.1 MPa) and 315–450°C (600–900°F). The catalyst used in the Selectoforming process is nonnoble metal with a low potassium content. As with the large-pore hydrocracking catalysts, the cracking activity increases with decreasing alkali metal content.

There are two configurations of the Selectoforming process that are being used commercially. The first Selectoformer was designed as a separate system and integrated with the reformer only to the extent of having a common hydrogen system. The reformer naphtha is mixed with hydrogen and passed into the reactor containing the shape-selective catalyst. The reactor effluent is cooled and separated into hydrogen, liquid petroleum, gas, and high-octane gasoline. The removal of *n*-paraffins reduces the vapor pressure of the reformate since these paraffins are in higher concentration in the front end of the feed. The separate Selectoforming system has the additional flexibility of being able to process other refinery streams.

The second process modification is the terminal reactor system. In this system, the shape-selective catalysts replace all or part of the reforming catalyst in the last reforming reactor. Although this configuration is more flexible, the high reforming operating temperature causes butane and propane cracking and consequently decreases the liquid petroleum gas yield and generates higher ethane and methane production. The life of a Selectoforming catalyst used in a terminal system is between 2 and 3 years, and regeneration only partially restores its catalytic activity.

Moving-bed processes. In moving-bed processes, the catalyst moves downward through the reactor by gravity flow and returns to the top by means of a solids-conveying technique (hyperflow)—this moves the catalyst at a low velocity and with minimum loss.

Hyperforming. Hyperforming is a moving-bed reforming process that uses catalyst pellets of cobalt molybdate with a silica-stabilized alumina base. Feedstock (naphtha vapor) and recycle gas flow upward, countercurrent to the catalyst, and regeneration of catalyst is accomplished in either an external vertical lift line or a separate vessel.

Hyperforming naphtha (65–230°C; 150–450°F) can result in improvement of the gasoline component; in addition, sulfur and nitrogen removal is accomplished. Light gas oil stocks can also be charged to remove sulfur and nitrogen under mild hydrogenation conditions for the production of premium diesel fuels and middle distillates. Operating conditions in the reactor are 400 lb/in.^{2} (2.8 MPa) and 425–480°C (800–900°F), the higher temperature being used for a straight-run naphtha feedstock; catalyst regeneration takes place at 510°C (950°F) and 415 lb/in.^{2} (2.9 MPa).

Thermofor catalytic reforming (TCR). This is a moving-bed process that uses a synthetic bead, coprecipitated chromia (CrO_2) and alumina (Al_2O_3) catalyst. Catalyst-to-naphtha ratios have little effect on product yield or quality when varied over a wide range. The catalyst flows downward through the reactor, and the naphtha-recycle gas feed enters the center of the reactor. The catalyst is transported from the base of the reactor to the top of the regenerator by bucket-type elevators.

Fluid-bed processes. In catalytic reforming processes using a fluidized bed, the catalyst is continuously regenerated with a separate or integrated reactor to maintain catalyst activity by coke and sulfur removal. Cracked or virgin naphtha is charged with hydrogen-rich recycle gas to the reactor. A molybdena (Mo_2O_3, 10.0%) on alumina catalyst, not materially affected by normal amounts of arsenic, iron, nitrogen, or sulfur, is used. Operating conditions in the reactor are about 200–300 lb/in.^{2} (1.4–2.0 MPa) and 480–950°C (900–950°F). *See* FLUIDIZATION.

Fluidized-bed operation with its attendant excellent temperature control prevents over- and underreforming operations, resulting in more selectivity in the conditions needed for optimum yield of the desired product.

Catalysts. Reforming reactions typically proceed through a number of elementary steps. For example, a straight-chain paraffin is converted into an isoparaffin by first being converted into an olefin, which is isomerized to an isoolefin, and then converted into an isoparaffin. Correspondingly, the catalyst has two functions: a hydrogenation-dehydrogenation function for the paraffin-olefin conversions and an isomerization function, which is associated with the catalyst acidity. The catalysts used until the early 1950s were chromium oxide or molybdenum oxide supported on alumina, which incorporated both the catalytic functions on the surface of the metal oxide. More recently developed reforming catalysts have crystallites of a metal such as platinum on an acidic support such as alumina, and the two functions are present in separate phases. The metal (palladium, Pd; platinum, Pt; or rhenium, Re; or a noble-metal-containing trimetallic alloy) provides the hydrogenation-dehydrogenation activity, and the promoted acidic alumina provides the isomerization activity. The hydrogenation-dehydrogenation activity of the supported metal and the isomerization activity of the alumina are much greater than the respective activities of the early-generation metal oxides.

The composition of a reforming catalyst is dictated by the composition of the feedstock and the desired reformate. The catalysts used are principally molybdena-alumina (MoO_2-Al_2O_3), chromia-alumina (Cr_2O_3-Al_2O_3), or platinum on a silica-alumina (SiO_2-Al_2O_3) or alumina (Al_2O_3) base. The nonplatinum catalysts are widely used in regenerative process for feeds containing sulfur, which poisons platinum catalysts, although pretreatment processes (such as hydrodesulfurization) may permit platinum catalysts to be used.

The purpose of platinum on the catalyst is to promote dehydrogenation and hydrogenation reactions, that is, the production of aromatics, participation in hydrocracking, and rapid hydrogenation of carbon-forming precursors. For the catalyst to have an activity for isomerization of both paraffins and naphthenes—the initial cracking step of hydrocracking—and to participate in paraffin dehydrocyclization, it must have an acid activity. The balance between these two activities is most important in a reforming catalyst.

In the production of aromatics from cyclic saturated materials (naphthenes), it is important that hydrocracking be minimized to avoid loss of the desired product. Thus, the catalytic activity must be moderated relative to the case of gasoline production from a paraffinic feed, where dehydrocyclization and hydrocracking play an important part.

The acid activity can be obtained by means of halogens (usually fluorine or chlorine up to about 1% by weight in catalyst) or silica incorporated in the alumina base. The platinum content of the catalyst is normally in the range 0.3–0.8 wt %. At higher levels there is some tendency to effect demethylation and naphthene ring opening, which is undesirable; at lower levels the catalysts tend to be less resistant to poisons.

Most processes have a means of regenerating the catalyst as needed. The time between regeneration, which varies with the process, the severity of the reforming reactions, and the impurities of the feedstock, ranges from every few hours to several months. Several processes use a nonregenerative catalyst that can be used for a year or more, after which it is returned to the catalyst manufacturer for reprocessing. In processes that have moving beds, the catalysts are continuously regenerated in separate reactors.

The processes using bauxite (cycloversion) and clay (isoforming) differ from other catalytic reforming processes in that hydrogen is not formed and hence none is recycled through the reactors. Since hydrogen is not concerned in the reforming reactions, there is no limit to the amount of olefin that may be present in the feedstock. The cycloversion process is also used as a catalytic cracking process and as a desulfurization process. The isoforming process causes only a moderate increase in octane number.

The modern reforming catalysts in use today contain platinum supported on a silica or silica-aluminum base. In most cases, rhenium is combined with platinum to form a more stable catalyst that permits operation at lower pressures. Platinum is thought to serve as the catalytic site for hydrogenation and dehydrogenation reactions, and chlorinated alumina provides the acidic site for isomerization, cyclization, and hydrocracking reactions. Reforming catalyst activity is a function of surface area, pore volume, and active platinum and chlorine content. Catalyst activity is reduced during operation by coke deposition and chloride loss. In a high-pressure process, up to 20 barrels (3.2 m^3) of charge can be processed per pound of catalyst before regeneration is needed. The activity of the catalyst can be restored by high-temperature oxidation of the carbon followed by chlorination. This type of process is referred to as semiregenerative and is able to operate for 6- to 24-month periods between regenerations. The activity of the catalyst decreases during the on-stream period and the reaction temperature is increased as the catalyst ages to maintain the desired operating severity. Normally the catalyst can be regenerated at least three times before it has to be replaced and returned to the manufacturer for reclamation. *See* HETEROGENEOUS CATALYSIS; PETROLEUM PROCESSING AND REFINING.

James G. Speight

Bibliography. G. J. Antos, A. M. Aitani, and J. M. Parera (eds.), *Catalytic Naphtha Reforming: Science and Technology*, Marcel Dekker, New York, 1995; J. G. Speight, *The Chemistry and Technology of Petroleum*, 3d ed., Marcel Dekker, New York, 1999.

Refraction of waves

The change of direction of propagation of any wave phenomenon which occurs when the wave velocity changes. The term is most frequently applied to visible light, but it also applies to all other electromagnetic waves, as well as to sound and water waves.

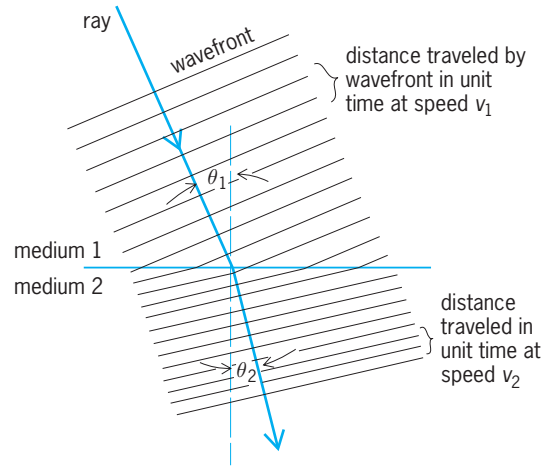

Fig. 1. Physical basis for Snell's law.

The physical basis for refraction can be readily understood with the aid of **Fig. 1**. Consider a succession of equally spaced wavefronts approaching a boundary surface obliquely. The direction of propagation is in ordinary cases perpendicular to the wavefronts. In the case shown, the velocity of propagation is less in medium 2 than in medium 1, so that the waves are slowed down as they enter the second medium. Thus the direction of travel is bent toward the perpendicular to the boundary surface (that is, $\theta_2 < \theta_1$). If the waves enter a medium in which the velocity of propagation is faster than in their original medium, they are refracted away from the normal.

Snell's law. The simple mathematical relation governing refraction is known as Snell's law. If waves traveling through a medium at speed υ_1 are incident on a boundary surface at angle θ_1 (with respect to the normal), and after refraction enter the second medium at angle θ_2 (with the normal) while traveling at speed υ_2, then Eq. (1) holds. The index of

$$\frac{\upsilon_1}{\upsilon_2} = \frac{\sin\theta_1}{\sin\theta_2} \tag{1}$$

refraction n of a medium is defined as the ratio of the speed of waves in vacuum c to their speed in the medium. Thus $c = n_1\upsilon_1 = n_2 n_2$, and therefore Eq. (2)

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \tag{2}$$

holds. The refracted ray, the normal to the surface, and the incident ray always lie in the same plane.

The relative index of refraction of medium 2 with respect to that of medium 1 may be defined as $n = n_1/n_2$. Snell's law then becomes Eq. (3). For

$$\sin\theta_1 = n \sin\theta_2 \tag{3}$$

sound and other elastic waves requiring a medium in which to propagate, only this last form has meaning. Equation (3) is frequently used for light when one medium is air, whose index of refraction is very nearly unity.

When the wave travels from a region of low velocity (high index) to one of high velocity (low index), refraction occurs only if $(n_1/n_2)\sin\theta_1 \leqq 1$. If θ_1 is too large for this relation to hold, then $\sin\theta_2 > 1$, which is meaningless. In this case the waves are totally reflected from the surface back into the first medium. The largest value that θ_1 can have without total internal reflection taking place is known as the critical angle θ_c. Thus $\sin\theta_c = n_2/n_1$. When the angle of incidence $\theta_1 < \theta_c$, refraction occurs, as in **Fig. 2***a*. When $\theta_1 = \theta_c$, the emergent ray just grazes the surface (Fig. 2*b*). Total internal reflection (Fig. 2*c*) represents the only practical case for which 100% of the incident energy is reflected and none is absorbed. When it is desired to change the direction of a beam of light without loss of energy, totally reflecting prisms are often used, as in prism binoculars.

If waves travel through a medium having a continuously varying index of refraction, the rays follow smooth curves with no abrupt changes of direction. Suppose (**Fig. 3**) that $n = n(y)$, and that the incident ray lies in the xy plane. If θ is the angle between the direction of the ray and the y axis, then Snell's law can be written in the differential form given by Eq. (4). In a particular case Eq. (4) can be integrated

$$\frac{d\theta}{dn} = -\frac{1}{n}\tan\theta \tag{4}$$

to give the path of the ray.

Visible light. Many interesting cases of refraction occur for visible light. The refraction of light by a prism in air affords a particularly simple and useful example. *See* LIGHT.

As a ray passes through the prism of **Fig. 4**, its total deflection or deviation $D = \theta_1 + \theta_4 - A$, where A is the vertex angle of the prism. Also, by Snell's law, Eq. (5) holds. It is found that the deviation is a minimum when the ray passes through the prism symmetrically (that is, when $\theta_1 = \theta_4$). For minimum deviation, Eq. (6) holds. For a given prism, the dispersion,

$$n = \frac{\sin {}^1/_2(A + D)}{\sin {}^1/_2 A} \tag{5}$$

$$n = \frac{\sin\theta_1}{\sin\theta_2} = \frac{\sin\theta_4}{\sin\theta_3} \tag{6}$$

or lateral spread of the spectrum formed is maximum for that wavelength of light which passes through the prism at minimum deviation. *See* OPTICAL PRISM.

For most optical materials, the dispersion $dn/d\lambda$ (λ is the wavelength) is negative; therefore red light is bent less than blue light. Typical values of n for optical materials range from 1.5 for ordinary crown glass, 1.7 or 1.8 for dense flint glass, and up to 2.42 for diamond. For water, n is 1.33. Some special substances have even higher values. Many substances show anisotropy in the refraction of light, with different indices of refraction in different directions. *See* OPTICAL MATERIALS.

For a lens (**Fig. 5**), refraction occurs at both surfaces. If the lens is thin and the rays all make small angles with the axis of the system, application of Snell's law to the two spherical surfaces yields the well-known lens formula, Eq. (7), where s is the object distance from the lens, s' the image distance, and f the focal length of the lens. Magnifying instruments such as binoculars, telescopes, microscopes, and projectors make use of refraction by lenses or prisms in their operation. *See* LENS (OPTICS).

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f} \tag{7}$$

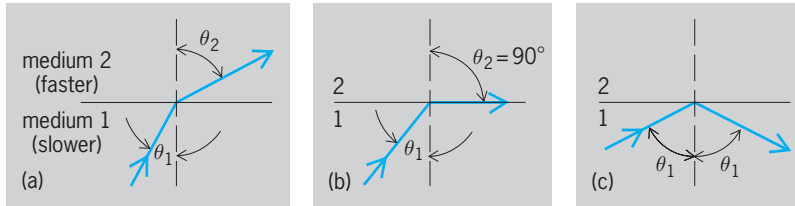

Fig. 2. Behavior of ray traveling from medium of high refractive index to medium of low refractive index. (*a*) When $\theta_1 < \theta_c$, ray is refracted. (*b*) When $\theta_1 = \theta_c$, ray grazes surface. (*c*) When $\theta_1 > \theta_c$, ray is reflected.

Double refraction. Some anisotropic single crystals such as those of calcite and quartz are birefringent, or doubly refracting. If one looks through such a crystal at a dot on a piece of paper, two images are seen. As the crystal is rotated in the plane of the paper, one image remains stationary while the other appears to rotate about it. *See* BIREFRINGENCE.

Two separate rays propagate through the crystal; they are called the ordinary ray and the extraordinary ray. These rays are linearly polarized at right angles to each other. The ordinary ray obeys Snell's

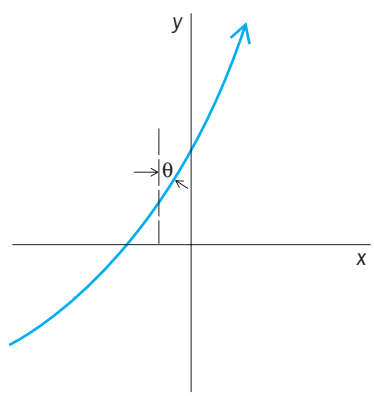

Fig. 3. Path of light in medium having continuously varying index of refraction $n = n(y)$.

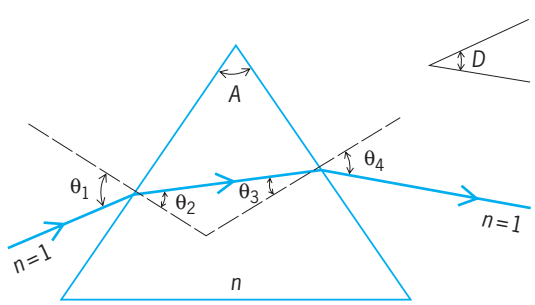

Fig. 4. Refraction of light by a prism.

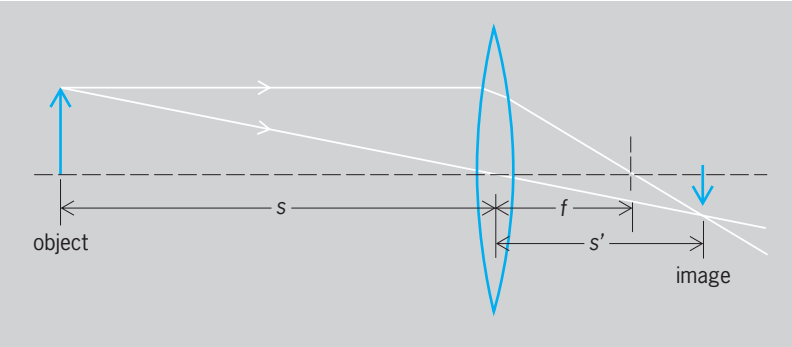

Fig. 5. Refraction of light by a lens.

law; the extraordinary ray in general does not. The extraordinary ray does not propagate perpendicularly to its wavefronts. The separation between the two rays depends upon the direction in which the light travels through the crystal relative to that of the optic axis of the crystal. Light traveling parallel to the optic axis is only singly refracted.

Birefringent crystals are either uniaxial or biaxial, depending upon whether they have one optic axis or two. They are said to be positive or negative, depending upon whether the velocity of propagation (within the crystal) of the extraordinary wave is greater or less than that of the ordinary wave. Calcite is a uniaxial negative crystal, quartz a uniaxial positive crystal. The most commonly used biaxial crystal is mica. Doubly refracting crystals are frequently employed as polarizers, such as the nicol prism. *See* CRYSTAL OPTICS; POLARIZED LIGHT.

Refractometry. The measurement of indices of refraction, called refractometry, can be made in several ways. A very accurate technique is to determine in a prism spectrometer the minimum deviation D for a prism made from the material in question. The value of n is then calculated from Eq. (6). Hollow prisms can be used in this manner to determine the values of indices of refraction of various liquids. Alternatively, the critical angle for total internal reflection may be measured. Another method is to observe visually the apparent thickness of a slab of material by looking straight through it, and to compare the apparent thickness with the real thickness as measured with a micrometer. Then Eq. (8) holds.

$$n = \frac{\text{real thickness}}{\text{apparent thickness}} \tag{8}$$

Interferometric methods are particularly convenient for gases. In the Jamin refractometer, for example, a simple count of fringes as the gas is slowly admitted to an initially evacuated tube in the optical path yields n. These techniques can also be used for solids, particularly when the material is available in the form of thin films. *See* INTERFEROMETRY.

Refractometry is an important tool in analytical chemistry. For example, information about the composition of an unknown solution can frequently be obtained by measurement of its index of refraction. *See* REFRACTOMETRIC ANALYSIS.

Atmospheric refraction. Gases have indices of refraction only slightly greater than unity. In general, $n - 1$ is proportional to the density of the gas, or to the ratio of pressure to absolute temperature. The index of refraction of the Earth's atmosphere increases continuously from 1.000000 at the edge of space to 1.000293 (yellow light) at 32°F (0°C) and 760 mmHg (101.325 kilopascals) pressure. Thus celestial bodies as seen in the sky are actually nearer to the horizon than they appear to be. The effect decreases from a maximum of about 35 minutes of arc for an object on the horizon to zero at the zenith, where the light enters the atmosphere at perpendicular incidence. Thus the Sun (and all other bodies) appears to rise 2 or more minutes earlier (depending upon latitude) and to set 2 or more minutes later than would be the case without refraction. This must be taken into account when the altitude of a celestial body is observed for navigational purposes.

Other manifestations of atmospheric refraction are the mirages and "looming" of distant objects which occur over oceans or deserts, where the vertical density gradient of the air is quite uniform over a large area. The twinkling of stars is caused by the rapid small fluctuations in density along the light path in the atmosphere. Rainbows are produced by the multiple reflections, refraction, and dispersion of sunlight by spherical raindrops. *See* METEOROLOGICAL OPTICS; MIRAGE; RAINBOW; TWINKLING STARS.

Other electromagnetic waves. Although refraction is most frequently encountered for the visible portion of the spectrum, it is of importance for other electromagnetic radiation. For very long-wavelength radiation, the index of refraction of many materials is equal to the square root of the relative permittivity, ϵ_r. In general, $dn/d\lambda$ is negative except in the regions of so-called anomalous dispersion near absorption bands. On the short-wavelength side of an absorption band, n can be less than 1.00. Since it is the phase velocity of the wave rather than the group velocity which is involved in the definition of the index of refraction, this does not represent a violation of the principle of relativity, that is, that energy cannot be propagated at a velocity faster than the velocity of light in vacuum. At very high frequencies, the index of refraction of all materials is also slightly less than unity. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; GROUP VELOCITY; PERMITTIVITY; PHASE VELOCITY; RELATIVITY.

Refraction plays a role in the propagation beyond the line of sight of radio waves in the Earth's atmosphere. *See* RADIO-WAVE PROPAGATION.

The interaction of electromagnetic radiation with more or less opaque substances is often described in terms of a complex index of refraction. The real part of this quantity has the usual meaning for the small amount of light which penetrates into the material before it is absorbed. The imaginary part is a measure of the absorption.

Sound waves. The velocity of sound in a gas is proportional to the square root of the absolute temperature. Because of the vertical temperature gradients in the atmosphere, refraction of sound can be quite

pronounced. As in mirage formation, to allow large-scale refraction the temperature at a given height must be uniform over a rather large horizontal area. If the temperature decreases with altitude (the usual situation), sound waves initially traveling at a small angle with the horizontal are refracted upward. A sound out of doors is thus not normally audible at a great distance. However, if there is a temperature inversion (as over a body of water on a calm sunny day), the waves would be refracted downward. This is the main reason that sound carries long distances across water on a calm day. On a windy day the horizontal temperature strata are broken up and the sound is dissipated.

Refraction accompanied by reflection accounts for the fact that large explosions are sometimes heard in several distinct regions at surprisingly large distances, with zones of silence in between. A temperature inversion at high levels can refract the waves downward into a zone of audibility. The sound is then reflected from the ground, and must again be refracted downward to give the next zone of audibility. *See* SOUND; WAVE MOTION IN FLUIDS.

Seismic waves. The velocity of elastic waves in a solid depends upon the modulus of elasticity and upon the density of the material. Waves propagating through solid earth are refracted by changes of material or changes of density. Worldwide observations of earthquake waves enable scientists to draw conclusions on the distribution of density within the Earth. These waves may be totally internally reflected at the boundary of the core. It was through such observations that the existence of the much denser core of the Earth was first postulated.

Refraction of compressional waves from explosions set off on the ground is (combined with reflection) used in prospecting for oil, natural gas, and minerals which have large differences in density and elastic constants from the surrounding rocks. *See* SEISMOLOGY.

Water waves. The speed of water waves in shallow water is proportional to the square root of the depth. As the waves enter shallower water they travel more slowly. As a train of waves approaches a coastline obliquely, its direction of travel becomes more nearly perpendicular to the shore because of refraction. *See* GEOMETRICAL OPTICS; NEARSHORE PROCESSES; WAVE MOTION IN LIQUIDS. John W. Stewart

Bibliography. M. Born and E. Wolf, *Principles of Optics*, 7th ed., 1999; B. D. Guenther, *Modern Optics*, 1990; D. Halliday and R. Resnick, *Physics*, 4th ed., 1992; F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976; J. Meyer-Arendt, *Introduction to Classical and Modern Optics*, 4th ed., 1995.

Refractometric analysis

A method of chemical analysis based on the measurement of the index of refraction of a substance. As shown in the **illustration**, when light impinges on the surface of a material at an angle i to the normal to the surface, its direction is changed on passing into the material so that it then travels at an angle r to the normal. The index of refraction is defined in the equation below. It varies as a function

$$\text{Index of refraction} = \frac{\sin i}{\sin r} = n_D^{20}$$

of temperature and wavelength of light, and also of pressure in gases. Refractive indices are usually measured at 20°C (68°F), using the yellow D line of the sodium spectrum. The indices of refraction of a few substances are water, 1.333; benzene, 1.5014; chloroform, 1.4464; and acetone, 1.3589.

The most common type of refractometer is the Abbe refractometer. It is simple to use, requiring but a drop or two of sample and allowing a measurement of refractive index to be made in 1–2 min, with a precision of 0.0001. More precise measurements of refractive indices, within 0.00003, may be made by using a dipping or immersion refractometer, the prism of which is completely immersed in the sample. This requires about 15 ml of sample and is widely used to detect trace impurities or to control the quality of a product. The most precise measurements of the refractive indices of gases or solutions containing small traces of impurities are made with an interferometer, based on the interference of light. Its precision is 0.000001. For measurements in flowing systems, differential refractometers, capable of detecting a difference of 10^{-7} refractive index unit, are used.

The measurement of refractive index is used to identify compounds whose other physical constants are quite similar. Because minute amounts of impurities often cause a measurable change in the refractive index of a pure material, refractive index is often used as a criterion for purity. A measurement

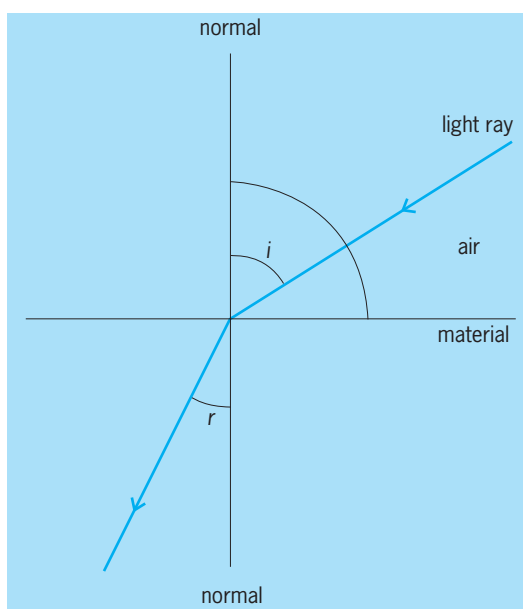

Refraction of light. The light ray changes direction as it passes into the material from the air.

of refractive index gives information as to the gross amount of impurity; it does not serve to identify the impurity. To give qualitative information, measurements would have to be made at different wavelengths, a rare procedure. In systems containing only two components, such as water and alcohol mixtures or aqueous salt solutions, refractive index is a sensitive and rapid method of determining the composition. Refractometric analysis has also been used for the determination of the water content in milk, the amount of sulfur in rubber, and the isotopic analysis of D_2O–H_2O mixtures. *See* REFRACTION OF WAVES.

Robert F. Goddu; James N. Little

Bibliography. J. D. Winefordner, *Treatise on Analytical Chemistry*, 2d ed., 1989.

Refractory

One of a number of ceramic materials for use in high-temperature structures or equipment. The term high temperatures is somewhat indefinite but usually means above about 1800°F (1000°C), or temperatures at which, because of melting or oxidation, the common metals cannot be used. In some special high-temperature applications, the so-called refractory metals such as tungsten, molybdenum, and tantalum are used.

The greatest use of refractories is in the steel industry, where they are used for construction of linings of equipment such as blast furnaces, hot stoves, and open-hearth furnaces. Other important uses of refractories are for cement kilns, glass tanks, nonferrous metallurgical furnaces, ceramic kilns, steam boilers, and paper plants. Special types of refractories are used in rockets, jets, and nuclear power plants. Many refractory materials, such as aluminum oxide and silicon carbide, are also very hard and are used as abrasives; some applications, for example, aircraft brake linings, use both characteristics. *See* ABRASIVE.

Refractory materials are commonly grouped into (1) those containing mainly aluminosilicates; (2) those made predominately of silica; (3) those made of magnesite, dolomite, or chrome ore, termed basic refractories (because of their chemical behavior); and (4) a miscellaneous category usually referred to as special refractories. *See* METAL COATINGS.

Aluminosilicate refractories. Fireclay is the raw material from which the bulk (about 70%) of refractories is manufactured. Different grades are distinguished according to the softening temperature or the pyrometric cone equivalent (PCE), the number of the standard pyrometric cone which deforms under heat treatment in the same manner as the fireclay. Thus, the minimum PCEs for low, intermediate, high, and superduty fireclays are 19, 29, 31/32, and 33, respectively. Fireclays are also classified by their working properties into two other classes: plastic (those which form a moldable mass when mixed with water), and flint (a hard, rocklike clay that does not become plastic when mixed with water). In general, flint clays have higher PCEs than plastic clays and are mixed with them to form higher-grade fireclay brick. *See* CLAY, COMMERCIAL.

High alumina refractories are made from clays which contain, in addition to the alumina (Al_2O_3) in the clay minerals, hydrates of aluminum oxide. These raise the total Al_2O_3 content and make the material more refractory. Different grades are distinguished on the basis of the total Al_2O_3 content (50, 60, 70% alumina refractories).

Sillimanite and kyanite are anhydrous aluminosilicate minerals used to make special refractory objects, such as crucibles, tubes, and muffles, or as an addition to fireclay to increase refractoriness and to control its shrinkage during firing.

Silica refractories. These account for about 15% of total production. They are made from crushed and ground quartzite (ganister) to which about 2% lime has been added to assist in bonding, both before and after firing. The quality of silica refractories is to a great extent determined by the amount of Al_2O_3 impurity, even small amounts having a deleterious effect on refractoriness. This is just opposite to the case of alumina in fireclays, where a higher alumina content means greater refractoriness. High-grade silica brick contains less than 0.6% Al_2O_3, and even the standard grade contains less than 1%. During firing, the mineral quartz transforms to cristobalite and tridymite, the high-temperature forms of silica. Since these are less dense than quartz, the brick expands on firing and the true density of the solid is often taken as a test of adequate firing. An example would be the case in which the density of acceptably fired material must be below 1.36 oz/in^3 (2.35 g/cm^3). The outstanding characteristic of silica is its ability to withstand high loads at elevated temperatures, for example, as a sprung-arch roof 30 to 40 ft (9–12 m) wide over an open hearth. The hearth may be operated within 120°F (50°C) of the melting point of silica. *See* SILICON.

Semisilica refractories are made from clay with a high silica (sand) content (over 70% total silica); their main advantage is their dimensional stability when heated, or fired. Apparently the expansion of the silica, as sand, offsets the contraction of the clay.

Basic refractories. Magnesite refractories are so named because magnesium carbonate mineral was for many years the sole raw material. Since World War II seawater has become a significant source of magnesium oxide refractory, and such material is often called seawater magnesite. In any case, the raw material is calcined to form a material largely magnesium oxide, MgO; about 5% iron oxide is usually added before calcining. *See* MAGNESITE.

Chrome refractories are made from chrome ore, a complex mineral containing oxides of chromium, iron, magnesium, aluminum, and other oxides crystallized in the spinel structure. These crystals are usually embedded in a less refractory matrix called gangue. *See* CHROMITE.

In an attempt to combine the best properties of each, magnesite and chrome are often mixed to form chrome-magnesite or magnesite-chrome refractories (the first-named is the dominant constituent).

Dolomite is a mixed calcium-magnesium carbonate, $CaMg(CO_3)_2$, which, when calcined to a

mixture of MgO and CaO, is used in granular form to patch the bottoms of open hearths and also to make bricks.

Miscellaneous materials. Special refractories are made of a great many materials, and it is possible to mention here only a few of the more important.

Silicon carbide, SiC, is used for many refractory shapes, its outstanding properties being good thermal and electrical conductivity (it is used to make electric heating elements for furnaces), good heat-shock resistance, strength at high temperatures, and abrasion resistance. The first silicon carbide refractories were bonded with clay, so that the refractory properties of the bond placed the ultimate limit on the material. A method of making self-bonded silicon carbide has been developed to remove this limitation. Although silicon carbide tends to oxidize to form SiO_2 and either CO or CO_2, the silica-oxidation product forms a glassy coating on the remaining material and to a certain extent protects it from further oxidation.

Insulating firebrick is made from refractory clays to which a combustible material (sawdust, cork, coal) has been added; when this burns during the firing operation, it leaves a brick of high porosity. The low thermal conductivity of insulating brick reduces heat losses from furnaces, and the low bulk density and consequent low heat capacity reduce the amount of heat needed to bring the furnace itself up to temperature. The main disadvantage of such bricks is their low strength, but even this is useful in that they can be cut or ground to shape quite readily.

Pure oxides, of which alumina, Al_2O_3, is the prime example, are used for many special refractories. Zircon, $ZrSiO_2$, and zirconia, ZrO_2, are finding increased and significant uses as refractory materials. Some, such as beryllia, BeO, thoria, ThO_2, and uranium oxide, UO_2, are of particular interest for nuclear applications.

Carbides, nitrides, borides, silicides, and sulfides of various sorts have been considered as refractory materials, and some study made of them; aside from a few carbides and nitrides, however, none have found much use.

Cermets are an intimate mixture of a metal and a nonmetal, for example, Al_2O_3 and chromium. Although the nonmetal may be an oxide, it is more commonly a carbide or nitride (as in cemented tungsten carbide). *See* CERMET.

Carbon, generally in the form of graphite, is used for such equipment as crucibles and as stopper nozzles in ladles for steel casting. A potentially very large use of carbon is in blocks for construction of blast-furnace hearths. Graphite has very good thermal-shock resistance and moderate electrical conductivity, does not melt but rather sublimes at a significant rate only at temperatures well above 5400°F (3000°C), is quite inert chemically, and is wet by very few molten materials. The main disadvantage of graphite, common to all nonoxide materials at high temperatures, is that it oxidizes; since the products are all gaseous, they offer no protection against further oxidation. *See* GRAPHITE.

Manufacture. Standard ceramic techniques are used. Hand molding, once widely used, is used only for special shapes and small orders. The extrusion or stiff mud process is used for plastic fireclays; very often the extruded blanks are repressed or hydraulically rammed to form special shapes, for example, T-sections of refractory pipe. Power pressing of simple shapes is the most widely used forming method. Hot pressing and hydrostatic pressing are used for some special refractories. Slip casting is used for special refractory shapes. Fusion casting is commonly used for glass tank blocks; these are mainly either Al_2O_3 or Al_2O_3, with significant amounts of SiO_2, ZrO_2, or both. *See* CERAMICS.

Refractories are generally fired in tunnel kilns, but some periodic kilns are still used, particularly for special shapes. *See* KILN.

Some types of basic refractories, known as chemically bonded, are pressed with a chemical binder, such as magnesium oxychloride, and installed without firing. Some of these, the steel-clad refractories, are encased in a metal sheath at the time of pressing. When the refractory is heated after installation, the iron oxidizes and reacts with the refractory, forming a tight bond between the individual bricks.

In all refractory products and in unfired brick in particular, the maximum possible formed density is desired. To this end, careful crushing and sizing of raw materials are carried out so that, as far as possible, the gaps between large pieces are filled with smaller particles, and the space between these with still smaller, and so on. In the case of clay refractories, it is customary to use prefired (calcined) clay or crushed, fired rejects (both are known as grog) to increase the density and to reduce the firing shrinkage.

Properties. A high melting point is of course necessary in a refractory, but many other properties must be considered in choosing a refractory for a specific application.

A definite melting point is characteristic of pure materials; actual minerals from which refractories are made, for example, clay, are far from pure and hence do not melt at a specific temperature. Rather, they form increasing amounts of liquid as the temperature is increased above a certain minimum temperature at which liquid first appears. This characteristic of gradual softening is indicated by the PCE of the material and the underload test.

High-temperature strength is important for refractories, but most materials become plastic and flow at elevated temperatures. Therefore, the rate of flow (creep rate) at a given temperature under a given load is a more important design criterion.

A knowledge of the thermal expansion of high-temperature materials is important, first, so that allowance can be made in furnace construction [long tunnel kilns must be built with expansion joints of several inches every 10 ft (3 m) or so], and second, because of its relation to thermal-shock resistance.

Thermal conductivity determines the amount of heat that will flow through a furnace wall under

given conditions, and a knowledge of this property is essential to furnace design.

Thermal-shock resistance is the ability of a specimen to withstand, without cracking, a difference in temperature between one part and another. For example, if a red-hot brick is dropped into cold water, it is likely to shatter since the outside cools and contracts while the center is still hot. This cracking is often referred to as thermal spalling, the term spalling meaning any cracking off of large pieces of brick. Other causes of spalling are mechanical (hitting the brick and knocking off a piece) and structural (a reaction in the brick which changes the mineral structure and causes cracking). Thermal-shock resistance is enhanced by high strength, low Young's modulus, low thermal expansion, and sometimes, depending on conditions, high thermal conductivity. Whether or not a given specimen cracks under heat shock depends not only on the material of which it is made, but also on its size and shape and on the test conditions, for example, whether it is dropped into water or into still air at the same temperature.

Various chemical properties are important in refractories. For example, the tendency of the magnesium oxide in basic brick to hydrate, that is, to react with water to form $Mg(OH)_2$, should be as low as possible. Turning to high-temperature chemistry, the rate of corrosion of refractories by molten slags and iron oxide fumes is vital to the length of service. Reference to the appropriate phase equilibrium diagrams may give some indication of which combinations of slag and refractory will react; but usually, actual tests are needed to make any precise predictions. The rate of corrosion depends to a great extent on such physical factors as the porosity of the refractory and whether or not the refractory is wet by the slag.

Carbon deposition is another chemical reaction which affects the life of refractories. The reaction is not with the refractory, but is catalyzed by substances in it. When carbon monoxide, perhaps in the top of a blast furnace, comes in contact with certain iron compounds which can occur in fireclays, its reduction to carbon is catalyzed. This carbon deposits at the site of the catalyst in the brick, and causes the brick to shatter. The effect is most pronounced around 950°F (500°C); much below this temperature, the rate of reaction is too slow, and much above it, the equilibrium oxygen pressure necessary for the reduction is lower than is found in practice. Although the reaction is not completely understood, it has been found that high-temperature firing of the fireclay refractories converts the iron to a form which does not catalyze the carbon deposition.

The bursting of spinel (chrome) refractories in contact with iron oxide is another high-temperature chemical reaction; it is not thoroughly understood, but appears to be related to oxidation and reduction reactions in the refractory. John F. McMahon

Bibliography. J. I. Duffy (ed.), *Refractory Materials Since 1977*, 1981; *Iron and Steel Society Refractories for Modern Steelmaking Systems*, 1987; C. Krause, *Refractories: The Hidden Industry*, 1987; P. Kumar and R. L. Ammon (eds.), *Refractory Metals: State of the Art*, 1989.

Refrigeration

The cooling of a space or substance below the environmental temperature. The art was known to the ancient Egyptians and people of India, who used evaporation to cool liquids in porous earthen jars exposed to dry night air; and to the early Chinese, Greeks, and Romans, who used natural ice or snow stored in underground pits for cooling wine and other delicacies. In the eighteenth and early nineteenth centuries natural ice cut from lakes and ponds in winter was stored underground for use in summer. The technique of mechanical refrigeration began with the invention of machines for making artificial ice. Great strides have been made in the twentieth century in the application of mechanical refrigeration to fields other than ice making, including the direct cooling and freezing of perishable foods and air conditioning for industry and human comfort.

Mechanical refrigeration is primarily an application of thermodynamics wherein the cooling medium, or refrigerant, goes through a cycle so that it can be recovered for reuse. The commonly used basic cycles, in order of importance, are vapor-compression, absorption, steam-jet or steam-ejector, and air. Each cycle operates between two pressure levels, and all except the air cycle use a two-phase working medium which alternates cyclically between the liquid and vapor phases.

The term "refrigeration" is used to signify cooling below the environmental temperature to lower than about 150 K (−190°F; −123°C). The term "cryogenics" is used to signify cooling to temperatures lower than 150 K. *See* CRYOGENICS.

Vapor-compression cycle. The vapor-compression cycle (**Fig. 1**) consists of an evaporator in which the liquid refrigerant boils at low temperature to produce cooling, a compressor to raise the pressure and temperature of the gaseous refrigerant, a condenser in which the refrigerant discharges its heat to the environment, usually a receiver for storing the liquid condensed in the condenser, and an expansion valve through which the liquid expands from the high-pressure level in the condenser to the low-pressure level in the evaporator. This cycle may also be used for heating if the useful energy is taken off at the condenser level instead of at the evaporator level. *See* HEAT PUMP.

The theoretical vapor-compression cycle can best be analyzed on the pressure-enthalpy or temperature-entropy coordinates for a two-phase fluid (**Fig. 2**). Enthalpy is a parameter that replaces heat content. The specific enthalpy, that is, the enthalpy per unit mass (Btu/lbm; kJ/kg), is equal to the specific internal energy (Btu/lbm; kJ/kg) plus the pressure (lbf · in.$^{-2}$; kPa)–volume (ft^3/lbm; m^3/kg) product divided by the mechanical equivalent of heat (778 ft · lbf/Btu; 1 J).

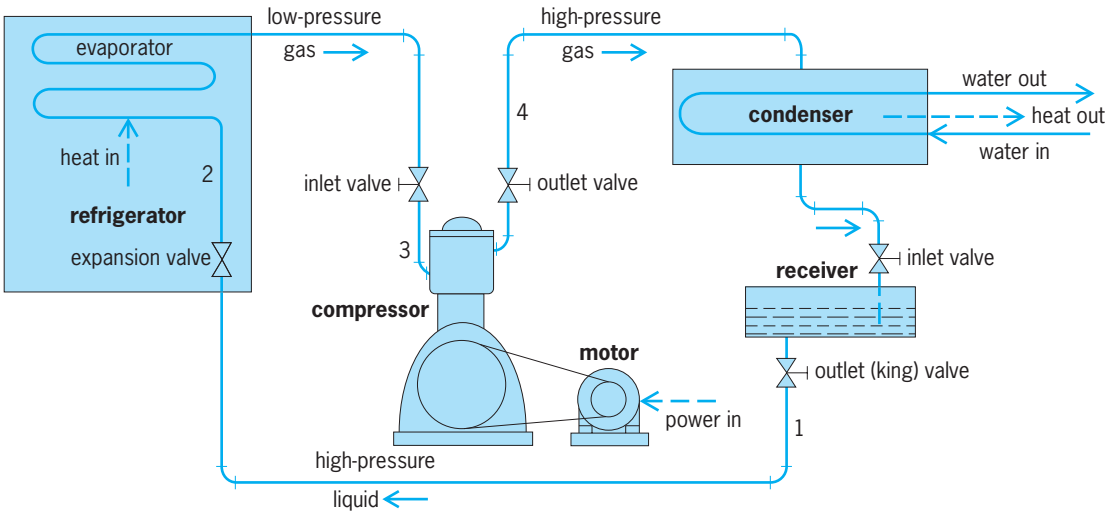

Fig. 1. Diagram of the vapor-compression cycle.

Entropy is a parameter that is obtained by dividing the heat flow by the average absolute temperature during the change. It is expressed in units of Btu per pound mass per degree Rankine (°R = °F + 460) or kilojoule per kilogram degree Kelvin [kJ/(kg · K)]. In Fig. 2, process 1–2 represents adiabatic (constant enthalpy) expansion; 2–3′, constant temperature (and pressure) evaporation; 3′–3, suction superheating at constant pressure; 3–4, ideal frictionless adiabatic (constant entropy) compression; 4–4′, removal of discharge superheat at constant pressure; 4′–1′, condensation at constant pressure (and temperature); and 1′–1, liquid subcooling at constant pressure.

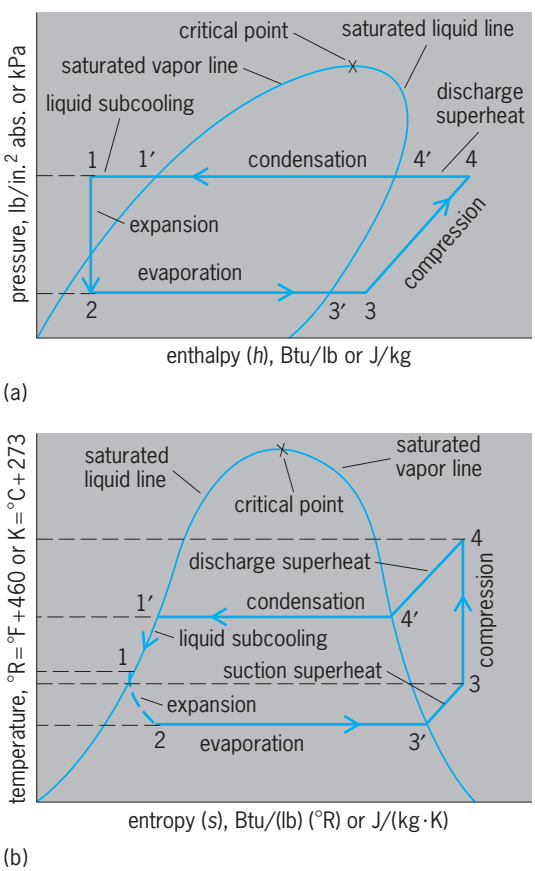

Fig. 2. Diagrams showing theoretical analysis of the vapor-compression cycle. (*a*) Pressure-enthalpy diagram. (*b*) Temperature-entropy diagram.

The efficiency of a heat-power cycle is defined as the ratio of useful output to energy input. For a heat engine, efficiency is less than unity. Efficiency is not very meaningful for the refrigeration and heat-pump cycles, for which the term "coefficient of performance" (CP) is used instead. Referring to the theoretical cycle (Fig. 2), the refrigeration CP is the ratio of cooling effect in evaporator 2–3 to compressor energy input 3–4, and the heat-pump CP is the ratio of heating effect in condenser 4–1 to compressor energy input 3–4. The coefficient of performance may be considerably greater than unity, and the theoretical heat-pump CP is 1 plus refrigeration CP. The cooling effect in evaporator 2–3, q_L, is $h_3 - h_2$, and the compressor energy input 3–4, w_C, is $h_4 - h_3$, so that the coefficient of performance of the system considered as a refrigerator, CP_R, is equal to $q_L/W_C = (h_3 - h_2)/(h_4 - h_3)$. The cycle CP is thus determinable from a knowledge of the four enthalpies h_1, h_2, h_3, and h_4.

Compound cycles using two or more compressors in series with a common refrigerant can be used to extend the temperature range with a single compressor stage. However, practical difficulties can occur, and a cascade cycle is used instead, consisting of two or more separate refrigerants each working in its own closed cycle. Such cycles reduce the power consumption and hence increase the thermodynamic efficiency.

Absorption cycle. The absorption cycle accomplishes compression by using a secondary fluid to absorb the refrigerant gas, which leaves the evaporator at low temperature and pressure. Heat is applied, by means such as steam or gas flame, to distill the refrigerant at high temperature and pressure. The most-used refrigerant in the basic cycle (**Fig. 3**) is

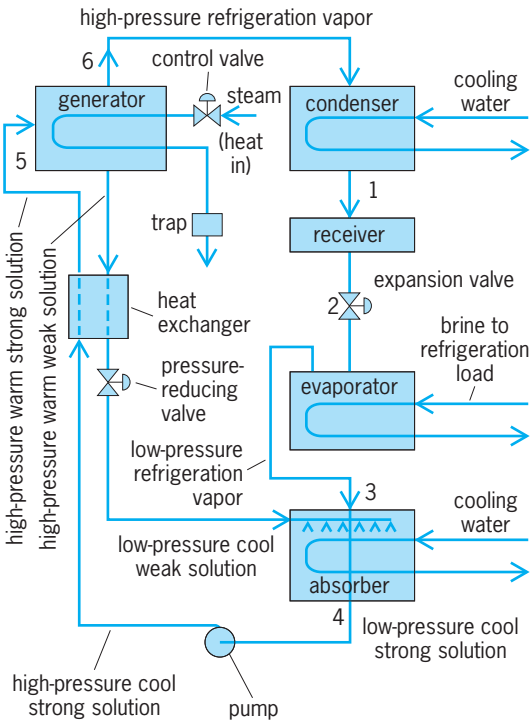

Fig. 3. Basic absorption cycle.

ammonia; the secondary fluid is then water. This system is used for the lower temperatures. Another system is lithium bromide–water, where the water is used as the refrigerant. This is used for higher temperatures. Due to corrosion, special inhibitors must be used in the lithium bromide–water system. The condenser, receiver, expansion valve, and evaporator are essentially the same as in any vapor-compression cycle. The compressor is replaced by an absorber, generator, pump, heat exchanger, and controlling-pressure reducing valve.

The operation of the cycle is based on the principle that the vapor pressure of a refrigerant is lowered by the addition of an absorbent having a lower vapor pressure; and the greater the quantity of absorbent used, the more the depression of the vapor pressure of the refrigerant. *See* DALTON'S LAW.

By maintaining the solution in the absorber at the proper temperature and concentration, the vapor pressure of the solution can be kept lower than that of the refrigerant in the evaporator. Spraying the weak solution in the absorber then causes the refrigerant vapor to flow from the evaporator to the absorber. The strong solution thus formed in the absorber is then pumped through a heat exchanger to the generator, where heat is applied to release the refrigerant vapor. Then follows condensation, expansion, and evaporation, as in the standard vapor-compression cycle. Except for small units, an indirect system is used wherein brine is cooled and circulated to the actual refrigeration load.

For air conditioning, water is the refrigerant and lithium bromide is the absorbent. In terms of the basic absorption cycle (Fig. 3), from 1 to 2 the high-pressure liquid refrigerant is expanded into the evaporator where brine is usually cooled, from 2 to 3 the low-pressure refrigerant vapor is drawn into the absorber, from 3 to 4 the low-pressure refrigerant vapor is absorbed in the weak solution, from 4 to 5 the low-pressure strong solution is pumped through the heat exchanger to the high-pressure generator, and from 5 to 6 heat is applied to drive off the refrigerant vapor and force it into the condenser. The hot weak solution drains back to the absorber through the heat exchanger and pressure-reducing valve.

Steam-jet cycle. The steam-jet cycle uses water as the refrigerant. High-velocity steam jets provide a high vacuum in the evaporator, causing the water to boil at low temperature and at the same time compressing the flashed vapor up to the condenser pressure level. Its use is limited to air conditioning and other applications for temperatures above 32°F (0°C).

The basic steam-jet or ejector cycle (**Fig. 4**) is usually analyzed on temperature-entropy coordinates (**Fig. 5**). High-pressure motive steam (Fig. 4) at 1 is expanded to a low absolute pressure at 2 through a converging-diverging nozzle. Path 1–2 (Fig. 5) is the ideal expansion and 1–2′ the actual expansion allowing for nozzle friction. Water vapor in the evaporator at 3 is entrained by the motive steam at 4, the steam having lost some of its energy from 2′ to 4 because

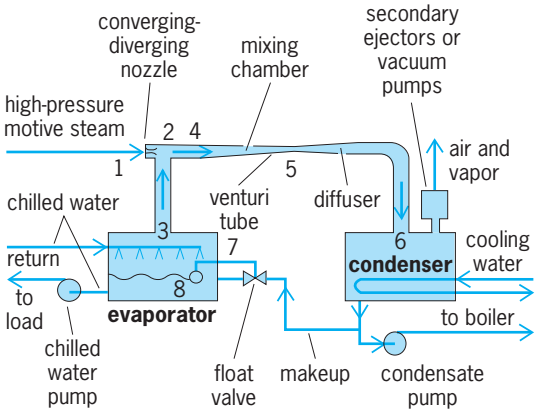

Fig. 4. Steam-jet water-vapor cycle.

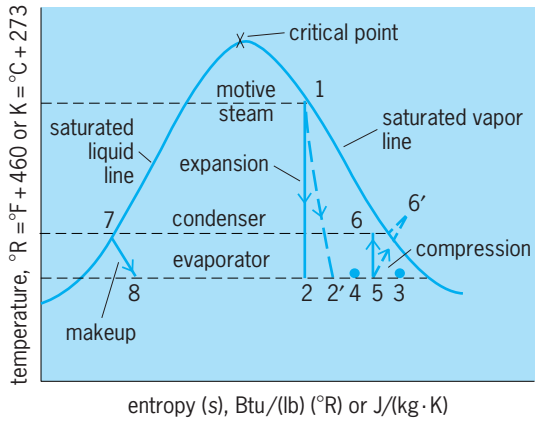

Fig. 5. Temperature-entropy coordinates for analyzing basic steam-jet or ejector cycle.

the entrainment efficiency is rather low. The motive steam at 4 plus the entrained moisture at 3 are forced through the venturi tube, in which the velocity of the incoming mixture is reduced and converted into pressure head in the condenser at 6. Path 5–6 is the ideal compression and 5–6′ the actual compression, taking into account compression efficiency. Typical efficiencies are nozzle 88%, entrainment 65%, and compression 80%.

In the evaporator or flash chamber, part of the water is evaporated to cool the rest of the water, which is circulated to the cooling load. Makeup water must be added from 7 to 8. Typical operating conditions are 100°F or 38°C (2 in. Hg or 6.8 kilopascals absolute pressure) in the condenser and 40°F or 4°C (0.25 in. Hg or 850 pascals absolute pressure) chilled water in the evaporator. The condensate from the condenser is pumped back to the boiler; secondary ejectors, or vacuum pumps, are required to remove the air and maintain the high vacuum. Considerably more water is required for condensing than for a vapor-compression system of the same capacity.

The steam jet refrigerator is most widely used for processes where direct vaporization is used for drying of heat-sensitive materials, or in processes using simultaneous desorption and chilling where hard water or seawater can be used. Examples include freeze drying of foods, the concentration of fruit juices, and chilling vegetables.

Air cycle. The air cycle, used primarily in airplane air conditioning, differs from the other cycles in that the working fluid, air, remains as a gas throughout the cycle. Air coolers replace the condenser, and the useful cooling effect is obtained by a refrigerator instead of by an evaporator. A compressor is used, but the expansion valve is replaced by an expansion engine or turbine which recovers the work of expansion. Systems may be open or closed. In the closed system, the refrigerant air is completely contained within the piping and components, and is continuously reused. In the open system, the refrigerator is replaced by the space to be cooled, the refrigerant air being expanded directly into the space rather than through a cooling coil.

One of the typical open air-cycle systems used on airplanes is called the "bootstrap" system (**Fig. 6**). It may be analyzed theoretically on the temperature-entropy coordinates (**Fig. 7**). From 1 to 2 ambient air is compressed ideally in the engine or supercharger of the airplane. Part of this high-pressure air is bled through a primary heat exchanger, where it is cooled from 2 to 3 by ram air, that is, ambient air compressed by the forward motion of the airplane. From 3 to 4 this air is further increased in pressure by the compressor of the refrigeration machine; from 4 to 5 the air is cooled by ram air in the cooler or secondary heat exchanger; and from 5 to 6 the air is further cooled in the expansion turbine, ideally without moisture. However, there is entrained moisture, about 70% of which is removed by the water separator, resulting in an approximate path 5–6′ as the air is warmed by the heat given up by the condensation and removal of this moisture. The balance of the

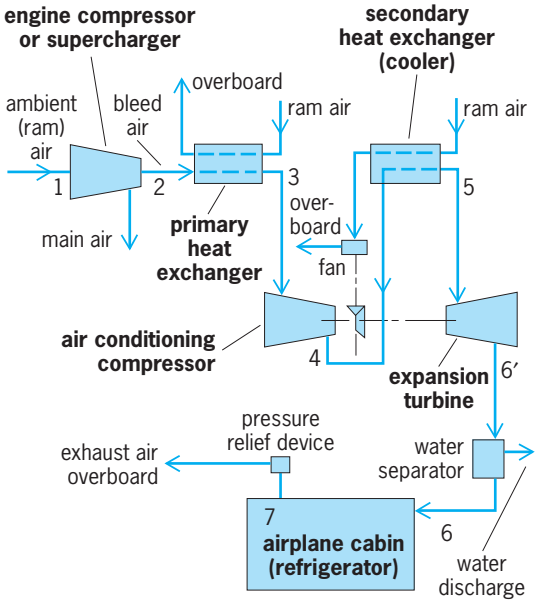

Fig. 6. Open air-cycle bootstrap system for airplanes.

moisture evaporates in the cabin and contributes to the cooling effect 6′–7, where 6′ is the dry-air rated temperature of the air as it enters the cabin. Dry-air rated temperature is the temperature which would be attained by the expansion of the refrigerant air in the absence of any moisture condensation. Refrigerant air at 7, having absorbed the heat load in the cabin, is released overboard through a pressure-relief device. The equipment is proportioned so that the work of expansion is recovered and is sufficient to drive the compressor and the cooler fan to approximate constant entropy expansion and compression. However, compressor and turbine efficiencies and heat-exchanger pressure drops, neglected in this analysis, reduce the ideal performance.

Thermoelectric refrigeration. The effect of heat application to one junction of a couple composed of two dissimilar conductors is to produce a thermoelectric voltage on the circuit (Seebeck effect).

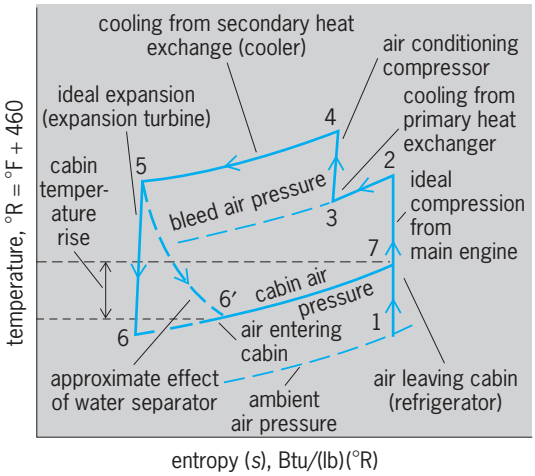

Fig. 7. Temperature-entropy coordinates for analyzing theoretically the air-cycle bootstrap system.

Fluids in REFPROP* Database

Name or R number	Chemical formula	Melting point, °C	Boiling point, °C	Critical temp., °C	Critical press., bar	Molecular weight
Ammonia	NH_3	−77.9	−33.3	133.0	114.2	17.03
Butane	C_4H_{10}	−159.5	−0.4	152.0	38.0	58.13
Carbon dioxide	CO_2	−56.7	−78.4	31.1	73.7	44.01
Ethane, R 170	C_2H_6	−174.2	−88.8	32.2	48.9	30.07
Isobutane, R 600a	C_4H_{10}	−136.2	−11.7	135.0	36.5	58.13
Propane, R 290	C_3H_8	−187.7	−42.1	96.8	42.5	44.10
Propylene, R 1270	C_3H_6	−185	−47.7	91.8	50.5	42.08
R 11	CCl_3F	−111	23.8	198.0	44.1	137.37
R 12	CCl_2F_2	−157.8	−29.8	112.0	41.1	120.91
R 13	$CClF_3$	−181.2	−81.4	28.8	38.7	104.46
R 14	CF_4	−183.7	−127.9	−45.7	37.5	88.00
R 22	$CHClF_2$	−160	−40.8	96.2	49.9	86.47
R 23	CHF_3	−155.2	−82.1	26.3	48.7	70.01
R 32	CH_2F_2	−136	−51.7	−78.2	58.0	52.02
R 41	CH_3F	−114.8	−78.4	44.3	58.8	34.03
R 113	$C_2Cl_3F_3$	−35.0	47.6	214.1	34.4	187.38
R 114	$C_2Cl_2F_4$	−93.9	3.8	145.7	32.5	170.92
R 115	C_2ClF_5	−102	−39.1	79.9	31.5	154.47
R 116	C_2F_6	−100.6	−78.2	19.9	30.4	138.01
R 123	$C_2HCl_3F_3$	−107	27.9	183.7	36.7	152.93
R 124	C_2HClF_4	−100	−12.0	122.5	36.4	136.48
R 125	C_2HF_5	−103	−48.1	66.3	36.3	120.02
R 134a	$C_2H_2F_4$	−101	−26.1	101.1	40.6	102.03
R 141b	$C_2H_3Cl_2F$	−103.5	32.2	204.4	42.5	116.95
R 142b	$C_2H_3ClF_2$	−130.8	−9.8	137.2	41.2	100.50
R 143a	$C_2H_3F_3$	−111	−47.4	73.6	38.3	84.04
R 152a	$C_2H_4F_2$	−117	−24.0	113.3	45.2	66.05
R 227ea	C_3HF_7		−18.3	103.5	29.5	170.03
R 236ea	$C_3H_2F_6$		6.6	141.2	35.3	152.04
R 236fa	$C_3H_2F_6$		−1.1	130.7	31.8	152.04
R 245ca	$C_3H_3F_5$		25.5	178.5	38.6	134.05
R 245fa	$C_3H_3F_5$		15.3	157.6	36.4	134.05
R C318	C_4F_8		−5.8	115.4	27.8	200.04

*Fluids listed in M. O. McLinden et al., NIST Standard Reference Database 23, 1998. Many of these fluids are being evaluated for use as binary mixtures—see M. O. McLinden et al., *Int. J. Refrig.*, 21(4):322–338, 1998. For fluoroethers see D. B. Bivens and B. H. Minor, *Int. J. Refrig.*, 21(7):567–576, 1998. Tables and charts for many of these fluids are contained in the *Handbook of Fundamentals*, ASHRAE, Atlanta, 1997, and in *Perry's Chemical Engineers Handbook*, McGraw-Hill, New York, 1997.

Inversely, this is matched by a cooling effect at one junction of an equivalent couple when a direct-current electrical flow is imposed on the circuit. This is known as the Peltier effect, discovered in 1834. The amount of the refrigerating effect is multiplied by the number of couples in parallel; the heat absorption at the cold junction is rejected at the second junction of each couple (that is, heat-pump effect).

With ordinary materials generally available prior to World War II, the Peltier effect, though recognized, was not considered particularly practical. Since the war and with the advent of the transistor and other semiconductor devices leading to more practical and efficient equipment, there has been considerable worldwide interest in the possibilities of this system. Its application is to special-purpose processes, since its low efficiency has to be compensated by its special features, such as elimination of moving parts and noiseless operation. *See* PELTIER EFFECT; SEEBECK EFFECT; THERMOELECTRICITY.

Refrigerants and equipment. The working fluid in a two-phase refrigeration cycle is called a refrigerant. A useful way to classify refrigerants is to divide them into primary and secondary. Primary refrigerants are those fluids (pure substances, azeotropic mixtures which behave physically as a single pure compound, and zeotropes which have temperature glides in the condenser and evaporator) used to directly achieve the cooling effect in cycles where they alternately absorb and reject heat. Secondary refrigerants are heat transfer or heat carrier fluids.

A number of commonly used refrigerantsm are listed in the **table**. Ammonia and Freon-22 are most important for industrial refrigeration, Freon-11 and Freon-12 being used most often for commercial and air conditioning work in which nontoxic refrigerants are necessary. A secondary cooling liquid that does not change from the liquid phase is called a brine. Solutions of sodium chloride or calcium chloride in water are frequently used as circulating brines in refrigeration systems.

Compressors. Refrigeration compressors may be positive-displacement of the reciprocating, rotary, or gear type for high- and medium-pressure differentials, or of the centrifugal type for low-pressure differentials. Early ammonia compressors were horizontal, double-acting, slow-speed units built like steam engines. Modern reciprocating compressors are vertical, single-acting, multicylinder, high-speed units built like automobile engines. Ammonia and other large compressors require water jacketing, whereas most Freon compressor cylinders are air-cooled.

Condensers. Refrigeration condensers may be air-cooled for small and medium capacities; or

water-cooled, as in the case of the shell-and-tube, shell-and-coil, or double-pipe types. Because of the large quantities of condensing water required, cooling towers or spray ponds are commonly used to recool the water for reuse. An evaporative condenser is a device combining a condensing coil and a forced-draft cooling tower in a single unit.

Evaporators. Refrigerant evaporators are the cooling units placed in the room or fluid to be cooled. Plain pipe coils or finned coils, with or without forced circulation of the fluid being cooled, are commonly used. Shell-and-tube coolers, or tanks with wetted or submerged cooling coils, are frequently used where water or brine is circulated as a secondary cooling medium.

Expansion valve. The main flow control in a vapor-compression system is the expansion valve. It permits the liquid refrigerant to expand from the high pressure in the condenser to the lower pressure in the evaporator. The expansion causes part of the liquid to evaporate and thereby to cool the remainder to the evaporator temperature. A float valve with a throttling orifice is often used instead of an expansion valve to provide flooded control and maintain a fixed liquid level in the evaporator. In domestic refrigerators, a capillary tube restricts flow from condenser to evaporator. For completely automatic operation, additional controls are required to maintain the desired temperature in the evaporator, to regulate the compressor operation and the flow of the condensing medium, and to provide safety protection. *See* AIR CONDITIONING; AIR COOLING; AUTOMOTIVE CLIMATE CONTROL; COLD STORAGE; COOLING TOWER; DRY ICE; ICE MANUFACTURE; MARINE REFRIGERATION; REFRIGERATOR. Peter E. Liley; Carl F. Kayan

Bibliography. Air Conditioning and Refrigeration Institute Staff, *Refrigeration and Air Conditioning*, 2d ed., 1987; R. J. Dossat, *Principles of Refrigeration*, 1996; B. C. Langley, *Fundamentals of Refrigeration*, 1995; B. Parsons (ed.), *1998 ASHRAE Handbook: Refrigeration* (SI or Inch-Pound editions), 1998.

Refrigeration cycle

A sequence of thermodynamic processes whereby heat is withdrawn from a cold body and expelled to a hot body. Theoretical thermodynamic cycles consist of nondissipative and frictionless processes. For this reason, a thermodynamic cycle can be operated in the forward direction to produce mechanical power from heat energy, or it can be operated in the reverse direction to produce heat energy from mechanical power. The reversed cycle is used primarily for the cooling effect that it produces during a portion of the cycle and so is called a refrigeration cycle. It may also be used for the heating effect, as in the comfort warming of space during the cold season of the year. *See* HEAT PUMP; THERMODYNAMIC PROCESSES.

In the refrigeration cycle a substance, called the refrigerant, is compressed, cooled, and then expanded. In expanding, the refrigerant absorbs heat from its surroundings to provide refrigeration. After the refrigerant absorbs heat from such a source, the cycle is repeated. Compression raises the temperature of the refrigerant above that of its natural surroundings so that it can give up its heat in a heat exchanger to a heat sink such as air or water. Expansion lowers the refrigerant temperature below the temperature that is to be produced inside the cold compartment or refrigerator. The sequence of processes performed by the refrigerant constitutes the refrigeration cycle. When the refrigerant is compressed mechanically, the refrigerative action is called mechanical refrigeration.

There are many methods by which cooling can be produced. The methods include the noncyclic melting of ice, or the evaporation of volatile liquids, as in local anesthetics; the Joule-Thomson effect, which is used to liquefy gases; the reverse Peltier effect, which produces heat flow from the cold to the hot junction of a bimetallic thermocouple when an external emf is imposed; and the paramagnetic effect, which is used to reach extremely low temperatures. However, large-scale refrigeration or cooling, in general, calls for mechanical refrigeration acting in a closed system. *See* PARAMAGNETISM; REFRIGERATION.

Reverse Carnot cycle. The purpose of a refrigerator is to extract as much heat from the cold body as possible with the expenditure of as little work as possible. The yardstick in measuring the performance of a refrigeration cycle is the coefficient of performance, defined as the ratio of the heat removed to the work expended. The coefficient of performance of the reverse Carnot cycle is the maximum obtainable for stated temperatures of source and sink. **Figure 1** depicts the reverse Carnot cycle on the *T-s* plane. *See* CARNOT CYCLE.

The appearance of the cycle in Fig. 1 is the same as that of the power cycle, but the order of the cyclic processes is reversed. Starting from state 1 of the figure, with the fluid at the temperature T_H of the hot body, the order of cyclic events is as follows:

1. Isentropic expansion, 1–2, of the refrigerant fluid to the temperature T_c of the cold body.

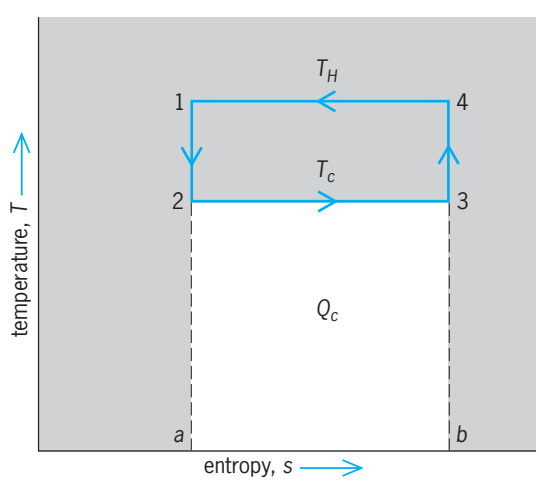

Fig. 1. Reverse Carnot cycle.

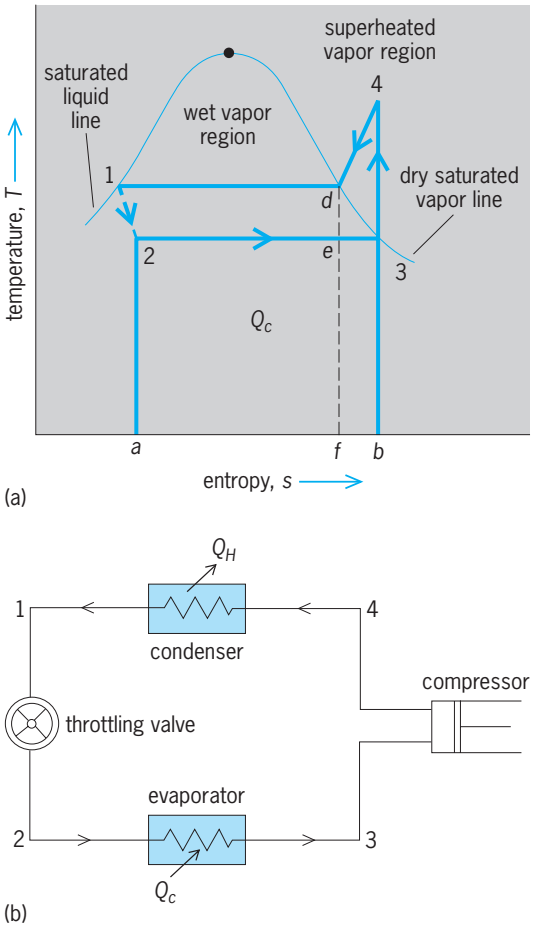

Fig. 2. Vapor-compression refrigeration cycle. (*a*) Temperature-entropy diagram. (*b*) Diagram of the components of a vapor-compression system.

2. Isothermal expansion, 2–3, at the temperature T_c of the cold body during which the cold body gives up heat to the refrigerant fluid in the amont Q_c, represented by the area 2–3–*b*–*a*.

3. Isentropic compression, 3–4, of the fluid to the temperature T_H of the hot body.

4. Isothermal compression, 4–1, at the temperature T_H of the hot body. During this process, the hot body receives heat from the refrigerant fluid in the amount Q_H represented by the area 1–4–*b*–*a*. The difference Q_H–Q_c represented by area 1–2–3–4 is the net work which must be supplied to the cycle by the external system.

Figure 1 indicates that Q_c and the net work rectangles each have areas in proportion to their vertical heights. Thus the coefficient of performance, defined as the ratio of Q_c to net work, is $T_c/(T_H - T_c)$.

The reverse Carnot cycle does not lend itself to practical adaptation because it requires both an expanding engine and a compressor. Nevertheless, its performance is a limiting ideal to which actual refrigeration equipment can be compared.

Modifications to reverse Carnot cycle. One change from the Carnot cycle which is always made in real vapor-compression plants is the substitution of an expansion valve for the expansion engine. Even if isentropic expansion were possible, the work delivered by the expansion engine would be very small and the irreversibilities present in any real operations would further reduce the work delivered by the expanding engine. The substitution of an expansion valve (throttling orifice), with constant enthalpy expansion, changes the theoretical performance but little, and greatly simplifies the apparatus. A typical vapor-compression refrigeration cycle is shown in **Fig. 2**.

Another practical change from the ideal Carnot cycle substitutes dry compression 3–4 for wet compression *e*-*d* in Fig. 2, placing state 4 in the superheat region above ambient temperature; the process is called dry compression in contrast to the wet compression of the Carnot cycle. Dry compression introduces a second irreversibility by exceeding the ambient temperature, thus reducing the coefficient of performance. Dry compression is usually preferred, however, because it simplifies the operation and control of a real machine. Vapor gives no readily observable signal as it approaches and passes point *e* in the course of its evaporation, but it would undergo a temperature rise if it accepted heat beyond point 3. This cycle, using dry compression, is the one which has won overwhelming acceptance for refrigeration work.

Reverse Brayton cycle. The reverse Brayton cycle constitutes another possible refrigeration cycle; it was one of the first cycles used for mechanical refrigeration. Before Freon and other condensable fluids were developed for the vapor-compression cycle, refrigerators operated on the Brayton cycle, using air as their working substance. **Figure 3**

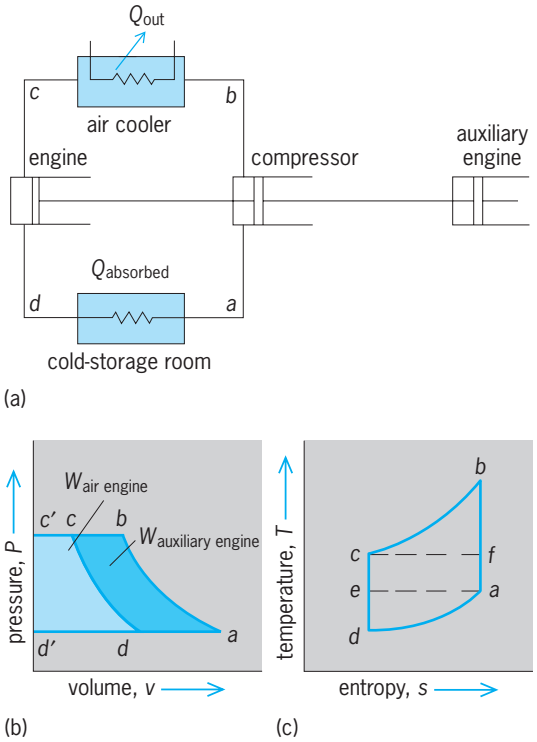

Fig. 3. Reverse Brayton cycle. (*a*) Schematic diagram of components. (*b*) Pressure-volume diagram. (*c*) Temperature-entropy diagram.

Ideal refrigeration cycle performance*

Refrigerant	Saturation pressure, lb/in.^{2} absolute		Refrigerating effect, Btu/lb	Refrigerant required per ton refrigeration		Compressor horsepower per ton refrigeration	Coefficient of performance
	Evaporator	Condenser		lb/min	cfm		
Carnot-any fluid							6.6
Freon, F12	47	151	49	4.1	3.6	0.89	5.31
Ammonia	66	247	445	0.44	1.9	0.85	5.55
Sulfur dioxide	24	99	135	1.5	4.8	0.83	5.67
Free air	15	75	32	6.2	7	2.7	1.75
Dense air	50	250	32	6.2	23	2.7	1.75
Steam	0.1	1.275	1000	0.2	590	0.91	5.18

*Performance is based on evaporator temperature = 35°F (2°C), condenser temperature = 110°F (44°C). In vapor-compression cycles, the pressures and temperatures are for saturation conditions; vapor enters the compressor dry and saturated; no subcooling in the condenser.
1 lb/in.^{2} absolute = 6.895 kPa.
1 Btu/lb = 2326 J/kg.
1 (lb/min)/ton refrigeration = 2.15×10^{-6} (kg/s)/W.
1 cfm/ton refrigeration = 1.34×10^{-7} (m^3/s)/W.
1 compressor horsepower per ton of refrigeration = 0.212 watt of compressor power per watt of refrigeration.

presents the schematic arrangement of this cycle. Air undergoes isentropic compression, followed by reversible constant-pressure cooling. The high-pressure air next expands reversibly in the engine and exhausts at low temperature. The cooled air passes through the cold storage chamber, picks up heat at constant pressure, and finally returns to the suction side of the compressor. *See* BRAYTON CYCLE.

The temperature-entropy diagram (Fig. 3*c*) points up the disadvantage of the dense-air cycle. If the temperature at *c* represents the ambient, then the only way that air can reject a significant quantity of heat along the line *b-c* is for *b* to be considerably higher than *c*. Correspondingly, if the cold body service temperature is *a*, the air must be at a much lower temperature in order to accept heat along path *d-a*. If a reverse Carnot cycle were used with a working substance undergoing changes in state, the fluid would traverse path *a-f-c-e* instead of path *a-b-c-d*. The reverse Carnot cycle would accept more heat along path *e-a* than the reverse Brayton cycle removes from the cold body along path *d-a*. Also, since the work area required by the reverse Carnot cycle is much smaller than the corresponding area for the reverse Brayton cycle, the vapor-compression cycle is preferred in refrigeration practice. *See* THERMODYNAMIC CYCLE.

Comparative performance of refrigerants. The **table** gives the significant theoretic performance data for a selected group of refrigerants when used in the ideal cycles as outlined above. These data include not only the requisite pressures in the evaporator and condenser for saturation temperatures of 35°F (2°C) and 110°F (44°C), respectively, but also the refrigerating effect, the weight and volume of refrigerant to be circulated, the horsepower required, and the coefficient of performance. The vapor compression cycles presuppose that the refrigerant leaves the evaporator dry and saturated (point 3 in Fig. 2) and leaves the condenser without subcooling (point 1 in Fig. 2). The heat absorbed, Q_c, in the evaporator is called the refrigerating effect and is measured in Btu/lb or joules/kg of refrigerant. The removal of heat in the evaporator at the rate of 200 Btu/min (3.517 kW) is defined as a ton refrigeration capacity. The weight and volume of refrigerant, measured at the compressor suction, and the theoretic horsepower required to drive the compressor per ton refrigeration capacity are also given in the table. These data show a wide diversity of numerical value. Each refrigerant has its advantages and disadvantages. The selection of the most acceptable refrigerant for a specific application is consequently a practical compromise among such divergent data.

Recent developements. The pulse tube refrigerator is a simple device containing a tube with one closed end and a movable piston at the other end. The closed end contains a heat exchanger, while the open end contains a heat exchanger through which a piston can move. With an orifice inserted in the tube, the device is called an orifice pulse tube refrigerator.

Theodore Baumeister; Peter E. Liley

Bibliography. American Society of Heating, Refrigerating and Air-Conditioning Engineers, *Heating, Ventilating, and Air Conditioning Guide*, revised periodically; E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996.

Refrigerator

An insulated, cooled compartment. If it is large enough for the entry of a person, it is termed a walk-in box; otherwise it is called a reach-in refrigerator. Cooling may be by mechanical or gas refrigeration, by water or dry ice, or by brine circulation. Temperatures maintained depend upon the requirements of the product stored, generally varying from 55°F (13°C) down to 0°F (−18°C), and sometimes lower.

A household or domestic refrigerator is a factory-built, self-contained cabinet. The range of storage capacities is wide and varies among manufacturers. Modern designs have a main compartment for holding food above freezing, a second compartment for storage below freezing, and trays for the freezing of

ice cubes. The cabinets are usually all metal with 2–3$^1/_2$ in. (5–9 cm) of insulation. The refrigeration unit is usually driven by electric motor, but gas refrigerators motivated by the thermal energy of fuel gas have been used extensively where cheap natural gas is available. Low-temperature household refrigerators, or home freezers, for the storage of frozen foods are manufactured in both the chest and the upright, or vertical, types.

A commercial refrigerator is any factory-built refrigerated fixture, cabinet, or room that can be assembled and disassembled readily, in contrast to a built-in refrigerator. Commercial or built-in refrigerators are used in restaurants, markets, hospitals, hotels, and schools for the storage of food and other perishables. A meat cooler is a refrigerator held at about 33°F (1°C) for the storage of fresh meats. Refrigerators for lower temperatures down to 0°F (−18°C) and below are called freezer boxes. Insulation thicknesses vary from 2 to 8 in. (5 to 20 cm), depending upon the service. In markets and stores, commercial display refrigerators may be of the self-service type from which customers help themselves. Both vertical types with glass doors, which the customer opens, and chest types with open tops are used. Electric refrigeration units may be built into each fixture, or they may be remotely located. *See* COLD STORAGE; REFRIGERATION.

Carl F. Kayan

Regeneration (engineering)

Any of the various processes for restoring a system to its original state by restoring some property of a system to its original value, or for using the properties of a system at one point in a cycle to modify the properties of the same system at another point in that cycle. The term has a wide variety of uses in different fields of engineering.

In mechanical engineering, regeneration is the process of transferring heat from the working fluid at one part of a cycle to the working fluid at a different part of the same cycle to increase cycle efficiency. An example is the use of extraction steam to preheat incoming boiler water. *See* BOILER.

In analog electronics, regeneration is the process of sampling the output variable, usually a voltage, and adding it to the input variable, again usually a voltage, so that an amplifier senses an increased signal level. Regeneration, in this sense, is called positive feedback. Low-level, controlled regeneration can result in increased amplification at the expense of narrowed bandwidth and decreased signal-to-noise ratio. Higher-level regeneration can result in oscillation, producing an output with no input. *See* FEEDBACK CIRCUIT; OSCILLATOR.

In digital electronics, regeneration is the process of restoring a set of data in a data storage device to a working level or restoring a degraded pulse-code-modulation (PCM) waveform to its original shape. In the data storage device, decay of the value of the variable representing a binary number, 0 or 1, can result in loss of data if the values are not refreshed or regenerated. In a pulse-code-modulation transmission system, a regenerative repeater can restore a degraded pulse waveform to its original shape, thus eliminating the effects of acquired noise or introduced distortion. *See* COMPUTER STORAGE TECHNOLOGY; PULSE MODULATION.

In higher-power electric motor applications, regeneration refers to the recapture of kinetic energy of the system by converting motor action to generator action when the motor is required to reduce speed. An example of this process is when the drive motors of an electrically propelled train are used as generators for dynamic braking, returning energy to the system during braking operations. *See* DYNAMIC BRAKING.

In nuclear engineering, regeneration refers to the process of increasing the fissile content of used nuclear fuel, a process that is achievable but seldom used. *See* NUCLEAR FUELS REPROCESSING.

William L. Beasley

Regenerative biology

The field within developmental biology that focuses on regeneration, the process by which an animal restores a lost part of its body. Broadly defined, the term regeneration can include many kinds of restorative activities shown by a wide variety of organisms. However, regeneration in this context connotes reproduction of a more perfect or complete replacement of a missing tissue or organ than usually results from processes of tissue repair or wound healing. Cells of the epidermis, blood, and digestive and reproductive tracts are regenerated continuously or periodically in all vertebrates. Within the field of developmental biology, however, most research in regeneration involves systems in which removing a complex structure or major part of an organism initiates a chain of events that produces a structure that duplicates the missing part both functionally and anatomically. *See* DEVELOPMENTAL BIOLOGY.

The best-known and most widely studied examples of regeneration are those involving epimorphosis, in which the lost structure is reproduced directly by a combination of cell proliferation and redifferentiation of new tissue. Examples can be found throughout the animal kingdom, from the extensive regenerative capacities of *Hydra* and *Planaria* to the regrowth of antlers in deer. Research on regenerating systems such as in annelid worms, hydra, and amphibian tails and limbs has yielded information regarding basic mechanisms of animal development. Regeneration can involve growth of an existing population of unspecialized cells or cells arising by dedifferentiation of tissues at the site of injury. Both processes are currently very active areas of investigation within the field of regenerative biology, holding promise for new approaches to various medical problems. *See* STEM CELLS; CELL DIFFERENTIATION.

Tissue and organ regeneration in mammals. Mammals, birds, and reptiles have a much more poorly

developed ability than amphibians and fish to regenerate complete organs, but nevertheless can reform missing tissue and restore function after partial removal of certain organs.

Liver. If part of the liver is cut away, the remaining portion increases in size to compensate for the missing tissue and to restore the normal functional capacity of the organ. The process of liver regeneration involves the triggering of active growth in the remaining liver cells, in cells of bile ductules, and in unspecialized stem cells, all of which are usually quiescent in the normal liver. Proliferation of these cells and their subsequent differentiation are key events in the process by which the missing liver mass is replaced and adequate hepatic function restored. *See* LIVER.

Musculoskeletal system. In the musculoskeletal system, different populations of quiescent stem cells allow efficient replacement of damaged or partially removed bones and muscles. Such undifferentiated progenitor cells are present within the inner and outer connective tissue coverings of bones, and are the major source of new proliferating cells from which new bone mass is formed during bone repair or restoration after surgery. Growth and patterning in these cells during bone regeneration are influenced by an important class of growth factors called bone morphogenetic proteins and by patterns of stress and mechanical stimulation placed on the existing bone. In skeletal muscle, undifferentiated cells called satellite cells are normally present along with the highly specialized muscle cells or fibers. When muscle fibers are killed by injury, the accompanying satellite cells are activated to proliferate and give rise to new muscle cells. The lack of satellite cells in heart muscle is one reason for the relatively poor potential for regeneration or repair exhibited by that tissue compared to other types of muscle. *See* BONE; MUSCLE.

While connective tissue cells of the bone marrow have long been recognized as a potential source of stem cells during bone repair, it has recently been shown that bone marrow–derived cells can also participate in regeneration of injured skeletal muscle and a variety of other tissues. Migrating from the marrow cavity to damaged muscle apparently via the circulation, these stem cells proliferate and contribute to the newly formed muscle fibers, supplementing the contribution of local satellite cells. Because these bone marrow cells can be isolated easily and grown in the laboratory, regenerative biologists are investigating whether such stem cells can be developed as grafts for restoring healthy cells in degenerative muscle or brain disorders.

Central nervous system. Brain and spinal cord research is of prime importance within regenerative medicine because of the high frequency and devastating effects of central nervous system (CNS) injuries. Progress remains slow but has been sufficient to dispel the long-held belief that mammalian CNS tissues are inherently unable to regenerate after injury. Limited reorganization and compensation do occur within the injured adult spinal cord, and such tissues now appear to be amenable to therapeutic manipulation. Moreover, new populations of stem cells have been discovered in the CNS. In peripheral nerves regenerative growth of cut or injured nerve fibers (axons) usually occurs readily with only partial interference from scar tissue. In the injured CNS, however, axonal regeneration appears to be blocked physically by a scar produced by several types of local cells and inhibited chemically by exposure to factors produced by these cells. Current attempts to circumvent these blocks to neural regeneration in the CNS include (1) local treatment with factors designed to reduce scar formation, inactivate inhibitors of axon growth, and stimulate nerve fiber regeneration and (2) grafting of stem cells, peripheral nerve tissue, or cells capable of ensheathing and facilitating regrowth of axons. *See* NERVE.

Appendages. Mammals have little capacity to regenerate missing appendages. One exception is the annual regrowth of shed antlers in deer. Regeneration involves a rapidly growing bud of skin-covered tissue that forms cartilage, then bone, until fully grown, at which point the skin is lost. In humans, especially children, clinical experience has shown that fingertips can regenerate fully if the wound is not closed with sutures but is allowed to heal normally with skin. Such observations of regenerative ability hold the promise that regrowth of missing appendages may be inducible in mammals once the various required factors are known.

Role of the immune system. Another model of mammalian appendage regeneration, the reconstitution of tissue to close large holes produced in the cartilaginous portion of the ear, which was thought to occur only in rabbits, was recently found in a strain of mice with specific defects in the immune system. This regenerative capacity in mice with immunological deficits is consistent with the hypothesis that the decline in regenerative ability during the evolution of mammals may correlate with the rise of more complex and efficient adaptive immune systems which allow mammals to mount rapid inflammatory responses in wounds to avoid infection. Such inflammation usually culminates in scar formation, which has been shown in many systems to be incompatible with a normal regenerative response. Animals that lack highly developed adaptive immunity, including invertebrates and primitive vertebrates, generally have greater regenerative potential than mammals. Regeneration of ear tissue in genetically well-characterized immune-deficient mice is important since it may allow molecular dissection of the genes required for the regenerative response in mammals. *See* CELLULAR IMMUNOLOGY; IMMUNITY.

Regeneration of damaged skin, without scarring and with good restoration of hair follicles, occurs if surgical injury is performed at an embryonic stage of skin development, and this has become an experimental model for regeneration in mammals. If skin injury is delayed until later stages of development, regeneration becomes progressively more imperfect as evidenced by scar formation. Research in scar-free wound healing using mice with various genes

"knocked out" again points to roles for cells of the immune system in the response to injury that leads to scarring. The immune system forms during the same embryonic period in which the potential for scarring appears. Greater understanding of the molecular events that lead to scar formation is yielding new treatments to mitigate postsurgical scarring and may yield additional clues regarding the loss of regenerative ability in skin and other organs.

Limb regeneration in amphibians. Of all vertebrates, amphibians have the most highly developed capacity for regeneration. Certain species have the ability to regenerate not only limbs and tails but also parts of the eye, lower jaw, intestine, and heart.

Impact of metamorphosis. Complete regeneration of amputated limbs can occur throughout the lifetime of most urodeles (salamanders and newts) but usually occurs much more rapidly in larval animals. In anurans (frogs and toads) the ability to regenerate limbs is lost during metamorphosis to the adult form. This transition from a regenerative to nonregenerative capacity in frogs has been investigated for insights into the more general inability of other vertebrates to regenerate missing appendages. As with the ontogeny of skin scarring in the mammalian embryo, the evidence suggests a primary role for the developing immune system in the loss of regenerative ability. Comparative immunologists have shown that many aspects of immunity in larval anurans are primitive and that during metamorphosis the adaptive immune system undergoes changes that result in adult frogs having immunity similar to that of mammals. Urodeles, which retain well-developed regenerative potential throughout life, display the more primitive type of adaptive immunity found in frog larvae. Evolution of the adaptive immune system, many aspects of which are involved in inflammation and defense against pathogens during wound healing, may have produced mechanisms that lead to scarring and failure of epimorphic regeneration, a possibility currently under active investigation. *See* METAMORPHOSIS.

Mechanism. All amphibian limbs that are capable of regeneration show essentially the same sequence of events following amputation (**Fig. 1**). The end of the limb stump is rapidly covered with epidermis that migrates out from the cut edge of the skin. Injured tissues immediately beneath this newly formed wound epithelium begin to dissociate, while much of the extracellular material is broken down and most specialized characteristics of the cells in these tissues are lost. This process, termed dedifferentiation, results in the appearance of cells similar to those in the embryonic limb bud. As they dedifferentiate, the cells begin to proliferate and soon form a mass of cells, the blastema, below the epidermis at the end of the limb stump. Continued proliferation of the blastema cells causes a budlike outgrowth on the stump. The regeneration bud gradually elongates into a conical structure that flattens and becomes paddle-shaped. During elongation of the blastema, proximal cells adjacent to tissues in the stump begin to differentiate. Capillaries grow into the blastema from the stump, and new cartilage begins to form in the area contiguous with cut skeletal elements of the stump. By the paddle stage, cartilage formation is under way in central areas throughout most of the outgrowth. Muscle and connective tissue also redifferentiate from blastemal cells in a proximal-to-distal direction. Reformation of the amputated parts of the limb is completed as digits develop along the distal end of the flattened structure and articulate with skeletal elements differentiating in that region. The typical regenerated limb is morphologically and functionally identical to the amputated appendage. *See* EMBRYONIC DIFFERENTIATION.

Role of electric currents. As in simple wound healing and tissue repair, the initial local stimuli for limb regeneration are generated as a result of injury and include enzymes and growth factors released from damaged tissue and from immigrating white blood cells. In addition to these biochemical factors, electrical properties may be important during the onset of regeneration. Currents driven by the uptake of sodium ions across the skin can be measured leaving the ends of amputated limbs and are much greater in regenerating species, such as newts and salamanders, than in adult frogs. Moreover, reducing the flow of current by various methods inhibits limb regeneration, while increasing the current associated with frog limb stumps by means of implanted batteries produces a regenerative response in this animal. The cellular basis for the electrical effects on growth is

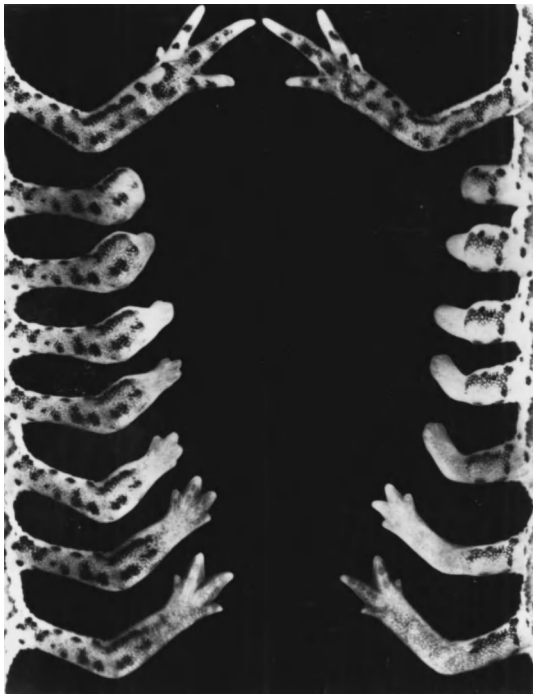

Fig. 1. Amputation of the forelimb of a newt, either through the level of the radius and ulna (left) or through the humerus (right), begins a process of regeneration that includes several morphological stages and results in an anatomically complete and fully functional new limb. The new four-digit limb formed by regeneration will continue to grow until the size of the original limb is attained. (*From R. J. Goss, Principles of Regeneration, Academic Press, 1969*)

not known, but experimental evidence suggests that endogenous currents are important for initiation of the limb regeneration process. *See* BIOPOTENTIALS AND IONIC CURRENTS.

Tissue interactions. The population of undifferentiated cells formed in the aftermath of amputation injury grows and develops the missing limb components through interactions with other limb tissues. As with limb development in embryos, the process of amphibian limb regeneration depends on an inductive influence from the thickened epithelium at the distal end of the growing appendage. In both developmental processes, proliferation of underlying cells is inhibited, and growth of the limb is blocked if the epithelium is removed or prevented from forming. Grafts of full-thickness skin over an amputation site preclude the interaction between epidermis and internal tissues and thereby inhibit regeneration. Deviation of the wound epithelium from its normal location results in asymmetrical outgrowth of the regenerated portion; grafting additional wound epithelia to a limb stump causes supernumerary limbs to develop. The epithelial effect in both embryonic and regenerating limbs is likely to involve release of growth-promoting factors and enzymes to break down extracellular components. The effect of these factors during regeneration is to transform a tissue-repair process into the process of blastema formation and epimorphic regeneration.

Another tissue interaction that is critical for appendage regeneration in amphibians involves the nerves in the limb stump. Cut ends of axons in motor, sensory, and autonomic nerves grow into the developing blastema and are present throughout this tissue during its period of active growth. If this elongation of axons is prevented or curtailed by cutting the nerves at a more proximal level, further growth of the early blastema is inhibited. Limbs of larval salamanders will resume regeneration when the nerves themselves regenerate to distal areas of the limb stump. Moreover, limbs without their normal supply of nerves can regenerate if neurons are grafted into the stump.

The mechanism by which nerves exert their stimulatory influence on regeneration involves secretion by axons of nutritive or growth-triggering factors. Although these factors have not yet been identified with certainty, the effect is clearly not dependent on a specific type of nerve: sensory, motor, and sympathetic nerves can all support blastema formation if present in sufficient quantity. The quantitative requirement is indicated by the observation that a limb will not regenerate unless it contains a certain minimum number of nerve fibers at the level of amputation. The number of nerve fibers necessary to support regeneration varies with the size of the limb and level of amputation. The neural influence is important primarily for mitotic activity of the dedifferentiated cells leading to establishment and growth of the blastema. Removal of nerves from a limb does not inhibit epidermal closure of an amputation site or the onset of internal tissue dissociation, nor does cutting the nerves after a well-developed blastema has appeared prevent redifferentiation and formation of the replacement limb. Most available evidence, both from analyses of proliferative activity of blastema cells and from studies on proteins transported and released from regenerating axons, is consistent with the view that nerves release factors specifically required for DNA replication in proliferating cells of the early blastema.

Artificial induction of limb regeneration in vertebrates. Various degrees of regenerative growth have been stimulated in certain nonregenerating animals by manipulation of the factors important for regeneration in salamanders. For example, regeneration has been reported in the limbs of adult frogs following amplification of the current leaving the stumps, after localized applications of growth factors, and after augmentation of the nerve supply to the limb. Grafts of neural tissue to limb stumps have also been shown to induce regeneration in lizards.

Most attempts to induce regeneration in warm-blooded vertebrates have been relatively unsuccessful, with the exception of the regeneration of lost fingertips in young children. Restorative growth of the digit depends on how the injury is treated: the wound must be allowed to close itself by simple epithelial growth and contraction, rather than being surgically closed with sutures. The importance of how the amputation wound heals suggests that human fingertip regeneration involves a positive influence from the epidermis on the underlying dissociated tissues, like that normally operative during amphibian limb regeneration.

Morphogenesis during limb regeneration. The manner in which specific limb tissues, such as individual muscles and bones, assume their unique morphological structures and locations during reconstitution of the complete appendage is a major unanswered question in regeneration biology.

Research into such questions of morphogenesis during amphibian limb regeneration is relevant to understanding similar events of pattern formation that occur during embryonic limb development. An experimental tool for analyzing these events was discovered with the observation that retinoids, a class of compounds related to vitamin A, are able to respecify in a predictable manner the positions in which muscles and skeletal elements develop in both embryonic and regenerating vertebrate limbs. Local application of retinoic acid at a specific site on the embryonic limb bud leads to a duplication in the anterior-posterior axis of the limb, so that six digits develop instead of the normal three. Application of retinoic acid during amphibian limb regeneration affects all axes of the developing limb, but the effect on the proximal distal axis, which leads to the duplication of structures present proximal to the level of amputation, has been best studied. For example, amputation through the wrist of a salamander and treatment with vitamin A can result in a blastema at the wrist that will give rise to structures of the upper arm and lower arm, as well as to a complete new hand (**Fig. 2**). Observations that concentration gradients of retinoic acid and binding proteins for

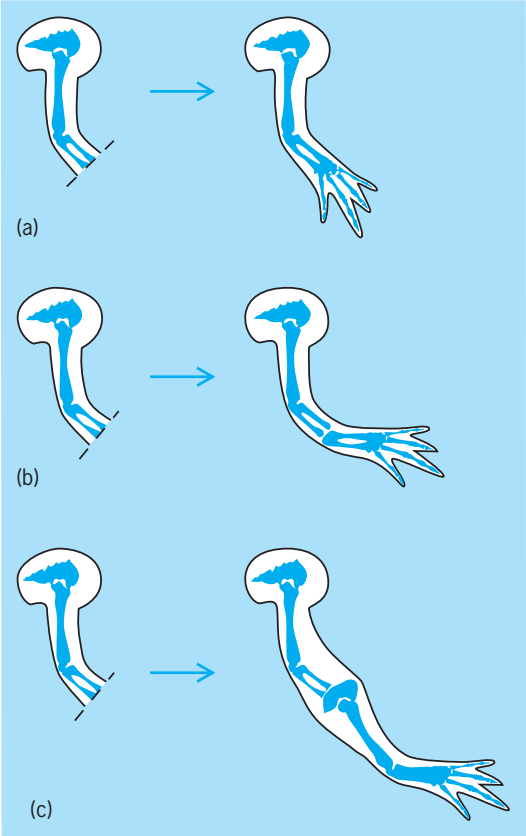

Fig. 2. Effect of a retinoid such as vitamin A on the proximal-distal axis of a regenerated amphibian forelimb. (*a*) Normal regeneration of the forelimb after amputation through the radius and ulna (broken line). (*b*) Regeneration after treatment with a low dose of vitamin A; the duplicate radius and ulna have formed in the regenerated limb. (*c*) Regeneration after treatment with a high dose of vitamin A; the entire forelimb, including the scapula, has been duplicated in the regenerated limb.

retinoic acid exist in developing limb buds suggest that retinoids may be among the endogenous factors, or morphogens, involved in the control of limb pattern formation. *See* PATTERN FORMATION (BIOLOGY).

The apparent involvement of retinoids in the control of limb morphogenesis has provided an opportunity for applying the methods of molecular biology in studies of regeneration. Genes for retinoic acid receptors, proteins to which retinoids bind before exerting their effect on expression of a cell's genetic program, have been cloned from salamanders. These genes are expressed in dedifferentiating cells of the limb stump as the blastema begins to form, and are important in establishing the positional identity of the cells within the new limb, although the mechanism by which retinoic acid and its receptors effect positional information in the regenerating limb remains unclear. Other genes implicated in pattern formation in other developing systems have also been shown to be active in regeneration blastemas, and differences in the rate of expression related to the proximal-distal location of the blastema cells have been observed. The ability to measure expression of such genes by molecular methods, together with the use of retinoid treatments and surgical methods to manipulate morphogenesis in limbs, should lead to an increased understanding of pattern formation during limb regeneration and development. *See* ANIMAL MORPHOGENESIS; VITAMIN A.

Tail regeneration. A well-developed capacity for tail regeneration is found in larval amphibians, as well as in certain species of adult salamanders. Many lizards can also regenerate tails lost through injury or by autotomy during escape from a predator. Autotomy is produced by a strong contraction of the segmental muscles of the tail, causing one of the vertebrae to come apart along a preformed breakage plane and separating the muscles and other tissues so that the distal portion of the tail is detached from the body.

Whether lost by autotomy or injury, the tail is regenerated by a process superficially similar to that of amphibian limbs. Tissues dissociate locally in the injured area of the stump, during which cells dedifferentiate and begin the period of proliferation that eventually gives rise to the regeneration blastema. Formation of new muscles, axial skeleton, and connective tissue from blastemal cells proceeds in a proximal-to-distal direction. Depending on the species, the morphology and tissue composition of the regenerated tail are abnormal to various degrees. In the lizard, vertebrae are not reconstituted; instead, the regenerated spinal cord is surrounded by a simple tube of flexible cartilage.

An important developmental process unique to tail regeneration is reconstitution of the spinal cord. Cells lining the central neural canal at the severed end of the cord grow out to form a tubular structure that extends into the developing blastema. Neurons and nerve fibers from proximal, intact regions of the spinal cord grow along this tubular outgrowth, and a functional though anatomically abnormal cord is reestablished. In salamanders, but not in lizards, paired spinal ganglia are also formed along the new cord, developing from cells that migrate out of the regenerating cord when it is a simple growing tube of cells.

In all lizards and amphibians (with the apparent exception of larval frogs), an inductive influence from the spinal cord is necessary for tail blastema formation and regeneration of axial structures. If the spinal cord is removed from the site of amputation and prevented from regrowing into that area, regeneration of the tail will not occur. The neurons and other cells around the central canal of the spinal cord appear to be the most important source for this growth-promoting influence. Spinal ganglia and peripheral nerves deviated into the tail stump are unable to replace the spinal cord in allowing growth of a tail regeneration blastema. This property of the spinal cord is similar to the activity of peripheral nerves during limb regeneration, but whether the same chemical factors and cellular interactions are involved is not known. *See* NERVOUS SYSTEM (VERTEBRATE).

Replacement of structures in the eye. Other noteworthy examples of regeneration found primarily in amphibian species involve certain structures of the eye. Surgical removal of the neural retina can be followed by reconstitution of this tissue from the

deeper pigmented layer of the retina. From the new retina a new optic nerve is regenerated. Removal of the iris is followed by its complete replacement, including new muscles and connective tissue, from the edge of the adjacent retina. With various species of salamanders, lens removal initiates a remarkable process by which epithelial cells on the dorsal rim of the iris dedifferentiate and proliferate as a mass of cells that develops into a new lens. This phenomenon is dependent on a neural influence from the retina. Certain larval frogs regenerate a missing lens from epithelium of the cornea, which more closely resembles lens formation in vertebrate embryos. Such transformations of one specialized ocular tissue into another make possible the reconstitution of a fully functional visual system. *See* EYE; VISION.

Regeneration in invertebrates. Protozoa and simple multicellular animals, including sponges, cnidaria (for example, *Hydra*), and flatworms (for example, *Planaria*), display remarkable capacities for regeneration following various experimental manipulations. Regenerative ability in such organisms correlates closely with their capacity to reproduce asexually, most commonly by fission or by budding, and the mechanisms of growth involved in regeneration are often very similar to those of asexual propagation. For example, just as complete ciliated protozoa will develop after fission, which divides the nucleus and organelles between daughter cells, intact individuals will also be reconstituted from fragments of a single organism if the fragment contains a complete set of the genetic material and a portion of the original cell's cortical cytoplasm. Similar rules regarding the importance of the nucleus apply to regenerative processes in all protozoa.

Porifera. Regeneration in sponges can be considered a variant of the method of asexual reproduction used by these primitive multicellular organisms. If part of a sponge is cut away, that portion will usually not grow back to any significant extent. However, if a sponge is mechanically or chemically dissociated into its component cells, a population of reserve cells called amebocytes will reaggregate and give rise to an intact new organism. This regenerative phenomenon is similar to the asexual process of gemmule formation in sponges. In this reproductive procedure, cells are packaged in a special structure (gemmule), which enables them to survive the death of the organism during certain adverse conditions and to emerge later and reconstitute an intact sponge. *See* PORIFERA.

Cnidaria. Among cnidaria, hydroids and sea anemones have considerable regenerative capacity, whereas jellyfish regenerate poorly. Regeneration in fresh-water *Hydra* has been studied extensively. Such animals have a definite anatomical polarity, with a base for attachment to a substrate and a hypostome (mouth) surrounded by tentacles at the opposite end. If a hydra is cut in half, the part containing the hypostome and tentacles will form a new basal half for attachment, and the original basal part will form a mouth and tentacles for a new apical end. A dissected central portion of the animal, having neither tentacles nor basal structures, will regenerate a new base at one end and tentacles at the other. Even tiny fragments of an individual hydra can regenerate all the parts for complete new organisms. Polarity is maintained not only during regeneration but also during the process that allows individuals to reconstitute themselves when several hydra are grafted surgically end to end.

Regeneration in *Hydra* usually involves growth and differentiation of interstitial cells situated among the cells of the organism's external layer. These cells represent a reserve population important for the steady replacement of certain specialized cells and for asexual reproduction by budding. Polarity appears to be maintained during regeneration and budding by the action of two morphogens, a head activator and a head inhibitor, which are present in gradients along the length of the organism and which together control formation of the hypostome and tentacles. *See* CNIDARIA.

Planaria. Gradients of regulative factors also appear to control regeneration in flatworms, such as *Planaria*. An adult flatworm is highly organized, having a head with eyes and a simple brain, complex digestive, reproductive, and excretory organs, and a tail. Regeneration of complete flatworms from dissected pieces occurs in a polarized fashion, provided a head-tail gradient of morphogens is present (**Fig 3**). New body parts are formed during regen-

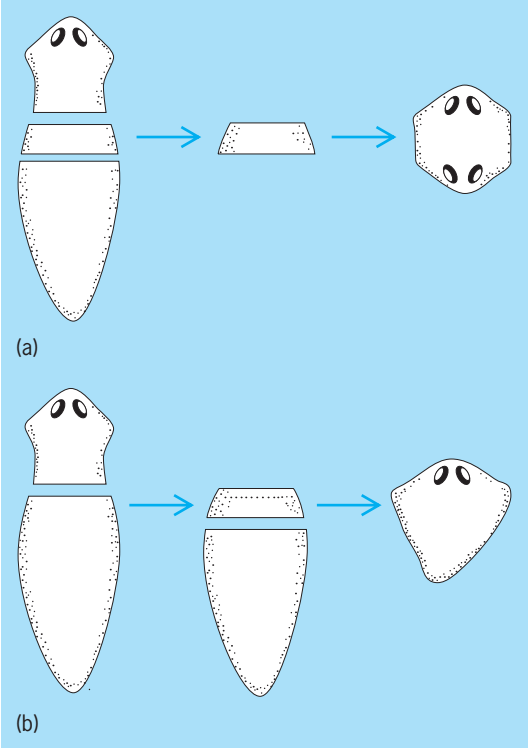

Fig. 3. Relationship of the anterior-posterior gradient to the product of *Planaria* sectioning. (*a*) A thin section of *Planaria* too narrow to establish an anterior-posterior gradient may form two heads facing opposite directions. (*b*) If the posterior cut is delayed, the thin section will form a worm of normal polarity because the delay has allowed an anterior-posterior gradient to be established. (*After R. J. Goss, Principles of Regeneration, Academic Press, 1969*)

eration from a population of reserve cells called neoblasts, which reside in connective tissue below the epidermis. This mechanism of growth in flatworms is similar to that following reproductive fission in these animals. Species incapable of asexual propagation also have poorly developed regenerative capacities. *See* TURBELLARIA.

Annelida. Most annelids, such as the earthworm, can readily regenerate segments after their removal: some species can regenerate whole organisms from any fragment. Like more primitive invertebrates, certain annelids can reproduce asexually by transverse fission. The capacities for fission and for reconstitution from fragments in annelids are remarkable, considering the anatomical complexity of animals in this phylum. Annelids have well-developed muscular systems associated with primitive exoskeletons and efficient digestive systems. The nervous system consists of an anterior brain, complex sense organs, and a highly organized network of motor and sensory nerves. When an earthworm is cut transversely into two parts, the anterior part can regenerate several posterior segments. The ability of the posterior half to regenerate anterior segments is, however, more limited and is absent altogether in some species. Experiments in which components of the nervous system are removed surgically have revealed the importance of a neural influence for segment regeneration, presumably mediated by a growth-stimulating hormone secreted from neural cells. In most species of annelids, cells of the regenerated segments are derived both from populations of reserve cells and from dedifferentiation of existing tissues. *See* ANNELIDA.

Echinodermata. The ability of certain echinoderms, such as starfish, to regenerate missing arms is well known. Cutting such an animal into several pieces results in each piece forming a new organism, a phenomenon that usually requires the presence of at least some of the central portion of the body. Some starfish species, however, can actually regenerate an entire body from the cut end of an isolated arm. Equally remarkable is a regenerative response shown by another echinoderm, the sea cucumber. When this animal is strongly irritated, it undergoes the unusual behavior of eviscerating itself through its anus or through a rupture of its body wall. This phenomenon, which may serve as a protective mechanism, produces a nearly empty sack of skin and muscle, which then proceeds to regenerate all the internal organs, beginning with the digestive tract. *See* ECHINODERMATA.

Arthropoda. The capacity for appendage regeneration is widespread among the many diverse members of the phylum *Arthropoda*. In these complex animals with well-developed exoskeletons and no asexual mode of reproduction, regeneration shows a close correlation with the molting process. Insects regenerate legs and antennae readily during their long larval and pupal stages, but lose this capacity as adults. Crustaceans, such as lobsters, crayfish, and crabs, continue to molt throughout life and retain regenerative ability as adults, regrowing missing appendages at each molt. Many species of crustaceans have capitalized on this regenerative ability by developing the capacity for autotomy of legs caught by predators. Regeneration in other forms of arthropods, including centipedes, millipedes, and spiders, occurs to varying degrees but has not been as well studied as that in crustaceans and insects. *See* ARTHROPODA; DEVELOPMENTAL BIOLOGY. Anthony Mescher

Bibliography. C. E. Dinsmore (ed.), *A History of Regeneration Research: Milestones in the Evolution of a Science*, 1991; P. Ferretti and J. Geraudie (eds.), *Cellular and Molecular Basis of Regeneration: From Invertebrates to Humans*, 1998; H. G. Gard and M. T. Langaker (eds.), *Scarless Wound Healing*, 2000; L. V. Polazhaev, *Loss and Regulation of Regenerative Capacity in Tissues and Organs of Animals*, 1990; D. L. Stocum, *Wound Repair, Regeneration and Artificial Tissues*, 1994; P. A. Tsonis, *Limb Regeneration*, 1996; I. V. Yannas, *Tissue and Organ Regeneration in Adults*, 2001.

Regolith

The mantle or blanket of unconsolidated or loose rock material that overlies the intact bedrock and nearly everywhere forms the land surface. The regolith may be residual (weathered in place), or it may have been transported to its present site. The undisturbed residual regolith may grade from agricultural soil at the surface, through fresher and coarser weathering products, to solid bedrock several meters or more beneath the surface (see **illus.**). The

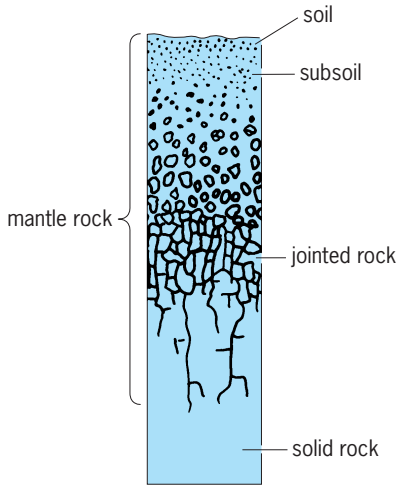

Diagram showing the components of residual regolith.

transported regolith includes the alluvium of rivers, sand dunes, glacial deposits, volcanic ash, coastal deposits, and the various mass-wasting deposits that occur on hillslopes. Scientific results from the Apollo space program show that the lunar surface also has a regolith. This layer of fragmental debris is believed to derive from prolonged meteoritic and secondary fragment impact. *See* SOIL; WEATHERING PROCESSES. Victor R. Baker

Regular polytopes

The n-dimensional analogs of the regular polygons ($n = 2$) and platonic solids ($n = 3$). They are conveniently denoted by their Schläfli symbols $\{p, q, \ldots\}$; for instance, the pentagon, hexagon, octagon, tetrahedron, octahedron are denoted by {5}, {6}, {8}, {3,3}, {3,4}. The cube is {4,3} because its faces are squares {4} and there are three of them at each vertex. The five platonic solids $\{p,q\}$, which are the subject of Euclid's *Elements*, Book XIII, are determined by inequality (1). The numbers of vertices, edges, faces (V,E,F), as listed in the **table**, can be deduced from the obvious relations $pF = 2E = qV$ with the help of Euler's formula $V - E + F = 2$.

$$(p-2)(q-2) < 4 \qquad (1)$$

The general polytope [sometimes loosely called a polyhedron (**Fig. 1**) regardless of the number of dimensions] is a finite region of n-dimensional space enclosed by a finite number of hyperplanes. When any redundant hyperplanes have been discarded, those that remain contain $(n - 1)$-dimensional polytopes called cells or "facets." For instance, the cells of a polygon are its sides, those of a polyhedron are its faces, and those of a four-dimensional polytope are solids.

The platonic solid $\{p,q\}$ is said to be regular because its faces are regular and its vertices are all surrounded alike. The four-dimensional regular polytope $\{p,q,r\}$ has three-dimensional solid cells $\{p,q\}$, r of which surround each edge; the five-dimensional regular polytope $\{p,q,r,s\}$ has cells $\{p,q,s\}$, s of which surround each plane face; and so on.

The six regular four-dimensional polytopes $\{p,q,r\}$ are determined by inequality (2). The 5-cell {3,3,3}

$$p - \frac{4}{p} + 2q + r - \frac{4}{r} < 12 \qquad (2)$$

(**Fig. 2***a*) may be drawn in perspective as a pentagon with its 5 diagonals, though in reality its 10 edges are all equal. The simplest drawing of the 8-cell {4,3,3} (Fig. 2*b*) consists of an octagon with a square placed inward on each side. The squares on two alternate sides are easily visualized as two opposite faces of a cube. The 8 such cubes are the cells of the 8-cell.

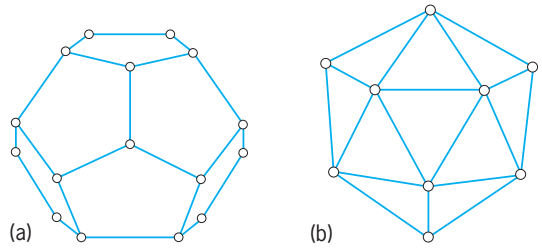

Fig. 1. Two regular polyhedrons. (*a*) Dodecahedron {5,3}. (*b*) Icosahedron {3,5}.

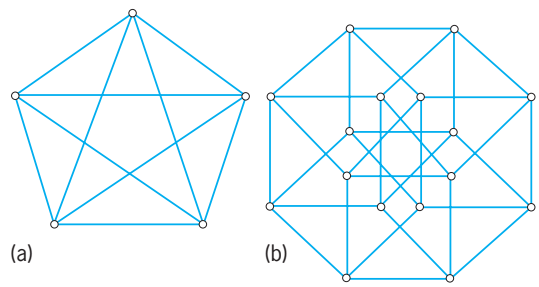

Fig. 2. Two of the six regular four-dimensional polytopes. (*a*) The 5-cell {3,3,3}. (*b*) The 8-cell {4,3,3}.

The regular tetrahedron can be inscribed in the cube, in the sense that the 4 vertices of the former occur among the 8 vertices of the latter. In the same sense, the cube can be inscribed in the dodecahedron, the 16-cell in the 8-cell, the 8-cell in the 24-cell, the 24-cell in the 600-cell, the 600-cell (and also the 5-cell) in the 120-cell.

The vertices of the n-dimensional simplex $\{3,3,\ldots,3\}$ consists of $n + 1$ points, all equidistant from one another. Those of the cross polytope $\{3,\ldots,3,4\}$ are at equal distances from the origin in both

Regular polytopes in *n* dimensions

Polytope	Schläfli symbol	Vertices	Edges	Faces	Solid cells	Hypersolid cells
$n=2$						
p-gon	$\{p\}$	p	p			
$n=3$						
Tetrahedron	{3,3}	4	6	4		
Cube	{4,3}	8	12	6		
Octahedron	{3,4}	6	12	8		
Dodecahedron	{5,3}	20	30	12		
Icosahedron	{3,5}	12	30	20		
$n=4$						
5-cell	{3,3,3}	5	10	10	5	
8-cell	{4,3,3}	16	32	24	8	
16-cell	{3,3,4}	8	24	32	16	
24-cell	{3,4,3}	24	96	96	24	
120-cell	{5,3,3}	600	1200	720	120	
600-cell	{3,3,5}	120	720	1200	600	
$n > 4$						
Simplex	$\{3,3,\ldots,3\}$	$n + 1$	$\frac{1}{2}n(n + 1)$	....		$n + 1$
Hypercube	$\{4,3,\ldots,3\}$	2^n	$2^{n-1}n$	....		$2n$
Cross polytope	$\{3,\ldots,3,4\}$	$2n$	$2n(n - 1)$	....		2^n

directions along the n coordinate axes; thus their coordinates (for a cross polytope of edge $\sqrt{2}$) are the permutations of $(\pm 1,0,\ldots,0)$. Similarly, the 2^n vertices of the hypercube $\{4,3,\ldots,3\}$, of edge 2, are $(\pm 1,\pm 1,\ldots,\pm 1)$. *See* ANALYTIC GEOMETRY; EUCLIDEAN GEOMETRY. H. S. M. Coxeter

Bibliography. A. Bronosted, *An Introduction to Convex Polytopes*, 1982; H. S. M. Coxeter, *Regular Complex Polytopes*, 2d ed., 1991; H. S. M. Coxeter, *Regular Polytopes*, 2d ed., 1963, reprint 1973; K. Miyazak, *An Adventure in Multidimensional Space: The Art and Geometry of Polygons, Polyhedra, and Polytopes*, 1986; G. M. Ziegler, *Lectures on Polytopes*, 1994.

Regularia

The name given by G. Cuvier in 1817 to an assemblage of echinoids in which the anus and periproct lie within the apical system. The test is globular and preponderantly radially symmetrical, and the ambulacral plates are commonly compound. The group included, in effect, all those echinoids which did not fall in the Irregularia. The Irregularia, however, were shown to be polyphyletic by J. Durham and R. Melville in 1957, and so neither the Irregularia nor the Regularia constitute valid taxa. *See* ECHINOIDEA; IRREGULARIA. Howard B. Fell

Regulator

A control device designed to maintain the value of some quantity substantially constant. Thus, a temperature regulator is a device designed to maintain the temperature of some environment at a constant value. The value to be maintained can usually be established at any value within the range of the regulator by making an appropriate setting.

A regulated system is a feedback control system employing a regulator to maintain some quantity of the system at a constant value. Another example is the voltage-regulator system of an automobile. *See* CONTROL SYSTEMS. John A. Hrones

Reheating

The addition of heat to steam of reduced pressure after the steam has given up some of its energy by expansion through the high-pressure stages of a turbine. The reheater tube banks are arranged within the setting of the steam-generating unit in such relation to the gas flow that the steam is restored to a high temperature. Under suitable conditions of initially high steam pressure and superheat, one or two stages of reheat can be advantageously employed to improve thermodynamic efficiency of the cycle. *See* STEAM-GENERATING UNIT; STEAM TURBINE; SUPERHEATER; VAPOR CYCLE. R. A. Miller

Reinforced concrete

Portland cement concrete containing higher-strength, solid materials to improve its structural properties. Generally, steel wires or bars are used for such reinforcement, but for some purposes glass fibers or chopped wires have provided desired results.

Unreinforced concrete cracks under relatively small loads or temperature changes because of low tensile strength. The cracks are unsightly and can cause structural failures. To prevent cracking or to control the size of crack openings, reinforcement is incorporated in the concrete. Reinforcement may also be used to help resist compressive forces or to improve dynamic properties.

Steel usually is used in concrete. It is elastic, yet has considerable reserve strength beyond its elastic limit. Under a specific axial load, it changes in length only about one-tenth as much as concrete. In compression, steel is more than 10 times stronger than concrete, and in tension, more than 100 times stronger. *See* STEEL.

As reinforcement for concrete, steel may be used in the form of bars or rods, wire, fibers, pipe, or structural shapes such as wide-flange beams (see **illus.**). Bars are the most commonly used form. The size of a bar usually is specified by a number that is about eight times the nominal diameter; sizes range from No. 2, a nominal $^1/_4$-in. (0.6-cm) bar, to No. 18, a nominal $2^1/_4$-in. (5.7-cm) bar.

During construction, the bars are placed in a form and then concrete from a mixer is cast to embed them. After the concrete has hardened, deformation is resisted and stresses are transferred from concrete to reinforcement by friction and adhesion along the surface of the reinforcement. The bonding may be improved mechanically by giving reinforcing bars raised surfaces; such bars are known as deformed bars. They are generally manufactured to standard specifications, such as those of the American Society for Testing and Materials (ASTM). Similarly, deformed wires used as reinforcement provide greater bond than smooth wires. Deformed wires

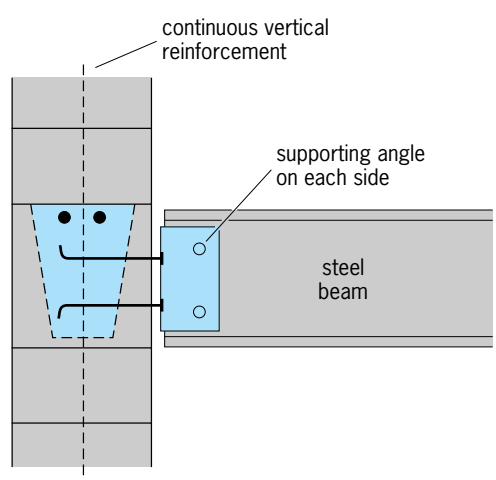

Continuous vertical reinforcement in concrete. (*After C. Beall, Masonry Design and Detailing for Architects, Engineers, and Contractors, 3d ed., McGraw-Hill, 1993*)

are designated by D followed by a number equal to 100 times the nominal area in square inches.

Individual wires or bars resist stretching and tensile stress in the concrete only in the direction in which such reinforcement extends. Tensile stresses and deformations, however, may occur simultaneously in other directions. Therefore reinforcement must usually be placed in more than one direction. For this purpose, reinforcement sometimes is assembled as a rectangular grid, with clips or welded joints at the intersections of the wires or bars. For example, prefabricated wire grids, called welded-wire fabric, often are used for slab reinforcement in highway pavement and in buildings.

Bars, grids, and fabric have the disadvantage that the principal effect of reinforcement occurs primarily in the plane of the layer in which they are placed. Consequently, the reinforcement often must be set in several layers or formed into cages.

Under some conditions, fiber-reinforced concrete is an alternative to such arrangements. The fiber, made of fine steel wires, glass fibers, or plastic threads, is embedded in the concrete in short lengths, often only about 1 in. (2.5 cm). They are added to the concrete mixer along with the other ingredients. To achieve crack control in all directions, the fibers should be uniformly spaced, at close intervals, and randomly oriented throughout the hardened concrete. Such reinforcement also improves concrete tensile strength, ductility, and dynamic properties. *See* COMPOSITE BEAM; CONCRETE; CONCRETE BEAM; CONCRETE COLUMN; CONCRETE SLAB; PRESTRESSED CONCRETE. Frederick S. Merritt

Bibliography. C. Beall, *Masonry Design and Detailing: For Architects, Engineers, and Contractors*, 4th ed., 1997; K. C. Hover, *Construction with Reinforced Concrete*, 1994; R. R. Schneider and W. L. Dickey, *Reinforced Masonry Design*, 3d ed., 1994.

Relapsing fever

An acute infectious disease characterized by recurring fever. It is caused by spirochetes of the genus *Borrelia* and transmitted by the body louse (*Pediculus humanus humanus*) and by ticks of the genus *Ornithodoros*.

Louse-borne relapsing fever, caused by *Borrelia recurrentis*, is typically epidemic. Its occurrence depends upon ecological and socioeconomic conditions favoring heavy infestations by body lice. Humans are the only reservoir. Epidemics, once widespread on all continents, are rare but still occur in certain parts of South America, Africa, and Asia.

Tick-borne relapsing fevers are endemic. They are more widely distributed throughout the Eastern and Western hemispheres. At least 15 species of *Borrelia* have been recognized as causative agents.

After incubation of 2–10 days, the initial attack begins abruptly with chills, high fever, headache, and pains in muscles and joints, and lasts 2–8 days, ending by crisis. A remission period of 3–10 days is followed by a relapse similar to the initial attack but milder. There may be 4–5 relapses, although occasionally 10 or more have been recorded as a result of antigenic variations in the causative *Borrelia*. Mortality varies from 2 to 5% but may be considerably higher during epidemics.

Diagnosis and treatment. Definite diagnosis depends upon detecting spirochetes in the blood during fevers. They may be seen by dark-field examination of wet blood preparations, or in Giemsa-stained thick and thin blood films. Where symptoms are present but spirochetes are absent or scanty, intraperitoneal inoculation of adult or suckling Swiss mice or rats with blood is useful to amplify the spirochetes. Detection of spirochetes in lice or ticks is by dark-field examination of hemolymph or tissues, and by inoculation of tissue suspensions into suckling and adult Swiss mice or rats.

Ability of borreliae to undergo antigenic phase variations make the serodiagnosis of relapsing fevers difficult. Of the many tests proposed, indirect immunofluorescence with cultured spirochetes gives the most dependable results. *See* IMMUNOFLUORESCENCE.

Chlortetracycline is the most effective antibiotic drug, but penicillin, oxytetracycline, and streptomycin also have therapeutic value. *See* ANTIBIOTIC.

Transmission and prevention. In lice, *B. recurrentis* produces an infection limited to the hemolymph. Thus, transmission does not occur by bite but by contamination of the bite wound with infectious hemolymph of lice that are smashed by scratching. Borreliae may penetrate normal, unbroken skin.

Tick-borne borreliae, on the other hand, produce in their tick vectors a generalized infection, including tissues of salivary glands. Therefore, transmission is via infected saliva or infectious body fluids (coxal fluid) excreted during feeding. In several tick vectors, the presence of spirochetes in ovarian tissues leads to the passage of organisms via eggs (transovarian infection).

The best way to prevent relapsing fever is to control louse and tick populations with effective insecticides and acaricides. *See* MEDICAL BACTERIOLOGY; PEDICULOSIS. Willy Burgdorfer

Bibliography. M. Kusinitz, *Tropical Medicine*, 1990.

Relative atomic mass

The ratio of the average mass per atom of the natural nuclidic composition of an element to $^1/_{12}$ of the mass of an atom of nuclide ^{12}C. For example, $\mu(Cl) = 35.453$. Relative atomic mass replaces the concept of atomic weight. It is also known as relative nuclidic mass. *See* ATOMIC MASS; NUCLIDE. Thomas C. Waddington

Relative molecular mass

The ratio of the average mass per formula unit of the natural nuclidic composition of a substance to $^1/_{12}$ of the mass of an atom of nuclide ^{12}C. For example,

μ(KCl) = 74.555. Relative molecular mass replaces the concept of molecular weight. *See* MOLECULAR WEIGHT; NUCLIDE. Thomas C. Waddington

Relative motion

All motion is relative to some frame of reference. The simplest laboratory frame of reference is three mutually perpendicular axes at rest with respect to an observer. Such a system is commonly used in the laboratory when various types of motion are being studied. The general effects of other motions to which the system as a whole is subjected are then neglected and the system is said to be isolated. In terms of the frame of reference of an observer some distance from Earth, the laboratory frame of reference would be moving with Earth as it rotates on its axis and as it revolves about the Sun. A simple form of motion in the laboratory frame of reference would appear to be a much more complicated motion in the frame of reference of the distant observer. *See* FRAME OF REFERENCE.

Motion means continuous change of position of an object with respect to an observer. To another observer in a different frame of reference the object may not be moving at all, or it may be moving in an entirely different manner. The motions of the planets were found in ancient times to appear quite complicated in the laboratory frame of reference of an observer on Earth. By transferring to the frame of reference of an imaginary observer on the Sun, Johannes Kepler showed that the relative motion of the planets could be simply described in terms of elliptical orbits. The validity of one description is no greater than the other, but the latter description is far more convenient. *See* PLANET.

Relative velocity. That motion is relative to an observer must have been implicit in the earliest ideas of motion. In the mechanics of Galileo and Isaac Newton these ideas became clarified, and methods were developed for finding the relative velocity of two bodies, each moving with a different velocity. If the velocity of one body is $\mathbf{v}_1$, represented in **Fig. 1** by the magnitude and direction of the vector $\mathbf{v}_1$, and if a second body has a velocity $\mathbf{v}_2$ represented by the vector $\mathbf{v}_2$, then $\mathbf{v}_3$ is the vector to be added to $\mathbf{v}_1$ to make the sum equal to $\mathbf{v}_2$, and, consequently, the vector $\mathbf{v}_3$ is equal to the vector difference $\mathbf{v}_2 - \mathbf{v}_1$. Therefore, the relative velocity of the second body with respect to the first is that velocity represented in magnitude and direction by the vector $\mathbf{v}_3$, all vectors being drawn to a suitable scale.

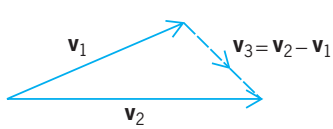

Fig. 1. Relative velocity of two bodies moving at an angle with respect to one another.

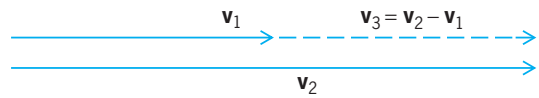

Fig. 2. Relative velocity of two bodies moving in the same direction; that is, velocities are parallel.

Fig. 3. Relative velocity of two bodies moving in opposite directions; that is, velocities are antiparallel.

In the simplest case the velocities $\mathbf{v}_1$ and $\mathbf{v}_2$ are parallel (**Fig. 2**). The relative velocity of the second body with respect to the first is again $\mathbf{v}_3$, and its magnitude is the difference in the numerical values of $\mathbf{v}_2$ and $\mathbf{v}_1$. When the two velocities are antiparallel (in opposite directions, **Fig. 3**), one is negative with respect to the other, and the magnitude of the relative velocity of $\mathbf{v}_2$ with respect to $\mathbf{v}_1$ is again the vector difference $\mathbf{v}_3$, which is now the numerical sum of the two. To obtain the relative velocity of $\mathbf{v}_1$ with respect to $\mathbf{v}_2$, the arrowhead on $\mathbf{v}_3$ should be reversed. For example, when two automobiles move with velocities of 40 and 60 mi/h (20 and 30 m/s) respectively in the frame of reference of an observer stationed by the roadside, the relative velocity of the two is 20 mi/h (10 m/s). If the automobiles were moving in opposite directions, the relative velocity would be 100 mi/h (50 m/s).

The flight of an airplane illustrates the principles of relative motion. The common statement that the speed of an airplane is, for instance, 300 mi/h (150 m/s) is essentially a meaningless statement since the speed is not designated with any particular frame of reference in mind and the listener must make an assumption that it is perhaps the ground speed in still air. For the airplane to sustain itself with normal lift, it is the airspeed or speed with respect to the air that is important. In a head wind of 100 mi/h (50 m/s) the airplane would have a ground speed of 200 mi/h (100 m/s). In a similar tail wind it would have a ground speed of 400 mi/h (200 m/s). If the wind velocity is at an angle to the direction of flight, the relative velocities must be considered. The velocity with respect to the ground would be the vector sum of the velocity of the plane with respect to the air and the velocity of the wind with respect to the ground. *See* FLIGHT CHARACTERISTICS.

Relative acceleration. Acceleration, like velocity, is relative to the observer's frame of reference. An automobile starting from rest is accelerated with respect to the Earth, but the driver of the car does not see himself or herself or the car accelerated forward. The driver sees objects at rest with respect to the roadway accelerated backward with respect to himself or herself and the car. Persons by the roadside see the driver and the car accelerated forward. In the driver's frame of reference, the car is at rest.

Acceleration is a vector quantity involving both magnitude and direction, just as velocity is. Just as the velocities of two objects may be represented by vectors and their relative velocity obtained by subtracting one vector from the other, so may the accelerations of two bodies be represented by vectors and the relative acceleration of one with respect to the other obtained by taking the vector difference. *See* ACCELERATION.

Relativity. Since, according to Einstein's theory of relativity, the velocity of light is the limiting velocity that any physical object can attain, the addition or subtraction of very large velocities cannot be accomplished by galilean-newtonian methods, and the rules to be followed are derived from relativistic theory. *See* MOTION; RELATIVITY. Rogers D. Rusk

Relativistic electrodynamics

The study of the interaction between electrically charged particles and electromagnetic fields when the velocities involved are comparable to that of light.

Development. A group of charged particles in motion can be represented by a distribution in charge and distribution in current. During the latter part of the eighteenth century and the early part of the nineteenth century, experiments by C. A. Coulomb, M. Faraday, A. M. Ampère, and others showed that electric and magnetic fields are produced by charge and current distributions. These fields, in turn, act on other charges and currents. The interaction between charges and currents on the one hand and electric and magnetic fields on the other is the topic of study of electrodynamics. This field of study was established as a quantitative and self-contained subject in 1864 when J. C. Maxwell formulated his equations for the electromagnetic field. Maxwell conjectured that a time-varying electric field is equivalent to an electric current in its effect of producing a magnetic field, and named it the displacement current. The inclusion of this displacement current enabled Maxwell to combine all the previously established laws of electromagnetism into a coherent whole in his equations. *See* CLASSICAL FIELD THEORY; DISPLACEMENT CURRENT; MAXWELL'S EQUATIONS.

With the inclusion of the displacement current, the Maxwell equations are relativistically covariant, meaning that they are valid for all velocities, even those approaching the velocity of light. However, the implications of the covariance of the equations were not fully appreciated until A. Einstein formulated the special theory of relativity in 1905. Relativistic electrodynamics was then rapidly developed into a powerful and precise field of physics. It describes and predicts all macroscopic electrodynamic phenomena to the minutest detail and with perfect accuracy, and now forms the foundation on which the entire electrical industry is based. However, its limitations soon became evident when attempts were made to apply it to atomic phenomena: Straightforwardly applied, relativistic electrodynamics failed to explain many of these phenomena, and its predictions frequently disagreed with experimental observations. For these microscopic phenomena, quantum electrodynamics (QED) was developed in the 1930s to replace classical relativistic electrodynamics. In 1967 quantum electrodynamics was further unified by S. Weinberg and A. Salam with the theory of weak interactions to form the electroweak theory. *See* QUANTUM ELECTRODYNAMICS; RELATIVISTIC MECHANICS; RELATIVITY; WEAK NUCLEAR INTERACTIONS.

Types of problems. Electrodynamic problems generally fall into one of two categories:

1. Finding the electromagnetic field produced by prescribed charge and current distributions. For example, one may want to determine the electromagnetic field produced or radiated by a given oscillatory electric current in a transmitting antenna, or the field radiated by an accelerating electron.

2. Finding the effect of a predetermined electromagnetic field on the motion of charges and currents. This is the inverse problem corresponding to that of the receiving antenna or of the motion of charged particles in an accelerator.

All other electrodynamic problems are combinations or iterations of these two basic types. For instance, the scattering of light (electromagnetic radiation) by a charged particle is composed of, first, the incident light shaking the charge and, second, the subsequent emission of the scattered light by the shaken charge. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

Electromagnetic field. The electromagnetic field produced by a given charge-current distribution is determined by Maxwell's equations. When written in separate space-time form, these are Eqs. (1), where

$$\begin{aligned} \vec{\nabla}\cdot\vec{E} &= 4\pi\rho & \vec{\nabla}\cdot\vec{B} &= 0 \\ \vec{\nabla}\times\vec{E} &= -\frac{1}{c}\frac{\partial\vec{B}}{\partial t} & \vec{\nabla}\times\vec{B} &= \frac{1}{c}\frac{\partial\vec{E}}{\partial t}+\frac{4\pi}{c}\vec{J} \end{aligned} \tag{1}$$

$\vec{E}$ and $\vec{B}$ are the electric and the magnetic fields; ρ and $\vec{J}$ are the charge and the current density distributions; $\vec{\nabla}$, given by Eq. (2), is the space differential

$$\vec{\nabla} = \vec{i}\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z} \tag{2}$$

operator; and $c \cong 3 \times 10^8$ m/s is the speed of light. As customary in this subject, all quantities are expressed in gaussian units. *See* ELECTRICAL UNITS AND STANDARDS; LIGHT.

The fact that these equations are relativistically covariant is demonstrated by showing that they can be written in a four-dimensional (three space dimensions and one time dimension) tensor form, because tensor equations have the same identical form in all frames of reference. For this, the coordinate 4-vector with imaginary time is used, namely $x_\mu = (x, y, z, ict)$ or (x_1, x_2, x_3, x_4). The field 4-tensor is defined by Eq. (3), and the charge-current density 4-vector is

given by Eq. (4), meaning that, for example, F_{12} is

$$F_{\mu\nu} = \begin{pmatrix} 0 & B_z & -B_y & -iE_x \\ -B_z & 0 & B_x & -iE_y \\ B_y & -B_x & 0 & -iE_z \\ iE_x & iE_y & iE_z & 0 \end{pmatrix} \tag{3}$$

$$J_\mu = (\vec{J}, ic\rho) \tag{4}$$

B_z, J_1 is J_x, and J_4 is $ic\rho$. Maxwell's equations (1) can now be written as Eqs. (5), as may be verified by

$$\begin{gathered} \frac{\partial F_{\mu\nu}}{\partial x_\nu} = \frac{4\pi}{c} J_\mu \\ \frac{\partial F_{\mu\nu}}{\partial x_\lambda} + \frac{\partial F_{\lambda\mu}}{\partial x_\nu} + \frac{\partial F_{\nu\lambda}}{\partial x_\mu} = 0 \end{gathered} \tag{5}$$

taking special values of the indices. A sum from 1 to 4 is understood to be made on indices repeated in a term, as the index ν in the first of Eqs. (5). *See* CALCULUS OF VECTORS.

If the antisymmetric field tensor is expressed in terms of the 4-potential $A_\mu = (\vec{A}, i\phi)$ by Eq. (6),

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x_\mu} - \frac{\partial A_\mu}{\partial x_\nu} \tag{6}$$

the second of Eqs. (5) is identically satisfied. Equation (6), however, does not define A_μ uniquely, and in addition the Lorentz condition, Eq. (7), can

$$\frac{\partial A_\nu}{\partial x_\nu} = 0 \tag{7}$$

be imposed. The first of Eqs. (5), then, becomes Eq. (8). It can easily be verified that Eq. (8) and

$$\frac{\partial^2 A_\mu}{\partial x_\nu^2} = \left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) A_\mu = -\frac{4\pi}{c} J_\mu \tag{8}$$

the Lorentz condition are indeed consistent with the continuity equation for the 4-current, Eq. (9). Equation (8) shows that the potential wave $A_\mu(\vec{x}, t)$ generated by the current source $J_\mu(\vec{x}', t')$ travels with a phase velocity equal to the velocity of light. This led Maxwell to infer that light is an electromagnetic wave. Furthermore, electromagnetic waves span a very broad frequency spectrum from a few hertz for alternating-current waves, to kilohertz and megahertz radio waves, 10^{10} Hz for microwaves, 10^{15} Hz for visible light, 10^{20} Hz for x-rays, to 10^{23} Hz for gamma rays and beyond. *See* ELECTROMAGNETIC RADIATION; EQUATION OF CONTINUITY; POTENTIALS.

$$\frac{\partial J_\mu}{\partial x_\mu} = 0 \tag{9}$$

Electromagnetic waves always propagate at the speed of light. Only the motion of the sources which are massive materials can be slow compared to the speed of light. Thus, contrary to the existence of the newtonian mechanics as the low-velocity approximation of the relativistic mechanics, it is impossible to derive a nonrelativistic approximation for Maxwell's equations. Indeed, it is through the investigation and understanding of the transformation properties of the Maxwell equations that Einstein was led to the discovery of the special relativity theory.

Transformation. The covariance of the theory states that the electromagnetic field must transform by the Lorentz transformation from one reference frame to another. This transformation property is illuminating. For example, a pure electric field at rest will appear as an electric field plus a magnetic field when observed in a moving frame. This can be shown by writing the Lorentz transformation from one frame, K, to another, K', moving relative to K at velocity $c\beta$ along the z axis, as Eq. (10), where the coefficients $a_{\mu\nu}$ are given by Eq. (11), and γ is defined by Eq. (12). Transforming $F_{\mu\nu}$ from K to K' by the tensor transformation rule $F'_{\mu\nu} = a_{\mu\lambda} a_{\nu\rho} F_{\lambda\rho}$ yields Eqs. (13). This shows that the electric and the magnetic fields are just different space-time projections of a more general object, electromagnetic field. *See* LORENTZ TRANSFORMATIONS; SPACE-TIME.

$$x'_\mu = a_{\mu\nu} x_\nu \tag{10}$$

$$a_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & -i\gamma\beta \\ 0 & 0 & -i\gamma\beta & \gamma \end{pmatrix} \tag{11}$$

$$\gamma \equiv (1 - \beta^2)^{-1/2} \tag{12}$$

$$\begin{aligned} E'_x &= \gamma(E_x - \beta B_y) & B'_x &= \gamma(B_x + \beta E_y) \\ E'_y &= \gamma(E_y + \beta B_x) & B'_y &= \gamma(B_y - \beta E_x) \\ E'_z &= E_z & B'_z &= B_z \end{aligned} \tag{13}$$

Energy and momentum. The electromagnetic field contains both momentum and energy, which is why electromagnetic waves can be used to transmit messages or power. The momentum and energy content of an electromagnetic field is given by the energy-momentum 4-tensor of Eq. (14), where the double overarrow signifies a 3-tensor, $\delta_{\mu\nu} = 1$ when $\mu = \nu$ and $\delta_{\mu\nu} = 0$ when $\mu \neq \nu$. This is expressed in separate space and time components by Eqs. (15);

$$T_{\mu\nu} = \frac{1}{4\pi}\left[F_{\mu\nu}F_{\lambda\nu} + \frac{1}{4}\delta_{\mu\nu}F_{\lambda\sigma}F_{\lambda\sigma}\right] = \left(\begin{array}{c|c} \vec{\vec{T}} & -ic\vec{g} \\ \hline -ic\vec{g} & u \end{array}\right) \tag{14}$$

$$\begin{aligned} \vec{\vec{T}} &= \frac{1}{4\pi}\left[\vec{E}\vec{E} + \vec{B}\vec{B} - \frac{1}{2}\vec{\vec{I}}(E^2 + B^2)\right] \\ &= \text{energy} - \text{momentum 3} - \text{tensor} \\ \vec{g} &\frac{1}{4\pi c}\vec{E} \times \vec{B} = \frac{1}{c^2}\vec{S} = \text{momentum density} \\ u &= \frac{1}{8\pi}(E^2 + B^2) = \text{energy density} \end{aligned} \tag{15}$$

here $\vec{S}$ is called the Poynting vector. Conservation of energy and momentum is expressed by Eq. (16),

$$\frac{\partial T_{\mu\nu}}{\partial x_\nu} = \frac{1}{c} F_{\mu\nu} J_\nu \tag{16}$$

which can be derived from Eqs. (5) and (14). The expression on the right side is the 4-force density exerted by the charge-current density J_ν on the field. Hertz carried out an experiment in 1887 which succeeded in demonstrating the transmission of energy across space on a radio wave produced by an oscillatory spark discharge, consistent with this aspect of Maxwell's theory. *See* CONSERVATION OF ENERGY; CONSERVATION OF MOMENTUM; ENERGY; POYNTING'S VECTOR.

Solution of wave equation. To obtain the electromagnetic field, it is necessary to solve Eq. (1) for $\vec{E}$ and $\vec{B}$ or, equivalently, Eq. (8) for A_μ and the solution of Eq. (8) requires the Green's function of the equation. This is given by Eq. (17), where R is given by Eq. (18),

$$G(\vec{x}, t; \vec{x}, t') = \frac{\delta\left(t' + \dfrac{R}{c} - t\right)}{R} \tag{17}$$

$$\begin{aligned} R &= \left|\vec{x} - \vec{x}'\right| \\ &= \text{distance between source point } \vec{x}' \\ &\quad \text{and field point } \vec{x} \end{aligned} \tag{18}$$

and the Dirac δ-function specifies that at the source point $\vec{x}'$ the only values that matter are at the retarded time t' given by Eq. (19). Hence the general solution

$$t' = t - \frac{R}{c} \tag{19}$$

for A_μ is given by Eq. (20), where the brackets on the

$$A_\mu(\vec{x}, t) = \frac{1}{c} \int d^3x' \left[\frac{J_\mu(\vec{x}', t')}{R}\right]_{\text{ret}} \tag{20}$$

right emphasize that t' is given by Eq. (19). The retarded time t' assures causal behavior, such that the effect of the source propagates only with velocity c. This now gives the general and complete solution of all problems of finding the electromagnetic field produced by prescribed charge and current distributions. With a given charge-current distribution J_μ, Eq. (20) yields the 4-potential, A_μ which in turn gives the electromagnetic field $\vec{E}$ and $\vec{B}$ through Eqs. (6) and (3). Other interesting physical quantities such as the momentum and energy distributions are obtained by substituting $\vec{E}$ and $\vec{B}$ in Eq. (15). *See* GREEN'S FUNCTION; WAVE EQUATION.

Radiation from general localized source. The most interesting special case of the electromagnetic field is the radiation, or the faraway electromagnetic field, generated by a localized oscillating source. If only a single frequency component, ω, of the source is chosen [that is, if $J\mu$ has the form $J_\mu(\vec{x}, t) = J_\mu(\vec{x})e^{-iwt}$], and if it is assumed that the wavelength c/ω is much larger than the size of the source but much smaller than the distance to the field point, then the expanded form obtained for the 3-vector potential $\vec{A}$ is given by Eq. (21), with $\vec{A}_m$ given by Eq. (22), where

$$\vec{A}(\vec{x}) = \sum_{m=0}^{\infty} \vec{A}_m(\vec{x}) \tag{21}$$

$$\vec{A}_m(\vec{x}) = \frac{e^{ikr}}{cr} \frac{(-ik)^m}{m!} \int \vec{J}(\vec{x}')(\vec{n} \cdot \vec{x}')^m \, d^3x' \tag{22}$$

$r \equiv |\vec{x}|$, $\vec{n} = \vec{x}/r$ and $k \equiv \omega/c$. The integral contains the electric $2(m+1)$-pole moment and the magnetic $2m$-pole moment of the source. Thus, A_0 gives the electric dipole radiation. This is the case of the simple dipole antenna. The term $\vec{A}_1$ gives the electric quadrupole and the magnetic dipole radiation. Its radiated power is lower than that of $\vec{A}_0$ by an order of (source size)^2/(wavelength)^2. *See* ANTENNA (ELECTROMAGNETISM); ELECTROMAGNETIC WAVE TRANSMISSION; MULTIPOLE RADIATION.

This part of relativistic electrodynamics clearly plays a central role in telecommunications, and in addition forms the basis for the development of radar, waveguides, radio telescopes, and so forth. *See* RADAR; RADIO; RADIO TELESCOPE; WAVEGUIDE.

Radiation from moving charge. In the case of a single charge e moving with given velocity $c\vec{\beta}$ along the given trajectory $\vec{x}_e(t)$, the 4-current density is given by Eq. (23), where $\beta_v = (\vec{\beta}, i)$. With this J_v, Eq. (20)

$$J_\nu(\vec{x}, t) = ec\beta_\nu \delta[\vec{x} - \vec{x}_e(t)] \tag{23}$$

gives, in separate space-time form, Eqs. (24), where κ is defined by Eq. (25). Here $R(t)$ denotes

$$\begin{aligned} \vec{A}(\vec{x}, t) &= e\left[\frac{\vec{\beta}}{\kappa R}\right]_{\text{ret}} \\ \phi(\vec{x}, t) &= e\left[\frac{1}{\kappa R}\right]_{\text{ret}} \end{aligned} \tag{24}$$

$$\kappa \equiv \frac{d}{dt}\left[t + \frac{R(t)}{c}\right] \tag{25}$$

$|x - x_e(t)|$ and so now is a function of time. The significance of the square brackets on the right is that the functions of t are to be evaluated at the retarded time t' found from Eq. (26). These are called

$$t' = t - \frac{1}{c} R(t') \tag{26}$$

Liénard-Wiechert potentials and give in the radiation zone (far away, where the fields depend on $1/R$) the fields of Eqs. (27), showing that a charged particle

$$\begin{aligned} \vec{E}(\vec{x}, t) &= \frac{e}{c}\left[\frac{\vec{n}}{\kappa^3 R} \times \{(\vec{n} - \vec{\beta}) \times \dot{\vec{\beta}}\}\right]_{\text{ret}} \\ \vec{B} &= [\vec{n}]_{\text{ret}} \times \vec{E} \end{aligned} \tag{27}$$

radiates when it is accelerating ($\dot{\vec{\beta}} \neq 0$). At this point $\vec{n}$ indicates a unit vector in the direction $\vec{x} - \vec{x}_e(t)$.

For a charged particle moving in a circle with radius ρ, the resulting radiative power is given by Eq. (28), and the energy loss per revolution is given

by Eq. (29). For high-energy electrons ($\beta \cong 1$) this has the numerical value given by Eq. (30).

$$P = \frac{2}{3}\frac{e^2 c}{\rho^2}\beta^4\gamma^4 \tag{28}$$

$$\Delta E = \frac{2\pi\rho}{c\beta}P = \frac{4\pi}{3}\frac{e^2}{\rho}\beta^3\gamma^4 \tag{29}$$

$$\Delta E\ (\text{keV}) = 88.5\frac{[E\ (\text{GeV})]^4}{\rho\ (\text{meter})} \tag{30}$$

In a circular accelerator, for example, a synchrotron, the electrons are deflected by dipole magnets to travel in a circular orbit. While traveling in this orbit, they emit the radiation given by Eq. (28). This synchrotron radiation was first observed and identified in 1948. At the time, it was considered a nuisance because it causes energy loss while one is trying to accelerate the electrons. Almost 20 years later it was realized that, because of its very high brightness, very broad spectrum (10^{-3} to 10^5 eV, or from infrared through visible and ultraviolet to soft and hard x-rays), very small source size, and special polarization, synchrotron radiation makes a very useful research tool for studying all types of atomic, molecular, solid, and surface structure problems. There are now about 20 electron storage rings with beam energies ranging from several hundred MeV to over 10 GeV operated for synchrotron radiation research. In addition, another 20 such facilities are being planned or in construction. *See* SYNCHROTRON RADIATION.

Charged particle motion. The problem of finding the motion of a charged particle in a specified electromagnetic field is more a mechanical problem, and hence has been largely formulated and solved in relativistic mechanics. All that is required is the "force." The relativistic equation of motion of a charged particle acted on by a given electromagnetic field is Eq. (31), where m is the rest mass of the particle;

$$\frac{dp_\mu}{d\tau} = \frac{e}{mc}F_{\mu\nu}p_\nu \tag{31}$$

$d\tau = dt/\gamma$ is the proper time differential; p_μ, given by Eq. (32), is the 4-momentum; and the expression

$$p_\mu = m\gamma c\beta_\mu = \left(\vec{p}, \frac{iE}{c}\right) \tag{32}$$

on the right side of Eq. (31) is the 4-force exerted on the particle by the field. Equation (32) can be derived from the relativistic lagrangian of Eq. (33).

$$L = \frac{1}{\gamma}\left(-mc^2 + \frac{e}{mc}p_\mu A_\mu\right) \tag{33}$$

See LAGRANGE'S EQUATIONS.

In the separated space-time form, the equation of motion becomes Eqs. (34) and (35). The expression

$$\frac{d\vec{p}}{dt} = e(\vec{E} + \vec{\beta} \times \vec{B}) \tag{34}$$

$$\frac{dE}{dt} = ec\vec{\beta} \cdot \vec{E} \tag{35}$$

on the right side of Eq. (34), the momentum equation, is the force exerted on the particle by the field and is known as the Lorentz force. The right-side expression of Eq. (35), the energy equation, is the rate of energy imparted to the particle by the field. With a prescribed electromagnetic field, solving Eqs. (34) and (35) for the particle motion is straightforward and gives precise particle orbits at velocities arbitrarily close to the velocity of light. In high-energy particle acclerators, as used for research in particle physics, electron velocities within 10^{-10} of the velocity of light have been attained.

In addition to the application to particle accelerators, the solution of this second category of relativistic electrodynamic problems is indispensable to the design of magnetic confinement fusion machines such as the tokamak and the mirror machine, magnetohydrodynamic devices, and other plasma devices. The theory is also useful for describing the aurora borealis, the Van Allen belt, and other phenomena involving interplanetary and interstellar particle motions. *See* AURORA; MAGNETOSPHERE; NUCLEAR FUSION; PARTICLE ACCELERATOR; PLASMA (PHYSICS); SOLAR WIND; VAN ALLEN RADIATION.

Self-consistent charge-field theory. It is very tempting to try combining the formulations of the two categories described above to form a consistent picture of charged particles interacting among themselves through their own electromagnetic field. For a single electron this could give a total description of its electromagnetic behavior (the theory of the electron). But while the two parts of relativistic electrodynamics give quantitatively correct predictions to very high degrees of accuracy for macroscopic phenomena, the combined self-consistent theory when applied to microscopic phenomena runs into a variety of insurmountable mathematical and physical difficulties. For example, the electron orbiting the proton in a hydrogen atom does not radiate despite its acceleration. The electromagnetic energy of the electron computed from the theory is much greater than the energy corresponding to its mass. The coupled equations also give an unphysical self-accelerating solution according to which an electron could spontaneously accelerate from rest to infinite velocity. All these difficulties indicate that there is some fundamental defect in the theory making it invalid when applied to microscopic phenomena. This was resolved by the advent of quantum (relativistic) electrodynamics. The additional ingredient in quantum electrodynamics is the appreciation of the minimal effect of the action of making a measurement which becomes significant or even dominant for microscopic phenomena. *See* QUANTUM MECHANICS. Lee C. Teng

Bibliography. D. F. Griffiths, *Introduction to Electrodynamics*, 3d ed., 1998; J. D. Jackson, *Classical Electrodynamics*; 3d ed., 1998; E. Konopinski, *Electromagnetic Fields and Relativistic Particles*, 1981; L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th ed., 1976; H. C. Ohanian,

Classical Electrodynamics, 1988; P. B. Visscher, *Fields and Electrodynamics*, 1988.

Relativistic heavy-ion collisions

Collisions between heavy atomic nuclei at relative velocities close to the speed of light. These high-energy nuclear collisions are usually divided into two different domains, relativistic and ultrarelativistic collisions, depending on whether the kinetic energy per nucleon (the generic name for protons and neutrons) is either close to the rest mass of the nucleon (relativistic collisions) or much larger than the nucleon rest mass (ultrarelativistic collisions).

Nuclear and hadronic matter. By utilizing high-energy nuclear collisions, it is possible to study nuclear matter under conditions of very high temperatures and densities. The most common form of nuclear matter, at least under terrestrial conditions, is found in the atomic nucleus, which consists of protons and neutrons bound together by the strong nuclear force. If nuclear matter is heated up to temperatures comparable to the rest mass of the pion, it becomes a mixture of nucleons, pions, and various other particles, collectively denoted hadrons. Under these circumstances, nuclear matter is referred to as hadronic matter. *See* HADRON; NEUTRON; NEUTRON STAR; NUCLEAR STRUCTURE; PROTON.

Quark-gluon plasma. According to the quantum chromodynamics (QCD) theory, all hadrons are bound states of a set of more fundamental entities called quarks. The quarks are confined within the hadrons by the exchange of gluons. Quantum chromodynamics calculations using the most powerful computers available show that if hadronic matter is further heated or compressed to very high densities it will undergo a phase transition into a new phase of matter, called the quark-gluon plasma. In this phase the hadrons will lose their identity, and the quarks and gluons will be deconfined within volumes much larger than the typical hadron volume of 0.1–0.5 cubic femtometer. Quantum chromodynamics calculations indicate that the phase transition will occur at a critical density around 5–10 times the normal nuclear matter density of approximately 0.2 nucleon/fm^3, or at a critical temperature around 150 MeV. These conditions should have been fulfilled during the first few microseconds of the big bang, and might be attained in the core of high-mass neutron stars. Strong interest in ultrarelativistic heavy-ion collisions is caused by the realization that these collisions seem to be the only way to create a quark-gluon plasma under laboratory conditions. *See* BIG BANG THEORY; GLUONS; QUANTUM CHROMODYNAMICS; QUARK-GLUON PLASMA; QUARKS.

Space-time evolution. When two nuclei collide at high energies, some of the nucleons in each nucleus, called spectators, will continue their motion unaffected, while other nucleons, called participants, will strike one or several nucleons in the other nucleus (see **illus.**). In the overlap volume a hot and dense fireball will develop. If the temperature or density of the fireball becomes larger than the critical values, a quark-gluon plasma will be created with an estimated lifetime of 1–5 $\times$ 10^{-23} s. The fireball will start to expand and cool, and the quarks in the plasma will eventually be reconfined into a large number of hadrons (hadronization). After further expansion the hadrons will cease interacting with each other (freeze out) and leave the collision zone without further mutual interactions.

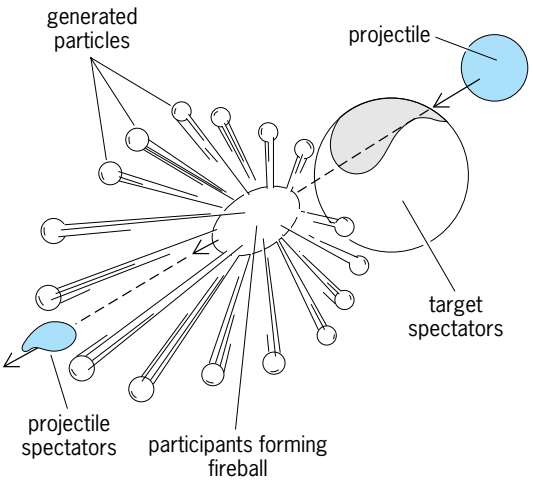

High-energy nuclear collision. The participant nucleons in the projectile and target transiently form a hot and dense fireball, while the spectator nucleons remain relatively undisturbed. The fireball will subsequently expand and emit a large number of particles, primarily pions and kaons.

Global variables. The probability of creating the quark-gluon plasma is largest in central collisions, where the collision is head-on or close to it. It is therefore important to distinguish these central collisions from peripheral collisions. This characterization is best done by detecting global variables like the transverse energy or the charged-particle multiplicity. The transverse energy is a measure of how much energy has been redirected from the initial beam direction into the directions transverse to the beam. Central collisions will generate the largest values of transverse energy. Values as large as 50% of the total available center-of-mass energy have been observed in the most central collisions of 200 GeV/nucleon sulfur nuclei with lead nuclei. Such large values of transverse energy show that the nuclei are not transparent to each other but that the participating nucleons are being significantly slowed down. This high nuclear stopping power implies that a large fraction of the participants' initial kinetic energy is transformed into the production of many new particles at high temperatures. Based on transverse energy measurements, the highest fireball energy density attained is estimated to be approximately 10 times the density of normal nuclear matter.

The charged-particle multiplicity, which is the number of charged particles emerging from the nuclear collision, can also be used as a centrality indicator. The charged-particle multiplicity will, in addition, carry information on the amount of entropy generated during the collision. Experimentally, the

transverse energy and the charged-particle multiplicity have been shown to be very strongly correlated.

Theoretical models based on the assumption that collisions between two nuclei can be viewed as a superposition of a large number of binary collisions between the individual nucleons in the nuclei have proven to be very successful in estimating the observed values of the transverse energy and the charged-particle multiplicities. String models are in especially wide use. In these models the kinetic energy of two colliding nucleons is transformed into a string-shaped gluon field stretching between the quarks in the nucleons as they move past each other. The string will then subsequently be stretched and snap in several pieces, with a quark and an antiquark situated at opposite ends of each string fragment. Each string fragment will eventually evolve into a hadron.

Signatures of the quark-gluon plasma. In the search for the quark-gluon plasma, a fundamental problem is that even if the plasma is created in the early phases of the collisions, the subsequent hadronization and scattering of the hadrons before freeze-out might mask any traces of the plasma. In order to circumvent this problem, many plasma signatures have been proposed. The most promising signature is based on the conjecture that certain particles consisting of a charm quark and an anticharm quark (such as the J/ψ and ψ' particles) will be dissolved in the plasma and will not be regenerated in the later phases of the collision. This suppression in the production of the J/ψ particle has been observed. *See* J/PSI PARTICLE.

Another important property of the quark-gluon plasma in chemical equilibrium will be the high strange and antistrange quark content. Most of these strange quarks will be bound in strange particles (kaons, lambda, hyperons, and so forth) during the later stages of the collision. An unusually high content of these strange particles could therefore also be a signature for the creation of the quark-gluon plasma. Several experiments have actually detected unusually high yields of strange particles. *See* ELEMENTARY PARTICLE; NUCLEAR REACTION; STRANGE PARTICLES.

Soren P. Sorensen

Bibliography. G. F. Bertsch, Searching for the quark-gluon plasma, *Science*, 265:480–481, 1994; B. Schwarschild, Nuclear matter in extremis sought with ultrarelativistic heavy ion beams, *Phys. Today*, 41(3):17–21, March 1988.

Relativistic mechanics

An extension of newtonian mechanics conforming to the principles of special relativity. Suitably generalized, its results are also incorporated in the general theory. The energy-momentum conservation laws of relativistic mechanics enter in the development of relativistic quantum mechanics, and they find important application in high-energy physics.

Of Newton's laws of motion, it is the third law whose relativistic formulation presents particular challenge. Its application is simplest in the idealized context of strictly continguous action, whether in continuous media or in discrete particle systems—the only case discussed here. Without this assumption, one must cope with the problem of retarded action at a distance, on which some interesting exploratory work has been done. *See* NEWTON'S LAWS OF MOTION; RELATIVITY.

Equations of motion of a particle. The first correct relativistic equation of motion was that of a charged particle in an electromagnetic field developed by Max Planck in 1906. The restriction on the applied force was later removed by Hermann Minkowski, who was also the first to use the four-dimensional space-time formalism systematically. This formalism greatly facilitates a concise discussion of the subject, and will be employed here from the start. *See* RELATIVISTIC ELECTRODYNAMICS; SPACE-TIME.

The relativistic generalization of Newton's second law of motion fora particle, Eq. (1), must satisfy two

$$m\mathbf{a} = \mathbf{f} \tag{1}$$

conditions: it must include the newtonian equation as a limiting case for small particle speeds; it must be covariant (that is, retain its form) under Lorentz transformations, which is automatically assured when it is expressed in four-vector form. As to Newton's first law, its relativistic extension is immediate, because it serves only to define the family of inertial frames relative to which the second law is enunciated, and that family plays an identical role in the special theory of relativity. *See* LORENTZ TRANSFORMATIONS.

To arrive at the relativistic generalization of Eq. (1), it is first rewritten, using the particle momentum, $\mathbf{p} = m\, d\mathbf{x}/dt$, as $d\mathbf{p}/dt = \mathbf{f}$ (which is actually a closer representation of the law as stated in Newton's *Principia*). One then replaces dx, dy, dz by dx^α ($\alpha = 0, 1, 2, 3$; $x^0 \equiv ct$), m by the rest mass m_0 (that is, the mass of the particle as measured in the reference frame in which it is instantaneously at rest), and dt by its four-scalar extension $d\tau = dt\sqrt{1-(v^2/c^2)}$. Retaining, as is customary, the symbol p for the generalization of $\mathbf{p}$, so that $p^\alpha = m_0 dx^\alpha/d\tau$ ($\alpha = 0,1,2,3$), the resulting four-vector equation assumes the form of Eq. (2), which may be called Minkowski's equation of motion.

$$\frac{dp^\alpha}{d\tau} = F^\alpha \qquad (\alpha = 0, 1, 2, 3) \tag{2}$$

The significance of Eq. (2) becomes apparent when the time t is reintroduced, resulting in Eqs. (3) and (4), where v is the instantaneous speed

$$\frac{dp^\alpha}{dt} = f^\alpha$$

$$p^\alpha = m\frac{dx^\alpha}{dt} \equiv mv^\alpha \qquad (\alpha = 0, 1, 2, 3) \tag{3}$$

$$m = \gamma m_0$$

$$f^\alpha = \frac{F^\alpha}{\gamma} \quad (\alpha = 0, 1, 2, 3) \tag{4}$$

$$\gamma \equiv 1/\sqrt{1-(v^2/c^2)}$$

of the particle. Since $\gamma = 1 + v^2/2c^2 \ldots$ it follows that

when v^2/c^2 can be neglected, Eqs. (3) for $\alpha = 1, 2, 3$ coincide indeed with the corresponding newtonian equations. Moreover, it is F^α and not f^α that are the components of a four-vector. This four-vector generalization of the newtonian force is known as the Minkowski force.

The meaning of Eqs. (3) for $\alpha = 0$ can be established with the aid of the four-scalar equation (5)

$$(p^0)^2 - \mathbf{p}^2 = m_0^2 c^2 \tag{5}$$

derivable from Eqs. (4) and (3), where $\mathbf{p} \equiv (p^1, p^2, p^3)$ (not to be confused with the previous $\mathbf{p}$). Differentiating this equation with respect to t and using Eqs. (3), one finds: $d(mc^2)/dt = cf^0 = \mathbf{f} \cdot \mathbf{v}$. Hence the quantitydefined by Eq. (6) is an energy, and the

$$E = cp^0 = mc^2 \tag{6}$$

four-vector p, the four-momentum, is thus also commonly called the energy-momentum vector. By Eqs. (5) and (6), it follows that E can also be expressed in terms of the "relativistic" $\mathbf{p}$ as in Eq. (7). In addi-

$$E = \sqrt{m_0^2 c^4 + c^2 \mathbf{p}^2} \tag{7}$$

tion to E, which is usually called the total energy, it is useful to introduce a relativistic kinetic energy (vanishing with v): $m_0 c^2 (\gamma - 1)$, where γ is defined in Eqs. (4). The expansion of γ shows that when v^2/c^2 can be neglected, this expression reduces indeed to the newtonian kinetic energy $mv^2/2$.

When $m_0 \neq 0$, Eqs. (4) and (6) show that for $v = 0$ the energy E has the nonvanishing value $E_0 = m_0 c^2$. This rest energy is thus an intrinsic energy of a particle associated entirely with its rest mass. The universality of this inertia-of-energy result, as established principally by Albert Einstein, and extensively confirmed by experiment, constitutes one of the most far-reaching conclusions of the theory of relativity.

Another important conclusion from Eqs. (4), which accords with other deductions from the relativistic principles, is that a particle of nonvanishing rest mass may approach but can never attain the speed c. On the other hand, the speed of a particle of vanishing rest mass, such as a photon, can only be c; otherwise, contrary to theory and experience, the particle's energy and momentum would be zero, as is seen, for instance, by taking the limit $m_0 \to 0$ in Eqs. (3). Equation (7) shows, moreover, that the energy and momentum of such a particle are related by the simple equation $E = c|\mathbf{p}|$; and also, that when $m_0 \neq 0$ this relation holds as an ultrahigh-energy approximation ($m_0 c \ll |\mathbf{p}|$).

Special systems of particles. Under the restrictive assumption that the particles of the system are involved only in spatially and temporally localized interactions, it follows easily that Newton's third law can be formulated in a relativistically covariant fashion. Moreover, if the system is closed, that is, subject to no external action, then as in the case of newtonian mechanics the total energy and momentum of the system are conserved quantities: $\Sigma E =$ constant, $\Sigma \mathbf{p} =$ constant. Or, in four-vector notation Eq. (8) is satisfied.

$$P^\alpha \equiv \Sigma p^\alpha = \text{constant} \tag{8}$$

$$(\alpha = 0, 1, 2, 3;\ \text{summation over the particles})$$

It is possible to extend this result to allow for the (quantum) processes of the creation and annihilation of particles, so that the number of particles of the system need not be conserved, and the summation symbol in Eq. (8) refers to the particles existing at the instant under consideration. In particular, for a particle decay process or a localized binary collision of any type, Eq. (8) can be written more explicitly as Eq. (9), with $m = 1$ or 2, respectively. An early

$$\sum_{i=1}^{m} p_i^\alpha \text{ (initial)} = \sum_{j=1}^{n} p_j^\alpha \text{ (final)} \tag{9}$$

and historically very important application of this formula was made by Arthur Compton in 1922 to the collision of an x-ray photon and a practically free electron. The extended formula has been applied to numerous high-energy particle reactions, and it appears to represent one of the most firmly established laws in physics. *See* COMPTON EFFECT.

As in the treatment of low-energy collisions, so also in the treatment of high-energy particle reactions it is helpful to deal not only with the laboratory frame but also with the center-of-mass (or center-of-momentum) frame of the mechanical system under consideration. The latter is defined by the condition that the spatial part of the total four-momentum of the system vanishes. In addition to the usual advantages of the center-of-mass frame, it is, in a practical sense, especially useful in the case of very high-energy collisions. With the aid of the Lorentz transformations connecting the components of the total four-momentum P defined by Eq. (8) in the laboratory and center-of-mass frames, it can be shown that if instead of bombarding some type of particles, which are effectively stationary in the laboratory, by a beam of similar particles, one directs two such beams against each other so that for the latter system the laboratory serves as a center-of-mass frame, the energy released in the collision is substantially greater in the second case, the advantage increasing with the rise in particle energy. This has been utilized in positron-electron and proton-proton colliding-beam devices called storage rings. *See* PARTICLE ACCELERATOR.

The center-of-mass frame has also independent interest. If it is denoted by S_0, and S is an arbitrary inertial frame, it can be shown from Lorentz transformations of the total four-momentum that $\mathbf{P} = \gamma P^0_{(0)} \mathbf{v}/c = P^0 \mathbf{v}/c$, where $\mathbf{v}$ represents the velocity of S_0 relative to S, and P is defined in Eq. (8), P_0^α being the components of this four-vector in S_0. It follows that one can consider the quantity $M_0 \equiv P^0_{(0)}/c$ as the rest-mass of the system, and that $M \equiv P^0/c = \gamma M_0$, $\mathbf{P} = M\mathbf{v}$. The latter relations are identical with the appropriate relations in Eqs. (3) and (4) for a single particle of rest mass M_0 and velocity $\mathbf{v}$ relative to S. One can show, further, that $M_0 = \Sigma m_0 + \Sigma K_0/c^2$, the sums

being taken over the particles of the system, with m_0 and K_0 representing, respectively, the rest mass and the relativistic kinetic energy of a particle relative to S_0. The last term is an interesting manifestation of the inertia of energy. H. M. Schwartz

Bibliography. A. Einstein, *The Meaning of Relativity*, 5th ed., 1956; R. A. Mould, *Basic Relativity*, 1994; R. Resnick and D. Halliday, *Basic Concepts in Relativity*, 1992; W. Rindler, *Introduction to Special Relativity*, 2d. ed, 1991; H. M. Schwartz, *Introduction to Special Relativity*, 1968, reprint 1977; E. Tocaci, *Relativistic Mechanics*, *Time and Inertia*, 1984.

Relativistic quantum theory

The quantum theory of particles which is consistent with the special theory of relativity, and thus can describe particles moving arbitrarily close to the speed of light. It is now realized that the only satisfactory relativistic quantum theory is quantum field theory; the attempt to relativize the Schrödinger equation for the wave function of a single particle fails, as described below. However, with a change of interpretation, relativistic wave equations do correctly describe some aspects of the motions of particles in an electromagnetic field. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM FIELD THEORY; QUANTUM MECHANICS; RELATIVITY.

Invariance. The Schrödinger equation for the wave function $\psi(\mathbf{r},t)$ of a particle $[|\psi(\mathbf{r},t)|^2$ is the density of probability of finding the particle at point $\mathbf{r}$ at time $t]$ is Eq. (1), where E is the energy operator $i\hbar(\partial/\partial t)$, $\mathbf{p}$ is the momentum operator $-i\hbar\nabla$, $H(\mathbf{p},\mathbf{r})$ is the classical hamiltonian, and $\hbar$ is Planck's constant divided by 2π. For a nonrelativistic free particle, $H = \mathbf{p}^2/2m$. The naive way to relativize Eq. (1) would be to use the relativistic hamiltonian, Eq. (2).

$$E\psi = H(\mathbf{p},\mathbf{r})\psi \tag{1}$$

$$H = \sqrt{(mc^2)^2 + \mathbf{p}^2c^2} \tag{2}$$

Although Eq. (1) with (2) gives the correct relativistic relation between frequency and wave number, and hence between energy and momentum, the equation itself is not relativistically invariant, essentially because E and $\mathbf{p}$ do not occur in it in a similar manner. A concrete difficulty is this: Suppose that at $t = 0$ the particle is localized at $\mathbf{r} = 0$, that is, $\psi(\mathbf{r},0) = 0$ for $\mathbf{r} \neq 0$. Then it can be shown that, according to Eq. (1) with (2), at any later time $\psi(\mathbf{r},t) \neq 0$ for all $\mathbf{r}$. But, according to relativity, the particle could not have traveled faster than light, speed c, so that $\psi(\mathbf{r},t)$ should be zero for $|\mathbf{r}| > ct$. *See* SYMMETRY LAWS (PHYSICS).

Klein-Gordon equation. An equation without the above defects is the so-called Klein-Gordon equation, Eq. (3). But $\varphi(\mathbf{r},t)$, which satisfies this equation, can-

$$E^2\varphi = [(mc^2)^2 + \mathbf{p}^2c^2]\varphi \tag{3}$$

not not be a wave function, for two related reasons: (i) Equation (3) is second-order in $\partial/\partial t$, and so not only $\varphi(\mathbf{r},0)$ but also $\partial\varphi/\partial t$ is needed to determine the future values of φ; (ii) the only possible density of a conserved quantity formed from φ is of the form shown in relation (4). But this cannot be a proba-

$$\rho \propto \varphi^* E\varphi - \varphi E\varphi^* \tag{4}$$

bility density, because it is not positive definite (it changes sign when φ is replaced by φ^*).

But ρ, in relation (4), can be interpreted as a charge density (when multiplied by a unit charge e); φ is then to be interpreted as a matrix element of a field operator Φ of a quantized field whose quanta are particles with mass m and charge e or $-e$ and zero spin. [Equation (3) is obeyed by $\Phi(\mathbf{r},t)$, and hence by any of its matrix elements.] The same is true in an electromagnetic field, where relations (3) and (4) are to be modified as shown in Eq. (5), where $\phi(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$

$$E \to E - e\phi \qquad \mathbf{p} \to \mathbf{p} - e\mathbf{A} \tag{5}$$

are the scalar and vector potentials, respectively, of the electromagnetic field. Eigenstates of φ in a static field can be found in the usual way, which yields the energy levels of a charged spinless particle, say a π^- meson in the electric field of a nucleus, correctly except for the effects of radiative corrections (the electromagnetic interactions of the particle with itself and other virtually present particles) and of any internal structure of the particle. *See* EIGENFUNCTION; ENERGY LEVEL (QUANTUM MECHANICS).

Dirac equation. P. A. M. Dirac found a relativized form of Eq. (1), Eq. (6), which is both linear in E and

$$E\psi = [\beta mc^2 + \alpha \cdot \mathbf{p}c]\psi \qquad \rho \propto \psi^*\psi \tag{6}$$

has a positive definite density form ρ, where β and $\boldsymbol{\alpha}$ are constants which obey Eqs. (7). Applying the

$$\begin{gathered} \alpha_i\alpha_j + \alpha_j\alpha_i = 0 \qquad i \neq j \\ \alpha_i\beta + \beta\alpha_i = 0, \quad \alpha_i^2 = 1, \beta^2 = 1 \quad i, j = 1, 2, 3 \end{gathered} \tag{7}$$

equation a second time, Eq. (8) is obtained, that is,

$$\begin{aligned} E^2\psi &= [\beta mc^2 + \alpha \cdot \mathbf{p}c]^2\psi \\ &= [(mc^2)^2 + p^2c^2]\psi \end{aligned} \tag{8}$$

the Klein-Gordon equation, which assures that the energy and momentum of the particle are correctly related. Obviously the four constants β and α_i cannot be numbers; however, they can be 4×4 matrices, and ψ is then a four-component object called a Dirac spinor. *See* MATRIX THEORY.

If plane wave solutions of Dirac's equation (6) are considered, then $\mathbf{p}$ is now a number. Taking Eq. (6) as an eigenequation for E, four eigenstates are found (because H is a 4×4 matrix), two with

$$E = +\sqrt{(mc^2)^2 + p^2c^2}$$

and two with

$$E = -\sqrt{(mc^2)^2 + p^2c^2}$$

the interpretation of the two positive energy states is that they are the two spin states of a particle with spin $^1/_2\hbar$; in fact an operator representing the spin angular momentum can be constructed out of $\boldsymbol{\alpha}$.

But the two negative energy states are an embarrassment; even a particle that was initially in a positive energy state would quickly make radiative transitions down through the negative energy states. Dirac's solution was to observe that if the particle described by ψ obeyed the Pauli principle (as the electron or any other spin-$^1/_2\hbar$ particle must), then it can be supposed that all the negative energy states are already filled with particles, thus excluding any more. There are still four so-called single-particle states for a given momentum $\mathbf{p}$: the two spin states of a particle with positive energy, and the two states obtained by removing a negative energy particle (of momentum $-\mathbf{p}$). These last states (hole states) have positive energy and a charge opposite the charge of the particle. The hole is in fact the antiparticle; if the particle is an electron, the hole is a positron. *See* ANTIMATTER; ELECTRON-POSITRON PAIR PRODUCTION; POSITRON.

With the filling up of the negative energy states, the system in question is no longer a single-particle system, and ψ, just as in the Klein-Gordon case, no longer can be interpreted as a wave function but must be interpreted as a matrix element of a field operator $\mathbf{\Psi}$. Equation (6), with the modifications in Eq. (5), gives correct results for the eigenstates of a charged spin-$^1/_2\hbar$ particle in an electromagnetic field, except for the effects of radiative corrections and internal structure (compositeness); the latter is present only in hadrons such as the proton, but absent for the electron and muon. Particularly significant is the fact that the correct magnetic moment (Dirac moment) is given for the electron and muon (up to radiative corrections). *See* ELECTRON SPIN; ELEMENTARY PARTICLE; HADRON.

For a massless spin-$^1/_2\hbar$ particle, β does not appear in Dirac equation (6), and the α_i, satisfying Eqs. (7) with β omitted, can be 2×2 matrices; this is known as the Weyl equation. The particle, and likewise the antiparticle, then have only one spin state. But it can be shown that the particle has spin $^1/_2\hbar$; the particle (antiparticle) can only have its spin parallel (antiparallel) to its momentum. In an inverted coordinate system these statements would reverse, so that the two-component Weyl equation is not invariant to space inversion (parity). This "economical" version of the Dirac equation can be used to describe massless neutrinos. More fundamentally, the traces of chiral symmetry seen in weak interaction theory indicate that all spin-$^1/_2\hbar$ fundamental particles (leptons and quarks) satisfy the Weyl equation; the fact that the physical particles do not, except for massless neutrinos, is a consequence of the interactions of the particles. *See* LEPTON; NEUTRINO; QUARKS; SYMMETRY LAWS (PHYSICS).

Still another two-component variant of the Dirac equation is the Majorana equation, which describes a self-conjugate spin-$^1/_2\hbar$ particle. Self-conjugate means that the antiparticle is identical to the particle. The particle must be uncharged; it has only two intrinsic states, spin up and spin down. No particles are known to satisfy the Majorana equation, although massive neutrinos might.

Higher-spin particles. Wave equations for higher spin particles also exist. A well-known example is Maxwell equations for a (massless) spin-$1\hbar$ particle. The equations become more complicated as the spin becomes higher, and they will not be written here. When the spin is equal to or greater than $^1/_2\hbar$, the wave equation is not uniquely fixed by the particle's mass and spin; those higher electromagnetic moments which are allowed by the value of the spin (for spin $^1/_2\hbar$, magnetic dipole; for spin $1\hbar$, electric quadrupole also; and so forth) are arbitrary. However, if the particles are meant to be fundamental (point) particles, then the condition of renormalizability should be imposed, that is, the condition that radiative corrections be finite and calculable. In that case, the magnetic dipole moment of a spin-$^1/_2\hbar$ particle is fixed to be the Dirac value (before radiative corrections). A spin-$1\hbar$ particle must either be completely neutral (no electromagnetic interaction at all) or else be massless and be coupled as is a gauge particle. (Such particles can acquire mass through spontaneous symmetry breaking, as do the w and z bosons, for example, according to the standard model. Their magnetic dipole and electric quadrupole moments are predictable.) Finally, particles of spin $^3/_2\hbar$ or greater are forbidden to have any interactions at all. *See* MAXWELL'S EQUATIONS; QUANTUM ELECTRODYNAMICS; RELATIVISTIC MECHANICS; RENORMALIZATION; STANDARD MODEL; SYMMETRY BREAKING; WAVE EQUATION.

Charles J. Goebel

Relativity

A general theory of physics, primarily conceived by Albert Einstein, which involves a profound analysis of time and space, leading to a generalization of physical laws, with far-reaching implications in important branches of physics and in cosmology. Historically, the theory developed in two stages. Einstein's initial formulation in 1905 (now known as the special, or restricted, theory of relativity) does not treat gravitation; and one of the two principles on which it is based, the principle of relativity (the other being the principle of the constancy of the speed of light), stipulates the form invariance of physical laws only for inertial reference systems. Both restrictions were removed by Einstein in his general theory of relativity developed in 1915, which exploits a deep-seated equivalence between inertial and gravitational effects, and leads to a successful "relativistic" generalization of Isaac Newton's theory of gravitation.

Special Theory

The key feature of the theory of special relativity is the elimination of an absolute notion of simultaneity in favor of the notion that all observers always measure light to have the same velocity, in vacuum, c, independently of their own motion. The impetus for the development of the theory arose from the theory of electricity and magnetism developed by J. C. Maxwell. This theory accounted for all observed phenomena involving electric and magnetic fields and also predicted that disturbances in these fields would propagate as waves with a definite speed, c, in vacuum. These electromagnetic waves predicted

in Maxwell's theory successfully accounted for the existence of light and other forms of electromagnetic radiation. However, the presence of a definite speed, c, posed a difficulty, since if one inertial observer measures light to have velocity c, it would be expected that another inertial observer, moving toward the light ray with velocity v with respect to the first, would measure the light to have velocity $c + v$. Hence, it initially was taken for granted that there must be a preferred rest frame (often referred to as the ether) in which Maxwell's equations would be valid, and only in that frame would light be seen to travel with velocity c. However, this viewpoint was greatly shaken by the 1887 experiment of A. A. Michelson and E. W. Morley, which failed to detect any motion of the Earth through the ether. By radically altering some previously held beliefs concerning the structure of space and time, the theory of special relativity allows Maxwell's equations to hold, and light to propagate with velocity c, in all frames of reference, thereby making Maxwell's theory consistent with the null result of Michelson and Morley. *See* ELECTROMAGNETIC RADIATION; LIGHT; MAXWELL'S EQUATIONS.

Simultaneity in prerelativity physics. The most dramatic aspect of the theory of special relativity is its overthrowing of the notion that there is a well-defined, observer-independent meaning to the notion of simultaneity. To explain more precisely what is meant by absolute simultaneity, the following terminology will be introduced: An event is a point of space at an instant of time. Since it takes four numbers to specify an event—one for the time at which the event occurred and three for its spatial position—it follows that the set of all events constitutes a four-dimensional continuum, which is referred to as space-time.

A space-time diagram (**Fig. 1**) is a plot of events in space-time, with time, t, represented by the vertical axis and two spatial directions (x, y) represented by the horizontal axes. (The z direction is not shown.) For any event A shown in the diagram, there are many other events in this diagram—say, an event B—having the property that an observer or material body starting at event B can, in principle, be present at event A. (The world line—that is, path in space-time—of an observer going from B to A can be illustrated in the diagram.) The collection of all such events constitutes the past of event A. Similarly, there are many events—say, an event C—having the property that an observer or material body starting at event A can, in principle, be present at event C. These events constitute the future of A. Finally, there remain some events in space-time which lie neither to past nor future of A. In prerelativity physics, these events are assumed to make a three-dimensional set, and are referred to as the events which are simultaneous with event A. This notion of simultaneity in prerelativity physics is described as being absolute, because it does not depend upon any constructions associated with a particular observer and all observers agree upon which events are simultaneous with event A.

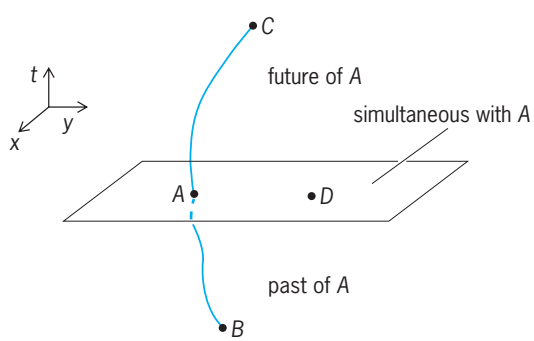

Fig. 1. Space-time diagram illustrating the causal relationships with respect to an event, *A*, in prerelativity physics. Event *B* lies to the past of *A*, event *C* lies to the future of *A*, and event *D* is simultaneous with *A*.

In both prerelativity physics and special relativity, an inertial observer is one who is not acted upon by any external forces. In both theories it is assumed that any inertial observer, $\mathcal{O}$ can build a rigid cartesian grid of meter sticks, all of which intersect each other at right angles. Observer $\mathcal{O}$ may then label the points on this cartesian grid by the coordinates (x, y, z) representing the distance of the point from $\mathcal{O}$ along the three orthogonal directions of the grid. Observer $\mathcal{O}$ may carry the grid without rotating it, so that each point on the grid also undergoes inertial motion. A clock may then be placed at each grid point. In prerelativity physics, these clocks may be synchronized by requiring that they start simultaneously with each other (using the absolute notion of simultaneity described above). Any event in space-time may then be labeled by the four numbers t, x, y, z as follows: The numbers x, y, z assigned to the event are the spatial coordinates of the grid point at which the event occurred, whereas t is the time of the event as determined by the synchronized clock situated at that grid point. This labeling of events is referred to as the global inertial coordinate system associated with observer $\mathcal{O}$. *See* FRAME OF REFERENCE.

It is of interest to compare the coordinate labelings given to events in space-time by two inertial observers, $\mathcal{O}$ and $\mathcal{O}'$, who are in relative motion. The relationship occurring in prerelativity physics is called a galilean transformation. In the simple case where $\mathcal{O}'$ moves with velocity v in the x direction with respect to $\mathcal{O}$, and these observers meet at the event A labeled by $(t, x, y, z) = (t', x', y', z') = (0, 0, 0, 0)$, with the axes of the grid of meter sticks carried by $\mathcal{O}'$ aligned (that is, not rotated) with respect to those of $\mathcal{O}$, the transformation is given by Eqs. (1). The

$$t' = t \tag{1a}$$

$$x' = x - vt \tag{1b}$$

$$y' = y \tag{1c}$$

$$z' = z \tag{1d}$$

galilean transformation displays in an explicit manner that the two inertial observers, $\mathcal{O}$ and $\mathcal{O}'$, agree upon the time labeling of events and, in particular,

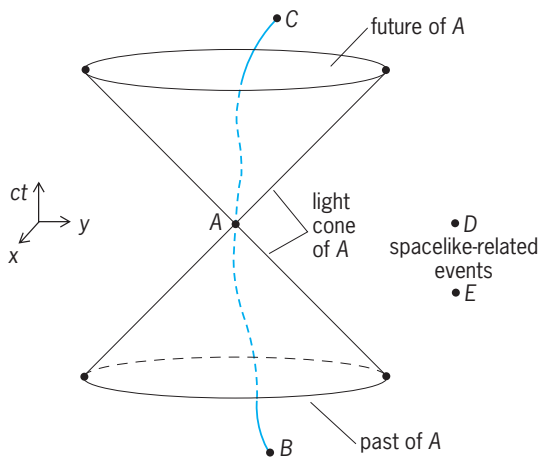

Fig. 2. Space-time diagram illustrating the causal relationships with respect to an event, *A*, in special relativity. Event *B* lies to the past of *A*, event *C* lies to the future of *A*, and events *D* and *E* are spacelike-related to *A*.

agree upon which events are simultaneous with a given event.

Causal structure in special relativity. The structure of space-time assumed in prerelativity physics as described above is both mathematically and logically self-consistent. Remarkably, however, nature does not appear to conform to this structure. Most dramatically, the causal structure of space-time in prerelativity physics (Fig. 1) is not correct. Rather, in special relativity there is a different causal relationship between an arbitrary event A and other events in space-time (**Fig. 2**). As in prerelativity physics, there are many events, B, which lie to the past of A; that is, an observer or material body starting at B can reach event A. There also are many events which lie to the future of A. However, there is now a much larger class of events which lie neither to the past nor to the future of A. These events are referred to as being spacelike-related to A.

The most striking feature of this causal structure (Fig. 2) is the absence of any three-dimensional surface of simultaneity. Indeed, the closest analog to the surface of simultaneity in prerelativity physics is the double-cone-shaped surface that marks the boundaries of the past and future of event A. This surface comprises the paths in space-time of all light rays which pass through event A, and for this reason it is referred to as the light cone of A. Thus, the statement that the events lying to the future of A are contained within the light cone of A is equivalent to the statement that a material body present at event A can never overtake a light ray emitted at event A. In special relativity, the light cone of an event A replaces the surface of simultaneity with event A as the absolute, observer-independent structure of space-time related to causality.

As in prerelativity physics, it is assumed in special relativity that an inertial observer, $\mathcal{O}$, can build a rigid grid of meter sticks, place clocks at the grid points, and label events in space-time by global inertial coordinates (t, x, y, z). The only difference from the procedure used for the construction of the similar coordinates in prerelativity physics is that the synchronization of clocks is now nontrivial, since the causal structure of space-time no longer defines an absolute notion of simultaneity. Nevertheless, any pair of clocks in $\mathcal{O}$'s system can be synchronized—and thereby all of $\mathcal{O}$'s clocks can be synchronized—by having an assistant stationed half-way between the clocks send a signal to the two clocks in a symmetrical manner. This synchronization of clocks allows $\mathcal{O}$ to define a notion of simultaneity; that is, $\mathcal{O}$ may declare events A_1 and A_2 to be simultaneous if time readings, t_1 and t_2, of the synchronized clocks at events A_1 and A_2 satisfy $t_1 = t_2$. However, since there is no longer any absolute notion of simultaneity defined by the causal structure of space-time, there is no guarantee that a different inertial observer, $\mathcal{O}'$, will agree with $\mathcal{O}$ on the simultaneity of events. Indeed, events judged by $\mathcal{O}$ to be simultaneous ($t_1 = t_2$) will, in general, be judged by $\mathcal{O}'$ to be nonsimultaneous ($t_1' = t_2'$). *See* LORENTZ TRANSFORMATIONS.

The key assumptions of special relativity are encapsulated by the following two postulates.

Postulate 1: The laws of physics do not distinguish between inertial observers; in particular, no inertial observer can be said to be at rest in an absolute sense. Thus, if observer $\mathcal{O}$ writes down equations describing laws of physics obeyed by physically measurable quantities in her global inertial coordinate system (t, x, y, z), then the form of these equations must be identical when written down by observer $\mathcal{O}'$ in his global inertial coordinates (t', x', y', z').

Postulate 2: All inertial observers (independent of their relative motion) must always obtain the same value, c, when they measure the velocity of light in vacuum. In particular, the path of a light ray in space-time must be independent of the motion of the emitter of the light ray. Furthermore, no material body can have a velocity greater than c.

It is a direct consequence of postulate 2 that two inertial observers, $\mathcal{O}$ and $\mathcal{O}'$, who are moving with respect to each other must disagree over the notion of simultaneity. The precise relationship between the labeling of events in space-time by the global inertial coordinate systems of two inertial observers, $\mathcal{O}$ and $\mathcal{O}'$, in special relativity is given by the Lorentz transformation formulas. In the simple case where $\mathcal{O}'$ moves with velocity v in the x direction with respect to $\mathcal{O}$ and crosses $\mathcal{O}'$s world line at the event labeled by $(t, x, y, z) = (t', x', y', z') = (0, 0, 0, 0)$ with spatial axes aligned, the Lorentz transformation is given by Eqs. (2).

$$t' = \frac{t - \dfrac{xv}{c^2}}{\sqrt{1 - \dfrac{v^2}{c^2}}} \qquad (2a)$$

$$x' = \frac{x - vt}{\sqrt{1 - \dfrac{v^2}{c^2}}} \qquad (2b)$$

$$y' = y \qquad (2c)$$

$$z' = z \qquad (2d)$$

Equation (2*a*) shows explicitly that $\mathcal{O}$ and $\mathcal{O}'$ disagree over simultaneity.

One immediate consequence of the disagreement between $\mathcal{O}$ and $\mathcal{O}'$ over simultaneity is the failure of the addition of velocities law of prerelativistic physics to hold. If $\mathcal{O}$ determines that $\mathcal{O}'$ moves (with respect to $\mathcal{O}$) with velocity v_1 in the x direction, and if $\mathcal{O}'$ similary determines that another inertial observer $\mathcal{O}''$ moves (with respect to $\mathcal{O}'$) with velocity v_2 in the x' direction, then velocity v of $\mathcal{O}''$ with respect to $\mathcal{O}$ as determined by $\mathcal{O}$ can be computed by composing the Lorentz transformations between $\mathcal{O}$ and $\mathcal{O}'$ and between $\mathcal{O}'$ and $\mathcal{O}''$ to get the Lorentz transformation between $\mathcal{O}$ and $\mathcal{O}''$, from which the velocity v can be read out. Equation (3) is thereby obtained. When either v_1 or v_2 is much less than c,

$$v = \frac{v_1 + v_2}{\left(1 + \dfrac{v_1 v_2}{c^2}\right)} \tag{3}$$

this reduces to the law $v = v_1 + v_2$ that would be obtained from the galilean transformation, Eq. (1).

Space-time geometry. A key question both in prerelativity physics and in special relativity is what quantities, describing the space-time relationships between events, are observer independent. Such quantities having observer-independent status may be viewed as describing the fundamental, intrinsic structure of space-time.

In prerelativity physics, the time interval $\Delta t \equiv t_2 - t_1$ between two events, labeled by an inertial observer $\mathcal{O}$ as (t_1, x_1, y_1, a_1) and (t_2, x_2, y_2, z_2), is observer independent. In other words, although a different inertial observer $\mathcal{O}'$ would attach different labels (t_1', x_1', y_1', z_1') and (t_2', x_2', y_2', z_2') to these two events, they always satisfy $\Delta t' = \Delta t$, where $\Delta t' \equiv t_2' - t_1'$. Similarly, in prerelativity physics any two inertial observers, $\mathcal{O}$ and $\mathcal{O}'$, would agree upon the spatial interval

$$\sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}$$

between two simultaneous events ($\Delta t = 0$).

It has already been seen that in special relativity the time interval, Δt, between two events is no longer observer independent. Furthermore, since different inertial observers disagree over simultaneity, the spatial interval between two simultaneous events is not even a well-defined concept, and cannot be observer independent. Remarkably, however, in special relativity, all inertial observers agree upon the value of the space-time interval, I, between any two events, where I is defined by Eq. (4).

$$I = (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2 - c^2(\Delta t)^2 \tag{4}$$

What is most remarkable about this formula for I is that it is very closely analogous to the formula for squared distance in euclidean geometry. Indeed, for ordinary euclidean geometry on a four-dimensional space, the cartesian coordinates (a_1, a_2, a_3, a_4) may be introduced to label the points. Many different cartesian coordinate systems may be introduced by making rotations of the axes of the original system, so the labels a_1, a_2, a_3, a_4 themselves have no intrinsic significance. However, the squared distance, D_2, between two points, given by the four-dimensional version of the pythagorean theorem, Eq. (5), is independent of the choice of cartesian

$$D^2 = (\Delta a_1)^2 + (\Delta a_2)^2 + (\Delta a_3)^2 + (\Delta a_4)^2 \tag{5}$$

coordinate system, in exactly the same way as I is independent of the choice of global inertial coordinate system.

The only mathematical difference between Eqs. (4) and (5) is The minus sign occurring in the last term in Eq. (4). This minus sign is of considerable importance, since it distinguishes between the notions of time and space in special relativity. Nevertheless, this minus sign turns out not to be a serious obstacle to the mathematical development of the theory of lorentzian geometry based upon the space-time interval, I, in a manner which parallels closely the development of euclidean geometry based upon D_2. In particular, notions such as geodesics (straightest possible lines) can be introduced in lorentzian geometry in complete analogy with euclidean geometry. The Lorentz transformation between the two inertial observers is seen from this perspective to be the mathematical analog of a rotation between two cartesian frames in euclidean geometry.

The formulation of special relativity as a theory of the lorentzian geometry of space-time is of great importance for the further development of the theory, since it makes possible the generalization which describes gravitation. The lorentzian geometry defined by Eq. (4) is a flat geometry, wherein initially parallel geodesics remain parallel forever. The theory of general relativity accounts for the effects of gravitation by allowing the lorentzian geometry of space-time to be curved. *See* SPACE-TIME.

Consequences. The theory of special relativity makes many important predictions, the most striking of which concern properties of time. One such effect, known as time dilation, is predicted directly by the Lorentz transformation formula (2*a*). If observer $\mathcal{O}$ carries a clock, then the event at which $\mathcal{O}'$s clock reads time τ would be labeled by her as $(\tau, 0, 0, 0)$. According to Eq. (2*a*), the observer $\mathcal{O}'$ would label the event as $(t', x', 0, 0)$ where t' is given by Eq. (6). Thus, $\mathcal{O}'$ could say that a clock carried by $\mathcal{O}$

$$t' = \frac{\tau}{\sqrt{1 - \dfrac{v^2}{c^2}}} > \tau \tag{6}$$

slows down on account of $\mathcal{O}'$s motion relative to $\mathcal{O}'$. Similarly, $\mathcal{O}$ would find that a clock carried by $\mathcal{O}'$ slows down with respect to hers. This apparent disagreement between $\mathcal{O}$ and $\mathcal{O}'$ as to whose clock runs slower is resolved by noting that $\mathcal{O}$ and $\mathcal{O}'$ use different notions is simultaneity in comparing the readings of their clocks. *See* CLOCK PARADOX.

The decay of unstable elementary particles provides an important direct application of the time dilation effect. If a particle is observed to have a decay lifetime T when it is at rest, special relativity predicts

that its observed lifetime will increase according to Eq. (6) when it is moving. Exactly such an increase is routinely observed in experiments using particle accelerators, where particle velocities can be made to be extremely close to c. *See* PARTICLE ACCELERATOR.

An even more striking prediction of special relativity is the clock paradox: Two identical clocks which start together at an event A, undergo different motions, and then rejoin at event B will, in general, register different total elapsed time in going from A to B. This effect is the lorentzian geometry analog of the mundane fact in euclidean geometry that different paths between two points can have different total lengths.

Another effect predicted by special relativity which is closely related to the time dilation effect is the relativistic Doppler shift for light (or other electromagnetic) waves. It can be derived directly from the Lorentz transformation formulas (2*a*) and (2*b*). *See* DOPPLER EFFECT.

Many other ramifications of special relativity arise directly from the assumption that the laws of physics cannot prefer any global inertial coordinate system (postulate 1) and, thus, in particular, cannot make use of any notion of simultaneity. In prerelativity physics, the momentum three-vector of a particle, $\vec{p} = m\vec{v}$, appears in many laws. However, from a space-time perspective, $\vec{p} = (p^x, p^y, p^z)$ points in a spatial direction; that is, it is tangent to a surface of simultaneity. A law of physics involving $\vec{p}$ alone could not satisfy postulate 1. However, $\vec{p}$ could appear in laws of physics if it were the spatial components of a space-time four-vector, $p^\mu = (p^t, p^x, p^y, p^z)$. The only natural choice of such a space-time vector is given by Eq. (7), where m is the rest mass of the par-

$$p^\mu = mu^\mu \tag{7}$$

ticle and u^μ denotes the tangent vector to the curve describing the world line of the particle in space-time, with this curve parametrized by the time on a clock carried by the particle. The time component, p^t, of p^μ in the global inertial coordinate system of observer $\mathcal{O}$ then has the natural interpretation of representing the energy, E, divided by c, of the particle as measured by $\mathcal{O}$; that is $pt = E/c$. Equation (7) then yields Eq. (8) for the energy and momentum of a par-

$$E = \frac{mc^2}{\sqrt{1 - \dfrac{v^2}{c^2}}} \tag{8a}$$

$$\vec{p} = \frac{m\vec{v}}{\sqrt{1 - \dfrac{v^2}{c^2}}} \tag{8b}$$

ticle in special relativity. *See* ENERGY; MOMENTUM; RELATIVISTIC MECHANICS.

The most striking feature of Eq. (8) is that an energy given by Eq. (9) must be assigned to a particleat

$$E = mc^2 \tag{9}$$

rest. This contrasts sharply with the situation in prerelativity physics, where the energy is undefined up to the addition of an arbitrary constant and, by convention, this constant is normally chosen to make the energy of a particle at rest be zero. The existence of the relationship (9) implies that rest mass should be viewed as a form of energy, and thereby suggests the possibility that rest mass could be converted to other forms of energy. The fact that this possibility can be realized is perhaps the most significant ramification of the theory of special relativity. *See* REST MASS.

Robert H. Wald

General Theory

General relativity is the geometric theory of gravitation developed by Einstein in the decade before 1915. It is a generalization of special relativity, and includes the classical gravitational theory of Newton as its limit for weak gravitational fields and low velocities. The most important applications of the theory are to cosmology and to the structure of neutron stars and black holes. Due to astronomical observations since the 1990s, general relativity is now in the forefront of fundamental physics research. *See* GRAVITATION.

Need for a relativistic theory of gravity. One of the basic results of special relativity is that no physical effect can propagate with a velocity greater than that of light c; nature has a universal speed limit. However, classical gravitational theory describes the gravitational field of a body everywhere in space as a function of its instantaneous position, which is equivalent to the assumption that gravitational effects propagate with an infinite velocity. Thus, special relativity and classical gravitational theory are inconsistent, and a relativistic theory of gravity is necessary.

Principle of equivalence. Dating to Galileo, it is a fundamental question as to why bodies of different mass fall with the same acceleration in a gravitational field. This is now called the universality of free fall or the weak equivalence principle, and has been well tested experimentally. The best tests at present involve measuring the distance to the Moon with extreme accuracy by using laser beams bounced from reflectors on the lunar surface; these have been used to verify the principle to better than a part in 10^{12}, and this should be improved in the near future.

The universality of free fall follows from classical Newtonian gravitational theory since both the gravitational force on a body and its inertial resistance to acceleration are proportional to its mass. Thus the mass cancels out of the equations for the motion. However, this is not really an explanation but only an ad hoc description.

A deeper and more natural explanation occurred to Einstein. If an observer in the gravitational field of the Earth and another in an accelerating elevator or rocket in free space (**Fig. 3**) both drop a test body, they will both observe it to accelerate relative to the floor. According to classical theory, the Earth-based observer would attribute this to a gravitational force, while the elevator-based observer would attribute it to the accelerated floor overtaking the uniformly moving body. For both observers the motion is identical, and in particular the acceleration is independent

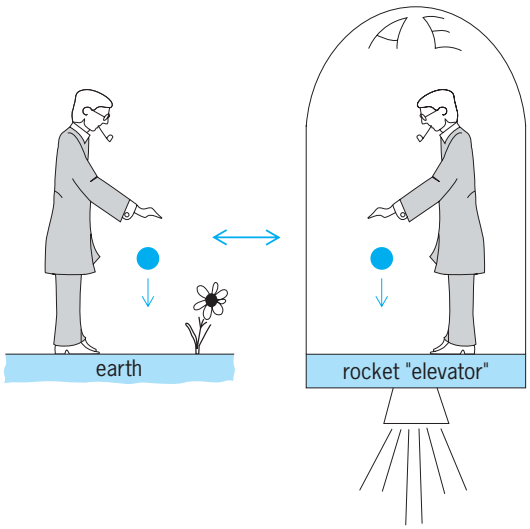

Fig. 3. Einstein's elevator, illustrating the principle of equivalence. The elevator acceleration is equal to *g*, the acceleration of gravity.

of the mass of the test body. Einstein elevated this fact to a general principle, called the Einstein principle of equivalence. This states that on a local scale all physical effects of a gravitational field are indistinguishable from the physical effects of an accelerated reference frame. This profound principle is a cornerstone of the theory of general relativity. From the point of view of the principle of equivalence, it is obvious why the motion of a test body in a gravitational field is independent of its mass. But the principle applies not only to gravitational effects but to all physical phenomena, and thereby has profound consequences (discussed below).

The principle of equivalence is true in general relativity, but not in modifications or in other gravity theories; for example, string theory suggests a violation of very roughly a part in 10^{17}. For the next decade a satellite test of the equivalence principle (STEP) is planned, which may achieve an accuracy of about 10^{18} and thereby provide a rare experimental test of string theory. *See* SUPERSTRING THEORY.

Tensors and Einstein's field equations. The observational fact that the motion of a body in a gravitational field is independent of its mass suggested to Einstein that gravity is fundamentally a geometric phenomenon, and that bodies should follow a special path determined only by geometry (see below). Because of this, gravity is naturally described by the mathematics of four-dimensional geometry. This geometry involves tensors, which are a simple generalization of vectors, and systems of tensor equations that are manifestly independent of the coordinate system. *See* TENSOR ANALYSIS.

Suppose that the points in four-dimensional space-time are labeled by two essentially arbitrary systems of coordinates x^μ and x'^μ (the indices μ and ν range from 0 to 3) and that the two coordinate systems are related to each other by a differentiable set of four functions, $x'^\mu = x'^\mu(x^\alpha)$ and $x^\nu = x^\nu(x'^\mu)$. Then contravariant tensors of rank 0, rank 1, rank 2, and so forth are defined as sets of 1, 4, 16, and so forth numbers or functions whose values at the same point in the two systems of coordinates are related by the transformation Eqs. (10). Covariant tensors are similarly defined by Eqs. (11).

$$\begin{aligned} &\text{rank 0:} && \phi' = \phi \\ &\text{rank 1:} && \eta'^\mu = \frac{\partial x'^\mu}{\partial x^\alpha}\eta^\alpha \\ &\text{rank 2:} && T'^{\mu\nu} = \frac{\partial x'^\mu}{\partial x^\alpha}\frac{\partial x'^\nu}{\partial x^\beta}T^{\alpha\beta} \end{aligned} \tag{10}$$

$$\begin{aligned} &\text{rank 1:} && \sigma'_\mu = \frac{\partial x^\alpha}{\partial x'^\mu}\sigma_\alpha \\ &\text{rank 2:} && \omega'_{\mu\nu} = \frac{\partial x^\alpha}{\partial x'^\mu}\frac{\partial x^\beta}{\partial x'^\nu}\omega_{\alpha\beta} \end{aligned} \tag{11}$$

In these equations the indices that appear twice in an expression are to be summed over; this is the Einstein summation convention. Tensors of rank 0 are also called scalars or invariants since they have the same value in all coordinate systems, and are thus of great importance in the description of physical quantities. Tensors of rank 1 are also called four-vectors.

The space-time of relativity contains one rank-2 tensor of particular importance, called the metric tensor $g_{\mu\nu}$, which is a generalization of the Lorentz metric of special relativity, introduced in Eq. (4). The metric tensor completely describes the geometry of space-time. Nearby points in space-time, also called events, which are separated by coordinate distances dx^μ, have an invariant physical separation whose square, called the line element, is defined by Eq. (12).

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu \tag{12}$$

This quantity is a generalization of the space-time interval in special relativity. When ds^2 is positive, it represents the square of the time interval between nearby events (multiplied by c^2) measured by an observer who moves at approximately constant velocity in such a way as to be present at both events. [A space with a quadratic metric form as in Eq. (12) is called a Riemann space.] *See* RIEMANNIAN GEOMETRY.

In a tensor equation one tensor of a given type is set equal to another of the same type, for example, $T^\mu = S^\mu$. In a different coordinate system, denoted by a prime, the definition of a tensor implies that both sides transform in the same way, so that the partial derivatives of the transformation cancel from the equation and it has the same form as in the original system, for example, $T'^\mu = S'^\mu$. Tensor equations are thus called form-invariant, or covariant.

The field equations of general relativity are tensor equations for the metric tensor. They will be presented very briefly. A Christoffel symbol (not a tensor) is first defined by Eq. (13); here $g_{\mu\tau}$ is the inverse of the matrix $g_{\mu\tau}$.

$$\left\{ \begin{matrix} \mu \\ \alpha\beta \end{matrix} \right\} = \frac{1}{2}g^{\mu\sigma}\left(\frac{\partial g_{\sigma\alpha}}{\partial x^\beta} + \frac{\partial g_{\sigma\beta}}{\partial x^\alpha} - \frac{\partial g_{\alpha\beta}}{\partial x^\sigma}\right) \tag{13}$$

The Riemann tensor (or curvature tensor) is then defined in terms of the Christoffel symbols by Eq. (14):

$$R^{\alpha}_{\mu\beta\nu} = \frac{\partial}{\partial x^{\nu}} \left\{ \begin{matrix} \alpha \\ \beta\mu \end{matrix} \right\} - \frac{\partial}{\partial x^{\mu}} \left\{ \begin{matrix} \alpha \\ \beta\nu \end{matrix} \right\} + \left\{ \begin{matrix} \alpha \\ \tau\nu \end{matrix} \right\} \left\{ \begin{matrix} \tau \\ \beta\mu \end{matrix} \right\} - \left\{ \begin{matrix} \alpha \\ \tau\beta \end{matrix} \right\} \left\{ \begin{matrix} \tau \\ \mu\nu \end{matrix} \right\} \quad (14)$$

This tensor plays a central role in the geometric structure of space-time; if it is zero, the space-time is termed flat and has no gravitational field; if nonzero, the space-time is termed curved and a gravitational field is present. In terms of the Ricci tensor, which is the Riemann tensor summed over $\alpha = \beta$, the Einstein field equations for empty space-time are given in Eq. (15):

$$R^{\alpha}_{\mu\alpha\nu} \equiv R_{\mu\nu} = 0 \qquad \text{(empty space)} \qquad (15)$$

The field equations are a set of 10 second-order partial differential equations, since $R^{\mu\nu}$ is a four-by-four symmetric tensor and has 10 independent components; the equations are to be solved for the metric tensor in a coordinate system that we are free to choose. Since the curvature of this space-time corresponds to the intrinsic presence of a gravitational field, the idea of a force field in classical gravitational theory is replaced in general relativity by the geometric concept of curved space-time. *See* DIFFERENTIAL EQUATION.

In a nonempty region of space-time the field equations (15) must be modified to include a tensor that represents the matter or energy which is the source of gravity, called the energy-momentum tensor $T_{\mu\nu}$. The modified equations are given in Eq. (16), where G is the gravitational constant, equal to 6.67×10^{-11} $\mathrm{N \cdot m^2 \cdot kg^{-2}}$: [For empty space, $G_{\mu\nu} = 0$, which can be shown to imply $R_{\mu\nu} = 0$ also, so this is consistent with Eq. (15).]

$$G_{\mu\nu} \equiv R_{\mu\nu} - {}^{1}/_{2}\, g_{\mu\nu} R^{\alpha}_{\alpha} = -\frac{8\pi G}{c^4} T_{\mu\nu} \qquad \text{(nonempty space)} \qquad (16)$$

The tensor $G_{\mu\nu}$, called the Einstein tensor, represents the space-time geometry, and the tensor $T_{\mu\nu}$ represents the mass or energy that produces gravity. These equations automatically imply the conservation of energy and momentum, which is an extremely important result. Moreover, in the limiting case when the fields are weak and velocities small compared to c, the equations reduce to the equations of classical Newtonian gravity.

Cosmological term. The field equations were given in the form of Eq. (16) by Einstein in 1916. However, they can be consistently generalized by the addition of another term on the left side, which he called the cosmological term, $\Lambda g_{\mu\nu}$. The more general equations are given in Eq. (17).

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = -\frac{8\pi G}{c^4} T_{\mu\nu} \qquad (17)$$

The constant Λ is called the cosmological constant, and has the dimension of an inverse length squared. The extra term corresponds to a repulsive force between bodies that increases with distance.

Einstein introduced the cosmological term in 1917 in order to balance the usual attractive force of gravity and obtain static models of the universe, since it was then believed that the universe was static. When it was discovered by E. P. Hubble in 1929 that the universe is expanding, as evidenced by the Doppler shifts of distant galaxies, Einstein abandoned the cosmological term. However, beginning in 1998 measurements of the distance and recession velocity of very distant supernovae indicated that the cosmological term, or something that behaves just like it, is present and is causing the expansion of the universe to accelerate. The observed value of the constant is about $(10^{10}\ \text{light-years})^{-2}$, so the cosmological term is only relevant on the cosmological scale. *See* COSMOLOGICAL CONSTANT; SUPERNOVA.

In the more modern view the cosmological term may be taken to the right side of the field equations and interpreted as the energy-momentum tensor of the vacuum; the energy density of the vacuum is $\rho_v = \Lambda c^4/8\pi G$. Just such a vacuum energy is predicted by quantum field theory, but its magnitude is, very roughly, an astounding 120 orders of magnitude greater than allowed by observation. This nonsensical result (with related questions) is called the cosmological constant problem (discussed below).

The energy of the vacuum, in whatever guise, is now generally referred to as dark energy (discussed below). Its magnitude, as inferred from observation, is about 3.5 GeV/m^3, and it comprises about 73% of the cosmic energy density; although it constitutes the dominant stuff of the universe, its nature is unknown, which is one of the biggest unresolved questions in physics today. *See* COSMOLOGY; DARK ENERGY.

Motion of test bodies. Test bodies are objects with very small mass and size and no internal structure. In general relativity the path of a test body in space-time is a generalized straight line called a geodesic; a geodesic may be thought of as a curve that is "everywhere parallel to itself," or as an "extremal curve," either the shortest or longest curve between two given points. For example, a straight line is a geodesic in Euclidean space, and a great circle is a geodesic on the surface of the Earth. In relativity a geodesic is the longest curve between two points in space-time, with the distance being defined as the integral of ds along the curve. In the special case of no gravitational field, with the metric equal to the Lorentz metric of special relativity, a geodesic is a simple straight line.

The scenario for gravitational influence in general relativity is that matter curves space in its vicinity, in accord with the Einstein equations, and in this curved space test bodies move on geodesics. The motion is thereby clearly independent of the mass of the test body, in accord with the principle of equivalence.

General relativity theory possesses an extraordinary property regarding geodesics: because the field

equations are nonlinear (unlike those of classical theory), the motion of a test body in a gravitational field cannot be specified independently of the field equations, since the body has mass and contributes a small amount to the field. Indeed the field equations are so restrictive that the equations of motion may be derived from them and need not be treated as a postulate.

Schwarzschild's solution. A very important solution of the field equations was obtained by K. Schwarzschild in 1916, surprisingly soon after the inception of general relativity. This solution represents the field in empty space around a spherically symmetric body such as the Sun. It is the basis for a relativistic description of the solar system and most of the experimental tests of general relativity to date. In spherical coordinates r, θ, ϕ and time coordinate $x^0 = ct$ the solution is represented by the line element given by Eq. (18), which is valid only outside the body:

$$ds^2 = \left(1 - \frac{2GM}{c^2 r}\right) c^2 dt^2 - \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad (18)$$

Here M is the mass of the body. In the limit of $M = 0$ this is the line element of special relativity written in spherical coordinates. This solution is the relativistic analog of the classical gravitational potential field of a spherical body, $\phi = -GM/r$. The quantity in each of the parentheses becomes zero at a radius $2GM/c^2$, the Schwarzschild radius, and infinite at the origin, which is a singular point.

Gravitational redshift. Electromagnetic radiation of a given frequency emitted in a gravitational field will appear to an outside observer to have a lower frequency; that is, it will be redshifted. The formula for the redshift can be derived from the principle of equivalence without the use of the Einstein field equations or the Schwarzschild solution, so redshift experiments test the basic idea of relativity but do not test the field equations. The most accurate test of the redshift to date was performed using a hydrogen maser on a rocket. Comparison of the maser frequency with Earth-based masers gave a measured redshift in agreement with theory to about 1 part in 10^4. *See* GRAVITATIONAL REDSHIFT.

Perihelion shift of Mercury. The geodesic equations of motion can be solved for the orbit of a planet, considered as a test body in the Schwarzschild field of the Sun. As should be expected, the planetary orbit obtained is very similar to the ellipse of classical theory. However, according to relativity the ellipse rotates very slowly in the plane of the orbit so that the perihelion, the point of closest approach of the planet to the Sun, is at a slightly different angular position for each orbit. This shift is extremely small for all the planets of the solar system, but it is greatest for Mercury, whose perihelion advance is predicted to be 43 seconds of arc in a century. This agrees with the discrepancy between classical theory and observation that was well known for many years before the discovery of general relativity. Results of very precise planetary and asteroid perihelion position measurements using radar are consistent with general relativity theory. *See* CELESTIAL MECHANICS; RADAR ASTRONOMY.

Deflection of light. The principle of equivalence suggests an extraordinary phenomenon of gravity. Light crossing the Einstein elevator (Fig. 3) horizontally will appear to be deflected downward in a parabolic arc because of the upward acceleration of the elevator. The same phenomenon must therefore occur for light in the gravitational field of the Sun; that is, it must be deflected toward the Sun. A calculation of this deflection gives 1.75″ for the net deflection of starlight grazing the edge of the Sun. The first measurements of this effect in 1919 verified the qualitative correctness of the deflection phenomenon. Modern measurements, made by tracking quasars with radio telescopes as they pass near or behind the Sun, find the deflection to be within 0.1% of the value predicted by general relativity.

In 1936 Einstein observed that if two stars were exactly lined up with the Earth, the more distant star would appear as a ring of light, distorted from its point appearance by the lens effect of the gravitational field of the nearer star. It was soon pointed out that a very similar phenomenon was much more likely to occur for entire galaxies instead of individual stars. Many examples of such gravitational lens systems have now been found. Some involve a quasar at a very large distance, nearly aligned with a normal galaxy nearer to us. Much effort is being given to gravitational lenses as tools to make absolute measurements of the distance to quasars, to measure the Hubble constant, and to seek and study the nonluminous dark matter. *See* DARK MATTER; GRAVITATIONAL LENS; QUASAR; RADIO ASTRONOMY.

Radar time delay. In the curved space around the Sun the distance between points in space, for example between two planets, is not the same as it would be in flat space. In particular, the round-trip travel time of a radar signal sent between the Earth and another planet will be increased by the curvature effect when the Earth, the Sun, and the planet are approximately lined up. Using radar ranging to planets and spacecraft, the time delay has been found to agree with the predictions of general relativity to an accuracy of better than 0.1%

Precession of a gyroscope. Relativity theory predicts that a spinning gyroscope in orbit around the Earth will precess by a small amount. Most of this relativistic precession is due to the curvature of space-time described by the Schwarzschild metric in Eq. (18); it is called the geodetic precession, and amounts to about 6.6 arcseconds per year for a low Earth orbit.

However, there is an additional smaller precession of great interest. For weak gravitational fields the Einstein field equations resemble Maxwell's equations of electromagnetism, and contain two types of fields: one resembles the electric field and is called the gravito-electric field, since it is due simply to the presence of matter; the other resembles the magnetic field and is called the gravito-magnetic

field, since it is due to matter in motion, that is, a mass current. The Schwarzschild metric describes the gravito-electric field of the Earth. The spin of the Earth produces a mass current which causes a gravito-magnetic field, and this gives rise to further precession of a gyroscope in orbit. Effects of the gravito-magnetic field are generally referred to as Lense-Thirring effects (following the work of J. Lense and H. Thirring in 1918) or as frame dragging effects. For an appropriately oriented low polar orbit and gyroscope spin direction the Lense-Thirring precession is about 0.042 arcseconds per year, and is orthogonal to the geodetic precession, so the two precessions may be separately measured. *See* MAXWELL'S EQUATIONS.

The existence of the gravito-magnetic field may be quite reliably inferred from the existence of the gravito-electric field and the Lorentz transformation, but before 2006 had never been directly verified in a clean experiment. The Gravity Probe B (GP-B) spacecraft was launched in April 2004 and has made a high-precision measurement of the geodetic precession and the first direct measurement of the Lense-Thirring precession—and thus the gravito-magnetic field of the Earth. Data taking was completed in 2005, and data analysis and release of results should be completed in 1 or 2 years. *See* GYROSCOPE.

Neutron stars. Pulsars are astronomical sources of electromagnetic radiation emitted in a series of precisely repeated pulses at a rate ranging from about 1 per second to over 100 per second. They are generally believed to be rotating neutron stars, the residual cores of heavy stars left behind after supernova explosions.

For ordinary stars, such as the Sun, the gravitational field is sufficiently weak that classical gravitational theory is a good approximation for studying the internal structure. This is not true for neutron stars, which contain a core composed largely of neutrons, and a few other elementary particles, at densities of about 2×10^{14}g/cm^3, greater than nuclear density; neutron stars pack more than a solar mass into a radius of only about 10 km (6 mi). The gravitational field in these stars can thus be quite large, and it is necessary to describe them using hydrostatic equations derived from the Einstein field equations in the presence of matter. These are called the Tolman-Oppenheimer-Volkov (TOV) equations.

One of the most remarkable theoretical predictions, found from the TOV equations, is an upper limit for the mass of a neutron star, above which it cannot be stable. This limit is similar to the well-known white dwarf mass limit of 1.4 solar masses, found by S. Chandrasekhar. One reason for neutron star instability above the limit is that, according to general relativity, pressure in the star is a source of gravity and contributes to the downward force of gravity. The exact value of the mass limit depends on the theoretical equation of state assumed, but is about a few solar masses. Many neutron stars have been observed; the masses cluster around 1.4 solar masses and none is much greater. For example, the binary pulsar (discussed below) consists of a pulsar and a neutron star companion, both with about 1.4 solar masses. Many astronomical x-ray sources are powered by the strong gravitational fields of neutron stars, according to current theory, and this is presently a very active field of astronomy. *See* NEUTRON STAR; PULSAR; X-RAY ASTRONOMY.

Binary pulsar. The binary pulsar PSR 1913+16, discovered by R. Hulse and J. Taylor in 1974, is a pulsar in orbit about a neutron star companion. The orbital period is about 8 h, the orbital radius is about 10^6 km (6×10^5 mi), both stars have about 1.4 solar masses, and the pulse period is about 59 ms. Both stars move at about 300 km/s (186 mi/s), and the system is clean in that it has few nongravitational complications. Timing of its pulses to an accuracy of about a part in 10^{10} has been used to obtain much information about the system. Its orbital period is obtained from the Doppler shift of the pulses as the pulsar moves toward and away from us, which in turn gives the periastron advance, which is over 4° per year. Many of the general relativistic effects measured in the solar system have also been measured for the binary pulsar, and all are consistent with general relativity.

Perhaps the most interesting feature of the binary pulsar is that for over three decades of observation the period has decreased precisely as predicted by general relativity for energy loss to gravitational radiation (discussed below). The data agree with theory to better than 1%, which is convincing evidence for the existence of gravitational radiation.

Gravitational radiation. Gravitational radiation is analogous to electromagnetic radiation; whereas electromagnetic radiation is emitted by charges in accelerated motion, gravitational radiation is emitted by masses in accelerated motion. Such motion produces ripples or waves in the gravitational field that propagate at the velocity of light. The usual method of studying these ripples is with linearized general-relativity theory; in this theory the metric tensor is expressed as the constant Lorentz metric of special relativity plus a small perturbation term representing small gravitational ripples in an almost flat space. The forces exerted by plane gravitational waves are at right angles to the direction of propagation of the wave; that is, they are transverse (**Fig. 4**). As with electromagnetic waves, there are two polarizations possible for gravitational waves; the force field of one polarization is rotated 45° to the other. *See* ELECTROMAGNETIC RADIATION; PERTURBATION (MATHEMATICS); POLARIZATION OF WAVES.

Most relativists consider the binary pulsar observations to be convincing evidence for the existence of gravitational radiation, but direct detection is still of the utmost importance. There are many expected astrophysical sources of gravitational radiation. These include violent events such as supernova explosions, the gravitational collapse of stars, and the collisions of neutron stars and black holes. Waves from even such violent events are expected to produce only a fractional length distortion of 10^{-20} or less when they reach the Earth.

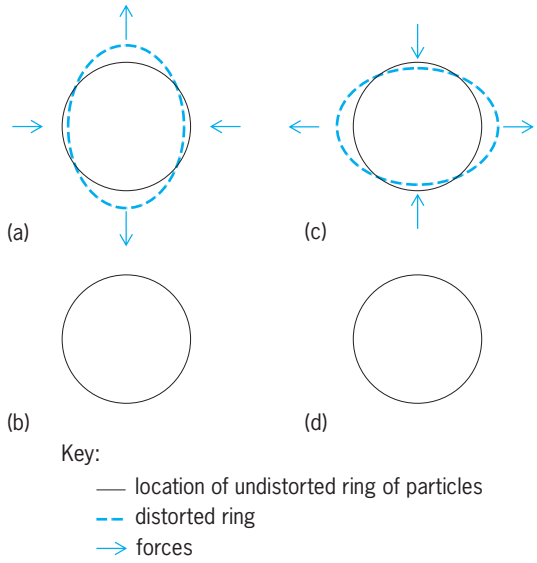

Fig. 4. The forces and distortions produced by a gravitational wave. (*a*) Begin cycle. (*b*) $^1/_4$ cycle (forces equal zero). (*c*) $^1/_2$ cycle (forces reverse). (*d*) $^3/_4$ cycle (forces equal zero).

Detection of astrophysical gravitational waves is difficult, but in principle the operation of real detectors is nearly as simple as shown in Fig. 4. The most sensitive modern detectors use laser interferometers; for example, the Laser Interferometer Gravitational-wave Observatory (LIGO) is a pair of L-shaped interferometers about 4 km (2.5 mi) in size, in which beams of laser light measure the difference in length between the two legs. The laser light bounces back and forth to give an effective length much greater than 4 km. For a typical expected wave the effective displacement is only about the size of a nucleus, but is measurable. LIGO has only been in operation since 2002, and no waves have yet been detected as of 2006, but as sensitivity is increased with various upgrades the first detection may be expected in a few years. *See* LIGO (LASER INTERFEROMETER GRAVITATIONAL-WAVE OBSERVATORY).

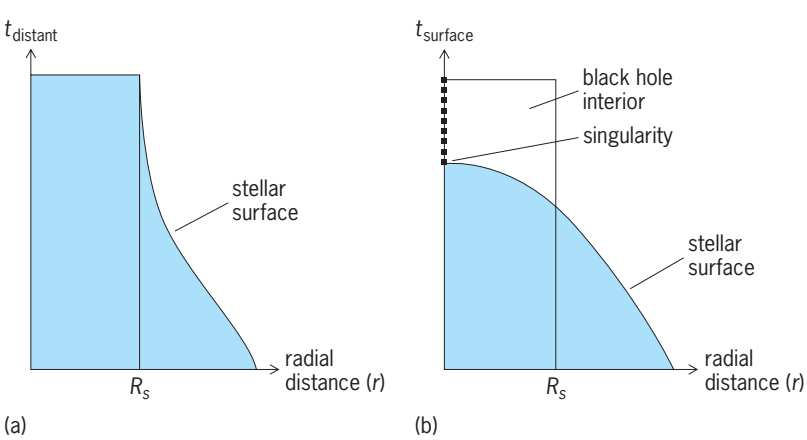

Fig. 5. Gravitational collapse. (*a*) From the point of view of a distant observer, the surface of a collapsing star approaches the Schwarzschild radius R_s or black hole surface. (*b*) As seen by an observer riding on its falling surface, it collapses through the black hole surface to reach zero size at a singularity.

Detection of gravitational radiation may open a new observational window that could be as important as the windows opened by neutrino astronomy and gamma-ray astronomy, which have changed our view of the universe. With gravitational radiation it might be possible to observe such things as the interiors of exploding supernovas and gamma-ray bursters, and see directly back to the early stages of the universe, none of which are accessible with light. *See* GAMMA-RAY BURSTS; GRAVITATIONAL RADIATION.

Black holes. If the core of an exploding supernova is heavier than the mass limit for a neutron star, as discussed above, there is no known stable end state. In 1937 J. R. Oppenheimer and H. Snyder predicted that such a core, when viewed from outside, will asymptotically collapse toward the Schwarzschild radius, $R_s = 2GM/c^2$. At that radius the redshift is infinite so no light can escape; effectively, time on the surface stops. This state is now known as a black hole, a name introduced by J. A. Wheeler. The prediction was verified in the 1970s when astrophysical x-ray sources were discovered. Many x-ray sources are clearly very compact objects with masses too large to be neutron stars, and black holes are the obvious candidates. One example is the source GRO J 1655-40 with 7.02 ± 0.22 solar masses. Moreover, most galaxies are believed to contain supermassive black holes of millions or billions of solar masses, and the huge energy output of quasars and radio galaxies are explained in terms of matter falling into black holes and emitting radiation.

Astronomical observations are limited to black hole exteriors, but theorists are also interested in their surfaces and interiors. To an outside observer the surface of a collapsing star (or stellar core) asymptotically approaches the Schwarzschild or black hole radius on a time scale of typically 10^{-5} s. Since the star is filled with burnt-out stellar material, the interior is not described by the Schwarzschild solution, which is only valid for the empty space outside the star. Thus there is no singularity inside the star.

However an observer sitting on the surface of the collapsing star measures a different time from that of an exterior observer (**Fig. 5**). For the surface observer the surface of the star passes through the Schwarzschild radius and continues on to the origin, where the density becomes infinite and a space-time singularity forms. At the Schwarzschild radius the surface observer's time differs by an infinite amount from that of a distant outside observer; the surface observer passes the Schwarzschild radius only after an infinite external time. In passing the surface of the black hole at the Schwarzschild radius, the surface observer enters the black hole interior, a region of space-time that is not accessible to any outside observer. Neither particles nor light can pass outward through the black hole surface. The surface thus acts as a one-way membrane or horizon, and outside observers cannot see inside.

If a collapsing star has angular momentum, the asymptotic black hole state is a generalization of the Schwarzschild solution to rotating systems,

discovered by R. Kerr. This solution is believed to be the general final asymptotic state for collapsing stars with angular momentum but no electric charge. A rotating black hole, described by the Kerr solution, has a gravito-magnetic field due to its angular momentum.

Gravitational singularities are a serious theoretical problem; they can produce paradoxical effects such as a flux of unlimited energy. But the singularity inside a black hole is inside a horizon and cannot affect the exterior world, which is thereby protected from such anomalies. R. Penrose and S. Hawking have conjectured a "cosmic censorship" principle, such that in relativity all singularities are similarly hidden by a horizon. Other theorists believe that singularities should not occur if quantum effects are taken into account, making cosmic censorship irrelevant.

Unlike an astrophysical black hole as discussed above, an "eternal" black hole is a theoretical object described by the Schwarzschild geometry everywhere in space-time. It is a region of empty space, containing a singularity, which has existed forever and was not formed by the collapse of a star. Part of the space-time is a black hole interior, part is a time-reversed black hole called a white hole, and part is a second exterior region. Eternal black holes have "wormholes" that connect the exterior regions, and which might allow time travel. Real black holes, formed by the collapse of matter in the past, do not have such regions, and there is no reason to expect that eternal black holes exist in nature.

Quantum effects change the above classical view of a black hole. In particular J. Bekenstein and Hawking (and many others) have shown that black holes have an intrinsic temperature, and radiate energy like a blackbody at that temperature. Thus they are not truly black and are not totally stable, but can evaporate away (as discussed below). *See* BLACK HOLE; GRAVITATIONAL COLLAPSE; HEAT RADIATION.

Cosmology. Observations of the last few decades have revolutionized the field of cosmology and put general relativity in the forefront of fundamental physics. But they have also raised many questions, and made it clear that we do not understand the dominant constituents of the universe.

The geometric viewpoint of relativity provides an elegant description of the universe on the cosmological scale of billions of light years. On this scale the visible material of the universe appears to be spatially isotropic and homogeneous; that is, it is substantially the same in all directions and at all points in space. (On a scale of less than 100 million light years there is much clumping of matter into galaxies, clusters of galaxies, and "great walls" and filaments of galactic clusters.) It is thus natural to assume that the large-scale geometry of the universe is isotropic and homogeneous. *See* GALAXY, EXTERNAL; UNIVERSE.

Consistent with this assumption there are essentially three kinds of geometry possible for the three-dimensional space of the universe (assuming a simple topology). These are illustrated by two-dimensional analogs, which are the surfaces of a plane, a sphere, and a saddle (**Fig. 6**). They are labeled with a curvature parameter k and called flat ($k = 0$), positively curved ($k = 1$), or negatively curved ($k = -1$). The negatively curved space is not really possible to visualize, but is everywhere analogous to the center of the two-dimensional saddle. Of these three geometries only the positively curved one is finite and closed, the others being open and having infinite extent.

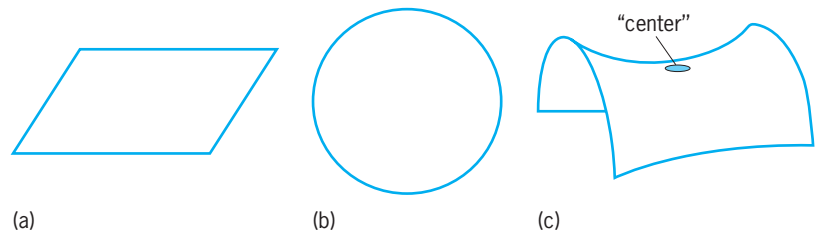

Fig. 6. Two-dimensional analogs of the three-dimensional space of the universe. All are homogeneous and isotropic. (*a*) Flat space ($k = 0$). (*b*) Positively curved space ($k = 1$). (*c*) Negatively curved space ($k = -1$). This space is like the circled center region everywhere.

Expansion of the universe may be visualized by picturing grid lines on the analog surfaces, with galaxies considered as particles of the cosmic fluid attached at the intersections. These move apart as the surfaces expand. In the case of the sphere the radius becomes larger with expansion. This picture emphasizes the role of coordinates in relativity as only markers of position: the galaxies remain at rest in the coordinate system but move apart physically as the universe expands.

The metric that most conveniently describes these geometries is called the FRW metric, after A. Friedmann (who first used it in 1922), H. P. Robertson and A. G. Walker; it is given in Eq. (19)

$$ds^2 = c^2dt^2 - a(t)^2 \times [dr^2/(1 - kr^2) - r^2(d\theta^2 + \sin^2\theta d\phi^2)] \quad (19)$$

in terms of spherical coordinates. The scale function $a(t)$ has the dimension of length, r is a dimensionless radial coordinate, and t is a universal time.

A primary task of theoretical cosmology is to determine the scale function, but the curvature parameter k must be determined from observation. If the FRW metric in Eq. (19) is substituted into the Einstein field equations (17), two equations for the scale function result, often called the cosmological equations (20);

$$-\left(\frac{8\pi G}{c^4}\right)\rho = \Lambda - 3\left[\frac{k}{a^2} + \frac{\dot{a}^2}{c^2a^2}\right],$$
$$\left(\frac{8\pi G}{c^4}\right)p = \Lambda - \left[\frac{k}{a^2} + \frac{\dot{a}^2}{c^2a^2} + \frac{2\ddot{a}}{c^2a}\right] \quad (20)$$

ρ denotes the density of the material of the universe, considered to be a fluid, p denotes its pressure, and the dot denotes a time derivative.

These equations must be supplemented by a relation between the density and pressure called the equation of state. The equation of state is often assumed to be a linear relation $p = w\rho$, with constant w; $w = 1/3$ corresponds to radiation (or hot matter in rapid thermal motion); $w = 0$ corresponds to cold

matter (with negligible pressure and slow thermal motion); and the rather curious case of $w = -1$ corresponds to vacuum or dark energy, which is equivalent to a cosmological constant.

Various assumptions about the nature of the cosmic fluid and the values of Λ and k may be made, and the cosmological equations solved accordingly. The favored combination for the present era is that the universe is spatially flat ($k = 0$), Λ is nonzero, and the fluid is cold matter with negligible pressure; this is the "concordance model" discussed below. Then the solution for the scale function is given in Eq. (21).

$$a(t) = A \sinh^{2/3}\left((\sqrt{3\Lambda}/2)ct\right) \qquad A = \text{arbitrary constant} \quad (21)$$

For early times the scale function is very small, and for late times it approaches accelerated exponential expansion. However, this solution is not valid for very early times when the universe contained radiation and high-density material.

Observational cosmology relates the scale function $a(t)$ and the parameters k and Λ and ω to the real universe. Hubble's law for the observed recession velocity of galaxies is $v = HL$, where L is the distance to the galaxy and H is the "Hubble constant"; H is actually a parameter that varies slowly with time and is related to the scale function by $H = \dot{a}/a$. Its measured value at present is $H = 71 \pm 4$ km/(s Mpc), where 1 megaparsec (Mpc) = 3.3 million light years 3.1×10^{19} km, in the units favored by astronomers; H^{-1} is a rough measure of the lifetime of the universe, and is 14×10^9 years. It is clear from Hubble's law that the material of the universe must have been very dense and hot about 14×10^9 years ago. *See* HUBBLE CONSTANT.

The curvature parameter k is related to the total energy density of the universe, including the energy of the vacuum: if the density is greater than a critical value given by $\rho_c = 3H^2c^2/8\pi G \cong 8.3 \times 10^{-10}$ J/m^3, $\cong 5.2$ GeV/m^3, then the universe is positively curved ($k = 1$); if it is equal to critical, then the universe is spatially flat ($k = 0$); if it is less than critical, then the universe is negatively curved ($k = -1$). Diverse measurements of the density, deviations from linearity in the Hubble expansion from recent supernova data, and measurements of the cosmic background radiation or cosmic microwave background from the *Wilkinson Microwave Anisotropy Probe (WMAP)* have led astronomers to the present concordance model. This is spatially flat ($k = 0$), with about 73% of the cosmic energy being vacuum, about 23% cold dark matter of unknown nature, and only about 4% ordinary baryonic matter. The scale function is that in Eq. (21). *See* COSMIC BACKGROUND RADIATION; WILKINSON MICROWAVE ANISOTROPY PROBE.

The nature of the dark matter and the vacuum energy, now generally called dark energy, are both unknown. One leading candidate for the dark matter is the neutralino, a leptonic partner of the photon suggested by supersymmetry theory (SUSY), but never observed in the laboratory. It is one type of WIMP (weakly interacting massive particle). There has as yet been no direct detection of dark matter particles. There are many speculative theories of the dark energy, which acts like vacuum energy or the cosmological constant. For example, it may be a slowly varying self-interacting scalar field that permeates the universe, analogous to the inflaton field of inflationary cosmology theory (discussed below). None of the theories of dark energy is generally accepted. *See* INFLATIONARY UNIVERSE COSMOLOGY; SUPERSYMMETRY; WEAKLY INTERACTING MASSIVE PARTICLE (WIMP).

In summary, cosmology is now in a somewhat paradoxical state. The general properties of the present universe are understood rather well, in terms of the concordance model of cosmology, but almost nothing is known about its main ingredients, the dark matter and dark energy.

Early universe. According to observation and general relativity, the universe began very small and very hot with a big bang, and expanded to become larger and cooler. While the present universe is dominated by dark energy and dark matter, the presence of the cosmic microwave background indicates that radiation and a hot plasma of electrons and nuclei dominated at earlier times, before about 100,000 years after the big bang, when the scale function was about 1,000 times smaller than at present. The cosmic microwave background is itself the cooled residual radiation from that time; it has been moving largely decoupled from interaction with matter since about 300,000 years after the big bang, when the electrons and nuclei in the previously existing plasma combined to form transparent neutral atoms. *See* BIG BANG THEORY.

The radiation era was in fact cosmically fairly simple, with a scale function proportional to $t^{1/2}$, and is fairly well understood theoretically. Rather early in the radiation era, at a minute or so, the nuclei of the light elements like helium and lithium formed, essentially by condensing out of a previously existing plasma of nucleons and electrons. The theory of this nucleosynthesis is quite successful at explaining the presently observed abundance of the light elements (70% hydrogen, 28% helium, 2% other nuclei), which gives us confidence in our understanding of the cosmology of the radiation era. *See* NUCLEOSYNTHESIS.

Most cosmologists believe that before the radiation era the universe went through a period of exponentially rapid expansion, or inflation, by a large factor of about e^{60} (10^{26}), during which time it was dominated by material that behaved much like a large cosmological constant, often called the inflaton field. Inflation is used to explain the present homogeneity of the universe and the isotropy of the cosmic microwave background. It also provides a possible explanation of the flatness (or near flatness) of the present universe ($k = 0$), the small fluctuations of about 10^{-5} in the temperature of the cosmic microwave background, and the clumpiness of the present universe in the form of galaxies and galactic clusters.

General relativity and quantum theory. In general relativity the space-time continuum is dynamical

and is the dominant entity. But in quantum theories such as quantum electrodynamics and quantum chromodynamics it serves as the unchanging arena in which objects interact. Moreover, in general relativity it is implicitly assumed that space-time distances can be measured to any desired degree of accuracy, whereas in quantum theory observables such as the position of an electron generally involve an intrinsic uncertainty. Thus it is not surprising that combining general relativity and quantum theory is difficult, and that there is as yet no generally accepted quantum theory of gravity and space-time. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM CHROMODYNAMICS; QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; QUANTUM MECHANICS; RELATIVISTIC QUANTUM THEORY; STANDARD MODEL; UNCERTAINTY PRINCIPLE.

Indeed it is clear that the concept of a space-time continuum as used in general relativity must break down at a sufficiently small distance scale; this follows heuristically from general ideas of quantum theory. To measure a distance L an observer would use light of a wavelength λ smaller than L, and the measurement would have an intrinsic uncertainty of about λ (according to general principles of optics). But light has energy that creates a gravitational field by distorting space-time in the region being measured, which leads to a further uncertainty in the measurement. The minimum energy of the light is that of one photon $E = hc/\lambda$ (where h is Planck's constant), the effective photon mass is $M = E/c^2 = h/c\lambda$, and the fractional distortion of the space-time is very roughly $GM/c^2L \approx Gh/c^3\lambda L$, which may be inferred from the Schwarzschild metric of Eq. (18). Thus the total uncertainty is about $\lambda + (G\hbar/c^3)/\lambda$. (At the desired level of approximation it is permissible to ignore factors like 2π and to make no distinction between h and $\hbar = h/2\pi$.) This expression has a minimum value $L_p = (G\hbar/c^3)^{1/2}$, which is called the Planck length. The Planck length is some 20 orders of magnitude smaller than distances presently measured in high-energy physics experiments. The classical concept of distance is widely believed to lose meaning at the Planck scale, and perhaps the intrinsic idea of distance itself does so as well. *See* PLANCK'S CONSTANT.

Max Planck recognized in 1899 that a basic scale of mass, length, and time is set by the constants G, c, and $\hbar$; these combine uniquely to give the Planck scale quantities in Eqs. (22). To the quantities in

$$L_P = \left(\frac{G\hbar}{c^3}\right)^{1/2} = 1.61 \times 10^{-35} \text{ m (Planck length)}$$

$$T_P = \left(\frac{G\hbar}{c^5}\right)^{1/2} = 5.37 \times 10^{-44} \text{ s (Planck time)}$$

$$M_P = \left(\frac{\hbar c}{G}\right)^{1/2} = 2.17 \times 10^{-8} \text{ kg (Planck mass)} \tag{22}$$

Eq. (22) a Planck energy may be added, defined by $E_p = M_pc^2 = (\hbar c^5/G)^{1/2} = 1.22 \times 10^{19}$ GeV in units appropriate to high-energy physics.

While gravitational theory needs to include quantum effects at the Planck scale the converse is also true, that quantum theories such as quantum electrodynamics and quantum chromodynamics need to include gravitational effects at the Planck scale. This is because the gravitational interaction becomes roughly comparable in strength to the electromagnetic and strong (quantum chromodynamics) forces at the Planck scale. The electric force between two particles with charges e is given classically by $F_{el} = e^2/r^2 = \hbar c\alpha/r^2$, where $\alpha = e^2/\hbar c \cong 1/137$ is the fine-structure constant. The gravitational force between two particles of energy E is given roughly by $F_g = G(E/C^2)^2/r^2$, so the two forces become equal at about $E \approx \sqrt{\hbar c^5/G}\sqrt{\alpha}$. To about an order of magnitude $\sqrt{\alpha} \approx 1$, so this energy is approximately the Planck energy. Thus, to describe physics at the Planck scale (the realm of small distances, short times, and high energy), a unified quantum treatment of gravity and space-time is required. Two such realms are the behavior of the hot dense material of the early universe, and the behavior of virtual particles of arbitrarily high energy in quantum field theory. Of course, laboratory experiments or astronomical observations at the Planck scale are difficult to even imagine. *See* GRAVITY.

Despite the lack of an accepted quantum theory of gravity, some interesting results have been obtained involving gravity and the quantum. The simplest way to combine ideas of gravity and quanta is to assume that quantum field theory is approximately valid in a classical curved background space-time obtained from the field equations of general relativity. The best-known example of this involves the thermodynamics of black holes. Bekenstein showed that for consistency in the interaction of radiation with black holes a black hole should have an entropy proportional to its area. This appeared to be inconsistent with the black hole property of absorbing radiation but not emitting it; but Hawking applied quantum field theory in the curved space-time of a black hole and showed that a black hole should indeed emit radiation (from very near its surface) as if it were a black body with a temperature given (in units of energy) by $T_{BH} = \hbar c/8\pi GM$. Hawking radiation makes black hole thermodynamics consistent, and is widely accepted by theorists, although it has as yet not been verified by observation.

However, it is not consistent to have a classical gravitational field coexisting with other quantized force fields. The uncertainty principle could then be subverted, in principle, by gravitational measurements. One approach to a true quantum theory of gravity is to ignore the geometric interpretation of gravity and develop the theory in analogy with quantum electrodynamics and other successful quantum field theories. To do this the metric tensor is written as the sum of the Lorentz metric of flat space-time plus a quantity $h_{\mu\nu}$ that describes the gravitational field in a flat background space-time. Then $h_{\mu\nu}$ may be quantized in the standard way, and Feynman diagrams and amplitude rules obtained for a perturbative quantum gravity theory. The perturbative

theory is successful until diagrams with internal loops are encountered. Divergent integrals associated with these diagrams cannot be reinterpreted and removed as they are in quantum electrodynamics; that is, the theory is not renormalizable, or involves an infinite number of renormalization constants, depending on the viewpoint of the theorist. *See* RENORMALIZATION.

The cosmological constant problem, mentioned previously, has perplexed theorists for decades. It is the result of combining gravity and quantum ideas, although its resolution may not require either a theory of quantum gravity or the Planck scale. In quantum field theory the vacuum is the lowest energy state of the field, which is not in general zero. Indeed it is generally given by a divergent integral, which cannot be removed by renormalization like other divergent integrals; the energy of the vacuum does have observable gravitational effects, equivalent to a cosmological constant as already discussed. If the upper limit of the divergent integral is taken to be the Planck energy instead of infinity, which is quite reasonable, the energy of the vacuum turns out to be approximately the Planck density, E_p/L_p^3, which is to be expected on dimensional grounds. But this has the nonsensical value 10^{124} GeV/m^3, which is about 124 orders of magnitude larger than allowed by observation. There has been no convincing resolution of this paradox, which is considered by many to be the most troublesome problem in theoretical physics.

Many interesting attempts have been made to develop a complete quantum gravity theory. The best known is superstring theory, which purports to describe all particles and all forces, including gravity, in terms of one-dimensional segments of string in a space of 10 or more dimensions. Another popular approach is called loop quantum gravity, and recasts the variables and mathematical structure of general relativity in a way that is more amenable to quantization. Neither of these theories has produced definitive predictions that may be tested by experiment. More drastic theoretical ideas involve constructing space-time in a way that is consistent with quantum theory, from so-called spin networks or from a quantum fluid ether. Probably the most drastic idea is that space-time as a fundamental entity simply does not exist, and only emerges as an approximation at scales well above the Planck scale. *See* QUANTUM GRAVITATION.

Ronald J. Adler

Bibliography. R. J. Adler, M. Bazin, and M. M. Schiffer, *Introduction to General Relativity*, McGraw-Hill, 1975; R. J. Adler, Gravity, Chap. 3 in G. Fraser (ed.), *The New Physics for the Twenty-first Century*, Cambridge, 2005 (in press); H. Bondi, *Relativity and Common Sense*, 1964, reprint, Dover, 1980; A. Einstein, *Relativity: The Special and the General Theory: The Masterpiece Edition*, Pi Press, 2005; N. D. Mermin, *Space and Time in Special Relativity*, 1968, reprint, Waveland Press, 1989; C. Misner, K. Thorne, and J. Wheeler, *Gravitation*, W. H. Freeman, 1973; B. Schutz, *A First Course in General Relativity*, Cambridge, 1985; C. M. Will, *Theory and Experiment in Gravitational Physics*, Cambridge, 1993; P. J. E. Peebles, *Principles of Physical Cosmology*, Princeton, 1993; E. W. Kolb and M. S. Turner, *The Early Universe*, Westview, 1993.

Relaxation time of electrons

The characteristic time for a distribution of electrons in a solid to approach or "relax" to equilibrium after a disturbance is removed. A familiar example is the property of electrical conductivity, in which an applied electric field generates an electron current which relaxes to an equilibrium zero current after the field is turned off. The conductivity of a material is directly proportional to this relaxation time; highly conductive materials have relatively long relaxation times. The closely related concept of a lifetime is the mean time that an electron will reside in a given quantum state before changing state as a result of collision with another particle or intrinsic excitation. This lifetime is related to equilibrium properties of the material, whereas the relaxation time relates to the thermal and electrical transport properties. The average distance that an electron travels before a collision is called the mean free path. Although typical collision times in metals are quite short (on the order of 10^{-14} s at room temperature), mean free paths range from about 100 atomic distances at room temperature to 10^6 atomic distances in pure metals near absolute zero temperature. Considering the very dense packing of atoms in a solid, these surprisingly long electron path lengths are analogous to the unlikely event that a rifle bullet might travel for miles through a dense forest without hitting a tree. The detailed explanation of the electron mean free path in metals is a major success of the modern theory of solids.

Transport property. A relaxation time appears in the simplest expression for the transport property of electrical conductivity, which states that the electrical conductivity equals the product of the relaxation time, the density of conduction electrons, and the square of the electron charge, divided by the electron effective mass in the solid. *See* BAND THEORY OF SOLIDS; ELECTRICAL CONDUCTIVITY OF METALS.

The conduction process is a steady-state balance between the accelerating force of an electric field and the decelerating friction of electron collisions which occur on the time scale of the relaxation time. This process may be described in terms of the probability distribution function for the electrons, which depends on the electron momentum (proportional to the wave vector, **k**, which labels the quantum state of the electrons), the position, and the time. Viewed in **k**-space, the entire distribution will shift from an equilibrium state under the influence of a perturbation such as an electric field (**Fig. 1**). For example, in the ground state, the collection of occupied electron states in **k**-space is bounded by the Fermi surface centered at the origin, while in an electric field this region is shifted. Because of electron collisions with impurities, lattice imperfections, and vibrations (also called phonons), the displaced surface may

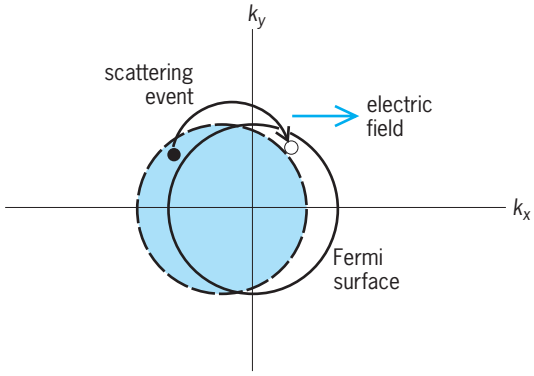

Fig. 1. Effect of the electric field on electron distribution in a solid, viewed in k-space, where k is the electron wave vector. The shaded area indicates occupied states in distribution which result when the field is applied.

be maintained in a steady state in an electric field. These collisions also restore the equilibrium distribution after the field is turned off, and the relaxation time is determined by the rate at which the shifted distribution returns to equilibrium. Specifically, the contribution of collisions to the rate of change of the shifted distribution after the field is turned off equals the difference between the shifted and equilibrium distributions divided by the relaxation time. This statement is referred to as the relaxation-time approximation, and is a simple way of expressing the role of collisions in the maintenance of thermodynamic equilibrium. The details of the various collision mechanisms are lumped into the parameter of the relaxation time. For example, in the case of mixed scattering by impurities and phonons, the inverse of the relaxation time can be determined from the sum of collision rates as the sum of the inverses of the electron-impurity and electron-phonon scattering times. *See* CRYSTAL DEFECTS; FERMI SURFACE; LATTICE VIBRATIONS; PHONON.

Constructive interference. In pure metals at low temperatures, the long mean free path of conduction electrons results from their large velocity (on the order of 10^6 m/s near the Fermi surface) and relatively long relaxation time, on the order of 10^{-9} s. From a practical standpoint, this is what makes metals useful as electrical conductors even at room temperature, where a relaxation time on the order of 10^{-14} s and a mean free path (equal to the product of the velocity and the relaxation time) on the order of 10^{-8} m (or about 100 atomic distances) is typical.

As compared to poorly conducting solids or insulators, an excess of so-called free electrons in a metal contributes to the long mean free paths. From a quantum mechanical standpoint, the free-electron wave function readjusts in a perfectly periodic atomic lattice to avoid the atomic ion cores and spend most of the time in the spaces between. In the analogy of the rifle fired into a forest, the bullet will not travel far, but the sound of the gunshot can, because the sound waves bend around the trees in a constructive manner. In a perfectly periodic lattice, the electron waves scatter constructively from the atomic ion cores, resulting in screening of the cores and coherent transmission of the waves over large distances. Any disturbance to the lattice periodicity tends to destroy this wave phenomenon, resulting in lower transmission or energy loss. At room temperature, the conductivity is usually limited by scattering from lattice vibrations (phonons). At lower temperatures, vibrations are greatly reduced, but the conductivity is still limited by scattering from impurities and imperfections. *See* ELECTRICAL RESISTIVITY; FREE-ELECTRON THEORY OF METALS; QUANTUM MECHANICS.

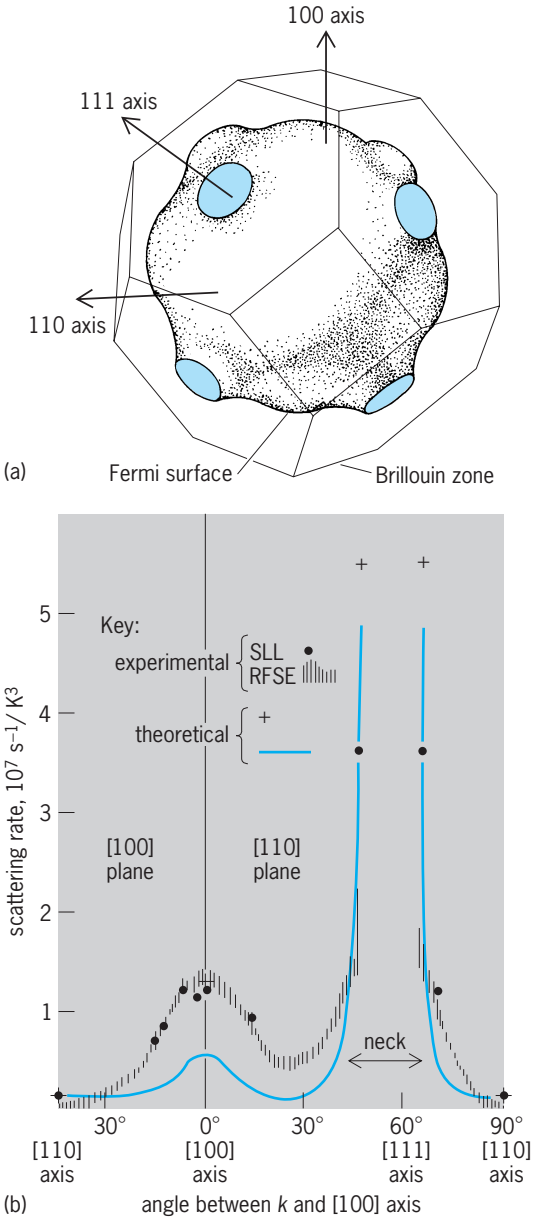

Fig. 2. Scattering of conduction electrons from lattice vibrations in copper. (*a*) Fermi surface and Brillouin zone of copper. (*b*) Dependence of the scattering coefficient on the direction of the conduction electron wave vector, according to theoretical calculations and measurements using the radio-frequency size effect (RFSE) and the cyclotron resonance of surface Landau levels (SLL). (*After J. F. Koch and R. E. Dozema, The anisotropy of thermal scattering on the Fermi surface of metals, in Proceedings of the 14th International Conference on Low Temperature Physics, pp. 314–338, 1975*)

In the superconductive state, observed in many metals and certain complex compounds at sufficiently low temperatures, impurities, imperfections, and lattice vibrations become completely ineffective in retarding current flow, leading to persistent electric currents even after the driving electric field is removed. This state of resistanceless conduction is described as an ordered quantum state of pairs of electrons resulting from lattice vibrations which deform the ion core potential in such a way as to provide an attractive electron-electron interaction. The relaxation time thus becomes infinite (as nearly as can be measured), a rare macroscopic manifestation of a quantum effect. *See* SUPERCONDUCTIVITY.

Local measurements. A direct measurement of electron lifetimes for specific electron states in **k**-space is possible. The techniques work best at low temperatures and generally involve some kind of resonance in space (for example, radio-frequency size effect), time (cyclotron resonance), or energy (de Haas–van Alphen effect). The richness of detail made possible by such measurements has greatly aided the understanding of the interactions between conduction electrons, impurities, lattice imperfections, and vibrations. *See* CYCLOTRON RESONANCE EXPERIMENTS; DE HAAS-VAN ALPHEN EFFECT.

For example, measurements of conduction electron scattering from lattice vibrations reveal that the strongest interaction occurs when the electron's wave vector (proportional to the inverse wavelength) is nearly commensurate with the lattice periodicity (**Fig. 2**). For example, the wave vectors of the conduction electrons in copper lie near the Fermi surface (Fig. 2*a*), and the scattering is strongest for wave vectors on the Fermi surface in the vicinity of the [111] axis (Fig. 2*b*). In this direction, the electron wavelength coincides with the lattice periodicity at points where the Fermi surface touches the Brillouin zone (Fig. 2*a*). *See* BRILLOUIN ZONE; CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Measurements of impurity scattering have shown that the symmetry character of both the impurity and the host are important in scattering. The scattering is strongest along directions where the electronic band structure symmetry is most nearly similar and where a "resonant" scattering is most nearly achieved. Thus studies of impurity scattering rates allow the symmetry of the scattering to be deduced. For example, the scattering of copper conduction electrons from nickel impurity atoms is largest near the "belly" region of the Fermi surface on the [110] axis (Fig. 2*a*). In this direction, the symmetry of the conduction electron wave function is very *d*−like, establishing that the scattering is likely due to resonance with the nickel *d*−band.

Measurements of conduction electron scattering from lattice dislocations (a disturbance of lattice periodicity which permits metal deformability) show that the measured collision lifetime differs by more than three orders of magnitude, depending upon the extent to which a particular type of measurement is sensitive to the small-angle scattering. For a dislocation "forest" in which 1 out of every 10,000 atoms in a row is out of place, the scattering rate is about 10^{11} s^{-1} in the de Haas–van Alphen effect, but only 10^{8} s^{-1} in resistivity, and intermediate in the radio-frequency size effect. These differences arise because small-angle scattering is a small perturbation to current flow, whereas for quantum-effect measurements even a tiny scattering event can destroy constructive interference of the incident and scattered electron waves.

Energy relaxation. At higher temperatures and in disordered or polycrystalline metals, such refined **k**-space measurements are difficult, if not impossible, to perform. Even transport measurements of relaxation times are complicated by mixed scattering from defects, phonons, and so forth. However, direct time-resolved measurements of inelastic electron scattering are possible by means of ultrashort-duration (10^{-14} s to 10^{-12} s) laser pulses. In simple metals, a rapid increase in the electron temperature is produced by absorption of a laser pulse whose duration is shorter than the time required for hot electrons to equilibrate with the surrounding lattice temperature. Various methods are used to monitor the relaxation of this nonequilibrium temperature differential, including photoemission of the hot electrons into energy analyzers, and optical measurements of the heating-induced reflectivity change. In such measurements, a sequentially time-delayed probing laser pulse is used to sample the electron temperature or energy distribution. *See* OPTICAL PULSES; PHOTOEMISSION; REFLECTION OF ELECTROMAGNETIC RADIATION.

Not only can the effect of inelastic electron-phonon collisions be observed, but the early evolution of a nonequilibrium electron energy distribution can also be resolved by time-resolved photoemission (**Fig. 3**). Soon after heating of the conduction electrons by a laser pulse, the nonequilibrium nature of the electron system is evidenced by a high-energy distribution extending beyond the best-fit equilibrium Fermi distribution. The energy distribution then relaxes back to an equilibrium electronic temperature via electron-electron and electron-phonon

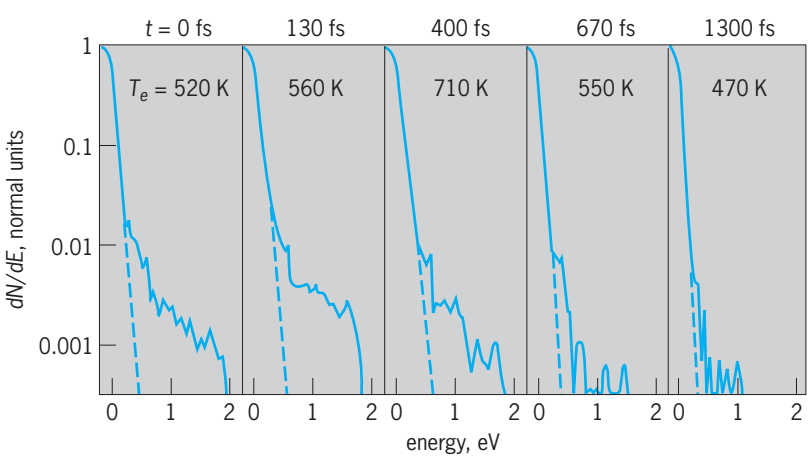

Fig. 3. Electron energy distribution function (density of states) versus energy at five time delays after absorption of an 180-femtosecond laser pulse by a gold film. The broken line is the best Fermi distribution fit at the electron temperature shown. (*After W. S. Fann et al., Electron thermalization in gold, Phys. Rev. B, 46(20):13592–13595, 1992*)

scattering. Subsequent electron-phonon scattering cools the electron distribution to the lower lattice temperature. When analyzed by suitable models, time-resolved optical measurements can yield values for both the electron-electron and electron-phonon energy relaxation rate parameters.

Semiconductor measurements. A measure of the relaxation time may be extracted from the expression for the conductivity stated above by combining the conductivity with a measurement of the Hall coefficient. It turns out that the product of the conductivity, the Hall coefficient, and the speed of light equals the mobility, which is a measure of the speed of response and power dissipation in semiconductor devices, and is thus an important figure of merit for evaluating the performance of integrated circuits. The mobility, in turn, equals the product of the relaxation time and the ratio of the electron charge to effective mass. *See* HALL EFFECT; INTEGRATED CIRCUITS.

The behavior of semiconductors is strongly influenced by another lifetime, the minority-carrier recombination time, which is the mean time an electron can survive before recombining with a hole in *p*-type material (or vice versa). Recombination times range from 10^{-4} to 10^{-9} s, many orders of magnitude longer than the transport relaxation time discussed above. They are measured by methods such as photocurrent decay, the characteristic time for the excess current created by a light pulse to decay. The dynamics of minority carrier transport may be studied in some detail by injecting excess carriers into a material with an applied electric field and observing the decay of the current pulse downstream. Ultrafast optical methods are also applied to the measurement of a variety of relaxation processes in semiconductors. This is especially relevant to gallium arsenide (GaAs) and related materials, which offer the prospect of very high speed devices for communications and computers. *See* MICROWAVE SOLID-STATE DEVICES.

The key to inventing the junction (or bipolar) transistor was fabricating a base region thin enough that a minority carrier injected from the adjacent emitter could survive recombination long enough to reach the collector. Recombination also limits solar-cell efficiency, since the electron-hole pairs must survive long enough to contribute to the current at the cell terminals. Recombination centers or traps may be impurities, lattice imperfections, or surface states. A clean silicon surface contains one dangling bond per atom which acts as a trap. Fabrication of metal-oxide-semiconductor (MOS) transistors and integrated circuits includes passivation of these surface states during oxide formation. An internal version of surface states is the dangling bond on lattice imperfections such as grain boundaries and dislocations. This is why semiconductor devices are ordinarily fabricated on single crystals of high perfection. The use of amorphous or polycrystalline materials, particularly in large-area solar cells, depends on methods that neutralize the trapping action of dangling bonds and thereby increase minority-carrier recombination time. *See* ELECTRON-HOLE RECOMBINATION; GRAIN BOUNDARIES; JUNCTION TRANSISTOR; SEMICONDUCTOR; SOLAR CELL; TRANSISTOR; TRAPS IN SOLIDS.

Gary L. Eesley

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics,* 1976; M. Ya Azbel', M. I. Kaganov, and I. M. Lifshitz, Conduction electrons in metals, *Sci. Amer.*, 228(1):88–98, 1973; R. H. Bube, *Electrons in Solids: An Introductory Survey*, 3d ed., 1992; M. L. Cohen and J. R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors*, 2d ed., 1989; G. Grimvall, *The Electron-Phonon Interaction in Metals*, 1981; S. M. Sze, *Physics of Semiconductor Devices*, 2d ed., 1981.

Relay

An electromechanical, solid-state, or digital device operated by variations in the input that, in turn, operate or control other devices connected to the output. Relays are used in a wide variety of applications throughout industry, such as in telephone exchanges, digital computers, motor and sequencing controls, and automation systems. Highly sophisticated relays are utilized to protect electric power systems against trouble and power blackouts as well as to regulate and control the generation and distribution of power. In the home, relays are used in refrigerators, automatic washing machines and dishwashers, and heating and air-conditioning controls. Although relays are generally associated with electrical circuitry, there are many other types, such as pneumatic and hydraulic. Input may be electrical and output directly mechanical, or vice versa. *See* ELECTRIC PROTECTIVE DEVICES; SWITCHING SYSTEMS (COMMUNICATIONS).

Relays using discrete solid-state components, operational amplifiers, or microprocessors can provide more sophisticated designs. For industrial, commercial, or residential control, they are used particularly in applications where the relay and associated equipment are packaged together. The basic operation may be complex, but frequently it is similar or equivalent to the units described. For industrial or utility power systems they are used separately in conjunction with appropriate switching devices. *See* AMPLIFIER; MICROPROCESSOR; OPERATIONAL AMPLIFIER.

Basic classifications. A basic, simplified block diagram of a relay is shown in **Fig. 1**. The actual physical embodiment of the blocks varies widely, and one or more blocks may not exist or may be combined in practice. Classifying electrical-type relays by function and somewhat in the order of increasing complexity, there are (1) auxiliary relays, (2) monitoring relays, (3) regulating relays, (4) reclosing, synchronism check, and synchronizing relays, and (5) protective relays.

While the application and protection principles are the same, the relay units may be electromechanical, solid-state, or based on digital microprocessors. Originally, all relays were electromechanical. Electromechanical relays are still used in many

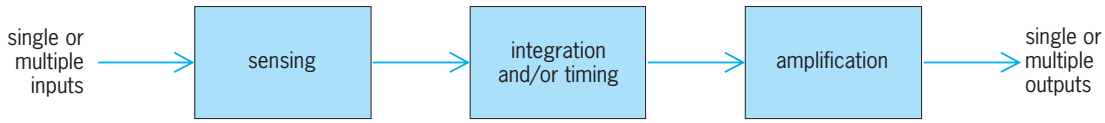

Fig. 1. Basic block diagram of a relay.

applications because of their relative simplicity, long life, and proven high reliability. However, in the late 1960s, solid-state components such as transistors and diodes began to be used. In this rapidly changing technology, operational amplifiers were used briefly. Relays based on digital technology are now preferred, particularly for protective relays, and are replacing the electromechanical or solid-state relays as existing applications or facilities are modified or new ones are installed.

Electromechanical relays. Relays of this type can be illustrated by several common basic units. An auxiliary relay (clapper or telephone type) is shown in **Fig. 2***a*. The sensing unit is the electric coil; when the applied current or voltage exceeds a threshold value, the coil activates the armature, which operates either to close the open contacts or to open the closed contacts. Relays of this type are used as contact multipliers, and to provide the necessary contact outputs for operating the trip coils of circuit breakers for solid-state digital-type relays. Other applications are to isolate circuits or to use a low-power input to control higher-power input.

The plunger type (Fig. 2*b*) operates on the principle of electromagnetic attraction. This relay is composed of a coil, plunger, and set of contacts. When current I flows in the coil, a force is produced that causes the plunger to move and close the relay contacts. Both the clapper and plunger types can be designed to operate on either direct or alternating current. *See* ELECTROMAGNET; SOLENOID (ELECTRICITY).

The balanced polar unit (Fig. 2*c*) is used as a monitoring relay. The unbalanced type is shown in Fig. 2*d*. A permanent magnet normally polarizes the pole faces on either side of the armature, and direct current is applied to the coil that is around the armature. At nominal values of current, the armature can be adjusted to float in the air gap between the pole pieces by the magnetic shunts, but when this current decreases or increases significantly the armature is magnetized and moves to one of the pole pieces to close the contacts. Thus this relay monitors the current level in an electric circuit. The sensing, integrating, and amplifying functions indicated in the block diagram of Fig. 1 are combined in the single mechanical element and its contacts for this polar unit.

The relay unit known as the induction-disk type (Fig. 2*e*) is widely used in regulating and protective relays. With alternating current or voltage applied to the main coil, transformer action induces current in the secondary circuit connected to the upper poles. Fluxes produced by the currents flowing in the upper pole circuit induce eddy currents in the rotor disk. Interaction between rotor eddy currents and the flux from the lower pole produces mechanical torque to rotate the disk, which is of the order of 3 in. (8 cm) in diameter, thus closing the contacts. This is the split-phase motor principle, where two out-of-phase fluxes produce torque in a rotor. The upper pole may be supplied from another source to permit comparison of two quantities. A spring automatically resets the disk after the relay has operated. By use of the principles of electromagnetic attraction and electromagnetic induction, protective relays can be built to respond to all abnormal conditions that may occur. *See* EDDY CURRENT; ELECTROMAGNETIC INDUCTION; INDUCTION MOTOR; TRANSFORMER.

By design of the coils, damping magnet, and spring, a variety of inverse time overcurrent and under- or overvoltage characteristics are produced. With reference to Fig. 1, the sensing unit is approximated by the coil, the armature and its gap by the timing unit, and the contacts by the amplifier.

For regulating, the output is arranged to adjust or regulate the input within prescribed limits; for monitoring, the output is used to alarm or to shut down the input equipment.

Reclosing, synchronism check, and synchronizing relays provide an automatic sequence of operations. Good examples of these are the timer control relays on automatic washing machines and dishwashers, which provide the necessary program in the various washing operations.

Another widely used basic relay unit is the cylinder or cup type (Fig. 2*f*). The operation is similar to a two-phase induction motor. With out-of-phase currents, voltages, or current and voltage applied to the two windings, flux is produced to rotate the thin-walled rotor in the form of a cylinder or cup. Contacts are attached to the rotor. The contacts and stops permit only a very small movement of the rotor, just enough to provide opening or closing of the contacts. This unit is used in directional sensing and in distance-type relays.

Solid-state relays. The growing complexity of industrial control and modern power systems has resulted in a need for relays, and particularly protective relays, with a higher level of performance and more sophisticated characteristics. This has been made possible by the development of semiconductors and other associated components, which can be utilized in relay designs. These are referred to as solid-state or static relays. All of the functions and characteristics of the electromechanical relays can be performed by solid-state devices, either as discrete components or integrated circuits. They may be divided into two categories: analog circuits that are either fault-sensing or measuring circuits, and digital logic circuits for operation on logical variables. *See* INTEGRATED CIRCUITS; LOGIC CIRCUITS.

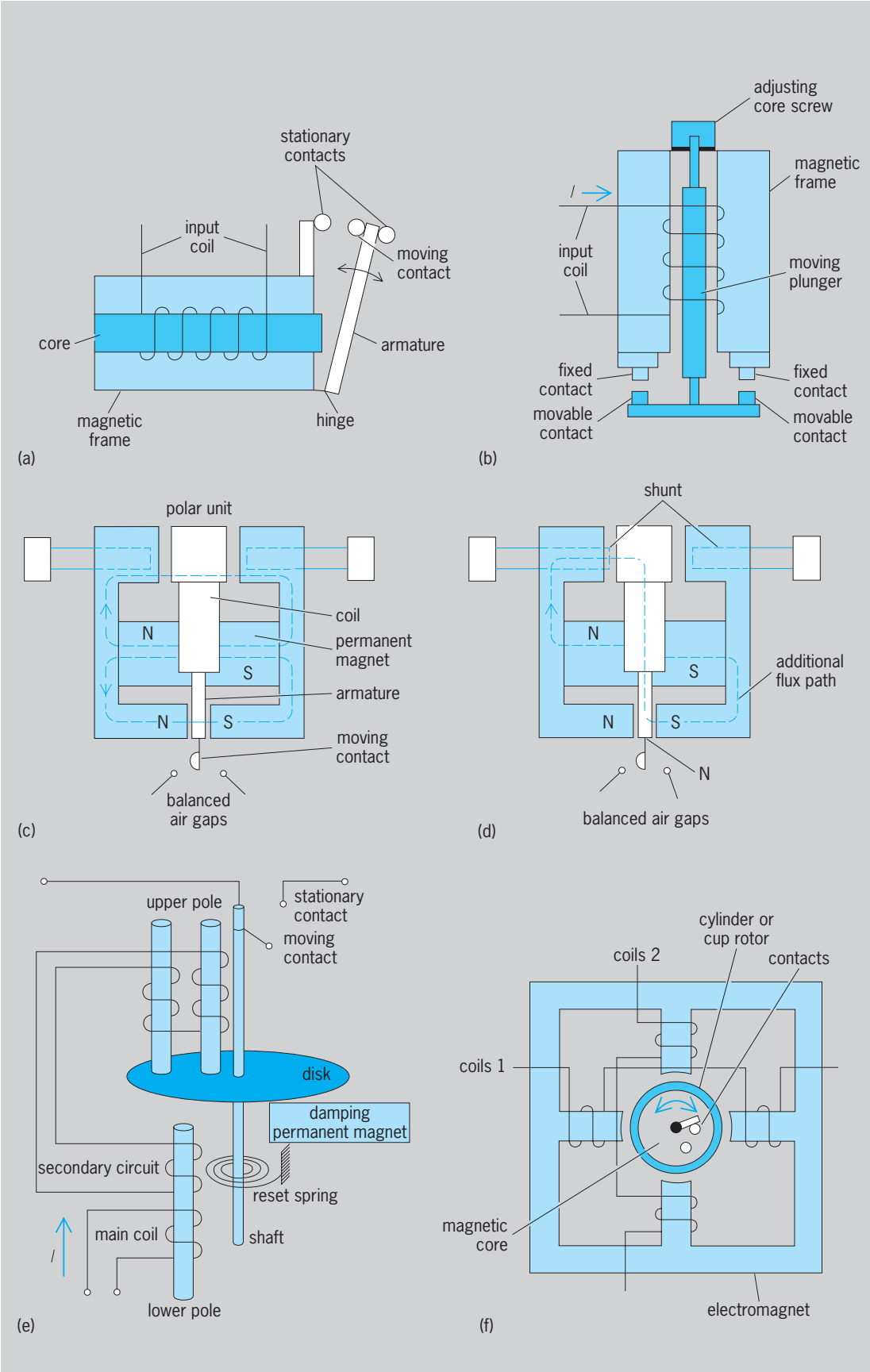

Fig. 2. Electromechanical relays. (*a*) Clapper or telephone type. (*b*) Plunger type. (*c*) Balanced polar unit. (*d*) Unbalanced polar unit. (*e*) Induction-disk type. (*f*) Cylinder or cup type (top view). (*Parts a, b, and f after J. L. Blackburn, Protective Relaying, Principles and Application, Marcel Dekker, 1987*)

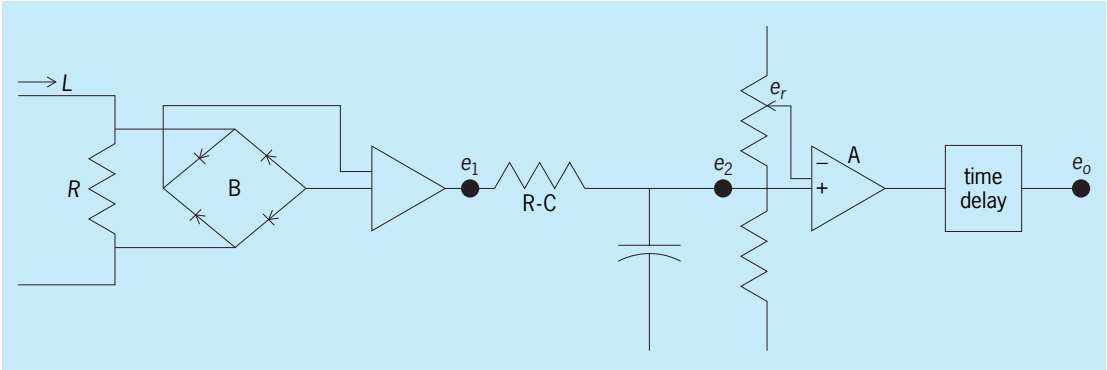

Fig. 3. A possible circuit configuration for a solid-state overcurrent relay. (*After S. H. Horowitz and A. G. Phadke, Power System Relaying, Research Studies Press, 2d ed., 1995*)

In the circuit shown in **Fig. 3**, the input current I is passed through the resistive shunt R, full-wave rectified by the bridge rectifier B, filtered to remove the ripple by the R-C filter, and applied to a high-gain summing amplifier A. The other input of the summing amplifier is supplied with an adjustable reference voltage e_r. When the input on the positive input of the summing amplifier exceeds the reference setting, the amplifier goes high. This step change is delayed by a time-delay circuit, in order to provide immunity against transient signals. By making the time-delay circuit adjustable, and making the amount of the delay depend on the magnitude of the input current, an instantaneous or a time-delay overcurrent relay characteristic can be obtained.

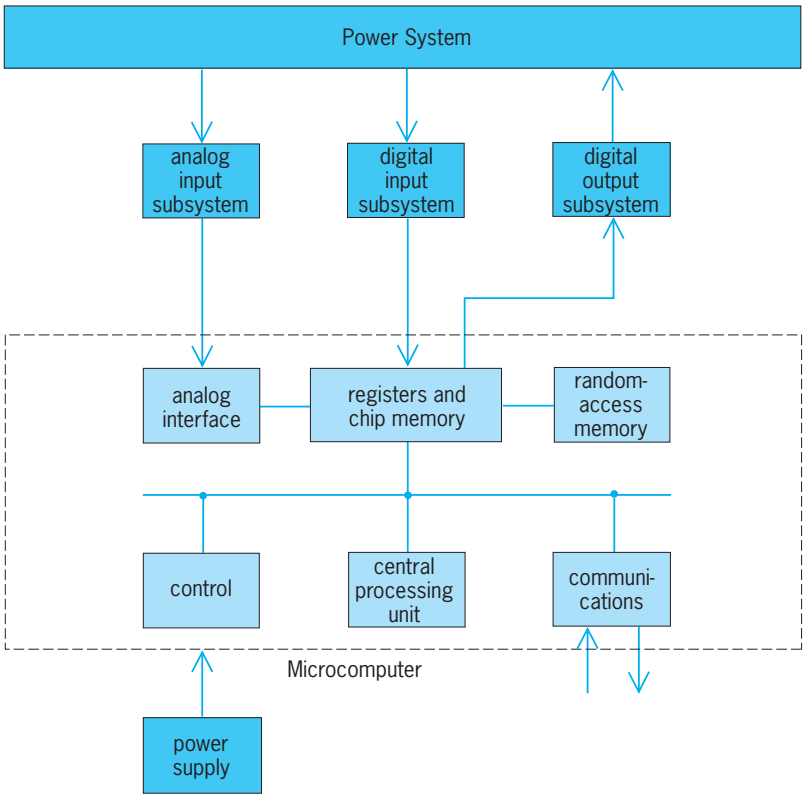

Fig. 4. Typical block diagram of a microprocessor relay. (*After Institute of Electrical and Electronics Engineers, Power Engineering Society, Microprocessor Relays and Protection Systems, Report 88EH0269-1-PWR, 1988*)

Microprocessor relays. Early computer relays were designed to replace existing protective functions, including transmission line, transformer, and bus protection. Some relays employed microprocessors to make the relaying decisions, using digitized analog signals, while others continued to use analog concepts to make relaying decisions and digital techniques for the necessary logic and auxiliary functions. However, both these types of digital relay provided two major advantages: (1) the ability of the relay to diagnose itself, a capability that could be obtained with an analog relay, if at all, only with great effort, cost, and complexity; and (2) a communication capability that allowed the relay to warn system operators when it was not functioning properly, permitted remote diagnostics and possible correction, and provided local and remote readout of its settings and operations. As research on digital relays continued, and confidence was gained, another dimension was added to the reliabilty of the protective system. An inherent feature of the software-dominated digital relay is its ability to adapt itself to changing system conditions in real time. These changes can be initiated by local inputs or by signals sent from a central computer. The digital relay is likely to become the preferred protective system due to the self-diagnostic capability, the hierarchical nature, and the data-sharing abilities of microprocessors, and their ability to adapt settings and other characteristics to actual system conditions in real time.

The major functional blocks for a microprocessor relay are shown in **Fig. 4**. Some of the blocks shown may not be required in certain relay designs.

The analog input subsystem translates the power-system information of current, voltage, frequency, temperature, and so forth, as required for relay fault detection and operation. It provides protection from power-system transients and overvoltages, converts currents to equivalent voltage signals, reduces voltages levels, and prevents high-frequency components from distorting the fault quantities.

The analog interface provides the analog-to-digital conversion and the multiplexing hardware. The status of the power-system circuit breakers, associated isolators, relay targets, and voltage levels is supplied to the relay through the digital input subsystem. The

relay program, settings, and logic are stored in chip memory. The control, central processing unit, and registers operate together to execute the program via chip and random-access memory. The relay output is via the digital output subsystem.

Where the relay is part of a pilot relaying system, the communications block provides the required channel information both to and from the remote terminals of the protected line. The random-access memory provides storage of pre- and postfault or disturbance information with appropriate time information. This is accessed through the digital output subsystem for system analysis and fault location.

J. Lewis Blackburn; Stanley H. Horowitz

Bibliography. J. L. Blackburn, *Protective Relaying, Principles and Applications*, 2d ed., Marcel Dekker, 1997; W. A. Elmore (ed.), *Protective Relaying: Theory and Applications*, 2d ed., Marcel Dekker, 2004; S. H. Horowitz and A. G. Phadke, *Power System Relaying*, 2d ed., Research Studies Press Ltd, Taunton, Somerset, England, 1995; Institute of Electrical and Electronics Engineers, Power Engineering Society, *Microprocessor Relays and Protection Systems*, Report 88EH0269-1-PWR, 1988; C. R. Mason, *The Art and Science of Protective Relaying*, Wiley, New York, 1955; L. P. Singh, *Digital Protection: Protective Relaying from Electromechanical to Microprocessor*, Wiley, 1994.

Reliability, availability, and maintainability

Reliability is the probability that an engineering system will perform its intended function satisfactorily (from the viewpoint of the customer) for its intended life under specified environmental and operating conditions. Maintainability is the probability that maintenance of the system will retain the system in, or restore it to, a specified condition within a given time period. Availability is the probability that the system is operating satisfactorily at any time, and it depends on the reliability and the maintainability. Hence the study of probability theory is essential for understanding the reliability, maintainability, and availability of the system. *See* PROBABILITY.

Reliability is basically a design parameter and must be incorporated into the system at the design stage. It is an inherent characteristic of the system, just as is capacity, power rating, or performance. A great deal of emphasis is placed on quality of products and services, and reliability is a time-oriented quality characteristic. There is a relationship between quality or customer satisfaction and measures of system effectiveness, including reliability and maintainability (**Fig. 1**). Customers are concerned with the performance of the product over time.

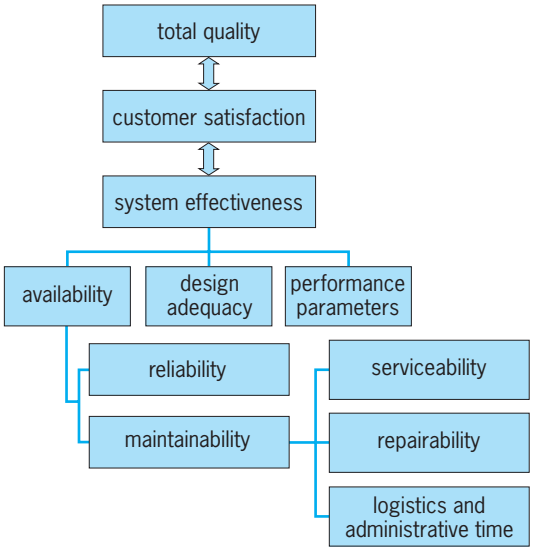

Fig. 1. Relationship between reliability, maintainability, and customer satisfaction.

General Principles

The reliability level must be established at the design phase, and subsequent testing and production will not raise the reliability without a basic design change. With increasing system complexity, reliability becomes an elusive and difficult design parameter to define and achieve. It also becomes more difficult to control, demonstrate, and ensure as an operational characteristic under the projected conditions of use by the customer. However, history has demonstrated that where reliability is recognized as a necessary program development component, with the practice of various reliability engineering methods throughout the evolutionary life cycle of the system, reliability can be quantified during the specification of design requirements, predicted by testing, controlled during production, and sustained in the field.

System reliability. To analyze and measure the reliability and maintainability characteristics of a system, there must be a mathematical model of the system that shows the functional relationships among all the components, the subsystems, and the overall system. The reliability of the system is a function of the reliabilities of its components. A system reliability model consists of some combination of a reliability block diagram or a cause-consequence chart, a definition of all equipment failure and repair distributions, and a statement of spare and repair strategies. All reliability analyses and optimizations are made on these conceptual mathematical models of the system.

Reliability block diagram. A reliability block diagram is obtained from careful analysis of the manner in which the system operates. An analysis has to be done of the effects on overall system performance of failure of the various components; the support environment and constraints, including such factors as the number and assignment of spare parts and repair persons; and the mission for the system.

The engineering analysis, which has to be done on the system in order to develop a reliability model, consists of four steps: (1) development of a functional block diagram of the system based on physical principles governing the operations of the system; (2) development of the logical and topological

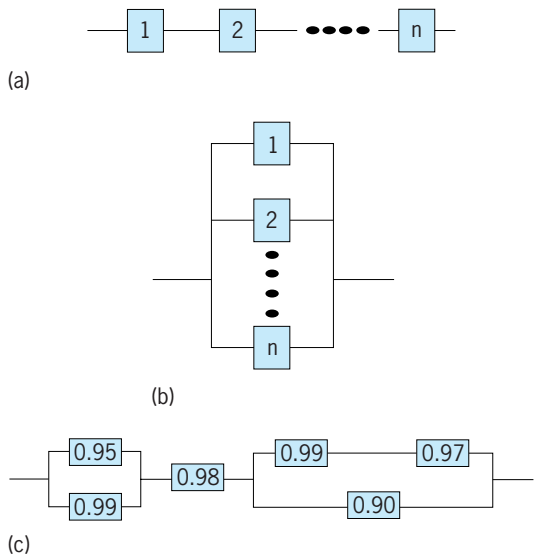

Fig. 2. System reliability block diagrams. (*a*) Series system. (*b*) Parallel system. (*c*) Combination system.

relationships between functional elements of the system; (3) use of performance evaluation studies to determine the extent to which the system can operate in a degraded state; and (4) definition of the spares and repairs strategies (for maintained systems).

Based on the engineering analysis, a reliability block diagram is developed that is used for calculating various measures of reliability and maintainability. The reliability block diagram depicts all the success or failure combinations for the system. Two of the guidelines for drawing these diagrams are: A group of components that are essential for the performance of the system and its mission are drawn in series (**Fig. 2***a*); and components that can substitute for other components are drawn in parallel (Fig. 2*b*). The failure behavior of all the redundant components must be specified. Some of the common types of redundancies are (1) active redundancy or hot standby, where the component has the same failure rate as if it were operating in the system; (2) passive redundancy, spare or cold standby, where the standby component cannot fail—generally assumed of spare or shelf items; (3) warm standby, where the standby component has a lower failure rate than the operating component—usually a realistic assumption.

Structure. The structure is called a series structure when the system functions if and only if all the n components of the system function (Fig. 2*a*). The components are assumed to fail or function independently of one another.

The structure is said to be parallel when the system functions if at least one of the n components of the system functions (Fig. 2*b*).

The reliability of a system that consists of both series and parallel subassemblies can also be computed (Fig. 2*c*). The reliability values (probability of success) for each component is given in each block.

Reliability measures. Various measures are used to quantify and evaluate the reliability of a system. The reliability function is used to evaluate reliability over time t. Reliability is concerned with the successful performance of a system, and it is evaluated by the time-to-failure random variable, T. Then reliability (R) at time t, $R(t)$, is the probability that the system will not fail by time t. This definition can be expressed mathematically by Eq. (1). Here $P[T > t]$ is the probability that the time-to-failure T is greater than t, and $f(t)$ and $F(t)$ are the probability density function and cumulative distribution function for T, respectively. For example, if the time to failure is exponentially distributed, then these functions are given by Eqs. (2) and (3), and reliability at time t is given by Eq. (4). If

$$\begin{aligned} R(t) &= P[T > t] \\ &= 1 - P[T \le t] \\ &= 1 - F(t) \\ &= 1 - \int_0^t f(\tau)\,d\tau \end{aligned} \tag{1}$$

$$f(t) = \lambda e^{-\lambda t} \qquad t \ge 0, \lambda > 0 \tag{2}$$

$$F(t) = \int_0^t \lambda e^{-\lambda t} d\tau = 1 - e^{-\lambda t} \qquad t \ge 0 \tag{3}$$

$$R(t) = e^{-\lambda t} \qquad t \ge 0 \tag{4}$$

there is some specified time or mission duration, t^*, for which the sytem must perform its intended function, then the mission reliability is given by $R(t^*)$. For many products, it is convenient to express reliability in terms of the mean life or expected life, given by Eq. (5). The quantity $E[T]$ is also called the mean

$$E[T] = \int_0^\infty t f(t)\,dt = \int_0^\infty R(t)\,dt \tag{5}$$

time to failure (MTTF) or mean time between failures (MTBF) for repairable systems. For the exponential distribution, it is given by Eq. (6).

$$E[T] = \frac{1}{\lambda} = \text{MTBF} \tag{6}$$

Another common measure of reliability is B_α life, which is the α percentile for the life of the product. Thus, $\alpha/100 = F(B_\alpha)$. B_{10} is the 10th percentile for the life of the product, and many times is used to specify the life requirements for the product.

If there is a large population of the items whose reliability is of interest, then for replacement and maintenance purposes it becomes important to determine the rate at which the items in the population, which have survived at any point in time, will fail. This is called the failure rate or hazard rate $h(t)$, and is given by $h(t) = f(t)/R(t)$.

The failure rate for most components follows the life characteristic curve (**Fig. 3***a*). There are three types of failures, namely quality failures, stress-related failures, and wear-out failures (Fig. 3*b*). The sum total of these failures gives the overall failure rate (Fig. 3*a*). The failure rate curve, also known as the bathtub curve (Fig. 3*a*), has three distinct periods.

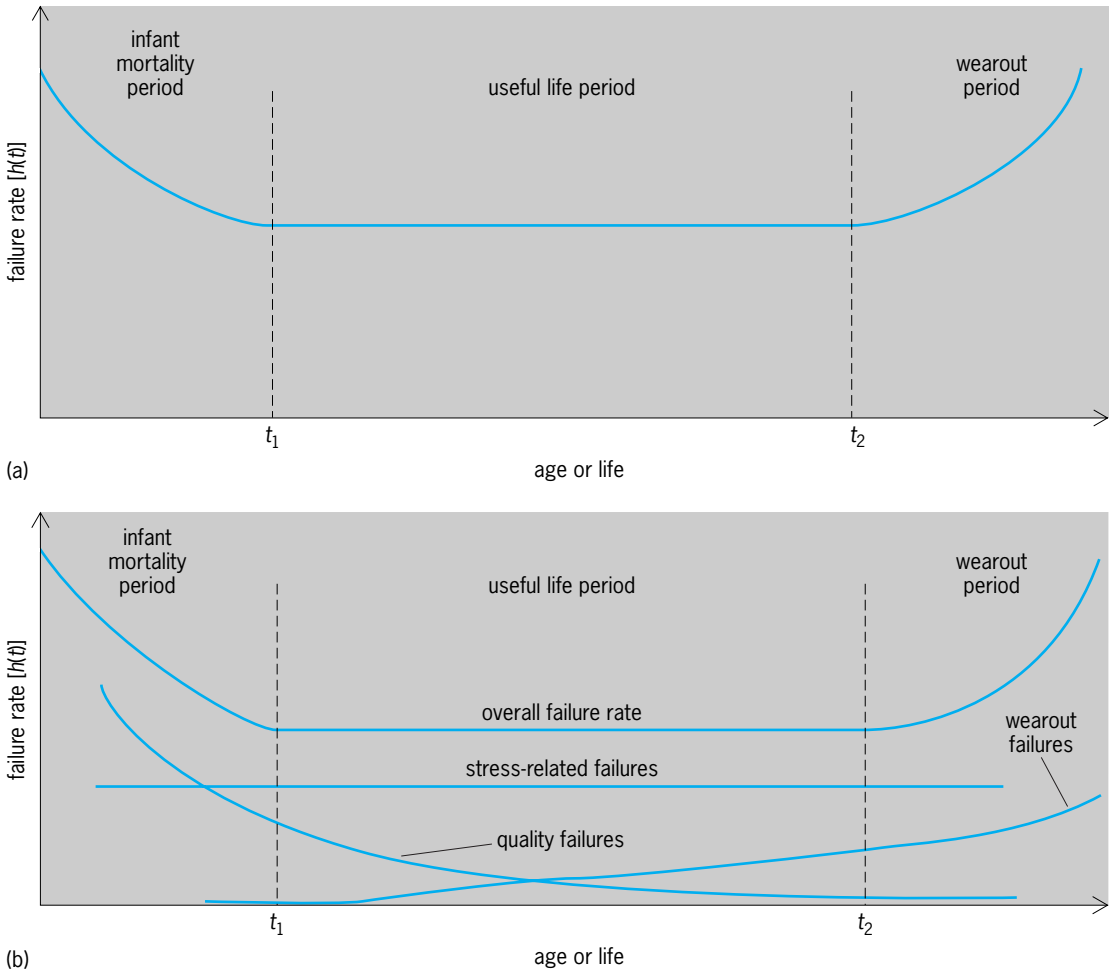

Fig. 3. Failure rate curves. (*a*) Failure rate–life characteristic curve (bathtub curve). (*b*) Failure rate curves based on components of failure.

The initial decreasing failure rate is termed infant mortality and is due to the early failure of substandard products. Latent material defects, poor assembly methods, and poor quality control can contribute to a high initial failure rate. A short period of in-plant product testing, termed burn-in, is used by manufacturers to eliminate these early failures from the consumer market. The flat, middle portion of the failure-rate curve represents the design failure rate for the specified product as used by the consumer market. During the useful-life portion, the failure rate is relatively constant. It might be decreased by redesign or by using a higher safety margin. Finally, as products age they reach a wear-out phase characterized by an increasing failure rate.

Durability. Durability is a special case of reliability. Durability is the probability that an item will successfully survive its projected service life, overhaul point, or rebuild point (whichever is the most appropriate durability measure for the item) without a durability failure. A durability failure is considered to be a malfunction that precludes further operation of the item and is great enough in cost, safety, or time to restore that the item must be replaced or rebuilt.

Maintainability. Maintainability is a measure of the ease and rapidity with which a system or equipment can be restored to operational status following a failure. It is a characteristic of equipment design and installation, personnel availability in the required skill levels, adequacy of maintenance procedures and test equipment, and the physical environment under which maintenance is performed. Maintainability is expressed as the probability that an item will be retained in or restored to a specific condition within a given period of time, when the maintenance is performed in accordance with prescribed procedures and resources.

One measure of maintainability is the mean time to repair (MTTR), which is defined as the mean of the distribution of system repair time. If a system has N subsystems, the MTTR for the system is given by Eq. (7), where λ_i is the failure rate of the ith repairable subassembly in the system, and t_i is the expected time to repair the system when its ith subassembly fails. The MTTR is also given by the total corrective maintenance time (the time required for actions performed, as a result of failure, to restore the equipment or system to operable conditions)

$$\mathrm{MTTR} = \frac{\sum_{i=1}^{N} \lambda_i t_i}{\sum_{i=1}^{N} \lambda_i} \tag{7}$$

divided by the total number of corrective maintenance actions (actions resulting from the nonoperablity of the equipment or system) during a given period of time.

Several other measures for maintainability are in use. The mean time between unscheduled maintenance actions (MTBUMA) is the average number of time units for which a system functions successfully between successive unscheduled maintenance actions. Active maintenance time is the time during which preventive and corrective maintenance work is actually being done on the item. Different types of maintenance are distinguished: Periodic maintenance is maintenance performed on equipment on the basis of hours of operation or calendar time elapsed since the last inspection; and preventive maintenance is the systematic care, servicing, and inspection of equipment and facilities for the purpose of maintaining them in serviceable condition and detecting and correcting incipient failures. Finally, the maintenance ratio (MR) is given by Eq. (8),

$$\mathrm{MR} = \frac{\mathrm{MPH}}{\mathrm{OH}} \tag{8}$$

where MPH is the number of maintenance person-hours and OH is the number of system operating hours.

Availability. Availability is defined as the probability that a system is operating satisfactorily at any point in time. This definition considers only the operating time and the downtime excluding idle time. The steady-state availability (A) of a system with known MTBF and MTTR is given by Eq. (9).

$$\mathrm{A} = \frac{\mathrm{MTBF}}{\mathrm{MTBF} + \mathrm{MTTR}} \tag{9}$$

Design reliability methodology. The best way to improve reliability is at the concept and design stage by using various failure prevention techniques. Two common techniques are failure mode, effects, and criticality analysis (FMECA), which represents a bottom-up approach; and fault-tree analysis (FTA), which represents a top-down approach. Both represent qualitative as well as quantitative approaches for assessing the consequences of failure and its prevention.

Failure mode, effects, and criticality analysis. This is an iterative documented process of a systematic nature performed to identify basic faults at the part level and to determine their effects at high levels of assembly (**Fig. 4**). Failure mode, effects, and criticality analysis can be performed by utilizing either actual failure modes from field data or hypothesized failure modes derived from design analysis, reliability prediction activities, and experiences relative to the manner in which parts fail. In their most complete form, failure modes are identified at the part level, which is usually the lowest level of direct concern to the designer of the product or process. In addition to providing insight into failure cause-and-effect relationships, the failure mode and effects analysis provides a disciplined method for proceeding part by part through the system to assess failure consequences.

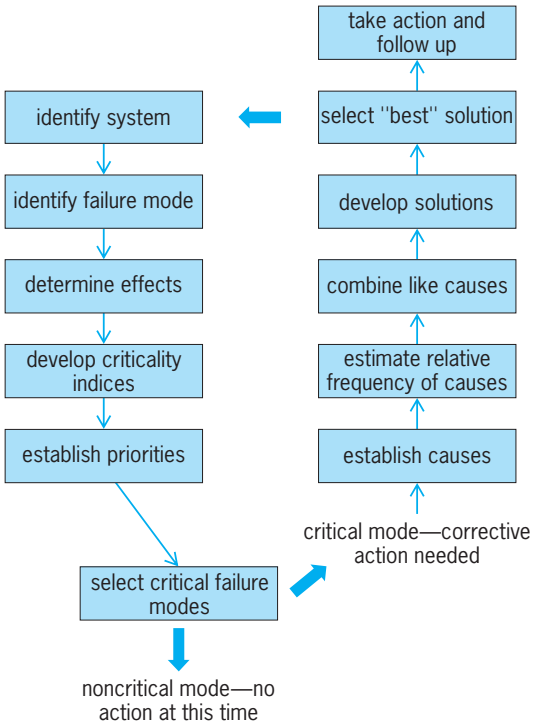

Fig. 4. Diagram of methodology for failure mode, effects, and criticality analysis.

Failure modes are analytically induced into each component, and failure effects are evaluated and noted, including severity and frequency (or probability) of occurrence. As the first mode is listed, the corresponding effect on performance at the next higher level of assembly is determined. The resulting failure effect becomes, in essence, the failure mode that impacts the next higher level. Iteration of this process results in establishing the ultimate effect at the system level. Once the analysis has been performed for all failure modes, it is usually the case that each effect or symptom at the system level is caused by several different failure modes at the lowest level. This relationship to the end effect provides the basis for grouping the lower-level failure modes.

By using this approach, probabilities for the occurrence of the system effect can be calculated, based on the probability occurrence of the lower-level failure modes. Based on these probabilities, and a severity factor assigned to the various system effects, a criticality number can be calculated. Criticality numerics provide a method of ranking the system-level effects derived previously. Criticality numerics also provide the basis for corrective action priorities, engineering changes, and resolution of problems in the field.

Fault-tree analysis. This is a tool that lends itself well to analyzing failure modes found during design, factory test, or field data returns. The fault-tree analysis procedure can be characterized as an iterative documented process of a systematic nature performed to identify basic faults, determine their causes and effects, and establish their probabilities of occurrence.

The approach involves several steps, among which is the structuring of a highly detailed logic diagram that depicts basic faults and events that can lead to system failure and safety hazards. The next steps are collecting basic fault data and failure probabilities for use in computation; and using computational techniques to analyze the basic faults, determine failure mode probabilities, and establish criticalities. The final step involves formulating corrective suggestions that, when implemented, would eliminate (or minimize) those faults considered critical.

Reliability growth, demonstration, and testing. As the product goes through the various life cycles, reliability should be estimated and predicted. These values, when plotted at selected points in the life cycle, result in a growth curve that reflects comparative reliability levels. The improvement in reliability is dependent upon the nature of the product. Desirable reliability growth results from planning, designing, testing, producing, and using the product according to specified rules.

The purpose of reliability demonstration and testing is to determine the current reliability levels of the product. Tests must be designed in such a manner that maximum information can be obtained from a minimum amount of testing. A major problem in the design of adequate tests is that of simulating the real-world environment. During its lifetime the product is subjected to many environmental factors, such as temperature, vibrations and shock, and rough handling. These stresses may be encountered singly, simultaneously, or sequentially. In addition, there are many other random factors. *See* OPTIMIZATION; RISK ASSESSMENT AND MANAGEMENT.

Software Reliability

Software reliability is the probability that software will not cause the failure of a product for a specified time under specified conditions. This probability is a function of the inputs to and use of the product as well as a function of the existence of faults in the software. The inputs to the product determine whether existing faults, if any, are encountered. Software reliability is also sometimes defined as the ability of a program to perform a required function under stated conditions for a stated period of time.

As computers spread through society, the size and complexity of software systems continually increase. Many institutions rely heavily on computers to provide vital services. International banking, police investigations, and airline reservations are dependent on computers. Computer software systems support safety-critical functions in the control of nuclear reactors, defense systems, air traffic, space exploration, and many consumer products including medical devices. The complexity of software systems, the trust placed in them, and the potentially catastrophic consequences of software failure demand quantitative guarantees of software quality.

Software faults and software failure. A fault is the manifestation of an error in software. A fault, if encountered, may cause a failure and sometimes is synonymous with a "bug." Failure is the inability of a product or component to perform a required function within specified limits; failures will be observed during testing or actual use of the software.

Whereas hardware failures are frequently due to physical wearout, software failures are due to design faults (bugs). These faults are defects, which may be due to incorrect or incomplete specifications in the design phase of the software, or due to simple programmer error. When the software is executed, it may encounter an input for which the software was not designed or tested. Under some of these inputs, a fault will induce failure.

While the number of faults remaining in a program affects its performance, a more informative measure of the reliability of software is its failure rate (failure intensity), the number of failures occurring per unit time. Indeed, studies have found that systems with many faults can still be highly reliable, if each fault occurs very infrequently.

Design for software reliability. The source of a software fault is a design fault, and hence when a software fault is removed correctly it will never recur. Thus, if all the faults were eventually removed, the software would be fault-free and would execute without failure forever. This is in contrast to hardware reliability where, due to physical causes and degradation, all systems will eventually fail. The techniques to achieve software reliability can be classified under three categories: fault avoidance, fault removal, and fault tolerance.

Fault-avoidance techniques are engineering and management tools to prevent the creation of faults. In addition to correct software requirements and specifications, some of the approaches are control of complexity, performance of top-down functional decomposition, modularization, hierarchical design, and adherence to structured approaches. *See* SOFTWARE ENGINEERING.

Fault-removal techniques attempt to discover and correct faults before software is released. Discovery is done through software testing for quality, reliability, safety, maintainability, and overall qualification. *See* SOFTWARE TESTING AND INSPECTION.

Since it is virtually impossible to guarantee fault-free software, fault-tolerance techniques are used to detect and recover from faults when the software is in use. N-version programming is one such approach. Several independently written programs are run in parallel; if their outputs are in conflict, a predetermined scheme such as majority voting determines the output to be used. *See* FAULT-TOLERANT SYSTEMS.

Software reliability engineering (SRE) activities to achieve high reliability should be applied throughout the life cycle of the product starting from concept, through delivery, to the end of the product's useful life (see **table**).

Software reliability modeling. Models are a major statistical tool for assessing software reliability. Typically, a model combines a probabilistic description of the software failure process with a statistical inference procedure to make predictions of software failure based on past performance data. The data are

Software reliability engineering activities

Stage of product life cycle	Activities
Concept, feasibility, and requirements	Determine functional profile Define and classify failures Identify customer reliability needs Conduct trade-off studies Set reliability objectives
Design and implementation	Allocate reliability among components Engineer to meet reliability objectives Focus resources based on functional profile Manage fault introduction and propagation Measure reliability of acquired software
System test and field trial	Determine operational profile Conduct reliability growth testing Track testing progress Project additional testing needed Obtain certification for reliability objectives
Delivery, upgrades, and maintenance	Project postrelease staff needs Monitor filed reliability versus objectives Track customer satisfaction with reliability Time introduction of new features Guide product and process improvement

obtained from testing and are usually a sequence of execution times between successive failures (interfailure times).

The earliest model for measurement of software reliability was proposed in 1972 independently by Z. Jelinsky and P. Moranda and by M. Shooman. The model makes the following assumptions: (1) failures occur randomly, (2) the program's failure rate is proportional to the number of remaining bugs it contains, and (3) at each failure a fault is fixed. Using the statistical method of maximum likelihood, the interfailure time data are used to estimate the total number of bugs and the constant of proportionality.

More advanced software reliability models and customer-centered tools have been developed. These tools allow software reliability to be measured and improved. One of the first software reliability measurement tools was developed in 1977 and has been commercially available since 1990. It was followed by SMERFS (Statistical Modeling and Estimation of Reliability Functions of Software), developed at the U.S. Naval Surface Warfare Center. Some other tools are SRMP (Software Reliability Modeling Program) and CASRE (Compiler Aided Software Reliability Estimation).

Unlike earlier models, which assume that the random discovery of faults (that is, the input to the program) is the only source of uncertainty in the failure process, more sophisticated models view the software failure process to be the result of two sources of uncertainty: the selection of input to the program, and the effect of debugging. The uncertainty in debugging may be due both to the size (rate and occurrence) of the fault removed and to the fact that debugging may not be perfect. Advanced models also attempt to estimate the probability of software failure when testing reveals no errors. These methods are based on bayesian estimation, and incorporate the testing results, information about the input distribution during use, and prior assumptions about the probability of failure of the software. Software reliability measurement is a useful tool in assessing when or whether software is ready for release, and for estimating software reliability for future maintenance costs. *See* BAYESIAN STATISTICS; QUALITY CONTROL.

Kailash C. Kapur

Bibliography. S. Bittanti (ed.), *Software Reliability Modelling and Identification*, 1989; E. A. Elsayed, *Reliability Engineering*, Addison Wesley, 1996; W. G. Ireson, C. F. Coombs, and R. V. Moss (eds.), *Handbook of Reliability Engineering and Management*, 2d ed., 1996; K. C. Kapur and L. R. Lamberson, *Reliability in Engineering Design*, 1977; D. Kececioglu, *Reliability Engineering Handbook*, vols. 1 and 2, 1991; K. B. Klaasen and J. C. L. van Peppen, *System Reliability: Concepts and Applications*, 1989; M. R. Lyu (ed.), *Handbook of Software Reliability Engineering*, 1996.

Reluctance

A property of a magnetic circuit analogous to resistance in an electric circuit.

Every line of magnetic flux is a closed path. Whenever the flux is largely confined to a well-defined closed path, there is a magnetic circuit. That part of the flux that departs from the path is called flux leakage.

For any closed path of length l in a magnetic field H, the line integral of $H \cos \alpha \, dl$ around the path is the magnetomotive force (mmf) of the path, as in

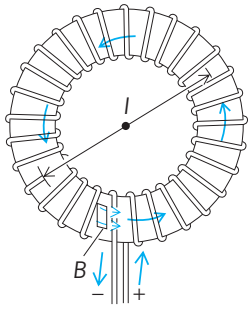

A toroidal coil, illustrating the concept of reluctance.

Eq. (1), where α is the angle between H and the path.

$$\text{mmf} = \oint H \cos\alpha \, dl \tag{1}$$

If the path encloses N conductors, each with current I, Eq. (2) holds.

$$\text{mmf} = \oint H \cos\alpha \, dl = NI \tag{2}$$

See MAGNETISM.

Consider the closely wound toroid shown in the **illustration**. For this arrangement of currents, the magnetic field is almost entirely within the toroidal coil, and there the flux density or magnetic induction B is given by Eq. (3), where l is the mean cir-

$$B = \mu \frac{NI}{l} \tag{3}$$

cumference of the toroid and μ is the permeability. The flux ϕ within the toroid of cross-sectional area A is given by either form of Eqs. (4), which is sim-

$$\begin{aligned} \Phi &= BA = \frac{\mu A}{l} NI \\ \Phi &= \frac{NI}{l/\mu A} = \frac{\text{mmf}}{l/\mu A} = \frac{\text{mmf}}{\mathcal{R}} \end{aligned} \tag{4}$$

ilar in form to the equation for the electric circuit, although nothing actually flows in the magnetic circuit. The factor $l/\mu A$ is called the reluctance $\mathcal{R}$ of the magnetic circuit. The reluctance is not constant because the permeability μ varies with changing flux density. From the defining equation for reluctance, it is seen that when the mmf is in ampere-turns and the flux is in webers, the unit of reluctance is the ampere-turn/weber.

Reluctances in series. For the simple toroid, all parts of the magnetic circuit have the same μ and the same A. More complicated circuits may include parts that differ in permeability, in cross section, or in both. Suppose a small gap were cut in the core of the toroid. The flux would fringe out at the gap, but as a rough approximation, the area of the gap may be considered the same as that of the core.

The magnetic path then has two parts, the core of length l_1 and reluctance $l_1/\mu_1 A$, and the air gap of length l_2 and reluctance $l_2/\mu_2 A$. Since the same flux is in both core and gap, this is considered a series circuit and Eq. (5) holds. Since the relative permeability

$$\mathcal{R} = \mathcal{R}_1 + \mathcal{R}_2 = \frac{l_1}{\mu_1 A} + \frac{l_2}{\mu_2 A} \tag{5}$$

of the ferromagnetic core is several hundred or even several thousand times that of air, the reluctance of the short gap may be much greater than that of the much longer core. For any combination of paths in series, $\mathcal{R} = \Sigma l/\mu A$. Then Eq. (6) holds.

$$\Phi = \frac{\text{mmf}}{\Sigma \mathcal{R}} = \frac{\text{mmf}}{\Sigma l/\mu A} \tag{6}$$

Reluctances in parallel. If the flux divides in part of the circuit, there is a parallel magnetic circuit and the reluctance of the circuit has the same relation to the reluctances of the parts as has the analogous electric resistance. For the parallel circuit Eq. (7) is

$$\frac{1}{\mathcal{R}} = \frac{1}{\mathcal{R}_1} + \frac{1}{\mathcal{R}_2} + \cdots \tag{7}$$

valid. *See* RELUCTANCE MOTOR. Kenneth V. Manning

Reluctance motor

An alternating current motor with a stator winding like that of an induction motor, and a rotor that has projecting or salient poles of ferromagnetic material. When connected to an alternating-current source, the stator winding produces a rotating magnetic field, with a speed of $4\pi f/p$ radians per second ($120f/p$ revolutions per minute), where f is the frequency of the source and p the number of magnetic poles produced by the winding. When the rotor is running at the same speed as the stator field, its iron poles tend to align themselves with the poles of that field, producing torque. If a mechanical load is applied to the shaft of the motor, the rotor poles lag farther behind the stator-field poles, and increased torque is developed to match that of the mechanical load. This torque is given by the equation below,

$$\tau = \frac{p}{2} \phi^2 \frac{dR}{d\delta}$$

where ϕ is the flux per pole, determined largely by the applied voltage. The quantity $dR/d\delta$ is the rate of change of magnetic reluctance per pole with respect to δ, the angle of lag in mechanical radians. This quantity typically varies as $\sin p\delta$. Here, $p\delta = 2\delta_e$, where δ_e is the lag angle in electrical radians. Therefore, at constant torque load the rotor runs in synchronism with the stator field, with the rotor poles lagging the field poles by a constant angle. *See* ELECTRICAL DEGREE.

This motor develops torque only at synchronous speed, and thus no starting torque is produced. For that reason, induction-motor rotor bars are usually built into the pole faces, and the motor starts as an induction motor. When the rotor speed approaches that of the magnetic field, the pole pieces lock in step with the magnetic poles of the field, and the rotor runs at synchronous speed.

Single-phase reluctance motors may be started by the methods used for single-phase induction motors, such as capacitor, split-phase, or shaded-pole starting. *See* ALTERNATING-CURRENT MOTOR; ELECTRIC ROTATING MACHINERY; INDUCTION MOTOR; MOTOR; SYNCHRONOUS MOTOR. George McPherson, Jr.

Bibliography. T. J. Miller, *Brushless Permanent-Magnet and Reluctance Motor Drives*, 1989; T. J. Miller, *Switched Reluctance Motors and Their Control*, 1993; J. Rosenblat and M. H. Friedman, *Direct and Alternating Current Machinery*, 2d ed., 1998; C. G. Veinott and J. E. Martin, *Fractional and Subfractional Horsepower Electric Motors*, 4th ed., 1986.

Remipedia

A unique class of the Crustacea whose members, at first glance, resemble polychaete worms in overall body form. This impression comes from their multiple serial segments and lack of obvious body divisions (tagmosis; see **illus.**). The lack of body organization suggests that remipeds are the most primitive crustaceans yet recognized. When compared with the other presumably primitive crustacean class, the Cephalocarida, remipeds are relatively large, with mature individuals ranging from 0.6 to 1.7 in. (15 to 42 mm) in body length. Although the trunk appendages are biramous and paddle-shaped in both classes, endites and epipods that are present in cephalocarids are lacking in remipeds. *See* CEPHALOCARIDA.

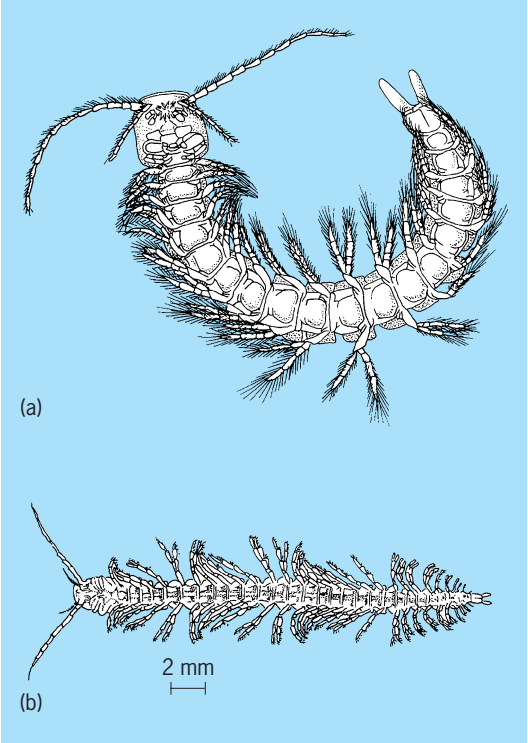

Representative Remipedia. (*a*) *Speleonectes lucayensis* (*after F. R. Schram, Crustacea, Oxford University Press, 1986*). (*b*) *Lasionectes entrichoma* (*after J. Yager and F. R. Schram, Lasionectes entrichoma, new genus, new species* (*Crustacea: Remipedia*) *from anchialine caves in the Turks and Caicos, British West Indies, Proc. Biol. Soc. Wash., vol. 99, 1986*).

Discovery. The first remiped species was discovered by divers in 1981 in an achialine cave on Grand Bahama Island. In subsequent years, five more Recent species have been described. Additionally, one previously described fossil species has been assigned to this class. As now constituted, the Remipedia comprise two orders, three families, and four genera; however, two or three potentially new species have already been collected, and as more marine caves are explored additional species will be discovered.

Morphology. Recent remipeds (order Nectiopoda) are characterized by the lack of a carapace but with a distinct cephalic shield covering the head. A pair of frontal processes of undetermined function is located just in front of the biramous, multiarticulated antennules. The antennae are also biramous and paddlelike in form. A lacinia mobilis separates the incisor and molar processes of the mandibles, which are at least partially enclosed in the atrium oris formed by the large labrum, or upper lip. Three pair of postoral food-handling appendages (maxillule, maxilla, and maxilliped) are raptorial in structure, and distinctive among the four genera. The trunk consists of approximately 29–32 somites in mature adults, with no discernible division into thorax and abdomen. All of the serially similar trunk limbs are biramous, paddlelike, and directed ventrolaterally. Eyes are lacking. The anal segment is provided with a pair of caudal rami. Remipeds are hermaphroditic, with female gonopores present on the seventh trunk segment and male gonopores on the fourteenth. Excretion is by means of a pair of maxillary glands located in the head. The digestive system consists of a distensible anterior and tubular posterior foregut, a long midgut with paired diverticula in each segment, and a short hindgut.

Ecology. Only limited ecological information on remipeds is available. All species known to date have been found in caves where oxygen levels of the water are low. It is not uncommon for several species to coexist in a single cave. Little is known about feeding habits of remipeds; however, comparisons of various adaptations of the raptorial feeding appendages suggest that species competition is avoided by prey selectivity. At present the distribution of remipeds appears restricted to the subtropical and tropical Atlantic Ocean islands of the Bahamian Archipelago and Canary Islands.

Phylogeny. The evolutionary position of the Remipedia and the relationship to other crustacean taxa are speculative at this time. Clearly the lack of trunk specialization is a primitive condition not seen in other crustacean taxa. Serial homology of trunk appendages is also generally accepted as the primitive state. Similarly, biramous appendages, such as those found in remipeds, have long been postulated as the precursor of the more specialized stenopodus walking leg. However, in view of incorporation of the anterior-most thoracic appendage (maxilliped) into the cephalon (head), the specialized feeding

appendages, and restricted habitat, it seems premature to postulate the Remipedia as being at the base of the crustacean phylogenetic tree. *See* CRUSTACEA.

Patsy A. McLaughlin

Bibliography. F. R. Schram, J. Yager, and M. J. Emerson, *Remipedia,* pt. I: *Systematics*, San Diego Soc. Nat. Hist. Mem. 15, 1986; J. Yager, *Cryptocorynetes haptodiscus*, new genus, new species, and *Speleonectes benjamini*, new species, of remipede crustaceans from anachialine caves in the Bahamas, with remarks on distribution and ecology, *Proc. Biol. Soc. Wash.,* vol. 100, 1987; J. Yager, Remipedia, a new class of Crustacea form a marine cave in the Bahamas, *J. Crust. Biol.*, vol. 1, 1981.

Remote-control system

A control system in which the issuing of the control command and its execution are separated by a relatively significant distance. The system normally includes a command device where the control command is entered, and an actuator that executes it. These are connected by a transmission medium that transmits the command, usually in a coded format.

The transmission medium may be a mechanical link, where the command is transmitted as force; a pneumatic or hydraulic line, where pressure represents the command; an electrical line with a voltage or current signal; or radio or infrared waves that are modulated according to the command. *See* MODULATION.

The simplest remote-control systems are limited to switching-type functions. These systems operate basically in an open loop, that is, without relying on feedback (although an acknowledgment of the command's execution may be provided). Some typical examples are a ceiling lamp turned on and off by a light switch via an electrical wire; the on/off function of a television receiver and videocassette recorder with an infrared remote controller; power switches in distribution networks operated from a control center; and railway switches operated from a remote-control room.

The most characteristic remote-control systems involve feedback that is provided by the human operator. The person issuing the control command senses (usually sees) the result of the control action and guides the system accordingly. This kind of operation can be found, for example, in remote control of the aspects of a television picture, remote control of toy cars and airplanes by wire or radio, remote operation of large construction cranes, and cockpit control of an airplane's engines and control surfaces, usually by means of a hydraulic system. *See* FLIGHT CONTROLS.

Teleoperation represents an important class within remote-control systems with human feedback. In contrast to the systems noted above where the control action is limited to precisely defined functions, teleoperators (or remote manipulators) act as extensions of the human hand. They are employed in situations where access is difficult or impossible or where the environment is hazardous for humans, such as in underwater and space operations, or in the presence of radiation, chemical, or biological contamination. *See* REMOTE MANIPULATORS.

Strictly speaking, many automatic control systems (that is, feedback control systems with no human operator in the loop) may also be considered as remote controllers. This is the case whenever the sensing of the controlled variable and the automatic formation of the control command are removed from the actuator. A typical example is the heating and air-conditioning system of a building, where room thermostats operate remotely located furnaces, compressors, and fans. Similar systems are frequently found in the control of large industrial plants. *See* CONTROL SYSTEMS.

Janos Gertler

Remote manipulators

Mechanical, electromechanical, or hydromechanical devices which enable a person to perform manual operations while separated from the site of the work. Remote manipulators are designed for situations where direct contact would be dangerous to the human (working with radioactive material), where direct human contact is ill-advised or impossible (certain medical procedures), and where human force-producing capabilities are absent (the disabled) or need to be amplified to complete some task (industrial assembly or construction).

Basic defining elements are common to almost all remote manipulators (**Fig. 1**). An input device or control handle allows the operator to command the remote manipulator. The movement of the input device is received by a control station that translates

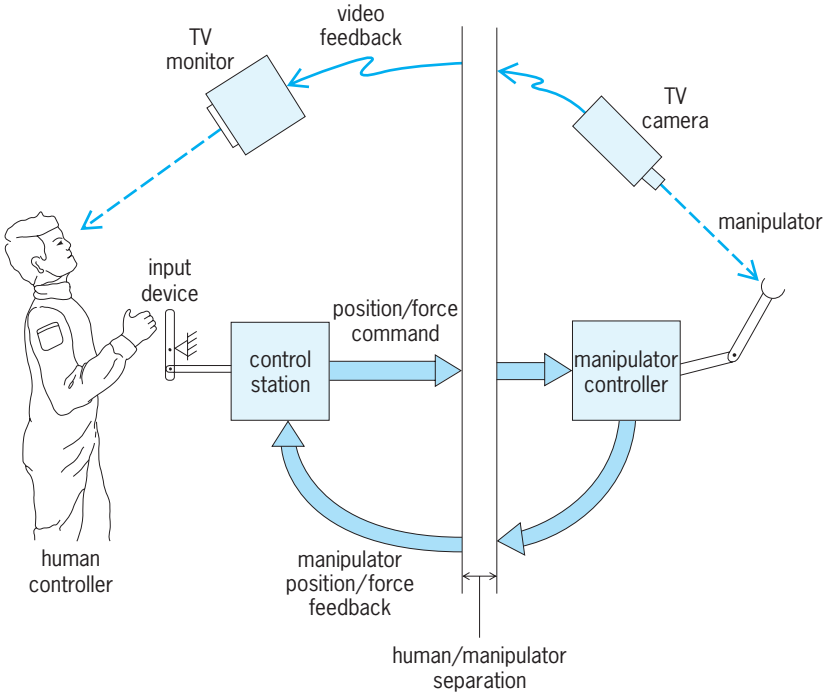

Fig. 1. Elements of a typical remote manipulator system.

the inputs into a form that can be transmitted over the distance separating the human and remote manipulator. This translation can be mechanical, using cables and linkages, or electrical/electromagnetic, using the movement of the input device to generate an electrical or electromagnetic signal that is easily transmitted to the remote manipulator. Since vision is an important cue that humans use in direct manipulation, visual feedback of a remote manipulator's actions typically must be provided. In some remote manipulation systems, tactile feedback to the human operator is provided; that is, forces proportional to those being exerted by the remote manipulator on the object are fed back to the human through the input device. Such force feedback is important in certain tasks where the possibility of damage to the manipulator or object can occur.

Historical development. Perhaps the earliest example of remote manipulation occurred in the mid-1940s, when master-slave devices were designed and constructed to aid in the handling of radioactive materials. The term "master-slave" refers to the parallel construction and operation of the human input device (master) and the remote manipulator (slave). Typically, the remote manipulator was a mechanical copy of the input device used by the human, although the remote manipulator could be scaled in size. Communication between these early devices was by direct mechanical linkages. Soon, however, electronic communication was used. Later, digital computers were used to develop sophisticated interfaces between operator and remote manipulator. *See* RADIOCHEMICAL LABORATORY.

The master-slave implementation for remote manipulation implies a nearly one-to-one correspondence between human controller input and remote manipulator motion. In some applications this is impractical, such as where the manipulator needs to be moved a considerable distance when performing a task. In such applications, the remote manipulator has been designed to provide a rate of change of position or velocity in response to human controller position changes.

Applications. The growth in the number and variety of remote manipulator applications has been aided by enabling technologies such as digital computers, lightweight materials, and video communication links. Space applications include the space shuttle remote manipulator arm, which has been used to retrieve and launch large satellites (**Fig. 2**). The arm is operated by astronauts in the shuttle orbiter cabin, and employs graphic displays of the forces and

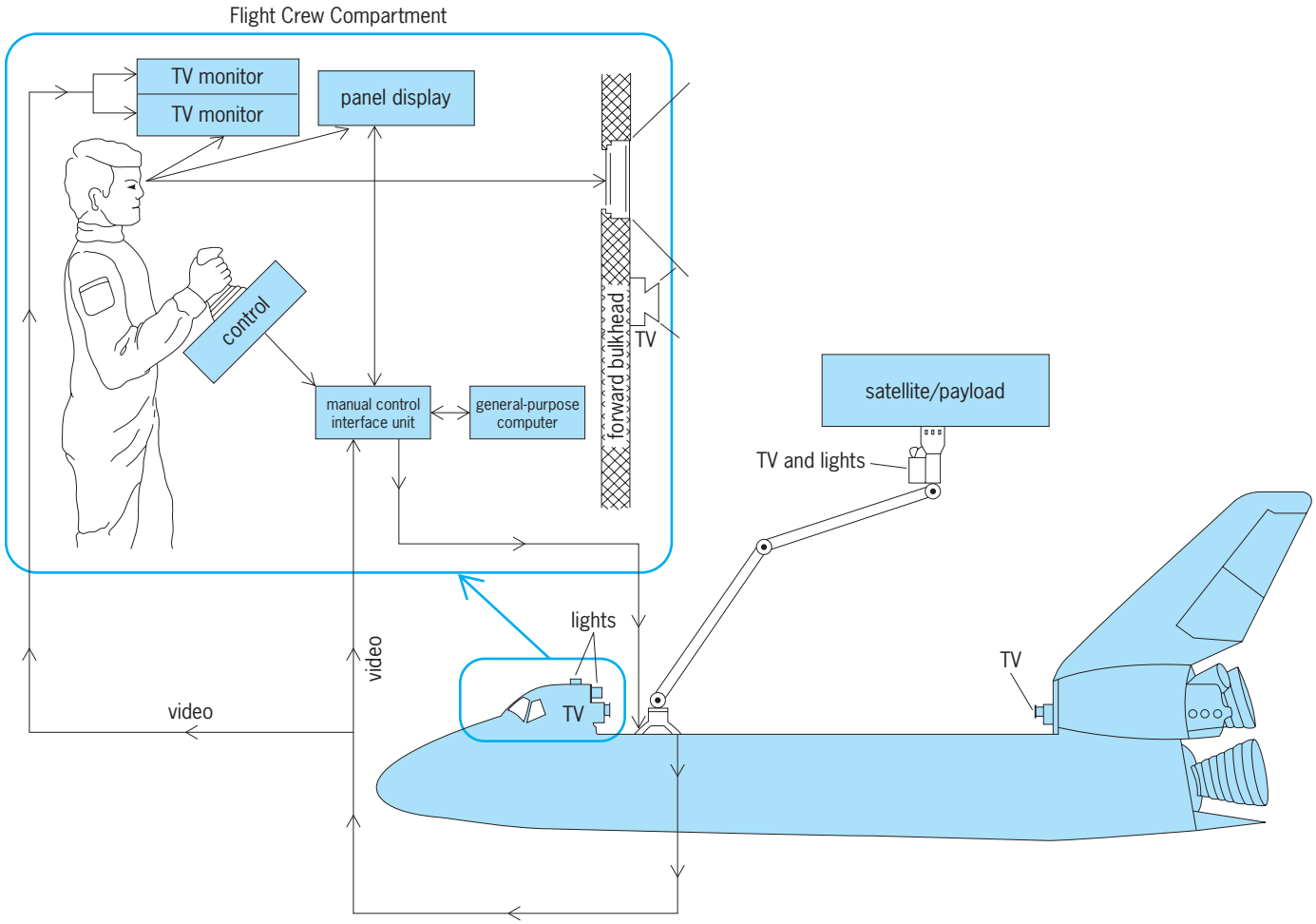

Fig. 2. Space shuttle orbiter with a remote manipulator arm.

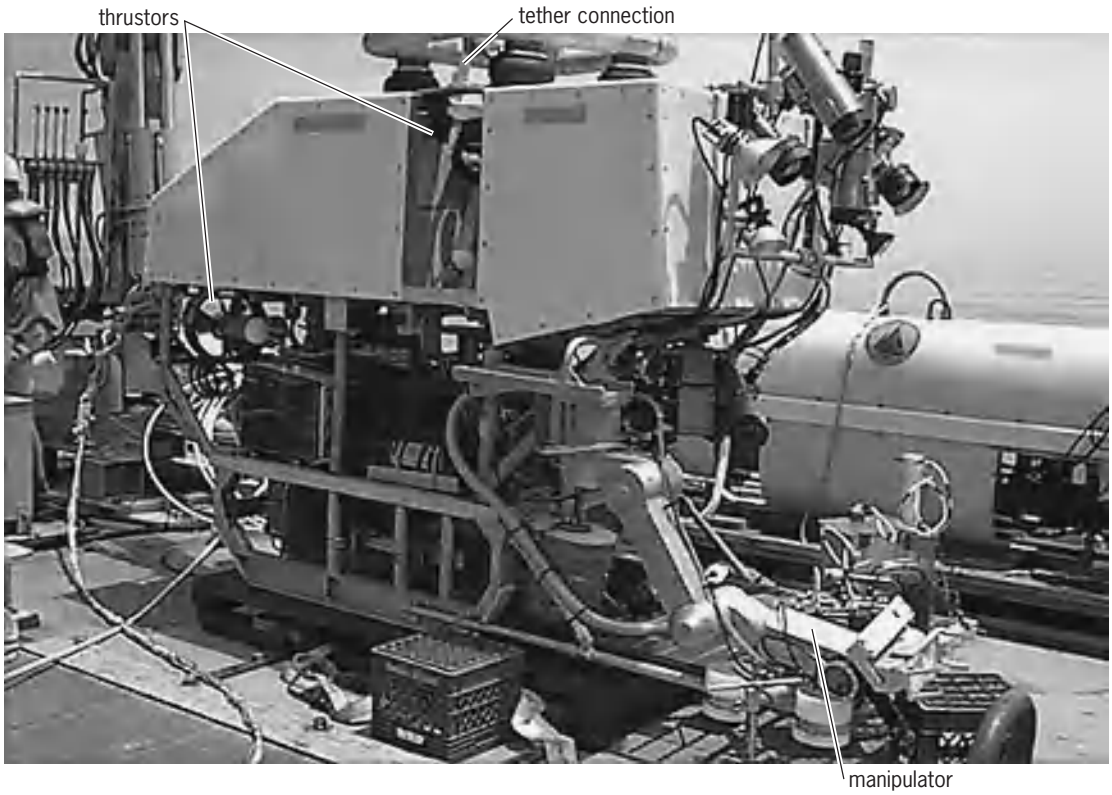

Fig. 3. ***JASON*** **remotely operated underwater vehicle. (*Woods Hole Oceanographic Institution*)**

torques being applied by the manipulator arm to the satellite. *See* SPACE SHUTTLE.

Crewless, remotely operated underwater vehicles are one of the fastest-growing applications of remote manipulators. These vehicles are used in oceanographic research and in the petroleum industry for inspecting and maintaining underwater well heads. The remotely operated underwater research vessel *JASON*, which is used in oceanographic research (**Fig. 3**), can operate at depths of 6000 m (20,000 ft). This vehicle has a manipulator that is remotely operated from a control station on a surface ship. The vehicle is connected to the surface ship by a tether that carries both command and feedback signals. *See* UNDERWATER VEHICLE.

Biomedical applications represent another rapidly growing area for remote manipulators. For example, abdominal surgery using minimally invasive (laparoscopic) procedures is increasingly common. A simplified representation of the procedure is shown in **Fig. 4**. The parallels between the general description of a remote manipulator (Fig. 1) and the surgical device are evident. Another biomedical application involving arthroscopic surgery is for human joint disease and injury, especially in professional sports. In this area, an endoscope, a device combining visual feedback, control, and remote manipulation, is used. This device typically consists of a fiber-optic bundle with additional tubes for moving fluids to or from the end point, and includes additional wires operating an end-point manipulation device to grip, incise, or cauterize diseased or injured tissue. Such endoscopic devices can perform surgery without large-scale tissue damage. *See* FIBER-OPTICS IMAGING; MEDICAL IMAGING; OPTICAL FIBERS; SURGERY.

Aiding the disabled is an important use of remote manipulators. One example is devices that

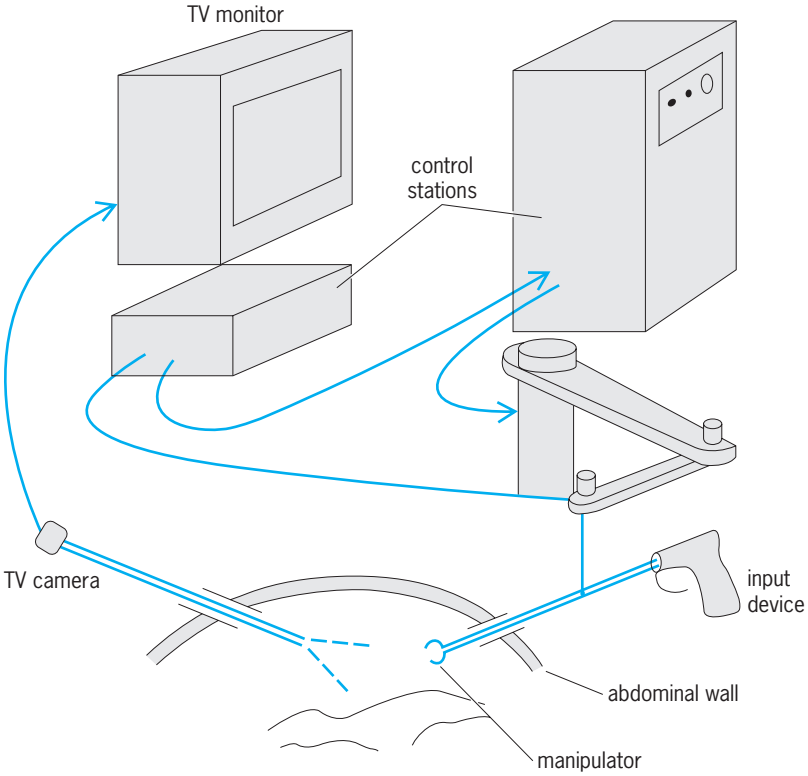

Fig. 4. Remote manipulation used in laparoscopic surgery.

allow individuals with little or no control of their upper extremities to feed themselves. Such devices have been referred to as teletheses (alluding to the extrasensory perception of distant objects). These devices are closely coupled to the user, acting as extensions of the human body. For example, an upper-body telethesis uses movements of the leg and foot as a control input. The manipulator itself is mechanical in nature, with an end effector (working component of a robot) consisting of a spoonlike utensil. Head-controlled feeding mechanisms have also been proposed for quadriplegics whereby head and neck movements will provide command inputs to the feeding manipulator. *See* ROBOTICS.

Devices that augment the strength of the human have been proposed for industrial applications. For example, a load-sharing manipulator has been envisioned in which human arm-manipulator coordination effectively allows the human to work with a "partner" that has considerably more strength. Another concept involves a "force extender." While not "remote" in the usual sense of the word, the manipulator is directed by the human and magnifies the forces supplied by the human operator. The human is thus able to move and manipulate masses considerably larger than those possible using muscle force alone.

Recreation of sensory input. The most crucial aspect of remote manipulation is the ability to recreate, with the appropriate realism, sensations that a human would perceive at a remote site—that is, to create a virtual reality so that the remote nature of the manipulating task is completely transparent to the operator. Significant progress is being made in this area. Appropriate re-creation of the physical cues (visual, aural, and tactile sensations) that the human would sense in the task is a major hurdle. If significant distance separates the remote manipulator from the human operator, delays can occur between the time that human actions are sent to the manipulator and the resulting manipulated-object changes are sensed and returned to the human. Delays of more than a few tenths of a second have been shown to destabilize a remotely manipulated task involving continuous human control inputs. Accommodating these delays is a difficult issue. *See* VIRTUAL REALITY.

Supervisory control. The necessity of continuous human inputs in the control of remote manipulators often results in a high degree of operator workload. Partial automation of some of these remote tasks is possible through the use of supervisory control by the human. Taking the remotely operated underwater vehicle as an example, supervisory control might be implemented by the human commanding the vehicle to move to a specified position near the ocean floor (perhaps at the head of an oil well). The vehicle's navigation to this position would be done without direct human intervention. More precise tasks, such as positioning the vehicle's manipulator, would then be performed by continuous human control. In such supervisory remote manipulation, the control station must be more sophisticated than in the case of continuous manipulation, and must receive and operate upon feedback information from the remote location. *See* CONTROL SYSTEMS; HUMAN-MACHINE SYSTEMS.

Ronald A. Hess

Bibliography. N. Moray, W. R. Ferrell, and W. B. Rouse (eds.), *Robotics Control and Society*, 1990; *NASA/Air Force Annual Conferences on Manual Control*, 1964–1986; T. B. Sheridan, *Telerobotics, Automation, and Human Supervisory Control*, 1992; L. I. Slutski, *Remote Manipulation Systems*, 1998.

Remote sensing

The gathering and recording of information about terrain and ocean surfaces without actual contact with the object or area being investigated. Remote terrain sensing is part of the large subject of remote sensing, which deals with the gathering and recording of information on many types of natural phenomena from a distance (see **table**). Remote sensing uses the visual, infrared, and microwave portions of the electromagnetic spectrum (**Fig. 1**).

Humans have always used remote sensing in a primitive form. In ancient times, they climbed a tree or stood on a hill and looked and listened, or sniffed for odors borne on the wind. Although it is possible to sense the environment without the use of instruments, it is not possible to record the sensed information without the aid of instruments referred to as remote sensors. These instruments are a modern refinement of the art of reconnaissance, an early example being the first aerial photograph taken in 1858 from a balloon floating over Paris.

The eye is sensitive only to visible light, a very small portion of the electromagnetic spectrum (Fig. 1). Cameras and electrooptical sensors, operating like the eye, can sense and record in a slightly larger portion of the spectrum. For gathering invisible data, instruments operating in other regions of the spectrum are employed. Remote sensors include devices that are sensitive to force fields, such as gravity-gradient systems, and devices such as antennas that record the reflection or emission of electromagnetic energy. Both passive electromagnetic sensors (those that rely on natural sources of illumination, such as the Sun) and active ones (those that utilize an artificial source of illumination such as radar) are considered to be remote sensors. Several remote-sensing instruments and their applications are listed in the table. *See* COLOR VISION.

Terrain Characteristics

Each type of surface material (for example, soils, rocks, vegetation, and ocean waves) absorbs and reflects solar energy in a characteristic manner depending upon its atomic and molecular structure (Fig. 1*c*). In addition, a certain amount of internal energy is

Some areas of application for remote sensor instruments

Technique	Agriculture and forestry	Geology and planetology	Hydrology	Oceanography	Geography
Visual photography	Soil types Plant vigor Disease	Surface structure Surface features	Drainage patterns	Sea state Erosion Turbidity Hydrography	Cartography Land use Transportation Terrain and vegetation characteristics Thematic mapping
Multispectral imagery	Same as above	Lithological units Formation boundaries	Soil moisture	Sea color Biological productivity	Same as above
Infrared imagery and spectroscopy	Vegetation extent Soil moisture Plant conditions	Thermal anomalies Faults	Areas of cooling Soil moisture	Ocean currents Sea-ice type and extent Sea-surface temperature	Thermal activity in urban areas Land use
Radar: imagery, scatterometry, altimetry	Soil characteristics Plant conditions	Surface roughness Structural framework	Soil moisture Runoff slopes	Sea state Tsunami warning	Land-ice type and extent Cartography Geodesy
Passive microwave radiometry and imagery	Thermal state of terrain Soil moisture		Soil moisture Snow and ice extent	Sea state Sea-surface temperature Sea-ice extent	Snow and ice extent

emitted which is partially independent of the solar flux.

The absorbed, reflected, and emitted energy can be detected by remote sensing instruments in terms of characteristic spectral signatures and images. These signatures can usually be correlated with known rock, soil, crop, terrain, or ocean surface features. Chemical composition, surface irregularity, degree of consolidation, and moisture content are among the parameters known to affect the records obtained by electromagnetic remote sensing devices. Selection of the specific parts of the electromagnetic spectrum to be utilized in terrain or ocean sensing is governed largely by the photon energy, frequency, and atmospheric transmission characteristics of the spectrum (Fig. 1).

Types of Sensing

Because remote sensing is a composite term which includes many types of sensing, its meaning can be best understood by describing several of the types. Remote sensing is generally conducted by means of remote sensors installed in aircraft and satellites, and much of the following discussion refers to sensing from such platforms. *See* SCIENTIFIC SATELLITES.

Photography. Photography is probably the most useful remote sensing system because it has the greatest number of known applications, it has been developed to a high degree, and a great number of people are experienced in analyzing terrain photographs. Much of the experience gained over the years from photographs of the terrain taken from

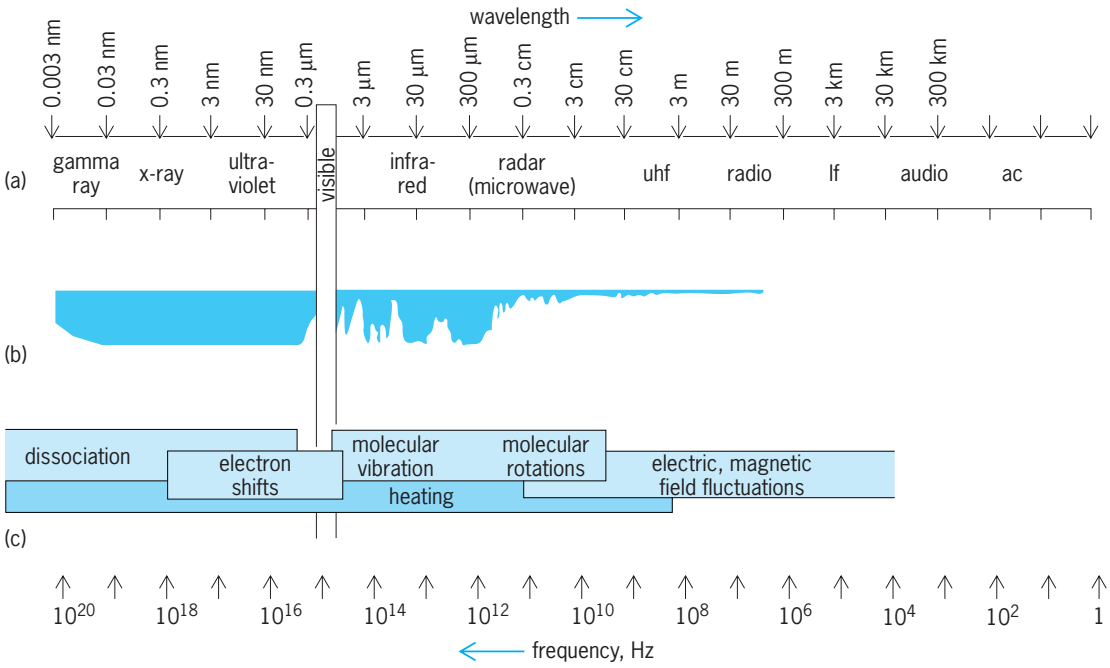

Fig. 1. Characteristics of the electromagnetic energy spectrum which are of significance in remote sensing. (*a*) Regions of electromagnetic spectrum. (*b*) Atmospheric transmission. The colored areas are regions through which electromagnetic energy is not transmitted. (*c*) Phenomena detected.

aircraft is being drawn upon for use in space. *See* PHOTOGRAPHY.

Conventional. Results of conventional photography experiments carried out in the short-duration crewed spacecraft such as a *Gemini*, *Apollo*, and *Skylab*, and in the long-duration uncrewed spacecraft such as *ERTS*, *Landsat*, *Nimbus*, *NOAA*, and *DMSP*, have vividly shown the applicability of these systems in space. These results indicate that space photography provides valuable data for delineating and identifying various terrain features such as those shown in **Fig. 2**.

The photomap in Fig. 2 was made by rectifying and fitting the photographs to the cultural plate of the existing line map of the area. The result combines all the original photographic detail with the cultural line data important for geographic orientation and location. This combination provides considerably more terrain data than conventional topographic or raised relief maps. A number of major copper deposits occur in this region, and their location is in part controlled by the presence of intense fracturing, as shown on the photomap. Also important is the presence of intrusive rocks, which sometimes show up as circular areas. Fault intersections, intrusive centers, and alteration halos are important clues in the search for new mineral deposits of this copper type. Detailed study of such high-quality spacecraft imagery is very useful for unraveling the structural framework of potential mineral districts.

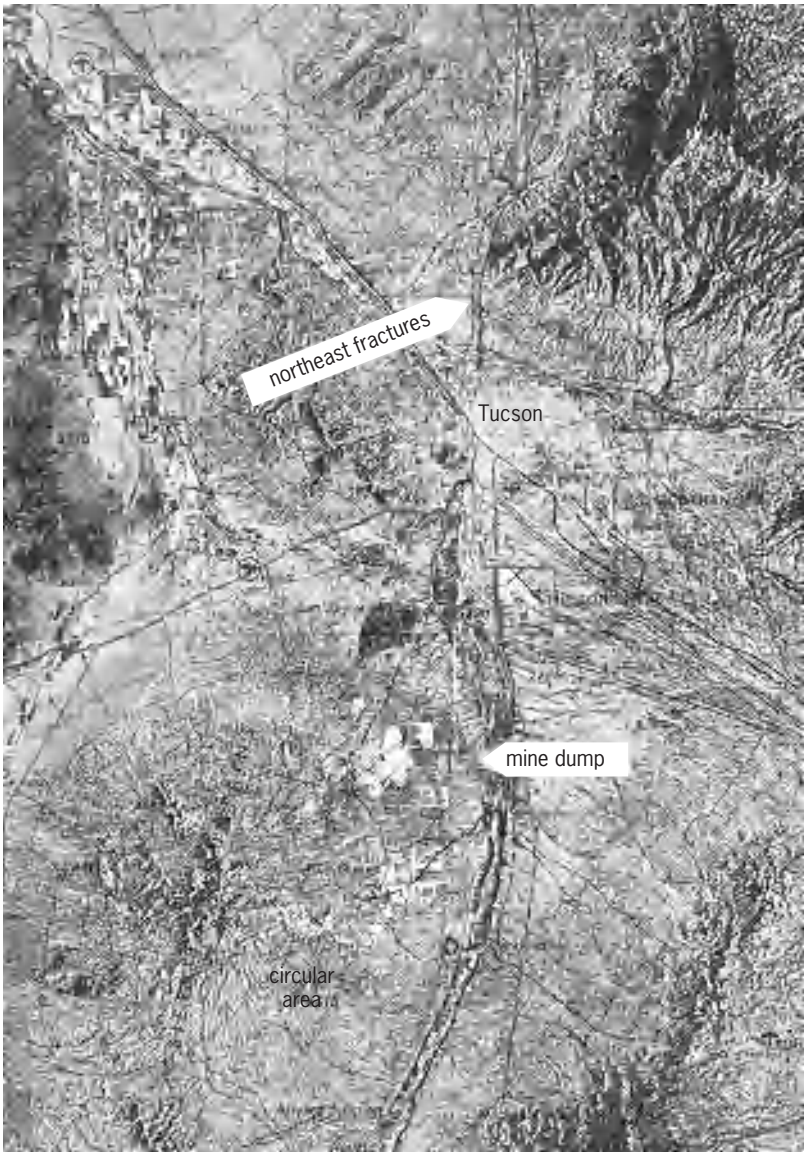

Fig. 2. Space photomap of the Tucson, Arizona, area, using 80-mm-focal-length Hasselblad cameras on the *Apollo 9* flight. The strong northeast-trending fractures in the mountains east of Tucson are apparent. The white patches on the northeast flank of the circular area are open pits and mine dumps, associated with already discovered mineral deposits.

Multispectral. Multispectral photography can be defined as the isolation of the reflected energy from a surface in a number of given wavelength bands and the recording of each spectral band separately on film. This technique allows the scientist to select the significant bandwidths in which a given area of terrain displays maximum tonal contrast and, hence, increases the effective spectral resolution of the system over conventional black-and-white or color systems.

Because of its spectral selectivity capabilities, the multispectral approach provides a means of collecting a great amount of specific information. In addition, it has less sensitivity to temperature, humidity, and reproduction variables than conventional color photography and retains the high resolution associated with broadband black-and-white mapping film.

Multispectral imagery. An advance in multispectral sensing is the use of multispectral scanning systems which record the spectral reflectance by photoelectric means (rather than by photochemical means as in multispectral photography) simultaneously in several individual wavelengths within the visual and near-infrared portions of the electromagnetic spectrum (0.47–1.11 micrometers). Such instruments have been used in aircraft since the early 1960s and were used in the *Earth Resources Technology Satellite* (*ERTS 1*) and in *Landsat B*.

In the satellite cases, optical energy is sensed by an array of detectors simultaneously in four spectral bands from 0.47 to 1.1 μm. As the optical sensors for the various frequency bands sweep across the underlying terrain in a plane perpendicular to the flight direction of the satellite, they record energy from individual areas on the ground. The size of these individual areas (instantaneous fields of view) is determined by the resolution capability (spot size) of the optical scanner. The smallest individual area distinguished by the scanner is called a picture element or pixel, and a separate spectral reflectance is recorded in analog or digital form for each pixel. A pixel covers about 1 acre (approximately 0.4 hectare, or 4047 m^2) of the Earth's surface in the case of the *Landsat*, with approximately 7.5×10^6 pixels composing each *Landsat* image (115 × 115 mi or 185 × 185 km in area).

Such multispectral scanning systems have a number of advantages over conventional photography: the spectral reflectance values for each pixel can be transmitted electronically to ground receiving stations in near-real time, or stored on magnetic tape in

the satellite until it is over a receiving station. When the signal intensities are received on the ground, they can be reconstructed almost instantaneously into the virtual equivalent of conventional aerial photographs (provided the resolution size of the individual pixels is fine enough). In the case of conventional photography, the entire film cassette generally has to be returned physically to the ground, before the film can be developed and studied. Such physical return of film cassettes is not practical for long-duration satellites such as *ERTS* which are continuously recording vast amounts of data. *See* FACSIMILE.

Computer enhancement of imagery. There are a number of other advantages of multispectral scanning systems. Since they record reflected radiation energy in a number of discrete wavelengths, these radiance values can be used singly or combined by digital computer processing to provide a response optimized for particular terrain features.

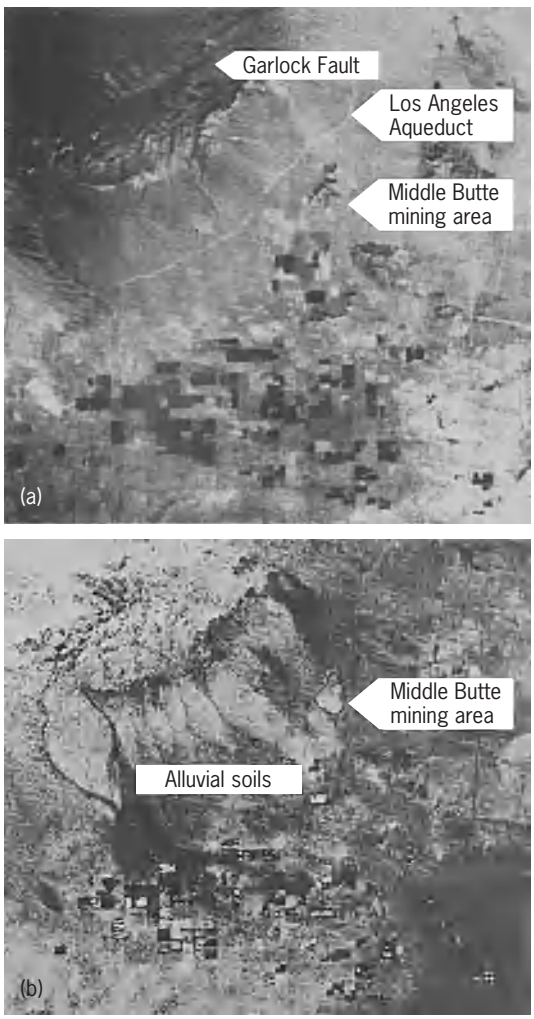

Fig. 3. Images taken by the earth resources technology satellite over the Mojave Desert of California on October 21, 1972: (*a*) band 4 (0. 5–0.6 μm); (*b*) a ratio of band 5 (0.6–0.7 μm) to band 4. The alluvial soils and Middle Butte mining area appear as bright areas on image *b*, while they are difficult to distinguish on image *a*. (*From P. M. Merifeld et al., Enhancement of Geological Features near Mojave, California by Spectral Band Ratioing of ERTS MSS Data, report prepared under U.S. Geological Survey Contract 14-08-0001-13911, 1974*)

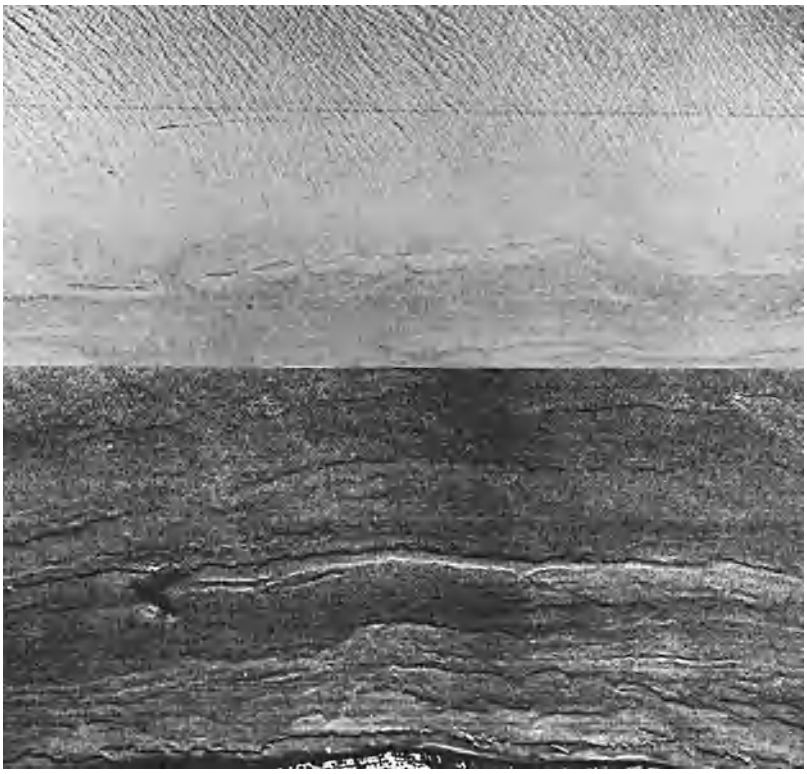

Fig. 4. Ratio image (0.50–0.54 μm)/(0.48–0.52 μm) from the Florida coast, acquired by an airborne multispectral scanner operating in the visible wavelengths. The reflection from the sea surface has been filtered out in the lower image, and the absorption effects of the water column have been removed by means of an algorithm. The residual reflections are from the sea floor and are indicative of surface roughness. (*Environmental Research Institute of Michigan, under contract to the U.S. Navy*)

One type of such computer enhancement which has proved to be very useful involves the preparation of ratio images. The individual spectral responses are ratioed, picture element by picture element using a computer, and they can also be contrast-stretched to further enhance the spectral differences. Such ratio images are very useful in detecting such items as subtle soil changes (**Fig. 3**) and hydrothermally altered areas (important for mineral exploration), as well as in discriminating most major rock types, and in detecting the nature of the sea floor (**Fig. 4**).

Another advantage of the multispectral scanner is its ability to extend the spectral coverage (beyond the 0.9-μm cutoff available with conventional photographic films) into the infrared wavelengths of the electromagnetic spectrum. *See* IMAGE PROCESSING.

Infrared. Infrared is electromagnetic radiation having wavelengths of 0.7 to about 1400 μm. All materials continuously emit infrared radiation as long as they have a temperature above absolute zero in the Kelvin scale. This radiation involves molecular vibrations as modified by crystal lattice motions of the material. The total amount and the wavelength distribution of the infrared radiation are dependent on two factors: the temperature of the material and its radiating efficiency (called emissivity). *See* EMISSIVITY; INFRARED RADIATION.

Thermal infrared radiation is mapped by means of infrared scanners similar to the multispectral

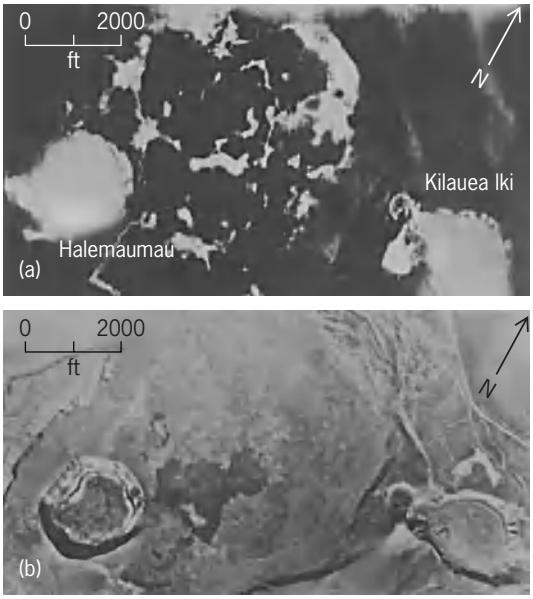

Fig. 5. Kilauea Volcano, Hawaii. (*a*) Infrared image. Tonal variation shows distribution of radiant heat; the brighter the tone, the warmer the surface. (*b*) Photograph by AT-11 mapping camera. Two vents left and right are within caldera. (*U.S. Geological Survey*)

scanners described previously, but in this case radiated energy is recorded generally in the 8–14-μm portion of the electromagnetic spectrum.

The imagery provided by an infrared scanning system gives information that is not available from ordinary photography or from multispectral scanners operating in the visual portion of the electromagnetic spectrum. The brightness with which an object appears on an infrared image depends on its radiant temperature. The hotter the object, the brighter it will appear on a positive image (**Fig. 5**). Radiant temperature is largely dependent upon chemistry, grain size, surface roughness, and thermal properties of the material. The time of day or night and the season of the year at which the infrared energy is recorded are also important factors, particularly when water or moisture exists in the near-surface terrain materials, since the water has a different thermal inertia relative to rock and soil materials and can therefore appear as a thermal anomaly. Because moisture frequently collects in geologic faults and fractures, infrared images are useful in detecting such features (**Fig. 6**).

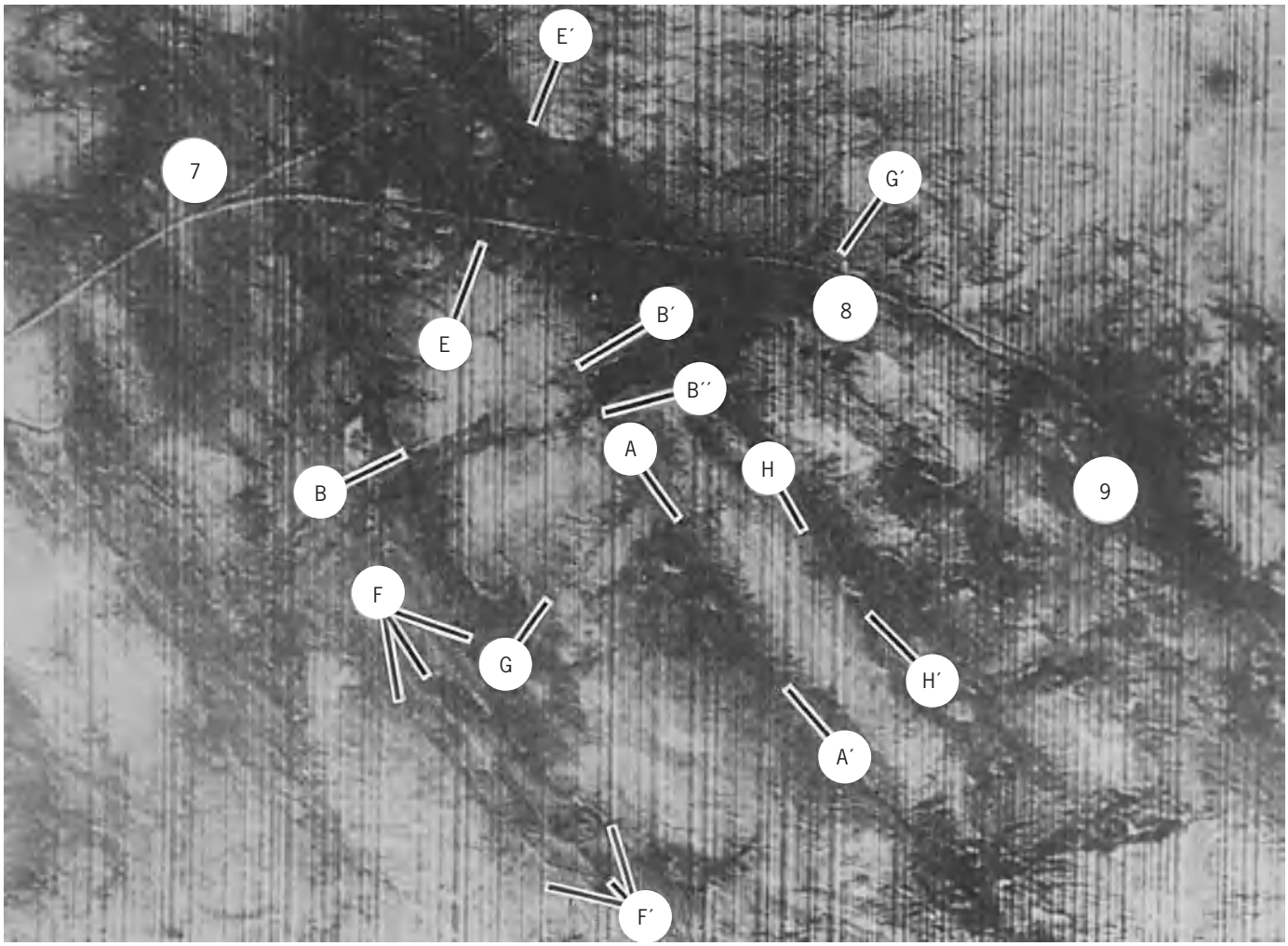

Fig. 6. Thermal infrared image of part of the Goldfields, Nevada, mineral district. Moisture is greater in the fracture zones of this district, which show up as dark bands mainly trending NW-SE (labeled FF′, AA′, HH′). A few less well-developed fractures trend ENE (labeled BB′ and BB″). In this district, commercial mineralization is best developed in fracture zones and in areas where hydrothermal alteration is most intense.

Although it is possible to obtain relatively high resolutions (spot sizes of a few feet) from infrared imagers in low-flying aircraft (Fig. 6), it has been more difficult to achieve such fine resolutions from satellite altitudes, partially because of the lower sensing-element sensitivity of infrared scanners (whose sensing elements must be cooled, in contrast to visible spectrum sensors), and partially because there is always a trade-off between field of view (generally large for satellite imagery as compared with aircraft imagery) and resolution. The larger the field of view, the more difficult it is to obtain fine resolution. Despite these restrictions, routine thermal mapping from long-duration satellites is now on the order of 2300–3000 ft (700–900 m) per picture element on the ground (*Nimbus 5* and *NOAA 2* satellites). The multispectral scanner on the short-duration *Skylab* space flight had an 260-ft (80-m) instantaneous field of view and recorded temperature differences of 0.40 K.

In the past, thermal infrared images were generally recorded on photographic film. Videotape records are replacing film as the primary recording medium and permit better imagery to be produced and greater versatility in interpretation of data.

Thermal mapping from satellite altitudes is proving to be useful for a number of purposes, one of which is the mapping of thermal currents in the ocean (**Fig. 7**). Such currents change their position quite frequently; in many cases these changes occur daily or even more frequently. The positions of such thermal boundaries are important for a variety of uses (fishery, naval, and so forth).

Thermal infrared mapping (thermography) from aircraft and satellite altitudes has many other uses also, including the mapping of volcanic activity (Fig. 5) and geothermal sites, location of groundwater discharge into surface and marine waters, and regional pollution monitoring.

Microwave radar. This type of remote sensing utilizes both active and passive sensors. The active sensors such as radar supply their own illumination and record the reflected energy. The passive microwave sensors record the natural radiation. A variety of sensor types are involved. These include imaging radars, radar scatterometers and altimeters, and over-the-horizon radar using large ground-based antenna arrays, as well as passive microwave radiometers and images. *See* MICROWAVE; RADAR.

One of the most significant advantages of these instruments is their all-weather capability, both day and night. The active radars also possess a certain amount of foliage-penetration ability which is valuable in jungle terrains (**Fig. 8**). This imagery illustrates the unique capability of side-looking radar airborne to penetrate both cloud cover and dense jungle vegetation and to map the structural fabric of the underlying terrain with considerable detail. It has been demonstrated that radar return amplitude is affected by the composition of the illuminated area, its moisture content, vegetation extent and type, surface roughness, and even temperature in certain circumstances.

Radar returns are recorded in various forms to aid in their analysis. The forms having primary geoscience interest can be placed in one of the following categories: (1) scattering coefficients, which are a powerful tool in studying the nature of radar return and which have been correlated with different terrain types and directly applied to oceanic surface studies; (2) altimetry data; (3) penetration measurements, which have been utilized for mineral exploration, for example, detection of faults through moisture associated with them; and (4) radar images.

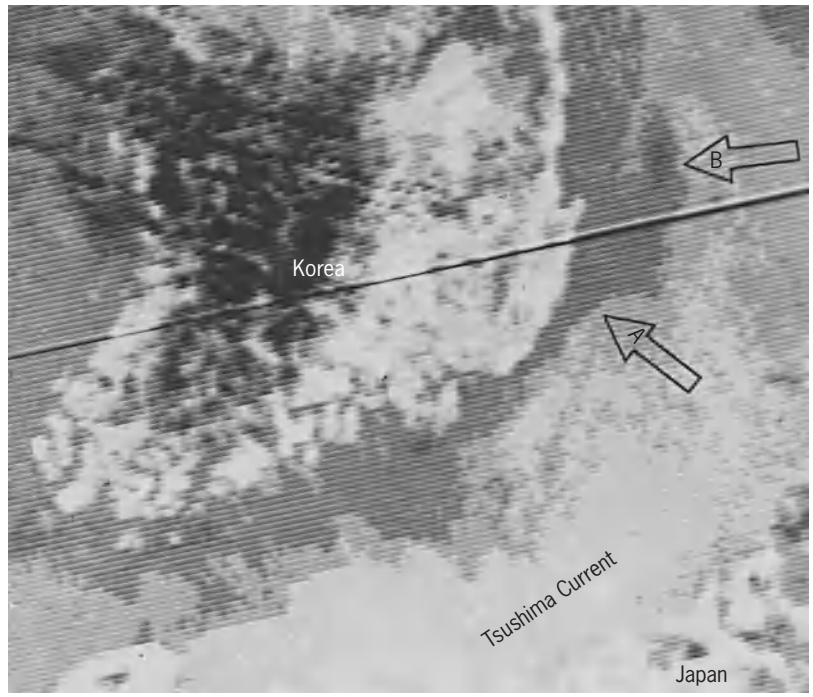

Fig. 7. Thermal infrared image of the Korea Straits, acquired by the Defense Meteorological Satellite Program (DMSP) on October 13, 1972. DMSP infrared-sensor ground resolution is 2 nautical miles (4 km). The light-colored waters (warm) are part of the Tsushima Current, while the darker gray areas (cool) are the coastal shelf waters. An oceanic "front" forms the boundary between the continental shelf waters and the Tsushima Current. The front turns seaward at A, and a cyclonic eddy is located at B.

Of these techniques, imagery generally presents the optimum information content for terrestrial geologic purposes, partly because well-developed photographic interpretation techniques are applicable to radar image analysis. These images are especially valuable for delineating various structural phenomena, as shown in Fig. 8.

The image record of the terrain return is affected by the frequency, angle of incidence, and polarization of the radar signal. For example, if the terrain being imaged is covered by vegetation, a K-band (35 GHz) signal records the vegetation, whereas a P-band (0.4 GHz) signal is likely to penetrate the vegetation and thus record a combination of vegetation and soil surface. In general, each frequency band represents a potential source of unique data. The angle of incidence of the incident wave affects the image because of radar shadowing on the back side of protruding objects. The angle of incidence can also show differences because of changes in orientation of the many facets on a surface which affect the return strength. Although information is lost in the shadow region, the extent of the shadow indicates the height of the object and therefore has been

Fig. 8. Side-looking radar (SLAR) imagery of a portion of the jungle-covered Guayana Precambrian Shield of Venezuela. This imagery was acquired by a Goodyear synthetic-aperture X-band (3.12-cm) radar system mounted on board an Aero Service Caravelle aircraft. The flight altitude was 46,000 ft (14,030 m). The east-west dimension of the image mosaic is approximately 104 mi (168 km).

useful in emphasizing linear features such as faults and lineations reflecting joint systems.

Although imaging radars have been flown on aircraft for a number of years, it was not until 1978 that the first space-borne radar imagery became available with the launching of *Seasat* on June 28, 1978. This research satellite carried a radar altimeter, a radar scatterometer, a visible and infrared radiometer, and a scanning multichannel microwave radiometer, in addition to a synthetic-aperture radar (SAR) imaging in the L-band (1.275 GHz) with a swath width of 62 mi (100 km). *Seasat* failed in orbit on October 10, 1978, as a result of a massive short circuit in the elec- trical system. Fortunately, a considerable amount of data was recorded from several experimental test areas. **Figure 9** is a mosaic of *Seasat* radar imagery taken over the Cape Cod, Massachusetts, area. Internal waves are readily apparent. These are waves with lengths of several hundred meters or more between crests, and they originate below the surface of the ocean due to internal slippage between water layers of different density. The amplitude of the waves is greatest at the density discontinuity. They are generally well developed in coastal waters. Such waves have been noted on a number of the *Seasat* SAR images. The brighter areas on the bottom of Fig. 9 are believed to be due to higher winds in these areas, although no detailed analysis of this image supported by accompanying sea surface truth data has been published. Additional airborne and satellite flights are planned with SAR over controlled test sites because the optimum flight trajectories (azimuths), depression angles, wavelengths (X-, L-band, and so on), and polarizations (HH, HV, and so on) needed for extracting the full data content from the signals received are not fully understood. Although the analysis of oceanographic data returns from SAR are still in their infancy, it is already clear that SAR will yield very valuable information because of its high resolution and all-weather (cloud-penetrating) capability.

High-frequency (hf) radar. Such radars utilize frequencies in the 3-30-MHz portion of the electromagnetic spectrum (median wavelength of about 20 m) and are thus not within the microwave part of the spectrum. The energy is transmitted by ground-based antennas in either a sky-wave or surface-wave mode. In the sky-wave mode, the energy is refracted by the various ionospheric layers back down to the Earth's surface some 500–1800 mi (800–3000 km) (on a single-hop basis) away from the hf radar antenna site. The incident waves are reflected from such surface features as sea waves, and the reflections are enhanced when the wavelength of the

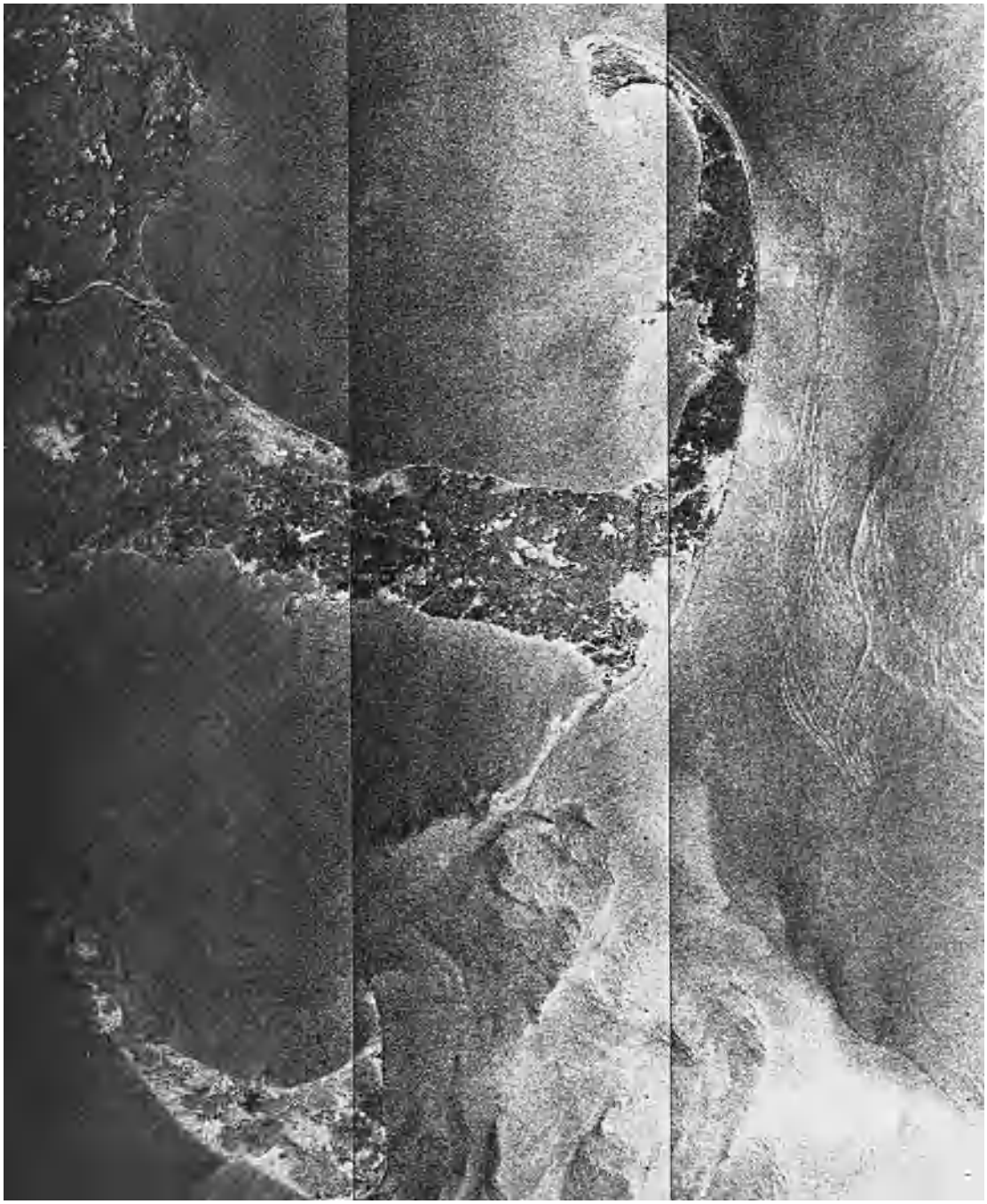

Fig. 9. Mosaic view of part of the Cape Cod, Massachusetts, area acquired by synthetic-aperture radar on board *Seasat*, 1978. The top of the figure is north. Nantucket Island is in the lower left portion of the figure. Internal waves are visible east of Cape Cod. The bright area in the lower part of the figure is believed to be roughened sea due to stronger wind action in this area. (*Johns Hopkins University Applied Physics Laboratory, under contract to the U.S. Navy*)

sea waves is twice the wavelength of the incident radar waves (Bragg reflection effect). In addition, the Doppler principle is applied to determine the component of sea-wave velocities moving toward or away from the radar. Spot sizes observed on the ocean surface by this sky-wave method are now as fine as 9 mi (15 km) in azimuth by 2 mi (3 km) in range, and it is expected that these resolution cells can be reduced by further development work.

In addition to this experimental use of hf radars for remote sensing of sea state, they are being used as telemetry (communication) and direction and range-finding links in conjunction with disposable buoys or drogues. The buoys are used to detect the sea state, surface and subsurface temperature, salinity, and so forth. The information is relayed back to the interrogating site by either sky-wave or surface-wave modes. The surface waves are that portion of the hf wavefront that follows closely along the Earth's surface. Ranges of 60–180 mi (100–300) km or so are possible with this mode, and the buoys can be interrogated and located by low-cost direction-finding

loops and commercial battery-powered receivers. The big advantage of these ground-based hf radars is that they can maintain continuous monitoring of an ocean area, while satellite-borne radars can monitor only the areas that they happen to be flying over at the time. In the case of the *Landsat* system, the satellite comes over (revisits) the same area about every 18 days, while in the case of the *Defense Meteorological Satellite Program (DMSP)* craft the frequency of revisits can be up to several times daily, depending upon how many satellites are in orbit at one time. Neither the *Landsat* nor *DMSP* carry all-weather radar systems, however. *See* METEOROLOGICAL SATELLITES.

Peter C. Badgley

Bibliography. American Society of Photogrammetry, and Remote Sensing Staff and American Congress on Surveying and Mapping, *Remote Sensing*, vol. 1, 1987; E. A. Beaumont and N. H. Foster (eds.), *Remote Sensing*, 1992; A. Chedin (ed.), *Microwave Remote Sensing of the Earth System*, 1989; A. P. Cracknell and L. W. Hayes, *Introduction to Remote Sensing*, 1990; P. Curran, K. Kondratyev, and V. Kozogerov, *Remote Sensing of Soils and Vegetation*, 1990; G. P. De Loor (ed.), *Radar Remote Sensing of the Atmospheres and the Oceans*, 1984; S. A. Drury, *A Guide to Remote Sensing: Interpreting Images of the Earth*, 2d ed., 1998; G. M. Foody and P. J. Curran (eds.), *Environmental Remote Sensing Data Systems and Networks*, 1995; M. T. Halbouty, Geologic significance of Landsat data for 15 giant oil and gas fields. *Amer. Ass. Petrol. Geol.*, 64(1):8–36, 1980; P. M. Mather (ed.), *TERRA-2, Understanding the Terrestrial Environment: Remote Sensing Data Systems and Networks,*, 1990; R. M. Measures, *Laser Remote Sensing: Fundamentals and Applications*, 1984, reprint 1992; P. Pamploni, *Microwave Radiometry and Remote Sensing Applications*, 1989: J. Vernberg (ed.), *Processes in Marine Remote Sensing*, 1981.

Renewable resources

Agricultural materials used as feedstocks for industrial processes. For many centuries agricultural products were the main sources of raw material for the manufacturing of soap, paint, ink, lubricants, grease, paper, cloth, drugs, and a host of other nonfood products. During the early 1900s, the advances in organic synthesis in western Europe and the United States led to the use of coal as an alternative resource; in the 1940s, oil and natural gas were added as starting materials as a result of great advances in catalysis and polymer sciences. Since then the petrochemical industry has grown rapidly as the result of the abundance and low price of the starting materials as well as the development of new products, such as nylon and other polymers, water-based paints, and detergents. However, with the rapidly increasing economies of the nations of the world, these developments did not ever result in reduction in the utilization of agricultural products as industrial materials. Although specific nonfood crops, such as linseed or cotton, may have suffered from these developments, animal fats, vegetable oils, starch, cellulose, and other renewable agricultural resources still are used extensively in industry, and many new uses have been developed since the 1950s.

Price increases for oil and natural gas, as well as the recognition that many major industrial nations depend on imported agricultural raw materials, have resulted in an upsurge in research and development. Private companies have been disposed to reconsider domestic agricultural resources as alternative materials for manufacturing processes.

Fats and oils. Animal fats, marine and vegetable oils, and their fatty acid derivatives have always played a major role in the manufacturing of many industrial products. Some of these commodities are produced solely for industrial end uses; examples are linseed, tung, castor (not counting minor amounts used for medicinal purposes), and sperm whale oils. Others, such as tallow and soybean oil, are used for both edible and industrial products.

Most fats and oils are triglycerides; that is, they consist of a glycerine molecule esterified with three identical or different fatty acids. Some are used in their original state, with minor chemical modification or formulation. Examples are drying oils in paints (linseed, tung) and castor and rapeseed oils as lubricants. Hydrogenation of fats and oils leads to solid products for candles and waxes. Soybean and linseed oils are also epoxidized for use as stabilizers in poly(vinyl chloride), or made into alkyds for paints.

A large proportion of fats and oils are split to glycerine and fatty acids or converted into fatty methyl esters. They, in turn, are used directly or converted to fatty amines or fatty alcohols. Monounsaturated fatty acids, such as oleic or erucic acid, are split further by oxidation of the double bond to produce a mixture of mono- and dicarboxylic acids. Oleic acid is thus converted to pelargonic and azeleic acids, and erucic acid to pelargonic and brassylic acids. All of these products serve as starting materials for manufacture of soaps, detergents, lubricants, and plasticizers. The energy crises of the 1970s led to serious consideration of vegetable oils and their simple esters as emergency fuels to replace diesel fuel, especially for farming operations. *See* DETERGENT; FAT AND OIL; LUBRICANT; SOAP.

Carbohydrates. Starch, cellulosics, and gums also have been used for many centuries as industrial materials, whereas sugar crops, such as sugarcane and sugarbeet, have mainly satisfied world food needs.

Starch. Starches from potato, maize, and many other major crops have been used as thickening agents and adhesives for a variety of industrial applications and as substrates in fermentation and enzyme-mediated processes for the manufacturing of many products. Unmodified as well as derivatized starches play a major role in the production

of paper and textiles. The most prominent fermentation product from starch is industrial ethanol, used both as solvent and as fuel extender in gasoline (gasohol). Economic and technical developments have led to fermentation alcohol becoming cheaper than ethanol from petrochemical ethylene. Research since the 1940s has created new industrial products, including the so-called Super Slurper and related water-absorbing polymers (graft copolymers of starch and acrylonitrile), insoluble starch xanthate (used to clean up wastewater from the metal-plating industry), water-soluble laundry bags, biodegradable plastics, and encapsulating agents for pesticides. *See* ETHYL ALCOHOL; STARCH.

Cellulose. The main use of cellulose has been in fiber form for manufacturing paper and textile products, and the principal sources are hard- and softwoods. Other prominent sources are cotton and flax, as well as bagasse and straw from agricultural residues. Although petrochemical research has led to the production of many types and large quantities of synthetic fibers, manufacture of cellulose-based products has not decreased appreciably worldwide. Cellulose has also been modified chemically for manufacturing plastics (celluloid), fibers (rayon viscose), and thickening agents (carboxy methyl cellulose and so on). However, the development of synthetic plastics and fibers has drastically cut the use of cellulose for these purposes. *See* MANUFACTURED FIBER; PAPER; TEXTILE.

Since cellulose is so abundant in nature, much effort has been expended to explore the potential of this material for production of glucose as the starting material for fermentation processes. Although this is easy to accomplish with starch, so far no economical route for preparing glucose from cellulose has been developed. In addition to the cellulose-to-glucose conversion being more difficult, cellulose normally is closely associated with lignin, which in itself is a rather inaccessible polymer. However, once achieved, such conversion would revolutionize utilization of cellulose as well as of hemicellulose, a polymer of pentose sugars, and also a major constituent of plant materials. *See* CELLULOSE; HEMICELLULOSE; LIGNIN.

Gum. Plant gums, including guar, tragacanth, and gum arabic, have served as thickening agents for a long time and for many purposes. Most industrial nations depend on imports for these materials. They can now be supplemented or replaced by microbial polysaccharides, such as xanthan gum, which are made by fermentative conversion of starch. One prominent use of gums is in drilling fluids for secondary oil recovery. *See* GUM.

Hydrocarbons. Natural rubber and turpentine are excellent examples of plant-derived hydrocarbons. The development of synthetic rubbers during and after World War II has never threatened the demand for natural rubber; there is generally a world shortage. Turpentine is a product of the wood and paper pulp industry and is used as a solvent and thinner in paints and varnishes. *See* PINE TERPENE; RUBBER.

Novel crops. The threat that industrial nations might be separated from part or all of their traditional sources of raw materials through political and economic upheavals or natural calamities has resulted in a renewed effort to develop additional crops for local agriculture. In the United States, research has provided a number of candidate species that either are now in commercial development or are ready for the time when circumstances warrant such development. Examples are jojoba (liquid wax ester to replace sperm whale oil), guayule (alternate source of natural rubber), kenaf (paper fiber with annual yields much higher than available from trees), and crambe and meadowfoam (long-chain fatty acids, since erucic acid is no longer available from rapeseed oil). There is also active research involving *Cuphea* species (alternate source of lauric and other medium-chain fatty acids, to augment coconut oil), *Vernonia* (source of epoxy oil), and several other promising plants. For example, the Chinese tallow tree has the potential of producing 2.2 tons per acre (5 metric tons per hectare) of seed oil that could be used for manufacturing fuel and other chemicals.

L. H. Princen

Bibliography. K. R. Anthony, J. Meadley, and G. Robbelen (eds.), *New Crops for Temperate Regions*, 1993; K. T. Bird and P. H. Benson, *Seaweed Cultivation for Renewable Resources*, 1987; F. D. Gunstone, *The Chemistry Of Oils And Fats: Sources, Composition, Properties, and Uses*, 2004; D. L. Klass, *Biomass for Renewable Energy, Fuels, and Chemicals*, 1998; G. Pahl, *Biodiesel: Growing a New Energy Economy*, 2005; J. Tickell, K. Tickell, and K. Roman (eds.), *From the Fryer to the Fuel Tank: The Complete Guide to Using Vegetable Oil as an Alternative Fuel*, 2000.

Renner-Teller effect

The splitting, into two, of the potential function along the bending coordinate in degenerate electronic states of linear triatomic or polyatomic molecules. Most of the areas and methods of molecular physics and spectroscopy assume the validity of the Born-Oppenheimer approximation. The nuclei generally move much more slowly than the electrons, the frequencies associated with electronic transitions are much higher than vibrational frequencies, and one can consider separately the three types of molecular motion: electronic, vibrational, and rotational. These statements are no longer necessarily valid for electronic states which are degenerate or at least close to degeneracy, and the Born-Oppenheimer approximation breaks down.

Degenerate electronic states usually occur in molecules having a high degree of symmetry. The symmetric equilibrium geometry which causes the electronic degeneracy is, in general, lowered in the course of molecular vibrations, and this may lead

to splitting of the potential. The molecular potential is usually expressed in terms of a polynomial expansion in displacements r, and, in nonlinear molecules, the linear terms may lead to coupling of the electronic and vibrational degrees of freedom. The resulting breakdown of the Born-Oppenheimer approximation is in this case known as the Jahn-Teller effect. In linear molecules the symmetry is lowered during bending vibrations. In the bending potential the linear (and other odd) terms are zero by symmetry. The first nonvanishing terms which can couple the degenerate electronic states are quadratic in the bending coordinate. The results of this coupling in linear molecules are referred to as the Renner-Teller effect, or simply the Renner effect. *See* JAHN-TELLER EFFECT.

Splitting of electronic potential. In a linear tri-atomic molecule in a Π electronic state, the twofold degeneracy will be lifted whenever, in the course of the bending vibration, the nuclei assume a nonlinear configuration. As was first pointed out by G. Herzberg and E. Teller and shown in detail by R. Renner, the electronic potential function will split into two (see **illus.**). As noted above, the potential energy expansion in terms of displacement r along the bending coordinate can, for reasons of symmetry, contain only even terms, as in Eq. (1). In the

$$V = ar^2 + br^4 + \cdots \tag{1}$$

presence of the vibronic coupling of the electronic states, the potential will be split into components usually denoted by V^+ and V^-, as in Eq. (2), where

$$V^{\pm} = \left(a \pm \frac{\alpha}{2}\right) r^2 + \left(b \pm \frac{\beta}{2}\right) r^4 \tag{2}$$

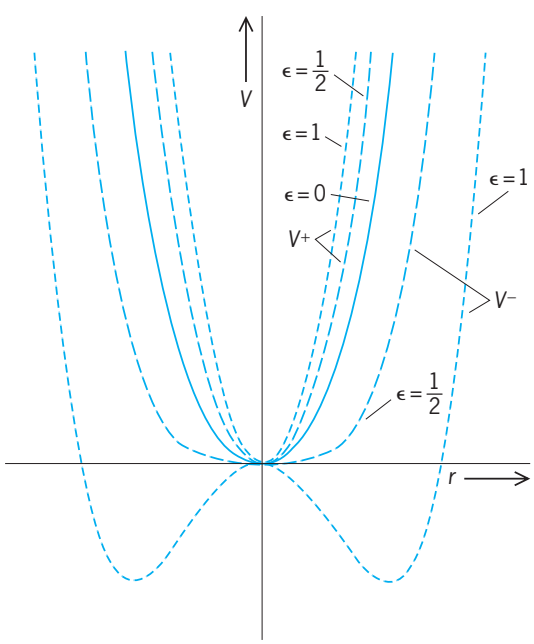

Bending potential function V for a degenerate electronic state of a linear molecule as a function of displacement r, and for three values (0, 0.5, 1.0) of the Renner parameter ϵ. In the absence of Renner-Teller splitting ($\epsilon = 0$) there is a single curve. When $\epsilon \neq 0$ the degenerate state splits into two components, V^+ and V^-.

α and β are the coupling terms. Often it is sufficient to consider only the quadratic coupling and neglect higher terms in the expansion.

Splitting of vibronic levels. The splitting of the electronic potential will, of course, affect the molecular vibrations and will result in splitting of the individual vibrational levels. Renner obtained approximate formulas for calculating the vibronic term values for a given level with vibrational quantum number v_2 and vibronic angular momentum K when the vibronic interaction is relatively small and the quartic terms negligible, Eqs. (3)–(5). Here ω_2 is the bending frequency and the Renner parameter, $\epsilon = \alpha/2a$, is a measure of the strength of the vibronic interaction.

$$G^{\pm}(v_2, K) = \omega_2\sqrt{1 \pm \epsilon}(v_1 + 1) \qquad K = 0 \tag{3}$$

$$G^{\pm}(v_2, K) = \omega_2[(v_2 + 1) - {}^1\!/_8\epsilon^2 K(K + 1)] \qquad v_2 = K - 1, K \neq 0 \tag{4}$$

$$G^{\pm}(v_2, K) = \omega_2(1 - {}^1\!/_8\epsilon^2)(v_2 + 1) \pm {}^1\!/_2\epsilon\sqrt{(v_2 + 1)^2 - K^2} \qquad v_2 > K - 1 \tag{5}$$

A well-studied example of Renner-Teller effect is the excited $A^1\Pi_u$ state of the C_3 radical. The observed vibronic levels conform relatively well to those calculated from Eqs. (3)–(5), even though the interaction is rather large in this case. When the vibronic interactions are very large, the vibronic levels will not be well described by Eqs. (3)–(5). In this case the energy levels of the molecule are best determined by setting up and diagonalizing the appropriate hamiltonian matrix.

Bent equilibrium configuration. It is easily seen that for $\epsilon > 0.5$ the lower potential V^- will possess potential minima at $r \neq 0$, corresponding to a bent configuration. (This is shown in the illustration for $\epsilon = 1$.) In extreme cases these minima in the lower potential may be so deep that a number of vibronic levels will lie below the central, linear local maximum. Under these circumstances it may be expedient to treat the lower component in terms of a nondegenerate electronic state of a bent molecule.

States of higher multiplicity. More complicated cases occur in Π and Δ states of higher multiplicity. In the presence of nonzero electron spin the effects of both vibronic and spin-orbit interactions must be considered. Particularly when both interactions are of comparable magnitude, their effects are not simply additive, but have to be treated simultaneously. Approximate formulas for treating these cases have also been developed. This situation is quite common, particularly in free radicals and molecular ions, and numerous examples of the Renner-Teller effect in $^2\Pi$ and $^3\Pi$ states have been studied. *See* MOLECULAR STRUCTURE AND SPECTRA.

Vladimir E. Bondybey; Terry A. Miller

Bibliography. G. Duxburg and R. N. Dixon, The Renner effect in a bent triatomic molecule using the adiabatic approach, *Mol. Phys.*, 43:225–274, 1981; G. Herzberg, *Electronic Spectra of Polyatomic Molecules*, 1966; G. Herzberg and E. Teller,

Schwingungstruktur der Elektronenübergänge bei mehr atomigen Molekülen, *Z. Phys. Chem.*, B21:410–466, 1933; Ch. Jungen and A. J. Merer, Orbital angular momentum in triatomic molecules, *Mol. Phys.*, 40:1–23, 1980; R. Renner, Zur Theorie der Wechselwirkung Zwischen Elektronen- und Kernbewegung bei drei atomigen, strabförmigen Molekülen, *Z. Phys.*, 92:172–193, 1934.

Rennin

The common name for chymosin, a proteolytic enzyme that is used to coagulate milk in cheesemaking. Rennin participates in the cheese-ripening process through its proteolytic activity.

Preparation. The traditional source of rennin is the fourth stomach (abomasum) of calves that have been fed only milk. The stomachs are dried or salted, cut into small pieces, and soaked in 10% salt brine to extract enzyme components from the stomach lining. The enzyme extract, commonly known as rennet, contains a mixture of around 80–95% rennin and 5–20% bovine pepsin, another proteolytic enzyme with milk-coagulating activity. Rennin also is produced commercially by using genetically engineered microorganisms. Recombinant deoxyribonucleic acid (DNA) techniques are used to clone and express the animal gene for rennin in the bacterium *Escherichia coli* and the yeast *Saccharomyces cerevisiae.*

Rennet may be used as a coagulant for cheesemaking or may be purified further by ion-exchange chromatography to yield pure rennin. Rennet or rennin is preserved with sodium benzoate and then standardized for milk-clotting activity.

Milk coagulation. Casein, the major milk protein, consists of three protein subunits, α-, β-, and k-caseins, arranged in a macromolecular structure known as a micelle. Casein micelles remain dispersed in milk because of the stabilizing action of k-casein. Milk coagulation occurs when casein micelles are destabilized by rennin. The process takes place in two distinct phases. Rennin catalyzes the first, or enzymatic, phase by cleaving one specific bond in k-casein. This results in the liberation of a glycomacropeptide and the loss of micelle stability. The second, or nonenzymatic, phase occurs after approximately 80% of total micellar k-casein has been cleaved by rennin. Casein micelles then agglomerate in the presence of ionic calcium, entrapping milk fat globules in the process. A three-dimensional spongelike matrix is formed that becomes the structural backbone of the resulting cheese curd. *See* CASEIN; MICELLE; MILK.

Cheese ripening. Approximately 6% of the rennin used to coagulate milk is retained in active form in cheese curd. During cheese ripening, rennin modifies the curd protein structure through its proteolytic action on α-casein, leading to textural changes described as a loss of curdiness. Casein peptides resulting from rennin action become precursors for flavor compounds in some cheeses such as Cheddar. *See* CHEESE; ENZYME.

Paul Kindstedt

Bibliography. P. F. Fox (ed.), *Cheese: Chemistry, Physics and Microbiology*, vol. 1: *General Aspects*, 1987; P. Walstra and R. Jenness, *Dairy Chemistry and Physics*, 1984; N. P. Wong (ed.), *Fundamentals of Dairy Chemistry*, 3d ed., 1988.

Renormalization

A program in quantum field theory consisting of a set of rules for calculating S-matrix amplitudes which are free of ultraviolet (or short-distance) divergences, order by order in perturbative calculations in an expansion with respect to coupling constants. *See* SCATTERING MATRIX.

Divergences in quantum field theory. To describe the nature of the problem, it is useful to consider the simple example of a ϕ^4 theory defined by the lagrangian density $L(x)$ in Eq. (1) in four-dimensional

$$L(x) = \frac{1}{2}\partial_\mu\phi(x)\partial^\mu\phi(x) - \frac{1}{2}m^2\phi(x)^2 - \frac{1}{4!}\lambda\phi(x)^4 \quad (1)$$

Minkowski space-time. Here, $\phi(x)$ is a quantum field operator depending on the four-vector x, ∂_μ and ∂^μ represent differentiation with respect to space-time coordinates, and m and λ are parameters. *See* LAGRANGIAN FUNCTION; RELATIVITY.

If one attempts to calculate any physical process in a perturbative expansion with respect to the coupling constant λ, in terms of Feynman diagrams, one has to confront divergent integrals. An example is the one-loop contribution to two-body scattering which is of order λ^2 (the second diagram in the **illustration**), and is proportional to the logarithmically divergent momentum integral given by expression (2),

$$\lambda^2 \int d^4k[(p_1+k)^2 - m^2 + i\epsilon]^{-1} \cdot [(p_2-k)^2 - m^2 + i\epsilon]^{-1} \quad (2)$$

where p_1 and p_2 are the initial momenta of the two particles, and k is the loop momentum. *See* FEYNMAN DIAGRAM.

These ultraviolet (large-momentum) divergences have their origin in short-distance singularities occurring in the product of the quantum field operators (or their matrix elements) such as $\phi(x_1)\phi(x_2)$ as x_1 approaches x_2 [see expression (5)]. Therefore,

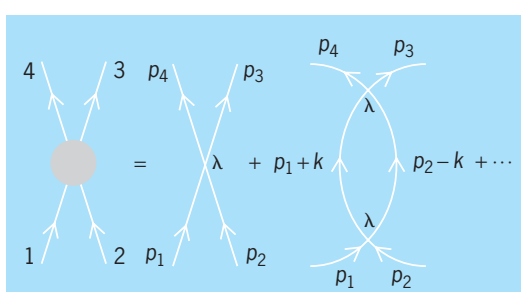

Two-body scattering in lowest-order perturbative expansion.

the quantum theory is not well defined since the lagrangian (or hamiltonian) contains these fields multiplied at the same point $x_1 = x_2 = x$. Renormalization is the procedure for constructing a well-defined finite quantum field theory, and deals with the proper definition of such singular terms in the lagrangian.

Regularization procedure. First the nature and kind of singularities must be identified, and then they should be removed to define the physical theory. The theory is first rendered finite by introducing a regularization parameter so that as it approaches a limiting value the divergences appear as definite singularities in this parameter. This process is called regularization. There are many methods of regularization. For example, in the cutoff method the integral in Eq. (2) will be rendered finite if it is cut off at $k^2 = \Lambda^2$, where Λ is finite. The result of integration, which is a function of Λ, will be seen to diverge as $\log \Lambda$ when $\Lambda \to \infty$. Another, more modern method is dimensional regularization: The integral is first performed in n dimensions rather than four. It will converge for $n < 4$ and will have a definite dependence on n. An example of this dependence is given by Eq. (3), where $\Gamma(\alpha)$ is the Euler

$$\int d^n k(-k^2 + c)^{-\alpha} = i\pi^{n/2} c^{(n/2)-\alpha} \frac{\Gamma[\alpha - (n/2)]}{\Gamma(\alpha)} \quad (3)$$

gamma function. All integrals, including the one in expression (2), $\Gamma(\alpha)$ is the Euler gamma function. All integrals, can be put in the form of Eq. (3) by using Feynman parameters to combine denominators. The answer is then analytically continued to complex values of n and the four-dimensional theory is defined as the limit of n approaching 4. In this way, as long as $n \neq 4$ the gamma functions in Eq. (3) are finite, and only finite well-defined quantities are manipulated. In this method the singularities reappear as poles of the form $(n-4)^{-k}$ as n approaches 4. The behavior of the integral of expression (2) is given by expression (4). The remainder of this article is restricted to

$$\frac{2i\pi^2}{n-4} + (\text{finite}) + 0(n-4) \quad (4)$$

dimensional regularization. *See* GAMMA FUNCTION.

Renormalization of parameters. In a regularized quantum field theory, one can, in principle, calculate finite expressions for all the Feynman diagrams that contribute to any physical process, to any desired order in perturbation theory. The answer will be a well-defined function of the momenta of the external legs, the parameters of the theory, such as m and λ in Eq. (1), and the number of dimensions n, which is taken as a complex number. As n approaches 4, the singularities can be studied and the method for removing them can be given. The degree of divergence of a graph is determined by Weinberg's theorem. In a field theory such as ϕ^4 of Eq. (1), one finds only two types of Feynman graphs that diverge as $n \to 4$. Those are the graphs with two external legs (propagator) and four external legs (scattering amplitude) and any other graph that contains these two types of graphs as subgraphs. Therefore, to further study the singularities it is sufficient to concentrate only on the two- and four-point functions of expression (5), whose Fourier transforms give the scattering amplitude. The singularities occur at short distances in the product of the field operators as points approach each other. This suggests that they can be removed from all physical processes by properly defining the local product of operators that appear in the lagrangian of Eq. (1). It turns out that indeed it is possible to absorb and remove all the infinities by a simple renormalization of the finite number of constants that define the theory. A theory is called nonrenormalizable when this procedure fails, and it can be rendered finite only by introducing an infinite number of renormalization parameters. *See* FOURIER SERIES AND TRANSFORMS.

$$\begin{gathered}\langle 0|T(\phi(x_1)\phi(x_2))|0\rangle \\ \langle 0|T(\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4))|0\rangle\end{gathered} \quad (5)$$

Nonphysical nature of parameters. The possibility of renormalization by a redefinition of parameters such as masses and coupling constants that appear in the lagrangian hinges on the fact that these are not the physical quantities that would be observed as the prediction of the theory. For example, the physical finite coupling constant is defined as the strength of the interaction that an experimentalist will observe in a two-body scattering process with initial four momenta p_1 and p_2, and final four-momenta p_3 and p_4 (see illus.). The strength of the interaction is defined by the measured value of the physical probability amplitude, $A(p_1,p_2,p_4,\lambda,m)$, at an agreed value of the external momenta. This clearly is a complicated function of the parameter λ when the perturbation series is summed. The strength of the interaction is equal to λ only in lowest order and in general differs from it in the full theory. Similar remarks apply to the parameter m. The true mass of the particle in the theory differs from m. It is in general a complicated function of m and λ and is defined as the location of the pole in the full propagator. Similarly, the field $\phi(x)$ does not create the correctly normalized interacting particle when applied on the vacuum. It differs from the correctly normalized field ϕ_R by the multiplicative wave function renormalization constant $\sqrt{Z}$. Renormalization theory is the process of rewriting systematically the field theory in terms of the physical, finite, renormalized coupling constant λ_R, mass m_R, and field ϕ_R. If the theory is renormalizable and rewritten in terms of renormalized quantities, then all infinities cancel and all physical quantities, such as S-matrix elements, are finite and meaningful, as the regularization parameter is removed.

Power series in the renormalized parameters. An outline of this procedure can be given as follows. Since it has been agreed that λ, m and the renormalization of the field are not the observed physical quantities, they need not be finite. They are reexpressed as power series in terms of finite renormalized parameters λ_R and m_R and the regularization parameter. In the dimensional regularization scheme they take the form of

Eqs. (6), where the renormalized field ϕ_R is defined

$$\lambda = \mu^{4-n}\left(\lambda_R + \frac{a_1(\lambda_R)}{n-4} + \frac{a_2(\lambda_R)}{(n-4)^2} + \cdots\right)$$
$$m^2 = m_R^2\left(1 + \frac{b_1(\lambda_R)}{n-4} + \frac{b_2(\lambda_R)}{(n-4)^2} + \cdots\right) \qquad (6)$$
$$Z = 1 + \frac{c_1(\lambda_R)}{n-4} + \frac{c_2(\lambda_R)}{(n-4)^2} + \cdots$$

as $\phi = \sqrt{Z}\phi_R$. Here $a_i(\lambda_R)$, $b_i(\lambda_R)$, $c_i(\lambda_R)$ are power series in the finite parameter λ_R; that is, $a_i(\lambda_R) = \Sigma a_{ik}\lambda_R^k$, and so forth, where the coefficients a_{ik}, b_{ik}, c_{ik} are to be determined so as to render the theory finite as described below. Here μ is a mass parameter called the renormalization point, whose role is better understood by studying the renormalization group (discussed below). All Green's functions or probability amplitudes such as the one in the illustration will be renormalized by setting the momenta at $p^2{}_i = \mu^2$. Furthermore, in Eq. (6), λ_R and m_R implicitly depend on μ, while λ and m are independent of it. *See* GREEN'S FUNCTION.

The perturbation series is now arranged as an expansion in the finite parameter λ_R and not in λ. This can be done by rewriting the lagrangian in terms of renormalized field ϕ_R and the parameters λ_R and m_R by substituting directly from Eqs. (6) into Eq. (1). Thus, the lagrangian takes the form of Eq. (7), where

$$L = \frac{1}{2}\partial_\mu\phi_R\partial^\mu\phi_R - \frac{1}{2}m_R^2\phi_R^2 - \frac{1}{4!}\mu^{4-n}\lambda_R\phi_R^4 + \Delta L_{\text{counterterm}} \qquad (7)$$

$\Delta L_{\text{counterterm}}$, the "counterterm" piece, has the same functional dependence on ϕ_R as the explicitly written terms in Eq. (7) but with coefficients that correspond to the pole terms in Eq. (6). This piece would diverge as $n \to 4$ and is a power series in λ_R involving the coefficients a_{ik}, b_{ik}, c_{ik} mentioned above.

Minimal subtraction. These coefficients are determined in order to cancel all infinities by the following procedure: The amplitude $A(p_i,\lambda_R,m_R,\mu,n)$ and the propagator $\Delta_F(p_i,\lambda_R,m_R,\mu,n)$ are calculated to any desired order in λ_R in terms of Feynman graphs [these are related to the Green's functions in expression (5)]. To these there are contributions from the explicit part of the lagrangian in Eq. (7) and also from the counterterms which involve the coefficients a_{ik}, b_{ik}, and c_{ik} up to the desired order in λ_R. Then, order by order, one requires that the poles $(n-4)^{-i}$ must completely cancel in the three measurable quantities given in Eqs. (8). This is called the minimal subtraction scheme, and it completely fixes the coefficients a_{ik}, b_{ik}, and c_{ik}. There are other subtraction schemes which require the above quantities to be equal to a fixed observed value rather than being simply finite as $n \to 4$. The observable quantities such as the S-matrix elements are independent of the particular subtraction scheme adopted.

$$A(p_i, \lambda_R, m_R, \mu, n)|_{p_2=\mu_2} = \text{finite as } n \to 4$$
$$\Delta_F^{-1}(p, \lambda_R, m_R, \mu, n)|_{p_2=\mu_2} = \text{finite as } n \to 4 \qquad (8)$$
$$\frac{\partial^2}{\partial p^2}\Delta_F^{-1}(p, \lambda_R, m_R, \mu, n)|_{p_2=\mu_2} = \text{finite as } n \to 4$$

Finite S-matrix. As mentioned above, the only "primitive" infinities in the ϕ^4 theory appear in the Green's functions of expression (5) containing two or four external legs and in those graphs containing them as subgraphs. With the above renormalization procedure (choosing a_{ik}, b_{ik}, c_{ik}) all such infinities in these two- and four-point functions have been arranged to cancel. It remains to be shown that as $n \to 4$ (removal of cutoff) the S-matrix is finite. This is done by mathematical induction: Assume that it is true to order m in perturbation theory, then prove it to order $m + 1$. This is carried out successfully if the theory is renormalizable. Therefore, all infinities can be removed by a proper definition of the products of the field operators appearing in the lagrangian. This amounts to properly identifying the observable finite quantities which led to a renormalization of the field, the coupling constant and the mass as in Eqs. (6). Any other physical quantity can now be calculated and will be a function of only the renormalized finite parameters λ_R and m_R, in addition to momenta.

Examples of renormalizable fields. So far the only field theories known to be renormalizable in four dimensions are those which include spin-0, spin-$^1/_2$ and spin-1 fields such that no term in the lagrangian exceeds operator dimension 4. The operator dimension of any term is calculated by assigning dimension 1 to bosons and derivatives ∂_μ, and dimension $^3/_2$ to fermions. Spin-1 fields are allowed only if they correspond to the massless gauge potentials of a locally gauge-invariant Yang-Mills-type theory associated with any compact Lie group. The gauge invariance can remain exact or can be allowed to break via spontaneous breakdown without spoiling the renormalizability of the theory. In the latter case the spin-1 field develops a mass. The successful quantum chromodynamics theory describing the strong forces and the SU(2) × U(1) Weinberg-Salam-Glashow gauge model of unified electroweak particle interactions are such renormalizable gauge models containing spin 0, $^1/_2$, and 1 fields. The renormalization procedure in a gauge theory is much more complicated than in the simple ϕ^4 theory because of gauge fixing and lack of either manifest unitarity or Lorentz invariance, but gauge theories have been shown to be renormalizable. *See* ELECTROWEAK INTERACTION; FUNDAMENTAL INTERACTIONS; GAUGE THEORY; LIE GROUP; QUANTUM CHROMODYNAMICS; QUANTUM ELECTRODYNAMICS; WEAK NUCLEAR INTERACTIONS.

Renormalization group. An important topic in renormalization theory is the renormalization group. This is the study of the dependence of the theory on the renormalization point μ that appeared in the subtraction procedure of Eqs. (6) and (8). Observable quantities such as the S-matrix elements, the measured strength of the interaction in the illustration, and so forth, do not depend on μ, but the finite

expansion parameter λ_R or the mass parameter m_R do. All computations depend on the lagrangian in Eq. (1) which has no information about μ. Changing the value of the renormalization point in Eq. (6) induces new values of λ_R and m_R in such a way as to keep the measurable S-matrix elements unchanged. Since changing the mass parameter μ can be viewed as a change of scales, the renormalization group is intimately connected to scale transformations in the theory. It makes it possible to study the high- or low-energy behavior of the field theory by applying scale transformations to the momenta of the particles involved in a particular reaction. It is then found that the value of the effective expansion parameter $\lambda_R(\mu)$ of the theory depends on the energy scales of the reaction under study. It becomes convenient to choose the value of μ such that $\lambda_R(\mu)$ is small enough for a valid perturbative expansion. For certain gauge theories the effective coupling constant, which is a measure of the interaction, decreases as the energy scale increases. This behavior is called asymptotic freedom, and it explains why quarks act as free particles (small effective coupling constant) when they collide at very high energies and come to within very short distances of each other. *See* QUARKS; SYMMETRY LAWS (PHYSICS). Itzhak Bars

Effective field theory. Effective field theory is a general and powerful method for analyzing quantum field theories over a wide range of length scales. Together with a closely related idea, the Wilson renormalization group, it places renormalization theory on a more general, physical, and rigorous basis. This method is most naturally developed in the Feynman path integral formulation of quantum field theory, where amplitudes are given by an integral over all histories. Each history is weighted by a phase equal to the classical action divided by Planck's constant. *See* ACTION.

An effective field theory is a quantum field theory which describes physics only at distances greater than some small length ℓ. The integral over histories is cut off, wavelengths less than ℓ being omitted. The physics on scales shorter than ℓ is parametrized by taking a completely general effective action. Factors of ℓ are inserted into the action so as to make the constants in the lagrangian dimensionless (in units where the speed of light and Planck's constant are equal to 1). The action has a series of terms that contain higher and higher powers of ℓ.

There are two evident differences between an effective field theory and the older idea of a renormalizable field theory: The integral over histories is cut off at a length ℓ, and the action contains an infinite number of terms. Because of the cutoff, ultraviolet divergences are no longer present, and the renormalization theorem takes a more physical form. That is, amplitudes at distances much greater than ℓ depend on the unknown physics at distances below ℓ only through a finite number of parameters. This is very plausible from dimensional analysis, because all but the first few terms in the lagrangian contain positive powers of the small length ℓ. These few terms are precisely the ones retained in renormalizable field theory. More generally, there may be measurable effects proportional to ℓ or ℓ^2, or even higher powers. These are again parametrized by a finite number of constants. In applications to statistical mechanics, a similar expression for the free energy (instead of the action) is used. *See* FEYNMAN INTEGRAL.

Wilson renormalization group. It is useful to organize the path integral by scales, integrating first wavelengths between the cutoff ℓ and a slightly greater scale ℓ^1. The result is an effective field theory with the new cutoff ℓ^1 and a new action, which now includes the effects of physics on scales between ℓ and ℓ^1. The action is thus a function of the cutoff scale, and the process of integrating out a narrow band of wavelengths gives rise to a flow, a differential equation for the couplings. This is the Wilson renormalization group equation, which has broad applications in particle physics, condensed-matter physics, and statistical mechanics.

A point where all the parameters in the equation vanish is a fixed point of the flow; if the flow at this point is inward, it is termed an infrared fixed point. The physics at very long distances will be governed by an infrared fixed point, so such points define the possible phases of the theory. This is the basis of K. Wilson's theory of critical phenomena. *See* CRITICAL PHENOMENA; QUANTUM FIELD THEORY. Joseph Polchinski

Bibliography. J. C. Collins, *Renormalization*, 1986; R. J. Creswick, H. A. Farach, and C. M. Poole, *Introduction to Renormalization Group Methods in Physics*, 1991; C. Itzykson and J. B. Zuber, *Quantum Field Theory*, 1980; W. Marciano and H. Pagels, *Phys. Rep.*, C36:137, 1978; K. Wilson, Problems in physics with many scales of length, *Sci. Amer.* 241:(2)158–179, August 1979; K. Wilson, The renormalization group and critical phenomena, *Rev. Mod. Phys.*, 55:583–600, 1983.

Reproductive behavior

Behavior related to the production of offspring; it includes such patterns as the establishment of mating systems, courtship, sexual behavior, parturition, and the care of young. Successful reproductive efforts require the establishment of a situation favorable for reproduction, often require behavior leading to the union of male and female gametes, and often require behavior that facilitates or ensures the survival and development of the young; the mere union of gametes is not generally sufficient for successful reproduction. For each species, there is a complex set of behavioral adaptations that coordinate the timing and patterning of reproductive activity. Typically, this entails integration of both overt behavioral and internal physiological events in both male and female, all of which are intricately enmeshed in manners adapted to the environment in which the animals live. The behavioral patterns related to reproduction tend to be relatively stereotyped within a species, but diverse among different species—especially

distantly related species. The end products of cycles of reproductive activity are viable, fertile offspring which, in turn, will reproduce and thus perpetuate the species.

Underlying the integrated and cooperative behavior leading to the perpetuation of the species are the interests of individual animals. Natural selection is most often seen as working at the level of the individual; those individuals that will be favored by natural selection are those that are able to survive and reproduce effectively. If there is genetic variation within a population, it is the genes that relate to successful reproduction by individuals that will be perpetuated and come to be best represented in future generations. The concept of fitness relates to the relative contribution to the genetic structure of future populations by individual organisms. Thus, what appears to drive the evolution of integrated and cooperative reproductive systems is the action of natural selection favoring those individuals behaving in ways that maximize their reproductive success. *See* ORGANIC EVOLUTION.

The emphasis in studies of reproductive behavior has been upon sexual reproduction—that form of reproduction that involves substantial genetic recombination. It should be remembered, however, that many species reproduce asexually, via autogamy, budding, parthenogenesis, or similar processes. The essence of asexual reproduction is that the offspring generally bear genotypes that are close, if not exact, copies of those of their parents. Reproductive behavior may nevertheless be quite complex. For example, in the parthenogenically reproducing lizards of the genus *Cnemidophorus*, females court, mount, and ride each other in a manner remarkably similar to that of closely related species that reproduce sexually. *See* REPRODUCTION (ANIMAL).

Each successful episode of reproductive activity takes place in a given context, begins with some preliminary events, includes the actual mating, has important sequelae, and leads to parental behavior. The events encompassing each aspect of the sequence must be described before they can be analyzed with respect to underlying mechanisms, development in the individual, evolutionary history, and adaptive significance.

Context

The context of reproductive behavior includes such factors as the season of breeding, the mating system, and social factors such as territoriality and dominance.

Seasonality. Seasonal breeding is the rule in many species—especially those living in temperate zones. The reason appears related to the seasonal availability of resources for the developing young. Species tend to bear their young at the time that will be most favorable, often in the late spring or early summer. Because gestation periods are so variable, the time of breeding tends to vary greatly among species. For example, many species of deer and elk have a fall rut, coyotes and wolves often breed in winter, and many seals and sea lions breed in late spring.

Along with reproductive behavior, a whole complex of behavioral and physiological changes occurs seasonally. Such changes are most dramatic in migratory species, such as the short-tailed shearwater, a bird that arrives on islands near Tasmania and lays most of its eggs within a period of a few days in November. Some annual cycles, such as those of ground squirrels, are driven by endogenous processes and thus show a cycle of about a year even when the animal is deprived of any external cues relating to season.

Cycles often are tied to external stimuli. Annual breeding cycles can be tied to such stimuli as rainfall, temperature, plant growth, or day length. Because it tends to be the most reliable cue, daylength is used by a large number of species. Syrian golden hamsters, for example, show a regression of their gonads in the fall when there are 12.5 h or less of light per day. However, they recrudesce about 4 or 5 months later, regardless of the photoperiod.

The roles of the pineal gland, melatonin, and the hypothalamic-pituitary-gonadal axis in the control of annual cycles in hamsters have been delineated. In canaries, parts of the brain related to singing are sexually dimorphic (larger in males than in females), but these too vary seasonally and are considerably smaller in the fall than in the spring.

Field studies of hormones in sparrows have shown that seasonal changes in male territorial behavior are tightly linked to changes in testosterone levels induced by social and environmental stimuli. Changes in male copulatory behavior are less testosterone-dependent, being triggered by female solicitation displays.

In sum, as the result of the action of natural selection, largely on the timing of births, whole complexes of behavioral, physiological, and morphological adaptations are integrated into seasonal patterns.

Mating systems. The relationships between individual males and females and the degree of exclusivity in mating are part of the mating system of a species or population. There are three basic mating system types: monogamy, polygamy, and promiscuity.

Monogamy. In this mating system the reproductive unit is generally a single male and a single female, the partners copulate only with each other, there may be shared parental care, and there is some kind of prolonged pair bond. Monogamy is very common among birds (it has been estimated that 90% of avian species are monogamous). There are many species, however, that fall in the gray areas between a monogamous system and some other form. In some species, there is a prolonged pair bond and shared parental care, but occasional matings outside the pair. In others, there may be mating exclusivity but no prolonged bond. Animal behaviorists disagree as to which of these characteristics should be defining attributes of a monogamous system.

Polygamy. In polygamous mating systems, there is again a prolonged association, but more than two individuals are involved in the relationship. When there is prolonged association and an essentially

Fig. 1. Sage grouse lek. Males stake out and defend territories on a traditional breeding ground. Females breed with males in selected territories. (*Courtesy of R. Haven Wiley*)

exclusive mating relationship between one male and two or more females, the system is termed polygynous. For example, bats of the species *Phyllostomus hastatus* roost in cave colonies, in clusters that consist of a single adult male and a harem of about 18 females. The harem male actively defends the cluster from other males, many of which live in bachelor groups, and he sires all or most of the offspring born to the harem females.

If there is a prolonged association but it is between a single female and two or more males, the system is termed polyandrous. In the American jaçana, a pond-dwelling bird, a female defends a large territory that encompasses the territories of several males with which she associates. Females are larger and more aggressive than males and lay clutches of eggs in nests in each of the males' territories. Males incubate the eggs.

Promiscuity. In this mating system, there is no prolonged bond formed, and there are multiple matings by members of at least one of the sexes. Some species of grouse, bears, and wildebeest appear to be promiscuous.

Territoriality and dominance. Territoriality or dominance occurs in many kinds of mating systems. A territory is an area that is defended against conspecific animals (those of the same species). It may be occupied by a single individual, a bonded male-female pair, or a larger group. If an intruder and a resident engage in an aggressive interaction in a territory, it is the resident that generally wins. Interestingly, if the same two animals are in conflict in the other individual's territory, the tables are turned, and the one that is now within its own territory generally wins even though it lost to the first animal when off its own territory. An example is provided by male speckled wood butterflies, which compete for spots of sunlight on the ground that serve as mating territories. Resident males almost always win encounters in their territories. *See* TERRITORIALITY.

The resident of a territory generally has privileged access to the resources on that territory. Where the territory is relatively large, as in many diurnal songbirds, it may include sufficient resources to support a bonded pair and their offspring. By contrast, in many colonially nesting marine birds the territories may encompass little more than a nest site, while food and other resources are collected at a distance. In a special form of territoriality, a lek, males of some species, such as sage grouse, defend small territories that are used only for breeding (**Fig. 1**). After male territories have been staked out, females approach the lek and breed with some of the males. The females then depart; males play no further role in reproduction but continue contests for territories and matings with additional females.

Whereas in a territorial system the outcome of a contest is generally predictable given only the location of the encounter, in a dominance relationship an individual wins regardless of location. In troops of various species of primates, for example, there may be a single dominant male and a hierarchy of males ranking below him with each dominant to males of lower rank. Such relationships can be quite stable over time. Often the dominant males have privileged access to resources, such as food and mating partners, but the evidence for such effects needs expansion. There are, of course, many varieties of dominance relationships, with male hierarchies, female hierarchies, mixed-sex hierarchies, and triangular relationships that are departures from linearity. Dominance-related contests may occur seasonally, generally peaking in intensity during the breeding season, and can be of great importance in determining which individuals reproduce.

The dividing line between dominance and territoriality is not always clear, and both can occur within a species. In some species of lizards, for example, several males may defend territories against each other, but with one dominant male that can move about the entire area. With a system based on dominions, there are zones of increasing and decreasing dominance. There is not overt defense of borders, but individuals nearest their core area generally secure access to resources.

Not all mating systems are based on dominance or territoriality. For example, scramble competition polygyny is common in the thirteen-lined ground squirrels (*Spermophilus tridecemlineatus*). Males compete in locating and mating with females but do not display dominance or territoriality; competitive mate searching is favored.

Preliminary Events

Much of reproductive behavior takes place prior to mating. Mate choice and then courtship are both essential to successful reproduction in animals.

Mate choice. Pairs often appear to result from an active choice of partners. In *The Descent of Man and Selection in Relation to Sex*, Charles Darwin differentiated sexual selection from natural selection. Sexual selection is related to the differential ability of individuals to secure mating partners, and is now generally agreed to be a subcategory of natural selection. Darwin noted that males generally compete among themselves for access to mating partners. He interpreted such structures as the elaborate plumage of some birds and mammalian horns and

antlers as functioning in such a context. Females, however, were viewed as rarely lacking mating partners. Rather, the problem for females was taken to be selection of a partner or partners from among several males. Thus, there developed a notion of the active and ardent male coupled with the coy and choosy female. Such differences are often related to the relative availability of potential mating partners for males and females and to the amount of reproductive investment characteristic of each sex. As will be discussed below, investments by females are generally regarded as larger than those by males, and hence females may be a limiting resource for the reproductive activity of males. There are some species, however, such as the American jaçana, where males appear to invest more and the relationship is reversed.

Female choice. The importance of female choice remains controversial. Ultimately, females can benefit by mating with males that are exceptional either in their ability to accrue resources or in the possession of "good genes," or both. The animals' basis for mate choice may be quite straightforward and correlated with these ultimate benefits but there is no implication that the animals understand these relationships. The female that mates preferentially with large males gains benefits without understanding the forces of natural selection that have shaped her preference.

A female could potentially increase her reproductive success by pairing with a male that provides more resources than do other males. In some avian species, for example, males arrive at the breeding ground first, establish territories, and display to females. The female that pairs with a male in a prime territory containing a protected nest site and an ample supply of food may fare better than one selecting a male on a marginal territory. In the black-tipped hangingfly, as in some other insects, courting males present a prey item to females. The female choosing a male with a large prey item gains a larger resource and may have greater reproductive success even though there is no prolonged male-female bond. In addition to the material benefits in these examples, however, the female's offspring will possess genes that could lead to fitness levels superior to those of alternatives. Benefits gained through genes and resources in such cases may be very difficult to separate.

In many species, however, males may contribute no obvious resources beyond their ejaculates. Mating is brief, there is no prey presentation, and pair bonds are nonexistent. One might suppose that selection would favor females mating with males carrying genes relating to traits that are adaptive in their habitat. For example, if large size might allow her offspring to sequester a differentially large number of mating partners or resources and if size has a genetic component, females would be better off mating with large males than with smaller ones. There are many studies demonstrating such choice in a variety of situations.

Kinship. Kinship may also play a role in mate choice. The offspring of matings between close kin are often of lower viability or fertility than those of animals more distantly related. There appear to be many adaptations that function to minimize incestuous breeding and the decrements in fitness that may follow. On the other hand, interspecific matings may be infertile and populations may become adapted to local conditions, making matings with individuals from other populations disadvantageous. Females thus should not mate with individuals too much different from themselves. There is some indication that females may prefer mates with an optimal level of discrepancy from themselves—similar but not too similar. In Japanese quail, for example, individuals reared with their siblings mate preferentially with first cousins. The conditions of rearing often play a major role in determining which individuals are treated as kin; although the ultimate benefit may be avoidance of inbreeding, the proximate cue may be familiarity.

The major histocompatibility complex (MHC) genes of house mice code for individuality of the immune systems in the body. Mice can discriminate individuals of different MHC genotypes, apparently on the basis of odor differences. They have been shown to display mate choice related to MHC genes, generally preferring mates different from themselves. This may provide the offspring with the capacity to respond to a wider range of antigens than would otherwise be the case. *See* HISTOCOMPATIBILITY.

Male choice. Males forming prolonged pair bonds invest much, and it is not surprising that they may exercise mate choice as well. Even males of promiscuous species, however, may show some evidence of mate choice. The male's ejaculate represents a nontrivial investment; it may take days for a male to replenish his supply of sperm or the secretions of his reproductive accessory glands. The spermatophores of some cricket species may constitute up to 30% of body weight. With a limited number of ejaculates to deliver, some exercise of male choice is to be expected and has been observed. In a variety of species of invertebrates, for example, males prefer large females to small ones. In both *Drosophila* and prairie voles, males prefer unmated females to recently mated females. Male red-spotted newts, with limited sperm-producing ability, prefer large females as\break mates.

Courtship. Courtship entails a sequence of behavioral patterns that eventually may lead to the completed mating. Patterns of courtship are quite diverse among different species but generally entail reciprocal signaling between male and female.

A favorite behavioral pattern of classical ethologists, the courtship behavior of the three-spined stickleback, provides a case in point (**Fig. 2**). At the onset of the breeding season, males move away from schools, adopt coloration typical of the breeding season (including a red belly), and become territorial. They build a nest within the territory. When the female enters the territory, the male initiates a zigzag dance, in which he darts back and forth toward and away from the female and the nest. Females signal their acceptance by adopting a head-up posture. The male leads the female to the nest

and she follows. He hovers above the nest, pointing at its entrance; the female enters the nest. The male then nudges the base of the female's tail with his snout which stimulates the female to spawn. The male fertilizes the eggs, the female leaves, and the sequence begins again with a new female until several females have spawned in the nest. The male cares for the eggs and the young. The courtship sequence can be seen to involve coordination and reciprocal stimulation between male and female as the encounter progresses stepwise to its completion with mating.

Mate choice is an important function of courtship. Many bouts of courtship break off without going to completed matings, often as a result of choice on the part of one or both partners. An important basis for choice is the species of the potential mate. Because of the disadvantages of interspecific matings, a whole array of adaptations, called reproductive isolating mechanisms, function to prevent such matings. Courtship isolation is one of the more important of these. A courtship sequence may be initiated but broken off when the partner fails to signal, or to respond to a signal in the manner characteristic of the species.

There is often much assessment of potential mates beyond that of species identity during courtship. It is during courtship that a male hangingfly presents his prey that may be accepted or declined by the female

Fig. 3. Pair of copulating scorpionflies (*Hylobittacus apicalis*, also called black-tipped hangingflies). The female on the right is feeding on a nuptial gift (blowfly). (*Courtesy of Randy Thornhill*)

(**Fig. 3**). In common terns it is the male that provides most of the food for the hatchlings. During courtship the male presents fish to the female. It has been found that the amount of food a male feeds a female during courtship is correlated with the amount of food he will later provide for the young, and with the total weight of the clutch. The female appears to select a male on the basis of the food presented and, in so doing, is apparently able to assess his effectiveness as a resource provider. There is no implication that it is done consciously.

Another function of courtship relates to synchronization. The gametes must be shed at a time when sperm are viable, eggs are ripe, male and female are in the appropriate state of readiness, and the environment is supportive of reproductive effort. The progressive interactive sequence of the courtship episode allows for coordinated events to occur at times appropriate for successful reproduction.

Arthropods. Courtship patterns are diverse among arthropods. Initial detection can occur in a number of ways. Silk moths, for example, are extremely sensitive to a sex attractant, called bombykol, produced by the female. A single molecule of bombykol can initiate a nerve impulse in a receptor cell; males can orient to females even in response to very low concentrations of the substance. More familiar to humans are the auditory signals of species such as crickets and katydids, which function to attract females and, in addition, sometimes to repel other males. They may thus aid in defense of territories. Signal patterns are species-characteristic. Females prefer the calls of males of their own species, while hybrid females may even prefer the calls of hybrid males. *See* CHEMICAL ECOLOGY.

Many species of arthropods use visual signals, the most spectacular of which are perhaps the courtship signals of fireflies. In many species the males fly above the ground and emit flashes in a species-typical pattern. The females are generally in the grass below and respond to a male's signal with a characteristic flash delivered with a specific latency after a male's signal. The male approaches the female, and pair members continue to signal as they approach and eventually mate. A skilled field investigator can

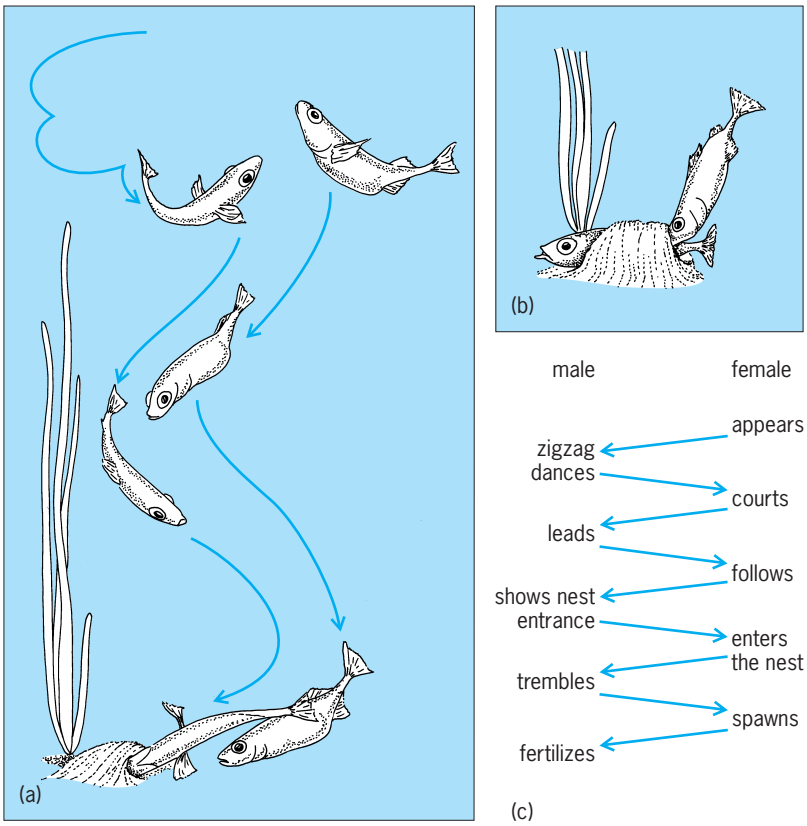

Fig. 2. Mating behavior of sticklebacks. (*a*) Zigzag dance of a male leading a female to the nest. (*b*) The male stimulating the female in the nest to lay eggs. (*c*) Schematic portrayal of the sequential events in a courtship sequence. (*After N. Tinbergen, The Study of Instinct, Oxford University Press, 1951*)

attract a male by mimicking the female's response pattern. Females of predatory firefly species have adopted the same trick: by mimicking the pattern of the females of the male's species, the predatory female attracts the male and devours him. This is but one instance of a predator cuing in on a courtship signal as a means of obtaining prey. Signaling male crickets, to take another example, also attract predators as well as females. *See* ANIMAL COMMUNICATION.

Male-male competition occurs during firefly courtship as well. In species in which the male signal involves a series of flashes, another male of his species may "inject" a flash into a signal. Such injected flashes may disrupt the courtship interaction. Patterns of such flashes provide further examples of the ways in which behavior is adapted to serve the interests of individuals, and cooperation in the perpetuation of the species occurs only so long as it is in the interests of the individuals to act in such a manner.

Fishes. The courtship pattern of sticklebacks, described above, is but one of the many diverse courtship patterns of fishes. Those of the tropical fishes kept by aquarium enthusiasts are perhaps the most familiar. Male guppies are among the most persistent courters. They often adopt a sigmoid (S-shaped) posture with the tail deflected, and a quivering motion. The display is usually given in the presence of a female, and during the display the male may move toward her. The male's gonopodium is a modified anal fin and may be swung laterally and forward, and thrust toward the female during courtship.

The beautiful patterns of the anabantoids, such as Siamese fighting fishes and gouramis, can be observed in a home aquarium. After the male builds a bubble nest on the surface of the water, there are sequences of patterns such as tail beating, leading, lateral displays, sigmoid displays, and circling prior to spawning.

Amphibians and reptiles. The calls of bullfrogs and other anurans are familiar to many. Males typically arrive at the breeding area first and establish a territory. Their species-typical mating calls generally attract females, stimulate female reproduction, and function in reproductive isolation. They may also attract predators. Some males, called satellites, may escape the risk of predator attraction by not calling at all. Rather, they may adopt a post near the territory of a calling male and attempt to mate with females approaching the calls. Such males often meet with some reproductive success. Such alternative mating strategies within a single stable breeding population are not uncommon.

Courtship in many species of anoles and other lizards involves characteristic patterns of head bobbing, together with other bodily movements. In the American "chameleon," *Anolis carolinensis*, there is a bright red fold of skin, the dewlap, that is exposed during such displays (**Fig. 4**).

A female red-sided garter snake emerging from hibernation typically encounters a mating ball of as many as 100 males competing to inseminate her. The male displays patterns of tongue flicking and chin rubbing prior to coupling.

Fig. 4. The extended dewlap of a male anole lizard (*Anolis carolinensis*) is part of the striking display pattern in this and related species. (*Courtesy of Thomas A. Jenssen*)

Birds. The courtship displays of ducks and geese have been the subject of intensive analysis. Patterns are similar in closely related species and, indeed, courtship patterns can be used to clarify taxonomic relationships that may be ambiguous when only morphological characters are used in classification. Such patterns as bill shaking, the tail shake, the grunt-whistle, and the head-up tail-down display can be seen in many duckponds.

Males of various species of Australian bowerbirds build structures (bowers) that attract females. In the golden bowerbird, males build a maypole bower that often is 8 ft (2.5 m) high. The bowers are decorated with mosses, flowers, and fruits, and the males display and vocalize in them. Satin bowerbirds paint and decorate their bowers elaborately, especially with items that are blue.

Bird songs may function in courtship, territorial advertisement, and other contexts.

Mammals. Courtship in mammals is generally less spectacular than in other taxa. Olfactory cues, which are important in many taxa, predominate in mammals. Thus, there is often much mutual grooming, sniffing, and licking, especially of the partner's genital area. Males of some species display a Flehmen response, in which the upper lip is curled and the neck is extended. The posture probably facilitates exposure of the odor to the vomeronasal organ via a duct in the front of the mouth. Animals of many mammalian species can discriminate an animal's species, sex, reproductive condition, and sometimes other characteristics as well, solely on the basis of olfactory stimuli. A variety of marking patterns make the odors particularly accessible.

Mating

Mating represents the consummation of sexual activity wherein the sperm are transferred from male to female so that the fertilization of ova can occur. Fertilization may occur internally (within the female's body) or externally (outside the body of the female). *See* FERTILIZATION.

Patterns of mating behavior. Different taxa display not only different modes of fertilization but also different patterns of mating behavior.

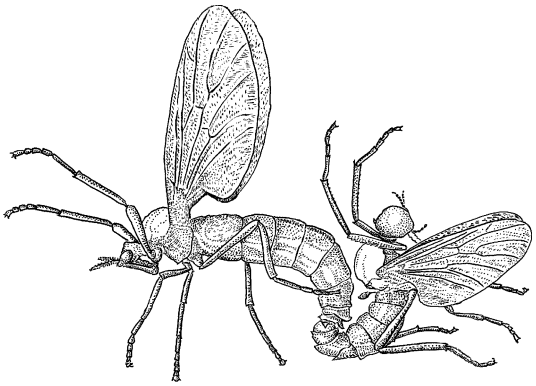

Fig. 5. Pair of lovebugs (*Plecia nearctica*) copulating. The female is on the left. These animals often remain in copulation for 2 days or more. (*After R. Thornhill and J. Alcock, The Evolution of Insect Mating Systems, Harvard University Press, 1983*)

Arthropods. In many species of arthropods the sperm are transferred in a spermatophore, a saclike structure in which they are enclosed. The spermatophore may also serve to plug the vaginal orifice and may have nutritional value. In some species of mites, pseudoscorpions, and millipedes, the male simply deposits the spermatophore on the substrate and the female comes along and picks it up; the pair may never meet. External fertilization occurs in a number of aquatic species, such as horseshoe crabs.

Copulation may be quite prolonged in species with internal fertilization; the mean duration of copulation in Florida "lovebugs" (*Plecia nearctica*) in the laboratory was found to be 56 h (**Fig. 5**). In many species the male clasps and locks the female during copulation. These clasps are maintained by a remarkable array of adaptations of mandibles, genital claspers, antennae, and modified legs. In the beetle *Lytta nuttalli* the penis contains dorsal and ventral penile spines that catch onto folds in the vaginal wall to secure a lock. Copulation in the southern green stink bug may last anywhere from 5 min to 14 days.

Male giant water bugs brood the fertilized eggs on their backs. Females deposit eggs a few at a time, and males insist on repeated copulations before each oviposition. In one case a pair coupled over 100 times in 36 h during the transfer of 144 eggs. In contrast, copulation in damselflies lasts just a little more than a minute. Sperm are transferred only during the last part of the copulation. Prior to that, males engage in undulating movements that, because of the special structure of the damselfly penis, serve to displace any sperm from males that mated previously with the female.

Fishes. In various species of the anabantoid bubble-nest builders mentioned above, as in many other species, fertilization is external. Sperm and eggs are released simultaneously, and fertilized eggs float up to the nest. Fertilization is also external in the Agnatha (jawless fishes). In the sea lamprey the female attaches herself to a rock with her sucker mouth; the male attaches himself similarly to the female and wraps his body around hers. Both vibrate their hindquarters and release gametes simultaneously.

Fertilization among sharks and other cartilaginous fishes is internal with sperm transfer effected via the male's clasper, a specialized intromittent organ derived from the basal portion of the pelvic fins. Sperm are propelled down a groove in the clasper. *See* COPULATORY ORGAN.

Fertilization is internal in many species of bony fishes as well. Copulations in guppies last just a few seconds. In the platyfish, copulations last 1 s or more; during longer copulations, pairs appear locked together as they swim but then they pull apart violently.

Oral fertilization has been found in the cichlid fish, *Pseudocrenilabrus multicolor*. The females brood eggs in their mouths, and the males have conspicuous anal-fin markings that resemble eggs. When females snap at the egglike markings as if to retrieve them, they suck up sperm that can fertilize the eggs within their mouths. External fertilization may also occur.

Amphibians and reptiles. In most anurans the basic mating pattern is the mating clasp, or amplexus, in which the male secures a firm hold on the female with his forelegs; fertilization is external (**Fig. 6**). Tailed frogs (*Ascaphus truei*) have an intromittent organ; during amplexus it is inserted into the female's cloaca, an opening near the base of the tail. In African toads there is also internal fertilization, effected by direct contact between the cloacae of the male and female.

In salamanders, fertilization is generally accomplished with a spermatophore. Typically the interaction begins with courtship, and the male deposits a spermatophore on the substrate. The female walks around and picks up the spermatophore between the lips of her cloaca. In some species, such as the red-spotted newt, the male captures the female in amplexus prior to depositing the spermatophore—especially if the female is unresponsive.

Many species of reptiles have internal fertilization, which is effected via one of the male's paired hemipenes. In anoles, for example, the male gains a neck grip on the female, twists his tail around her body, and inserts a hemipenis.

Birds. Although the courtship patterns of birds are remarkably diverse, there is little variability at

Fig. 6. Amplexus. In frogs and toads the male secures a firm clasp on the female. (*Photograph by D. Kim Sawrey*)

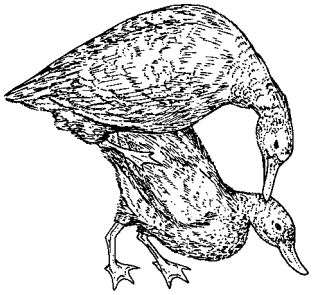

Fig. 7. Copulatory stance with the "cloacal kiss" in whistling ducks. (*After E. G. Bolen, Copulatory behavior in Dendrocygna, Southwest. Natural., 18:341–357, 1973*)

mating. Males of most species lack a penis; fertilization is effected via a "cloacal kiss," wherein the male and female cloacal glands are brought together and sperm are passed from male to female. Typically, the male stands on the female's back, sometimes with a neck grip, and the cloacae are brought together (**Fig. 7**).

Mammals. Mammalian copulatory patterns are quite diverse, although fertilization is always internal. Patterns can be usefully examined in relation to four characteristics: locks, intravaginal thrusting, multiple intromissions, and multiple ejaculations. In locking species, there is a mechanical tie, resistant to separation, between the genitalia of the male and female. Locking is found in some marsupials, short-tailed shrews, some rodents (such as golden mice, grasshopper mice, and some woodrats), and many carnivores (such as dogs and wolves).

In some species, pelvic thrusting ceases when the male gains insertion; this is true of many rodents, lagomorphs, and ungulates. In other species, however, as in virtually all primate species, there is repetitive intravaginal thrusting.

The third characteristic is the necessity of multiple intromissions as a prerequisite to ejaculation. In many species of rodents and primates the male repeatedly mounts the female and gains vaginal insertion only to dismount without ejaculating. Such multiple intromissions are prerequisite to ejaculation. In other species, ejaculation can occur the first time the male mounts and gains vaginal insertion.

The occurrence of ejaculation appears to terminate the copulatory interaction of a particular male and female in some species, such as certain shrews, squirrels, primates, pinnipeds, and ungulates. In other species, the pair may mate for several ejaculations.

Because each of these characteristics can be either present or absent, there are 16 theoretically possible patterns of copulatory behavior, if that behavior is described solely by the four characteristics mentioned here. The copulatory behavior of laboratory rats has been particularly well studied. The male mounts the female from behind. The female then adopts a stereotyped posture, lordosis, with a concave flexion of the spine, elevation of the head and tail region, and deflection of the tail (**Fig. 8**). During intromissions, the male mounts, gains penetration, and dismounts after an insertion of just 250–300 ms. He may engage in other behavior for a minute or so before remounting. About 10 such intromissions may be spaced a minute or so apart before the male displays a longer insertion that signals ejaculation. After a refractory period of approximately 5 min, the sequence begins anew—this time generally with fewer intromissions that are spaced more closely together. After about seven such ejaculations, refractory periods become so long that males can be regarded as satiated. This pattern can be seen as one with no locking, no intravaginal thrusting, multiple intromissions prerequisite to ejaculation, and multiple ejaculations. It is also characteristic of several species of gerbils, deer mice, and harvest mice.

The copulatory pattern of dogs contrasts with that of laboratory rats. Although males may mount (without insertion) several times prior to completed matings, they are capable of ejaculation the first time they mount. On gaining intromission, the male makes distinctive postural adjustments and begins a pattern of stepping with the rear legs. The bulbus glandis of the penis becomes engorged with blood and locks the male and female together. After an intense ejaculatory reaction the male typically dismounts while still locked to the female. In a study of beagles, the mean duration of locks was 14 min. Dogs display locking, intravaginal thrusting, ejaculation on a single insertion, and multiple ejaculations.

Japanese macaques display a pattern with no lock, intravaginal thrusting, multiple intromissions, and multiple ejaculations. By contrast, in black-tailed deer there is no lock, no thrusting, ejaculation on a single insertion, and just one ejaculation.

In many species the female is far from passive. In rodents, for example, the female often solicits copulations by approaching the male and then darting away. Female rats will learn an arbitrary task when the reward is the presentation of a sexually active male that will copulate with her. In primates the female uses a "present" posture to invite copulations; the same posture is used in other contexts, such as in indicating social subordination.

Mechanisms underlying mating behavior. Sensory control of mammalian copulatory behavior is generally complex, with no single sensation critical to the normal occurrence of behavior. Visual cues, such as the primate present posture, are important in some species. Olfactory cues play an especially important

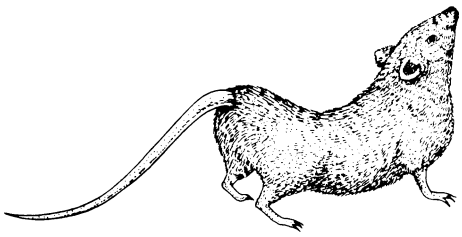

Fig. 8. Lordosis in a female rodent. When mounted, the females of many species adopt the lordotic posture rendering copulation possible. Note the concave arching of the spine. (*After L. C. Drickamer and S. H. Vessey, Animal Behavior: Concepts, Processes, and Methods, Willard Grant, 1982*)

role in a wide range of species. As might be expected, tactile sensitivity is critical to normal behavior; interference with an animal's sense of touch, especially that from the genital region, disrupts behavior. Auditory cues are important in many species. In various species of rodents, for example, males emit ultrasonic calls, inaudible to human observers, which are important in the interaction.

Hormones are critical to the display of copulatory behavior in most species. In general, hormonal effects in adults are reversible. If an animal's gonads are surgically removed, mating behavior will often disappear. Such behavior reappears, however, after treatment with appropriate hormone regimens—generally androgen injections for males and injections of estrogen followed by progesterone for females. In most species, female sexual behavior appears more closely tied to hormones than does male behavior. Males may continue to copulate for weeks after castration, whereas females cease to mate almost immediately after a surgery. This may be termed an associated reproductive pattern, as mating coincides with both gamete production and maximum levels of steroid hormones.

Species with dissociated patterns generally live in harsh environments, where there is little time for mating. Mating may occur in the absence of gamete production at a time when steroid hormone levels are low. In Canadian red-sided garter snakes (*Thamnophis sirtalis parietalis*), for example, mating occurs within about 3 weeks after emergence from hibernation. Later, testosterone levels increase and gametes are formed, to be stored until the next breeding season. Gonadal sex hormones do not play a major role in the control of mating in these species.

Many parts of the nervous system are important to the mediation of copulatory behavior. The reflexes that compose normal sexual behavior are partially organized in the spinal cord and hindbrain. A male dog with a transaction in the thoracic region of the spinal cord displays a complete set of spinal reflexes in response to tactile stimulation even though the neural connections with the brain have been completely severed. Within the brain, structures in the region of the hypothalamus are particularly important for reproductive behavior. This region lies at the base of the brain and sits atop the pituitary gland—the master control gland of the endocrine system. Various structures within the region are important in regulating the pituitary gland and, via both direct and indirect routes, controlling copulatory behavior as well. A variety of structures in the midbrain and forebrain have been found to be part of complex networks involved in the regulation of copulatory behavior. *See* NERVOUS SYSTEM (VERTEBRATE); PITUITARY GLAND.

Development of copulatory behavior. Copulatory behavior is not static, but changes, as the result of the interactions of genes and the environment, as an animal matures and ages. In vertebrates, sexual activity is delayed until a time of puberty. This is generally hormone-dependent; precocious copulatory behavior can be induced with exogenous hormone treatment. The behavior of the animal maturing naturally can be seen to unfold gradually. In the first copulatory episode of a male rhesus monkey, for example, the orientation of the mount is poor. Similarly, the first time a male rat copulates he may take longer to initiate mating and require more intromissions before ejaculating than he will later. Both age and experience can be important in effecting these changes. After a period of relative stability in adulthood, copulatory behavior often declines in old age. Often this takes the form of an increasing threshold for copulation; behavior may be quite normal once it occurs.

The early social environment is important to the development of normal copulatory behavior. In rodents, for example, deprivation of contact with the mother and siblings during development can lead to a complete lack of sexual activity or to aberrant behavior, such as mounting that is inappropriately directed. Similarly, monkeys and chimpanzees need to interact with others if adult behavior is to be normal; experience may be even more important for rhesus monkeys than for chimpanzees.

The most direct evidence for a genetic basis for copulatory behavior comes from studies related to behavioral development. Although raising an animal with an alien species may alter its mate preferences and raising it in isolation may eliminate the behavior altogether, the motor patterns of copulatory activity generally are species-typical when they occur. There is good direct evidence of genetic influences on copulatory behavior from studies of various strains and their crosses and from genetic selection experiments. Such studies have been done in house mice, laboratory rats, Japanese quail, and chickens. The importance of genetic influences has been repeatedly confirmed.

The actions of hormones during development are different from those in the adult, especially in that their actions appear less reversible. In mammals, sexual development in the absence of gonadal hormones is female; male characteristics fail to develop. If a newborn rat, for example, is deprived of its gonads, as an adult it will display the lordotic female mating posture in response to estrogen, progesterone, and mounting, but it will not show male-like behavior even if given androgen injections. By contrast, if the animal is exposed to gonadal steroids during the first few days of life, as an adult it will display normal male behavior in response to androgens, but the expression of female behavior will be blocked. Such relationships vary with taxon. In birds, for example, where males rather than females are the homogametic sex, development in the absence of hormones is male and early hormones feminize and demasculinize rather than masculinize and defeminize as in mammals.

The development of copulatory patterns is a continuous, dynamic process reflecting both social and organismic interactions. Thus, newborn rats exposed to androgens produce hormone-dependent odors, which stimulate a specific pattern of licking by the mother. That behavior, in turn, contributes to the development of masculine copulatory behavior.

The mother's licking is stimulated, in part, by her salt and water appetite. Mothers recover significant amounts of water by licking, and in turn stimulating, the young. The mother–young relationship in development is thus a reciprocal one.

Satiation. Although multiple ejaculations occur in many species, the rate of occurrence of copulatory behavior typically declines as the interaction continues. Some species, such as many species in the cat family, may continue to copulate regularly for several days. In many other species, however, a point of satiation is reached after a given number of ejaculations in a day.

Recovery from satiety also occurs at variable rates. It takes about a week or more for measures of the copulatory behavior of satiated male rats to return to the baseline levels characteristic of rested males. Recovery in male hamsters is much faster.

In many species, such as laboratory rats and various species of farm mammals, the copulatory activity of satiated males can be reactivated by replacing the original mating partner with a novel female—a phenomenon known as the Coolidge effect. Under some conditions a male that appeared satiated with one female may deliver several additional ejaculations when presented with a novel mating partner.

Sequelae to Mating

The most obvious consequence of mating activity is the fertilization of ova. With external fertilization the location of fertilization is obvious, but in species with internal fertilization, sperm and ova often come together in a specialized location, such as an oviduct or an analogous structure, and fertilization is effected. The sequelae to copulatory behavior are often somewhat more complicated than this, however.

Immediate consequences. The stimuli derived from mating can have important immediate physiological consequences. In some species, such as rabbits and cats, the ova are not released until triggered by copulation or similar vaginocervical stimulation. This is a kind of built-in "hold" and functions to broaden the range in time that a female can mate effectively. The postcoital afterreactions of female cats are correlated with intense physiological responses to copulation.

In other species, such as laboratory rats, mice, and hamsters, ovulation is spontaneous but there is an induced pseudopregnancy. A neuroendocrine reflex, operating via the nervous system, pituitary gland, and ovary, stimulates functional levels of progesterone in preparation of the uterus for the implantation of the fertilized ovum.

The proliferation of the uterine wall occurs only in response to vaginocervical stimulation. In different species of rodents the multiple intromissions preceding ejaculation and multiple ejaculations appear critical in triggering this reflex.

Sperm transport. Sperm transport is an active process in many mammalian species. One function of the repetitive multiple intromissions preceding ejaculation in rats is to facilitate sperm transport. Sperm that are delivered to the female without a full complement of preejaculatory intromissions fail to reach the uterus.

Sperm transport can be disrupted if copulation occurs too soon after a female receives an ejaculate. Female rats require about 5 min for appreciable sperm transport. If a female receives vaginal stimulation during that 5-min period, transport can be disrupted and, if she receives no more ejaculates, pregnancy may not occur. Male deer mice appear able to "cancel out" another male's ejaculation by copulating with a female within a minute or so after delivery of the ejaculate. This is presumably the result of disruption of sperm transport and possibly actual removal of sperm. The process appears analogous to that in damselflies discussed above.

Sperm competition. In many species the females copulate with more than one male during a single receptive period. Sperm competition is the competition between the ejaculates of two or more males within the reproductive tract of a single female. Because natural selection acts at the level of the individual and via differential reproduction, strong selection on effects related to sperm competition would be expected. Sperm competition was first studied extensively in insects, because of their great sperm storage capacity, but it is clear that in other species, such as sheep, deer mice, ground squirrels, and Norway rats, multimale mating may be quite common, and thus sperm competition occurs.

The pattern of sperm competition varies with the species and prevailing conditions. In the giant water bugs discussed above, for example, there is a strong last-male advantage. The last male to deposit sperm sires a disproportionately large number of young. This is a common pattern in insects. However, there are some species of insects and mammals (for example, house mice) in which the first male to ejaculate gains a disproportionate reproductive advantage. In others, such as laboratory rats and deer mice, there are no order effects. Males can gain a reproductive advantage in these species by delivering a higher percentage of the ejaculates received by a female than other males. The systems that regulate sperm competition are dynamic, and interactions are variable within species as well. The delay between males, relative numbers of ejaculations, and timing of the matings all affect which male prevails in sperm-competition interaction.

Parturition and Parental Behavior

There are various means used by animals to nurture developing young and to discharge them into the environment. In oviparous (egg-laying) species, fertilized eggs develop outside the body of the adult—even though fertilization may be internal. In viviparous (live-bearing) species, by contrast, the embryo develops within the reproductive tract of the female and derives nourishment from her. Ovoviviparous species bear live young, but they hatch from eggs that develop within the parent's body; this pattern is found in some snakes and fishes.

Patterns of parental behavior. Prior to birth or hatching, the parent may engage in behavioral patterns

that will aid the young when they arrive. This may entail preparation of a burrow or nest, provision of stored food, or acquisition of other resources. In some species the parent's aid ends with such preparations, but in others parental care may be extensive and prolonged.

Arthropods. In most arthropods, there is no parental care beyond preparation of an appropriate site and deposition of eggs in a favorable location. The most dramatic exceptions to this general rule can be found in the eusocial species—bees, wasps, ants, and termites, in which there is typically a division of labor, cooperation in the rearing of young, and an overlap of at least two generations at a time. Often there are sterile castes and care for the young is elaborate—typically in an extensive hive, nest, or burrow. Although sterile workers do not reproduce, they aid their close kin, which is to their genetic advantage even though the offspring produced are not their own. This view is compatible with the view that natural selection works at the level of the individual. However, this is only part of the story, and factors other than the close genetic relationship among females in some eusocial species appear essential in the evolution of this pattern. *See* SOCIOBIOLOGY.

Less elaborate parental care is found in a variety of species, such as many true bugs, thrips, webspinners, and hydrophilid beetles. Typically, mothers guard the young, shield them from predators and parasites, and may even aggressively attack predators. In some species the young may be guarded through several developmental stages.

In giant water bugs it is the males that care for the eggs, which are cemented to their backs. Males of several species of tropical assassin bugs straddle eggs and guard them from parasites and predators.

Examination and comparison of the insect species that do and do not show parental care and of the environments in which they live leads to the conclusion that parental behavior represents an adaptation to harsh living conditions of one type or another.

Fishes. The fishes may be the most variable taxon with respect to parental behavior and are ideal for tests of hypotheses regarding the evolutionary pressures creating variability in parental behavior. Eggs may be scattered on plants, over shoals, gravel, sand, or boulders, or in the open ocean, but in many species there is elaborate male care, female care, or biparental care. In seahorses and pipefishes the female deposits eggs in a special brood pouch on the male. The male stickleback aerates the nest and tends young. Active herding occurs in some species and others are mouthbreeders; they guard the young within the oral cavity during times of danger. The aquarist's *Tilapia* provide a familiar example of a mouthbreeding species.

Amphibians and reptiles. Although parental behavior in amphibians and reptiles is generally less elaborate than in other taxa, there are also many parental species. Among different species of salamanders the males or females may protect the young—often from attacks by conspecifics. Male midwife toads wrap eggs around their hindlegs after fertilization and carry them until they hatch.

Although snakes do not generate their own heat, they do absorb heat; in some species, adults coil themselves about eggs to keep them warm. Parental behavior is elaborate in many crocodilians, which guard the nest during incubation of the eggs and herd and defend a pod of young through early posthatching development.

Birds. Parental behavior is highly developed among birds. The diversity and elaborateness of bird nests are widely known. Although there are colonial species in which large numbers of birds build the nest, most commonly a single mated pair builds the nest for the eggs and defends it against intruders. In some species, "helpers," typically the young from previous years, may aid the parents in raising the young. The division of labor between male and female in nest building and incubation varies widely.

Temperatures above ambient levels are important for the development of avian eggs and are provided by body heat as one or both parents sit on the eggs. Typically, the feathers are erected so that the skin of the parent comes into direct contact with the eggs. In some species, there are specialized brood patches, areas in which blood vessels proliferate, skin is thickened, and feathers are lost during the incubation period.

The interactions of male and female in sharing parental care can be subtle. Male and female ring doves (*Streptopelia risoria*), for example, take turns sitting on the nest, with the male sitting for a period in the middle of the day and the female sitting the rest of the time. The timing of parental care is regulated by biological clocks, internal timing mechanisms that are entrained by the occurrence of light and dark during the 24-h cycle. Although the timing mechanisms differ in the male and female, the end result is an integrated pattern of shared brooding of the young. *See* BIOLOGICAL CLOCKS.

The American cowbird and various species of cuckoos are brood parasites. Rather than incubate their own eggs, they lay eggs in the nests of other species, which incubate the eggs. A variety of adaptations ensure that the parasitized host is deceived into caring for the parasites. The eggs of the parasite resemble those of the host. Typically, the eggs of parasites hatch before those of the host, thus giving the parasite an advantage in gaining parental attention. In some species the young parasite may roll the host's eggs from the nest; in others the parasite grows faster and larger than the host's young and again is able to monopolize parental resources.

The young of precocial bird species require little care; in such species as chickens and ducks the mother can be seen leading the young about, directing them to food and protecting them. In altricial species, the young are quite helpless when hatched as they are blind, unable to locomote, and unable to feed themselves. Prolonged parental care is required for normal development. It is generally believed that biparental care has evolved to meet the energetic and other demands of the developing altricial young.

Mammals. Parental care, especially maternal care, is highly developed in all species of mammals. By definition, the females of all mammalian species possess mammary glands for the nourishment of young. It is probably for this physiological reason that role reversal is less common among mammals than in other taxa.

Parturition in mammals is characterized by periods of uterine contractions, emergence of the young, delivery, and care or consumption of the placenta. Often there is appreciable squatting or straining during delivery. Prior to, during, and after delivery, there is much licking of the female's body, the newly delivered young, previously born young, and the umbilical cord. Often the female ingests the fetal membranes and the cord.

Nursing is initiated soon after birth. Typically the mother approaches and adopts the nursing posture. In cats, for example, the female lies near the kittens and arches her body about them. The young may develop specific nipple preferences. Later in development there may be mutual approach by the mother and young. The young continue to become more active in initiating nursing until finally the mother weans them.

Mammalian parental behavior also involves behavioral patterns other than nursing. Nests of varying degrees of complexity are built by various species. In rabbits, for example, the female digs or excavates a burrow and collects straw or other material with which to line it. At this stage, the female's hair begins to loosen and she plucks it and uses it to line the inside of the nest. Mothers may visit the young to nurse them only once per day, but the young are kept warm and safe. Of course, the mother must also feed, maintain social relationships, and engage in other activities necessary for survival. Her activities may be conceptualized as a kind of time sharing, the pattern of which changes as the young mature.

The manner of transporting the young varies with different species. Typically, a young rodent or canid is carried in its mother's mouth. Baboon mothers carry young on their chests, with a transition to riding on the mother's back, jockey style, that begins at about 5 weeks of age. The young of precocial mammalian species, such as most ungulates, are not transported by the mother.

Parents not only defend young from predators and conspecifics but also play an active role in tutoring them. This is most obvious in carnivorous species. A mother cat, for example, uses a method of shaping in encouraging her young to hunt. She first brings prey back to the nest and eats in front of them. Later, she leaves prey for the young to eat and still later presents them with live prey to kill for themselves. Developing carnivores accompany parents on some hunts and learn to catch prey in their company. Similarly, sea otters stay close to their mothers as they dive for food and eventually develop skills of diving and locating prey themselves. Young rodents often accompany adults on feeding trips and learn what to eat via cues from them.

The role of male mammals in caring for young has been increasingly recognized. In many species of rodents, for example, males engage in virtually all parental behaviors with the exception of nursing. The males of various species of primates and carnivores play an important and active role in the care and tuition of young.

Parental responsibilities may not end with the maturation of the young; grandparents play a role in development in some species. In vervet monkeys, for example, the presence of a maternal grandmother in a troop is associated with an increase in the survival of the infants of young adult females. Grandmothers play an active role in the development of the young, with the interaction being more prolonged with granddaughters than with grandsons.

Mechanisms of parental behavior. The occurrence and integration of cycles of parental behavior are a function of reciprocal interactions between the internal and external environments of the parents and young. Fluctuating hormone levels appear important in stimulating parental behavior and are, in turn, affected by the behavior of the partner and the young. In ring doves, a pair-forming species, the birds select a nest site, build a nest, and copulate. The male gathers most of the nesting material, and the female does most of the building. As the female's oviducts grow, she becomes progressively more attached to the nest site. When the eggs are laid, the members of the pair take turns incubating them, with the female on the eggs about three-quarters of the time. Once they are hatched, the young are fed with crop "milk," a regurgitated substance produced in the lining of the crop. The parents become progressively less interested in the young as they develop until the young fledge and a new cycle begins.

The various stages of the ring dove cycle are correlated with dramatic motivational changes. The bird that engaged in intensive nest building at one stage is equally intense about incubation at the next. These motivational changes are correlated with fluctuating hormone levels. In females, estrogen is associated with nest building and progesterone with incubation. These hormones have been shown to act on specific brain regions in females.

In mammals, too, fluctuating hormone levels are related to maternal behavior. Nest building, retrieval of pups, nursing, and grooming can be induced in laboratory rodents with various schedules of administration of estrogen, progesterone, and prolactin. Cross-transfusion of blood from a mother to a nonmaternal female stimulates maternal behavior in the recipient.

Also in mammals, various brain structures are involved in maternal behavior. With lesions of the medial preoptic area of the hypothalamus, for example, maternal rats fail to build nests, retrieve pups, or nurse. With lesions in certain portions of the limbic system of the forebrain, the behavioral patterns related to maternal behavior occur but in disrupted sequences, so that a mother may pick up and deposit pups but never complete their return to the nest. The timing of mammalian parental behavior is the

product of integration of the nervous and endocrine systems and external stimuli.

Parental investment. Parental investment entails any investment by the parent that increases the ability of the young to survive and reproduce at some cost to the parent. Much parental investment, like milk, cannot be shared; that which is given to one offspring cannot be given to another. Other kinds of parental investment, like defense against predators, is shareable. In addition, the parent making any investment is prevented from engaging in other activities, such as searching for his or her own food or seeking additional mates.

Because they contribute larger gametes and often engage in more extensive parental behavior, the females of most species display a higher level of parental investment than males. Members of the sex investing more (typically females) thus become a limiting resource for the sex investing less (typically males). It is generally agreed that this gives rise to the disparity between female and male reproductive strategies—with males more often competing for access to females and females more choosy than males.

In many nonmonogamous species, most females reproduce whereas some males may sire many offspring and others very few. Thus, variance in reproductive success is much greater among males than females. This difference in variance implies that sexual selection acts more strongly on males than females and relates to differences in parental investment and mating strategies.

Because natural selection works at the level of the individual, the interests of parent and young do not always coincide. In some situations it may be in the parent's best interest to spread resources equally among her young. Up to a point, however, it may be in the interest of a juvenile to demand a higher proportion of these resources for itself. This discrepancy of interests can generate parent-offspring conflict and the changing relationship between parent and young that leads to weaning.

Thus, the orderly integrated cycles of reproductive behavior generally require the cooperation of different individuals and result in the perpetuation of the species. Driving this activity, however, is the action of natural selection shaping individual organisms to behave in ways that will maximize their levels of fitness—contributions to the gene pools of future generations. Much of reproductive behavior can be understood in relation to these conflicting forces. The result, however, is a pattern that for each surviving species is adapted to its environment and social structure and results in the successful production of young. *See* BEHAVIORAL ECOLOGY; ETHOLOGY.

Donald A. Dewsbury

Bibliography. J. B. Becker, S. M. Breedlove, and D. Crews (eds.), *Behavioral Endocrinology*, 1992; E. M. Blass (ed.), *Developmental Psychobiology and Behavioral Ecology*, 1988; D. Crews (ed.), *Psychobiology of Reproductive Behavior: An Evolutionary Perspective*, 1987; D. A. Dewsbury (ed.), *Mammalian Sexual Behavior: Foundations for Contemporary Research*, 1981.

Reproductive system

The structures concerned with the production of sex cells (gametes) and perpetuation of the species. The comparative anatomy, human histology, embryology, physiology, endocrinology, and biochemistry of this system are treated in this article.

The reproductive function constitutes the only vertebrate physiological function that necessitates the existence of two morphologically different kinds of individuals in each animal species, the males and the females (sexual dimorphism).

The purpose of the reproductive function is fertilization, that is, the fusion of a male and a female sex cell produced by two distinct individuals. In each sex the reproductive system comprises a sex gland or gonad, which produces sex cells, or gametes, and ducts, which permit the passage of the gametes. In some animals, such as mammals, copulatory organs permit the male germ cells to be introduced into the female ducts and fertilization is internal, but in many vertebrates, such as anuran amphibians and many fishes, no copulatory organ exists and fertilization is external.

Comparative Anatomy

Egg cells, or ova, and sperm cells, or spermatozoa, are formed in the primary reproductive organs, which are collectively known as gonads. Those of the male are called testes; those of the female are ovaries. Besides giving rise to reproductive cells, both ovaries and testes give off endocrine secretions, or sex hormones, which pass into the blood or lymphatic streams and are carried to all parts of the body, where they bring about profound effects, not only on the rest of the reproductive system but on several other systems of the body as well. The gonads are paired structures, although in some forms what appears to be an unpaired gonad is the result either of fusion of paired structures or of unilateral degeneration.

The reproductive elements formed in the gonads must be transported to the outside of the body. In most vertebrates, ducts are utilized for this purpose. These ducts, together with the structures that serve to bring the gametes of both sexes together, are known as sex organs. The structures used to transport the reproductive cells in the male are known as deferent ducts and those of the female as oviducts. In a few forms no ducts are present in either sex, and eggs and sperm escape from the body cavity through genital or abdominal pores. The deferent ducts are usually the mesonephric or Wolffian ducts, which in some cases also serve to carry urinary wastes in those vertebrates in which opistonephros or mesonephros function either during embryonic or adult life. In vertebrates whose functional adult kidney is the metanephros, the Wolffian duct on each side persists as the ductus deferens. *See* KIDNEY.

In most vertebrates the reproductive ducts in both sexes open posteriorly into the cloaca. In some, modifications of the cloacal region occur and the ducts open separately to the outside or, in the male,

join the excretory ducts to emerge by a common orifice. *See* COPULATORY ORGAN; REPRODUCTION (ANIMAL).

The sex of an individual is dependent upon the chromosomes received from both parents at the time that the egg is fertilized. That the balance between maleness and femaleness is a delicate one, however, is reflected in the fact that environmental factors may assume an influential role in sexual development, and hormonal secretions may modify the extent to which various structures and even behavioral characteristics develop and are maintained.

Ovaries. A typical mammalian ovary is a solid, irregularly shaped structure indistinctly separated into an inner medulla and an outer cortex. The cortex contains numbers of ovarian follicles in various stages of development. *See* OVARY.

In certain fishes and in amphibians, snakes, and lizards, the ovaries are hollow, saccular structures. The cavity within the teleostean fish ovary is actually a closed-off portion of the coelom (body cavity) into which ripe ova are shed. This is not true of the saccular ovaries of other forms.

Cyclostomes. The adult female lamprey has a single ovary, representing a fusion of two, which runs the length of the body cavity, suspended from the middorsal body wall by a single mesovarium. At the height of the breeding season it fills the greater part of the body cavity; ripe eggs are shed into the coelom and fertilization is external. The hagfish is hermaphroditic; the anterior part of the single gonad is ovarian and the posterior part testicular. Usually only one or the other region matures.

Fishes. The ovaries of most fishes are paired, although in some cases they have fused into a single organ. The large eggs of elasmobranchs are discharged from the anteriorly located ovaries directly into the body cavity. In oviparous and ovoviviparous species, following ovulation the ovarian follicles become transformed into corpora lutea, structures which presumably have an endocrine function and may play a role in the extended retention of eggs in the oviducts of oviparous species and of young in the uteri of ovoviviparous forms. Peritoneal folds form in connection with each ovary in teleosts, closing off all connection with the coelom. The anterior end of the ovarian cavity ends blindly, but in most cases continuations of the folds at the posterior end form an oviduct which opens directly to the outside. Ripe ova, sometimes numbered in the millions, are discharged into the central ovarian cavity, which is actually a part of the body cavity, and thence pass down the oviducts to the outside. The garpike *Lepisosteus* is the only ganoid fish with a saccular ovary. The ovaries are usually solid, flat elongate structures from which mature ova break out into the coelom.

Amphibians. Although the paired amphibian ovaries are saccular structures, ripe ova are liberated into the body cavity through their external walls. The shape of the ovaries varies with the shape of the body. They are long and narrow in caecilians, elongated to a lesser degree in salamanders, and short and more compact in frogs and toads. Fat bodies are associated with amphibian ovaries. They serve for the storage of nutriment and undergo profound changes during the year. A peculiar structure in the male toad, known as Bidder's organ, may under certain conditions develop into a true ovary.

Reptiles. The saccular ovaries of snakes and lizards are similar to those of amphibians, whereas turtles and crocodilians have solid ovaries. In snakes and lizards they are elongated but not symmetrically disposed. Only the yolk of reptilian eggs is formed in the ovaries, and this represents the true ovum. The size of the eggs is generally in proportion to that of the animal. In certain ovoviviparous snakes and lizards, corpora lutea form from ruptured follicles after ovulation and persist throughout pregnancy. These probably secrete a hormone necessary for maintenance of pregnancy.

Birds. Although both ovaries are present during embryonic development in most birds (except many birds of prey) the right ovary degenerates and only the left is functional. In birds, stalks extend from the surface of the ovary, and each stalk contains many ovarian follicles in various stages of development. A mature ovum escapes from the ovarian follicle through a preformed nonvascular band, the stigma or cicatrix, located on the surface of the follicle opposite the stalk. As in reptiles, only the yolk of the egg represents the true ovum. Increase in the number of hours of daylight stimulates ovarian activity in many birds. Even brief exposure to intense and bright light during hours of sleep increases production of eggs in the domestic fowl. This effect is undoubtedly mediated through stimulation of the pituitary gland. If the functional ovary of the domestic fowl is removed, the rudimentary right gonad will develop into a testislike organ, but without germ cells. *See* SEXUAL DIMORPHISM.

Mammals. The ovaries of mammals are located in the lumbar or pelvic regions and are small in comparison to the size of the body. The relationship of the microscopic mammalian ovum to the ovarian follicle differs somewhat from conditions in other vertebrates. Follicles in various stages of development, with the youngest near the surface of the ovary, are depicted in **Fig. 1**. At periodic intervals one or more follicles grow to maturity, rupture, and liberate their ova into the body cavity. In such animals as the rabbit, cat, and ferret, ovulation will not occur unless the animal copulates. Following ovulation certain cells of the follicle undergo a transformation and the entire structure becomes a more or less solid body, the corpus luteum.

If pregnancy does not occur, the corpus luteum persists only for a short time. If pregnancy does ensue, the corpus luteum usually persists throughout pregnancy. In either case it ultimately degenerates. The corpus luteum is of primary importance as an endocrine gland secreting a hormone called progesterone. *See* PREGNANCY.

Oviducts and associated structures. Oviducts, except in teleosts and a few other fishes, are modifications of Müllerian ducts formed early during

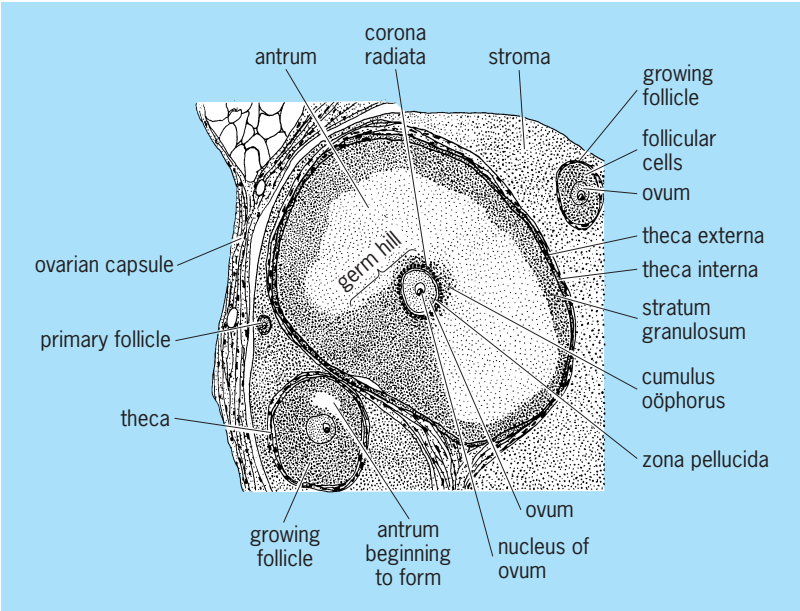

Fig. 1. Section of portion of cortex of rat ovary, showing ovarian follicles in various stages of development, the youngest being near the surface of the ovary. (*After C. K. Weichert, Elements of Chordate Anatomy, 3d ed., McGraw-Hill, 1967*)

embryonic development. In most cases each duct is formed by an invagination of the peritoneum that covers the ventrolateral part of the opisthonephros or mesonephros. The edges of the groove fuse to form a tube which joins the cloaca or forms a uterus posteriorly and remains open anteriorly to become the ostium tubae. Although Müllerian ducts also form in the male, they ordinarily degenerate except for a few vestigial remnants. In some they persist as prominent but nonfunctional structures.

Cyclostomes. Oviducts are lacking in cyclostomes. Ova pass from the coelom through genital pores into a urogenital papilla and then to the outside.

Fishes. Much diversity exists in the oviducts of fishes. In some teleosts, and in a few other fish, eggs escape from the body cavity through modified abdominal pores. In elasmobranchs the two Müllerian ducts may fuse at their anterior ends so that only a single ostium tubae connects with the coelom. An enlargement, known as the shell gland, is present in the upper part of each oviduct. Beyond the shell gland the Müllerian duct enlarges on each side to form a uterus which opens into the cloaca.

The oviducts of most teleosts are short and continuous with the cavities of the saccular ovaries. It is doubtful whether they are true Müllerian ducts because they are formed in a different manner. A cloaca is lacking in teleosts, and the oviducts open independently to the outside. The two oviducts usually fuse, continuing posteriorly as a single structure which may open to the outside through a genital pore or else at the tip of a genital papilla. Most teleosts are oviparous, but many are ovoviviparous. Although the young may develop within the cavities of the ovaries, intrauterine development is more common. The size of the oviducts fluctuates markedly with the seasons. They are naturally largest during the breeding period.

Amphibians. Oviducts in amphibians are paired elongated tubes, each with an ostium tubae situated well forward in the body cavity. The posterior end of each oviduct is enlarged slightly to form a short uterus which opens into the cloaca. In almost all forms the uteri serve only as temporary storage places for ova that are soon to be laid. In some toads the two uteri unite before entering the cloaca by a common orifice. Marked fluctuation in size of the oviducts is apparent at different seasons (**Fig. 2**). During the breeding period they become elongated and coiled. The glandular lining secretes a clear gelatinous substance (jelly) which is deposited about each ovum as it passes down the oviduct. External fertilization is the general rule in frogs and toads, but in salamanders, with few exceptions, internal fertilization takes place. No copulatory organs are present. A diverticulum, the spermatheca, of the salamander's cloaca serves as a storage place for spermatozoa. The males deposit spermatophores (small packets of spermatozoa) which are taken into the cloaca of the female by muscular movements of the cloacal lips. Internal fertilization occurs in caecilians, the male of which has an eversible cloaca which may serve as a copulatory organ.

Reptiles. The paired oviducts of reptiles open into the coelom through large, slitlike ostia. Each oviduct is differentiated into regions which mediate different functions in forming the layers of materials deposited about the ova prior to laying. The eggshell, which is of parchmentlike consistency, except in some lizards and in crocodilians in which it is hard, is formed in the uterus. Fertilization is always internal in reptiles, copulatory organs being present in the male, with the exception of the primitive *Sphenodon*. Most reptiles are oviparous, but many snakes and lizards are ovoviviparous.

Birds. The right oviduct usually degenerates in birds and only the left one is functional (birds of prey being exceptions). The egg enters the coiled oviduct through the ostium and passes through a glandular region, an isthmus, and a uterus. Two layers of albumen are deposited about the ovum in the glandular region; inner and outer shell membranes and more albumen are laid down in the isthmus; the hard calcareous shell is formed in the uterus. In the hen these

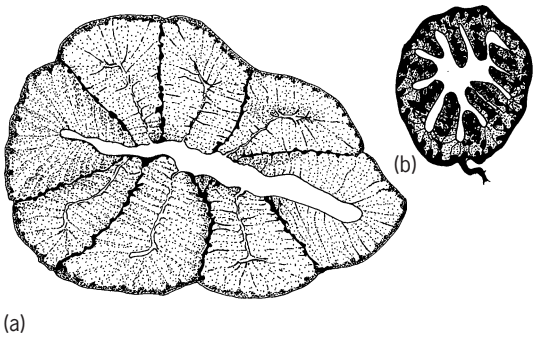

Fig. 2. Cross section of oviduct of the salamander *Eurycea bislineata*. (*a*) During breeding season. (*b*) After breeding season. (*After C. K. Weichert, Elements of Chordate Anatomy, 3d ed., McGraw-Hill, 1967*)

processes take 21–23 h. Fertilization is internal in birds and, in a majority of species, is accomplished by cloacal apposition of the two sexes.

Mammals. Paired Müllerian ducts develop in all mammalian embryos. Each differentiates into an anterior, nondistensible Fallopian tube and a posterior, expanded uterus. In all mammals except monotremes the uterus leads to a terminal vagina which serves for the reception of the penis of the male during copulation.

The lower part, or neck, of the uterus is usually telescoped into the vagina to a slight degree. This portion is referred to as the cervix.

In the primitive monotremes a cloaca is present in the adult, and each uterus terminates independently in a urogenital sinus anterior to the region where ureters and bladder enter. Marsupials retain the primitive paired condition, but two vaginas open into a common urogenital sinus. In some mammals, the kangaroo for example, the vaginas fuse at their upper ends to form a vaginal sinus which extends posteriorly as a blind pocket or tube. It is sometimes referred to as a third vagina and may serve as the birth canal. Its posterior end connects with the urogenital sinus. In the event that this pouchlike structure has no opening, its wall ruptures at the time of delivery, thus permitting the young, lodged in it, to pass directly into the urogenital sinus. Placental mammals have a single vagina which represents a fusion of two; only the pika, which is a lagomorph, has a cloaca in the adult form.

The uterine portions of the Müllerian ducts fuse to varying degrees, resulting in different types of uteri (**Fig. 3**). The simplex type, found in apes and humans, represents the greatest degree of fusion. Only the bilaterally disposed Fallopian tubes indicate the paired origin of this type of uterus. Anomalous uteri of the duplex, bipartite, and bicornuate types are occasionally encountered in the human being.

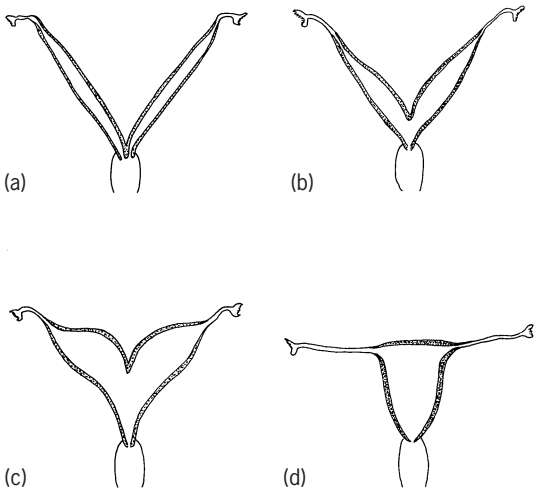

Fig. 3. Degrees of fusion of the uterine portions of the Müllerian ducts in four types of mammalian uteri. (*a*) Duplex (rat). (*b*) Bipartite (pig). (*c*) Bicornuate (horse). (*d*) Simplex (human). (*After C. K. Weichert, Elements of Chordate Anatomy, 3d ed., McGraw-Hill, 1967*)

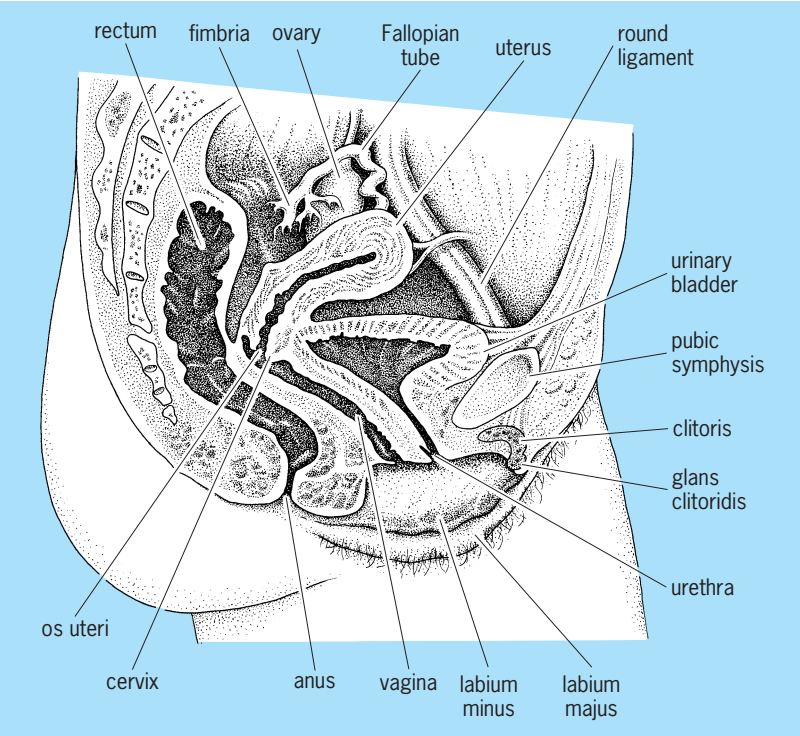

Fig. 4. Urogenital system of the human female. (*After C. K. Weichert, Elements of Chordate Anatomy, 3d ed., McGraw-Hill, 1967*)

The urethra, coming from the bladder, may join the vagina to form a urogenital sinus, or vestibule, which opens to the outside. Often, a more or less complete fold of mucous membrane, called the hymen, marks the border of the vagina where it opens into the vestibule. The external part of the female reproductive system is called the vulva. In primates, two folds of skin, the labia minora, are located about the margins of the opening of the vestibule. In some apes, and in the human female (**Fig. 4**), two additional outerfolds, the labia majora, make up part of the vulva. A small erectile organ, the clitoris, homologous with the penis of the male, lies ventral to the vaginal orifice. It differs from the penis, however, in that it has no connection with the urethra except in a few forms such as rats and mice.

In addition to the above structures, certain glands, as well as rudiments remaining from the degenerated mesonephros, are associated with the female reproductive system. Glands of Bartholin correspond to Cowper's glands (bulbourethral glands) of males. They open into the vestibule near the hymen and secrete a clear, viscid fluid under sexual excitement. This serves as a lubricant during copulation. Paraurethral glands (of Skene), corresponding to the prostate glands of males, are occasionally encountered in females but are inconstant in appearance and of doubtful function. Small, mucus-secreting, vestibular glands, homologous with the glands of Littré of the male, are located around the opening of the urethra and on the clitoris. Among the remnants of the degenerated mesonephros that are associated with the reproductive organs of the female

is the epoophoron, a complex of degenerate anterior mesonephric tubules connecting to a persistent portion of the Wolffian duct. The paraoophoron, a similar group of more posterior tubules, is located farther caudad. Both epoophoron and paraoophoron are situated in the broad ligament near the ovaries. A canal of Gärtner, located in the wall of the uterus or vagina, represents a vestige of the Wolffian duct proper. None of these structures is functional.

The epithelium that lines the Fallopian tube consists of two kinds of cells: ciliated cells, which are most numerous on the fimbriated funnel surrounding the ostium tubae, and others of a glandular nature. No true glands are present here, however. The beat of the cilia is abovarian in direction, and this, together with the peristaltic muscular contractions, serves to transport ova down the Fallopian tube. In most species of mammals it takes 3 or 4 days for fertile ova to pass down the Fallopian tube to the uterus. If fertilization does not occur, the ova die and probably disintegrate before they have even reached the uterus.

Testes. The typical testis is a compact organ which varies greatly in shape in different groups of vertebrates. In all but a few primitive forms each testis is composed of numbers of seminiferous ampullae or seminiferous tubules which connect by means of ducts to the outside. Spermatozoa are formed within the tubules or ampullae. *See* TESTIS.

In seasonal breeders the size of the testes fluctuates with the seasons, being largest just before the breeding season. After this period they shrink to only a fraction of their former size. Each testis is suspended from the middorsal body wall by a membrane, the mesorchium.

Cyclostomes. The gonad of the male lamprey differs little in general appearance from that of the female. It is a single organ, representing a fusion of two, and is suspended from the middorsal body wall by a single mesorchium. Even when fully developed the testis does not become as voluminous as does the ovary in the female. Spermatozoa are discharged into the body cavity from which they escape via genital pores through a urogenital papilla. The hermaphroditic gonad of the hagfish has already been mentioned.

Fishes. In elasmobranchs, the paired testes are relatively small compact structures located at the anterior end of the coelom. In most other fishes they are elongated and often lobulated. Ducts transport spermatozoa to the outside. Male gonads vary greatly in size and are extremely large at the breeding season. In some teleosts, interstitial cells lie in spaces among the seminiferous ampullae or tubules; in others they form part of the walls of ampullae and show seasonal variation in appearance and activity.

Amphibians. The shape of the testes of amphibians is roughly correlated with that of the body. Thus, in caecilians, each testis is elongated and resembles a string of beads. The enlargements consist of masses of seminiferous ampullae connected by a longitudinal collecting duct. In salamanders the testes are shorter and irregular in outline; in frogs and toads they are small, oval, compact structures. A pronounced difference in size is apparent during the breeding season. Fat bodies are associated with the gonads of male as well as of female amphibians.

Reptiles. Reptilian testes are round, oval, or pyriform in shape and contain seminiferous tubules that are long and convoluted. In snakes and lizards one testis usually lies farther forward in the body cavity than the other. Periodic fluctuations in size of the testes are typical.

Birds. The round or oval shape of bird testes is characteristic. In the domestic fowl, the testes function throughout the year and no periodic variations in size are obvious. Most birds, however, are seasonal breeders, the testes enlarging conspicuously at the approach of the breeding season. Increase in the number of hours of daylight stimulates spermatogenesis in certain birds and hence brings about testicular enlargement, this generally occurring in spring.

Mammals. In all mammals except monotremes the oval-shaped testes move from their place of origin to the pelvic region, where they may remain permanently, or they may descend farther into a pouchlike scrotum. In many seasonal breeders the testes are located in the scrotum only during the breeding period. The scrotum serves as a temperature regulator, providing an environment for the testes several degrees below that of the body. This seems to be a requirement for normal development of spermatozoa. In marsupials the scrotum lies anterior to the penis but in other mammals it is posterior to that organ. In several species a relation between the number of hours of daylight and testicular activity has been demonstrated.

Male ducts. The ducts which in most vertebrates serve to transport spermatozoa to the outside of the body are the archinephric ducts or Wolffian ducts formed in connection with the opisthonephros or mesonephros, respectively. Their original function is elimination of urinary wastes. In some fishes and amphibians certain modified kidney tubules are employed in carrying spermatozoa from the testis to the archinephric duct. The tubules are known as efferent ductules, and the duct is termed the ductus deferens. The male ducts undergo profound changes in size in seasonal breeders (**Fig. 5**). Reproductive ducts are lacking in cyclostomes.

Fishes. A variety of conditions is encountered in the reproductive system of male fishes. In elasmobranchs small efferent tubules course from the testis through the mesorchium, connecting with some anterior kidney tubules along the medial border of the opisthonephros. The kidney tubules lead to the archinephric duct, which now serves almost entirely as a ductus deferens. The ductus deferens courses along the ventral side of the opisthonephros. In young specimens it is a straight tube with a urinary function; in older individuals it is highly convoluted. The posterior end is markedly dilated to form a seminal vesicle. The two seminal vesicles open into a common urogenital sinus which enters the cloaca through an aperture at the tip of a urogenital papilla. A pair of blind sperm sacs projects forward from the

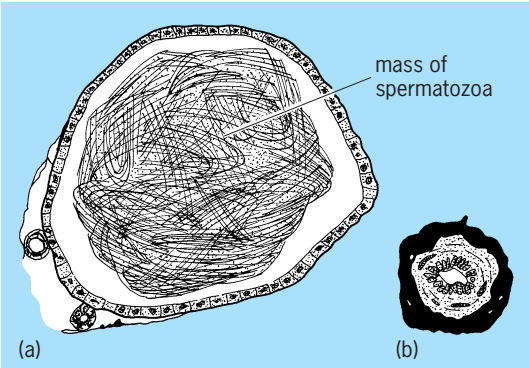

Fig. 5. Cross section of ductus deferens of male salamander *Eurycea bislineata*. From specimen obtained (*a*) just prior to breeding season; (*b*) several weeks after breeding season. (*After C. K. Weichert, Elements of Chordate Anatomy, 3d ed., McGraw-Hill, 1967*)

ventral wall of the urogenital sinus. These sacs may be remnants of the Müllerian ducts which persist in the male.

In most other fishes the kidney ducts serve only for the passage of urinary wastes. The sperm duct, which is not a true ductus deferens because it is formed in a different manner, may be entirely independent of the kidney duct, although the two may have a common opening to the outside.

Amphibians. The arrangement of the male ducts in amphibians rather closely resembles that of elasmobranch fishes. In salamanders the ductus deferens courses outside the lateral border of the kidney (**Fig. 6**). Small efferent ductules from the testis join Bidder's canal, a duct lying medial to the kidney. This in turns connects to the ductus deferens by means of modified kidney tubules. In frogs and toads, both Bidder's canal and the ductus deferens lie within the opisthonephros. A Bidder's canal is also present in the female, but its function, if it has any, is obscure. In caecilians the longitudinal collecting duct, which connects the lobules of the testis, gives off small, transverse canals between successive lobules. These pass to the kidney to join another longitudinal duct which runs along the lateral edge of the opisthonephros. Spermatozoa then pass through a second series of transverse canals to join kidney tubules which transport them to the archinephric duct which joins the cloaca posteriorly. In males of several species of frogs and toads a dilation of the archinephric duct, or ductus deferens, as it nears the cloaca, forms a seminal vesicle in which spermatozoa may be stored temporarily. Ductus deferens, and seminal vesicles when present, undergo striking seasonal variations in size, being controlled by changes in the level of testosterone, the hormone secreted by the testes. The cloacal glands of male urodeles, which secrete a jellylike material used in forming the spermatophores deposited by the male, become markedly enlarged as the breeding season approaches.

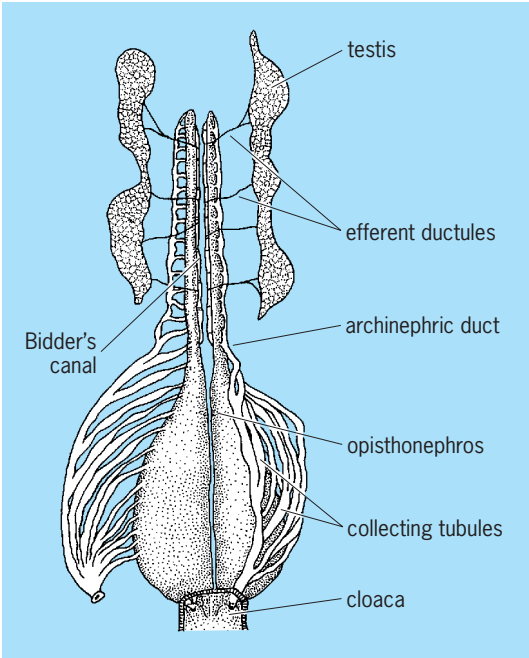

Fig. 6. Urogenital organs of male salamander, ventral view. The collecting ducts on the right side are shown detached from the cloaca and spread out for clarity. (*After C. K. Weichert, Elements of Chordate Anatomy, 3d ed., McGraw-Hill, 1967*)

Reptiles. When the embryonic mesonephros degenerates in male reptiles, its duct (archinephric or Wolffian duct) persists as the reproductive duct. The end near the testis becomes greatly convoluted and is called the epididymis. This connects with the testes by means of a few persistent mesonephric tubules which serve as efferent ductules. The remainder is the ductus deferens, sometimes straight, sometimes convoluted. In snakes and lizards this joins the ureter of the metanephros before entering the cloaca. In turtles and crocodilians these ducts open at the proximal end of a groove which carries spermatozoa to the free end of the penis. Seasonal variations under endocrine control are obvious in the epididymides and deferent ducts of most reptiles. In many, some of the posterior urinary tubules of the metanephros enlarge during the breeding season, producing an albuminous secretion which contributes to the seminal fluid. Glandular secretions from the cloacal walls in snakes and lizards pass into the grooves of the hemipenes. Nonfunctional Müllerian ducts, usually much reduced in size, commonly persist in male reptiles.

Birds. The reproductive ducts of male birds are essentially similar to those of reptiles but open independently into the cloaca. In the few birds that possess a penis, a groove on the upper surface carries spermatozoa to the apex. In some passerine birds, a nodule, composed of a tightly coiled portion of the ductus deferens, protrudes into the cloaca. The temperature of the nodule is somewhat lower than body temperature. This may be of importance in the maturation of spermatozoa.

Mammals. A few persistent mesonephric tubules connect the seminiferous tubules of each testis to a compactly coiled epididymis which is continuous with the ductus deferens. In man the portion of the

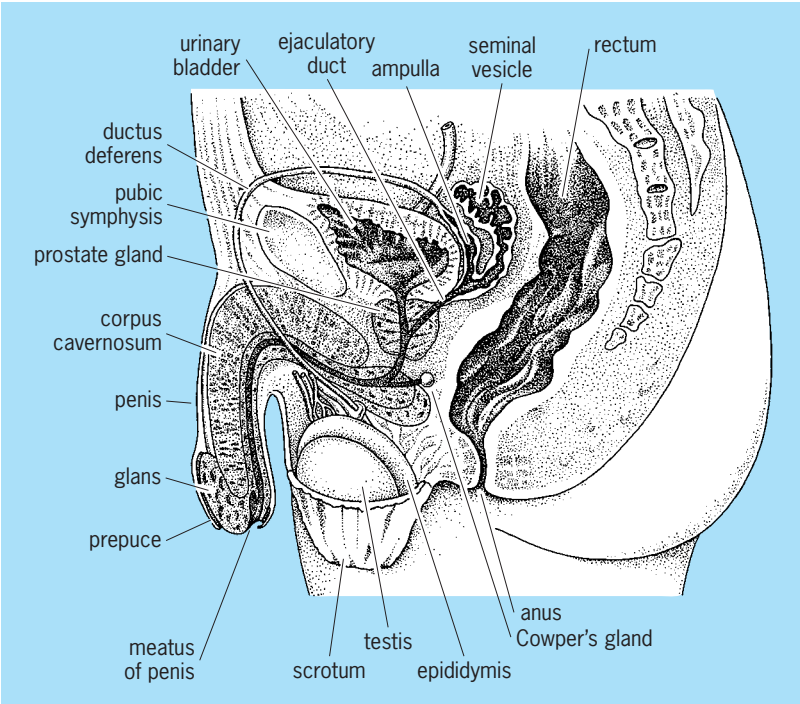

Fig. 7. Urogenital system of the human male. (*After C. K. Weichert, Elements of Chordate Anatomy, 3d ed., McGraw-Hill, 1967*)

Wolffian duct included in the epididymis averages 20 ft (6 m) in length. Experiments have demonstrated that, during their passage through the epididymis, spermatozoa, although remaining quiescent, acquire the capability of becoming fully motile. In mammals having scrotal testes, the ductus deferens enters the body cavity lying between the peritoneum and the body wall. It crosses in front of the ureter, loops over that structure, and then courses posteriorly for a short distance before joining the urethra (**Fig. 7**). This peculiar arrangement is the result of the descent of testes from their original abdominal location. In many mammals the ductus deferens enlarges at its posterior end to form an ampulla. A glandular seminal vesicle may arise on each side as a saccular diverticulum of the ductus deferens. It does not store spermatozoa, but its secretion contributes to the seminal fluid in which spermatozoa are suspended. Seminal vesicles are absent in monotremes, marsupials, carnivores, and whales. The lower portion of the ductus deferens, between seminal vesicle and urethra, is often called the ejaculatory duct. The urethra, coming from the bladder, extends the length of the penis, opening at the tip of that organ through a small meatus. It serves both for passage of urine and seminal fluid. Accessory glands associated with the urethra include the prostate gland, which contributes to the seminal fluid; Cowper's glands, which secrete a clear, viscid fluid during sexual excitement; and small, mucus-secreting urethral glands. In some mammals, as in the rat, when the secretions of the seminal vesicles and prostate gland are intermingled they coagulate, so that just after mating a so-called copulation plug is present in the vagina. This aids in retaining spermatozoa in the reproductive tract of the female. The development and functioning of the accessory sex organs in mammals are clearly under control of the hormone testosterone. A few remnants of the mesonephros may persist in male mammals in close relation to the reproductive systems. Among these are the paradidymis, a few aberrant ductules, and the appendix of the epididymis. Persistent homologs of the Müllerian ducts may also be present in males. The small appendix of the testis and the prostatic utricle represent such remnants.

Copulatory organs. Fertilization always takes place in a fluid medium. In many aquatic vertebrates external fertilization takes place, with water furnishing the medium in which sperm travel to gain access to ova. Terrestrial forms and even numerous aquatic species have internal fertilization. Secretions supplied by both sexes furnish the necessary fluid medium. In a number of terrestrial vertebrates, spermatozoa are transferred from male to female by cloacal apposition. In most, however, as well as in many aquatic species, intromittent, or copulatory, organs are employed to deposit spermatozoa, suspended in seminal fluid, within the reproductive tract of the female.

Fertilization usually takes place in the upper part of the oviduct. In forms in which the egg is surrounded by a shell, fertilization must take place before the shell has been formed about the ovum.

Fishes. In those fishes having internal fertilization, the copulatory organs are modifications of fins, pelvic fins in elasmobranchs and anal fins in teleosts. Claspers, which are modifications of pelvic fins, are used as intromittent organs by male elasmobranchs. Each clasper is essentially a scroll-like tube through which seminal fluid may be ejected with some force into the cloaca of the female. It is probable that only one clasper is inserted at a time. In teleosts having internal fertilization, the anterior border of the anal fin is elongated posteriorly to form an intromittent organ, the gonopodium. In some teleosts, modifications of the hemal spines of certain caudal vertebrae form a copulatory organ.

Amphibians. Although internal fertilization occurs in most salamanders, copulatory organs are lacking. By muscular action of the cloacal lips, the female is able to pick up spermatophores, or packets of spermatozoa, deposited by the male. The eversible and protrusible cloaca of some male caecilians may be used as a copulatory organ when the cloacae of both sexes are in apposition.

Reptiles. *Sphenodon* is the only reptile lacking copulatory organs. In others, two types of structures are evident. Snakes and lizards employ paired hemipenes, which are saclike structures, devoid of erectile tissue, lying under the skin adjacent to the cloaca. Each bears a groove for the passage of spermatozoa. These organs are everted during copulation and spermatozoa are deposited in the cloaca of the female. Hemipenes are not homologous with the single penis of turtles and crocodilians, which is basically similar to that of mammals, and contains erectile tissue which becomes distended with blood during sexual excitement. A homologous structure,

the clitoris, is present in rudimentary form in females. Paired, thickened ridges in the anterior and ventral walls of the cloaca are called corpora cavernosa. They are composed of connective and erectile tissue. A groove along the dorsal surface provides for the passage of spermatozoa. During the act of mating the corpora cavernosa are filled and distended with blood and the penis is firm and greatly enlarged. It is then said to be erect. This property of erectile tissue makes it possible for the penis to serve as an intromittent organ. Without erection, copulation is impossible. During erection the veins which normally drain blood from the erectile tissue become compressed. There is an increased supply of arterial blood supply accompanied by an impeded venous drainage. Return of the penis to its flaccid state after erection is termed detumescence.

Birds. Most birds copulate by cloacal apposition. A penis is present only in ducks, geese, swans, and ostriches. It is a single structure of the same type as that of turtles and crocodilians. A clitoris is present in females of these species. In many other birds a rudimentary penis can be identified.

Mammals. In monotremes the single penis lies on the floor of the cloaca. It is similar to the organ in turtles, crocodilians, and birds except that the groove on the dorsal side has become a closed tube. This is surrounded by another mass of erectile tissue known as the corpus spongiosum. The canal in monotremes possibly carries only seminal fluid since the urethra has a separate opening into the cloaca. The monotreme urethra thus differs from that of all other mammals. Marsupials lack a cloaca. The marsupial penis is covered by a sheath which opens to the outside of the body just beneath the anus. It may be protracted and retracted. The scrotum in marsupials is anterior to the penis. There are two erectile corpora cavernosa, separated by a septum, and a single erectile corpus spongiosum surrounding the urethra. Usually in marsupials three pairs of Cowper's glands open into the urethra at the base of the penis.

In higher mammals there is a tendency for the penis to be directed forward, and in all forms possessing a scrotum the penis is located anterior to the scrotum. In most mammals it lies within a sheath from which it can be protracted and retracted. In primates, the penis is permanently exserted. Its distal end bears a sensitive, swollen glans. Among mammals there are many variations in shape and structure of the glans, which contains erectile tissue continuous with the corpus spongiosum. In some forms, as in the cat, the glans bears numerous horny papillae or spines on its surface; these undoubtedly function as a sexual irritant. The glans is covered by a sheath, the prepuce, or foreskin, which may be retracted. Preputial glands secrete a sebaceous material called smegma about the base of the glans underneath the prepuce.

A bone is present in the penis of many rodents, carnivores, bats, whales, and lower primates. It is situated in the connective tissues between the corpora cavernosa and corpus spongiosum. In whales there is but a single corpus cavernosum. Anomalous development of the vertebrate penis is occasionally encountered. The most common anomalies are referred to as hypospadias and epispadias, in which the urethra opens on the under or upper sides, respectively, of the penis, rather than being continued to the tip of the glans. Charles K. Weichert

Histology in Humans

The human reproductive system comprises many different organs, and therefore all the primary types of tissue are identified.

Male reproductive system. In the following sections the histological features of the testes, scrotum, excretory ducts, and auxiliary glands are discussed.

Testes. The testes have two important structural elements, the contorted seminiferous tubules and the interstitial cells which are situated between the tubules. The contorted seminiferous tubules have two kinds of cells, nutrient cells, which also serve as supporting cells (the cells of Sertoli), and germinal cells, which develop into spermatozoa by the process of spermatogenesis. The interstitial cells of the testis are seemingly modified connective-tissue cells. They resemble epithelial cells and hence are correctly designated epithelioid cells. They produce testosterone, the male sex hormone, which chemically is a steroid compound, in mammals and perhaps in other vertebrates. *See* SPERMATOGENESIS.

Scrotum. The scrotum consists of skin, smooth muscle, and connective tissue. The scrotal skin is rich in sweat glands; it is thus able to act as a thermoregulator for the testes. The human scrotal temperature is about 13°F (7°C) lower than that of the abdominal cavity. The smooth muscle is a "skin muscle" called the dartos tunic; its relaxation lengthens the scrotum and promotes loss of heat, whereas its contraction shortens the scrotum and reduces loss of heat. *See* SWEAT GLAND; THERMOREGULATION.

Excretory ducts. The excretory ducts which convey spermatozoa from the contorted seminiferous tubules to the urethra are either intratesticular (tubuli seminiferi recti and rete testis) or extratesticular (epididymis, ductus deferens, and ejaculatory duct). From the seminiferous tubules, spermatozoa enter the epididymis, a tubular structure consisting of head, body, and tail. The head of the epididymis has 12 or more efferent ductules (ductuli efferentes testis), the epithelium of which bears cilia, or little hairlike projections, which assist in moving sperm into the coiled part of the epididymis. The body and tail of the epididymis constitute the ductus epididymis, a coiled tube which, when uncoiled, is about 15–20 ft (4.5–6 m) long. The ductus deferens in humans is about 18 in. (45 cm) when uncoiled. Its proximal end is dilated to form the ampulla ductus deferentis.

Auxiliary glands. The auxiliary glands in humans include two seminal vesicles, one prostate, and two bulbourethral glands. Their secretions together with spermatozoa and the small amount of secretion from the excretory ducts constitute the semen. In the male, one ejection of semen, known as ejaculation,

has a volume of 3–5 ml, and each milliliter contains about 6×10^7 spermatozoa. The seminal vesicles, named in accordance with the erroneous belief that they are receptacles for spermatozoa (after death, sperm may be found in the vesicles), are hollow glands which lie along the back wall of the prostate. They are about 2 in. (5 cm) in length, and, developmentally, each gland is an outgrowth of the ampulla of the ductus deferens. The right and left seminal vesicles open into the prostatic urethra via the ejaculatory ducts. The prostate is a solid gland which surrounds the urethra (prostatic urethra) at its origin from the urinary bladder. It is about 1–1.5 in. (2.5–4 cm) in diameter, and has several small ducts which open individually into the prostatic urethra. The bulbourethral glands, or glands of Cowper, are about 0.25 in. (6 mm) in diameter. Each opens into the urethra at a site slightly below the prostate.

In such species as the rat, guinea pig, and rhesus monkey, the mixture of the secretions from the auxiliary glands coagulates, as in the case of ejaculated semen. In the mating of rats and guinea pigs, the coagulated semen produces the vaginal plug which temporarily partly occludes the vagina of the female partner.

Female reproductive system. The histological discussion of the female system includes the ovaries, uterine tubes, uterus, vagina, and external genitalia.

Ovary. The ovary (ovarium), or egg producer, is largely covered by peritoneum; it has follicles with egg cells (ova), a stroma of connective tissue, and, in maturity, corpora lutea.

Maturation of the ovum. The process of the origin, growth, and formation of the ovum in its preparation for fertilization is known as oogenesis. The young sex cell is an oogonium. Growth produces the primary oocyte; only the beginning of the first maturation division occurs before ovulation. After ovulation, the primary oocyte divides to produce two unequal elements: a secondary oocyte and a smaller body, the first polar body. Then the secondary oocyte divides to yield the mature ovum (rarely called ootid) and the second polar body. The mature ovum has the haploid number of chromosomes; the two polar bodies eventually disintegrate. *See* OOGENESIS.

Corpus luteum. The corpus luteum is a yellow endocrine body which originates at the site of a ruptured Graafian follicle and which produces progesterone, one of the female sex hormones. It originates by the metamorphosis of granulosa cells and thecal cells into lutein cells which are epithelioid. The corpus luteum of ovulation persists about 14 days, whereas the corpus luteum of pregnancy lasts several months. The degeneration of a corpus luteum produces a white, fibrous scar in the ovary, the corpus albicans.

Fallopian tube. The human Fallopian tube is a muscular structure about 4.5–5 in. (11.4–12.7 cm) long, and has three regional subdivisions: a funnellike upper part, the infundibulum (infundibulum tubae uterinae); a dilated part below the infundibulum, the ampulla; and a constricted part near the uterine junction, the isthmus. The mouth of the tube, or ostium, communicates with the peritoneal cavity, and is guarded by a fringe, the fimbriae tubae. Fertilization of the ovum occurs in the tube.

Uterus. The uterus is covered in part by peritoneum. It has an outer layer of smooth muscle, the myometrium, and an inner layer of connective tissue and epithelium, the endometrium. Thickness of the endometrium varies during the menstrual cycle. *See* MENSTRUATION.

Vagina. The musculomembranous organ situated between the cervix uteri and the external genitalia is the vagina. During copulation, it ensheathes the penis. Ventrally and dorsally, the anterior and posterior fornices, respectively, overlap the uterine cervix. In youth, the lower end of the vagina is constricted a bit by a membranous shelf of mucosa, the hymen vaginae.

External genitalia. The external genitalia include the clitoris, vestibule of the vulva (fossa vestibuli vaginae), paired lips of the vulva (labium majus pudendi and labium minus pudendi), and a pair of small vestibular glands (glandulae vestibulares minores). The large lips of the vulva are essentially folds of skin, whereas the small lips are covered with a mucous membrane. The vestibular glands secrete mucus, and, in cases of gonorrhea, may become inflamed and then form large cysts. Lemen J. Wells

Embryology

The embryology of the reproductive system shows great similarities in all vertebrates, with the exception of teleost fishes, in which it is less specialized. It proceeds by three steps: sex determination; organogenesis of gonads; and later, during development, differentiation of the genital tract.

Sex determination. At the time of fertilization, the sex of each individual is genetically determined by certain genes carried by the sex chromosomes contained in the gametes. If both the male and the female germ cells have an identical X chromosome, the egg, upon fertilization, receives two X chromosomes (homozygous egg). If one parent germ cell contains an X and the other a Y chromosome, the egg cell receives an XY assortment (heterozygous egg). In mammals and in frogs, for instance, the male is the heterozygous XY sex. In birds and several urodeles the condition is reversed and the female sex is the heterozygous XY sex.

Sex determination should be considered as the first step in sexual differentiation of each individual. Techniques have been developed which permit direct study of the chromosomes, particularly the sex chromosomes, in dividing cells of any tissue. Chromosomal anomalies, such as the absence of one chromosome (for example, XO) or the presence of one or more supplementary chromosomes (for example, XXY) can be recognized in humans or in animals. Mosaicism results from the presence of two or more types of cells, for instance, cells with XX and cells with XY chromosomes, in the same individual. In such individuals the genetic basis of sex is disturbed, and embryological anomalies may occur. *See* CHROMOSOME; SEX DETERMINATION.

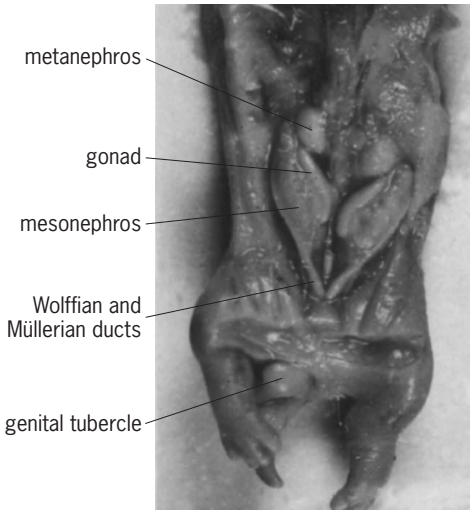

Fig. 8. Dissection of a 19-day-old rabbit fetus showing the undifferentiated condition of the genital tract.

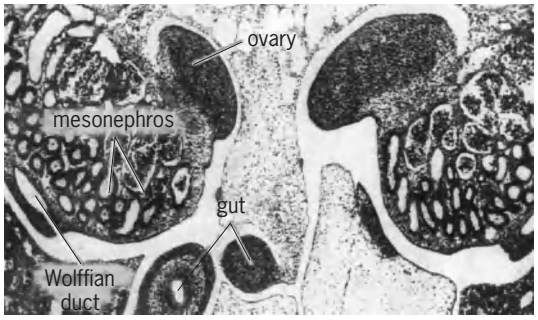

Fig. 9. Section through the body of a 17-day-old rabbit fetus, showing the ovaries at an early stage of differentiation located on the internal side of the mesonephroi.

Role of urinary system. The development of the genital tract is intimately correlated with development of the urinary system (**Figs. 8–11**); the urinary system itself is derived from successive excretory organs. The first pair of kidneys, the pronephroi of the early embryo, develops mainly in the future neck region. The excretory (drainage) duct of the pronephros grows posteriorly and reaches the cloaca. Shortly afterward the second pair of excretory organs, the mesonephroi, differentiates approximately in the middle of the trunk, in connection with the primitive pronephric duct. The pronephros then retrogresses, and its duct becomes the mesonephric duct, also called the Wolffian duct. In higher vertebrates (birds and mammals) the mesonephroi are later replaced by a third pair of kidneys, the metanephroi.

Among the vertebrates, except for the teleost fishes, the sex glands differentiate in connection with the mesonephros. They retain these connections in male amphibians, in which the mesonephros remains the adult kidney (Fig. 11). The testicular tubules are connected with specialized mesonephric tubules. An adult male frog displays conditions which are similar to those found in the early bird or mammalian embryo. *See* URINARY SYSTEM; UROGENITAL SYSTEM.

Development of gonads. The development of the gonads is a progressive process, which may be divided into three main phases: (1) appearance of a genital ridge, (2) organization of an undifferentiated gonadal anlage, and (3) sexual differentiation of this primordium.

Genital ridge. The genital ridge appears on the mesolateral side of the mesonephros as a thickening of the coelomic epithelium covering the mesonephros. In reptiles, birds, and mammals it consists of a layer

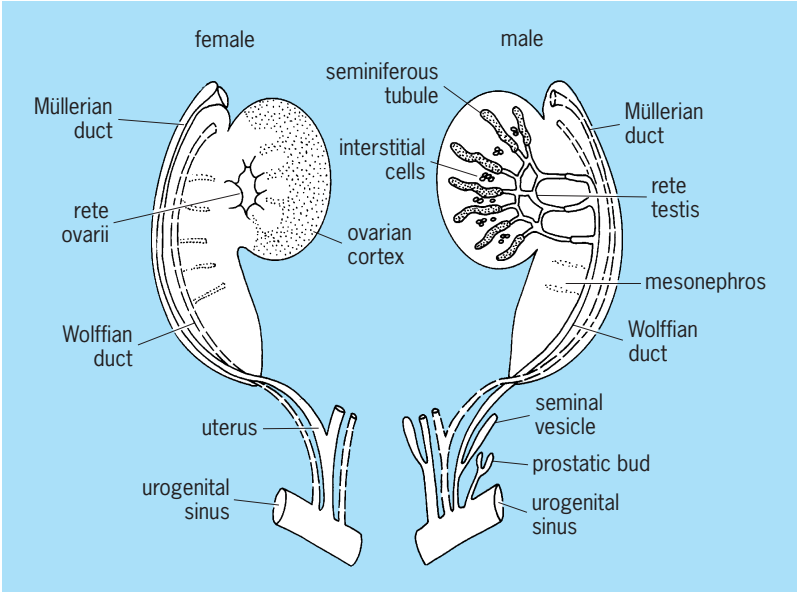

Fig. 10. Homologies in the development of the male and of the female genital tract of mammals. The gonads are shown on the inner side of the mesonephroi, and the ducts on their outer border. (*After A. Jost, General outline about reproductive physiology and its developmental background, in H. Gibian and E. J. Plotz, eds., Mammalian Reproduction, pp. 4–32, Springer, 1970*)

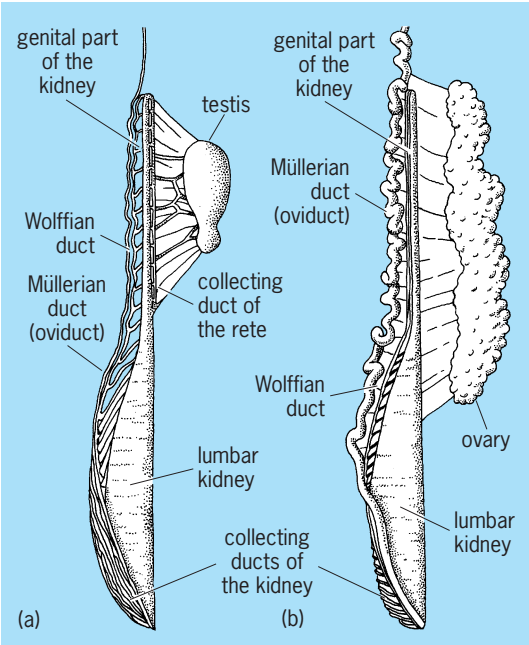

Fig. 11. Drawings of the right half of the reproductive system of the newt *Triton* (*Triturus*) *cristatus*. (*a*) Male. (*b*) Female. (*After J. de Beaumont, Wilhelm Roux Arch. Entwicklungsmech. Organ., 129:120, 1933*)

of enlarged coelomic cells, which at first is two or three cells thick. This contrasts with the flat cells of the other parts of the mesonephric coelomic wall. Primordial germ cells, the cells which will give rise to the germ cells, come in close contact with the so-called germinal epithelium and even penetrate between its cells. These primordial germ cells have an extragonadal origin.

Extragonadal origin of germ cells. In the chick embryo, the primordial germ cells are first seen as large undifferentiated cells, located in the germinal crescent, a part of the extraembryonic area of the blastoderm situated anteriorly to the embryo's future head. They reach the level of the gonadal anlage by way of the bloodstream. In sauropsidians and mammals their extragonadal origin has also been established. They originate in the primitive gut and usually migrate through the dorsal mesentery toward the gonadal area. In mammals, and in human embryos, they can be traced with histochemical techniques, showing their high content of alkaline phosphatase. In the frog the primitive germ cells have been traced from the very beginning of the cleavage of the egg; their early destruction led to the development of sterile gonads.

Undifferentiated gonad. In Amphibia, the primordial germ cells are first disposed in clumps near the inner border of the mesonephros. The genital ridge is formed when these primordial germ cells protrude into the abdominal cavity with some connective tissue and the covering coelomic epithelium. The anlage is invaded by clusters of cells coming from the mesonephric blastema. The gonadal anlage has then attained a stage known as the indifferent stage, which is the same irrespective of the genetic sex of the animal.

The indifferent sex gland consists of the coelomic epithelium at the periphery, the cells of which are flat and scattered and cover the voluminous primitive germ cells. Some connective tissue surrounds the central mass of small cells which came from the mesonephric blastema. The outer zone is often called cortex, and the inner part is called medulla. Both constituents participate in the formation of the definitive sex glands.

In mammals and in birds the undifferentiated primordia of the gonads are also composed of three major components: a superficial epithelium, inner cell masses, and germ cells. The limit between the superficial layer and the inner cells usually disappears and becomes poorly defined. Therefore, the origin and the significance of the constituents of the indifferent gonad gave rise to various interpretations which still prevail. In order to understand the issue, it must be kept in mind that in the mature female animal the ova are released from the ovary through its surface, into the abdominal cavity, in a centrifugal way, and that they are then collected by the tubes. Accordingly, in the development of the ovary the superficial layers play the major part. On the contrary, in males the sperm cells are released into an intratesticular and complex system of channels derived from the mesonephros, in a centripetal way. Accordingly, the seminiferous tubules develop in the center or in the depth of the organ (Fig. 10).

This explains why embryologists more than a century ago introduced the concept that the undifferentiated gonadal primordium was hermaphroditic. It was supposed to comprise a superficial female component and a central male component, the latter originating from the mesonephric renal structures. Only one of these sexually specialized parts was assumed to differentiate in each sex, while the other degenerated. Another long-lived classical theory assumed that the superficial epithelium of the genital ridge proliferated so-called sex cords into the underneath undifferentiated mesenchyme. A first set of these cords was held to constitute the male medullary cords, while in females the superficial epithelium, the cortex, proliferated a second set of sex cords, which developed in an ovarian cortex, after the involution of the medullary cords.

It is still difficult to assess the exact history of the various cell types involved in the development of the sex glands, since no cellular markers are available. In any case, there is no indication of a hermaphroditic constitution of the primordium, since no separate male or female component can be distinguished.

Sexual differentiation of the gonads. Testes differentiate earlier than ovaries, and for a while female embryos can be recognized only because they do not develop testes.

The initial step of testicular differentiation consists of the appearance of a new type of cells in the undifferentiated gonad, the primordial Sertoli cells. These cells aggregate together, and with germ cells they delineate the early testicular seminiferous cords. (In the adult testis, the Sertoli cells support the germ cells in the seminiferous tubules.) In the meantime, at the surface of the testis, the cells become rarefied, flatten, and form the albuginea of the testis. A second crucial step in testicular differentiation is the differentiation of the Leydig cells, between the seminiferous cords. These cells produce the steroid male hormone. *See* HORMONE; STEROID.

At the same time that these events occur in males, the presumptive ovaries in females maintain a morphologically undifferentiated condition. The first clear-cut sign of ovarian differentiation is the onset of the meiotic prophase. In fetal ovaries, the germ cells enter the preparatory phases of chromosomal reduction (or meiosis), a nuclear process which will be completed during adulthood when oocytes are released from the ovary. A long phase of nuclear quiescence takes place between the fetal preliminaries and their completion at the time of ovulation in adulthood. *See* MEIOSIS.

Meiosis seems to be a critical phase of ovarian development. Many germ cells entering meiosis degenerate; a few of them survive and become surrounded by follicular cells. The primary ovarian follicles are constituted in that way. It seems very likely that the follicular cells are homologous to the testicular Sertoli cells. The question as to why the ovarian germ cells enter the meiotic prophase whereas in males of the same age they do not, is not yet

solved, though it has been suggested that a meiosis-inhibiting substance is produced by the testis and a meiosis-inducing substance by the ovary. The reason why so many germ cells degenerate at incipient meiosis also remains unknown.

Gonadal differentiation is well established in reptiles or birds before hatching, or in placental mammals at birth. In the marsupials, such as the opossum, intrauterine pregnancy is very short and is followed by a period of development in the marsupial pouch. Sexual organogenesis takes place after birth during pouch life.

Exceptions to general scheme. Many animals have developmental patterns that deviate from the scheme outlined above. In a great number of birds, only one ovary becomes functional in the adult. In the female chick embryo, for instance, dissymmetry of the gonads is conspicuous at an early stage, because the left anlage is much larger than the right one. The right gonad, composed of some vestigial medullary tubules, remains, and is a nonfunctional rudiment in the hen. It may develop as a small testis if the left ovary is removed.

In toads of either sex, the anterior part of the gonadal anlage acquires the structure of a persistent rudimentary ovary, Bidder's organ, above the functional gonad. In adult animals, Bidder's organ may develop into a functional ovary when the actual gonads are surgically removed. Males deprived of their testes undergo a slow feminization and may lay eggs.

In frogs, there are racial differences. In animals living in cold countries or at high altitude, males and females both differentiate early. In most frogs from temperate climates, the process of sexual differentiation is different. During the months before metamorphosis, in all individuals, whatever their genetic sex, the gonads first differentiate in a feminine way, since testicular differentiation takes place and several germ cells increase in size and enter the meiotic prophase. However, at metamorphosis, in half the individuals the largest germ cells degenerate and testes differentiate, while the other animals follow the normal feminine line. Thus, males suffer from a delayed testicular organogenesis. Animal species or strains which show such a transitory feminine phase of the male gonads are known as indifferent strains. Several amphibians, teleosts, and cyclostomes also develop indifferent strains.

Control of gonadal differentiation. In mice and in humans, evidence that the presence of a Y chromosome is a prerequisite for morphological differentiation of testes has accumulated. Thus, exceptional XO mice are females and XXY mice are males. The problem of how the sex genes control gonadal organogenesis has not yet been fully explained on cellular or biochemical bases. *See* SEX DETERMINATION.

The effects of administering hormones to developing embryos are profound, but vary among organisms (**Table 1**). Such experiments do not, however, warrant the generalization that sex hormones control sex differentiation. In lower vertebrates, sex differentiation may be influenced by external factors. Experiments made on turtles, alligators, and the newt *Pleurodeles* showed that breeding the eggs at different temperatures resulted in the development of young which were all of the same sex, depending on the temperature. Genetically, male *Pleurodeles* which were sex-reversed into functional females by breeding at high temperature could be mated with normal males and the progeny could be studied.

TABLE 1. Effect of administered sex hormones on developing gonads

Animal	Estradiol (on males)	Testosterone (on females)
Fishes		
Lebistes	Feminization	Masculinization
Amphibians		
Frogs, various species	Feminization (low dosage) Masculinization (high dosage), paradox effect	Masculinization
Alytes	Feminization	No effect
Xenopus	Complete feminization	No effect
Bombina	No effect	No effect
Amblystoma	Complete feminization	Feminization, paradox effect
Pleurodeles	Complete feminization	Feminization, paradox effect
Birds		
Chick	Feminization	Slight masculinization
Duck	Slight feminization	No effect
Marsupials		
Opossum	Feminization	No effect
Placental mammals		
Several species	No effect	No effect

Sex differentiation of the gonads is controlled by genetic factors and the question arises of how and in which cells these genetic factors express themselves. As mentioned above, the gonadal primordium has long been assumed to be hermaphroditic, and it was accepted that the genetic factors gave prevalence to one sex component over the other, possibly via the dominance of one of the two antagonistic sex inductors produced by the sexually opposed gonadal components. In 1965, Jost introduced the concept that the sex glands were programmed to develop according to the homogametic sex (female in mammals, male in birds) and that in the heterogametic sex a special triggering mechanism imposes alternate sexuality on the primordium. This scheme was largely accepted, but the trigger has yet to be identified. A theory propounded in 1975 by Stephen Wachtel and others had some success. It proposed that the HY-histocompatibility antigen (discovered by grafting male skin in female mice and characteristic of individuals possessing the Y chromosome) was a triggering protein responsible for the differentiation of the sex gland in the heterogametic sex. This concept does not, however, fit with Humphrey's fundamental experiments mentioned above or with other observations. The chemical control of gonadal sex differentiation remains to be definitively elucidated.

Gonadal abnormalities. Important abnormalities include absence of the gonad, hermaphroditism, and sex reversal.

Absence of the gonad. The entire gonad may fail to develop or may retrogress at very early stages, leaving only some indistinct remnants which are impossible to recognize as testes or as ovaries. This condition, called gonadal agenesis or gonadal dysgenesis, is known among humans and other animals. In humans, gonadal dysgenesis is a part of the polymorphic Turner syndrome, which occurs in patients having the abnormal chromosomic formula XO. Since it also occurs in patients who have other normal or abnormal chromosomal formulas, the exact correlation between the sex chromosomes and gonadal dysgenesis is not yet clear.

Hermaphroditism. The presence of male and female gonadal tissue in the same individual is not normal among vertebrates except in some species of fishes. It can occur as an abnormal condition more or less frequently among other vertebrates. It is rather frequent in some frogs and toads, and was considered to be the result of an incomplete dominance of either the cortex or the medulla of the indifferent gonad. It might also, however, result from incomplete expression of the testicular determining system. In humans, several cases of true hermaphroditism have been reported in sterile individuals. Several cases have been correlated with chromosomal abnormalities such as XX/XY mosaicism, and in one case it was observed that the ovarian tissue contained mainly XX cells and the testicular tissue XY cells.

Genital tract development. The sex ducts become sexually specialized some time after the sexual differentiation of the sex glands. The male or female conditions develop from an indifferent condition which is identical in both sexes in early stages.

Indifferent stage. The gonads are already recognizable as presumptive ovaries or testes and are located on the anterior part of the mesonephros. The mesonephric or Wolffian duct is the ureter, and opens posteriorly into the cloaca in lower vertebrates or into the urogenital sinus in mammals. Another duct, the oviduct or Müllerian duct, parallels the mesonephric duct. The oviduct arises from a funnel which opens into the coelomic cavity. The blind end of this primordium proliferates and extends progressively caudally.

In selachians and urodeles, the funnel from which the Müllerian duct originates corresponds to a pronephric nephrostome, the coelomic opening of the primitive urinary tubules. Because the pronephros is located near the neck of the larva, and because the ostium of the oviduct retains this position, the oviducts open into an anterior part of the body cavity (Fig. 11).

In birds and mammals the origin of the oviduct from pronephric remnants is not as clear, but it is obvious that the oviduct develops in the region of the nephric field. The early funnel is located on the top of the mesonephros, and the ostia tubae open above the ovaries.

Differentiation of female genital tract. The Müllerian ducts differentiate into the female ducts. Depending upon the animal species, either a simple secretory oviduct (amphibians) or a more complicated structure develops from this simple unicellular layered duct. It is divided into several specialized sections involved in the secretion of albumen or shell, as in selachians, reptiles, and birds. In those birds in which the female has only one functional ovary, only one oviduct develops (**Fig. 12**).

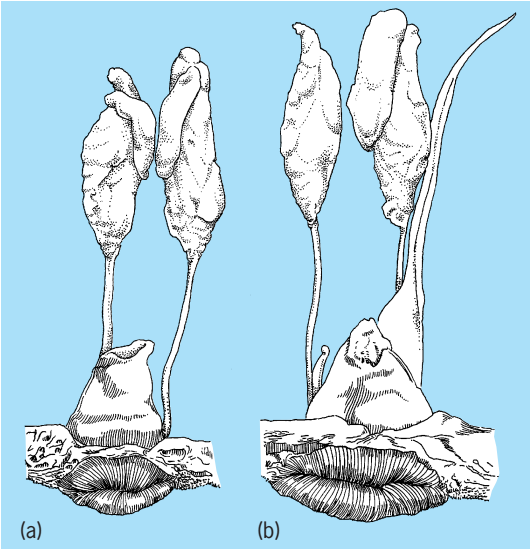

Fig. 12. Reproductive system of normal chick embryos, shortly before hatching. (*a*) Male (two testes present). (*b*) Female (sole left ovary and left oviduct). (*After E. Allen, ed., Sex and Internal Secretions, 3d ed., Williams and Wilkins, 1961*)

In mammals, the Müllerian ducts give rise to the oviduct, or tube, and the uterus. Usually the embryonic Müllerian ducts fuse posteriorly, but the extent of fusion is variable. In humans and monkeys a single uterus is formed from the fused part of the ducts; the tubes correspond to the nonfused part. In rodents, the two Müllerian ducts fuse only in the upper vagina, and form two uterine horns and two tubes. In ruminants, an intermediary condition is realized in which the uterus is composed of an inferior stem and two horns which largely communicate (Fig. 3).

The posterior part of the mammalian female genital tract is constituted by the vagina. The embryology of this organ displays great variability from one animal species to the other, and this makes an accurate interpretation difficult.

In the undifferentiated stage, the Müllerian ducts terminate blindly at the wall of the urogenital sinus (or embryonic urethra), between the two Wolffian ducts which open into the urogenital sinus. They may retain this primitive connection, as is the case in the rabbit (**Fig. 13***a*), and the urogenital sinus then becomes a urethrovaginal duct.

In other animals a cord of cells, the vaginal cord, may detach from the dorsal wall of the urogenital sinus and grow progressively caudally below the end of the Müllerian parts (Fig 13*b*). In the rat or the mouse this sinusary vagina finally opens independently of the definitive female urethra in a vaginal opening (Fig. 13*c*). In other animals, such as the mare, the vaginal cord remains connected with the posterior part of the urogenital sinus. The

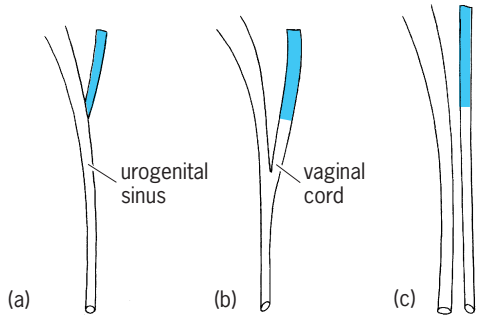

Fig. 13. Schematic interpretation of vaginal organogenesis in three species. Müllerian ducts are shaded. The three conditions are explained in the text.

vestibulum then remains as a common opening to both the vagina and urethra.

Comparative embryology of the vagina helps in understanding abnormalities of the human genital tract. Absence of the vagina and opening of the female ducts into the urethra may result from the preservation of the primitive connections.

The Wolffian or mesonephric duct of the female embryo is never incorporated into the genital tract. In animals that keep their mesonephros as a functional kidney in adulthood (selachians and amphibians), the mesonephric duct remains as the excretory channel (Fig. 11). In female birds and mammals, the mesonephric duct, as well as the mesonephros, disappears, leaving only some minor vestiges.

Differentiation of male genital tract. The male sex ducts are derived from the mesonephric or Wolffian ducts. In selachians and amphibians the mesonephros remains the functional kidney, and the mesonephric ducts function as pathways for urine as well as for sperm. Even in such animals, however, the anterior part of the kidney often becomes specialized as the sexual part. The posterior part of the kidney then produces urine, and several excretory tubes may bring this urine directly to the cloaca. In such cases the Wolffian duct of the male is only a genital canal, as in the newt *Triton cristatus* (Fig. 11).

The Wolffian ducts of birds differentiate into an undulated vas deferens, whereas in mammals the ducts differentiate into the epididymis at one end and the seminal vesicles at the other (Fig. 10).

The Müllerian duct has no function in males, and as a rule it disappears. In some amphibians, such as newts and toads, the oviducts persist in a rudimentary condition (Fig. 11); they may be activated under appropriate hormonal stimulation in adult males.

In male mammals the urogenital sinus becomes the definitive male urethra. Several accessory glands, such as the prostatic glands, bud from it, and display great variations from one animal species to another, but all open into the urethra (Fig. 10).

Copulatory organs. Copulatory organs are well developed in mammals and reptiles. In selachians they are often a specialized part of the fins, the claspers. As a rule no copulatory organ is present in amphibians and birds; only a few male birds, such as the duck, possess a penis.

In mammals the copulatory organ develops from an undifferentiated genital tubercle which is identical in both the male and female embryos and lies above the opening of the urogenital sinus. The male penis encloses the penile urethra and increases in size, whereas the homologous female clitoris remains more like the primitive tubercle.

Hormonal control. The genital tract differentiates after sexual differentiation of the sex glands, and time relationships support the view that sexual specialization of the genital tract is controlled by hormones produced by the developing sex glands. This has been shown experimentally by depriving embryos or young animals of their gonads. It was established that in the gonadless body, the genital tract becomes identical whatever the genetic sex of the individual. This means that in the absence of the sex glands, sexual dimorphism, which is a characteristic feature of almost all vertebrate species, does not appear. This identical, hormoneless aspect of the body is known as the neutral form. However, the gonadless sexual type is not the same in mammals and birds.

In mammals, gonadless sexual organization has been studied in rabbit fetuses which were surgically castrated in utero. It was found that the neutral condition is essentially feminine. This means that no embryonic gonad is necessary to produce the feminine sexual structures. In males, the testes prevent persistence and development of the female structures (tube and uterus) and impose masculinity on the whole genital apparatus (**Fig. 14**).

In a duck embryo castrated by a beam of x-rays, E. Wolff noticed masculine differentiation of such sex characters as the penis or the voice organ (syrinx). During normal sexual differentiation the embryonic ovary prevents such characters from becoming masculine. In addition, the neutral aspect of the

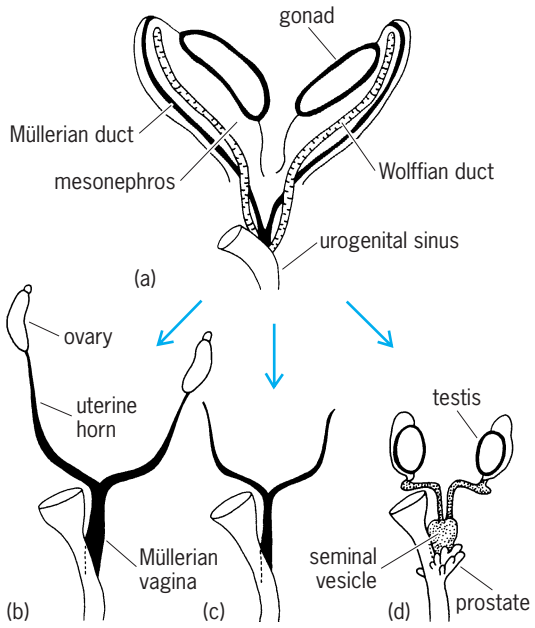

Fig. 14. Sexual differentiation of the genital tract of the rabbit fetus: (*a*) undifferentiated condition; (*b*) female; (*c*) castrate; (*d*) male. (*After A. Jost, Mem. Soc. Endocrinol., no. 7, 1960*)

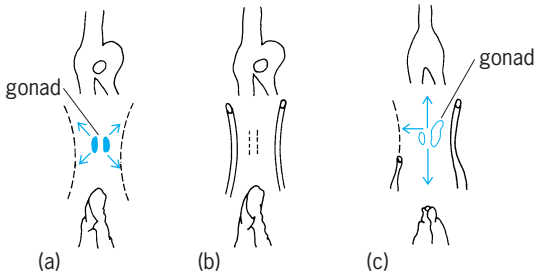

Fig. 15. Evolution of sex characters in the duck embryo: (*a*) male, (*b*) castrate, (*c*) female. Top sketch, voice organ or syrinx, well developed in males. Middle sketch, Müllerian duct (compare with Fig. 13); arrows symbolize inhibiting gonadal actions. Bottom sketch, genital tubercle. (*After E. Wolff, Compt. Rend., 229:428, 1949*)

genital tract is also characterized by the presence of oviducts. These structures are normally absent in males in which they are inhibited by an embryonic testicular hormone (**Fig. 15**). A similar trend toward maleness was observed in young newts, *Triturus cristatus*, castrated before sex differentiation. In that species females are heterozygous, as in birds.

It appears that in some vertebrates the neutral type is feminine (mammals) and in others predominantly masculine (birds and newts). In the absence of gonadal hormones, a basic developmental program prevails that corresponds to the homozygous sex, female in mammals, male in birds and newts. In the heterozygous sex, the gonads counteract this program and impose respectively maleness or femaleness in mammals and in birds and newts.

As mentioned above, in some amphibians complete sexual reversal of the gonads can be obtained under hormonal influences (testes developing in genetic females or ovaries in genetic males). In such animals the genital tract develops in agreement with the actual sexual structure, and function of these gonads and fertility are ensured.

Such a complete sexual reversion has not yet been produced in higher vertebrates. The condition prevailing for placental mammals deserves special attention. Sex hormones administered to the pregnant mother animal or injected into the fetus itself modify the development of the genital tract with the exception of the gonads themselves. The young thus become intersexes and definitely abnormal. Estrogenic hormones given to pregnant rats or mice feminize the male fetuses to a large extent. In addition to unaltered testes, large parts of the Müllerian ducts can be retained and a vagina and feminine external genitalia develop. Under the influence of androgenic hormones, female fetuses acquire several masculine structures, such as male external genitalia, urethra, accessory sex glands, vasa deferentia, and seminal vesicles. But they retain their ovaries, tubes, and uterus. Similarly, in castrated rabbit fetuses given testosterone or similar hormones, masculine organs develop but the Müllerian ducts are not suppressed. Androgenic steroid hormones do not inhibit the Müllerian ducts. This suggests that the fetal testis produces two kinds of substances, one stimulating the development of masculine characters, another inhibiting the female ducts. The dual nature of the hormonal control of the differentiation of the genital tract is now established. In addition to steroidal androgens such as testosterone, the fetal testis produces a Müllerian inhibitor called anti-Müllerian hormone or Müllerian-inhibiting substance. This hormone, as produced by the fetal calf testis in cell cultures, was shown to be a glycoprotein of a molecular weight of 124,000. The two fetal testicular hormones are produced by two different cell types, the Sertoli cells for the anti-Müllerian factor and the Leydig cells for androgens.

Another line of important facts concerns the target cells for androgens. It is established that androgens can masculinize undifferentiated primordia only if the cells are provided with appropriate androgen receptors. Moreover, in several target organs testosterone has to be converted into dihydrotestosterone by the intracellular enzyme 5α-reductase, before it acts in the nucleus. Absence of receptors or of proper conversion of testosterone results in the same effect as absence of hormone. A partial cellular defect produces an intermediary degree of masculinization.

Abnormalities of genital tract. Many types of genital developmental abnormalities are known in animals and humans. Genetically female fetuses may be partially virilized (masculine or ambiguous intermediary aspect of the external genitalia) under the influence of androgenic substances either given to their pregnant mother or produced by their own abnormal adrenals (adrenogenital syndrome).

Genetically male fetuses may suffer from incomplete masculinization by their own testes. The discovery of a dual hormonal control of the differentiation of the genital tract and of the necessity of hormonal receptors in target tissues gave explanations for many sex abnormalities. Males can be incompletely masculinized in the following cases:

1. The testes are absent. In this case the genital tract is completely feminine, but remains infantile because no ovarian hormones are produced. This condition resembles that of castrated rabbit fetuses.

2. The testes do not produce testosterone, or not enough. Several enzymes are involved in the synthesis of testosterone from acetate or cholesterol, and several enzymatic defects depending on discrete genes are known. This results in a condition of pseudohermaphroditism.

3. The target cells lack androgen receptors. Since androgens are not effective, no male sex duct (vas deferens), and no male organs are present, except endocrinologically normal testes. Since the Müllerian inhibitor is produced, the tubes and uterus do not persist. This condition prevails in the syndrome called testicular feminization in humans or in animals.

4. The target organs lack 5α-reductase, a defect that does not permit appropriate masculinization (such as at the level of external genitalia). This is another type of pseudohermaphroditism.

5. The testes fail to produce the Müllerian inhibitor (or possibly the Müllerian ducts lack

receptors). In these subjects testes and male sexual characters are well developed, but a set of female organs persists (persistent Müllerian ducts).

6. In true hermaphrodites, when testicular tissue is present only on one side of the body, masculinization of the internal genital structures may be limited to the same side.

In conclusion, multiple causes of genital abnormalities are known. In females, they usually result from androgens that should not be present. In males, partial or total lack of masculinization may be due either to lack of androgen secretion or action, or to lack of Müllerian inhibitor secretion or action, or to both.

Alfred Jost

Physiology

The physiological process by which a living being gives rise to another of its kind is considered one of the outstanding characteristics of plants and animals. It is one of the two great drives of all animals: self-preservation and racial perpetuation. In contrast to other physiological processes, reproduction in vertebrates can be achieved only by participation of two individuals, the male and the female. Each produces germ cells called gametes. The male produces spermatozoa and the female ova, which carry the biochemical information arranged as genes in the chromosomes for the transmission of inherited characters. No matter how discrepant the pairing gametes may be in size and in form within a species, they contribute the same number of chromosomes. Although reproductive devices are quite different from one species to another, all serve one end, the bringing together of the spermatozoon and the ovum, each containing half the number of chromosomes of the parent cells. After the union of the chromosomes at fertilization, the newly constituted embryo possesses the number of chromosomes of its race and then divides, differentiates, grows and develops into an individual either outside the mother (oviparity) or inside the mother (viviparity). The obvious advantage in bisexual reproduction lies in the fact that young produced from the mingled genes of two ancestral lines will not be a direct copy of either parent, but will represent different combinations of ancestral traits with great potential for the survival of the fittest. Reproduction is facilitated or inhibited by environmental and nutritional factors as well as by all the other physiological activities, but it is controlled predominantly by the endocrine system and mediated at least in part by the nervous system.

Breeding season. The season when animals perform their overt reproductive functions is known as the breeding season. This sexual periodicity is a general phenomenon common to plants and animals. In general, sexual periodicity is predominant in lower vertebrates, but all birds and wild mammals are seasonal breeders. As nutritional and environmental conditions improve, the reproductive season is not so restricted; cattle, the domestic rabbit, and humans all illustrate this fact; nevertheless their fertility is higher in spring than in winter. Males of many mammalian species are capable of sperm production and copulation at any time and rarely experience a true sexual period such as rut in the deer. In mice and rats, the testes do not descend to the scrotum until puberty and in certain other rodents they descend only during the breeding season. Most vertebrates breed only in spring or summer at the time when food, temperature, and light are optimal; they are then in the best physiological condition for reproduction. Others, for example, sheep, breed at a time which allows parturition to occur at the season of year most suitable for successful survival of the young. *See* REPRODUCTIVE BEHAVIOR.

All environmental factors play a part, but an increase of temperature for lower vertebrates and an increase or decrease of daylight for mammals play a major role in determining the onset of their breeding seasons. For instance, most fish and amphibians breed when temperature increases. The mink and horse breed at the time of increase of daylight, in the spring, and their offspring are born in the summer or the spring, respectively, of the following year. The armadillo, deer, goat, and sheep breed during the decrease of daylight in the autumn and their offspring are born the next spring. Transportation of sheep from the Northern to the Southern Hemisphere changes their breeding season to accord with the seasons of the new environment. The testes of the cottontail rabbit return to a completely immature condition in the fall and their germ cells are all of the spermatogonial type. In late November, the testes begin to grow; in early spring they reach 50–100 times their weight in the inactive stage and the growth of accessory glands follows that of the testes. Although artificial increase of daylight brings wild rabbits and ferrets into breeding condition in midwinter, permanent short-day lighting does not entirely prevent breeding in the spring. Furthermore, artificial increase of daylight has so far failed to show any definite effect in advancing the breeding season of the domestic rabbit, guinea pig, hedgehog, or ground squirrel.

Among fishes the duration of the breeding season varies considerably according to the group to which they belong and where they live. The ova of the elasmobranchs are deposited singly or in pairs at varying intervals throughout a great part of the year, but some species appear to have regular recurrent breeding seasons. Most bony fishes, producing millions of ova, breed only in the spring, and most of them migrate to a suitable locality for the deposition of their gametes. In the frog and many other amphibians the ova are produced during the winter hibernation when the animals eat very little. Similarly, the genital organs of salmon develop during migration, when the fishes cease to feed. The spermatogenesis of the amphibian *Rana temporaria* is determined by the gonadotrophic activity of the pituitary and by the sensitivity of the germinal epithelium to gonadotrophin during the spring; this activity and sensitivity are influenced by temperature rather than light. In the autumn and winter their testes are not sensitive to gonadotrophin under natural conditions. Reptiles that hibernate usually begin to breed shortly

TABLE 2. Average length of the ovulatory cycle of some familiar mammals

Animal	Length
Opossum	28 days
Guinea pig	16 days
Rat	5 days
Mouse	5 days
Sheep	16 or 17 days
Pig	20–22 days
Cow	21 days
Horse	21 days
Macaque monkey	24–26 days
Human	28 days
Chimpanzee	37 days

after the beginning of the warm weather that terminates the hibernating period. Spring and summer are the seasons when most birds pair, build their nests, and incubate their eggs. The migration of some birds is invariably associated with an increase in the size of ovaries and testes.

Estrous and menstrual cycles. The cyclic changes of reproductive activities in mammalian females are known as estrous or menstrual cycles (**Table 2**). Most mammalian females accept males only at estrus (heat). The follicles in the ovaries grow at a rather constant rate before estrus. Just before or at estrus the follicles grow rapidly, resulting in a higher output of estrogen, a female hormone. During or toward the end of estrus the follicle, or group of follicles in animals producing litters, ruptures and releases the ovum (ovulation). The follicle is transformed into the corpus luteum, a glandular tissue which produces another female hormone, progesterone. If fertilization occurs, the corpus luteum persists during all or most of the gestation period according to the species. Without mating or fertilization the corpus luteum persists for a relatively short time and then degenerates. The growth of the next crop of follicles begins again.

Estrus. Estrus in mammals can occur several times in one breeding season; the mare, ewe, and rat come to estrus every 21, 16, and 5 days respectively if breeding does not take place. This condition is called polyestrus. The bitch is monestrous; she has only one heat, or estrus, to the breeding season and if not served then, she does not come into heat again for a prolonged interval, 4–6 months according to different breeds. The duration of estrus varies according to the species, such as 4–9 days in the mare, 16–18 h in the cow, and 14 h in the rat. In the ferret, estrus is shown by predominant swelling of the vulva and lasts several weeks unless copulation takes place.

In monestrous and seasonally polyestrous species the period of sexual quiescence between seasons is called anestrus. In sheep and deer this occurs in the summer, while in the ferret and mink it is restricted to the winter. Proestrus denotes a short period just before the time of coming in heat. It is very marked in the bitch; the genital organs swell and become congested, and there is bleeding from the uterus. Following estrus, the ensuing period in which sexual activity declines is called metestrus. After metestrus a relatively longer period of diestrus occurs and occupies the bulk of the estrous cycle before proestrus recurs. Most females ovulate spontaneously, and shed their ova close to the end of estrus: about 2–4 days before the end of estrus in the mare, 13 h after the end of estrus in the cow, 10 h after the beginning of estrus in the rat. Some species, such as the rabbit, ferret, and cat, ovulate following the stimulation of copulation. If fertilization takes place, pregnancy occurs. If mating is unsuccessful, pseudopregnancy, a physiological condition very similar to pregnancy but which lasts a relatively short period, may occur in some species, particularly the dog and the rabbit. *See* ESTRUS.

Menstruation. The reproductive cycle of the female in the primate and human is well marked by menstruation, the period of blood flow from the vagina. Menstruation does not correspond to estrus but occurs between the periods of ovulation at the time the corpus luteum declines precipitously. As a result of the collapse of the superficial capillaries in the endometrium, blood is extravasated in quantity and about two-thirds of the endometrium is desquamated, or shed. After menstruation there is a period in which the uterus is replenished under the influence of estrogens by the growth of the epithelium and its capillaries. This is known as the proliferative stage and is equivalent to proestrus and estrus in other mammals. It is accompanied by growth of the vaginal epithelium and some cornification, but these changes are gradual and are not as well defined as in many other mammals. The Graafian follicle ruptures near the midcycle, about 15 days before the beginning of the next menstruation. Ovulation is spontaneous and is followed by the formation of the corpus luteum, whose secretion of progesterone causes great glandular growth of the endometrium. This is known as the luteal or secretory phase and is equivalent to diestrus in other mammals. In New World monkeys, menstruation is confined to the appearance in the vaginal lavage of a few red blood cells. Tissue destruction and hemorrhage in the uterus of the elephant shrew *Elephantulus* of South Africa at the end of diestrus has been described, but the bleeding at proestrus in the bitch, at metestrus in the heifer 14 h after ovulation, and at metestrus occasionally in the guinea pig is due to intensive congestion of blood vessels in the uterus or in the upper part of the vagina and is not true menstruation.

Mating. Mating, also called copulation or coitus, is the synchronized bodily activity of the two sexes which enables them to deposit their gametes in close contact. It is essential for successful fertilization because sperm and ovum have a very limited life-span; as examples, the fertilizing capacity of trout and salmon sperm lasts about 30 s in fresh water, that of rabbit sperm about 30 h in the female tract. Trout and frog ova become nonfertilizable soon after contact with water, whereas rabbit ova remain fertilizable for 6–8 h after ovulation. Human spermatozoa have a fertilizing life-span of about 72 h, and ova a fertilizable life of about 24 h. Mating behavior is a very

common phenomenon in the animal kingdom. Even in the primitive vertebrate the lamprey, some sort of mating takes place. The female is seized by the male, who winds his tail around her and discharges his sperm over the ova as they extrude from her body. In the dogfish, which practices internal fertilization, contact between male and female is necessary for the claspers of the male to convey his sperm into the cloaca of the female. In the majority of the bony fish physical contact may be slight, yet the male follows the female closely or presses his body to hers and deposits his sperm over the ova either soon after they are laid or at the same time. Certain salamanders walk slowly ahead of the female and deposit spermatophores which the female straddles and secures with the lips of her cloaca. In other low vertebrates a rudimentary penis exists either as a single intracloacal organ (crocodile, turtle, and duck) or as a paired structure placed behind the cloaca (snake and lizard) and sperm are deposited into the female's cloaca at mating. In mammals the copulatory organs are well developed, varying a great deal in structure. Patterns of mating are different; the time required for mating varies from 3 h (ferret) to a few seconds (sheep and rabbit), and the number of intromissions varies from 1 (bull) to 50 (hamster).

Mating in many birds and cold-blooded vertebrates is the culminating event of a more or less complicated series of love antics. It is quite probable that these sexual displays have a direct bearing upon ovulation and that in many species ovulation would not occur in the absence of such sexual experiences. Sexual excitement and violent physical exertions of both partners may also be essential for a fertile mating.

Ovulation. Ovulation can be considered as the culmination of a series of events in the process of oogenesis, which begins in the embryo and continues into adult life. Oogenesis is the formation, growth, and maturation of the oocyte within the ovary. The time of first ovulation at puberty varies among species, occurring as early as 30 days after birth in mice and as late as 11+ years in humans. The stages in the development of the ovum are, however, similar in most mammals. By birth, most of the primordial germ cells have become oogonia anad developed into primary oocytes. Most of these oocytes remain in this stage until, under the stimulation of gonadotropic hormones (especially follicle-stimulating hormone), one or more (depending on the species) follicles containing the primary oocytes begin to enlarge and differentiate into Graafian follicles. During this process, the primary oocyte undergoes a first meiotic division, in which the chromosome number is halved from the diploid to haploid state with the extrusion of the unwanted chromosomal material in the first polar body via unequal cytoplasmic division, and commences the second meiotic division. This secondary oocyte is released when the follicle is stimulated to rupture under the influence of the gonadotropin, known as luteinizing hormone. Ovulation has been likened to an inflammatory process, and it is thought that certain mediators of inflammation, such as prostaglandin $F_{2\alpha}$ 12-HETE, may be involved in the final rupture of the follicle wall, thus permitting release of the ovum. The ovum remains arrested at metaphase II of the second meiotic division until sperm activation. It is surrounded at ovulation by two cellular layers, the bulky and sticky cumulus oophorus and the corona radiata, the highly adherent radially arranged three-deep layer of cells with connective processes joined to the vitelline membrane. *See* ENDOCRINE MECHANISMS.

Sperm transport in mammals. The logistics of sperm transport to the site of fertilization in the oviduct present many interesting features in mammals, but it is important to distinguish between passive transport of sperm cells in the female genital tract and sperm migration, which clearly attributes significance to the intrinsic motility of the cell. The rapidity (a few minutes in most mammals) with which spermatozoa reach the upper portions of the oviduct after mating or artificial insemination has emphasized that sperm motility alone could not have made a major contribution to passage over the relatively long distances involved. It is generally accepted that the role of sperm motility becomes critical only in certain regions of the tract such as the cervix and uterotubal junction and during penetration of the egg membranes. The journey for the spermatozoa is largely dependent upon contractions in the smooth-muscle layers of the tract, which may be enhanced by the action of oxytocin released after coital stimulation. Smooth-muscle stimulants in the semen, such as prostaglandins of the E series, may also contribute to the process of sperm transport in species where semen is deposited in the anterior vagina (for example, humans and sheep). Nonetheless, in animals in which the duration of estrus is long and mating takes place early in the receptive period, migration of spermatozoa from storage regions such as the lower oviduct must be of physiological significance, especially since the excited state of the uterus and oviducts promoted by coitally released oxytocin would long since have ceased.

In comparison with the number of spermatozoa deposited in the female at mating, there is a massive reduction in sperm numbers toward the upper portions of the oviduct. This sperm gradient reduces the chances of polyspermy by restricting the numbers of spermatozoa that achieve the site of fertilization. During this marked diminution in sperm numbers the cells are progressively removed from the seminal plasma, and they become strongly influenced by the uterine and tubal fluids. This situation increases the metabolic activity of ejaculated spermatozoa, and may also be directly associated with development of the capacitated state, discussed below.

Explanations for the fact that spermatozoa "ascend" the oviducts while the eggs are "descending" arise in part from viewing the isthmus as a compartment functioning in a manner different from the ampulla, a view that is emphasized by the results of intraperitoneal insemination. However, viable spermatozoa are actively motile, and although myometrial contractions play a major role in sperm transport through the uterus, progressive motility does

contribute to migration into and within the oviducts. Waves of peristaltic contraction in the isthmus proceed principally toward the ampullary-isthmic junction, and this activity should therefore assist transport of sperm cells. These waves of contraction are propagated through the circular muscle, and may be sufficiently powerful to occlude the lumen and, for example, to transport droplets of oil or Indian ink upward to the peritoneal cavity, but the contractions tend to fade beyond the ampullary-isthmic junction. Whether the relatively high concentration of 19-hydroxyprostaglandin E_1 or other prostaglandins of the E series in primate semen benefit sperm transport by means of some alteration in the myosalpinx remains to be clarified.

Cilia in the rabbit and pig isthmus may beat actively toward the ampullary-isthmic junction at the time of ovulation, thereby creating local currents of fluid. However, such ciliary currents have not been detected in the isthmus of cat, cow, sheep, monkey, or human.

The rate and efficiency of transport of spermatozoa are influenced by the timing of coitus relative to the moment of ovulation. Transport of cells as far as the tubal isthmus usually follows semen deposition shortly before ovulation, but it is uncertain whether sufficient viable spermatozoa would have reached the ampulla before the actual time of ovulation to be immediately available for penetration of the freshly ovulated egg or eggs. Spermatozoa may be largely sequestered in the isthmus in the preovulatory interval, only to be released for further proovarian transport when ovulation has occurred and the products of ovulation have entered the tube. The inference here is that there may be some synchrony between release of the egg from the Graafian follicle and a redistribution of spermatozoa from the constricted lumen of the isthmus.

Even though a specific attractant substance for spermatozoa has not yet been demonstrated to be released from mammalian eggs or their investments, some form of chemotaxis may contribute to the final phase of sperm transport and orientation toward the egg surface. Alternatively, maturation of the spermatozoa within the oviduct may induce intrinsic motility, which is then responsible for further progression of the spermatozoa along the oviduct to the site of fertilization.

Capacitation of spermatozoa. Although in most mammalian species the oocyte is shed from the Graafian follicle in a condition suitable for fertilization, ejaculated spermatozoa must undergo some form of physiological change in the female reproductive tract before they can penetrate the egg membranes. The interval required for this change varies from about 1.5 to 6 h according to species, and the process is referred to as capacitation. The precise changes that constitute capacitation remain unknown, although there is strong evidence that they are—at least in part—membrane-associated phenomena, particularly in the region of the sperm head, that permit release of lytic acrosomal enzymes with which the spermatozoon gains access to the vitelline surface of the egg. However, increased respiratory and flagellar activity are also noted after capacitation, the changes in motility being presumed to facilitate sperm penetration of egg membranes.

Capacitation is now understood to precede and permit a coordinated acrosome reaction, thus being an essential preliminary to activation or release of inner acrosomal enzymes and penetration of the zona pellucida; it may involve some adjustment in the distance separating the plasma and outer acrosomal membranes to permit the vesiculation reaction to proceed.

An understanding of the posttesticular changes occurring in spermatozoa in the male reproductive tract has provided some insight into the possible nature of capacitation. During passage along the epididymal duct, spermatozoa undergo a series of modifications that confer increased stability on the head, including a compaction of the nuclear chromatin and a tighter application of the plasma and acrosomal membranes to the sperm nucleus. Secretions within the epididymal duct also influence the maturing spermatozoa, for the antigenic nature of the sperm surface changes during the 1 to 2 weeks spent in epididymal transit. There is good evidence that enzyme inhibitors are "added" to the spermatozoa while still in the male, although whether the principal phase of addition occurs in the epididymis, during the preejaculatory mixing with whole seminal plasma, or in the initial passage of spermatozoa within seminiferous and rete testis fluid remains to be clarified. Neutralization or removal of inhibitors of acrosomal enzymes in the female tract would seem a logical stage in the capacitation process, or an immediate sequel to it, and could explain how the enzymes of fundamental importance in fertilization are stabilized in the sperm head until spermatozoa are within reach of the eggs.

Although capacitation seemingly involves both membranous and metabolic changes in the sperm cell, evidence suggests that modification in the glycoprotein and other moieties of the plasma membrane is a critical step. For example, as uterine incubation of ejaculated rabbit spermatozoa proceeds, the ability of the cell surface to bind lectins disappears progressively from the tip of the sperm head, suggesting that carbohydrate-containing moieties are being removed progressively from the plasma membrane. This alteration in the cell surface would be expected to prepare it for fusion with the outer acrosomal membrane, possibly by reducing net negative charges or by increasing membrane fluidity. As a sequel to the acrosome reaction, at least in hamster spermatozoa, there is a modification in the postnuclear cap which facilitates sperm head incorporation by the egg plasmalemma.

Although conventional descriptions of the acrosome reaction emphasize that membrane fusion occurs as a means of releasing proteolytic enzymes critical for sperm penetration of the egg investments, there is evidence that hyaluronidase may escape from viable sperm cells before membrane vesiculation can be detected in the electron microscope.

A second point concerning acrosomal enzymes is that acrosin—the putative zona lysin—is apparently packaged in zymogen form and needs to be activated from proacrosin during the capacitation process.

There is evidence from at least three mammalian species (rabbit, hamster, pig) that different compartments of the female tract act synergistically to promote capacitation, and that sequential exposure to the uterus and then oviduct accelerates the process of capacitation. A specific action of the oviducal environment is therefore suggested and, in the golden hamster, the effects on stimulated respiration and achievement of the capacitated state are so distinct that they can be seen through the thin wall of the ampulla in terms of a massive increase in sperm motility; even though hamster sperm reach the oviducts in under 1 h, another 3 h are required for this dramatic change to be accomplished. Precise evidence for the capacitation of primate spermatozoa is not available; if it occurs it must be rapid, since human ejaculated spermatozoa after washing, but without any exposure to female reproductive tract fluids, can penetrate matured oocytes in culture within 4 h.

The evolutionary significance of capacitation can be viewed in two ways. Physiologically, it can be regarded as a process whereby the sperm cell becomes competent to undergo an orderly sequence of membrane changes, collectively termed the acrosome reaction, to permit release of the lytic enzymes. Biologically, capacitation probably evolved to meet requirements associated with internal fertilization in mammals, and might best be regarded as an effective means of gamete selection during sperm transport through the different regions of the female reproductive tract; only spermatozoa at the optimal stage of their maturation process are able to penetrate the egg.

Fertilization. Fertilization is a process that begins with penetration of the ovum surface layer (zona pellucida) and binding to the vitelline membrane by the fully matured and acrosome-reacted spermatozoon, followed by incorporation of the spermatozoon into the ovum cytoplasm and activation of the ovum. The haploid chromosome complements of the sperm nucleus and ovum are then transformed into male and female pronuclei, respectively. The pronuclei migrate to the center of the ovum, permitting the joining of the two sets of chromosomes, or syngamy, which is the final stage of the fertilization process. The ovum, now known as a zygote, begins to develop by cell division, or mitosis. Normally, after penetration of the vitelline membrane by one spermatozoon, the fertilizing sperm, the entry of additional spermatozoa is prevented by the cortical reaction, the breakdown of cortical granules that underline the membrane. *See* FERTILIZATION.

Even if the processes of gamete transport succeed in bringing capacitated spermatozoa and recently ovulated eggs into functional contact in the appropriate portion of the oviduct, a viable zygote may still not be formed. Anomalies of fertilization such as digyny (retention of the second polar body) or polyspermy (penetration of the egg cytoplasm by more than one sperm) are pathological in the higher mammals, and nearly always associated with very early embryonic death. Digyny or polyspermy arises especially following postovulatory aging of the egg, a situation in which control of the numbers of spermatozoa passing through the isthmus to the site of fertilization is also less effective.

Regulation of early embryonic development and transport. As noted above, fertilization takes place in the oviducts of mammals, usually in the region of the ampullary-isthmic junction, and the fertilized eggs or embryos do not descend to the uterus for some 3 to 4 days in most species. During this interval, the embryo undergoes a series of mitotic divisions until it comprises a sphere of 8 or 16 cells and is termed a morula. Formation of a blastocyst occurs when the cells of the morula rearrange themselves around a central, fluid-filled cavity, the blastocoele. As the blastocyst develops within the uterine environment, it sheds its protective coat, the zona pellucida, and undergoes further differentiation before developing an intimate association with the endometrium, which represents the commencement of implantation or nidation. Early development of embryos in most species is largely under the control of the maternal genome, with sequential activation and utilization of oocyte components. Expression of the embryonic genome does not occur until after the two-cell stage in mice, four-cell stage in cattle, four- to eight-cell stage in humans, and eight-cell stage in sheep zygotes.

During this preimplantation period, whose duration is approximately 4 days in small laboratory species, 7 days in humans, and 2 to 3 weeks in pigs and sheep, respectively, metabolic support of the embryo depends partially on its own cytoplasmic reserves of lipids and other substrates, but to a far greater extent on constituents of oviducal and uterine secretions. The nature and volume of the secretions are known to be closely regulated by the prevailing balance of ovarian hormones, estrogen and progesterone, and since the corpus luteum is developing during this period, it follows that the secretion of progesterone will also be increasing. Thus, the dynamic situation in the pattern of ovarian steroid secretion will be reflected in a changing pattern of oviducal and uterine fluids, which is critically adjusted to meet the nutritional requirements of the developing embryo. Many experiments show that the nature of the substrate available in oviducal fluid is closely regulated by the prevailing balance of ovarian hormones and is therefore changing to match the embryo's requirement for and ability to utilize the substrates. Mouse one-cell zygotes require pyruvate or oxaloacetate as an energy source, but at the two-cell stage can also utilize phosphoenolpyruvate and lactate; only after the eight-cell stage can glucose be utilized. Human zygotes appear much less dependent on a critical composition of the surrounding medium for growth and development, since many media, both simple and complex, have been used for culture of oocytes prior to the successful replacement of the

resulting zygotes into the oviduct or uterus of the infertile woman.

In nonprimate species, when developing zygotes are delayed in the oviduct beyond the time of normal tubal transport, they fail to implant; similarly, when they reach the uterus too early, subsequent development can be compromised. In primates, however, a zygote reaching the blastocyst stage in the oviduct can implant, resulting in an ectopic pregnancy. This condition is rare in nonhuman primates but not uncommon in humans, accounting for about 1% of all pregnancies. The reason for the occurrence of this condition, which is potentially life-threatening because of hemorrhage and shock subsequent to tubal rupture, is unknown; in nonprimates, oviducal secretions may be hostile to continued zygote development after the normal period of tubal transport, approximately 72 h. *See* PREGNANCY DISORDERS.

Coincident with their influence upon secretory activity, the ovarian hormones also regulate muscular activity in the oviducts, and so control passage of the embryo into the uterus. Any interference with the ratio of estrogens to progesterone during the free-living, preimplantation stages of the embryo will adversely influence fertility through dual effects on transport and nutrition of the embryo. Administration of suitable preparations of ovarian steroid hormones, such as those contained in the contraceptive pill, thereby provide potent means of regulating fertility even when taken after coitus. In fact, in some instances, the principal mechanism whereby ovarian steroidal hormones regulate fertility is not by blocking ovulation via steroidal feedback mechanisms, but by upsetting the delicate relationship between the egg and the reproductive tract, leading to failure of fertilization or degeneration of the egg.

The factors that control muscular activity of the oviduct, and hence modulate the transport of the developing zygotes to the uterus at the correct time, have not been defined. However, it is known that oviducal motility can be influenced by prostaglandins, catecholamines, peptides (such as vasoactive intestinal peptide, neuropeptide Y, substance P, oxytocin, and vasopressin), cyclic nucleotides, and gamma aminobutyric acid. Flow of oviducal fluid and ciliary activity of epithelial cells lining the oviducts are also thought to be involved in the transport process.

Implantation. Association of the embryo with the uterine epithelium, either by superficial attachment or specific embedding in or beneath the endometrium, leads in due course to the formation of a placenta and complete dependence of the differentiating embryo upon metabolic support from the mother. Implantation and placentation exhibit a variety of forms, but in all instances the hormonal status of the mother is of great importance in determining whether or not implantation can proceed. In several species of rodent such as rats and mice, implantation is held in abeyance during lactation, while in other mammals such as roe deer, badger, and certain bears, there is an obligatory period of delayed implantation known as embryonic diapause; the inappropriate endocrine conditions associated with delayed implantation render the blastocyst metabolically dormant. Ovariectomy shortly before the expected time of implantation can prevent implantation, whereas ovariectomy plus replacement therapy with a suitable ratio of progesterone and estrogen permits implantation.

Despite this dominance by maternal hormonal factors, studies have indicated that the embryo itself may be modifying the local environment between its trophoblastic surface and the endometrium by means of steroid hormone, prostaglandin, and platelet-activating factor secretion and perhaps also by production of carbon dioxide (CO_2). The cells of the trophoblast possess enzymes capable of synthesizing ovarian hormones, and embryonic membranes supplied in culture with suitable biochemical substrates will secrete estrogens and progesterone. Such a pattern of synthetic activity in the late preimplantation embryo, which has also been indicated by histochemical studies, may explain in part the regional changes in the endometrium corresponding specifically to the future implantation sites, and also the local changes in capillary permeability of the uterus as demonstrated by the Psychoyos pontamine-blue reaction.

The physical relationship between the implanting embryo and the uterus is best summarized according to whether the embryo remains in the lumen of the uterus, referred to as central implantation as in cattle, sheep, and pigs, or colonizes the glandular crypts and maternal tissues leading to eccentric or interstitial implantation, as in rodents and humans. The process of invasion by the embryonic trophoblast in the mouse, for example, promotes characteristic changes in the uterine stroma termed the decidual reaction. This cellular proliferation occurs on the fourth day of pregnancy in mice in response to the presence of a blastocyst, but the precise physiological nature of the stimulus leading to these changes is not understood.

One other feature of the implantation process that influences the form of the placenta is whether maternal cell layers are bypassed during development of the implantation association. In species where the embryo is invasive and penetrates the maternal tissues to reach the wall or lumen of the blood vessels, this gives rise to an endotheliochorial (carnivores) or hemochorial (rodents, primates) form of placenta. In other words, cells of the maternal epithelium no longer interpose between the maternal vascular supply and the developing embryo, a situation which facilitates transfer of nutrients to the fetus.

Prevention of implantation. It should be apparent that the process of implantation is susceptible to interference by hormonal or physical means, and provides a major avenue for the regulation of human fertility. Apart from contraceptive treatment with steroid hormones that disturb the oviducal transport of eggs or embryos (as described above) and hence the time relationships essential to implantation, the application of intrauterine devices (IUDs) in a variety of forms has also been used to prevent implantation

and development of the embryo. Such devices consist of threads; plastic coils; and T- or 7-shaped plastic devices that release copper or T-shaped devices that release progesterone or a more potent progestational steroid, levonorgestrel; they are inserted into the uterus via the cervical canal.

Their mechanism of action in inhibiting implantation remains to be clarified, but whatever form it takes, it does not prevent the rapid restoration of fertility upon loss or removal of the device from the uterus. A component of their action is almost certainly mechanical, disturbing the spatial relationship between the blastocyst and endometrium, but recent studies on laboratory and farm species indicate that the abnormal uterine fluids produced in the presence of IUDs must also play a role directly or indirectly in preventing development of the embryo and its subsequent implantation. The observation that the addition of copper or progesterone to inert IUDs markedly increases their contraceptive efficiency suggests the occurrence of toxic or inflammatory effects in the uterus acting on the blastocyst as well as on the endometrium itself. While systematic research continues to clarify the basic mechanisms involved, particularly the manner whereby uterine prostaglandins may mediate the observed effects, the use of IUDs by many women in Western societies and in developing countries continues to give substantial protection against pregnancy, without resort to daily or weekly regimes of pill taking. In the United States the availability of IUDs has been greatly restricted because of potential legal liability.

The most widely used method of fertility control is sterilization—vasectomy in the male and tubal ligation, resection, or cauterization in the female. In general, sterilization is nonreversible. Various plastic devices (capsules, IUDs, vaginal rings) or long-acting injectable biodegradable formulations that release potent progestional agents such as levonorgestrel, alone or in combinations with an estrogen, are possible alternatives to conventional oral contraceptives. Such devices and formulations are especially suitable for use in developing countries. The most advanced method available in certain countries comprises six silastic capsules that release levonorgestrel and are inserted under the skin of the forearm of the woman. They can provide effective contraception for up to 5 years. Methods to inhibit male fertility, other than condoms and sterilization, are not well advanced.

Much research has been devoted to the development of contraceptive vaccines. The most advanced method involves using the beta subunit of human chorionic gonadotropin (hCG) or a portion joined to a foreign protein to induce antibodies to hCG, which is required to maintain luteal function (secretion of progesterone) during early pregnancy in the human. Other antigens under consideration as the basis for a contraceptive vaccine involve proteins or glycoproteins isolated from spermatozoa or the zona pellucida (the outer coating) of the ovum. Antibodies to such antigens would prevent fertilization. Difficulties involved in development of contraceptive vaccines include risks of irreversibility, induction of autoimmune diseases, risk of teratological damage at low antibody titers, and the need to make each vaccination effective for at least 1 year.

Fertility regulation after implantation. Research is also attempting to develop a means of regulating fertility subsequent to the process of implantation. The luteotrophic support required for the early maintenance of pregnancy may be vulnerable to immunological blockade, although in treatments requiring antibody administration, the shortcomings mentioned above must be taken into account. The alternative approach to this form of hormonal regulation is the use of compounds that are directly luteolytic in nature or block the action of progesterone at the uterine receptors and that rapidly terminate the endocrine support of pregnancy. The prostaglandin $F_{2\alpha}$ is effective as a luteolytic agent in animals but not in primates. However, administration of prostaglandins alone during the first 49 days of pregnancy can terminate the pregnancy by directly stimulating uterine contractility with expulsion of the fetus. Side effects with this treatment are common and, although in general not life-threatening, make the method less acceptable. More promising is a potent progesterone antagonist, mifepristone, which combined with a low dose of a prostaglandin is very effective in terminating early pregnancy (up to 49 days) with minimal side effects. However, this method is available in only a few countries.

Endocrine function in reproduction. The endocrine glands secrete certain substances (hormones) which are necessary for growth, metabolism, reproduction, response to stress, and various other physiological processes. The endocrine glands most concerned with the process of reproduction are the pituitary and the gonads. Other contributions to the proper functioning of the reproductive system are provided by the thyroid, pineal, and adrenal glands. The hypothalamus is also an important organ for regulating reproduction. *See* ENDOCRINE SYSTEM (VERTEBRATE); ESTROGEN; PITUITARY GLAND.

The formation of gametes (spermatogenesis and oogenesis) is controlled by anterior pituitary hormones. The differentiation of male and female reproductive tracts is influenced, and mating behavior and estrous cycles are controlled, by male or female hormones. The occurrence of the breeding season is mainly dependent upon the activity of the anterior lobe of the pituitary, which is influenced through the nervous system by external factors, such as light and temperature. The ratio of light to dark during the daily cycle affects the secretion of melatonin by the pineal gland in the brain. Melatonin can be either stimulatory or inhibitory for reproductive activity in a wide variety of mammals, especially those with defined periods of estrus and anestrus (for example sheep and many rodents). Its role in human reproduction is still uncertain, but it can inhibit ovulation in certain circumstances. The transportation of ova from the ovary to the Fallopian tube and their subsequent transportation, development, and implantation in the uterus are controlled by a balanced ratio

between estrogen and progesterone. Furthermore, it is known that estrogens, androgens, and progesterone can all have the effect of inhibiting the production or the secretion, or both, of gonadotrophic hormones, permitting the cyclic changes of reproductive activity among different animals.

Mammary glands are essential for the nursing of young. Their growth, differentiation, and secretion of milk, and in fact the whole process of lactation, are controlled by pituitary hormones as well as by estrogen and progesterone. Other glands and physiological activities also influence lactation, although this is largely via the trophic support of other pituitary hormones.

M. C. Chang; Michael J. K. Harper; R. H. F. Hunter

Nonmammalian species. The successful radiation of recent vertebrates has depended upon the adaptation of their reproductive cycles to the environments in which they have evolved. Since almost all environments show some degree of fluctuation (water, food supplies, temperature, light), control of the reproductive cycle of males and females of a species is essential if young are to be born at a time of year that is advantageous to survival. Such control is brought about through the brain-pituitary-gonadal axis. In this system, information about the external environment is detected by appropriate sensory structures (for example, the lateral eyes) and released to the basal hypothalamus of the brain, which regulates pituitary production of gonadotropic hormones and thus synchronizes gonadal development with environmental cues. Evidence is available for all vertebrate groups (except the degenerate hagfishes, or lampreys) to suggest that control of the gonad by the brain-pituitary system exists. This involves a brain peptide, gonadotropin-releasing hormone, that stimulates the release of two pituitary gonadotropins. One of these, the homolog of mammalian follicle-stimulating hormone, is concerned with gamete production; the other, a homolog of mammalian luteinizing hormone, is primarily concerned with gonadal hormone production and ovulation in conjunction with follicle-stimulating hormone. With few exceptions, the known testicular and ovarian steroids that are synthesized and secreted by nonmammalian gonads are identical to the major gonadal steroids of mammals. These are progesterone and testosterone and their oxygenated and hydroxylated derivatives and the estrogens, of which the most important is estradiol-17β. The testicular and ovarian sites of synthesis and secretion of these hormones appear to be similar, if not identical, at least down to the elasmobranch level of vertebrate gonadal organization. Plasma levels of the gonadal steroids of both sexes reflect the stage of germ cell development and correlated steroidogenic cell function. The plasma steroids thus undergo marked fluctuations associated with the gonadal cycle and ensure proper central nervous function and development of the peripheral structures of the reproductive tract so that behavioral responses and internal (elasmobranchs, reptiles, and birds) or external (teleosts and amphibia) fertilization occur.

Spermatogenesis. Throughout the Vertebrata the cytological sequence of changes from immature, round germ cells through mature, flagellated spermatozoa is virtually identical. From early stages, germ cells form part of an anatomically and functionally related clone of cells, all derived from the same stem cell and synchronized through development. In addition to the germ cells, the primary spermatogenic units (spermatocysts) include somatic elements (the Sertoli cells) that provide structural and nutritive support and in all ways mediate between the soma and the developing germ cell. The spermatogenic process begins with a number of premeiotic spermatogonial divisions. These mitotic stages are followed by meiosis (spermatocyte stage), during which the $2n$ (diploid) number of chromosomes is reduced to the round spermatid stage. The spermatids enter a metamorphic stage during which they are transformed into elongate, flagellated spermatozoa.

The primary unit of testicular structure is the spermatoblast or spermatocyst, which consists of germ cell clones plus associated Sertoli cells. In anamniote species, germ cell clones are anatomically discrete spherical units. In contrast, in amniotes (reptiles, birds, mammals) the spermatocyst is not closed or spherical but irregular in form and open to other clones and the tubular lumen. The individual Sertoli cells of amniotes, unlike those of anamniotes, are shared by four or five germ cell clones in succeeding generations. These nonrandom cellular associations signify functional relationships that are the basis for the cycle of the seminiferous epithelium and the spermatogenic wave of amniotes. Apart from the differences in clonal association of the Sertoli cells of the amniotes compared to anamniotes, the Sertoli cells of some anamniotes (such as fish or urodele amphibians) proliferate, and they remain associated and develop synchronously with a given germ cell clone throughout spermatogenesis; at spermiation they degenerate or are otherwise lost into the semen. By contrast, Sertoli cells of reptiles, birds, and mammals are permanent elements of the seminiferous epithelium, nurturing generation after generation of developing germ cells. In these groups, Sertoli cell mitoses are thought to be restricted to prepubertal development.

The secondary germinal compartment is delineated by a boundary wall and composed of blind-ended sacs (lobules) or open-ended lobules, both of which are continuous with the intratesticular collecting duct system and surrounded by the endocrine Leydig cells. In the elasmobranchs, however, definitive Leydig cells appear to be absent; the spermatocysts are embedded in connective tissue stroma and joined to collecting ducts via a short stalk patent only at spermiation. In some teleosts and in urodele amphibians, a secondary germinal compartment can be identified in mid-late developmental stages, with stem cells and spermatogonial stage spermatocysts embedded directly in connective tissue stroma. The secondary germinal units are continuously formed adjacent to the germinal region and then regress after spermiation. Leydig cells are limited to

lobular-tubular regions. In some teleosts and anuran amphibians and in all amniotes, Sertoli cells and germ cells are permanent components of the adult testis, containing germinal elements at all developmental stages. Interstitial Leydig cells are present.

In all species, there is a strict temporal and spatial relationship among the different germ cell stages. Species variation is evident, however, in the timing and anatomical organization of cells in the spermatogenic sequence. The unique developmental sequence of spermatogenesis is regulated at several levels: at the organismal level, by circulating gonadotropic hormones; at the organ level, by local (paracrine-autocrine) factors; and at the cellular level, by direct cell–cell interactions. Comparative studies suggest that the Leydig cells evolved primarily for the secretion of steroids into the peripheral circulation. Although of secondary importance to the Leydig cell in the supply of steroid hormones, the nonmammalian Sertoli cell has a steroidogenic capacity that is quantitatively more important than that of mammals in providing the steroidal needs of germ cell production; in elasmobranchs, Sertoli cells are the primary steroidogenic component of the testis throughout the year. The elasmobranch testis is of particular interest, inasmuch as germ cell–Sertoli cell units can be harvested at specific stages of development, and it is possible to correlate steroidogenic function with specific stages. Androgen production is correlated with the more mature stages of sperm development, and androgen synthesis and specific estrogen receptors are significantly correlated with earlier spermatogenic stages. Studies with the mud puppy, *Necturus*, support a role for estrogen in the proliferation of early germinal elements because both androgen and estrogen receptors are found in testicular zones with immature germ cells. As in mammals, in nonmammals gonadal steroid hormones are transported within the testis and blood by androgen-binding proteins, as demonstrated in trout and shark. *See* TESTIS.

Oogenesis. As in mammals, oogonial proliferation in nonmammalian vertebrates is characterized by the presence of intercellular bridges, probably important in synchronizing of mitosis. Oogonial proliferation may be limited to the embryonic (larval) period (in lampreys, elasmobranchs, some bony fishes, and birds) or to the adult (in most teleosts, amphibians, and reptiles). Little is known of the role of hormones in the regulation of oogonial proliferation, but some evidence suggests that gonadotropins and possibly gonadal steroids increase oogonial division in some adult fish, amphibians, and reptiles. Primary oocyte growth appears to be gonadotropin-independent in cyclostomes; in contrast, in amphibians and reptiles the process is gonadotropin-dependent, and atresia of primary oocytes in hypophysectomized amphibians and reptiles is prevented by gonadotropins. Folliculogenesis, or the investment of the primary oocyte with follicle cells to form a primordial follicle, is similar in all nonmammalian vertebrates. As in mammals, the trigger for folliculogenesis is not known, but considerable evidence suggests that at least the vitellogenic stages of follicular growth in all nonmammalian species are dependent upon pituitary gonadotropins. In amphibia and teleost fish, the resumption of meiosis at the onset of oocyte maturation and subsequent ovulation is gonadotropin- and progestin-dependent, as in mammals. Although all nonmammals produce yolky eggs, the lampreys, teleosts, and amphibians generally produce large numbers of small yolky oocytes, and the elasmobranchs, reptiles, and birds produce fewer, larger yolky eggs. A hierarchical arrangement of follicles or cohorts of follicles is readily seen in some reptiles and birds. Smaller ovarian follicles represent future generations of oocytes to be ovulated sequentially.

The process of hepatic yolk protein synthesis, vitellogenesis, is regulated by the ovarian hormone 17(β)-estradiol and thus indirectly by the pituitary gland. The action of estradiol on the induction of yolk protein synthesis by the liver has been well studied. The process is steroid receptor–ediated and involves the hormonal induction of specific vitellogenin genes in liver cells. Vitellogenin, a complex phospholipoglycoprotein, is secreted into the blood by hepatic cells and is sequestered from the blood by developing oocytes under the influence of gonadotropin. In reptiles, it appears that hormones other than estradiol are involved in vitellogenesis. Growth hormone is synergistic with estrogen and required for the full expression of the estradiol effect. In contrast, progesterone, testosterone, and prolactin appear to be inhibitory to vitellogenin synthesis. Of these hormones, the progesterone is of particular interest in the vertebrate groups in which viviparity (live-bearing) is a common reproductive strategy (as opposed to egg laying, such as in elasmobranchs or squamate reptiles).

The evolution of live-bearing from the primitive egg-laying mode of reproduction has occurred in parallel in all vertebrate groups except birds. Physiologically, this process requires the retention of eggs, loss of the outer eggshell membranes, and the ultimate development of a placental mechanism. The placenta, when present, is involved in fetal nutrition and ultimately supplants the process of vitellogenesis. In elasmobranchs and reptiles, species have been described in which the dependency of the nutrition of intrauterine young upon the placenta has reached a near-mammalian condition. A corollary of this development is the reduction or elimination of yolk from the eggs. Some evidence suggests that the ovarian hormone progesterone is responsible for the switch from oviparity to viviparity through an inhibitory action on the smooth muscle of the reproductive tract. This tends to favor egg retention and thus viviparity. In mammals, progesterone is of primary importance for the maintenance of a quiescent uterus during gestation. Since even elasmobranch corpora lutea (formed from follicles after ovulation) synthesize and secrete progesterone, which inhibits uterine muscle activity, it appears that many of the endocrine mechanisms associated with reproduction and reproductive tract adaptations for egg, embryo, and

fetal care were evolved early in vertebrate evolution. Coincident with the effect of progesterone on the reproductive tract and egg retention, the function of the hormone to switch off cyclic vitellogenesis is extended during evolution to a progressive inhibition of vitellogenesis as placentation takes on the role of embryo-fetal nutrition.

In general, the two layers of the ovarian follicle wall (theca and granulosa) cooperate to carry out the enzymatic conversions from cholesterol through progesterone and androgens to estrogens. However, the two components have different roles in different species. Thus, androgens are converted to estrogens by the granulosa cells of fish and mammals, but this function is performed by the theca in birds. Conversely, androgens are produced by the mammalian theca, but by the avian granulosa. In elasmobranchs, the granulosa elements are the primary steroidogenic components, being able to synthesize estrogens from cholesterol. In a squamate reptile (snake), thecal synthesis of estradiol is low in the absence of substrate in the form of progesterone supplied by the granulosa. This pattern of enzymatic activities in the snake ovarian follicle is unique in vertebrates. In all vertebrate groups, ovarian follicular steroid synthesis and secretion is gonadotropin-dependent. In at least elasmobranchs and reptiles, corpora lutea formed from the ovulated follicle are steroidogenic and produce progesterone in response to gonadotropins in the living organism and in cell culture. In general, in all vertebrates gonadal steroids have been identified in plasma, and cyclic fluctuations of plasma levels of the principal ovarian steroids—progesterone, testosterone, and estradiol—correlate well with ovarian follicular development and ovulation. There is evidence that the circulating hormones have extensive actions on the peripheral target tissues that are important in the synchronization of the reproductive process in nonmammalian species. Ian P. Callard

Neuroendocrine function. The associated physiological activities of the nervous system and of the endocrine system that influence animal reproduction are briefly dealt with here to illustrate the intrinsic mechanism involved. In the lower forms of life, especially in those forms without a nervous system, the rhythm of reproduction may be controlled metabolically by the direct action of environmental factors—food, temperature, light, humidity, and chemical composition of the environment. In the higher forms certain external factors act through the intermediation of the nervous system. In the bird, the number of eggs in a clutch is generally constant within narrow limits; if the eggs are withdrawn shortly after they are laid, many birds will go on laying, making an attempt to lay the right number. In the pigeon, ovulation often is induced by courtship with another pigeon. The number of eggs laid has been reported to be increased if a mirror is placed in front of a pigeon cage. These instances illustrate the influence of the nervous system on reproduction. *See* NEUROSECRETION.

Stimuli. Experimental study on birds and on ferrets has shown that breeding can be induced in midwinter by artificial light. Hypophysectomized and blind ferrets do not ordinarily react to light as expected, and it is obvious that the stimulus must be passed through the eye, optic nerve, or some receptors in the pineal gland or brain region, and via changes in secretion of melatonin or other indoles, and thence to the anterior pituitary. Moreover, the rabbit, ferret, and ground squirrel normally ovulate in response to the stimulation of copulation. Since this stimulation to switch from the follicular phase to the luteal phase cannot be effected in the absence of the pituitary but can be brought about by injecting pituitary extracts or pituitarylike extracts, pregnant woman's urine (PU) or pregnant mare serum (PMS), it would seem that this stimulus is normally due to nervous reflexes mediated through the hypothalamus and pituitary. The stimulus, however, may be carried by several nervous paths, because local anesthesia of vagina and vulva, complete thoracosympathectomy, absence of any nerve pathway to the ovaries, or cervical sympathectomy do not inhibit ovulation after coitus. Because stimulation of the brain, of the lumbosacral part of the spinal cord, of the cervical sympathetic ganglion, or of the hypothalamus will induce ovulation to a certain extent, it seems that more than one nervous path and more than one mechanism for the initiation of ovulation must be involved. Furthermore, the rat, unlike the rabbit, ovulates spontaneously, but a prolongation of the life of the corpora lutea with subsequent pseudopregnancy can be induced by sterile mating, mechanical stimulation of the cervix, or electrical stimulation of the brain. Pseudopregnancy in the rat also seems to be mediated by the pituitary through nervous pathways. There is additional evidence to show that the stimulus for luteinizing hormone release in a spontaneously ovulating animal is controlled by nervous mechanisms employing cholinergic and adrenergic components.

Lesions. It is well established that lesions in the basal tuber or median eminence induce ovarian atrophy in the cat, dog, and rabbit. The role of the nervous system in establishing cyclic pituitary activity has also been emphasized. Localization of an erection center and of an ejaculation center in the hypothalamus has been reported. Appropriately placed hypothalamic lesions in the female guinea pig sometimes result in anestrus or prolonged estrous periods with sexual behavior in keeping with the gross changes of the cycle. In the male guinea pig, similar lesions induce sexual impotence, without genital regression. Hypothalamic production of oxytocin and the derivation of vasopressin from neurosecretory process have been postulated. As for mating behavior, it is assumed that advancing evolutionary status is accompanied by a progressive dominance of the nervous system and a corresponding reduction of endocrine control. To interpret the effect of sex hormones on mating behavior it has been suggested that their activity may increase the excitability of the central excitatory mechanism.

Fertility and sterility. The ability or inability to produce offspring is termed fertility or sterility. Fertility

and sterility occur in different grades among various species and among individuals of the same species. Absolute sterility is rare, but infertility of all degrees is very common, especially among higher vertebrates. The rate of reproduction in any species depends upon the average number of young born in each litter, the frequency of recurrence of breeding season, the duration of the reproductive period, and the age at which the animal starts to breed. The age as a general rule is earlier in small species than in large ones. In general, the number of young in a litter of mammals is inversely proportional to the size of the animal. For instance, a cow rarely produces twins, whereas the rat occasionally bears as many as 16 young. A theory of fertility has been proposed that states that individuation and genesis vary inversely; that is, the power to sustain individual life and the power to produce new individuals are inversely proportional. Where there is abundant food supply and a favorable environment, and the necessary expenditure of energy is relatively slight, the cost of individuation is much reduced and the rate of genesis is correspondingly increased. *See* INFERTILITY.

Factors controlling fertility. J. Hammond proposed that three factors control fertility:

1. The number of ova shed. In accordance with the genetic constitution of the species and the nutritional status, the number of ova shed is controlled by gonadotrophic hormones through the pituitary gland, but influenced by external factors. Before puberty, at old age, and during pregnancy, pseudopregnancy, and early lactation, very few mammals ovulate. Although by administering gonadotrophic hormones the possibility of increasing the number of ova shed in mammals was demonstrated, the actual number of young produced was rather low, probably as a result of limitations to support excessive embryos.

2. The number of ova fertilized. This depends upon the number of spermatozoa produced by the male, and the morphological and physiological integrity of the gametes. It also depends upon the probability of meeting between gametes provided by the male and female in the lower vertebrates, the efficiency of sperm transport to the site of fertilization, and the time of mating in the higher vertebrates. Aging phenomena in gametes have a deleterious influence on fertility.

3. The number of embryos developing into self-sustaining individuals. The probability of normal development depends upon where ova are deposited and the protection that the parents give to the zygotes in the lower vertebrates. In the higher animals, it depends upon the transportation of zygotes to the prepared uterus at the right time, and the physiological activities of embryo and of mother for proper implantation, for the maintenance of pregnancy, and for proper parturition.

Control of reproduction. The topic of controlling fertility, for example, the regulation of implantation in humans described above, should not be left without portraying the considerable technology available to modify fertility in domestic animals.

While artificial breeding of cattle, and to a lesser extent pigs, sheep, and goats, by means of insemination of deep-frozen or diluted semen samples has been extensively applied in recent years, enabling the exploitation of males of superior genetic merit, other features of contemporary reproductive technology are less widely known; attention will be devoted to two of these. The first can be termed controlled breeding, as seen in the various techniques of synchronization of estrus. There are obvious advantages in being able to decide when an animal is to be bred, not only for convenience at the time of artificial insemination or mating, but also because the time of birth, weaning, and marketing of the offspring can be predicted quite accurately. The key to controlled breeding lies in reprogramming the ovary, and treatments attempt to modify, in the short term, the influence of the secretions of the corpus luteum and the Graafian follicle expressed, respectively, in the luteal and follicular phases of the cycle. Inducing synchronized estrus in a population of animals is achieved most effectively by causing regression of the corpus luteum so that animals enter the follicular phase and estrus approximately synchronously.

The demonstration in the late 1960s that prostaglandin $F_{2\alpha}$($PGF_{2\alpha}$) was almost certainly the luteolytic factor in domestic species, informing the ovary that the uterus did not contain a viable embryo and inducing regression of the corpus luteum, has caused this hormone to become the basis of extensive estrous synchronization programs, replacing the progestagenic compounds previously used to simulate and extend the luteal phase prior to a synchronized follicular phase. Potent synthetic analogs of $PGF_{2\alpha}$ are now available that cause luteolysis following a single intramuscular injection. However, the corpus luteum in cattle is only susceptible to injected $PGF_{2\alpha}$ between days 5 and 15 of the 21-day estrous cycle, and so to achieve synchronization in all animals in a herd, two injections are given 10 days apart. Nearly all animals will be in estrus within 48–72 h of the second injection, when they can be inseminated as a group. Prostaglandins are being used in this manner extensively.

Another area of reproductive technology in domestic animals receiving attention is that of storage and transplantation of embryos, particularly in the context of exploiting rare or exotic breeds of cattle after procedures of superovulation to increase egg numbers. Transplantation of embryos is, of itself, not a new technique, for it was successfully performed in rabbits before the turn of the century, and since about 1950 embryo transfer has been a leading experimental tool in reproductive physiology. Two of the biological facts permitting such manipulations of the embryo are the relatively long period during which the blastocyst is free-living in the uterine lumen before implantation, especially in pigs, sheep, and cattle (ungulates), and also the fact that the embryo is not rejected as foreign material from the host uterus.

Apart from the care needed during handling and storage of embryos in the culture medium, many

experiments have indicated that the single most important factor determining the success of a transfer operation is the degree of synchronization of estrous cycles between the donor and recipient animals. In cattle and sheep, there is some degree of tolerance in the timing, but once development of the embryo is more than 24 h out of phase with that of the recipient uterus, the pregnancy rate commences to decline sharply. Transfer of single, freshly recovered embryos by a surgical means involving an incision to expose the uterus can lead to 90% of the transferred embryos developing into calves. Transfer of two embryos, one to each horn of the uterus, has been used to induce twinning in cattle, although here the success rate is closer to 70%. Nonsurgical transfer of embryos, that is, their introduction into the uterus from the vagina by means of a suitable pipette, has been much less satisfactory, rarely giving pregnancy rates of greater than 20–40%. Apart from the greater risk of introducing infection by this approach, there is also the problem of ejection of embryos which usually follows manipulation of the cervix during the luteal phase. Nevertheless, now that bovine embryos can be preserved at −321°F (−196°C), as has been the case for spermatozoa for many years, the emphasis is on overcoming the problems of nonsurgical transplantation. With synchronization of estrus a practical proposition, and deep-frozen banks of embryos in the offing, the attraction of breeding programs involving this technically straightforward means of propagation is considerable, although it can be economically justified only for valuable animals with desirable genetic traits.

In several species of farm animals (sheep, cattle, and pigs), zygotes up to the 8–16-cell stage can be halved or quartered. The separated clumps of cells can then be placed in homologous zonae, from which all cytoplasmic material has been removed, and can develop into normal young after retransfer to a suitable recipient. Not only does this technology rapidly create additional copies of the traits of a particular mother or father or both, but it also provides for identical twins, triplets, and so forth for studies of the impact of environmental and nutritional factors on the same genetic background. Such a technique will be useful primarily as a research tool rather than for production purposes, where transgenic animals will be more useful.

Such advances in control of animal reproduction have given rise to the application of these techniques to humans for the alleviation of infertility. Men with low sperm counts or poor spermatozoal motility can be assisted by collection of the semen specimen and concentration of the spermatozoa. The chance of pregnancy can be improved if the female partner undergoes artificial insemination with such treated sperm samples at the time of midcycle ovulation.

Infertile women can be induced to mature larger numbers of follicles by treatment with hormones. Ova from such follicles can then be recovered and placed in the oviduct with a prepared sperm sample; matured and fertilized in a culture dish and replaced in the oviduct; or matured, fertilized, and cultured in a dish and then replaced into the uterus. Many thousands of children have been born worldwide following such procedures.

In addition, zygotes can still develop normally after removal of one or more blastomeres. Through the use of polymerase chain reaction amplification of gene sequences from such a blastomere, prenatal diagnosis of genetic defects can be detected prior to replacement of an externally fertilized embryo in the mother. Such techniques can revolutionize the field of prenatal diagnosis, and have been used to detect a number of hereditary disorders. *See* PRENATAL DIAGNOSIS.

In animals, research has focused on the production of transgenic animals. Most experiments have been conducted in mice, although the techniques developed are being tried in farm animals. Genetic material with desirable characteristics can be isolated with the techniques of molecular biology, and the deoxyribonucleic acid (DNA) then microinjected into the ovum, or inserted into embryonic stem cells in a culture dish, which cells can then be microinjected into the cavity of a blastocyst and become commingled with the cells of the host inner cell mass. The foreign DNA can integrate into the host chromosomal DNA and be carried into both germ and somatic cells. The transgenic animal passes the foreign DNA to its offspring in a mendelian manner. After introduction of genes that are normally expressed in a tissue-specific manner, the location of DNA sequences important in developmental programming of each gene can be determined. The gene for rat growth hormone injected into mouse ova gave rise to mice that grew much larger than their noninjected siblings.

Other areas in which contemporary reproductive technology is attempting to increase animal productivity include induction of ovulation at a precise time, increasing the numbers of eggs shed (superovulation), extending the breeding season of sheep and goats, and perhaps most important of all, regulating the time of parturition to avoid perinatal losses. In the future this technology will doubtless also include the application of pheromones to modify in diverse ways the reproductive activity of farm livestock, and the possibility of predetermining the sex of offspring by insemination of X- or Y-bearing spermatozoa in fractionated semen samples. *See* GENETIC ENGINEERING; REPRODUCTIVE TECHNOLOGY.

M. C. Chang; Michael J. K. Harper; R. H. F. Hunter

Bibliography. E. Y. Adashi and P. C. Lang (eds.), *The Ovary*, 1993; W. Andrew, *Textbook of Comparative Histology*, 1959; C. R. Austin and R. V. Short (eds.), *Reproduction in Mammals*, 2d ed., vols. 1–5, 1982–1986; G. F. Bastian, *Reproductive System*, 1993; V. Blum, *Vertebrate Reproduction*, 1985; R. H. F. Hunter, *Physiology and Technology of Reproduction in Female Domestic Animals*, 1980; G. Lamming (ed.), *Marshall's Physiology of Reproduction*, vol. 2, 4th ed., 1990; W. R. Lyons et al., The hormonal control of mammary growth and lactation, *Recent Prog. Hormone Res.*, 14:219–248, 1958;

P. Pang and M. Schreibman (eds.), *Vertebrate Endocrinology*, 1991; J. R. Pasqualini, F. A. Kinel, and C. Sumida, *Hormones in the Fetus*, vol. 2, 1991; E. C. Roosen-Runge, *The Process of Spermatogenesis in Animals*, 1977; J. Rossant and R. A. Pedersen (eds.), *Experimental Approaches to Mammalian Embryonic Development*, 1988; M. Shemesh and B. J. Weir (eds.), *Maternal Recognition of Pregnancy and Maintenance of the Corpus Luteum*, *J. Reprod. Fertil.*, suppl. 37, 1989; D. Shoupe (ed.), *Contraception*, 1993; C. K. Weichert and W. Presch, *Elements of Chordate Anatomy*, 4th ed., 1975; R. G. Wolff, *Functional Chordate Anatomy*, 1991; R. M. Wynn and W. P. Jollie (eds.), *Biology of the Uterus*, 2d ed., 1989; S. S. C. Yen and R. B. Jaffee (eds.), *Reproductive Endocrinology: Physiology, Pathophysiology, and Clinical Management*, 4th ed., 1999; K. Yoshinaga and T. Mori (eds.), *Development of Preimplantation Embryos and Their Environment*, vol. 294 of *Progress in Clinical Biological Research*, 1989.

Reproductive system disorders

Those disorders which involve the structures of the human female and male reproductive systems.

Female System

Disorders of the female reproductive system may involve the ovaries, Fallopian tubes, uterus, cervix, vagina, or vulva.

Ovaries. Failure of the ovaries to form normally results in short stature, sterility, and lack of development of female secondary sex characteristics, such as breast growth, fat deposition in buttocks and thighs and mons pubis, and female escutcheon. Destruction of the ovaries after puberty results in loss of fertility, cessation of menses, loss of secondary sex characteristics, and osteoporosis. *See* OSTEOPOROSIS.

Abnormal ovarian development, or dysgenesis, includes chromosomal mosaics: true hermaphrodites, who possess both ovarian and testicular tissue, and pseudohermaphrodites. Male pseudohermaphrodites have undescended male gonads (testicles), but their external genitalia appear female because those tissues are not sensitive to the male hormone testosterone. Female pseudohermaphrodites have female gonads, but their external genitalia appear masculine because they have been exposed to male hormones. This might occur, for example, in certain enzyme defects of the adrenal cortex in which male hormone levels that are higher than normal are present in the circulation.

The ovaries undergo cyclic stimulation of pituitary hormones to produce an egg, and in some women ovulation is accompanied by a day or two of sharp unilateral lower abdominal pain. Either the follicle which produces the egg or the corpus luteum which forms after the egg is liberated can persist as nonneoplastic cysts. These may be asymptomatic except for altering the menstrual cycle, or they may be quite painful. *See* MENSTRUATION.

Physiologic cysts may be quite large and either single or multiple. Single cysts usually resolve spontaneously in the menstruating woman, but they may persist and require surgical removal. In polycystic ovarian syndrome, multiple cysts develop in both ovaries. Women with this condition may also have infertility, obesity, hirsutism (excessive or unusually situated hair growth), absent or unusually infrequent menstruation (respectively, amenorrhea or, oligomenorrhea), or heavy menses. These changes are all due to anovulation (the absence of ovulation) and unopposed estrogen stimulation. This condition is usually treated by clomiphene citrate, which produces ovulation; by oral contraceptive tablets, which suppress hormone production; or by progesterone, which opposes the action of estrogen. Surgical procedures used to be performed on the ovaries; although this may produce ovulation, it was also found to cause adhesions around the ovaries, and is no longer common practice.

Neoplastic enlargement of the ovary can be cystic or solid, benign or malignant. Adenomatous cysts are filled with watery (serous) or thick (mucinous) material, and the nature of the lining cells determines the kind of fluid produced and whether the cyst is malignant. The cysts are frequently bilateral, and even benign ones can grow large enough to interfere with the functions of the gastrointestinal or urinary systems. One of the most common ovarian tumors occurring in young women is the benign adult cystic teratoma (dermoid). This has both solid and cystic elements, and contains tissue from many different organ systems. Solid tumors can arise from the outer or germ-cell layer or the inner stroma, and vary in malignant potential and hormone production. Those tumors which produce male or female hormones can affect the menstrual cycle and produce symptoms such as hirsutism. Malignant ovarian tumors are commonly asymptomatic in their early stages, and may be quite widely spread in the pelvic and abdominal cavity before they are discovered. Surgical removal is still the mainstay of therapy for this disease. However, chemotherapy regimes, both intravenous and intraperitoneal, as well as whole abdominal radiation have been used to extend survival time and improve quality of life. The ovaries can also be the site of metastatic cancer from such locations as the stomach, breast, and colon. *See* OVARIAN DISORDERS; OVARY.

Endometriosis. Endometriosis, which is a condition involving the presence of ectopic endometrial tissue, can affect the ovaries, and is thought to be maintained by cyclic ovarian hormone production. Endometriosis can be found as large "chocolate" cysts of the ovary, called endometriomas, or small "blue domed" cysts; both are filled with old, dark brown blood. Tiny implants of endometriosis commonly called powder burns also occur on the surface of the ovaries, on other pelvic peritoneum, and over the Fallopian tubes and uterus as well. Endometriomas of the ovary may be painful, or may rupture and cause diffuse pelvic pain, while the smaller endometrial implants may cause severe pain with

menstrual periods, generalized pelvic pain, pain with intercourse, and infertility. To treat this disorder it is necessary to first consider the individual's age, reproductive goals, symptoms, and the extent of her disease. Medical management involves using chemicals that suppress ovulation such as oral contraceptive tablets; a weak antiestrogen drug; or gonadotropin-releasing hormone agonist, a synthetic brain hormone that suppresses ovulation and estrogen production. Each of these medications acts in a different manner to suppress ovulation and has different side effects. There are no medical measures available that cure endometriosis. Surgical treatment can be done through the laparoscope with removal of the endometriosis using electrocautery or the laser. An open surgical procedure can be done to cut adhesions, remove endometriosis, and clear the Fallopian tubes. Surgical cure is produced by removal of uterus, tubes, and ovaries. This will lead to menopausal symptoms, but they can usually be managed by continuous low-dose estrogen replacement.

Fallopian tubes (oviducts). Inflammation is the most common disorder of these organs, and if it is repeated or severe, destruction of the tubal lining with closure of the outer ends of the tubes can occur. A pyosalpinx may then develop, in which the tube is distended with pus, and subsequently a hydrosalpinx is left when the purulent material is replaced by a thin watery fluid. This inflammation can be caused by several organisms, such as gonococcus, *Escherichia coli*, bacteroides, and *Chlamydia trachomatis* agent. Sterility commonly results because the tube is permanently closed to passage of egg and sperm.

The second most common problem involving the tube is pregnancy. The egg is fertilized in the outer portion of the tube and descends to implant in the uterus, but in some cases the passage is delayed and the conceptus attaches to the wall of the tube. The tube has a small lumen and thin wall, and the growing pregnancy quickly enlarges and grows through the tube, leading usually to rupture and hemorrhage into the peritoneal cavity. In rare circumstances, the conceptus is aborted out the end of the tube or between the folds of the broad ligament and continues to grow. Usually, however, maternal bleeding is significant enough to cause pain and fainting, and investigation leads to surgical treatment of the problem. *See* PREGNANCY DISORDERS.

Uterus. Developmental abnormalities can occur during formation of the uterus. The Fallopian tubes might not join at all, or might join partially from the cervical end upward. Septae or walls in the vagina and uterus can also occur. These abnormalities are more common in females who are exposed to diethylstilbesterol (DES) during the mother's pregnancy.

The muscle (myometrium) and lining (endometrium) of the uterus are susceptible to various problems, including tumors, infections, and hormonal derangements. Benign tumors of the myometrium (fibroids) are a common disorder, producing irregular enlargement of the uterus and sometimes causing pain, obstruction of the urinary tract, and heavy vaginal bleeding. They are estrogen-dependent and usually regress with menopause. They can be treated with GnRH compounds to shrink them temporarily; they can be surgically removed (myomectomy); or the uterus can be removed (hysterectomy). Muscle tumors are rarely malignant. The uterine lining can form polyps, or can undergo malignant transformation, producing heavy, irregular bleeding. These tumors can occur during the reproductive years or postmenopausally, when they are thought to be stimulated by continuous high doses of estrogen. Treatment is removal of the uterus, tubes, and ovaries.

The uterus can be one site of a significant infection which produces fever and pain and usually involves other organs, such as ovaries and tubes. A more localized form of infection can occur, following delivery, in the endometrium alone, with low-grade fever and irregular bleeding.

Hormonal abnormalities resulting in anovulation can lead to overgrowth and irregular shedding of the endometrium. Exogenous hormones such as those administered to regulate bleeding or those contained in oral contraceptive tablets also affect the uterine lining and can cause irregular bleeding in some women. Hormonal replacement therapy after menopause can also cause regular uterine bleeding like menstrual periods, or irregular bleeding, depending on the way the hormones are administered. However, treatment of abnormal bleeding depends on its cause.

Pregnancy usually proceeds uneventfully, but certain accidents of pregnancy, such as threatened, incomplete, or missed abortion, can produce irregular bleeding and some uterine discomfort. A rare tumor involving the uterus is caused by the abnormal development of trophoblastic tissue into a hydatidiform mole (benign) or choriocarcinoma (malignant). Chemotherapy is currently an effective mode of treatment. *See* PREGNANCY; UTERINE DISORDERS; UTERUS.

Cervix. Infection of the endocervical glands with gonococcus or chlamydia trachomatis agent can occur. This may be asymptomatic, except for producing a mucopurulent discharge, or it may cause pain when the cervix is manipulated, particularly during intercourse. The infection can ascend from this area into the internal genital organs and adjacent structures, so that these organisms should be treated if detected. *See* GONORRHEA.

The cervix can be affected by malignant tumors which can be adenocarcinomas, or tumors of the glandular cells, or more commonly squamous tumors. Cervical cancer is associated with several factors, including early age at first coitus, multiple sex partners, especially at an early age, smoking, and infection with certain subtypes of human papilloma virus. There is no single factor that "causes" this cancer, but rather a combination of these factors is critical. Papanicolaou smears, in which exfoliated cervical cells are specially stained and examined, are used

to detect early changes suggestive of this cancer. Further evaluation is by low-power magnification, called colposcopy, and biopsy, which allows the cancer to be detected in tissue samples. Treatment includes surgery or radiation, depending on the extent and location of the disease. Malignant tumors confined to the surface of the cervix can be removed with a cold knife or laser.

Vagina. Developmental abnormalities of the vagina may include imperforate hymen, septae, both vertical and horizontal, and complete failure of development. These conditions are usually diagnosed by a combination of visual inspection, chromosome evaluation, radiologic studies, and exploratory surgery. If the problem is causing symptoms, surgical correction can be undertaken.

Inflammation of the lining of the vagina can be due to several common organisms: *Monilia*, a fungus; *Trichomonas*, a protozoon; and *Mobiluncus*, a bacterium. Specific chemical treatment is available for each of these entities. Postmenopausal women can experience atrophy or shrinkage of the vagina and mucosa secondary to estrogen deprivation. This produces itching, bleeding, and pain with intercourse. Estrogen therapy can relieve these symptoms. The vaginal support can be weakened through childbirth, or may simply be naturally poor, allowing the bladder and rectum to bulge inward, and the cervix and uterus to protrude from the introitus. This condition can cause a variety of symptoms, including involuntary loss of urine with a cough or laugh, inability to move the bowels without mechanically pushing the stool out, and a sensation of pelvic heaviness. Intravaginal devices called pessaries may improve support, but surgical correction may be necessary.

Vaginal cancer can occur, but it is rare, and cervical cancer can also involve the vagina. Treatment is usually surgical.

Vaginal trauma can lead to significant bleeding. Penetrating straddle injuries can occur, and trauma with intercourse or childbirth can produce lacerations. Surgical repair is usually undertaken.

Vaginismus, or painful spasm of the muscular sidewalls of the vagina, can occur, making intercourse painful. This involuntary muscle spasm then recurs with each episode of intercourse. Treatment involves use of desensitization, muscle relaxation techniques, and, in some cases, psychotherapy.

Chronic spasm of these muscles or muscle pain due to tension can produce chronic pelvic pain. This condition may have no secondary cause or may be caused by musculoskeletal trauma such as falls. It is treated by the use of measures for muscle relaxation, including medication, massage, diathermy, and ultrasound; and by behavioral measures, such as relaxation training, guided imagery, and hypnosis. *See* VAGINAL DISORDERS.

Vulva. Infection of the female external genitalia can be diffuse or localized, involving bacteria, viruses, or fungal agents. Local infections such as a Bartholin's abscess are treated by surgical drainage. Infections of hair follicles may require topical or systemic antibiotics. Fungus infections should be treated by topical antifungal agents. Viral ulcers caused by herpes can be reduced in frequency by oral acyclovir, an acyclic nucleoside; the virus cannot be cured. Viral warts can be treated by chemical application, freezing, or laser vaporization; the wart is removed, but the virus may still live in adjacent tissue.

Inflammation can be caused by allergic reactions to soap, powders, semen, lubricant, or even clothing. Local measures to improve perineal hygiene and decrease irritation usually reduce symptoms. These measures include use of loose cotton underwear; soaking with water and mild agents such as colloidal oatmeal or artificial seawater; air drying; and use of nonocclusive clothing. The most serious disorder of this area is carcinoma arising in the epidermis. This tumor can be unifocal or multifocal, and is usually slow-growing and relatively asymptomatic. It can, however, spread to the lymph nodes, both in the groin and abdomen. Surgical removal is the mainstay of treatment.

Male System

The principal organs of the male reproductive system are the testicles, epididymis, vas deferens and seminal vesicles, prostate gland, urethra, and penis.

Testicles. During fetal life the testicles form in the abdominal cavity near the kidneys and migrate down into the scrotum. Failure of descent permanently damages the sperm-producing cells, but allows the interstitial cells which produce hormones to survive. Sometimes only one testicle remains undescended, in which case fertility is usually normal because gametes form in the other gonad. Undescended testicles develop malignant neoplasms more commonly than normal testicles.

Other causes of male infertility are organic problems which block passage of seminal fluid or interfere with sperm production. Some common examples are metabolic derangements such as hypothyroidism or malnutrition, exposure to radiation or toxic chemicals, infection (most common with mumps virus), and destruction of testicular tissue by twisting and obstruction of the blood supply (torsion).

Absence of testosterone production by the interstitial cells can occur because of castration or disease. If it happens before puberty, there is a failure of development of male secondary sex characteristics, and the individual is eunuchoid, lacking a beard, male escutcheon, deep voice, and libido. Since testosterone is responsible for maintaining these characteristics in the adult male, a decrease in production (or an increase in binding proteins) results in a gradual decrease in these characteristics, which can be seen in men as they pass into old age.

The testes are sites of malignant tumors which carry a high mortality rate. Undescended testicles are more susceptible to malignant transformation, and for that reason should be removed when discovered, and appropriate hormonal replacement instituted. Seminomas are the most common tumors, but the cure rate approaches 80% with surgery and

radiation. Chemotherapy is utilized for recurrent or advanced disease. Other tumors such as embryonal cell carcinomas and teratocarcinomas are of more primitive cell types, occur in young men and metastasize widely. However, about 75% of these individuals are responsive to chemotherapy.

The membranes covering the testicle can become filled with fluid, a condition known as hydrocele. *See* TESTIS.

Epididymis. The major disorder of these paired organs is infection and inflammation, which can lead to scarring and permanent blockage of the ducts. Tuberculous epididymitis is a chronic inflammation, in which the organisms apparently gain access from the bloodstream, while the gonococcus causes an acute inflammation acquired during sexual intercourse and ascending the urethra, vas deferens, and finally the epididymis. Nonspecific inflammation can also occur, usually due to introduction of organisms during instrumentation of the urethra or surgery on the prostate.

Vas deferens and seminal vesicles. These organs are rarely afflicted by disease, although the network of veins surrounding the vas deferens can become engorged and tortuous, and is called a variococoele. This condition may be a cause of lowered sperm production and male infertility. Treatment is operative removal of the veins. The vasa can be surgically interrupted by a simple procedure to produce sterility in the male.

Prostate. Inflammation and development of small calculi are relatively minor ailments of this organ. Bacteria reach the prostate gland by migrating upward from the urethra or down from the kidneys, ureters, and bladder. Acute infection of this gland can cause chills, fever, and bladder irritability and may lead to abscess formation. Gonorrhea is a common cause of this problem, and appropriate antibiotic therapy should be administered. Chronic infections may produce only mild symptoms but can be accompanied by calculi formation.

Benign overgrowth of this gland is the most common and troublesome complaint. This enlargement results in constriction of the urethra, which obstructs urinary outflow and leads to an increasing residual of urine in the bladder. The symptoms of benign prostatic hypertrophy are difficulty and frequency of micturition and increased urinary tract infections because of the residual urine. The causes of this condition are unknown, and treatment is surgical removal of portions of the gland.

Cancer of the prostate is another relatively frequent problem in older males. This tumor spreads primarily to bones, where it is markedly painful, and it produces a large quantity of the enzyme known as acid phosphatase. This growth is stimulated by male sex hormones and inhibited by female sex hormones, so that treatment for advanced disease is removal of the testicles and administration of female hormones, or use of drugs to suppress adrenal androgens. Localized disease can be treated with brachytherapy (the placement of a solid radioisotopic source on the surface of the growth) using radioactive iodine, external beam irradiation, radial surgical removal of the prostate gland, or in some cases cryosurgery. *See* PROSTATE GLAND DISORDERS.

Urethra. Inflammation is the chief disorder of this organ, causing dysuria and a discharge. The gonococcus can cause an infection of this site, and before the age of antibiotic treatment, gonococcal urethritis or its treatment often resulted in scars or structures which obstructed urinary outflow. Other agents can cause urethritis as well, and it now appears that the chlamydia trachomatis organism accounts for a large proportion of those infections once thought to be "nonspecific."

Penis. Inflammation (balanitis) or narrowing (phimosis) of the foreskin, which covers the glans penis, can occur and may interfere with urination. These conditions are unknown in parts of the world where the foreskin is removed for social or medical reasons. Viral infections can produce warty growths (condyloma acuminata) or painful ulcers (herpetic lesions) of the glans or shaft of the penis, and a syphilitic infection can result in the firm, painless ulcer known as a chancre. *See* HERPES; SYPHILIS.

Disorders of erectile capability range from tumescence unaccompanied by sexual desire (Peyronie's disease, priapism) to impotence, both psychologic and secondary to old age. A number of prosthetic devices are available to surgically treat erectile dysfunction of organic cause. Psychosexual counseling to treat psychogenic disorders has had variable success. Carcinoma can arise from the surface epithelium of the penis and spread both locally and via the lymphatic system to the nodes of the groin. This disease is virtually unknown in populations where males are circumcised at birth. *See* INFERTILITY; ONCOLOGY; REPRODUCTIVE SYSTEM; REPRODUCTIVE TECHNOLOGY.

Gay M. Guzinski

Bibliography. L. Copeland, *Textbook of Gynecology*, 2d ed., 1999; A. Herbst et al., *Comprehensive Gynecology*, 4th ed., 2001; N. Kase et al., *Principles and Practice of Clinical Gynecology*, 2d ed., 1990; D. B. Seifer and L. Speroff et al., *Clinical Gynecologic Endocrinology and Infertility*, 6th ed., 1999; P. Walsh et al., *Campbell's Urology*, 7th ed., 1997.

Reproductive technology

Any procedure undertaken to aid in conception, intrauterine development, and birth when natural processes do not function normally. The most common are in vitro fertilization and gamete intrafallopian transfer.

Infertility, which is the inability to conceive during at least 12 months of unprotected intercourse, is an increasingly common problem. Hormonal therapy and microsurgery are used to overcome many hormonal and mechanical forms of infertility, but many types of infertility do not respond to such treatment. The birth of the first infant conceived by in vitro fertilization—in England in 1978—marked the beginning of reproductive technology as a clinical

tool. Thousands of infants have been born as a result of in vitro fertilization, and numerous centers have been established to perform the procedure. The related technological procedure of gamete intrafallopian transfer was introduced in 1984. *See* INFERTILITY.

In vitro fertilization. In vitro fertilization bypasses the Fallopian tubes. Immediately prior to ovulation the mature oocyte is removed from the ovary and placed, together with prepared sperm, in a petri dish for 2–3 days. Fertilization takes place during this time and the fertilized oocyte develops into a two- to eight-cell embryo, which is then transferred into the uterus.

In vitro fertilization is useful for females with absent or severely damaged Fallopian tubes; couples in which the female has endometriosis; when the male has severely reduced sperm counts; or when the couple has immunologic or unexplained infertility for a period of 2 or more years.

The four principal steps of in vitro fertilization are induction and timing of ovulation, oocyte retrieval, fertilization, and embryo transfer.

Induction and timing of ovulation. Prior to ovulation, oocytes are in a resting stage of the first meiotic division (prophase). Under the influence of the midcycle surge of luteinizing hormone, an oocyte matures to the metaphase II stage of meiosis and is released into the oviduct approximately 38 h later. An immature oocyte that has not reached metaphase II is not receptive to fertilization. In order to obtain a mature oocyte, it must be retrieved immediately prior to ovulation.

Although the natural cycle develops only a single oocyte, results can be comparable to medicated cycles if the woman is 35 years of age or less and has very regular cycles. Women over 35 and women who have irregular cycles become less successful with increasing age and are poorer candidates for a natural cycle. Women of any age whose partners have low sperm counts are also poor candidates.

The majority of centers use agents to induce ovulation for in vitro fertilization. These medications allow the simultaneous development and maturation of multiple oocytes. In addition, ovulation can be forced to occur at a specified time so that oocyte retrieval can be precisely scheduled.

Many ovulatory inducing agents are available. Human menopausal gonadotropins can be used either alone or in sequence or combination with clomiphene citrate. A purified compound of follicle-stimulating hormone is also used. The availability of a gonadotropin-releasing hormone agonist has provided another approach.

The various regimens are monitored by obtaining daily serum estradiol levels and frequent pelvic ultrasound examination to determine the size, number, and potential adequacy of the developing follicles. Ideally, estradiol levels will steadily increase, and at least two, or preferably three or more, follicles will reach a diameter of approximately 0.72 in. (18 mm). Before the natural midcycle rise in luteinizing hormone occurs, an injection of human chorionic gonadotropin is administered intramuscularly to simulate the luteinizing hormone surge and initiate resumption of meiosis. Ovulation will occur approximately 36–38 h following an injection of human chorionic gonadotropin, just as it does after a natural rise in serum luteinizing hormone. Failure is typically due to a poor response of estradiol, an insufficient number of developing follicles, or a spontaneous luteinizing hormone surge.

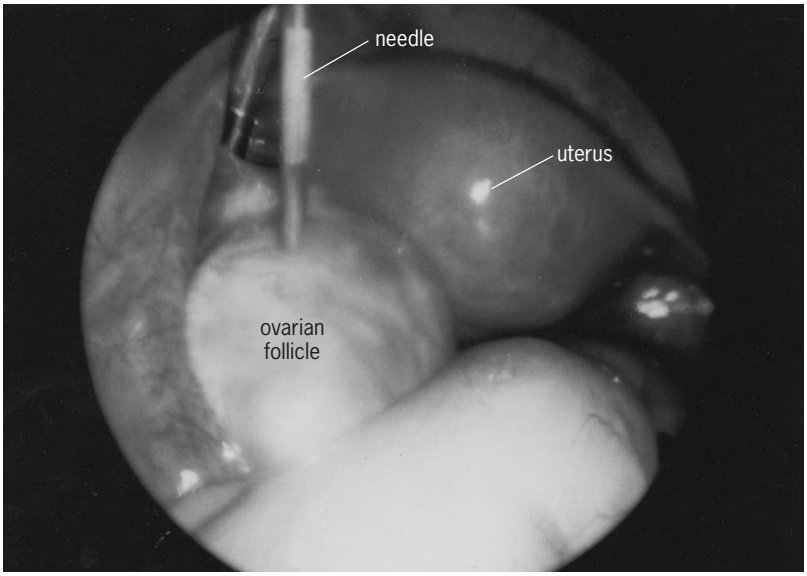

In laparoscopic oocyte retrieval the needle is inserted into the ovarian follicle. The uterus is interior. (*Courtesy of M. M. Seibel*)

Oocyte retrieval. Following the adequate development of preovulatory follicles, a mechanical ovulation or oocyte retrieval is performed. Oocytes are successfully retrieved more than 95% of the time. The traditional method of performing oocyte retrieval has been by laparoscopy (see **illus.**). The laparoscope is inserted in the routine fashion through a subumbilical incision that is 0.5 in. (13 mm) in length. A second small incision is made in the pubic hairline through which a grasping instrument is inserted to stabilize the ovary. A needle is then inserted either through the laparoscope itself or through a small trocar inserted 1–2 in. (25–50 mm) above the second incision. With the aid of the laparoscope, the needle is directed into the follicle, and the follicular fluid and oocyte are aspirated out by means of gentle suction.

Although laparoscopic oocyte retrieval is a highly successful technique, increasing use has been made of ultrasound-guided oocyte retrieval for several reasons. First, many individuals requiring in vitro fertilization have extensive pelvic adhesions that make the ovaries difficult or impossible to visualize by laparoscopy. Without ultrasound, a preparatory laparotomy would be required to make the ovaries accessible for laparoscopic retrieval. A second advantage is that the procedure is less invasive and can, therefore, be performed under local anesthesia augmented with intravenous analgesic agents. These advantages allow ultrasound-guided oocyte retrieval to be performed outside the operating room in an

ambulatory surgery or even an office setting. However, concealed hemorrhage can occur, so that ideally the procedure should be performed in a hospital setting or surgical facility.

The most commonly used ultrasound technique is transvaginal. A transvaginal transducer with an attached needle guide is placed into the vagina and the needle is guided through the posterior cul-de-sac directly into the follicle.

Fertilization. Following retrieval, the oocytes are allowed to complete maturation in culture because they are removed from the follicle several hours prior to ovulation. The male is asked to produce a semen specimen during this incubation period. The spermatozoa are washed in special insemination medium to separate them from the seminal plasma, which is the fluid surrounding the spermatozoa. This process allows the spermatozoa to undergo capacitation, that is, changes in the sperm plasma membrane that must be induced to allow fertilization to occur. Approximately 50,000–100,000 sperm are added to the culture medium containing the oocyte. It is not until the spermatozoa enter the oocyte that the oocyte resumes meiosis and completes its second meiotic division. The oocyte is visualized the following morning for evidence of fertilization. If two pronuclei are seen, fertilization has occurred. By 37 h (range 31–43 h), cleavage will have occurred, resulting in a two-celled stage. Embryos are usually transferred into the uterus 48–72 h following the fertilization. *See* FERTILIZATION.

Embryo transfer. Although the transfer of embryos into the uterus is technically simple, it is the stage at which most in vitro fertilization procedures fail. A speculum is placed in the vagina and the cervix cleansed with sterile saline. One or more embryos, combined with a small volume of culture medium (20–30 microliters), are loaded into a thin catheter, which is threaded through the cervical catheter to within a few centimeters of the uterine fundus, where the embryos are released. The catheter is held in place for 1 min or less and then slowly removed. The woman is asked to remain in the same position for a few minutes.

Gamete intrafallopian transfer. Gamete intrafallopian transfer, or GIFT, is similar to in vitro fertilization with a few important distinctions. The medications used to induce ovulation, and the daily blood testing and pelvic ultrasound are virtually identical to those of in vitro fertilization. The male provides a semen specimen, and it is prepared in a manner similar to that used in in vitro fertilization. However, in gamete intrafallopian tube transfer, the spermatozoa and oocytes are placed into the fimbriated end of the Fallopian tube during the laparoscopy. Indications for gamete intrafallopian transfer include unexplained infertility of two or more years' duration, cervical stenosis, immunologic infertility, oligospermia, and endometriosis. At least one Fallopian tube must appear normal; where there has been severe pelvic adhesions or distorted tubal anatomy from any cause, gamete intrafallopian transfer should not be considered. Individual consideration should be exercised when a patient has had a previous ectopic pregnancy. *See* PREGNANCY; REPRODUCTIVE SYSTEM DISORDERS.

Machelle M. Seibel

Bibliography. J. Bernstein, J. Lewis, and M. M. Seibel, Effect of previous infertility on maternal-fetal attachment, coping styles and self-concept during pregnancy, *J. Women's Health*, 3:125–133, 1994; P. Devroey, Normal fertilization of human oocytes after testicular sperm extraction and intracytoplasmic sperm injection, *Fertil. Steril.*, 62:639–641, 1994; M. M. Seibel et al., *Technology and Infertility: Clinical, Psychosocial, Legal and Ethical Aspects*, 1993; M. M. Seibel, M. Kearnan, and A. Kiessling, Parameters that predict success for natural cycle in vitro fertilization, *Fertil. Steril.*, 63:1251, 1995.

Reptilia

As traditionally defined, a class of vertebrates composed of four living orders—Testudines (turtles), Rhycocephalia (tuataras), Squamata (lizards and snakes), and Crocodylia (alligators and crocodiles)—as well as numerous extinct orders. Using recent scientific findings, however, many experts are now giving support to a very different arrangement of the reptile group that asserts that birds have a far closer evolutionary relationship to reptiles than once thought, and in fact should be considered reptiles. A classical taxonomy for the class is given below, followed by the classification scheme for extant reptiles currently accepted by most professional herpetologists. All orders other than the Testudines, Rhycocephalia, Squamata, and Crocodylia are extinct (see separate articles on each taxon listed).

Classical taxonomy:

- Class Reptilia
  - Subclass Anapsida
    - Order Captothinida
      - Mesosauria
      - Testudines (turtles)
  - Subclass Diapsida
    - Infraclass Sauropterygia
      - Order Nothosauria
        - Plesiosauria
        - Placodontia
    - Infraclass Lepidosauria
      - Order Araeoscelida
        - Eosuchia
        - Rhyncocephalia (tuatara)
        - Squamata (lizards and snakes)
    - Infraclass Archosauria
      - Order Protosauria
        - Rhynchosauria
        - Thecodontia
        - Crocodylia (alligators and crocodiles)
        - Pterosauria
        - Saurischia
        - Ornithischia
  - Subclass Ichthyopterygia
  - Subclass Synapsida

Order Pelycosauria
Therapsida

Current classification scheme for extant reptiles:

Subclass Anapsida
Order Testudines (turtles)
Subclass Lepidosauria
Order Rhyncocephalia (tuataras)
Squamata
Suborder Sauria (lizards and skinks)
Amphisbaenia
(amphisbaenids)
Ophidia (snakes)
Subclass Archosauria
Order Crocodylia (alligators and crocodiles)

Globally, there are approximately 8200 extant species of reptiles described (excluding birds), including 300 species of turtles, 2980 species of snakes, and 4760 species of lizards and skinks. Reptiles are found on all continents excluding Antarctica, with their greatest diversity centered in tropical and subtropical ecosystems.

Fundamental characteristics. Reptiles are considered to have certain fundamental characteristics. Most basic is an integument (skin) with well-developed scales. These scales cover most of the body and serve to reduce water loss in the terrestrial environment. The scales are made mostly of keratin, a protein-based substance that is relatively light and durable. The evolution of scales is not thought to have been driven by the need for protection, but was simply an adaptation to prevent rapid water loss through the skin. However, scales quickly began to find utility as excellent dermal protective coverings. Another fundamental characteristic of reptiles is that they are ectothermic, meaning that the predominant amount of their internal body temperature comes from external sources such as sunlight. Reptiles do generate metabolic heat, but most rely on thermoregulatory behaviors to maintain their optimal temperature, such as can be seen in the behavior of the basking turtle. Thermoregulatory behavior also means that reptiles must retreat to refugia when it becomes too hot. Reptile reproduction is as varied as their diversity and forms. Most species are oviparous, laying eggs in the sand, soil, or other materials such as decaying wood, but some species of lizards and snakes are viviparous, giving birth to live young.

Evolution. The evolution of reptiles originates with their diversification from early amphibians. Adaptations of reptiles are rooted in a more general early vertebrate move from a strong reliance on aquatic environments. One of the central events in this process was the evolution of the amniotic egg, hence the Amniota, which includes reptiles, birds, and mammals. The amniotic egg is far more complex than its anamniotic predecessor (represented today by the eggs of fish and modern amphibians). The amniotic egg has three layers: the chorion, allantois, and amnion. These layers are involved in protection, nutrient supply, and waste removal, all important functions when moving away from a dependence on the aquatic environment for the development of the larvae, a predominant condition in all fish and most amphibians today.

The earliest reptiles are best discussed within the context of the earliest amniotes, which date from the mid-Carboniferous period beginning approximately 330 million years before present (mybp). Based on the fossil record to date, there are two important time frames of reptile radiation. The first, from the late Carboniferous (290 mybp) through the Permian period (250 mybp), was an important phase of evolution and radiation of the ancestral groups of species we are familiar with today (turtles, snakes, and lizards), and the mammal-like reptiles (the reptile group that is the ancestral group to modern-day mammals). The second is often referred to as the Age of Reptiles. This period begins in the mid-Triassic (~225 mybp), with the increasing diversity of groups like the Ichthyosaurs, Pterosaurs, Plesiosaurs, Synapsids, Ornithischia, and Saurischia (the latter two orders are commonly referred to as dinosaurs). The success of these groups, along with other groups such as the squamates and testudines, continued until a very abrupt decline (particularly among the dinosaurs) at the end of the Cretaceous Period (65 mybp). The cause of this decline has been a topic of debate for half a century, and two more widely accepted hypotheses have emerged. First, it is proposed that increased volcanic activity globally led to excessive gases and particulate matter entering the Earth's upper atmosphere, thereby causing rapid and drastic climate shifts that affected all organisms via a cascade effect on the food web and caused drastic declines of many, particularly the large-bodied, dinosaurs. The second hypothesis asserts that a large meteor struck the Earth near the tip of the Yucatan Peninsula (often referred to as the Chicxulub Bolide Impact). This would have caused a significant plume of gas and particulate matter to enter the upper atmosphere, with the same result associated with the first hypothesis. In any case, while the outcome was catastrophic for dinosaurs, other groups began to be more successful, among them the evolutionary lines that ultimately led to modern mammals, birds, turtles, snakes, and lizards.

Living reptiles. Extant reptile groups can be distinguished easily by basic characteristics. Lizards, the most species-rich of the living reptile groups, have four limbs (with some interesting exceptions in the case of legless lizards), external ear openings, and eyelids that close. Lizards also have postanal tails that have the ability to break off when grabbed by a predator (tail autotomy).

Conversely, snakes have elongate bodies with no limbs, although members of the group have vestiges of limbs associated with their pelvic girdle. Unlike lizards, snakes have no external ear opening and no eyelids capable of closing. Snakes do, however, have a specialized transparent protective scale (called a spectacle) that covers the eye.

One highly specialized group, the amphisbaenids,

are similar in some respects to lizards and snakes, but have unique characteristics that cause some specialists to consider them a separate suborder of Squamata (Amphibaenia), apart from Sauria and Ophidia. Amphisbaenians have no limbs, with the exception of one family, and reticulate scales giving them a ringed appearance. They are highly adapted burrowers, and their head shape, often blunt or spatulate, belies their subterranean lifestyle. The highest species richness for the group occurs in South America and Africa.

Turtles are distinguished by their external shell, made up of a carapace (dorsal) and a plastron (ventral). This external covering is actually an outgrowth of the underlying vertebrae, complete with integrated ribs for support. The bone is covered by keratinized scutes for further protection. There are species of aquatic turtles, like the smooth softshell turtle (*Apalone mutica*), that have reduced bony shells covered with thick skin instead of keratinized scutes.

Crocodylians have distinctive elongate bodies, long tails and snouts, jaws with well-developed teeth, and bony plates covered with keratinized skin. Crocodylians are typically large in size and found in riverine, swamp, and marine habitats in a tropical and subtropical distribution.

Ecology. Living reptile species as a group are very successful in that they make up a significant proportion of the biodiversity of vertebrates in certain regions, and many have extraordinary adaptations that allow them to inhabit very harsh environments (such as the lizard assemblages of the Australian deserts). In fact, the most successful reptiles are snakes and lizards, making up 94% of the total number of extant reptile species (excluding birds). Reptile species have found ways to exploit many ecological niches. The gopher tortoise (*Gopherus polyphemus*) of the southeastern United States burrows meters underground, creating its own refuge and thereby creating refuge for many other (300+) species of plants and animals. Many species of snakes and amphisbaenids have fossorial or semifossorial lifestyles—that is, they either stay underground entirely or spend a significant portion of time there. Conversely, many snake species have made use of the arboreal niche and spend their time entirely in the canopy of trees.

Although one evolutionary trend in reptiles is a movement away from reliance on water, many have found great success in aquatic habitats. Some snakes, often referred to as sea snakes, are adapted to marine ecosystems, spending time in nearshore areas and estuaries or in coral reef ecosystems. Other reptiles have also demonstrated great success in marine habitats. The sea turtles, represented by two families, Cheloniidae (5 species) and Dermochelyidae (1 species), are found globally in temperate and tropical oceans, and many species migrate long distances to reach nesting beaches. The Galápagos marine iguana (*Amblyrhynchus cristatus*) lives in colonies along the rocky shorelines of the Galápagos Islands and dives into the surf to graze on dense algal mats.

Position of birds and other controversies. The study of the evolutionary course, or phylogeny, of various vertebrate groups has seen an era of change recently. The use of cladistic analysis (grouping organisms by shared common descent), particularly using molecular data (DNA) as a variable, has yielded very different hypotheses regarding the relationships of the vertebrate groups. Recent hypotheses place birds within the reptile clade (a single group of organisms related by a single common ancestor). Although this idea has met resistance, most scientists are now beginning to accept this significant step. This results in a fundamental change in our perspectives of what it means to be reptile. Specifically, scales have always been considered a fundamental characteristic of reptiles, but evolutionary morphologists point out that bird feathers are, in essence, nothing more than highly modified scales. Our view of the reptiles as ectotherms is also being modified and expanded. The metabolic physiology of reptiles is now best described as a continuum, from ectothermic to endothermic (the predominant amount of internal body heat generated by physiological processes). Another proposal that has generated great debate places turtles among the diapsid reptiles alongside squamates, crocodylians and birds. At present this step is a point of continued debate and research. Due to these new hypotheses, the taxonomy of the Reptilia is currently unresolved and in transition.

W. Ben Cash

Bibliography. S. B. Hedges and L. L. Poling, A molecular phylogeny of reptiles, *Science*, 283:998–1001, 1999; M. Laurin and R. R. Reisz, A re-evaluation of early amniote phylogeny, *Zoolog. J. Linnaean Soc.*, 113:105–225, 1995; D. W. Linzey, *Vertebrate Biology*, McGraw-Hill, 2001; F. H. Pough et al., *Herpetology*, 3d ed., Prentice-Hall, Upper Saddle River, New Jersey, 2003; O. Rieppel and R. R. Reisz, The origin and early evolution of turtles, *Ann. Rev. Ecol. Systematics*, 30:1–22, 1999; G. R. Zug, L. J. Vitt, and J. P. Caldwell, *Herpetology*, Academic Press, San Diego, 2001.

Repulsion motor

An alternating-current (ac) commutator motor designed for single-phase operation. The chief distinction between the repulsion motor and the single-phase series motors is the way in which the armature receives its power. In the series motor the armature power is supplied by conduction from the line power supply. In the repulsion motor, however, armature power is supplied by induction (transformer action) from the field of the stator winding. For discussion of the ac series motor *see* UNIVERSAL MOTOR. *See also* ALTERNATING-CURRENT MOTOR.

The repulsion motor primary or stationary field winding is connected to the power supply. The secondary or armature winding is mounted on the motor shaft and rotates with it. The terminals of the armature winding are short-circuited through a commutator and brushes. There is no electrical contact between the stationary field and rotating armature (**Fig. 1**). *See* WINDINGS IN ELECTRIC MACHINERY.

If the motor is at rest and the field coils are energized from an outside ac source, a current is induced in the armature, just as in a static transformer. If the brushes are in line with the neutral axis of the magnetic field, there is no torque, or tendency to rotate. However, if they are set at a proper angle (generally 15–25° from the neutral plane), the motor will rotate.

Repulsion motors may be started with external resistance in series with the motor field, as is done with dc series motors. A more common method is to start the motor with reduced field voltage and increase the voltage as the motor increases speed. This can be done with a transformer having an adjustable tapped secondary or a variable autotransformer.

It is also possible to doubly excite the motor; that is, the armature may receive its power not only by induction from the stator winding but also by conduction from a transformer with adjustable taps, as shown in **Fig. 2**.

Repulsion-start, induction-run motor. This motor (**Fig. 3***a*) possesses the characteristics of the repulsion motor at low speeds and those of the induction motor at high speed. It starts as a repulsion motor. At a predetermined speed (generally at about two-thirds of synchronous speed) a centrifugal device lifts the brushes from the commutator and short-circuits the armature coils, producing a squirrel-cage rotor. The motor then runs as an induction motor. In Fig. 3*b* the curve *AB* represents the characteristics of an induction motor and curve *CD* a repulsion motor. The solid curve *AD* is the combined characteristic of a repulsion-start, induction-run motor. *See* INDUCTION MOTOR.

Repulsion-induction motor. This motor (**Fig. 4***a*) is very similar to the standard repulsion motor in construction, except for the addition of a second separate high-resistance squirrel cage on the rotor. Both rotor windings are torque-producing, and the total torque produced is the sum of the individual torques developed in these two windings. In Fig. 4*b*, curve *A* is the characteristic of the repulsion-motor torque developed in this motor. Curve *B* represents the induction-motor torque. Curve *C* is the combined total torque of the motor. The advantages of this machine are its high starting torque and good speed regulation. Its disadvantages are its poor commutation and high initial cost. Irving L. Kosow

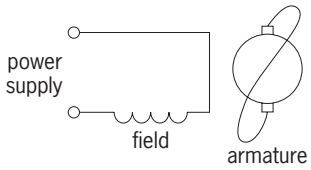

Fig. 1. Schematic diagram of a repulsion motor.

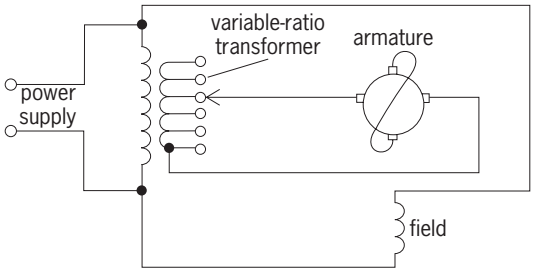

Fig. 2. Doubly excited repulsion motor.

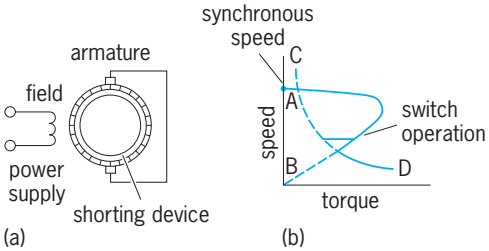

Fig. 3. Repulsion-start, induction-run motor. (*a*) Schematic diagram. (*b*) Speed-torque-characteristic.

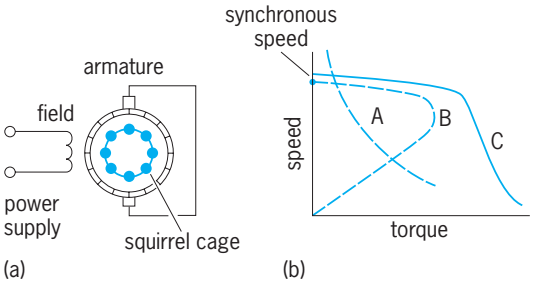

Fig. 4. Repulsion-induction motor. (*a*) Schematic diagram. (*b*) Speed-torque characteristic.

Bibliography. I. L. Kosow, *Electric Machinery and Transformers*, 2d ed., 1991; G. McPherson and R. D. Larramore, *An Introduction to Electrical Machines and Transformers*, 2d ed., 1990; M. G. Say, *Alternating-Current Machines*, 5th ed., 1984.

Reservoir

A place or containment area where water is stored. Where large volumes of water are to be stored, reservoirs usually are created by the construction of a dam across a flowing stream. When water occurs naturally in streams, it is sometimes not available when needed. Reservoirs solve this problem by capturing water and making it available at later times. *See* DAM.

In addition to large reservoirs, many small reservoirs are in service. These include varieties of farm ponds, regulating lakes, and small industrial or recreational facilities. In some regions, small ponds are called tanks. Small reservoirs can have important cumulative effects in rural regions.

Purposes. Reservoirs can be developed for single or multiple purposes, such as to supply water for people and cities, to provide irrigation water, to lift water levels to make navigation possible on streams, and to generate electricity.

When used to supply water for a city, a reservoir can provide water during all parts of the year, and can even make it possible to populate dry regions. If cities have underground water available, they may rely on it; otherwise, they must have stored surface water from a reservoir, or they risk running out of drinking water.

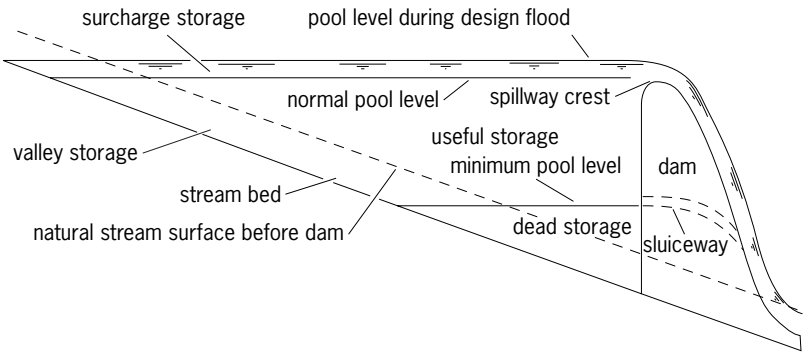

Fig. 1. Zones of storage in a reservoir. (*After R. K. Linsley et al., Water-Resources Engineering, McGraw-Hill, 1992*)

A reservoir can supply water for irrigation of crops at precisely the time required for the maximum plant growth. Many reservoirs in dry regions capture water during wet periods and release it to farm fields at times of plant growth and stress.

Reservoirs may provide for energy generation (hydropower) through run-of-river hydroelectric plants which generate power when the water is available; or if the storage is adequate, they may provide water for energy generation at any time it is needed. Another way that reservoirs aid in energy management is through pumped storage, where electric energy is stored by pumping water to a reservoir, and releasing the water later to flow downward through a generating system. Water from reservoirs is also used to cool thermal or nuclear electricity-generating plants. *See* PUMPED STORAGE.

Reservoirs also make navigation possible by maintaining river depths at required levels, usually by means of locks and dams. *See* RIVER ENGINEERING.

Another purpose of reservoirs is to control floods by providing empty spaces for flood waters to fill, thereby diminishing the rate of flow and water depth downstream of the reservoir. Such use of reservoirs has saved many lives and goods from flood damage, but reservoirs cannot control all flood problems.

Reservoirs also provide for environmental uses of water by providing water to sustain fisheries and meet other fish and wildlife needs, or to improve water quality by providing dilution water when it is needed in downstream sections of rivers. Reservoirs may also have esthetic and recreational value, providing boating, swimming, fishing, rafting, hiking, viewing, photography, and general enjoyment of nature.

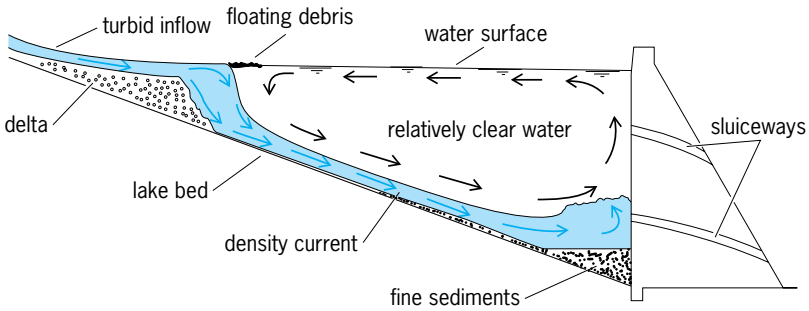

Fig. 2. Sediment accumulation in a typical reservoir. (*After R. K. Linsley et al., Water-Resources Engineering, McGraw-Hill, 1992*)

Characteristics and configuration. Generally, a reservoir is divided into zones of water that are reserved for different uses (**Fig. 1**). The zone for useful storage (also called working storage, multiple-purpose capacity, or operating storage) provides space for water storage applications. In this zone it is normal for the water level to rise and fall during the year. Sometimes people who rely on reservoirs only for recreational use complain about this, because they do not understand the necessity of using the full storage zone.

The surcharge storage zone enables the reservoir to be used for flood control. If flood control is an important purpose, part of the useful storage could also be devoted to that use. Surcharge storage occurs above the spillway crest, which is the elevation where water just begins to flow over the spillway's upper surface. The dam is designed so that the design flood, a maximum flood event, safely passes over the spillway without placing the dam at risk. The dead storage, normally reserved to be filled with sediment, occurs below the normal release elevation for water.

Planning. Planning for reservoir size takes into account the statistical variation of future inflows and water demands.

The reservoir sizing process carries the risk that the capacity will be too small to meet the purposes or too large for the reservoir to fill. Therefore, the planner compiles as much historical data as possible and makes studies of how the planned reservoir would have performed if it had been in place during the historical period. These are sometimes referred to as "what if" studies. Reservoirs are complex, however, because the water supplied by the stream system is not fully predictable and is subject to variations from high to low flow periods. Thus, the water reservoir is made large enough to compensate for the risk factors.

After the location of a reservoir is determined, many aspects of the construction process must be settled through the design process, an effort that involves engineers, geologists, hydrologists, and other professionals. A major consideration is the necessity to make the dam as safe as possible to avoid placing people downstream at risk due to dam failure. While there have been few dam failures, when they do occur they may cause unacceptable levels of damage.

After a dam is built, it must be operated correctly. The key person is the operator who makes decisions about when to release or store water. In the past, reservoir operating decisions were made by rule curves, which provided the operator with simple guidelines about how much water to release and what lake levels to maintain. As the science of forecasting and the use of computers became more complex, however, reservoir operation became more sophisticated. It is not uncommon to have a reservoir control center where operators use computers to monitor weather forecasts furnished from satellite data and to simulate future demands for water

in order to make decisions about water releases. They may also be bound by legal requirements to release water for downstream users, including fish and wildlife. As the demands for water increase in developed areas, conflicts over use of the water arise, especially in water-short areas. This requires the reservoir owner or management agency to develop conflict-resolution procedures.

Reservoirs require continuing attention from owners and managers. Problems may include excessive deposition of sediments on the lake bottom which reduce the capacity to store water (**Fig. 2**), pollution of lake waters, eutrophication (aging of waters with excessive growth of algae), shoreline protection, and associated issues such as dam leakage or settlement. Sedimentation, a naturally occurring phenomenon, is a particularly difficult problem that is impossible to fully control, and it may ruin a reservoir's capacity to store water. *See* EUTROPHICATION; WATER POLLUTION.

Controversies. The construction or alteration of reservoirs is often resisted by some people because of perceived negative side effects. One side effect is reservoir evaporation, which can consume waters that could otherwise be used in the river environment for fish and wildlife or for flushing salts through the stream system. As a result of evaporation, the microclimate around the reservoir might even be altered. Seepage caused by the reservoir impoundment can change the local patterns of underground water. In the valley floor, settlement can occur because of the weight of the water in storage. Ecosystems can be changed because the schedule of water release and the quality of the water can be altered. Also, when reservoirs are built, large numbers of residents may have to be resettled. *See* WATER SUPPLY ENGINEERING.

Neil S. Grigg

Bibliography. N. S. Grigg, *Water Resources Management*, 1996; R. K. Linsley et al., *Water-Resources Engineering*, 1992; D. R. Maidment (ed.), *Handbook of Hydrology*, 1993.

Resin

Originally a category of vegetable substances soluble in ethanol but insoluble in water, but generally in modern technology an organic polymer of indeterminate molecular weight. The class of flammable, amorphous secretions of conifers or legumes are considered true resins, and include kauri, copal, dammar, mastic, guaiacum, jalap, colophony, shellac, and numerous less well-known substances. Water-swellable secretions of various plants, especially the Burseraceae, are called gum resins and include myrrh and olibanum. The official resins, according to the *United States Pharmacopeia*, are benzoin, guaiac, mastic, and resin. Other natural resins are copal, dammar, dragon's blood, elaterium, lac, and sandarac. The natural vegetable resins are largely polyterpenes and their acid derivatives, which find application in the manufacture of lacquers, adhesives, varnishes, and inks.

The synthetic resins, originally viewed as substitutes for copal, dammar, and elemi natural resins, have a large place of their own in industry and commerce. Phenol-formaldehyde, phenol-urea, and phenol-melamine resins have been important commercially for a long time. Any unplasticized organic polymer is considered a resin, thus nearly any of the common plastics may be viewed as a synthetic resin. Water-soluble resins are marketed chiefly as substitutes for vegetable gums and in their own right for highly specialized applications. Carboxymethylcellulose, hydroxyalkylated cellulose derivatives, modified starches, polyvinyl alcohol, polyvinylpyrrolindone, and polyacrylamides are used as thickening agents for foods, paints, and drilling muds, as fiber sizings, in various kinds of protective coatings, and as encapsulating substances. *See* POLYMER.

Frank Wagner

Resistance heating

The generation of heat by electric conductors carrying current. The degree of heating for a given current is proportional to the electrical resistance of the conductor. If the resistance is high, a large amount of heat is generated, and the material is used as a resistor rather than as a conductor. *See* ELECTRICAL RESISTANCE.

Resistor materials. In addition to having high resistivity, heating elements must be able to withstand high temperatures without deteriorating or sagging. Other desirable characteristics are low temperature coefficient of resistance, low cost, formability, and availability of materials. Most commercial resistance alloys contain chromium or aluminum or both, since a protective coating of chrome oxide or aluminum oxide forms on the surface upon heating and inhibits or retards further oxidation. Some commercial resistor materials are listed in **Table 1**.

Heating element forms. Since heat is transmitted by radiation, convection, or conduction or combinations of these, the form of element is designed for the major mode of transmission. The simplest form is the helix, using a round wire resistor, with the pitch of the helix approximately three wire diameters. This form is adapted to radiation and convection and is generally used for room or air heating. It is also used in industrial furnaces, utilizing forced convection up to about 1200°F (650°C). Such helixes are stretched over grooved high-alumina refractory insulators and are otherwise open and unrestricted. These helixes are suitable for mounting in air ducts or enclosed chambers, where there is no danger of human contact. *See* ELECTRIC FURNACE; INDUCTION HEATING.

For such applications as water heating, electric range units, and die heating, where complete electrical isolation is necessary, the helix is embedded in magnesium oxide inside a metal tube, after which the tube is swaged to a smaller diameter to compact the oxide and increase its thermal conductivity. Such units can then be formed and flattened to desired shapes. The metal tubing is usually copper for water

TABLE 1. Electric furnace resistor materials and temperature ranges

Material, major elements	Maximum resistor temperature, °F (°C)	
	In air	In reducing or neutral atmosphere
35% nickel, 20% chromium*	1900 (1040)	2100 (1150)
60% nickel, 16% chromium*	1800 (980)	2000 (1090)
68% nickel, 20% chromium, 1% cobalt*	2250 (1230)	2250 (1230)
78% nickel, 20% chromium	2250 (1230)	2250 (1230)
15% chromium, 4.6% aluminum*	2050 (1120)	Not used
22.5% chromium, 4.6% aluminum*	2150 (1180)	Not used
22.5% chromium, 5.5% aluminum*	2450 (1340)	Not used
Silicon carbide	2800 (1540)	2500 (1370)
Platinum	2900 (1590)	2900 (1590)
Molybdenum†	Not used	3400 (1870)
Tungsten†	Not used	3700 (2040)
Graphite†	Not used	5000 (2760)

*Balance is largely iron, with 0.5–1.5% silicon.
†Usable only in pure hydrogen, nitrogen, helium, or oxygen or in vacuum because of inability to form protective oxide.

heaters and stainless steel for radiant elements, such as range units. In some cases the tubes may be cast into finned aluminum housings, or fins may be brazed directly to the tubing to increase surface area for convection heating.

Modification of the helix for high-temperature furnaces involves supporting each turn in a grooved refractory insulator, the insulators being strung on stainless alloy rods. Wire sizes for such elements are $^3/_{16}$ in. (5 mm) in diameter or larger, or they may be edge-wound strap. Such elements may be used up to 1800°F (980°C) furnace temperature. *See* REFRACTORY.

Another form of furnace heating element is the sinuous grid element, made of heavy wire or strap or casting and suspended from refractory or stainless supports built into the furnace walls, floor, and roof. Some of these forms are shown in **Fig. 1**.

Silicon carbide elements are in rod form, with low-resistance integral terminals extending through the furnace walls, as shown in **Fig. 2**.

Direct heating. When heating metal strip or wire continuously, the supporting rolls can be used as electrodes, and the strip or wire can be used as the resistor.

In **Fig. 3** the electric current passes from roll A through the strip or wire to roll B. Heating by this method can be very rapid. Disadvantages are that the electric currents are large, and uniform contact between the strip and the rolls is difficult to

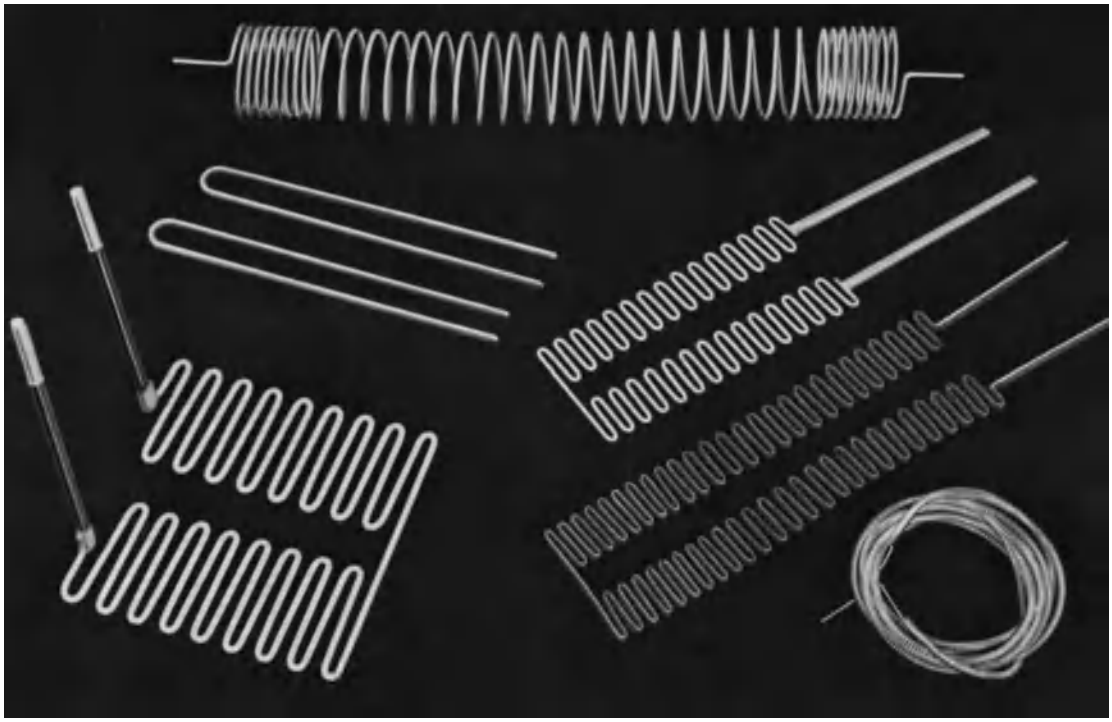

Fig. 1. Metallic resistor forms.

maintain, since both surfaces must be clean and free of oxides. For these reasons, direct heating is not used extensively.

Molten salts. The electrical resistance of molten salts between immersed electrodes can be used to generate heat. Limiting temperatures are dependent on decomposition or evaporization temperatures of the salt. Parts to be heated are immersed in the salt. Heating is rapid and, since there is no exposure to air, oxidation is largely prevented. Disadvantages are the personnel hazards and discomfort of working close to molten salts.

Major applications. A major application of resistance heating is in electric home appliances, including electric ranges, clothes dryers, water heaters, coffee percolators, portable radiant heaters, and hair dryers. Resistance heating has also found application in home or space heating; some homes are designed with suitable thermal insulation to make electric heating practicable. A general rule for such insulation is that heat loss be restricted to 10 W/ft^2 (108 W/m^2) of floor living area at 75°F (42°C) temperature difference, inside to outside. This applies to areas having 5000–7500 degree days annually. The degree day is the difference between the 24-h average outdoor temperature and 65°F (18.3°C). For instance, if the 24-h average were 40°F (4.4°C), there would be 25 degree days for that day. This includes also a three-quarter air change per hour, except in the basement, where one-quarter air change per hour is calculated. The values in **Table 2** are applied to determine overall heat loss. *See* DEGREE-DAY.

Heating system capacity. For rooms normally kept warm, the installed capacity should be 20% higher than the calculated losses. For rooms intermittently heated or for entries with frequent door openings, the installed capacity should be 50% more than the calculated losses.

A variety of electric heating units may be used for home heating. The simplest is a grid of helical resistance coils mounted in the central heating unit, with a blower to circulate the warm air through the rooms, and adjustable louvers in each room to control the temperature. Another type is the baseboard heater, consisting of finned sheathed helical resistance coils, with individual room control (thermostat in each room). Still another room unit is the glass or ceramic panel in which resistance wires are embedded. Such panels are designed to radiate infrared wavelengths and are particularly suitable for glassed-in areas or bathrooms. Mounted usually in the ceiling, such panels neutralize the chilling effect of large glass exposures, thus preventing drafts. *See* THERMOSTAT.

TABLE 2. Maximum recommended heat loss

Structure	Heat loss, W/(ft^2)(°F) [W/(m^2)(°C)]
Ceiling	0.015 (0.29)
Wall	0.021 (0.41)
Floor	0.021 (0.41)
Floor slab perimeter	12.0/lineal foot (71/lineal meter)
Basement walls	0.045 (0.77)
Basement walls, 25% of wall aboveground	0.026 (0.50)

Requisites:
Windows—double-glazed or storm
Doors—storm doors required
Attic—ventilating area of 1 ft^2/150 ft^2 or 1 m^2/150 m^2 ceiling area

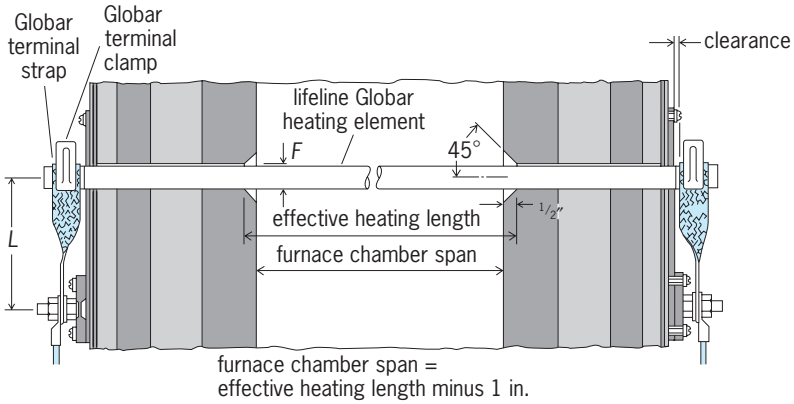

Fig. 2. Silicon carbide heating element. 1 in. = 2.5 cm.

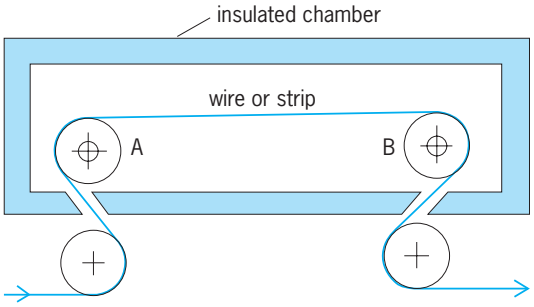

Fig. 3. Direct heating of metal strip or wire.

For isolated rooms or work areas, integrated units which combine resistors, circulating fans, and magnetic contactors operated by thermostats open and close the power circuits to the heating elements to maintain desired temperature. Where central heating units involving several kilowatts of energy are used, the contactors are arranged in multiple circuits and close in sequence to minimize the transient voltage effects of sudden large energy demands.

Ovens and furnaces. If the resistor is located in a thermally insulated chamber, most of the heat generated is conserved and can be applied to a wide variety of heating processes. Such insulated chambers are called ovens or furnaces, depending on the temperature range and use.

The term oven is generally applied to units which operate up to approximately 800°F (430°C). Ovens use rock wool or glass wool between the inner and outer steel casings for thermal insulation. Typical uses are for baking or roasting foods, drying paints and organic enamels, baking foundry cores, and low-temperature treatments of metals.

The term furnace generally applies to units operating above 1200°F (650°C). In these the thermal

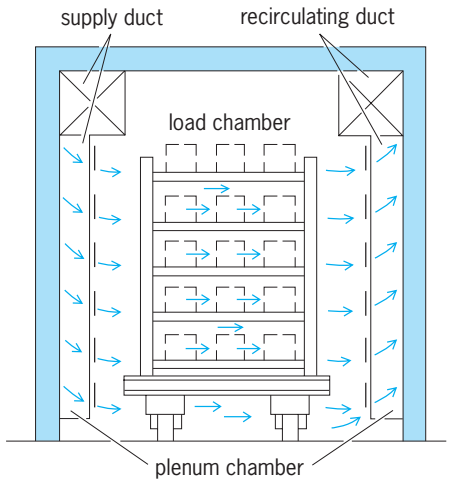

Fig. 4. Typical oven for maximum temperature of 800°F (430°C); heating is by forced convection. (*Carl Mayer Corp.*)

insulation is made up of an inner wall of fireclay, kaolin, or high-alumina or zirconia brick, depending on the temperature, with secondary insulating blocks made from such base materials as rock wool, asbestos fiber, and diatomaceous earth. Typical uses of furnaces are for heat treatment or melting of metals, for vitrification and glazing of ceramic wares, for annealing of glass, and for roasting and calcining of ores. *See* ELECTRIC HEATING.

Actually, ovens and furnaces overlap in temperature range, with ovens being used at temperatures as high as 1000°F (540°C), and furnaces as low as 250°F (120°C). Electrically heated ovens and furnaces have advantages over fuel-fired units. These advantages often compensate for the generally higher cost of electric energy. The main advantages are (1) ease of distributing resistors, or heating elements, to obtain a uniform temperature in the product being heated; (2) ease of operation, since adjustments by operators are usually unnecessary; (3) cleanliness; (4) comfort, since heat losses are low and there are no waste fuel products; (5) adaptability to the use of controlled furnace atmospheres or vacuum, and (6) high temperatures beyond the range attainable with commercial fuels.

Heat transfer in ovens and furnaces. Heating elements mounted in ovens and furnaces may be located to radiate directly to the parts being heated or may be located behind baffles or walls so that direct radiation cannot take place. The heat then is transferred by circulating the furnace air or gas. Determination of which method or combination of methods to use is based on temperature uniformity, speed of heating, and high-temperature strength limitations of fans or blowers. At temperatures below 1200°F (650°C) radiation is slow, and virtually all ovens and furnaces in this temperature range use forced convection or circulation. From 1200 to 1500°F (650 to 820°C), radiation is increasingly effective, while reduced gas density and lower fan or blower speeds (because of reduced strength at high temperatures) make forced convection less effective. Therefore in this temperature range, combinations of radiation and forced convection are used. Above 1500°F forced convection is employed only when direct radiation cannot reach all parts of the loads being heated. An example of this is a container filled with bolts or a rack filled with gears.

Obviously, in vacuum furnaces, heat transfer can be only by radiation and conduction.

Ovens. **Figure 4** shows a typical oven, for a maximum temperature of 800°F (430°C); heating is by forced convection. The heated air enters through the supply duct, passing down into the plenum chamber and distributing louvers into and across the load chamber and returning through the right-side plenum to the recirculating duct. The external heater and blower, completing the circuit, are shown.

Low-temperature furnace. **Figure 5** shows a pit-type furnace used for tempering steel. It operates at temperatures of 250–1400°F (120–760°C) and uses forced convection.

The steel parts to be tempered are placed in the load chamber at the right. The electrical heating elements are in the heating chamber at the upper left. The fan, at the lower left, circulates air upward over the heating elements and into the load chamber. The hot air is forced down through the load and back to

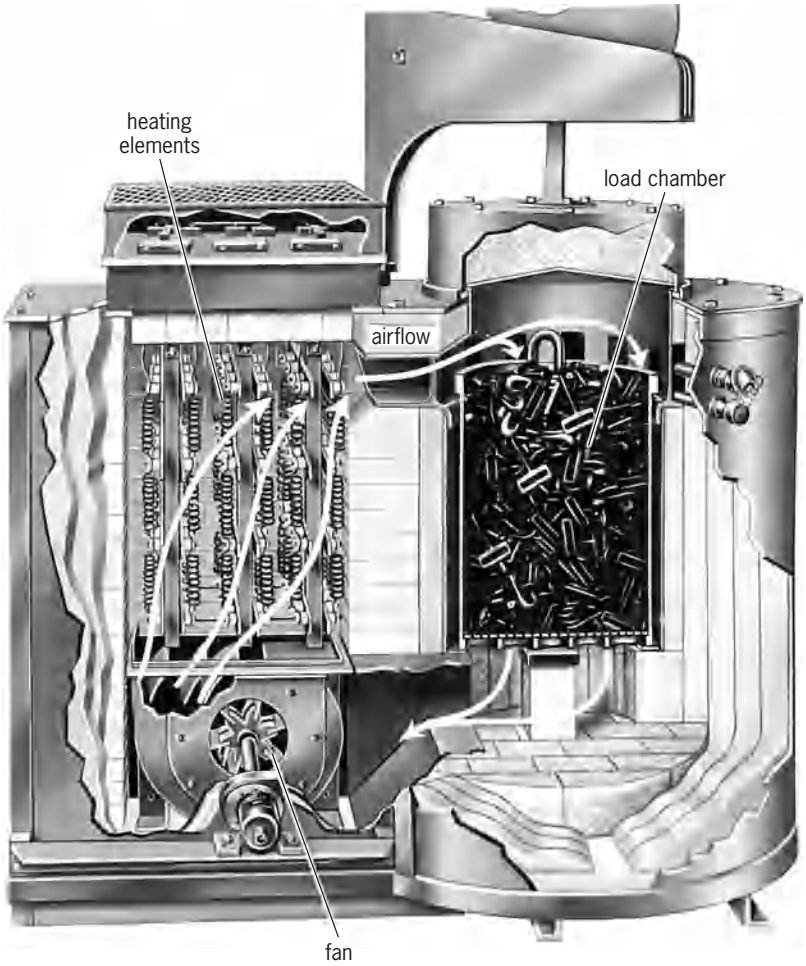

Fig. 5. Low-temperature pit-type furnace, using forced convection and operating at temperatures of 250–1400°F (120–760°C). (*Lindberg Industrial Corp.*)

the fan. The heating chamber is thermally insulated from the load chamber so there is no direct radiation. Such furnaces are designed commercially to hold the load temperature within 10° of the control temperature, and can be designed for closer control if desired.

High-temperature furnace. **Figure 6** shows the interior of a pit-type carburizing furnace, which uses radiant-heating elements in the form of corrugated metal bands mounted on the inside of the brick walls. The work basket (not shown) rests on the load support at the bottom. These elements are designed to operate at low voltage (approximately 30 V) so that soot deposits from the carburizing gases will not cause short circuits. *See* FURNACE CONSTRUCTION.

Electrical input. Electric ovens and furnaces are rated in kilowatts. The electrical input is determined from the energy absorption Q of the load and the thermal losses.

The average rate of energy flow into load during the heating period is Q/t, where Q is in kilowatt-hours and t is the heating time in hours.

However, the rate of energy flow into the load is high when the load is cold and decreases as the load temperature approaches furnace temperature. Therefore the energy input must be high enough to take care of the high initial heating rate. This leads to the following approximate formulas for the input, in which L is the thermal losses in kilowatts at the operating temperature:

$$\text{Batch furnaces: input in kW} = L + 1.5\ Q/t$$

$$\text{Continuous furnaces: input in kW} = L + 1.25\ Q/t$$

Operating voltage for heating elements. Heating elements are usually designed to operate at standard service voltages of 115, 230, or 460 V, if two conditions can be satisfied. These are, first, that the heating element is sufficiently heavy in cross section to avoid sagging or deformation in service, and second, that there is no appreciable electrical leakage through the furnace refractories tending to short-circuit the heating elements. The latter consideration generally limits voltages to 260 volts at 2100°F (1150°C) and to approximately 50 volts at 3200°F (1700°C), because the refractory walls of the furnace become increasingly better conductors at higher temperatures.

In vacuum furnaces, voltages are limited to 230 V by the tendency to break down into glow discharge at low pressure.

Temperature control. Almost all commercial electric ovens and furnaces have automatic temperature control. The simplest control uses magnetic contactors which open and close the circuit to the heating elements in response to temperature signals from control thermostats or thermoelectric pyrometers. A refinement of this ON-and-OFF control is to modulate it with a timer, with the ON period becoming a progressively smaller percentage as the control temperature is approached. This prevents the overshooting which results from thermal lag in the control thermocouple and the furnace.

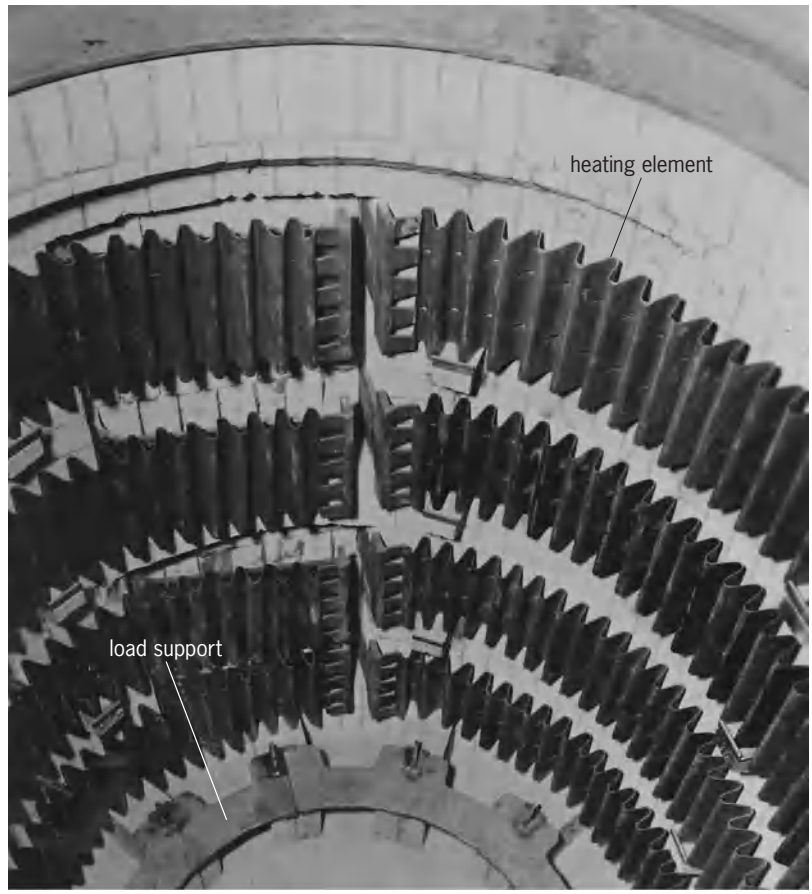

Fig. 6. Interior of high-temperature pit-type carburizing furnace, using radiant heating. (*Lindberg Industrial Corp.*)

True proportioning control is achieved through the use of saturable reactors in series with the heating elements. **Figure 7** shows the arrangement used.

The pyrometer, through the amplifier, controls the direct current to the control winding of the reactor. When the direct current is maximum, the reactor offers virtually no impedance to the alternating current flowing through the heating element, and the normal amount of heat is generated. As control temperature is approached, the direct current is decreased, increasing the impedance of the reactor and reducing the current to the heating elements until equilibrium is reached.

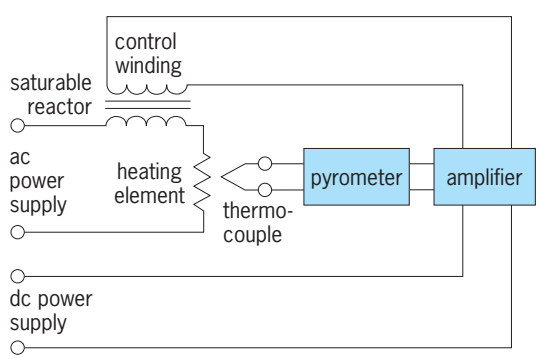

Fig. 7. Temperature control using saturable reactor.

The solid-state silicon controlled rectifier is now used as a heating contactor. Compared to magnetic contactor, the SCR has no moving parts, lower maintenance costs, and silent, vibration-free operation. Moreover, it is directly adaptable to proportioning control, obviating the need for saturable reactors. In this latter case it is cost-competitive. *See* SEMICONDUCTOR RECTIFIER.

Snow and ice melting. Low-temperature resistors with moisture-proof insulation are embedded in concrete sidewalks or strung in roof gutters to melt snow and ice as needed. Such cables also are used to prevent freezing and exposed water pipes. Analogously, very thin resistance wires may be cast in automobile windows to melt snow and ice.

Metal-sheathed units are used at railroad switches and dam locks to keep the moving parts free in freezing temperatures.

Die and platen heating. Metal-sheathed elements are embedded in grooves or holes in dies or platens to maintain these parts at desired temperatures for hot processing or molding of plastics.

Cryogenics. Metal-sheathed immersion-type units are used to supply heat of evaporation for liquefied gases such as argon, helium, nitrogen, hydrogen, and oxygen. Metal sheathing is usually of stainless steel or aluminum to withstand low temperatures. Gaskets are of poly(tetrafluoroethylene) for the same reason.

Willard Roth

Bibliography. Chemical Rubber Publishing Co., *Handbook of Chemistry and Physics*, annually; C. J. Erikson, *Handbook of Electrical Heating forIndustry*, IEEE 1994; D. G. Fink and H. W. Beatty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; B. Parsons (ed.), *ASHRAE Handbook: Heating, Ventilating, and Air-Conditioning Systems and Equipment*, 1992.

Resistance measurement

The quantitative determination of that property of an electrically conductive material, component, or circuit called electrical resistance. The ohm, which is the International System (SI) unit of resistance, is defined through the application of Ohm's law as the electric resistance between two points of a conductor when a constant potential difference of 1 volt applied to these points produces in the conductor a current of 1 ampere. Ohm's law can thus be taken to define resistance R as the ratio of dc voltage V to current I, Eq. (1). For bulk metallic conductors, for ex-

$$R = \frac{V}{I} \tag{1}$$

ample, bars, sheets, wires, and foils, this ratio is constant. For most other substances, such as semiconductors, ceramics, and composite materials, it may vary with voltage, and many electronic devices depend on this fact. The resistance of any conductor is given by the integral of expression (2), where l is the length, A the cross-sectional area, and ρ the resistiv-

$$\int_0^l \frac{\rho\, dl}{A} \tag{2}$$

ity. *See* ELECTRICAL CONDUCTIVITY OF METALS; ELECTRICAL RESISTANCE; ELECTRICAL RESISTIVITY; OHM'S LAW; SEMICONDUCTOR.

Realization of the ohm. The ohm can be realized by means of a calculable capacitor and a series of ac and dc bridges. Following international agreement, the ohm has been maintained since January 1, 1990, with reference to a quantized Hall resistance.

Calculable capacitor. The calculable capacitor enables the ohm to be realized independently of voltage and current. The dielectric constant of free space ϵ_0 is given by Eq. (3), where $\mu_0 = 4\pi \times 10^{-7}$ henry per

$$\epsilon_0 = \frac{1}{c^2 \mu_0} \tag{3}$$

meter (a consequence of the definition of the ampere), and c is the velocity of light, which has been declared by the International Committee for Weights and Measures (CIPM) to be 299,792,458 m/s. Given this value, the capacitance in farads of any assembly of conductors can be calculated through the use of electromagnetic theory. The Thomson-Lampard theorem in electrostatics gives the result that the capacitance per unit length between the opposite members of cylinders arranged in a symmetrical quadrupole (**Fig. 1***a*) is given by Eq. (4), a result which is

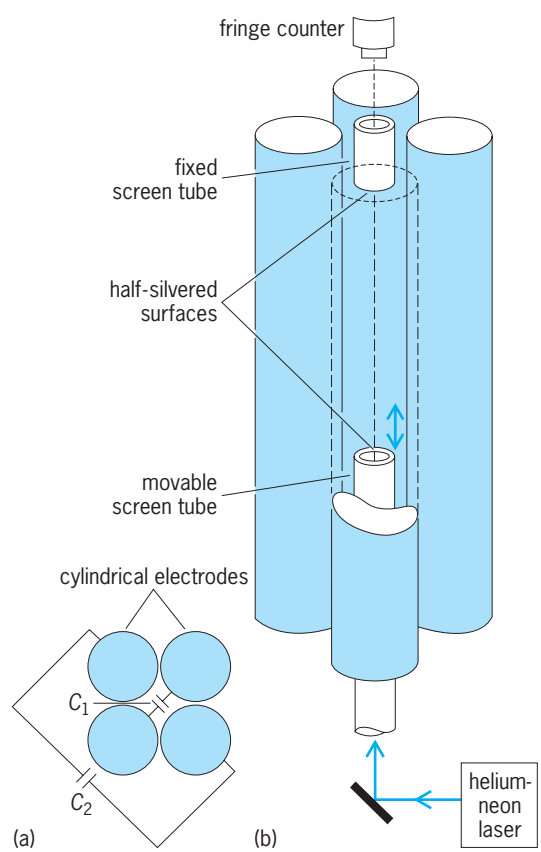

Fig. 1. Thomson-Lampard calculable capacitor. (*a*) Cross section of electrodes. (*b*) Arrangement of components.

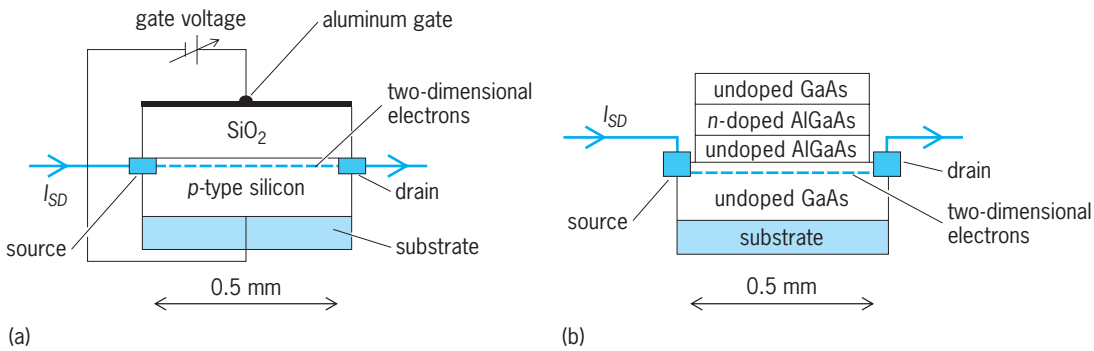

Fig. 2. Two types of semiconductor devices having quasi-two-dimensional electron layers. I_{SD} = current between source and drain electrodes. (*a*) Silicon MOSFET. (*b*) Heterostructure. (*After A. Hartland, Quantum standards for electrical units, Contemp. Phys., 29:477–498, 1988*)

independent of constant cross-sectional dimensions

$$C = \frac{\epsilon_0}{\pi}(\log_e 2) \qquad \text{farads/meter} \qquad (4)$$

of the apparatus. This result is used to construct a calculable capacitor, in which the active length is defined by two central screen tubes, one of which can be moved along the axis. The change in the capacitance between either of the opposite pairs of conducting cylinders is then given by Eq. (3). The change in effective length is measured by arranging a Fabry-Perot interferometer between the faces of the screen tubes (Fig. 1*b*).

Several such instruments exist, and produce a change in capacitance of about 0.4 picofarad with an uncertainty of about 3 parts in 10^8. Through the use of accurate transformer bridges, this change in capacitance is used to establish the value of a 1-nanofarad capacitor, and the reactance of this at a frequency of 1592 Hz (10,000 radians/s) is then used to establish the resistance value of 100 kilohms at the same frequency. Again using transformer bridges, the value of 1000 ohms (Ω) is obtained, and this in turn is used to establish a 1000-Ω dc value by using a coaxial resistor having a single thin-wire element which permits accurate calculation of the difference between the resistance values at 1592 Hz and dc. The final step from 1000 Ω to 1 Ω is made by dc comparison techniques, for example, through the use of a series-parallel Hamon network of 32 × 32 Ω resistors, discussed below, or a cryogenic current comparator. The overall uncertainties of this type of experiment are such that, for example, the realizations of the ohm in the United States and the United Kingdom are in agreement to within 3 parts in 10^7. *See* BRIDGE CIRCUIT; CAPACITANCE MEASUREMENT; CURRENT COMPARATOR; ELECTRICAL UNITS AND STANDARDS; INTERFEROMETRY; RADIO-FREQUENCY IMPEDANCE MEASUREMENTS; RESISTOR.

Quantized Hall resistance. The quantum Hall effect occurs at the interface between two suitable semiconducting layers at low temperatures (less than 4.2 K or −452°F). Under these conditions, a narrow potential well is formed and a two-dimensional layer of electrons is confined at the interface. Although the effect was first discovered in silicon metal-oxide-semiconductor field-effect transistor (MOSFET) devices, the most commonly used system is a gallium arsenide–aluminium gallium arsenide heterostructure (**Fig. 2**). The inset in **Fig. 3** shows a typical device configuration. Any of the three electrode pairs, 1–4, 2–5, and 3–6, permit measurement of the transverse Hall voltage, and the Hall resistance R_H is defined as the ratio of this voltage to the current between the source electrode and the drain electrode. If a current of less than 100 microamperes flows between the source and drain electrodes, the Hall voltage and hence the Hall resistance when measured as a function of magnetic field exhibit plateaus (Fig. 3). These plateaus occur at values of the Hall resistance given by Eq. (5), where i is an integer, and

$$R_H = \frac{R_K}{i} \qquad (5)$$

R_K is the von Klitzing constant and is believed to be equal to h/e^2, h being the Planck constant and e the elementary charge. The International Committee for

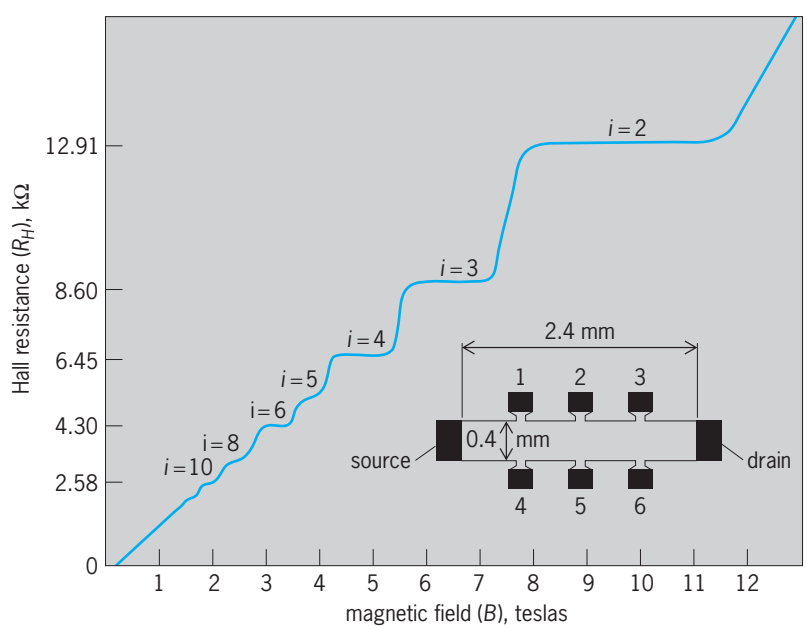

Fig. 3. Quantized Hall resistance in a gallium arsenide–aluminum gallium arsenide heterostructure at a temperature of 0.33 K. Inset is a plan view of the device. (*After A. Hartland, Quantum standards for electrical units, Contemp. Phys., 29:477–498, 1988*)

Weights and Measures has recommended that R_K be assigned the value 25,812.807 $(1 \pm 0.2 \times 10^{-6})$ Ω (one standard deviation uncertainty). This value was obtained mainly from measurements referred to a calculable capacitor and partly from measurements of fundamental constants such as the fine-structure constant. *See* FUNDAMENTAL CONSTANTS; SEMICONDUCTOR HETEROSTRUCTURES; TRANSISTOR.

Since January 1, 1990, all resistance measurements worldwide have been referred to the quantized Hall resistance standard, which is used to maintain the ohm in all national standards laboratories. Conventional wire-wound working standards are measured in terms of the quantized Hall resistance and then used to disseminate the ohm through the normal calibration chain. These working standards can be measured in terms of the quantized Hall resistance with a one-standard-deviation uncertainty of about 1 part in 10^8. *See* HALL EFFECT.

Resistance standards. Various resistance standards are in common use.

Wire-wound standard resistors. The day-to-day primary reference standards of national laboratories are 1-Ω resistors wound either from a quaternary alloy, which is a nickel-chrome alloy with small amounts of other metals added to produce better stability and reduced variation with temperature, or of manganin, which is a copper-nickel-manganese alloy. Both materials have temperature coefficients of zero at some temperature close to 68°F (20°C). The wire is best wound and located in a strain-free manner, and is immersed in oil to assist temperature uniformity. In use, the power dissipated is normally kept below a few milliwatts. The best resistors of this type have been found to change by less than 1 part in 10^7 per year. Basically similar construction is used for standard resistors at decade values from 1 milliohm to at least 100 kilohms, and at 25 Ω, a value at which a standard is needed as the basis for platinum resistance thermometry. These are all four-terminal resistors (**Fig. 4**); the ends of the resistive element are connected by pairs of leads, one of which is used to conduct the current, while the other carries only the much smaller current needed by the detector.

The resistive elements are often wound in a bifilar or hairpin fashion to reduce the self-inductance. All resistors necessarily have some inductance and capacitance, and can be simply represented by the

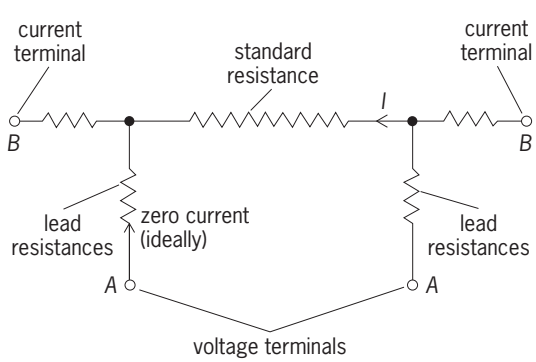

Fig. 4. Four-terminal resistor. Resistance = V_{AA}/I.

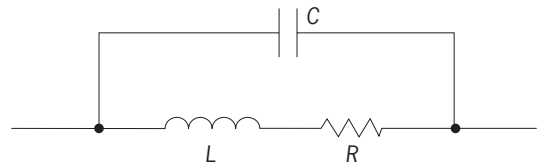

Fig. 5. Equivalent circuit of two-terminal resistor with resistance *R*. The capacitor represents the stray capacitance *C* of the resistor, and the inductor represents its stray inductance *L*.

circuit of **Fig. 5**. To first order, the impedance Z is given by Eq. (6), where C is the stray capacitance, L is

$$Z = R(1 - \omega^2 C^2 R^2) + j\omega(L - R^2 C) \qquad (6)$$

the stray inductance, and ω is the angular frequency. For resistors below about 300 Ω the inductive effect predominates, while resistors of more than this value are essentially capacitative. *See* ALTERNATING-CURRENT CIRCUIT THEORY.

Standard resistors below 100 kΩ in value are usually measured in thermostatic oil baths, permitting temperature stability to within a few millidegrees, and under these conditions it is possible to achieve repeatability of a few parts in 10^8. Higher values of resistance, above 1 megohm, cannot be oil-immersed because the conductance of the oil becomes significant. These are usually loosely wound on bobbins, and are two-terminal resistors. They may be fitted with a guard, which can be used to reduce the effects of leakage or stray conductance. Such resistors extend as far as 10^7 Ω. Beyond that level they become prohibitively large and expensive because of the quantity of wire involved.

Thick-film resistors. These resistors consist of elements of dimensions of only a fraction of an inch (a few millimeters), in which thick nichrome films are deposited on glass or ceramic substrates. These provide good stability with temperature by acting as compensated strain gages, in which the increase of resistivity due to change in temperature is compensated by the effect of strain produced by the differential expansion between the metal and the substrate. The chips are made and adjusted by printed-circuit techniques, and assemblies of very good overall performance can be obtained by selection of suitable elements. They have zero voltage coefficients, and overall temperature coefficients which may be as good as a few parts per million per degree Celsius. Their stability with time is also in the parts-per-million range. Because of their size, they have low inductance and capacitance, and so can be used over a wider range of frequency than wire-wound resistors. Resistors of similar form, printed onto a single substrate, are widely used in measuring instruments where accurate and constant ratios are required, since the maintenance of the ratio will normally be much better than the actual stability of the resistance.

Very high resistances. Above 10^9 Ω, the best available resistors are of protected thin-metal films, often completely encapsulated in glass. At these low conductances, any stray conductance due to moisture or dirt

becomes increasingly important. These resistances must therefore be kept clean and dry. Their resistance value is voltage-dependent to some extent, and so the voltage at which measurements are made must be specified.

Current-carrying resistors. These are needed for the calibration of resistive shunts which are used for measurement of direct current in the range from 10 to 10,000 A or higher. They are four-terminal resistors made from rods or strips of manganin, using several elements in parallel for the lower-resistance high-current values. They are mounted in oil-filled containers, fitted with some means of continuously stirring the oil. The range of resistance is from about 0.01 mΩ to 0.01 Ω, and in those intended for the larger currents the case is fitted with a water-cooled jacket. Because of the less well-defined thermal conditions, the accuracy obtainable is limited to 0.01% at the low current values and degrades to about 0.1% for high current values.

Decade resistors. These are switched assemblies of resistors, usually wire-wound, to provide variable resistances for laboratory work. They are necessarily two-terminal instruments, and hence have inbuilt errors due to lead resistances and variation of switch contacts. As in most switched decade resistance elements, each dial includes a 10 position in addition to the values from 0 to 9. This permits a ready self-calibration procedure to check linearity.

Kelvin-Varley dividers. These are used as variable voltage dividers for potentiometric work, for example, for the calibration of digital voltmeters from a fixed voltage standard. They consist (**Fig. 6**) of a series of sets of equal resistors connected by switches. In each set there are 11 resistors, 2 of which are spanned by the resistors of the next decade. This arrangement maintains an input resistance which is independent of the setting, and a low output resistance. *See* POTENTIOMETER.

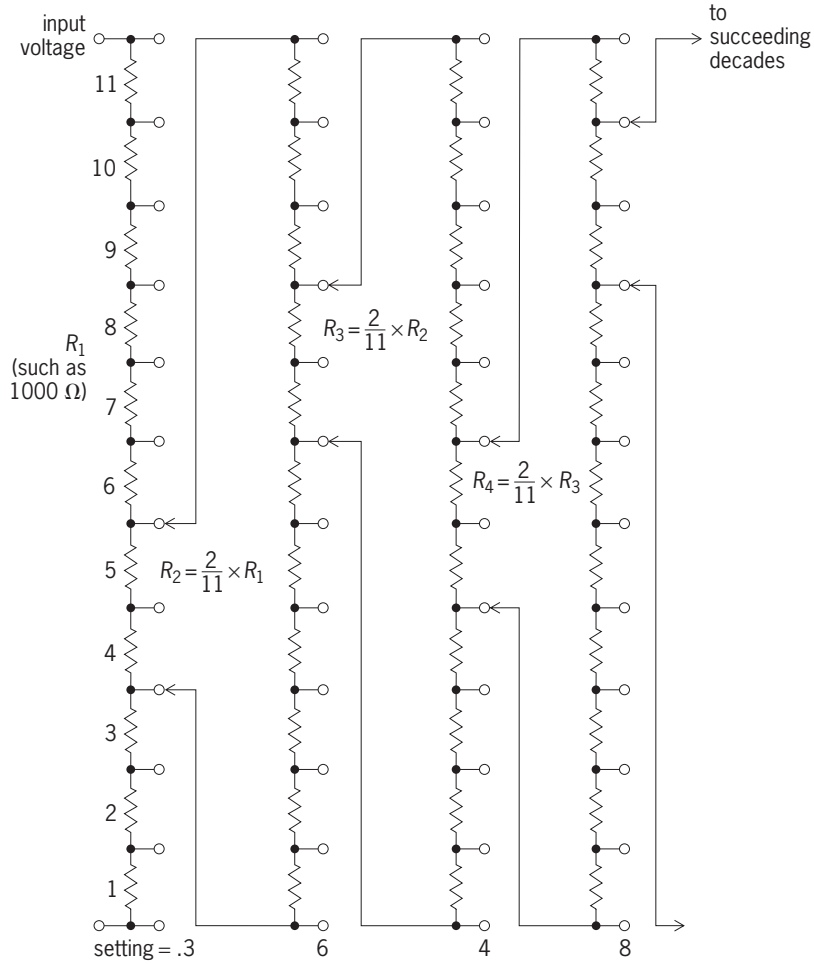

Fig. 6. Kelvin-Varley divider. Setting of divider with taps in the positions shown is 0.3648.

Comparison methods. The value of an unknown resistance R_x is determined by comparison with a standard resistor R_s.

Wheatstone bridge. This bridge (**Fig. 7**) is perhaps the most basic and widely used resistance- or impedance-comparing device. Its principal advantage is that its operation and balance are independent of variations in the supply. At balance, Eq. (7) is satisfied. The

$$R_x = R_S \times \frac{R_1}{R_2} \tag{7}$$

greatest sensitivity is obtained when all resistances are similar in value, and the comparison of standard resistors can then be made with a repeatability of about 3 parts in 10^8, the limit arising from thermal noise in the resistors. In use, the direction of supply is reversed periodically to eliminate effects of thermal or contact emf's.

The bridge is normally arranged for two-terminal measurements, and so is not suitable for the most accurate measurement at values below about 100 Ω, although still very convenient for lower resistances if the loss of accuracy does not matter. However, a Wheatstone bridge has also been developed for the measurement of four-terminal resistors. This involves the use of auxiliary balances, and resistors of the same value can be compared with uncertainties of a few parts in 10^8.

Typically a bridge will have two decade-ratio arms, for example, of 1, 10, 100, 1000, and 10,000 Ω, and a variable switched decade arm of 1–100,000 Ω, although many variations are encountered. For the measurement of resistors of values close to the

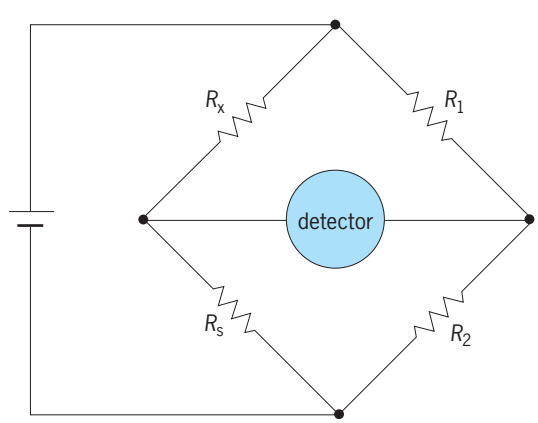

Fig. 7. Wheatstone bridge. At balance, $R_1/R_2 = R_x/R_s$.

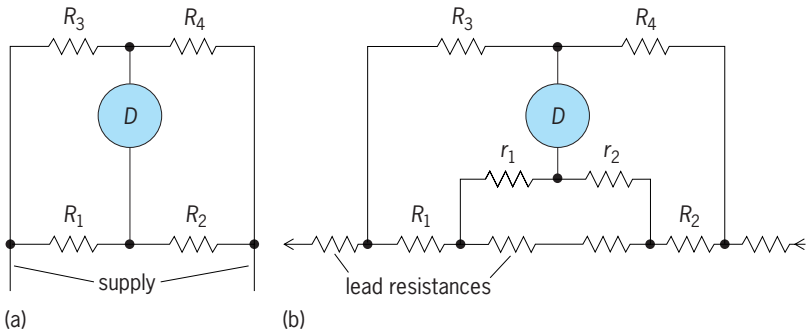

Fig. 8. Comparison of bridge circuits. (*a*) Wheatstone bridge. At balance, $R_1/R_2 = R_3/R_4$. (*b*) Kelvin double bridge. At balance, $R_1/R_2 = R_3/R_4 = r_1/r_2$.

decade values, a considerable increase in accuracy can be obtained by substitution measurement, in which the bridge is used only as an indicating instrument. The resistors being compared can be brought to the same value by connecting a much higher variable resistance across the larger of them, and the accuracy of this high-resistance shunt can be much less than that of the resistance being compared. *See* WHEATSTONE BRIDGE.

Kelvin double bridge. This is a double bridge (**Fig. 8**) for four-terminal measurements, and so can be used for very low resistances. The double condition for balance is given by Eq. (8). In addition to its use for accurate laboratory measurement of resistances below 100 Ω, it is very valuable for finding the resistance of conducting rods or bars, or for the calibration in the field of air-cooled resistors used for measurement of large currents. *See* KELVIN BRIDGE.

$$\frac{R_3}{R_4} = \frac{R_1}{R_2} = \frac{r_1}{r_2} \tag{8}$$

High-resistance measurements. Measurements of resistances from 10 megohms to 1 terohm (10^{12} Ω) or even higher with a Wheatstone bridge present additional problems. The resistance to be measured will usually be voltage-dependent, and so the measurement voltage must be specified. The resistors in the ratio arms must be sufficiently high in value that they are not overloaded. If a guard electrode is fitted, it is necessary to eliminate any current flowing to the guard from the measurement circuit. As an example, for measurement of 1 TΩ at 100 V, the ratio arms could be as shown in **Fig. 9**. The power dissipated in the 1-MΩ resistor is then 10 mW, and the bridge ratio is 10^6. The guard is connected to a subsidiary divider of the same ratio, so that any current flowing to it does not pass through the detector. Automated measurements can be made by replacing the ratio arms of the Wheatstone bridge by programmable voltage sources. An alternative method that can also be automated is to measure the *RC* time constant of the unknown resistor *R* combined with a capacitor of known value *C*. *See* INSULATION RESISTANCE TESTING.

Magnetic current comparator. This is a very effective instrument, which offers simple four-terminal measurements of resistors up to a value of 100 kΩ (or to 1 MΩ by modification). The essential feature of the comparator (**Fig. 10**) is that the ratio of the currents in the two circuits is measured by means of balancing the fluxes they produce in a high-permeability toroidal comparator, a device like a transformer but connected so that the dc fluxes due to each winding are in opposition. The balance condition of zero flux is sensitively detected by a third winding. This balance condition is then maintained by a servo circuit controlling one of the power supplies, and is obtained when the ampere-turns products of the two windings are equal. The resistors to be compared are connected to the two circuits, and the turns' ratio adjusted by rotary switches until the emf's across the

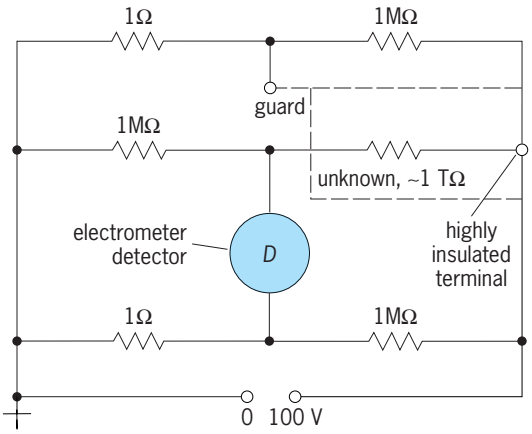

Fig. 9. Wheatstone bridge for high-resistance measurement.

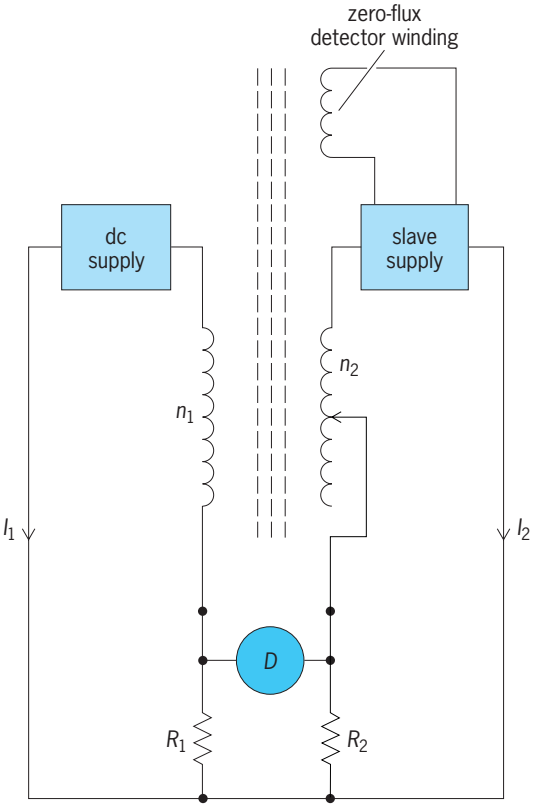

Fig. 10. DC current comparator bridge. $I_1/I_2 = n_2/n_1$; $R_1/R_2 = n_1/n_2$.

resistors balance. As with the Wheatstone bridge, a quick check can be made by reversing the positions of the two resistors being compared. An important merit of the system is that the ratio measured depends only on numbers of turns on the comparator, and not on ratio resistors, which will inevitably drift to some extent. *See* CURRENT COMPARATOR.

Cryogenic current comparator. This system operates in a similar way to the magnetic current comparator, except that the device that provides the accurate current ratio is a superconducting current comparator. The device relies on the Meissner effect in superconductivity. The principle can be described as follows, but the actual topology is more complex. A superconducting cylinder is threaded by two coils, one of one turn and one of N turns. The currents in the two coils are in inverse ratio to the number of turns. Any imbalance will give rise, via the Meissner effect, to current flowing in the surface of the superconducting cylinder. This current can be detected by another coil connected to a very sensitive superconducting quantum interference device (SQUID) detector, which acts as the feedback element in a servo system that maintains the two currents in the same precise ratio as the ratio of the turns in the two coils. The excellent shielding properties of superconductors make possible the provision of very accurate and stable ratios that can be measured to a few parts in 10^{10}. The small difference in the voltages across the two resistors being compared can then be measured by using a sensitive voltmeter. Resistance ratios can be measured with uncertainties approaching 1 part in 10^9. *See* MEISSNER EFFECT; SQUID; SUPERCONDUCTING DEVICES; SUPERCONDUCTIVITY.

Series-parallel resistance networks. The values of standard resistors at decade values are usually established from the 1-Ω level by means of buildup boxes. If n similar resistors are connected first in parallel and then in series, they provide the ratio n^2 between the series and parallel values. It can readily be shown that if each resistor is different from the mean by no more than some small amount δ (for example, 1 in 10^4), the ratio is achieved with an accuracy no worse than δ^2 (in this case, 1 in 10^8). This readily provides ratios of 100:1. A 10:1 ratio can be provided by connecting 10 resistors first in series, and then in three groups of 3 resistors in parallel connected in series with the remaining resistor, although this does not give as good accuracy as does the 100:1 arrangement. In order to connect four-terminal resistors so that only the resistive elements are truly in series, it is necessary to use special symmetrical four-terminal junctions, which are made so that the transresistance defined by the ratio of the voltage across any two connections to the current passing between the other two connections is less than 10^{-9} Ω. Resistance networks made in this way are called Hamon resistors, after their inventor B. V. Hamon. The nano-ohm junctions are made from solid copper, usually in cylindrical or cubic form, with the connecting wires brazed symmetrically into them.

Voltage and current measurement. An obvious and direct way of measuring resistance is by the simultaneous measurement of voltage and current, and this is usual in very many indicating ohmmeters and multirange meters. In most digital instruments, which are usually also digital voltage meters, the resistor is supplied from a constant-current circuit and the voltage across it is measured by the digital voltage meter. This is a convenient arrangement for a four-terminal measurement, so that long leads can be used from the instrument to the resistor without introducing errors (**Fig. 11**). The simplest systems, used in passive pointer instruments, measure directly the current through the meter which is adjusted to give full-scale deflection by an additional resistor in series with the battery. This gives a nonlinear scale of limited accuracy, but sufficient for many practical applications. *See* CURRENT MEASUREMENT; VOLTAGE MEASUREMENT.

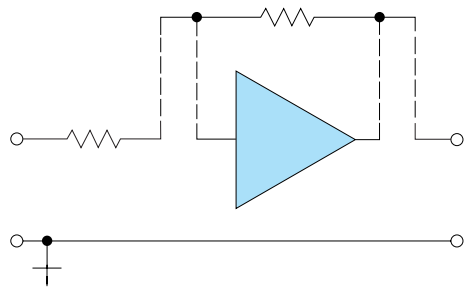

Fig. 11. Constant current resistance measurement circuit.

AC resistance measurements. The use of inductive voltage dividers makes it possible to create simple ac resistance bridges in which the ratio arms of the bridge are composed of the two parts of the inductive voltage divider. Inductive voltage dividers are inherently very accurate and stable devices, and so do not need the extensive regular recalibration needed by resistive bridges. The simple bridge shown in **Fig. 12** is principally suitable for the comparison of resistors when their reactance (due to their stray inductance or capacitance) is negligible in comparison with the resistance, for example, at low frequencies. An important field of application is for platinum resistance thermometry, in which the resistance to be measured is usually in the range 1–1000 Ω. *See* INDUCTIVE VOLTAGE DIVIDER; THERMOMETER.

At high frequencies, two effects occur which

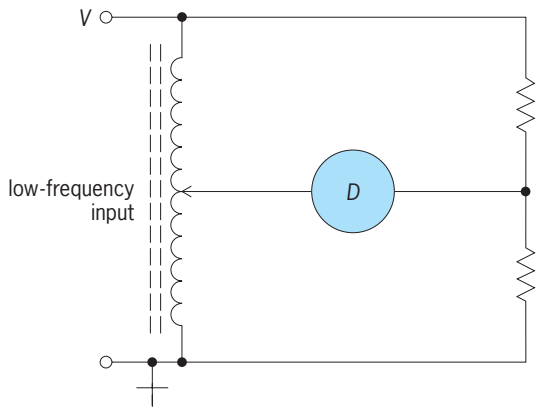

Fig. 12. Inductive divider resistance bridge.

change the measured value of resistance. The stray shunt capacitance C reduces the measured value of resistance by a factor $1/(1 + \omega^2C^2R^2)$ in addition to introducing resistance, and the resistance is increased due to the fact that the current flows only in the surface of the conductor. At radio and microwave frequencies, resistance becomes of importance principally in limiting the resolution or Q factor of circuits, and its measurement is often made by the measurement of this factor. *See* Q (ELECTRICITY).

Radio-frequency resistance is also used to evaluate the state of annealing of metals used in aircraft construction, and for this purpose it is measured by the coupled resistance induced in an inductor placed close to the metal. *See* EDDY CURRENT.

Cyril H. Dix; R. Gareth Jones

Bibliography. D. A. Bell, *Electronic Instrumentation and Measurements*, 2d ed., 1997; C. H. Dix and A. E. Bailey, Electrical standards of measurement, *Proc. IEE*, vol. 122, no. 1OR, October 1975; F. K. Harris, *Electrical Measurements*, 1952, reprint 1975; A. Hartland, Quantum standards for electrical units, *Contemp. Phys.*, 29:477–498, 1988; L. Schnell (ed.), *Technology of Electrical Measurements*, 1993; L. M. Thompson, *Electrical Measurements and Calibration*, 2d ed., 1994.

Resistance welding

A process in which the heat for producing the weld is generated by the resistance to the flow of current through the parts to be joined. The application of external force is required; however, no fluxes, filler metals, or external heat sources are necessary. Most metals and their alloys can be successfully joined by resistance welding processes.

Methods. Several methods are classified as resistance welding processes: spot, roll-spot, seam, projection, upset, flash, and percussion (**Fig. 1**).

Resistance spot welding. In this process, coalescence at the faying surfaces is produced in one spot by the heat obtained from the resistance to electric current through the work parts held together under pressure by electrodes. The size and shape of the individually formed welds are limited primarily by the size and contour of the electrodes (**Fig. 2***a*). *See* SPOT WELDING.

Roll resistance spot welding. Separated resistance spot welds are made with one or more rotating circular electrodes. The rotation of the electrodes may or may not be stopped during the making of a weld.

Resistance seam welding. Coalescence at the faying surfaces is produced by the heat obtained from resistance to electric current through the work parts held together under pressure by electrodes. The resulting weld is a series of overlapping resistance spot welds made progressively along a joint by rotating the electrodes (Fig. 2*b*).

Projection welding. Coalescence is produced by the heat obtained from resistance to electric current through the work parts held together under pressure by electrodes. The resulting welds are localized at predetermined points by projections, embossments, or intersections (Fig. 2*c*).

Upset welding. Coalescence is produced, simultaneously over the entire area of abutting surfaces or progressively along a joint, by the heat obtained from resistance to electric current through the area of contact of those surfaces. Pressure is applied before heating is started and is maintained throughout the heating period.

Flash welding. Coalescence is produced simultaneously over the entire area of abutting surfaces by the heat obtained from resistance to electric current between the two surfaces and by the application of pressure after heating is substantially completed. Flash and upsetting are accompanied by expulsion of the metal from the joint. *See* FLASH WELDING.

Percussion welding. Coalescence is produced simultaneously over the entire abutting surfaces by the heat obtained from an arc produced by a rapid discharge of electrical energy with pressure percussively applied during or immediately following the electrical discharge.

Fundamentals of process. In all resistance welding processes, heat is generated at a concentrated position and is controlled by adjusting the current magnitude and the duration of current flow. The spot welding process will be used to illustrate the fundamentals involved.

The welding current from the secondary of the transformer flows through a flexible conductor, an electrode, and the workpieces, and back through an electrode and a flexible conductor to the transformer. The application of force to the electrodes is used to position the electrodes in intimate contact with the workpieces, thus completing the electrical circuit. The amount of heat generated during the current flow through the resistance of the parts may be expressed as $H = I^2Rt$, where H is the heat generated in the workpieces in joules (watt-seconds), I is the current in amperes, R is the resistance of the workpieces in ohms measured between the electrodes, and t is the time of current flow in seconds. In a typical spot welding application for joining two sheets of 0.04-in.-thick (1.0-mm) steel, the current is high (typically 10,000 A), the resistance is low (typically 0.0001 ohm), and the time is short (typically 0.15 s). In this case the heat generated in the parts, as calculated by the formula, is 1500 joules.

The values of current and time are selected such that sufficient base metal is melted at the desired position to produce a weld of the size and strength appropriate for the application.

Electrode force. The electrode force, which is used to provide intimate contact between the workpieces, reduces porosity and cracking in the fusion zone of the weld. However, when the electrode force is excessive, the weld region becomes severely indented by the electrodes.

Welding current. In the majority of cases, the current during resistance welding is in the range from 5000 to 50,000 A. Because of the high values of current required, resistance welding machines are designed

as low-voltage sources, usually in the range from 1 to 25 V. These low voltages are obtained from a step-down transformer which usually has a single- or two-turn, cast-copper, water-cooled secondary. Power supplied to the primary of the transformer is usually obtained from public utility single-phase, alternating-current sources at 440 V and 60 Hz. The step-down ratio of the transformer is approximately 100; thus the current requirements of the primary are lowered to reasonable values, which generally range from 50 to 500 A.

Power sources. Alternating current of 60 Hz is used in about 90% of the installations, although three-phase frequency converters are used to supply 3- to 25-Hz voltage to single-phase transformers. Welding current may also be supplied by a direct-current or stored-energy source. Direct current may be obtained from various low-voltage sources, such as rectifiers, homopolar generators, or storage batteries. The energy may be stored during a relatively long period and released suddenly from capacitors, magnetic fields, storage batteries, or heavy flywheels on homopolar generators. These types of power supplies eliminate large transient loads on power lines.

Welding time. The duration of the welding time is short, generally in the range from $^1/_2$ cycle ($^1/_{\sqrt{20}}$ of a second) to a few minutes. In the great majority of applications, the time is in the range of 5–120 cycles of a 60-Hz source ($^1/_{\sqrt{12}}$ to 2 s).

Welding electrodes. Resistance welding electrodes serve several purposes in the process: the application of the force required to bring the workpieces into alignment and into intimate contact; the production of proper contact resistance between electrodes and the workpieces; the conduction of welding current; the prevention of weld porosity and cracking; the prevention of "spitting," which is liquid-metal expulsion during welding; dissipation of heat developed at the electrode-workpiece interface; and dissipation of heat developed in the workpieces after the weld is made.

Electrodes are made of copper alloys which provide a satisfactory combination of mechanical properties and electrical and thermal conductivities. Electrodes generally are water-cooled to dissipate the heat developed, to minimize deformation of the electrode tips, and to prevent them from sticking to the workpieces.

Controls. The functions of these controls are to initiate and terminate the welding current, to control the magnitude of the current, and to sequence the operations. Controls are classified into two groups: synchronous and nonsynchronous.

In synchronous controls the timer initiates and terminates each half-cycle of the input current to the primary of the welding transformer at specific times with respect to the input voltage wave. Thus each half-cycle of welding current is identical, and the power delivered is the same for consecutive welds.

With nonsynchronous controls, the timer initiates the current to the welding transformer at random points with respect to the input voltage wave. In the worst case the nonsynchronous closure of the circuit causes a current transient with a duration of about 2 cycles and a magnitude of current twice that of the steady-stage value. The differences in current and power for nonsynchronous and synchronous controls are shown in **Fig. 3**. The variations in timing and current, which become less important as the time of welding increases, can be neglected for times

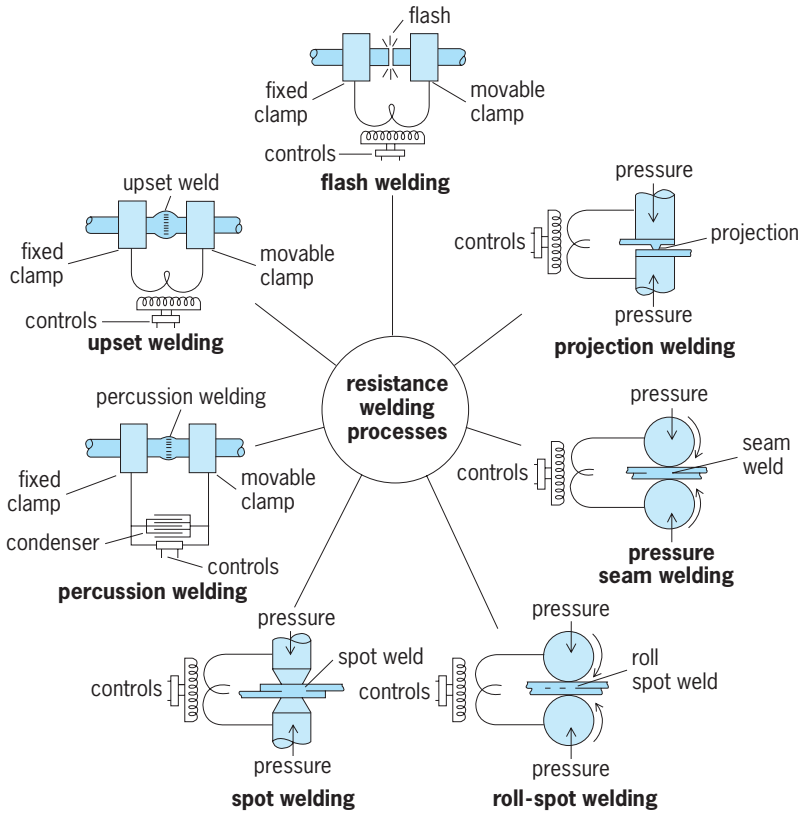

Fig. 1. Comparison of basic resistance welding processes. (*After B. E. Rossi, Welding Engineering, McGraw-Hill, 1954*)

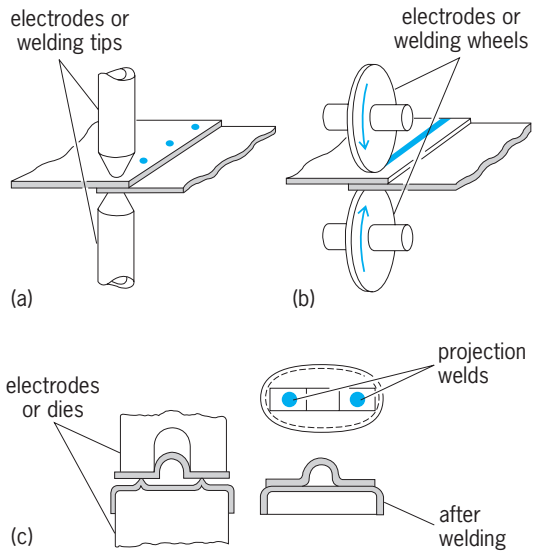

Fig. 2. Three typical resistance welding methods. (*a*) Spot welding. (*b*) Seam welding. (*c*) Projection welding. (*After W. H. Kearns, ed., Welding Handbook, 7th ed., vol. 3, American Welding Society, 1976*)

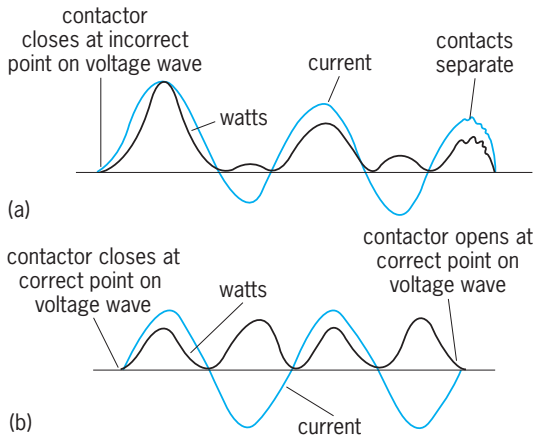

Fig. 3. Curves showing difference in current and power waves when circuit is closed nonsynchronously and synchronously. (*a*) Closure of circuit at incorrect point, with resultant decaying transient. The opening is also nonsynchronous. (*b*) Synchronous opening and closing of circuit at point of zero current corresponding to minimum transient. (*After A. L. Phillips, ed., Welding Handbook, 6th ed., American Welding Society, 1969*)

greater than about 20 cycles. However, when the welding time is less than 10 cycles, synchronous controls must be used for producing consistent welds.

The control of the magnitude of the welding current is accomplished, in coarse steps, by changing the taps on the welding transformer and, in a continuous manner, by adjustment of an electronic heat control. Synchronous controls are provided with electronic heat controls; however, these heat controls may be added to nonsynchronous controls. When a spot or projection weld is made by more than one impulse of current, the process is termed multiple-impulse welding. Multiple-impulse welding is also used to produce overlapping welds in the seam welding process.

Various levels of impulses can be used to preheat the metal, make the weld, refine the grain structure of the weld, and temper the hardened weld zone.

Machines. Welding machines are selected for use on the basis of the type of weld to be produced, the required weld quality, and production schedules. Simple, standard equipment is often utilized for production operations. However, because of economic considerations, complex resistance welding machines are used in the automotive and appliance industries, where high hourly production rates with minimum labor costs are essential.

Modern machines are designed with low-inertia systems which enable the electrodes to maintain the proper force on the workpieces at all times. If the electrodes fail to maintain adequate force during the welding operation, increased contact resistance will occur at the electrode-workpiece interface and at the workpiece interface. This increased resistance may cause excessive heating and result in deformed electrodes and expulsion of molten metal.

The clamping force of the electrodes is obtained by manual, mechanical, hydraulic, or pneumatic means. Prior to welding, electrodes should approach rapidly but in a controlled manner, so that they are not deformed by impacting the workpieces.

Surface preparation. Preweld surface preparation is usually required to control the magnitude and the variation in magnitude of the contact resistance at the electrode-workpiece interfaces and at the workpiece interface. Thus surface preparation, which removes dirt, oil, and oxide films by mechanical or chemical methods, provides for the production of consistent welds.

Metallurgical considerations. Most metals and alloys can be resistance-welded to themselves and to each other. The weld properties are determined by the metal and by the resultant alloys which form during the welding process. Stronger metals and alloys require higher electrode forces, and poor electrical conductors require less current. Copper, silver, and gold, which are excellent electrical conductors, are very difficult to weld because they require high current densities to compensate for their low resistance. Medium- and high-carbon steels, which are hardened and embrittled during the normal welding process, must be tempered by multiple impulses. *See* WELDING AND CUTTING OF MATERIALS. Ernest F. Nippes

Bibliography. R. Carr and R. O'Con, *Welding Practices and Procedures*, 1983; A. C. Davies, *The Science and Practice of Welding*, 2 vols., 10th ed., 1993; W. H. Kearns (ed.), *Welding Handbook*, 7th ed., vol. 3, American Welding Society, 1980; Resistance Welders Manufacturers Association, *Resistance Welding Manual*, 4th ed., 1989.

Resistor

One of the three basic passive components of an electric circuit that displays a voltage drop across its terminals and produces heat when an electric current passes through it. The electrical resistance, measured in ohms, is equal to the ratio of the voltage drop across the resistor terminals measured in volts divided by the current measured in amperes. *See* OHM'S LAW.

Resistors are described by stating their total resistance in ohms along with their safe power-dissipating ability in watts. The tolerance and temperature coefficient of the resistance value may also be given. *See* ELECTRICAL RESISTANCE; ELECTRICAL RESISTIVITY.

All resistors possess a finite shunt capacitance across their terminals, leading to a reduced impedance at high frequencies. Resistors also possess inductance, the magnitude of which depends greatly on the construction and is largest for wire-wound types. *See* CAPACITANCE; ELECTRICAL IMPEDANCE; INDUCTANCE.

Classification by use. Resistors may be classified according to the general field of engineering in which they are used.

Power resistors. Such resistors range in size from about 5 W to many kilowatts and may be cooled by air convection, air blast, or water. The smaller sizes, up to several hundred watts, are used in both the power and electronics fields of engineering.

Resistor color code

Color	Figures	Multiplier	Tolerance, %	Temperature coefficient, ppm/°C
No band	—	—	20	—
Silver	—	0.01	10	—
Gold	—	0.1	5	—
Black	0	1	—	200
Brown	1	10	1	100
Red	2	10^2	2	50
Orange	3	10^3	—	15†
Yellow	4	10^4	*	25†
Green	5	10^5	0.5	—
Blue	6	10^6	0.25	10
Violet	7	10^7	0.1	5
Gray	8	10^8	—	1
White	9	10^9	—	—

*Yellow sometimes used instead of gold to indicate 5% tolerance.
†Orange or yellow sixth bands may indicate reliability ratings when used on a 5% tolerance device.

Instrument resistors. Direct-current (dc) ammeters employ resistors as meter shunts to bypass the major portion of the current around the low-current elements. These high-accuracy, four-terminal resistors are commonly designed to provide a voltage drop of 50–100 mV when a stated current passes through the shunt. *See* AMMETER.

Voltmeters of both the dc and the ac types employ scale-multiplying resistors designed for accuracy and stability. The arc-over voltage rating of these resistors is of importance in the case of high-voltage voltmeters. *See* VOLTMETER.

Standard resistors. These devices are used for calibration purposes in resistance measurements and are made to be as stable as possible, in value, with time, temperature, and other influences. Resistors with values from 1 ohm to 10 megohms are wound by using wire made from special alloys. The best performance is obtained from quaternary alloys, which contain four metals. The proportions are chosen to give a shallow parabolic variation of resistance with temperature, with a peak, and therefore the slowest rate of change, near room temperature. *See* ELECTRICAL UNITS AND STANDARDS.

Resistors for electronic circuits. By far the greatest number of resistors manufactured are intended for use in the electronics field. The major application of these resistors is in transistor analog and digital circuits which operate at voltage levels between 0.1 and 200 V, currents between 1 μA and 100 mA, and frequencies from dc to 100 MHz. Their power-dissipating ability is small, as is their physical size.

Preferred values. Since their exact value is rarely important, resistors are supplied in decade values (0.1, 1, 10, 100 ohms, and so forth) with the interval between these divided into a geometric series, thus having a constant percentage increase. For noncritical applications, values from a series with intervals of 20% (12 per decade) are appropriate. A series with 10% intervals (24 per decade) is often used for resistors having a tolerance of 1%. Where the precise value of a resistor is important, a series with 2.5% intervals (96 per decade) may be used.

Resistor marking. Most electronic-type resistors are marked by using a standard code consisting of a set of closely spaced colored bands which start near one end of the resistor. The code evolved from a three-digit system in which the first two bands define the two significant figures of the value, and the third band indicates the number by which the significant figures are to be multiplied in order to obtain the value, in ohms (see **table** and **illus.**). A tolerance of ±20% is assumed unless a fourth band is applied, spaced more widely from the previous bands. The code has been extended to allow three significant figures and a multiplier to be displayed in four equally spaced bands. These are always followed, after a wider gap, by a tolerance band, and there may be a sixth band to indicate the temperature coefficient.

Resistors may also be marked by using a figure code. The symbols Ω, K, and M can be used both to indicate the units in which the value is defined and to effectively serve as a decimal point. Thus 2Ω7 and 18K7 signify 2.7 ohms and 18,700 ohms. Decimal digits are sometimes used in the same way as colored

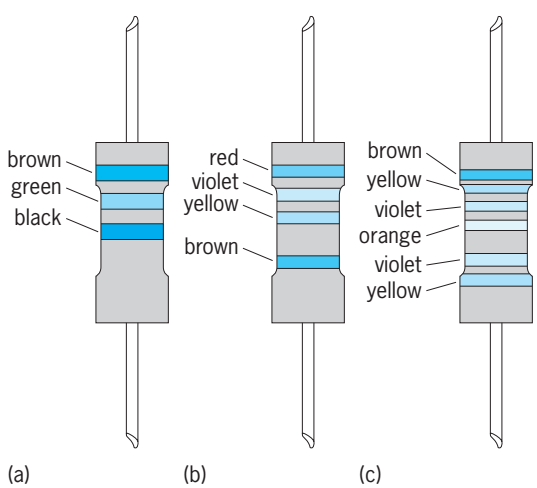

Examples of color-coded resistors. (*a*) 15 Ω ± 20%. (*b*) 270,000 Ω ± 1%. (*c*) 147,000 Ω ± 0.1%, 25 ppm/°C.

bands, so that the final digit is to be interpreted as a multiplier; for example, a 392-kΩ resistor may be marked 3923.

Classification by construction. Resistors are also classified according to their construction, which may be composition, film-type, wire-wound, or integrated circuit.

Composition resistor. This resistor is in wide use because of its low cost, high reliability, and small size. Basically it is a mixture of resistive materials, usually carbon, and a suitable binder molded into a cylinder. Copper wire leads are attached to the ends of the cylinder, and the entire resistor is molded into a plastic or ceramic jacket. The overall length of the jacketed resistor excluding the leads is $^1/_2$–$1^1/_3$ in. (13–34 mm) for resistors varying in power rating from 0.25 to 2 W. Composition resistors are commonly used in the range from several ohms to 10–20 MΩ, and are available with tolerances of 20, 10, or 5%.

Film-type resistor. This resistor is now the preferred type for most electronic applications because its performance has surpassed that of composition resistors and mass-production techniques have reduced the cost to a comparable level. The design of the film resistor further lends itself to the controlled manufacture of precision resistors of any desired value. Basically this resistor consists of a thin conducting film of carbon, metal, or metal oxide deposited on a cylindrical ceramic or glass former. The resistance is controlled by cutting a helical groove through the conducting film. This helical groove increases the length and decreases the width of the conducting path, thereby determining its ohmic value. By controlling the conductivity, thickness of the film, and pitch of the helix, resistors over a wide range of values can be manufactured. Since the resistor value may be measured while cutting is in progress, close manufacturing tolerances are possible. Standard tolerances of ±0.1% and temperature coefficients of resistance (TCR) better than 15 parts per million (ppm)/°C are readily available. Temperature coefficients as low as 1 ppm/°C are available at higher cost.

Film construction is used for very high value resistors, up to and even beyond 1 TΩ (10^{12} ohms). To avoid the problems caused by humidity, such resistors require special packaging in silicone-treated glass envelopes. A guard electrode is sometimes applied, to both the inner and outer glass surfaces. High-value resistors should never be touched by bare hands for their value will be affected by contamination.

The spiral-cut resistor displays a small inductive effect, which may be a problem at high frequencies. In order to obtain the widest possible bandwidth, film resistors are constructed by producing a serpentine path of resistance material on a flat substrate by using photographic methods, followed by mechanical or laser trimming. Such resistors can closely approach the performance of the best wire-wound types.

Wire-wound resistor. Wire remains the most stable form of resistance material available; therefore, all high-precision instruments rely upon wire-wound resistors. Wire also will tolerate operation at high temperatures, and so compact high-power resistors use this construction. Power resistors are available in resistance values from a fraction of an ohm to several hundred thousand ohms, at power ratings from one to several thousand watts, and at tolerances from 10 to 0.1%. The usual design of a power resistor is a helical winding of wire on a cylindrical ceramic former. After winding, the entire resistor is coated in vitreous enamel. Alternatively, the wound element may be fitted inside a ceramic or metal package, which will assist in heat dissipation. The helical winding results in the resistor having significant inductance, which may become objectionable at the higher audio frequencies and all radio frequencies.

Precision wire-wound resistors are usually wound in several sections on ceramic or plastic bobbins and are available in the range from 0.1 Ω to 10 MΩ. They can be supplied with a tolerance as small as 25 ppm and temperature coefficients less than 1 ppm/°C. Special selection can provide groups of components with even more closely matched temperature coefficients and values or ratios. This form of construction gives large and unpredictable inductive and capacitive effects, and so the use of devices of this type is limited to dc and frequencies below about 50 kHz. For ac standards applications, special constructional methods are used to control the reactive properties, for example, by winding the wire on a flat card, thereby reducing the cross-sectional area of the coil.

Integrated circuit resistor. These resistors must be capable of fabrication on a silicon integrated circuit chip along with transistors and capacitors. There are two major types: thin-film resistors and diffused resistors.

Thin-film resistors are formed by vacuum deposition or sputtering of nichrome, tantalum, or Cermet (Cr-SiO). Such resistors are stable, and the resistance may be adjusted to close tolerances by trimming the film by using a laser. Typical resistor values lie in the range from 100 Ω to 10 kΩ with a matching tolerance of ±0.2% and a temperature coefficient of resistance of ±10 to ±200 ppm/°C.

Diffused resistors are based upon the same fabrication geometry and techniques used to produce the active transistors on the silicon chip or die. A diffused base, emitter, or epitaxial layer may be formed as a bar with contacts at its extremities. The resistance of such a semiconductor resistor depends upon the impurity doping and the length and cross section of the resistor region. In the case of the base-diffused resistor, the emitter and collector regions may be formed so as to pinch the base region to a very small cross-sectional area, thereby appreciably increasing the resistance. The relatively large impurity carrier concentration in *n*- and *p*-type regions limits the resistance value. Resistor values between 100 Ω and 10 kΩ are common with a matching tolerance lower than ±2% but with a temperature coefficient of resistance approaching 2000 ppm/°C. The epitaxial layer resistor has a much lower impurity carrier concentration and offers resistance values up to 50 kΩ but with a higher temperature coefficient of resistance

of 3000 ppm/°C and a matching tolerance of about ±.6%. *See* INTEGRATED CIRCUITS.

Adjustable resistors. The deposited-film and wire-wound resistors lend themselves to the design of adjustable resistors or rheostats and potentiometers. Adjustable-slider power resistors are constructed in the same manner as any wire-wound resistor on a cylindrical form except that when the vitreous outer coating is applied an uncovered strip is provided. The resistance wire is exposed along this strip, and a suitable slider contact can be used to adjust the overall resistance, or the slider can be used as the tap on a potentiometer. *See* POTENTIOMETER; RHEOSTAT.

Where continuous adjustment of the resistor is intended, a ring-shaped form is generally used. For power resistors, the ring is wound with resistance wire. For compact 0.5- and 1-W resistors, the ring is coated on one surface with a resistance film. Adjustable resistors have the problem of maintaining a good, noise-free, electrical contact at the wiper, which is mounted on a shaft concentric with the ring.

For discussion of nonlinear resistors *see* THERMISTOR; VARISTOR. R. B. D. Knight

Bibliography. A. B. Glaser and G. E. Subak-Sharpe, *Integrated Circuit Engineering*, 1977; P. R. Gray and R. G. Meyer, *Analysis and Design of Integrated Circuits*, 4th ed., 2001; F. K. Harris, *Electrical Measurements*, 1952, reprint 1975; P. Horowitz and W. Hill, *The Art of Electronics*, 2d ed., 1989; F. Mazda (ed.), *Electronics Engineer's Reference Book*, 7th ed., 1993; M. G. Say (ed.), *Electrical Engineer's Reference Book*, 15th ed., 1993.

Resolving power (optics)

A quantitative measure of the ability of an optical instrument to produce separable images. The images to be resolved may differ in position because they represent (1) different points on the object, as in telescopes and microscopes, or (2) images of the same object in light of two different wavelengths, as in prism and grating spectroscopes. For the former class of instruments, the resolving limit is usually quoted as the smallest angular or linear separation of two object points, and for the latter class, as the smallest difference in wavelength or wave number that will produce separate images. Since these quantities are inversely proportional to the power of the instrument to resolve, the term resolving power has generally fallen into disfavor. It is still commonly applied to spectroscopes, however, for which the term chromatic resolving power is used, signifying the ratio of the wavelength itself to the smallest wavelength interval resolved. The figure quoted as the resolving power or resolving limit of an instrument may be the theoretical value that would be obtained if all optical parts were perfect, or it may be the actual value found experimentally. Aberrations of lenses or defects in the ruling of gratings usually cause the actual resolution to fall below the theoretical value, which therefore represents the maximum that could be obtained with the given dimensions of the instrument in question. This maximum is fixed by the wave nature of light and may be calculated for given conditions by diffraction theory. *See* DIFFRACTION; OPTICAL IMAGE.

Chromatic resolving power. The chromatic resolving power R of any spectroscopic instrument, including prisms, gratings, and interferometers, is defined by Eq. (1), where $\delta\lambda$ represents the difference

$$R = \frac{\lambda}{\delta\lambda} \tag{1}$$

in wavelength of two equally strong spectrum lines that can barely be separated by the instrument, and λ the average wavelength of these two lines. It is necessary to specify more precisely the term "barely separated," and for prisms and gratings, in which the width of the lines is determined by diffraction, this is done by use of Rayleigh's criterion. **Illustration** *a* shows the contours of two similar spectrum lines which are at the limit of resolution according to this criterion. The lighter curve represents the line shape due to Fraunhofer diffraction for the wavelength λ, the dashed curve that for $\lambda + \delta\lambda$, and the heavy curve the sum of the two. Rayleigh's criterion specifies that the lines are resolved when the principal maximum of one falls exactly on the first minimum (zero intensity) of the other. Diffraction theory shows that the intensity I of either pattern at the central crossing point is $4/\pi^2$ of that at the maximum, so that the curve representing the sum dips to 81% at the center. The theory also shows that the angular separation $\delta\theta$ of the rays forming the two maxima is λ/a, where a is the linear width of the

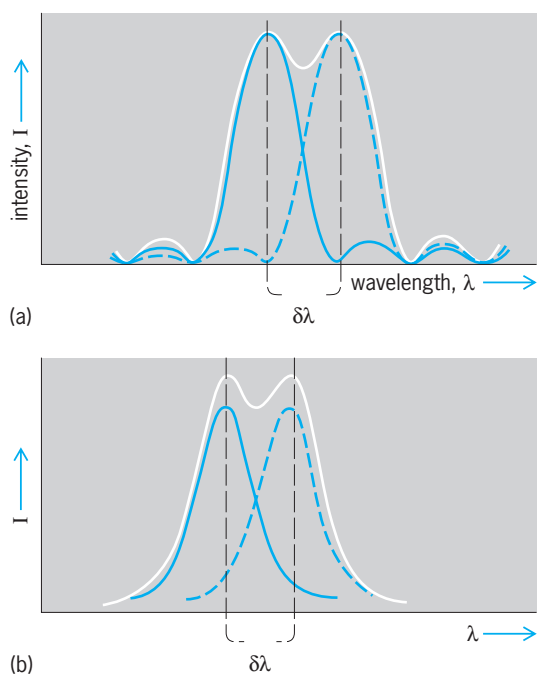

Resolution of two spectrum lines (*a*) when the shape is determined by diffraction (Rayleigh criterion), and (*b*) when the shape follows the Airy formula. The latter is applicable to multiple-beam interferometers.

beam of light emerging from the prism or grating. Hence, quite generally for such an instrument, the resolving power may be defined by Eq. (2*a*), which can be expressed in words as Eq. (2*b*). In a given

$$R = \frac{\lambda}{\delta\lambda} = \frac{\lambda}{\delta\theta}\frac{d\theta}{d\lambda} = a\,\frac{d\theta}{d\lambda} \qquad (2a)$$

$$\text{Chromatic resolving power} = \text{width of emergent beam} \times \text{angular dispersion} \qquad (2b)$$

instrument, the calculation of resolving power thus involves finding the last two quantities.

Resolving power of prisms. When a prism is used at minimum deviation, the resolving power depends on the length b of the base of the prism and the slope $dn/d\lambda$ of the dispersion curve giving the wavelength variation of the refractive index n. Thus Eq. (3) holds.

$$R = b\frac{dn}{d\lambda} \qquad (3)$$

Here the assumption is made that the prism is completely filled by the beam of light. If it is not, b must represent the difference in path length between the longest and shortest rays through the prism. *See* OPTICAL PRISM.

Resolving power of gratings. This equals the product of the order of interference m and the total number of rulings N. The order m may be expressed in terms of the grating space s and the angles α and β of incidence and diffraction. Thus Eq. (4) holds.

$$R = mN = \frac{Ns(\sin\alpha + \sin\beta)}{\lambda} = \frac{w(\sin\alpha + \sin\beta)}{\lambda} \qquad (4)$$

Here w is the width of the ruled area of the grating. For the limiting case of grazing angles of incidence and diffraction, the maximum possible R is seen to be $2w/\delta$, or the number of wavelengths in twice the width of the grating. *See* DIFFRACTION GRATING.

Resolving power of interferometers. For the type of interferometer most commonly used, the Fabry-Perot interferometer, the resolving power may be expressed as the product of the order of interference $m = 2t/\lambda$, where t is the separation of the interferometer mirrors, and an effective number N_{eff} of interfering beams. For interferometers the line contour of the spectrum lines is not that of Fraunhofer diffraction, but is given by a relation called the Airy formula. This contour has no points of zero intensity but has the general shape shown in illus. *b*. Therefore, the Rayleigh criterion cannot be applied in the usual way. If, however, the two curves are made to cross at the half-intensity point of each, it is found that there is a dip of approximately 20% in the resultant curve. The value of N_{eff} is thereby specified, and the resolving power R is given by Eq. (5), where

$$R = mN_{\text{eff}} = m\left(\frac{\pi\sqrt{\rho}}{1-\rho}\right) \qquad (5)$$

ρ designates the reflectance of the interferometer plates. *See* INTERFEROMETRY.

Resolving power of telescopes. This depends on the size of the diffraction maximum produced when light from a distant point source passes through a circular aperture of size equal to that of the objective lens or mirror. A graph of the intensity in the diffraction pattern plotted against radial distance closely resembles one of the curves of illus. *a*, and hence the pattern consists of a central spot surrounded by faint rings. The angular radius of the first dark ring corresponds, by the Rayleigh criterion, to the angular separation of two point sources that are barely resolved. Theory gives this angle, which represents the resolving limit, as defined by Eq. (6) for

$$\alpha = \frac{1.220\lambda}{d}\ \text{radians} = \frac{14.1}{d}\ \text{seconds of arc} \qquad (6)$$

$\lambda = 560$ nanometers and d, the diameter of the objective lens, in centimeters. *See* TELESCOPE.

Resolving power of microscopes. This is determined by diffraction of a circular aperture representing the exit pupil of the microscope objective. There are two important differences between the resolving power of microscopes and that of telescopes. First, the resolving limit of microscopes is expressed in terms of the smallest distance l between two points on the object that are just resolved. Second, this limit depends on the mode of illumination of the object. If the illumination is incoherent, so that there is no constant phase relation between light from adjacent points, the resolving limit is given by Eq. (7), where

$$l = \frac{0.61\lambda}{n\ \sin\alpha} \qquad (7)$$

n is the refractive index of the material (for example, oil) in the object space, and α is the angle that the extreme ray entering the objective makes with the axis of the instrument. The quantity $n \sin \alpha$ is called the numerical aperture of the objective. With coherent illumination the resolving limit is given by this formula, with 1.0 in place of 0.61, provided the illumination is central. When the object is illuminated from a point slightly to one side, the factor may be reduced to 0.5. *See* OPTICAL MICROSCOPE; SPECTROSCOPY.

Francis A. Jenkins; George R. Harrison

Bibliography. F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976; Optical Society of America, *Handbook of Optics*, 2 vols., 2d ed., 1995; W. J. Smith, *Modern Optical Engineering*, 3d ed., 2000.

Resonance (acoustics and mechanics)

When a mechanical or acoustical system is acted upon by an external periodic driving force whose frequency equals a natural free oscillation frequency of the system, the amplitude of oscillation becomes large and the system is said to be in a state of resonance.

When a simple oscillator of mass m, stiffness constant s, and mechanical damping constant R is driven by a periodic driving force $F \cos 2\pi ft$, it vibrates with a velocity amplitude given by the equation below, which implies that (1) the amplitude be-

$$V = \frac{F}{[R^2 + (2\pi fm - s/2\pi f)^2]^{1/2}}$$

comes a maximum when the driving frequency is $f = (1/2\pi) \cdot \sqrt{s/m}$, that is, at the natural free oscillation frequency of the oscillator, (2) small damping constants R are associated with large amplitudes of vibration at resonance, and (3) the smaller R, the more rapidly the amplitude decreases as the driving frequency departs from the resonance frequency. In addition, driving any vibrating system at its resonance frequency is characterized by a maximum dissipation of power.

A knowledge of both the resonance frequency and the sharpness of resonance is essential to any discussion of driven vibrating systems. When a vibrating system is sharply resonant, careful tuning is required to obtain the resonance condition. Mechanical standards of frequency must be sharply resonant so that their peak response can easily be determined. In other circumstances, resonance is undesirable. For example, in the faithful recording and reproduction of musical sounds, it is necessary either to have all vibrational resonances of the system outside the band of frequencies being reproduced or to employ heavily damped systems. *See* ACOUSTIC RESONATOR; OSCILLATOR; SYMPATHETIC VIBRATION; VIBRATION.

Lawrence E. Kinsler

Resonance (alternating-current circuits)

A condition in a circuit characterized by relatively unimpeded oscillation of energy from a potential to a kinetic form. In an electrical network there is oscillation between the potential energy of current in inductance. This is analogous to the mechanical resonance seen in a pendulum.

Three kinds of resonant frequency in circuits are officially defined. Phase resonance is the frequency at which the phase angle between sinusoidal current entering a circuit and sinusoidal voltage applied to the terminals of the circuit is zero. Amplitude resonance is the frequency at which a given sinusoidal excitation (voltage of current) produces the maximum oscillation of electric charge in the resonant circuit. Natural resonance is the natural frequency of oscillation of the resonant circuit in the absence of any forcing excitation. These three frequencies are so nearly equal in low-loss circuits that they do not often have to be distinguished.

Phase resonance is perhaps the most useful in many practical situations, as well as being slightly simpler mathematically. The following discussion considers phase resonance in passive, linear, two-terminal networks.

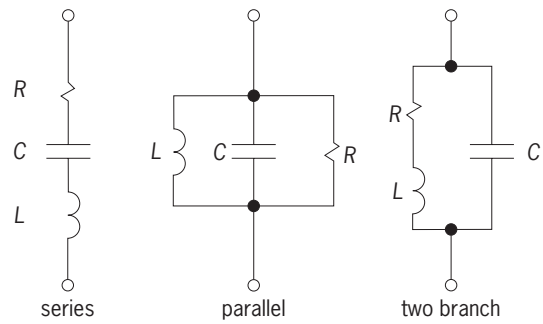

Fig. 1. Resonant circuits.

Resonance can appear in two-terminal networks of any degree of complication, but the three circuits shown in **Fig. 1** are simple and typical. The first illustrates series resonance and the second, parallel resonance; the third is a series-parallel resonant circuit of two branches (sometimes referred to as antiresonance). Series resonance is highly practical for providing low impedance at the resonant frequency. Parallel resonance is the dual of series resonance, but it is not practical because it assumes an inductive element with no resistance. The third example, however, shows an eminently practical means of providing the typical characteristic of parallel resonance, which is high impedance at the resonant frequency.

Use. Resonance is of great importance in communications, permitting certain frequencies to be passed and others to be rejected. Thus a pair of telephone wires can carry many messages at the same time, each modulating a different carrier frequency, and each being separated from the others at the receiving end of the line by an appropriate arrangement of resonant filters. A radio or television receiver uses much the same principle to accept a desired signal and to reject all the undesired signals that arrive concurrently at its antenna; tuning a receiver means adjusting a circuit to be resonant at a desired frequency.

Many frequency-sensitive circuits are not truly resonant, and oscillations of a certain frequency can be produced or enhanced by networks that do not involve inductance. It is difficult and expensive to provide inductance with integrated circuits, but frequency selection can be provided by the use of capacitance and resistance, a large amount of amplification being obtained from the semiconductor material employed.

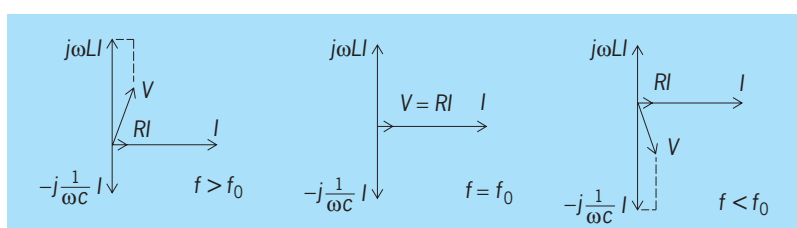

Fig. 2. Phasor diagrams at frequencies near resonance ($Q = 5$).

Series resonance. **Figure 2** shows phasor diagrams of the voltage, resulting from a given current (steady alternating current) in a series-resonant circuit, such as shown in Fig. 1. The component voltages across the three circuit elements add to give the total applied voltage V, as shown for a frequency slightly above resonance, for the resonant frequency, and for the frequency below resonance. It is of course possible in low-loss (high Q) circuits for the voltage across the capacitance and the voltage across the inductance each to be many times greater than the applied voltage.

Analytically, the impedance of the series-resonant circuit is given by Eq. (1). The resonant frequency

$$Z = R + j\omega L + \frac{1}{j\omega C} = R + j\left(\omega L - \frac{1}{\omega C}\right) \qquad (1)$$

f_0 is the frequency at which Z is purely real (phase resonance), so $\omega_0 L = 1/\omega_0 C$, or $2\pi f_0 L = 1/2\pi f_0 C$, from which Eq. (2) obtains. A more convenient notation is expressed by Eq. (3). In Eqs. (1)–(3), $Z =$

$$f_0 = \frac{1}{2\pi\sqrt{LC}} \qquad (2)$$

$$Z = R_0\left(\frac{R}{R_0} + jQ_0\delta\frac{2+\delta}{1+\delta}\right) \qquad (3)$$

impedance at the terminals of the series-resonant circuit; R, L, $C =$ the three circuit parameters; $R_0 =$ resistance (effective) at resonant frequency; $Q_0 = \omega_0 L/R_0$; $\delta = (\omega - \omega_0)/\omega_0$; $\omega = 2\pi f$, where f is frequency (hertz); $\omega_0 = 2\pi f_0$, where f_0 is resonant frequency.

Equation (3) is true for all series-resonant circuits, but interest is mainly in circuits for which Q_0, the quality factor at the resonant frequency, is high (20 or more) and for which δ, the fractional detuning, is low (perhaps less than 0.1). Assuming high Q_0 and low δ, which means a low-loss circuit and a frequency near resonance, Eq. (4) is very nearly the relative

$$\frac{Y}{Y_0} = \frac{Z_0}{Z} = \frac{1}{1 + j2Q_0\delta} \qquad (4)$$

admittance of the series-resonant circuit.

Universal resonance curve. The magnitude and the real and imaginary components of Eq. (4) are usefully plotted in the universal resonance curve of **Fig. 3**. Since Y/Y_0 is plotted as a function of $Q_0\delta$, this curve can be applied to all series-resonant circuits. (If $Q_0 =$ 20, the error in Y barely exceeds 1% of Y_0 for any δ, and is less for small δ.)

Moreover, because of the duality of the network, the curve can also be applied to any parallel-resonant circuit (Fig. 1) provided Q_0 is now interpreted as $Q_0 = R_0/\omega_0 L$. When used for a parallel-resonant circuit, the curve of Fig. 3 gives not Y/Y_0 but the relative input impedance Z/Z_0.

Finally, the universal resonance curve of Fig. 3 can also be applied (with the same slight approximations) to the two-branch resonant circuit of Fig. 1. For this purpose the curve shows Z/Z_0 (as for the three-branch parallel-resonant circuit), but the value of Q to be used is $Q_0 = \omega_0/R_0$, exactly as with the series-resonant circuit. Note that Z_0 for this circuit is given by Eq. (5) [instead of being equal to R_0 as it is

$$Z_0 = (\omega_0 L)Q_0 = R_0 Q_0{}^2 \qquad (5)$$

in the other two circuits of Fig. 1].

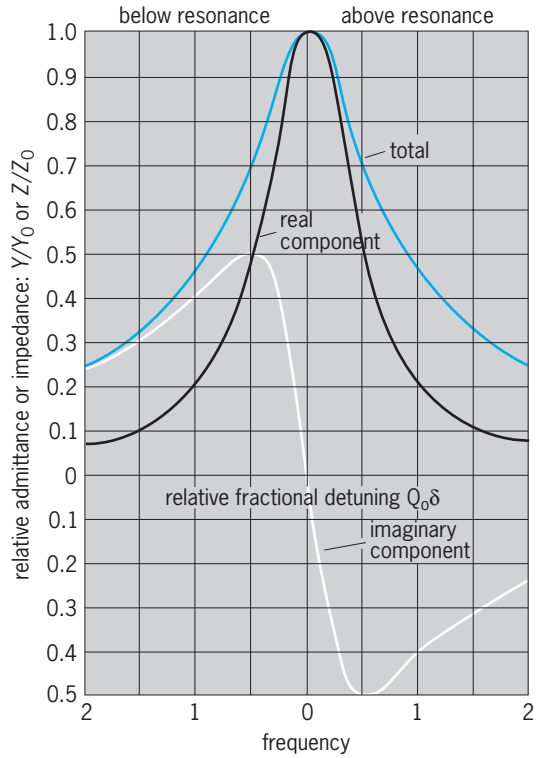

Fig. 3. Universal resonance curve. (*After H. H. Skilling, Electrical Engineering Circuits, 2d ed., John Wiley and Sons, 1955*)

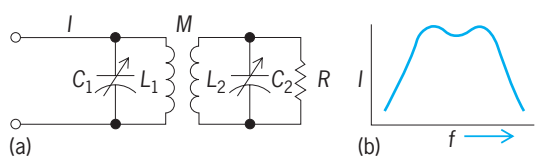

Fig. 4. Multiple resonance. (*a*) Double-tuned network with mutual inductance *M*. (*b*) Current *I* in *R* as function of frequency *f*.

Multiple resonance. If two or more coupled circuits are resonant at slightly different frequencies, many valuable characteristics can be obtained. **Figure 4** shows a double-tuned network and a typical curve of current in R, the load, as a function of frequency. *See* ALTERNATING-CURRENT CIRCUIT THEORY; NETWORK THEORY.

Hugh Hildreth Skilling

Bibliography. N. Balabanian, *Electric Circuits*, 1994; J. R. Duff and S. L. Herman, *Alternating Current Fundamentals*, 6th ed., 1999; Institute of Electrical and Electronics Engineers, *IEEE Standard Dictionary of Electrical and Electronics Terms*, 5th ed., 1992.

Resonance (molecular structure)

A feature of the valence-bond method, which is a mathematical procedure to obtain approximate solutions to the Schrödinger equation for molecules. The term came into use because the procedure is similar to that describing how weakly coupled tuning forks, pendulums, and such resonate, that is, transfer energy back and forth to one another. The valence-bond method is based on the theorem that if two or more solutions to the Schrödinger equation are available, certain linear combinations of them will also be solutions. It has this basis in common with its rival, the molecular orbital method. The valence-bond and molecular orbital approaches are both approximations and, if carried out to their logical and exact extremes, must yield identical results; nevertheless, both are often described as theories. In the valence-bond theory, combinations of solutions represent hypothetical structures of the molecule in question. These structures are said to be resonance (or contributing) structures, and the real molecule is said to be the resonance hybrid (or just simply the hybrid) of these structures. *See* MOLECULAR ORBITAL THEORY; SCHRÖDINGER'S WAVE EQUATION.

After it was demonstrated that the Schrödinger equation provided an exact three-dimensional description of the hydrogen atom, and thus that it was the key to the description of all chemical species, the resonance concept was developed to deal with the much more complex problem of finding solutions for multielectron atoms and for the multiatom molecules. The resonance theory was initially more popular than the molecular orbital approach, since it was applicable and understandable in qualitative terms so that the chemist could forego the difficult mathematics altogether. Also, the contributing structures that could be used were the classical ones that had been in use all along, and the resonance theory provided a solution for a molecule which had baffled and preoccupied chemists for a century—benzene. The principal use of resonance still lies in the qualitative description of molecules whose properties would otherwise be difficult to understand. *See* BENZENE.

Structural formulas. It had long been a tradition in chemistry, especially in organic chemistry, to portray molecules in terms of the symbols of the atoms that are part of the molecules, with lines drawn between them to indicate chemical (covalent) bonds. Thus, methane, which is the principal component of natural gas and which has the formula CH_4, can be pictured to indicate that four hydrogen atoms are bound to a carbon atom (**1**); alternatively, the slightly more complex structure (**2**) is often drawn to indicate that

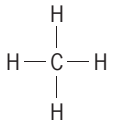

(**1**)

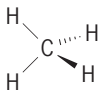

(**2**)

the real molecule is not planar, but tetrahedral; that is, the four hydrogen atoms are at the corners of a (hypothetical) tetrahedron which has the carbon atom at the center. The lines drawn between the atoms denote the electron pairs which provide the bonds. Almost all molecules can be represented by such structures, and conversely, correctly devised structures may (with certain exceptions) be expected to correspond to molecules even if as yet they are not known either in nature or in the laboratory.

Formulating a correct structure requires knowledge of the number of electrons each atom has in its outer (valence) shell, and application of the so-called octet rule, which states that most atoms strive in their compounds to reach a total of eight electrons in this shell. Exceptions to this rule are the lightest few atoms—especially hydrogen (H), which can accommodate only two electrons—and the heavier ones which may have more than eight. Thus, the rule is not rigid, but is a generalization that is especially useful in biologically significant molecules, which are primarily made of atoms which obey it, namely, carbon (C), nitrogen (N), and oxygen (O). The isolated atoms, H, C, N, and O, have one, four, five, and six electrons, respectively, in their valence shells. Thus methane could also be represented as structure (**3**). The carbon atom in this molecule is

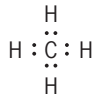

(**3**)

surrounded by four hydrogen atoms; it shares a pair of electrons with each of these, one contributed by the hydrogen atom and one by the carbon. In this way, the needs for all five atoms (two electrons for each H—, and eight for the C atom) are satisified. In a similar way, the correct structures can be readily devised for such molecules as ammonia, NH_3 (**4**), water, H_2O (**5**), and ethane, C_2H_6 (**6**), another constituent

(**4**)

(**5**)

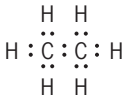

(**6**)

of natural gas. The unshared electron pairs in such representations are often simply omitted, but understood to be there. The structures of such molecules as ethylene, C_2H_4 (**7**), acetylene, C_2H_2 (**8**), and formaldehyde, CH_2O (**9**), represent slight

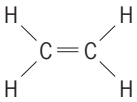

(**7**)

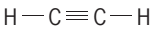

(**8**)

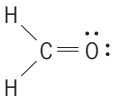

(**9**)

extensions in that two or even three electron pairs may be used to bind first-row atoms such as C, N, and O together. One of the interesting chemical properties of such molecules containing multiple bonds is

their tendency to form single bonds; thus, formaldehyde is readily converted into the new molecule trioxane (**10**) as shown in the reaction below.

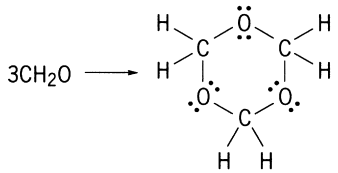

(**10**)

Benzene structure. Chemists had learned to deduce the atomic composition of molecules and the rudiments of the structural theory in the first half of the nineteenth century. The extension to three dimensions came in the 1870s, and that of the electron pair followed in the beginning of the twentieth century. But through most of this period the substance benzene posed a baffling challenge to organic chemists: in spite of its relatively simple formula, C_6H_6, they were unable to conceive of a suitable structure for it. While a great many structures were proposed, the properties of benzene corresponded to none of them. Thus, it was known that there were three and only three disubstituted benzenes (that is, benzenes in which two of the six hydrogen atoms had been replaced by other monovalent atoms such as chlorine, Cl). Furthermore, although it is difficult to avoid writing several multiple bonds, benzene and its derivatives have little or no tendency to undergo reactions known as additions, which typically convert such multiple bonds into single ones. Thus, structure (**11**) could be written for benzene, but

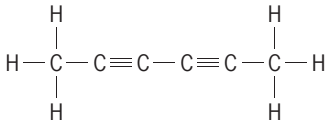

(**11**)

would not be correct. Such a substance should have only two disubstituted derivatives (with the substituting atoms either at the same or at opposite ends of the six-carbon chain). A compound that does have this structure is known, but it is strongly sensitive to addition reagents.

In the early 1870s F. A. Kekulé proposed a revolutionary idea: benzene must be represented by two structures (**12**) and (**13**) rather than one, and all com-

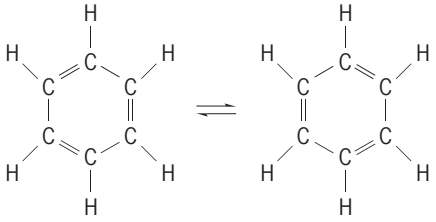

(**12**) (**13**)

pounds containing the benzene skeleton must be subject to a rapid equilibration (oscillation) between the two. Thus, whereas four disubstituted benzenes (**14**)–(**17**) would be expected on the basis of a

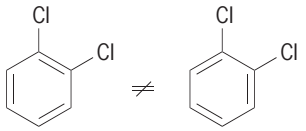

(**14**) (**15**)

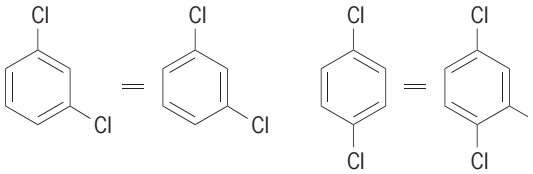

(**16**) (**17**)

single structure for benzene, the equilibration between (**14**) and (**15**) was proposed to be so facile that only a mixture of the two can be isolated. This was a new development; chemists at that time assumed that every structure that can be drawn must correspond to an isolable compound, and that its chemical conversion to another compound requires some action on the part of a chemist: heating, the addition of chemical reagents, and so on. Many instances of such labile structures are now known; they are described as tautomers. However, benzene is not among them. *See* TAUTOMERISM.

Hybrid structure. Kekulé's description of benzene was not completely satisfactory. While it accounted for the number of isomers, it did not explain why the compound failed to exhibit reactivity indicating the presence of multiple bonds. The problem therefore continued to attract attention, but its final solution had to wait until the advent of quantum mechanics in the early part of the twentieth century. In a sense, this solution is an expansion of Kekulé's oscillating pair: the so-called activation energy (the energy which must be imparted to a molecule in order to make it overcome the barrier that keeps it from being converted into another molecule) is negative in the case of benzene with respect to the oscillation, and this molecule therefore exists neither as (**12**) nor as (**13**) at any time, but it is in an intermediate form (**18**) all the time. This intermediate structure of

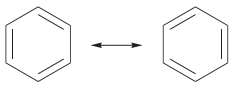

(**18**)

benzene is described in terms of Kekulé's structures with the symbol ↔ between them; this is intended to signify that benzene has neither structure, but in fact is a hybrid of the two. The properties of benzene are thereby indicated to be those of neither (**12**) nor (**13**), but to be intermediate between the two. Thus, since single bonds are longer than double bonds, it might be expected from the Kekulé structures that the molecule should be irregularly shaped; in fact, benzene is known to have the shape of a regular hexagon. All six of the carbon-carbon bonds are equally long, and their length is intermediate

between those of normal single and double bonds. The only property of the hybrid which is not intermediate between those of the hypothetical contributing structures is the energy: the energy of a resonance hybrid is by definition always at a minimum. This fact is responsible for the abnormal reluctance of benzene to undergo addition reactions; such reactions would lead to products that no longer have the resonance energy.

Other structures. Although benzene is the classical example of resonance, the phenomenon is certainly not limited to it; other instances include carboxylate anions such as formate (**19**) and guanidinium cation (**20**). Furthermore, the properties of all compounds

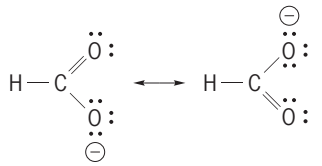

(**19**)

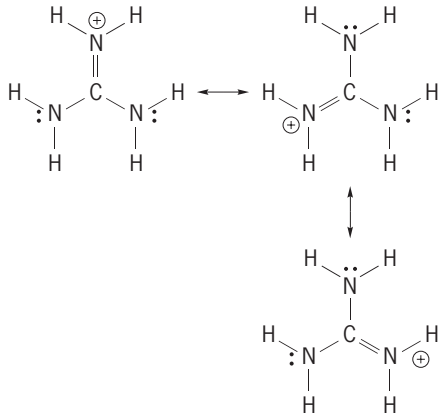

(**20**)

are affected by resonance to some degree. Thus, the properties of acrolein (**21**) suggest that the molecule

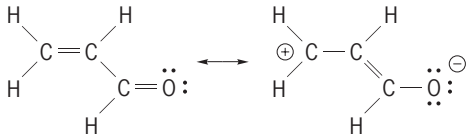

(**21**)

cannot be adequately described by a single structure: the single carbon-carbon bond length is rather short, and the double bonds are longer than normal. Thus, the dipolar structure is making a contribution. The main difference in the preceding cases is that the contributions are not equal; the weight of the dipolar structure is much less than half. In general, this inequality of weight will occur whenever the two structures drawn are different. *See* STRUCTURAL CHEMISTRY.

Bond order. This is a useful numerical device to indicate the occurrence of resonance in various molecules. The bond orders of single, double, and triple bonds are 1, 2, and 3, respectively; that of the carbon-carbon bonds in benzene and of the carbon-oxygen bonds in carboxylate anions is 1.50, and that of the carbon-nitrogen bonds in guanidinium ion is 1.33. The importance of this concept is its connection with experimental observables such as the stretching frequencies of infrared absorption spectroscopy, bond lengths as determined by means of x-ray diffraction, and dipole moments. On such grounds, the bond order of the carbon-oxygen bond in acrolein may be estimated to be 1.80; in other words, the zwitterion contributes about 20% to the structure of this molecule.

No-bond resonance. Bond orders can in principle have values less than one. An example of this phenomenon is furnished by the molecule diborane, B_2H_6. Boron has three valence electrons and hydrogen has one; it might therefore be supposed that these two elements should be able to form the borane BH_3, with the structure (**22**). However, this structure

(**22**)

would clearly constitute a violation of the octet rule for the boron atom. Accordingly, two of the entities combine to form the the molecule diborane, which in resonance symbolism can be represented as (**23**). This molecule differs from the apparently

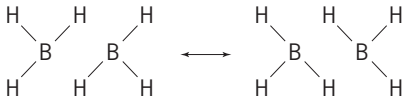

(**23**)

analogous ethane (**6**). In diborane, there are four hydrogen atoms bound to boron by bonds of order 1, and two that are bound to boron by bonds of order 0.5; in ethane, all six are bound to carbon by simple single bonds. Both contributing structures show atoms bound together but without a bond or electron pair—hence the name, no-bond resonance. The best-known examples occur in the boron hydrides and in carbocations such as the *tert*-butyl cation (**24**),

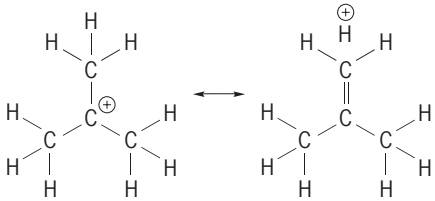

(**24**)

where any of the other eight hydrogen atoms could be chosen as the nonbonded one. *See* REACTIVE INTERMEDIATES.

Equilibration versus resonance. The distinction between equilibrating structures and resonance structures is not always as clearly in favor of the latter as it is in benzene, the resonance energy of which is about 36 kcal/mole. The cyclooctatetraene structures (**25**) and (**26**) furnish a clear example of the absence of

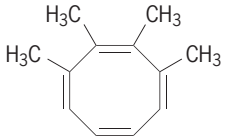

(25)

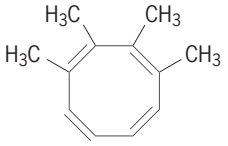

(26)

resonance. These so-called bond-shift isomers correspond to different, stable, and isolable compounds that equilibrate only slowly, even at elevated temperatures; the energy barrier between them is about 30 kcal/mole. Between such extremes as benzene and cyclooctatetraene, there are many instances of degenerate equilibration (reactions that give a product identical with the reactant) in which the rates are extremely fast; cyclobutadiene **(27)**

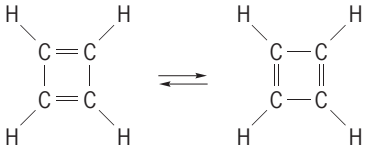

(27)

is an example of a species that equilibrates many times per second, even at temperatures as low as a few kelvins. On the other hand, there are also compounds that are hybrids of identical structures but that are not highly stablized; some members of the thiothiophthene family **(28)** exemplify this class.

(28)

The question of resonance versus rapid equilibration has in fact not been settled in many instances; the so-called nonclassical carbocations such as **(29)**, the semibullvalenes such as **(30)**, and the mixed valence complexes **(31)** are among them, and this general question has become a flourishing research

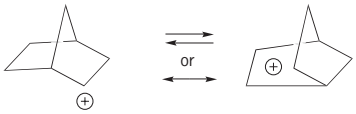

(29)

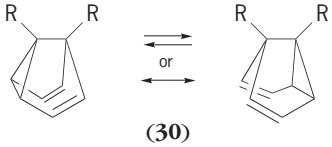

(30)

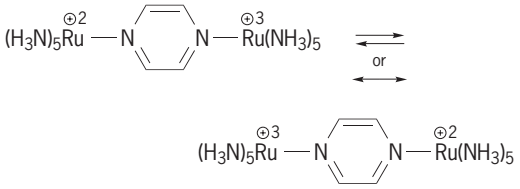

(31)

area. *See* ALICYCLIC HYDROCARBON; FLUXIONAL COMPOUNDS.

Limitations. The theory of resonance owed its initial success primarily to the fact that molecules could continue to be described in terms of the same structures that chemists had used for a century. However, many organic chemists have looked askance at no-bond structures. Moreover, use of the concept led to new questions that eventually began to undermine its popularity. Thus, its simple language cannot explain why cyclobutadiene is extremely unstable: it resisted all attempts at synthesis for many years, and it readily converts into other substances at temperatures as low as 20 K (−424°F). It is not obvious why cyclobutadiene is apparently not subject to resonance stabilization just as benzene is. The molecular orbital treatment can readily account for this observation; it also provides a much simpler account for the spectral properties of organic compounds. While it represented a more radical departure from the traditional structures and language of classical chemistry, it has been accepted. Although the molecular orbital approach has largely supplanted the valence-bond method, the resonance language remains so convenient that it is still used. *See* CHEMICAL BONDING; CONJUGATION AND HYPERCONJUGATION; MOLECULAR STRUCTURE AND SPECTRA; ORGANIC CHEMISTRY; QUANTUM CHEMISTRY.

William J. le Noble

Resonance (quantum mechanics)

An enhanced coupling between quantum states with the same energy. The concept of resonance in quantum mechanics is closely related to resonances in classical physics.

Frequency and energy matching. A child's swing provides a simple example of resonances. The swing is a mechanical pendulum with a natural frequency that depends only on the length of the suspending rope. Only rhythmic pushing with the same frequency as the natural frequency of the swing, that is, at resonance, is effective in building up a swinging motion of a reasonable amplitude. Other examples of resonances, in acoustics and in alternating-current circuits (tuning a radio or television to a particular broadcasting channel's frequency), display this same feature. *See* PENDULUM; RESONANCE (ACOUSTICS AND MECHANICS); RESONANCE (ALTERNATING-CURRENT CIRCUITS).

The matching of frequencies is central to the concept of resonance. An example is provided by waves, acoustic or electromagnetic, of a spectrum of

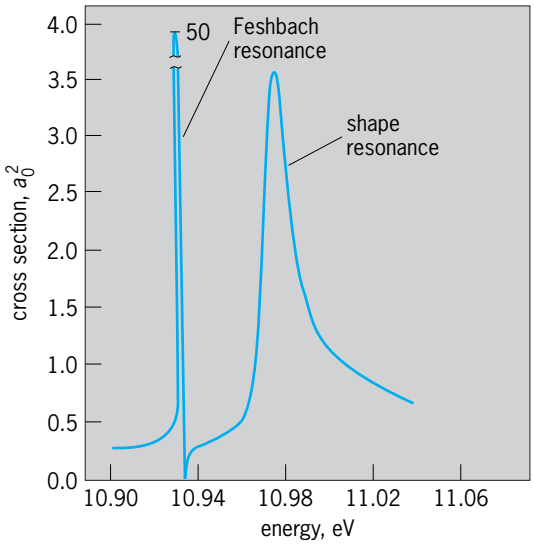

Fig. 1. Two resonances of the electron + hydrogen atom complex seen when the H^- ion absorbs light of energy shown. Cross section is measured in terms of the square of the Bohr radius, $a_0 \cong 0.5 \times 10^{-10}$ m. (*After H. C. Bryant et al., Observation of resonances near 11 eV in the photodetachment cross section of the H^- ion, Phys. Rev. Lett., 38:228–230, 1977*)

frequencies propagating down a tube or waveguide. If a closed side tube is attached, its characteristic natural frequencies will couple and resonate with waves of those same frequencies propagating down the main tube. This simple illustration provides a description of all resonances, including those in quantum mechanics. The propagation of all quantum entities, whether electrons, nucleons, or other elementary particles, is represented through wave functions and thus is subject to resonant effects. Indeed, nanometer-size channels for electron flow in semiconductors, fabricated by modern lithographic techniques, directly realize the above example. The electron current exhibits resonances. *See* ACOUSTIC RESONATOR; CAVITY RESONATOR; HARMONIC (PERIODIC PHENOMENA); WAVEGUIDE.

An important allied element of quantum mechanics lies in its correspondence between frequency and energy. Instead of frequencies, differences between allowed energy levels of a system are considered. In the presence of degeneracy, that is, of different states of the system with the same energy, even the slightest influence results in the system resonating back and forth between the degenerate states. These states may differ in their internal motions or in divisions of the system into subsystems. The above example of wave flow suggests the terminology of channels, each channel being a family of energy levels similar in other respects. These energies are discretely distributed for a closed channel, whereas a continuum of energy levels occurs in open channels whose subsystems can separate to infinity. The energy of quantum systems replaces frequency in the classical counterpart. If all channels are closed, that is, within the realm of bound states, resonance between degenerate states leads to a theme of central importance to quantum chemistry, namely, stabilization by resonance and the resulting formation of resonant bonds. *See* DEGENERACY (QUANTUM MECHANICS); ENERGY LEVEL (QUANTUM MECHANICS); RESONANCE (MOLECULAR STRUCTURE).

Resonances in scattering. Resonances occur in scattering when at least one channel is closed and one open. Typically, a system is divided into two parts: projectile + target, for example, electron + atom, or nucleon + nucleus. One channel consists of continuum states with their two parts separated to infinity. The other, closed channel, consists of bound states. In the atomic example, a bound state of the full system would be a state of the negative ion and, in the nuclear example, a state of the larger nucleus formed by incorporating one extra nucleon in the target nucleus. Degeneracy implies that the bound state is an excited state of the complex. Thus, in the

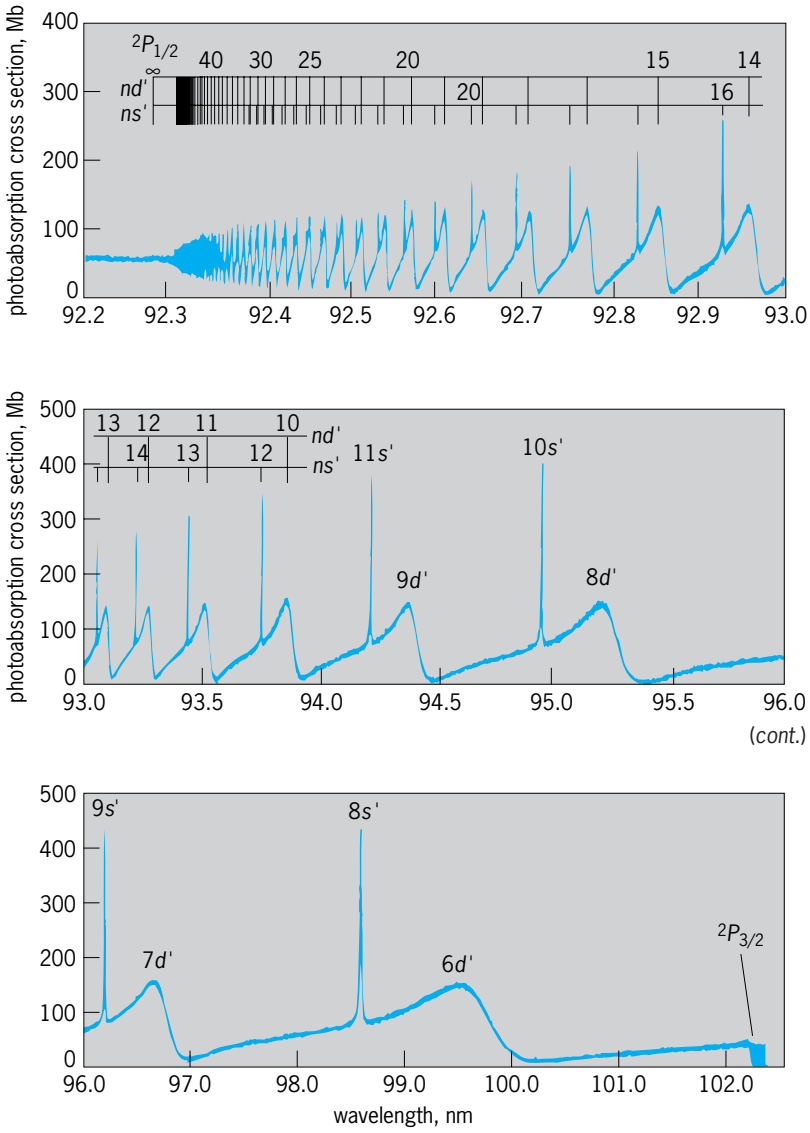

Fig. 2. Two prominent sequences of resonances seen in photoabsorption by the krypton atom between its two lowest ionization limits, marked as $^2P_{3/2}$ and $^2P_{1/2}$. Resonances are labeled ns' or nd', $n = 1, 2, 3, \ldots$. Decreasing photon wavelength translates to increasing energy. 1 Mb = 10^{-22} m^2. (*After K. Maeda, K. Ueda, and K. Ito, High-resolution measurement for photoabsorption cross sections in the autoionization regions of Ar, Kr, and Xe, J. Phys. B: At. Mol. Opt. Phys., 26:1541–1555, 1993*)

example of electron + hydrogen atom, the complex's state is denoted by H^{-**}, the two asterisks showing that both electrons are excited and have principal quantum numbers n larger than 1. For the lowest such H^{-**}, the electron is bound to excited hydrogen with $n = 2$, its energy lying slightly below the threshold energy with hydrogen in $n = 2$ and an electron at infinity. This state lies in the $n = 2$ channel. At that same energy lies a continuum state of the $n = 1$ channel, of elastic electron + hydrogen scattering, with hydrogen in the ground state ($n = 1$) and an electron with asymptotic kinetic energy of $\simeq$10 eV. *See* ATOMIC STRUCTURE AND SPECTRA; QUANTUM NUMBERS.

The complete and correct description of the state of the system is as a superposition of the bound and the continuum states. It is neither a pure bound state nor a pure continuum state, and different labelings are possible. Thus, starting from the bound-state end, a quasibound state is referred to. The mixing of the continuum state indicates that the system initially prepared as H^{-**} falls apart (decays), leaving behind a hydrogenation ($n = 1$), with a 10-eV electron escaping to infinity. A characteristic lifetime describes this decay, called autodetachment. Starting from the other end of electron + hydrogen ($n = 1$) scattering, at the energy corresponding to the bound state, the system resonates from continuum state into bound state where the electron is trapped in the H^{-**} state. This trapping is only temporary because of the continuum state character within the superposition, the electron returning eventually into the continuum channel. In the elastic scattering of electrons from hydrogen ($n = 1$), the temporary trapping when the energy approximates that of the bound state leads to an anomaly in the scattering cross section, the common manifestation of the resonance. *See* SUPERPOSITION PRINCIPLE.

A resonance may also be observed in quantities other than the cross section, such as the phase shift which typically rises through π radians. The energy range encompassing this rise (that is, the range covered by the cross-section anomaly) is called the resonance width, which is inversely related to the lifetime for decay of the quasibound state in the complementary description. A closely related parameter that measures the rate of rise of the phase shift is called the time delay, namely, the duration of the electron's temporary trapping in the negative-ion state as it scatters from the atom. Besides the energy position and the width, the shape of the cross section is characterized by a so-called profile index. *See* SCATTERING EXPERIMENTS (ATOMS AND MOLECULES); SCATTERING EXPERIMENTS (NUCLEI).

Feshbach and shape resonances. Such a resonance in the electron + hydrogen system appears as a sharp feature lying just below the hydrogen ($n = 2$) threshold at 10.95 eV (**Fig. 1**). This feature is called a Feshbach resonance, after the discussion by H. Feshbach of analogous features in nuclear physics. A second type of resonance, referred to as a shape resonance, appears as a second, broader feature. Its name arises because the bound and continuum aspects are now associated with a subdivision of the real three-dimensional space of motion reflecting the shape of the potential between projectile and target. This shape consists typically of a barrier separating an inner deep well from a shallow, asymptotically vanishing potential at large separation distances. Atomic examples arise from combinations of centrifugal and Coulomb (or dipole) potentials, whereas a nuclear example results in alpha-particle radioactivity from

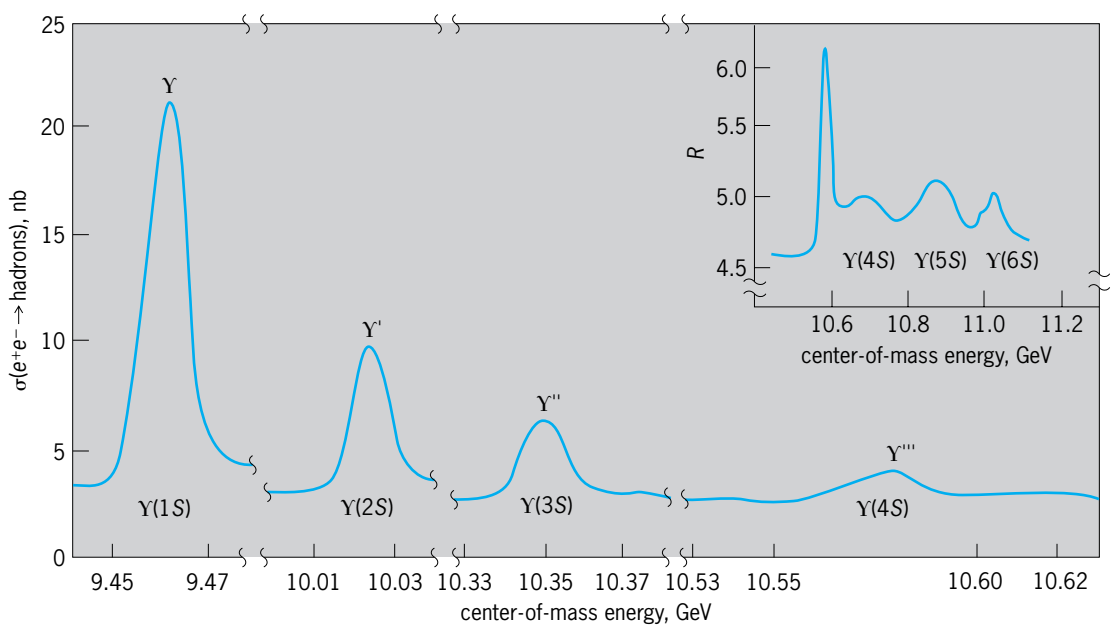

Fig. 3. Quasibound quark-antiquark states called upsilon, seen as resonances in the collision cross section (σ) of electrons and positrons. Data from two different detectors, one of which is shown in an inset as the ratio $R = \sigma(e^+e^- \rightarrow \text{hadrons})/\sigma(e^+e^- \rightarrow \text{muons})$, show six resonances. 1 nb = 10^{-37} m^2. (*After D. Besson et al., Observation of new structure in the e^+e^- cross section above the $\Upsilon(4S)$, Phys. Rev. Lett., 54:381–384, 1985*)

the combination of attractive nuclear and repulsive Coulomb forces between the alpha particle and the residual daughter nucleus. When the inner well supports a bound state above the zero of the energy, this state appears as a quasibound resonance because the alpha particle or the electron can tunnel quantum-mechanically through the barrier and escape to infinity. The distinguishing characteristic of a shape resonance is, therefore, that it lies just above the threshold to which it couples, in contrast to a Feshbach resonance, which lies below it. *See* ALPHA PARTICLES; RADIOACTIVITY.

Multiple resonances. Whole sequences of resonances are often observed because a closed channel typically supports many quasibound energy levels (**Figs. 2** and **3**).

Another direction for generalization involves many coupled channels, some open and others closed. Energy levels in any closed channel resonate with degenerate energy states of all the open channels and show up as cross-section anomalies, elastic or inelastic. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS. A. R. P. Rau

Bibliography. A. de Shalit and H. Feshbach, *Theoretical Nuclear Physics*, vol. 1, 1974, reprint 1992; U. Fano and A. R. P. Rau, *Atomic Collisions and Spectra*, 1986; E. Merzbacher, *Quantum Mechanics*, 3d ed., 1997; R. G. Newton, *Scattering Theory of Waves and Particles*, 2d ed., 1982; D. H. Perkins, *Introduction to High Energy Physics*, 4th ed., 2000.

Resonance ionization spectroscopy

A form of atomic and molecular spectroscopy in which wavelength-tunable lasers are used to remove electrons from (ionize) a given kind of atom or molecule. Laser-based resonance ionization spectroscopy (RIS) methods have been developed and used with ionization detectors, such as proportional counters, to detect single atoms. Resonance ionization spectroscopy can be combined with mass spectrometers to provide analytical systems for a wide range of applications, including physics, chemistry, materials sciences, medicine, and the environmental sciences.

Theory. When an atom or molecule is irradiated with a light source of frequency ν, photons at this selected frequency are absorbed only when the energy $h\nu$ (h is Planck's constant) is almost exactly the same as the difference in energy between some excited state (or intermediate state) and the ground state of the atom or molecule. If a laser source is tuned to a very narrow bandwidth (energy width or frequency width) at a frequency that excites a given kind of atom (**Fig. 1**), it is highly unlikely that any other kind of atom will be excited. An atom in an excited state can be ionized by photons of the specified frequency, ν, provided that $2h\nu$ is greater than the ionization potential of the atom. Even though the final ionization step can occur with any energy above a threshold, the entire process of ionization is a resonance one, because the intermediate state must first be excited in a resonance photon absorption. In sharp contrast to all other ionization means, such as x-rays, energetic charged particles, or radioactive sources, resonance ionization spectroscopy is a selective process in which only those atoms that are in resonance with the light source are ionized. Modern pulsed lasers can provide a sufficient number of photons in a single pulse to remove one electron from each atom of any selected type. The first transition from the ground state to the intermediate state is easily saturated as these states come into equilibrium with a moderate photon flux. To saturate the final transition from the intermediate state to the ionization continuum (to remove one electron from each atom) requires laser pulses of considerable energy per pulse and is generally difficult. These frequency-tunable (or wavelength-tunable) lasers have made resonance ionization spectroscopy a potential method for the sensitive (and highly selective) detection of nearly every type of atom in the periodic table. *See* ATOMIC STRUCTURE AND SPECTRA; IONIZATION POTENTIAL; LASER; LASER SPECTROSCOPY; PHOTOIONIZATION; RESONANCE (QUANTUM MECHANICS).

Laser schemes. The simple type of resonance ionization spectroscopy discussed above (Fig. 1) was used to illustrate the resonance ionization process; more elaborate schemes are required in practice. Fortunately, many others are possible. An atom already in an excited state can absorb other photons of another wavelength to create a more highly excited state. Each kind of atom has its own unique set of possible excited states, much like a ladder with steps that are widely spaced near the ground and are more closely spaced as the steps converge to the ionization continuum. (An energy equal to the ionization potential is required to ionize a given atom; the liberated electron is no longer bound, and is thus in the ionization continuum.) After a few of these steps have been taken, another photon may be used to complete the resonance ionization process by the ejection of an electron with a small kinetic energy into the ionization continuum. Alternatively, if the final bound state

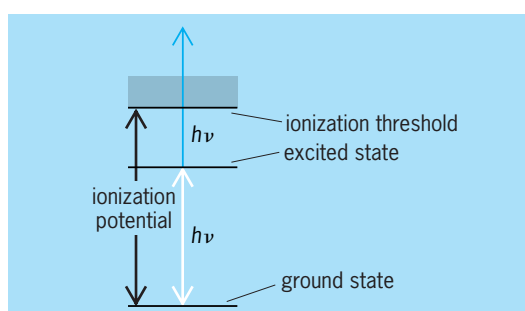

Fig. 1. Basic laser scheme for resonance ionization spectroscopy (RIS). The atom or molecule is irradiated by a light source with frequency ν, or photons of energy $h\nu$, where h is Planck's constant. The shading shows the beginning of the ionization continuum, which starts at several electronvolts above the ground state, depending on the type of atom. An electron from the ground state of the atom or molecule is being excited by two photons, thus resonantly ionizing the atom or molecule.

is very close to the edge of the continuum, a pulsed electrical field can be applied to eject the electron. Thus, a rich variety of laser schemes are in use, so that any atom in the periodic table can be detected with resonance ionization spectroscopy. However, those atoms with high ionization potentials, such as the inert gases, require the more difficult resonance ionization schemes. *See* ENERGY LEVEL (QUANTUM MECHANICS).

Ultrasensitive detection. Both high selectivity and the ultrasensitive detection of atoms have been demonstrated by pulsing a laser directly through a proportional counter. Proportional counters, like Geiger-Muller counters, will detect a single electron. Therefore, if a laser beam is directed through a proportional counter to remove one electron from each of the atoms of a selected type, selective detection of a single atom is possible. In the original demonstration of one-atom detection, it was proved that one atom of cesium could be selected out of 10^{19} atoms of the counting gas (argon and methane). *See* GEIGER-MÜLLER COUNTER; IONIZATION CHAMBER.

Applications. Resonance ionization spectroscopy is used in a wide variety of applications such as the trace analysis of low levels of elements in extremely pure materials, for example, semiconductors in the electronics industry. For these analyses, sputter-initiated resonance ionization spectroscopy (SIRIS) is achieved with the apparatus shown in **Fig. 2**. An argon ion beam sputters a tiny cloud of atoms from a sample placed in a high-vacuum system and a pulsed laser is tuned to detect the specified impurity atom. These lasers create positive ions (using the resonance ionization process) of a given atomic number within the cloud of neutral atoms released from the solid material. Ions are then further selected for atomic mass by using a mass spectrometer. The combination is doubly selective and eliminates isomers (atoms having the same mass but different atomic numbers) as well as certain backgrounds resulting from fragmentation of molecules by electron impact in the usual mass spectrometer. Trace elements present at the level of parts per billion and, in some cases, parts per trillion can be detected. Further, imaging capability provides maps of these impurities on the surface of materials, while depth profiling gives the concentration as a function of depth into the materials. This kind of information is of increasing importance to the development of very large scale integrated devices. *See* INTEGRATED CIRCUITS; MASS SPECTROMETRY; MASS SPECTROSCOPE; SPUTTERING.

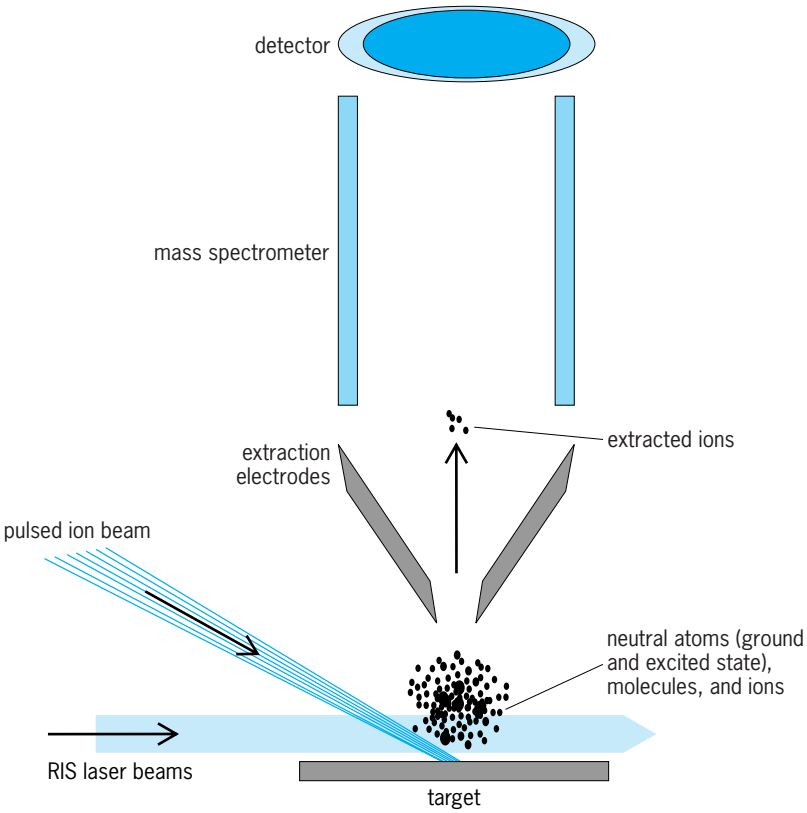

Fig. 2. Sputter-initiated resonance ionization spectroscopy (SIRIS). (*After H. F. Arlinghaus, M. T. Spaar, and N. Thonnard, Quantitative and sensitive profiling of dopants and impurities in semiconductors using sputter-initiated resonance ionization spectroscopy, J. Vac. Sci. Technol. A, 11:2317–2323, 1993, updated by T. Whitaker, Atom Sciences, Inc.*)

The sputter-initiated resonance ionization spectroscopy method is also used for chemical and materials research, geophysical research and explorations, medical diagnostics, biological research, and environment analysis. When imaging capability is not needed, a simpler technique known as thermal-atomization resonance ionization spectroscopy (TARIS) may be used (instead of SIRIS) for the bulk analysis of materials. By simply using resonance ionization spectroscopy with ionization chambers or proportional counters, gas-phase work can be done to study the diffusion of atoms, measure chemical reaction rates, and investigate the statistical behavior of atoms and molecules. *See* CHEMICAL DYNAMICS; DIFFUSION.

Resonance ionization spectroscopy is used in sophisticated nuclear physics studies involving high-energy accelerators. The spacing of atomic energy levels in atoms is governed mainly by the protons in the atom's nucleus and the orbiting electrons. However, the spin of electrons causes a perturbation of these levels known as the fine structure. Further, there is a minute effect due to the spin of the nucleus, which gives rise to a hyperfine structure. Resonance ionization spectroscopy is used as an on-line detector to record the hyperfine structure of nuclei with short lifetimes and hence to determine several nuclear properties such as nuclear spin and the shape of nuclei. *See* FINE STRUCTURE (SPECTRAL LINES); HYPERFINE STRUCTURE; NUCLEAR STRUCTURE.

In the usual ion source, even with magnetic separation, the ion beam is contaminated with other isobars (atoms with the same mass number but different atomic numbers). The use of resonance ionization spectroscopy in the ion source (**Fig. 3**) essentially eliminates the isobars and leads to much more definitive determination of nuclear parameters. Thus, resonance ionization spectroscopy is used at both ends of major accelerators for nuclear physics research. *See* ION SOURCES; ISOBAR (NUCLEAR PHYSICS).

An example of the potential of resonance

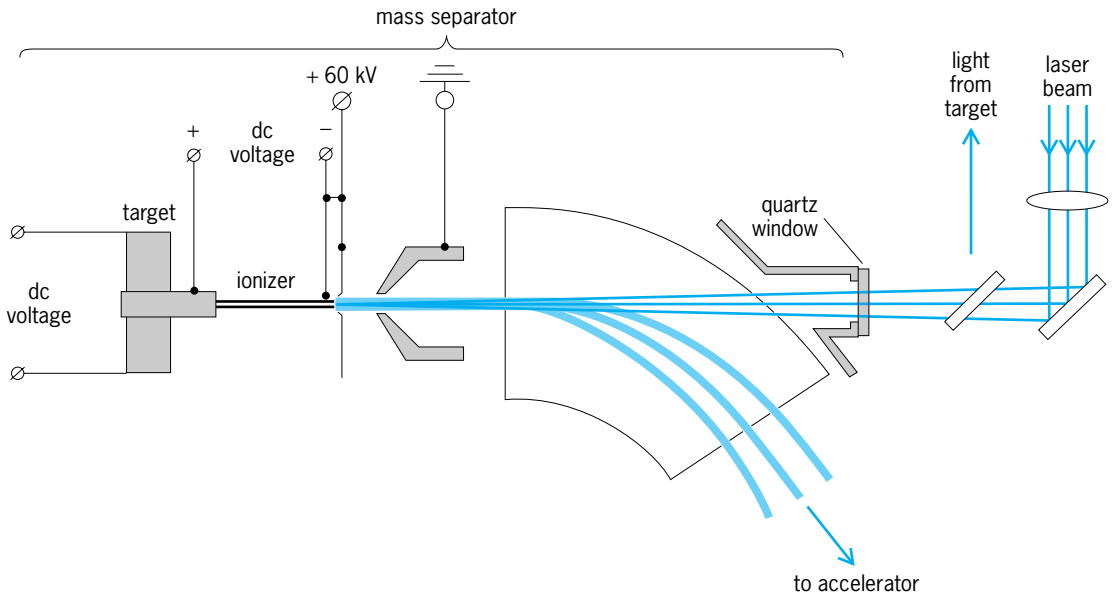

Fig. 3. Selective ion source using resonance ionization spectroscopy. The laser beam comes to a focus at the ionizer after being reflected by the right diagonal mirror. The left diagonal mirror is used for precision timing of the laser pulse. (*After V. I. Mishin et al., Chemically selective laser ion-source for the CERN-ISOLDE on-line mass separator facility, Nucl. Instrum. Meth. Phys. Res., B73:550–560, 1993*)

ionization spectroscopy for weak-interaction physics is the detection of krypton-81, which results from neutrino capture in bromine-81. A detector using large tanks of a bromine solution could be placed underground (reducing cosmic-ray backgrounds) to study neutrinos from the Sun. Resonance ionization spectroscopy could be used to count the few atoms of krypton-81 directly, before radioactive decay. Decay counting of isotopes with long half-lives (for krypton-81 it is about 200,000 years), for samples of 100–1000 atoms, is obviously impossible. *See* NEUTRINO; SOLAR NEUTRINOS; WEAK NUCLEAR INTERACTIONS.

Resonance ionization spectroscopy is used for measurements of krypton-81 in the natural environment to determine the ages of polar ice caps and old ground-water deposits. Oceanic circulation and the mixing of oceans could also be studied by measuring the concentrations of noble-gas isotopes by resonance ionization spectroscopy. *See* DATING METHODS; RADIOISOTOPE (GEOCHEMISTRY).

G. Samuel Hurst

Bibliography. G. S. Hurst and M. G. Payne, *Principles and Applications of Resonance Ionization Spectroscopy*, 1988; G. S. Hurst et al., Resonance ionization spectroscopy and one-atom detection, *Rev. Mod. Phys.*, 51:767–819, 1979; V. S. Letokhov, *Laser Photoionization Spectroscopy*, 1987.

Respiration

Commonly, the processes by which living organisms take up oxygen and eliminate carbon dioxide in order to provide energy. A more formal, comprehensive definition is the various processes associated with the biochemical transformation of the energy available in the organic substrates derived from foodstuffs to energy usable for synthetic and transport processes, external work, and, eventually, heat. This transformation, generally identified as metabolism, most commonly requires the presence of oxygen and involves the complete oxidation of organic substrates to carbon dioxide and water (aerobic respiration). If the oxidation is incomplete, resulting in organic compounds as end products, oxygen is typically not involved, and the process is then identified as anaerobic respiration. The oxygen required in aerobic respiration is ultimately derived from the atmosphere or from oxygen dissolved in water. The process of the exchange and transport of oxygen from an animal's environment to the intracellular sites involved in metabolism, and the reverse transport and exchange of the end product carbon dioxide, is often called respiration, although it is only a part of the overall process.

The term external respiration is more appropriate for describing the exchange of oxygen and carbon dioxide between the organism and its environment. In most multicellular organisms, and nearly all vertebrates (with the exception of a few salamanders lacking both lungs and gills), external respiration takes place in specialized structures termed respiratory organs, such as gills and lungs.

Diffusion and convection. The ultimate physical process causing movement of gases across living tissues is simple passive diffusion. Diffusion is the random movement of molecules from regions of high concentration (activity) to regions of low concentration. Fick's law describing this process states that the rate of transfer of a gas is directly proportional to its concentration gradient and to the surface area. The concentration gradient is the difference in the concentrations at the two ends of the diffusion path,

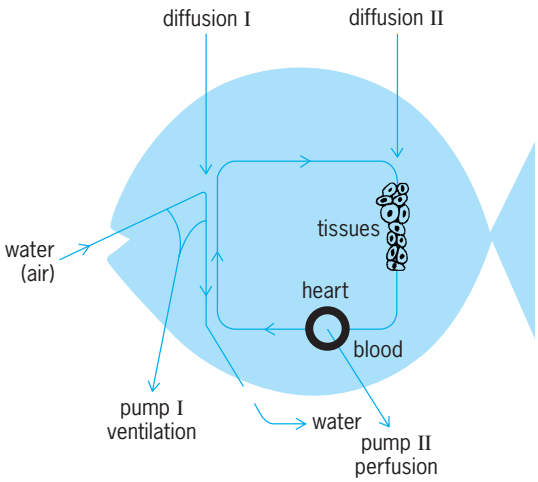

Fig. 1. Transport path for respiratory gases in a vertebrate. At diffusion site I, water or air exchanges gases with blood; at diffusion site II blood exchanges gases with metabolizing tissues. The rate of diffusion is increased by bulk convective transport of air or water and blood effected by a ventilatory pump (pump I) and a perfusion pump (pump II). The presence of hemoglobin in blood greatly increases its gas-transporting capacity and also facilitates the rate of diffusion across diffusion sites I and II. (*After A. G. Kluge, ed., Chordate Structure and Function, 2d ed., Macmillan, 1977*)

divided by the length of the path. Thus, the rate of gas transfer is inversely proportional to the tissue thickness (diffusion distance). Moreover, the rate of diffusion of gases through tissues is much slower than through water or air.

Respiratory gas exchange also depends on two convective fluid movements. The first is the bulk transport of the external medium, air or water, to and across the external respiratory exchange surfaces. The second is the transport of coelomic fluid or blood across the internal surfaces of the respiratory organ. These two convective transports are referred to as ventilation and circulation (or perfusion), respectively. They are active processes, powered by muscular or ciliary pumps. In addition to providing bulk transport of the respiratory gases, the convection pumps are important in maximizing the concentration gradients for diffusion across the gas-exchanging tissue surfaces. Aquatic organisms living in air-saturated water, and supplied with oxygen by diffusion alone, cannot have a radius greater than about 0.02 in. (0.5 mm). Consequently, larger organisms that depend on aerobic energy production must be equipped with external and internal convective pumps.

In almost all vertebrates and many invertebrates, the circulating internal medium (blood, hemolymph, or coelomic fluid) contains a respiratory pigment, for example, hemoglobin or hemocyanin, which binds reversibly with oxygen, carbon dioxide, and protons. Respiratory pigments augment respiratory gas exchange, both by increasing the capacity for bulk transport of the gases and by influencing gas partial pressure (concentration) gradients across tissue exchange surfaces. *See* HEMOGLOBIN; RESPIRATORY PIGMENTS (INVERTEBRATE).

Respiration in a vertebrate illustrates well the transport chain that moves respiratory gases between the external medium and the intracellular metabolic sites (**Fig. 1**). The links in this chain are the two bulk flows, ventilation and perfusion, with their convective pumps; diffusion across the specialized exchange surfaces in the respiratory organs (lungs or gills) connecting ventilation and perfusion; and diffusion across the tissue capillaries to and from the sites of metabolism. The delivery, by this transport chain, of the molecular oxygen needed for aerobic metabolism is often referred to as the oxygen cascade (**Fig. 2**). In addition to the steps in the oxygen cascade, needed chemical reactions may be rate-limiting in respiratory gas exchange, particularly at the lower body temperatures of many poikilotherms. Two of these are the reaction of oxygen with hemoglobin and the dehydration of carbonic acid (H_2CO_3), the latter depending on the enzyme carbonic anhydrase.

The physiological adjustment of organisms to variations in their need for aerobic energy production involves regulated changes in the exchange and transport of respiratory gases. The adjustments are effected by rapid alterations in the ventilatory and circulatory pumps and by longer-term modifications in the respiratory properties of blood.

Respiratory media. Air and water differ qualitatively as respiratory media, and the differences are reflected in the structural and functional qualities of the external gas exchangers (respiratory organs) of aerial and aquatic breathers.

The primary quality of a respiratory medium is its capacitance for oxygen and carbon dioxide. The capacitance coefficient is defined as the increment in

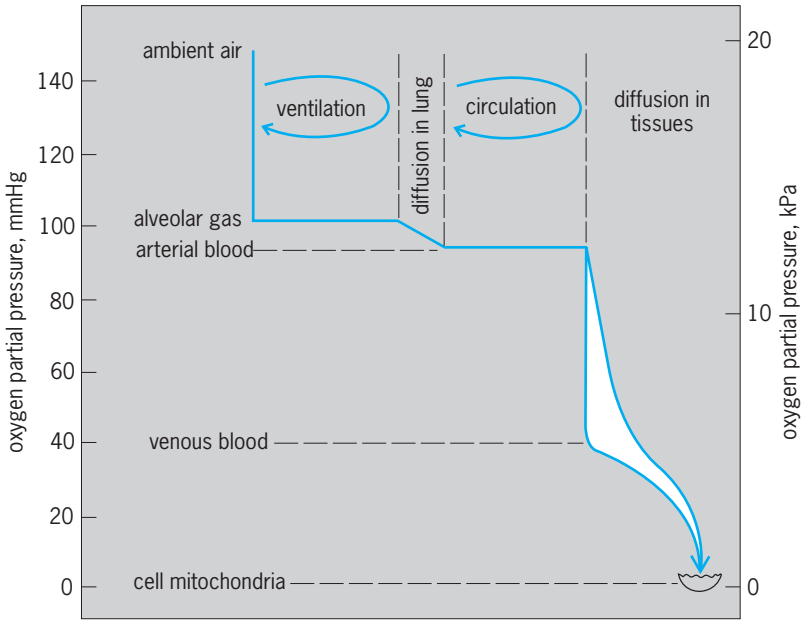

Fig. 2. Oxygen cascade from the ambient medium to the cellular metabolic sites in the mitochondria for a lung breathing animal. The oxygen partial pressure gradient is determined by the ventilation and circulation and by diffusion in the lung and tissues. 1 mmHg = 0.04 in. Hg. (*After S. C. Wood and C. Lenfant, eds., Evolution of Respiratory Processes, vol. 13 in Lung Biology in Health and Disease, Marcel Dekker, 1979*)

concentration per increment in partial pressure. The capacitance coefficient for oxygen in water (equal to the physical solubility) is very low compared to air and depends on temperature and salinity. Due to the low capacitance of water for oxygen, aquatic breathers must ventilate 20–30 times more medium than air breathers to obtain the same amount of oxygen. The carbon dioxide capacitance of water and internal fluids, like blood, exceeds that for oxygen. Therefore, an aquatic breather will increase carbon dioxide tension in the ambient medium far less than it decreases oxygen tension, assuming that the changes are caused only by the animal's respiratory gas exchange. Similarly, the high ventilatory rate characteristic of aquatic gas exchange will be accompanied by very small differences between inspired and expired carbon dioxide tensions, and will cause low internal carbon dioxide tensions in body fluids and tissues. In seawater, and in some special freshwater habitats, carbonate and bicarbonate, or other buffer substances, increase the carbon dioxide capacitance by varying in concentration with changes in the carbon dioxide tension. Similarly, the chemical binding of the respiratory gases, oxygen and carbon dioxide, in blood containing a respiratory pigment will greatly increase the capacitance coefficient of both gases. Due to a number of properties of respiratory pigments, the capacitance coefficient for oxygen in blood will depend on the oxygen tension. Similarly, the capacitance coefficient for carbon dioxide will depend nonlinearly on the buffering properties of blood, and on the carbon dioxide tension. *See* DIVING.

Oxygen consumption. Most animals satisfy their energy requirements by oxidation of organic molecules obtained in food. Thus, the rate of oxygen uptake reflects the energy metabolism, or the metabolic rate, of an aerobic organism. The energy made available and transformed to heat when food stuffs are completely oxidized corresponds to 5.0 kilocalories (20.9 kilojoules) per liter oxygen for carbohydrate, 4.5 kcal (18.8 kJ) for protein, and 4.8 kcal (20.1 kJ) for fat. The respiratory quotient (RQ), which is calculated as the ratio between carbon dioxide produced and oxygen consumed in metabolism, varies between 0.7 for fat and 1.0 for carbohydrate. The RQ value is therefore an index to the foodstuffs being used by an organism. In the context of dynamics of external gas exchange, the RQ is often called the respiratory exchange ratio (RE). In these circumstances it implies how oxygen uptake and carbon dioxide output are divided between respiratory organs in animals depending on multiple modes of external gas exchange (such as amphibians).

The oxygen consumption of organisms at rest varies with their phylogenetic position, with their size, and with oxygen availability in their immediate environment. In addition, oxygen consumption increases with their food intake and during their physical activity. Homeothermic vertebrates, including birds and mammals, have much higher oxygen uptake rates than poikilothermic vertebrates, even when the comparison is made at similar body temperatures (98.6–104°F or 37–40°C). *See* HIBERNATION AND ESTIVATION.

Organisms with stable oxygen uptake rates over wide ranges of external oxygen concentration are classified as oxygen regulators, or as being oxygen-independent. In contrast, an oxygen uptake more or less linearly related to external oxygen availability is termed oxygen dependence, and the animal an oxygen conformer. Whereas homeotherm vertebrates are usually oxygen-independent, many aquatic invertebrates show total oxygen dependence. The critical oxygen tension is the environmental partial pressure at which an organism shifts from oxygen independence to oxygen dependence. For poikilotherm organisms, the critical oxygen tension varies with temperature.

For nearly all animal groups, the metabolic rate per unit body weight, or the weight-specific metabolic rate, decreases with increasing body weight. This implies that the metabolic requirements for oxygen and organic substrates in food are relatively much higher for smaller organisms than for larger ones. Resting metabolic rate (M) is defined by the equation $M = aW^b$, where W is body weight, and a and b are constants, with b varying between 0.6 and 0.9. The weight-specific metabolic rate is therefore expressed by $M/W = aW^{b-1}$.

The aerobic scope for activity is defined as the difference between maximal oxygen uptake and resting oxygen uptake. For poikilotherms, the scope for activity is greatly increased by temperature, up to a certain level. Homeotherms have aerobic scopes for activity many times greater than poikilotherms.

Respiratory organs. Nearly all respiratory organs of water-breathing animals are characterized by a unidirectional passage of water across the respiratory exchange surfaces. This design is basically different from that of the respiratory organs of air breathers where air typically enters and exits through the same channels as a tidal or bidirectional flow. The significance of this difference is that anatomical dead space is virtually eliminated in water breathers, and much higher oxygen tensions occur on the external respiratory surfaces than is possible in a tidal system where the conducting airways constitute dead space.

The gas exchange in the respiratory organs of postembryonic vertebrates and many invertebrates can conveniently be classified according to the four exchange types illustrated in **Fig. 3**.

The unidirectional ventilatory current in gills of fishes and many invertebrates offers a structural basis for a countercurrent exchange of gases between water and blood. Such a system is highly efficient and allows arterial oxygen partial pressure (Fig. 3*a*) to exceed the expired oxygen partial pressure (*e*) and, in theory, to approach the oxygen partial pressure in external water (*i*). The respiratory organs of birds are unique in design among vertebrates; the system includes air sacs that serve to ventilate the gas exchange units of the lung, the parabronchi. *See* COUNTERCURRENT EXCHANGE (BIOLOGY).

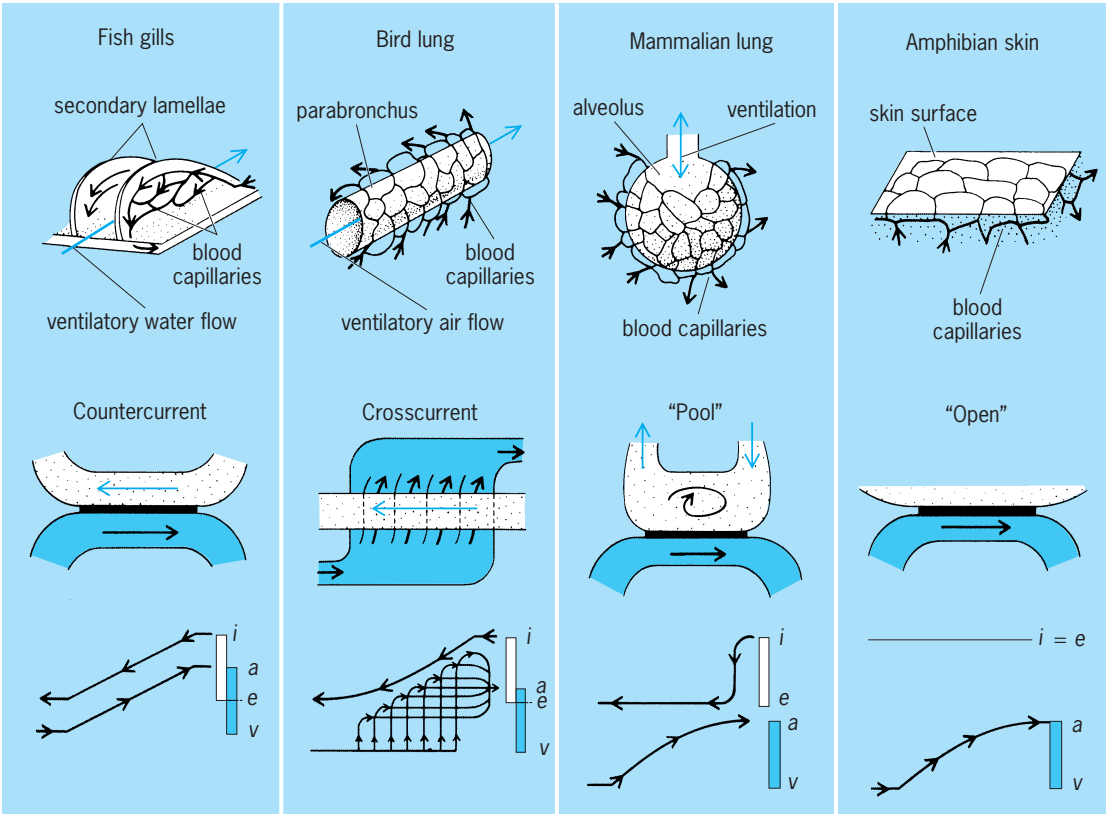

Fig. 3. Most common external gas exchangers with models of their exchange types. At the bottom, typical curves of oxygen tension changes in the external medium and blood along the exchange areas *i*, inspired; *e*, expired; *a*, arterial; *v* venous. (*After J. Piiper and P. Scheid, Comparative physiology of respiration: Functional analysis of gas exchange organs in vertebrates, Int. Rev. Physiol., vol. 14, 1977*)

Gas exchange in the parabronchi of the bird lung can best be analyzed as a crosscurrent exchange system (Fig. 3). In the parabronchi, as in fish gills, ventilation occurs in a flow-through, one-way manner. Gas exchange occurs by diffusion along blind-ending air capillaries radiating from the parabronchial lumen. Blood capillaries are profusely intertwined with the air capillaries. The partial-pressure profile for gas exchange in the parabronchial lung documents a high gas exchange efficiency, with larger arterial oxygen partial pressures than at the end of the parabronchus.

Mammalian lungs are referred to as alveolar lungs, and are homologous with the lungs in lungfishes, amphibians, and reptiles. These lungs have bidirectional conducting airways that, in the most advanced mammalian design, terminate in the blind-ending alveoli after 20–25 subdivisions (branchings). The lungs of humans may contain about 3×10^8 alveoli, representing a surface area of more than 860 ft^2 (80 m^2). Gas in the alveoli, and blood inside the profuse network of capillaries investing the alveolar walls (Fig. 3), are separated by less than 0.5 micrometer of intervening tissue. In lungfishes, amphibians, and reptiles an ascending degree of internal septation of the lung is the basis for the corresponding surface expansion of the alveolar surfaces. A comparison of vertebrate lungs shows lung volume varying in direct proportion to body mass, and lung surface area positively correlated with metabolic rate, physical activity, and environmental factors. Increased lung surface area results when the number of alveoli increases and their individual size diminishes. The gas-to-blood diffusion distance also gets shorter in smaller or more active animals, correlating with their higher weight-specific metabolism. The bidirectional airways in lung breathers create a nonfunctional air space, called the anatomical dead space. It causes oxygen partial pressures in alveoli to be lower, and carbon dioxide partial pressures higher, than in outside air. Compared to fish gills, the gas exchange surfaces of lungs may be subject to smaller, more stable gas partial pressure gradients. The gas exchange in alveolar lungs can be modeled as a ventilated pool, and the oxygen partial pressures in the blood leaving a lung can never, as in fish gills or bird lungs, be higher than that in the medium leaving the gas exchange unit (Fig. 3). *See* LUNG.

The muscular basis for ventilation in lungfish and amphibians is a buccal force pump bringing air into the lungs under positive pressure; in reptiles, birds, and mammals the lungs are ventilated by aspiration or suction.

Breathing through the general body surface is phylogenetically the most primitive method, and it remains an important complement to external gas exchange for most aquatic invertebrates, fishes, and amphibians. In some amphibians, the skin is the sole site

of gas exchange. Skin gas exchange can be modeled as an infinite or open pool, typically with no organized convective movement of the external medium. The capillary network in the skin is not as well developed for exchanging gases with the external medium as are the capillaries in lungs, or in the lamellar blood spaces of fish gills. Moreover, the diffusion distance between the blood and the external medium in skin is longer than that in gills or lungs. Since skin is an unspecialized gas exchanger, it also lacks the vascular connections that are used for optimizing internal gas transport and that are characteristic of specialized organs. Of the four gas exchange models in Fig. 3, the exchange efficiency declines in the order: countercurrent exchange (fish gills), crosscurrent exchange (bird lungs), ventilated pool (alveolar lungs), and infinite pool (general body surface).

Tracheal respiration. In insects, the largest and most diversified order of terrestrial animals, gas exchange depends largely on diffusion. The basic design is a highly branching internal system of air-filled tubes (trachea) connected to the ambient air through controllable openings (spiracles). The smallest air tubes (tracheoles) may be only 1 μm in diameter; they ramify extensively, and even penetrate individual muscle fibers. In smaller insects, oxygen is transported through the tracheal system solely by diffusion; larger insects depend on convective air movement produced by ventilation. Insect flight muscles are strictly aerobic and during flight may consume more oxygen than any other animal tissue; the oxygen uptake increases 400-fold when activity begins. In insects, internal transport by the hemolymph is of no importance for respiratory gas exchange, but the organic substrates that are needed to support the high and variable aerobic metabolism must be delivered to the cells by hemolymph circulation. *See* INSECT PHYSIOLOGY.

Smaller insects may have an air-filled internal tracheal system with no external openings, so gas exchange must occur by diffusion through the body wall. In some aquatic insects the closed-off air-filled tracheal system may project into gill-like structures. These tracheal gills expand the general body surface and thereby augment the diffusive exchange of gas between the surrounding water and internal tracheal air. In the absence of an effective microvascular circulation of hemolymph, the usefulness of the system of closed air tubes for internal gas transport lies in the 250,000-times-more-rapid diffusion of gases in air than in water.

Other aquatic insects with an air-filled tracheal system make regular excursions to the surface to capture a bubble of air that is connected to the tracheal system and held in place by a specialized structure (plastron). When the insect dives and the air bubble becomes compressed, gases will exchange between the insect and the bubble. This exchange, and the large difference in the solubilities of oxygen and carbon dioxide in water, will set up partial pressure differences between the bubble and the surrounding water, favoring efflux of carbon dioxide over influx of oxygen. This system will function for a long time as a diffusion lung, but since carbon dioxide escapes the bubble more easily than oxygen that enters it from the water, the bubble will shrink, and its N_2 pressure will rise. The bubble will therefore gradually lose N_2 to the water, and the insect must return to the surface, to renew its N_2 supply and not its oxygen supply.

Blood circulation and gas transport. In animals using blood circulation for exchange and transport of respiratory gases, the efficiency of the transport function is highest if the vascular system conveys the most deoxygenated, carbon dioxide–rich blood to the respiratory organs, gills or lungs, and if the flow path of oxygen-rich blood, from the respiratory organ to the metabolizing cells, is direct and with no admixture of venous blood. The carbon dioxide–rich blood increases the oxygen affinity of hemoglobin and therefore assists loading of oxygen in the lung, while the carbon dioxide–depleted blood with its reduced oxygen affinity enhances oxygen delivery in peripheral tissues. Among vertebrates, this scheme is present in water-breathing fish, birds, and mammals. Fish have a single circulation, with the gill perfusion in direct series with the perfusion that serves metabolism (systemic perfusion). Birds and mammals have a double circulation, with the pulmonary and systemic circulations connected in series by the two cardiac pumps, that is, the right and left ventricles. In vertebrate phylogeny, the transition from a single to a double circulation arose concomitantly with the transition from aquatic to aerial breathing. *See* CIRCULATION.

The hearts and vascular systems of lungfishes, amphibians, and reptiles are incompletely divided, providing for selective passage of blood in the pulmonary and systemic circulations. The regulated shunting of blood between these two circuits is associated with the phenomenon of periodic breathing. During breathing a nearly complete selective passage of blood occurs between the pulmonary and systemic circuits; but during the often long periods of breath-holding (apnea), blood is increasingly shunted from the pulmonary to the systemic circuit, and the circulation to the lungs is gradually bypassed. This pattern of alternating breathing and shunting ensures that blood circulation is adjusted to the oxygen availability at the gas-exchanging surfaces of the lungs. Such matching of blood perfusion and ventilation (that is, oxygen availability) is important for gas exchange efficiency in all animals in which the exchange depends on active ventilation and circulation.

The higher metabolic rate in birds and mammals accords with a higher pulmonary ventilation that stabilizes alveolar oxygen availability. A complete separation of the pulmonary and systemic circulation is therefore optimal for these animals. *See* CARDIOVASCULAR SYSTEM.

Diffusing capacity. Since surface areas and diffusion distances cannot readily be measured in live animals, gas exchange efficiency is often expressed by the diffusing capacity of the individual gases. For aquatic animals, diffusing capacity is often referred

to as the transfer factor. Both terms are defined as the amount of a gas (such as oxygen) transferred across the respiratory exchange surfaces per unit time and per unit of pressure difference across the gas exchange surfaces. The diffusion capacity of both gills and lungs are positively related to the rates of oxygen uptake of an animal. And this relationship holds whether the oxygen uptake level is set by the genetic or phylogenetic status of the animal, or is responding to behavioral patterns (such as physical activity) or to environmental factors (such as oxygen availability). *See* METABOLISM; RESPIRATORY SYSTEM.

Kjell Johansen; John B. West

Bibliography. R. G. Crystal et al. (eds.), *The Lung: Scientific Foundations*, 2d ed. (2 vols.), Raven Press, 1997; E. R. Weibel, *The Pathway for Oxygen*, Harvard University Press, Cambridge, 1984; J. B. West, *Respiratory Physiology: The Essentials*, 7th ed., Lippincott Williams & Wilkins, Philadelphia, 2004.

Respirator

A device designed to protect the wearer from noxious gases, vapors, and aerosols or to supply oxygen or doses of medication to the wearer. Respirators are used widely in industry to protect workers against harmful atmospheres, and in the military to protect personnel against chemical or biological warfare agents. **Figure 1** shows the general classification of respirators according to whether they are atmosphere-supplying or air-purifying.

The hood, helmet, and facepiece provide the essential interface between the respirator and the wearer. Quarter-mask facepieces cover the wearer's nose and mouth above the chin; half-mask facepieces cover the nose and mouth and extend below the chin; and full facepieces cover the nose, mouth, and eyes. Facepieces made of elastomeric material or flexible fabric are equipped with a head harness or straps designed to secure the device to the wearer.

Important to the industrial application of respirators is a program that includes the monitoring of the respiratory hazard, physiological and psychological limitations of the wearers, selection of properly approved equipment, training of personnel, and equipment maintenance. Facepiece selection for a satisfactory fit is necessary to guarantee low face-seal leakage for negative-pressure respirators.

Atmosphere-supplying respirators. Atmosphere-supplying respirators are used in atmospheres deficient in oxygen or extremely hazardous to the health of the wearer. Such atmospheres can occur in unventilated cellars, wells, mines, burning buildings, and enclosures containing inert gas. In the United States the minimum legal requirement for such a respirator is a supply of oxygen of 19.5% or greater by volume to the wearer, compared with 21% existing naturally in pure air.

Self-contained breathing apparatus (SCBA). This is a completely self-contained unit with the air supply or the oxygen-generating material being carried by the wearer. It may be of the open-circuit (**Fig. 2**) or closed-circuit type. A closed-circuit SCBA (rebreather) involves recycling the exhaled air through sodium peroxide to remove the exhaled carbon dioxide, and adding oxygen to the air from an oxygen bottle.

Air-supplied respirators. Equipped with the same variety of facepieces as the SCBA, these respirators can have the air supplied to the facepiece by means of a hose and a blower—the hose mask—or from a compressed-air source equipped with proper airflow and pressure-regulating equipment—the air-line mask. Helmet or hood air-line respirators can afford temperature control with an unencumbered wide-angle view. A highly specialized version of an air-line respirator is the medical respirator used for artificial respiration, or for the delivery of oxygen and medication. The aviator mask is another example.

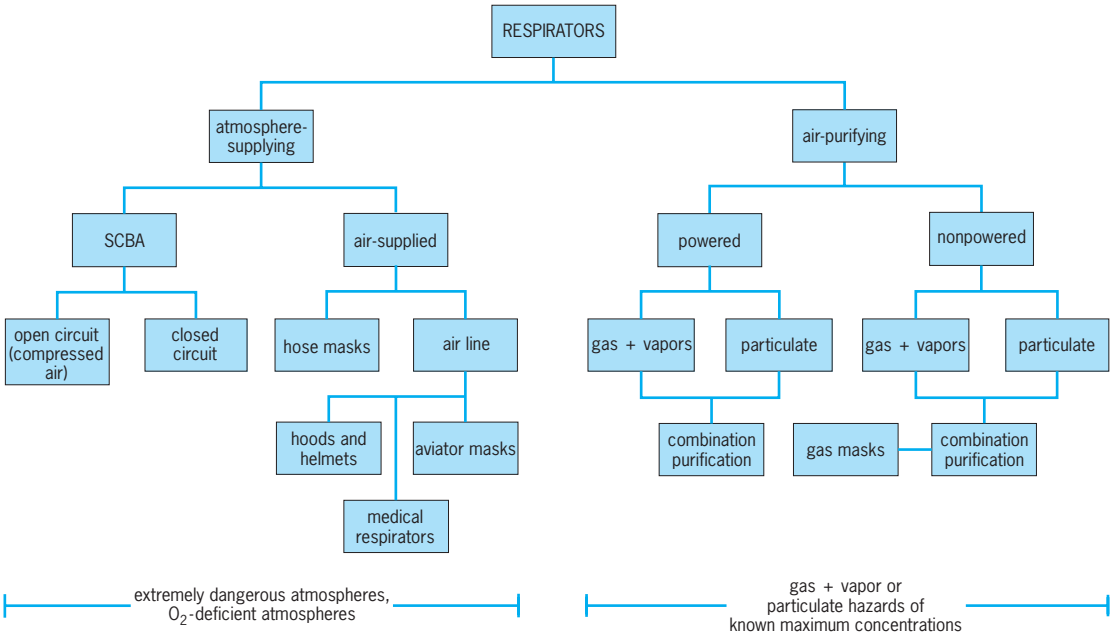

Fig. 1. General classification of respirators according to design and application.

Air-purifying respirators. In an air-purifying respirator, ambient air is passed through a purifying medium to remove the contaminants. However, these devices do not provide oxygen or protect against oxygen-deficient atmospheres.

Nonpowered. A widely used air-purifying respirator is the nonpowered, or negative-pressure, respirator (**Fig. 3**). Ambient air is inhaled through the purifying medium in the replaceable cartridges and exhaled through an exhaust valve. Sometimes the air is exhaled back through the medium if an exhaust valve is not provided. Because of the negative air pressure created inside the facepiece during inspiration, proper fit to the wearer is important to prevent air leakage through the face seal.

An increasingly popular form of nonpowered, air-purifying respirator is the disposable mask (**Fig. 4**). Here, the entire half-mask facepiece is made of the respirator filter material. It is light in weight and can be discarded after use.

Powered. In the case of the powered air-purifying respirator, an external blower, usually powered by a belt or helmet-mounted battery pack, forces air through the purifying medium and supplies it to the wearer under positive pressure, thus minimizing the problem of face-seal leakage. The purified air can be supplied to the wearer through either a hood or a helmet-face shield, which provides ample room for beards, eyewear, and hair styles that may otherwise cause face-seal leak problems in facepiece devices. While the design of powered respirators eliminates the inhalation airflow resistance of nonpowered respirators for easy breathing, powered respirators are larger and more complicated to use. Belt or helmet-mounted models provide added mobility over atmosphere-supplying respirators.

Particle removal. The filter medium used in a respirator for particle removal is usually selected according to the type of hazard against which the respirator is designed to protect.

Particulate hazards. These include dusts, mists and sprays, and fumes.

Dusts are solid particles that are mechanically produced by activities such as grinding, drilling, or sanding. Dusts range in size from submicrometer to tens of micrometers and can cause lung disease such as fibrosis (silica), chronic irritation (acids) and allergies (pollen, spices), and diseases from systemic poisons (cadmium) and carcinogens (asbestos, radioactive particles).

Mists and sprays are liquid particles produced by condensation (mists) or mechanical means (sprays). They may range in size from submicrometer to a few micrometers, and can cause respiratory illness such as major chemical irritation (smoke containing acidic condensate), systemic poisoning (spray paint), and chronic respiratory ailments (cigarette smoke, smog).

Fumes are solid particles produced by condensation, generally of metal or metal oxides with diameters in the submicrometer range. They can cause systemic poisoning (lead-welding fumes) or chills followed by fever (zinc or copper fumes).

Fig. 2. Open-circuit self-contained breathing apparatus (SCBA) atmosphere-supplying respirator.

Filters. Respirator air filters for particle removal fall into three general classes. Filters for dust or mist removal generally consist of a filter medium of moderate efficiency. They are made of fiber mats designed to provide particulate protection while giving low breathing resistance. A common dust/mist filter material is wool felt mechanically impregnated with modified phenolic resin to produce electrostatic capture sites. Another material is fine-fiber (about 2 μm in diameter) polyolefin web formed over a supporting facepiece shell.

Filters intended to remove dust, mist, and fume require a filter medium of moderate to high efficiency. Cartridge fume filters are made of fine-fiber matting or paper pleated to give a large extended filter area, since the loading of the filter with fume particles can cause rapid increases in breathing resistance. Noncartridge dust/mist/fume filters (Fig. 4) utilize

Fig. 3. Negative-pressure air-purifying respirator.

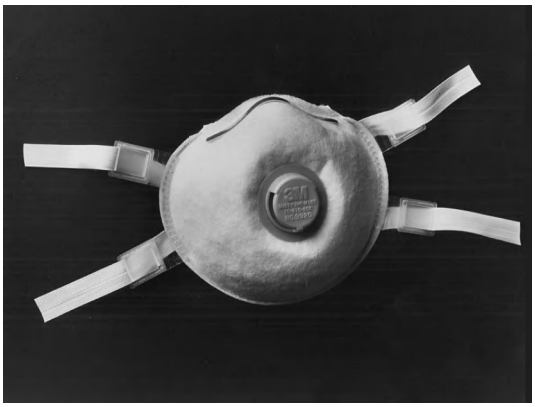

Fig. 4. Disposable negative-pressure air-purifying respirator.

a combination of very fine fibers and electrostatics in a very open matrix to provide high efficiency with low breathing resistance during particle loading on the filter.

High-efficiency particulate air (HEPA) filters are made of a special filter medium intended to remove nearly all airborne particles. The filters are generally made of very fine fibers, and the filter paper is pleated to provide a large extended surface area while giving an acceptable but moderately high breathing resistance. Such filters capture virtually all airborne particles and are used in nuclear installations or other critical applications and in the presence of extreme hazard. *See* AIR FILTER.

Gas and vapor removal. Media for gas and vapor removal are selected according to their service life performance against the specific type of gas or vapor hazard present. Gas and vapor hazards generally fall into the following major categories: (1) acidic (substances that are acids or react with water to produce acids, for example hydrogen chloride, fluorine, and hydrogen sulfide); (2) alkaline (substances that are alkalies or react with water to produce alkalies, for example ammonia, amines, and phosphenes); (3) organic (compounds of carbon such as hydrocarbons, alcohols, halides, isocyanates, and aromatics); (4) organometallic (metals chemically bonded to organic groups, for example tetraethyllead and organic phosphates); and (5) hydride (compounds in which hydrogen is chemically bonded to metals or metalloids, for example diborane and tetraborane).

Activated carbon in a granular form is used widely as an adsorbent to remove organic vapors. Vapors of high-molecular-weight substances are readily adsorbed onto the carbon surface, while low-molecular-weight gases or vapors, such as carbon monoxide, must be decomposed to be rendered harmless.

For the removal of acidic or alkaline gases, activated carbon is often modified to provide sites for a chemical reaction to occur with the gas. Metal oxides or caustic impregnation of the carbon can make it effective for removing acidic gases, while acidic impregnation of the carbon can make it effective against alkaline gases. Sometimes, activated carbon is simply mixed with a chemical adsorbent in a granular form to provide the necessary protection against the gas.

Gases or vapors that are not readily adsorbed either physically or chemically, such as carbon monoxide or formaldehyde, must be decomposed to render them harmless. This can be achieved by impregnating carbon with a catalyst such as a metallic salt or oxide. *See* ACTIVATED CARBON; ADSORPTION.

Cartridges and combination filters. To provide protection against both vapor and particles, activated carbon is often combined with a particulate filter to form a cartridge which can then be attached to the facepiece. Specially designed nonwoven facepiece respirators are available which use particulate filter media that are layered or loaded with gas- and vapor-adsorbent material for both vapor and particle removal.

Military gas masks. These are specialized devices which protect the wearer from the chemical, biological, and radioactive challenges possible in warfare. Some designs provide the wearer with tubes and other means to eat and drink without doffing the respirator. Included with the military mask is a storage bag or carrier that keeps the respirator and adsorbent media from contamination when not in use.

Benjamin Y. H. Liu; Daniel A. Japuntich

Bibliography. American National Standards Institute, *Practice for Respiratory Protection*, ANSI788.2-1980, 1980; D. S. Blackwell and G. S. Rajhans, *Practical Guide to Respirator Usage in Industry*, 1988; Los Alamos Scientific Laboratory, *Guide to Industrial Respiratory Protection*, 1977; U.S. Nuclear Regulatory Commission, *Manual of Respiratory Protection Against Airborne Radioactive Materials*, 1976.

Respiratory pigments (invertebrate)

Colored, metal-containing proteins that combine reversibly with oxygen, found in the body fluids or tissues of multicellular invertebrate animals and microorganisms. The role of these pigments is primarily to aid in the transport of molecular oxygen. Thus they are distinguished from respiratory enzymes, which are concerned with the metabolic consumption of oxygen. Four distinctly colored groups of respiratory pigments exist among invertebrates: hemoglobins (purple, become orange-red with oxygen), chlorocruorins (green, become red with oxygen), hemocyanins (colorless, become blue with oxygen), and hemerythrins (colorless, become red with oxygen). Formerly, invertebrate hemoglobins were called erythrocruorins to distinguish them from functionally similar yet structurally distinct pigments of vertebrate bloods. Those hemoglobins confined to muscle cells are called myoglobins. *See* HEMOGLOBIN.

Each pigment is composed of two parts, a large protein molecule to which is bound one or more small moieties called prosthetic groups, each of which is or contains a metal. The metal binds the oxygen, and this binding imparts the characteristic

color to the pigment. In hemoglobins the prosthetic group is an iron porphyrin compound called heme. Chlorocruorin contains a similar iron porphyrin which differs from heme only in that a vinyl group in the molecule is replaced by formyl. The prosthetic group of hemerythrin consists of two adjacent iron atoms which bind an oxygen molecule between them. The prosthetic group of hemocyanin is analogous and consists of two adjacent copper atoms. Pigments containing vanadium have been found in tunicates, but these substances do not combine reversibly with oxygen and so cannot be considered respiratory pigments.

The protein part of the pigment confers reversibility upon the combination of the metal with oxygen. In the absence of protein, the prosthetic groups lose their capacity to combine with oxygen reversibly. Instead, the metals are irreversibly oxidized: Electrons are transferred from metal to oxygen. The bonds between metal and protein so alter the electronic energy levels of the metal that this transfer, if it occurs, is reversible. For this reason, the combination of hemoglobin with oxygen is described as oxygenation rather than oxidation. The protein is also responsible for certain physiological adaptations of the pigment to the environment. Thus the affinity of the pigment for oxygen is often highest in animals that inhabit environments with the lowest oxygen content.

Hemoglobins. Hemoglobin molecules of invertebrates are usually relatively small, with molecular weights of 65,000 or less, and enclosed in cells, or else they are extracellular and very large with molecular weights ranging from 400,000 (in *Daphnia*) to $8–12 \times 10^6$ (in *Cardita*, a marine bivalve mollusk). The size of the minimum functional subunit of many of these pigments has not been accurately determined, but the hemoglobin from the perienteric fluid of *Ascaris* has a molecular weight of about 41,000 and has only one heme. The very large hemoglobins are composed of a large number of subunits and may contain as many as 200 or more heme-containing protein subunits. The stability of these assemblages often depends on the presence of divalent cations such as calcium. The confining of the smaller molecules to cells may prevent their loss and may also provide a local ionic environment favorable to their stability. It is also possible that the large hemoglobin molecules in free solution in the body fluids of various marine animals help to maintain an osmotic balance of salts. *See* ANNELIDA.

Distribution and function. Hemoglobins are widely distributed among invertebrates as well as among almost all vertebrates. Those pigments found in the body fluids, whether in cells or in free solution in the plasma, function primarily as agents that transport molecular oxygen to the tissues. Two barriers exist to this transport: One is at the surface of the animal (skin and gill membranes) and the other is within the tissue cells, where the oxygen must move from the cell surface to the respiratory enzymes of the mitochondria that consume the oxygen. Transport across these barriers depends on the physical process of diffusion, which is extremely slow and often is entirely inadequate by itself to supply the metabolic needs for oxygen. However, if hemoglobin is present within these barriers, diffusion of oxygen is greatly facilitated because the total capacity of the cell for oxygen is much greater. Hemoglobins that function in this way usually have high affinities compared to the affinities of blood or fluid hemoglobins. Myoglobins apparently work in this way: They pick up oxygen at the muscle cell surface and facilitate its diffusion to the mitochondria adjacent to the contractile muscle fibers. Cellular hemoglobin is found in the brain of the predatory nemertean worm *Cerebratulus*, where it acts as an oxygen store to enable the worm to invade anoxic marine muds for short periods of time in search of other invertebrates. *See* MITOCHONDRIA.

Oxygen-carrying capacity. The oxygen-carrying capacity of blood is greatly increased if it contains hemoglobin. In the absence of any pigment the maximum amount of oxygen which 100 cm^3 of blood can physically dissolve is only about 0.3–0.5 cm^3. The blood of the annelid worm *Arenicola* has a capacity of up to 10 cm^3 of oxygen per 100 cm^3 of blood. Thus the hemoglobin accounts for at least 95% of the oxygen carried by the blood of this animal. If a blood pigment is to be efficient as a transport agent, it must not only bind oxygen but be able to release it at sufficiently high pressures to meet the needs of the animal. The higher the oxygen pressure at which the pigment unloads its oxygen, the better will be the supply of oxygen in the tissue capillaries. The more active a particular species of animal, the greater, in general, is the oxygen pressure at which its blood unloads its oxygen. An increased temperature lowers the affinity with which all blood pigments bind oxygen. Thus an increased temperature not only accelerates metabolism generally but also tends to make more oxygen available by raising the pressure at which the hemoglobin delivers oxygen to the tissues.

Many hemoglobins aid in the transport of carbon dioxide as well as oxygen. Vertebrate blood hemoglobins have a lower oxygen affinity in the presence of carbon dioxide than in its absence. This phenomenon, found in many invertebrates, is called the Bohr effect after its discoverer, the physiologist, Christian Bohr, the father of the atomic physicist Niels Bohr. The effect facilitates the unloading of oxygen in the tissue capillaries, and the discharge of carbon dioxide at the gills or other gas-exchanging surfaces. A pronounced Bohr effect is absent from most high-affinity hemoglobins such as myoglobin, which appear to function by facilitating diffusion.

The distribution of hemoglobins among invertebrate animals does not fit any phylogenetic scheme. It appears as if hemoglobins have arisen many times in the course of evolution. In some species the hemoglobins appear to be essential for the animal's survival; in others the advantages appear marginal at best, and perhaps their occurrence is the fortuitous result of the evolution of an enzyme whose primary function is quite different from the more familiar hemoglobins. This supposition is strengthened

by the discovery that hemoglobin-like proteins have now been found in virtually all cells. Furthermore, hemoglobin-like molecules occur that have spectra resembling those of certain cytochromes (b_5). Strong evidence for this possibility comes from the discovery that cytochrome b_5 (calf liver) is homologous in its primary structure (sequence of amino acids) to that of vertebrate hemoglobin.

Multicellular organisms. Hemoglobins are distributed in every major phylum of invertebrates except those of the sponges and cnidarians. The pigments occur in the parenchyma of the flatworm *Phaenocora* and in the blood cells of a few nemertean worms. Hemoglobins occur in several parasitic nematodes, including *Ascaris, Nippostrongylus*, and *Strongylus*. Each pigment possesses an extremely high affinity for oxygen and dissociates its oxygen very slowly. Thus, it takes about 150 s to deoxygenate the hemoglobin of the perienteric fluid of *Ascaris*; under the same conditions the dissociation time of sheep hemoglobin is 0.008 s. *Strongylus* hemoglobin binds oxygen so tightly that the worms die of anoxia when placed in an anaerobic environment before their hemoglobin loses its oxygen. This fact suggests that the function of the hemoglobin might be catalytic and not to transport oxygen. In contrast, the hemoglobin in the body wall of *Ascaris* may serve a transport or diffusion function, because all body motion stops when the deoxygenation of the pigment starts. *See* NEMATA; PLATYHELMINTHES.

Among mollusks, hemoglobin occurs chiefly in the gastropods, but it has been found in a few bivalved mollusks. It occurs in the radular muscle of the whelk *Busycon*, whose blood pigment is hemocyanin. This muscle controls the radula, an organ in the buccal cavity that is used for scraping and gathering of food. The hemoglobin probably facilitates oxygen supply to this muscle. *See* MOLLUSCA.

Hemoglobin is found in the ctenidia (gills) of the bivalve mollusk *Phacoides*, which lives in the mud of mangrove swamps. The pigment may facilitate transport of oxygen across the gill to the associated blood vessels. It has a very high oxygen affinity; an oxygen pressure of only 0.007 in. Hg (0.19 mmHg or 25 pascals) is sufficient to half-saturate the pigment with oxygen. In contrast, the hemoglobin of the blood of another bivalve, *Anadara*, requires an oxygen pressure of 10 mmHg to achieve the same degree of oxygen binding. This 50-fold difference in oxygen affinity may be related to the fact that *Phacoides* dwells deep in the hypoxic mud, whereas *Anadara* lives on the surface. Hemoglobin occurs in the blood of many annelid worms. In those worms with a closed circulation the hemoglobin is usually found dissolved in the plasma; in those with less well-developed circulatory systems the hemoglobin is usually in cells confined to the coelomic fluid. Hemoglobin apparently occurs both in the plasma and in the cells of the coelomic fluid of two annelids, *Terebella* and *Travisia*.

Hemoglobins occur in the blood plasma of several crustaceans, including certain species of the water flea, *Daphnia*. Its presence in *Daphnia* depends on the oxygen content of the water: the lower the oxygen content, the greater the quantity of hemoglobin synthesized. The mechanism of this relationship is not known but may depend on the fact that one of the steps in the synthesis of heme from δ-aminolevulinic acid is strongly dependent upon oxygen, and that the synthesis of hemoglobin protein is apparently controlled in part by the quantity of heme available. The sea cucumber *Thyone*, an echinoderm, contains hemoglobin in distinct blood cells. *See* CRUSTACEA.

Hemoglobin has been found in only a few adult insects, all of which are aquatic, and in two larval insects, each of which develops in an low-oxygen aquatic environment. The tracheal tubes of the back swimmers *Buenoa* and *Anisops* are surrounded by clusters of cells that contain hemoglobin. It occurs in the tracheal cells of the larva of the botfly *Gastrophilus*, which is found in the stomach of horses. It also occurs in the water boatman *Macrocorixa* and in the body fluid of the larva of the midge *Chironomus*. No hemoglobin has been discovered in any protochordate. *See* INSECTA.

Curiously, hemoglobins have been found in the nerve tissues of a wide variety of invertebrate animals, including nematode and annelid worms, mollusks, and crustaceans. Indeed, the nerve ganglia of the annelid *Aphrodite* is bright crimson. When the hemoglobins in the tissues of certain worms and mollusks are deprived of oxygen, they exhibit a peculiar optical absorption spectrum quite different from that of deoxyhemoglobin. These spectra suggest the possible presence of ferryl hemoglobin, in which the oxidation state of the iron is 4+. The hemoglobins may serve a catalytic function, as has been suggested for similar pigments in microorganisms. A hemoglobin-like pigment occurs in the ganglia of the mollusk *Spisula*, which when deoxygenated exhibits an absorption spectrum characteristic of cytochrome b_5; the absorption spectrum of the oxygenated pigment resembles that of oxyhemoglobin. *See* HEMOGLOBIN.

Microorganisms. Hemoglobins and hemoglobin-like pigments are also found in microorganisms, being widely distributed in Protozoa, Fungi, and bacteria. The root nodules that fix molecular nitrogen in leguminous plants are pink as a result of the presence of hemoglobin. Both the fixation process and the occurrence of hemoglobin require the presence of the symbiotic bacterium *Rhizobium*. The hemoglobin is present only during periods of active fixation, when it appears to sequester oxygen for delivery to mitochondria and to protect the nitrogen fixation enzymes from the toxic effects of oxygen. Yeast and many bacteria contain flavohemoglobins, which have both heme- and flavin-containing parts. These proteins are enzymes that catalyze the oxygen-dependent conversion of nitric oxide to nitrate. *See* NITROGEN CYCLE.

Hemocyanins. Hemocyanins are found only in mollusks and in arthropods other than insects. However, they are not the only copper proteins that combine reversibly with oxygen. Ascorbic acid oxidase, an enzyme found in plants, has about the same copper content. In the presence of oxygen this enzyme is blue,

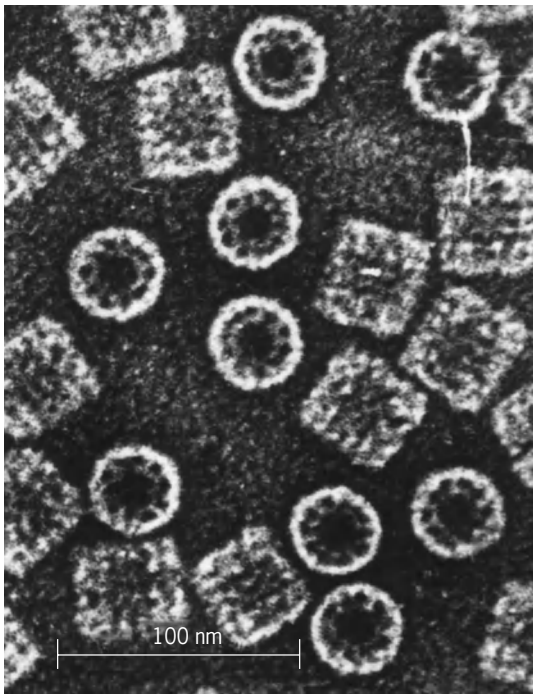

Electron micrograph of hemocyanin molecules from the snail *Helix pomatia*, showing "pillbox" subunit structure. Circular forms show subunits in a tenfold symmetrical array. There are six parallel rows of subunits in the rectangular forms.

as is hemocyanin, and when oxygen is removed, the pigment becomes colorless, as does hemocyanin. Hemocyanin itself has no oxidase activity. Hemocyanins always occur in the plasma, never in cells. The molecular weights are very high, ranging up to 7×10^6. They are sufficiently large to be seen readily under the electron microscope, where they appear as pillbox-like aggregates of many subunits (see **illustration**).

This pigment is found in many of the higher crustaceans such as crabs and lobsters, in the horseshoe crab *Limulus*, and in scorpions. Among the mollusks it occurs both in the cephalopods (squid and octopus) and in certain gastropods. The squid is an active carnivore and requires well-aerated seawater for survival. The blood of both the squid and the octopus has about the same oxygen capacity; yet the two hemocyanins have very different properties, the hemocyanin of the squid being half-saturated at an oxygen pressure of only 3 mmHg (400 Pa). This large difference reflects the fact that the squid is a more active animal than the octopus. About 92% of the oxygen of the arterial blood of the squid is normally removed during circulation through the tissues; in humans no more than about 30% is so removed. Thus the squid is poorly adapted to survive even very short periods in an environment with little oxygen. *Limulus*, on the other hand, survives for weeks with all of its hemocyanin removed.

Hemocyanins are of two distinct classes. Those from crustaceans have a minimum functional subunit of 68,000–70,000 (the unit combining with one molecule of oxygen and containing two atoms of copper). In contrast, the analytical data for molluscan hemocyanins indicate a minimum unit of about 50,000, although no functional subunit of this size has yet been prepared. Divalent cations strongly influence both the aggregation of subunits and the oxygen equilibrium; their removal from crayfish hemocyanin results in a reversible decrease in oxygen affinity. Lobster (*Homarus*) hemocyanin has a molecular weight of about 825,000 and probably consists of 12 subunits.

Chlorocruorin. Chlorocruorin has properties very similar to those of hemoglobin. It is restricted to certain sessile marine annelids. One species of the genus *Spirorbis* contains chlorocruorin in its blood, while another has hemoglobin. A third has no pigment at all. Another closely related worm, *Serpula*, has both chlorocruorin and hemoglobin dissolved in the blood. This is the only animal known to have two blood pigments. No functional reason for these differences has been suggested; each of these worms lives in a similar environment. The oxygen affinity of chlorocruorins is much lower than that of most hemoglobins.

Hemerythrin. Hemerythrin is similarly restricted and is found in a few annelid and sipunculid worms and in the brachiopod *Lingula*. It is found only in cells. The hemerythrin of the sipunculid worm *Golfingia gouldii* consists of eight identical subunits that form a molecule of 107,000. Each subunit is composed of 112 amino acids and contains two atoms of iron. The oxygen is believed to be bound between each pair of iron atoms. The subunits are held together by noncovalent bonds. *See* RESPIRATION.

Austen F. Riggs

Bibliography. P. W. Hochachka and G. N. Somero, *Biochemical Adaptation*, 2002; C. L. Prosser, *Adaptional Biology: Molecules to Organisms*, 1986; E. E. Ruppert and R. D. Barnes, *Invertebrate Zoology*, 1993; R. E. Weber and S. N. Vinogradov, *Physiolog. Rev.*, 81:569–628, 2001.

Respiratory syncytial virus

A virus belonging to the Paramyxoviridae, genus *Pneumovirus*. This virus, although unrelated to any other known respiratory disease agent and differing from the parainfluenza viruses in a number of important characteristics, has been associated with a large proportion of respiratory illnesses in very young children, particularly bronchiolitis and pneumonia. It appears to be one of the major causes of these serious illnesses of infants. It is the only respiratory virus that occurs with its greatest frequency in infants in their first 6 months of life. In older infants and children, a milder illness is produced. Respiratory syncytial virus infection in adult volunteers gave rise to a coldlike illness; adult infection took place readily, even though the adults had moderate to high levels of antibody.

The respiratory syncytial virus is 90–140 nanometers in size. It shares with measles, mumps, and parainfluenza 2 the property of producing a

characteristic syncytial effect (pseudogiant cell) in cell cultures.

The clinical disease in young infants may be the result of an antigen-antibody reaction that occurs when the infecting virus meets antibody transmitted from the mother. For this reason respiratory syncytial vaccines that stimulate production of antibodies in the serum but not in the nasal secretions may do more harm than good. Efforts to develop an attenuated vaccine that infects subclinically and produces nasal antibody are encouraging. *See* ANIMAL VIRUS; VIRUS CLASSIFICATION. Joseph L. Melnick; M. E. Reichmann

Bibliography. R. M. Chanock, M. A. Mufson, and K. M. Johnson (eds.), Comparative biology and ecology of human virus and mycoplasma respiratory pathogens, *Progr. Med. Virol.*, 7:208–272, 1965; A. S. Evans (ed.), *Viral Infections of Humans: Epidemiology and Control*, 4th ed., 1997.

Respiratory system

The system of organs involved in the acquisition of oxygen and the elimination of carbon dioxide by an organism. The lungs and gills are the two most important structures of vertebrates involved in the phase known as external respiration, or gaseous exchanges, between the blood and environment. Internal respiration refers to the gaseous exchanges which occur between the blood and cells. Certain other structures in some species of vertebrates serve as respiratory organs; among these are the integument or skin of fishes and amphibians. The moist, highly vascular skin of anuran amphibians is important in respiration. Certain species of fishes have a vascular rectum which is utilized as a respiratory structure, water being taken in and ejected regularly by the animal. Saclike cloacal structures occur in some aquatic species of turtles. These are vascular and are intermittently filled with, and emptied of, water. It is thought that they may function in respiration. During embryonic life the yolk sac and allantois are important respiratory organs in certain vertebrates.

Structurally, respiratory organs usually present a vascular surface that is sufficiently extensive to provide an adequate area of absorption for gaseous exchange. This surface is moist and thin enough to allow for the passage of gases. This article treats the embryology and physiology of the gills and embryology, anatomy, and histology of the vertebrate lung and vertebrate respiratory physiology. *See* ALLANTOIS; HEMOGLOBIN; LUNG; YOLK SAC.

Gill Development and Anatomy

In the most primitive chordates, the Acrania, the gills are really part of the digestive, rather than the respiratory, system. However, in all these animals, except one genus, *Rhabdopleura*, there is a pharyngeal region in which there are one or more (usually many) paired visceral clefts with delicate intervening strands of tissue which constitute the primary bars or arches; these are greatly augmented in some cases by the down growth of tongues of tissue between them, which are termed secondary bars. There are also transverse connecting strands termed synapticulae in some species, such as *Branchiostoma*. The bars and synapticulae usually contain supporting gelatinous or chitinoid rods; in *Branchiostoma* the primary bars contain prolongations of the dorsal coelom. Most important, blood circulates through the bars and synapticulae during respiration and is aerated by water drawn in through the oral opening by the ciliated endoderm and passed out through the interstices of the basketlike structure. This arrangement acts primarily as a strainer to retain food particles (**Fig. 1**).

The gills of chordates are located in the pharyngeal region. They are always related to the visceral arches and clefts. The latter structures are so defined because, in vertebrates that respire entirely with lungs, the arches and clefts, or incipient clefts, lack gills, which exist only in the embryo and disappear or are highly modified in the adult. They always arise by the outpushing of endodermal pouches, which are usually dorsoventrally elongated, through the mesoderm, which contact corresponding inpushings of ectoderm. Where these meet and break through, clefts are formed, and the intervening concentrations of mesoderm form the arches, which are covered internally by endoderm and externally by ectoderm. The number of clefts, the eventual content of the arches, and the extent to which the latter develop gills of one type or another are indicated in the following discussion.

Visceral arches in craniates. In this group the typical number of paired visceral arches and pouches or clefts is six, including the mandibular and hyoid arches with the intervening spiracle. However, the number may be greater in some elasmobranchs and cyclostomes, and less in some animals where these structures occur only in the embryos. In the latter cases the number of pouches which actually become clefts is usually very limited; for example, there are three pairs in the chick, one pair in the cow, and

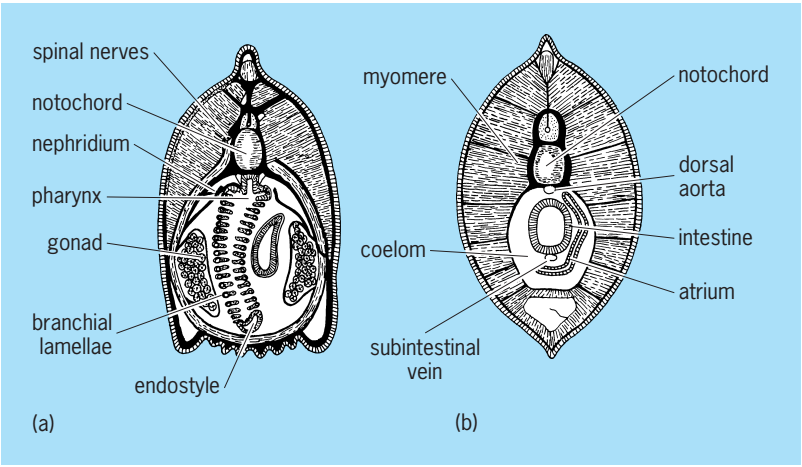

Fig. 1. ***Branchiostoma lanceolatus*. (*a*) Transverse section of the pharyngeal region. (*b*) Transverse section of the intestinal region. (*After T. J. Parker and W. A. Haswell, A Textbook of Zoology, 6th ed., Macmillan, 1956*)**

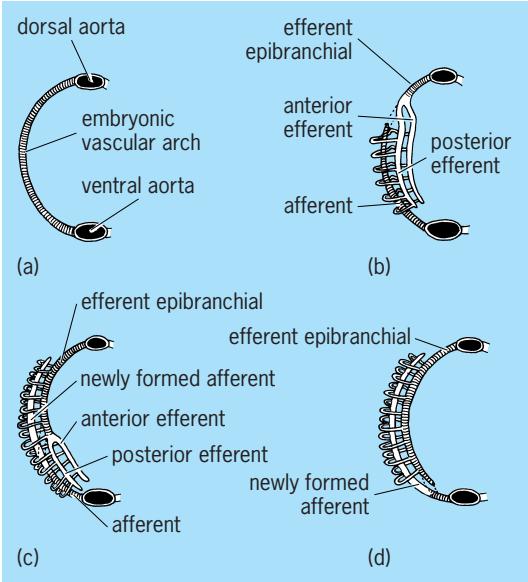

Fig. 2. Diagrams of the development of adult branchial vessels in various fishes. (*a*) Original continuous embryonic arch cross-lined. (*b*) Selachian. (*c*) Intermediate form, such as *Acipenser*. (*d*) Teleost: newly developed vessels, white. In *b*, *c*, and *d* the original arch is interrupted. (*After E. S. Goodrich, Studies on the Structure and Development of Vertebrates, Macmillan, 1930*)

normally none in the pig and human, in which the number of arches is reduced to five, with the last vestigial. The blood vessels found in the arches in animals other than fish are of considerable phylogenetic interest. Thus the vessels in the third visceral arches become the roots of the internal carotids; those of the fourth form a pair of aortae in amphibians and reptiles, or one of this pair an aorta in birds and mammals; and those in the fifth pair give rise to the pulmonary arteries. The pulmonary arteries are really the sixth pair of aortic vessels because a vestigial fifth pair occurs, together with the sixth, in the fifth visceral arches.

In fishes the aorta in each of the arches which bear gills, commonly the posterior four pairs, becomes altered and augmented to form the afferent and efferent vessels of the gills of that arch. The precise method of alteration and final condition vary somewhat in different groups of fishes.

In the elasmobranchs the ventral two-thirds of the original aortic arch persists as the afferent vessel. This becomes disconnected from the dorsal third, from which two vessels grow ventrally. These and the dorsal third thus become the efferent aortic arteries and are connected with the afferent vessel through gill capillaries in a manner to be described below (**Fig. 2***b*). The more anterior of the two growing vessels drains the posterior half of the gill pouch preceding it, while the posterior half drains the anterior half of the gill pouch following. In each arch the two efferent vessels have transverse connections midway between them, and all the efferent vessels on a side are eventually united by dorsal and ventral vessels. The teleost condition represents essentially the reverse of what occurs in the elasmobranchs with respect to the origin of the afferent and efferent vessels. Thus in the teleost it is a new upgrowth from the base of the original arch which forms the afferent vessel, whereas most of the original aortic arch forms the efferent vessel. Teleosts also have only one efferent vessel in each visceral arch (Fig. 2*d*). An intermediate form (Fig. 2*c*) with respect to the origins and character of these vessels occurs in *Acipenser*, *Amia*, and *Lepisosteus*, although in detail these forms are not precisely as shown in the diagram.

Branchial gills. The term branchial gills is used because there are other gills and structures functioning as gills, which are not in the pharyngeal region. There are generally considered to be two kinds of branchial gills, external and internal. Because the actinopterygian internal gill is apparently derived from the elasmobranch type, the elasmobranch gill is discussed first.

Elasmobranch gill. In these fishes, there develops within each branchial arch a delicate cartilaginous bar slightly proximal to the vessels, the median afferent and two lateral efferent vessels, which have already been described. The arches also become supplied by branches of the ninth and tenth cranial nerves (the hyoid by the seventh) and acquire muscle fibers derived from the walls of coelomic extensions by which the arches are temporarily invaded. From the outer side of each arch there grows along its dorsoventral extent a thickish sheet of tissue which develops along its border two indentations or clefts, tending to divide the single sheet into anterior, median, and posterior layers. The indentations never become very deep, however, so that these layers are essentially one. The middle layer or part of this sheet becomes the septum, consisting of connective tissue and a little muscle and covered, where it is free, with epithelium. The anterior and posterior layers on each side of the septum become half gills or hemibranchs, and the whole structure a holobranch (**Fig. 3**).

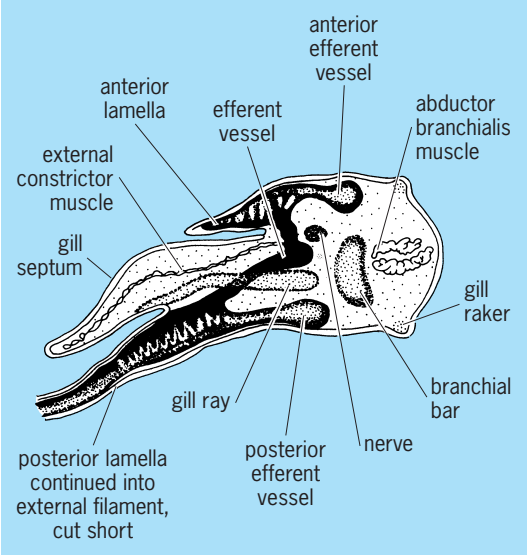

Fig. 3. Section across gill-bar of *Scyllium canicula* late embryo 1.3 in. (32 mm) long, showing blood supply to lamellae. (*After E. S. Goodrich, Studies on the Structure and Development of Vertebrates, Macmillan, 1930*)

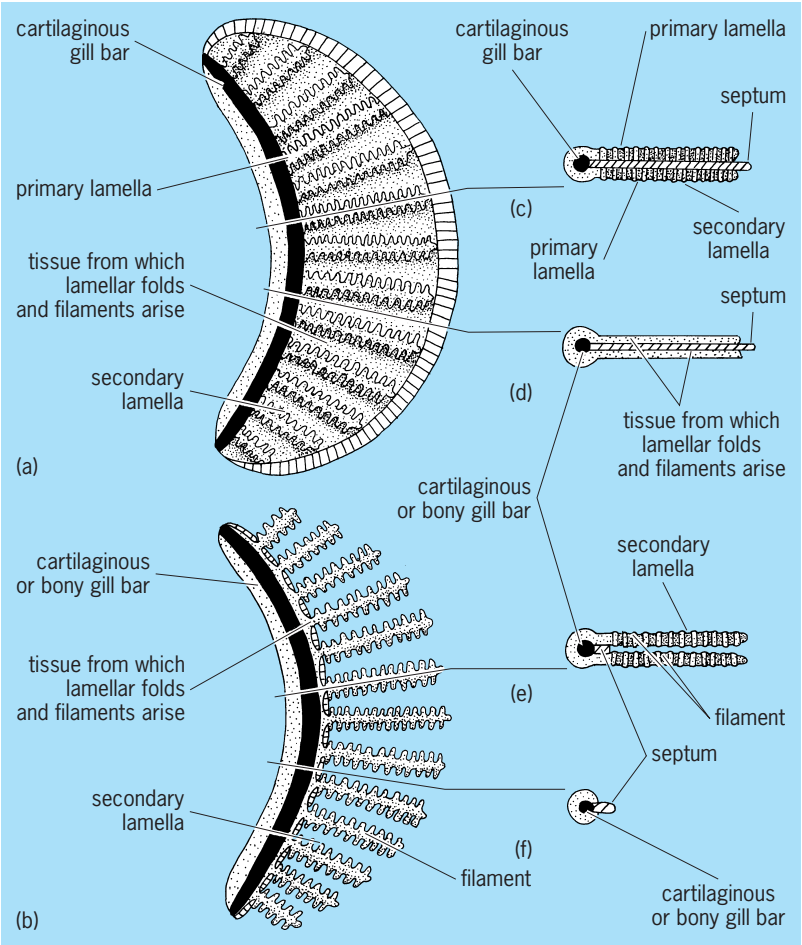

Fig. 4. Diagrams of elasmobranch and teleost gill structure. (*a*) Gill arch and gill (hemibranch) from the right side of the head of an elasmobranch; posterior view. (*b*) Same view of teleost arch and gill. (*c*) Horizontal section along the length of a pair of primary lamellae and the septum of the holobranch of an elasmobranch. (*d*) Horizontal section through three tissue layers between primary lamellae of an elasmobranch. (*e*) Horizontal section along the length of a pair of filaments of a teleost holobranch. (*f*) Horizontal section between the filaments of a teleost holobranch, hence through the arch and remains of the septum only.

Along the surface of each anterior and posterior layer there arises a series of transverse folds of tissue extending outward from the arch (**Fig. 4**). These are the primary lamellae covered by an epithelium, flat for the most part but with some columnar or cuboidal secretory cells. Between the epithelial cells there is connective tissue which contains, within each lamella, a loop from the afferent and adjacent efferent vessel of the arch. Another series of folds transverse to these lamellae develops next on the dorsal and ventral surface of each primary lamella. They are the secondary lamellae, whose internal surfaces are connected by numerous columnar cells which send out delicate protoplasmic strands beneath the epithelium of the folds. These cells are termed pilaster cells, and a network of capillaries runs among them connecting the afferent and efferent loops in the respective lamella (**Fig. 5**). A cartilaginous ray extends outward along the base of each primary lamella at the boundary between the posterior layer or hemibranch and the septum and helps to support the entire gill (Figs. 3 and 4*c*). In unhatched elasmobranchs the primary lamellae of the posterior hemibranch grow outward as filaments beyond the external openings of the gill clefts (Fig. 3). These temporary filaments are therefore essentially external gills, although in both origin and structure they are unlike the "true" external gills, described below. Because they occur only in the unhatched fishes, where they float in the albuminous fluid of the egg, they are probably as much for absorption of food as respiration. The septum of each holobranch grows out beyond the gill in these fishes and turns posteriorly beneath the outer skin to the posterior edge of the following cleft, thus forming for the latter an individual cover analogous to the operculum. Each branchial arch develops an anterior and posterior row of papillae, sometimes covered with enamel, on its internal border. These are the gill rakers, which act as strainers to prevent the escape of food (Fig. 3).

Holocephalan gills. In the subclass Holocephali the gills are similar to those in the elasmobranchs, although arranged slightly differently with respect to the arches, and are covered by an operculum.

Cyclostome gill. Although this class is considered the most primitive of all the craniates, its consideration has been delayed because the gills present a highly specialized condition most easily described as a variant of the elasmobranch type. The early development is similar to that in elasmobranchs. The septa are soon drawn out, however, so that the clefts become elongated tubes that sometimes open separately, as in *Bdellostoma*, or unite to open through a single orifice, as in *Myxine*. The gills proper have primary and secondary lamellae, as in the elasmobranchs, and occur as hemibranchs on the sides of the septa near the pharynx where the tubes become enlarged into spherical pouches. The pouches open directly either to the pharynx or through very short ducts. The pharyngeal region becomes divided during development in the lampreys into a dorsal

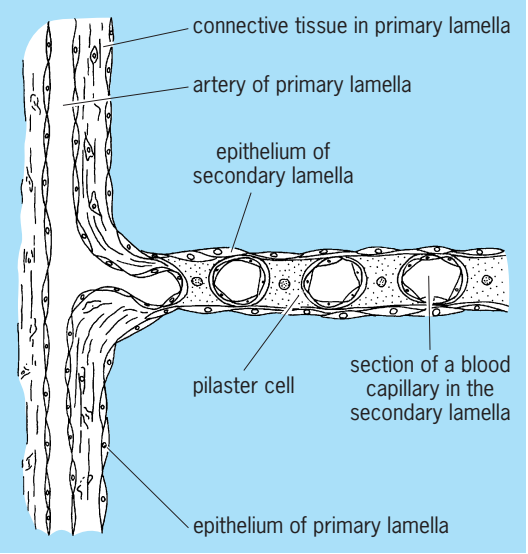

Fig. 5. Diagram of a section through a part of a secondary lamella of the gill of an elasmobranch fish, showing connection of the secondary lamella with a primary lamella.

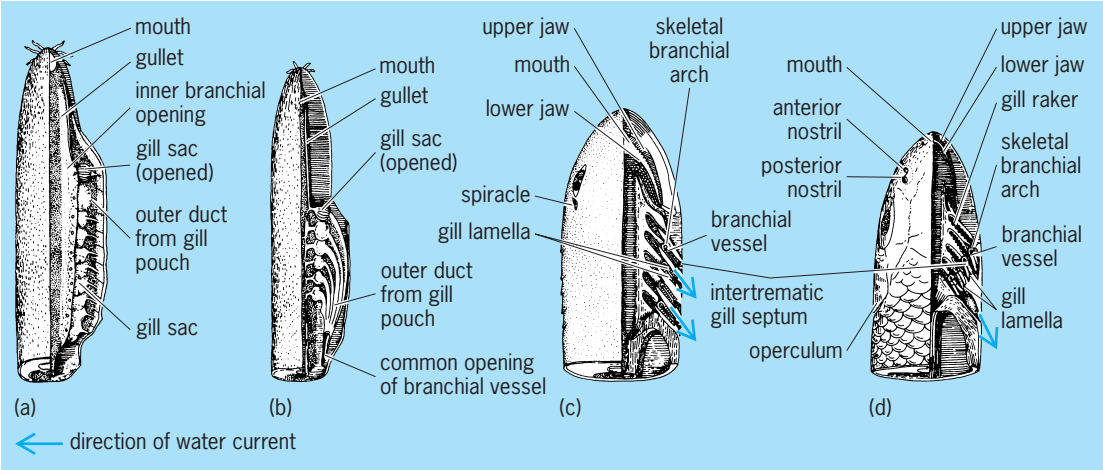

Fig. 6. Diagrams of head and gill region in (*a*) *Bdellostoma*, (*b*) *Myxine*, (*c*) a selachian, and (*d*) a teleost. Dorsal view partly dissected to show arrangement of gills. (*After E. S. Goodrich, Studies on the Structure and Development of Vertebrates, Macmillan, 1930*)

esophageal portion and a ventral respiratory tube with a blind posterior ending. It is this ventral portion which gives rise to the gill slits. The number of slits, eventually pouches and ducts, varies greatly in this class, ranging from 6 to 14 in different species of *Bdellostoma* (**Fig. 6**).

Actinopterygian gill. With a few exceptions to be noted below, it is possible to describe the gill situation in most bony fishes by indicating the ways in which it differs from that of the elasmobranchs. Although the holobranchs start to develop from the four branchial arches, as in the latter group, actinopterygians generally lack an open spiracle and the hemibranch related to it; instead, there is often a so-called pseudobranch which is probably glandular. Usually no temporary filamentous external gills occur. The septum grows out only a very short distance, and the anterior and posterior hemibranchs continue to grow out, not as sheets, but as numerous free filaments, corresponding to the primary lamellar folds (Fig. 4*b*, *e*, *f*). The secondary lamellae develop from these in the same way and with the same histological characters, as in elasmobranchs. Although there is only a single efferent artery in the arches of most bony fishes (Fig. 2*d*), the circulation in the filaments and secondary lamellae is similar to that in the primary and secondary lamellae of the former group. Another difference is that instead of a single ray extending between each primary lamella of the posterior hemibranch and the septum there is a ray in each pair of filaments of a holobranch (Fig. 4*b*, *e*). Lastly there is a single operculum on each side, attached anteriorly to the hyoid arch and covering all the gills. Each operculum is a sheet of tissue in which are embedded three flat bones, the operculars. Gill rakers are present in one or more (often two) rows, sometimes supported by ossifications (Fig. 6*d*).

There are a few exceptions to the situation just described. The gills of *Acipenser* are about halfway between those of elasmobranchs and typical actinopterygians, and the gill filaments of the lophobranchs consist of tufted processes.

Dipnoan gill. The Dipnoi may be considered as a sort of connecting link between the fishes and the amphibians. In Dipnoi there are both internal and apparently true external gills. These latter gills exist in the embryos of all the Dipnoi, and vestiges of them persist in the adult where they are attached to the last three pairs of arches. The internal gills in this group are reduced in correlation with the accessory respiration furnished by the lung or lungs. The septa are somewhat diminished, causing the primary lamellae to become partly filamentous, but without rays. There may be hemibranchs on the hyoids, on the last pair of branchial arches, or on both; there are thus two to four pairs of holobranchs.

Other respiratory devices of fishes. In addition to pharyngeal gills of the types indicated, there are other aquatic respiratory mechanisms, which sometimes occur in quite different locations. One such organ, still associated with the pharynx, is found in *Anabas*, the climbing perch; in this fish some of the pharyngeal bones are developed into folded plates covered with vascular epithelium, through which respiration can occur. This arrangement is covered by the operculum and can be kept moist for extended periods out of water. In one of the phyostomes, *Amphinous*, another pharyngeal derivative acting as a gill consists of vascular sacs opening out through the spiracles. Another peculiar structure used for respiration is the tail fin of an Indian Ocean fish which is immersed while the fish basks out of water. One of the Dipnoi, *Lepidosiren*, despite its lungs, apparently develops respiratory filaments on the paired fins. One of the siluroids employs rectal respiration by sucking in and expelling water from the anus.

Amphibian gills. Almost all amphibians have gills at some time during development and may retain them as adults. At least part of the time these are of the true external type as compared with the filamentous variety previously mentioned. Although some authorities consider that the differences between the internal gills and what are termed true external gills are not very significant, it seems that in this class of

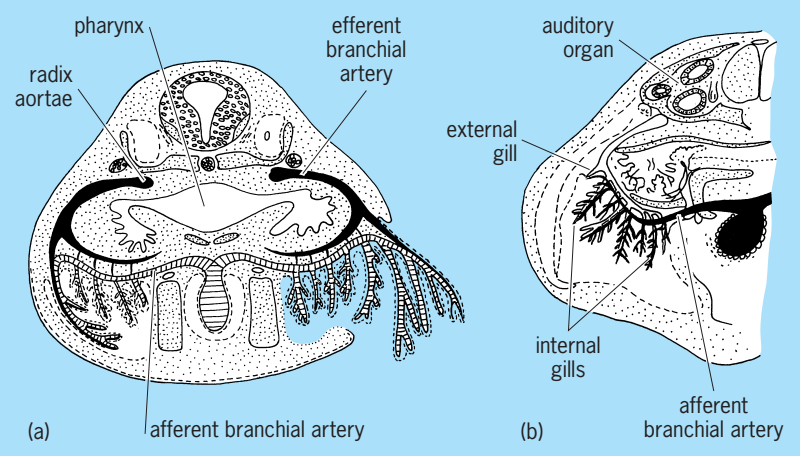

Fig. 7. External gills. (*a*) Relations with internal gills; transverse section. (*b*) External gill development. (*After E. S. Goodrich, Studies on the Structure and Development of Vertebrates, Macmillan, 1930*)

animals these differences are sufficient to demand some notice. Generally speaking, true external gills of amphibians may be distinguished from the gills, internal or external, already discussed in that they arise from the outer borders of the gill arches rather than from the sides and differ in certain details of structure. They are sometimes covered by an operculum.

Each gill of the type under consideration is composed of a rather heavy main stem, or rachis, which arises first. This is much thicker than the filaments of real internal gills and has been compared to the modified septum of such gills, which is otherwise entirely lacking in the variety being described. In those amphibians in which external gills are entirely larval, this rachis usually gives rise to rather blunt, short, fingerlike processes. The epithelium of such gills is ciliated, and within the rachis, at least, are muscle fibers so that the gills can be moved. The rachis also contains an arterial loop with extensions into each branch (**Fig. 7**). Gills of this type, which are typically found in anurans, such as the frog, spring from the upper parts of the first three pairs of posthyoid branchial arches; those from the first pair are the most prominent, overlapping and concealing the other two pairs.

These external gills in the anurans and many urodeles soon begin to be absorbed and are covered by back-growths of tissue from the hyoid arches which are continuous across the ventral side of the throat. This tissue contains no cartilaginous or bony plates and constitutes an operculum which fuses with the body wall everywhere posterior to the gill slits, except for one or two small openings termed spiracles through which water from the clefts leaves the branchial chamber. These spiracles are not homologous with those in fish previously mentioned.

Meanwhile, as the operculum is completed, the original external gills are absorbed, and in the Anura double rows of outgrowths on the borders of the first three pairs of branchial arches and a single row on the anterior border of the fourth are developed in their stead. These processes are quite similar to the processes which arose from the main rachis of the external gills, except they are shorter. They are sometimes termed internal gills, chiefly because they are covered by an operculum, but they are still not highly filamentous and lack any part of a septum. Hence, they may be regarded as reduced external gills which are covered (Fig. 7). There are, nevertheless, parts connected with these gills which appear to be closely related to similar parts of the internal gills of fish. The inner margins of the gill arches bear double rows of papillae which correspond, at least in function (that is, acting as strainers), to the gill rakers of the former group.

In certain of the urodeles, such as *Necturus*, the original gills are never covered or absorbed, but persist throughout life as actual external structures. In these cases, however, they differ somewhat from the larval external gills just described. The rachis is larger and heavier, and the secondary branches give rise to numerous fine, short, tertiary outgrowths, giving the whole structure a bushy effect (**Fig. 8**). Although apparently unnecessary, some of the clefts between these sets of gills remain open, for example, one pair in *Pseudobranchus*, two pairs in *Necturus*, and three in *Siren*. This retention of the gills and clefts in such forms has been regarded as an example of neoteny, that is, the persistence of larval characters in sexually mature animals. *See* NEOTENY.

Other amphibian respiratory devices. In the amphibians as in fish there are various peculiar structures which act as gills. Thus in the marsupial frog the two anterior pairs of gills are transformed into vascular wrappings which surround the body, whereas the balancers of some salamanders have been said to have a partly respiratory effect, although this is doubtful. The skin in amphibians certainly has a respiratory function, and that of *Cryptobranchus* is thrown into vascularized folds which are waved about. In the "hairy" frog the skin of the breeding male develops filamentous processes which aid in respiration.

Gill ectoderm and endoderm. It has been claimed that the epithelium of internal gills is endodermal, but because the boundary between endoderm and ectoderm is indistinguishable by the time the gills arise, this is difficult to prove. It has also been stated that the ectoderm grows inward and covers the

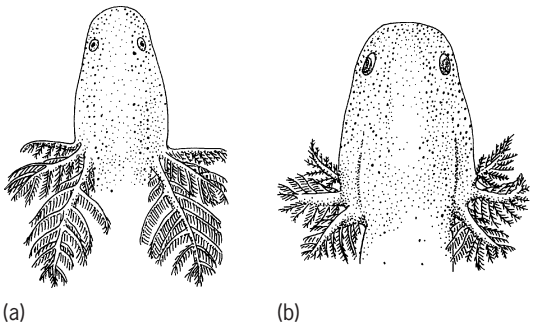

Fig. 8. External gill form in adult urodeles. (*a*) *Pseudobranchus striatus*. (*b*) *Siren lacertina*. Drawn from living specimens about 6 in. (15 cm) total length. (*After G. K. Noble, The Biology of the Amphibia, Dover, 1931*)

areas from which the gill lamellae and rakers develop, thus making the epithelium of these parts ectodermal. This would at least account for occasional rakers with enamel, a substance supposed to be derived only from ectoderm. The point of origin of gills designated as truly external marks their epithelium as clearly ectodermal; this is also probably true of the later internal gills of those amphibians which have them. Robert S. McEwen; Thomas S. Parsons

Physiology of Gills

Gills are involved in many functions, including the exchange of oxygen, carbon dioxide, water, acid, inorganic ions, ammonia, and urea between the body and the surrounding water. Much of the transfer of these compounds is passive, but some ions are moved against an electrochemical gradient requiring the expenditure of energy. As a result, the gills are a metabolically active tissue; they consume up to 10–25% of the oxygen uptake of a resting fish.

Functionally, the gills can be considered as a large but thin, epithelial surface layer separating the inside (blood) of the animal from the water. In mammals and other air-breathing vertebrates, oxygen and carbon dioxide are transferred across a lung epithelium. The lung has a very different gross structure than gills and is ventilated in a different way. This is because gills are specialized for exchange of material in water whereas lungs have evolved for gas exchange in air. Water is about 1000 times more dense and viscous than air and contains only 1/30 as much molecular oxygen; oxygen diffuses 10,000 times more slowly in water than air. To ensure adequate oxygen transfer from water into the animal, most aquatic animals maintain a continual, unidirectional flow of water over the gills, with only very small diffusion distances. The gills are essentially a fine sieve placed in the path of the water flow so that a thin layer of water passes over each region of the gill epithelium.

The exchange units of the gills of fishes are the lamellae, which are arranged in rows on filaments attached to gill arches (**Fig. 9**). Blood flows through the lamellae in the opposite direction to water flow,

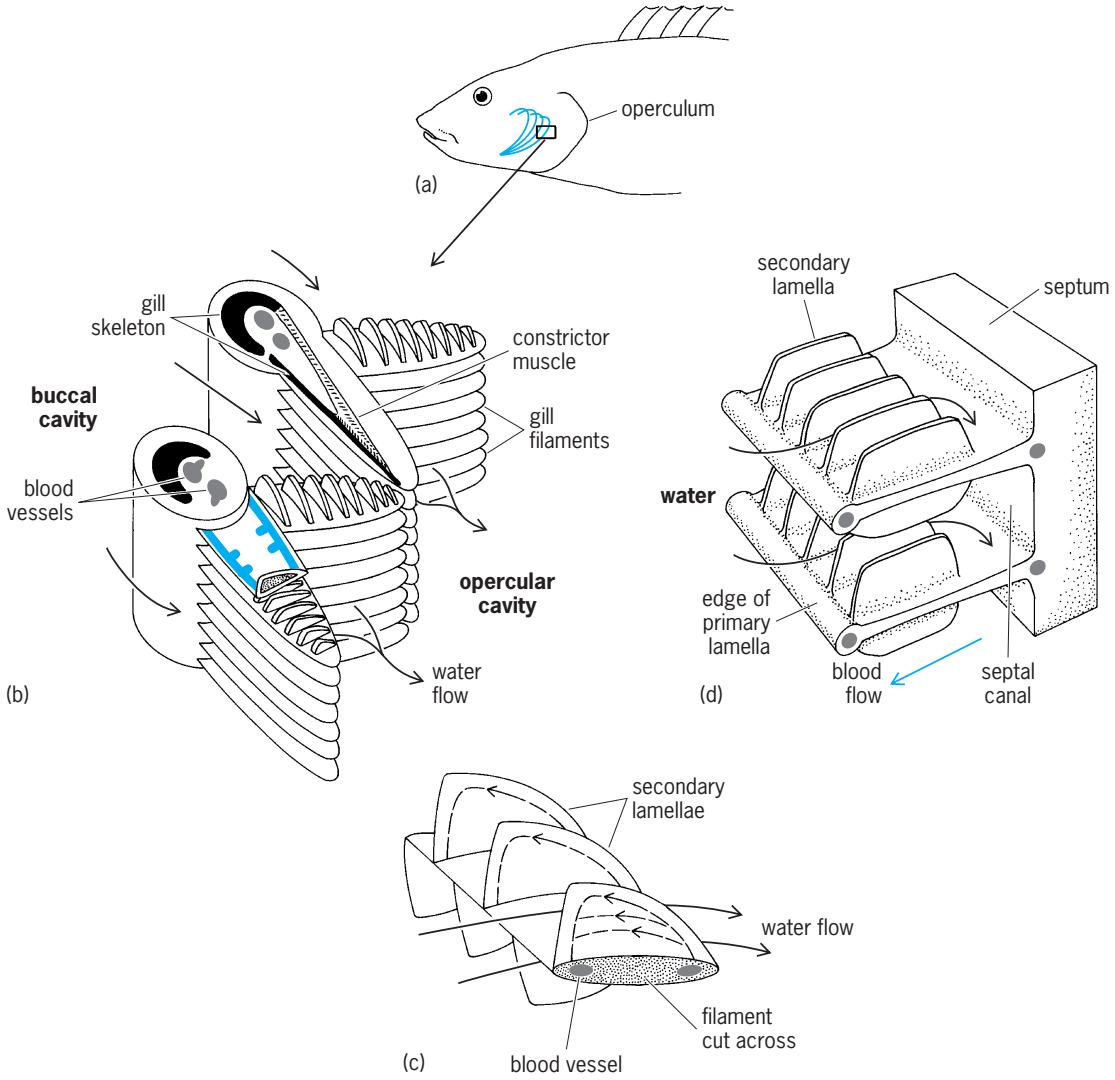

Fig. 9. Structure of fish gills. (*a–c*) Structure of a bony fish, with increasing magnification. (*d*) Elasmobranch gill at the same magnification as *c*. Black arrows show path and direction of water flow. Compare with Fig. 4.

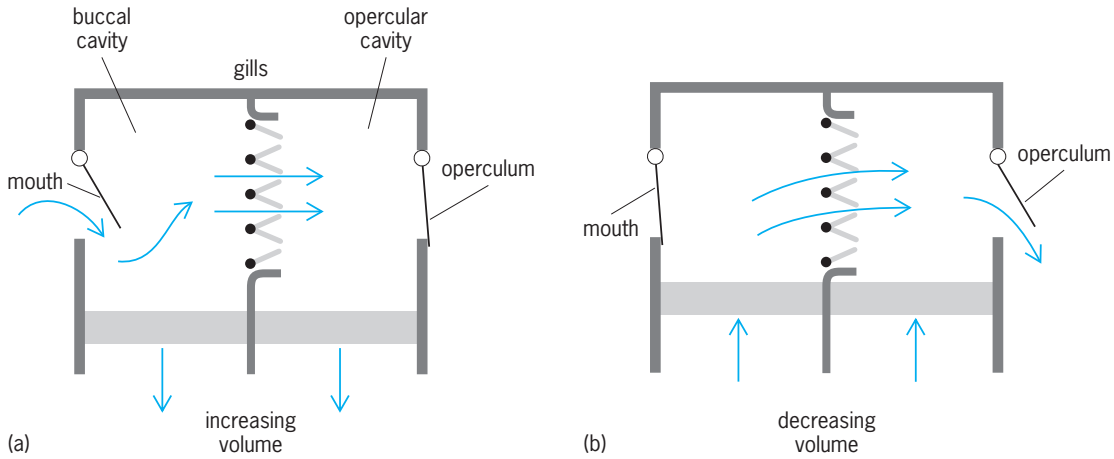

Fig. 10. Diagram of the mechanism of generating water flow over the gills of a fish (*a*) with mouth open and (*b*) with mouth closed.

creating a countercurrent exchange system between blood and water across the gill epithelium. Oxygen is delivered to the gill surface by water flow, diffuses across the gill epithelium, and is carried away in the blood bound to hemoglobin. Oxygenated blood, because of the presence of hemoglobin, contains much more oxygen per unit volume than water. The flow of water over the gills is much greater than the flow of blood through the gills, but the volume flow times oxygen content of blood leaving and water entering the gills is about the same.

Gill ventilation. Water flow over the gills of fishes is maintained by the action of muscular pumps and unidirectional flap valves situated at the mouth and the opercular cleft in bony fishes, and the mouth and gill slits of elasmobranches. The valves are skin folds that permit flow in only one direction, into the mouth and out via the gill slit (**Fig. 10**). The mouth (buccal) cavity expands, drawing water into the mouth; at the same time the postgill or opercular cavity expands with the outlet valve closed, drawing water over the gills; the opercular and buccal cavities then contract, forcing water out via the gill slit. The buccal force and opercular suction pumps operate at the same frequency but slightly out of phase so that, although water velocity over the gills varies with each breathing cycle, there is nearly always some water flow over the gills, with little or no backflow. The relative importance of the buccal versus the opercular pump differs among species.

Fast-swimming fishes, for example mackerel and tuna, swim forward with an open mouth so that water is forced over the gills by the forward motion of the fishes. This is referred to as ram ventilation: clearly there is no requirement for, nor are there any, buccal or opercular movements. If, however, the animal stops swimming, then the buccal and opercular pumps are used to generate gill water flow. Some fishes, for example tuna, swim all their lives and maintain water flow by ram ventilation. The actual water flow rate depends on the swimming speed and is adjusted at any one speed by changing the size of the mouth gap.

Some amphibians have external gills, that is, the gills are not enclosed in a buccal cavity. These gills are not ventilated but are waved around in the water to ensure adequate oxygen delivery. Some other amphibians have neither gills nor lungs but obtain oxygen from the water simply by diffusion across the general body surface.

Low oxygen levels often occur in rivers and lakes and even in the sea. Under these conditions, many fish supplement their oxygen supply by breathing air. The air-breathing organ is usually the swimbladder or a buccal or gut cavity and not the gills, which are retained for gas transfer in water. Some fish are obligate air breathers and die if denied access to air. In this case, the gills are retained for excreting carbon dioxide into water and maintaining ionic balance. The air-breathing organ is for oxygen uptake. This is associated with changes in the circulatory system (**Fig. 11**) culminating in the completely divided circulation of mammals and birds.

Blood supply. The fish heart is situated ventral to the gills. Blood, pumped through the gills by the heart, spreads into a thin sheet within the lamellae. Blood pressure within the gill lamellae is high, but the lamellae do not expand because the two sides of the structure are held parallel to each other by a number of cells that contain collagen fibers, called pillar cells. Their collagen fibers are extracellular but contained in lateral folds of the pillar cell and extend along the lateral face of the lamellae, enclosing and supporting the blood space and retaining its thin, sheetlike nature in the face of a high blood pressure. Such high pressures are required to drive blood around the body after it has left the gills.

Level of oxygen transfer. If oxygen levels in the water are low, fishes attempt to maintain oxygen supply by increasing water flow across the gills. If the oxygen requirements of the fishes increase (for example, during exercise), then both water flow and blood flow are increased. The blood is more evenly distributed throughout the gills, and the epithlium is stretched and thinned, enhancing the area and reducing the thickness of the gill epithelium. This

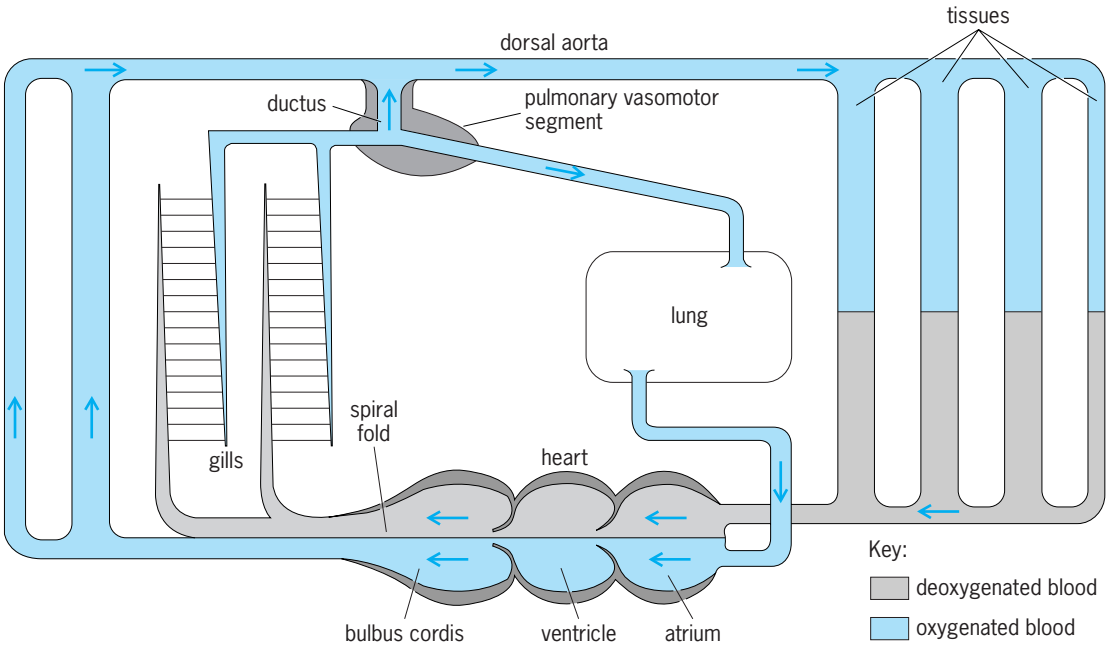

Fig. 11. Circulation of the African lungfish, *Protopterus*. Separation of oxygenated blood and deoxygenated blood is achieved by a septum which divides the atrial and ventricular chambers. Blood flows directly to the systemic circulation via the dorsal aorta. The blood is directed to the dorsal aorta or lungs depending on whether the fish is breathing in air or water. (*After D. Randall et al., eds., Eckert Animal Physiology, 4th ed., W. H. Freeman, New York, 1997*)

increases the capacity of the gills to transfer gases. The fish can increase oxygen transfer across the gills by 10–15 times and so maintain high levels of aerobic exercise. There are oxygen receptors on the gills that respond to reduced oxygen levels in the water and in the blood. These receptors when activated result in an increase in gill water flow. They may also cause a slowing of the heart rate, but cardiac output is not much affected because stroke volume increases. It seems that fishes adjust water and blood flow in the gills to meet the oxygen requirements of the animal rather than to regulate the carbon dioxide excretion of the fishes.

CO_2 delivery to gills. The tissues of animals produce about the same amount (or a little less) of CO_2 as the amount of O_2 consumed. The exact ratio of CO_2 produced to oxygen utilized (the respiratory quotient) depends on whether fats, proteins, or carbohydrates are used as the metabolic substrate. Carbon dioxide is transferred more rapidly than oxygen in tissues and water, so if the respiratory system is set to deliver adequate supplies of oxygen from water to the tissues, there will be no problem excreting carbon dioxide in the reverse direction. The only problem concerns the fact that CO_2 reacts with water to form carbonic acid which dissociates to form bicarbonate. At body pH most of the CO_2 exists as bicarbonate, but membranes are much more permeable to CO_2 than bicarbonate. Thus bicarbonate in tissues is converted to CO_2, diffuses into blood, forms bicarbonate, and is carried to the gills, where it is converted to CO_2 to diffuse across the gills into water. The conversion of CO_2 to bicarbonate and the reverse reactions occur slowly unless catalyzed by the enzyme carbonic anhydrase. Teleost fish red blood cells and fish gills have high levels of this enzyme but no activity is found in the plasma. Bicarbonate in plasma enters the red blood cell and is converted to CO_2 before diffusing across the gills. Carbonic anhydrase in the gill epithelium appears not to play an important role in CO_2 excretion but rather couples CO_2 excretion to ion transfer (Fig. 5). Elasmobranchs have carbonic anhydrase activity in the plasma as well as the red blood cells and the gills.

Ion transfer. Animals in fresh water lose NaCl across their body surfaces. This loss is counterbalanced by an uptake of NaCl in the gills. Influx of Na^+ is coupled to H^+ or NH_4^+ excretion, whereas Cl^- is coupled to HCO_3^- excretion in a separate mechanism. The HCO_3^- excreted via this pathway represents only a small portion (perhaps 5–10%) of the total CO_2 excretion. Sodium uptake via a sodium channel is maintained by the activity of a proton ATPase on the apical membrane of the gill epithelium. The excretion of protons generates a potential across the apical membrane, the inside negative. This creates an electrochemical gradient for sodium and draws sodium into the cell (**Fig. 12**). The sodium is then transferred into the blood via a Na^+/K^+ ATPase on the basolateral border of the epithelium. Carbonic anhydrase catalyzes CO_2 hydration and supplies protons for the proton pump. The associated bicarbonate production is exchanged for chloride. Seawater fishes have evolved specialized chloride cells which remove excess NaCl from within the fishes (**Fig. 13**). *See* OSMOREGULATORY MECHANISMS.

Fishes like the salmon that can tolerate a wide range of salinity increase the number of chloride cells within the gill epithelium as they migrate from rivers to the sea. An associated accessory cell develops

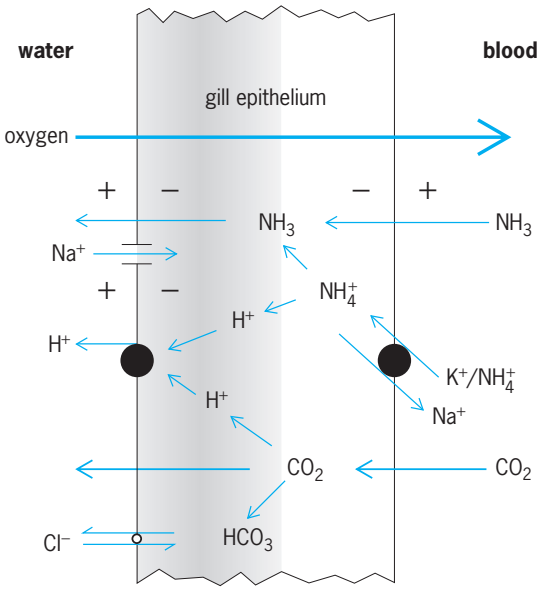

Fig. 12. Model of gas, sodium, and proton transfer across the gill epithelium of a fresh-water teleost. Shading indicates carbonic anhydrase activity. (*After H. Lin and D. Randall, Proton pumps in fish gills, in C. M. Wood and T. J. Shuttleworth, eds., Cellular and molecular approaches to fish ionic regulation, Fish Physiology, vol. 14, pp. 229–255, Academic Press, New York, 1995*)

alongside each chloride cell with loose connections forming a gap between the cells (paracellular channel). The chloride cell is much larger than the epithelial cells and extends from the outside across the gill epithelium. The epithelial cells and the chloride cells are tightly cemented together; the only open channel is between the chloride and accessory cell. The basal (blood) border of the chloride cell is infolded, and the membrane is associated with high levels of the enzyme Na^+/K^+ ATPase. This enzyme is associated with Na^+ removal from the chloride cell into the blood. This creates a Na^+ gradient from blood into

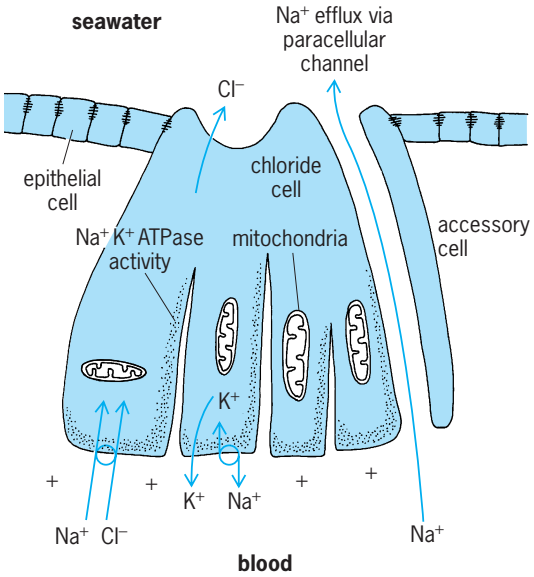

Fig. 13. Sodium chloride (NaCl) excretion in seawater fish occurs via the chloride cell found within the gill epithelium.

the chloride cell. Na^+ diffuses back into the chloride cell, involving a carrier that couples chloride movement to Na^+ movement. In this way chloride builds up in the cell and, because the cell is negatively charged, there is a large electrochemical gradient from the chloride cell into seawater and chloride is excreted from the cell. This net movement of chloride from blood to seawater results in the development of a positive potential across the gill, creating an electrochemical gradient for sodium, which diffuses from blood to seawater mainly via the paracellular channels. If a fish returns to fresh water, this paracellular channel closes, the accessory cells disappear, sodium loss is reduced, and chloride efflux also declines. Thus the gills of fishes are very much involved in maintaining internal NaCl levels.

Ammonia excretion. The main excretory end product of protein metabolism in fishes is ammonia (NH_3), but some urea is also produced. Ammonia and urea are both lost by passive diffusion across the gills. In elasmobranchs the body surfaces are not so permeable to urea, which is retained at high levels along with trimethylamine oxide to create conditions in the body fluid which are iso-osmotic with seawater.

Most of the ammonia excreted can be accounted for by passive diffusion of ammonia gas. Membranes are very permeable to NH_3 but less so to ammonium ion, which is sometimes moved by a carrier molecule through membranes. Ammonia is toxic, and if levels in the water increase as a result, for example, of the deposition of biological waste, then fishes cannot excrete ammonia and it builds up in their bodies, and they die. Under some circumstances fishes can excrete ammonium ions in exchange for sodium via a carrier-mediated process which may occur against a net ammonia gradient. Most ammonia excreted across the gills is produced in the liver, but some may be produced in the gills, which have the capacity to deaminate adenylates. The gill tissue also has the capacity to produce glutamine, but there is no evidence that ammonia is ever excreted as glutamine.

pH regulation. Fish have difficulty excreting ammonia into alkaline waters or when air-exposed. Some fish have a complete ornithine urea cycle and can convert ammonia to urea for excretion. For example, the Lake Magadi (Kenya) tilapia, living in very alkaline conditions, can produce urea via the ornithine urea cycle at about the same rate as a rat. Other fish store ammonia as glutamine or other amino acids during air exposure and then convert the amino acids back to ammonia for excretion after returning to water. The gill epithelium also appears to transport hydrogen ions. Thus any acid produced by the fish could result in the movement of H^+ across the gills into the water. The environment can be considered, therefore, as a very large H^+ sink, ameliorating pH changes within the animal. When, however, the water is acidic, proton excretion is inhibited, and this in turn inhibits sodium uptake and sodium levels in the fish and the animal dies. As a result, acid waters are a danger to many fish. Calcium levels in the

water decrease the permeability of the gill epithelium to ions, including H^+, so acid conditions in the water are better tolerated if calcium levels in the water are high. Industrial pollution in the form of acid rain quickly reduces the pH of soft-water rivers and lakes. These waters are poorly buffered and also have low Ca^{2+} levels; as a result the fishes in these waters are very susceptible to acidification. Large numbers of lakes in North America and Europe no longer contain fishes because of acidification due to industrial pollution. *See* ACID RAIN; PH REGULATION (BIOLOGY).

David J. Randall

Anatomy of the Lung

The shape and volume of the lung, because of its pliability, conforms almost completely to that of its cavity. The lungs are conical; each has an apex and a base, two surfaces, two borders, and a hilum.

Thoracic cavity. The apex extends into the superior limit of the thoracic cavity. The base is the diaphragmatic surface. The costal surface may show bulgings into the intercostal spaces. The medial surface has a part lying in the space beside the vertebral column and a part imprinted by the form of structures bulging outward beneath the mediastinal pleura (**Fig. 14**). The hilum and pulmonary ligament descending from it are notable. The cardiac impression is deeper on the left lung because of the position of the heart. The aorta arches over the left hilum, and the azygos vein over the right. Joining the right cardiac impression are the groove for the superior vena cava in front of the hilum and that for the inferior vena cava in front of the pulmonary ligament. Other impressions shown in the figures are those of the esophagus and trachea, the left subclavian artery, and brachiocephalic or innominate vein. The borders of the lung are pinched extensions between the pericardium and body of the sternum (anterior border) and between the diaphragm and body wall (inferior border). The inferior limit of the thoracic cavity on both sides is related to ribs 8, 10, and 12, whereas the lower borders of the lungs stop at ribs 6, 8, and 10, and the unoccupied space is the costodiaphragmatic recess on both sides. The cardiac notch is the absence of lung because of pressure from the heart. A similar, but smaller, notch is made in the underlying pleura, which gives rise to the left costomediastinal recess.

The oblique fissure cuts through the costal, diaphragmatic, and medial surfaces to the root of the lung. In the right lung the horizontal fissure runs backward from the anterior border and meets the oblique fissure in the midlateral line. Thus the right lung has three lobes, superior, middle, and inferior, whereas the left lung has two, the superior and inferior. The lingula, or anteroinferior part of the left upper lobe, and the cardiac notch above it correspond to the middle lobe of the right lung.

Bronchopulmonary segments. For convenience in exploration and study of the lung, it may be divided into anatomical areas. The bronchial tree branches mainly by dichotomy (**Fig. 15**). The ultimate generations, that is, the respiratory bronchioles, alveolar ducts, and alveoli constitute all of the respiratory portion of the lung. This respiratory portion (**Fig. 16**) consists of 10 segments in the right lung and 8 in the left, each of which is supplied by a tertiary branch of the bronchial tree (Fig. 15). Two or more of the segments make up a lung lobe. Further, the primary bronchi (from the trachea) divide as secondary bronchi, three on the right side and two on the left, corresponding to the lobes of the lung. Because the upper left lobe results from the fusion of two lobes, the prospective left upper and middle lobar bronchi become partially fused. As a result, the four segmental bronchi of the upper lobe are not tertiary, but bronchi of the fourth division of the tree. However, there is great similarity in the structure of the right and left bronchi.

The trachea and extrapulmonary bronchi are kept open by C-shaped bars of hyaline cartilage. Within the lung the bars are arcs staggered in the bronchial walls at different levels. When in their branching the bronchi and bronchioles are reduced to a diameter of 0.04 in. (1 mm) or less, they are then free of

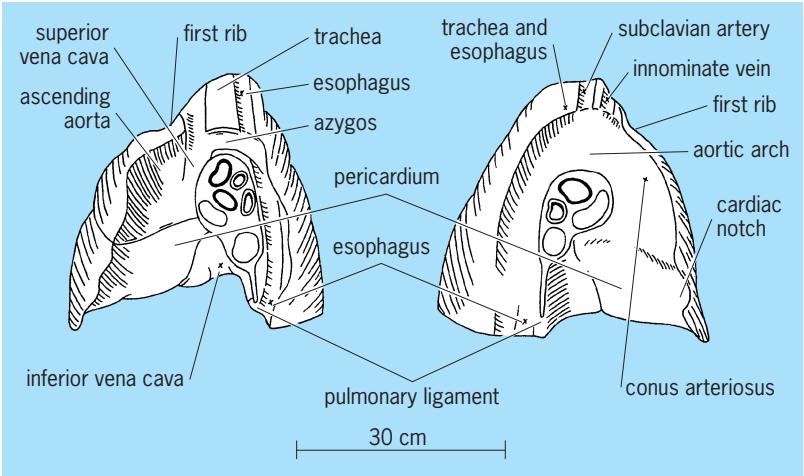

Fig. 14. Impressions on the mediastinal surfaces of the lungs.

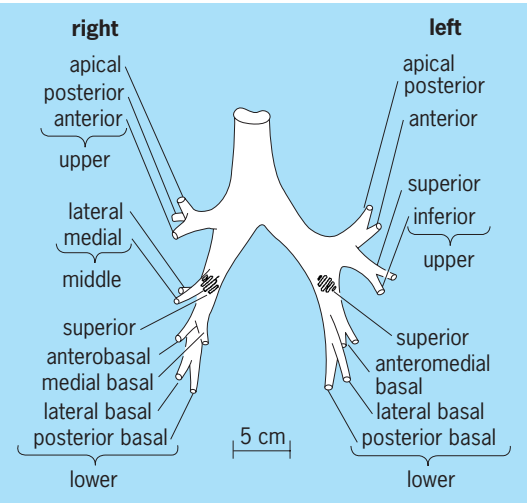

Fig. 15. Dichotomous branching of bronchial tree, showing 10 right and 8 left segmental bronchi.

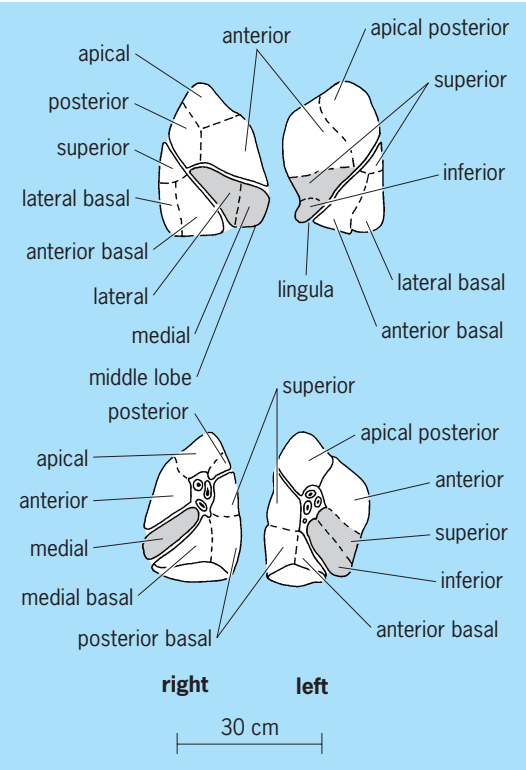

Fig. 16. The 10 right and 8 left bronchopulmonary segments.

cartilage and are called terminal bronchioles. One of the terminal bronchioles enters the apex of a secondary lobule of the lung. These secondary lobules are anatomic units of the lung, whose hexagonal bases, 0.4–0.8 in. (10–20 mm) in diameter, rest on the pleura (Fig. 16) or next to a bronchiole or blood vessel, and whose apices point toward the hilum. Finer lines divide the bases of the secondary lobules into smaller areas (**Fig. 17**). These are the bases of primary lobules, each served by a respiratory bronchiole.

Blood supply. The blood supply to the lung is provided by the pulmonary and the bronchial arteries. The right pulmonary artery runs dorsally between the bronchi of the right upper and middle lobes so that the upper bronchus arises 1 in. (25 mm) from the trachea. The left pulmonary artery and the arch of the aorta pass dorsally between the trachea and the bronchus of the left upper lobe bronchus, so that the latter is 2 in. (50 mm) from the trachea. Each pulmonary artery divides into 10 branches, which follow closely the posterosuperior walls of the segmental bronchi. They take the names of the 10 right segmental bronchi.

The bronchial arteries arise on the left side from the aorta, and on the right from either an intercostal or the left bronchial artery. They supply the walls of the bronchi, pulmonary vessels, and lymph nodes. They pass with the radicles of the pulmonary vein through the interlobular septa and supply the pulmonary pleura. Blood delivered to the lung by the bronchial arteries is returned by the radicles of the pulmonary veins, except that to the largest bronchi, which is returned by the bronchial veins to the azygos veins.

The pulmonary veins have 10 branches; their main stems run in the medial or inferior sides of the bronchi. Because their tributaries run intersegmentally and drain adjacent segments, and because the arteries may cross intersegmental boundaries, a bronchopulmonary segment would not be a morphologic bronchovascular unit. The upper and lower lobes of the left lung send an upper and a lower pulmonary vein, respectively, into the left side of the left atrium. The upper and middle right lobes provide the upper vein to the right side of the atrium.

Lymphatic channels are not found in interalveolar partitions but occur everywhere else in the lung, except in cartilage. There are two main sets of lung lymphatics, those in the pleura and those within the lung. The latter begin at the alveolar ducts and follow the bronchi and pulmonary vessels to the lymph nodes at the hilum. Those in the pleura form secondary and primary networks in the lobular septa and drain into the hilar lymph nodes.

Nerve supply. The nerves which supply the lung are branches of the vagus and of the thoracic sympathetic ganglia 2, 3, and 4. Efferent vagal fibers are bronchoconstrictor and secretory, whereas the afferents are part of the arc for the breathing reflex. Efferent sympathetic fibers are bronchodilators; hence, the use of adrenalin for relief of bronchial spasm resulting from asthma.

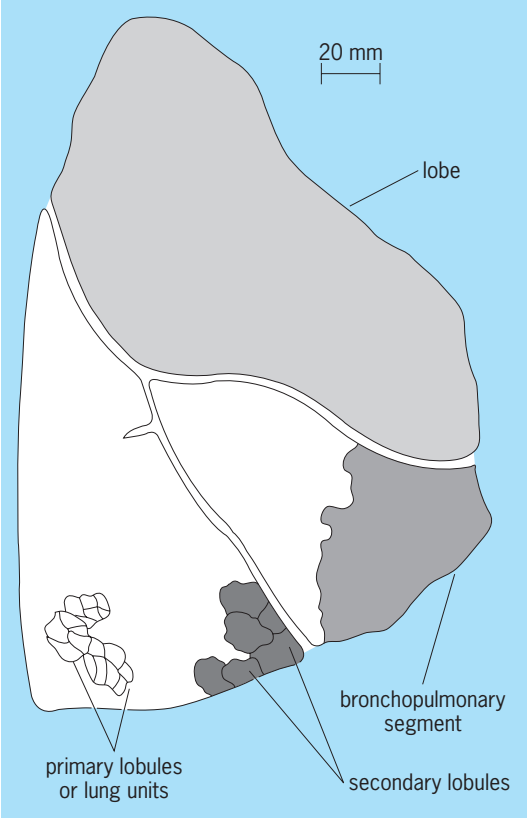

Fig. 17. Subdivisions of the lung.

Histology of the Lung

Essential to lung function, histologically, are the respiratory passages, pleura, blood vessels, lymphatics, and respiratory surfaces proper.

Epithelium. Between the cartilage in the walls and the surface epithelium of the air passages are serous and mucous glands, which continuously pour their secretion through ducts to the surface of the epithelium. Also between the cartilage and epithelium below the trachea are smooth muscles which increase relatively in amount with decrease in size of the bronchioles. In the terminal bronchioles, which are free of cartilage, smooth muscle is the main constituent and may close the bronchiole, as in asthma. The epithelium is the respiratory type; that is, the cells are tall and capped by motile cilia. The cilia wave toward the throat, carrying along the secretions of the glands laden with inspired dust.

Pleura. The pleura lines the outer surface of the lung and the inner surface of the pleural cavity. It consists of a layer of flattened mesothelial cells and an underlying layer of collagenous and elastic fibers which support numerous blood and lymphatic capillaries, fibroblasts, and macrophages. The pleura continuously pours out a mucoid exudate, which lubricates the opposed surfaces. The pleura is pink in the newborn child.

Respiratory portion. This portion includes the respiratory bronchioles, which branch from the terminal bronchioles, the respiratory ducts, alveolar sacs, and alveoli. The respiratory ducts are similar to terminal bronchioles, except that they are smaller and have a few scattered alveoli protruding from their walls. Alveolar ducts, the next order of branching, have their walls studded with contiguous alveoli so that the openings of the alveoli occupy the greater area of the wall. Alveolar sacs are terminal dilations of the alveolar ducts. Alveoli are the smallest functional units of the lung. Their walls are made up of one or two capillaries, a few elastic fibers, and more reticular fibers, among which are a few fibroblasts, dust cells or lung macrophages, and septal cells. The septal cells give rise to some of the macrophages, and others are carried in by the blood. Macrophages from both sources wander freely in the alveolar spaces. Two alveoli frequently have a single wall in common. The epithelial lining of the alveoli is so thin that it cannot be demonstrated with ordinary techniques. The greater volume of the lung is taken up by air space of the respiratory portion. Thus the lung feels spongy and, when seen in sections, looks like fine lace.

Embryology of the Lung

The lung is an organ adapted to respiratory exchange of gases between air and the blood of vertebrates. It ranges from the simple swim bladder of dipnoid fish to the large spongy, compound air sac of humans. The story of the lung includes the evolution of certain vital accessories to respiration, the diaphragm and thoracic wall. All lungs correspond in originating as a pocket from the pharynx near the level of the sixth aortic arches which furnish the blood supply; in innervation by the vagus nerve; and in their embryonic position in the abdomen. This position is retained in the adult of fish, amphibians, and reptiles and partially retained in birds, but in mammals the lung is segregated from the abdomen in fetal life into a new-formed thoracic cavity (**Fig. 18**). *See* PHARYNX; SWIM BLADDER.

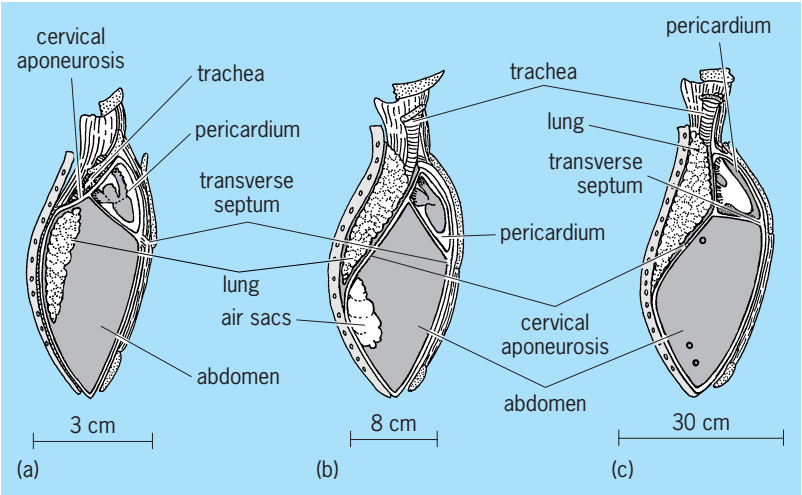

Fig. 18. Position of the lung in vertebrates. (*a*) Fishes and amphibians. (*b*) Birds. (*c*) Mammals.

Extrusion of lung. In the extrusion of the lung from the abdominal cavity it first invaginates the cervical

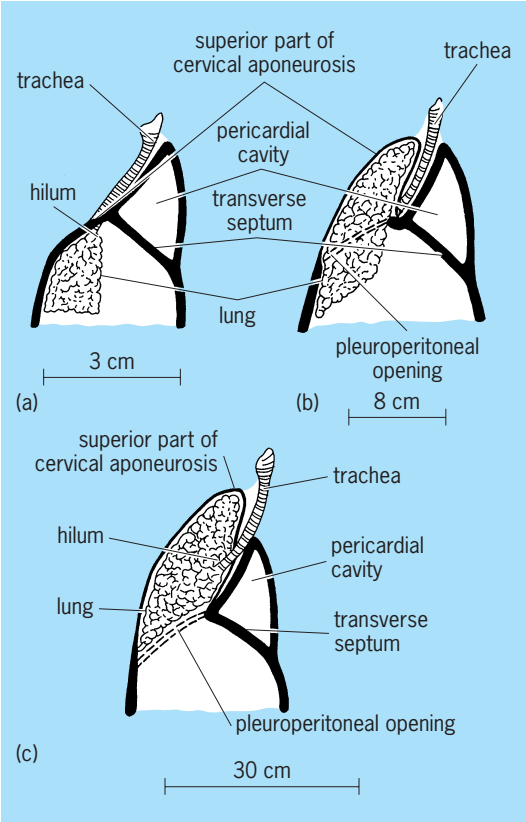

Fig. 19. Lung extrusion in (*a*) amphibians, (*b*) birds, and (*c*) mammals.

aponeurosis of the abdomen and in so doing enters the space being formed between the inner and middle layers of the body wall, the pleural cavity. These events bring the lung to an anterior, thoracic position, and also allow it to develop an apex, both important in respiration. In amphibians (**Fig. 19***a*) the lung comes to lie in the pleuroperitoneal cavity. It is infrapericardial, with both the root and apex in the same plane. Figure 19*b* represents the condition in birds in which an apex grows cranially beyond the root of the lung. Thus the entire lung remains abdominal, because the apex carries forward the entire cervical aponeurosis. The human lung becomes supradiaphragmatic (Fig. 19*c*). The diaphragm is represented by only a thin layer of the aponeurosis that is carried forward (Gibson's fascia), and the remainder of the aponeurosis serves to close the passageway. Moreover, the root also is carried forward, and in humans it becomes closely bound to the pericardium (**Fig. 20**).

The pericardial bond is of primary importance in humans and the great anthropoids because the pericardium is bound also to the central tendon and diaphragm (Fig. 20). The part of the lung above the pulmonary root is relatively large, and the apex, located at a fixed point at the level of the neck of the first rib, can expand when the diaphragm contracts and pulls the root and heart in a downward and forward direction (**Fig. 21**). Upon expiration they move in the reverse direction. In the typical mammal there

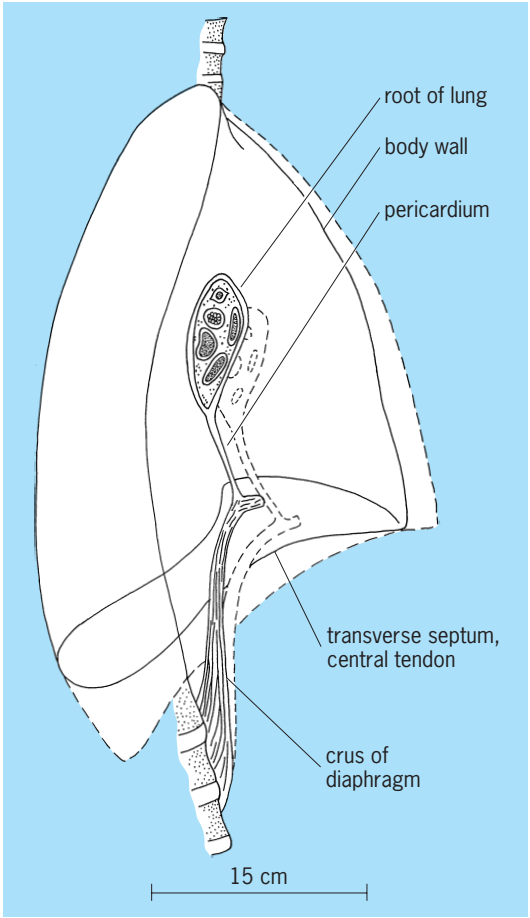

Fig. 21. Diagram showing movements of the diaphragm, pericardium, and root of the lung during respiration in higher mammals. Broken lines show position of thorax and viscera during inspiration.

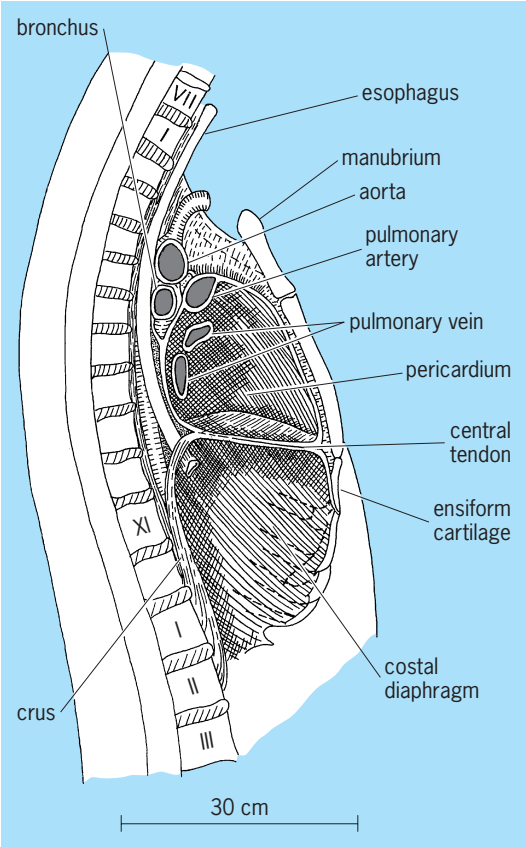

Fig. 20. Diagram showing closely bound connections of the diaphragm, pericardium, and root of the lung in a human.

is little or no contact of the diaphragm with the pericardium, and the part of the lung above the root is relatively small (**Fig. 22**). Here the descent of the diaphragm has little effect on the movement of the heart and pulmonary roots. Instead, the azygos lobe expands between the esophagus and inferior vena cava into the subpericardial space beside the pericardial mesentery.

The body wall becomes an inspiratory mechanism in reptiles as a result of the ribs, sternum, and intercostal muscles which evolve in the middle layer of the wall. In phylogeny the first intermuscular septum to extend completely to the median ventral line is that behind the seventh segment of the rectus abdominis in the amphibian *Necturus maculatus*. In this septum the first sternal rib develops in typical reptiles and mammals.

Amphibia. The amphibian lung is bifid, and each half projects into the abdomen above the pericardium and liver (**Fig. 23**). The inner surface of the lung presents folds and ridges which increase respiratory surfaces. Air is pumped into the lung by pharyngeal muscles and in the absence of ribs is forced out by abdominal muscles.

Reptilia. The collapsed lung of *Iguana* (**Fig. 24**) extends forward to the cervical aponeurosis. The 8th

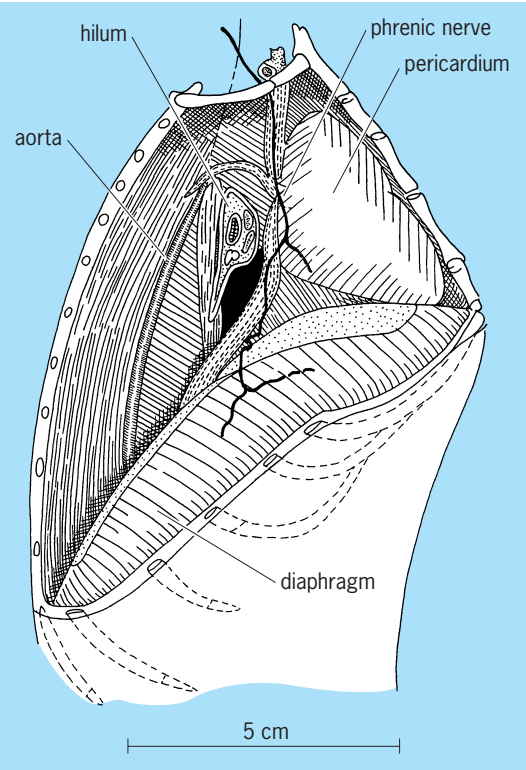

Fig. 22. Diagram showing discontinuity of the diaphragm with the pericardium and hilum of the lung in the typical mammal (rabbit).

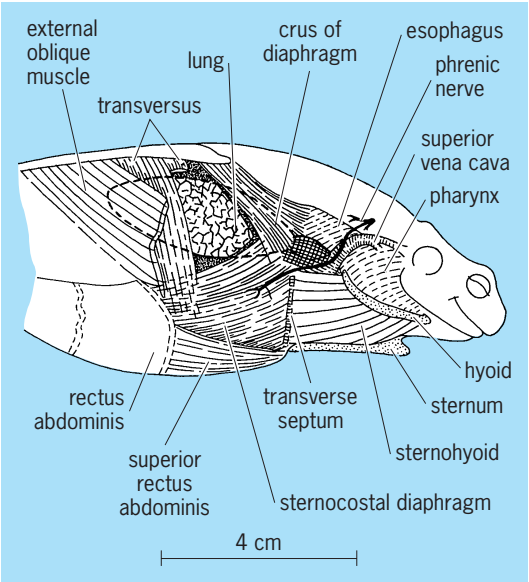

Fig. 23. Diagram showing the lung in an amphibian (Surinam toad) and the muscles from which the diaphragm has evolved.

to 14th ribs are complete, and three cervical ribs can both expand and compress the apex of the lung. The sternum is a strong fulcrum for the costal cartilages. The cervical part of the trunk is increased in length, so that the head and pharynx are carried away from the heart. Whereas in amphibians the swallowing of air tends to force pulmonary blood out of the lungs, in the lizard the negative pressure produced by its costal mechanism draws both air and blood into the lungs.

Birds. Lungs in adult birds are extra-abdominal, except the posterior ends, which are dilated to form abdominal air sacs. However, the avian lungs have grown through the cervical aponeurosis of the abdomen into a new space in the body wall, the pleural cavities (**Fig. 25**). Air sacs develop in this aponeurosis, and the septum is divided by these into two layers, the dorsal or pulmonary and ventral or abdominal (Fig. 25). If these lungs were removed from the thorax and the avian septum were replaced against the dorsal wall, the three-layered body wall of the amphibian would be restored. Although the lungs are in the thorax, the air sacs remain in the abdomen.

Human embryo. The lung in the 5-week human embryo lies in the abdomen. The septum transversum in humans consists of a ventral part (**Fig. 26*b*, *c***) representing the entire septum of amphibians (Fig. 19*a*) and a dorsal part representing a fusion of the amphibian cervical aponeurosis with the dorsal wall of the pericardium. The common cardinal vein (Fig. 26*b*, *c*)

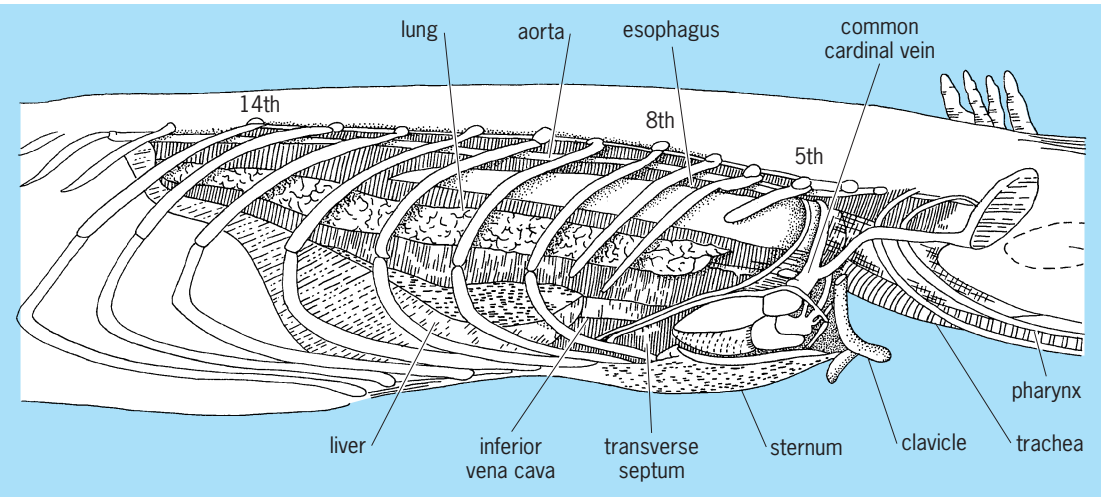

Fig. 24. Diagrammatic sketch of a reptile, *Iguana* sp., showing the collapsed abdominal lung and the rib structures.

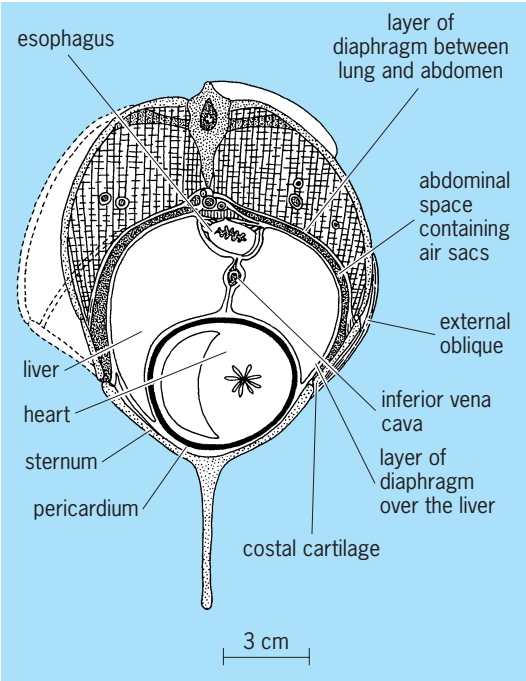

Fig. 25. Cross-sectional diagram of thorax (lung), abdomen, and pericardium in pigeon.

runs in the anterior part of the septum next to the wall of the pericardium. The lung lies medial to the Wolffian fold, which is attached to the mesonephros dorsally, the liver ventrally, and the septum in front,

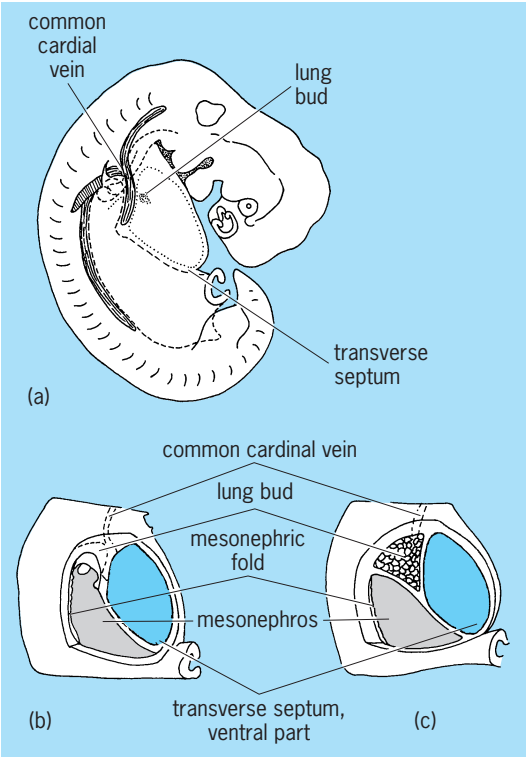

Fig. 26. Human embryo. (*a*, *b*) Diagram of embryo of 5 weeks showing initial process of lung extrusion. (*c*) Embryo of about 7 weeks showing extrusion into the cervical aponeurosis.

with its free border facing backward. In the amphibians the lungs grow backward in the abdomen and are attached to the mesentery between the aorta and esophagus above and liver and stomach below. In the human embryo they advance medially to the septum (Fig. 26*b*, *c*), then evaginate laterally into the dorsal part of the septum between its anterior part, which contains the common cardinal vein next to the pericardial roof, and the posterior part, which develops to become the pleuroperitoneal membrane.

After this extrusion of the human lung from the abdominal cavity, the lung lies next to the medial surface of the middle layer of the body wall. Upon removal of the body wall, the lungs are exposed. The lungs can grow in a forward or backward direction and ventrally between the inner and middle layers of the body wall at equal rates, as these open up. The site of extrusion (pleuroperitoneal opening) need only close, and the lung will occupy new, completely enclosed, pleural cavities.

Leo P. Clements

Physiology of the Lung

Simple molecular diffusion of gases underlies respiratory exchange in all animals. The lungs of air-breathing vertebrates require mechanical ventilation to sustain this diffusion. The molecular characteristics of respiratory gases (**Table 1**), the physical properties of atmosphere (**Table 2**) and of body fluids (**Table 3**), the physiological limitations on lung structure and location, and the large requirement of the active animal for O_2 uptake and CO_2 removal (**Table 4**), all restrict the effectiveness of diffusion respiration alone. Diffusion suffices only within a few millimeters of those membranes across which interchange of O_2 and CO_2 can occur. Refinements in ventilation respiration intervene as continuing speciation involves increased O_2 requirements. The functional as well as structural changes follow recognizable evolutionary trends. *See* METABOLISM; RESPIRATION.

Other vital requirements. Besides exchange of O_2 and CO_2, there are other vital requirements for breathing processes in air. Among these are (1) limitation of the evaporation of water in general and control of the dissipation of heat in birds and mammals in particular; (2) ventilation of the olfactory membrane; (3) production of high intrapulmonary pressure for such acts as coughing, defecation, and parturition; and low intrapulmonary pressure for regurgitation; (4) provision for large pulmonary volume in such acts as sneezing, yawning, and phonation; and (5) allowance for control of buoyancy in aquatic animals. All these ancillary processes involve corresponding variations in the respiratory mechanisms. As animals are compared, differences are found in buccopharyngeal, pulmonary interstitial, and diaphragmatic musculature for inducing and directing movement of air in valvular muscle of nares, glottis, and pulmonary ducts or apertures for controlling flow and pressure; in accessory air sacs to increase capacity for ventilation itself, or for phonation; in heart for adequate pulmonary blood supply and in circulation, for efficient perfusion; in brain and peripheral nerve

TABLE 1. Characteristics of the respiratory media

Variable	Water		Atmosphere (N_2)	
	Ocean	Fresh	Sea level	6000-m altitude
Temperature, °C	−2.0 to 30.0	2.0–32.0	0.7–15.7	−28.1 to −15.1
Pressure, total mmHg	760–760,000	760–20,000	760	347.5 to 360.2
Density, g/liter	1027* (20°C)	1000* (4°C)	1.223–1.290	0.649–0.659
pH	7.5–8.4	3.2–10.6		
Concentration, vol %				
H_2O	94–97	97–99	1.00†	1.00†
N_2	1.03* (15°C)	1.33* (15°C)	78.03 (STP)	78.03 (STP)
CO_2	0.02* (15°C)	0.03* (15°C)	0.03 (STP)	0.03 (STP)
O_2	0.58* (15°C)	0.72* (15°C)	20.99 (STP)	20.99 (STP)
Salts	3.46*	0.18*		
Inert gases	Trace	Trace	0.95 (STP)	0.95 (STP)
Partial pressure (tension), mmHg				
H_2O	12.79 (15°C)	6.10 (4°C)	6.40‡ (15°C)	0.72‡ (−15°C)
N_2	593.02 (STP)	593.02 (STP)	593.02 (STP)	281.06 (STP)
CO_2	0.23* (STP)	0.23* (STP)	0.23 (STP)	0.11 (STP)
O_2	159.52* (STP)	159.52* (STP)	159.52 (STP)	75.61 (STP)
Inert gases	7.46 (STP)	7.46 (STP)	7.46 (STP)	3.42 (STP)
Total pressure	760.00	760.00	760.00	360.20
Diffusion coefficient, ml/(min)(cm^2)(cm), at 760 mmHg, 20°C				
H_2O				
N_2		0.000018 (0.53)§		
CO_2		0.000785 (23.1)§		
O_2		0.000034 (1.0)§	11.0	

* Average of many determinations; varies with conditions of measurement.
† Varies, but never absent and always of biological significance.
‡ Calculated for 50% relative humidity.
§ Values in parentheses are relative coefficients with O_2 as unity.

patterns for initiation, maintenance, and adjustment of ventilatory processes, control of circulation, and regulation of ancillary functions; and in skeletal and connecting structures. A brief review of comparative respiration can only delineate the basic trends in phylogeny of respiration. Certain general features of the breathing process must be considered before comparisons among species can be made. *See* SPEECH.

Breathing. The act of breathing induces mass flow of air in and out of the respiratory organs. This inspiration and expiration mechanically ventilates the entire lung of those animals, such as lungfish, salamander, and frog, in which the organ is a simple sac. Here the inner lining or pulmonary epithelium, where diffusion itself occurs, is directly exposed to the tidal airflow caused by breathing. In these animals, breathing is intermittent or periodic; that is, cycles of inspiration and expiration occur in groups after many minutes of nonbreathing or apnea. The periods of apnea may exceed those of ventilating. The lung is normally closed from the atmosphere by a valve (glottal valve) at the opening into the trachea supplemented by valves (nasal valves) at the nares. The lung, however, does not remain a simple sac. Beginning with toads and climaxing in birds and mammals, more elaborate secondary and tertiary sacculations, the alveoli and the alveolar or air sacs, evolve. This evolution increases the diffusion surfaces tremendously and also removes these surfaces from direct exposure to ventilatory airflow. Thus, a uniform gaseous exchange medium comes

TABLE 2. Characteristics of respiratory molecules*

Type	Weight (O = 16)	Diameter,† cm × 10^{-8}	Density, g/liter	Mean free path, cm × 10^{-6} (750 mmHg)	Collision frequency (20°C)	Average velocity, cm/s	Water solubility STP	Water solubility 20°C	Vol % (40°C)
N_2	28.02	3.15–3.53	1.251	8.50	5070	45,400	2.35	1.54	1.18
H_2O	18.02	3.0–5.0	0.0005–0.030‡			56,600			
CO_2	44.01	3.34–3.40	1.977	5.56	6120	36,200	171.3	87.8	53.0
O_2	32.00	2.92–2.98	1.429	9.05	4430	42,500	4.89	3.10	2.31

* Unless otherwise indicated, values are for standard conditions (STP) of temperature (0°C) and pressure (760 mmHg).
† Range indicates variability with method of measurement (such as viscosity, heat conductivity).
‡ Water vapor in saturated air, that is, in equilibrium with water at 0°C and 30°C.

TABLE 3. Respiratory exchange characteristics in humans

	VENTILATION					
	Inspired air		Alveolar air[a]		Expired air[a]	
Gas	Composition, vol %[b]	Partial pressure, mmHg[c]	Composition, vol %	Partial pressure mmHg[d]	Composition, vol %[b]	Partial pressure mmHg[d]
H_2O	0.00	5.7	00.0	47	00.0	47
N_2	79.02	596.0	80.4	573	79.2	565
O_2	20.95	158.0	14.0	100	16.3	116
CO_2	0.03	0.3	5.6	40	4.5	32

	TRANSPORT[e]							
	Arterial		Capillary		Tissue fluid		Venous	
Gas	vol %	mmHg	vol %	mmHg	vol %	mmHg	vol %	mmHg
H_2O	83 (81–86)	47	83 (81–86)	47	83 (81–86)	47	83 (81–86)	47
N_2	0.975	573	0.975	573	0.975	573	0.975	573
O_2	19.6 (17.3–22.3)	94	1–22.3[f]	1–94[f]	0.185[f]	30[f]	12.9 (11.0–16.1)[g]	40
CO_2	48.2 (44.6–50.4)	40	44.6–57.7[f]	40–50[f]	3.046[f]	50[f]	54.8 (51.0–57.7)[g]	46

[a] Alveolar air, actually last part of expired samples.
[b] Dry air, partial pressure in mmHg = (vol %)/100 × 760 mmHg (Dalton's law).
[c] Ambient air (slight variations exist), in vol %, = (100 × mmHg)/760 (Dalton's law).
[d] Physiological air, normal temperature (37°C) and standard pressure (760 mmHg).
[e] Values in parentheses are ranges.
[f] Variable, depending on blood flow, tissue activity, and relation of sample to capillary length or field.
[g] Internal jugular.

to prevail at the respiratory membranes of birds and mammals, the breathing cycles become continuous and shallow (eupneic), and the lung remains normally open to the atmosphere. Nasal valves are usually absent in animals above reptiles, and the glottis closes only for such occasional acts as swallowing and coughing.

Breathing in vertebrates. Breathing induces air movement resulting from a mechanically imposed difference in pressure between a compressible cavity and the atmosphere. In mammals and reptiles this cavity is the lung itself. In birds and amphibians, however, secondary cavities are responsible, namely, special air sacs in the former and the mouth (buccopharyngeal cavity) in the latter. These secondary cavities communicate with both the lung and the atmosphere and accommodate the bellowslike action. Two basic processes of air breathing reflect the structural differences. The more primitive type, buccopharyngeal breathing, is found in lungfish and amphibians. It involves the same basic neuromuscular elements of mouth and throat as water breathing in fishes. The more specialized type, thoracoabdominal breathing (in humans, for example), involves trunk musculature to supply pulmonary ventilation in all reptiles, birds, and mammals. A rather continuous buccopharyngeal ventilation, not always involving pulmonary ventilation itself, is characteristic of amphibians (**Table 5**) and persists in lizards, turtles, and other reptiles. In amphibians, therefore, buccopharyngeal activity subserves both breathing and olfaction; in reptiles it subserves only the olfactory sense (smell).

TABLE 4. Combined measurements of respiratory values in pigeon, duck, and chicken to compare roles of various sacs and lung

		Composition	
Category	Volume, % of total	O_2 %	CO_2 %
Tidal air	10–15		
Inspired	10–15	21.0	0.03–0.04
Expired	10–15	13.5	5.1–6.5
Interclavicular	20–25	14.6	5.0–6.9
Abdominal	50–70	19.0	1.9–2.7
Lung	10–12	20.0*	1.0*

*Estimated for parabronchi.

Buccopharyngeal breathing. Buccopharyngeal breathing is indirect, when compared with thoracoabdominal. It involves two distinct stages: ventilation of the mouth and ventilation of the lungs. The necessary pressure gradients between mouth and atmosphere and between mouth and lungs are generated by muscles which raise and lower the hyoid apparatus and floor of mouth and throat; the same mechanism is used for water breathing in fishes. Pulmonary inspiration in buccopharyngeal breathing is more descriptively an injection or adspiration. The volume of air inspired or expired per breathing cycle is called tidal volume.

In the frog, movement of air between mouth and atmosphere requires only about 3–5 mm of water (H_2O) pressure; that between mouth and lungs reaches 25–35 mm H_2O pressure during the peak of pulmonary inspiration. A gradient is directed toward the lung of about 20 mm on inspiration, because a volume of gas (functional residual volume) remains in the lung from the preceding expiration under about 10 mm H_2O pressure. The glottis closes at the end of inspiration, and a positive intrapulmonary pressure persists during apnea, of about 20 mm H_2O, as a result of elastic recoil and muscular

TABLE 5. Some respiratory values in vertebrates

Animal*	Weight, kg	Breathing rate,† cycles		Tidal volume, ml	Minimum volume, liter/min	O_2 consumption, $mm^3/(g)(h)$	Intrapleural pressure, cm H_2O	Compliance, liter/cm H_2O
		Vent./min	Comp./h					
Frog (*Rana fusca*)		41	17 (26°C)			210.0 (20°C)		0.001
Turtle (*Malaclemys centrata*)		3.7 (24°C)		14.0	0.051	35.0 (24°C)		
Alligator (*Alligator mississippiensis*)						8.9 (22°C)		
Canary (*Serinus canarius*)		108				2900.0		
Chicken (*Gallus domesticus*)		17		45.0		497.0		
Duck (*Anas* sp.)		42		36.5		800.0		
Rat (*Rattus norvegicus*)	0.273	60	26	1.4	0.074	770.0	−2 to −8	0.0012
Dog (*Canis familiaris*)	20	17	5	302.0	5.30	580.0	−5.4 to −13.5	0.09
Horse (*Equus caballus*)	696	12		9060.0	107.0	250.0	−8.0 to −22.0	0.80
Human (*Homo sapiens*)	66	14	3	372.0	5.04	220.0	−3.8 to −9.3	0.20

* Adult male at rest.
† Vent. = ventilatory, except buccopharyngeal in frog; comp. = complementary, except pulmonary in frog (see text).

tonus in the lung itself and in the body wall. These forces also cause expiration when the glottis opens. Whether buccal ventilation alone or pulmonary ventilation or a combination of these occurs depends upon neuroregulatory processes which determine the relationship of nasal and glottal valves with each other and with the breathing musculature. Thus a frog's lungs and body can be distended greatly beyond normal dimensions by successive inspirations alone.

Vital capacity. The excess capacity of any animal to inspire beyond normal tidal volume is called inspiratory reserve volume, and the total breathing capacity of the lungs as measured by the volume which can be completely expired after maximum filling is the vital capacity. Included in this is an amount, called expiratory reserve volume, which can be expired from the functional residual volume. Total expiration might completely empty the simple lungs of some amphibians. However, as the lungs of animals elaborate with alveolar development, it is not possible to expire all lung contents. The remainder after limit of vital capacity is reached is called residual air, and its presence at metamorphosis, hatching, or birth always indicates that breathing has started. The various volumes and pressures of ventilation have not been measured in most species. Some representative values are given in Table 5.

Dead space. Buccopharyngeal ventilation continues in reptiles, but this mechanism no longer provides for pulmonary ventilation. It remains an important adjunct to breathing, however, because it serves to reduce the dead-space volume. Dead-space average normal volume for an adult human, for example, is about 150 ml and tidal volume about 500 ml, which means an actual ventilatory volume of about 350 ml. In a few air-breathing animals which occupy an aquatic habitat, respiration is apparently supplemented by buccopharyngeal breathing of water. Such a process has been described for a few species of turtles which have an especially vascular pharynx, but whether an important amount of O_2 is thus derived has not yet been ascertained. This ventilation with water may actually subserve olfactory and gustatory senses.

Thoracoabdominal breathing. Pulmonary ventilation in reptiles utilizes a more familiar process than buccopharyngeal breathing; that is, it depends upon an aspiration or sucking inspiration, such as in birds and humans. This involves development of a movable rib basket and elaboration of the intercostal musculature which, by enlargement of the body cavity, produces on inspiration a negative pressure in the lungs (intrapulmonary) with reference to the atmosphere. The abdominal muscles and myoelastic tissue of the lungs and air sacs remain, as in amphibians, important in expiration, but these are augmented by striated muscular membranes which form diaphragms and also ensheath the lungs in some turtles.

Turtles are exceptional among reptiles and air breathers generally because the ribs are fused into a shell which prohibits expansion of the body wall for inspiration. In these animals, muscular membranes which enclose the viscera and others which form diaphragms at the leg pockets in the shell produce expiratory and inspiratory force, respectively.

Reptiles. In reptiles generally, when at rest, breathing cycles occur in groups which are interspersed among long intervals of apnea. During apnea the glottis is closed, the lung air is under a few millimeters of mercury positive pressure, and buccopharyngeal ventilation waxes and wanes to a degree associated at least in part with the extent of sensory disturbance. This buccopharyngeal activity resembles olfactory sniffing, as seen in dogs. The reptilian cycle of pulmonary inspiration and expiration is much simpler than the frog's, because the glottis and nares remain open while breathing movements occur, and air moves freely between lung and atmosphere in direct response to the action of breathing muscles in trunk and viscera. These muscles act alternately to enlarge and reduce the body cavity, exerting changes in pressure through tissue fluids directly on the lung air itself. Inspiration is clearly by suction. This is in contrast to the injection action of the buccopharyngeal muscles in amphibians. Like amphibians, however, the breathing remains periodic, and the glottis is normally closed.

Birds and mammals. Birds and mammals utilize strictly thoracoabdominal breathing to ventilate both nasal

and pharyngeal cavities and lungs. The same structures used in reptiles continue to operate, namely, trunk musculature and diaphragms. The glottis does not close the lung from atmosphere normally, however; nasal valves are absent, and buccopharyngeal movements cease. Because the lung and airway are greatly elaborated over those in previously considered classes of animals, certain central regions develop more critical and specific control than glottal valves over airflow and diffusion. Such control regions involve, in mammals, the conducting bronchi and bronchioles and the valvelike sphincters at alveolar sac openings; and in birds, the corresponding ducts and orifices, including parabronchi and air capillaries. Also, other discrete structures develop as bellows, to change breathing forces into ventilation pressures. These structures are the extrapulmonary air sacs in birds (air sacs also occur in some reptiles) and the distal ducts and alveolar sacs in mammals. The alveolus becomes more strictly a diffusion exchange unit. As a result of all these specializations, great stability of physical conditions at diffusion surfaces is achieved in a system which, at the same time, accommodates high and variable rates of exchange.

The muscles of bronchioles, like those of arterioles in the circulatory system, are strategically situated to modify the distribution of inspired air in lungs by affecting duct caliber. However, not only distribution but also direction of airflow is altered in the bronchi of birds, a fact long obscured by emphasis on the simple, uniform mammalian tidal pattern.

It is now well established for breathing in birds that the tidal flow of air in and out of the conducting airways between atmosphere and air sacs is interrupted in certain bronchi leading to the lung itself; the flow is changed from tidal to unidirectional before reaching most air capillaries. A detailed explanation of this change of flow is still obscure, but the advantages of the change are apparent, namely, (1) continual exposure of respiratory surfaces to atmospheric air and the highest available O_2 concentration; concomitant with (2) optimum conditions for diffusion control of H_2O and CO_2 in pulmonary fluids. Birds thus surpass mammals in physiological access to oxygen.

There are other elements of breathing besides respiratory exchange which require control of the air movement itself as a mass flow along pressure differentials, as distinguished from control of movement in molecular migration along diffusion gradients. For example, whenever panting or phonation affects breathing, or when pulmonary pathology occurs, the amount of air going to different parts of the lung might greatly affect underlying respiratory exchange and disrupt diffusion control. But homeostasis requires a stable gaseous composition within alveoli and a minimal evaporation of water from alveolar surfaces; and diseased or injured lung regions require immobilization. Consequently, panting thermoregulation depends primarily on conduction and evaporation only from the upper air passages, and also from peripheral air sacs in birds. The cooling fluid is provided mainly through controlled secretion from serous membranes or glands, such as salivary glands in dogs. The membranes of passages and air sacs are not adapted for respiratory exchange, and insignificant amounts of O_2 or CO_2 diffuse across them. Phonation likewise uses the segregated air capacity of the ducts and sacs to move large volumes and produce sounds without affecting gaseous molecular exchange.

The structural differences between the respiratory units of bird and mammal are subordinate to their common functional characteristics. The alveolus of mammals is usually a terminal membranous pouch with a porelike orifice (the postmortem diameter is about 70 micrometers in the rabbit) into an alveolar sac; occasionally they are appendant on a respiratory bronchiole. The corresponding structure in birds, called an air capillary or cylindrical alveolus, is an appendant membranous tubule with an orifice (diameter about 50 μm in the chicken) into a tubular parabronchus. These alveolar structures frequently interconnect, via pores in mammals and anastomoses in birds. They both provide, in the aggregate, a reservoir with tremendous surface area (at least 50 m^2 in humans) which sequesters a mechanically stable intra-alveolar atmosphere based on nitrogen. It is through this medium that large quantities of oxygen, carbon dioxide, and water vapor molecules diffuse according to pressure gradients established, as in all vertebrates, on one hand by ventilation and on the other by pulmonary circulation. This intrapulmonary atmosphere, made up of the slightly varying composition among millions of alveoli, is often called alveolar air. It stays remarkably uniform in total composition despite ten- to twentyfold variations in oxygen and carbon dioxide exchange, such as occur during exercise. Some standard resting values for ventilation and diffusion in respiratory exchange of humans are shown in Tables 3–5.

The microscopic size of alveoli exaggerates a persistent tendency of all such bubblelike or balloonlike structures to collapse. This tendency in alveoli is partly the result of the surface tension of a fluid film which bathes their gas-exposed surface and partly the result of the elasticity of stretched fibers widely distributed within pulmonary tissues. Actual collapse (atelectasis) will occur at once as a result of these forces, for example, if much gas (pneumothorax) or fluid (hydrothorax) accidentally enters the pleural cavity.

A lipoprotein material is secreted onto the alveolar surface, a material characterized as surfactant, which reduces the surface tension at the gas-liquid interphase where respiratory exchange begins. This action lowers the alveolar rate of shrinkage and consequently helps to maintain an effective area of respiratory surface.

Alveolar size and effective surface area are further supported by periodic deep inspirations which are an integral part of the respiratory patterns of mammals, with the possible exception of a few very large representatives like the horse. These are called complementary cycles, or sighs, in humans, and they arise from sensory monitoring and nervous control of the functional residual volume of alveolar air.

These cycles occur about nine times an hour in humans; in other animals they occur more or less frequently, depending on the characteristic size of the adult (Table 5). F. Harold McCutcheon

Avian respiration. The respiratory system of birds has a unique and complex structure. The lungs are relatively rigid and do not change volume significantly during ventilation; air passes through them and into a group of usually nine air sacs. There is no diaphragm separating the thoracic and abdominal cavities, and both phases of ventilation are active, even in birds at rest.

Structure. The primary bronchus originates from the trachea, runs along the entire length of the lung, and divides into two groups of secondary bronchi (**Fig. 27***a*). As the two groups of secondary bronchi enter the lung, they divide into four branches that spread across the medial and ventral surface of the lung. The primary bronchus also forms branches that extend along the lateral and dorsal surface and join the secondary bronchi by means of straight tubes (tertiary bronchi).

This network system of secondary and tertiary bronchi is known as the paleopulmo system, and it is present in all birds (Fig. 27*a*). In the penguins and emus, it is the only system, whereas in all other birds there is also a tertiary bronchial network between the lateral side of the primary bronchus, the laterobronchi, and the posterior air sacs. This network system is known as the neopulmo system (Fig. 27*b*). A meshwork of fine air capillaries, which intertwine with equally fine blood capillaries, radiates from the lumen of each tertiary bronchus. These blood capillaries arise from arterioles and form collecting venules near the tertiary bronchial lumen. Gas exchange occurs in the air capillaries leading from the tertiary bronchi.

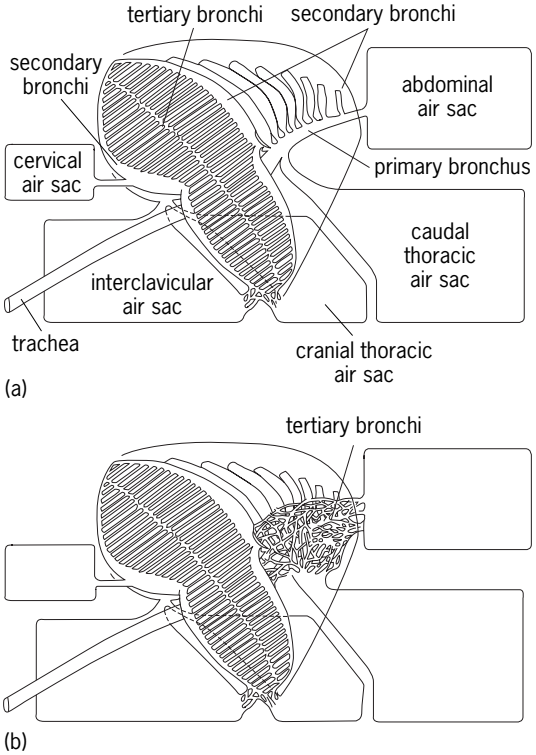

Fig. 27. Overall structure of the lung of birds (*a*) without and (*b*) with the neopulmo system (tertiary bronchi), showing the connections to the air sacs. (*After H.-R. Duncker, Structure of the avian respiratory tract, Respir. Physiol., 22:1–19, 1974*)

The cervical air sacs arise from the first medioventral bronchus of each lung, and the (single) clavicular air sac is joined to the first, second, and third medioventral bronchi of both lungs, whereas the cranial thoracic sacs connect predominantly to the third medioventral bronchus of each lung. These form the cranial (anterior) group of air sacs. The second lateroventral bronchi of each lung form large connections to the caudal thoracic air sac, and the primary bronchi continue to the caudal edge of each lung and empty into the abdominal air sacs. These latter two pairs of air sacs form the caudal (posterior) group.

Pathway of air. Airflow through the tertiary bronchi of the paleopulmo system is in the same direction during both phases of ventilation (**Fig. 28**). During inspiration, some air travels through the primary and tertiary bronchi of the neopulmo system to the posterior group of air sacs, while the remaining air travels through the mediodorsal secondary bronchi and the tertiary bronchi of the paleopulmo system to the anterior group of air sacs. During expiration, the gas in the posterior air sacs travels again through the tertiary bronchi of the neopulmo system, then into the mediodorsal secondary bronchi, through the tertiary bronchi of the paleopulmo system, and into the medioventral secondary bronchi where, together with gas from the anterior group of air sacs, it enters the primary bronchus and trachea.

There is no evidence of any mechanical valves in the airways, and the unidirectional airflow through the tertiary bronchi of the paleopulmo system is thought to result from the aerodynamic conditions in the lung. During inspiration, the velocity of airflow is important; the greater the velocity, the more likely is the inspired gas to flow pass the medioventral bronchi and continue along the primary bronchus. There are indications that the diameter of the primary bronchus decreases close to its junction with the first medioventral bronchus. This region is called the segmentum accelerans because it causes an increase in the velocity of airflow at that point.

Functional implications. Bulk airflow through the tertiary bronchi of the paleopulmo system is in the same direction during both phases of ventilation, and it has been demonstrated that bulk blood flow through the arterioles in the peritertiary bronchial tissue is at right angles to the bulk airflow. This is known as a cross-current arrangement, and it is more effective at exchanging gases (CO_2, in particular) than is the mammalian lung. It has been suggested that the unique structure of the avian lung is an adaptation to the high oxygen requirements of flight and an explanation of the better ability of birds to tolerate the hypoxia of high altitude compared with mammals. However, the typically mammalian lungs of bats have equally high mass-related oxygen consumptions during flight, and although there are indications that the

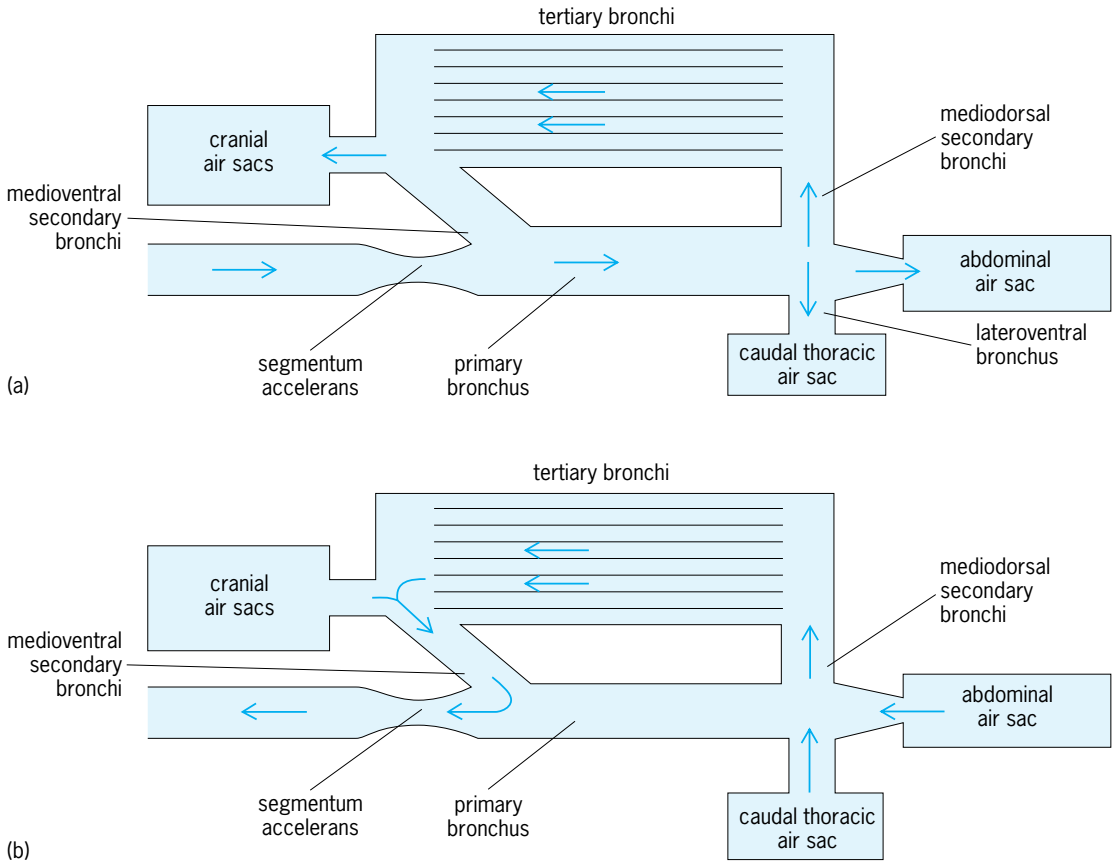

Fig. 28. Pathway of air through the lung of a bird. (*a*) Inspiration. (*b*) Expiration. (*After R. B. Banzett et al., Pressure profiles show features essential to aerodynamic valving in geese, Respir. Physiol., 84:295–309, 1991*)

lungs of birds confer some advantage during hypoxia compared to that of mammals, this decreases at the highest altitudes. P. J. Butler

Mammalian respiration. Ventilation in mammals is more nearly like that in reptiles than in birds. In mammals, alveoli are terminal sacculations in a serial arrangement of ventilatory ducts and sacs, rather than appendant tubules along a parallel duct arrangement as in birds. The only nondiffusion regions of mammalian lungs are the major airways down through the terminal bronchioles; there are no ventilatory sacs in the pulmonary system. However, the alveolar sacs of mammals may be considered the ventilatory analogs of the air sacs of birds, whereas the alveoli themselves incorporate most of the functionally specialized exchange surface. Alveolar sacs histologically are diffusion structures, unlike avian air sacs; however, the size and structural location relative to the alveoli indicate that they and not the alveoli accommodate much of the lung volume change during eupnea. Thus enlargement of the lung during inspiration brings air down to the mouths of alveoli, including a small remainder of expiratory air in the ducts from the preceding breath. This unavoidable dilution of inspired air is like that found in the neopulmo of birds. But unlike the bird, there is no counterpart of a paleopulmo system to provide undiluted atmosphere to any part of the exchange surface. Consequently, the pulmonary system of mammals has H_2O and CO_2 diffusion control like birds, but it lacks the O_2 supply efficiency. The important question of the specific intrapulmonary distribution of air in mammals is under intensive study by histological as well as physiological techniques. A major organ of ventilation found only in mammals is the muscular diaphragm which divides thorax and abdomen. Because it lies as a dome with its convex face toward the thorax, when it contracts and flattens it augments thoracic muscles in increasing the capacity of the thoracic cavity. The abdominal wall relaxes at the same time to accommodate viscera which are displaced by the diaphragm. As a result of these movements, air under pressure differential of a few millimeters of mercury (for example, −1.5 mmHg in the nasal cavity of the horse) passes into the pulmonary system. The abdominal muscles cannot contribute directly to inspiration in mammals or any other vertebrates except turtles, in which a special arrangement is associated with the shell as previously mentioned.

Breathing mechanism. It is in the nature of muscular membrances that they must be oriented as the mammalian diaphragm is; that is, they must insert along their periphery into a resistant structure and bulge into the cavity if contraction is to produce negative pressure in the cavity. A positive expiratory force is a different matter, however, and abdominal muscles contribute such force in all vertebrates. The extent varies with the degree of ventilation,

a much greater contribution in exercise than at rest; with the species; and even with the sex. In the human female during eupneic breathing, for example, costal movement predominates and expiration results largely from the passive recoil of lung and chest, whereas in males there is somewhat more abdominal involvement. In mammals generally, abdominal muscles are more involved in expiration than is true for humans, and in the larger quadrupeds the work of displacing heavy pendant viscera requires their continuous activity. Forced breathing (hyperpnea) with increased amplitude, as during exercise; labored breathing (dyspnea), as during strenuous exercise or at high altitude; and compressatory acts all require the abdominal muscles. These supplement the internal intercostal muscles which pull the ribs to resting position. On the other hand, they are little involved in panting (polypnea) or sniffing.

Myoelastic fibers. Myoelastic fibers of the lung itself not only provide a passive component of expiration, but also account for collapse of the lung if the chest cavity is opened to the atmosphere. Because from the first filling at birth these fibers are stretched during the entire life of the animal, the surface of the lung always tends to recoil from adjacent structures. This recoil is limited in birds because the lung is structurally attached, but in mammals and all other animals the lung surface is free from attachment. The visceral pleura of mammals, a covering membrane, adheres to the lungs, but this is in turn separated from all adjacent structures, which are covered by a parietal pleura, only by a film of mucoid fluid. Measurement of the elastic pulling force exerted by the lung against the surface tension of this fluid reveals the equivalent of a pressure which is negative with respect to atmosphere, when the animal is at rest or during eupnea. It is negative with respect to intrapulmonic pressure at all times. This is called intrapleural pressure (Table 5).

Pressure change and ventilation. The relationships between various pressure changes and ventilation activity in mammalian breathing are illustrated in **Fig. 29**. Here are shown tracings from instruments which recorded abdominal movements in a horse during one breathing cycle, along with concomitant pressure changes in the designated areas. The shapes of such graphs vary in detail among different species of mammals, as can be expected from a comparison of some ventilation parameters for a few animals (Table 5). The forces of respiratory exchange are apparent from such measurements as those in Table 3, made on air and blood entering, within, and leaving the lungs of humans.

Vertebrate respiratory regulation. Ventilation and circulation through the lungs are controlled and adjusted to maintain an efficient exchange of gases in relationship to the needs of the animal and to the suitability of the environment. Sensory units (tension receptors) for pressure or distension occur in lung and breathing elements and in related circulatory organs. These signal the medulla oblongata of the brain, from which the breathing rhythm itself originates, as well as other parts of the brain, and breathing is

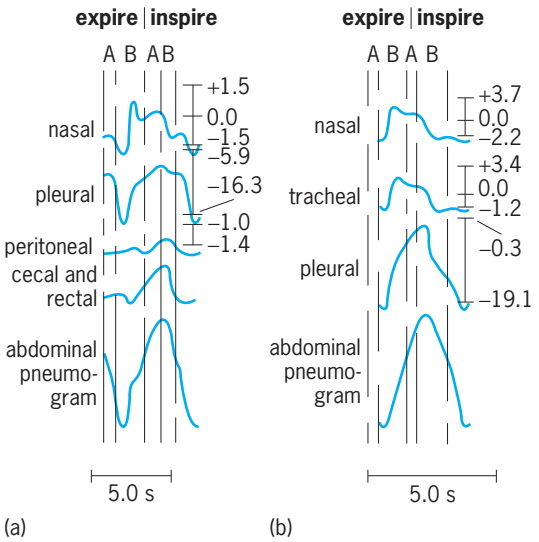

Fig. 29. Composite diagram of the factors in the breathing cycle of the horse, analyzed to show the correlation of intracoelomic and intrapassage pressures with the external musculature. (*a*) Eupnea. (*b*) Hyperpnea. Numbers indicate mmHg.

adjusted to physiological changes and physical conditions. Other receptors, especially in certain major blood vessels near the heart and in the medulla itself, signal the respiratory areas of the brain about blood chemical conditions, in particular, about oxygen, carbon dioxide, and acidity. Signals from other parts of the body, such as the olfactory and tracheal membranes and the joints, and from other parts of the brain, such as heat-regulating, olfactory, and phonation areas, also reach respiratory areas. Here the correlating and integrating role of the brain determines how breathing will proceed: whether to speed up, stop, cough, or to do whatever is appropriate. Oxygen tension in blood is a basic regulatory element for ventilation rate in lower vertebrates, but it becomes secondary to carbon dioxide in birds and mammals. Because carbon dioxide is a major factor in blood acidity, ventilation also comes to play a major role in critical regulation of hydrogen ion concentration (blood acidity) in higher animals. *See* CAROTID BODY.

Respiration in aquatic animals. The critical need for oxygen in amounts available only by continuous air breathing becomes more pronounced the higher an animal is in the evolutionary scale. This reflects the increasing energy requirements of such biological advantages as rapid sustained locomotion and critical continuous neuroregulatory activity. Through such attributes animals gain increasing independence from environmental limitations. Because there is about 30 times less oxygen in an aquatic environment, volume for volume, than in the atmosphere above it, no air-breathing animal except a few fish and most amphibians can maintain vital processes with oxygen gained by aquatic respiration. In amphibians the skin is free from such protective structures as scales, corneum, and hair; thus diffusion is much less restricted than in other animals, and the skin is an important respiratory

organ in water as well as in air. Permeability to respiratory gases includes permeability to water, however, and this restricts even terrestrial amphibians, such as toads, to a very humid atmosphere. Some aquatic amphibians, such as the mud puppy (*Necturus*), possess lungs as well as gills, but no such animals are capable of the high energy output which characterizes each of the strictly terrestrial and air-breathing vertebrate classes: the reptiles, birds, and mammals. Even these classes do include some species which are partly and some strictly aquatic in habitat. All the aquatic species remain strictly air-breathing, but they have modified respiratory and circulatory processes which accommodate diving and underwater activity during long suspension of breathing.

Suspension of breathing. In aquatic reptiles, such as snapping turtles and alligators, breathing suspension for dives is simply a special instance of the apnea which characterizes normal periodic breathing of reptiles on land. In diving birds and mammals, however, the suspension is a departure from the eupneic pattern, for eupnea is normally interrupted only at rest by an occasional deep breath and a pause of about 5 s duration, such as the sigh in humans. A dive of more than 2–3 min duration in birds and mammals requires special adaptations, because no animal has means of storing oxygen to last more than a few minutes; however, seals may submerge for 15–25 min and whales for 1 h or more. Important means of submergence in such animals include (1) restriction of blood flow from the bulk of the muscles, thus sequestering lactic acid and building up a debt for oxygen to be paid on access to air; (2) higher tolerance for carbon dioxide, which is always toxic in excess, and which must accumulate until breathing is resumed; (3) conservation of movement while submerged; (4) reflexes around nostrils which stop breathing on contact with water; and (5) great reduction in lung volume when the high pressure of great depths is involved, which reduces diffusion area of the lung. The last of these does not concern respiration itself; rather it serves to limit nitrogen buildup in tissues, with its attendant hazard of bubble formation, or bends. Such bubbles may form on too rapid ascent from depths, and they can cause immobilization and death. Also, lungs and air sacs may serve as buoyancy organs for swimming and diving, because their contents can be altered in volume to change the displacement of the animal.

Adaptations in terrestrial animals. The adaptations for diving in birds and mammals have been found in counterpart among some nonaquatic animals such as rabbits, which may become immobile and suppress breathing under certain conditions, for example, in evading predators. For active locomotion in air and for highest brain activity, however, oxygen is the most urgently needed of all substances required from the environment; therefore, essentially continuous ventilation is a necessity for strictly terrestrial animals such as the chicken, cat, and human. Such animals survive without actual ventilation only before the time of birth, or hatching, although nonrespiratory periodic breathing movements do occur in egg or uterus as appropriate structures develop. Early respiration utilizes special diffusion respiratory structures, such as the allantois and placenta. In birds an actual pulmonary ventilation, utilizing the air chamber of the egg, begins shortly before hatching. In mammals it can begin only when the fetus gains direct access to air at birth. *See* PLACENTATION; RESPIRATORY SYSTEM DISORDERS. F. Harold McCutcheon

Bibliography. W. Bloom and D. W. Fawcett, *A Textbook of Histology*, 11th ed., 1986; J. N. Cameron, *The Respiratory Physiology of Animals*, 1989; R. M. Cherniack (ed.), *Lung Disease: State of the Art, 1988–1989*, 1990; P. Dejours et al. (eds.), *Comparative Physiology: Life in Water and on Land*, 1987; J. Graham, *Air-Breathing Fishes*, 1997; W. S. Hoar and D. J. Randall (eds.), *Fish Physiology*, vol. 10A and B: *Gills*, 1984; D. J. Randall, The control of circulation in fish during exercise and hypoxia, *J. Exp. Biol.*, 100:275–288, 1982; I. V. Seller (ed.), *Bird Respiration*, vol. 1, 1987; N. B. Slonim, *Respiratory Physiology*, 5th ed., 1987; P. D. Sturkie (ed.), *Avian Physiology*, 4th ed., 1986.

Respiratory system disorders

Dysfunction of the respiratory system, which supplies the body with the oxygen needed for metabolic activities in the cells and removes carbon dioxide, a product of cellular metabolism. For this purpose the right ventricle of the heart pumps venous blood, containing carbon dioxide, through the pulmonary arteries into the lung capillaries, where gaseous exchanges between the blood and air occur. The arterial or oxygenated blood is then drained through the pulmonary veins into the left atrium. *See* RESPIRATION.

The respiratory system includes the nose, mouth, throat, larynx, trachea, bronchi, lungs, and the muscles of respiration such as the intercostal muscles and the diaphragm. The inspiratory movements of the chest wall and diaphragm suck air through the nose, throat, larynx, trachea, and bronchi into the lung alveoli, where it comes into close contact with the blood circulating in the capillaries and gas exchange occurs. During expiration this air is expelled through the same pathway.

Symptoms and signs. The lung has a great reserve capacity, and therefore a significant amount of disease usually must be present to produce clinical signs and symptoms. Shortness of breath (dyspnea) on exertion is the most common symptom of a respiratory disorder. Shortness of breath while at rest is indicative of severe respiratory disease and usually implies a severe abnormality of the lung tissue. If the respiratory system is so diseased that normal oxygenation of the blood cannot occur, blood remains dark, and a bluish color can be seen in the lips or under the fingernails; this condition is referred to as cyanosis. If the process continues and insufficient oxygen is supplied to vital organs such as the brain and heart, a person will become unconscious and die. Other

signs and symptoms of respiratory disorder can include fever, chest pain, coughing, excess sputum production, and hemoptysis (coughing up blood). Most of these signs and symptoms are nonspecific. For example, hemoptysis can result from various conditions, though in smokers it is commonly a sign of lung cancer. *See* HYPOXIA.

Diseases of air passages. Most diseases of the airways increase the resistance against which air is sucked in and pushed out of the lungs. The effect of a disease of the airways depends largely on its localization. Diseases of the nose usually have little influence since collateral respiration through the mouth compensates easily. Diseases of the throat, larynx, and trachea can significantly inhibit the flow of air into the lungs. Infections in the back of the throat, such as in diphtheria, can cause marked swelling of mucous membranes, resulting in air obstruction. Edema (swelling) of the mucosal lining of the larynx can also cause a reduction in air flow. Likewise, air flow can be inhibited in asthma, in which the smooth muscle in the trachea and bronchi episodically constricts. Chronic bronchitis, which is caused predominantly by cigarette smoking, results in inflammation of and excess mucus production by the bronchi and this also can lead to a reduction in air flow. Bronchiolitis, a condition that usually occurs in children and is often caused by a respiratory virus, results in narrowing and inflammation of small airways and a decrease in air flow.

Diseases of the lung. Pneumonia, cancer, and emphysema are the most common lung diseases and are a major cause of morbidity and mortality in the United States. Pneumonia is caused by bacteria, viruses, and other infectious agents, that result in inflammation and consolidation (solidification) of the gas exchange units of the lung (alveoli). Persons with pneumonia frequently develop fever, chest pain, and rapid breathing. When the disease involves a significant portion of the lung, shortness of breath may develop. *See* PNEUMONIA.

There are four main types of pulmonary emphysema. The most common is centrilobular emphysema, and it is caused by cigarette smoke. This disease usually involves the upper lobes of the lungs and is characterized by the breakdown and permanent loss of lung tissue, resulting in large cystic spaces in the lung referred to as blebs. Emphysema is usually associated with chronic bronchitis, and individuals with the combination of diseases will often have shortness of breath on exertion and usually have excess sputum production. *See* EMPHYSEMA.

Of the four major types of lung cancer, approximately 90% can be attributed to the carcinogens present in cigarette smoke. Lung cancer may be detected in asymptomatic persons with a routine chest x-ray, or it may be discovered because of pain, excess coughing, or hemoptysis. *See* CANCER (MEDICINE).

Pulmonary circulatory diseases. Among the diseases of pulmonary circulation, congenital malformations of the heart and pulmonary artery account for many cases of respiratory insufficiency in newborn and younger children; those affected are called blue babies because they appear cyanotic. In adults, acquired heart diseases, such as narrowing of the valve between left atrium and left ventricle (mitral stenosis), can cause backup of blood into the lung and an increased pressure in the pulmonary circulatory system. Also, blood clots, which usually develop first in the deep veins of the legs, can break free and flow to the heart. There they enter the pulmonary arteries and wedge in their small branches, where the clots are referred to as pulmonary thromboemboli, and can cause areas of death in the lung tissue (pulmonary infarcts). Persons who develop thromboemboli usually have chest pain and shortness of breath, and some have hemoptysis. *See* CARDIOVASCULAR SYSTEM; CIRCULATION DISORDERS.

Neuromuscular diseases. Some neuromuscular diseases, such as poliomyelitis and amyotrophic lateral sclerosis (Lou Gehrig's disease), can cause dysfunction of the muscles of respiration. The resulting inability of these muscles to move air into the lungs can cause severe shortness of breath and predispose the patient to pneumonia. Artificially assisted ventilation may be necessary for breathing. *See* MUSCULAR SYSTEM DISORDERS; POLIOMYELITIS; RESPIRATORY SYSTEM.

Samuel P. Hammar

Bibliography. D. H. Dail and S. P. Hammar (eds.), *Pulmonary Pathology*, 2d ed., 1994; A. P. Fishman, *Pulmonary Diseases and Disorders*, 3d ed., 1998.

Response

The reaction or result produced by applying an input signal to a system. Given a system, or interconnection of devices to perform a desired function, the signals processed by that system can be divided into several categories, some of which are referred to as causes, stimuli, or input signals; and others, as responses, effects, or output signals. Although there can be multiple input and output signals for a given system, attention will be focused on single-input single-output systems. Often, the input and output signals of interest are functions of time only, although the discussions given can be generalized to more than one independent variable, such as space and time. Also, attention will be limited to continuous-time signals and systems, although the discussion can be easily put in terms of discrete-time signals and systems.

Linear, time-invariant systems. It is convenient to represent the relationship between stimulus and response of a system as Eq. (1), which is read "$y(t)$ is

$$y(t) = \mathscr{H}[x(t)] \tag{1}$$

the response of $\mathscr{H}$ to the input $x(t)$," where initial conditions are included if pertinent. The symbol $\mathscr{H}$ serves the dual role of identifying the system and specifying the operation to be performed on $x(t)$ to produce $y(t)$.

It is difficult to say anything explicit about the response of a general system, as described by Eq. (1), and usually attention is limited to linear, time-invariant systems. A system is linear if an arbitrary

linear combination of two inputs produces the same linear combination of the outputs to each input applied separately. A system is time-invariant if the response to an input delayed by a constant amount is the response to the undelayed output delayed by the same amount; that is, if $x(t)$ in Eq. (1) is replaced by $x(t - \tau)$, where τ is a constant, the output is $y(t - \tau)$. *See* LINEAR SYSTEM ANALYSIS; LINEARITY.

Complete, free, and forced responses. Many linear, time-invariant systems can be described by constant-coefficient, linear differential equations. When a system is described in this way, the derivative of highest degree on $y(t)$ in the differential equation is called the order of the system. It is well known from the theory of ordinary differential equations that the solution to a constant-coefficient, linear differential equation consists of two parts: the complementary and particular solutions. The complementary solution is found by setting $x(t) = 0$ and solving for the roots of the characteristic equation (that is, the equation which results by assuming that $y(t) = e^{pt}$, and canceling the common factor e^{pt} after differentiation and substitution). The complementary solution, also known as the homogeneous solution, is then a sum of n exponential functions, if the n roots to the characteristic equation are distinct. Since this solution results by setting the input to zero, it represents the natural behavior of the system in response to an initial stored energy when the input is removed. Thus the complementary solution is also called the free response of the system. The other part of the solution, the particular integral, is of the same form as the input and can be determined by the method of undetermined coefficients. Since it depends on the form of the forcing function, or $x(t)$, the particular solution is also called the forced response of the system. The complete response of the system is the sum of the homogeneous solution and particular integral of the describing differential equation of the system, or the sum of the free and forced responses. *See* DIFFERENTIAL EQUATION.

Gain and phase response. It is useful to consider the forced response of a system described by a constant-coefficient, linear differential equation to an input of the form $e^{j\omega t} = \cos \omega t + j \sin \omega t$, where j is the square root of -1, and ω is a parameter, called the radian frequency of the input, which is to be varied later. The forced response must be of the same form as the input, so it is assumed that $y(t) = Ae^{j(\omega t+\theta)}$. After differentiation of both the input and assumed response, substitution, and cancellation of the factor $e^{j\omega t}$ on both sides of the equation, Eq. (2) is obtained,

$$\left.\frac{y(t)}{x(t)}\right|_{x(t)=e^{j\omega t}} = H(j\omega) \tag{2}$$

where $H(j\omega)$ is a rational function of $j\omega$. The function $H(j\omega)$, called the transfer function or frequency response, of the system, is a complex function of the variable ω, which is the radian frequency of the sinusoidal input. (Although a complex exponential was assumed, the response to $\cos \omega t$ of the system can be obtained by multiplying $H(j\omega)$ by $e^{j\omega t}$ and taking the real part.) Writing the complex function $H(j\omega)$ in terms of its magnitude and complex argument yields Eq. (3), where $A(j\omega)$ is called the amplitude response, or gain, and $\theta(j\omega)$ is called the phase response. *See* GAIN; PHASE (PERIODIC PHENOMENA).

$$H(j\omega) = A(j\omega)e^{j\theta(j\omega)} \tag{3}$$

These are important functions for a number of reasons. The amplitude response shows how the sinusoidal frequencies that are used to resolve an input (that is, its spectral components) are changed in amplitude, or filtered, in going through a system. Filter theory deals with how to synthesize transfer functions that give certain desired filter characteristics such as low-pass, band-pass, or high-pass. The phase response can be used to obtain the group delay of the system (group delay is the negative derivative of the phase response), which can be used to determine how the various spectral components of the input are delayed with respect to one another. For no distortion of the input, the system should impose the same gain and delay on all spectral components that compose the input. In the field of control theory, the frequency response of the open loop of a feedback control system plays an important role in determining system stability. It can be shown that oscillations within the closed loop are possible if the open-loop phase response is an odd multiple of 180° at a frequency or frequencies and the closed-loop gain is unity at any of these frequencies. The amounts by which phase and gain of the open-loop system do not meet these criteria are called the phase and gain margins, respectively. *See* CONTROL SYSTEM STABILITY; CONTROL SYSTEMS; ELECTRIC FILTER.

Superposition integrals. The impulse response is of special interest in the description of a linear time-invariant system. For such a system, it is the response to a unit impulse function applied as the input at time zero. The impulse response provides a means to obtain an explicit expression for the response in terms of the input via the superposition integrals, as opposed to the implicit relationship given in a differential equation.

Rodger E. Ziemer

Bibliography. C. M. Close and D. K. Frederick, *Modeling and Analysis of Dynamic Systems*, 3d ed., 2001; E. Kreyszig, *Advanced Engineering Mathematics*, 8th ed., 1998; R. E. Ziemer, W. H. Tranter, and D. R. Fannin, *Signals and Systems: Continuous and Discrete*, 4th ed., 1998.

Rest mass

A constant intrinsic to a body which determines its inertial and energy-momentum properties. It is a fundamental concept of special relativity, and in particular it determines the internal energy content of a body. It is the same as the inertial mass of classical mechanics. According to the principle of equivalence, the basic physical principle of general relativity, the inertial mass of a body is also equal to its gravitational

mass. *See* CLASSICAL MECHANICS; GRAVITATION; RELATIVISTIC MECHANICS; RELATIVITY.

Newton's law. The rest mass or inertial mass of a body, m, is a measure of its resistance to being accelerated at $\mathbf{a}$ by a force $\mathbf{F}$; in classical mechanics the relation between inertial mass, acceleration, and force is given by Newton's law, Eq. (1). In special rel-

$$\mathbf{F} = m\mathbf{a} \tag{1}$$

ativity Newton's law holds exactly only in the body's rest frame, that is, the frame in which the body is instantaneously at rest. *See* NEWTON'S LAWS OF MOTION.

Energy-momentum. Rest mass serves to define the energy-momentum of a body. In special relativity the trajectory of a body is specified by the position 4-vector, given in Eq. (2), where t denotes time, c

$$(x^0, x^1, x^2, x^3) = (ct, x, y, z) \tag{2}$$

is the speed of lights, and (x,y,z) is the body's position in cartesian coordinates. The 4-velocity is the derivative of the position 4-vector with respect to the proper time, that is, the time measured in the body's rest frame. A body's energy-momentum 4-vector is defined as the product of its 4-velocity and its rest mass. *See* VECTOR METHODS (PHYSICS).

This 4-vector is the relativistic generalization of both the energy and vector momentum of classical mechanics, with the zeroth component, p^0, being the energy divided by c, and the components 1 to 3 being the momentum vector. Thereby, the separate conservation laws for energy and momentum in classical mechanics are combined into one conservation law for the energy-momentum 4-vector in special relativity. *See* CONSERVATION LAWS (PHYSICS); CONSERVATION OF ENERGY; CONSERVATION OF MOMENTUM; ENERGY; MOMENTUM.

Rest energy. Associated with the rest mass of a body, there is an internal or rest energy, which follows from the above definition. In the system where the body is at rest, the zeroth component of the velocity 4-vector is equal to c since the proper time and the laboratory time, t, are then the same. The energy of the body is given by Eq. (3).

$$p^0 c = E = mc^2 \qquad \text{(body at rest)} \tag{3}$$

A more physical explanation of this mass-energy equivalence based on the Doppler-shift relation of special relativity, follows that of A. Einstein. If a slowly moving body with velocity v emits one light pulse in the forward direction and one in the backward direction, each of energy $\Delta E/2$ as seen in its rest frame, then by symmetry the body does not change velocity, but it does lose energy, ΔE, in its rest frame. However, as seen in the laboratory frame, the light pulses will be Doppler-shifted and will carry a larger total energy, $\Delta E(1 - v^2/c^2)^{-1/2}$, so that the body's energy is there seen to decrease by this larger amount. A difference between the energies of the body in the rest and laboratory frames must correspond to a difference in kinetic energy, $mv^2/2$. Since the body's velocity does not change as the light pulses are emitted, this decrease in kinetic energy must reflect a decrease in the body's rest mass, Δm. Comparison of kinetic energy loss, $\Delta mv^2/2$, with the difference between the energy loss in the two frames, $\Delta E[1 - (1 - v^2/c^2)^{-1/2}]$, yields Eq. (4), in the limit of low velocity, in accord with Eq. (3).

$$\Delta m = \frac{\Delta E}{c^2} \tag{4}$$

Reactions. The experimental realization of the interconversion of mass and energy is accomplished in the reactions of nuclei and elementary particles. In particular, the energy source of nuclear bombs and nuclear fission reactors is a small decrease in the total mass of the interacting nuclei, which gives rise to a large energy release because of the large numerical value of c^2. *See* ELEMENTARY PARTICLE; NUCLEAR FISSION.

Ronald J. Adler

Bibliography. H. A. Lorentz et al., *The Principle of Relativity*, 1923; H. M. Schwartz, *Introduction to Special Relativity,* 1968, reprint 1977.

Restionales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Commelinidae of the class Liliopsida (monocotyledons). The order consists of 4 families and about 450 species, some 400 of them belonging to the Restionaceae. The vast majority of the species grow in temperate regions of the Southern Hemisphere. The Restionales are wind- or self-pollinated, with reduced flowers and a single, pendulous, orthotropous ovule in each of the 1–3 locules of the ovary. The Restionaceae are much like the Gramineae in the related order Cyperales and are sometimes referred to as Southern Hemisphere grasses. *See* COMMELINIDAE; CYPERALES; LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Restoration ecology

A field in the science of conservation that is concerned with the application of ecological principles to restoring degraded, derelict, or fragmented ecosystems. The primary goal of restoration ecology (also known as ecological restoration) is to return a community or ecosystem to a condition similar in ecological structure, function, or both, to that existing prior to site disturbance or degradation. *See* ECOLOGY.

Most restorationists trace their discipline to Aldo Leopold, a professor at the University of Wisconsin who began the first restoration project in the United States in 1935. Working with the Civilian Conservation Corps, a federal relief work program during the Dust Bowl era, Leopold replanted 24 hectares (60 acres) of North American tallgrass prairie, a formerly widespread ecosystem that to this day remains highly fragmented and endangered. By working to

restore this species-rich grassland, Leopold established a national conservation ethic that still serves as the basis of this ecological discipline. Restoration ecology today, however, involves more than replanting native species. It is a highly synthetic discipline that draws upon older ecological fields such as applied ecology and wildlife management, and newer ones such as invasion biology and landscape ecology. *See* CONSERVATION OF RESOURCES; ECOLOGY, APPLIED; ECOSYSTEM.

Ecologists engage in restoration activities for different reasons. For example, wetlands are restored as a consequence of legal mandates, such as the federal Clean Water Act. This legislation requires the restoration or creation of wetlands to mitigate the unavoidable impacts of human activities such as commercial construction or habitat conversion. Other ecologists are involved in restoration-related activities because they recognize that conservationists cannot afford to be concerned only with the preservation of pristine areas. Many ecosystems have nearly disappeared altogether, and what remains is highly fragmented or degraded. The only alternative in such cases is to restore the fragmented patches that remain. Some ecologists restore specific segments of an ecosystem (for example, a rare prairie orchid or a top carnivore) to reestablish the processes that control the structure and function of the food web. *See* FOOD WEB.

Restoration projects. A reference framework is needed to guide any restoration attempt—that is, to form the basis of the design (for example, desired species composition and density) and monitoring plan (for example, setting restoration milestones and success criteria for restoration projects). Such a reference system is derived from ecological data collected from a suite of similar ecosystems in similar geomorphic settings within an appropriate biogeographic region. Typically, many sites representing a range of conditions (for example, pristine to highly degraded) are sampled, and statistical analyses of these data reveal what is possible given the initial conditions at the restoration site. Two examples of restoration deserve mention, because they represent the range of activities in which restoration ecologists are engaged: Guanacaste National Park and San Diego Bay.

Guanacaste National Park. Perhaps the most well known (and most ambitious) restoration project of the 1990s was the restoration of thousands of hectares of tropical dry forest in Guanacaste National Park in Costa Rica. The original forest cover of Guanacaste was composed of canopy trees 20–30 m in height, with many understory shrubs and vines. However, much of the dry forest ecosystem had been converted in the 1800s to pasturelands of exotic African grasses. These grasses appear to be maintained by fire, so the cessation of annual burning was key to the restoration process. Costa Rican biologists and United States scientists worked together using ecological principles and standard land management techniques to keep fire out and increase seed dispersal. Horses were used to disperse one of the dominant tree species, the Guanacaste tree (*Enterolobium cyclocarpum*, Fabaceae), by allowing them to wander through the restoration site and deposit in manure piles the seeds they had eaten. As of 2001, a significant portion of Guanacaste National Park consisted of a secondary forest that formerly was only pastureland. Scientists believe that this restoration project continues to be successful because some dry forest patches still exist that serve as sources of restoration propagules (units of dispersal), and because the long time frame (decades) is appropriate for this kind of long-term restoration endeavor. *See* FOREST AND FORESTRY; FOREST ECOSYSTEM; REFORESTATION; LAND RECLAMATION.

San Diego Bay. Since 1984, the U.S. Fish and Wildlife Service has put considerable effort into the restoration of the wetlands in San Diego Bay, California, to provide a habitat for the light-footed clapper rail (*Rallus longirostris levipes*), a bird on the Endangered Species List. Two wetland ecosystems totaling 29 acres were constructed for the rails, but as of early 2001 rails had not nested at either site. The scientists involved in this restoration project believe that the constructed wetlands were perceptively different from natural systems in such ecological characteristics as plant height, abundance of epibenthic (living on the soil surface at the bottom of a body of water) invertebrates, and concentrations of organic matter and nitrogen in the soil. Lessons learned from this project include the sobering realization that restorationists cannot guarantee long-term success. Ecologists may not know of, much less understand, all of the ecological linkages in the ecosystem being restored. This lack of understanding can lead to the creation of a different, but not necessarily restored, ecosystem. *See* ENDANGERED SPECIES; WETLANDS.

Challenges. Restoration ecology faces both practical and scientific challenges. Practical issues involve the costs associated with restoration projects, as well as the scale. Typically restoration is expensive, and land acquisition, nursery stock, soil preparation, grading, irrigation, monitoring, maintenance, and labor can contribute to a total construction cost that is prohibitive. This is especially relevant in developing countries, where national budgets for conservation-related activities are often minimal. In terms of scale, most restoration projects involve the restoration of less than 10 acres. A 50-acre project is considered quite large, and there are fewer than 50 restoration projects in the United States that involve more than 100 acres. The ability of restoration ecologists to effectively restore ecosystems on a long-term basis is unknown; therefore concerns about the effectiveness of restoration efforts are justified. Restoration ecologists are turning to creative ways to cut costs and expand their efforts in a land-based metric. Utilization of the large volunteer base of nonprofit conservation organizations such as The Nature Conservancy is one way that labor costs have been diminished.

Scientifically, restoration ecologists face equally challenging issues. Perhaps the most important guiding ecological principle applied to restoration

ecology is that of plant succession. Many restoration projects must accelerate the species replacement process or simply "jump-start" communities without allowing time for natural succession to occur. Ecologists are now turning to experimental approaches to determine which species should be planted when, and how best to assemble a community in a restored environment. *See* ECOLOGICAL SUCCESSION.

Peggy L. Fiedler

Bibliography. A. D. Bradshaw, Ecological principles and land reclamation practice, *Landsc. Plan.*, 11:35–48, 1984; J. Cairns, Jr., *Rehabilitating Damaged Ecosystems*, 2d ed., CRC Press, Boca Raton, 1995; R. J. Hobbs and D. A. Norton, Towards a conceptual framework for restoration ecology, *Restor. Ecol.*, 4:93–110, 1996; W. R. Jordan III, M. E. Gilpin, and J. D. Aber (eds.), *Restoration Ecology: A Synthetic Approach to Ecological Restoration*, Cambridge University Press, 1987.

Restriction enzyme

An enzyme, specifically an endodeoxyribonuclease, that recognizes a short specific sequence within a deoxyribonucleic acid (DNA) molecule and then catalyzes double-strand cleavage of that molecule. Restriction enzymes have been found in bacteria, where they serve to protect against the deleterious effects of foreign DNA, and in viruses that infect *Chlorella*-like organisms, the only eukaryotes from which those enzymes have yet been isolated. *See* DEOXYRIBONUCLEIC ACID (DNA).

Types. There are three known types of restriction enzymes.

Type I enzymes recognize a specific sequence on DNA but cleave the DNA chain at random locations with respect to this sequence. They have an absolute requirement for the cofactors adenosine triphosphate (ATP) and *S*-adenosylmethionine, and during cleavage they hydrolyze massive amounts of ATP. Because of the random nature of the cleavage, the products are a heterogeneous array of DNA fragments. *See* ADENOSINE TRIPHOSPHATE (ATP).

Type II enzymes also recognize a specific nucleotide sequence but differ from the type I enzymes in that they do not require cofactors and they cleave specifically within or close to the recognition sequence, thus generating a specific set of fragments. It is this exquisite specificity which has made these enzymes of great importance in DNA research, especially in the production of recombinant DNAs. More than 1200 such enzymes have been characterized, and almost 150 different specific sequences are recognized. In those cases where enzymes from different sources recognize the same sequence, the enzymes are called isoschizomers. The sequences that are recognized range from four to eight nucleotides in length and often contain one or two positions at which several sequence combinations can be present. Many type II enzymes recognize sequences that contain a dyad axis of symmetry (commonly called palindromes) and cleave to leave fragments with cohesive ends.

Type III enzymes have properties intermediate between those of the type I and type II enzymes. They recognize a specific sequence and cleave specifically a short distance away from the recognition sequence. However, it is often difficult to obtain complete digestion. They have an absolute requirement for the cofactor ATP, but they do not hydrolyze it. Their activity is stimulated by *S*-adenosylmethionine, although it is not an absolute requirement.

Biological importance. Given the capacity of a restriction enzyme to cleave DNA, it is interesting that the bacterium containing such an enzyme does not kill itself as a result of the restriction enzyme cutting the bacterial DNA. The reason is found by considering the circumstances under which the restriction enzyme is present in the bacterium, and relates to the biological phenomenon of restriction modification.

A restriction modification system contains two components, the restriction enzyme and the modification enzyme. The modification enzyme recognizes exactly the same sequence as is recognized by the restriction enzyme and then specifically modifies that sequence so that it can no longer be cleaved by the restriction enzyme. In all known cases, this modification is achieved by methylation. In the bacterium containing the restriction enzyme, all of the recognition sequences on the bacterial DNA are modified by the companion modification enzyme, and hence the bacterium is immune to the effects of its own restriction enzyme. However, if DNA from some foreign source enters the bacterium, as might occur during bacteriophage infection, then it will not usually be premodified in the specific manner required to protect its recognition sequences from the restriction enzyme, and so it will be cleaved and effectively destroyed. In this way the restriction modification system serves to protect the bacterium from the deleterious effects of incoming foreign DNA.

Use in research. Restriction enzymes are important to research in molecular biology. Their discovery was honored by the award of the 1978 Nobel Prize in Physiology or Medicine to Werner Arber, Hamilton O. Smith, and Daniel Nathans. Arber was the first to postulate the existence of these enzymes, and subsequently isolated the first I enzyme. Smith was responsible for the discovery and characterization of the first II enzyme, *Hin*dII from *Haemophilus influenzae*. Nathans was the first to use that enzyme to cleave SV40 DNA into discrete fragments and then to use the fragments generated to study the molecular biology of the monkey virus SV40.

Naturally occurring DNA molecules are usually rather large and therefore inaccessible to direct biochemical experimentation. For instance, one of the smallest known DNA molecules is the chromosome of SV40, which contains 5243 nucleotide pairs; a typical human chromosome may contain 10^8 nucleotide pairs. The II restriction enzymes allow molecular biologists to cut these large DNA molecules into a series of smaller specific fragments, each of which

contains a unique section of the larger molecule. For example, the enzyme *Hae*III, which recognizes the sequence GGCC, cuts the SV40 chromosome into 19 specific fragments. These fragments can be separated by size by using gel electrophoresis and then can be purified individually. The relative order of these fragments on the genome can be determined, so that a map can be produced that shows the location of each restriction enzyme site. The construction of such a physical map for one or more enzymes is often the first step in a more detailed study of a DNA molecule. Restriction enzyme cleavage sites provide a convenient set of markers along the chromosome by which genes and their regulatory signals may be ordered and their positions assigned. The aim of this kind of approach is a direct correlation of physical and genetic maps, ultimately taken to the nucleotide sequence level. In more complex cases where very large numbers of fragments are present, they can still be separated effectively by gel electrophoresis, and individual fragments can be detected by hybridization using radioactivity labeled probes. *See* GENETIC MAPPING.

Genetic engineering. A key feature of the fragments produced by restriction enzymes is that when mixed in the presence of the enzyme DNA ligase, the fragments can be rejoined. For instance, when the bacterial plasmid pBR322 is cleaved with the restriction enzyme *Eco*RI, for which there is a single site, the original circular DNA molecule is converted to a linear form. Upon addition of DNA ligase, the ends of the linear molecule are joined to reform a circular pBR322 molecule that is indistinguishable, chemically and biologically, from the starting material. This rejoining is very efficient when the restriction enzyme used produces fragments with cohesive ends. If two different linear fragments produced by cleavage with *Eco*RI are mixed in the presence of DNA ligase, then in addition to producing circular versions of the two original DNA molecules there will result recombinant molecules in which the two linear molecules have been joined together. If one of the molecules is the plasmid pBR322, which contains the genetic information necessary for replication, then the new recombinant molecule, containing the additional fragment, will be able to replicate when placed in a bacterial cell. Thus, by simple laboratory manipulation, a new molecule has been created that comprises the genetic information of the plasmid pBR322 together with new genetic information present on the additional fragment. If, instead of a single *Eco*RI fragment, a mixture of 10 fragments was present plus the linear pBR322, then 10 different recombinant plasmids would result from the experiment, each deriving its replicative ability from the pBR322 sequences. By using standard techniques of bacterial genetics, colonies of bacterial cells containing individual recombinant plasmids can be easily isolated. Indeed, extremely complex mixtures of restriction fragments can be resolved into individual components, each of which is in a readily accessible form by this same technique.

Should the new fragment carry genetic information that can be interpreted by the bacterial cell containing the recombinant molecule, then the information will be expressed as a protein, and the bacterial cell will serve as an ideal source from which to obtain that protein. For instance, if the DNA fragment carries the genetic information encoding the hormone insulin, the bacterial cell carrying that fragment will produce insulin. Because many plasmids are present in high copy number in bacterial cells, the amount of the protein produced can be extremely high. By using this method, the human gene for insulin has been cloned into bacterial cells and used for the commercial production of human insulin. The potential impact of this technology forms the basis of the genetic engineering industry. *See* ENZYME; GENETIC ENGINEERING.

Richard Roberts

Bibliography. F. C. Neidhardt et al. (eds.), *Escherichia coli and Salmonella typhimurium*: *Cellular and Molecular Biology*, 1987; R. J. Roberts, S. M. Linn, and R. S. Lloyd (eds.), *Nucleases, Monograph 25*, 2d ed., 1994; G. G. Wilson, Type II restriction-modification systems, *Trends Genet.*, 4:314–318, 1988.

Resultant of forces

A system of at most a single force and a single couple whose external effects on a rigid body are identical with the effects of the several actual forces that act on the body. For analytic purposes, forces are grouped and replaced by their resultant. Forces can be added graphically (**Fig. 1**) or analytically. The sum of more than two vector forces can be found by extending the method of Fig. 1*c* to a three-dimensional vector polygon in which one force is drawn from the tip of the previous one until all are laid out. The resultant force is the force vector required to close the polygon directed from the tail of the first force vector to the tip of the last. A force system has a zero force resultant if its vector polygon closes. *See* CALCULUS OF VECTORS.

The resultant of force along an axis may be desired. In that case, all forces are resolved into their

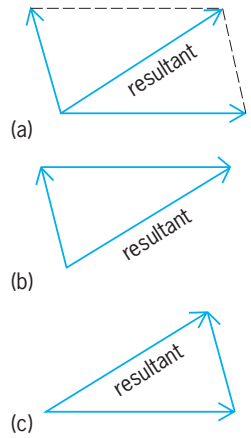

Fig. 1. Resultant of two forces acting through common center. (*a*) Diagonal of parallelogram. (*b*, *c*) Hypotenuse of triangle.

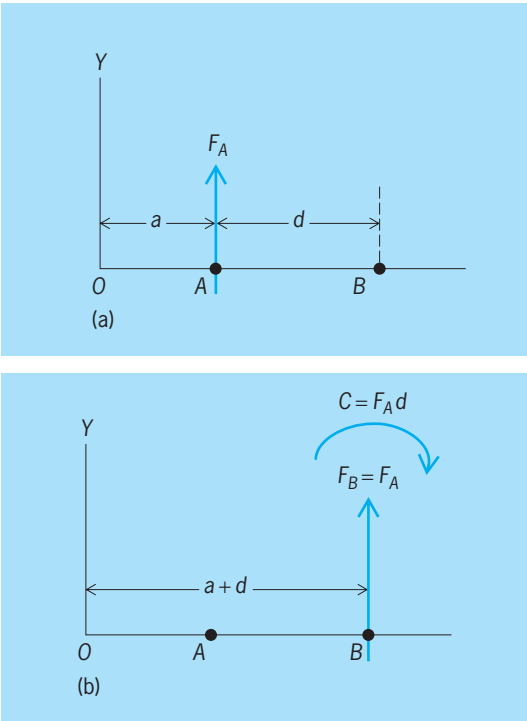

Fig. 2. Resolution of a force. (*a*) Force F_A. (*b*) Resolved force F_B and couple *C*.

respective components along the axis, and the components are added algebraically to obtain the desired resultant.

The resolution of a force into components is an inverse operation to the composition of multiple forces into a resultant. However, whereas the resultant force is unique to a given force system, a force can be resolved into any variety of components.

Two force systems are equivalent if their resultant forces, as described above, are equal and if their total vector moments about the same point are also equal. Vector moments are combined in the same manner as forces, that is, by parallelograms, triangles, or polygons.

In **Fig. 2** the force F_A acting at point A may be resolved into an equivalent system consisting of equal parallel F_B acting at point B and couple C; the magnitude of C is the moment of F_A about B. Also, the total moment of F_B and C about A is the moment of F_A about A, namely zero.

A resultant is the equivalent force system having the fewest possible forces and couples. The resultant of concurrent forces is a force equal to their vector sum and acting through the point of concurrence. A special case is a collinear system in which the resultant is collinear with the forces of the system as well as equal to their vector sum.

When all forces of a system are coplanar, the resultant may be a force or a couple. If it is a force, the resultant is positioned to produce the same moment about a reference point as the system. Should the vector sum of forces be zero, the resultant is a couple that develops the same moment as the system. *See* COUPLE.

In a three-dimensional force system, the resultant consists of a force element passing through an arbitrary point and equal to the total force of the system and a couple element that produces a moment equal to the total moment of the system about any point on the line of action of the force element. *See* FORCE; STATICS.

Nelson S. Fisk

Retaining wall

A generic structure that is employed to restrain a vertical-faced or near-vertical-faced mass of earth. The earth behind the wall may be either the natural embankment or the backfill material placed adjacent to the retaining wall. Retaining walls must resist the lateral pressure of the earth, which tends to cause the structure to slide or overturn.

Types. There are several types of retaining walls. A gravity wall is typically made of concrete and relies on its weight for stability (**illus.** *a*). The mass of the structure must be sufficient to develop enough frictional resistance to sliding, and the base or footing of the structure must be wide enough to develop sufficient moment to resist overturning earth forces. The base width of the gravity wall must also be sufficient that the stresses on the soil below the base are not greater than the bearing capacity of this soil. Resistance to sliding may be increased by including a key that projects down into the soil.

A cantilever retaining wall (illus. *b*) gains a larger effective mass by virtue of the soil placed on the horizontal cantilevered section of the wall. Reinforced counterforts are spaced along the wall to increase its strength. A key may be included to resist sliding. *See* CANTILEVER.

A variation of the gravity retaining wall is the crib wall (illus. *c*) is usually constructed of prefabricated interlocking concrete units. The crib is then filled with soil before the backfill adjacent to the crib is placed. Crib retaining walls require less concrete than gravity structures, and the individual units can be prefabricated for quicker construction.

Bulkhead retaining walls (illus. *d*) consist of vertical sheet piling that extends down into the soil and is stabilized by one or more tiebacks and anchors periodically spaced along the structure. The sheet piling may be made of reinforced concrete, steel, or aluminum.

A novel method for stabilizing excavations and freshly cut slopes in place is soil nailing. As excavation proceeds, the segmented face of the retaining wall is secured by many long steel bars (nails) that are driven into the embankment or placed in predrilled holes and grouted. This type of retaining wall is popular because of the economy of time and materials used in construction.

Marine installations. Retaining walls are often used in the marine environment, where they separate the retained soil from the water. Gravity walls (known as seawalls) can be constructed where strong wave and current forces are exerted on the wall. Bulkheads

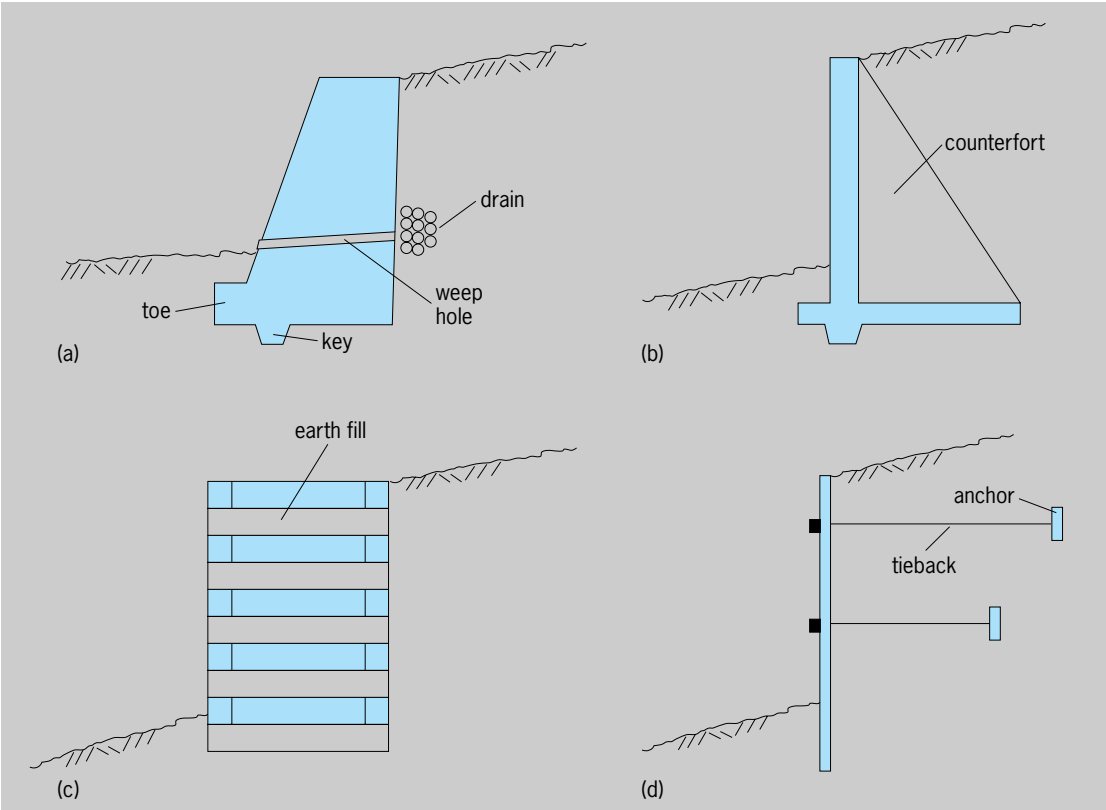

Common types of retaining walls. (*a*) Gravity wall. (*b*) Cantilever wall. (*c*) Crib wall. (*d*) Bulkhead.

are more commonly found in sheltered areas such as harbors and navigation channels. Where there is strong wave or current action, there is a potential for scour at the toe of the structure and eventual undermining of the retaining wall. To counter scour, a protective blanket of stone is placed at the structure toe. Also, the structure may be extended more deeply into the earth to allow for toe scour without causing the structure to be undermined.

Stability. The retaining wall must not fail by sliding or rotating under the combined action of all of the forces acting on the wall. These forces primarily include the vertical gravitational force acting through the center of mass of the structure, the lateral earth pressure on the wall including any dead loads (for example, buildings) or live loads (such as roadway traffic) that act on the embankment and transmit to the wall, and the frictional and bearing forces of the foundation soil.

The embankment earth pressure may be increased because of hydrostatic pressures caused by water in the soil if the embankment is not adequately drained. Water in the embankment may also cause swelling if the soil is cohesive, and ice thrust when the water freezes and thaws repeatedly. If the foundation soil is saturated, there will be a hydrostatic uplift force on the structure base.

Soil pressures at the toe of the wall will generally increase structure stability by resisting sliding of the structure. These forces should be included in the stability analysis only if the designer can be certain that this soil will remain at the toe. Lateral earth pressures may be increased by earthquake loadings or other vibratory loading from machines or traffic.

Drainage. The embankment behind and under a retaining wall should be well drained. This can be accomplished by paving the top surface of the embankment and installing a gutter at the toe of the structure to carry water runoff away from the structure. A drain along the lower section of the embankment and weep holes through the structure (illus. *a*) will prevent water from accumulating in the embankment. If fill is placed in the embankment next to the structure, it should be of highly permeable material. *See* SOIL MECHANICS. R. M. Sorensen

Bibliography. R. Brown, *Practical Foundation Engineering Handbook*, 2d ed., 2000; D. D. Driscoll, *Retaining Wall Design Guide*, Forest Service, U.S. Department of Agriculture, 1979; P. C. Lambe and L. A. Hansen, *Design and Performance of Earth Retaining Structures*, Geotech. Spec. Publ. 25, American Society of Civil Engineers, 1990; G. P. Tschebotarioff, *Foundations, Retaining and Earth Structures*, McGraw-Hill, 1978.

Reticular formation

Characteristic clusters of nerve cell bodies (gray matter) and their meshwork, or reticulum, of fibers which are found in the brainstem and the diencephalon. The reticular formation is thought to be a complex, highly integrated mechanism which exerts both inhibition and facilitation on almost

every type of activity of the central nervous system.

Anatomy. The brainstem, composed of the medulla oblongata, the pons, and the midbrain or mesencephalon, is the basic integrating and connecting unit of the central nervous system. The long ascending sensory fibers pass upward from all parts of the body and are distributed largely in the brainstem to the cerebellum, cerebral cortex, and other related higher brain centers. Similarly, the descending, or motor, fibers pass downward from higher centers through the stem to be distributed to appropriate lower levels, such as the spinal cord. In addition, the brainstem is itself the site of many important structures, notably the cell body clusters, or nuclei, of cranial nerves. *See* BRAIN; NERVOUS SYSTEM (VERTEBRATE).

The reticular formation lies in and around these better-defined structures, which break the reticular formation into many small islands of gray matter connected by large numbers of relatively short nerve fibers. These fibers pass in an apparently haphazard manner in all directions; but investigation has revealed that this reticulum is in reality a highly complex, intricately organized, master system of communication which alters many body activities.

Many authorities disagree on the exact anatomic or physiologic extent of the reticular formation because of differences in study methods in the several disciplines. Although certain reticular formation cell clusters are well circumscribed, others are not. This fact, in addition to the multiplicity of connections that are difficult to trace, has prevented a more detailed analysis.

Influence on behavior. For many years the reticular formation was largely ignored, except for a vague descriptive acknowledgment of its presence. However, increased attention has focused on this portion of the brain, mainly because of increasing evidence of its vital, though often subtle, role in many body activities. Sleep, wakefulness, attention, and other aspects of consciousness, as well as effects on muscular coordination, vascular tone, blood pressure, and many other aspects of everyday adjustments to environment are all affected by the condition and responses of the reticular system. *See* SLEEP AND DREAMING.

The discovery of drugs, such as chlorpromazine, which seem to affect selectively the reticular structures created an upsurge of interest and investigation, in which many pharmacological methods were added to those of anatomy and physiology to gain further insight in this field.

The use of electrical stimulation and electroencephalography were added to the microscopic study of neurons, fibers, and degeneration phenomena following experimental injury to selected sites. *See* ELECTROENCEPHALOGRAPHY.

These refinements in techniques led to observations which support the general view that the reticular formation may exert a facilitatory or an inhibitory action on at least three major areas of nervous activity.

Sensory impulses. One area of activity so affected is the modification of ascending impulses received from almost any sensory receptor in the body. The passage of such impulses apparently stimulates the reticular system through a collateral system. The level of activation depends upon previous correlation between cortical areas and the reticular structures involved, so that a kind of presetting mechanism is achieved. An illustration may be helpful. A person goes to bed with the conscious or unconscious knowledge that waking will be signaled by the noise of an alarm clock. Despite many other stimuli, some of a high degree of intensity, the person sleeps until the ring of the alarm. Often, in fact, the alarm need not actually ring because the warning to wakefulness may be elicited by the preliminary click of the alarm control. In a similar manner, a mother may sleep through the crying of other babies but wake at the first cry of her child.

This illustrates at least two principles involved in the modification of ascending impulses: first, the inhibitory or damping effect of the reticular formation on most sensory impulses; and second, the primary importance of some level of cortical association and control over the reticular alarm system.

Similar reactions have been repeatedly demonstrated in animal experiments, wherein electroencephalographic records can be correlated with cortical patterns in many states relating to consciousness.

The various tranquilizers are thought to work through a similar mechanism to produce an inhibition of ascending stimuli and also a decrease in reverberation within the brain itself. *See* TRANQUILIZER.

Motor activity. The second area in which nervous activity is altered by the reticular formation is in its relations with the cerebellum. Under particular conditions, the reticular formation may facilitate or inhibit certain actions involving motor movement and coordination. Somewhat conflicting observations have been reported, but it is apparent that many cerebellar functions are dependent upon intact nuclei in the reticular area, notably the paramedian reticular nuclei, the lateral reticular nuclei, and others.

Much of this relationship is a complex feedback system, wherein continuing motor activity requires continuing adjustment or coordination of both sensory and motor impulses by the modifying reticular centers. Inhibition of extensor tone in decerebrate animals is an example of one aspect of this relationship. Many more subtle influences, particularly those involving localization of motor functions between specific sites in the cerebellum and reticular nuclei, have been reported. The implications of this are of prime significance in terms of the so-called primitive reflex acts and other motor activities.

Integrative impulses. The third area influenced by the reticular formation is that of impulses mediated by the spinal cord. Facilitation or inhibition of cortical or reflex movements by reticular elements has been demonstrated repeatedly. In addition, general and specific effects on vasomotor tone, muscular tonus, and the inspiratory and expiratory phases of

respiration have been traced to alterations induced by reticular activity, or lack of it.

Summary. It may be useful to visualize the reticular formation as a complex, highly integrated mechanism which exerts some degree of inhibition or facilitation on almost every type of nerve-body activity. The dualistic response cannot be overemphasized, particularly because the cortically arranged selectivity of the inhibition-facilitation mechanism is a primary feature. The mediation of sensory, motor, and integrative impulses touches upon specific somatic states of wakefulness, sleep, attention, and related conditions of whole body activity. Finally, coordination and reflex activity require reticular formation participation. Thus, the reticular formation emerges from obscurity to become a fascinating regulatory mechanism, at present only dimly perceived. *See* CONDITIONED REFLEX; PSYCHOLOGY.

Douglas B. Webster

Reticulosa

An order of the subclass Hexasterophora in the class Hexactinellida. This is a group of Paleozoic hexactinellids with a branching form. Each branch is provided with dermal, parenchymal, and gastral spicule reticulations. *Titusvillia* from the Mississippian Period is an example. *See* HEXACTINELLIDA; HEXASTEROPHORA.

Willard D. Hartman

Retinoid receptor

A protein in the cell nucleus that mediates the actions of retinoids (the natural and synthetic analogs of vitamin A) by regulating the rates of transcription of retinoid-responsive genes. Retinoid receptors facilitate the normal biological actions of vitamin A within the body. Thus, they play an important role in the maintenance of normal growth and development; in the immune response; in male and female reproduction; in blood cell development; and in the maintenance of healthy skin and bones; and ultimately in the general good health of the organism. *See* VITAMIN A.

The biochemical details of how retinoid receptors regulate gene transcription have been worked out over the last decade and are well understood. The retinoid receptors recognize specific deoxyribonucleic acid (DNA) sequences, called retinoid response elements, within the regulatory (promoter) regions of genes. The transcription of a retinoid-responsive gene is either increased or diminished through interactions with these response elements, with coactivator or corepressor proteins present in the nucleus, and with the proteins that compose the basal transcription machinery. In general, retinoid binding to a retinoid receptor facilitates gene transcription by making the gene more accessible to enzymes and other factors responsible for the synthesis of ribonucleic acid (RNA) from the gene, whereas the absence of such binding lessens RNA synthesis.

As many as 500 genes may be regulated by retinoids and retinoid receptors. While some of these genes are likely direct targets of retinoid-dependent transcription, many others are undoubtedly indirect targets of retinoid receptor action, reflecting retinoid involvement in modulating communications within the nucleus that are needed to coordinate gene expression. *See* CELL NUCLEUS; DEOXYRIBONUCLEIC ACID (DNA); GENE; PROTEIN.

Receptor structure and distribution. The retinoid receptors are members of the steroid/thyroid/retinoid superfamily of ligand-dependent transcription factors. Each of the proteins that compose this superfamily shares certain common structural features, including a ligand-binding domain and a DNA-binding domain (**Fig. 1**). Often the members of this family of proteins are referred to as hormone nuclear receptors or, more simply, nuclear receptors. Each member of the steroid/thyroid/retinoid superfamily recognizes a specific natural ligand (or ligands), and in response to the availability of this ligand acts to influence either positively or negatively the expression of genes that are responsive to the ligand. *See* NUCLEAR HORMONE RECEPTOR.

Six different retinoid nuclear receptors have been identified, and each is the product of its own individual gene. Based on similarities in their protein structure, three of these are classified as retinoic acid receptors (RARs) and three are classified as retinoid

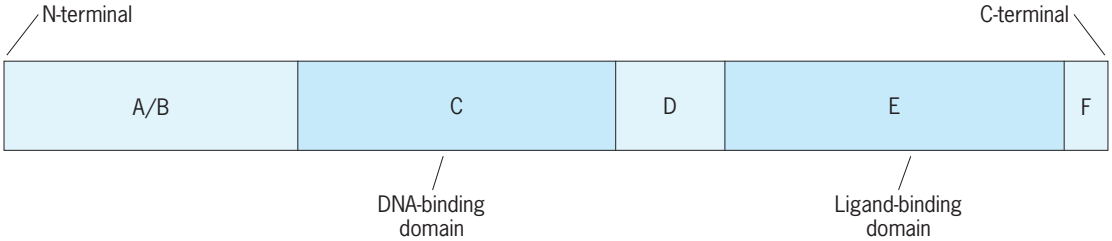

Fig. 1. Schematic of the structural and functional organization of retinoid receptors. The C region of the receptor interacts with response element DNA [DNA-binding domain (DBD)], and the E region with the retinoid ligand [ligand-binding domain (LBD)]; both are highly conserved evolutionarily. All members of the steroid/thyroid/retinoid superfamily of nuclear receptors have the same structure-function organization as illustrated here. This conserved structure-function relationship is central for identifying and classifying nuclear receptors to this superfamily. The A/B, D, and F regions are divergent and not conserved evolutionarily across the members of the steroid/thyroid/retinoid superfamily. However, these regions do play important roles in mediating the functions of the retinoid receptors and other superfamily members. Modifications to these nonconserved regions can have a significant impact on the transcription-modulating activity of the retinoid receptor.

X receptors (RXRs). The three distinct RAR subtypes are termed RAR-α, RAR-β, and RAR-γ; and the RXR subtypes are termed RXR-α, RXR-β, and RXR-γ. Furthermore, for each subtype there are at least two isoforms that are generated by differential promoter usage and alternative splicing. All-*trans*- and 9-*cis*-retinoic acid stereoisomers of retinoic acid are naturally occurring retinoid forms that have physiologically important roles in activating the retinoid receptors. All-*trans*-retinoic acid activates members of the RAR family of nuclear receptors and 9-*cis*-retinoic acid activates members of both the RAR and RXR family. Thus, different retinoid species can activate differentially the RAR and RXR retinoid receptor families.

Different combinations of RAR and RXR species are found in tissues and cells. Thus, not all tissues and cells have the same complement of RARs or RXRs. It is not fully understood what significance these differences in RAR and RXR distribution have for mediating retinoid actions in the body. It is clear, however, that almost all tissues and cells within the human body have one or more of the RARs or RXRs. This implies that retinoids and retinoid receptors play an important and diverse role in maintaining the health of all tissues within the body.

Mechanism of transcriptional regulation. Upon binding to a cognate response element, the retinoid receptors must interact with other nuclear proteins that function as either coactivators or corepressors of transcription. These interactions are essential for regulating transcriptional rates of retinoid responsive genes because they influence the degree of acetylation or deacetylation of histone proteins. The binding of retinoids to their receptors enables the receptors to interact with coactivators that have histone acetyltransferase activity (**Fig. 2**). This results in the local acetylation of histone proteins and their dissociation from the gene and an opening up of chromatin structure. This allows the transcriptional machinery access to the DNA and results in enhanced rates of gene expression. In the absence of retinoid, unliganded retinoid receptors can interact with corepressors that have histone deacetylase activity (**Fig. 3**). This results in the local deacetylation of histones and a closure of chromatin structure, thus lessening the activity of the transcriptional machinery for that gene. Consequently, corepressor binding by retinoid receptors serves to diminish transcription rates.

The response elements within genes to which the retinoid receptors bind are well characterized for many retinoid-responsive genes. Most, but not all, of these retinoid response elements share certain common characteristics, including a direct repeat (DR) of a six-nucleotide sequence, in which the two repeated sequences are separated by either two or five nucleotides (base pairs). These response elements, which consist of two 6-nucleotide half-sites separated by two or five nucleotides, are referred to as DR-2 or DR-5 motifs. The strongest, or most retinoid responsive, of these elements has a consensus six nucleotide half-site sequence of AGGTCA (A = adenylate, C = cytidylate, G = guanylate, and T = thymidylate). Although minor variations in this half-site nucleotide sequence are found in different genes across different species, the presence of two repeated stretches of a sequence consisting of AGGTCA separated by two or five nucleotides in the promoter region of a gene suggests that the gene

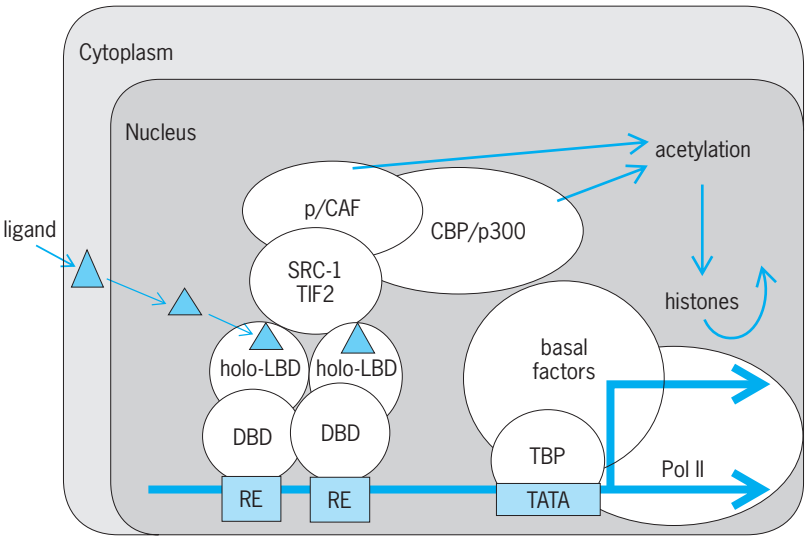

Fig. 2. Model for retinoid receptor-mediated gene activation. The DBD is able to bind the retinoid response element (RE), and an activating retinoid is available to bind the LBD, yielding a holo-LBD form. This displaces corepressors that may have bound the receptor in the absence of ligand, and facilitates binding of coactivators (p/CAF, SRC-1, TIF2, and CBP/p300 in this example). These bound coactivators catalyze the local acetylation of histones associated with the DNA composing the gene. This has the effect of releasing the histones from the DNA and making the gene more accessible to the basal transcriptional machinery (the basal factors, TBP and Pol II). This has the overall effect of enhancing the actions of this gene product within the cell.

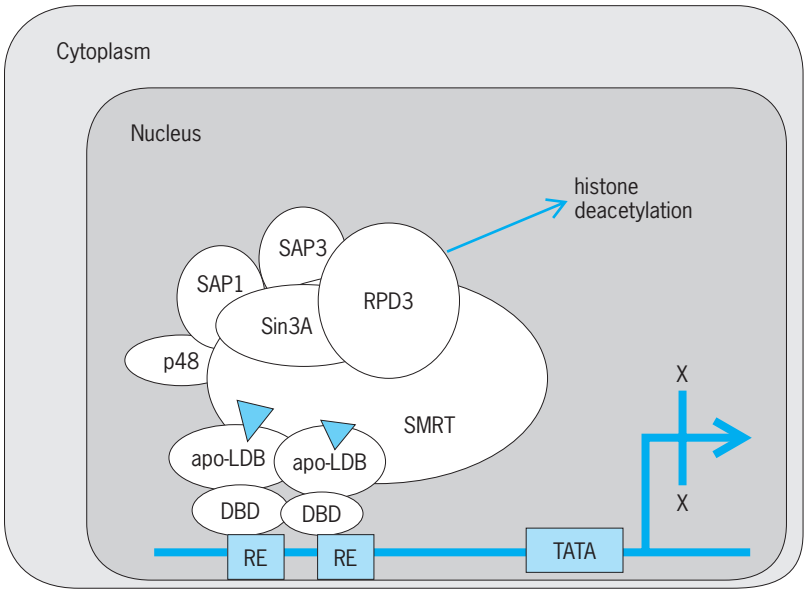

Fig. 3. Model for retinoid receptor-mediated gene silencing. The DBD is able to bind the RE. In the absence of retinoid, the LBD is unoccupied or in the apo-LBD form. This allows corepressors (p48, Sin3A, SAP1, SAP3, RPD3, and SMRT) to bind to the retinoid receptor and to catalyze local deacetylation of histones. As a result, the histones bind the DNA more tightly, thus preventing the basal transcription machinery from efficiently transcribing the gene. This has the overall effect of lessening the actions of this gene product within the cell.

is responsive to retinoid receptor action. *See* GENE ACTION; NUCLEIC ACID; NUCLEOPROTEIN.

Physiologic actions. The actions that retinoids and their receptors have in regulating gene activity affect many essential physiologic processes within the body. Living organisms require retinoids to maintain normal embryonic growth and differentiation, epithelial linings throughout the body, reproduction, an immune response, and normal brain function, in addition to assuring normal blood cell development and proper bone growth and remodeling. At the cellular level, retinoids are needed to maintain a balance between cellular differentiation and proliferation and cell death. A review of even a partial listing of the more than 500 genes reported to be responsive to retinoids provides insight into the diversity of responses regulated by retinoid receptors and their importance in maintaining the health of living organisms (see **table**).

This importance is further underscored by targeted gene disruptions (gene knockouts) in mice, in which one or more of the genes encoding RARs or RXRs have been deleted. Single RAR or RXR gene knockouts have resulted in mice that show severe developmental or functional impairments in organ systems such as the cardiovascular system, the central nervous system, the reproductive system, and the gastrointestinal system. Disruption of combinations of two RAR and/or RXR genes results in embryonic lethality. Presumably, the absence of marked embryonic lethality for single-gene knockouts arises from a functional redundancy in the actions of the RARs and/or RXRs.

Clinical applications. Because retinoids are such potent regulators of cellular processes, there has been considerable interest in the use, or potential use, of retinoids to treat chronic diseases. This interest centers especially on disease states involving aberrant cell proliferation and differentiation. Over 10,000 synthetic retinoids have been produced by medicinal chemists in the hope of developing effective pharmacological agents for use in the treatment of disease. Some of these synthetic retinoids can act as retinoid receptor agonists and bring about the same biologic actions as all-*trans*- or 9-*cis*-retinoic acid. Others act as retinoid receptor antagonists and block the actions of all-*trans*- or 9-*cis*-retinoic acid.

Examples of genes regulated by retinoid receptors

Substance encoded in gene	Action
Oxytocin	Needed for reproduction
Growth hormone	Needed to maintain growth
Phosphoenol pyruvate carboxykinase	Involved in glucose metabolism
Class I alcohol dehydrogenase	Needed for alcohol oxidation
Tissue transglutaminase	Needed to regulate cell growth and death
Laminin B1	Needed to facilitate cell-cell interactions
Matrix gla-protein	Needed for bone formation
Several keratin genes	Needed for healthy skin formation
Cellular retinol-binding protein, type I	Active in retinoid metabolism
RAR-β	Affects retinoid action
Cytochrome 26	Catalyzes catabolism of all-*trans*-retinoic acid
Several hox genes	Needed to control embryologic development
Dopamine D2 receptor	Needed in central nervous system function

Skin disease. At present, both naturally occurring and synthetic retinoids are used clinically to treat skin disease. Among the retinoids commonly used in clinical dermatology are 13-*cis*-retinoic acid (also known as isotretinoin or Accutane®) and all-*trans*-retinoic acid (also known as tretinoin or RetinA®). Natural and synthetic retinoids have been used effectively in the treatment of skin disorders such as severe cystic acne, psoriasis, many cutaneous disorders of keratinization, several dermatoses, acne vulgaris, and the damaging effects of sun exposure. Although retinoids are highly effective for use in clinical dermatology, their systemic use is associated with a high incidence of birth defects (teratogenic effects). They must never be taken by pregnant women or women considering pregnancy.

Cancer. Since retinoids have an important role in regulating cellular proliferation and differentiation, there has long been interest in their use as cancer chemopreventive and chemotherapeutic agents. The most successful use of retinoids as cancer therapeutic agents has been for treatment of some forms of acute promyelocytic leukemia (APL). Retinoids, administered either as all-*trans*- or 13-*cis*-retinoic acid, have been reported to have clinical activity in cancer therapy and prevention in several other settings. These include treatment of the premalignant lesions of oral leukoplakia, actinic keratosis, cervical dysplasia, and xeroderma pigmentosum. *See* CANCER (MEDICINE); LEUKEMIA.

Cardiovascular disease and diabetes. Much recent work in animal models suggests that the pharmacological use of retinoids to target retinoid receptor action will be of benefit to patients with cardiovascular disease and with type II diabetes. Data from these studies indicate that rexinoids have efficacy for lowering plasma lipid levels, for preventing restenosis injury, and for increasing insulin responsiveness. (A rexinoid is an RXR-specific agonist or antagonist.) Clinical trials to demonstrate rexinoid efficacy in human patients are being initiated. *See* DIABETES; HEART DISORDERS.

Toxicity. Because retinoids are such potent modulators of cellular actions and because they influence many diverse physiologic actions, retinoids can have many undesirable toxic effects when administered in pharmacologically effective doses. The toxicity of retinoids is the major factor that impedes their widespread use for treating disease. Nevertheless, there is a large body of clinical research focused on using retinoids pharmacologically to treat or to prevent chronic disease.

Solveig Halldorsdottir; William S. Blaner

Bibliography. P. Chambon, A decade of molecular biology of retinoic acid receptors, *FASEB J.*, 10:940–954, 1996; A. Chawla et al., Nuclear receptors and lipid physiology: Opening the X-files, *Science,* 294:1866–1870, 2001; IARC Working

Group on the Evaluation of Cancer Prevention Agents, *IARC Handbooks of Cancer Prevention*, vol 3: *Vitamin A*, World Health Organization, International Agency for Research on Cancer, Lyon, 1998; IARC Working Group on the Evaluation of Cancer Prevention Agents, *IARC Handbooks of Cancer Prevention*, vol 4: *Retinoic Acid*, World Health Organization, International Agency for Research on Cancer, Lyon, 2000; H. Nau and W. S. Blaner (eds.), *Retinoids: The Biochemical and Molecular Basis of Vitamin A and Retinoid Action*, vol. 139 of *Handbook of Experimental Phyarmacology*, Springer Verlag, Heidelberg, 1999; L. Packer et al. (eds.), *Carotenoids and Retinoids: Molecular Aspects and Health Issues*, AOCS Press, Champaign, 2005; M. B. Sporn, A. B. Roberts, and D. S. Goodman (eds.), *The Retinoids: Biology, Chemistry, and Medicine*, 2d ed., Raven Press, New York, 1994; L.-N. Wei, Retinoid receptors and their coregulators, *Annu. Rev. Pharmacol. Toxicol.*, 43:47–72, 2003.

Retortamonadida

An order of parasitic flagellate protozoa belonging to the class Zoomastigophorea. P. P. Grassé suggested that there should be eight orders of these parasitic (and symbiotic) zooflagellates, all in a superorder Metamonadina, but this may be an oversimplification. All retortamonads are medium to large in size and have a complicated blepharoplast-centrosome-axo-style apparatus. Retortamonadida have two or four flagella, one turned ventrally into a cytostomal depression. The nucleus, containing a distinct endosome, is located at the anterior tip. The body is twisted. These organisms are actually symbionts, ingesting bacteria in the digestive tracts of their hosts.

Retortamonas has several species, some of which infest insects and vertebrates. *Chilomastix* also has a number of species, found in vertebrates and invertebrates (see **illus.**). *Chilomastix mesnili* is probably the best-known species; it is found in humans, but is believed to be harmless. It shows the sinistral spiraling very well, and the cytostome with the fourth flagellum lying therein. The cell divides by a well-defined mitosis; the nucleus migrates to midcell, where it forms a nuclear spindle. The poles are blepharoplasts, which give rise to flagella. *See* PROTOZOA; SARCOMASTIGOPHORA; ZOOMASTIGOPHOREA.

James B. Lackey

Bibliography. J. A. Pechenick, *Biology of the Invertebrates*, 5th ed., McGraw-Hill, 2005.

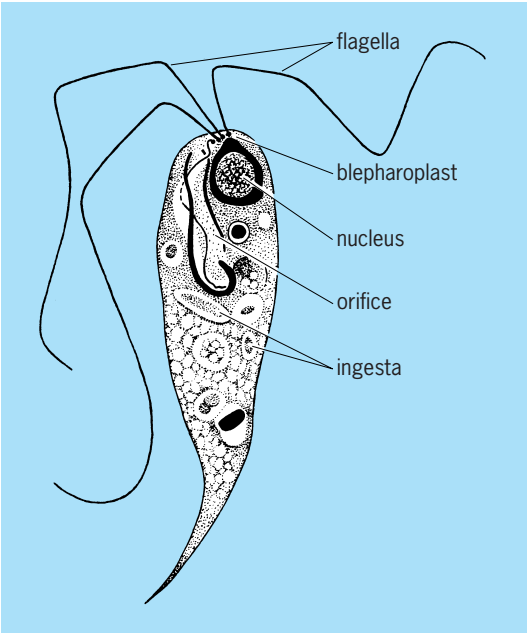

A retortamonad, *Chilomastix aulastomi*.

Retrograde motion (astronomy)

In astronomy, either an apparent east-to-west motion of a planet or comet with respect to the background stars or a real east-to-west orbital motion of a comet about the Sun or of a satellite about its primary. The majority of the objects in the solar system revolve from west to east about their primaries. However, near the time of closest approach of Earth and a superior planet, such as Jupiter, because of their relative motion, the superior planet appears to move from east to west with respect to the background stars. The same apparent motion occurs for an inferior planet, such as Venus, near the time of closest approach to Earth.

Actual, rather than apparent, retrograde motion occurs among the satellites and comets; the eighth and ninth satellites of Jupiter and the ninth satellite of Saturn are examples. *See* ORBITAL MOTION.

Raynor L. Duncombe

Retrovirus

A family of viruses distinguished by four characteristics: (1) genetic information is in the form of ribonucleic acid (RNA); (2) each single virus particle or virion (**illus.** *c*, arrow) possesses the enzyme reverse transcriptase; (3) virion morphology consists of two proteinaceous structures, a dense core (illus. *a*) and an envelope (illus. *a*) that surrounds the core; and (4) the viral reproductive cycle requires the synthesis of a deoxyribonucleic acid (DNA) copy of the viral genome and insertion of that DNA into the genome of the host cell. *See* ANIMAL VIRUS; RIBONUCLEIC ACID (RNA).

The genome is composed of two identical molecules of single-stranded RNA. This genomic RNA is similar in function to messenger RNA (positive polarity). The genome contains two structural genes, termed *gag* and *env*, that code for the structural coat of the virus. *Env* codes for the viral proteins of the outer envelope surface of the virus and *gag* codes for the core proteins. The core contains the RNA genome of the virus.

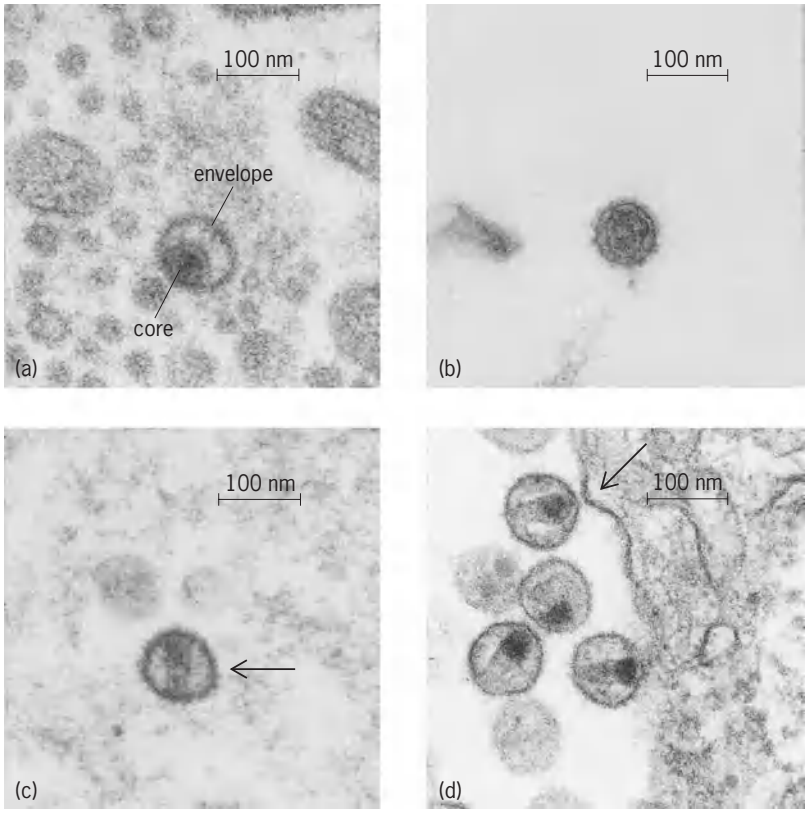

Different morphological types of retroviruses. (*a*) Murine mammary tumor virus, a betaretrovirus, shown with its off-center core surrounded by an envelope. (*b*) Feline leukemia virus, a gammaretrovirus, shown with its centrally located core. (*c*) SRV, a betaretrovirus of Asian macaque monkeys, shown with its rodlike core. (*d*) Human immunodeficiency virus (HIV), a lentivirus that causes acquired immune deficiency syndrome (AIDS), shown with its conical core.

Other genes, such as *pol*, encode enzymes such as the reverse transcriptase that is needed for viral multiplication. Some retroviruses have additional genes as well. DNA is not present in the virions of retroviruses. The reverse transcriptase in each virion copies the RNA genome into DNA shortly after entry of the virus into the animal or human host cell. The discovery of this enzyme changed thinking in biology. Previously, the only known direction for the flow of genetic information was from DNA to RNA, yet retroviruses make DNA copies of their genome by using an RNA molecule. This reversal of genetic information was "backwards," and hence the family name retrovirus, meaning backward virus. *See* REVERSE TRANSCRIPTASE.

The DNA copy of the RNA genome is inserted directly into one of the chromosomes of the host cell. This results in new genetic information being acquired by the host cell. Acquiring new genetic information can induce changes in that cell that may lead to tumors. The study of reverse transcriptase led to other discoveries of how retroviruses may add a variety of new genetic information into the host. One such class of new genes carried by retroviruses is oncogenes (tumor genes). Oncogenes are carried either in place of or in addition to one of the normal genes of retroviruses. Retroviral oncogenes appear to be involved in the induction of tumors in animals. *See* ONCOGENES.

Multiplication. Viral multiplication occurs in six steps: adsorption, penetration, uncoating, synthesis of viral proteins and nucleic acids, assembly and maturation, and release. In the first step, the external envelope of the virus attaches itself to the surface cellular membrane of an animal or human cell (illus. *d*, arrow). The cell takes up the bound virus (penetration) and the virus releases its genome (uncoating) and begins to multiply in the host cell. Viral synthesis begins when the reverse transcriptase copies the RNA genome into DNA. A viral integrase enzyme then inserts the newly made DNA into the genome of one of the host chromosomes in the nucleus of the infected cell. The viral DNA, called a provirus in this state, is transcribed by cell enzymes to form viral messenger RNAs as well as new viral RNA genomic molecules. The messenger RNAs are then translated into viral proteins such as the reverse transcriptase, gag, and envelope proteins. These newly made components assemble into new progeny virus, signaling the end of viral synthesis and the beginning of maturation. The newly made virions are released from the cell surface by budding through the outer cell membrane. The newly made virions will continue the cycle by infecting new cells in that host or by transmission to a new individual of that species.

Specific genera of retroviruses and diseases. The family Retroviridae contains all of the Retroviruses and is divided into seven genera; Alpharetrovirus, Betaretrovirus, Gammaretrovirus, Deltaretrovirus, Epsilonretrovirus, Lentivirus, and Spumavirus. Retroviruses are either exogenous or endogenous. Exogenous retroviruses are transmitted horizontally by infection of new cells or hosts, just as other viruses are transmitted. Endogenous viruses, unlike other viruses, are transmitted vertically, as their DNA proviruses are inherited through the genome of a species.

1. Alpharetroviruses contain the avian leucosis virus and the Rous sarcoma virus. All members infect avian species and some are associated with tumors.

2. Betaretroviruses contain viruses that infect mice, certain species of monkeys, and sheep. Mouse mammary tumor virus is associated with mammary tumors in mice. Mason-Pfizer monkey virus, SRV-1 and SRV-2, are natural infections of Asian macaque monkeys. Infected monkeys develop an AIDS-like disease and also may have fibrosarcoma, a tumor similar to Kaposi sarcoma, a skin cancer seen in humans with AIDS. SRV-1, SRV-2, and Mason-Pfizer monkey virus are transmitted by a bite that pierces the skin. The betaretroviruses of langur and squirrel monkeys are endogenous retroviruses, meaning that they are not acquired by infection but are inherited. Ovine pulmonary adenosarcoma, a tumor of sheep, is associated with the Jaagsiekte sheep retrovirus.

3. Gammaretrovirus contains retroviruses of diverse species, including snakes and the domestic cat. Feline leukemia virus (FeLV) (illus. *b*) infects the domestic cat and is of significant economic importance as a cause of an AIDS-like disease and leukemia in that species. There is no evidence of infection of humans

by FeLV. Vaccines against FeLV are available and are widely used. Retroviruses that are associated with leukemia in gibbon apes and in mice are also known. Endogenous retroviruses of mice are also included in this genus.

4. Deltaretroviruses contain leukemia viruses of cattle, humans, and simians. The human T cell leukemia virus (HTLV) is one of two retroviruses known to cause disease in humans. The other is HIV, the cause of AIDS. HTLV is associated with human leukemias of T cell origin and neurologic diseases. The closest known relative of HTLV is simian T cell leukemia virus (STLV). STLV naturally infects a number of simian species in Africa and Asia. Genetic comparisons between HTLV and STLV indicate that HTLV originated from cross-species transmission of STLV to humans. Leukemia in humans develops in about 2% of infected persons.

5. Epsilonretrovirus contains exogenous retroviruses of fish. The Walleye dermal sarcoma virus, perch hyperplasic virus, and snakehead retrovirus have been described.

6. The genus Lentivirus is named for the slow progression to disease in the infected host. Only exogenous lentiviruses are known. Members include Visna-Maedi virus, associated with pulmonary and central nervous system diseases of sheep, and the human immune deficiency virus (HIV), which is the cause of AIDS in humans.

Two lentiviruses found in chimpanzees and sooty mangabey monkeys, SIVcpz and SIVsm, are the ancestral viruses of HIV type 1 and HIV type 2, respectively. Although HIV has its origins in African simian species, it is not correct to think of AIDS as a zoonosis (a disease acquired directly from animals, like rabies); there is no evidence that SIV directly causes AIDS in human beings. Over 30 SIVs have been identified in simian species on the African continent. It is noteworthy that SIV rarely causes AIDS in its natural simian hosts in Africa.

Other animal lentiviruses are known, including feline immunodeficiency virus, bovine immunodeficiency virus, and equine infectious anemia virus.

7. Spumavirus contains the foamy viruses, so called because they appear to cause foam when infecting tissue culture cells. These viruses are found in humans, chimpanzees, monkeys, and domestic cats. The virus is present in saliva. There is evidence that the simian foamy viruses gave rise to human infections. There is no known disease associated with the spumavirus group.

Epidemiology. Endogenous retroviruses are inherited as part of the genome of the host species. Exogenous retroviruses are acquired by infection. The exogenous retroviruses gain entry into the host through the respiratory system, circulatory system, genital tract, or by a bite or other action that pierces the skin. Retroviruses may be sexually transmitted, as is HIV. HIV can also be passed from mother to infant during gestation, at birth, or through infected breast milk. Treatment of the infected mother may effectively prevent transmission to her child. Infection through the respiratory tract occurs when a virus enters the mouth or nose. Entry into the circulatory system can occur from a blood transfusion, use of contaminated needles, or from insect bites, as is the case for equine infectious anemia. However, HIV is not transmitted through insect bites.

Origin of the AIDS virus. AIDS was identified in Los Angeles in 1981. The causative agent, HIV-1, was discovered by scientists at the Pasteur Institute in Paris in 1983. Thus far, all available evidence indicates that the AIDS virus is spreading worldwide in humans for the first time. There are no known past epidemics. Computer models that attempt to calculate the age of HIV based on its genetic makeup place its origin from 50 to 75 years ago. The SIV viruses that are ancestral to HIV are found in naturally infected simian hosts (the chimpanzee and sooty mangabey) that live along the west coast of Africa. Yet, despite 300 years of slave trading from that same African coast, HIV did not spread to North or South America, indicating HIV's origin after the end of the slave trade. This fact makes AIDS a new disease, relatively speaking.

Although the African origin of HIV from infected monkeys and apes is certain, it is not known what launched HIV and how it became a potent human pathogen. HIV probably adapted to humans in the twentieth century by mechanisms still unknown.

Preston A. Marx

Bibliography. C. Apetrei, D. L. Robertson, and P. A. Marx, The history of SIVs and AIDS: Epidemiology, phylogeny and biology of isolates from naturally SIV infected non-human primates (NHP) in Africa, *Frontiers Biosci.*, 9:225–254, 2004; E. Drucker, P. G. Alcabes, and P. A. Marx, The injection century: Consequences of massive unsterile injecting for the emergence of human pathogens, *Lancet*, 358:1989, 2001; D. M. Knipe et al. (eds.), *Fields Virology*, 2 vols., 4th ed., pp. 1871–2123, 2001.

Reverberation

After sound has been produced in, or enters, an enclosed space, it is reflected repeatedly by the boundaries of the enclosure, even after the source ceases to emit sound. This prolongation of sound after the original source has stopped is called reverberation. A certain amount of reverberation adds a pleasing characteristic to the acoustical qualities of a room. However, excessive reverberation can ruin the acoustical properties of an otherwise well-designed room. A typical record representing the sound-pressure level at a given point in a room plotted against time, after a sound source has been turned off, is given in the decay curve shown in **Fig. 1**. The rate of sound decay is not uniform but fluctuates about an average slope.

Reverberation time. Because of the importance of the proper control of reverberation in rooms, a standard of measure called reverberation time (abbreviated t_{60}) has been established. Reverberation time is the time required for sound to die away to one-thousandth of its initial pressure, that is, to drop 60 dB in sound-pressure level.

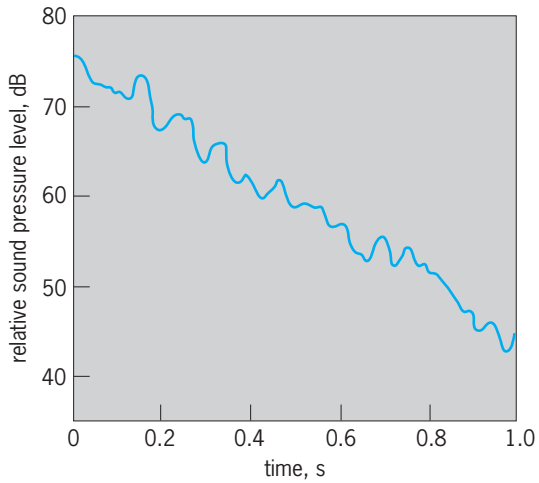

Fig. 1. Typical decay curve illustrating reverberation.

Optimum reverberation time is a matter of individual preference. A critical study of empirical data based upon preference evaluations in the United States and abroad has been made by V. O. Knudsen and C. M. Harris (**Fig. 2**). Since the optimum reverberation time for music depends on the type of music, it is represented in the figure by a broad band. The optimum reverberation time for a room used primarily for speech is considerably shorter; reverberation times longer than those shown for speech result in a decrease in speech intelligibility. The optimum reverberation time at frequencies other than 500 Hz is obtained by multiplying the 500-Hz value by the ratio R, which is given in **Fig. 3**. Note that R is unity for frequencies above 500 Hz and is given by a band for frequencies below 500 Hz. For large rooms R may have any value within the indicated band; for small rooms preferred ratios are in the lower part of the band.

Mean free path. According to the principles of geometrical acoustics, sound radiated from a source in an enclosure is successively reflected by its boundaries. The average distance between reflections is defined as the mean free path. The mean free path of a sound ray in a room depends on the shape and size of the room, and to some extent on the distribution and nature of the absorptive material. However, in most cases, it is approximately $4V/S$, where V is the volume of the room and S the total surface area.

Decay rate. The number of reflections per second of a decaying sound wave is numerically equal to the distance sound will travel in 1 s, that is, the velocity of sound c, which is about 343 m/s (1130 ft/s) in air at 20°C (68°F), divided by the average distance between reflections, or the mean free path. Hence the number of reflections per second is $cS/4V$. Each time a wave strikes one of the boundaries, on the average, a fraction (α) of the energy is absorbed, and a fraction $(1 - \overline{\alpha})$ is reflected. Here $\overline{\alpha}$ is the average absorption coefficient given by Eq. (1), where α_1 is

$$\overline{\alpha} = \frac{\alpha_1 S_1 + \alpha_2 S_2 + \alpha_3 S_3 + \cdots}{S_1 + S_2 + S_3 + \cdots} \qquad (1)$$

the coefficient of absorption of surface S_1, and so forth. Because sound pressure is proportional to the square root of sound intensity, the ratio of the average reflected pressure to incident pressure is given by $(1 - \overline{\alpha})^{1/2}$, and the average decrease in the sound pressure level is therefore given by notation (2).

$$10 \log_{10}\left(\frac{1}{1-\overline{\alpha}}\right) \text{dB/reflection} \qquad (2)$$

Since there are $cS/4V$ reflections per second, the average decay rate is given by Eq. (3*a*), where S is in square meters and V is in cubic meters, or (3*b*), where S' is in square feet and V' is in cubic feet.

$$\text{Decay} = 373(S/V) \cdot [-2.30 \log_{10}(1-\overline{\alpha})] \text{ dB/second} \qquad (3a)$$

$$\text{Decay} = 1230(S'/V') \cdot [-2.30 \log_{10}(1-\overline{\alpha})] \text{ dB/second} \qquad (3b)$$

Reverberation-time formulas. From Eqs. (3) for decay rate, it follows that the time it takes for the second-pressure level to decay 60 dB, that is, the reverberation time, is given by Eqs. (4).

$$t_{60} = \frac{0.161V}{S[-2.30 \log_{10}(1-\overline{\alpha})]} \text{ seconds} \qquad (4a)$$

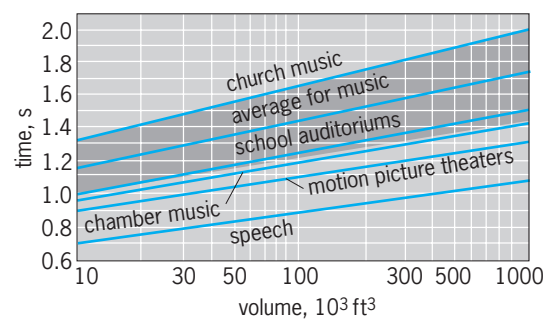

Fig. 2. Optimum reverberation time at 500 Hz for different types of rooms as a function of room volume. This figure should be used in conjunction with Fig. 3 to obtain optimum reverberation time as a function of frequency. 10^3 ft = 28.3 m^3. (*After V. O. Knudsen and C. M. Harris, Acoustical Designing in Architecture, John Wiley and Sons, 1950*)

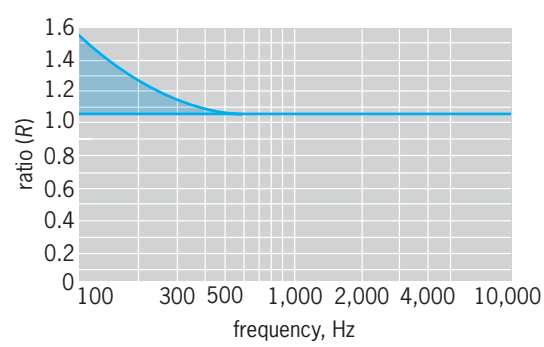

Fig. 3. Chart for computing optimum reverberation time. The time at any frequency is given in terms of the ratio R, which should be multiplied by the optimum time at 500 Hz (from Fig. 2) to obtain the optimum time at that frequency. (*After V. O. Knudsen and C. M. Harris, Acoustical Designing in Architecture, John Wiley and Sons, 1950*)

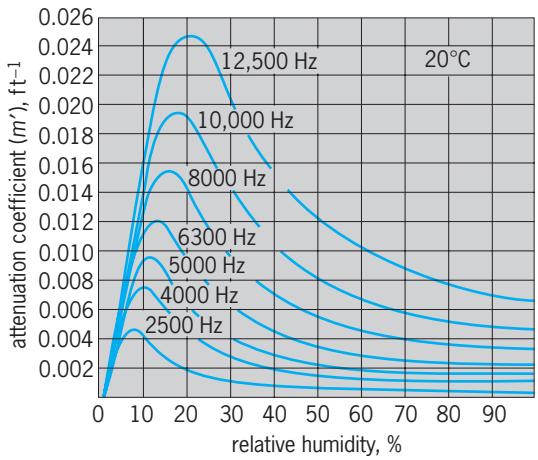

Fig. 4. Values of the attenuation coefficient *m* as a function of relative humidity for different frequencies. 1 ft^{-1} = 3.28 m^{-1}. 20°C = 68°F. (*After C. M. Harris, Absorption of sound in air versus humidity and temperature, J. Acoust. Soc. Amer., 40:148–159, 1966*)

$$t_{60} = \frac{0.049V'}{S'[-2.30\log_{10}(1-\overline{\alpha})]} \text{ seconds} \qquad (4b)$$

When $\overline{\alpha} \ll 1$, Eqs. (4) become Eqs. (5). For frequencies above 2000 Hz, especially in large auditoriums, the effects of air absorption must be included in the reverberation time formulas. The corresponding equations are then Eqs. (6) and (7), where m is the attenuation coefficient in inverse meters and m' is the attenuation coefficient in inverse feet, given in **Fig. 4** . It can be shown that air absorption is molecular in origin.

$$t_{60} = \frac{0.161V}{S\overline{\alpha}} \text{ seconds} \qquad (5a)$$

$$t_{60} = \frac{0.049V'}{S'\overline{\alpha}} \text{ seconds} \qquad (5b)$$

$$t_{60} = \frac{0.161V}{S[-2.30\log_{10}(1-\overline{\alpha})+4mV]} \text{ seconds} \qquad (6a)$$

$$t_{60} = \frac{0.049V'}{S'[-2.30\log_{10}(1-\overline{\alpha})+4m'V']} \text{ seconds} \qquad (6b)$$

$$t_{60} = \frac{0.161V}{S\overline{\alpha}+4mV} \text{ seconds} \qquad (7b)$$

$$t_{60} = \frac{0.049V'}{S'\overline{\alpha}+4m'V'} \text{ seconds} \qquad (7b)$$

The above reverberation time formulas apply only in rooms in which sound is diffuse, a condition that was assumed in their derivation. Thus, large discrepancies may be noted between the calculated and observed values of reverberation time in a room whose shape or absorptive treatment (or both) do not promote diffuse conditions of sound. *See* ARCHITECTURAL ACOUSTICS; SOUND.

Cyril M. Harris

Bibliography. T. Cox and P. D'Antonio, *Acoustic Absorbers and Diffusers*, Spon Press, London, 2004; M. D. Egan, *Architectural Acoustics*, McGraw-Hill, 1988; C. M. Harris (ed.), *Handbook of Acoustical Measurements and Noise Control*, 3d ed., McGraw-Hill, 1991; L. E. Kinsler et al., *Fundamentals of Acoustics*, 4th ed., Wiley, 2000; V. O. Knudsen and C. M. Harris, *Acoustical Design in Architecture*, Wiley, 1978; Z. MacKawa and P. Lord, *Environmental and Architectural Acoustics*, E & FN Spon, London, 1994.

Reverse transcriptase

Any of the deoxyribonucleic acid (DNA) polymerases present in particles of retroviruses, that is, in virions of ribonucleic acid (RNA) tumor viruses and related viruses. These DNA polymerases are coded by the retroviruses and are able to carry out DNA synthesis using an RNA template. This reaction is called reverse transcription since it is the opposite of the usual transcription reaction, which involves RNA synthesis using a DNA template. Thus, the viral DNA polymerases that carry out this reverse reaction have been called reverse transcriptases, and the viruses with this enzyme are called retroviruses because they reverse the usual flow of genetic information. Such viruses are much discussed now, since the etiologic agent of acquired immune deficiency syndrome (AIDS) is a retrovirus. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); RETROVIRUS.

Genetic information flow. DNA carries the genetic information of all living systems except RNA viruses. As shown in **Fig. 1**, the genetic information in DNA for proteins is first transferred to an RNA molecule, messenger RNA, by a process called transcription and then to the final product, protein, by a process called translation. The other process of information transfer in living systems except for RNA viruses is the multiplication of DNA by a process known as replication. In all living systems except some viruses, genetic information flows only from DNA to DNA and from DNA to RNA to protein. *See* GENETIC CODE; PROTEIN.

Virus classification. Viruses are simple genetic systems that consist of genes coding for viral structural components and for proteins necessary for virus replication. Viruses can replicate only in living cells.

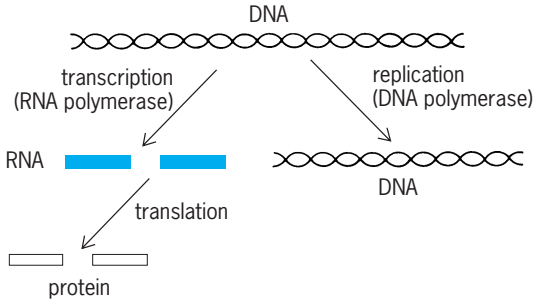

Fig. 1. Modes of genetic information transfer found in all living systems except some viruses.

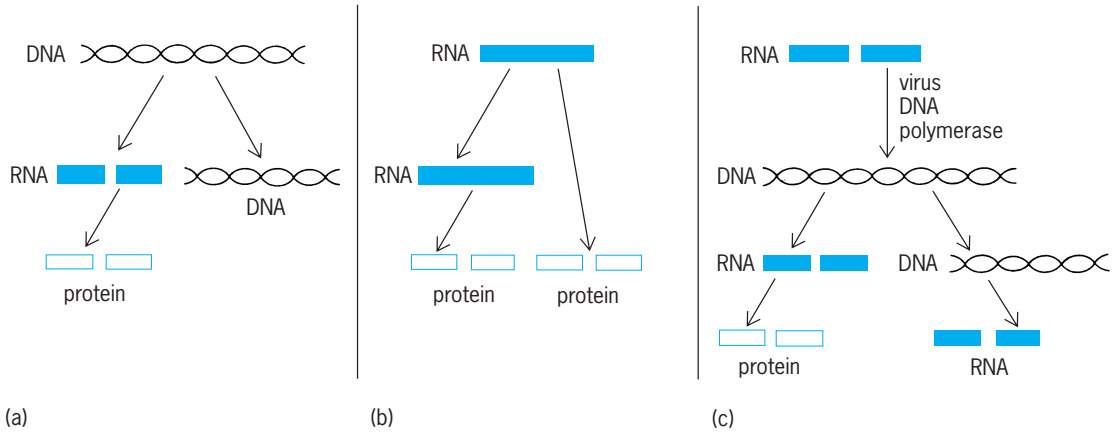

Fig. 2. Modes of information transfer that characterize the replication of viruses in three major classes. (*a*) DNA viruses or deoxyviruses (example: smallpox virus); (*b*) RNA viruses or riboviruses (example: poliovirus); and (*c*) retroviruses (example: Rous sarcoma virus and human immunodeficiency virus).

Viruses are grouped into three major classes according to the kind of nucleic acid contained in the extracellular and intracellular viral genomes (**Fig. 2**). There are viruses whose genome is only DNA (DNA viruses or deoxyviruses), viruses whose genome is only RNA (RNA viruses or riboviruses), and viruses whose genome is RNA in extracellular virus particles and DNA inside cells (retroviruses). There are also viruses whose genome is DNA in extracellular virus particles and RNA inside cells (pararetroviruses, that is, hepadnaviruses such as hepatitis B virus and caulimoviruses). Deoxyviruses use the same modes of information transfer as do cells. Riboviruses use information transfer from RNA to RNA in addition to information transfer from RNA to protein. Retroviruses use an additional mode of information transfer from RNA to DNA. *See* VIRUS.

DNA provirus. The transfer of genetic information from RNA to DNA in retrovirus replication was proposed in 1964 by H. M. Temin in the DNA provirus hypothesis for the replication of Rous sarcoma virus, an avian retrovirus which causes tumors in chickens and transformation of cells in culture. Rous sarcoma virus is the prototype retrovirus.

The DNA provirus hypothesis was based on the results of studies of the effects of inhibitors of nucleic acid synthesis on the replication of Rous sarcoma virus. The replication of Rous sarcoma virus is inhibited by actinomycin D, an inhibitor of DNA-dependent RNA synthesis. In addition, when DNA synthesis is inhibited by compounds such as cytosine arabinoside early after infection of nondividing cells by Rous sarcoma virus, no Rous sarcoma virus replication or cell transformation is found. These studies indicated that a DNA intermediate is involved in the replication of Rous sarcoma virus. The discovery in 1970 of an enzyme, reverse transcriptase, that carries out RNA-directed DNA synthesis in the particles of retroviruses provided convincing evidence in support of the DNA provirus hypothesis. (Reverse transcriptase also carries out DNA-directed DNA synthesis.) Direct evidence for the DNA provirus hypothesis was then obtained by nucleic acid hybridization experiments that detected new Rous sarcoma virus–specific DNA sequences in the DNA of Rous sarcoma virus–infected cells, and by the isolation of infectious DNA from Rous sarcoma virus–infected cells which, after addition to cells in culture (by transfection), resulted in the production of complete infectious progeny Rous sarcoma virus.

The role of reverse transcriptase during retrovirus replication was established by the study of temperature-sensitive mutants of Rous sarcoma virus with a temperature-sensitive reverse transcriptase. These virus mutants do not replicate at temperatures that inactivate the DNA polymerase activity in their virions. In addition, retrovirus particles that lack reverse transcriptase are not infectious. Therefore, the reverse transcription of information from the viral RNA genome to the viral DNA intermediate by the reverse transcriptase in retrovirus virions is an essential early step in retrovirus replication.

Now, RNA-directed DNA polymerase activity or genes coding for proteins homologous to reverse transcriptase have been reported in several cellular genetic elements. These are listed in the **table** and are discussed below.

Properties of reverse transcriptase. Reverse transcriptase has been purified from virions of many retroviruses. The avian, murine, and human retrovirus DNA polymerases have been extensively studied.

All reverse transcriptases have ribonuclease H activity on the same molecule as the DNA polymerase activity. Ribonuclease H activity is an enzymatic activity that degrades only the RNA strand of RNA-DNA hybrid molecules. It is involved in retrovirus DNA synthesis. The ribonuclease H activity of the retrovirus DNA polymerase is endonucleolytic. A DNA endonuclease activity, in the form of an enzyme that can cut DNA and is involved in retrovirus integration, is also coded by the reverse transcriptase gene.

Most of the other biochemical properties of the purified reverse transcriptases are common to them and to other DNA polymerases. For example, like cellular DNA polymerases, reverse transcriptases

Names and properties of DNA sequences encoding reverse transcriptase (retroelements) or resulting from it

Reverse transcriptase	Transposition or integration	Long terminal repeats	Virions	Name
Present	Absent	Absent	Absent	Retron
Present	Present	Absent	Absent	Retroposon or LINE element
Present	Present	Present	Absent	Retrotransposon
Present	Present	Present	Present	Retrovirus
Present	Absent	Present	Present	Pararetrovirus (hepadnavirus; caulimovirus)
Absent	Absent	Absent	Absent	Retrosequence (cDNA genes; *Alu* sequences)

require for DNA synthesis a primer with a free 3′-hydroxyl group, a template, a divalent cation (Mg^{2+} or Mn^{2+}), and all four deoxyribonucleoside triphosphates (**Fig. 3**). The direction of DNA synthesis for reverse transcriptases and all other DNA polymerases is from the 5′ to the 3′ end of the newly synthesized DNA. However, the retrovirus reverse transcriptases, in contrast to most cellular DNA polymerases, are able to utilize natural RNA templates for DNA synthesis as efficiently as DNA templates. Moreover, retrovirus reverse transcriptases do not seem to have strict requirements for specific RNA templates, unlike other RNA virus replicases or RNA-directed RNA polymerases, which usually recognize only their own genome as a template.

One consequence of this lack of strict template specificity of reverse transcriptases is that they can be used to make DNA copies of any RNA. This property of reverse transcriptase has become very useful in molecular biology, especially in genetic engineering. DNA copies of purified RNAs are made by reverse transcriptase, and then the DNA is cloned in bacteria. Several potentially important molecules have been produced in bacteria as a result of this cloning. *See* GENETIC ENGINEERING.

It has also been determined that a reverse transcriptase activity is involved in the formation of telomeres (chromosome ends). This is an essential cellular enzyme activity.

Mechanisms of synthesis of retrovirus DNA. Much has been learned about the detailed mechanism of synthesis of retroviral DNA. This mechanism depends upon the properties of the reverse transcriptase and ribonuclease H described above and properties of the viral RNA and DNA described in **Figs. 4** and **5**. The viral RNA contains a 15- to 250-base direct repeat (R) at its 5′ end next to the cap and at its 3′ end next to the poly (A) tail. At the 5′ end of the viral RNA, about 150–350 bases from the end is a sequence (PBS or PB) complementary to the eighteen 3′ nucleotides of a specific transfer RNA molecule which acts as a binding site for that tRNA. The tRNA serves as a primer for minus-strand DNA synthesis. The cellular tRNA is $tRNA^{trp}$ for avian retroviruses and $tRNA^{pro}$ for some mammalian retroviruses. In the viral DNA there are additional sequences outside the sequences that are directly repeated in the viral RNA such that the viral DNA has a direct repeat of 300–1200 base pairs, depending upon the species of virus. This direct repeat is called the long terminal repeat (LTR) and consists of sequences called U_3-R-U_5. The U_3 sequences are from the 3′ end of the RNA, and the U_5 sequences are from the 5′ end of viral RNA.

Figure 4 shows a possible mechanism for the synthesis of this DNA structure from viral RNA, making use of the RNA direct repeats, the tRNA primer, the ribonuclease H activity of the reverse transcriptase, movement of the DNA product from one end of the viral RNA to the other end of another viral RNA (there are two identical molecules of viral RNA in a retrovirus virion), and specific priming of the plus strand of the viral DNA, by a polypurine tract next to U_3. The steps in this model for the synthesis of retrovirus DNA are as follows: (*a*) Viral RNA. RU_5PB and U_3R represent two contiguous sequences at the 5′ and 3′ ends of the RNA, respectively. R is a terminal repeat in the RNA; PB (primer binding) is a sequence complementary to the end of

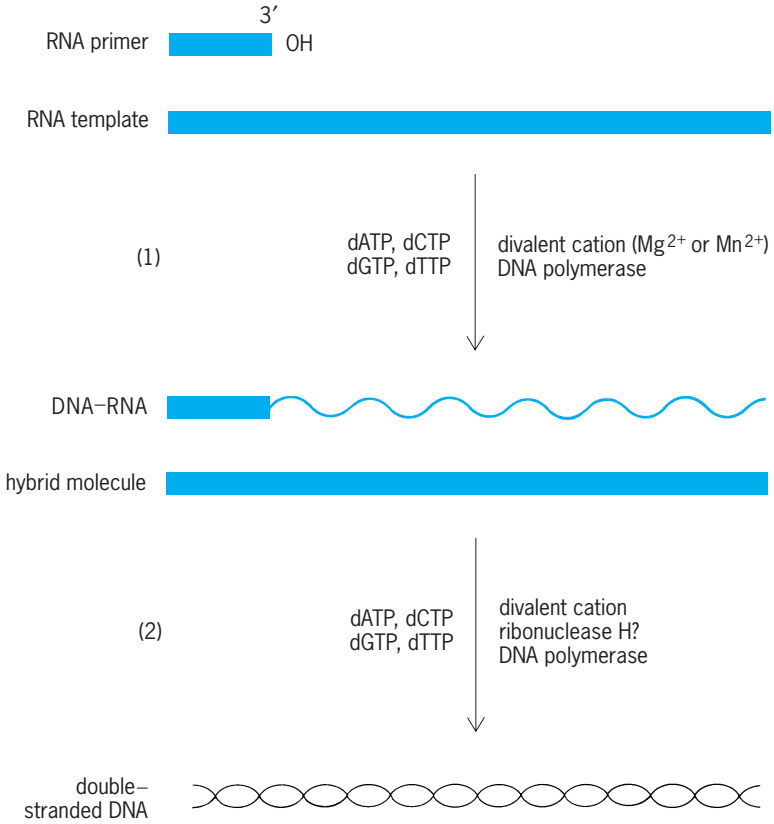

Fig. 3. Two reactions in retrovirus RNA-directed DNA synthesis. The viral RNA primer is actually attached close to the 5′ end of the viral RNA template.

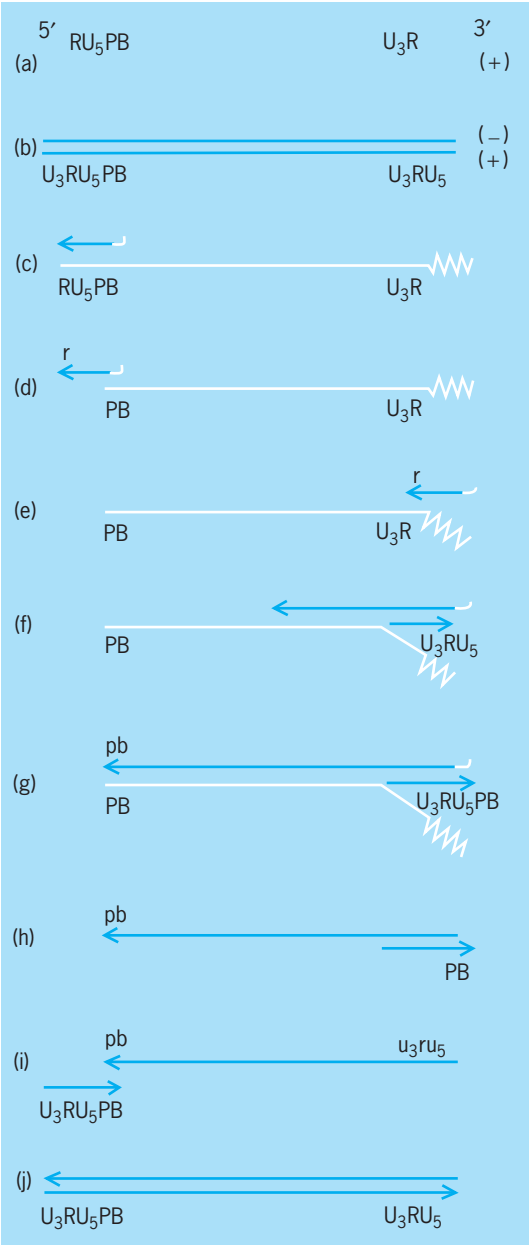

Fig. 4. Model for the synthesis of retrovirus DNA. Uppercase letters represent the plus strand; lowercase letters represent the minus strand, RNA is shown in white, DNA in color.

a cellular tRNA, the primer. (*b*) Unintegrated viral DNA. The steps in its synthesis from *a* are as follows: (*c*) Viral DNA synthesis begins with minus-strand synthesis using the tRNA as a primer and proceeding through U_5 and R to the 5′ end of the RNA. (*d*) Viral ribonuclease H digests the part of the RNA template (viral RNA) that is hydrogen-bonded to minus-strand DNA. (*e*) The newly synthesized minus-strand DNA hydrogen- bonds to the repeat (R) at the 3′ end of the other molecule of viral RNA. (*f*) Minus-strand DNA synthesis proceeds using viral RNA as a template. Plus-strand DNA synthesis is initiated by using ribonuclease H-cut viral RNA as a primer and minus-strand DNA as a template. (*g*) Minus- and plus-strand viral DNA synthesis proceed to the end of their respective templates. (*h*) Viral ribonuclease H further degrades the template RNAs, exposing homologous sequences, pb and PB on the minus- and plus-strand DNAs, respectively. (*i*) The plus-strand and minus-strand DNAs hydrogen-bond at PB. (*j*) Synthesis of the unintegrated viral DNA is completed.

Evolution. Studies indicate that reverse transcriptase is widely distributed in living organisms and that all reverse transcriptases are evolutionarily related. For example, the organization of the nucleotide sequence of integrated retroviral DNA described in Fig. 5 has a remarkable resemblance to the structure of bacterial transposable elements, in particular, transposons. Transposons have a five-base-pair repeat of bacterial DNA next to an inverted repeat of transposon DNA, which often is in a direct repeat of transposon DNA. The similarity of this structure to that of retroviruses indicates that these genetic elements are very similar and that it is not possible to determine from the structure of an integrated retrovirus or transposon whether this structure was formed as a result of reverse transcription or of transposition involving DNA only. *See* TRANSPOSONS.

These organizational similarities supported the hypothesis that retroviruses evolved from cellular movable genetic elements such as transposons. More direct support for this hypothesis has come from cloning and sequencing of reverse transcriptase genes and movable genetic elements from *Drosophila*, yeast, and many other organisms, including even myxobacteria (see table). Reverse transcriptase genes are present in the eukaryotic organisms in retrotransposons, which are DNA sequences with long terminal repeats (LTRs) and reverse transcriptase coding sequences, and in retroposons or long interspersed (LINE) elements, which are DNA sequences without LTRs but with reverse transcriptase coding sequences. Both of these types of elements can transpose in cells. In bacteria, coding sequences for reverse transcriptase are found as part of DNA sequences coding for unusual small molecules composed of both RNA and DNA and

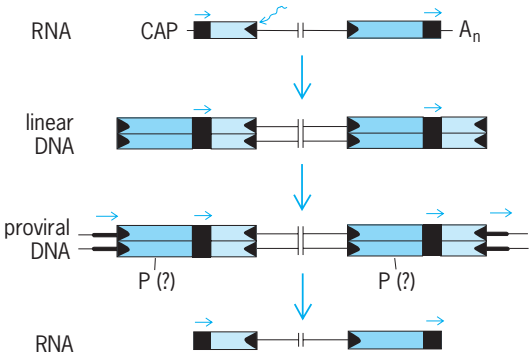

Fig. 5. Schematic representation of the structures at the termini of retroviral nucleic acid molecules. Solid areas represent terminal sequences at either terminus which are directly repeated in the RNA (as indicated by the arrows above these areas in the bottom RNA diagram). Shaded areas represent sequences unique to either end of the RNA. Heavy lines outside the long terminal repeats in the provirus represent direct repeats of cell DNA formed in the integration process.

called msDNA. The DNA in these molecules is synthesized by reverse transcriptase. The function of msDNA molecules is unknown. *See* DEOXYRIBONUCLEIC ACID (DNA); RIBONUCLEIC ACID (RNA).

Howard M. Temin

Bibliography. A. J. Davison and R. M. Elliott (eds.), *Molecular Virology*, 1993; R. F. Doolittle et al., Origins and evolutionary relationships of retroviruses, *Quart. Rev. Biol.*, 64:1–30, 1989; S. P. Goff, Retroviral reverse transcriptase: Synthesis, structure, and function, *JAIDS*, 3:817–831, 1990; D. Shippen-Lentz and E. H. Blackburn, Functional evidence for an RNA template in telomerase, *Science*, 247:546–552, 1990; H. M. Temin, Retrons in bacteria, *Nature*, 339:254–255, 1989; H. Varmus, Retroviruses, *Science*, 240:1427–1435, 1988; R. Weiss et al. (eds.), *RNA Tumor Viruses: Molecular Biology of Tumor Viruses*, 2d ed., Cold Spring Harbor (New York) Laboratory, 1982.

Revetment

A facing or veneer of stone, concrete, or other materials constructed on a sloping embankment, dike, or beach face to protect it against erosion caused by waves or currents. The revetment may be a rigid cast-in-place concrete structure; but more commonly it is a flexible structure constructed of stone riprap or interlocking concrete blocks. It is sometimes an articulated block structure where the armor blocks are set in a form known as a flexible carpet; that is, the blocks interlock for stability, but the interlocking makes them flexible enough to respond to settlement of the underlying soil. A flexible revetment provides protection from exterior hydraulic forces, and it also can tolerate some settlement or consolidation of the underlying soil.

Components. A typical revetment might employ stone riprap as the armor material (see **illus.**). A revetment typically has three major components: (1) the armor layer, which resists the wave or current-induced hydraulic forces; (2) a filter layer under the armor layer to allow water seepage out of the underlying soil without the removal of fine soil particles; and (3) a mechanism to stabilize the structure toe. Toe stabilization is particularly important where waves break on the structure, but may not be necessary if the revetment extends to sufficient depths where hydraulic forces will not erode the toe of the slope. The design water level (see illus.) for the structure may be higher than the normal water level during nonstorm conditions. If the revetment is exposed to waves that will break and run up the face of the revetment, the upper extent of the revetment must be sufficiently high to counter the force exerted by the waves.

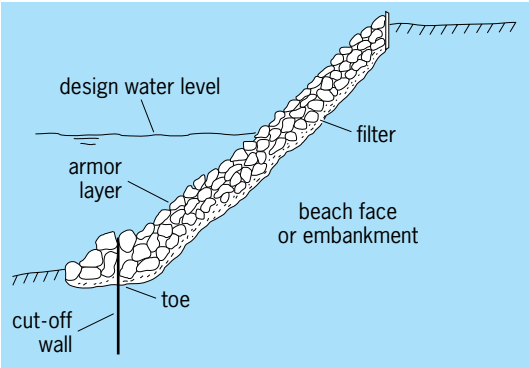

Cross-sectional profile of a typical stone revetment.

Materials. Although stone riprap is the most commonly used material for revetment armor layers, a wide variety of other materials have been used, including cast-in-place concrete and poured asphalt, wire bags filled with stone (gabions), interlocking concrete blocks, soil cement, cement-filled bags, interlocked tires, woven wooden mattresses, and vegetation. Vegetation would only be used for surfaces exposed to very low waves or slow-moving currents.

A rough porous armor layer with a surface that is as uniform as possible has advantages when the revetment is exposed to wave action. This surface condition will decrease the elevation of wave runup on the revetment face and thus allow for a lower structure for the top elevation. Armor units that protrude above the surface of the revetment are exposed to stronger hydraulic forces and thus are more prone to removal from the structure, leading to possible structure failure.

An advantage that a revetment with a stone riprap armor layer has over an armor layer consisting of interlocking concrete units is that it is self-healing to some extent. If a few armor stones are dislodged, others will shift into the vacated position. Stone revetments also are generally more esthetically pleasing as they are more natural in appearance.

Design geometry. Typical reveted slopes vary in vertical to horizontal proportion between 1:2 and 1:4. A steeper slope requires less material for the revetment. However, the armor layer will be more stable on a flatter slope, as will the underlying soil embankment. Other considerations such as revetment constructability, slope accessibility for maintenance and repair, and movement over the revetment to access a beach may also affect the design geometry.

Filter layer. The filter layer typically consists of one or more layers of finer stone or a woven plastic fabric or a combination of both. While the filter must be sufficiently fine to prevent embankment material from being pulled through the armor layer, it must not be so fine that there is a buildup of excessive pore water pressure in the embankment. Stone filter layers also distribute the load from larger armor stone in the armor layer. *See* GEOSYNTHETIC.

Stabilization. The revetment toe may need to be stabilized. This can be accomplished by maintaining sufficient soil or beach sand at the toe. This would involve placing sand fill with or without groins at a wave-exposed shore or constructing spur dikes perpendicular to the bank of a river to trap river sediment. Or the toe may be stabilized by placement of a cut-off wall (see illus.) or larger armor stone that prevents the toe from being undermined. If the

revetment toe fails, the entire structure is prone to failure.

At some point along the river or shoreline, the revetment must terminate or merge with other structures. This point is often critical to the stability of the entire structure. If the revetment just ends at a natural unprotected section of embankment, the revetment must be stabilized to allow for the possible erosion of the adjacent natural embankment without undermining the revetment. This stabilization may involve the placement of a more massive segment of revetment or the construction of a cut-off wall perpendicular to the embankment surface that extends well back into the embankment. *See* RETAINING WALL; RIVER ENGINEERING.

R. M. Sorensen

Bibliography. M. S. Peterson, *River Engineering*, 1986; L. L. Whiteneck and L. A. Hockney, *Structural Materials for Harbor and Coastal Construction*, 1989.

Reynolds number

The number Re, a dimensionless parameter that determines the behavior and characteristics of viscous flow patterns. It is defined by Eq. (1), where ρ is

$$\mathrm{Re} = \frac{\rho V L}{\mu} \tag{1}$$

fluid density, V is stream velocity, L is a characteristic length scale, and μ is fluid viscosity. This parameter, the dominant factor in viscous flow analysis, was formulated in 1883 by Osborne Reynolds and named in his honor about 40 years later.

The form of Eq. (1) is such that Re has the same value regardless of the unit system used to define its constitutents. In the International System (SI), for example, if brackets denote the dimensions of a variable, Eq. (2) is obtained. Thus the Reynolds number

$$[\mathrm{Re}] = \frac{[\mathrm{kg/m^3}][\mathrm{m/s}][\mathrm{m}]}{[\mathrm{kg/m \cdot s}]} = [1] \tag{2}$$

is not subject to any ambiguity.

Laminar and turbulent flow. It had been well known since the eighteenth century that viscous flows were sometimes smooth (laminar) and sometimes disorderly (turbulent). Reynolds revealed the reason in 1883 in a classic experiment which introduced a dye streak down the centerline of water flowing in a glass tube (see **illus.**). For values of Re less than 2000 (illus. *a*) the dye streak remained smooth and undisturbed—that is, laminar flow resulted. For Re greater than 3000 (illus. *b*), the dye streak seemed to burst into turbulence and fill the whole tube with color. However, a spark-flash photograph (illus. *c*) showed that the turbulence structure was really a series of minute whirls and eddies, oscillating too fast for the eye to see. This classic demonstration laid the foundation for what is now a rich field of turbulent flow analysis and experimentation. *See* LAMINAR FLOW; TURBULENT FLOW.

For Re between 2000 and 3000 (not shown in the illustration), the flow pattern cycled erratically be-

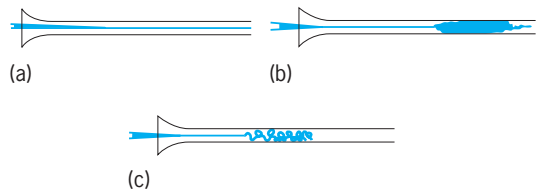

Reynolds's original sketches of dye streak behavior in pipe flow: (*a*) laminar flow, Re > 2000; (*b*) turbulent flow, Re > 3000; (*c*) a spark-flash photograph of condition *b*. (*After O. Reynolds, An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous and of the law of resistance in parallel channels, Phil. Trans. Roy. Soc., London, A174:935–982, 1883*)

tween laminar and turbulent regimes. This is now known as the transition region between the two more stable flow patterns.

The transition Reynolds number is not always equal to 2000 but rather varies somewhat with flow geometry. It is less than 2000 for unstable patterns, such as jet flow, and greater than 2000 for highly stable flows such as converging nozzles. *See* BOUNDARY-LAYER FLOW; JET FLOW.

Model testing. In experimenting with model tests of viscous flows, the Reynolds number is the primary parameter by which the model and prototype are matched. Fulfillment of Eq. (3) ensures the dynamic

$$\left(\frac{\rho V L}{\mu}\right)_{\text{model}} = \left(\frac{\rho V L}{\mu}\right)_{\text{prototype}} \tag{3}$$

similarity of the model test. One can vary density, velocity, length, or viscosity in any manner as long as the overall grouping ($\rho VL/\mu$) remains the same. Of course, the model must also be geometrically similar to the prototype, that is, must have exactly the same proportions and orientation to the fluid stream. Then model test data can be used to predict the drag, pressure distribution, friction, and even the convection heat transfer from the prototype, if the data are properly interpreted. An exception is free surface flow—such as ship resistance, ocean wave motion, and buoy dynamics—where the Froude number of the flow is also extremely important. *See* DIMENSIONAL ANALYSIS; DYNAMIC SIMILARITY; FLUID FLOW; FROUDE NUMBER; MODEL THEORY.

Frank M. White

Bibliography. H. Rouse and S. Ince, *History of Hydraulics*, 1963; H. Schlichting, *Boundary Layer Theory*, 8th ed., 2000; I. H. Shames, *Mechanics of Fluids*, 3d ed., 1992; F. M. White, *Fluid Mechanics*, 4th ed., 1998.

Rh incompatibility

A condition in which red blood cells of the fetus become coated with an immunoglobulin (IgG) antibody [anti-Rh D antibody] of maternal origin which is directed against an antigen (Rh D antigen) of paternal origin that is present on fetal cells. *See* ANTIBODY; ANTIGEN; IMMUNOGLOBULIN.

Etiology. A pregnant woman may develop an antibody to a red blood cell antigen that the fetus has but

that she does not possess. This occurs because the fetus has inherited the antigen from the father; fetal red blood cells pass into the maternal circulation and are recognized as foreign by the mother. The mother develops antibodies to the foreign protein; the antibodies pass across the placenta and bind to the fetal red blood cells. The fetal mononuclear phagocytic system recognizes the antibody-sensitized red blood cells as foreign and removes them from the circulation producing a hemolytic anemia. The maternal immune response is known as red blood cell alloimmunization. Alloimmunization of the mother results in hemolytic disease of the fetus and newborn (also known as erythroblastosis fetalis). It is important to understand that hemolytic disease of the fetus and newborn can occur from an incompatibility to a number of blood group antigens. The incidence of Rh incompatibility leading to hemolytic disease of the newborn has decreased dramatically from the development of passive immunization to prevent maternal Rh D alloimmunization. *See* ANEMIA; AUTOIMMUNITY.

Blood group nomenclature. Rh antigens are essential to maintain structural stability of the red cell membrane and are composed of protein and lipid (proteolipid). The Rh blood group system has a two gene loci (RHD and RHCE), and a combination of Rh alleles is inherited from each parent. The Rh antigen of greatest importance is the D antigen. If both D genes are inherited, the individual is said to be homozygous Rh positive. If one D gene is inherited, the blood type will be Rh positive and the individual is said to be hemizygous due to the deletion of one D gene from one chromosome. If neither D gene is present, the individual will be Rh negative. Total absence of both the RHD and RHCE genes results in the condition known as Rh null; it is very rare. *See* BLOOD GROUPS.

Risk of sensitization. The development of Rh incompatibility is determined by the Rh types of the parents. If the mother is Rh positive, there is no risk of Rh D incompatibility. If the mother is Rh negative, the risk to the fetus depends upon its blood type, which in turn is related to the blood type of the father. If the father is Rh negative, there is no risk to the fetus because all fetuses will be Rh negative. If the father is hemizygous Rh positive, half of the offspring will be Rh positive and, following sensitization of the mother, will be in danger of developing erythroblastosis. If the father is homozygous Rh positive, all fetuses will be Rh positive; if the mother has been sensitized after the first pregnancy, there will be the risk of hemolytic disease of the fetus and newborn for each subsequent pregnancy. The risk of alloimmunization to Rh is 16% with each ABO-compatible pregnancy and 1.5–2% with each Rh-positive ABO-incompatible pregnancy.

Some Rh-incompatible pregnancies do not result in sensitization. Factors influencing the occurrence of sensitization include the timing and extent of the transplacental hemorrhage, the degree and strength of antibody development in the mother, and the ABO status of mother and fetus. Most immunizations result from transplacental hemorrhage during placental separation at delivery. Smaller numbers of cells enter the mother's circulation during the last half of pregnancy. If the amount of cells is less than 0.1 milliliter throughout and after delivery, sensitization is less likely. ABO-incompatible Rh-incompatible fetal red blood cells are more easily recognized as foreign, due to preexisting ABO antibodies, so that the fetal red blood cells are rapidly destroyed in the maternal circulation and are less likely to elicit an antibody response. Because most sensitization occurs as a result of bleeding at delivery, hemolytic disease of the newborn almost never occurs during the first pregnancy, but subsequent Rh-positive pregnancies may result in an affected infant. Once immunization has occurred, successive Rh-positive pregnancies will result in hemolytic disease of the fetus and newborn of increasing severity. Because Rh negativity is less common in Blacks and is rare in Asians, Rh incompatibility is less likely to occur in these groups.

Pathogenesis. The symptoms and signs of hemolytic disease of the newborn from Rh incompatibility are due to the destruction (hemolysis) of the antibody-coated fetal red blood cells. The fetus responds to the hemolysis and anemia in a number of different ways. The products of hemolysis, specifically bilirubin, pass through the placenta and are cleared by the maternal circulation. The fetus compensates for the anemia by producing more red blood cells. The liver and spleen become the major sites of red blood cell production, and this is known as extramedullary hematopoiesis, the formation of blood cells outside the bone marrow. As the liver and spleen enlarge, their function decreases and serum protein levels fall. This results in anasarca, or body edema; hence the term hydrops fetalis—literally, water baby. Anemia results in heart failure, which contributes to the hydrops. In the severest forms, infants are stillborn or die soon after birth. Another major manifestation is jaundice, caused by increased amounts of the blood pigment bilirubin, derived from the breakdown of hemoglobin. Although much of the bilirubin passes across the placenta and is metabolized by the mother's liver, the newborn still may be perceptibly jaundiced, and this jaundice rapidly deepens after birth because the placenta is no longer present to clear the bilirubin. In less severe cases, the infant may be in good condition at birth but become jaundiced as the bilirubin increases in concentration. Indirect bilirubin that is not bound to albumin may cross the blood–brain barrier and affect certain brain cells, resulting in permanent brain damage. This condition of bilirubin encephalopathy is termed kernicterus. *See* BILIRUBIN; HEMATOPOIESIS; HEMOGLOBIN; JAUNDICE.

Diagnosis and management. The at-risk mother is monitored for the development and progression of Rh incompatibility with serial serum antibody titers to anti-D. The presence of anti-D and its titer do not necessarily correlate with the status of the fetus. To monitor the status, a procedure known as amniocentesis is performed in which amniotic fluid is collected by introducing a needle into the amniotic sac through the mother's abdominal wall. This

procedure permits measurement of bile pigment concentration and assessment of lung maturity. These tests are done serially to decide when to deliver the fetus or whether to transfuse the fetus until it is mature enough to be delivered safely.

Because invasive procedures to obtain fetal blood for prenatal blood group analysis may present a risk to the fetus, other methods have been explored. During pregnancy, cell-free fetal deoxyribonucleic acid (DNA) circulates in maternal blood plasma. Molecular techniques, especially real-time polymerase chain reaction, or RT-PCR, can be used to identify copies of the product of the RHD gene and to accurately predict fetal D status because the genome of the Rh-negative mother does not contain an RHD gene. In another technique, Doppler assessment of fetal middle cerebral artery blood flow is used as a reliable noninvasive predictor of fetal anemia. Peak systolic velocity increases with decreased red cell volume (that is, anemia), which is assessed in the proximal region of the artery identified by cross-sectional ultrasonography. *See* PRENATAL DIAGNOSIS.

Transfusions. Treatment of the fetus and infant depends upon the severity of the disease. Intrauterine transfusions can be administered as early as 18 weeks' gestation and are given by introducing a needle into the fetal abdominal cavity or by catheterizing a fetal umbilical vein, using ultrasound to guide the needle placement. This procedure is called PUBS (percutaneous umbilical blood sampling); both sampling and intrauterine transfusions can be administered through PUBS. Such transfusions may be done weekly or bimonthly. The blood used for the transfusion is washed, type-O, Rh-negative, cytomegalovirus (CMV) "safe" blood, and it may be specially processed in other ways.

Following delivery, the infant may require transfusion immediately or during the postnatal period. This may involve simple replacement or exchange transfusion, in which the infant's blood volume is replaced with compatible blood. The aim of exchange is to remove bilirubin and antibody-coated red blood cells and to correct the anemia while replacing the infant's blood with blood that is antigen negative for the offending antibody. Other medical treatment for jaundice may also be required. *See* TRANSFUSION.

Rh immune globulin. Rh immune globulin (RhIG) is a purified solution of IgG containing anti-D obtained from pooled human plasma of deliberately alloimmunized donors or previously pregnant women alloimmunized to the Rh D antigen. A 1-ml dose contains 300 micrograms of RhIG, and it will counteract the immunizing effects of 30 ml of Rh-positive fetal red blood cells. The mechanism of action is not well understood but is thought to interfere with antigen recognition by the maternal immune system. Because of its method of preparation, it will not transmit hepatitis or human immunodeficiency virus (HIV) infection. Special tests can be performed on the mother's blood after delivery to quantify the amount of transplacental hemorrhage. RhIG should be administered to any Rh-negative mother of an Rh-positive infant within 72 h after delivery. In cases where there is bleeding of more than 30 ml of fetal cells, additional RhIG is usually administered. Administration of RhIG at the twenty-eighth week of pregnancy followed by a second dose at birth has been very successful in all but eliminating Rh hemolytic disease. Anti-Rh immunoglobulin should also be administered in a lower dose to Rh-negative women following amniocentesis, abortion, ectopic pregnancy, and fetal death if the status of the father is Rh positive or unknown. If the pregnancy is terminated after 13 weeks of gestation, a full dose (300 mg) should be administered. *See* PREGNANCY.

Naomi L. C. Luban; Chrysanthe Gaitatzes; Aimee M. Barton

Bibliography. J. M. Bowman, The prevention of Rh immunization, *Trans. Med. Rev.*, 2:129, 1988; P. A. Grannum and J. A. Copel, Prevention of Rh isoimmunization and treatment of the compromised fetus, *Seminar Perinatol.*, 12:324, 1988; K. H. Nicolaides, Studies on fetal physiology and pathophysiology in Rhesus disease, *Seminar Perinatol.*, 13:328, 1989.

Rhabditida

An order of nematodes in which the number of labia varies from a full complement of six to three or two or none. The tubular stoma may be composed of five or more sections called rhabdions. The three-part esophagus always ends in a muscular bulb that is invariably valved. The excretory tube is cuticularly lined, and paired lateral collecting tubes generally run posteriorly from the excretory cell; some taxa have anterior tubules also. Females have one or two ovaries; when only one is present the vulva shifts posteriorly. The cells of the intestine may be uninucleate, binucleate, or tetranucleate, and the hypodermal cells may also be multinucleate. Caudal alae, when present, contain papillae.

There are eight superfamilies in the order: Rhabditoidea, Alloionematoidea, Bunonematoidea, Cephaloboidea, Panagrolaimoidea, Robertioidea, Chambersiellolidea, and Elaphonematoidea. Of these, the Rhabditoidea not only is one of the largest nematode superfamilies but also contains many important parasites of humans and domestic animals. The group is distinguished by the well-developed cylindrical stoma and three-part esophagus that ends in a valved terminal bulb. In parasitic species, adult stages and some larval stages lack the valved terminal bulb. Though most species are free-living feeders on terrestrial bacteria, others are important in the biological control of insects or as parasites of mammals.

There are eight superfamilies in the order: Rhabditoidea, Alloionematoidea, Bunonematoidea, Cephaloboidea, Panagrolaimoidea, Robertioidea, Chambersiellolidea, and Elaphonematoidea. Of these, the Rhabditoidea not only is one of the largest nematode superfamilies but also contains many important parasites of humans and domestic animals. The group is distinguished by the well-developed cylindrical stoma and three-part esophagus that ends in a valved terminal bulb (see **illustration**). In

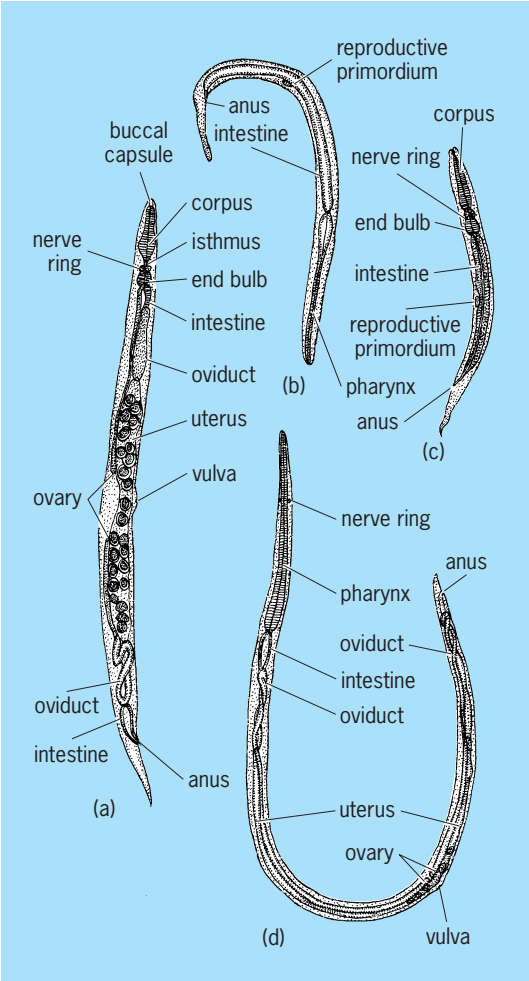

***Strongyloides ransomi*. (*a*) Free-living female, 1 mm long. (*b*) Filariform young. (*c*) Rhabdiform young. (*d*) Parasitic female, 4 mm long. (*After L. H. Hyman, The Invertebrates, vol. 3, McGraw-Hill, 1951*)**

parasitic species, adult stages and some larval stages lack the valved terminal bulb. Though most species are free-living feeders on terrestrial bacteria, others are important in the biological control of insects or as parasites of mammals.

Neoaplectana, a genus whose known species are vectors of bacterial insect diseases, are of great interest in biological control schemes. A common bacterium transmitted is *Xenorhabdus nematophilus*. This bacterium is carried from generation to generation of nematodes in a pouch or expanded region at the anterior end of the nematode's intestine. The nematode, upon penetrating into the insect's hemocoel, ejects the bacteria from its alimentary canal. Once released, the bacteria multiply rapidly in the insect, killing it in less than 48 h. The nematode undergoes no development until this point; then it begins to feed on the bacteria, and development proceeds to adulthood. The nematode may go through several generations; eventually, when food is depleted, development is arrested in the infective stage, which once agains retains a bacterial bolus. These infective larvae remain in this state until another insect is invaded.

Strongyloides stercoralis, commonly called the threadworm, is an important parasite of humans. It can sustain itself as either a free-living bacterial feeder or as a parasite. In the free-living state it is a fecal associate feeding on bacteria. When the fecal substrate or bacteria are exhausted, the nematode arrests its development in the third (infective) stage. Morphologically it is distinguished by its lack of the valved rhabditoid esophagus. *Strongyloides stercoralis* enters the definitive host (humans) by direct skin penetration. Upon penetration it makes its way through the circulatory system to the lungs; from the lungs it is coughed up and swallowed. Upon reaching the small intestine it develops to adulthood. The eggs hatch in the intestine, and the released larvae may remain rhabditiform (free living) or metamorphose to the infective filariform that can immediately reinfect the same victim or be passed out to seek new victims. The rhabditiform larvae, when passed out with the feces, begin the free-living cycle. Human symptoms may include itching, from skin penetration by larvae; pneumonia, from lung penetration; or anemia, from blood loss due to adult penetration into the intestinal mucosa. *See* MEDICAL PARASITOLOGY; NEMATA.

Armand R. Maggenti

Rhenium

A chemical element, Re, with atomic number 75 and atomic weight 186.2. Rhenium is a transition element. It is a dense metal (21.04) with the very high melting point of 3440°C (6220°F). *See* PERIODIC TABLE.

Rhenium is similar to its homolog technetium in that it may be oxidized at elevated temperatures by oxygen to form the volatile heptoxide, Re_2O_7; this in turn may be reduced to a lower oxide, ReO_2. The compounds ReO_3, Re_2O_3, and Re_2O, are well known. Perrhenic acid, $HReO_4$, is a strong monobasic acid and is only a very weak oxidizing agent. Complex perrhenates, such as cobalt hexammine perrhenate, $[Co(NH_3)_6(ReO_4)_3]$, are also known.

The halogen compounds of rhenium are very complicated, and a large series of halides and oxyhalides have been reported. Rhenium forms two well-characterized sulfides, Re_2S_7 and ReS_2, as well as two selenides, Re_2Se_7 and $ReSe_2$. The sulfides have their counterparts in the technetium compounds, Tc_2S_7

1												13	14	15	16	17	18
1 H	2																2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

and TcS_2. *See* TECHNETIUM; TRANSITION ELEMENTS.

Sherman Fried

Bibliography. R. Alberto, *Technetium Rhenium*, vol. 176 in *Topics in Current Chemistry*, Springer, 1995; R. Albrecht, *Re Rhenium-Organorhenium Compounds: Binuclear Compounds* (*Gmelin Handbooks of Inorganic and Organometallic Chemistry*, pt. 7), 8th ed., Springer, 1997; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004.

Rheology

In the broadest sense of the term, that part of mechanics which deals with the relation between force and deformation in material bodies. The nature of this relation depends on the material of which the body is constituted. It is customary to represent the deformation behavior of metals and other solids by a model called the linear or hookean elastic solid (displaying the property known as elasticity) and that of fluids by the model of the linear viscous or newtonian fluid (displaying the property known as viscosity). These classical models are, however, inadequate to depict certain nonlinear and time-dependent deformation behavior that is sometimes observed. It is these nonclassical behaviors which are the chief interest of rheologists and hence referred to as rheological behavior. *See* ELASTICITY; VISCOSITY.

Rheological behavior is particularly readily observed in materials containing polymer molecules which typically contain thousands of atoms per molecule, although such properties are also exhibited in some experiments on metals, glasses, and gases. Thus rheology is of interest not only to mathematicians and physicists, who consider it to be a part of continuum mechanics, but also to chemists and engineers who have to deal with these materials. It is of special importance in the plastics, rubber, film, and coatings industries. *See* FLUID MECHANICS; PAINT; PLASTICS PROCESSING; POLYMER; RUBBER; SURFACE COATING.

This article deals with three useful nonclassical models: linear viscoelasticity, nonlinear elasticity, and nonlinear viscoelasticity. Plasticity, of interest in metals, is rarely studied by rheologists. *See* PLASTIC DEFORMATION OF METAL; PLASTICITY.

Models and properties. To make clear the nature of the various rheological models, the discussion will focus on some very simple types of deformation. Simple shear will be discussed in most detail; similar phenomena and properties, for the most part, also occur in extension. *See* STRESS AND STRAIN.

Consider a block of material of height h deformed in the manner indicated in **Fig. 1**; the bottom surface is fixed and the top moves a distance w parallel to itself. A measure of the deformation is the shear strain γ given by Eq. (1).

$$\gamma = \frac{w}{h} \tag{1}$$

To achieve such a deformation if the block is a linear elastic material, it is necessary to apply uniformly distributed tangential forces on the top and bottom of the block as shown in Fig. 1*b*. The intensity of these forces, that is, the magnitude of the net force per unit area, is called the shear stress S. For a linear elastic material, γ is much less than unity and is related to S by Eq. (2), where the proportionality

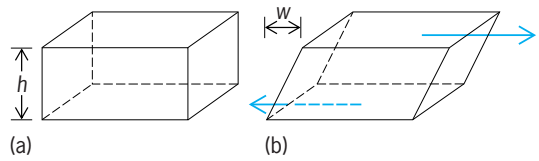

Fig. 1. Simple shear. (*a*) Undeformed block of height *h*. (*b*) Deformed block after top has moved a distance *w* parallel to itself. The arrows indicate the net forces acting on the top and bottom faces. The forces which must be applied to left and right faces to maintain a steady state are not indicated.

$$S = G\gamma \tag{2}$$

constant G is a property of the material known as the shear modulus.

If the material in the block is a newtonian fluid and a similar set of forces is imposed, the result is a simple shearing flow, a deformation as pictured in Fig. 1*b* with the top surface moving with a velocity dw/dt. This type of motion is characterized by a rate of shear $\dot{\gamma} = (dw/h)/dt$, which is proportional to the shear stress S as given by Eq. (3), where η is a

$$S = \eta\dot{\gamma} \tag{3}$$

property of the material called the viscosity.

Linear viscoelasticity. If the imposed forces are small enough, time-dependent deformation behavior can often be described by the model of linear viscoelasticity. The material properties in this model are most easily specified in terms of simple experiments: creep, stress relaxation, and sinusoidal deformation.

Creep and relaxation. In a creep experiment a stress is suddenly applied and then held constant; the deformation is then followed as a function of time. This stress history is indicated in the solid line of **Fig. 2***a* for the case of an applied constant shear stress S_0. If such an experiment is performed on a linear elastic solid, the resultant deformation is indicated by the full line in Fig. 2*b* and for the linear viscous fluid in Fig. 2*c*. In the case of elasticity, the result is an instantly achieved constant strain; in the case of the fluid, an instantly achieved constant rate of strain. In the case of viscoelastic materials, there are some which eventually attain a constant equilibrium strain (Fig. 2*d*) and hence are called viscoelastic solids. Others eventually achieve constant rate of strain (Fig. 2*e*) and are called viscoelastic fluids. If the material is linear viscoelastic, the deformation $\gamma(S_0,t)$ is a function of the time t since the stress was applied and also a linear function of S_0; that is, Eq. (4) is satisfied,

$$\gamma(S_0, t) = S_0 J(t) \tag{4}$$

where $J(t)$ is independent of S_0. The function $J(t)$ is

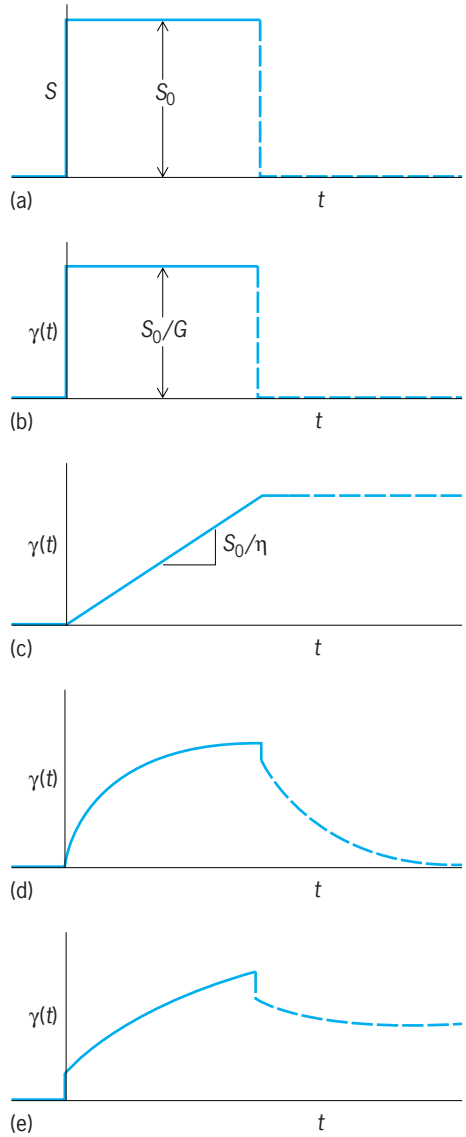

Fig. 2. Creep and recovery; solid lines indicate creep; broken lines indicate recovery. (*a*) Applied stress history. (*b*) Corresponding strain history for linear elastic solid, (*c*) linear viscous fluid, (*d*) viscoelastic solid, and (*e*) viscoelastic fluid.

a property of the material known as the shear creep compliance. Similar results are obtained for creep experiments performed in extension. *See* CREEP (MATERIALS).

The broken lines in Fig. 2 indicate the result when the stress is removed after a creep experiment. This is called a recovery experiment. The linear viscoelastic solid eventually goes back to its original undeformed shape while the fluid does not. However, the linear viscoelastic fluid recovers somewhat. The amount of recovery is taken as a measure of the elastic character of the fluid. Some of the energy which was expended on the fluid during the creep experiment is dissipated, but some is recovered. In contrast, with the newtonian fluid all the energy is dissipated and is converted to heat.

Stress relaxation. In a stress relaxation experiment, a deformation is suddenly imposed and then held constant; the stress is observed as a function of time. The stress decreases with time until it reaches an equilibrium value. This equilibrium value is zero for a viscoelastic fluid, and some finite value for a viscoelastic solid. If the imposed deformation is a simple shear with a constant shear strain γ_0, the required shear stress $S(t,\gamma_0)$ depends on the time t since the strain was imposed, and is proportional to γ_0; that is, Eq. (5) is satisfied, where $G(t)$ is a property of the

$$S(t, \gamma_0) = \gamma_0 G(t) \tag{5}$$

material, called the shear relaxation modulus, which does not depend on the strain. If the experiment performed is an extension, the corresponding property is the tensile relaxation modulus $E(t)$. Both $G(t)$ and $E(t)$ are functions which decrease as t increases.

Sinusoidal deformation. Linear viscoelasticity is useful for describing experiments where the deformation is a sinusoidal function of time. For example, if a sound wave (or other mechanical vibration) is transmitted through a thick sheet of a polymeric solid, the amplitude of the oscillation decreases as the wave progresses, and the speed of propagation and the rate of damping depend on the frequency of the oscillation and on the temperature. There usually is at least one frequency at which the rate of damping has a maximum.

Behavior of linear polymers. To illustrate some of the characteristics of linear viscoelastic properties, **Fig. 3** shows a schematic logarithmic plot of $J(t)$ for a high-molecular-weight polystyrene over a range of temperatures. This set of curves is typical of high-molecular-weight linear amorphous polymers, which include the commercial transparent thermoplastics such as poly(methyl methacrylate) (known under the trademarks Lucite and Plexiglas). These polymers are incapable of crystallization. However, they have a characteristic temperature, the glass transition temperature T_g, where the thermal expansion coefficient and the specific heat undergo a sudden change. Below T_g it takes an appreciable time to achieve an equilibrium volume if the material is subjected to a sudden decrease of temperature or a sudden change of pressure; the material is called

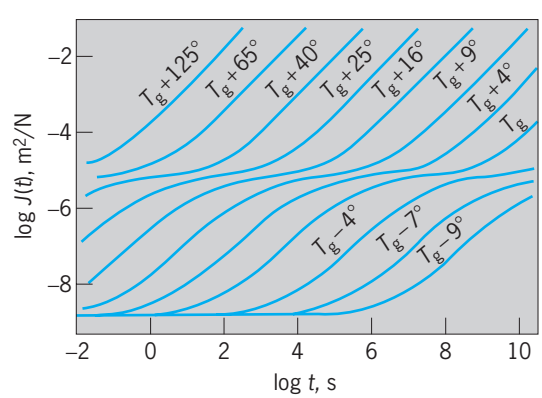

Fig. 3. Schematic shear creep compliance $J(t)$ and its temperature dependence for a high-molecular-weight polystyrene. Temperatures are in degrees Celsius (1°C = 1.8°F) above or below the glass transition temperature T_g, which is 105°C (221°F).

a glass. At temperatures appreciably below T_g the polymer behaves like a hard, brittle elastic solid. *See* GLASS; GLASS TRANSITION; POLYACRYLATE RESIN; POLYSTYRENE RESIN.

In Fig. 3, it is seen that the main effect of temperature is to shift the curve along the log time axis. When a viscoelastic material has this property, the material is thermorheologically simple. Increasing the temperature does not change the processes which are occurring; it just makes them happen faster.

At long times, $J(t)$ is essentially proportional to time (the slope is unity in a logarithmic plot) and the polymer behaves much like a newtonian fluid. At temperatures considerably above T_g, this newtonian behavior begins at times less than a second. At such temperatures, the fluid polymer can be readily shaped by forcing it into a mold or extruding it through an orifice of the desired cross section. If the polymer is then cooled considerably below T_g, it becomes a stiff solid with very little creep over a time scale of hours or days. It behaves almost like an elastic solid with a $J(t)$ of about 10^{-9} m^2/N, which corresponds to a Young's modulus of about 3×10^9 N/m^2 (4×10^5 lb/in.^{2}). For critical and some long-term uses, even this small amount of creep may not be tolerable. Addition of solid fillers (for example, glass fibers) or replacement by a highly crystalline polymer (for example, high-density polyethylene) can reduce the amount of creep. *See* COMPOSITE MATERIAL; POLYOLEFIN RESINS.

It is in the neighborhood of T_g that viscoelastic effects are most readily evident and where they are most sensitive to temperature for experiments in ordinary laboratory time scales.

Polymers have hundreds or thousands of atoms chemically bonded to one another in long sequences. In most synthetic polymers, the presence of single C-C (or similar) bonds makes rotation about these bonds possible. As a result, such macromolecules can have many geometric shapes, most of them coiled conformations—like exceedingly long, thin coiled-up snakes. In an amorphous polymer, these snakelike molecules are highly intertwined. When forces are applied, they cannot immediately attain the steady-state response. Increasing the temperature does not change the process; it can, however, speed it up.

Nonlinear elasticity. A strip of lightly vulcanized rubber (for example, a rubber band) can be stretched to many times its undeformed length by application of a tensile force without breaking; when the force is removed, the strip returns to its original length. The equilibrium force-deformation curve (**Fig. 4**) is clearly nonlinear. Linear elasticity is obviously not applicable, and the model of nonlinear elasticity must be employed.

If a torque is applied to the ends of a rod, the rod is simply twisted if it is a linear elastic material. However, if the rod is made of rubber, it also increases in length. This normal stress effect is a natural consequence of nonlinear elasticity.

Addition of carbon black to a rubber increases the stiffness (Fig. 4). Such a rubber compound is called filled or reinforced rubber. After stretching a strip of reinforced rubber for the first time and then removing the tensile force, the rubber is less stiff when it is stretched again. This is known as the Mullins effect and is attributed to the slippage and detachment of rubber molecules from the surface of the carbon black.

Because such high recoverable deformations are possible, a great deal of energy can be stored in rubber. This property has been used not only in toys (slingshots and motors for model airplanes) but also in supports for automobile bumpers designed so that they will absorb the impact of a low-velocity collision without damage to the vehicle.

Nonlinear viscoelasticity. If stresses become too high, linear viscoelasticity is no longer an adequate model for materials which exhibit time-dependent behavior. In a creep experiment, for example, the ratio of the strain to stress, $\gamma(t,S_0)/S_0$, is no longer independent of S_0; this ratio generally decreases with increasing S_0.

An important consequence of this nonlinearity occurs in the steady flow behavior of nonlinear viscoelastic fluids. In a steady simple shearing flow, the shear rate is not proportional to the shear stress S. In analogy with the case of the newtonian fluid [Eq. (3)], the ratio $S/\dot{\gamma}$ is also used to characterize the fluid in this nonlinear behavior, as in Eq. (6), and the

$$\frac{S}{\dot{\gamma}} = \eta(\dot{\gamma}) \tag{6}$$

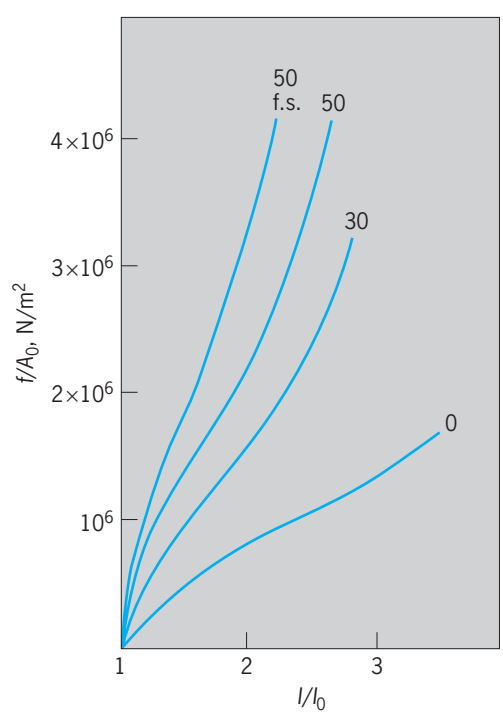

Fig. 4. Force-elongation data for a vulcanized synthetic rubber containing various amounts of carbon black (the number on the graph indicates parts of carbon black per 100 parts of rubber). The ordinate is the force f per initial cross-sectional area A_0. The abscissa is the ratio of the stretched length l to undeformed length l_0. The designation f.s. on one curve indicates that these data were obtained on the first stretching of the rubber strip; other curves show results after repeated stretching. (*After E. A. Meinecke and S. Maksin, Rubber Chemistry and Technology, 54:857, 1981*)

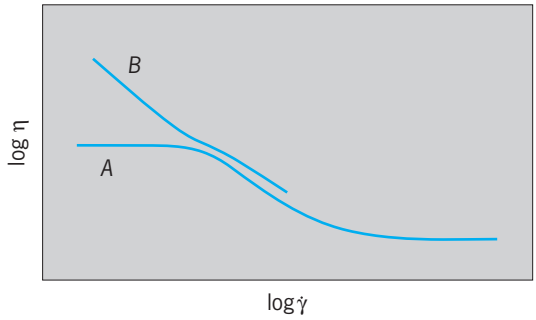

Fig. 5. Schematic logarithmic plot of the viscosity function $\eta(\dot{\gamma})$ for a polymer melt or concentrated polymer solution (curve A) and for the same polymer fluid with solid particles added (curve B).

symbol η has been adopted for this ratio, although now it is a function of $\dot{\gamma}$ as the notation indicates. It is called the viscosity function or sometimes simply the viscosity.

Shear thinning. For polymer melts, solutions, and suspensions, generally speaking, as indicated schematically in the logarithmic plot of **Fig. 5**, $\eta(\dot{\gamma})$ decreases as $\dot{\gamma}$ increases. This type of behavior, called shear thinning, is of considerable industrial significance. For example, paints are formulated to be shear-thinning. A high viscosity at low flow rates keeps the paint from dripping from the brush or roller and prevents sagging of the paint film newly applied to a vertical wall. The lower viscosity at the high deformation rates while brushing or rolling means that less energy is required, and hence the painter's arm does not become overly tired.

As $\dot{\gamma}$ goes to very small values for polymer melts and solutions, $\eta(\dot{\gamma})$ goes to a limiting value η_0, called the zero shear viscosity, as shown in curve A of Fig. 5. For high-polymer melts, η_0 is a very sensitive function of molecular weight; for a homologous series of linear polymers, η_0 is proportional to $M^{3.4}$; if the molecular weight is doubled, for example, η_0 is increased by a factor of about 10. Temperature also has a great effect on η_0, particularly in the neighborhood of T_g, where it can decrease by 75% per degree Celsius (45% per degree Fahrenheit).

On addition of solid particles (such as glass fibers or finely divided silica) to molten polymers or concentrated polymer solution, $\eta(\dot{\gamma})$ changes character as indicated in curve B of Fig. 5. A much larger effect is seen at low $\dot{\gamma}$ than at higher γ. Instead of approaching a limit at low $\dot{\gamma}$, $\eta(\dot{\gamma})$ keeps increasing as $\dot{\gamma}$ decreases. Some scientists believe that such suspensions will not flow unless the shear stress exceeds a value called the yield stress.

Normal stress effect. For newtonian fluids, steady simple shearing can be maintained by simple shear stress alone (Fig. 1). For nonlinear viscoelastic fluids, normal stresses must also be applied. If they are not, the height of the block of fluid will increase. This normal stress effect is also responsible for the observation that, when a rod is rotated in a container of some fluids (for example, some cake batters in a kitchen mixer), the fluid "climbs" around the rotating rod. With newtonian fluids, the liquid level becomes depressed near the rotating rod as the fluid moves outward due to centrifugal effects.

Die swell. When some molten plastics are forced through a tube at a slow rate, the diameter of the exiting stream can be considerably larger than that of the tube. Under similar flow conditions, newtonian fluids show a diameter increase of only a few percent. This phenomenon is known as die swell or extrudate expansion. In extrusion equipment for making plastic tubing and rods, the orifice must be made smaller than the desired size of the finished product. *See* EXTRUSION.

Drag reduction. When water is pumped through a pipe at rates high enough for the flow to be turbulent, the flow rate, for a given driving pressure, can be greatly increased by adding a very small amount (a few parts per million) of a very high-molecular-weight (greater than a million) polymer. This phenomenon has been named drag reduction. It has been used to increase the range of the water stream from a firehose while using the same pumping equipment on the firetruck. Drag reduction also was used to decrease the number of pumping stations required for the Trans-Alaskan oil pipeline. *See* TURBULENT FLOW.

Thixotropy. There are suspensions (for example, bentonite clay in water) which, after remaining at rest for a long time, act as solids; for example, they cannot be poured. However, if it is stirred, such a suspension can be poured quit freely. If the suspension is then allowed to rest, the viscosity increases with time and finally sets again. This whole process is reversibly; it can be repeated again and again. This phenomenon is called thixotropy. (The word thixotropy may be used in many different senses; a standard nomenclature for rheological phenomena and properties has not been universally adopted.) *See* GEL; NON-NEWTONIAN FLUID. Hershel Markovitz

Bibliography. J. Ferguson, *Applied Fluid Rheology*, 1991; A. I. Lurie, *Non-Linear Theory of Elasticity*, 1990; C. W. Macosko, *Rheology: Principles, Measurements, and Applications*, 1994; R. I. Tanner, *Engineering Rheology*, 2d ed., 2000; Y. G. Yanovsky, *Polymer Rheology: Theory and Practice*, 1993.

Rheostat

A variable resistor constructed so that its resistance value may be changed without interrupting the circuit to which it is connected. It is used to vary the current in a circuit. The resistive element of a rheostat may be a metal wire or ribbon, carbon disks, or a conducting liquid. *See* POTENTIOMETER; RESISTOR.

Frank H. Rockett

Rheumatic fever

An illness that follows an upper respiratory infection with the group A streptococcus (*Streptococcus pyogenes*) and is characterized by inflammation of the joints (arthritis) and the heart (carditis). Arthritis typically involves multiple joints and may migrate from

one joint to another. The carditis may involve the outer lining of the heart (pericarditis), the heart muscle itself (myocarditis), or the inner lining of the heart (endocarditis). A minority of affected individuals also develop a rash (erythema marginatum), nodules under the skin, or Sydenham's chorea (a neurologic disorder characterized by involuntary, uncoordinated movements of the legs, arms, and face). These clinical manifestations may occur singly or in combination, and although most disappear completely, damage to heart valves may be permanent and progressive, leading to severe disability or death from rheumatic heart disease years after the initial attack.

Only about 3% of individuals with untreated acute streptococcal pharyngitis ("strep throat") develop rheumatic fever. The disease occurs an average of 19 days after the infection and is thought to be the result of an abnormal immunologic reaction to the group A streptococcus. Initial attacks of rheumatic fever generally occur among individuals aged 5 to 15. Those who have had one attack are highly susceptible to recurrences after future streptococcal infections.

The incidence of rheumatic fever declined steadily in the United States and other developed countries following World War II. During the 1980s, however, outbreaks of rheumatic fever were reported in a number of communities, the largest occurring in Salt Lake City and the surrounding area. In the late 1980s, two outbreaks occurred among recruits in military training camps. Streptococcal strains associated with these civilian and military outbreaks had biological characteristics similar to the strains that were responsible for rheumatic fever earlier in the century, suggesting that the resurgence of the disease may be partially explained by the reappearance of "rheumatogenic" streptococci into the population. Although outbreaks of rheumatic fever appear to have subsided for now in the United States, the disease remains rampant in developing countries and continues to be the most common cause of heart disease in children around the world.

Treatment of rheumatic fever is largely symptomatic and supportive. Aspirin is very effective in controlling the symptoms of arthritis, whereas corticosteroids may be required for management of severe carditis. Neither of these drugs prevents the development of chronic rheumatic heart disease. Initial attacks of rheumatic fever can be prevented by treatment of strep throat with penicillin for at least 10 days. Patients who have had an episode of rheumatic fever should continue taking antibiotics for many years to prevent group A streptococcal infections that may trigger a recurrence of rheumatic fever. Vaccines to prevent strep throat, and thus rheumatic fever, are in the early stages of human testing. *See* HEART DISORDERS; STREPTOCOCCUS.

A. L. Bisno; James B. Dale

Bibliography. B. F. Massell, *Rheumatic Fever and Streptococcal Infection: Unraveling the Mysteries of a Dread Disease*, Francis A. Countway Library of Medicine, Boston, 1997.

Rheumatism

Any combination of muscle or joint pain, stiffness, or discomfort arising from nonspecific disorders. It is generally used as a lay expression to indicate a chronic or recurrent condition affecting a certain area and precipitated by cold, dampness, or emotional stress.

Fibromyositis refers to acute or chronic symptoms of tenderness, stiffness, and pain in a joint and related structures which follows exposure, strain, trauma, infection, or stress. Myositis denotes an inflammation of a muscle; myalgia refers to muscle pain or tenderness without inflammation. Fibrositis is an inflammation of fibrous connective tissue, usually that of a joint region.

Rheumatism includes all of the above nonspecific disorders and is best reserved for complaints not related to a specific disease such as arthritis, rheumatic fever, muscle disorders, or others that may cause the same or similar symptoms.

Lumbago, wryneck, charley horse, and shin splints are commonly used expressions included under the catchall category of rheumatism. Robert Searles

Rhinoceros

An odd-toed ungulate (order Perissodactyla) belonging to the family Rhinoceratidae. The bodies and limbs of these mammals are massive and thick-skinned. Also characteristic of the family are the horns (epidermal derivatives), of which there may be one or two, composed of a solid mass of hairs attached to a bony prominence on the skull. The feet are tridactyl, with three hooves; the middle one is the most developed. Five species are found in Asia and Africa (see **table**).

Sight is poorly developed in these animals; however, hearing is quite good and the sense of smell is excellent. Most species are solitary and aggressive, and are extremely dangerous since they are unpredictable in behavior. All species are vegetarians. The great Indian rhinoceros (*Rhinoceros unicornis*) is about 14 ft (4.3 m) long and nearly 6 ft (1.8 m) high, with a single horn that may be 2 ft (0.6 m) long in both sexes. This species, found in southern Asia, is usually solitary and inoffensive. The gestation period is 16–17 months and the single offspring weighs about 100 lb (45 kg) at birth. The White rhinoceros (*Ceratotherium simum*) is found in Africa and does not stray far from water, which it requires not only for drinking but also for bathing. It is a diurnal, peaceful, and sociable species and is usually seen grazing in small groups. It is the only rhinoceros that is entirely a grazer. The Black rhinoceros (*D. bicornis*) is becoming scarce in its natural range near the East African lakes and is found mainly in game preserves (see **illustration**).

Some species show a strict territoriality, having definite tracks through the grass, definite sites for dung deposition, and personal mud wallows. Mating

Distribution and characteristics of the rhinoceros

Name	Characteristics and distribution
Rhinoceros unicornis (*great Indian rhinoceros*)	Largest Asian species; thick skin with deep folds covered with scalelike granulations; South Asia
Rhinoceros sondaicus (Javan rhinoceros)	Rare; single short horn; granular skin; Java
Dicerorhinus sumatrensis (Sumatran rhinoceros)	Small (about $4^1/_2$ ft or 1.4 m high); two horns; Sumatra, Thailand, and Borneo
Dicerorhinus simus (white rhinoceros)	Two horns; square upper lip; flattened nose; Africa
Diceros bicornis (black rhinoceros)	Large (11 ft or 3.4 m long, 5–6 ft or 1.5–1.8 m high); two horns, the one on the nose being larger (about 2 ft or 0.6 m long) than the one behind it; Africa

may occur at any time of the year. Female White and Indian rhinoceroses breed for the first time when 6 to 8 years old; female Black rhinoceroses breed for the first time when 5 to 6 years of age. Males are usually several years older than females before they begin breeding. Gestation in most species lasts for 16 to 17 months with females normally giving birth to a single young every two to four years. Average longevity is 40 to 45 years. Major predators include humans, lions, and hyenas.

Black rhinoceros (*Diceros bicornis*). (*Dr. Lloyd Glenn Ingles, © 1999 California Academy of Sciences*)

The gradual disappearance of the rhinoceros is attributed to habitat destruction and poaching, especially the unrestrained hunting for the horns (which were used as a medicine, although of no medicinal value), the meat, and the hide. The known Sumatran rhinoceros population now consists of only about 300 animals; only about 60 Javan rhincoeroses remain. The Javan, Black, and Sumatran species are classified as critically endangered by the IUCN (International Union for Conservation of Nature) and Endangered by the U. S. Department of the Interior; the Indian rhinoceros is classifed as endangered; and one population of White rhinoceros in the Democratic Republic of Congo is classified as critically endangered. *See* PERISSODACTYLA; TERRITORIALITY.

Donald W. Linzey

Bibliography. *Grzimek's Encyclopedia of Mammals*, McGraw-Hill, 1990; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford Ltd., 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., The Johns Hopkins University Press, 1999.

Rhinovirus

A genus of the family Picornaviridae. Members of the human rhinovirus group include at least 113 antigenically distinct types. Like the enteroviruses, the rhinoviruses are small (17–30 nanometers), contain ribonucleic acid (RNA), and are not inactivated by ether. Unlike the enteroviruses, they are isolated from the nose and throat rather than from the enteric tract, and are unstable if kept under acid conditions (pH 3–5) for 1–3 h. Most strains can be recovered by culture only in cells of human origin; they grow better when cultures are rolled (rather than held stationary) and are kept at a temperature of 91.5°F (33°C). *See* ENTEROVIRUS; TISSUE CULTURE.

Rhinoviruses have been recovered chiefly from adults with colds and only rarely from patients with more severe respiratory diseases. *See* COMMON COLD.

Rhinoviruses occur throughout the world. Natural immunity exists but is transitory. Only 30–50% of volunteers can be infected; yet the "resistant" volunteers may catch cold at other times. After being infected by one type of rhinovirus, volunteers resist the infecting type for as much as 16 weeks, and other types of rhinovirus for less than 5 weeks. *See* IMMUNITY.

In a single community, different rhinovirus types predominate during different seasons and during different outbreaks in a single season, but more than one type may be present at the same time.

Although efforts have been made to develop vaccines, none is available. Problems that hinder development of a potent and useful rhinovirus vaccine include the short duration of natural immunity even to the specific infecting type, the large number of different antigenic types of rhinovirus, and the variation of types present in a community from one year to the next. *See* ANIMAL VIRUS; PICORNAVIRIDAE; VIRUS CLASSIFICATION.

Joseph L. Melnick; M. E. Reichmann

Bibliography. R. W. Compans et al. (eds.), *Current Topics in Microbiology and Immunology*, vol. 16: *Picornaviruses*, 1990; D. Hamre, Rhinoviruses, *Monographs in Virology*, vol. 1, 1968; F. L. Horsfall and I. Tamm (eds.), *Viral and Rickettsial Infections of Man*, 4th ed., 1965.

Rhizocephala

An order of crustacean parasites related to the barnacles. Worldwide in distribution, they prey on other crustaceans, principally Decapoda, such as crabs, shrimp, and their allies. In 1949 the rhizocephalan *Loxothylacus texanus* (see **illus.**) infested 16% of the blue crab population in Arkansas Bay, Texas. Rhizocephala produce modifications affecting the abdomen of the crab, making males resemble females and causing immature females to acquire precociously the adult form. These parasites have become so modified by their mode of life that, as adults, they are no longer recognizable as barnacles, or even as crustaceans.

An adult rhizocephalan is a thin-walled sac enclosing a visceral mass, composed chiefly of ovaries and testes. It shows no trace of segmentation, appendages, or sense organs. Even an alimentary tract is missing. Instead, it possesses a threadlike root system which penetrates the interior of the host in all directions and absorbs nourishment from the body fluids of the crab. Fertilized eggs develop in a brood chamber. An opening at the summit of the sac allows seawater to enter for purposes of respiration and provides an avenue of escape for the larvae. A short stalk securely rivets the base of the sac to the abdomen of the crab.

Fertilized eggs become nauplii. These larval parasites, expelled from the mother's brood pouch, measure 0.01 in. (0.25 mm) in length. They have the appearance of barnacle nauplii, except that they lack an alimentary tract. Free-swimming nauplii metamorphose into cypris larvae which appear sexually indifferent. Those that settle on crabs become kentrogen larvae. A cell mass injected by the kentrogen develops into a tumor and root system on the intestine of the host. After 8 weeks the endoparasite emerges as a small sac. Cypris that settle on immature Rhizocephala become larval males. A cell mass, injected into the brood pouch by these males, migrates to the "testes" and differentiates into sperm. Six weeks later the sac is sexually mature.

Rhizocephala, with the exception of the family Sylonidae, were considered hermaphroditic, possessing ovaries and testes in the same individual. In 1958, A. Ichikawa and R. Yanagimachi proved that the "testes" in *Peltogasterella socialis* were really seminal receptacles. The sperm in these female repository organs is produced exclusively by cypris male cells. Adults of *Peltogasterella* are thus true females carrying hyperparasitic larval males. The hermaphroditic nature of other Rhizocephala is now questionable. *See* DECAPODA (CRUSTACEA).

Placidus G. Reischman

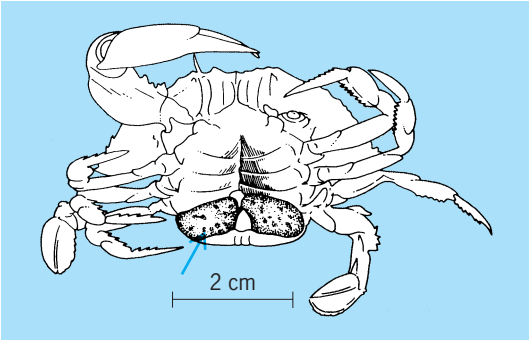

Loxothylacus texanus **(arrow) attached to the abdomen of the blue crab (*Callinectes sapidus*). (*After E. G. Reinhard, Parasitic castration of Callinectes, Biol. Bull., 98(3):277–288, 1950***

Bibliography. H. Boschma, The species of the genus *Sacculina* (Crustacea Rhizocephala), *Zool. Mededeel.*, 19(3–4):187–328, 1937; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; G. Smith, *Rhizocephala, Fauna und Flora des Golfes von Neapel*, vol. 29, 1906.

Rhizophorales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Rosidae of the class Magnoliopsida (dicotyledons). The order contains a single family, Rhizophoraceae, with about 100 species widely distributed in the tropics. The plants are mostly tanniferous trees and shrubs with the leaves opposite, simple, and entire. The flowers are regular, mostly perfect, and variously perigynous or epigynous. The sepals are four or five and commonly fleshy or leathery; the petals are the same number as the sepals and likewise small and fleshy. The stamens are twice as many as the petals or sometimes more. The pistil has 2–6 fused carpels with two or rarely more ovules per carpel; the fruit is berry-like or rarely a capsule. Most members of the family are inland species, but the most conspicuous group are some 17 species of shoreline shrubs, the mangroves. The fruits of mangroves are viviparous and have distinctive enlarged hypocotyls. The family has been subject to diverse interpretations, and has been associated with the Myrtales and with the Cornales. *See* MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM.

T. M. Barkley

Rhizopodea

A class of Sarcodina including both parasitic and free-living species found in fresh and salt water and the soil. No species forms true axopodia; instead, pseudopodia may be filopodia, lobopodia, or reticulopodia; or there may be no pseudopodia in some cases. Rhizopodea include five subclasses: Lobosia, Fiosia, Granuloreticulosia, Mycetozoia, and Labyrinthulia. Lobosia are amebas and testate rhizopods with lobopodia. Filosia are naked or testate species with filopodia. Granuloreticulosia are noted for their reticulopodia, which often fuse into extensive networks. A bidirectional flow of cytoplasm typically moves granules into opposite directions along the two sides of a single pseudopodium. Mycetozoia are a heterogeneous collection, including both cellular and true slime molds. Labyrinthulia are unique protozoa which form no obvious pseudopodia, but move by still undetermined methods on or inside secreted tubules. Many Rhizopodea

are phagotrophic, but ingestion has not been described in Labyrinthulia. A test may be present or absent in different genera of Lobosia, Filosia, and Granuloreticulosia. *See* FILOSIA; GRANULORETICULOSIA; LABYRINTHULIA; LOBOSIA; PROTOZOA; SARCODINA; SARCOMASTIGOPHORA. Richard P. Hall

Bibliography. J. J. Lee et al., *An Illustrated Guide to the Protozoa*, 2d ed., Society of Protozoologists, 2000.

Rhizosphere

The volume of soil around living plant roots that is influenced by root activity, either directly or via the activities of the microbiota that are stimulated by rhizodeposits (materials released by roots). Root–soil–microbiota interactions ultimately alter the biological, biochemical, chemical, and physical properties of the surrounding soil. The rhizosphere is thus a complex, dynamic, and heterogeneous microenvironment that plays a key role in plant nutrition and health, diversity of plant and soil microbial communities, soil functioning and genesis, and biogeochemical cycling of elements. *See* ROOT (BOTANY); SOIL.

Rhizodeposition and trophic relationships. The prime rhizosphere effect is a consequence of rhizodeposition. On average, 20% of the carbon assimilated by plants via photosynthesis is released by roots in the soil as a range of rhizodeposits: sloughed-off cells and decaying root hairs, respired CO_2, and organic compounds that are released via active secretion (for example, root mucilage) or passive diffusion (for example, root exudates such as sugars, amino acids, or organic acids). Rhizodeposition is a significant component of carbon cycling in ecosystems. Although sometimes viewed as a net loss of carbon for the plant, it is a major source of carbon for soil microbiota, making the rhizosphere a hot spot for trophic relationships. Rhizodeposits consist largely of rather small molecules that can be easily metabolized by microorganisms, either beneficial or deleterious (in the case of pathogenic microorganisms) for plants. Consequently, considerable stimulation of microorganisms is known to occur around roots. This rhizosphere effect is evidenced by an increased microbial biomass in the rhizosphere relative to bulk soil, particularly for those microorganisms that are cultivable in the laboratory, which are, however, known to represent only a minor proportion of the total number of microbial communities. Rhizosphere communities are especially enriched in rather fast-growing, opportunistic microorganisms (quite many of which are easily cultivable).

Recent advances in molecular microbial ecology have shown that the composition of microbial communities is considerably altered in the rhizosphere because the stimulation of microbial growth by rhizodeposits subsequently stimulates the activity of microbial grazers such as protozoa and nematodes in the rhizosphere. Thus, rhizodeposition enhances the turnover of rhizosphere microorganisms and soil microbiota, more so than the total microbial biomass. Combined with other changes in the structure of microbial communities, this will stimulate or depress a whole range of microbial activities in the rhizosphere, which will ultimately be either beneficial or detrimental to plants.

Such microbial activities and rhizodeposits are also involved in biochemical processes that have received less attention than those involved in carbon cycling, in spite of their importance: root-borne or microbial extracellular enzymes such as phosphatases, proteases, and arylsulfatases exhibit greater activity in the rhizosphere relative to bulk soil, and may have a dramatic effect on the cycle of nutrients such as phosphorus, nitrogen, and sulfur. In addition, the release of rhizodeposits exhibiting very different carbon/nitrogen ratios than bulk soil organic matter has been sometimes shown to induce a priming effect and enhance the mineralization of soil organic matter. Rhizodeposition thus significantly impacts the fate and cycle of major nutrients and trophic relationships in the soil surrounding living plant roots. *See* BIOGEOCHEMISTRY.

Communication and biological interactions. Alteration of the structure of rhizosphere microbial communities can also result from complex communication pathways between plant roots and rhizosphere microorganisms or between these microorganisms. Minute amounts of some signaling molecules can have a dramatic effect on microbial ecology. Such processes have been much studied with the advent of new tools and approaches in molecular ecology. Communication pathways and plant-microorganism interactions have been especially studied for associations such as legume–*Rhizobia* or mycorrhizal symbioses: such interactions result in distinct morphological features at the host root–microsymbiont interface, such as the nitrogen-fixing nodules and mycorrhizae that are involved in the exchanges of matter and energy between the microsymbiont (supplying inorganic nutrients) and the host plant (supplying carbon compounds). Prior to the formation of these specific, symbiotic organs, complex interactions occur which imply a communication pathway to establish adequate recognition of the two partners. The identification of flavonoids as signaling molecules has been well documented in legume–*Rhizobia* interactions, which explains the high specificity of this symbiosis. Genes involved in the regulation of symbiosis formation on both plant and microbial sides have now been identified and, interestingly, some of these are implicated in both rhizobial and mycorrhizal symbioses.

Beside the much-studied symbioses, interactions between roots and a whole range of plant growth–promoting rhizobacteria (PGPR) have received considerable interest. The production of phytohormones by PGPR has been well documented, for instance. Antibiotics have been shown to play a key role in the interactions between rhizosphere microorganisms, with potential applications in the field of biocontrol of root-borne diseases. Quorum sensing as related to molecules such as *N*-acetylated homoserine lactones is another type of interaction that has strong implications in the regulation of bacterial activities at the population level. Helper bacteria

implied in ectomycorrhizal symbiosis are an example of interactions which involve more than two partners. Many of these intriguing processes are somehow connected to the various strategies developed by roots and microbes for sensing their biotic and abiotic environment. Deciphering the molecular basis of the crosstalk between roots and microorganisms, as well as between various rhizosphere microorganisms, is a major challenge for managing rhizosphere ecology. *See* MYCORRHIZAE; SOIL MICROBIOLOGY.

Physical interactions. Although it is much less documented, root and microbial activities are also responsible for physical processes in the rhizosphere. These affect the movement of water and solutes in and out of the root, and thus the habitat of rhizosphere microorganisms. Among root activities that directly alter physical properties of the rhizosphere are primary functions such as root growth and uptake of water. These can result in changes in bulk density and porosity, soil strength, and water potential that can have a dramatic effect on the structure of microbial communities as well as on the uptake of water and nutrients by the plant. The production of mucilage by roots as well as microbial exopolysaccharides has also been shown to alter viscosity and surface tension properties of the soil, and thereby soil–water relationships in the rhizosphere; as well as the aggregation of soil around roots and the formation of rhizosheaths, which are unique structural features of the rhizosphere. Root-borne and microbial exopolysaccharides are also implicated in chemical processes such as the protection of root apex from aluminum toxicity and the acquisition of nutrients such as phosphorus. *See* PLANT MINERAL NUTRITION.

Chemical interactions. Chemical processes occurring in the rhizosphere as a direct consequence of a range of root functions such as absorption, respiration, and exudation are well documented. Combined with similar functions of rhizosphere microorganisms, these can result in steep gradients of concentrations of nutrients, potentially toxic elements, complexing or chelating compounds, pH and redox potential, and partial pressures of gases such as O_2 and CO_2. These chemical processes have a dramatic influence on the composition of the soil solution, and on the whole environment of the root, including the soil solid phase which is also the habitat of rhizosphere microorganisms. Some of these reactions play a role in the weathering of soil minerals and soil formation processes, e.g., in the dissolution of minerals such as calcium carbonate or phosphates as a result of rhizosphere acidification. These reactions also play a major role in the fate of nutrients. In the case of phosphorus, protons and exudates such as carboxylic anions (citrate, malate, or oxalate, especially) produced by roots and rhizosphere microorganisms have been shown to play a key role in some plant species such as legumes and in phosphorus-solubilizing bacteria or ectomycorrhizal fungi.

Other well-documented chemical interactions are those involved in the fate of iron. Plants have evolved two distinct strategies of iron acquisition, based either on localized rhizosphere acidification combined with enzymatic iron-reductase activity, or on the secretion of efficient iron chelators called phytosiderophores in grasses. Microbes also produce iron chelators, which represent a broader range of molecules called siderophores. These processes not only play a key role in the iron nutrition of the plant and microbiota: Given the poor availability of iron in many soils, the rhizosphere is the site for a strong competition, favoring the most iron-efficient organisms—via the production of siderophores, fluorescent *Pseudomonas* bacteria play an antagonistic role in controlling the population of the less iron-efficient *Fusarium* fungi. As the fungi is pathogenic, such chemical interactions occurring in the rhizosphere affect not only iron nutrition but also plant health.

Changes of pH, redox potential, and the production of complexing or chelating compounds in the rhizosphere also have a dramatic effect on the speciation and bioavailability of potentially hazardous trace elements such as arsenic, cadmium, and copper. *See* IRON; IRON METABOLISM; SIDEROPHORES; SOIL CHEMISTRY.

Knowledge gaps and prospects for management. The outer boundary of the rhizosphere is loosely defined, the spatial extension of the rhizosphere gradients extending for 1 mm or less when considering the depletion of phosphorus, for instance, while extending beyond several centimeters from the root surface when considering the outward diffusion of respired CO_2. Depending on which measure of root activity is considered, the rhizosphere volume can thus represent from much less than 1% up to 100% of the soil volume colonized by roots. In the case of phosphorus, mycorrhizal hyphae in symbiotic plants can considerably expand the depletion zone, sometimes described as the mycorrhizosphere. Thus, one cannot draw a single borderline between the rhizosphere and the bulk soil, and most rhizosphere processes do not show a sharp boundary but a more or less steep gradient. The rhizosphere is not only a site of radial heterogeneity around roots but also shows considerable spatial variation along the root axes and substantial temporal variations. These arise from the dynamic nature of root growth and the short- (diurnal rhythm) and long-term (seasonal rhythm) patterns of plant activities. These spatial and temporal heterogeneities combined with the multiplicity of biological, biochemical, chemical, and physical interactions make the rhizosphere a unique and complex microenvironment. Considerable variation also occurs between plant species or even between genotypes of a given species. Inoculation of PGPR, nitrogen-fixing bacteria, either symbiotic (*Rhizobia*) or nonsymbiotic (*Azospirilum*), or ectomycorrhizal fungi has already proved efficient for improving plant growth and nutrition. Applications of our knowledge of microbial interactions in the rhizosphere have also been successful in the field of biological control. In the future, better use of adequately selected or constructed plant genotypes may prove another efficient tool for managing those intimate rhizosphere interactions that ultimately determine

plant health, nutrition, and growth. *See* MICROBIAL ECOLOGY.

Philippe Hinsinger

Bibliography. E. A. Curl and B. Truelove, *The Rhizosphere,* Springer-Verlag, Berlin, 1986; P. Hinsinger, How do plant roots acquire mineral nutrients? Chemical processes involved in the rhizosphere, *Adv. Agron.*, 64:225–265, 1998; D. L. Jones, A. Hodge, and Y. Kuzyakov, Plant and mycorrhizal regulation of rhizodeposition, *New Phytologist*, 163:459–480, 2004; J. M. Lynch, *The Rhizosphere*, Wiley, Chichester, 1990; B. T. Tinker and P. H. Nye, *Solute Movement in the Rhizosphere*, Oxford University Press, 2000.

Rhizostomeae

An order of the class Scyphozoa with the most highly organized features of this class. The umbrella is generally higher than it is wide, except in *Cassiopea* and *Cephea*. The margin of the umbrella is divided into many lappets but is not provided with tentacles. Many radial canals, which are connected with each other to form a complicated network, issue from the cruciform stomach. The oral part is very complicated and is eight-sided, with many suctorial mouths surrounded by numerous small tentacles (see **illus.**). Usually there is no central mouth.

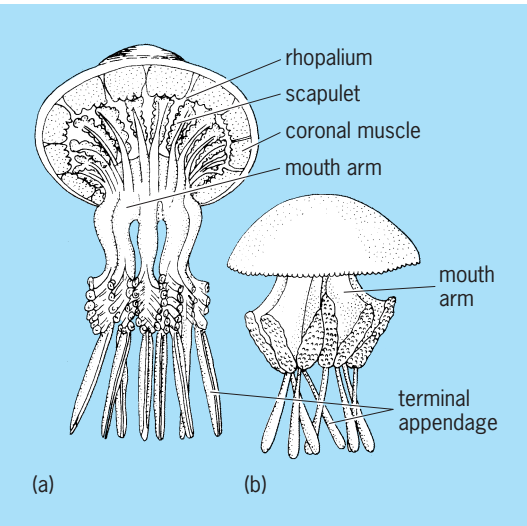

Rhizostomeae. (*a*) *Rhizostoma*; the bell may reach 2 ft (0.6 m) in diameter. (*b*) *Mastigias*. (*After L. Hyman, The Invertebrates, vol. 1, McGraw-Hill, 1940*)

The development of eggs into medusae, through the stages of planula, scyphopolyp, strobila, and ephyra, is similar to that of the Semaeostomeae. However, the metamorphosis of the ephyrae is more complicated than in the Semaeostomeae. No species of this group is injurious to the human skin. Some large forms, such as *Rhopilema*, are used as food in China and Japan.

A fair number of fossils of this order were found in the strata of the Jurassic Period. *See* SCYPHOZOA.

Tohru Uchida

Rhodium

A chemical element, Rh, atomic number 45, relative atomic weight 102.905. Rhodium is a transition metal and one of the group of platinum metals (ruthenium, osmium, rhodium, iridium, palladium, and platinum) that share similar chemical and physical properties. *See* METAL; PERIODIC TABLE; PLATINUM.

1												13	14	15	16	17	18
1 **H**	2																2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

The terrestrial abundance of rhodium is exceedingly low; it is estimated to be 0.0004 part per million in the Earth's crust. It is found as a single isotope, ^{103}Rh, with a nuclear spin of 12. Since the platinum metals share common reactivities and are mined from a common source, there is an involved chemical process that is used to separate the individual elements, including rhodium. *See* ISOTOPE.

Metallic rhodium is the whitest of the platinum metals and does not tarnish under atmospheric conditions. Its surface is normally covered by a thin, firmly bound layer of rhodium(IV) oxide (RhO_2). Rhodium is insoluble in all acids, including aqua regia. It dissolves in molten potassium bisulfate ($KHSO_4$), a useful property for its extraction from platinum ores, since iridium, ruthenium, and osmium are insoluble in this melt. Important physical properties of metallic rhodium are given in the **table**. *See* ACID AND BASE; AQUA REGIA; HALOGEN ELEMENTS.

Metallic rhodium is available as powder, sponge, wire, and sheets. It is ductile when hot and retains its ductility when cold. However, it work-hardens rapidly. Molten rhodium dissolves oxygen. Upon cooling, the oxygen gas is liberated, and this can lead to ruptures in the external surface of the crust of the metal. As a result, molten rhodium is best handled under an inert atmosphere of argon, which does not dissolve in rhodium.

Complexes of Rh(III), including $RhCl_3(pyridine)_3$, $Rh(CO)Cl_3[P(C_6H_5)_3]_2$, and $RhCl_6^{3-}$, are diamagnetic six-coordinate with octahedral geometry. The most common chemical form of rhodium is $RhCl_3 \cdot 3H_2O$, a red-brown, deliquescent material that is a useful starting material for the preparation of other rhodium compounds. In contrast to the hydrated material, red anhydrous rhodium(III) chloride ($RhCl_3$) is a polymeric, paramagnetic compound that does not dissolve in water. *See* DIAMAGNETISM.

Physical properties of rhodium metal

Property	Value
Crystal structure	Face-centered cubic
Lattice constant *a*, at 25°C (77°F), nm	0.38031
Thermal neutron capture cross section, barns (10^{-28} m^2)	149
Density at 25°C (77°F), g/cm^3	12.43
Melting point	1963°C (3565°F)
Boiling point	3700°C (6700°F)
Specific heat at 0°C, cal/g (J/kg)	0.0589 (246)
Thermal conductivity, 0–100°C, cal cm/cm^2s, °C (J · m/m^2 · s · °C)	0.36 (151)
Linear coefficient of thermal expansion, 20–100°C, μin./(in./°C) or m/(m · °C)	8.3
Electrical resistivity at 0°C, microhm-cm	4.33
Temperature coefficient of electrical resistance, 0–100°C/°C	0.00463
Tensile strength, 10^3 lb/in.^{2} (6.895 MPa)	
Soft	120–130
Hard	200–230
Young's modulus at 20°C (68°F), lb/in.^{2} (GPa)	
Static	46.2×10^6 (319)
Dynamic	54.8×10^6 (378)
Hardness, diamond pyramid number	
Soft	120–140
Hard	300
ΔH_{fusion}, kJ/mol	21.6
$\Delta H_{vaporization}$, kJ/mol	494
ΔH_f monoatomic gas, kJ/mol	556
Electronegativity	2.2

The low natural abundance and high cost of rhodium limit its uses to specialty applications. The major use is in catalysis, which accounts for over 60% of its production. Rhodium is a component of catalytic converters used in the control of exhaust emissions from automobiles. *See* CATALYTIC CONVERTER.

Rhodium is also used in the hydrogenation of olefins to alkanes. For hydrogenation, both heterogeneous catalysis and homogeneous catalysis are used. Heterogeneous conditions are achieved with rhodium metal finely dispersed on an inert support (activated carbon, charcoal, or alumina).

Rhodium complexes have been developed as catalysts for the synthesis of one optical isomer of L-dopa (used in treatment of Parkinson's disease). Greater selectivity makes rhodium catalysts more useful in hydroformylation or oxo reactions than the less expensive cobalt catalysts. A platinum-rhodium alloy is an efficient commercial catalyst for the formation of nitric acid through ammonia oxidation. *See* CATALYSIS; HETEROGENEOUS CATALYSIS; HOMOGENEOUS CATALYSIS; HYDROFORMYLATION; HYDROGENATION.

Rhodium-platinum alloys are favored for high-temperature applications. The International Temperature Scale over the range 630.5–1063°C (1134.9–1945.4°F) is defined by a thermocouple using a 10% rhodium-platinum alloy. Electroplated rhodium retains its bright surface under atmospheric conditions and finds use as electrical contacts and reflective surfaces. The reflectivity of rhodium surfaces is high (80%) and does not tarnish. About 6% of the rhodium production goes into jewelry manufacturing. *See* ELECTROPLATING OF METALS; TRANSITION ELEMENTS.

Alan A. Balch

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; R. S. Dickson, *Organometallic Chemistry of Rhodium and Iridium*, 1983; A. Earnshaw and N. Greenwood, *Chemistry of the Elements*, 2d ed., Butterworth-Heinemann, 1997; J. A. McCleverty and T. J. Meyer (eds.), *Comprehensive Coordination Chemistry II*, 2d ed., Elsevier Science, 2003.

Rho-theta system

A navigation system in which one or more signals are emitted from a facility (or colocated facilities) to produce simultaneous indication of bearing and distance. Since a bearing is a radial line of position and a distance is a circular line of position, the rho-theta system always ensures a position fix produced by the intersection of two lines of position which are at right angles to each other. This produces a minimum geometric dilution of position (GDOP), a figure of merit for all radio navigation systems, and is one of the chief advantages of a rho-theta system. Another major advantage is that it is a single-site system and can thus be installed on a ship or an island. This has also made it attractive politically, enabling small countries to have their own navigation systems. *See* ELECTRONIC NAVIGATION SYSTEMS; VOR (VHF OMNIDIRECTIONAL RANGE).

VOR. In May 1949 the Council of the International Civil Aviation Organization (ICAO), an agency of the United Nations, adopted the vhf (very high-frequency) omnidirectional radio range, commonly referred to as the VOR, to become standard on March 1, 1950. The ICAO also recommended that research and development work be actively pursued to produce a complete specification for distance-measuring equipment (DME) operating in the uhf (ultrahigh-frequency) band. The VOR was to operate in the band of 112 to 118 MHz. Equipment designed to the specification listed in the ICAO document was installed on the ground in many places

throughout the world. The location of these installations was agreed to by contracting states at their various ICAO regional meetings. This equipment operating with companion equipment on board aircraft produced static-free directional guidance to distances of over 150 mi (240 km). More than 900 VOR ground installations are in use in the United States and an equal number in the rest of the world.

It is difficult to specify the accuracy of the VOR system; accuracy of the airborne equipment varies considerably, from the highly engineered units which are utilized by the air carriers to equipment which the owner of a light pleasure craft purchases for one-tenth of the price. VOR accuracy also varies with terrain. In the early 1960s a wide baseline version of the VOR was developed and installed in the United States and Europe wherever a difficult site was encountered. This development was termed the Doppler VOR and has improved accuracy in some cases by as much as 10 times. *See* DOPPLER VOR.

According to the *2001 Federal Radionavigation Plan*, VOR/DME will continue tp provide navigation services for en route through nonprecision approach phases of flight throughout the transition to satellite-based navigation. The Federal Aviation Administration (FAA) plans to reduce VOR/DME services provided in the national airspace based on the anticipated decrease in its use for en route navigation and instrument approaches. The Phase-down is scheduled to begin in 2010. *See* SATELLITE NAVIGATION SYSTEMS.

Tacan. By the mid-1950s the military services of the United States had completed the development of a novel rho-theta system called Tacan for tactical air navigation. The system consists of ground equipment which produces a rotating radio field pattern in space very similar to that produced by the VOR; however, instead of the simple cardioid produced by the VOR, the Tacan antenna produces a cardioid plus a strong ninth harmonic which enhances its bearing accuracy. Unlike the VOR, the emission is not a continuous wave but a constant series of pulses. These pulses are emitted in time in accordance with the signals received from an airborne interrogator, so that by measuring round-trip time, distance to the ground beacon is displayed on a meter in the cockpit. The miltary need for and use of land-based TACAN will continue until aircraft are properly integrated with the Global Positioning System (GPS) and GPS is approved for all operations in national and international controlled airspace. The phase-down of TACAN usage is scheduled to begin in 2010 unless TACAN is determined to be necessary for long-term navigation services. *See* TACAN.

Vortac. The Tacan system operates in the band of 962–1213 MHz, which was the band that ICAO had recommended to be occupied by the civil DME. With the completion of the system, a facility combining the VOR and the Tacan was developed. Installations called vortacs were developed and installed where both the military and civil air operations had requirements for service. A vortac installation consists of the VOR beacon which furnishes directional guidance to civil aviation aircraft or to military aircraft that happen to be equipped with the VOR receiver.

The civil aviation aircraft carry a DME interrogator which interrogates the Tacan beacon that is colocated at the VOR site. The civil DME airborne interrogator measures round-trip elapsed time of the transmission and displays it as distance. A military aircraft carrying an airborne interrogator-responsor interrogates the ground colocated beacon and in reply receives both bearing and distance information from it. The phase-down of VORTAC sites is planned to begin in 2010.

Both the civil airborne VOR with its DME and the military Tacan can generate arbitrary courses to any point within the coverage of the ground vortac beacons by the use of simple rho-theta computers.

Other systems. The instrument landing system (ILS) and the microwave landing system (MLS) are, in essence, also rho-theta systems when they are used with DME. However, they do not normally provide full 360° coverage. The FAA has terminated the development of the MLS based on favorable GPS tests. The United States does not anticipate installing MLS equipment in the national airspace. *See* INSTRUMENT LANDING SYSTEM (ILS); MICROWAVE LANDING SYSTEM (MLS).

The advantage of the rho-theta concept has led to many other proposals, but the only one to have been implemented on a large scale is the Soviet Union's RSBN, developed before the Soviet Union joined the ICAO in 1970. It comprises a secondary radar, rotating at 100 revolutions per minute, which receives replies from airborne transponders. The aircraft derives bearing by comparing the arrival time of the radar beam with an omnidirectional pulse transmission from the radar each time the radar beam sweeps through the direction due north. The two-way link between radar and aircraft is also used by a conventional air-derived DME. RSBN, while standard throughout the Soviet Union, is not used elsewhere. Sven H. Dodington; Richard L. Greenspan

Bibliography. B. Forssell, *Radionavigation Systems*, Prentice Hall, 1991; M. Kayton and W. Fried, *Avionics Navigation Systems*, 2d ed., Wiley, 1997; G. Pakholkov and G. Grimov, Short-range radio navigation systems, *IEEE/AESS Magazine*, January 1988; U.S. Department of Defense and U.S. Department of Transportation, *2001 Federal Radionavigation Plan*, 2002.

Rhodochrosite

The mineral form of manganese carbonate. Calcium, iron, magnesium, and zinc have all been reported to replace some of the manganese. The equilibrium replacement of manganese by calcium has been determined and found to increase with the temperature of crystallization.

Rhodochrosite is sometimes found in low-temperature veins near deposits of copper, lead, zinc, and silver, or it may occur with other manganiferous minerals of higher temperature origin.

Rhodochrosite crystals, Alicante, Colorado. (*Specimen from Department of Geology, Bryn Mawr College*)

It has also been found in sediments and in pegmatites. Well-known occurrences of rhodochrosite are in Europe, Asia, and South America. In the United States large quantities occur at Butte, Montana. As a source of manganese, rhodochrosite is also important at Chamberlain, South Dakota, and in Aroostook County, Maine.

Rhodochrosite has hexagonal (rhombohedral) symmetry and the calcite-type crystal structure (see **illus.**). It occurs more often in massive or columnar form than in distinct crystals. The color ranges from pale pink to brownish pink. Hardness is 3.5–4 on Mohs scale, and specific gravity is 3.70. *See* CARBONATE MINERALS; MANGANESE.

Robert I. Harker

Rhodonite

A mineral inosilicate with composition $MnSiO_3$. Rhodonite crystallizes in the triclinic system in crystals that are commonly tabular parallel to the base (see **illus.**). More often it is in cleavable to compact masses or in embedded grains. Crystallographically, rhodonite is closely related to the pyroxenes and thus has two cleavage directions at about 88 and 92°. Hardness is 5.5–6 on Mohs scale, and specific gravity is 3.4–3.7. The luster is vitreous and the color is rose red, pink, or brown. Rhodonite is similar in color to rhodochrosite, manganese carbonate, but it may be distinguished by its greater hardness and insolubility in hydrochloric acid. It has been found at Langban, Sweden; near Sverdlovsk in the Ural Mountains; and at Broken Hill, Australia. Fine crystals of a zinc-bearing variety, fowlerite, are found at Franklin, New Jersey. *See* SILICATE MINERALS.

Cornelius S. Hurlbut, Jr.

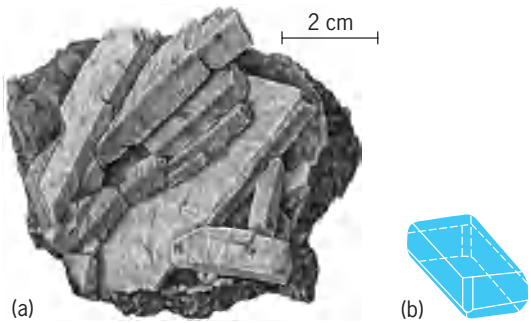

Rhodonite. (*a*) Crystals of variety fowlerite with limestone, Franklin, New Jersey (*specimen from Department of Geology, Bryn Mawr College*). (*b*) Crystal habit (*after C. Klein and C. S. Hurlbut, Jr., Dana's Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*).

Rhodophyceae

A large class of plants, commonly called red algae, coextensive with the division Rhodophycota. Most red algae are found in the ocean, growing on rocks, wood, other plants, or animals in the intertidal zone and to depths limited by the availability of light. A few genera and species occur in fresh water, and these are usually found in rapidly flowing, well-aerated, cold streams. Some, however, grow in quiet warm water, while a few are subaerial. Most red algae are photosynthetic, but some grow on other algae with varying degrees of parasitism. Approximately 675 genera and 4100 species are recognized. *See* ALGAE.

Characteristics. Rhodophyceae are characterized by a unique combination of biochemical, reproductive, and ultrastructural features. The primary photosynthetic pigment is chlorophyll *a*. In addition, small amounts of chlorophyll *d* are found in some species. Water-soluble tetrapyrrolic compounds called phycobilins serve as accessory photosynthetic pigments. The phycobilins, which include phycoerythrin (red) and phycocyanin and allophycocyanin (blue), are bound to proteins. The resulting pigment-protein complexes (phycobiliproteins) aggregate to form phycobilisomes, which are arranged in an orderly pattern on the outer surfaces of the chloroplast lamellae. The carotenoid pigments are similar to those of green plants. Varying ratios of chlorophylls, carotenoids, and phycobilins provide a spectrum of color from blue-green through olive-green and reddish brown to bright pink. The production of phycoerythrin is adversely affected by nitrogen deficiency and intense illumination, so that the brightest pink or red colors are found in plants growing subtidally in nutrient-rich water. The absorption by the phycobilins of energy in the green portion of the spectrum compensates for the loss of red rays, which are absorbed in the upper few meters of water. The chloroplasts, which may be stellate, discoid, or linear, are bounded by a double membrane. Each photosynthetic lamella comprises a single thylakoid. A pyrenoid is present in some genera.

The chief food reserve, floridean starch, is a branched polymer of glucose (α-1,4-glucan) similar to amylopectin of green plants. It occurs as granules in the cytoplasm. Multinucleate cells are common. In some genera, specific areas of nominally haploid or diploid plants are polyploid or polytenic, the orderly progression from one condition to another being related to ontogeny. The mitochondria have flattened cristae. The cell walls are composed of an inner fibrillar layer of cellulose and an outer mucilaginous layer of sulfated polymers of galactose that have economic importance. Throughout the order Corallinales, the cell walls are

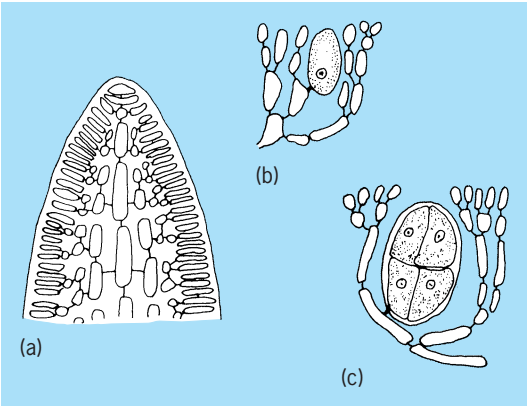

Fig. 1. ***Gelidium*. (*a*) Longitudinal section of apex of branch, showing uniaxial filamentous construction. (*b*) Tetrasporangium mother cell, before meiosis. (*c*) Tetrasporangium, following meiosis, cleaved into tetraspores. (*After G. M. Smith, Cryptogamic Botany, 2d ed., McGraw-Hill, 1955*)**

impregnated with the calcite form of calcium carbonate. In two common tropical noncoralline red algae (*Galaxaura* and *Liagora*), the cell walls are impregnated with the aragonite form of calcium carbonate. Rhodophyceae are distinctive among eukaryotic algae in their lack of flagella, a feature shared among major groups only by the chlorophycean order Zygnematales. Some unicellular forms (in the order Porphyridiales) and many spores and male gametes of multicellular forms are capable of gliding or feeble ameboid motion. *See* CORALLINALES.

Structure. Unicellular red algae, which may form mucilaginous colonies, are considered primitive. Most red algae have multicellular thalli of microscopic or macroscopic size, including individual filaments, blades, and complex plants of distinctive form produced by the interplay of filamentous systems (**Fig. 1***a*). The largest red algae rarely exceed a meter (3.3 ft) in length. In all but a few families, filaments grow by elongation of existing cells and cutting off of new cells by apical cells. Typically, intercalary cells initiate branches, but never divide transversely. In all members of the subclass Florideophycidae and in a few members of the subclass Bangiophycidae, the derivatives of a dividing cell remain linked. This link, the primary pit connection, results from incomplete cell plate formation. A hole left in the center subsequently becomes plugged with a proteinaceous substance. Contiguous cells of different lineages (or rarely the same lineage) may be linked secondarily by pit connections formed when a very small cell cut off from the initiating cell fuses with an adjoining cell. Secondary pit connections strengthen the thallus. In some members of the Corallinales, they are formed directly, without an intermediate conjunctive cell. They may also be formed between a host and its red algal parasite. During their formation, a nucleus is transferred from the parasite to the host. Fusions between and among cells is frequently encountered during development of the female reproductive system for the apparent purpose of increasing the supply of nutrients to the maturing cystocarp. Fusions are also found in vegetative tissues of many Corallinales, where their function is unknown.

Life histories and reproduction. Most red algae have a life history involving a succession of somatic and cytological phases. A free-living diploid phase (tetrasporophyte) bears sporangia that undergo meiosis to form tetrads of haploid spores (tetraspores; Fig. 1*b*, *c*). The tetraspores germinate, and the germlings develop into haploid thalli (gametophytes). Male and female gametes are usually produced on different plants. The tetrasporophyte and the gametophytes may be similar in size and form (isomorphic) or dissimilar to a marked degree (heteromorphic). The three free-living phases often exhibit vegetative differences that are so slight as to be appreciated only by specialists; such life histories are considered to be isomorphic. In the most common heteromorphic life history, the tetrasporophyte is a crust or a minute filamentous structure while the gametophytes are macroscopic blades. In one order (Palmariales) the male gametophyte resembles the tetrasporophyte, but the female gametophyte is greatly reduced, in at least one species to a single cell. A dwarf male gametophyte is produced by one anomalous species in a family characterized by isomorphy.

The contents of a male sex organ (spermatangium) function as a male gamete (spermatium; **Fig. 2***a*). Spermatia, which measure 2–6 micrometers in diameter, are spherical or ovoid, colorless, and produced in great abundance. Although they have no flagella, some spermatia have mucilaginous appendages that may assist in dispersal or attachment to the female organ (carpogonium), whose contents function as the female gamete. The carpogonium is usually a flask-shaped cell with an elongate receptive extension (trichogyne; Fig. 2*b*). The nucleus of an attached spermatium migrates down the trichogyne to the base of the carpogonium, where it fuses with the female nucleus to produce a zygote. In some red algae the zygote develops directly into a compact, diploid somatic phase, the carposporophyte (or gonimoblast), which remains an integral part of the female plant (Fig. 2*c*). A protective envelope (pericarp) may be produced by the female gametophyte; the carposporophyte with its pericarp is termed

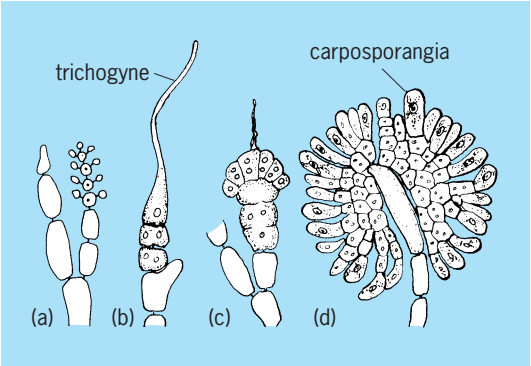

Fig. 2. ***Nemalion*. (*a*) Terminal group of spermatangia. (*b*) Three-celled carpogonial branch terminating in carpogonium with trichogyne. (*c*) Early stage in direct development of carposporophyte from zygote. (*d*) Longitudinal section through mature carposporophyte, showing terminal carposporangia. (*After G. M. Smith, Cryptogamic Botany, 2d ed., McGraw-Hill, 1955*)**

the cystocarp. In most red algae, by contrast, the zygote nucleus is transferred to another, often cytologically differentiated cell (auxiliary cell) which then undergoes a series of divisions to produce the carposporophyte. The carposporophyte in all red algae is partly or entirely converted into carposporangia (Fig. 2*d*). Carpospores are released singly from carposporangia and germinate to give rise to tetrasporophytes. In a few species, meiosis is involved in the production of carpospores, which therefore are haploid and germinate to give rise to gametophytes, thereby eliminating the tetrasporophyte phase. Some red algae bypass sexual reproduction by producing apomeiotic (unreduced) tetraspores, mitotically derived monospores, or vegetative propagules.

Classification. Two subclasses are traditionally recognized. In the Bangiophycidae the thallus is relatively simple: unicells, mucilaginous colonies, branched and unbranched discrete filaments, and thin blades. Only asexual reproduction is known for most members of this subclass, although the widespread and well-known genus *Porphyra* has a heteromorphic sexual life history. Growth is diffuse in the Bangiophycidae, and pit connections are formed in only a few genera. The Florideophycidae comprise multicellular forms, most of which are known or suspected to have sexual life histories. Growth is initiated by apical cells and all members form pit connections. Classification within the Bangiophycidae is based largely on vegetative and asexual reproductive features, while that of the Florideophycidae is based primarily on details of the development of the female reproductive system and carposporophyte, secondarily on vegetative features.

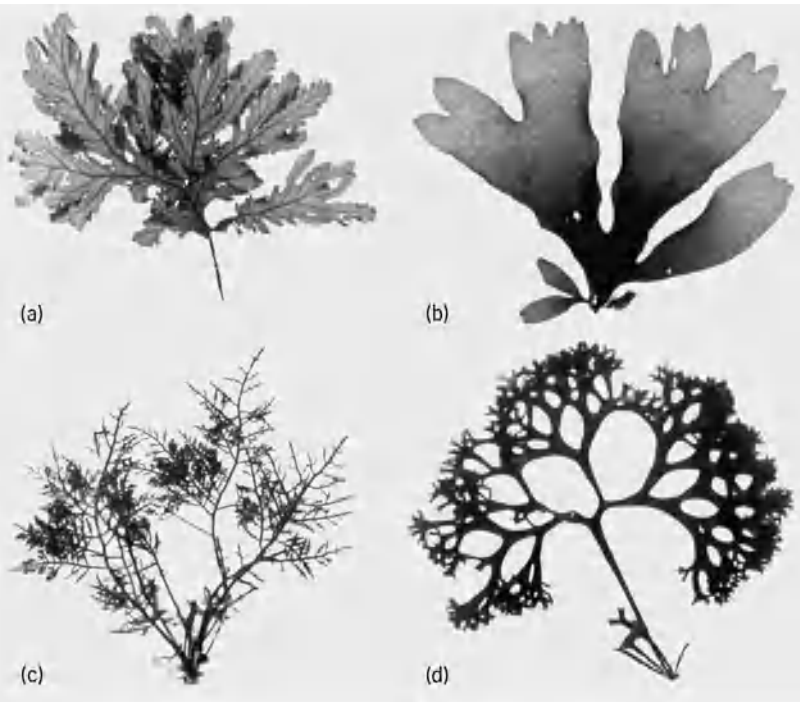

Fig. 3. Representative species of Rhodophyceae. (*a*) *Phycodrys fimbriata*, or sea oak, an alga of beauty but of no commercial importance. (*b*) *Palmaria palmata*, or dulse, eaten by maritime peoples in the North Atlantic. (*c*) *Gelidium robustum*, a commercial source of agar. (*d*) *Chondrus crispus*, or Irish moss, the chief commercial source of carrageenan.

Phylogeny. Rhodophyceae seem most closely related to Cyanophyceae in their use of chlorophyll *a* and phycobilins as photosynthetic pigments and the absence of flagella. They probably did not evolve directly from Cyanophyceae, however, but from a colorless, nonflagellate, eukaryotic ancestor that acquired pigments from an endosymbiotic blue-green alga. Fronds, presumably of red algae, are found in Cambrian shales. A more certain fossil record is provided by the reef-forming calcareous Corallinales, which extend back to the Devonian (about 400 million years ago).

Economic importance. The most salient feature of red algae is their beauty (**Fig. 3***a*), which has drawn admiration from generations of seaside visitors. A popular and sometimes profitable Victorian pastime, in both Britain and the United States, was to arrange red algae artistically on cards, which were sold as souvenirs. Their greatest significance, however, is their role in the formation of coral reefs, the Corallinales being responsible for cementing together various animal and algal components. Of more apparent economic importance is their use as food, a centuries-old tradition of maritime peoples in many parts of the world. In the North Atlantic, *Palmaria palmata* (Fig. 3*b*, formerly, but incorrectly, placed in the genus *Rhodymenia*) was relished under the name dulse (English) or dillesk (Irish), while *Chondrus crispus* (Irish moss; Fig. 3*d*) was used in making a pudding by the extraction of a jelling agent. In the Far East and Indonesia, jelling agents were extracted from *Gelidium* (Fig. 1*d*), *Gracilaria*, and *Eucheuma*. The reddish-purple blades of *Porphyra* (asakusi-nori or simply hoshi) have long been considered a delicacy by the Japanese. *See* REEF.

Many of these folk customs have been capitalized into substantial commercial enterprises. Carrageenan, the jelling agent of Irish moss, has been found to have numerous industrial applications and is extracted in factories from *Chondrus*, the closely related genus *Gigartina*, and the less closely related genus *Eucheuma*. Natural sources of *Eucheuma* are greatly supplemented by crops grown in shallow water, especially in the Philippines, where the economy of coastal villages has been significantly strengthened. Agar, the jelling agent of *Gelidium* and *Gracilaria*, is best known for its use in preparing microbiological media. Its manufacture, although still dependent upon naturally occurring plants, has spread from Japan to nations (such as Spain) without a traditional agar industry. *See* AGAR; CARRAGEENAN.

The great economic importance of producing nori in Japan is commensurate with the social importance of eating it. Specially prepared kinds command astoundingly high prices in the luxury food market, while the ordinary product is a staple in the Japanese household. *Porphyra* has been cultivated in Japan for nearly three centuries, originally on twigs secured in the mud of shallow bays, later on rope nets suspended at tide levels optimal for

growth. After the discovery in 1949 that the life history of *Porphyra* includes a microscopic spore-producing stage (previously described as an independent shell-boring alga, *Conchocelis*), the cultivation was modified to take advantage of natural cycles of sporulation. The "seeding" of nets, formerly done by randomly towing them in open-coast areas where *Porphyra* occurred naturally, is done in the factory by dipping them in a culture of the conchocelis stage. Paul C. Silva; Richard L. Moe

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae: Structure and Reproduction*, 2d ed., 1997; P. S. Dixon, *Biology of the Rhodophyta*, 1973; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; R. G. Sheath, The biology of freshwater red algae, *Prog. Phycol. Res.*, 3:89–157, 1984.

Rhombifera

A diverse class of extinct, marine, stalked echinoderms whose stratigraphic range extends from the Late Cambrian (505 MA) through Upper Devonian (360 MA). The class is named for rhombus-shaped structures (rhombs) that circulated seawater or body fluids through the body for the respiratory extraction of dissolved oxygen. The main body, or theca, held the visceral mass and was attached to the seafloor by a short to long stem. Atop the theca were variously developed small appendages that were used for suspension feeding in slowly moving currents. This class is composed of three distantly related groups: Glyptocystitida, Hemicosmititida, and Caryocystitida (see **illustration**). A fourth group, Polycosmititida, is too poorly known to assign to Rhombifera with certainty.

Glyptocystitids (Late Cambrian to Upper Devonian) comprise the bulk of rhombiferans. This group is characterized by a large theca, comprised of 25–27 skeletal elements, which bore numerous hairlike feeding appendages called brachioles. The short, stout stem is enlarged and flexible near the thecal attachment. The rhombs have elongate slits with little differentiation between incurrent and excurrent sides. Glyptocystitids include most of the variation with sessile and mobile species, and thecal shapes including barrel-shaped, spherical, and flattened triangular.

Hemicosmitids (Early Ordovician to Early Devonian) are characterized by a large theca with about 36 skeletal elements bearing stout erect arms upon which hairlike brachioles were born. The stem is long and stout and lacks the enlarged upper end of glyptocystitids. Rhombs have small sievelike clusters of incurrent pores and larger spoutlike excurrent pores.

Caryocystitids (Early to Late Ordovician) are characterized by an irregularly elongate to globular theca formed from numerous skeletal elements, bearing a few stout feeding brachioles. The stem is short, gracile, and irregularly plated. Rhombs of caryocystitids were fully internal, circulating body fluids instead of seawater. *See* ECHINODERMATA.

Colin D. Sumrall

Bibliography. R. V. Kesling, Cystoids, pp. S85–S267, in R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, Part S: *Echinodermata* 1(1), Geological Society of America and University of Kansas, New York and Lawrence, 1968; J. Sprinkle, *Morphology and Evolution of Blastozoan Echinoderms*, Spec. Publ., Museum of Paleontology, Harvard University, Cambridge, 1973; J. Sprinkle, Cylindrical and globular rhombiferans, pp. 231–273, in J. Sprinkle (ed.), *Echinoderm Faunas from the Bromide Formation (Middle Ordovician) of Oklahoma*, University of Kansas, Lawrence, 1982.

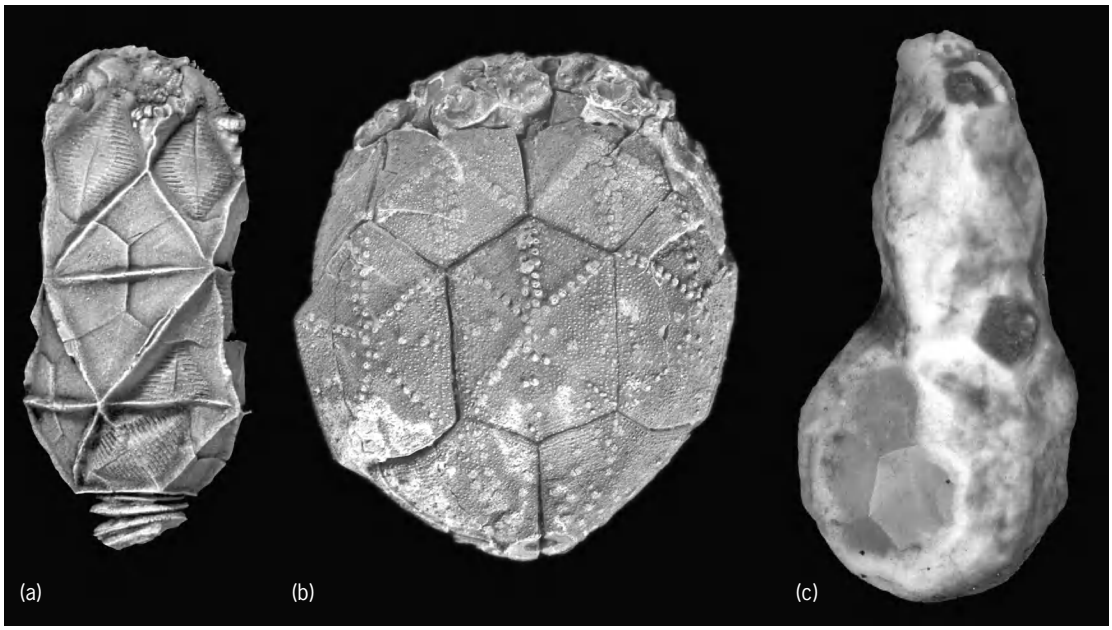

Rhombiferans of three groups. (*a*) The glyptocystitid *Cheirocystis fulotonensis*. (*b*) The hemicosmitid *Caryocrinites ornatus*. (*c*) The caryocystitid *Caryocystites angelini*.

Rhubarb

A herbaceous perennial, *Rheum rhaponticum*, of Mediterranean origin, belonging to the plant order Polygonales. Rhubarb is grown for its thick petioles which are used mainly as a cooked dessert (see **illus.**); it is frequently called the pie-plant. The leaves, which are high in oxalic acid content, are not commonly considered edible. Propagation is by division of root crowns. Victoria, Macdonald, and Valentine are popular varieties (cultivars). Commercial production is generally limited to areas where crowns may become dormant for 2–3 months each year. Outdoor rhubarb is a common garden vegetable in most areas of the United States except the South. Harvesting begins in the spring and continues for 6–10 weeks. Commercial plantings are renewed every 4–8 years. Michigan and Washington are important centers for forced or hothouse rhubarb. Two- or three-year-old field-grown crowns are moved into darkened forcing structures in late winter and forced at 55–60°F (12–15°C) to obtain petioles of a bright-red color. *See* POLYGONALES.

H. John Carew

Cut stalks of rhubarb with leaves removed. (*W. Atlee Burpee Co.*)

Rhynchobdellae

An order of the class Hirudinea. These leeches possess an eversible proboscis and lack hemoglobin in the blood. They may be divided into two families, the Glossiphoniidae and the Ichthyobdellidea.

Glossiphoniidae are flattened, mostly small leeches occurring chiefly in fresh water. *Theromyzon* is a parasite of waterfowl, sucking blood from the nasal cavity, the mouth, or the eye. A heavy infestation on a young bird may be fatal. *Hemiclepsis* sucks the blood of fishes and amphibians and is found in stagnant waters in Europe and northern Asia. *Placobdella* is a common American parasite of aquatic turtles, frogs, and fishes, and *Glossiphonia* sucks the blood of aquatic invertebrates, especially snails (see **illus.**). The genus *Haementeria* contains reptile parasites chiefly, but also includes parasites of mammals, including humans. *Haementeria officinalis* is used in Mexico for medicinal purposes.

Ichthyobdellidae typically have cylindrical bodies with conspicuous, powerful suckers used to attach themselves to passing fish. They frequently have lateral appendages which aid in respiration. There is a tendency for a particular species to be confined to a specific host. For instance, *Callobdella lophii* is confined to the angler fish *Lophius piscatorius*, and *Abranchus microstomus* to the small shore fish *Cottus scorpius*. On the other hand, *Pontobdella muricata* attacks a variety of skates and rays. *Crangonobdella* is an example of a parasite of crustaceans. *Piscicola* is one of the few fresh-water forms. *See* ARHYNCHOBDELLAE; HIRUDINEA.

Kenneth H. Mann

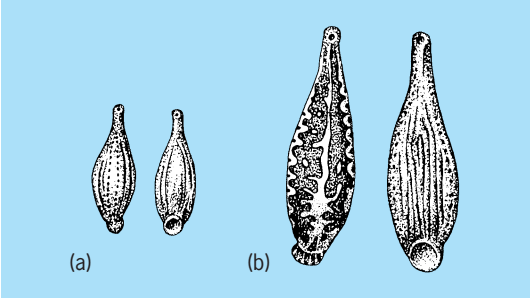

Examples of Rhynchobdellae. (*a*) *Glossiphonia complanata*. (*b*) *Placobdella parasitica*.

Rhynchocephalia

One of the two surviving clades of lepidosaurian reptiles, represented today by only two species, *Sphenodon punctatus* and *Sphenodon guntheri*, commonly known as tuatara. *Sphenodon* (a name meaning wedge-tooth) has a fully diapsid skull (**Fig. 1***a*) that was secondarily acquired from an ancestral state represented by *Gephyrosaurus*, the most primitive known member of the clade Rhynchocephalia, in which the lower temporal bar was incomplete, due to a loss of contact between jugal and quadratojugal (Fig. 1), as in lizards and basal lepidosauromorphs. The extinct *Gephryosaurus* shows some but not all of the advanced characters of more typical rhynchocephalians. Most notably, it has a primitive tooth implantation, that is, pleurodont (attached to the inside of the jaw). Typically, the teeth of rhynchocephalians are fused to the edges of the jaws (a condition known as acrodont). There are two upper rows of teeth, one of which is enlarged and palatine (on the palate), and one row of dentary (lower jaw) teeth. The dentary teeth bite between the two upper rows in a vertical manner (as in some extinct taxa, for example, *Clevosaurus*; Fig. 1*c*) or in a horizontal sliding movement known as propaliny (as in the living *Sphenodon*). *See* DIAPSIDA; LEPIDOSAURIA.

Living Sphenodon. The living New Zealand tuatara, or *Sphenodon*, is a medium-sized (2–2$^1/_2$ ft, or 60–80 mm) reptile, rather like a heavy-bodied iguana in appearance but very different in its lifestyle (**Fig. 2**). Tuatara is a Maori word meaning "peaks on the back," a reference to the jagged dorsal crest. The reptiles have an important place in Maori tradition, but, unfortunately, were also part of the cuisine. Once widespread on the main islands of New Zealand, they are now restricted to about 30 small offshore islets, threatened mainly by introduced animals such as pack-rats and cats.

Sphenodon is unique among lepidosaurs in having only very rudimentary hemipenes (the paired

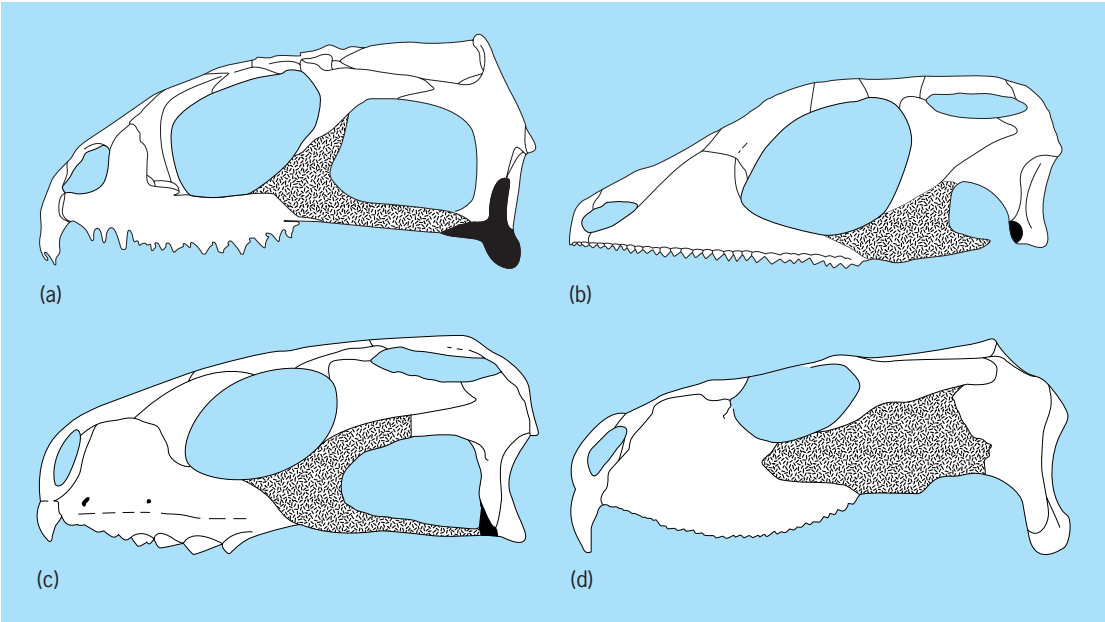

Fig. 1. Rhynchocephalian skulls in simplified lateral view. In each skull, the jugal is shown in stipple and the quadratojugal in black. (*a*) *Sphenodon* (Recent, New Zealand). (*b*) *Gephyrosaurus* (Lower Jurassic, Britain) [*redrawn from S. E. Evans, At the feet of the dinosaurs: The early history and radiation of lizards, Biolog. Rev., 78:513–551, 2003*]. (*c*) *Clevosaurus* (Upper Triassic, Britain) [*redrawn and simplified from N. C. Fraser, The osteology and relationships of Clevosaurus (Reptilia: Sphenodontida), Phil. Trans. Roy. Soc. Lond., B321:125–178. 1988*]. (*d*) The herbivorous *Priosphenodon* (Upper Cretaceous, Argentina) [*redrawn and simplified from S. Apesteguia and F. E. Novas, Large Cretaceous sphenodontian from Patagonia provides insight to lepidosaurian evolution in Gondwana, Nature, 425:609–612, 2003*].

male copulatory organs that characterize lizards and snakes) and an extended, tortoiselike lifespan. It also possesses a well-developed parietal eye, a light-sensitive organ on the top of the head that helps in the control of seasonal cycles. By day, tuatara live in shallow burrows excavated by seabirds; at night they emerge to feed. The typical diet consists of worms, snails, and insects, most notably large local crickets called wetas. However, seasonally the larger males will feed on bird chicks, small lizards, and young tuatara. Fertilization is internal. The eggs are buried and take 12–14 months to develop.

Gephyrosaurus. The earliest known rhynchocephalians date from the beginning of the Upper Triassic in Europe, but their relatively derived skulls suggest a sustained period of evolution before this. The primitive rhynchocephalian morphology is represented by *Gephryosaurus*, a small superficially lizardlike reptile from the Lower Jurassic of Britain (it was already a relict, or living fossil, by that time). The basal position of *Gephyrosaurus* is recognized by the fact that it is the only rhynchocephalian that is not included within the higher group Sphenodontia (that is, *Gephyrosaurus* plus Sphenodontia form Rhynchocephalia).

Other fossil groups. By the end of the Triassic, rhynchocephalians had spread globally and are known from India, Madagascar, South Africa, Brazil, North America, China, and Europe. They seem to have become extinct in Asia in the Early to Middle Jurassic, but elsewhere they survived into the Cretaceous. These survivors show greater anatomical diversity than their insectivorous ancestors, giving rise to derived long-bodied swimmers (for example, *Palaeopleurosaurus* and *Pleurosaurus*) and herbivores with specialized dentitions (for example, *Priosphenodon*: see Fig. 1*d*). Despite this, the group seems to have become extinct in Europe and North America by the end of the Lower Cretaceous (approximately 110 million years ago). This created something of a paradox, given their survival in New Zealand to the present day. Recent finds of rhynchocephalians in Late Cretaceous deposits in South America (*Priosphenodon*, Argentina) have helped to

Fig. 2. Young *Sphenodon* (tuatara) at Chester Zoo, U.K. (*Photo by Susan Evans*)

bridge the gap, hinting at the possibility of a long, unrecorded post-Cretaceous history in southern continents. *See* REPTILIA. Susan E. Evans

Bibliography. M. J. Benton (ed.), *The Phylogeny and Classification of the Tetrapods*, vol. 1: *Amphibians, Reptiles and Birds*, Clarendon Press, Oxford, 1988; T. R. Halliday and K. Adler (eds.), *The New Encyclopedia of Reptiles and Amphibians*, Oxford University Press, Oxford, 2002; G. R. Zug, L. J. Vitt, and J. P. Caldwell, *Herpetology*, Academic Press, San Diego, 2001.

Rhynchonellida

An extant order of brachiopods that has been an important component of marine benthic communities since the Ordovician. As one of only two Paleozoic orders in the subphylum Rhynchonelliformea that survives to the present, the rhynchonellids exemplify the fate of the Paleozoic evolutionary fauna. Despite their stratigraphic longevity, rhynchonellids have been a remarkably conservative group morphologically.

Rhynchonellids possess unequally biconvex valves typically with a fold and sulcus; many species contain strong radial ribs that produce deflections in the commissure, the line of junction between the two valves. Their shells are generally impunctate and also typically lack a hinge line parallel to the hinge axis, resulting in a pointed beak or umbo when viewed in lateral profile (**illus.** *a–c*). Internally, rhynchonellids possess calcareous processes (crura) that in extant species provide support for the lophophore. The presence of crura is a key feature of rhynchonellid evolution. Rhynchonellids were derived from a group of pentamerid brachiopods that also possessed crura. Rhynchonellids are also thought to be the ancestral group for the spire-bearing brachiopods and then ultimately the terebratulids by elaboration of the crura into spiralia and then into a complex loop.

Rhynchonellids are sessile, attached, epifaunal suspension feeders. They have a functional pedicle that they use to attach to the substrate (illus. *a*). Although rhynchonellids were never diverse compared to other brachiopod orders, they were commonly important members of local communities. They achieved a diversity peak in the Devonian and again in the Jurassic (illus. *d*).

Rhynchonellids have shifted their habitat preference in the oceans since their origin in the Middle Ordovician when they originated in shallow low-latitude seas; however, presently they are more common in deep-water habitats from middle and high latitudes and are rare members of benthic communities in low-latitude shallow seas. This major shift in habitat preference represents a common pattern among marine orders over the last 550 million years and especially in the last 250 million years. Marine orders have tended to originate in shallow-water habitats, have expanded offshore, and in many cases have ultimately become restricted to deep-water environments. In the case of the rhynchonellids, their restriction from low-latitude shallow-water environments apparently occurred as early as the Jurassic and may reflect differential survival from the late Permian mass extinction. An alternative explanation suggests that interactions with other diversifying taxonomic groups dampened the rebound of the rhynchonellids from the late Permian mass extinction in low latitudes. *See* ARTICULATA (ECHINODERMATA); BRACHIOPODA. Mark E. Patzkowsky

Bibliography. J. A. Walsh, No second chances? New perspectives on biotic interactions in post-Paleozoic brachiopod history, in P. Copper and J. Jin (eds.), *Brachiopods*, A. A. Balkema, Rotterdam, 1996.

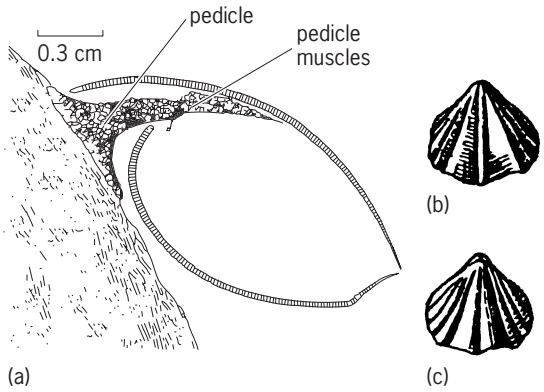

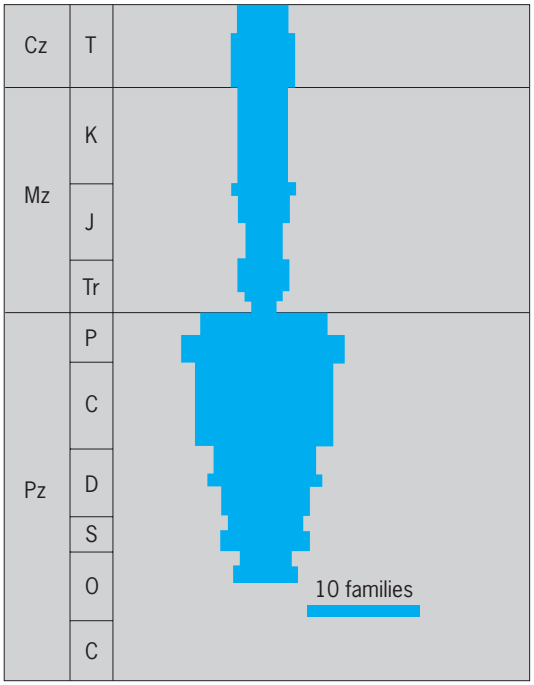

Rhynchonellida. (*a*) Lateral view of the modern species *Notosaria nigricans* with shell removed showing the pedicle system. (*b–c*) Ventral and dorsal views of *Rhynchotrema*, Ordovician. (*d*) Diversity history of rhynchonellid families. (*a after J. R. Richardson, Brachiopods, Sci. Amer., 255:100–106, 1986; b and c after R. C. Moore, C. G. Lalicker, and A. G. Fischer, Invertebrate Fossils, McGraw-Hill, New York, 1952; d after J. J. Sepkoski, Jr., A Compendium of Fossil Marine Animal Families, 2d ed., Milwaukee Public Mus. Contrib. Biol. Geol., no. 83, 1992*)

Rhynchonelliformea

One of three subphyla currently recognized in the phylum Brachiopoda, a clade of lophophore-bearing protostome metazoan animals. The name is derived from the stratigraphically oldest subclade with extant representatives, the order Rhynchonellida (**Fig. 1**). Rhynchonelliforms constitute over 90% of the more than 4500 named brachiopod genera, most of which (over 95%) are extinct. Rhynchonelliformea includes

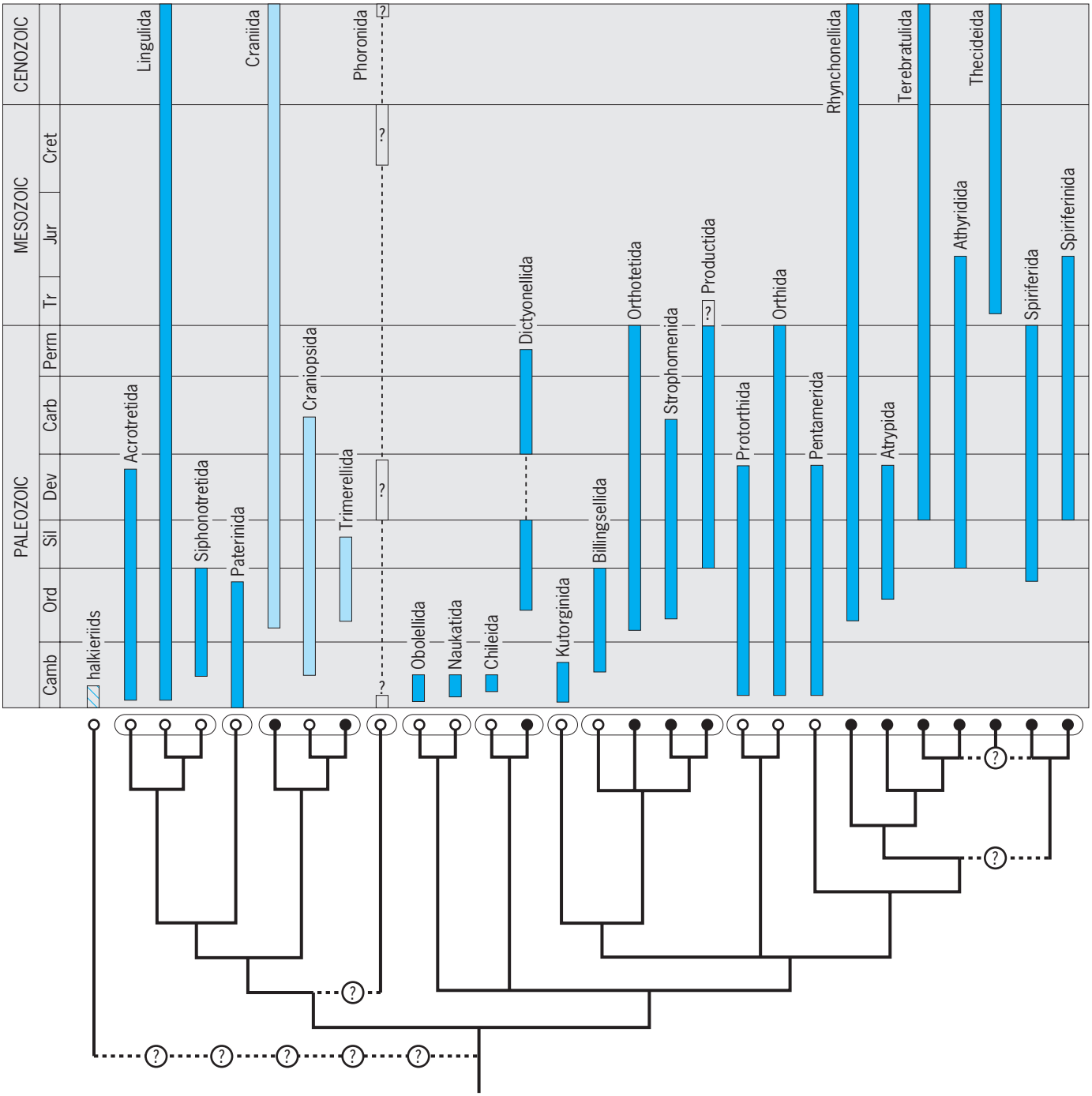

Fig. 1. Stratigraphic ranges and consensus cladogram of Brachiopoda, illustrating one hypothesis of phylogenetic relationship among the 26 orders of brachiopods currently recognized, and phoronids, constructed from data derived primarily from analyses of Cohen and Weydmann (2005) using molecular sequence data from living brachiopods and phoronids; Williams et al. (2000), Holmer and Popov (in Williams et al., 2000), and Popov et al. (in Williams et al., 2000), using morphological data from mostly Cambrian and Ordovician taxa; Carlson (1995) using morphological data from Recent brachiopods; and Carlson and Leighton (in Brunton et al., 2001) using morphological and stratigraphic data together for all rhynchonelliform suborders. Dark color of stratigraphic ranges indicates rhynchonelliforms; light color, linguliforms; and cross-hatched bar at far left, craniiforms. Open circles immediately below stratigraphic ranges indicate those orders first appearing in the Cambrian; closed circles, those first appearing in the post-Cambrian; elongated ellipses surrounding the circles identify the eight brachiopod classes and phoronids; encircled question marks indicate uncertainties in topology. (*Adapted from A. Williams and S. J. Carlson, in press, in R. L. Kaesler, ed., Treatise on Invertebrate Paleontology, pt. H, vol. 6, Geological Society of America and University of Kansas*)

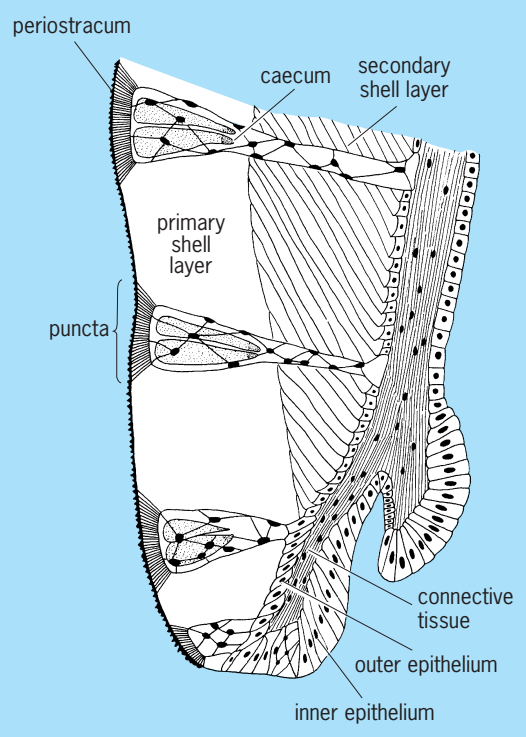

Fig. 2. Generalized section of a punctate rhynchonelliform shell and epithelium at anterior mantle margin. (*After R. L. Kaesler, ed., Treatise on Invertebrate Paleontology, pt. H, vol. 1, Geological Society of America and University of Kansas, 1997*)

all taxa formerly referred to the class Articulata, in addition to several groups newly discovered as fossils, formerly placed in class Inarticulata, or of uncertain taxonomic affiliation. Living Rhynchonelliformea are monophyletic, as confirmed by molecular and morphological data. Phylogenetic relationships among the basal, extinct rhynchonelliforms (Fig. 1) are currently being evaluated.

All rhynchonelliforms have an organic-rich, bivalved calcite shell and can be distinguished from other brachiopods by the following shared, derived characters: the presence of a fibrous secondary shell layer (that transforms to other shell fabrics in some derived rhynchonelliform clades) (**Fig. 2**), a pedicle without a coelomic core (**Fig. 3**), a diductor (valve opening) muscle system (**Fig. 4**), and the presence of articulatory structures between the two valves (**Fig. 5**).

The rhynchonelliform shell is composed of a thin, outer, cryptocrystalline primary layer and a thicker, inner secondary layer of membrane-bound fibers (anvil-shaped in cross section), both mineralized by mantle epithelium. Regularly arranged cylindrical perforations of the shell called punctae (Fig. 2) have evolved more than once among the rhynchonelliforms (as in Terebratulida and Spiriferinida), and contain papillose mantle-tissue extensions called caeca. An unmodified, impunctate fibrous shell structure appears to be primitive for the subphylum (as in Rhynchonellida). A cross-bladed laminar shell structure evolved secondarily from a fibrous microstructure within the strophomenates; it commonly contains conical deflections of the secondary shell layer called pseudopunctae, which point inwardly and appear as tubercles on the valve interior.

The pedicle is a tough but flexible, nonmineralized, cuticle-covered, stalklike appendage by which the animal attaches to the substrate (Fig. 3). It emerges through an indentation of the posterior margin of the ventral valve called a delthyrium (or foramen in derived rhynchonelliforms). The

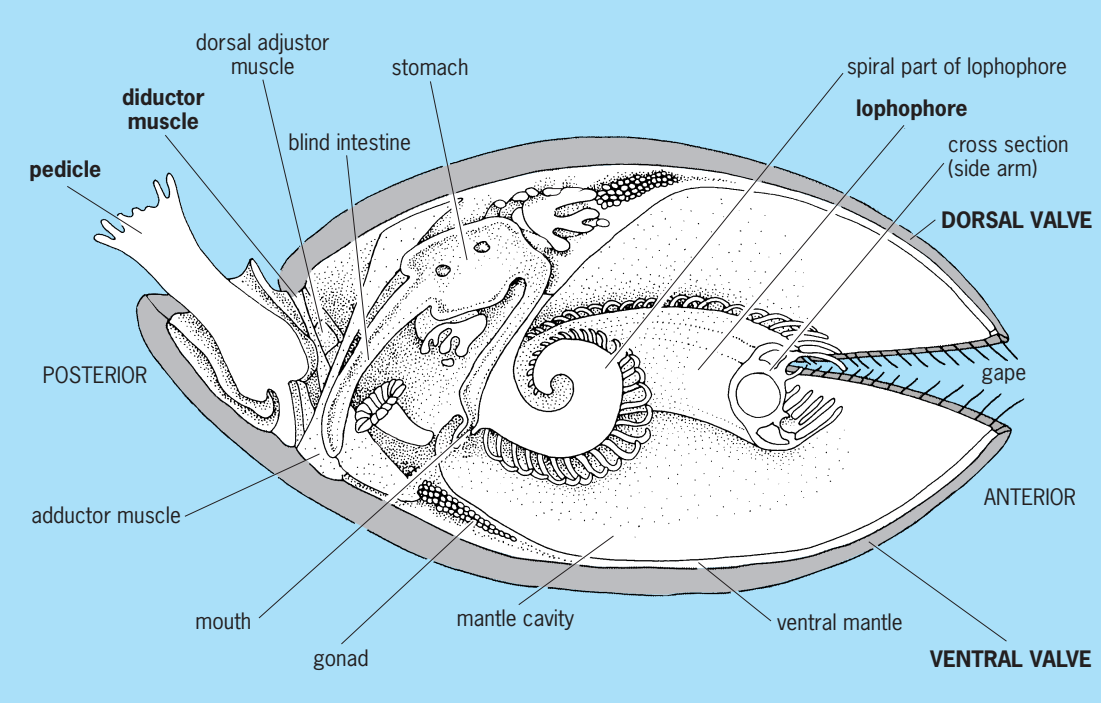

Fig. 3. Diagrammatic representation of the principal organs of *Terebratulina caputserpentis*. (*After R. L. Kaesler, ed., Treatise on Invertebrate Paleontology, pt. H, vol. 1, Geological Society of America and University of Kansas, 1997*)

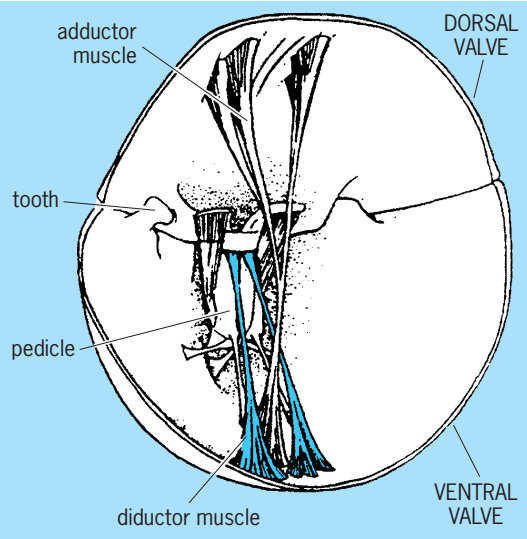

Fig. 4. Musculature of *Gryphus vitreus* in anterolateral view, dorsal valve above. (*After R. L. Kaesler, ed., Treatise on Invertebrate Paleontology, pt. H, vol. 1, Geological Society of America and University of Kansas, 1997*)

pedicles of all extant rhynchonelliforms contain a core of connective tissue rather than a coelomic cavity, develop from a larval rudiment (not as an extension of the ventral body wall), and can be moved relative to the shell by external adjustor muscles. It is inferred that most extinct rhynchonelliforms possessed similar pedicles. The productides and thecideides, however, lost the pedicle as adults as evidenced by the lack of any delthyrial opening.

Diductor muscles (most commonly in two pairs) contract to open the valves by rotation about a hinge axis defined either by a straight, strophic or curved, astrophic hinge line (Figs. 4 and 5). They typically originate at a posterior projection of the dorsal valve called the cardinal process and insert broadly on either side of the paired adductor (closing) muscles on the ventral valve interior.

The oldest rhynchonelliforms possessed a rudimentary valve articulation (as in Kutorginida). The valves were in contact and rotated about a hinge axis, but dentition was poorly developed. Articulation among derived rhynchonelliforms is developed as two ventral teeth that fit into dorsal sockets on either side of the delthyrium. Tooth-and-socket articulation may be either noninterlocking (the primitive deltidiodont state, as in Orthida) or interlocking (the derived cyrtomatodont state, as in all extant rhynchonelliforms). Some strophomenates possess a row of denticles rather than a single pair of teeth, and most productides have lost dentition entirely. *See* ORTHIDA; STROPHOMENIDA.

Recent revision of the classification of brachiopods has attempted to incorporate more interpretable phylogenetic structure. Phylogenetic analyses have been completed on the extant taxa (with a broad range of characters, but very few taxa), the extinct taxa (with shell characters only, but many more taxa), and a combination of both. Disagreements about relationships still remain, but disparate sources of information are converging on a general pattern of evolutionary relationships (Fig. 1). Rhynchonelliformea includes 19 orders organized in five classes.

Phylum Brachiopoda
- Subphylum Linguliformea
- Craniiformea
- Rhynchonelliformea
  - Class Chileata
    - Order: Chileida
    - Dictyonellida
  - Class Obollellata
    - Order: Obolellida
    - Naukatida
  - Class Kutorginata
    - Order: Kutorginida
  - Class Strophomenata
    - Order: Strophomenida

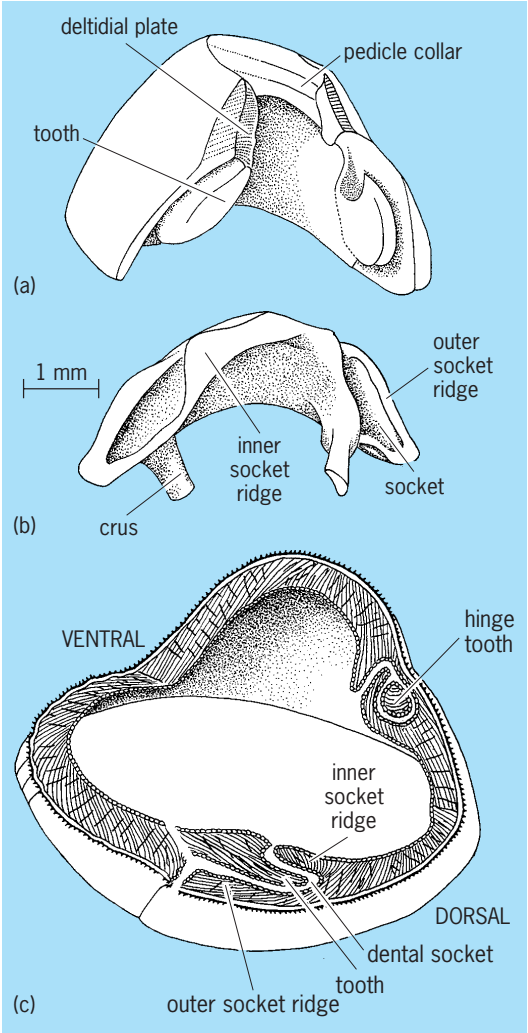

Fig. 5. Articulatory features of partial, posteriormost (*a*) ventral valve and (*b*) dorsal valve of *Terebratulina caputserpentis*; (*c*) view of combined transverse and submedian sections through valves. (*After R. L. Kaesler, ed., Treatise on Invertebrate Paleontology, pt. H, vol. 1, Geological Society of America and University of Kansas, 1997*)

Order: Productida
Orthotetida
Billingsellida
Class Rhynchonellata
Order: Protorthida
Orthida
Pentamerida
Rhynchonellida
Atrypida
Spiriferida
Spiriferinida
Thecideida
Athyridida
Terebratulida

Rhynchonelliform morphologies in the Cambrian are already rather diverse, as evidenced by the presence of eight different orders shortly after their first stratigraphic appearance. Molecular phylogeny research provides strong evidence for a clade of craniiforms and linguliforms together as the sister group to rhynchonelliforms.

Living rhynchonelliforms are distributed globally, and enjoy their greatest abundance and diversity in temperate latitudes. They live from rather deep intertidal to abyssal depths, reaching greatest abundance and diversity in the shallow to relatively deep (few hundred meters) subtidal. Extinct rhynchonelliforms, particularly in the Paleozoic Era, were far more abundant and diverse than today in virtually all geographic realms, and lived at nearly all depths. Fossil rhynchonelliforms have a continuous stratigraphic record from the Early Cambrian Period (Atdabanian Stage), over 540 million years ago, up to the present day. Cambrian diversity was low but significant. After a dramatic diversification in the Ordovician Period, generic diversity remained high through the Paleozoic Era. Rhynchonelliform diversity plummeted in the Permian extinction event, and has remained low (but very slowly increasing) ever since. *See* BRANCHIOPODA; BRACHIOPODA; PENTAMERIDA; SPIRIFERIDA; TEREBRATULIDA.

Sandra J. Carlson

Bibliography. C. H. C. Brunton, L. R. M. Cocks, and S. L. Long (eds.), *Brachiopods Past and Present*, Systematics Association Spec. Vol. Ser. 63, Taylor and Francis, London, 2001; S. J. Carlson, Phylogenetic relationships among extant brachiopods, *Cladistics*, 11:131–197, 1995; B. L. Cohen and A. Weydmann, Molecular evidence that phoronids are a subtaxon of brachiopods (Brachiopoda: Phoronata) and that genetic divergence of metazoan phyla began long before the Early Cambrian, *Organisms, Diversity Evolut.*, 5(3), 2005 (in press); M. A. James et al., Biology of living brachiopods, *Adv. Mar. Biol.*, 28:175–387, 1992; M. J. S. Rudwick, *Living and Fossil Brachiopods*, Hutchinson, London, 1970; A. Williams et al., Part H: *Brachiopoda*, revised, vols. 2–5: Linguliformea, Craniiformea, and Rhynchonelliformea (part), in R. L. Kaesler (ed.), *Treatise on Invertebrate Paleontology*, Geological Society of America and University of Kansas, Boulder and Lawrence, 2000–(in press) 2005.

Rhynchosauria

An order of herbivorous diapsid reptiles in the infraclass Archosauria, limited to the Triassic System but with a worldwide distribution. Rhynchosaurs were pig- or sheep-sized quadrupedal reptiles that were common in the Middle and Upper Triassic Series of India, South America, Europe, and eastern Africa. Except for the earliest South African genus, *Mesosuchus*, they are characterized by multiple rows of teeth on both the upper and lower jaws that were fused into deep sockets. Teeth were not replaced as they were worn but were added posteriorly as the jaw grew. In most genera, the premaxillae are devoid of teeth and overhang the front of the lower jaw like a beak. The external nostril is median rather than paired (see **illus.**).

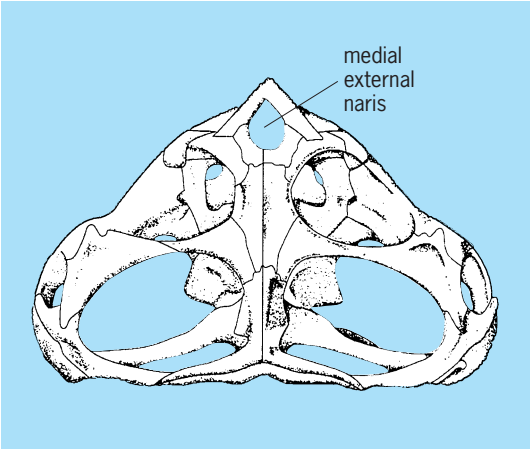

Skull (dorsal view) of the rhynchosaur *Paradapedon*.

Rhynchosaurs were elements of the early archosauromorph radiation that included the protorosaurs and primitive archosaurs. The structure of the ankle is nearly identical in the early members of these groups, but later rhynchosaurs enlarged the centrale which contributed to a simple hinge joint between the lower leg and the tarsals.

It was long thought that the Triassic rhynchosaurs were related to the primitive living genus *Sphenodon*, the New Zealand tuatara, which also has a beaklike premaxilla and a diapsid pattern of the temporal region. However, the teeth of *Sphenodon* differ significantly in being fused to the margin of the jaw rather than being set in sockets, and the beak is formed by premaxillary teeth, not by the bones of the premaxillae. The nature of rhynchosaurian ankle structure and dentition indicates affinities with archosaurs and protorosaurs. *See* PROTOROSAURIA.

Rhynchosaurs were locally common in the Late Triassic but are unknown in the Jurassic. Their rapid extinction may have resulted from changes in the vegetation on which they fed, or from predation from the expanding community of predaceous archosaurs, including early dinosaurs. *See* ARCHOSAURIA; DIAPSIDA; REPTILIA.

Robert L. Carroll

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1987.

Rhyniophyta

The first division under the subkingdom Embryobionta, in the classificatory system of A. Cronquist, A. Takhtajan, and W. Zimmermann. The bryophytes and vascular plants are included in this subkingdom. The category Rhyniophyta was devised for the relatively simple Silurian-Devonian vascular plants long held to be ancestral to other groups of vascular plants and usually referred to as Psilophytales. These plants have leafless stems and lack roots; their general morphological structure is not complex. Nevertheless, studies have shown that they fall into distinct groups on the basis of such characteristics as maturation of primary xylem, complexity of primary xylem in the stem, position of sporangia (terminal versus lateral), shape of sporangia (fusiform versus reniform), dehiscence of sporangia (longitudinal versus distal), and presence or absence of extensive lateral branch systems bearing large masses of terminal sporangia. The three classes of Rhyniophyta currently recognized are Rhyniopsida, Zosterophyllopsida, and Trimerophytopsida. *See* EMBRYOBIONTA; PSILOPHYTALES; RHYNIOPSIDA; TRIMEROPHYTOPSIDA; ZOSTEROPHYLLOPSIDA. Harlan P. Banks

Rhyniopsida

The earliest demonstrable vascular land plants, appearing in Silurian (mid-Ludlovian) time. Their small, leafless axes were usually branched dichotomously in three planes; adventitious and perhaps pseudomonopodial branching also occurred. Sporangia were usually terminal on the main axes. Some terminated in lateral branches, and some were subtended by adventitious branches. The sporangia were globose, almost reniform, elongate-ellipsoidal, or cylindrical; some were columellate and branched. Dehiscence was unspecialized; some opened by an apical slit, and some by a modified distal region. The spores were simple and produced in tetrads. A possible gametophyte plant, associated with *Rhynia*, was erect, leafless, and terminated by a gametangiophore bearing antheridia and possible archegonia. The basal parts, where known, were rhizomatous, with rhizoids but no true roots. Stomata have been found on some axes and sporangia. The xylem, where known, was centrarch, and consisted of annular elements. Evidence of phloem sieve cells is unconvincing.

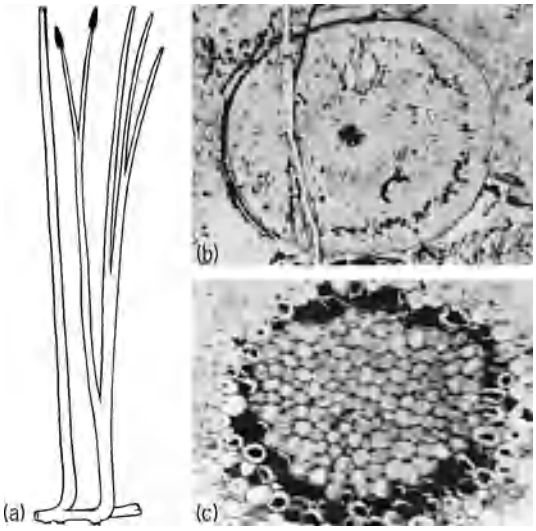

***Rhynia*. (*a*) Reconstruction showing leafless axis, basal rhizome with rhizoids (hairs), and fusiform, terminal sporangia. (*b*) Photomicrograph of transverse section of petrified stem. (*c*) Central xylem (vascular) strand. (*After Walton, An Introduction to the Study of Fossil Plants, A. and C. Black, 1940*)**

Rhynia (see **illus.**) and *Horneophyton* are best known because of their cellular preservation. *Cooksonia*, the oldest genus, is numerically the most abundant. The group flourished in Late Silurian–Early Devonian time and may then have disappeared, although one genus is reported in Upper Devonian strata. *Cooksonia*, *Rhynia*, *Horneophyton*, and *Renalia* are the best-established genera. *Steganotheca*, *Salopella*, *Eogaspesiea*, *Dutoitea*, and some *Psilophytites* and *Hostinella* are probably rhyniophytes. *Taeniocrada*, *Hedeia*, and *Yarravia*, when better known, may belong elsewhere. Green algae were probable precursors, and the Trimerophytopsida appear to represent a higher evolutionary level. *See* PALEOBOTANY; RHYNIOPHYTA. Harlan P. Banks

Rhyolite

A very light-colored, aphanitic (not visibly crystalline), volcanic rock that is rich in silica and broadly equivalent to granite in composition. Migration of rhyolitic magma through the Earth's crust, which causes much of the Earth's explosive and hazardous volcanic activity, represents a major process of chemical fractionation by which continental crust grows and evolves. *See* GRANITE.

Structure and eruptive style. Rhyolites are formed by the process of molten silica-rich magma flowing toward the Earth's surface. Small differences in this process, notably those related to the release of gas from the magma at shallow depth, produce extremely diverse structural features. Air-fall tephra eruptions form when molten rock foams as it rises to the surface, as a result of the gas being released; the foam breaks into a spray that is ejected into the air at high speeds. Large porous fragments making up the tephra layer are known as pumice; very small particles are known as ash. Some of the magma may not become airborne during an explosive event. Instead, the magmatic spray of ash particles, larger pumice fragments, and gas forms a dense, rapidly moving, ground-hugging fluid. Rhyolitic ash flow sheets formed in this way exhibit a number of characteristic features; these include glass shards, collapsed pumices, densely welded basal vitrophyres, welded devitrified interiors, and vapor-phase-altered zones. If the gas either has a low concentration or can escape easily in the shallow environment, rhyolitic magma sometimes remains intact and flows along the surface as a lava flow. The very viscous and slow-moving magma forms very thick flows and

domes that typically exhibit flow banding. *See* LAVA; MAGMA; PUMICE; TUFF; VOLCANIC GLASS; VOLCANO.

Composition, mineralogy, and texture. The high silica content gives rhyolitic lava a correspondingly high viscosity; this hinders crystallization and often causes young rhyolite to be a mixture of microcrystalline aggregates and glassy material. Rhyolites are generally porphyritic with numerous large crystals (phenocrysts) of sanidine (high-temperature form of potassium feldspar) and quartz disseminated throughout the aphanitic matrix. The crystalline portion of this groundmass consists of quartz intergrown with alkali feldspar. Rhyolite also contains minor amounts of dark minerals, including biotite, hornblende, or pyroxene. Many lavas of rhyolitic composition occur in the glassy state as obsidian or pitchstone. Obsidian is a black natural glass with a conchoidal fracture, whereas pitchstone has a dull appearance. Pitchstone contains up to 10% water and owes its appearance to secondary hydration and devitrification, the conversion of glass to a crystalline aggregate. Because of the glassy nature of most rhyolites, they are best characterized by chemical analysis. They typically have 70–75 wt % silicon dioxide (SiO_2) and more potassium oxide (K_2O) than sodium oxide (Na_2O). *See* FELDSPAR; OBSIDIAN; PITCHSTONE; PORPHYRY; QUARTZ.

Occurrence and origin. Rhyolite is one of the most common volcanic rocks in continental regions; it is virtually absent in the ocean basins. The rock often occurs in large quantities associated with andesite and basalt. It is common in environments ranging from accretionary prisms at continental margins to magmatic arcs related to subduction zones. Rhyolite is also prevalent in extensional regions and hot spots in continental interiors.

Noteworthy are the large and explosive caldera-forming ash flow eruptions. Although no eruption of this magnitude has been observed in historical times, they were quite common in the recent geological past, especially during the middle Tertiary time when vast areas of North, Central, and South America were blanketed with ash-flow sheets covering thousands of square kilometers with thicknesses of hundreds of meters.

The size and scope of these very explosive eruptions have been documented with studies at Yellowstone National Park (northwest-central United States), where three such eruptions occurred on a fairly regular schedule starting roughly 2×10^6 years ago. The first eruption was the largest. Within a few days, 2500 km^3 (600 mi^3) of dust, ash, and lava erupted, an amount 15,000 times greater than the amount blown out by Mount Saint Helens, Washington, in 1980. Similar but smaller eruptions occurred again 1.2×10^6 years ago and 600,000 years ago. Long Valley Caldera, California, and Valles Caldera, New Mexico, are the sites of other relatively recent, similar devastating eruptions. *See* CALDERA.

A variety of theories have been developed to explain the origin of rhyolite. Because it is largely confined to continental crust, many geologists believe that its origin is related to partial melting of preexisting continental crust. In addition, because it is often associated with basalt and andesite, its origin may be tied to the evolution of these mantle-derived magmas, perhaps reflecting differentiation of the parent magma. More complicated models involving melting of the continental crust by intrusion of hot basaltic magma, fractional crystallization of the basalt, magma mixing, and assimilation of crustal material have been developed to explain the origin of rhyolite and associated rocks. *See* ANDESITE; BASALT; EARTH CRUST.

Research drilling. The area around Mammoth Lakes in northern California contains some of the youngest and most interesting volcanic features in the United States. The largest of these is the Long Valley Caldera, which formed 700,000 years ago during a catastrophic eruption of pyroclastic pumice and ash. The Inyo Domes chain is a much smaller, younger, and simpler feature. It is a line of rhyolitic lava domes, craters, and fractures that runs north-south across the rim of the caldera. The Inyo Domes chain last erupted 600 years ago with an event that rivaled the Mount Saint Helen's eruption of 1980. Six hundred years is a very short time with respect to geological processes; volcanoes that were active 600 years ago have changed very little, so that this region is an ideal place to study how rhyolite magma is formed and erupts.

As part of the Continental Scientific Drilling Program, research holes were drilled into the volcanoes at Inyo to sample the actual intrusions that fed these young eruptions. A special ring-shaped drill bit cut out a core of rock from the hole. Analyses of this core have provided information for constructing chemical and mathematical models for the intrusion and eruption of rhyolitic magmas. The drilling program has elucidated the shallow plumbing systems that feed young eruptions. The evidence collected also suggested that the nature of volcanic eruptions, quiet or explosive, may be controlled by the manner in which gas escapes at shallow depths, not by the original gas content of the magma, as was previously believed. Finally, identification of striking chemical zonation in the conduit to the dome may ultimately reveal a great deal about how magma flows in the Earth's crust. *See* IGNEOUS ROCKS.

Leland W. Younker

Bibliography. F. Barker, Granites and rhyolites, *J. Geophys. Res.*, special issue, 86(B11):10131–10638, 1981; I. S. E. Carmichael, F. J. Turner, and J. Verhoogen, *Igneous Petrology*, 1974; J. H. Fink (ed.), *The Emplacement of Silicic Domes and Lava Flows*, Geolog. Soc. Amer. Spec. Pap. 212, 1987; A. R. McBirney, *Igneous Petrology*, 2d ed., 1997; C. S. Ross and R. L. Smith, *Ash-Flow Tuffs: Their Origin, Geologic Relations, and Identification*, USGS Prof. Pap. 366, 1961.

Riboflavin

Also known as vitamin B_2, riboflavin is widely distributed in nature, and is found mostly in milk, egg white, liver, and leafy vegetables. It is a water-soluble

yellow-orange fluorescent pigment with the structural formula shown below. It is stable to acid and

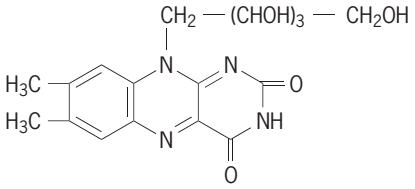

oxidation, but is rapidly destroyed by alkali at elevated temperatures and by light. Although little riboflavin is usually destroyed during cooking, half of the riboflavin in bottled milk exposed to the sun may be destroyed in 2 h.

Riboflavin is determined by microbiological methods using the lactic acid–forming organism *Lactobacillus casei*, which requires it for growth. Riboflavin is also determined by chemical methods in which its fluorescent properties are used.

Dietary requirement. A dietary source of riboflavin is required by all nonruminant animal species studied. Riboflavin deficiency results in poor growth and other pathologic changes in the skin, eyes, liver, and nerves. Riboflavin deficiency in humans is usually associated with a cracking at the corners of the mouth called cheilosis; inflammation of the tongue, which appears red and glistening (glossitis); corneal vascularization accompanied by itching; and a scaly, greasy dermatitis about the corners of the nose, eyes, and ears.

Riboflavin is found in all tissues with its concentration usually paralleling metabolic activity. The vitamin functions biochemically in two coenzymes, flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). Enzymes containing these coenzymes are called flavoproteins. In general, these enzymes participate in oxidation-reduction reactions by accepting hydrogen ions (protons) and electrons from one substrate and transferring them to another. Over 14 riboflavin-containing enzymes have been studied. *See* COENZYME; ENZYME.

Most of the energy-yielding oxidations of food to carbon dioxide and water occur through the cytochrome system, in which the flavoprotein cytochrome reductase is essential. L-Amino oxidase, another flavoprotein, is important in the interconversion of amino acids to nonnitrogenous metabolities of carbohydrate and fat metabolism. *See* CYTOCHROME.

Riboflavin requirements appear to be related to caloric requirements and muscular activity, and are affected by heredity, growth, environment, age, and health. Evidence in animals suggests a need for increased riboflavin in low-protein diets because of a decreased ability of the liver to retain the vitamin. There is also evidence that as a result of intestinal synthesis by bacteria, less riboflavin is required on diets high in carbohydrate than on those high in fat. Human requirements for riboflavin are based primarily on urinary excretion data and studies in which riboflavin deficiency has been experimentally produced. Riboflavin allowances of the National Research Council have been set at 0.07–0.1 mg/0.75 kg of body weight. Former allowances using a factor of 0.025 mg/g of protein allowance provided similar amounts of the vitamin. Stanley N. Gershoff

Industrial production. Riboflavin is produced commercially by either direct chemical synthesis or fermentation.

Chemical synthesis. Industrial syntheses of riboflavin generally proceed along the lines of the original Karrer method with modifications by individual producers. A typical manufacture starts with D-ribose, which is condensed with 1,3,4-xylidine and simultaneously hydrogenated to form *N*-D-ribitylxylidine. The ribitylxylidine is coupled with diazotized aniline to produce 1,2-dimethyl-4-D-ribitylamino-5-phenyl azo benzene. This azo compound is hydrogenated to 1,2-dimethyl-4-D-ribitylamino-5-aminobenzene, and then is condensed with a mixture of alloxan and alloxantin to riboflavin. By another method, the azo compound is condensed directly with barbituric acid. Leo A. Flexser

Fermentation. Riboflavin produced commercially by fermentation utilizes the synthetic ability of bacteria, yeasts, or fungi. The outstanding organisms for production of riboflavin are the two closely related ascomycete fungi, *Eremothecium ashbyii* and *Ashbya gossypii*. The inoculum is started from slants or from spores dried on sand, and after one or two flask stages is carried through one or two tank inoculum stages.

The final fermentation is carried out in tanks with a capacity of 10,000–100,000 gal (40,000–400,000 liters) with a medium suitable for the organism being used. For *Eremothecium* usually stillage (still slops) from the ethyl alcohol fermentation with skim milk, soybean meal, or casein added is used as a proteinaceous source; a carbohydrate source, such as maltose, sucrose, or glucose, is added. For *Ashbya*, commercially used media may contain corn-steep liquor, and usually also some animal protein such as crude peptones, animal-stick liquor, or fish-stick liquor; the carbohydrate sources are similar to those used for *Eremothecium*. The medium is aerated and usually agitated during fermentation for 96–120 h; optimum titers may be 3–6 g of riboflavin or more per liter. *See* ASCOMYCOTA; CARBOHYDRATE; CASEIN; ETHYL ALCOHOL; INDUSTRIAL MICROBIOLOGY; SOYBEAN.

Riboflavin may be recovered for animal feed supplements by evaporation of the whole broth in multiple-effect evaporators, followed by drum or spray drying. For drug and fine foods uses, pure crystalline riboflavin is isolated by heating the fermentation broth, filtration, and precipitation of the riboflavin with dithionite (hydrosulfite) followed by several purification steps, including crystallization. Ralph E. Bennett

Bibliography. G. F. Combs, Jr., *The Vitamins: Fundamental Aspects in Nutrition and Health*, 1997; H. J. Peppler and D. Perlman (eds.), *Microbial Technology*, vol. 2: *Fermentation Technology*, 2d ed., 1979; B. Volesky and J. Votruba, *Modeling and Optimization of Fermentation Processes*, 1992.

Ribonuclease

A group of enzymes, widely distributed in nature, which catalyze hydrolysis of the phosphodiester bonds in ribonucleic acid (RNA). The sites of hydrolysis may vary considerably, depending upon the specificity of the particular enzyme. Ribonuclease isolated from the pancreas appears to cleave RNA in two steps (**Fig. 1**). Cleavage occurs predominantly at P-O bonds which are attached to the 3′-OH groups of pyrimidine nucleotides. A number of plant ribonucleases catalyze similar reactions but are relatively nonspecific with regard to the nature of the base moiety. On the other hand, a ribonuclease from *Bacillus subtilis* appears to require purine rather than pyrimidine bases as part of the susceptible nucleotides. Even more stringent specificity requirements are found in two ribonucleases from *Aspergillus oryzae*, ribonuclease T_1 and ribonuclease T_2, which specifically hydrolyze the phosphodiester bonds adjacent to 3′-guanylate and 3′-adenylate groups, respectively. Differences in specificity for the site of cleavage have led to the use of these various ribonucleases as tools in determining the structure and chemistry of RNA.

The most thoroughly studied of the ribonucleases has been that isolated from bovine pancreas. This enzyme is the subject of the following presentation.

The systematic name for pancreatic ribonuclease proposed by the Enzyme Commission is polyribonucleotide 2-oligonucleotidotransferase (cyclizing). The small size and great stability of ribonuclease, coupled with the ease of isolation of pure enzyme in large quantities, have contributed notably to its successful characterization. Since the elucidation of the amino acid sequence of ribonuclease, much information has been compiled with regard to the three-dimensional structure of the enzyme and to specific regions of the molecule that are catalytically important. *See* ENZYME; NUCLEIC ACID.

Discovery, isolation, and purification. Pancreatic ribonuclease was discovered in 1920 by W. Jones, who described the hydrolysis of yeast nucleic acid by a heat-stable pancreatic extract. It has now been established that the enzyme is synthesized on the microsomes of the acinar cells in the pancreas and is then distributed throughout the cytoplasm. The techniques employed in the isolation of ribonuclease take advantage of the enzyme's unique stability to extreme conditions of temperature and pH; indeed, this stability has largely facilitated subsequent chemical studies of the protein. M. Kunitz was the first to crystallize ribonuclease (1940), and most commercial preparations of the enzyme have been obtained by means of his procedure or modifications thereof. These crystalline preparations contain at least six distinguishable, enzymatically active components. The major constituent, ribonuclease A, accounts for about 90% of the total activity and may be prepared from the crystalline material by chromatography. *See* CHROMATOGRAPHY; PANCREAS.

Covalent structure. The determination of the primary structure of ribonuclease A represents a milestone in protein chemistry, as ribonuclease was the first enzyme for which such information was obtained. This work was accomplished in 1962 largely through the efforts of W. H. Stein and Stanford Moore and their collaborators, using techniques which have since become routine tools in the structural analysis of a great many proteins. A two-dimensional representation of the covalent structure of ribonuclease A is shown in **Fig. 2**. This drawing indicates only the sequence of amino acids and the positions of the disulfide cross-links; it has no three-dimensional significance. The enzyme consists of a single polypeptide

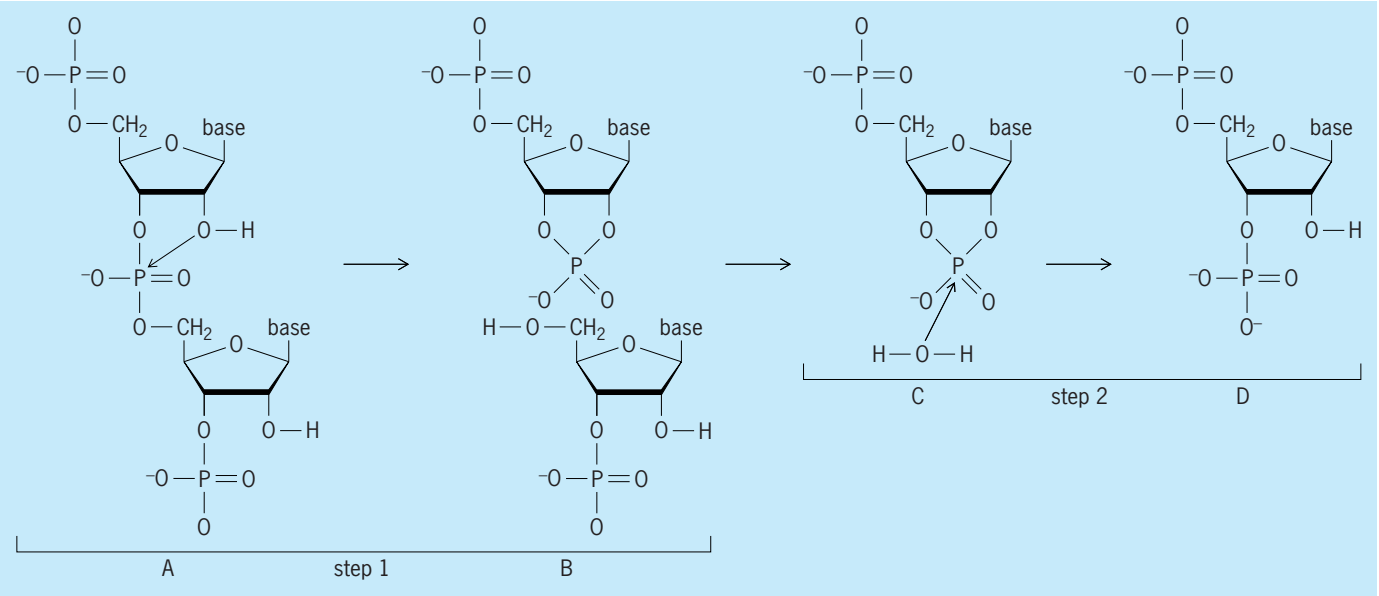

Fig. 1. Cleavage of RNA by pancreatic ribonuclease. In the first step, the oxygen atom of the 2′-OH group of the original, uncut RNA strand (species A) attacks the phosphorus atom of the adjacent phosphate group, cleaving the RNA strand and generating the cyclic 2′,3′-phosphate intermediate (species B). Part of the former RNA strand then detaches from the enzyme, leaving the remaining part of the RNA strand with the 2′,3′-phosphate cyclic end still bound to the enzyme (species C). In the second step, the oxygen atom of a water molecule attacks the phosphorus atom of the cyclic 2′,3′-phosphate group to form a noncyclic 3′-phosphate end (species D).

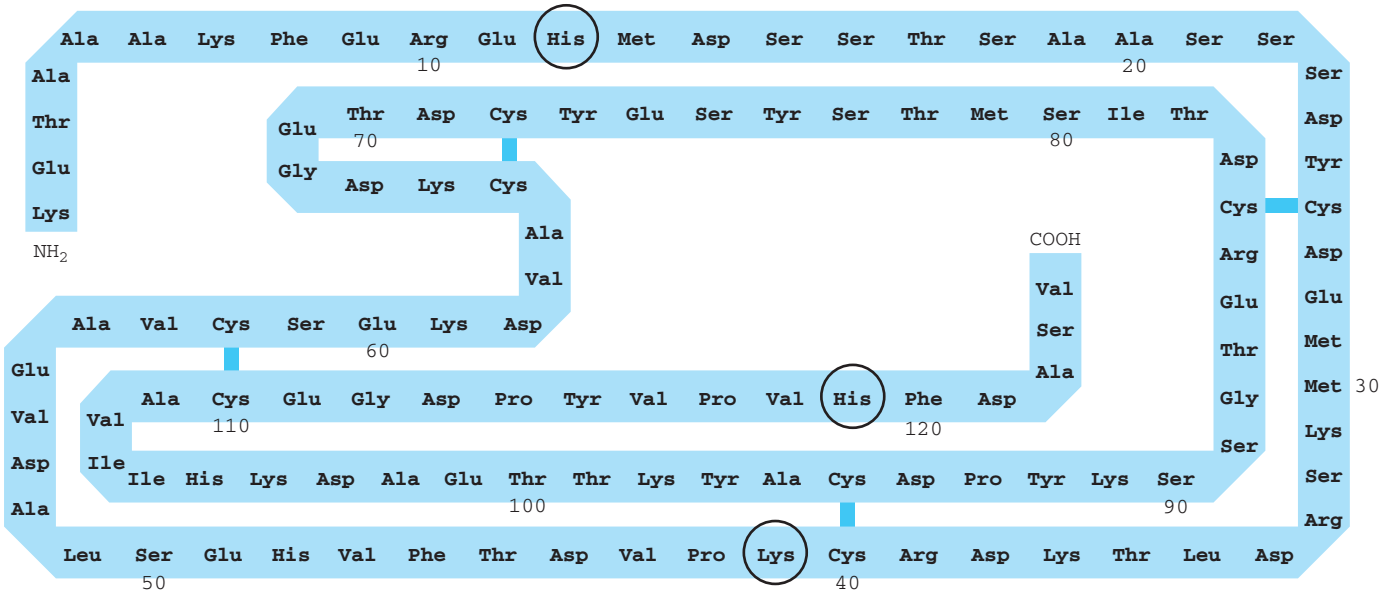

Lys	amino acid lysine	His	histidine	Leu	leucine
Glu	glutamic acid	Met	methionine	Pro	prolamine
Thr	threonine	Asp	aspartic acid	Val	valine
Ala	alanine	Ser	serine	Ileu	isoleucine
Phe	phenylalanine	$AspNH_2$	asparagine	Gly	glycine
Arg	arginine	Tyr	tyrosine		
$GluNH_2$	glutamine	Cys	cysteine		

Fig. 2. Covalent structure of bovine pancreatic ribonuclease A.

chain of 124 amino acid residues containing 8 disulfide bridges between its 8 half-cystine residues. There are no free sulfhydryl groups or tryptophan residues in the enzyme. Particularly noteworthy are the histidines at positions 12 and 119 and lysine-41, all of which are encircled in Fig. 2. Evidence suggests that these residues form part of the active site of ribonuclease. The molecular weight of ribonuclease A calculated from its amino acid composition is 13,683, and this figure is in good agreement with values obtained by physical methods.

Research on ribonuclease has played a prime role in advancing the understanding of protein structure and function; also, it was the first protein to be totally synthesized from its component amino acids.

Tertiary structure. A chemical approach has been utilized with some success in pinpointing various aspects of the three-dimensional conformation of ribonuclease. The amino groups of all 10 lysine residues appear to lie at the surface of the enzyme, although the molecule is resistant to tryptic hydrolysis. Moreover, the enzyme in its native state is not hydrolyzed by a number of proteinases, suggesting that the peptide bonds may participate in hydrogen bonding of some sort which contributes to their stability to enzymatic attack. Tyrosine residues 25 and 97 are buried in the molecule, as are all four methionine residues, the side chains of which probably contribute to hydrophobic binding.

The most definitive information concerning the tertiary structure of the enzyme has been obtained by x-ray crystallography. Ribonuclease S is a fully active enzyme formed from ribonuclease A by hydrolysis with subtilisin, a proteinase which cleaves a single peptide bond in the enzyme between alanine-20 and serine-21. The tertiary structures of the two are essentially the same except that, in ribonuclease S, the residues involved in the subtilisin cleavage are separated by 10–15 angstroms (1–1.5 nanometers). A computer-generated illustration of the polypeptide chain folding in the model of ribonuclease A is shown in **Fig. 3**. Thus far, the x-ray models are compatible with, and confirm most of the data obtained by, chemical means. The positions of the four disulfide cross-links are the same as in the covalent representation (Fig. 2). Histidine-12, histidine-119, and lysine-41

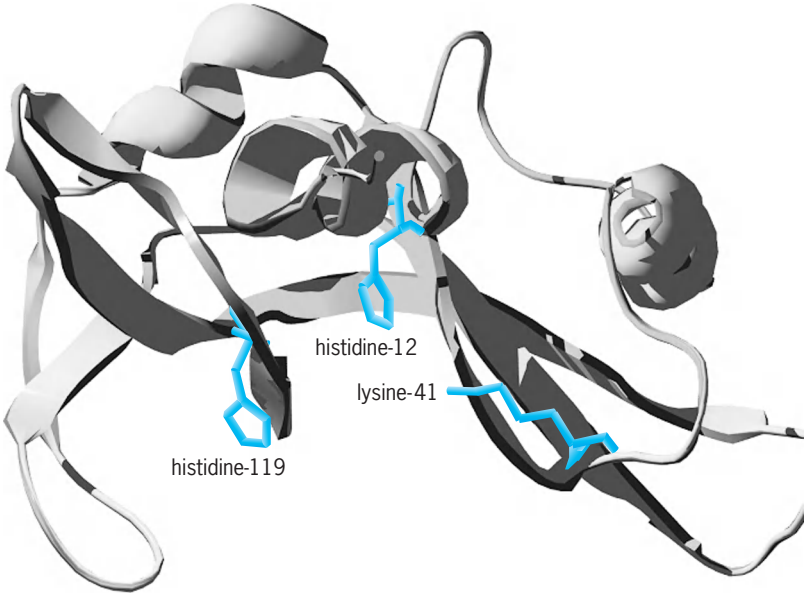

Fig. 3. Computer-generated illustration of the main chain folding in the ribonuclease A molecule.

are close together, in the region of the molecule where substrate is bound. Ribonuclease S contains 15% helix, in close agreement with the value predicted by optical methods (17%). The ribonuclease A molecule is kidney-shaped, with approximate dimensions of 38 × 28 × 22 Å (3.8 × 2.8 × 2.2 nm). The disposition of charged side chains is toward the surface of the enzyme, while neutral groups are directed toward the interior. Much of the structure is exposed, and there is no part of the molecule which is shielded from the surrounding medium by more than a single layer of the main chain.

Activity. It has been postulated that the cleavage of RNA by ribonuclease occurs in two steps (Fig. 1), and activators and inhibitors of the enzyme may exert their effects on either phase or both phases of the overall reaction. Most of the kinetic studies of the enzyme have been carried out with pyrimidine nucleoside-2′,3′-cyclic phosphates as substrates, in which cases only the second step of the reaction has been assayed. Ribonuclease is active over a broad range of pH from 5 to 9, with optimal activity near 7. The enzyme is irreversibly inhibited by a number of alkylating agents and reversibly inhibited by Zn^{2+}, Cu^{2+}, phosphate, pyrimidine 2′- and 3′-phosphates, and a number of polyanions.

Active site. The active site is that region of the enzyme that interacts with the substrate and at which catalysis takes place. Definitive evidence has been presented which suggests that two of the four histidine residues in ribonuclease are at the active site and that they participate in the function of the enzyme. Different chemical groups may be introduced at a number of positions in ribonuclease without destroying, or even altering, its activity. However, if the histidine residues at positions 12 or 119 are chemically modified by a variety of alkylating agents, catalytic activity is abolished. The chemical modification of these highly reactive histidines is blocked by substances that bind to the active site of the enzyme. Moreover, if histidine-12 is chemically modified, it is no longer possible to react at histidine-119 and vice versa. On the basis of such chemical evidence it was postulated that histidine-12 and histidine-119 are close together even though widely separated in the polypeptide chain, and this has since been confirmed by x-ray analysis (Fig. 3). Evidence based upon chemical modification studies and x-ray analysis indicates that the lysine residue at position 41 is also near the active site, although it is not known whether its amino group participates in catalysis.

The active site would thus appear to consist of a cationic cluster involving the side chains of amino acid residues 119, 12, and 41. The positively charged site undoubtedly plays some role in the electrostatic attraction of binding of the negatively charged substrate. *See* PROTEIN; X-RAY CRYSTALLOGRAPHY.

Robert L. Heinrikson; M. Todd Washington

Bibliography. R. H. Abeles, P. A. Frey, and W. P. Jencks, *Biochemistry*, Jones and Bartlett, Boston, 1992; T. M. Devlin (ed.), *Textbook of Biochemistry with Clinical Correlation*, 5th ed., Wiley, New York, 2001; D. Voet and J. G. Voet, *Biochemistry*, 3d ed., Wiley, New York, 2004.

Ribonucleic acid (RNA)

One of the two major classes of nucleic acid, mainly involved in translating the genetic information carried in deoxyribonucleic acid (DNA) into proteins. Various types of ribonucleic acids function in protein synthesis: transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs) function in the synthesis of all proteins, while messenger RNAs (mRNAs) are a diverse set, each member of which acts specifically in the synthesis of one protein. Messenger RNA is the intermediate in the usual biological pathway DNA → RNA → protein. Ribonucleic acid is a very versatile molecule, however. Other types of RNA serve other important functions for cells and viruses, such as the involvement of small nuclear RNAs (snRNAs) in mRNA splicing. In some cases, RNA performs functions typically considered DNA-like, such as serving as the genetic material for certain viruses, or roles typically carried out by proteins, such as RNA enzymes or ribozymes. *See* DEOXYRIBONUCLEIC ACID (DNA).

Structure and synthesis. RNA is a linear polymer of four different nucleotides (**Fig. 1**). Each nucleotide is composed of three parts: a five-carbon sugar known as ribose, a phosphate group, and one of four bases attached to each ribose, adenine (A), cytosine (C), guanine (G), or uracil (U). The structure of RNA is basically a repeating chain of ribose and phosphate moieties, with one of the four bases attached to each ribose. The structure and function of the RNA vary depending on its sequence and length. *See* NUCLEOTIDE; RIBOSE.

In its basic structure, RNA is quite similar to DNA. It differs by a single change in the sugar group (ribose instead of deoxyribose) and by the substitution of uracil for the base thymine (T). Typically, RNA does not exist as long double-stranded chains as does DNA, but as short single chains with higher-order structure due to base pairing and tertiary interactions within the RNA molecule. Within the cell, RNA usually exists in association with specific proteins in a ribonucleoprotein (RNP) complex.

The nucleotide sequence of RNA is encoded in genes in the DNA, and it is transcribed from the DNA by a complementary templating mechanism that is catalyzed by one of the RNA polymerase enzymes. In this templating scheme, the DNA base T specifies A in the RNA, A specifies U, C specifies G, and G specifies C.

Transfer RNA. These small RNAs (70–90 nucleotides) act as adapters that translate the nucleotide sequence of mRNA into a protein sequence. They do this by carrying the appropriate amino acids to ribosomes during the process of protein synthesis. Each cell contains at least one type of tRNA specific for each of the 20 amino acids, and usually several types. The tRNAs exist in a precisely folded, three-dimensional L-shaped structure (**Fig. 2**). Complementary base-pairing interactions between different parts of the tRNA play a major role in specifying this structure, with additional contributions from higher-order interactions. At one end of the L shape is the three-nucleotide anticodon sequence that binds

its complementary sequence in the mRNA, and at the other end is the amino acid specified by the mRNA codon. The base sequence in the mRNA directs the appropriate amino acid–carrying tRNAs to the ribosome to ensure that the correct protein sequence is made. About 25% of tRNA bases are unusual in being modifications (such as methylations) of the four standard bases. Among other roles, these modified bases are thought to facilitate recognition of the mRNA codon by tRNA. In the tRNA tertiary structure, a variable-length extra arm is accommodated within the hinge region that joins the two parts of the L-shaped structure. Few functions have been ascribed to the extra arm, although in a few cases it is known to contribute essential identity elements (in the form of its overall structure and/or specific nucleotide residues within it) for correct amino acid acceptor activity. *See* PROTEIN; RIBOSOMES.

Ribosomal RNA. Ribosomes are complex ribonucleoprotein particles that act as workbenches for the process of protein synthesis, that is, the process of linking amino acids to form proteins. Each ribosome is made up of several structural rRNA molecules and more than 50 different proteins, and it is divided into two subunits, termed large and small. A bacterial ribosome contains one copy each of 5S rRNA (~120 nucleotides) and 23S rRNA (~2900 nucleotides) in its large subunit, and one copy of 16S rRNA (~1540 nucleotides) in its small subunit. These rRNAs are also present in somewhat larger versions in the ribosomes of higher organisms, which also contain an additional rRNA of 5.8S (~160 nucleotides) in their large subunit.

The RNA components of the ribosome account for more than half of its weight. Like tRNAs, rRNAs are stable molecules and exist in complex folded structures. Each of these rRNAs is closely associated with ribosomal proteins and is essential in determining the exact structure of the ribosome. In addition, the rRNAs, rather than the ribosomal proteins, are likely the basic functional elements of the ribosome. They are actively involved in most, if not all, stages of protein synthesis. In this process, they make direct and sometimes transient contacts, via short stretches of base pairing, with the mRNA, tRNA, and each other. Recently, atomic-scale three-dimensional structures have been determined for the large and small ribosomal subunits. This work has shown the absence of any protein moiety within 18 Å of the site of peptide bond formation (within the peptidyltransferase center of the large subunit), confirming that the ribosome is indeed an RNA enzyme (ribozyme).

The nucleotide sequences of many rRNAs from different organisms have been characterized. The secondary stem-loop structures adopted by the RNAs are more highly conserved than the overall primary sequence, indicating that the higher-order structure is the functional one. Because these RNAs are present and perform the same central function in all known organisms, they are well suited for defining evolutionary relations. Sequence comparisons of 16S-type RNAs have been largely responsible for establishing the existence of the third major branch in the evolutionary tree, that of the archaebacteria, and in showing that archaebacteria, eubacteria, and eukaryotes represent independent lines of descent from a common ancestor. *See* ARCHAEBACTERIA; EUKARYOTAE.

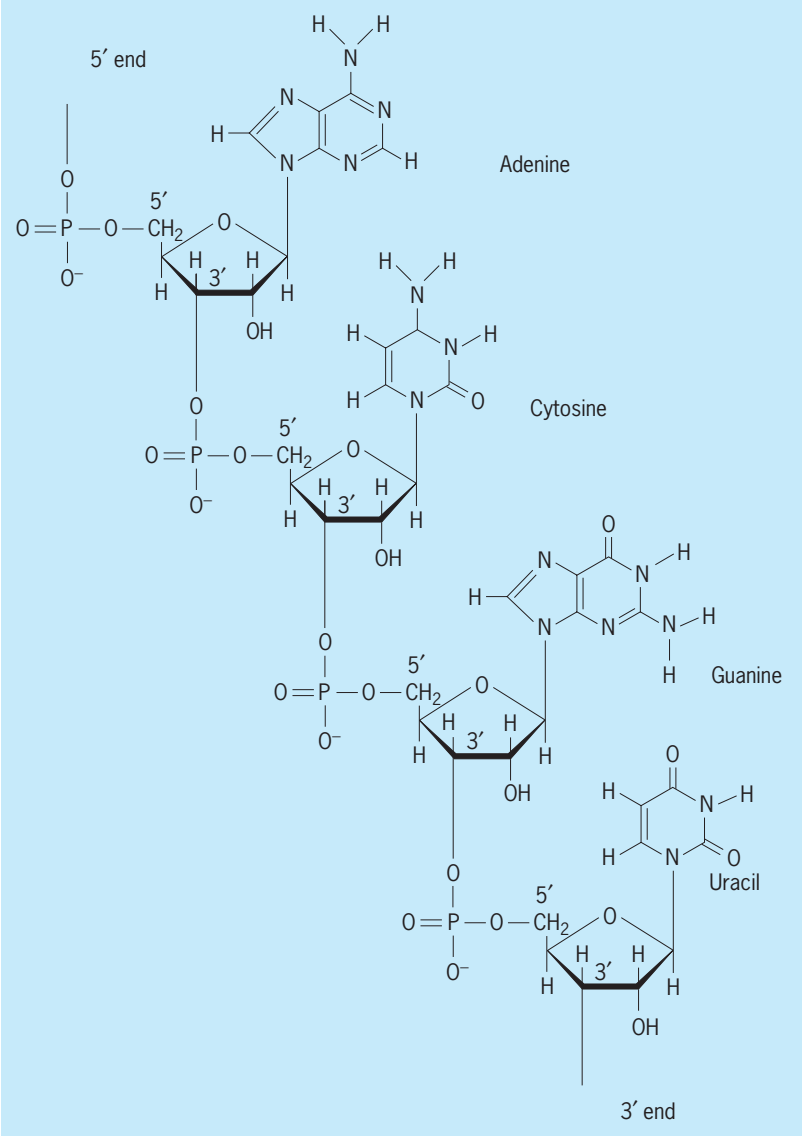

Fig. 1. Portion of an RNA molecule showing four nucleotides, each representing one of the four bases present in RNA.

Messenger RNA. Whereas most types of RNA are the final products of genes, mRNA is an intermediate in information transfer. It carries information from DNA to the ribosome in a genetic code that the protein-synthesizing machinery translates into protein. Specifically, mRNA sequence is recognized in a sequential fashion as a series of nucleotide triplets by tRNAs via base pairing to the three-nucleotide anticodons in the tRNAs. There are specific triplet codons that specify the beginning and end of the protein-coding sequence. Thus, the function of mRNA involves the reading of its primary nucleotide sequence rather than the activity of its overall structure. A mammalian cell contains up to 20,000 different mRNAs, some of which are present in only a few copies and others in many thousands of copies, depending on how much of the protein product is needed. Most mRNAs range in length between 700

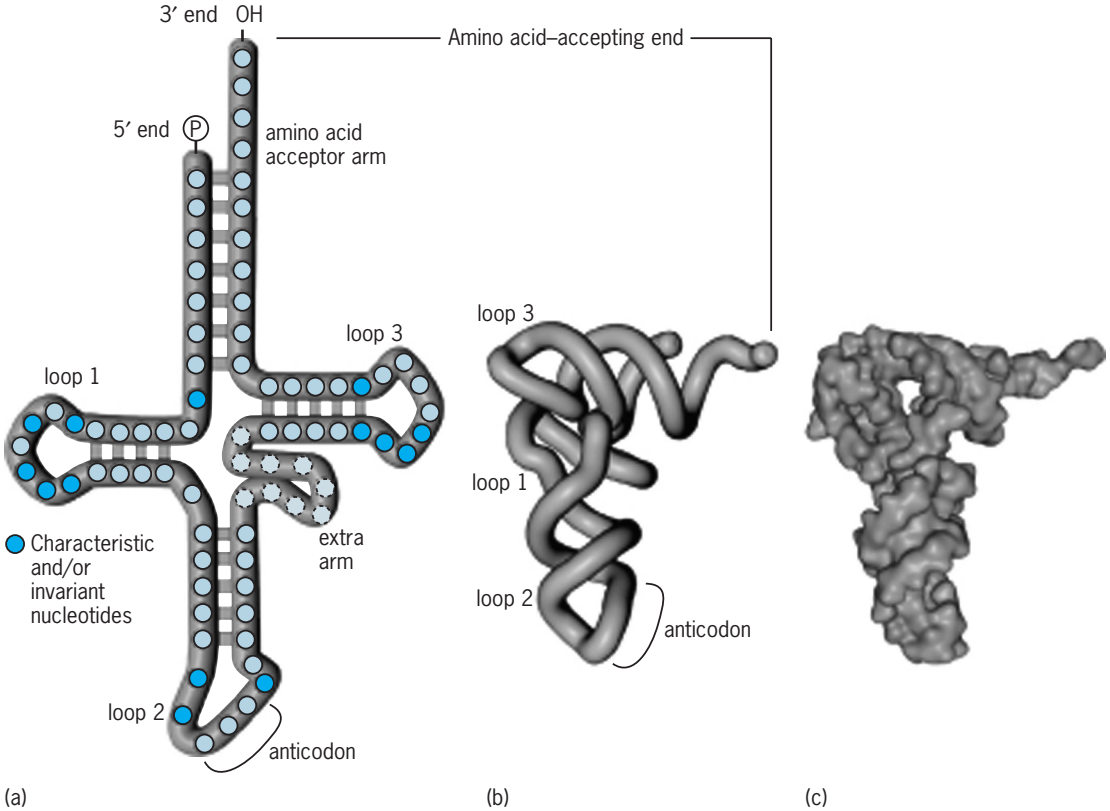

Fig. 2. Structural representations of a generalized tRNA molecule. (*a*) Secondary structure model showing base-pairing interactions within the tRNA. The characteristic and/or invariant nucleotides are the same regardless of the type of tRNA; the remainder vary. The extra arm varies in size among tRNAs. (*b*) Outline of the overall three-dimensional conformation of the tRNA. (*c*) Surface model of the tRNA.

and 10,000 nucleotides, with the length determined primarily by the size of the protein encoded. Messenger RNAs are typically shorter-lived than the more stable structural RNAs, such as tRNA and rRNA. *See* GENETIC CODE.

In eukaryotes, mRNAs are derived from larger precursor RNAs by a series of RNA processing steps that take place within the cell nucleus. Both ends of the pre-mRNA are modified: the 5′ end by addition of a methylated guanine nucleotide cap and the 3′ end by the addition of 50–200 adenosine residues in a nontemplated reaction to form a poly(adenosine) tail. Additionally, the informational, or coding, sequences (exons) of most genes are interrupted by long stretches of noncoding sequences known as introns. A typical mRNA precursor consists of alternating exons and introns, with most of the RNA being intron. These long pre-mRNAs must be spliced to remove noncoding intervening sequences. The mRNA splicing process, which consists of RNA cleavage at junctions between exons and introns and ligation of exons, is mediated by another class of cellular RNAs, the small nuclear RNAs (snRNAs). *See* EXON; INTRON.

Small nuclear RNA. Small RNAs, generally less than 300 nucleotides long and rich in uridine (U), are localized in the nucleoplasm (snRNAs) and nucleolus (snoRNAs) of eukaryotic cells, where they take part in RNA processing. These RNAs have characteristic folded structures and function as ribonucleoproteins (snRNPs and snoRNPs), in association with specific proteins. Five snRNPs (U1, U2, U4, U5, U6) participate in intron removal during eukaryotic mRNA splicing, whereas a sixth (U7) has a role in 3′-end formation in histone mRNAs. Several snoRNPs (in particular, U3) are involved in specific steps of precursor rRNA (pre-rRNA) cleavage, whereas many others participate in the extensive posttranscriptional modification that occurs during production of mature rRNA. *See* CELL NUCLEUS.

snRNA. Intron splicing occurs in the spliceosome, a particle assembled from precursor mRNA, snRNPs, and proteins (splicing factors). Intron sequences up to many thousands of nucleotides long are precisely removed, with the flanking exon (coding) sequences being joined in the correct reading frame for translation. The precision and efficiency of splicing are achieved through complementary base pairing between short conserved sequences in the snRNAs and at the intron-exon boundaries (splice junctions). For example, U1 snRNA pairs with the 5′ splice junction, whereas U2 snRNA recognizes and binds to a sequence in the intron just upstream of the 3′ splice junction and encompassing the so-called "branchpoint A" residue. Base-pairing interactions between U2 and U6 snRNAs in the spliceosome lead to activation of the 3′-hydroxyl of the branchpoint A residue so that it functions as a nucleophile in an attack on the 5′ splice junction, thereby initiating splicing. Two successive cuts in the pre-mRNA chain (first at the 5′-splice junction, then at the 3′-splice junction) are

required for intron removal, which is followed by joining of the flanking exons. Although not definitive, mounting evidence indicates that splicing is catalyzed by snRNA rather than by protein, so that the spliceosome (like the ribosome) may be fundamentally an "RNA machine."

snoRNA. One class of modification-guide snoRNP, box C/D, mediates ribose methylation through base pairing with pre-rRNA, with different snoRNAs selecting different residues within a targeted region of pre-rRNA for 2′-O-methylation. The sequence in the snoRNA that is complementary to the modification target in pre-rRNA is immediately upstream of the 5′-end of box D, one of the short conserved sequence motifs that characterize a box C/D snoRNA. The precise site of 2′-O-methylation is specified by the "N + 5" rule, such that the pre-rRNA residue that is modified is always the one that is paired to the fifth snoRNA nucleotide upstream of the D box, within the helical region formed by base pairing between pre-rRNA and snoRNA. Box C/D snoRNPs contain a common set of proteins, one of which (fibrillarin) is the methyltransferase that carries out the formation of 2′-O-methylribose.

A second class of modification-guide snoRNP (box H/ACA) is responsible for formation of pseudouridine (5-ribosyluracil) residues in rRNA. The site of pseudouridine formation is also determined by complementary base pairing between a box H/ACA snoRNA and its target sequence in pre-rRNA. In this case, however, the U residue selected for modification remains unpaired within a pseudouridylation pocket flanked by short pre-rRNA:snoRNA helices. Box H/ACA snoRNPs also contain a common set of proteins, different than the box C/D set, one of which (Cbf5p; dyskerin in humans) is the pseudouridine synthase that carries out the modification.

Although most snoRNAs function in cleavage and modification of pre-rRNA in eukaryotic cells, a subset directs formation of 2′-O-methylnucleosides and pseudouridine residues in spliceosomal snRNAs. This subset of small nuclear RNAs, designated "scaRNAs," is localized in a nucleoplasmic structure known as the Cajal body. Except for the fact that they function in a nonnucleolar milieu, scaRNAs are structurally and functionally indistinguishable from box C/D and box H/ACA snoRNAs.

Catalytic RNA. RNA enzymes, or ribozymes, are able to catalyze specific cleavage or joining reactions either in themselves or in other molecules of nucleic acid. For example, one of the ribonucleases involved in tRNA processing, RNase P, is composed of one protein and one RNA molecule, and under certain conditions the enzymatic activity can be supplied entirely by the RNA moiety. Another example is self-splicing introns found in various genes, such as the rRNA gene of the protozoon *Tetrahymena*. These introns mediate their own removal from an RNA molecule in a reaction that is dependent on the specific higher-order structure in the intron RNA. Presumably the three-dimensional surface of the RNA results in an arrangement of atoms such that certain covalent bonds are strained and unusually reactive, becoming susceptible to cleavage catalyzed by a bound divalent metal ion [usually magnesium (Mg^{2+})]. Several other types of catalytic RNA are categorized by specific conserved features of primary and secondary structure. *See* CATALYSIS; RIBOZYME.

Viral RNA. While most organisms carry their genetic information in the form of DNA, certain viruses, such as polio and influenza viruses, have RNA as their genetic material. The viral RNAs occur in different forms in different viruses. For example, some are single-stranded and some are double-stranded; some occur as a single RNA chromosome while others are multiple. In any case, the RNA is replicated as the genetic material, and either its sequence, or a complementary copy of itself, serves as mRNA to encode viral proteins. The RNA viruses known as retroviruses contain an enzyme that promotes synthesis of complementary DNA in the host cell, thus reversing the typical flow of information in biological systems. Many transforming viruses, which convert healthy cells to tumor cells, are retroviruses, as is human immunodeficiency virus (HIV), the causative agent of acquired immune deficiency syndrome (AIDS). *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME; ANIMAL VIRUS; RETROVIRUS; VIRUS.

As is evident from the above discussion, RNA molecules may function both as carriers of genetic information and as enzymes. The discoveries of RNA catalysis and of the central role of rRNA in protein synthesis have led to an enhanced appreciation of RNA as the probable original informational macromolecule, preceding both the more specialized DNA and protein molecules in evolution. *See* MOLECULAR BIOLOGY; NUCLEIC ACID.

Other types of RNA. RNA molecules serve other important and diverse cellular functions. For example, a ribonucleoprotein enzyme (telomerase) is responsible for replication of chromosome ends, with the RNA component functioning as a template for the reverse transcriptase activity of telomerase. An essential RNA component is also present in the signal recognition particle (SRP), a ribonucleoprotein complex that directs ribosomes synthesizing membrane and secreted proteins to the endoplasmic reticulum, where synthesis and posttranslational processing are completed. Finally, transfer-messenger RNA (tmRNA) functions during translation to rescue stalled ribosomes, targeting aberrant, partially synthesized proteins for proteolytic degradation.

Editing. RNA editing describes several different posttranscriptional processes that alter the nucleotide sequence of a mature RNA species compared to that of the gene that encodes it. In the case of mRNA, RNA editing results in codon substitutions that change the final protein sequence and that may also create new translation initiation and termination signals. Noncoding RNAs (rRNA, tRNA) sometimes undergo editing, as do certain viral RNAs.

Several different types of RNA editing have been described, such as insertions and deletions of uridine residues in trypanosome mitochondria, cytidine to uridine changes in plant mitochondria and chloroplasts, and adenosine to inosine substitutions in

animal nuclei. Editing ranges from a single cytidine deamination to generation of more than 50% of the final sequence by uridine addition. In trypanosome mitochondria, small guide RNAs (gRNAs) base-pair with the pre-mRNA in such a way as to determine the positions at which insertion or deletion of uridine residues occurs, thereby creating translatable open reading frames that did not exist prior to editing. Through approaches involving biochemical and genetic characterization, considerable progress has been made recently in identifying the RNA and protein components of editing complexes (editosomes) in several editing systems.

Antisense RNA: micro RNAs and small interfering RNAs. Antisense RNA is an RNA transcript (or portion of one) that is complementary to another nucleic acid, usually another RNA molecule. Naturally occurring examples of antisense RNA include, as mentioned earlier, the snoRNAs that mediate site-specific cleavage and modification of rRNA precursors, and gRNAs that supply the information for the uridine insertion/deletion type of RNA editing. In both bacterial and eukaryotic cells, other naturally occurring antisense RNAs function to regulate gene expression by binding to a target RNA molecule and interfering with its normal function. In eukaryotic cells, introduction of antisense RNA (for example, through activation of an appropriate transgene) has been shown to modulate expression of the complementary gene. This observation has encouraged attempts to use antisense RNA to manipulate gene expression ("gene silencing") for therapeutic and other purposes.

An unexpected and exciting consequence of antisense RNA studies has been the discovery of small interfering RNAs (siRNAs) that mediate the phenomenon of RNA interference (RNAi). An enzyme called Dicer generates siRNA via endonucleolytic cleavage of long double-stranded RNA molecules, which yields siRNA products approximately 22 nucleotides in length. Small interfering RNAs that are complementary to a long RNA molecule are able to target that molecule for degradation, thereby affecting RNA stability. Introduction or expression of mRNA-specific siRNAs in target cells is proving to be a powerful means of selectively and comprehensively inactivating gene expression. *See* GENE ACTION.

Micro RNAs (miRNAs) are another class of small RNA species that constitute a natural counterpart of siRNAs. They are the same size as siRNAs and are also generated by Dicer, although miRNAs are processed from longer single-strand precursors that are capable of assuming a hairpin configuration. Micro RNAs specifically regulate the activity of target genes, usually through translational arrest as a result of complementary base pairing with a sequence within the 3′ untranslated region of the corresponding mRNA. They appear to be key components of a novel RNA-based system of gene regulation that is evolutionarily conserved among at least multicellular eukaryotes, and that controls timing of development, stem cell maintenance, and other developmental and physiological processes in plants and animals. *See* RNA INTERFERENCE.

Ann L. Beyer; Michael W. Gray

Bibliography. B. Alberts et al., *Molecular Biology of the Cell*, 4th ed., 2002; T. R. Cech, RNA as an enzyme, *Sci. Amer.*, 255(5):64–75, 1986; R. F. Gesteland, T. R. Cech, and J. F. Atkins (eds.), *The RNA World*, 2d ed., Cold Spring Harbor Laboratory Press, 1999; H. Grosjean and R. Benne (eds.), *Modification and Editing of RNA*, American Society for Microbiology, Washington, DC, 1998; A. Rich and S. H. Kim, The three-dimensional structure of transfer RNA, *Sci. Amer.*, 238(1):52–62, 1978; J. A. Steitz, "Snurps," *Sci. Amer.*, 258(6):58–63, 1988.

Ribosomes

Small particles, present in large numbers in every living cell, whose function is to convert stored genetic information into protein molecules. In this synthesis process, a molecule of messenger ribonucleic acid (mRNA) is fed through the ribosome, and each successive trinucleotide codon on the messenger is recognized by complementary base-pairing to the anticodon of an appropriate transfer RNA (tRNA) molecule, which is in turn covalently bound to a specific amino acid. The successive amino acids become linked together on the ribosome, forming a polypeptide chain whose amino acid sequence has thus been determined by the nucleic acid sequence of the mRNA. The polypeptide is subsequently folded into an active protein molecule. Ribosomes are themselves complex arrays of protein and RNA molecules, and their fundamental importance in molecular biology has prompted a vast amount of research, with a view to finding out how these particles function at the molecular level. *See* GENETIC CODE; PROTEIN; RIBONUCLEIC ACID (RNA).

Components. Ribosomes are composed of two subunits, one approximately twice the size of the other. In the bacterium *Escherichia coli*, whose ribosomes have been the most extensively studied, the smaller subunit (named 30S, according to its sedimentation coefficient) contains 21 proteins and a single 16S RNA molecule. The larger (50S) subunit contains 32 proteins, and two RNA molecules (23S and 5S). The proteins vary in size from about 50 to 500 amino acids; their amino acid sequences have all been determined. Each protein is present in a single copy per ribosome, with the exception of one of the large subunit proteins, for which there are four copies. The RNA molecules have also been sequenced and have lengths of 120 nucleotides (5S RNA), about 1540 nucleotides (16S), and 2900 nucleotides (23S). The overall mass ratio of RNA to protein is about 2:1, and the combined molecular weight of the two subunits is about 2.5×10^6 daltons. Cations, in particular magnesium and polyamines, play an important role in maintaining the integrity of the ribosomal structures.

Comparative studies on the ribosomal components from other prokaryotic species have shown

that the sequences of both proteins and RNA have been highly conserved during evolution. Furthermore, the secondary structure of the RNA, that is, the way in which the RNA is folded into loops by complementary base pairing, has remained virtually identical in all the species studied. The ribosomes are considerably larger in the cytoplasm of higher organisms (eukaryotes). They have a higher protein content, and the two principal RNA molecules are up to 1900 and 4700 nucleotides in length for the small and large subunits, respectively. In contrast, ribosomes from organelles such as mitochondria are much smaller, mammalian mitochondrial ribosomes having RNA molecules of only about one-half the size of their bacterial counterparts. Nevertheless, it has become clear that, despite these gross differences in overall size, all ribosomal RNA molecules have a central core of conserved structure, which presumably reflects the universality of the ribosomal function. *See* CELL (BIOLOGY); ORGANIC EVOLUTION.

Arrangement of components. An understanding of ribosomal function requires a knowledge of the spatial arrangement of the individual ribosomal components; many techniques have been applied with this end in view, mostly using the ribosome of *E. coli* as a substrate. Electron microscopy of the ribosomal subunits has revealed that both subunits have a characteristic morphology or shape (**Figs. 1** and **2**), and that this property can be exploited by using antibodies to the individual proteins; the antibody connects two subunits together, forming a dimer, and the site of antibody attachment can be observed in the electron microscope. In this way a number of ribosomal proteins have been mapped on the subunit surface (Fig. 2). Antibodies to modified nucleotides in the RNA (the RNA contains a few such modifications) can be used in the same way. Other mapping techniques involve the use of physicochemical methods to measure the distance between two selected components, which can be labeled with fluorescent markers or with deuterium. In the latter case, neutron scattering measurements are made to determine the distance between the two components under consideration. Crystals have also been obtained from isolated ribosomal proteins or even whole subunits, which opens the possibility of making detailed structural analyses via x-ray diffraction techniques. *See* ANTIBODY; X-RAY DIFFRACTION.

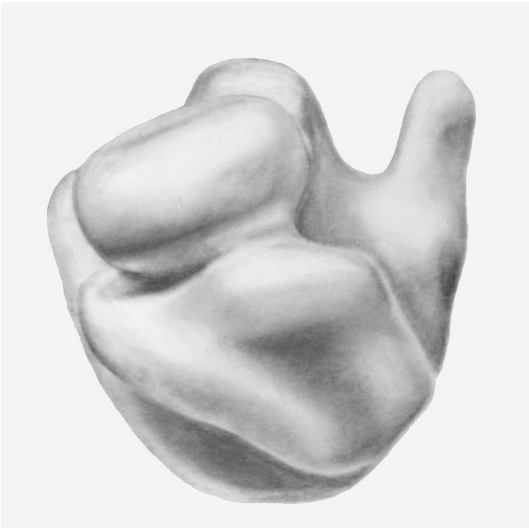

Fig. 1. Model of the *Escherichia coli* ribosome, based on electron microscopic data, showing the two subunits.

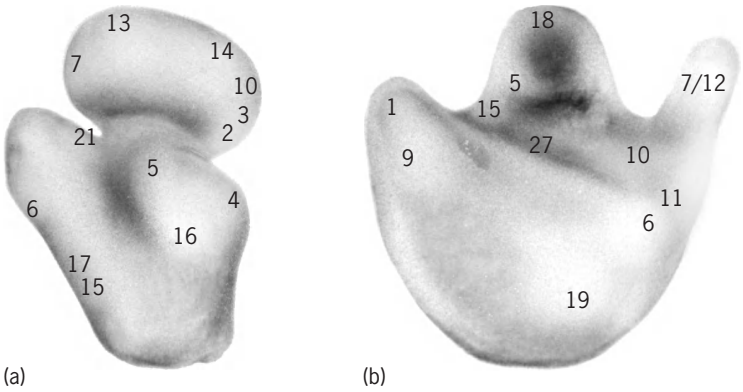

Fig. 2. Models, from electron microscopic data, of (*a*) 30S ribosomal subunit and (*b*) 50S subunit. Numbers identify some proteins that have been localized on subunit surface.

Chemical cross-linking techniques have been widely applied to investigate "neighborhoods" (as opposed to absolute distances) between different parts of the ribosome. Here the subunits are treated with a suitable bifunctional reagent, and the subsequent analysis determines which parts of the ribosome are close enough to become cross-linked to one another. The technique can be applied equally to the study of protein-protein neighborhoods, protein-RNA neighborhoods, or neighborhoods between different regions of the RNA. It also lends itself well to the study of the ribosomal interface—the region of contact between the large and small subunits. Chemical modification or susceptibility to enzyme digestion can be used in a similar manner to determine which regions of the ribosome are accessible on the surface of the particles.

A number of the methods just outlined rely on the fact that functionally active bacterial ribosomes can be reconstituted from their isolated RNA and protein components. These reconstitution techniques have proved to be a powerful tool in both structural and functional studies, since one or more components can be either omitted from the reconstitution mixture or replaced by a suitably modified counterpart. This type of approach can be complemented by genetic studies, since many ribosomal mutants have been isolated in which individual proteins are either altered or even absent.

In 2000, high-resolution structures of the 30S and 50S subunits were solved using x-ray diffraction. Remarkably, the structure of the 50S subunit indicates that the 23S RNA is responsible for peptidyl transferase activity of the ribosome.

Biosynthesis. Ribosomal RNA is produced in the cell by transcription of the appropriate deoxyribonucleic acid (DNA) gene, and the primary transcription

product is a large precursor molecule containing all three ribosomal RNA species (although in eukaryotes the 5S RNA is not contained in this molecule). The precursor is subsequently cleaved in several stages to yield the mature ribosomal RNA molecules. In prokaryotes the ribosomal proteins begin to be added while the RNA is still being transcribed; in eukaryotes the RNA is synthesized in the cell nucleolus and the completed precursor is then transported to the cytoplasm. The ribosomal proteins are themselves synthesized via mRNA on ribosomes in the cytoplasm, and in the case of *E. coli* it is known that the majority of the genes coding for ribosomal proteins are grouped in a few large clusters on the genome. Modification of the RNA and proteins (for example, by methylation) occurs at various stages during the assembly process. *See* GENE; GENE ACTION.

Role in protein synthesis. Once the assembly of the ribosomal subunits is complete, they are ready to take part in protein biosynthesis. This process is essentially very similar in both prokaryotes and eukaryotes; what follows is a brief summary of what happens in *E. coli*.

The first step (initiation of protein synthesis) is that an initiator tRNA molecule attached to the amino acid *N*-formyl methionine recognizes its appropriate codon on a mRNA molecule, and binds with the mRNA to the 30S subunit. One end (the 3′ end) of the 16S RNA is also involved in the formation of this complex. A 50S subunit then joins the complex, forming a complete 70S ribosome. A number of proteins (initiation factors, which are not ribosomal proteins) are also involved in the process. At this stage, the initiator aminoacyl tRNA occupies the P-site (peptidyl site) on the ribosome, while a second tRNA binding site (the A-site, or aminoacyl site) is free to accept the next coded aminoacyl tRNA molecule.

In the subsequent steps, the elongation process (**Fig. 3**), aminoacyl tRNA molecules are brought to the A-site as ternary complexes together with guanosine triphosphate (GTP) and a protein factor (elongation factor Tu). Once an aminoacyl tRNA is in the A-site, the initiator amino acid (or at later stages the growing polypeptide chain) is transferred from the P-site tRNA to the A-site aminoacyl tRNA. It is not clear whether this peptidyl transferase activity requires the active participation of a ribosomal component, or whether merely the correct placement of the two tRNA molecules is sufficient for the process to occur spontaneously.

After peptide transfer has taken place, the peptide is attached to the A-site tRNA, and an "empty" tRNA molecule is at the P-site. The peptidyl tRNA complex must now be translocated to the P-site, in order to free the A-site for the next incoming tRNA molecule. Here again protein factors and GTP are involved, and it is probable that the empty tRNA occupies a third ribosomal site (exit or E-site) before finally leaving the complex. Studies show that elongation factor G hydrolyzes GTP to bring about translocation of the tRNA molecules.

The protein chain is completed by the appearance of a "stop" codon in the mRNA. This is recognized by yet another set of protein factors, which causes the completed polypeptide chain to be released from the ribosome. At any one time a number of ribosomes are engaged in the reading of a single mRNA molecule, which leads to the appearance of polyribosomes (polysomes). Each step of the protein biosynthetic process can be disrupted by various antibiotics, and these (or their analogs) can be used for studying different aspects of the ribosomal function, as well as for locating functional sites on the ribosomal subunits by utilization of electron microscopy. *See* MOLECULAR BIOLOGY.

R. Brimacombe; H. G. Wittmann; S. Joseph

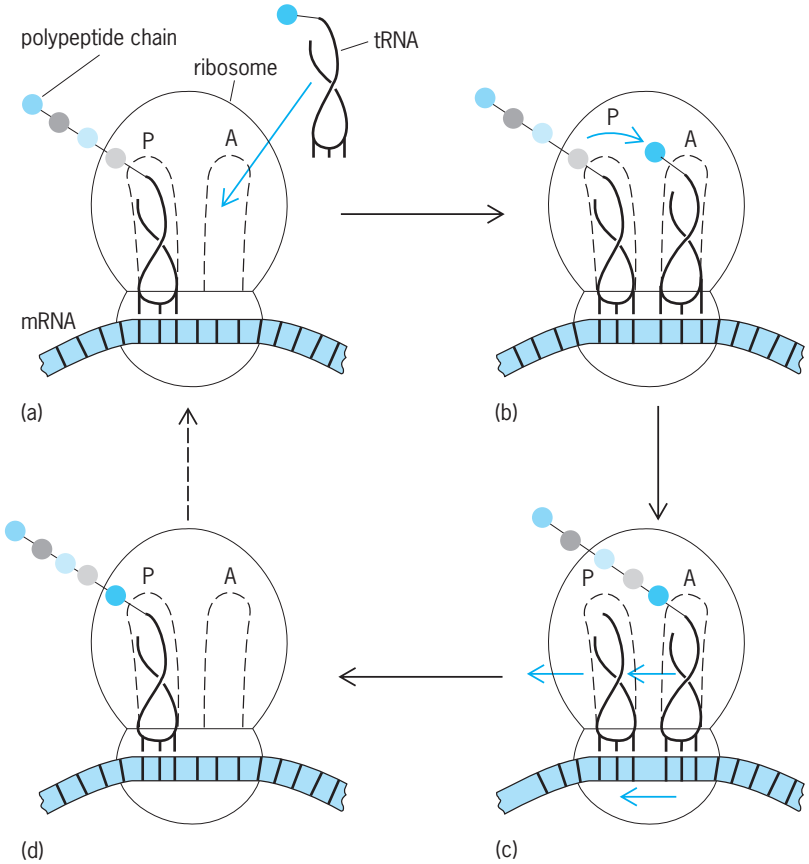

Fig. 3. Simplified diagram of a ribosome synthesizing a protein. It shows the messenger RNA (ribbon), the transfer RNA molecules binding to the A- and P-sites, and the growing polypeptide chain emerging from the large subunit. The steps are: (*a*) Aminoacyl tRNA (as ternary complex) enters A-site. (*b*) Peptide is transferred to the new amino acid at the A-site. (*c*) mRNA and peptidyl tRNA are translocated, with subsequent release of "empty" tRNA. (*d*) Ribosome, with polypeptide one amino acid longer, is ready for the next elongation cycle. Colored arrows indicate movement.

Bibliography. R. Green and H. Noller, Ribosomes and translation, *Annu. Rev. Biochem.*, 66:679-716, 1997; W. E. Hill (ed.), *The Ribosome: Structure, Function, and Evolution*, 1990; J. D. Puglisi, S. C. Blanchard, and R. Green, Approaching translation at atomic resolution, *Nature Struct. Bio.*,7(10): 855–861, 2000; G. Spedding (ed.), *Ribosomes and Protein Synthesis: A Practical Approach*, 1990; H. G. Wittmann, Architecture of prokaryotic ribosomes, *Annu. Rev. Biochem.*, 52:35–65, 1983; R. A. Zimmerman and A. E. Dahlberg, *Ribosome RNA: Structure, Evolution, Processing and Function in Protein Biosynthesis*, 1995.

Ribozyme

A ribonucleic acid (RNA) molecule that, like a protein, can catalyze specific biochemical reactions. *See* RIBONUCLEIC ACID (RNA).

Self-splicing rRNA. The first example of RNA molecules with catalytic activity came from observations of RNA processing in the protozoon *Tetrahymena thermophila*. Processing reactions, that is, the biochemical reactions that convert a newly synthesized RNA molecule to its mature form, may result in the removal of terminal sequences of an RNA molecule or in the removal of internal sequences that interrupt the mature RNA segments. The interruptions are called intervening sequences and are removed by an RNA processing reaction called splicing. Splicing involves bond breakage to excise the intervening sequence and bond formation to religate the mature RNA segments. *See* INTRON.

The intervening sequence of *T. thermophila* ribosomal RNA (rRNA) belongs to a large family of functionally related intervening sequences that were isolated from phylogenetically diverse organisms. The self-splicing reaction is inhibited by biochemicals that denature the native RNA structure, thereby providing evidence for a role of a specific folded RNA structure in functional activity. This folded structure is important to the ribozyme's ability to efficiently bind a substrate molecule in the active site, to accelerate the conversion of the substrate molecule to a product molecule by lowering the activation energy for chemical bond cleavage or religation, and to enable release of the reaction product from the ribozyme's active site. Furthermore, specific mutated-nucleotide residues play important roles in binding substrates in the active site, and in stabilizing the folded three-dimensional structure of the ribozyme. Similar mutations within the primary amino acid sequence of a protein enzyme can also destroy its structure and function. Thus, biological catalysts composed of RNA behave very much like protein enzymes.

RNase P. Ribonuclease P (RNase P) is the RNA processing enzyme responsible for the maturation of the 5′ ends of precursor transfer RNA (tRNA) molecules. The intact enzyme from bacteria and eukaryotes is composed of both RNA and protein. Both moieties are required for function if the enzyme's activity is measured under physiological conditions. The protein component RNase P is thought to facilitate the ability of the ribozyme to fold into the appropriate three-dimensional structure that is necessary for substrate binding and catalysis.

The RNase P RNA molecules from divergent bacteria differ considerably in the nucleotide sequence composition. In contrast to self-splicing RNA molecules, no single nucleotide in the RNase P ribozyme is critical to the function of RNase P. Thus, the RNA processing reaction catalyzed by RNase P appears less dependent on the identity of specific nucleotide residues in the ribozyme than is evident for RNA-catalyzed self-splicing. The notion that radically different ribozyme structures catalyze quite distinct RNA processing reactions is not new: protein enzyme families that are composed of different structures catalyze different types of biochemical reactions as well.

Ribozymes versus enzymes. Ribozymes share many similarities with protein enzymes, as assessed by two parameters that are used to describe a biological catalyst. The Michaelis-Menten constant K_m relates to the affinity that the catalyst has for its substrate. Ribozymes possess K_m values in the micromolar range, which is comparable to K_m values of protein enzymes. The catalytic rate constant describes how efficiently a catalyst converts substrate into product. The values of this constant for ribozymes are markedly lower than those values observed for protein enzymes. Nevertheless, ribozymes accelerate the rate of chemical reaction with specific substrates by 10^{11} compared with the rate observed for the corresponding uncatalyzed, spontaneous reaction. Therefore, ribozymes and protein enzymes are capable of lowering to similar extents the activation energy for chemical reaction. *See* ENZYME.

Primordial catalyst. Before the discovery of ribozymes, biochemists debated which biological molecule could behave as the primordial catalyst. Deoxyribonucleic acid (DNA), which is a stable, information-rich molecule, was previously not believed to encode any functional activity. Proteins are functionally rich molecules, however, they may not have readily lent themselves to being reproduced in the primordial environment because of their diverse amino acid composition. Because RNA molecules can encode both information and catalytic activity, many biochemists now believe that RNA was the likely primordial biological catalyst. In that role, RNA molecules may have catalyzed the elementary biochemical reactions that were essential to the early forms of life in the primordial biosphere. Polypeptides may have supplanted RNA-based enzymes as a greater need for more stable and functionally diverse biological catalysts evolved.

Daniel W. Celander

Bibliography. D. W. Celander and T. R. Cech, Visualizing the higher order folding of a catalytic RNA molecule, *Science*, 251:401–407, 1991; C. Guerrier-Takada et al., The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme, *Cell*, 35:849–857, 1983; D. Herschlag and T. R. Cech, Catalysis of RNA cleavage by the *Tetrahymena thermophila* ribozyme: 1. Kinetic description of the reaction of an RNA substrate complementary to the active site, *Biochemistry*, 29:10159–10171, 1990; K. Kruger et al., Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena, Cell*, 31:147–157, 1983; F. Michel and E. Westhof, Modelling of the three-dimensional architecture of Group I catalytic introns based on comparative sequence analysis, *J. Mol. Biol.*, 216:585–610, 1990; H. F. Noller, V. Hoffarth, and L. Zimniak, Unusual resistance of peptidyl transferase to protein extraction procedures, *Science*, 256:1416–1419, 1992; D. S. Waugh and N. R. Pace, Gap-scan deletion analysis of *Bacillus subtilis* RNase P RNA, *FASEB J.*, 7:188–195, 1993.

Rice

The plant *Oryza sativa* is the major source of food for nearly one-half of the world's population. In China, Japan, Korea, the Philippines, India, and other countries of Asia, rice is far more important than wheat as a source of carbohydrates. In some countries of the Orient the consumption of rice per capita is estimated at 200–400 lb (90–180 kg) per year. In contrast, the yearly per capita consumption of rice in the United States is only about 8 lb (3.6 kg). The most important rice-producing countries are mainland China, India, and Indonesia, but in many smaller countries rice is the leading food crop. In the United States, rice production is largely concentrated in selected areas of Arkansas, California, Louisiana, and Texas. *See* CARBOHYDRATE; CYPERALES; WHEAT.

Use. Over 95% of the world rice crop is used for human food. Although most rice is boiled, a considerable amount is consumed as breakfast cereals. Rice starch also has many uses. Broken rice is used as a livestock feed and for the production of alcoholic beverages. The bran from polished rice is used for livestock feed; the hulls are used for fuel and cellulose. The straw is used for thatching roofs in the Orient and for making paper, mats, hats, and baskets. Rice straw is also woven into rope and used as cordage for bags. This crop serves a multitude of purposes in countries where agriculture is dependent largely upon rice.

Origin and description. Rice apparently originated more than 7000 years ago in Southeast Asia, in the areas that are now eastern India, Indochina, and southern China. Rice was introduced to North America as early as 1609, and became established in South Carolina about 1690. Until about 1890, rice was grown mainly in the southeastern states. Louisiana became an important rice-producing state in the late nineteenth century. Early in the twentieth century, rice production spread to southeastern Texas, eastern Arkansas, and north-central California. Rice is a comparatively new crop in its present areas of greatest production in the United States.

Rice is unlike many other cereal grains in that all cultivated varieties belong to the same species and have 12 pairs of chromosomes, as do most wild types. The extent of variation in morphological and physiological characteristics within this single species is greater than for any other cereal crop. Although the chromosome number is the same, many of the ancient types have become so widely differentiated that hybrids between them are only partially fertile. *See* GENETICS.

Rice is an annual grass plant varying in height from 2 to 6 ft (0.6 to 1.8 m). Plants tiller, that is, develop new shoots freely, the number depending upon spacing and soil fertility. Among the many types grown, some mature in 80 days; others require over 200 days. The inflorescence is an open panicle (**Fig. 1**). Flowers are perfect and normally self-pollinated, with natural crossing seldom exceeding 3–4%. A distinct characteristic of the flower is the six anthers rather than the customary three of other grasses. Spikelets have a single floret, lemma and palea completely enclosing the caryopsis or fruit, which may be yellow, red, brown, or black. Lemmas may be awnless, partly awned, or fully awned. Threshed rice, which retains its lemma and palea, is called rough rice or paddy. *See* FLOWER; FRUIT; GRASS CROPS; INFLORESCENCE; REPRODUCTION (PLANT).

Fig. 1. Panicles of (*a*) short-grain and (*b*) long-grain rice. Each spikelet has a single caryopsis enclosed in the lemma and palea.

Varieties. In the United States, only about 25 varieties are in commercial production. Cultivated rices are classified as upland and lowland. Upland types, which can be grown in high-rainfall areas without irrigation, produce relatively low yields. The lowland types, which are grown submerged in water for the greater part of the season, produce higher yields. In contrast to most plants, rice can thrive when submerged because oxygen is transported from the leaves to the roots. All rice in the United States is produced under lowland or flooded conditions. Rice varieties are also classified as long- or short-grain (**Fig. 2**). Most long-grain rices have high amylose content and are dry or fluffy when cooked, while most short-grain rices have lower amylose content

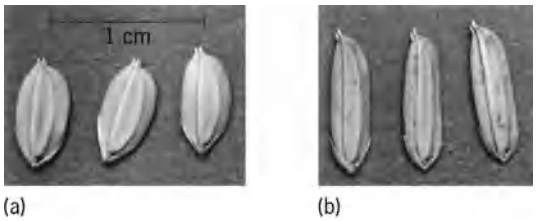

Fig. 2. Rice grains range from (*a*) short to (*b*) long.

and are sticky when cooked. In the United States a third grain length is recognized: medium-grain. The medium-grain rices have cooking qualities similar to those of short-grain varieties.

Plant breeders have been successful in developing high-yielding varieties that combine the vigor and disease resistance of typical tropical rices with the short stature of Taiwan varieties. The "architecture" of these plants ensures their ability to stay erect until harvesting, thus preventing crop loss due to lodging (falling over) into the water.

Cultural practices. Because of the peculiar conditions of growing a crop submerged in water, rice land seldom becomes a part of a regular rotation system. Often two or more successive crops of rice are grown, and the land is then pastured or fallowed to control weeds. Rice is grown on heavy soils underlain by an impervious subsoil to prevent seepage of water. *See* SOIL.

In Asian countries and elsewhere, rice fields are established by transplanting seedlings from beds when the plants are 20–50 days old. Fields to be transplanted are flooded and worked into a soft mud. Clumps of three to four seedlings are pushed into the mud in rows to permit hand cultivation for control of weeds. This system of transplanting seedlings saves irrigation water and permits the field in which they are to be established to grow another crop while the smaller-sized seedling bed is being grown.

In the United States, transplanting is impractical because of high labor costs. In California, almost all rice seed is sown from airplanes directly into fields flooded with 4–6 in. (10–15 cm) of water. In the southern states, rice is seeded either by airplane or by grain drill. In the latter case the rice is seeded into dry soil, which is gradually flooded to a depth of 4–6 in. (10–15 cm) as the rice emerges and grows. In all rice areas of the United States the water is kept at the 4–6 in. (10–15 cm) depth until the land is drained shortly before harvest.

The rice crop in Asian countries and some other countries is harvested by hand (**Fig. 3**). In the United States, rice is harvested with large self-propelled combines similar to those used for wheat and other grain crops. The best stage for combine harvesting is at grain moisture content of 23–28%. Rice dried to a lower moisture content while standing in the field may break up in milling, resulting in lower grain quality. When harvested at this high moisture content, the grain is dried artificially to 14% moisture, care being taken to keep the drying temperature below 110°F (43°C) to avoid damage to the grain. *See* AGRICULTURAL MACHINERY; AGRICULTURAL SCIENCE (PLANT); GRAIN CROPS. J. N. Rutger

Cultivars and genetics. The genus *Oryza* has unusual diversity and great antiquity because it originated in the Gondwana supercontinent and is now distributed over Africa, Oceania, Asia, and Latin America. The genus consists of about 20 wild species and two cultigens (*O. sativa* of Asia and *O. glaberrima* of western Africa). The enormous diversity in *O. sativa* results from ecogenetic diversification following its wide dispersal by human populations.

At the time of the Green Revolution, which for rice began in the late 1960s, probably more than 100,000 cultivars existed in different parts of the world. These numerous unimproved cultivars (land races) in irrigated areas were displaced by the high-yielding modern varieties, and many disappeared in the absence of continuous cultivation. Even the wild rices are increasingly threatened by development projects. Since 1972, however, massive international efforts in field collection have netted more than 40,000 samples and saved many cultivars from extinction. The base collection maintained at the International Rice Research Institute (IRRI) in the Philippines totals 86,000 accessions. Holdings of national and regional gene banks total over 100,000 but include many duplicates.

Genetic conservation of rice germ plasm has relied almost entirely on preserving seed or plants in gene banks. Most of the major rice-growing countries have established refrigerated storage facilities, but genetic erosion inside gene banks in the humid tropics continues when cold, dry storage conditions are not maintained or when seeds are not frequently regenerated. The IRRI collection is backed up by a duplicate storage site at the U.S. National Seed Storage Laboratory in Fort Collins, Colorado. Conservation in natural habitats has been fostered, but it is difficult to implement.

In spite of some setbacks due to pest epidemics, rice germ plasm is the most intensively evaluated and used crop collection in the world.The potential genetic vulnerability in *O. sativa* is considerably higher than that of other major cereals: About 1000 improved varieties bred since the 1960s and grown on nearly 60% of the world's rice lands largely share the same Chinese semidwarfing gene (sd_1); many tropical semidwarfs have a common cytoplasm; and more than 90% of the 3.2×10^7 acres (1.3×10^7 hectares) of hybrid rice grown in China have a common

Fig. 3. Hand harvesting is common in many countries.

male-sterile cytoplasm. Moreover, farmers in tropical Asia tend to use the same modern variety twice in a year in staggered plantings, adding to pest incidences and quick breakdown of genetic resistance. Although no major disaster has occurred, farmers using the above practices and heavy fertilization have suffered substantial losses in a cyclic manner. Te-Tzu Chang

Disease. A comprehensive disease control program is used in rice culture to ensure against yield and quality losses in fields and deterioration of quality during drying and storage. Key elements in this program are preventive measures such as crop rotation and sanitation to suppress carryover of pathogens from crop to crop in the soil or crop residues, use of good-quality seeds treated with an effective seed protectant fungicide, selection of adapted varieties with greatest degree of resistance to potentially epidemic diseases of the area, weed control, timely water management, and fertilization practices that consider both proper balance and timely application. In some disease situations, applications of foliar fungicides may be used to protect against buildup of disease to economically important levels. *See* FUNGISTAT AND FUNGICIDE.

Fig. 4. Rice blast disease.

Fig. 5. Brown bordered sheath spot disease.

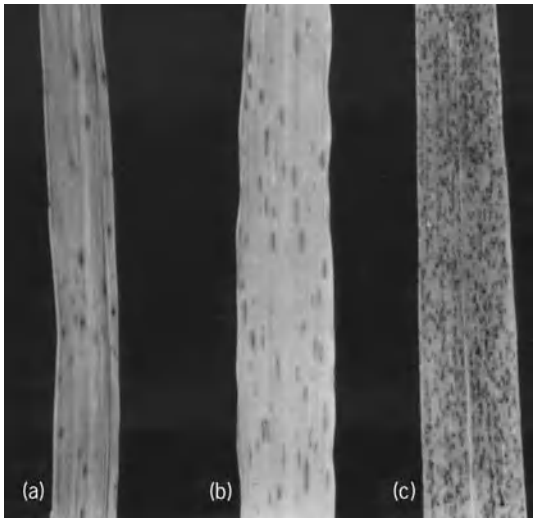

Fig. 6. Fungal diseases of rice. (*a*) Brown spot. (*b*) Narrow brown leaf spot. (*c*) Leaf smut.

Control measures. The variety of measures required for disease control is necessitated by the diversity of pathogens affecting the crop. Pathogens may be classed as fungi, bacteria, nematodes, viruses, or physiologic agents. There are numerous unique disease control strategies within each class. It is essential, therefore, to make correct identification to select the most appropriate control measures.

It is not possible to predict with much certainty what diseases are likely to be most prevalent in a given season, and their intensity may vary in different areas or years. However, certain conditions are known to favor certain diseases.

Fungi. Rice blast (**Fig. 4**), one of the most widely distributed and serious rice diseases, is caused by the airborne fungus *Pyricularia oryzae* and is favored by high levels of native or applied nitrogen such as are found in newly cleared land or following fish culture. In the seedling stage the disease is also favored by delayed flooding. Host plant resistance is a major means of control. Two culm rotting diseases, bordered sheath spot (**Fig. 5**) and stem rot, are caused by the soil-invading fungi *Rhizoctonia solani* and *Magnaporthe salvanii*, respectively, and are favored by continuous cropping of rice and very thick stands that hold moisture on the vegetation for several days near heading time. Stem rot is also favored by soils deficient in potassium. Rotation, proper fertilization, and proper seeding rates are the major means of controlling these two diseases. Brown leaf spot, caused by *Helminthosporium oryzae* (**Fig. 6**), is another serious and widespread airborne fungal disease; it is most serious on upland rice or where plants are stressed by cold or inadequate water, low nitrogen levels, or damaging herbicide applications.

Physiologic disorder. Straighthead (**Fig. 7**), a physiologic disorder, is likely to occur in low, poorly drained

Fig. 7. Straighthead, a physiologic disorder of rice.

Fig. 8. White tip disease of rice.

spots in the field or on highly fertile, newly cleared land not properly drained.

Nematodes. White tip disease (**Fig. 8**) is caused by a seed-borne nematode that overseasons almost exclusively in the rice kernel; it builds up to damaging levels if seeds are used from successive crops, season after season. It can be controlled by a seed certification program that excludes this pathogen.

Bacteria. Bacterial blight, caused by *Xanthomonas oryzae*, is most serious in low, wet, poorly drained locations that are easily flooded during the seedling stages and are often subject to mist.

Virus. Viral diseases are mostly transmitted by leaf hoppers and plant hoppers and thus build up in areas where two or three successive crops are grown each year.

Identification. Proper identification of diseases and the circumstances that favor each of them permits selection of varieties, cultural practices, or other control measures to prevent their occurrence at economic levels. *See* PLANT PATHOLOGY.

George E. Templeton

Processing. The rice kernel has four primary components: the hull or husk, the seedcoat or bran, the embryo or germ, and the endosperm. The main objective of milling rice is to remove the indigestible hull and additional portions of bran to yield whole unbroken endosperm. Rice milling involves relatively uncomplicated abrasive and separatory procedures which provide a variety of products dependent on the degree of bran removal or the extent of endosperm breakage.

Rice milling. For rice milling (**Fig. 9**), rough rice, or paddy, which is dried grain with the husk intact and with no adhering stalk, is cleaned by selective screening, disk separators, and air aspiration systems to remove dirt, stones, chaff, and other extraneous material. Clean paddy may be converted by parboiling and drying prior to milling.

Hull removal. The coarse husk of paddy is loosened from the intact kernel by either disk or rubber

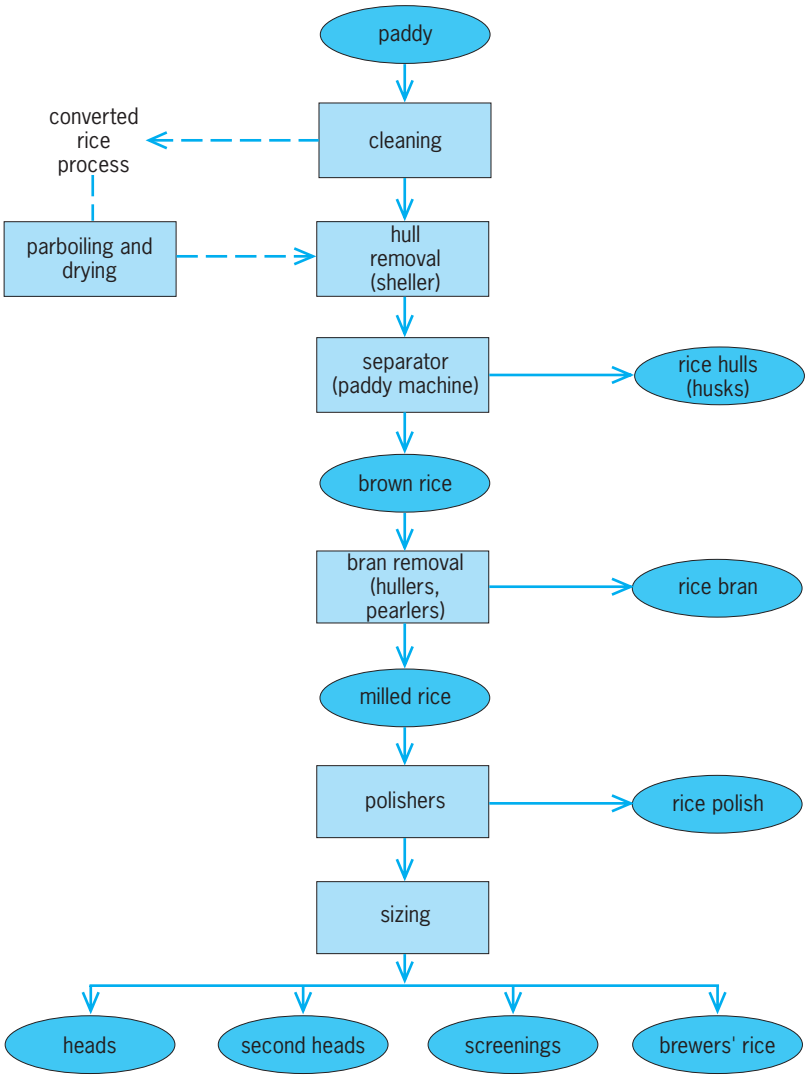

Fig. 9. General rice milling operations and products.

roll shellers. Disk shellers comprise two large horizontal emery-faced abrasive plates (disks) positioned with parallel surfaces. The upper disk is stationary and possesses a center hole into which a stream of paddy enters, while the lower disk is solid and can be rotated at designated speeds. The distance between these disks is adjusted to a width greater than that of the rice kernel length. Paddy passes from the center of the sheller to its periphery by a tumbling centrifugal action. This abrasive action removes the husk from about 80% of the kernels in a single pass. Disk shellers are adjusted to minimize kernel damage; however, they are being replaced increasingly by more efficient rubber roll shellers. Rubber-surfaced rolls are positioned in parallel at a specified gap and rotate toward each other at differential speeds. Rough rice fed between the rolls results in a gentle removal of husk by tearing action. This method yields greater quantities of hulled and undamaged kernels than do disk shellers.

Following the sheller operation, the low-density freed hulls are lifted and separated by air aspiration. Rice hulls comprise about 17–20% of the paddy and contain about 20% ash, primarily silica. Hulls have been used as boiler feed stock and as a press aid in fruit juice extraction.

The remaining paddy and dehulled kernels are conveyed to a separator which is commonly referred to as a paddy machine. This unit separates the less dense paddy from dehulled kernels on a vibrating, inclined gravity platform in which paddy migrates upward through a series of vertical baffle plates while the dehulled kernels migrate down the platform. Segregated paddy is returned to additional shellers. The dehulled kernel with seedcoat intact is termed brown rice.

Bran removal. Brown rice is further milled to remove the bran and thus to produce a whitened product by gentle and progressive abrasion in a series of machines referred to as hullers or pearlers. Actually the term huller is misleading, as the hull has already been removed. Hullers are constructed with an abrasive roller cylinder rotating inside a stationary cylindrical sieve. As this cylinder rotates, brown rice passes between the abrasive surface and the screen and is scoured to remove the outer bran layer. The fine bran particles are removed through the sieve. The adjustment, sizing, speed, and rice feed rates are critically controlled to minimize kernel breakage. Typically successive movement through two or three hullers, each adjusted to an increasingly finer setting, is desired. Pearlers are also used to remove bran by passing kernels between a cone-shaped rotor and a screen. Abrasive action is obtained among kernels and bran separated by screening. Heat generated during bran removal must be dissipated by air currents to avoid damage to rice cooking characteristics. Further whitening of rice may be accomplished by brushing or polishing to remove the inner white-bran, aleurone layer, and small portions of endosperm. This procedure develops light-colored fines termed rice polish which are frequently used as an ingredient in formulated baby foods.

Sizing of milled rice. Milled kernels are separated by size by using a variety of systems. Finest fragments are removed by reel screening, and remaining broken kernels by disk separation. Brewers' rice is the smallest kernel fragment screened and is used as fermentation feed stock. Medium-sized pieces (less than one-third kernel) are termed screenings, second heads include large pieces up to three-fourths of a kernel, while the whole unbroken kernels are classified as heads. Polished rice may be surface-coated with glucose by tumbling to improve its gloss, sheen, and uniformity.

Parboiled or converted rice. The soaking and cooking of paddy prior to milling has been practiced in many cultures. The greatest concentration of water-soluble B-complex vitamins (thiamine, riboflavin, and niacin) is found in the bran layers of the kernel, which are lost during dry-milling operations. During parboiling, the soluble vitamins diffuse into the endosperm and greatly enhance the nutritive value of the milled rice. Soaked and boiled paddy is dried prior to conventional milling. Parboiled rice starch is gelatinized, and the kernel acquires a darker, creamier color upon milling. The endosperm of parboiled rice is firmer and denser, and requires additional cooking time compared to regular rice; however, since the surface starch is removed, the cooked product resists sticking and clumping. *See* NIACIN; RIBOFLAVIN; THIAMINE.

Many industrial systems have been developed to accomplish the parboiling of paddy. Variations include long-time steeping, heating in live steam, and the use of vacuum infusion. The converted rice produced in the United States is made from cleaned paddy under an exacting process. Paddy is placed in large chambers and subjected to high vacuum to evacuate air from the kernel tissue. Hot water is directly percolated into the evacuated tank, and external pressure is applied to infuse the soluble nutrients of the hull and bran into the endosperm. Following sufficient hot water treatment, the steep water is drained and the rice is subjected to live steam. Drying of the parboiled paddy may include vacuum or conventional air treatment and a tempering period to assure uniformity of kernel moisture, which is maintained at about 12%.

Prepared rice products. Instant rice is made from whole grain rice by pretreating under controlled cooking, cooling, and drying conditions to impart the quick-cooking characteristic. Typically, instant rices require less than 5 min preparation time, as compared to 20–25 min for regular or parboiled rice products. A variety of processing schedules, including use of soaking, heating, cooking, vacuum or steam pressure, and controlled dehydration, are employed to render milled rice suitable for rapid preparation. Use of these techniques enables hydration and gelatinization of starch and yields an open, porous kernel surface with minimum kernel damage. This precooked product readily rehydrates and softens to a palatable grain in boiling water.

Ready-to-eat breakfast rice cereals are prepared

from milled rice as flakes or puffs. Rice is frequently precooked under steam pressure, conditioned to uniform kernel moisture, and passed through high-pressure, smooth flaking rolls, and toasted. Vacuum puffing of cooked grains or flaked rice is common. Instant hot baby cereal is produced by drum-drying a slurry of rice flour. Thin sheets of cooked dehydrated cereal are removed by surface-scraping, rotating, steam-heated drums. The sheets are ground to yield thin flakes which readily hydrate and form a characteristic, soft, pasty porridge. *See* CEREAL; FOOD MANUFACTURING.

Rice bran oil. This product was developed as a result of increased extraction of lipids from rice bran. Edible oils obtained from rice bran must be treated to inactivate lipase enzymes that will hydrolyze the lipids to form free fatty acids and rancid odors during storage. The crude lipid content of rice bran is about 10% by weight, and has been commercially extracted in several Asian countries, including India, China, and Japan. Traditionally, rice bran oil was obtained through hydraulic pressing, which provided a relatively low yield (10–12%). This procedure has been replaced by solvent extraction, which yields 16–18%. Bran is typically preheated by using steam pressure to sufficiently inactivate hydrolytic lipases prior to hexane extraction. The crude rice bran oil possesses a dark greenish-brown to light yellow pigmentation, which is removed during subsequent refining processes. Refined rice bran oil possesses a low content of linolenic acid (highly unsaturated) and a high content of tocopherols (having antioxidant properties). Both constituents contribute to increased oxidative stability. These properties enable rice bran oil to be utilized as an edible-grade oil in a variety of applications as well as an industrial feedstock for soap and resin manufacture. Improved utilization of rice bran oil has been achieved through improved extraction, refining, and formulation in food products; thus, this by-product utilization has the potential for increasing the value added to rice as a food and food ingredient. *See* FAT AND OIL (FOOD); SOLVENT EXTRACTION.

Mark A. Uebersax

Bibliography. T. T. Chang, Conservation of rice genetic resources: Luxury or necessity?, *Science*, 224:251–256, 1984; D. G. Dalrymple, *Development and Spread of High-Yielding Rice Varieties in Developing Countries*, 1986; S. K. DeDatta, *Principles and Practices of Rice Production*, 1987; H. C. Dethloff, *A History of the American Rice Industry, 1685–1985*, 1988; S. Gangopadhyay and S. Y. Padmanabhan, *Breeding for Resistance in Rice*, 1987; T. H. Hargrove et al., Twenty years of rice breeding, *BioScience*, 38:675–681, 1988; B. O. Juliano (ed.), *Rice Chemistry and Technology*, 1985; W. E. Marshall and J. I. Wadsworth, *Rice Science and Technology*, 1993; A. Prakash et al., *Rice Storage and Insect Pest Management*, 1987; B. M. Shepard et al., *Helpful Insects, Spiders, and Pathogens: Friends of the Rice Farmer*, 1987; S. Tsunoda and N. Takahashi (eds.), *Biology of Rice*, 1984; H. Waibel, *The Economics of Integrated Pest Control in Irrigated Rice*, 1986.

Ricinulei

An order of extremely rare arachnids, also known as the Podogona, with a body less than 1 in. (25 mm) in length. Superficially, they resemble ticks in general appearance and movement, and are found only in tropical Africa and in the Americas, from the Amazon to Texas. The two anterior pairs of appendages are chelate. The terminal segments of the third legs of the male are modified as copulatory structures. Less than 25 modern species are known. The occurrence of several fossils from Carboniferous time suggests that the group was formerly more common. Virtually nothing is known about the reproduction, growth, and ecology of the Ricinulei. *See* ARACHNIDA.

C. Clayton Hoff

Rickettsioses

Often severe infectious diseases caused by several diverse and specialized bacteria, the rickettsiae and rickettsialike organisms. The best-known rickettsial diseases infect humans and are usually transmitted by parasitic arthropod vectors (see **table**).

Agents. Rickettsiae and rickettsialike organisms are some of the smallest microorganisms visible under a light microscope. Although originally confused with viruses, in part because of their small size and requirements for intracellular replication, rickettsiae and rickettsialike organisms are characterized by basic bacterial (gram-negative) morphologic features. Their key metabolic enzymes are variations of typical bacterial enzymes. These organisms apparently exploit the normally rich pool of host intracellular nutrients for their own, perhaps abbreviated, synthetic pathways. This, at least partial, dependence on host-cell molecules helps explain why rickettsiae (with two exceptions) replicate only in living eukaryotic cells.

The genetic material of rickettsiae and rickettsialike organisms likewise seems to conform to basic bacterial patterns. The genome of all rickettsialike organisms consists of double-stranded deoxyribonucleic acid (DNA). In addition to the genomic DNA, *Coxiella burnetii*, the causative agent of Q fever, may also have extrachromosomal, circular DNA plasmids.

Rickettsiae enter host cells by phagocytosis and reproduce by simple binary fission. The site of growth and reproduction varies among the various genera. Rickettsiae that cause spotted fever appear to move easily through the membranes of infected cells and can often be seen in infected cell nuclei as well as in the cytoplasm. Other rickettsiae, for example, the typhus-group rickettsiae, do not appear able to pass easily out of the host cell through the cytoplasmic membranes. Instead of being budded and released to the extracellular environment, they may pack the host cytoplasm until the host cell is destroyed. Members of the rickettsialike genus *Ehrlichia* replicate within cytoplasmic vacuoles instead of replicating directly within the cytoplasm. Similarly, *C. burnetii* replicates only within the phagolysosomes of the

Causes and modes of transmission for some common rickettsial diseases

Disease	Causative agent	Mode of transmission
Epidemic typhus	*Rickettsia prowazekii*	Louse vector
Murine typhus	*Rickettsia typhi*	Flea vector
Trench fever	*Rochalimaea quintana*	Louse vector
Rocky Mountain spotted fever	*Rickettsia rickettsii*	Tick vector
Q fever	*Coxiella burnetii*	Tick vector, airborne transmission
Scrub typhus	*Rickettsia tsutsugamushi*	Mite vector

host cells under acidic conditions that otherwise would be considered quite hostile to microbial growth.

Although once grouped within the genus *Rickettsia*, the causative agent of trench fever and one other related organism isolated only from rodents have been grown on artificial media. Thus, they have been assigned to the genus *Rochalimaea*, despite apparently significant DNA homologies with typhus-group rickettsiae and very similar epidemiologic characteristics.

How other rickettsiae are related is less clearly defined. *Rickettsia tsutsugamushi*, the causative agent of scrub typhus, replicates in the cytoplasm of infected cells, is transmitted by arthropod vectors, and is included in the *Rickettsia* genus. However, *R. tsutsugamushi* shares no significant common protein antigens with the spotted fever or typhus-group rickettsial agents that would indicate close evolutionary histories; several other features of *R. tsutsugamushi* are equally distinct.

Human disease. Clinically, the rickettsial diseases of humans are most commonly characterized by fever, headache, and some form of cutaneous eruption, often including diffuse rash, as in epidemic and murine typhus and Rocky Mountain spotted fever, or a primary ulcer or eschar at the site of vector attachment, as in Mediterranean spotted fever and scrub typhus. Signs of disease may vary significantly between individual cases of rickettsial disease. Q fever is clinically exceptional in several respects, including the frequent absence of skin lesions.

All of the human rickettsial diseases, if diagnosed early enough in the infection, can usually be effectively treated with the appropriate antibiotics. Tetracycline and chloramphenicol are among the most effective antibiotics used; they halt the progression of the disease activity, but do so without actually killing the rickettsial organisms. Presumably, the immune system, even that of patients who have been treated with antibiotics, is ultimately responsible for ridding the body of infectious organisms. Penicillin and related compounds are not considered effective. *See* ANTIBIOTIC.

Most rickettsial diseases are maintained in nature as diseases of nonhuman vertebrate animals and their parasites. Although human disease may be extremely serious, human infection may usually be regarded as peripheral to the normal natural infection cycles, and human-to-human transmission is not the rule. However, two noticeable exceptions to the zoonotic pattern of rickettsial disease can occur: the organism responsible for epidemic typhus (*Rickettsia prowazekii*) and the agent responsible for trench fever (*Rochalimaea quintana*) have the potential to spread rapidly within louse-ridden human populations. *See* ZOONOSES.

Historically, epidemic typhus has been an important factor in shaping Western civilization. Throughout most of European recorded history, epidemic typhus has been associated with famine and war and has often played a dominant role in determining the outcome of military campaigns. Epidemic typhus is transmitted between persons by the human body louse. Poor sanitation and inadequate living conditions create an ideal environment for human louse infestations, and hence for potential human epidemic typhus transmission. One estimate suggests that approximately 25 million persons suffered from epidemic typhus infections in Russia alone from the beginning of World War I through the Bolshevik Revolution; the mortality rate prior to the advent of appropriate antibiotics is estimated to have ranged from 10 to 66%. With the advent of higher standards of personal hygiene, the use of insecticides to control lice, and the recognition of the cycle of human epidemic typhus infections, the dominance of this disease over human life and political affairs has largely receded. However, human epidemic typhus still persists in some isolated human populations.

Even in louse-free human environments, recrudescent human typhus (Brill-Zinsser disease) may occur in individuals who have experienced a primary infection many years in the past. This disease complicates attempts at typhus eradication. In addition, an organism essentially identical to the causative agent of human epidemic typhus (*Rickettsia prowazekii*) exists in an almost totally separate, nonhuman cycle among North American flying squirrels. Sporadic human disease, usually linked to close human association with flying squirrels, occurs every year, and the disease syndrome is compatible with classic human epidemic typhus. Like most Brill-Zinsser disease in developed countries, secondary spread of the flying squirrel-associated disease between humans has not been documented; however, in the absence of an effective human epidemic typhus vector (the body louse), human-to-human transmission would not be expected.

The causative agent of another human disease, murine typhus (*Rickettsia typhi*), is closely related to the agent responsible for epidemic typhus. Murine typhus persists as a worldwide disease of rats and is sporadically vectored from rats to humans by fleas.

Human-to-human spread of the disease does not occur.

During World War I and, to a lesser extent, World War II, trench fever (caused by *Rochalimaea quintana*), like epidemic typhus, was a serious problem among military troops. Large numbers of troops were affected, although the clinical signs of the disease were relatively mild. Trench fever is infrequently diagnosed today.

The spotted fever group of rickettsiae represents an interesting collection of closely related organisms. All known spotted fever group organisms are transmitted by ticks. Despite a global distribution in the form of various diseases, nearly all spotted fever group organisms share close genetic, antigenic, and certain pathologic features. Examples of human diseases include Rocky Mountain spotted fever (in North and South America), fièvre boutonneuse or Mediterranean spotted fever (southern Europe), South African tick-bite fever (Africa), Indian tick typhus (Indian subcontinent), and Siberian tick typhus (northeastern Europe and northern Asia). If appropriate antibiotics are not administered, Rocky Mountain spotted fever, for example, is a life-threatening disease. As many as 1000 cases of Rocky Mountain spotted fever are diagnosed each year in the United States. Several other well-documented and rather common spotted fever group rickettsiae appear not to be pathogens of humans, although they can often be found infecting the same tick species that may transmit virulent human spotted fever disease.

Coxiella burnetii, the agent responsible for Q fever, commonly infects certain domestic animals, especially cattle, sheep, and goats, although disease is often inapparent. The organism is present only at low levels in these animals, except during parturition. This organism can also often be found in the mammary glands of these animals and is frequently shed in milk as well as in urine and feces. The organism is especially resistant to inactivation and can form highly infective aerosols. Thus, Q fever is a common disease of slaughter-house workers and others in close contact with the animals and animal products in which the organism is present. The acute stage of the disease in humans is usually self-limiting; however, if chronic endocarditis develops as a result of Q fever infection, the disease is often fatal.

Scrub typhus disease occurs throughout Asia, Australia, and the Pacific islands. The natural reservoirs of scrub typhus are native rats, and the disease is transmitted by chiggers. Considerable variation occurs among the protein antigens of *Rickettsia tsutsugamushi*, the causative agent; unlike other rickettsiae, this species appears to consist of a constellation of related disease-causing organisms.

Human disease associated with infection by an organism identical or very similar to *Ehrlichia canis*, a disease agent first recognized in dogs, has been recognized in the United States. The disease is believed to be transmitted by ticks, and the basic ecology of the disease may be similar to that of the spotted fever organisms. *See* INFECTIOUS DISEASE.

Russell L. Regnery

Bibliography. E. H. Lennette, P. Halonen, and F. A. Murphy (eds.), *Laboratory Diagnosis of Infectious Diseases: Principles and Practice*, vol. 2: *Viral, Rickettsial, and Chlamydial Diseases*, 1988; K. Maeda et al., Human infection with *Ehrlichia canis*, a leukocytic rickettsia, *N. Engl. J. Med.*, 316(14):853–856, 1987; D. H. Walker (ed.), *Biology of Rickettsial Diseases*, vols. 1 and 2, 1988.

Riemann surface

A generalization of the complex plane that was originally conceived to make sense of mathematical expressions such as $\sqrt{z}$ or $\log z$. These expressions cannot be made single-valued and analytic in the punctured plane $\mathbf{C}\backslash\{0\}$ (that is, the complex plane with the point 0 removed). The difficulty is that for some closed paths the value of the expression when reaching the end of the path is not the same as it is at the beginning. For example, the closed path can be chosen to be the unit circle centered at $z = 0$ and followed counterclockwise from $z = 1$. If $\sqrt{z}$ is assigned the value $+1$ at $z = 1$, its value at the end of the circuit is -1. Similarly, if $\log z$ is assigned the value 0 at $z = 1$, at the end of the circuit, allowing the values to change continuously, the value is $2\pi i$. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

Construction. The construction of an abstract surface for $\sqrt{z}$ on which $\sqrt{z}$ has a single value at each point and on which the values of $\sqrt{z}$ vary continuously around each closed path, always ending with the starting value, proceeds as follows. From the complex plane $\mathbf{C}$ the negative half of the real axis removed. The slit $\{z = x + iy : y = 0, 0 \le x\}$ is thought of as having two edges: an upper or $+$ edge and a lower or $-$ edge ($z = 0$ is not counted). Another copy of this slit plane is then placed over, and parallel to, the original, as the first and second floors in a building. Next, the two sheets are connected by attaching the $-$ edge of the bottom sheet to the $+$ edge of the top sheet, and the $+$ edge of the bottom sheet to the $-$ edge of the top sheet. The abstract surface thus constructed is not embedded in euclidean 3-space because the attachments cannot be done in 3-space without creating self-intersections. The expression $\sqrt{z}$ is defined at the point over z on the upper sheet to have positive imaginary part and on the lower to have negative part. If a point is imagined traveling counterclockwise around the unit circle centered at the origin in the bottom sheet, it hits the negative edge of the slit, ascends to the top sheet, leaves from the positive edge of the top slit, travels counterclockwise around the circle in the upper sheet, and finally returns to the initial point after descending from the negative edge of the top slit to the positive edge of the bottom. The values of $\sqrt{z}$ change continuously around this path and return to the initial value.

The abstract surface constructed is called the Riemann surface for $\sqrt{z}$. It is two-sheeted over the z plane, except that only one point lies over $z = 0$, which is therefore called a branch point. Actually the surface lies not just over $\mathbf{C}$ but over the extended

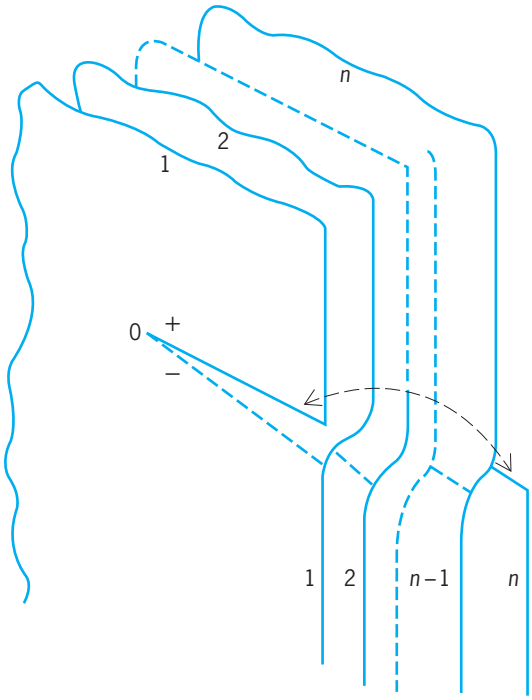

Fig. 1. Riemann surface for $\sqrt[n]{z}$. (*After A. G. Sveshnikov and A. N. Tikhonov, The Theory of Functions of a Complex Variable, Mir Publishers, 1973*)

plane $\mathbf{C} \cup \{\infty\}$ (or the Riemann sphere), formed by attaching the so-called point at infinity (∞) to the complex plane. There is also a branch point of order 2 over $z = \infty$, and the surface constructed is compact. The Riemann surface for $\sqrt[n]{z}$, which has n branches instead of 2, is shown in **Fig. 1**. *See* TOPOLOGY.

Likewise a Riemann surface can be reconstructed for expression (1). It is compact of genus g. A surface

$$[z(z - a_1) \ldots (z - a_{2g})]^{1/2} \qquad (1)$$
$$a_i \neq a_j \neq 0 \text{ for } i \neq j$$

can also be constructed for $\log z$. It is a doubly infinite spiral over $\mathbf{C}/\{0\}$, and is not compact.

Construction for algebraic curves. More generally, there is a compact Riemann surface R_w corresponding to any irreducible algebraic curve defined by Eq. (2),

$$w^n + p_1(z)w^{n-1} + \cdots + p_n(z) = 0 \qquad (2)$$

where $p_i(z)$, $1 \leq i \leq n$, are polynomials, $n > 2$. On R_w, $w = w(z)$ becomes a single-valued meromorphic function. Except for a finite number of points of the extended plane $\mathbf{C} \cup \{\infty\}$, for each $z = z_0$ there are n distinct values of w, $w_1(z_0), \ldots, w_n(z_0)$, and so R_w has n sheets over the extended plane. These can be fastened together according to the nature of the curve to form a single surface. The global function $w(z)$ has the values $w_1(z_0), \ldots, w_n(z_0)$ at the n points over z_0. The construction for algebraic curves was first carried out by B. Riemann. *See* ALGEBRAIC GEOMETRY.

Construction for power series. Alternatively, it is possible to start with a convergent power series about, for example, $z = 1$, given by Eq. (3) [for example,

$$w = \sum_{n=0}^{\infty} a_n(z - 1)^n \qquad (3)$$

a local solution of an algebraic equation as above]. The power series together with all its analytic continuations determines a Riemann surface R_w which in general is not compact but on which $w = w(z)$ becomes a single-valued meromorphic function.

Topological properties. In the early 1920s, H. Weyl defined a Riemann surface in the context of the newly developing field of topology. This definition does not depend on having a particular function in hand, and does not require the use of such seemingly artificial devices as slits.

A Riemann surface is a two-dimensional connected topological manifold (surface) with a complex structure. This means there is a covering of the surface with open neighborhoods $\{U_i\}$ with the following properties: (1) Each neighborhood U_i corresponds (is homeomorphic to) to a neighborhood V_i in $\mathbf{C}$; this correspondence can be denoted as z_i: $U_i \to V_i$. (2) If two neighborhoods, U_i and U_j, overlap, then the composition $z_j \circ z^{-1}{}_i$, which is defined on $z_i(U_i \cap U_j)$ in $\mathbf{C}$, is a conformal map onto $z_j(U_i \cap U_j)$. *See* MANIFOLD (MATHEMATICS).

For the surface for $\sqrt{z}$ constructed above, the correspondence z_i is given by the projection to the complex plane except in neighborhoods of the branch points over 0 and ∞. About those points the correspondence is described by $\sqrt{z}$ or $\sqrt{z^{-1}}$, which determines a one-to-one mapping of neighborhoods onto neighborhoods of the origin in $\mathbf{C}$. Every Riemann surface has meromorphic functions and therefore is associated with power series.

A Riemann surface can be thought of as a topological surface together with a rule for measuring angles at points. The angles about a point must sum to 2π. An angle-preserving orientation-preserving mapping between Riemann surfaces is called conformal. In the theory, no distinction is made between conformally equivalent surfaces. *See* CONFORMAL MAPPING.

Any smooth surface embedded in euclidean space can be made into a Riemann surface. Angles between intersecting smooth curves on the surface can be measured simply by using the euclidean metric of the surrounding space to measure the angle of intersection of their tangent lines.

Conversely, it is known that any Riemann surface, compact or noncompact, can be conformally mapped onto a surface embedded in euclidean 3-space as a closed subset. In fact, each Riemann surface can be so embedded in infinitely many ways.

Likewise, a polyhedral surface can be made into a Riemann surface. On the flat pieces the angles are as given. But before measuring angles at a vertex, the surface must be flattened out about the vertex, for the angles about each point must add to 2π. This is done by applying a map of the form z^α, where $\alpha = 2\pi/\theta$, θ being the sum of the angles at the vertex.

A Riemann surface that is compact without boundary can be conformally mapped onto the

Riemann surface associated with some algebraic curve.

A 2-dimensional Riemannian manifold can be made into a Riemann surface. Local conformal mappings into the plane are determined by solving the Beltrami equation; these determine the conformal structure. *See* RIEMANNIAN GEOMETRY.

The most familiar Riemann surfaces are the regions of the plane or the sphere (the Riemann sphere), such as the 3-punctured sphere, an annulus, and the complement of the Cantor set. Via steographic projection, the complex plane is mapped onto the unit sphere so that the point at infinity becomes the north pole. Thus, the complex plane is alternately described as the once-punctured sphere.

Uniformization. Perhaps the most important single theorem of complex analysis is the uniformization theorem. It says that a simply connected Riemann surface can be mapped conformally onto exactly one of: the sphere, the complex plane, or the unit disk $\{z : |z| < 1\}$.

If M is a manifold, then, corresponding to each subgroup H of the fundamental group of M, there is another manifold called a covering manifold of M. If H is the trivial group, then the covering is simply connected and is called the universal covering. This is a standard construction in topology, but its application to Riemann surfaces is important because the covering surface of a Riemann surface is also a Riemann surface. In particular, uniformization may be applied to the universal covering surface of a Riemann surface. The universal cover of the Riemann surface R is the unit disk except in the cases that R can be mapped conformally onto the sphere itself (spherical case), or the complex plane, the 1-punctured plane, or a torus (euclidean cases). As a consequence, every Riemann surface has an intrinsic geometry: spherical, euclidean, or hyperbolic. *See* GROUP THEORY.

Almost all Riemann surfaces fall into the third category: their universal covering surfaces can be mapped conformally onto the unit disk, and they have an intrinsic hyperbolic geometry coming from the hyperbolic geometry of the unit disk in the hyperbolic metric given by Eq. (4), which has constant

$$ds = \frac{2|dz|}{1 - |z|^2} \tag{4}$$

gaussian curvature -1. *See* DIFFERENTIAL GEOMETRY; GEOMETRY; NONEUCLIDEAN GEOMETRY.

Put another way, each such Riemann surface R can be "unrolled" onto the unit disk. In the unit disk, R can be modeled as a polygon in the hyperbolic geometry. The sides of the polygon are arranged in pairs. A pair of sides are identified with each other by an orientation-preserving isometry that maps one side onto the other. When the polygon is "rolled up" by identifying each pair of sides, a Riemann surface is obtained which is conformally equivalent to R. **Figure 2** shows an ideal hyperbolic square (it has hyperbolic area 2π) and the first few generations of the tiling of the hyperbolic plane by isometric copies

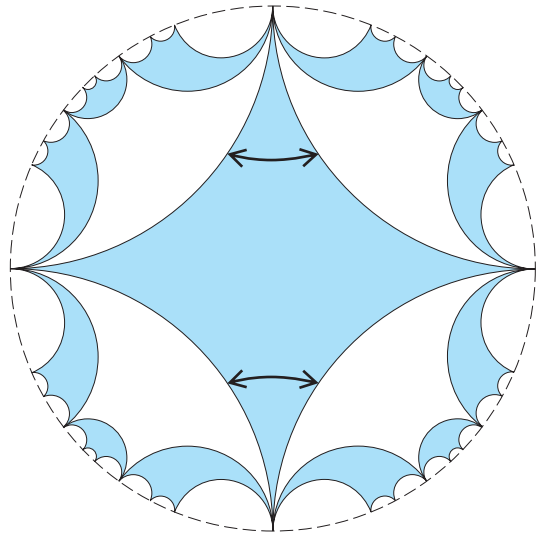

Fig. 2. Rolling up an ideal hyperbolic square. (*After W. Dyck, Ueber Untersuchen und Aufstellung von Gruppe und Irrationalitaet regulaerer Riemannscher Flaechen, Math. Ann., 17:473–509, 1880*)

of itself; "ideal" means its vertices are "at infinity" in terms of the metric and its vertex angles there are zero. (Figure 2 is an approximation; there are really infinitely many white pieces along the edge of the disk, getting smaller and smaller.) When the pairs of sides indicated by double-headed arrows in Fig. 2 are identified, the square "rolls up" to become the 3-punctured sphere. (In the other direction, one can take a 3-punctured sphere and indicate how a tile is placed within it: Take one of the punctures and draw arcs from that puncture to each of the two others. This procedure cuts the 3-punctured sphere into a 4-sided figure that represents the original tile. This operation is analogous to making a torus from a sheet of paper by first rolling the sheet up to get a cylinder and then pasting the two ends of the cylinder together to get a torus.)

Moduli. The conformal equivalence classes of all compact Riemann surfaces of fixed genus $g \geq 0$ with $b \geq 0$ boundary components and $n \geq 0$ punctures such that $d = 6g + 3b + 2n - 6 \geq 1$ can be parametrized: The collection forms a simply connected space $M_{g,b,n}$ with singularities (an orbifold), called moduli space, of real dimension d. Distinct points correspond to surfaces which cannot be mapped conformally onto each other. If a point is imagined moving through $M_{g,b,n}$, the corresponding Riemann surface changes continuously. If $b = 0$, the space $M_{g,0,n}$ has a compact boundary $\partial M_{g,0,n}$. As a point approaches it, the Riemann surface degenerates, that is, gets "pinched." Thus, a 4-punctured sphere can become pinched so as to result in two 3-punctured spheres. A compact surface of genus 2 can become pinched to form two 1-punctured tori, or a single surface of genus 1 with two punctures, or two 3-punctured spheres.

Thus, a Riemann surface can be studied from three different vantage points: its complex analysis, its geometry in the hyperbolic metric (for most surfaces), and for compact surfaces, the algebraic geometry

of its associated polynomial. In addition, Riemann surfaces can be gathered together in spaces $M_{g,b,n}$ of topologically similar surfaces. The study of these spaces is called Teichmüller theory. Albert Marden

Bibliography. L. V. Ahlfors, *Lectures on Quasiconformal Mappings*, 1966, reprint 1987; L. V. Ahlfors and L. Sario, *Riemann Surfaces*, 1960, reprint, Princeton University Press, 1996; C. Chevalley, *Introduction to the Theory of Algebraic Functions of One Variable*, 1951, reprint, American Mathematical Society, 1987; H. M. Farkas and I. Kra, *Riemann Surfaces*, 2d ed., Springer-Verlag, 1991; O. Forster, *Lectures on Riemann Surfaces*, 1981, reprint, Springer-Verlag, 1993; F. Kirwan, *Complex Algebraic Curves*, Cambridge University Press, 1992; G. Springer, *Introduction to Riemann Surfaces*, 2d ed., Chelsea Publishing, 1981.

Riemannian geometry

The geometry of riemannian manifolds. Riemannian geometry was initiated by B. Riemann in 1854, following the pioneering work of C. F. Gauss on surface theory in 1827. Riemann introduced a coordinate space in which the infinitesimal distance between two neighboring points is specified by a quadratic differential form, given below. Such a space is a natural generalization of euclidean geometry and Gauss's geometry of surfaces in three-dimensional euclidean space, as well as the noneuclidean geometries: hyperbolic geometry (previously discovered by J. Bolyai and N. I. Lobachevsky) and elliptic geometry. A riemannian manifold is a topological space that further generalizes this notion. Riemannian geometry derives great importance from its application in the general theory of relativity, where the universe is considered to be a riemannian manifold. *See* DIFFERENTIAL GEOMETRY; EUCLIDEAN GEOMETRY; NONEUCLIDEAN GEOMETRY; RELATIVITY.

Riemannian spaces. An n-dimensional riemannian space is a space whose points can be characterized by n coordinates u^i ($i = 1, 2, \ldots, n$), and where an infinitesimal distance ds between two points with coordinate differences du^i is given by a quadratic differential form, Eq. (1), to be called the metric. In

$$ds^2 = \sum_{i,j} g_{ij}\, du^i\, du^j \qquad (1)$$
$$g_{ij} = g_{ji} \qquad 1 \le i, j \le n$$

general, the quantities g_{ij} are functions of the coordinates u^i. This allows the definition of the arc length of a curve $u^i = u^i(t)$, $t_0 \le t \le t_1$. Indeed, taking the square root of both sides of Eq. (1) and integrating over t from t_0 to t_1 yields Eq. (2) for the arc length s.

$$s = \int_{t_0}^{t_2} \sqrt{\sum_{i,j} g_{ij} \frac{du^i}{dt} \frac{du^j}{dt}}\, dt \qquad (2)$$

Other geometric notions, such as the angle between intersecting curves and the volume of a domain, can also be defined. Generally the ds^2 is supposed to be positive definite, meaning that it is greater than zero for all values of the du^i, unless they are all zero; or at least nondegenerate, meaning that inequality (3) is

$$\det(g_{ij}) \neq 0 \qquad (3)$$

satisfied, where $\det(g_{ij})$ is the determinant of the matrix whose elements are the g_{ij}. *See* DETERMINANT; INTEGRATION.

Examples of riemannian spaces include the following:

1. The n-dimensional euclidean space E^n has the metric given by Eq. (4).

$$ds^2 = (du^1)^2 + \cdots + (du^n)^2 \qquad (4)$$

2. More generally, elliptic and hyperbolic spaces are given by the noneuclidean metric of Eq. (5), where K is a constant and r^2 is given by Eq. (6).

$$ds^2 = \frac{(du^1)^2 + \cdots + (du^n)^2}{\left(1 + \frac{K}{4} r^2\right)^2} \qquad (5)$$

$$r^2 = (u^1)^2 + \cdots + (u^n)^2 \qquad (6)$$

The space is called elliptic if $K > 0$, and hyperbolic if $K < 0$; K is the curvature of the space.

3. If the assumption of positive definiteness is dropped, the n-dimensional lorentzian space with the metric given by Eq. (7) is a generalization of a

$$ds^2 = (du^1)^2 + \cdots + (du^{n-1})^2 - (du^n)^2 \qquad (7)$$

riemannian space. A four-dimensional lorentzian space characterizes the physical universe in the special theory of relativity. *See* SPACE-TIME.

4. If E^N is the euclidean space of dimension N with the coordinates $x_1, \ldots, x_N$, a submanifold M of dimension $n < N$ is defined by Eq. (8), where the functional

$$x_A = x_A(u^1, \ldots, u^n) \qquad 1 \le A \le N \qquad (8)$$

matrix $(\partial x_A/\partial u^i)$ is of rank n. Then M has the induced riemannian metric given by Eq. (9). The case $N = 3$, $n = 2$, is treated in the surface theory of Gauss.

$$ds^2 = \sum_{i,j,A} \frac{\partial x_A}{\partial u^i} \frac{\partial x_A}{\partial u^i}\, du^i\, du^j \qquad (9)$$

Tensors. Riemannian geometry is local in the sense that it is valid in a neighborhood of the u^i coordinates. It studies properties that are invariant under coordinate transformations of the form of Eq. (10),

$$u'^i = u'^i(u^1, \ldots, u^n) \qquad 1 \le i \le n \qquad (10)$$

where the functions are smooth with the functional determinant $\det(\partial u'^i/\partial u^j) \neq 0$. Thus, in the new coordinates, the metric is given by Eq. (11). This implies that the g'_{ij} are related to the g_{ij} by Eq. (12).

$$ds^2 = \sum_{i,j} g'_{ij}\, du'^i\, du'^j \qquad g'_{ij} = g'_{ji} \qquad (11)$$

$$g'_{ij} = \sum_{k,l} g_{kl} \frac{\partial u^k}{\partial u'^i} \frac{\partial u^l}{\partial u'^j} \qquad (12)$$

Following Einstein, from now on the convention will be adopted that repeated indices mean summation. It will also be assumed that all small Latin indices run from 1 to n, unless otherwise specified. A system of quantities $T^{i_1\cdots i_r}{}_{j_1\cdots j_s}$ associated with each coordinate system u^i is called a tensor, contravariant of order r and covariant of order s, if under the transformation of Eq. (10) they are transformed according to the rule given by Eq. (13). It is called a contravariant

$$T'^{i_1\cdots i_r}{}_{j_1\cdots j_s} = T^{k_1\cdots k_r}{}_{l_1\cdots l_s} \times \frac{\partial u'^{i_1}}{\partial u^{k_1}}\cdots\frac{\partial u'^{i_r}}{\partial u^{k_r}}\frac{\partial u^{l_1}}{\partial u'^{j_1}}\cdots\frac{\partial u^{l_s}}{\partial u'^{j_s}} \tag{13}$$

vector if $r = 1$ and $s = 0$; it is called a covariant vector if $r = 0$ and $s = 1$. Thus, the g_{ij}, called the fundamental tensor, is a covariant tensor of order two and is symmetric. *See* TENSOR ANALYSIS.

Covariant differentiation. It is useful to define the matrix (g^{ij}) to be the inverse matrix of (g_{ij}), so that Eq. (14) is satisfied. Another set of useful quantities are the Christoffel symbols, given by Eq. (15). They

$$g_{ij}g^{jk} = \delta_i^k = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases} \tag{14}$$

$$\Gamma_{ijk} = \frac{1}{2}\left(\frac{\partial g_{ij}}{\partial u^k} + \frac{\partial g_{jk}}{\partial u^i} - \frac{\partial g_{ik}}{\partial u^j}\right) \tag{15}$$

can be used to define a covariant differentiation. For $r = 2$, $s = 1$, the covariant derivatives of $T^{i_1 i_2}{}_j$ are given by Eq. (16). A similar formula is valid for arbitrary values of r and s. A fundamental fact is that $T^{i_1 i_2}{}_{j|m}$ is still a tensor, of one covariant order higher than $T^{i_1 i_2}{}_j$. The Christoffel symbols are characterized by the symmetry condition of Eq. (17) and the property that the covariant derivatives of the fundamental tensor are identically zero.

$$T^{i_1 i_2}{}_{j|m} = \frac{\partial T^{i_1 i_2}{}_j}{\partial u^m} + \Gamma_k{}^{i_1}{}_m T^{ki_2}{}_j + \Gamma_k{}^{i_2}{}_m T^{i_1 k}{}_j - \Gamma_j^k{}_m T^{i_1 i_2}{}_k \tag{16}$$

$$\Gamma_{ik}^j = \Gamma_{ki}^j \tag{17}$$

In particular, a contravariant vector T^i has the covariant differential given by Eq. (18), where ω_j^i is given by Eq. (19).

$$DT^i = dT^i + \omega_j^i T^j \tag{18}$$

$$\omega_j^i = \Gamma_{jk}^i\, du^k \tag{19}$$

Parallelism and geodesics. If the vectors $T^i(t)$ are given along a curve $u^i = u^i(t)$, they are parallel along the curve in the sense of T. Levi-Civita if Eq. (20) is

$$DT^i(t) = 0 \tag{20}$$

satisfied. This curve is called a geodesic if its tangent vectors du^i/dt are parallel. Geodesics are therefore the integral curves of the system of ordinary differential equations (21) of the second order.

$$\frac{d^2u^i}{dt^2} + \Gamma_{jk}^i \frac{du^j}{dt}\frac{du^k}{dt} = 0 \tag{21}$$

Geometrically, geodesics are the shortest curves joining two points. In other words, Eqs. (21) are the Euler equations of the variational problem given by Eq. (22).

$$\delta \int ds = 0 \tag{22}$$

See CALCULUS OF VARIATIONS.

For the euclidean space given in Eq. (4), the fundamental tensor and Christoffel symbols are given by Eqs. (23). In this case, parallelism of a vector means

$$g_{ik} = \delta_{ik} \qquad \Gamma_{ijk} = 0 \tag{23}$$

constancy of its components T^i. Also, the geodesics are the straight lines defined by Eqs. (24). Thus,

$$\frac{d^2u_i}{dt^2} = 0 \tag{24}$$

the above definitions of parallelism and geodesics are generalizations of standard notions of euclidean geometry. An essential difference is that Eqs. (20), being ordinary differential equations, are defined only along a curve, so that parallel displacement of a vector from one point to another depends on the curve joining them.

Riemann-Christoffel tensor. The properties of a riemannian space are described by its Riemann-Christoffel curvature tensor. It is defined by Eq. (25),

$$R_{ijrs} = \frac{\partial \Gamma_{ijs}}{\partial u^r} - \frac{\partial \Gamma_{ijr}}{\partial u^s} + g^{kb}(\Gamma_{ikr}\Gamma_{jbs} - \Gamma_{iks}\Gamma_{jbr}) \tag{25}$$

and satisfies the symmetry relations of Eqs. (26).

$$\begin{aligned} R_{ijrs} &= -R_{jirs} = -R_{ijsr} \\ R_{ijrs} &= R_{rsij} \\ R_{ijkl} &+ R_{iklj} + R_{iljk} = 0 \end{aligned} \tag{26}$$

The tensor has the following geometrical interpretation: If ξ^i and η^j are two vectors, and σ is the plane spanned by them, then the Plücker coordinates of the plane σ are given by Eq. (27), and the quantity K defined by Eq. (28) depends only on the point u

$$p^{ij} = \xi^i\eta^j - \xi^j\eta^i \tag{27}$$

$$K(u,\sigma) = -\frac{R_{ijkl}p^{ij}p^{ij}}{(g_{ik}g_{jl} - g_{il}g_{jk})p^{ij}p^{ij}} \tag{28}$$

and the plane σ through u. It is called the sectional or riemannian curvature at (u, σ). Geometrically, it is equal to the gaussian curvature of the surface generated by the geodesics through u and tangent to the vectors of σ.

The simplest riemannian spaces are those of constant curvature, for which the sectional curvature is independent of (u,σ). Their condition is Eq. (29),

$$R_{ijkl} = -K(g_{ik}g_{jl} - g_{il}g_{jk}) \tag{29}$$

where K is a constant. The metric given by Eq. (5) has this property.

Other important tensors can be constructed from the Riemann-Christoffel tensor. Thus, Eq. (30) defines a tensor called the Ricci tensor, and the quantity R defined by Eq. (31) is a scalar, called the scalar curvature.

$$R_{ik} = g^{jl} R_{ijkl} = R_{ki} \tag{30}$$

$$R = g^{ik} R_{ik} \tag{31}$$

The four-dimensional lorentzian space characterizes the universe of the theory of general relativity. It is restricted by the Einstein field equations (32),

$$R_{ik} - {}^1\!/_2\, g_{ik} R = 8\pi\kappa T_{ik} \tag{32}$$

where κ is a constant and T_{ik} is the energy-stress tensor.

Manifolds and bundles. At the foundation of riemannian geometry is the notion of a riemannian manifold. A differentiable manifold is a topological space with a covering by coordinate neighborhoods, such that in the domain where the u^i coordinates and the u'^i coordinates are both valid, they are related by the transformation of Eq. (10) with the conditions mentioned there. The manifold is called riemannian if a riemannian metric is defined in each neighborhood that gives the same value of ds^2 in the intersection of any two neighborhoods; that is, Eq. (12) is satisfied. Riemannian geometry is divided between local and global: the former dealing with properties in a neighborhood, the latter with properties of the manifold as a whole. *See* MANIFOLD (MATHEMATICS); TOPOLOGY.

Examples of local properties include the Levi-Civita parallelism defined by Eq. (20) and the Riemann-Christoffel tensor. An example of a global property is the notion of completeness: A riemannian manifold is called complete if every geodesic can be indefinitely extended, with the possibility of a closed geodesic repeating itself (such as a great circle on the sphere) not excluded. On a complete riemannian manifold any two points can be joined by a shortest geodesic.

The tensors have a global formulation, in the notion of tensor bundles. If the riemannian manifold M is of dimension n, the contravariant vectors, or tangent vectors, at a point x in M, form the tangent space T_x and the covariant vectors, or covectors, form the cotangent space T_x^*. They are dual vector spaces in the sense that there is a pairing that associates to each X in T_x and ω in T_x^* a quantity $\langle X,\omega\rangle$, which is a real number and is linear in each of the arguments. In fact, if X has the components X^i and ω has the components ω_i in the same local coordinate system, then $\langle X,\omega\rangle$ is given by Eq. (33).

$$\langle X, \omega\rangle = X^i \omega_i \tag{33}$$

More generally, the tensors $T_x(r,s)$ of contravariant order r and covariant order s at a point x in M form a vector space of dimension n^{r+s}, which is given by Eq. (34). A tensor bundle over M is defined to be the

$$T_x(r, s) = \underbrace{T_x \otimes \cdots \otimes T_x}_{r} \otimes \underbrace{T_x^* \cdots \otimes T_x^*}_{s} \tag{34}$$

union of the vector spaces $T_x(r,s)$ for each point x in M, denoted

$$\bigcup_{x\in M} T_x(r, s)$$

together with the projection mapping π defined by Eq. (35). For $r = 1$, $s = 0$ and for $r = 0$, $s = 1$,

$$\pi : \bigcup_{x\in M} T_x(r, s) \to M \pi T_x(r, s) = x \tag{35}$$

these are, respectively, the tangent and cotangent bundles over M. A mapping given by Eq. (36) such

$$s : M \to \bigcup_{x\in M} T_x(r, s) \tag{36}$$

that $\pi \circ s =$ identity, that is, the assignment of a tensor to every point of M, is called a section of the tensor bundle or a tensor field.

Space-forms and symmetric spaces. Some results on global riemannian geometry will now be described. The riemannian manifolds of constant curvature are locally characterized by Eq. (29). They are called space-forms, because they are locally the euclidean and noneuclidean spaces. The metric is given by Eqs. (5). For $K = 0$, such a space is a euclidean space E^n with the coordinates $u^1, \ldots, u^n$. For $K < 0$, the space is restricted to the domain of points satisfying Eq. (37), and is called the hyperbolic space

$$(u^1)^2 + \cdots + (u^n)^2 < -\frac{4}{K} \tag{37}$$

H^n. For $K > 0$, E^n can be compactified by the addition of a so-called point at infinity, making it into the n-dimensional sphere S^n; the resulting space is called the elliptic or spherical space. All these spaces, the euclidean, the hyperbolic, and the spherical, are simply connected (that is, any closed curve in the space can be deformed continuously to a point).

A fundamental theorem says that a complete, simply connected riemannian manifold of constant curvature K is isometric to the euclidean space, the hyperbolic space, or the spherical space, according to whether $K = 0$, $K < 0$, or $K > 0$. (An isometry is a one-to-one mapping that preserves the metric ds^2.) It follows that a complete riemannian manifold of constant curvature (which is not necessarily simply connected) is covered by one of the above three spaces. Two important examples are the following:

1. The manifold constructed by identifying in E^n the points $(u^1, \ldots, u^n)$ and $(u^1 + 1, \ldots, u^n + 1)$. The result is an n-dimensional torus, having a metric with $K = 0$.
2. The manifold constructed by identifying the antipodal points of S^n. The result is the real projective space, having a metric with $K > 0$.

A more general family of riemannian spaces consists of the symmetric spaces. They are characterized

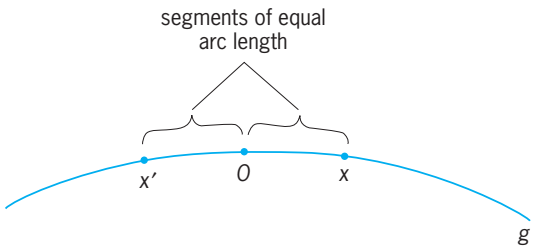

Fig. 1. Configuration of points in definition of a symmetry about a point *0*.

by the vanishing of the covariant derivatives of the Riemann-Christoffel tensor, that is, by Eq. (38). Ge-

$$R_{ijkl|m} = 0 \tag{38}$$

ometrically this can be expressed by the property that the symmetries about the points of the space are isometries. (A symmetry about a point 0 of a riemannian space is a mapping that takes any nearby point x into a point x' such that x, 0, and x' lie on a single geodesic g in the configuration shown in **Fig. 1**, and the arc lengths of the segments $x'0$ and $0x$ of g are equal.) Symmetric riemannian spaces are of great importance in the applications of riemannian geometry.

Gauss-Bonnet formula. In two dimensions, this formula states the following: If D is a domain with a sectionally smooth boundary, then Eq. (39) is satis-

$$\sum(\pi - \alpha) + \int_{\alpha D} k_g \, ds + \int\int_D K \, dA = 2\pi\chi(D) \tag{39}$$

fied, where the first term is the sum of the exterior angles at the corners, the second term is the integral of the geodesic curvature along the sides, the third term is the integral of the gaussian curvature over D, and $(D)_\chi$ is the Euler characteristic of D. The Euler characteristic can be defined by dividing up D into cells (**Fig. 2**); then the Euler characteristic is given by Eq. (40). For a rectilinear triangle in the euclidean

$$\chi(D) = \text{number of vertices} - \text{number of edges} + \text{number of cells} \tag{40}$$

plane, $k_g = 0$, $K = 0$, and Eq. (39) indicates that the angle-sum is equal to π.

In higher dimensions, for the sake of simplicity, the theorem will be given only for a compact oriented riemannian manifold M of dimension $n = 2m$ without boundary. The so-called curvature form is given by Eq. (41), where the wedge symbol $\wedge$ de-

$$\Omega_{ij} = \sum R_{ijkl} du^k \wedge du^l \tag{41}$$

notes exterior or wedge multiplication. The pfaffian is defined by Eq. (42), where $\epsilon_{i_1 \ldots i_n}$ is $+1$ or -1 ac-

$$Pf = \sum \epsilon_{i_1 \ldots i_n} \Omega_{i_1 i_2} \wedge \ldots \wedge \Omega_{i_{n-1} i_n} \tag{42}$$

cording as its indices form an even or odd permutation of $1, \ldots, n$, and is otherwise zero, and the sum is extended over all indices from 1 to n. Then the Gauss-Bonnet theorem states that Eq. (43) holds,

$$(-1)^m \frac{1}{2^{2m}\pi^m m!} \int_M Pf = \chi(M) \tag{43}$$

where $\chi(M)$ is the Euler-Poincaré characteristic of M. The Euler-Poincaré characteristic, an important topological invariant, can be defined by dividing up M into cells; then the Euler-Poincaré characteristic is given by Eq. (44), a generalization of Eq. (40). The

$$\chi(M) = \sum_{0 \le k' \dim M} (-1)^k \text{ (number of } k\text{-cells)} \tag{44}$$

Gauss-Bonnet theorem gives a relation between local and topological invariants.

DeRham and Hodge theorems. A geometrical idea that leads to important topological invariants is that of bounding or connectivity. For instance, every closed curve on the two-dimensional sphere bounds a domain, while there are closed curves on the two-dimensional torus that do not bound. This idea is developed into the homology theory in algebraic topology. Dual to this is the notion of exterior differential forms, given by Eq. (45), where the $a_{i_1 \ldots i_p}$ are smooth

$$\alpha = a_{i_1 \ldots i_p} du^{i_1} \wedge \cdots \wedge du^{i_p} \tag{45}$$

functions in the coordinates u_i, and the wedge multiplication of differentials is antisymmetric, that is, Eq. (46) holds. The form α is said to be of degree p.

$$du^i \wedge du^j = -du^j \wedge du^i \tag{46}$$

The exterior derivative is defined by Eq. (47), and is

$$d\alpha = da_{i_1 \ldots i_p} \wedge du^{i_1} \wedge \cdots \wedge du^{i_p} \tag{47}$$

of degree $p + 1$. The exterior differentiation, as an operator on exterior differential forms, can be said to be adjoint to the boundary operator, in view of

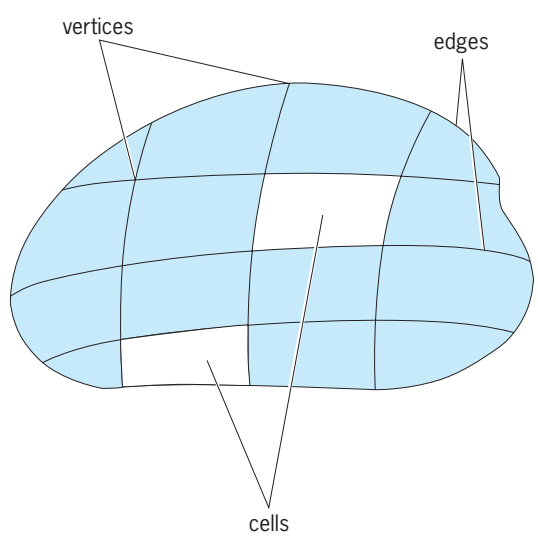

Fig. 2. Division of a domain *D* into cells in the definition of the Euler characteristic.

Stokes' theorem, which states that Eq. (48) holds,

$$\int_D d\alpha = \int_{\alpha D} \alpha \tag{48}$$

where D is a $(p+1)$-dimensional domain and αD is its boundary. It is easily verified that $dd\alpha = 0$. *See* STOKES' THEOREM.

All the smooth forms of degree p on a manifold M constitute a vector space A^p (of infinite dimension). The form α is said to be closed if $d\alpha = 0$. If C^p is the space of closed forms of degree p, then dA^{p-1} is a substance of C^p. DeRham's theorem states that if the manifold M is compact, the quotient space C^p/dA^{p-1} (also denoted C^p mod dA^{p-1}) is of finite dimension b^p, which is the pth Betti number of M, an important topological invariant of M.

When M is a riemannian manifold of dimension n, this theorem can be refined. Based on the inner product in the cotangent space T^*_x, x in M, the $*$-operator can be defined, which maps a p-form α into an $(n-p)$-form $*\alpha$. The codifferential δ defined by Eq. (49) maps a p-form into a $(p-1)$-form. The op-

$$\delta = \pm * d* \tag{49}$$

erator Δ defined by Eq. (50) maps a p-form into a

$$\Delta = d\delta + \delta d \tag{50}$$

p-form, and is called the Laplace operator. This is a generalization, to a riemannian manifold, of the classical Laplace operator in analysis. In analogy to the classical case, a form α is called harmonic if $\Delta\alpha = 0$. *See* LAPLACE'S DIFFERENTIAL EQUATION.

Hodge's theorem states that for every compact riemannian manifold a coset of C^p mod dA^{p-1} contains exactly one p-form that is harmonic.

Generalizations. Riemannian geometry has several interesting generalizations, including finslerian geometry, symplectic geometry, and the complex manifolds.

Finslerian geometry. Riemannian geometry can be considered to be based on the integral in Eq. (2), giving the arc length of a curve. The geometry based on the more general integral of Eq. (51), where F

$$s = \int_{t_0}^{t_1} F\left(u^i, \frac{du^i}{dt}\right) dt \tag{51}$$

is positively homogeneous of degree one, that is, Eq. (52) holds, is called finslerian geometry. Besides

$$F(u^i, \lambda v^i) = \lambda F(u^i, v^i) \qquad \lambda > 0 \tag{52}$$

riemannian geometry, another important special case of finslerian geometry is minkowskian geometry, where $F(du^i/dt)$ is a function of du^i/dt only. In this case, the geodesics are straight lines, and a theory of convex bodies can be developed, which plays an important role in number theory.

Symplectic geometry. Another interesting object is the replacement of the symmetric covariant tensor g_{ij} by an antisymmetric one, that is, a covariant tensor a_{ij} that satisfies Eq. (53). The matrix (a_{ij}) is taken to be

$$a_{ij} = -a_{ji} \tag{53}$$

of maximum rank, so that the space is of even dimension. The tensor is the same as the exterior quadratic differential form given by Eq. (54). The space is called

$$\alpha = a_{ij}\, du^i \wedge du^j \tag{54}$$

symplectic if such an α is given, satisfying $d\alpha = 0$. The resulting geometry is called symplectic geometry. It is a fundamental structure in mechanics. *See* MATRIX THEORY.

Unlike riemannian geometry, symplectic geometry has no local invariants, for a theorem of J. G. Darboux states that there are local coordinates where α takes the normal form given by Eq. (55).

$$\alpha = du^1 \wedge du^2 + \cdots + du^{n-1} \wedge du^n \tag{55}$$

Complex manifolds. Finally, there is the important generalization when the coordinates u^i are complex numbers, and the change of coordinates of Eq. (10) is given by holomorphic functions. Such a manifold is called a complex manifold. It is of complex dimension n and even real dimension $2n$. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

Even the case $n = 1$ is of great importance. A complex manifold of (complex) dimension one is called a Riemann surface. It is a basic object in complex function theory and plays a fundamental role in string theory in theoretical physics. *See* RIEMANN SURFACE; SUPERSTRING THEORY.

A complex manifold is called hermitian if there is a hermitian differential form, given by Eq. (56), where

$$ds^2 = h_{ij}\, du^i\, d\overline{u}^j \qquad \overline{h_{ij}} = h_{ji} \tag{56}$$

the bar denotes complex conjugation. Hermitian geometry is the counterpart of riemannian geometry for complex manifolds. S. S. Chern

Bibliography. E. Cartan, *Geometry of Riemannian Spaces*, 1983; I. Chavel, *Riemannian Geometry: A Modern Introduction*, 1995; S. Chern, *Complex Manifolds Without Potential Theory*, 1978; L. P. Eisenhart, *Riemannian Geometry*, 2d ed., 1950; W. Klingenberg, *Riemannian Geometry*, 2d ed., 1995; S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, vol. 1, 1963, vol. 2, 1969; C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, 1973.

Rift valley

One of the geomorphological expressions between two tectonic plates that are opening relative to each other or sliding past each other. The term originally was used to describe the central graben structures of such classic continental rift zones as the Rhinegraben and the East African Rift, but the definition now encompasses mid-oceanic ridge systems with central valleys such as the Mid-Atlantic Ridge. *See* MID-OCEANIC RIDGE; PLATE TECTONICS.

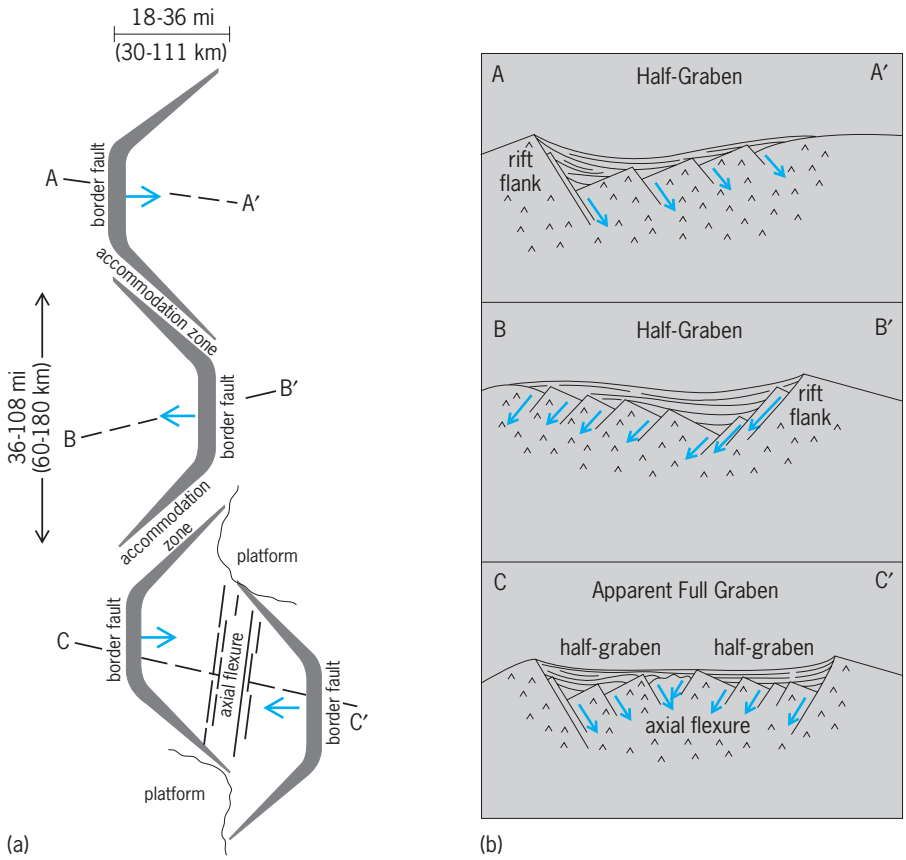

Sketch showing several ways in which half-graben units can link to make rift zones. (*a*) Map view. (*b*) Cross sections. A-A′ and B-B′ show simple half-grabens that alternate polarity along strike. C-C′ shows a typical morphology across an overlapping, facing arrangement of half-grabens.

Continental and oceanic rift valleys are end members in what many consider to be an evolutionary continuum. In the case of continental rift valleys, plate separation is incomplete, and the orientation of the stress field relative to the rift valley can range from nearly orthogonal to subparallel. Strongly oblique relationships are probably the norm. Also, volcanism is not a universal phenomenon in continental rift valleys, and when it does occur, the positioning, expressions, and chemistries of the effluents are complicated and variable. In contrast, oceanic rift valleys mark the place where the trailing edges of two distinctly different plates are separating. The separation is complete, and the spreading is organized and focused, resulting in rift valleys that tend to be oriented orthogonal or suborthogonal to the spreading directions. Basaltic magmas with a tholeiitic chemistry upwell from the asthenosphere into oceanic rift valleys, essentially filling the void that otherwise would develop between the separating plates. This is how new oceanic lithosphere is generated. *See* ASTHENOSPHERE; MAGMA.

The basic cross-sectional form of rift valleys consists of a central graben surrounded by elevated flanks. Considerable debate exists regarding how much of the elevation of the flanks is residual and how much represents actual uplift, but it is almost universally accepted that the central grabens of continental rift valleys are subsidence features. The crystalline basement floors of some parts of the Tanganyika and Malawi (Nyasa) rift valleys in East Africa lie more than 5 mi (9 km) below elevated flanks. *See* GRABEN.

In continental rift valleys the true cross-sectional form is typically asymmetric, with the rift floors tilted toward the most elevated flank. Most of the subsidence is controlled by one border fault system, and most of the internal faults parallel the dip of the border faults. Sedimentation and volcanism tend to obscure this basic symmetry, especially where deposits are very thick. The polarity of these half-graben structures alternates along the strike of many rift zones (see **illus.**). This creates a kind of segmentation in which half-grabens of opposite dip domains are separated by structures referred to as accommodation zones (or transfer faults). It is possible that these polarity alternations may occur when ancient dislocation systems cross the rift axis. *See* FAULT AND FAULT STRUCTURES.

Oceanic rift valleys are also distinctly separated into segments by structures known as transform faults. The cross-sectional form of oceanic rift valleys can be markedly asymmetric, but this may not be the norm, and it is unclear if any pattern occurs to the variations along strike. Regardless, it is unlikely that the cross-sectional form of oceanic rift valleys is related genetically to that of continental rift valleys,

except in the broadest possible terms. *See* TRANSFORM FAULT.

Many modern continental rift valleys are marked by spectacular, deep lakes whose floors can be covered with many miles of sediment. Lakes Tanganyika and Malawi (Nyasa) in East Africa and Lake Baikal in Russia are examples. Ancient continental rift valleys are marked usually by long, narrow strips of sedimentary rocks or by volcanic rocks that fill in what were originally basins. The fact that there are many continental rifts argues that the fate of most rift valleys is fossilization prior to successful plate separation. The Mid-Continent Rift in the midwestern United States is an example of a rift that failed more than a billion years ago. The North Sea is an example of a system of rift basins that were subjected to much extension prior to abandonment about 100 million years ago. The North Sea and the Gulf of Suez are examples of failed rifts that were subjected to marine incursion. However, these are not oceanic rift valleys in the true sense because neither reached the stage of complete plate separation (that is, no oceanic lithosphere was generated). The North Sea, Gulf of Suez, and Reconcavo Basin in Brazil demonstrate that rift valley settings can be associated with prolific deposits of hydrocarbons. *See* BASIN; LITHOSPHERE.

Bruce R. Rosendahl

Bibliography. D. W. Burbank and R. S. Anderson, *Tectonic Geomorhology: A Frontier in Earth Science*, 2001; P. M. LeTourneau and P. E. Olsen (eds.), *The Great Rift Valleys of Pangea in Eastern North America, Volume II: Sedimentology, Stratigraphy, and Paleontology*, 2003; B. R. Rosendahl, Architecture of continental rifts with special reference to East Africa, *Annu. Rev. Earth Planet. Sci.*, 15:445–503, 1987; B. R. Rosendahl et al., Structural expressions of rifting: Lessons from Lake Tanganyika, Africa, in *Sedimentation in the African Rifts*, Spec. Publ. GSSP 23, Quart. J. Geol. Soc. Lond., 1986.

Rift Valley fever

An arthropod-borne (primarily mosquito), acute, febrile, viral disease of humans and numerous species of animals, particularly ruminants. Rift Valley fever is caused by a ribonucleic acid (RNA) virus in the genus *Phlebovirus* of the family Bunyaviridae. In sheep and cattle, it is also known as infectious enzootic hepatitis. First described in the Rift Valley of Africa, the disease presently occurs in west, east, and south Africa and has extended as far north as Egypt. Outside of Africa, Rift Valley fever has spread to the island of Madagascar and the Arabian Peninsula (Saudi Arabia and Yemen).

Clinical signs and pathology. In humans, clinical signs of Rift Valley fever are influenzalike, and include fever, headache, muscular pain, weakness, nausea, epigastric pain, and photophobia. Most people recover within 4–7 days, but some individuals may have impaired vision or blindness in one or both eyes; a small percentage of infected individuals develop encephalitis or a hemorrhagic syndrome and die.

The clinical signs of Rift Valley fever in animals depend on the species infected, age, and physiological condition. The most susceptible are neonatal lambs, calves, and puppies, which develop a high fever; become weak, and die 24–48 h after infection. The mortality rate in lambs that are less than a week of age can be over 90%. Mature sheep and cattle develop a high fever; abortion at any stage of gestation is the most prominent sign. Mortality rates in aborting ewes may reach 30%, and 10% in aborting cows.

The prominent lesion in aborted lambs and calves results from the death of many liver cells, causing the organ to become yellow and friable. Lesions in older animals that die from the disease consist of widespread hemorrhages on serous membranes and on intestines, and the localized death of liver cells.

Diagnosis. Rift Valley fever should be suspected when the following observations are made in a disease outbreak: (1) high abortion rates (possibly approaching 100%) in ewes, cows, and female dogs, with lower rates in goats and in other ruminants; (2) high mortality (possibly approaching 100%) in lambs and calves less than 7 days old, and lower rates of disease and mortality in older animals; (3) extensive liver lesions in aborted fetuses and newborn animals; (4) an influenzalike disease in people, particularly in individuals associated with livestock; (5) occurrence of the disease during a period of high insect activity following heavy seasonal rainfall with persistent flooding; (6) rapid spread. The diagnosis is confirmed by isolating the virus from tissues of an infected animal or human. Viral antigen may be detected in the laboratory from tissues or tissue smears using a number of techniques. Antibodies (IgM or IgG) can be detected in serum.

Epidemiology. Historically, outbreaks of Rift Valley fever have occurred at 10–15-year intervals in normally dry areas of Africa subsequent to a period of heavy rainfall. During interepidemic periods the virus is present in the eggs of the mosquito *Aedes lineatopinnis*, which can survive for many years in the dried soil of grassland depressions called dambos. After a period of heavy rain, the dambos fill with water, the mosquito eggs hatch, the larvae mature, and then the infected mosquitoes can feed on a ruminant animal, which becomes infected. A key factor in the epizootic (outbreak of disease among animal populations) of Rift Valley fever is sustained flooding, which leads to the emergence of large numbers of secondary (that is, other species of) mosquitoes. Once the virus proliferates in the blood of the ruminant animal, most species of mosquito feeding on this animal can also become infected and transmit the virus to other animals. Humans can also become infected by an aerosol that can be produced during autopsy; ritual slaughter, as occurred in Saudi Arabia, by contact with an infected animal that has aborted; or by laboratory procedures. Moreover, they can develop a sufficiently high blood infection to be a source of infection for mosquitoes; thus, a sick individual could accidentally transmit the virus to other parts of the world. In Egypt, human

infection has occurred through mosquito bites in an urban peridomestic mosquito cycle.

Prevention and control. Animals should not be moved from an infected area to a Rift Valley–free area if there is any indication of active disease. Control of mosquitoes by spraying insecticides may be of value in reducing human exposure in limited areas. In addition, control of the disease is best accomplished by widespread vaccination of susceptible animals to prevent amplification of the virus and, thus, infection of vectors. Any individual that works with infected animals or live virus in a laboratory should be vaccinated. An inactivated cell-culture virus vaccine is available for use in both animals and humans. This vaccine has the disadvantages of requiring two inoculations over a 30-day period and a large quantity of antigen. A modified live Smithburn strain vaccine is available for use in nonpregnant animals. Other vaccines still under development are based on a mutagen-induced MP12 strain and a Clone 13 strain. *See* ANIMAL VIRUS; VACCINATION. Charles A. Mebus

Bibliography. H. Caplan, C. J. Peters, and D. H. L. Bishop, Mutagen-directed attenuation of Rift Valley fever virus as a method for vaccine development, *J. Gen. Virol.*, 66:2271–2277, 1985; F. G. Davies, K. J. Linthicum, and A. D. James, Rainfall and epizootic Rift Valley fever, *Bull. WHO*, 63:941–943, 1985; G. Geran (ed.), *CRC Handbook Series in Zoonosis*, 1981; A. Jouan et al., Analytical study of a Rift Valley fever epidemic, *Res. Virol.*, 140(2):175–186, 1989; K. J. Linthicum et al., Climate and satellite indicators to forecast Rift Valley fever epidemics in Kenya, *Science*, 285:397–400, 1999; C. J. Peters and J. M. Meegan, Rift Valley fever, in G. W. Beran and J. H. Steele (eds.), *Handbook of Zoonoses*, Sec. B: *Viral Zoonoses*, 2d ed., pp. 403–420, CRC Press, Boca Raton, 1994; W. C. Reeves et al., The potential effect of global warming on mosquito-borne arboviruses, *J. Med. Entomol.*, 31(3):323–332, 1994; A. Shimshony and R. Barzilai, Rift Valley fever, *Adv. Vet. Sci. Comp. Med.*, 27:347–425, 1983; WHO/FAO Working Group on Rift Valley fever, *Rift Valley Fever: An Emerging Human and Animal Problem*, WHO Publ. 63, 1982; H. Zeller et al., Enzootic activity of Rift Valley fever virus in Senegal, *Am. J. Trop. Med. Hyg.*, 56(3):265–272, 1997.

Rigel

The bright star in the southwest corner of the constellation Orion (apparent magnitude +0.12), also referred to as β Orionis. It is a blue-white supergiant of spectral type B8, one of the most luminous stars known. Its intrinsic brightness is estimated to be more than 50,000 times that of the Sun. Rigel is a very young star by stellar standards, with such an enormous rate of energy output that its life span is only a few million years. By comparison, the Sun is approximately 5×10^9 years old, and is still only half way through its main-sequence evolution. Rigel is the brightest member of the Orion OB1 Association, a large molecular cloud complex in which active star formation is currently taking place. In addition to its high luminosity, Rigel has a photosphere which is so large that if the star were placed where the Sun is, it would almost fill the orbit of Mercury. *See* MOLECULAR CLOUD; SPECTRAL TYPE.

The spectrum of Rigel reveals pronounced absorption lines of neutral helium and hydrogen. There is evidence that the star is loosing mass in the form of strong winds emanating from its surface, at a rate of 10^{-6} solar mass per year, or about 6×10^{13} metric tons (7×10^{13} tons) per second. It rotates around its axis with a projected velocity of 42 km/s (26 mi/s). The effective temperature is approximately 12,000 K (21,100°F). Rigel also displays small irregular light variations and erratic changes in its radial velocity that seem to be quite common among evolved stars and may be due to pulsation. *See* ORION; STELLAR ROTATION; STAR; SUPERGIANT STAR. David W. Latham

Rigid body

An idealized extended solid whose size and shape are definitely fixed and remain unaltered when forces are applied. Treatment of the motion of a rigid body in terms of Newton's laws of motion leads to an understanding of certain important aspects of the translational and rotational motion of real bodies without the necessity of considering the complications involved when changes in size and shape occur. Many of the principles used to treat the motion of rigid bodies apply in good approximation to the motion of real elastic solids. *See* NEWTON'S LAWS OF MOTION; RIGID-BODY DYNAMICS. Dudley Williams

Rigid-body dynamics

The study of the motion of a rigid body under the influence of forces. A rigid body is a system of particles whose distances from one another are fixed. The general motion of a rigid body consists of a combination of translations (parallel motion of all particles in the body) and rotations (circular motion of all particles in the body about an axis). Its equations of motion can be derived from the equations of motion of its constituent particles. *See* RECTILINEAR MOTION; ROTATIONAL MOTION.

The location of a mass point m_i can be specified relative to a fixed-coordinate system by a position vector $\vec{r}_i$ with cartesian components (x_i, y_i, z_i). The vector force $\vec{F}_i$ which acts on the mass points has corresponding components (F_{ix}, F_{iy}, F_{iz}). Newton's second law for the motion of m_i is stated in Eq. (1).

$$\vec{F}_i = m_i \ddot{\vec{r}}_i \tag{1}$$

Here $\ddot{\vec{r}}_i \equiv d^2\vec{r}_i/dt^2 = \vec{a}$ is the acceleration of m_i. *See* ACCELERATION; FORCE; NEWTON'S LAWS OF MOTION.

Translational motion. If Eq. (1) is summed over all particles in the rigid body, the left-hand side becomes

the total force

$$\vec{F} = \sum_i \vec{F}_i$$

acting on the rigid body. If the internal forces satisfy Newton's third law (to each action there is an equal but opposite reaction), the contributions of the internal forces cancel in pairs and $\vec{F}$ is the total external force on the rigid body, $\vec{F}^{\text{ext}}$. The right-hand side can be expressed in terms of the center-of-mass (CM) position vector defined by Eq. (2), where M is the total

$$\vec{R} = \frac{1}{M}\sum_i m_i \vec{r}_i \tag{2}$$

$$M = \sum_i m_i$$

mass of the body. [In continuous systems the mass elements m_i are replaced by differential mass elements $dm = \rho(\vec{r})dV$. where $\rho(\vec{r})$ is the mass density and dV is the differential volume element. For a rigid body with a continuous mass density, the center-of-mass vector is calculated by integrating dm over the volume of the body as in Eq. (3).] Then the sum of

$$\vec{R} = \frac{1}{M}\int \vec{r}\rho(\vec{r})\,dV \tag{3}$$

the equations of motion, Eq. (1), takes the form of Eq. (4). This equation of motion for the center of

$$\vec{F}^{\text{ext}} = M\ddot{\vec{R}} \tag{4}$$

mass of the rigid body has exactly the form of the equation of motion for a particle of mass M and position $\vec{R}$, under the influence of an external force $\vec{F}^{\text{ext}}$. Consequently, the second law of motion holds, not just for a particle, but for an arbitrary rigid body, if the position of the body is interpreted to mean the position of its center of mass. *See* CENTER OF MASS.

The momentum of a mass point is given by the product of the mass and the velocity, $\vec{p}_i = m_i\dot{\vec{r}}_i$, where $\dot{\vec{r}}_i \equiv d\vec{r}_i/dt$. The total momentum of the center of mass of the rigid body, obtained by summing over the momenta $\vec{p}_i$ of its constituent masses, is given by Eq. (5). In terms of the center-of-mass momentum

$$\vec{P} = M\dot{\vec{R}} \tag{5}$$

$\vec{P}$, the equation of motion for the center of mass is expressed by Eq. (6). For an isolated rigid body, the

$$\dot{\vec{P}} = \vec{F}^{\text{ext}} \tag{6}$$

external force is zero and therefore $\vec{P}$ is constant. According to Eq. (5), this implies that the center of mass moves with constant velocity $\vec{V} = \vec{p}/M$. *See* CONSERVATION OF MOMENTUM; MOMENTUM.

In fact, the preceding equations for translational motion hold for any body, rigid or nonrigid.

Rotational motion. The total angular momentum of a rigid body about a point O with coordinate $\vec{r}_O$ is the sum of the angular momenta of its constituent masses, and is given by Eq. (7). Here $\times$ denotes the

$$\vec{L}_O = \sum_i (\vec{r}_i - \vec{r}_O) \times m_i(\dot{\vec{r}}_i - \dot{\vec{r}}_O) \tag{7}$$

cross-product of the coordinate vector $(\vec{r}_i - \vec{r}_O)$ with the momentum vector $m_i(\dot{\vec{r}}_i - \dot{\vec{r}}_O)$. The time derivative of $\vec{L}$ is given in Eq. (8). Hereafter the point O

$$\dot{\vec{L}}_O = \sum_i (\vec{r}_i - \vec{r}_O) \times m_i(\ddot{\vec{r}}_i - \ddot{\vec{r}}_O) \tag{8}$$

is taken to be either a fixed point (in which case $\dot{\vec{r}}_O = \ddot{\vec{r}}_O = 0$) or the center-of-mass point. Using the equation of motion (1), $m_i\ddot{\vec{r}}_i$ can be replaced by $\vec{F}_i$. Thus the rotational equation of motion (9) is ob-

$$\dot{\vec{L}}_O = \sum_i (\vec{r}_i - \vec{r}_O) \times \vec{F}_i \tag{9}$$

tained. The right-hand side of Eq. (9) is known as the torque, $\vec{N}$. The contribution of the internal forces to the torque vanishes if the "extended third law" holds; namely, action equals reaction and is directed along a line between the particles. In this circumstance the rotational equation of motion is given by Eqs. (10),

$$\dot{\vec{L}}_0 = \vec{N}_0^{\text{ext}} \tag{10a}$$

$$\vec{N}_0^{\text{ext}} = \sum_i (\vec{r}_i - \vec{r}_O) \times \vec{F}_i^{\text{ext}} \tag{10b}$$

where $\vec{N}^{\text{ext}}$ is the total torque associated with external forces that act on the rigid body. *See* ANGULAR MOMENTUM; TORQUE.

It is straightforward to show from Eq. (7) that the angular momentum about an arbitrary point O is related to the angular momentum about the center of mass by Eq. (11). The torque about an arbitrary point

$$\vec{L}_O = \vec{L}_{\text{CM}} + (\vec{R} - \vec{r}_O) \times \vec{P} \tag{11}$$

O can also be easily related to the torque about the center of mass, by Eq. (12).

$$\vec{N}_O^{\text{ext}} = \vec{N}_{\text{CM}}^{\text{ext}} + (\vec{R} - \vec{r}_O) \times \vec{F}^{\text{ext}} \tag{12}$$

The position of every particle in a rigid body is specified by the position of any one point of the body (such as the center of mass) plus the orientation of the body about that point. Three coordinates are needed to specify one particle in the body. Two angular coordinates can specify the position of a second particle, since it lies at a fixed distance from the first. Only one further coordinate is needed to locate the position of a third particle because its distances from the first and second particles are fixed. The positions of any other particles in the rigid body are completely determined by their distances from the first three particles. Hence six coordinates determine the positions of all particles in a rigid body, and the motion of a rigid body is described by six equations of motion. The translational motion of the center of mass is determined by Eq. (6), and the rotational motion about the center of mass, or a fixed point, is determined by Eq. (10). These six equations, which hold for any system of particles, completely describe the motion of a rigid body.

Motion of an isolated system. The equation $\dot{\vec{L}} = \vec{N}^{\text{ext}}$ has the same form as $\dot{\vec{P}} = \vec{F}^{\text{ext}}$. Both $\vec{L}$ and $\vec{p}$ are constants for an isolated system since $\vec{N}^{\text{ext}} = 0$

and $\vec{F}^{\text{ext}} = 0$. Even though the two conditions $\vec{N}^{\text{ext}} = 0$ and $\vec{F}^{\text{ext}} = 0$ appear similar, there are some important differences for systems in which internal motion is possible. If $\vec{F}^{\text{ext}} = 0$, a center of mass which is at rest will remain so, regardless of internal forces or internal motion. If $\vec{N}^{\text{ext}} = 0$, the total angular momentum is constant, and if initially zero, will remain zero. However, $\vec{L} = 0$ does not exclude changes in orientation of the system by the use of merely internal forces. There is no rotational analog to the equation $\vec{r}(t) = (\vec{p}/M)t + \vec{r}(0)$ for linear motion of the center of mass.

Static equilibrium. In the design of permanent structures, the conditions under which a rigid body remains in steady motion under the action of a set of forces are of great importance. The six conditions for complete equilibrium of a rigid body are given in Eqs. (13). However, in many circumstances equilibrium is desired only for a subset of the six independent directions of motion. To illustrate, the external force in the direction of motion of an accelerating automobile is nonzero, but equilibrium is maintained in all other directions. For rotational equilibrium, the torque about the center of mass of the automobile vanishes. However, by Eq. (12) the torque about another point O may differ from zero. *See* STATICS.

$$\begin{aligned} \vec{F}^{\text{ext}} &= \sum_i \vec{F}_i^{\text{ext}} = 0 \\ \vec{N}_{\text{CM}}^{\text{ext}} &= \sum_i (\vec{r}_i - \vec{R}) \times \vec{F}_i^{\text{ext}} = 0 \end{aligned} \tag{13}$$

Single-axis rotations. Consider first the case of a rigid body that is constrained to rotate about a fixed axis. For definiteness, the z axis is chosen as the rotation axis, $\vec{\omega} = \omega_z \hat{z}$. The rotational velocity of a mass point m_i at a distance $\mathscr{R}_i = (x_i^2 + y_i^2)^{1/2}$ from the z axis is $v_\phi = \mathscr{R}_i \omega_z$, where ϕ is the azimuthal angle (**Fig. 1**). The angular momentum of this mass point is given by Eq. (14). Summing over all masses, the

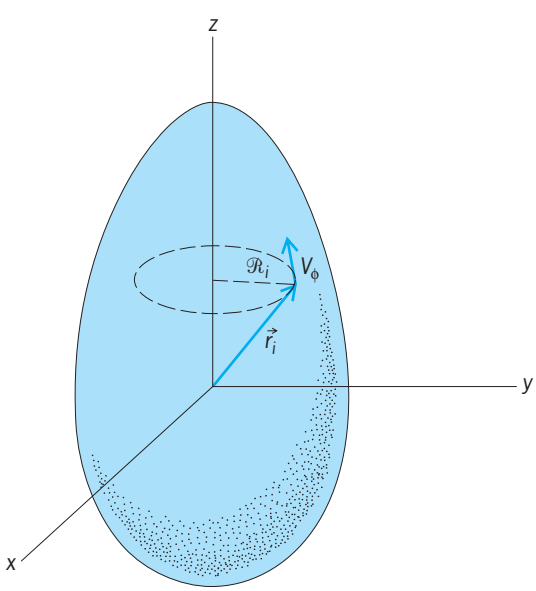

Fig. 1. Rotation of a symmetric rigid body around a fixed axis through the center of mass.

Moments of inertia of some simple bodies

Body	Axis through center of mass	Moment of inertia (I_{CM})
Rod, length l	Perpendicular to rod	$^1/_{12}Ml^2$
Rectangular plate, sides a, b	Parallel to side b	$^1/_{12}Ma^2$
	Perpendicular to plate	$^1/_{12}M(a^2 + b^2)$
Cube, sides a	Perpendicular to face	$^1/_6Ma^2$
Hoop, radius a	Perpendicular to plane	Ma^2
Disk, radius a	Perpendicular to plane	$^1/_2Ma^2$
	Parallel to plane	$^1/_4Ma^2$
Solid cylinder, radius a	Along cylinder axis	$^1/_2Ma^2$
Spherical shell, radius a	Any axis	$^2/_3Ma^2$
Solid sphere, radius a	Any axis	$^2/_5Ma^2$

$$L_z = \mathscr{R}_i p_\phi = \mathscr{R}_i(m_i v_\phi) = m_i \mathscr{R}_i^2 \omega_z \tag{14}$$

angular momentum of the rigid body is given by Eq. (15). The quantity of proportionality between L_z

$$L_z = \left(\sum_i m_i \mathscr{R}_i^2 \right) \omega_z \tag{15}$$

and ω_z is called the moment of inertia, which will be denoted here by I_z. The moment-of-inertia expression for a continuous mass distribution is given in Eq. (16).

$$I_z = \int (x^2 + y^2) \rho(\mathscr{R}) dV \tag{16}$$

The rotational motion about the z axis is accelerated by the external torque according to Eq. (10). The moment of inertia I_z is time-independent, since the perpendicular distance $\mathscr{R}$ from the rotation axis of a given mass point is fixed in the rigid body. Hence, equation of motion (17) is obtained.

$$N_z = I_z \dot{\omega}_z \tag{17}$$

Moments of inertia about an axis through the center of mass of some simple bodies of uniform density are given in the **table**. The moment of inertia I'_z about the parallel axis is related to the moment of inertia I_z^{CM} by the parallel axis rule of Eq. (18), where d is the

$$I'_z = I_z^{\text{CM}} + Md^2 \tag{18}$$

perpendicular distance between the two axes and M is the total mass of the rigid body. N_Z and L_Z have different values about the parallel axis than about the center of mass, as given by Eqs. (11) and (12).

Since Eq. (17) has the same mathematical structure as the equation for linear motion in one direction, $F_z = M\dot{v}_z$, the techniques for solving one-dimensional linear motion problems can also be applied to solve Eq. (17) for rotations about a fixed axis.

Energy conservation. If the external force on a rigid body is conservative, energy conservation methods

may be used in determining the motion. The conserved mechanical energy is $E = T + V$, where the kinetic energy T and the potential energy V are given in Eqs. (19), where $\vec{r}_s$ is a reference point which

$$T = \frac{1}{2}\int v^2 \rho(\vec{r}')\, dV' \qquad (19a)$$

$$V(\vec{r}) = -\int_{\vec{r}_s}^{\vec{r}} \vec{F}(\vec{r}') \cdot d\vec{r}' \qquad (19b)$$

specifies the zero of the potential. The velocity of a mass point in the rigid body is the vector sum of the center-of-mass velocity and the velocity of rotation about the center of mass, Eq. (20). By using this fact, the kinetic energy can be expressed by Eq. (21)

$$\vec{v}_i = \vec{V} + \mathscr{R}_i \omega_z \hat{\phi} \qquad (20)$$

$$T = {}^1\!/_2 MV^2 + {}^1\!/_2 I_z^{\mathrm{CM}} \omega_z^2 \qquad (21)$$

as the sum of the translational kinetic energy and the rotational kinetic energy about an axis through the center of mass. The gravitational potential energy of a body near the surface of the Earth is $V(\vec{r}) = Mgh$, where h is the vertical distance of the body's center of mass from the Earth's surface and g is the gravitational acceleration. *See* CONSERVATION OF ENERGY; ENERGY.

Impulsive motion. For forces such as a sharp blow that act only during a very short time, it is convenient to use an integrated form of the laws of motion. The translational motion of the center-of-mass point is determined by Eq. (6). When both sides of this equation are integrated with respect to t over the time interval $\Delta t = t_1 - t_0$, during which the force acts, Eq. (22) is obtained. The time integral of the force

$$\Delta\vec{P} = \vec{P}^1 - \vec{P}^0 = \int_{t_0}^{t_1} \vec{F}^{\mathrm{ext}}\, dt \qquad (22)$$

on the right is called the impulse. For angular motion the integrated form of the equation of motion, Eq. (23), results. The time integral of the torque is

$$\Delta\vec{L} = \vec{L}^1 - \vec{L}^0 = \int_{t_0}^{t_1} N^{\mathrm{ext}}\, dt \qquad (23)$$

called the angular impulse. For rotations of a rigid body about a symmetry axis which passes through the center of mass, the angular-impulse equation can be expressed as in Eq. (24), where z is the symmetry axis.

$$\Delta L_z = I_z^{\mathrm{CM}}(\omega_z^1 - \omega_z^0) = \int_{t_0}^{t_1} N_z^{\mathrm{ext}}\, dt \qquad (24)$$

See IMPACT; IMPULSE (MECHANICS).

The dynamics of billiard shots is an example of the usefulness of the impulse formulation of the equations of motion. For simplicity, consider shots in which the cue hits the ball in its vertical median plane in a horizontal direction. The cue imparts an impulse to the stationary ball at a vertical distance h above the table (**Fig. 2**). The linear impulse is given in Eq. (25),

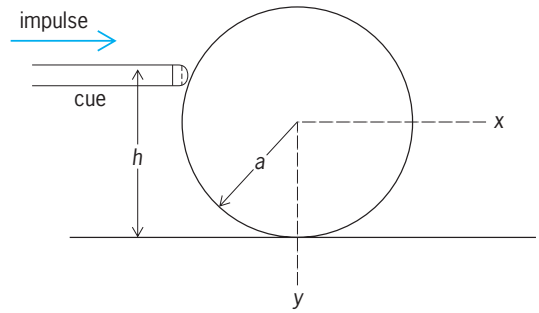

Fig. 2. Impulse imparted to a billiard ball by a cue stick. (*After V. Barger and M. Olsson, Classical Mechanics: A Modern Perspective, McGraw-Hill, 1973*)

$$M\Delta V_x = MV_x^1 = \int_{t_0}^{t_1} F_x\, dt \qquad (25)$$

where V_x^1 is the velocity of the center of mass just after impact. The angular impulse about the z axis which passes through the center of mass of the ball is given in Eq. (26), where a is the radius of the

$$\Delta L_z = I_z^{\mathrm{CM}}\omega_z^1 = \int_{t_0}^{t_1} (b - a)F_x\, dt \qquad (26)$$

ball. By elimination of the force integral and substitution of the sphere moment of inertia, the relation of Eq. (27) is obtained between the spin and velocity of

$$\omega_z^1 = \frac{5}{2}\left(\frac{b-a}{a^2}\right) V_x^1 \qquad (27)$$

the ball immediately after the impulse. The velocity of the ball at the point of contact with the table is given by Eq. (28). If the ball is to roll without slipping

$$V_c = V_x^1 - a\omega_z^1 = V_x^1\left(\frac{7a - 5b}{2a}\right) \qquad (28)$$

($V_c = 0$), the height at which the ball must be hit is given by Eq. (29). Only if this condition is satisfied

$$b = \frac{7}{5}a \qquad (29)$$

does pure rolling take place from the very start.

General rotations. A rotation is a motion such that a point i located at $\vec{r}_i$ has a constant distance $|\vec{r}_i - \vec{r}_O|$ to each point O on a line, the axis of rotation (**Fig. 3**). The path traced out by the point i is a circle whose symmetry axis is the axis of rotation. If the z axis is chosen to be parallel to the rotation axis, the circular path lies in the x, y plane and the vector $\vec{r} - \vec{r}_O$ is given by Eq. (30). Here a is the radius of the circle,

$$\vec{r} - \vec{r}_O = a\cos\phi\hat{x} + a\sin\phi\hat{y} + b\hat{z} \qquad (30)$$

ϕ is the rotation angle, and b is the distance from O to the center of the circle. The time derivative of Eq. (30) gives the velocity of point i in Eq. (31), which

$$\vec{v} = \dot{\phi}(-a\sin\phi\hat{x} + a\cos\phi\hat{y}) \qquad (31)$$

is perpendicular to $\vec{r} - \vec{r}_O$. Defining an angular velocity vector $\dot{\omega} \equiv \dot{\phi}\,\hat{z}$, the velocity of rotation can be written as in Eq. (32). Equation (32) is generally valid

$$\vec{v} = \vec{\omega} \times (\vec{r} - \vec{r}_O) \tag{32}$$

for any coordinate system if $\vec{\omega}$ is defined as $\vec{\omega} \equiv \omega\hat{n}$ with the unit vector $\hat{n}$ directed along the rotation axis. If the point O is moving with a translational velocity $\vec{v}_O$ relative to a fixed reference frame, the velocity of the point i in the fixed frame is the vector sum of the translational velocity $\vec{v}_O$ and $\vec{\omega}$ may change with time. The expression, Eq. (33), is the

$$\vec{v}(r) = \vec{v}_O + \vec{\omega} \times (\vec{r} - \vec{r}_O) \tag{33}$$

most general for the velocity of the points of a rigid body. The vectors $\vec{v}_O$ and $\vec{\omega}$ amount to six velocity components, agreeing with the number of coordinates of a rigid body.

As a single-axis example, consider a wheel of radius R which rolls without slipping on a level surface (**Fig. 4**). For no slipping, $dx = Rd\theta$, where dx and $d\theta$ are infinitesimal horizontal and angular displacements, respectively. Dividing both sides by the time interval dt, $\vec{v}_O = R\omega\hat{x}$, where $\vec{v}_O$ is the velocity of the center of mass. The velocity of a point on the wheel

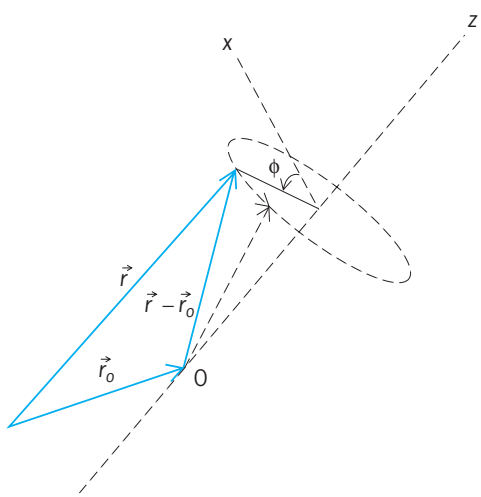

Fig. 3. Rotation of a point located at $\vec{r}$ about the z axis.

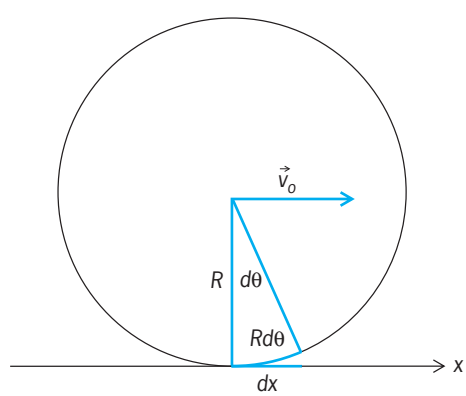

Fig. 4. Wheel rolling without slipping on a level surface. (*After V. Barger and M. Olsson, Classical Mechanics: A Modern Perspective, McGraw-Hill, 1973*)

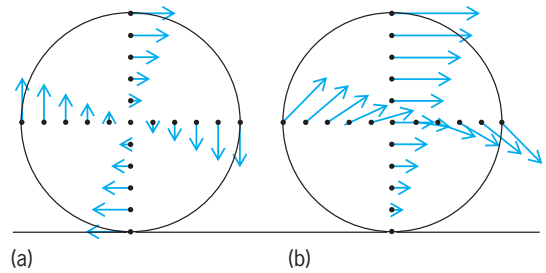

Fig. 5. Velocity of points on a wheel (*a*) relative to the center of mass and (*b*) relative to the ground. (*After V. Barger and M. Olsson, Classical Mechanics: A Modern Perspective, McGraw-Hill, 1973*)

relative to the center of mass is given in Eq. (34),

$$\vec{v}_{\text{rot}} = \vec{\omega} \times \vec{r} = v_O\hat{\omega} \times \left(\frac{\vec{r}}{R}\right) \tag{34}$$

where $\hat{\omega} = \vec{\omega}/\omega$. The rotational velocities of points on the wheel are illustrated in **Fig. 5**. With respect to a reference frame at rest on the level surface, the velocity of a point on the wheel is given by Eq. (35).

$$\vec{v} = \vec{v}_O + \vec{v}_{\text{rot}} = \vec{v}_O + \vec{v}_O\hat{\omega} \times \left(\frac{\vec{r}}{R}\right) \tag{35}$$

The velocity $\vec{v} = 0$ at the point of contact, since no slipping means that there is no relative motion of the wheel and surface at the contact point.

Moments and products of inertia. The dynamics of rigid-body rotations are contained in Eq. (10), which relates the time rate of change of the total angular momentum to the external torque. The angular momentum $\vec{L}$ about a point O can be computed in terms of the angular velocity $\vec{\omega}$ in Eq. (21). The location of a point mass m_i in the rigid body relative to O is hereafter denoted by $\vec{r}_i$. The velocity of the mass m_i relative to O is then $\vec{v}_i = \vec{\omega} \times \vec{r}_i$, and the angular momentum about O is expressed by Eq. (36). The triple

$$\vec{L} = \sum_i m_i(\vec{r}_i \times \vec{v}_i) = \sum_i m_i\vec{r}_i \times (\vec{\omega} \times \vec{r}_i) \tag{36}$$

cross-product can be expanded by using the identity $\vec{a} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{a} \cdot \vec{C}) - \vec{C}(\vec{a} \cdot \vec{B})$, leading to the result in Eq. (37). It can be seen that the angular

$$\vec{L} = \vec{\omega}\left(\sum_i m_i\,|\vec{r}_i|^2\right) - \sum_i m_i\vec{r}_i(\vec{r}_i \cdot \vec{\omega}) \tag{37}$$

momentum vector $\vec{L}$ will not necessarily be parallel to the angular velocity vector $\vec{\omega}$.

In cartesian components of $\vec{\omega} = (\omega_x,\omega_y,\omega_z)$ and $\vec{r} = (x,y,z)$, the components of $\vec{L} = (L_x,L_y,L_z)$ have the form of Eqs. (38). The coefficients of the angu-

$$\begin{aligned} L_x &= I_{xx}\omega_x + I_{xy}\omega_y + I_{xz}\omega_z \\ L_y &= I_{yx}\omega_x + I_{yy}\omega_y + I_{yz}\omega_z \\ L_z &= I_{zx}\omega_x + I_{zy}\omega_y + I_{zz}\omega_z \end{aligned} \tag{38}$$

lar velocity components are written in Eqs. (39). The

$$\begin{aligned} I_x &\equiv I_{xx} = \sum_i m_i(y_i^2 + z_i^2) \\ I_y &\equiv I_{yy} = \sum_i m_i(x_i^2 + z_i^2) \\ I_z &\equiv I_{zz} = \sum_i m_i(x_i^2 + y_i^2) \\ I_{xy} &= I_{yx} = -\sum_i m_i x_i y_i \\ I_{xz} &= I_{zx} = -\sum_i m_i x_i z_i \\ I_{yz} &= I_{zy} = -\sum_i m_i y_i z_i \end{aligned} \tag{39}$$

three quantities I_x, I_y, and I_z are known as moments of inertia and the other six quantities are called products of inertia. For a continuous mass distribution

$$\sum_i m_i$$

is replaced by

$$\int dm$$

where $dm = \rho dV$ as in Eq. (3). The moments and products of inertia will be denoted henceforth by I_{jk}, where the subscripts j, k take on the values x, y, z. The products of inertia are symmetric: $I_{jk} = I_{kj}$. *See* MOMENT OF INERTIA; PRODUCT OF INERTIA.

The nine quantities I_{jk} form a symmetric tensor which can be written as the 3×3 matrix in Eq. (40).

$$\mathbf{I} \equiv \begin{pmatrix} I_{xx} & I_{xy} & I_{xz} \\ I_{yx} & I_{yy} & I_{yz} \\ I_{zx} & I_{zy} & I_{zz} \end{pmatrix} \tag{40}$$

In matrix notation, $\vec{L}$, $\vec{\omega}$, and $\vec{N}$ are written as column vectors, with their x, y, and z components forming the elements, as in Eq. (41). Then the an-

$$\vec{L} = \begin{pmatrix} L_x \\ L_y \\ L_z \end{pmatrix} \quad \vec{\omega} = \begin{pmatrix} \omega_x \\ \omega_y \\ \omega_z \end{pmatrix} \quad \vec{N} = \begin{pmatrix} N_x \\ N_y \\ N_z \end{pmatrix} \tag{41}$$

gular momentum and the equation for rigid-body rotational motion can be written in compact form as in Eqs. (42).

$$\vec{L} = \mathbf{I} \cdot \vec{\omega} \tag{42a}$$

$$\vec{N} = \frac{d\vec{L}}{dt} = \frac{d}{dt}(\mathbf{I} \cdot \vec{\omega}) \tag{42b}$$

See MATRIX THEORY.

The kinetic energy of a rigid body can likewise be expressed in terms of the moments and products of inertia. The steps leading to this form for T are given in Eq. (43). In tensor notation the kinetic energy is then given by Eq. (44).

$$\begin{aligned} T &= \frac{1}{2}\sum_i m_i \vec{v}_i \cdot \vec{v}_i \\ &= \frac{1}{2}\sum_i m_i(\vec{\omega} \times \vec{r}_i) \cdot (\vec{\omega} \times \vec{r}_i) \\ &= \frac{1}{2}\vec{\omega} \cdot \left[\sum_i m_i \vec{r}_i \times (\vec{\omega} \times \vec{r}_i)\right] = \frac{1}{2}\vec{\omega} \cdot \vec{L} \end{aligned} \tag{43}$$

$$T = \frac{1}{2}\vec{\omega} \cdot \mathbf{I} \cdot \vec{\omega} = \frac{1}{2}\sum_{j,k} \omega_j I_{jk} \omega_k \tag{44}$$

Euler's equations. The moments and products of inertia I_{jk} relative to the fixed coordinate system change as a function of time as the rigid body rotates, making the description of the motion directly from Eq. (40) difficult. Great simplification can often be achieved by choosing instead a coordinate system that rotates with the rigid body. In the rotating reference frame the moments and products of inertia are time-independent. The equation of motion with respect to the moving "body" axes is given by Eq. (45).

$$\vec{N} = \frac{\delta \vec{L}}{\delta t} + \vec{\omega} \times \vec{L} \tag{45}$$

The time derivative $d\vec{L}/dt$ in a fixed reference frame consists of a part $\delta\vec{L}/\delta t$ from the time rate of change of $\vec{L}$ relative to the axes of the moving frame, and part $\vec{\omega} \times \vec{L}$ from the rotation of these axes relative to the fixed axes.

It is always possible to make a choice of axes in the body so that all the products of inertia vanish (that is, $I_{jk} = 0$ for $j \neq k$), thus making the moment of inertia matrix diagonal. These axes are called the principal axes of the rigid body. The angular momentum components along these axes (which are labeled 1, 2, 3) are given in terms of the angular velocity components along these axes and the principal moments of inertia I_1, I_2, I_3 by Eqs. (46). The resulting Euler's

$$\begin{aligned} L_1 &= I_1\omega_1 \\ L_2 &= I_2\omega_2 \\ L_3 &= I_3\omega_3 \end{aligned} \tag{46}$$

equations of motion for a rigid body in terms of principal-axes components of $\vec{N}$ and $\vec{\omega}$ are given in Eqs. (47). These equations are a convenient starting

$$\begin{aligned} N_1 &= I_1\dot{\omega}_1 + (I_3 - I_2)\omega_3\omega_2 \\ N_2 &= I_2\dot{\omega}_2 + (I_1 - I_3)\omega_1\omega_3 \\ N_3 &= I_3\dot{\omega}_3 + (I_2 - I_1)\omega_2\omega_1 \end{aligned} \tag{47}$$

point for many discussions of rigid-body rotations. *See* EULER'S EQUATIONS OF MOTION.

Poinsot's method. The motion of a freely rotating rigid body can be described by a geometrical approach known as Poinsot's construction, which is based on conservation of kinetic energy and angular momentum. In the coordinate system defined by the principal axes, the kinetic energy T is given by

Eq. (48). A vector $\vec{\rho}$ is defined as in Eq. (49). Rewriting Eq. (48) in terms of the components of $\vec{\rho}$, Eq. (50) is obtained. In the space specified by the

$$2T = I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2 \tag{48}$$

$$\vec{\rho} \equiv \frac{\vec{\omega}}{\sqrt{2T}} \tag{49}$$

$$I_1\rho_1^2 + I_2\rho_2^2 + I_3\rho_3^2 = 1 \tag{50}$$

coordinates (ρ_1,ρ_2,ρ_3), this describes the surface of an ellipsoid, known as the inertial ellipsoid. The inertial ellipsoid is imagined to be rigidly attached to the body and rotating with it. Taking the gradient of the left-hand side of Eq. (50) with respect to $\vec{\rho}$ gives the direction of the normal to the ellipsoid surface. According to Eq. (51), the normal direction is

$$\nabla_{\vec{\rho}}(I_1\rho_1^2 + I_2\rho_2^2 + I_3\rho_3^2) = 2\frac{\vec{L}}{\sqrt{2T}} \tag{51}$$

along the constant vector $\vec{L}$. Thus the tangent plane to the ellipsoid at $\vec{\rho}$ gives the direction of the normal to the ellipsoid surface. Furthermore, the dot product $\vec{\rho}\cdot\vec{L}/L$ is also constant, as demonstrated in Eq. (52).

$$\frac{\vec{\rho}\cdot\vec{L}}{L} = \frac{2T}{L} \tag{52}$$

Since $\vec{\rho}\cdot\vec{L}/L$ is the distance from the center of the ellipsoid to the tangent plane, the center of the ellipsoid remains at a constant height above the plane. Consequently the force-free motion of the rigid body is represented by the rolling motion of the inertial ellipsoid on the tangent plane, with the center of mass at a fixed height above the tangent plane (**Fig. 6**). The point of contact with the plane is specified by $\vec{\rho}$, which is along the direction of the rotation axis. Thus the point of contact is at rest and the ellipsoid rolls without slipping. The tangent plane is known as the invariable plane, the curve traced out by $\vec{\rho}$ on the inertial ellipsoid is called the polhode, and the corresponding curve on the invariable plane is called the herpolhode.

Tennis racket theorem. The solution of Euler's equation's for a free rigid body with unequal principal moments of inertia can be simply illustrated with a tennis racket. A tennis racket has three principal axes (**Fig. 7**); the moments of inertia around these three principal axes are ordered as in Eq. (53). In a

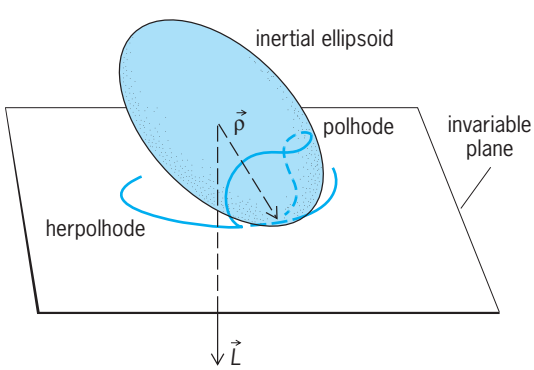

Fig. 6. Rolling motion of inertial ellipsoid on the tangent plane describing the force-free motion of a rigid body.

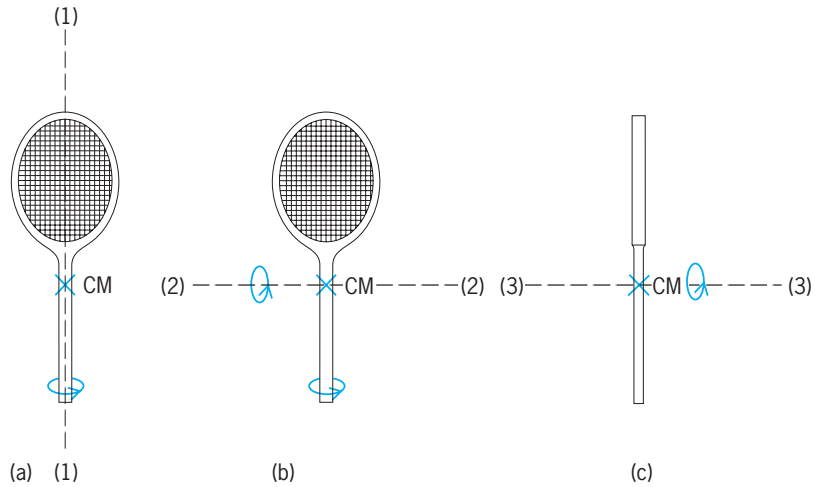

Fig. 7. Principal axes of a tennis racket (*a*) along the handle, (*b*) perpendicular to the handle in the plane of the strings, and (*c*) perpendicular to the handle and strings. (*After V. Barger and M. Olsson, Classical Mechanics: A Modern Perspective, McGraw-Hill, 1973*)

$$I_1 < I_2 < I_3 \tag{53}$$

coordinate system with the racket in the x, y plane and x and y axes along the (1) and (2) principal axes, the principal moments of inertia are given by Eqs. (54), where σ is the mass per unit area. Thus the

$$\begin{aligned} I_1 &= \int y^2\sigma\, dx\, dy \\ I_2 &= \int x^2\sigma\, dx\, dy \\ I_3 &= \int (x^2+y^2)\sigma\, dx\, dy \end{aligned} \tag{54}$$

three moments of inertia for this plane body satisfy the relation of Eq. (55).

$$I_3 = I_1 + I_2 \tag{55}$$

Suppose that a racket is tossed into the air with a spin about one of the principal axes. In the approximation that the Earth's gravity is a uniform force, there are no gravitational torques about the center of mass of the racket, and Euler's equations with $\vec{N} = 0$ apply. They can be written as in Eqs. (56),

$$\begin{aligned} \dot{\omega}_1 + \omega_3\omega_2 &= 0 \\ \dot{\omega}_2 - \omega_1\omega_3 &= 0 \\ \dot{\omega}_3 + K\omega_2\omega_1 &= 0 \end{aligned} \tag{56}$$

where $K = (I_2 - I_1)/(I_2 + I_1)$; for a tennis racket $I_2 > I_1$ and thus K is positive.

If the initial spin is nearly around the axis (1), the product $\omega_3\omega_2$ is small and ω_1 is nearly constant. Then by time differentiation and substitution, equations for ω_2 and ω_3 separately can be found; they are of simple harmonic form, Eqs. (57). These equations

have the solution given in Eqs. (58). By the initial

$$\begin{aligned} \ddot{\omega}_2 + K\omega_1^2\omega_2 &= 0 \\ \ddot{\omega}_3 + K\omega_1^2\omega_3 &= 0 \end{aligned} \tag{57}$$

$$\begin{aligned} \omega_2(t) &= a \sin(\sqrt{K}\omega_1 t + \alpha) \\ \omega_3(t) &= a \cos(\sqrt{K}\omega_1 t + \alpha) \end{aligned} \tag{58}$$

conditions the amplitude a is small and the ω_2 and ω_3 components of the angular velocity remain small. Thus when the tennis racket is tossed into the air with a spin nearly about the (1) axis, the racket continues to spin nearly about this axis and can be easily recaught. A similar stability of the motion applies for axis (3).

For initial motion nearly about axis (2), the second-order differential equations for ω_1 and ω_3 are given in Eqs. (59). Here the solutions for $\omega_1(t)$ and $\omega_3(t)$ grow

$$\begin{aligned} \ddot{\omega}_1 - K\omega_2^2\omega_1 &= 0 \\ \ddot{\omega}_3 - K\omega_2^2\omega_3 &= 0 \end{aligned} \tag{59}$$

exponentially with time, and the motion quickly develops spin around all three principal axes, which makes it difficult to recatch the racket. The instability of rotations about the principal axis with the intermediate moment of inertia (I_2) and the stability of rotations about the other two principal axes is true for any free rigid body. *See* DIFFERENTIAL EQUATION; HARMONIC MOTION. Vernon D. Barger

Bibliography. V. Barger and M. Olsson, *Classical Mechanics: A Modern Perspective*, 2d ed., McGraw-Hill, 1995; H. Goldstein, C. P. Poole, and J. L. Safko, *Classical Mechanics*, 3d ed., Addison-Wesley, 2002; D. Halliday, R. Resnick, and K. S. Krane, *Physics*, 5th ed., Wiley, 2002; K. R. Symon, *Mechanics*, 3d ed., Addison-Wesley, 1971.

Rinderpest

An acute or subacute, contagious viral disease of ruminants and swine, manifested by high fever, lachrymal discharge, profuse diarrhea, erosion of the epithelium of the mouth and of the digestive tract, and high mortality.

Infectious agent. Rinderpest (also known as cattle plague) is caused by a ribonucleic acid (RNA) virus classified in the genus *Morbillivirus* within the Paramyxoviridae family. This virus is closely associated with the viruses of human measles, peste des petits ruminants of sheep and goats, canine distemper, phocine distemper, and the dolphin morbillivirus. Although only one serotype of rinderpest virus is known, the various field strains, grouped by molecular characterization into three lineages known as the Asian lineage and African lineages 1 and 2, vary significantly in virulence. The rinderpest virus is easily inactivated by heat and survives outside the host for a short time. Therefore, transmission of rinderpest is by direct contact between animals. *See* ANIMAL VIRUS; MEASLES; PARAMYXOVIRUS.

Pathogenesis. Rinderpest is one of the oldest diseases known to affect cattle. It is characterized by the development of high fever that lasts for several days until just prior to death. Animals become depressed and have profuse clear serous nasal and lachrymal secretions. The mucosal membranes of the eyes and mouth become congested, and as the disease progresses, there is development of severe erosions on the inside of lips, gums, tongue, and hard and soft palate areas. Cattle suffering from rinderpest develop a very severe, profuse diarrhea and a drop in the number of white blood cells. The morbidity in susceptible cattle and buffalo is greater than 90%, with death of almost all clinically affected animals. *See* DIARRHEA.

The pathological characteristics of rinderpest are related to the ability of the virus to grow in lymphoid and epithelial tissues. In addition to the oral erosions, there is a destruction of epithelium over aggregated lymphoid follicles (Peyer's patches) in the submucosa of the small intestine, and mucosal areas (rich in lymphoid cells) in the large intestine and colon. *See* LYMPHATIC SYSTEM.

Epidemiology. Rinderpest affects virtually all cloven-hoofed animals, both domesticated and wild. Cattle, Asian buffaloes, yaks, African buffaloes, lesser kudus, elands, warthogs, and giraffes are particularly susceptible. Wildlife does not appear to act as a long-term reservoir of infection independently of cattle. All swine are susceptible and may suffer severe disease; they have been important in disease transmission in the Far East. Sheep and goats undergo infection but rarely exhibit clinical signs. Earlier confusing reports of rinderpest in Indian sheep are now mainly considered to have referred to the related virus disease "peste des petits ruminants." Strains of virus causing only mild clinical reactions have been reported, as have others causing peracute (severe and brief) disease. The clinical signs of disease appear in as little as 3 days after exposure to strains causing acute, severe disease, however, the incubation period can extend to 15 days for milder strains. In most fatal cases the animal collapses and dies within 6 to 12 days after the onset of fever; otherwise a protracted convalescence ensues. If the animal survives, it will develop a life-long sterile, solid immunity; there is no carrier state.

Immunity and prevention. Modified (attenuated), live rinderpest vaccines generate a strong, protective immune response that lasts for life. One such vaccine developed in East Africa has been widely used in Africa and Asia and has proved safe and efficacious in large-scale vaccination campaigns and outbreak control. Another vaccine used for decades in Russia might have reverted to virulence on several occasions. A vaccinia recombinant vaccine has been described but has not been licensed for general use. There is now growing confidence that rinderpest has been eliminated from Asia. The last reservoirs of infection persisted in Yemen until 1997 and in the Indus River buffalo tract of Pakistan until late 2000. Thus, the Asian lineage of rinderpest virus is now almost certainly extinct.

Similarly African lineage 1 persisted in southern Sudan until 2001 but was eliminated by vaccination of all cattle belonging to the last tribe to be affected.

The rest of Africa is also now considered to be free from rinderpest except, possibly, for a relatively small area of the Somali pastoral ecosystem spanning the border of Somalia and Kenya, where a reservoir of African lineage 2 could still be present in mild form. However, even there the last confirmation of virus presence was in 2001 in wild African buffaloes in a Kenyan national park. Investigations continue to determine whether or not rinderpest is still present.

In the rest of the world, progress is being made in accreditation of rinderpest freedom by the Office International des Epizooties. The Global Rinderpest Eradication Programme hosted by the United Nations Food and Agriculture Organization has provided international coordination since 1993. *See* VACCINATION.

Alfonso Torres; Peter Roeder

Bibliography. E. P. J. Gibbs (ed.), *Virus Diseases of Food Animals*, vols. 1 and 2, 1981; W. Plowright, Rinderpest, *Ann. N.Y. Acad. Sci.*, 101:327–382, 1962; P. L. Roeder and W. P. Taylor, Rinderpest in Veterinary Clinics of North America, Food Animal Practice, *Animal Health Emergency Diseases*, 18:515–547, 2002; P. B. Rossiter, Rinderpest in J. A. W. Coetzer et al. (eds.), *Infectious Diseases of Livestock with Special Reference to Southern Africa*, Chap. 74, Oxford University Press, pp. 735–757, 1994; G. R. Scott Rinderpest, *Adv. Vet. Sci.*, 9:113–124, 1964.

Ring

A tie member or chain link. Tension or compression applied through the center of a ring produces bending moment, shear, and normal force on radial sections. Because shear stress is zero at the boundaries of the section where bending stress is maximum, it is usually neglected. A quadrant of the ring is a curved bar with moments M_A and M_B at the ends (see **illus.**).

An approximate solution for moments at the ends of the quadrant neglects curvature and employs the procedures used for statically indeterminate straight beams. The numerically larger moment occurs at the load section A where $M_A = 0.318PR$ and $M_B = 0.181PR$, where P is load and R is radius. The variation of moment is found by statics in terms of angle θ defining the section. For tensile load, $M_x = PR(0.318 - {}^1/_2 \sin\theta)$. Moment at A tends to increase the curvature; the ring becomes flatter at B. The end moments have opposite signs and $M = 0$ near $\theta = 39^\circ$.

Important stresses occur at the inside and outside of sections where the normal and bending stresses reach maximum values. At any section, the relationship is as shown in Eq. (1), where S_b is found by

$$S = S_b + \frac{P\sin\theta}{2A} \qquad (1)$$

curved bar theory and A is sectional area. Or stresses can be calculated from the corrected straight beam formula, shown in Eq. (2), where K depends on end

$$S = K\frac{Mc}{I} + \frac{P\sin\theta}{2A} \qquad (2)$$

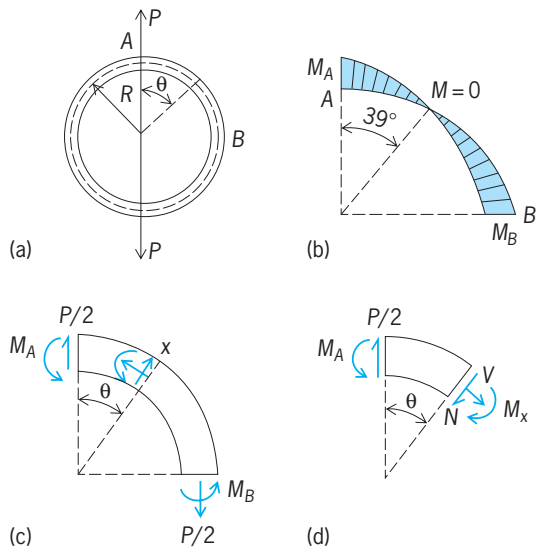

Stresses in rings. (*a*) Entire ring, showing relation of quadrant *AB* and angle θ to load *P* and radius *R*. (*b*) Stresses on quadrant *AB*. (*c*) Variation of moment with angle θ. (*d*) Stresses on arbitrary section.

restraints, I is moment of inertia, and c is distance from centroids to extreme element. The greatest stress (compressive for tensile load) occurs at the inside of the section where load is applied. Where the greatest tensile stress occurs depends on the ring dimensions. For a thick ring, whose radius of curvature is small compared to the dimensions, maximum tension may occur at top and bottom sections, whereas for thin rings with relatively large radius of curvature, greatest tension is at the side sections perpendicular to the load line. Thin rings subjected to external hydrostatic pressure may fail by compressive buckling. Deflection of the ring is most conveniently found by energy methods, which include the moments, shearing, and normal forces acting in the ring. Bending has greater influence in thin rings.

W. J. Krefeld; W. G. Browman

Ring-opening polymerization

The formation of macromolecules from cyclic monomers such as cyclic hydrocarbons, ethers, esters, amides, siloxanes, and sulfur (eight-membered ring). Thus, ring-opening polymerization is of particular interest, since macromolecules of almost any chemical structure can be prepared.

Polymerization is initiated with the breaking of a single (sigma) bond in the cyclic monomer. The driving force of ring-opening polymerization comes from the strain of the rings. The major source of the ring strain is the angular strain, with the release of the strain being energetically favorable.

As the ring size increases, ring strain decreases, although this change is not absolute. Some cyclic monomers, such as six-membered cyclic ethers, are virtually strain-free, and their polymerization to a high-molecular-weight polymer is not possible. By contrast, six-membered cyclic esters and siloxanes

(for example, hexamethylcyclotrisiloxane) polymerize easily, due to their particular conformation. Conformational strain (related to the opposition of hydrogen atoms across the rings) is responsible for the strain for some of the larger cyclic monomers.

The majority of heterocyclic monomers polymerize by ionic mechanisms. That is, either anions or cations located at the end of a growing macromolecule attack the monomer molecule, breaking a bond between the heteroatom (for example, nitrogen or oxygen) and the adjacent atom so that both bonding electrons remain with one of the atoms (heterolytic cleavage). *See* HETEROCYCLIC COMPOUNDS; POLYMER; POLYMERIZATION.

Anionic polymerization. In the anionic polymerization of cyclic ethers or esters, mostly alkoxide anions are involved. Examples of monomers and propagating species are shown in reactions (1)–(3), where

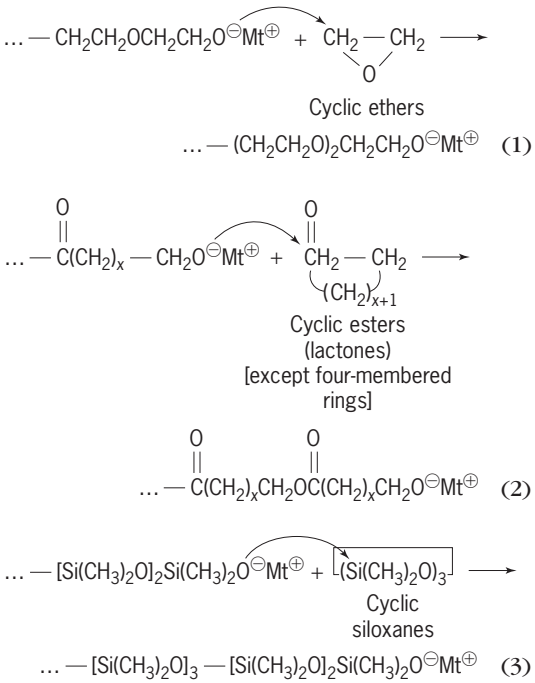

Mt^+ denotes metal cations, mostly Na^+ or K^+, and less often divalent cations.

The polymerization of some of these monomers (for example, ethylene oxide or some lactones) are called "living polymerizations" because there is no termination to stop chain growth.

Polymers of ethylene oxide find numerous applications. Very high molecular weight polymer, up to several million, is used to increase the fluidity of water. Water containing a fraction of a percent of poly(ethylene oxide) can be pumped (for example, by firefighters) with much less energy and therefore for much longer (higher) distances. Lower-molecular-weight products are used as water-soluble surface-active agents. In addition, poly(ethylene oxide) is biocompatible, and it is finding applications in biomedicine, such as drug delivery systems and gels (highly swelling in water) used for healing wounds. *See* POLYETHER RESINS; POLYETHYLENE GLYCOL.

Polysiloxanes are very versatile polymers used in engineering applications due to their outstanding heat resistance, as well as in biomedicine because of their biocompatibility. Silicone rubber (sealant) is another popular product. It can be made as a liquid which solidifies on contact with air. *See* SILICONE RESINS.

Another technically important family of polymers made by anionic ring-opening polymerization is based on ε-caprolactam. Polyamides derived from this monomer are used as fibers for textiles known as nylon (for example, Nylon 6). *See* MANUFACTURED FIBER; POLYAMIDE RESINS.

Coordinated polymerization. The most recently developed living polymerizations of cyclic ethers and cyclic esters are based on multicenter concerted mechanisms of propagation. This mechanism of coordinated polymerization was elaborated in Lodz, Poland, and is described in reaction (4) for the poly-

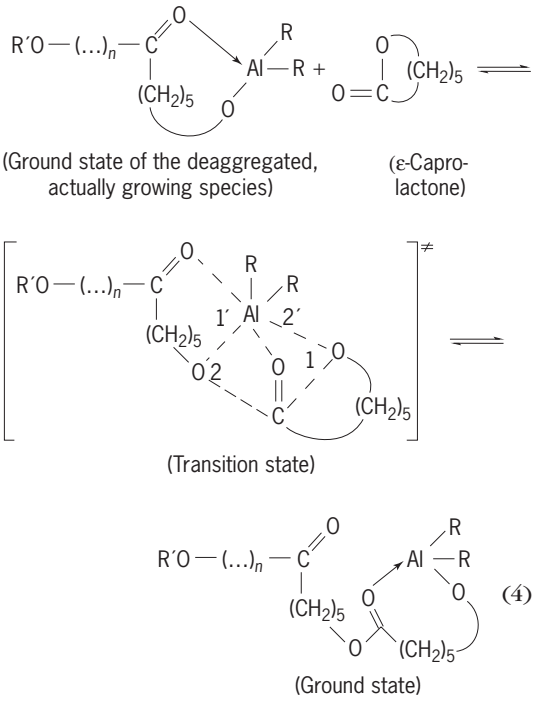

merization of ε-caprolactone initiated with dialkylalkoxyaluminum (R and R′ = C_2H_5), where in a concerted manner two bonds (1 and 1′) are broken and simultaneously two new bonds (2 and 2′) are formed. The coordinated propagating species are less reactive than ions and therefore are much more selective. Alkoxides of iron (Fe), tin (Sn), titanium (Ti), and several others polymerize cyclic esters in the same way. Only with these initiators is it possible to avoid formation of the unwanted cyclic oligomers, which notoriously contaminate the desired linear macromolecules.

The most comprehensively studied cyclic ester is L,L-dilactide, giving the novel industrial polymer poly(L-lactide) [reaction (5)]. It meets several criteria of future polymers; namely, it is based on renewable starting materials, and its biodegradation products

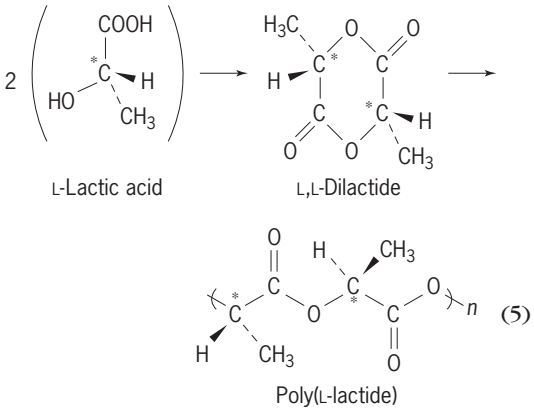

(5)

are water and carbon dioxide. Municipal waste is used as the starting material and converted in a high yield into L,L-dilactide. L,L-dilactide is a six-membered ring with two chiral centers in the molecule, and is prepared from L-lactic acid, a product of bacterial fermentation of mono- or polysaccharides.

The most common initiator is tin octoate [$Sn(Oct)_2$ or $Sn(OC(O)CH(CH_2CH_3)(CH_2)_3CH_3]$. It has recently been shown that $Sn(Oct)_2$ is not an initiator by itself, and like other carboxylates reacts with impurities present in monomer (for example, with lactic acid or water) and gives tin alkoxide, which initiates and then propagates in the same way, as described above for $ROAlR_2$ (that is, on the —SnOR bond).

Although L,L-dilactide is a six-membered cyclic compound that has relatively low ring strain, its thermodynamics of polymerization are favorable enough to ensure high monomer conversion at the polymer-monomer equilibrium. The polymerization proceeds with almost complete retention of configuration (at the stereocenter), in contrast to the anionic polymerization, in which substantial loss of optical activity (racemization) occurs.

The mechanical properties of poly(L-lactide) [containing approximately 5% of D,L-dilactide units] resemble those of polystyrene. Poly(L-lactide) is colorless and transparent, and is mostly used as a commodity plastic for films and fibers. When prepared with special precautions and at the high purity conditions, poly(L-lactide) can be used as a biocompatible material (for example, surgical sutures) whenever bioresorption is required.

In the recent years, the number of published works on, and related to, poly(L-lactide) has been fast increasing, reaching 500 papers in the year 2002.

Cationic ring-opening polymerization. A majority of the monomers that polymerize by anionic mechanisms can also be polymerized cationically. However, several monomers that undergo cationic ring-opening polymerization do not polymerize anionically, such as cyclic acetals, some cyclic ethers (oxetanes, tetrahydrofurans and larger rings), cyclic 1,3-oxaza monomers (for example, oxazolines), and cyclic imines (for example, aziridines). Cationic ring-opening polymerization can be initiated by protonic acids, carbocations, onium ions, or Friedel-Crafts catalysts, most often initiating in the form of complex acids with water. Polymerization proceeds with onium ions as active species, shown in notations (6) and (7). Polyoxazolines upon hydrolysis can be converted into the linear polyamines.

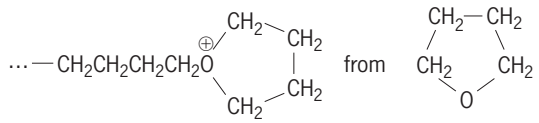

(6)

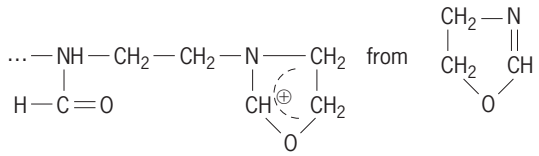

(7)

See POLYACETAL.

Cationic ring-opening polymerization is used in industry mostly for polymerization of 1,3,5-trioxane and its copolymerization with ethylene oxide or 1,3-dioxolane. This is the largest single polymeric material of technical importance made by cationic ring-opening polymerization. It has a number of applications as a high-melting polymer with excellent mechanical properties and high dimensional stability, such as electronic components and automotive parts. *See* COPOLYMER.

The poly(1,3,5-trioxane) [that is, polyformaldehyde] hemiacetal end groups resulting directly from polymerization are unstable, and on heating produce formaldehyde by an unzipping reaction (8).

$$\ldots\text{—}CH_2OCH_2OCH_2OCH_2OH \longrightarrow \ldots\text{—}CH_2OCH_2OCH_2OH + CH_2O \quad (8)$$

The homopolymer is stabilized by blocking the end groups, thus increasing the energy of activation for the first step of depolymerization. Another way to increase the polyacetal stability is to copolymerize 1,3,5-trioxane with ethylene oxide (or 1,3-dioxolane). At the end groups or at their immediate vicinity are units, $\ldots CH_2CH_2O$—, that are stable toward depropagation, since removal of the monomer molecule from the chain's end would require much more energy—in this instance a highly strained three-membered ring would have to be formed.

Another industrially important product is polytetrahydrofuran (polyTHF), mostly used as an oligomer with a molecular weight in the range of a few thousand. The α,ω-dihydroxypolyTHF are used as soft, elastic blocks in elastoplastic multiblock copolymers, namely with aromatic polyesters or polyamide rigid blocks, as well as in elastic fibers. Elastic fibers (such as spandex), which are additives to many modern textiles, are polyurethanes which contain soft and elastic poly(THF) blocks.

Activated monomer mechanism. Cationic ring-opening polymerization of heterocyclic compounds is plagued by inevitable side reactions. Recently a

novel mechanism of cationic ring-opening polymerization was elaborated. It was based on the observation that formation of cyclic side products by backbiting is suppressed in the presence of alcohols or water. Backbiting is the only way (except end-to-end biting) of macrocyclics formation when onium ions' active centers are located at the chain ends [reaction (9)]. In the traditional process, a macromolecule

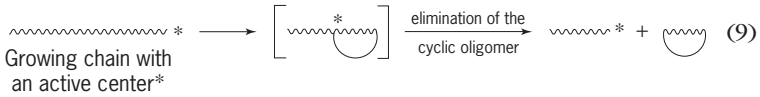

(9)

having active cationic species at its end (∼∗) attacks the nucleophilic unit in the chain (for example, an ether or an ester bond) and expels an oligocyclic compound in the next step.

The observed absence of macrocyclics when polymerization is conducted in the presence of alcohols led to the formulation of a mechanism in which cationic species are no longer at the end of the macromolecules, but are located on the monomer molecules. This is called the activated monomer mechanism (AMM). The simplest system, namely AMM in the cationic ring-opening polymerization of ethylene oxide, is as follows.

1. Protonation of the monomer [reaction (10)].

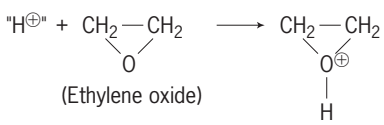

(10)

("$H^{\oplus}$" is a proton, linked ionically to a counterion; counterions are omitted, however.)

2. Initiation: addition (S_N2 substitution) of the initiator (for example, an alcohol molecule ROH) to the activated (that is, protonated) monomer molecule [reaction (11)].

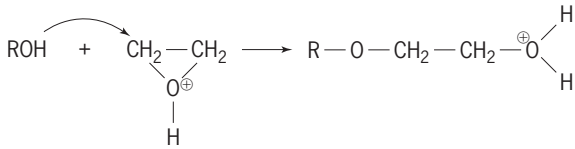

(11)

3. Transfer of a proton from the protonated end group to the next monomer molecule, followed by addition of thus activated monomer molecule to the . . . —OH end group in the growing macromolecule.

Propagation proceeds by repeating this sequence. The tertiary oxonium ions at the end of macromolecules are not formed, eliminating formation of macrocyclics. S. Penczek

Bibliography. G. Allen and J. C. Bevington (eds.), *Comprehensive Polymer Science*, vol. III, Chap. 31–37 (Anionic ROP) and 46–52 (Cationic ROP), Pergamon Press, Oxford, 1984; A. Duda and S. Penczek, Mechanisms of aliphatic polyester formation, Chap. 12 in Biopolymers, vol. 3b: *Polyesters II—Properties and Chemical Synthesis*, edited by A. Steinbüchel and Y. Doi, Wiley-VCH, Weinheim, 2002; K. J. Ivin and T. Saegusa (eds.), *Ring-Opening Polymerization*, vols. I–III, Elsevier Applied Science, London, 1984; S. Penczek, P. Kubisa, and K. Matyjaszewski, *Cationic Ring-Opening Polymerization*, Springer, Berlin, vol. I, 1980, vol. II, 1985; M. Szwarc and M. van Beylen, *Ionic Polymerization and Living Polymers*, Chapman & Hall, New York, 1993.

Ring theory

The mathematical term ring is used to designate a type of algebraic system with two compositions satisfying most but not all the properties of addition and multiplication in the system of integers, 0, ±1, ±2, ... In precise terms a ring is a set R with two binary compositions called addition and multiplication whose results on an ordered pair (a,b), a,b in R, are denoted by $a + b$ and ab, respectively. These compositions must satisfy the following conditions:

C. $a + b$ and ab belong to R (closure).
A1. $a + b = b + a$ (commutative law).
A2. $(a + b) + c = a + (b + c)$ [associative law].
A3. There exist an element 0 (called zero) in R satisfying $a + 0 = a$ for every a in R.
A4. For each a in R there exist an element $-a$ (called the negative of a) in R such that $a + (-a) = 0$.
M1. $(ab)c = a(bc)$.
D. $a(b + c) = ab + ac$; $(b + c)a = ab + ca$ (distributive laws).

In the ring I of integers (addition and multiplication as usual) there are further conditions, for example, the commutative law of multiplication ($ab = ba$) and the cancellation law that if $a \neq 0$ and $ab = ac$, then $b = c$. *See* SET THEORY.

Types of rings. The importance of the concept of a ring stems from the fact that it embraces many special cases which are fundamental in all branches of mathematics. Thus it includes the ring I of integers, the ring R_0 of rational numbers, the ring $R^{\#}$ of real numbers, the ring C of complex numbers, various rings of functions, rings of matrices, and so on. An example of matrix rings is the following: Let $R_n^{\#}$ denote the collection of all the n by n arrays or matrices

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

whose entries a_{ij} are taken in the ring $R^{\#}$ of real numbers. Two such matrices $A = (a_{ij})$ and $B = (b_{ij})$ are regarded as equal if and only if $a_{ij} = b_{ij}$ for every $i,j = 1, 2, \ldots, n$. If $A = (a_{ij})$ and $B = (b_{ij})$, then $A + B$ is defined to be $S = (s_{ij})$ where $s_{ij} = a_{ij} + b_{ij}$, and AB is defined to be $P = (p_{ij})$ where $p_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$. The conditions C, A1–A4, M, and D are fulfilled, so that $R_n^{\#}$ is a ring. If $n = 1$, this is essentially the same as the ring $R_n^{\#}$, but if $n > 1$, then $R_n^{\#}$ differs from $R^{\#}$ because the commutative law of multiplication does not hold; that is, there are instances in which $AB \neq BA$. The example $R_n^{\#}$ has a geometric counterpart, namely, the ring of linear transformations of an n-dimensional vector

space. In fact these rings are isomorphic in the sense defined below. The first example of a ring in which the cummutative law of multiplication does not hold is the system Q of quaternions introduced by W. R. Hamilton as an appropriate tool for studying rotations in three-dimensional space. This system is the set of all the expressions of the form $a1 + bi + cj + dk$ where a, b, c, d are in $R^{\#}$. Addition is defined by adding corresponding coefficients as in ordinary algebra, and multiplication is defined to satisfy the associative and distributive laws and the following rules for the quaternion units 1, i, j, k: $1u = u = u1$ for $u = 1, i, j, k$; $i^2 = k^2 = -1$: $ij = -ji = k$, $jk = -kj = i, ki = -ik = j$. *See* QUATERNIONS.

The conditions A1–A4 on the addition composition are exactly equivalent to the statement that any ring is a commutative group relative to its addition composition. This group is called the additive group of the ring. The algebraic system consisting of the set of elements of a ring together with its multiplication composition is called the multiplicative semigroup of the ring. *See* GRAPH THEORY.

Various classes of rings are singled out by imposing conditions on the multiplicative semigroup. Thus integral domains are rings in which the product of nonzero elements is nonzero. Division rings are groups relative to the multiplication composition, and fields are division rings satisfying the commutative law of multiplication. The system of quaternions Q is a noncommutative division ring, $R^{\#}$ and C are fields. Important instances of integral domains are the rings $F[x_1, x_2, \ldots x_r]$ of all formal polynomials in the letters x_i with coefficients taken out of some field F.

If R is a ring, then a subring of R is a subset S of R, which is a ring with respect to the addition and multiplication defined in R. The conditions for this are that if s_1 and s_2 belong to S, then so do $s_1 - s_2$ $[= s_1 + (-s_2)]$ and s_1s_2. In a similar manner the concept of a subfield of a field is defined.

Ideals, difference rings, homomorphism. In elementary number theory it is often important to separate the ring I of integers into subsets defined by a divisibility condition. For example, there are the sets of even and odd integers. More generally, if m is a positive integer, then the set of integers decomposes into m subsets $I_0, I_1, \ldots, I_{m-1}$ where I_j is the set of integers which gives the remainder j on division by m. Two integers a and b are in the same I_j if and only if $a - b$ is divisible by m. If this condition holds, one writes, following Gauss, $a \equiv b \pmod{m}$ which is read, "a is congruent to b modulo m." Congruences can be added and multiplied, leading to a new ring whose elements are the m sets $I_0, I_1, \ldots, I_{m-1}$. The study of this ring gives a natural setting for important results of number theory.

The construction just indicated can be carried over to any ring R. One begins with an ideal M in R which is defined to be a subset of R having the following closure properties:

1. If m_1, m_2 are in M, then $m_1 - m_2$ is in M.
2. If m is in M and a is any element of R, then am and ma are in M. If $R = I$ the ring of integers, then the set M of multiples of a fixed positive integer m is an ideal.

As in this special case, two elements a,b of any ring R are said to be congruent modulo an ideal M if $a - b$ is in M. The ring M can be decomposed into nonoverlapping congruence classes where such a class $[a]$ is the complete set of elements x of R which are congruent to a fixed a. The congruence classes can be added and multiplied and are the elements of a new ring R/M called the difference (or factor or quotient) ring. This is the ring analog of the quotient group defined in the theory of groups. Also, as in group theory, the mapping which associates with every a of R the congruence class $[a]$ of R/M is the prime instance of a homomorphism. If R and R' are any two rings, a mapping $a \to f(a)$ of R into R' is a homomorphism if $f(a + b) = f(a) + f(b)$ and $f(ab) = f(a)f(b)$. If distinct elements have distinct images, then f is called an isomorphism, and R and its image are said to be isomorphic. The image under any homomorphism is isomorphic to the difference ring R/M, where M is the ideal of elements mapped into 0. This basic result is the fundamental theorem of homomorphism for rings. *See* GROUP THEORY.

Advanced aspects. Ideals play an important role in higher arithmetic, which deals with the arithmetic aspects of number fields. Ideals are also basic in algebraic geometry. In fact, the theory of algebraic curves makes considerable use of a type of ideal theory, called Dedekind ideal theory, which includes the ideal theory of number fields. On the other hand, higher-dimensional algebraic geometry can be founded on the ideal theory of so-called noetherian rings.

The structure theory of rings is essentially an analysis of the anatomy of rings. The general idea is that of reducing the study of classes of rings to that of simpler classes. Intertwined with this is the theory of representations of rings. A representation of a ring is defined to be a homomorphism of a ring into a ring of endomorphisms of a commutative group. *See* BOOLEAN ALGEBRA; LATTICE (MATHEMATICS); NUMBER THEORY. Nathan Jacobson

Bibliography. N. Jacobson, *Lectures in Abstract Algebra*, vols. 1 and 2, 1951 and 1953, reprint 1981 and 1984; N. Jacobson, *Structure of Rings*, 2d ed., 1964, reprint 1990; I. Kaplansky, *Commutative Rings*, 1970, reprint 1994; S. MacLane and G. Birkhoff, *Algebra*, 3d ed., 1987; D. Passman, *A Course in Ring Theory*, 1991; L. H. Rowen, *Ring Theory*, student edition, 1991; O. Zariski and P. Samuel, *Commutative Algebra*, 2 vols., 1958 and 1960, reprint 1986 and 1991.

Ripple voltage

The time-varying part of a voltage that is ideally time-invariant. Most electronic systems require a direct-current voltage for at least part of their operation. An ideal direct-current voltage is available from a battery, but batteries are impractical for many applications. To obtain a direct-current voltage from

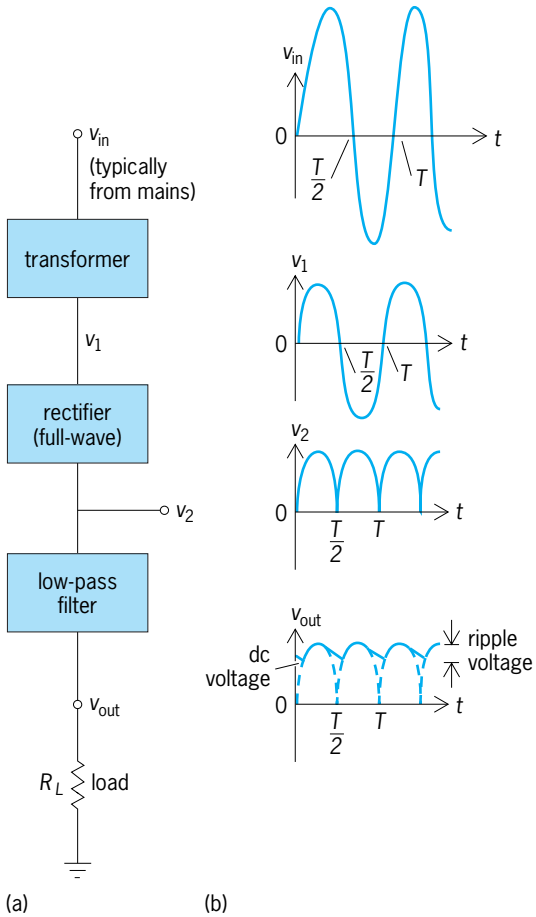

Power supply system with ripple voltage defined. (*a*) System configuration. (*b*) Voltage waveforms. *T* = period of input voltage.

the alternating-current power mains requires using some type of power supply.

A typical linear power supply system configuration (see **illus.**) consists of a transformer to change the voltage at the mains to the desired level, a rectifier to convert the alternating-current input voltage v_1 to a pulsating direct-current voltage v_2, followed by a low-pass filter. The output voltage v_{out} of the filter consists of a large direct-current voltage with a superimposed alternating-current voltage. This remaining superimposed alternating-current voltage is called the ripple voltage. The percent ripple (in contrast to the ripple voltage) is defined by the equation below.

$$\%\ \text{ripple} = \frac{\text{root-mean-square (rms) value of the ac component}}{\text{dc or average value of } v_{out}} \times 100\%$$

The root-mean-square value of the alternating-current component (that is, the ripple voltage) is typically what is measured by an alternating-current voltmeter, while the direct-current value of v_{out} is measured by a direct-current voltmeter. From these two measurements, this formula can be used to obtain a quantitative value for the percent ripple. The ripple voltage has a fundamental frequency of twice the input voltage frequency for a single-phase, full-wave rectifier, and a fundamental frequency equal to that of the input voltage for a single-phase, half-wave rectifier. The ideal ripple would be 0%, which would mean that v_{out} consists of only a direct-current component. Ripple is often less than 1% in a well-designed system. The ripple voltage can be reduced at the expense of increasing the size of some of the components of the low-pass filter or by reducing the load-current drain on the circuit by increasing the value of R_L. The low-pass filter is more effective in reducing ripple voltage if the frequency of the input voltage is higher. A higher input frequency is not practical for operation from 60-Hz alternating-current mains but is used in switched-mode power supplies.

Practical linear power supplies often include a voltage regulator between the low-pass filter and the load. The voltage regulator is usually an electronic circuit that is specifically designed to provide a very stable dc output voltage even if large variations occur in the input. Nonlinear power supplies, which are often termed switching power supplies or switched-mode power supplies, are a practical alternative for producing a low-ripple dc output. *See* ELECTRIC FILTER; ELECTRONIC POWER SUPPLY; RECTIFIER; VOLTAGE REGULATOR.

Stanley G. Burns

Risk assessment and management

The scientific study of risk, the potential realization of undesirable consequences from hazards arising from a possible event, the assessment of the acceptability of the risks, and the management of unacceptable risks. For example, the probability of contracting lung cancer (unwanted consequence) is a risk caused by carcinogens (hazards) contained in second-hand tobacco smoke (event). The risk is estimated using scientific methods and then the acceptability of that risk is assessed by public health officials. Risk management is the term for the systematic analysis and control of risk, such as prohibiting smoking in public places. Risks are caused by exposure to hazards. Sudden hazards are referred to as acute (for example, a flash flood caused by heavy rains); prolonged hazards are referred to as chronic (for example, carcinogens in second-hand tobacco smoke and polluted air).

The definition of risk contains two components: the probability of an undesirable consequence of an event and the seriousness of that consequence. In the example of a flash flood, risk can be defined as the probability of having a flood of a given magnitude. Sometimes the probability is expressed as a return period, which means, for instance, that a flood of a specified magnitude is expected to occur once every 100 years. The scope of a flood can be expressed as the level or stage of a river, or the dollar amount of property damage.

Most human activities involve risk. The risk of driving, for example, can be subdivided according to property damage, human injuries, fatalities, and harm to the environment. Even the stress and lack of exercise due to driving create health risks. Although

risk pervades modern society and is widely acknowledged, it continues to cause unending controversy and debate. In the example of driving a car, the issues include which risks should be studied and how they should be measured and managed; which safety features should be incorporated into cars in order to reduce certain risks; whether governments should require stronger automobile frames to reduce the severity of accidents; and whether the use of safety belts should be required in the front and rear seats. People perceive risks in different ways; decision makers can exercise more effective risk management by improved public risk communication.

Risk analysis and assessment are applied to new technologies, chemicals, air-quality standards, the adequacy of emergency response forces, and other risk problems. However, risk estimates are seldom accurate to even two orders of magnitude, and widely varying perceptions of risk by different interest groups can add confusion and conflict to the risk management process. As a result, risk analysis and assessment are often carried out conservatively with a safety factor as high as 1000. This, in turn, leads to nonoptimal decision making, because the risks are exaggerated.

Environmental risk assessment is laden with uncertainty, particularly with respect to the quantification of chemical emissions; the nature of contaminant transport (such as the region over which a chemical may spread and the velocity of movement) in the water, air, and soil; the type of exposure pathway (such as inhalation, ingestion, and dermal contact); the effects on people based on dose-response studies (which are extrapolated from animals); ecological impacts; and so forth. Moreover, the public perception of risk is shaped by powerful psychological and social processes such as the dread of cancer, the involuntary nature of the exposure, and the risks to future generations. This outlook may lead to biases when judging risks (such as overestimating the frequency of very uncommon events), public demands for zero-risk criteria (such as the no-threshold criterion), and improved compliance with environmental regulations; and also may raise a myriad of socioeconomic issues, from social equity and competiveness to who bears the costs for improved environmental stewardship. To this end, the use of effective risk communication and risk principles such as ALARA (as low as reasonably achievable) and de minimus (the risk is so small that most people are uninterested in giving up the risk-producing activity, or the risk does not warrant regulatory enforcement) may help to promote more effective environmental risk management, particularly for the most noxious environmental risks.

Risk management. Thousands of natural and other hazards are subjected to the statistical analysis of mortality and morbidity data. Society selects a small number of risks to manage, but often some high risks (such as radon in houses) may not be managed, while some low risks (such as movement of dangerous goods) may be selected for management. Management alternatives include banning of the hazard (drugs), regulating the hazard (drivers' tests and licensing), controlling the release and exposure of hazardous materials, treatment after exposure, and penalties for damages. Each management alternative may be analyzed to estimate the impact on risk, which may be known only within a few orders of magnitude (**Fig. 1**).

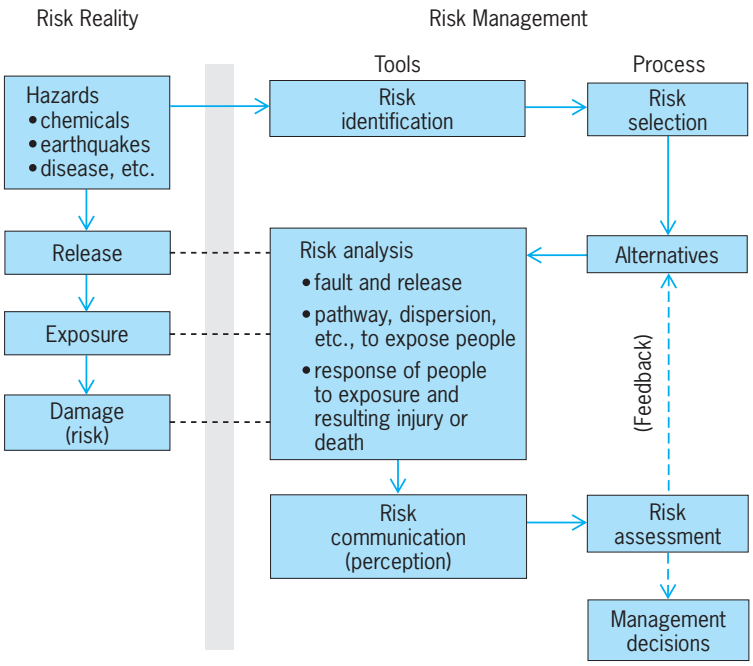

Fig. 1. Risk analysis, assessment, and management.

Risk estimates are uncertain, are described in technical language, and are outside the general understanding or experience of most people. Perception plays a crucial role, tending to exaggerate the significance, for example, of risks that are involuntary, catastrophic, or newsworthy. Effective risk management therefore requires effective risk communication.

Risk assessment is the evaluation of the relative importance of an estimated risk with respect to other risks faced by the population, the benefits of the activity source of the risk, and the costs of managing the risk. For risks due to long-term exposure to chemicals, the risk assessment activity generally incorporates the estimation of the response of people to the exposure (that is, risk analysis is a part of risk assessment). The methods used include studies on animals, exposure of tissues, and epidemiology.

Examples of applications. The general approach to risk management outlined above is needed for solving risk problems arising in areas ranging from the disposal of nuclear wastes to the design of safe toys. Below, examples of risk analysis and risk assessment are used to illustrate how risk management is carried out in practice.

Risk analysis. The nuclear and chemical industries have developed methods of probabilistic risk assessment, using techniques such as event trees and fault trees. The top event in a fault tree is the fault, and the tree indicates how it can occur and with what frequency. In a fault tree for an overflowing gasoline

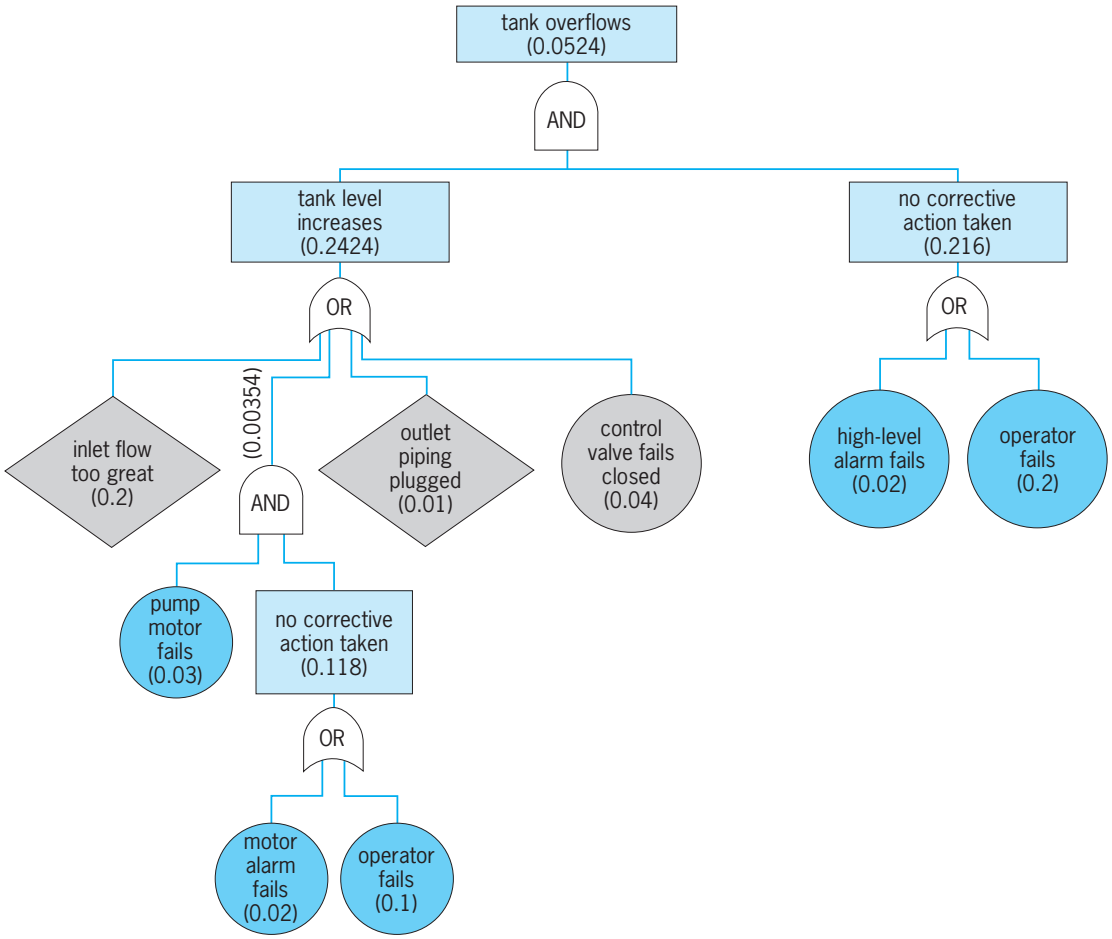

Fig. 2. Fault tree for an overflowing tank. The numbers in parentheses represent probabilities. (*After Chemical Manufacturers Association, Process Safety Management* (*Control of Acute Hazards*), *Washington, D.C., May 1985*)

tank, one of the most common causes of industrial accidents (**Fig. 2**), the tank overflows if it is being filled too fast AND the filling is not stopped. The probability associated with the AND logical operator can be calculated by the product of the relevant probabilities. An OR gate indicates that any of the input events is sufficient, and the probabilities are found by the complement of the probability that none of the input events would occur. For example, "no corrective action" can occur if the high-level alarm fails OR the operator fails to respond. Basic events for which data exist on failure rates are shown in circles. For example, data on thousands of motors can be used to estimate the mean failure rate per year and the variation in failure rates. Events that are specific to the situation, and whose probabilities must be estimated, are shown in diamond-shaped boxes. Events whose probabilities are calculated internal to the fault tree are shown in squares. Fault trees permit probability calculations of rare events through the application of logic and the knowledge of failure rates for relatively common events.

If a basic event such as "electrical power supply fails" is used as input to a number of roots of the failure tree, the probability of the fault or the top event is calculated by reducing the tree to a number of independent cut sets using boolean logic. Fault trees can also be used to evaluate alternative risk management strategies such as using an automatic shut-off in place of operator response to the high-level alarm.

An event tree generally starts from a fault such as the tank overflow problem. The event tree then continues as a typical decision tree to list the occurrence or nonoccurrence of subsequent events that can lead to the risk of loss of life. *See* DECISION THEORY.

Risk assessment. Health risks in the home or office are caused by many kinds of hazardous air pollutants, of which cigarette smoke is one of the best known. Another pollutant is radon, a radioactive gas occurring naturally by the nuclear decay of radium in the ground, which can enter houses through the basements. Less familiar indoor air pollutants include the great variety of complex organic chemicals that are given off by building materials, paints, furniture, cleaning fluids, and other manufactured products. Because buildings are often designed to be relatively airtight, air pollutants can build up to fairly high levels. Many indoor air pollutants not only are carcinogenic but can also lead to respiratory problems. Unfortunately, causal relationships are usually difficult to establish and consequently risk analyses are subject to great uncertainty.

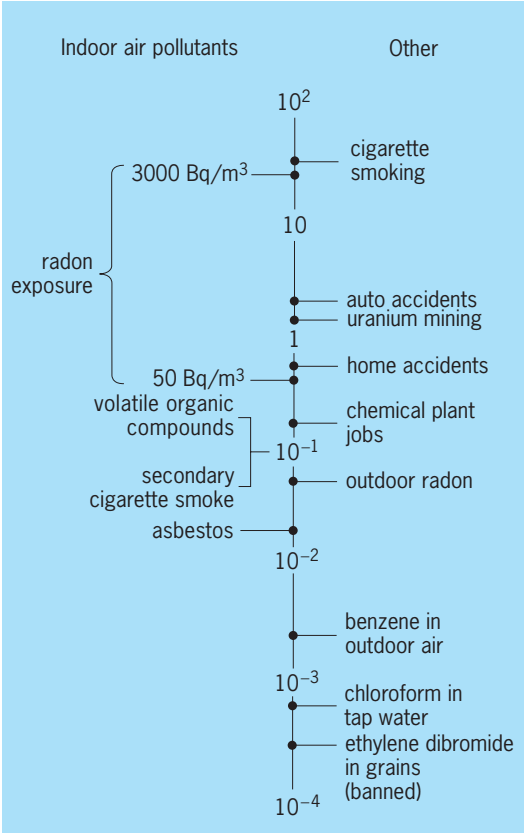

Fig. 3. Risk assessment scale, showing percentage lifetime risk of premature death from indoor radon exposure compared to other hazards; 1 becquerel per cubic meter = 0. 027 picocurie per liter. (*After A. V. Nero, Jr., Controlling indoor air pollution, Sci. Amer., 258*(5):42–48, May 1988)

A risk comparison (**Fig. 3**) shows that indoor radon exposure varies widely but is more dangerous than many familiar hazards. Such a comparison also shows that the estimated probability of suffering a fatal disease is significantly higher for exposure to indoor air pollutants than for exposure to pollutants in outdoor air, drinking water, and food; the average indoor concentration of radon in the United States translates to about 10,000 lung cancer deaths per year. The outdoor air pollution is kept low by regulation, and the resulting risk of cancer is actually two orders of magnitude below the risk due to chemicals in the indoor environment. *See* AIR POLLUTION; ENVIRONMENTAL ENGINEERING; ENVIRONMENTAL TOXICOLOGY; HAZARDOUS WASTE; NUCLEAR POWER; ONCOLOGY; RADIOACTIVE WASTE MANAGEMENT; RADON.

Keith W. Hipel; John Shortreed

Bibliography. P. L. Bernstein, *Against the Gods: The Remarkable Story of Risk*, 1996; Canadian Standards Association, *Risk Management: Guideline for Decision-Makers*, 1997; Y. Y. Haines, *Risk Modeling, Assessment and Management*, 1998; K. W. Hipel and L. Fang (eds.), *Effective Environmental Management for Sustainable Development*, 1994; G. Holmes, B. R. Singh, and L. Theodore, *Handbook of Environmental Management and Technology*, 1993; S. Kaplan, The words of risk analysis, *Risk Analysis*, 17(4): 407-417, 1997; M. G. Morgan, Risk analysis and management, *Sci. Amer.*, 269(1):32–41, July 1993; National Research Council, *Improving Risk Communication*, 1989; National Research Council, *Science and Judgment in Risk Assessment*, 1994; Royal Society of London, *Risk: Analysis, Perception and Management*, 1992.

Ritz's combination principle

The empirical rule, formulated by W. Ritz in 1905, that sums and differences of the frequencies of spectral lines often equal other observed frequencies. The rule is an immediate consequence of the quantum-mechanical formula $hf = E_i - E_f$ relating the energy hf of an emitted photon to the initial energy E_i and final energy E_f of the radiating system; h is Planck's constant and f is the frequency of the emitted light. For example, the **illustration** shows the photon

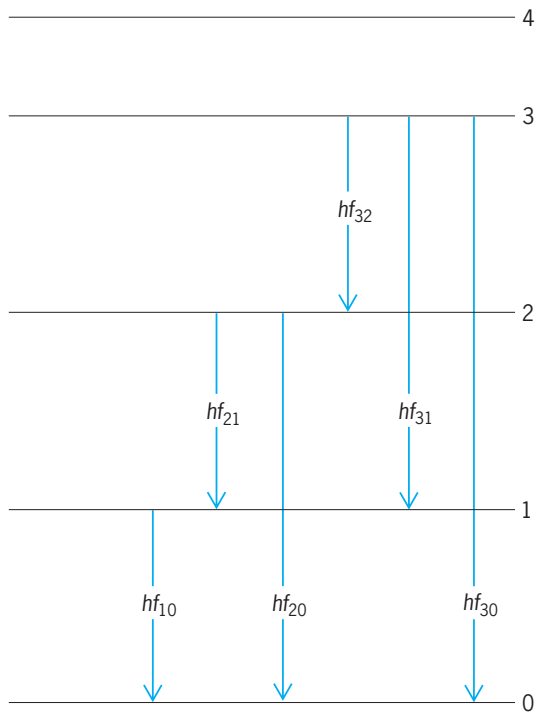

Energy levels and emitted frequencies.

energies hf_{32}, hf_{31}, hf_{30} associated with transitions from level 3 to lower-lying levels, and so on. Level 3 may radiate directly to the ground state 0, emitting f_{30}, or it may first make a transition to level 2, which subsequently radiates to the ground state, and so on. Since the total energy emitted in these two alternative means of making transitions from 3 to 0 is exactly the same, namely $E_3 - E_0$, it follows that $hf_{30} = hf_{32} + hf_{20}$. Similarly, $hf_{30} = hf_{32} + hf_{21} + hf_{10}$, and so forth. *See* ATOMIC STRUCTURE AND SPECTRA; ENERGY LEVEL (QUANTUM MECHANICS); QUANTUM MECHANICS.

Edward Gerjuoy

River

A natural, fresh-water surface stream that has considerable volume compared with its smaller tributaries. The tributaries are known as brooks, creeks, branches, or forks. Rivers are usually the main stems and larger tributaries of the drainage systems that convey surface runoff from the land. Rivers flow from headwater areas of small tributaries to their mouths, where they may discharge into the ocean, a major lake, or a desert basin.

Rivers flowing to the ocean drain about 68% of the Earth's land surface. The remainder of the land either is covered by ice or drains to closed basins (common in desert regions). Regions draining to the sea are termed exoreic, while those draining to interior closed basins are endoreic. Areic regions are those which lack surface streams because of low rainfall or lithoogic conditions.

Sixteen of the largest rivers (see **table**) account for nearly half of the total world river flow of water. The Amazon River alone carries nearly 20% of all the water annually discharged by the world's rivers. Rivers also carry large loads of sediment. The total sediment load for all the world's rivers averages about 22×10^9 tons (20×10^9 metric tons) brought to the sea each year. Sediment loads for individual rivers vary considerably. The Yellow River of northern China is the most prolific transporter of sediment. Draining an agricultural region of easily eroded loess, this river averages about 2×10^9 tons (1.8×10^9 metric tons) of sediment per year, one-tenth of the world average. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; LOESS.

River floods. River discharge varies over a broad range, depending on many climatic and geologic factors. The low flows of the river influence water supply and navigation. The high flows are a concern as threats to life and property. However, floods are also beneficial. Indeed, the ancient Egyptian civilization was dependent upon the Nile River floods to provide new soil and moisture for crops.

Floods are a natural consequence of the spectrum of discharges exhibited by a river. Rivers in humid temperate regions often exhibit less variable flood behavior than rivers in semiarid regions of rugged terrain. Floods in the humid regions tend to occur on average once each year. Rarer, larger floods are often no more than one or two times the magnitude of more common annual floods. The semiarid floods, however, are usually very small for common events, such as the average annual flood. However, rare, high-magnitude floods may be catastrophic.

The greatest floods in the geologic record occurred during the Pleistocene about 13,000 years ago in the northwestern United States. Lake Missoula, an ice-dammed lake in western Montana, released several catastrophic floods across the Channeled Scabland region of eastern Washington. The largest of these floods discharged as much as 20×10^6 ft^3 (570,000 m^3) of water per second. Flow velocities in the Scabland channelways ranged from 33 to 100 ft/s (10 to 30 m/s) for water 100–330 ft (30–100 m) deep. These phenomenal floods created a bizarre landscape of anastomosing channels, abandoned cataracts, streamlined hills, and immense gravel bars. So much water entered the preflood river valleys that they filled to overflowing, and floodwater scoured the divide areas between the valleys.

The Missoula floods were certainly among the most spectacular fluvial phenomena of all time. However, it should be remembered that most rivers do their work very slowly. Rivers are mainly transport agents, removing debris produced by the prolonged action of rainsplash, frost action, and mass movement. The many smaller floods probably accomplish much more of this work of transport than the rare large flood.

Floods are but one attribute of rivers that affect human society. Means of counteracting the vagaries of river flow have concerned engineers for centuries. In modern times many of the world's rivers are managed to conserve the natural flow for release at times required by human activity, to confine flood flows

Characteristics of some of the world's major rivers

River	Average discharge, ft^3/s (m^3/s)	Drainage area, 10^3 mi^2 (10^3 km^2)	Average annual sediment load, 10^3 tons (10^3 metric tons)	Length, mi (km)
Amazon	6,390,000 (181,000)	2770 (7180)	990,000 (900,000)	3899 (6275)
Congo	1,400,000 (39,620)	1420 (3690)	71,300 (64,680)	2901 (4670)
Orinoco	800,000 (22,640)	571 (1480)	93,130 (86,490)	1600 (2570)
Yangtze	770,000 (21,790)	749 (1940)	610,000 (550,000)	3100 (4990)
Brahmaputra	706,000 (19,980)	361 (935)	880,000 (800,000)	1700 (2700)
Mississippi-Missouri	630,000 (17,830)	1240 (3220)	379,000 (344,000)	3890 (6260)
Yenisei	614,000 (17,380)	1000 (2590)	11,600 (10,520)	3550 (5710)
Lena	546,671 (15,480)	1170 (3030)	—	2900 (4600)
Mekong	530,000 (15,000)	350 (910)	206,850 (187,650)	2600 (4180)
Parana	526,000 (14,890)	1200 (3100)	90,000 (81,650)	2450 (3940)
St. Lawrence	500,000 (14,150)	564 (1460)	4,000 (3,630)	2150 (3460)
Ganges	497,600 (14,090)	451 (1170)	1,800,000 (1,600,000)	1640 (2640)
Irrawaddy	478,900 (13,560)	140 (370)	364,070 (330,280)	1400 (2300)
Ob	440,700 (12,480)	1000 (2590)	15,700 (14,240)	2800 (4500)
Volga	350,000 (9,900)	591 (1530)	20,780 (18,840)	2320 (3740)
Amur	338,000 (9,570)	788 (2040)	—	2900 (4670)

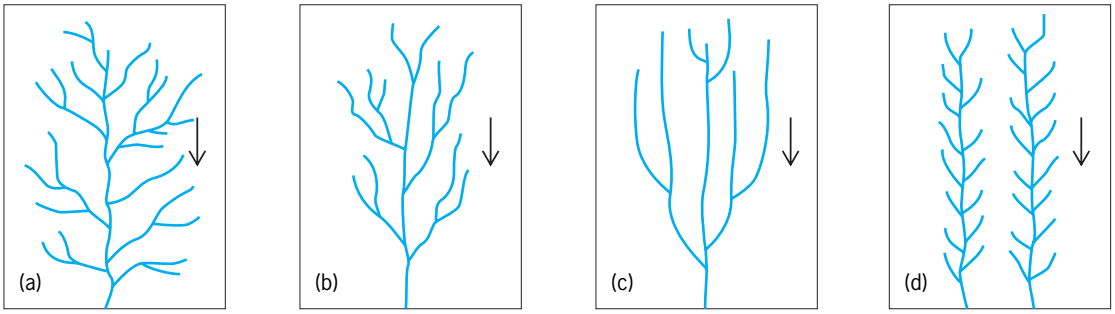

Fig. 1. Typical slope angles and associated drainage patterns forming on relatively homogenous planar beds: (*a*) surface gradient 1%, dendritic; (*b*) surface gradient 3%, subparallel; (*c*) surface gradient 5%, parallel; and (*d*) surface gradient 5%, pinnate. Arrows indicate direction of flow. (*After L. F. Phillips and S. A. Schumm, Effects of regional slope on drainage networks. Geology, 15:813–816, 1987*)

to the channel and to planned areas of floodwater storage, and to maintain water quality at optimum levels. *See* FLOODPLAIN; RIVER ENGINEERING.

Geologic history. Some rivers possess a long heritage related to their location along relatively stable continental cratons or their correspondence to structural lows, such as rift valleys. More commonly, rivers have been disrupted through geologic time by the Earth's active tectonic and erosional processes. The most recent disruptions were caused by the glaciations of the Pleistocene. Many rivers, like the Mississippi, became heavily loaded with coarse glacial debris delivered to their headwaters by glaciers. Since the last glacial maximum about 18,000 years ago, most of the world's rivers have adjusted their channel sizes, patterns, and gradients to the new environmental conditions of postglacial time.

In tropical regions the effects of glaciation were indirect. During full-glacial episodes of the Pleistocene, tropical areas of the Amazon and Congo river basins experienced relative aridity. Forests were replaced by savanna and grassland. The greater erosion rates on the land contributed large amounts of coarse sediment to the rivers. Today these regions have returned to their high-rainfall condition. The rivers receive relatively little sediment from interfluves that are stabilized by a dense forest cover. However, human exploitation of the tropical forests seems to be effectively returning the landscape to its glacial condition. This will undoubtedly induce profound changes in many tropical rivers. Victor R. Baker

Stream pattern analysis. The patterns of intersecting stream channels provide valuable information about the geology, topography, climate, and hydrology of the region, and about some of the ways in which humans have altered the land surface. The alignment of the streams (for example, branched, rectangular, or parallel) is controlled by bedrock type, topography, and the locations of faults, folds, and joint patterns. The density of the network (that is, how closely spaced the channels are) is controlled by the climate, vegetation, age, permeability, and slope of the surface. Stream patterns are thus a product of the physical geography of a region and provide valuable clues to that geography.

Geologists and geographers have long used stream patterns for landscape analysis. An interesting application of pattern analysis is in extraterrestrial studies; scientists have used stream patterns to identify surficial processes on planets where on-site collection of data is impossible.

Form, topography, and geology. The general form of stream patterns reflects the underlying topography and geology of a region. Dendritic, or branchlike, drainage (**Fig. 1***a*) is the most commonly occurring pattern; it indicates a lack of strong structural controls (that is, faults, joints, or folds) or complex topography. Dendritic patterns thus commonly occur on relatively horizontal sediments or beveled plains of any rock type. Large drainage networks such as the Mississippi River almost always display dendritic patterns, with other patterns occurring at smaller area scales.

If there are no structural controls or significant variations in rock type and the surface gradient is increased to 3% (that is, a rise of 3 m in 100 m of horizontal distance), the dendritic pattern typically assumes a subparallel pattern (Fig. 1*b*). At slopes of 5% and higher on planar surfaces, the subparallel pattern usually goes through a transition to a parallel pattern (Fig. 1*c*) on most rock surfaces, or a pinnate pattern on silts and clays (Fig. 1*d*).

The shape of the topography as well as the gradient also plays a role in controlling the form of drainage patterns. Volcanoes create radial drainage patterns (**Fig. 2***a*). Topographic lows, such as occur in the closed basins of Nevada, create the centripetal pattern shown in Fig. 2*b*. The hummocky terrain of permafrost areas, regions of limestone solution, and recently deposited materials from glaciers, volcanic explosions, and slides creates a deranged pattern, which consists of many ponds and unconnected tributaries (Fig. 2*c*). Deranged patterns may evolve to become dendritic as streams erode the divides between ponds and eventually form an interconnected network of tributaries (Fig. 2*d*).

Structural controls can also override the natural tendency for drainages to assume a dendritic form, or a parallel form on slopes with gradients greater than 5%. For example, parallel drainages can occur on slopes of less than 5% in areas such as the south shore of Lake Superior, where ancient glacial grooves restrict flow to parallel channels. Trellis patterns (Fig. 2*e*) generally indicate parallel fractures or

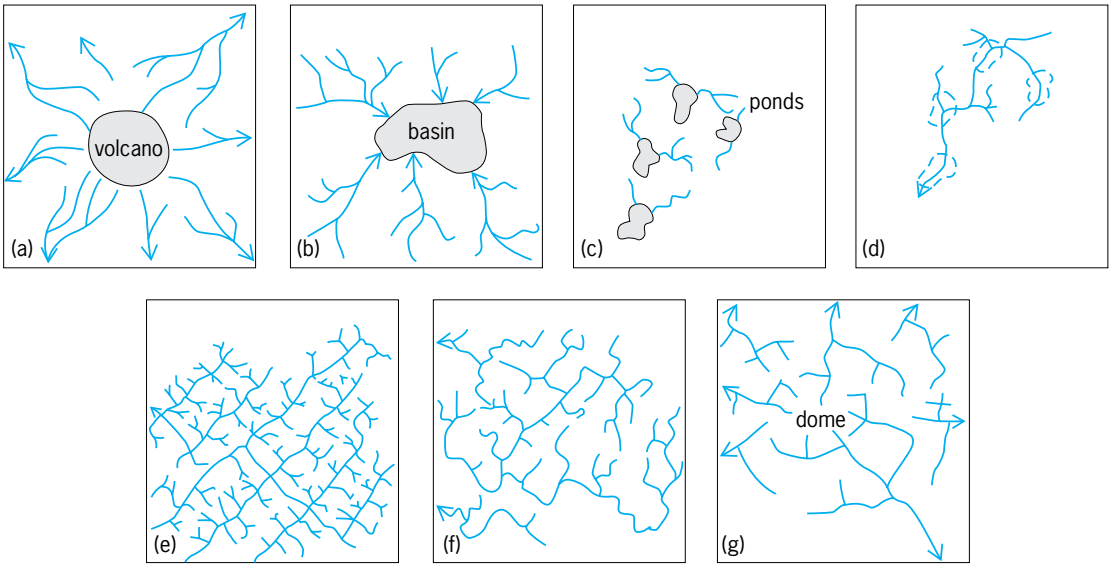

Fig. 2. Diagrams of broad categories of structurally and topographically controlled drainage pattern forms. (*a*) Radial. (*b*) Centripetal. (*c*) Deranged. (*d*) Dendritic. (*e*) Trellis. (*f*) Rectangular. (*g*) Annular. The dendritic drainage (*d*) shows the pattern that evolves at the deranged pattern (*c*) location after sufficient time has passed for erosion of the divides between ponds and formation of a connected channel network. Broken outlines in *d* show old pond locations in the deranged pattern (*c*). (*After A. D. Howard, Drainage analysis in geologic interpretation: A summation, Amer. Ass. Petrol. Geol. Bull., 51:2246–2259, 1967*)

parallel dipping rock beds (for example, old beaches that form resistant sandstone), which create long reaches of parallel streams with smaller tributaries to either side. The rectangular pattern shown in Fig. 2*f* reflects an underlying structure or joints or faults at right angles. Annular patterns occur in areas where the rocks have been bent to form domes and basins (Fig. 2*g*).

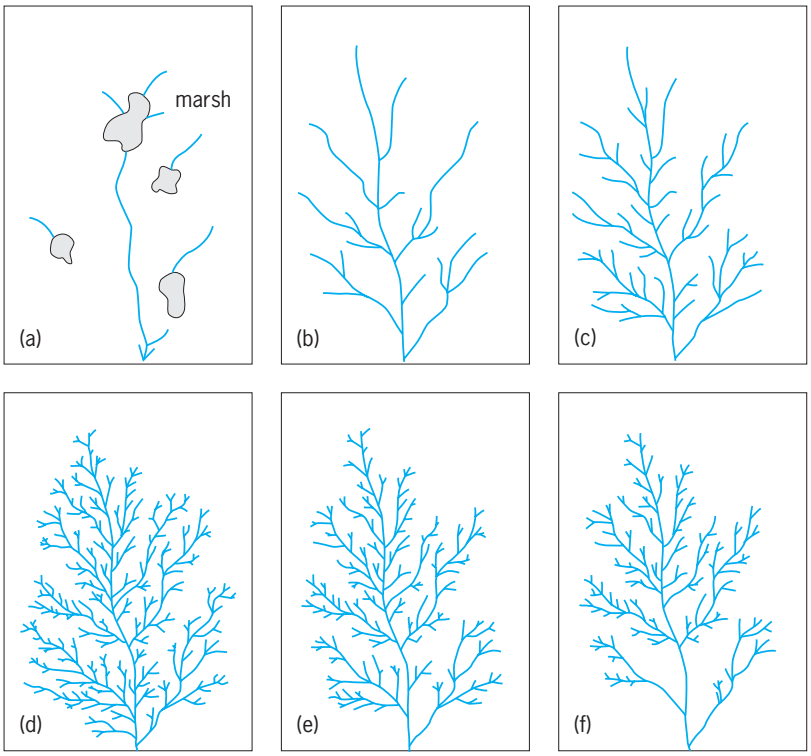

Fig. 3. Drainage pattern evolution from (*a*) Initiation as a deranged pattern through (*b*)–(*d*) maximum channel development in a dendritic pattern to (*e, f*) erosional lowering of the basin and loss of channels from the dendritic pattern. (*After W. S. Glock, The development of drainage systems: A synoptic view, Geog. Rev., 21:475–482, 1931*)

Pattern density. The various stream patterns can occur at different densities; in other words, channels may range from being closely to widely spaced. Density can be controlled by climate, vegetation, bedrock and soil permeability and strength, slope and shape of surface, and the age of a surface and stream. In turn, human disruption of the landscape can affect all these factors at a variety of scales, from a garbage dump to a river basin.

The evaluation of controls on channel density is important because channel density provides key clues to the physical geography of drainage basins. Furthermore, erosion and flash floods are controlled by drainage density, because more channels mean more runoff and excavation of sediment. Drainage densities therefore have been used by hydrologists to estimate mean annual flows, annual flood sizes, low flows due to ground-water discharge into streams, and sediment yields. Finally, by understanding the controls on drainage density, engineers can design mine tailing piles, landfills, and other artificial landscapes that help maintain drainage densities that reduce erosion and flash floods. *See* GROUND-WATER HYDROLOGY; HYDROLOGY.

Climate and vegetation. Drainage density is most strongly controlled by the intensity of precipitation (the amount of precipitation falling per unit time) at regional scales; other factors generally play a role at local scales. The widest variation in drainage densities occurs in semiarid climates, where local factors often override climatic controls.

Despite the impact of local factors, climate-driven factors generally cause densities to be greatest in semiarid areas such as the western Great Plains the United States, where occasional, intense thunderstorm activity coupled with sparse vegetation cover

promotes erosion and the formation of many channels. Densities typically decrease in more arid climates because of the absence of rainfall and runoff. Densities also typically decrease in more humid climates, probably because of (1) increased vegetation, which holds soil in place and reduces erosion, and (2) increased soil cover, which acts in a spongelike manner to absorb and store water, thus reducing the surface runoff that erodes channels.

Rock type and soil permeability. Surfaces possessing high permeability, such as sands or loose volcanic debris, generally produce low stream densities, because precipitation can soak into the surface rather than generating runoff and erosion. Similarly, surfaces such as clays or silts that have low permeability produce high drainage densities, because water cannot sink into the surface, and so runs off and erodes channels. Badlands, which are barren, heavily eroded and dissected landscapes, almost invariably consist of shales or silts that cannot absorb precipitation. On occasion, however, this relationship can be confusing. Granites, for example, are impermeable to water, but are also sufficiently strong to resist extensive erosion and channel formation. In addition, the soils that form from granites are very porous and can absorb significant quantities of rainfall.

Age of the surface. The older the surface, the more time has been available for erosion and network formation. Thus, if there are identical surfaces that have experienced identical climates, the surface will progress through stages of little erosion, tributary development and high drainage density, and finally reduction in density as erosion reduces lower portions of the basin to a relatively flat plane (**Fig. 3**). This process may be repeated if the basin is uplifted or the base level (for example, sea level or the level of a reservoir) drops, both events starting a new cycle of erosion.

At the scale of small areas (for example, one hillside) in soft bedrock or soils, the drainage density may also change on a seasonal basis. During the cold or dry season, channels may be destroyed as frost action loosens dry sediment that fills in the channel, while during the wet season, channels are reestablished by runoff.

Slope and shape of the surface. If all other factors are equal, drainage densities increase with gradient up to a material-dependent threshold value. Densities increase because steeper slopes have less soil cover to absorb rain, and runoff flows more quickly and has more erosive power on steeper gradients. At lower gradients, upwardly convex slopes tend to produce slightly lower drainage densities than planar or concave slopes, but this difference in densities disappears as gradients increase to the threshold value.

W. Andrew Marcus

Bibliography. E. Derbyshire (ed.), *Geomorphology and Climate*, 1976; W. L. Graf, *Fluvial Processes in Dryland Rivers*, 1988; L. Leopold, M. G. Wolman, and J. P. Miller, *Fluvial Processes in Geomorphology*, 1995; L. F. Phillips and S. A. Schumm, Effects of regional slope on drainage networks, *Geology*, 15:813–816, 1987; S. A. Schumm, *The Fluvial System*, 1977; S. A. Schumm, *River Morphology*, 1982; S. A. Schumm, M. P. Mosley, and W. E. Weaver, *Experimental Fluvial Geomorphology*, 1987; M. M. Smart et al. (eds.), *Ecological Perspectives of the Upper Mississippi River*, 1986; B. A. Whitton (ed.), *Ecology of European Rivers*, 1989.

River engineering

River engineering involves the control and utilization of rivers for the benefit of humankind. Its scope, in the broad sense, may include river training, channel design, flood control, water supply, navigation improvement, hydraulic structure design, hazard mitigation, and environmental enhancement. River engineering is also necessary to provide protection against floods and other river disasters.

Throughout recorded history, rivers have been a subject of study by engineers and scientists who have been fascinated by the geometric shapes that rivers form and their responses to natural and human interference. River engineering is somewhat different from other aspects of civil engineering because the emphasis is often on river responses, long-term and short-term, to changes in nature, and stabilization and utilization, such as damming, channelization, diversion, bridge construction, and sand or gravel mining. Evaluation of river responses is essential at the conceptual, planning, and design phases of a project. The use of fundamental principles of river and sedimentation engineering is required in the analysis of special conditions for each river project. *See* CANAL; DAM; RIVER.

River training. This, in a broad sense, covers all the engineering works to regulate river flow and sediment transport for the sake of flood control, navigation, irrigation, and channel stabilization. River training by embankments can be traced to early human history, exemplified by the works on the Nile, Yellow, Euphrates, Tigris, and Ganges rivers. Such practice continues in modern times. Rivers are often trained because of urbanization; land reclamation from flood plains has resulted in the channelization of many rivers in different parts of the world. *See* FLOODPLAIN; STREAM TRANSPORT AND DEPOSITION.

The principle of river training is twofold: (1) The training works must be designed to be strong enough to withstand the design flow, and (2) the impacts on the river should be understood and evaluated whenever feasible. The commonly used types of training works can be broadly classified as bank protection, dikes, and grade-control structures.

Bank protection. Different types of bank protection are in use, including rock riprap, mattress, gabion, soil cement, concrete blocks, cribs, Armorflex (**Fig. 1**), and so on. Because of environmental concerns, materials that also allow vegetation growth are preferred (**Fig. 2**). Soil cement blocks are used as bank protection as a suitable and inexpensive alternative to rock riprap. They are used mainly in places where riprap is not readily available at a reasonable cost. The adequate cement content was

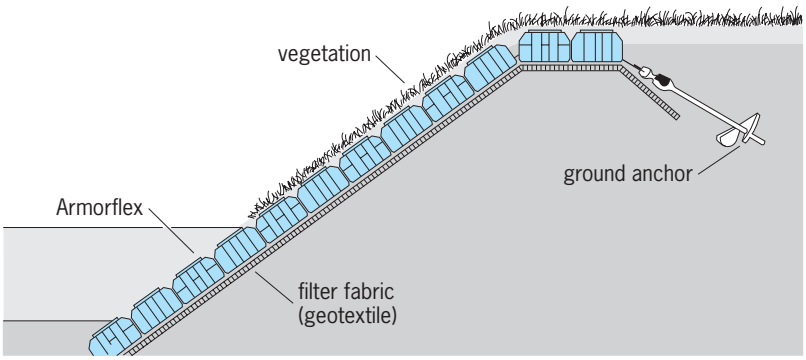

Fig. 1. Armorflex mats made of concrete blocks connected by nylon cables are used for bank protection. (*Armortec Concrete Erosion Control Systems*)

determined to range 9–10% by weight of the soil. The soil-cement bank protection is popular in the semiarid southwestern United States, where the construction is facilitated by the usually dry river bed. Such bank protection is also desirable from the environmental viewpoint because it preserves the soil appearance. *See* GEOSYNTHETIC.

Dikes. These are training structures that extend from the bank to the river at an angle, or perpendicular, to the flow. Most dikes are aligned in a slightly downstream direction. They are also known as groins, spurs, spur dikes, or transverse dikes and represent a very widely used type of training works. Dikes are often used to form a system covering a certain river reach. Dikes serve one or more of the following functions: (1) training a river along a desired course, (2) creating a region of low velocity to induce siltation, (3) protecting the bank by keeping the flow away, and (4) contracting a wide river channel usually for the improvement of depth for navigation. There are two principal types of dikes: permeable and impermeable. Permeable dikes, such as timber pile dikes and jetty fields, permit flow through the dikes at reduced velocities, thereby preventing bank erosion and causing deposition of suspended sediment from the flow. Experience has shown that permeable dikes are more effective than solid ones as a bank protection, especially in silt and sand rivers. Jetties, or steel jacks, consist of angles fastened together and strung with wire. The jetty field is typically used in shallow streams to train the stream into a single narrower channel; it reduces the velocity near the bank and hence protects the bank from erosion. The jetty field is more effective in streams that have a considerable amount of debris and a high concentration of suspended load.

Grade-control structures. These are also called drop structures, stabilizers, weirs, barrages, or check dams, and are generally constructed normal to the channel flow and traverse the channel bed (Fig. 2). Grade-control structures are used in river channels to maintain a slope flatter than the slope of the terrain. Stabilizers refer to sediment control structures that are used primarily to stabilize the upstream channel bed where scour may endanger certain structures such as bridge foundations. The crest of a structure usually extends across the channel, and the side walls should extend into the bank and have adequate bank protection to prevent flanking at high flows. Each structure should also have adequate upstream and downstream protection. Dumped stones should be placed on the downstream side to the anticipated scour depth. While a grade-control structure stabilizes the upstream channel bed, it usually induces downstream changes. When major gradation change is expected, a series of grade-control structures may be used to limit the extent of change at each structure.

River responses. Alluvial rivers are self-regulatory in that they adjust their characteristics in response to any change in the environment. These environmental changes may occur naturally, as in the case of climatic variation or changes in vegetative cover, or may be a result of such human activities as river training, damming, diversion, sand and gravel mining, channelization, bank protection, and bridge and highway construction. Such responses usually occur at places where the bed is mobile (**Fig. 3**) or where the bank is

Fig. 2. Channelization of the Santa Gertrudis Creek in California showing bank protection and grade control structure traversing the channel bed.

unprotected. If a river reach is straightened and channelized, its freedom for meandering is constrained. In response to this change in slope due to straightening, changes in the adjacent natural reaches may occur in order to restore the equilibrium of the river system. In that process, the river adjusts to the new conditions by changing its slope, roughness, bed-material size, cross-sectional shape, or meandering pattern. Within the existing constraints, any one or a combination of these characteristics may adjust as the river seeks to maintain the balance between its ability to transport sediment and the load provided.

Any plan for river training needs to be evaluated with regard to channel responses. In fact, channel responses should be the criteria for determining the adequacy of a design scheme. Modifications in design are often necessary in arriving at the final plan. In modern river engineering, a very challenging task is the prediction of river channel responses. Studies of river hydraulics, sediment transport, and river channel changes may be conducted through physical modeling or mathematical modeling, or both. Traditionally, physical modeling has been relied upon to obtain the essential design information. It nevertheless often involves large expenditures and is time-consuming in model construction and experimentation. What limits the accuracy of physical modeling is the scale distortion, which is almost unavoidable whenever it involves sedimentation. Mathematical modeling of erodible channels has been advanced with the progress in the physics of fluvial processes and computer techniques. Since the actual size of a river is employed in mathematical modeling, there is no scale distortion. The applicability and accuracy of a model depends on the physical foundation and numerical techniques employed. There exist several mathematical models that simulate river flow, sediment transport, and river channel changes, such as the HEC-6 model by the U.S. Army Corps of Engineers and the FLUVIAL-12 model by Chang. The computer-aided approach is illustrated by case studies for two types of problems: general scour at bridge crossings, and fluvial design of river bank protection. *See* COMPUTER-AIDED ENGINEERING; MATHEMATICAL GEOGRAPHY.

Bridge scour. Since a bridge may be damaged by river bed scour (Fig. 3), the scour potential should be determined in bridge design or restoration. River bed scour at bridge crossings consists of local and general scour. Local scour that occurs around bridge piers and abutments is caused by the local obstruction to flow. General scour develops when more sediment is removed from an area than is supplied from upstream. For example, the proposed Rancho Santa Fe Road Bridge on San Marcos Creek in California is located a few miles below a dam. The deficit in sediment supply caused by the dam will induce general scour at the bridge crossing. General scour at bridge crossings is related to the flow and sediment-transport processes of the adjacent river as a system; therefore, evaluation of such scour requires modeling of a river reach.

Fig. 3. River bed scour at a bridge crossing induced by sand mining in the river channel.

Channelization. The major trend at present is to stay away from totally lined channels for the sake of preserving ground-water recharge, wildlife habitat, and so on. Channelization is therefore accomplished by constructing bank protection while leaving an alluvial bed (Fig. 2). The bank protection must be designed to withstand the design flood, and it must be entrenched beyond the potential channel-bed scour. Therefore, an important task in channel design is to arrive at a hydraulic geometry which will minimize potential channel-bed changes. The water-surface profile and bed topography during the passage of the design flood must also be determined for the design of bank height and toe-down (entrenchment) depth. The design for channelization is exemplified by the Santa Cruz River. The river reach in Tucson, Arizona, underwent dramatic changes during the October 1983 flood, which reached the 100-year magnitude at many locations. River channel changes resulted in severe damages. For the purposes of flood control and hazard mitigation, Pima County took the responsibility of channelizing several reaches of the river. Channelization of the Santa Cruz River was designed on the basis of the pattern of river channel responses simulated using the FLUVIAL-12 model. Scour and fill of the channel bed related to the longitudinal imbalance in sediment load were simulated; they were also tied in with the effects of the spiral motion that induces an uneven bed topography in channel bends, characterized by greater scour near the concave bank. Variable top and toe elevations for hydraulic structures were determined and employed to make more efficient use of construction materials while providing effective protection against erosion.

Howard H. Chang

Bibliography. H. H. Chang, *Fluvial Processes in River Engineering*, Wiley, 1988; H. H. Chang and Z. Osmolski, Computer-aided design for channelization, *J. Hydraulic Eng.*, 114(11):1377–1389, 1988;

M. S. Petersen, *River Engineering*, Prentice Hall, Englewood Cliffs, NJ, 1986; C. R. Thorne, R. D. Hey, and M. D. Newson (eds.), *Applied Fluvial Geomorphology for River Engineering and Management*, 1998.

River tides

Tides that occur in rivers with an open connection to a tidal estuary or ocean. These tides are highly modified from coastal ocean tides as they propagate landward into shallow water, often through a brackish estuary and then into a fresh-water river. Ocean tides form in response to the gravitational forces exerted by the Sun and Moon acting on the water masses on the surface of the rotating Earth. Within estuaries or rivers, water masses are small enough that tidal motions are not generated directly from lunar and solar forces; rather, the motions occur because of tide-generating forces at the estuary entrance or river mouth. However, variations in the tide-generating forces in a coastal ocean produce corresponding variations in tides within the upland estuary/river system. In addition to the well-known diurnal or semidiurnal occurrences of high and low waters, numerous other important cycles occur that are necessary to describe tidal oscillations in estuaries and tidal rivers. Tides have their greatest amplitude when the Sun, Moon, and Earth are in alignment during full and new moon and the tidal forces of the Moon and Sun reinforce each other. These tides are referred to as spring tides. Alternatively, neap tides occur when the moon is in quadrature (the solar and lunar forces oppose rather than reinforce each other) and tidal amplitudes are much smaller; that is, higher low tides and lower high tides occur during neap periods. In addition to the spring/neap cycle that has an approximately fifteen-day period, many other astronomical cycles occur. Two of the more important cycles are the ones responsible for diurnal inequalities between the heights of the two high and two low waters experienced daily in some locations and the effect on tidal amplitude that results from the moon's varying distance from earth. *See* ESTUARINE OCEANOGRAPHY; RIVER; TIDE.

Tidal range. Tide-generated waves are known as "long" or "shallow water" waves because of their long wavelengths and typically small-wave amplitudes relative to water depths. Their wave speed, given approximately by the time of travel of a long progressive wave, is defined as $\sqrt{gD}$, where $g = 9.81$ m/s^2 (32.2 ft/s^2) is gravitational acceleration and D is water depth. Thus, wave speed is affected by local bottom topography, which can significantly modify incoming waves in shallow bays, estuaries, and rivers. Shoaling within estuaries and rivers slows progress of the tidal wave; the speed/depth relation creates a condition where the tide crest moves faster than the trough causing the wave front to steepen. This results in a situation where the duration of the rising tide is generally less than that of the falling tide and the slope of the tide curve becomes more asymmetrical as it propagates further inland. Under certain conditions usually associated with large tide range, shallow depth, and sharp increase of slope of riverbed, the distortion of the tidal curve between low water and high water becomes sufficiently large that a discontinuity of water level occurs. The condition is referred to as a tidal bore. Most tidal bores are small, possibly only a few centimeters (0.1 ft), but extreme bores may reach heights of 3 m (10 ft) or more and travel at speeds of 25 km/h (15.5 mi/h) during spring tides. *See* TIDAL BORE.

Whereas tidal bores are often spectacular, they are generally rare. Typically, tidal range increases as the width of the estuary or river narrows. Both high- and low-water levels rise with bed level, and tidal range begins to decrease. River tides propagate upstream until friction dissipates the tidal energy or some other physical impediment is encountered. Each estuary and river system has different characteristics because of its unique topography and geometry. As an example, 36 h of predicted tide heights starting January 1, 2005, at four locations within the San Francisco Estuary and Sacramento River system in California are shown (see **illustration**). Delay in arrival time of the tide and general damping of tide amplitude are observed as the tide propagates from the bay entrance at Golden Gate near San Francisco to an upstream location at Walnut Grove on the Sacramento River south of Sacramento. Tide tables indicate that by the time the tide reaches Rio Vista at 100 km (62 mi) up river, the time of high water occurs about 4 h later and the time of low water is delayed more than 5 h. In addition, the range of tide is reduced; high waters are reduced by about 20% and low waters are only about half as low as the corresponding values of high and low waters at Golden Gate.

While the effects of astronomical tides on estuary/river systems are periodic and predictable, other nonperiodic factors can greatly modify the astronomical tides in these upland systems. Meteorological influences such as variations in air pressure and wind and hydrologic influences such as river inflows, including the secondary influence that river flows have on salinity gradients within the estuary, can affect river tides. Air pressure changes, often referred to as the inverse barometer effect, create a rise in water level of about 1 cm (0.4 in.) for every 1 millibar (0.1 kPa) drop in pressure. Sustained winds may cause water to "pile up" along a coastline or at the downwind end of open water bodies, such as estuaries, creating changes in water level. Finally, runoff from rain or snowmelt may cause river flows to rise and lead to an increase in water level heights. River flows can retard the progress of rising tides, while enhancing falling tides.

Tidal currents. Tidal currents in rivers are the horizontal motions of water produced by the river tides. In the lower reaches of an estuary or river, tidal currents are usually bidirectional but become unidirectional (but still oscillatory) further upstream. In a river, tidal current reversals may occur far upstream under conditions of large tidal amplitudes and low

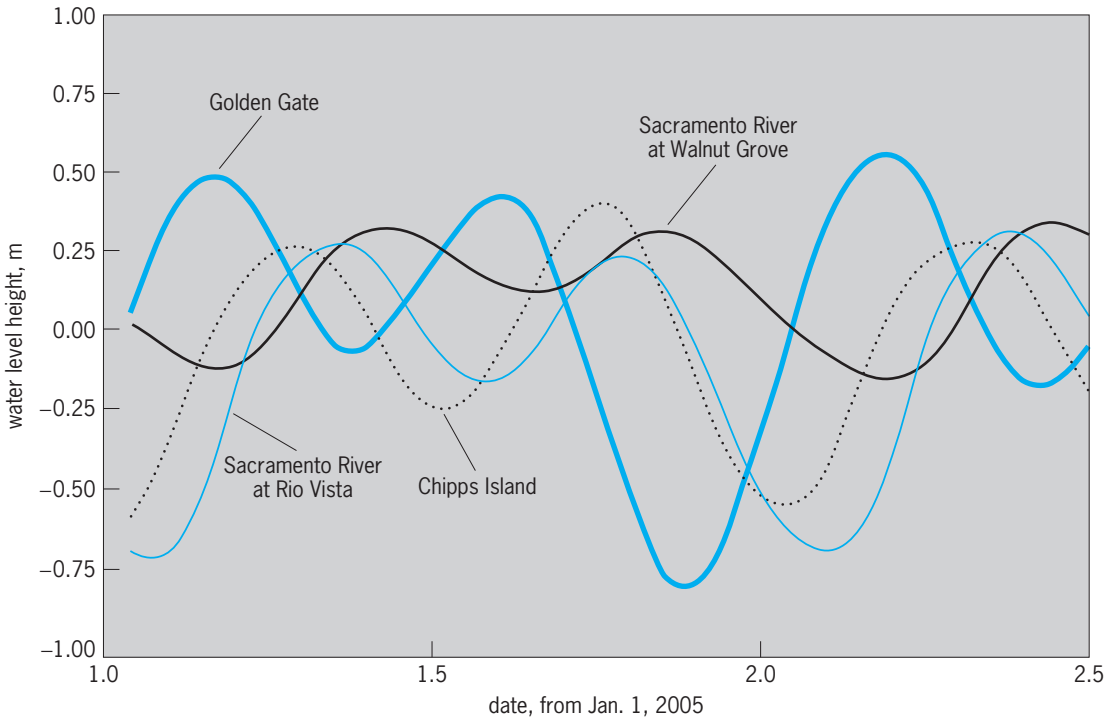

Time series of predicted water level heights referenced to mean water level at four locations in the San Francisco Bay and Sacramento River system. Golden Gate is at the entrance to bay (km = 0); Chipps Island is 72 km (45 mi) up-estuary; Sacramento River at Rio Vista is 100 km (62 mi) from Golden Gate; and Sacramento River at Walnut Grove is 130 km (81 mi) from Golden Gate.

river flows. Tidal currents are highly modified by geometry and bathymetry, with stronger currents occurring in deep, narrow channels and weaker currents occurring in wide, shallow areas. Upstream-directed currents that occur on rising tides are referred to as flood currents and downstream-directed currents that occur on falling tides are referred to as ebb currents. Slack water occurs at times when the flow changes direction and there is little or no water motion.

Times of maximum and minimum currents and their relation to high and low tides are related to how the tide wave progresses within the estuary/river system. The tide is a progressive wave that advances with respect to a fixed point of observation. Tides and tidal currents are in phase in a pure progressive wave, which means that maximum flood currents occur at high waters and maximum ebb currents occur at low waters. Slack waters occur at midtide levels. In contrast to a progressive wave, a standing wave appears stationary with respect to a fixed point of observation. Standing waves occur when two progressive waves of equal amplitude and period traveling in opposite directions are superimposed. Standing waves can occur in estuaries with closed ends or in tidal rivers that are blocked by a barrier such as a dam or falls from which the incident tide wave is reflected. A pure standing wave is characterized by tides and tidal currents that are 90° out of phase; maximum currents occur at midtide level and slack waters occur at times of high and low water. Reflections of the incident tide wave from channel sidewalls can result in a wave progressing down the estuary or river. This ordinarily leads to tides that propagate in some combination of progressive and standing waves. Under this type of combined motion, the time delay as the tide progresses landward is generally less than for a pure progressive wave. Where rivers flow into estuaries, tides are generally more progressive than standing and are subject to decreases in amplitude from attenuation by bottom friction and channel constrictions, as well as attenuation from reflection from sides of channels. Jeffrey W. Gartner

Bibliography. A. Defant, *Ebb and Flow: the Tides of Earth, Air, and Water*, 1958, University of Michigan Press, 1958; C. B. Officer, *Physical Oceanography of Estuaries (and Associated Coastal Waters)*, Wiley, 1976; D. T. Pugh, *Tides, Surges and Mean Sea-level, a Handbook for Engineers and Scientists*, Wiley, 1987.

Rivet

A short rod with a head formed on one end. A rivet is inserted through aligned holes in two or more parts to be joined; then by pressing the protruding end, a second head is formed to hold the parts together permanently. The first head is called the manufactured head and the second one the point. In forming the point, a hold-on or dolly bar is used to back up the manufactured head and the rivet is driven, preferably by a machine riveter. For high-grade work such as boiler-joint riveting, the rivet holes are drilled and reamed to size, and the rivet is driven to fill the hole completely. Structural riveting uses punched holes.

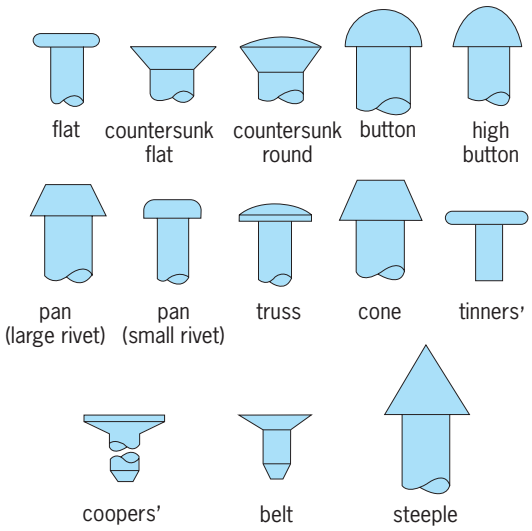

Fig. 1. Standard rivet heads.

Small rivets 7/16 in. or 11 mm and under) are used for general-purpose work with head forms as follows: flat, countersunk, button, pan, and truss (**Fig. 1**). These rivets are commonly made of rivet steel, although aluminum and copper are used for some applications. The fillet under the head may be up to 1/32 in. or 0.8 mm in radius.

Large rivets (1/2 in. or 13 mm and over) are used for structural work and in boiler and ship construction with heads as follows: roundtop countersunk, button (most common), high button or acorn, pan, cone (truncated), and flattop countersunk.

Boiler rivets have heads similar to large rivets with steeple (conical) added but have different proportions from large rivet heads in some cases.

Special-purpose rivets are tinners' rivets, which have flatheads for use in sheet-metal work; coopers' rivets, that are used for riveting hoops for barrels, casks, and kegs; and belt rivets, used for joining belt ends.

Blind rivets are special rivets that can be set without access to the point. They are available in many designs but are of three general types: screw, mandrel, and explosive (**Fig. 2**). In the mandrel type the rivet is set as the mandrel is pulled through. In the explosive type an explosive charge in the point is set off by a special hot iron; the explosion expands the point and sets the rivet.

Standard material for rivets is open-hearth steel (containing Mn, P, S) with tensile strength 45,000–55,000 lb/in.^{2} or 310–380 megapascals. Standards include acceptance tests for cold and hot ductility and hardness. Materials for some special-purpose rivets are aluminum and copper. Paul H. Black

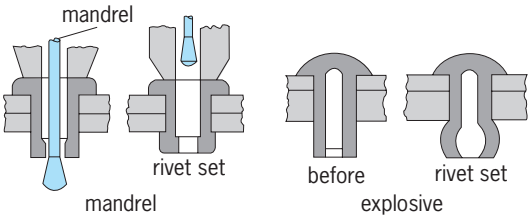

Fig. 2. Two types of blind rivet.

Bibliography. American National Standards Institute (ANSI), *Semitubular Rivets, Full Tubular Rivets, Split Rivets and Rivet Caps*, B18.7–1972; ANSI, *Small Solid Rivets 7/16 Inch Nominal Diameter and Smaller*, B18.1.1–1972; E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; Institute for Power System, *Handbook of Industrial Fasteners*, 3d ed., 1997.

Robotics

A field of engineering concerned with the development and application of robots, and computer systems for their control, sensory feedback, and information processing. There are many types of robotic systems, including robotic manipulators, robotic hands, mobile robots, walking robots, aids for disabled persons, telerobots, and microelectromechanical systems.

The term "robotics" has been broadly interpreted. It includes research and engineering activities involving the design and development of robotic systems. Planning for the use of industrial robots in manufacturing or evaluation of the economic impact of robotic automation can also be viewed as robotics. This breadth of usage arises from the interdisciplinary nature of robotics, a field involving mechanisms, computers, control systems, actuators, and software. *See* BIOMECHANICS; COMPUTER; CONTROL SYSTEMS; CYBERNETICS; ELECTRICAL ENGINEERING; INDUSTRIAL ENGINEERING; MECHANICAL ENGINEERING; SOFTWARE ENGINEERING.

Mechanisms. Robots produce mechanical motion that, in most cases, results in manipulation or locomotion. For example, industrial robots (**Fig. 1**) manipulate parts or tools to perform manufacturing tasks such as material handling, welding, spray painting, or assembly; automated guided vehicles are used for transporting materials in factories and warehouses. Telerobotic mechanisms provide astronauts with large manipulators for remotely performing spacecraft maintenance. Walking machines explore active volcanoes. Mechanical characteristics for robotic mechanisms include degrees of freedom of movement, size and shape of the operating space, stiffness and strength of the structure, lifting capacity, velocity, and acceleration under load. Performance measures include repeatability and accuracy of positioning, speed, and freedom from vibration.

Manipulator geometries. An important design for robotic manipulators is the articulated arm, a configuration with rotary joints that resembles a human arm (Fig. 1*a*). An extension of this idea is the selective compliance assembly robot arm (SCARA) configuration (Fig. 1*b*), which has been extensively applied to assembly in manufacturing. SCARA manipulators have high stiffness during vertical motions, but they

Fig. 1. Industrial robots. (*a*) Articulated, with rotary joints (*GM Fanuc Robotics Corp.*). (*b*) Selective compliance assembly robotn arm (SCARA) configuration (*Adept Technology, Inc.*) (*c*) Overthead gantry.

may be fitted with mechanical and sensory control devices that allow freedom in lateral motions during assembly operations.

Cylindrical configurations having a rotary joint at the base and a prismatic joint at the shoulder are useful for simple material transfer and assembly. One of the earliest manipulators, marketed in the late 1950s, was based on a configuration that had two rotary axes at the shoulder; this configuration has disappeared from commercial application. Overhead gantry (cartesian geometry) robots are used for high-payload applications and those requiring linear positioning over large work spaces (Fig. 1*c*).

Degrees of freedom. Some manipulators have simple mechanical configurations involving only two or three degrees of freedom of movement. Most robotic manipulators have six degrees of freedom so that they can position a part or approach a part with any desired position and orientation. The wrist is positioned at a desired x,y,z position in the work space. Then, the end effector is rotated to a desired orientation (roll, pitch, yaw). In effect, the wrist represents the origin of a three-axis coordinate system fixed to the gripper. Moving the first three joints of the arm translates this origin to any point in a three-axis coordinate system fixed to the work space; motion of the final three joints (in the wrist) orients the gripper coordinate system in rotation about an origin at the wrist point. While not all robots are constructed this way, the mathematics of mechanical motion (kinematics) is considerably simplified when these axes of the wrist do not intersect. *See* DEGREE OF FREEDOM (MECHANICS); KINEMATICS.

Types of joints. Robotic mechanisms can have joints that are prismatic (linear motion), articulated (rotary motion), or a combination of both types. While many robots use only articulated joints to imitate the human arm, limited actions can be produced by using prismatic joints alone. Robot locomotion can also be obtained by the use of wheels and treads. Walking robots have been developed that use articulated legs. *See* MECHANISM.

Joint actuators. Actuators for moving joints of a robotic mechanism are usually electric motors. In the past, larger robots were built with hydraulic actuation due to their high payload capacity and ability to work in explosion-prone environments. In the 1990s there was a trend by the major manufacturers to use electric actuators in most of their industrial robots.

Actuators can be placed directly at the joints; however, the weight and bulk of these motors and associated gear transmissions limit the performance of the robot, particularly at the wrist of industrial robots. Another design involves placement of the actuators in the base of the robot and transmission of motion to the joints through mechanical linkages such as shafts, belts, cables, or gears. This approach overcomes many of the problems associated with locating actuators at the joints, but requires the design of backlash-free mechanical linkages that can transmit power effectively through the arm in all of its positions and orientations. An innovation in the design of assembly robots is the use of direct-drive motors, which eliminate the need for gear transmissions. This type of robot has become popular for high-speed assembly tasks. *See* BELT DRIVE; CHAIN DRIVE; GEAR; LINKAGE (MECHANISM).

End effectors. A two-jaw parallel gripper actuated by a pneumatic cylinder is a common type of end effector for industrial robots. End effectors are specialized for particular applications, such as parts handling, welding, and drilling. Using a quick-disconnect mechanism, robots can change their own end effectors to suit the task from a selection of special-purpose attachments. Since position errors of robots, fixtures, and grippers may occur during high-precision assembly tasks, compliance is needed to compensate for misalignment between parts during operations such as inserting a piston into an

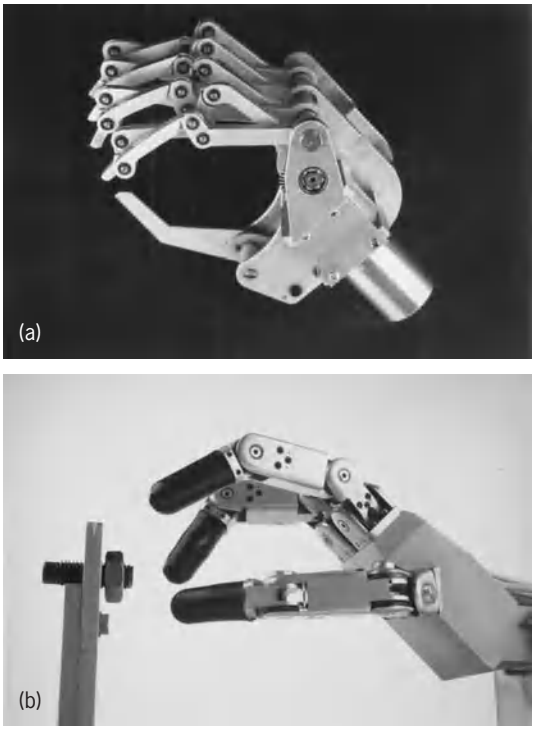

Fig. 2. Robotic hands. (*a*) Five-finger mechanical hand actuated by a single direct-current motor. (*b*) Three-finger, nine-degree-of-freedom dextrous mechanical hand. (*Intelligent Robotics Corp.*)

engine head or fitting a gland on a valve assembly. A spring-actuated mechanism, called a remote center compliance (RCC) wrist, may be attached between the robot arm and the end effector for this purpose.

Robotic hands. Dexterity is as basic to manufacturing assembly as it is to prosthetic and human hands. Multifingered robot hands can accomplish fine manipulation, achieve robotic hand cooperation for fixtureless assembly, operate in hazardous environments, explore remote locations, and perform other important functions. The usefulness of multifingered robot hands to perform fine manipulation and grasping tasks is widely acknowledged, but it is costly to implement. Mechanisms for robotic hands are commercially available with two to four fingers and a thumb, each with one to four degrees of freedom and actuated by servomotors (**Fig. 2**).

Mobile robots. Mobile robots can have wheels, tracks, legs, or a combination of these parts. They also can crawl, for example, for inspecting pipes. Tracked mobile robots developed for military and surveillance applications are among the most stable and rugged. Automated guided vehicles (AGV) are wheeled mobile robots that deliver parts and tools in factories by following a guide path embedded in the floor. The payload capacity of these vehicles varies from 220 lb (100 kg) to several tons. An AGV for light-duty work such as mail delivery in office buildings can follow a stripe attached to the floor. An AGV can be programmed like an industrial robot to follow a preprogrammed trajectory and to make decisions along the way depending on inputs from external signals. The path of an AGV can be guided by a wire underneath the floor or a stripe painted on the surface. Free-ranging AGVs are also available for industrial and commercial use to avoid the need for installing a fixed path.

Walking machines. Walking robots have been built with two, four, and six legs. One company marketed a four-legged walking machine that could walk over obstacles, lift up to five times its weight while stationary, and lift two times its weight while moving (**Fig. 3**). The control of gait stability for legged locomotion systems is a difficult implementation issue, particularly for two-legged (biped) machines. Research in locomotion has provided results on kinematic and dynamic modeling, control algorithms, and other forms of stability. Solutions to these research issues have resulted in a better understanding of high-level planning, control requirements, and their implementation.

Robot control and planning. A robot control system directs the motion and sensory processing of a robot or system of cooperating robots. The controller may consist of only a sequencing device for simple robots, although most multiaxis industrial robots today employ servo-controlled positioning of their joints by a microprocessor-based system.

Servo systems. A measure of the actual joint position is obtained from a transducer, such as an optical shaft encoder, and this signal is compared with the position specified for the joint. If the desired position (called the set point) and the measured position differ, the circuitry applies a correcting drive signal through the digital-to-analog converter (DAC) to the power amplifier of the joint motor. Such servo systems employ analog and digital circuitry; in addition

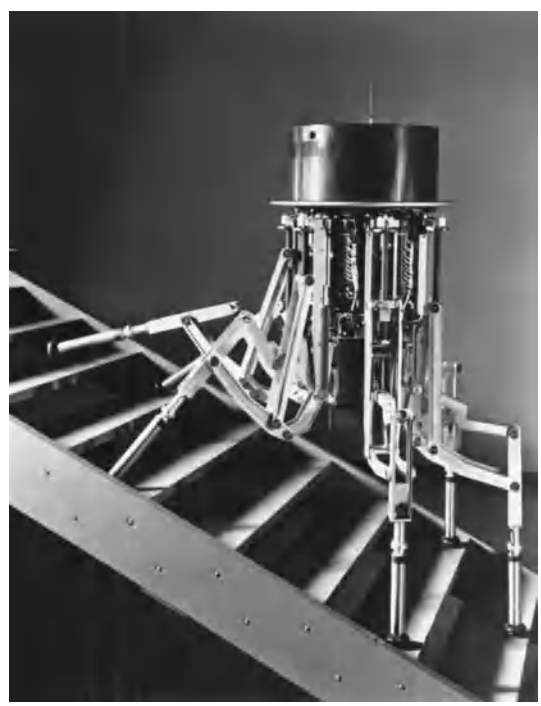

Fig. 3. Six-legged walking robot. (*Odetics*)

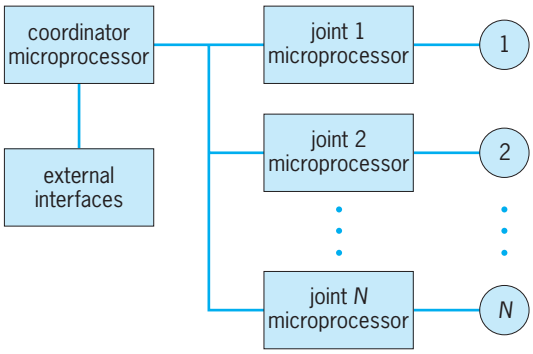

Fig. 4. Joint-parallel controller.

to position, they usually control joint velocity. *See* DIGITAL-TO-ANALOG CONVERTER.

Servo systems allow the robot to be moved through any selected sequence of positions without the need to preset mechanical stops. The sequence of positions defining the robot's trajectory may be preprogrammed by numerically specifying the positions (and orientations) in the robot's coordinate frame, or they may be taught to the system by moving the robot to the desired point and recording the outputs of the joint position transducers. Such teaching methods allow the user to direct a complex action once, and to have the robot repeat it under preprogrammed computer control as long as the motion is required. *See* SERVOMECHANISM; TRANSDUCER.

Robot controllers. Because of the computing power and low cost of microprocessors, most robot controllers are computer-based systems, often with individual microprocessors for each joint (**Fig. 4**). The controller generates the robot's trajectory based on mathematical representation of the work space, work-space objects, and tasks stored in the computer's database. The computer may also generate trajectories for the robot that are not fixed but vary with the state of the external world as monitored by the robot's sensors. This type of sensory-interactive control permits the robot to act appropriately in relation to conditions, rather than relying on assumptions about the world. For example, without custom fixtures and timing of the workflow, the actual location of a part may vary from one instance to the next; without sensory-interactive control in such cases, the robot could proceed blindly through a set of actions at preprogrammed but incorrect positions. *See* COMPUTER PROGRAMMING; DIGITAL CONTROL; MULTIPROCESSING.

When separate microprocessors are used for control of each joint of the robot, coordination of these axes can be obtained by an upper-level processor that also generates the path for each joint and converts a description of the end-effector position to joint angles (**Fig. 5**). The robot controller receives inputs from a teach pendent that records joint positions of the desired path. An input/output (I/O) module monitors on-off signals from sensors, such as proximity switches, indicating part presence prior to gripping. Programs are stored on hard or floppy disk drives. Another function of the upper-level control is to interpolate the path so that the robot moves smoothly. This requirement, called continuous-path control, is needed for applications such as arc welding and spray painting. *See* CONTROL SYSTEMS; MICROPROCESSOR.

Mobile robot navigation. Navigation of mobile robots requires constructing a map of where a robot has been, a plan to tell it where to go, and a controller to drive it there. As in the case of manipulator-motion planning, there is a need for the automatic creation of plans or programs for guiding a mobile robot to its desired state. Path planners for mobile robots find a path through a mapped environment so that the robot can travel along it without colliding with another object. Methods designed for expert systems have been used to combine the vast amount of poorly structured knowledge obtained for control. *See* EXPERT SYSTEMS.

Robot path planning. Automatic path generation algorithms have been developed to enable robots to avoid obstacles, such as those encountered in a cluttered environment, or to coordinate the actions of multiple robots to lift a heavy object. One category of path-planning algorithms is based on an explicit representation of free space, subsets of robot configurations that are free of collisions. Such an approach will find a path if it exists, within the known subset of free space. Unfortunately, it is difficult to find a representation for a multidimensional path-planning problem that can be solved. In cartesian space, a manipulator is easy to visualize, but its description is very complex. Regions of free space are easy to find and model, but planning through them can be very difficult. A configuration of a moving object is a set of parameters that completely specifies every point on the object. Configuration space (C-space) is a graphical representation that helps the systems designer to visualize the motion of robots. The C-space for a robot manipulator is a rectangular coordinate space in which each axis represents one joint angle. In C-space, path planning is reduced to planning the

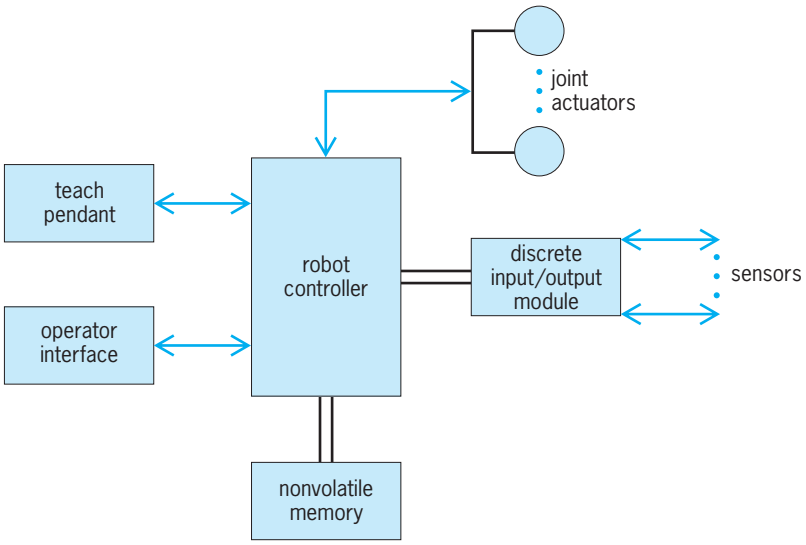

Fig. 5. Robot controller architecture.

path of a point through intersecting configuration space volumes. Computational algorithms for solving C-space plans have been based on "A* -Search," a graph search algorithm for selecting the best path from many possible paths. Other techniques based on nonlinear programming and heuristics have also been developed. *See* ALGORITHM; NONLINEAR PROGRAMMING.

Sensory systems. The robot sensory system gathers specific information needed by the control system and, in more advanced systems, maintains an internal model of the environment to enable prediction and decision making. The joint position transducers on industrial robots provide a minimal sensory system for many industrial applications, but other sensors are needed to gather data about the external environment. Sensors may detect position, velocity, acceleration, visual, proximity, acoustic, force-torque, tactile, thermal, and radiation data.

Tactile sensors. Tactile sensors may be mounted on the robot gripper to detect contact with objects in the work space. These sensors may be simple mechanical switches or transducers indicating the degree and direction of force. Arrays of such transducers may be used to discriminate types and orientations of objects. Force-torque sensors, mounted in the robot's wrist or end effector, are often used to sense the degree and direction of resistance due to the object being handled. These resistive forces may be due to the weight of the object being manipulated, or contact with other objects or surfaces. Such sensors are used to adjust gripping pressure, avoid applying destructive forces, and guide proper mating of surfaces and parts. In combination, these senses allow a robot to feel the proper fit of parts as does a human worker. *See* PRESSURE TRANSDUCER.

Robot vision. An inexpensive means of noncontact sensing for robots is to use a photoelectric diode that provides on-off signals such as "gripper open" or "part missing." These devices can also determine proximity. In robot vision, solid-state video cameras are used to capture an image such as a part to be grasped or a distance to be measured. The image can be processed by computer analysis using pattern recognition techniques, and can be used to evaluate grasp positions and automatically determine flaws based on comparison with stored images in the computer. Computer vision systems may use ambient light or structured light sources. Systems with ambient light sources rely on normal sources of illumination, while structured light sources provide special patterns of illumination whose shape and orientation are known to the sensory system. The advantage of structured light is that special patterns of illumination may be chosen to simplify and speed up the processing needed to interpret the image. Procedures for determining depth in images rely fundamentally on geometric calculations associated with triangulation procedures, but in structured light systems the triangulation is between the camera and the light source. In ambient light systems, corresponding points in two images taken from different viewing positions must be triangulated, and this is much more difficult to achieve. Speed is important in robot vision, because visual information can be used in real time by the control system to correct the robot's movements, a process called visual servoing. Sensor fusion allows the combining of independent and qualitatively different sensor data (for example, from vision and force sensing) to reinforce consistent interpretations, and disregard inconsistent interpretations. Research in this area has led to methods for extracting three-dimensional shapes from two-dimensional images; such methods require dealing with uncertain and imprecise data. *See* COMPUTER VISION.

Multisensor systems. The robot's world model can be continually updated on the basis of new sensory inputs, and can also generate expectations about how the view will transform under movements in progress. The world model's hypotheses in turn assist the interpretation of incoming sensory data. The most advanced robot sensory systems process information from many different sensor modalities, and are also structured to generate successive levels of description, at increasing degrees of complexity and decreasing rates, suited to use by the successive levels of the control system (**Fig. 6**). At the lowest level are the motor controllers and processes that describe and use feedback from sensor signals. These processes occur with cycle times on the order of 10–100 milliseconds. At the next level, perceptual information is brought together and overall control is computed to direct the robot. These processes occur at 100 milliseconds to 1 second. The third level in the architecture consolidates perception and locomotion. Cycle times in this level operate at 100 milliseconds for road following to 10 seconds for local path planning. Approaches for planning these actions have been based on preprogrammed algorithmic methods, while other methods have employed artificial intelligence rule-based decision making. At the highest level, the supervisory computer operates at control cycles of 1–100 seconds. *See* ARTIFICIAL INTELLIGENCE.

Intelligent systems. As information moves up from the device, the amount of information increases and the speed of data acquisition decreases. These control architectures form the basis for computer integrated manufacturing (CIM), a hierarchical approach to organizing automated factories. Recently, a new paradigm has emerged, based on the interconnection of intelligent system elements that can learn, reason, and modify their configuration to satisfy overall system requirements. One of the most important of these approaches is based on holonic systems. *See* AUTOMATION; COMPUTER-INTEGRATED MANUFACTURING; INTELLIGENT MACHINE.

Holonic systems. A holon is an intelligent, autonomous, cooperative unit of a system for transforming, transporting, storing, or validating information and physical objects. Holons are capable of processing information and performing physical tasks. A holon is a building block for holonic systems (**Fig. 7**). It can be either a fully self-functioning entity or an integral part of another holon. From an

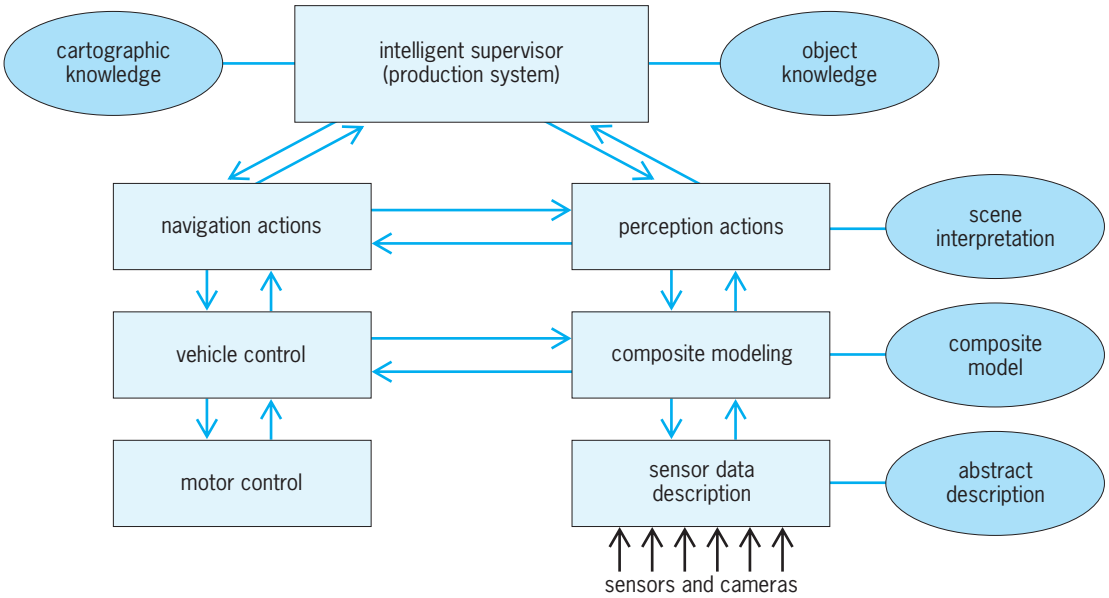

Fig. 6. Typical framework or architecture for mobile robot control. (*After R. C. Dorf, ed., International Encyclopedia of Robotics: Applications and Automation, vol. 3, Wiley, 1988*)

implementation viewpoint, holons are simply system hardware and software components that are modular, reusable, and reconfigurable. Holons may be robots, software packages, machine tools, part feeders, automated guided vehicles, or human operators. Object-oriented programming languages, such as C++, and computer-aided software engineering (CASE) tools provide a development environment to implement holonic systems. *See* OBJECT-ORIENTED PROGRAMMING.

A holarcy is a structure composed of holons to achieve a given objective or to execute a task optimally—for example, in terms of the amount of time taken or resources consumed. It can be distinguished from a hierarchical top-down structure and a group of loosely organized free agents in that actions take place in a cooperative and autonomous manner. Cooperation is obtained through the holons' interaction with the environment, the communication between holons themselves, and the communication from an upper-level controller. Autonomy is achieved through decentralized and distributed intelligence across the overall system, thereby enabling holons to interpret commands as well as sensory feedback to maintain a robust operation and handle unforeseeable disturbances. *See* DISTRIBUTED SYSTEMS (CONTROL SYSTEMS).

With traditional hierarchical architectures, high investments may be needed for system implementation to accommodate the required variety in production lot sizes and product mixes. Although related to multiagent systems, holonic systems exhibit not only autonomy but also a high degree of cooperation. Holonic systems provide a means for designing and integrating robotic and manufacturing systems. They offer the user quick installation and start-up, as well as easy development and operation with flexibility to meet changes in the environment. A holonic manufacturing system exhibits a high degree of strategic flexibility while maintaining a fixed or relatively stable set of rules for its internal organization and functioning. Human experiences and interaction may also be incorporated into holonic systems.

Neural systems. A neural network is an information-processing system that has performance characteristics in common with biological neural systems. This class of systems is useful for implementing difficult control and sensing tasks in robotics without requiring an accurate representation of the system. A neural net consists of a large number of simple processing elements called neurons. Each neuron is connected to other neurons by means of directed communication links, each with an associated weight. The weights represent information being used by the net to solve a problem. Neural nets can be used to classify patterns obtained from a digital camera image and perform general mappings from input patterns to output patterns, such as in fusing information from multiple sensors in map building. They

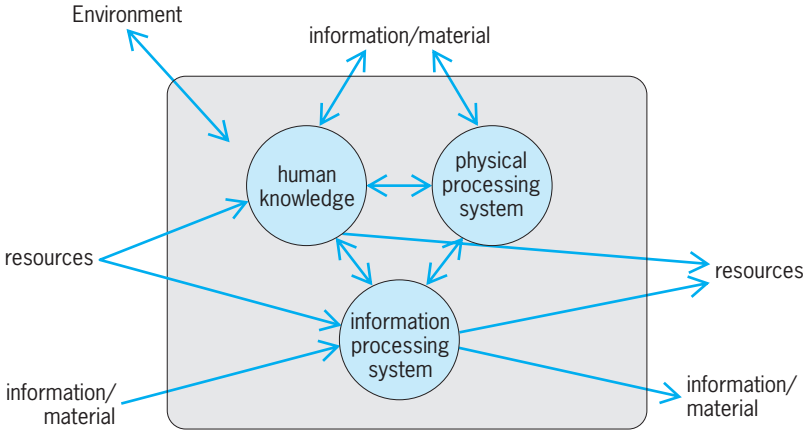

Fig. 7. Holonic system element. (*Holonic Manufacturing Systems Consortium*)

can also be used to find solutions to constrained optimization problems such as those required for robot path planning. *See* NEURAL NETWORK.

Fuzzy systems. Fuzzy sets, fuzzy logic, and fuzzy control of complex systems are characterized by high degree of nonlinearity, strong coupling, and time variance. Fuzzy logic is a superset of boolean logic that has been extended to handle the concept of partial truth that falls between "completely true" and "completely false." A fuzzy system uses knowledge that experts have accumulated over time on topics for which it is very hard to establish precise mathematical models, and expresses this knowledge in explicit "if-then" linguistic rules. The design of a fuzzy system involves the following steps. Fuzzification maps the actual input variables to their fuzzy subsets according to membership functions. Fuzzy inference and reasoning evaluate each fuzzy subset and integrate all the variables to obtain the output fuzzy subset based on fuzzy operators and control rules. Defuzzification is the process of converting fuzzified values into crisp variables to command the physical plants. Fuzzy systems have been applied in many engineering fields, including process control, artificial intelligence, system diagnosis, and pattern recognition, as well as in economics, sociology, management, and medical research. Fuzzy systems provide an important tool for implementing the control of sensory interactive robotic systems. *See* FUZZY SETS AND SYSTEMS.

Human-machine interfaces. A telerobotic system augments humans by allowing them to extend their ability to perform complex tasks in remote locations. It is a technology that couples the human operator's visual, tactile, and other sensory perception functions with a remote manipulator or mobile robot. These systems are useful for performing tasks in environments that are dangerous or not easily accessible for humans. Telerobotic systems are used in nuclear handling, maintenance in space, undersea exploration, and servicing electric transmission lines. Perhaps the most important sensory data needed for telepresence are feedback of visual information, robot position, body motion and forces, as well as tactile information. Master-slave systems have been developed in which, for example, a hand controller provides control inputs to an articulated robotic manipulator. These systems are capable of feeding back forces felt by the robot to actuators on the exoskeletal master controller so that the operator can "feel" the remote environment. *See* HUMAN-MACHINE SYSTEMS; REMOTE MANIPULATORS.

Programming languages. Software development is a significant part of the cost of engineering a robotic system. System architectures and programming environments are generally not able to handle the demands of flexibility, performance, and complexity that are required by modern production processes. For this reason, high-level, textual programming languages are provided for specifying a sequence of actions that allow the operator to program external sensors, such as vision and force, and to use computer-aided design (CAD) systems to program the robot off-line. Programming languages have been developed for use in factory automation as well as robots in unstructured environments. Some of these languages have attained several generations of development, and they now include concepts from knowledge-based systems, model-based systems, discrete-event systems and Petri nets (a graphical method for analyzing and modeling discrete-event dynamic systems), graphical languages, and object-oriented programming languages, as well as related methodologies and tools from database engineering and communication networks. Building on textual languages developed for industrial robots, these programming environments provide a common interface to devices and sensors that would otherwise require specialized programming by people with different levels of skills. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING; COMPUTER PROGRAMMING; PROGRAMMING LANGUAGES.

Telepresence. This technology is a combination of interface technologies that enables users to intuitively interact with immersive and dynamic computer-generated environments. It frequently involves the use of computer-generated visuals to enhance the perception of the physical environment. Such systems, called virtual reality (VR), have cameras at the worksite that are slaved to the operator's head movement. The operator views the scene at the worksite through a binocular display mounted on a helmet. While virtual reality has been used by flight simulators and military displays for many years, the low cost of computers and peripherals, such as data gloves, is helping to make this technology available to new applications of robotics. *See* AIRCRAFT TESTING; COMPUTER GRAPHICS; VIRTUAL REALITY.

Robotic system applications. An early justification for industrial robots was to replace humans in dangerous and repetitive labor, and to replace fixed-function automation with more flexible equipment that could be reprogrammed for new situations and adapted through the use of sensory inputs from the external world. Because of their initial limited sensory abilities, robots were best suited to those tasks with clearly prescribed paths of motion. For these reasons, most applications of robots were found in tasks such as material transfer, machine loading and unloading, spot welding, and spray painting.

The robots employed for these purposes range from small models with lifting capacities of several pounds to very large types that can lift many hundreds of pounds. Overhead gantry robots with three or four axes have been developed for heavy lifting tasks involving payloads up to 1000 lb (450 kg). Designs also vary according to the size and shape of the work envelope. For example, many robots are designed to reach sideways and backwards so that they can achieve higher productivity by tending several machines at the same time. Graphical simulation is used to design and evaluate a workcell layout before it is built. The robot motion can be programmed on the simulation and downloaded to the robot controller. Companies market software systems that include libraries of commercially available robots and

postprocessors for off-line robot programming. *See* SIMULATION.

Automated manufacturing. Advances in robotics stimulated the development of industrial applications that exploited the robot's speed and accuracy in such tasks as precision drilling and other light machining, gluing and fastening, simple assembly, packing and sealing, and testing and measurement. Some of the machines developed for these tasks have repeatability of movement better than 0.001 in. (25 micrometers), yet are able to move at high rates throughout the work envelope. Usually they are not capable of heavy payloads. These robots are often fitted with highly specialized end effectors that are automated tools themselves. Their control systems are more elaborate and frequently include external sensory capabilities. Devices of this type may, for example, place odd-sized components on electronic circuit boards, with very small clearances, after visually locating positioning marks on the circuit board.

Flexible manufacturing requires high levels of sensory interaction and processing. Robots must be capable of locating and positioning three-dimensional parts, mating them to high tolerances, and performing complex joining operations, such as screwing bolts and riveting skins on aircraft wings. Robots with redundant degrees of freedom (more than six for a manipulator) can reach a specific position and orientation in more than one way, for example, as needed for inspection inside the body of a car that is being assembled. Robots may be equipped with sensors that enable them to feel and adaptively respond to forces from sticking and jamming of parts during positioning and mating. Since assembly requires the presence of many parts in the work space, intelligent robots have been developed with visual sensing that can identify and locate parts without the need for complex, inflexible part-presentation equipment such as feeders, positioners, and custom end effectors. Vision capabilities required for assembly range from simple to complex. In many cases, parts have only a few stable positions in which they rest, so that recognition and location of two-dimensional outlines is sufficient to acquire and orient them. *See* FLEXIBLE MANUFACTURING SYSTEM.

Autonomous mobile robots. Simple mobile robots have long been employed for material movement in factories, mail delivery in office buildings, and similar tasks requiring intelligent choice of path and a limited ability to avoid obstacles. Mobile robots for service in the home and office are becoming available. Developments in robot control and sensing, combined with advances in mechanisms for robot locomotion, are expanding the applications for mobile robots. Among these applications are robots for guard and sentry duty, hospital care, and commercial cleaning.

Research and development has resulted in robots for agriculture, house cleaning, construction, mining, fire fighting, rescue, surveillance, handling of hazardous materials encountered in waste cleanup, military applications on land, and undersea exploration. In some cases, this involves the provision of sensory, world-modeling, and control devices for existing machines which already possess mobility and actuators, such as earth-moving equipment. In other cases, entirely new devices and system architectures must be developed. The holonic system paradigm described previously is well suited for these applications because of its ability to adapt to changing environments.

One of the difficulties in using autonomous mobile robots is the complex sensory requirement imposed by navigation in unstructured environments. Control requirements for mobile robots are considerably more difficult than those for industrial robots. Since the control and perception processes in a mobile robot decompose naturally across a set of layers, and may be implemented on different processors, a holonic systems architecture can be used for control. Because the control of autonomous mobile robots requires decision making with large amounts of poorly structured knowledge, expert systems and fuzzy logic are useful tools.

Prosthetic devices and aids for disabled persons. Prosthetic devices have been developed to provide replacements for lost or damaged human functions. Prostheses for upper limbs must be reliable and versatile so that they can partially replace the functionalities of a human arm and hand. The earliest upper-limb prostheses offered simple motional functionalities without providing anthropomorphic features. Command signals for electrical prostheses may be obtained from electromyographic (EMG) signals on the surface of the skin resulting from muscle action or directly from nerve signals with implanted sensors. Commercially available prostheses, however, have not exploited modern robotic technologies. Most of these devices lack fingertip tactile sensing, a feature that humans routinely use in most daily tasks. The largest number of disabilities involves the loss or impairment of lower limbs. An active exoskeletal walking mechanism can significantly improve a person's capability for movement in daily living. Prototype mechanisms have been built to assist a disabled person in walking. With the aid of such mechanisms, a person can walk smoothly using a rolling support or crutches. The device also can be used as a training device for patients with functional disabilities. Exoskeletel mechanisms can also provide therapeutic exercise of a patient's arm and shoulder during stroke rehabilitation. *See* BIOMEDICAL ENGINEERING; MEDICAL CONTROL SYSTEMS; PROSTHESIS.

Micromachines. Miniature mechanical and sensor systems are being developed using the technology of integrated circuit fabrication to etch micro-electromechanical systems (MEMS) directly onto silicon chips. Through the use of multiple structural polysilicon layers, it is possible to generate structures on the scale of a few hundred micrometers, such as a 16-joint anthropomorphic skeleton of a hand or a cantilever for measuring gas flow. By integrating thin-film transistors into the same process, interface electronics may be fabricated simultaneously with the mechanical structure. Pressure sensors,

accelerometers, resonators, and micromotors have been built. Another process generates a photoresist pattern using deep x-ray lithography to create a design on a conductive substrate. This process is useful for making small gears, linkages, or other parts. *See* MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS). William A. Gruver

Bibliography. J. Craig, *Introduction to Robotics: Mechanics and Control*, 2d ed., 1989; R. Dorf (ed.), *International Encyclopedia of Robotics: Applications and Automation*, 3 vols., 1988; J. Engelberger, *Robotics in Practice*, AMACOM, New York, 1980; J. Engelberger, *Robotics in Service*, MIT Press, Cambridge, MA, 1989; L. Fausett, *Fundamentals of Neural Networks*, Prentice Hall, 1994; J. W. Gardner, *Microsensors: Principles and Applications*, Wiley, 1994; W. A. Gruver and J. Boudreaux, *Intelligent Manufacturing: Programming Environments for CIM*, Springer-Verlag, 1993; P. McKerrow, *Introduction to Robotics*, Addison Wesley, 1991; Micromachines on the march, *IEEE Spectrum*, pp. 20–31, May 1993; T. Sheridan, Telerobotics, *Automatica*, 23(4):487–507, 1988; Virtual reality is for real, *IEEE Spectrum*, pp. 22–39, October 1993; P. H. Winston, *Artificial Intelligence*, Addison Wesley, 1984; E. H. van Leeuwen and D. Norrie, Inteligent manufacturing: Holons and holarchies, *Manuf. Eng.*, 76(2):86–88, April 1997.

Roche limit

The closest distance which a satellite, revolving around a parent body, can approach the parent without being pulled apart tidally. The simplest formal definition is that the Roche limit is the minimum distance at which a satellite can be in equilibrium under the influence of its own gravitation and that of the central mass about which it is describing a circular orbit. If the satellite is in a circular orbit and has negligible mass, the same density as the primary, and zero tensile strength, the Roche limit is 2.46 times the radius of the primary.

The concept of an equilibrium state was first given by Édouard Roche in 1849. The radius r of the Roche limit appropriate to satellite formation is given by the equation below, where B is the radius and ρ the

$$r = 2.46\left(\frac{\rho}{\rho_s}\right)^{1/3} B$$

mean density of the parent planet, and ρ_s is the mean density of the satellite. This is the radius at which the surface of a homogeneous, fluid satellite can remain in stable hydrostatic equilibrium.

The popular definition of Roche's limit, that is, the distance from the planet at which a satellite would suffer tidal loss of particles, depends on properties of the material and the shape and density of the satellite. If, for instance, the material can flow to adopt the hydrostatic equilibrium shape, then the body will be elongated in the direction of the radius vector. For a binary system, the Roche limit is defined by the value of the Jacobi constant such that the zero velocity surfaces around the two bodies intersect in the lagrangian libration point L_2 between the finite masses. This implies an upper limit to the size of the components. While magnetohydrodynamic forces act on the outer layers of stars and the theory holds only for the circular restricted problem with the masses acting as point masses, the Roche limit seems to correspond to the observational data. When a star has exhausted the supply of hydrogen in its core, its radius will increase by a factor of 10 to 100. A star in a binary system may then exceed its Roche limit, material will thus escape from that star, and its companion will receive the excess material. *See* BINARY STAR; CELESTIAL MECHANICS. P. K. Seidelmann

Rock

A relatively common aggregate of mineral grains. Some rocks consist essentially of but one mineral species (monomineralic, such as quartzite, composed of quartz); others consist of two or more minerals (polymineralic, such as granite, composed of quartz, feldspar, and biotite). Rock names are not given for those rare combinations of minerals that constitute ore deposits, such as quartz, pyrite, and gold. In the popular sense rock is considered also to denote a compact substance, one with some coherence; but geologically, friable volcanic ash also is a rock. A genetic classification of rocks follows:

- Igneous
  - Intrusive
    - Plutonic (deep)
    - Hypabyssal (shallow)
  - Extrusive
  - Flow
    - Pyroclastic (explosive)
- Sedimentary
  - Clastic (mechanical or detrital)
  - Chemical (crystalline or precipitated)
  - Organic (biogenic)
- Metamorphic
  - Cataclastic
  - Contact metamorphic and pyrometasomatic
  - Regional metamorphic (dynamothermal)
- Hybrid
  - Metasomatic
  - Migmatitic

Exceptions to the requirement that rocks consist of minerals are obsidian, a volcanic rock consisting of glass; and coal, a sedimentary rock which is a mixture of organic compounds. *See* COAL; VOLCANIC GLASS.

Igneous rocks. Igneous rocks are those that have solidified from a molten condition. The parent material is magma—a natural, hot, mutual solution of silicates with minor amounts of water and other volatiles. Igneous rocks are divided into those which crystallized before magma reached the Earth's

TABLE 1. Simplified classification of major igneous rocks on the basis of composition and texture

	SiO_2-rich (acidic) ←				→ SiO_2-poor (basic)
	Light colored ←		Gray	Dark colored →	Black
Mineral composition:	Quartz, potash feldspar, biotite	Potash feldspar, biotite, or amphibole	Sodic plagioclase, hornblende, or augite	Augite, olivine, hypersthene, calcic plagioclase	Olivine, enstatite, augite
Intrusive					
Medium-grained	Granite*	Syenite	Diorite	Gabbro	Peridotite
Extrusive					
Fine-grained to aphanitic	Rhyolite	Trachyte	Andesite	Basalt	
	←	Felsite	→		
Porphyritic	Rhyolite porphyry	Tracyte porphyry	Andesite porphyry	Basalt porphyry	
Glassy	Obsidian				
Vesicular	Pumice			Scoria	
Fragmental	Tuff and agglomerate of each type				

*Exceptionally coarse-grained rock of general granitic composition is pregmatite.

surface (intrusive rocks) and those that solidified at the surface, some as layers of lava (the extrusive flow rocks) and others as pyroclastic debris in explosive eruption (**Table 1**). *See* MAGMA; PYROCLASTIC ROCKS.

Chief elements in igneous rocks are oxygen, O; silicon, Si; aluminum, Al; iron, Fe; magnesium, Mg; calcium, Ca; sodium, Na; and potassium, K. Compositions range from about 40% SiO_2 (peridotites) to as much as 70% SiO_2 (granites). Silica-poor rocks contain relatively large amounts of Ca, Mg, and Fe[2] (basic rocks), whereas silica-rich types contain larger amounts of Na and K (acidic rocks). *See* IGNEOUS ROCKS.

Sedimentary rocks. Clastic sedimentary rocks (consisting of mechanically transported particles) are subdivided on the basis of particle size (**Table 2**). Those having intermediate and fine-grain sizes are further subdivided on the basis of composition (see **illus.**). Other significant clastic rocks are those consisting of detrital calcite, the calcarenites and calcilutites.

Textures of clastic rocks derive from grain size, sorting, form, and arrangement. Form includes sphericity (shape), the degree to which a particle approximates a sphere; and roundness, the measurement of the sharpness of edges and corners. Only glacial sedimentary rocks (tillite) do not show layering or stratification.

Chemical sedimentary rocks are those precipitated from ocean, lake, and ground water. The most important ones are shown in the following list:

Rock	*Chief Mineral*
Chert	Chalcedony, SiO_2
Limestone	Calcite, $CaCO_3$
Travertine (spring deposit)	Calcite, $CaCO_3$
Dolomite	Dolomite, $CaMg(CO_3)_2$
Phosphorite	Apatite, $Ca_{10}(PO_4)_5(CO_3)F_3$
Salines (evaporites)	
Rock salt	Halite, NaCl
Rock anhydrite	Anhydrite, $CaSO_4$
Rock gypsum	Gypsum, $CaSO_4 \cdot 2H_2O$

Organic sedimentary rocks include (1) siliceous types made up of opaline tests of diatoms, diatomite, or radiolaria, radiolarite; (2) calcareous types, consisting of calcite shell, namely, shell limestone and coquina; and (3) carbonaceous types, namely, coal and other accumulations of altered plant debris. *See* SEDIMENTARY ROCKS; SEDIMENTOLOGY.

Metamorphic rocks. Metamorphic rocks owe their complexity of composition and texture not only to the existence of several types of metamorphism but also to the application of these types under different intensities to a variety of parent rocks; thus, both sedimentary and igneous rocks may be metamorphosed (**Table 3**). Regional metamorphic rocks are distinguished by foliation, a parallel orientation of platy or

TABLE 2. Size classification of clastic sedimentary particles and aggregates

Size, in. (mm)	Particle	Aggregate
>10 (>256)	Boulder	Gravel, conglomerate (psephite, rudite)
10–2.5 (256–64)	Cobble	Breccia (angular)
2.5–0.16 (64–4)	Pebble	
0.16–0.079 (4–2)	Coarse sand	Sandstone (psammite, arenite)
0.079–0.0025 (2–0.63)	Sand	
0.0025–0.00015 (0.63–0.0039)	Silt	Siltstone (pelite, lutite)
0.00015 (<0.0039)	Clay	Clay
		Shale

TABLE 3. Simplied classification of metamorphic rocks, with selected examples

Parent rocks	Contact metamorphism	Regional metamorphism: Low grade	Regional metamorphism: Intermediate grade	Regional metamorphism: High grade
Sandstone, arkose	Quartzite (quartz)	Quartzite and quartz-feldspar gneiss		
Shale	Hornfels (andalusite, cordierite)	Slate, phyllite, (chlorite, muscovite)	Mica schist (biotite, garnet, kyanite, staurolite)	Sillimanite gneiss (sillimanite, almandite)
Limestone	Marble (calcite)	Marble and calc-silicate gneiss		
		(Calcite, tremolite)	(Calcite, wollastonite)	(Calcite, diopside, anorthite)
Basalt	Hornfels (plagioclase, hypersthene)	Greenschist (chlorite, albite, epidote)	Amphibole schist, (plagioclase, actinolite)	Amphibolite (andesine, hornblende, garnet)

TABLE 4. Physical properties of some common rocks

Rock	Specific gravity	Porosity, %	Compressive strength, lb/in.^{2}*	Tensile strength, lb/in.^{2}*
Igneous				
Granite	2.67	1	30,000–50,000	500–1000
Basalt	2.75	1	25,000–30,000	
Sedimentary				
Sandstone	2.1–2.5	5–30	5,000–15,000	100–200
Shale	1.9–2.4	7–25	5,000–10,000	
Limestone	2.2–2.5	2–20	2,000–20,000	400–850
Metamorphic				
Marble	2.5–2.8	0.5–2	10,000–30,000	700–1000
Quartzite	2.5–2.6	1–2	15,000–40,000	
Slate	2.6–2.8	0.5–5	15,000–30,000	

*1 lb/in.^{2} = 6.9 kPa.

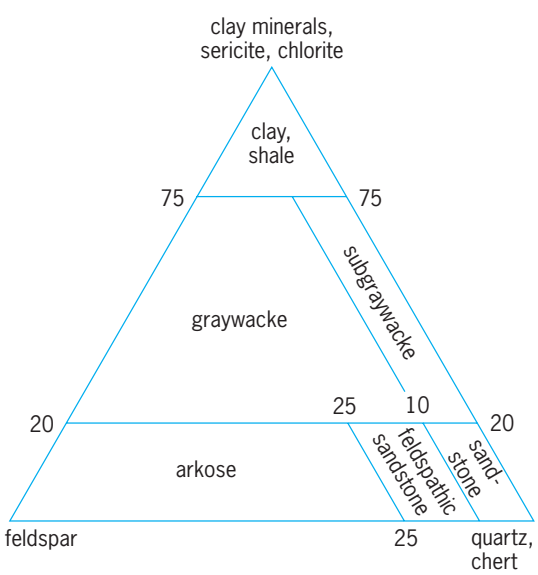

Classification of psammitic and pelitic rocks based on proportions of the three most common clastic minerals: quartz, feldspars, and clays—excluding calcilutites and calcarenites. (*After E. W. Heinrich, Microscopic Petrography, McGraw-Hill*, 1956)

prismatic minerals. *See* METAMORPHIC ROCKS; METAMORPHISM.

Physical properties and behavior. When stresses are applied to rocks, either natural (those active in mountain building) or human-made, resulting from loading with structures (such as dams), deformation (strain) may result. Rocks subjected to stress normally undergo three deformation stages: elastic—the rock returns to its original size or shape upon withdrawal of stress; plastic—beyond a limiting stress (elastic limit) there is only partial restoration upon stress removal; and fracture—breakage with further increase in stress. With increases in confining pressure (load of overlying rocks) and temperature, the interval between the elastic limit and fracture increases. Thus rocks that behave as brittle substances near the Earth's surface, failing by fracturing, will at depth be deformed plastically by solid flow. The effects of stress depend on physical properties (**Table 4**). *See* ENGINEERING GEOLOGY; HIGH-PRESSURE MINERAL SYNTHESIS; ROCK MECHANICS; STRUCTURAL GEOLOGY.

Rock cycle. Igneous rocks exposed at the Earth's surface are subject to weathering, which alters them chemically and physically. Such material when transported, deposited, and consolidated becomes

sedimentary rock, which, through heat and pressure, may be converted to metamorphic rock. Both sedimentary and metamorphic rocks also may be weathered and transformed into younger sediments. Deeply buried metamorphic rocks may be remelted to yield new igneous material. *See* LITHOSPHERE; PETROGRAPHY; PETROLOGY; STONE AND STONE PRODUCTS.

E. William Heinrich

Rock, electrical properties of

The effect of changes in pressure and temperature on electrical properties of rocks. There has been increasing interest in the electrical properties of rocks at depth within the Earth and the Moon. The reason for this interest has been consideration of the use of electrical properties in studying the interior of the Earth and its satellite, particularly to depths of tens or hundreds of kilometers. At such depths pressures and temperatures are very great, and laboratory studies in which these pressures and temperatures are duplicated have been used to predict what the electrical properties at depth actually are. More direct measurements of the electrical properties deep within the Earth have been made by using surface-based electrical surveys of various sorts. An important side aspect of the study of electrical properties has been the observation that, when pressures near the crushing strength are applied to a rock, marked changes in electrical properties occur, probably caused by the development of incipient fractures. Such changes in resistivity might be used in predicting earthquakes, if they can be measured in the ground. *See* GEOELECTRICITY; GEOPHYSICAL EXPLORATION.

Electrical zones in the Earth. Attempts to measure the electrical properties of rocks to depths of tens or hundreds of kilometers in the Earth indicate that the Earth's crust is zoned electrically. The surface zone, with which scientists are most familiar, consists of a sequence of sedimentary rocks, along with fractured crystalline and metamorphic rocks, all of which are moderately good conductors of electricity because they contain relatively large amounts of water in pore spaces and other voids. This zone, which may range in thickness from a kilometer to several tens of kilometers, has conductivities varying from about $^1/_2$ ohm-m in recent sediments to 1000 ohm-m or more in weathered crystalline rock.

The basement rocks beneath this surface zone are crystalline, igneous, or metamorphic rocks which are much more dense, having little pore space in which water may collect. Since most rock-forming minerals are good insulators at normal temperatures, conduction of electricity in such rocks is determined almost entirely by the water in them. As a result, this part of the Earth's crust is electrically resistant. There are numerous experimental difficulties involved in trying to measure the electrical properties of this part of the Earth's crust, but it appears that the resistivity lies in the range 10,000–1,000,000 ohm-m.

At rather moderate depths beneath the surface

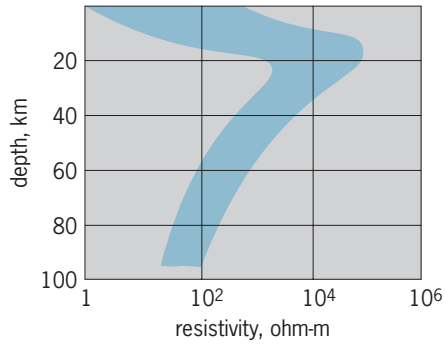

Fig. 1. Probable range of electrical resistivity through the crust and into the upper mantle. 1 km = 0.6 mi.

of the second zone, resistivity begins to decrease with depth. This decrease is considered to be the result of higher temperatures, which almost certainly are present at great depths. High temperatures lead to partial ionization of the molecular structure of minerals composing a rock, and the ions render even the insulating minerals conductive. As temperature increases with depth, this effect becomes more pronounced and resistivity decreases markedly. The maximum resistivity in the Earth's crust probably occurs at depths of only 3–6 mi (5–10 km) below the surface of the crystalline basement. At depths of 18–24 mi (30–40 km), corresponding to the top of the mantle, resistivity is only about a few hundred ohm-meters. Beyond this depth, resistivity seems to decrease slightly with increasing depth, but again drops sharply at depths around 420 mi (700 km). At greater depths the resistivity is less than 1 ohm-m. A profile of resistivity as it varies downward into the Earth is shown in **Fig. 1**.

At relatively shallow depths information is gained by lowering equipment into bore holes to measure the electrical properties of rocks directly. Another approach is the use of low-frequency currents introduced into the Earth at the surface. These methods are restricted to use in studying resistivities at depths less than 6 mi (10 km). Information from greater depths is obtained by studying slow variations in the Earth's magnetic field and the currents induced in the Earth by these variations. Fluctuations in the magnetic field with periods measured in months penetrate to distances up to 600 mi (1000 km), and they may be used in estimating conductivity to these depths.

Temperature effects. Laboratory investigations have also been pursued to find how the electrical properties of rocks depend on the pressure and temperature deep in the Earth. Inasmuch as the temperature at depths of 24–36 mi (40–60 km) is probably in the range 1470–2200°F (800–1200°C) and the pressure is in the range 140,000–210,000 lb/in.^{2} (10,000–15,000 kg/cm^2), these conditions are difficult to duplicate in the laboratory. Both electrical conductivity and dielectric constant have been measured by several researchers at pressures to 700,000 lb/in.^{2} (50,000 kg/cm^2) and temperatures to 2200°F (1200°C), using dry rock samples. The most obvious conclusion from such studies is that

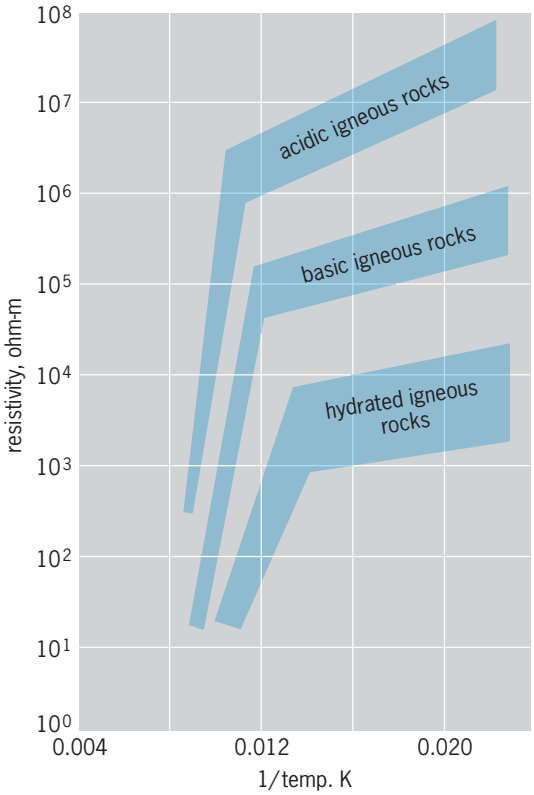

Fig. 2. Variation of resistivity with temperature in dry igneous rocks. Hydrated samples contained only water of hydration, not free water. At the melting point, conduction increases abruptly. °F = (K × 1.8) − 459.67.

for this range of conditions the effect of temperature on electrical properties is much more profound than that of pressure. An increase in temperature from 68 to 1830°F (20 to 1000°C) will reduce the resistivity of a dry rock by a factor of 10^6. On the other hand, an increase in pressure from 0 to 700,000 lb/in.^{2} (50,000 kg/cm^2) changes the resistivity of a dry rock by less than 100%. Typical behavior for the resistivity of dry rocks as temperature is raised is shown by curves in **Fig. 2**. At temperatures up to about half the melting point, conduction increases relatively slowly with increasing temperature, and it is thought that conduction is due to impurities, which contribute ions under weak thermal excitation. At temperatures above half the melting point, conduction increases rapidly with increasing temperature, with conduction being caused by ions torn from normal lattice positions in the crystals by violent thermal agitation. At the melting point, conduction increases abruptly as a rock melts and becomes highly ionized.

It has been observed that silicic rocks, such as granite, granodiorite, and rhyolite, are poorer conductors of electricity at all temperatures than other rock types. Rocks, such as gabbro, basalt, and dunite, which are rich in ferromagnesian minerals, are more conductive than the silicic rocks by about an order of magnitude. This is believed to result from the fact that the principal ions which contribute to conduction are the small ferromagnesian ions, which are more abundant in the dark rocks.

Pressure effects. Typical curves for the behavior of resistivity and of dielectric constant with the application of pressure are shown in **Figs. 3** and **4**. The dielectric constant is found to increase by 10–20% as the pressure on a rock sample is increased to 1400–2800 lb/in.^{2} (100–200 kg/cm^2) and then to increase more slowly as the pressure is elevated further. The increase in dielectric constant is due in part to an increase in density of the rocks under pressure.

Two forms of behavior have been recognized in studies of the behavior of resistivity at pressures up to 700,000 lb/in.^{2} (50,000 kg/cm^2). With some rocks resistivity decreases uniformly with increasing pressure; in other rocks resistivity decreases at low applied pressures but increases at high pressures. A uniform decrease has been observed in rocks such as basalt, diabase, and amphibolite, while an increase in resistivity over part of the pressure range has been observed in microcline-rich rocks. The increase in resistivity may be associated with a decrease in the mobility of charge carriers as the crystal lattice is compressed.

The effect of pressure on the resistivity of rocks in which conduction is determined by water content rather than by the crystal matrix is quite different.

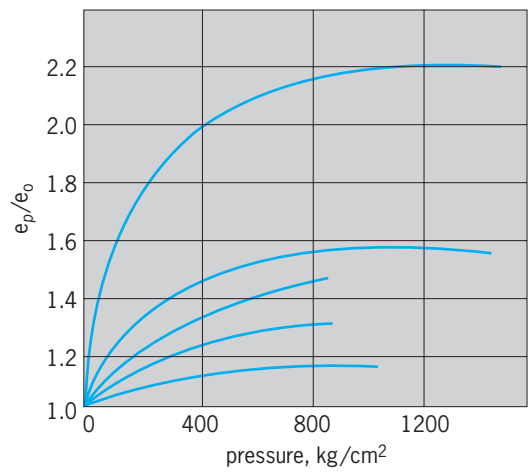

Fig. 3. Effect of pressure on the dielectric constants of five samples of granite. The ratio e_p/e_0 is the ratio of the dielectric constant measured at elevated pressure to dielectric constant measured at zero pressure. 1 kg/cm^2 = 0.98 × 10^5 Pa = 14 lb/in.^{2}.

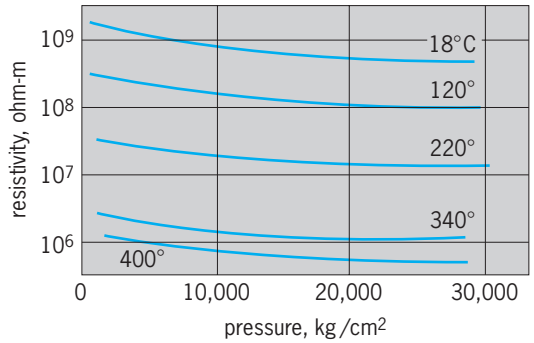

Fig. 4. Resistivity of a basalt sample measured as a function of pressure at various temperatures. 1 kg/cm^2 = 0.98 × 10^5 Pa = 14 lb/in.^{2}; °F = (°C × 1.8) + 32.

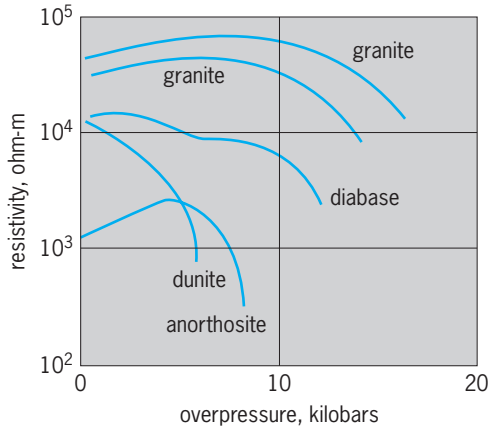

Fig. 5. Effect of greater pressure on framework of a rock than on fluid in pores in water-bearing rocks. 1 kilobar = 10^8 Pa.

Several investigators have studied the effect of having greater pressure on the crystal framework than on the pore fluid. As an overpressure is applied to the rock skeleton, the framework starts to close in on the void spaces, reducing their volume and increasing the resistivity. It has also been observed that, if the overpressure is made great enough, resistivity begins to increase again, a phenomenon which is attributed to the opening of small cracks prior to rock failure, with the movement of water into these cracks to contribute to conduction. The behavior of resistivity for a number of rock specimens as a function of frame overpressure is shown by the curves in **Fig. 5**. According to W. F. Brace and A. S. Orange, observations of such changes in resistivity might be used to detect minute changes in the structure of rock under stress. They observe that the buildup of strain in the Earth preceding earthquakes ought to be accompanied by changes in resistivity, and that these changes might warn of an impending earthquake. *See* HIGH-PRESSURE MINERAL SYNTHESIS; PETROLOGY.

Electrical zones in the Moon. In planning the exploration of the Moon, some thought has been given to the use of measurements of electrical properties as a means for studying the Moon's interior. Although there is little information on the electrical properties of the Moon, the knowledge gained about the Earth can be used to make an intelligent guess. The major difference is the lack of an atmosphere and moisture, which means that the surface rocks on the Moon must be quite different from those on Earth. Measurements with radar and infrared sensors have indicated that the surface of the Moon must be covered with a rock froth or soil which has a very low density—the rocks may be only about 10% solid material. Inasmuch as there is probably no moisture in this rock, it seems very likely that the electrical resistivity must be very large, perhaps 10^{10} ohm-m.

It is believed that temperature increases with depth inside the Moon, though perhaps not as rapidly as in the Earth. The temperature at the center of the Moon may be 3140°F (2000 K) and, as a result, the rocks in the interior can be expected to be conductive, just as they are on the Earth. The resistivity at the center of the Moon may be of the order of 1–100 ohm-m.

The change in resistivity between the highly resistant surface and the conductive interior depends very much on the composition of the Moon. Some scientists believe that the interior contains as much water as the rocks within the Earth, at depths beyond which evaporization into space has been able to take place. Others believe that the Moon had a different origin from the Earth, and thus never contained very much water. If water is present, one would expect the resistivity to drop very rapidly over the first few kilometers in depth, as water contributes to electrical conduction. If the Moon is devoid of water, the resistivity should change more gradually with depth and become low only at depths of many hundreds of kilometers, where the temperatures are high. The presence of water at moderate depths would be very important to future lunar exploration—it might be used as fuel for trips deeper into space, or even to provide a tenuous atmosphere to protect humans' installations on the Moon. Therefore it is likely that electrical surveys for water on the Moon will have a high priority in the early stages of lunar exploration. *See* MOON.

George V. Keller

Bibliography. M. B. Dobrin and C. H. Savit, *Introduction to Geophysical Prospecting*, 4th ed., 1988; P. Kearey and M. Brooks, *Introduction to Geophysical Exploration*, 2d ed., 1991; G. V. Keller, *Electrical Methods of Geophysical Prospecting*, 2d ed., 1982; E. S. Robinson and C. Coruh, *Basic Exploration Geophysics*, 1988; R. E. Sheriff, *Geophysical Methods*, 1989.

Rock age determination

Finding the age of rocks based on the presence of naturally occurring long-lived radioactive isotopes of several elements in certain minerals and rocks. Measurements of rock ages have enabled geologists to reconstruct the geologic history of the Earth from the time of its formation 4.6×10^9 years ago to the present. Age determinations of rocks from the Moon have also contributed to knowledge of the history of the Moon, and may someday be used to study the history of Mars and of other bodies within the solar system. *See* RADIOACTIVITY.

Principles of dating. Many rocks and minerals contain radioactive atoms that decay spontaneously to form stable atoms of other elements. Under certain conditions these radiogenic daughter atoms accumulate within the mineral crystals so that the ratio of the daughter atoms divided by the parent atoms increases with time. This ratio can be measured very accurately with a mass spectrometer, and is then used to calculate the age of the rock by means of an equation based on the law of radioactivity.

According to this law, the number of radiogenic daughters (D) in a unit weight of rock is related to the number of radioactive parent atoms remaining

Parent-daughter pairs used for dating rocks and minerals

Parent	Daughter	Half-life, 10^9 years
Potassium-40	Argon-40	11.8
Potassium-40	Calcium-40	1.47
Rubidium-87	Strontium-87	48.8
Samarium-147	Neodymium-143	107
Rhenium-187	Osmium-187	43
Thorium-232	Lead-208	14.008
Uranium-235	Lead-207	0.7038
Uranium-238	Lead-206	4.468

(P) by Eq. (1). Here e is a mathematical constant ap-

$$\frac{D}{P} = e^{\lambda t} - 1 \qquad (1)$$

proximately equal to 2.7182, λ is the decay constant of the parent, and t may be regarded as the age of the rock or mineral. The value of the decay constant is characteristic of each radioactive atomic species, and must be known from experimental determinations. It is related to the half-life ($t_{1/2}$) of the parent by the equation $\lambda = 0.693/t_{1/2}$, where the half-life is defined as the time required for one-half of a given number of radioactive atoms to decay. The radioactive atoms used for dating rocks and minerals have very long half-lives, measured in billions of years. They occur in nature only because they decay very slowly. The pairs of parents and daughters used for dating are listed in the **table**. *See* DATING METHODS.

Rubidium-strontium method. Rubidium (Rb) is an alkali metal whose chemical properties are similar to those of potassium. Therefore, rubidium occurs in small amounts in all minerals in which potassium is a major constituent. The most common rubidium-bearing minerals are the micas (biotite and muscovite), the feldspars (orthoclase and microcline), and certain of the clay minerals (illite and glauconite). *See* CLAY MINERALS; FELDSPAR; MICA.

The radioactivity of rubidium is due to rubidium-87, which decays to stable strontium-87 (^{87}Sr) by emitting a beta particle from its nucleus. The abundance of the radiogenic strontium-87 therefore increases with time at a rate that is proportional to the Rb/Sr ratio of the rock or mineral. This relationship is expressed by Eq. (2), where $^{87}Sr/^{86}Sr$ and $^{87}Rb/^{86}Sr$

$$\frac{^{87}Sr}{^{86}Sr} = \left(\frac{^{87}Sr}{^{86}Sr}\right)_0 + \frac{^{87}Rb}{^{86}Sr}(e^{\lambda t} - 1) \qquad (2)$$

are the ratios of these atoms in the sample to be dated and $(^{87}Sr/^{86}Sr)_0$ is the value of this ratio in the rock or mineral at the time of its formation. *See* BETA PARTICLES.

The effectiveness of the Rb-Sr method of dating is enhanced by the difference in the geochemical properties of strontium and rubidium. Strontium is an alkaline earth element that replaces calcium in minerals such as plagioclase feldspar and apatite, a calcium phosphate. The crystallization of a silicate liquid called magma may produce different varieties of igneous rocks whose Rb/Sr ratios have a wide range of values depending on whether they are composed of Rb-rich or Sr-rich minerals. Such comagmatic rocks all have the same age (t) and the same initial $^{87}Sr/^{86}Sr$ ratio. When their $^{87}Sr/^{86}Sr$ and $^{87}Rb/^{86}Sr$ ratios are plotted (**Fig. 1**), each rock forms a point that lies on a straight line called an isochron. The age of the rocks that form the isochron is indicated by its slope (m) because $m = e^{\lambda t} - 1$ and therefore $t = (1/\lambda) \ln (m + 1)$. The intercept of the isochron yields the value of the initial $^{87}Sr/^{86}Sr$ ratio. *See* MAGMA.

The whole-rock Rb-Sr isochron method outlined above has been used to date not only igneous but also metamorphic rocks. The method is particularly well suited to the dating of very old rocks such as the ancient gneisses near Godthaab in Greenland, which are almost 3.8×10^9 years old. This method has also been used to date rocks from the Moon and to determine the age of the Earth by analyses of stony meteorites.

The Rb-Sr method has had limited success in the dating of sedimentary rocks. Such rocks commonly contain fossils that indicate their geologic ages. However, only dating methods based on radioactivity give an absolute time scale. The clay mineral glauconite has yielded the only reliable dates for sedimentary rocks by the Rb-Sr and the potassium-argon (K-Ar) method. Certain other clay minerals called smectite may also be suitable for dating. Nevertheless, the dating of sedimentary rocks remains a challenge for the future.

Potassium-argon method. The decay of potassium-40 to argon-40 was one of the first methods of dating to be developed in the 1950s. It is based on the assumption that all of the atoms of radiogenic argon-40 that form within a potassium-bearing mineral accumulate within it. This assumption is satisfied only by a few kinds of minerals and rocks, because argon is an inert gas that does not readily form bonds with other atoms. In addition, the mineral or rock to be dated must not contain any excess argon-40 atoms that did not actually form by decay of potassium-40 within it. These conditions are satisfied only by the

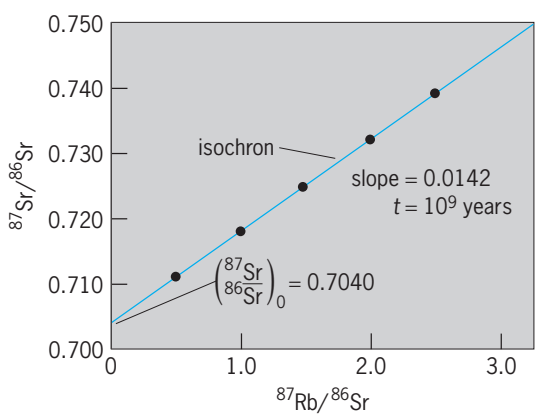

Fig. 1. Isochron diagram used for dating by the Rb-Sr method. The isochron is fitted to the data points representing rocks or minerals. The slope of the isochron yields the age of samples, and its intercept is the initial $^{87}Sr/^{86}Sr$ ratio. The samples in this hypothetical example are 10^9 years old.

micas, sanidine feldspar, hornblende, and the clay minerals glauconite and illite. In addition, certain fine-grained volcanic rocks of basaltic composition have been dated successfully. Age determinations by this method require measurements of the concentrations of argon-40 and potassium in the sample and are based on Eq. (3), where $\lambda = 5.305 \times 10^{-10}$ per year

$$^{40}\text{Ar} = 0.1102 \quad ^{40}\text{K}(e^{\lambda t} - 1) \qquad (3)$$

is the decay constant of potassium-40. The alternate decay of potassium-40 to calcium-40 is not useful for dating because that atomic species of calcium has a very high natural abundance.

Potassium-argon dates of micas in igneous and metamorphic rocks indicate the time elapsed since the mineral cooled to about 300°C (570°F), when radiogenic argon is effectively trapped in the crystal lattice. Therefore K-Ar dates of micas may underestimate the crystallization age of the igneous or metamorphic rocks in which they occur. This property of K-Ar dates has been used to measure the rate of cooling of such rocks whose crystallization age can be measured by the whole-rock Rb-Sr method. Potassium-argon dates of hornblende are commonly older than those of biotite extracted from the same rocks because hornblende retains argon-40 at a higher temperature than biotite. The K-Ar method of dating has been used to establish a chronology of mountain building events in North America beginning about 2.8×10^9 years ago and continuing to the present. In addition, the method has been used to date reversals of the polarity of the Earth's magnetic field during the past 1.3×10^7 years. *See* OROGENY; PALEOMAGNETISM.

Argon-40/argon-39 method. The problem of argon loss at elevated temperatures can be avoided by the so-called argon 40-39 method. The sample to be dated is irradiated with energetic neutrons in a nuclear reactor. Some of the potassium-39 atoms in the sample are thereby converted to argon-39. The sample is subsequently heated in steps, and the ratio of argon-40 to argon-39 is measured as the gas is released. Experience has shown that argon released at low temperatures is deficient in radiogenic argon-40 compared to argon-39, and yields dates that are less than the age of the sample. Gas fractions released at higher temperatures close to the melting point of the mineral originate from the centers of grains and have argon-40 to argon-39 ratios that yield older dates approaching the crystallization age of the sample. The explanation is that radiogenic argon-40 formed in the interior of mineral grains is less likely to be lost than argon-40 formed near grain boundaries and microfractures. The $^{40}\text{Ar}/^{39}\text{Ar}$ method yields reliable results when it is used to date highly purified concentrates of potassium-bearing minerals, including plagioclase, hornblende, and pyroxene.

Uranium, thorium-lead method. All of the atoms of uranium and thorium are radioactive and decay through a series of radioactive daughters to stable atoms of lead (Pb). Minerals that contain both elements can be dated by three separate methods based on the decay of uranium-238 to lead-206, uranium-235 to lead-207, and thorium-232 to lead-208. The three dates agree with each other only when no atoms of uranium, thorium, lead, and of the intermediate daughters have escaped. Only a few minerals satisfy this condition. The most commonly used mineral is zircon ($ZrSiO_4$), in which atoms of uranium and thorium occur by replacing zirconium. Other minerals that have been used for dating are monazite [$(Ce,La,Dy)PO_4$], titanite ($CaTiSiO_5$), apatite [$(CaF)Ca_4(PO_4)_3$], uraninite (UO_2), and others that are less common. The necessary measurements are the concentrations of uranium, thorium, and lead and the abundances of the radiogenic lead atoms. As an example, the date based on the decay of uranium-238 to lead-206 is calculated from Eq. (4), where λ_1

$$\frac{^{206}\text{Pb}}{^{204}\text{Pb}} = \left(\frac{^{206}\text{Pb}}{^{204}\text{Pb}}\right)_0 + \frac{^{238}\text{U}}{^{204}\text{Pb}}(e^{\lambda_1 t} - 1) \qquad (4)$$

is the decay constant of uranium-238 and the atomic ratios have the same meaning as in the Rb-Sr method.

The dates obtained from uranium-bearing minerals are frequently discordant; that is, they are not in agreement, because lead atoms may be lost from the dated mineral during episodes of thermal metamorphism or by diffusion. Such discordant dates do not indicate the age of the mineral and must be corrected by means of the concordia diagram (**Fig. 2**). Minerals whose uranium-lead decay schemes have been altered lie on straight lines called discordia that intersect the concordia curve in two points. The upper intersection yields the age of the mineral, whereas the lower intersection may date the metamorphic event. The diagram can be used to treat both gain or loss of uranium and loss of lead. The points below concordia have experienced loss of lead or gain of uranium, whereas those above concordia have lost uranium. In this hypothetical example the samples are 3.8×10^9 years old and have been altered recently. Uranium-thorium concordia diagrams have also been used, but complications may arise from

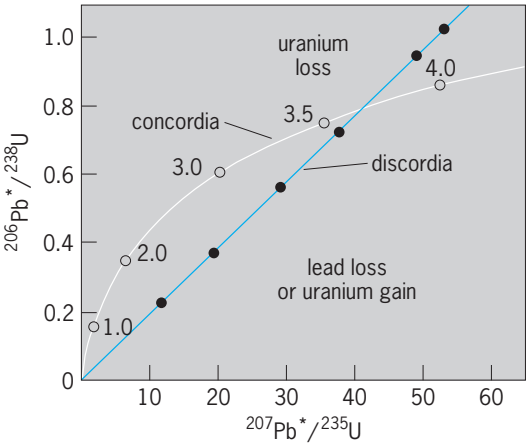

Fig. 2. Concordia diagram used for dating by the U-Pb method. Minerals that yield concordant dates plot on concordia as indicated by the dates expressed in billions of years. Discordant samples form the discordia line which intersects concordia at 3.8×10^9 years and at the origin.

the differences in the geochemical properties of uranium and thorium. When metamorphosed igneous rocks are dated by different methods, the U-Pb concordia dates of zircon usually agree with whole-rock Rb-Sr isochron dates, whereas the K-Ar dates of biotite mica agree with the time of episodic lead loss indicated by the concordia diagram. *See* LEAD ISOTOPES (GEOCHEMISTRY); RADIOACTIVE MINERALS.

Common-lead method. Lead in the common ore mineral galena (PbS) consists of primordial lead that dates from the time of formation of the Earth and varying amounts of radiogenic lead that formed by decay of uranium and thorium in the Earth. The amount of radiogenic lead that was available for mixing with primordial lead increased with time since formation of the Earth. Consequently, young galenas contain more radiogenic lead than old ones. The common-lead method has been used to date ore deposits that cannot readily be dated by other methods. In addition, the theoretical models required for the interpretation of common lead have provided insight into the early history of the solar system and into the relationship between meteorites and the Earth.

Fission-track method. Uranium-238 can decay both by emitting an alpha particle from its nucleus and by spontaneous fission with a decay constant of 8.46×10^{-17} per year. The two fragments of the nucleus leave a trail of damage in the solid through which they move as they dissipate their kinetic energies. The resulting tracks can be enlarged by etching a polished surface of the solid with a suitable reagent. They can then be counted visually by use of a microscope at a magnification of about 1000. The number of spontaneous fission tracks per square centimeter is proportional to the concentration of uranium and to the age of the sample. When the uranium content is known, the age of the sample can be calculated.

This method is suitable for dating a variety of minerals and both natural and manufactured glass. Its range extends from less than 100 years to hundreds of millions of years. Fission track dates are interpreted as cooling ages, because the tracks tend to fade at elevated temperature due to annealing. Minerals such as titanite, zircon, and apatite have different annealing temperatures and therefore yield different dates from which the cooling history of the rock can be determined. The method has also been used for dating archeological objects of glass. *See* FISSION TRACK DATING.

Thorium-230/uranium-238 method. Thorium-230 (^{230}Th) is a radioactive isotope that forms when uranium-238 (^{238}U) decays by emitting alpha particles. Since thorium-230 is an isotope of thorium whose chemical properties differ from those of uranium, the link between uranium-238 and thorium-230 may be broken in the course of geological processes such as partial melting of rocks in the mantle of the Earth or the weathering of uranium-bearing minerals in rocks at the surface of the Earth. When uranium and thorium atoms enter the oceans, uranium tends to be deposited in calcium carbonate skeletons of corals and other organisms, whereas thorium is deposited with sediment particles. Consequently, modern coral reefs contain uranium-238 but lack thorium-230. As time passes, the abundance of thorium-230 in coral reefs increases because of the decay of uranium-238 until parent and daughter reestablish radioactive equilibrium as their decay rates become equal. The time-dependent change in the decay rates of thorium-230 and uranium-238 is expressed by Eq. (5), where the subscript A means

$$\left(\frac{^{230}\mathrm{Th}}{^{238}\mathrm{U}}\right)_{\mathrm{A}} = 1 - e^{-\lambda t} \tag{5}$$

activity and λ is the decay constant of thorium-230 (9.217×10^{-6} per year, half-life: 7.52×10^4 years). This method of dating is useful because the calcium carbonate skeletons of corals have recorded changes in the temperature of ocean water during and after the great Pleistocene ice ages that started about 1 million years ago. *See* ALPHA PARTICLES; RADIOISOTOPE (GEOCHEMISTRY).

Samarium-neodymium method. Both samarium and neodymium belong to the rare earths and occur in low concentrations in many rock-forming minerals. Samarium-147 (^{147}Sm) is radioactive and decays by alpha emission to neodymium-143 (^{143}Nd). The rate of decay is exceedingly slow as indicated by its half-life of 1.07×10^{11} years. The Sm-Nd method of dating separated minerals or whole-rock specimens is similar to the Rb-Sr method [Fig. 1; Eq. (2)]. Specimens of comagmatic igneous rocks or their separated minerals form isochrons whose slope is used to calculate the age of the rocks or minerals. The Sm-Nd method is even more reliable than the Rb-Sr method of dating rocks and minerals, because samarium and neodymium are less mobile than rubidium and strontium.

The isotopic evolution of neodymium in the Earth is described by comparison with stony meteorites. Accordingly, this model can be used to calculate how long ago the neodymium in an igneous, metamorphic, or sedimentary rock separated from the reservoir represented by stony meteorites. *See* METEORITE.

Rhenium-osmium method. Rhenium (Re) occurs primarily in minerals of molybdenum such as molybdenite (MoS_2). Osmium (Os) is found as the native metal together with other platinum-group metals. The Re-Os method of dating is based on the beta decay of naturally occurring rhenium-187 to stable osmium-187. It has been used to date iron meteorites and sulfide ore deposits containing molybdenite.

Improvements in analytical techniques permit a wider use of the Re-Os method of dating for studying iron meteorites and identifying meteorite fragments in rock glasses that occur in some large terrestrial craters formed by impacts of meteorites. In addition, the isotope composition of osmium in the rocks of the crust of the Earth differs appreciably from that of young basalt derived from the mantle. Consequently, the isotope composition of osmium in volcanic rocks is a sensitive indicator of assimilation of crustal rocks by basalt magma. *See* EARTH, AGE OF; RADIOCARBON DATING.

Gunter Faure

Bibliography. G. B. Dalrymple and M. A. Lanphere, *Potassium-Argon Dating*, 1972; D. J. De Paolo, *Neodymium Isotope Geochemistry*, 1988; B. R. Doe, *Lead Isotopes*, 1970; G. Faure, *Principles of Isotope Geology*, 2d ed., 1986; D. R. Givens, *Professional Papers in Archiometric Dating*, 1982; R. E. Taylor, *Radiocarbon Dating*, 1987; D. York and R. M. Farquhar, *The Earth's Age and Geochronology*, 1972.

Rock burst

A seismic event that is caused by the mining of underground or surface openings in a high-stress environment. The Mine Safety and Health Administration (MSHA) defines a rock burst as "a sudden and violent failure of a large volume of overstressed rock, resulting in the instantaneous release of large amounts of accumulated energy." Underground, a rock burst sounds like a loud blast. On the surface, a rock burst sounds like a sonic boom, while the ground and structures on the surface shake as if a small earthquake had occurred.

Rock bursts mainly occur in deep underground hard-rock mines. They also occur in some soft-rock mines, such as coal, potash, and trona ($NaHCO_3 \cdot Na_2CO_3$). A rock burst is a very serious hazard to underground mining personnel. Worldwide, there are over 100 fatalities per year from rock bursts. Because of the violent nature of a rock burst, its occurrence often causes significant damage to nearby underground openings. *See* ORE AND MINERAL DEPOSITS; UNDERGROUND MINING.

Classification and types. In mining, rock bursts are categorized as strain, pillar, and fault-slip bursts. Strain bursts are caused by local high stresses at the edge of a mine opening. Generally, the Richter magnitude of these bursts is less than 2.0, and the resulting damage to the opening is less than 100 tons of displaced rock. Pillar bursts occur when the stress or load on a support pillar exceeds its strength (see **illustration**). Pillar bursts can have Richter magnitudes up to 3.5, and damage to the pillar and adjacent openings can be in the hundreds of tons. A fault-slip burst occurs when the shear stress along a nearby geologic structure is influenced by the overall mine geometry, resulting in a movement along the structure, generally toward the mined-out opening. Richter magnitudes of fault-slip bursts can be very large, up to 5.2, and the resulting damage can be spread over large areas of the mine, involving thousands of tons of displaced rock. *See* MINING; ROCK MECHANICS; SEISMOLOGY; STRESS AND STRAIN.

Rock burst damage to a mine opening from a small pillar burst. Note the track and rail heaved up almost 8 ft (2.5 m).

Causes. As an excavation progresses, the stresses change and the rock deforms around the mine opening. For any type of seismic phenomena to occur, including rock bursting, the stress changes caused by the excavation must result in an unstable equilibrium of the rock mass.

Most hard-rock mining districts throughout the world that experience rock bursts have many similarities. They are in an initially high-stress environment, because of depth, tectonic forces, or both. The shape of the orebody is tabular and the thickness of the vein or width of the seam is usually less than 5 m (16 ft). A very high percentage (over 80%) of the orebody is mined. The wall rocks, and often the vein or seam material, are very hard, brittle, and strong, as well as very old geologically—greater than a billion years old. This combination of high stresses, narrow veins, high extraction ratios, and strong, brittle rocks in a complex geologic environment almost always leads to rock bursting in deep mines.

Dealing with rock bursts. While there is no practical way to eliminate rock bursting, measures have been developed to minimize both the occurrence and effects of bursting. The most usual means of minimizing high stresses around mining openings is by using geometries and excavation sequences that do not create burst-prone pillars and that transfer stress to abutments in a manner that reduces stress concentrations. In practice, this is achieved by using a long-wall-type mining geometry. As mining experience in an orebody or district is gained, the relationship between rock bursting and certain geologic features become evident. Mining through or along such features should be done so that stress increases are minimized or dealt with. Mining rates can be adjusted to allow time for the ground around an advancing opening to settle after blasting. Burst-prone faces are usually mined on a single-shift basis. Most mines in North America have converted to cut-and-fill mining to minimize the geometry change caused by the daily face advance. More recently, mines have converted to long-hole mining to reduce miner exposure at the face and increase productivity.

Standard roof and wall support for an opening will not hold up under the effects of a nearby rock burst. Such reinforcement is designed to support dead loads and not the dynamic shock load resulting from a rock burst. Yielding support, which absorbs energy during deformation, has been found to best contain rock-burst damage. Containment of damage maintains the integrity of the opening and prevents miners from being buried.

In mining, it is not always possible to prevent a rock burst. For some special geometric relationships, it has been possible to avoid the occurrence of

or trigger a burst by destressing. The idea of destressing is to blast/fracture the rock in a high-stress or potential rock-burst zone to reduce the stress. Destress blasting reduces the stress and dissipates the stored strain energy in the rock by inducing fractures that allow yielding.

Most mines that experience rock bursting have installed some type of seismic or microseismic monitoring system. Data from seismic monitoring are used to assess the relative seismicity and potential rock-burst hazard associated with advancing mining faces. Knowledge of a problem area allows rock-burst-control measures to be taken to minimize the occurrence and effects of a burst. *See* SEISMOGRAPHIC INSTRUMENTATION.

During the last 40 years, numerous government-industry cooperative research groups have investigated rock bursts in the United States, Canada, South Africa, Poland, and Australia. This research has resulted in mines being better able to manage their rock-burst problems. However, rock bursts are still the most serious and least understood problem facing deep-mining operations worldwide.

Wilson Blake

Bibliography. W. Blake and D. G. F. Hedley, *Rockbursts: Case Studies from North American Hard-Rock Mines*, 2003; D. G. F. Hedley, *Rockburst Handbook for Ontario Hardrock Mines*, 1992; M. Hudyma, *Mining-Induced Seismicity in Underground Mines: Results of World Wide Survey*, 2004.

Rock cleavage

A secondary, planar structure of deformed rocks. A cleavage is penetrative and systematic, as opposed to fractures and shear zones which may occur alone or in widely spaced sets. It is generally better developed in fine grained rocks than in coarse ones. Application of the term derives from the ability to split rocks along the structure.

Simple cleavages are generally parallel to the axial surfaces of folds. This is true for folds formed by uniform flow of the rock mass where primary layering behaves as a passive marker. When rock layers are buckled, the primary layering behaving as mechanical discontinuities, cleavage that develops early may be fanned by subsequent growth of the fold. The mean orientation of the fanned structures remains a measure of the orientation of the axial surface. Cleavage orientation may change from layer to layer in folds when the rock layers possess contrasting mechanical properties. This phenomenon, known as refraction of cleavage, probably results from local variations in stress orientation because of interlayer shear or from cleavage forming at different times in the various layers during fold growth. The angular relationship between bedding planes and cleavage is an important indicator for the interpretation of fold geometry. In normal situations, cleavage dips more steeply than the bedding planes of the layers in which it occurs. If the cleavage is the more steeply dipping structure, the strata are overturned. In certain completely overturned structures, the structural facing may be determined by examining the younging direction of the strata as observed on the cleavage surface. Cleavage is also associated with faults. *See* FAULT AND FAULT STRUCTURES; FOLD AND FOLD SYSTEMS.

Continuous (microscopically penetrative) cleavages, as in slates, are the earliest tectonic fabric elements that can be recognized in rocks. They may be deformed subsequently by disjunctive (mesoscopically penetrative) structures called crenulation cleavage, a periodic corrugation of the previous foliation. These domainal structures can occur in several generations, reflecting different episodes of deformation.

H. C. Sorby observed the microscopic preferred orientation of phyllosilicate grains in slates. He suggested that platy grains in the shale protolith were rotated into the cleavage plane. A. March showed theoretically that preferred orientation derived in this manner is a function of the strain in the rock. Other early workers reasoned that minerals should grow with their most compliant directions parallel to the maximum compressive stress; phyllosilicate grains would align their basal planes perpendicular to the maximum compression. D. S. Wood has shown that cleavage develops perpendicular to the maximum shortening direction, and that the degree of preferred orientation reflects the strain as predicted by March. It appears that both growth and reorientation may play important parts. As temperature and duration of metamorphism increase, growth processes become dominant, and slaty cleavage is transformed into the coarser foliation of schists. *See* SCHIST; SLATE.

Cleavage also develops in rocks of lower deformational and metamorphic grade than slates. In these rocks, the cleavage occurs as discrete surfaces or seams, often coated with a film of clay or carbonaceous material; the cleavage surfaces are separated by zones of undeformed sedimentary rock. The surfaces may be smooth, anastomosing, or stylolitic. Pressure solution is a major cleavage-forming process in these rocks. Because no macroscopic inhomogeneity exists, these cleavages were thought to result from brittle failure and were termed fracture cleavage. However, it has been determined that compressive stresses cause elastic strains which allow preferential solution that is roughly perpendicular to the direction of the applied stress. *See* METAMORPHISM; ROCK MECHANICS.

David B. Bieler

Bibliography. A. J. Barker, *Metamorphic Textures and Microstructures*, 1990; G. J. Borradaile, M. B. Bayly, and C. McA. Powell, *Atlas of Deformational and Metamorphic Rock Fabrics*, 1982; G. H. Davis, Cleavage, foliation and lineation, *Structural Geology of Rocks and Regions*, 2d ed., 1996; F. J. Turner and L. E. Weiss, *Structural Analysis of Metamorphic Tectonites*, 1963; D. S. Wood, Current views of the development of slaty cleavage, *Annu. Rev. Earth Sci.*, 2:1–35, 1974; D. S. Wood et al., Strain and anisotropy in rocks, *Phil. Trans. Roy. Soc. London*, A283:27–42, 1976.

Rock magnetism

The permanent and induced magnetism of rocks and minerals on scales ranging from the atomic to the global, including applications to magnetic field anomalies and paleomagnetism. Natural compasses, concentrations of magnetite (Fe_3O_4) called lodestones, are one of humankind's oldest devices. W. Gilbert in 1600 discovered that the Earth itself is a giant magnet, and speculated that its magnetism might be due to subterranean lodestone deposits. Observations by B. Brunhes in 1906 that some rocks are magnetized reversely to the present Earth's magnetic field, and by M. Matuyama in 1929 that reversely and normally magnetized rocks correspond to different geological time periods, made it clear that geomagnetism is dynamic, with frequent reversals of north and south poles. Nevertheless, permanent magnetism of rocks remains important because it alone provides a memory of the intensity, direction, and polarity of the Earth's magnetic field in the geological past. From this magnetic record comes much of the evidence for continental drift, sea-floor spreading, and plate tectonics. *See* CONTINENTAL DRIFT; GEOMAGNETISM; MAGNET; MAGNETISM; PALEOMAGNETISM; PLATE TECTONICS.

Origin of rock magnetism. The magnetism of rocks arises from the ferromagnetism or ferrimagnetism of a few percent or less of minerals such as magnetite. The magnetic moments of neighboring atoms in such minerals are coupled parallel or antiparallel, creating a spontaneous magnetization M_s. All magnetic memory, including that of computers, permanent magnets, and rocks, is due to the spontaneous and permanent nature of this magnetism. Spontaneous magnetization requires no magnetic field to create it, and cannot be demagnetized.

The magnetism can be randomized on the scale of magnetic mineral grains because different regions of a crystal tend to have their M_s vectors in different directions. These regions are called magnetic domains. Grains so small that they contain only one domain (single-domain grains) are the most powerful and stable paleomagnetic recorders. Larger, multidomain grains can also preserve a paleomagnetic memory, through imbalance in the numbers, sizes, or directions of domains, but this memory is more easily altered by time and changing geological conditions. *See* FERRIMAGNETISM; FERROMAGNETISM; MAGNETISM; MAGNETITE.

Magnetization processes. By applying a large enough magnetic field, domains can be enlarged or reoriented. When the magnetic field is reduced to zero, the original domain sizes and directions are usually not recovered. This phenomenon is called hysteresis (**Fig. 1**). It leaves the entire rock sample with a permanent magnetism, called remanent magnetization M_r. Although M_r is much smaller than M_s of an individual domain, it is large enough to be measured easily with modern magnetometers. *See* MAGNETOMETER.

The permanent magnetism of rocks does not normally respond to a magnetic field as small as that of the Earth, except for a small component called viscous remanent magnetization (VRM) which is acquired slowly over very long times. Instead, the main part of an igneous rock's magnetic memory is acquired as a result of temperature changes, and is called thermoremanent magnetization (TRM). As a rock is heated, its hysteresis loop shrinks (Fig. 1). The magnetization decreases and eventually disappears at the Curie temperature. At the same time, the loop shrinks along the field axis, until at sufficiently high temperatures even a field as small as the Earth's can redirect the domains and create a new M_r. By cooling from this temperature, called the blocking temperature, M_r is stabilized against further field changes. *See* CURIE TEMPERATURE; MAGNETIZATION.

Laws of partial TRM. The thermoremanent magnetization creation process was demonstrated by J. Koenigsberger, E. Thellier, and T. Nagata in the 1930s and 1940s. Thellier in particular described some

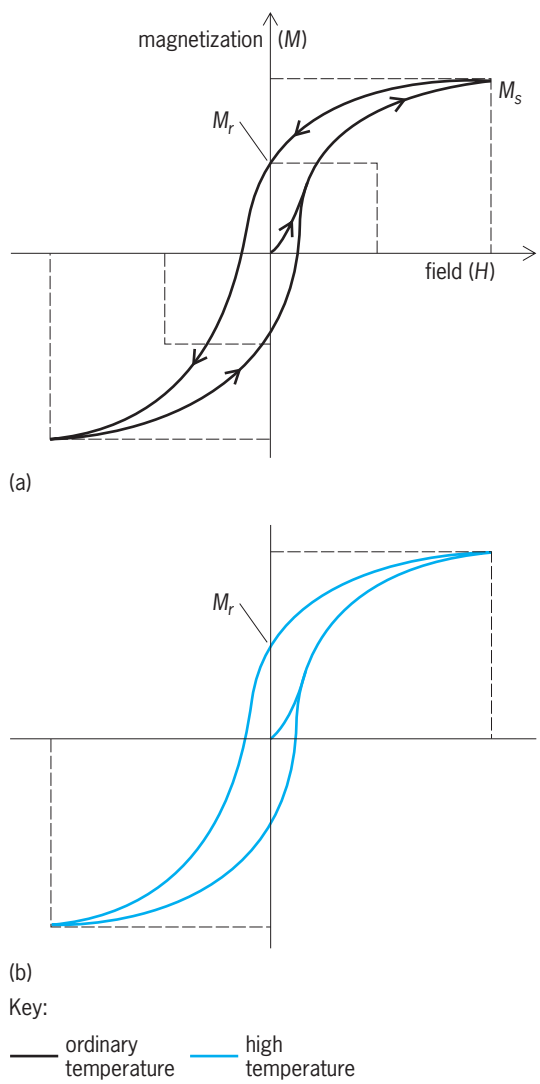

Fig. 1. Magnetic hysteresis curves at ordinary temperature and high temperature approaching the Curie point. M_r = remanent magnetization; M_s = saturation magnetization. The *M-H* curve contracts along both the *M* and *H* axes with heating.

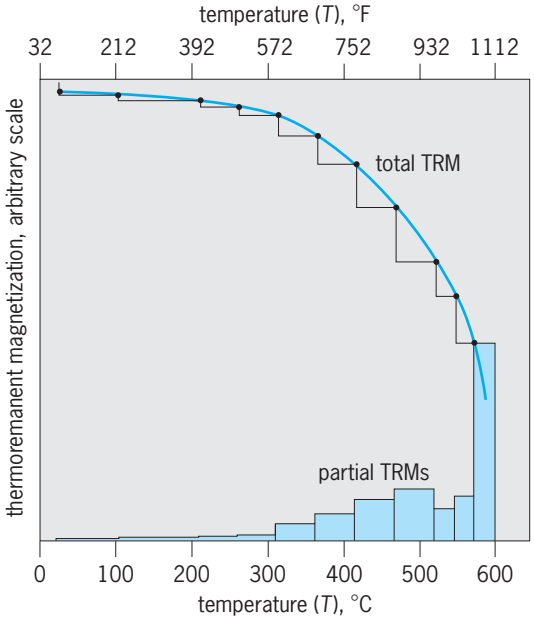

Fig. 2. Illustration of the Thellier laws. The partial thermoremanent magnetizations sum up to give the total thermoremanent magnetization. When the total thermoremanent magnetization is demagnetized by heating in steps, the decrements match the partial thermoremanent magnetizations over corresponding temperature intervals.

remarkable experimental properties of thermoremanent magnetization. The magnetic field may change during cooling. For example, the geomagnetic field may reverse during slow cooling of a rock. Thellier discovered that each blocking temperature interval over which the field is constant, has associated with it a partial thermoremanent magnetization that is independent of all other partial thermoremanent magnetizations created in other intervals. The partial thermoremanent magnetizations can be measured, then erased by heating in a zero field to randomize the domains, without affecting one another. They sum together to give the total thermoremanent magnetization created by cooling from the Curie temperature in a single field (**Fig. 2**).

These observations, called the Thellier laws, are the basis of the paleomagnetic method. They allow paleomagnetists to unravel the often complicated magnetic history of a rock that may have cooled through a series of field reversals, or have been reheated in nature at different times, by stripping away the successive generations of partial thermoremanent magnetizations and isolating the M_r vector of each one. By observing how the natural thermoremanent magnetization of the rock (its natural remanent magnetization, or NRM) compares with the partial TRMs imparted in laboratory experiments, paleomagnetists can deduce the intensity of ancient magnetic fields as well as their vector directions.

The elegant simplicity of Thellier's laws were explained by L. Néel in 1949, the year after his Nobel prize-winning work on ferrimagnetism and antiferromagnetism. Néel showed that a single-domain grain has one and only one blocking temperature. Each partial thermoremanent magnetization is associated with a separate population of single-domain grains in a rock. The independence and additivity of partial thermoremanent magnetizations is made clear by this simple picture. However, the thermoremanent magnetization of multidomain grains, which also approximately obey the Thellier laws, is not so easily understood because the domains in any grain can contribute to partial thermoremanent magnetizations in several different blocking temperature intervals. *See* ANTIFERROMAGNETISM.

Induced magnetization and magnetic anomalies. Although rock magnetism emphasizes remanent or permanent magnetization, the temporary magnetization induced by the Earth's field is also of importance. For most types of rocks in the Earth's crust, induced magnetization outweighs natural remanent magnetization. This is sometimes described by a Koenigsberger ratio of remanent to induced magnetization, Q, of less than 1. Another reason for the importance of induced magnetization is that it remains constant or increases with heating, whereas M_r decreases with heating (Fig. 1). Rocks in the deep crust, where temperatures approach the Curie point, have smaller Q values, that is, more important induced magnetizations, than similar rocks at the surface. *See* GEOPHYSICAL EXPLORATION.

The large magnetic anomalies recently discovered by the *Mars Global Surveyor* satellite cannot be due to induced magnetization because the present global magnetic field of Mars is very small compared to that of the Earth. Remanent magnetization must be responsible. The magnetic mineral or minerals involved have not yet been identified.

Magnetic minerals. Five groups of magnetic minerals are significant sources of induced and remanent magnetization. On the Moon and in most meteorites, metallic iron is most important. On Earth, titanomagnetites ($Fe_{3-x}Ti_xO_4$) or their oxidized equivalents, titanomaghemites, are usually dominant. Fresh basaltic lavas erupting at mid-ocean ridges, the source of new sea floor, contain titanomagnetites with about 60% titanium (Ti). Older sea floor and continental rocks tend to contain almost titanium-free magnetite (Fe_3O_4) or its oxidized equivalent, maghemite (γFe_2O_3). Titanohematites ($Fe_{2-y}Ti_yO_3$) occur as primary minerals in some lavas and plutonic rocks. Almost titanium-free hematite (αFe_2O_3) is a common oxidation product in rocks, sediments, and soils, as are iron oxyhydroxides such as goethite (FeOOH), which ultimately dewaters to hematite. Both hematite and goethite are weakly magnetic antiferromagnets, but are significant sources of natural remanent magnetization because they occur in such large quantities in weathered materials. The iron sulfides pyrrhotite (Fe_7S_8) and greigite (Fe_3S_4) are fairly common in terrestrial igneous and sedimentary rocks. Troilite (FeS) occurs in lunar rocks and meteorites. *See* GOETHITE; HEMATITE; IGNEOUS ROCKS; METEORITE; PYRRHOTITE; SEDIMENTARY ROCKS.

Not all magnetic minerals are of inorganic origin. Magnetotactic bacteria produce chains of single-domain magnetite or greigite crystals, which they use as compasses. These crystals become magnetic fossils after the bacteria die. When deposited in soils and sediments, the crystals are an important source of stable natural remanent magnetization of an origin quite different from thermoremanent magnetization. Many other organisms, among them bees, birds, and fish, produce magnetic crystals, but their navigational or other functions are not always clear. *See* BIOMAGNETISM.

Sea-floor spreading and linear magnetic anomalies. When lavas erupt at mid-ocean ridges, the titanomagnetites they contain cool from above the Curie temperature. The geomagnetic field produces thermoremanent magnetization during this cooling (**Fig. 3**). Periodically, at intervals of about 1 million years, the geomagnetic field reverses. The thermoremanent magnetization acquired by lavas erupted after such a reversal is opposite in direction to thermoremanent magnetization acquired before the reversal. Because the sea floor is slowly spreading away from mid-ocean ridges, older lavas are found farther from ridges than are recently erupted lavas. For this reason, the spreading sea floor acts like a tape recorder, with successive bands of oppositely directed thermoremanent magnetization recording successive field reversals (Fig. 3). Mid-ocean ridges are hundreds or thousands of kilometers long while the bands are typically only tens of kilometers wide, because spreading rates are only a few centimeters per year. The bands therefore form "magnetic stripes" paralleling the ridge axis on either side.

This picture of sea-floor magnetism was proposed by F. Vine and D. Matthews in 1963. Their main evidence was the widespread linear pattern of magnetic field anomalies measured over the oceans, which is quite different from the irregular shapes of anomalies over the continents. The linear magnetic anomalies arise from fields produced by the magnetic stripes on the sea floor, which alternately add to or subtract from the main geomagnetic field, producing measurable anomalies. This picture convinced most geologists of the reality of sea-floor spreading and ushered in the era of plate tectonics. *See* MARINE GEOLOGY; MID-OCEANIC RIDGE.

Paleomagnetic evidence for continental drift and plate tectonics. Once geologists were convinced that thermoremanent magnetization of sea-floor basalts faithfully records sea-floor spreading, and therefore represents a trustworthy magnetic memory of the geomagnetic field millions of years earlier when the basalts formed, they looked more favorably on paleomagnetic measurements of the natural remanent magnetization of continental rocks. This natural remanent magnetization is much more varied than that of the sea floor. It includes thermoremanent magnetization of igneous rocks and depositional remanent magnetization (DRM) of sedimentary rocks produced by partial alignment of magnetized grains in wet sediments. These are primary natural remanent magnetizations dating from the formation of the rock. Rocks of all types can also acquire secondary natural remanent magnetizations. Partial thermoremanent magnetization due to later burial and reheating and chemical remanent magnetization (CRM) resulting from changes in the magnetic minerals are the most common types. Metamorphic rocks which have experienced a turbulent tectonic history may have several generations of partial thermoremanent magnetization and chemical remanent magnetization. *See* METAMORPHIC ROCKS.

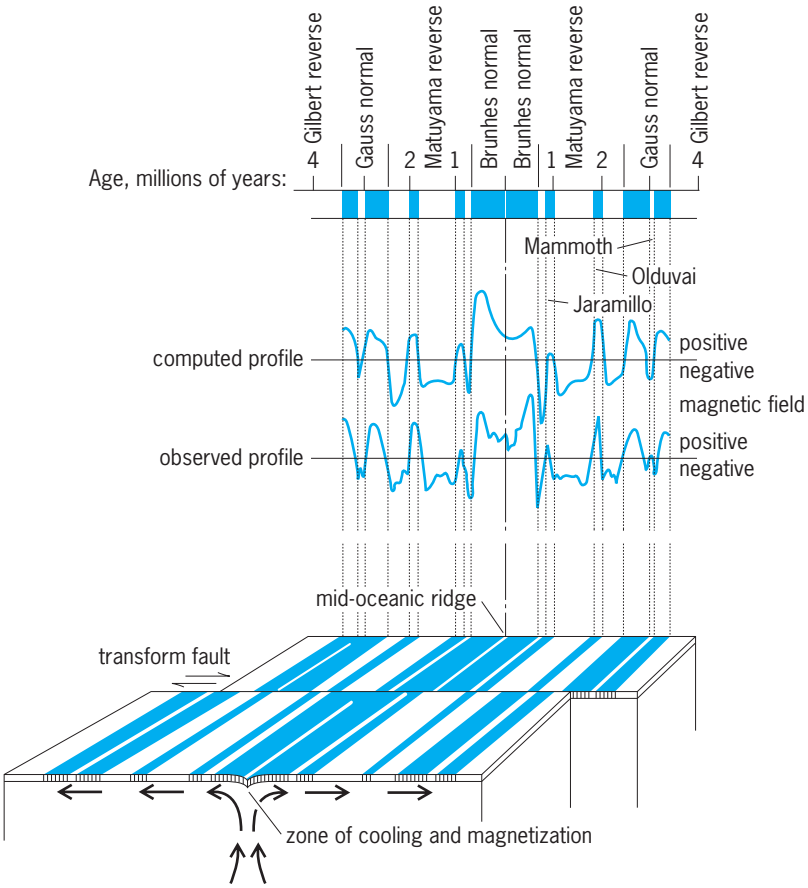

Fig. 3. Vine and Matthews' model of linear magnetic anomalies. Lavas cooling at a mid-ocean ridge acquire thermoremanent magnetization. The combination of spreading away from the ridge and periodic reversals of the geomagnetic field imprints magnetic stripes, alternately normal and reverse, on the sea floor. At the sea surface, the magnetic stripes create alternating positive and negative magnetic field anomalies, shown in the profiles.

Part of the earlier skepticism about paleomagnetic data arose from the complexity of natural remanent magnetization in continental rocks. The paleomagnetist's task is to separate and successively erase different layers of natural remanent magnetization, associating an age with each. The usual methods for such "cleaning" are demagnetization, in steps, of the natural remanent magnetization by either heating or alternating magnetic fields in the laboratory. *See* DEMAGNETIZATION.

The methodology is different from that used for measuring sea-floor spreading recorded by magnetic

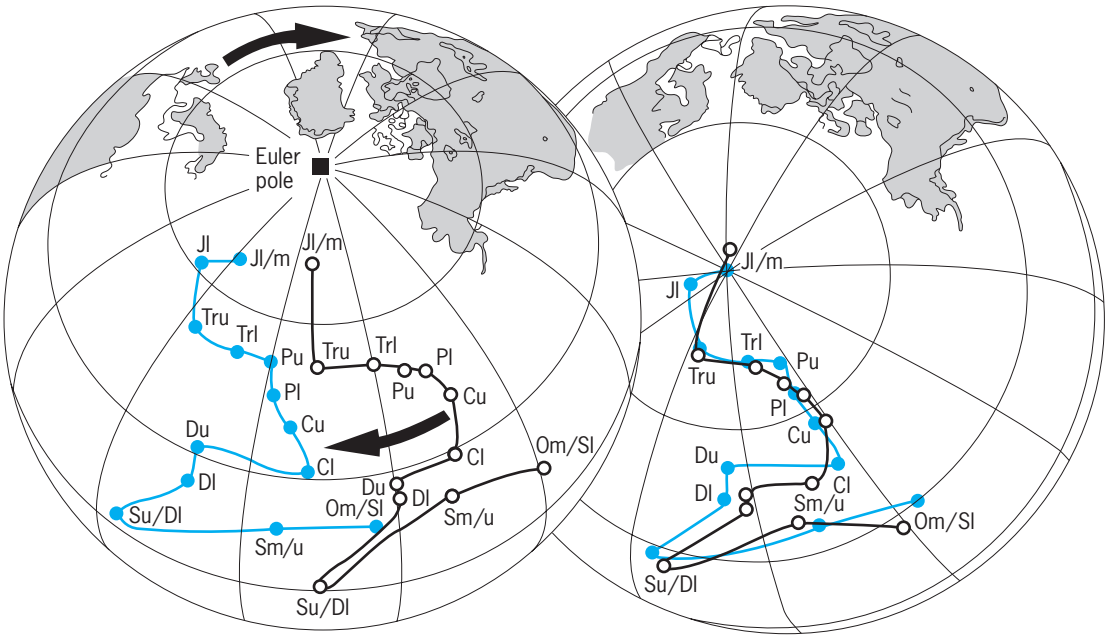

Fig. 4. Tracks of the positions of paleomagnetic poles for European and North American rocks with these continents in their present positions (left) and after closing the Atlantic Ocean (right). The polar tracks diverge before closure but become coincident with the continents in their Pangea configuration. The symbols along the tracks refer to geological time periods: J, Jurassic; Tr, Triassic; P, Permian; C, Carboniferous; D, Devonian; S, Silurian; O, Ordovician; u, upper; m, middle; l, lower. (*After R. Van der Voo, Phanerozoic paleomagnetic poles from Europe and North America and comparisons with continental reconstructions, Rev. Geophys., 28:167–206, 1990*)

field anomalies. Instead of remote sensing of a magnetic signal, in paleomagnetism the magnetized rocks themselves are sampled and their magnetic moments measured. Each sample acts as a fossil compass, recording both the direction and the intensity of the geomagnetic field. Except for very recent rocks, the natural remanent magnetization is generally not parallel to the present geomagnetic field vector. Either the magnetic poles of the Earth have moved (polar wander) or the continents have moved relative to fixed poles (continental drift). Rocks from a single continent define a consistent magnetic pole track with increasing age, but different continents have different pole tracks. However, when the continents are fitted together into a Pangea supercontinent, as suggested by A. Wegener in 1912, large sections of the pole tracks for different continents come together as well (**Fig. 4**).

Continental drift and sea-floor spreading are different manifestations of plate tectonics. Each continent and its adjacent sea floor form a lithospheric plate, which moves rigidly as a unit over the Earth's surface. The plate boundaries do not necessarily correspond to continental margins. Instead they form a continuous linkage of mid-ocean ridges, subduction zones, and transform faults. *See* CONTINENTS, EVOLUTION OF; LITHOSPHERE; SUBDUCTION ZONES; TRANSFORM FAULT.

Magnetic field reversals and magnetostratigraphy. Reversals of the Earth's magnetic field are synchronous worldwide. This fact was first shown by A. Cox, R. Doell, and B. Dalrymple in the 1960s by dating lava flows from different parts of the world recording the same sequences of reversals. Reversals therefore are time markers in the geological record and can be used for correlation of rocks and sediments locally, regionally, or globally. This branch of paleomagnetism is called magnetostratigraphy. Long sections of ocean sediments cored by drilling ships and similar sections of older sedimentary rocks now exposed on the continents have yielded sequences of normal and reversed natural remanent magnetizations that match the magnetic stripe sequence from linear magnetic anomalies (**Fig. 5**). Many important geological boundaries, including the Cretaceous-Tertiary boundary around 65 million years ago, at which the dinosaurs became extinct, and the Brunhes-Matuyama boundary (the most recent geomagnetic reversal, 780,000 years ago), an important time marker in human evolution, are now much more accurately dated geomagnetically than they were previously from fossil records. *See* CRETACEOUS; GEOLOGIC TIME SCALE; TERTIARY.

Magnetic records from lake sediments, soils, and loess. When sedimentary deposition is rapid, as in lakes, soil formation, and windblown loess deposits, it is possible to measure the fine-scale variation of geomagnetic field direction and intensity, called paleosecular variation. In some rapidly deposited marine sediments, it is also possible to measure magnetic field variations between and even during field reversals. The geomagnetic signal is often modulated by climatic influences such as paleotemperature, rainfall variation, and paleowind direction. Therefore, paleomagnetic records from sediments and soils are

proving to be valuable archives of information about global climate changes in the past. *See* CLIMATE HISTORY; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; LOESS.

Paleointensity variation. Although depositional remanent magnetization of sediments gives some indication of the past variation of the Earth's magnetic field intensity, precise records rely on the thermoremanent magnetization of rocks or baked archeological materials such as pottery, bricks, and kilns. Such records are a method of archeological and geological dating. Another important purpose of paleointensity studies is to take the pulse of the geodynamo in the Earth's core, which is believed to be the source of most of the geomagnetic field. The changing intensity of the field before, during, and after reversals helps advance the understanding of how the geodynamo works. *See* DATING METHODS; GEODYNAMO; MAGNETOHYDRODYNAMICS.

Paleomagnetism of other planets. The natural remanent magnetizations of meteorites bring to Earth a record of magnetic fields from other times and places in the solar system. Direct sampling of rocks on the Moon during the *Apollo* scacecraft missions of the 1970s demonstrated that many carry very ancient natural remanent magnetizations that record a declining lunar magnetic field strength over time (**Fig. 6***a*). On the basis of magnetic field anomalies measured by the *Mars Global Surveyor* satellite, portions of the crust of Mars are intensely magnetized, about ten times more strongly than on Earth, with a lineated pattern reminiscent of magnetic stripes in the Earth's oceans. As with the Moon, the present global field of Mars is very small. Only the Earth, among the terrestrial planets, has preserved a vigorous geodynamo as evidenced by its paleointensity record covering practically all of geological history (Fig. 6*b*).

David J. Dunlop

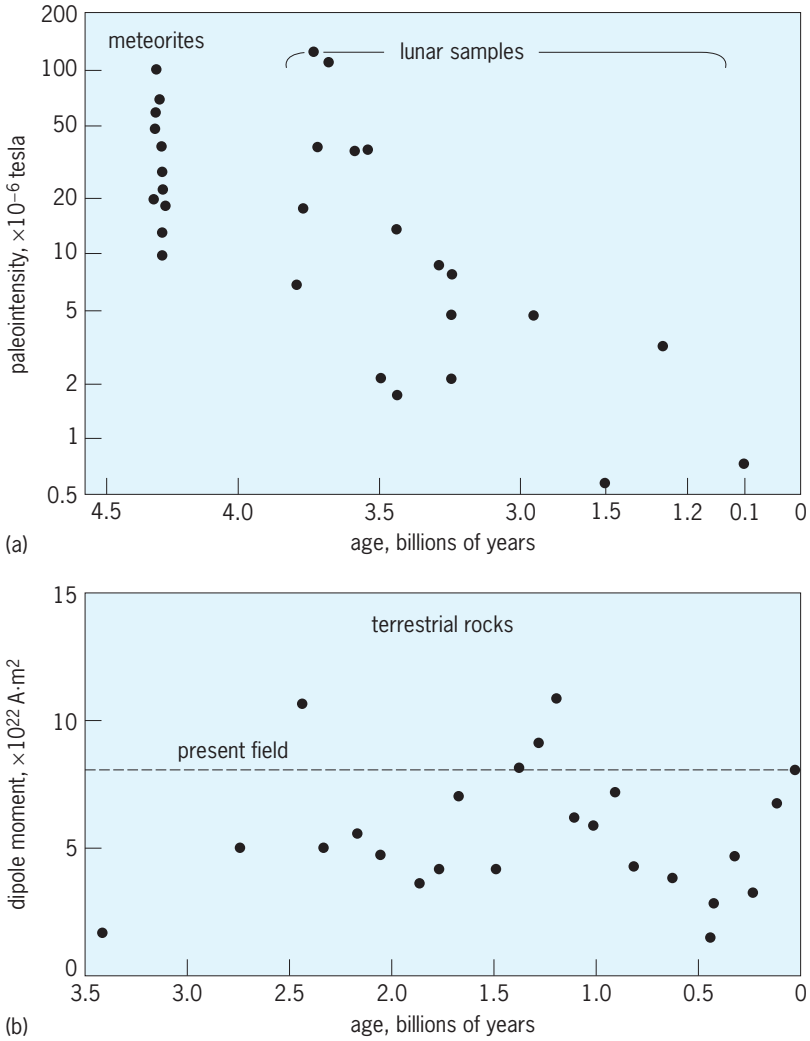

Fig. 6. Paleointensities recorded by meteorites, lunar samples, and terrestrial rocks throughout geological history. The Earth's magnetic field has fluctuated with time, but its present strength is comparable to the range of values 2.5–3 billion years ago. The Moon's global field has decreased with time and is negligible today. For comparison between the two graphs, the present field of the Earth is about 30 × 10^6 tesla near the Equator and about 60 × 10^{-6} tesla near the Poles, similar to average meteorite and early lunar values.

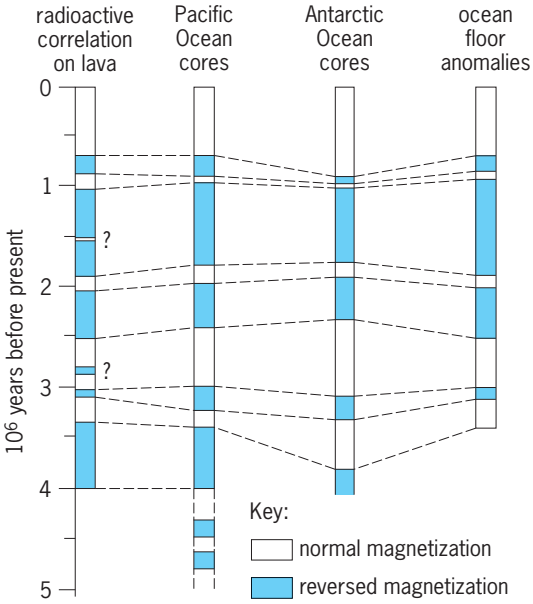

Fig. 5. Sequence of reversals of the geomagnetic field in late Cenozoic time.

Bibliography. R. F. Butler, *Paleomagnetism*, Blackwell, 1992; J. E. P. Connerney et al., Magnetic lineations in the ancient crust of Mars, *Science*, 284: 794–798, 1999; A. Cox, G. B. Dalrymple, and R. R. Doell, Reversals of the Earth's magnetic field, *Sci. Amer.*, 216:44–54, 1967; D. J. Dunlop and Ö. Özdemir, *Rock Magnetism: Fundamentals and Frontiers*, Cambridge University Press, 1997; W. Gilbert, *De Magnete*, Dover Reprints, 1958; B. A. Maher and R. Thompson (eds.), *Quaternary Climates, Environments and Magnetism*, Cambridge University Press, 2000; R. T. Merrill, M. W. McElhinny, and P. L. McFadden, *The Magnetic Field of the Earth*, Academic Press, 1996; L. Néel, Some theoretical aspects of rock magnetism, *Adv. Phys.*, 4:191–243, 1955; N. D. Opdyke and J. E. T. Channell, *Magnetic Stratigraphy*, Academic Press, 1996.

Rock mechanics

The scientific discipline that deals with understanding how rocks deform and fail due to natural and human-induced forces at scales ranging from micrometers to kilometers. Rock mechanics is also the engineering discipline for designing and stabilizing underground and surface excavations and rock foundations.

The deformation and failure of rocks influence our lives in many ways. Structures that are built with or into rocks include highways, dams, bridges, tunnels, mines, and water and petroleum wells, to name some. These structures must be designed to remain stable over their expected lifespan. Natural occurrences of rock failure include earthquakes, volcanic eruptions, and landslides. These geologic hazards must be understood and methods for predicting rock failure events must be developed. *See* ENGINEERING GEOLOGY; GEODYNAMICS.

Intact rock, discontinuities, and rock mass. Rock, such as granite, sandstone, or gneiss, is defined as an aggregate of minerals. Small samples of rock that are tested in the laboratory are referred to as intact rock, in contrast to the larger volumes of rock that accommodate highways, tunnels, and other engineering structures. These larger volumes of rock are referred to as rock masses and contain discontinuities as well as intact rock. Discontinuities in rock include joints, faults, bedding planes, and other types of large-scale rock fractures. A typical rock mass is shown in **Fig. 1**. Joints are the most common type of discontinuity and are found in almost all rocks. Joints usually occur in sets defined by their common orientation. Several joint sets can be seen in Fig. 1 (one set sloping to the right and one set sloping to the left). A stereonet is often used to display joint orientation information and assist in identifying joint sets. *See* JOINT (GEOLOGY); PETROFABRIC ANALYSIS; PETROLOGY; ROCK.

Fig. 1. Typical rock mass containing numerous discontinuities. The person in picture is conducting a scanline survey, where a tape is placed across the rock and detailed information is gathered on the fractures that intersect the scanline.

Rock-mass characterization and classification. Discontinuities cause the strength of a rock mass to be much weaker than the intact rock. Before an engineering structure is made in a rock mass, it is very important to characterize the properties of the discontinuities, a process called rock-mass characterization. Rock-mass characterization involves the careful measurement of discontinuity type, orientation, spacing, length, roughness, fill, and other properties. This information can be collected in the field from existing rock outcrops or from borehole drill core. Figure 1 shows one method of collecting rock-mass information in the field, referred to as a scanline survey. *See* DRILLING, GEOTECHNICAL.

Rock-mass classification is another aspect of rock-mass characterization. Rock-mass classification uses subsets of the rock characterization data to determine numerical or descriptive ratings. These ratings can then be used to estimate underground support requirements, rock-mass strength and deformation, and other engineering design criteria. Rock-mass classification schemes represent a convenient way to assimilate the rock characterization data into a single number or distribution of numbers. The rock quality designation (RQD) considers one type of information: the total percent of a rock core containing pieces of core greater than 10 cm (4 in.). It gives a rating from 0 to 100 (worst to best). The rock mass rating (RMR) system considers 10 types of information: intact rock strength, rock quality designation, joint spacing, joint length, joint roughness, joint fill, joint weathering, joint aperture, water, and joint orientation. Individual ratings are determined for each information type. The rock-mass rating, which is the sum of these individual ratings, is a number that ranges from 0 to 100 (worst to best). Other popular rock-mass classification schemes used today include the mining rock-mass rating (MRMR), the slope mass rating (SMR), the geologic strength index (GSI), the rock quality index (Q), and the rock-mass index (RMi).

Stress and strain. Rocks deform and fail due to forces, including gravity, tectonic, seismic, water, and thermal. Force is not a useful measure of rock strength because it depends on the size of the rock specimen subjected to the force; that is, it takes a small force to break a small rock sample and a larger force to break a larger sample of the same rock type. Instead, we use stress as a measure of rock strength. Rock failure will occur at a point in a rock when the stress at that point exceeds the strength of the rock. Stress σ has the unit of force per unit area. In SI units the stress unit is pascals (newtons/m^2) and in English units the stress unit is psi (pounds/in.^{2}).

Strain is a measure of the deformation of a rock. Strain is defined as the change in length divided by the original length, and is unitless. The symbol ε is normally used for strain. Under conditions in the shallow crust, rock is brittle and cannot strain very much before failure. A strain of 0.05 (0.5%) is usually

sufficient to cause most rocks to fail. *See* STRESS AND STRAIN.

Strength of intact rock. The strength of a small sample of rock can be determined from laboratory tests. Many tests usually are done to determined the strength of a particular rock type, because the strength of rock in tension is different than its strength in compression and because the strength varies from one location to another (heterogeneity). Common laboratory tests to determine the strength of rock include the uniaxial and triaxial compression tests, the Brazilian disk tension test, and the point load test. **Figure 2** shows a point load test.

In the uniaxial compression test, a force is applied in only one direction, as shown in **Fig. 3***a*. The rock strength under this type of loading is referred to as the uniaxial compressive strength (UCS). The uniaxial compressive strength is the most common index property for intact rock strength. **Table 1** shows the uniaxial compressive strength for different rock types and UCS ranges for the categories of very weak, weak, medium, strong, and very strong rock.

TABLE 1. Classification of rock hardness

Strength classification	UCS (MPa)	Typical rock types
Very weak	10–20	weathered and weakly compacted sedimentary rocks
Weak	20–40	weakly cemented sedimentary rocks, schist
Medium	40–80	competent sedimentary rocks; some low-density coarse-grained igneous rocks
Strong	80–160	competent igneous rocks; some metamorphic rocks and fine-grained sandstones
Very strong	160–320	quartzites; dense fine-grained igneous rocks

Adapted from P. B. Attewell and I. W. Farmer, *Principles Of Engineering Geology*, Chapman and Hall, London, 1976.

The uniaxial compression test does not adequately represent the actual stresses in the ground. In an actual rock mass, the stress state is a result of forces in more than one direction. The triaxial compression test is a better representation for the actual stresses in the ground. In the triaxial test, a confining stress is applied in addition to the axial stress as shown in Fig. 3*b*. During the test, the confining stress is held at a constant value and then the axial stress is increased until the sample fails. Rock strength increases with increasing confining stress and the friction angle ϕ is the rock parameter that describes the increase in strength with increasing confining stress. A typical friction angle for a strong rock is 40–60° but can be as low as 20° for weak rock. With the two strength parameters UCS and friction angle, the Mohr-Coulomb criterion predicts a linear relationship between the triaxial compressive strength (σ_1) and the confining stress (σ_3), as shown in Eq. (1).

$$\sigma_1 = \text{UCS} + \sigma_3 \tan^2(45 + \phi/2) \qquad (1)$$

Fig. 2. Point load test on a core sample of rock. The results of the point load test can be used to estimate the uniaxial compressive strength (UCS) of the intact rock.

Strength of discontinuities. Discontinuities represent preexisting weakness planes in the rock. Sliding will occur along discontinuities when the shear stress τ along the discontinuity plane exceeds the shear strength of the discontinuity. The strength of a discontinuity is frictional and depends on the normal stress acting perpendicular to the plane, σ_n. The Coulomb criterion predicts a linear relationship between the shear strength of a discontinuity (τ)

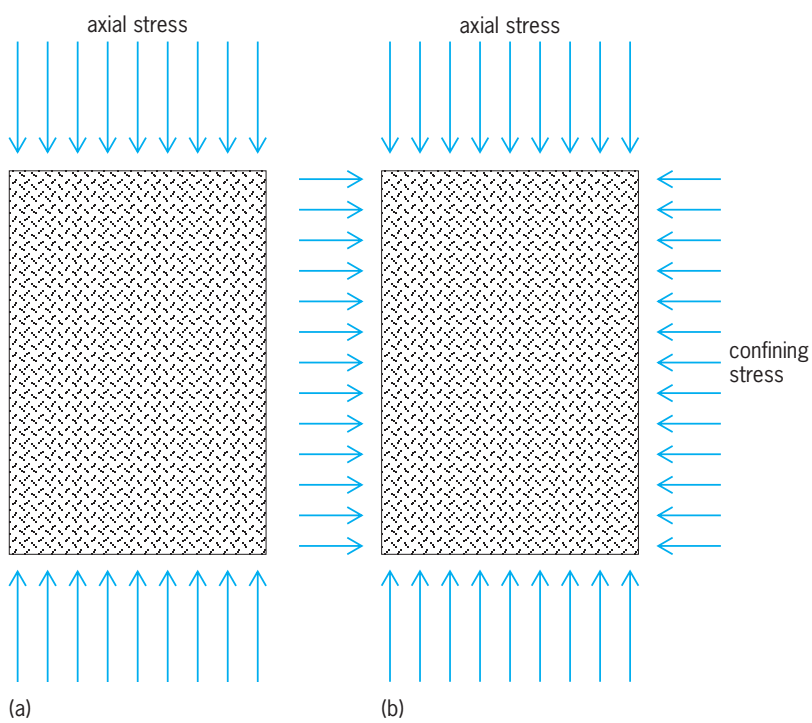

Fig. 3. (*a*) Uniaxial compression test, where an axial stress is applied to a core sample of rock and increased until failure. (*b*) Triaxial compression test, where a constant confining stress is first applied, then the axial stress is increased until failure.

and the normal stress (σ_n) as shown in Eq. (2).

$$\tau = \sigma_n \tan \phi_j \tag{2}$$

Note that the friction angle for a discontinuity is not the same as that for the intact rock, and thus the symbol ϕ_j is used. A typical friction angle for a smooth unweathered discontinuity is 30°. Faults often have frictional angles as low as 10°, which can result in stability problems encountered in excavations or foundations. The roughness of the discontinuity can significantly increase the friction angle for a discontinuity to values above 50°. However, under very high normal stresses, the roughness does not contribute as much to the friction angle since the asperities that make up the roughness can be sheared off during sliding.

Strength of rock mass. Both the strength of the intact rock and the strength of the discontinuities influence the strength of a rock mass. If only one or two sets of joints occur in a rock mass, the strength of the rock mass may be anisotropic, which means that the strength is orientation dependent. Shale is an example of a commonly encountered rock that has anisotropic strength properties. If three or more joint sets occur in a rock mass, then it is usually assumed that the rock-mass strength is isotropic (same strength in all directions). The strength of a rock mass will depend on all the information that was collected during rock-mass characterization, such as the number of joint sets, joint spacing, joint roughness, joint fill, and other properties. Because of this, empirical relationships for rock-mass strength often use the rock-mass classification directly. For instance, the 1992 version of the popular Hoek and Brown rock-mass strength criterion uses the rock mass rating (RMR), as given in Eq. (3), where m are parameters that depend on the rock type and condition.

$$\frac{\sigma_1}{\mathrm{UCS}} = \frac{\sigma_3}{\mathrm{UCS}} + \sqrt{m\frac{\sigma_3}{\mathrm{UCS}} + s} \tag{3}$$

Because the strength of a rock mass cannot be measured directly, empirical methods for estimating rock-mass strength are widely used today. With the advent of sophisticated numerical methods, it is now possible to estimate rock-mass strength using numerical modeling.

In-situ stress. To determine how a rock mass will behave in an excavation or foundation, it is necessary to estimate the stresses in the ground prior to engineering works, referred to as in-situ stresses. In general, the stresses in the ground increase with depth below the surface, and the vertical stress can be estimated by the weight of the material above a location of interest. Most rocks have a density of about .025 MN/m^3 (which corresponds to 0.025 MPa per m depth), and therefore at a depth of 1000 m (3300 ft), the vertical stress will be about 25 MPa. The horizontal stresses can be less than or greater than the vertical stress at a particular depth. Normal faulting tectonic regimes are associated with stress states where the horizontal stress is less than the vertical stress, and thrust faulting regimes are associated with stress states where the horizontal stress is greater than the vertical stress. In addition, high horizontal stresses can occur in regions where a large amount of geologic removal of rock has occurred. There are many techniques available for measuring the in-situ stress state, including overcoring, hydraulic fracturing, and borehole breakouts.

Induced stresses around underground excavations. An excavation made in a rock mass will perturb the in-situ stress state. It will cause a concentration of stress at the boundaries of the excavation, particularly in the case of an underground excavation [more precisely, referring to Eq. (1) or (3), it results in an increase in σ_1 and a decrease in σ_3 around the boundary]. These induced stresses may result in rock failure. Failure around an underground excavation usually starts at the boundary and progresses into the rock mass, as shown in **Fig. 4*a***. In some cases, the

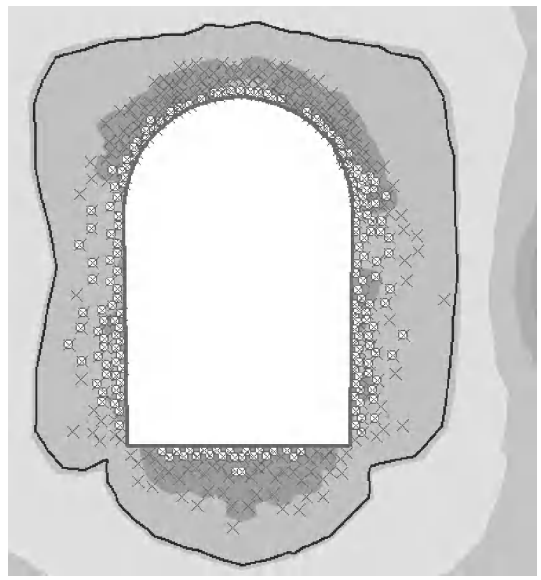

(a)

(b)

Fig. 4. (*a*) Example of the failure zone around an underground excavation (×, shear failure; o, tensile failure), calculated using the Rocscience Phase2 computer program. (*b*) Example of a shallow underground excavation where the failure mode is either gravity-induced block falls or block sliding along discontinuities.

TABLE 2. Recommendations for rock-support and excavation sequence based on the rock-mass rating (RMR) for the rock mass

Rock-mass class	Excavation	Rock bolts (20 mm diameter, fully grouted)	Shotcrete	Steel sets
I. Very good rock RMR: 81–00	Full face, 3 m advance	Generally no support required except spot bolting		
II. Good rock RMR: 61–80	Full face, 1–1.5 m advance; complete support 20 m from face	Locally, bolts in crown 3 m long, spaced 2.5 m with occasional wire mesh	50 mm in crown where required	None
III. Fair rock RMR: 41–60	Top heading and bench 1.5–3 m advance in top heading; commence support after each blast; complete support 10 m from face.	Systematic bolts 4 m long, spread 1.5–2 m in crown and walls with wire mesh in crown	50–100 mm in crown and 30 mm in sides	None
IV. Poor rock RMR: 21–40	Top heading and bench 1.0–1.5 m advance in top heading; install support concurrently with excavation, 10 m from face.	Systematic bolts 4–5 m long, spaced 1–1.5 m in crown and walls with wire mesh	100–150 mm in crown and 100 mm in sides.	Light to medium ribs spaced 1.5 m where required
V. Very poor rock RMR: < 20	Multiple drifts 0.5–1.5 m advance in top heading; install support concurrently with excavation; Shortcreate as soon as possible after blasting	Systematic bolts 5–6 m long, spaced 1–1.5 m in crown and walls with wire mesh; bolt invert	150–200 mm in crown, 150 mm in sides, and 50 mm on face	Medium to heavy ribs spaced 0.75 m with steel lagging and forepoling if required; close invert

Adapted from http://www.rocscience.com/hoek/PracticalRockEngineering.asp.

induced stresses may increase the stability of the excavation. This is often the case in relatively competent rock mass in a shallow excavation where the dominant mode of failure is gravity, causing blocks of rock to fall from the roof or slide from the walls, as shown in Fig. 4*b*. In this case, the induced stresses squeeze the rock blocks together and prevent them from sliding [more precisely, referring to Eq. (2), it results in an increase in J_b on the discontinuities]. The induced stresses around the opening depend on the in-situ stresses and the shape of the excavation. Sharp corners have particularly high stress concentrations, as do tunnel intersections. It is often possible to increase the stability of an underground excavation by changing its size, shape, or orientation. A technique used in weak rocks is to excavate multiple drifts rather than a single "full face" excavation to minimize induced stresses and to allow rock support to be installed between excavations. For simple geometries, closed- form solutions are available for the induced stresses. One of the most-used formulas in rock mechanics is the Kirsch solution for the induced stresses around a circular underground opening. For more complicated geometries (for example, the three-dimensional geometry formed by two intersecting tunnels), numerical methods are used to estimate induced stresses and subsequent rock failure.

Supporting unstable rock masses. In many circumstances, a rock engineering design will be unstable without artificial rock support. Rock support technologies include rock bolts, cable bolts, shotcrete, concrete, steel sets, wire mesh, and other methods. Rock bolts are the most popular method for underground stabilization. They consist of steel or aluminum rods 1–6 m (3–20 ft) in length inserted into predrilled holes, with bolt spacing ranging 0.5–3 m (1.5–10 ft). Rock-bolt-support forces range 0.04–0.3 MN per bolt. Friction bolts support rock though frictional forces between the bolt and the drill hole. Mechanical anchor bolts are anchored to the rock at the end of the drill hole and often grouted for additional support. Resin bolts use a quick-setting resin and steel dowels to support the rock. Wire mesh is often used in addition to bolts to contain small rock blocks between the bolts. Shotcrete is a form of sprayed concrete that is very effective in stabilizing rock excavations. The shotcrete mix contains cement, aggregate, and various agents to improve strength and ductility. Shotcrete thickness varies 25–300 mm (1–12 in.) and has a final (28 day) strength of 25–35 MPa. Various techniques are available to determine the type and amount of support required to stabilize a specific surface and underground design. Empirically based tables are commonly used to determine the support requirements, and most of these tables are based on rock-mass classification schemes. An example of an underground-excavation support table using the rock mass rating is shown in **Table 2**. The other method for estimating the support requirements for a particular rock engineering design is using numerical methods. Modern numerical programs allow detailed support designs to be modeled and the resulting deformation and failure of the rock mass to be analyzed. *See* MODEL THEORY.

Modern technologies and research. New technologies and research are essential elements of the field of rock mechanics. The most important new technologies that have been developed in the past 20 years are sophisticated numerical programs for modeling rock-mass deformation and failure. The numerical programs that are used in rock mechanics include finite-element models, boundary-element models, distinct-element models, finite difference models, micromechanical models, and many others. These include two- and three-dimensional models, constitutive relationships such as strain hardening and strain softening, large-strain, and coupled hydrothermomechanical models. New technologies for rock-mass characterization are just emerging.

In general, these are existing technologies that are being applied to rock-mass characterization for the purpose of increasing accuracy, safety, access, and human bias. These technologies include digital image processing, ground-based lidar, Geographical Information Systems (GIS), Geographical Positioning Systems (GPS), remote sensing, and geophysics. The important research areas include neural networks, scale effects (including fractals), coupled mechanical/hydraulic/thermal/chemical models, fracture mechanics, fracture flow, block theory, constitutive modeling, wellbore stability, contaminant transport/multiphase flow, and rock mechanics for long-term underground nuclear waste storage. *See* FINITE ELEMENT METHOD; GEOGRAPHIC INFORMATION SYSTEMS; GEOPHYSICS; IMAGE PROCESSING; LIDAR; REMOTE SENSING.

John M. Kemeny

Bibliography. B. H. G. Brady and E. T. Brown, *Rock Mechanics for Underground Mining*, 3d ed., Allen and Unwin, London, 2005; R. E. Goodman, *Introduction to Rock Mechanics*, 2d ed., Wiley, 1989; E. Hoek and E. T. Brown, *Underground Excavations in Rock*, Institution of Mining and Metallurgy, London, 1980; E. Hoek and J. W. Bray, *Rock Slope Engineering*, 3d ed., Institution of Mining and Metallurgy, London, 1981; E. Hoek, P. K. Kaiser, and W. F. Bawden, *Support of Underground Excavations in Hard Rock*, Balkema, Rotterdam, 1995; J. A. Hudson and J. P. Harrison, *Engineering Rock Mechanics: An Introduction to the Principles*, Pergamon, 1997; J. C. Jaeger and N. G. W. Cook, *Fundamentals of Rock Mechanics*, 2d ed., Chapman and Hall, London, 1979; S. D. Priest, *Discontinuity Analysis for Rock Engineering*, Chapman Hall, London, 1993.

Rock varnish

A dark to dark-brown coating on rock surfaces exposed to the atmosphere (**Fig. 1**). Rock varnish is probably the slowest-accumulating sedimentary deposit, growing at only a few micrometers to tens of micrometers per thousand years. Its thickness ranges from less than 5 μm to over 600 μm, and is typically 100 μm or so. Although found in all terrestrial environments, varnish is mostly developed and well preserved in arid to semiarid deserts; thus, another common name is desert varnish. Rock varnish is composed of about 30% manganese and iron oxides, up to 70% clay minerals, and over a dozen trace and rare-earth elements. The building blocks of rock varnish are mostly blown in as airborne dust. Although the mechanism responsible for the formation of rock varnish remains unclear, two hypotheses have been proposed to explain the great enrichment in manganese within varnish (typically over 50 times compared with the adjacent environment such as soils, underlying rock, or dust). The abiotic hypothesis assumes that small changes in pH can concentrate manganese by geochemical processes. The biotic hypothesis posits that bacteria, and perhaps other microorganisms, concentrate manganese; this is supported by culturing experiments and direct observations of bacterial enhancement of manganese. *See* SEDIMENTARY ROCKS.

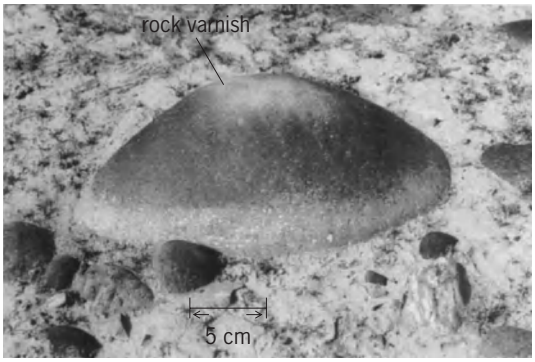

Fig. 1. Varnished river cobble in western China.

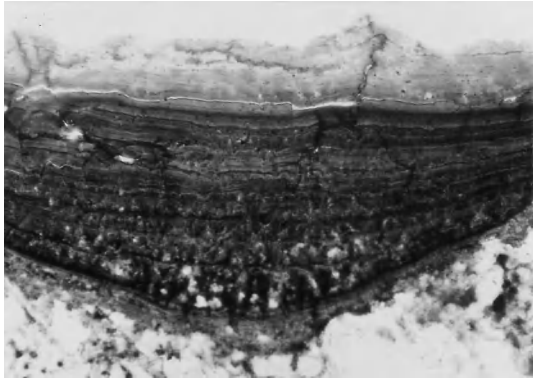

Fig. 2. Microlaminations seen in a varnish ultrathin section under a light microscope. Maximum varnish thickness is 150 μm.

Rock varnish has been scientifically studied since the 1950s largely because of the possibility of using its development as an indicator of time or environmental change. One outstanding technique in recent research is varnish microlaminations. Microlaminations are seen only when the varnish is shaved thin enough (5–10 μm) to see through with a light microscope (**Fig. 2**). The dark layers are rich in manganese (but poor in silicon, aluminum, and iron), while the lighter layers are poor in manganese (but rich in silicon, aluminum, and iron). Manganese-rich dark layers in varnish appear to form under relatively humid climate, and manganese-poor lighter layers under relatively dry climate. These two types of layers alternate to form distinct sequences. Thus, rock varnish can be used to reconstruct the moisture history of deserts where other records of climatic change are often hard to obtain or difficult to read. *See* VARNISH.

Since rock varnish records environmental, especially climatic, events that are regionally or even globally synchronous, varnish microstratigraphy can be used as a tephrachronology-like tool for age dating. Without radiometric calibration, varnish microstratigraphy itself may be used to estimate relative ages of varnished geomorphic or archeological features in deserts. Once calibrated, varnish microstratigraphy can provide numerical age estimates for geomorphic and archeological features. Specifically, for

petroglyphs and geoglyphs, the layering patterns of rock varnish hold the greatest potential for assigning ages. *See* ARCHEOLOGICAL CHRONOLOGY; DATING METHODS; ROCK AGE DETERMINATION. Tanzhuo Liu

Bibliography. R. I. Dorn and T. M. Oberlander, Rock varnish, *Progr. Phys. Geog.*, 6:317–367, 1982; C. D. Elvidge and C. B. Moore, A model for desert varnish formation, *Geol. Soc. Amer. Abstr. Prog.*, 11:271, 1979; M. Fleisher et al., A clue regarding the origin of rock varnish, *Geophys. Res. Lett.*, 26:103–106, 1999; T. Liu et al., Terminal Pleistocene wet event recorded in rock varnish from southern Nevada, western United States, *Geol. Soc. Amer. Abstr. Prog.*, 30(7):361, 1998; R. M. Potter and G. R. Rossman, Desert varnish: The importance of clay minerals, *Science*, 196:1446–1448, 1977.

Rocket

Either a propulsion system or a complete vehicle driven by such a propulsive engine. A rocket engine provides the means whereby chemical matter is burned to release the energy stored in it and the energy expended, in this example, by ejection at high velocity of the products of combustion (the working fluid). The ejection imparts motion to the vehicle in a direction opposite to that of the ejected matter. A rocket vehicle is propelled by rocket reaction and includes all components necessary for such propulsion, a payload such as an explosive charge, scientific instruments, or a human crew. A rocket vehicle also includes guidance and control equipment mounted in a structural airframe or spaceframe. *See* ROCKET PROPULSION; ROCKET STAGING; SATELLITE (SPACECRAFT); SPACE FLIGHT; SPACE PROBE. George P. Sutton

Rocket astronomy

The discipline that makes use of sounding rockets (**Fig. 1**) that fly near-vertical paths carrying scientific instruments to altitudes ranging from 25 to more than 900 mi (40 to 1500 km). Altitudes up to 30 mi (48 km) can be reached by balloons, so sounding rockets are typically used for higher altitudes in order to measure emissions from the Sun or other celestial sources that do not penetrate the Earth's atmosphere. Sounding rockets do not achieve orbital velocity; after completion of the launch phase, the payload follows a ballistic trajectory that permits 5–15 min of data taking before reentry. *See* ROCKET.

Mission profile. Rocket astronomy began in the United States in 1946, using captured German V-2 rockets with no pointing capability and very simple instrumentation. A typical modern launch is comparatively sophisticated, involving engine ignition, vehicle spinup, payload separation and despin, target acquisition, attitude control system (ACS) operation, telemetry and data command via the ground station, door closure or boom retraction, reentry and chute deployment, impact and recovery, data retrieval, and reuse of the scientific payload.

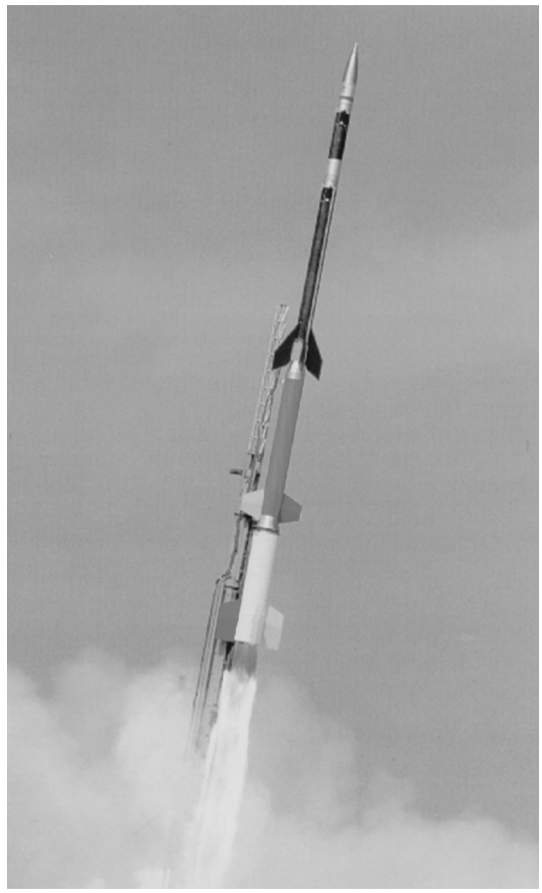

Fig. 1. Launch of a NASA Black Brant XII sounding rocket. (*NASA Goddard Space Flight Center*)

In comparison with experiments launched on satellites, rockets offer the advantages of simplicity, in that the payload is usually composed of a single instrument package; relatively frequent access to launch opportunities, typically once per year in solar physics; a shorter time scale from conception to reality, 3–5 years for a rocket versus 5–20 years for satellites; lower cost; and recoverability of the payload and the possibility of postflight instrument calibration, refurbishment, and reflight. Major disadvantages are short observing time, minutes versus months or years for satellites; localized coverage, restricted to an area adjacent to the launch site; and size and weight restrictions on payloads. *See* SATELLITE ASTRONOMY.

Over a dozen different launch vehicles are in common use (**Fig. 2**), ranging in length from 10 to 65 ft (3 to 20 m) and in diameter from 4.5 to 22 in. (11 to 56 cm). Payloads up to 2200 lb (1000 kg) can be carried to altitudes that are typical of those at which the space shuttle operates, 150–200 mi (250–320 km). For each launch vehicle the maximum altitude is strongly dependent upon payload weight (**Fig. 3**). For instance, the Black Brant X can carry a 200-lb (91-kg) payload to an altitude of 620 mi (1000 km) and can boost an 800-lb (364-kg) payload to 250 mi (400 km).

Sounding rockets are routinely launched in the United States from Wallops Island, Virginia; Poker

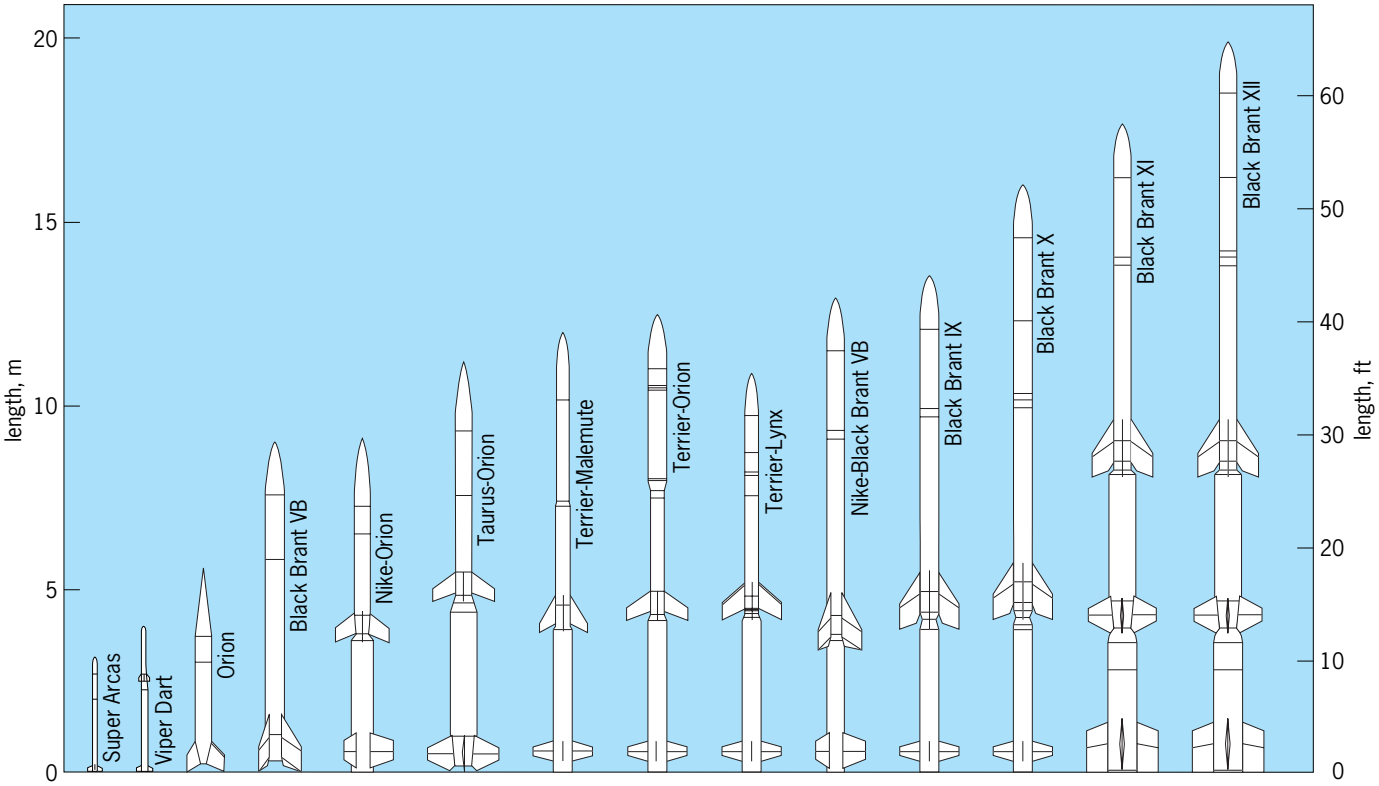

Fig. 2. NASA sounding rockets. (*NASA Sounding Rockets Working Group*)

Flats Research Range, Alaska; and White Sands Missile Range, New Mexico. Sites in Norway, Sweden, and Canada are also used. Temporary launch ranges are arranged when needed, and scientific requirements may dictate heavy usage of a particular facility. This was the case in 1988 and 1989 when numerous observations of Supernova 1987*a* were carried out from Woomera, in Australia. The United States sounding rocket program makes extensive use of surplus military rockets as an economical source of engines. NASA's Orion was formerly a Hawk ground-to-air missile, while Nike, Taurus, and Terrier engines are used as boosters in two- and three-stage flights. These components are usually supplied without guidance control systems.

The United States, Canada, Japan, and many European countries, including Germany, France, Sweden, Norway, Switzerland, and the United Kingdom, maintain vigorous scientific rocket programs. These programs focus on the disciplines of aeronomy, magnetospheric physics, meteorology, and material sciences, as well as astronomy and astrophysics. The ability to carry out vertical profile measurements of relevant atmospheric parameters at heights of 25–125 mi (40–200 km) is essential in many of these scientific disciplines. To the extent that the Sun influences or controls conditions in the upper atmosphere and magnetosphere of the Earth, there is a strong connection between solar astronomy and the more local research areas. *See* AERONOMY; IONOSPHERE; MAGNETOSPHERE.

Solar rocket astronomy. The Earth's atmosphere is opaque to wavelengths shorter than ~300 nanometers. Observations in the ultraviolet and x-ray regions of the spectrum therefore require that instruments be placed at altitudes exceeding 25–87 mi (40–140 km), depending upon wavelength. The outer portion of the Sun's atmosphere, the corona, is at a substantially higher temperature than is the photosphere, several million kelvins versus 6000 K. The emission from this region is dominated by ultraviolet and x-ray photons, whose total luminosity is very weak in comparison to the Sun's visible light output because of the extremely low density of the corona. During the late 1950s and early 1960s, techniques were developed for focusing x-rays and thereby providing direct imaging of the corona. These early studies revealed the highly structured nature of the atmosphere, with approximately semicircular loops of hot plasma outlining the shape of the underlying magnetic field, which confines the hot gas. *See* X-RAY TELESCOPE.

Areas of the solar surface at which strong magnetic fields emerge are known as active regions; in addition to producing strong x-ray and ultraviolet emission, they are the sites of explosive solar flare events. In such events the plasma is temporarily heated to temperatures exceeding 10^7 K by the rapid release of stored magnetic energy; in addition, material is often ejected into interplanetary space, and these events may have associated with them bursts of high-energy particles, microwave emission, gamma rays, and magnetic storms. Flares have long been known to cause disturbances of the Earth's atmosphere, and the largest flares pose a serious threat to astronauts situated above the atmosphere; forecasting of flares

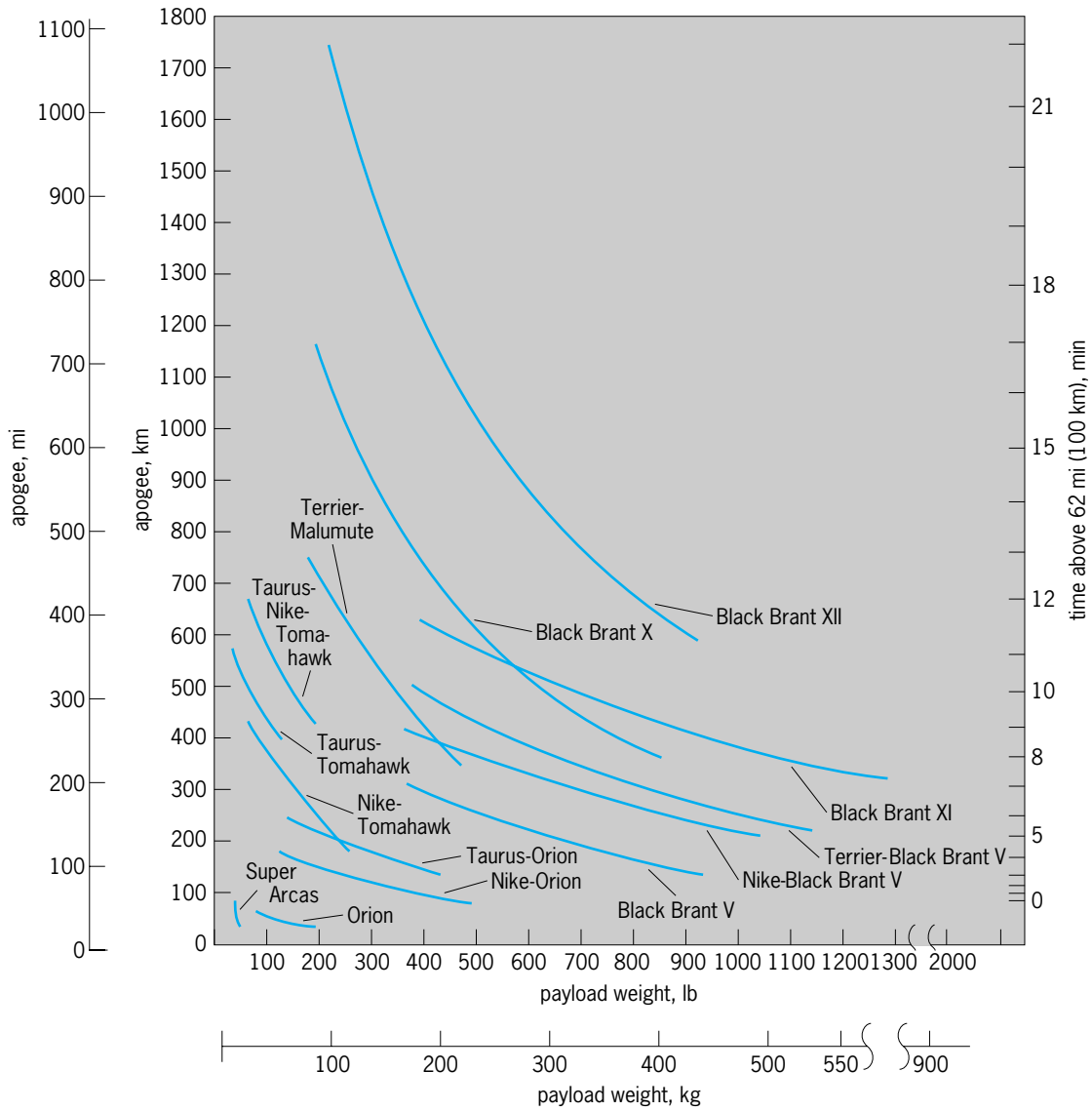

Fig. 3. Lift capability of NASA sounding rockets. Apogee is shown as function of payload weight.

is thus a major area of research. Analysis of the physical conditions leading to and during a flare involves observation of the entire three-dimensional atmospheric structure before a flare and the location and evolution of the hot plasma during the flare. Such studies require recording of all temperature regimes from 5000 to 10^7 K, which implies the ability to observe wavelengths from the infrared, through the ultraviolet and soft x-ray, into the hard x-ray (10-keV) regime, plus microwave and other radio emission. Such data cannot be obtained by a single technique, and solar research satellites are therefore composed of multiple co-observing instruments combining both imaging and spectroscopy, and including ground-based observations as a necessary part of the overall scientific program.

Among the notable advances in coronal studies from rockets has been the development of a new technique for x-ray imaging, with concomitant improvements in imaging spectroscope methods. The use of multilayer coatings for enhanced x-ray reflectivity may be viewed as depositing an artificial Bragg crystal onto a figured substrate. This technique offers the advantage of substantially higher image sharpness, and provides as well simultaneous spectroscopy because of the narrow x-ray bandwidth of the coatings. The major technological limitation of this technique is the inability to produce adequate normal-incidence reflectivity at short wavelengths (less than 2 nm); in this spectral region, grazing-incidence methods continue to be the most appropriate. *See* SUN; X-RAY DIFFRACTION.

Other celestial sources. Observations of nonsolar sources from sounding rockets encounter two major difficulties: the low intensity of the emission and the technological problem of pointing the payload at the source during the flight. The two problems are related in that the low flux levels generally lead to a requirement for long exposure or integration times, and the faintness of the source leads to difficulty in constructing a sensor that can provide the appropriate detection signal to an attitude control system.

The development of a gas-jet attitude control system for Aerobee rockets in 1964 and the addition of gyroscopic stabilization in 1965 permitted ultraviolet spectroscopy of stars with high enough dispersion to record interstellar absorption features in the 130–160-nm range. Such data permit measurement of interstellar hydrogen, carbon, nitrogen, oxygen, silicon, and sulfur. Spectra in the 100–250-nm region permit studies of stars with surface temperatures of 10,000–30,000 K; in addition to total-luminosity determination, a major discovery from line-profile measurements of hot-star absorption features was the finding that such stars have large mass loss due to high-velocity (greater than 600 mi/s or 10^3 km/s) stellar winds. In 1970, a sounding rocket detected the absorption line of interstellar molecular hydrogen; this detection was the forerunner of later extensive studies from the *Copernicus* satellite, which have had a major impact on the field of interstellar chemistry. *See* ASTRONOMICAL SPECTROSCOPY; COSMOCHEMISTRY; INTERSTELLAR MATTER; SPACECRAFT PROPULSION; ULTRAVIOLET ASTRONOMY.

In 1962, a rocket experiment detected nonsolar x-rays for the first time, at a level many orders of magnitude higher than would be produced by an equivalently placed solar-strength source. Subsequent flights confirmed the initial result and discovered other x-ray sources, as well as a diffuse x-ray background and an emission source in the Crab Nebula. The field of x-ray astronomy is now dominated by observations from satellites, which have the ability to carry out all-sky surveys and to integrate for many hours in order to study low-intensity sources. *See* X-RAY ASTRONOMY. Leon Golub

Bibliography. American Institute of Aeronautics and Astronautics, *Proceedings of the 7th Conference on Sounding Rockets, Balloons, and Related Space Systems*, 1986; G. V. Groves (ed.), *Dynamics of Rockets and Satellites*, 1965; H. E. Newell, Jr., *Beyond the Atmosphere*, NASA Publ. SP-4211, 1980; H. E. Newell, Jr., *Sounding Rockets*, 1959; A. I. Skoog (ed.), *History of Rocketry and Astronautics*, American Astronautical Society, 1990.

Rocket propulsion

The process of imparting a force to a flying vehicle, such as a missile or a spacecraft, by the momentum of ejected matter. This matter, called propellant, is stored in the vehicle and ejected at high velocity. In chemical rockets the propellants are chemical compounds that undergo a chemical combustion reaction, releasing the energy for thermodynamically accelerating and ejecting the gaseous reaction products at high velocities. Chemical rocket propulsion is thus differentiated from other types of rocket propulsion, which use nuclear, solar, or electrical energy as their power source and which may use mechanisms other than the adiabatic expansion of a gas for achieving a high ejection velocity. *See* ELECTROTHERMAL PROPULSION; INTERPLANETARY PROPULSION; ION PROPULSION; PLASMA PROPULSION; PROPULSION; SPACECRAFT PROPULSION.

In this article the performance criteria, the basic features, the application, and the preparation of various rocket propulsion systems that are driven by the combustion energy of chemicals are described. Propulsion systems using liquid propellants (such as kerosine and liquid oxygen) have traditionally been called rocket engines, and those that use propellants in solid form have been called rocket motors.

Performance. The performance of a missile or space vehicle propelled by a rocket propulsion system is usually expressed in terms of such parameters as range, maximum velocity increase of flight, payload, maximum altitude, or time to reach a given target. Propulsion performance parameters (such as rocket exhaust velocity, specific impulse, thrust, or propulsion system weight) are used in computing these vehicle performance criteria. **Table 1** gives typical performance values.

The momentum imparted to the vehicle by a rocket depends directly on the exhaust velocity v at which the hot gases are expelled from the rocket in a supersonic nozzle. This supersonic exhaust velocity is a good indicator of the effectiveness of any one propellant combination and the gas expansion in the nozzle. It is given by Eq. (1), where R is the

$$v = \sqrt{\frac{2gkR}{k-1}\frac{T}{M}\left[1-\left(\frac{p_2}{p_1}\right)^{(k-1)/k}\right]} \qquad (1)$$

universal gas constant, k is the specific heat ratio of the reaction product gases, and g is the acceleration of gravity (32.2 ft/s^2 or 9.8 m/s^2). When SI units are used, the factor g can be deleted. The pressure of the nozzle exit at optimum nozzle expansion area ratio is p_2 and the chamber pressure is p_1. Because of this

TABLE 1. Typical performance values of rocket propulsion systems*

Propulsion system parameter	Typical range of values
Specific impulse at sea level	180–390 s
Specific impulse at altitude	215–470 s
Exhaust velocity at sea level	5800–15,000 ft/s (1800–4500 m/s)
Combustion temperature	4000–7200°F (2200–4000 °C)
Chamber pressures	100–3000 lb/in.^{2} (0.7–20 MPa)
Ratio of thrust to propulsion system weight	20–150
Thrust	0.01–6.6 × 10^6 lb (0.05–2.9 × 10^7 N)†
Flight speeds	0–50,000 ft/s (0–15,000 m/s)

*Exact values depend on application, propulsion system design, and propellant selection.
†Maximum value applies to a cluster; for a single rocket motor it is 3.3 × 10^6 lb (14,700 kN).

dependence on the pressure ratio, there is a slight increase (10–20%) in exhaust velocity as the exit pressure is decreased (increase in altitude) or as the chamber pressure is increased. The exhaust velocity increases as the molecular weight of the combustion gases M decreases and the combustion temperature T increases, but exhaust velocity also increases if the pressure ratio across the exhaust nozzle increases.

The specific impulse is a key rocket propulsion performance parameter, and it can be considered the thrust obtained from an equivalent rocket propulsion system that has a propellant mass flow of unity. It is expressed in pounds of force per pound of mass per second, or the exhaust velocity v divided by the acceleration of gravity. In the foot-pound-second system of units it is sometimes expressed simply as seconds. In SI units it is expressed in newtons per kilogram per second. *See* SPECIFIC IMPULSE.

Exhaust velocities and specific impulses can be calculated from thermochemical relations; the values so obtained are known as the theoretical specific impulses and theoretical nozzle exhaust velocities. Actual values are 92–98% of their theoretical values, because the losses due to friction, nozzle exhaust, divergence, and incomplete combustion are relatively small.

Thrust F is defined by Eq. (2) in terms of a momen-

$$F = \dot{m}v + (p_2 - p_3)A_e \qquad (2)$$

tum and a pressure term, where $\dot{m}$ is the mass flow through the nozzle, v is the average nozzle exit velocity, A_e is the nozzle exit area, and pressure difference $(p_2 - p_3)$ is that between nozzle exit pressure p_2 and ambient pressure p_3. As the mass flow increases, so does the thrust. Because p_2 is usually close to p_3, the second term of the thrust equation is small, and the thrust is roughly equal to the product of mass flow $\dot{m}$ and exhaust velocity v. As the altitude increases (p_3 decreases), the velocity and the thrust increase (**Fig. 1**).

The nozzle area ratio is the nozzle exit area divided by the nozzle throat area. If this ratio is too large, then the gas flow in the nozzle overexpands, the value of p_2 is less than p_3 (the last term in the thrust equation becomes negative), and the thrust decreases. The values of nozzle area ratio are low (between 3 and 20) for rocket propulsion systems operating at

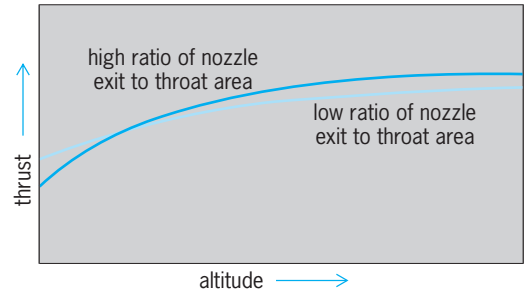

Fig. 1. Typical variation of thrust with altitude.

low altitudes. For high-altitude operation, where p_3 is close to zero, a high area ratio (between 20 and 400) allows a 10–25% increase in performance over sea level; this means a relatively large exit nozzle. *See* NOZZLE.

Clever designs and highly stressed materials are used to make propulsion system weights as low as possible so as to minimize the amount of energy necessary to accelerate inert mass. The best vehicle performance is obtained when the propellant mass is a maximum and the inert mass of the propulsion system, the residual propellants, structure, controls, nonpropulsive upper stage, or payload is a minimum.

Applications. Rocket propulsion is used for different military missiles or space-flight missions. Each requires different thrust levels, operating durations, and other capabilities. **Tables 2** and **3** show some of these and also the preference for liquid or solid propellants, but in some applications both types are used. In addition, rocket propulsion systems are used for rocket sleds, jet-assisted takeoff, principal power plants for experimental aircraft, or weather sounding rockets. For some space-flight applications, systems other than chemical rockets are used or are being investigated for possible future use. *See* GUIDED MISSILE; MISSILE; SATELLITE (SPACECRAFT); SPACE FLIGHT; SPACE PROBE.

The space shuttle is an example of a multistage vehicle that uses a variety of propulsion systems. It has 67 individual propulsion units, located in different places on the vehicle (**Fig. 2**). Each of the several rocket propulsion systems has its own application requirements (**Table 4**). *See* ROCKET STAGING; SPACE SHUTTLE.

TABLE 2. Space flight application of rocket propulsion systems

Type	Liquid (L) or solid (S) propellant	Thrust range, lbf (N)	Special capabilities
Launch vehicle			
Booster	S/L	20,000–7,000,000 (90,000–31,000,000)	Reuse
Sustainer(s)	L/S	2000–1,000,000 (9000–4,500,000)	Throttling, reuse
Attitude control	L	1–1000 (4.5–4500)	Many restarts, reuse
Separation	S	200–20,000 (900–90,000)	
Spacecraft*			
Main propulsion	L	200–15,000 (900–67,000)	Throttling (variable thrust)
Attitude/maneuver	L	0.01–300 (0.045–1350)	Many restarts
Planetary/lunar landing	L	100–10,000 (450–45,000)	Throttling, restarts

*For reusable, orbiting space shuttle, see Table 4.

TABLE 3. Military applications of rocket propulsion systems

Type	Liquid (L) or solid (S) propellant	Thrust, lbf (N)	Maximum acceleration* and special capabilities	Example
Surface-to-surface long-range or antisatellite (1, 2, or 3 stages)				
Booster	S/(L)	20,000–2,500,000	1.2–6 *g*	Titan (2 or 3 stages)
Sustainer(s)	S/(L)	2000–500,000 (90,000–11,000,000)	0.5–5 *g*	Minuteman (3 stages) Poseidon (2 stages)
Attitude/maneuvers	L	1–3000 (4.5–13,500)	0.01–1 *g*	Pershing (2 stages)
Air-to-surface (1 or 2 stages)				
Boost/sustain	S	1000–35,000 (4500–155,00)	Two thrust levels	SRAM, Shrike
Surface-to-air (1 or 2 stages)				
Boost/sustain	S	800–30,000 (3600–135,000)	Regressive up to 15*g*	Redeye, Chapparell, Standard
Air-to-air				
Boost/sustain	S	1000–35,000 (4500–155,000)	Two thrust levels	Phoenix, Sparrow, Falcon, Sidewinder
Infantry support				
Antitank	S	500–8000 (2200–36,000)	1–15 *g*	Lance (L), Shillelagh,
Short range	S/(L)	500–15,000 (2200–67,000)		Hellfire
Ballistic missile defense				
Booster	S	50,000–500,000 (220,000–2,200,000)	3–100 *g*	Being developed
Sustainers	L/(S)	5000–50,000 (22,000–220,000)	2 to 50 *g*	Being developed
Experimental aircraft	L	3000–7000 (13,500–31,000)	0.5–4.0 *g*	X-1 and X-15 research aircraft
Assisted takeoff	S	250–10,000 (1100–45,000)		

*1 g = 32.2 ft/s^2 = 9.8 m/s^2.

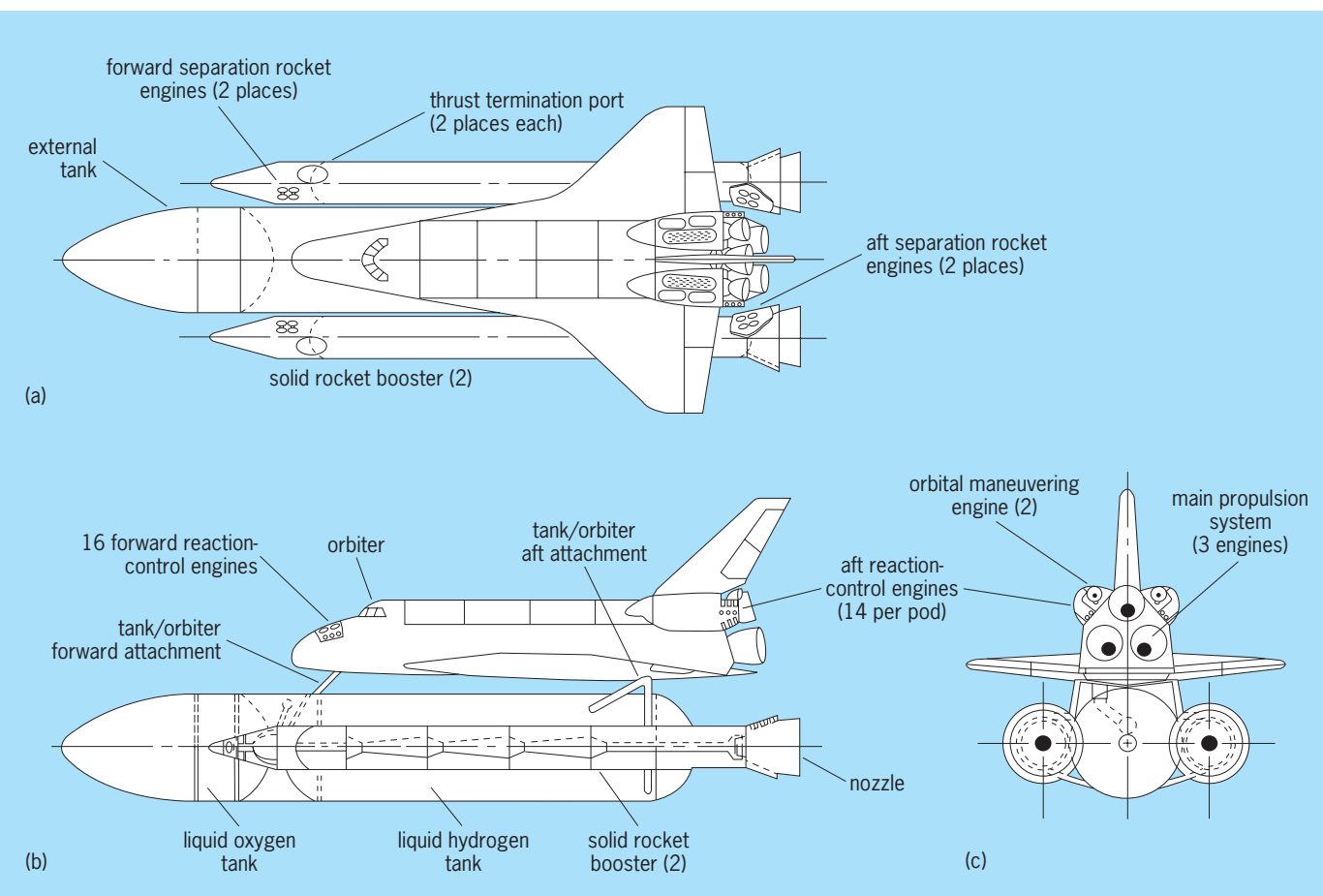

Fig. 2. Space shuttle and its propulsion systems. (*a*) Top view. (*b*) Side view. (*c*) Rear view. (*NASA*)

TABLE 4. Rocket propulsion systems used on the space shuttle

Type	Number of units	Propellants	Thrust, lbf (kN)	Specific impulse,* lbf·s/lbm	Gimbal angle, degrees	Absolute chamber pressure, lbf/in.^{2} (MPa)	Nozzle area ratio
Main engine	3	Liquid oxygen and liquid hydrogen Mixture ratio = 6	At 100% thrust: 370,000 (1650) at sea level, 470,000 (2090) at altitude; throttled between 65 and 109%	454 at altitude	±10.5	~3000 (~21)	77.5
Orbital maneuvering system engine	2	Nitrogen tetroxide and monomethyl-hydrazine Mixture ratio = 1.65	6000 (26.7) in orbit	313 at altitude	±6° pitch, ±7° yaw	125 (0.86)	55
Reaction control system:							
Primary engine	38	Nitrogen tetroxide and monomethyl-hydrazine	870 (3.87) at altitude	~280	None	~152 (1.05)	20 to 30
Vernier engines	6	Mixture ratio = 1.6	24 (0.106)	~265	None	110 (0.76)	~22
Solid rocket booster motor	2	Composite solid: 16% aluminum, 69.8% ammonium perchlorate, 12% polymer, 1.96% epoxy curing agent, 0.17% burn-rate controller	3.3×10^6 (14.680) at sea level; 2.3×10^5 (10,230) at altitude	274	±8°	950 to 600 (6.5 to 4.1)	~7
Booster separation rocket motor	16	Composite solid: ammonium perchlorate, polybutadiene, 2% aluminum	~20,050 (~89) Lift-off thrust: ~7.7×10^6 (34,300)	225	None	1800 (12.4)	5.83
	Total number of units: 67						

*Specific impulse is usually expressed in the units given here, even where SI is used.

The three main engines are high-performance liquid-propellant rocket engines, which provide the major velocity increase to the space shuttle. During a typical shuttle flight they may operate from 8 to 9 min, which is rather long for any rocket propulsion system. Throttling of the thrust down to 65% allows the vehicle to limit its aerodynamic forces during ascent and also to limit the maximum vehicle acceleration for the last minute or two of its powered flight.

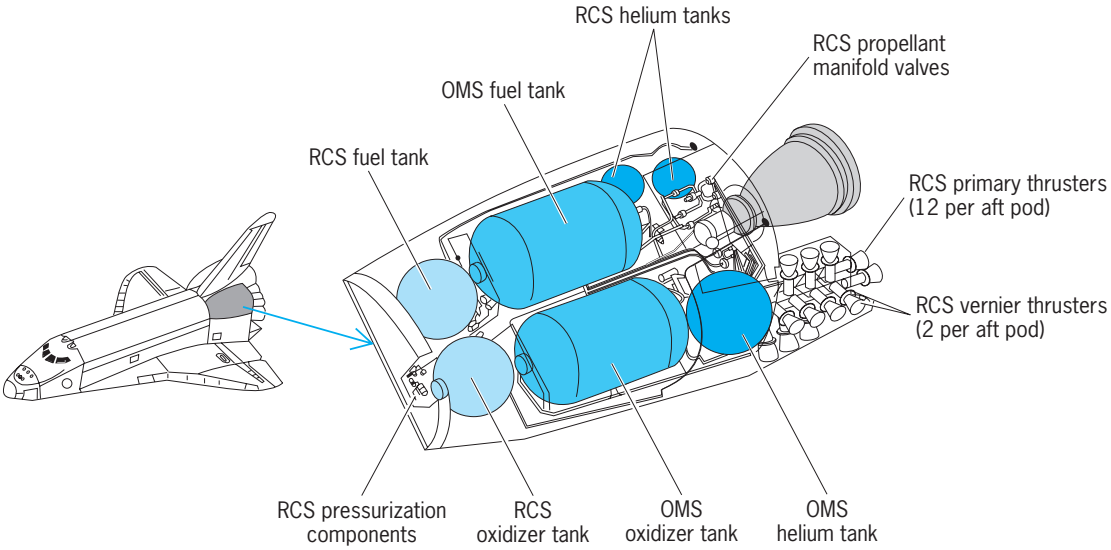

Fig. 3. One of two space shuttle aft pods containing the rocket engines of the orbital maneuvering system (OMS) and reaction control system (RCS). (*NASA*)

The orbital maneuvering system (OMS) has two liquid-propellant multipurpose engines that provide the thrust for orbit injection, orbit circularization, orbit transfer, rendezvous, deorbit, and several potential abort maneuvers. The reaction control system (RCS) has two sizes of liquid-propellant engines (38 primary engines and 6 smaller vernier engines) at three different locations. Any one engine may need to operate for thousands of pulses during a flight. The reaction control system provides attitude control of the orbiter vehicle in pitch, yaw, and roll; provides precise vehicle velocity corrections in three directions; and can assist in orbital maneuvers. Multiple primary engines provide for redundancy. There are two engine pods at the aft end of the space shuttle orbiter vehicle (**Fig. 3**), each of which contains one engine of the orbital maneuvering system with 6000 lb (26,700 newtons) vacuum thrust, and several engines of the reaction control system, together with their tanks of storable liquid propellants, valves, piping, and helium pressurization systems.

The two solid rocket boosters (SRBs) provide the principal thrust for lifting the space shuttle off the launch pad to about 150,000 ft (45,000 m). They are the largest solid-propellant rocket motors ever flown and the first designed for reuse. Each motor is made of four matched segments, which are assembled at the launch site. The pairs of boosters are matched by loading each of the four segments in pairs from the same batches of propellant ingredients so as to minimize any difference in burning rate between the two motors and thus also any thrust imbalance. The lift-off thrust is reduced to about two-thirds of its maximum value to prevent excessive stresses during ascent at maximum dynamic pressure. Each booster assembly has eight booster separation motors, four in the nose and four in the aft skirt, which burn for 0.8 s and push the spent boosters from the flying vehicle.

Liquid-propellant rocket engines. These use liquid propellants stored in the vehicle for their chemical combustion energy. The principal hardware subsystems are one or more thrust chambers, a propellant feed system, which includes the propellant tanks in the vehicle, and a control system.

Liquid propellants. **Table 5** lists several common liquid propellants. Bipropellants have a separate oxidizer liquid (such as liquefied oxygen or nitrogen tetroxide) and a separate fuel liquid (such as liquefied hydrogen or hydrazine). Those propellant combinations, which ignite spontaneously when the oxidizer liquid comes in contact with the fuel, are called hypergolic. Those that mix without self-ignition could result in a potentially explosive mixture, unless an ignition system (sparks, glowplugs, or a solid-propellant igniter) provides the initial energy to start the combustion. Cryogenic propellants (for example, liquid oxygen and liquid hydrogen) are those that are liquid only at very low temperature, evaporate rapidly at ambient temperature, and require

TABLE 5. Theoretical performance of liquid rocket propellant combinations

Propellant type (oxidizer/fuel)	Mass mixture ratio (oxidizer-to-fuel)	Bulk density, g/cm^3	Combustion temperature, °F(°C)	Molecular mass of chamber gases, lb/mole or kg/mole	Specific impulse,* lbf·s/lbm	Typical application	Vehicle where currently used	Hypergolic ignition
Cryogenic liquid propellant								
Oxygen/kerosine	2.56	1.02	6150 (3400)	23.2	294	Space launch vehicles	Atlas missile and first stage of Saturn	No
Oxygen/hydrogen	4.02	0.25	4935 (2724)	10.0	390	Space launch vehicles	Space shuttle main engine	No
Fluorine/hydrogen	7.60	0.35	6505 (3596)	11.8	405	Experimental		Yes
Storable liquid propellant								
Red fuming nitric acid/hydrazine mixture	1.47	1.23	5200 (2871)	22	280	Target drone	Sounding rocket	Yes
Red fuming nitric acid/kerosine	4.8	1.35	5590 (3088)	22.6	268	Surface-to-surface missile	Non-U.S. missiles	No
Nitrogen tetroxide/ 50% hydrazine with 50% dimethyl-hydrazine	2.0	1.21	5725 (3163)	22.6	283	Ballistic missiles, attitude control	Titan missile Agena vehicle	Yes
Nitrogen tetroxide/ monomethyl-hydrazine	1.6	1.23	5180 (2860)	22.5	282	Spacecraft attitude control	Space shuttle RCS engines	Yes
Liquid monopropellant								
90% hydrogen peroxide (at 300 lb/in.^{2} abs or 2.07 MPa abs)	—	1.44	1370 (743)	21.5	132	Gas generator for turbopump		Needs catalyst
Hydrazine (at 150 lb/in.^{2} abs or 1.03 MPa abs)	—	1.01	1115 (602)	10.7	155	Spacecraft attitude control	Deep space mission	Needs catalyst

*At a combustion chamber pressure of 1000 lb/in.^{2} abs (6.9 MPa abs) and full expansion at sea level to a nozzle exit pressure of 14.7 lb/in.^{2} abs (101 kPa abs). Special impulse is usually expressed in units given here, even where SI is used.

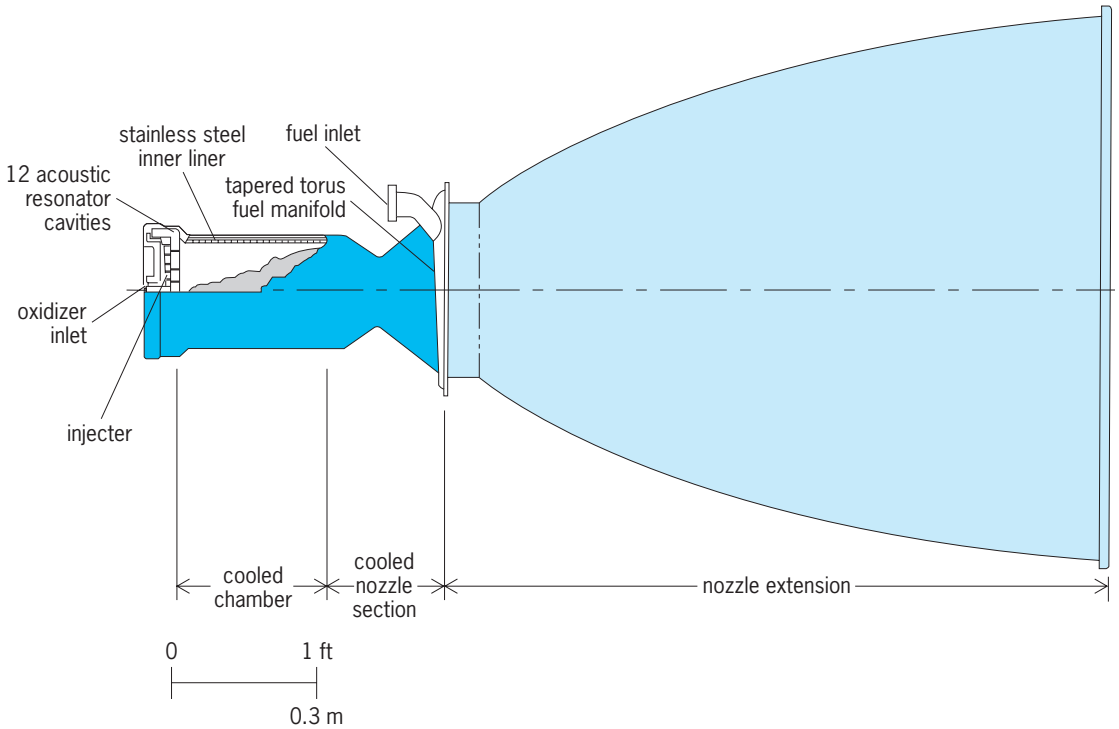

Fig. 4. Simplified half-section of one of the two thrust chambers of the orbital maneuvering engines used on the space shuttle. (*Aerojet Tech Systems*)

thermal insulation or refrigeration for prolonged storage. Storable propellants (for example, nitrogen tetroxide, hydrazine, and kerosine) are liquid at ambient temperatures.

Monopropellants consist of a single liquid that contains both oxidizer and fuel ingredients. A catalyst is required to decompose the monopropellant into gaseous combustion products. Bipropellant combinations allow higher preformance (higher specific impulse) than monopropellants. Hydrazine can be used as a monopropellant or, in combination with an oxidizer liquid, it can be a fuel for a bipropellant.

Liquid propellants have some undesirable properties and require careful handling. They each have their own peculiar types of hazard and all are either corrosive, toxic, flammable, or explosive; some propellants have all four of these problems. Special procedures, compatible containment materials, various safety precautions, and, in some cases, gas masks and protective clothing are required. *See* PROPELLANT.

Thrust chamber. The three principal components (**Fig. 4**) are the combustion chamber, where rapid, high-temperature combustion takes place; the converging-diverging nozzle, where the hot reaction-product gases are accelerated to supersonic velocities; and an injector (**Fig. 5**), which meters the flow of propellants in the desired mixture of fuel and oxidizer, introduces the propellants into the combustion chamber, and causes them to be atomized or broken up into small droplets. If the propellants are not hypergolic, then the thrust chamber also has an ignition system.

The burning of the liquid propellants takes place inside the combustion chamber at high chamber pressure, typically between 60 and 300 lb/in.^{2} (0.4 and 2 megapascals) for pressurized feed systems and between 300 and 3000 lb/in.^{2} (2 and 20 MPa) for pump-fed systems. The combustion is an intense, turbulent, unsteady, and noisy process with many strong, very rapid chemical reactions occurring simultaneously. Special design provisions, such as cavities or injector baffles, are often included to prevent combustion vibration from building up excessive pressure amplitudes that could damage or destroy the engine.

Some thrust chambers (such as the space shuttle's main engines and orbital maneuvering engines) are gimbaled or swiveled to allow a change in the direction of the thrust vector for vehicle flight motion control (**Fig. 6**). The angular movement of the nozzle is controlled by actuators; this enables controlling the vehicle in pitch, yaw, and roll. However, roll control requires at least two nozzles (or jet vanes).

Since the combustion temperature is roughly twice as high as the melting point of steel, bipropellant rocket engines must be cooled if they operate for longer than a few seconds. Cooling can be accomplished by (1) circulating one of the propellants in a cooling jacket around the chamber and nozzle and absorbing heat in the propellant prior to its introduction into the chamber (best for long duration); or (2) by using ablative liners or layers bonded to the inside of the chamber and nozzle; these liners absorb heat by melting, charring, or vaporizing some of the composite fiber-reinforced liner material, which then forms a cool-gas boundary layer near the wall surfaces. For monopropellant thrust chambers and for areas near the nozzle exit of a bipropellant engine (where the local propellant reaction-gas temperature is relatively low, 1600–2800°F or 900–1500°C), a radiation-cooled, simple, formed,

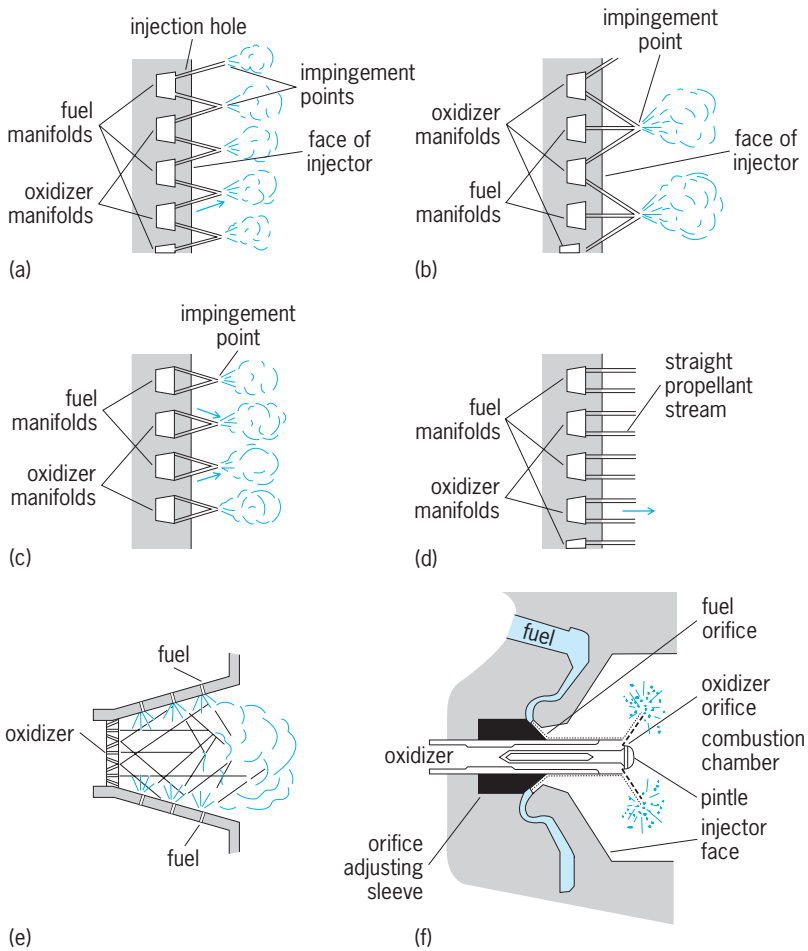

Fig. 5. Several types of injectors. (*a*) Doublet-impinging stream pattern. (*b*) Triplet-impinging stream pattern. (*c*) Self-impinging stream pattern. (*d*) Shower-head stream pattern. (*e*) Nonimpinging stream pattern. (*f*) Variable-injection-area concentric-tube injector. (*After G. P. Sutton, Rocket Propulsion Elements, 5th ed., John Wiley and Sons, 1986*)

heat-resistant metal such as coated niobium is usually satisfactory.

The chamber and nozzle throat region of each engine of the space shuttle's orbital maneuvering system (Fig. 4) are fuel-cooled. There is a stainless steel inner liner with 120 grooves for the coolant passage and an electroformed nickel closure. The nozzle exit skirt is a thin niobium alloy shell. The injector is of the platelet type and has 12 acoustic resonator cavities at its periphery to suppress combustion vibration.

Feed systems. Liquid-propellant rocket engine transfers the liquid propellants from the vehicle storage tanks to one or more thrust chambers. There are two basic types: gas-pressure feed systems and turbopump feed systems.

In a gas-pressure feed system, propellants are displaced with a high-pressure gas which is fed into the tanks under a regulated pressure (**Fig. 7**). The stored high-pressure gas furnishes the pressurization energy. The thrust of a pressurized-gas rocket propulsion system is determined by the magnitude of the propellant flow as controlled by the gas pressure regulator setting. For low thrust and short duration this feed system is generally lighter and superior to other more complicated ones.

For spacecraft or satellite attitude control and maneuvering, the use of a pressurized feed system, together with storable liquid propellants, permits rapid pulsing of several thrust chambers for spacecraft stabilization, station keeping, or rendezvous and position control. Small thrust chambers of perhaps 1–1000-lb (4.5–4500-N) thrust are pulsed for short duration (as little as 0.01 s of operation) and many thousands of pulses during a given mission. Several thrust chambers (2–16) are fed propellants from a single pressurized feed system.

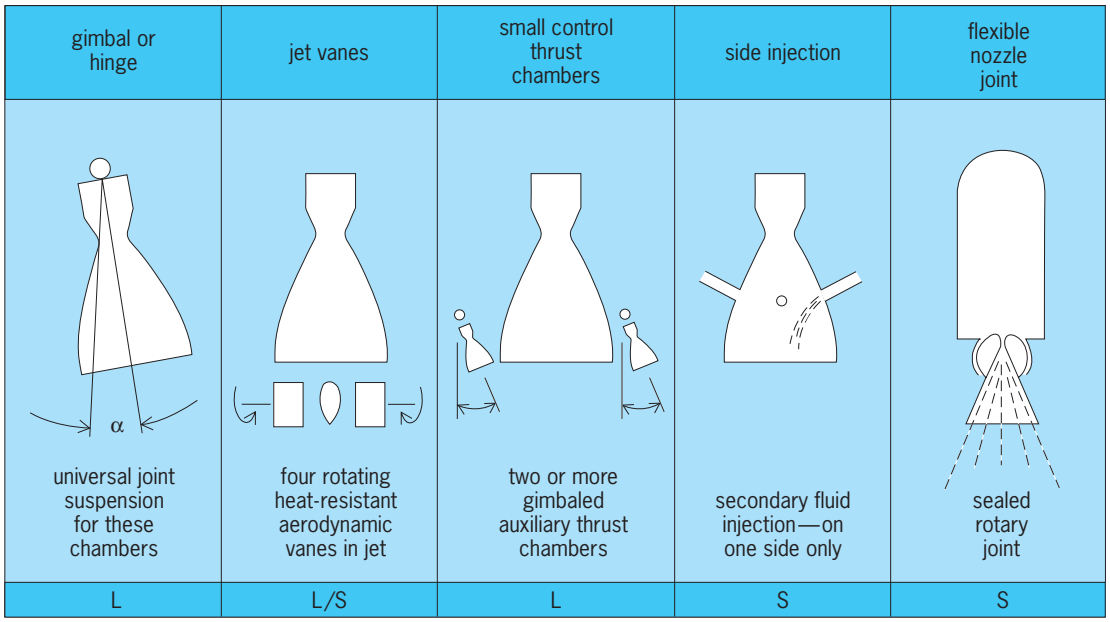

Fig. 6. Several thrust-vector control methods. L indicates use with liquid propllant engines, and S indicates use with solid-propellant motors. Actuators are not shown.

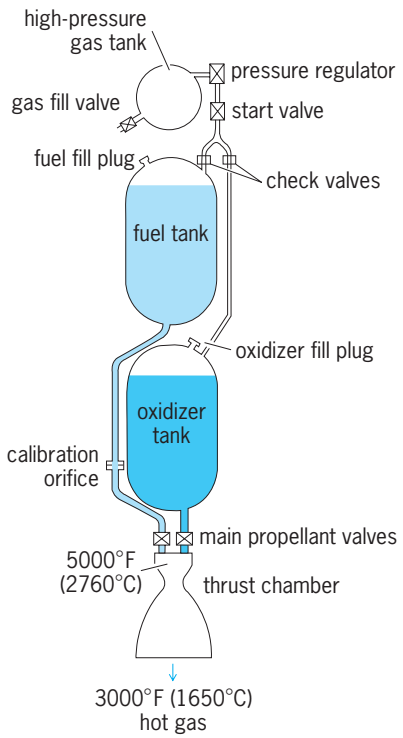

Fig. 7. Liquid-propellant rocket engine with gas-pressurized feed system.

In a field of zero gravity, as exists in outer space, the propellant liquids float around in a partially emptied tank; it is therefore necessary to include provisions for assuring that the outlet pipe opening in the propellant tank is covered with liquid (and not with the pressurized gas, which is also inside the tank) to permit proper starting and operation of the rocket. This has been accomplished by surface-tension devices, such as screens, or by including a positive expulsion device, such as a flexible bladder (made usually of thin metal or plastic), which separates the propellant liquid at all times from the pressurized gas. Pressure-fed systems are used on both the orbital maneuvering system and the reaction control system of the space shuttle.

In a turbopump feed system the propellant is pressurized by means of pumps driven by one or more turbines (**Fig. 8**) which derive their power from the expansion of hot gases. A separate gas generator ordinarily produces these gases in the required quantities and at a temperature which will not damage or partially melt the turbine buckets (1200–1800°F or 650–980°C). This is achieved in a chemical combustion chamber similar to the main thrust chamber, but the propellants are burned at a different and cooler mixture ratio. The assembly of pumps and turbine is termed a turbopump. Like other rotating machinery, it has bearings, seals, high-speed shafts, and lubrication devices. Turbopump feeds are generally superior to and lighter than other feed systems for high-thrust and long-duration applications. *See* GAS TURBINE.

For the large propellant flow and the high pressures involved in a rocket propulsion system, a centrifugal pump is generally considered superior to other types of pumps and is economical in weight and space. A multiblade centrifugal impeller is used with a volute casing. For fluids of low density, like liquid hydrogen, it is necessary to use multiple impeller stages of axial-flow pumps to develop the required inlet pressures at the thrust chamber. It is important that the pressure at the inlet of the pump always be sufficiently high to avoid low-pressure boiling, which induces cavitation. Special impellers are often used to prevent cavitation because of its detrimental effects on the operation and combustion stability of the rocket. *See* CAVITATION; CENTRIFUGAL PUMP.

The hot exhaust gases from the turbine are often used for heating pressurizing gas or evaporating cryogenic propellant in order to pressurize a propellant tank or for producing a small additional thrust by ejecting these gases through a nozzle. A turbopump system can be started in several ways. In one method, a low-temperature solid-propellant cartridge is ignited and starts the turbine rotation, which then allows the pumps to raise the propellant pressures and to feed propellant to the combustion chambers.

The space shuttle main engine (**Fig. 9**) uses a turbopump feed. The cryogenic propellants, liquid hydrogen and liquid oxygen, are carried in separate slightly pressurized tanks in the external tank (Fig. 2), and are transferred to the shuttle orbiter, where the engine's four turbopumps feed it to the combustion chambers. The gases in the preburners that drive the turbopumps flow through the injector into the main combustion chamber and are mixed and burned with the liquid oxygen and gaseous hydrogen supplied to the thrust chamber. This unique

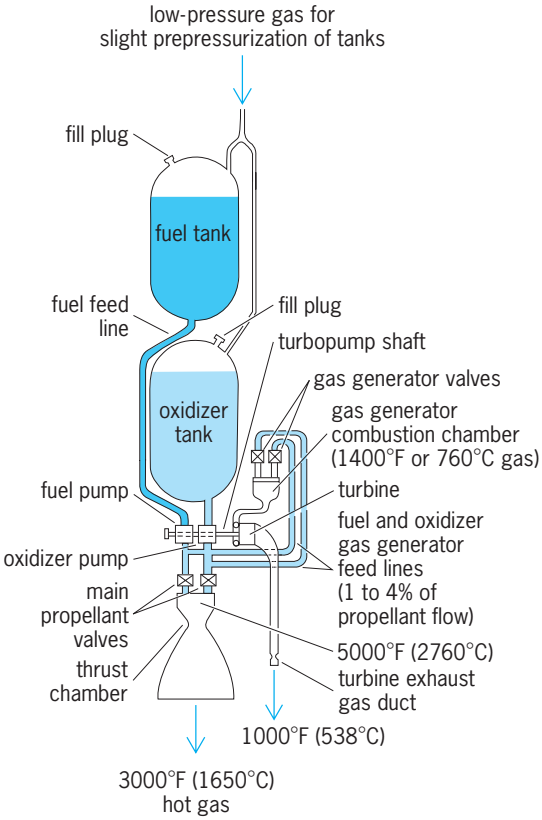

Fig. 8. Liquid-propellant rocket with turbopump feed system.

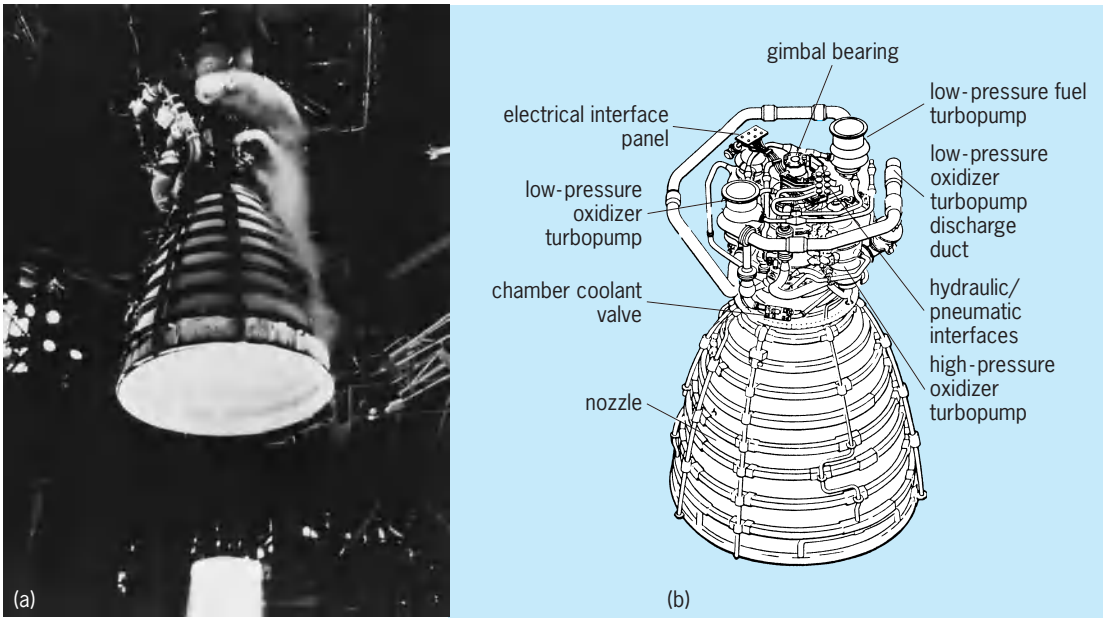

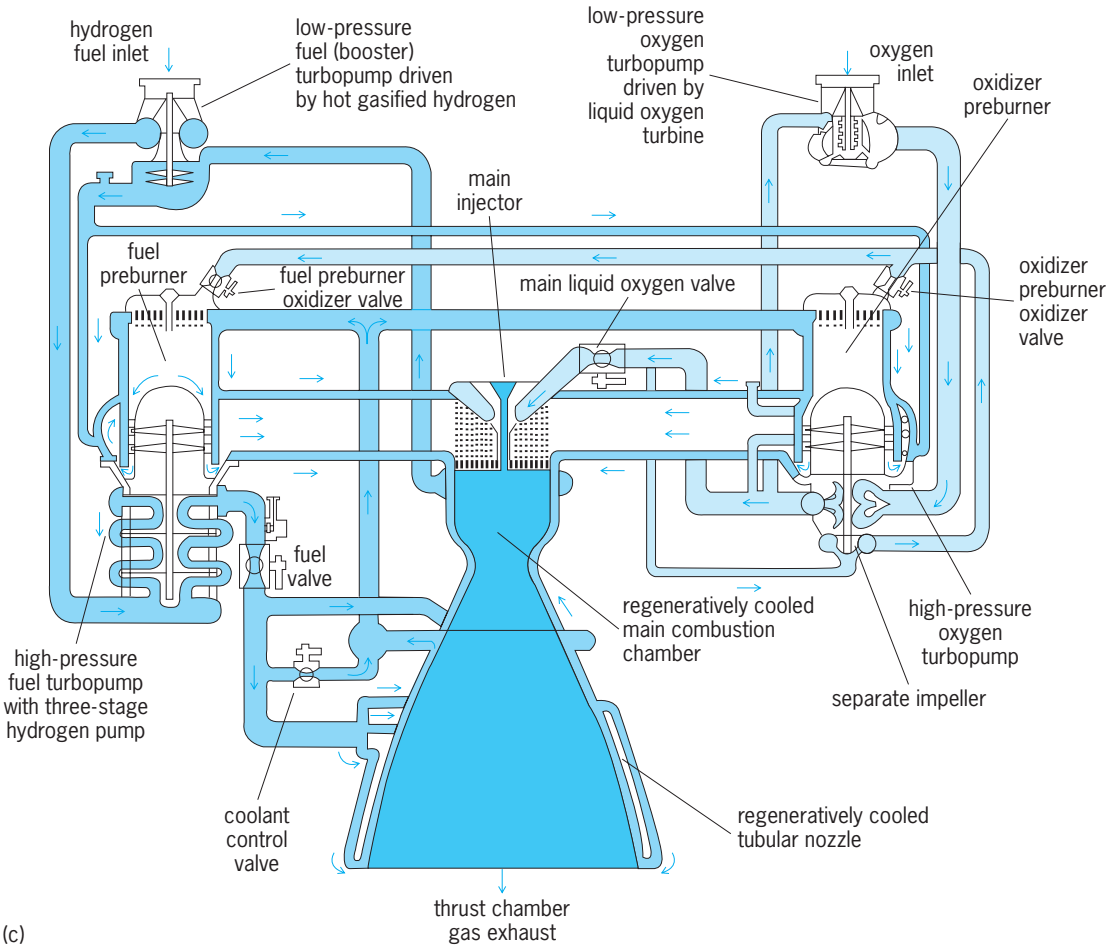

Fig. 9. Space shuttle main engine. (*a*) Static firing test. Hydrogen/oxygen flame is invisible except for shock wave at figure bottom (*Rocketdyne Division*, *Rockwell International*). (*b*) Oblique view. Main injector, chamber, and high-pressure fuel turbopump are behind other hardware (*NASA*). (*c*) Simplified flow diagram. Part of oxygen flow is pressurized with a separate impeller.

flow cycle contributes to the engine's high performance.

Controls and operation. All liquid-propellant rocket engines have controls that accomplish some or all of the following functions: start and shut down rocket operation, restart, maintain a predetermined constant or variable thrust, make emergency shutdowns when safety devices sense a malfunction, fill and drain propellants, and permit functional checkout of critical components without actual rocket engine operation. For example, each space shuttle main engine is controlled by an electronic computer/controller, together with several sensors, valves, and actuators, which provide flight readiness verification, start and shutdown sequencing, closed-loop thrust control and mixture-ratio control, on-board engine monitoring and checkout, response to vehicle commands, data storage, and malfunction detection with triggering of remedial actions.

Starting, stopping, and restarting are generally accomplished by valves in the main propellant lines. Because of the high pressures and the large flows involved, these valves are often operated from separate pilot valves that control the energy going to these main valves. A hydraulic, pneumatic, or electrical system is often used to actuate the principal valves to the thrust chamber and the gas generator and to furnish power to the thrust vector control actuators.

A propellant utilization system ensures that both the fuel and the oxidizer tank are emptied simultaneously. This is achieved by continuously measuring the amount of propellant that remains in both of these tanks and by adjusting throttling valves in both of the main propellant lines leading to the thrust chamber. This minimizes the amount of residual propellant remaining in the tanks, and minimizes the inert mass that remains to be accelerated.

In the operation of liquid-propellant rocket engines it is often desirable to vary the thrust or to throttle the rocket engine. In multiple rocket engines this can be achieved by shutting down some of the engines while letting others continue to operate. In a single-thrust-chamber rocket engine the throttling can be achieved by reducing the chamber pressure and the flow. Throttling is necessary for an accurately reproducible flight trajectory, limiting maximum acceleration, and for lunar descent and landing maneuvers. For example, the lunar landing engine of the Apollo space program has a thrust variation of 10:1.

Solid-propellant rocket motors. In rocket motors the propellant is a solid material that feels like a soft plastic or soap. The solid propellant cake or body is known as the grain. It can have a complex internal geometry and is fully contained inside the solid motor case, to which a supersonic nozzle is attached (**Figs. 10** and **11**).

The propellant contains all the chemicals necessary to maintain combustion. Once ignited, a grain will burn on all exposed surfaces until all the usable propellant is consumed; small unburned residual propellant slivers often remain in the chamber. As

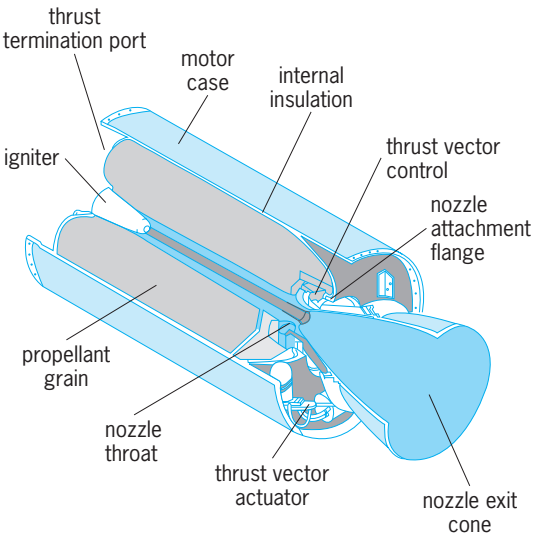

Fig. 10. Solid-propellant rocket motor. (*Hercules, Inc.*)

the grain surface recedes, a chemical reaction converts the solid propellant into hot gas. The hot gas then flows through internal passages within the grain to the nozzle, where it is accelerated to supersonic velocities. A pyrotechnic igniter provides the energy for starting the combustion.

Solid propellants. Of the many different solid propellants, the three most common types are composite propellants, double-base propellants, and composite double-base propellants (**Table 6**).

The largest number of motors use composite propellants. They contain crystalline, finely ground oxidizer particles, such as ammonium perchlorate or ammonium nitrate, dispersed in a fuel matrix of polymer. Several different polymers have been used successfully. Aluminum powder may be added as fuel; in the combustion it is burned into aluminum oxide, a white powdery solid material visible in the flame. There are minor quantities of other additives to improve the fabrication processing and the physical properties, control the burning rate, or limit chemical deterioration. Most composite solid-propellant grains are cast (as a viscous liquid) into the cavity of the case, where they polymerize and solidify. An internal rubberlike thin liner allows good bonding of the solid propellant to the motor case walls.

The next most common class are double-base propellants, which are based on two major ingredients: solid nitrocellulose and a liquid solvent, namely nitroglycerine. Sometimes other organic compounds that have oxidizing radicals attached to them are added, including compounds that are normally explosives. Double-base propellants can be extruded or cast by a solvent process. There are other minor additives, whose purposes are similar to those mentioned above. Aluminum powder may or may not be added as an extra fuel or as a combustion vibration suppressant.

Composite double-base propellants combine ingredients from both the above types. This gives a slightly higher performance, but the propellant is more difficult and hazardous to fabricate or handle.

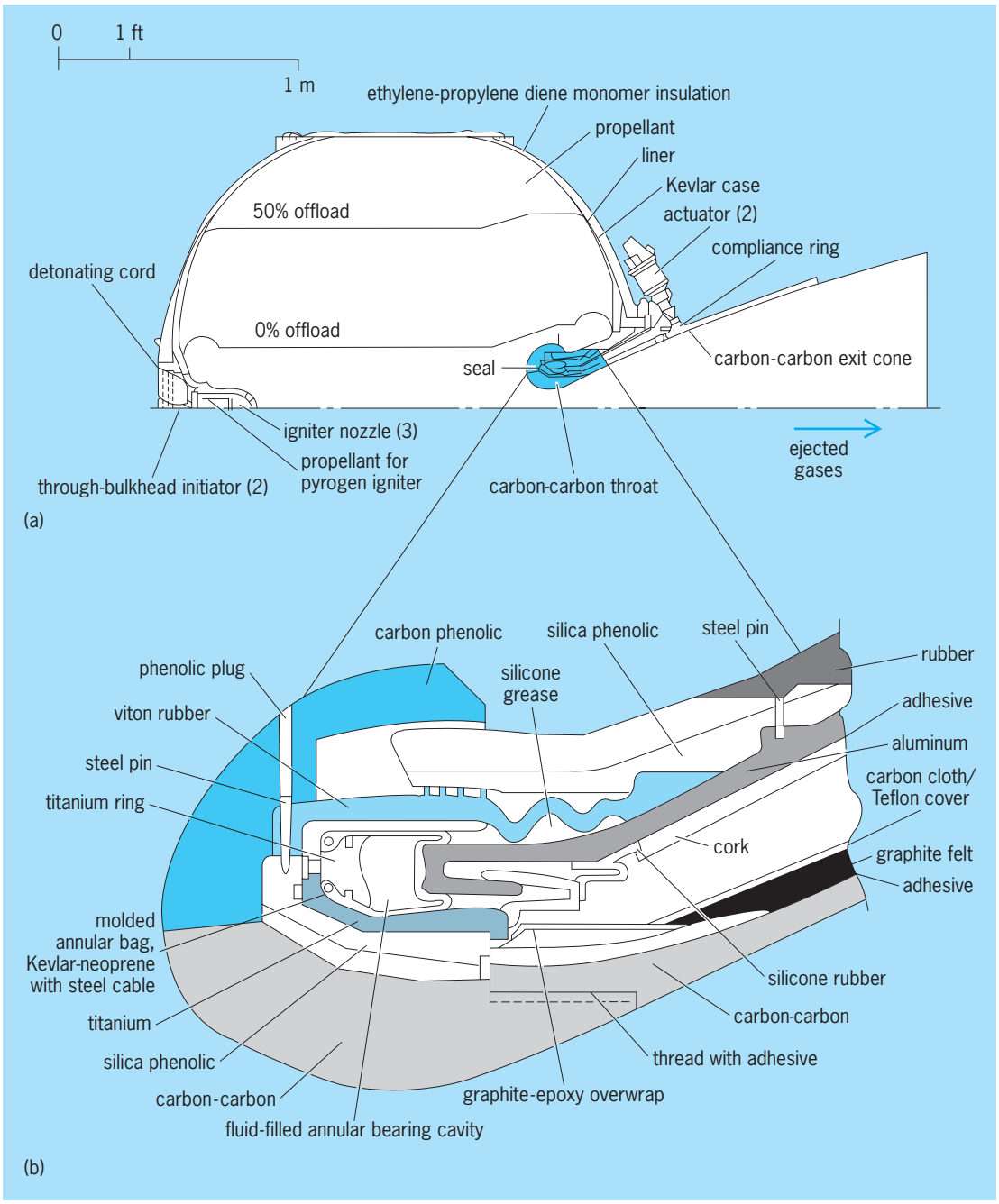

Fig. 11. Simplified cross sections of an upper-stage solid-propellant rocket motor (inertial upper stage; IUS) using an insulated carbon-fiber/carbon-matrix nozzle, an insulated carbon-filament-wound case, a pyrogen igniter, hydroxyl-terminated polybutadiene propellant, forward and aft stress-relieving boots, and a fluid-filled bearing and elastomeric seal assembly in the nozzle to allow 4.5° of thrust-vector deflection. (*a*) Half cross section along length of motor. (*b*) Enlarged cross section of nozzle package assembly. (*United Technologies Chemical Systems*)

All solid propellants, and many of the ingredients that are used in formulating them, are hazardous and have to be processed and handled using rigorous safety procedures. All solid propellants can be inadvertently ignited and thus provide thrust at an inopportune time and make the vehicle mobile, discharge hot flames that will start local fires, or cause areas to be exposed to toxic exhaust products. Under certain unfavorable circumstances, many solid rocket motors can explode.

Many solid propellants undergo an aging process consisting of a series of low-order chemical reactions between the fuel and the oxidizer ingredients (or sometimes between the propellant and the liner) during storage. Deterioration can often be minimized by the addition of stabilizers or inhibitors to the propellant. Nevertheless, the useful life of some solid rocket motors is limited to several years because aging and stress cycling degrade the physical properties or performance of the propellant, and may even make the rocket motor unsafe to operate.

Typical strengths of cast solid propellants vary between 25 and 500 lb/in.^{2} (0.17 and 3.4 MPa). This strength usually decreases as the ambient

TABLE 6. Characteristics of some operational solid propellants

Propellant type*	Specific impulse range, lbf·s/lbm (kN·s/kg)	Flame temperature, °F (°C)	Specific gravity	Metal content, weight %	Stress/strain, (lbf/in.^{2})/% (MPa/%) −60 °F (−51 °C)	Stress/strain, (lbf/in.^{2})/% (MPa/%) +150 °F (66 °C)	Processing method
DB	220–230 (2.15–2.25)	4100 (2250)	1.61	0	4600/2 (31.7/2)	490/60 (3.38/60)	Extruded
DB/AP-HMX/Al	265–270 (2.60–2.65)	6700 (3700)	1.80	20	2375/3 (16.4/3)	50/33 (0.34/33)	Solvent cast
PU/AP/Al	260–265 (2.55–2.60)	5400–6000 (3000–3300)	1.77	16–20	1170/6 (8.1/6)	75/33 (0.52/33)	Cast
Polymer/AN	180–196 (1.77–1.92)	2600 (1400)	1.55	0	400/2 (2.75/2)	50/15 (0.34/15)	Extruded
CTPB/AP/Al	260–265 (2.55–2.60)	5600–5800 (3100–3200)	1.77	15–17	325/26 (2.25/26)	88/75 (0.61/75)	Cast
HTPB/AP/Al	260–265 (2.55–2.60)	5600–5800 (3100–3200)	1.86	4–17	910/50 (6.3/50)	90/33 (0.62/33)	Cast
PBAA/AP/Al	260–265 (2.55–2.60)	5400–6000 (3000–3300)	1.77	14	500/13 (3.45/13)	41/31 (0.28/31)	Cast

*Al = aluminum; AN = ammonium nitrate; AP = ammonium perchlorate; CTPB = carboxy-terminated polybutadiene; DB = double base; HMX = cyclotetramethylene tetranitramine; HTPB = hydroxy-terminated polybutadiene; PBAA = polybutadiene-acrylic acid polymer; PU = polyurethane. The first propellant is typical of a double-base type, the second of a composite modified double base, and the other five are composite propellants.

temperature rises, and therefore most solid-propellant grains have a tendency to become soft on hot days or brittle on cold days. It is therefore necessary to restrict the temperature range over which certain propellants and grain designs may operate.

The finished, cast-in-place grain must absorb a variety of loads caused by thermal expansion and contraction due to ambient temperature variations, shocks from the sudden pressurization at start, rough handling, vibrations during transport, and flight accelerations. Solid propellant is a viscoelastic material that can suffer damage (cracks) and can become weaker by repeated applications of load at low stress such as in temperature cycling or transport vibrations. If the propellant should crack from an excessive temperature variation or rough handling, additional burning surfaces would be created in the cracks and an unregulated increase in the operating chamber pressure would cause failure of the chamber. It is necessary to handle propellants and solid rocket motors gently, so as not to subject them to sudden shocks by dropping or jarring.

As mentioned above, composite solid-propellant grains are frequently cast and case-bonded to the combustion chamber. Casting requires a fuel-binder mixture that can be safely handled as a viscous liquid at elevated temperatures (150–300°F or 65–150°C) and that cures and solidifies after it is poured into the case. Case bonding places stringent requirements on the mechanical properties of the propellant. About 65–85% of the propellant is dispersed small solids (namely, oxidizer crystals and aluminum powder), the mechanical properties are controlled by the 10–20% of binder and minor additives. This binder must absorb a wide range of stresses and differences in thermal expansion.

Special categories of solid propellants are smokeless propellants (which make detection more difficult) and propellants with nontoxic exhaust gas. Smokeless propellants contain little or no powered aluminum, which ordinarily would form solid particles of aluminum oxide in the exhaust gas. Most composite propellants contain a perchlorate oxidizer, which forms gaseous hydrochloric acid, a very toxic gas. Double-base and ammonium nitrate–based propellants can be both smokeless and nontoxic.

Components. These include the case, nozzle, igniter hardware, air seal, and safety devices. Sometimes there are also instruments, thrust vector control devices, and provisions for emergency shutdown, such as blow-out diaphragms and other devices.

The case confines the propellant grain and the reaction products of combustion. It is generally made of high-strength material, such as alloy steel, or high-strength fiber-reinforced plastics applied by a computer-controlled filament-winding machine. The case is usually protected on its exposed surfaces from excessive heating by internal insulating layers. Fiber-reinforced plastic case structures are usually lighter than comparable steel cases. The case has attachments to the vehicle structure for transmitting the thrust forces and a seal to prevent leakage between the case and the nozzle, which is usually fabricated separately. *See* COMPOSITE MATERIAL.

The nozzle must be protected from excessive heat transfer, from high-velocity hot gases, from erosion by small solid or liquid particles in the gas (such as aluminum oxide), and from chemical reactions with aggressive rocket exhaust products. The highest heat transfer and the most severe erosion occur at the nozzle throat and immediately upstream from there. Special composite materials, called ablative materials, are used for heat protection, such as various types of graphite or reinforced plastics with fibers made of carbon or silica. The development of a new composite material, namely, woven carbon fibers in a carbon matrix, has allowed higher wall temperatures and higher strength at elevated temperatures; it is now used in nozzle throats, nozzle inlets, and exit cones. It is made by carbonizing (heating in a nonoxidizing atmosphere) organic materials, such as rayon or phenolics. The fiber orientation influences the erosion resistance and strength. Multiple layers of different heat-resistant and heat-insulating materials are often

particularly effective. A three-dimensional pattern of fibers created by a process similar to weaving gives the nozzle extra strength. The divergent nozzle section does not experience such severe heat loads or erosion and often can be manufactured with thinner and less complex or less costly ablative materials.

The ablative materials are partly burned, charred, eroded, and progressively vaporized, thus absorbing some heat and forming a relatively cooler gas boundary layer. During this process the throat area can actually increase slightly as wall material is consumed. If the erosion is unsymmetrical, it may cause a slight misalignment of the thrust direction, which can result in vehicle flight control problems.

The nozzles shown in Figs. 10 and 11 have sophisticated thrust-vector control mechanisms. In Fig. 11 the nozzle forces are absorbed by a doughnut-shaped, confined, liquid-filled bag, in which the liquid moves as the nozzle is canted. This is a more detailed version of the flexible nozzle joint shown in Fig. 6. The space shuttle solid rocket boosters have gimbaled nozzles for thrust-vector control, with actuators driven by auxiliary power units and hydraulic pumps.

Most igniters use a pyrotechnic powder charge and an electric initiator to cause ignition of a separate propellant inside the igniter. In some motors there are also thrust termination devices, such as protechnically activated blow-out ports at the forward end of the case, that would allow gases to be ejected in a direction opposite to the nozzle.

There is usually a seal or diaphragm at the nozzle that is burst when the rocket motor starts. It prevents access of moisture and atmospheric oxygen which can cause a deterioriation of the propellant in the case or the igniter.

In many rocket motor cases the opening of the nozzle attachment or the igniter attachment is much smaller than the outside case diameter. During fabrication a mandrel is inserted through this relatively small hole prior to casting to ensure the proper geometry of the perforations and port area. This mandrel or fixture, which can have a complex geometry, is coated with an inert material [such as poly(tetrafluoroethylene)] to allow withdrawal after propellant casting. In some units the mandrel is made of porous, foamlike propellant that stays permanently in the motor and has a very high burning rate.

Grain configuration and burning. The thrust of the rocket is about equal to the product of mass flow and effective exhaust velocity, as shown in Eq. (2). The mass flow rate is equal to the product of the exposed burning area, the rate of burning, and the density of the propellant. By designing the grain so that more or less surface is exposed, the designer can increase or decrease the thrust during rocket motor operation (**Fig. 12**). It is also possible to limit the exposed burning surfaces by applying inhibitors (special inert or slow-burning chemicals) to parts of the surface. Grains that have large quantities of inhibitor applied to their exposed surfaces, called restricted-burning grains, are differentiated from grains that burn on all exposed surfaces. Holes put into the grain to increase the burning surface are known as perforations.

If the grain design is such that the thrust increases during the operation of the rocket, it is called progressive burning; if thrust decreases during operation, it is regressive burning; if the thrust or the chamber pressure stays approximately constant, it is neutral burning.

In restartable solid rocket units, the combustion chamber is divided into segments or wafers separated by a flame barrier. When all the propellant of the first segment is burned, the flame is stopped at the barrier. The barrier is destroyed upon ignition of the next segment.

The burning rate of solid propellant is measured in a direction normal to the burning surface and is expressed as a linear distance per unit time. Different propellants have burning rates between 0.2 and 0.8 in./s (5 and 20 mm/s), but some burn as slowly as 0.02 in./s (0.5 mm/s) or as fast as 3.5 in./s (90 mm/s). The burning rate usually varies as the chamber pressure to a power between 0.4 and 0.9, and is also a function of the initial ambient temperature. The same grain gives more thrust but a proportionally

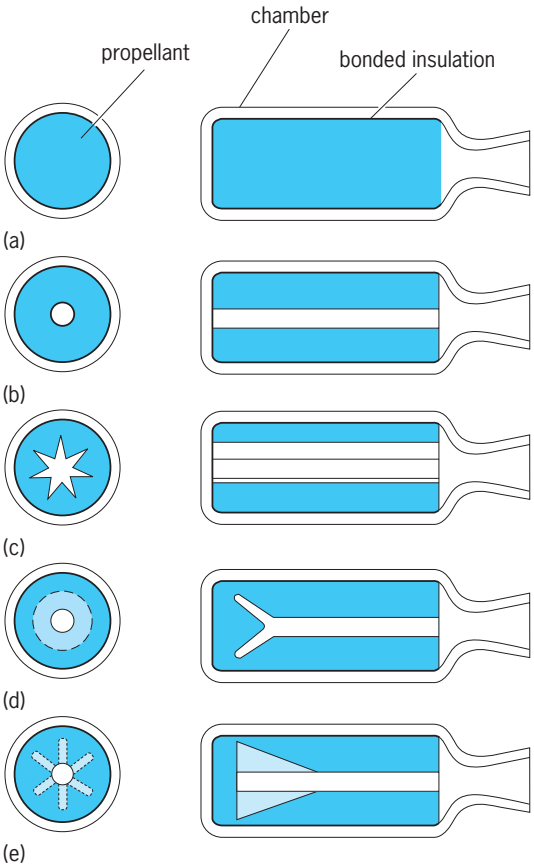

Fig. 12. Several solid-propellant grain configurations. In each part of the figure the diagram at the left is the cross section perpendicular to the length of the motor, and the diagram at the right is the cross section along the length of the motor. All configurations shown are case-bonded. (*a*) End burner, neutral burn. (*b*) Internal tube (end restricted), progressive burn. (*c*) Internal star, neutral burn. (*d*) Cone and cylinder, neutral burn. (*e*) Fins and cylinder, neutral burn.

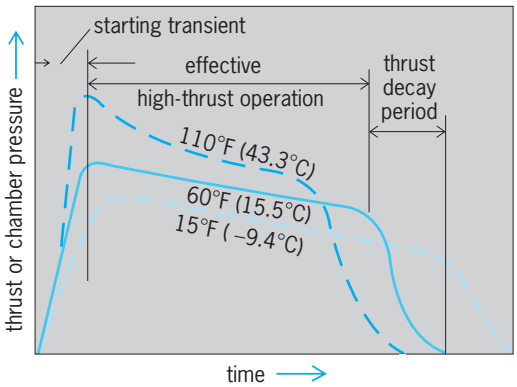

Fig. 13. Thrust-time variation of a typical solid-propellant rocket with slightly regressive burning characteristics. Broken alternative curves show effect of ambient grain temperature. Design of grain and varieties of propellants affect thrust and chamber pressure.

shorter firing duration at higher temperatures (**Fig. 13**). The burning rate also depends on the chemical and physical composition of the propellant; it can be changed by adding special burning-rate accelerators (such as iron compounds) or inhibitors, and is influenced by the erosive action of high-speed hot gases flowing past the internal grain surfaces. There is a maximum pressure above which smooth combustion can no longer be sustained and detonations may occur, and a minimum pressure below which stable, consistently smooth burning does not seem possible.

Comparison of propulsion systems. Both solid- and liquid-propellant systems have several important advantages that must be considered in the selection of a propulsion system. Solid-propellant motors have a simpler design, with few moving parts; are easier to operate, requiring little preflight checkout; have propellants that usually do not leak or spill; often require less overall weight for low-impulse applications; can be throttled or stopped and restarted a few times if preprogrammed (not at random); have propellants that can be stored for 5–20 years; and can be designed for recovery and reuse. Liquid-propellant engines have higher specific impulse; can be randomly throttled; have an accurately controllable thrust profile; can be randomly stopped and restarted, and pulsed many times; have a more controllable cutoff impulse; can be largely checked out just prior to operation; have easier thrust-vector control; can be designed for reuse after field services and checking; have storable propellants that can be kept in the vehicle for a very long time; and are only rarely subject to catastrophic failure.

Both types of propulsion system also have several important disadvantages that must be considered. In the case of solid-propellant motors, propellants can deteriorate in storage; the explosion and fire potential is larger; failure is often catastrophic; the grain displays cumulative damage with temperature cycling or rough handling, and therefore must be handled carefully; factory rework is required before reuse; each restart requires a separate ignition system; the thrust-vector control system may include liquid propellant; exhaust gases are usually toxic; and the system can be stopped at random only by disabling the rocket, and, once ignited, cannot change thrust or duration. In the case of liquid-propellant engines, the design is relatively complex, with more components; cryogenic propellants cannot be stored; propellant spills can be hazardous, corrosive, and toxic, and can cause fires; nonhypergolic propellants require an ignition system; combustion instability is more difficult to control; more overall weight is required for short-duration low-total-impulse applications; sloshing of liquids in the tank can cause flight stability problems; and, if the tank outlet is uncovered during sloshing, gas in the propellant lines can cause combustion interruption or instabilities.

Hybrid rocket propulsion. A hybrid uses a liquid propellant together with a solid propellant in the same rocket engine (**Fig. 14**). The arrangement of the solid fuel is similar to that of the grain of a solid-propellant rocket; however, no burning takes place directly on the surface of the grain because it contains little or no oxidizer. Instead, the fuel on the grain surface is heated, decomposed, and vaporized, and the vapors burn with the oxidizer some distance away from the surface. The combustion is therefore inefficient. Hybrids usually require a separate ignition device. Hybrids can be restarted and throttled, although at a sacrifice in performance, and they usually lend themselves to safe and rugged construction. Successful experimental target drones have been powered by liquid nitric oxides and a solid fuel consisting of an acrylic plastic with a magnesium compound as additive.

Testing. Because flights of rocket-propelled vehicles are usually fairly expensive and because it is sometimes difficult to obtain sufficient and accurate data from fast-moving flight vehicles, it is accepted practice to test rocket propulsion systems and components extensively on the ground under simulated flight conditions. Components such as an igniter or a turbine are tested separately. Complete engines are tested in static engine test stands; the complete vehicle stage is also tested statically. In the latter two tests the engine and vehicle are adequately secured by suitable structures. Only in flight tests are they allowed to leave the ground.

The static test facilities must be capable of handling and disposing of the hot gases (5000°F or 2760°C) that are expelled from the rocket. For those rockets that must be tested in a vertical position, water-cooled metal flame buckets are used to prevent

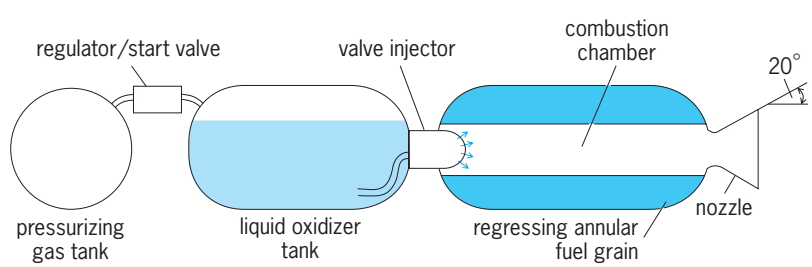

Fig. 14. Hybrid rocket propulsion system.

ground erosion. Propulsion systems which must work in a vacuum, such as space-vehicle engines, are usually also tested in a chamber with an artificial vacuum.

Since rocket propellants are dangerous to handle (they can explode, releasing an energy higher than that of an equivalent weight of high explosive), special precautions and safety provisions must be observed in fabrication, shipping, and testing. These include reinforced buildings of heavy construction that can withstand a blast, remote operation and control, an extensive series of personnel warning devices, special fire-fighting equipment, limitations on the amount of propellant in any one area and on the distance between buildings, the use of sparkproof tools, and rigorous training of personnel.

The instrumentation is usually of the remote indicating or remote recording type. Pressures, temperatures, forces, operating sequence, flows, vibrations, and strains are some of the more common parameters measured. Special instruments are needed for many of these, for example, a pressure recorder with unusually high-frequency response for combustion gas vibrations and optical temperature indicators for high-temperature gas.

Many tests are computer controlled, in part because a computer can monitor parameters and cause a parameter variation or an emergency shutdown faster than a human operator. Most test facilities use computers for data recording, data handling, and instrument calibration corrections. *See* LAUNCH COMPLEX.

George P. Sutton

Bibliography. J. W. Cornelisse et al., *Rocket Propulsion and Spaceflight Dynamics*, Pitman, 1979; A. Davenas (ed.), *Solid Rocket Propulsion Technology*, Pergamon, 1993; National Aeronautics and Space Administration, *National Space Transportation Reference*, vol. 1: *Systems and Facilities*, 1989; G. P. Sutton and O. Biblarz, *Rocket Propulsion Elements*, 7th ed., Wiley, 2001; M. J. L. Turner, *Rocket and Spacecraft Propulsion*, 2d ed., Springer, 2005.

Rocket staging

The use of successive rocket vehicle sections, each having its own propulsion system, such as one or more engines. One way to minimize the mass of large missiles, or space vehicles, is to use multiple stages. The first or initial stage is usually the heaviest and biggest and is often called the booster. It requires the highest thrust and largest rocket propulsion system. The next few stages are successively smaller, have much lower thrusts, and are generally called sustainers. Each stage is a complete vehicle in itself and carries its own propellant (either solid or liquid; both fuel and oxidizer) and its own propulsion system, and has its own tankage and control system.

Once the propellant of a given stage is expended, the dead or inert mass of that stage, including the hardware of a liquid-propellant rocket engine or a solid-propellant rocket motor with flight controls, is no longer useful in contributing additional kinetic energy to the succeeding stages. By dropping off this useless inert mass, the mass that remains to be accelerated is made smaller; therefore it is possible to accelerate the payload to higher velocity than would be attainable if multiple staging were not used.

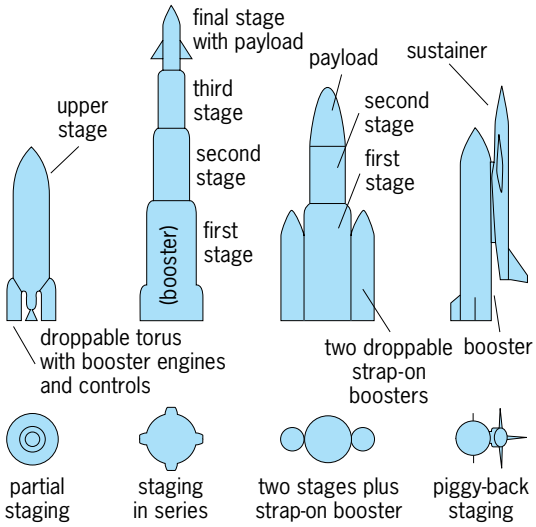

Typical schemes for staging missiles.

Use. Various staging configurations can be used. Multiple-stage arrangements are used in large space-launch vehicles, in some long-range ballistic missiles, various research rockets, and certain antiaircraft and antimissile military weapon systems. It is quite possible to employ different types of power plants, different types of propellants, and entirely different configurations in successive stages of any one multistage vehicle (see **illus.**). Because staging adds separation mechanisms and extra structure and reduces overall reliability, it is impractical to have more than four to six stages in any one vehicle. The Saturn V/Apollo Moon landing and return flight had six stages.

Advantages. By staging a missile or space vehicle, it is possible to improve its performance either by providing more range, more altitude, faster flight, or more payload, or by reducing the gross weight and the size of the vehicle for a given specific mission without diminishing the payload.

For a given initial launch vehicle configuration and mass, the payload can be increased, but at a decrease in flight performance, such as ballistic missile range or spacecraft orbit altitude. Furthermore, the payload also depends on the specific flight trajectory and the vehicle design. The maximum payload that can be carried by any particular rocket-propelled vehicle therefore has to be evaluated individually by using trajectory analyses and design variations. The use of additional stages makes possible an increase in the mass of the payload (such as instruments, passengers, and equipment) or warheads (see **table**). For a fixed payload the range of a ballistic missile can be increased by adding a stage.

By redesigning existing space-launch vehicles and incorporating an additional stage, it is possible to achieve more ambitious space missions. For example, the Titan 34D, a particular two-stage version of

Total payload as a function of the number of stages for specific missions (payload values relative to each other)

Mission	Missile takeoff weight, lb*	Payload, lb, for number of stages			
		1	2	3	4
1500-mi (2400-km) ballistic missile	50,000	2000	2700	†	†
6000-mi (9700-km) ballistic missile	500,000	‡	6000	9000	†
Moon impact	500,000	‡	100	1000	3000

* 1 lb = 0.45 kg.
† Gain from further staging is small.
‡ This mission cannot be achieved with a single-stage missile and chemical rocket propulsion. The dry weight is so large a percentage of the total weight that the missile will fall short of its target even with zero payload.

the Titan family of space-launch vehicles with two strap-on solid-propellant rocket motor boosters, can carry a payload of about 31,300 lb (14,200 kg) into a 100-nautical-mile-high (185.3-km) circular orbit, and a 15,000-lb (6818-kg) payload into a 5000-nmi (9265-km) circular orbit. This configuration cannot put payloads into high earth orbits, such as a synchronous orbit at 22,700 nmi (42,000 km) or a Moon or planetary mission. However, adding a third stage (in this case powered by an IUS solid-rocket motor) allows the three-stage Titan 34D to carry about 4000 lb (1820 kg) into a synchronous orbit.

Stage separation. Ideally it would be desirable to stop the propulsion unit of the operating stage at the same time that the operation of the propulsion unit of the next stage is initiated. This close timing is not possible. There are inherent difficulties in the separation of the stages caused by drag forces, the lack of gravity in space, the noninstantaneous starting and stopping of rocket engines, and the possible interference of various stages. Often a positive mechanical separation mechanism is included in the design that forces the two stages to separate. *See* ROCKET; ROCKET PROPULSION; SPACECRAFT PROPULSION.

George P. Sutton

Bibliography. M. D. Griffin and J. R. French, *Space Vehicle Design*, 1991.

Rodentia

The mammalian order consisting of the rodents, often known as the gnawing mammals. This is the most diverse group of mammals in the world, consisting of over 2000 species, more than 40% of the known species of mammals on Earth today. Rodents range in size from mice, weighing only a few grams, to the Central American capybara, which is up to 130 cm (4 ft) in length and weighs up to 79 kg (170 lb) [**Fig. 1**]. Rodents have been found in virtually every habitat, from arctic tundra to tropical rainforests, and on every continent except Antarctica. Rodents have adapted to nearly every mode of life, including semiaquatic swimming (beavers and muskrats), gliding ("flying" squirrels), burrowing (gophers and African mole rats), arboreal (dormice and tree squirrels), and hopping (kangaroo rats and jerboas). Nearly all rodents are herbivorous, with a few exceptions that are partially insectivorous to totally omnivorous, such as the domestic rat. Rodents usually live in small family groups, but the naked mole rats of Africa live in large colonies with a distinct social structure that is similar to that of insect colonies with different classes of members. The great adaptability and rapid evolution and diversity of rodents are mainly due to their short gestation periods (only 3 weeks in some mice) and rapid turnover of generations.

Anatomy. The most diagnostic feature of the Rodentia is the presence of two pair of ever-growing incisors (one pair above and one below) at the front of the jaws. These teeth have enamel only on the front surface, which allows them to wear into a chisellike shape, giving rodents the ability to gnaw. Associated with these unique teeth are a number of other anatomical features that enhance this ability. Behind the incisors is a gap in the jaws where no teeth grow,

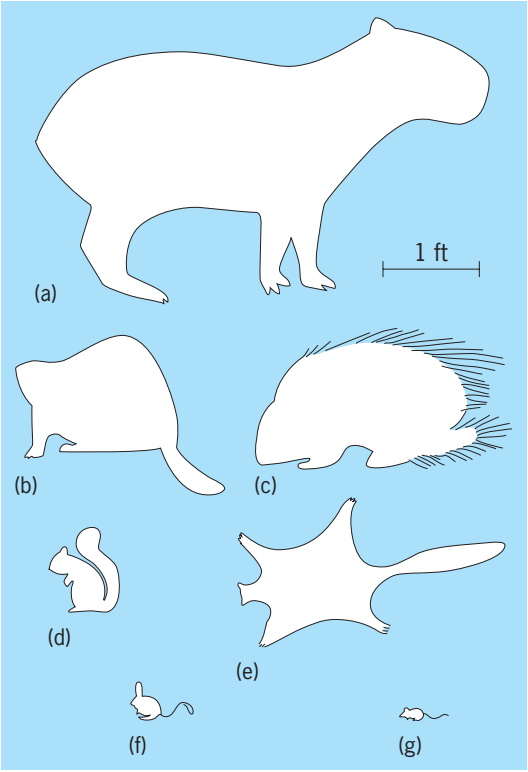

Fig. 1. Silhouettes of modern rodents, showing their wide range in size and shape. (*a*) Central American capybara (*Hydrochoerus*). (*b*) Beaver (*Castor*). (*c*) Old World porcupine (*Hystrix*). (*d*) Tree squirrel (*Sciurus*). (*e*) Asian giant flying squirrel (*Petaurista*). (*f*) Asian jerboa (*Allactaga*). (*g*) Common house mouse (*Mus*).

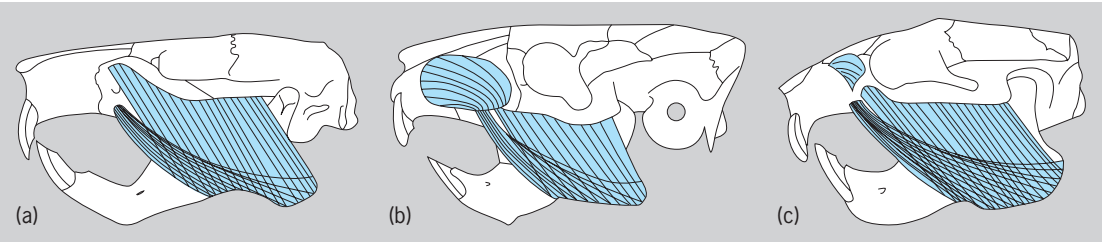

Fig. 2. Skull musculature of three fossil rodents exemplifying the three major groupings of rodents used in the most common classification scheme. (*a*) Sciuromorpha: *Eutypomys thompsoni*, a primitive beaverlike rodent from the Oligocene of North America. (*b*) Hystricomorpha: *Neoreomys australis*, a Miocene South American rodent. (*c*) Myomorpha: *Eumys elegans*, an early "field mouse" from the Oligocene of North America. Skulls are not to scale. (*Redrawn from A. E. Wood, Symposium of the Zoological Society of London, 34:21–60, 1974*)

called a diastema. The diastema of the upper jaw is longer than that of the lower jaw, which allows rodents to engage their gnawing incisors while their chewing teeth (molars and premolars) are not being used. The reverse is also true; rodents can use their chewing teeth (also called cheek teeth) while their incisors are disengaged.

The entire skull structure of rodents is designed to accommodate this task of separating the use of the different types of teeth. Rodent skulls have long snouts; the articulation of the lower jaw with the skull is oriented front to back rather than sideways as in other mammals; the jaw muscles (masseter complex) are extended well forward into the snout; and the number of cheek teeth is less than in most other mammals—all features unique to rodents.

The cheek teeth of rodents vary greatly, from very low crowned teeth only a few millimeters high (squirrels or hamsters) to cheek teeth that (like the incisors) continue to grow throughout the life of the animal (voles, beavers, and chinchillas). In the species with low-crowned cheek teeth, enamel covers the entire tooth surface until it is worn away in extreme old age. Those species with very high crowned (or ever-growing) cheek teeth usually have teeth that are surrounded by enamel on the sides only and have a chewing surface made mostly of dentine. The maximum number of cheek teeth in rodents is two premolars and three molars (on one side) on the upper jaw and one premolar and three molars in the lower jaw. The maximum dental formula for rodents is 1-0-2-3/1-0-1-3, where first is the number of incisors, second the number of canines, third the number of premolars, and fourth the number of molars. This dental formula refers to only one side of the jaw, so the maximum number of teeth in any rodent is 22. Some species of African naked mole rats have only two molars and no premolars, reducing their total number of teeth to 12. *See* DENTITION.

The bodies and skeletons of rodents vary greatly, as do their habitats and modes of life. Those that swim have webbed feet and even paddlelike tails. Those that glide have a special flap of skin that attaches the front and hind limbs together for a sort of wing. Those that are burrowers develop enlarged forearms and hands for digging, and those that hop have elongated hindlimbs and smaller forelimbs. Other features are unique to certain families of rodents, such as the fur-lined cheek pouches for storing food in the pocket mice and gophers. The naked mole rats of Africa have virtually no fur at all.

Classification. The classification of rodents has always been difficult because of the great diversity of both Recent and fossil species. Traditionally, there are two ways that rodents have been divided: into three major groups based on the structure of the attachment of the jaw muscle (masseter) on the skull (Sciuromorpha, Hystricomorpha, Myomorpha) [**Fig. 2**]; or into two groups based on the structure of the lower jaw (Sciurognathi, Hystricognathi). The difficulty in using these groups (usually considered suborders or infraorders) is that the distinctive adaptations of one group of rodents are also present in others, derived in completely separate ways. Thus, any classification of rodents has always had a number of families that cannot be designated to any one of the suborders. In some cases, new suborders of rodents are made to accommodate the families that do not fall into the traditional categories, but these groups usually consist of a number of unrelated families and are referred to as "waste basket" groups. Still, the most commonly used classification is that of the three major groups based on the musculature of the skull.

Fossil history. The earliest record of rodents is from the late Paleocene of both North America and Inner Mongolia. These earliest examples already have all of the distinguishing features of rodents but belong to primitive extinct families. By the early Eocene, rodents are also known from Europe, India, and Africa. In the early Oligocene, the first rodents appear in South America. Rodents do not occur in Australia until the Pleistocene (**Fig. 3**).

The pattern of rodent evolution is characterized by the development of endemic lineages on each of the continents, with numerous intercontinent emigration events involving the more successful groups. There are more than a dozen extinct families of rodents along with the 30 or more recognized extant families. The Eocene is dominated by species of extinct families in North America and Europe; the still-extant family of ctenodactyloid rodents (gundis) in eastern Asia and India; and an extinct family, the Phyomyidae, in Africa (believed to be ancestral to later families that originated in Africa). By the end of the Eocene, the earliest a members of many modern

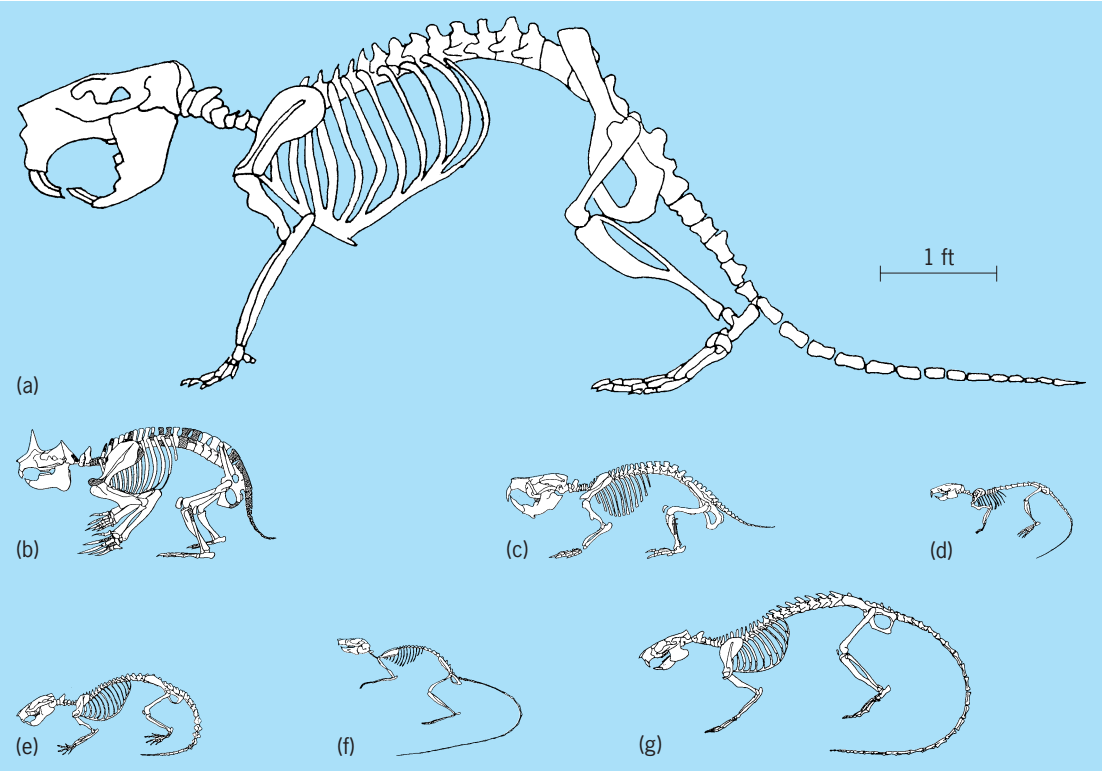

Fig. 3. Reconstructions of the skeletons of fossil rodents from North America. (*a*) Pleistocene giant beaver *Castoroides ohioensis*. (*b*) Late Miocene "horned" rodent *Epigaulus hatcheri*. (*c*) Late Oligocene to early Miocene burrowing beaver *Palaeocastor fossor*. (*d*) Late Oligocene to early Miocene "field mouse" *Geringia mcgregori*. (*e*) Late Miocene to early Oligocene primitive rodent *Ischyromys typus*. (*f*) Late Eocene "jumping" rodent *Protoptychus hatcheri*. (*g*) Early Eocene primitive rodent *Paramys delicatus*. (*Part a redrawn from A. S. Romer, Vertebrate Paleontology, University of Chicago Press, 1966; b from J. W. Gidley, A new horned rodent from the Miocene of Kansas, Proc. U.S. Nat. Mus. 32:627–636, 1907; c–g from W. W. Korth, The Tertiary Record of Rodents in North America, Plenum Press, New York, 1994*)

families had appeared: in North America, the ancestors of pocket mice, gophers, and jumping mice; in Europe, the earliest dormice; in Asia, the cricetid rodents (field mice, hamsters); and in Africa, still geographically isolated, the endemic species of phiomyids and the extant family Anomaluridae (scaly tailed flying squirrels).

At the beginning of the Oligocene, there was a great turnover in the rodent faunas of the world, due to both migrations and evolution of new families. The cricetids from Asia spread to both North America and Europe. The first appearance of squirrels was nearly simultaneous in North America and Europe. Similarly, the earliest beavers occurred in both North America and Asia. In Asia, an extinct family of very specialized burrowing rodents that is restricted to that continent, the Tsagonomyidae, appeared. The earliest jerboas also appeared in Asia at this time. By the end of the Oligocene, most of the Eocene families of primitive rodents had become extinct. The Oligocene is significant in South America because the first record of rodents on that continent dates to that period; they are most similar to those from the Oligocene of Africa. This group of unique rodents in South America diversified without interaction with rodents from any other continent until the very end of the Miocene, when there was a great faunal interchange with North America due to the forming of the isthmus of Panama joining the two continents.

During the Miocene, all extant families of rodents appeared. A notable event was the first occurrence of the Muridae (true rats and mice) in central Asia. This group rapidly diversified and invaded Africa and Europe. Specialized members of the Muridae (voles, lemmings, and muskrats) developed near the end of the Miocene and appeared in both Eurasia and North America. The connection of North and South America allowed the spread of squirrels and cricetid rodents into South America, and the northward expansion of the endemic South American rodents (porcupines and nutria). In Africa, the endemic modern families appeared (springhares, naked mole rats) at the time of the extinction of the primitive African rodents (phiomyids) and the introduction of squirrels, Old World porcupines, and cricetids. During the Miocene, a family of unusual rodents, the Mylagaulidae, appeared and diversified in North America. These rodents are well adapted for burrowing, and some species developed horns on their snouts much like that of a rhinoceros. In Asia, the ctenodactyloid rodents, which had been dominant since the early Eocene, migrated to Africa and became extinct in Asia.

By the Pleistocene, nearly all the extant genera of rodents around the world had appeared in the

fossil record, producing a rodent fauna much like that of today. A notable exception is the gigantic beaver *Castoroides* from North America, which was up to 2.3 m (7 ft) long and weighed hundreds of pounds. These large creatures disappeared by the end of the Pleistocene, along with the other large mammals that were part of the late Pleistocene extinctions (mammoths, giant rhinos, giant bears). *See* EOCENE; MIOCENE; OLIGOCENE; PALEOCENE; PLEISTOCENE.

Economic and historical importance. A few rodent species have been trapped or farmed for their fur. Most notably, the beaver (*Castor*) was trapped in North America, which led to European exploration of the continent. Currently, muskrats (*Ondatra*), the South American chinchillas (*Chinchilla*), and nutria (*Myocastor*) are the main sources of rodent furs. Domestic rats and mice (*Rattus* and *Mus*) are almost universally used as laboratory subjects for medical research. Guinea pigs (*Cavia*) were domesticated in South America as a food source a few thousand years ago.

In the twentieth century a number of rodents were domesticated as pets: gerbils, hamsters, rats, mice (all Old World members of the mouse family), guinea pigs (from South America), and even prairie dogs (a member of the squirrel family from western North America).

Rodent damage to harvested crops is high around the world. Domestic rats consume large amounts of stored food. The amount of stored grains destroyed by these rodents has been estimated at as much as $1 billion per year in the United States alone. Other rodents have been responsible for the destruction of unharvested crops and grazing lands around the world.

Rodents have also been the source of the spread of disease. In the fourteenth century, tens of millions of people in Europe and Asia died from the bubonic plague, which was transmitted by the fleas that infested rats. Other diseases, such as Lyme disease (Lyme borreliosis) and Rocky Mountain spotted fever, are spread similarly, not by rodents directly but by the parasites that they carry. Hemorrhagic fever viruses (such as hantaviruses) are also spread by rodents of many species worldwide. The diseases caused by these viruses, spread in air that carries microscopic particles from rodent secretions or excretions, are severe and often result in death.

William W. Korth

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, H. Freeman, New York, 1988; J. E. Childs, Special feature: Zoonoses, *J. Mammal.*, 76:663, 1995; W. W. Korth, *The Tertiary Record of Rodents in North America*, Plenum Press, New York, 1994; M. C. McKenna and S. K. Bell, *Classification of Mammals above the Species Level*, Columbia University Press, New York, 1997; R. M. Nowak, *Walker's Mammal Species of the World*, Johns Hopkins University Press, Baltimore, 1991; D. E. Savage and D. E. Russell, *Mammalian Paleofaunas of the World*, Addison-Wesley, London, 1983; D. E. Wilson and D. M. Reeder, *Mammal Species of the World*, Smithsonian Institution Press, Washington, DC, 1993.

Rodenticide

A toxic chemical that is used to kill pest rodents and sometimes other pest mammals, including moles, rabbits, and hares. Rodenticides are the least widely used of the pesticides, which also include insecticides, fungicides, and herbicides.

Most rodenticides are used to control rats and house mice, the common rodents that infest urban, suburban, and rural buildings and feed on a wide variety of edible commodities. Their droppings and urine contaminate several times more food than is actually consumed. Rats and mice cause destruction by gnawing on buildings and valuable items such as books, paintings and other art works, and rare manuscripts. They also can transmit a number of serious diseases to people and domestic animals. It is a common practice to make buildings rodent-proof in an effort to reduce the need for rodenticides.

Certain field and forest rodents, such as pocket gophers, ground squirrels, prairie dogs, and voles, become pests when they conflict with crop production, reforestation, and environmental quality. When nonlethal methods either fail to resolve the problem or are impractical, these animals must be controlled with rodenticides.

Use formulations. Rodenticides are generally combined with some rodent-preferred food item such as grain (corn, wheat, oats) or a combination of grains in low yet effective amounts. Bait formulations may be in pellet forms or incorporated in paraffin blocks of varying sizes. Loose-grain kernels and ground-meal baits are also available. As a safeguard against accidental ingestion by nontarget species, baits are placed either where they are inaccessible to children, domestic animals, or wildlife, or within tamper-resistant bait boxes designed to exclude all but rodent-size animals. Some exceptions exist in agricultural situations. To prevent ingestion by children, a bitter substance, such as denatonium benzoate, is sometimes added at concentrations that do not adversely affect rodent consumption.

Types. As a group, anticoagulant rodenticides dominate the market and are sold under a wide variety of trade names. In order of their development, they are warfarin, pindone (Pival), diphacinone, and chlorophacinone. When small amounts of these anticoagulants are consumed over several days, death results from internal bleeding. The newer, second-generation anticoagulant rodenticides, such as brodifacoum, bromadiolone, and difethialone, were developed to counteract the growing genetic resistance in rats and house mice to the earlier anticoagulants, especially warfarin. The second-generation compounds are more potent and capable of being lethal following a single night's feeding, although death is generally delayed by several days. Vitamin K_1 is the antidote for all anticoagulant rodenticides, if administered soon enough after ingestion.

Rodenticides that do not belong to the anticoagulant group include zinc phosphide, bromethalin, and cholecalciferol (vitamin D_3). The feeding and lethal characteristics differ among them. Acutely

toxic strychnine baits are also available but are restricted to underground application, primarily for pocket gophers and moles.

A number of toxic agents have been used in the past as rodenticides but are not currently registered for use in the United States. These include arsenic, thallium sulfate, antu, red squill, sodium fluoroacetate (compound 1080), fluoroacetamide (1081), and cyanide. The reasons for their discontinuation vary, but environmental concerns and the availability of more efficacious rodenticides are the most significant.

Several lethal fumigants or materials that produce poisonous gases are used to kill rodents in burrows and within other confined areas such as unoccupied railway cars or buildings. Lethal fumigants include aluminum phosphide, carbon dioxide, chloropicrin, and smoke or gas cartridges, which are ignited to produce carbon monoxide and other asphyxiating gases.

Regulations. Because of their high toxicity, rodenticides are inherently hazardous to people, domestic animals, and wildlife. They are highly regulated, as are certain other types of pesticides. Rodenticide products must be proven effective and most demonstrate that they can be used without presenting unacceptable hazards to nontarget species or to the environment. Some rodenticides can be purchased and used only by trained certified or licensed pest control operators, while others with a greater safety margin can be used by the general public. Rodenticide products, prior to marketing, are registered under the comprehensive rigid criteria of the U.S. Environmental Protection Agency (EPA). States may also require pesticide registration and may demand safety and other additional pesticide data beyond that required by the EPA. Certain rodenticides may be registered with EPA approval for specific uses only in a particular state. All rodenticides must be used in accordance with the label directions and may be prohibited where they may jeopardize certain endangered species. *See* PESTICIDE; RODENTIA.

Rex E. Marsh

Bibliography. L. L. Bergeson (ed.), *FIFRA: Federal Insecticide, Fungicide and Rodenticide Act* (*Basic Practice Series*), 2000.

Roentgenium

The nineteenth of the synthetic transuranium elements, symbol Rg, atomic number 111. Discovered in late 1994, roentgenium should be a homolog of the elements gold, silver, and copper. It is the ninth element in the *6d* shell. *See* COPPER; GOLD; HALF-LIFE; RADIOISOTOPE; SILVER.

The element was discovered on December 17, 1994, at GSI (Gesellschaft für Schwerionenforschung), Darmstadt, Germany, by detection of the isotope ^{272}Rg (the isotope of roentgenium with mass number 272), which was produced by fusion of a nickel-64 projectile and a bismuth-209 target nucleus after the fused system was cooled by emission of one neutron. The optimum bombarding energy for producing ^{272}Rg corresponds to an excitation energy of 15 MeV for the fused system. Sequential alpha decays to meitnerium-268, bohrium-264, dubnium-260,

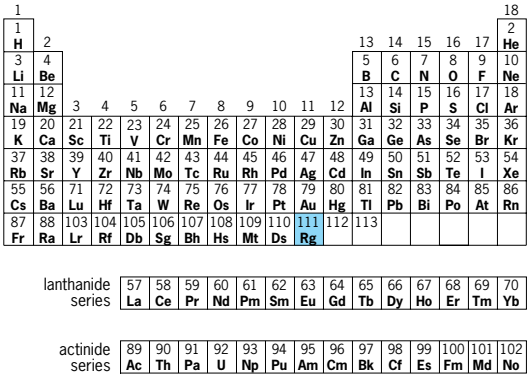

and lawrencium-256 allowed identification from the known decay properties of ^{260}Db and ^{256}Lr. In the decay chain (see **illustration**), the first three members are new isotopes. The isotope ^{272}Rg has a half-life of 1.5 ms, and is produced with a cross section of 3.5×10^{-36} cm^2. Altogether, three chains were observed during the 17 days of irradiation. The new isotopes meitnerium-268 and bohrium-264 with their half-lives of 70 ms and 0.4 s, respectively, show a trend toward longer half-lives than those of the previously known isotopes of these elements. *See* ALPHA PARTICLES; NEUTRON; NUCLEAR REACTION.

The discovery of roentgenium followed the discovery of ^{269}Ds within 5 weeks. Both elements were produced by the same techniques. The production

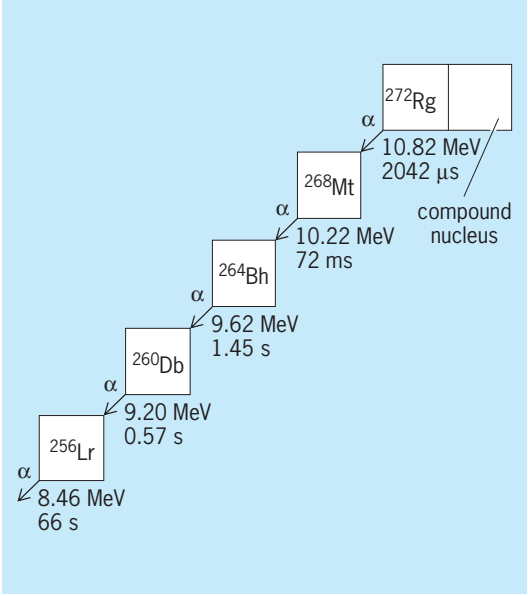

Sequence of decay chains that document the discovery of roentgenium. Numbers below boxes are alpha energies and correlation times. Roentgenium is produced in the reaction ^{64}Ni + ^{209}Bi → ^{272}Rg + 1*n*.

cross section decreases by a factor of 4 compared to the even darmstadtium. This factor is smaller than the corresponding factor of 9 between the cross sections to produce elements meitnerium and hassium. *See* DARMSTADTIUM; HASSIUM; NUCLEAR STRUCTURE.

The group of RIKEN, Japan, in 2004 observed a single decay chain of $^{278}113$, which shows as first descendant the new isotope ^{274}Rg. Peter Armbruster

Bibliography. S. Hofmann et al., The new element 111, *Z. Phys. A*, 350:281–282, 1995.

Roll mill

A series of rolls operating at different speeds. Roll mills are used to grind paint or to mill flour. In paint grinding, a paste is fed between two low-speed rolls running toward each other at different speeds. Because the next roll in the mill is turning faster, it develops shear in the paste and draws the paste through the mill. The film is scraped from the last high-speed roll. For grinding flour, rolls are operated in pairs, rolls in each pair running toward each other at different speeds. Grooved rolls crush the grain; smooth rolls mill the flour to the desired fineness. The term roller mill is applied to a ring-roll mill. *See* GRINDING MILL. Ralph M. Hardgrove

Rolling contact

Contact between bodies such that the relative velocity of the two contacting surfaces at the point of contact is zero. Common applications of rolling contact are the friction gearing of phonograph turntables, speed changers, and wheels on roadways. Rolling contact mechanisms are, generally speaking, a special variety of cam mechanisms. An understanding of rolling contact is essential in the study of antifriction bearings. The concepts are also useful in the study of the behavior of toothed gearing. *See* ANTIFRICTION BEARING; GEAR; MECHANISM.

Pure rolling contact can exist between two cylinders rotating about their centers, with either external or internal contact. Two friction disks (**Fig. 1**) have external rolling contact if no slipping occurs between them. The rotational speeds of the disks are then inversely proportional to their radii.

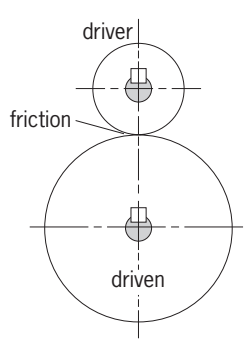

Fig. 1. Rolling friction disks.

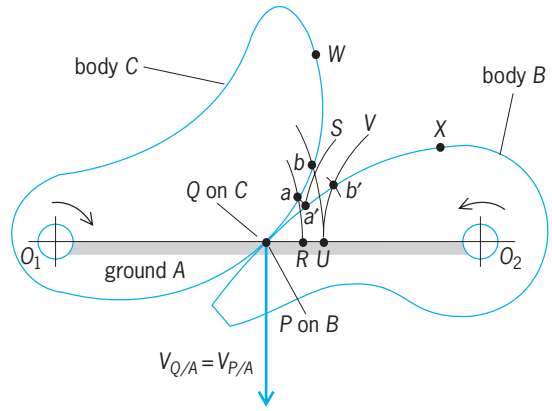

Fig. 2. Pure rolling between bodies in contact. (*After H. A. Rothbart, Cams, John Wiley and Sons, 1956*)

The general case of pure rolling between contacting bodies is illustrated in **Fig. 2**. Particle Q on body C and particle P on body B coincide at the moment. The velocities of these particles relative to ground frame A must be the same if there is to be no slip. It can be shown, using the theory of kinematics of mechanisms, that for two bodies to have rolling contact at an instant the point of contact must lie on the line joining the two centers of rotation. *See* MECHANISM; VELOCITY ANALYSIS.

Figure 2 shows that if the contacting particles Q and P were anywhere except on the line of centers O_1O_2 their velocities would necessarily be different, for they would have different directions. Because the centers of rotation are fixed, the sum of the distances from the centers of rotation to the point of contact must equal the center distance. If, in Fig. 2, W and X are two points that ultimately will come into contact, then the bodies must be designed to satisfy Eq. (1).

$$O_1W + O_2X = O_1O_2 \tag{1}$$

Finally, because there is to be no slip between the contacting surfaces, the length of the arc of action on each body must be the same. In Fig. 2, for example, arc length QW must equal arc length PX.

At an instant the angular velocities of the two links are inversely proportional to the distances from the centers of rotation to the point of contact, stated mathematically as Eq. (2).

$$\frac{\omega_C}{\omega_B} = \frac{O_2P}{O_1Q} \tag{2}$$

A graphical technique for designing the shape of body B so that it will have pure rolling contact with given body C is illustrated in Fig. 2. Select a sequence of points along the profile of body C such as a, b, and so on. Then the corresponding points a', b', and so on of body B must be found. Revolve point a about O_1 to the line of centers, giving point R. With O_2 as center and O_2R as radius, draw an arc RS of indefinite length. Because arc length Pa' must equal arc length Qa, the location of a' can be found approximately by using chord Qa as a compass distance and swinging an arc from R to intersect arc RS at a'. The same

construction is used for finding b'. By repeating this procedure for a number of closely spaced points, the profile of body B can be determined.

Pure rolling occurs between the bodies with profiles that are the basic plane curves, that is, logarithmic spirals, ellipses, parabolas, and hyperbolas. To achieve pure rolling between these curves, the centers of rotation must be at the foci of the curves; the curves must be properly proportioned and positioned with respect to one another (**Fig. 3**).

Logarithmic spirals of equal obliquity are shown pivoting about their foci in Fig. 3*a* and *b*. In the first instance the links turn in opposite directions; in the second they turn in the same direction. The latter arrangement is the more compact. A logarithmic spiral rolling against a straight-sided follower is shown in Fig. 3*c*. Here the follower has a reciprocating motion. For the pair of ellipses shown in Fig. 3*c*, pure rolling contact can occur for complete rotations of both members. Rolling parabolas, as in Fig. 3*e*, provide a reciprocating output motion. Where space is limited, the equal hyperbolas of Fig. 3*f* are an especially effective choice.

A difficulty with most of these arrangements is that only partial rotations are possible. The lobe wheels shown in Fig. 3*g*, on the other hand, make

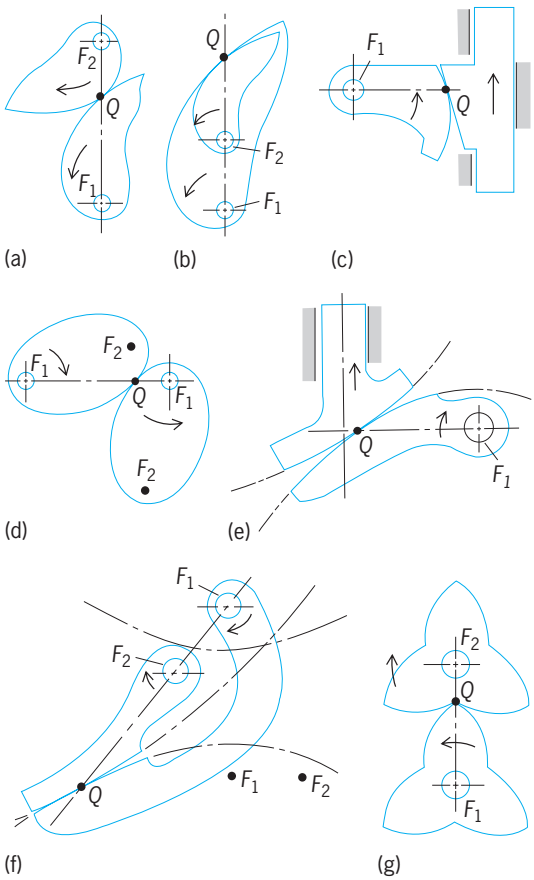

Fig. 3. Pure rolling between bodies in contact. (*a*) Externally rolling logarithmic spirals. (*b*) Internally rolling logarithmic spirals. (*c*) Logarithmic spiral and straight-sided follower. (*d*) Ellipses. (*e*) Equal parabolas. (*f*) Equal hyperbolas. (*g*) Lobe wheels contoured either from sectors of ellipses or logarithmic spirals. (*After H. A. Rothbart, Cams, John Wiley and Sons, 1956*)

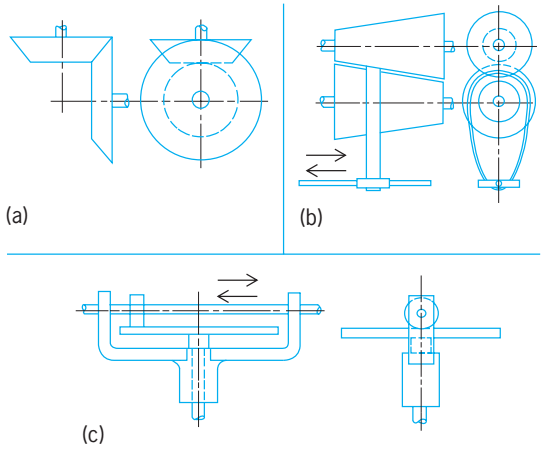

Fig. 4. Friction gears. (*a*) Cones. (*b*) Variable speed. (*c*) Mechanical integrator. (*International Textbook Co.*)

it possible to have continuous action. The lobes are made up of sectors of logarithmic spirals or ellipses. Friction gearing consists mainly of rolling cylinders, cones, and disks. The limited areas of surface contact permit only small quantities of power to be transmitted. Slippage can be often reduced by the use of spring-loaded bearings. Friction devices include friction cones (**Fig. 4***a*); Evan's friction cones, which permit variations in output speed by shifting the belt position (Fig. 4*a*); and the Bush wheel and plate arrangement (Fig. 4*c*) used as an integrator for the original Vannevar Bush analog computer. The Bush arrangement permits not only variations in output speed but also reversal of the direction of the output shaft. *See* ANALOG COMPUTER; INTEGRATION.

John R. Zimmerman

Bibliography. F. T. Barwell, *Bearing Systems: Principles and Practice*, 1979; B. Bolt, *Mathematics Meets Technology*, 1990; C.-H. Chiang, *Kinematics of Spherical Mechanisms*, 1988; V. M. Faires and R. M. Keown, *Mechanism*, 5th ed., 1960, reprint 1980; H. B. Kepler, *Basic Graphical Kinematics*, 2d ed., 1973; Society of Automotive Engineers, *Bearing Technology: Analysis, Development and Testing*, 1985; C. H. Suh and C. W. Radcliffe, *Kinematics and Mechanisms Design*, 1978, reprint 1983.

Roof construction

An assemblage to provide cover for homes, buildings, and commercial, industrial, and recreational areas. Roofs are constructed in different forms and shapes with various materials. A properly designed and constructed roof protects the structure beneath it from exterior weather conditions, provides structural support for superimposed loads, provides diaphragm strength to maintain the shape of the structure below, suppresses fire spread, and meets desired esthetic criteria. Roofs have been constructed in many shapes by using structural assemblies and materials including beams, arches, cut stone slabs, cast-iron segments, masonry or terra-cotta tiles, and thatched grasses.

Modern roof construction, evolving from older technologies, usually consists of an outer roofing assembly that is attached atop a deck or sheathing surface, which in turn is supported by a primary framework such as a series of beams, trusses, or arches. A hung ceiling may be placed below the roof structure to give a closed or finished look and provide a space for modern utilities such as heating ducts and electrical and communication lines. The shape of the roof and type of roof construction are usually determined by, and consistent with, the materials and deck of the primary structure underneath. *See* ARCH; BEAM; TRUSS.

Roof shapes include flat; hipped, where two sloping deck surfaces intersect in a line, the ridge or hip; pyramidal, which involves three of more sloping planes; domed, or other three-dimensional-surface, such as spherical, parabolic, or hyperbolic, shells; and tentlike, which are suspended fabric or membrane surfaces. Some famous old surviving roofs include the steep buttressed Gothic church roofs such as Notre Dame Cathedral in Paris; large span domes such as St. Peter's Church in Rome; the dome over the U.S Capitol in Washington, D.C.; and the mosque-like dome such as that over the Taj Mahal in Agra, India.

Roof assembly. A roof assembly (**Fig. 1**) is a series of layers of different materials placed on and attached to the roof deck. Each type of roof assembly—related to protection against water entry from rain, snow, or ice; and insulation for temperature change, fire propagation, wind uplift, and moisture migration—has its own design requirements and methods of construction and attachment.

Materials. Roofing materials vary, and are selected according to design requirements to keep water out, to provide adequate strength against wind uplift, to sustain light impact loads, and to reduce wear if the roof is to sustain human activity, and according to availability of local materials.

Shingles. Usually used for hipped or pyramidal roofs, these are small water shedding pieces made of wood, slate, asphaltic felts, and other materials. The shingles are overlapped downslope to shed water, and they are usually nailed or stapled onto the roof deck made of membrane-coated wood. A variation of the shingle used in warmer climates is either flat or half-round ceramic or concrete, known as Spanish tiles.

Built-up roofing. Built-up roofing is a traditional type of roof cover; it is usually used on flat or low-slope roofs. It is constructed of several layers of asphalt or tar-saturated fabric sheets, reinforced with plastic or fiberglass, with each sheet being called a ply. Each ply is bonded to the surface below by hot or cold applications of asphalt-based adhesives in order to form a multi-ply built-up roof. The top ply may have granules of sand or stone embedded into it, making it a mineral-surfaced ply that can provide some resistance to wear and tear from human activity. Flat roofing may be ballasted by being covered with loose stone or gravel of $^1/_2$–3 in. (15–75 mm) diameter to depths of 4 in. (100 mm) to resist wind uplift forces.

Single-ply roofing. Single-ply roofing utilizes large factory-fabricated sheets of flexible plastic or rubberlike materials, such as neoprene rubber, chlorosulfonated polyethylene elastomer, or poly(vinyl chloride), about 0.03–0.50 in. (0.76–12.7 mm) thick.

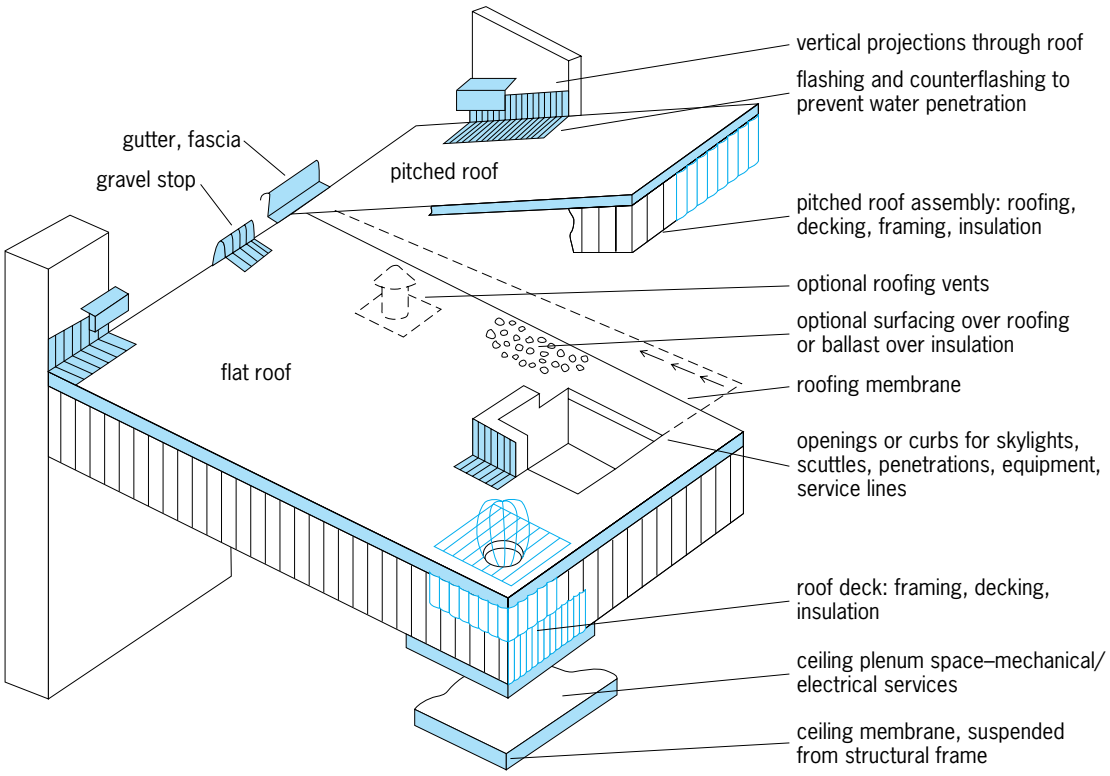

Fig. 1. Common components of roof assembly. (*After Sweets Catalog, The McGraw-Hill Companies, 1993*)

If one factory-produced sheet does not cover the whole roof, additional single-ply sheets are laid down, with adjacent edges or ends lapped, glued, and sealed to form watertight seams between adjoining sheets. Some single-ply systems are loose-laid and held down by gravel ballast or a waterproof batten system, where thin structural planks are spaced atop the single ply and fastened to the roof structural deck below, or are adhered to the insulation and roof deck below.

Other occasionally used roofing materials include corrugated, flat, or standing seam sheets of copper, terne plate metal, painted or galvanized steel, or aluminum. In a standing seam configuration, the two metal pieces are joined above the roof surface (**Fig. 2**).

Vapor barriers. Air, which contains moisture, generally moves from a warm area to a cooler one, and the moisture within condenses. Condensation can cause water corrosion or other damage to a roof structure, and reduce the effectiveness of insulation. Vapor barriers are a ply or membrane of plastic film, aluminum foil, foil-laminated paper, or other prepared materials usually placed on the warm side of insulation to prevent or reduce the passage of moisture, and subsequent condensation.

Venting sheets. Vapor barriers are not totally effective. In areas where large amounts of water vapor may be encountered, a special type of roofing ply is installed above the insulation to wick or expedite the drying out of any moisture that reaches the roofing assembly.

Insulation. Usually composed of materials such as fiberglass, urethane, or styrofoam, this is placed on or under modern roof decks in order to reduce energy losses for both winter and summer. Some insulation is fabricated with a vapor barrier on one side for rapid application.

Edges. The outer edges of roofing are embedded, attached, and sealed to the edge structure of the roof

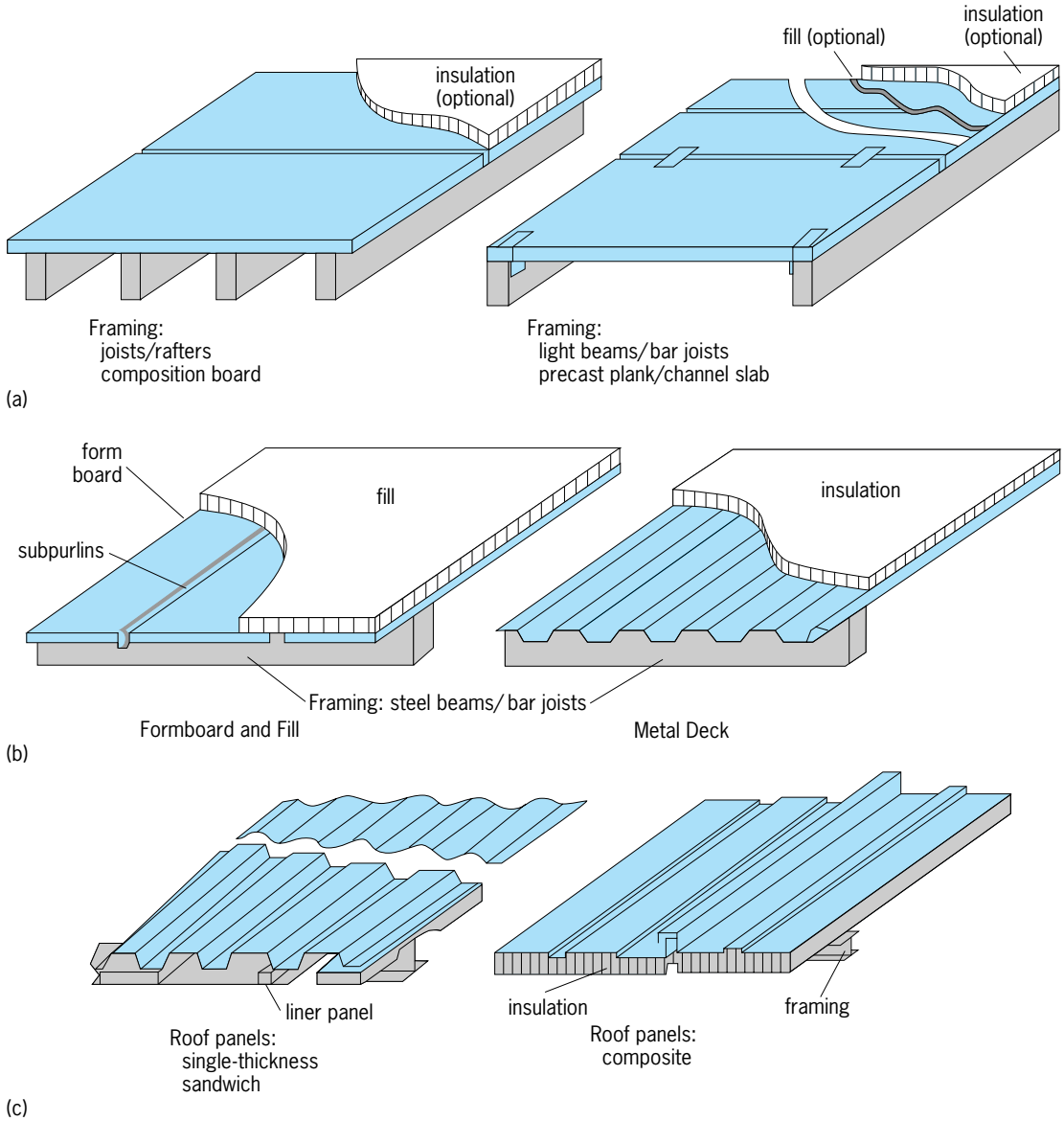

Fig. 2. Types of roof deck structure. (*a*) Site-assembled roof. (*b*) Flat roof. (*c*) Combined decking/roofing for a pitched roof. (*After Sweets Catalog, The McGraw-Hill Companies, 1993*)

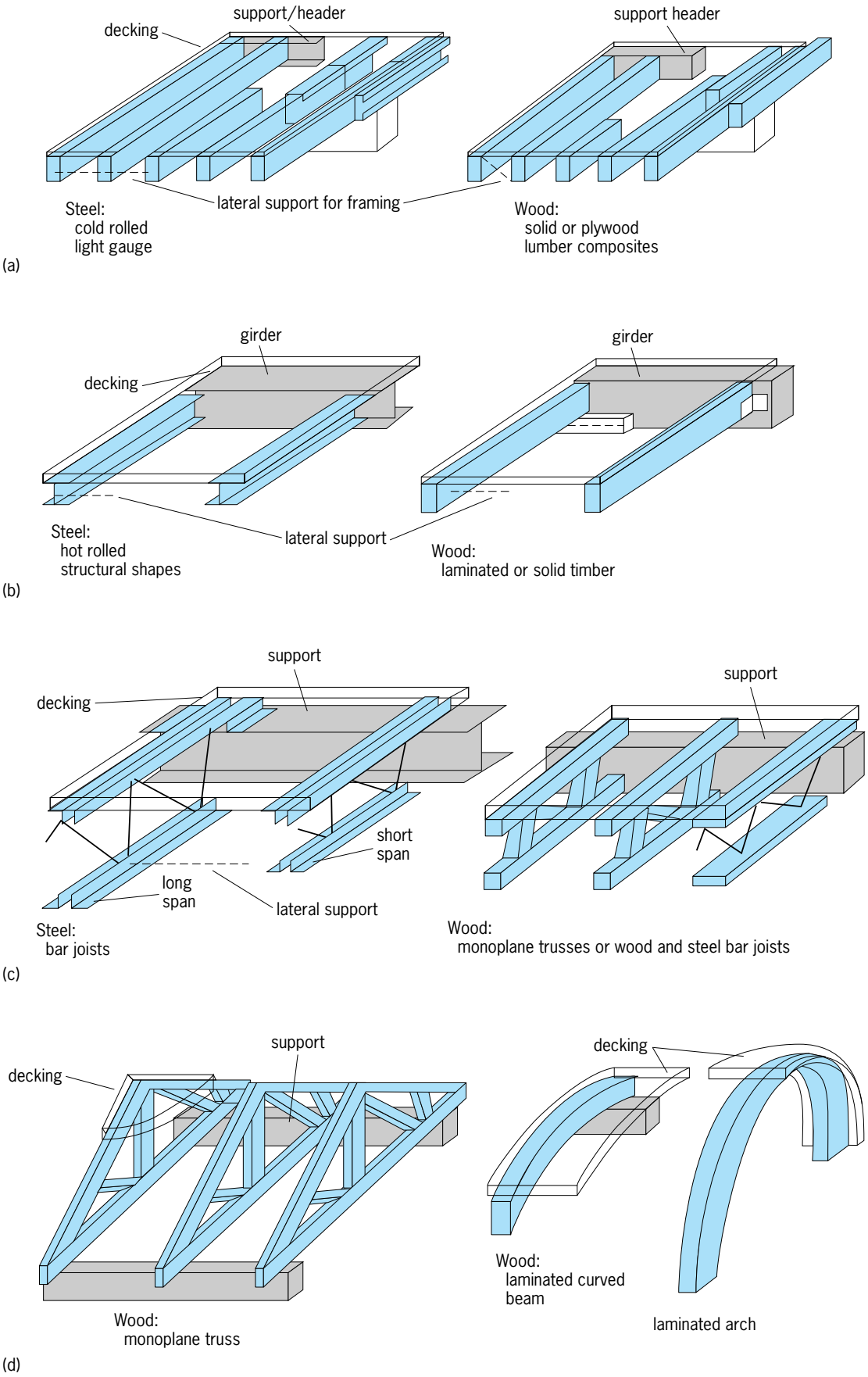

Fig. 3. Types of roof primary structural framing. (*a*) Joists/rafters, close spaced. (*b*) Beams. (*c*) Trusses, flat top chord. (*d*) Trusses/beams, pitched/curved. (*After Sweets Catalog, The McGraw-Hill Companies, 1993*)

as well as where there is a penetration through the roof. Eave constructions using gravel stops or metal fascia plates are designed to physically protect the roofing edge and prevent the entry of rainwater or snow melt under the roofing or into the top edge of the building wall construction. Where there are penetrations through the roof, cant strips, flashings, curbs, and other means are used to control water and see page (Fig. 1).

Structural components. A roof, which is the closure, is built upward from the structure below. The framework, or primary structural components, rest on the walls and columns of the structure, and these support the roof deck or the sheathing, which in turn carries the roofing assembly. The walls and columns may have girders framing into them. Beams rest on or are connected to girders. For flat roofs, these beams are called joists or purlins, while for hipped roofs they are termed rafters. Beams that are larger and made up of rigid geometric shapes, usually triangles, are termed trusses. Trusses can be made of wood or steel. Special types of trusses, formed from prefabricated light steel members, are called open web joists. The primary components may also be arches made of reinforced concrete, steel, wood, or laminated and glued wood sections. Under certain conditions a cast-in-place reinforced concrete deck, made integral with reinforced-concrete beams or arch sections, may be employed. *See* REINFORCED CONCRETE.

The roof deck or sheathing, the components that provide the basic support for the roofing assembly, span between and are anchored to the primary structural framing (**Fig. 3**). The various types of decking include wood planking, often tongue and groove, or plywood or special composition board; corrugated decking of steel or aluminum, sometimes covered with and structurally integral with the reinforced-concrete slab; precast planks of reinforced concrete or gypsum; cast-in-place reinforced concrete or gypsum, often placed on top of a sheet-rock-type ceiling material to provide a finished underside; and flat plates, sometimes with drop panels, which are reinforced-concrete slabs that are supported directly on columns without the use of separate beams.

Special roof designs. Where a long, clear space is required within a building, where a unique shape is desired for a structure, or where other constraints are demanding, special engineering designs are necessary for both the roof primary framing and the roof assembly. In many of these structures the deck and primary framework are effectively combined. Long-span roofs use space trusses, usually of steel; but reinforced-concrete domes or other shell shapes, including folded plates, may be employed. The TWA Terminal Building at Kennedy Airport in New York City, and the U.S. Air Force Academy Dining Hall in Colorado are examples of long-span roofs.

Cable supports. In cable-supported roofs, the primary framework is composed of cables in tension that are slung between separate posts or from the top of the surrounding building perimeter. The cables are restrained both vertically and horizontally by the building walls or other structural framing. The deck, which is supported by the tensioned cables, may be of any of the previously described assemblies or may be of a specially fabricated waterproof fabric. Tentlike or membrane roofs are a special application of a cable-supported roof. Here the membrane or the tentlike fabric is placed in tension directly, and it serves as both the support and the covering.

Air-supported roofs. Air-supported roofs utilize a waterproof coated fabric that is inflated to its rigid shape by developing and maintaining a positive air pressure inside the structure, which keeps the roof surface under tension. Tennis-court "bubbles" utilize this design to enable the fabric surfaces to resist the bending moments caused by external roof loadings.

Design loadings. These design parameters have evolved from much engineering study, field measurements, testing, experimentation, and construction practices; and they are formalized into codes. The basic design loadings include dead loads, live loads, wind loads, and seismic loads. The dead loads are essentially the constant loads due to the weights of the roofing assembly, the deck, the primary framework, and the hung ceiling loads, including any piping and ductwork in the space formed by the hung ceiling. The live loads are variable gravity loading, which include human occupancy or utilization of the roof surface for maintenance and repair; snow loads, which vary with the geographic location and the shape and slope of the roof; rainfall or snow melt water loads, which are the loads the roof must sustain if it is not properly drained; and special or concentrated loads due to roof-mounted equipment. Wind loads are generally complex and will vary with location, with variations due to roof height, slopes, and shape-caused turbulence. Wind loads may be applied both inward and outward on the roof surface, and they may oscillate. Seismic loading, which adds loading factors because of rapid earth movement during earthquakes, is a very specialized consideration, since it produces increases in both vertical and horizontal forces on the roof, requiring special design and construction considerations.

Roofs must be constructed under proper conditions, for if the deck is not dry, clean, or properly prepared, and the roof assembly is not properly put in place and anchored, with attention to flashing and eave details, the cover over the building will not last. *See* BUILDINGS; LOADS, DYNAMIC; SEISMIC RISK.

Milton Alpern

Bibliography. D. K. Ballast, *Architect's Handbook of Construction Detailing*, 1990; H. J. Cowan, *Design of Reinforced Concrete Structures*, 2d ed., 1988; K. F. Faherty, *Wood Engineering and Construction Handbook*, 3d ed., 1998; C. W. Griffin and R. Fricklas, *The Manual of Built-Up Roof Systems*, 1996; S. Hardy, *Time Saver Details for Roof Design*, 1997; F. S. Merritt, M. K. Loftin, and J. T. Ricketts, *Standard Handbook for Civil Engineers*, 4th ed., 1995; National Fire Protection Association Staff, *Roof Coverings and Roof Deck Constructions*, 1992.

Root (botany)

The absorbing and anchoring organ of vascular plants. Roots are simple axial organs that produce lateral roots, and sometimes buds, but bear neither leaves nor flowers. Elongation occurs in the root tip, which is usually 0.004–0.04 in. (0.1–1.0 mm) in diameter and 0.4–4 in. (1–10 cm) long. The older portion of the root, behind the root tip, may thicken through cambial activity. Some roots, grass for example, scarcely thicken, but tree roots can become 4 in. (10 cm) or more in diameter near the stem. Roots may be very long. The longest maple (*Acer*) roots are usually as long as the tree is tall, but the majority of roots are only a few centimeters long. The longest roots may live for many years, while small roots may live for only a few weeks or months.

Function. Root tips and the root hairs on their surface take up water and minerals from the soil. They also synthesize amino acids and growth regulators (gibberellins and cytokinins). These materials move up through the woody, basal portion of the root to the stem. The thickened, basal portion of the root anchors the plant in the soil. Thickened roots, such as carrots, can store food that is later used in stem growth. *See* CYTOKININS; GIBBERELLIN.

Growth environments. Roots usually grow in soil where it is not too dense to stop root tip elongation; there is enough water and oxygen for root growth; and temperatures are high enough (above 39°F or 4°C) to permit root growth, but not so high that the roots are killed (above 104°F or 40°C). In temperate zones, most roots are in the uppermost 4 in. (10 cm) of the soil; root numbers decrease so rapidly with increasing depth that few roots are found more than 6.5 ft (2 m) below the surface. Roots grow deeper in areas where the soil is hot and dry; roots from desert shrubs have been found in mines more than 230 ft (70 m) below the surface. In swamps with high water tables the lack of oxygen restricts roots to the uppermost soil layers.

Roots may also grow in the air. Poison ivy (*Toxicodendron*) vines form many small aerial roots that anchor them to bark or other surfaces. Fig (*Ficus*) and mangrove (*Rhizophora*) trees have aerial roots which may grow several meters through the air before reaching the soil. Aerial roots often have some characteristic that prevents desiccation; orchid roots have a special surface layer, the velamen, and screw pine (*Pandanus*) roots are up to 0. 4 in. (10 mm) in diameter with a thick cortex protecting the central vascular tissues.

Origin. The primary root originates in the seed as part of the embryo, normally being the first organ to grow. It grows downward into the soil and produces lateral second-order roots that emerge at right angles behind the root tip. Sometimes it persists and thickens to form a taproot. The second-order laterals produce third-order laterals and so on until there are millions of roots in a mature tree root system. In contrast to the primary root, most lateral roots grow horizontally or even upward. In many plants a few horizontal lateral roots thicken more than the primary, so no taproot is present in the mature root system.

Roots can also form laterals just behind the cut surface when they are severed. Unlike regular lateral roots, these roots bend to grow in the same direction as the parent, thus replacing the parent root. This type of root formation occurs after root pruning of woody root systems.

Adventitious roots originate from stems or leaves rather than the embryo or other roots. They may form at the base of cut stems, as seen in the horticultural practice of rooting cuttings. Adventitious roots may also form on branches which have been covered by fallen leaves or soil, a layering process common in shrubs like mountain laurel (*Kalmia*), and some trees like black spruce (*Picea*) where branches grow near the ground.

Bud formation. Roots of many plants can also form buds. When a root is cut into segments, buds can form at the end toward the stem and roots at the end toward the tip so that each segment develops into a new plant. Aspen (*Populus*), locust (*Robinia*), beech (*Fagus*), and other species produce root suckers, stems that grow out from buds on the roots when the parent stem is cut or injured. These new stems form large clones of genetically identical plants. *See* BUD.

Special forms. Woody roots can graft together. In grafts between individuals of the same species the roots grow next to each other, so that as they increase in diameter the two roots are pressed against each other and eventually grow together. Grafting may also occur between roots of the same individual, as in the case of the strangler fig (*Ficus*), which starts as a small plant growing in another tree. The roots grow down along the stem of the host tree, and after reaching the soil they thicken and graft together to form a woody cylinder that seems to be the stem of the fig tree when the host tree dies.

Some tree species that grow in flooded areas, for example, mangrove (*Avicennia*), form roots called pneumatophores that grow vertically upward from horizontal roots to reach above the water surface. Pneumatophores contain loosely packed cells that permit air movement which may help to aerate the underwater roots.

Fungal and bacterial relations. Hyphae of some fungi can invade root tips and form mycorrhizae. Almost all plants have mycorrhizae which help them grow on mineral-deficient soils. Apparently the hyphae provide minerals to the plant, and the plant provides carbohydrates to the fungus. Many form characteristic structures such as many-forked short roots in pines (*Pinus*) and a series of beads in maples (*Acer*). *See* FUNGI; MYCORRHIZAE.

Some plants can form root nodules that fix inert organic nitrogen into usable organic compounds. Leguminous plants, such as peas (*Pisum*) and black locust trees (*Robinia*), form nodules that contain *Rhizobium* bacteria. Nonleguminous nitrogen-fixing plants occur in many different plant groups, such as

gale (*Myrica*) and alder (*Alnus*), which form nodules with actinomycete fungi. *See* NITROGEN FIXATION.

B. F. Wilson

Root systems. The extent of the system or root mass varies in relation to inherited growth characteristics and factors such as soil porosity, aeration, and the availability of water (**Fig. 1**). In general, root systems are extensive, spreading widely or penetrating deeply, the numerous finer branches and their root hairs being in contact with a large volume of soil. Their physical attributes are evaluated in different ways (**Tables 1** and **2**). For example, as much as 30 ft (9 m) of root with an external surface up to 65 in.^{2} (420 cm^2) have been found in 1 in.^{3} (16 cm^3) of soil from under grass. In many species the root system constitutes the bulk of the plant body. Although most of the roots of many trees may be found in the top 4–5 ft (1–1.5 m) of the Earth's crust, their lateral spread may extend through the soil considerably beyond the tips of the longest branches of the crown. The form of root systems differs with the origin and manner of growth of their members. Nearly all are variations of taproots and fibrous roots.

Taproots. In taproot systems the primary root forms a dominant central axis. The main root penetrates rather deeply into the soil and is generally larger than its branches. Plants with taproots range from trees, in which the primary root and its main branches are thick and woody, to herbs, in which the taproot may be rather slender or develop into a fleshy food-storage organ as in a carrot plant.

Fibrous roots. Several to many main roots of equal dominance typify fibrous root systems. Most commonly these main roots arise adventitiously from the stem as in grasses, but sometimes fibrous roots are composed of branches of a primary root that ceased to be dominant. In some species, root systems organized like fibrous roots are composed of thick and fleshy units, for example, those of dahlia.

Root cuttings. Roots of many plants normally form buds from which shoots develop, or they may be induced to form buds by injury or severe pruning. Cuttings, or short lengths, of roots may be planted to propagate such plants. *See* REPRODUCTION (PLANT).

Primary tissues. Roots are characterized by a pattern of apical growth similar to that of stems. Through the persistent activities of meristematic (rapidly dividing) cells at the root apex, new cells are added continually to the root. A terminal root cap covers this meristematic region (**Fig. 2**). *See* STEM.

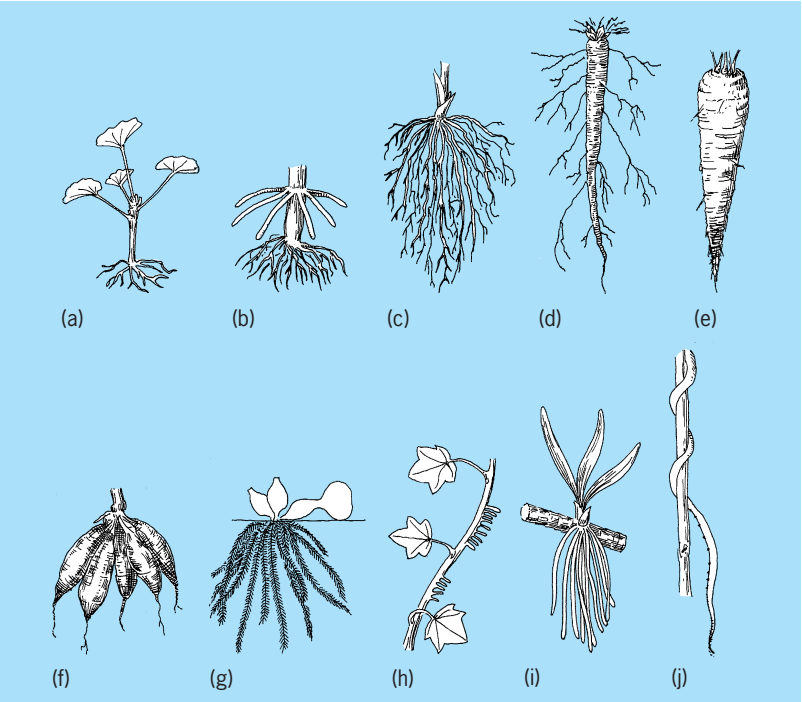

Fig. 1. Kinds of roots. (*a*) Adventitious, on a cutting or slip (geranium). (*b*) Prop and other adventitious (corn). (*c*) Fibrous adventitious (grass). (*d*) Primary, or taproot (dandelion). (*e*) Fleshy taproot (carrot). (*f*) Fleshy, fascicled adventitious (dahlia). (*g*) Aquatic adventitious (water hyacinth). (*h*) Aerial adventitious (English ivy). (*i*) Aerial adventitious (orchid). (*j*) Parasitic, or haustorial, adventitious (dodder).

Apical meristem. In a rapidly growing root, two distinct meristematic regions are evident. They are separated from each other by a group of nonmeristematic cells known collectively as the quiescent center (**Fig. 3**). On the basal or proximal side of the quiescent center is located the proximal meristem, which produces most of the new cells in a root. On the opposite or distal side of the quiescent center is the distal meristem. Both meristems may fluctuate in size, and often can be shown to have specific developmental relationships to the primary tissues or to the root cap. Depending upon the plant species, approximately 500–2000 cells can compose the quiescent center. Three-dimensionally these cells often assume the shape of an inverted cup, along the periphery of which are located cells of the proximal and distal meristems. The quiescent center is visualized most clearly by the techniques of autoradiography, which demonstrate that cells of the quiescent center divide rarely. A quiescent center is absent in embryonic roots, but is evident very soon after seedling

TABLE 1. Dimensions of root systems

Plant	Lateral spread, ft (m)	Depth, ft (m)
Little bluestem (*Andropogon scoparius*)	1.0 (0.3)	4.0 (1.2)
Blazing star (*Liatris punctata*)	9.0 (3.0)	16.0 (5.6)
Comanche cactus (*Opuntia camanchica*)	9.0 (3.0)	2.8 (0.8)
Wheat (*Triticum aestivum*)	2.0 (0.6)	5.0 (1.5)
Corn (*Zea mays*)	8.0 (2.5)	7.0 (2.4)
Sugarbeet (*Beta vulgaris*)	3.0 (0.9)	6.0 (1.8)

TABLE 2. Quantitative characteristics of root systems

Characteristic	Crested wheat grass	Winter rye	Coffee
Age	2 years	4 months	3 years
Soil mass	56 ft^3 (1.59 m^3)	2 ft^3 (0.056 m^3)	270 ft^3 (7.65 m^3)
Total root length	315 mi (507 km)	387 mi (623 km)	14.6 mi (23.5 km)
Number of roots		13,000,000	
Total surface		2554 ft^2 (237 m^2)	

germination, and may fluctuate in size throughout the life of an individual root. Developmentally the function of the quiescent center is not known, though it has been implicated in the production of plant growth substances in the root.

In roots of ferns a quiescent center is absent, and new cells are thought to arise from the meristematic activities of a single enlarged apical cell and its immediate derivatives. *See* APICAL MERISTEM.

Root cap. The root cap is composed of cells at various stages of differentiation and, because of its terminal position, is considered to be a protective structure. It serves as the site of gravity perception in the root, and is believed to facilitate the movement of the root through the soil by the production of mucilaginous substances.

Primary meristems. Cell divisions are most frequent in the primary tissue regions of the root, which usually become distinct immediately behind the apical meristem. In recognition of their meristematic nature at this root level, these regions are often designated as primary meristems. These meristems are the protoderm, ground meristem, and procambium. Protoderm represents the surface layer which develops into the epidermis; procambium designates the central region which forms the vascular cylinder (stele); and ground meristem constitutes the remaining tissue which gives rise to the root cortex.

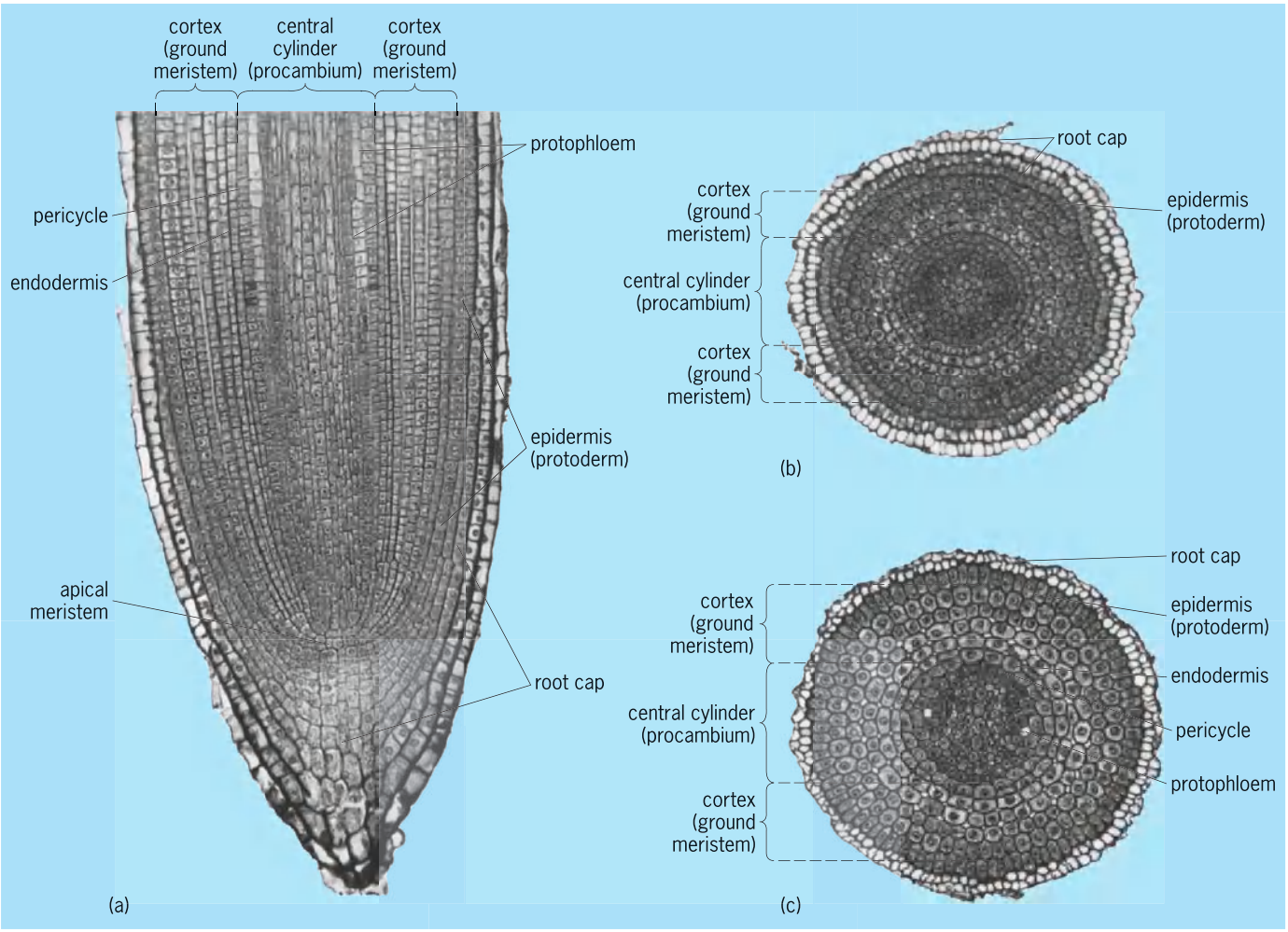

Fig. 2. Root-tip organization of a tomato plant. (*a*) Median longitudinal section, showing meristematic region, which is covered by a terminal root cap. (*b*, *c*) Transverse sections of same root at 160 and 480 μm behind root apex (apical meristem), respectively.

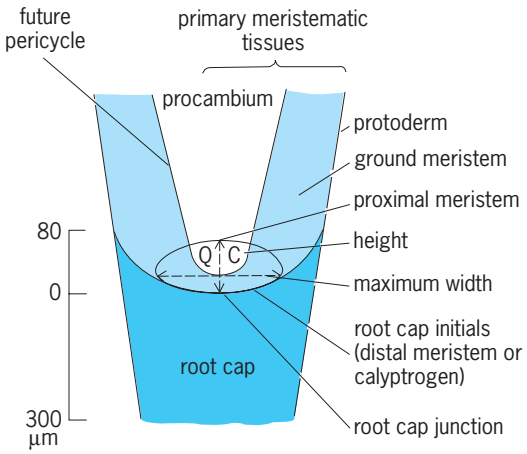

Fig. 3. Distribution of tissues in a typical root apex.

Epidermis. The root epidermis usually consists of a single layer of elongated, thin-walled cells in which the cuticle is difficult to demonstrate with certainty (**Fig. 4***a* and *b*). In many older roots the epidermis is sloughed off. If the layer persists in mature roots, more or less conspicuous cutinization or other wall modification may develop. The production of root hairs is characteristic of the root epidermis. *See* EPIDERMIS (PLANT).

Root hairs. These structures develop as narrow, tubular outgrowths of epidermal cells, their walls and contents being continuous with those of the main body of the cell. They increase the absorbing surface of the root by contact with a much greater surface of soil than would be touched by the root surface without hairs. All epidermal cells may be capable of producing root hairs, or in some species they may be formed only from special short cells, called trichoblasts. Root hairs are usually confined to a short length of root behind the zone of elongation. Those toward the root tip are youngest, and toward the base they are progressively older and longer. As a root grows in length, new root hairs are initiated toward the tip while the oldest ones usually degenerate. The newly developing root hairs make contact with additional soil particles and maintain an extensive absorbing surface.

Cortex. Thin-walled cells among which intercellular spaces develop to varying degrees are common features of the root cortex. Cells may be arranged more or less irregularly, or they may show radial or concentric patterns, separate or in combination. Cell arrangement may reflect some aspects of cortical growth; for example, radial seriation arises as the result of centripetal growth involving successive periclinal (parallel with the circumference) divisions of the innermost cells. If these divisions are synchronized, rings of cells or a concentric pattern may also be apparent. Small intercellular spaces, schizogenous (splitting of adjacent cell walls) in origin, develop near the root apex, particularly in the inner cortex. Larger lysigenous (dissolution of cells) cavities may also develop in the older root. One or more layers of the outermost cortical cells may develop suberized or lignified walls which serve protective functions. Those of the innermost layer usually become specialized as an endodermis. *See* CORTEX (PLANT).

Endodermis. Cells of the endodermis exhibit wall modifications which involve thickening and deposition of suberin or lignin. After the centripetal growth of the cortex is completed, suberin and lignin are deposited in the form of a band, the Casparian strip, on and within the transverse and radial walls of the endodermal cells. In many roots the endodermal cells undergo no further changes; in others an extensive thickening of the radial and inner tangential walls follow. *See* ENDODERMIS.

Pericycle. The outer portion of the central cylinder, or stele, is designated as the pericycle, which may be one or more cell layers in thickness. Although generally continuous, it may be interrupted by xylem elements as in roots of some grasses (Fig. 4*c*). Cell divisions in the pericycle initiate the growth of lateral roots and contribute to the formation of the vascular cambium and periderm in roots with secondary growth. *See* PERICYCLE.

Vascular system. The vascular system, consisting of primary xylem (water-conducting) and phloem (food-conducting) tissues, constitutes the bulk of the central cylinder in many roots. Phloem occurs in strands near the periphery the periphery; xylem alternates with phloem, either as strands or as ridges of a central xylem mass. This distribution, that is, the arrangement of xylem and phloem on separate radii, is characteristic of roots. Whereas a central pith occurs in the roots of some dicotyledons, it is common among monocotyledons, although in many of the latter the central core consists of lignified parenchyma, including one or more strands of water-conducting cells.

Maturation of the primary vascular tissues is a complex process and occurs at various levels in the root. The outer xylem in the ridges or strands is the first xylem to mature, and is designated as protoxylem; the inner xylem matures after the protoxylem, and is called the metaxylem. Thus the maturation of the primary xylem occurs in a centripetal direction, and the xylem having such a pattern of maturation is termed exarch. Protophloem is likewise external to metaphloem. Phloem usually matures closer to the root tip than the xylem. As seen in cross sections, the points of origin of the protophloem and protoxylem may be called phloem and xylem poles, respectively. Roots are described as monarch, diarch, and triarch, on the basis of whether there are one, two, or three poles; those with many poles are designated as polyarch. The number of poles may vary among roots of different plants, of different roots of the same plant, or at different levels of an individual root. This variation in pole number is now believed correlated with changes in the dimensions of the quiescent center. A large quiescent center is suggested to be associated with complex vascular tissue patterns. Dicotyledon roots generally have fewer poles than those of monocotyledons. *See* PHLOEM; XYLEM.

Lateral roots. Root branches or lateral roots are generally produced at some distance behind the apex,

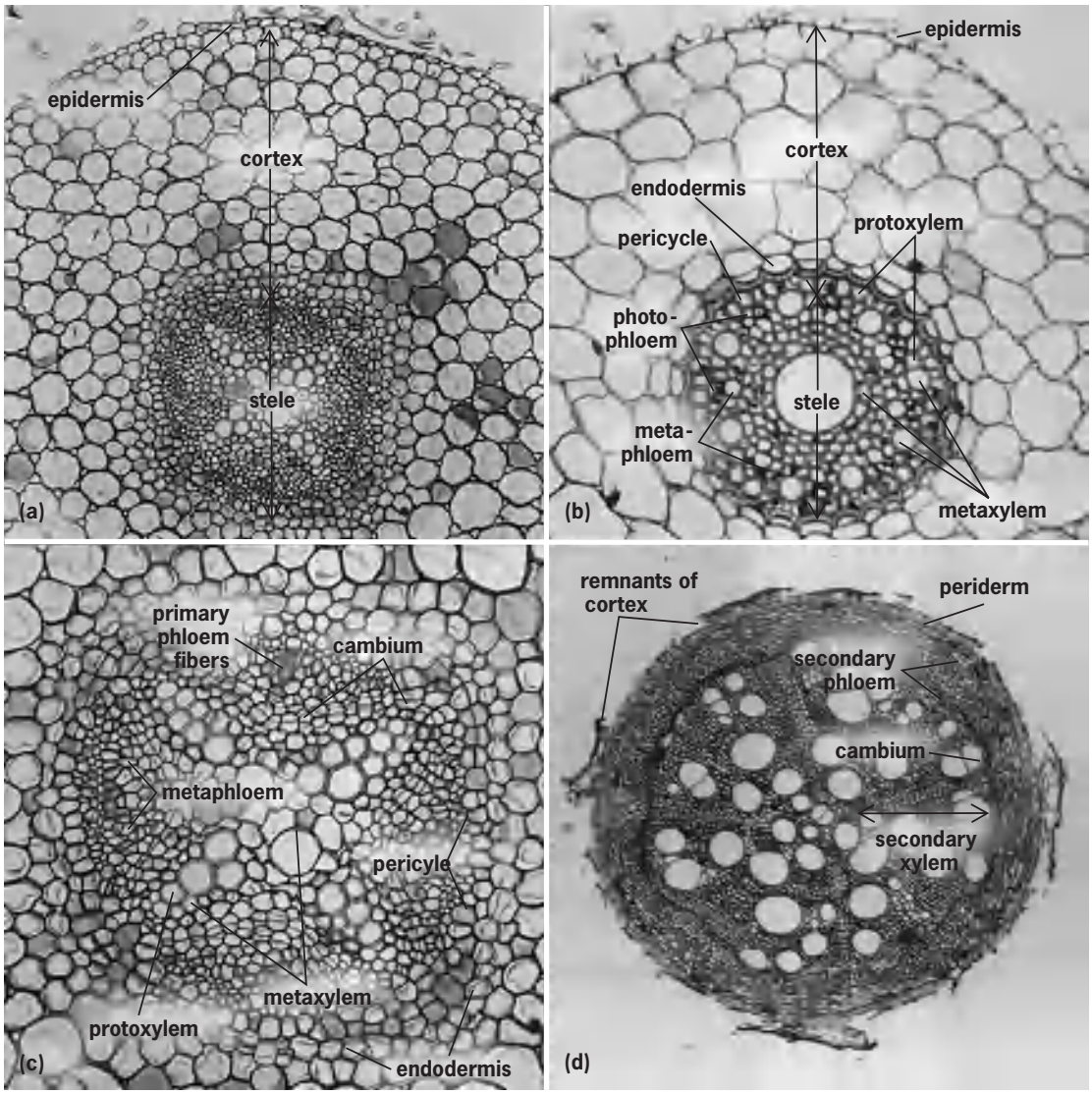

Fig. 4. Transverse sections of roots, showing primary and secondary tissues. (*a*) Soybean. (*b*) Barley. (*c*) Soybean; enlarged part of *a*, showing early formation of cambium. (*d*) Soybean with abundant secondary tissue.

and they originate endogenously, or deep within the tissues. Depending on the number of poles in the parent root, lateral roots arise in distinctive positions in relation to the xylem and phloem. A lateral root is commonly initiated by localized cell divisions in the pericycle; however, in lower vascular plants the initial divisions may involve endodermal cells. Continued divisions produce a primordium in which the root cap, apical meristem, and primary tissue regions are soon organized. The young root grows through the cortex, and by the time it emerges, vascular connections with the main axis are established.

Elongation. Some elongation characteristic of primary growth of the root results from the increase in number of cells in the meristematic tip, but most of it is produced by the elongation of cells behind the meristematic region. The location and extent of the zone of elongation varies, in part with root size and stage of growth. For example, in relatively thin, rapidly growing roots of timothy (*Phleum*) the zone of cell elongation occurs 450–1400 micrometers behind the root apex, whereas in larger primary roots of corn (*Zea*) the zone lies 2000–9000 μm behind the root tip.

Differentiation. The primary tissues of the root begin to differentiate during the elongation but complete their differentiation, or become mature, after the elongation ceases. Differentiation consists in a series of changes through which the cells become more or less specialized with regard to the functions they carry out in the mature state. The differentiation follows characteristic patterns in both longitudinal and transverse directions. Thus, for example, the protophloem matures more closely to the apex than the protoxylem, and both differentiate in the direction from the base of the root toward the apex, that is, acropetally. The metaphloem and metaxylem follow in the same direction. Because the phenomena of growth, that is, cell division and cell enlargement, overlap with the earlier phenomena of differentiation, the sequence of maturation of a root is highly complex.

Secondary tissues. The formation of secondary xylem and secondary phloem by the vascular cambium is the chief characteristic of secondary organization in roots. Other structural modifications invariably accompany cambial activity and thus are also features of the secondary state. *See* LATERAL MERISTEM.

The extent of secondary tissue formation in roots generally parallels that in stems. Trees and shrubs with woody stems are known to develop woody roots. Herbaceous plants with a smaller volume of secondary tissues in the stem also show less cambial activity in the roots. However, roots of plants that have little or no secondary vascular tissues in the stems have not been investigated sufficiently to determine the presence or absence of a cambium.

Development of vascular cambium. Formation of the vascular cambium in the dicotyledons begins at root levels where primary development is approaching completion. When viewed in transverse section, the cambium first appears as short disconnected arcs of periclinally dividing cells internal to the phloem strands. As vascular cells are formed by the periclinal divisions in the arcs of cambium, pericycle cells opposite the xylem poles also divide periclinally. Inner cells formed by divisions in the pericycle differentiate as cambium and unite the arcs internal to the phloem. A circular or cylindrical distribution of the cambium results from the accumulation of secondary xylem internal to the phloem strands, whereby these strands and adjacent cambium are displaced outwardly. Subsequently, complete cylinders of secondary xylem and secondary phloem are deposited inside and outside the cambium, respectively (Fig. 4*d*). As cambial activity progresses, the outer tissues are subjected to tension and pressure. Primary phloem is crushed and individual cells degenerate, except for those which in some plants differentiate into primary phloem fibers. The cortex may be split and completely shed, as described in the subsection on periderm, which follows. The original cambium adjusts in circumference to diameter increase and normally continues to function throughout the life of the root.

Although monocotyledons generally lack cambial activity, in some, secondary growth results from the activity of a cambiumlike thickening meristem that originates in the pericycle or cortex and produces a secondary body of vascular bundles embedded in parenchyma tissue.

Secondary vascular tissues. The cambium and secondary tissues derived from it are similar in general organization to those of the stem, but certain quantitative differences between the root and stem tissues and their constituent cells are present. One of the most common characteristics of roots is a relatively large volume of parenchyma, a feature considered to be a result of high degree of specialization for storage. The abundant parenchyma results from normal cambial activity or from anomalous growth involving irregular cell proliferation in the secondary xylem or formation of additional layers of cambium outside the normal.

Periderm. Periderm, a secondary protective tissue, is characteristic of woody roots, but it may not develop in those with only a small degree of cambial activity. This tissue is derived from a lateral meristem called phellogen, or cork cambium, which most commonly has its origin in the pericycle. Its characteristic component is the cork tissue. *See* BARK; PERIDERM.

After the vascular cambium becomes complete outside the xylem poles, periclinal divisions in the pericycle continue around the vascular cylinder, and the outer cells thus formed function as a phellogen. It begins to form periderm as the vascular core enlarges and stretches the cortex, which is ultimately split and shed. If the original periderm is ruptured by growth in diameter, new layers are formed from new phellogen which differentiates in parenchyma of the secondary phloem.

Periderm formation may also follow a different course. In some dicotyledons it first arises from subepidermal cells, but subsequently another is formed from a proliferated pericycle. In others, it originates only in the outer cortex. Many roots with little or no secondary thickening lack a periderm, the outer protective tissue being the persisting cutinized epidermis or an exodermis of subepidermal origin. The exodermis comprises layers of cells with suberized or lignified wall thickenings.

Root and shoot connection. The more or less cylindrical vascular system of the root is morphologically distinct from that of the shoot, which is composed of interconnected vascular bundles. In the hypocotyl (seedling axis below the cotyledons) the vascular tissues are arranged in a pattern that is intermediate or transitional between those of the two systems, root and stem. Because of its intermediate structure, the hypocotyl is often designated as the transition region. In this region, the rather solid vascular core of the root, with its radial tissue arrangement and exarch xylem, is merged with separate collateral, vascular bundles having endarch xylem. These bundles are the leaf traces of the cotyledons, the first foliar organs of the shoot. Structural features of the transition region are related to spatial adjustments which provide tissue continuity between the root and shoot systems of contrasting arrangements of tissues and patterns of differentiation.

The organization of the transition region varies greatly in different species. Essential features of a pattern common among dicotyledons are shown in **Fig. 5**. In this example the diarch xylem plate of the root is continuous with two leaf traces of the cotyledons, each obviously double in structure. At successively higher levels the relative position of protoxylem and metaxylem changes. Each phloem strand of the root is connected with each of the two cotyledonary traces. Initial union with the vascular system of the epicotyl (shoot above the cotyledons) is established indirectly with the root, the traces of the first leaves remaining distinct throughout the hypocotyl.

Other types of transition occur in plants with more than one cotyledonary trace or with three or more

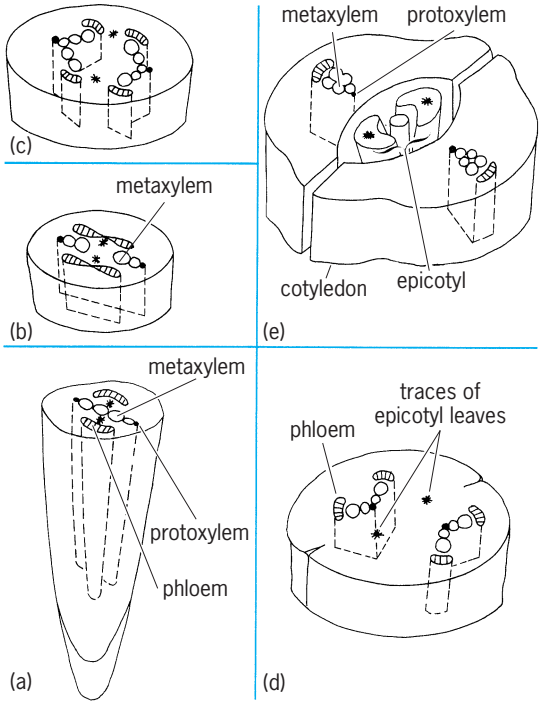

Fig. 5. Diagrams of vascular transition in dicotyledonous seedling (beet, *Beta vulgaris*) from root tip upward to level of cotyledons. (*a*) Root is diarch; its vascular tissues are continuous with those of (*e*) the cotyledons through (*b–d*) the transition region in the hypocotyl. (*After K. Esau, Plant Anatomy, John Wiley and Sons, 1953*)

poles of xylem and phloem in the root. Differences in the timing of epicotyl development and in the pattern of germination (epigean or hypogean cotyledons) also effect variations in the transition region. In plants that develop a cambium, secondary tissues are continuous between the root and stem.

The general organization of the transition region in gymnosperms is similar to that of dicotyledons, any variations being related to the usually greater number of cotyledons in gymnosperms and to the number of traces to each cotyledon. In monocotyledons the vascular system of the root is commonly connected with those of the single cotyledon and the epicotyl; transitional characteristics may occur in both connections or only in that with the cotyledon. A particularly complex transition characterizes the grass family, Gramineae. Vascular strands of the root join in a nodal plate at about the level of the scutellum or cotyledon. Also joined with the plate are vascular strands of the scutellum, coleoptile (first leaflike structure above the cotyledon), and first leaves. In this complex pattern the transition between vascular systems of root and shoot is relatively abrupt, and the root system is connected with more than one leaf above the scutellum. *See* SEED.

Charles Heimsch; Lewis Feldman

Root density characterization. The density of roots in a volume of soil is important in determining the ability of the plant to extract water and nutrients located there. The density is characterized by measuring either the length or mass of roots in a specific soil volume. A traditional means for obtaining root length density (feet of roots per cubic foot of soil or meters of roots per cubic meter of soil) or root mass density (pounds of roots per cubic foot of soils or grams of roots per cubic meter of soil) has been to extract representative samples from the soil volume occupied by the roots, wash the soil from the roots, and measure the root length or root mass contained in the sample. This procedure is very time-consuming and labor-intensive if the plants are large. In addition, part or all of the root system is destroyed by the sampling process, making successive measurements impossible on the same roots.

Minirhizotrons, rhizotrons, computer-aided tomography, neutron radiography, and nuclear magnetic resonance are techniques which estimate root length density without disturbing the plant root system. Computer-aided analyses and enhancement of images obtained by these techniques also allow root mass density to be determined, sometimes without any destructive sampling of the root system.

Minirhizotrons. This technique involves installing a clear tube into the soil and lowering into the tube a viewing device that allows the operator to see roots that intersect the tube. Tubes are installed at 30–45° from the vertical a few days before measurements are to be started. The tubes are marked with permanent reference points prior to installation. A battery-powered light allows the roots to be observed through a mirror, a fiber-optic lens, and a borescope or a color video camera with a right-angle lens. Photographic or video tape images can be stored for later processing. Root length density can be determined by measuring the length of roots at the interface and estimating the volume of soil which these roots would have occupied if the tube had not been present, or by estimating the length of roots that would have grown in a specific volume of soil occupied by the tube. Root mass densities can be estimated from root length density, root diameter distribution, and specific weight values. Several tubes are required to estimate root densities accurately.

Rhizotrons. These are covered underground cellars or walkways with clear windows on one or both sides. The windows may be in contact with natural soil that surrounds the cellar or may be one wall of a soil-filled box. Plants are grown in soil adjacent to the window. Roots are observed either at the window–soil interface or by inserting a viewing device into clear tubes placed horizontally into the soil at different depths. Empirical relationships can be developed between the length of roots at the soil–window interface and root length density in soil away from the window when the horizontally placed tubes are used to view the roots. The number of roots that intersect the tubes are counted, and empirical relationships are developed between root counts and root length densities. Roots tend to be more concentrated near the window than in bulk soil, but light shining on the roots through the window tends to reduce root concentration at the interface, especially for most legumes.

Computer-aided tomography (CAT). This technique uses x-rays or gamma rays to measure the density of a soil

matrix and to infer the distribution of water in that matrix. Living plant material, such as a root, has a high concentration of water and can be seen readily in a computer-generated image. The soil sample is placed in a fixed position between a radiation source and a detector. The radiation source and detector, which are a fixed distance apart, travel a 360° circular path around the stationary sample. The radiation beam passes through the sample, and the transmitted radiation is measured by the detector. Electrical impulses from the detector are integrated by a computer that constructs an image representing the internal structure of the sample. *See* COMPUTERIZED TOMOGRAPHY.

Neutron radiography. This technique is based on the fact that hydrogen atoms will slow (thermalize) high-energy neutrons to a greater extent than most other atoms and that roots in soil are locally high concentrations of water, which is composed of hydrogen and oxygen. Water in roots effectively removes more neutrons from a collimated beam than does water in moist soil. The beam produces an image on a rare-earth transfer screen, which can be further transferred to x-ray film.

Small sample sizes, safety concerns for the users, and the need for access to a neutron source of at least 1.4×10^6 neutrons cm^{-2} s^{-1} are the principal problems associated with neutron radiography. *See* RADIOGRAPHY.

Nuclear magnetic resonance. This technique uses a strong magnetic field and a sequence of radio-frequency pulses and magnetic-field gradients to create a digital image of the spatial distribution of water in a soil sample that contains roots. These images allow the determination of time-dependent changes in water contents of the root, the rhizosphere, and adjacent soil. Resolution of the image increases as sample size decreases. *See* NUCLEAR MAGNETIC RESONANCE (NMR); RHIZOSPHERE. Howard M. Taylor

Bibliography. H. M. Taylor (ed.), *Minirhizotron Observation Tubes: Methods and Applications for Measuring Rhizosphere Dynamics*, Amer. Soc. Agron. Spec. Publ. 50, 1987.

Root (mathematics)

If a function $f(x)$ has the value 0 for $x = a$, a is a root of the equation $f(x) = 0$. The fundamental theorem of algebra states that any algebraic equation of the form $a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n = 0$, where the a_k's are real numbers ($k = 0, 1, \ldots, n$), has at least one root. From this it follows readily that such an equation has roots, real or complex, in number equal to the index (here n) of the highest power of x.

Furthermore, if $a + ib$ (where $i = \sqrt{-1}$) is a complex root of the given equation, so is $a - ib$, the conjugate of $a + ib$. Equations of degrees up to four may be solved algebraically. This statement means that the roots may be expressed as functions of the coefficients, the functions involving the elementary arithmetical processes of addition, multiplication, raising a number to a power, or extracting the root of a certain order of a given number. It was proved by H. Abel and by E. Galois that it is not possible to solve algebraically the general algebraic equation of degree higher than four. However, it is possible to determine the real roots of an algebraic equation to any desired degree of approximation.

The term zero is sometimes used in lieu of root when dealing with functions which are defined as infinite power series. Thus, one talks about the zeros of $\sin x$, $\cos x$, and $J_0(x)$ [the Bessel function of first kind and zeroth order]. Each of these three functions may be expressed as an infinite power series.

Numbers are called transcendentals if they cannot be the roots of any algebraic equation with integral or rational coefficients. The most important transcendental numbers are π, the ratio between the circumference of a circle and its diameter; and e, the base of the system of natural logarithms. *See* CALCULUS; EQUATIONS, THEORY OF; NUMERICAL ANALYSIS.

Arnold N. Lowan; Salomon Bochner

Root-mean-square

The square root of the arithmetic mean of the squares of a set of numbers is called their root-mean-square. If the numbers are $x_1, x_2, x_3, \ldots, x_n$, then

$$\text{Root-mean-square} = \sqrt{\frac{x_1^2 + x_2^2 + x_3^2 \cdots + x_n^2}{n}}$$

It is valuable as an average of the magnitudes of quantities, and it is not affected by the signs of the quantities.

Among applications of root-mean-square the most important is the standard deviation from the arithmetic mean $\bar{x}$:

$$\bar{x} = \frac{x_1 + x_2 + x_3 \cdots + x_n}{n}$$

Then the standard deviation

$$s = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + (x_3 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n}}$$

Thus standard deviation from the mean is the root-mean-square of the deviations from the mean. It has a great advantage over other measures of deviation because its computation can be greatly facilitated by the following transformation. If a set of values $x_1, x_2, x_3, \ldots, x_n$ are each reduced by the same number k and then the results are divided by the same number p, the root-mean-square of the deviations of the resulting quantities from their mean is equal to $1/p$ times the root-mean-square of the deviations of the x's from their mean. Thus if $x_i = pu_i + k$, Standard deviation of the x's $= p \times$ (standard deviation of the u's).

The most common method of minimizing errors in the use of a formula to fit empirical observations is that of determining the coefficients in the formula so as to make the mean of the squares of the errors

least. This is the principle of least squares. Its relation to root-mean-square is obvious since minimizing the root-mean-square of the errors accomplishes the same purpose. *See* CURVE FITTING; STATISTICS.

Hollis R. Cooley

Rope

A long flexible structure consisting of many strands of wire, plastic, or vegetable fiber such as manila. Rope is classified as a flexible connector and is used generally for hoisting, conveying, or transporting loads; transmitting motion; and occasionally transmitting power. For flexibility and to reduce stresses as the rope bends over the sheave (pulley), a rope is made of many small strands.

Wire rope. In making wire rope, the wires are preformed and wound into strands which are then twisted together to form the rope. A hemp center is usually used to hold lubricant for reducing friction and wear as the wires rub over each other when the rope is bent. *See* WIRE.

Figure 1 shows a 6 × 7 construction (6 strands, each with 7 wires). The hemp center is shown in black. Usual constructions are 6 × 7 coarse laid, 6 × 19 standard hoisting, and 8 × 19 and 6 × 37 extraflexible hoisting for elevator service. To keep bending stresses low, sheaves of large diameter must be used. Factors of safety based on ultimate strength of the rope vary from 3 to 8 depending on the application. Some manufacturers' data allow for bending stresses in normal size sheaves; otherwise these stresses must be allowed for separately.

There are two common ways of winding wire rope indicated by lay of the rope. Standard rope is called regular lay; the wires in the strand are laid to the left and the strands to the right. The wires are at right angles to the motion of the rope as it slides into the sheave groove so that the wire may wear excessively, particularly if the sheaves are not well aligned. In lang lay, wires in the strand and the strands themselves are laid in the same direction. This type of rope wears better, is more flexible, and lasts longer. Flat rope is made of right and left strands placed alternately beside each other and sewed together with soft steel wire. This rope can be used on a reel slightly wider than the rope; it is used to save space.

For aircraft service, 6 × 7, 7 × 7, and 7 × 19 cord and 19-wire strand have been developed. They are made of tinned or galvanized carbon steel or stainless steel.

Materials for commercial wire ropes in order of strength and cost are iron, mild plow steel, plow steel, and improved plow steel. Galvanized-wire rope

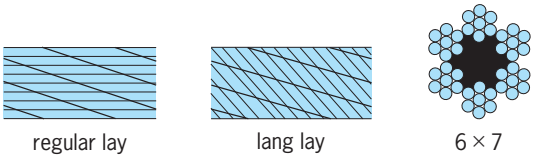

Fig. 1. Examples of wire-rope construction.

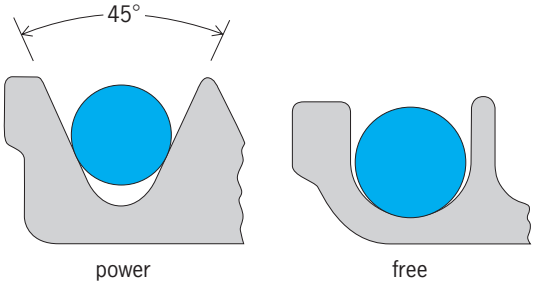

Fig. 2. Groove profiles for rope.

may be used for standing service, as for guy wires, but should not be used in hoisting because the galvanizing wears off. Failure of wire rope is usually caused by wear of wires as they rub on each other during bending or by high compressive stress as they bear against the grooves in power transmission (**Fig. 2**). Internal corrosion due to water or other corrosive substances causes failure. Profiles of grooves on the sheaves are semicircular at the bottom when the rope runs free for hoisting and transportation (Fig. 2). The rope should not be pinched at the bottom of the groove. For power transmission with wire ropes, the contacting part of the semicircular groove should be lined with rubber or similar resilient material.

Wire ropes may be attached to a hoisting drum by using a rope socket at the end of the rope; the socket is slipped into a pocket cast in the drum at the end of the last groove. The socket is held in place by a screwed plug, which can be easily removed to detach the rope. To attach the rope to its load or to a block-and-tackle anchor, various sockets, clips, and thimbles are available.

Manila rope. Manila fiber is twisted into yarn, the yarn is twisted into strands, and the strands are laid up into rope. Manila rope is used mainly for light service hoisting but sometimes for long-distance transmission of power. Being flexible, this rope accommodates itself to small-size pulleys and absorbs starting and shock loads, but the fiber at the center of the strand powders and pulverizes, although the rope may appear sound at the surface. This deterioration, from the rubbing of strands on each other as the rope flexes, can be retarded by such lubricants as acid-free paraffin or graphite worked into the strands during laying or twisting of the strands to form the rope. *See* HEMP; NATURAL FIBER.

For power transmission and for hoisting with manila rope see the groove profiles shown in Fig. 2. The 45° groove angle that is used for power transmission provides sufficient friction to prevent slippage between the rope and the surface of the groove. *See* PULLEY.

Paul H. Black

Rosales

An order of flowering plants in the eurosid I group of the rosid eudicots. In previous modern systems of classifications, this order was considered to be

much larger and to contain up to 25 families, but recently on the basis of DNA sequence studies the number of families was greatly reduced and changed. The order as now recognized contains only 11 families, many of which are small. Many families in this order are wind-pollinated (the former Urticales, for example Moraceae, Ulmaceae, and Urticaceae) and exhibit the typical syndrome of small petalless flowers with dangling anthers, whereas the other families are insect-pollinated with large, showy flowers in which the carpels are sometimes free. *See* FLOWER; MAGNOLIOPHYTA; MAGNOLIOPSIDA; POLLINATION.

The largest family is Rosaceae with nearly 3000 species. The great number of economically important trees, shrubs, and herbs in this family include apples (*Malus*), pears (*Pyrus*), almonds, cherries, plums, and prunes (*Prunus*), strawberries (*Fragaria*), blackberries, raspberries, and their relatives (*Rubus*), as well as many minor fruits such as quinces (*Cydonia*), loquat (*Eriobotrya*), and medlars (*Mespilus*). Many are grown as ornamental plants, including roses (*Rosa*), avens (*Geum*), cinquefoil (*Potentilla*), firethorn (*Pyracantha*), redbush (*Photinia*), and spirea (*Spiraea*). *See* ALMOND; APPLE; BLACKBERRY; CHERRY; ORNAMENTAL PLANTS; PEAR; PLUM; QUINCE; RASPBERRY; STRAWBERRY.

The second-largest family, Rhamnaceae (900 species), includes mostly woody species (shrubs and trees) that differ from Rosaceae in having fused carpels and only five stamens (rather than many) that are placed in the same position as the petals. A number of species in Rhamnaceae are of minor economic importance as timbers, some are of medicinal use or as dyes, and one is a minor fruit crop, the jujube (*Zizyphus*). A few are ornamentals, such as *Ceanothus* and *Colletia*.

Of the families formerly placed in Urticales, the largest are Moraceae (950 species) and Urticaceae (700 species), the exact limits of which are problematic. Many of these are large forest trees, but some are forest herbs and weeds. In Moraceae, many species are sources of timber and fruit, including figs (*Ficus*), mulberries (*Morus*), and breadfruits (*Artocarpus*). In Urticaceae, fiber-producing species are common, including hemp (*Cannabis*), ramie or China grass (*Boehmeria*), and bast-fiber (*Urtica*), the last infamous due to the stinging hairs present on many species. Ulmaceae and Celtidaceae are small families of north temperate forest trees, many of which provide useful timbers, including elm (*Ulmus*), hackberries (*Celtis*), and *Zelkova*. *See* ELM; FIBER CROPS; FIG; HACKBERRY; HEMP; MULBERRY; RAMIE.

Mark W. Chase

Rose curve

A type of plane curve that consists of loops (leaves, petals) emanating from a common point and that has a roselike appearance. Taking the common point O as the pole of a polar coordinate system (see **illus.**), these curves have equations of the form $\rho = a \cdot \sin n\theta$, where $a > 0$ and n is a positive integer (also

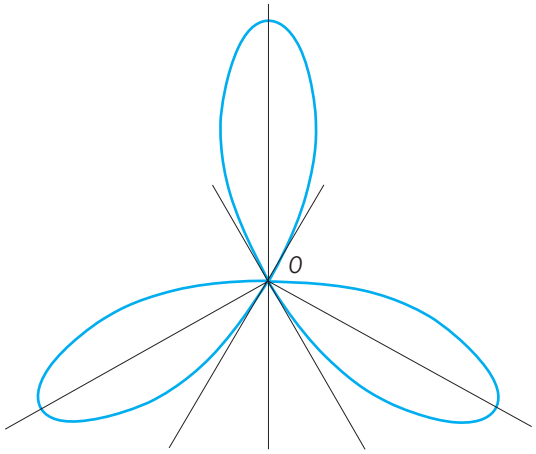

Diagram of rose curve.

$\rho = a \cdot \cos n\theta$, with a different choice of the initial line of the coordinate system). The curve is a circle of diameter a for $n = 1$. It has n or $2n$ leaves, according to whether n is an odd or even integer, respectively. The area enclosed by one leaf is $\pi a^2/n$; therefore the entire area of the three-leaved rose $\rho = a \sin 3\theta$ is $\pi a^2/4$, and that of the four-leaved rose $\rho = a \sin 2\theta$ is $\pi a^2/2$. The lemniscate is sometimes called a two-leaved rose, though its equation, $\rho^2 = a^2 \cos 2\theta$, is not of the form given above. *See* LEMNISCATE OF BERNOULLI.

Leonard M. Blumenthal

Rosemary

Rosmarinus officinalis, a member of the mint family, grown for its highly aromatic leaves and as an ornamental. Rosemary is an evergreen and perennial which can live as long as 20 years under favorable conditions. Although many varieties exist, *R. officinalis* is the only species. Most varieties are suitable for culinary use, although some (such as Pine Scented) contain high levels of terpenes and have a turpentinelike scent. Other varieties such as Forestii, Collingwood Ingram, and Lockwood de Forest have been developed as ornamentals which have large leaves and showy flowers. *See* LAMIALES.

Growth habit varies greatly. Some prostrate varieties form mats reaching only 6 in. (15 cm) in height, while other upright types, such as Erectus, become large shrubs as tall as 15 ft (4.5 m). Leaf size and shape also vary from short and needlelike to long flat blades. Plants commercially cultivated for their leaves or oil are generally upright in habit, which aids in mechanical harvesting.

Rosemary is native to the Mediterranean area, and much of the world production is harvested from wild plants growing there. Though it is widely cultivated mainly in France, Spain, and California, it will grow in most temperate areas not subject to hard frosts.

Rosemary can be propagated both by rooted cuttings and by seeds. In both instances the plants are usually started in flats and transplanted. Germination of rosemary seeds is low (only 5–20%), and it is slow

and erratic. Because of this and the genetic variability of seeds, they are little used for propagation.

Rosemary oil has been found to contain chemicals such as rosmanol which inhibit oxidation and bacterial growth. Rosemary is used in poultry seasoning and for flavoring soups and vegetables. *See* SPICE AND FLAVORING. Seth Kirby

Bibliography. S. Arctander, *Perfume and Flavor Materials of Natural Origin*, 1960; L. H. Bailey, *Manual of Cultivated Plants*, rev. ed., 1975; F. Rosengarten, *The Book of Spices*, 1969.

Rosidae

A former widely recognized subclass of Magnoliophyta (angiosperms or flowering plants). Deoxyribonucleic acid (DNA) and morphological studies have demonstrated that this group includes many more families than previously thought (for example, a large number of families previously referred to the subclass Dilleniidae). Whereas formal reclassification awaits further evidence, it is clear that there is a large group that can be considered rosid. Approximately 136 families (roughly 30% of the total angiosperms) in 11 orders are included in the rosids. With the asterids (often recognized as Asteridae), the rosids represent the two most advanced lines of dicots.

The primary distinguishing features of rosids are their diplostemenous flowers (anthers in two whorls, each whorl typically numbering the same as the perianth parts, although in some groups the number of stamens has increased) with unfused parts, except for the carpels (asterids have haplostemenous flowers with fused parts, typically with the stamens fused at least to the bases of the petals). Rosids have ovules with two integuments (that is, they are bitegmic, as opposed to the unitegmic asterids) and a several-layered nucellus (that is, they are crassinucellate, as opposed to the tenuinucellate asterids, which have a single-layered nucellus). Ellagic acid is almost universally present (but this character occurs also in the asterid orders Cornales and Ericales). There is nuclear endosperm formation (cellular in asterids), a reticulate pollen layer, mucilaginous leaf epidermis, and generally simple perforation vessel end-walls in the wood (although the orders Malpighiales, Celastrales, and Oxalidales have more advanced scalariform perforations). The largest rosid orders are Malpighiales (16,100 species), Fabales (17,500), and Myrtales (10,600). *See* APIALES; CELASTRALES; CORNALES; EUPHORBIALES; FABALES; GERANIALES; HALORAGALES; LINALES; MAGNOLIOPSIDA; MYRTALES; PLANT KINGDOM; PODOSTEMALES; POLYGALALES; PROTEALES; RAFFLESIALES; RHIZOPHORALES; ROSALES; SANTALALES; SAPINDALES. Mark W. Chase

Rosin

A brittle resin ranging in color from dark brown to pale lemon yellow and derived from the oleoresin of pine trees. Rosin is insoluble in water, but soluble in most organic solvents. It softens at about 80–90°C (180–190°F). Rosin consists of about 90% resin acids and about 10% neutral materials such as anhydrides, sterols (mainly sitosterol), and diterpene aldehydes and alcohols. The resin acids are diterpene acids of the abietic type—abietic (**1**), neoabietic, palustric, and dehydroabietic—and the pimaric type—pimaric (**2**), isopimaric, and sandaracopimaric. Isomerization

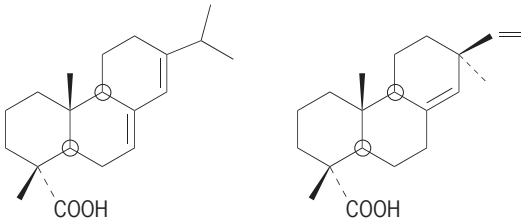

(**1**) (**2**)

among these acids is facile, and rosin composition depends upon its origin and prior processing. However, the abietic resin acids predominate over the pimaric type by a 4:1 ratio.

Rosin is obtained by wounding living trees and collecting the exudate (gum rosin), by extraction of pine stumps (wood rosin), and as a by-product from the kraft pulping process (sulfate or tall oil rosin).

Most applications of rosin involve the use of a chemically modified form. Earlier destructive distillation to viscous rosin oils has given away to hydrogenation, disproportionation (dehydrogenation), esterification, polymerization, salt formation, or reaction with maleic anhydride or formaldehyde.

The largest single use of rosin is in sizing paper to control water absorption, an application in which fortified rosin (the maleic anhydride adduct) is important. Earlier use of rosin in laundry soaps evolved into application of rosin soaps as emulsifying and tackifying agents in synthetic rubber manufacture. Glycerol and pentaerythritol esters (ester gum) are used in coatings and lacquers. Other rosin uses include adhesives, printing inks, and chewing gum. *See* PAPER; PINE TERPENE; TALL OIL. Irving S. Goldstein

Rotary engine

Internal combustion engine that duplicates in some fashion the intermittent cycle of the piston engine, consisting of the intake-compression-power-exhaust cycle, wherein the form of the power output is directly rotational.

Four general categories of rotary engines can be considered: (1) cat-and-mouse (or scissor) engines, which are analogs of the reciprocating piston engine, except that the pistons travel in a circular path; (2) eccentric-rotor engines, wherein motion is imparted to a shaft by a principal rotating part, or rotor, that is eccentric to the shaft; (3) multiple-rotor engines, which are based on simple rotary motion of two or more rotors; and (4) revolving-block engines, which combine reciprocating piston and rotary motion. Some of the more interesting engines of each type are discussed in this article.

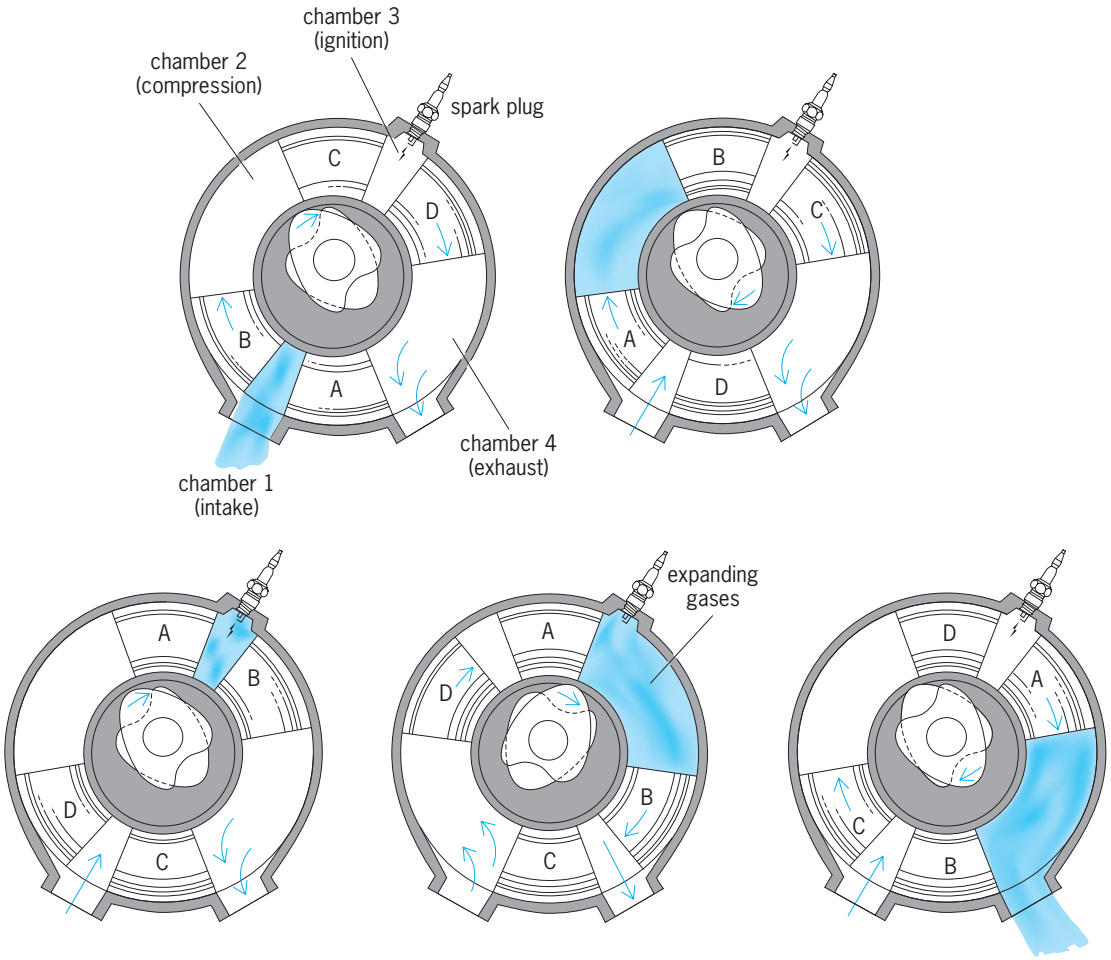

Fig. 1. Intake-compression-power-exhaust cycle of the Tschudi engine.

Cat-and-mouse engines. Typical of this class is the engine developed by T. Tschudi, the initial design of which goes back to 1927. The pistons, which are sections of a torus, travel around a toroidal cylinder. The operation of the engine can be visualized with the aid of **Fig. 1**, where piston A operates with piston C, and B with D. In chamber 1 a fresh fuel-air mixture is initially injected while pistons C and D are closest together. During the intake stroke the rotor attached to pistons B and D rotates, thereby increasing the volume of chamber 1. During this time the A-C rotor is stationary. When piston D reaches its topmost position, the B-D rotor becomes stationary, and A and C rotate so that the volume of chamber 1 decreases and the fuel-air mixture is compressed.

When the volume of chamber 1 is again minimal, both rotors move to locate chamber 1 under the spark plug, which is fired. The power stroke finds piston D moving away from piston C, with the A-C rotor again locked during most of the power stroke. Finally, when piston D has reached bottom, the B-D rotor locks, the exhaust port has been exposed, and the movement of piston C forces out the combustion product gas. Note that four chambers exist at any time, so that at each instant all the processes making up the four-strike cycle (intake-compression-power-exhaust) are occurring.

The motion of the rotors, and hence the pistons, is controlled by two cams which bear against rollers attached to the rotors. The cams and rollers associated with one of the rotors disengage when it is desired to stop the motion of that rotor. The shock loads associated with starting and stopping the rotors at high speeds may be a problem with this engine, and lubrication and sealing problems are characteristic of virtually all the engines discussed herein. However, the problem of fabricating toroidal pistons does not appear to be as formidable as was once believed.

An engine similar to the Tschudi in operation is that developed by E. Kauertz. In this case, however, the pistons are vanes which are sections of a right circular cylinder. Another difference is that while one set of pistons is attached to one rotor so that these two pistons rotate with a constant angular velocity, the motion of the second set of pistons is controlled by a complex gear-and-crank arrangement so that the angular velocity of this second set varies. In this manner, the chambers between the pistons can be made to vary in volume in a prescribed manner. Hence, the standard piston-engine cycle can be duplicated. Kauertz tested a prototype which was found to run smoothly and to deliver 213 hp (160 kW) at 4000 rpm. Here again, however, the varying angular velocity of the second set of pistons must produce

inertia effects that will be absorbed by the gear-and-crank system. At high speeds over extended periods, problems with this system are likely to be encountered.

An advanced version of the cat-and-mouse concept called the SODRIC engine has been developed by K. Chahrouri. Unlike the Tschudi engine, in which the four processes of intake-compression-power-exhaust are distributed over 360° of arc, the SODRIC engine performs these same four processes in 60° of arc. Hence, six power strokes per revolution are achieved, resulting in very substantial improvements in engine performance parameters. For example, Chahrouri has estimated that 225 hp (168 kW) can be achieved at only 1000 rpm using an engine having a toroid radius of 8 in. (20 cm) and 1-in.-radius (25-mm) pistons. Chahrouri has also improved upon the method by which alternate acceleration-deceleration of the pistons is achieved, and power is transmitted to the output shaft by using noncircular gears.

The "cat-and-mouse" and "scissors" characterizations of these engines should be clear once the picture of pistons alternately running away from, and catching up to, each other is firmly in mind. Other engines of this type, including the Maier, Rayment, and Virmel designs, differ principally in the system used to achieve the cat-and-mouse effect.

It should be noted that since the length of the power stroke is readily controlled in these engines, good combustion efficiencies (close to complete combustion) should be attainable.

Eccentric-rotor engines. The rotary engine which has received by far the greatest development to date is the Wankel engine, an eccentric-rotor type. The basic engine components are pictured in **Fig. 2**. Only two primary moving parts are present: the rotor and the eccentric shaft. The rotor moves in one direction around the trochoidal chamber, which contains peripheral intake and exhaust ports.

The operation of this engine can be visualized with the aid of Fig. 2. The rotor divides the inner volume into three chambers, with each chamber the analog of the cylinder in the standard piston engine. Initially, chamber AB is terminating the intake phase and commencing its compression phase, while chamber BC is terminating its compression phase and chamber CA is commencing its exhaust phase. As the rotor moves clockwise, the volume of chamber AB approaches a minimum. When the volume of chamber AB is minimal, the spark plug fires, initiating the combustion phase in that chamber. As combustion continues, the point is reached where the exhaust port is exposed, and the products of combustion are expelled from chamber AB.

To increase the chamber volumes, each segment

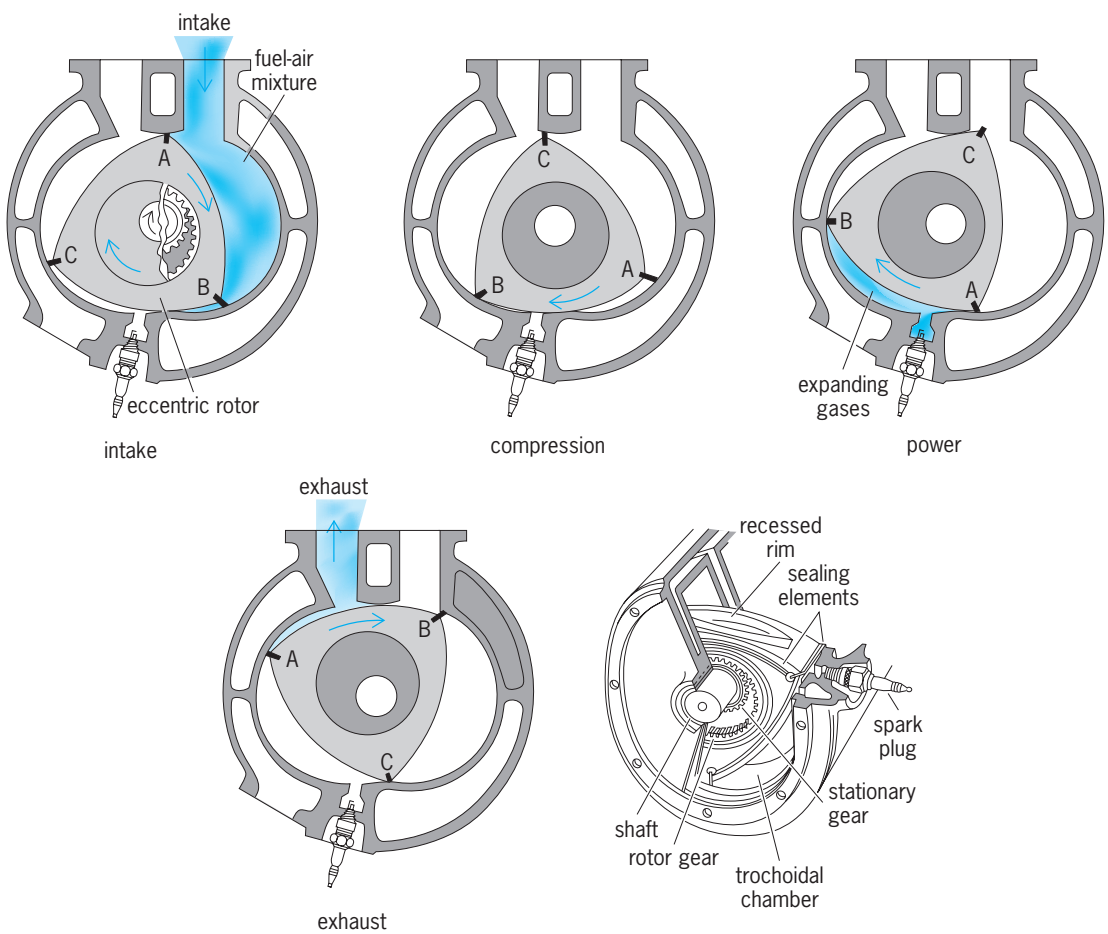

Fig. 2. Operation and basic components of the Wankel engine.

of the rotor rim is recessed (Fig. 2). During the combustion-expansion phase, unburned gas tends to flow at high velocity away from the combustion zone toward the opposite corner. As a result, this engine has a tendency to leave a portion of the charge unburned, similar to the problem encountered in ordinary piston engines. In addition to reducing the engine performance, this unburned gas is a source of air pollution. Efforts have been directed toward increasing the turbulence in each chamber, thereby improving the mixing between the burned and unburned gases, leading to better combustion efficiency.

However, Wankel engines have demonstrated impressive advantages when contrasted with standard engines. Some of these are listed below.

1. The Wankel has superior power-to-weight ratio; that is, it generally produces more or at least comparable horsepower per pound of engine weight when compared with conventional piston engines.

2. To increase power output, additional rotor-trochoidal chamber assemblies can readily be added, which occupy relatively little space and add little weight. In piston engines, cylinder volumes must be increased, leading to substantial increases in weight and installation space.

3. The rotor and eccentric shaft assembly can be completely balanced; since they usually rotate at constant velocity in one direction, vibration is almost completely eliminated and noise levels are markedly reduced.

4. As with the cat-and-mouse engines, the intake and exhaust ports always remain open, that is, gas flow into and from the engine is never stopped, so that surging phenomena and problems associated with valves which open and close are eliminated.

5. Tests indicate that Wankel engines can run on a wide variety of fuels, including ordinary gasoline and cheaper fuels as well.

6. After considerable development, reasonably effective sealing between the chambers has been achieved, and springs maintain a light pressure against the trochoidal surface.

7. The Wankel has so few parts, relative to a piston engine, that in the long run it will probably be cheaper to manufacture.

The initial application of the Wankel engine as an automotive power plant occurred in the NSU Spider. In the early 1970s, however, the Japanese automobile manufacturer Mazda began to use Wankel engines exclusively. However, relatively high pollutant emissions, coupled with low gasoline mileage for automobiles of this size and weight, resulted in poor sales in the United States. Mazda ceased marketing Wankel-powered automobiles in the United States in the mid-1970s. Several American automobile manufacturers have experimented with Wankel-powered prototypes, but no production vehicles have emerged.

The Wankel engine is being used as a marine engine and in engine–electric generator installations, where its overall weight and fuel consumption have proved to be superior to those of a diesel engine or gas turbine generating equivalent power. Other

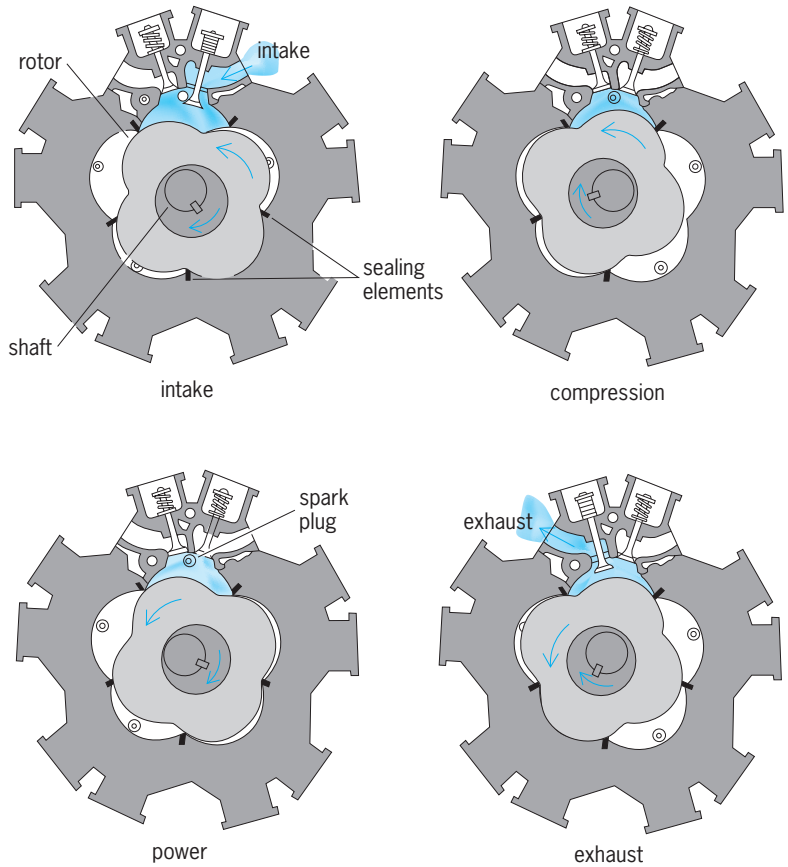

Fig. 3. Renault-Rambler engine.

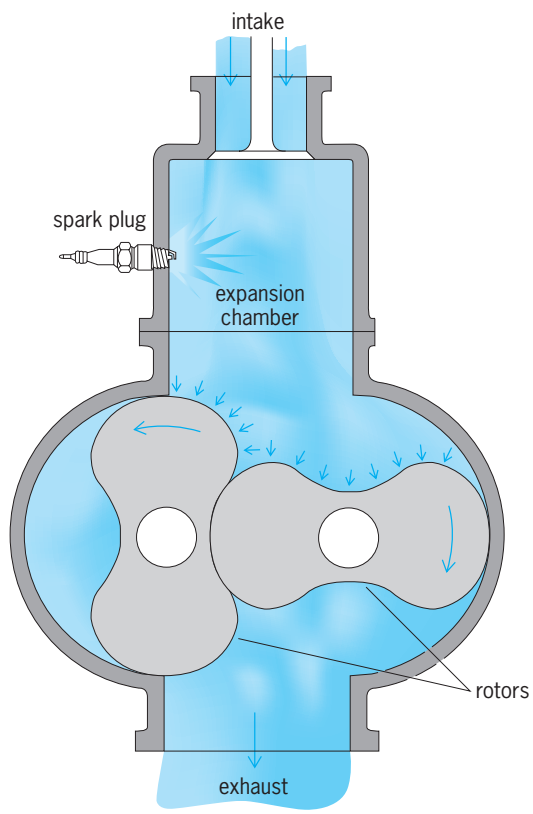

Fig. 4. Simple multirotor engine.

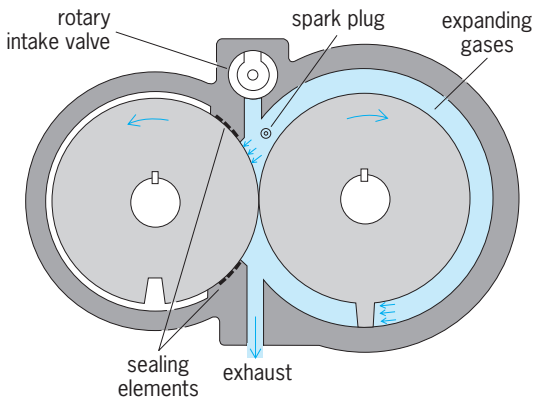

Fig. 5. Unsin engine, with two circular rotors.

projected applications include lawnmower and chainsaw engines. This wide range of applications is made possible by the fact that almost any size of Wankel engine is feasible.

An engine conceptually equivalent to the Wankel was developed jointly by Renault, Inc., and the American Motors Corp. It is sometimes called the Renault-Rambler engine. In this case, however, the rotor consists of a four-lobe arrangement, operating in a five-lobe chamber (**Fig. 3**). When a lobe moves into a cavity, which is analogous to the upward motion of the piston in the cylinder, the gas volume decreases, resulting in a compression process. The operation of this engine is detailed in Fig. 3. The fact that each cavity has two valves (intake and exhaust) represents a significant drawback. However, sealing between chambers may be simpler than in the Wankel; since each cavity acts as a combustion chamber, heat is evenly distributed around the housing, resulting in little thermal distortion.

It can be concluded that engines of the eccentric-rotor type are an integral part of the internal combustion engine scene. Their inherent simplicity, coupled with their advanced state of development, make them attractive alternatives to the piston engine in a number of applications.

Multirotor engines. These engines operate on some form of simple rotary motion. A typical design, shown in **Fig. 4**, operates as follows. A fuel-air mixture enters the combustion chamber through some type of valve. No compression takes place; rather, a spark plug ignites the mixture which burns in the combustion chamber, with a consequent increase in temperature and pressure. The hot gas expands by pushing against the two trochoidal rotors. The eccentric force on the left-hand rotor forces the rotors to rotate in the direction shown. Eventually, the combustion gases find their way out the exhaust.

The problems associated with all engines of this type are principally twofold: The absence of a compression phase leads to low engine efficiency, and sealing between the rotors is an enormously difficult problem. One theoretical estimate of the amount of work produced per unit of heat energy put into the engine (by the combustion process), called the thermal efficiency, is only 4%.

The Unsin engine (**Fig. 5**) replaces the trochoidal rotors with two circular rotors, one of which has a single gear tooth upon which the gas pressure acts. The second rotor has a slot which accepts the gear tooth. The two rotors are in constant frictional contact, and in a small prototype engine sealing apparently was adequate. The recommendation of its inventor was that some compression of the intake charge be provided externally for larger engines.

The Walley and Scheffel engines employ the principle of the engine in Fig. 4, except that in the former, four approximately elliptical rotors are used, while in the latter, nine are used. In both cases the rotors turn in the same clockwise sense, which leads to excessively high rubbing velocities. (The rotors are in contact to prevent leakage.) The Walter engine uses two different-sized elliptical rotors.

Revolving-block engines. These engines combine reciprocating piston motion with rotational motion of the entire engine block. One engine of this type is the Mercer (**Fig. 6**). In this case two opposing pistons operate in a single cylinder. Attached to each piston are two rollers which run on a track that consists of two circular arcs. When the pistons are closest together, the intake ports to the chambers behind the pistons are uncovered, admitting a fresh charge (Fig. 6*a*). At this moment a charge contained between the pistons has achieved maximum compression, and the spark plug fires. The pistons separate as combustion takes place between them, which results in a compression of the gases behind the pistons (Fig. 6*b*). However, the pistons moving apart force the rollers to move outward as well. This latter motion can only occur if the rollers run on their circular track, which consequently forces the entire engine block to rotate. When the pistons are farthest apart, the exhaust ports are uncovered and the combustion gases purged (Fig. 6*c*). At this same time the compressed fresh charge behind the pistons is transferred to the region between the pistons to prepare for its recompression and combustion, which must occur because of the continuing rotation of the block.

No doubt some of the fresh charge is lost to the exhaust during the transfer process. Stresses on the roller assembly and cylinder walls are likely to be

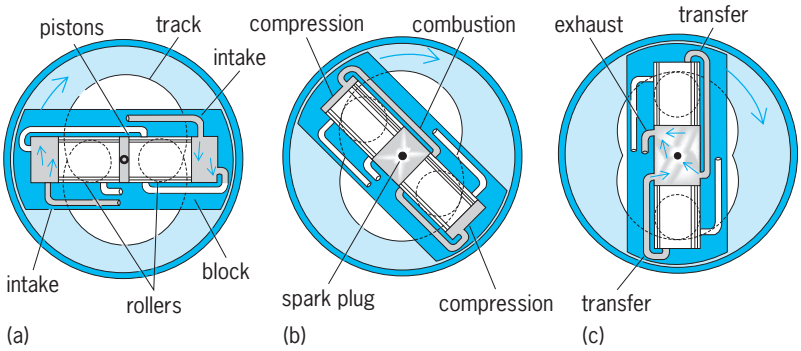

Fig. 6. Cycle of the Mercer engine. (*a*) Intake of fresh charge. (*b*) Compression of charge, following firing of spark plug. (*c*) Exhaust of combustion gases and transfer to compressed fresh charge.

quite high, which poses some design problems. Cooling is a further problem, since cooling of the pistons is difficult to achieve in this arrangement.

One the other hand, the reciprocating piston motion is converted directly to rotary motion, in contrast with the connecting rod–crank arrangement in the conventional piston engine. Also, no flywheel should be necessary since the entire rotating block acts to sustain the rotary inertia. Vibration will also be minimal.

The Selwood engine is similar in operation, except that two curved pistons 180° opposed run in toroidal tracks. This design recalls the Tschudi cat-and-mouse engine, except that the pistons only travel through 30° of toroidal track. This motion forces the entire block to rotate. The Leath engine has a square rotor with four pistons, each 90° apart, with a roller connected to each. As in the Mercer engine, the reciprocating motion of the pistons forces the rollers to run around a trochoidal track which causes the entire block to rotate. The Porshe engine uses a four-cylinder cruciform block. Again, rollers are attached to each of the four pistons. In this arrangement power is achieved on the inward strokes of the piston. Finally, the Rajakaruna engine (**Fig. 7**) uses a combustion chamber whose sides are pin-jointed together at their ends. Volume changes result from distortion of the four-sided chamber as the surrounding housing, which contains a trochoidal track, rotates. The huge pins are forced against the track. As usual, cooling and lubrication problems will be encountered with this engine, as will excessive wear of the hinge pins and track.

The Ma-Ho engine, invented by G. Hofmann, uses four cylinders welded concentrically around a central shft. As the pistons oscillate in the cylinders as a result of the intake-compression-power-exhaust processes undergone by the fuel-air mixture, the pistons rotate a barrel cam to which they are connected by cam followers. Hence, the entire block, including the central shaft, is forced to rotate.

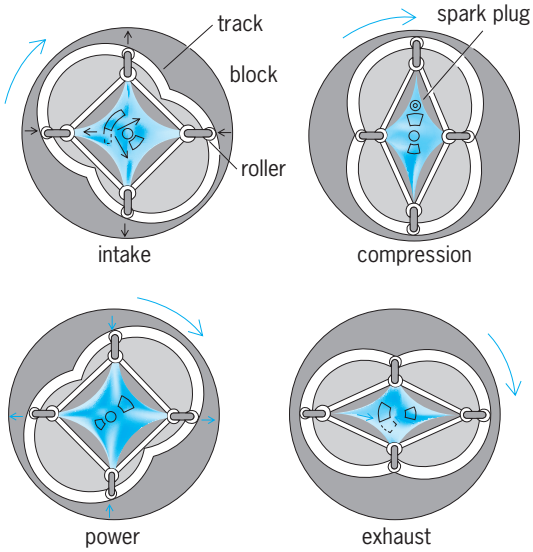

Fig. 7. Cycle of the Rajakaruna engine, showing distortion of the four-sided chamber as the surrounding housing rotates.

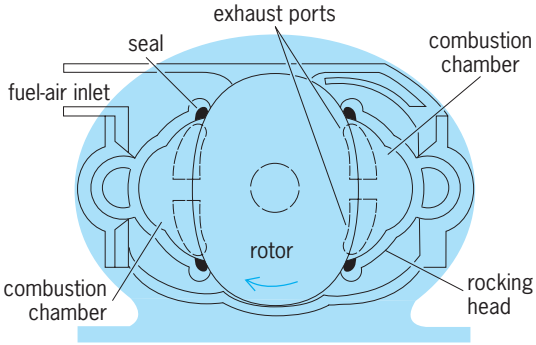

Fig. 8. Walker engine, with elliptical rotor and two C-shaped rocking.

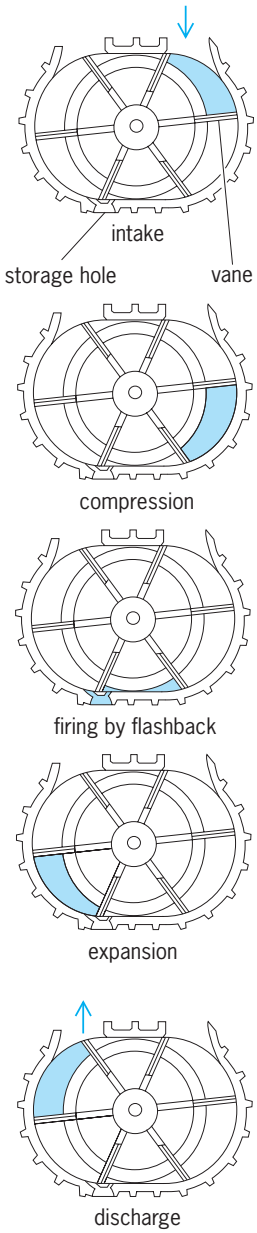

Fig. 9. Cycle of one pocket of combustion gases in the Heydrich vane-type rotary engine.

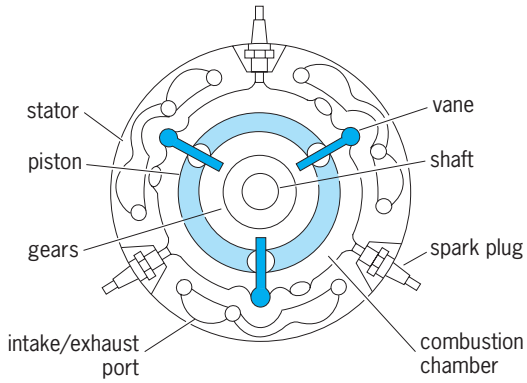

Fig. 10. Leclerc-Edmon-Benstead engine.

Engines of other types. Although the vast majority of rotary engines fall into one of the categories discussed above, several ingenious designs which do not are worthy of mention.

The Walker engine (**Fig. 8**) involves an elliptical rotor which rotates inside a casing containing two C-shaped rocking heads. The fuel-air mixture is drawn into combustion chambers on each side of the rotor; mixture cut-off occurs when the rotor is in the vertical position. As the rotor turns, the mixture undergoes compression, and combustion is initiated by spark plugs. Rotor momentum coupled with the expansion of the combustion gases forces the rotor to continue turning. Compensation for seal wear is made by adjusting the rocking heads closer to the rotor.

The Heydrich engine (**Fig. 9**) is a vane-type rotary engine which utilizes a small hole to "store" a quantity of high-temperature combustion gases from a previous firing. This gas is then used to ignite the subsequent fuel-air charge. The first charge is ignited by a glow plug. Floating seals make contact with the chamber wall.

The Leclerc-Edmon-Benstead engine (**Fig. 10**) has three combustion chambers defined by stationary vanes, a cylindrical stator, and end flanges. An output shaft passing through the chamber is geared to a circular, eccentrically mounted piston. As the chambers are fired sequentially, the movement of the piston forces the shaft to rotate. Slots in the flange control porting of the intake and exhaust gases. *See* COMBUSTION CHAMBER; DIESEL CYCLE; DIESEL ENGINE; GAS TURBINE; INTERNAL COMBUSTION ENGINE; OTTO CYCLE.

Wallace Chinitz

Bibliography. R. F. Ansdale, Rotary combustion engines, *Auto. Eng.*, vol. 53, no. 13, 1963, and vol. 54, nos. 1 and 2, 1965; W. Chinitz, Rotary engines, *Sci. Amer.*, vol. 220, no. 2, 1969; H. E. Dark, *The Wankel Rotary Engine*, 1974; W. Froede, *The Rotary Engine and the NSU Spider*, Soc. Automot. Eng. Pap. 650722, October, 1965; K. Matsumoto et al., *The Effects of Combustion Chamber Design and Compression Ratio on Emissions, Fuel Economy and Octane Number Requirement*, SAE Pap. 770193, 1977; R. H. Thring, Gasoline engines and their future, *Mech. Eng.*, vol. 15, no. 10, 1983; R. Wakefield, Revolutionary engines, *Road Track*, vol. 18, no. 3, 1966; F. Wankel, in R. F. Ansdale (ed.), *Rotary Piston Machines: Classification of Design Principles for Engines, Pumps and Compressors*, 1965.

Rotational motion

The motion of a rigid body which takes place in such a way that all of its particles move in circles about an axis with a common angular velocity; also, the rotation of a particle about a fixed point in space. Rotational motion is illustrated by (1) the fixed speed of rotation of the Earth about its axis; (2) the varying speed of rotation of the flywheel of a sewing machine; (3) the rotation of a satellite about a planet, in which both the speed of rotation and the distance from the center of rotation may vary; (4) the motion of an ion in a cyclotron, where the angular speed of rotation remains constant, but the radius of the circular motion increases; and (5) the motion of a pendulum, in which case the particles describe harmonic motion along a circular arc.

This discussion of rotational motion is limited to circular motion such as is exhibited by the first and second examples. For information concerning the other examples *see* CELESTIAL MECHANICS; EARTH ROTATION AND ORBITAL MOTION; HARMONIC MOTION; PARTICLE ACCELERATOR; PENDULUM.

Circular motion is a rotational motion in which each particle of the rotating body moves in a circular path about an axis. The motion may be uniform, that is, with constant angular velocity, or nonuniform, with changing angular velocity.

Uniform circular motion. The speed of rotation, or angular velocity, remains constant in uniform circular motion. In this case, the angular displacement θ experienced by the particle or rotating body in a time t is $\theta = \omega t$, where ω is the constant angular velocity.

Nonuniform circular motion. A special case of circular motion occurs when the rotating body moves with constant angular acceleration. If a body is moving in a circle with an angular acceleration of α radians/s^2, and if at a certain instant it has an angular velocity ω_0, then at a time t seconds later, the angular velocity may be expressed as $\omega = \omega_0 + \alpha t$, and the angular displacement as $\theta = \omega_0 t + {}^1/_2\alpha t^2$. *See* ACCELERATION; VELOCITY.

Banking of curves. When a car travels around a horizontal curve on a highway, the path is a circular arc of radius R, where R is the radius of curvature of the roadway. In order to have the car move in this circular arc, a horizontal external force must be applied to give the car an acceleration perpendicular to its path, that is, toward the center of rotation. This force must equal $M\upsilon^2/R$, where M is the mass of the car and υ its speed. This centripetal force is supplied by the friction between the tires and the road. If the force of friction is not great enough to produce this acceleration, the inertia of the car will tend to make it continue with its speed in a straight line, tangent to the road rather than around the curve, and this will cause the car to slide off the road. *See* CENTRIPETAL FORCE.

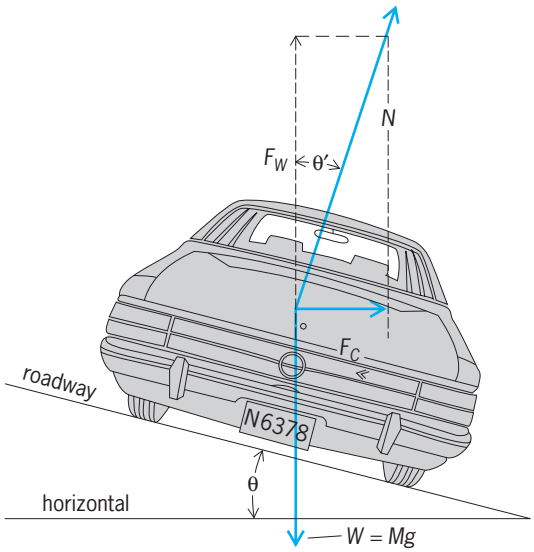

Fig. 1. Banking of a curve.

To reduce the probability of skidding roadways are customarily banked as illustrated in **Fig. 1**, which shows a car of mass M going away from the reader with a speed v and making a right-hand turn along an arc of radius R. The roadway must exert a vertical force F_W upward, equal and opposite to the weight $W = Mg$ of the car (g is the acceleration of gravity), and a horizontal centripetal force $F_C = Mv^2/R$ to make the car move in a circular arc. The net force N of the road on the car is the vector sum of these two forces.

From the diagram, it can be seen that the angle θ' which N makes with the vertical is given by Eq. (1). If

$$\tan\theta' = \frac{F_C}{F_W} = \frac{Mv^2/R}{Mg} = \frac{v^2}{gR} \qquad (1)$$

this angle θ' is equal to the bank angle θ of the road, the force N of the road on the car is perpendicular to the roadway, and there will be no tendency to skid. Equation (1) shows that the correct bank angle is proportional to the square of the speed and inversely proportional to the radius of the curve. For a given curve, there is no correct bank angle for all speeds; thus roadways are banked for the average speed of traffic. Bank angle enters into the design of railroads and into the banking of an airplane when it executes a turn.

Work and power relations. A rotating body posesses kinetic energy of rotation which may be expressed as $T_{\text{rot.}} = {}^1/_2 I\omega^2$, where ω is the magnitude of the angular velocity of the rotating body and I is the moment of inertia, which is a measure of the opposition of the body to angular acceleration. The moment of inertia of a body depends on the mass of a body and the distribution of the mass relative to the axis of rotation. For example, the moment of inertia of a solid cylinder of mass M and radius R about its axis of symmetry is ${}^1/_2 MR^2$.

To impart kinetic energy to a rotating body, work must be done. In **Fig. 2** there is represented a solid cylinder of mass M and radius R, capable of rotation without friction about an axis perpendicular to the plane of the page through O. By means of a cord wrapped around the cylinder, a constant force F is applied, thus imparting angular acceleration to the cylinder. If the cylinder is originally at rest and the force F acts through a distance $s = R\theta$ equal to the arc PP', thus rotating the cylinder through the angle θ, the work W done is $W = Fs = FR\theta = L\theta$, where $L = FR$ is called the torque or moment of force. The action of this torque L is to produce an angular acceleration α according to Eq. (2), where $I\omega$, the

$$L = I\alpha = I\frac{d\omega}{dt} = \frac{d}{dt}(I\omega) \qquad (2)$$

product of moment of inertia and angular velocity, is called the angular momentum of the rotating body. This equation points out that the angular momentum $I\omega$ of a rotating body, and hence its angular velocity ω, remains constant unless the rotating body is acted upon by a torque. Both L and $I\omega$ may be represented by vectors.

It is readily shown by Eq. (3) that the work done by

$$\begin{aligned} W &= L\theta = I\alpha\theta = I\alpha\,{}^1/_2(\alpha t^2) \\ &= {}^1/_2 I(\alpha t)^2 = {}^1/_2 I\omega^2 \end{aligned} \qquad (3)$$

the torque L acting through an angle θ on a rotating body originally at rest is exactly equal to the kinetic energy of rotation, because, for the case at hand, $\theta = {}^1/_2(\alpha t^2)$ and $\omega = \alpha t$.

Power is defined as rate of doing work, and the power P in rotational motion is given by Eq. (4).

$$P = \frac{dW}{dt} = \frac{d}{dt}(L\theta) = L\frac{d\theta}{dt} = L\omega \qquad (4)$$

See ANGULAR MOMENTUM; MOMENT OF INERTIA; MOTION; POWER; RIGID-BODY DYNAMICS; TORQUE; WORK.

Carl E. Howe; R. J. Stephenson

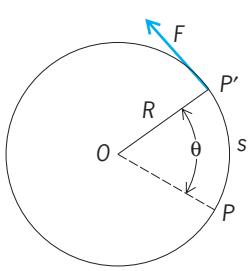

Fig. 2. Work in angular motion.

Rotifera

A phylum of microscopic, mainly free-living aquatic animals, characterized primarily by an anterior ciliary apparatus, the corona. When the cilia of the corona are in motion the animals appear to have a pair of rapidly rotating wheels on their heads, thus the name "Rotifera," Latin for "wheel bearers." Rotifers were first observed by Anton Leeuwenhoek more than 300 years ago; because they are easy to find and of a convenient size for observation with a

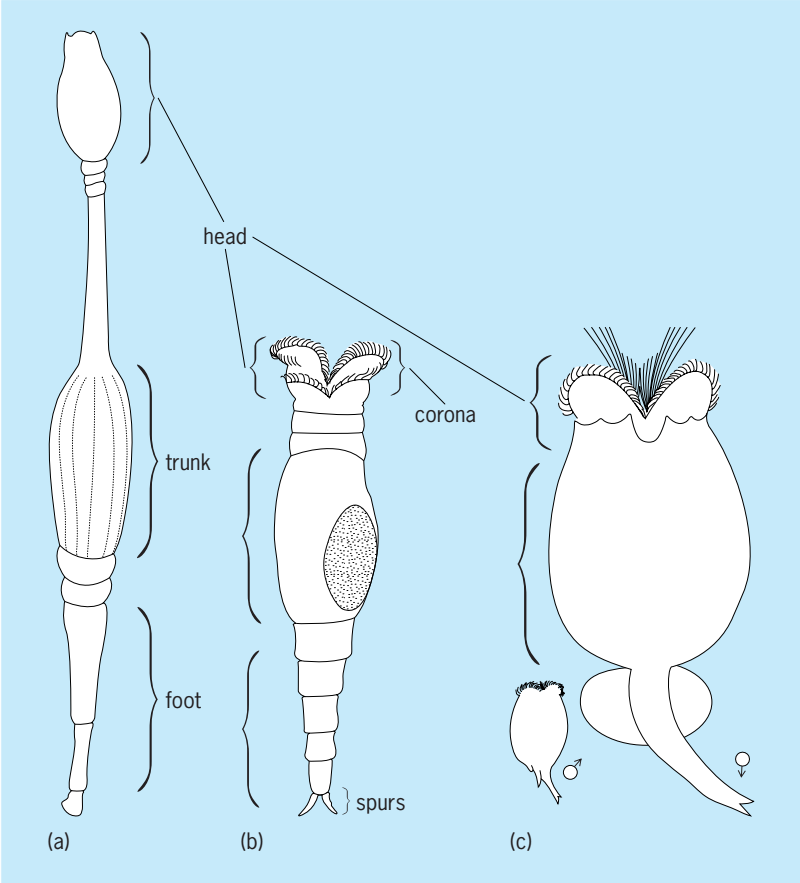

Fig. 1. Common species from the three rotifer groups. (*a*) The seisonid *Seison nebaliae*. (*b*) The bdelloid *Philodina roseola*. (*c*) The monogonont *Brachionus plicatilis*.

simple microscope, they have been avidly studied by amateur naturalists ever since.

Rotifers show considerable diversity in form and structure. They are bilaterally symmetrical animals with complete digestive, excretory, nervous, and reproductive systems but lack separate respiratory and circulatory systems. Most have about 1000 nuclei and are 0.1–1.0 mm in length. They possess three characteristic morphological features: the corona; a syncytial epidermis with an intracytoplasmic lamina; and the mastax, a gizzardlike structure derived from a modified pharynx.

Our understanding of the evolutionary relationships within Rotifera and between Rotifera and related phyla have been undergoing extensive revision. Older classifications placed rotifers in the phylum Aschelminthes, an incongruous assemblage including also the Nematoda, Priapulida, Kinorhyncha, and Gastrotricha, and divided rotifers into three groups: Monogononts, Bdelloids, and Seisonids. Because of new molecular and ultrastructural information, Aschelminthes is no longer considered a valid group and Rotifera is generally considered to be a phylum, with the three groups now given class status. Only Bdelloidea and Monogononta possess all the classical rotifer features and are therefore sometimes referred to as the Eurotatoria. The obligately parasitic Acanthocephala is a closely related phylum or may be a fourth group of Rotifera. Rotifera and Acanthocephala are often referred to as Syndermata, which in turn is part of an assemblage of microscopic animals including Micrognathizoa and Gnathostomulida called Gnathifera, because of the shared characteristic of jawlike structures made of sclero-protein.

Morphology. The body of a rotifer is usually divided into three parts: the head, trunk, and foot (**Fig. 1**). The head carries the corona, the mouth, and mastax of the digestive system, as well as the central ganglion (brain) of the nervous system. Most organs are located in the trunk, including the stomach, intestine, cloaca, anus, and gastric glands of the digestive system, the simple excretory system, and the reproductive organs. The excretory system consists primarily of paired nephridial tubes, provided with flame bulbs and opening posteriorly either into a bladder or directly into the cloaca. The female reproductive organs consist of paired ovaries and vitellaria; the ovaries open into the cloaca and are fused in monogononts. The foot contains cement glands whose secretions enable many species to attach themselves to the substratum. Sensory organs consist of one or more eyespots, absent in some species, as well as sensory bristles, papillae, and, in most species, a dorsal antenna. The majority of rotifer tissues and organs are syncytial.

Rotifers are covered by a thin; nonchitinous cuticle and an underlying epidermal intracytoplasmic lamina, unique to rotifers and acanthocephalans, which in some species is thickened into a hard case, or lorica, over much of the exterior epidermis. Loricate rotifers often have characteristic patterns and ornamentations, which are useful for species identification. The trunk and foot of many species, particularly bdelloids, have a segmented appearance; however, rotifers are not segmented in the classical sense, nor do they represent an early stage of segmentation. Sessile monogononts and some bdelloids construct a gelatinous tube around themselves, which in some species may be extensively ornamented with small debris and waste particles. Species in the genus *Floscularia* have modified mouthparts used to concentrate particles into pellets that are then attached to the tube in highly ordered rows.

The muscular system is complex. It includes bands of circular muscles in the body wall that allow the rotifer to elongate, longitudinal muscles in the body wall that control expansion and retraction of the head and foot, muscles associated with viscera, and others. Many rotifers can retract their head or foot sections into the trunk; when disturbed, bdelloids in particular will contract into a nondescript ellipsoid.

The pharynx is modified into a structure peculiar to rotifers, the muscular mastax, which contains a masticatory apparatus, the trophus. The terms mastax and trophus are used as synonyms by many workers. Typical trophi contain seven parts: three central pieces, making up the incus, and four lateral pieces, forming the paired mallei. The three parts of the incus include a posterior supporting structure, the fulcrum, and two rami. Each malleus consists of a head, or uncus, and a handlelike piece, or manubrium.

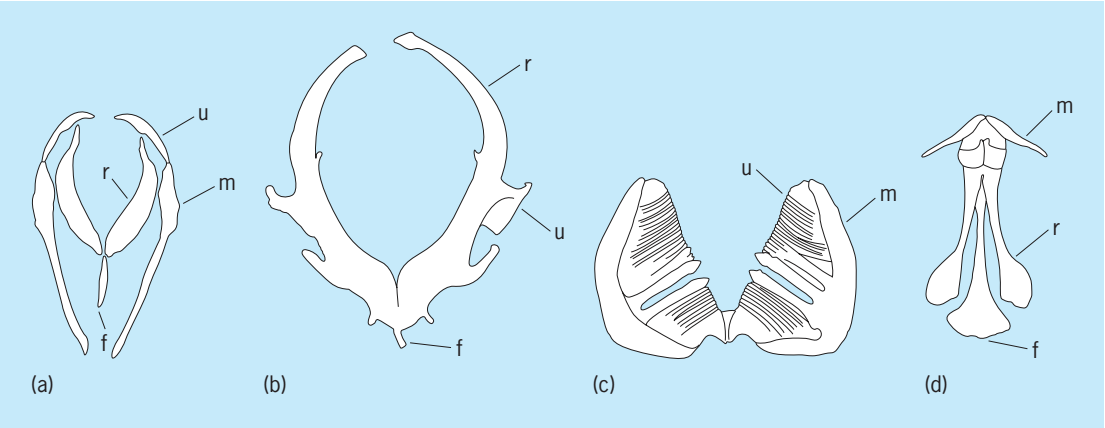

Fig. 2. Four of the 8 or 9 recognized types of rotifer mastax. (*a*) Forcipate mastax of the monogonont *Albertia vermiculus* has all parts of the mallei and incus. (*b*) Incudate mastax of the monogonont *Asplanchna brightwellii* has enlarged, pincerlike rami and reduced mallei. (*c*) Ramate mastax of the bdelloid *Macrotrachela quadricornifera* has a reduced incus and toothlike unci. (*d*) Fulcrate mastax of *Seison nebaliae* has an enlarged fulcrum and reduced mallei. Here u, unci; r, rami; f, fulcrum; m, mallei.

There is extensive diversification of mastax morphology between rotifer groups (**Fig. 2**). Some predatory rotifers have pincerlike rami and extend their trophus out of their mouth to seize prey; others have an enlarged fulcrum supporting a muscular piston that acts to suck out the contents of prey captured by toothed unci. In rotifers that feed exclusively on microphytoplankton and particulate matter, the unci are adapted for grinding and are often quite large, while other parts may be degenerate.

The most notable feature of rotifers is the corona, composed of a ventral, ciliated buccal field, which surrounds the mouth, and a more anterior circumapical band of cilia that extends around the head, enclosing an unciliated apical field. Large, powerful outer cilia of the circumapical band form the anterior trochus and posterior cingulum. As with the mastax, there is great diversity of form of the corona, with particular regions enlarged or reduced, reflecting habitat, means of locomotion, and feeding. Only a relatively small number of rotifer groups possess a corona that actually presents the impression of rotating wheels. The effect is most clearly seen among bdelloid species, where the apical field and trochus is divided into two trochal disks raised on pedicels. When the animal is disturbed or crawling, the pedicels are retracted into the head. When extended, rhythmic rapid strokes of the trochal cilia, followed by a slower recovery, give the striking appearance of a pair of spinning disks.

Feeding and locomotion. Most rotifers are free-living and feed on bacteria, algae, and organic debris. Water currents created by the beating of the corona draw particles into the mouth area, where each is examined by sensory cilia and either passed on to the mouth or rejected. Most bdelloids are of this type. Predatory rotifers may trap prey in an enlarged corona (for example, *Collotheca*) or, more commonly (for example, *Asplanchna, Dicranophorus, Philodinavus*), extrude their modified pincerlike mastax to seize prey such as ciliates, nematodes, and smaller rotifers. Some rotifers, particularly in the genus *Proales*, are parasitic on aquatic animals such as cladocerans, oligochaetes, insect larvae, or snail eggs, or on algae such as *Vaucheria* and *Volvox*. Those in the genera *Albertia, Balatro*, and *Drilophage* are nearly all obligately parasitic, feeding on body fluids of oligochaetes, leeches, or earthworms. They remain attached to their host by their mastax and have highly reduced corona and feet. Many more species are epizoic commensal or opportunistically parasitic symbionts: The only marine bdelloid species, *Zelinkiella synaptae*, is found on the surface or in the body cavity of sea cucumbers, and many species of *Embata* and *Proales* are found on the gills and legs of crustaceans and cladocerans. The only three described species of Seisonid rotifers are found on the gills and legs of the marine crustacean Nebalia; one of these, *Paraseison annulatus*, is parasitic, feeding on the body fluids and eggs of the host, while the other two feed on bacteria.

Rotifers move about by crawling in a leechlike manner (common in many bdelloids) or by swimming at speeds up to 0.5 mm/s using the water currents created by their coronal cilia. Many species in the monogonont groups Flosculariidae and Collothecidae are sessile: after hatching they rapidly settle on a substrate and produce a gelatinous tube, followed by extensive changes in the morphology of the corona (often incorrectly described as metamorphosis).

Reproduction and development. The three major groups of rotifers each have a characteristic form of reproduction (**Fig. 3**). Seisonids reproduce only sexually: males are fully developed and only slightly smaller than females. The monogonont life cycle is a complex combination of asexual and sexual forms of reproduction: Monogononts generally reproduce by parthenogenesis (females produce diploid eggs by mitosis that develop, without fertilization, into females) but in response to certain environmental cues, such as crowding, females are produced that produce haploid eggs by meiosis. If unfertilized, these eggs develop into haploid males (generally much small than females and often with a highly reduced or absent digestive tract). If a male

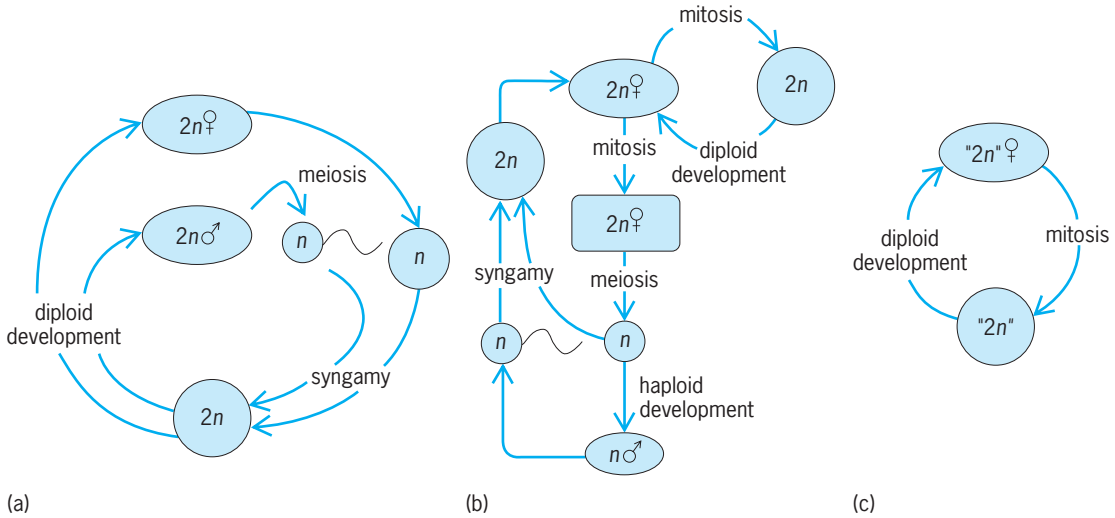

Fig. 3. Life cycles of the three rotifer groups. (*a*) Gonochoristic reproduction in seisonids. (*b*) Parthenogenesis alternating with sexual reproduction in monogononts. (*c*) Obligate parthenogenesis in bdelloids. 2*n* represents diploidy, and *n*, haploidy; bdelloids are designated "2*n*" because the concept of diploidy loses meaning in the absence of meiosis.

mates with a female and fertilizes a haploid egg, the female pumps additional nutrients and cytoplasm into the developing diploid embryo, which develops a thick shell and is called a resting egg. The resting egg can remain dormant (diapause) for years until, in response to environmental cues, it hatches into a female that will reproduce parthenogenetically. Bdelloid rotifers reproduce parthogenetically; no males or hermaphrodites are known and cytological and molecular evidence suggest that the group has evolved without sexual reproduction over many millions of years. Monogononts and Bdelloids represent two of the most successful examples of facultative and obligate parthenogenesis.

Seisonids are oviparous; monogononts and bdelloids are oviparous or viviparous depending on species. Embryonic development proceeds through a modified spiral cleavage and is eutelic; juveniles develop directly into adults. Fecundity and lifespan are highly variable between species: monogonont and bdelloid females possess approximately 10–50 oocytes and live for approximately 5–45 days.

Ecology. Rotifers are among the most widely distributed of aquatic animals. Monogononts are the most common nonarthropods in most benthic and pelagic fresh- and brackish water environments, and are a food source for copepods, insect larvae, and fish fry. Bdelloids are extremely widely distributed in all forms of fresh water, from the littoral zones of large permanent lakes to temporary pools, mud holes, water films on mosses or lichens, and similar ephemerally moist environments. They are often the most commonly encountered invertebrates in these habitats. Bdelloids can be found in hot springs, withstanding temperatures as high as 45°C (113°F) in Antarctic lakes where they can survive being frozen, and in highly acidic or alkaline streams. *See* MARINE ECOLOGY.

Most bdelloid species are capable of surviving extreme desiccation by a process known as anhydrobiosis. As water slowly evaporates from their surroundings, bdelloids contract into tight oblong balls and exude nearly all free water. In this state bdelloids are metabolically quiescent and can survive vacuum, large doses of ionizing radiation, and temperatures of −270 to 150°C (−454° to −320°F). A bdelloid can remain desiccated for years and, upon rehydration, resumes its normal lifespan with little or no loss in fecundity. *See* ACANTHOCEPHALA; BDELLOIDEA; MONOGONONTA; SEISONACEA. J. B. Jennings; David B. Mark Welch

Bibliography. L. H. Hymen, *The Invertebrates*, vol 3, 1951; Th. Nogrady (ed.), *Guides to the Identification of the Microinvertebrates of the Continental Waters of the World: Rotifera*, 1993; S. D. Rundle, A. L. Robertson, and J. M. Schmid-Araya, *Freshwater Meiofauna: Biology and Ecology*, 2002; D. G. Smith, *Pennak's Freshwater Invertebrates of the United States: Porifera to Crustacea*, 4th ed, 2001; J. H. Thorp and A. P. Covich (eds.), *Ecology and Classification of North American Freshwater Invertebrates*, 2nd ed., 2001.

Rous sarcoma

The first filterable agent (virus) known to cause a solid tumor in chickens. It was discovered in 1911 by P. F. Rous, who won the Nobel prize in 1967 for his discovery. Little advance was made in the study of this virus until technological advances made more detailed studies possible. Certain strains of the virus cause tumors in hamsters, rabbits, monkeys, and other species, thus effectively eliminating an age-old notion that tumor viruses are more or less species-specific. Although the virus cannot be demonstrated in the tumors in foreign species, transfer of the tumors back to the chick results in reappearance of the active virus. Tissue culture of chick embryos inoculated with Rous sarcoma virus produce foci of

proliferating cells composed of minute aggregations of tumor cells. It was found that the Rous virus by itself was incapable of bringing about this change but needed another closely related virus of the avian leukosis group to act as a "helper" for the production of the foci. Thus the Rous virus is known as a "defective" virus in that it is incapable of producing tumors by itself. *See* AVIAN LEUKOSIS; TISSUE CULTURE.

It is a ribonucleic acid virus and belongs to the avian leukosis group. When inoculated into chick embryos, it causes fatal hemorrhagic disease, and produces sarcomas in older chickens, pheasants, and ducklings. *See* ANIMAL VIRUS; TUMOR VIRUSES; VIRUS, DEFECTIVE. Alice E. Moore

Bibliography. J. W. Beard (ed.), *International Conference on Avian Tumor Viruses*, Nat. Cancer Inst. Monogr. 17, 1964.

Rubber

Originally, a natural or tree rubber, which is a hydrocarbon polymer of isoprene units. With the development of synthetic rubbers having some rubbery characteristics but differing in chemical structure as well as properties, a more general designation was needed to cover both natural and synthetic rubbers. The term elastomer, a contraction of the words elastic and polymer, was introduced, and defined as a substance that can be stretched at room temperature to at least twice its original length and, after having been stretched and the stress removed, returns with force to approximately its original length in a short time.

Three requirements must be met for rubbery properties to be present in both natural and synthetic rubbers: long threadlike molecules, flexibility in the molecular chain to allow flexing and coiling, and some mechanical or chemical bonds between molecules.

A useful way of visualizing rubber structure is to consider as a model a bundle of wiggling snakes in constant motion. When the bundle is stretched and released, it tends to return to its original condition. If there were no entanglements, stretching the bundle would pull it apart. The more entanglements, the greater the tendency to recover, corresponding to cross-links in rubber. In rubbers, the characteristic property of reversible extensibility results from the randomly coiled arrangement of the long polymer chains. Upon extension the chains are elongated in a more or less orderly array. The tendency to revert to the original coiled disarray upon removal of the stress accounts for the elastic behavior. Vulcanizing rubber increases the number of cross-links and improves the properties.

Natural rubber and most synthetic rubbers are also commercially available in the form of latex, a colloidal suspension of polymers in an aqueous medium. Natural rubber comes from trees in this form; many synthetic rubbers are polymerized in this form; some other solid polymers can be dispersed in water. *See* POLYMER.

Latexes are the basis for a technology and production methods completely different from the conventional methods used with solid rubbers.

Processing

In the crude state, natural and synthetic rubbers possess certain physical properties which must be modified to obtain useful end products. The raw or unmodified forms are weak and adhesive. They lose their elasticity with use, change markedly in physical properties with temperature, and are degraded by air and sunlight. Consequently, it is necessary to transform the crude rubbers by compounding and vulcanization procedures into products which can better fulfill a specific function.

Curatives and vulcanization. After the addition of curing or vulcanizing agents to rubber, the application of heat causes cross-linking, yielding a durable product by binding the long chains together. Sulfur, the first successful curing ingredient, is the basis of vulcanization. However, various chemicals or combinations of chemicals are also capable of vulcanizing rubber. These include oxidizing agents, such as selenium, tellurium, organic peroxides, and nitro compounds, and also generators of free radicals, such as organic peroxides, azo compounds, and certain organic sulfur compounds such as the alkyl thiuram disulfides.

Sulfur alone results in very slow cure rates and develops less than optimum physical properties in the rubber. Several classes of compounds have been found which accelerate the rate of vulcanization, improve the efficiency of the vulcanization reaction, and reduce the sulfur requirements, in general enhancing the physical properties of the vulcanized rubber.

Accelerators are substances which act as catalysts in the vulcanization process, initiating free radicals. The first accelerators to be used in the rubber industry were inorganic chemicals, such as the basic carbonates and oxides of lead supplemented by magnesium or lime. Since the introduction of the first organic accelerator in 1906, several thousand have been patented. Only a small number of these are in general use. Most accelerators can be grouped into the following general classifications: aldehyde-amines, guanidines, thiuram sulfides, thiazoles, thiazolines, dithiocarbamates, and mercaptoimidazolines. *See* FREE RADICAL.

Some rubbers can be vulcanized by gamma radiation, but it is not an economically feasible process. For example, butyl rubber cannot be vulcanized by irradiation because degradation rather than cross-linkage occurs. The structure of the chain molecule is the determining factor. Cross-linking predominates if the polymer has the structure $(—CH_2CH_2—)_n$ or $(—CH_2CHR—)_n$. If the structure is $(—CH_2CR_1R_2—)_n$, degradation is the rule. There is great variation in the effects of gamma-radiation vulcanization, depending on the type of rubber being treated, the nature of the compounding ingredients, and the process conditions.

Although the vulcanization process was discovered by Charles Goodyear in 1839, an understanding of the mechanism of this solid-state reaction is still incomplete. There are formations of loose ends, rings, and other structures, but the main vulcanization reaction is a cross-linking of the long-chain polymer molecules by sulfur bridges or covalently bonded carbon-carbon linkages to form an extensive three-dimensional structure. This network structure modifies the essentially thermoplastic properties of the raw rubber and confers predominantly elastic properties on the resulting vulcanizate by preventing slippage of the long-chain molecules past each other.

Pigments. Vulcanization improves the elasticity and aging properties of rubber, but in most cases it is necessary to enhance further such properties as tensile strength, abrasion resistance, and tear resistance by incorporation of fillers. Those which improve specific physical properties are known as reinforcing fillers; those which serve primarily as diluents are classed as inert fillers. The physical properties of the resulting vulcanizates are affected by both the type and amount of filler.

The most universally used reinforcing filler in the rubber industry is carbon black. The types used commercially in the greatest bulk for this purpose include furnace and thermal blacks. Each may be further classified according to particle size and surface structure, and then selected according to specific properties which are required in the end product. *See* CARBON BLACK.

In addition to the carbon blacks, inorganic reinforcing agents, such as zinc oxide and the silicas, are used for reinforcement of light-colored end products. Although zinc oxide is not used extensively in the rubber industry, its incorporation does enable the resulting product to withstand extended exposure at high temperatures, and it also functions as an activator during the vulcanization process. The silicas are used in those products in which high abrasion resistance is an essential requirement.

The inert fillers, such as whiting and various types of clays, serve merely as extenders and to facilitate the processing of the compound. In general, they do not improve the tensile strength of the resulting vulcanizate, but do increase hardness and modulus.

Protective agents. The aging stability of rubber compounds is influenced by such factors as heat, light, and atmospheric conditions. The chemical unsaturation (presence of —C═C— groups) of natural and synthetic rubber compounds provides a focal point for oxidation. The initial oxidative degradation involves hydroperoxide formation and scission of polymer chains. These reactions are accompanied by network-forming or cross-linking reactions which occur to an extent, depending on the process conditions and the composition of the polymer. Chain scission is dominant in natural rubber, which becomes soft and sticky and loses tensile strength on degrading. In SBR (styrene-butadiene rubber), the predominant reaction is cross-linking, and so it becomes hard and brittle when attacked by oxygen. With either chain scission or cross-linking, the intermediates are free radicals which set off many repetitive reaction cycles. Some synthetic rubbers such as butyl have low unsaturation and so are resistant to oxygen and ozone.

While vulcanization improves resistance to attack by oxygen, further improvements have been accomplished by the incorporation of chemical compounds, known as antioxidants and antiozonants, into the rubber compound. Broadly speaking, antioxidants may be classified into two general groups: aromatic amines and phenols. The choice of the antioxidant is governed by the rubber and the purpose for which the rubber is intended. Antioxidants function by combining with one or more of the free radicals which arise and thus prevent further degradation of the rubber chains. Amine antioxidants have staining characteristics which do not interfere with their use in rubber compounds for tires, hose, and belts. However, light-colored end products or rubber compounds which come in contact with light-colored surfaces require nonstaining antioxidants. These compounds are usually highly hindered phenols, obtained by alkylation of phenols or cresols, or are derivatives of aromatic phosphite esters. Unfortunately, the most effective anoxidants usually tend to stain or discolor the vulcanizate. *See* ANTIOXIDANT.

The cracking of rubber compounds under stress when exposed to very low ozone concentrations, such as exist in the atmosphere, also originates from the unsaturated chemical structure of the rubber. The deteriorating effects of ozone can be diminished by the incorporation of various types of insoluble wax which bloom to the surface to form a physical barrier. If rubber is deformed, however, the protective surface is broken and is useless in many applications such as tires.

Certain accelerators have antiozonant properties, but the majority of commercial antiozonants are substituted *p*-phenylenediamines. These can be subdivided into the diaryl-, aryl-alkyl-, and dialkyl-*p*-phenylenediamines. Another type is formed by condensation of amines and ketones. The mechanism of antiozonant action has not definitely been established, but there is evidence to indicate that it is different from the reaction mechanism of antioxidants. The protective effect is believed to be the result of preferential combination of the ozone with the antiozonant instead of the rubber. Two parameters are believed to define ozone cracking behavior: the critical energy necessary to rupture the rubber surface, and the rate of crack growth. A large number of antiozonants decrease the rate of crack growth but only the *N*,*N*′-dialkyl-*p*-phenylenediamines increase the critical energy needed for the crack growth to occur.

Photooxidation of rubber compounds, resulting from exposure to bright sunlight, causes a chemical reaction which gives the surface a resinous or crazed appearance. This deleterious effect, caused mainly by ultraviolet light, is not reduced to any appreciable extent by antioxidants; indeed, some amine antioxidants are known to accelerate the degradation of natural rubber. Although complete protection against photooxidation is difficult, stocks containing

carbon black are sufficiently resistant for most applications. Ultraviolet-absorbing materials, such as salicylates and benzotriazoles, provide a certain degree of protection for light-colored products.

Gamma radiation causes degradation by initiating free-radical cross-linking and chain-scission processes in the rubber. The net change in physical properties depends on the extent of each process, which in turn is dependent on the molecular structure of the rubber, radiation dosage, and exposure conditions. Since the changes induced by radiation are the result of free-radical mechanism, the type of compounding can affect the results.

The degradative effects of radiation exposure may be reduced by incorporation of protective agents known as antirads, whose efficiency does not correlate with their behavior as antioxidants or antiozonants. Test results have shown that the following compounds impart radiation stability to various rubbers: *N*,*N*′-cyclohexylphenyl-*p*-phenylenediamine, a mixture of di-phenyl-*p*-phenylenediamine and phenyl-α-naphthylamine, quinhydrone, diphenylamine, and *p*-methoxyphenol. These represent just a few of the numerous compounds which can function as radiation inhibitors.

Physical testing. The physical testing of rubber compounds plays an essential role in rubber technology. Because available rubbers exhibit a wide range of properties which can be further varied by compounding techniques, standard methods of evaluation based upon physical properties have been established. Physical tests screen, measure, and evaluate the desired physical properties of a product, according to the information requested. In the rubber industry, tests are conducted on natural and synthetic rubbers, rubberlike materials, plastics, fabrics, and cords. What seem like trivial changes in the formulation, compounding, or processing of a material such as rubber are often quite discernible in the physical testing data. Frequent physical testing helps to ensure a standard product.

Commercial articles are often so complicated in structure that only actual service tests will suffice to prove their utility. In the most notable instance, new tire designs are introduced only after literally millions of miles of road testing.

Standards of the American Society for Testing and Materials (ASTM) are used to test materials whenever possible. This standardization of testing procedures simplifies the demands of the manufacturer and the compounder, and guarantees universal, comparable data.

In screening for physical properties, the first step is the evaluation of unvulcanized compounds. The raw polymers may be evaluated by techniques appropriate to any scientific investigation. These include chemical analyses and examination by infrared spectroscopy, nuclear magnetic resonance, pyrolysis, gel permeation chromatography, thermal analysis, and others. Polymer molecular weight may be estimated by a viscosity or relative plasticity measurement such as that done by the Mooney plastometer. Other instruments such as the Instron extrusion rheometer measure plasticity at different shear rates.

In vulcanized compounds, stress-strain properties, that is, the relation of the force required to the sample elongation, are of importance. The elastic modulus, the ratio of stress to strain at low elongations, and the ultimate elongation to sample rupture are usually reported. Stress-strain properties are evaluated by the use of a constant-power-driven machine such as the Scott tensile tester. With the advancement in electronics, the Instron tester and similar apparatus employing load cells have come into practical use for obtaining more detailed information regarding such physical testing. *See* STRESS AND STRAIN.

All physical testing is conducted at a standard temperature of 25°C (77°F) and 50% relative humidity, except where higher or lower temperature testing is requested. Accelerated aging and conditioning of materials prior to testing is conducted in specially designed and automatically controlled ovens and refrigerators. Physical tests may be requested for properties such as hardness or softness, tear resistance, abrasion, flex, resistance to ozone, ultraviolet, salt spray and water, compression and recovery, permeability to gases, volatility, rebound, elasticity, dynamic modulus and resilience, and low-temperature properties. Each specific test employs a specifically designed apparatus for measuring the physical property or properties desired.

An example of a modern piece of equipment is the Gehman torsion apparatus. It has automatic controlling devices whereby measurements can be made at any desired temperature from ambient to −320°F (−195.5°C) by the use of liquid gases such as nitrogen. The Gehman torsion test determines the change in twist modulus with temperature and may be used to estimate the glass transition temperature of the rubber. This is the temperature at which there is enough free volume in the rubber that the polymer chain segments may slip or rotate past one another. This temperature is of fundamental importance since many rubbery properties depend on how far above it tests are conducted. Extensive research has attempted to uncover the complicated relation between many different physical properties and the chemical structure of a polymer. Rubber testing involves so many factors and variables that computer techniques have been developed to cope with their complexity. *See* INSPECTION AND TESTING.

Specific Types of Rubber

One rubber is obtained from natural sources in commercial quantities. In addition, a number of synthetic compounds are classified as rubbers.

Smoked sheet and pale crepe represent the forms in which the major portion of natural rubber is commercially available. The smoked sheet is obtained by acid coagulation of the tree latex, sheeting the coagulum, and then drying and smoking the sheets of rubber. The chemicals in the smoke preserve the rubber from molds and other organisms. Pale crepe is obtained by treating the latex, before or after coagulation, with sodium bisulfite and then washing, drying, and sheeting. Synthetic rubbers are available

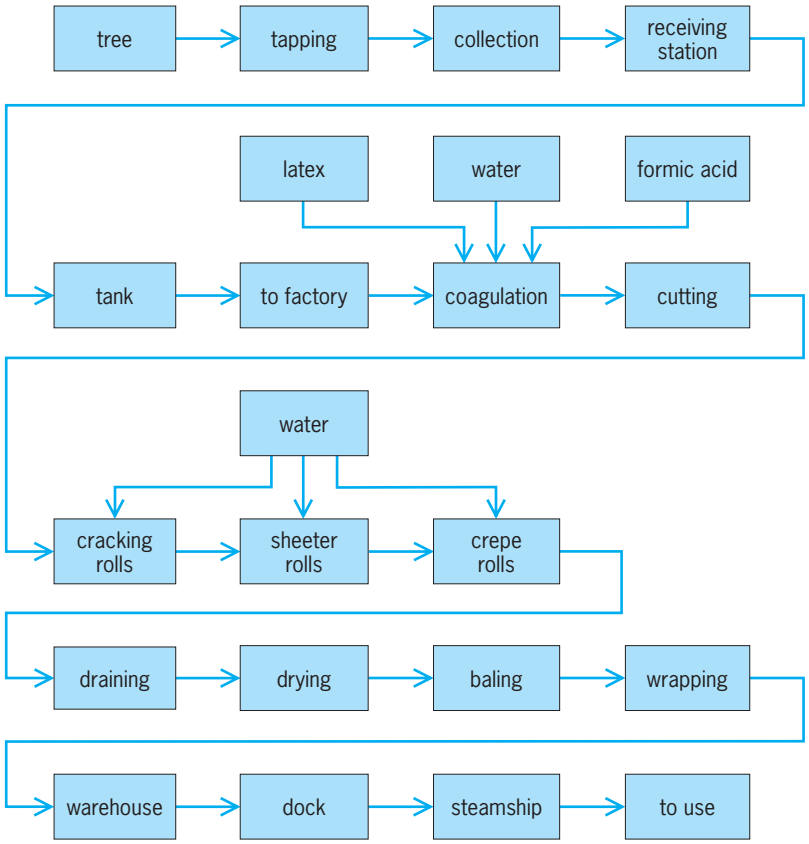

Fig. 1. Steps in natural-rubber production.

as bales, pellets, or flakes made from the polymerization process.

Natural rubber (NR). Although natural rubber may be obtained from hundreds of different plant species, the most important source is the rubber tree (*Hevea brasiliensis*). Natural rubber is *cis*-1,4-polyisoprene (**1**), containing approximately 5000 isoprene units (**2**) in the average polymer chain. World annual production of natural rubber increased from 50,500 tons (45,900 metric tons) in 1900 to over 3.9×10^6 tons (3.5×10^6 metric tons) in 1980. *See* RUBBER TREE.

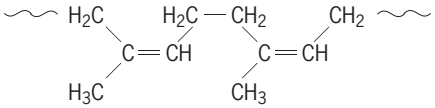

(1)

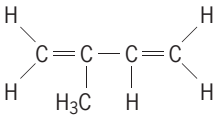

(2)

The economic competition from synthetic rubbers has stimulated research and development in natural rubber by increasing productivity in the field, improving uniformity and quality of the product and packaging, and developing modified natural rubbers with specific properties.

Increased productivity has been achieved by increasing the yield of the trees by cross-pollination of high-rubber-yielding clones of *H. brasiliensis*, use of chemical stimulants, and better tapping and collection methods. These methods have resulted in large increases in yields on Malaysian estates, with much greater increases anticipated.

With improved productivity and better processing methods and controls, much better quality and uniformity have been reached. This has made possible the development of standard Malaysian rubbers (SMR) to meet specifications on a number of properties, including dirt and ash content, viscosity, and copper and manganese content.

A number of modified forms of natural rubber have been developed, including the following: superior processing (SP), oil extended (OENR), special tire rubber, deproteinized (DPNR), pellet form, filler master batches, and chemically modified.

In addition, experimental work has been done on improved collection of latex by micro tapping and collecting in polyethylene bags, and trees having three parts: a high-yielding trunk system grafted onto a strong root system and an improved, more prolific leaf system grafted onto the trunk.

Despite competition from synthetic rubbers and a great reduction in percentage use worldwide of natural rubber, the tonnage of natural rubber continues a steady growth due to the growth of the industry. Steps in the production of natural rubber are shown in **Fig. 1**.

During World War II, considerable development work was done on cultivation of the Guayule shrub as a source of rubber, and some excellent rubber was produced experimentally from this plant in the United States but the process was not economically feasible. Improved cultivation and processing of Guayule coupled with the sharply increased cost of natural and synthetic rubber revived interest in this type of rubber. *See* GUAYULE.

Styrene-butadiene rubbers (SBR). The extensive development of the synthetic rubber industry originated with the World War II emergency, but continued expansion has been the result of the superiority of the various synthetic rubbers in certain properties and applications. The most important synthetic rubbers and the most widely used rubbers in the entire world are the styrene-butadiene types.

Formerly designated GR-S, SBRs are obtained by the emulsion polymerization of butadiene and styrene in varying ratios. However, in the most commonly used type, shown in reaction (1), the ratio

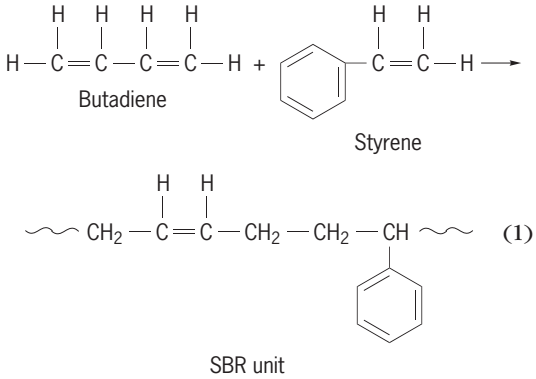

(1)

of butadiene to styrene is approximately 78:22. The polymerization is conducted in water emulsion also containing soap, a molecular weight regulator, and an initiator which generates the free radicals necessary to promote the polymer chain growth. At 122°F (50°C) conversion to polymer occurs at about 5–6%/h. The reaction is terminated at 70–75% conversion, since higher conversions lead to cross-linked polymer formation in the emulsion particle. The termination is effected by adding a so-called shortstop such as hydroquinone (about 0.1%), which reacts rapidly with free radicals and oxidizing agents. This destroys the initiator and prevents formation of any new polymer chains. Antioxidant is then added, the remaining butadiene and styrene are removed, and the latex is coagulated by adding salt and dilute sulfuric acid or aluminum sulfate. The coagulated crumb is then washed, dried, and baled for shipment. This general procedure is basic for production of all emulsion polymers. *See* POLYMERIZATION.

The SBR formed is a copolymer of styrene and butadiene units linked randomly in long chains. The butadiene units enter the polymer chain randomly in three different conformations known as *cis*-1,4; *trans*-1,4; and 1,2 or vinyl. The original or hot SBR is polymerized at a temperature of 122°F (50°C) or higher and is used primarily for mechanical goods because of its better color retention. The cold type, polymerized at 41°F (5°C), can be made to a higher molecular weight and therefore exhibits an improvement over the hot types in most physical properties. When used in light-duty tires such as passenger car tires, the cold-rubber compounds have proved equal or superior to natural rubber treads. However, they are inferior to natural rubber for truck tires because of the greater heat buildup during flexing. The cold, oil-extended type, prepared by the replacement of a portion of the polymer by a heavy-fraction oil, accounts for more than 50% of the cold-rubber production. Its advantage is primarily economic. Master batches, containing both oil and carbon black, have the advantage of process simplification.

In general, the compounding and processing methods for all the SBR types are similar to those of natural rubber. Although natural rubber is superior with respect to lower heat buildup, resilience, and hot tear strength, the SBR types are more resistant to abrasion and weathering. However, carbon black or some other reinforcing fillers must be added to the SBR to develop the best physical properties. Unlike natural rubber, SBR does not crystallize on stretching and thus has low tensile strength unless reinforced. The major use for SBR is in tires and tire products. Other uses include belting, hose, wire and cable coatings, flooring, shoe products, sponge, insulation, and molded goods.

cis-1,4-Polybutadiene (BR). Work on the duplication of natural rubber stimulated interest in stereoregulated polymerization of butadiene, particularly of the high *cis*-1,4 structure. Structure (3) shows the *cis*-1,4-polybutadiene units in the polymer chain. Several different catalyst systems were developed to give polybutadiene rubbers of varying *cis*-1,4 content. A

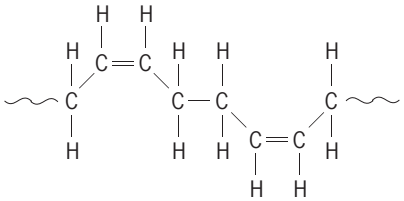

(3)

cobalt-containing catalyst, for example, will result in a polymer of 98% *cis*-1,4; a catalyst of an aluminum alkyl and titanium tetraiodide, 90–95% *cis*-1,4; and a catalyst of an alkyl lithium, 40–50% *cis*-1,4. All these systems employ air-free, moisture-free polymerization in hydrocarbon solvents.

Black-loaded vulcanizates of *cis*-1,4-polybutadiene rubbers exhibit good physical properties, such as lower heat generation, higher resilience, improved low-temperature properties, and greatly improved abrasion resistance. Processing properties are rather poor, but can be greatly improved by employing blends with natural rubber or SBR. A 1:1 blend of *cis*-1,4-polybutadiene and natural rubber displayed hysteresis properties equivalent to the natural rubber control and exhibited satisfactory modulus, tensile strength, and hardness. Tests on retreaded passenger tires gave outstanding abrasion resistance and increased resistance to cracking as compared to natural rubber. In passenger tires, *cis*-1,4-polybutadiene improves treadwear by about 1% for each percent of polybutadiene in the tread compound. In truck tires, a blend with natural rubber gives 14% more wear than natural rubber alone.

Butyl rubber. Isobutylene and isoprene or butadiene obtained from cracked refinery gases are the primary raw materials required for the manufacture of butyl rubber, as shown in reaction (2). Butyl is es-

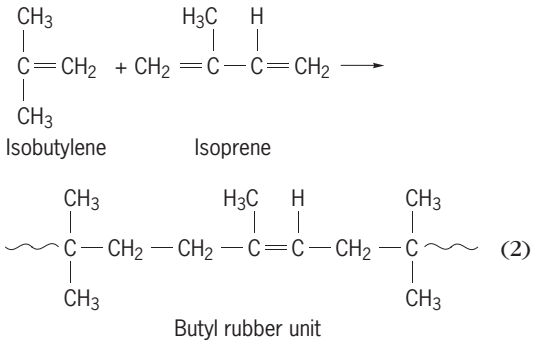

(2)

sentially a polyisobutylene except for the presence of 0.5–4.5% diolefin, usually isoprene, to provide the unsaturation necessary for vulcanization. *See* PETROCHEMICAL.

A solution polymerization system is used, the reaction being conducted at approximately −150°F (−101°C) in methyl chloride with aluminum chloride ($AlCl_3$) and a trace of water as the cationic catalyst. The low temperature is necessary to obtain a high-molecular-weight product and to control the almost explosive reaction rate. Butyl rubbers are processed according to conventional methods used for natural rubber and SBR. An antioxidant, sulfur, and an ultraaccelerator are compounded into the stock.

However, because of the low degree of unsaturation in the butyl rubber molecule, it is more difficult to complete the cure and, consequently, the vulcanization requires higher temperatures and longer times. Also because of the very low unsaturation, small amounts of unsaturated materials such as other rubbers, or plasticizers will interfere with proper vulcanization, and contamination by such materials during processing must be avoided.

Neither neoprene nor butyl rubber requires carbon black to increase its tensile strength, but the reinforcement of butyl rubber by carbon black or other fillers does improve the modulus and increases the resistance to tear and abrasion.

The excellent resistance of butyl rubbers to oxygen, ozone, and weathering can be attributed to the smaller amount of unsaturation present in the polymer molecule. In addition, these rubbers exhibit good electrical properties and high impermeability to gases. The high impermeability to gases results in use of butyl as an inner liner in tubeless tires. Other widespread uses are for wire and cable products, injection-molded and extruded products, hose, gaskets, and sealants, and where good damping characteristics are needed.

Chlorobutyl rubber, which contains about 1% chlorine, is more readily mixed and cured than butyl. It exhibits improved adhesion to, and compatibility with, natural rubber and the synthetic rubbers. Brominated butyl (containing 1–3.5% bromine) is improved similarly. Vulcanization of butyl with modified phenol resins, that is, resin-cured butyl, greatly increases the useful life of articles such as curing bladders, which are exposed to heat for long periods of time.

Ethylene-propylene polymers (EPM, EPDM). Stereospecific catalysts are employed to make synthetic rubbers by the copolymerization of ethylene and propylene. Either monomer alone polymerizes to a hard, crystallizable plastic, but copolymers containing 35–65% of either monomer are amorphous, rubbery solids. Special catalysts must be employed because ethylene polymerizes many times faster than propylene. Best results seem to be obtained with complex catalysts derived from an aluminum alkyl and a vanadium chloride or oxychloride.

Since these copolymers are free of double bonds, they exhibit outstanding resistance to heat, oxygen, ozone, and other aging and degrading agents. Abrasion resistance in tire treads is excellent. Special curative systems, such as organic peroxides, must be employed because of the absence of double bonds. Processing techniques and factory equipment used with other rubbers can also be applied to these copolymers. The mechanical properties of their vulcanizates are generally approximately equivalent to those of SBR.

Terpolymers containing ethylene, propylene, and a third monomer, such as dicyclopentadiene, have become more popular because they contain unsaturation and thus may be sulfur-cured by using more or less conventional curing systems. Because the unsaturation is in the side chains rather than the main molecular chain, it does not appreciably impair the inherent chemical resistance.

Neoprene (CR). One of the first synthetic rubbers used commercially in the rubber industry, neoprene is a polymer of chloroprene, 2-chlorobutadiene-1,3. In the manufacturing process, acetylene, the basic raw material, is dimerized to vinylacetylene and then hydrochlorinated to the chloroprene monomer, as shown in reaction (3). An emulsion system is used

$$\underset{\text{Acetylene}}{2(CH{\equiv}CH)} \xrightarrow[CuCl_2]{} \underset{\text{Vinylacetylene}}{CH_2{=}CH{-}C{\equiv}CH} \xrightarrow[CuCl_2]{HCl} \underset{\text{Chloroprene}}{CH_2{=}CCl{-}CH{=}CH_2} \qquad (3)$$

$$\underset{\text{Neoprene unit}}{\sim\!\sim CH_2{-}\overset{\overset{Cl}{|}}{C}{=}CH{-}CH_2\sim\!\sim}$$

for the polymerization, and the resulting polymer is isolated by freeze coagulation. The products are about 85% *trans*-1,4 in structure and about 100,000–180,000 in molecular weight. Several grades are available, differing because of varying polymerization conditions. Copolymer types of neoprene, in which the chloroprene monomer predominates [reaction (3)], are also available.

Sulfur is used to vulcanize some types of neoprene, but most of the neoprenes are vulcanized by the addition of basic oxides such as magnesium oxide and zinc oxide. The cure proceeds through reaction of the metal oxide with the tertiary allylic chlorine that arises from the small amount of 1,2 polymerization that occurs. Other compounding and processing techniques follow similar procedures and use the same equipment as for natural rubber. One of the outstanding characteristics of neoprene is the good tensile strength without the addition of carbon black filler. However, carbon black and other fillers can be used when reinforcement is required for specific end-use applications that require increased tear and abrasion resistance.

The neoprenes have exceptional resistance to weather, sun, ozone, and abrasion. They are good in resilience, gas impermeability, and resistance to heat, oil, and flame. They are fairly good in low temperature and electrical properties. This versatility makes them useful in many applications requiring oil, weather, abrasion, or electrical resistance or combinations of these properties, such as wire and cable, hose, belts, molded and extruded goods, soles and heels, and adhesives.

Nitrile rubber (NBR). Much of the basic pioneering research on emulsion polymerization systems was with nitrile-type rubbers. These rubbers, first commercialized as the German Buna N types in 1930, are copolymers of acrylonitrile and a diene, usually butadiene [reaction (4)]. The ratios of the monomers can vary from 80:20 to 55:45 butadiene/acrylonitrile. The oil-resistant properties of the rubber increase with increasing nitrile content, but there is a corresponding decrease of the low-temperature flexibility

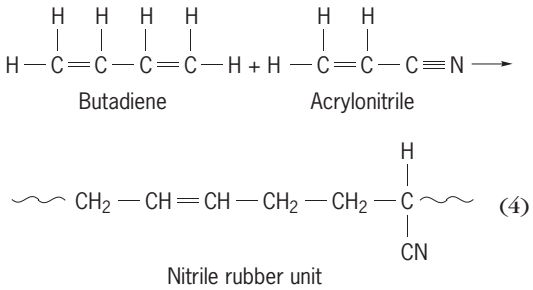

(4)

of the rubber because the glass transition temperature increases with nitrile content. Both sulfur and nonsulfur vulcanizing agents may be used to cure these rubbers. Carbon black or other reinforcing agents are necessary to obtain the optimum properties. If proper processing methods are followed, the nitrile rubbers can be blended with natural rubber, polysulfide rubbers, and various resins to provide characteristics such as increased tensile strength, better solvent resistance, and improved weathering resistance.

The nitrile rubbers have outstanding oil, grease, and solvent resistance. Consequently, the commercial usage of these rubbers is largely for items in which these properties are essential. Another major usage is the utilization of the latex form for adhesives and for the finishing of leather, impregnation of paper, and the manufacture of nonwoven fabrics.

cis-1,4-Polyisoprene (IR). In 1954, synthetic *cis*-1,4-polyisoprene was made from isoprene with two different classes of catalysts. The first class includes lithium and the lithium alkyls. The second class uses a mixture of an aluminum alkyl and titanium tetrachloride, the system first used for the low-pressure polymerization of ethylene by Karl Ziegler. Both catalyst system polymerizations are carried out in hydrocarbon solution and require highly purified monomer and solvent. Traces of air, moisture, and most polar compounds adversely affect reaction rates, polymer properties, and structure.

The *cis*-1,4 polymer structure obtained with these catalysts is also characteristic of natural *Hevea* rubber. The presence of high *cis* content appears necessary for the desirable physical properties with *Hevea* in contrast to the inferior properties of emulsion-type polyisoprene, which contains mixed *cis*-, *trans*-, 1,2-, and 3,4-isoprene units [structures (4)–(7)] in the

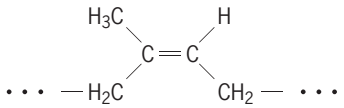

(4)

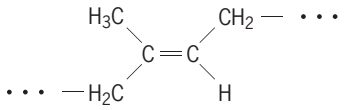

(5)

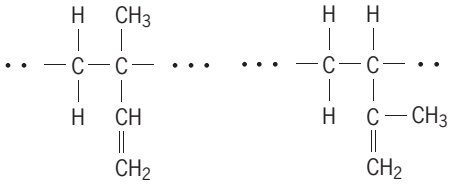

(6) (7)

polymer chain. When polymerization is complete, the catalyst is destroyed, for example, by treating with alcohol, and the solvent removed. The solid polymer thus obtained may be handled and compounded as any other dry rubber is.

This polymer and the corresponding butadiene polymer discussed above are called stereorubbers because of their preparation with stereospecific catalysts. The emulsion polymers are formed by a free-radical mechanism that does not permit control of the molecular structure; stereorubbers are formed by anionic mechanisms that permit nearly complete control of the structure of the growing polymer chain in stereoregular fashion. *See* STEREOSPECIFIC CATALYST.

The type of catalyst employed influences the structure of the polymer. The aluminum-titanium-catalyzed rubbers exhibit, on stretching, a gradual crystallization somewhat slower than natural, and they are readily processed since the molecular weight range tends to be less than that of natural rubber. The lithium-catalyzed rubbers contain about 93% *cis*-1,4 structure; they exhibit very little tendency to crystallize and cannot be processed satisfactorily until they have undergone substantial mastication to reduce their very high molecular weight values to a level more nearly resembling that of well masticated natural rubber.

The differences in structure influence the properties of these rubbers. The aluminum-titanium-catalyzed rubber, since it crystallizes more readily, exhibits better hot tensile strength. The lithium-catalyzed polymer, being of higher molecular weight, exhibits higher resilience and less heat generation. Both types, when compounded and cured, produce physical properties which closely approach, or are equivalent to, natural rubber.

Tire tests with heavy-duty tires for trucks, buses, and airplanes have shown that with respect to wear and heat buildup the isoprene rubbers are comparable to natural rubber during tire operation. Polyisoprene rubber has passed qualification tests in high-speed jet aircraft tires to withstand landing speeds as high as 250 mi/h (112 m/s).

Other rubbers. There are other specialty rubbers available which are important because of specific properties, but in the aggregate they make up only approximately 2% of the world production of synthetic rubbers.

Silicone rubbers. Silicone rubber is a linear condensation polymer based on dimethyl siloxane. In the preparation, dimethyl dichlorosilane is hydrolyzed to form dimethyl silanol, which is then condensed to dimethyl siloxane, and this, upon further

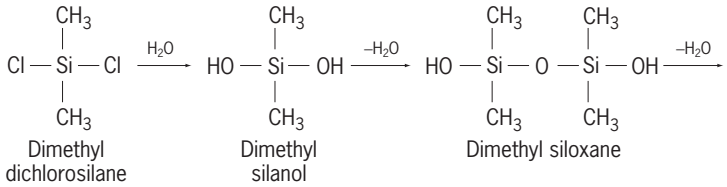

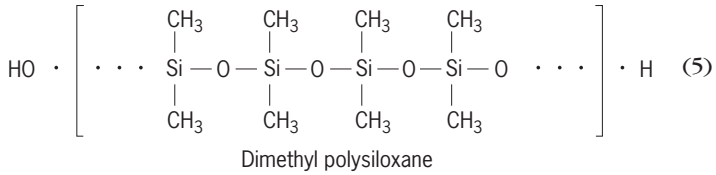

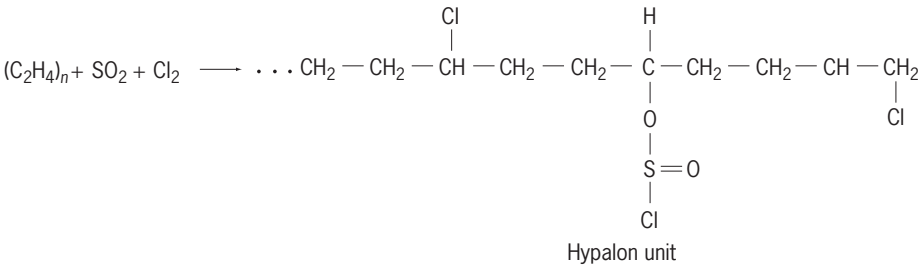

condensation, yields dimethyl polysiloxane, the standard silicone rubber [reaction (5)].

Various types of silicone rubbers are produced by substituting some of the methyl groups in the polymer with other groups such as phenyl or vinyl. Advantages of this type of substitution are evidenced by improvements in specific properties. For example, the presence of phenyl groups in the polymer chain gives further improvement in low-temperature properties. Fluorine-containing side groups improve chemical resistance. Many types are commercially available, ranging from fluid liquids to tough solids.

Because sulfur is not effective for the vulcanization of most silicone rubbers, a strong oxidizing agent, such as benzoyl peroxide, is used; the cross-linking produced is random. However, when unsaturated groups, such as vinyl or allyl, are present, these products are vulcanized with sulfur, the vinyl groups serving as a control of the degree of cross-linking.

Although the standard silicone rubbers are not reinforced by carbon black, the physical properties can be improved by the incorporation of various inorganic fillers such as titania, zinc oxide, iron oxide, and silica, which act as reinforcing and modifying agents. The physical, chemical, and electrical properties can be altered by varying the type and amount of these fillers. Carbon black can be used as a filler with vinyl-containing polymers.

In general, the silicone rubbers have relatively poor physical properties and are difficult to process. However, they are the most stable of rubbers and are capable of remaining flexible over a temperature range of −130 to 600°F (−90 to 316°C). They are unaffected by ozone, are resistant to hot oils, and have excellent electrical properties. Their most extensive uses are for wire and cable insulation, tubing, packings, and gaskets in aerospace and aircraft applications. In the form of dispersions and pastes, they are used for dip-coating, spraying, brushing, and spreading. Silicone rubbers are important in medical and surgical devices because of their property, unique among elastomers, of being compatible with body tissues. Fast, automatic, economical injection molding of liquid silicones has been developed. *See* SILICON; SILICONE RESINS.

Hypalon. Hypalon is the Du Pont trade name for a family of chlorosulfonated polyethylenes prepared by treating polyethylene with a mixture of chlorine and sulfur dioxide, whereby a few scattered chlorine and sulfonyl chloride groups are introduced into the polyethylene chain [reaction (6)]. By this treatment, polyethylene is converted to a rubberlike material in which the undesirable degree of crystallinity is destroyed but other desirable properties of polyethylene are retained. The outstanding chemical stability of Hypalon results from the complete absence of unsaturation in the polymer chain.

Vulcanization is accomplished by means of metallic oxides, such as litharge, magnesia, or red lead, in the presence of an accelerator. The sulfonyl chloride groups provide the sites of reactivity at which the bases react with the chlorine of the sulfonyl chloride group to give cross-links. The incorporation of an organic acid, such as hydrogenated wood rosin or stearic acid, is also necessary for optimum cure. Carbon black or other fillers are not needed to obtain optimum strength properties, and antioxidants are required only for maximum resistance to heat. In addition to its use as a stock in itself, Hypalon can be blended with other types of rubbers to provide a wide range of properties.

Hypalon has extreme resistance to ozone. Its chemical resistance to strong chemicals, such as nitric acid, hydrogen peroxide, and strong bleaching

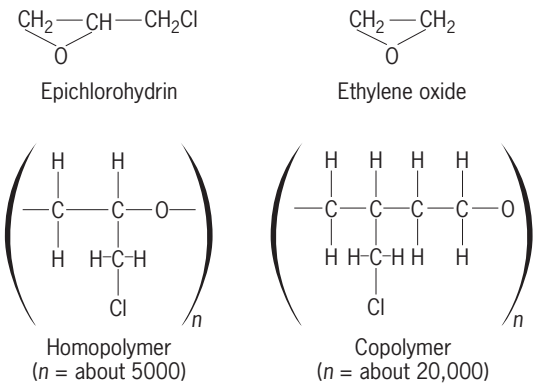

Fig. 2. Two main types of epichlorohydrin elastomers, showing the monomer and the repeating unit.

agents, is superior to any of the commonly used rubbers. These vulcanizates also have good heat resistance, mechanical properties, and unlimited colorability. Typical applications include white sidewall tires and a variety of sealing, waterproofing, insulating, and molded items.

Epichlorhydrin elastomers. These elastomers are polymers of epichlorhydrin. The two main types of epichlorohydrin elastomers are the homopolymer (CO) and the copolymer (ECO) of epichlorohydrin and ethylene oxide (**Fig. 2**). The copolymer has a lower brittle point, better resilience, and a lower specific gravity than the homopolymer, but is not as good in high-temperature properties.

Since the backbone of the molecule is saturated, these elastomers cannot be vulcanized with sulfur. Cross-linking is achieved by reaction of the chloromethyl side group with diamines or thioureas. A metal oxide is also required.

They exhibit outstanding ease of processing; extreme impermeability to gases; moderate tensile strength and elongation; high modulus; good abrasion resistance; good heat aging; excellent resistance to hot oil, water, perchloroethylene, acids, bases, and ozone; good low-temperature properties; and electrical properties ranging between those of a poor insulator and a good conductor, depending on compounding.

These properties make the epichlorhydrin elastomers useful in gaskets for oil field specialties, diaphragms, pump and valve parts; hose for low-temperature flexibility, oil and fuel resistance, or gas impermeability applications, and mechanical goods such as belting, wire, and cable.

Thermoplastic elastomers. Proper choice of catalyst and order of procedure in polymerization have led to development of thermoplastic elastomers. The leading commercial types are styrene block copolymers having a structure which, unlike the random distribution of monomer units in conventional polymers, consists of polystyrene segments or blocks connected by rubbery polymers such as polybutadiene, polyisoprene, or ethylene-butylene polymer. These types can be designated SBS, SIS, and SEBS, respectively. Other types of thermoplastic elastomers, such as the polyester type, have also been developed.

SBS, SIS, and SEBS polymers, when heated to a temperature above the softening point of the styrene blocks, can be processed and shaped like thermoplastics but, when cooled, act like vulcanized rubber, with the styrene blocks serving as cross-links. Thus, it is unnecessary to vulcanize. Also scrap or excess material can be reused.

Thermoplastic elastomers are very useful in providing a fast and economical method of producing a variety of products, including molded goods, toys, sporting goods, and footwear. One of the disadvantages for many applications is the low softening point of the thermoplastic elastomers. Attempts to modify the composition and structure to raise the softening temperature have generally resulted in inferior rubbery properties.

Fluoroelastomers. These are basically copolymers of vinylidene fluoride and hexafluoropropylene. Because of their fluorine content, they are the most chemically resistant of the elastomers and also have good properties under extremes of temperature conditions. They are useful in the aircraft, automotive, and industrial areas.

Polyurethane elastomers. Polyurethane elastomers are of interest because of their versatility and variety of properties and uses. They can be used as liquids or solids in a number of manufacturing methods. The largest use has been for making foam for upholstery and bedding, but difficulties have been encountered in public transportation due to flammability. The basic repeating structure (**8**) is shown for

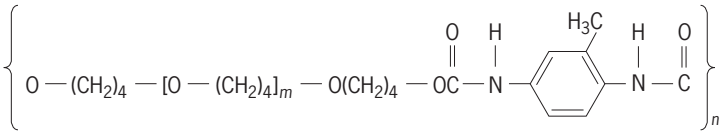

(8)

urethane rubber. *See* POLYURETHANE RESINS.

Polysulfide rubbers. These rubbers have a large amount of sulfur in the main polymer chain and are therefore very chemically resistant, particularly to oils and solvents. They are used in such applications as putties, caulks, and hose for paint spray, gasoline, and fuel. *See* POLYSULFIDE RESINS.

Polyacrylate rubbers. Polyacrylate rubbers are useful because of their resistance to oils at high temperatures, including sulfur-bearing extreme-pressure lubricants.

Edward G. Partridge

Products

A wide variety of products have a rubber, or elastomer, as an essential component. Most rubber products contain a significant amount of nonrubber materials, used to impart processing, performance, or cost advantages.

The automotive industry is the biggest consumer of rubber products. About 60% of all rubber used goes into the production of passenger, bus, truck and off-the-road tires. In addition, a typical automobile contains about 150 lb (68 kg) of rubber products such as belts, hose, and cushions. Outside the automotive area, a wide variety of rubber products are

produced, including such familiar items as rubber bands, gloves, and shoe soles.

The formulation and manufacture of rubber products is as diverse as the wide variety of applications for which they are intended. This discussion is limited to the more common methods of manufacture and those which apply to the widest variety of products. In general, the manufacture of rubber products involves compounding, mixing, processing, building or assembly, and vulcanization.

Compounding. The technical process of determining what the formulation components of a rubber product should be is called compounding. This term is also applied to the actual process of weighing out the individual components in preparing for mixing.

Rubber is the backbone of any rubber product. Natural rubber, obtained from the latex of rubber trees, accounts for about 23% of all rubber consumed in the United States. The balance is synthetic rubber. Of the synthetic rubbers, SBR (styrene-butadiene rubber) and polybutadine rubber are most important, accounting for 71% of synthetic rubber production. A variety of specialty synthetic rubbers, such as butyl, EPDM, polychloroprene, nitrile, and silicone, account for the balance of synthetic rubber production.

Choice of the rubber to be used depends on cost and performance requirements. The specialty rubbers often give superior performance properties but do so at higher product cost.

Many rubber products contain less than 50% by weight of rubber. The balance is a selection of fillers, extenders, and processing or protective chemicals. A typical formulation for the tread of a passenger car tire is shown in the **table**.

In the formulation shown in the table, the mixture of SBR and polybutadiene rubber represents a balance between cost, wear, grip, and aging factors. The carbon black adds strength to the rubber (reinforcing effect). A variety of specialty carbon blacks are available to give a range of processing and performance characteristics. Oil is used as an extender for the elastomer and may be used as a softener or processing aid. In many cases, part of the oil may be incorporated by the synthetic rubber producer directly in the rubber and supplied as a rubber oil master batch.

Formulation for passenger car tire tread

Component	Parts
Styrene-butadiene rubber	65
Polybutadiene rubber	35
Carbon black	67
Oil	40
Zinc oxide	3
Stearic acid	1
Accelerator	1.2
Sulfur	2
Wax	2
Antidegradant	2

Zinc oxide, stearic acid, and an accelerator are used to control the rate and degree of vulcanization. Mercaptobenzothiazoles and benzothiazole sulfenamide derivatives are the most commonly used accelerators for sulfur-cured products, although a wide variety of other chemical classes may be used. Sulfur is used to provide the cross-linking between rubber molecules. Wax and antidegradant are used to protect the finished product against thermal and oxidative aging, fatigue cracking, and ozone attack. Substituted quinolines, amines, or phenolics provide protection against thermal or oxidative aging.

In products other than tires, clays may be added as extenders, silicas as reinforcing agents, and plasticizers for flex or fire-retardant properties; and colors or brightening agents may also be used.

Mixing. This step accomplishes an intimate and homogeneous mix of the formulation components. Most large-volume stocks are mixed in internal mixers; these operate with two winged rotors with the compound ingredients forced into the rotors by an external ram.

In many cases, sulfur and accelerator systems are added in a separate mixing operation on a two-roll mill (**Fig. 3**); this mixing operation is normally carried out at lower temperatures than in an internal mixer to prevent the cross-linking reaction between sulfur and the rubber.

A rubber mill consists of two steel rollers. The rollers rotate at different surface speeds toward each other, thus giving a high shear condition that aids in mixing of the formulation.

There are thousands of different recipes, each designed for different purposes and each requiring a mixing procedure of its own. In addition, various manufacturers differ in their ideas as to precisely how the mixing operation should be conducted, and the equipment in various plants differs widely. The trend is toward more automation in this process, both in weighing and charging the ingredients into the mixer and in handling the stock after discharge.

Forming. Usually, forming operations involve either extrusion into the desired shape or calendering

Fig. 3. Two-roll mill. (*B. F. Goodrich Co.*)

to sheet the material to some specified gage or to apply a sheet of the material to a fabric (**Fig. 4**).

Building. Building operations, varying from simple to complex, are required for products such as tires, shoes, fuel cells, press rolls, conveyor belts, and life rafts. These products may be built by combining stocks of different compositions or by combining rubber stocks with other construction materials such as textile cords, woven fabric, or metal. For a few products no building operations are required. For example, many molded products are made by extruding the rubber, cutting it into lengths, and placing them in the mold cavities. Also, extruded products that are given their final shape by the die used in the extruder are ready for the final step of vulcanization, with no intermediate building operations.

In building tires, the "green" tire is built on a rotatable, collapsible drum, the diameter of which is slightly larger than the diameter of the bead rings. In producing bias-ply tires, a ply of coated cord fabric cut on a bias is applied to the drum so that the cords form an angle of about 40° with a circumferential line around the drum. A second ply is then applied, with the cords running at about 90° to those in the first ply. The bead assemblies are then applied, and the edges of the first two plies are folded over the beads. A third and then a fourth ply are added, with alternation of the direction of the cords, and their edges are folded over the beads. The extruded tread strip, which has previously been cut to the proper length and spliced, is then applied and rolled down. The drum is then collapsed and the green tire removed from the drum. **Figure 5** shows a cross section of a building drum with a green tire on it. Following the building, the tire may be shaped or "lifted" by inserting a rubber curing bag. This operation increases the diameter of the green tire in the crown or tread area and draws the two beads together, thus forming a section resembling that of the finished tire. Sometimes this operation is carried out automatically in the curing press.

The building operations for tires vary widely, depending on the kind of tire to be built and the equipment available in the particular plant. For example, the cord reinforcement may be of polyester, nylon, or steel wire; the number of plies required increases as the size and service requirements increase; the tread proper may be of a different composition than the tread base and sidewalls; one of the sidewalls may be made of a white stock; the construction may be of the tubeless variety, in which case an air-barrier ply of rubber is applied to the inside surface of the first ply; a puncture-sealing layer may be applied to the inside surface; or there may be several bead assemblies instead of only one.

Fig. 4. A calender. (*B. F. Goodrich Co.*)

Increasing popularity of radial tires has caused some basic changes in tire manufacturing operations. In a radial tire, the reinforcement cords in the body, or carcass, of the tire are placed in a radial direction from bead to bead. Radial tires require special building drums and result in a green tire of a different appearance from the barrel-shaped green bias-ply tire.

Vulcanization. The final process, vulcanization, follows the building operation or, if no building operations are involved, the forming operation.

Vulcanization is the process that converts the essentially plastic, raw rubber mixture to an elastic state. It is normally accomplished by applying heat for a specified time at the desired level. The most common methods for vulcanization are carried out: in molds held closed by hydraulic presses and heated by contact with steam-heated platens, which are a part of the press in open steam in an autoclave; under water maintained at a pressure higher than that of saturated steam at the desired temperature; in air chambers in which hot air is circulated over the product; or by various combinations of these methods.

The vast majority of products are sulfur-cured; that is, sulfur cross-links join the rubber chains together. For special applications, vulcanization may take place without the use of sulfur, for example,

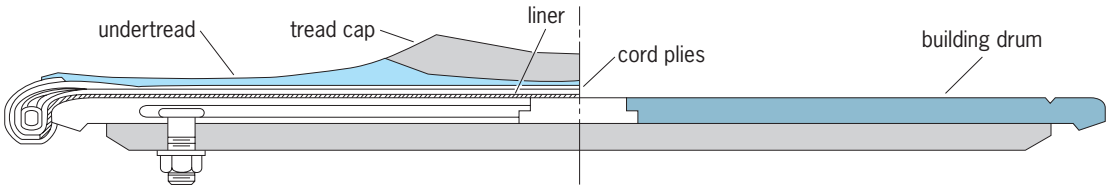

Fig. 5. Cross section of building drum and green tire.

in resin- or peroxide-initiated cross-links or metal-oxide-cured polychloroprene.

The time and temperature required for vulcanication of a particular product may be varied over a wide range by proper selection of the vulcanizing system. The usual practice is to use as fast a system as can be tolerated by the processing steps through which the material will pass without "scorching," that is, without premature vulcanization caused by heat during these processing steps. Rapid vulcanization effects economies by producing the largest volume of goods possible from the available equipment. This is particularly the case for products made in molds, because molds are costly, and their output is determined by the number of heats which can be made per day.

The rate of vulcanization increases exponentially with an increase in temperature; hence, the tendency is to vulcanize at the highest temperature possible. In practice, this is limited by many factors, and the practical curing temperature range is 260–340°F (127–171°C). There are numerous exceptions both below and above this range, but it probably covers 95% of the products made.

Finishing operations following vulcanization include removal of mold flash, sometimes cutting or punching to size, cleaning, inspection for defects, addition of fittings such as valves or couplings, painting or varnishing, and packing.

Latex technology. In addition to the technology of solid rubber, there is a completely different, relatively small but important part of the rubber industry involving the manufacture of products directly from natural rubber latex from the tree, synthetic latex from emulsion polymerization, or aqueous dispersions made from solid rubbers.

In latex technology, materials to be added to the rubber are colloidally dispersed in water and mixed into the latex, a process involving the use of lighter equipment and less power than the mixing of solid rubber compounds. The latex compound can then be used in a variety of processes such as coating or impregnating of cords, fabrics, or paper; in paints or adhesives; molding (such as in toys); dipping (for thin articles like balloons, or household and surgeon's gloves); rubber thread (for garments); and production of foam. Latex technology is particularly important in producing articles for medical and surgical uses. John J. Leucken; Edward G. Partridge

Bibliography. M. Ash, *Handbook of Plastic and Rubber Additives*, 2004; A. K. Bhowmick and H. L. Stephens (eds.), *Handbook of Elastomers (Plastics Engineering)*, 2000; D. C. Blackley, *Synthetic Rubbers: Their Chemistry and Technology*, 1983; J. Brydson, *Rubbery Materials and Their Compounds*, 1989; C. A. Harper, *Handbook of Plastics, Elastomers & Composites*, 2002; C. Hepburn (ed.), *Rubber Technology*, 3d ed., 1990; W. Hofmann, *Rubber Technology Handbook*, 1990; J. E. Mark, B. Erman, and F. R. Eirich (eds.), *Science and Technology of Rubber*, 2d ed., 1994; A. D. Roberts (ed.), *Natural Rubber Science and Technology Handbook*, 1988; M. R. Sethuraj and N. M. Mathew (eds.), *Natural Rubber: Biology, Cultivation, and Technology*, 1992; J. R. White and S. K. De (eds.), *Rubber Technologist's Handbook*, 2001.

Rubber tree

Hevea brasiliensis, a member of the spurge family (Euphorbiaceae) and a native of the Amazon valley. It is the natural source of commercial rubber. The tree may become 60–100 ft (18–30 m) tall and 8–10 ft (2.4–3 m) in circumference. *See* EUPHORBIALES.

Economic importance. It has been introduced into all the tropical countries supporting the rainforest type of vegetation, and is grown extensively in established plantations, especially in Malaysia. The latex from the trees is collected and coagulated. The coagulated latex is treated in different ways to produce the kind of rubber desired. Rubber is made from the latex of a number of other plants, but *Hevea* (see **illus.**) is the rubber plant of major importance. *See* RAINFOREST; RUBBER. Perry D. Strausbaugh; Earl L. Core

Foliage and flowers of the rubber tree.

Diseases. South American leaf blight, restricted to Central America and northern South America, remains a limiting major factor in the development of a rubber industry in the Western Hemisphere. Symptoms caused by the fungus *Microcyclus ulei* (three spore forms) are strictly foliar but vary according to leaf age. Young, expanding, reddish leaves (to 10 days of age) are most susceptible, but leaves 16 days old are effectively immune. Airborne conidia germinate immediately and initiate a disease cycle completed in about a week. Incidence and severity of the disease have been correlated with occurrence

and length of an annual dry season combined with seasonal and cumulative rainfall. Control of the disease in nurseries has been achieved with a fungicide (zineb) but is not economically feasible in the field. Field control has been accomplished through the use of resistant material crown-grafted onto susceptible stocks that give high yields of latex.

A complex of rubber diseases caused by *Phytophthora palmivora* includes vertical black lines in renewing bark (black thread or black stripe), discolored areas of untapped bark 0.4–0.8 in. (1–2 cm) above the soil line (patch canker), and defoliation of green pods and leaves of all ages (Phytophthora leaf fall, the only disease of mature rubber leaves). All aspects of this complex are most severe in wet weather. Incidence of black thread is reduced by opening panels in dry weather, extending the frequency of tapping (renewing bark becomes increasingly resistant to infection after 4–6 days), and applying fungicide after each tapping cut. Excision of diseased tissues and application of a fungicide has been practiced to control patch canker. Foliar application of cuprous oxide or copper oxychloride has been used to reduce leaf fall.

Secondary leaf fall caused by *Glomerella cingulata*, troublesome only under conditions of excessive moisture, can be controlled by repeated application of a fungicide (zineb, copper oxychloride, or chlorothalonil). Powdery mildew (*Oidium hevea*) affects young leaves in cool humid conditions, but can be controlled by early application of sulfur dust. Pink disease (*Botryobasidium salmonicolor*) is more prevalent in wet weather, and serious loss may result from bark damage. Brush application of 2% tridemorph in a natural latex base has provided control).

Bird's-eye spot, caused by *Drechslera hevea*, occurs in most rubber-growing areas but is more serious in nurseries. The disease is less severe in shade, and repeated applications of zineb or maneb have provided satisfactory control.

Symptoms of root diseases on rubber trees often are not apparent until affected trees have died. Infections by *Rigidoporus lignosus* (white root), *Ganoderma philippii* (red root), *Phellinus noxius* (brown root), or *Armillariella mellea* (*Armillaria* root disease) result from root contact with debris remaining after clearing (diseased) endemic flora. A rationale has been adopted in some areas that scattered, early, root disease losses of young rubber trees are acceptable if the disease complex is eliminated by the time that trees reach maturity. Control in recent years involves killing endemic trees in place by ring barking and the application of a herbicide, which also kills roots and renders them nonsusceptible to infection by root-disease fungi. Rubber is planted after clearing. Early collar and root infections by root-disease fungi may be controlled by excision of diseased tissues and the application of 20% pentachloronitrobenzene in bitumen (white root disease), 10% drazoxolon (red root disease), and 10% tridemorph (brown root disease). *See* PLANT PATHOLOGY.

A. A. Cook

Rubella

A benign, infectious virus disease of humans characterized by coldlike symptoms and transient, generalized rash. This disease, also known as German measles, is primarily a disease of childhood. However, maternal infection during early pregnancy may result in infection of the fetus, giving rise to serious abnormalities and malformations. The congenital infection persists in the infant, who harbors and sheds virus for many months after birth.

Virus characteristics. Rubella virus is a ribonucleic acid–containing, ether-sensitive virus about 50–70 nanometers in diameter. It has an internal nucleoid with a double membrane; virus grown in certain kinds of cell cultures will agglutinate red blood cells. Rubella virus can be propagated in tissue cultures, but is difficult to detect directly since it produces no cytopathic effect in cells in which it multiplies best—cercopithecus monkey kidney cells. Various tests for detection of the virus or its antibody are based upon the ability of the virus to interfere with the growth of certain other viruses which are cytopathogenic. The presence of unneutralized rubella virus prevents destruction of the cultures by the challenge virus. *See* ANIMAL VIRUS; ANTIBODY; TISSUE CULTURE.

Immunology. In rubella infection acquired by ordinary person-to-person contact, the virus is believed to enter the body through respiratory pathways. Antibodies against the virus develop as the rash fades, increase rapidly over a 2–3-week period, and then fall during the following months to levels that are maintained for life. One attack confers life-long immunity, since only one antigenic type of the virus exists. Immune mothers transfer antibodies to their offspring, who are then protected for approximately 4–6 months after birth. *See* IMMUNITY.

In the congenital infection, however, even though the fetus produces antibody in response to virus transmitted from the mother, the infant continues to harbor and shed the virus for as long as 18–24 months after birth, and is capable of transmitting the infection to susceptible contacts. Congenitally infected infants may have cataracts, deafness, mental retardation, microcephaly, congenital heart disease, growth retardation, and disorders of blood, liver, spleen, long bones, and skull, as well as interstitial pneumonitis. As a result of the severe epidemic of rubella in the United States in the spring of 1964, about 30,000 babies born in that fall and winter were afflicted with the congenital rubella syndrome.

Epidemiology. Rubella is a worldwide disease. Epidemics occur less frequently than those of some of the other childhood diseases, such as measles, and therefore more members of the population remain susceptible into young adulthood. Serological studies indicate that in most areas of the world, about 50% of the female population have already experienced infection by the age of 6–8 years and are therefore immune, and 80–85% are immune by 17–22 years of age.

Diagnosis. Since rubella infection in adults resembles a number of other illnesses, diagnosis cannot

be made with certainty on the basis of symptoms alone. Laboratory-proved rubella in the crucial period of pregnancy, the first 10 weeks, is associated with fetal infection in virtually 100% of cases, and there is a very high probability of serious abnormality or malformation of the fetus. Developments in the cultivation and testing of the rubella virus and its antibody make it feasible for pregnant women to be tested for preexisting rubella antibody and consequent immunity to infection, or for laboratory confirmation of suspected rubella infection.

Immunization. Live attenuated rubella vaccines have been available since 1969. The vaccine, which is now grown in human diploid cells, induces high antibody titers and an enduring and solid immunity. It may also induce secretory immunoglobulin (IgA) antibody in the respiratory tract and thus interfere with establishment of infection by wild virus. This vaccine is available as a single antigen or combined with measles and mumps vaccines (MMR vaccine).

The vaccine virus multiplies in the body and is shed in small amounts, but it does not spread to contacts. Vaccinated children pose no threat to mothers who are susceptible and pregnant. In contrast, nonimmunized children can contract and spread wild virus to susceptible family contacts. The vaccine induces immunity in at least 95% of recipients, and that immunity endures for at least 10 years.

The vaccine is safe and causes few side effects in children. In postpubertal females, the vaccine may produce self-limited arthralgia and arthritis in about one-third of vaccinees.

In the United States, control of rubella is being attempted by routine vaccination of children aged 1–12 years and selective immunization of adolescents and women of childbearing age. Before the vaccine became available, about 70,000 cases were being reported annually. Vaccination decreased the incidence of rubella to only 550 cases in 1986, a reduction of 99%. However, the decrease occurred primarily in children; in persons 15 years of age and older, only a small reduction in incidence was observed. Since the introduction of vaccine, scattered outbreaks of rubella still occur, chiefly among nonvaccinated adolescents in high school and college.

If vaccine is inoculated into pregnant women, the living vaccine virus may cross the placenta and infect the fetus. However, it is not teratogenic. In a study conducted between 1971 and 1986, more than 1000 women were vaccinated within 3 months of conception. None gave birth to children with congenital rubella syndrome. Therefore, accidental immunization during pregnancy is not an indication for termination of the pregnancy. Nevertheless, it is still considered prudent to avoid vaccination during pregnancy, and nonpregnant women vaccinees should be advised to delay conception for at least 3 months.

Because immunity may wane in those who were vaccinated as children, some women who become pregnant may be at risk of infection. Therefore it has been proposed that vaccine be given to prepubertal girls and to women in the immediate postpartum period. It may be wise for all pregnant women to undergo a serum antibody test for rubella and, if found to be susceptible, receive a vaccination immediately after delivery. *See* BIOLOGICALS; VACCINATION.

Joseph L. Melnick

Bibliography. A. S. Evans, *Viral Infections of Humans*: *Epidemiology and Control*, 4th ed., 1997; D. M. Knipe et al. (eds.), *Fields Virology,* 4th ed., 2001; W. A. Orenstein et al., The opportunity and obligation to eliminate rubella from the United States, *J. Amer. Med. Ass.*, 251:1988–1994, 1984; M. K. Serdula et al., Serological response to rubella revaccination, *J. Amer. Med. Ass.*, 251:1974–1977, 1984.

Rubellite

The red to red-violet variety of the gem mineral tourmaline. Perhaps the most sought-for of the many colors in which tourmaline occurs, it was named for its resemblance to ruby. The color is thought to be caused by the presence of lithium. Fine gem-quality material is found in Brazil, Madagascar, Maine, southern California, the Ural Mountains, and elsewhere. Gem material is almost exclusively a product of pegmatite dikes. Although rubellite is relatively inexpensive, it is regarded by many as one of the loveliest of gemstones. It has a hardness of 7–7.5 on Mohs scale, a specific gravity near 3.04, and refractive indices of 1.624 and 1.644. *See* GEM; TOURMALINE.

Richard T. Liddicoat, Jr.

Rubiales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Asteridae of the class Magnoliopsida (dicotyledons). The order consists of the large family Rubiaceae, with about 6500 species and the family Theligoniaceae with only 3 species. The Rubiales are marked by their inferior ovary; regular or nearly regular corolla with the petals grown together by their margins; stamens (equal in number to the petals) which are attached to the corolla tube alternate with the lobes; and opposite leaves with interpetiolar stipules or whorled leaves without stipules. The most familiar species of temperate regions are herbs with whorled leaves, such as madder (*Rubia tinctorium*, the traditional source of red dye), but opposite-leaved tropical shrubs such as *Coffea* (the source of coffee) are more typical of the group. *See* COFFEE; IPECAC; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

Rubidium

A chemical element, Rb, atomic number 37, and atomic weight 85.47. Rubidium is an alkali metal. It is a light, low-melting, reactive metal. The physical properties of rubidium metal are summarized in the **table**. *See* PERIODIC TABLE.

Physical properties of rubidium metal

Property	Temperature °C	Temperature °F	Metric (scientific) units	British (engineering) units
Density	20	68	1.53 g/cm^3	95.5 lb/ft^3
Melting point	39	102		
Boiling point	688	1270		
Heat of fusion	39	102	6.1 cal/g	10.95 Btu/lb
Heat of vaporization	688	1270	212 cal/g	381 Btu/lb
Viscosity	50	122	6.26 millipoises	4.1 kinetic units
	220	428	3.23 millipoises	
Vapor pressure	294	561	1 mm	0.019 lb/in.^{2}
	628	1162	400 mm	7.75 lb/in.^{2}
Thermal conductivity	39	102	0.07 cal/(s)(cm^2)(cm)(°C)	16.9 Btu/(h)(ft^2)(°F)
	50	122	0.075 cal/(s)(cm^2)(cm)(°C)	18.1 Btu/(h)(ft^2)(°F)
Heat capacity	39–126	102–259	0.0913 cal/(g)(°C)	0.0913 Btu/(lb)(°F)
Electrical resistivity	50	122	23.15 microhm-cm	
	100	212	27.47 microhm-cm	

Most uses of rubidium metal and rubidium compounds are the same as those of cesium and its compounds. The metal is used in the manufacture of electron tubes, and the salts in glass and ceramic production.

Rubidium is a fairly abundant element in the Earth's crust, being present to the extent of 310 parts per million (ppm). This places it just below carbon and chlorine and just above fluorine and strontium in abundance. Sea water contains 0.2 ppm of rubidium, which (although low) is twice the concentration of lithium. Rubidium is like lithium and cesium in that it is tied up in complex minerals; it is not available in nature as simple halide salts as are sodium and potassium.

Rubidium has a density of 1.53 g/cm^3 (95.5 lb/ft^3, a melting point of 39°C (102°F), and a boiling point of 688°C (1270°F).

Rubidium is so reactive with oxygen that it will ignite spontaneously in pure oxygen. The metal tarnishes very rapidly in air to form an oxide coating, and it may ignite. The oxides formed are a mixture of Rb_2O, Rb_2O_2, and RbO_2. The molten metal is spontaneously flammable in air.

Rubidium reacts violently with water or ice at temperatures down to −100°C (−148°F). It reacts with hydrogen to form a hydride which is one of the least stable of the alkali hydrides. Rubidium does not react with nitrogen. With bromine or chlorine, rubidium reacts vigorously with flame formation. Organorubidium compounds can be prepared by techniques similar to those used for sodium and potassium. *See* ALKALI METALS; CESIUM.

Marshall Sittig

Bibliography. D. R. Lide, *CRC Handbook of Chemistry and Physics*, 85th ed., CRC Press, 2004.

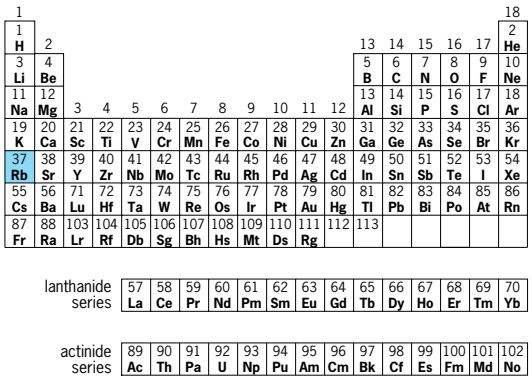

Ruby

The red variety of the mineral corundum, in its finest quality the most valuable of gemstones. Only medium to dark tones of red to slightly violet-red or very slightly orange-red are called ruby; light reds, purples, and other colors are properly called sapphires. In its pure form the mineral corundum, with composition Al_2O_3, is colorless. The rich red of fine-quality ruby is the result of the presence of a minute amount of chromic oxide, usually well under 1%. The chromium presence permits rubies to be used for lasers producing red light. *See* LASER; SAPPHIRE.

The mineral corundum is commonly a constituent of basic igneous rocks, but it rarely occurs in a transparent form suitable for gem use. It is also known as a constituent of the type of marble formed in the contact zone of an igneous intrusion into an impure limestone. It is in this type of deposit that the finest rubies (those of Mogok, Burma) were formed. Today, most of those mined in the Mogok region are taken from the famous gem gravels of that area. There are only two other sources of significance, Ceylon and Thailand. In each of these countries, the finest quality obtained is far less valuable than fine Burma material, for the Ceylon ruby is too light in color and the so-called Siam ruby is a darker red. The finest ruby is the transparent type with a medium tone and a high intensity of slightly violet-red, which has been likened to the color of pigeon's blood. Star rubies do not command comparable prices, but they, too, are in great demand.

The ruby was among the first of the gemstones to be duplicated synthetically and the first to be used extensively in jewelry. A French chemist, A. Verneuil, announced successful reproduction in 1902 by a

flame-fusion process. (However, it is certain that flame-fusion synthetic rubies were available much earlier.) Many years later, C. Chatham produced synthetic ruby by flux fusion. Several others have also been successful. Bell Laboratories made synthetic rubies by a hydrothermal process. *See* CORUNDUM; GEM.

Richard T. Liddicoat, Jr.

Rugosa

One of the two principal orders of Palaeozoic corals, the other being the Tabulata. Rugose corals first appeared in the Middle Ordovician. The small number of known Cambrian corals include at least one species with superficial similarity, but it is not considered a direct ancestor of the Rugosa. The Rugosans diversified steadily during the Early and Middle Paleozoic to reach their greatest diversity in the Middle Devonian. A series of extinction events in the Late Devonian almost resulted in the disappearance of the group, but a handful of survivors rediversified to almost match peak diversity by the Late Mississippian. Thereafter, the group declined, slowly at first but more rapidly in the Permian, and became extinct at the end of the Permian. *See* PALEOZOIC; PERMIAN.

The rugose coral skeleton (or corallum) was secreted by a polyp (coral animal) assumed to have been similar in many respects to the polyps of living corals. Rugose corals existed in either solitary or colonial form, though solitary forms were more common.

Solitary forms. The skeleton of a solitary rugose coral (corallite) consists of a cup, cylinder, or (more rarely) a plate that is almost always defined by a solid outer wall (epitheca), in or on which radial vertical blades or spines (septa) are disposed (**Figs. 1***a*, *b*, **2***a*, *b*). The upper surface, to which the soft parts were attached, is the calice. Subhorizontal elements are usually deposited between or beneath the septa and vary, including simple, more or less flat plates only (tabulae), a peripheral series of smaller vesicular plates (dissepiments), and an axial series of larger tabulae which may vary in form considerably from genus to genus (Figs. 1*c* and 3*b*). An axial structure, constructed of modified septal ends and/or axial tabulae, may be present (Fig. 2*d*).

Colonial forms. In about one-third of rugose corals, the initial corallite (protocorallite) produces offsets to form a colony. There was no specialization of individuals in rugose coral colonies and all the corallites are essentially similar in appearance. The offsets formed either fasiculate colonies, characterized by discrete cylindrical corallites, either parallel or divergent, that were held in a cluster by the skeleton beneath them (Fig. 2*c*) or massive colonies, in which the corallites remained in contact. In many massive colonies, the corallites retained their individual epithecae, and packed together to form a polygonal (often hexagonal) cross section (Fig. 2*d*). However, in a small number of these colonies, the epithecae are lost, and are replaced by a continuous holotheca (external wall), which surrounds the entire colony. The corallites then merge into one another in various ways (Figs. 2*e* and 3*b*). Often, the peripheral ends of septa are confluent between corallites, or (more rarely) withdrawn from the intercorallite areas so that corallite centers are joined by dissepimental tissue only. This is thought to reflect integration of the polyps secreting the skeleton with continuous neural networks across the colony surface and, possibly in some rare cases, confluent gastric cavities.

Classification. Classification of rugose corals can be difficult, as the skeletons are often relatively simple and the same structures may appear again in different lineages. The presence and arrangement of septa, dissepiments, and tabulae; the form of any axial structure that may be present; and to some extent the microstructure of the skeletal elements are all important in classification. The number of major subdivisions is relatively stable, but the phylogenetic relationships between them are disputed.

Rugosa versus Scleractinia. The superficially similar scleractinian corals of the post-Paleozoic are not directly related to the Rugosa, although the

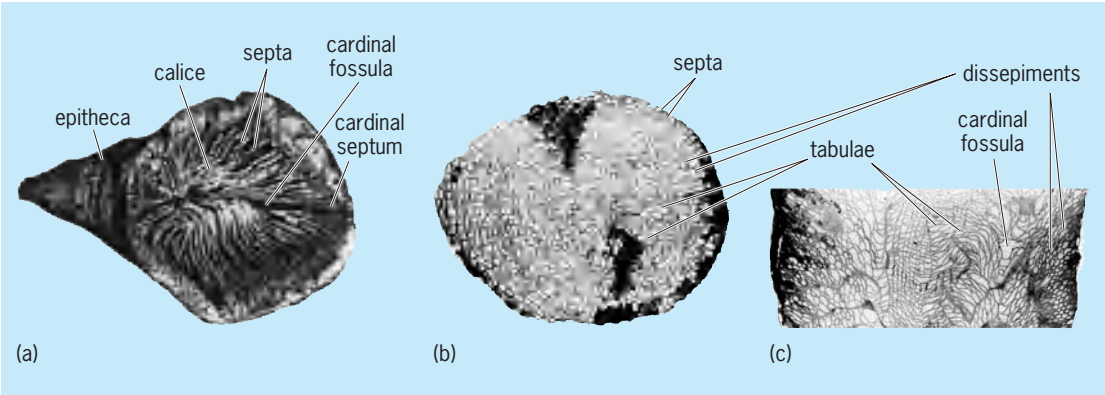

Fig. 1. Rugose corals. (*a*) *Aulacophyllum* sp of horn coral showing the major external features of a solitary coral skeleton; septa pinnately arranged about the cardinal fossula. Devonian, North America, ×0.75. [Most rugose corals need to be studied in thin section for accurate identification, as shown in parts *b* and *c*.] (*b*) Cross section of *Phaulactis angusta*, showing septa and the intersects of dissepiments and tabulae between the septa, ×1. (*c*) Longitudinal section of *P. angusta*, showing dissepiments, tabulae, and the depression that is the cardinal fossula. Silurian, U.K., ×1.

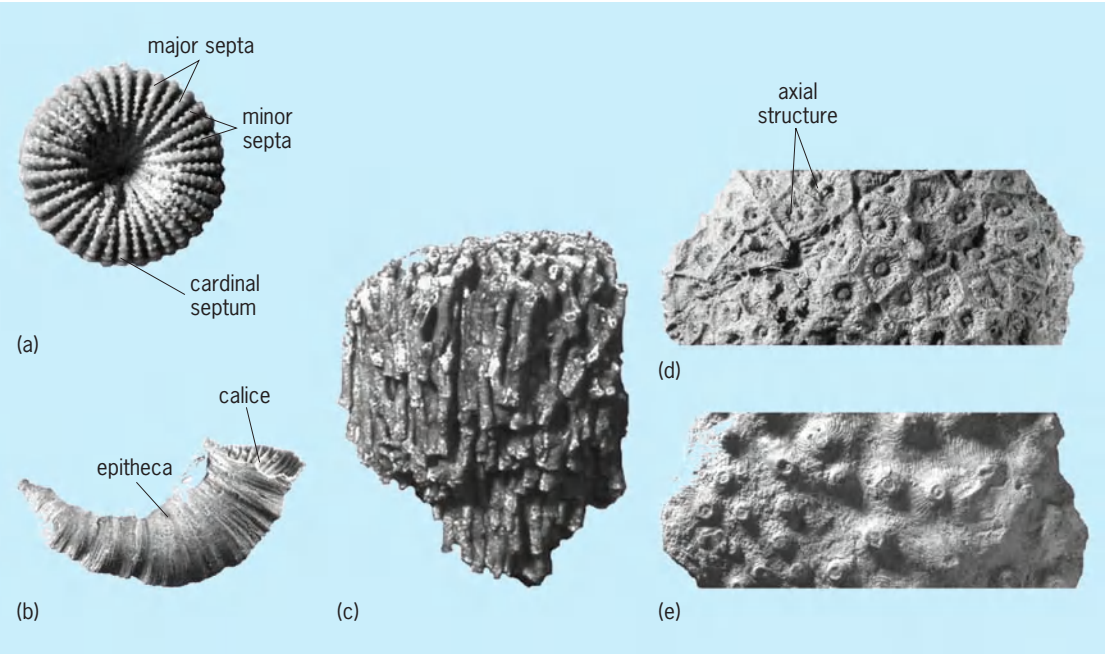

Fig. 2. Rugose coral skeletons. (*a*) *Palaeocyclus porpita*, discoidal corallite, calical view with platelike epitheca beneath, cardinal septum shortened. Silurian, Sweden, ×2. (*b*) *Caninia* sp., typical horn coral in growth orientation with septa visible in the calice. Carboniferous, Ireland, ×1. (*c*) *Siphonodendron* sp., fasciculate colony. Carboniferous, UK, ×0.5. (*d*) *Actinocyathus floriformis*, massive colony, individual corallite walls retained. Carboniferous, UK, ×0.6. (*e*) *Arachnophyllum murchisoni*, massive colony, individual corallite walls lost; peripheral ends of septa either abut or are continuous between corallites. Silurian, U.K., ×0.4.

relationship of the polyp to the skeleton in living scleractinian corals can be used to some extent to interpret the biology of the rugose corals.

Morphology. The main differences between the Rugosa and Scleractinia are (1) a contrasting mode of formation of the septa, together with a second order of minor septa (and rarely additional orders) in Rugosa; (2) a different mineralogy of the skeleton, calcitic in Rugosa, aragonitic in Scleractinia; and (3) the relative rarity of an epitheca surrounding the corallite in Scleractinia. In the Scleractinia, septal insertion is related to the sequence of appearance of fleshy radial partitions (mesenteries) in the polyp; a similar relationship is assumed for the Rugosa. Both orders have six protosepta (the initial septa); however, septal insertion in scleractinians is cyclical in all six sectors defined by the protosepta, whereas in the rugosans, septa are inserted serially in only four of the sectors. This can result in a pinnate arrangement of septa, particularly around the protoseptum (known as the cardinal septum), which may be shortened (Figs. 1*a* and 2*a*) and associated with a pit (known as the cardinal fossula) in rugose corals. However, in most mature corallites, the arrangement of septa is essentially radial.

Paleoecology and ecology. Another major difference between the Rugosa and the Scleractinia is in their ecology and paleoecology. The vast majority of reef-building scleractinian coral polyps are symbiotic with algae (zooxanthellae), which play an important role in their metabolism and reef-building potential. Indirect evidence suggests that this relationship was not developed in the rugosans. The Rugosa were not able to dominate reef building in the Paleozoic as the scleractinians do in the post-Paleozoic. This appears to be the result of the almost universal presence of an epitheca or holotheca in the Rugosa, which seems to have inhibited secure attachment to a hard substrate. Rugosans mostly have very small attachment scars, and almost all were essentially free-living as adults.

Paleoecological adaptations. Paleoecological adaptations predominantly relate to ensuring stability on a soft substrate.

Solitary rugosans. Among solitary rugosans, long cylindrical corallites had much of their length buried in soft sediment to hold them upright. Storm events resulted in uprooting, survivors often showing sharp changes in their direction of growth. Most cylindrical corals were ultimately preserved lying on their sides (Fig. 3*a*). The common horn-shaped coralla lived convex-side down partially buried in soft sediment (Fig. 2*b*). Some of these developed noncircular cross sections to aid stability. Discoidal corallites of low relief could spread their weight across the substrate (snow-shoe strategy).

Colonial rugosans. Fasciculate colonial rugosans relied on sediment building up between their corallites to ensure stability. Massive rugosans may show the effects of partial mortality due to the movement of soft sediment, with breaks in growth followed by recolonization over a layer of sediment from surviving areas of the colony (Fig. 3*b*). Massive corals followed two principal strategies: (1) spreading, low relief, disk- to domelike colonies adapted to areas of episodic sedimentation and (2) less-spreading, high relief, conical colonies adapted to areas with more

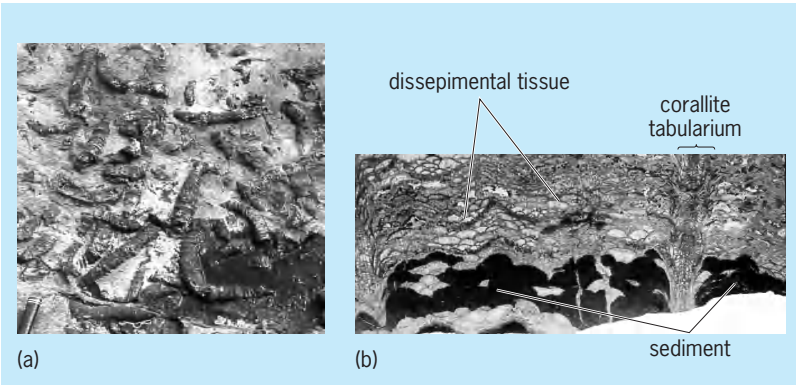

Fig. 3. Rugose coral skeletons. (*a*) *Siphonphyllia gigantea*, solitary cylindrical corallites prone on bedding plane; individual bottom-right showing sharp change of growth direction. Carboniferous, Ireland; width of photo, 1 m. (*b*) *Arachnophyllum pentagonum*, massive colony with extensive sediment fouling in lower part, as indicated by the darker sediment infill. Corallite tabularia connected by continuous dissepimental tissue. Silurian, North America, ×1.

continuous sedimentation. In both cases, sediment fouling was the most likely cause of death. Colonial corals, particularly fasciculate forms, are much more likely to be found in position of growth than solitary corals.

Distribution. Thus, rugose corals played only a minor part in Paleozoic reef-building, although they were more common in peri-reefal environments. Middle Paleozoic reef core facies were dominated by stromatoporoids. Massive tabulate corals often contributed to the structure and rugose corals might be found scattered through the fabric. In some cases, fasciculate rugosans were overgrown by stromatoporoids in an apparently symbiotic relationship. Massive rugosans occasionally occupied niches in the reef core but were more common on the reef flat, and specialized solitary and fasciculate rugosans were often present in more restricted back-reef environments. Rugose corals were most abundant in bioherms (lenslike to moundlike organic structures) and biostromes (sheetlike accumulations of skeletal organisms) usually associated with fine-grained carbonate sediments or calcareous muds. Some biostromes are close to monospecific, while others, and particularly many bioherms, may contain a great diversity of species and growth forms. Rugosans remain reasonably common in deposits interpreted as formed in deeper shelf environments, but little is known of their penetration to off-shelf depths. Their diversity was highest in warm, tropical and subtropical waters, decreasing markedly toward higher temperate and polar regions. Faunas in high latitudes and in deeper waters tend to be dominated by solitary corals, and any colonial corals are relatively small. There was also a specialized fauna of small, thick-walled, solitary rugose and tabulate corals (the laccophyllid fauna) that colonized marginal environments. *See* CNIDARIA; SCLERACTINIA; TABULATA.

Colin Scrutton

Bibliography. D. Hill, in C. Teichert (ed.), *Treatise on Invertebrate Paleontology*, pt. F: *Coelenterata*, suppl. 1, vols. 1, 2, 1981; W. A. Oliver, Jr., Origins and relationships of Paleozoic corals groups and the origin of the Scleractinia, ed. by G. D. Stanley, Jr., *Paleobiology and Biology of Corals: The Paleontological Society Papers*, vol. 1, pp. 107–134, 1996; C. T. Scrutton, *Corals and Other Cnidaria*, in R. C. Selley, L. R. M. Cocks, and I. R. Plimer (eds.), *Encyclopedia of Geology*, vol. 2, pp. 321–334, Elsevier, Oxford, 2004; C. T. Scrutton, The Palaeozoic corals, *Proc. Yorkshire Geol. Soc.*, 51:177–208, 1997, 52:1–57, 1998; J. E. Sorauf, Biocrystallisation models and skeletal structure of Phanerozoic corals, in G. D. Stanley, Jr. (ed.), *Paleobiology and Biology of Corals: The Paleontological Society Papers*, vol. 1, pp. 159–185, 1996; G. E. Webb, Morphological variation and homoplasy: The challenge of Paleozoic coral systematics, in G. D. Stanley, Jr. (ed.), *Paleobiology and Biology of Corals: The Paleontological Society Papers*, vol. 1, pp. 135–157, 1996.

Runge vector

The Runge vector describes certain unchanging features of a nonrelativistic two-body interaction for which the potential energy is inversely proportional to the distance r between the bodies or, alternatively, in which each body exerts a force on the other that is directed along the line between them and proportional to r^{-2}. Two basic interactions in nature are of this type: the gravitational interaction between two masses (called the classical Kepler problem), and the Coulomb interaction between like or unlike charges (as in the hydrogen atom). Both at the classical level and the quantum-mechanical level, the existence of a Runge vector is a reflection of the symmetry inherent in the interaction. *See* COULOMB'S LAW; KEPLER'S EQUATION; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; SYMMETRY LAWS (PHYSICS).

Classical Kepler problem. The nonrelativistic two-body problem can be transformed into an equivalent problem for one body attracted to a fixed center. In the equivalent one-body problem, the motion is governed by the relation in Eq. (1), where $\hat{\mathbf{r}} = \mathbf{r}/r$,

$$\frac{d\mathbf{p}}{dt} = -\lambda r^{-2}\hat{\mathbf{r}} \tag{1}$$

$\mathbf{p} = m\, d\mathbf{r}/dt$, m is the reduced mass, and λ is a constant (positive for attractive force) that characterizes the strength of the interaction. It follows that the energy E and the angular momentum $\mathbf{L}$ are constants of the motion and that Eqs. (2) hold. The Runge vector, defined by Eq. (3), is also conserved (that is, does

$$E = \frac{p^2}{2m} - \frac{\lambda}{r} \qquad \mathbf{L} = r \times p \tag{2}$$

$$\mathbf{R} = \mathbf{p} \times \mathbf{L} - \lambda m \hat{\mathbf{r}} \tag{3}$$

not change with time). Since $\mathbf{R} \cdot \mathbf{L} = 0$, it follows

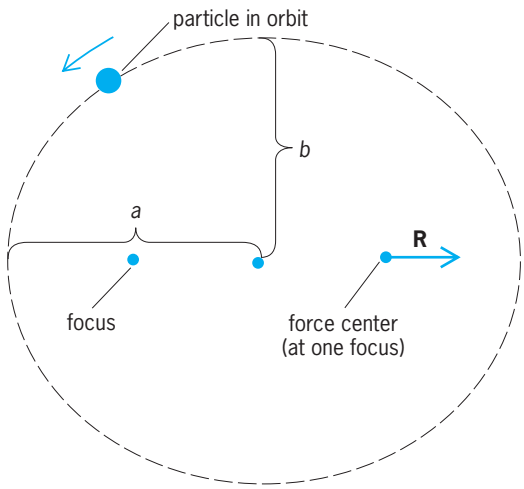

Runge vector for elliptical motion ($R = \lambda m \cdot \sqrt{a^2 - b^2}/a$).

that **R** is a fixed vector in the plane of the orbit. If **R** is taken to be the fixed direction from which the azimuthal angle θ is measured, Eq. (4) holds. Here

$$\frac{1}{r} = \frac{\lambda m}{L^2}\left(1 + \frac{R}{\lambda m}\cos\theta\right) \tag{4}$$

$R = |\mathbf{R}| = [2mEL^2 + \lambda^2 m^2]^{1/2}$. Hence **R** points in the apsidal direction and its magnitude determines the eccentricity $\epsilon = R/\lambda m$. The **illustration** shows the situation for an elliptical orbit. Since the particle in orbit endlessly retraces the same path (broken line in the illustration), the shape of the elliptical path (described by the eccentricity) and its orientation (described by the direction $\mathbf{R}/R$) are constant. A single-valued constant quantity giving a complete description of the orbit is possible only because of the periodic nature of the motion; that is, the orbit is reentrant. Moreover, a single-valued conserved vector in the plane of motion occurs only in the nonrelativistic problem in which the force varies as r^{-2}. In the corresponding relativistic problem, the ellipse precesses.

This symmetry (constancy of the orbit path in the bound case), which leads to the Runge vector, is reflected in other aspects of the Kepler problem. The fact that the corresponding Hamilton-Jacobi equation is separable not only in spherical coordinates but also in parabolic coordinates implies the existence and form of the Runge vector. The components of the Runge vector may also be related to the generating functions of canonical transformations that transform a given orbit pertaining to a fixed energy into a different orbit of the same energy. The group of transformations comprising those generated by rotation in three-dimensional space (namely, generated by **L**) as well as by the Runge vector is isomorphic to rotations in four-dimensional space. *See* CANONICAL TRANSFORMATIONS.

Quantum-mechanical problem. The hamiltonian of the Schrödinger equation analogous to the classical problem is given by Eq. (5), where **p** is the operator $-i\hbar\nabla$. In addition to the angular momentum operator $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, there exists a conserved hermitian Runge vector operator given by Eq. (6). Its

$$H = \frac{\mathbf{p}\cdot\mathbf{p}}{2m} - \frac{\lambda}{r} \tag{5}$$

$$\mathbf{R} = {}^1/_2(\mathbf{p}\times\mathbf{L} - \mathbf{L}\times\mathbf{P}) - \lambda m\hat{\mathbf{r}} \tag{6}$$

time independence is assured by the fact that it commutes with H. Its existence and form can be inferred from the fact that the time-independent Schrödinger equation is separable in both spherical and parabolic coordinates.

Instead of the Runge vector given above, the Runge-Lenz vector operator $\mathbf{A} = (-2mH)^{-1/2}\,\mathbf{R}$ is usually used. It is well defined on all eigenstates and is hermitian for bound states ($E' < 0$). The square of the Runge-Lenz operator is given by Eq. (7), where

$$\mathbf{A}\cdot\mathbf{A} = -\lambda^2 m(2H)^{-1} - (\mathbf{L}\cdot\mathbf{L} + \hbar^2) \tag{7}$$

$\hbar$ is Planck's constant divided by 2π. *See* EIGENFUNCTION; EIGENVALUE (QUANTUM MECHANICS).

The Runge-Lenz vector operator **A** and the angular momentum operator **L** have the commutation properties characteristic of the Lie algebra 0(4) [orthogonal group of rotations in four dimensions]. Explicitly, Eqs. (8)–(10) hold, where ϵ_{ijk} is antisymmetric in its

$$A_iL_j - L_jA_i = i\hbar\epsilon_{ijk}A_k \tag{8}$$

$$A_iA_j - A_jA_i = i\hbar\epsilon_{ijk}L_k \tag{9}$$

$$L_iL_j - L_jL_i = i\hbar\epsilon_{ijk}L_k \tag{10}$$

indices with $\epsilon_{123} = 1$. Moreover, since both A_3 and L_3 commute with each other as well as with the hamiltonian, a simultaneous eigenfunction of these operators can be found. This corresponds to separating the Schrödinger equation in parabolic coordinates $\xi = r(1 + \cos\theta)$, $\eta = r(1 - \cos\theta)$, and ϕ, instead of the usual spherical coordinates r, the polar angle θ, and the azimuthal angle ϕ. These simultaneous eigenfunctions are not also eigenfunctions of the angular momentum since L_2 does not commute with A_3. For energy $E = k^2(2m)^{-1}$ which is greater than zero, one particular eigenfunction is of special interest. It is the one whose L_3 and A_3 eigenvalues are respectively zero and $(\hbar - im\lambda k^{-1})$. It is defined by Eq. (11), where F_{11} is the confluent hypergeometric function. As $r \to \infty$, this wave function behaves as if it were a plane wave plus an outgoing spherical wave, and hence this single eigenfunction has the asymptotic properties of a scattering solution for the Schrödinger equation. The scattering solution is useful in a discussion of scattering of a beam of noninteracting charged nonrelativistic particles off

$$\psi = e^{ikr\cos\theta}\,{}_1F_1[im\lambda(\hbar k)^{-1}, 1, ik\eta] \tag{11}$$

a Coulomb force center since, asymptotically, the plane wave represents that which is scattered by the Coulomb force center. The same eigenfunction is expressible as an infinite sum of angular momentum eigenfunctions, that is, the so-called partial wave expansion of the scattering eigenfunction. *See* DISPERSION RELATIONS.

Symmetry and invariance group. Since the potential for the hydrogen atom is spherically symmetric, it is apparent that both the classical and quantum-mechanical problems enjoy three-dimensional rotational symmetry. The additional symmetry peculiar to the $1/r$ potential extends that symmetry to a four-dimensional symmetry. The space of the four dimensions is that of the variables $2(-2mE)^{1/2}\mathbf{p}/(p^2 - 2mE)$ and $(2mE + p^2)/(p^2 - 2mE)$, where $\mathbf{p}$ is the ordinary three-dimensional momentum. It can be shown that the Schrödinger equation in momentum space can be converted into an integral equation symmetric in these four variables by projecting the momentum space stereographically onto a four-dimensional unit sphere. The abstract Kepler problem in n spatial dimensions has been analyzed with similar results.

The additional internal symmetry is responsible for the degeneracy (that is, number of distinct states of a single energy) characteristic of the hydrogen atom. The group 0(4) is an invariance group of the hydrogen atom since all the bound states of a given energy constitute the basis for an irreducible representation of that group. In turn, all the bound states constitute an irreducible representation of the noncompact DeSitter group 0(4,1). This larger group (whose generators do not all commute with the hamiltonian) is called the noninvariance group of the problem. The conceptual analogy between energy levels and mass, hydrogen atom states and elementary particle states, has given hope that perhaps the spectrum of elementary particles can also be understood in terms of an appropriate invariance and noninvariance group. *See* ELEMENTARY PARTICLE; ENERGY LEVEL (QUANTUM MECHANICS). D. M. Fradkin

Bibliography. I. M. Bander and C. Itzykson, Group theory and the hydrogen atom, *Rev. Mod. Phys.*, 38:330–346, 1966; G. F. Bassani, M. Inguscio, and T. W. Hansch (eds.), *The Hydrogen Atom*, 1989; H. Goldstein, *Classical Mechanics*, 3d ed., 2001; L. D. Landau and E. M. Lifshitz, *Mechanics*, 3d ed., 1976; E. C. G. Sudarshan, N. Mukunda, and L. O'Raifeartaigh, Group theory of the Kepler problem, *Phys. Lett.*, 19:322, 1965.

Running fit

The intentional difference in dimensions of mating mechanical parts that permits them to move relative to each other. A free running fit has liberal allowance; it is used on high-speed rotating journals or shafts. A medium fit has less allowance; it is used on low-speed rotating shafts and for sliding parts. Running fits are affected markedly by their surface finish and the effectiveness of lubrication. *See* ALLOWANCE.

Running and sliding fits are standardized into nine classes. Close sliding fits accurately locate parts with some sacrifice in free motion; they permit no perceptible play. Sliding fits permit the parts to move but are not intended for freely running parts or moving parts subject to appreciable temperature change. Precision running fits permit parts to run freely at low speeds and at light journal pressures, provided temperature differences are limited. Close, medium, and free running fits are intended for progressively higher surface speeds, journal pressures, and temperature ranges. Loose running fits are for use with cold-rolled shafting and tubing made to commercial tolerances. Paul H. Black

Rural sanitation

Those procedures, employed in areas outside incorporated cities and not governed by city ordinances, that act on the human environment for the purpose of maintaining or improving public health. Their purpose is the furtherance of community cleanliness and orderliness for esthetic as well as health values.

Water. Purification of water supplies since 1900 has helped to prolong human life more than any other public-health measure. Organisms which produce such diseases as typhoid, dysentery, and cholera may survive for a long time in polluted water, and prevention of water contamination is imperative to keep down the spread of diseases. Watertight covers for wells are important means of preventing surface contamination to the water supply in rural areas.

Purification of a surface water supply, such as from a lake or a farm pond, is accomplished by sedimentation, filtering, and chlorination. Sedimentation can be effected in a storage chamber by adding aluminum sulfate, which flocculates the finer particles of soil and other undesirable matter in suspension in the water. Filtering through fine sand removes the flocculated particles. Finally, the water is purified by a chlorine solution at the rate of $^1/_2$–1 part of chlorine to 10^6 parts of water. The system of treatment for farm-pond water in **Fig. 1** comprises (1) hedge post or pipe; (2) screen suspended from post 3 ft or 90 cm under water; (3) flexible pipe; (4) hand valve; (5) aspirator, or alum feeder; (6) float valve; (7) hinged wood cover; (8) hand valve; (9) reinforced concrete top; (10) foot valve and strainer; (11) insulated pump house; (12) automatic pump; (13) automatic chlorinator; (14) pressure tank; (15) 2-in. or 5-cm iron pipe or plastic pipe muting; (16) concrete

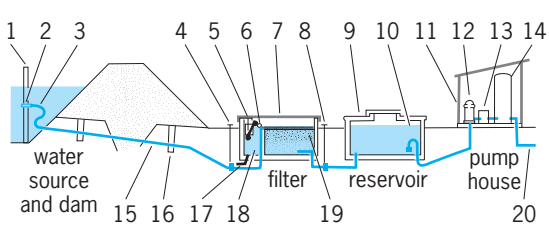

Fig. 1. Farm-pond water-treatment system. The numbers are explained in text.

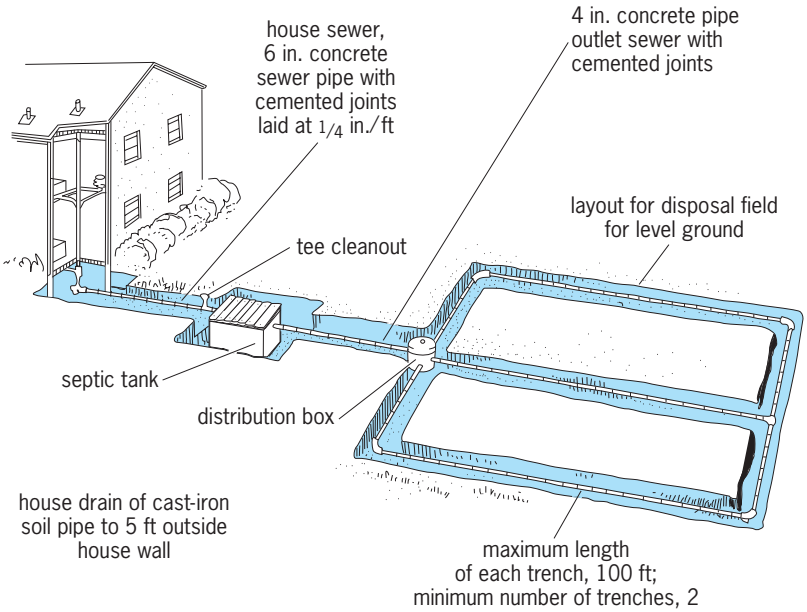

Fig. 2. Typical family-size sewage-disposal system. 1 in. = 2.5 cm; 1 ft = 0.3 m.

cutoff collar; (17) drain when needed; (18) coagulation sedimentation chamber; (19) washed river sand screened through $^1/_8$-in. or 3-mm sieve; and (20) purified water to house, below frost line.

The colon bacillus is the usual indicator of pollution of water supplies by human waste. Chlorine kills such organisms, and therefore it is widely used for purification of water supplies.

Sewage disposal. The problem of safe disposal of sewage becomes more complex as population increases. The old practice of piping sewage to the nearest body of water has proved to be dangerous. Sanitary engineering techniques are now being used in rural areas, as well as in cities, for the disposal of household and human wastes. Where sewage-plant facilities are not available, the most satisfactory method of sewage disposal is by means of the septic tank system (**Fig. 2**).

The septic tank system makes use of a watertight tank for receiving all sewage. Bacterial action takes place in the septic tank and most of the sewage solids decompose, are given off as gases, or go out into the drainage lines as liquid. The gas and liquid are then released from the top 2 ft (60 cm) of soil without odor or sanitary problems. The solids that do not decompose settle to the bottom of the tank, where they can be easily removed and disposed of safely. Such a sewage-disposal system has made it possible for all farm homes and rural communities to have modern bathroom equipment and sanitary methods of sewage disposal. *See* SEPTIC TANK; SEWAGE DISPOSAL.

Harold E. Stover

Rutabaga

The plant *Brassica napobrassica*, a cool-season, hardy biennial crucifer of European origin, belonging to the order Capparales and probably resulting from the natural crossing of cabbage and turnip. Unlike turnip, it has smooth nonhairy leaves and 38 chromosomes. Propagation is by seed, commonly sown in early summer. The fleshy roots (see **illus.**) are cooked and eaten mashed as a vegetable. Rutabagas have been widely grown as a livestock feed in northern Europe and eastern Canada. Popular yellow-fleshed varieties (cultivars) are Laurentian and American Purple Top; a leading white variety is Macomber. Rutabagas have a high requirement for boron. High temperatures cause misshapen root growth. Commercial production is limited to Canada and the northern part of the United States.

Fleshy roots and leaves of rutabaga plants. (*Asgrow Seed Co., Subsidiary of The Upjohn Company*)

Harvesting generally begins after frost and when the roots are 4–6 in. (10–15 cm) in diameter, commonly 90–100 days after planting. The disease clubroot and the root maggot insect are the most common problems. For diseases of rutabaga *see* TURNIP; CABBAGE; CAPPARALES.

H. John Carew

Ruthenium

A chemical element, Ru, atomic number 44. The element is a brittle gray-white metal of low natural abundance, usually found alloyed with other platinum metals in nature. Ruthenium occurs as seven

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

stable isotopes, and more than ten radioactive (unstable) isotopes are known. Four allotropes of the metal are known. Ruthenium is a hard white metal, workable only at elevated temperatures. It can be melted with an electric arc or an electron beam. *See* ISOTOPE; METAL; PERIODIC TABLE; RADIOISOTOPE; TRANSITION ELEMENTS.

The metal is not oxidized by air at room temperature, but it does oxidize to give a surface layer of ruthenium dioxide (RuO_2) at about 900°C (1650°F); at about 1000°C (1830°F) the volatile compounds ruthenium tetraoxide (RuO_4) and ruthenium monoxide (RuO) form, which can result in loss of the metal. Metallic ruthenium is insoluble in common acids and aqua regia up to 100°C (212°F). The principal properties of ruthenium ore given in the **table**. *See* AQUA REGIA.

Ruthenium is relatively rare, having a natural abundance in the Earth's crust of about 0.0004 part per million. It is always found in the presence of other platinum metals. The major commercial sources of the element are the native alloys osmiridium and iridiosmium and the sulfide ore laurite. The element is also separated from other platinum metals by an intricate process, involving treatment with aqua regia (in which ruthenium, osmium, rhodium, and iridium are insoluble), to yield the pure metal.

Ruthenium is used commercially to harden alloys of palladium and platinum. The alloys are used in electrical contacts, jewelry, and fountain-pen tips. Application of ruthenium to industrial catalysis (hydrogenation of alkenes and ketones) and to automobile emission control (catalytic reduction of nitric oxide) and detection have been active areas of research. In medicine, ruthenium complexes have attracted some attention as potential antitumor reagents and imaging reagents. Ruthenium tetraoxide is finding increasing use as an oxidant for organic compounds. *See* CATALYSIS; CATALYTIC CONVERTER; HYDROGENATION; OSMIUM; PALLADIUM; PLATINUM; TECHNETIUM.

Carol Creutz

Bibliography. C. Bruneau et al. (ed.), *Ruthenium Catalysts and Fine Chemistry*, Springer, 2005; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; J. R. Davis (ed.),

Principal properties of ruthenium

Property	Value
Atomic number	44
Atomic weight	101.07
Crystal structure	Hexagonal close-packed
Lattice constant *a* at 25°C (77°F), nm *c/a* at 25°C (77°F)	0.27056
	1.5820
Density at 25°C (77°F), g/cm^3	12.37
Thermal neutron capture cross section, barns (10^{-28} m^2)	2.50
Melting point	2310°C (4190°F)
Boiling point	4080°C (7380°F)
Specific heat at 0°C, cal/g (J/kg)	0.0551 (231)
Thermal conductivity, 0–100°C, cal cm/cm^2 °C	0.25
Linear coefficient of thermal expansion, 20–100°C, μin./(in.)(°C) or μm/(m)(°C)	9.1
Electrical resistivity at 0°C, microhm-cm	6.80
Temperature coefficient of electrical resistance, 0–100°C/°C	0.0042
Tensile strength (annealed), kN · m^{-2}	4.96×10^5
Young's modulus at 20°C (68°F), lb/in.^{2} (Pa)	
Static	60×10^6 (4.1×10^{11})
Dynamic	69×10^6 (4.75×10^{11})
Vickers hardness number (diamond pyramid hardness)	200–350

Metals Handbook: Desk Edition, 2d ed., ASM International, 1998; E. A. Seddon and K. R. Seddon, *The Chemistry of Ruthenium*, Elsevier, 1984.

Rutherfordium

A chemical element, symbol Rf, atomic number 104. Rutherfordium is the first element beyond the actinide series. In 1964 G. N. Flerov and coworkers at the Dubna Laboratories in Russia claimed the first identification of rutherfordium. A. Ghiorso and coworkers made a definitive identification at the Lawrence Radiation Laboratory, University of California, Berkeley, in 1969. *See* PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

The Dubna group claimed the preparation of rutherfordium, mass number 260, by irradiating plutonium-242 with neon-22 ions in the heavy-ion cyclotron. The postulated nuclear reaction was $^{242}Pu + {}^{22}Ne \rightarrow {}^{260}Rf + 4$ neutrons. By 1969 the Berkeley group had succeeded in discovering two alpha-emitting isotopes of rutherfordium with mass numbers 257 and 259 by bombarding ^{249}Cf with ^{12}C and ^{13}C projectiles from the Berkeley heavy-ion linear accelerator (HILAC).

A number of years after the discovery at Berkeley, a team at Oak Ridge National Laboratory confirmed discovery of the isotope ^{257}Rf by detecting the characteristic nobelium x-rays following alpha decay. *See* ACTINIDE ELEMENTS; NOBELIUM; TRANSURANIUM ELEMENTS.

Albert Ghiorso

Bibliography. S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; G. T. Seaborg and W. D. Loveland, *The Elements Beyond Uranium*, 1990.

Rydberg atom

An atom which possesses one valence electron orbiting about an atomic nucleus within an electron shell well outside all the other electrons in the atom. Such an atom approximates the hydrogen atom in that a single electron is interacting with a positively charged core. Early observations of atomic electrons in such Rydberg quantum states involved studies of the Rydberg series in optical spectra. Electrons jumping between Rydberg states with adjacent principal quantum numbers, n and $n - 1$, with n near 80 produce microwave radiation. Microwave spectral lines due to such electronic transitions in Rydberg atoms have been observed both in laboratory experiments and in the emissions originating from certain low-density partially ionized portions of the universe called HII regions. *See* ELECTRON CONFIGURATION; INTERSTELLAR MATTER; MICROWAVE.

The valence electron's classical orbit radius increases as n^2, while the orbital velocity decreases as n^{-1}. The binding energy between the electron and the core decreases as n^{-2}. Thus Rydberg atoms with $n = 80$ have diameters of 10^{-6} m and are as large as some bacteria. Yet they are very delicate, being easily distorted by weak electric and magnetic fields. They are also easily destroyed in collisions with other atomic particles. The natural lifetime of an undisturbed Rydberg atom increases as n^3 for a given electronic angular momentum and has millisecond values for n near 80.

Production and applications. The advent of the laser has made possible the production of sizable numbers of Rydberg atoms within a bulb containing gas at low pressures, 10^{-2} torr (1.3 pascals) or less. The rapid energy-resonance absorption of several laser light photons by an atom in its normal or ground state results in a Rydberg atom in a state with a selected principal quantum number. Aggregates of Rydberg atoms have been used as sensitive detectors of infrared radiation, including thermal radiation. They have also been observed to collectively participate in spontaneous photon emission, called superradiance. Such aggregates form the active medium for infrared lasers that operate through the usual laser mechanism of collective stimulated photon emission. All these developments are based upon the great sensitivity of Rydberg atoms to external electromagnetic radiation fields. Atoms with n near 40 can absorb almost instantaneously over a hundred microwave photons and become ionized at easily achievable microwave power levels. Isotope separation techniques have been developed that combine the selectivity of laser excitation of Rydberg states with the ready ionizability of Rydberg atoms. Such applications have been pursued for atoms ranging from deuterium through uranium. *See* INFRARED RADIATION; ISOTOPE SEPARATION; LASER.

Classical behavior. Quantum systems that have large values for their quantum numbers have many properties that are adequately described by the equations of classical physics. Thus the picture of an atomic electron orbiting in an elliptical trajectory about the nucleus can be increasingly appropriate as n is increased. Experiments can produce localized Rydberg electron wave packets that actually orbit a nucleus. These employ a short laser pulse to initially create an appropriate coherent combination of Rydberg atom energy eigenstates. Although the center of the packet follows the classical electron trajectory, the shape of the packet exhibits quantum wave

interference effects. These include time recurrences of the initial packet localization. The wave effects take longer times to develop as n is increased. This also occurs theoretically when the classical limit is approached by letting Planck's constant pass to zero. As classical physics is considered to be a limiting case of quantum physics, there should be correspondences between quantum and classical quantities and phenomena. *See* QUANTUM MECHANICS.

Quantum chaos. The field of quantum chaos is concerned with those properties of quantum systems that are related to chaos in the classical limit. In this limit, several perturbed Rydberg atom systems exhibit deterministic chaos, a randomlike property characterized by an extreme sensitivity of time evolution to initial conditions. In these chaotic classical systems, it is impossible to predict the future of the Rydberg atom's electron without actually following its evolution. The classical differential equations of motion must be solved continually, by using a computer that must be increasingly precise as longer and longer times are considered.

When a highly excited hydrogen atom is in a sufficiently strong external static magnetic field or microwave field, the competition between the external force on the electron and the atomic Coulomb force on it is so great that the trajectory of the electron becomes chaotic. The optical absorption spectrum of the real quantum-mechanical atom in the external field then exhibits a special distribution of nearest-neighbor energy-level spacings. This distribution is one manifestation of chaos in the classical limit. Another arises when a strong microwave field is applied for a very short time of about 100 microwave oscillation periods. During this time the atom with an initially well-defined value for n quickly evolves into a probability distribution in n that then broadens with time according to a classical diffusion equation, the second manifestation of chaos. A third is that the diffusion stops when a destructive quantum wave interference becomes complete. The interference is highly destructive because of the random character of the underlying classical chaos. All three effects related to classical chaos have been observed in quantal computer calculations that have been confirmed by laboratory experiments. However, the property of quantum systems that corresponds to chaotic behavior in the classical limit is yet undiscovered. *See* CHAOS.

Related systems. The positively charged core attracting a Rydberg valence electron may be a molecular ion rather than an atomic ion. Even more complex Rydberg-like quantum states exist in solids, in which a valence electron can be weakly bound either to an impurity ion or to a hole vacancy in a semiconductor. Thus many of the phenomena observed in studies with Rydberg atoms also can occur in molecules and in the solid state. The consequences of this for physics, chemistry, biology, and technology largely remain unexplored. *See* CRYSTAL DEFECTS; SEMICONDUCTOR.

Experiments have been carried out on doubly excited atoms, in which laser radiation is used to drive simultaneously two atomic electrons out of their normal atomic shells into different Rydberg valence shells. The resulting planetary atoms have two or more valence electrons in widely separated classical orbits. *See* ATOMIC STRUCTURE AND SPECTRA.

James E. Bayfield

Bibliography. G. F. Bassani, M. Inguscio, and T. W. Hansch, *The Hydrogen Atom,* 1989; G. Casati et al., Relevance of classical chaos in quantum mechanics: The hydrogen atom in a monochromatic field, *Phys. Rep.,* 154:77–123, 1987; T. G. Gallagher, *Rydberg Atoms*, 1994; P. W. Milonni, M. L. Shih, and J. R. Ackerhalt, *Chaos in Laser-Matter Interactions*, 1987; R. F. Stebbings and F. B. Dunning, *Rydberg States of Atoms and Molecules*, 1983.

Rydberg constant

The most accurately measured of the fundamental constants; it is a universal scaling factor for any spectroscopic transition and an important cornerstone in the determination of other constants.

Empirical definition. This constant was introduced empirically. J. Balmer's formula described the visible spectral lines of atomic hydrogen, while J. Rydberg's formula applied to the spectra of many elements. Their results may be summarized by Eq. (1), where

$$\frac{1}{\lambda} = R\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right) \tag{1}$$

λ is the wavelength of the spectral line and R is a constant. In Balmer's account of the visible hydrogen spectrum, n_1 was equal to 2, while n_2 took on the integer values 3, 4, 5, and so forth. In Rydberg's more general work, n_1 and n_2 differed slightly from integer values. A remarkable result of Rydberg's work was that the constant R was the same for all spectral series he studied, regardless of the element. This constant R has come to be known as the Rydberg constant.

Atomic theory. Applied to hydrogen, Niels Bohr's atomic model leads to Balmer's formula with a predicted value for the Rydberg constant given by Eq. (2), where m_e is the electron mass, e is the elec-

$$R_\infty = \frac{m_e e^4}{8h^3\epsilon_0^2 c} \tag{2}$$

tron charge, h is Planck's constant, ϵ_0 is the permittivity of vacuum, and c is the speed of light. The equation expresses the Rydberg constant in SI units. To express it in cgs units, the right-hand side must be multiplied by $(4\pi\epsilon_0)^2$. The subscript ∞ means that this is the Rydberg constant corresponding to an infinitely massive nucleus. To account for the fact that the electron and the nucleus of finite mass m_N orbit around a common center of gravity, the electron mass m_e must be replaced by the reduced mass given by Eq. (3). This changes the energy levels by

$$\mu = \frac{m_e m_N}{m_e + m_N} \tag{3}$$

the ratio of electron mass to nuclear mass, which

is about 1/2000 in the case of hydrogen. This prediction was found in excellent agreement with the empirical value found by Rydberg.

E. Schrödinger's wave mechanics predicts the same energy levels as the simple Bohr model, but the relativistic quantum theory of P. A. M. Dirac introduces small corrections or fine-structure splittings of order α^2, where α is the fine-structure constant and is approximately 1/137. The modern theory of quantum electrodynamics predicts further corrections of order α^3 and higher. Additional small hyperfine-structure corrections account for the interaction of the electron and nuclear magnetic moments. *See* FINE STRUCTURE (SPECTRAL LINES); HYPERFINE STRUCTURE; NONRELATIVISTIC QUANTUM THEORY; QUANTUM ELECTRODYNAMICS; RELATIVISTIC QUANTUM THEORY.

The Rydberg constant plays an important role in its relationship to other fundamental constants. It may, for instance, be expressed in terms of the fine-structure constant α or the Bohr radius a_0 by Eqs. (4) and (5). Because of its extreme accuracy, the

$$R_\infty = \frac{m_e c \alpha^2}{2h} \tag{4}$$

$$R_\infty = \frac{\alpha}{4\pi a_0} \tag{5}$$

Rydberg constant gives precise relations between these constants that must be satisfied by any set of constants that may be adopted.

Measurement. The Rydberg constant is determined by measuring the wavelength or frequency of a spectral line of a hydrogenlike atom or ion. The energy levels of other atoms are also proportional to the Rydberg constant, but the difficulty of many-body calculations precludes computations with the same precision as is possible for simple hydrogenic atoms.

Conventional measurements. Prior to 1974, the value of the Rydberg constant was based on wavelength measurements of spectral lines (hydrogen, deuterium, tritium, and singly ionized helium) observed in the emission of gas discharges. The wavelengths are measured interferometrically by comparing the fringe patterns they produce with that of a known wavelength standard. The major limitation has been the Doppler broadening of the spectral lines due to the random thermal motion of the gas atoms. This broadening makes it impossible to resolve the different fine-structure components (**illus.** *a*) even when a deuterium discharge is cooled to liquid-nitrogen temperature (illus. *b*), and it has prevented any significant improvement of such wavelength measurements beyond 1 part in 10^7. A Rydberg value $R_\infty = 10{,}973{,}731.77 \pm 0.83\ \mathrm{m}^{-1}$, based on conventional measurements, was adopted in the 1973 adjustment of the fundamental constants. *See* DOPPLER EFFECT; WAVELENGTH STANDARDS.

Laser spectroscopy. Major improvements in the accuracy of the Rydberg constant became possible when techniques of laser spectroscopy were introduced which could overcome the Doppler broadening of hydrogen spectral lines. In the first laser measurement of the Rydberg constant in 1974, saturation spectroscopy was introduced to resolve single fine-structure components of the red Balmer-α line (illus. *c*). An intense, highly monochromatic beam from a tunable dye laser was used, in effect, to label atoms of slow velocity in a glow discharge, by exciting and thus removing them temporarily from the absorbing lower level. A signal from these atoms was registered as a bleaching effect for a weaker counterpropagating probe beam. An almost tenfold better Rydberg value, $R_\infty = 10{,}973{,}731.406 \pm 0.100\ \mathrm{m}^{-1}$, was obtained.

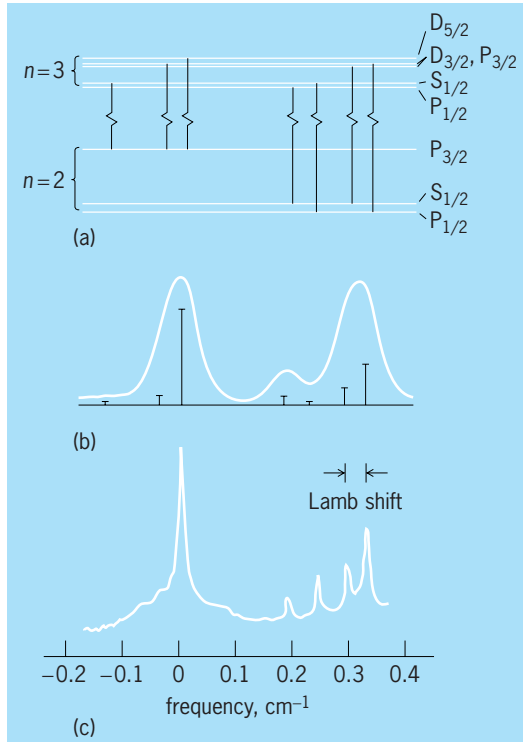

Deuterium Balmer line D_α. (*a*) Energy levels with fine-structure transitions. (*b*) Emission-line profile of a cooled deuterium gas discharge (temperature 50 K or −370°F) and theoretical fine-structure lines with relative transition probabilities (*from B. P. Kibble et al., J. Phys., B6:1079–1089, 1973*). (*c*) Saturation spectrum with optically resolved Lamb shift (*from T. W. Hänsch et al., Precision measurement of the Rydberg constant by laser saturation spectroscopy of the Balmer α line in hydrogen and deuterium, Phys. Rev. Lett., 32:1336–1340, 1974*).

The highest resolution and accuracy has been achieved by the method of Doppler-free two-photon spectroscopy, which permits the observation of very sharp resonance transitions between long-living states. The atoms are excited by two counterpropagating laser beams so that two photons, one from each beam, provide the required excitation energy. From a moving atom, the two beams show opposite Doppler shifts, so that the Doppler effect is canceled to first order. This technique was first applied around 1975 to the transition from the hydrogen 1*S* ground state to the metastable 2*S* state, which has a natural linewidth of only 1.3 Hz. Such laser experiments were the basis for a Rydberg value,

$R_\infty = 10{,}973{,}731.534 \pm 0.013\ \mathrm{m}^{-1}$, that was adopted in the 1986 adjustment of the fundamental constants.

Precision measurements. Since 1986, several Rydberg constant measurements have been performed in different laboratories. Aside from improvements in spectroscopic techniques, the experimental precision has benefited from the development of methods to phase-coherently measure and compare optical frequencies. With this increased precision, the accurate knowledge of quantum-electrodynamic contributions and corrections due to the finite size of the proton has been increasingly important in the determination of the Rydberg constant. These contributions are summarized in the Lamb shift of an energy level. While hydrogenic Lamb shifts, notably the $2S_{1/2}$–$2P_{1/2}$ splitting (illus. *c*), were traditionally measured by radio-frequency spectroscopy, more accurate Lamb-shift values have been obtained by comparison of different optical transition frequencies of the hydrogen atom. By themselves, Lamb-shift measurements provide stringent tests of quantum-electrodynamic theory, approaching the limits imposed by the uncertainty of the proton charge radius.

In one series of experiments, the hydrogen 1*S*-2*S* transition was excited in a cold beam of ground-state atoms in a collinear excitation geometry with the frequency-doubled light of a continuous-wave dye laser. In one of these experiments, the laser frequency was compared to the standard radio frequency of a cesium atomic clock with the help of a wide comb of precisely equidistant laser frequencies emitted by a mode-locked laser. This enabled a determination of the frequency of the hydrogen 1*S*-2*S* transition to a precision of 1.8 parts in 10^{-14}. In other experiments, Doppler-free two-photon spectroscopy on a metastable hydrogen beam was used to observe the transitions 2*S*-8*S*, 2*S*-8*D*, and 2*S*-12*D*. With a harmonic frequency chain, accuracies near 8 parts in 10^{-12} were obtained for the absolute frequencies of the transitions 2*S*-8*D* and 2*S*-12*D*. *See* ATOMIC CLOCK; MOLECULAR BEAMS; PHYSICAL MEASUREMENT.

By combining the result of, for example, the hydrogen 1*S*-2*S* two-photon transition frequency with that of 2*S*-8*D*, values for both the Rydberg constant and the 1*S* Lamb shift are obtained. The 2002 adjustment of the fundamental constants, taking into account different measurements, adopted the value $R_\infty = 10{,}973{,}731.568{,}525 \pm 0.000{,}073\ \mathrm{m}^{-1}$ for the Rydberg constant. The measurements provide an important cornerstone for fundamental tests of basic laws of physics. *See* ATOMIC STRUCTURE AND SPECTRA; FUNDAMENTAL CONSTANTS; LASER; LASER SPECTROSCOPY.

Theodor W. Hänsch; Martin Weitz

Bibliography. G. F. Bassani, M. Inguscio, and T. W. Hänsch (eds.), *The Hydrogen Atom*, 1989; T. W. Hänsch, A. L. Schawlow, and G. W. Series, The spectrum of atomic hydrogen, *Sci. Amer.*, 240(3):94–110, 1979; P. J. Mohr and B. N. Taylor, CODATA recommended values of the fundamental physical constants: 2002, *Rev. Mod. Phys.*, 77:1–107, 2005; M. Niering et al., Measurement of the hydrogen 1*S*-2*S* transition frequency by phase coherent comparison with a microwave cesium fountain clock, *Phys. Rev. Lett.*, 84:5496–5499, 2000; C. Schwob et al., Optical frequency measurement of the 2*S*-12*D* transitions in hydrogen and deuterium: Rydberg constant and Lamb shift determinations, *Phys. Rev. Lett.*, 82:4960–4963, 1999; G. W. Series (ed.), *The Spectrum of Atomic Hydrogen: Advances*, 1988; *Units and Fundamental Constants in Physics and Chemistry*, subvol. b, *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology*, 1992.

Rye

A winter-hardy and drought-resistant cereal plant, *Secale cereale*, in the grass family (Graminae). It resembles wheat, with which it intercrosses to a limited extent. Rye is propagated almost completely by cross-pollination; it is partly sterile if a plant is made to self-pollinate. The inflorescence is a spike or ear (**Fig. 1***a*). Spikelets are arranged flatwise against a zigzag rachis (**Fig. 2**); they usually have two flowers, enclosed by a lemma and palea with two adjacent glumes. The young florets contain three stamens and a pistil. The fertilized pistil develops into a naked grain, or kernel (Fig. 1*b* and *c*), that is easily threshed. There are several recognized species of *Secale*, most of which have shattering spikes and small kernels. There are both perennial and winter-annual species of rye, with winter forms being favored over spring types for production. The only commercially cultivated species is the nonshattering *S. cereale*. Ergot sclerotia often are evident in the field and in threshed grain (Fig. 1*d*). *See* CYPERALES; WHEAT.

Production. Rye is more important in Europe and Asia than in the Western Hemisphere. Russia is the leading world producer, followed by Poland and Germany. Canada and Argentina produce significant

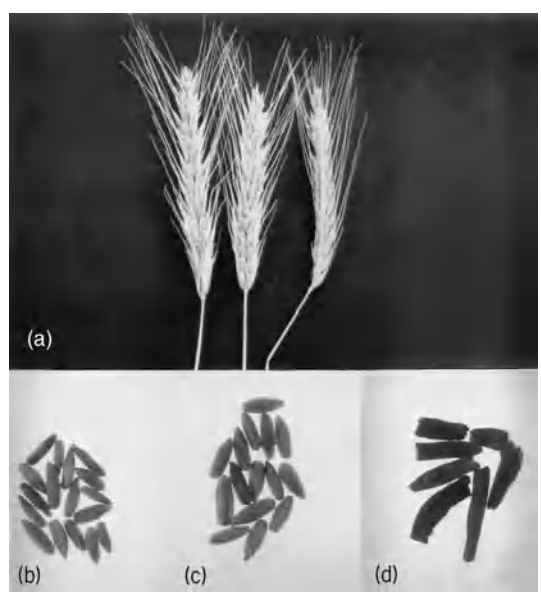

Fig. 1. Rye. (*a*) Spikes or ears. (*b*) Diploid kernels. (*c*) Tetraploid kernels. (*d*) Ergot.

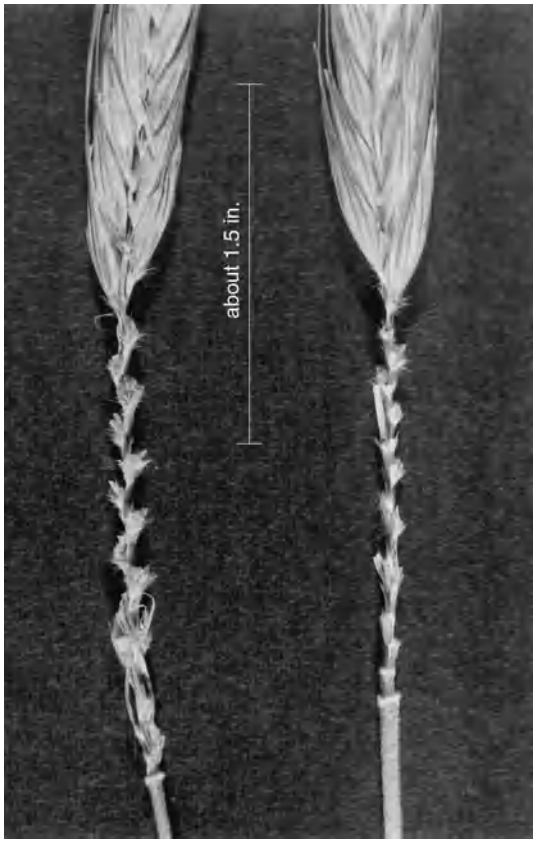

Fig. 2. Florets and zigzag rachis of rye. (*Wisconsin Agricultural Experiment Station*)

amounts, and Switzerland and northwest Europe have high yields. Rye production in the United States is mostly in South Dakota, North Dakota, Minnesota, and Georgia.

Uses. Rye grain is used for animal feed, human food, and production of spirits. Ground rye is mixed with other feeds for livestock. It is often fall-sown to provide soil cover and pasturage for livestock. Egg yolks of chickens and butter from cows fed on rye have a rich yellow color. *See* DISTILLED SPIRITS.

Next to wheat, rye has the most desirable gluten for breadmaking. Rye bread is a staple food in certain sandy land areas of Europe and is favored by peoples of Slavic origin. In the United States, rye bread is made from blends of rye and wheat flour. Rye doughs lack gas retention, resulting in compact loaves or heavy bread. Rye flour contains several B vitamins and as many as 18 amino acids. Rye of different production conditions has 30% more lysine than spring wheats and durums, grown mostly in North Dakota, and somewhat more lysine than Triticales. Dark flours contain a higher content of amino acids than the lighter, refined flours.

Origin. N. I. Vavilov, a Russian plant scientist, believed that rye was introduced into cultivation simultaneously and independently at many localities in central Asia or Asia Minor. Great botanical diversity is found in Transcaucasia, Iran, Turkestan, and Asia Minor, where rye persists without being cultivated. Weedy rye types have closely investing lemmas with partially shattering spikes. Others resemble cultivated rye and are said to have high grain protein; their kernels are not well filled. Rye probably spread as mixtures of wheat.

Varieties. Compared to other small grains, rye has a fewer number of cultivars (agricultural varieties). Some common varieties grown in Europe and Russia are Belta, Danae, Dominant, Golden Dankow, Petkus, Petkus short straw, Sangaste, Toivo, Von Lochow, Vyatka, and Vyatka 2. Some common varieties in North America are Abruzzi, Adams, Cougar, Gator, Rymin, Von Lochow, and Weser. Short-strawed types are gaining favor. Plant and kernel characteristics of rye are variable, partly because of cross-pollination. Height may range from 4 to 6 ft (120 to 180 cm) under moderately fertile conditions. Kernel color may be amber, gray, green, blue, brown, or black.

Tetraploid forms, whose chromosome number has been doubled, are available. These kernels are large (Fig. 1*c*) and the straw is usually stiff, yet this type has not found much commercial use. Tetraploid wheat and rye have been hybridized and chromosomes doubled to form Triticales, which is increasing in usage. *See* BREEDING (PLANT).

Hybrid rye. Heterosis in rye, though not readily usable, has long been recognized. Attempts to find cytoplasmic male sterility in cultivated types have been made in Russia and Europe. Fertility restorers have been found. Hybridizing proper combinations of cytoplasmic male sterility lines and re storers could enhance productivity and use of this crop. *See* HETEROSIS.

Cultural practices. Culture of rye is similar to that of wheat. Seed-bed preparation is kept to a minimum, usually utilizing sandy soil. Plowing the land is preferable, but rye is sometimes drilled directly into stubble. Seed can be sown by hand or with a regular drill at a rate of 56–85 lb/acre (63–95 kg/ha). Rye seed deteriorates during storage; therefore, it should be tested for germinability before sowing. Sowing in the Northern Hemisphere may be done from August to November, but before sustained freezing temperatures. Spring-sown rye is less important than winter rye. Rye may be plowed under for green manure, harvested for hay, or threshed for grain and straw. Grain should be 13.5% or lower in moisture and stored under dry conditions. *See* AGRICULTURAL SOIL AND CROP PRACTICES; GRAIN CROPS. H. L. Shands

Diseases. Ergot is the most important disease of rye (occurs also in wheat) because of the presence of toxic alkaloids in the sclerotia that may be present in the grain. Grain containing ergot sclerotia is discounted heavily when marketed in the United States and many other countries. Grain-cleaning equipment can be used to remove the sclerotia to within tolerance levels of marketing grades and pure food laws. The disease-causing organism, *Claviceps purpurea*, is usually most severe on rye and Triticale, but is also found on wheat, barley, and such grasses as *Bromus* and *Agropyron* species. The disease is recognized by the purple-black, curved sclerotia which protrude from the glumes of infected flowers. *See* ERGOT AND ERGOTISM.

Fig. 3. Ergot of rye. Sclerotium is "germinating" to produce perithecial stromata.

The organism is carried from one season to the next by the sclerotia (ergot bodies), which fall to the ground before harvest or are planted with the seed grain, usually in the fall. In the next spring, when the host plants are flowering, the sclerotia in the ground "germinate" to produce stalks and spherical heads (**Fig. 3**) from which spores (ascospores) are ejected. These spores are carried by air currents to open grain flowers, where infection of the ovaries occurs. As a result of this infection, a second spore stage (conidia) is produced in honeylike droplets. Insects are attracted to these droplets and then visit other grain flowers, secondarily spreading the disease organism. Eventually, the sclerotia replace the kernels in infected flowers. The sclerotia may be up to 2 in. (5 cm) in length and are horny in texture.

Rye is particularly susceptible to ergot because it is normally cross-pollinated, and the flowers tend to be open longer than in self-pollinated grains such as barley and wheat. The longer the flowers are open, the more chance there is for the disease to develop. This is particularly true in thin plantings (where each plant develops many tillers that flower over a long period of time) or in plantings of mixed varieties (which flower at different times).

Control of ergot therefore involves a relatively heavy planting rate with seed of a pure variety. Seed should not contain sclerotia, and rotation of fields may be needed to avoid sclerotia in the soil. Deep plowing would bury the sclerotia so that primary inoculum would not be produced. Sclerotia formed in grasses in fence rows or meadows adjacent to rye fields may also serve as sources of inoculum. Mowing of these grasses to prevent flowering will eliminate this source.

Several other diseases caused by fungi are found on rye and, although usually of minor importance, may be problems at times. These include rusts of stems and leaves (caused by *Puccinia graminis* and *P. recondita*), scab (caused by *Gibberella zeae*) on heads and kernels, and leaf scald (caused by *Rhynchosporium secalis*). Control of these diseases involves resistance, rotation, and seed treatment. Bacteria and nematode diseases have also been described. *See* PLANT PATHOLOGY.

Deane C. Arny

Processing. Rye grain is milled into flour in a manner similar to that used for wheat flour. Several noted variations are made based on the compositional and structural differences between these grains. The careful sizing of rye kernels prior to breaking is important due to greater size and shape variability of this grain. Rye grain of sound, plump character without presence of mold or ergot is essential to yield a high-quality flour. Cleaning operations for rye are more extensive than those required for wheat. Removal of dirt, stones, and foreign material is accomplished by gravity separation and washing. Washed grain is commonly tempered to a moisture content of approximately 14.5% prior to milling.

Break roll operations typically release the greatest portion of flour. The break roll operation results in a limited quantity of rye middlings (intermediate-size rye particles) due to the softness and starchiness of the endosperm. Middling purifiers are not utilized extensively in rye milling. Sifting operations for rye require greater sifting surface area than is typically found in comparable wheat flour sifting. Reduction rolls are also corrugated to minimize flaking of rye break streams (rye particles produced during the first break roll operation) which would inhibit release of flour. The yield of rye flour will generally range from 75 to 85% extraction.

Grades of rye flour are based primarily on color rather than on ash or protein content. Rye flours are commonly described by color, ranging from white (light) to cream (medium) to dark, as the extraction rate increases from 60% to 85%. The desirable rye flavor also increases with extraction rate; thus high-extraction flours possess desirable attributes which have increasingly been utilized in variety breads.

Rye flour doughs lack mixing tolerance and do not develop strong gluten. Rye bread production requires blending of rye flours with wheat flours (commonly clear grades are utilized because color is generally not a limitation) to provide sufficient dough strength. Rye breads are commonly produced from doughs ranging from 20 to 40% rye flour without appreciable loaf volume depression. Specialty varieties of rye breads are classified according to ethnic origins or as sweet or sour doughs. Sour rye breads may be developed from natural lactic fermentations or through the incorporation of cultured milk.

Swedish rye crisp breads are generally prepared from whole ground meal. A mixture of whole meal, yeast, salt, and water may be fermented, thinly sheeted, and cut prior to baking. A short-time–high-temperature baking at 450°F (230°C) for 10 min, followed by additional drying at 200°F (93°C) for 3 h, results in the firm crisp product. *See* FOOD ENGINEERING.

Mark A. Uebersax

Bibliography. W. Bushuk, *Rye: Production, Chemistry and Technology*, 2d ed., 2001; J. G. Dickson, *Diseases of Field Crops*, 2d ed., 1956; R. Gair, J. E. Jenkins, and E. Lester *Cereal Pests and Diseases*, 4th ed., 1987; H. H. Geiger and F. W. Schnell, Cytoplasmic male sterility in rye (*Secale cereale* L.), *Crop Sci.*, 10:590–593, 1970; B. Godon and C. Willin (eds.), *Primary Cereal Processing: A Comprehensive*

Sourcebook, 1993; N. L. Kent, Rye, triticale, rice: Processing, nutritional attributes, technological uses, in *Technology of Cereals with Special Reference to Wheat*, 2d ed., pp. 256–276, 1975; N. L. Kent and A. D. Evers, *Technology of Cereals*: *An Introduction for Students of Food Science and Agriculture*, 4th ed., 1994; W. H. Leonard and J. H. Martin, Rye, in *Cereal Crops*, 1963; K. Lorenz and K. Kulp (eds.), *Handbook of Cereal Science and Technology*, 1990; H. H. Schopmeter, Rye and rye milling, *Cereal Sci. Today*, 7(5):138, 1962; M. V. Wiese, *Compendium of Wheat Diseases*, 2d ed., 1987.

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

16 S - SON

www.MHEST.com

Sable

A carnivore classified in the family Mustelidae along with martens and fishers. The sable (*Martes zibellina*) has a very slender, long, and supple body (see **illustration**). Coloration varies but is usually dark brown on the back and slightly lighter on the flanks, belly, and head. The winter pelage (the coat, which includes fur plus longer, coarser guard hairs) is long, silky, and luxurious with a particularly strong, thick undercoat; the summer pelage is shorter, coarser, duller, and darker. The legs and tail are dark. The limbs are short and terminate in five toes with semiretractile claws. The soles of the feet are covered with extremely dense, stiff hairs. The bushy tail serves as a balancing rod. The dental formula is I 3/3, C 1/1, PM 4/4, M 1/2 × 2, for a total of 38 teeth. Adult sables have a head and body length of 35–55 cm (14–22 in.), including a tail length of 12–19 cm (5–7 in.). They weigh 0.5–2 kg (1–4 lb). *See* DENTITION.

Ecology. Sables are generally solitary and inhabit both coniferous and deciduous forests, preferably near streams. Though they prefer dense, dark coniferous forests, they occur also in stonebirch or larch forests, as well as in forest tundras and swamps. They avoid cultivated areas. Sables move by leaping and are very agile and quick. They spend most of their time on the ground, although at times they climb into trees and rocks. Burrows several meters long that begin under or among rocks, logs, or roots lead to the enlarged nest chamber, which is lined with fur and dry vegetation.

The sable (*Martes zibellina*).

Sables may be active at any time during the day or night. Food consists principally of rodents, but pikas, birds, fish, honey, nuts, and berries may also be consumed. In some regions, they are known to catch salmon migrating upstream during spawning.

Adult sables have few enemies, although young animals may be taken by larger carnivores and predatory birds. The mortality rate of the young is high, with 75–85% dying in their first year of life due mainly to climate and lack of food.

Reproduction and development. Breeding occurs from June to August, with the single annual litter being born the following April or May. The total period of pregnancy lasts 245–298 days due to the process of delayed implantation, in which the fertilized egg remains inactive in the uterus for a considerable time before attaching to the uterine wall and continuing its development. Implantation occurs in February or March; thus, actual embryonic development takes approximately 25–35 days. Usually 3 or 4 (ranging from 1 to 7) young comprise a litter. Newborn sables are blind, deaf, helpless, and weigh 25–35 g (0.9–1.3 oz). The auditory canals open between 23 and 24 days; the eyes open between 30 and 36 days; and the young leave the nest for the first time at about 45 days of age. Weaning occurs at about 7 weeks. Sexual maturity is not attained until about 27 months. The maximum known longevity of sables in captivity is 15 years.

Population status. The sable originally ranged throughout the entire taiga (boreal coniferous

forest) zone from Scandinavia (northern Finland) to eastern Siberia and North Korea including Sakhalin and Hokkaido islands. Because of excessive trapping for its valuable fur, it disappeared from much of its range by the early 1900s. Hunting sable was completely prohibited from 1936 until 1940. However, programs of protection and reintroduction have allowed the species to increase its numbers and distribution, so that exceptionally dense populations are once again found in various areas, chiefly in the mountains of the northernmost parts of its former range. *See* CARNIVORA. Donald W. Linzey

Bibliography. D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford Ltd., 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Saccharin

The sodium salt of *o*-sulfobenzimide. The free imide, called insoluble saccharin because it is insoluble in water, has limited use as a flavoring agent in pharmaceuticals. The sodium and calcium salts are very soluble in water and are widely used as sweetening agents.

Saccharin was first synthesized in 1879 by I. Remsen and C. Fahlberg. Fahlberg discovered the sweet taste of the chemical by accident. Commercial production of saccharin starting from toluene began in 1900. A process for making saccharin from phthalic anhydride was developed about 1950, and saccharin is manufactured by both of these processes, which are shown in the following reaction schemes:

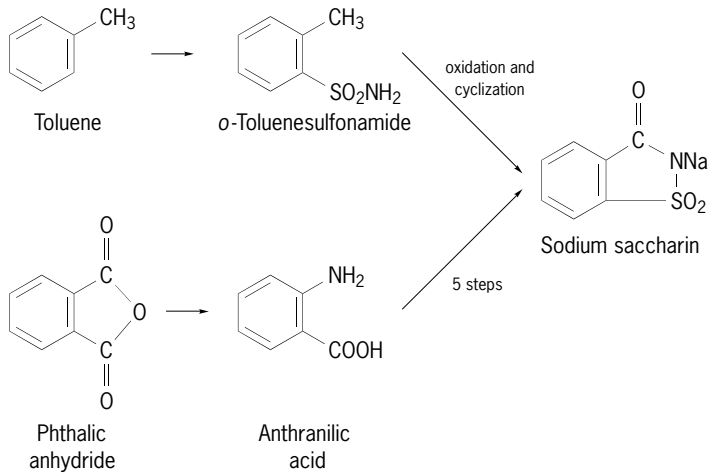

Sodium saccharin is 300–500 times sweeter than cane sugar (sucrose). The saccharin salts are used to improve the taste of pharmaceuticals and toothpaste and other toiletries, and as nonnutritive sweeteners in special dietary foods and beverages. Using noncaloric saccharin in place of sugar permits the formulation of low-calorie products for people on calorie-restricted diets and of low-sugar products for diabetics. Some people experience a bitter aftertaste from saccharin. This can be minimized by proper flavor formulation and by using saccharin as a cosweetener along with another sweetening agent such as aspartame or cyclamate. *See* ASPARTAME.

The regulatory status of saccharin varies from country to country. The United States allows its use in special dietary foods but requires a label statement: "Use of this product may be hazardous to your health. This product contains saccharin which has been determined to cause cancer in laboratory animals." Canada does not allow the use of saccharin in foods or drugs. Most countries permit some uses of saccharin but restrict applications and levels of use as a means of controlling the amount that people consume. *See* FOOD MANUFACTURING. Karl M. Beck

Bibliography. T. E. Furia (ed.), *CRC Handbook of Food Additives*, 2d ed., vol. 2, 1980; *Kirk-Othmer Encyclopedia of Chemical Technology*, 3d ed., vol. 22, 1983.

Saccopharyngiformes

Teleost fishes in the subdivision Elopomorpha whose member taxa have a leptocephalous (elongated and flatterned side to side) larval stage. They are the most aberrant and bizarre of all fishes, having lost a number of structures common to most fishes, such as the symplectic bones, opercular bones, branchiostegal bones, scales, pelvic fins, ribs, pyloric caeca, and swim bladder. The caudal fin is absent or rudimentary, the gill opening small and ventral, the dorsal and anal fins long, and the jaw and hyomandibular bones greatly elongate and attached to the neurocranium by only one condyle (see **illustration**).

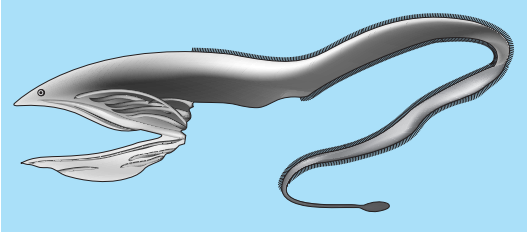

Gulper eel (*Saccopharynx ampullaceus*).

These fishes live at great depths in total darkness. They are adapted to take advantage of any prey that comes their way, large or small. The enormous jaws and distensible gut allows them to capture and swallow prey much larger than themselves.

Twenty-six species in five genera and four families are known from the bathypelagic areas of the Atlantic, Indian, and Pacific oceans. *See* OSTEICHTHYES; TELEOSTEI. Herbert Boschung

Bibliography. E. B. Bohlke (ed.), Orders Anguilliformes and Saccopharygiformes, pp. 1–655, in *Fishes of the Western North Pacific*, pt. 9, vol. 1, Sears Foundation for Marine Research Memoir (Yale University).

Sacoglossa

An order of the gastropod subclass Opisthobranchia containing a thousand living species of herbivorous sea slugs; sometimes called Ascoglossa. They occur

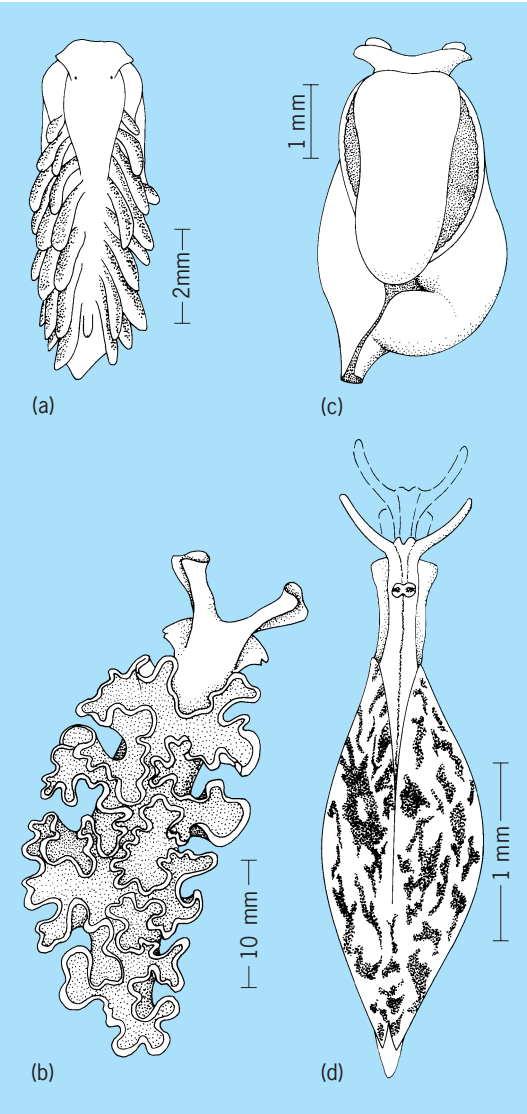

Representative sacoglossans: (*a*) *Stiliger*, (*b*) *Tridachia*, (*c*) *Volvatella*, and (*d*) *Berthelinia*.

in all the oceans, at shallow depths, reaching their greatest size and diversity in tropical seas.

Members of this abundant order have two common features: the herbivorous habit [except for *Olea* and a species of *Stiliger* (**illus.** *a*) which eat the eggs of other mollusks] and the possession of a uniseriate radula in which the oldest, often broken, teeth are usually retained within the body, not discarded. The radular teeth are stilettolike and are used to puncture individual plant cells so that the fluid contents may be swallowed. Dissociated chloroplasts from the prey may be retained and "farmed" in cells of the digestive gland, persisting in *Tridachia* (illus. *b*) for up to 6 weeks. This remarkable endosymbiosis occurs also in the widespread genus *Elysia*.

Sacoglossans possess varied shells; there may be a single shell, capacious and coiled (*Volvatella*, illus. *c*; *Cylindrobulla*), a flattened open dorsal shell (*Lobiger*), or two lateral shells (the bivalved gastropods, *Berthelinia*, illus. *d*). In the highest sacoglossans (*Elysia*, *Limapontia*) the true shell is completely lost after larval metamorphosis.

The association between sacoglossan species and the algae of the sea bottom may be extraordinarily precise, as evidenced in the shallow waters of Florida. *See* OPISTHOBRANCHIA. T. E. Thompson

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; T. E. Thompson, *Biology of Opisthobranch Molluscs*, I, Ray Society, London, 1976.

Safe

An extremely strong container, usually made of metal and sometimes built into a wall or vault, designed to protect money or other valuables. A bank safe is fireproof, strongly built, and equipped with a well-locked door. A furrier's safe provides, in addition, cooling, ventilation, and protection against insects. A safe or vault is a complete structure, the walls, floor, and ceiling being as much a part of the protective features as the more conspicuous door. For utmost protection, a vault is so large and heavy that it cannot be removed. Combination locks eliminate the keyhole as a weakness, and eliminate the key, which could be stolen or duplicated. A time clock built into the locking mechanism unlocks the vault entirely from the inside. The time lock operates only after the elapsed time, set when the vault was closed. Walk-in vaults are furnished with shelves, locked drawers, and cabinets for systematic and compact storage of valuables. The vault may be air conditioned and may have a telephone extension inside. *See* STRUCTURAL MATERIALS. Frank H. Rockett

Safety glass

A unitary structure formed of two or more sheets of glass between each of which is interposed a sheet of plastic, usually polyvinyl butyral. In usual manufacture, two clean and dry sheets of plate glass and a sheet of plastic are preliminarily assembled as a sandwich under slight pressure in order to produce a void-free bond. The laminate is then pressed at 75–225 lb/in.^{2} (0.5–1.5 megapascals) under heat at 239–302°F (115–150°C) long enough to unite. For use in surface vehicles the finished laminated glass is approximately $^1/_4$ in. (6 mm) thick; for aircraft it is thicker.

Because of the ability of the plastic interlayer to yield rather than rupture under load, and because of the adherence of the glass to the plastic, the laminated glass presents less hazard from flying or scattered glass chips in event of damage than would an equal thickness of glass. Alternatively, with chemically tempered glass, the laminate shatters into blunt granules. From this feature, such laminates are termed safety glass. However, the nonrigid plastic layer lowers the modulus of rupture of the laminate to approximately 60% of that of monolithic glass at room temperature. The thermoplastic interlayer

also makes other properties of safety glass appreciably temperature-dependent. Optically, safety glass has substantially the properties of the glass sheets of which it is made. However, above about 160°F (70°C) the organic plastic may deteriorate.

Most safety glass is used in automotive vehicles where the plastic interlayer is 0.030 in. (0.75 mm) thick. It is also used in ships, locomotives, railroad cars, aircraft, safety goggles, and viewing windows. Formed from tempered glass and in multiple layers built up to greater thickness, the laminate is used in bullet-resisting structures in banks, jewelry display cases, military vehicles, and test-chamber windows. The glass or plastic can be tinted to provide color filtering. Where resistance to shattering due to fire is a consideration, wire mesh is embedded in plate glass during manufacture. *See* GLASS. Frank H. Rockett

Bibliography. G. McLellan and E. B. Shand, *Glass Engineering Handbook*, 1984.

Safety valve

A relief valve set to open at a pressure safely below the bursting pressure of a container, such as a boiler or compressed air receiver. Typically, a disk is held against a seat by a spring; excessive pressure forces the disk open (see **illus.**). Construction is

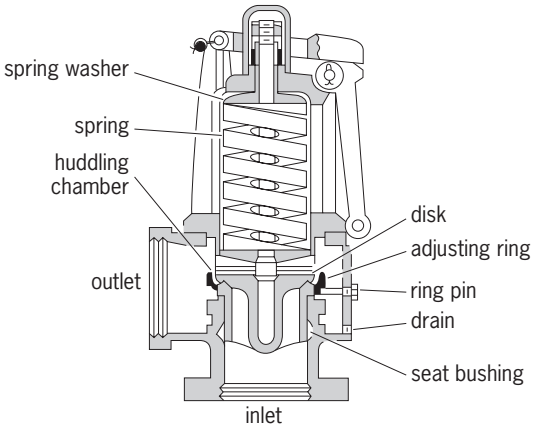

Typical safety valve.

such that when the valve opens slightly, the opening force builds up to open it fully and to hold the valve open until the pressure drops a predetermined amount, such as 2–4% of the opening pressure. This differential or blow-down pressure and the initial relieving pressure are adjustable. Adjustments must be set by licensed operators, and settings must be tamperproof. The ASME Boiler Construction Code gives typical requirements for safety valves. *See* VALVE.

Theodore Baumeister

Safflower

An oilseed crop (*Carthamus tinctorius*) that is a member of the thistle (Compositae) family and produces its seed in heads (see **illus.**). Flowers vary in

Safflower (*Carthamus tinctorius*). (*USDA*)

color from white through shades of yellow and orange to red. Most cultivars of safflower are spiny. The seed is shaped like a small sunflower seed, and is covered with a hull that may be white with a smooth surface or off-white to dark gray with a ridged surface. Depending on hull thickness, the oil content varies from 25 to 45%. *See* ASTERALES.

Origin. Safflower was domesticated several thousand years ago in the Middle East, and spread at an early date in Indian and also along the Nile valley to modern-day Sudan. At first it was used as a source of dye, obtained from the flowers, and later as a source of oil. Until the discovery of aniline dyes, safflower and indigo served as the main source of dye for textiles. Over much of the Middle East, safflower is grown on a small scale for its dried flowers, which are used to color foods. *See* DYE.

Adaptation. Safflower is adapted to an environment that provides abundant supplies of moisture to the roots from seeding through to the end of flowering, and little or no moisture to the aboveground parts of the plant during and after flowering. Such an environment is provided in south-central India, where the summer monsoon rains fill the soil with moisture, which carries the crop from planting in October–November through to harvest in February–March, when the weather is very dry. In areas with a Mediterranean climate (California, Australia, and Mexico) safflower develops vegetatively during or just after the wet winters and matures in the dry summers. Deep soils with high-moisture-holding capacity have been best for safflower. At least 120 days are required from seeding to maturity.

Culture. Where safflower has been grown as a dryland crop, it has been sown in a flat seedbed with rows 12–30 in. (30–75 cm) apart. If the rows are 24–30 in. (60–75 cm) apart, there is room for mechanical cultivation. Under irrigation, safflower has been sown on raised beds, with one row on each bed if the centers are 24–30 in. (60–75 cm) apart, and two rows on beds if the centers are more than 30 in. (75 cm) apart. Frequently the herbicide trifluralin is incorporated into the seedbed before planting to prevent the development of several species of weeds. Nitrogen and phosphorus fertilizers should be applied at about the same rates used for cereal crops.

Plants may be harvested with a grain combine adjusted to provide no damage to the seed. Moisture

content of the seed should be 8% or less for safe storage.

Utilization. There are two types of safflower oil, both with 6–8% palmitic acid and 1–2% stearic acid. One type, the standard or polyunsaturated type, has 76–79% linoleic acid and 11–17% oleic acid. This high-linoleic type has been used in soft margarines marketed in small tubs, in salad oils, and in the manufacture of paints and varnishes. It has had limited use for frying foods, because heat causes it to polymerize and form a tough film on the cooking vessel. The second type of oil, called the high–oleic-acid or monounsaturated type, has 76–79% oleic acid and 11–17% linoleic acid. Its fatty-acid composition is similar to that of olive oil, but the flavor is bland. High-oleic safflower oil is a premium frying oil, and has been reported to be used in some cosmetics. *See* FAT AND OIL (FOOD).

The meal left after the extraction of the oil may contain 20–45% protein, depending on the amount of hull removed from the seed before processing. The meal is used as a poultry and livestock feed. *See* ANIMAL FEEDS.

There is a small but continuing market for dried safflower to be used for ornamental purposes. Types with no spines, long erect branches, and red flowers are preferred. Plants are harvested when in full bloom, and dried in a shaded location.

Production. In the United States, fluctuations in production have been mostly a consequence of the relative prices of safflower and wheat. Production data in other countries are not known definitely, but those in Mexico are estimated to be in the same range as those in the United States, and those in Australia about 20% of those in the United States. California has been the leading state in the production of safflower.

Pests. Insect pests have been a serious limiting factor to safflower production in the Old World, the larvae of the safflower fly (*Acanthiophilus helianthi*) being the most serious. In areas where safflower has been recently introduced, only a few insects such as western flower thrips (*Frankliniella occidentallis*) and lygus bugs (*Lygus hesperus*) have caused moderate damage.

Diseases which cause serious damage but which are controlled by resistant cultivars are rust due to *Puccinia carthami* and wilts due to *Fusarium oxysporum carthami* and *Verticillium dahliae*. Only moderate resistance has been found to a root rot caused by species of *Phytophthora*, so that control has been achieved by careful management of irrigations. Where humidities are high during flowering, gray mold caused by *Botrytis cinerea* and leaf spot caused by *Alternaria carthami* have been serious diseases. *See* PLANT PATHOLOGY. P. F. Knowles

Saffron

The plant *Crocus sativus*, a member of the iris family (Iridaceae). A native of Greece and Asia Minor, it is now cultivated in various parts of Europe, India, and China. This crocus (see **illus.**) is the source of a potent yellow dye used for coloring foods and medicine. The dye is extracted from the styles and stigmas of the flowers, which appear in autumn. It takes 4000 flowers to produce 1 oz (28 g) of the dye. *See* DYE; LILIALES. Perry D. Strausbaugh; Earl L. Core

Saffron (*Crocus sativus*), of the family Iridaceae.

Sage

A shrubby perennial plant in the genus *Salvia* of the mint family (Limiaceae). There are several species, including garden, or true, sage (*S. officinalis*), the sage most commonly used in foods. Many varieties of garden sage are known, but the Dalmatian type possesses the finest aroma. Mediterranean, or Greek, sage (*S. fructicosa*) is also used in foods, but is of inferior quality. Clary sage (*S. sclarea*) and Spanish sage (*S. lavandulifolia*) are used in perfumes and cosmetics. *See* LAMIALES.

Garden sage is native to southern and eastern Europe, and is still cultivated extensively there and in the United States and Russia. It is a plant of low stature (2 ft or 60 cm), with hairy, oblong grayish-green leaves about $1^1/_2$–2 in. (4–5 cm) long. Sage does best in warm, dry regions, with full sun. Sage plants are established by seed or, for better uniformity, by vegetative propagation (stem cuttings). Two or three harvests of sage, with yields of 1200–2000 lb/acre (1050–1800 kg/hectare) of dried leaves, are possible by the second year of cultivation. Plantings last 2–6 years.

The quantity of essential oil is higher if the sage plant is harvested before flowering. To preserve the essential oil content and leaf color, sage is dried, as are most other herbs. Once dried, sage is separated

from the stems and made available to consumers as whole, rubbed (crushed), or ground leaves.

The dried leaves are among the most popular spices in western foods, and are used as an essential ingredient in sausages, pork dishes, and poultry stuffings. Sage is highly aromatic and fragrant, with a pungent, slightly bitter and astringent taste. Both the dried leaves and essential oil of sage are used in flavoring and for antioxidant properties in cheeses, pickles, processed foods, vermouth, and bitters. *See* SPICE AND FLAVORING.

Seth Kirby

Sagittarius

The Archer, a large, zodiacal, southern summer constellation (see **illustration**). It is identified with the centaur Chiron of Greek mythology, knowledgeable in music and medicine and skilled in hunting. (Other myths identify Chiron with the constellation Centaurus.) Sagittarius contains the center of the Milky Way Galaxy, so it is especially rich in star clusters and star clouds. However, it never rises high in the sky in the Northern Hemisphere. Visitors to the Southern Hemisphere marvel at Sagittarius's overall brightness and richness. *See* CENTAURUS; MILKY WAY GALAXY; STAR CLOUDS; STAR CLUSTERS; ZODIAC.

At the center of this large constellation, there is an asterism (star pattern) known as the Milk Dipper, since the Milky Way can be seen through its dipper shape. Several Messier objects (noted in Charles Messier's eighteenth-century catalog of nonstellar objects) include the Lagoon Nebula (M8), the globular star cluster M22, and the open star cluster M24. *See* MESSIER CATALOG.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Sahel

A semiarid bioclimatic zone bordering the southern margins of the Sahara Desert in Africa. It extends between 13 and 18°N and stretches across the width of the African continent from the Atlantic Ocean to the Red Sea, an area of about 5 million square kilometers (1.9 million square miles). Within this area are large portions of the countries of Mauritania, Senegal, Gambia, Mali, Niger, Chad, Sudan, and Eritrea, as well as the northern parts of Burkina Faso, Nigeria, and Ethiopia (**Fig. 1**). The actual boundary of the

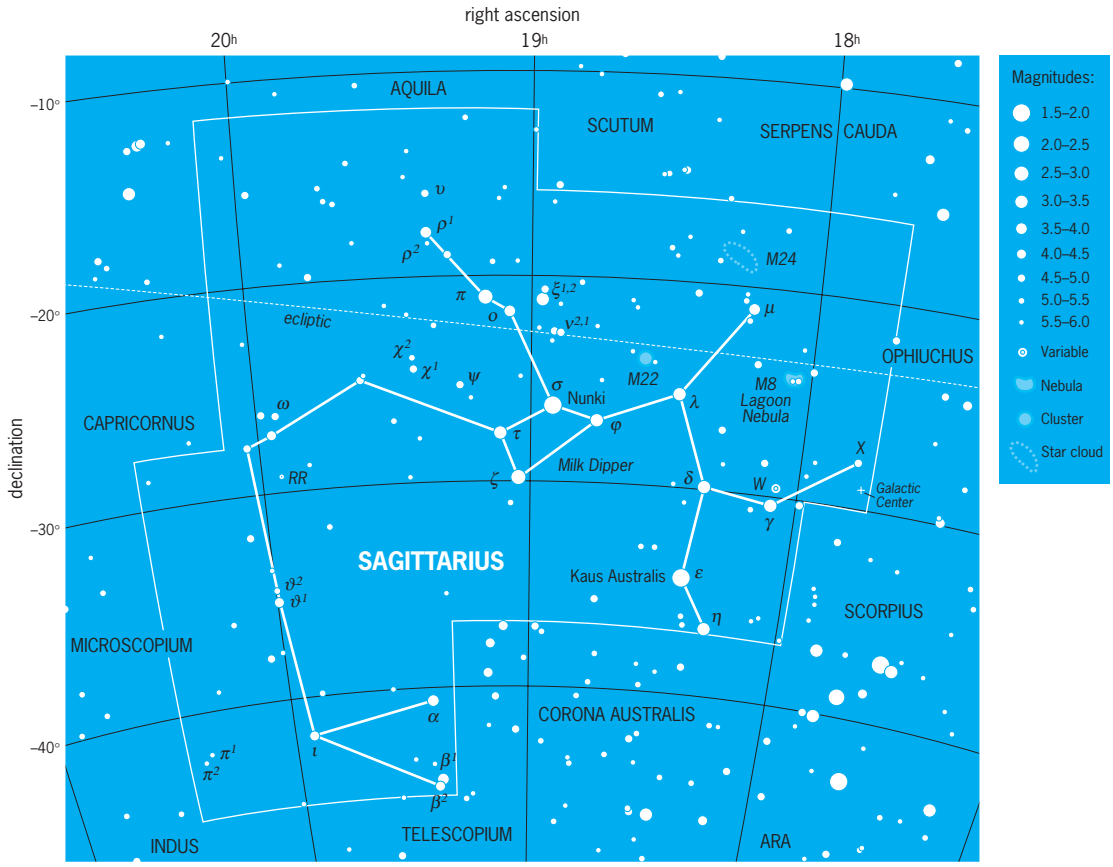

Modern boundaries of the constellation Sagittarius, the Archer. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

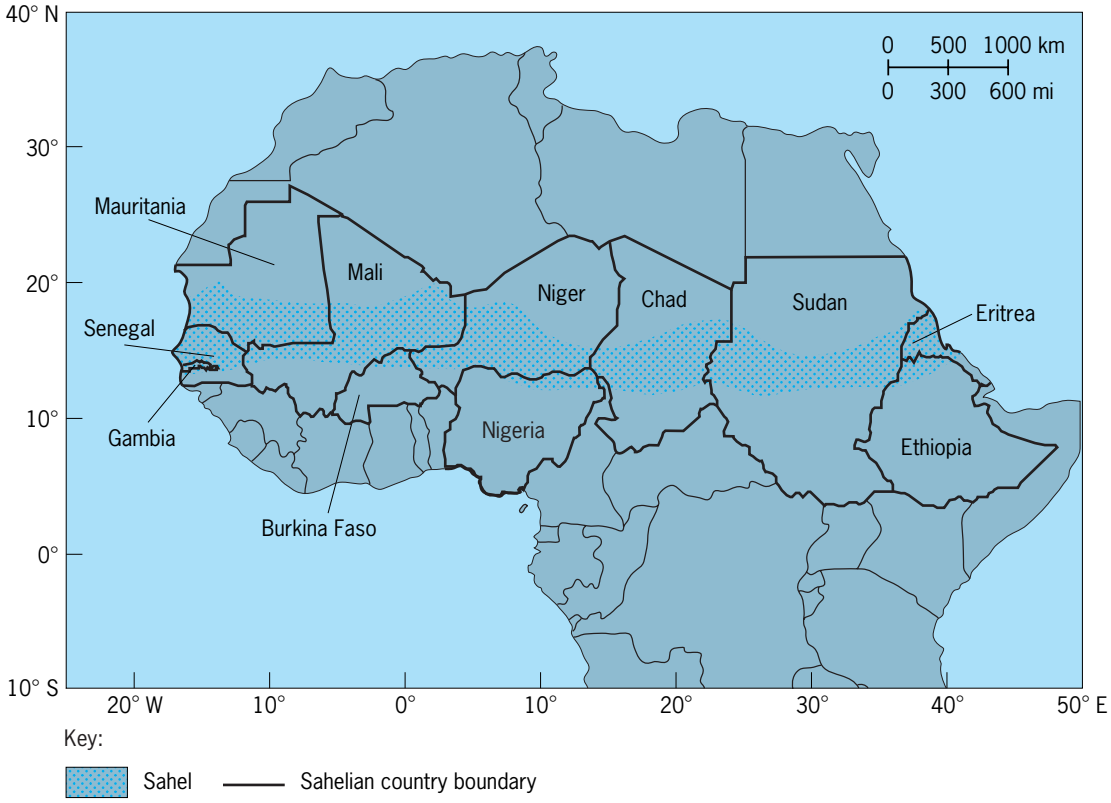

Fig. 1. Location of the Sahel. The Sahelian countries are named.

Sahel is neither as precise nor as rigid as suggested by lines of latitude. For practical purposes, the Sahel is usually defined in terms of its climate and vegetation. *See* AFRICA.

Climate. Two major but contrasting air masses dominate the climate of the Sahel: the dry, tropical continental air from the northeast and the humid, tropical maritime air from the southwest. These air masses converge in a zone of humidity and temperature contrast known as the Inter-Tropical Discontinuity (ITD; over the oceans this boundary is known as the Inter-Tropical Convergence zone, ITCZ). The ITD migrates south to north in response to the apparent position of the overhead sun and the relative intensity of the subtropical high-pressure systems located approximately over the Azores in the North Atlantic Ocean and St. Helena in the South Atlantic. In December–January, a strengthening of the Azores high and a corresponding weakening of the St. Helena causes the ITD to shift south to the coastal margins of West Africa. In the eastern Sahel, where the moderating effect of a coastline is lacking, the ITD is displaced much farther south (below the Equator). As a result, most of North Africa, including the Sahel, comes under the influence of the tropical continental air mass. The wind, known locally as the harmattan, is dry, dusty, and frequently chilly during early morning. In July–August, the dipole reverses and the ITD penetrates deep into West Africa, bringing monsoon rainfall to the Sahel. The length of the rainy season is about 2 or 3 months in the northern Sahel (July–August/September), increasing to 4 months in the southern Sahel (June–September). The mean annual rainfall also shows pronounced south–north gradients, decreasing from about 600–750 mm (24–30 in.) at the boundary of the Sahel and the Sudan savanna in the south to 100–200 mm (4–8 in.) in the north. *See* AIR MASS; CLIMATOLOGY; MONSOON METEOROLOGY; TROPICAL METEOROLOGY.

Rainfall varies significantly from one year to the next. At Niamey, Republic of Niger, for example, the total rainfall received in 1998 was 1159 mm (46 in.), more than double the average of 544 mm (21.4 in.) for the 30-year period from 1970 to 2000, and nearly four times the 324 mm (12.8 in.) received in the severe drought year of 1983. Such high variability complicates agricultural planning and is responsible for the droughts that have plagued the region. *See* DROUGHT.

Temperature is high throughout the year, but tends to be highest between March and June with mean monthly temperatures between 30 and 35°C (86 and 95°F), although daily maximum temperatures frequently exceed 40°C (104°F). The coldest months are December to February with mean monthly temperatures between 20 and 25°C (68 and 77°F). Potential evapotranspiration is also high throughout the year and exceeds rainfall amounts in all months except during the core of the rainy season (July–August) in the southern Sahel.

The Sahel is of interest climatically because it provides the most dramatic and sustained decline in rainfall recorded anywhere in the world (**Fig. 2**).

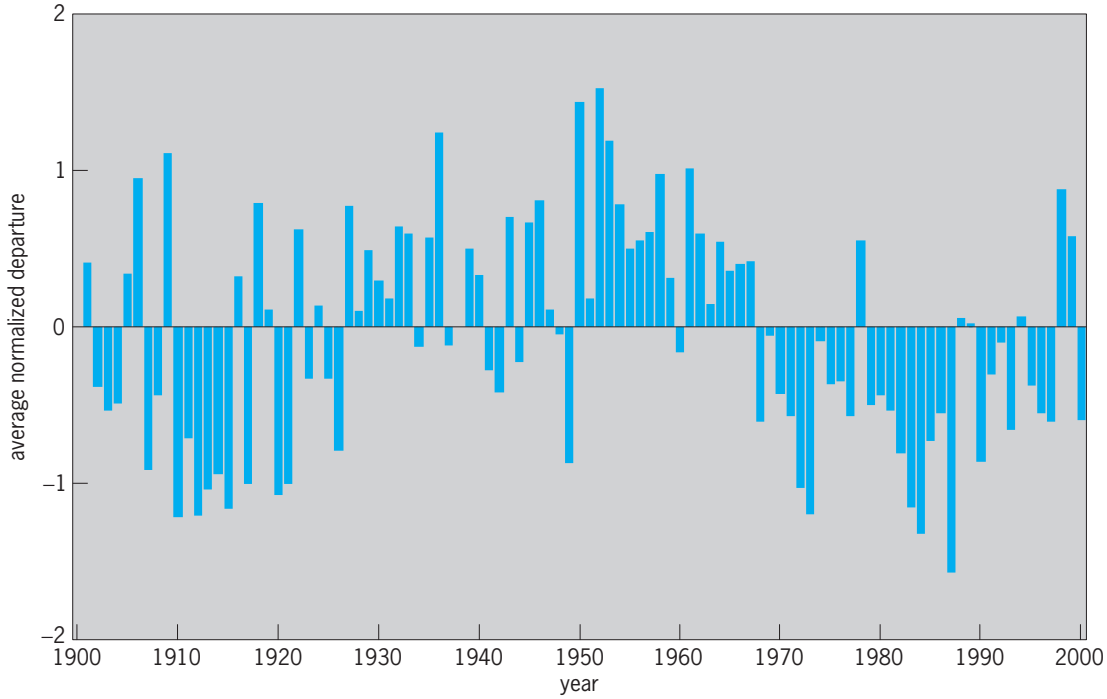

Fig. 2. Standardized rainfall variability over the Sahel during the twentieth century.

Moreover, the livelihoods of its inhabitants are closely linked to the climate, principally rainfall, and finely balanced so that fluctuations in rainfall have serious consequences on agricultural and socioeconomic systems. During the late 1960s and early 1970s, sharply decreased rainfall resulted in devastating drought and famines that caused the deaths of 100,000 to 200,000 people and approximately 12 million head of livestock. The poor rains continued throughout the 1970s and 1980s, producing another drought throughout the region in 1983–1985 that forced millions of people to abandon their homes in search of food. Rainfall conditions recovered somewhat during the mid-1990s, but they have not fully returned to predrought levels.

The causes of these droughts and climate variability in the Sahel generally are not fully understood. Major research efforts utilizing computer modeling are focusing on the remote forcing of Sahelian rainfall through ocean-atmosphere interaction or regional feedback processes involving changes in vegetation and land cover characteristics.

Vegetation. Sahel vegetation is transitional between the Sudan savanna woody grassland to the south and the open desert to the north. The vegetation consists mostly of open grassland and thorny woody species. Annual grasslands dominate the northern Sahel, while wooded grasslands occur on sandy soils in the southern part. Among the most important woody plant species are various species of acacia, combretum, and baobab (*Adansonia digitata*). In some locations, the vegetation is concentrated in strips separated by patches of bare soil. This formation produces a unique pattern called tiger brush because viewed from above it resembles the strips on a tiger's back. The reasons for this peculiar arrangement are not known. *See* DESERT; SAVANNA.

In recent decades, the vegetation in the Sahel has come under intense pressure mainly from increased demand for farmland and fuel wood to satisfy rapidly increasing urban populations. Around major cities, such as Niamey, the landscape has been largely cleared of woody species to a radius of more than 30 km (19 mi). Such changes have raised concerns about desertification—permanent land degradation from the effects of climatic variability and human activities. Such rapid land use changes may be reinforcing the effects of natural climatic variability and helping to perpetuate droughts and increasing aridity. *See* DESERTIFICATION.

Landscape features. Sahel landscape lacks dramatic topographic expression. Apart from the Red Sea hills (1270 m; 4200 ft) on the Eritrean border, sand dunes, and isolated erosion relics, most of the region consists of flat monotonous plains or gently undulating landscape generally below 600 m (2000 ft). The highest points are formed by two late Tertiary volcanic mountains in the Western Sudan: the Jebel Mara (3071 m; 10,075 ft) and the Jebel Gurgeil (2400 m; 7900 ft). A little farther west, the Devonian sandstone-capped Ennedi Plateau in eastern Chad stands 1450 m (4760 ft) high. In east-central Mali, the Dogon sandstone plateau descends gently westward to the River Niger floodplains but terminates abruptly in the southeast, forming cliffs 300 m (1000 ft) high. Known as the Bandiagara escarpment, these cliffs have been inhabited for the past millennium by the Dogon people, famous for their traditional beliefs, art, and cosmology. Farther north,

the Adrar des Iforas, a southern extension of the Ahaggar massif, reaches heights of 700 m (2300 ft). *See* MASSIF.

Rivers and surface water resources. The Niger is the most important river in the Sahel. Originating in the Fouta Djallon highlands in Guinea, the river flows northeast toward the fringe of the Sahara Desert before turning southeast in a great bend to empty into the Gulf of Guinea. This unusual course takes the river through semiarid West Africa for much of its 4200 km (2600 mi) length, providing a source of water and livelihood that helps to mitigate the effects of droughts. The original Tuareg name *egerou n-igereou*, meaning river of rivers, bears eloquent testimony to the significance that the inhabitants have attached to the river historically.

The Bani River joins the River Niger in central Mali, forming a vast inland delta. During wet periods, the delta may attain a surface area of about 80,000 km^2 (30,900 mi^2). The inland delta is important hydrologically because it helps to dissipate a significant portion of the river flow through evaporation and infiltration. *See* DELTA.

Other major rivers in the Sahel include the Senegal in the west, and in the east the Nile, which traverses the zone on its way to the Mediterranean Sea. East of Niamey are the Dallols, a system of fossil river channels that evidently supported a sizable river during the recent geologic past.

Lake Chad is the only natural lake in the Sahel. Once the sixth largest lake in Africa, this inland drainage lake has shrunk from a surface area of about 25,000 km^2 (9650 mi^2) in 1963 to less than 1400 km^2 (540 mi^2) in 2001; it is $^1/_{20}$ of the size that it was 40 years ago. As a result, Lake Chad, which supports nearly 20 million people, is increasingly referred to as the disappearing lake. The primary cause is reduced rainfall throughout the Sahel that began during the late 1960s. However, the construction of dams and increased abstraction of irrigation water on the major river systems that discharge into the lake have worsened the problem. These river systems include the Chari and Logone in Cameroon and the Komadougou-Yobe in Nigeria.

People and economy. The Sahel has traditionally been a cross roads for people and culture from North and West Africa. Indeed, some of the largest and best-organized indigenous African empires and kingdoms, including Nubia, Ghana, Mali, Songhay, and Kanem-Bornu, were located either wholly or substantially within the Sahel. These kingdoms engaged in the trans-Saharan trade with North Africa, which helped the spread of Islam into Sub-Saharan Africa, created a cultural mosaic, and gave rise to some of the legendary cities of Africa, including Tombouctou.

Rain-fed farming, nomadic livestock herding, and agropastoralism (a mixture of sedentary agriculture and livestock keeping) are the major economic activities. Unreliable rainfall makes agriculture, whether cultivated or livestock herding, very risky. Additionally, limited use of technology and modern agricultural aides such as fertilizer or pesticides results in very poor yields, making the zone susceptible to drought-induced famines. Small-scale irrigated agriculture on river floodplains is becoming widespread, especially near large urban centers.

Economically, the countries in the Sahel are poor. Of the 173 countries ranked on the 2002 United Nations Human Development Index, Sudan is the highest-rated Sahelian country at 139. The remaining eight countries rank among the 25 poorest nations in the world. These countries have primarily agrarian underdeveloped economies, and the strain of dealing with persistent drought, famines, and desertification has made many of them dependent on external aid.

Aondover Tarhule

Bibliography. M. Glantz, The West African Sahel, in M. Glantz (ed.), *Drought Follows the Plow*, pp. 33–43, Cambridge University Press, 1994; M. Hulme, Climatic perspectives on Sahelian desiccation: 1973–1998, *Global Environ. Change*, 11:19–29, 2001; M. Tiffen and M. Mortimore, Questioning desertification in dryland Sub-Saharan Africa, *Nat. Resources Forum*, 26:218–233, 2002; United Nations Development Program, *Human Development Report 2002: Deepening Democracy in a Fragmented World*, Oxford University Press, 2002; F. White, *The Vegetation of Africa*, Natural Resources Research, UNESCO, 1983.

Saint Elmo's fire

A type of corona discharge observed on ships under conditions approaching those of an electrical storm. The charge in the atmosphere induces a charge on the masts and other elevated structures. The result of this is a corona discharge which causes a spectacular glow around these points. This effect was accentuated by the contours of early sailing vessels, which usually had several masts and much rigging. The display thus created under darkened conditions was quite striking. The effect was not understood through much of that era and led to much superstition on the part of sailors. *See* CORONA DISCHARGE.

Glenn H. Miller

Salenioida

An order of Echinacea in which the apical system includes one or several large angular plates covering the periproct, with the other characters similar to those of the hemicidaroid urchins. There are two families. The Acrosaleniidae, an extinct group confined to the Jurassic and Cretaceous, had the anus displaced to one side of the periproct as a result of the unequal growth of the suranal plates covering the periproct; the tubercles were perforate and crenulate. The Saleniidae, in which the tubercles are imperforate, ranged from the Jurassic onward, with two surviving deep-sea genera. *See* ECHINACEA; ECHINODERMATA; ECHINOIDEA; HEMICIDAROIDA.

Howard B. Fell

Salicales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Dilleniidae of the class Magnoliopsida (dicotyledons). The order consists of the single family Salicaceae, with about 350 pieces. There are only two genera, *Salix* (willow) and *Populus* (poplar and cottonwood), the former by far the larger. The Salicales are dioecious, woody plants, with alternate, simple, stipulate leaves (see **illus.**) and much reduced flowers that are aggregated into catkins. They have numerous ovules on parietal placentae in a compound, unilocular pistil, which ripens into a usually two-valved capsule. The mature seeds are plumose-hairy and are distributed by the wind. *See* DILLENIIDAE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; PLANT KINGDOM; POPLAR; WILLOW. Arthur Cronquist; T. M. Barkley

Eastern American cottonwood (*Populus deltoides*), showing female catkins with ripening capsules and cottony seeds. (*John H. Gerard, National Audubon Society*)

Salicylate

A salt or ester of salicylic acid having the general formula shown below and formed by replacing the carboxylic hydrogen of the acid by a metal (M) to give a salt or by an organic radical (R) to give an ester. Alkali-metal salts are water-soluble; the others, insoluble. Sodium salicylate is used in medicines as an antirheumatic and antiseptic, in the manufacture of dyes, and as a preservative (illegal in foods). Salicylic acid is used in the preparation of aspirin. The methyl ester, the chief component of oil of wintergreen, occurs free and as the glycoside in many plants. This ester is used in pharmaceuticals as a component of rubbing liniment for its counterirritant effect on sore muscles. It is also used as a flavoring agent and an odorant. The phenyl ester (salol) and others are used medicinally. *See* ASPIRIN. Elbert H. Hadley

$$(o)\text{-}C_6H_4(OH)C(=O)-O-M \quad \text{or} \quad -R$$

Saline evaporites

Deposits of bedded sedimentary rocks composed of salts precipitated during solar evaporation of surface or near-surface brines derived from seawater or continental waters. Dominant minerals in ancient evaporite beds are anhydrite (along with varying amounts of gypsum) and halite, which make up more than 85% of the total sedimentary evaporite salts. Many other salts make up the remaining 15% (**Table 1**); their varying proportions in particular beds can be diagnostic of the original source of the mother brine. *See* ANHYDRITE; EVAPORATION; GYPSUM; HALITE; SALT (CHEMISTRY); SEAWATER; SEDIMENTARY ROCKS.

TABLE 1. Major evaporite minerals*

Mineral	Formula
Anhydrite	$CaSO_4$
Antarcticite	$CaCl_2 \cdot 6H_2O$
Aphthitalite (glaserite)	$K_2SO_4 \cdot (Na,K)SO_4$
Aragonite[LS]	$CaCO_3$
Bassanite	$CaSO_4 \cdot 1/2\ H_2O$
Bischofite	$MgCl_2 \cdot 6H_2O$
Bloedite (astrakanite)	$Na_2SO_4 \cdot MgSO_4 \cdot 4H_2O$
Burkeite	$Na_2CO_3 \cdot 2Na_2SO_4$
Calcite[LS]	$CaCO_3$
Carnallite	$MgCl_2 \cdot KCl \cdot 6H_2O$
Dolomite[LS]	$Ca_{(1+x)}Mg_{(1-x)}(CO_3)_2$
Epsomite	$MgSO_4 \cdot 7H_2O$
Gaylussite	$CaCO_3 \cdot Na_2CO_3 \cdot 5H_2O$
Glauberite	$CaSO_4 \cdot Na_2SO_4$
Gypsum	$CaSO_4 \cdot 2H_2O$
Halite	$NaCl$
Hanksite	$9Na_2SO_4 \cdot 2Na_2CO_3 \cdot KCl$
Hexahydrite	$MgSO_4 \cdot 6H_2O$
Ikaite[LS]	$CaCO_3 \cdot 6H_2O$
Kainite	$4MgSO_4 \cdot 4KCl \cdot 11H_2O$
Kieserite	$MgSO_4 \cdot H_2O$
Langbeinite	$2MgSO_4 \cdot K_2SO_4$
Leonhardtite	$MgSO_4 \cdot 4H_2O$
Leonite	$MgSO_4 \cdot K_2SO_4 \cdot 4H_2O$
Loewite	$2MgSO_4 \cdot 2Na_2SO_4 \cdot 5H_2O$
Mg-calcite[LS]	$(Mg_xCa_{1-x})CO_3$
Magnesite[LS]	$MgCO_3$
Mirabilite	$Na_2SO_4 \cdot 10H_2O$
Nahcolite	$NaHCO_3$
Natron	$Na_2CO_3 \cdot 10H_2O$
Pentahydrite	$MgSO_4 \cdot 5H_2O$
Pirssonite	$CaCO_3 \cdot Na_2CO_3 \cdot 2H_2O$
Polyhalite	$2CaSO_4 \cdot MgSO_4 \cdot K_2SO_4 \cdot 2H_2O$
Rinneite	$FeCl_2 \cdot NaCl \cdot 3KCl$
Sanderite	$MgSO_4 \cdot 2H_2O$
Schoenite (picromerite)	$MgSO_4 \cdot K_2SO_4 \cdot 6H_2O$
Shortite	$2CaCO_3 \cdot Na_2CO_3$
Sylvite	KCl
Syngenite	$CaSO_4 \cdot K_2SO_4 \cdot H_2O$
Tachyhydrite	$CaCl_2 \cdot 2MgCl_2 \cdot 12H_2O$
Thermonatrite	$NaCO_3 \cdot H_2O$
Thernadite	Na_2SO_4
Trona	$NaHCO_3 \cdot Na_2CO_3$
Van'thoffite	$MgSO_4 \cdot 3Na_2SO_4$

*After J. K. Warren, *Evaporite: Their Evolution and Economics*, Blackwell Scientific, Oxford, 1999.

†Less saline alkaline earth carbonates or evaporitic carbonates are indicated by LS; the remainder are evaporite salts. Less common minerals, such as monohydrocalcite, hydromagnesite, the borates, nitrates, and strontium salts, are not listed.

Today, brines deposit their salts within continental playas or coastal salt lakes and lay down beds a few meters thick and tens of kilometers across. In contrast, ancient, now-buried evaporite beds are often much thicker and wider; they can be up to hundreds of meters thick and hundreds of kilometers wide. Most ancient evaporites were formed by the evaporation of saline waters within hyperarid areas of huge seaways typically located within arid continental interiors. The inflow brines in such seaways were combinations of varying proportions of marine and continental ground waters and surface waters. There are few modern depositional counterparts to these ancient evaporites, and none to those beds laid down when whole oceanic basins dried up, for example, the Mediterranean some 5.5 million years ago. *See* BASIN; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; MEDITERRANEAN SEA; PLAYA.

Some of the thicker ancient salt beds, especially those dominated by the salt mineral halite (common table salt), have flowed and folded like toothpaste and have been squeezed through the adjacent sediments to form large slugs of salt unconnected to the original salt bed. In the Gulf of Mexico, such slugs of Jurassic age salt have climbed upward through more than 10 km (6 mi) of Tertiary-age sediment and moved laterally more than 60–70 km (37–44 mi) into the gulf. *See* GULF OF MEXICO; JURASSIC; TERTIARY.

Evaporite minerals and brine evolution. Evaporite salts precipitate by the solar concentration of seawater, continental water, or hybrids of the two. The chemical makeup, salinity (35‰), and the proportions of the major ions in modern seawater are near-constant in all the world's oceans, with sodium (Na) and chloride (Cl) as the dominant ions and calcium (Ca) and sulfate (SO_4) ions present in smaller quantities [a Na-(Ca)-SO_4-Cl brine]. It is a matter of ongoing scientific debate whether the same ionic proportions have characterized seawater in times past. Halite and gypsum anhydrites have been the major products of seawater evaporation for at least the past 2 billion years, but the proportions of the more saline minerals, such as sylvite/ magnesium sulfate ($MgSO_4$) salts, appear to have been more variable. *See* CALCIUM; CHLORINE; ION; MAGNESIUM; SODIUM; SULFUR.

Modern marine evaporites. When modern seawater evaporates, a predictable sequence of evaporite salts precipitates from increasingly concentrated hypersaline waters (**Table 2**). The first mineral to precipitate is calcium carbonate ($CaCO_3$), usually as aragonite; this begins in mesohaline waters when the brine reaches twice the concentration of seawater (40–60‰) and achieves a density of 1.10 g/cm^3. As the brine continues to concentrate and approaches about five times the concentration of seawater (that is, 130–160‰), gypsum (calcium sulfate dihydride; $CaSO_4 \cdot 2H_2O$) precipitates from penesaline waters with brine densities of 1.13 g/cm^3. At surface temperatures higher than 45°C (113°F) but in the same salinity range, anhydrite ($CaSO_4$), the anhydrous form of calcium sulfate, tends to precipitate in preference to gypsum. At 10–12 times the original seawater concentration (340–360‰) and at densities of 1.22 g/cm^3, halite (NaCl) drops out of the supersaline waters. After halite, the bittern salts (potassium or magnesium sulfates/chlorides) precipitate from supersaline waters at concentrations more than 70–90 times that of the original seawater. If calcium sulfate is still precipitating at these higher salinities, it tends to be as anhydrite rather than gypsum. By this bittern stage of seawater concentration, the brine density is in excess of 1.30 g/cm^3 and the brines have a viscosity similar to that of liquid honey. *See* ARAGONITE; WATER DESALINATION.

The increasing density of the concentrating brine increases the likelihood that it will slowly seep into underlying sediments where it alters its surroundings via a process known as brine reflux. A common outcome of brine reflux in many ancient seawater-fed evaporite basins is the conversion of underlying limestone into dolomite via the process of reflux dolomitization. *See* DOLOMITE; LIMESTONE.

Evaporation of continental waters (from both river and ground-water inflow) precipitates a less predictable suite of evaporite salts. Unlike marine-derived brines, evaporating continental waters draw on a more diverse reservoir, and so ionic proportions in the various brines are far more variable (**Table 3**). The varying ionic proportions in the final brine depend on the types of rock that have been leached by the inflowing waters. Flow through limestone aquifers produces inflow waters rich in calcium and bicarbonate (HCO_3); dolomite weathering and dissolution generate magnesium (Mg); igneous and metamorphic matrices yield silica-rich Ca-Na-HCO_3 waters. Pyritic shales and other sulfide-rich sediments will contribute sulfate ions, whereas basic and ultrabasic rocks tend to produce alkaline Mg-HCO_3 waters. *See* IGNEOUS ROCKS; METAMORPHIC ROCKS; PYRITE.

In all, H. P. Eugster and L. A. Hardie (1978) distinguished five major water types that precipitate

TABLE 2. Mineral paragenesis and classification of concentrating seawater*

Brine stage (hypersaline)	Mineral precipitate	Salinity, ‰	Degree of evaporation	Water loss, %	Density, g/cm^3
Normal marine	Seawater	35	1x	0	1.04
Mesohaline or vitahaline	Alkaline-earth carbonates	35–140	1–4x	0–75	1.04–1.10
Penesaline	$CaSO_4$ (gypsum/anhydrite)	140–250	4–7x	75–85	1.10–1.126
Penesaline	$CaSO_4 \pm$ halite	250–350	7–11x	85–90	1.126–1.214
Supersaline	Halite (NaCl)	>350	>11x	>90	>1.214
Supersaline	Bittern salts, K-Mg Salts	Variable	>60x	>99	>1.29

*After J. K. Warren, *Evaporites: Their Evolution and Economics*, Blackwell Scientific, Oxford, 1999.

TABLE 3. Typical brine compositions of seawater and various salt lake waters

	Na	Ca	Mg	K	Cl	SO_4	HCO_3
Seawater	10,762	410	1,293	399	19,353	2,709	142
Bristol Dry Lake, U.S.A.	57,400	43,330	1,070	3,330	173,000	210	—
Coastal salina, South Australia	99,700	600	10,700	2,700	171,100	7,200	200
Dallol Salt Pan, Danakil Africa	46,020	23,290	16,570	8,860	168,560	0	—
Dead Sea, Middle East (lower brine)	39,700	17,180	42,430	7,590	219,250	420	220
Freefight Lake, Canada (lower brine)	48,531	395	15,192	3,734	10,316	118,318	15,360
Great Salt Lake, U.S.A.	105,120	420	10,420	8,480	181,730	21,730	—
Laguna de Tirez, Southeast Spain	68,000	1,100	45,000	2,600	27,500	116,000	440
Lake Assal, northeast Africa	50,300	8,310	6,760	3,100	113,600	860	52
Lake Magadi, Africa	119,500	—	—	1,470	71,750	2,100	3,250
Lake Van, Turkey	7,740	10	90	510	5,450	2,340	2,190

evaporite beds in closed continental evaporite basins: (1) Ca-Mg-Na-(K)-Cl, (2) Na-(Ca)-SO_4-Cl, (3) Mg-Na-(Ca)-SO_4-Cl, (4) Na-CO_3-Cl, (5) Na-CO_3-SO_4-Cl. As any one of these waters concentrates within a particular evaporite basin, it deposits a characteristic suite of evaporite minerals (**Table 4**). First precipitates are the alkaline-earth carbonates: low-magnesian calcite, high-magnesian calcite, aragonite, and dolomite. The mineralogy of this initial precipitate depends on the Mg/Ca ratio of the parent brine. The proportions of calcium, magnesium, and bicarbonate ions in the brackish inflow waters determine the subsequent evaporation pathway of the brine. *See* CALCITE; CARBONATE MINERALS.

Very ancient marine evaporites. It is chemically impossible for sodium carbonate minerals to form by evaporation of modern seawater. Assuming that the proportions of calcium to sulfate and bicarbonate in seawater have not changed much in the last 600–800 million years, the presence of sodium carbonate minerals such as trona or its pseudomorphs (minerals of the same outward crystal form) in ancient evaporite beds would indicate nonmarine settings. In the oceans of the very early Earth (Archean), however, the proportions of dissolved ionic components were different than today as levels of methane and carbon dioxide in the atmosphere were much higher and the atmosphere was highly reducing (little or no free oxygen). Back then, solar evaporation of the earliest oceans precipitated trona along with halite; gypsum did not form. The various types of evaporite pseudomorphs retained in Archean sediments support this rather different seawater brine evolution. Halite pseudomorphs occur in many evaporitic early-Precambrian rocks 2–4 billion years ago. Widespread gypsum/anhydrite pseudomorphs become commonplace only around 1700 million years ago as rising levels of atmospheric oxygen meant that the oxidized form of sulfur, the sulfate ion, became a stable ionic phase in the world's oceans. Only pseudomorphs, rather than beds of the actual salts, are preserved in rocks this old. The oldest actual evaporite minerals are found in Neoproterozoic beds deposited some 600–700 million years ago in Oman and Australia. *See* ARCHEAN; ATMOSPHERE; CARBON DIOXIDE; EARTH; LAVA; METHANE; PRECAMBRIAN; PROTEROZOIC.

TABLE 4. Major evaporite minerals associated with the five brine types

Brine type	Saline mineral*
Ca-Mg-Na-(K)-Cl	Antarcticite
	Bischofite
	Carnallite
	Halite
	Sylvite
Na-(Ca)-SO_4-Cl	Techyhydrite
	Glauberite
	Gypsum
	Halite
	Mirabilite
	Thernadite
Mg-Na-(Ca)-SO_4-Cl	Bischofite
	Bloedite
	Epsomite
	Glauberite
	Gypsum
	Halite
	Hexahydrite
	Kieserite
	Mirabilite
	Thernadite
Na-CO_3-Cl	Halite
	Nahcolite
	Natron
	Thermonatrite
	Trona
Na-CO_3-SO_4-Cl	Burkeite
	Halite
	Mirabilite
	Nahcolite
	Natron
	Thernadite
	Thermonatrite

*For formulas, see Table 1.

Modern distribution. Modern bedded evaporite deposits typically accumulate in saline lakes and playa mudflats within ground-water discharge regions in the arid and semiarid deserts of the world. Coastal deposits of evaporites also occur in the same desert belts in areas fed by seepage of seawater into isolated coastal depressions or mudflats (coastal sabkhas). Continental playa deposits, such as Lake Eyre and Salar de Uyuni, typically contain much larger areas and volumes of salts than modern coastal deposits, but still do not approach the aerial extent or thicknesses of their ancient counterparts. Evaporite salts also form as lake precipitates and efflorescences in the cold polar deserts of Antarctica, but the amount of salt in these regions pales to insignificance when compared with the volumes forming in settings closer to the Equator. Brine mixing and freezing, rather than direct solar concentration, play

a much more important role during crystallization of the polar deposits.

Salt-accumulating playas define discharge areas within internal drainage basins (endoheic basins); they are areas where more water is leaving the basin via evaporation than is entering it as rainfall, surface, or subsurface inflow. Modern evaporite-filled playas and lakes mostly form in deserts, in two belts located between 15 and 45° north and south of the Equator. Due to changing atmospheric circulation and climate distribution, almost all modern continental playas, in both the Northern and Southern hemispheres, have experienced numerous water-full (humid) versus dry (arid playa) stages in the last few hundred thousand years of glacial-interglacial interactions. *See* DESERT; GLACIAL EPOCH.

Ancient distribution. Due to the high solubility of evaporite minerals, almost all sizable ancient evaporite deposits are less than 600–800 million years old. This is not a result of different intensities of dissolution; otherwise, evaporites would be more common today than, say, in the Permian. Rather, there have been times and tectonic configurations in the past when evaporite deposition was much more commonplace and much more widespread than it is today. At distinct times in the past—early Cambrian of Siberia, Permian of the southern United States and northwest Europe, Jurassic of the Gulf of Mexico and the Arabian Gulf, the Late Miocene in the Mediterranean—whole seaways were characterized by the deposition of widespread evaporites. The scale and depositional diversity of these ancient evaporites far exceed that seen in modern deposits; at times evaporites were precipitated across whole sedimentary basins as units hundreds of meters thick and thousands of kilometers wide. *See* CAMBRIAN; EUROPE; MIOCENE; PERMIAN.

For example, the Late Miocene some 5.5 million years ago was the last time that a large evaporite-covered basin formed on the Earth's surface; it occupied the position of the present Mediterranean Sea. At that time, the Mediterranean Sea had dried up into four major subbasins where thick sequences (0.3–2.0 km; 0.2–1.2 mi) of subaqueous gypsum, halite, and bittern salts were precipitating from seawater-derived brines. Almost all these evaporite beds are dominated by sedimentary textures whose present-day counterparts can only be found precipitating in shallow waters (<10 m or 33 ft deep) of much less extensive modern saline lakes.

The documentation in the 1970s of an ocean-wide shoalwater depositional setting for most Mediterranean evaporites revolutionized the understanding of how widespread ancient saline evaporites were deposited. Geologists realized that there is no modern counterpart of a desiccated oceanic basin; yet in the Late Miocene (Messinian) salinity crisis, the brine-filled depressions on the Mediterranean floor were filled by 2-km-thick (3-mi) units of halite and gypsum in less than 300,000 years. These widespread evaporites were laid down on the floors of a chain of shallow saline seas/lakes, with some of the water surfaces as much as 2000 m (6600 ft) below ambient sea level. The lakes were fed by a combination of seawater seepage and river inflow, but there was no surface connection with the adjacent Atlantic and Tethyan oceans. At that time a gigantic desert extended across the old sea floor from Spain to Israel and from northern Italy to Libya. The isolation that allowed the Mediterranean Ocean to dry up was created by the collision of Africa with Eurasia via plate tectonics. *See* PLATE TECTONICS.

Compared with evaporite deposits of today, the wider extent, greater depositional and chemical diversity, and greater thickness of evaporites deposited on ancient platforms and across whole basins were due to three main factors:

1. Times of warmer worldwide climate created wider latitudinal belts for evaporite deposition and preservation. For example, there is a clear diminution in evaporite-depositing latitudes from the Jurassic to the Cretaceous, coinciding with a change from an arid to a more humid worldwide climate. Yet, even in the Cretaceous, with its transition from evaporites to coal (indicating humid climate) at around 50° paleolatitude, the extent of evaporite deposition was still wider than today.

2. Huge shallow epeiric seas often formed large shelf and shallow inland seaways in arid areas, creating areas of extensive platform and intracratonic evaporites. This style of evaporite deposit was commonplace during times of long-term tectonic quiescence and during low-amplitude sea-level fluctuations of greenhouse Earth. It was often associated with deposition of extensive carbonate platforms. Shoal-water evaporite beds in the backreef of the Permian Basin of west Texas and New Mexico were deposited in this fashion, as was the Cretaceous Ferry Lake Anhydrite of east Texas; the Red River Formation of the Williston Basin, Wyoming; and the Red Heart Dolomite of the Georgina Basin, Australia.

3. The correct tectonic and climatic conditions existed for the formation of thick (>100 m or 330 ft) evaporite deposits extending across whole depositional basins (for example, the thick salt beds of the Castile/Salado formations of west Texas, and the Zechstein strata of northwest Europe). These basinwide domains formed during tectonically active times, when a combination of climate and tectonism restricted seawater inflow to vast ocean-margin and lacustrine basins that had formed by incipient rifting, continental collision, or transtensional faulting. The last such saline giant formed 5.5 million years ago, when the collision of Africa with Eurasia converted the Mediterranean-Tethyan Sea into a chain of shoal-water evaporite lakes and the Zagros Basin into an evaporite-filled depression. John Warren

Bibliography. G. I. Alsop, D. J. Blundell, and I. Davison, *Salt Tectonics*, Geol. Soc. Spec. Publ., vol. 100, London, 1996; N. P. James and A. C. Kendall, Introduction to carbonate and evaporite facies models, in R. G. Walker and N. P. James (eds.), *Facies Models: Responses to Sea Level Change*, pp. 265–275, Geological Association of Canada, 1992; J. R. Li et al., Responses of evaporite mineralogy to inflow

water sources and climate during the past 100 Ky in Death Valley, California, *Geol. Soc. Amer. Bull.*, 109(10):1361–1371, 1997; J. L. Melvin (ed.), *Evaporites, Petroleum and Mineral Resources*, Developments in Sedimentology Series, vol. 50, Elsevier, Amsterdam, 1991; M. G. Rowan, Structural styles and evolution of allochthonous salt, central Louisiana outer shelf and upper slope, in M. P. A. Jackson, D. G. Roberts, and S. Snelson (eds.), *Salt Tectonics: A Global Perspective*, pp. 199–228, Amer. Ass. Petrol. Geol. Mem., no. 65, 1995; B. C. Schreiber (ed.), *Evaporites and Hydrocarbons*, Columbia University Press, New York, 1988; J. K. Warren, *Evaporites: Their Evolution and Economics*, Blackwell Scientific, Oxford, 1999.

Salmonelloses

Diseases caused by *Salmonella*. These include enteritis (90–95% of all salmonelloses in the United States) and septicemia with or without enteritis (5–10%). *Salmonella typhi*, *S. paratyphi A*, *B*, and *C*, and occasionally *S. cholerae suis* cause particular types of septicemia called typhoid and paratyphoid fever, respectively; while all other types may cause enteritis or septicemia, or both together.

Typhoid fever. This type has an incubation period of 5–14 days. Typhoid fever is typified by a slow onset with initial bronchitis, diarrhea or constipation, a characteristic fever pattern (increase for 1 week, plateau for 2 weeks, and decrease for 2–3 weeks), a slow pulse rate, development of rose spots, swelling of the spleen, and often an altered consciousness; complications include perforation of the bowel and osteomyelitis. *Salmonella typhi* can be isolated from the blood in the first 10 days and later from the feces and urine. The ileum shows characteristic ulcerations. Fecal excretion of *S. typhi* usually ends by the sixth week, but 2–5% of convalescents, mostly women, become chronic carriers. The organism may also persist in the gallbladder.

Typhoid fever leaves the individual with a high degree of immunity. Antibodies in the serum are detectable from the second week on: Anti-O antibodies signify a present infection, whereas anti-H antibodies, which appear later, are observed in convalescence, after vaccination, or after an earlier infection with an antigenically (H) related *Salmonella* species. Anti-Vi is often observed in carriers. Both anti-H and anti-Vi tend to persist. *See* ANTIBODY; IMMUNOLOGY.

Vaccination with an oral vaccine gives an individual considerable protection for about 3 years, which may be broken by large oral challenge doses. *See* VACCINATION.

The only effective antibiotic is chloramphenicol, although a few resistant strains have been reported in Mexico. Ampicillin may be helpful but is more useful in treating carriers. If gallbladder disease is present, cholecystectomy should be considered in carriers.

Preventive measures must concentrate on sanitation, since this exclusively human salmonella is transmitted most often by fecal contamination of water or foodstuffs. *See* ANTIBIOTIC; WATER-BORNE DISEASE.

Paratyphoid fever has a shorter course and is generally less severe than typhoid fever. Vaccination is an ineffective protective measure.

Enteric fevers. Enteric fevers, that is, septicemias due to types of *Salmonella* other than those previously mentioned, are more frequent in the United States than typhoid and paratyphoid fever but much less frequent than *Salmonella* enteritis. In children and in previously healthy adults, enteric fevers are most often combined with enteritis and have a favorable outlook. In certain predisposed individuals (for example, those under adrenocortical steroid treatment or those suffering from sickle cell anemia, malaria, or leukemia), septicemia with or without enteritis and metastatic infection in other organs may occur; the prognosis then depends on the underlying illness. The organisms involved are the same as those causing *Salmonella* enteritis. Chloramphenicol or ampicillin are used in treatment. However, strains resistant to both drugs have been observed. *See* DRUG RESISTANCE.

Enteritis. Inflammation of the small bowel due to *Salmonella* is one of the most important bacterial zoonoses. The most frequent agents are *S. typhimurium*, *S. enteritidis*, *S. newport*, *S. heidelberg*, *S. infantis*, and *S. derby*. The incubation period varies from 6 h to several days. Diarrhea and fever are the main symptoms; the intestinal epithelium is invaded, and early bacteremia is probable. Predisposed are persons with certain preexisting diseases (the same as for enteric fevers), very old and very young individuals, and postoperative patients. Chronic carriers exist but are rare in comparison with post-typhoid carriers. The main reservoir is animals, with transmission occurring chiefly through foodstuffs.

As would be expected, enteritides occur more frequently during the summer months. There is no immunity to this salmonellosis. Antibody determinations in the patient's serum may or may not yield positive results. In spite of laboratory effectiveness of many antibiotics, antimicrobial treatment serves only to prolong the carrier state and has no effect on the disease; neither has vaccination. Prevention must concentrate on improving sanitation in commercial food production (for example, improvements in slaughtering methods, slaughterhouse conditions, and transport facilities; and such measures as exclusion of carriers and decontamination of offal and powdered bulk food such as egg powder).

Alexander von Graevenitz

Bibliography. E. J. Bowmer, The challenge of salmonellosis, *Amer. J. Med. Sci.*, 247:467, 1964; F. Cabello (ed.), *Biology of Salmonella*, 1993; J. I. Cohen et al., Extraintestinal manifestations of *Salmonella* infections, *Medicine*, 66:349, 1987; E. van Oye (ed.), *The World Problem of Salmonellosis*, 1964.

Salmoniformes

An order of generalized fishes, including the salmons and trouts, characterized by soft or articulated fin rays; the presence (usually) of an adipose dorsal fin; cycloid scales (scales absent in few species); pelvic fins in the abdominal region and each with more than six rays and an associated axillary process (absent in Osmeridea); a pelvic girdle free from the pectoral girdle; a pectoral fin placed low on the side and more or less horizontal; an upper jaw usually bordered by premaxillae and maxillae; nonprotractile premaxillae; the presence (usually) of an orbitosphenoid bone as well as a mesocoracoid arch in the pectoral girdle; the absence of Weberian apparatus connecting the swim bladder with the inner ear; a physostomous (connected to the esophagus) air bladder; the larvae are not leptocephalous-like; and there are no luminescent organs.

In the past several groups presently considered distinct orders—namely, Argentiniformes, Osmeriformes, Salmoniformes, Esociformes, Stomiiformes, and Myctophiformes—were placed, in various combinations, in Salmoniformes. The last two named orders are Neoteleosts; the remaining four orders constitute the superorder Protacanthopterygii. Herein the osmerids and salmonids are arbitrarily united in Salmoniformes as follows:

Order Salmoniformes
- Superfamily Osmeridea
  - Family Osmeridae
    - Salangidae
    - Sundasalangidae
- Superfamily Galaxioidea
  - Family Retropinnidae
    - Lepidogalaxiidae
    - Galaxiidae
- Superfamily Salmonoidea
  - Family Salmonidae (trouts and salmons)
    - Subfamily Coregoninae
      - Thymallinae
      - Salmoninae

Osmeridae (smelts). Smelts are small, silvery, minnow-like fishes, distinguished by a large mouth, large teeth in the jaws, few pylonic caeca or none, and deciduous scales. Absence of an axillary process associated with the pelvic fins readily distinguishes smelts from the salmons and trouts. They occupy cold and temperate marine waters of the Northern Hemisphere, namely the Arctic, Atlantic, and Pacific oceans. Several species are known to complete their life cycle in freshwater. Others spawn in the lower reaches of rivers from where the larvae drift seaward. Rainbow smelt (*Osmerus mordax*), a species in North America, spawn on wave-swept ocean beaches where they create excitement among locals, who find the fish easy to harvest. There are seven genera and 13 species.

Salanagidae (icefishes). Icefishes are transparent or translucent fishes with a poorly ossified skeletons, flattened head, and adipose fin, and are virtually scaleless. They are anadromous (live in salt water but spawn in freshwater) and freshwater fishes known from Sakhalin Island, off the coast of Japan, and from Korea, China, and northern Vietnam, where they are locally important food fishes. There are four genera and 11 species.

Sundasalangidae (noodlefishes). Noodlefishes are similar to icefishes but lack an adipose fin. They occur in the fresh and brackish waters of Borneo and southern Thailand. Sexually mature at less than 15 cm (6 in.), Noodlefishes are among the smallest protacanthopterygian. There is one genus and two species.

Retropinnidae (New Zealand smelts). Small plankton-feeding fishes rarely exceeding 13 cm (5 in.), New Zealand smelts differ from the above families by having after an adipose fin, a hard keel on their belly, and lacking the right gonad. They emit a cucumber-like odor when captured. Occuring in southeastern Australia, Tasmania, and New Zealand, they complete their life cycle in freshwater; however, some species spend early stages in the marine environment and return to freshwater to spawn. There are six genera and 10 species.

Lepidogalaxiidae (salamanderfishes). The single species (*Lepidogalaxis salamandroides*) in this family is an extraordinary fish. Its eyes lack muscles. Compensating for the inability to rotate the eye, it has the ability to move its head downward and sideways, which is most unusual for a fish. Endemic to a small area in southwestern Australia, salamanderfishes can survive periodic droughts by burrowing in the sand.

Galaxiidae (galaxiids). The following combination of characteristics serves to distinguish galaxiids from other salmoniforms: absence of maxillary, vomer, palatine, and basibranchial teeth; presence of paired gonads; absence of an adipose fin in most species; and lack of a cucumber-like odor. While most species spend their entire lives in freshwater, all species breed in freshwater. A few are catadromous (live in freshwater and spawn in salt water), the larvae and juveniles spending some time at sea. Some survive dry season by aestivating. *See* AESTIVATION AND HIBERNATION.

Galaxiidae is strictly a Southern Hemisphere family, with an interesting distribution. On a distribution map their range indications appear as small blobs skipping from Lord Howe Island and New Caledonia to New Zealand, across wide spans of ocean to Southern Australia, then to the tip of South Africa, and farther across the south Atlantic to the tip of South America. Galaxiids are the dominant freshwater fishes where they occur. There are eight genera and 40 species.

Salmonidae. Salmonids are distinguished from the osmerids and galaxiids by the following combination of characters: large adult size; more robust body; dorsal fin near the middle of the back, the origin of the dorsal fin is over or in advance of a vertical line drawn from the origin of the pelvic fins; presence of pelvic axillary process and adipose fin; gill membranes extending far forward and free from the

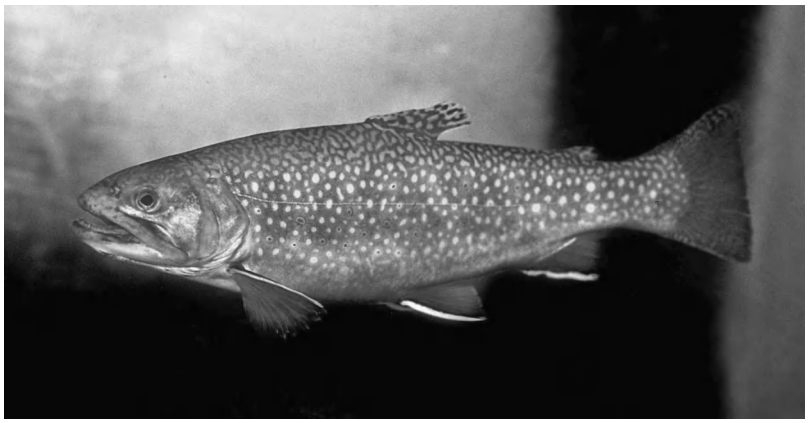

Brook trout (*Salvelinus fontinalis*). (*Photo courtesy of Johnny Jensen; © www.jjphoto.dk*)

isthmus; and parr with vertical bars in most species. The family consists of three subfamilies.

Coregoninae (whitefishes and ciscoes). Coregonids have fewer than 16 dorsal fin rays; large scales, fewer than 110 in lateral series; toothless maxilla; and usually small toothless vomer. They range from Asia and Siberia to northern North America. There are three genera and 32 species, three of which are endemic to Bear Lake in Utah-Idaho.

Thymallinae (graylings). A long, high dorsal fin with more than 17 rays distinguishes the graylings. Four species represent the single genus, *Thymallus*, in the freshwaters of Europe, Mongolia, northern Asia, and North America.

Salmoninae (salmons and trouts). The salmonids are distinguished by having fewer than 16 dorsal fin rays, small scales numbering 110 or more (usually many more), and maxillary teeth.

Generally the large seagoing anadromous species that return to their natal stream to spawn and die are called salmons, whereas the smaller species that live to spawn several seasons are called trouts (see **illustration**). All species spawn in freshwater. An ultrasensitive sense of smell and the ability to fix in their memory the odors of streams they traverse on their journey to sea serve as a "road map" for salmons to return to their place of birth.

Salmons and trouts are Holarctic in native distribution; however, species such as *Salvelinus fontinalis*, *Salmo trutta*, and *Oncorhynchus mykiss* (long known as *Salmo gairdneri*) have been introduced throughout the world. Worldwide salmons and trouts are probably more studied, more fished for recreation purposes, more exploited commercially, more genetically manipulated, and more introduced in foreign countries than any other group of fishes. These fishes are very important economically. *See* AQUACULTURE.

There are seven genera and about 30 species. Ongoing studies will probably add to the numbers.

Herbert Boschung

Bibliography. R. J. Behnke, *Native Trout of Western North America.* Amer. Fish. Soc. Monog. 6, American Fisheries Society, Bethesda, Maryland, 1992; R. J. Behnke, *Trout and Salmon of North America*, Free Press, New York, 2002; T. M. Berra, *An Atlas of Distribution of Freshwater Fish Families of the World*, University of Nebraska Press, Lincoln, 1981; R. H. Smith, *Native Trout of North America*, 2d ed., Frank Amato Publications, Portland, Oregon, 1994.

Salt (chemistry)

A compound formed when one or more of the hydrogen atoms of an acid are replaced by one or more cations of the base. The common example is sodium chloride in which the hydrogen ions of hydrochloric acid are replaced by the sodium ions (cations) of sodium hydroxide. There is a great variety of salts because of the large number of acids and bases which has become known.

Classification. Salts are classified in several ways. One method—normal, acid, and basic salts—depends upon whether all the hydrogen ions of the acid or all the hydroxide ions of the base have been replaced:

Class	*Examples*
Normal salts	$NaCl$, NH_4Cl, Na_2SO_4, Na_2CO_3, Na_3PO_4, $Ca_3(PO_4)_2$
Acids salts	$NaHCO_3$, NaH_2PO_4, Na_2HPO_4, $NaHSO_4$
Basic salts	$Pb(OH)Cl$, $Sn(OH)Cl$

The other method—simple salts, double salts (including alums), and complex salts—depends upon the character of completeness of the ionization:

Class	*Examples*
Simple salts	$NaCl$, $NaHCO_3$, $Pb(OH)Cl$
Double salts	$KCl \cdot MgCl_2$
Alums	$KAl(SO_4)_2$, $NaFe(SO_4)_2$, $NH_4Cr(SO_4)_2$
Complex salts	$K_3Fe(CN)_6$, $Cu(NH_3)_4Cl_2$, $K_2Cr_2O_7$

In general, all salts in solution will give ions of each of the metal ions; an exception is the complex type of salt such as $K_3Fe(CN)_6$ and $K_2Cr_2O_7$. In such salts the ionization is entirely as shown by reactions (1) and (2). No detectable quantities of Fe^{3+} or Cr^{3+} from

$$K_3Fe(CN)_6 \rightarrow 3K^+ + Fe(CN)_6{}^{3-} \qquad (1)$$

$$K_2Cr_2O_7 \rightarrow 2K^+ + Cr_2O_7{}^{3-} \qquad (2)$$

these salts exist in solution because of the strong bonding of these ions in the complex ions. However, in those complex salts where the bonding is weak, ions of the metal can be detected; for example, in Na_2CdCl_4, the cadmium complex ion ionizes appreciably as in reaction (3). The elements with unfilled

$$CdCl_4{}^{2-} \rightarrow Cd^{2+} + 4Cl^- \qquad (3)$$

inner electron shells form complex salts readily. The alum type of salt is a sulfate including the univalent

cation of a relatively strong base and a trivalent metal ion such as Al^{3+}, Fe^{3+}, or Cr^{3+}.

Double salts include ions of nearly enough the same size to fit into the same crystal lattice.

Hydrolysis. The solutions of some normal salts are neutral, but those of others are acidic or basic. This results from the reactor of the ions of salt with water. This reaction is called hydrolysis. Examples are shown in reactions (4)–(6). The resulting solu-

$$CO_3^{2-} + HOH \rightleftharpoons HCO_3^- + OH^- \quad (4)$$

$$CH_3COO^- + HOH \rightleftharpoons CH_3COOH + OH^- \quad (5)$$

$$NH_4^+ + HOH \rightleftharpoons NH_4OH + H^+ \quad (6)$$

tion will be acidic or basic, depending upon whether the hydrolysis produces an excess of hydrogen or hydroxide ion. *See* HYDROLYSIS.

Theories of acids and definition of salts. The development of more general theories of acids and bases in the twentieth century has required a broadening of the concept of salts.

The Brönsted theory of acids lays emphasis on the process of the reaction between acids and bases, and not so much on the product except that the products are other acids and bases, as in reactions (7) and (8).

$$\text{acid}_1 + \text{base}_2 \rightleftharpoons \text{acid}_2 + \text{base}_1 \quad (7)$$

$$HCl + H_2O \rightleftharpoons H_3O^+ + Cl^- \quad (8)$$

Since the Brönsted theory extends the proton theory of acids to solvents other than water, the original definition of a salt must be expanded as follows: A salt is an electrovalent compound that contains some cation other than the solvated proton and some anion other than the anion which is the conjugate base of the solvent. In the water system the salt should not contain the H_3O^+ and OH^- ions alone; in the liquid ammonia system it should not contain the NH_4^+ and NH_2^- ions alone.

In terms of the Lewis theory of acids and bases, compount (**1**) is an acid, and compount (**2**) is a base. Hence the salt should be compound (**3**). Here the salt

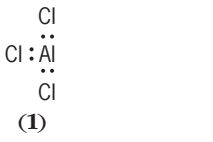

(**1**)

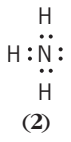

(**2**)

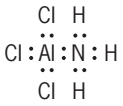

(**3**)

is not limited to replacement of the H^+ with a cation of a base. Rather, it is any aggregate of molecules, atoms, or ions joined together with a coordinate covalent bond. Such compounds correctly can be called salts; however, by common parlance, the term salt usually refers to an electrovalent compound, the classical example of which is sodium chloride. *See* ACID AND BASE; CHEMICAL BONDING. Alfred B. Garrett

Bibliography. S. Radel and M. Navidi, *Chemistry*, 2d ed., 1994.

Salt (food)

The chemical compound sodium chloride. It is used extensively in the food industry as a preservative and flavoring, as well as in the chemical industry to make chlorine and sodium. Historically, salt is one of the oldest materials used in people's food. *See* CHLORINE; SALT (CHEMISTRY); SODIUM.

Manufacture. Salt was originally made by evaporating seawater (solar salt). This method is still in common usage; however, impurities in solar salt make it unsatisfactory for most commercial uses, and these impurities also lead to clumping. Salt, freshly produced from seawater evaporation ponds, may contain large numbers of halophilic (salt-loving) microorganisms. These occasionally cause spoilage of meat, fish, vegetables, and hides when salt has been used in preservation.

In the United States refined salt is obtained from underground mines located in Michigan and Louisiana. Salt is usually handled during the refining processes as brine. These processes are discussed below.

Grainer salt. This type of salt is made by evaporation of brine in long shallow pans, as large as 18 ft (5.5 m) wide, 1.5 ft (4.6 m) deep, and 150 ft (46 m) long. The daily capacity of such a grainer may be 80 tons (73 metric tons). A scraping conveyor continually removes the crystallizing salt from the bottom of the grainer. The salt is then filtered, dewatered, dried, cooled, and rolled to break clumps. Grainer salt is usually the coarsest in grain and highest in impurities.

Vacuum pan salt. Salt brine is boiled at reduced pressure. A triple-effect evaporator is used; the first stage uses relatively light vacuum, but this is increased until in the third it is quite high and the salt solution boils at about 110°F (43°C). Production is continuous and the production cycle takes 48 h. A 20% salt slurry is brought out from the bottom of the third evaporator at the same time fresh brine is admitted; thus impurities are washed from the surface of the outgoing crystals. The salt slurry is filtered, dewatered, and high-temperature-dried at 350°F (177°C) before screening and packing.

Alberger process. Salt brine is heated to high pressures in heaters and then is passed to a graveller. A graveller is a large cylindrical vessel filled with stones which serves as a deposition site for calcium sulfate. The brine proceeds to flashers, where the pressure is gradually reduced to that of the atmosphere, and salt begins to crystallize. The brine and salt mixture is discharged to a large open pan; the crystallized salt is pumped to a centrifuge for dewatering, then dried in a rotary dryer.

Use. Large users of salt in the food industry are pickle makers and meat packers. In the pickle industry salt is used as brine to which fresh cucumbers are added. A selective fermentation then proceeds

which is governed by salt concentration. Cost and time are dictating a movement toward "fresh-pack" pickles, in which salt and spices are added with the fresh cucumber, the jar is sealed, and the whole package is heated for pasteurization preservation. In the meat packing industry salt is added to fresh meat as a preservative, as in salt pork, or along with nitrates or nitrites as the first step in the production of cured meats, such as hams or bacon. *See* FOOD ENGINEERING.

Additives. Salt is liable to clumping during periods of high humidity, so preventives are added. Materials used include magnesium carbonate and certain silicates. Iodides are added in those areas where iodine deficiencies exist. Roy E. Morse

Bibliography. T. P. Coultate (ed.), *Food: The Chemistry of Its Components*, 3d ed., 1996; B. Kraus, *Complete Guide to Sodium*, 1990; Y. Pomeranz and C. E. Meloan, *Food Analysis: Theory and Practice*, 3d ed., 1994.

Salt dome

An upwelling of crystalline rock salt and its aureole of deformed sediments. A salt pillow is an immature salt dome comprising a broad salt swell draped by concordant strata. A salt stock is a more mature, pluglike diapir of salt that has pierced, or appears to have pierced, overlying strata. Most salt stocks are 0.6–6 mi (1–10 km) wide and high. Salt domes are closely related to other salt upwellings (**Fig. 1**), some of which are much larger. Salt canopies, which form by coalescence of salt domes and tongues, can be more than 200 mi (300 km) wide. *See* DIAPIR.

Distribution. Exploration for oil and gas has revealed salt domes in more than 100 sedimentary basins that contain rock salt layers several hundred meters or more thick. The salt was precipitated from evaporating lakes in rift valleys, intermontaine basins, and especially along divergent continental margins. Salt domes are known in every ocean and continent (except little-explored Antarctica) but are scarce in the Pacific Ocean, where few salt-bearing basins have escaped destruction by plate convergence. *See* BASIN; PLATE TECTONICS.

Composition. Salt domes consist largely of halite (NaCl, common table salt). Other evaporites, such as anhydrite ($CaSO_4$), gypsum ($CaSO_4 \cdot 2H_2O$), and more rarely sylvite (KCl), carnallite ($KCl \cdot MgCl_2 \cdot 6H_2O$), and other potash and magnesian sulfates and chlorides, form thinner layers within the rock salt. Nonevaporite inclusions range from disseminated grains to massive rafts 2 mi (4 km) wide. Salt is highly soluble, so it is remarkable that salt as much as 800 million years old is still preserved in domes in Iran, Oman, and Pakistan. *See* HALITE; SALINE EVAPORITES.

Subsurface dissolution by ground water leaves a compacted residue, known as cap rock, which mantles the crest and shoulders of the salt dome. Residual anhydrite typically undergoes reactions involving water, hydrocarbons, and bacteria to yield gypsum, calcite ($CaCO_3$), sulfur, and sulfides of iron, zinc, lead, and barium. *See* ANHYDRITE; CALCITE; GYPSUM.

Formation. Crystalline salt flows by creep as a viscous or power-law fluid. Creep is promoted by fine

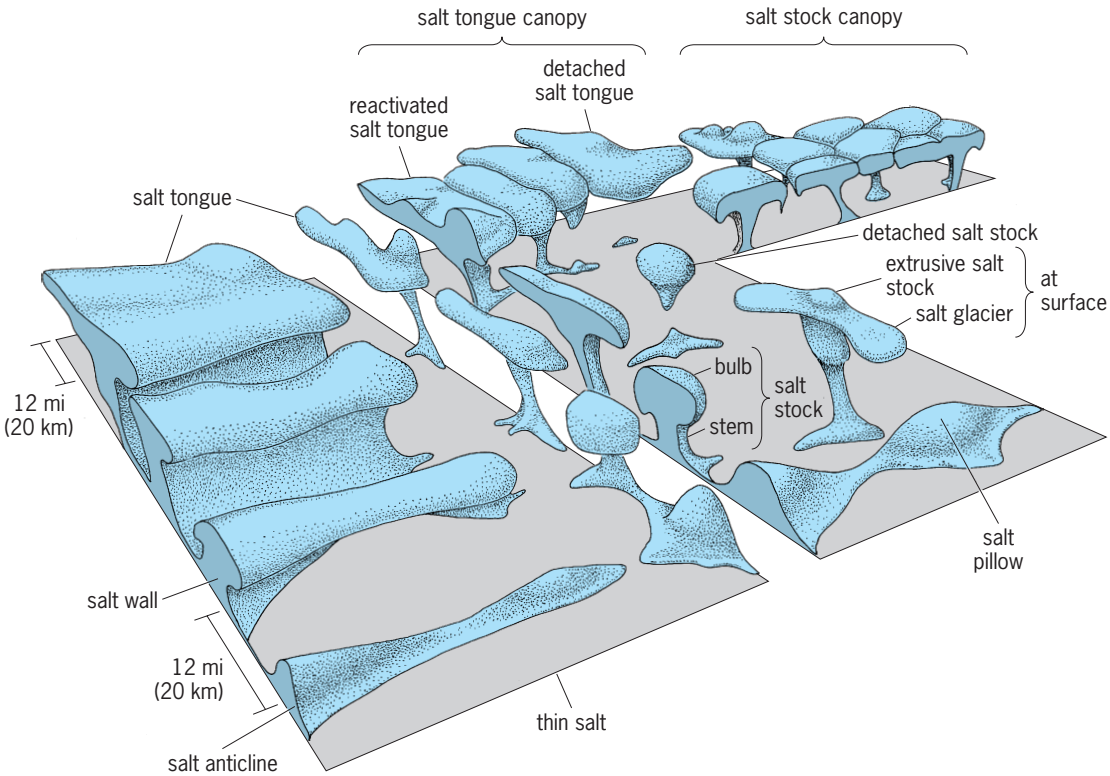

Fig. 1. Salt domes and related structures; the overlying and surrounding sediments have been removed for clarity.

grain size, heat, and traces of water. Salt acts as a weak but pressurized geologic lubricant that allows stronger strata above and below to deform partly independently of each other. Salt thus facilitates deformation, whether driven by plate tectonics or merely by gravity (the process of halokinesis). *See* CREEP (MATERIALS).

Salt diapirs are initiated and driven upward by differential loading: laterally varying sedimentary loads are created by erosion, deformation, or uneven sedimentation. As salt diapirs pierce overlying strata, they become discordantly encased in brittle rocks. Three piercement modes are recognized: reactive, active, and passive (**Fig. 2**). Reactive diapirs are emplaced while a basin is stretched (Fig. 2*a*). Salt upwells into structurally thinned zones, such as grabens. Reactive rise stops when regional stretching ceases. If this diapir becomes tall enough and its roof strata thin enough, the diapir becomes active (Fig. 2*b*). The pressure of salt is enough to break through and shoulder aside the thin roof. Salt then emerges under air or water (Fig. 2*c*). This passive diapir remains at or near the surface while sediments accumulate around it. The passive diapir grows taller because its base sinks below the surface along with the surrounding strata, in a process known as downbuilding. A passive diapir is a slow-moving fountain that can extrude, even under water, to form salt glaciers. These are flowing today in the Zagros Mountains of Iran. The many salt tongues and sheets in the Gulf of Mexico are thought to have originated as submarine salt glaciers, partly protected from dissolution by residual cap rock and veneers of marine sediment. Spreading salt diapirs can coalesce to form a variety of canopies (Fig. 1). These buried extrusions can act as source layers for younger generations of salt domes. The lower parts of salt walls can be squeezed together by compression: the salt is displaced upward, where it bulges its roof. The closing sides of the squeezed salt wall can eventually meet to form a salt weld in which only a smear of salt remains or vanishes entirely. *See* GRABEN.

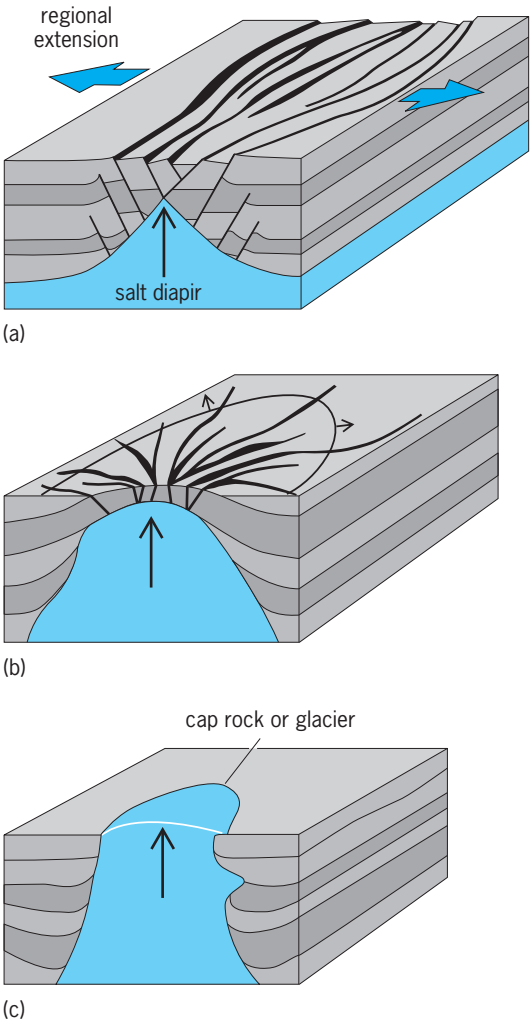

Fig. 2. Three modes of salt diapirism and their characteristic structures. (*a*) Reactive piercement. (*b*) Active piercement. (*c*) Passive piercement.

Detection. Salt domes exposed at the surface have been known for millennia as sources of salt, a historically prized commodity. In the search for salt or hydrocarbons, buried salt domes were discovered by increasingly sophisticated means: by searching for mounds, saline or sulfurous springs, and hydrocarbon seeps; by drilling; by conducting gravity surveys that respond to the low density of salt and the high density of cap rock, and magnetic surveys, where salt is less magnetic than surrounding sediments; and by seismic refraction and reflection surveys that recognize salt by its high sonic velocity and lack of internal reflections. *See* PETROLEUM GEOLOGY; PROSPECTING; SEISMIC EXPLORATION FOR OIL AND GAS.

Economic importance. Salt domes supply industrial commodities, including fuel, minerals, chemical feedstock, and storage caverns. Giant oil or gas fields are associated with salt domes in many basins around the world, especially in the Middle East, North Sea, and South Atlantic regions. The most intensively explored of these basins is the United States Gulf Coast and Gulf of Mexico, where 80% of the oil and gas fields are associated with salt structures and form the United States' largest exploration province. Salt canopies obscure vast regions of little-explored potential hydrocarbons. The 1901 oil gusher at Spindletop salt dome in Texas initiated the world's continuing dependence on—and exploration for—prolific liquid fossil fuels. Oil and gas are trapped around salt domes by arching, fracturing, or faulting of surrounding strata, or by sedimentary facies changes caused by dome rise, and are sealed in traps by the impermeable salt itself.

The cap rock of salt domes provides the main source of elemental sulfur, which is readily extracted by the Frasch process. Salt domes on the United States Gulf Coast have yielded 65% of the world's sulfur production. About 65% of the United States salt is mined as rock or as brine dissolved from salt domes. Potassium salts in German salt domes are a major source of potash for fertilizers. *See* FERTILIZER; SULFUR.

Salt domes are also used to store crude oil, natural gas (methane), liquefied petroleum gas, and

radioactive or toxic wastes. Bottle-shaped caverns up to 2000 ft (600 m) tall can be created by dissolving salt with water injected down a well into the dome and extracting the resulting brine. More than 500 million barrels of crude oil are stored in United States salt domes as an emergency supply called the Strategic Petroleum Reserve. *See* OIL AND GAS STORAGE.

M. P. A. Jackson

Bibliography. G. I. Alsop, D. J. Blundell, and I. Davison, Salt Tectonics, *Geol. Soc. London Spec. Publ.*, no. 100, 1996; M. T. Halbouty, *Salt Domes: Gulf Region, United States and Mexico*, 2d ed., 1979; M. P. A. Jackson, D. G. Roberts, and S. Snelson (eds.), Salt Tectonics: A Global Perspective, *Amer. Ass. Petrol. Geol. Mem.*, no. 65, 1995; M. P. A. Jackson et. al., Structural dynamics of salt systems, *Annu. Rev. Earth Planet. Sci.*, 1994.

Salt-effect distillation

A process of extractive distillation in which a salt that is soluble in the liquid phase of the system being separated is used in place of the normal liquid additive introduced to the extractive distillation column in order to effect the separation.

Process. Extractive distillation is a process used to separate azeotrope-containing systems or systems in which relative volatility is excessively low. An additive, or separating agent, that is capable of raising relative volatility and eliminating azeotropes in the system being distilled is supplied to the column, where it mixes with the feed components and exerts its effect. The agent is subsequently recovered from one or both product streams by a separate process and recycled for reuse. In conventional extractive distillation, the agent is a liquid, usually one having fairly low volatility. An example is ethanol-water separation using ethylene glycol as the added agent. *See* AZEOTROPIC MIXTURE.

In salt-effect distillation, the flow sheet is essentially the same as for a liquid agent, although the subsequent process used to recover the agent for recycling is different; that is, evaporation is used rather than distillation. A typical flow sheet is shown in the **illustration**. The salt is added to the system by being dissolved in the reentering reflux stream at the top of the column. Being nonvolatile, it will reside in the liquid phase, flowing down the column and out in the bottom product stream. (An alternative method would be to use regular distillation up to the azeotrope composition, salt-effect distillation to cross it, and regular distillation to complete the separation.)

Liquid and salt agents work in a similar manner: their molecules or ions exert attractive forces, which result in the formation of association complexes in the liquid phase with the molecules of the feed components. An agent is used whose attraction is preferential toward the molecules of the less volatile component of the system being separated. Thus, the volatility of the higher-boiling component is reduced by an amount greater than that of the lower-boiling component. As a result, the relative volatility is increased, and the azeotrope is eliminated if the effect is sufficiently strong and sufficiently selective. *See* COORDINATION COMPLEXES.

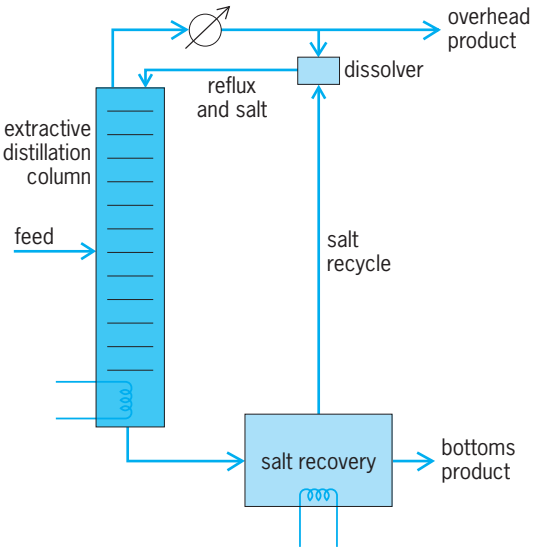

Flow sheet for salt-effect distillation. (*After W. F. Furter, Salt effect in distillation: A technical review, Chem. Eng., 219:173–177, 1968*)

Advantages and disadvantages. The advantage of salts over liquids as separating agents is that the forces of attraction exerted by their ions toward molecules of the higher-boiling feed component may be much stronger than those that can be exerted by the molecules of a liquid agent. With a salt, a much smaller quantity of separating agent may have to be fed, recovered, and recirculated for a particular separation than in the case of a liquid. This results in savings in costs of energy and capital, particularly in the separating-agent recovery process. For example, 6 mol % of potassium acetate present in the liquid phase will completely eliminate the ethanol-water azeotrope, whereas conventional azeotropic distillation of this system using a liquid agent such as benzene or *n*-pentane may require an agent-to-feed ratio ranging anywhere from 1:1 to 5:1.

A second advantage is that the salt, being completely nonvolatile, does not distribute itself partially into the vapor phase in the extractive distillation column. Therefore, no provision is required to recover agent from the overhead product stream. With a liquid agent, however, a so-called knockback section of column is usually required above the agent's feed point to recover that portion of the agent that has vaporized in the column and to prevent it from contaminating the overhead product.

The main disadvantage of a salt over a liquid agent is found in solubility limitations. It is generally much more difficult to find a salt that is both effective and soluble in the system being distilled than it is to find a liquid with these properties. Yet, in that limited number of systems for which a salt exists that has a strong and selective effect and is sufficiently soluble, the advantages of a salt over a liquid can be substantial, particularly in the saving of energy.

Applications. The first large-scale use of salt-effect distillation was in producing ethanol for blending

with gasoline to produce gasohol. Most gasoline sold in Europe from the late 1920s to World War II contained 10% ethanol, which was produced from sugarbeets and refined by the HIAG process, a salt-effect distillation that uses a mixture of potassium and sodium acetates as the separating agent. Over 100 HIAG plants were built worldwide, with the last ceasing operation in 1965, in Brazil. Interest in salt-effect distillation revived in the late 1970s and throughout the 1980s among gasohol producers in the United States and elsewhere. *See* ALCOHOL FUEL.

The major commercial use of salt-effect distillation is in the concentration of aqueous nitric acid, using the salt magnesium nitrate as the separating agent instead of sulfuric acid, which was the liquid agent used in earlier extractive distillation processes. Other commercial applications exist, including acetone-methanol separation using calcium chloride and isopropanol-water separation using the same salt.

Salt-effect distillation has not achieved a wider use in the past for several reasons: the chemical theory is very complex; engineers much prefer handling of fluids over handling of solids when there is a choice; and, most importantly, when energy is cheap there is little incentive for improved energy efficiency. With its strong advantages in energy economy over conventional extractive distillation, salt-effect distillation can be expected to grow as energy becomes increasingly costly. *See* AZEOTROPIC DISTILLATION; DISTILLATION. William F. Furter

Bibliography. W. F. Furter, Salt effect in distillation, Pt. II, *Can. J. Chem Eng.*, 55:229–239, 1977; H. Z. Kister, *Distillation Design*, 1992.

Salt gland

A specialized gland located around the eyes and nasal passages in marine turtles, snakes, and lizards, and in birds such as the petrels, gulls, and albatrosses, which spend much time at sea. Salt glands are compound tubular glands which apparently arise as invaginations of the epithelium of the nasal passage or from the developing conjunctival sac of the eye. In the marine turtle it is an accessory lacrimal gland which opens into the conjunctival sac. In seagoing birds and in marine lizards it opens into the nasal passageway. Salt glands copiously secrete a watery fluid containing a high percentage of salt, higher than the salt content of urine in these species. As a consequence, these animals are able to drink salt-laden seawater without experiencing the dehydration necessary to eliminate the excess salt via the kidney route. *See* EXCRETION; GLAND. Olin E. Nelsen

Salt marsh

A maritime habitat characterized by grasses, sedges, and other plants that have adapted to continual, periodic flooding. Salt marshes are found primarily throughout the temperate and subarctic regions.

Fig. 1. Plant zonation in a salt marsh, Canary Creek Marsh, Lewes, Delaware: smooth cordgrass (*Spartina alterniflora*) at the left; salt-meadow hay in the middle; and marsh elder (*Iva frutescens*) and groundsel tree (*Baccharis halimifolia*) at the upland edge on the right. (*Courtesy of F. C. Daiber*)

The tide is the dominating characteristic of a salt marsh. The salinity of the tide defines the plants and animals that can survive in the marsh area. The vertical range of the tide determines flooding depths and thus the height of the vegetation, and the tidal cycle controls how often and how long vegetation is submerged. Two areas are delineated by the tide: the low marsh and the high marsh. The low marsh generally floods and drains twice daily with the rise and fall of the tide; the high marsh, which is at a slightly higher elevation, floods less frequently. *See* MANGROVE.

Marsh formation. Salt marshes usually are developed on a sinking coastline, originating as mud flats in the shallow water of sheltered bays, lagoons, and estuaries, or behind sandbars. Where tidal flow is impeded, sediments accumulate with the help of submerged vegetation. Eventually, the surface becomes exposed at low tide, stranding seeds or fragments of plant roots. The first colonizers are *Spartina*, *Salicornia*, and *Puccinellia;* and with continuing sedimentation, the soil matures and secondary invaders such as *Spartina patens* (salt-meadow hay) and *Distichlis spicata* (salt grass) arrive. The marsh surface then takes on a slope, with the landward edge higher, and finally plant zonation becomes evident (**Figs. 1** and **2**). *See* ESTUARINE OCEANOGRAPHY.

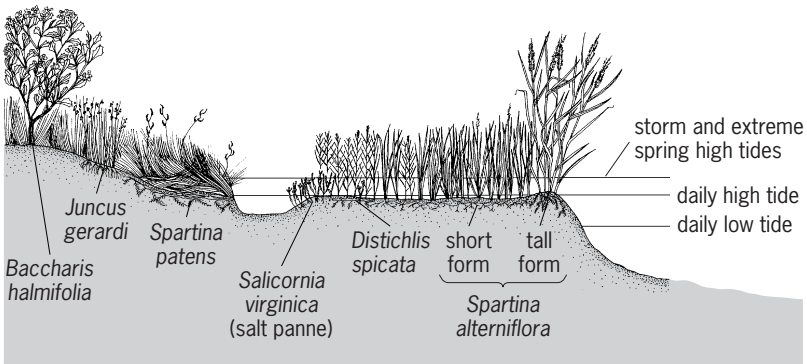

Fig. 2. Salt-marsh profile of the Delaware estuary. Such salt marshes occur primarily in the lower estuary, where they are dominated by smooth cordgrass (*Spartina alterniflora*). The typical profile from creek bank to upland features a levee next to the creek bank that is colonized by the tall form of *Spartina*. Away from the creek bank, the broad expanse of marsh is dominated by short-form *Spartina* intermixed with several subdominant species, including salt-meadow hay (*Spartina patens*) and glasswort (*Salicornia virginica*). (*Courtesy of Karen Grosz*)

Physiography. The primary physiographic features of salt marshes are creeks and pans. Typically, during the development of mud flats, minor surface irregularities cause water to be deflected into channels. Once channels are formed, scouring takes place, and the constant flooding and ebbing of the tide prevents vegetation from colonizing them. Salt-marsh creeks are an example of a stream bed produced by two conflicting processes—sedimentation and erosion.

In the early stages of marsh development, plant colonization is irregular, and parts of the bare surface, or pans, become surrounded by vegetation. As the marsh surface rises, such bare spots lose their outlets for tidal water. Filled with water, pans prevent the growth of vegetation, but if they are drained, vegetation invades. The so-called rotten spot is an irregularly shaped pan formed by mats of debris deposited by storm tides. This debris, combined with excessive salinity, kills the vegetation below.

Plants. Salt marshes are formed where salinity is high, ranging from 20 to 30 parts per thousand of sodium chloride. Proceeding up the estuary, there is a transitional zone where salinity ranges from 20 to less than 5 ppt. In the upper estuary, where river input dominates, the water has only a trace of salt. This varying salinity produces changes in the marsh—in the kinds of species and also in their number. Typically, the fewest species are found in the salt marsh and the greatest number in the fresh-water tidal marsh.

The salt marshes along the Atlantic shore of the temperate latitudes of North America are dominated by smooth cordgrass (*Spartina alterniflora*) and salt-meadow hay, and at the landward edge by black grass (*Juncus gerardi* or *J. roemarianus*), marsh elder (*Iva frutescens*), and groundsel tree (*Baccharis halimifolia*). The brackish-water marsh is dominated by big cordgrass (*Spartina cynosuroides*) and three-square sedge (*Scirpus americanus*). The fresh-water marsh is dominated by spadder dock (*Nuphar lutem*), pickerelweed (*Pontederia cordata*), broadleaf arrowhead (*Sagittaria latifolia*), and water hemp (*Amaranthus canabinus*).

Animals. Animals—from protozoa to mammals—abound in salt marshes. Just as the plants do, these creatures display a zonation determined by salinity, frequency, and duration of tidal submergence; examples are the ribbed mussel (*Geukensia demissa*), salt-marsh snail (*Melampus bidentatus*), and salt-marsh mosquito (*Aedes sollicitans*). The fauna of a tidal marsh can be placed in three major subdivisions: marine, terrestrial, and marsh species. The marine species, such as the bay anchovy (*Anchoa mitchilli*) and blue crab (*Callinectes sapidus*), have a center of distribution in the estuary. The terrestrial species, such as the meadow vole (*Microtus pennsylvanicus*) and raccoon (*Procyon lotor*), are air breathers; they live on the marsh or at the landward edge and make excursions onto the marsh. The marsh species are unique to the marsh. Most invertebrate species, such as the ribbed mussel, salt-marsh snail, and fiddler crab (*Uca* sp.), are derived from marine ancestors with aquatic larvae; as part of their life cycle, they require two or more adjacent but different ecological communities. Other marsh species include the greenhead fly (*Tabanus nigrovittatus*), mummichog (*Fundulus heteroclitus*), clapper rail (*Rallus longirostris*), seaside sparrow (*Ammospiza maritima*), and, in brackish-water marshes, the muskrat (*Ondatra zibethica*).

The large variety of plants and animals in the marsh leads to an assortment of plant–animal relationships. Direct interactions include trampling of vegetation and feeding, reproductive, and residential activities. Indirect interactions, which become evident some time after the occurrence of plant–animal interplay, include nutrient cycling and seed dispersal. The greater the intensity of interaction, the sooner and the greater the effect.

Productivity. The salt marsh is one of the most productive ecosystems in nature. This productivity is made possible by the involvement of solar energy in the photosynthetic process of higher rooted plants of the marsh and the algae growing on the surface muds. In addition to solar energy, tidal energy repeatedly spreads nutrient-enriched waters over the marsh surface, subsidizing the production process.

Some of this enormous supply of live plant material may be consumed by marsh animals, but the most significant values are realized when the vegetation dies and is decomposed by microorganisms to form detritus. Dissolved organic materials are released, providing an essential energy source for bacteria that mediate wetland biogeochemical cycles (carbon, nitrogen, and sulfur cycles). *See* BIOGEOCHEMISTRY.

Detritus helps form the base of the marsh–estuarine food web. It is fed upon directly by detritivores, small invertebrates and fishes living in the tidal creeks, which are in turn fed upon by coastal fishes such as menhaden (*Brevoortia tyrannus*), croaker (*Micropogonias undulatus*), spot (*Leiostomus xanthurus*), and summer flounder (*Paralichthys dentatus*), which spend a portion of their larval and juvenile stages in tidal creeks or in the adjoining estuary. These fishes then move out of the tidal creeks and estuary into the coastal waters, where they in turn are consumed by larger predators. This sequence represents an important transfer of marsh-derived production to the coastal ocean. *See* BIOLOGICAL PRODUCTIVITY; FOOD WEB.

Roles. The salt marsh has two general roles: one is human-based and the other is natural. It has been a source of hordes of biting flies, of pasturage for livestock, and of income from agricultural crops, shellfish, fish, animal pelts, gravel, and clay. It has been ditched, drained, impounded, and filled with trash from human activity, and it has served as a site for homes and factories.

The salt marsh also serves as a sediment sink, a nursery habitat for fishes and crustaceans, a feeding and nesting site for waterfowl and shorebirds, a habitat for numerous unique plants and animals, a nutrient source, a reservoir for storm water, an erosion control mechanism, and a site for esthetic

pleasures. An understanding of and an appreciation for the importance of salt marshes has led to federal and state legislation aimed at their protection. *See* PLANTS, SALINE ENVIRONMENTS OF; WATER CONSERVATION.

Franklin C. Daiber

Bibliography. P. Adam, *Saltmarsh Ecology*, 1990; R. H. Chabreck, *Coastal Marshes: Ecology and Wildlife Management*, 1988; V. J. Chapman, *Salt Marshes and Salt Deserts of the World*, 2d suppl. reprint ed., 1974; V. J. Chapman, *Wet Coastal Ecosystems*, 1977; F. C. Daiber, *Animals of the Tidal Marsh*, 1982; F. C. Daiber, *Conservation of Tidal Marshes*, 1986; L. R. Pomeroy and R. G. Wiegert (eds.), *The Ecology of a Salt Marsh*, 1981; D. S. Ranwell, *Ecology of Salt Marshes and Sand Dunes*, 1972.

Salviniales

A small order of heterosporous, leptosporangiate ferns (division Polypodiophyta) which float on the surface of the water. The delicate, branching stem becomes as much as 1 in. (2.5 cm) or so long and is provided with small, simple to bifid or more or less dissected leaves. The sporangia are enclosed in specialized appendages of the leaves, called sporocarps. A sporocarp contains numerous microsporangia or a single megasporangium, but not both types. The order contains only a single family, with two widely distributed genera, *Salvinia* and *Azolla*, with only about 20 species in all. *Azolla* is of some interest because of its symbiotic relationship with *Anabaena azollae*, a nitrogen-fixing blue-green alga which inhabits specialized chambers in the leaves. *See* POLYPODIOPHYTA.

Arthur Cronquist

Samarium

A chemical element, Sm, atomic number 62, belonging to the rare-earth group. Its atomic weight is 150.35, and there are seven naturally occurring isotopes; ^{147}Sm, ^{148}Sm, and ^{149}Sm are radioactive and emit α particles. *See* PERIODIC TABLE.

Samarium oxide is pale yellow, is readily soluble in most acids, and gives topaz-yellow salts in solutions. Samarium has found rather limited use in the ceramic industry, and it is used as a catalyst for certain organic reactions. One of its isotopes has a very high cross section for the capture of neutrons, and therefore there has been some interest in samarium in the atomic industry for use as control rods and nuclear poisons. *See* LANTHANUM; RARE-EARTH ELEMENTS.

Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; N. Kaltsoyannis and P. Scott (eds.), *The F Elements (Oxford Chemistry Primers, 76)*, Oxford University Press, 1999.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Sampled-data control system

A type of digital control system in which one or more of the input or output signals is a continuous, or analog, signal that has been sampled. There are two aspects of a sampled signal: sampling in time and quantization in amplitude. Sampling refers to the process of converting an analog signal from a continuously valued range of amplitude values to one of a finite set of possible numerical values. This sampling typically occurs at a regular sampling rate, but for some applications the sampling may be aperiodic or random. The sampling process consists of an ideal switch which passes the time-varying analog signal (**Fig. 1***a*) to a hold device which stores the instantaneous analog amplitude of the signal while it is converted to a binary number. The ideal sampler allows the signal to pass to the hold device at regular intervals for a very short period of time. The effective

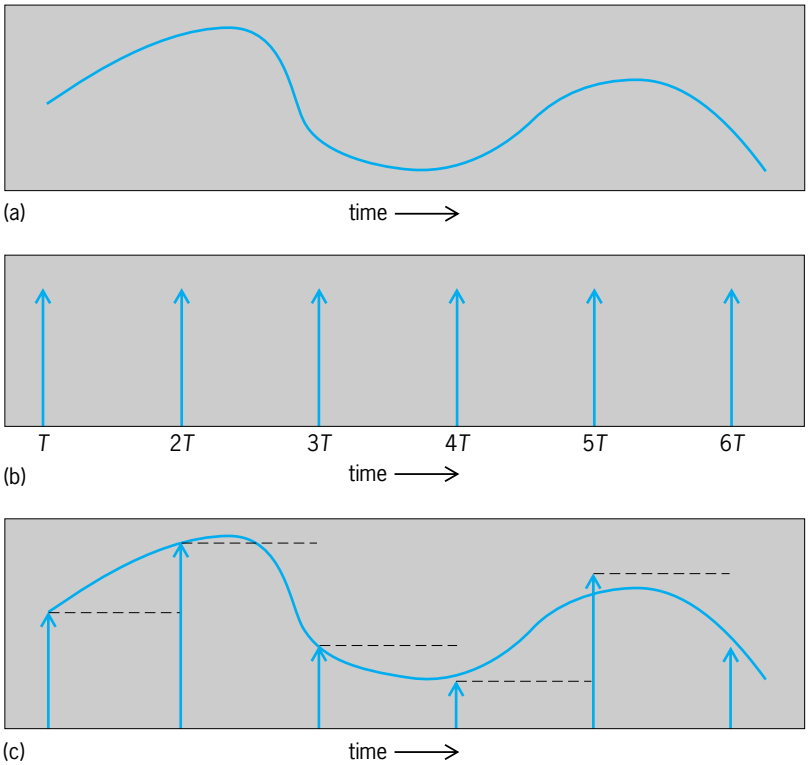

Fig. 1. Sampling process. (*a*) Continuous, analog signal to be sampled. (*b*) Sampling signals, corresponding to an ideal sampler that allows the signal to pass to the hold device at regular intervals of *T* seconds. (*c*) Resulting quantized values occurring at discrete instants of time. The horizontal lines represent the values of the analog signal as seen by the digital compensator, as well as the typical output of a digital-to-analog converter.

sampling signal is thus a series of delta functions (Fig. 1*b*). The resulting sampled signal (Fig. 1*c*) has the appearance of a pulse-amplitude-modulated signal and, for theoretical analysis, is often considered to be the product of a sequence of Kronecker delta functions with the original analog signal. *See* PULSE MODULATION.

A sampled-data control system can be decomposed into four components (**Fig. 2**): the plant which is to be controlled, the analog-to-digital converter, the digital device which implements the compensator, and the digital-to-analog converter.

Plant. While the device to be controlled is usually referred to as the plant, sampled-data control systems are also used to control processes. The term plant refers to machines or mechanical devices which can usually be mathematically modeled by an analysis of their kinematics, such as a robotic arm or an engine. A process refers to a system of operations such as a batch reactor for the production of a particular chemical, or the operation of a nation's economy. The output of the plant which is to be controlled is called the controlled variable. A regulator is one type of sampled-data control system, and its purpose is to maintain the controlled variable at a preset value (for example, the robotic arm at a particular position, or an airplane turboprop engine at a constant speed) or the process at a constant value (for example, the concentration of an acid, or the inflation rate of an economy). This input is called the reference or setpoint. The second type of sampled-data control system is a servomechanism, whose purpose is to make the controlled variable follow an input variable. Examples of servomechanisms are a robotic arm used to paint automobiles which may be required to move through a predefined path in three-dimensional space while holding the sprayer at varying angles, an automobile engine which is expected to follow the input commands of the driver, a chemical process that may require the pH of a batch process to change at a specified rate, and an economy's growth rate which is to be changed by altering the money supply. *See* PROCESS CONTROL; REGULATOR; SERVOMECHANISM.

Analog-to-digital converter. The analog-to-digital converter changes the sampled signal into a binary number so that it can be used in calculations by the digital compensator. The word length (number of bits of resolution) of the analog-to-digital converter limits the fundamental accuracy of the control system and determines its maximum speed of operation. Typical word lengths are 10–12 bits, but greater accuracies are available for some applications. For example, a 10-bit analog-to-digital converter quantizes the analog signal into 2^{10}, or 1024, discrete levels, which yields approximately 0.1% resolution. The conversion of a continuous valued signal into one of 2^n allowed values (where n is the word length) creates the equivalent of an additive noise called quantization noise. This noise is the difference in amplitude between the continuous signal and the amplitudes of the sampled signal (Fig. 1*c*). The word length of the analog-to-digital converter and subsequent computations is selected to keep this noise to an acceptable level.

The analog-to-digital converter also sets the maximum speed of operation of the sampled-data control system since it takes some time, usually microseconds, to effect the conversion. The Nyquist criterion requires that a system be sampled at greater than twice its maximum frequency component in order to properly represent a signal. Since the analog-to-digital converter is used to convert plant feedback signals, it is the dynamics of the plant which determine the minimum acceptable sampling rate, and this is reflected in the selection of the analog-to-digital converter conversion speed. Typical practical sampling rates are 8 to 10 times the maximum plant frequency. *See* ANALOG-TO-DIGITAL CONVERTER.

Digital-to-analog converter. Since a digital controller computes the control signal used to drive the plant, a digital-to-analog converter must be used to change this binary number to an analog voltage. The output of a digital-to-analog converter followed by a simple zero-order hold would be a step function (Fig. 1*c*), with the plant's dynamics being slow relative to the discontinuities in the output of the zero-order hold. *See* DIGITAL-TO-ANALOG CONVERTER.

Digital compensator. The digital compensator in the typical sampled-data control system (Fig. 2) takes

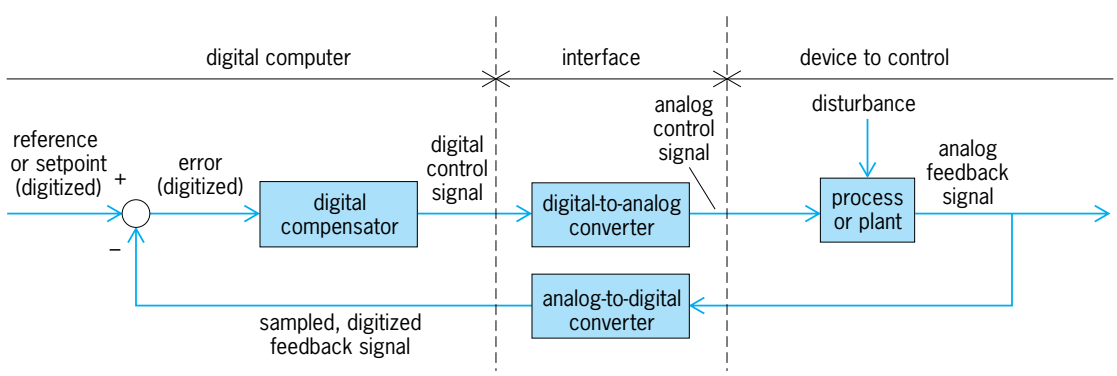

Fig. 2. Typical closed-loop, sampled data control system. The device to be controlled has continuously valued feedback signals which are sampled by an analog-to-digital converter, operated on by the digital compensator, and then converted back to analog signals for control of the plant.

the digitized values of the analog feedback signals and combines them with the setpoint or desired trajectory signals to compute a digital control signal, to actuate the plant through the digital-to-analog converter. A compensator is used to modify the feedback signals in such a way that the dynamic performance of the plant is improved relative to some performance index. In general, control systems are used to reduce output errors, maintain proper plant output in response to disturbances, compensate for plant wear or changing parameters, or make an uncontrollable plant controllable.

Sampled-data control systems are designed by first developing a mathematical model for the plant or process to control. From this model, inherent capabilities can be computed and performance deficiencies identified. A sampling rate can be selected and a compensator designed through well-established procedures in order to meet the desired performance measure. Compensators can be designed through a direct digital approach operating on sampled data signals, or they can be designed as continuous systems and then converted to sampled-data systems, with some trial and error required in both cases to meet the desired performance goals. *See* CONTROL SYSTEMS; DIGITAL COMPUTER; DIGITAL CONTROL. Kenneth J. Hintz

Bibliography. C. H. Houpis and G. B. Lamont, *Digital Control Systems*, 2d ed., 1992; K. Ogata, *Discrete-Time Control Systems*, 2d ed., 1995; G. A. Perdikaris, *Computer Controlled Systems*, 1991; J. R. Ragazzini and G. F. Franklin, *Sampled-Data Control Systems*, 1958.

Sand

Unconsolidated granular material consisting of mineral, rock, or biological fragments between 63 micrometers and 2 mm in diameter. Finer material is referred to as silt and clay; coarser material is known as gravel. Sand is usually produced primarily by the chemical or mechanical breakdown of older source rocks, but may also be formed by the direct chemical precipitation of mineral grains or by biological processes. Accumulations of sand result from hydrodynamic sorting of sediment during transport and deposition. *See* CLAY MINERALS; GRAVEL; MINERAL; ROCK; SEDIMENTARY ROCKS.

Origin. Most sand originates from the chemical and mechanical breakdown, or weathering, of bedrock. Since chemical weathering is most efficient in soils, most sand grains originate within soils. Rocks may also be broken into sand-size fragments by mechanical processes, including diurnal temperature changes, freeze-thaw cycles, wedging by salt crystals or plant roots, and ice gouging beneath glaciers. *See* WEATHERING PROCESSES.

Composition. Because sand is largely a residual product left behind by incomplete chemical and mechanical weathering, it is usually enriched in minerals that are resistant to these processes. Quartz not only is extremely resistant to chemical and mechanical weathering but is also one of the most abundant minerals in the Earth's crust. Many sands dominantly consist of quartz. Other common constituents include feldspar, and fragments of igneous or metamorphic rock. Direct chemical precipitation or hydrodynamic processes can result in sand that consists almost entirely of calcite, glauconite, or dense dark-colored minerals such as magnetite and ilmenite. Depending on local source rocks, sand may consist of more unusual minerals, such as gypsum, halite, or phosphate. Sand that consists largely of calcite in the form of shell fragments or skeletal debris is common in environments where biological activity is great but input from other sediments is low. *See* FELDSPAR; QUARTZ.

The compositions of sand most closely resembles the composition of source rocks when chemical and mechanical weathering processes are subdued, as in arid (desert) or tundra (arctic) settings. Where high relief allows for little soil development and rapid transport of sediment, sand is exposed to the weathering environment only briefly, and sands also are little altered by chemical or mechanical weathering. Conversely, in low-relief settings in humid or tropical climates, sand composition may differ greatly from the compositions of source rocks.

Texture. Sand can be described texturally in terms of the distribution of grain sizes, the roundness and shapes of grains, and the surface textures of grains. The mean size of sand grains is largely determined by the energy of the transport medium; higher current velocities or greater wave energy can transport grains of larger size. The coarsest grains are deposited at the first point where current velocity decreases, generally near the source area, and so grain size generally decreases with increased transport distance. The range of grain sizes present is referred to as sorting; sand containing a large variation in grain sizes is poorly sorted, whereas sand of nearly a single grain size is well sorted. Sorting is strongly influenced by transport and deposition processes; river sand tends to be poorly sorted, beach sand is intermediate, and eolian sand is generally well sorted.

The shape of grains is usually inherited from the source rock. For example, sand grains derived from a schist tend to be flat and tabular, whereas sand derived from a granite may consist of grains that are nearly equidim e nsional. Grain shape is not to be confused with grain roundness, a measure of the angularity of the grain edges. Well-rounded grains are smooth, while angular grains have sharp edges. Mechanical weathering in high-energy environments (such as beaches) or in low-viscosity transport mediums (air) can round sand grains, as can chemical dissolution in soils. The surface textures of sand grains, as seen under extreme magnification, can be useful indicators of the origin and transportational history of sand. For example, sand from an intense chemical weathering environment may be covered with dissolution etch pits (**Fig. 1**), and sand from eolian environments tends to be frosted.

Transport. Sand can be transported by any medium with sufficient kinetic energy to keep the grains

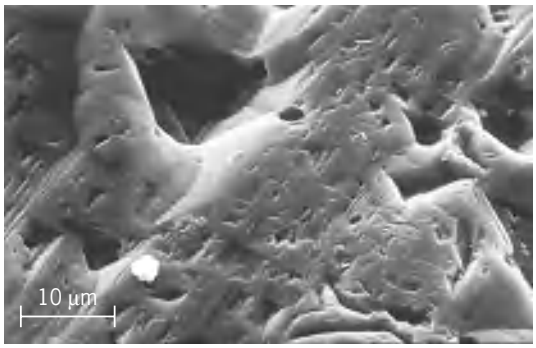

Fig. 1. Scanning electron photograph of the surface of a quartz grain from an intense chemical weathering environment. The surface texture shows numerous dissolution etch pits.

in movement. In nature, sand is commonly transported by rivers, ocean currents, wind, or ice. Rivers are responsible for transporting the greatest volume of sand the greatest distances. When rivers deposit sand at the oceans edge, it is commonly remobilized by waves or transported along shore by powerful longshore currents. Sand, like coarser and finer sediments, may also be entrained by glacial ice and transported to the sea by glacial flow. Where sand is widely exposed to wind currents, such as in river bars, on beaches, and in glacial outwash plains, it may be picked up and transported by the wind. As it moves, whether subaqueously or subaerially, sand commonly is organized into complex moving structures (bedforms) such as ripples and dunes (**Fig. 2**). *See* DUNE.

Occurrence. Accumulations of sand result from hydrodynamic sorting of sediment during transport. Most dramatic are the vast inland sand seas (ergs) in desert areas and accumulations at beaches. Probably at least as much sand is present, however, in river beds and temporarily stored in river floodplains as alluvium. Sand accumulations occur where a sudden decrease in current velocity causes the current to drop any carried sand. Thus, points where rivers or currents suddenly decelerate, such as deltas or the shallow continental shelf, are likely sites for the accumulation of sand. Sand is rare in the deep sea because currents have been robbed of their sand by nearshore deposition. Nevertheless, catastrophic subsea avalanches of mixtures of sediment and water, known as turbiditity currents, can carry appreciable amounts of sand to the deep sea. *See* DELTA; TURBIDITY CURRENT.

Fig. 2. Dunes produced by the transport of sand by wind, White Sands National Monument, NM. The sand in these dunes is made up almost entirely of gypsum ($CaSO_4$). Despite their appearance, the dunes are only a thin veneer of sand that is steadily moving across the countryside.

Geologic significance. Sand composition and texture can be used to tell much about the geologic history of the region in which it is found. Sand composition, for example, reflects the composition of source rocks, which in turn is controlled by the plate tectonic setting. With the advent of the plate tectonic model, it was noted (especially by W. R. Dickinson and colleagues) that the compositions of sands and sandstones could be related to the tectonic environments of source terranes, and sandstone composition has since been used as a tool for inferring the tectonic setting of ancient sequences. However, sand composition also is modified by chemical and mechanical processes. Thus, if source rocks remain unchanged through geologic time, changes in the composition of sands produced from them may be used as a tool for assessing paleoclimate. Finally, grain texture (roundness, surface markings, and so on) may be used to judge the nature (eolian, fluvial, glacial) and distance of transport, or the extent of reworking in the depositional environment, such as at beaches. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; PLATE TECTONICS.

Economic importance. Sand and gravel production is second only to crushed stone among nonfuel minerals in the United States. Although sand and gravel has one of the lowest average per ton values of all mineral commodities, the vast demand makes it among the most economically important of all mineral resources. Sand and gravel is used primarily for construction purposes, mostly as concrete aggregate. Pure quartz sand is used in the production of glass, and some sand is enriched in rare commodities such as ilmenite (a source of titanium) and in gold. *See* CONCRETE.

Mark J. Johnsson

Bibliography. M. J. Johnsson and A. Basu, Processes Controlling the Composition of Clastic Sediments, *Geol. Soc. Amer. Spec. Pap.*, no. 284, 1993; A. C. Morton, S. P. Todd, and P. D. W. Haughton (eds.), Developments in Sedimentary Provenance Studies, *Geol. Soc. London. Spec. Publ.*, no. 57, 1991; F. J. Pettijohn, P. E. Potter, and R. Siever, *Sand and Sandstone*, Springer-Verlag, New York, 1987; R. Siever, *Sand*, W. H. Freeman, New York, 1988; G. G. Zuffa, *Provenance of Arenites*, NATO Advanced Studies Institute, Reidel, Dordrecht, 1985.

Sand dollar

An echinoderm belonging to the order Clypeasteroida in the class Echinoidea. Sand dollars have a flat, disk-shaped body (see **illus.**), with the mouth in a mid-ventral position and with the anus, also on the ventral surface, laterally displaced. There are several species, but the best known is *Echinarachnius parma*, common along the Atlantic Coast of

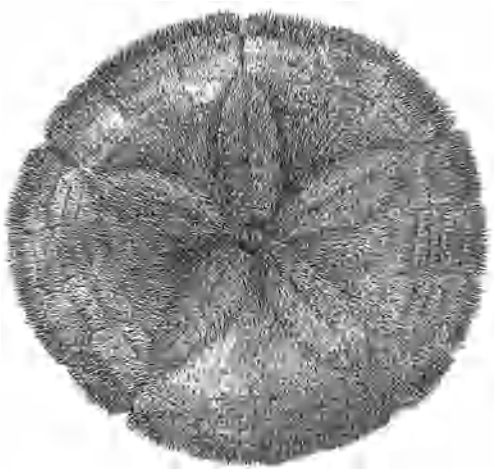

Sand dollar. The nearly circular body may reach 3 in. (7.5 cm) in diameter. The characteristic pattern of a five-part flower is seen on the aboral surface.

the United States, as well as the Pacific Coast south of Puget Sound. Other species are *Clypeaster subdepressus* and *Dendraster excentricus* of the Pacific Coast and *Mellita quinquiesperforata* found from Nantucket to Brazil.

Sand dollars live in sand, on the surface or partly buried, from the low-tide mark to depths of 4800 ft (1460 m). Burrowing and locomotion are assisted by the short spines which cover the body. Sand dollars ingest sand grains covered with diatoms or other algae. Little is known about their breeding habits. *See* ECHINOIDEA; SEA URCHIN. Charles B. Curtin

Sandalwood

The name applied to any species of the genus *Santalum* of the sandalwood family, Santalaceae (see **illus.**). However, the true sandalwood is the hard, close-grained aromatic heartwood of a parasitic tree, *S. album*, of the Indo-Malayan region. This fragrant wood is used in ornamental carving, cabinet work, and as a source of certain perfumes. The odor of the wood is an insect repellent, and on this account the wood is much used in making boxes and chests. The fragrant wood of a number of species in other families bears the same name, but none of these is the real sandalwood. *See* SANTALALES.

Perry D. Strausbaugh; Earl L. Core

Sandalwood branch with foliage and fruit.

Sandstone

A clastic sedimentary rock comprising an aggregate of sand-sized (0.06–2.0-mm) fragments of minerals, rocks, or fossils held together by a mineral cement. Sandstone forms when sand is buried under successive layers of sediment. During burial the sand is compacted, and a binding agent such as quartz, calcite, or iron oxide is precipitated from ground water which moves through passageways between grains. Sandstones grade upward in grain size into conglomerates and breccias; they grade downward in size into siltstones and shales. When the proportion of fossil fragments or carbonate grains is greater than 50%, sandstones grade into clastic limestones. *See* BRECCIA; CONGLOMERATE; LIMESTONE; SAND; SHALE.

Components. The basic components of a sandstone are framework grains (sand particles), which supply the rock's strength; matrix or mud-sized particles, which fill some of the space between grains; and crystalline cement. The composition of the framework grains reveals much about the history of the derivation of the sand grains, including the parent rock type and weathering history of the parent rock. Presence or absence of mud matrix is a function of the viscosity of the transporting agent and the rate of deposition and burial of the sand. Crystalline cement is largely responsible for the color of sandstone. Very small amounts (less than 1%) of iron oxide cement, for example, can impart vivid reds to ancient sandstones. Cements can be studied to learn about the temperature and depth of burial of sand. Textural attributes of sandstone are the same as those for sand, and they have the same genetic significance. *See* SAND. Lee J. Suttner

Classification. Sandstones are composed of mineral grains and rock fragments of many different kinds. Individual crystals of quartz and feldspar (the most common rock-forming minerals) and polycrystalline fragments of volcanic, sedimentary, and metamorphic rocks are typically most abundant. The quartz and feldspar grains are derived from the erosion of bedrock sources, such as granite, with constituent crystals large enough to yield sand grains composed of single minerals. The polycrystalline sand grains, called lithic fagments, are derived from rocks so finely crystalline that derivative sand grains are composed of multiple internal crystals. The nature of the sand grains is thus controlled partly by the nature of the bedrock from which the sand grains were derived, and partly by the degree of weathering of bedrock sources prior to erosion. Weathering tends to concentrate quartz, which is resistant to weathering, by dissolving feldspar and most other rock-forming minerals or converting them to clays at varying rates, depending upon the climate and topographic relief of source areas. In general, warm humid climates and low relief promote more

intense weathering of source rocks, while cool dry climates and high relief allow deeper erosion of fresh bedrock without the comparable weathering of susceptible minerals. Consequently, sandstones are classified according to the relative proportion of quartz to other grain types, and according to the ratio of feldspar grains to finely crystalline lithic fragments. Quartz-rich sandstones are commonly called quartz-arenite. Sandstones poor in quartz are commonly called arkose, when feldspar grains are more abundant than lithic fragments, and litharenite (or graywacke) when the reverse is true. Subarkose and sublitharenite (or subgraywacke) refer to analogous sandstones of intermediate quartz content. The interpretation of sandstone composition in terms of parent source rock is complicated by the fact that many sand grains are successively reworked by sedimentary transport during sequential cycles of weathering and erosion, meaning that the source of some sand is preexisting sandstone. Some sandstones also contain grains that are produced by either organic or inorganic processes at or near the site of deposition, and do not reflect any bedrock sources. Chief among these intrabasinal grains are fragments of fossils and other calcareous particles composed of calcium carbonate. Sandstones composed dominantly of calcareous grains are called calcarenite, and represent a special variety of limestone. Other sandstones composed exclusively of volcanic debris are called volcanic sandstone, and are gradational, through the interplay of eruptive and erosional processes, to tuff, the fragmental volcanic rocks produced by the disintegration of magma during explosive volcanic eruptions. *See* ARENACEOUS ROCKS; ARKOSE; FELDSPAR; GRAYWACKE; QUARTZ; TUFF. William R. Dickinson

Structures. Much can be learned about the history of a particular sandstone deposit through careful study of its sedimentary structures and fossils. The most common structure found in sandstones is crossbedding, which is produced by the wind- or water-driven migration of ripples or dunes (see **illus.**). Crossbeds are inclined at angles less than 35° to the primary parallel bedding visible in most outcrops. The crossbeds slope downward in the direction in which the current that formed them was flowing. Consequently they are important indicators of paleoflow directions. Detailed study of the size and shape of the crossbeds also can aid in interpreting the velocity of the ancient current flow which formed them, and in the case of subaqueously formed crossbeds, the depth of the flow.

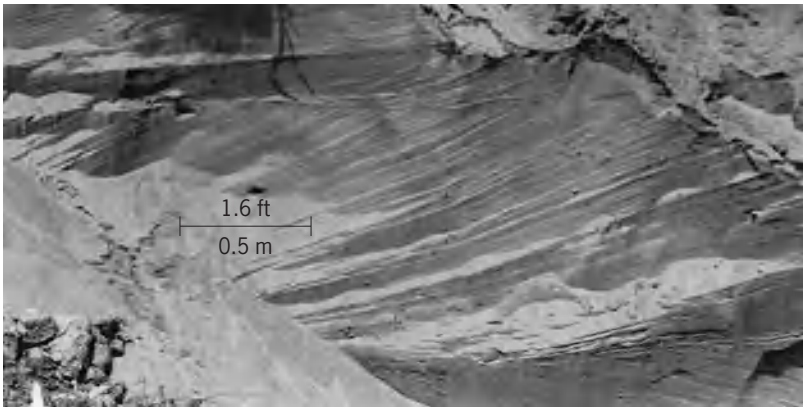

Crossbeds in sandstone. The water which deposited the sand was flowing from right to left as indicated by the direction of inclination of the crossbeds.

Structures also can be observed on the tops and bottoms of sandstone beds. For example, many different shapes of ripples have been found on the tops of sandstone beds and have been used to distinguish between movement of the sand by water currents, tidal currents, waves, and wind. Linear groove and scour marks occur as casts on the bottoms of some sandstone beds. These features are formed by very swift currents which gouge the underlying sand or mud. The gouges are then filled with sand when the current slows down. Generally, fossil shells are not common in sandstone. However, both within some sandstone beds and on the tops and bottoms of some beds, a variety of fossil trails and tracks of organisms may be present.

Interpretation. Many consider sandstones to be the most useful sedimentary rocks to interpret Earth history. They are widespread, constituting between 10 and 20% of the total volume of sedimentary rocks in the Earth's crust, and they extend almost back to the beginning of geologic time. Furthermore, they are easier to study than fine-grained rocks such as shale and are less susceptible to change after deposition, which obscures their original properties, than limestone.

Unlike limestones, sandstones also can reveal important paleogeographic information about the areas surrounding their sites of deposition. As an example, an area where arkose or feldspathic sandstone accumulated more than likely was located near mountains, as abundant sand-sized feldspar can be derived only from crystalline rocks. Significant volumes of such rocks normally are exposed only within mountain ranges. In contrast, quartz arenites typically are formed where relief is low or where older sandstones served as parent rocks. Lithic sandstones have been deposited in a wide range of shallow- to deep-water environments in structurally active parts of the Earth's crust such as subduction zones and zones of continental collision. *See* PLATE TECTONICS.

Uses. Because sandstone can possess up to 35% connected pore space, it is the most important reservoir rock in the Earth's crust. By far the greatest volume of water, oil, and gas is extracted from sandstone, where it occurs in pore spaces. Although in the past fluid has usually been extracted from sandstone, in the future sandstone may serve as a reservoir into which hazardous fluids, such as nuclear wastes, are injected for storage. *See* HAZARDOUS WASTE.

Sandstone which is easily split (flagstone) and has an attractive color is used as a building stone. Sandstone is also an important source of sand for the glass industry and the construction industry, where it is used as a filler in cement and plaster. Crushed sandstone is used as road fill and railroad ballast. Silica-cemented sandstone is used as firebrick in industrial

furnaces. Some of the most extensive deposits of uranium are found in sandstones deposited in ancient stream channels. *See* GLASS; REFRACTORY; SEDIMENTARY ROCKS; STONE AND STONE PRODUCTS; URANIUM. Lee J. Suttner

Bibliography. H. Blatt, *Sedimentary Petrology*, 2d ed., 1991; S. Boggs, Jr., *Petrology of Sedimentary Rocks*, 1991; S. Boggs, Jr., *Principles of Sedimentology and Stratigraphy*, 2d ed., 1996; R. L. Folk, *Petrology of Sedimentary Rocks*, 4th ed., 1980; J. T. Greensmith, *Petrology of the Sedimentary Rocks*, 7th ed., 1988; G. Klein, *Sandstone Depositional Models for Exploration*, 1988; F. J. Pettijohn, P. E. Potter, and R. Siever, *Sand and Sandstone*, 2d ed., 1987.

Santalales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Rosidae of the class Magnoliopsida (dicotyledons). The order consists of 10 families and about 2000 species. The largest families are the Loranthaceae (about 900 species), Santalaceae (about 400 species), Viscaceae (about 350 species), and Olacaceae (about 250 species). A few of the more primitive members of the Santalales are autotrophic, but otherwise the order is characterized by progressive adaptation to parasitism, accompanied by progressive simplification of the ovules. In the Loranthaceae the ovules have no integument and are embedded in the large, central placenta. Some other families have a free central placenta with pendulous, terminal ovules; in any case there are only a few ovules, in contrast to the very numerous ovules of the Rafflesiales. Some members of the Santalales, such as sandalwood (*Santalum album*, a small tree of southern Asia), are rooted in the ground and produce small branch roots which invade and parasitize the roots of other plants. Others, such as mistletoe (*Viscum* and other genera of the Viscaceae), grow on trees, well above the ground. *See* FLOWER; MAGNOLIOPHYTA; MAGNOLIOPSIDA; MISTLETOE; PLANT KINGDOM; RAFFLESIALES; SANDALWOOD.

Arthur Cronquist; T. M. Barkley

Sapindales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Rosidae of the class Magnoliopsida (dicotyledons). The order consists of 15 families and about 5400 species. The largest families are the Rutaceae (about 1500 species), Sapindaceae (about 1500 species), Meliaceae (about 550 species), Anacardiaceae (about 600 species), and Burseraceae (about 600 species). Most of the Sapindales are woody plants, with compound or lobed leaves and polypetalous, hypogynous to perigynous flowers with one or two sets of stamens and only one or two ovules in each locule of the ovary. A large proportion of the species have glandular cavities in the leaves, or resin ducts in the bark and wood, or other sorts of secretory structures. Some of the Anacardiaceae, including poison ivy (*Toxicodendron radicans*; see **illus.**) and the lacquer tree (*T. vernicifluum*) of the Orient, are notoriously allergenic to humans. Some other well-known members of the order are orange, lemon, lime, and grapefruit (all species of *Citrus* in the Rutaceae); the various kinds of maple (*Acer*, in the Aceraceae); and the horse chestnut (*Aesculus* hippocastanum, in the Hippocastanaceae). *See* BUCKEYE; CASHEW; CITRON; GRAPEFRUIT; KUMQUAT; LEMON; LIGNUMVITAE; LIME (BOTANY); MAGNOLIOPHYTA; MAGNOLIOPSIDA; MAHOGANY; MANDARIN; MAPLE; ORANGE; PISTACHIO; PLANT KINGDOM; POISON IVY; QUEBRACHO; SECRETORY STRUCTURES (PLANT); TANGERINE; VARNISH TREE. Arthur Cronquist; T. M. Barkley

Poison ivy (*Toxicodendron radicans*), a characteristic member of the family Anacardiaceae in the order Sapindales, division Magnoliophyta. (*John Gerard*, *National Audubon Society*)

Sapphire

Any gem variety of the mineral corundum (Al_2O_3), except those called ruby because of their medium- to dark-red color. Sapphire has a hardness of 9 (Mohs scale), a specific gravity near 4.00, and refractive indices of 1.76–1.77.

Although blue sapphires are most familiar to the public, transparent sapphires occur in many other colors, including yellow, brown, green, pink, orange, and purple, as well as colorless and black. Some sapphires change color, from violet to purple or green to red, depending on the type of light under which they are viewed. A lovely, slightly light red-orange variety is called padparadscha. Star sapphires are fashioned from translucent corundum that contains many silk-like crystals of the mineral rutile. The rutile crystals

form in three sets, and when the corundum is cut en cabochon (domed, without facets) and oriented correctly, light reflected from the rutile creates a six-rayed star. Corundum belongs to the trigonal crystal system and is isostructural with hematite, Fe_2O_3. *See* CORUNDUM; HEMATITE; RUBY.

Kashmir, in northern India, once the most famous source of blue sapphires, has become a minor source; another former source is Burma. Sri Lanka, Thailand, and Australia are now considered the most important sources. Many Sri Lankan and Australian sapphires are sent to Bangkok, where they are heat-treated to improve their color. One type, which is a cloudy white, becomes a magnificent, transparent rich blue when the heating dissolves the rutile needles, clarifying the stone. Australia produces enormous quantities of very dark blue sapphires, as well as yellow-to-golden sapphires. Sapphires are also mined in Montana in the United States and in East Africa. Most sapphires are mined either in gravels resulting from the weathering of basic igneous rocks, or in the type of marble resulting from the intrusion of an igneous mass into an impure limestone.

Blue sapphire is most valuable when it is a medium to medium-dark tone of a slightly violet-blue, or cornflower blue. The Kashmir grade has a slightly "sleepy" appearance, caused by inclusions that very slightly reduce transparency. Blue sapphires and the padparadscha variety are generally more valuable than other colors. Colorless sapphire is of little value. The golden sapphires are much in demand and thus quite expensive.

Sapphire in blue and other colors is synthesized by the Verneuil flame-fusion process, and some blue sapphires and the padparadscha variety have been made by other methods. The only flux-melt blue sapphires are made in California. *See* GEM.

Richard T. Liddicoat, Jr.

Sapropel

A term possessing genetic implications, originally defined as an aquatic sediment rich in organic matter that formed under reducing conditions (lack of dissolved oxygen in the water column) in a stagnant water body. This contrasts with the term gyttja, which is also a sediment high in organic carbon content but which formed under inferred oxygenated conditions in the water column down to the sediment-water interface (thus benthic organisms may be present). Such inferences about water-column dissolved-oxygen contents are not always easy to make for ancient environments. Therefore, the term sapropel or sapropelic mud has been used loosely to describe any discrete black or dark-colored sedimentary layers (>1 cm or 0.4 in. thick) that contain greater than 2 wt % organic carbon. Sapropels may be finely laminated (varved) or homogeneous, and may less commonly exhibit structures indicating reworking or deposition of the sediment by currents. Sapropels largely contain amorphous (sapropelic) organic matter derived from planktonic organisms (such as planktonic or benthic algae in lakes or plankton in marine settings). Such organic matter possesses a large hydrogen-to-carbon ratio; therefore, sapropelic sequences are potential petroleum-forming deposits. The enhanced preservation of amorphous organic matter in sapropels may indicate conditions of exceptionally great surface-water productivity, extremely low bottom-water dissolved-oxygen contents, or both. Some sapropels may, however, contain substantial amounts of organic matter derived from land plants. *See* ANOXIC ZONES; MARINE SEDIMENTS; ORGANIC GEOCHEMISTRY; PETROLEUM; VARVE.

Michael A. Arthur

Sarcodina

A superclass of Protozoa (in the subphylum Sarcomastigophora) in which movement involves protoplasmic flow, often with recognizable pseudopodia. Gametes may be flagellated, as in certain Foraminiferida. Most species are floating or creeping; a few, sessile. The pellicle is thin, and the body is apt to be plastic unless restrained by skeletal structures. Sarcodina live in fresh, brackish, or salt waters; soil or sand; and as endoparasites in animals and plants. A group may be limited to a specific habitat, but many have a rather wide range. *See* FORAMINIFERIDA.

Sarcodina include two major classes: Rhizopodea and Actinopodea. In addition, for supposed lack of a better taxonomic position, Piroplasmea (*Babesia*), as a third class, have been assigned to the group by some protozoologists.

Rhizopodea. Pseudopodia show two different types of activity: contractile-hydraulic movement (*Amoeba proteus*) and bidirectional movement (Foraminiferida and Heliozoia). Such pseudopodial dimorphism has led to the suggestion that a basic revision of the Sarcodina is needed. *See* RHIZOPODEA.

The Rhizopodea include five subclasses (Lobosia, Filosia, Granuloreticulosia, Mycetozoia, and Labyrinthulia), with such varied characteristics that interrelationships are somewhat uncertain. However, Rhizopodea do not show distinctly radial arrangement of slender pseudopodia. Skeletal elements range from none to complicated tests (**Fig. 1**) in many Granuloreticulosia. Simpler testate rhizopods build a one-chambered test (**Fig. 2**). In some species the test is composed entirely of secreted material (often siliceous), but many have a test composed of sand grains or other foreign particles (Fig. 2*e*) held together by a secreted cement. A few have a moderately flexible chitinous test. Rarely (for example, *Paraquadrula*), shell plates may be composed of calcium salts. Skeletal plates of *Euglypha* (Fig. 2*d*) are secreted during growth and stored until the next fission, when one organism receives them for building its test. Other rhizopods may ingest and store sand grains for use at the next fission.

Multichambered tests are formed by many Foraminiferida (Fig. 1*a–l*), and simpler tests by primitive species (Fig. 1*o*). In some genera the adult retains a chitinous test. More often, a primary test is

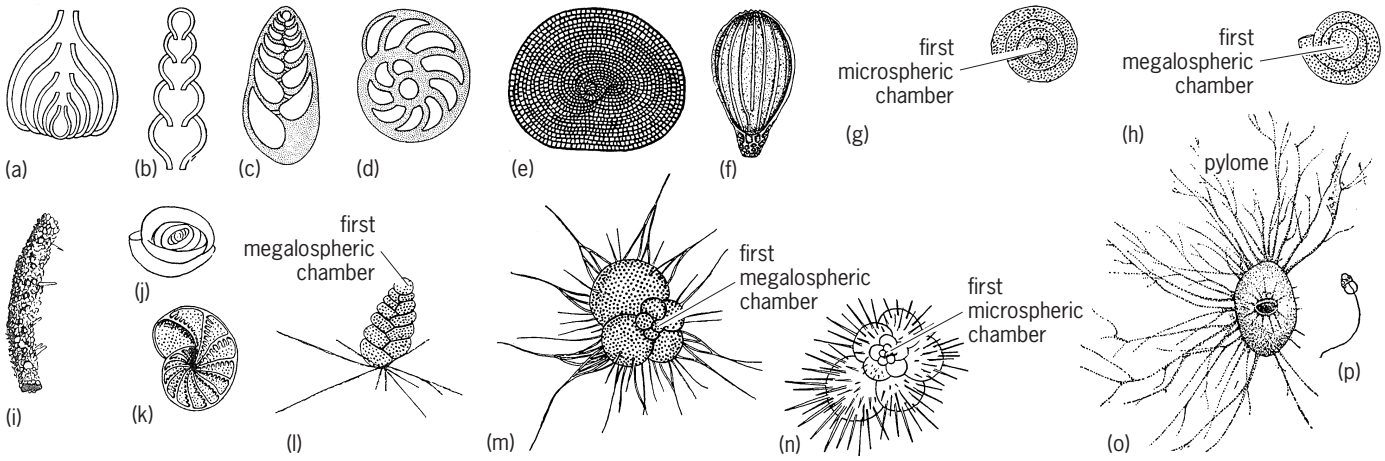

Fig. 1. Foraminiferida. (*a*, *b*) Foraminiferan tests, nodosaroid types. (*c*) Textularid test. (*d*) Spiral test. (*e*) Cycloid test (*Discospirulina*). (*f*) Test of *Lagena*. (*g*, *h*) Microspheric and megalospheric tests of *Cornuspira*. (*i*) Arenaceous test (*Saccorhiza*). (*j*) Test representative of family Miliolidae. (*k*) Spiral test (*Elphidium*). (*l*) *Textularia* (living). (*m*, *n*) *Globigerina* (living) megalospheric and microspheric. (*o*) *Gromia*. (*p*) Supposed flagellated gamete of *Gromia*. (*After L. H. Hyman*, *The Invertebrates*, *vol. 1*, *McGraw-Hill*, *1940*)

reinforced by the addition of inorganic salts or foreign particles. In the formation of arenaceous tests (Fig. 1*i*), sand grains, sponge spicules, or other materials are cemented onto a primary chitinous wall. However, tests of most species are mainly calcareous (rarely siliceous) secretions, the materials resembling the cements used in arenaceous tests. During growth of certain simple species the original test is discarded and then replaced by a larger one. Forms with multilocular tests typically add new chambers in succession as growth continues (Fig. 1*a–b* and *j–n*). The protoplasm remains continuous from one chamber to the next, since the intervening apertures remain at least partly open. These openings (foramina) suggested the taxonomic name for the group. The reticulopodia tend to anastomose into a sticky net (Fig. 1*o*), often greatly exceeding the test in area. The net is an efficient food trap; also, partial digestion may occur outside the test. Reticulopodia also may construct the walls of new chambers or cysts. In addition, some Foraminiferida can creep at rates up to several millimeters an hour.

Actinopodea. Members of this class have slender pseudopodia, with or without axonemes, commonly extended radially with little anastomosing or branching (**Fig. 3**). Bidirectional cytoplasmic flow is known in reticulopodia of representative genera. There are four subclasses: Radiolaria, Acantharia, Heliozoia, and Proteomyxidia. *See* ACTINOPODEA.

Life cycles. Sexual activities include syngamy (the fusion of gametes) in certain Heliozoia, Mycetozoia, and Foraminiferida. Other reported cases are less well substantiated. In certain Foraminiferida there is an alternation of a haploid gamete-producing generation with a diploid generation, a plantlike cycle apparently unique in Protozoa. Species with such cycles may show dimorphism with respect to initial chambers of the tests (Fig. 1*g*, *h*, *m*, and *n*). A diploid adult has a microspheric (small) initial chamber; a haploid adult is megalospheric. It is suggested that in *Gromia* a naked multinucleate asexual

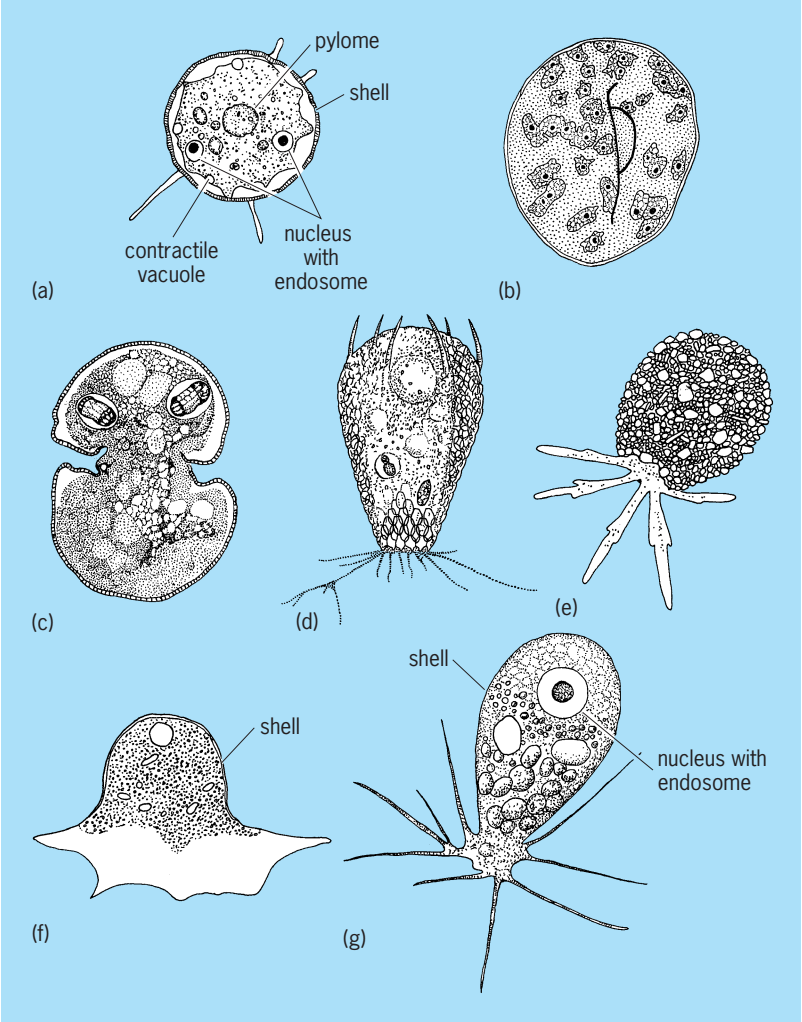

Fig. 2. Some shelled Lobosia. (*a*) *Arcella*, seen from above; (*b*) sporulating into numerous ameboid young (*c*) in binary fission, with the two nuclei in mitosis with polar caps part of the cytoplasm exudes from the pylome and secretes a new shell. (*d*) *Euglypha* with a shell of little siliceous scales. (*e*) *Difflugia*, with a case made of rock grains. (*f*) *Cochliopodium*, with a membranous case. (*g*) *Chlamydophrys*. (*After L. H. Hyman*, *The Invertebrates*, *vol. 1*, *McGraw-Hill*, *1940*)

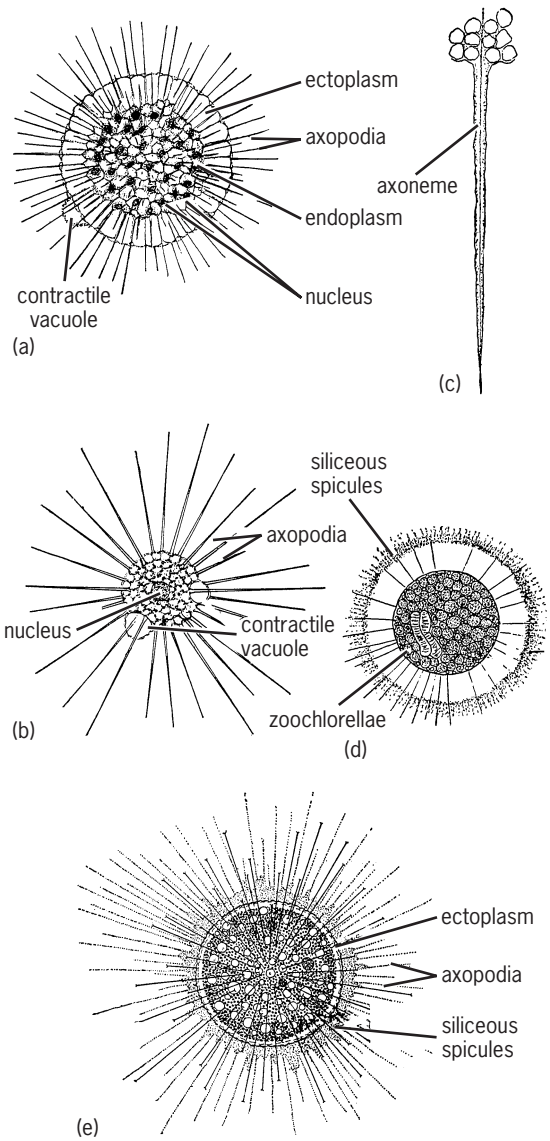

Fig. 3. Heliozoia. (*a*) *Actinosphaerium*. (*b*) *Actinophrys*. (*c*) Axopodium of *Actinosphaerium*, highly magnified. (*d*) *Heterophrys*. (*e*) *Acanthocystis*. (*After L. H. Hyman, The Invertebrates, vol. 1, McGraw-Hill, 1940*)

generation may alternate with a testate sexual generation.

In typical alternation of generation cycles, ordinary reproduction may occur in both haploid and diploid phases in sequences ranging from 2–3 weeks to 1 year or so in duration. *See* PROTOZOA; SARCOMASTIGOPHORA.

Richard P. Hall

Sarcomastigophora

A subphylum of Protozoa. Sarcomastigophora replaces Rhizoflagellata for those protozoa that possess flagella or pseudopodia or both, according to a classification set up by an international committee of the Society of Protozoologists. Included organisms have a single type of nucleus, except the developmental stages of some Foraminiferida. Sexuality, if present, is syngamy, the fusion of two gametes. Spores typically are not formed. Flagella may be permanent or transient or confined to a certain stage in the life history; this is true also of pseudopodia. Both flagella and pseudopodia may be present at the same time.

Classification. Three superclasses are included: (1) Mastigophora, commonly flagellates, contains 19 orders. (2) Opalinata includes 1 order; these organisms were once considered as ciliates, but further research has indicated flagellate kinships. (3) Sarcodina comprises organisms which normally possess pseudopodia and are flagellated only in the developmental stages; 13 orders possess irregularly distributed lobose or filose and branching pseudopodia, while 7 orders have radially distributed axopodia, often with axial filaments. *See* MASTIGOPHORA; SARCODINA.

Mastigophora. This superclass is often divided into two classes: Phytamastigophorea and Zoomastigophorea. The Phytamastigophorea, or plant flagellates, usually have green or variously colored chromatophores. However, the order Ebriida is altogether colorless, and eight other orders contain few to many species without chromatophores, especially Chrysomonadida, Dinoflagellida, and Euglenida. Besides chlorophylls, many other pigments have been identified, and reserve materials vary from order to order. Thus Volvocida elaborate starch; Euglenida, paramylum; Chloromonadida, lipids and glycogen; Dinoflagellida, starch, lipids, and unknown materials; and Chrysomonadida, leucosin and lipids. *See* PHYTAMASTIGOPHOREA.

These plant flagellates are very diverse in form and structure. Some are naked, and some live in loricas or tests which they manufacture. Some are fresh-water and others marine forms; most Cryptomonadida and colorless Euglenida tolerate either milieu. Most of them are unicellular, but some form colonies, and in the Volvocida the *Volvox* colonies are quite complex. However, this whole class shows decided relationships to one another.

Members of the class Zoomastigophorea, all nonpigmented, are referred to as the animal flagellates. Three of the orders, Choanoflagellida, Bicosoecida, and Rhizomastigida, are free-living; the remaining six orders are predominantly parasitic. The choanoflagellates have been used to relate the protozoa to sponges by virtue of the common protoplasmic collar. This group also manufactures very complex loricae, as shown by R. E. Norris for *Pleurasiga* or *Sportelloeca*. Relationships between these three orders are quite evident, and through genera such as *Mastigella* also, to the ameboid protozoa. The increase in number of flagella from one in Choanoflagellida to many in Hypermastigida conceivably leads to opalinid and ciliate relationships.

The remaining six orders are partially or wholly parasitic, the parasites (or in some cases symbionts) of wood roaches and termites attaining a very high degree of structural organization and adaptation to their mode of life. *See* ZOOMASTIGOPHOREA.

Opalinata. These are practically all inhabitants of the digestive tract of amphibians, and M. M. Metcalf, who made very extensive studies of them in 1923, referred to them as "opalinid ciliates," which they

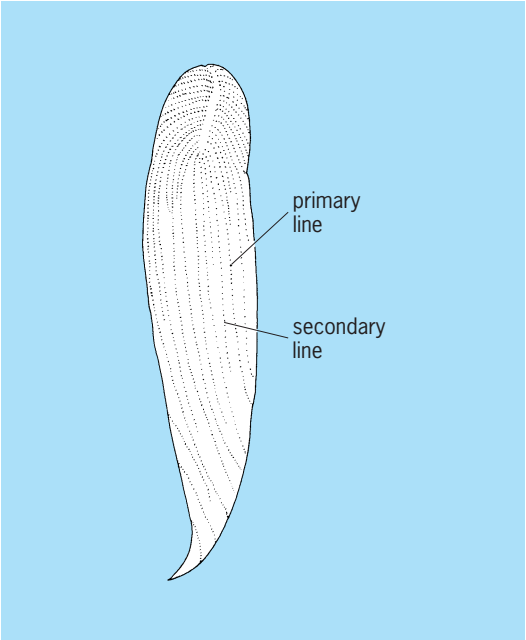

Fig. 1. An opalinid (*Protoopalina axonucleata*) showing primary and secondary lines of blepharoplasts.

certainly resemble (**Fig. 1**). Since about 1950, however, considerable questioning of their phylogenetic position has arisen, especially since anisogamy (fusion of dissimilar gametes) has been demonstrated. They are regarded as a highly specialized group, and the protozoologists' committee regards their taxonomic position as needing further clarification. In this respect they are important, but as to the size of the group and their ecology, they are less so.

Sarcodina. No common name can apply to this group, which includes the amebas, foraminiferids, slime molds, radiolarians, acantharians, and heliozoans. Each of these is well known to perhaps a few specialists. The superclass Sarcodina comprises 20 orders which appear to differ widely but commonly have pseudopodia, restriction of flagella to developmental stages, a cortical zone of relatively undifferentiated ectoplasm, and asexual reproduction by fission. Most of them are free-living. *See* SARCODINA.

The class Rhizopodea includes the large *Amoeba*, *Chaos*, and *Pelomyxa* (of beginning courses in biology), which are common in fresh waters. It also includes the large *Gromia*, usually marine, and the very abundant and often large Foraminiferida. The subclass Mycetozoia embraces slime molds, which may be parasitic (order Plasmidiophora) or free-living and which may live in an aerial instead of an aqueous environment. The Mycetozoia forms peculiar and diagnostic fruiting bodies on decaying logs and leaves; and these organisms have a complicated life cycle. *See* RHIZOPODEA.

Rhizopodea even includes blood parasites (order Piroplasmida), in which flagella are unknown. This order is placed in a separate class, Piroplasmea, according to the new classification.

The class Actinopodea contains four subclasses which are usually floating and have radiating filose pseudopodia, often with an axial filament. Included are the marine Radiolaria, which have a central capsule containing endoplasm, perforated by one to many pores. The subclass Acantharia differs in having a skeleton of regularly arranged radial spines of strontium sulfate. The central capsule is thin, membranous, and without pores. *See* ACTINOPODEA.

The three orders of Heliozoia are common freshwater genera, some large (*Actinophrys*), some small (**Fig. 2**), and all free-living. The subclass Proteomyxidia is principally parasitic and is poorly known.

Ecology and economic significance. Most of the plant flagellates will live either in fresh water or in both fresh and salt water. The orders Silicoflagellida, Coccolithophorida, and Ebriida are marine forms. The green Euglenida have only four or five species of *Euglena* in salt water, while species of the genera *Eutreptia* and *Eutreptiella* are confined to salt water. The plant flagellates may be very abundant and are of general ecologic interest in being reoxygenators, as a lower step in the food chain, and in the decomposition of dissolved organic matter, since some of them are saprophytic in nutrition. Economically, some of them, such as the dinoflagellate *Gymnodinium breve*, cause enormous fish kills. *See* FOOD WEB.

The zooflagellates are small and are not sufficiently abundant to enter markedly into the food chain. But the parasites and symbionts are of considerable interest economically and theoretically, for example, the trypanosomes and the peculiar xylophagous

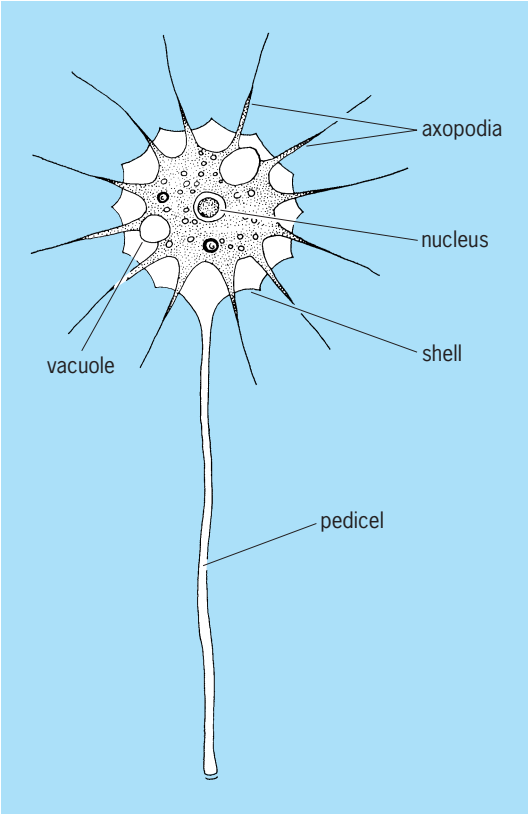

Fig. 2. *Hedriocystis reticulata*, a small heliozoan of the superclass Sarcodina.

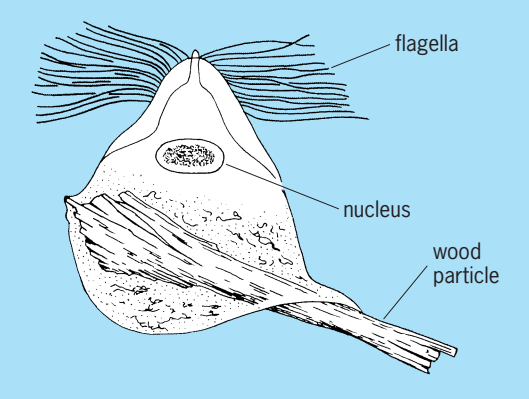

Fig. 3. A xylophagous symbiotic flagellate (*Trichonympha* sp.) from a termite.

(wood-eating) symbionts of termites (**Fig. 3**). In the termites these parasites actually digest the wood eaten by the host.

The rhizopods are typically free-living and rarely overabundant. Oil geologists make use of the shells of foraminiferids to determine what strata may have been penetrated. *See* FORAMINIFERIDA.

In the Mycetozoia are important parasites of cruciferous ciferous plants and the organism which is claimed, but not proven, to be the cause of the eelgrass (*Zostera*) decline along the Atlantic coast in the 1930s.

Conspicuous in the ecology of marine waters are the dinoflagellates, radiolarians, and acantharians, especially in tropical waters of otherwise low productivity. *See* PROTOZOA. James B. Lackey

Bibliography. B. M. Honigberg et al., A revised classification of the phylum Protozoa, *J. Protozool.*, 11(1):7–20, 1964; M. M. Metcalf, The opalinid ciliate infusorians, *Bull. U.S. Nat. Mus.*, no. 120, pp. 1–484, 1923; M. M. Metcalf, Further studies on the opalinid ciliate infusorians and their hosts, *Proc. U.S. Nat. Mus.*, 87:465–635, 1940; R. E. Norris, Neustonic marine craspedomonadates (choanoflagellates) from Washington and California, *Protozoology*, 12(4):589–602, 1965.

Sarcopterygii

By consensus of contemporary scholars, Sarcopterygii is a class of vertebrates that contains the orders Coelacanthiformes (lobefin fishes) and Ceratodoniformes (lungfishes), as well as the Tetrapoda.

Following is a recent classification of the extant sarcopterygians:

Class Sarcopterygii
- Subclass Coelacanthimorpha
  - Order Coelacanthiformes (lobefin fishes)
    - Family Coelacanthidae
- Subclass Dipnotetrapodomorpha
  - Order Ceratodontiformes (lungfishes)
    - Family Ceratodontidae
      - Lepidoserinidae
      - Protopteridae
- Infraclass Tetrapoda

The names Amphibioidei and Choanichthyes also appear in the literature, both being equivalent to Sarcopterygii less the tetrapods. The latter name is certainly inappropriate, as not all sarcopterygians have choanae (internal nares which allow passage from the nasal cavity to the mouth). The Sarcopterygii of only what are conventionally called crossopterygians and dipnoans is also still a popular classification. Debates concerning the relationships of the sarcopterygians and the ancestry of tetrapods—that is, whether the origin of the tetrapods occurred within the lungfishes, coelacanths, or rhipidistians—have persisted for many years, and will probably continue into the future. *See* COELACANTHIFORMES; CROSSOPTERYGII; DIPNOI (CERATODONIFORMES); OSTEICHTHYES; RHIPIDISTIANS (OSTEOLEPIMORPHA); TETRAPODA.

Herbert Boschung

Bibliography. W. E. Bemis, W. W. Burggren, and N. E. Kemp (eds.), *The Biology and Evolution of Lungfishes.* J. Morphol. Suppl. 1, New York, Alan R. Liss, 1987; M.-M. Chang, Rhipidistians, dipnoans, and tetrapods, in H.-P. Schultze and L. Trueb (eds.), *Origins of the Higher Groups of Tetrapods*, pp. 3–28, Cornell University Press, Ithaca, New York, 1991; R. Cloutier, Patterns, trends, and rates of evolution within the Actinistia, *Environ. Biol. Fishes*, 32(1):23–58, 1991; J. G. Maisey, Heads and tails: A chordate phylogeny, *Cladistics*, 2(3):201–256, 1986; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 1994; D. E. Rosen, Lungfishes, tetrapods, paleontology and plesiomorphy, *Bull. Amer. Mus. Nat. Hist.*, 167:159–276, 1981; H.-P. Schultze, Dipnoans and sarcopterygians, in W. E. Bemis, W. W. Burggren, and N. E. Kemp (eds.), *The Biology and Evolution of Lungfishes*, pp. 39–47, J. Morphol. Suppl. 1, New York, Alan R. Liss, New York, 1987; K. S. Thompson, The origin of the tetrapods, in P. Dodson and P. Gingerich (eds.) *Functional Morphology and Evolution*, pp. 33–62, Amer. J. Sci. Spec. Vol. 293-A, 1993; G. C. Young, J. A. Long, and A. Richie, Crossopterygian fishes from the Devonian of Antarctica: Systematics, relationships and biogeographical significance, *Rec. Aust. Mus.*, *Suppl. 14*, 1992.

Sarcosporida

An order of Protozoa of the class Haplosporea which comprises parasites in skeletal and cardiac muscle of vertebrates. The organisms have a very wide distribution both geographically and in host species, infecting reptiles, birds, and mammals (including marsupials). Humans are occasionally infected by the parasite referred to as *Sarcocystis lindemanni.* It is doubtful whether Sarcosporida found in different hosts are themselves always different, though this has often been assumed. Host specificity is known not to be strict; however, it is unlikely that there is only one species—*S. miescheriana*—as proposed by some investigators.

Sarcocystis. This genus is now usually placed with the Protozoa (subphylum Sporozoa), but it has given

taxonomists considerable difficulty in the past. By some it was thought to be a fungus, and by others it was simply relegated to a group of "undetermined position." The Committee on Taxonomic Problems of the Society of Protozoologists recommended that the parasites collectively referred to as the genus *Sarcocystis* be put in an order to be called Sarcosporida, class Haplosporea (this suggestion is followed here); *Sarcocystis* would then be the only genus in the order.

Species descriptions. The criteria of species are very hard to define for *Sarcocystis*. Species descriptions have usually been based on an assumed host specificity, size of spores, and cyst characteristics. But host specificity has seldom been proved, spore size is variable (degree of variation often in part dependent on the host species), and cyst morphology is frequently inconstant. Nevertheless, cyst morphology is probably more stable than anything else about the parasite.

Morphology. In general, the cyst wall is said to have one or two layers, the outer one being either smooth or provided with spines or villi. The genesis of the wall is in some dispute; some think it is formed by the host as a reaction to the parasite, but the majority opinion is that the parasite itself forms the cyst.

The morphology of both cyst wall and spores has been studied under the electron microscope. Examination with the ordinary optical microscope shows the cyst to be divided internally into compartments and the outer wall to be smooth or rough, depending on the species. Electron microscopy shows numerous, minute villuslike projections on the outer surface (at least for *S. fusiformis*, a parasite of cattle). The spores are crescent-shaped and have a somewhat superficial resemblance to *Toxoplasma* under the light microscope. When viewed with the electron microscope, however, this resemblance is much more marked and extends even to malaria sporozoites. All three have a minute projection, termed a conoid, at the more pointed or anterior end, from which fibrils pass anteriorly. These fibrils are called toxonemes at the point of origin, and sarconemes (in *Sarcocystis*) farther back. The nucleus and several mitochondria are in the posterior half.

Immune reactions. Additional evidence for thinking that there may be a close relationship between *Toxoplasma* and *Sarcocystis* is the frequent occurrence of cross-immune reactions when the Sabin-Feldman dye test is used. This test is ordinarily used for the diagnosis of toxoplasmosis, present or past, but is often positive in cases of sarcosporidiosis.

Life cycle. The life cycle of *Sarcocystis* is quite typical of the Sporozoa, since sexual stages are lacking and schizogony (multiple fission), though sometimes claimed, apparently does not occur. Instead, reproduction is by binary fission, with the eventual development of cysts (Miescher's tubules) in the muscles. These cysts are relatively large structures, easily visible with the unaided eye, and contain myriads of the minute crescentic spores (see **illus.**).

Infection of a new host is believed to be by the ingestion of these spores with food and water contaminated by feces from an infected animal. Most natural infections appear to be mild and to do the host little harm. Of course, some muscle tissue is destroyed but probably too little to be significant. However, sheep with heavy infections of *S. tenella*, one of the best-known species, suffer from loss of weight and give an impression of general unthriftiness. Though a quite potent endotoxin (sarcocystin) has been isolated from the parasites, there is little evidence that this usually does the host much harm. *See* HAPLOSPOREA; PROTOZOA; SPOROZOA.

Reginald D. Manwell

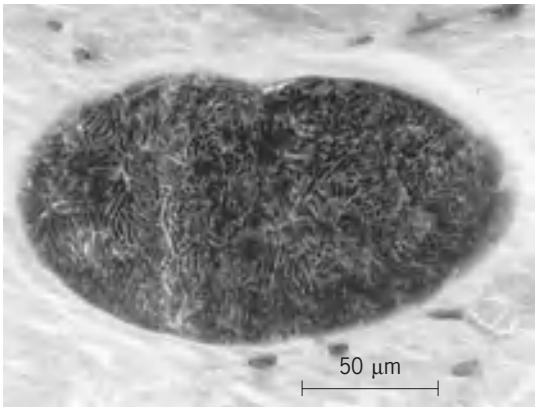

Cross section of Miescher's tube filled with *Sarcocystis* spores (probably *S. lindemanni*) from monkey.

Sarsaparilla

A flavoring material obtained from the roots of at least four species of the genus *Smilax* (Liliaceae). These are *S. medica* of Mexico, *S. officinalis* of Honduras, *S. papyracea* of Brazil, and *S. ornata* of Jamaica, all tropical American vines found in the dense, moist jungles (see **illus.**). The flavoring is used mostly in combination with other aromatics such as wintergreen. *See* LILIALES; SPICE AND FLAVORING.

Perry D. Strausbaugh; Earl L. Core

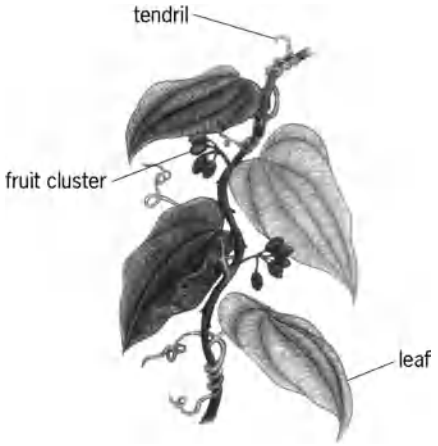

***Smilax aristolochiaefolia*, which yields Mexican sarsaparilla.**

Sassafras

A medium-sized tree, *Sassafras albidum*, of the eastern United States, extending north as far as southern Maine. Sometimes it is only a shrub in the north, but from Pennsylvania southward heights of 90 ft (27 m) or more with diameters of 4–7 ft (1.2–2.1 m) have been reported for this plant. Sassafras is said to live from 700 to 1000 years. It can be recognized by the bright-green color and aromatic odor of the twigs and leaves. The leaves are simple or mitten-shaped (hence a common name "mitten-tree"), or they may have lobes on both sides of the leaf blade (see **illus.**). The wood is soft, brittle, coarse-grained and

***Sassafras albidum*: twig, terminal bud, and leaf.**

somewhat aromatic, and the heartwood is reddish. Because it is durable in contact with moisture, it has been used for fence posts, rail fences, sills for houses, and boats. The aromatic substance, found especially in the roots, is a volatile oil (oil of sassafras) used medicinally as a stimulant and a diaphoretic, and also as a flavoring agent. *See* MAGNOLIALES; STEM; TREE.

Arthur H. Graves; Kenneth P. Davis

Satellite (astronomy)

A relatively small body orbiting a larger one that in turn orbits a star. In the solar system, all of the planets except Mercury and Venus have satellites. Over 150 were known by November 1, 2005, with 154 in well-established orbits and distributed as follows: Earth, 1; Mars, 2; Jupiter, 63; Saturn, 47; Uranus, 27; Neptune, 13; Pluto, 1. More satellites are being discovered each year as a result of deliberate searches by increasingly powerful Earth-based techniques and by the *Cassini-Huygens* mission to the Saturn system. The close flyby of the *Galileo* spacecraft (enroute to Jupiter) of the asteroid 243 Ida in 1993 revealed the presence of a 0.9-mi-diameter (1.5-km) satellite, now known as Dactyl. This unexpected discovery has been followed by the detection of several other asteroid satellites by Earth-based observers. Several Kuiper Belt objects (distant comet nuclei) have also been observed to be binaries. *See* ASTEROID; BINARY ASTEROID; KUIPER BELT.

For planets, it is customary to distinguish between regular satellites that have nearly circular orbits lying essentially in the plane of a planet's equator and irregular satellites whose orbits are highly inclined, very elliptical, or both. The former almost certainly originated with the parent planet, while the latter must be captured objects. The Earth's Moon is a special case. The most widely favored hypothesis for its origin invokes an impact with Earth by a Mars-sized planetesimal, and ejection of material that first formed a ring around the Earth and then coalesced to form the Moon. Pluto's Charon may have formed through a similar collision. *See* MOON; PLUTO; SOLAR SYSTEM.

The two satellites of Mars are tiny objects, less than 9 mi (15 km) in their longest dimension. They are irregularly shaped because the gravitational fields produced by their small masses are too weak to deform the material into spheres. The same is true of the small satellites of the outer planets. *See* MARS.

Jupiter has six groups of small outer satellites, all of which are irregular. Three groups have direct orbits; the other three have retrograde (clockwise) orbits. All six groups have high eccentricities. These six families of satellites presumably represent the capture of six separate objects in an early stage of Jupiter's history when the planet was surrounded by an extended subnebula. The capture process resulted in fragmentation, leading to objects from less than 2 km (1.2 mi) to 185 km (115 mi) in diameter. So far, 32 of these fragments have well-established orbits; there may be over 100 more minor fragments awaiting discovery. The four largest satellites were discovered by Galileo and Simon Marius in 1610. One of these, Io, exhibits an astonishing amount of volcanic activity, driven by tidal forces. Another, Europa, may have an ocean of liquid water trapped beneath an icy crust. The remaining two, Ganymede and Callisto, are both larger than the planet Mercury. Unlike the planet, however, they consist of 50% ice by mass, surrounding a rocky core. *See* JUPITER.

Saturn's inner, icy satellites have the brightest surfaces in the solar system, due in part to a coating of tiny ice crystals from the E ring, which is centered on (and presumably originates from) the satellite Enceladus. This tiny (diameter 500 km or 310 mi) satellite has an unusually warm region near its south pole, where the surface is split by a series of roughly parallel fissures. A plume of gas dominated by water vapor fills the space nearby. The leading hemisphere of Iapetus has a reflectivity of only 4%. It appears to be coated with dark, nitrogen-rich organic material whose origin remains unknown. Titan, the largest of the family, is only slightly smaller than Jupiter's Ganymede and has a nitrogen-dominated atmosphere that is denser than Earth's. Methane and nitrogen in that atmosphere are involved in photochemical reactions that produce organic compounds, some of which contribute to a ubiquitous aerosol layer that hides the surface of the

satellite at visible wavelengths. The *Huygens* probe successfully landed on Titan's surface on January 14, 2005, finding evidence of rivers of liquid methane and deposits of organic aerosols. Saturn has four distinct groups of irregular satellites, implying four discrete capture events, a situation similar to Jupiter's satellite families. *See* ALBEDO; SATURN.

Voyager 2 discovered 10 small regular satellites of Uranus close to the planet during its 1985–1986 flyby. Twelve distant, irregular satellites have subsequently been discovered from Earth. The five large previously known satellites of Uranus are unusually regular, except for Miranda, whose surface reveals evidence of a surprisingly vigorous geological history. *See* URANUS.

Neptune has two highly irregular satellites, Triton and Nereid, plus six small regular moons discovered by *Voyager 2*, in 1989, close to the planet, and five irregular satellites found from Earth. Additional observations of the latter group are required to establish their orbits. Triton has a tenuous atmosphere of nitrogen (a few millionths of the Earth's surface pressure) with traces of methane and carbon monoxide. All three of these gases, plus carbon dioxide and water, are frozen on its surface. That frigid (38 K or −391°F) surface reveals streaks of windblown material and active, eruptive plumes. *See* NEPTUNE.

Ice-covered Charon orbits Pluto, the closest member of the Kuiper Belt, with a period of revolution that is the same as the planet's period of rotation. Thus, Charon is like one of the Earth's artificial geosynchronous satellites: it would appear to hang motionless in the dark skies of Pluto. *See* PLANET.

Tobias C. Owen

Bibliography. J. K. Beatty, C. C. Petersen, and A. Chaikin (eds.), *The New Solar System*, 4th ed., Cambridge University Press, 1999; J. Burns and M. Matthews, *Planetary Satellites*, University of Arizona Press, 1987; D. Morrison and T. Owen, *The Planetary System*, 3d ed., Addison-Wesley, 2003; P. Nicholson, Natural satellites of the planets, in R. Gupta (ed.), *Observer's Handbook*, Royal Astronomical Society of Canada, published annually; D. A. Rothery, *Satellites of the Outer Planets: Worlds in Their Own Right*, 2d ed., Oxford University Press, 1999; F. W. Taylor, *The Cambridge Photographic Guide to the Planets*, Cambridge University Press, 2001.

Satellite (spacecraft)

A spacecraft that is in orbit about a planet (usually the Earth) or a moon. Spacecraft are devices intended for observation, research, or communication in space. Even those spacecraft which are on the way to probe the outer reaches of the solar system usually complete at least a partial revolution around Earth before being accelerated into an interplanetary trajectory. Devices such as sounding rockets follow ballistic (approximately parabolic) paths after fuel exhaustion, but they are not satellites because they do not achieve velocities great enough to avoid falling back to Earth before completing even one revolution. *See* ROCKET ASTRONOMY; SPACE PROBE.

Orbital motion. An observer looking from a remote body might see artificial satellites in orbit around Earth. Their paths would be circular at a nearly constant altitude above Earth's land and water surfaces, or elliptical with Earth's center of mass at one focus. Once in orbit at an altitude above perceptible atmospheric drag, without further propulsive maneuvers a satellite remains in a plane through Earth's center called the orbital plane. This plane travels with Earth in its orbit around the Sun but does not rotate with Earth about its polar axis. Earth rotates around its center in the satellite's orbital plane with a period relative to the fixed stars of one sidereal day, which is 3 min 56 s shorter than the 24-h solar day.

As seen from the remote body, the motion of a satellite in its orbital plane depends on the satellite's orbital period and the inclination of its orbital plane relative to Earth's equatorial plane. In prograde (direct) orbits the satellite moves with an eastward velocity component, in the direction of Earth's rotation. By definition, the inclination of an orbital plane is 0° when it coincides with Earth's Equator and the satellite is traveling eastward. Other prograde orbits have inclinations between 0 and 90°. In retrograde orbits, which have inclinations between 90 and 180°, the satellite moves with a westward velocity component, opposite to that of Earth's rotation. A polar orbit has an inclination of exactly 90° and is neither prograde nor retrograde. The maximum inclination of 180° characterizes an equatorial retrograde orbit. *See* CELESTIAL MECHANICS; ORBITAL MOTION.

Low Earth orbits. The space shuttle, the International Space Station (ISS), and many automated (crewless, robotic) satellites travel in low Earth orbits (LEO) about 100 mi (160 km) or more above Earth's surface. They have typical orbital periods of about 90 min. These satellites have lifetimes of days, weeks, months, or years, depending on their altitudes, their mass-to-drag ratios, and atmospheric drag variations caused by solar activity. The International Space Station would have an orbital lifetime of only a few years without the periodic orbital boosts provided by the space shuttle or other rocket-powered space vehicles such as the Russian *Progress*. Most LEO satellites spend up to nearly half of their time in Earth's shadow. The space shuttle provides its electric power from fuel cells, but almost every other LEO spacecraft depends on solar cells for its power and batteries to operate through the sixteen 35-min "nights" which occur during each 24-h terrestrial day. *See* SPACE SHUTTLE.

Sun-synchronous orbits. Earth is not a perfect sphere. Its rotation causes its equatorial diameter to be 26 mi (42 km) greater than its corresponding polar dimension. For LEO satellites at altitudes below 3700 mi (6000 km), a retrograde inclination slightly greater than 90° may be selected which will cause the orbital plane to rotate eastward at exactly one revolution

per year. This equatorial bulge phenomenon has the desirable result of permitting the plane of such Sun-synchronous orbits, as viewed from the Sun, to remain in the same apparent orientation throughout the year. In more practical terms, if such a Sun-synchronous satellite crosses the Equator in Brazil at 10:00 a.m. on January 1, it will also do so on June 30 or on any other day of the year. Since the orbital plane remains fixed relative to the Earth-Sun axis, the equatorial crossing time also occurs at the same local time at any longitude. This is ideal for weather, Earth resources monitoring, and reconnaissance purposes, because shadows will fall with the same relative length in the same direction and daily weather buildups will be imaged at essentially the same stage from each orbit to the next. *See* METEOROLOGICAL SATELLITES; MILITARY SATELLITES; REMOTE SENSING.

Geosynchronous and geostationary orbits. Earth has a period of rotation relative to the fixed stars of 23 h 56 min 4 s, which is one sidereal day. A satellite orbit of this period is said to be a geosynchronous orbit (GEO). If this orbit is also circular and equatorial, the spacecraft is said to be geostationary, because it remains in a fixed position relative to any observer on the approximately one-third of Earth from which the satellite is visible. The geostationary orbit is also known as the Clarke orbit, after the science fiction writer Arthur C. Clarke, who in 1945 first suggested the use of this particular orbit for communications purposes. Its principal advantage for communications is that, once pointed at the GEO spacecraft, an antenna on Earth never needs to be repointed. *See* COMMUNICATIONS SATELLITE.

A nearly stationary circular GEO satellite with a slight inclination appears to hover over a fixed longitude centered on the Equator and moves north and south in latitude in a very narrow symmetric figure-eight. It takes one sidereal day to complete the figure. The maximum latitude excursion, north or south, equals the orbital inclination. A perfectly positioned GEO satellite is stationed directly over Earth's Equator at its assigned longitude, but is quickly displaced from this location by the combined gravitational pull of the Moon and Sun.

Operators of a geostationary communications satellite must monitor its position regularly to be sure that it remains within its prescribed longitudinal slot, which is assigned internationally by the International Telecommunication Union (ITU) and domestically in the United States of America by the Federal Communications Commission (FCC). Typical orbital slot widths are a few tenths of a degree of longitude. This dimension must be small compared with the spacing between adjacent satellites, so that the ground antenna beamwidth can be wide enough to ignore satellite position variations within the assigned slot but narrow enough to reject interfering signals from adjacent satellite transmitters. The geostationary orbit altitude is 22,237 mi (35,786 km) above Earth's equatorial surface. Below this altitude the satellite will drift eastward; above it, westward. This fact is used when a satellite must be repositioned to a different slot along the geostationary arc.

When a GEO spacecraft approaches a boundary of its assigned position, its operator must fire station-keeping jet thrusters to prevent crossing that boundary. This is no small matter. Historically up to half of the fueled spacecraft mass, upon arrival at the assigned orbital position, has consisted of station-keeping propellant which was exhausted in about a decade, thereby ending that satellite's useful life and requiring the launch of a replacement spacecraft.

In addition to communications, the GEO arc is used for weather observation spacecraft. Three such spacecraft evenly spaced along the Equator can monitor continuously severe weather around the entire globe, with the exception of regions within about 10° of the North and South poles, where hurricanes and tornadoes are absent.

Other types of satellite orbits. There are certain space science phenomena whose investigation requires direct sampling of the local environment, notably magnetic fields and associated ionized plasma particles in Earth's magnetosphere and radiation (Van Allen) belts. Many *Explorer*-class spacecraft [special-purpose smaller satellites, typically 150–500 lb (70–230 kg) mass] have been devoted to such studies. Their orbits have been quite varied. Most have traveled in highly eccentric Earth orbits, characterized by perigees (lowest altitudes) of 100–200 mi (160–320 km) and apogees (highest altitudes) out to lunar distances. *See* SCIENTIFIC AND APPLICATIONS SATELLITES.

During the 1990s, radio navigation satellite systems assumed global importance, from missile mid-course guidance and satellite, aircraft, ship, land vehicle, and hiker position location, to precision aircraft terminal guidance. The two leading systems are the U.S. Global Positioning Satellite (GPS) constellation and the Russian GLONASS system. GPS employs eight satellites in each of three highly inclined circular orbits at altitudes intermediate between LEO and GEO. GLONASS is similar. **Table 1** enumerates and gives examples of various satellite missions and payloads. *See* SATELLITE NAVIGATION SYSTEMS.

Configuration and structure. The size and shape of a spacecraft is almost always dictated primarily by its mission requirements. The principal constraints are usually imposed by the dimensions and shape of the satellite payload provisions of the launch vehicle. An important requirement of virtually all powered automated spacecraft is sufficient solar cell mounting area both to power the payload in sunlight and to charge its batteries to continue payload operations during solar eclipse periods. Another requirement is to provide spacecraft attitude stabilization and control so that sensors and antennas can be pointed in the required directions. *See* SPACECRAFT STRUCTURE.

Spin-stabilized satellites. In the 1960s and 1970s, many smaller *Explorer* and communications satellites were launched by vehicles such as the Delta, which used a spin-stabilized (same principle as a spinning rifle bullet) upper stage. This led to a spin-stabilized cylindrical (drum-shaped) satellite, symmetric about its longitudinal axis for ease of spin balancing (similar to automobile tire spin balancing) before launch, and

TABLE 1. Types of satellite missions and payloads

Mission	Payload	Example
Communications		
Full-duplex broadband	Transceiver	Intelsat, Milstar
Message broadcast	Transmitter	DirecTV, GPS
Personal communication	Transceiver	Globalstar
Remote sensing		
Imaging	Imagers and cameras	Landsat, Hubble Space Telescope
Intensity measurement	Radiometer	Chandra X-Ray Observatory
Topographic mapping	Altimeter	TOPEX/Poseidon
Navigation		
Ranging	Transceiver	TDRS
Navigation signal	Clock and transmitter	GPS, GLONASS
In situ science		
Crewed	Physical and life sciences	International Space Station
Robotic	Sample collection and return	LDEF/space shuttle
Other		
Microgravity manufacturing	Physical plant and raw materials	International Space Station
Space medicine	Medical laboratory, centrifuge	International Space Station

SOURCE: After B. Chesley, R. Lutz, and R. F. Brodsky, Space payload design and sizing, Ch. 9 in J. R. Wertz and W. J. Lawson (eds.), *Space Mission Analysis and Design*, 3d ed., Kluwer Academic, 1999.

axially oriented in the desired direction (for example, parallel to Earth's polar axis) by the launch vehicle before separation. The usual choice for solar cell mounting was the curved sides of the cylinder, leaving the flat ends free for sensors and antennas with no obstructions over an entire hemisphere. The cylindrical satellite exterior was also ideal for locating the spacecraft inside the launch vehicle protective fairing (nose cone), which usually consists of a cylindrical section of the same diameter as the rocket's fuel tanks topped by a conical tip for thermal protection from the nose stagnation temperature.

A serious disadvantage of spin-stabilized cylindrical satellites is the limitation on solar-cell area. Because of its cylindrical shape, at any given time half of the cells are shadowed by the satellite itself, and the remainder lose about one-third of their efficiency as compared to a flat solar panel, both perpendicular to the satellite-Sun line, resulting in an overall efficiency of the cylindrical geometry of 33% of that of the flat solar panel of the same area.

Three-axis-stabilized satellites. As payloads became larger, heavier, and more power-hungry, the body-mounted solar cell area limitation became more troublesome because satellite volume and weight increase as the cube of its length or diameter while its surface area increases only as its length squared. The solution to this dilemma for the larger satellites is to switch from spin stabilization to three-axis stabilization (without spin), usually with internal gyroscopes to maintain stability about each axis. These requirements lead to a rectangular or cubic box configuration, with symmetric solar panels mounted at each end of a shaft running through the satellite's center of mass, to minimize drag-induced torques. This shaft is usually perpendicular to the orbital plane, and is rotated to increase solar cell illumination by making the line perpendicular to the solar panels fall into the plane defined by the shaft and the satellite-Sun line. This box configuration has other manufacturing advantages associated with the ability to assemble each box face individually to check out subsystems thoroughly before final assembly of the entire spacecraft with its payload. A major disadvantage of the three-axis stabilized configuration is the necessity to fold the solar panels and shafts for launch, followed by deployment after satellite injection into its final orbit. Such on-orbit mechanical deployments are a major cause of mission failures and consequent insurance cost increases.

Because the space shuttle is reusable dozens of times, it was designed with a cargo bay which can accommodate payloads in a cylindrical configuration 15 ft (4.5 m) in diameter and up to 60 ft (18 m) long.

In the free-fall conditions of orbital flight, objects are essentially weightless but possess their usual mass. Consequently it is customary to refer to spacecraft component masses in kilograms, rather than in pounds-mass. Early satellites were limited to a few kilograms, but launch vehicles such as the space shuttle, the Titan IV, and the Russian Proton can place more than 20,000 kg (44,000 lbm) into low Earth orbit.

Spacecraft construction materials. Aluminum is the most common spacecraft construction material because of its favorable combination of light weight, low cost, adequate rigidity and strength, and good electrical and thermal conductivity. Certain esoteric materials such as beryllium, magnesium, and graphite-reinforced fiber plastics (GRFP) have special properties such as lighter weight, better strength-to-weight ratio, or superior rigidity. However, they are expensive and difficult to fabricate. In addition, beryllium dust is hazardous (carcinogenic) to humans, magnesium with oxygen is quite exothermic (burns at a high temperature if ignited), and GRFPs may require special treatment to prevent electrical interference or thermal hot spots because of their relatively low conductivity.

Thermal control. Satellite exteriors are usually covered partly with thermal insulation (multiple layers

of aluminized plastic) to isolate the spacecraft from the radiated temperature extremes of the hot Sun (6000 K or 10,000°F) or cold deep space (few kelvins). The remaining area is usually covered by passive thermal control surfaces designed to maintain the satellite interior at about 68°F (20°C) for optimum operation and lifetime of solid-state electronics and about 18°F (10°C) cooler for nickel-hydrogen batteries. For example, satellite exterior white paint with its high emissivity and low absorptivity works to reduce the satellite temperature, while black interior paint with its high emissivity and absorptivity helps to maintain a more uniform temperature throughout the satellite's interior.

Power systems. Except for the brief trips of the space shuttle, most satellites obtain their electrical power from the Sun through large areas of solar cells. The solar panels are sized to carry the payload electrical load in sunlight plus enough additional power, stored in batteries, to power the payload during solar eclipse periods. Because solar cells degrade over time because of solar bombardment by ultraviolet radiation and energetic particles, the panels must be sized to produce the amount of power which will be required at the planned satellite end of life, rather than at the beginning of the mission. *See* SOLAR CELL; SPACE POWER SYSTEMS.

Launch vehicles. All current launch vehicles use chemical combustion of a fuel and oxidizer to generate a hot, high-pressure gas which is expanded through a supersonic nozzle to generate thrust. Solid-propellant motors cast the fuel and oxidizer together in a rubbery matrix. Liquid fuels and oxidizers are stored in separate tanks and pressure-fed into the combustion chamber just ahead of the nozzle. Liquid cryogenic hydrogen and oxygen constitute the most

TABLE 2. Launch systems characteristics[1]

Country	Launch system	Upper stage (if any)	Payload mass to low Earth orbit (LEO)[2], kg[5]	Payload mass to geosynchronous transfer orbit (GTO)[3], kg[5]	Payload mass to geosynchronous orbit (GEO), kg[5]	Payload mass to polar orbit,[4] kg[5]	Payload accommodations		
							Launch site[7]	Diameter, m[6]	Length, m[6]
United States	Atlas I	Centaur-1	—	2,255	—	—	A	3.3	10.4
	Atlas II	Centaur-2	6,580	2,810	570	5,510	A	4.2	12.0
	Atlas IIAS	Centaur-2A	8,640	3,606	1,050	7,300	A	4.2	12.0
	Delta II								
	6920/25	Star 48B	3,990	1,450	730	2,850	A, B	2.9	8.5
	7920/25	Star 48B	5,089	1,840	910	3,890	A, B	2.9	8.5
	Pegasus	—	375	—	—	—	Aircraft	1.3	4.4
	Pegasus XL	—	460	—	—	345	Launch[8]	1.3	4.4
	Shuttle		24,400	—	—	—	A	4.5	18.0
		Inertial Upper Stage	—	5,900	2,360	—			
	Taurus	STAR 37	1,400	450	—	1,060	A, B	1.4	2.8
	Titan II	No upper stage	—	—	—	1,905	B	2.8	3.7, 5.2, 6.7
	Titan IV	No upper stage	—	—	—	14,110	A, B	4.5	9.7
		Centaur	—	8,820[9]	4,540	—	A, B		12.8
		Inertial Upper Stage	—	8,350	2,380	—	A, B		15.8
		No upper stage[9]	21,645	—	—	18,600	A, B		18.0
France (European Space Agency)	Ariane 4	H-10	4,900	2,050	—	3,900	C	3.7	3.9 S
	42P	H-10	6,100	2,840	—	4,800			4.9 L
	42L[10]	H-10	7,400	3,380	—	5,900			6.5 XL
	44P	H-10	6,900	3,320	—	5,500			
	44LP[11]	H-10	8,300	4,060	—	6,800			
	44L	H-10	9,600	4,520	—	7,700			
	Ariane 5	L9	18,000	6,800	—	12,000	C	4.5	12.0
Japan	H-2	—	10,500	4,000	2,200	6,600	D	3.7, 4.8	3.5–9.0
China	Long March								
	CZ38	—	13,600	4,500	2,250	—	E	3.8	6.0
	CZ4	—	4,000	1,100	550	—	F	3.0	3.9
Russia	Proton	D1	20,900	—	—	—	G	4.1	15.6
		D1e	—	5,500	2,200	—	G	4.0	7.5
	Proton K	DM	20,100	4,815	2,100	—	G	—	—
	Proton M	Breeze M	22,000	5,100	2,500	—	G	—	—
	Zenit 2	—	13,740	5,180	1,535	11,380	G	3.4	5.9–8.4

[1]Characteristics for existing systems to 28.5° inclination, unless specified otherwise.
[2]Approximately 185 km (115 mi) circular orbit.
[3]Geosynchronous transfer orbit from the LEO altitude to the GEO altitude.
[4]90° inclination and approximately 185 km (115 mi) circular orbit.
[5]1 kg = 2.20 lbm.
[6]1 m = 3.28 ft.
[7]A = Kennedy Space Center. B = Vandenberg Air Force Base. C = Kourou Launch Center. D = Tanegashima Space Center. E = Xichang (Sichuan) Satellite Launch Center. F = Taiyuan Satellite Launch Center. G = Tyuratam (Baikonur).
[8]Carrier aircraft can stage from various locations.
[9]With solid rocket motor upgrade.
[10]With two liquid rocket boosters.
[11]With two liquid and two solid rocket boosters.
SOURCE: After J. P. Loftus, Jr., C. Teixera, and D. Kirkpatrick, Launch systems, ch. 18 in J. R. Wertz and W. J. Lawson (eds.), *Space Mission Analysis and Design*, 3d ed., Kluwer Academic, 1999.

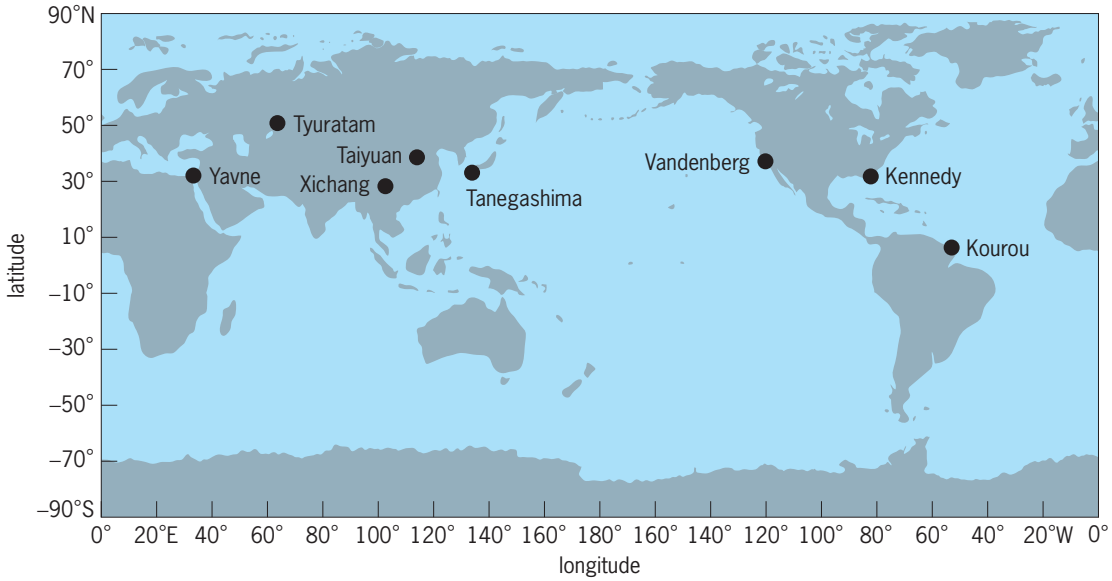

Seven major launch sites that can do orbital launches. Their location and range safety considerations determine acceptable azimuths. (*After J. P. Loftus, Jr., C. Teixera, and D. Kirkpatrick, Launch Systems, Ch. 18 in J. R. Wertz and W. J. Larson, eds., Space Mission Analysis and Design, 3d ed., Kluwer Academic, 1999*)

powerful fuel and oxidizer combination in present use. Kerosene is sometimes substituted for hydrogen in a launch vehicle first stage, to reduce the fuel tank size because of the higher density of kerosene as compared with hydrogen.

The space shuttle uses a combination of liquid and solid motors. At liftoff the three main liquid hydrogen–liquid oxygen motors are assisted by two large solid-rocket boosters, which burn out 123 s after launch and are recovered from the ocean, refurbished, and refueled for future use. The liquid engines are used to complete the orbit insertion process, near the end of which the large external hydrogen-oxygen tank is jettisoned. Unlike the reusable space shuttle, expendable (automated, crewless) vehicles are used only once. *See* ROCKET PROPULSION.

Table 2 summarizes the capabilities of the most important launch systems from the United States, Russia, China, Japan, and France, the five world leaders in the field. The **illustration** provides a map of the launch sites referred to in the table.

A few spacecraft have been launched from international waters. Italy used *Scout* to launch *Uhuru*, an early x-ray astronomy satellite, from a converted oil-drilling platform in the Indian Ocean east of Kenya. Heavier spacecraft have been launched from the floating *Odyssey* platform in the low-latitude Pacific Ocean.

Space stations. The Russian space station *Mir* was deorbited in a controlled reentry into the South Pacific Ocean on March 23, 2001. Meanwhile the International Space Station construction proceeded so well that the station was occupied by its first permanent crew of two cosmonauts and one astronaut mission commander by the end of 2000. In March 2001 the space shuttle made the first three-person crew exchange, and it was planned to make such exchanges at similar intervals for the foreseeable future. However, the 2003 *Columbia* tragedy initiated a two-year hiatus before a shuttle returned to the International Space Station, during which the station crew was limited to one astronaut and one cosmonaut transported by the three-person Russian *Soyuz*. The crew size may be increased if and when a United States Crew Return Vehicle has been developed or an additional *Soyuz* is dedicated to this use. *See* SPACE FLIGHT; SPACE STATION.

John F. Clark

Bibliography. *Jane's All the World's Aircraft*, biennially; J. R. Wertz and W. J. Larson (eds.), *Space Mission Analysis and Design*, 3d ed., Kluwer Academic, 1999.

Satellite meteorology

The branch of meteorological science that uses meteorological sensing elements on satellites to define the past and present state of the atmosphere. Meteorological satellites can measure a wide spectrum of electromagnetic radiation in real time, providing the meteorologist with a supplemental source of data.

Since the 1960 launching of *TIROS 1*, the first satellite designed for direct visual sensing of the atmosphere, sensors on satellites have become more sophisticated. Modern satellites are sent aloft with multichannel high-resolution radiometers covering an extensive range of infrared and microwave wavelengths. Radiometers sense cloudy and clear-air atmospheric radiation at various vertical levels, atmospheric moisture content, ground and sea surface temperatures, and ocean winds, and provide visual imagery as well. *See* METEOROLOGICAL SATELLITES.

Satellite types. There are two satellite platforms used for satellite meteorology: geostationary and polar. Geostationary (geo) satellites orbit the Earth

at a distance that allows them to make one orbit every 24 hours. By establishing the orbit over the Equator, the satellite appears to remain stationary in the sky. This is important for continuous scanning of a region on the Earth for mesoscale (approximately 10–1000 km horizontal) forecasting. Geo platforms have been operating since the first *GOES* satellite launch in the late 1970s. The biggest liability to geo platforms is their cost (it takes a lot of fuel to get the satellite to this orbit), and their distance from Earth (22,000 mi; 35,000 km). At this distance, it takes much more sophisticated navigation and control to scan the Earth with accuracy to enable images to be looped (animated) smoothly and to be useful for co-locating satellite-sensed data with data gathered on the Earth. This sophistication also adds significantly to cost. The biggest advantage of geosatellites is that they can image a specific place on the Earth as frequently as every 2 minutes, allowing for unprecedented animation of meteorological phenomena, such as severe weather.

Polar satellites orbit the Earth in any range of orbital distances with a high inclination angle that causes part of the orbit to fly over polar regions. The orbital distance of 100–200 mi (160–320 km) is selected for meteorological applications, enabling the satellite to fly over a part of the Earth at about the same time every day. With orbital distances of a few hundred miles, the easiest way to visualize the Earth-satellite relationship is to think of a satellite orbiting the Earth pole-to-pole while the Earth rotates independently beneath the orbiting satellite. The advantage of polar platforms is that they eventually fly over most of the Earth. This is important for climate studies since one set of instruments with known properties will view the entire world. Another advantage is the low orbit. Since the instruments are closer to the Earth, the optics and physical constraints (especially in the microwave region) allow the instrumentation to achieve high spatial resolution at much lower cost than geo platforms. Fuel costs for launch are lower, allowing more and heavier instrumentation to be flown. The biggest liability for polar satellites is that they pass over a given spot on the Earth only twice per day. The high-resolution imagery is spectacular, but continuity is lost when the imagery is looped. For this reason, three polar satellites are flown in staggered orbits, so approximately six polar overflights occur every day over a midlatitude location. With only six passes per day over a location, this satellite is used more for climate studies than mesoscale analysis. The one exception is in the Polar Regions. There the spacecraft passes over the high latitudes on each orbit. This is very useful for high-latitude regions such as Alaska, northern Canada, and arctic regions as well as the Antarctic since geo satellites can scan only to about 75°N and 75°S latitude.

Data presentation. The enormous aerial coverage by satellite sensors bridges many of the observational gaps over the Earth's surface. Satellite data instantaneously give meteorologists up-to-the minute views of current weather phenomena.

Visual. Images derived from the visual channels are presented as black and white photographs. The brightness is solely due to the reflected solar light illuminating the Earth. Visible images are useful for determining general cloud patterns and detailed cloud structure, and they provide inherently higher resolution than the infrared channels. In addition to clouds, visible imagery shows snowcover, which is useful for diagnosing snow amount by observing how fast the snow melts following a storm. Cloud patterns defined by visual imagery can give the meteorologist detailed information about the strength and location of weather systems, which is important for determining storm motion and provides a first guess or forecast as to when a storm will move into a region. *See* CLOUD.

Infrared. More quantitative information is available from infrared sensors, which measure radiation at longer wavelengths (from infrared to microwave). Through most of the infrared spectrum, the Earth's surface and clouds absorb and reradiate as a blackbody, that is, they emit energy at infrared wavelengths with both a peak wavelength and intensity dependent on the temperature of the radiator (atmosphere, ground, or cloud). The Earth is also a very efficient absorber of infrared radiation so that infrared radiation observed by the satellite even during the daytime is coming from the Earth's radiation and is not reflected from the Sun. Thus by analyzing the infrared data, the ground surface, cloud top, and even intermediate clear air temperatures can be determined 24 hours a day. By relating the cloud top temperature in the infrared radiation to an atmospheric temperature profile from balloon data, cloud top height can be estimated. This is a very useful indicator of convective storm intensity since more vigorous convection will generally extend higher in the atmosphere and appear colder. *See* BLACKBODY; PLANCK'S RADIATION LAW.

At a few specific infrared and microwave bands, the atmospheric constituent gases absorb and reradiate. Carbon dioxide (CO_2), water vapor (H_2O), nitrous oxide (N_2O), oxygen (O_2), and ozone (O_3) have significant absorption bands at the wavelengths, shown in **Fig. 1**. These infrared radiation active regions are generally due to quantum-mechanical vibrational-rotational molecular absorption. As shown, the bands do not have sharp, steep sides; instead, absorption falls off gradually. The absorption band edges are called tails. The tails are important since they are modified by pressure through a quantum-mechanical process called pressure broadening. The result is that absorption occurring far out on a tail is associated with trace-gas absorption low in the atmosphere (at high pressures). Toward the center of the absorption band are the absorption properties of the gas higher and higher in the atmosphere. As energy radiates from the Earth's surface, it interacts with the trace gases, which modify the radiation by absorption and reradiation at each point up to the top of the atmosphere. The net result is sensed by the satellite. Infrared radiation sensors are tuned to look at different portions

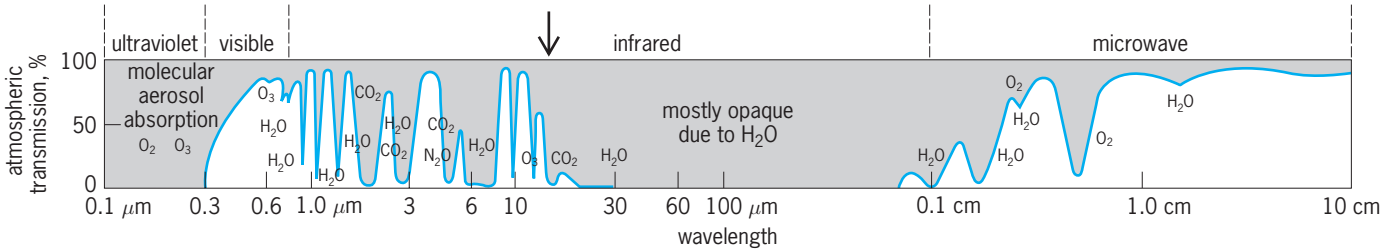

Fig. 1. Radiational transmission coefficient for a 1-km-deep (0.6-mi) atmospheric layer (100 to 0%) from ultraviolet to microwave for various atmospheric gases. 100% indicates total transmission (all wave energy passes out of the atmosphere). 0% indicates that the gas is opaque to radiation (the energy is absorbed by the gases indicated on the diagram). The arrow shows wavelengths where CO_2 is a partial absorber. These are the bands for which sensors (shown in Fig. 2) are designed to recover vertical temperature profiles. (*After D. Houghton, ed., Handbook of Applied Meteorology, Wiley, 1985*)

of the absorption tail, making it possible to establish the thermal characteristics of the atmosphere through different vertical regions or layers. Infrared radiometry is the science used to determine the thermal atmospheric profile based in part on the observed radiation intensity at the specific wavelengths that correspond to the different atmospheric levels (**Fig. 2**). Geostationary satellites have about 13 channels devoted to measuring CO_2 temperature at various atmospheric levels. The CO_2 absorption spectra is used to measure the air temperature because CO_2 exists in uniform proportion to air throughout the atmosphere. In addition to CO_2, O_2 absorption is used for thermal profiling in the microwave region. In general, the sensors mounted on both polar orbiting and geosynchronous satellites give highly useful results. *See* ELECTROMAGNETIC RADIATION; HEAT RADIATION; INFRARED RADIATION; QUANTUM MECHANICS.

Microwave. Microwave radiation is not impeded by clouds, as in the infrared bands, since radio waves can propagate through clouds without attenuation. However, the layer of atmosphere sensed by the microwave channels is greater than can be resolved by the infrared radiation, and microwave sensors have poorer horizontal resolution. The reason for the lower horizontal resolution is that longer microwave wavelengths require a large radio receiving antenna instead of optics, which are used for visible and infrared radiation measurements. Microwave sensing is constrained to polar satellites only, since it is currently impractical to fly the enormous microwave antenna to achieve comparable horizontal resolution of a polar spacecraft at geosynchronous altitudes. *See* MICROWAVE.

Wind measuring. The advent of geosynchronous satellites allowed the position of cloud elements to be traced over time. These cloud movements can be converted to winds, which can provide an additional source of data in an otherwise unobserved region. These techniques are most valuable for determination of mid- and high-level winds, particularly over tropical ocean areas. Other applications have shown that low-level winds can be determined in more spatially limited environments, such as those near thunderstorms, but those winds become more uncertain when the cloud elements grow vertically into air with a different speed and direction (a sheared environment).

Microwave data from polar satellites can measure sea surface roughness since the emissivity of the ocean changes when it contains foam, small waves, or ripples. The degree of the emissivity change can be correlated with surface wind speed. Wind direction is not possible using the microwave technique, but the speed can be inferred. *See* WIND.

Moisture. Like the other constituent gases of the atmosphere, water vapor has its unique set of absorption bands. Unlike CO_2, water vapor is not a fixed percentage of total mass in an air column. With sensors that measure radiation in the moisture bands, a complete vertical moisture profile can be developed by the same basic methods used to retrieve temperatures by studying the tails of the absorption bands. There is a very strong water vapor absorption band

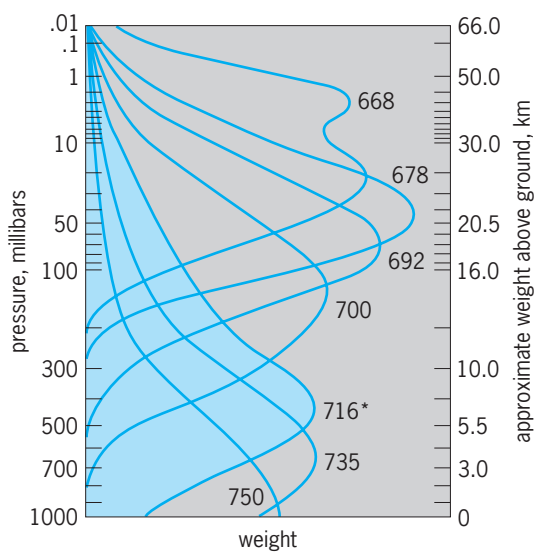

Fig. 2. Plots of transmissivity variation with height (weighting factors) for seven wavelengths (or wavenumbers, given in cm^{-1}) in the 15-m band for CO_2. The shaded weight factor representing wavelength 13.966 μm starred wavenumber 716 cm^{-1}) shows the atmospheric layers, which influence the radiance measurement at the satellite. Retrieved temperature is most representative of the layer where this weight is maximum (about 6.2 km or 3.9 mi), but note that atmospheric temperatures at levels as high as 30 km (19 mi) and as low as the surface significantly influence this value. Other curves indicate weighting factors for the other six sensor channels. Seven sensors will yield seven layer temperatures, which can be reassembled into a vertical temperature profile. 1 millibar = 102 Pa; 1 km = 0.6 mi. (*After D. Houghton, ed., Handbook of Applied Meteorology, Wiley, 1985*)

around 6.8–7.0 micrometers (Fig. 1). It is so strong that it acts like a cloud in the infrared radiation, and virtually all the sensed radiation is coming from the topmost water vapor layer in the atmosphere. Image loops of this water vapor infrared radiation imagery reveal many of the air motions in clear air. Work is under way to derive water vapor winds much like cloud drift winds only by sensing water vapor in this part of the infrared radiation spectrum.

At infrared radiation wavelengths near 12 μm, there exists a very weak water vapor absorption band that can be exploited to measure the total water vapor in cloud-free areas since water will only slightly attenuate the infrared radiation. Near this water vapor band is the infrared radiation window at 11.2 μm. At 11.2 μm, no absorption takes place from any atmospheric gas. Total water vapor in an atmospheric column can be estimated by comparing radiances from these two very close wavelengths—since if there were no water vapor, the radiation in the 12 μm region would show no absorption, as would the measurement at 11.2 μm. A decrease in intensity in the 12 μm measurements is directly related to total water vapor in an atmospheric column.

Applications. By using a wide variety of sensors, satellite data provide measurements of phenomena from the largest-scale global heat and energy budgets down to details of individual thunderstorms. Having both polar orbiting and geosynchronous satellites allows coverage over most Earth locations at time intervals from 3 minutes to 3 hours.

Global (large) scale. Polar orbiting satellites have had an impact on the monitoring of global climate and climate change. Long-term observational programs, notably the Earth Radiation Budget Experiments (ERBE), active since the late 1970s, have produced information in the form of colorized images of the outgoing long-wave radiation, a quantity that can aid in detection of long-term warming or cooling. Satellites have also been used to monitor atmospheric ozone, a gas that protects the Earth from harmful solar radiation. The ozone hole that forms over the Antarctic during October probably would not be known even today had it not been for polar orbiter satellite data.

Satellites can closely monitor land and sea temperatures and show anomalies that produce phenomena such as El Niño, which shifts the storm track of the eastern Pacific region. Satellite monitoring of the polar ice packs helps identify important seasonal variations. Ice caps and areas where vegetation is decreasing (owing to drought or human exploitation) modify the albedo (fraction of incident radiation reflected) and influence the effective solar radiation reaching the Earth. *See* ALBEDO; INSOLATION.

The great semipermanent cloud bands of the Earth, such as those along the Intertropical Convergence Zone, can be monitored by either polar orbiting or geosynchronous satellites. Long-term perturbations in the tropical regions can lead to excessive rain or drought and influence the frequency of tropical storms. Long-term averaged satellite cloud imagery is a basis for estimating rainfall in the tropics.

With observations using space-borne instruments carried out for many years, satellites are the prime tools for monitoring global climate and climatic change. *See* CLIMATE HISTORY.

Storm (synoptic) scale. Superimposed on the great atmospheric circulations are the transient eddies or cyclones, which ultimately maintain the heat balance of the Earth by moving warm air northward and cold air southward. Of course, these eddies and swirls are the major weather-producing phenomena over most of the globe. Storms are regions of concentrated air circulation (or relative vorticity), with rotation counterclockwise in the Northern Hemisphere and clockwise in the Southern Hemisphere. The centers of the eddies are generally associated with relative vorticity maxima. As these centers move along in the background global circulation, certain adjustments of the atmosphere produce vertical lift of the air, leading to the clouds and weather experienced at a ground location. The spiral or comma-shaped patterns often seen in visual imagery are in reality reflections of these vorticity centers moving through the atmosphere (**Fig. 3**). As the intensity of the circulation increases, the cloud material, which has been generated ahead of the circulation center by lifting, begins to deform, eventually resulting in a spiral or comma-shaped cloud band. In clear air, the spiral shape can be seen in the water vapor imagery. The cloud band or tail elongates in regions of horizontal speed variation (Fig. 3*f*). Elongated cloud bands are common in regions where winds stretch neighboring air parcels apart (equatorward of jet stream bands) in regions of atmospheric deformation. Geostationary satellites globally monitor temperature, moisture, and cloud

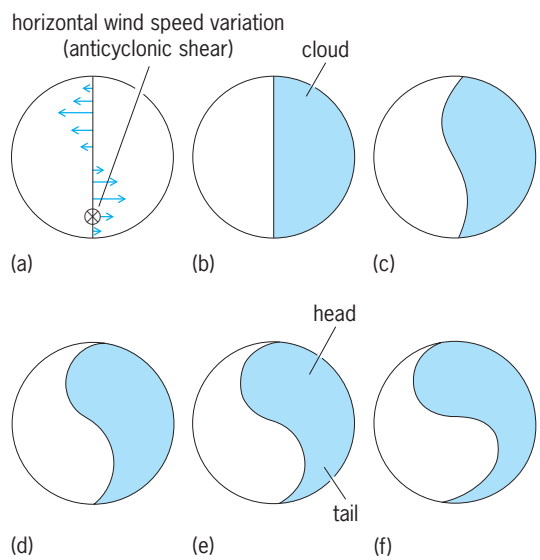

Fig. 3. Schematic diagram of a circulation center or vorticity maximum and the formation of the comma-shaped cloud pattern. (*a*) Relative wind circulation (wind vectors along a north-south axis). (*b*) Lifting forms cloud (shaded region) ahead of the center, and with time (*c*) 6 h, (*d*) 12 h, (*e*) 18 h, and (*f*) 24 h later, the circulation transforms the cloud into the comma shape seen typically in satellite imagery. The comma head is in the north, the tail is in the south. (*After R. Weldon, Cloud Pattern and Upper Level Winds, NESDIS Satellite Training Course, NOAA, 1979*)

patterns associated with these systems. The satellite measurements are combined with other weather data to initialize computerrun mathematical models, which routinely forecast major cyclones.

In tropical latitudes, widespread convective clouds or thunderstorms leave much cloud debris and moisture at very high levels. Hadley cell circulation moves this air poleward; the northward limit of this air at high levels is the subtropical jet stream. Satellite imagery typically shows extensive elongated bands of high cirrus clouds at these latitudes (10–30°N) with very sharp cloud edges. Many cloud bands seen on satellite imagery correspond closely to regions of the atmosphere where neighboring air parcels are pulled apart along the so-called axis of dilation in a zone of atmospheric deformation. Jet streams are characteristically associated with deformation zones, so the position of cloud bands can often be used to locate these strong wind currents. *See* ATMOSPHERIC GENERAL CIRCULATION; FRONT; JET STREAM.

Figure 3*f* shows how the elongation of the tail of the comma cloud appears in a region where neighboring air parcels are in horizontal shear. Thus, the same general rule applies in locating the major jet stream bands existing in the intense westerly currents at higher latitudes. Midlatitude jet stream flows indicate the position of the polar front, where there exists an intense temperature contrast through the depth of the troposphere. The importance of locating these frontal zones or jets is that the eddies or cyclones previously discussed typically have their origin here. The meteorologist, by noting past locations and apparent qualitative strength of jet streams and storms, and by using satellite-improved, model-generated forecasts, is able to make more accurate weather predictions for a particular location.

The information that can be extracted from water vapor imagery is important for analysis of synoptic-scale upper-level flow patterns. Color infrared radiation images are used primarily for cloud top and surface temperature (see **colorplate**). **Figures 4** and **5** illustrate the application of water vapor (WV) imagery in analyzing 500-hectopascal

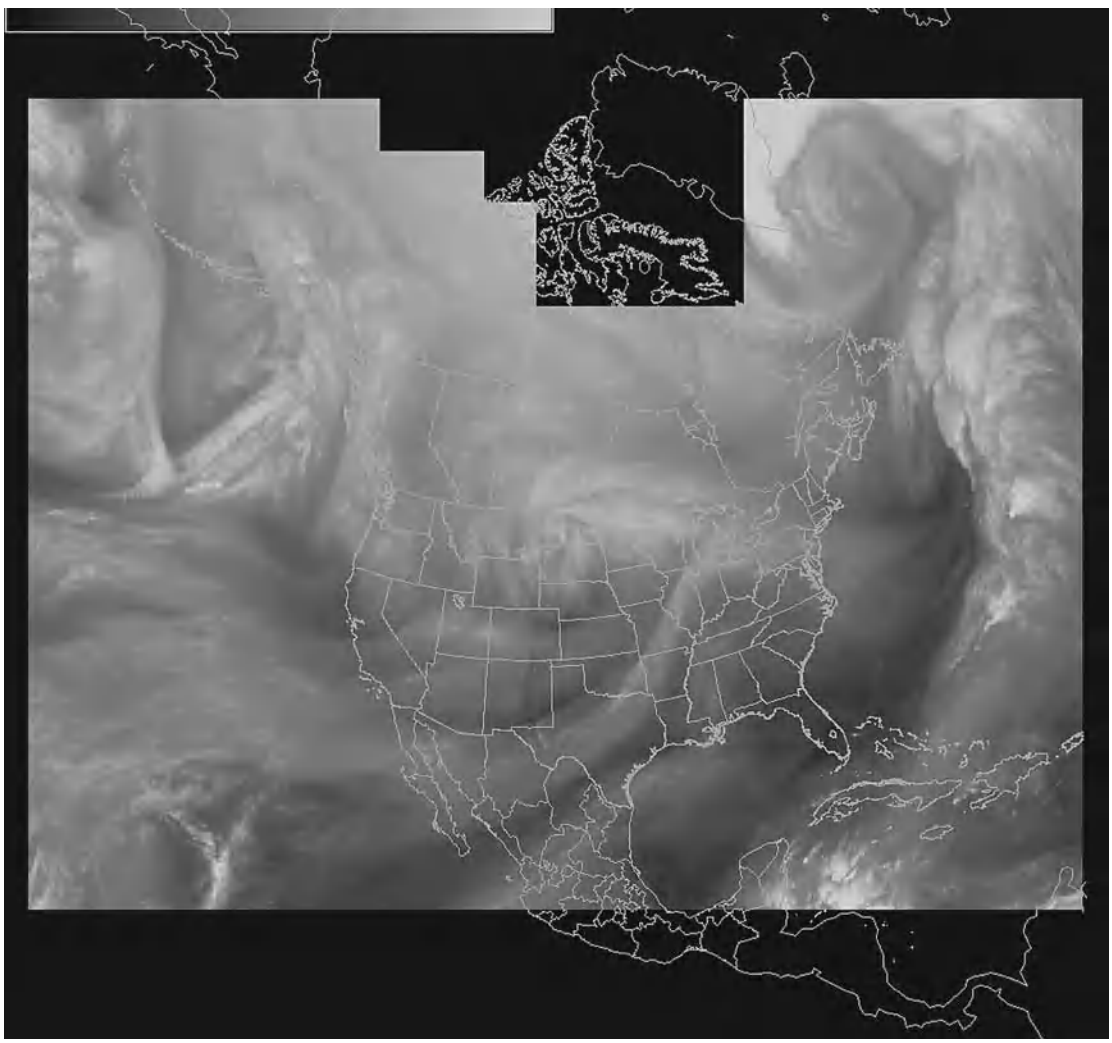

Fig. 4. Water vapor image at the same time and location as in the colorplate. Water vapor imagery can show features in clear areas such as the dark variations imaged over northern Kansas, central California, and another extending from northern Baja, southern Arizona, and New Mexico, up through the Oklahoma panhandle. Each of these areas indicates an upper-level "short wave" that is not evident at all in the cloud image in the window infrared radiation (see colorplate). The closed low pressure in southern Greenland and the system in the Gulf of Alaska are also evident in this image.

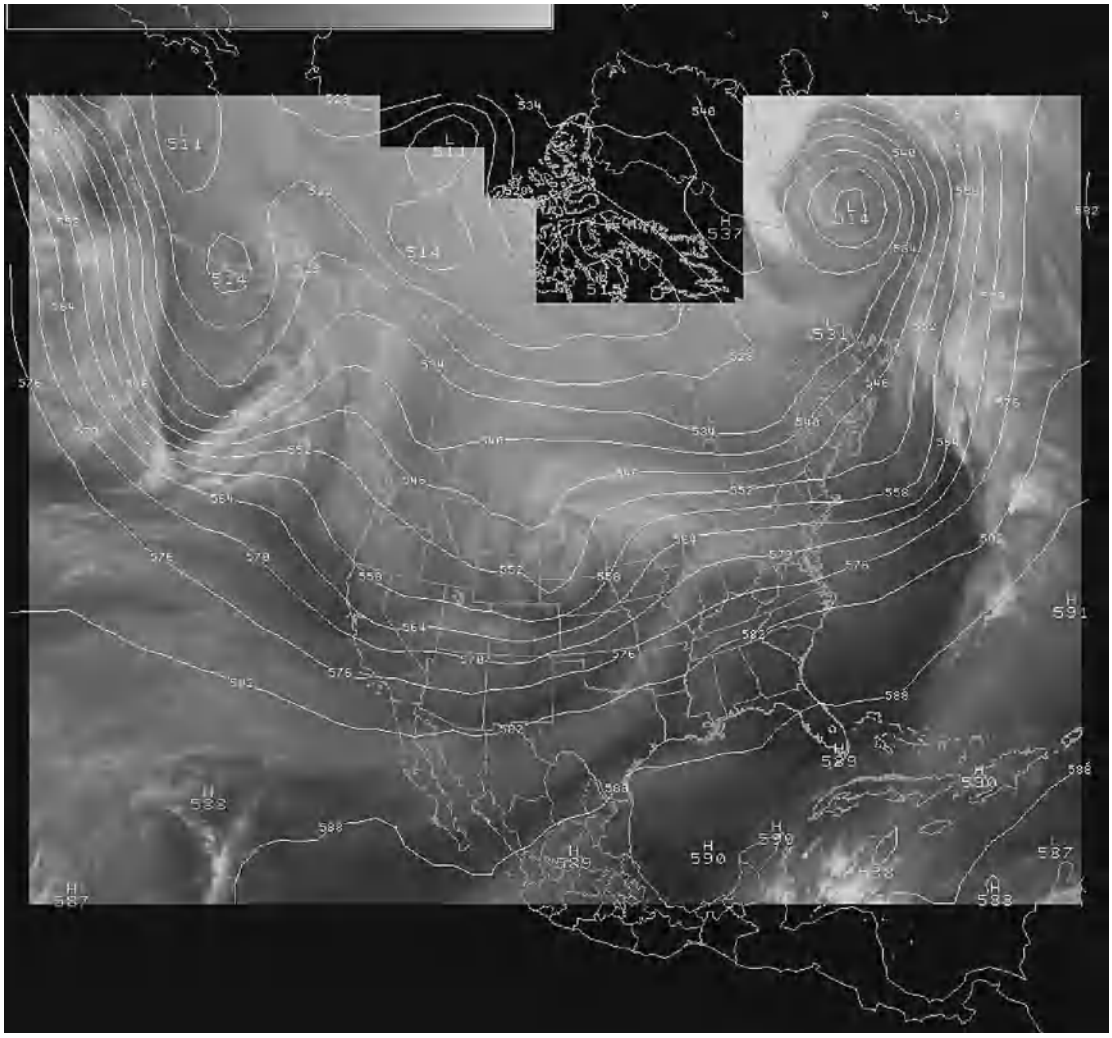

Fig. 5. Same water vapor image as shown in Fig. 4 but with contours depicting 500-hectopascal heights. The analysis has a resolution of about 60 km, whereas the image can show up to 8-km detail. The upper-level height contours align well with the water vapor imagery features. The deep trough in the Gulf of Alaska also is depicted well; the analysis indicates a very weak closed low. The imagery shows the complete circulation in the closed low over southern Greenland. South of that location can be seen another front and corresponding upper-level trough over the Atlantic Ocean. Other subtle features are the short wave over Nebraska with the dark (warm) band at its eastern edge. Another short wave is seen extending from central Missouri south through Oklahoma and the Texas panhandle. In addition, the imagery shows a similar short wave over California; in fact, the short wave might be composed of three parts, which would be nearly impossible to resolve in the height analysis. The extra detail over California is typical of the added resolution that can be inferred from satellite data. (*NOAA Forecast Systems Laboratory*)

(millibar) heights. Image and analysis overlays are very useful for determining whether the analysis is in agreement with atmospheric circulations (Fig. 5). If the image features and analysis contours are not well aligned, it implies that model forecasts initialized from the analysis may not produce an accurate forecast. This is especially important for systems over oceanic regions where conventional data are not available for the analysis. Research is focusing on methods to extract image features and automatically incorporate them into the analysis used for model forecasts. However, for now the imagery is primarily used to assess the analysis quality as has been demonstrated here.

Tropical storms. The greatest gain with the introduction of weather satellites was in early detection, positioning, and monitoring of the strength of tropical storms (hurricanes, typhoons). Lack of conventional meteorological data over the tropics (particularly the oceanic areas) makes satellite data indispensable for this task. The hurricane is one of the most spectacular satellite images. The exact position, estimates of winds, and qualitative determination of strength are possible with continuous monitoring of satellite imagery in the visible channels. In addition, infrared sensors provide information on cloud top height, important for locating rain bands. Microwave sensors can penetrate the storm to provide an indication of the interior core's relative warmth, closely related to the strength of the hurricane, and sea surface temperature to assess its development potential. *See* HURRICANE; TROPICAL METEOROLOGY.

Regional (meso) scale. Most significant weather events experienced by society—heavy rain or snow, severe thunderstorms, or high winds—are organized by systems that have horizontal dimensions of about

60 mi (100 km). These weather systems, known as mesoscale convective systems, often fall between stations of conventional observing networks. Hence, meteorologists might miss them were it not for satellite sensing.

Satellite data are critical in the period 2–3 hours before the development of regional weather phenomena. Experienced satellite meteorologists can recognize cloud patterns that mark areas of potential development (converging winds, rapidly growing clouds), and can issue forecasts or advisories long before the weather event strikes. Even in situations where no clouds exist initially, satellites can estimate the instability of air columns by measuring the thermal stability and vertical moisture distribution. Measurement of surface temperatures, as derived from infrared channels in clear-sky situations, significantly aids in accurately determining the surface air temperature, particularly when combined with a few observations from sites at ground level. This can help the forecaster locate potential severe storm regions up to 6 hours before development, such as the development of thunderstorm squall lines. Hours before storm development, upper-level systems can trigger mesoscale convective events, which are often revealed by patterns in water vapor imagery and can be tracked. By frequently monitoring the visual or infrared satellite imagery, the meteorologist can recognize lines of convergence (where air is flowing together) and note where low clouds seem to be merging and organizing. Later, as the forced upward motion allows the release of instability, vertical development of the cloud can be easily seen in visual imagery and confirmed by infrared. Finally, when the storm tops reach the tropopause, environmental winds and storm-generated pressure forces cause the tops to spread horizontally in cumulonimbus anvils. By noting the rapidity of development, cloud top heights, and anvil spreading rates, the forecaster can estimate the strength of the thunderstorm complex. At low levels, air cooled by precipitaion also spreads rapidly outward away from the squall line. Positions of associated arc cloud lines (named for their shape) are important for determining when a location will experience gusty winds and where new storms may develop. When severe storm systems remain stationary, flash floods can result. **Figure 6** shows the rapid development of a squall line.

Heavy snows or rains associated with large cyclones are often produced by mesoscale-sized thunderstorm systems that can be easily seen and tracked by (visual and infrared) satellite imagery. *See* MESOMETEOROLOGY; SQUALL LINE; STORM; THUNDERSTORM; TROPOPAUSE.

Small (cloud) scale. Although weather radar is the major source of data at this scale, individual clouds can be closely watched with respect to structure, height, and location (and hence movement) by frequent satellite scans. The geo satellites are set up to scan a particular location every few minutes. When the data are combined in computer-based weather workstations, which can rapidly process, display, and animate the imagery, near-real-time monitoring of individual storms is possible. Large thunderstorms with tornadic potential often exhibit a tilted or sheared appearance, with rapid growth on their rear flank, which often can be seen from satellites before tornado development. The rate of growth of single cells can be determined. Rain-cooled, low-level outflow from a single thunderstorm can often be seen. Intersections of these outflow air masses indicate a strong potential for new cell development. Satellite research has shown that at the anvil level, tracking

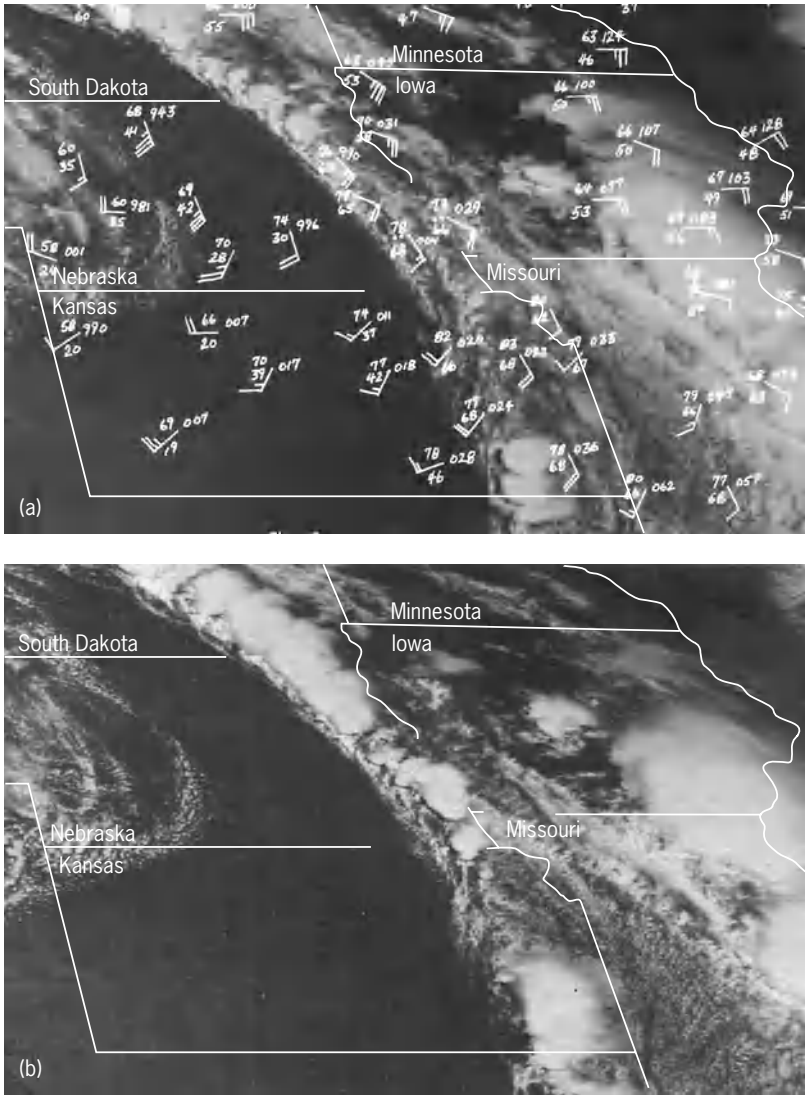

Fig. 6. Use of satellite imagery for severe thunderstorm forecasting. The two visual images were taken 1 hour apart at 18 GMT (13 EST) and 19 GMT (14 EST) on May 6, 1975. The region includes Nebraska, Kansas, Iowa, and Missouri. (*a*) Surface weather observations superimposed over the visual cloud image at 18 GMT. Surface parameters plotted are wind direction and speed [each large barb is 10 knots (5 m/s), half barb 5 knots (2.5 m/s)], temperature (°F) in upper left of a barb, dewpoints (°F) in lower left, surface pressure in upper right (for example, 029 is 1002.9 millibars or 100.29 kilopascals). Note newly developed thunderstorms in northeastern Nebraska. At least three cloud lines extend southward into a region of converging airflow. Dry air is in central Nebraska and western Kansas (shown by clear skies and low dewpoints). Forecasters issued a tornado watch for most of eastern Kansas and Nebraska based on conventional weather information and forecast models. Hardening of clumping of clouds on the line shows that this region is ready to explode. (*b*) At 19 GMT, massive storms developed over southeastern Nebraska. Soon after, the storms just northwest of the Missouri-Iowa-Nebraska juncture spawned a tornado, which devastated Omaha. In northern Nebraska, note how quickly the anvils (smooth-textured clouds) have expanded in 1 hour, indicating strong thunderstorm updrafts. Tornadoes occurred here as well. (*J. Purdom, National Earth Satellite Service, NOAA*)

cloud elements can lead to estimates of cloud top divergence (airflow spreading) and, with a few assumptions, an estimate of the vertical air speed or updraft of the storm. This is an important development since the magnitude of the updraft is related to hail production and tornadic potential. Tops that overshoot the anvil level can be seen on both visual (at low sun angles) and infrared imagery. Such occurrences indicate intense updrafts capable of forcing air well above the tropopause. Quite frequently, the anvil exhibits an enhanced V shape (open end downwind), which has been correlated with severe weather occurrences at ground level.

Rainfall estimates derived from satellite imagery, although crude, can guide the forecaster in issuing quantitative precipitation forecasts. Satellite data can be a valuable complement to weather radar by allowing simultaneous observation of the entire cloud and its associated features. Intense precipitation, like other thunderstorm-produced weather events, is related to updraft strength and hence rapid anvil spreading. The storm-relative location of the heaviest precipitation lies typically beneath strong gradients of the cloud top temperature on the upwind (usually western) edge of the infrared cloud image. Empirical quantitative methods under development show that it may be possible to estimate a rainfall rate. Research indicates that radiances measured by satellites may be used to classify cloud water drops by size. *See* HAIL; METEOROLOGY; PRECIPITATION (METEOROLOGY); TORNADO; WEATHER FORECASTING AND PREDICTION. Daniel L. Birkenheuer; John A. McGinley

Bibliography. D. Houghton (ed.), *Handbook of Applied Meteorology*, 1985; S. O. Kidder and T. H. Vonder Haar, *Satellite Meteorology: An Introduction*, 1995; U.S. Department of Commerce, Environmental Research Laboratories, *Thunderstorms: A Social, Scientific and Technological Documentary*, vol. 3, chap. 11, 1982; R. Weldon, *Cloud Pattern and Upper Level Winds*, NESDIS Satellite Training Course, NOAA, 1979.

Satellite navigation systems

Radionavigation systems using artificial satellites as sources of radio signals and position references. Development of satellite-based systems for global positioning and navigation began almost immediately after the launch of *Sputnik I* by the Soviet Union in 1957. The U.S. Navy Navigation Satellite System, better known as *Transit*, became operational in 1964. The Soviet Union responded with its own satellite navigation systems, *Parus* and *Tsikada*. All three systems were based on measurement of Doppler shift in signals received from a satellite, and were aimed mainly at offering navigational guidance to ships and submarines. *See* DOPPLER EFFECT.

Starting in the mid-1990s, the term satellite navigation system became synonymous with the U.S. Navstar Global Positioning System (GPS), but this would change. GPS is used every day by millions around the world for accurate estimates of position, velocity, and time. These estimates can be obtained instantaneously and continuously anywhere and by anyone equipped with an inexpensive, pocket-sized GPS receiver. The only requirement is that the receiver must be able to "see" at least four GPS satellites in the sky.

The success of GPS has led to efforts by other countries to develop systems modeled after it. The development of GLONASS began in the Soviet Union a few steps behind GPS. The system was acquired by Russia following the breakup of the Soviet Union, but it is not yet ready for operational use. GPS and GLONASS were designed during the Cold War and both started out as military systems which offered a subset of their signals for civil use. By contrast, Galileo, a European system on the drawing boards in 2005, is being designed primarily as a civil system. GPS, GLONASS, and Galileo are based on the same principles and exploit the same technologies. The generic name for each is Global Navigation Satellite System (GNSS).

Principles of satellite navigation. The basic idea behind position estimation with GPS (as well as GLONASS and Galileo) is very simple: One's position can be determined if one knows, or can measure, distances to objects whose position coordinates are known. Estimation of a position based on measurement of distances is referred to as trilateration. The technology for accurate measurement of long distances did not become available until the middle of the twentieth century, and before then the more common technique of positioning was known as triangulation. It was based on measurement of angles to objects whose positions were known.

The basic idea of trilateration is easily illustrated for positioning on a plane (**Fig. 1**). An observer measures his or her distance (or range) r_1 from station S_1 whose position coordinates are known. The observer must be located somewhere on a circle with a radius of r_1 centered at S_1. Range measurement r_2 from station S_2 defines another circle, and reduces the uncertainty in the observer's position to the two points (P and P') where the circles intersect. The observer may be able to reject one of the estimates on the basis of prior information. If not, a third measurement r_3 from station S_3 determines the position uniquely.

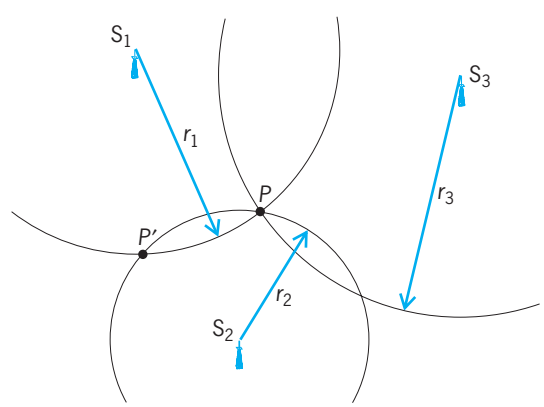

Fig. 1. Positioning by trilateration. The measured ranges to three position references, S_1, S_2, and S_3, determine the position of *P*.

This idea extends to positioning in three dimensions, with the circles replaced by spheres. The problem is formulated algebraically. Each distance r_1, r_2, and r_3 can be related to the known position coordinates of the three stations and the three unknown position coordinates (x, y, z) in the form of an equation. The three equations can be solved for the three unknowns. *See* ANALYTIC GEOMETRY.

Key enabling technologies. While the idea of trilateration is not new, a global navigation satellite system based on this idea would not have been feasible without several technologies which were developed or came to maturity in the second half of the twentieth century. The key technologies are stable space-based platforms in predictable orbits, extremely stable atomic clocks, spread spectrum signals, global coordinate frames, and microelectronics.

Recall that trilateration requires measurement of ranges to three position references (that is, objects at known positions). A satellite navigation system uses satellites as the position references. A satellite in a medium Earth orbit (for example, GPS) moves at a speed of about 4 km/s (2.5 mi/s), but its position at any instant can now be predicted 24–48 h earlier with an error no larger than a few meters.

The next two key technologies concern measurement of ranges in terms of transit times of signals from the satellites to a receiver. Roughly speaking, if meter-level positioning accuracy is required, measurement of ranges must be accurate to meter level and, therefore, measurement of transit times must be accurate to the nanosecond (10^{-9} s) level. (Radio waves travel at the speed of light, which is approximately 0.3 m/ns or 1 ft/ns.) Such precision is obtained by marking the times of transmission on the signals in accordance with ultrastable cesium and rubidium atomic clocks carried aboard the satellites. These clocks have stability of better than one part in 10^{13} and can be kept synchronized at the nanosecond level. The precise measurement of arrival times is made possible by broadcasting spread spectrum signals, which have wide bandwidths but allow each satellite to transmit its unique signal over a common frequency band. *See* ATOMIC CLOCK; SPREAD SPECTRUM COMMUNICATION.

A global positioning system requires a global coordinate frame in which to express the positions of its satellites and the users. Previously, the coordinate frames were defined at most on a country or regional basis. GPS provided the impetus for development by the U.S. Department of Defense of the World Geodetic System 1984 (WGS 84), which has now become a de facto international standard. *See* GEODESY.

Finally, the spectacular developments in microelectronics in the last quarter of the twentieth century have made possible the widespread use of GPS by producing receivers that are light, compact, and an order-of-magnitude less inexpensive than thought possible in 1980.

Pseudorange. In order to measure range to a satellite (that is, transit time of a signal from a satellite to a receiver), clearly, the clocks in the satellite and the receiver must be maintained in synchronism.

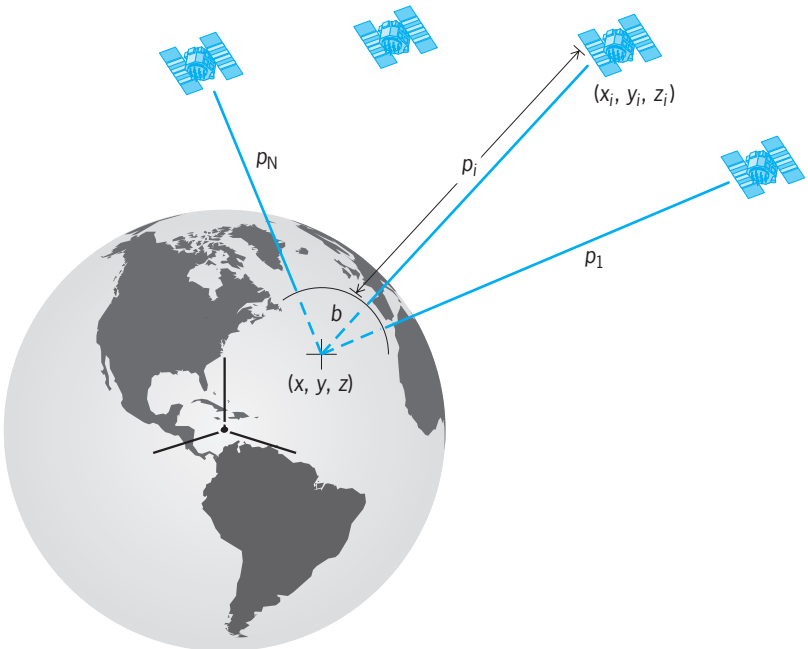

Fig. 2. Principles of satellite navigation. User-satellite range measurements based on transit times of signals are biased by a common amount and are called pseudoranges. For a constellation of N satellites, the pseudoranges p_i, $i = 1, 2, \ldots, N$, are given by $p_i = \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2} - b$, where the (x_i, y_i, z_i) are the known satellite positions. Pseudorange measurements are needed from at least four satellites ($N \geq 4$) to estimate the user position (x, y, z) and receiver clock bias b (in units of length).

Fortunately, this onerous requirement is easily sidestepped, allowing use of inexpensive quartz oscillators in GPS receivers. The bias in the receiver clock at the instant of the measurements affects the observed transit times for all satellites equally. The corresponding measured ranges are thus all too short, or all too long, by a common amount, and are called pseudoranges. The receiver clock bias, thus, becomes the fourth unknown to be estimated, in addition to the three coordinates of position. A user, therefore, needs a minimum of four satellites in view to set up four equations in the four unknowns: three coordinates of spatial position plus time (**Fig. 2**).

The pseudorange constitutes the main observable of a satellite navigation system. A second, more precise observable is the change in pseudorange, called accumulated delta range, used in applications requiring centimeter-level accuracy in relative positioning.

Range measurement errors and positioning accuracy. The quality of a position estimate obtained from a GNSS basically depends upon two factors: the accuracy of the pseudorange measurements and the spatial distribution of the satellites around the user. Multiple sources contribute to pseudorange errors. There are system errors: errors in the predicted satellite position and the satellite clock synchronization error. There are also errors in relating the measured signal transit time to pseudorange: While a radio signal travels in a straight line at the speed of light in space, the last 1000 km (600 mi) of its journey is through the variable environments of charged particles (ionosphere) and the neutral, gaseous atmosphere (troposphere), where the signal speed varies. Closer to

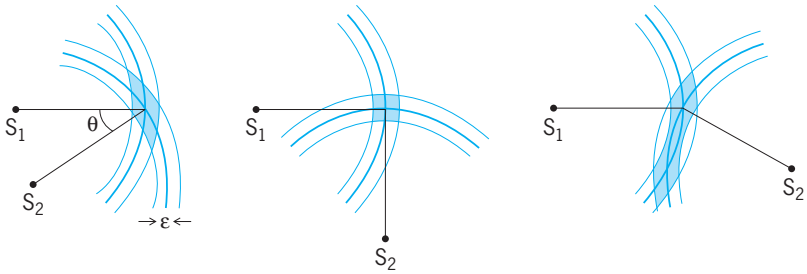

Fig. 3. Role of measurement geometry in positioning. The quality of a position estimate depends upon both the quality of the range measurements (ε) and their geometry (θ). Shaded regions represent the uncertainties in the position estimates.

the receiver, a signal may reach the antenna through multiple paths due to reflections from nearby surfaces and, finally, there is the inescapable thermal noise corrupting the signals. With careful control of the system errors and compensation for the atmospheric propagation conditions, pseudoranges can now be measured with an accuracy of several meters.

The spatial distribution (or geometry) of the satellites around the user also affects the quality of the position estimate. Satellite geometry is often characterized in terms of a parameter called dilution of precision (DOP). The more favorable the geometry, the lower the DOP. The basic idea can be illustrated simply for positioning on a plane (**Fig. 3**). For best three-dimensional positioning, a user would be "surrounded" by the satellites, several of which would be low in the sky and some high. With a 24-satellite constellation, poor satellite geometries are infrequent and change quickly.

GPS. GPS was developed in the 1970s and 1980s by the U.S. Department of Defense, and was declared ready for operational use in 1995. The system was intended primarily for the U.S. military; civil use was a secondary objective. GPS has proven to be a great success on both fronts. Its military value as a "force enhancer" surpassed expectations in the Gulf War (1990–1991), U.S interventions in Kosovo (1999), and Afghanistan (2001–2002), and the Iraq War (2003). And through its many civil applications in industry, commerce, and outdoor recreational activities, GPS is becoming essential to daily life as a new utility.

System architecture. GPS can be thought of in terms of its three elements, generally referred to as segments: space segment, control segment, and user segment. The space segment comprises the satellite constellation. The baseline constellation consists of 24 satellites in circular orbits with a radius of 26,560 km (16,504 mi) and orbital period of about 12 h. The satellites are arranged in six planes containing four satellites each and inclined at 55° relative to the equator (**Fig. 4**). With this satellite constellation, more than 99% of the GPS users worldwide with a clear view of the sky have at least four satellites in view at all times.

The first generation of GPS satellites, called Block I, was launched starting in 1978 to demonstrate the feasibility of the system. The operational system was subsequently populated with the second-generation satellites called Block II and Block IIA. The satellites launched in 2004–2005 to replace these operational satellites at the end of their service lives are called Block IIR.

The control segment encompasses the facilities and functions required for operating GPS. The facilities include crewless monitoring stations and uplink radar sites distributed around the globe, and the Master Control Station located at Schriever Air Force Base in Colorado. The main functions include satellite health monitoring, and prediction and upload of satellite orbits and clock parameters.

The user segment encompasses all aspects of design, development, production, and utilization of user equipment for the different applications.

Positioning services and signals. GPS offers two kinds of service: Standard Positioning Service (SPS) for open civil use and Precise Positioning Service (PPS) for users authorized by the Department of Defense. Access to the full capability of the system (that is, PPS) is restricted through cryptographic techniques. It rankled the civil user community that the SPS signal was degraded throughout the 1990s by introducing controlled errors to reduce its precision and limit the positioning accuracy to 60–100 m (200–330 ft). This feature, called selective availability (SA), was deactivated in 2000. Since then, the horizontal positioning error for most users has been less than 10 m (33 ft).

As of June 2005, each GPS satellite transmits continuously on two radio frequencies in the L-band (1 GHz–2 GHz), referred to as Link 1 (L1) and Link 2 (L2), with center frequencies of 1575.42 and 1227.60 MHz, respectively. Two signals are transmitted on L1, one for the civil users and the other for the Department of Defense authorized users. The lone signal on L2 is intended for the Department of Defense authorized users only (**Fig. 5**).

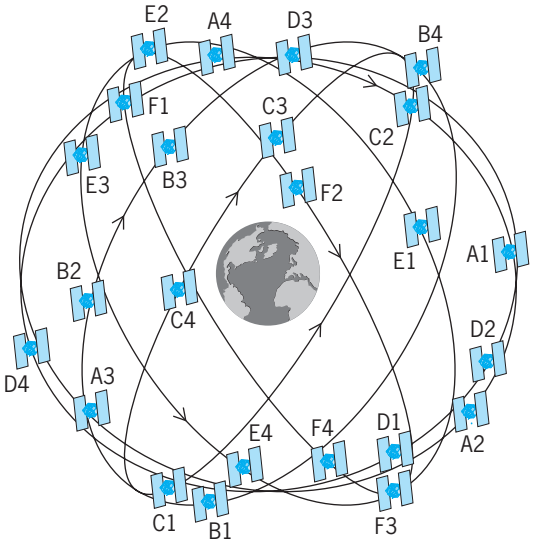

Fig. 4. GPS baseline constellation. Each satellite is identified by a two-character code: A letter identifies the orbital plane (A through F), and a number identifies the satellite slot number in the plane (1 through 4).

A signal consists of three components: the radio-frequency (RF) carrier, the ranging code, and navigation data (**Fig. 6**).

Ranging code. Associated with each Service (that is, SPS and PPS) is a family of binary codes called pseudorandom noise (PRN) codes. The PRN codes have special mathematical properties that allow all satellites to transmit in the same frequency band without interfering with each other. These codes also enable precise range measurements. The SPS codes are called coarse/acquisition codes (C/A codes) and the PPS codes are referred to as precision/encrypted codes, or P(Y) codes. Each satellite transmits a C/A code on L1 and P(Y) codes on both L1 and L2.

Each C/A code is a unique sequence of 1023 bits, called chips in this context, which is repeated each millisecond. The duration of each C/A-code chip is about 1 microsecond. Equivalently, the chip width or wavelength is about 300 m. The rate of the C/A-code chips, called the chipping rate, is 1.023 megachips per second (Mcps). A coarse/acquisition code is so-called because it is coarse as compared to a precision code, and it was originally intended to serve as a necessary steppingstone for acquisition of the much longer P(Y) code. Newer technology now allows direct acquisition of the P(Y) codes.

Navigation data. The navigation message, transmitted at a leisurely 50 bits per second (bps), consists of data on the satellite health status, ephemeris (satellite position and velocity), clock bias parameters, and an almanac giving reduced-precision ephemeris data for the entire constellation. It takes 12.5 min for the entire message to be received. The essential satellite ephemeris and clock parameters are repeated each 30 s.

Signal composition and modulation. The three components of a signal (carrier, code, and navigation data) are derived coherently from one of the atomic standards aboard the satellite. The code is combined with the binary navigation data using modulo-2 addition: If the code chip and the data bit are the same (both are 0s or both are 1s), the result is 0; and if they are different, the result is 1. The composite binary signal is then impressed upon the carrier in a process called modulation. The specific form of modulation used is called binary phase-shift keying (BPSK): a 0 bit leaves the carrier signal unchanged; and a 1 bit multiplies the carrier by -1, which is equivalent to shifting the phase of the sinusoidal signal by 180° (Fig. 5). *See* MODULATION; PHASE MODULATION.

The modulation of a carrier by a binary code spreads the signal energy, initially concentrated at a single frequency, over a wide frequency band: over 2 MHz for the C/A code and about 20 MHz for the P(Y) code, centered at the carrier frequency (**Fig. 7**). While the signal power is unchanged, this step reduces the power spectral density below that for the background RF radiation. Such signals, referred to as spread spectrum signals, have many properties which make them attractive for use in communication and navigation. The signal energy can be "despread" in a receiver if the code is known. In principle, keeping a code secret limits access to the signal.

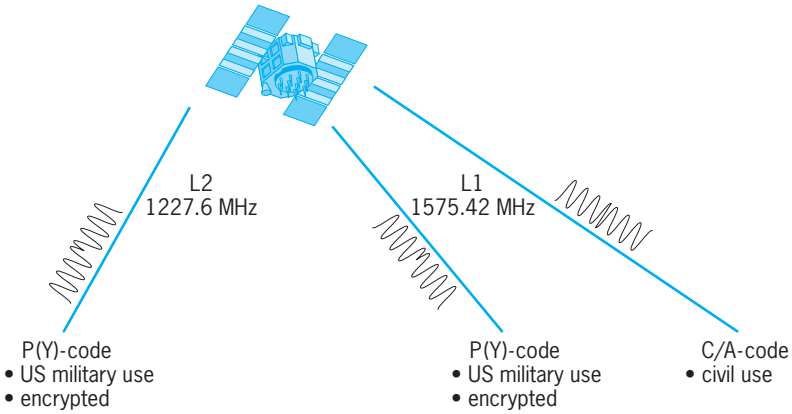

Fig. 5. GPS signals in June 2005. Each satellite transmits three signals: one for open civil use on L1, and one each on L1 and L2 for users authorized by the U.S. Department of Defense. Satellites launched starting in July 2005 transmit additional signals.

The GPS signals are extremely weak. The signal power reaching a terrestrial antenna is only about 10^{-16} W (-160 dBW). The ambient radio noise power associated with the signal is 100 to 1000 times (20–30 dB) stronger. Extraction of a signal buried so deep in noise is made possible by the knowledge of the spread spectrum ranging code. But there is little margin and attenuation (by foliage, for example) can be great enough to make a signal irretrievable.

Receiver architecture. Given a recent almanac and a rough idea of user location and time, the receiver determines which satellites are in view. Given a satellite ID, the receiver knows the structure of the C/A code transmitted by it. The receiver "tunes" to the C/A code to acquire the signal, and from then on tracks changes in it continuously.

To acquire a signal, the receiver generates a replica of the known C/A code, and attempts to align it with the incoming code by sliding the replica in time and computing the correlation between the two. The correlation function exhibits a sharp peak when the code replica is aligned with the code received from the satellite. The uncertainty in matching the replica with the incoming code is limited to only 1023 code chips, and the process of aligning them is generally quick. Direct acquisition of a P(Y) code is difficult by design due to the length of the code.

The measurement of transit time for a signal modulated by a C/A code is conceptually quite simple: The time shift required to align the receiver-generated code replica and the signal received from the satellite is the apparent transit time of the signal (modulo 1 millisecond) (Fig. 6).

The Doppler shift, caused by the relative motion of a satellite and the user, is a measure of the projection of the relative velocity on the line of sight, and can be converted into pseudorange rate. Given the pseudorange rates corresponding to four satellites, and the satellite velocity (obtained from the navigation message), a user can compute his or her velocity. A GPS receiver does this automatically, continuously, and virtually instantaneously.

Differential GPS (DGPS). The quality of the GPS position estimates can be improved by reducing

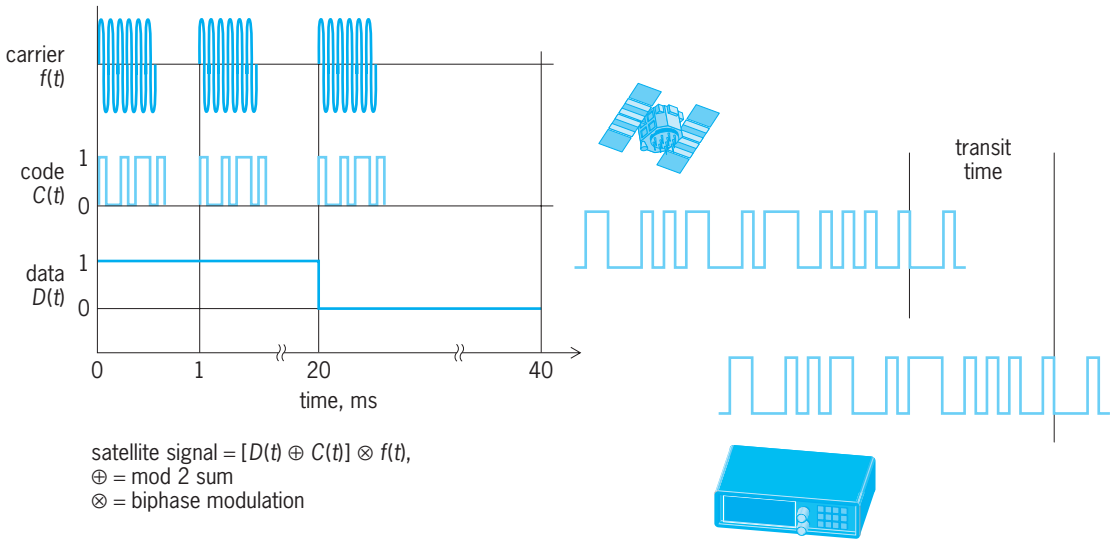

Fig. 6. Structure of the signal available for civil use and estimation of transit time from the satellite to the user.

pseudorange measurement errors. The errors associated with satellite ephemeris and clock, and atmospheric propagation delays are similar for users located near each other, and change slowly in time. These errors can be estimated from measurements with a receiver at a known location, and can be made available as differential corrections to the GPS users in the area, allowing them to mitigate errors in their measurements (**Fig. 8**).

The U.S. Coast Guard broadcasts differential corrections on marine radiobeacon frequencies (285–325 kHz) from about 50 broadcast sites in U.S. coastal areas and around the Great Lakes. This service, called Maritime DGPS, has been operational since 1999 and provides horizontal positioning accuracy of 1–3 m (3–10 ft) at a distance of 100 km (60 mi), or more, from a reference station. A number of countries have implemented systems compliant with the Maritime DGPS standards to enhance safety in their waterways.

An alternative to local-area DGPS is a centralized system which provides differential corrections usable over continent-wide areas or oceans. Such systems, referred to as wide-area DGPS, collect measurements from a network of reference stations covering the area of interest and process them to decompose the measurement errors into their constituent parts: satellite orbit, satellite clock, and ionospheric delay. The differential corrections are then transmitted in a form so that each user can determine his or her corrections on the basis of his or her (rough) position estimate.

The U.S. Wide Area Augmentation System (WAAS), operational since 2003, broadcasts corrections applicable to the United States from geostationary satellites. These corrections are coded into the navigation messages modulated on GPS-like signals on L1 which can be received with a GPS receiver with some software modifications. The Canada-Wide Differential GPS (CDGPS), which became operational in 2004, also uses geostationary satellites to broadcast corrections which are available to all who can invest in a separate radio receiver. Global Differential GPS (GDGPS), operated by the National Aeronautics and Space Administration, transmits corrections on the Internet (to authorized users). A number of commercial services also provide differential corrections via communication satellites and FM subcarrier, and

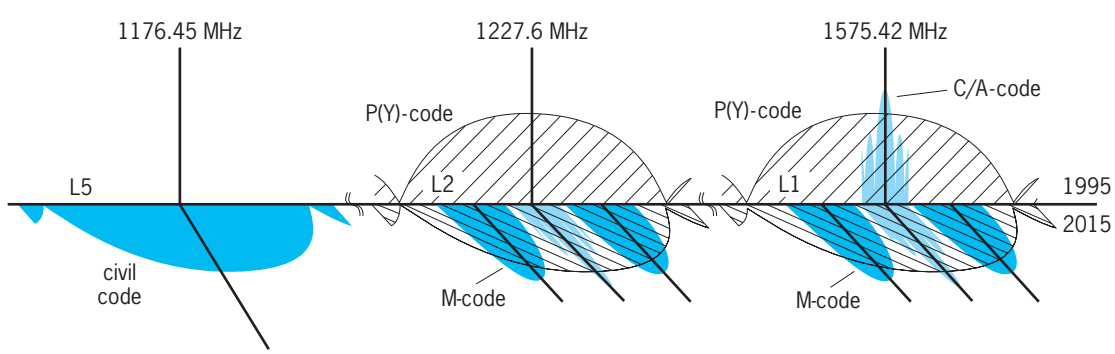

Fig. 7. Power spectral densities of the current and future GPS signals. Each satellite transmits three signals: one for civil users and two for Department of Defense–authorized users. Satellites launched starting in 2005 transmitted seven signals: three for the civil users and four for the Department of Defense-authorized users.

use of such services is common in offshore oil exploration and fleet management.

Modernization. GPS modernization was launched in the late 1990s, soon after the system became operational and its full potential for military and civil applications became clear. The civil users lobbied successfully for expanded benefits and secured two new signals: a signal on L2, and another at L5 (1176.45 MHz) to benefit civil aviation and other applications with safety-of-life considerations.

With its primary mission as a warfare support system, new measures were also deemed necessary to protect the signals for military use from interference, to prevent the use of any GPS signals by adversaries during hostilities, and to preserve peaceful civil use outside the area of military operations. The Navigation Warfare (Navwar) program of the Department of Defense addresses these requirements as a part of modernization.

There are plans to modify the design of satellites scheduled for launch during and after 2005 to include the new signals, and to design a whole new GPS of the future, called GPS III. The number of ranging signals for civil users will increase from one to three. The military will get a pair of new, Navwar-compatible signals called M-codes (Fig. 7). There will be equally important changes behind the scenes in the Control Segment to operate the modernized system. It will be 2015 or later before a full constellation of the expanded-capability satellites is deployed.

Systems under development. Besides GPS, other satellite navigation systems are under development.

GLONASS. While GPS was under development, the Soviet Union undertook to develop a similar system called GLONASS. Like GPS, GLONASS was designed primarily for the military, and a subset of its signals was offered for civil use in the late 1980s. Since the dissolution of the Soviet Union, the Russian Federation has assumed responsibility for GLONASS.

GLONASS, like GPS, fielded a 24-satellite constellation in medium earth orbits but, unlike GPS, placed the satellites in three planes with an inclination of 65°. Each satellite transmits in two frequency bands in the L-band and offers GPS-like services for the civil and military users.

The civil user community was excited at the prospect of having access to two autonomous systems as GLONASS built up to a full constellation in 1996. Unfortunately, the system has since declined as it has suffered from a lack of resources in a changed political and economic climate. Between 2000 and 2004, the Russians managed only one launch per year, each placing three new satellites in orbit, and the constellation strength has stayed mostly below 12. While a couple of GPS+GLNASS receiver models have been available since the mid-1990s, the user community and receiver manufacturers appear uncertain of the viability of GLONASS.

Galileo. A more promising GNSS is the European system, Galileo, planned as "an open, global system, fully compatible with GPS, but independent from it." Galileo was conceived as a joint public-private enterprise under civilian control, to be financed and managed by the European Commission, the European Space Agency, and industry. The deployment phase was scheduled to begin in 2006 and an operational system is planned for 2008.

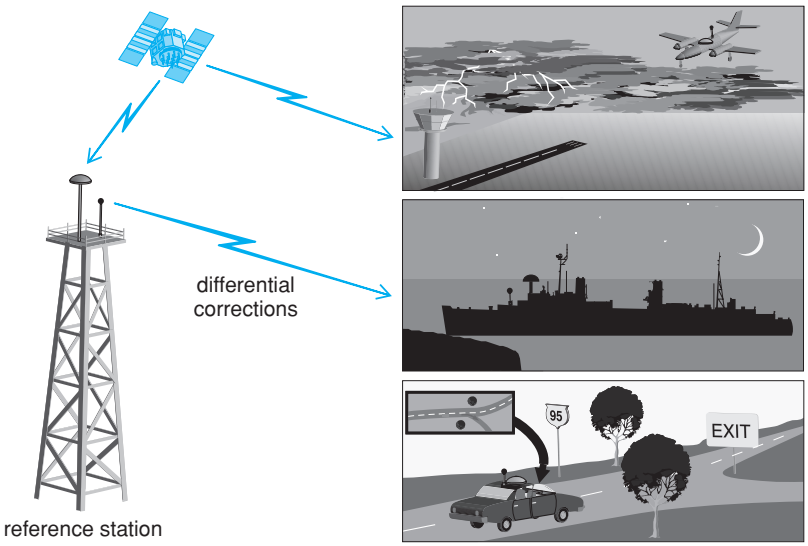

Fig. 8. Local-area differential GPS. A reference receiver at a known location measures pseudoranges, determines errors therein, and transmits corrections on a radio link to the users in the area. In this model, each user is equipped with a saucer-shaped antenna to receive the GPS signals and a bulb-shaped antenna to transmit or receive the differential corrections on a radio link. The DGPS users in the picture are (top) an aircraft in clouds preparing to land, (middle) a ship entering a harbor at night, and (bottom) a motorist on a highway in an unfamiliar city.

Galileo would be a second-generation GNSS, following the path of GPS and benefiting from technological advances since the mid-1970s when GPS was designed. Galileo would offer an Open Service without any user fees, like the GPS Standard Positioning Service, and three different, value-added services for which a user would pay.

Beidou. With the launch of two geostationary satellites in 2000 and another in 2003, China has apparently completed a regional satellite navigation system called Beidou (Big Dipper). This system provides positioning capability over China and neighboring areas. Unlike GPS, GLONASS, and Galileo, which are passive or listen-only systems, Beidou interacts with its users: The system determines the position estimate and transmits it to each user. The positioning accuracy is reported to be dozens of meters.

Compatibility and interoperability of systems. GPS, GLONASS, and Galileo are autonomous systems. From a user's perspective, at a minimum, these systems should be compatible or noninterfering: None should degrade the performance of the others in a significant way. Ideally, the systems would be interoperable, allowing a user to combine measurements from the three systems so that the performance is as least as good as the best of the three at that place and time. A user could then choose between the three independent, redundant services, or one robust integrated service.

The three systems use different time scales and coordinate frames. The offsets among the time scales can be estimated and accounted for at the user level. The differences between the coordinate frames,

however, must be characterized at system level. A user's position as determined from GPS and GLONASS was found to differ systematically by several meters because of the differences in their coordinate frames: GPS expresses the user position in WGS 84, and GLONASS in an independently realized global coordinate frame called PZ-90. A transformation between the two coordinate frames was not determined until the mid-1990s. The difference between WGS 84 and the independently realized Galileo Terrestrial Reference Frame (GTRF) is expected to be at centimeter level, insignificant for most users.

Applications. The civil applications of GPS grew rapidly in the 1990s and new applications continue to appear. The number of GPS receivers manufactured each year exceeds one million. The emergence of GLONASS and Galileo is bound to accelerate the pace of development of new applications.

While its positioning capability receives most attention, GPS is also a precise timing service. For example, before GPS, it would have been a challenge to synchronize clocks at multiple sites to conduct simultaneous observations of events (perhaps celestial) at distant locations. An inexpensive GPS receiver can now serve as a precise clock, keeping time with an accuracy of better than 0.1 microseconds relative to Coordinated Universal Time (UTC). Timing is critical in today's telecommunication systems, and GPS is being used increasingly to synchronize the elements of networks. *See* ATOMIC TIME; TIME.

The applications of GPS for navigation and fleet tracking in land transportation, aviation, and maritime commerce continue to grow. The convergence of GPS and personal computer technology has made it possible to collect vast amounts of positional data and organize them into geographic information systems (GIS), showing, for example, the location and status of every utility pole or access-hole cover in a town, or a map showing concentrations of toxic or radioactive wastes at a site. The centimeter-level positioning capability of GPS has revolutionized surveying and mapping, and has created novel applications in construction, mining, agriculture, and the earth sciences. *See* GEOGRAPHIC INFORMATION SYSTEMS.

Millions of bikers, runners, hikers, sailors, and fishermen have come to rely on GPS for their position and velocity. GPS receivers are now routine in cell phones and personal digital assistants. Indoor GPS positioning in conjunction with a cell phone, now under development, would lead to a whole new class of applications.

One's position expressed in a well-defined coordinate frame is a valuable piece of information. Determining position indoor and outside would soon become as easy as determining precise time is today, and position would come to occupy the same important place in daily life as time does today.

Acknowledgment. Figures 1–8 are reproduced by permission of Ganga-Jamuna Press, Lincoln, MA.

Pratap Misra

Bibliography. T. Logsdon, *Understanding the Navstar: GPS, GIS, and IVHS*, 2d ed., Van Nostrand Reinhold, New York, 1995; P. Misra and P. Enge *Global Positioning System: Signals, Measurements, and Performance*, 2d ed., Ganga-Jamuna Press, Lincoln, Mass., 2005; B. W. Parkinson et al. (eds.), *Global Positioning System: Theory and Applications*, vols. I and II, American Institute of Aeronautics and Astronautics, Washington, DC, 1996; D. Wells (ed.), *Guide to GPS Positioning*, Canadian GPS Associates, 1986.

Satellite radio

Direct audio broadcasting from satellites to mobile and stationary receivers. This concept originated in the late 1980s as an outgrowth of direct television broadcasting from satellites to home receivers. There are currently two satellite radio systems (also known as satellite digital audio radio service) operating in the continental United States: one is providing service to users in Japan and Korea, and another is in advanced planning to serve Europe. *See* SATELLITE TELEVISION BROADCASTING.

The useful range of frequencies for radio broadcast transmission from satellites is 1–3 GHz, the most sought-after frequency band by all transmission services due to its superior propagation characteristics. Above 3 GHz, propagation losses and rain (precipitation) attenuation pose difficulties, and below 1 GHz, satellite antenna size and galactic noise (radio-frequency noise that originates outside the solar system) pose difficulties. The Federal Communications Commission (FCC) has allocated a frequency band of 2320– 2345 MHz for satellite radio, which has been coordinated internationally. This frequency band cannot be used in the United States for any other service because the sensitivity, mobility, and antenna omnidirectionality of the ground receivers require an interference-free environment. Of the available 25 MHz, a satellite radio system requires between 10 and 15 MHz bandwidth to support a reasonable number of audio programs. As a result, the FCC auctioned two 12.5-MHz-wide bands for satellite radio service, and granted licenses for two such systems in October 1997 to Sirius Satellite Radio Inc. and to XM Satellite Radio, Inc. *See* RADIO SPECTRUM ALLOCATIONS; RADIO-WAVE PROPAGATION.

Technology. The technical challenge of satellite radio is in keeping service outages very infrequent and short. Service outages are caused by blockages, such as a building, structure, or mountain that obstructs the line-of-sight between a satellite and receiving antenna; by multipath, where received reflected signal components have sufficient amplitude and phase offset to corrupt (distort) the unreflected received signal; by foliage attenuation; and by interference, particularly the out-of-band emissions from transmitters in adjacent frequency bands.

The quality goal is to achieve continuity of service (that is, no outages) better than 99.9% of the time, excluding the time when a vehicle is indoors since the satellite transmissions do not substantially penetrate building walls. High service availability is accomplished by using high-power transmission satellites and by providing an unprecedented degree of transmission diversity in the system design.

Spatial diversity. This is accomplished by simultaneously broadcasting the same transmission from two satellites located in space so that their azimuth and elevation angles from the radio receiver are significantly different. Since the radio is designed to receive both transmissions, if the path from the radio to one of the satellites is blocked by a building or mountain, there will be no outage if the path to the other satellite is unobstructed.

Time diversity. This is accomplished by providing a 4–5-s offset either between transmissions from two satellites by a time delay or within each satellite transmission by a long interleaving span. If one or both transmissions are lost, the receiver continues to play using the portion of the transmission stored in a 4–5-s buffer memory, and no outage occurs if a transmission path is restored before the end of the stored signal. This technique is particularly effective in combating outages from roadway overpasses.

Frequency diversity. In this technique, two satellites transmit at different radio frequencies, several megahertz apart. Frequency diversity is effective in minimizing a form of multipath called frequency selective fading.

Modal diversity. Despite the above diversity techniques, it is impossible to provide high-quality service in dense urban areas since the blockage from buildings is overwhelming. However, this can be overcome by providing terrestrial repeaters. High-power broadcast transmitters located in dense urban areas are fed the same signal as transmitted from the satellite and then retransmit this signal in a frequency band between the two satellites' transmissions. A receiver with three channels, one for the terrestrial repeater and two for the diversity satellites, allows virtually uninterupted service and seamless transition between satellite and terrestrial transmissions. The three signals, which are simultaneously received when available, are added together in a maximal ratio (diversity) combiner.

In addition to diversity, satellite elevation angle and the use of equalizers are important techniques for attaining high service availability.

Elevation angle. The elevation angle from the satellite to the radio is extremely important in mitigating blockage, multipath, and foliage attenuation. For geostationary orbits, the elevation angle is reduced as the location of the radio moves further north or south of the Equator. An alternative is to use a constellation of satellites in highly inclined, elliptical geosynchronous orbits so that high elevation angles can be provided to radios located at latitudes above 30°N (**Fig. 1**). In the Sirius geosynchronous satellite constellation, three satellites follow each other at 8-h intervals and spend approximately 16 h north of the Equator and 8 h below. *See* COMMUNICATIONS SATELLITE.

The importance of elevation angle with regard to foliage attenuation is shown in **Fig. 2**. The difference between a 30° and a 60° elevation angle for a 99% service availability is over 13 dB, which means a satellite must transmit more than 20 times additional power at 30° than at 60° for equal service availability through foliage. **Figure 3** shows the importance of higher elevation angles for blockage reduction, where the blockage distance with a 60° elevation angle is only one-third of that with a 30° elevation angle. A high elevation angle also alleviates blockage outages from passing trucks.

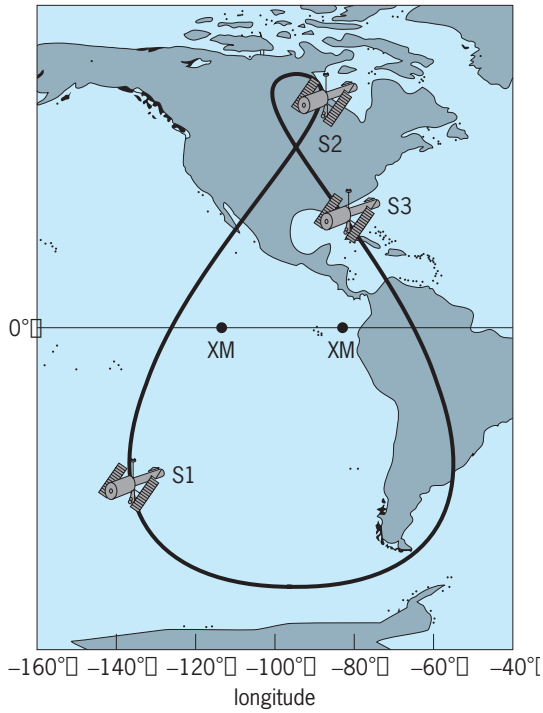

Fig. 1. Orbital locations of the XM geostationary satellites located at longitudes 85°W and 115°W, and the satellite ground track of the Sirius (S) geosynchronous satellite constellation.

Equalization. A technique for reducing outage from multipath is to have an equalizer in the receiver that compensates for a significant portion of the signal distortion in amplitude and phase caused by the multipath components. This is particularly effective in satellite radio for handling short multipath propagations that are caused, for example, by reflections from large, overhead roadway signs. *See* EQUALIZER.

System design. For satellite radio, each satellite system generally provides 65 music channels and 55 voice channels. This mix can be changed and, if

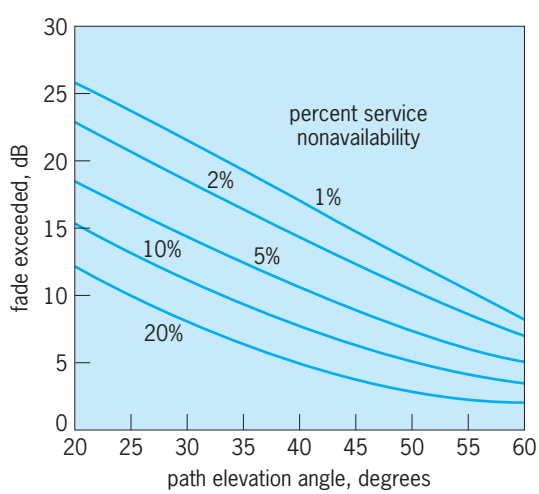

Fig. 2. Foliage attenuation.

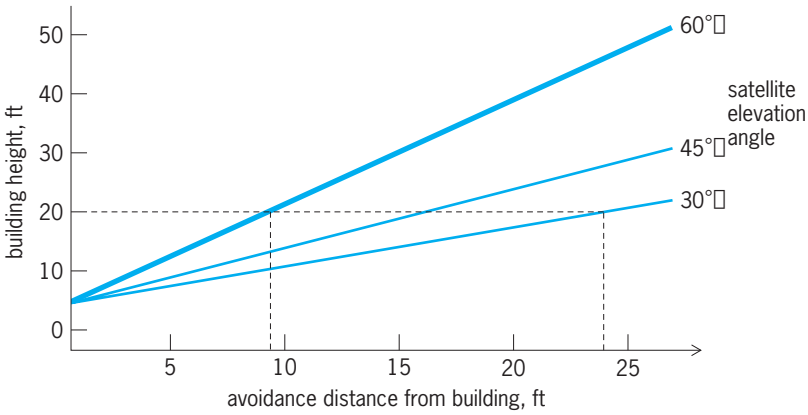

Fig. 3. Blockage avoidance. A vehicle (with antenna height 5 ft or 1.5 m) must be a minimal distance from a building to avoid blockage in the worse case. 1 ft = 0.3 m.

a market develops, one-way data channels can be substituted for one or more audio channels. The major elements of the system are the satellites, radios, and the broadcast studio/uplink (**Fig. 4**). The telemetry, tracking, and command (TT&C) subsystems that take care of the satellites are standard and use no new technology. Also, the terrestrial repeater network uses standard digital audio broadcasting (DAB) techniques, which include coded or orthogonal frequency division modulation (COFDM). *See* MODULATION.

Satellites. The satellites used are modifications of existing high-power communications satellites. The important technical modifications are conversion to 2.3-GHz transmitter operation, and concentration of the transmitter power to achieve a few megawatts (63–65 dBW) of effective isotropic radiated power (EIRP) over the relatively narrow transmission bandwidth (2.5–4.2 MHz). The high transmitter power is achieved by paralleling many traveling-wave-tube amplifiers.

Radios. Radios have been produced for aftermarket installation in motor vehicles, boats, planes, and homes, as well as for installation during the assembly of motor vehicles. The receiver is a three-channel type with intelligent combining of the two satellite and terrestrial transmissions, when available. An application-specific integrated circuit (ASIC) chipset contains most of the required electronic circuits, excluding the antenna (omnidirectional with 3–5 dBi gain), radio-frequency amplifiers (typical system noise temperature of 160 K or −172°F), bulk memories, and power supply. The received digital signal is encrypted at the broadcast studio to prevent theft of service. In addition to having a decryptor in the receiver, each receiver has an embedded unique digital address. This allows the broadcast studio, through a service channel, to send messages and commands, which are recognized solely by the addressed receiver. Among other functions, this is

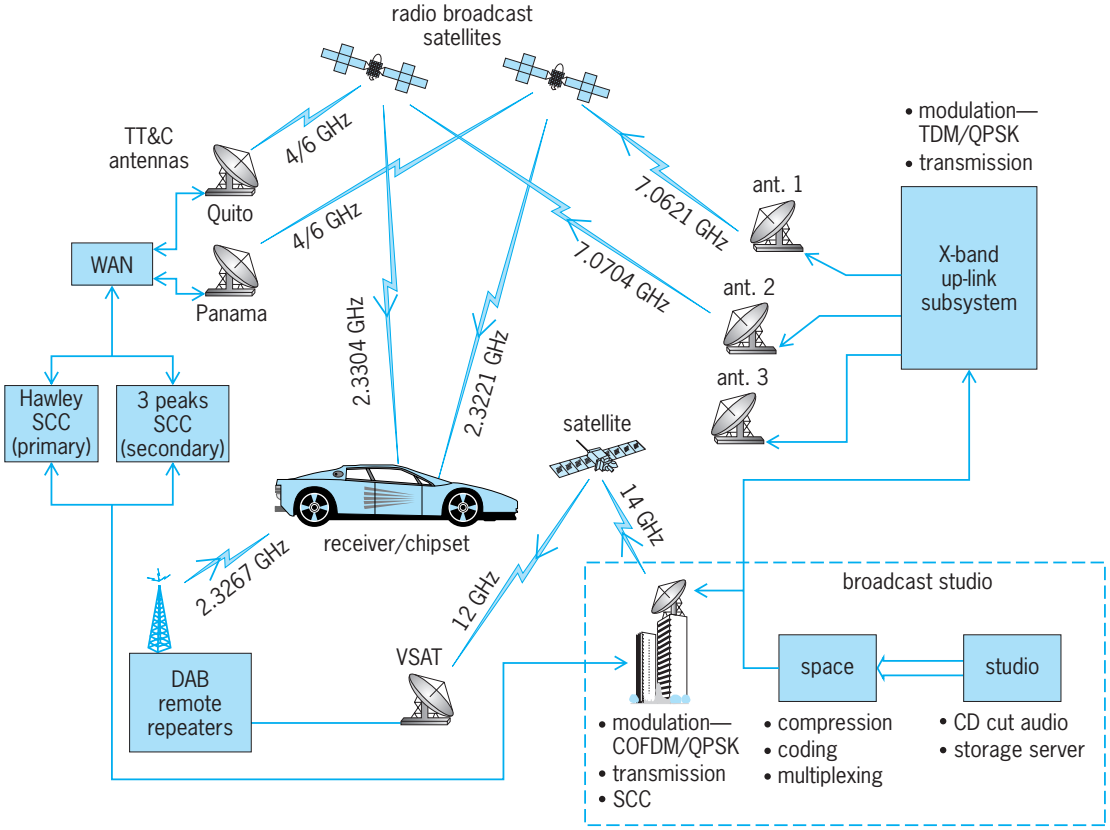

Fig. 4. Typical system for providing satellite radio service. TT&C = telemetry, tracking, and command; DAB = digital audio broadcasting; VSAT = very small aperture terminal; SCC = satellite control center; COFDM = coded othogonal frequency division modulation; QPSK = quadrature phase shift keying; TDM = time division modulation.

used to turn subscriber service on and off. *See* INTEGRATED CIRCUITS; RADIO RECEIVER.

Broadcast studio/uplink. These facilities include those used for generating the audio programming, as well as the resulting program signals' multiplexing, coding, modulation, and transmission (Fig. 4). The programming facilities are audio studios, using both recorded and live performances, which output in a digital form as a music or a voice channel. The channels are compressed, coded (the inner code is convolutional, the outer code is a Reed Solomon block code), multiplexed together using time division into one digital stream, modulated using quadrature phase shift keying, split into two identical signals, one of which is then time-delayed, translated to 7 GHz, and transmitted to the two active satellites. Other than the audio compression, the facilities and techniques are those used by newer communications satellite, terrestrial broadcast, and satellite television systems.

Audio compression. An audio compression technique, called perceptual audio coding (PAC), is the first extensive use of a technology developed over the past decade without which satellite radio would not be practical or economic. It allows compact disk recordings (whose digital output is 1.2 mb/s) to be compressed to approximately 50 kb/s at the transmit studio and expanded in the radio receiver in a manner that the listener hears a music quality close to the uncompressed music. This is accomplished by compression/expansion algorithms which take advantage of the characteristics of the human ear (for example, its inability to hear certain sounds and artifacts). Similarly, a voice channel is compressed to approximately 25 kb/s. *See* DATA COMPRESSION.

Outlook. Satellite radio is fast developing with regard to commercial success, which requires a few million subscribers for each system. During mid-2005, the two United States satellite radio systems had 7 million subscribers combined. New services and uses will unquestionably develop over the next several years. The possibilities of providing data, multimedia, and personalized services are under investigation. Navigation services were scheduled to be introduced in 2005 and television to the rear seats of sports utility vehicles in 2006. The Canadian government authorized (with some conditions on programming content) the extension of the Sirius and XM satellite radio service to subscribers in that country during 2005. *See* RADIO; RADIO BROADCASTING.

Robert D. Briskman

Bibliography. R. D. Briskman and R. J. Prevaux, S-DARS broadcast from inclined, elliptical orbits, *52d International Astronautical Congress*, IAF-01M.5.05, October 1–15, 2001; R. D. Briskman and S. P. Sharma, DARS satellite constellation performance, *AIAA 20th International Communications Satellite Systems Conference*, AIAA-2002-1967, May 12–15, 2002; Efficient high latitude service arc satellite mobile broadcasting systems, U.S. Patent and Trademark Office, 6,223,019, April 24, 2001; R. A. Michalski, An overview of the XM satellite radio system, *AIAA 20th International Communications Satellite Systems Conference*, AIAA-2002-1844, May 12–15, 2002; Mobile radio receivers using time diversity to avoid service outages in multichannel broadcast systems, U.S. Patent and Trademark Office, 5,592,471, January 7, 1997; Radio frequency broadcasting systems and methods using two low-cost geosynchronous satellites, U.S. Patent and Trademark Office, 5,319,673, June 7, 1994.

Satellite television broadcasting

The transmission of television and other program material via satellite directly to individual homes and businesses. Direct broadcasting satellite (DBS) systems, or direct-to-home (DTH) satellite systems, are operational in many nations and regions around the world. The highly successful "small dish" DTH systems are characterized by all-digital transmission with the links to the customer dishes at frequencies above 10 GHz using radio spectrum sometimes referred to as Ku-band. Two major small-dish systems in the United States, DIRECTV and Dish, have been operational since the mid-1990s and serve over 20% of all American households.

Direct television broadcasting is the most widely used consumer application for communications satellites. The related application, direct audio broadcasting, has also shown significant growth since its launch in the United States in 2001. Other "direct" consumer applications include voice, data, and Internet access, but these services have not yet achieved mass market success. *See* SATELLITE RADIO.

Technology. DBS systems use a satellite in geostationary orbit to receive television signals sent up from a broadcasting center, amplify them, and transmit them down to the customer receiving antennas called dishes. The satellite also shifts the signal frequency, so that, for example, a signal sent up to the satellite in the 17.3–17.8-GHz uplink band is transmitted back down in the 12.2–12.7-GHz downlink band. The downlink signal is picked up by a receive antenna located atop an individual home or office; these antennas are usually in the form of a small parabolic dish. The receive antenna is permanently pointed at the satellite, which is at a fixed point in the sky, in an equatorial geosynchronous or geostationary orbit, a circular orbit in the plane of the equator with a period synchronized to the 24-h rotational period of the Earth. *See* ANTENNA (ELECTROMAGNETISM).

Receiver technology. The electronics for handling microwave frequencies is expensive, so the signal from the dish antenna is first passed to a downconverter, mounted outdoors on the antenna, which shifts it to (typically) the 0.95–1.45-GHz band. This signal is then conducted by cable to the receiver near the television set. The receiver contains the channel selector or tuner, a demodulator, and a decryptor (descrambler) as well as a video-audio decoder to permit the user to view authorized channels. With the broadcasting industry's evolution to high-definition television (HDTV), the more advanced indoor

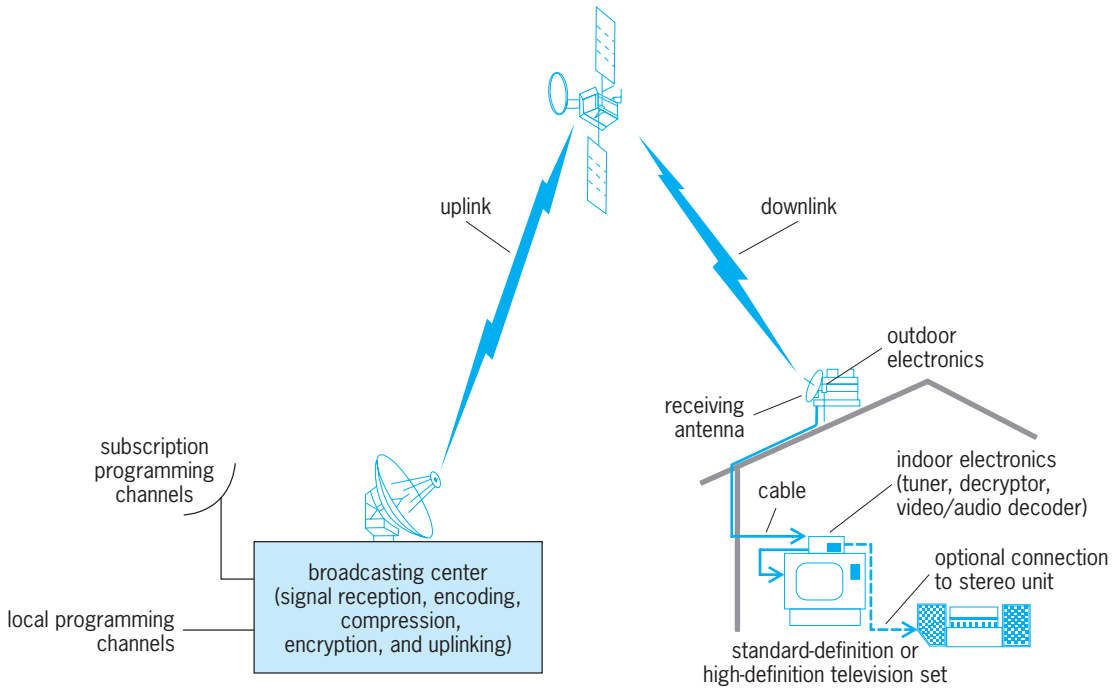

Direct broadcasting television satellite system.

receivers can now output either standard-definition or high-definition television. Additionally, the receivers have the capability to output monaural, stereo, or multichannel audio to a television set, home audio system, or home theater system (see **illustration**). *See* TELEVISION RECEIVER.

Spectrum allocations. Television broadcasting by satellite may use radio-frequency spectrum licensed as either the broadcasting satellite service (BSS) or the fixed satellite service (FSS). In most cases the system design differences due to the use of these alternative bands are due to international and domestic regulation and not technical or environmental differences. The two major U.S. DBS systems have used BSS spectrum with a downlink frequency of 12.2–12.7 GHz. Other major systems in the Western Hemisphere use FSS spectrum with a downlink frequency around 11 GHz. The term DBS is generally used to denote systems using the BSS band. The term DTH is more general and applies to television delivery systems employing any band. *See* RADIO SPECTRUM ALLOCATION.

Satellite technology. A direct broadcasting satellite typically contains 32 or more transponders, each with a radio-frequency power output in the range 120–240 W. Each transponder acts as a separate amplifier, frequency translator, and transmitter chain for a signal uplinked from the broadcasting facility. The transmitter of each transponder drives one of the satellite's downlink antennas. Each downlink antenna creates either a broad national beam or spot beams. In the major U.S. DBS systems, each service provider uses multiple satellites, with a given satellite broadcasting on a national beam or on one or more "local market" beams. Each local spot beam delivers television signals from the very high frequency (VHF) and ultrahigh frequency (UHF) stations of a particular urban area back into that geographic area.

As each satellite orbits the Earth, its communications antennas are maintained precisely pointed toward the target service areas. Large panels of photovoltaic cells are continuously oriented toward the Sun to provide power for the signal transmitters. Each satellite also carries rechargeable batteries for continuous operation during occasional "eclipse" periods when the satellite is in the Earth's shadow.

Digital signal technology. All recently deployed DBS and DTH systems use digital signals. Typically in these systems, a single 24-MHz-wide satellite transponder can carry an error-corrected digital signal of 30 megabits per second or greater. A wide variety of communications services can be converted to digital form and carried as part of this digital signal, including conventional television, high-definition television, multichannel audio, information services (such as financial and news retrieval services), and other forms of digital data.

Modern digital signal compression technology greatly increases the capacity of a satellite transponder and hence the entire system. It is possible to compress up to perhaps 10 or more standard-definition television signals into the bandwidth of a single DBS transponder. The technology used for this signal compression, from the Motion Picture Experts Group (MPEG) standards body, is very similar to that used in production of DVD optical disks. In a satellite transponder the bandwidth is fixed, but multiple signals can be "statistically multiplexed" into this bandwidth in real time. In a DVD the total storage capacity is fixed, but the processing of the movie can

be modified to assure that the entire capacity is used, but not exceeded, to achieve the target playback quality. *See* COMPACT DISK; DATA COMPRESSION; OPTICAL RECORDING.

Signal attenuation by rain. DBS systems, like all satellite systems operating above 10 GHz, are subject to attenuation of their signals by rain. Rainfall absorbs some of the signal power transmitted by the satellite. For a small percentage of the time, the rainfall may be so heavy in a particular geographical area that the remaining power is not sufficient for reception of the signal. The amount of rain-induced attenuation can be predicted, and the system is usually designed to work for the highest percentage of the time that is economically practical. The combination of satellite power and receive-dish antenna size is chosen to enable reception for all but the heaviest rainfall periods of the year at any particular location. The DBS customer can further reduce this expected outage period by purchasing a slightly larger dish antenna.

Uplink centers. Each DBS digital uplink center accepts multiple video and audio channels, processes and compresses the signals, and transmits them to the digital broadcasting satellites for relay to individual home-receive systems. Typically the subscription programming channels arrive at the broadcasting center via satellite systems provided by other firms and intended for national channel distribution to satellite and cable companies. The local channels are typically delivered to an uplink center using dedicated digital terrestrial facilities. Each incoming channel is converted to a common digital form and is then compressed to reduce the number of bits per second required for its transmission. These compressed channels are then individually encrypted to prevent unauthorized reception and combined into composite digital bit streams by digital multiplexers. Finally, each bit stream is protected by forward error control (FEC) coding and modulated, amplified, and transmitted to the satellite via a large uplink antenna.

Comparison with other satellite systems. Other satellite systems are used to provide "big dish" DTH services, notably communications satellites operating in the 3.7–4.2-GHz band (the C band). These satellites are principally designed to offer general telecommunications services, including the transmission of television programming to cable television companies and television broadcast stations. These satellites operate at lower power levels, and thus much larger antennas, typically 6–10 ft (1.8–3 m) in diameter, are needed to receive the signals. Direct-to-home service was initiated during the 1980s via these systems using analog transmission, but has since been largely supplanted by small dish systems.

Comparison with other technologies. DBS systems have expanded very rapidly in the United States and elsewhere because of their superior quality, expansive programming lineup, and centralized, efficient technical and administrative facilities. DBS technical quality is due to its all-digital delivery and the use of microwave frequencies and highly directive antennas. Cable companies are now achieving wide deployment of their own all-digital solutions and competing favorably with direct broadcast. Telephone companies have announced new video delivery networks using fiber-optic cable or Digital Subscriber Line (DSL) transmission over telephone wiring. All these competing systems use similar underlying technology for television signal processing and for the consumer's set top box. The primary network differences are in the means of transmission to the home. *See* CABLE TELEVISION SYSTEM; COMMUNICATIONS SATELLITE; OPTICAL COMMUNICATIONS; TELEVISION.

John P. Godwin

Bibliography. T. T. Ha, *Digital Satellite Communications*, 2d ed., McGraw-Hill, 1990; G. Maral and M. Bousquet, *Satellite Communication Systems*, 4th ed., Wiley, 2002; M. Robin and M. Poulin, *Digital Television Fundamentals*, 2nd ed., McGraw-Hill, 2000; B. Sklar, *Digital Communications: Fundamentals and Applications*, 2d ed., Prentice-Hall, 2001.

Saturable reactor

An iron-core inductor in which the effective inductance is changed by varying the permeability of the core. Saturable-core reactors are used to control large alternating currents where rheostats are impractical. Theater light dimmers often employ saturable reactors.

In the **illustration** of two types of saturable-core reactors, illus. *a* shows two separate cores, while in illus. *b* a three-legged core is formed by placing two two-legged cores together. The load winding, connected in series with the load, carries the alternating

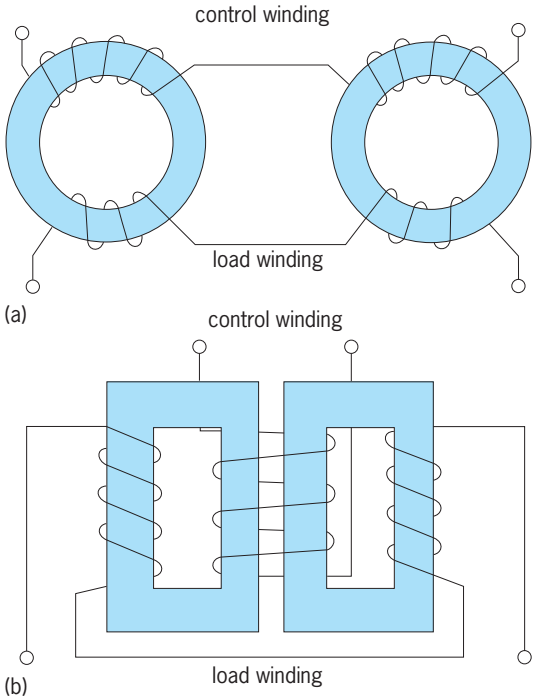

Typical construction of saturable-core reactors. (*a*) Two separate cores. (*b*) Three-legged cores.

current and acts as an inductive element. The control winding carries a direct current of adjustable magnitude, which can saturate the magnetic core. When the core is saturated by the direct current, the effective inductance (and therefore the reactance) of the coils in the ac circuit is small. Little voltage can therefore be induced in the coil to oppose the applied voltage and therefore reduce the voltage that is supplied to the load.

Reducing the magnitude of the control current reduces the intensity of saturation. This increases the reactance of the load winding. As the reactance increases, the voltage drop in the load winding increases and causes a reduction in the magnitude of the voltage applied to the load. *See* INDUCTANCE; MAGNETIZATION. Wilson S. Pritchett

Bibliography. T. Croft, J. Watt, and W. I. Summers, *American Electricians' Handbook*, 13th ed., 1996; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000.

Saturation

The condition in which, after a sufficient increase in a causal force, further increase in the force produces no additional increase in the resultant effect. Many natural phenomena display saturation. For example, after a magnetizing force becomes sufficiently strong, further increase in the force produces no additional magnetization in a magnetic circuit; all the magnetic domains have been aligned, and the magnetic material is saturated.

After a sponge has absorbed all the liquid it can hold, it is saturated. In thermionic vacuum tubes thermal saturation is reached when further increase in cathode temperature produces no (or negligible) increase in cathode current: anode saturation is reached when further increase in plate voltage produces substantially no increase in anode current. *See* DISTORTION (ELECTRONIC CIRCUITS); VACUUM TUBE.

In an *npn* transistor when the collector voltage becomes more negative than the base voltage, the base-to-collector diode becomes forward-biased: the collector then emits and the transistor is saturated. Current flow is controlled by a stored charge of carriers in the base region. *See* SEMICONDUCTOR; TRANSISTOR.

In colorimetry the purer a color is, the higher its saturation. Radiation from a color of low saturation contains frequencies throughout much of the visible spectrum. *See* COLORIMETRY.

In an induced nuclear reaction, saturation exists when the decay rate of a given radionuclide is equal to its rate of production. In an ionization chamber, saturation exists when the applied voltage is high enough to collect all the ions formed by radiation but not high enough to produce ionization by collision. When addition of the dissolved species produces no further increase in the concentration of a solution at a certain temperature and pressure, the solution is said to be saturated. *See* SOLUTION.

Frank H. Rockett

Saturation current

A term having a variety of specific applications but generally meaning the maximum current which can be obtained under certain conditions.

In a simple two-element vacuum tube, it refers to either the space-charge-limited current on one hand or the temperature-limited current on the other. In the first case, further increase in filament temperature produces no significant increase in anode current, whereas in the latter a further increase in voltage produces only a relatively small increase in current. *See* VACUUM TUBE.

In a gaseous-discharge device, the saturation current is the maximum current which can be obtained for a given mode of discharge. Attempts to increase the current result in a different type of discharge. Such a case would be the transition from a glow discharge to an arc discharge. *See* ELECTRICAL CONDUCTION IN GASES.

A third case is that of a semiconductor. Here again the saturation current is that maximum current which just precedes a change in conduction mode. *See* SEMICONDUCTOR. Glenn H. Miller

Saturn

The second largest planet in the solar system and the sixth in order of distance to the Sun. The outermost planet known prior to 1781, Saturn is surrounded by a beautiful system of rings, more than 30 moons, and a complex magnetosphere. Despite the planet's huge size, its mean density is so low it could float in water. Saturn is also the only planet that has a satellite (Titan) with a dense atmosphere. This distant planetary system has been visited by four spacecraft: *Pioneer 11* in 1979, *Voyagers 1* and *2* in 1980–1981, and *Cassini-Huygens* since 2004. The first three were flybys; the last is a joint United States–European mission consisting of a Saturn orbiter and a probe into the atmosphere of Titan.

Orbit and physical elements. The orbit of Saturn has a semimajor axis or mean distance to the Sun of 8.95×10^8 mi (1.43×10^9 km), an eccentricity of 0.056, and its plane is inclined to the plane of the ecliptic at an angle of 2.5°. With a mean orbital velocity of 6 mi/s (9.67 km/s), Saturn makes one revolution about the Sun in 29.42 years. *See* PLANET.

The equatorial diameter of Saturn is about 75,000 mi (120,540 km), and the polar diameter about 67,600 mi (108,700 km). The volume is 769 (Earth = 1) with a few percent uncertainty. The polar flattening caused by the rapid rotation is the largest of all the planets, and the ellipticity, $(r_e - r_p)/r_e = 0.098$, is 30 times the value for Earth (r_e is the equatorial radius and r_p is the polar radius).

The mass, about 95.2 (Earth = 1) or 1/3500 (Sun = 1), is accurately determined from the motions of the planet's brighter satellites. The mean density is 0.70 g/cm^3, the lowest mean density of all the planets. The corresponding value of the mean gravity

at the visible surface (superimposed cloud layers) is 1.15 (Earth = 1) or 11.3 m/s^2.

Photometric properties. The apparent visual magnitude of Saturn at mean opposition is +0.7 when the ring is seen edgewise, and the corresponding value of the reflectivity (geometric albedo) is 0.47. *See* ALBEDO.

Appearance. Observed through a telescope, Saturn appears as an elliptical disk, darkened near the limb and crossed by a series of barely discernible bands parallel to the equator. Frequently only the bright equatorial zone and the two darker tropical bands on either side of it are visible (**Fig. 1**). The rings may be seen even with a relatively small telescope, but their visibility changes with the position of the planet in its orbit, because the rotation axis (of both the planet and the rings) is inclined 27° to the perpendicular to the orbital plane. Thus Saturn has seasons similar to those on Earth, but each season lasts 7.32 years (**Fig. 2**).

The rotation period of Saturn's interior is 10^h40^m, as determined by radio emissions controlled by the planet's magnetic field. The visible cloud layers of Saturn are much more homogeneous in appearance than those of Jupiter. There is no feature comparable to the Great Red Spot, and the contrast of the features (huge storm systems in the cloud decks) that are visible is very low. The contrast may be enhanced by the use of special filters. By studying the movement of these storms it has been possible to determine that the circulation patterns on Saturn are also very different from those on Jupiter. At a latitude of ±40°, the atmosphere rotates with the same velocity as the interior, but wind velocities increase smoothly toward the equator, where, at the time of the *Voyager* observations (1980–1981), they reached a value of 1100 mi/h (500 m/s), about four times faster than Jupiter's equatorial jet. Twenty-five years later, the *Cassini Orbiter* found the wind speed at the equator to be considerably reduced, for reasons that are not yet clear. There is no alternation of easterly and westerly currents corresponding to transitions between belts and zones, as there is on Jupiter, except at latitudes above ±40°. This difference between Saturn and Jupiter represents a fundamental difference in global circulation, perhaps related to the relative sizes of the cores of the two planets. *See* JUPITER.

Atmosphere. The optical spectrum of Saturn is characterized by strong absorption bands of methane (CH_4) and by much weaker bands of ammonia (NH_3). Absorption lines of molecular hydrogen (H_2) have also been detected. The estimated quantities of these gases that are present above the clouds are equivalent to STP path lengths of about 30 mi or 50 km (H_2), 200 ft or 60 m (CH_4), and 6 ft or 2 m (NH_3). Here STP refers to standard temperature (273 K or 32°F) and pressure (1 atm or 1.01325×10^5 pascals). In these units, the STP path length from sea level through the Earth's atmosphere would be 5 mi (8 km). The presence of about 5 mi (8 km) of helium (He) has been deduced indirectly from infrared observations of pressure-broadened hydrogen emission lines.

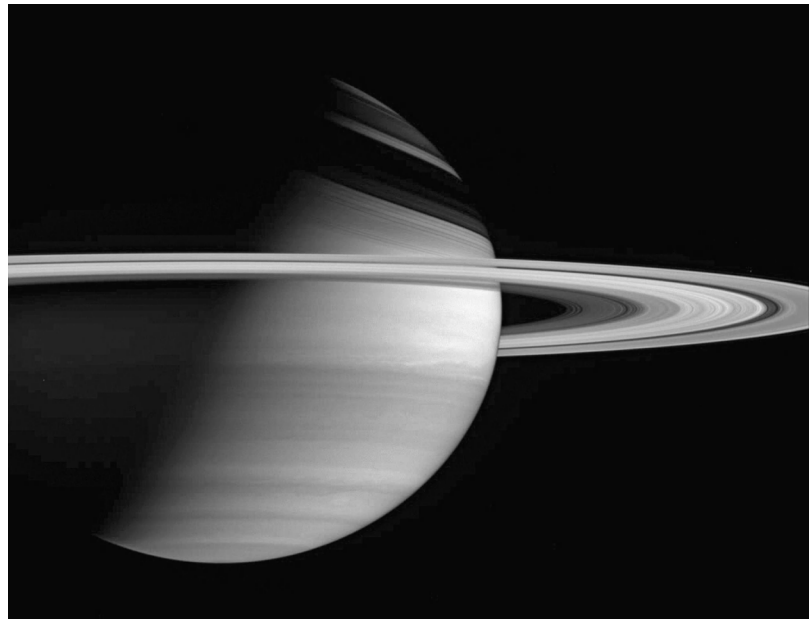

Fig. 1. Saturn viewed from the *Cassini Orbiter*. The soft velvety appearance of the low-contrast banded structure is due to scattering by the haze layer above the planet's cloud deck. The image has been rotated so that north is up. The Sun illuminates Saturn from below and Saturn's tilt throws shadows of the rings onto the northern hemisphere during the current season. The image was taken on January 23, 2005 at a distance of 1.7×10^6 mi (2.8×10^6 km) from Saturn. (*NASA/JPL/Space Science Institute*)

The temperature the planet should assume in response to solar heating is calculated to be about 76 K (−323°F), somewhat lower than the measured value of 92 K (−294°F). This suggests that Saturn has an internal heat source of roughly the same magnitude as that on Jupiter. Relatively intense infrared emission lines near a wavelength of 12 micrometers have been identified, resulting from ethane (C_2H_6) and acetylene (C_2H_2) formed in the upper atmosphere from the dissociation of methane. As in the case of Jupiter, a thermal inversion exists in this region of Saturn's atmosphere. Emission bands of methane and phosphine have also been detected here, as well as bands of water (H_2O) and carbon dioxide (CO_2), thought to arise from in-falling material. This region is well above the main cloud layer, which should consist primarily of frozen ammonia crystals, with an admixture of some other substances such as sulfur to provide the yellowish color sometimes observed in the equatorial zone. The ammonia cirrus on Saturn is apparently denser and more ubiquitous than on Jupiter, since one does not see

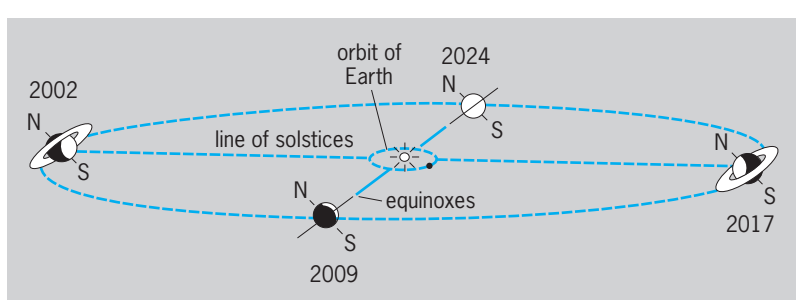

Fig. 2. Presentations of Saturn's ring system.

through it to lower cloud layers (Fig. 1). This difference probably results from the lower atmospheric temperature and smaller gravity of Saturn, which will act together to spread out and increase the depth of the cloud layer.

Internal structure and radiation belts. Observations of Saturn at radio frequencies indicate that the temperature steadily increases with depth into the atmosphere. Theoretical models for the internal structure of Saturn are similar to those for Jupiter, that is, a dense core surrounded by hydrogen compressed to a liquid metallic state that gradually merges into an extremely deep atmosphere. As on Jupiter, the gradual precipitation of helium from solution in the metallic hydrogen surrounding the core liberates heat from Saturn's interior and produces the subsolar abundance of helium (relative to hydrogen) found in Saturn's atmosphere. The existence of a magnetic field and belts of trapped electrons was initially deduced from observations of nonthermal radiation at dekameter wavelengths and was mapped out in detail by in situ measurements from instruments on the spacecraft that have visited the planet. The *Cassini Orbiter* is continuing these investigations. Saturn's magnetic field has the same polarity as Jupiter's; a terrestrial compass taken to either planet would point south instead of north. The magnetic moment of Saturn is 4.3×10^{28} gauss $\cdot$ cm^3 (4.3×10^{18} teslas $\cdot$ m^3), 500 times Earth's and 34 times smaller than Jupiter's. But the magnetic field at the cloud tops near the equator is only 0.2 gauss (2×10^{-5} tesla), two-thirds the value of Earth's field, because of Saturn's greater size. The magnetosphere contains a plasma of charged particles, but no distinct torus like the one associated with Io in the Jupiter system. The charged particle belts are absent in the region of the rings. Radio signals resembling radiation from lightning discharges in the atmosphere were detected by the *Voyager* spacecraft. *See* MAGNETOSPHERE.

Origin. Jupiter and Saturn are relatively similar bodies. Both seem to have bulk compositions close to that of the Sun and the other stars, and both are rich in hydrogen and helium. In that sense, a large fraction of the mass they contain consists of the same primitive material from which the entire solar system was formed, whereas the inner planets such as Earth and Mars never accumulated a solar proportion of the light gases.

This conclusion has been strengthened by the detection of the heavy isotope of hydrogen known as deuterium, in the form of HD (mixed with ordinary hydrogen, H_2) and CH_3D (mixed with ordinary methane, CH_4). The derived value of D/H appears to be identical to the value of this ratio measured in Jupiter's atmosphere by the *Galileo* entry probe, which is 10 times lower than the value in seawater on Earth and twice as high as the value in local interstellar hydrogen. This demonstrates that the hydrogen that makes up most of the mass of both Jupiter and Saturn was captured directly from the solar nebula when these planets formed 4.6 billion years ago. The hydrogen in seawater came from water delivered to the Earth by comets and rocks and has a higher value of D/H, while the current interstellar ratio reflects the result of nucleosynthesis in stars, which converts deuterium into helium and has therefore depleted this isotope during the last 4.6 billion years by a factor of 2. *See* DEUTERIUM; NUCLEOSYNTHESIS.

However, both Jupiter and Saturn show an enhancement of C/H (as determined from methane and hydrogen) compared with the Sun. On Jupiter, C/H is three times the solar ratio, while on Saturn the enhancement is a factor of 6. This suggests that both planets formed in a two-stage process that led initially to formation of a large core of approximately the same size that generated an outgassed, secondary atmosphere, followed by the attraction of an envelope of gases from the surrounding nebula. *See* SOLAR SYSTEM.

Ring system. The most remarkable feature associated with Saturn is the complex ring system that surrounds the planet (**Fig. 3; color plate 1**). The system is divided into four main regions, designated A through D. The narrow F ring is located just beyond the edge of ring A, and there are G and E rings still farther out. Each of the four main regions is subdivided into many individual "ringlets," so that Saturn is actually surrounded by thousands of rings.

The ring system is made up of myriad separate particles that move independently in flat, mostly circular orbits in Saturn's equatorial plane. The discontinuous, meteoric nature of the rings is demonstrated directly by spectroscopic observations, which show that the inner edge of the ring revolves about the planet faster than the outer edge, and the material in each ringlet moves precisely with the velocity that independent satellites would have at the same distance from the planet.

Appearance from Earth. The nearly circular ring system looks like an ellipse whose appearance changes with

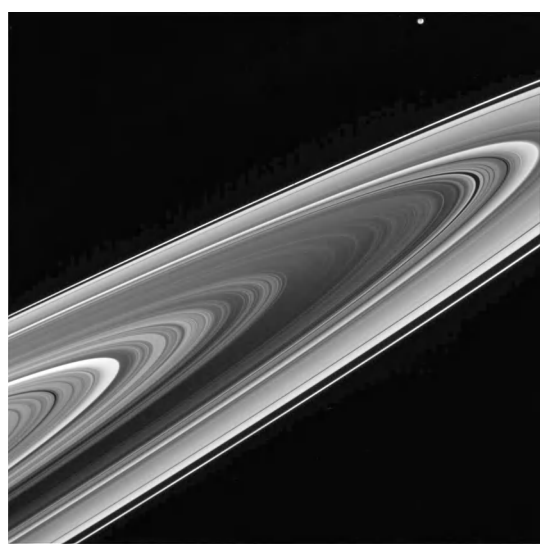

Fig. 3. Visible-light image of Saturn's rings from the *Cassini Orbiter*. The Sun is shining through the rings so that they take on the appearance of a photonegative; the dense B ring at the center blocks much of the incoming light, while the less dense regions scatter and transmit light. The image was taken on April 15, 2005 at a distance of 350,000 mi (570,000 km) from Saturn. (*NASA/JPL/Space Science Institute*)

the relative positions of Earth and Saturn (Fig. 2). The maximum opening of the rings occurs when the system is tilted 27° to the line of sight. In contrast, the rings appear as a thin line when Earth crosses Saturn's equatorial plane. Depending on the position of Earth in its orbit, the slightly variable point of view causes a slight annual oscillation of the tilt angle. Thus, when the tilt is near 0°, Earth may cross the plane of the ring system either once or three times. During these edge-on configurations, it is possible to search for faint inner satellites and distant rings that ordinarily are hidden by the bright light reflected by the main ring system.

Nomenclature and structure. The bright outer ring, A, has an outside diameter of 170,000 mi (274,000 km) and an inner diameter of 150,000 mi (244,000 km). This ring contains many wavelike patterns, indicating gravitational disturbances by the satellites. A dark gap in the A ring, often called the Encke division, was found to contain a narrow, elliptical, discontinuous and kinky ring, together with a small satellite, Pan. The broad (2800 mi or 4500 km wide) division separating ring A from B is called the Cassini division after its discoverer and contains material comprising at least five discrete rings, each of which shows some internal structure.

Ring B, the brightest component of the system, has an outer diameter of 146,000 mi (235,000 km) and an inner diameter of 114,000 mi (184,000 km). This ring contains the largest density of material in the system. It is slightly less bright in its inner regions, where it is also more transparent.

Ring C, sometimes called the crepe ring, is much fainter and more transparent. It appears as a dusky band in projection against the disk of the planet and only faintly against the sky. Its outer diameter is 114,000 mi (184,000 km) and its inner diameter 92,000 mi (149,000 km). This means the inner edge of ring C is just 8700 mi (14,000 km) above the visible cloud deck on the planet.

A fourth zone, D, of the rings of Saturn between C and the globe is fainter than the crepe ring, and still fainter rings lie beyond the A ring.

The E ring begins inside the orbit of Enceladus at a distance of about 112,000 mi (180,000 km) from the center of Saturn, and extends outward past the orbit of Dione, fading from view at about 300,000 mi (480,000 km) from the planet's center. The ring is very much brighter just at the orbit of Enceladus, suggesting that this satellite is somehow responsible for the production of the material seen in the ring.

The F ring (**Fig. 4**) lies just outside the A ring with an average diameter of 174,600 mi (281,000 km), and actually consists of more than five separate strands, which may not all be in the same plane. The two brightest strands deviate markedly from ellipses, showing the effects of nongravitational forces and the influence of embedded moonlets. The F ring is held in place by two small satellites, one inside and one outside. A third guards the outer edge of the A ring, and a fourth is responsible for the Encke gap.

Finally, the tenuous G ring has an average diameter of 211,000 mi (340,000 km), 21,000 mi (33,400 km) outside the A ring. Although this feature resembles the E ring, no known satellites are associated with it, so both its origin and its ability to persist are not understood.

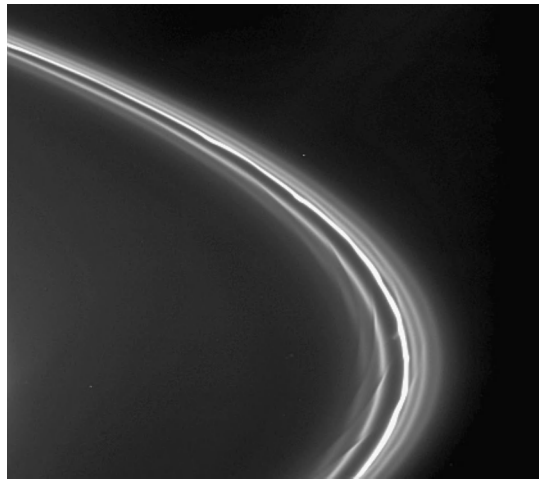

Fig. 4. Saturn's F ring viewed from the *Cassini Orbiter*. Features resembling drapes and kinks are visible, and several distinct ringlets are present, in addition to the bright knotted core of the ring. The structure in the ring and its strands has been caused by Prometheus, the inner F ring shepherd moon that swept by this region shortly before the image was taken on January 19, 2005 at a distance of 1.2×10^6 mi (1.9×10^6 km) from Saturn. (*NASA/JPL/Space Science Institute*)

To place this complex system in perspective, the mean distance from the Earth to its Moon is 240,000 mi (384,000 km), so Saturn and its rings (except for the outer edge of the E ring) would just fit in the space between the Earth and the Moon.

In addition to these nearly circular rings, transient radial features exist in the B ring. These features, called spokes, appear to be clouds of micrometer-size charged particles that are initially controlled by the planet's magnetic field but soon begin moving in keplerian orbits like the larger particles. The change in solar illumination, hence charging, of the ring particles between the *Voyager* and the beginning of the *Cassini* missions changed the charging process sufficiently that initially no spokes were visible to the *Cassini* cameras.

Origin and nature. Édouard Roche proved in 1849 that a liquid satellite with the same density as its planet cannot form if it is closer than 2.44 times the planetary radius. Within this distance, which is known as the Roche limit, disruptive tidal forces (the gradient of the planet's gravitational field) will be greater than the self-gravity of the satellite. But a rocky or icy satellite is held together by stronger forces, and will, therefore, not be disrupted by the planet's field outside a distance of 1.35 radii. In fact, the F ring and its guardian satellites are at 2.33 radii, just inside the Roche limit. If dispersed matter exists within the Roche limit, for example, in the form of a planetary ring, the gravitational attraction of individual ring particles for one another will be insufficient to overcome the planet's tidal forces. Thus a satellite cannot form from a ring that lies within the Roche limit. The ring will persist. *See* ROCHE LIMIT.

In 1859 J. C. Maxwell was able to show that a

TABLE 1. Saturn's satellites[a]

Satellite	Mean distance from center of Saturn, 10^3 km	Revolution period d	h	m	Mean density, g/cm^3	Diameter, km	Visual magnitude at mean opposition
XVIII Pan	134	0	13	49	?	(20)	(19)
XV Atlas	137	0	14	26	?	30[b]	(18)
XVI Prometheus	139	0	14	43	?	100[b]	(16)
XVII Pandora	142	0	15	05	?	90[b]	(16)
X Janus	151	0	16	41	?	190[b]	(14)
XI Epimetheus	151	0	16	41	?	120[b]	(15)
I Mimas	187	0	22	37	(1.2)	390	12.9
XXXII Methone	194	1	0	13	?	3	(23)
XXXIII Pallene	211	1	3	27	?	4	(22)
II Enceladus	238	1	08	53	1.6	500	11.7
III Tethys	295	1	21	18	(1.2)	1060	10.2
XIII Telesto	295	1	21	18[c]	?	25[b]	(19)
XIV Calypso	295	1	21	18[d]	?	25[b]	(19)
XXXIV Polydeuces	377	2	17	41[e]	?	3.5	(23)
IV Dione	378	2	17	41	1.4	1120	10.4
XII Helene	378	2	17	41[f]	?	30[b]	(18)
V Rhea	526	4	12	25	1.3	1530	9.7
VI Titan	1221	15	22	41	1.88	5550[g]	8.3
VII Hyperion	1481	21	06	38	?	255[b]	14.2
VIII Iapetus	3561	79	07	56	1.2	1460	10.2–11.9
IX Phoebe	12,940	550	11		?	220	16.4

[a]Values in parentheses still uncertain. The 13 irregular satellites discovered in 2000 and 2004 are listed separately in Table 2. Four satellites have not been observed well enough for their orbits to be determined, and therefore they are not listed in this table. 1 km = 0.62 mi.
[b]Irregular in shape
[c]Librates about trailing (L_5) lagrangian point of Tethys's orbit.
[c]Librates about leading (L_4) lagrangian point of Tethys's orbit.
[e]Librates about trailing (L_5) lagrangian point of Dione's orbit.
[f]Librates about leading (L_4) lagrangian point of Dione's orbit.
[g]This diameter refers to top of haze layer. Diameter of solid body is 5150 km.

ring system of small mass, formed of a large number of particles, is quite stable against external perturbations such as those caused by the larger satellites. The aggregate mass of Saturn's ring system is quite small, probably less than one-half the mass of the Earth's Moon. Periodic perturbations by the major satellites are responsible, in part, for the main divisions of Saturn's rings in the same way that perturbations by Jupiter cause the Kirkwood gaps in the asteroid belt. The cause of many small divisions (Fig. 3) is still obscure. It has been proposed that tiny moonlets may be present in these gaps, stabilizing individual rings even as the F ring is stabilized by its two satellites. But a careful search has failed to reveal any small satellites (with diameters of 5–10 km) in any of the gaps except Encke's. *See* ASTEROID; PERTURBATION (ASTRONOMY).

The reflection spectrum of most of the ring system is identical with that of water ice, indicating that these ring particles are either ice or covered with ice. However, there are discrete ringlets with differing reflective properties, suggesting the presence of other materials such as silicates or organic compounds. The ring system is a surprisingly good reflector of radar waves, suggesting that some relatively large (diameter approximately 0.6 mi or 1 km) particles may be present. An average effective particle size of 3 ft (1 m) has been indicated. The mean thickness of the ring system is not well determined, but *Voyager* observations of the occultation of a star by the rings demonstrated that the outer edge of the A ring is only 500 ft (150 m) thick.

Satellites. As of June 1, 2005, Saturn had 34 well-established satellites, with at least 13 more whose orbits were being confirmed. Still more are expected to be discovered (**Tables 1** and **2**). The largest and brightest, Titan, is visible with small telescopes; the other satellites are much fainter. Two inner satellites (X Janus and XI Epimetheus) are very nearly in the same orbit. Two small satellites are, respectively, in the leading and trailing lagrangian points of Dione's orbit, and two are in the leading and trailing lagrangian points of Tethys's orbit. All of these objects are small (less than 60 mi or 100 km long) and irregular in shape. The *Cassini* spacecraft found two even smaller objects (0.6–1.2 mi or 1–2 km) in independent orbits between Mimas and Enceladus (Table 1). More discoveries may follow. *See* CELESTIAL MECHANICS.

Titan. Titan shows a measurable disk in large telescopes; the mean apparent diameter corresponds to a linear diameter of approximately 3440 mi (5550 km). But this diameter refers to the satellite's atmosphere, which contains several layers of smog consisting of aerosols produced photochemically by incident sunlight. The solid surface of Titan has a diameter of 3200 mi (5150 km), making this satellite larger than Mercury but slightly smaller than Jupiter's giant Ganymede. This large satellite has a mass about two times that of the Moon, with a corresponding mean density of 1.9 g/cm^3. The low density (Moon = 3.3) means that this object contains a large fraction of icy material and is thus quite different from the Moon or the inner planets in composition. Furthermore, it is large and cold enough to retain an atmosphere of gases with relatively high molecular weights.

In fact, Titan has a thick, nitrogen (N_2)-dominated atmosphere that contains a few percent of methane

Fig. 5. Image from the *Huygens* probe from an altitude of about 5 mi (8 km) above Titan's surface, showing the boundary between lighter uplifted terrain, marked with what appear to be drainage channels, and darker, lower areas. (*ESA/NASA/JPL/University of Arizona*)

(CH_4) and exerts a surface pressure of 1.5 bars (1.5×10^5 Pa), or 1.5 times the sea-level pressure on Earth. Photochemical reactions in the atmosphere produce traces of many interesting compounds, such as acetylene (C_2H_2), hydrogen cyanide (HCN), ethane, and propane (C_3H_8), which condense and polymerize in the stratosphere, forming thick surface-shielding haze layers. The atmosphere is sufficiently thick to have allowed the safe entry and 2 h, 20 min descent by parachute of the *Huygens* probe on January 14, 2005. Instruments on the probe successfully measured the atmospheric temperature, pressure, and composition along the descent trajectory while obtaining pictures (**Fig. 5**) and spectra of the surface. After landing, the probe continued to transmit data for another hour (**color plate 2**). The surface of Titan is so cold (94 ± 2 K or -290 ± 4°F) that rivers and lakes of liquid methane must be present. Branching channel systems resembling rivers on Earth were seen by the probe, which also detected liquid methane in the surface material at the landing site. The cameras and the radar on the orbiter have found very few craters, indicating that Titan has some active processes at work that are burying and erasing them. A search for lakes, seas, and geysers of methane is continuing. The atmospheric chemistry and the solid surface on which liquids and aerosols can accumulate make Titan an attractive natural laboratory in which future investigators will be able to test ideas about chemical evolution on the primitive Earth.

Other inner satellites. The other inner satellites encompass a variety of characteristics. Most have densities near unity, indicating a predominantly icy composition. The variations in density that do occur appear random, rather than showing a radial trend as is the case for the Jupiter system. These density differences must reflect variations in the proportion of rocky material embedded in the ices. The surfaces of these objects are covered with impact craters, with the exception of Enceladus. Large regions of the surface of Enceladus are free of craters, indicating reworking of the surface in relatively recent times. Coupled with the unusually high reflectivity of this satellite (close to 100%), the modified surface suggests internal activity leading to partial melting or pulverization of the icy surface, and the production and expulsion of the tiny ice grains that populate the E ring. The *Cassini* spacecraft discovered the emission of water vapor from fissures at the satellite's south pole. Iapetus is unique in the solar system in that its trailing hemisphere is six times brighter than the leading one. The dark material on the leading hemisphere appears to be a mixture of organic compounds, but the origin of this material and the process that coated the satellite's surface with it remain unknown. All of these satellites except Hyperion keep the same hemisphere facing Saturn as they revolve around it, meaning that their rotational periods are equal to the periods of revolution as a result of tidal friction, just as is the case for the Moon. Hyperion exhibits an aperiodic, "chaotic" rotation, which, with its irregular shape, may be an indication of a relatively recent collision. *See* MOON.

Irregular satellites. Saturn also has a number of more distant satellites that are classified as "irregular" because of the large eccentricities or inclinations (or both) of their orbits. These satellites must have been captured by Saturn early in its history rather than forming together with it from the solar nebula. The largest of these is Phoebe (**Fig. 6**), which moves in a retrograde direction (opposite to that of the inner satellites and to the direction of the planets about the Sun) in an orbit of relatively high eccentricity (0.16).

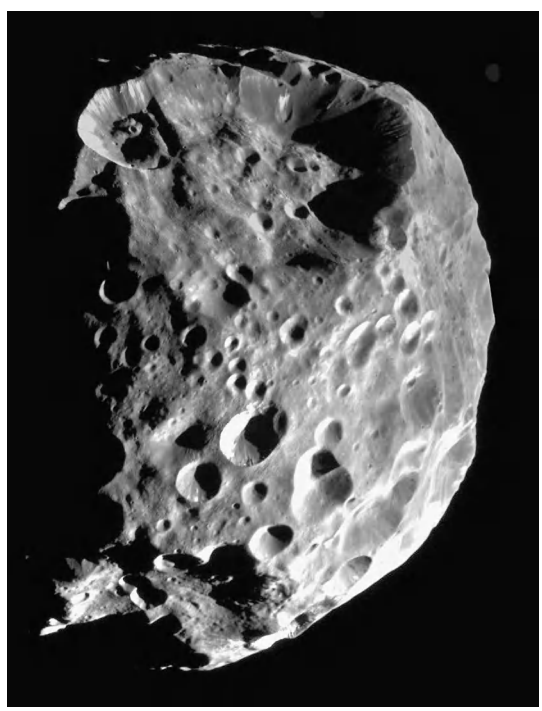

Fig. 6. Mosaic of two images of Saturn's outer satellite Phoebe taken during *Cassini*'s flyby on June 11, 2004. The image shows evidence for the theory that Phoebe is an ice-rich body coated with a thin layer of dark material. (*NASA/JPL/Space Science Institute*)

TABLE 2. Irregular satellites of Saturn*

Satellite	Mean distance from center of Saturn, 10^6 km†	Eccentricity	Inclination	Revolution period, years	Visual magnitude at mean opposition	Radius, km†
XXVI Albiorix	17.9	0.38	33.1°	2.43	20.5	13
XX1 Tarvos	18.2	0.54	33.5°	2.53	22.1	7
XXVIII Erriapo	18.2	0.48	33.8°	2.50	23.0	4
XXIX Siarnaq	16.8	0.26	44.8°	2.18	20.1	16
XX Paaliaq	15.2	0.36	45.1°	1.88	21.3	10
XXIV Kiviuq	11.3	0.33	46.1°	1.23	22.0	7
XXII Ijiraq	11.4	0.32	46.7°	1.24	22.6	5
XXXI Narvi	18.8	0.43	145.8°	2.76	24.0	3.3
XXVII Skathi	15.5	0.27	152.6°	1.97	23.6	3
XXV Mundilfari	18.4	0.23	167.4°	2.56	23.8	3
XIX Ymir	23.1	0.34	173.1°	3.59	21.7	8
IX Phoebe	12.94	0.16	174.7°	1.51	16.4	110
XXX Thrymr	20.1	0.44	175.9°	2.92	23.9	3
XXIII Suttungr	17.6	0.12	176.0°	2.40	23.9	3

*Data provided by the leader of one of the discovery teams, Dr. Brett Gladman.
†1 km = 0.62 mi.

Phoebe's rotation period is only 9 h, completely out of synchrony with its 550-day period of revolution. *See* RETROGRADE MOTION (ASTRONOMY).

Twelve additional irregular satellites were confirmed in 2000 and one in 2004 (Table 2) as a result of deliberate searches with sensitive new detectors on Earth-based telescopes. More discoveries are expected. These small objects (the largest is only 20 mi or 32 km across) can be classified into four distinct groups by the inclinations of their orbits: three near 34°, four near 46°, two near 150°, and four near Phoebe's inclination of 175°. All the orbits have high eccentricities, and this plus the high inclinations mark these satellites as irregular. The distinct groups suggest separate capture events in which the original object was broken apart during the capturing process. Similar families exist in the Jupiter satellite system. The exploration of the Saturn system by the *Cassini Orbiter* is continuing. *See* SATELLITE (ASTRONOMY).

Tobias C. Owen

Bibliography. J. K. Beatty, C. C. Petersen, and A. Chaikan (eds.), *The New Solar System*, 4th ed., Sky Publishing/Cambridge University Press, 1999; K. R. Lang, *The Cambridge Guide to the Solar System*, Cambridge University Press, 2003; D. Morrison and T. Owen, *The Planetary System*, 3d ed., Addison Wesley, 2003; Special issue on *Voyager 1* encounter with the Saturnian system, *Science*, 212(4491):159–243, 1981; Special issue on *Voyager 2* encounter, *Science*, 215(4532):499–594, 1981; Special issue on *Cassini* at Titan, *Science*, 308(5724):969–995, 2005.

Saurischia

One of two major monophyletic clades of Dinosauria, the other being Ornithischia. Saurischian dinosaurs take their name (meaning "lizard hipped") from the structure of their pelvis. In saurischians, the pubis bone points downward and forward from the hip joint, and the ischium extends downward and backward, so that the two bones are separated by a wide angle in side view (see **illustration**). This feature allows the majority of saurischians to be easily distinguished from ornithischian dinosaurs (in which the pubis has rotated so that it points downward and backward and thus lies alongside the ischium). However, the saurischian condition is actually a retention of the primitive reptilian condition (which is also present in crocodilians, squamates, as well as many extinct groups) and, therefore, cannot be regarded as a unique feature of saurischians. To add to the confusion, some saurischian dinosaurs (various advanced theropod groups and their direct descendants, the birds) have independently evolved an ornithischian-type pelvis. However, Saurischia can be satisfactorily defined as a monophyletic clade of animals on the basis of many other detailed anatomical features shared by all members of the group, largely pertaining to the structure of the hand and vertebral column, such as the possession of a stout, robust thumb and the presence of

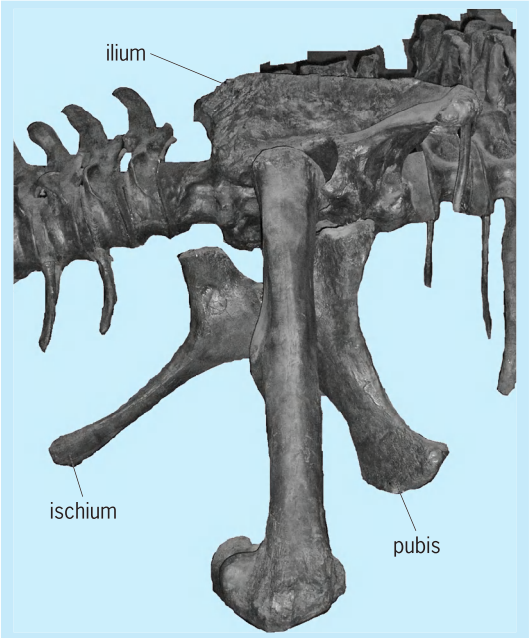

Pelvic girdle of the sauropod dinosaur *Camarasaurus*, showing the hip orientation typical of saurischian dinosaurs in which the pubis and ischium diverge from the hip joint. The front of the animal is to the right.

incipient lamination on the vertebrae. *See* DINOSAURIA; ORNITHISCHIA.

The earliest saurischians—including *Staurikosaurus*, *Herrerasaurus*, and *Eoraptor*—are known from deposits of Late Triassic age (approximately 225 million years ago) in Brazil and Argentina. Early in its evolutionary history, Saurischia split into two major monophyletic clades: the Sauropodomorpha and Theropoda. Both groups diversified rapidly in the Late Triassic–Early Jurassic interval, each increasing in species richness and abundance to become the dominant vertebrates in terrestrial ecosystems. They quickly achieved a global distribution, and the remains of early saurischians are known from all continents.

Sauropodomorpha. Sauropodomorpha can be defined on the basis of many anatomical characteristics, including a relatively small skull (length is less than 50% of femur length) and various features of the skull and dentition (for example, tooth crowns with coarse marginal serrations set at 45° to the long axis of the tooth) and the postcranial skeleton (for example, a "funnel-shaped" fifth metatarsal). Traditionally, sauropodomorphs have been divided into two groups: Sauropoda and Prosauropoda. Sauropods are generally recognized as a monophyletic group on the basis of numerous synapomorphies (shared derived traits), which are largely concentrated in the vertebral column (for example, the presence of a complex system of bony laminae on the cervical vertebrae), and limbs (for example, elongated forelimbs relative to other dinosaurs). The relationships of prosauropods are more controversial: some paleontologists suggest that they form a monophyletic clade, whereas others believe they form an array of forms that are ancestral to the later sauropods. In general, all sauropodomorphs share the same body plan: a small head, elongate neck, barrel-shaped body, and long, counterbalancing tail.

Prosauropods and basal sauropodomorphs first appeared early in the Late Triassic, represented by small (probably bipedal) animals such as *Saturnalia* and *Thecodontosaurus*, and quickly achieved a global distribution (excluding Australia). These forms were largely herbivorous, but may have taken animal prey and/or carrion occasionally. Later prosauropods increased in size, with some forms reaching 10 m (33 ft) in length: these larger animals were probably entirely herbivorous. The majority of prosauropods were facultatively bipedal, spending most of their time on all fours, but occasionally walking bipedally: very large prosauropods (such as *Riojasaurus*) were obligate quadrupeds. Prosauropods were confined to the Late Triassic–Early Jurassic, an interval in which they were the dominant terrestrial herbivores. The reason for their extinction is unknown, but coincides with the diversification of the ecologically similar sauropods.

Sauropods also appeared in the Late Triassic, but initially were rare members of early dinosaur faunas. In general, sauropod evolution through the Jurassic and Cretaceous followed several trends, including increased body size [culminating in animals such as *Argentinosaurus*, from the Late Cretaceous of Patagonia, which may have reached a weight of 70 metric tons (77 tons)] and increased neck length, the latter allowing feeding over wide vertical and horizontal ranges. All sauropods were obligate quadrupeds and were more advanced herbivores than prosauropods, having acquired a number of cranial and dental features (such as interlocking dentitions) that allowed more thorough processing of plant food than could be achieved by prosauropods. By the Middle Jurassic, sauropods had become the dominant terrestrial vertebrates and had achieved a global distribution (excluding Antarctica). Sauropods reached their acme in the Late Jurassic, after which they declined in diversity and abundance, though a second peak in diversity also occurred in the middle to Late Cretaceous. The group survived until the end of the Cretaceous Period. *See* CRETACEOUS; JURASSIC; TRIASSIC.

Theropoda. All theropods possess an intramandibular joint within their lower jaws, which allowed the front end of the mandible to flex slightly with respect to the back of the jaw. In addition, various features of the hands and vertebrae (such as elongated processes on the tail vertebrae) help to define Theropoda as a monophyletic group, which also includes birds. All nonavian theropods were bipedal, and the majority were carnivorous, as evidenced by the presence of recurved teeth with fine, steak knife–like serrations, grasping hands, and large claws on both hands and feet. A few Cretaceous forms (such as therizinosaurs and ornithomimosaurs) may have been omnivores or herbivores. Some authors regard *Eoraptor* and *Herrerasaurus* as the earliest theropods, though others argue that these animals are more primitive saurischians. Nevertheless, it is generally accepted that the earliest definite theropods (for example, *Coelophysis*) appeared in the Late Triassic and quickly radiated to become the most abundant and diverse predators of Late Mesozoic terrestrial ecosystems. Soon after its origin the group achieved a global distribution. Nonavian theropods survived until the end of the Cretaceous Period; avian theropods survived the Cretaceous/Tertiary (K/T) extinction event and, therefore, represent the only group of living dinosaurs.

Theropods ranged in size from animals less than a meter in length (for example, *Microraptor*), which were probably largely insectivorous, to gigantic superpredators (such as *Tyrannosaurus rex*, *Giganotosaurus*, and *Spinosaurus*) over 12 m (39 ft) long that weighed up to 6.5 metric tons (7.2 tons) and preyed upon other dinosaurs. A few theropods, such as *Deinonychus*, are believed to have been pack-hunters.

The dominant theme in theropod evolution relates to their transformation from scaly, fast-running, terrestrial bipeds to flying birds, with the subsequent transformation of forelimbs into wings and many other modifications associated with flight, including changes in brain anatomy, modifications to the hindlimbs and pelvic girdle, and the origin of feathers. The evolution of feathers is now known to have occurred in nonavian theropods, with filamentous "protofeathers" (as in *Sinosauropteryx*)

giving rise to true feathers (with vanes and barbs, as in *Sinornithosaurus*) to asymmetrical flight feathers (as in *Archaeopteryx*, the earliest-known bird). It is likely that feathers initially appeared for intraspecific display, or egg brooding or insulation, and were later co-opted for flight. *See* AVES; FEATHER.

Paul M. Barrett

Bibliography. P. J. Currie, and K. Padian (eds.), *The Encyclopedia of Dinosaurs*, Academic Press, San Diego, 1997; J. O. Farlow and M. K. Brett-Surman (eds.), *The Complete Dinosaur*, Indiana University Press, Bloomington, 1997; D. B. Weishampel, P. Dodson, and H. Osmólska (eds.), *The Dinosauria*, 2d ed., University of California Press, Berkeley, 2004.

Sauropterygia

An infraclass of Mesozoic reptiles that are, without exception, adapted to the marine environment. The infraclass includes the nothosaurs, plesiosaurs, and placodonts. These reptiles, along with the ichthyosaurs, played a significant role as predators within the marine animal community of the Mesozoic Era.

Placodontia. The placodonts are a distinctive but highly varied assemblage of aquatic reptiles. They had short bodies, paddlelike limbs, and flat cheek teeth designed for crushing hard-shelled prey. The genus *Helodus* was covered by a dorsal bony armor and a roofing of dermal scutes, but most other genera lacked armor. Placodonts have come only from rocks of the Middle and Upper Triassic of Europe, North Africa, and the Middle East.

Nothosauria. The Nothosauria are the relatively generalized stem group from which the plesiosaurs evolved. With the exception of a single New World species and a record from Japan, nothosaurs are known primarily from Europe and the Near East (Israel) in rocks of Triassic age. The nothosaurs are notably diverse in the mode and degree of secondary aquatic modification. The early phase of this process manifests itself (1) by a change in the histological structure of the bones, characterized by a total absence of Haversian systems and often a swollen appearance of the bones (pachyostosis); (2) by a notable reduction in size of the lateral girdle bones, scapulae (shoulder), and ilia (pelvis); and (3) by enlargement of the ventral elements and by the loss of digital claws. The directions of aquatic specialization involve shortening, or more often lengthening, of the neck (see **illus.**); enlargement of the orbits or the temporal fenestrae; and reduction, or more commonly increase, in the number of phalanges in manus (hand), pes (foot), or both. A feature of considerable evolutionary significance in the light of plesiosaurian differentiation is the great individual variability in the number of presacral vertebrae (32–42) in *Pachypleurosaurus edwardsi*. *See* SKELETAL SYSTEM.

Although most genera of nothosaurs are not in direct line of ancestry of the pleisosaurs, such a relationship has been suggested for *Pistosaurus*. The North American form, *Corosaurus alcovensis*, is probably the most advanced nothosaur known. *See* NOTHOSAURIA.

Plesiosauria. The Plesiosauria are the successful, compact, and highly specialized offshoot of the nothosaurs that attained worldwide distribution. The early steps of aquatic adaptation, initiated by the nothosaurs, have led to extensive anatomical modifications: The region comprising chest and abdomen became short, stout, and inflexible; the ventral bones of shoulder girdle and pelvis increased in area enormously; and the limbs, transformed into large flippers, became the principal organs of propulsion.

Two major trends of plesiosaur evolution may be discerned since the Early Jurassic: In the one group there was a tendency toward a shortening of the neck from 27 to 13 vertebrae and an increase in skull size: in the other group the opposite trend led to forms of bizarre body proportions, for example, *Elasmosaurus*, with a neck containing 76 vertebrae. The plesiosaurs were carnivorous. The precise affinities of placodonts, nothosaurs, and plesiosaurs remain uncertain. No earlier group of reptiles from which they might have come is known, and they left no descendants. *See* PLACODONTIA; PLESIOSAURIA; REPTILIA.

Rainer Zangerl; Everett C. Olson

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1987.

Savanna

The term savanna was originally used to describe a tropical grassland with more or less scattered dense tree areas. This vegetation type is very abundant in tropical and subtropical areas, primarily because of

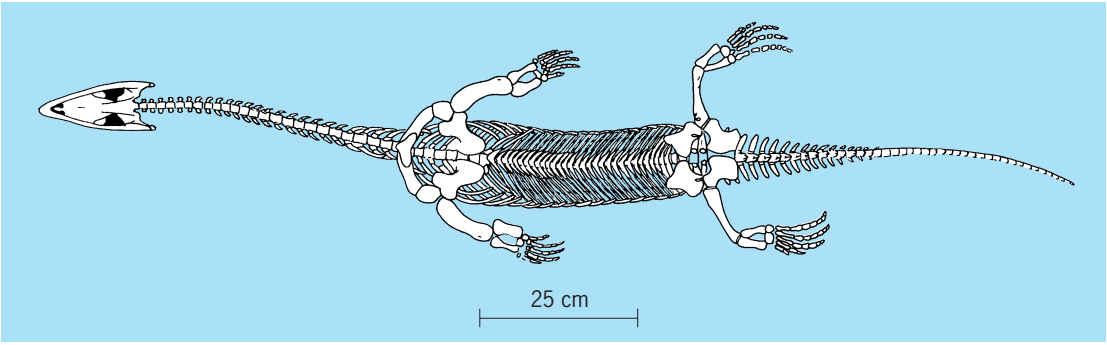

***Ceresiosaurus calcagnii*, ventral view.**

climatic factors. The modern definition of savanna includes a variety of physiognomically or environmentally similar vegetation types in tropical and extratropical regions.

In the widest sense savanna includes a range of vegetable zones from tropical savannas with vegetation types such as the savanna woodlands to tropical grassland and thornbush. In the extratropical regions it includes the "temperate" and "cold savanna" vegetation types known under such names as taiga, forest tundra, or glades. The physiognomically savannalike extratropical vegetation types differ greatly in environment and species composition.

During the growing season the typical tropical savanna displays a short-to-tall, green-to-silvery shiny cover of bunch grasses, with either single trees or groups of trees widely scattered. This is followed by a rest period of several months during which, because of severe drought, the vegetation appears quite different, with the brown-gray dead grasses bent over and the trees either without leaves or with stiff or wilted gray-green foliage. The heat and drought during this season of the year exert a high selective pressure upon the floral and faunal composition of the savanna.

Floral and faunal composition. The physiognomic similarity of the tropical savannas is underlined by the similarity among certain floristic components. All savannas contain members of the grass family (Gramineae) in the herbaceous layer. Most savannas of the world also have one or more members of the tree family (Leguminosae), particularly of the genus *Acacia*. Also included among the trees are the families Bombacaceae, Bignoniaceae, Palmae, and Dilleniaceae, and the genera *Prosopis* and *Eucalyptus*; these are abundant when they occur. One of the most outstanding savanna trees is *Adansonia digitata* (Bombacaceae), which achieves one of the biggest trunk diameters known for all trees (**Fig. 1**). The grass species, although mostly from the genera *Panicum*, *Paspalum*, and *Andropogon*, include numerous other genera such as *Aristida*, *Eragrostis*, *Schmidtia*, *Trachypogon*, *Axonopus*, *Triodia*, and *Plectrachne*, all of regional importance.

Fig. 1. ***Adansonia digitata*** **of Senegal. (*Courtesy of H. Lieth*)**

The fauna of the savannas is among the most interesting in the world. Savannas shelter herds of mammals such as the genera *Antelopus*, *Gazellus*, and *Giraffus*, and the African savannas are especially famous for their enormous species diversity, including various members of the Felidae (for example, the lion), and the elephant.

Numerous species of birds are indigenous to the savannas. Among these is the biggest bird, the ostrich (*Struthio camelus*), found in Africa. Many birds from extratropical regions migrate into the tropical savannas when the unfavorable season occurs.

Among the lower animals, the ants and termites are most abundant. Termite colonies erect large, conical nests above the ground, which are so prominent in some savannas that they partly dominate the view of the landscape, especially during the dry season.

Environmental conditions. The climate of the tropical savannas is marked by high temperatures with more or less seasonal fluctuations. Temperatures rarely fall below 32°F (0°C). The most characteristic climatic feature, however, is the seasonal rainfall, which usually comes during the 3–5 months of the astronomic summertime. Nearly all savannas are in regions with average annual temperatures from 59 to 77°F (15 to 25°C) and an annual rainfall of 32 in. (81 cm).

The soil under tropical savannas shows a diversity similar to that known from other semiarid regions. Black soils, mostly "chernozem," are common in the moister regions. Hardpans and occasional surface salinities are also found, and lateritic conglomerations occur along the rivers. The soils vary in mineral nutrient level, depending on geologic age, climate, and parent material. Deficiencies of minerals, specifically trace elements, as well as aluminum toxicity, are reported from many grassland areas in the continents of Africa, South America, and Australia. *See* SOIL.

Certain savanna areas suffer from severe erosion and no soil can be accumulated. Plant growth in these areas is scarce, often depending on cracks and crevices in the ground material to support tree roots. The herbaceous cover provides only a thin cover for the otherwise rocky surface.

Water is the main limiting environmental factor in the majority of the tropical savannas. Total amount and seasonality of precipitation are unfavorable for tree growth. Additional stresses to forest vegetation are caused by frequent fires, normal activities of animals, excess of salt, or nutrient deficiencies. Wherever a river flows, most of these factors change in favor of tree growth. This explains the existence of extensive gallery forests along the rivers and creeks (**Fig. 2**). The gallery forest is missing only where severe local floods after storms cause soil erosion and the formation of canyons.

Geographic distribution. Tropical savannas exist between the areas of the tropical forests and deserts; this is most apparent in Africa and Australia. In the

Fig. 2. Aerial view of savanna with gallery forest north of Guiaba, Brazil. (*Courtesy of H. Lieth*)

New World tropics, different conditions exist because of the circumstances created by the continental relief, and the savannas are situated between tropical forests and mountain ranges. In Madagascar and India there is a combination of both conditions.

The transitional (ecotonal) position of the savannas between forest and grassland or semidesert is the basis for the differences in opinion among authors about the size and geographical distribution of savannas. Most authors include, however, the savanna in East Africa and the belt south of the Sahara, the bush veld in South Africa, the Llanos in northern South America, and some types of scrub vegetation in Australia. Some areas in southern Madagascar, on several tropical islands (in leeward position), in Central America, in southern North America, and in India are usually included. The two latter regions, however, are subtropical. Still other areas included in the savanna concept are the Campo Cerrado and parts of the Chaco in South America; portions of the Miombo in southern Africa; wide portions of northern Africa; south of the Sahara; and wide portions of Madagascar, the Indian peninsula, and Australia. Because of the variations in the savanna concept, it is difficult to give a correct estimate of the total surface area covered by savanna vegetation. The Food and Agriculture Organization (FAO) considers that about one-third of the total land surface is covered by predominantly grassland vegetation. Of this area, one-third can be assumed to be tropical grassland, most of which can be called savanna.

Agricultural practices. Most of the original savanna areas throughout the world are currently farmed. Ranch farming is the predominant type, with sheep being raised in the drier areas and cattle in the moister regions. In the hotter regions zebus are raised, along with several hybrids of zebus and European cattle, which are the preferred stock for this climate. The yield in meat per unit area is low, and even under extensive management it seems to be lower than the meat production of the natural animal herds of the savanna, including antelopes, giraffes, and zebras. Ranch farming does not change the character of the vegetation very much if it is well managed. The adjacent dry woodlands very often resume a physiognomy similar to the natural savannas if good farm management is applied.

Agricultural crops are of many varieties in the savanna areas, where with careful protection and management any crop can be cultivated, provided that enough water is available. Drought-tolerant crops are usually preferred among the perennials.

The majority of human settlements are small in the savanna regions of the world; the hot temperatures during part of or the entire year, together with problems of water supply, limit interest in larger settlements. Settlements are usually found along the rivers, close to the coast, or in the higher elevations. Much of the land is managed by small tribal villages or large plantation or ranch owners, with separate groups of tenants, sharecroppers or employees, or single families. *See* AGROECOSYSTEM.

Extratropical types. The main structural character of tropical savannas is the scattered trees standing within a close cover of herbaceous vegetation. This structural character is also found in several extratropical vegetation types, but these differ greatly in the forces that limit a close tree cover, including drought (areas intermediate between steppe or prairie and forest); excess water or water combined with soil that has a shortage of oxygen and nutrients (peatbogs, marshes, or glades); short vegetation periods because of extended cold temperatures below the freezing point; excessive, long snow covers; and low light intensity (forest tundra, taiga, and cold savanna).

An intermediate condition is exhibited by the savannas and everglades of the southeastern United States. The intermittent soaked or dry conditions of a peaty soil, the tropically hot summers and mild winters (with frost periods, however), and the generally low nutrient level of the soil give these areas the characteristics of the cold savannalike vegetation types and the tropical and subtropical types.

The economic potential of the three extratropical savannalike areas also varies greatly. The conditions in the southeastern United States allow orchards, tree plantations, and cattle ranching. Economic considerations control whether a given area should be developed.

Some savanna regions in many parts of the world are the last survival territories for many plant and animal species. This implies the need for conserving some savanna pieces, both tropical and extratropical. Some of the most interesting wildlife sanctuaries are set aside for this reason (such as Serengeti, Tanzania; Ruhung, Sri Lanka; and Krüger National Park, South Africa). *See* FOREST AND FORESTRY; GRASSLAND ECOSYSTEM; TAIGA; TUNDRA. Helmut Lieth

Bibliography. M. Cole, *The Savannas: Biogeography and Geobotany*, 1986; R. T. Coupland (ed.), *Grassland Ecosystems of the World*, 1979; P. M. Dansereau, *Biogeography*, 1957; A. Engler and O. Drude (eds.), *Vegetation der Erde*, 15 vols., 1976; L. T. Gutierrez and W. R. Fey, *Ecosystem Succession: A General Hypothesis and a Test Model of a Grassland*, 1980.

Savory

A herb of the mint family in the genus *Satureja*. There are more than 100 species, but only *S. hortensis* (summer savory) and *S. montana* (winter savory) are grown for flavoring purposes. A number of varieties of each species are available, although no varietal names have been assigned to them. *See* LAMIALES.

Summer savory, an annual herb with an upright growth habit reaching about 2.5 ft (0.75 m), is characterized by long thin wiry stems with long internodes between small leaves. Winter savory is a perennial in most climates and, unlike summer savory, will tolerate some freezing weather. It can become woody after one or two growing seasons.

Savory is indigenous to areas surrounding the Mediterranean Sea. *Satureja montana* occurs wild from North Africa to as far north as Russia but is little cultivated. *Satureja hortensis*, which is widely cultivated, is native to Europe.

Savory is a rapid grower that home gardeners can plant any time during the growing season up to 60 days before frost. Winter savory will resume growth in the spring, but summer savory must be replanted. Commercial cultivation employs direct seeding into the field at 10–20 lb/acre (9–18 kg/hectare). It has relatively large seeds, and germination of 60–80% can be expected. Both types of savory are harvested or cut two or three times a year, after which the leaves are dehydrated and separated from the stems to be used as a spice in various foods, including poultry seasoning and beans. *See* SPICE AND FLAVORING. Seth Kirby

Bibliography. S. Arctander, *Perfume and Flavor Materials of Natural Origin*, 1960; L. H. Bailey, *Manual of Cultivated Plants*, rev. ed., 1975.

Saxifragales

An order of flowering plants near the base of the eudicots; DNA sequence analysis has indicated that this order includes few families that have been associated together in previous classifications. The order's exact relationships to other orders are still under study, but Saxifragales appear to be related to the rosid eudicots. They include 13 small to moderately sized families and perhaps around 2300 species. There are two categories: woody species that are sometimes wind-pollinated, and insect-pollinated herbs. The members are difficult to characterize morphologically but fall into several sets of families that have been considered closely related by most authors. *See* MAGNOLIOPHYTA; MAGNOLIOPSIDA; POLLINATION; ROSIDAE.

The families that belong here were formerly considered to be in either Saxifragales in the broad sense or the largely wind-pollinated Hamamelidales. None of these are of particular economic importance except as ornamentals, although *Cercidiphyllum* (Cercidiphyllaceae), *Liquidambar* (sweet gums), *Altingia* (both Altingiaceae), and a few others produce timbers. The largest family is Crassulaceae (1500 species), which are plants of arid zones, including the stonecrops (*Sedum*) and kalanchoes (*Kalanchoe*). A number of species are commonly cultivated ornamentals, including peony (*Paeonia*, Paeoniaceae), saxifrages (*Saxifraga*, Saxifagaceae), and witch hazel (*Hamamelis*, Hamamelidaceae); a few, the currants and gooseberries (*Ribes*, Grossulariaceae), are cultivated as fruits. *See* HAMAMELIDALES; ORNAMENTAL PLANTS. Mark W. Chase

Scalar

A term synonymous in mathematics with "real" in real number or real function. The magnitude of a vector two units in length is the real number or scalar 2. The dot or scalar product of two vectors is the product of three real numbers associated with them (the magnitude of the first vector times the magnitude of the second, times the cosine of the angle between them) and is therefore a scalar. If in the functional relationships $S = S(s,y,z)$, $\mathbf{F} = \mathbf{F}(x,y,z)$, S is a real number and $\mathbf{F}$ is a vector, then $S(s,y,z)$ is a scalar function but $\mathbf{F}(x,y,z)$ is a vector function. More generally, if V is a vector space over a field F, the elements of F are often referred to as scalars. *See* CALCULUS OF VECTORS; LINEAR ALGEBRA. Homer V. Craig

Scale (music)

As a piece of music progresses, it typically outlines a set of pitches by repeatedly sounding a subset of all the possible notes. When these notes are rearranged into ascending or descending order, they are called a musical scale.

There is often a range of pitches that will be heard as the same. Perhaps the trumpet plays middle C a bit flat, while the guitar attacks a bit sharp, in accordance with the artistic requirements of the musical context. The mind hears both pitches as the same note, C, and the limits of acceptability are far cruder than the ear's powers of resolution. This suggests a

kind of categorical perception, where a continuum of possible stimuli (in this case, pitch) is perceived as consisting of a small number of disjoint classes. Thus scales partition pitch space into disjoint chunks. *See* MUSICAL INSTRUMENTS; PITCH; SENSATION.

Western tradition. In the modern Western tradition, scales are standardized subsets of the 12-tone equal temperament (abbreviated 12-tet, and also called the chromatic scale) in which each octave is divided into 12 (approximately) equal sounding divisions. These are further classified into major and minor depending on the exact ordering of the intervals, and are classified into modes depending on the starting point. Historically, however, scales based on 12-tet are fairly recent. *See* TUNING.

The ancient Greeks theorized about scales which are based on what they called pure or just intervals, those in which the fundamental frequencies of two tones are related by simple integer ratios such as 2:1 (the octave), 3:2 (the just fifth), and 4:3 (the just fourth). Pythagoras is often credited with designing the scale that contains the largest possible number of just fourths and fifths, though there are a variety of other ancient scales based on tetrachords (where a pure fourth is divided into three subintervals). Just intonation refers to a family of scales which utilize a variety of intervals based on small integer ratios. Harry Partch championed a scale with 43 unequal tones per octave, and created a family of instruments on which to perform his music. Other historical scales, such as the well temperaments and various meantone tunings, can be viewed as different ways of approximating the just intervals while attempting to maintain the ability to modulate between keys.

Worldwide variety. Other cultures use musical scales that are quite different. For example, traditional Thai music is played in a scale that is close to 7-tone equal temperament. Various Arab musics use scales with a large number of tones; one with 22 unequally spaced tones per octave is frequently cited. The gamelan orchestras of Indonesia consist of a large family of inharmonic metallophones that are tuned to either the 5-note slendro or the 7-tone pelog scales; neither lies close to 12-tet. Moreover, each gamelan is tuned differently, and tunings often have "octaves" that differ significantly from the standard 2:1 frequency ratio.

One explanation for the large variety of intervals and scales used throughout the world is cultural: each musical tradition has followed a unique path that reflects its own history. Another explanation observes that the overtones of string and wind instruments form a harmonic series, and that these overtones overlap best when two notes are played in simple integer ratios. This is one basis for the just intonation approach. Similarly, it has been observed that the inharmonic overtones of certain gamelan instruments overlap best when two tones are played in intervals of the slendro and pelog scales, suggesting that the instruments and scales of a musical tradition may have coevolved to their present state. *See* MUSICAL ACOUSTICS.

William A. Sethares

Bibliography. D. Deutsch (ed.), *The Psychology of Music*, 2d ed., Academic Press, San Diego, 1999; M. Harrison, *Contemporary Music Theory*, 2 vols., Hal Leonard Publishing, 1999; H. Partch, *Genesis of a Music*, 2d ed., Da Capo Press, New York, 1974; W. A. Sethares, *Tuning, Timbre, Spectrum, Scale*, Springer-Verlag, 1998.

Scale (zoology)

The fundamental unit of the primary scaled integument of vertebrates, of which the epidermal or dermal component may be the more conspicuous or elaborated. Scale may have different meanings in different contexts; the semantic confusion has arisen partly for historical reasons, partly because of inherent problems in understanding fossil material, and partly due to increased knowledge of the development of the integument and teeth.

If a scale is defined as an overlapping fold of the epidermal-dermal boundary of the body, it is clear that some vertebrates (for example, a herring or a snake) should be described as having a scaled integument. Similarly, portions of the integument of some vertebrates comprise similar units, such as the leg of a bird or the tail of a rat. However, comparable integumentary units are absent from vertebrates such as sharks, eels, amphibians, and over most of the body of birds and mammals. The integumentary organization of vertebrates is very diverse; to understand the evolutionary relationships between scaled and nonscaled integuments, it is necessary to combine data from paleontology and studies of integumentary morphogenesis.

The fossil record implies that the vertebrate dermis has always possessed the capacity to house bone-forming cells (osteocytes) which give rise to dermal bone by direct deposition of extracellular calcium salts, without a cartilaginous precursor. Much of the skull roof of all vertebrates is formed in this manner. Studies of amphibian and avian embryos have shown that these osteocytes derive from the neural crest. In amphibians it has been shown that the odontoblasts, which form the dentine of the teeth, also derive from the neural crest. Studies of mammalian tooth development demonstrate that odontoblasts, lying beside dermal bone, interact with the overlying buccal ectoderm which then gives rise to ameloblasts, the cells responsible for enamel formation. Because neither odontoblasts nor ameloblasts grown separately from one another are capable of synthesizing dentine or enamel, respectively, it is recognized that odontogenesis involves a complex series of interactions between mesenchyme of neural crest origin and ectoderm. The amount and precise histology of the enamel and dentine (dentomorphic tissues) in various mammalian teeth varies in different mammals. However, a wide variety of investigations has shown that odontogenesis is basically similar in all toothed vertebrates and also that integumentary histogenesis in many fishes is characterized by similar, if not identical, cellular activities. When the histology of the

integumentary sclerifications in Paleozoic fish fossils is examined, there seems little doubt that the odontogenic capacities now confined to the oral region in tetrapods once characterized the entire body surface of their piscine ancestors.

Fishes. There is no way of knowing the precise epidermal structure of fossils. Available data suggest that in the majority of extant fishes the epidermis is extremely thin, consisting largely of mucus-synthesizing cells. The term "scale" when applied to fishes usually refers to the relatively large prominent, dermal ossification, on the outer surface of which lie so-called dentomorphic tissues. However, these sclerifications either may be parts of overlapping units in a scaled integument or may be discrete integumentary appendages where overlapping units are absent.

A full complement of bone plus both types of dentomorphic tissue characterizes the integumentary sclerifications of many "primitive" fishes: many jawless ostracoderms, the acanthodians, and certain bony fishes. The dentomorphic tissues may be seen as tuberculate units protruding irregularly from the surface of the sclerifications as in ostracoderms (**Fig. 1**) or as flattened sheets in acanthodians (**Fig. 2**) or certain bony fishes. The bony tissues consist of deep laminated materials (basal layer) and a superficial spongy vascular region (middle layer). Growth of the integumentary sclerifications involves periodic resorption and redeposition of the dental tissues, as well as the addition of laminate bone. *See* ACANTHODII; OSTRACODERM.

The thelodont ostracoderms and contemporary sharks, two completely unrelated lineages, do not have overlapping scales. Their body surfaces are covered by "dermal denticles" which strongly resemble small teeth. As such, they are true integumentary appendages, although they are frequently referred to as placoid scales.

The evolutionary history of dermal sclerifications varied in different lineages. In general, the tendency is to reduce or lose various of the fundamental constituents. Thus sharks have no bony constituents in their "dermal denticles," while in teleosts only a thin layer of acellular, bonelike material is seen in the dermis. A variety of terms describing "fish scales" appear in the literature; some, for example, cosmoid and ganoid, refer to the detailed histological structure (in sarcopterygians and primitive actinopterygians respectively; **Fig. 3**), while others, for example, rhomboid, cycloid, or ctenoid (**Fig. 4**), refer to the overall shape of the units.

The scaled integument of fishes shows a definite pattern which is governed by the orientation of the underlying myomeres and by changes in body depth and length. The general evolutionary tendency

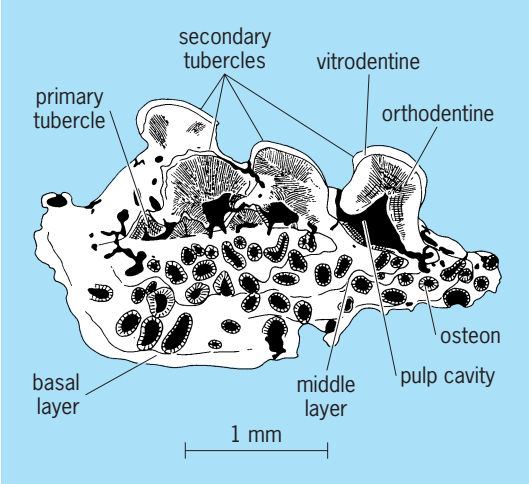

Fig. 1. Cross section of dermal plate of Ordovician ostracoderm. (*After R. H. Denison, The early history of the vertebrate calcified skeleton, Clin. Orthop., no. 31, 143, 1963*)

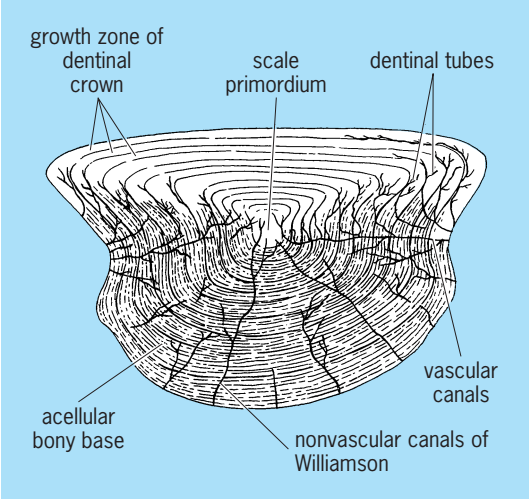

Fig. 2. Cross section of Permian acanthodian scale, showing structural details.

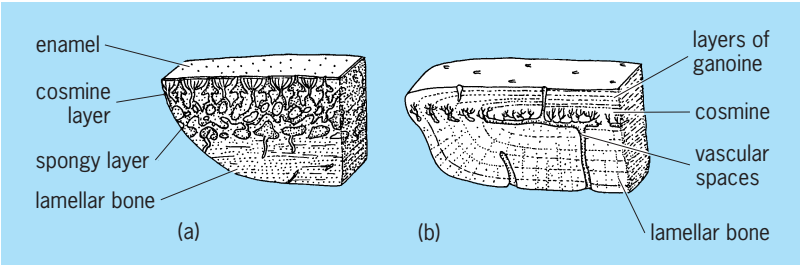

Fig. 3. Cross sections of primitive vertebrate scales. (*a*) Cosmoid scale. (*b*) Ganoid scale. (*After A. S. Romer, The Vertebrate Body, Saunders, 1962*)

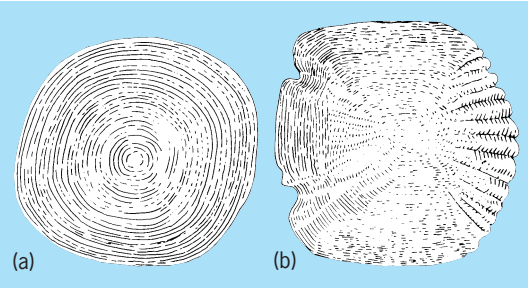

Fig. 4. Scale types. (*a*) Cycloidal scale from burbot. (*b*) Ctenoid scale from carp. (*After K. F. Lagler et al., Ichthyology, John Wiley and Sons, 1962*)

toward thinning, or size reduction, of integumentary sclerifications seen in various fish lineages is probably associated with weight reduction and increased locomotor efficiency.

Tetrapods. Scaled integuments are absent from modern amphibians, although nonoverlapping dermal ossifications are seen in some apodans. It seems probable that the absence of scales in other modern species is associated with the secondary utilization of the integument as a respiratory surface. While direct evidence for the presence of patterned arrangements of dermal ossification is available from a variety of Paleozoic amphibian genera, suggesting a scaled integument, it is uncertain if the overlying epidermis was heavily cornified, a feature which might have been expected in terrestrial forms. Certain chemical features of the matrix surrounding the fossilized remains of the lower Triassic *Uranocentrodon* have been interpreted as indicating a localized decomposition of sulfur-containing epidermal proteins.

While dermal ossifications are known from many early reptiles (captorhinomorphs), it cannot be stated with certainty whether these were retained from their amphibian ancestors or independently evolved. This doubt springs from the fact that in the well-documented lineages of captorhinomorph to therapsid to mammal, there is early reduction and loss of dermal ossifications (suggesting a loss of a scaled integument), and yet numerous fossil and extant mammals, for example, ground sloths, glyptodonts, and armadillos, have very well-developed dermal ossifications over much of the body. The characteristic mammalian hairs did not evolve from scales, but their patterned arrangement has been interpreted as indicating an evolutionary origin associated with developmental fields surrounding scales in the ancestral reptiles. In general, it would appear that the evolutionary history of the mammalian integument involved early loss of scales, concomitant with increasing reliance on initially a thickened epidermis and later a pelage, for both physiological and mechanical protection. *See* HAIR.

In all other reptilian lineages, data from fossils and from study of living species suggest a ubiquity of scaled integuments, with or without dermal ossifications, but always with elaboration of epidermal tissues. This fact has led to the formulation of the generic terms "reptilian" and "epidermal" scales, both of which are best avoided. The distribution of keratinaceous proteins is highly variable in both reptilian and avian scales. Because modern crocodilian scales exactly resemble the scutate (overlapping) units on a bird's leg, an excellent case can be made for the direct evolution of feathers from archosaurian scales. However, the reticulate scales on the plantar surface of the avian foot do not overlap, and their epidermal cells synthesize alpha rather than beta keratin.

Birds, of course, totally lack the dermal ossifications which reached such impressive proportions in their Mesozoic archosaurian cousins such as *Ankylosaurus* and *Stegosaurus*. Scaled integuments prompted the subclass name for lepidosaurs (*lepidos* is Greek for flake or scale)—the tuatara, lizards, and snakes, the last two constituting the order Squamata (*squama* is Latin for scale).

Lepidosaurian scales may or may not possess dermal ossifications. Among extant reptiles the extraordinary development of the dermal skeleton (obviously associated with epidermal patterning) in turtles forms the characteristic carapace, a remarkable parallel with the mammalian glyptodonts, but the carapace is reduced or lost (along with visible scalation) in a variety of aquatic forms.

P. F. A. Maderson

Bibliography. L. B. Halstead and R. Hill, *Vertebrate Hard Tissues*, 1974; E. Jarvik, *Basic Structure and Evolution of Vertebrates*, 2 vols., 1980.

Scandium

A chemical element, Sc, atomic number 21, atomic weight 44.956. The only naturally occurring isotope is ^{45}Sc. The electronic configuration of the ground-state, gaseous atom consists of the argon rare-gas core plus three more electrons in the $3d14s2$ levels. It has an unfilled inner shell (only one $3d$ electron) and is the first transition metal. It is one of the elements of the rare-earth group. *See* PERIODIC TABLE; RARE-EARTH ELEMENTS; TRANSITION ELEMENTS.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

The principal raw materials for the commercial production of scandium are uranium and tungsten tailings and slags from tin smelters or blast furnaces used in cast iron production. Wolframite (WO_3) concentrates contain 500–800 ppm scandium. *See* WOLFRAMITE.

Scandium is the least understood of the $3d$ metals. The major reason has been the unavailability of high-purity scandium metal, especially with respect to iron impurities. Many of the physical properties reported in the literature vary considerably, but the availability of electrotransport-purified scandium has allowed measurement of the intrinsic properties of scandium (see **table**).

Scandium increases the strength of aluminum. It also strengthens magnesium alloys when added to magnesium together with silver, cadmium, or yttrium. Scandium inhibits the oxidation of the light

Room-temperature properties of scandium metal (unless otherwise specified)

Property	Value
Atomic number	21
Atomic weight	44.9559 ($^{12}C = 12$)
Lattice constant (hcp, α–Sc), a_0	0.33088 nm
c_0	0.52680 nm
Density	2.989 g/cm^3
Metallic radius	0.16406 nm
Atomic volume	15.041 cm^3/mol
Transformation point	1337°C (2439°F)
Melting point	1541°C (2806°F)
Boiling point	2836°C (5137°F)
Heat capacity	25.51 J/mol K
Standard entropy, $S^{\circ}_{298.15}$	34.78 J/mol K
Heat of transformation	4.01 kJ/mol
Heat of fusion	14.10 kJ/mol
Heat of sublimation (at 298 K)	377.8 kJ/mol
Debye temperature (at 0 K)	345.3 K
Electronic specific heat constant	10.334 mJ/mol K^2
Magnetic susceptibility, χ_A^{298} (*a*)	297.6×10^{-6} emu/mol
χ_A^{298} (*c*)	288.6×10^{-6} emu/mol
Electrical resistivity, ρ_a^{300}	70.90 μohm-cm
ρ_c^{300}	26.88 μohm-cm
Thermal expansion, $\alpha_{a,i}$	7.55×10^{-6}
$\alpha_{c,i}$	15.68×10^{-6}
Isothermal compressibility	17.8×10^{-12} m^2/N
Bulk modulus	5.67×10^{10} N/m^2
Young's modulus	7.52×10^{10} N/m^2
Shear modulus	2.94×10^{10} N/m^2
Poisson's ratio	0.279

rare earths and, if added along with molybdenum, inhibits the corrosion of zirconium alloys in high-pressure steam. The addition of ScC to TiC has been reported to form the second-hardest material known. Sc_2O_3 can be used in many other oxides to improve electrical conductivity, resistance to thermal shock, stability, and density. Scandium is used in the preparation of the laser material $Gd_3ScGa_4O_{12}$, gadolinium scandium gallium garnet (GSGG). This garnet when doped with both Cr^{3+} and Nd^{3+} ions is said to be $3^1/_2$ times as efficient as the widely used Nd^{3+}-doped yttrium aluminum garnet (YAG:Nd^{3+}) laser. Ferrites and garnets containing scandium are used in switches in computers; in magnetically controlled switches that modulate light passing through the garnet; and in microwave equipment. Scandium is used in high-intensity lights. Scandium iodide is added because of its broad emission spectrum. Bulbs with mercury, NaI, and ScI_3 produce a highly efficient light output of a color close to sunlight. This is especially important when televising presentations indoors or at night. When used with night displays, the bulbs give a natural daylight appearance. Scandium metal has been used as a neutron filter. It allows 2-keV neutrons to pass through, but stops other neutrons that have higher or lower energies. *See* YTTRIUM. Jennings Capellen; Karl A. Gschneidner, Jr.

Bibliography. E. W. Abel et al., *Comprehensive Organometallic Chemistry II: Scandium, Yttrium, Lanthanides and Actinides, and Titanium, Zirconium, and Hafnium*, Pergamon Press, 1995; C. T. Horovitz et al., *Scandium, Its Occurrence, Chemistry, Physics, Metallurgy, Biology and Technology*, Academic Press, 1975; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004.

Scanning electron microscope

An electron microscope that builds up its image as a time sequence of points in a manner similar to that employed in television. While the idea of the scanning electron microscope (SEM) dates back to the 1930s, a great deal of the pioneering development of the instrument took place at Cambridge University, England, during the years following World War II. The SEM became commercially available in 1966 and has rapidly taken its place as a useful addition to the battery of electron optical instruments used in both biological and physical research. A part of the attractiveness of the SEM lies in its ability to utilize both analytic and subjective or intuitive channels in attempts to understand the microscopic world.

Imaging method. The imaging method of the scanning electron microscope allows separation of the two functions of a microscope, localization and information transfer. The SEM utilizes a very fine probing beam of electrons which sweeps over the specimen and at each instant excites the material of the specimen to emit a variety of radiations (**Fig. 1**). The signal, which is proportional to the amount of radiation leaving an individual point of the specimen at any instant, can be used to modulate the brightness of the beam of the display cathode-ray tube as it rests on the corresponding point of the image. When the radiation from the point on the specimen is high, the intensity of the point on the display cathode-ray tube is also high. If the probing beam is then moved to an adjacent spot on the specimen, information from that point can be used to modulate the corresponding point on the display cathode-ray

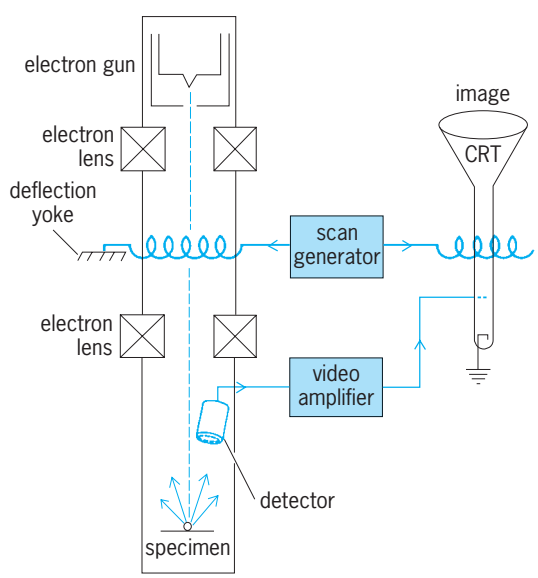

Fig. 1. Schematic diagram of scanning electron microscope showing twin-electron-beam imaging system and some of the radiations used for the information signal.

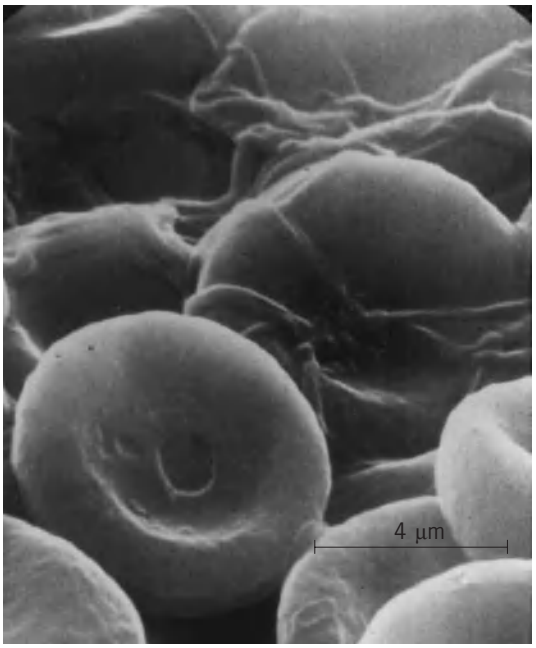

Fig. 2. Human blood clot. Red blood cells are held in fibrin network. (*From L. W. McDonald and T. L. Hayes, Correlation of scanning electron microscope and light microscope images of individual cells in human blood and blood clots, Exp. Molec. Pathol., 10:186–198, 1969*)

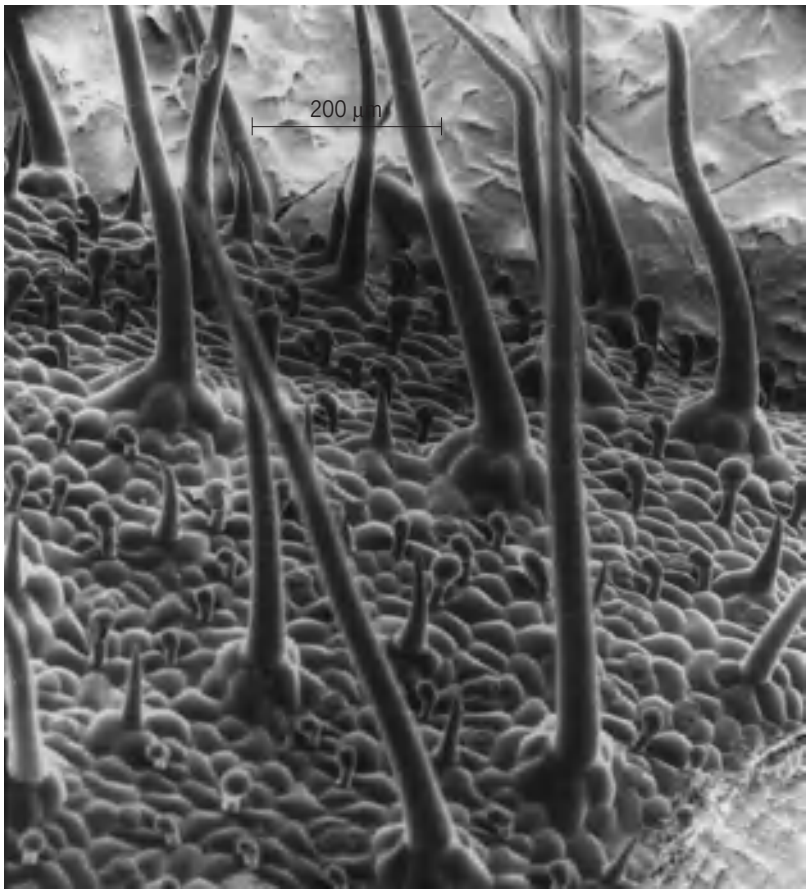

Fig. 3. Leaf surface viewed while held at −238°F (−150°C) in the scanning electron microscope. Specimen is surrounded by ice field and is still hydrated. (*From T. L. Hayes and G. Koch, Some problems associated with low temperature micromanipulation in the SEM, Proceedings of the Scanning Electron Microscopy Symposium, pp. 35–42, 1975*)

tube. In a similar fashion a third and fourth point can be added in sequence, and so on until all parts of the specimen have been covered in a regular array. The one-to-one correspondence that is necessary for the formation of an image occurs because a single position of the probing beam corresponds to a single position of the display cathode-ray-tube beam at each instant in time. In practice, the points follow one another with great rapidity so that the image of each point becomes an image of a line, and the line in turn can move down the screen so rapidly that the human eye sees a complete image as in television. The image can also be recorded in its entirety by allowing the point-by-point information to build up in sequence on a photographic film.

Both high resolution and a variety of kinds of information can be achieved with this system. Albert Crewe developed ultrahigh resolution scanning transmission electron microscopes (STEM) capable of resolving single atoms. Secondary electrons, backscatter electrons, characteristic x-rays, visible light (cathodoluminescence) and specimen-induced current have been used by investigators as SEM signals to provide information-rich images that relate to the size, shape, and chemical and electrical properties of the specimen.

Sensory codes. In addition to a very powerful array of objective analytic information, the SEM offers possibilities for use as an instrument to extend the human senses. Operated in the secondary electron detection mode, the SEM mimics some of the codes that people use to assess depth and shape in the large macroscopic world of everyday life. Since this ability to experience the microsystem through the use of an SEM permits the application of synthetic, intuitive processing, perhaps subjective channels can be added to enhance understanding of the classical disciplines of physics, chemistry, mathematics, or biology. Some of the fascination of the SEM may be related to its ability to provide intuitive contact with the microspecimen (**Fig. 2**).

Gun intensity. Since the resolution of the SEM is controlled in part by the brightness of the electron gun, experimental designs which increase gun intensity have been pursued for some years. Crewe, using a field emission source to increase the gun intensity, has produced instruments with resolutions well below 0.5 nanometer. Bimacromolecular staining as developed by Michael Beer has enhanced the contrast differences available and has resulted in a specific location of heavy-atom components within a macromolecule. In addition to the field emission gun, lanthanum hexaboride guns developed by Alec Broers provide greater brightness than the standard heated filaments.

Specimen preparation. As with all microscopy research, it is very often the preparative methodology that determines the success or failure of the research. A specific requirement of scanning electron microscope preparation is that the material be dried. In conventional electron microscopy or light microscopy, the specimen is carried through a

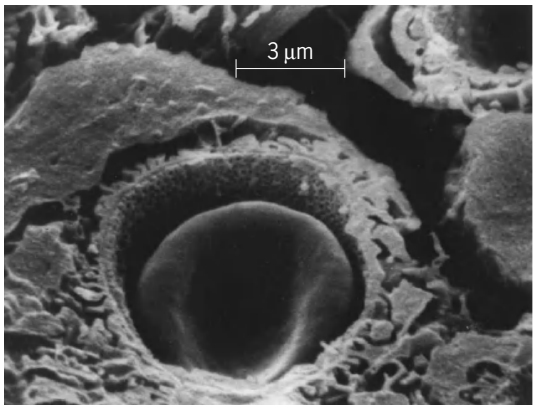

Fig. 4. Red blood cell in a capillary of the kidney. Ethanol cryofracture technique was followed by critical-point drying. (*From W. J. Humphreys, B. O. Spurdock, and J. S. Johnson, Critical point drying of ethanol-infiltrated, cryofractured biological specimens for scanning electron microscopy, Proceedings of the Scanning Electron Microscopy Symposium, pp. 275–282, 1974*)

series of liquid stains, fixatives, dehydrating agents, and embedding materials, but is never dried. In scanning microscopy, on the other hand, the material must be dried, and this preparative step must be done very carefully to minimize surface tension effects. The two methods most often used—critical-point drying and freeze drying—have yielded excellent results even for very fragile biological tissue. Alternatively, the drying step can be avoided completely by observing the specimen while it is still frozen (**Fig. 3**). Interior structure can be revealed by using techniques in which the specimen is broken at a low temperature (**Fig. 4**). *See* MICROTECHNIQUE.

Channels of information transfer. In microscopy there has always been a very useful aspect of image perception associated with the observer's intuitive or subjective contact in "seeing" the specimen. It is this kind of contact that has been particularly valuable when the specimen under study becomes extremely complex and difficult to analyze by strictly objective approaches. The SEM provides both objective and experiential channels of information transfer. In summary, the SEM is an excellent tool for analysis. It also provides valuable channels for human contact with the microworld through extension of the human sense. *See* ELECTRON MICROSCOPE; MICROSCOPE.

Thomas L. Hayes

Bibliography. W. R. Bottoms, *Scanning Electron Microscopy and X-Ray Microanalysis: A Text for Biologists, Materials Scientists, and Geologists*, 2d ed., 1992; P. Echlin and D. E. Newbury, *Scanning Electron Microscopy*, 1977; J. I. Goldstein and H. Yakowitz (eds.), *Practical Scanning Electron Microscopy: Electron and Ion Microprobe Analysis*, 1975; O. Johari and R. P. Becker (eds.), *Scanning Electron Microscopy*, 14th ed., 1981; J. Ohnsorge and R. Holm, *Scanning Electron Microscopy*, 2d ed., 1978; L. Reimer, *Image Formation in Low-Voltage Scanning Electron Microscopy*, 1993.

Scanning tunneling microscope

A device for imaging the surface of conductors and semiconductors with atomic resolution. Due to its relatively simple construction and operation and its ability to achieve atomic resolution with relative ease, the scanning tunneling microscope (STM) has gained worldwide acceptance by scientists studying surface phenomena.

Principles of operation. The operating principles of the scanning tunneling microscope are illustrated in **Fig. 1**. A wire made of tungsten, gold, or some other noble metal is etched electrochemically at one end down to a fine tip whose radius is 0.1 micrometer, or less; it is then attached onto a piezoelectrically scanned (x,y,z) stage. Under a powerful microscope, such as a transmission electron microscope, the tip of the wire typically looks as illustrated in Fig. 1, with one atom always protruding farther from the surface. The piezoelectric tripod can be scanned in x, y, or z over distances of up to 100 μm with a resolution far better than 0.1 nanometer (approximately one-half the diameter of typical atoms) by applying a voltage across its x, y, or z electrode, respectively. A voltage V in the range of 10 mV is applied between the wire tip and the conducting sample to be imaged, and the piezoelectric z-scanner is used to move the tip gently toward the sample. As the end atom on the tip approaches to within a few atomic diameters of a sample atom, electrons can tunnel from the tip to the sample (or vice versa) through the potential barrier imposed by the work functions of the tip and sample, thereby generating a current, which is typically of the order of nanoamperes. *See* ELECTRON MICROSCOPE; WORK FUNCTION (THERMODYNAMICS).

The tunnel current I_t increases exponentially with decreasing spacing z between tip and sample atom.

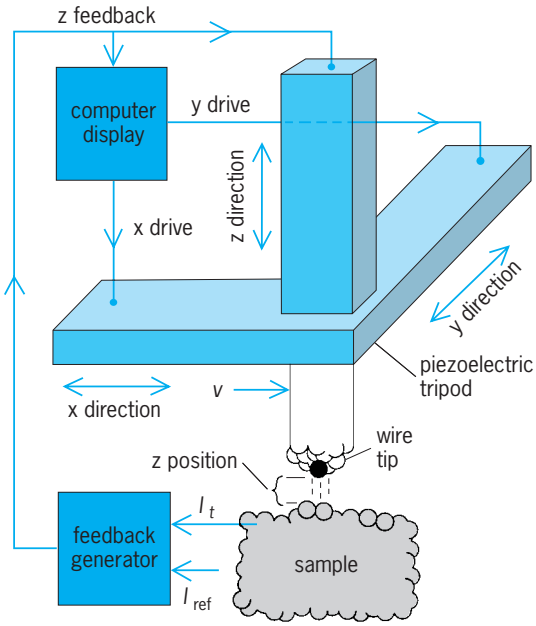

Fig. 1. Schematic diagram of the scanning tunneling microscope showing the configuration of the piezoelectric tripod, wire tip, and sample.

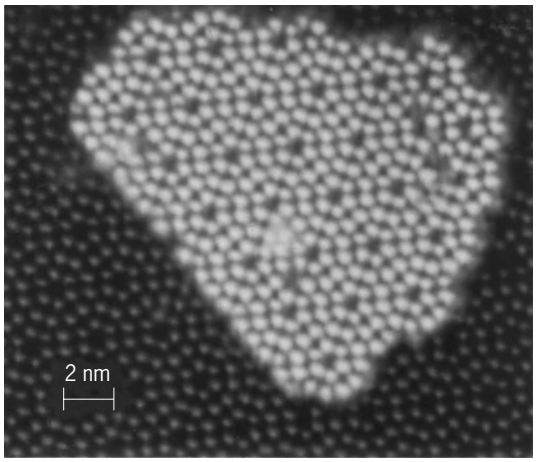

Fig. 2. Silicon ⟨111⟩ surface as imaged with the scanning tunneling microscope. The image extends over a lateral area whose dimensions are 23 × 20 nm. The image shows the (7 × 7) reconstruction of the unit cell on the crystal surface together with the registration of an epitaxially grown single layer. (*Courtesy of U. Kohler, J. E. Demuth, and R. J. Hamers*)

For small applied bias voltages υ, the tunnel current is given by the equation below, where ϕ is the mean

$$I_t = f(\upsilon)\exp -(A\phi^{1/2}z)$$

of the work functions of the tip and sample, A is a constant equal to $(4\pi/h)(2m)^{1/2}$, m is the electronic mass, h is Planck's constant, and $F(\upsilon)$ is the overlap integral of the local density of states of tip and sample taken over the energies between the Fermi levels of the tip and the sample. For a typical tip–sample combination, the tunnel current I_t decreases by a factor of 10 for every 0.1 nm increase in spacing z. For this reason, a negligible contribution to the tunnel current I_t is provided by electron tunneling from other atoms on the tip since the other atoms lie farther from the sample than the atom that protrudes from the tip.

In the scanning tunnel microscope, the tunnel current I_t is compared with a constant reference current (I_{ref}); the error signal so generated is used to move the z-piezoelectric scanner up or down (thereby increasing or decreasing z) in order to maintain a constant tunnel current as the tip is raster-scanned in x and y to record an image. The tunneling image is usually recorded on a computer as variations in the z position (measured through the z-piezoelectric voltage) corresponding to each (x,y) coordinate of the tip. For metallic surfaces such as gold, the local density of states is a weak function of the applied bias voltage υ, and the overlap integral $f(\upsilon)$ has no structure; it varies linearly with voltage, yielding a voltage-independent tunneling conductance. In these cases the variations in z closely approximate the atomic topography of the surface. However, for more complex surfaces, $f(\upsilon)$ is no longer a linear function of υ and furthermore can vary from one (x,y) location on the sample to the next, depending on the particular atom being probed. Similarly, the mean work function ϕ could also vary as a function of (x,y) on the sample. It can be shown that local-density-of-states spectroscopy can be performed on the sample by recording $(d \ln I_t)/(d \ln V)$ as a function of V. If local barrier height is desired, $A\phi^{1/2}$ can be measured by recording $(d \ln I_t)/dz$. Such measurements are useful in differentiating atomic features on the surface being imaged. *See* COMPUTER GRAPHICS; PIEZOELECTRICITY; TUNNELING IN SOLIDS.

Applications and extensions. One application of the scanning tunneling microscope is in the mapping of the atomic reconstructions of crystal surfaces. The image in **Fig. 2** demonstrates the potential of the scanning tunneling microscope for studying nucleation and growth of thin films. The brighter region in the center of the image is an epitaxially grown single layer of silicon ⟨111⟩ that shows the near-perfect registration with the substrate.

Other areas of application of the scanning tunneling microscope are in the study of electrochemical processes and biological imaging. It is also being used for surface modifications on an atomic scale.

The success of the scanning tunneling microscope has led to the development of many novel scanned probe microscopes that use the same piezoelectric scanning and feedback principles but rely on different mechanisms of tip–sample interaction to generate the error signal that controls the tip–sample spacing and hence the image. One scanning probe microscope that has been particularly successful is the atomic force microscope. In this microscope (**Fig. 3**) the interaction mechanism sensed is the repulsive force between the charge clouds of the tip atom and the sample atoms. Because this interaction does not rely on a tunneling current, the atomic force microscope can image both conducting and insulating surfaces with atomic resolution. In the early versions of the atomic force microscope, a diamond tip was attached onto a gold foil, and the repulsive force between tip and sample was detected through

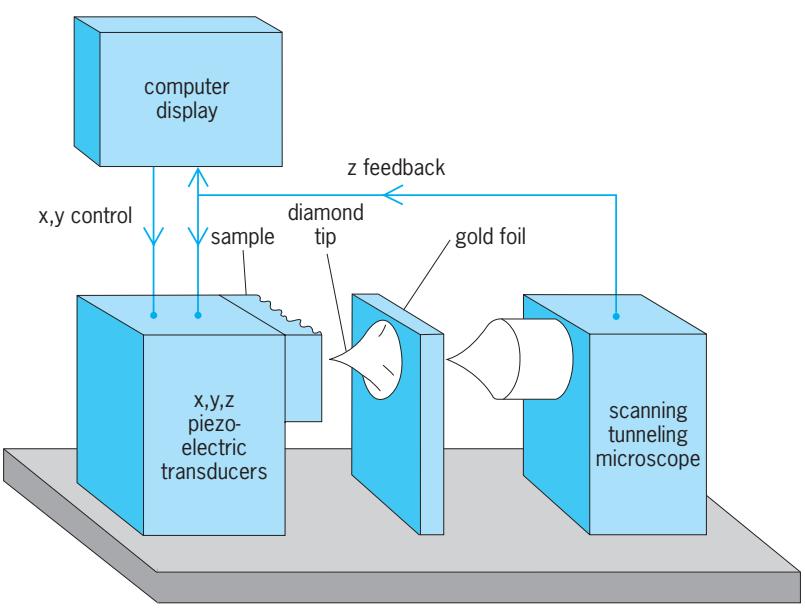

Fig. 3. Schematic diagram of the atomic force microscope.

the deflection of the gold foil, which in turn was sensed by using a scanning tunneling microscope. If the gold foil has a spring constant of 10 newtons/m, a force of 10^{-9} N will result in a deflection of 0.1 nm, which is easily detectable with a tunneling microscope. Later experiments have used silicon micromechanical cantilevers with integrated silicon tips and laser probes for sensing the deflection. Typical tracking forces are in the range of 10^{-7} N for instruments operating in air and 10^{-9} N under water. Applications include the imaging of the atomic arrangement of insulating crystal surfaces and the high-resolution imaging of biological surfaces and processes. Extensions of the atomic force microscope are the magnetic force microscope, which uses a magnetic tip to image magnetic surfaces, and the electrostatic force microscope, which measures the local dielectric polarizability of a sample through the attractive force caused by applying a voltage between the tip and the sample. *See* MICROSCOPE. H. K. Wickramasinghe

Bibliography. G. Binnig and H. Rohrer, The scanning tunneling microscope, *Sci. Amer.*, 253(8):50–56, 1985; G. Binnig and H. Rohrer, Scanning tunneling microscopy: From birth to adolescence, *Rev. Mod. Phys.*, 59(3):651–625, 1987; C. J. Chen, *Introduction to Scanning Tunneling Microscopy*, 1993; P. K. Hansma et al., Scanning tunneling microscopy and atomic force microscopy: Applications to biology and technology, *Science*, 242(4876):209–216, 1988; U. Kohler, J. E. Demuth, and R. J. Hamers, Scanning tunneling microscopy study of low-temperature epitaxial growth of silicon on Si(111)-(7 × 7), *J. Vac. Sci. Technol.*, A7(4):2860–2867, 1989; C. F. Quate, Vacuum tunneling: A new technique for microscopy, *Phys. Today*, 39(8):26–33, 1986; H. K. Wickramasinghe, Scanned probe microscopes, *Sci. Amer.*, 260(10):98–105, 1989.

Scaphopoda

A class in the phylum Mollusca, comprising two orders, Gadilida and Dentaliida. The class is wholly marine; it is common and probably most diverse in the deep sea, and poorly represented in brackish, littoral, or estuarine habitats. Scaphopods are characterized by a tapering tubular shell with two apertures: the large ventral aperture that the foot extends out of and a small dorsal aperture. The tapering shells found in the Dentaliida give the class its common names tusk or tooth shells. Adult scaphopods range in length from about 3 to 150 mm. The shell consists of aragonite arranged in three or four layers.

Classification. Systematically the class is related to bivalves, and probably arose from a member of the extinct molluscan class, the Rostroconchia. The two orders can be separated on the basis of shell structure. Typically gadilids have the widest part of the shell more toward the middle of the animal, while in the dentaliids the widest part of the shell is the ventral aperture. Gadilid shells are typically highly polished, while dentaliid shells lack this luster. In addition, there are significant internal morphological differences between the two orders.

Anatomical features. Scaphopods are bilaterally symmetrical and surrounded by the mantle, which forms a tube. The body is suspended in the tube from the dorsal-anterior part of the shell. The largest part of the mantle cavity is adjacent to the ventral aperture and contains the foot, head, and feeding tentacles, called captacula. Each captaculum consist of a long extendable stalk and a terminal bulb. The captacular bulb is covered with cilia, and the stalk may have a ciliated band or tufts of cilia along it. Captacula are hollow and have longitudinal muscles running from their base to the bulb.

The mantle cavity extends to the dorsal shell aperture along the ventral part of the body. Scaphopods lack the typical molluscan gill, the ctenidium. Respiration occurs through the mantle wall and is probably facilitated by the presence of large ciliated ridges found in the middle of the mantle cavity.

The foot is large and robust and capable of significant extension from the shell. In gadilids, where the foot is extended hydrostatically and retracted by inversion, the animals are often capable of rapid prolonged burrowing. The gadilid foot ends in a terminal disk. In dentaliids, where the foot is extended and retracted by the use of a muscular hydrostat and flexion, the animals move much more slowly. The dentaliid foot has a conical tip, and lateral lobes a bit proximal to the end.

The head, which is reduced and incapable of extension from the shell, bears a pair of lateral ridges and a proboscis which is tipped by the mouth surrounded by lips. The captacula arise from the lateral ridges; generally several hundred captacula are present. Scaphopods lack eyes and the specialized molluscan sensory organ, the osphradium.

The body is elongated dorsally and terminates internally near the dorsal shell aperture; however, the gut curves back on itself, and the anus is normally found at about the midpoint of the body and empties into the mantle cavity. The radula is mineralized and functions either to crush prey or to ratchet them into the rest of the gut through the esophagus. A single unmineralized jaw is anterior to the radula in the buccal region. There are five radular teeth per row. The stomach is capacious and is connected to one (in gadilids) or two (in dentaliids) digestive glands. The intestine is coiled, and the rectum is just proximal to the anus. There are rectal glands of unknown function near the anus.

The large dorsal retractor muscles pass laterally and ventrally to the gut. These originate dorsally near the aperture and extend to the base of the muscular dentaliid foot or to the tip of the hollow gadilid foot. They retract the foot or, when the foot is extended into sediments, pull the animal into the sediments.

The nervous system is relatively large and is built on the basic molluscan pattern. The largest ganglia are the paired cerebral and pedal ganglia; the latter bear statocysts. Other small ganglia are present, as are major nerves to the main body regions. Each captaculum also contains a small sensory ganglion.

The circulatory system consists of a series of blood

sinuses and channels. There is no heart that is homologous to other molluscan hearts, although contractile tissues are found near where a heart should be. These contractile elements cause fluid movement in the central blood sinus but are incapable of moving the colorless blood to the periphery of the animal. Most blood movement is due to the movement of the musculature of the foot. There are paired excretory organs, but no evidence of the urine formation by ultrafiltration that is typical of other mollusks.

The gonads are positioned dorsally and empty into the mantle cavity; sexes are separate. Early embryonic development and complete larval development have been observed in the dentaliids, but little is known about any aspect of embryonic or larval development in the gadilids.

Feeding mechanism. Scaphopods eat foraminiferans, and other small shelled organisms such as bivalves or shelled eggs. Dentaliids can also ingest sediment. Prey are captured by the use of the captacula, which have a chemical adhesive system to attach to prey. The foot is used to burrow and construct a feeding cavity. The captacula move out into the feeding cavity through the ventral aperture by the motion of the cilia covering the bulb. The cilia pull the captacular stalk along behind.

When a captaculum encounters an acceptable food item, it attaches to that item with a duo-gland adhesive system. After several other captacula attach, the longitudinal muscles within them contract, pulling the food item out of the feeding cavity and into the mantle cavity. Food is manipulated internally by the captacula and the lips. If the food item is acceptable, it is ingested and held in the proboscis prior to radular maceration. Up to 300 foraminiferan prey have been counted in the proboscis of one species of scaphopods.

Habitat. Scaphopods live in unconsolidated sediments, and most members of any population are typically completely buried at any one time. However, scaphopods back up to the surface and extend the dorsal shell apex to release eggs or sperm, and presumably to flush the mantle cavity with fresh water. They are vulnerable to predation during these periods, which are of limited duration.

In some localities along the Canadian Pacific coast, scaphopods are abundant and possibly constitute the major predators on organisms living in the sediments of the fiord bottoms. Scaphopods, in turn, are eaten by fishes, particularly chimerids, and crabs. *See* MOLLUSCA.

Ronald Shimek

Bibliography. F. Harrison and A. J. Kohn (eds.), *Mollusca II: Microscopic Anatomy of the Invertebrates*, vol. 6B, 1996; E. N. Kozloff, *Invertebrates*, 1990, reprint 1997; R. L. Shimek, Diet and habitat utilization in a Northeastern Pacific Ocean scaphopod assemblage, *Amer. Malacol. Bull.*, 7:147–169, 1990.

Scapolite

An aluminosilicate mineral containing sodium (Na), calcium (Ca), chlorine (Cl), and carbonate (CO_3). It is commonly found as light-colored, translucent tetragonal prisms (**Fig. 1**). Scapolite is normally white, but many other colors are known, including some used as semiprecious gems resembling amethyst and citrine. It has good 90° cleavage parallel to its

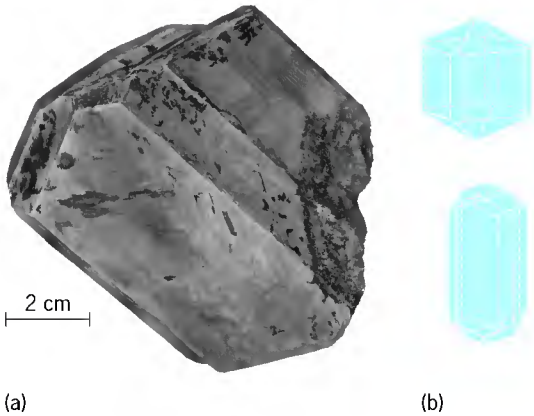

Fig. 1. Scapolite. (a) Crystal taken from Pierrepont, New York (*specimen from Department of Geology, Bryn Mawr College*). (*b*) Crystal habits (*after C. S. Hurlbut, Jr., Dana's Manual of Mineralogy, 17th ed., John Wiley and Sons, 1959*)

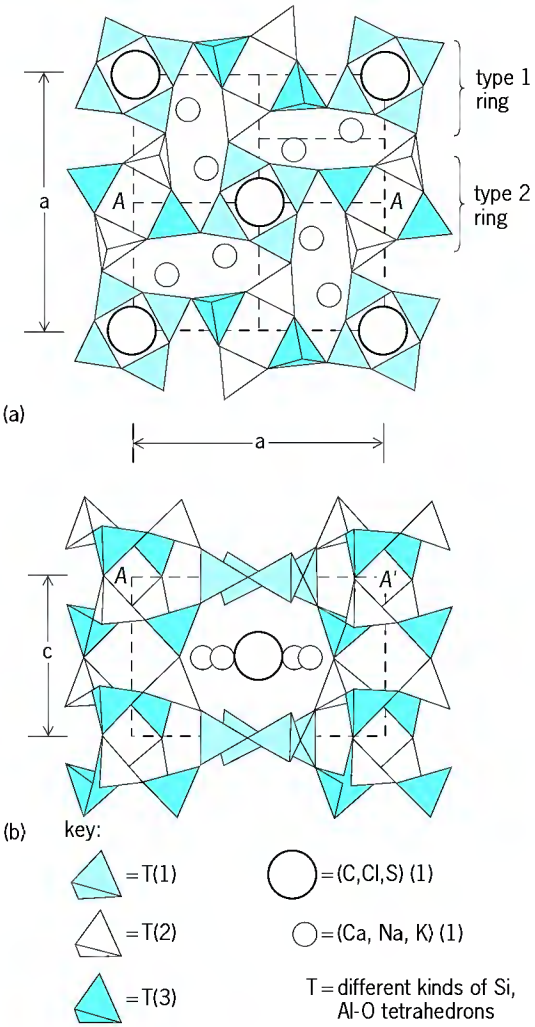

Fig. 2. Crystal structure of scapolite, (a) perpendicular to and (*b*) parallel to the tetragonal *c* axis. The dimensions designated a and c are lengths of the unit cell edges. (*After S. B. Lin and B. J. Burley, The crystal structure of an intermediate scapolite-wernerite, Tschermaks Min. Petr. Mitt., 21:196–215, 1974*)

c axis, and has a Mohs hardness of 5–6.

The formula of scapolite is $(Na,Ca)_4(Al,Si)_6Si_6O_{24}(Cl,CO_3,SO_4)$. Natural scapolites have compositions which vary from $Ca/(Ca + Na) = 0.1–0.9$ and $Cl/(Cl + CO_3 + SO_4) = 0.0–0.7$. In nature significant amounts of K and SO_4 substitute for Na and CO_3 (up to 2% K and 4% SO_4). Crystal-structure studies indicate that scapolite crystallizes in the space group $P4_2/n$. The structure (**Fig. 2**) consists of a framework of Al-O and Si-O tetrahedra, with oval cavities which hold Na and Ca and large cavities in which the Cl, CO_3 and SO_4 occur. Optically scapolite is uniaxial negative, with moderate to high birefringence. Unit cell dimensions, density, and refractive indices all increase with Ca content.

Scapolite is a common mineral in metamorphic rocks, particularly in those which contain calcite. It is found in marbles, gneisses, skarns, and schists. Feldspar, quartz, calcite, pyroxene, amphibole, and other Na-Ca minerals are common associates. Scapolite probably forms about 0.1% of the Earth's upper crust. It is commonly found as inclusions in igneous rocks derived from deep within the Earth's crust, and probably makes up several percent of the lower crust. Experiments have shown that scapolite is stable up to 2700°F (1500°C) at 20 kbar (2 gigapascals) pressure, and that the range of compositions can be much greater at those pressures. Assemblages of scapolite, plagioclase, calcite, and halite simulating those found in nature have been synthesized at 750°C (1400°F) and 4 kbar (400 MPa). Many scapolitic rocks undoubtedly formed at much lower temperatures and pressures, under conditions not yet experimentally reproduced. *See* SILICATE MINERALS. David E. Ellis

Bibliography. D. E. Ellis, Stability and phase equilibria of chloride and carbonate bearing scapolites at 750°C and 4000 bar, *Geochim. Cosmochim. Acta*, 42:1271–1281, 1978; J. R. Goldsmith, Scapolites, granulites, and volatiles in the lower crust, *Geol. Soc. Amer. Bull.*, 87:161–168, 1976; S. B. Lin and B. J. Burley, The crystal structure of an intermediate scapolite-wernerite, *Tschermaks Min. Petr. Mitt.*, 21:196–215, 1974; D. M. Shaw, The geochemistry of scapolite, *Geochim. Cosmochim. Acta*, 1:218–285, 1960.

Scarlet fever

An acute contagious disease that is a consequence of infection with *Streptococcus pyogenes* (group A streptococci). It most often accompanies pharyngeal (throat) infections with this organism but is occasionally associated with wound infection or septicemia. Scarlet fever is characterized by the appearance, about 2 days after development of pharyngitis, of a red rash that blanches under pressure and has a sandpaper texture. Usually the rash appears first on the trunk and neck and spreads to the extremities. The rash fades after a week, with desquamation, or peeling, generally occurring during convalescence. Like streptococcal pharyngitis without such a rash, the disease is usually self-limiting, although severe forms are occasionally seen with high fever and systemic toxicity. Appropriate antibiotic therapy is recommended to prevent the onset in susceptible individuals of rheumatic fever and rheumatic heart disease.

Scarlet fever occurs after infection with streptococci that produce the extracellular products, erythrogenic toxins A, B, and C (also known as streptococcal pyrogenic exotoxins). The genes for the toxins are carried by lysogenic bacteriophages that, on infection of streptococci, can convert them to toxin-producing strains. The toxins are believed to be responsible for the rash of scarlet fever through interaction with the host defense system. Formerly, scarlet fever was associated with toxin A–producing streptococcal strains. Then there was a shift to a prevalence of toxin B and toxin C producers. A resurgence of toxin A–producing streptococci, however, is again being noted. The incidence of scarlet fever is declining, but this decline is apparently not a consequence of a reduction in streptococcal infections. The decline remains unexplained. *See* MEDICAL BACTERIOLOGY; RHEUMATIC FEVER; STREPTOCOCCUS. Ernest D. Gray

Bibliography. G. L. Mandell, G. Douglas, Jr., and J. E. Bennett (eds.), *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases*, 5th ed., 2000; L. W. Wannamaker, Streptococcal toxins, *Rev. Infect. Dis.*, 5(suppl. 14):S723–S732, 1983; G. P. Youmans, P. Y. Paterson, and H. M. Sommers (eds.), *Biologic and Clinical Basis of Infectious Diseases*, 1985.

Scattering experiments (atoms and molecules)

Experiments in which a beam of incident electrons, atoms, or molecules is deflected by collisions with an atom or molecule. Such experiments provide tests of the theory of scattering as well as information about atomic and molecular forces. Scattering experiments can be designed to simulate conditions in planetary atmospheres, electrical discharges, gas lasers, fusion reactors, stars, and planetary nebulae. *See* ELECTRICAL CONDUCTION IN GASES; GAS DISCHARGE; LASER; NUCLEAR FUSION; PLANET; PLANETARY NEBULA; STAR.

Classification of collisions. A collision between two atomic systems is termed elastic if there is no transfer of energy between the internal motions of the two systems and the kinetic energy of their relative motion. A collision is termed inelastic or superelastic, depending on whether kinetic energy is given to or taken from the internal motion.

If electromagnetic radiation is emitted during the impact, the collision is termed radiative. In an ionizing collision, one or more electrons are set free by the colliding partners. Free electrons may also become bound to other particles in collisions by a process called recombination (with positive ions to form a neutral particle) or attachment (with atoms

or molecules to form a negatively charged ion). In a rearrangement collision, there is a redistribution of particles between the colliding systems during the impact. This rearrangement may involve the transfer of atoms or electrons from one system to the other. The latter process is called electron transfer or charge exchange.

In general, in any type of collision, scattering occurs, which causes the direction of relative motion of the two systems to be rotated to a new direction after the collision. More than two systems may also result from such an impact. A complete description of a collision event requires measurement of the directions, speeds and internal states of all the products. *See* COLLISION (PHYSICS).

Types of scattering experiments. There are two basic types of scattering experiments. The simpler (**Fig. 1***a*) involves passing a collimated beam of particles (electrons, atoms, molecules, or ions) through a dilute target gas (in a cell or a jet) and measuring the fraction of incident particles that are deflected into a certain angle relative to the incident beam direction. In the second method (Fig. 1*b*), a collimated beam of particles intersects a second beam. The scattering events are usually registered by measuring the deflection or internal-state change of the beam particles. *See* MOLECULAR BEAMS.

Collision cross sections. A scattering event is usually quantified in terms of a collision cross section, which may be simply thought of as an effective area surrounding one system into which the second system must pass in order for a specific process to occur. It is helpful to consider the simple case of a beam of electrons passing through a dilute target gas. An electron may be considered lost from the beam because of a change of either its direction or speed (energy). Through these processes, the beam current will be reduced by a fraction αx of the total beam in passing an infinitesimal distance x through the gas. Here, the quantity α is the product of the number of gas atoms or molecules per unit volume and $\sigma(E)$, the total effective cross section for collisions between electrons of energy E and the gas atoms or molecules.

Effective cross sections for different classes of collisions follow by introducing probabilities for a particular type of collision to occur. The product of this probability and the total effective cross section is

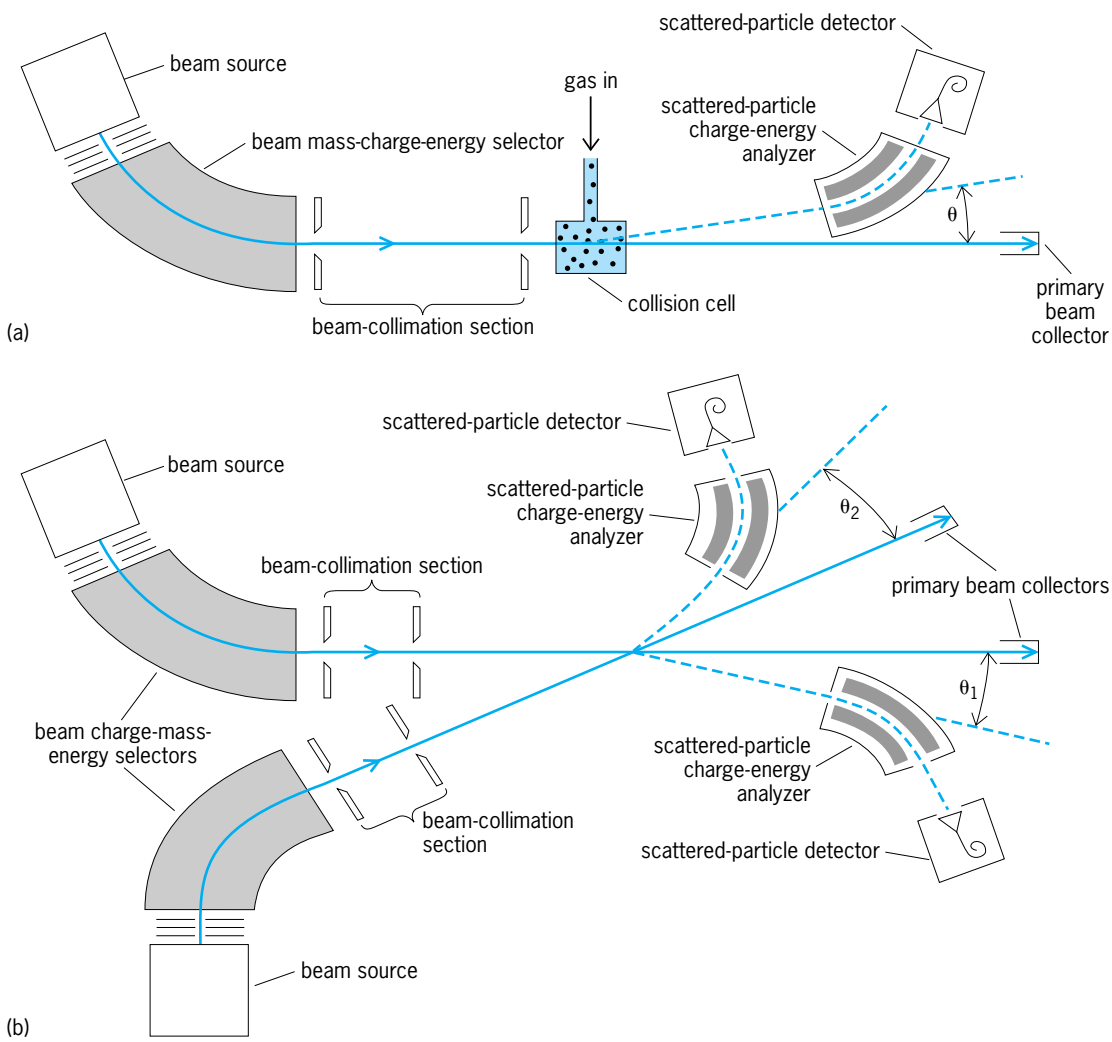

Fig. 1. Two different types of scattering experiments. (*a*) Beam-gas-target experiment. θ = scattering angle. (*b*) Colliding-beams experiment. θ_1, θ_2 = scattering angles.

then the effective total cross section for the particular type of collision.

Scattering in a particular type of collision is specified in terms of a differential cross section. The probability that, in a particular type of collision, the direction of motion of the electron is turned through a specified angle θ (called the scattering angle) into a specified solid angle is proportional to the corresponding differential scattering cross section.

Collision cross sections are most often expressed in square centimeters, and can be measured with appropriately designed experimental apparatus (Fig. 1). Depending on the type of collision process, that apparatus may measure the scattering angle, energy, charge, or mass of the scattered systems.

Scattering theory. For the simplest case, the scattering of a beam of structureless particles of specified mass and speed by a structureless scattering center, the differential cross section may be calculated exactly by using the quantum theory. In the special case where the Coulomb force fully describes the interaction, both the quantum and classical theory give the same exact value for the differential cross section at all values of the scattering angle. *See* NONRELATIVISTIC QUANTUM THEORY.

For scattering of systems with internal structure (for example, molecules, and their ions), no exact theoretical calculation of the cross section is possible. Methods of approximation specific to different types of collisions have been developed. The power of modern high-speed computers has greatly increased their scope and effectiveness, with scattering experiments serving as benchmarks. *See* ATOMIC STRUCTURE AND SPECTRA; SUPERCOMPUTER.

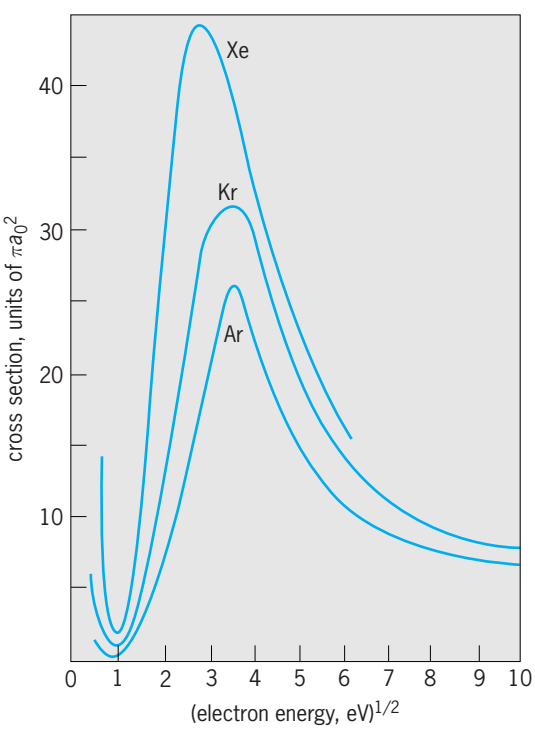

Fig. 2. Measured variation of total cross sections for collisions of slow electrons in argon (Ar), krypton (Kr), and xenon (Xe), as a function of the square root of the electron energy in electronvolts, which is proportional to electron velocity; $a_0 = 5.3 \times 10^{-9}$ cm.

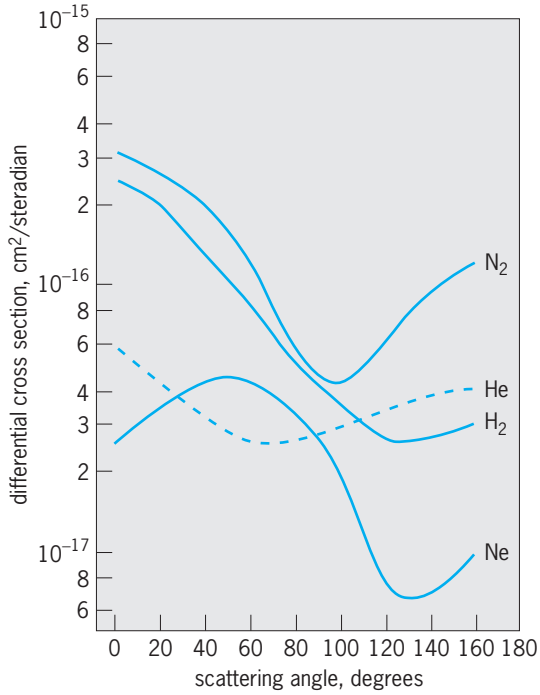

Fig. 3. Measured differential cross sections as a function of scattering angle for the elastic scattering of electrons from neon atoms (Ne), hydrogen molecules (H_2), helium atoms (He), and nitrogen molecules (N_2) at an energy of 10 eV.

Electron-atom collisions. The different types of collisions that can occur are elastic, inelastic involving excitation of discrete atomic states, ionization, radiative electron capture, and radiative collisions in which the electron is not captured. Radiative collisions are much less probable than nonradiative collisions, whereas cross sections for inelastic collisions are comparable to those for elastic collisions at electron energies that are large compared to the minimum (threshold) energy necessary to produce excitation or ionization.

Elastic scattering. The cross section for electron scattering from atoms varies with electron energy (**Fig. 2**). The measured differential cross sections in this energy range exhibit maxima and minima as a function of the scattering angle θ (**Fig. 3**). These effects are due to diffraction of the electrons by the scattering atoms. *See* ELECTRON DIFFRACTION.

If the energy homogeneity of the electron beam is sufficient, fine structure is observed in the variation of the total cross section with electron energy. For example, if the fraction of electrons in a beam that pass undeviated through an argon gas target is plotted against electron energy (**Fig. 4**), the transmitted fraction fluctuates rapidly over a narrow 0.5-eV energy range. This fine structure is due the existence of unstable atomic states called resonances, in which the electron is temporarily bound to an excited atom. After a characteristic lifetime, τ, of typically 10^{-14} s, the captured electron will be reemitted in some direction that is uncorrelated with its incident direction; that is, it will be scattered. Such scattering is

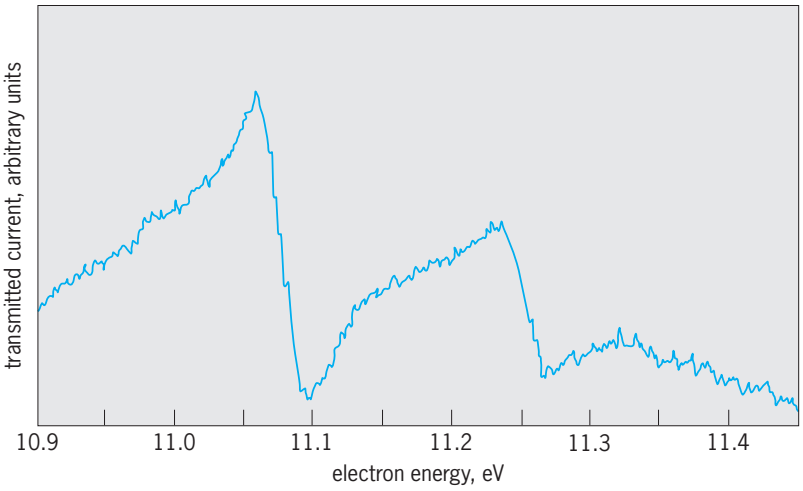

Fig. 4. Observed variation with electron energy of the intensity of an electron beam after transmission through argon gas; the greater the transmitted intensity, the smaller the scattering cross section.

probable only if the incident electron has an energy within ΔE, given by Eq. (1), of an unstable

$$\Delta E = \frac{h}{2\pi\tau} \tag{1}$$

excited state, where h is Planck's constant (4.14×10^{-15} eV-s). *See* RESONANCE (QUANTUM MECHANICS); UNCERTAINTY PRINCIPLE.

Inelastic scattering. A variety of experimental techniques have been used to study inelastic collisions. The excitation of atomic states has been investigated by both optical and energy-loss methods. In the optical method, measurement is made of the intensity of radiation emitted at a particular wavelength by gas atoms excited along the electron beam path. More detailed information about the scattering process has been obtained by detecting the scattered electron and emitted photon in coincidence. Simultaneous measurement of the electron scattering angle and light polarization in such an experiment allows the relative population of angular momentum states of the excited atom to be determined. *See* ANGULAR MOMENTUM.

The energy-loss method involves measurement of the fraction of electrons that have lost specific amounts of energy after passing through a gas target. The energy lost is the energy difference between the ground and excited atomic states. For energies not much greater than the excitation threshold, the scattered electron current exhibits fluctuations (**Fig. 5**) similar to that described above for resonances in elastic scattering, except that in this case the unstable excited state breaks up to leave the atom still excited, so that the ejected electron has a reduced energy.

Ionizing collisions have been studied primarily by measuring the number of positive ions produced by an electron beam of a definite energy passing through an atomic gas target. The relative cross sections for collisions in which one, two, or more electrons are removed from the atom have been determined by performing a charge analysis of the product ions. Fluctuations in the ionization cross section also arise from the excitation of inner-shell electrons, which produces unstable doubly excited states of the atom. Such states usually decay with a characteristic time of about 10^{-14} s by a radiationless process called Auger decay or autoionization, which involves two outer-shell electrons. One electron drops down to fill the inner-shell vacancy, and the other is ejected from the atom with a well-defined energy. It is also possible for such a doubly excited state to decay radiatively, and many experimental measurements have been made of x-ray emission resulting from atomic collisions. *See* AUGER EFFECT; IONIZATION.

There are more sophisticated ionization experiments in which both the scattered and ejected electrons are detected in coincidence. With a suitable choice of incident-electron energy and scattering geometry, the momentum distributions of electrons in atoms and molecules can be probed.

The cross section for radiative capture of an electron by an atom to form a negative ion is quite small (of the other of 10^{-21} cm^2 or less), and has been studied in shock-wave excited discharges and high-pressure arcs. The inverse process, called photodetachment of an electron from a negative ion, has been studied extensively by using lasers, and is of major importance in the solar atmosphere. The H^- ions determine the wavelength distribution of the continuous emission from the Sun in the visible region of the spectrum. *See* ARC DISCHARGE; SHOCK WAVE; SUN.

The experimental study of the scattering of electrons by atoms in short-lived excited states has been made possible by the use of lasers to prepare the atoms in specific states. Lasers also permit the study of collisions in the presence of strong electromagnetic fields.

Electron-ion collisions. Cross sections for collisions between electrons and positive or negative ions have been measured primarily by the colliding-beams method. Such measurements are often more

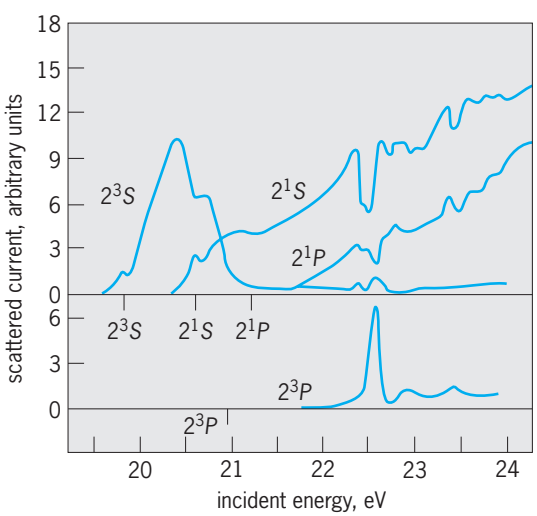

Fig. 5. Observed variation with incident electron energy of the current of electrons scattered without deviation after exciting the 2^1S, 2^1P, 2^3S, and 2^3P electronic states of helium. Excitation threshold energies are indicated. The values for the 2^3S curve are multiplied by 0.3.

accurate than those for atoms because the ions carry a charge and are more easily registered.

Excitation. The behavior with electron energy of cross sections for excitation of ions by electron impact are similar to those for atoms, except that for ions the cross section is finite and often largest at the threshold energy. This phenomenon is a consequence of the Coulomb attraction between the electron and ion, and is predicted by quantum theory. Total excitation cross sections have been measured for approximately a dozen singly and multiply charged ions, and differential cross sections have been measured for a few.

Ionization. Total cross sections for ionization of ions by electron impact have been measured for a large number of singly and multiply charged positive ions by using the crossed-beams method, in which the further-ionized ions are separated and detected after the collision. The ionization of singly charged ions is dominated by the direct "knock-on" ejection of outer-shell electrons, whereas an indirect mechanism involving excitation of inner-shell electrons can be important for multiply charged ions. In the latter case, doubly excited levels are formed that decay rapidly by autoionization, resulting in the net loss of one or more electrons from the ion. Cross sections have been measured for electron-impact single ionization of iron atoms and ions in charge states as high as +15. Iron ions play an important role in the solar corona and in high-temperature reactors used in fusion energy research. Cross sections for electron detachment from negative ions such as H^- have also been measured by the crossed-beams method.

Electron-ion recombination. An electron and a positive ion may recombine to form a neutral system or systems, if there is a mechanism available by which the recombination can be stabilized and the total energy conserved. This may be accomplished via the processes given in reactions (2)–(5).

Radiative recombination:

$$e^- + A^+ \rightarrow A + h\nu \qquad (2)$$

Dielectronic recombination:

$$e^- + A^+ \rightarrow A^{**} \rightarrow A^* + h\nu \qquad (3)$$

Dissociative recombination:

$$e^- + AB^+ \rightarrow A + B \qquad (4)$$

Three-body recombination:

$$e^- + A^+ + B \rightarrow A + B \qquad (5a)$$

$$e^- + e^- + A^+ \rightarrow A + e^- \qquad (5b)$$

Radiative recombination, given in reaction (2), where $h\nu$ represents an emitted photon, is important only at very low electron energies, and cross sections have been measured with a merged-beams technique. It has been predicted theoretically that this process, which is the inverse of photoionization, can be enhanced by an external radiation field. The stimulated radiative recombination process has been observed experimentally with a collinear laser beam. *See* PHOTOIONIZATION; PHOTON.

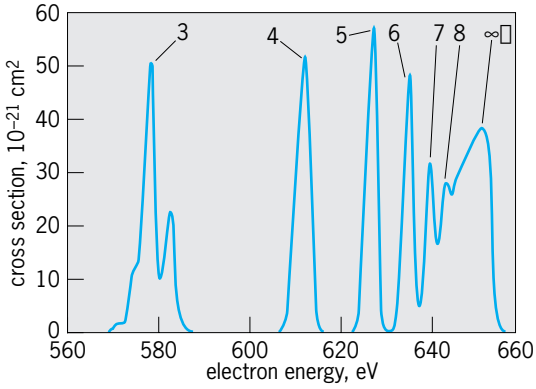

Fig. 6. Theoretical cross section for dielectronic recombination of electrons and hydrogenlike O^{7+} ions. The number above each resonance is the principal quantum number of the captured electron. Measurements made in a heavy-ion storage ring were generally in agreement with these theoretical values.

Dielectronic recombination, given in reaction (3), where A^* and A^{**} refer to a singly excited and doubly excited atom, respectively, is important in stellar atmospheres and in fusion-reactor plasmas. It is a resonant process resulting in the formation of an unstable, doubly excited state, A^{**}, of the electron-ion system. If the unstable state decays by a radiative transition rather than by autoionization, the recombination is stabilized. The branching ratio for these two pathways is critical in determining the cross section for dielectronic recombination. Intense electron beams have been merged with ion beams in heavy-ion synchrotron storage rings in order to measure cross sections for this process with extremely high energy resolution and precision (**Fig. 6**). Laser-stimulated dielectronic recombination has also been observed. *See* PARTICLE ACCELERATOR; PLASMA (PHYSICS).

Dissociative recombination, reaction (4) has been studied experimentally for many molecular ions by using crossed-beams, merged-beams, and flowing-afterglow techniques. Cross sections for this process are largest at near-thermal electron energies and, at their maximum, range from 10^{-15} cm^2 for simple diatomic ions to 10^{-12} cm^2 for complex polyatomic ions. This process is believed to play an important role in planetary ionospheres and in the formation of molecules in interstellar space. *See* INTERSTELLAR MATTER; IONOSPHERE.

Three-body recombination processes, reactions (5*a*) and (5*b*), have received less experimental study, and are important in high-density environments such as laser-produced plasmas.

Electron-molecule collisions. If an electron collides with a molecule rather than an atom, additional possibilities arise. Molecular vibration or rotation may be excited, and the molecule may dissociate into two or more neutral or ionized fragments. More than one of these processes may occur in the same collision event. It has been found experimentally that vibrational excitation of some molecules occurs most efficiently via the formation of a temporary unstable negative molecular ion. The excitation of rotation

is much more likely for molecules with permanent electric dipole or quadrupole moments. *See* MOLECULAR STRUCTURE AND SPECTRA.

Extensive studies have been made of the production of different ion fragments due to collisions of 100-eV electrons with various molecules. Such analysis using mass spectrometry is useful for industrial gases and vapors, especially hydrocarbons. Negative ion formation is most likely to occur via the process of dissociative electron attachment, given in reaction (6).

$$e^- + AB \rightarrow A + B^- \tag{6}$$

Collisions of atomic systems. If both the colliding systems are structured (for example, atoms, molecules, orions), elastic, excitation, or ionization processes may occur. In general, electronic excitation and ionization will be likely only if the relative speed of the colliding systems exceeds that of the bound electrons, and the cross sections will be similar to those for impact of electrons of the same speed. Either or both of the colliding systems may be excited in the collision.

Transfer collisions. There are additional possibilities called transfer collisions. They may involve the transfer of excitation, as in reactions (7) and (8). Reaction

$$A^* + B \rightarrow A + B^* \tag{7}$$

$$A^* + B \rightarrow A + B^+ + e^- \tag{8}$$

(8) is called Penning ionization.

Transfer of electrons from a neutral atom or molecule to an ion is most often studied by passing an ion beam through a gas target and measuring the fraction of ions whose charge has been changed. Slow ions produced in the gas may also be collected. At very low collision energies, the process has been studied by using merged beams. The electron transfer process may be represented in general by reaction (9), where q is the initial ion charge, k elec-

$$A^{q+} + B \rightarrow A^{(q-k)+} + B^{k+} + \Delta E \tag{9}$$

trons are transferred, and ΔE is the potential energy difference between the initial and final states of the system. If $\Delta E > 0$, the process is exoergic. For an endoergic process, $\Delta E < 0$, and this energy must be supplied by the initial kinetic energy of the reactants. The process is resonant if $\Delta E = 0$. Cross sections are large (10^{-14} cm^2) at low energies if ΔE is zero or small. For multiply-charged ions, the process is often exoergic ($\Delta E > 0$). Such a collision is characterized as high- or low-energy, depending on whether the relative velocity is large or small compared to the orbital velocity of the transferred electron (2.2×10^8 cm/s for hydrogen). The initial ionic charge determines the cross section at intermediate and high collision energies. At low energies the cross section depends on the structure of the transient quasimolecule formed during the collision.

Multiple electron transfer from multielectron atoms and molecules occurs predominantly into multiply excited levels of the ion which stabilize either radiatively, leading to true electron transfer, or via autoionization, leading to transfer ionization. Electron transfer processes are important in determining the ionization and energy balance in fusion plasmas and in the interstellar medium.

Radiative collisions. Extensive studies have been made of the emission of radiation by complex atoms and ions in collision at energies in the kiloelectronvolt range. This radiation is characteristic of transitions between states of the quasimolecule formed upon close approach of the colliding systems.

Ronald A. Phaneuf

Bibliography. E. W. McDaniel, *Atomic Collisions: Electron and Photon Projectiles*, 1989; H. S. W. Massey, E. H. S. Burhop, and H. B. Gilbody, *Electronic and Ionic Impact Phenomena*, 2d ed., vols. 1–5, 1969–1974; N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions*, 3d ed., 1965, paper 1987.

Scattering experiments (nuclei)

Experiments in which beams of particles such as electrons, nucleons, alpha particles and other atomic nuclei, and mesons are deflected by elastic collisions with atomic nuclei. Much is learned from such experiments about the nature of the scattered particle, the scattering center, and the forces acting between them. Scattering experiments, made possible by the construction of high-energy particle accelerators and the development of specialized techniques for detecting the scattered particles, are one of the main sources of information regarding the structure of matter. *See* PARTICLE ACCELERATOR; PARTICLE DETECTOR.

In the broad sense, any nuclear reaction is an example of scattering. However, this article treats elastic scattering only in the more restricted sense given below, and only insofar as it involves atomic nuclei. *See* COULOMB EXCITATION; GIANT NUCLEAR RESONANCES; NEUTRON SPECTROMETRY; NUCLEAR REACTION; NUCLEAR SPECTRA; NUCLEAR STRUCTURE.

Definitions of elastic scattering. The word "elastic" is used to indicate the absence of energy loss. If particle A collides with particle B of finite mass, there is a loss in the energy of A even if no energy has been transferred to the internal degrees of freedom of either A or B. Sometimes such a collision is referred to as inelastic, in order to distinguish its character from that of a collision with a particle having an infinite mass or its idealization, a fixed center of force. This terminology is not useful in the present context, because in the center-of-mass system of the two particles the sum of kinetic energies after the collision is the same as before. The distinction between elastic and inelastic scattering is made therefore on the basis of whether there are internal energy changes in the colliding particles. The collision is said to be inelastic even if the energy changes of the two particles compensate so as to leave the sum of the kinetic energies in the center-of-mass system unaltered. The treatment of inelastic scattering involves nuclear reaction theory because nuclear reactions markedly

influence the scattering. *See* COLLISION (PHYSICS).

Cross sections. The results of scattering measurements and calculations are generally expressed in terms of differential cross sections, which furnish a quantitative measure of the probability that the incident particle is scattered through an angle, θ. Cross sections have units of area, and in nuclear physics the convenient measure of area is the barn and its subunits, the millibarn and microbarn (1 barn = 10^{-28} m^2).

Coulomb scattering by nuclei. The simplest type of elastic scattering experiment relevant to the study of nuclei involves the deflection of incident electrically charged particles by the Coulomb field of the target nucleus. Such experiments provided the first evidence for atomic nuclei. H. Geiger and E. Marsden observed in 1909 that low-energy alpha particles could be scattered through large angles in collisions with gold and silver targets, and Ernest Rutherford showed in 1911 that these results could be understood if the scattering center consisted of a positively charged region (the nucleus) considerably smaller in size than that occupied by an individual target atom. The yield of scattered alpha particles as a function of angle, called the angular distribution, was measured in detail in a subsequent (1913) experiment by Geiger and Marsden, and the results were entirely consistent with the Rutherford expression for the differential cross section given by Eq. (1), where Z_1 and

$$\frac{d\sigma}{d\Omega_{\text{point}}}(\theta) = \left(\frac{Z_1 Z_2 e^2}{16\pi\epsilon_0 E}\right)^2 \sin^{-4}\frac{\theta}{2} \qquad (1)$$

Z_2 are the atomic numbers of the target and projectile, e is the electronic charge, ϵ_0 is the permittivity of free space, E is the center-of-mass energy of the alpha particle projectile, and θ is the center-of-mass scattering, or observation angle.

Equation (1) is appropriate only when the classical distance of closest approach, $d = Z_1 Z_2 e^2/4\pi\epsilon_0 E$, is greater than the combined size of the colliding pair; that is, only if the two charge distributions never overlap during the course of their interaction. The 1913 experiments involved collisions of 7.68-MeV alpha particles (1 MeV = 1.6×10^{-13} joule) with gold targets, and for this situation $d = 2.96 \times 10^{-14}$ m (29.6 fermis or femtometers in conventional nuclear physics notation, with 1 fm = 10^{-15} m). The detailed agreement between the Rutherford prediction of Eq. (1) and the 7.68-MeV data thus implied that the nuclear charge is contained within a region smaller than 30 fm (1.2×10^{-12} in.) in radius. Subsequent experiments involving higher-energy projectiles have determined the actual half-density radius of the gold nucleus to be approximately 7 fm (3×10^{-13} in.).

Although nuclear charge radii can be measured by systematically increasing the energy of the incident alpha particle until Eq. (1) no longer describes the angular distributions, specifically nuclear interactions come into play when the colliding nuclei overlap. In that situation, nuclear size is no longer the only factor being probed. This complexity is avoided when high-energy electrons are used as projectiles, since electrons are not subject to the nuclear forces. At very low energies the elastic scattering of electrons by nuclei conforms to the predictions of Eq. (1), but much more information is revealed at higher energies where the electrons can penetrate into the target nucleus. In the latter case and neglecting magnetic effects for the moment, the finite extent of the nucleus requires that Eq. (1) be modified by the presence of a nuclear form factor, $F(q^2)$, as in Eq. (2), in

$$\frac{d\sigma}{d\Omega}(\theta) = \frac{d\sigma}{d\Omega_{\text{point}}}(\theta)[F(\vec{q}^2)]^2 \qquad (2)$$

which q^2 is the momentum transfer in the collision. The significance of $F(q^2)$ can be appreciated by noting that at nonrelativistic energies, and assuming the Born approximation to be valid, $F(q^2)$ is the Fourier transform of the nuclear charge distribution.

If sufficiently complete measurements over a wide range of momentum transfer are available, the form factor can be determined in detail, and the nuclear charge density, $\rho_{\text{ch}}(r)$, obtained directly through the inverse transform. In practice, a functional form is usually assumed for $\rho_{\text{ch}}(r)$, and its parameters are adjusted to fit the electron scattering data. In this manner it is found that the experiments are consistent with nuclear charge distributions specified by Eq. (3). The parameter c is the radius at which the

$$\rho_{\text{ch}}(r) = \rho_0[1 + e^{(r-c)/a}]^{-1} \qquad (3)$$

density falls to one-half its central value. It assumes the typical value $c \approx 1.1A^{1/3}$ fm, where A is the atomic number of the target nucleus. The quantity a (typically ≈ 0.55 fm) reflects the fact that nuclear surfaces are not sharply defined, so that the density falls off gradually with increasing radius. The so-called central charge density, ρ_0, is found to be approximately 1.1×10^{25} coulombs/m^3, which corresponds to 0.07 proton per cubic femtometer (1.1×10^{39} protons per cubic inch). That ρ_0 is approximately the same for all but the lightest nuclei and $c \propto A^{1/3}$ implies that nuclear matter is nearly incompressible. The charge distributions of a variety of nuclei, as determined by electron scattering measurements, are shown in **Fig. 1**.

Electron scattering measurements involving large angular momentum transfer are sensitive to the distribution of magnetism as well as charge, and Eq. (2) then must be generalized to reflect the contributions from the two distributions. In addition, very precise measurements show that Eq. (3) is only approximately correct and the form of $\rho_{\text{ch}}(r)$ varies slightly from nucleus to nucleus as a consequence of variations in the structure of the nuclei in question.

Electron-nucleon scattering. Scattering of electrons by hydrogen gives information regarding the electron-proton interaction. From measurements of the variation of the differential cross section, it has been found necessary to postulate that both the proton charge and its intrinsic magnetic moment are distributed through a finite volume. Existing work favors the assumption of similarity of shape of these

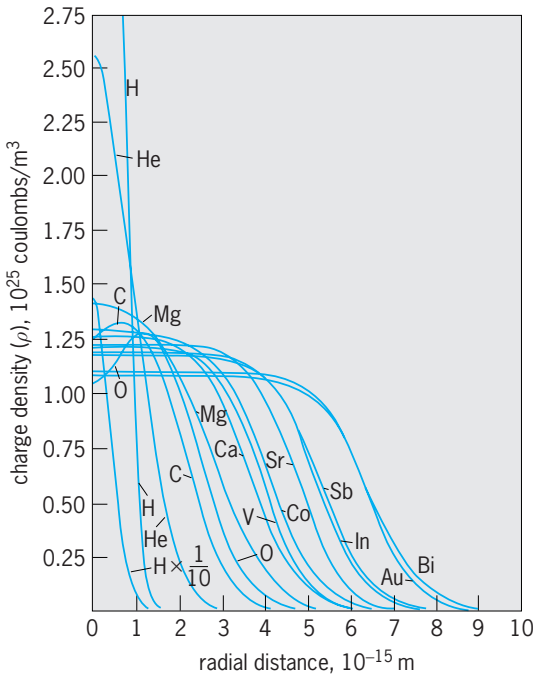

Fig. 1. Summary of the nuclear charge distributions found for various nuclei by electron scattering methods. (*After R. Hofstadter, Nuclear and nucleon scattering of high-energy electrons, Annu. Rev. Nucl. Sci., 7:231–316, 1957*)

distributions. Energies in excess of 200 MeV are required for the detection of these effects, and the anomalous magnetic moment of the proton plays an important part at high energies. The experiments make it probable that the charge density has a root-mean-square radius of approximately 8×10^{-14} cm (3×10^{-14} in.). Measurements on the scattering of electrons by deuterium at large angles and higher energies lead to the conclusion that the magnetic moment of the neutron is not concentrated at this point. If it is assumed that neutron and proton magnetic moments are distributed through nearly the same volumes, a good representation of the scattering measurements is obtained.

Analyses of scattering data at various angles and energies indicate that there is no net charge density within the volume occupied by the neutron. This result is in agreement with measurements of the neutron-electron interaction made by scattering very slow neutrons from atomic electrons and atomic nuclei. While there is an interaction equivalent to a potential energy of approximately −3900 eV through a distance of e^2/mc^2, where m is the electron mass and c is the speed of light, amounting to approximately 2.8×10^{-13} cm (1.1×10^{-13} in.), it is accounted for qualitatively as a consequence of what E. Schrödinger called the *Zitterbewegung* (tremblatory motion) expected for the neutron.

Electron scattering by nucleons has also provided valuable information concerning subnuclear processes. *See* ELEMENTARY PARTICLE; QUARKS.

Low-energy np scattering. Since the neutron is electrically neutral, an understanding of how it scatters in an interaction with a proton requires, even for low collision energies, a knowledge of the properties of the nuclear forces. Conversely, the scattering experiments probe these properties. The *np* force is responsible for the binding together of a neutron, n, a proton, p, to form a deuteron, and the detailed connection between *np* scattering cross sections and the properties of the deuteron was perceived in the early 1930s. For example, the assumption that the spatial extension, or range, of the *np* force is relatively small makes it possible to estimate the magnitude of the force, given the measured binding energy of the deuteron, and Eugene Wigner used this information in 1933 to calculate *np* scattering cross sections. The measured yield was found to be larger than would be expected if the forces between free neutrons and free protons were the same as those in the deuteron. To explain this difference, it was postulated by Wigner in 1935 that the *np* interaction is spin-dependent; that is, it depends on the relative orientation of the spins of the interacting particles. The proton and neutron are known to have a spin of $^1/_2$; that is, their intrinsic angular momenta are known to be $(h/2\pi)/2 = \hbar/2$, where h is Planck's constant. According to quantum mechanics, when two spins s_1, s_2 combine vectorially, only the values given by Eq. (4) are possible for the resultant s. For the *np* system, therefore, the resultant spins are 0 or 1. In the first case one speaks of a singlet, and in the second of a triplet. *See* ANGULAR MOMENTUM; SPIN (QUANTUM MECHANICS).

$$s = s_1 + s_2, s_1 + s_2 - 1, \ldots, |s_1 - s_2| \qquad (4)$$

The singlet state behaves much like a round and perfectly smooth object which has the same appearance no matter how it is viewed, corresponding to only one possibility of forming a state with $s = 0$. The state with $s = 1$, on the other hand, can have three distinct spin orientations. Measurement of the projection of s on an axis fixed in space can give only the three values (1, 0, −1), again in units $\hbar$. When protons with random spin directions collide with neutrons also having random spin directions, the triplet state is formed three times as often as the singlet. The deuteron, however, is in a triplet state. Thus the hypothesis of spin dependence can account for the difference between the forces in the deuteron and those in *np* scattering.

The neutron-hydrogen scattering experiments on which these conclusions were based were performed with slow neutrons having energies of a few electronvolts or less. The general quantum-mechanical theory of scattering is much simplified in this case. Because of the small range of nuclear forces, only the collisions with zero orbital angular momentum ($L\hbar = 0$) play a role, collisions with higher orbital angular momenta missing the region within which nuclear interactions take place. States with $L = 0$ (called S states) have the property of spherical symmetry, and nuclear forces matter in this case only inasmuch as they modify the spherically symmetric part of the wave functions.

The long-wavelength or low-energy scattering cross section can be described completely and sim-

ply by a quantity called the scattering length. For interparticle distances r greater than the range of nuclear forces, the spherically symmetric part of the wave function has the form given by expression (5),

$$C\left(\frac{1+a}{r}\right) \tag{5}$$

where a and C are constants. The constant a is called the scattering length. It has the following meaning. If $R(r)$ denotes the wave function describing the np relative motion, the product $rR(r)$, when plotted against r, is represented by a straight line which cuts the axis of r at a distance a from the origin of coordinates. If the intersection is to the left (right) of the origin, a is counted as positive (negative). The two conditions are illustrated in **Figs. 2** and **3**. The possibilities $a > 0$ and $a < 0$ are sometimes referred to as those of the virtual and real level, respectively. Sometimes the opposite convention regarding the sign of a is used. The convention adhered to here provides the simplest connection with phase shifts.

The scattering cross section for a state with well-defined a is $4\pi a^2$, and thus low-energy np scattering is described by Eq. (6), where σ_{np} is the total

$$\sigma_{np} = 4\pi({}^1/_4 a_s^2 + {}^3/_4 a_t^2) \tag{6}$$

cross section and a_s and a_t are the singlet and triplet scattering lengths, respectively. Comparison of the scattering yields with the deuteron data provides the values $a_s = 24$ fm and $a_t = 5.4$ fm.

There are several additional means by which np scattering can be calculated. For example, with the techniques outlined below ("Potential scattering"), effective interaction potentials can be found which simulate the np nuclear forces to the extent that calculations using the potentials reproduce the measured cross sections. Two potentials corresponding

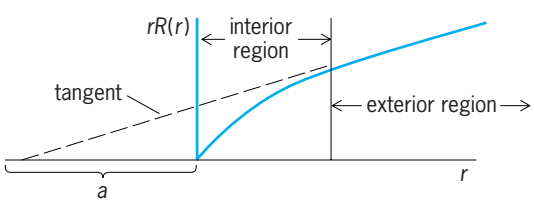

Fig. 2. Scattering length in the case of a virtual level.

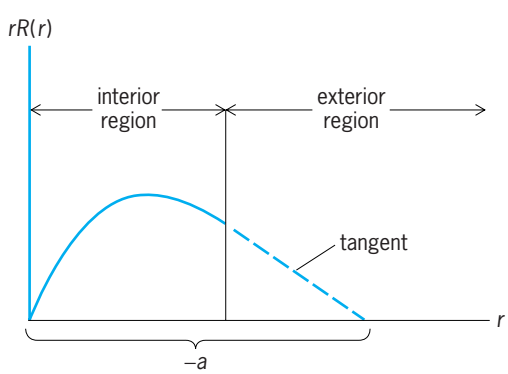

Fig. 3. Scattering length in the case of a real level.

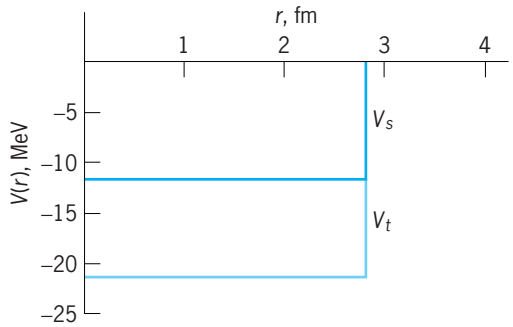

Fig. 4. Examples of singlet (V_s) and triplet (V_t) potential wells suitable for low-energy *np* scattering. The potentials are not uniquely determined; many other equally satisfactory possibilities exist.

to singlet and triplet $L = 0$ scattering are required for a description of low-energy np data, as is illustrated in **Fig. 4**. The term "potential well" is often used to describe these potential energy surfaces, because the system can be trapped in the region of space occupied by the potential somewhat similarly to the way in which water is trapped in a well. For nucleon-nucleon scattering, it is frequently useful to express the potential energy in the form $V(r) = V_0 f(r/b)$, where f is a function which determines the shape of the well. The constants V_0 and b are usually referred to as the depth and range parameters. The scattering length determines approximately the product $V_0 b^2$ for a potential well of assigned shape. The variation of the cross section with energy E through an energy range of a few MeV can be used for the determination of b. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Low-energy pp scattering. The pp and pn interactions are believed to be closely equal. Gregory Breit proposed in 1936 the hypothesis of charge independence of nuclear forces, which supposes that nuclear forces, acting in addition to the electrostatic (Coulomb) repulsion, in the pp, pn, and nn cases are equal to each other. Because of the limits of experimental accuracy and uncertainties in the theoretical interpretation, the hypothesis is not established with perfect accuracy, but it is believed to hold within a few percent for the depth parameter V_0 if the range parameter b is specified as the same in the three cases. The concept of charge independence also has important implications for the structure of nuclei. *See* I-SPIN.

Phase shifts. Scattering can be treated by means of phase shifts, which will be illustrated for two spinless particles. The wave function of relative motion will be considered first for a state of definite orbital angular momentum $L\hbar$, with L representing an integer. Outside the range R of nuclear forces, the wave function may be represented by Eq. (7), where k is

$$\psi_L = \frac{Y_{LM}(\theta, \phi)\mathscr{F}_L(kr)}{(kr)} \tag{7}$$

$$(r > R)$$

2π times the reciprocal of the wavelength, the so-called wave number, θ and σ are the colatitude

and azimuthal angles of a polar coordinate system, and Y_{LM} is the spherical harmonic of order L and azimuthal quantum number M. The form of ψ_L is determined by the Schrödinger wave equation which restricts $\mathscr{F}_L$ by the differential equation (8),

$$\left\{\frac{d^2}{dr^2} + \frac{k^2 - L(L+1)}{r^2}\right\}\mathscr{F}_L = 0 \qquad (r > R) \tag{8}$$

it being supposed that there is no Coulomb field. In the absence of nuclear forces, $\mathscr{F}_L$ satisfies the same equation at all distances and, aside from a constant factor, has its asymptotic form determined by the boundary conditions at $r = 0$ as in notation (9). In

$$\mathscr{F}_L \sim \sin\left(kr - \frac{L\pi}{2}\right) \text{ as } r \to \infty \tag{9}$$

the presence of nuclear interactions, the asymptotic form is given by notation (10), where δ_L is a constant, called the phase shift, which determines the scattering. *See* SPHERICAL HARMONICS.

$$\mathscr{F}_L \sim \sin\left(kr - \frac{L\pi}{2} + \delta_L\right) \tag{10}$$

When elastic scattering is the only process that can take place, as for low- and medium-energy nucleon-nucleon scattering, the phase shifts are real numbers. Above a few hundred megaelectronvolts, nucleon-nucleon scattering becomes strongly inelastic because of pion production, however, and the phase parameters become complex numbers reflecting the loss of flux from the elastic channel. Nucleon-nucleus and nucleus-nucleus interactions are characterized by the presence of many nonelastic scattering channels, except at very low energies.

The wave packet representing the scattered particle consists of a superposition of all possible partial waves, $\mathscr{F}_L$, and calculation of the scattering cross section requires a knowledge of all phase shifts. In practice, however, only partial waves with orbital angular momenta corresponding to impact parameters within the range of the nuclear forces suffer nuclear phase shifts different from zero. The direct method of varying phase shifts to reproduce experimental data is most useful, therefore, when only a few L values satisfy this condition, such as for low- and intermediate-energy nucleon-nucleon scattering and for low-energy nucleon-nucleus and nucleus-nucleus collisions.

For charged spinless particles, such as two alpha particles, the phase shifts caused by specifically nuclear forces add to the asymptotic phase of the functions $\mathscr{F}_L$ for the Coulomb case, which differs from the non-Coulomb case only through the replacement of $kr - L\pi/2$ by $kr - L\pi/2 - \eta \ln(2kr) + \arg\Gamma(L + 1 + i\eta)$, where $\eta = Z_1Z_2e^2/4\pi\epsilon_0\hbar v$, Z_1e and Z_2e represent the charges on the colliding particles, e represents the electronic charge, and v represents the relative velocity. *See* GAMMA FUNCTION.

As in the *np* scattering case discussed above, the presence of spin adds additional complications. For example, if two particles with spin $^1/_2$ collide, it is necessary, in general, to introduce phase shifts for each state with definite total angular momentum $J\hbar$.

Potential scattering. The direct determination of phase shifts through comparison with scattering data becomes difficult in situations where large numbers of partial waves, $\mathscr{F}_L$, are modified by the nuclear scattering center. The usual procedure in this case is to simulate the influence of the nuclear interactions by means of potentials, $V(r)$. For the nonrelativistic scattering of two spinless, uncharged particles, Eq. (8) is supplemented by the Schrödinger equation for $r < R$, Eq. (11), in which m is the reduced

$$\left\{\frac{d^2}{dr^2} + \left[k^2 - \frac{L(L+1)}{r^2} - \frac{2mV(r)}{\hbar^2}\right]\right\}\mathscr{F}_L(r) = 0 \qquad (r > R) \tag{11}$$

mass of the colliding pair. Equations (8) and (11) are then solved subject to appropriate boundary conditions to determine the $\mathscr{F}_L(r)$ and hence δ_L for each L. A functional form is generally assumed for $V(r)$, and its parameters are varied and Eqs. (8) and (11) solved iteratively until the measured cross sections, polarizations, and so forth are reproduced.

From the viewpoint of microscopic models of nucleon-nucleon interactions, it appears highly improbable that a description of nucleon-nucleon scattering in terms of two-body energy-independent local potentials can have fundamental significance. Nevertheless, in a limited energy range, it is practical and customary to represent scattering by means of such a potential. As noted above, the simple potentials illustrated in Fig. 4 suffice for the description of low-energy *np* scattering. At somewhat higher energies, but below the threshold of meson production, a real potential may still be used, but different potentials are required for triplet-even, triplet-odd, singlet-even, and singlet-odd states to account for the (even, odd) parity dependence of the interactions. *See* PARITY (QUANTUM MECHANICS).

Intermediate energy np and pp scattering. In the intermediate-energy region (10–440 MeV) the analysis of experimental material is more difficult than at low energies because of the necessity of employing many phase shifts and coupling constants. Analysis in terms of phase parameters involves fewer assumptions than that in terms of potentials. Except for approximations connected with the infrared catastrophe and related small inaccuracies in relativistic treatment of Coulomb scattering, it is based on very generally accepted assumptions, such as the validity of time reversal and parity symmetries for strong interactions. With infinite experimental accuracy, it should be possible to extract all the phase parameters from measurements of the differential cross section, the polarization spin correlation coefficients, and "triple scattering" quantities describing spin orientation which, for unpolarized incident beams and unpolarized targets, require three successive scatterings.

The analysis is usually carried out by assuming that for sufficiently high L and J the phase parameters may

be represented by means of the one-pion exchange approximation. The value of the pion-nucleon coupling constant g^2 is often varied in an attempt to improve the fit to experimental data, and values for best fits are compared with those from pion physics. Reasonable agreement usually results.

The consistency of values of g^2 from *pp* data with those from *np* measurements indicates approximate validity of charge independence at the larger distances.

Potentials to be used in a nonrelativistic Schrödinger equation and capable of representing *pp* and *np* scattering have been devised either on a purely phenomenological or semiphenomenological basis. The former way provides a more accurate representation of the data. Nonrelativistic local potentials required from 0 to 310 MeV are different according to whether the state is even or odd, singlet or triplet. It is necessary to use central, tensor, spin-orbit, and quadratic spin-orbit parts of the potential. Most of the accurately adjusted potentials employ hard cores within which the potential is infinite. The spin-orbit potential suggested by *pp* scattering data indicated the probable participation of vector-meson exchange in nucleon-nucleon scattering, anticipating the discoveries of the ω- and ρ-mesons, as well as fictitious mesons, the latter partly intended as a representation of simultaneous two-pion exchange. At short distances, the potentials are often modified in order to improve agreement with experiment. Superposition of the single-boson potentials combined with the short-range modifications of the resulting potential gives the so-called one-boson exchange (OBE) potentials. These provide a fair but not excellent reproduction of phenomenological fit phase shifts. *See* QUANTUM FIELD THEORY.

Nucleon-nucleus scattering. Nucleon-nucleus scattering at low and intermediate energies ($E <$ 300 MeV) is of interest primarily in connection with nuclear structure studies. The scattering wave functions, $\mathscr{F}_L(r)$, are required for quantitative interpretation of the inelastic scattering and reaction experiments which have been the main source of information regarding the properties of nuclear excited states. These scattering functions are obtained as the solutions of Eq. (11) when elastic data are reproduced by using potential models.

Nucleon-nucleus scattering experiments can be accounted for by a potential-well model with a complex potential (optical model) making use of a spin-orbit interaction term. The details of angular distributions of the cross section and polarization are reproduced remarkably well, and experiments favor some potential well shapes over others. Wells thus determined are wider than those obtained from electron-nucleus scattering experiments, and similar functional shapes work in both cases. The potential energy represented by the wells is added to the electrostatic potential energy in the calculations. The electrostatic potential energy is approximated by a central potential corresponding to the average distribution of nuclear charge. The quantitative success of the potential-well approach to the scattering of

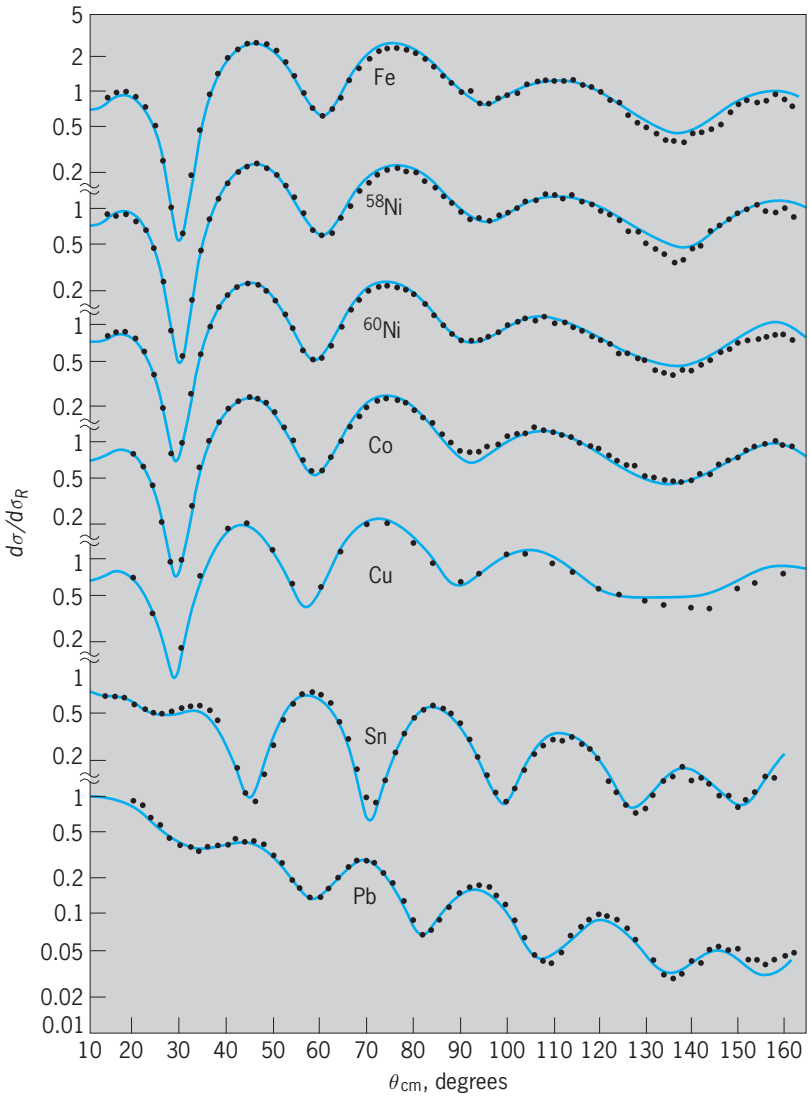

Fig. 5. Measured (points) and calculated (curves) differential cross sections for the scattering of 30-MeV protons from a variety of target nuclei, as a function of scattering angle in the center-of-mass system θ_{cm}. The theoretical calculations made use of optical potential wells. The vertical scale represents the differential cross section $d\sigma$ divided by the Rutherford differential cross section $d\sigma_R$. (*After G. R. Satchler, Optical model for 30 MeV proton scattering, Nucl. Phys., A92:273–305, 1966*)

30-MeV protons by a variety of nuclei is illustrated in **Fig. 5**.

Optical model potential fits to nucleon-nucleus data are in general agreement with the data in a wide energy range from several MeV to about 300 MeV, but the parameters of the potential have to be varied progressively. The spin-orbit potential found to represent the scattering data at the lower energies has the same value of its ratio to $(dV_{ct})/(rdr)$, where V_{ct} is the central potential, as it has in the shell theory of nuclear structure. At higher energies (300 MeV), it has been found that the spin-orbit potential has to be used with a smaller strength than in shell theory. The real part of the central potential has to be used with a smaller strength than in shell theory. The real part of the central potential decreases with energy and becomes almost zero at 300 MeV. Data at 1 GeV on proton scattering from carbon can be accounted

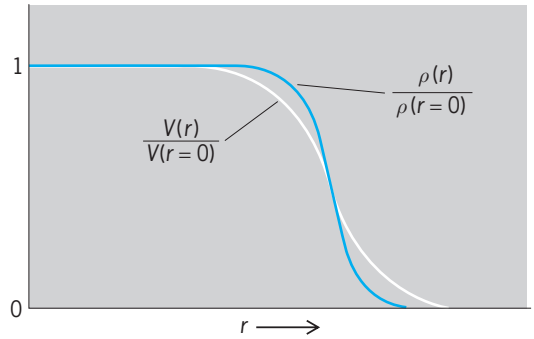

Fig. 6. Comparison between the shape of a nuclear density distribution, $\rho(r)$, and a potential, $V(r)$, obtained from it by folding in an interaction potential, $v(r)$, with a finite but short range as in Eq. (12). (*After G. R. Satchler, Introduction to Nuclear Reactions, John Wiley and Sons, 1980*)

for on the optical model by means of an imaginary central and real spin-orbit potential.

The real part of the volume potential $V(\vec{r}_p)$ can be calculated by folding the interaction potential between the projectile and a target nucleon, $v_{tp}(\vec{r})$, with the nuclear density of the target as in Eq. (12).

$$V(\vec{r}_p) = \int \rho(\vec{r}_t) v(\vec{r}_{tp}) d\vec{r}_t \qquad (12)$$

Effective, rather than free, nucleon-nucleon interactions are used for $v(\vec{r}_{tp})$ because the target nucleon is embedded in a nucleus. The short range of $v(\vec{r}_{tp})$ implies that the resulting central potential has a spatial distribution closely similar to, but somewhat more rounded than, that of the target nuclear density, as is illustrated in **Fig. 6**. The imaginary part of the optical potential is phenomenologically determined.

Nucleus-nucleus scattering. Composite nuclei ranging from deuterium through uranium have been accelerated to energies as high as several GeV per constituent nucleon. Elastic scattering is an important process only toward the lower end of this energy range, perhaps only below a few tens of MeV per nucleon; at the higher energies, the principal interest is in nonelastic collisions, which are expected to provide information concerning properties of nuclear matter under unusual conditions (like high temperature and high density).

Nucleus-nucleus scattering at low and intermediate energies is characterized by strong absorption associated with the relatively short mean free path for nuclei in nuclear matter. Only peripheral or glancing collisions are likely to lead to direct elastic scattering; more head-on collisions lead to more complicated processes such as compound nucleus formation which absorb flux from the incident beam. Since the de Broglie wavelengths of the incident nuclei are typically comparable to or smaller than the size of the target nuclei, the strong absorption for head-on collisions frequently causes the scattering to be diffractive in nature, and the elastic scattering angular distributions to resemble those observed in the scattering of light by small opaque obstacles. A typical example, corresponding to optical Fraunhöfer diffraction, is illustrated in **Fig. 7**.

Good fits to nucleus-nucleus scattering data are obtained by using potential wells. For collisions involving relatively light nuclei, the real potentials may be derived by folding an effective nucleon-nucleon interaction with the densities of the two colliding nuclei, in a slight generalization of Eq. (12). For more massive nuclei, analogies with the collisions of liquid droplets, provide guidance in determining the potentials. In many cases, purely phenomenological potentials provide the best fits to the data. There is considerable ambiguity in the potential parameters, because the scattering depends primarily on the values of the potentials at distances corresponding to glancing collisions, and is relatively insensitive to the values at smaller distances. The imaginary potentials are deeper than those used for nucleon-nucleus scattering because of the strong absorption property. Except for collisions involving light nuclei, no very strong evidence for the influence of spin-orbit potentials on elastic scattering yields has been found, but there is evidence suggesting that these may become important for more massive nuclei at energies in excess of 10 MeV per nucleon.

The optical model potential is not a potential in the ordinary sense. When inserted in the wave equation (11), it gives agreement with experiment by simulating the complicated interactions between the two colliding many-body systems in terms of a prescribed

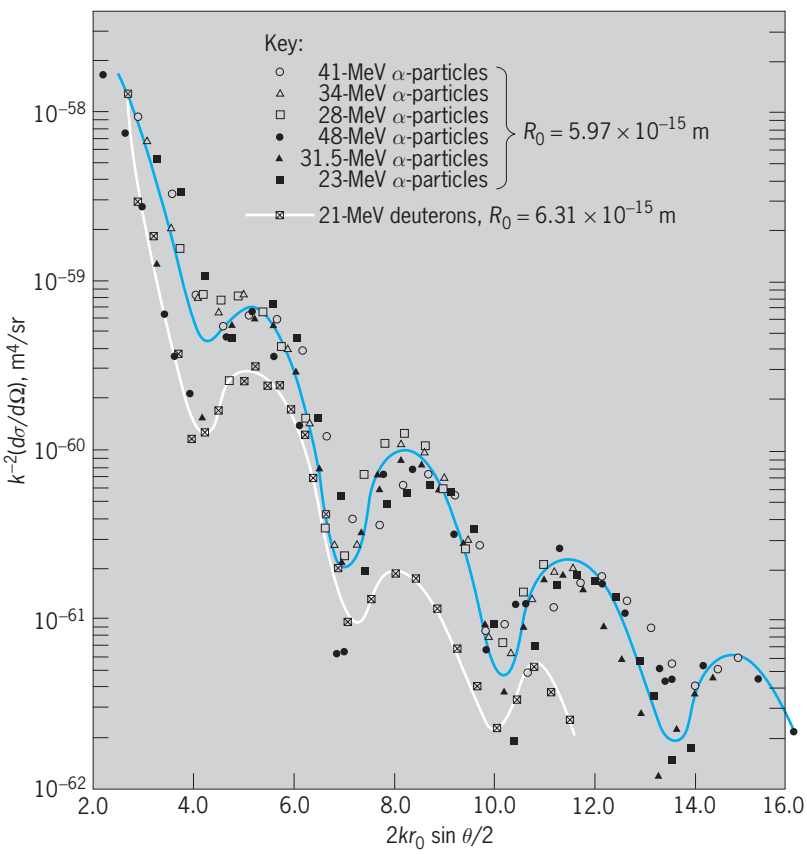

Fig. 7. Angular distribution for the elastic scattering of deuterons and alpha particles from magnesium, showing the typical Fraunhöfer-like diffraction pattern. (*After J. S. Blair et al., Diffraction analysis of elastic and inelastic scattering by magnesium, Nucl. Phys., 17:641–654, 1960*)

functional form. Possible variations in this form, associated with the varying proximity of the two nuclei during the collision, are generally neglected. Alternative methods, such as the time-dependent Hartree-Fock approach, are being developed to treat nucleus-nucleus interactions from a more fundamental perspective.

Nucleus-nucleus scattering sometimes exhibits anomalous behavior which cannot be reproduced with potential models and which is believed to reflect the transient formation of moleculelike configurations. These configurations, or resonances, signal their presence by producing a rapid variation in the energy dependence of the scattering yields and by modifying the behavior of the angular distribution data. *See* NUCLEAR MOLECULE.

Meson scattering. The scattering of π mesons (pions) by nucleons has been studied intensively as an example of a strong interaction between a boson and a baryon, and particularly because of the connection with nucleon-nucleon interactions. Prominent resonances occur in pion-nucleon scattering yields, the best known of which appears for pion bombarding energies of about 200 MeV. Charge independence in pion-nucleon interactions is confirmed by the scattering measurements and their phase-shift analyses.

The scattering of pions by nuclei has been studied over the energy range extending from below 20 MeV to beyond 60 GeV. At low energies the experiments probe nucleon-nucleon correlations within nuclei. Attempts to relate the scattering behavior to the properties of the nuclear wave functions have been partially successful. Comparison of π^+ with π^- scattering yields provides information concerning differences between proton and neutron distributions within nuclei. Future study of pion-nucleus interactions at high energies is expected to reveal properties of the strong interactions, modified by the nuclear movement, at short distances.

The study of kaon-nucleus interactions, still in its infancy, is currently of interest primarily in connection with the production of nuclei in which, for example, a neutron is replaced by a lambda particle (Λ), and the investigation of Λ-nucleon and related interactions. *See* MESON. K. A. Erb

Bibliography. R. C. Barrett and D. F. Jackson, *Nuclear Sizes and Structure*, 1977; G. Breit and R. D. Haracz, Nucleon-nucleon scattering, in E. H. S. Burhop (ed.), *High-Energy Physics*, vol. 1, 1967; D. M. Brink, *Semi-Classical Methods for Nucleus-Nucleus Scattering*, 1986; M. Goldberger and K. M. Watson, *Collision Theory*, 1964, reprint 1975; G. R. Satchler, *Introduction to Nuclear Reactions*, 2d ed., 1990.

Scattering layer

A layer of organisms in the sea which causes sound to scatter and returns echoes. Recordings by sonic devices of echoes from sound scatterers indicate that the scattering organisms are arranged in approximately horizontal layers in the water, usually well above the bottom. The layers are found in both shallow and deep water.

Shallow water. In the shallow water of the continental shelves (less than 600 ft or 200 m deep), scattering layers and echoes from individuals or compact groups are very irregularly distributed and are probably made up of a variety of sea animals and possibly plants. While some fishes, notably herring, and some zooplankton have been identified by fishing, many others have not been identified.

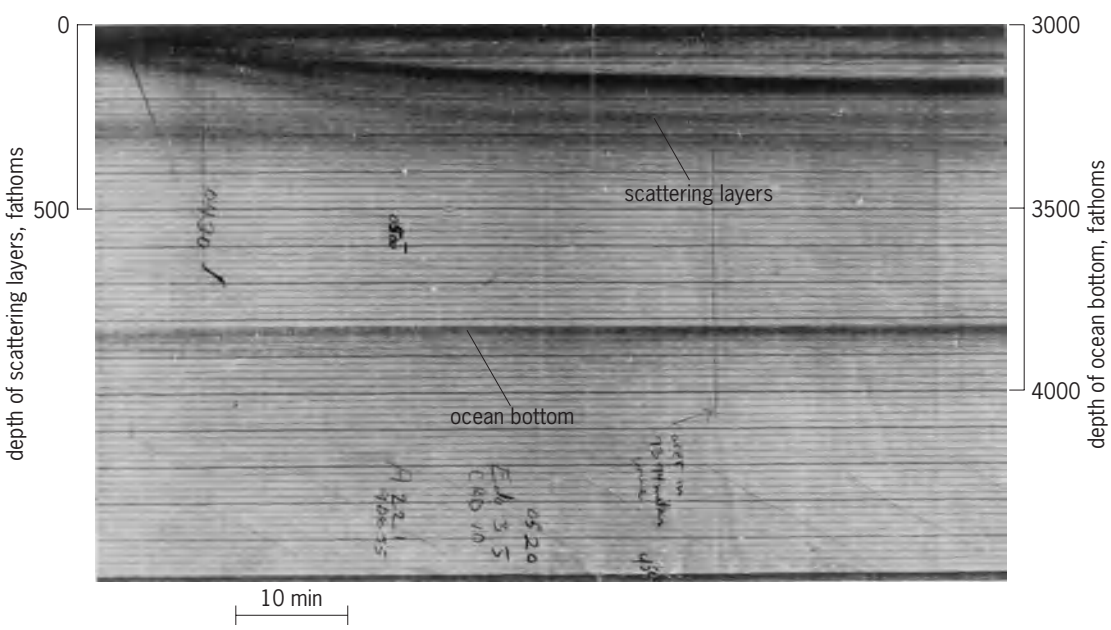

Fig. 1. Record by a 12-kHz echo sounder of the sunrise descent of deep scattering layers in eastern Pacific off northern Chile. A layer, which appears to have remained at depth throughout the night, is shown near 300 fathoms (540 m). 1 fathom = 1.8 m. (*From M. N. Hill, ed., The Sea, vol. 1, Interscience, 1962*)

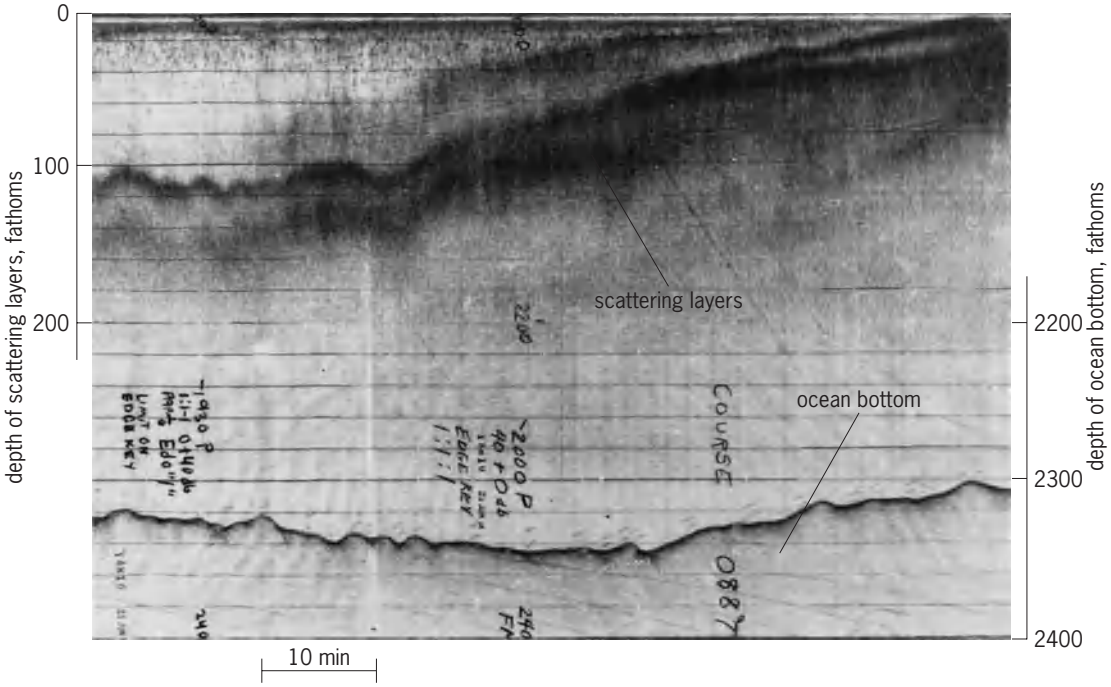

Fig. 2. Sunset ascent of deep scattering layers in western North Atlantic near 40°30′N, 50°W. 1 fathom = 1.8 m. (*From M. N. Hill, ed., The Sea, vol. 1, Interscience, 1962*)

Deep scattering layers. In deep water (greater than 600 ft or 200 m) one or more well-defined layers generally are present. Though commonly variable in detail, they are found to be very widely distributed. They are readily detected by echo-sounding equipment operating in the frequency range 3–60 kHz, the sound spectrum of each layer generally having maximum scattering at a somewhat different frequency than others found at the same place. Commonly, but not universally, the deep-water layers migrate vertically in apparent response to changes in natural illumination. The most pronounced migration follows a diurnal cycle, the layers rising at night, sometimes to the surface, and descending to greater depths during the day (**Figs. 1** and **2**). The common range of daytime depths is 600–2400 ft (200–800 m). The migration is modified by moonlight and has been observed to be modified during the day by heavy local cloud cover, for example, a squall. Occurrence of the layers in deep water was first demonstrated by C. Eyring, R. Christiansen, and R. Raitt. *See* ECHO SOUNDER; UNDERWATER SOUND.

Deep scattering organisms. All animals and plants, as well as nonliving detritus, contrast acoustically with seawater and, hence, any may be responsible for observed scattering in a particular instance. Many animals and plants have as part of their natural structure a gas-filled flotation organ which scatters sound many times more strongly than would be inferred from the sound energy they intercept. These are the strongest scatterers of their size. M. W. Johnson pointed out in 1946 that the layers which migrate diurnally must be animals that are capable of swimming to change their depth, rather than plant life or some physical boundary such as an abrupt

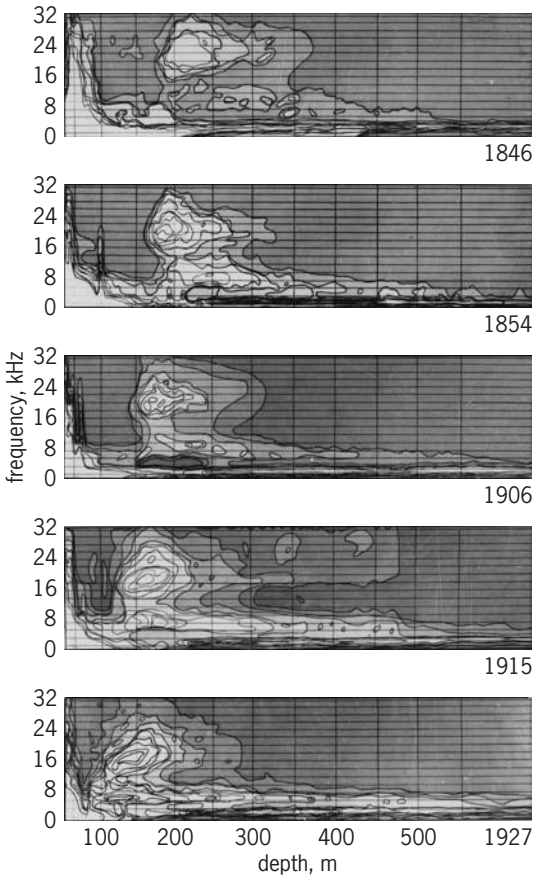

Fig. 3. Sequence of sunset observations showing scattering as a function of depth and frequency. Contours of equal sound level are 2 dB apart, with lightest areas denoting highest levels. Time of day of each observation is indicated by the number at the lower right-hand corner of each record. 1 m = 3.3 ft. (*From M. R. Hill, ed., The Sea, vol. 1, Interscience, 1962*)

temperature change in the water. In 1953 V. C. Anderson demonstrated that some of the deep-water scatterers have a much smaller acoustical impedance than seawater. This fact fits the suggestion by N. B. Marshall in 1951 that the scatterers may be small fishes with gas-filled swim bladders, many of which are known to be geographically distributed much as the layers are. In 1954 J. B. Hersey and R. H. Backus found that the principal layers in several localities migrate in frequency of peak response while migrating in depth, thus indicating that the majority of scatterers fit Marshall's suggestion (**Fig. 3**). In deep water, corroboration of Marshall's suggestion has come from several independent combined acoustical and visual observations made during dives of the deep submersibles *Trieste*, *Alvin*, *Soucoupe*, *Deep Star*, and others. Nearly all these dives were made within a few hundred miles of the east or west coasts of the continental United States. The fishes observed most commonly to form scattering layers are the lantern fishes, or myctophids, planktonic fishes that are a few inches long and possess a small swim bladder. The siphonophores (jellyfish) also have gas-filled floats and have been correlated with scattering layers by observation from submersibles. It is not clear whether fishes with swim bladders and siphonophores make up the principal constituents of the deep-water scattering layers in deep-ocean areas, but it is nearly certain that they are not exclusively responsible for all the widely observed scattering. *See* DEEP-SEA FAUNA.

John B. Hersey

Bibliography. M. N. Hill (ed.), *The Sea*, vol. 1, 1962, reprint 1982; J. S. Levinton, *Marine Biology: Function, Biodiversity, Ecology*, 2d ed., 2001; G. L. Pickard and W. J. Emery, *Descriptive Physical Oceanography*, 5th ed., 1990.

Scattering matrix

An infinite-dimensional matrix or operator that expresses the state of a scattering system consisting of waves or particles or both in the far future in terms of its state in the remote past; also called the S matrix. In the case of electromagnetic (or acoustic) waves, it connects the intensity, phase, and polarization of the outgoing waves in the far field at various angles to the direction and polarization of the beam pointed toward an obstacle. It is used most prominently in the quantum-mechanical description of particle scattering, in which context it was invented in 1937 by J. A. Wheeler to describe nuclear reactions. Because an analog of the Schrödinger equation for the description of particle dynamics is lacking in the relativistic domain, W. Heisenberg proposed in 1943 that the S matrix rather than the hamiltonian or the lagrangian be regarded as the fundamental dynamical entity of quantum mechanics. This program played an important role in high-energy physics during the 1960s but is now largely abandoned. The physics of fundamental particles is now described primarily in terms of quantum gauge fields, and these are used to determine the S matrix and its elements for the collision and reaction processes observed in the laboratory. *See* ELEMENTARY PARTICLE; GAUGE THEORY; NONRELATIVISTIC QUANTUM THEORY; NUCLEAR REACTION; QUANTUM MECHANICS; RELATIVISTIC QUANTUM THEORY; SCATTERING EXPERIMENTS (ATOMS AND MOLECULES); SCATTERING EXPERIMENTS (NUCLEI).

Definition and properties. A normalized quantum wave packet $\Psi_{in}(t)$ may be defined that describes a free particle, with specified quantum numbers (such as spin and isospin) and almost fixed energy and momentum (almost, because fixing the energy or momentum would preclude normalization); its development in time is governed by the free Schrödinger equation. In the presence of an obstacle, however, the actual time development of a physical wave packet is governed by a Schrödinger equation that includes the interaction force. Thus, a physical wave packet $\Psi_+(t)$ is uniquely determined by the full hamiltonian that governs its time development and by the freely developing packet $\Psi_{in}(t)$ from which it asymptotically originated. A physical wave packet $\Psi_-(t)$ can equally well be defined that develops according to the same full hamiltonian but that will asymptotically in the far future become equal to the free packet $\Psi_{out}(t)$ with specified quantum numbers. The quantum-mechanical probability amplitude for finding a particle with the quantum numbers, energy, and momentum described by Ψ_{out} in the far future, given that in the distant past it had those specified by Ψ_{in}, is therefore equal to the probability amplitude that the particle in the state $\Psi_+(t)$ now is in the state $\Psi_-(t)$, which is the inner product of $\Psi_+(t)$ and $\Psi_-(t)$. This amplitude may also be expressed as the matrix element corresponding to Ψ_{in} and Ψ_{out} of the scattering operator S, which maps the initial free state onto the final free state into which it develops.

If the situation is idealized to a sharp energy E in the initial free state (in which case the wave packet cannot be normalized), the energy will also be equal to E in the final free state; this is the quantum version of the conservation of energy. The S matrix is then defined after factoring out the energy-conserving Dirac delta function. The resulting matrix has partly discrete indices (for the discrete quantum numbers) and partly continuous ones (for the momenta). Since the number of particles in nonrelativistic quantum mechanics is conserved, the norm of the final state must equal that of the initial one, and no possible final states are excluded. This implies that the S matrix is unitary. (In the case of wave scattering, the unitarity property is the result of the conservation of energy.) *See* CONSERVATION OF ENERGY; MATRIX THEORY.

The quantity of greatest physical interest is the probability amplitude for the presence of the target to cause scattering in a given direction. This, up to a kinematical factor, is the scattering or reaction amplitude: it is the deviation of the S matrix from the unit matrix (no scattering). The unitarity of the S matrix manifests itself as the optical theorem, which expresses the imaginary part of the scattering

amplitude for forward scattering in terms of the total cross section.

Invariance properties of the physical system result in symmetries of the S matrix. An important example is the reciprocity relation, which, on the basis of invariance under a reversal of the sense of time, equates the scattering from the direction $\vec{k}$ to $\vec{k}'$ to the scattering from $-\vec{k}$ to $-\vec{k}'$. *See* SYMMETRY LAWS (PHYSICS); TIME REVERSAL INVARIANCE.

The mathematical properties of the S matrix in nonrelativistic quantum mechanics have been thoroughly studied and are, for the most part, well understood. If the potential energy in the Schrödinger equation, or the scattering obstacle, is spherically symmetric, the eigenfunctions of the S matrix are spherical harmonics and its eigenvalues are of the form $\exp(2i\delta_l)$, where the real number δ_l is the phase shift of angular momentum l. In the nonspherically symmetric case, analogous quantities are called the eigenphase shifts, and the eigenfunctions depend on both the energy and the dynamics. In the relativistic regime, without an underlying Schrödinger equation for the particles, the mathematical properties are not as well known. Causality arguments (no signal should propagate faster than light) lead to dispersion relations, which constitute experimentally verifiable consequences of very general assumptions on the properties of nature that are independent of the detailed dynamics. *See* ANGULAR MOMENTUM; CAUSALITY; DISPERSION RELATIONS; EIGENFUNCTION; SPHERICAL HARMONICS.

Calculation. In the nonrelativistic scattering of one particle by a fixed center of force, or of two particles by one another in their center-of-mass system, the S matrix is calculated by converting the Schrödinger equation with the scattering boundary condition into an integral equation known as the Lippmann-Schwinger equation. If the potential energy is invariant under rotation, angular momentum is conserved, each phase shift can be calculated separately by a separation of variables in the Schrödinger equation, and the S matrix can be computed as a Legendre series, which is known as the partial-wave series. For particles interacting via spin-dependent forces, the procedure is similarly based on solving systems of coupled Schrödinger equations that describe the spin-induced coupling between orbital angular momenta. Under certain assumptions on the structure of the forces, purely algebraic methods may be utilized for the calculation of S-matrix elements. In the more general case of noncentral forces, they have to be computed numerically from the solution of the three-dimensional Schrödinger equation. In the calculation of the S matrix for three or more interacting particles, the Lippmann-Schwinger equation does not determine the answer uniquely. This problem was overcome by the Faddeev equations and other subsequently developed methods. *See* LEGENDRE FUNCTIONS.

Actual calculations almost always must resort to approximation methods. These include low-energy approximations based on the rapid convergence of the partial-wave series at low energies; high-energy approximations analogous to ray optics with diffraction (WKB approximation); perturbation theory based on assumed weakness of the forces (Born approximation) or on a small deviation from forces for which the results are known (distorted-wave Born approximation); variational methods; and procedures based on physical intuition and insight. *See* PERTURBATION (QUANTUM MECHANICS); WENTZEL-KRAMERS-BRILLOUIN METHOD.

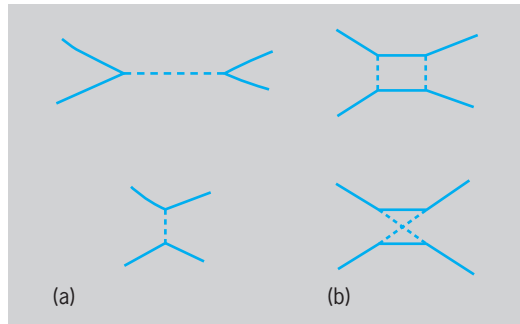

Typical Feynman diagrams for two-particle scattering. The solid lines are fermions, the broken lines bosons. (*a*) Second-order diagrams. (*b*) Fourth-order diagrams.

In quantum field theory, the S matrix, when it is practical to calculate it at all, is most commonly calculated by means of Feynman diagrams (see **illus.**). Each such diagram, which schematically represents the so-called virtual emission and reabsorption of force-transmitting bosons (in quantum electrodynamics, photons) can be translated by a general rule into a multiple integral that usually must be computed numerically. In the case of string theory, each line in such a diagram becomes a two-dimensional sheet. In all realistic cases, many of the integrals diverge and the S matrix has to be renormalized, a procedure, first introduced in quantum electrodynamics, in which a finite number of quantities that are infinite in the perturbation calculation are replaced by their experimental values. The strength of the coupling that describes the strong nuclear forces, or of the weak nuclear force at high energies, is such that straightforward expansions in terms of Feynman diagrams are impractical and it is necessary to resort to partial summations of infinitely many of them. In some cases, the requirement of unitarity of the S matrix is an important constraint. *See* FEYNMAN DIAGRAM; QUANTUM CHROMODYNAMICS; QUANTUM ELECTRODYNAMICS; QUANTUM FIELD THEORY; RENORMALIZATION; WEAK NUCLEAR INTERACTIONS.

Another method of calculation is to replace the continuous physical space-time by a lattice of discrete points. In that case, no infinities arise, but very large computers are required for meaningful results. The problem of approaching the continuum limit is difficult and not completely solved.

Inverse scattering. The inverse-scattering problem is to determine, if possible, the underlying force (or the shape of an obstacle) from the S matrix, which is assumed to be known from experiments. It was first solved for central potentials in the one-particle

Schrödinger equation on the assumption that what is known is one phase shift for all energies and the binding energies and so-called norming constants of the same angular momentum. The problem for central potentials and all phase shifts known at one energy was subsequently solved. The inverse-scattering problem for noncentral forces, in which the S matrix is given as a function of the energy and the directions of the incoming and scattered particles, is partially solved. *See* INVERSE SCATTERING THEORY.

Roger G. Newton

Bibliography. W. O. Amrein et al., *Scattering Theory in Quantum Mechanics*, 1977; K. Chadan and P. C. Sabatier, *Inverse Problems in Quantum Scattering Theory*, 1989; E. Leader and E. Predazzi, *An Introduction to Gauge Theories and the "New Physics,"* 1982; R. G. Newton, *Scattering Theory of Waves and Particles*, 2d ed., 1982; A. M. Polyakov, *Gauge Fields and Strings*, 1987; M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, vol. 3: *Scattering Theory*, 1979.

Scattering of electromagnetic radiation

The process in which energy is removed from a beam of electromagnetic radiation and reemitted with a change in direction, phase, or wavelength. All electromagnetic radiation is subject to scattering by the medium (gas, liquid, or solid) through which it passes. In the short-wavelength, high-energy regime in which electromagnetic radiation is most easily discussed by means of a particle description, these processes are termed photon scattering. At slightly longer wavelengths, the scattering of x-rays provides the most effective means of determining the structure of crystalline solids. In the visible wavelength region, scattering of light produces the blue sky, red sunsets, and white clouds. At longer wavelengths, scattering of radio waves determines their characteristics as they pass through the atmosphere. *See* LIGHT; METEOROLOGICAL OPTICS; RADIO-WAVE PROPAGATION.

It has been known since the work of J. Maxwell in the nineteenth century that accelerating electric charges radiate energy and, conversely, that electromagnetic radiation consists of fields which accelerate charged particles. Light in the visible, infrared, or ultraviolet region interacts primarily with the electrons in gases, liquids, and solids—not the nuclei. The scattering process in these wavelength regions consists of acceleration of the electrons by the incident beam, followed by reradiation from the accelerating charges. *See* ELECTROMAGNETIC RADIATION.

Scattering processes may be divided according to the time between the absorption of energy from the incident beam and the subsequent reradiation. True "scattering" refers only to those processes which are essentially instantaneous. Mechanisms in which there is a measurable delay between absorption and reemission are usually termed luminescence. If the delay is longer than a microsecond or so, the process may be called fluorescence; and mechanisms involving very long delays (seconds) are usually termed phosphorescence. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; FLUORESCENCE; LUMINESCENCE; PHOSPHORESCENCE.

Inelastic scattering. Instantaneous scattering processes may be further categorized according to the wavelength shifts involved. Some scattering is "elastic"; there is no wavelength change, only a phase shift. In 1928 C. V. Raman discovered the process in which light was inelastically scattered and its energy was shifted by an amount equal to the vibrational energy of a molecule or crystal. Such scattering is usually called the Raman effect. This term has been used in a more general way, however, often to describe inelastic scattering of light by spin waves in magnetic crystals, by plasma waves in semiconductors, or by such exotic excitations as "rotons," the elementary quanta of superfluid helium. *See* RAMAN EFFECT.

Brillouin and Rayleigh scattering. In liquids or gases two distinct processes generate inelastic scattering with small wavelength shifts. The first is Brillouin scattering from pressure waves. When a sound wave propagates through a medium, it produces alternate regions of high compression (high density) and low compression (or rarefaction). A picture of the density distribution in such a medium is shown in **Fig. 1**. The separation of the high-density regions is equal to one wavelength λ for the sound wave propagating. Brillouin scattering of light to higher (or lower) frequencies occurs because the medium is moving toward (or away from) the light source. This is an optical Doppler effect, in which the frequency of sound from a moving object is shifted up in frequency as the object moves toward the observer. *See* DOPPLER EFFECT.

The second kind of inelastic scattering studied in fluids is due to entropy and temperature fluctuations. In contrast to the pressure fluctuations due to sound waves, these entropy fluctuations do not generate scattering at sharp, well-defined wavelength shifts from the exciting wavelength; rather, they produce a broadening in the scattered radiation centered about the exciting wavelength. This is because entropy fluctuations in a normal fluid are not propagating and do not, therefore, have a characteristic frequency; unlike sound they do not move like

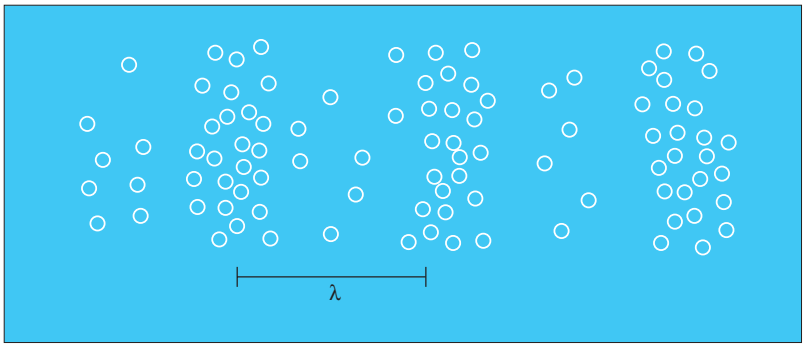

Fig. 1. Pressure-induced density fluctuations in a fluid.

waves through a liquid—instead, they diffuse. Under rather special circumstances, however, these entropy or temperature fluctuations can propagate; this has been observed in super-fluid helium and in crystalline sodium fluoride at low temperatures, and is called second sound. *See* ENTROPY; SECOND SOUND.

Scattering from entropy or temperature fluctuations is called Rayleigh scattering. In solids this process is obscured by scattering from defects and impurities. Under the assumption that the scattering in fluids is from particles much smaller than the wavelength of the exciting light, Lord Rayleigh derived in 1871 an equation, shown below, for such scattering; here $I(\theta)$ is the intensity of light scattered from an incident beam of wavelength λ and intensity I_0 at a distance r; d is the number of scattering particles; υ is the volume of the disturbing particle; and n is the index of refraction of the fluid. The $\cos\theta$ term is present for unpolarized incident light, where θ is the scattering angle. Measurement of the ratio in Rayleigh's equation allows the determination of either Avogadro's number N or the molecular weight M of the fluid, if the other is known. The dependence of scattering intensity upon the inverse fourth power of the wavelength given in Rayleigh's equation is responsible for the fact that daytime sky looks blue and sunsets red: blue light is scattered out of the sunlight by the air molecules more strongly than red; at sunset, more red light passes directly to the eyes without being scattered.

$$\frac{r^2 I(\theta)}{I_0} = \pi d \lambda^{-4} \upsilon^2 (1 + \cos^2\theta)(n-1)^2$$

Large particles. Rayleigh's derivation of his scattering equation relies on the assumption of small, independent particles. Under some circumstances of interest, both of these assumptions fail. Colloidal suspensions provide systems in which the scattering particles are comparable to or larger than the exciting wavelengths. Such scattering is called the Tyndall effect and results in a nearly wavelength-independent (that is, white) scattering spectrum. The Tyndall effect is the reason clouds are white (the water droplets become larger than the wavelengths of visible light). *See* TYNDALL EFFECT.

The breakdown of Rayleigh's second assumption—that of independent particles—occurs in all liquids. There is strong correlation between the motion of neighboring particles. This leads to fixed phase relations and destructive interference for most of the scattered light. The remaining scattering arises from fluctuations in particle density discussed above and was first analyzed theoretically by A. Einstein in 1910 and by M. Smoluchowski in 1908.

Rayleigh's basic theory has been extended by several authors. Rayleigh in 1911 and R. Gans in 1925 derived scattering formulas appropriate for spheres of finite size, and in 1947 P. Debye extended the theory to include random coil polymers. The combined results of these three workers, generally called the Rayleigh-Gans-Debye theory, is valid for any size of particle, provided the refractive index n is near unity; whereas the Rayleigh theory is valid for any index, provided the particles are very small. The index n of any medium would be unity if there were no scattering.

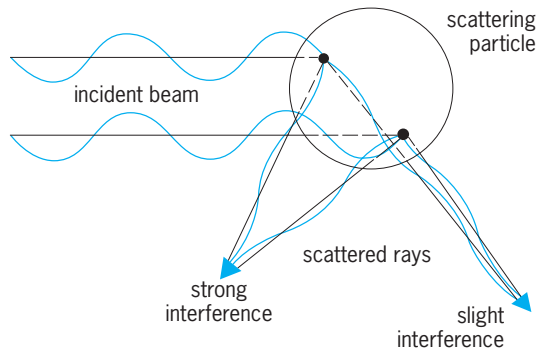

Fig. 2. Role of interference in determining the angular dependence of light scattering from particles that are comparable in size to the wavelength.

A more complete theory than Rayleigh-Gans-Debye was actually developed earlier by G. Mie in 1908; however, Mie's theory generally requires numerical solution. Mie's theory is valid even for particles larger than the wavelength of light used. For such particles large phase shifts occur for the scattered light. Mie scattering exhibits several maxima and minima as a function of scattering angle; the positions of these maxima depend upon particle size, as indicated in **Fig. 2**. These secondary maxima are essentially higher-order Tyndall scattering.

Critical opalescence. The most striking example of light scattering is that of critical opalescence, discovered by T. Andrews in 1869. As the critical temperature of a fluid is approached along the critical isochore, fluctuations in the density of the medium become larger than a wavelength and stay together for longer times. The independent particle approximation becomes very bad, and the intensity of light scattered in all directions becomes very large. The fluid, which might have been highly transparent at a temperature 1°C (1.8°F) higher, suddenly scatters practically all of the light incident upon it. The regions of liquidlike density in the fluid have molecules moving in phase, or "coherently," within them. Such coherent regions in fluids near their critical temperatures behave very much like the large spherical particles diagrammed in Fig. 2, thus establishing the relationship between Mie scattering of colloids and critical opalescence. The theory of scattering from correlated particles was developed primarily by L. S. Ornstein and F. Zernike in 1914–1926. *See* COHERENCE; CRITICAL PHENOMENA; OPALESCENCE.

Nonlinear scattering. At the power densities available with pulsed lasers, a variety of other scattering mechanisms can become important. (For nonlaser excitation these processes exist but are generally too weak to be detected.) Such processes include second-harmonic generation, in which the energy from two quanta (photons) in the incident beam produces one doubly energetic photon. A practical example is the production of green light from

the red light of a ruby laser. Third-harmonic generation (and similar higher-order processes) also occurs as an intense scattering mechanism at sufficiently high power densities. Whereas second-harmonic generation can occur only in certain crystals, for reasons of symmetry, third-harmonic generation occurs in any medium. Other more exotic nonlinear scattering mechanisms include sum generation and difference generation, in which two beams of light are combined to produce scattering having frequencies which are the sum or difference of the frequencies in the two incident beams, and rather weak processes such as the hyper-Raman effect, in which the scattered light has a frequency which is twice that of the incident beam plus (or minus) some vibrational energy. These effects are nonlinear because their intensities vary according to the square or higher power of the incident intensity. These nonlinear processes may be very important in astrophysics (for example, in stellar interiors), in addition to laboratory laser experiments. *See* NONLINEAR OPTICS. J. F. Scott

Bibliography. B. J. Berne and R. Pecora, *Dynamic Light Scattering*, 1976, reprint 1990; W. Brown (ed.), *Dynamic Light Scattering*, 1993; W. Hayes and R. Loudon, *Scattering of Light by Crystals*, 1978; K. S. Schmitz, *An Introduction to Dynamic Light Scattering by Macromolecules*, 1990.

Scent gland

A specialized skin gland of the tubuloalveolar or acinous variety found in many mammals. These glands produce substances having peculiar odors. In some instances they are large, in others small. Examples of large glands are the civet gland in the civet cat, the musk gland in the musk deer, and the castoreum gland in the beaver. The civet gland is an anal gland, whereas the musk and castoreum are preputial. Examples of small scent glands are the preputial or Tyson's glands in the human male which secrete the smegma, and the vulval glands in the female. The secretions in all of the above glands are sebaceous.

Many other vertebrates have glands whose secretions give off various types of odors. The mucus-secreting skin glands of fishes produce their fishy odor. Amphibia have glands which emit pungent, sweetish, or onionlike odors. Many urodeles have specialized courtship or hedonic glands. These amphibian glands are mucus-secreting. Scent glands are protective devices for many animals; in some species they serve to attract members of the same species or the opposite sex. The femoral glands on the inner aspect of the upper region of the hindlimbs of male lizards are specialized sebaceous glands associated with copulation. They give off a musty odor. The uropygial or preen glands opening upon the upper tail surface of birds produce an odorous, oily material used to waterproof feathers. *See* EPITHELIUM; GLAND; UROPYGIAL GLAND. Olin E. Nelsen

Scheduling

A decision-making function that plays an important role in most manufacturing and service industries. Scheduling is applied in procurement and production, in transportation and distribution, and in information processing and communication. A scheduling function typically uses mathematical optimization techniques or heuristic methods to allocate limited resources to the processing of tasks.

Application domains. There are many types of scheduling problems. The main categories include project scheduling, production or job shop scheduling, timetabling or rostering, and work-force scheduling. Each one of these categories can be subdivided.

Project scheduling is concerned with a set of activities that are subject to precedence constraints, specifying which jobs have to be completed before a given job is allowed to start its processing. All activities belong to a single (and typically large) project that has to be completed in a minimum time; for example, a large real estate development or the construction of an aircraft carrier.

Production or job shop scheduling is important in manufacturing settings, for example, semiconductor manufacturing. Customer orders have to be executed. Each order entails a number of operations that have to be processed on the resources or the machines available. Each order has a committed shipping date that plays the role of a due date. Production scheduling often also includes lot sizing and batching.

Timetabling occurs often in class room scheduling, scheduling of meetings, and reservation systems. In many organizations, especially in the service industries, meetings must be scheduled in such a way that all necessary participants are present; often other constraints have to be satisfied as well (in the form of space and equipment needed). Such problems occur in schools with classroom and examination scheduling as well as in the renting of hotel rooms and automobiles.

Work-force scheduling (crew scheduling, and so on) is increasingly important, especially in the service industries. For example, large call centers in many types of enterprises (airlines, financial institutions, and others) require the development of complicated personnel scheduling techniques.

Modeling. In order to determine satisfactory or optimal schedules, it is helpful to formulate the scheduling problem as a mathematical model. Such a model typically describes a number of important characteristics. One characteristic specifies the number of machines or resources as well as their interrelationships with regard to the configuration, for example, machines set up in series, machines set up in parallel. A second characteristic of a mathematical model concerns the processing requirements and constraints. These include setup costs and setup times, and precedence constraints between various activities. A third characteristic has to do with the objective that has to be optimized, which may be a

single objective or a composite of different objectives. For example, the objective may be a combination of maximizing throughput (which is often equivalent to minimizing setup times) and maximizing the number of orders that are shipped on time.

These mathematical models can often be formulated as mathematical programs, for example, linear programs, integer programs, or disjunctive programs. The mathematical programs typically exhibit a special structure that can be exploited when solving them. For example, the coefficient matrices of the integer programs used to formulate work-force scheduling problems typically consist of only zeroes and ones.

Computational complexity and algorithms. Most scheduling problems in theory as well as in practice are extremely hard to solve. In the theory of computational complexity, the hard problems are referred to as NP-hard problems (NP for nondeterministic polynomial time), whereas the easy problems are said to be polynomial time solvable. NP-hard implies that the computation time needed to find an optimal solution increases exponentially and not polynomially with the size of the instance of the problem (the size of an instance of a given problem could, for example, be measured by the number of jobs that have to be scheduled). The computation time needed to find an optimal solution for a problem that is polynomial time solvable increases polynomially in the size of an instance. The computational complexity of scheduling problems, that is, whether they are polynomial time solvable or NP-hard, can often be determined fairly easily. The majority of the scheduling problems studied in the literature are NP-hard.

Because of the many different types of scheduling problems and also because of the fact that any given problem type has instances of varying sizes, there are many different types of algorithms. The types can be categorized in several ways. One way is based on whether the algorithm is designed to find an optimal solution or whether it is designed to find just a good solution. An algorithm of the first type may have to consider all possible schedules and may be enumerative, whereas an algorithm of the second type is typically based on heuristics (algorithmic procedures that are based on common sense, but which do not guarantee an optimal solution).

A different way of categorizing the algorithms is based on the following: Some algorithms are of the constructive type, that is, they generate schedules one job at the time; whereas other algorithms are of the improvement type, that is, they start out with a complete schedule and then try to improve on it step by step, performing certain manipulations.

A class of algorithms that is widely used in practice consists of the so-called dispatching rules, typically heuristics that generate schedules constructively. A dispatching rule is a priority rule that sets priorities for the jobs available for processing according to a given function of the attributes of the jobs. A well-known priority rule is the so-called weighted shortest processing time first rule. This rule orders the jobs in a way that takes into account the processing time p of a job as well as the weight (or importance factor) w. The rule gives the job with the highest ratio of w/p the highest priority.

Another widely used class of algorithms uses so-called branch-and-bound procedures. Branch-and-bound is an established way of carrying out complete enumeration. An algorithm of this type is a constructive procedure designed to find an optimal solution.

A third class of algorithms consists of local search procedures. A local search routine is a heuristic of the improvement type. It starts out with a complete schedule that is feasible. Then it attempts to find better schedules through a series of well-defined manipulations, for example, adjacent pairwise interchanges. The three-best known types of local search procedures are simulated annealing, tabu search, and genetic algorithms. Simulated annealing attempts to find a better schedule than the best schedule generated so far by selecting a schedule in the neighborhood or proximity of the current schedule. Simulated annealing performs this neighborhood search in a random manner. Tabu search is similar to simulated annealing, but performs its neighborhood search in a nonrandom manner. A genetic algorithm selects, each time, multiple new schedules simultaneously.

Many algorithms that are implemented in practice are hybrids. For example, they may use a dispatching rule to find an initial feasible schedule and then feed this solution into a local search procedure that attempts to find solutions that are better than the initial solution.

Scheduling system design. The scheduling function is often incorporated in a system that is embedded in the information infrastructure of the organization. This infrastructure may be an enterprise-wide information system that is connected to the main databases of the company. Many other decision support systems may be plugged into such an enterprise-wide information system—for example, forecasting, order promising and due date setting, and material requirements planning (MRP).

The scheduling system may rely on its own piece of hardware, which may be a workstation with a monitor that has special graphics capabilities. The scheduling system may have its own local database that is linked with other databases in the company and possibly also with the Internet.

The database that the scheduling system relies on usually has some special characteristics. It has static data as well as dynamic data. The static data—for example, processing requirements, product characteristics, and routing specifications—are fixed and do not depend on the schedules developed. The dynamic data are schedule-dependent; they include the start times and completion times of all the operations on all the different machines, and the length of the setup times (since these may also be schedule-dependent).

The user interfaces of scheduling systems may often be very elaborate. They almost always include an electronic Gantt chart (horizontal bar chart displaying the time relationships between the different tasks in a project), which allows the user to click on operations, drag them, and drop them at other

machines and in other time periods. The system often has elaborate capabilities of readjusting itself after considering all the changes made by the scheduler.

The scheduling engine often has a library of algorithms. Such a library may contain priority rules and local search techniques, as well as more specialized techniques that take the bottleneck dynamics of the environment into account.

The scheduling engine often allows the user to provide some input. For example, the user may specify the amount of time (in seconds) that an algorithm, for example, a local search procedure, is allowed to run. The user may provide even more input such as specifying how one algorithm should be coupled with another. For example, first a priority rule is used in order to generate a schedule, and then this schedule is fed into a local search procedure that tries to find a better schedule.

Economic impact. The economic impact of scheduling is significant. In certain industries the viability of a company may depend on the effectiveness of its scheduling systems, for example, airlines and semiconductor manufacturing. Good scheduling often allows an organization to conduct its operations with a minimum of resources. In the case of a large airline, it may be possible to free-up the use of a plane just by more efficient scheduling. If a manufacturing company has established a reputation for on-time deliveries of customer orders, the company can charge a premium to its customers (which may be an additional 2–3%). *See* MATERIAL RESOURCE PLANNING; PRODUCTION PLANNING.

Michael Pinedo; Sridhar Seshadri

Bibliography. M. Pinedo, *Scheduling: Theory, Algorithms, and Systems*, 1995; M. Pinedo and X. Chao, *Operations Scheduling with Applications in Manufacturing and Services*, 1999.

Scheelite

A mineral consisting of calcium tungstate, $CaWO_4$. Scheelite occurs in colorless to white, tetragonal crystals (see **illus.**); it may also be massive and granular. Its fracture is uneven, and its luster is vitreous to adamantine. Scheelite has a hardness of 4.5–5 on Mohs scale and a specific gravity of 6.1. Its streak is white. The mineral is transparent and fluoresces bright bluish-white under ultraviolet light.

Scheelite may contain small amounts of molybdenum. It is an important tungsten mineral and occurs principally in contact metamorphosed deposits (tactite) associated with garnet, diopside, tremolite, epidote, wollastonite, sphene, molybdenite, and fluorite, with minor amounts of pyrite and chalcopyrite. It also occurs in small amounts in vein deposits. The most important scheelite deposit in the United States is near Mill City, Nevada. *See* TUNGSTEN.

Edward C. T. Chao

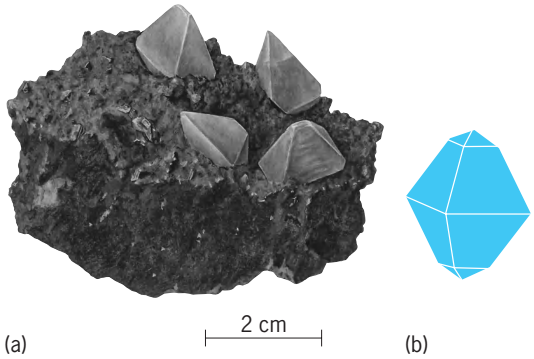

Scheelite. (*a*) Crystals with chalcopyrite ore from Piedmont, Italy (*specimen from Department of Geology, Bryn Mawr College*). (*b*) Crystal habit (*after C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*)

Schematic drawing

Concise, graphical symbolism whereby the engineer communicates to others the functional relationship of the parts in a component and, in turn, of the components in a system. The symbols do not attempt to describe in complete detail the characteristics or physical form of the elements, but they do suggest the functional form which the ensemble of elements will take in satisfying the functional requirements of the component. They are different from a block diagram in that schematics describe more specifically

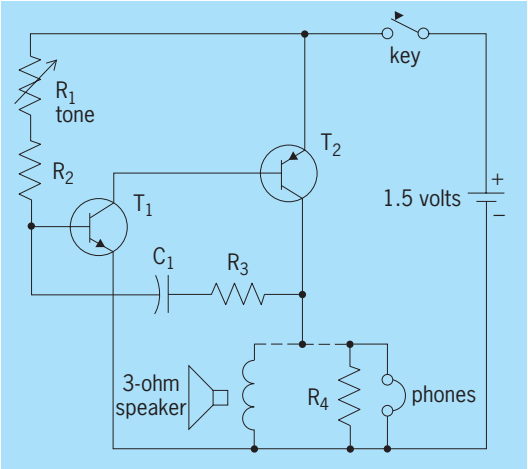

Fig. 1. Simple transistorized code practice oscillator, using standard symbols. (*After J. Markus, Sourcebook of Electronic Circuits, McGraw-Hill, 1968*)

the physical process by which the functional specifications of a block diagram are satisfied. Rather than expressing a mathematical relationship between, for example, an input and an output variable as in a block diagram, a schematic illustrates the physical principles and techniques by which the mathematical requirements of the element are realized. For instance, the schematic indicates whether electrical, hydraulic, mechanical, or pneumatic techniques are employed, and suitable symbols indicate the appropriate elements, such as batteries, resistors, valves, gearing, vacuum tubes, and motors.

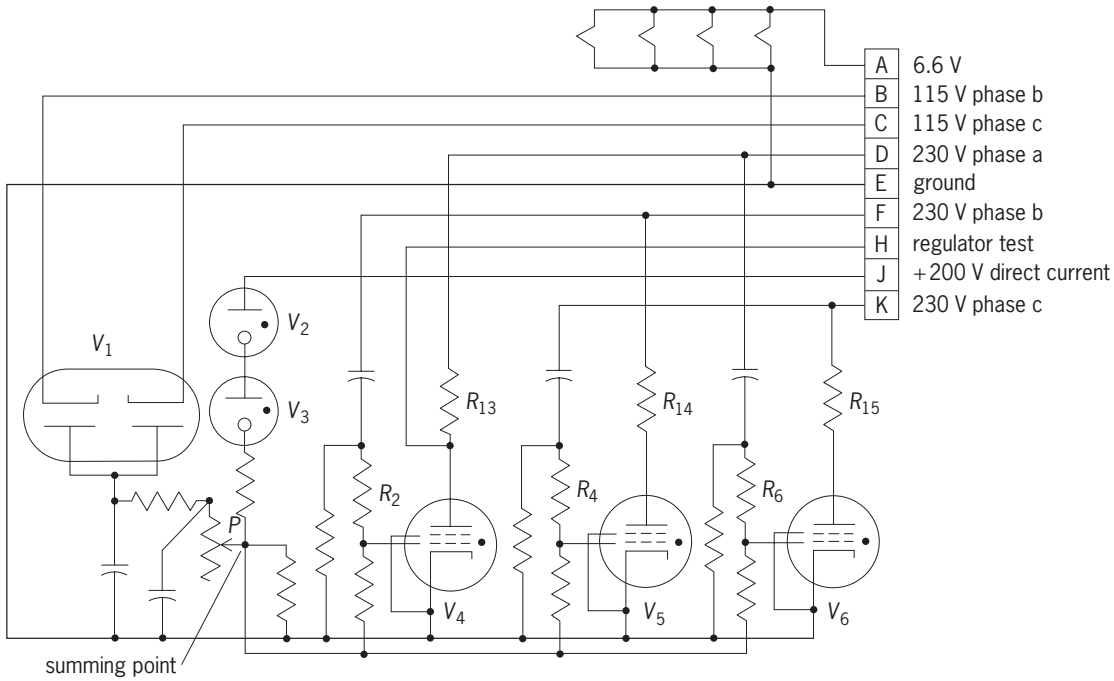

Fig. 2. Electrical schematic of voltage regulator.

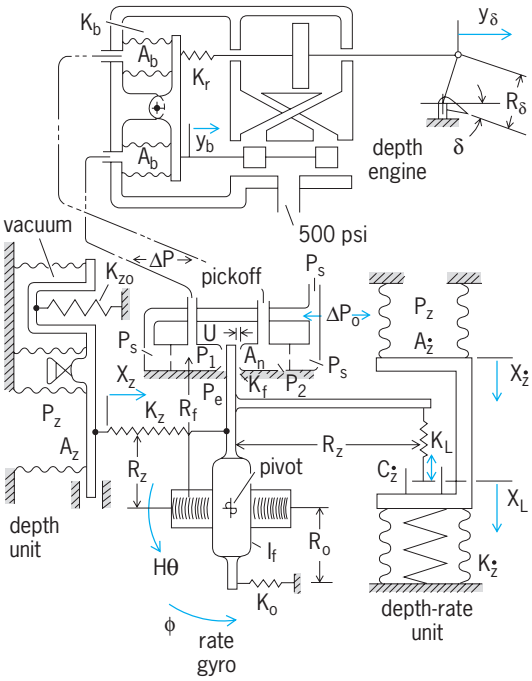

Subscripts:

b = differential–pressure bellows
e = environment
f = flapper
L = depth-rate linkage
n = nozzles
o = ground, or reference
r = ram feedback
s = supply
z = depth unit
ż = depth-rate unit
δ = elevator

Fig. 3. Mechanical schematic of the depth-control mechanism of a torpedo.

Electrical schematic. An electrical schematic is a functional schematic which defines the interrelationship of the electrical elements in a circuit, equipment, or system. The symbols describing the electrical elements are stylized, simplified, and standardized to the point of universal acceptance (**Figs. 1** and **2**).

The simple character of the element symbol makes it possible to represent in a small area the interrelationship of the electrical elements in complex systems. This has the double advantage of economy of space and an increased facility of understanding, because one experienced in the symbolism can easily follow the various functional paths in the electrical schematic. The tracing of a signal path through an electrical schematic is considerably enhanced by the existence of more or less accepted rules with regard to the arrangement of the symbols and of the interconnections between the symbols, all contrived to make more lucid the functional interrelationship of the elements.

Mechanical schematic. A mechanical schematic is also a functional schematic. The graphical descriptions of elements of a mechanical system are more complex and more intimately interrelated than the symbolism of an electrical system and so the graphical characterizations are not nearly as well standardized or simplified (**Fig. 3**). However, a mechanical schematic illustrates such features as components, acceleration, velocity, position force sensing, and viscous damping devices. The symbols are arranged in such a manner and with such simplification as to economize on space and to facilitate an understanding of the functional interrelationship of the various components which make up the system. *See* DRAFTING; ENGINEERING DRAWING.

Robert W. Mann

Schist

Medium- to coarse-grained, mica-bearing metamorphic rock with well-developed foliation termed schistosity. Schist is derived primarily from fine-grained, mica-bearing rocks such as shales and slates. The schistosity is formed by rotation, recrystallization, and new growth of mica; it is deformational in origin. The planar to wavy foliation is defined by the strong preferred orientation of platy minerals, primarily muscovite, biotite, and chlorite. The relatively large grain size of these minerals (up to centimeters) produces the characteristic strong reflection when light shines on the rock. *See* BIOTITE; CHLORITE; MUSCOVITE.

The metamorphic grade of schists typically is in the middle to upper part (~660–930°F or ~350–500°C) of the greenschist facies and the lower part (~930–1100°F or ~500–600°C) of the amphibolite facies. In the amphibolite facies, micas react to form the mineral feldspar, which results in a coarsely foliated rock type called a gneiss. The wavy foliation pattern that is common in many schists reflects the presence of large, secondary minerals (such as garnet, staurolite, and chloritoid), which are grouped under the term porphyroblast. *See* AMPHIBOLITE; FACIES (GEOLOGY); FELDSPAR; GNEISS; PORPHYROBLAST; SHALE; SLATE.

Schists are named by the assemblage of minerals that is most characteristic in the field; for example, a garnet-biotite schist contains porphyroblasts of garnet and a schistosity dominated by biotite. Under the microscope, schists commonly show a crenulation fabric (see **illus.**), reflecting the presence of an older foliation that may be depositional in origin or may represent an earlier deformation history. *See* PETROFABRIC ANALYSIS.

Schists can provide important information on the relationship between metamorphism, the occurrence of mineral reactions in response to changing pressure and temperature conditions, and deformation. When porphyroblasts grow, minerals that are either not involved in the reaction (accessory minerals such as graphite, zircon, and monazite) or minerals that are left over because the rock no longer contains the right mix of ingredients (such as quartz and feldspar) may become inclusions. These inclusions may form trails in the porphyroblast defining an internal foliation. The relationship between the internal foliation and the foliation in the matrix surrounding the blast, called the external foliation, gives information on the temporal relationship between deformation and metamorphism.

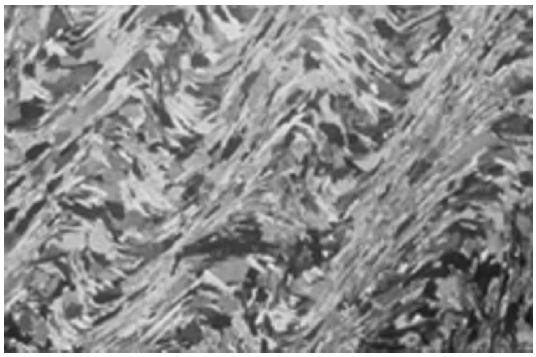

Photomicrograph of a mica schist from the Swiss Alps showing a crenulation fabric. The microscope image was taken under crossed polarized light with a field of view of ~1.5 mm across.

Predeformational metamorphism is characterized by an internal foliation whose shape is unrelated to the external foliation, and typically the two do not connect at the porphyroblast-matrix boundary. At the other extreme lies postdeformational metamorphism, in which an external foliation connects with the internal foliation within the grain without any disruption. For syndeformational growth, that is, when metamorphism and deformation overlap, the external foliation can be traced into a porphyroblast; but often the internal foliation is wavy to sigmoidal in shape. The classic example of the latter relationship is a snowball garnet. After the garnet nucleates, it rapidly grows by consumption and inclusion of minerals in the matrix. Simultaneously, the blast is affected by a shearing motion, much like ball bearings rotating between the hands. As the garnet continues to grow, it produces an internal foliation that appears to be rolled up, forming the characteristic snowball geometry. The rotation that is preserved from this inclusion pattern furthermore provides information on the type of deformation history of the rock (shear-sense indicator). *See* METAMORPHIC ROCKS; METAMORPHISM. Ben A. van der Pluijm

Bibliography. B. A. van der Pluijm and S. Marshak, *Earth Structure: An Introduction to Structural Geology and Tectonics*, Ch. 11, 1997.

Schistosomiasis

A disease in which humans are parasitized by any of four species of blood flukes: *Schistosoma mansoni*, *S. haematobium*, *S. japonicum*, and *S. mekongi*. In contrast to other trematodes, the sexes are separate; that is, male and female reproductive systems occur in individual worms. The disease is also known as bilharziasis. *See* DIGENEA.

The approximate dimensions of the egg and male and female adult in each of the four species are shown in the **table**.

Dimensions of human blood flukes, *Schistosoma*

Species	Egg, μm	Adult	Length, cm	Diameter, mm
S. mansoni	150 × 65	Female	1–1.5	0.25
		Male	2	1
S. haematobium	150 × 60	Female	1–1.5	0.25
		Male	2	1
S. japonicum	80 × 65	Female	1–2	0.25
		Male	1–2.5	1
S. mekongi	45 × 40	Female	1.2	0.23
		Male	1.5	0.41

Distribution. *Schistosoma mansoni* is found in Africa, Brazil, Venezuela, Surinam, the Lesser Antilles, Puerto Rico, and the Dominican Republic. *Schistosoma haematobium*, originally found in Africa, the Near East, and the Mediterranean basin, was introduced into India during World War II. *Schistosoma japonicum* is widely spread in eastern Asia and the southwestern Pacific; however, in Taiwan it infects only animals, not humans. *Schistosoma mekongi* is reported from Laos and Cambodia.

Biology. An embryonated egg passed in feces or urine hatches in fresh water, liberating a miracidium. The ciliated miracidium penetrates into specific gastropod snails. There the miracidium is transformed into a mother sporocyst that gives rise by pedogenesis, or reproduction by larval stages, to several hundred daughter sporocysts. These move away to settle elsewhere in the snail and produce several thousand cercariae. The larval cycle lasts for about 1 month. The cercaria emerges from the mollusk, swims for up to 48 h, and penetrates the skin of the final host upon coming in contact with it. After a phase of growth in the lungs, the schistosomule moves to the intrahepatic veins for further development. Then it moves to its final habitat.

Schistosoma mansoni prefers to live in the inferior mesenteric vein of the lower bowel, particularly the branches of the sigmoid and rectum; *S. haematobium*, in the veins of the urinary bladder; *S. japonicum* and *S. mekongi*, in both the superior and inferior mesenteric veins. *Schistosoma mekongi* is related to *S. japonicum* but differs in that it has smaller eggs, a different intermediate host, and a longer prepatent period in the mammalian host.

For all schistosomes of humans, eggs start to appear within 2 months after exposure. The adult may live for many years. The generalized life cycle is shown in the **illustration**.

Epidemiology. As an intermediate host, *S. mansoni* uses snails of the genus *Biomphalaria*; *S. haematobium*, the genera *Bulinus and Physopsis*; *S. japonicum*, the genus *Onchomelania*; and *S. mekongi*, the genus *Tricula*. *Schistosoma mansoni* seems incapable of infecting *Oncomelania* ssp. *Schistosoma mansoni* is essentially a human parasite. Monkeys may act as natural reservoirs in Africa, and rodents in Brazil and Guadeloupe. *Schistosoma haematobium* has no reservoirs, while *S. japonicum* and *S. mekongi* possess many. Various animals may be infected in the laboratory with the four species.

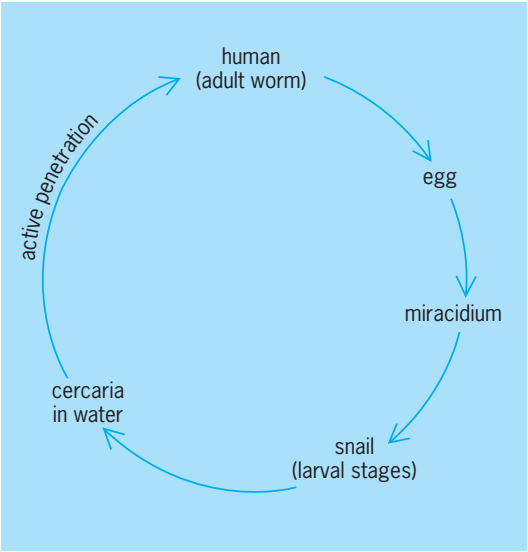

Life cycle of the schistosomes.

Schistosomiasis is an agricultural hazard for humans of all ages in irrigated lands or swamps. Elsewhere fluvial waters are the main source of infection, in which case incidence is marked in human beings who are less than 15 years old, and is higher among boys than among girls. Since human activity is what causes the spread of the disease, it is epidemiologically sensible to say that schistosomiasis is caused by humans, not by snails.

Schistosomiasis is endemic in 74 developing countries, infects 200 million persons, and daily threatens 600 million. It persists as one of the most important parasitic diseases of humans.

Pathology. The disease changes in schistosomiasis may be described as follows.

Prepatent period. The penetration of the cercariae may or may not produce skin irritation. Heavy initial exposures result in inflammatory disease of the skin (urticaria), toxic symptoms, an increase in the size of the liver and spleen, and eosinophilia before egg laying starts.

Patent period. Schistosoma mansoni may lay 350 eggs daily, *S. japonicum*, 3000. About 10% reach the feces; the rest remain imprisoned in tissues, particularly the liver and intestine or urinary bladder, where they provoke fibrotic reaction (scars). Egg extrusion may result in bloody diarrhea or urine over a protracted period. The liver and spleen enlarge, and the lungs and spinal cord may become involved.

Postpatent period. The fibrotic changes may result eventually even when eggs are absent in the feces. The changes interfere with intestinal, hepatic, and bladder functions with serious consequences. Liver fibrosis leads to splenomegaly and esophageal varices. Death may occur from profuse hematemesis. Lung fibrosis may cause cardiac failure. Spinal cord lesions may result in paralysis.

Diagnosis. Rectal biopsy is the best definitive method to show the presence of eggs in the intestinal schistosomiasis. Bladder biopsy, or cystoscopy, is indicated in the urinary type. Fecal examination is of value in the patent period. Serological methods are of great complementary value; the complement-fixation test is the most sensitive, while the circumoval-precipitin one is highly specific. The method consists in incubating live eggs with blood serum of the suspect and searching for a hyaline, ear-like precipitate protruding from the shell. Enzyme immunoassays have been developed that may prove superior to other methods as purified antigens become available. *See* COMPLEMENT-FIXATION TEST; IMMUNOASSAY.

Prevention. Measures to avoid contact with the infective stage are the only means of preventing

infection. In agricultural societies, this is virtually impossible. Latrines, rubber boots, skin repellants, education, and snail control have been tried, but with limited success. In more advanced industrialized countries, controlled urbanization has reduced exposure sites, with a consequent decrease in new infections.

Treatment. Chemotherapy has been difficult and dangerous, particularly in severe infections, but drugs have become available that promise effective control. Praziquantel, for example, with a 1-day oral administration is well tolerated and reportedly attains 95% cure of all schistosomes of humans at any stage of development. *See* EPIDEMIOLOGY; MEDICAL PARASITOLOGY. José F. Maldonado-Moll

Bibliography. P. C. Beaver and R. C. Jung, *Animal Agents and Vectors of Human Disease*, 5th ed., 1985; S. Cohen and K. S. Warren (eds.), *Immunology of Parasitic Infections*, 2d ed., 1982; M. Katz, D. D. Despommier, and R. W. Gadz, *Parasitic Diseases*, 1988; K. S. Warren, *Tropical and Geographical Medicine*, 1989.

Schistostegales

An order of the true mosses (subclass Bryidae), consisting of a single species, *Schistostega pennata*, the cave moss, which is especially characterized by its leaf arrangement and the form of its protonema. The plants grow in dimly lit, cavelike places, such as the undersides of upturned tree roots. The cave moss is widely distributed in north temperate regions. An unrelated moss in Australia and New Zealand, *Mittenia plumula*, has an identical protonema and habitat.

The protonema is much branched, with some branches consisting of spherical cells acting like lenses to concentrate weak rays of light on a backing of chloroplasts which reflect a greenish glow. The leafy plants are erect, unbranched, and dimorphous: sterile shoots have leaves wide-spreading in two rows and confluent because of broad basal decurrencies. Fertile plants have leaves in five rows and erect-spreading in a terminal tuft. All leaves are ecostate, entire, and unbordered, with thin-walled, laxly oblong-rhomboidal cells. The inflorescences are terminal, the setae elongate, and the capsules subglobose, with a flat operculum but no peristome. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA. Howard Crum

Bibliography. K. A. Lye, Studies in Norwegian bryophytes, 1: The family Schistostegaceae, *Lindbergia*, 1:205–213, 1972; H. Robinson, Observations on the origin of the specialized leaves of *Fissidens* and *Schistostega*, *Rev. Bryol. Lichen.*, 37:941–947, 1970.

Schizomida

An order of Arachnida with two families, Protoschizomidae (*Agastoschizomus*, *Protoschizomus*) and Schizomida (*Schizomus*, *Trithyreus*, *Megaschizomus*), including about 130 described species, mainly in the genus *Schizomus*. Schizomids occur

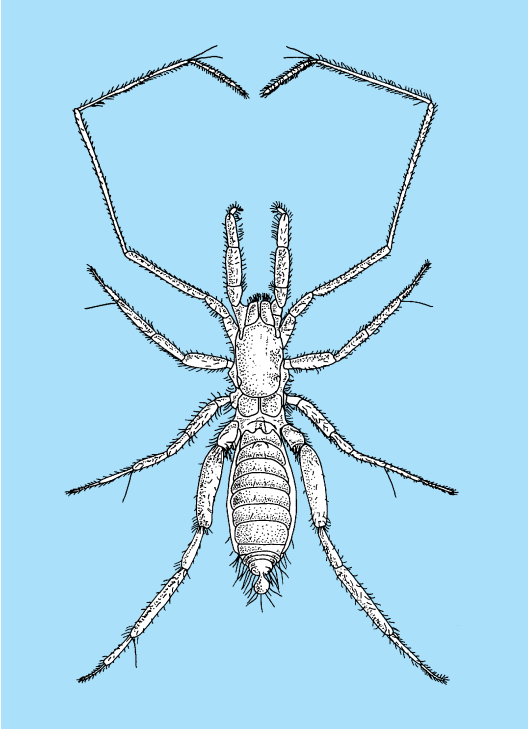

***Schizomus sawadai*. (After K. Sekiguchi and T. Yamasaki, A redescription of "Trithyreus sawadi" (Uropygi: Schizomidae) from the Bonin Islands, Acta Arachnol., 24(2):73–81, 1972)**

in most tropical and warm temperate regions, but each species has a restricted distribution. Many undescribed species probably occur in tropical leaf litter and caves. Schizomids are most closely related to Uropygi, with which they share several evolutionary novelties. *See* UROPYGI.

Schizomids are 0.12–0.44 in. (3–11 mm) in size, are white to tan, and lack eyes, although some retain vestigial "eyespots" (see **illus.**). The carapace is distinctively tripartite, and the abdomen has a short, terminal flagellum. Pedipalpal coxae are fused to form a trough, in which the animal digests its food externally. Pedipalps are raptorial, and first legs are elongated to serve as tactile organs. Pedipalps and flagellum are often sexually dimorphic.

Because schizomids are vulnerable to desiccation, they inhabit only moist places such as leaf litter, soil, caves, or beneath stones and logs. They are strict carnivores and fast runners. Sperm is transferred between sexes by a sperm capsule (spermatophore). The male deposits the spermatophore on the ground, the female grasps his modified flagellum in her chelicerae, and they advance in tandem until she contacts the spermatophore. Females apparently construct earthen brood chambers and stay with the young for a period of time. *See* ARACHNIDA. Jonathan A. Coddington

Bibliography. A. Kaestner, H. W. Levi, and L. R. Levi, *Invertebrate Zoology*, 2d ed., vol. 2, 1980; J. R. Reddell and J. C. Cokendolpher, Redescription of *Trithyreus grassii* (Arachnida: Schizomida: Schizomidae), *Oriental Insects*, 18:43–52, 1985;

J. M. Rowland, Breeding habits and early development of *Trithyreus pentapeltis* (Arachnida: Schizomida), *Entomol. News*, 83(3):69–74. 1972; J. M. Rowland and J. R. Reddell, The order Schizomida (Arachnida) in the New World, IV. *Goodnightorum* and *briggsi* groups and unplaced species (Schizomidae: *Schizomus*), *J. Arachnol.*, 9(1):19–46, 1981.

Schizophrenia

A brain disorder that is characterized by unusual mental experiences, such as hallucinations, changes in emotional experience, and severe decrements in social, cognitive, and occupational functioning. The popular belief that schizophrenia is characterized by multiple personalities is a misconception.

Schizophrenia is the most financially costly of all psychiatric conditions. Patients with schizophrenia use a disproportionately high percentage of mental health services. The prevalence of schizophrenia in the United States is approximately 1% of the population, but annual U.S. mental health care expenditures for its treatment have been estimated to be higher than 2.5% of the total cost of all health care in the country. Schizophrenia is one of the top five causes of disability for males and females, which is particularly striking since other causes of disability are far more common (for example, alcohol abuse and motor vehicle accidents). Schizophrenia is one of the top three most costly illnesses in the United States (following vascular illnesses and cancer). This high cost is underscored by the fact that many more Americans are affected annually by heart disease and cancer than schizophrenia. The indirect cost (lost wages and decreased productivity) of schizophrenia has been estimated to be approximately three times as great as the direct costs (medicine, hospitalization, and mental health care) of treating the illness. This article describes the characteristics of schizophrenia, its possible causes, and current treatments.

Diagnosis and Clinical Characteristics

Patients with schizophrenia demonstrate a series of biological differences when compared as a group with healthy individuals without schizophrenia. At present, however, there is no single biological marker available to indicate the presence of schizophrenia. A diagnosis is made on the basis of a cluster of symptoms reported by the patient and signs identified by the clinician. The most recent version of the *Diagnostic and Statistical Manual of Mental Disorders*, referred to as DSM-IV, includes criteria to help distinguish schizophrenia from other psychiatric disorders (see **table**).

Diagnostic criteria for schizophrenia (DSM-IV)

Criterion category	Description
Symptoms	Two or more of the following (one if hallucinations are characteristic or delusions are bizarre): delusions, hallucinations, disorganized speech, grossly disorganized or catatonic behavior, and negative symptoms
Social/occupational dysfunction	Significant decline in functioning in areas such as work, interpersonal relations, or self-care
Duration	Continuous signs for at least 6 months, including 1 month of symptoms
Exclusions	Symptoms are not better explained by other disorders such as mood disorders, other brain disorders, substance abuse, autism or pervasive developmental disorder

Hallucinations. People with schizophrenia may report perceptual experiences in the absence of a real environmental stimulus. The most common of these is auditory hallucinations, most often reported in the form of words spoken to the person with schizophrenia. The voices are often derogatory in nature, and can be tremendously frightening, especially in people experiencing hallucinations for the first time. Visual hallucinations, such as of human faces, are also possible. Tactile hallucinations, such as the experience of something moving on the skin, are less likely. Hallucinations through the senses of smell (olfactory) and taste (gustatory) are rare but can occur in some patients. *See* HALLUCINATION.

Delusions. People with schizophrenia often maintain beliefs that are not held by the overwhelming majority of the general population. To be considered delusions, the beliefs must be unshakable. In many cases, these beliefs may be "bizarre." A bizarre delusion is defined as an idea that is completely impossible (for example, "a computer has been inserted into my skull"), in contrast to nonbizarre delusions that are plausible ("The FBI is watching me"), but still not true. Many of these bizarre beliefs stem from odd experiences. For instance, a person may report that new thoughts are placed inside his or her head via an outside force (thought insertion), or that his or her movements are being controlled by an outside agency, such as a satellite (passivity experiences). In some instances, the delusions have an element of suspiciousness to them, such as the incorrect belief that others are planning to cause the person with schizophrenia harm. The delusions may or may not be related to hallucinatory experiences.

Negative symptoms. Delusions and hallucinations are often referred to as "positive" symptoms, since they represent features of experience that are present in people with schizophrenia and absent in the general population. The phrase "negative" symptoms refers to human behavior that is found in most people, yet absent in people with schizophrenia. Included among these negative symptoms are social isolation, lack of motivation, lack of energy, slow or delayed speech, and diminished emotional expression, often referred to as "blunted affect."

Disorganization. People with schizophrenia may manifest an odd outward appearance due to disorganization. This presentation may include speech that does not follow logically or sensibly, at times to the point of being incoherent. Facial expression may be odd or inappropriate, such as laughing for no reason.

In some cases, people with schizophrenia may move in a strange and awkward manner. The extreme of this behavior, referred to as catatonia, has become very rare since pharmacological treatments became available.

Cognitive deficits. Although less striking than delusions and hallucinations, perhaps the most devastating feature of schizophrenia is the cognitive impairment found in most people with the disorder. On average, people with schizophrenia perform in the lowest 2–10% of the general population on tests of attention, memory, reasoning and problem solving, motor skills, and language abilities. Performance on tests of intelligence is lower than the population average, but not as impaired as these other cognitive domains. These cognitive deficits are perhaps the most important explanation for the difficulties that people with schizophrenia have functioning in everyday society. In addition, people with schizophrenia often have reduced awareness (referred to as "insight") regarding their illness. In fact, some may deny all symptoms of their disorder to the point that they will refuse treatment. *See* COGNITION.

Mood symptoms. Schizophrenia is associated with depression or bipolar disorder (manic-depressive illness) in about 10–15% of people with the disorder. The presence of both of these illnesses is referred to as schizoaffective disorder, and is generally associated with a more favorable outcome. In addition, as many as 50% or more of patients with schizophrenia experience episodes of depression that do not meet the full criteria for major depression. Up to 12% of patients with schizophrenia commit suicide, often during periods when they are depressed. *See* AFFECTIVE DISORDERS.

Functional impairment. People with schizophrenia are far less likely than the general population to work, marry, have offspring, and live independently. About 10–15% are able to sustain full-time employment, while 25–40% marry and have children. While the deinstitutionalization movements of the 1950s and the 1980s drastically reduced the number of people in long-term inpatient facilities, only a small minority of people diagnosed with schizophrenia in the United States are able to live without some form of public assistance. This functional impairment develops very early in the illness, and as many as 50% of patients with schizophrenia receive disability assistance within 6 months of their first treatment for the illness.

Course. The onset of schizophrenia is generally in the late teens to early twenties; however, onset is possible throughout the life span, including childhood, which is referred to as childhood schizophrenia, and in the later stages of life (after age 40), which is referred to as late-onset schizophrenia. Recent research on the early predictors of schizophrenia suggest that social isolation and cognitive deficits appear to be present years prior to the onset of delusions and hallucinations. While the onset of symptoms is abrupt in some people, others experience a more insidious process, including extreme social withdrawal, reduced motivation, mood changes, and cognitive and functional decline prior to the onset of full-blown schizophrenia symptoms. Recent interest has focused on the early indicators of impending psychosis, referred to as the "prodrome." If the presence of a prodrome could be reliably identified, then treatments for patients with schizophrenia could potentially be offered earlier and maybe even prevent the onset of illness in some cases.

Following the onset of illness, the course of schizophrenia is normally characterized by episodes of relative remission, in which only subtle residual and negative symptoms remain, and episodes of exacerbation of symptoms, which are often caused by failure to continue with treatment. Some long-term studies have suggested that some patients may show a reduced tendency to have exacerbations beyond their fifties.

Medical comorbidities. Patients with schizophrenia have a reduced life expectancy, above and beyond the statistical effects of suicide. Much of the increased mortality is due to cardiovascular illnesses related to smoking and obesity, both of which are more common in patients with schizophrenia than the population as a whole. Approximately 50–80% of people with schizophrenia smoke cigarettes, in contrast to the general population estimate of 25%. Due to their functional impairments and poverty, many patients with schizophrenia also receive substandard medical care.

Causes

The cause of schizophrenia has been a matter of concern and controversy for the last century. In a sense, the argument of "nature versus nurture" has been acted out repeatedly in this domain. It is likely that there are various forms of schizophrenia, perhaps all with different causes. Although schizophrenia appears to be inherited, at least in some cases, the influence of genes is far from complete. Although many arguments have been put forth regarding environmental factors that could cause schizophrenia, very few of these theories are consistently supported.

Genetics. Schizophrenia runs in families. The children of a parent with schizophrenia have a risk of developing the illness of about 10%, which is greater than 10 times the risk of developing schizophrenia with no relatives with the illness. This risk appears to be similar regardless of whether the children are raised by their parents with schizophrenia or are adopted away. At the same time, if one identical twin has schizophrenia there is only about a 50% chance that the other twin (who has exactly the same genes) will also be affected. These data suggest that a strictly genetic explanation is not adequate to describe the causes of the illness.

There have been extensive searches for genetic markers for the predisposition to develop schizophrenia. Other illnesses, such as Huntington's disease, which follow a pattern of inheritance associated with a strictly genetic transmission (50% of the children of any affected parent will develop the condition), have been found to have specific genetic markers. Recently, several genetic markers have been

found to be associated with schizophrenia. However, some of these influences appear similar in the healthy population and individuals with schizophrenia. It is likely, therefore, that each of these markers predispose a certain subgroup of the population to schizophrenia, but that a combination of genetic and environmental influences are needed to lead to the actual disorder. At this time, the most reasonable conclusion is that schizophrenia has a genetic component, albeit one that is not a classic pattern of single-gene inheritance, either recessive or dominant in nature. Research is continuing the search for better indicators of the genetic predisposition for schizophrenia. At present, too little is known to warrant the use of genetic counseling or other interventions. *See* BEHAVIOR GENETICS; HUMAN GENETICS.

Obstetrical complications. Several studies have found that the birth and intrauterine development of someone who eventually develops schizophrenia is more likely than normal to be associated with complications, such as forceps delivery or second-trimester maternal illness such as influenza. In addition, children of parents with schizophrenia who experience these complications are more likely than their siblings to manifest schizophrenia when they grow up. However, the number of documented cases of schizophrenia associated with pregnancy and birth complications in individuals without a schizophrenic relative is quite low. Thus, the simple experience of complicated pregnancy and delivery does not mean that any individual child has any meaningful risk for developing schizophrenia.

Family and social interaction. Many theories have tried to explain schizophrenia in terms of patterns of family interaction. There is no credible evidence that interaction patterns cause schizophrenia in someone who would not have developed it anyway. A small minority of mental health professionals still adhere to these theories, but they have been discarded by the majority since the late 1970s. It is clear, however, that social interaction patterns in the environment of individuals who already have schizophrenia can have an effect on their course of illness. High levels of criticism directed at newly discharged people with schizophrenia, like other stressors, is associated with increased frequency of reemergence of symptoms and rehospitalization.

Regional brain dysfunction. One of the new developments in the study of schizophrenia is associated with imaging technologies. Magnetic resonance imaging (MRI) allows for the visualization of the brain at high levels of resolution, allowing the identification of structures as small as 1 mm. People with schizophrenia often have changes in the structure of their brain, such as enlargement of the cerebral ventricles (fluid-filled spaces in the brain close to the midline). Various brain regions have been found to be smaller in patients with schizophrenia, including the frontal cortex, temporal lobes, thalamus, and the hippocampi. These brain structure abnormalities appear not to be isolated to the gray matter in the cortex, but may include the white matter that connects brain regions.

An additional development in the study of schizophrenia is the ability to study brain functions. Using MRI machines that take a rapid series of images of the brain, it is possible to capture patterns of blood flow in the brain. Since blood flow correlates with brain activity, it is possible to examine brain activity while people perform mental tasks. Studies of patients with schizophrenia have found patterns of abnormal activation of the brain while performing tests of memory and problem solving. While healthy individuals' brains use the frontal cortex to perform these tasks, patients with schizophrenia tend to have a less organized and coherent pattern of activation. Consistent with their generally less organized approach to learning new information and solving problems, brain activity itself in schizophrenia appears to be organized in a less efficient manner. *See* BRAIN; MEDICAL IMAGING.

Neurotransmitter changes. Individual cells in the brain (that is, neurons) communicate with each other using electrical-chemical means. These communications occur when different chemicals (called neurotransmitters) produced by the brain are released in proximity to other neurons. For years it has been suspected that neurotransmitter activity in schizophrenia is abnormal, because the medications used to treat schizophrenia act on dopamine receptors in the brain. Also, psychotic conditions can be caused by overdoses of medications that stimulate dopamine, such as amphetamine. The medications that are used to treat psychotic conditions such as schizophrenia act, in part, to reduce the activity of dopamine in the corpus striatum. This brain region is located at the juncture of the newer (cortex) and older (subcortical) parts of the brain.

Although dopamine is clearly implicated in schizophrenia, it is not simple overactivity of this neurotransmitter that causes schizophrenia. Measures of dopamine functions often do not suggest increased activity in patients with schizophrenia, and in some patients there is evidence of decreased activity. In addition, genes that influence levels of dopaminergic transmission have similar effects in healthy individuals and in patients with schizophrenia. Finally, dopamine interacts with several other neurotransmitters, meaning that the effects of dopamine might be indirect. With advances in technology that allow for the direct measurement of neurotransmitter activity in the central nervous system, more information in this area will be available in the near future. The current view of the dopamine hypothesis is that levels of transmitter activity are improperly balanced, with too little activation in the cortex and too much in the striatum. *See* DOPAMINE; NEUROBIOLOGY; NEURON.

Treatment

Both pharmacological and behavioral approaches are used to treat schizophrenia.

Pharmacological treatment. Chlorpromazine was the first pharmacological intervention to substantially reduce delusions and hallucinations in patients with schizophrenia. Several other similar medications were developed in the following decades. This class of medications, referred to as typical neuroleptics or conventional antipsychotics, act primarily to block dopamine receptors in the brain. These medications were able to have an impact on delusions and hallucinations in some patients, but treatment was accompanied by a variety of side effects, including sedation, muscular rigidity, and restlessness. Adjunctive medications such as benztropine are used to reduce these side effects, but such anticholinergic medications have side effects of their own, such as dry mouth, dizziness, and further cognitive impairment. Perhaps the worst side effect of typical antipsychotics is tardive dyskinesia, an irreversible movement disorder. The most effective dosage of these medications was initially believed to be very high. More recent studies have suggested that lower doses (for example, 2–4 milligrams per day of haloperidol) may be as effective with fewer side effects. The disadvantage of these medications is that they have minimal impact on cognitive deficits and negative symptoms such as social isolation. In the United States, only a small group of patients with schizophrenia receive treatment with these medications. In other areas of the world, particularly in economically disadvantaged countries, these medications continue to be the norm.

Since the approval of clozapine in the United States in 1988, clinicians have referred to a new group of medications as atypical neuroleptics or novel antipsychotics. The side effects associated with these medications are far different than the conventional antipsychotic medications, and in many cases are believed to be minimized. While clozapine had been available in Europe for years, it was banned in the United States due to the presence of a severe side effect, agranulocytosis (loss of white blood cells), which can be fatal. However, clozapine was the first medication that improved negative symptoms and cognitive deficits as well as delusions and hallucinations. Its approval in the United States was dependent upon a strict monitoring procedure that required patients to receive frequent blood tests if they were to continue on the medication.

Over the past decade, the Food and Drug Administration (FDA) approved several newer, atypical antipsychotic medications for use in the United States, including risperidone (1994), olanzapine (1996), quetiapine (1999), ziprasidone (2000), and aripiprazole (2002). The advantage of these medications is that they appear to have few of the side effects of the conventional antipsychotic medications, and they improve negative symptoms and cognitive function. The disadvantage of these medications is that they are more expensive than the conventional medications, whose patents have expired. The novel antipsychotic medications also have side effects. At higher doses, some of them have some of the side effects of the older medications, and many of them cause significant weight gain and diabetes. The efficacy and side-effect profiles have not been fully determined for the newest of these medications. *See* PSYCHOPHARMACOLOGY.

Behavioral treatment. There are several targets for behavioral treatments in schizophrenia. Patients with schizophrenia have difficulty acquiring skills in social, occupational, and independent living domains. Structured training programs have attempted to teach patients with schizophrenia how to function more effectively in these areas. Family interventions have also been designed to provide a supportive environment for patients with schizophrenia. These interventions have been demonstrated to reduce risk of relapse. Direct treatment of cognitive impairments, referred to as cognitive remediation, has also been implemented for patients with schizophrenia. While these interventions have not had tremendous success, their efficacy in combination with newer antipsychotics or other medications that potentially enhance cognition is promising. Teaching patients how to cope with their hallucinations and delusions is another target of current behavioral treatment. Most patients with schizophrenia do not spontaneously recognize their symptoms as unusual and unreal. Cognitive-behavioral treatments have been employed to help patients realize the nature of their symptoms and to develop plans for coping with them. Given the numerous difficulties confronted by people with schizophrenia, they often benefit greatly from the opportunity to discuss the exigencies of their daily life with someone who does not turn away from the unusual nature of their disorder. *See* PSYCHOTHERAPY.

Richard S. E. Keefe; Philip D. Harvey

Bibliography. American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed. (DSM-IV), American Psychiatric Association, 1994; M. F. Green, *Schizophrenia Revealed*, Norton, New York, 2003; R. S. E. Keefe and P. D. Harvey, *Understanding Schizophrenia: A Guide to the New Research on Causes and Treatment*, Free Press, New York, 1994; E. F. Torrey, *Surviving Schizophrenia: A Manual for Families, Consumers, and Providers*, 4th ed., HarperCollins, New York, 2001; P. J. Weiden, *Breakthroughs in Antipsychotic Medications: A Guide for Consumers, Families, and Clinicians*, Norton, New York, 1999.

Schlieren photography

Any technique for the photographic recording of schlieren, which are regions or stria in a medium that is surrounded by a medium of different refractive index. Refractive index gradients in transparent media cause light rays to bend (refract) in the direction of increasing refractive index. This is a result of the reduced light velocity in a higher-refractive-index material. This phenomenon is exploited in viewing the schlieren, with schlieren photographs as the

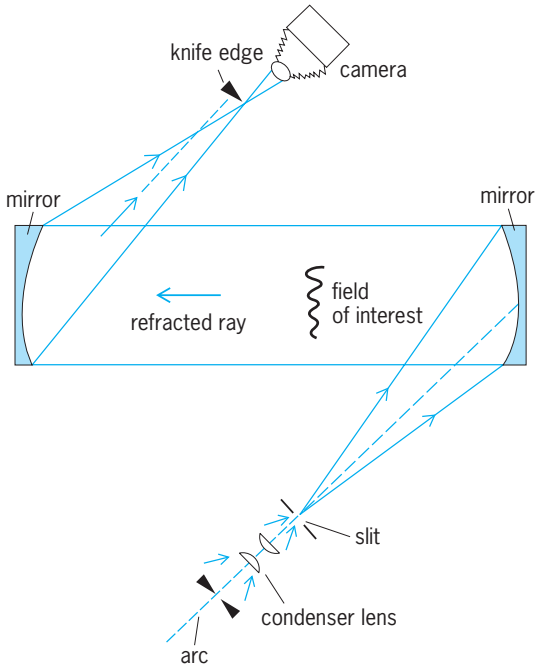

Knife-edge method of viewing schlieren, employing the "z" configuration.

result. Modern schlieren systems are rarely limited to photography; electronic video recorders, scanning diode array cameras, and holography are widely used as supplements. *See* HOLOGRAPHY; OPTICAL DETECTORS; PHOTOGRAPHY; REFRACTION OF WAVES; SHADOWGRAPH.

Knife-edge method. There are many techniques for optically enhancing the appearance of the schlieren in an image of the field of interest. In the oldest of these, called the knife-edge method (see **illus.**), a point or slit source of light is collimated by a mirror and passed through a field of interest, after which a second mirror focuses the light, reimaging the point or slit where it is intercepted by an adjustable knife edge (commonly a razor blade is used). The illustration shows the "z" configuration which minimizes the coma aberration in the focus. Mirrors are most often used because of the absence of chromatic aberration. *See* ABERRATION (OPTICS).

Rays of light that are bent by the schlieren in the direction of the knife edge are intercepted and removed from the final image of the region of interest, causing those regions to appear dark. Consequently, the system is most sensitive to the density gradients that are perpendicular to the knife edge. The knife edge is commonly mounted on a rotatable mount so that it can be adjusted during a measurement to optimally observe different gradients in the same field of interest. The intensity in the processed image is proportional to the refractive index gradient. A gradient in the same direction as the knife edge appears dark. Gradients in the opposite direction appear bright. This method, employed with arc light sources, is still one of the simplest ways to view refractive index changes in transparent solids, liquids, and gases. A well-designed schlieren system can easily detect the presence of a refractive index gradient that causes 1 arc-second deviation of a light ray.

Color schlieren. Except for locating and identifying schlieren-causing events such as turbulent eddies, shock waves, and density gradients, schlieren systems are usually considered to be qualitative instruments. Quantitative techniques for determining density are possible but are much more difficult to employ. The most common of these is color schlieren. The knife edge is replaced with a multicolored filter. Rays of light refracted through different angles appear in different colors in the final image.

Use of coherent light. The availability of lasers and new optical components has expanded the method considerably. When a coherent light source such as a laser is used, the knife edge can be replaced by a variety of phase-, amplitude-, or polarization-modulating filters to produce useful transformations in the image intensity. For example, replacing the knife edge with a 180° phase step produces the Hilbert transform. Adding a graded-amplitude wedge to this performs differentiation of the image. Phase-shifting only the central focused spot produces phase-contrast interferometry. A variety of filters produce images referred to as schlieren interferometry. These include the insertion of stops, gratings, and Wollaston prisms in place of the knife edge. *See* INTERFEROMETRY; LASER; POLARIZED LIGHT.

Applications. Schlieren systems are common in aerodynamic test facilities and optical test laboratories. A schlieren system was used in the space shuttle *Space Lab 3* flights to observe fluid densities around crystals growing in zero gravity. *See* SHOCK WAVE; TRANSONIC FLIGHT; WIND TUNNEL.

James D. Trolinger

Bibliography. G. S. Settles, *Schlieren and Shadowgraph Techniques*, Springer-Verlag, 2001.

Schmidt camera

A fast, compact, wide-field optical system that uses a thin aspheric front lens at the center of curvature of a larger concave spherical mirror. The image is focused onto a curved focal plane between the two elements (**Fig. 1**). Schmidt cameras are sometimes

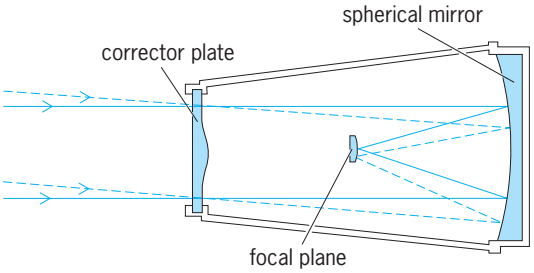

Fig. 1. Cross section of Schmidt camera with aspherical corrector plate. (*After J. M. Pasachoff, Astronomy: From the Earth to the Universe, 6th ed., Brooks/Cole Publishing, 2002*)

Largest Schmidt telescopes

Telescope or institution	Location	Diameter: Corrector plate, in. (m)	Diameter: Spherical mirror, in. (m)	Focal ratio	Date completed
Karl Schwarzschild Observatory	Tautenberg (near Jena), Germany	53 (1.3) (removable)	79 (2.0)	*f*/2	1960
Samuel Oschin Telescope, Palomar Observatory	Palomar Observatory, California	48 (1.2)	72 (1.8)	*f*/2.5	1948, 1987, 2000
United Kingdom Schmidt, Anglo-Australian Observatory	Siding Spring Observatory, New South Wales, Australia	48 (1.2)	72 (1.8)	*f*/2.5	1973
Tokyo Astronomical Observatory	Kiso Mountains, Japan	41 (1.1)	60 (1.5)	*f*/3.1	1975
European Southern Observatory Schmidt*	La Silla, Chile	39 (1.0)	64 (1.6)	*f*/3	1972

*Decommissioned in 1998.

used in microscopes, astronomical spectrographs, and projection televisions. The largest such devices can image several degrees of the sky and are known as Schmidt telescopes. The Estonian optician Bernhard Schmidt devised the scheme in 1930. The field of best focus is located midway between the lens and the mirror and is curved toward the mirror, with a radius of curvature equal to the focal length. Historically film or photographic plates were bent to match this curved focus, and used to make the first all-sky survey pictures. Field-flattening lenses permit the use of charge-coupled devices (CCDs) in newer applications. Large mosaics of CCDs have replaced photographic emulsions in some applications, and large numbers of optical fibers have also been used in multiobject spectrograph configurations. *See* GEOMETRICAL OPTICS; LENS (OPTICS); MIRROR OPTICS.

Schmidt telescopes are very fast, some having focal ratios in the vicinity of *f*/1. Their distinguishing feature is the very wide angle over which good images are obtained. Thus, under a dark sky they are sensitive to extended objects of low surface brightness. Schmidt telescopes have no coma; because the only lens element is so thin, they suffer only slightly from chromatic aberration and astigmatism. The front element, smaller than the main mirror, is sometimes known as a corrector plate. Often the outer surface of the corrector plate is plane, with the inner surface bearing the figure; however, a slight curvature is sometimes introduced to prevent misleading ghost images. This corrector plate almost eliminates the spherical aberration, giving extremely sharp images. For the most critical work, over a wide range of wavelengths, it must be made achromatic. *See* ABERRATION (OPTICS); OPTICAL SURFACES.

The Schmidt design became widely used after 1936, when the secret of the fabrication of the corrector plate was released. Schmidt is credited as much with his skill in figuring the fourth-degree curves on the corrector plate as with the design itself. The largest Schmidt telescopes (see **table**; **Fig. 2**) are 39–53 in. (1.0–1.3 m) in diameter. Larger corrector plates would sag too much, as they can only be supported around their circumference.

Increasingly, Schmidt telescopes, especially several 24 in. (0.6 m) in diameter, are equipped with

Fig. 2. Samuel Oschin Telescope, a 48-in. (1.2-m) Schmidt camera at Palomar Observatory. (*a*) Interior view (*Alain Maury*). (*b*) Exterior view. A large liquid nitrogen tank to the right of the open dome is used for cooling the instrument. (*Scott Kardel*)

charge-coupled devices. An upgrade to the Samuel Oschin Telescope (48-in. or 1.2-m Schmidt) at Palomar Observatory in California installed the Yale-QUEST mosaic of 112 CCDs, which images an area of the sky 4° square. While smaller than the original photographic field (6°), it is still enormous when compared with the fields for charge-coupled devices at most current large telescopes. The telescope is now fully robotic, and operates automatically and unattended every clear night. The Palomar-QUEST camera is used about half the time to survey the sky for time-variable objects such as Supernovae and Gamma-Ray Bursters, and the other half to search for moving objects such as near-Earth asteroids and distant solar system objects such as Sedna. The use of charge-coupled devices provides higher sensitivity and allows advanced image processing and easy distribution of data and results. These large-format cameras require cooling to reduce electronic and readout noise (Fig. 2*b*). *See* CHARGE DEVICES.

Sky surveys. The Schmidt telescope at Palomar Observatory, now the Samuel Oschin Telescope, was used in the 1950s to survey the northern two-thirds of the sky in the Palomar Observatory–National Geographic Society Sky Survey (POSS). The plate pairs—one photographed in red light and one in blue light—provided a first-epoch coverage on which many nebulae, galaxies, clusters of galaxies, and other objects were discovered. Each plate, about 14 in. (35 cm) square, covers an area about 6° square. The POSS and later Schmidt sky surveys are basic references. An infrared survey north of the equator was carried out in the 1970s. A smaller Schmidt telescope (18 in. or 46 cm) at the Palomar Observatory has discovered many asteroids with orbits that cross that of the Earth, as well as comet Shoemaker-Levy 9 and other comets. *See* ASTEROID; COMET.

The European Southern Observatory (ESO) Schmidt in Chile (decommissioned in 1998) and the United Kingdom Schmidt at the Anglo-Australian Observatory in Australia were used to survey the southern hemisphere, with improved emulsions, from declination +20° to the south celestial pole. The United Kingdom Schmidt was provided with an achromatic corrector plate in 1977. It completed an infrared survey of the Milky Way and Magellanic Clouds, and extended this survey to the whole southern sky, and to a blue-red study of the region just south of the celestial equator to provide images showing stars about 1.5 magnitudes fainter than those on the Palomar survey. In addition to mapping, the plates are used to help make optical identifications of southern radio and x-ray sources. A second-epoch red survey was carried out in the 1990s, largely for Hubble guide-star purposes. The United Kingdom Schmidt Telescope is carrying out a survey of the Milky Way Galaxy and Magellanic Clouds in the red light of hydrogen (H-alpha) to reveal supernova remnants, planetary nebulae, and the fine structure of gas clouds in the Galaxy. Study of the plates has been speeded by the development of automatic plate-measuring machines. Three-color U.K. Schmidt photographs of celestial objects by D. F. Malin are widely viewed.

In 1987, an achromatic doublet corrector plate was installed in the Samuel Oschin Telescope. With it, a second-epoch sky survey was undertaken. An infrared color (800–900 nanometers, limiting magnitude 19) was added to the blue (385–550 nm) and red (610–690 nm) colors, which are slightly different from those of the first survey but are similar to those used by the United Kingdom and European Southern Observatory Schmidts for their southern survey. Advances in film technology and thermal control of the telescope, to improve the focus, allowed a limiting magnitude fainter by a factor of 6 in the blue (to magnitude 22.5) and 2 in the red (to magnitude 20.8) to be reached and images to be recorded with finer resolution. The infrared survey was made possible by developments in hypersensitization of infrared emulsions, which increased their speed by a factor of 200. The survey includes 894 fields between the celestial equator and the north celestial pole. The second-epoch plates were scanned to produce the Digital Sky Survey (DPOSS). As parts of the sky were completely surveyed, the telescope was used to search for Kuiper Belt objects, and this work has continued with the large digital camera. *See* KUIPER BELT.

Second-epoch Schmidt survey plates were scanned to make the Hubble Space Telescope Guide Star Catalog of 16,000,000 objects, and also form the basis of the U.S. Naval Observatory's USNO-B astrometric catalog of over 10^9 stars. The catalogs, scanned images, and database are available on the World Wide Web. *See* ASTRONOMICAL ATLASES; ASTRONOMICAL CATALOGS; ASTRONOMICAL PHOTOGRAPHY; HUBBLE SPACE TELESCOPE.

Spectroscopy. In addition to direct photography, Schmidt telescopes can be used to obtain the spectra of celestial objects. The United Kingdom Schmidt can be used with two 48-in.-diameter (1.2-m) objective prisms, either singly or in combination, to spread each image out into a short spectrum. The telescope also has a facility called FLAIR, which uses optical fibers to couple the focal surface to a charge-coupled-device spectrograph, allowing detailed spectra to be obtained of up to 100 objects simultaneously in the wide field of view. FLAIR is widely used for galaxy redshift surveys and surveys of particular star types. A fully automated version of FLAIR, known as 6dF (for 6-degree field) is being built to allow a galaxy-redshift survey of the entire southern hemisphere to be undertaken. The United Kingdom Schmidt and the Curtis Schmidt Telescope, a Schmidt telescope with a 24-in.-diameter (0.61-m) corrector plate and a 36-in.-diameter (0.91-m) spherical mirror at the Cerro Tololo Inter-American Observatory in Chile, also have objective-prism capabilities. *See* ASTRONOMICAL SPECTROSCOPY.

Related systems. The success of the Schmidt design has led to many other types of catadioptric systems, with a combination of lenses and mirrors. The desire for a wide field has led to most modern large telescopes being built to the Ritchey-Chrétien

design, which use two hyperboloid mirrors instead of the traditional paraboloids to achieve coma-free fields of up to one degree, with a small amount of focal plane curvature. These fields are much smaller than those of Schmidts, but the telescope aperture can be larger, because no large corrector lens is needed.

Schmidt-Cassegrain design and amateur use. Telescopes employing some features of the Schmidt design are also available for amateurs. Many amateur astronomers use telescopes of the Schmidt-Cassegrain design, in which a small mirror attached to the rear of the corrector plate reflects and refocuses the image through a hole in the center of the primary mirror. These telescopes, in sizes ranging from 5 in. (12.5 cm) to 16 in. (41 cm), are portable and relatively inexpensive for the quality of image. Their fields are much narrower than a standard Schmidt design, however, and the focal ratios much slower, usually about *f*/10. A relatively inexpensive range of telescopes are manufactured based on a modified Ritchey-Chrétien design. Because aspheric hyperboloids are expensive to manufacture, these telescopes combine a spherical primary with full-aperture corrector lens to emulate the necessary primary figure. Together with a hyperboloidal secondary attached to the rear of the corrector, they produce a wide angle (curved) coma-free field. *See* TELESCOPE.

Jay M. Pasachoff; Andrew Pickles

Bibliography. M. Capaccioli (ed.), *Astronomy with Schmidt-Type Telescopes*, IAU Colloq. 78, D. Reidel, 1983; J. Chapman et al. (eds.), *The Future Utilisation of Schmidt Telescopes*, IAU Colloq. 148, Astronomical Society of the Pacific, 1995; D. di Cicco, A preview of Meade's RCX400, *Sky Telesc.*, 109(5):86–89, May 2005; H. T. MacGillivray and E. B. Thomson (eds.), *Digitized Optical Sky Surveys*, Kluwer Academic Publishers, 1992; D. F. Malin, *A View of the Universe*, Sky Publishing Co./Cambridge University Press, 1993; S. Marx and W. Pfau, *Astrophotography with the Schmidt Telescope*, Cambridge University Press, 1992.

Schottky anomaly

A contribution to the heat capacity of a solid arising from the thermal population of discrete energy levels as the temperature is raised. The effect is particularly prominent at low temperatures, where other contributions to the heat capacity are generally small. *See* SPECIFIC HEAT OF SOLIDS.

Discrete energy levels may arise from a variety of causes, including the removal of orbital or spin degeneracy by magnetic fields, crystalline electric fields, and spin orbit coupling, or from the magnetic hyperfind interaction. Such effects commonly occur in paramagnetic ions. *See* LOW-TEMPERATURE THERMOMETRY; PARAMAGNETISM.

The thermal population of an energy level E_i is proportional to the Boltzmann factor, $\exp(-E_i/k_BT)$,

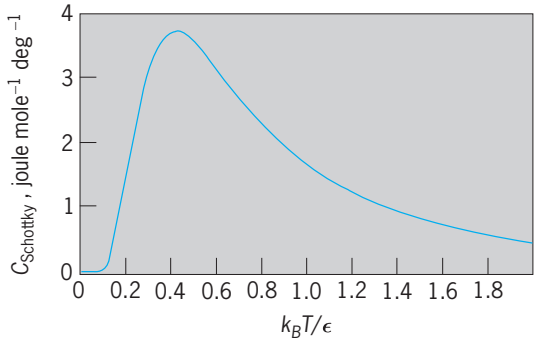

Theoretical curve for the Schottky heat capacity when there are two energy levels separated by a splitting ϵ. (*After H. M. Rosenberg, Low Temperature Solid State Physics, Clarendon Press, 1963*)

where T is the temperature and k_B is Boltzmann's constant ($k_B = 1.38 \times 10^{-23}\,\mathrm{J \cdot K^{-1}}$). The largest contributions to the Schottky heat capacity, C_{Schottky}, will therefore occur at temperatures where k_BT is comparable with the E_i. Quantitative estimates of C_{Schottky} may be made in terms of the Boltzmann factors. For an assembly of N, identical noninteracting systems, each with n low-lying energy levels E_i ($i = 1, \ldots, n$), the total internal energy is given by the equation below. The corresponding heat capacity is then

$$U = \frac{N \sum_{i=1}^{n} E_i \exp(-E_i/k_BT)}{\sum_{i=1}^{n} \exp(-E_i/k_BT)}$$

given by $C_{\text{Schottky}} = dU/dT$. The simplest example is the case $n = 2$, corresponding to a single splitting ϵ between the ground state ($i = 1$) and the first excited state ($i = 2$). The **illustration** shows the heat capacity for this case plotted as a function of k_BT/ϵ. For more complex energy-level schemes, the curves are generally similar, with a single broad peak near $k_BT/\bar{\epsilon} \sim 1$, where $\bar{\epsilon}$ is the mean energy above the ground state. However, more structured curves, with more than one peak, are also possible if the energy-level scheme includes excitation energies of very different relative magnitudes. *See* BOLTZMANN STATISTICS.

Corresponding to the Schottky heat capacity, there is a contribution to the entropy, $S = \int (C_{\text{Schottky}})/dT$. This can act as a barrier to the attainment of low temperatures if the substance is to be cooled either by adiabatic demagnetization or by contact with another cooled substance. Conversely, a substance with a Schottky anomaly can be used as a heat sink in experiments at low temperatures (generally below 1 K) to reduce temperature changes resulting from the influx or generation of heat. *See* ADIABATIC DEMAGNETIZATION; LOW-TEMPERATURE PHYSICS.

W. P. Wolf

Bibliography. E. S. Gopal, *Specific Heats at Low Temperatures*, 1966; C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1995; H. M. Rosenberg, *Low Temperature Solid State Physics*, 1963.

Schottky barrier diode

A metal-semiconductor diode (two terminal electrical device) that exhibits a very nonlinear relation between voltage across it and current through it; formally known as a metallic disk rectifier. Original metallic disk rectifiers used selenium of copper oxide as the semiconductor coated on a metal disk. Today, the semiconductor is usually single-crystal silicon with two separate thin metal layers deposited on it to form electrical contacts. One of the two layers is made of a metal which forms a Schottky barrier to the silicon. The other forms a very low resistance, so-called ohmic, contact. The Schottky barrier is an electron or hole barrier caused by an electric dipole charge distribution associated with the contact potential difference which forms between a metal and a semiconductor under equilibrium conditions. The barrier is very abrupt at the surface of the metal because the charge is primarily on the surface. However, in the semiconductor, the charge is distributed over a small distance, and the potential gradually varies across this distance (see **illus.**). *See* CONTACT POTENTIAL DIFFERENCE.

A basic useful feature of the Schottky diode is the fact that it can rectify an alternating current. Substantial current can pass through the diode in one direction but not in the other. If the semiconductor is *n*-type, electrons can easily pass from the semiconductor to the metal for one polarity of applied voltage, but are blocked from moving into the semiconductor from the metal by a potential barrier when the applied voltage is reversed. If the semiconductor is *p*-type, holes experience the same type of potential barrier but, since holes are positively charged, the polarities are reversed from the case of the *n*-type semiconductor. In both cases the applied voltage of one polarity (called forward bias) can reduce the potential barrier for charge carriers leaving the semiconductor, but for the other polarity (called reverse bias) it has no such effect. *See* DIODE; SEMICONDUCTOR; SEMICONDUCTOR RECTIFIER.

James E. Nordman

Bibliography. R. H. Bube, *Electrons in Solids: An Introductory Survey*, 3d ed., 1992; G. Neudeck, *The PN Junction Diode*, 2d ed., 1989.

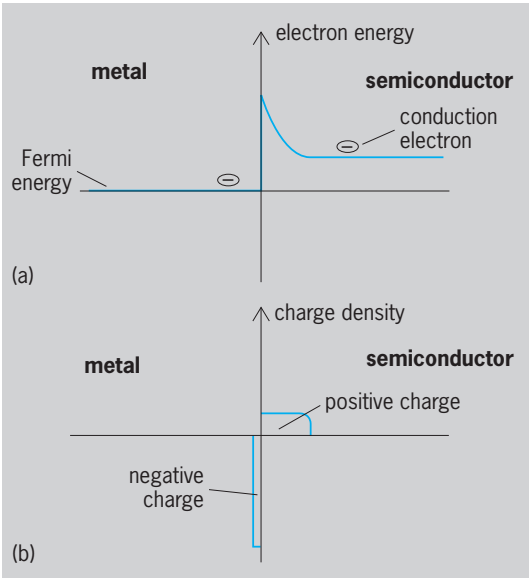

Distributions of (*a*) electric potential and (*b*) electric charge for a Schottky diode without voltage bias.

Schottky effect

The enhancement of the thermionic emission of a conductor resulting from an electric field at the conductor surface. Since the thermionic emission current is given by the Richardson formula, an increase in the current at a given temperature implies a reduction in the work function ϕ of the emitter. *See* THERMIONIC EMISSION.

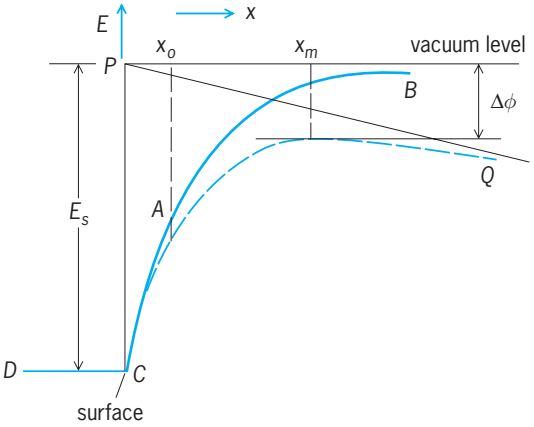

Fig. 1. Surface potential barrier and Schottky effect; the lowering of the work function $\Delta\phi$ is highly exaggerated.

With reference to **Fig. 1**, let the vacuum level represent the energy of an electron at rest in free space and let *CD* be the energy of a conduction electron at rest in a metal. If an electron approaches the metal surface from infinity, its potential energy *V* relative to the vacuum level is given by the well–known image potential $V(x) = -e^2/4x$, where x is the distance from the surface, and $e = 1.6 \times 10^{-19}$ coulomb is the magnitude of the electron's charge. The image potential is valid only for $x > x_0$, where x_0 is of the order of the distance between neighboring atoms in the metal; that is, x_0 is a few angstroms. In the absence of an applied field, *CAB* then represents the potential energy of an electron as a function of x. *AB* corresponds to the image potential; the exact shape of the curve between *C* and *A* is uncertain.

Suppose now a constant field *F* is applied externally between the surface of the emitting cathode and an anode; this produces a potential energy of an electron of $-eFx$ (line *PQ* in Fig. 1), and hence the total potential energy of an electron for $x > x_0$ is

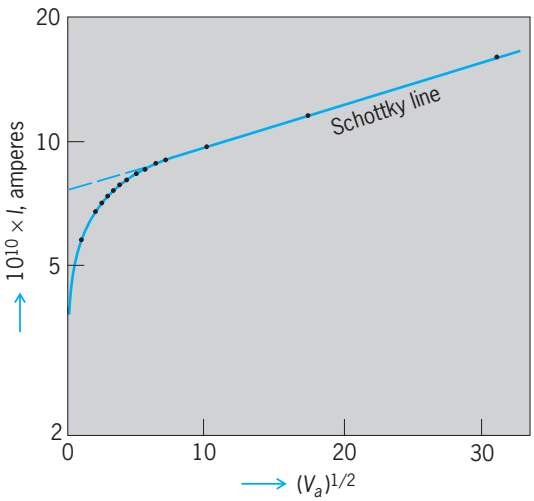

Fig. 2. Logarithm of thermionic emission current *I* of tungsten as function of square root of anode voltage V_a. (*After W. B. Nottingham, Phys. Rev., 58:927–928, 1940*)

given by Eq. (1), indicated by the broken line *CAQ*

$$V(x) = \frac{e^2}{4x} - eFx \tag{1}$$

in Fig. 1. This function has a maximum value given by Eq. (2).

$$x = x_m = {}^1/_2\left(\frac{e}{F}\right)^{1/2} \tag{2}$$

The maximum lies below the vacuum level by an amount $\Delta\phi$, given in Eq. (3), which represents the

$$\Delta\phi = V(x_m) = -e(eF)^{1/2} \tag{3}$$

reduction in the work function of the metal. For $F = 1000$ volts/cm, $x_m \equiv 10^{-5}$ cm and $\Delta\phi \equiv 10^{-2}$ eV; the actual change in the work function is thus small. If a field is present, the work function ϕ in the Richardson formula should be replaced by $(\phi - \Delta\phi)$. Hence, the current increases by the factor given by notation (4).

$$\exp\left[\frac{e}{kT}(eF)^{1/2}\right] \tag{4}$$

According to this interpretation, a plot of the logarithm of the current versus the square root of the anode voltage should yield a straight line. An example is given in **Fig. 2** for tungsten; the deviation from the straight line for low anode voltages is due to space-charge effects. *See* SPACE CHARGE.

The straight portion of the line (the Schottky line) confirms the interpretation; the true saturation current for zero field is obtained by extrapolation of the Schottky line as indicated. Detailed studies have shown extremely small periodic deviations with reference to the Schottky line; these deviations are interpreted on the basis of the wave-mechanical theory describing the motion of electrons across the image potential barrier shown in Fig. 1. Adrianus J. Dekker

Schrödinger's wave equation

A linear, homogeneous partial differential equation that determines the evolution with time of a quantum-mechanical wave function.

Quantum mechanics was developed in the 1920s along two different lines, by W. Heisenberg and by E. Schrödinger. Schrödinger's approach can be traced to the notion of wave-particle duality that flowed from A. Einstein's association of particlelike energy bundles (photons, as they were later called) with electromagnetic radiation, which, classically, is a wavelike phenomenon. For radiation of definite frequency f, each bundle carries energy hf. The proportionality factor, $h = 6.626 \times 10^{-34}$ joule-second, is a fundamental constant of nature, introduced by M. Planck in his empirical fit to the spectrum of blackbody radiation. It is the quintessential parameter of quantum mechanics. This notion of wave-particle duality was extended in 1923 by L. de Broglie, who postulated the existence of wavelike phenomena associated with material particles such as electrons. *See* PHOTON; WAVE MECHANICS.

Energy equation. This was the direction pursued by Schrödinger. There are certain purely mathematical similarities between classical particle dynamics and the so-called geometric optics approximation to propagation of electromagnetic signals in material media. For the case of a single (nonrelativistic) particle moving in a potential $V(\mathbf{r})$, this analogy leads to the association with the system of a wave function, $\Psi(\mathbf{r})$, which obeys Eq. (1). Here m is the mass of

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi = E\Psi \tag{1}$$

the particle, E its energy, $\hbar = h/(2\pi)$, and ∇^2 is the laplacian operator. *See* CALCULUS OF VECTORS; GEOMETRICAL OPTICS; LAPLACIAN.

This equation may be applied to the hydrogen atom, taking for V the Coulomb potential $V = -e^2/r$, where e is the electron charge. From a purely mathematical point of view, the equation of course has solutions for any value of the energy. However, even without a view as to what the wave function signifies, it may be postulated on intuitive grounds that, in order to be physically acceptable, solutions ought to be well behaved, that is, single-valued, nonsingular, and bounded as the distance r goes to infinity. With these conditions imposed, it is found, for energies less than 0, that only certain discrete energies are allowed. The energy spectrum that emerges in this way is exactly that of the older Bohr model of the atom. It is in good, though not quite perfect, agreement with experiment. *See* ATOMIC STRUCTURE AND SPECTRA; ENERGY LEVEL (QUANTUM MECHANICS).

The so-called energy equation (1) stands in analogy with the equation governing an optical (that is, electromagnetic) disturbance of definite frequency f propagating in a medium with a space-dependent index of refraction, $n(\mathbf{r})$. With $\Phi(\mathbf{r},t) = u(\mathbf{r}) \sin 2\pi ft$ representing the amplitude of the disturbance, the

equation for $u(\mathbf{r})$ is Eq. (2).

$$-\nabla^2 u = \frac{n^2}{c^2}\omega^2 u \qquad \omega = 2\pi f \tag{2}$$

See ELECTROMAGNETIC RADIATION.

Time-dependent equations. In the electromagnetic case, the disturbance of course need not be of definite frequency. For the general situation corresponding to an arbitrary superposition over frequencies, the amplitude obeys the wavelike equation (3). For

$$-\nabla^2 \Phi = \frac{n^2}{c^2}\frac{\partial^2}{\partial t^2}\Phi \tag{3}$$

the particle case, a superposition over frequencies would correspond to a superposition of motions of different energies. That is without meaning classically. Classically, a particle in its motions has some definite energy. Nevertheless, it is possible to set that problem aside and ask what more general equation a time- as well as space-dependent wave function, $\Psi(\mathbf{r},t)$, might obey. Here the analogy with electromagnetism beckons. Instead of a wave equation involving second-order time derivatives, what suggests itself is Eq. (4), which is now called the Schrödinger

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi \tag{4}$$

equation. It involves only a first-order time derivative.

The energy equation (1) corresponds to a monochromatic solution of the form $\Phi = u(\mathbf{r})e^{i\omega t}$, with $E = \hbar\omega$ the particle energy. A general solution of Eq. (4) corresponds to a superposition over energies. But such a superposition raises the question of what it means to superpose energies, and, for that matter, what is the physical significance of Ψ.

Multiparticle systems. The wave function can be generalized to a system of more than one particle, say N of them. A separate wave function is not assigned to each particle. Instead, there is a single wave function, $\Psi(\mathbf{r}_1,\mathbf{r}_2, \ldots, \mathbf{r}_N,t)$, which depends at once on all the position coordinates as well as time. This space of position variables is the so-called configuration space. The generalized Schrödinger equation is Eq. (5), where the potential V may now depend on

$$i\hbar\frac{\partial\Psi}{\partial t} = -\sum_{i=1}^{N}\frac{\hbar^2}{2m_i}\nabla_i^2\Psi + V\Psi \tag{5}$$

all the position variables. Three striking features of this equation are to be noted:

1. The complex number i (the square root of minus one) appears in the equation. Thus Ψ is in general complex. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

2. The time derivative is of first order. Thus, if the wave function is known as a function of the position variables at any one instant, it is fully determined for all later times.

3. The Schrödinger equation is linear and homogeneous in Ψ, which means that if Ψ is a solution so is $c\Psi$, where c is an arbitrary complex constant. More generally, if Ψ_1 and Ψ_2 are solutions, so too is the linear combination $c_1\Psi_1 + c_2\Psi_2$, where c_1 and c_2 are arbitrary complex constants. This is the superposition principle of quantum mechanics. *See* SUPERPOSITION PRINCIPLE.

Probability interpretation. The Schrödinger equation suggests an interpretation in terms of probabilities. Provided that the wave function is square integrable over configuration space, it follows from Eq. (5) that the norm, $\langle\Psi|\Psi\rangle$, is independent of time, where the norm is defined by Eq. (6). It is possible

$$\langle\Psi \mid \Psi\rangle = \int d^3x_1 d^3x_2 \ldots d^3x_N \Psi^*\Psi \tag{6}$$

to normalize Ψ (multiply it by a suitable constant) to arrange that this norm is equal to unity. With that done, the Schrödinger equation itself suggests that expression (7) is the joint probability distribution at

$$\Psi^*\Psi d^3x_1 d^3x_2 \ldots d^3x_N \tag{7}$$

time t for finding particle 1 in the volume element d^3x_1, particle 2 in d^3x_2, and so forth.

Quantum mechanics, that is, deals with probabilities. All that can be known about the state of a system at any instant is contained in its wave function; and the wave function, though it evolves deterministically with time, yields in general only probabilistic predictions about the outcomes of measurements that are to be made on the system. There are two broad classes of problems that arise within quantum mechanics: (1) to determine, in general, the possible outcomes of measurements of various physical observables (energy, angular momentum, linear momentum, particle location, and so forth); and (2) to determine, for any given state Ψ, the probability distribution over those outcomes. For position measurements all locations are possible, with the probability distribution as given above. For other kinds of observables the rules are more complicated, but they are definite. For instance, in the special case of energy the possible outcomes (as illustrated on the example of a one-particle system) are the energies, E, for which Eq. (1), Schrödinger's starting point, has well-behaved solutions. *See* PROBABILITY (PHYSICS).

Generalizations. The Schrödinger equation and the rules of quantum-mechanical interpretation can be easily extended to incorporate spin, a sort of intrinsic angular momentum property possessed by certain species of particles (for example, electrons, protons, and neutrons). Extensions that incorporate the dictates of special relativity pose much deeper issues. They lead to the Dirac equation for the relativistic electron, a generalization of the Schrödinger equation that embodies its fundamental characteristics (the superposition principle, probabilistic interpretations, and so forth). *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; RELATIVISTIC QUANTUM THEORY; SPIN (QUANTUM MECHANICS); WAVE EQUATION; WAVE MOTION.

Sam Treiman

Bibliography. N. Bohr, *Atomic Physics and Human Knowledge*, 1958, reprint 1987; W. Heisenberg, *Physics and Philosophy*, 1959; M. Jammer, *The Conceptual Development of Quantum Mechanics*, 1966, reprint 1989; E. Schrödinger, *Collected Papers on Wave Mechanics*, 1928; J. Wheeler and W. Zurek (eds.), *Quantum Theory and Measurement*, 1983.

Schuler pendulum

Any apparatus which swings, because of gravity, with a natural period of 84.4 min, that is, with the same period as a hypothetical simple pendulum whose length is the Earth's radius. In 1923 Max Schuler showed that such an apparatus has the unique property that the pendulum arm will remain vertical despite any motions of its pivot. It is therefore useful as a base for navigational instruments. Schuler also showed how gyroscopes can be used to increase the period of a physical pendulum to the desired 84.4 min.

Gyrocompasses employ the Schuler principle to avoid errors due to ship accelerations. The principle has become the foundation of the science of inertial navigation.

Basic principle. The principle is most easily explained with the simple pendulum of **Fig. 1***a*. The period T for one complete swing of this device is given by Eq. (1), where l is the length of the pendulum and g is the acceleration, due to gravity, of a freely falling object (32.2 ft/s^2 or 9.8 m/s^2). If length l could be made equal to the Earth's radius R, as in Fig. 1*b*, gravity would hold the bob fixed, and the pendulum arm would stay vertical for all motions of point P. If l equals R, the period becomes 84. 4 min.

$$T = 2\pi\sqrt{\frac{l}{g}} \tag{1}$$

While a pendulum of such length cannot be built, the same effect can be achieved using the solid pendulous rod of Fig. 1*c*. Although only a few feet long, this pendulum can be given a period of 84.4 min by placing the pivot close to the center of mass. (If the pivot were at the mass center, the period would be infinitely long; that is, the rod would remain indefinitely in any position.)

Figure 2*b* demonstrates that a Schuler-period pendulum will also stay locally vertical during accelerations of its pivot, just as the simple pendulum did in Fig. 1*b*. The proof is by comparison. In Fig. 2*a* the pivot is high, and the pendulum swings back from vertical when the pivot is accelerated. In Fig. 2*c* the pivot is at the mass center, and the pendulum does not swing at all, so that after traveling some distance around the Earth it is forward of the local vertical. In Fig. 2*b* the pivot is placed at just the right point, so that whatever the acceleration of the pivot, the pendulum swings back just enough to stay locally vertical.

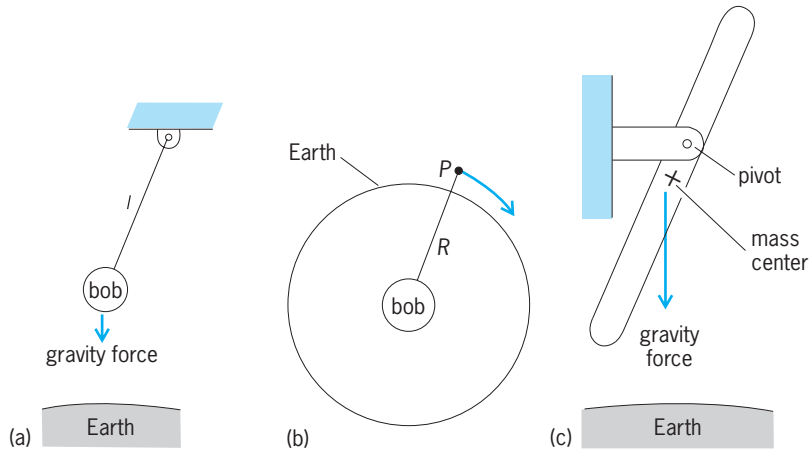

Fig. 1. Basic principles illustrated by (*a*) simple pendulum; (*b*) Earth-radius pendulum; (*c*) solid pendulous rod.

Unfortunately, for a 2-ft (0.6-m) pendulum like that of Fig. 1*c* to have an 84.4-min period, the pivot must be placed only about 0.2×10^{-6} in. (5 nanometers) above the mass center, an unrealizable physical tolerance.

Gyropendulum and gyrocompass. The period of a pendulous mass can be greatly increased by attaching to it a rapidly spinning gyro wheel as, for example, in **Fig. 3**. Extra freedom, provided by the thrust bearing in Fig. 3, is necessary to permit the gyroscopic action.

The gyrocompass is a special pendulous gyroscope that is sensitive to Earth rotation. If the pendulous mass is m, the horizontal spin axis has a natural period of oscillation, about north, given by Eq. (2),

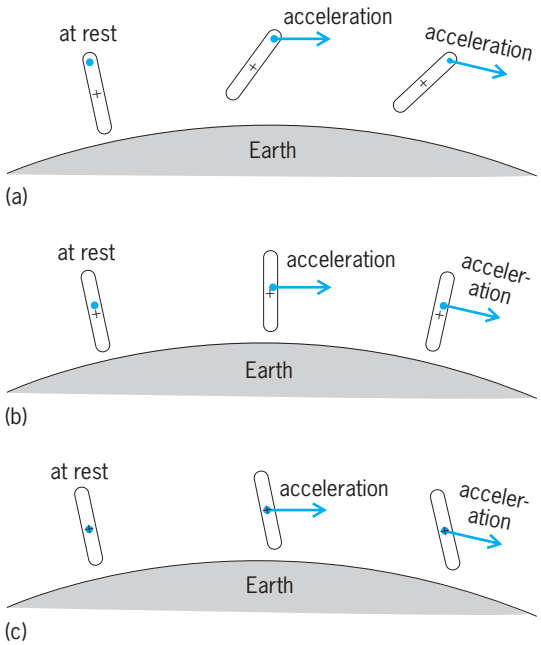

Fig. 2. Behavior of pendulous rod when accelerated about Earth. (*a*) Short-period pendulous rod. (*b*) Schuler-period pendulous rod. (*c*) Nonpendulous rod.

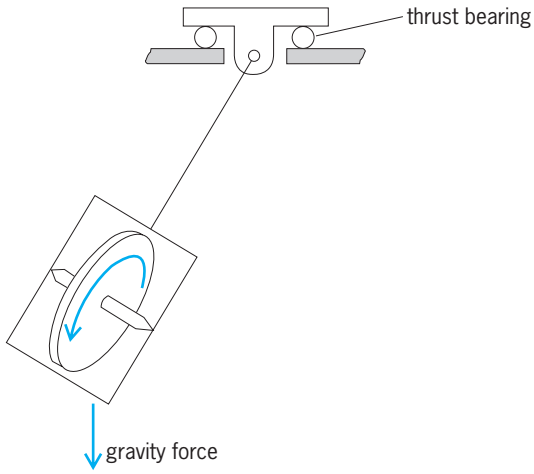

Fig. 3. Gyropendulum; thrust bearing provides extra freedom, necessary for gyroscopic action.

where H is gyro momentum, λ is latitude, and Ω_e is

$$T = 2\pi\sqrt{\frac{H}{mgl\Omega_e \cos\lambda}} \qquad (2)$$

Earth rate (360°/day). Ship-borne gyrocompasses are rendered insensitive to ship accelerations by making $T = 84.4$ min. *See* GYROCOMPASS.

Inertial guidance. The stable platform of an inertial guidance system is made to behave like a pendulum by the use of feedback from an accelerometer mounted on the platform. The accelerometer signal controls a restoring torque to the platform.

The principle is demonstrated in **Fig. 4**. The accelerometer consists of a spring-restrained mass. If the platform tilts, gravity pulls the mass to one side; the amount is measured electrically, integrated, and used to drive the platform motor. (Actual inertial platforms are controlled by a gyro, which is precessed at a rate by the integrator signal.) *See* ACCELEROMETER.

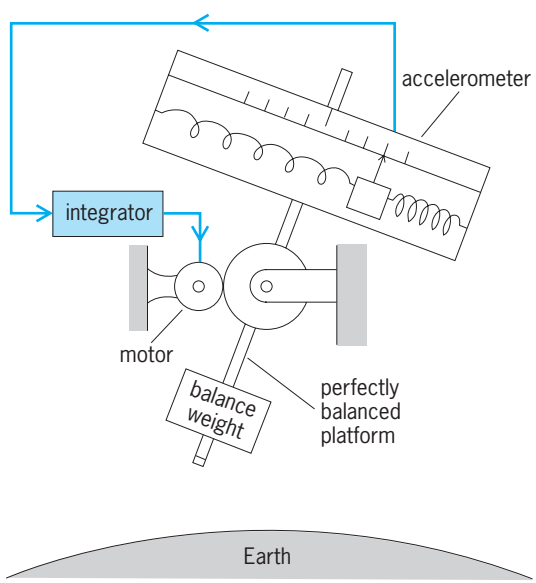

Fig. 4. Principle of Schuler-tuned stable platform.

The system of Fig. 4 behaves like the physical pendulum of Fig. 1*a*. When either device is tilted from vertical, gravity produces a restoring action so that the device swings back and forth about the vertical. The period of swing is $2\pi\sqrt{l/g}$ for the pendulum of Fig. 1*a*, and $2\pi\sqrt{1/Kg}$ for the platform of Fig. 4, where K is the loop gain from tilt angle to motor rate. The platform is accordingly given an 84.4-min period by making K equal to $1/R$.

If the platform of Fig. 4 is installed in a vehicle and is once leveled, it will always remain locally level despite motions of the vehicle about the Earth. Then the signal from the accelerometer will represent only vehicle acceleration, uncontaminated by gravity, and can be used to determine vehicle position. *See* INERTIAL GUIDANCE SYSTEM.

Robert H. Cannon

Sciatica

A specific pattern of leg pain that follows the course of the sciatic nerves. It is usually caused by injury to one or more of the sciatic root nerves within the lower spinal canal. The contributing nerves, the last two lumbar and first two sacral, exit sequentially through openings on each side of the spine. The individual nerves then enter common protective sheaths to become the sciatic nerves. Each one leaves the pelvis and runs down the back of the leg to the foot. Along the way branches supply the skin, particularly in the lower leg, and some of the muscles that move the hip and bend the knee. The sciatic nerves control all the muscles that enable walking on the heels and toes.

Sciatic pain is radicular pain, that is, pain that follows the distribution of a nerve. The typical pain locations are in the back of the thigh, behind the knee, over the calf, and into one or both sides of the foot as far as the toes. Due to the length of the sciatic nerve, the precise location of the pain varies, but it is always constant and dominant in the leg as opposed to the back. In the acute phase the intensity may fluctuate in response to movement or position, but the pain never disappears completely. Sciatica may be associated with diminished sensation or decreased power in the muscles supplied by the affected nerve roots. More often the pain occurs with signs of root irritation only but with otherwise normal function.

Sciatica results from the combination of mechanical pressure and chemical insult acting on a nerve root. Pressure alone produces a highly unpleasant tingling or burning sensation, but when accompanied by a local inflammatory response the result is excruciating pain. Because the physical deformation of a nerve root caused by a herniated disc is always associated with an inflammatory reaction, an acute disc rupture is by far the most frequent cause of sciatica. Other but rare possibilities include pyogenic infections with abscess formation, synovial cysts (cysts occurring in proximity to facet joints in the spine), diabetic neuropathy, vasculitis (inflammation of blood or lymph vessels), injected irritating agents,

and a range of metabolic and autoimmune disorders producing neuritis (nerve inflammation).

The term sciatica is frequently misapplied to a variety of other conditions that cause leg pain. It is most often confused with the referred leg pain that accompanies common backache. In these cases the back pain is dominant while the leg symptoms are typically intermittent and do not follow the path of the sciatic nerve. *See* PAIN; SPINE; VERTEBRA.

Hamilton Hall

Bibliography. D. Borenstein, *Back in Control*, M. Evans and Co., New York, 2001; J. Frymoyer (ed.), *The Adult Spine: Principles and Practice*, 2d ed., Lippincott–Raven, Philadelphia, 1997; H. Hall, *A Consultation with the Back Doctor*, McClelland & Stewart, Toronto, 2003.

Science

The study of nature and natural phenomena. Application of the term did not begin with any formal definition; rather the various disciplines arose independently, each in response to some particular need. It was then observed that certain of these disciplines had enough traits in common to justify classifying them together as one of the sciences. Usage is not, however, always unanimous as to whether some disciplines should be called sciences, and there is often lively controversy as to the propriety of speaking of the social or historical sciences.

Usually a science is characterized by the possibility of making precise statements which are susceptible of some sort of check or proof. This often implies that the situations with which the special science is concerned can be made to recur in order to submit themselves to check, although this is by no means always the case. There are observational sciences such as astronomy or geology in which repetition of a situation at will is intrinsically impossible, and the possible precision is limited to precision of description. There is also usually the implication that the subject matter of the individual science is something in the world of phenomena. Thus it is not usual to speak of the "science" of mathematics or the "science" of logic, even though both these disciplines are capable of the highest precision.

A common method of classifying sciences is to refer to them as either exact sciences or descriptive sciences. Examples of the former are physics and, to a lesser degree, chemistry; and of the latter, taxonomical botany or zoology. The exact sciences are in general characterized by the possibility of exact measurement. Measurement is fundamentally description by the use of numbers. Given the system of measurement and the measuring numbers for any special situation, that situation has been adequately described if it is possible to reconstruct a situation such that measurement on it gives the same numbers. Because mathematics operates to a large extent with numbers, systems that are subject to exact measurement are also susceptible of mathematical analysis; this susceptibility is one of the most important characteristics of the exact sciences. One of the most important tasks of a descriptive science is to develop a method of description or classification that will permit precision of reference to the subject matter. *See* PHYSICAL SCIENCE; SCIENTIFIC METHODS.

Percy W. Bridgman; Gerald Holton

Scientific and applications satellites

Scientific satellites are used to gain basic knowledge. They may be satellites of Earth or of other planets or their moons. Satellites used to apply space knowledge and techniques to practical purposes are called applications satellites. Many spacecraft are used for multiple purposes, combining space exploration, science, and applications in various ways. Scientific satellites are often called research satellites, and applications satellites are commonly designated by the application field, for example, navigation satellites. *See* SATELLITE (SPACECRAFT).

Many types of investigations are possible with scientific satellites. Some use the vantage point of space for Earth and astronomical observations. Others use the space environment for experiments not possible in the laboratory, such as determining the behavior of biological specimens or the growth of crystals under long periods of weightlessness. Astronomical telescopes, spectrometers, Geiger counters, ionization chambers, pressure gages, and magnetometers are among the instruments used in scientific satellites.

Space environment observations. The space environment was one of the first areas investigated in scientific satellites. In this environment are cosmic rays, dust, magnetic fields, and various radiations from the Sun and galaxies. A variable solar wind of charged particles blows continually throughout the solar system. More energetic solar particles are called solar cosmic rays. *See* COSMIC RAYS; SOLAR MAGNETIC FIELD; SOLAR WIND.

The famous Van Allen radiation belts were discovered by the first United States satellites. Soon thereafter satellites and space probes confirmed the existence of the solar wind. The continued investigation of these phenomena by satellites and space probes has led to the knowledge that a planet with a strong magnetic field, like Earth, is surrounded by a complex region, called a magnetosphere, which plays a major role in magnetic storms and auroras and in trapping solar particles to form radiation belts. *See* MAGNETOSPHERE; SPACE PROBE; VAN ALLEN RADIATION.

Meteoric particle studies seek to determine the size, velocity, and energy distributions of space dust particles, their composition and likely origin, and the physical processes that occur on impact of the particles with a surface. *See* INTERPLANETARY MATTER; METEOR.

Numerous investigations are made of the effect of the space environment on materials and processes, including effects of radiation damage, meteoric erosion processes, and extremely high vacuum for very

long periods. The influence of prolonged weightlessness on welding, alloying metals, growing crystals, and biological processes constitutes a promising field for future practical applications. *See* SPACE PROCESSING.

Earth science. Scientific satellites, sounding rockets, and space probes have had a profound effect on earth science. It is now possible to solve problems of the upper atmosphere and ionosphere that could not be resolved by observations from the ground. The Earth and its atmosphere can now be studied in comparison with the other planets and their atmospheres, providing greater insight into the formation and evolution of the solar system. *See* IONOSPHERE; PLANETARY PHYSICS; SOLAR SYSTEM.

The combined results of measurements from Earth, lunar, and other planetary satellites, space probes, and crewed lunar missions have added significantly to scientists' knowledge of Earth and other planets. For example, these observations have shown that for 5×10^8 years or more after their formation the planets were subjected to bombardment by huge asteroid-sized meteorites; that Mars is a very active planet; and that large bodies like the Moon and planets will necessarily undergo substantial, even violent, evolution after their formation, so that much of the record of the very earliest days of the solar system will have to be sought from smaller objects like comets and asteroids, primarily by using space probes such as *Deep Impact*, which triggered a substantial cometary explosion on July 4, 2005.

Satellites also make possible precision measurements of the size and shape of Earth and its gravitational field, and even measurements of the slow drifting of continents relative to each other. *See* CONTINENTAL DRIFT; EARTH, GRAVITY FIELD OF; GEODESY.

Astronomy. Satellites have also had a great influence on astronomy by making it possible to observe the Sun, stars, and other celestial objects in all the wavelengths that reach the vicinity of Earth, whereas the ionosphere and atmosphere prevent most of these radiations from reaching telescopes on the ground. Moreover, even in the visible wavelengths, the small-scale turbulence and continuously varying refraction in the lower atmosphere distort the image of a stellar object viewed at the ground, so that there is a limit to the improvement that can be achieved by increasing the size of a ground-based telescope. It has been estimated that a 12-in.-diameter (30-cm) optical telescope in a satellite is as effective in resolution as a 100-in.-diameter (254-cm) telescope at Earth's surface. Also, the Hubble Space Telescope with its 94-in. (2.4-m) mirror is living up to its promise of making observations of objects nearly 100 times fainter, or 10 times farther away, than can now be observed with the best ground-based instruments. *See* HUBBLE SPACE TELESCOPE; TELESCOPE.

The importance of being able to look deeper into space arises from the fact that the farther away an object is, the longer its light has taken to get to Earth. Since the radiation from a celestial object reveals the object as it was when emitting that radiation, the more distant objects will let scientists look farther back in time. By increasing the distance that can be seen into space, the chances are increased that the time when the universe was generated may be seen.

The importance of being able to observe in all wavelengths arises from the fact that theory shows that vital information about the birth, evolution, and demise of celestial objects is contained in the infrared, ultraviolet, and other wavelengths that do not reach the ground.

With satellite instruments a hitherto unobtainable ultraviolet map of the sky has been made. Hundreds of x-ray sources have been discovered. Huge clouds of hydrogen, millions of miles across, have been detected around comets. The Sun has been shown to be even more complex than previously supposed, and a great deal of information has been obtained about sunspots, solar flares, and solar magnetic fields. Numerous solar observatories have been launched, along with radio, infrared, ultraviolet, x-ray, and gamma-ray astronomy satellites. *See* ASTRONOMY; COMET; GAMMA-RAY ASTRONOMY; INFRARED ASTRONOMY; RADIO ASTRONOMY; SUN; ULTRAVIOLET ASTRONOMY; X-RAY ASTRONOMY.

Applications satellites. The applications satellites emphasize the continuing day-to-day practical utilization of the satellite. They are operational in nature, although some of them, either directly or as a by-product of their operational output, contribute significantly to research, for example, with meteorological or Earth survey satellites. An applied research program generally precedes establishment of an applications satellite system. Thus, *Tiros* and *Nimbus* laid the groundwork for operational meteorological satellites. Likewise, *Syncom* satellites preceded the *Intelsat* communications system.

Applications satellites fall generally into two distinct classes: (1) those using satellite radio communications as the essential element; and (2) those that are directed toward providing observational or survey data. In the first class are the now-familiar radio and television relay satellites like *Intelsat* and DirecTV; and navigation, geodetic, data collection and relay, disaster warning, air-traffic control, and maritime communications satellites. In the second class are meteorological satellites and satellites for the survey and monitoring of Earth's natural and cultural resources in such areas as geology and mining, hydrology, glaciology, sea ice, agriculture, forestry, land use, and civil engineering.

The earliest commercial space field to become fully established is that of communications satellites, with which countries all over the world are involved. In a very real sense the satellite for communications substitutes for a very high antenna mast. Because of its height the satellite is particularly useful for long-distance communications. Since stations far apart are within the line of sight of the satellite, microwave techniques can be extended from a distance of a few tens to many thousands of miles, eliminating dependence on the highly variable ionosphere and greatly extending the range of frequencies which may be used for long-distance communications. For

these purposes the geostationary orbit altitude of 22,237 mi (35,786 km) has been found especially effective. *See* COMMUNICATIONS SATELLITE.

Navigational satellites are used to determine position on the surface of Earth, supplementing celestial and radio navigation techniques. Marine and aeronautical navigational systems based on the Global Positioning System are in operation. *See* SATELLITE NAVIGATION SYSTEMS.

The meteorological satellite system obtains data for the preparation of weather forecasts and for the detection and surveillance of unusual weather conditions such as hurricanes and tornadoes. The complete system consists of a number of satellites in low-altitude—several-hundred-mile—Sun-synchronous, near-polar orbits, and a number in geosynchronous orbits over the Equator. The system affords a constant awareness of weather conditions over virtually the whole world. Optical and infrared imaging furnish large-scale pictures of cloud patterns and weather systems, while other instruments give temperature profiles of the lower atmosphere and stratosphere and some estimates of humidity and winds. *See* METEOROLOGICAL SATELLITES.

Cartographic and geodetic satellites are used as adjuncts to conventional techniques. With cartographic satellites the object is to prepare maps of Earth's surface and of the culture upon it. The geodetic satellite aids in determining the shape of Earth's surface—the geoid. Cartographic satellites use optical camera techniques, essentially identical to those of aerial photography. Geodetic satellites may use optical techniques in which angles of vision of the satellite as seen from Earth stations are measured; radio techniques in which both distance and angles are measured; or indirect techniques in which the shape of the geoid is inferred from the satellite orbit, in particular from the perturbations of this orbit resulting from the deviations of the geoid from a perfect sphere. *See* CARTOGRAPHY.

Imaging of Earth in various wavelengths from the ultraviolet to the infrared and radio regions yields a wide range of detail about Earth's surface structure and physical properties, vegetation, and cultural development. Most of the imaging is done by passive techniques using radiations emitted by the object being viewed, but much useful work can be done with active methods like radar. In the broad perspective of images covering thousands or tens of thousands of square miles, features stand out that were not recognized on foot or from aircraft. Numerous previously unknown faults and fracture systems of the kind associated with mineral ore and oil deposits have been identified. The combination of satellite and ground techniques for geological survey and prospecting promises to be especially fruitful.

A staging technique in which accurate information about some resource—a timber stand, for example—obtained on the ground over a small area is extrapolated to a larger area by aircraft imaging, and then further extrapolated to a vastly greater area by satellite imaging, is proving to be very accurate, economical, and time-saving for surveys of agricultural crops, grazing lands, and forests. Total crop productions can be estimated in advance. Tree or crop diseases can be spotted early and the total extent and location of the infection can be estimated. Watershed managers can use the satellite images to help in assessing the amount of water in reservoirs and in mountain snow and ice. Sea ice, hazardous to shipping, can be kept under surveillance. Many states use satellite images to produce land-use maps which are important for helping to guide the continuing development of the state and its resources.

Orbit. The orbits for scientific and applications satellites depend on the requirements of the investigation or use, and also on the nature of the local environment which may adversely affect the purpose or life of the satellite. For example, a satellite to study the radiation belts would be placed in an orbit traversing the belts, but an astronomical satellite with sensitive electronic instruments should be kept clear of the radiation in the belts.

Below about 100 mi (160 km) air drag on a satellite is prohibitively great. Between 100 and 200 mi (160 and 320 km) altitude air drag is still appreciable and can significantly affect the orbit of a satellite. Above 200 mi (320 km) air drag may generally be ignored except for precision measurements. Ionospheric effects on radio signals may have to be corrected for, when using satellites for such purposes as navigation or geodetic studies. Even solar radiation pressure may have to be taken into account in using satellites for air density or geodetic measurements.

In general, the remarks on orbits for Earth satellites apply to satellites of other planets as well, but specific characteristics of satellite orbits must be calculated for each planet.

Spacecraft. Scientific satellites divide conveniently into two groups: the Explorer class (named after the early *Explorer* satellites) and the observatory class. In general, the Explorer-class satellites are special-purpose and lighter than the observatory class. Typical weights are from 150 to 500 lb (70 to 230 kg), although some may be less, some more. Many dozens of Explorer-class scientific and applications satellites have been launched. Those that made the early discoveries in the radiation belts were of the Explorer class.

Observatory-class satellites tend to be heavier than Explorer-class satellites, ranging from below 1000 lb (450 kg) to many tons in weight. They also usually have more complicated systems for power, telemetering, and orientation, and often combine a large number of experimental objectives in one mission. The orbiting solar observatories (OSOs), the orbiting astronomical observatories, and the *Nimbus* meteorological satellites are all examples of the observatory class. *See* SPACE COMMUNICATIONS; SPACE POWER SYSTEMS; SPACE TECHNOLOGY; SPACECRAFT STRUCTURE.

International interest. Since the Soviet Union and the United States pioneered the launching of space rockets, interest in the use of satellites for science and applications has become worldwide. Other countries, including Italy, France, Japan, China, and

India, have launched their own satellites. Over 70 nations have been associated in various ways with the United States space program, many of them with scientific and applications satellites, sometimes furnishing instruments for United States satellites, other times providing complete satellites for launching by the United States. Homer E. Newell; John F. Clark

Bibliography. S. Q. Kidder and T. H. Vonder Haar, *Satellite Meteorology: An Introduction*, 1995; National Research Council, *Mission to Planet Earth*, 1988; National Research Council, *Solar and Space Physics*, 1988; C. B. Pease, *Satellite Imaging Instruments*, 1991; G. Seeber, *Satellite Geodesy*, 1993; K.-H. Szekielda, *Satellite Monitoring of the Earth*, 1989; J. Wertz and W. Larson (eds.), *Space Mission Analysis and Design*, 3d ed., 1999.

Scientific methods

Strategies or uniform rules of procedure used in some scientific research with a measure of success. Scientific methods differ in generality, precision, and the extent to which they are scientifically justified. Thus, whereas the experimental method can in principle be used in all the sciences dealing with ascertainable facts, the various methods for measuring the electron charge are specific. The search for increasing quantitative precision involves the improvement or invention of special methods of measurement, also called techniques. All scientific methods are required to be compatible with confirmed scientific theories capable of explaining how the methods work. The most general of all the methods employed in science is called the scientific method.

The scientific method. All the sciences referring to real or putatively real things are supposed to abide by the scientific method. The latter may be summarized as the following sequence of steps: identification of a knowledge problem; precise formulation or reformulation of the problem; examination of the background knowledge in a search for items that might help solve the problem; choice or invention of a tentative hypothesis that looks promising; conceptual test of the hypothesis, that is, checking whether it is compatible with the bulk of the existing knowledge on the matter (for it might be a wild conjecture not worth pursuing); drawing some testable consequences of the hypothesis; design of an empirical (observational or experimental) test of the hypothesis or a consequence of it; actual empirical test of the hypothesis, involving a search for both favorable and unfavorable evidence (examples and counterexamples); critical examination and statistical processing of the data (for example, calculation of average error and elimination of outlying data); evaluation of the hypothesis in the light of its compatibility with both the background knowledge and the fresh empirical evidence; if the test results are inconclusive, design and performance of new tests, possibly using different special methods; if the test results are conclusive, acceptance, modification, or rejection of the hypothesis; if the hypothesis is acceptable, checking whether its acceptance forces some change (enrichment or correction) in the background knowledge; identifying and tackling new problems raised by the confirmed hypothesis; and repetition of the test and reexamination of its possible impact on existing knowledge. The case of the introduction and checking of new procedures is parallel, except that in this case reliability and accuracy, not truth, are at stake. So is the case of the introduction and checking of artifacts such as machines or organizations, except that in this case the checking is for efficiency. *See* SYSTEMS ENGINEERING.

An alternative interpretation of the scientific method is as a sequence of problems. A typical example is the empirical problem of finding a particular value of a certain magnitude M, or the parallel theoretical problem of explaining why the value of M happens to be m. The experimenter's problem is to determine the measured value of M with relative error smaller than a preassigned number e. The theorist's problem is to determine what premises imply that the value of M is m, or whether m is to be found only by empirical means. The next experimental problem is to determine which experimental device would make it possible to measure M with error less than e. The corresponding theoretical problem is to determine which theory, subsidiary hypotheses, and data imply that M is worth m. A subsequent experimental problem is to determine the value of M that a measurement with the help of the chosen device yields. A corresponding theoretical problem is to determine the value of M according to the theory, subsidiary hypothesis, and data. From here on, the questions are addressed to both experimenter and theorist. They seek to find what the result implies or suggests; how the result could be corroborated independently, that is, by alternative procedures; whether the new result is more precise and plausible than the results obtained by alternative procedures; if so, what the result implies or suggests; and, if not, what may have to be altered in the preceding operations.

Experimental method. When the scientific method involves experiment, that is, the deliberate controlled modification of some properties (factors or variables), it is called the experimental method. This method involves the design and operation of an experimental setup. Such a setup often includes one or more measuring instruments. When the objects of experiment are very similar, they are compared before and after the variable or variables in question have been altered. When the objects of experiment exhibit significant variations, as is the case with organisms, in particular humans, a large collection of them must be studied. Every such collection is divided into two roughly equal parts: the experimental and the control groups. To avoid bias, the members of each collection are picked at random and, in the case of humans, the controls are given placebos. (In more refined experiments the control group is split into two, only one of which is given a placebo, and the experimenter does not know beforehand who will be given what.) Only the members of the experimental group are subjected to the stimulus whose

effects the experimenter wishes to find out. The simplest experimental design involves the variation of one variable at a time. More complex experimental designs allow for the simultaneous variation of two or more variables. After the stimulus has been applied, the variables of interest are observed or measured in the two groups. If a difference is observed, a statistical significance test is applied to find out whether the difference is genuine or due to small individual differences or to random errors. The procedure is usually repeated by the same observer or by an independent investigator to check for possible errors in design, execution, or interpretation. *See* EXPERIMENT; STATISTICS.

Discovering, inventing, and checking. The scientific method is not a recipe for making original discoveries or inventions; it does not prescribe the pathway that scientists must follow to attain success. Nor is it a substitute for creativity and resourcefulness. The actual pathway of scientific research is messy and partly at the mercy of unforeseen accidents, both lucky and unlucky. Moreover, this pathway depends on the subject matter as well as on the individual's imagination and experience. The goal of the scientific method is to ascertain whether a hypothesis is true to some degree. Indeed, the nucleus of the scientific method is the confrontation of an idea (hypothesis) with the facts it refers to, regardless of the source of the idea in question. In sum, the scientific method is a means for checking hypotheses for truth rather than for finding facts or inventing ideas.

The scientific method has not always been understood in this fashion. For instance, Francis Bacon and many others thought that the scientific method is a simple and guaranteed recipe for discovery and invention. But in fact no one has come up with any algorithms for having original ideas. (This is one of the reasons that it is impossible at present to design creative computers.) Others have denied the very existence of the scientific method.

Since this is a question concerning facts, it can only be resolved by an empirical investigation concerning the way that people actually conduct scientific research, regardless of what they say about method. Such investigation is bound to show that the answer depends on the kind of problem. Routine problems can indeed be handled with the help of well-tried special methods (techniques) that can be described almost exhaustively, so much so that their application can be entrusted to apprentices or technicians. Original problems are something utterly different. It requires originality and experience to find an open and interesting problem that can be investigated in a finite time with limited resources. Once a promising problem has been identified, it must be stated or perhaps reformulated in a precise way, and a tentative plan to work on it must be formulated—again, an original task. The invention of new hypotheses and new methods also takes inspiration and luck in addition to hard work and discussions with colleagues. Finally, the hypothesis must be checked in order to find out whether it is true to some degree, and the method must be tried out to ascertain whether it accomplishes what it purports do and, if so, whether it is better than rival techniques. It is here, in checking, that the scientific method plays a decisive and distinctive role. In short, the scientific method is the way that scientists proceed to check ideas and techniques, not invent them.

Reach of the scientific method. Whether the scientific method can legitimately and fruitfully be used in the social sciences, such as history, and in the socionatural or hybrid sciences, such as psychology and linguistics, has been a subject of controversy. The history of these sciences shows conclusively that the scientific method has been fruitful wherever it has actually been employed.

Test of scientificity. Since the use of the scientific method is one characteristic of scientific research, its absence is a sure indicator of nonscience. In other words, a discipline where the scientific method plays no role is not a science. Thus, such fields as theology, literary criticism, psychoanalysis, homeopathy, graphology, and palmistry can hardly be regarded as scientific. *See* EMPIRICAL METHOD; HYPOTHESIS; SCIENCE.

Mario Bunge

Bibliography. M. Bunge, *Epistemology and Methodology I: Exploring the World*, 1983; M. Bunge, *Epistemology and Methodology II: Understanding the World*, 1983; H. Feigl and M. Brodbeck (eds.), *Readings in the Philosophy of Science*, 1953; E. Nagel, *The Structure of Science*, 1961, reprint 1979; A. Zucker, *Introduction to the Philosophy of Science*, 1996.

Scintillation counter

A particle or radiation detector which operates through emission of light flashes that are detected by a photosensitive device, usually a photomultiplier or a silicon PIN diode. The scintillation counter not only can detect the presence of a particle, gamma ray, or x-ray, but can measure the energy, or the energy loss, of the particle or radiation in the scintillating medium. The sensitive medium may be solid, liquid, or gaseous, but is usually one of the first two. The scintillation counter is one of the most versatile particle detectors, and is widely used in industry, scientific research, medical diagnosis, and radiation monitoring, as well as in exploration for petroleum and radioactive minerals that emit gamma rays. Many low-level radioactivity measurements are made with scintillation counters. *See* LOW-LEVEL COUNTING.

The scintillation phenomenon was used by E. Rutherford and colleagues, who observed the light flashes from screens of powdered zinc sulfide struck by alpha particles. The scattering of alpha particles by various foils was used by Rutherford to establish the modern notion of the nuclear atom. Since 1947, photomultipliers or PIN diodes, which feed into amplifiers, have replaced the eye and have also proved to be very efficient in measuring the time of arrival of a particle, as well as its energy or energy loss. Scintillations are light pulses emitted in the electronic decays of excited atoms or molecules, the

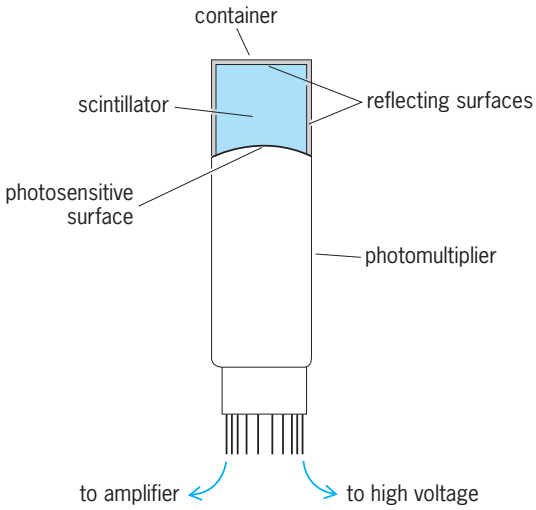

Diagram of a scintillation counter.

energy responsible for these excitations having been provided, directly or indirectly, by the detected particle. Scintillations are one example of luminescent behavior, and are related to the process called fluorescence. *See* FLUORESCENCE; LUMINESCENCE; PHOTOMULTIPLIER.

Operation of bulk counter. Following the innovation in 1947 of using bulk material, instead of screens of powdered luminescent materials, scintillation counters have been made of transparent crystalline materials, liquids, plastics, or glasses. In order to be an efficient detector, the bulk scintillating medium must be transparent to its own luminescent radiation, and since some detectors are quite extensive, covering meters in length, the transparency must be of a high order. One face of the scintillator is placed in optical contact with the photosensitive surface of the photomultiplier or PIN diode (see **illus.**). In order to direct as much as possible of the light flash to the photosensitive surface, reflecting material is placed between the scintillator and the inside surface of the container.

In many cases it is necessary to collect the light from a large area and transmit it to the small surface of a photomultiplier. In this case, a "light pipe" leads the light signal from the scintillator surface to the photomultiplier with only small loss. The best light guides and light fibers are made of glass, plastic, or quartz. It is also possible to use lenses and mirrors in conjunction with scintillators and photomultipliers. *See* OPTICAL FIBERS.

A charged particle, moving through the scintillator, loses energy and leaves a trail of ions and excited atoms and molecules. Rapid interatomic or intermolecular transfer of electronic excitation energy follows, leading eventually to a burst of luminescence characteristic of the scintillator material. When a particle stops in the scintillator, the integral of the resulting light output, called the scintillation response, provides a measure of the particle energy, and can be calibrated by reference to particle sources of known energy. Photomultipliers or PIN diodes may be operated so as to generate an output pulse of amplitude proportional to the scintillation response.

When a particle passes completely through a scintillator, the energy loss of the particle is measured. When a gamma ray converts to charged particles in a scintillator, its energy may also be determined. When the scintillator is made of dense material and of very large dimensions, the entire energy of a very energetic particle or gamma ray may be contained within the scintillator, and again the original energy may be measured. Such is the case for energetic electrons, positrons, or gamma rays which produce electromagnetic showers in the scintillator. Energy spectra can be determined in these various cases by using electronic equipment to convert amplitudes of the output pulses from the photomultiplier or PIN diode to digital form, for further processing by computers or pulse-height analyzers.

Characteristics. Scintillation counters have several characteristics which make them particularly useful as detectors of x-rays, gamma rays, neutrons, and other nuclear and high-energy particles.

Efficiency and size. Charged particles, x-rays, and low-energy gamma rays can be detected with efficiencies close to 100%. Small, thin detectors may be used to define a particle's position accurately. When detection of very energetic particles is desired, the scintillator can be made very large and massive, and viewed by many photomultipliers.

Speed. Useful scintillation materials emit light flashes which rise to peak intensity in less than 10^{-9} s and decay to half of this intensity in 10^{-9} s to 10^{-6} s. The rapid rise makes scintillators useful for fast timing measurements. Short decay times allow very fast counting while avoiding pileup, which is very characteristic of older and slower particle detectors such as the Geiger counter. The Cerenkov counter is similar to a scintillation counter, but depends only on the index of refraction of the transparent medium and the particles' velocity in the medium, and not on luminescence processes. *See* CERENKOV RADIATION.

Energy resolution. While the energy resolution attained by scintillation counters is useful, their performance in this respect is markedly inferior to that of semiconductor detectors such as silicon, germanium, or lithium-drifted germanium at low temperatures. However, semiconductor counters have been rather limited in size, so that they do not usually compete with the large inorganic scintillators needed in high-energy research. Often coincidence or anticoincidence combinations are made between semiconductor counters and scintillation counters. *See* JUNCTION DETECTOR.

Substances used. Scintillation counter materials are generally classified into crystal scintillators, glass scintillators, and organic scintillators. The organics include crystal, liquid, and plastic types.

Inorganic scintillators. Inorganic crystal scintillators are generally characterized by the presence of heavy elements. The most widely used inorganic scintillator is sodium iodide activated with a small amount of thallium salt. The usual designation of this material is NaI(Tl). Cesium iodide is also useful, and may be activated with thallium or sodium, or used in pure form. Other useful crystalline scintillators include calcium fluoride, barium fluoride, cadmium

tungstate, and bismuth germanate (BGO). NaI(Tl) and bismuth germanate are particularly useful for detecting gamma rays because of the presence of iodine and bismuth, respectively; the high atomic numbers of these elements (53 and 83, respectively) are important in both the photoelectric and pair production processes which result from the interaction of a gamma ray with the base material. *See* GAMMA-RAY DETECTORS.

Glass scintillators. Glass scintillators formed from lithium silicate and containing cerium as activator are used to detect thermal and low-energy (less than 0.1 MeV) neutrons. Glasses containing the isotope ^{6}Li are sensitive to neutrons through the charged particles released in the nuclear reaction ^{6}Li$(n,t)^4$He, which has a high cross section for thermal neutrons. Glasses of similar chemical composition but containing lithium depleted in the ^{6}Li isotope are insensitive to neutrons and can be used to monitor backgrounds due to other types of radiation.

Organic scintillators. Liquid and plastic scintillators are the most extensively used forms of organic scintillator. Organic crystals such as anthracene and stilbene are also used, but the liquids and plastics are favored for the ease with which they may be formed in arbitrary shapes, large or small, and for the economy they offer in achieving large, sensitive detecting volumes. The timing precision achieved by the faster liquids and plastics is typically in the range of 10^{-9} to 10^{-10} s. The solvents which form the bulk constituents of liquid and plastic scintillators are often xylene, toluene, benzene or mineral oil (liquids), or polystyrene or polyvinyltoluene (plastics). Other constituents of the liquid or plastic solutions are compounds such as terphenyl, diphenyloxazole, and 1,4-bis-{2-(5-phenyloxazolyl)}-benzene, known as POPOP.

Liquid scintillators are used in volumes ranging from a few tenths of a cubic inch or less through several hundreds of cubic feet. The largest scintillators are used in high-energy physics to study neutrino interactions. Liquid scintillators also have the capacity to assimilate other liquids or substances and form homogeneous media, thereby providing an efficient and simple means for measuring the products of nuclear reaction or radioactive decay in those substances. Plastic scintillators are used in a variety of forms ranging from films a few micrometers thick, or fibers less than 0.01 in. (0.25 mm) in diameter, to large sheets or blocks up to about 15 ft^3 (0.5 m^3) in volume. *See* NEUTRINO.

Pulse-shape discrimination. The response of most liquid scintillators can be improved significantly if dissolved oxygen is either removed from the solution or displaced by nitrogen or an inert gas. This procedure also endows many liquid scintillators with the ability to pulse-shape-discriminate, that is, to exhibit scintillation decay properties which depend on the type of ionizing particle causing the scintillation. Pulse-shape discrimination is also a strong effect in anthracene and stilbene crystals, but not in plastic scintillators. Most organic scintillators contain a high proportion of hydrogen, and therefore detect fast neutrons (energy greater than 0.1 MeV) via the proton recoils generated by internal neutron-proton scattering. Pulse shape discrimination may be used to select these proton scintillations while rejecting others, especially those caused by gamma rays. Liquid and crystalline organic scintillators are therefore very effective in detecting and identifying neutrons, even when the background of gamma radiation is high. The response of organic scintillators to electrons is linear, that is, proportional to electron energy. However, their response to heavier particles is nonlinear; and for different particles of the same energy, the response is smaller, the heavier the particle.

Additions of compounds and samples. Appreciable quantities of compounds of a number of elements can be incorporated in liquid scintillators (in appropriate chemical form) without degrading the scintillation performance excessively. Some examples are lithium, boron, gadolinium, or cadmium compounds, which facilitate the detection of low-energy neutrons (less than about 0.1 MeV) through neutron capture in these elements. Organic liquid scintillators are extensively used in the medical, biological, and environmental fields to detect radioactive samples and tracer compounds labeled with the beta emitters ^{14}C or ^{3}H, or with other radionuclides. A carrier solution containing the sample or tracer compound is assimilated or dissolved in the liquid scintillator to form a uniform counting sample or scintillating medium with a volume of a few milliliters. Such counting samples are often processed in large batches by using sophisticated instrumentation which incorporates automatic sample changing and the capacity to carry out all the necessary data reduction automatically. *See* RADIOACTIVE TRACER.

Manufacture and handling. The handling of liquid scintillators requires extreme care, because many are highly inflammable and because small concentrations of common impurities can drastically impair the scintillation performance by quenching the energy-transfer mechanism of the scintillation process. Commercial manufacturers offer a wide variety of crystal, glass, liquid, and plastic scintillators, which may also be tailored to specific requirements. Liquid scintillators can be packaged in a variety of forms, for example, bottled in bulk form or encapsulated in ready-to-use glass or metal-and-glass containers. NaI(Tl) crystals must be hermetically isolated from the atmosphere because they are hygroscopic and deteriorate rapidly if exposed to moisture. They are often supplied optically coupled to a photomultiplier, in a sealed canister.

Applications. Fundamental discoveries in physics and astrophysics have been made with scintillators, and scintillation counters have found application in fields as diverse as nuclear physics, elementary particle physics, atomic physics, solid-state studies, chemistry, nutrition studies, biology, nuclear medicine, astrophysics and space physics, geology, nuclear reaction technology and monitoring, petroleum and mineral resources exploration, and oceanography. *See* PARTICLE DETECTOR.

Frank D. Brooks; Robert Hofstadter

Bibliography. J. B. Birks, *The Theory and Practice of Scintillation Counting*, 2d ed., 1967; F. D. Brooks, Development of organic scintillators, *Nucl. Instrum. Meth.*, 162:477–505, 1979; R. L. Heath, R. Hofstadter, and E. B. Hughes, Inorganic scintillators: A review of techniques and applications, *Nucl. Instrum. Meth.*, 162:431–476, 1979.

Scleractinia

An order of the subclass Zoantharia which comprises the true or stony corals (**Fig. 1***a*). These are solitary or colonial anthozoans which attach to a firm substrate. They are profuse in tropical and subtropical waters and contribute to the formation of coral reefs or islands. Some species are free and unattached.

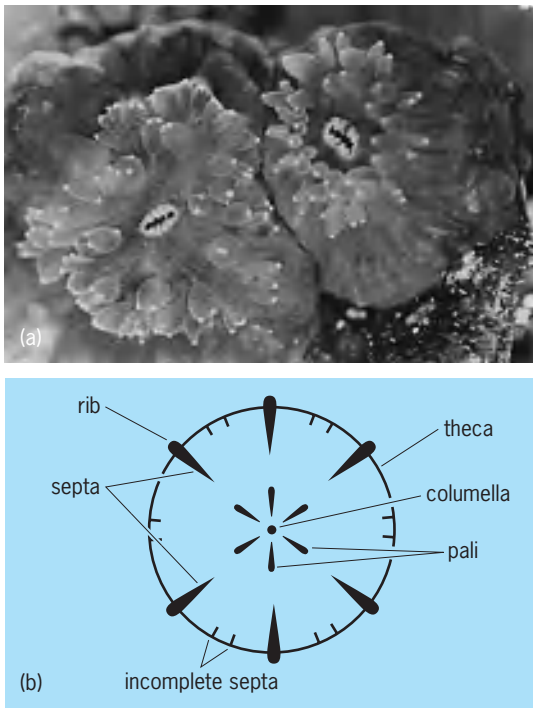

Fig. 1. Solitary corals. (*a*) Coral polyps, *Oulangia* sp. (*b*) Diagrammatic figure, cross section.

Fig. 2. Skeletons of aggregated colonies of *Pocillopora damicornis caespitosa* formed 15 days after fixing.

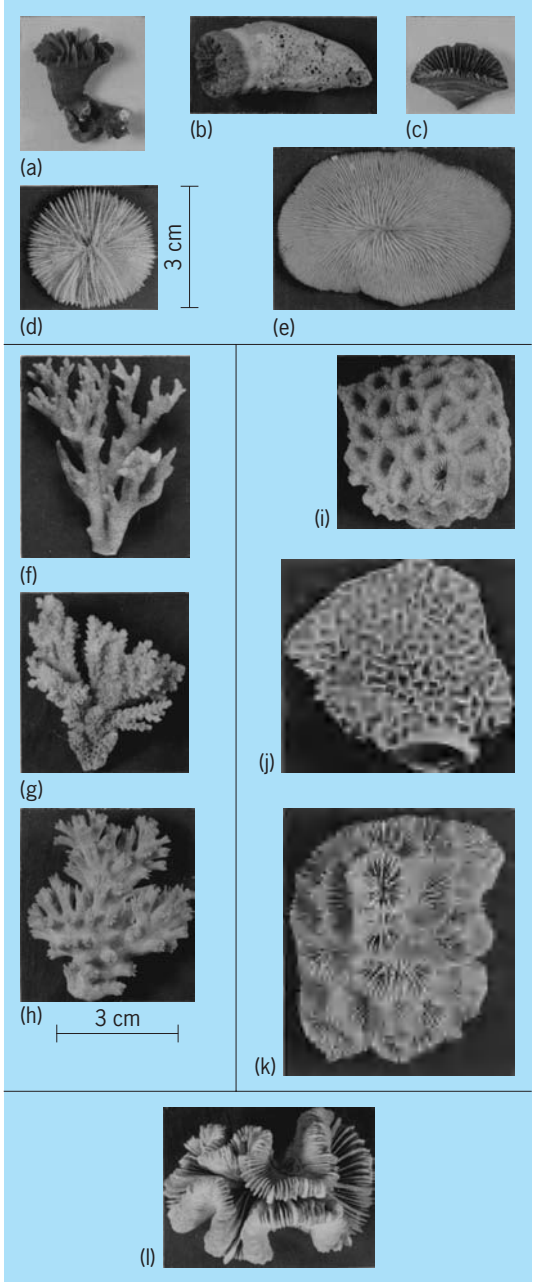

Fig. 3. Coral skeletons. (*a*) *Desmophyllum dianthus*. (*b*) *Balanophyllia gigas*. (*c*) *Flabellum distinctum*. (*d*) *Fungia actiniformia palauensis*. (*e*) *Fungia scutaria*. (*f*) *Pocillopora acuta*. (*g*) *Acropora hyacinthus*. (*h*) *Acrhelia horrescens*. (*i*) *Favia pallida*. (*j*) *Platygyra lamellina*. (*k*) *Acanthastrea echinata*. (*l*) *Trachyphyllia geoffroyi*.

Most of the polyp is impregnated with a hard calcareous skeleton secreted from ectodermal calcioblasts. As soon as the planula settles, it secretes a thin skeletal basal plate, from which radiating vertical platelike partitions or septa arise; then characteristic thecal formation follows. The basal parts of the inner septal edges are often fused to form a columella (**Fig. 2**), which is sometimes ridged by vertical thin plates or pali. The outer septal edges, projecting

beyond the theca, are the costae or ribs. Such is the outline of the corallum. Various other structures develop, such as the endo- or exothecae which comprise the dissepiments. They are inside and outside the wall of the corallum, respectively (Fig. 1*b*).

Skeleton. The solitary corals form cylindrical, discoidal, or cuneiform skeletons (**Fig. 3***a–e*), whereas colonial skeletons are multifarious. *Stylophora*, *Seriatopora*, *Pocillopora*, and *Acropora* form branching skeletons (Fig. 3*f–h*), while the skeletal systems of *Favia*, *Favites*, *Porites*, *Coeloseris*, and *Goniastrea* are massive (Fig. 3*i–k*). In others, such as *Echinopora*, *Plerogyra*, *Lobophyllia*, and *Trachyphyllia* (Fig. 3*l*), all sorts of shapes are found.

Colony formation. The polyps increase rapidly by intra- or extratentacular budding, and the skeletons of polyps which settle in groups may fuse to form a colony (Fig. 2). The pyriform, ciliated planula swims with its aboral extremity, which is composed of an ectodermal sensory layer, directed anteriorly. Planulation occurs periodically in conformity with lunar phases in many tropical species. *See* ZOANTHARIA.

Kenjai Atoda

Fossils. Scleractinian corals possess robust skeletons, so they have a rich fossil record. Because they are restricted mainly to tropical belts, they help indicate the position of the continents throughtout the Mesozoic and Cenozoic periods. They are also important for understanding the evolution of corals and the origin and maintenance of reef diversity through time. Pleistocene corals shows persistent reef coral communities throughout the last several hundred thousand years. Environmental degradation has led to the dramatic alteration of living coral communities during the past several decades.

Geologic record. The first Scleractinia appeared during the Middle Triassic (**Fig. 4**). Corals may have had multiple origins at the same time that they evolved skeletons, or they may have evolved much earlier (during the Paleozoic) and multiple groups acquired skeletons at the same time in the Triassic. Recent molecular work points to the evolution of the Scleractinia well before they first appeared some 240 million years before present. If true, soft-bodied Scleractinia existed in Paleozoic seas with the skeletonized rugose and tabulate corals. The Permian extinction event wiped away the rugose and tabulate corals, paving the way for scleractinians to evolve skeletons.

Scleractinians played a prominent role in Triassic reefs, where they were highly zoned. The end-Triassic extinction event led to the collapse of the first Mesozoic reef ecosystem with its new scleractinian components. The middle Oxfordian (Late Jurassic) marked an important phase of reef expansion. Reefs became so widespread that large and extensive reef belts of the size of the modern Australian Great Barrier Reef were developed in western Europe and the Caucasus belt. Late Jurassic reefs were zoned into fore-reef and back-reef facies, and atolls were common. These reefs are now preserved in geological sections up to several hundred meters thick. The all-time global maximum of coral diversity occurred in the Late Jurassic with over 200 genera recorded.

The Cretaceous continued the Upper Jurassic pattern, but in the Barremian (Early Cretaceous) coral-mimicking rudist bivalves began to compete with and replace the scleractinians as the dominant reef biota. Corals generally occupied separate habitats at greater depths than earlier. Corals returned to dominance in the Late Cretaceous, following extinction of the rudistids. By the close of the Cretaceous, extensive coral reefs occurred worldwide. Few coral reefs are recorded from the Paleocene. But from the Eocene to Recent, scleractinian coral-dominated reef tracts occur in tropical belts.

Environmental perturbation. Corals and reefs are subject to a number of stresses: seasonal temperatures above 30°C (86°F) or below 20°C (68°F); sharp seasonal temperature fluctuations (for example, El Niño); high input of land-derived muds, clays, ash, and such; low salinity or hypersalinity; low light levels caused by suspended detritus; high nutrient input; rapid sea-level change; excessive tidal range and storm events; changes in local shelf topography; and biological agents such as disease, plankton blooms, and outbreak predator populations (such as crown-of-thorns starfish). Triassic coral reefs were disrupted by the loss of shallow-water habitat, and climate change at the end of the Cretaceous appears to have been the driving force in the gradual extinction of corals. Cenozoic reefs have been most affected by ocean current and temperature fluctuations induced by climatic change. Successive Plio-Pleistocene glaciations of the past 2–3 million years and the accompanying multiple sea-level fluctuations suggest that the Cenozoic was a relatively turbulent setting for reef development.

Distribution. The earliest scleractinians were not reef-dwelling but occupied the ancient Tethys Sea near the present-day Bavarian, Austrian, and Italian Alps, and Indo-China. Middle and Late Triassic corals were widespread throughout the Tethys region and around most of the equatorial Panthalassa Ocean rim (including India, Australia, New Guinea, and western North and South America) [**Fig. 5**]. In the Middle Jurassic, reef development proliferated in present-day Europe and the Mediterranean but remained poorly developed in Panthalassa. Late Jurassic reefs extended from Greenland and Sakhalin in the north to Chile and Argentina in the south. By this time, there were well-established coral provinces. An Asiatic province extended along the northern margin of Tethys, with subprovinces along the east Asian coast; a southern Tethys province extended to Panthalassa South America; and a northern Tethys province extended to Panthalassa North America. The paleobiogeographic pattern that developed was the precursor to the Tethys/Atlantic/eastern Pacific pattern of the Paleogene. During the Miocene, reef development was at a global maximum for the Cenozoic. The continents were close to their present position, and reef growth occurred throughout the tropical seas. Pleistocene and modern reef

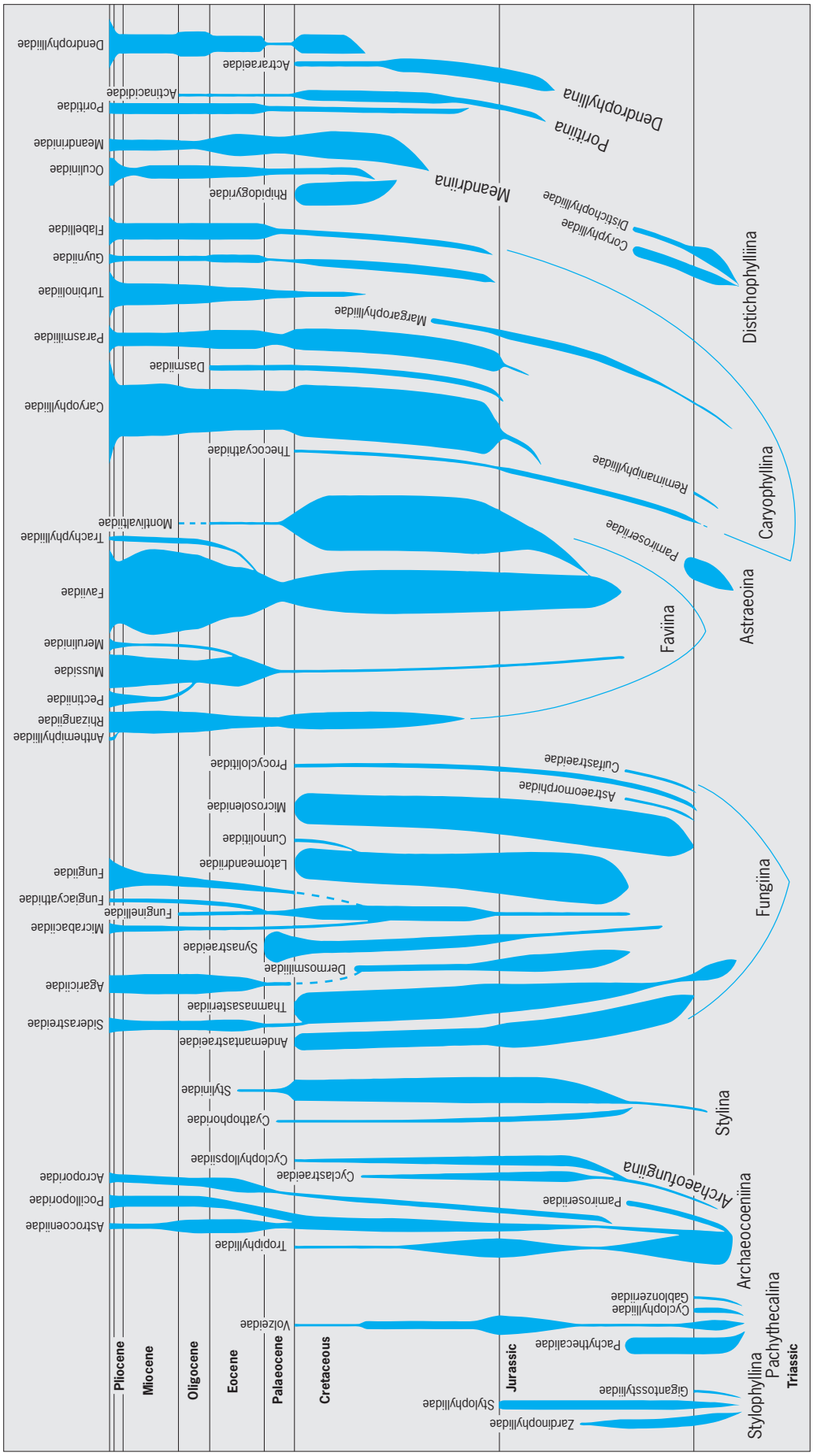

Fig. 4. Family-level phylogeny of the Scleractinia through geological time. Branch widths indicate numbers of genera per family for each geological interval. (*From J. E. N. Veron, Corals in Space and Time: The Biogeography and Evolution of the Scleractinia; copyright © 1995 J. E. N. Veron and the Australian Institute of Marine Science; used by permission of the publisher, Cornell University Press*)

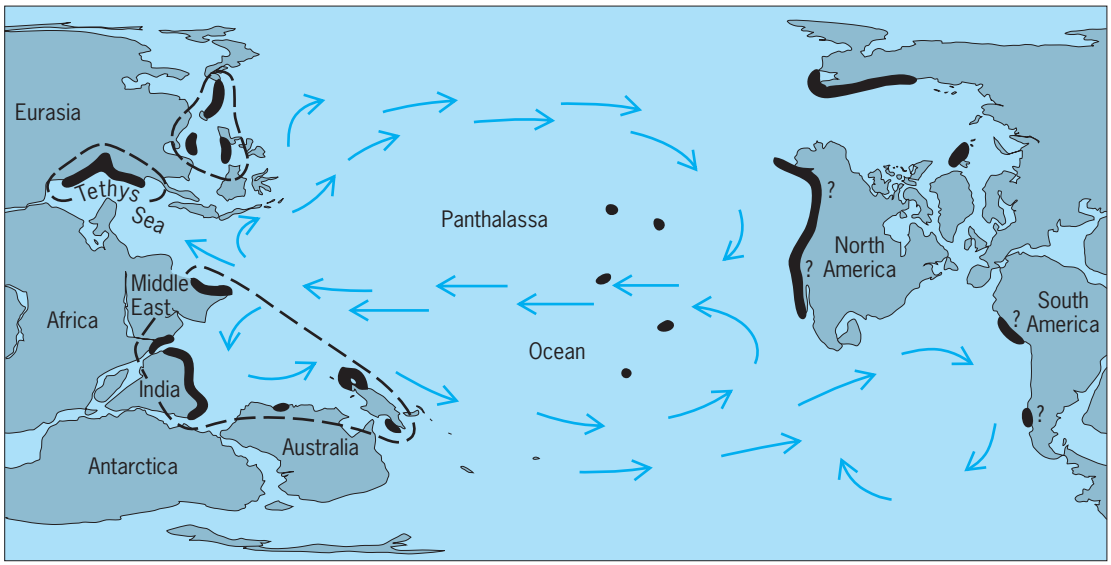

Fig. 5. Late Triassic distribution of corals and coral reefs. Broken lines encompass the major coral provinces. (*From J. E. N. Veron, Corals in Space and Time: The Biogeography and Evolution of the Scleractinia; copyright © 1995 by J. E. N. Veron and the Australian Institute of Marine Science; used with permission of the publisher, Cornell University Press*)

development is best developed between 30°N and 30°S latitude. John M. Pandolfi

Bibliography. P. Copper, Enigmas in Phanerozoic reef development, *Mem. Ass. Australasian Palaeontol.*, 8:371–385, 1989; J. A. Fagerstrom, *The Evolution of Reef Communities*, Wiley, New York, 1987; S. L. Romano and S. R. Palumbi, Evolution of scleractinian corals inferred from molecular systematics, *Science*, 271:640–642, 1996; J. E. N. Veron, *Corals in Space and Time: The Biogeography and Evolution of the Scleractinia*, University of New South Wales Press, Sydney, 1995.

Sclerenchyma

Single cells or aggregates of cells whose principal function is thought to be mechanical support of plants or plant parts. Sclerenchyma cells have thick secondary walls and may or may not remain alive when mature. They vary greatly in form and are of widespread occurrence in vascular plants. Two general types, sclereids and fibers, are widely recognized, but since these intergrade, the distinction is sometimes arbitrary.

Sclereids. These range from isodiametric to much elongated cells and may be branched. They may occur as isolated cells (idioblasts), groups of cells, or extensive tissues. Five major kinds of sclereids are described here: brachysclereids, macrosclereids, osteosclereids, astrosclereids, and trichosclereids.

Brachysclereids (stone cells) have the form of the parenchyma cells from which they are derived by secondary sclerosis (hardening). They have conspicuous, frequently ramiform (branched) pits in their thickened walls and are found in the shells of nuts, the pits of stone fruits, the bark of many trees, and the xylem and pitch of some plants. Nests of brachysclereids form the grit in the flesh of pears and quinces (**illus.** *a*).

Macrosclereids are rodlike cells formed from the protoderm (embryonic epidermis) of certain seed coats (illus. *b*). Their secondary walls are unevenly thickened.

Osteosclereids, or "bone" cells, are rodlike with swollen ends. They occur in seed coats and in some leaves.

Astrosclereids tend to be radiately branched but are otherwise quite variable (illus. *c*). They occur in the leaves of many plants (*Trochodendron*, *Pseudotsuga*, and *Mouriria*) and in the petioles of *Camellia*. Astrosclereids sometimes have crystals embedded in their walls (*Castalia*).

Trichosclereids (illus. *d*) are long and slender, resembling fibers, with which they intergrade (leaves of *Monstera* and olive).

In general, foliar sclereids are idioblastic. They often arise from initials in the ground meristem. Branched forms extend themselves into intercellular spaces and even between cells. Sometimes they are associated with the ends of veinlets.

Fibers. Typically, fibers have tapering ends and are very long in proportion to their width. They usually occur as coherent strands of tissue and are rarely idioblastic. Individual fiber cells vary in length from a few millimeters to more than half a meter (*Boehmeria nivea*). Their thick walls are often sparsely pitted, and the pits usually appear simple. Fibers are widely distributed both in the primary and in the secondary body of vascular plants. Two broad types are generally recognized: intraxylary, or wood, fibers occur in the xylem (illus. *e*); extraxylary, or bast, fibers are found in the phloem and cortex of dicotyledons (illus. *f*) and in association with vascular bundles in the stems and leaves of monocotyledons.

Libriform (elongated, thick-walled) wood fibers (illus. *f*), which may constitute as much as 50% of an angiosperm wood, are so called because of their structural resemblance to phloem fibers. They often intergrade with tracheary (water-conducting)

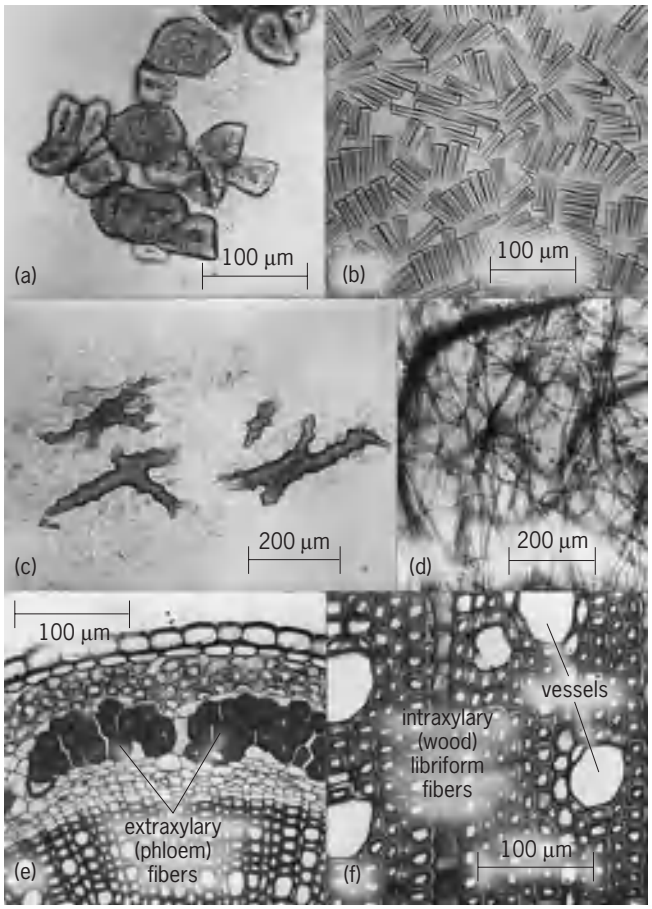

Sclerenchyma. (*a*) Brachysclereids (stone cells) from fruit of pear. (*b*) Macrosclereids (rod cells) from macerated seed coat of bean. (*c*) Branched sclereids from petiole of *Camellia*. (*d*) Fiberlike sclereids in cleared leaf of olive. (*e*) Transection of a portion of stem of *Linum* showing the extraxylary (phloem) fibers. (*f*) Transection of wood of *Cornus* showing the intraxylary (wood) libriform fibers and vessels.

elements. Intermediate forms are called fiber tracheids. Septate wood fibers and fiber tracheids have transverse partitions which develop after the secondary wall is laid down (*Vitis* and *Hypericum*). Gelatinous or mucilaginous fibers have hygroscopic cell walls. They are found in many angiosperm woods such as oak (*Quercus*) and black locust (*Robinia*).

The fibers of commerce, excepting cotton and kapok (which consist of unicellular hairs), are bast fibers. The leaf fibers from monocotyledons, such as sisal (*Agave sisalina*), henequen (*A. fourcroydes*), and abaca (*Musa textilis*), are known as hard fibers. They are lignified and coarse and are used in the manufacture of cordage. The soft, or stem, fibers include jute (*Corchorus*), true hemp (*Cannabis*), ramie (*B. nivea*), and flax (*Linum usitatissimum*). Flax fibers, the raw material for linen, contain little or no lignin. *See* NATURAL FIBER; PHLOEM; PLANT ANATOMY; XYLEM.

Norman H. Boke

Sclerosponge

A class of sponges that lay down a compound skeleton comprising an external, basal mass of calcium carbonate, either aragonite or calcite, and internal siliceous spicules and protein fibers. The living tissue forms a thin layer over the basal calcareous skeleton and extends into its surface depressions; the organization of the tissue is similar to that of encrusting demosponges. Sclerosponges are common inhabitants of cryptic habitats on coral reefs in both the Caribbean and Indo-Pacific biogeographic regions. Their discovery has reopened questions of the affinities of certain fossil reef–inhabiting organisms, the stromatoporoids and the chaetetid and favositid tabulate "corals," all of which have been generally regarded as members of the phylum Cnidaria. *See* DEMOSPONGIAE.

Astrosclera. This sponge, first described in 1900 by J. Lister, is notable for its reticulate spherulitic skeleton of aragonite, in the interstices of which the living tissues occur. Siliceous spicules in the form of acanthostyles are also present, and some of these may be incorporated into the calcareous moiety of the skeleton.

Stromatoporoid affinities. *Astrosclera* resembles the stromatoporoids (1) in general aspect as a calcareous, mound-shaped organism varying in diameter from 0.4 to 6 in. (1 to 15 cm), (2) in the reticulate pattern of its fibrous aragonitic skeleton, (3) in the occurrence of astrorhizae (**Fig. 1**) at the surface of the skeleton, (4) in the absence of regularly repeated zooidal units in the skeleton, and (5) in the presence of an epitheca showing concentric growth lines on the lower side of the animal. In postulating a relationship between *Astrosclera* and stromatoporoids it must be remembered that a single relict species is being compared with a host of organisms comprising about 2000 described species that have existed for 4×10^8 years, during which time they have radiated into many divergent evolutionary lines. *Astrosclera* is an end point of one of these lines.

The reticulate skeleton of *Astrosclera* resembles the skeletons of some Paleozoic and most Mesozoic stromatoporoids in general aspect except that it lacks tabulae. In large Pacific specimens of *Astrosclera*, 6 in. (15 cm) across and 3 in. (8 cm) high, living tissue extends to a depth of only about 0.2 in. (5 mm) into the interstices of the skeleton; the basal skeletal mass is filled in by secondary deposits of aragonite. Instead of the intermittent secretion of tabular partitions to wall off the interior of the skeleton, in *Astrosclera* a continuous secretion of aragonite occurs internal to the living tissue.

In Mesozoic stromatoporoids the skeletal microstructure is fibrous in nature, with crystalline units oriented at right angles or obliquely to linear centers of calcification. A fibrous microstructure is uncommon among Paleozoic stromatoporoids, in which compact and cellular types predominate. These compact and cellular microstructures may result from diagenetic processes, however, as may the occurrence of calcite in fossil forms. A fibrous aragonitic skeleton is probably the primary condition in stromatoporoids. In the spherulitic microstructure of *Astrosclera* the aragonitic crystalline units are oriented around a point (**Fig. 2**), whereas in

stromatoporoids they are formed around linear centers of calcification.

The most striking similarity between *Astrosclera* and stromatoporoids is the presence of astrorhizae on the surface of the skeletal mass. The astrorhizae of *Astrosclera* and Mesozoic stromatoporoids are comparable in minute detail. In both, into the base of the surface grooves that shelter exhalant channels leading to a central oscule, there open oblique skeletal tubes carrying smaller exhalant channels that join the main surficial ones. The walls of these tubes look like tabulae lying across the base of the grooves. Vertical skeletal tubes join the central oscular region from below, bringing exhalant channels to the surface from deeper areas. Vertical series of astrorhizae, each lying on a previous growing surface and joined by vertical tabulate tubes, characterize many Paleozoic stromatoporoids with seasonal growth patterns; such features are absent from the tropical, continuously growing *Astrosclera*.

The presence of siliceous spicules in *Astrosclera* finds no certain counterpart among stromatoporoids. N. D. Newell reported spicules in a Pennsylvanian stromatoporoid that he named *Parallelopora mira*, but C. W. Stearn doubts Newell's placement of this form. W. D. Hartman and T. F. Goreau have pointed out questionable pseudomorphs of spicules in Mesozoic stromatoporoids. Since northern Pacific populations of *Astrosclera* lack siliceous spicules, the thrust of this argument against a stromatoporoid affinity for this sponge loses some of its force.

Astrosclera may be regarded as a living descendant of the stromatoporoids, and this group may be classified as an order of the class Sclerospongiae. Stearn, however, prefers to consider the stromatoporoids as a subphylum of the phylum Porifera bearing no direct relationship with sclerosponges. Some authors, such as J. St. Jean, Jr., classify the late Paleozoic and Mesozoic stromatoporoids in a distinct order, the Sphaeractinoidea. The high probability of the occurrence of a similar fibrous aragonitic skeleton in "true" stromatoporoids as well as in "sphaeractinoid" forms largely obviates the necessity for a separation of the time series into different orders, however.

Ceratoporella. This sponge, first dredged off Cuba in 1878, was rediscovered by Goreau on Jamaican coral reefs in 1965. *Ceratoporella* secretes a massive skeleton of aragonite that is covered in life by a thin veneer of living tissue. The surface of the calcareous skeleton is marked by closely spaced calicles (**Fig. 3**) that share common walls and vary from 0.008 to 0.02 in. (0.2 to 0.5 mm) in lumen diameter. Each calicle continues to the base of the skeletal mass but is filled in secondarily by aragonite as the sponge grows upward so that only the distal 0.04 to 0.05 in. (1.0 to 1.2 mm) is filled with living tissue. The surface of the skeleton is marked by stellate impressions that indicate the position of the radiating exhalant canal systems. In addition to possessing a calcareous skeleton, *Ceratoporella* secretes siliceous spicules in the form of acanthostyles. These become embedded in the calcareous skeletal mass as it grows upward, and there they begin to erode away.

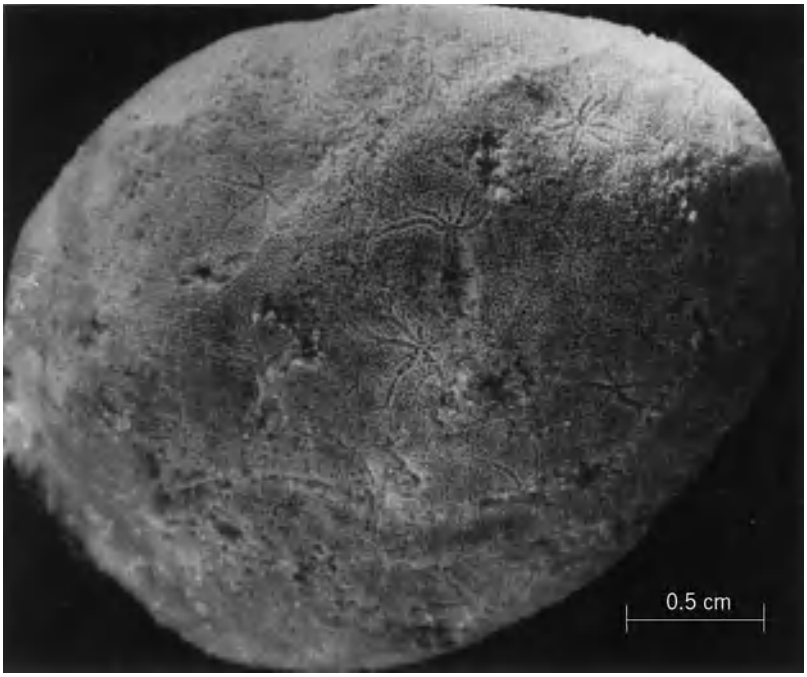

Fig. 1. Skeleton of *Astrosclera*, with astrorhizal patterns on surface.

Three genera related to *Ceratoporella* also occur on Caribbean reefs. In these the surface of the basal calcareous skeleton is marked by branched or lamellate upright processes, and the living tissues grow down into the irregularly shaped depressions between them. A species of one of these genera, *Stromatospongia*, is found on reefs of western Pacific islands.

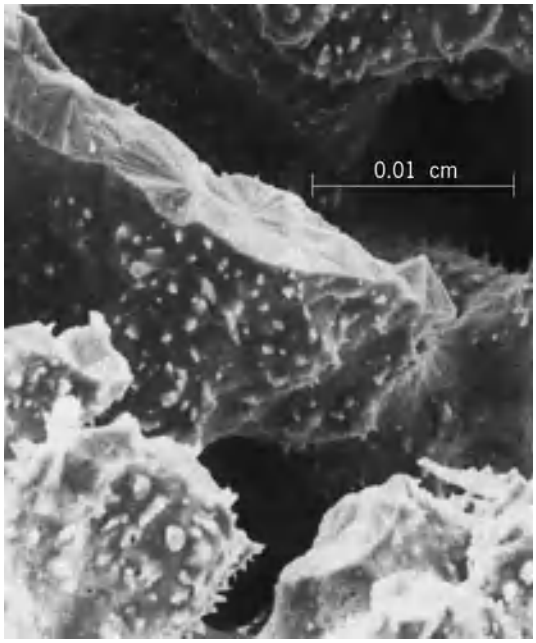

Fig. 2. Skeletal reticulum of *Astrosclera* showing spherulitic arrangement of crystalline units.

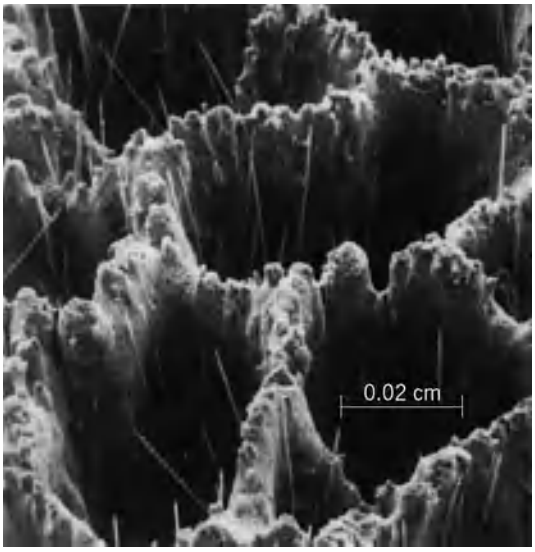

Fig. 3. Surface calicles of *Ceratoporella* showing protruding siliceous spicules. (*From W. D. Hartman and T. F. Goreau, Trans. Conn. Acad. Arts Sci., 44:131–148, 1972*)

Chaetetid affinities. The calcareous skeleton of *Ceratoporella* is remarkably similar to that of the group of fossil coralline animals known as chaetetids, ranging from Ordovician to Cretaceous times. Chaetetids, like *Ceratoporella*, have a calcareous skeleton composed of contiguous, aseptate, cylindrical tubes with shared walls between adjacent units. Another striking resemblance between the two groups is the mode of asexual reproduction by longitudinal fission, accomplished by the ingrowth of longitudinal "pseudosepta" to form a pair of daughter calicles. Further, the microstructure of the walls of the calicles in both is trabecular in nature, and the internal diameter of the calicles of *Ceratoporella* lies within the range of those found among chaetetids. In both groups the regular arrangement of contiguous longitudinal calicles breaks down in some species to a meandroid or irregular pattern.

Differences between *Ceratoporella* and its relatives on the one hand and chaetetids on the other include the incorporation of siliceous spicules into the calcareous skeleton of the former. No comparable occurrence is known among the chaetetids. Moreover, in *Ceratoporella*, the base of the skeleton is separated from the peripheral growing part by a continuous secretion of calcium carbonate rather than a periodic laying down of tabulae, as is characteristic of chaetetids. Since the characters shared by the two groups outweigh the differences, Hartman and Goreau have proposed placing *Ceratoporella* and its relatives in an order Ceratoporellida along with the order Chaetetida as related groups in the class Sclerospongiae. Fossil species of ceratoporellids are known from Triassic strata.

It seems likely that the rapid expansion of the scleractinian corals on Mesozoic and Cenozoic reefs brought the stromatoporoids and chaetetids into eclipse, with the surviving members relegated to cryptic reef habitats except at depths below those inhabited by hermatypic corals.

Acanthochaetetes wellsi. Unlike other sclerosponges described earlier, this Pacific sponge has a basal skeleton of calcite composed of contiguous vertical tabulate calicles ornamented within by vertical rows or irregular clumps of spines (**Fig. 4**). Living tissue extends to depths of 0.05–0.08 in. (1.2–2.0 mm) in the calicles and to the level of the outermost tabula in each. Ostia, located in the floor of shallow, curved grooves, open into the units of sponge tissue enclosed in the calicles; oscules occur at the center of slightly raised mounds and are converged upon by stellate configurations of exhalant channels that are impressed in the calcitic skeleton below as astrorhizae. Siliceous spicules of two types, tylostyles and modified spirasters, occur in the living tissue but are seldom incorporated into the calcitic skeleton.

Extinct species of *Acanthochaetetes* occur in Jurassic and Cretaceous strata. Hartman and Goreau have placed the family Acanthochaetetidae in a new order of sclerosponges, the Tabulospongida.

Favositid affinities. Certain characteristics of the calcareous skeleton of *Acanthochaetetes*, such as the presence of spines on the walls of the tabulate calicles and the method of asexual reproduction by means of budding walls between calicles, are strikingly similar to characteristics of the Favositida, an

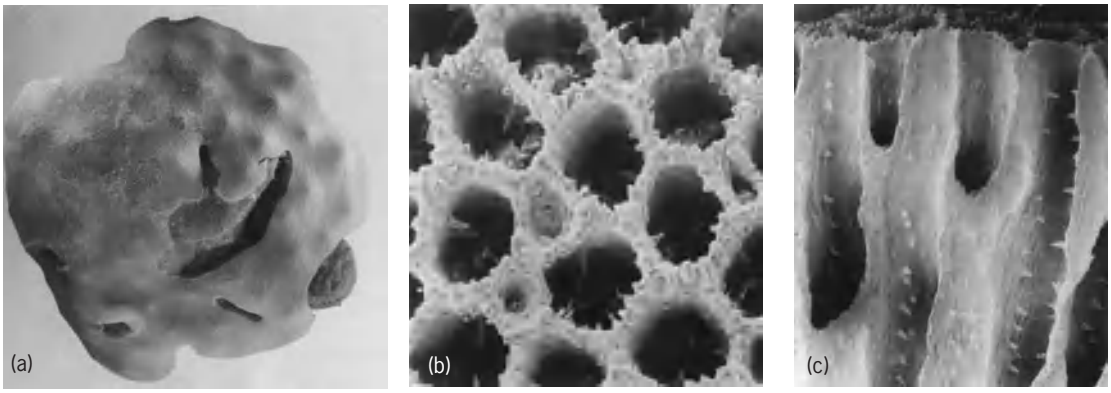

Fig. 4. *Acanthochaetetes wellsi*. (*a*) Skeleton with astrorhizal patterns on surface. (*b*) Calicles viewed from above; two developing asexual buds are seen in lower left. (*c*) Longitudinal fracture of calcitic skeleton showing rows of spines in calicles.

order of Paleozoic tabulate corals. However, the calicles of species of the favositids tend to bear vertical rows of pores that permit physical connections between the living tissue of adjacent calicles, while only submicroscopic pores, one-hundredth the size of those of favositids, occur in *Acanthochaetetes*. Further, whether the calcium carbonate of the calicle walls of favositids was aragonitic or calcitic is still disputed by paleobiologists, and astrorhizal patterns on the surface of favositid skeletons are unknown. Thus, despite some remarkable similarities between the Favositida and the Tabulospongida, the postulate that there was a phylogenetic relationship between the two groups, and its corollary that the Favositida were sponges instead of cnidarians, remains unresolved. *See* SPHAERACTINOIDEA; STROMATOPOROIDEA; TABULATA.

Willard D. Hartman

Bibliography. W. D. Hartman and T. F. Goreau, *Postilla*, 167:1–21, 1975; W. D. Hartman and T. F. Goreau, *Symp. Zool. Soc. London*, 25:205–243, 1970; W. D. Hartman and T. F. Goreau, *Trans. Conn. Acad. Arts Sci.*, 44:131–148, 1972; C. W. Stearn, *Lethaia*, 5:269–288, 1972.

Scorpaeniformes

Fishes of the order Scorpaeniformes, called the mail-cheeked fishes, are distinguished by a unique character, the suborbital stay. The stay is an extension of the hypertrophied third infraorbital bone, which crosses the cheek obliquely and usually is firmly attached to the preoperculum. The head is usually spiny, and the body may be spiny or have bony plates, or both, or be smooth and appear naked. Almost all scorpaenids have prominent pectoral fins, either broad and fanlike or winglike, with the lower rays deeply incised. The order consists of about 1271 species and 266 genera in 25 families.

Order Scorpaeniformes
- Suborder Dactylopteroidei (flying gurnards)
- Scorpaenoidei (scorpionfishes)
- Platycephaloidei (searobins and flatheads)
- Anoplopomatoidei (sablefishes)
- Hexagrammoidei (greenlings)
- Normanichthyiodei (no common name)
- Cottoidei
  - Superfamily Cottoidea (sculpins and poachers)
  - Cyclopteroidea (lumpfishes and snailfishes)

Dactylopteroidei. This suborder consists of one family and seven species. The flying gurnards are tropical fishes of the Indo-Pacific and Atlantic oceans. They resemble searobins in having large, colorful pectoral fins with the inner rays free. "Flying gurnards" is a misnomer; rarely, if ever, do these fishes leave the ocean floor where they "walk" by alternately moving their pelvic fins. Large pectoral fins notwithstanding, there is no evidence that flying gurnards ever break the surface of the water and glide. However, *gurnard*, meaning "grunting," is appropriate, because these fishes produce stridulations by using the hyomandibular bone.

Scorpaenoidei. Thirty-five percent of all scorpaeniform fishes are in the suborder Scorpaenoidei. Of its six families, the principal one is Scorpaenidae (scorpionfishes), which accounts for 80% of the 444 species in the suborder. These fishes are represented in all of the tropical and temperate seas of the world and are the most generalized scorpaeniforms.

Scorpaenidae. The family Scorpaenidae contains the world's most venomous fishes. Scorpaenids are identified by a compressed body; a head usually with conspicuous ridges and spines; an opercle and preopercle armed with spines; scales that, when present, are usually ctenoid; a usually single dorsal fin; well-developed pectoral fins, fanlike and rounded, rarely one to three free rays; venomous glands at the base of the dorsal, anal, and pelvic spines; internal fertilization, with some species oviparous and others ovoviviparous. Most scorpionfishes are bottom dwellers that sit motionless among coral, rock, or grass, and, depending on their color patterns for concealment, wait for a venturesome prey to come within striking distance.

Among the venomous fishes is the beautiful red lionfish (*Pterois volitans*). Its sting causes a number of debilitating symptoms in humans which could lead to death if not properly and promptly treated. The lionfish is an Indo-Pacific native but now is known in the western North Atlantic from Florida to New York where it has proliferated, undoubtedly from released aquarium specimens. A popular subject for underwater photographers, the lionfish should be approached with caution, as it is quite aggressive when threatened. However, far more dangerous are the stonefishes (for example, *Synanceia*). They lie quietly while partially buried in the substrate and are so well camouflaged that they appear to be a rock or piece of coral (see **illustration**). Their spines are like hypodermic needles, ready to deliver the most venomous neurotoxin (ichthyocanthotoxin = toxin from fish spines) known in the fish kingdom. Punctures by these fishes bring on agonizing pain and could lead to a quick death.

Platycephaloidei. These fishes have an elongate body and large head, usually with spines and ridges. The head varies from a steep anterior to a strongly depressed profile. The dorsal fin is in two separate parts, the first with 6 to 12 stiff spines well separated from the soft rays. The lower two or three pectoral rays are enlarged and free in many species. All species are benthic, with some inshore and others occupying depths to 1500 m [5000 ft] (*Hoplichthys*, for example). Species in the families Triglidae and Peristediidae occupy all tropical and temperate seas; the other species are essentially limited to the Indo-Pacific.

Triglidae. The triglids (searobins) are characterized by a head typically encased in bony plates and armed with numerous stout spines, and with no barbel on the lower jaw. The body is scaly, with paired rostral

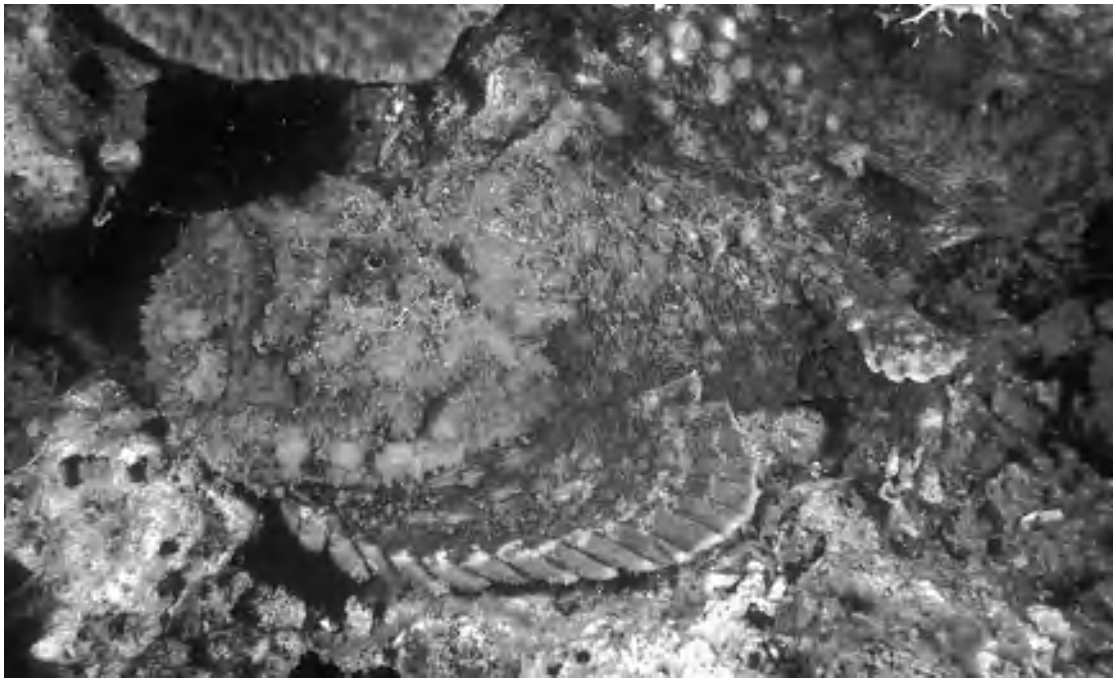

Stonefish (*Synanceia verrucosa*). (*Photo © John E. Randall*)

projections giving the snout a bilobed appearance in dorsal view. The lower three pectoral rays are enlarged, flexible, and free from each other and the rest of the fin, thus appearing like little legs. The free rays function as tactile organs as the triglids creep over the bottom in search of food, which consists mostly of crustaceans and mollusks. Triglids are good sound producers, both in and out of the water. The sound is produced by certain muscles vibrating the air bladder. There are ten genera and 70 species in this family.

Peristediidae. Armored searobins and triglids share some trenchant characters but differ in the former having the entire body encased in spine-bearing plates, two free pectoral rays, large barbels on the lower jaw, and usually conspicuously long and narrow rostral projections. There are four genera and 30 species in this family, all found in deep water.

Anoplopomatoidei. The sablefishes have little affinity with the other scorpaeniforms. The only uniting character, the suborbital stay, may have evolved independently. The head is without spines, ridges, or cirri, and the spinous dorsal fin is well separated from the soft dorsal. There is one family with two genera and two species ranging from Japan, across the Bering Sea, to California.

Hexagrammoidei. Greenlings have cirri on the head but lack ridges and spines. The dorsal fin is single but with a notch between the spinous portion and the soft rays. There are five genera and 11 species, all endemic to the North Pacific.

Normanichthyiodei. The head lacks spines and the body is covered with ctenoid scales. Previously in the suborder Cottoidei, its relationships remain very uncertain. There is only one species in this suborder, which is known to occur off Peru and Chile.

Cottoidei. The cottoidei are the largest group of scorpaeniform fishes, accounting for 49% of the species. The physiognomy of cottoids is quite variable, with overall shapes from elongate to globose. The head is usually large and lacking spines; the body either is naked (or appearing naked) or covered by scales or bony plates. The dorsal fin is either single or double; the pelvic fin has one spine and two to five rays, or may be completely absent. The pectoral fin is either continuous or the lower rays are free as in triglids. Cottoidei is recognized as two monophyletic lineages, the superfamilies Cottoidea and Cyclopteroidea.

Cottoidea. This superfamily consists of nine families. The five families described below account for 96% of the Cottoidea.

Cottidae (sculpins) are the largest family, with 70 genera and 300 species accounting for 75% of Cottoidea species. A lateral process of the hyomandibular bone is a unifying character of the family. Also, the skin appears naked, although the body may be covered with prickles and have a few ctenoid scales in the lateral line. The eyes are large and directed upward. There are two prominent preopercular spines, no spine in the anal fin, one spine and two or three soft rays in the pelvic fin, and large and fanlike pectoral fins. An air bladder is absent. Sculpins are essentially Holarctic in distribution. The marine species reach their greatest diversity in the coastal waters of the North Pacific. Freshwater species occur in North America and northern Eurasia, including a dozen or so species endemic to Lake Baikal. Four genera and 32 described species occur in North America, *Cottus* being the principal genus.

Agonidae (poachers) are mostly quite elongate and resemble armored searobins in having the body

covered with bony plates. Their antitropical distribution is unusual for scorpaeniforms, ranging in the southern hemisphere to the tip of South America and in the northern hemisphere in the Atlantic and Pacific oceans to near the Arctic Circle, with the greatest diversity in the North Pacific. Their vertical range is from shallow inshore waters to depths of over 1000 m (3,300 ft).

Psychrolutidae (fathead sculpins) are robust fishes with a large round head and a usually continuous dorsal fin. Some have well-ossified head bones with spines, while others have poorly ossified head bones, no spines, and a wide interorbital distance. These scorpaenids are found in both hemispheres of the Indo-Pacific and Atlantic oceans, at depths ranging from shallow inshore waters to depths of 2800 m (9,200 ft).

Comephoridae (Baikal oilfishes) and Abyssocottidae are freshwater families with elongate bodies, a large head, large terminal mouth, and a spinous dorsal fin separate from a soft dorsal fun. The two species of Comephoridae, which are endemic to Lake Baikal, lack body scales and prickles, have no pelvic fin, and possess very long pectoral fins. The Abyssocottidae, consisting of 20 species primarily from Lake Baikal, have pelvic fins and relatively short pectorals. The two families collectively contribute about half of the lake's known fish fauna.

Cyclopteroidea. In species of this superfamily the pelvic fins, when present, are thoracic in position and usually modified to form a suction-cup-like disk. The pectoral fins are broad and tucked under the operculum; the gill openings are rather restricted, and the lateral line is usually absent. Two families, Cyclopteridae and Liparidae, are the most bizarre scorpaeniforms.

Cyclopteridae (lumpfishes) possess a globose body and are covered with thick skin, which is usually tubercular but may be smooth or spinous. The dorsal spines are flexible and often hidden beneath skin, and the soft dorsal and anal fins are small and never confluent with the caudal fin. By means of the ventral sucking disc, these fishes can attach themselves to rock, which they may resemble. Seven genera and 29 species occupy the cold marine waters of the northern hemisphere.

Liparidae (snailfishes) are tadpole-shaped fishes with soft, smooth, jellylike skin. The spines of the dorsal fin are weak and flexible; the soft dorsal and anal fins are long and often confluent with the caudal fin. The pelvic fins, when present, form a sucking disk. The 19 genera and about 195 species are represented over the widest geographical range of all scorpaeniforms, from Arctic to Antarctic seas. Their vertical range is from tidal pools to depths of over 7000 m (23,000 ft). *See* ACTINOPTERYGII; OSTEICHTHYES; PERCIFORMES. Herbert Boschung

Bibliography. J. S. Nelson, *Fishes of the World*, 3rd ed, 1994; B. B. Washington, W. N. Eschmeyer, and K. M. Howe, Scorpaeniformes: Relationships, in H. G. Moser et al. (eds.), *Ontogeny and Systematics of Fishes*, pp. 438–447, Spec. Publ. no. 1, American Society of Ichthyology and Herpetology, 1984.

Scorpiones

An order of the Arachnida characterized by chelate pedipalps and chelicerae (pincerlike appendages), pectines (feathery chemo- and mechanoreceptors used to survey the texture of the ground surface and detect pheromones), and a narrow, flexible postabdomen bearing a venomous segment (telson) with a terminal sting (aculeus).

Morphology. The scorpion body is divided into a cephalothorax (prosoma), covered by an unsegmented carapace, and a segmented abdomen (opisthosoma) (**Figs. 1** and **2**). The opisthosoma is differentiated into an anterior preabdomen (mesosoma) and a postabdomen (metasoma), which, together with the telson, constitutes the "tail" or cauda. The cephalothorax bears the chelicerae, pedipalps, and four pairs of walking legs. The preabdomen contains seven segments, the postabdomen five. One pair of small, simple median eyes (ocelli) and, depending on the species, 2–5 (usually 3) pairs of anterolateral eyes are situated on the carapace. Troglobitic (cave-dwelling) scorpions lack all or some eyes (usually the median pair). Each of the seven preabdominal segments is covered by a tergite, the ventral segments by a sternite. The genital aperture, covered by the genital operculum, opens ventrally on the first segment. A small, quadrate basal piece articulates laterally with the pectines on the second segment. The lateral areas of sternites III–VI possess four pairs of oblique, oval slits, the openings (spiracles) of the internal respiratory organs (book lungs).

Classification. Scorpions are derived from amphibious ancestors that lived in the Silurian, more than 400 million years ago. Paleozoic scorpions closely resemble modern descendants in anatomical details, except that some were considerably larger. *Brontoscorpio anglicus* measured approximately 1 m (3.3 ft) in length. Scorpions were formerly considered the sister group of other arachnids because they closely resemble extinct marine eurypterids (the sister group of arachnids) but recent data suggest that scorpions are embedded in the arachnid lineage, and merely retain primitive features.

Approximately 1490 extant species of scorpions, placed in 170 genera, and between 13 and 20 families (depending on the authority), have been described worldwide. They vary in color from translucent through tan or brown to black, or combinations thereof, and in size from 10 mm to 21 cm (0.4–8.3 in.).

Distribution. Scorpions occur on all continents except Antarctica, but are most abundant and diverse in tropical and subtropical regions, especially in desert and semidesert habitats. They also occur in savannas and grasslands, in deciduous, coniferous, and tropical rainforests, on high mountain slopes (above 5500 m, 18,000 ft, elevation) in the Alps, Himalayas, and Andes; in some of the deepest caves (nearly 1000 m, 3300 ft, below the surface), and in the intertidal zone. Although most scorpions are terrestrial, some are arboreal. Favored habitats include burrows (up to 1 m, or 3.3 ft, deep, to escape hot diurnal

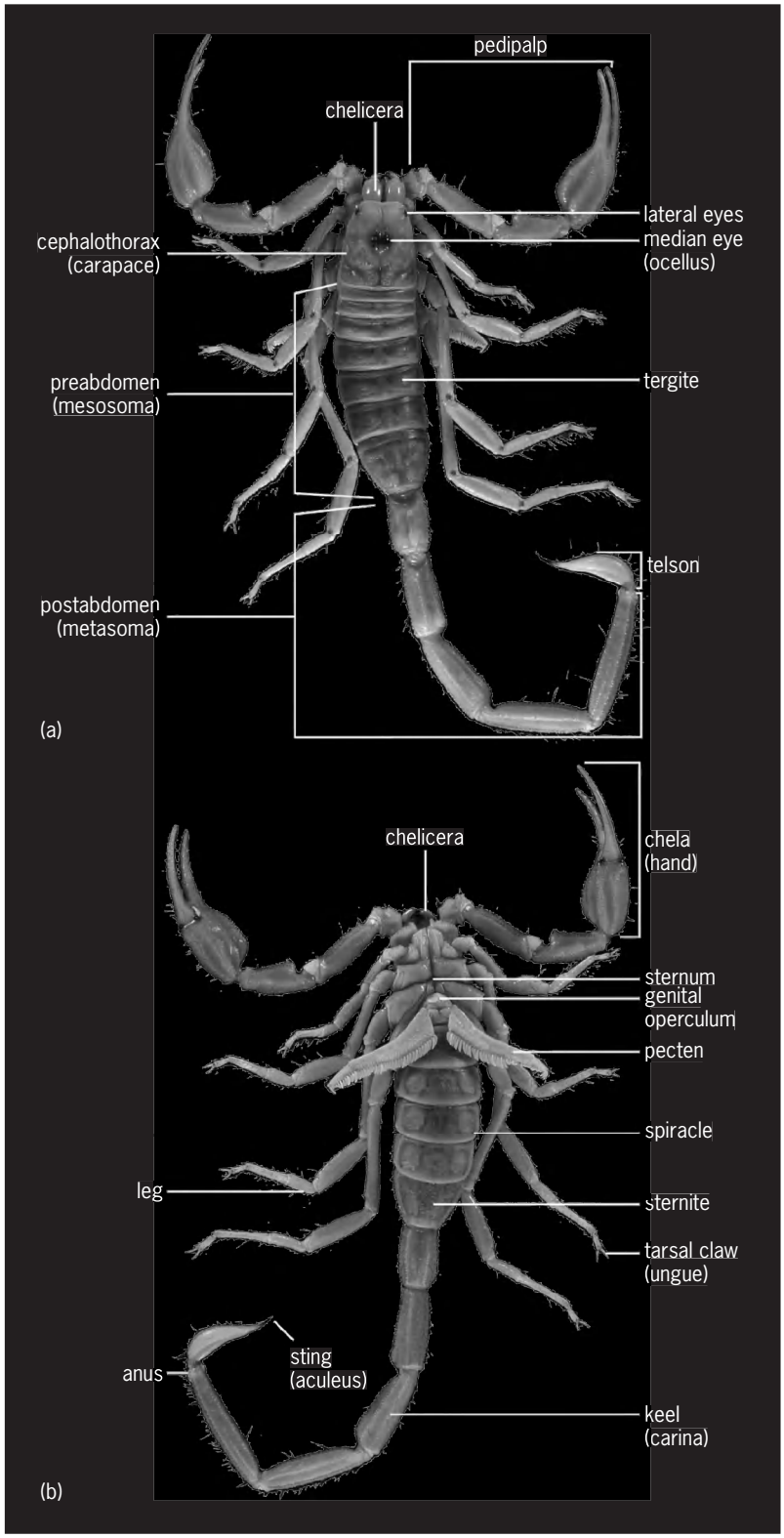

Fig. 1. Scorpion, *Smeringurus mesaensis* (Stahnke, 1957), adult male. (*a*) Dorsal view. (*b*) Ventral view. (*Photos by Randy Mercurio and Steve Thurston; copyright © American Museum of Natural History*)

temperatures), spaces under tree bark, logs, stones, and in rock crevices. Some species adapt well to human environments. Scorpions are often abundant in suitable habitats. Densities of 1 per square meter were reported for the Middle Eastern *Leiurus quinquestriatus* and of 8–12 per square meter for the intertidal *Serradigitus littoralis* from Baja California.

Ecology. Scorpions are primarily nocturnal. All fluoresce under long-wave ultraviolet light, facilitating their collection and observation at night. Their simple eyes detect luminosity, but little else. Prey are detected with slit sense organs in the tarsi, sensory setae (trichobothria) on the pedipalps, and the pectines, and are attacked with the chelate pedipalps or venomous sting. Scorpions with slender pedipalps are prone to sting their prey, those with robust pedipalps to crush prey mechanically, reserving the sting for large or strong prey. All use the pedipalps to manipulate prey, tearing pieces off with the chelicerae to be digested in a preoral cavity before being sucked into the gut.

Scorpions are important consumers in some communities. *Scorpio maurus* was reported to eat an annual average of 11% of the Israeli isopod population; *Urodacus yaschenkoi* consumes 7.9 kg/ha (7.8 lb/acre) of invertebrate prey in Australia. Cannibalism and predation by other scorpion species may be the most important sources of scorpion mortality, but other invertebrate predators (such as centipedes) and vertebrates are also important predators. Mortality is highest immediately after birth, lower for individuals of intermediate age, and high for adults (65%, 30%, and 60% per year, respectively, for the Australian *Urodacus manicatus*). Mortality is particularly high among males due to increased mobility during the breeding season and cannibalism by females. Skewed adult sex ratios of 1.2–1.4:1 are typical. Social behavior occurs rarely in species of *Heterometrus*, *Opisthacanthus*, and *Pandinus*, in which family groups with overlapping generations cooperate to construct and occupy communal burrows, inhabited by individuals of various ages.

Reproduction and development. Reproduction in scorpions is indirect. Intromission occurs via a spermatophore. The male attaches the spermatophore to the substrate while grasping the female by the pedipalps or chelicerae during a mating dance or "promenade-aux-deux." The species-specific spermatophore catapults the sperm mass into the female gonopore when a lever is touched (a lock-and-key mechanism). All scorpions are viviparous (livebearing). Embryos develop in the reproductive tract and receive nourishment from maternal tissues.

Fig. 2. Scorpion, *Hadrurus arizonensis* (Ewing, 1928), adult male. (*Photo by Randy Mercurio; copyright © American Museum of Natural History*)

Some are parthenogenetic. Scorpions have low reproductive rates in comparison to other terrestrial arthropods and are among the most long-lived. Gestation times are long (several months to more than a year) and litter sizes small (1–105). Young are fairly large at birth and helpless (altricial), clinging to the mother for the first few molts before dispersing. Time to sexual maturity varies from 2 to 8 years, depending on the species. Average longevity is 4–5 years, but larger species may live 25–30 years. Scorpions do not molt as adults.

Small litter sizes, long generation times, and low survivorship among sexually immature females contribute to a low rate of population increase for most scorpions. Many are habitat specific, range-restricted, and sensitive to environmental degradation, exacerbating their risk of extinction due to human activities, including habitat destruction and harvesting for the souvenir and exotic pet trades.

Toxicity. Scorpion venoms contain multiple low-molecular-weight proteinaceous neurotoxins that block sodium and potassium channels, preventing the transmission of nerve impulses across synapses. Scorpion envenomation represents a significant cause of morbidity and mortality in some regions (such as Mexico, North Africa, and the Middle East). Some 100,000 scorpion stings occur annually in Mexico, killing as many as 800 people (mostly young children and the elderly). The figures may be even higher in North Africa and the Middle East. However, most scorpions are harmless. The sting may be painful, but not dangerous. About 25 species, all in the family Buthidae, are considered medically important worldwide. Most of these occur in the New World genera *Centruroides* and *Tityus*, and the Old World genera *Androctonus*, *Buthus*, *Leiurus*, *Mesobuthus* and *Parabuthus*. *Centruroides exilicauda*, from Arizona, California, and New Mexico, is the only species known to be lethal in the United States. The venom of these scorpions has proved fatal to healthy children up to 16 years of age and to adults suffering from hypertension and general debility.

Morphology. Scorpions are prominent in mythology and folklore. A classical zodiacal constellation is named after scorpions. The toxicity and fearsome appearance of scorpions contribute to a fascination that has always and continues to surround them. *See* ARACHNIDA.

Lorenzo Prendini

Bibliography. P. Brownell and G. Polis, *Scorpion Biology and Research*, Oxford University Press, 2001; V. Fet et al., *Catalog of the Scorpions of the World (1758–1998)*, New York Entomological Society, 2000; V. Fet and P. A. Selden, *Scorpions 2001: In Memoriam Gary A. Polis*, British Arachnological Society, Burnham Beeches, Bucks, UK, 2001; G. A. Polis, *The Biology of Scorpions*, Stanford University Press, 1990.

Scorpius

The Scorpion, a large zodiacal constellation (see **illustration**). It probably represents the scorpion in Greek mythology sent by Apollo to sting and kill Orion, in order to protect Apollo's sister Artemis's virginity. Now Scorpius and Orion are separated by half the celestial sphere from each other, though Scorpius eternally pursues Orion. *See* ORION; ZODIAC.

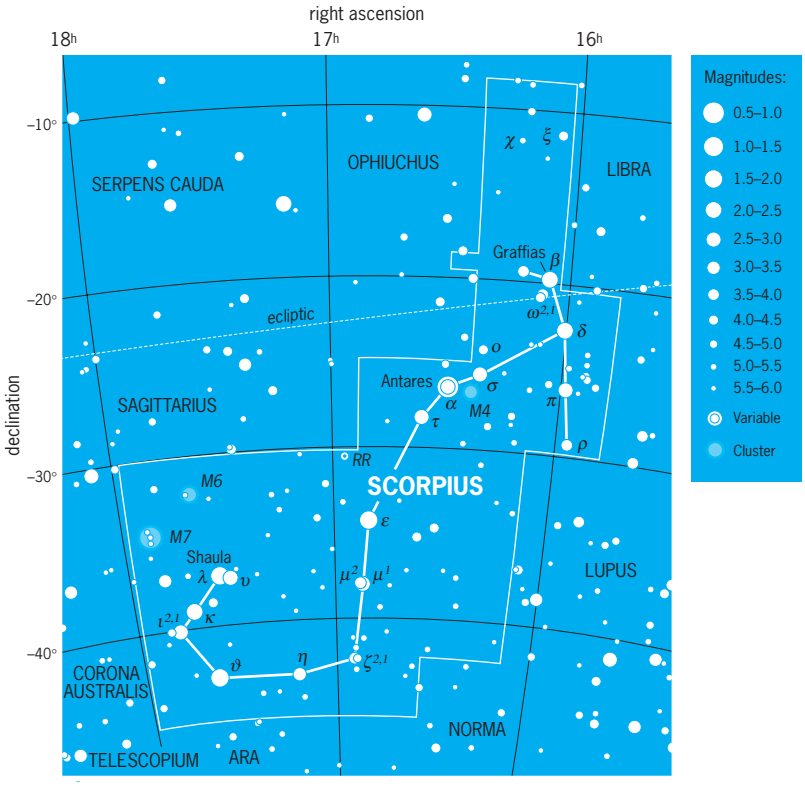

Modern boundaries of the constellation Scorpius, the Scorpion. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

The constellation loops through the sky like a giant fish-hook. The bright supergiant star Antares is the 15th brightest star in the sky. Its name means Rival of Mars, since both are bright and reddish in color, and Mars sometimes passes close to Antares. *See* ANTARES.

Scorpius lies along the Milky Way and contains many beautiful star clusters and star clouds. The globular cluster M4 and the open clusters M6 and M7 are especially noticeable. *See* MESSIER CATALOG; STAR CLOUDS; STAR CLUSTERS.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See also* CONSTELLATION.

Jay M. Pasachoff

Scramjet

An air-breathing jet engine that relies for propulsion on the compressing or ramming effect on air taken into the engine inlet at supersonic speeds, normally when the aircraft is traveling at speeds above Mach 4. The term is derived from "supersonic combustion ramjet." Scramjet technology now in development is expected to make global rapid travel and

affordable orbit access a reality in coming years. A breakthrough was realized in the spring of 2003 when a flight-weight scramjet engine was run at 4.5 and 6.5 times the speed of sound (Mach 4.5 and 6.5) using conventional hydrocarbon fuel. More significantly, this represents the first flight-weight, hydrocarbon-fueled scramjet engine. This test was conducted by the Pratt & Whitney Space Propulsion Hydrocarbon Scramjet Engine Technology (HySET) team, in partnership with and sponsored by the U.S. Air Force Research Laboratories (AFRL) Propulsion Directorate. *See* MACH NUMBER.

Scramjet operation. A ramjet is a jet engine that relies on the compressing or ramming effect on the air taken into the inlet while the aircraft is in motion. The geometry of the engine compresses the air rather than a piston, as in an internal combustion engine, or a compressor, as in a gas turbine. In a ramjet, the compressed air is slowed to subsonic speeds (less than Mach 1) as it is routed into the combustion chamber, where it is mixed with fuel, burned, and expanded through a nozzle to generate thrust.

A scramjet normally operates at speeds above Mach 4. To get a missile or aircraft to Mach 4, a rocket or jet engine is typically used for the initial acceleration. As the vehicle is traveling through the atmosphere on its way to Mach 4, the ignition condition for the scramjet, air is being compressed at hypersonic (greater than Mach 5) and supersonic (Mach 1–5) speeds on the lower surface of the aircraft and enters the engine at high pressure (**Fig. 1**). From the scramjet inlet, the air flows at supersonic speed into the combustor, where fuel is mixed with the air, burned, and expanded to make thrust. The mixing, burning, and expansion process takes place in less than 0.001 second. Engineers who developed this technology have referred to scramjet combustion as attempting to "light a match in a hurricane."

Advantages. In their operating speed regime, scramjets offer several advantages over conventional jet and rocket engines. Although not as efficient as a gas turbine engine, a scramjet has the ability to operate at much faster speeds. While burning hydrocarbon fuel, a scramjet engine can operate at speeds up to Mach 8. Using hydrogen as the fuel, a scramjet engine can operate at speeds in excess of Mach 12. Because of its air-breathing nature, the performance (specific impulse) of a scramjet is significantly higher than that of the best-performing rocket engines. One big advantage of a scramjet engine is the minimal number of moving parts required for operation. Since it relies on the shape of its flow path to compress the air by dynamic means, a scramjet engine does not require the use of high-pressure pumps, compressors, or turbines as do gas turbine or rocket engines. While reducing moving parts increases engine reliability and durability, benefits such as high speed, relative fuel efficiency, and range are the reasons that scramjets will revolutionize the aerospace industry.

The development of scramjet propulsion technology will enable affordable and reusable space transportation systems with operation much like a commercial airliner. Scramjet-powered access-to-space systems that will take off and land horizontally like commercial airplanes are being conceptualized today. These systems will be designed to achieve orbit while employing relatively low acceleration profiles, potentially opening access to orbital flight to a much larger section of the population.

Scramjets will be used in conjunction with either gas turbine engines or rockets in combined cycle

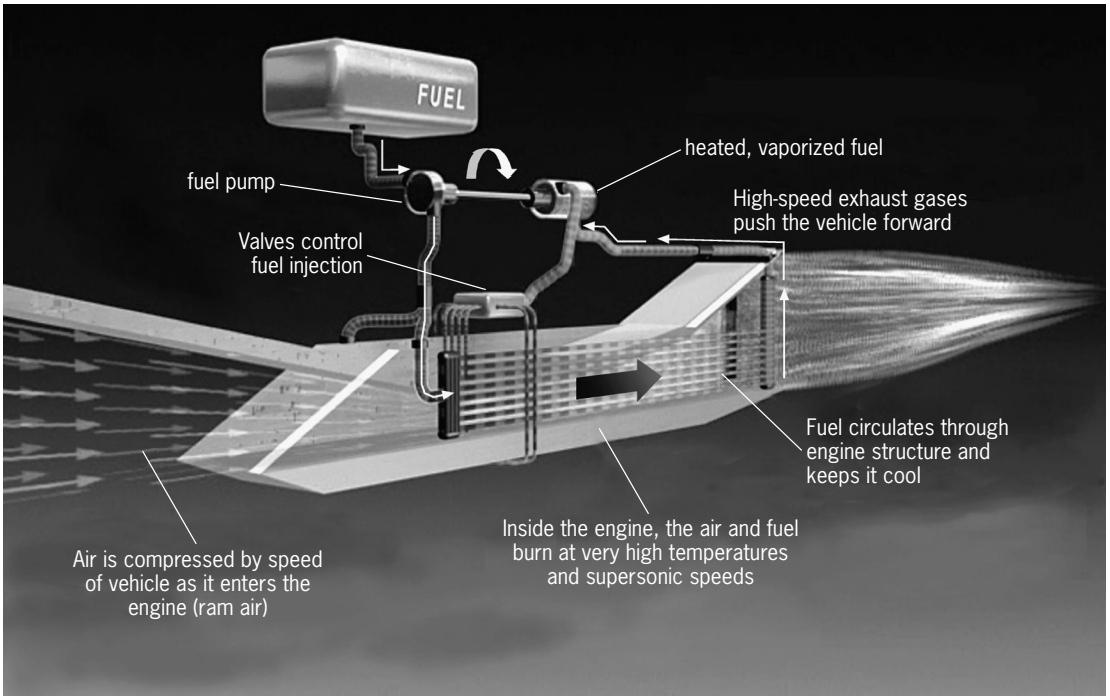

Fig. 1. Operation of a scramjet.

engines (CCEs). In a turbine-based combined cycle (TBCC) engine, a turbine engine will accelerate the aircraft from takeoff to a scramjet takeover speed (Mach 3–4), at which point the scramjet will be ignited and the turbine will be shut down. In a rocket-based combined cycle (RBCC), small rockets embedded in the flow path of the scramjet will first be used as ejectors to accelerate the vehicle to ramjet and then to scramjet operation. In both cases, the CCEs will be used to accelerate the space vehicle to escape velocity while flying through the atmosphere using the oxygen in the air. As the vehicle gains altitude and escapes the Earth's atmosphere, propulsion will transition to rockets since air will not be available to the scramjets. *See* ROCKET PROPULSION; TURBINE PROPULSION.

Future air-breathing access-to-space systems could increase the payload fraction by two-to-five times that of conventional rocket-powered systems. Since the scramjet engine takes its oxidizer from the air, a scramjet-powered vehicle does not require an oxidizer tank or a supply control system. This reduction in oxidizer volume and weight can be translated into smaller vehicles and reduced launch costs.

Initial technology developments for reusable applications will concentrate on developing variable geometry to optimize scramjet performance and extend the Mach number envelope. Since a hypersonic vehicle will spend prolonged periods traveling at high speed in the atmosphere, its surface will require significantly more thermal protection than the expendable launch systems in use today. Structural and thermal durability will be increased to allow multimission capability. Additionally, scientists and engineers are defining scalability factors that must be taken into consideration when designing the larger engines required for space applications.

Development. The U.S. Air Force established the Hypersonic Technology (HyTech) Program in 1995 after the National Aero-Space Plane's development was terminated. In 1996, Pratt & Whitney won a contract for HySET to demonstrate the performance and structural durability of a hydrocarbon-fueled scramjet engine. As a result of the HySET Program's success the X43-C, a joint USAF/NASA program, has emerged. The goal of this program is to flight-test a three-module derivative of the HySET engine simulating elements of a space mission. In parallel and as risk reduction for the X43-C Program, the Air Force is also considering flying a propulsion flight demonstration program: the Endothermically Fueled Scramjet Engine Flight Demonstrator (EFSEFD) Program. This flight demonstration will use a single HySET engine derivative module and will focus on the military application of the scramjet.

Building block approach. Once the hypersonic vehicle mission requirements were translated into scramjet engine component requirements, the HySET team generated the computational fluid dynamic codes. These codes define combustion products and optimize the fueling locations and concentrations required for top performance. Their results were used to design the hardware, allowing optimization of the engine design prior to test.

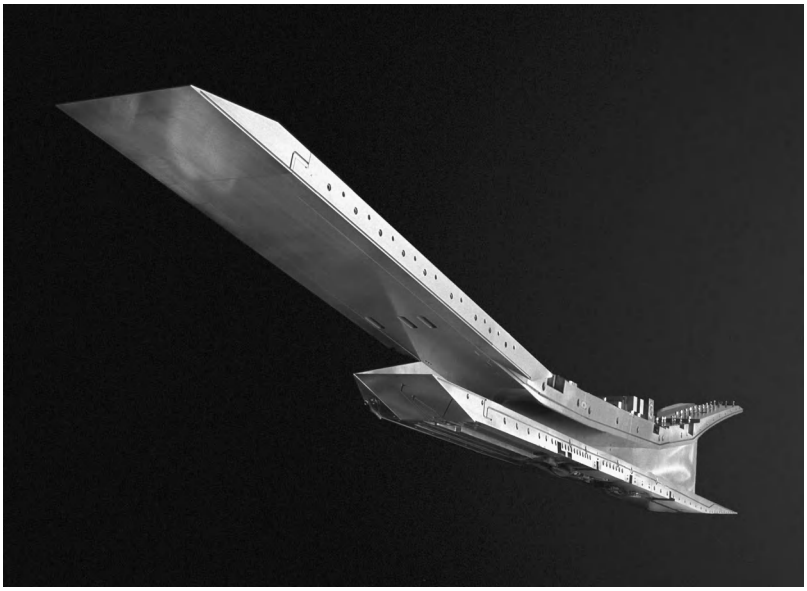

Fig. 2. Performance Test Engine (PTE).

From subscale testing, the team updated codes and defined the Performance Test Engine (PTE). This full-scale test of a heat sink engine validated the computer codes and demonstrated engine performance and operability. The PTE is a heavy-weight, heat-sink engine fabricated in solid copper (**Fig. 2**). The use of conventional jet fuel provided a major challenge to the team. Prior to this program, most developmental scramjets used hydrogen, which is much easier to mix and burn than jet fuel. In addition, the team developed a special fuel-conditioning system to simulate the fuel temperatures that would be generated by a fuel-cooled engine in flight. *See* AIRCRAFT FUEL.

The PTE was operated at Mach 4.5 and 6.5 consistent with scramjet takeover and cruise conditions for a missile application. The team tested and examined 95 test points, demonstrating excellent operability. The engine achieved 110% of its performance test objectives at Mach 4.5 conditions and 100% of its test objectives at Mach 6.5. Based on the performance and operability of the engine during test, the team went on to design and fabricate a more sophisticated full-scale, fuel-cooled engine. *See* AIRCRAFT FUEL.

As part of the building block approach, the team built on lessons learned, and used test results from the PTE to calibrate and improve analytical tools leading to the evolution of the Ground Demonstration Engine (GDE). With every test result, the analytical tools used for the design were calibrated, resulting not only in the ability to understand the behavior of the hardware at test but also in the definition of a design set of tools to be employed in subsequent iterations. Along with the development of the scramjet's operability and performance, the HySET team used the building block approach for developing the thermal and structural technologies. Multiple thermostructural components with increasingly larger sizes were designed, built, and tested in

scramjet operating conditions. In 2001, this culminated in the creation of the first hydrocarbon-fueled scramjet engine, GDE-1, weighing less than 150 lb (59 kg) and using readily available nickel-based alloys for its fabrication.

Operation. In the GDE-1 engine, the fuel must enter the combustor as a vapor in order to facilitate mixing and burning with the airflow. This is achieved through regenerative cooling of the hardware using a revolutionary endothermic cooling approach. The liquid fuel enters the front of the engine and is routed through the structure to the rear of the engine. The fuel cools the metallic structure as it passes through very small longitudinal passages. The passages in the structure are coated with a catalyst. While the fuel is absorbing the heat from the structure, the fuel temperature and pressure rise, causing the catalyst to "crack" the fuel in an endothermic reaction; that is, the fuel breaks down into gaseous lighter hydrocarbon components such as hydrogen, ethylene, and methane. This process gasifies the fuel for easier burning while allowing the fuel to absorb a significantly larger amount of heat from the scramjet structure than otherwise would be possible. The process is akin to a flying fuel refinery. At the rear of the engine structure, the cracked fuel is collected and metered into the scramjet combustor and burned. The combustion process expands through a nozzle, where it produces positive net thrust.

Testing. In the spring of 2003, the GDE-1 was tested initially at low speed (Mach 4.5) and then at Mach 6.5 (**Fig. 3**). As in the prior component tests, the results were used to calibrate and validate the design tools. By July 2003, GDE-1 had completed about 60 tests, successfully validating the thermostructural, performance, and operability predictions. The results of GDE-1 testing have proven that with today's technology scramjets fueled by hydrocarbon fuels can generate the thrust required to support hypersonic flight.

The next scramjet engine, GDE-2, was also fuel-cooled and flight-weight as was GDE-1. However, this was designed as a full system test to introduce control hardware and software so that the engine could run as a complete closed-loop system. Fabrication of the GDE-2 was completed in the spring of 2005.

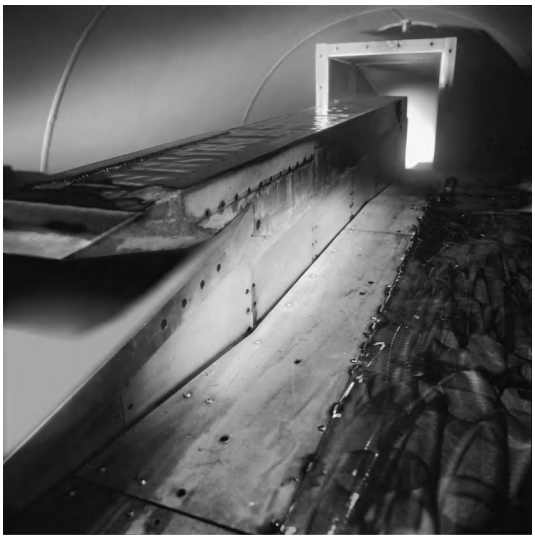

Fig. 3. GDE-1 test at Mach 4.5.

Outlook. Building on the lessons learned during PTE, GDE-1, and GDE-2 ground testing, the HySET team's goal is flight demonstration of this technology. The X43-C and EFSEFD programs will provide the first opportunities to demonstrate the operation of a hydrocarbon-fueled scramjet engine in flight. The lessons learned during these flight experiments will pave the way for hypersonic flight. *See* HYPERSONIC FLIGHT.

Joaquin Castro

Bibliography. E. T. Curran and S. N. B. Murthy (eds.), *Developments in High-speed Vehicle Propulsion Systems*, vol. 165, AIAA, 1996; E. T. Curran and S. N. B. Murthy (eds.), *High-speed Flight Propulsion Systems*, vol. 137, AIAA, 2000; E. T. Curran and S. N. B. Murthy (eds.), *Scramjet Propulsion*, vol. 189, AIAA, 2000.

Scrapie

A transmissible, invariably fatal disease of adult sheep and goats characterized by degeneration of the central nervous system. It is one of the transmissible spongiform encephalopathies, a disease family which includes bovine spongiform encephalopathy (BSE) in cattle (commonly known as mad cow disease), chronic wasting disease in deer, and Kuru and variant Creutzfeldt-Jakob disease in humans. Scrapie was first discovered in European sheep in the early eighteenth century and is now present in many sheep populations worldwide, excluding Australia and New Zealand. *See* NERVOUS SYSTEM DISORDERS.

Clinical signs. Scrapie has a long incubation period. It is widely believed that most animals are infected at or shortly after birth, with about 90% of sheep developing signs of disease at 2–5 years of age. Scrapie affects both sexes and is insidious and progressive in its onset, with a clinical course that lasts weeks or months. The clinical signs can be highly variable between flocks, depending on the strain of agent and breed of sheep, but often begin with hyperexcitability or dullness, and proceed to pruritus (itching), ataxia (incoordination), trembling, and loss of condition. Pruritus is accompanied by scratching or scraping (hence the name), leading to wool loss (**Fig. 1**). There is no treatment for scrapie, and affected animals will, if not euthanized, become recumbent and die.

Etiology. Scrapie is caused by a transmissible agent, which can be found in cell-free filtrates of the brain and other organs from scrapie-affected sheep. Importantly, inocula remain infectious in the absence of detectable nucleic acid, but lose infectivity after treatment with methods that denature protein. The prion (proteinaceous infectious particle) hypothesis, for which Stanley Prusiner was awarded the 1997 Nobel Prize in Physiology or Medicine, proposes that the etiologic agent of scrapie and other transmissible spongiform encephalopathies is

the host protein, PrP, which can exist in a nonpathological form, PrP^{C} or PrP_{Sen}, and a pathological form, PrP^{Sc} or PrP_{Res}, with the two forms differing in conformation only. According to the prion hypothesis, PrP^{Sc} converts PrP^{C} into more PrP^{Sc}, and this cascade reaction leads, ultimately, to disease pathology. However, PrP^{Sc} could simply be a marker for infection, rather than the infectious agent itself, as transmissible spongiform encephalopathy infections have been reported in the absence of detectable PrP^{Sc}. *See* PRION DISEASE.

Pathology. Infection is believed to occur mostly by the oral route, with entry to the body through the ileal Peyer's patch (a collection of lymphoid follicles in the ileum of the small intestine), and dissemination through the lymphatic/vascular system to tissues of the lymphoreticular system, including tonsil, spleen, retropharyngeal, and mesenteric lymph nodes. With modern immunohistochemical methods, the presence of PrP^{Sc} can be detected in the lymphoreticular system just months after infection, primarily within the follicular dendritic cells of primary and secondary B cell follicles. Infection spreads from the lymphoreticular system through the autonomic nervous system to the central nervous system, where PrP^{Sc} accumulates primarily in glial cells (astrocytes) and neurons. Within the brain there is astrocytic gliosis (production of astrocytes in area of neurodegeneration), the accumulation of deposits of PrP^{Sc}, vacuolation, and the characteristic spongiform degeneration (**Fig. 2**). *See* CELLULAR IMMUNOLOGY; LYMPHATIC SYSTEM.

It is possible that some infections are transmitted via nonoral routes. For example, scrapie can be efficiently transmitted by scarification (skin or tissue incision) and via the conjunctiva (mucous membrane covering the eyeball and lining the eyelids).

Transmission. There is very little evidence for vertical transmission (direct transmission from infected mother to offspring in utero) of any transmissible spongiform encephalopathy, including scrapie. By contrast, there is substantial evidence for horizontal (direct lateral; among related or unrelated individuals) transmission of scrapie between sheep housed together, although the precise mechanism is not clear. However, it is widely believed that many scrapie infections are acquired by lambs at or shortly after birth. Infectivity and PrP^{Sc} can be readily detected in the placenta and fetal membranes of infected ewes, and these tissues are believed to play a central role in transmission. One model for horizontal transmission is that an infected ewe gives birth to lambs that are themselves uninfected but may be accompanied by an infected placenta and other birth tissues/fluids. The newly born lambs and others that are in close association (within the same lambing pen, perhaps), if genetically susceptible to scrapie, then become infected by ingestion of infected material.

There are reports in the scientific literature of sheep, sourced originally from scrapie-free areas, acquiring scrapie after being kept on farmland that previously held infected stock. For example, in Iceland in the 1940s all sheep within the scrapie-affected north were slaughtered to help eradicate another sheep disease, maedi visna. Restocking occurred 1 to 3 years later with lambs sourced from the scrapie-free south. Reportedly, scrapie reappeared in farms that were previously scrapie-affected, but not in farms that were previously disease-free. These observations have led to the concept of the environmental persistence of the scrapie agent and the possibility of indirect lateral transmission between sheep mediated by a contaminated environment. How sheep actually acquire the infection from the environment is unknown.

Fig. 1. Poll Dorset sheep with severe clinical signs of scrapie.

The route of scrapie transmission between goats is also not known, but is probably similar to that of sheep.

Genetics. Whether a sheep can acquire scrapie infection or, following infection, proceed to clinical disease is under the control of the gene that encodes its PrP protein, called the *PrP* gene. Five polymorphic variants (alleles or haplotypes) of the *PrP* gene that exhibit this linkage to scrapie are commonly found in sheep. The variants encode the following amino acid sequences in PrP: alanine (A) at codon 136, arginine (R) at codon 154, and arginine (R) at codon 171 ($A_{136}R_{154}R_{171}$, or ARR for short), and, continuing this notation, ARQ, AHQ, ARH, and VRQ (in which Q is glutamine, H is histidine, and V is valine).

Each sheep inherits one haplotype from each of its parents, giving it 1 of 15 possible *PrP* genotypes often written as, for example, ARQ/VRQ or ARR/AHQ. The relationship between the haplotypes/genotypes and susceptibility to scrapie is complex. Briefly, ARR and AHQ confer resistance to classical scrapie, while ARQ and VRQ confer susceptibility. The ARR/ARR genotype of sheep is the most resistant to scrapie, while the VRQ/VRQ genotype is the most susceptible. The effect of the ARH haplotype

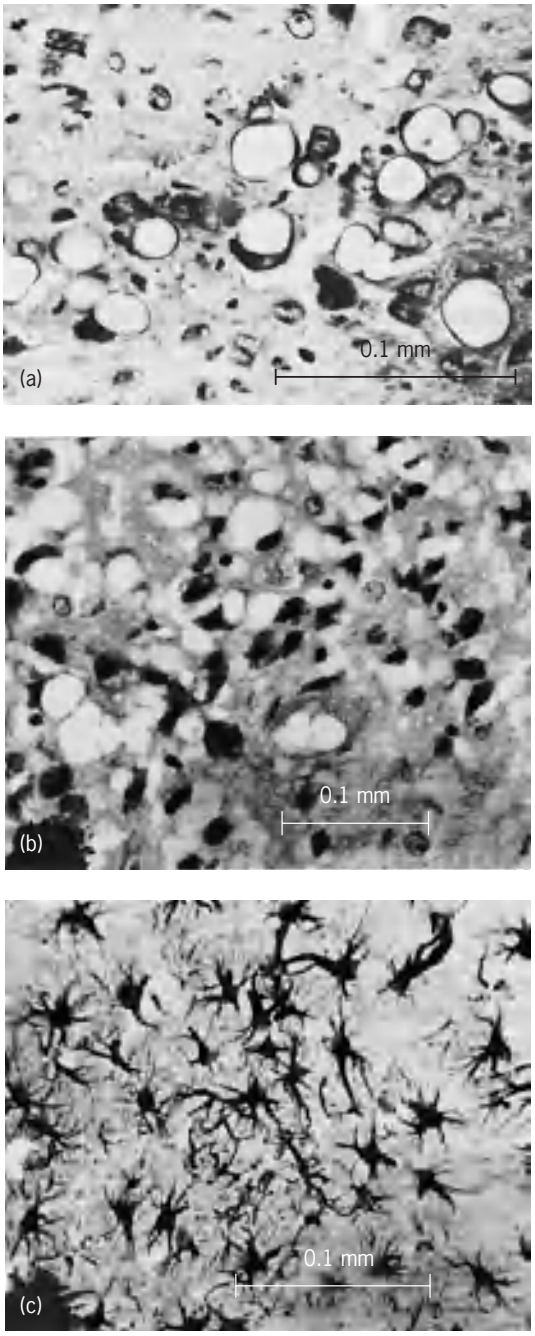

Fig. 2. Photomicrographs of the brain from scrapie-affected animals showing (*a*) vacuoles in neurons of the medulla, (*b*) spongy degeneration, (*c*) astrocytic hypertrophy (astrocytic gliosis in mice)

varies, depending on the other haplotype present. Small changes in genotype can have a remarkably large effect on susceptibility to scrapie. For example, British sheep of the ARQ/VRQ genotype are 300 times more likely to be reported with scrapie than sheep of the AHQ/VRQ genotype, despite differing only in arginine or histidine being encoded by codon 154 of one of the two haplotypes. The molecular mechanism that links *PrP* genotype to resistance/susceptibility is not understood.

Several strains of scrapie exist, and some exhibit patterns different from those described above. In many countries, including the United States, the VRQ haplotype is absent in the major sheep breeds, and sheep of the ARQ/ARQ genotype are most susceptible to the local scrapie strains. Sheep of the AHQ/AHQ genotype are the most susceptible to a scrapie strain discovered in Norway in 1998 (Nor98), while VRQ/VRQ sheep appear to be resistant to it. Novel diagnostic tests for scrapie have detected sheep with "atypical" scrapie infections (so-called because of an unusual distribution of PrP^{Sc} in the brain), including some of the ARR/ARR genotype which have hitherto been considered entirely or almost entirely resistant to scrapie. Under experimental conditions bovine spongiform encephalopathy can be transmitted to sheep and behaves as a novel scrapie strain; sheep of the ARQ/ARQ genotype are most susceptible, and sheep of the ARR/ARR genotype can succumb to infection and clinical disease if inoculated intracerebrally. *See* ALLELE; GENE; GENE ACTION; GENETIC CODE.

Control. Many approaches to scrapie control have been tried in the past, such as cleaning and disinfection of premises and culling of affected animals, entire populations of age-cohorts, genetically related animals, and even whole-flock slaughter. Such methods have generally not brought long-term success. The recent discovery of the strong association between scrapie and *PrP* genotype has, however, raised the possibility of controlling the disease at a national level by selective breeding. Genetic control programs are now under way in many countries, including the United States and the entire European Union. These programs are based on large-scale genetic testing, followed by culling of the most genetically susceptible sheep, and promoting the use of ARR/ARR sheep for breeding. The extreme resistance of sheep of the ARR/ARR genotype to most strains of transmissible spongiform encephalopathy suggests a high probability of success for these programs, although the recent discovery of "atypical" scrapie infections in ARR/ARR sheep is an important concern. *See* INFECTIOUS DISEASE; VETERINARY MEDICINE.

M. Baylis

Bibliography. H. B. Parry, *Scrapie Disease in Sheep: Historical, Clinical, Epidemiological, Pathological and Practical Aspects of the Natural Disease*, Academic Press, Oxford, 1983; S. B. Prusiner, *Prion Biology and Diseases*, 2d ed., Cold Spring Harbor Laboratory Press, 2004; R. M. Ridley and H. F. Baker, *Fatal Protein. The Story of CJD, BSE, and Other Prion Diseases*, Oxford University Press, 1998.

Screening

A mechanical method of separating a mixture of solid particles into fractions by size. The mixture to be separated, called the feed, is passed over a screen surface containing openings of definite size. Particles smaller than the openings fall through the screen and are collected as undersize. Particles larger than the openings slide off the screen and are caught as oversize. A single screen separates the feed into only two fractions. Two or more screens may be operated in series to give additional fractions. Screening

occasionally is done wet, but most commonly it is done dry.

Industrial screens may be constructed of metal bars, perforated or slotted metal plates, woven wire cloth, or bolting cloth. The openings are usually square but may be circular or rectangular. In rectangular openings the separation is controlled by the smaller dimension.

When the opening in a screen is larger than 1 in. (2.5 cm), the actual opening size in inches is specified. When the opening is 1 in. or less, the screen size is designated as the number of openings per linear inch; that is, a 20-mesh screen has 20 openings per inch. The actual size of an opening is less than that corresponding to the mesh number by the thickness of the metal between openings. Mesh sizes range from about 4 in. (10 cm) to 400-mesh, but screens finer than 100- or 150-mesh are seldom used industrially. Particles smaller than this are usually separated by other methods. *See* MECHANICAL CLASSIFICATION; SEDIMENTATION (INDUSTRY).

Testing sieves are used to measure the size distribution of mixtures of particles and to specify the performance of commercial screens and other equipment. They are made of woven wire screening. Mesh and wire size of each screen are standardized. The usual range is from 1- to 200-mesh. In use, a set of standard screens is arranged serially in a stack, with the finest mesh on the bottom and the coarsest on top. An analysis is conducted by placing a weighed sample on the top screen and shaking the stack mechanically for a definite time. Particles passing the bottom screen are caught in a pan. The particles remaining on each screen are removed and weighed, and the weights of the individual increments are converted to percentages of the total sample.

In most screens the particles are pulled through the openings by gravity only. Large particles are heavy enough to pass through the openings easily, but intermediate and smaller particles are too light to pass through the screen unaided. Most screens are agitated to speed the particles through the openings. Common methods are to revolve a cylindrical screen about an axis slightly inclined to the horizontal or to shake, gyrate, or vibrate a flat screen.

Two effects tend to restrict the fall of small particles through a screen: blinding of the screen by wedging of particles into the openings and sticking of individual particles to each other and to the screen. The motion of the screen reduces blinding, and sometimes positive means are provided to drive wedged particles back through the openings. Sticking is severe if the particles are damp, and may be reduced by drying the feed.

Two common screens are shown in the **illustration**. The trommel, a revolving screen, is a combination of four screens in series. This unit gives five products. The gyrating screen uses bouncing balls under the sizing screens to dislodge particles caught in the meshes. The balls are supported on coarse screens which offer no resistance to the materials passing the sizing screens. This screen gives four products.

Ideally, a screen should effect a sharp separation between undersize and oversize, and the largest particle in the undersize should be just smaller than the smallest particles in the oversize. An actual screen does not do this, and appreciable portions of both fractions have the same size range. Even testing screens shows this overlap between adjoining fractions. *See* MECHANICAL SEPARATION TECHNIQUES.

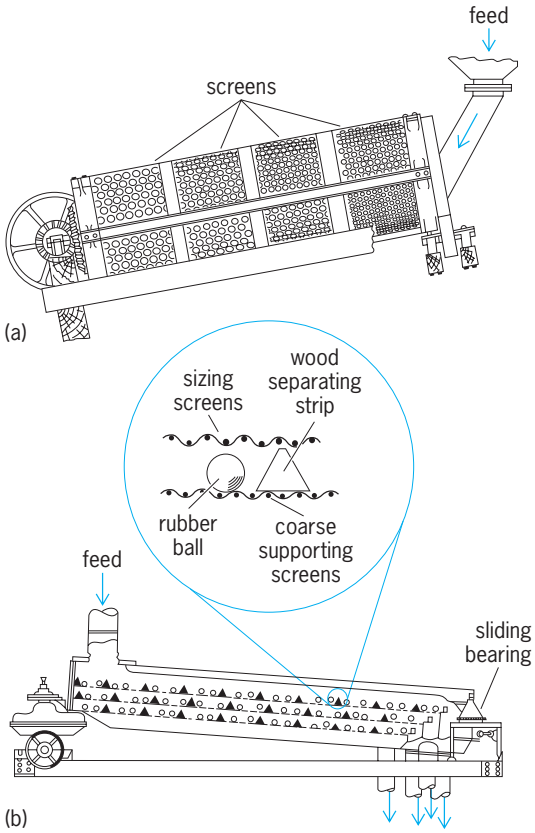

Types of screens. (*a*) Trommel. (*b*) Gyratory.

Warren L. McCabe

Bibliography. A. Grandison and M. Lewis, *Separation Processes*, 1993; W. L. McCabe, J. C. Smith, and P. Hariott, *Unit Operations in Chemical Engineering*, 5th ed., 1993.

Screening smoke

Smoke clouds used for screening purposes. Traditionally, smoke is defined as a visible aerosol (that is, a suspension of solid or liquid particles in a gas) resulting from incomplete combustion. Military screening smokes are a special case. They are generally composed of liquid droplets with a mean particle size around 1 micrometer, and produce a bright white cloud of long duration. The effectiveness of these aerosols results from their size and composition. The individual aerosol particles efficiently scatter light in all directions. This scattering, known as Mie scattering, reduces the intensity of any light passing through the cloud and produces considerable glare, the white cloud effect, to reduce the target-to-background contrast. *See* SCATTERING OF ELECTROMAGNETIC RADIATION; SMOKE.

In warfare, screening smokes have three major functions: blinding, covering, and deception.

Blinding smokes are placed on an enemy's position to cause confusion and reduce enemy effectiveness. Covering smokes are usually placed between the enemy and a combat unit or on the unit's position to deny the enemy knowledge of the unit's location and actions. Deception, that is, putting a smoke screen in a location of no immediate interest, has been used throughout history to misdirect enemy interests and to gain surprise.

Screening smoke compositions have remained unchanged since the mid-1940s. However, methods of delivery have been improved to keep pace with the rapid rise in the use of high-speed, sophisticated weaponry.

There are four well-established methods of smoke production: burning phosphorus in air, burning pyrotechnic compositions, vaporization and recondensation of oils, and dispersion of reactive liquids.

Phosphorus smokes. The most widely used and effective screening smokes are formed by burning elemental phosphorus in air. The reacting phosphorus rapidly forms a variety of phosphorus oxides [reactions (1) and (2)], which absorb atmospheric

$$P_4 + 5O_2 \rightarrow 2P_2O_5 \quad (1)$$

$$P_2O_5 + 3H_2O \rightarrow 2(H_3PO_4) \quad (2)$$

moisture to form aerosols of dilute phosphoric acid. Phosphorus smokes may be delivered as white phosphorus, plasticized white phosphorus, or red phosphorus. Various munition configurations have been developed to control the burning rate of the phosphorus oxidation reaction and thus eliminate the problem of thermal pluming found in early white phosphorus munitions (see **illus.**). Phosphorus smoke clouds are very efficient. They have a yield factor (that is, the ratio of the mass of smoke produced to the mass of starting material) of between 3 to 7 depending on the relative humidity. These high yields result from the phosphorus oxides being hygroscopic and thus absorbing significant amounts of atmospheric water. Phosphorus smokes are delivered from artillery projectiles, grenades, mortar rounds, and rockets. *See* PHOSPHORUS.

Thermal pluming from a white phosphorus munition. (*U.S. Army*)

Pyrotechnic compositions. Aerosols formed by burning pyrotechnic compositions generally are composed of hydrated chloride salts. Most NATO countries use a composition known as HC, a mixture of zinc oxide, hexachloroethane, and various burning rate modifiers, which produces a hydrated zinc chloride aerosol. Other nations use an alternate composition, often called the Yershov composition, which contains ammonium chloride plus fuels and modifiers and produces an ammonium chloride aerosol. Both formulations have a yield factor ranging from 0.9 to 2, depending on relative humidity, and both produce aerosols with similar optical properties. The principal military munitions containing these mixtures are grenades and smoke pots. *See* PYROTECHNICS.

Oil smokes. Oil smokes are used when large areas must be screened for long periods. These smokes are produced by vaporizing an oil and allowing it to cool in air, where it condenses into fine aerosol droplets that form a stable cloud. In general, they are much less expensive than other methods for producing screens. Oil smokes are produced by spraying diesel oil on the exhaust manifold of a vehicle and allowing the exhaust gas pressure to propel the cloud away from the vehicle, or by using fog oil, a specially produced smoke oil, with mechanical smoke generators. The optical properties of the smoke produced by the two oils are identical; however, fog oil clouds last much longer because of their lower vapor pressure. During World War II, the Allies used simultaneous operation of many fog oil generators to screen several square miles of terrain, port areas, and staging zones from air observation. Some of these screens lasted for over 30 days.

Oil smokes are also useful for nonmilitary purposes. Farmers have used oil smoke for many years to prevent the freezing of fruit and other crops. The blanket of smoke reduces the loss of heat by radiation from ground and plant surfaces. Oil smoke is also used as a carrier for insecticides and fungicides in agriculture and forestry.

Reactive liquid smokes. The most common reactive smokes are chlorosulfonic acid, which produces sulfuric acid smoke, and titanium tetrachloride, which produces a dilute titanium hydroxide and hydrochloric acid smoke mixture. Both have optical properties similar to phosphorus smokes. These reactive liquids are obsolete in the U.S. Army inventory for several reasons: they are highly corrosive and dangerous; smoke is produced by hydrolysis reactions, thus making them ineffective in extremely dry conditions; and their main method of employment, spray from tanks on the wings of slow-flying aircraft, is no longer militarily acceptable. Joseph J. Vervier

Bibliography. A. Baily and S. G. Murray, *Explosives, Propellants, and Pyrotechnics*, 1989; J. A. Compton, *Military Chemical and Biological Agents: Chemical and Toxicological Properties*, 1988; W. C. Hinds, *Aerosol Technology; Properties, Behavior, and Measurement of Airborne Particles*, 2d ed., 1998.

Screw

A cylindrical body with a helical groove cut into its surface. For practical purposes a screw may be considered to be a wedge wound in the form of a helix so that the input motion is a rotation while the output remains translation. The screw is to the wedge much the same as the wheel and axle is to the lever in that it permits the exertion of force through a greatly increased distance. *See* WEDGE.

Figure 1 shows a frictionless screw with square threads mounted on a ball thrust bearing being used to raise a load Q. The load and other forces may be considered to be concentrated at an effective radius called the pitch radius. The force diagram is similar to the diagram for a wedge, except that wedge angle θ is replaced by screw lead angle λ, which is the angle between the thread of the helix and a plane perpendicular to the axis of rotation. Subscript t is added to indicate that force F is applied tangent to a circle, giving Eq. (1). Because F_t acts at the pitch radius R, the torque on the screw is given by Eq. (2).

$$F_t = Q \tan \lambda \tag{1}$$

$$T = F_t R = QR \tan \lambda \tag{2}$$

When a practical screw is considered, friction becomes important, and the torque on a screw with square threads is given by Eq. (3), and the efficiency

$$T = QR \frac{\tan \lambda + \mu}{1 - \mu \tan \lambda} \tag{3}$$

of a screw with square threads by Eq. (4), where μ is the coefficient of friction.

$$\eta = \tan \lambda \frac{1 - \mu \tan \lambda}{\tan \lambda + \mu} \tag{4}$$

If a screw jack (**Fig. 2**) is used to raise a heavy object, such as a house or machine, it is normally desirable for the screw to be self-locking, that is, for the screw to not rotate and lower the load when the torque is removed with the load remaining on the jack. *See* SCREW JACK.

For a screw to be self-locking, the coefficient of friction must be greater than the tangent of the lead angle: $\mu > \tan \lambda$. *See* SCREW FASTENER.

The screw is by far the most useful form of inclined plane or wedge and finds application in the bolts and nuts used to fasten parts together; in lead and feed screws used to advance cutting tools or parts in machine tools; in screw jacks used to lift such objects as automobiles, houses, and heavy machinery; in screw-type conveyors (**Fig. 3**) used to move bulk materials; and in propellers for airplanes and ships. *See* SCREW THREADS; SIMPLE MACHINE.

Richard M. Phelan

Bibliography. J. Phillips, *Freedom in Machinery*, vol. 1: *Introducing Screw Theory*, 1984, vol. 2: *Screw Theory Exemplified*, 1990.

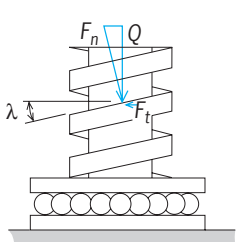

Fig. 1. Diagram of screw. See text for symbols.

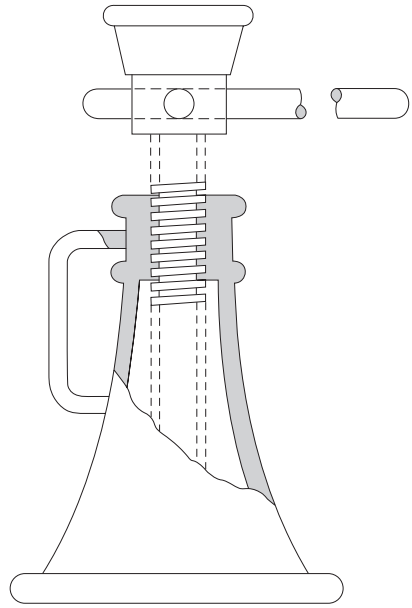

Fig. 2. Screw jack, used to raise heavy objects.

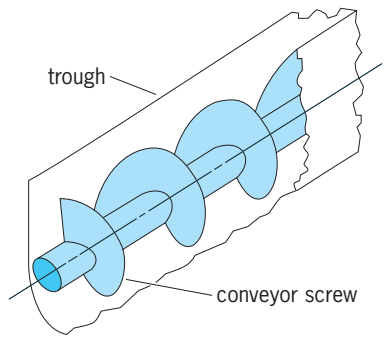

Fig. 3. Screw conveyor, used to move bulk objects.

Screw fastener

A threaded machine part used to join parts of a machine or structure. Screw fasteners are used when a connection that can be disassembled and reconnected and that must resist tension and shear is required. A nut and bolt is a common screw fastener. Bolt material is chosen to have an extended stress-strain characteristic free from a pronounced yield point. Nut material is chosen for slight plastic flow.

The nut is tightened on the bolt to produce a preload tension in the bolt (see **illus.**). This preload has several advantageous effects. It places the bolt under sufficient tension so that during vibration the relative stress change is slight with consequent improved fatigue resistance and locking of the nut.

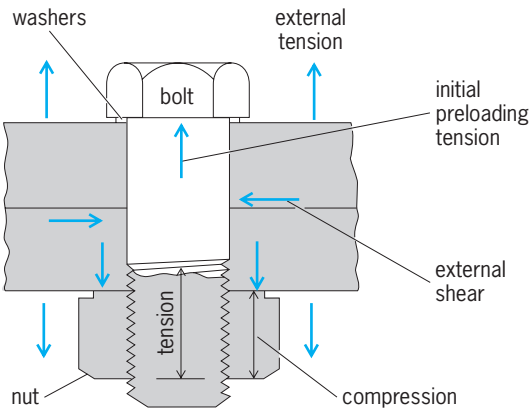

Forces on a bolt fastener under preloading.

Preloading also increases the friction between bearing surfaces of the joined members so that shear loads are carried by the friction forces rather than by the bolt.

The tightened nut is under compression; the bolt is under tension. The deformation that accompanies these forces tends to place the entire preload on the thread nearest the bearing surface. Thus, concentration of loading is counteracted if the nut is slightly plastic so as to set under load. This yielding may be achieved by choice of material or by special shape. If a soft gasket is used in the assembly, preloading is less effective; the bolt may carry the full tension and shear loads. *See* BOLT; JOINT (STRUCTURES); NUT (ENGINEERING); SCREW THREADS. Frank H. Rockett

Bibliography. American Society for Testing and Materials, *Section 15, General Products, Chemical Specialties, and End Use Products*, vol. 15.08: *Fasteners*, 1989; R. O. Parmley, *Standard Handbook of Fastening and Joining*, 3d ed., 1997; J. E. Shigley and C. R. Mischke, *Fastening, Joining, and Connecting: A Mechanical Designers' Handbook*, 1989; J. E. Shigley and J. J. Vicker, *Theory of Machines and Mechanisms*, 2d ed., 1995.

Screw jack

A mechanism for lifting and supporting loads, usually of large size. A screw jack mechanism consists of a thrust collar and a nut which rides on a bolt; the threads between the nut and bolt normally have a square shape. A standard form of screw jack has a heavy metal base with a central threaded hole into which fits a bolt capable of rotation under a collar thrusting against the load. Screw jacks are also used for positioning mechanical parts on machine tools in order to carry out manufacturing processes. These jacks are small in size and have standard V threads. Load and stress calculations may be performed on screw jacks in the same manner as in power screws and screw fastenings. *See* SCREW; SIMPLE MACHINE. James J. Ryan

Bibliography. N. Chironis, *Mechanisms and Mechanical Devices Sourcebook*, 2d ed., 1996; V. M. Faires and R. M. Keown, *Mechanism*, 5th ed., 1960, reprint 1980; S. Molian, *Mechanism Design: An Introductory Text,* 2d ed., 1997.

Screw threads

Continuous helical ribs on a cylindrical shank. Screw threads are used principally for fastening, adjusting, and transmitting power. To perform these specific functions, various thread forms have been developed. A thread on the outside of a cylinder or cone is an external (male) thread; a thread on the inside of a member is an internal (female) thread (**Fig. 1**). *See* SCREW.

Types of thread. A thread may be either right-hand or left-hand. A right-hand thread on an external member advances into an internal thread when turned clockwise; a left-hand thread advances when turned counterclockwise. If a single helical groove is cut or formed on a cylinder, it is called a single-thread screw. Should the helix angle be increased sufficiently for a second thread to be cut between the grooves of the first thread, a double thread will be formed on the screw. Double, triple, and even quadruple threads are used whenever a rapid advance is desired, as on fountain pens and valves. The helices on a double thread start 180° apart, while those on a triple begin 120° apart. A multiple thread produces a rapid advance without resort to a coarse thread.

Pitch and major diameter designate a thread. Lead is the distance advanced parallel to the axis when the screw is turned one revolution. For a single thread, lead is equal to the pitch; for a double thread, lead is twice the pitch. For a straight thread, the pitch diameter is the diameter of an imaginary coaxial cylinder that would cut the thread forms at a height where the width of the thread and groove would be equal.

Thread forms have been developed to satisfy particular requirements (**Fig. 2**). Those employed on fasteners and couplings and those used for making adjustments are generally of the modified 60° V type (unified and ISO metric). Where strength is required for the transmission of power and motion, a thread having faces that are more nearly perpendicular to the axis is preferred, such as the modified square and the acme. These threads, with their strong thread sections, transmit power nearly parallel to the axis of

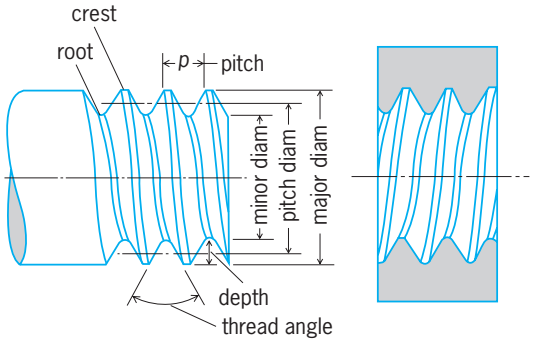

Fig. 1. Screw thread nomenclature.

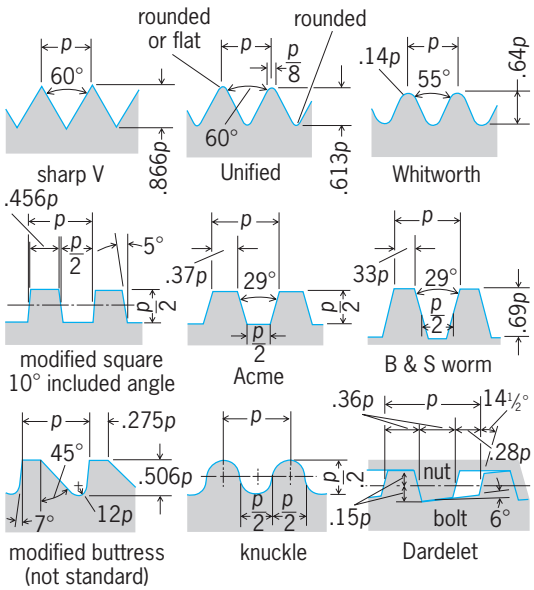

Fig. 2. Screw threads, for particular needs. (*After W. J. Luzadder, Fundamentals of Engineering, 8th ed., Prentice-Hall, 1977*)

the screw. The sharp V, formerly found on set screws, is now rarely used because of the difficulty of cutting sharp roots and crests in quantity production.

Thread forms. The unified thread form accompanied by the unified thread standards has in the past fulfilled the necessary requirements for interchangeability of threaded products between the United States, Great Britain, and Canada. This modified thread form represented two compromises between British Whitworth and United States National thread forms. The British accepted the 60° thread angle of the American thread, and the Americans accepted a rounded root and crest similar to the Whitworth form. The United States National unified form for external threads has a rounded root and may have either a rounded or flat crest. The unified thread contours provide advantages over both the earlier forms. The design provided for greater fatigue strength and permitted longer wear of cutting tools. The unified thread is a general-purpose thread that is particularly suited for fasteners.

Threads on bolts and screws may be formed by rolling, cutting, or extruding and rolling. Cut threads frequently produce a tighter fit than do rolled threads because of irregularities and burrs. To facilitate assembly, cut threads should have a finished point. Rolled threads ordinarily can be manufactured at less cost, but the range of sizes of bolts that can be produced is limited.

The Whitworth 55° thread, similar in many respects to the United States standard form, is the standard V form used in Great Britain. The unified form, identical with the United States standard form, is also standard in Great Britain.

A modified square thread, with an angle of 10° between the sides, is sometimes used for jacks and vises where strength is needed for the transmission of power in the direction of the thread axis.

When originally formulated, the acme thread was intended as an alternate to the square thread and other thread forms that were being used to produce transverse motion on machines. Used for a variety of purposes, the acme thread is easier to produce than the square thread, and its design permits the use of a split nut to provide on-off engagement. A modification of the acme is the 29° stub acme. Because the basic depth of the thread is reduced to 0.30 of the pitch, the thread section is strong and well suited to power applications where space limitations make a shallow thread desirable. A still further modification of the acme is the 60° stub having a 60° angle between the sides of the thread. This thread has a basic depth of 0.433 of the pitch. The Brown and Sharp worm thread, a modified form of the acme thread, is used for transmitting power to a worm wheel.

The buttress or breech-block thread is designed for transmitting an exceptionally high pressure in one direction only. In its original form the pressure flank was perpendicular to the thread axis, and the trailing flank sloped at 45°. To simplify the cutting of the face, modern practice is to give the pressure flank a 7° slope. This form of thread is applicable for assembled tubular members because of the small radial thrust. Buttress threads are used on breech mechanisms of large guns and for airplane propeller hubs.

The knuckle rolled thread is found on sheet-metal shells of lamp bases, fuse plugs, and in electric sockets. The thread form, consisting of circular segments forming crests and roots tangent to each other, has been standardized. The thread is usually rolled, but it may be molded or cast. The Dardelet thread is self-locking in assembly.

SI (ISO) metric thread system. The ISO (International Organization for Standardization) thread form (**Fig. 3**) has essentially the same basic profile as the unified thread and, when the changeover to the metric system has been completed in the United States, it will supplant the unified form. However, before the metric thread system is fully accepted in the United States, minor modifications of thread characteristics may be made to satisfy industry.

The principal differences between the SI metric thread system and the unified inch system are in the basic sizes, the magnitude and application of

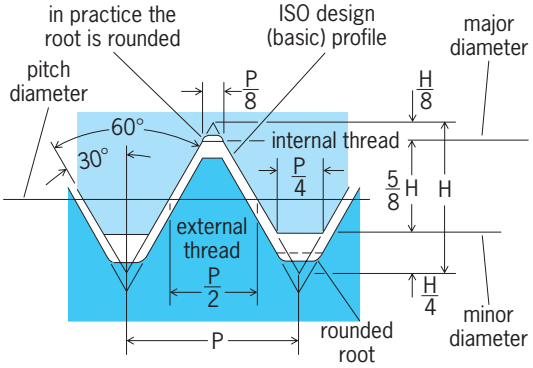

Fig. 3. ISO metric internal and external thread design profiles. (*After W. J. Luzadder, Fundamentals of Engineering Drawing, 8th ed., Prentice-Hall, 1981*)

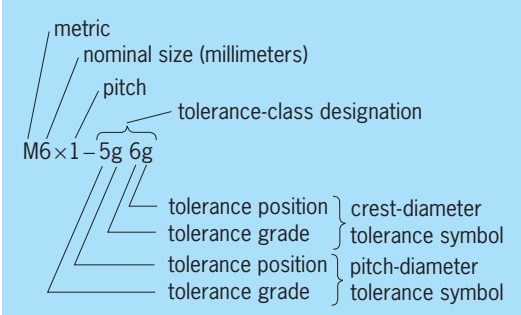

Fig. 4. Metric thread system designations.

allowances and tolerances, and the method for designating and specifying threads.

The basic designation (**Fig. 4**) consists of the letter M followed by the nominal size (basic major diameter in millimeters) and the pitch. The nominal size and pitch are separated by the sign × as in M8 × 1 (ISO designation) and M20 × 1.5.

For coarse series threads, the indication of the pitch may be omitted. The complete designation of an ISO metric thread includes the basic designation followed by the specification for tolerance class. The tolerance class designation includes the symbol for the pitch-diameter tolerance followed by the symbol for the crest-diameter tolerance. Each of these symbols consists of a number indicating the tolerance position. A dash separates the tolerance-class designation from the basic designation. When the pitch-diameter and crest-diameter tolerance symbols are identical, the symbol is given only once.

The ISO standard provides a number of tolerance grades that reflect various magnitudes. Basically, three metric tolerance grades are recommended: 4, 6, and 8. Grade 6 is commonly used for general-purpose threads. This is the closest ISO grade to the unified 2A and 2B fits. The numbers of the tolerance grades reflect the size of the tolerance. Tolerances below grade 6 are smaller and are specified for fine-quality requirements. Tolerances above grade 6 are larger and are recommended for coarse-quality and long lengths of engagement.

Tolerance positions establish the maximum material limits of the pitch and crest diameters for both internal and external threads. The series of tolerance-position symbols to reflect varying amounts of allowance are, for external threads, e = large allowance, g = small allowance, h = no allowance; and for internal threads, G = small allowance, H = no allowance.

A desired fit between mating threads may be specified by giving the internal thread tolerance-class designation followed by the external thread tolerance class. The two designations are separated by a slash, as in M6 × 1 − 6H/6g.

ISO metric thread series are selections of diameter-pitch combinations that are distinguished from each other by the different pitch applied to a specific diameter (M10 × 1.5 coarse series or M10 × 1.25 fine series). Depending upon design requirements, the thread for a particular application might be selected from the coarse-thread series, fine-thread series, or one of the contant-pitch series.

Unified and standard thread series. Unified and American National screw thread standards for screws, bolts, nuts, and other threaded parts consist of six series of threads that cover nonstandard combinations of diameter and pitch. Each series of standard threads has a specific number of threads per inch for a particular diameter. In general, the unified screw threads are limited to three series: coarse (UNC), fine (UNF), and extrafine (UNEF). A 1/4-in.-diameter (6-mm) thread in the UNC series has 20 threads/in., while in the UNF series it has 28.

In the unified and American National standards, the coarse thread series (UNC and NC) is recommended for general industrial use on threaded machine parts and for screws and bolts. Because a coarse thread has fewer threads per inch than a fine thread of the same diameter, the helix angle is greater, and the thread travels farther in one turn. A coarse thread is preferred when rapidity and ease of assembly are desired, and where strength and clamping power are not of prime importance.

The fine thread series (UNF and NF) is recommended for general use in the automotive and aircraft field for threads subject to strong vibration. In general, a fine thread should be used only where close adjustment, extra strength, or increased resistance to loosening is a factor.

The extrafine thread series (UNEF and NEF) is used principally in aeronautical structures where an extremely shallow thread is needed for thin-walled material and where a maximum practicable number of threads is required for a given length.

The 8-thread series (8N) with 8 threads/in. for all diameters is used on bolts for high-pressure flanges, cylinder-head studs, and other fasteners against pressure, where it is required that an initial tension be set up so that the joint will not open when steam or other pressure is applied. This series has come into general use for many types of engineering work. It is sometimes used as a substitute for the coarse thread series for diameters greater than 1 in. (2.5 cm).

The 12-thread series (12UN or 12N) is a uniform pitch series that is widely used in industry for thin nuts on shafts and sleeves. This series is considered to be a continuation of the fine-thread series for diameters greater than $1^1/_2$ in. (3.8 cm).

The 16-thread series (16UN or 16N) is a uniform pitch series for applications that require a very fine thread, as in threaded adjusting collars and bearing retaining nuts. This series is used as a continuation of the extrafine thread series for diameters greater than 2 in. (5 cm).

The manufacturing tolerance and allowance permitted distinguish one class of thread from another under the unified thread system. Because allowance is an intentional difference between correlated size dimensions of mating threads, and tolerance is the difference between limits of size, a thread class may be considered to control the looseness or tightness between mating threaded parts. The American National standard for unified and National screw

threads provides classes 1A, 2A, and 3A for external threads, classes 1B, 2B, and 3B for internal threads, and classes 2 and 3 for both internal and external threads. Classes 1A and 1B are for ordnance and other special uses. Classes 2A and 2B are the recognized standards for the bulk of screw-thread work and for the normal production of threads on screws, bolts, and nuts. Classes 3A and 3B are for applications where smaller tolerances than those afforded by class 2A and 2B are justified and where closeness of fit between mating threads is important. Class 3A has no allowance.

The screw-thread fit needed for a specific application can be obtained by combining suitable classes of thread. For example, a class 2A external thread might be used with a class 3B internal thread to meet particular requirements.

Designation of unified, square, and acme threads. Threads are designated by standardized notes (**Fig. 5**). Under the unified system, threads are specified by giving in order the nominal diameter, number of threads per inch, the initial letters such as UNC or UNF that identify the series, and class of thread (1A, 2A, and 3A or 1B, 2B, and 3B). Threads are considered to be right-hand and single unless otherwise noted; therefore the specification for a left-hand thread must have the letters LH included. To indicate multiplicity, the word DOUBLE, TRIPLE, and so on must follow the class symbol.

Thread fastener uses. Most threaded fasteners are a threaded cylindrical rod with some form of head on one end. Various types of threaded fasteners are available; some, such as bolts, cap screws, and machine screws, have been standardized; others are special designs. The use of removable threaded fasteners is necessary on machines and structures for holding together those parts that must be frequently disassembled and reassembled. Because standard threaded fasteners are mass produced at relatively low cost and are uniform and interchangeable, they are used whenever possible.

Fasteners are identified by names that are descriptive of either their form or application: set screw, shoulder screw, self-tapping screw, thumb screw,

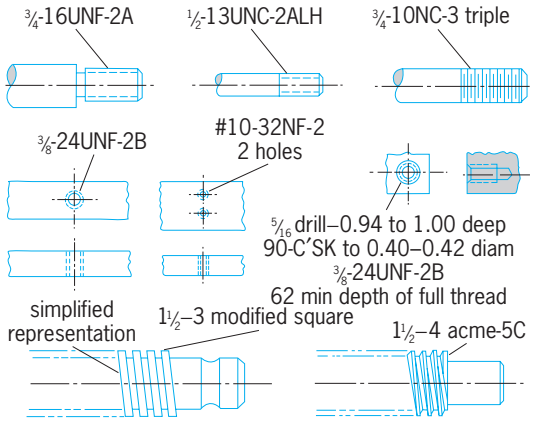

Fig. 5. Thread identification by standardized symbols. (*After W. J. Luzadder, Fundamentals of Engineering Drawing, 8th ed., Prentice-Hall, 1981*)

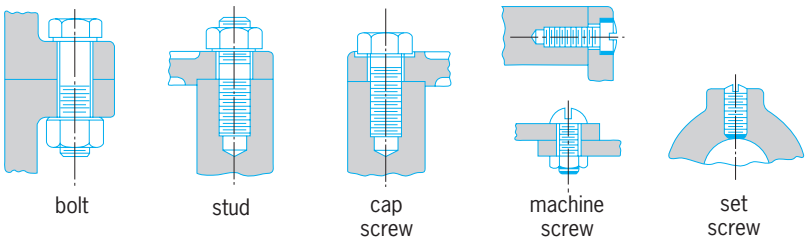

Fig. 6. Common types of fasteners. (*After T. E. French and C. J. Vierck, Engineering Drawing and Graphic Technology, 12th ed., McGraw-Hill, 1978*)

and eyebolt. Of the many forms available, five types meet most requirements for threaded fasteners and are used for the bulk of production work: bolt, stud, cap screw, machine screw, and set screw (**Fig. 6**). Bolts and screws can be obtained with varied heads and points (**Fig. 7**).

Fastener materials. Metallic materials and certain plastics are employed as thread fasteners. Fasteners most commonly used in commercial practice are made of alloys of aluminum, copper (brass and bronze), and steel. For special applications (when required to be heat or corrosion resistant, for example), they are made from a wide variety of materials—metallic or nonmetallic—suitable to the operating environment. Stainless steels, titanium alloys, high-nickel alloys, and nylon are among such materials.

A bolt is generally used for drawing two parts together. Having a threaded end and an integral head formed in manufacture, it is passed through aligned clearance holes in the two parts and a nut is applied.

A stud is a rod threaded on both ends. Studs are used for parts that must be removed frequently and for applications where bolts would be impractical. They are first screwed more or less permanently into one part before the removable member with corresponding clearance holes is placed into position. Nuts are used on the projecting ends to draw the parts together.

Cap screws (bare or plated) are widely used in machine tools and to assemble parts in all types of machinery and equipment. The most common heads are hexagonal, flat, round, fillister, and socket (Fig. 7 and **Fig. 8**). Torque is applied to the fastener by means of a turning device that has a tip designed to mate with a conforming contour or socket in the fastener head; the latter may take the form of a straight slot, hexagonal or fluted socket, crossed Phillips head, or any number of special configurations generally called "tamper proof."

When mating parts are assembled, the cap screws pass through clear holes in one member and screw into threaded holes in the other. The head, an integral part of the fastener along with the thread, holds the parts together. In the automotive industry, the hexagon-head cap screw in combination with a nut is often used as an automotive hexagon-head bolt.

Machine screws, which are similar to cap screws and fulfill the same purpose, are employed principally in the numbered diameter sizes on small work having thin sections or with a machine nut to function as a small bolt.

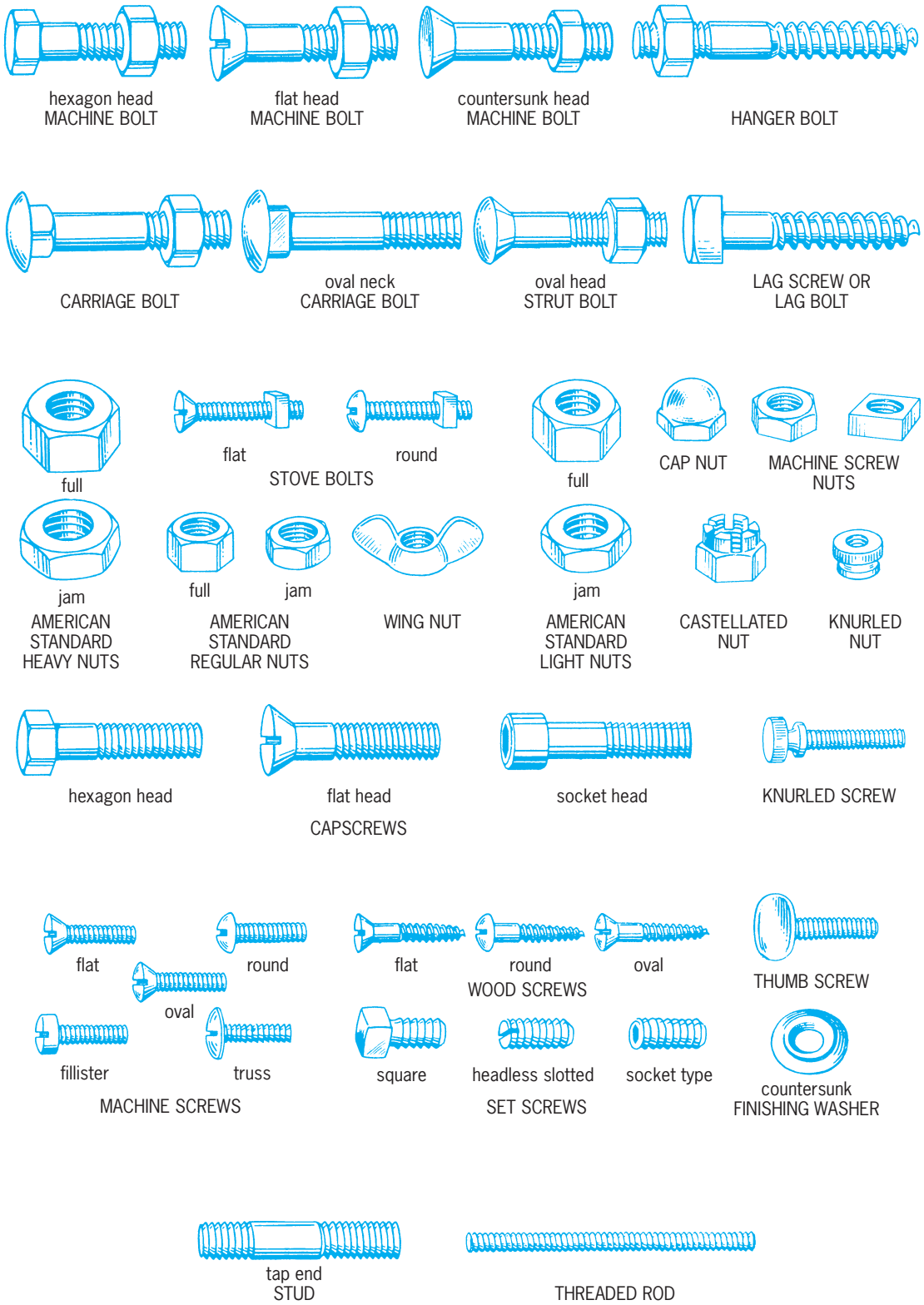

Fig. 7. Threaded fasteners are made in a wide variety of types and used for various purposes. Some of the most common are shown here. Screws may be slotted or hollow-headed. Hollow-headed screws may have a socket or Phillips head or any of a number of special contours for patented turning systems. (*Reynolds Metals Co.*)

Set screws made of hardened steel are used to hold parts in a position relative to one another. Normally, their purpose is to prevent rotary motion between two parts, such as would occur in the case of a rotating pulley and shaft combination. Set screws can be purchased with any one of six types of points in combination with any style of head (Fig. 7 and **Fig. 9**). In the application of set screws, a flat surface is formed on the shaft to provide not only a seat for the screw point, but also clearance space for any burr

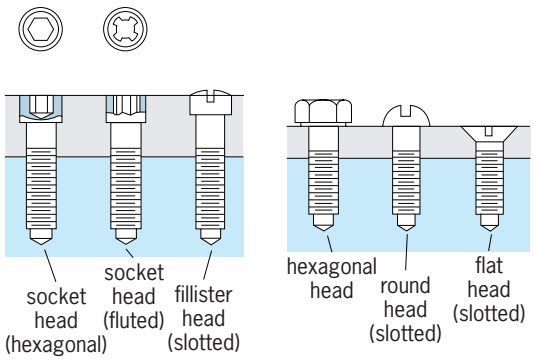

Fig. 8. Cap screws. (*After W. J. Luzadder, Fundamentals of Engineering Drawing, 8th ed., Prentice-Hall, 1981*)

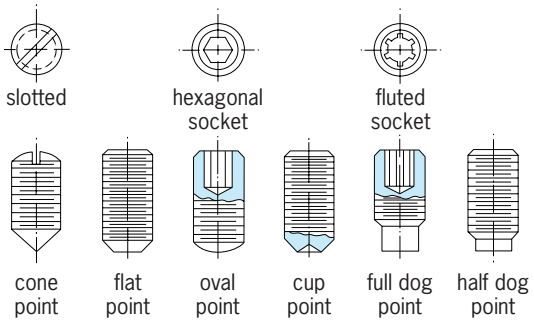

Fig. 9. Set screws. (*After W. J. Luzadder, Fundamentals of Engineering Drawing, 8th ed., Prentice-Hall, 1981*)

which may be raised when the screw is tightened. Without that clearance, the shaft is removed from the assembly with great difficulty, and may result in significant damage to the shaft as well as the bearing.

Wood screws, lag screws, and hanger bolts are used in wood (Fig. 7). Wood screws may have either slotted or recessed heads.

Self-tapping screws (**Fig. 10**) have a specially hardened thread that makes it possible for the screws to form their own internal thread in sheet metal and soft materials when driven into a hole that has been drilled, punched, or punched and reamed. The use of self-tapping screws eliminates costly tapping operations and saves time in assembling parts.

Metric fasteners (ISO). Existing standards for the production of threaded metric fasteners in the United States are in general agreement with ISO standards, with minor differences. When these differences have been resolved, final standards bearing the approval of both ISO and ANSI (American National Standards Institute) will become available for worldwide use. These approved metric fastener standards will assure international interchangeability. Being more restrictive in the choices of sizes and styles than the standards of the past, the new metric standards will make possible reduced inventories and permit bulk pricing advantages.

Metric threaded fasteners are available in only one series of diameter-pitch combinations that range from M1.6 × 0.35 to M100 × 6. The thread pitches are between those of the coarse thread and fine thread series of the present unified (inch) threads. Except for metric socket screws, all threaded fastener products have one tolerance grade, 6g for bolts, machine screws, and so forth, and 6H for nuts. The 6g/6H threads closely match the unified 2A/2B threads for fasteners, in that there is an allowance on the external thread and no allowance on the internal thread.

The standard lengths for short fasteners are 8, 12, 14, 16, and 20 mm. Fastener lengths from 20 to 100 mm increase in length by 5-mm increments. Metric fasteners are specified by giving in sequence the nominal size in millimeters, thread pitch, nominal length, product name (including type of head), material, and protective finish, if needed. Examples are M24 × 3 × 50 stainless steel hex bolt; M14 × 2 × 80 hex cap screw cadmium plated; M8 × 1.25 × 14 hex socket flat point set screw.

Pipe thread. Standard (Briggs) taper pipe threads (**Fig. 11**) are cut (on a cone) to a taper of $^1/_{16}$ in./in. to ensure a tight joint. Although a normal connection employs a taper external and a taper internal thread, an American National standard straight pipe thread having the same pitch angle and depth of thread as the corresponding taper pipe thread is used for pressure-tight joints for couplings, pressure-tight joints for fuel- and oil-line fittings, and for loose-fitting and free-fitting mechanical joints. Assemblies made with taper external threads and straight internal threads are frequently preferred to assemblies employing all taper threads; the assumption is made that relatively soft or ductile metals will adjust to the taper external pipe thread. A modified pipe thread, the United States standard Dryseal pipe thread (taper

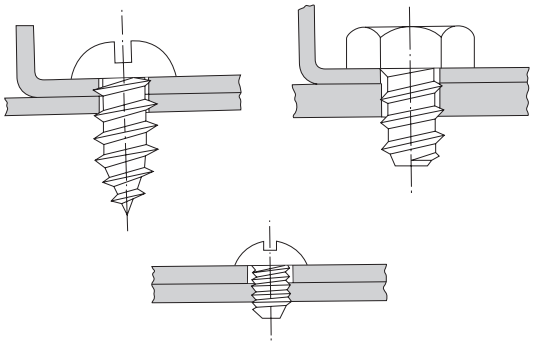

Fig. 10. Three applications of self-tapping screws.

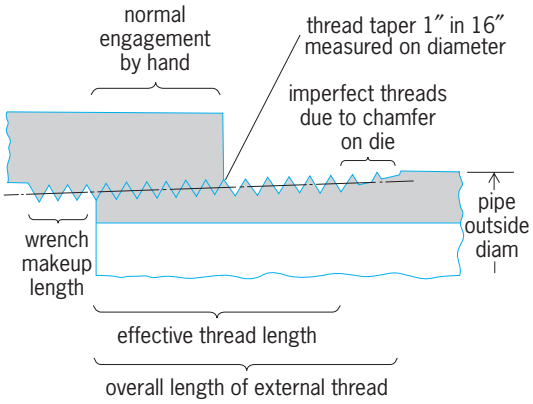

Fig. 11. Standard (Briggs) taper pipe thread.

and internal straight), is used for pressure-tight connections that are to be assembled without lubricant or sealer. Except for a difference in the truncation at the roots and crests, the general form and dimensions of this thread are the same as those of the standard taper pipe thread. The principal uses for the Dryseal thread are in refrigerant pipes, automotive and aircraft fuel-line fittings, and for gas and chemical shells (ordnance). For railing joints, where a rigid mechanical thread joint is needed, external and internal taper threads are used. Basically, the external thread is the same as the standard taper pipe thread, except that the length of the thread is shortened at the end to permit use of the larger-diameter portion of the thread. Warren J. Luzadder

Bibliography. T. E. French, C. J. Vierck, and R. J. Foster, *Engineering Drawing and Graphic Technology*, 14th ed., 1993; Industrial Fasteners Institute, *Metric Fastener Standards*, 1986; W. J. Luzasser and J. M. Duff, *Fundamentals of Engineering Drawing*, 11th ed., 1992; R. O. Parmley, *Standard Handbook of Fastening and Joining*, 3d ed., 1997; J. E. Shigley and C. R. Mischke, *Fastening, Joining, and Connecting: A Mechanical Designer's Handbook*, 1989.

Scrophulariales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Asteridae of the class Magnoliopsida (dicotyledons). The order consists of 12 families and more than 11,000 species. The largest families are the Scrophulariaceae (about 4000 species), Acanthaceae (about 2500 species), Gesneriaceae (about 2500 species), Bignoniaceae (about 800 species), and Oleaceae (about 600 species). The Scrophulariales are Asteridae, with a usually superior ovary and generally either with an irregular corolla or with fewer stamens than corolla lobes, or commonly both. They uniformly lack stipules.

The Scrophulariaceae are characterized by their usually herbaceous habit; irregular flowers with usually only two or four functional stamens; axile placentation; and dry, dehiscent fruits with more or less numerous seeds that have endosperm. Such well-known ornamentals as snapdragon (*Antirrhinum majus*; see **illus.**), foxglove (*Digitalis purpurea*), and calceolaria belong to the Scrophulariaceae.

The Acanthaceae are distinguished by their chiefly herbaceous habit, mostly simple and opposite leaves, lack of endosperm, and especially the enlarged and specialized funiculus, which is commonly developed into a jaculator that expels the seeds at maturity. The family includes several fairly common ornamentals, such as species of *Acanthus*, *Aphelandra*, *Beloperone* (shrimp plant), and *Strobilanthes*.

Some other well-known members of the Scrophulariales are the African violet (*Saintpaulia ionantha*, in the Gesneriaceae), catalpa (Bignoniaceae), and lilac (*Syringa vulgaris*, in the Oleaceae). *See* ASTERIDAE; FLOWER; MAGNOLIOPHYTA; MAGNOLIOPSIDA; OLIVE; ORNAMENTAL PLANTS; PLANT KINGDOM.

Arthur Cronquist; T. M. Barkley

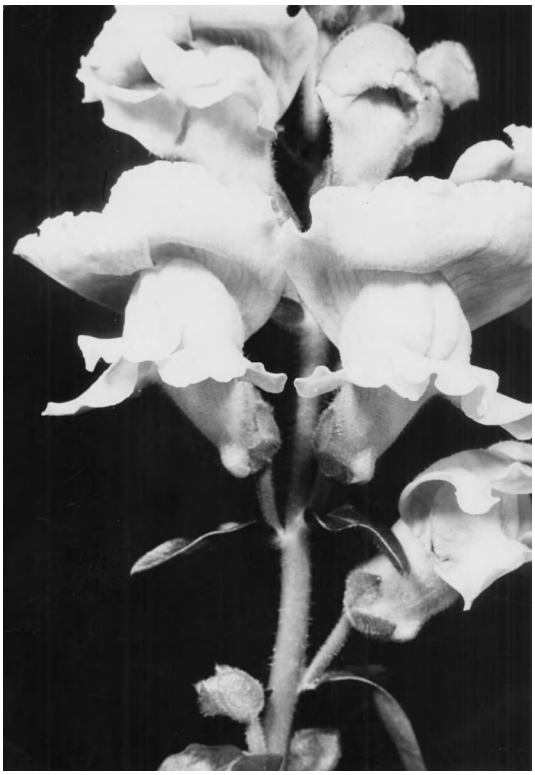

Flowers of the common snapdragon (*Antirrhinum majus*), a typical member of the family Scrophulariaceae and order Scrophulariales. (*J. Douglas Spillane, National Audubon Society*)

Scyphozoa

A class of the phylum Cnidaria traditionally containing four major groups—Coronatae, Rhizostomeae, Semaeostomeae, and Stauromedusae—all marine or estuarine. Most occur in two forms, the polyp, or scyphistoma, and the medusa, or scyphomedusa. However, some (the Stauromedusae) are polyplike and sessile throughout their lives, while others (such as the semaeostomes *Pelagia* and *Stellamedusa* and the coronate *Periphylla periphylla*) are always pelagic and lack the sessile scyphistoma stage. Among the Cnidaria, the Scyphozoa are characterized by having well-developed medusae of large size and fairly well organized scyphistomae of small size. The Scyphozoa are most closely related to the other medusae-containing classes—Cubozoa and Hydrozoa—within the Cnidaria, all of which evolved from a polyp-only ancestor currently represented by members of Class Anthozoa. Recent analyses of DNA and morphology indicate that the Stauromedusae group may represent an intermediate evolutionary stage between Class Anthozoa and Medusozoa (which would encompass Cubozoa, Hydrozoa, Coronatae, Rhizostomeae, and Semaeostomeae) and may, in fact, represent a distinct class of cnidarian (Class Staurozoa).

Morphology. Scyphistomae are usually solitary and small, measuring much less than an inch (only a

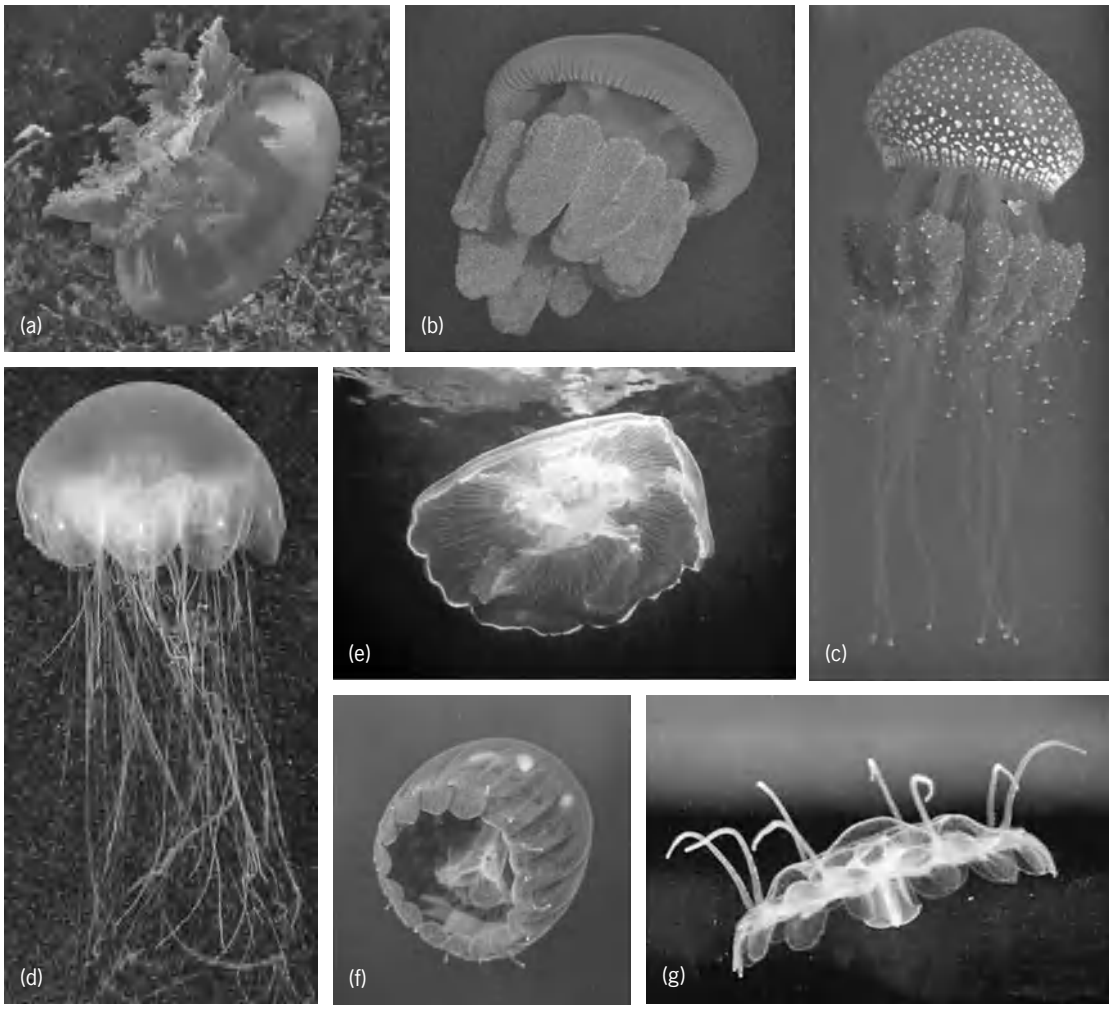

Fig. 1. Some common genera of scyphomedusae. Order Rhizostomeae. (*a*) *Cassiopea* (*photo by L. E. Martin*), (*b*) *Catostylus* (*photo by K. Pitt*), (*c*) *Phyllorhiza* (*photo by K. Pitt*). Order Semaeostomeae: (*d*) *Cyanea* (*photo by G. Edgar, reproduced with permission from Edgar, 2000*), (*e*) *Aurelia*. Order Coronatae: (*f*) *Linuche* (*photo by L. E. Martin*), (*g*) *Nausithoe* (*photo by A. Migotto*).

few millimeters) in height. A minority (including *Stephanoscyphus*, the polyp of *Nausithoe*) form colonies up to 4 to 10 in. (100–250 mm) high. They are sessile. Each solitary scyphistoma is composed of three major parts—pedal disk, stalk, and calyx—and is shaped, overall, something like a wine glass or vase. The narrow pedal disk has many gland cells that secrete a sticky fluid which attaches the scyphistoma to hard substrate. The stalk, along which run fine muscular strands enabling the polyp to contract, leads away from the substrate, then broadens into the calyx. The calyx is crowned by the oral disk with a circlet of tentacles around its perimeter and a four-sided mouth at its center. The mouth opens into the stomodeum, which leads to the stomach. The axes running from the four corners of the mouth are called perradii, and halfway between the perradii are the interraddii. At each interradius of the polyp there is an infundibulum, a canal which leads down into the stalk. The polypoid Stauromedusae are peculiar in having many of the features (for example, rhopalia, or sensory structures, and gonads) that are found only in the free-living medusae of Semaeostomeae, Rhizostomeae, and Coronatae.

Scyphomedusae, when fully grown, are generally large, sometimes attaining a diameter of 3 ft (1 m) and weighing over 300 lb (140 kg). Most scyphomedusae are natatory (adapted for swimming). The scyphomedusa is composed of the umbrella (or bell), which is the major propulsive structure, generally appended with oral arms. Differences in the shape, structure, ornamentation, and color of the bell and oral arms are the main features used to distinguish major groups and species of Scyphozoa (**Figs. 1** and **2**). The convex upper surface of the umbrella is called the exumbrella; the concave inner surface, the subumbrella. Muscles are distributed mainly in the subumbrella and the most important muscles are the ring muscles, which contract the umbrella for locomotion of the medusa. The umbrella may have various shapes, being disk-, cone-, or domelike. In the Coronatae, the central portion is partially separated from the skirt-like peripheral regions (the bell margin) by the coronal groove, from which the order gets its name. In all groups of scyphomedusae the margin of the umbrella is divided into many lappets (Fig. 2). Rhopalia, of

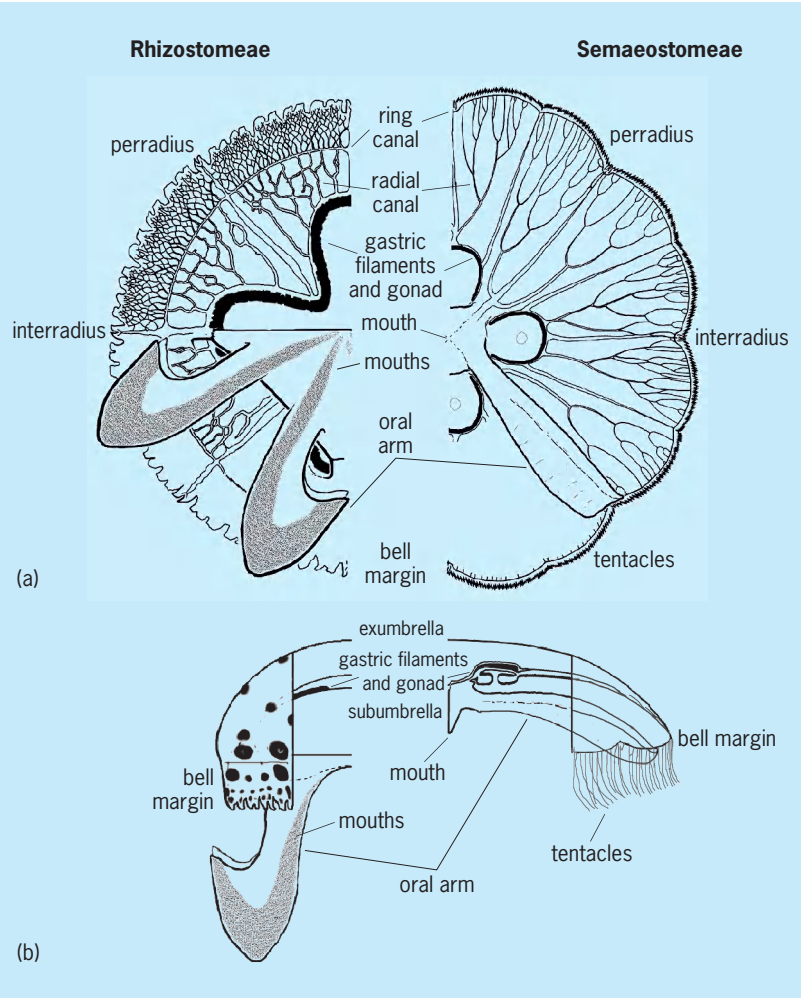

Fig. 2. Some of the major anatomical features of rhizostome (left) and semaeostome (right) medusae. (*a*) View of the subumbrella emphasizing features of the oral arms (lower portions) and the bell (cutaway, upper portions). (*b*) Side view of external and, with cutaway, internal features. The shape and size of these and other anatomical features can very considerably and are used to distinguish among taxa from class-level down to species.

which there are usually eight, are found between the lappets. Each rhopalium is composed of a statocyst or vesicle, containing a round inorganic concretion, and an ocellus, a structure composed of a lens, retina cells, and pigment cells. Tentacles also are found on the bell margin, or marginal regions of the subumbrella, and vary in number from eight to several hundred with the exception of the Rhizostomeae and some deep-water semaeostomes (for example *Deepstaria, Stellamedusa*, and *Tiburonia*) which have no tentacles. Tentacles, when present, have numerous cnidocytes, or "nettle cells," and longitudinal muscles which enable the tentacles to be extended and contracted to catch prey. The mouth, situated in the center of the subumbrella, is surrounded by four simple lips in some forms (such as Coronatae), but generally there are four well-developed oral arms (eight in Rhizostomeae). The oral arms may be used to catch and manipulate prey, to pass food from the tentacles to the mouth, and/or to brood larvae. The oral arms are most obvious in the Rhizostomeae, often being as substantial as the umbrella. In Rhizostomeae, the oral arms provide a large surface area covered by many small suctorial mouths that lead by way of canals to the gut, thus replacing the single larger central mouth adjacent to the stomach in semaeostome and coronate scyphomedusae. The stomach is often large and composed of four interradial pouches forming a cross-shape within the bell, particularly in the Rhizostomeae. Each pouch contains several rows of gastric filaments, involved in extracellular digestion, as well as the gonads.

Development and reproduction. The gonads develop from endoderm cells just above the gastric filaments. Medusae are generally gonochoristic, that is, have separate male and female individuals. The ripe ova, which are liberated from the broken gastrodermal wall of the ovary, enter the stomach cavity and then are discharged from the mouth into seawater. Fertilization may occur in the female gonad, in the stomach cavity, on the oral arms, or in open seawater. Cleavage is mainly radial, and the new endoderm is formed by ingression or epiboly. The egg develops into a ciliated planula that swims freely, except in Stauromedusae, in which the planula lacks cilia and creeps on the substratum. The planula attaches to seaweed or rocks and changes into a scyphistoma, which includes adding tentacles.

Scyphistomae reproduce asexually by various modes, including budding and pedal cysts, giving rise to other autonomous scyphistomae (or members of a colony). Under appropriate conditions, a scyphistoma will metamorphose into a strobila and, via an asexual process known as strobilation, which involves transverse differentiation and fission of the polyp, produce nascent scyphomedusae called ephyrae. An ephyra undergoes further metamorphosis into a young medusa. Thus most scyphozoans have a bipartite (two-stage) life history, alternating between sessile polyp and free-living medusa forms (**Fig. 3**). There are, however, exceptions. For example, all stauromedusae combine many features of the medusa into a polypoid stage (they have no free-living medusa phase); and the planulae of the semaeostome genus, *Pelagia*, and coronate species *Periphylla periphylla* develop directly into ephyrae.

Physiology. The major sensory structures are the rhopalia; ocelli permit medusae to sense and respond to changes in light intensity, and statocysts enable them to swim oriented relative to a vertical axis defined by the direction of gravity. The medusae may also respond to pressure, which changes with depth. In contrast, scyphistomae, which lack rhopalia, are largely indifferent to light.

Scyphomedusae and scyphistomae are sensitive to touch, often contracting the body (which alters the swimming direction of medusae) or tentacles. Chemical senses enable scyphozoans to respond to the presence of, swim toward, and catch food, although they have no structures dedicated to gustatory or olfactory functions. They are sensitive to organic acids, several fatty acids, and lipids but not to carbohydrates.

The Scyphozoa are ectothermic and their metabolic rate is influenced strongly by changes in the temperature of their environment. This is most evident in the rate of muscular contractions (pulses) of the bell, which slow dramatically when scyphomedusae are placed in colder water. However, species of scyphomedusae that live naturally at different temperatures may pulse at similar rates, indicating physiological adaptations to different environments. Scyphozoans are osmoconformers (that is, their body fluids change directly with a change in the concentrations of dissolved ions in their environment).

Biodiversity and biogeography. Important differences in morphology, DNA, and proteins used to identify jellyfish indicate that there are several hundreds of species of Scyphozoa. Most are typically found near the coast because scyphistomae are benthic. Exceptions are *Pelagia* and other wholly pelagic forms, including many of the mid- and deep-water scyphomedusae.

Order Rhizostomeae contains the most species; about 95 are currently recognized. Rhizostomes are particularly interesting because they are the most recently derived and also the most diverse scyphozoans, indicating comparatively recent rapid diversification. The rhizostomes are predominantly found in tropical regions, especially the Indo-West Pacific center of marine biodiversity. Order Semaeostomeae contains fewer species (approximately 75) which are more widely distributed, with representatives in polar, deep-sea, open-ocean, tropical, and temperate coastal regions. Coronatae is the least diverse of the three groups of scyphomedusae, with only approximately 50 species recognized, and is typically regarded as a deep-sea group although it does have a number of shallow-water representatives in both the tropics and temperate regions. There are only approximately 35 known species of stauromedusae, which are distributed predominantly in cold temperate and polar regions of northern and southern hemispheres.

Applied ecology. Scyphozoa are important members of marine ecosystems. Mostly thought of as predators (of zooplankton and sometimes fish), they also absorb dissolved organic matter and form symbioses with zooxanthellae. The number and size of medusae can increase very rapidly, a phenomenon known as a "jellyfish bloom." Blooms can disrupt human activities—blocking seawater intakes of desalination and power plants, overloading and tearing fishing trawls, and outcompeting or eating larvae of commercial fishes—and have therefore attracted considerable attention. Although rapid change in population size is a natural consequence of the jellyfish life-cycle, the concern (which there is little evidence for or against at this time) is that blooms are becoming more frequent and more intense possibly as a result of environmental degradation or invasive species. Some Semaeostomeae and Rhizostomeae are injurious to humans because of their cnidocytes. On the other hand, some medusae protect young fish and crustaceans which seek shelter among their oral arms. Some Rhizostomeae are eaten by several cultures of southeast Asia. *See* CNIDARIA.

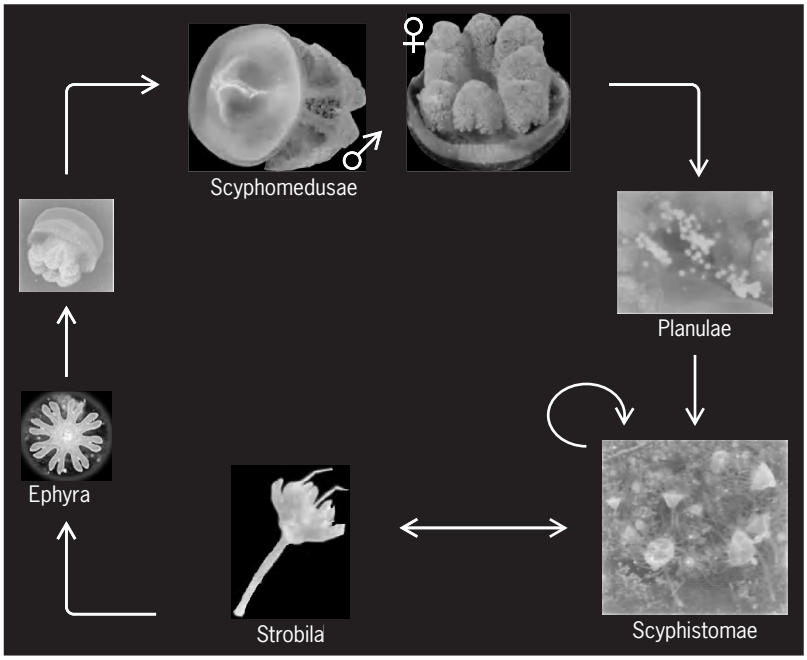

Fig. 3. Typical life history of species of Scyphozoa. Life stages not drawn to scale.

Michael N. Dawson

Bibliography. M. N. Arai, *A Functional Biology of Scyphozoa*, Chapman and Hall, London, 1997; T. Heeger, *Quallen: Gefährliche Schönheiten*, Wissenschaftliche, Stuttgart, 1998; C. E. Mills, Jellyfish blooms: Are populations increasing globally in response to changing ocean conditions?, *Hydrobiologia*, 451:55–68, 2001; F. S. Russell, *The Medusae of the British Isles. II Pelagic Scyphozoa with a Supplement to the First Volume on Hydromedusae*, Cambridge University Press, 1970; D. Wrobel and C. Mills, *Pacific Coast Pelagic Invertebrates—A Guide to the Common Gelatinous Animals*, Sea Challengers and the Monterey Bay Aquarium, 1998.

Sea anemone

Any polyp of the nearly 1000 anthozoan coelenterates belonging to the order Actiniaria. They occur intertidally and subtidally in marine and estuarine habitats, attached to solid substrates or burrowing into soft sediments. No fresh-water or truly planktonic species are known, though one genus, *Minyas*, secretes a float within its base and hangs suspended at the surface of the seas. Anemones may be very small, 0.04 in. (1 mm) long, to large individuals more than 3 ft (1 m) in length or diameter. *See* ACTINIARIA.

Each polyp is cylindrical with an adhesive base or a burrowing structure (physa) at its aboral end. The mouth occurs at the free end of the hollow body (column) and is generally surrounded by simple, unbranched tentacles (see **illus.**). Many species attach foreign material to their columns with

Sea anemone, a polyp belonging to the order Actiniaria.

varied adhesive structures. Food (fish, worms, mollusks, crustaceans, and dead organisms) is captured by the tentacles. Since sea anemones are sedentary, they depend on currents and wave action to carry their food to them.

Most anemones are capable of only limited locomotion. Attached forms creep on their basal disk, and a few move by somersaulting. Burrowing forms move through the substrate with the action of the physa. A few unrelated species swim by either writhing the body or beating the tentacles. Swimming is an escape response usually initiated by the attack of a predator such as a starfish or nudibranch.

Sea anemones carry out such movements as bending, elongating, and shortening by using muscles whose antagonist is the water within the body cavity and hollow tentacles. Water enters the body via special ciliated grooves in the actinopharynx, the siphonoglyphs. With all muscles relaxed, the cilia can create sufficient water pressure to reinflate the anemone. Elongation results from the contraction of circular muscles. Bending and shortening are caused by contraction of longitudinal muscles of the body wall, tentacular walls, and mesenteries. During normal movements, the volume of water contained by an anemone is held constant while the shape changes. The contained water is referred to as a hydrostatic skeleton. Water is released through the mouth and pores in the tips of the tentacles or body wall when the anemone is disturbed and contracts fully.

Sea anemones enter into a number of interesting symbiotic partnerships. Some are host to single-celled marine algae, which grow within the cells of the anemone. The algae provide organic materials which aid in the nutrition of the anemone. Other anemones live on gastropod shells inhabited by hermit crabs. Such anemones have special behavior which leads them to climb on gastropod shells, usually by somersaulting. Some hermit crabs can induce this behavior by stroking the anemone. Certain true crabs collect anemones and place them on their carapace and legs. With both crabs and hermit crabs, the presence of anemones tends to protect them from such predators as octopuses or other crabs. In the tropics, some anemones play host to anemone fishes which live around and among the tentacles. These fish, using a special behavior, cover themselves with mucus from the anemone and so are chemically disguised from the anemone. If the mucus is removed, the fish will be attacked by nematocysts and eaten by the host. Some anemones are very specific in their choice of substrate, occurring only on a certain species of seaweed or gorgonian. *See* ANTHOZOA; COELENTERATA. Cadet Hand

Sea breeze

A diurnal, thermally driven circulation in which a surface convergence zone often exists between airstreams having over-water versus over-land histories. The sea breeze is one of the most frequently occurring small-scale (mesoscale) weather systems. It results from the unequal sensible heat flux of the lower atmosphere over adjacent solar-heated land and water masses. Owing to the large thermal inertia of a water body, during daytime the air temperature changes little over the water while over land the air mass warms. Occurring during periods of fair skies and generally weak large-scale winds, the sea breeze is recognizable by a wind shift to onshore, generally several hours after sunrise. On many tropical coastlines the sea breeze is an almost daily occurrence. It also occurs with regularity during the warm season along midlatitude coastlines and even occasionally on Arctic shores. Especially during periods of very light winds, similar though sometimes weaker wind systems occur over the shores of larger lakes and even wide rivers and estuaries (lake breezes, river breezes). At night, colder air from the land often will move offshore as a land breeze. Typically the land breeze circulation is much weaker and shallower than its daytime counterpart.

Of all mesoscale phenomena, the sea breeze has been among the most studied, undoubtedly a result of its geographically fixed nature, its frequent occurrence, its ease of recognition using conventional observations, the concentration of observers in coastal zones, and its importance to local weather and climate. The sea breeze, among the simpler of the atmospheric circulation systems, has been frequently simulated by numerical models, expanding knowledge into areas where observations are lacking.

Structure. The occurrence and strength of the sea breeze is controlled by a variety of factors, including land-sea surface temperature differences; latitude and day of the year; the synoptic wind and its orientation with respect to the shoreline; the thermal stability of the lower atmosphere; surface solar radiation as affected by haze, smoke, and stratiform and convective cloudiness; and the geometry of the shoreline and the complexity of the surrounding terrain.

The classic sea breeze is represented as a closed circulation with a well-defined onshore flow layer, a distinct sea breeze front, (which can be very sharply defined, on the order of hundreds of meters in some cases), and a more diffuse return flow layer aloft in which air returns over the water and subsides in a

broad zone offshore. Sea breeze inflow layer depths vary from about 100 m (330 ft) to over 1000 m (3300 ft). Inland frontal penetration can vary from less than 1 km (0.6 mi) to over 100 km (60 mi), with propagation speeds ranging from nearly stationary to greater than 5 m/s (10 mi/h).

The offshore extent of the circulation is less well known, though it typically extends about twice as far offshore as the marine air has penetrated inland. Peak wind speeds are usually less than 10 m/s (20 mi/h). The overlying return flow layer depth is generally twice that of the inflow but is sometimes difficult to differentiate from the synoptic flow. Observational evidence has suggested that organized frontal zone upward motions can be as strong as 2–3 m/s (about 4–6 mi/h). Much less is known about the broader and weaker subsidence regions offshore. The initial sea breeze onshore flow is often nearly normal to the coastline, but because of the Coriolis effect the wind direction often tends to become nearly parallel to the shore by late afternoon. *See* CORIOLIS ACCELERATION; FRONT; WIND.

Impacts. The effects of the sea breeze are many, including significantly altering the direction and speed of the coastal zone winds; influencing low-level stratiform and cumuliform clouds; initiating, suppressing, and modifying precipitating convective storms; recirculating and trapping pollutants released into or becoming entrained into the circulation; perturbing regional mixing depths; and creating strong nearshore temperature, moisture, and refractive index gradients.

Cumulus suppression in the subsidence regions of the sea breeze cell offshore has been noted in satellite imagery. Especially in the tropics the sea breeze has been found to trigger the general development of thunderstorms. Sea breeze thunderstorms contribute approximately 40% of Florida's rainfall and probably more in many other coastal regions. Sea breeze circulations are major factors in exacerbating photochemical ozone pollution in cities as diverse as Los Angeles, Houston, Chicago, Oslo, Athens, and Tokyo. *See* ATMOSPHERIC GENERAL CIRCULATION; MESOMETEOROLOGY; METEOROLOGY; THUNDERSTORM.

Walter A. Lyons

Bibliography. B. W. Atkinson, *Meso-scale Atmospheric Circulations*, 1981; S. A. Hsu, *Coastal Meteorology*; 1988; T. F. Malone (ed.), *Compendium of Meteorology*, 1951; R. A. Pielke, *Mesoscale Meteorological Modeling*, 2d ed., 2001.

Sea-floor imaging

The process whereby mapping technologies are used to produce highly detailed images of the sea floor. Ever since large ships began sailing the oceans, there has been a need to know what the sea floor looks like. At first this need was simply to avoid dangerous shoals upon which a ship could wreck. In the mid-1800s, the need increased with the first laying of transocean communication cables. World War II and the initiation of submarine warfare elevated the need to a matter of national defense, and mapping of the world's ocean floors was greatly accelerated. The vast body of information accrued by this effort went far beyond aiding militaries. It also led to the theory of plate tectonics, the backbone of modern geology that explains volcanoes, earthquakes, mountain chains, and many other Earth features. At present, sea-floor mapping is also aiding a variety of societal endeavors. High-resolution images of the sea floor are being used to locate and manage marine resources such as fisheries and oil and gas reserves, identify offshore faults and the potential for coastal damage due to earthquakes, and map out and monitor marine pollution. The greatly improved sea-floor imaging is providing information on what processes are affecting the sea floor, where these processes occur, and how they interact. *See* MARINE GEOLOGY.

Mapping technologies. New, detailed images of the sea floor are revealing terrains as remarkable as those being radioed back to Earth by planetary probes. Modern technologies have revolutionized sea-floor mapping. The most commonly used are the four technologies described below, each of which provides a unique viewpoint of the sea floor.

Satellite altimetry. The technology that provides the broadest perspective but the lowest resolution is satellite altimetry (**Fig. 1***a*). A laser altimeter is mounted on a satellite and, in combination with land-based radars that track the satellite's altitude, is used to measure variations in sea-surface elevation to within 2 in. (5 cm). Removing elevation changes due to waves and currents, sea-surface height can vary up to 660 ft (200 m). These variations are caused by minute differences in the Earth's gravity field, which in turn result from heterogeneities in the Earth's mass. These heterogeneities are often associated with sea-floor topography. For example, the gravitational attraction of a massive undersea volcano pulls water toward it, producing a local bulge in the ocean surface. If the volcano is 6600 ft (2000 m) high and has an average radius of 12.5 mi (20 km), the sea surface above will typically form a bulge 6.6 ft (2 m) high. By using a mathematical function that equates sea-surface height to bottom elevations, global areas of the sea floor can be mapped within a matter of weeks. However, this approach has limitations. Sea-floor features less than 6–9 mi (10–15 km) in length are generally not massive enough to deflect the ocean surface, and thus go undetected. Furthermore, sea-floor density also affects the gravity field; and where different-density rocks are found, such as along the margins of continents, the correlation between Earth's gravity field and sea-floor topography breaks down. *See* ALTIMETER.

Side-scan sonar. This provides a higher-resolution view of the sea floor, but one which is a bit different from other technologies. In general, a side-scan sonar consists of two sonar units attached to the sides of a sled tethered to the back of a ship (Fig. 1*b*). Each sonar emits a burst of sound that insonifies a long, narrow corridor of the sea floor extending away from the sled. Sound reflections from the corridor that

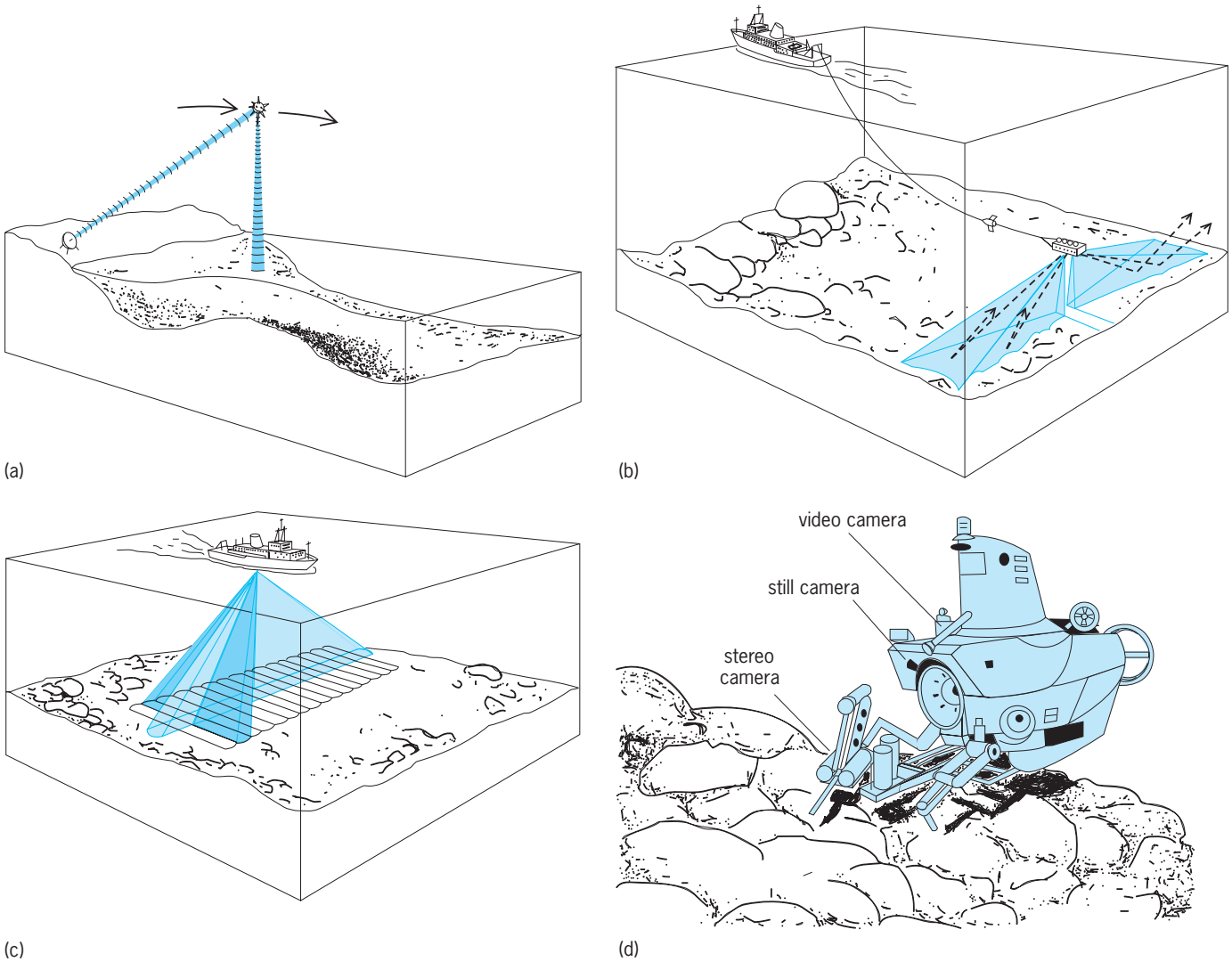

Fig. 1. Technologies that have revolutionized sea-floor mapping: (*a*) satellite altimetry; (*b*) side-scan sonar; (*c*) multibeam sonar; (*d*) submersibles. *(Courtesy of Roberto Osti)*

echo back to the sled are then recorded by the sonar in their arrival sequence, with echoes from points farther away arriving successively later. The sonars repeat this sequence of "talking" and listening every few seconds as the sled is pulled through the water so that consecutive recordings build up a continuous swath of sea-floor reflections, which provide information about the texture of the sea floor. For example, if the sea floor is flat and smooth, none of the sound from the sonars is reflected back to the instrument, just as a mirror does not reflect back light shown on it at an angle. But if the sea floor has a roughness, the sound will be reflected in multiple directions; and similar to light reflecting off broken glass, some of the reflections return to the sonar. By equating the amplitude of the returning echoes to different levels of gray, and the time that the echoes arrive back at the sled to distance, an image is obtained of sea-floor textures that looks like a black and white photograph. The textures in the image reflect a combination of sea-floor morphology and bottom sediment type; but like a single photograph, they cannot be used to determine sea-floor depths. *See* ECHO SOUNDER; SONAR; UNDERWATER SOUND.

Multibeam sonar. The best technology for mapping sea-floor depths or bathymetry is multibeam sonar (Fig. 1*c*). These systems employ a series of sound sources and listening devices that are mounted on the hull of a survey ship. As with side-scan sonar, every few seconds the sound sources emit a burst that insonifies a long, slim strip of the sea floor aligned perpendicular to the ship's direction. The listening devices then begin recording sounds from within a fan of narrow sea-floor corridors that are aligned parallel to the ship and that cross the insonified strip. Thus, sound reflections received at the listening devices are only from relatively small sea-floor regions where the insonified strip and listening corridors intersect. The timing of these reflections yields a profile of sea-floor depth, and successive profiles collected as the ship moves build up a continuous swath of sea-floor bathymetry along the ship's track. By running the survey the same way that one mows

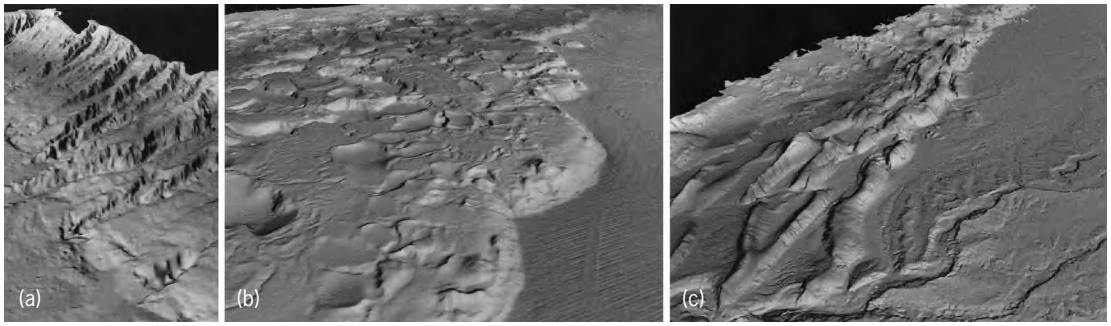

Fig. 2. In the original images, colors are used to portray sea-floor depth. Here, white represents the shallowest depths and dark gray the deepest; land is represented by black. (*a*) New Jersey continental slope. (*b*) Louisiana continental slope. (*c*) Oregon continental slope.

a lawn, adjacent swaths are collected parallel to one another to produce a complete sea-floor map of an area. *See* ACOUSTIC SIGNAL PROCESSING.

Direct visual imaging. The most modern swath mapping systems now collect both bathymetry and side-scan sonar imagery. But the most accurate and detailed view of the sea floor is provided by direct visual imaging through bottom cameras, submersibles (Fig. 1*d*), remotely operated vehicles, or if the waters are not too deep, scuba diving. Of these, bottom cameras were the first widely used means of getting an upclose view of the sea floor. However, because light is scattered and absorbed in waters greater than about 33 ft (10 m) deep, the sea-floor area that bottom cameras can image is no more than a few meters. This limitation has been partly overcome by deep-sea submersibles and remotely operated vehicles, which provide researchers with the opportunity to explore the sea floor close-up for hours to weeks at a time. But even the sea-floor coverage that can be achieved with these devices is greatly restricted relative to side-scan sonar, multibeam sonar, and satellite altimetry. Hence, with each technology there is a trade-off in sea-floor coverage with resolution, and which technology is used depends on what information is needed.

Sea-floor panoramas. Diverse sea-floor terrains, covering some 70% of the Earth's surface, are concealed from view by the oceans. Images have been produced that show the ocean stripped away and the sea floor laid bare. This has been accomplished with the use of computers by mathematically fitting a continuous surface to numerous, closely spaced, high-resolution depth measurements to render a detailed model of what the sea floor in various regions looks like (**Fig. 2**).

A wide range of submarine terrains border the continental United States. Continental slopes are where the sea floor drops at a relatively steep angle (on average, 2–6°) from the edge of the shallow shelf plateaus (on average, <400 ft or 120 m water depth) that rim the continents down to the gradually dipping (≤0.5°) rise and abyssal plains that underlie the deep sea (generally, >2 mi or 3000 m water depth). The continental slope off New Jersey (Fig. 2*a*) is scarred by numerous submarine canyons. A primary agent in cutting the canyons has been subsea avalanches of shelf-edge and slope sediments that became unstable and failed.

On the continental slope off Louisiana (Fig. 2*b*), the movement of buried salt has deformed the sea floor to create a lunarlike surface. The salt was deposited earlier in geologic time when the Gulf of Mexico was first forming. It became buried by the massive amounts of sediments eroded from North America and dumped into the Gulf of Mexico by the Mississippi River. Under the great weight of these sediments, pockets of salt are being squeezed upward and seaward, rearranging the seascape above them like groundhogs burrowing through soil.

On the continental slope off Oregon (Fig. 2*c*), the North American plate, which is one of the 12 lithospheric plates forming the Earth's outer shell and in which the United States is embedded, is overriding the Juan de Fuca plate, which lies beneath flat sea-floor sediments extending from the base of the slope. As the plates collide, the overthrusting North American plate is bulldozing the sediments off the downgoing Juan de Fuca plate and folding them into the ridges that run parallel to the continental slope.

Lincoln F. Pratson

Bibliography. L. F. Pratson and W. F. Haxby, Panoramas of the seafloor, *Sci. Amer.*, 276(6):66–71, 1997; L. F. Pratson and W. F. Haxby, What is the slope of the U.S. continental slope?, *Geology*, 24(1):3–6, 1996; D. T. Sandwell, Geophysical applications of satellite altimetry, *Rev. Geophys. Suppl.*, pp. 132–137, 1990; R. Vogt and B. E. Tucholke (eds.), *Imaging the Ocean Floor: History and State of the Art—Geology of North America: The Western North Atlantic Region*, 1986.

Sea ice

Ice formed by the freezing of seawater. Ice in the sea includes sea ice, river ice, and land ice. Land ice is principally icebergs which are prominent in some areas, such as the Ross Sea and Baffin Bay. River ice is carried into the sea during spring breakup and is important only near river mouths. The greatest part, probably 99% of ice in the sea, is sea ice. *See* ICEBERG.

Properties. The freezing point temperature and the temperature of maximum density of seawater

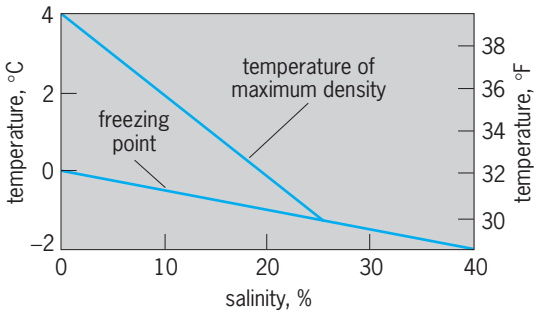

Fig. 1. Change of freezing point and temperature of maximum density with varying salinity of seawater.

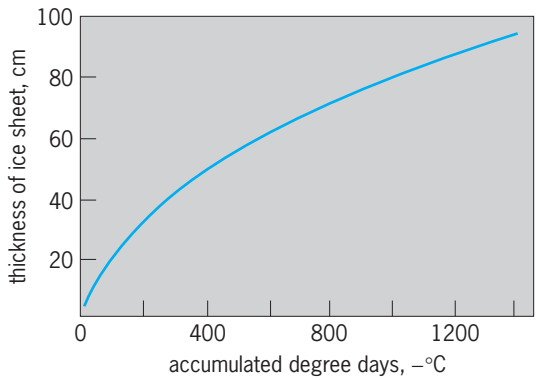

Fig. 2. Growth of undisturbed ice sheet. °F = (°C × 1.8) + 32. 1 cm = 0.4 in.

vary with salinity (**Fig. 1**). When freezing occurs, small flat plates of pure ice freeze out of solution to form a network which entraps brine in layers of cells. As the temperature decreases more water freezes out of the brine cells, further concentrating the remaining brine so that the freezing point of the brine equals the temperature of the surrounding pure ice structure. The brine is a complex solution of many ions. With lowering of temperature below −8°C (18°F), sodium sulfate decahydrate ($Na_2SO_4 \cdot 10H_2O$) and calcium sulfate dihydrate ($CaSO_4 \cdot 2H_2O$) are precipitated. Beginning at −24°C (−11°F), sodium chloride dihydrate ($NaCl \cdot 2H_2O$) is precipitated, followed by precipitation of potassium chloride (KCl) and magnesium chloride dodecahydrate ($MgCl \cdot 12H_2O$) at −34°C (−29°F), and the remaining ions with further lowering of temperature.

The brine cells migrate and change size with changes in temperature and pressure. The general downward migration of brine cells through the ice sheet leads to freshening of the top layers to near zero salinity by late summer. During winter the top surface temperature closely follows the air temperature, whereas the temperature of the underside remains at freezing point, corresponding to the salinity of water in contact. Heat flux up through the ice permits freezing at the underside. In summer freezing can also take place under sea ice in regions where complete melting does not occur. Surface melt water (temperature 0°C or 32°F) runs down through cracks in the ice to spread out underneath and contact the still cold ice masses and underlying colder seawater. Soft slush ice forms with large cells of entrapped seawater which then solidifies the following winter.

The salinity of recently formed sea ice depends on rate of freezing; thus sea ice formed at −10°C (14°F) has a salinity from 4 to 6 parts per thousand (‰), whereas that formed at −40°C (−40°F) may have a salinity from 10 to 15‰. Sea ice is a poor conductor of heat and the rate of ice formation drops appreciably after 4–6 in. (10–15 cm) are formed. An undisturbed sheet grows in relation to accumulated degree-days of frost. **Figure 2** shows an empirical relation between ice thickness and the sum of the mean diurnal negative air temperature (degrees Celsius). The thermal conductivity varies greatly with the air bubble content, ranging perhaps between 1.5 and 5.0×10^{-3} cal/(cm)(s)(°C) [0.6 and 2.0 J/(m)(s)(°C)].

The specific gravity of sea ice varies between 0.85 and 0.95, depending on the amount of entrapped air bubbles. The specific heat varies greatly because changing temperature involves freezing or melting of

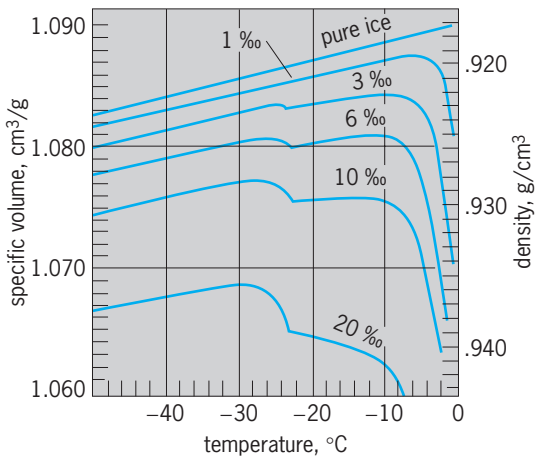

Fig. 3. Specific volume of sea ice for varying salinity and temperature, computed on basis of chemical model. °F = (°C × 1.8) + 32; 1 cm³/g = 1.730 in.³/oz; 1 g/cm³ = 0.578 oz/in.³ (*After D. L. Anderson, based on data in Arctic Sea Ice, NAS-NRC Publ. 598, 1958*)

Fig. 4. Open water in ice field. (*U.S. Navy*)

Fig. 5. Pressure ridge formation. (*U.S. Navy*)

ice. Near 0°C (32°F), amounts that freeze or melt at slight change of temperature are large and "specific heat" is anomalous. At low temperatures the value approaches that of pure ice; thus, specific heat for 4‰ saline ice is 4.6 cal/(g)(°C) [19 kJ/ (kg)(°C) at −2°C and 0.6 [2.5] at −14°C; for 8‰ saline ice, 8.8 [37] at −2°C and 0.6 [2.5] at −14°C.

High-saline sea ice may expand when cooled because further freezing out occurs with an increase of specific volume, for example, ice of salinity 8‰ at −2°C expands at a rate of about 93×10^{-4} cm^3/g per degree Celsius decrease in temperature, at −14°C expands 0.1×10^{-4}, but at −20°C contracts 0.4×10^{-4} per degree Celsius decrease. Change of specific volume with temperature and salinity is shown in **Fig. 3**.

Sea ice is viscoelastic. Its brine content, which is very sensitive to temperature and to air bubble content, causes the elasticity to vary widely. Young's modulus measured by dynamic methods varies from 5.5×10^{10} dynes/cm^2 (Pa) during autumn freezing to 7.3×10^{10} dynes/cm^2 (Pa) at spring breakup. Static tests give much smaller values, as low as 0.2×10^{10} dynes/cm^2 (Pa). The flexural strength varies between 0.5 and 17.3 kg/cm^2 over salinity range of 7–16‰ and temperatures −2 to −19°C. Acoustic properties are highly variable, depending principally on the size and distribution of entrapped air bubbles.

Electrical properties vary greatly with frequency because of ionic migration within the brine cells. For example, for sea salinity of 10‰ at −22°C, the dielectric coefficient is very large, about 10^6 at 20 Hz, and decreases with increasing frequency to about 10^3 at 10 kHz and to 10, or less, at 50 MHz. The effective electrical conductivity decreases with

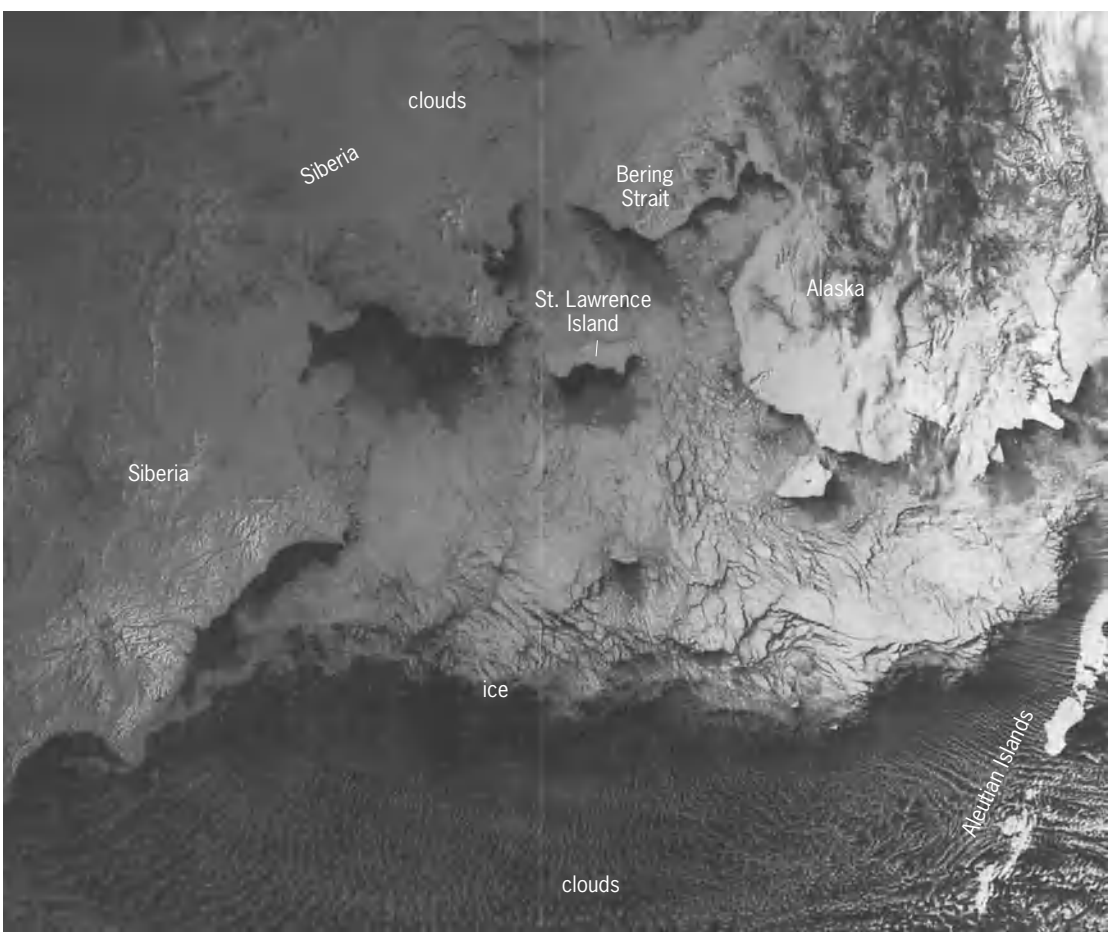

Fig. 6. Enhanced infrared image of Bering Sea by satellite, December 1976. (*Environmental Products Branch*, *NOAA*)

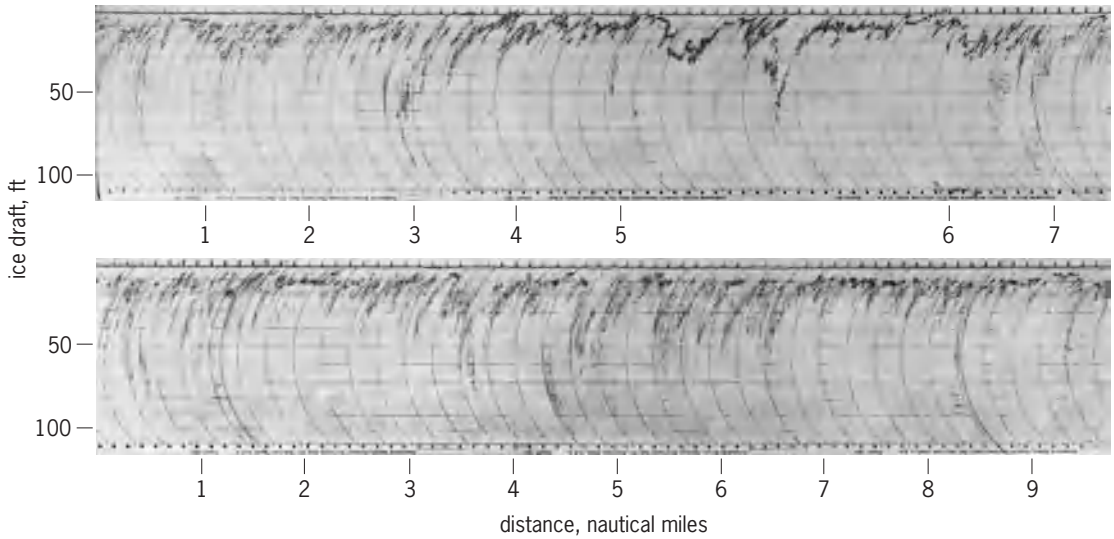

Fig. 7. Underside profile records of sea ice by submarine. Elongation of the distance scale between mile 5 and mile 6 on the upper record is due to the slower speed of the submarine during this portion of the observation. 1 nautical mile = 1.8 km; 1 ft = 0. 3 m. (*U.S. Navy*)

lowering of temperature, for example, from less than 10^{-3} mho/cm at $-5°C$ to 10^{-6} mho/cm at $-50°C$ (frequency 1 to 10 kHz).

Extent and observation. The sea ice in any locality is commonly a mixture of recently formed ice and old ice which has survived one or more summers. Except in sheltered bays, sea ice is continually in motion because of wind and current. The weaker parts of the sea ice canopy break when overstressed in tension, compression, or shear, pulling apart to form a lead (open water) or piling block on block to form a pressure ridge (**Figs. 4** and **5**). Ridges may form in ice of any thickness, from thin sheets (10 cm or 4 in. in thickness) to heavy blocks (3 m or 10 ft or more in thickness). The ridges may pile 13 m (43 ft) high above and extend 50 m (164 ft) below the sea surface. Massive ridges become grounded in coastal zones. The extent, movement, and type of sea ice are routinely observed by satellite or aircraft-carried sensor systems. **Figure 6** is an infrared photograph of ice covering the Bering Sea. The underside can be observed from a submarine transiting under the sea ice and using a sonar to record the draft profile of the ice canopy (**Fig. 7**). Waldo Lyon

Bibliography. J. C. Comiso, Sea ice microwave emissivities from satellite passive microwave and infrared observations, *J. Geophys. Res.*, vol. 88, no. C12, September 1983; *Proceedings of the Conference on Sea Ice*, NRC, Iceland, 1972; Symposium on Applied Glaciology, *J. Glaciol.*, vol. 19, no. 81, 1977; Symposium on Remote Sensing in Glaciology, *J. Glaciol.*, vol. 15, no. 73, 1975; P. Wadhams, *Ice Characteristics in the Seasonal Ice Zone*, *Cold Regions Science and Technology*, vol. 2, 1980.

Sea of Okhotsk

A semienclosed basin adjacent to the North Pacific Ocean, bounded on the north, east, and west by continental Russia, the Kamchatka Peninsula, and northern Japan (see **illus.**). On its southeast side the Sea of Okhotsk is connected to the North Pacific via a number of straits and passages through the Kuril Islands. The sea covers a surface area of approximately 590,000 mi^2 (1.5 million km^2), or about 1% of the total area of the Pacific, and has a maximum depth of over 9000 ft (3000 m). Its mean depth is about 2500 ft (830 m). Owing to the cold, wintertime Arctic winds that blow to the southeast, from Russia toward the North Pacific Ocean, the Sea of Okhotsk is partially covered with ice during the winter months, from November through April. *See* BASIN; PACIFIC OCEAN.

The basic circulation of surface water in the Sea of Okhotsk is thought to be generally counterclockwise, with water entering the basin south of Kamchatka through passages in the Kuril Islands, flowing north, west, and south through the sea, and exiting through other passages in the islands, roughly halfway between northern Japan and Kamchatka. Since most of the sills in the passages of the Kuril

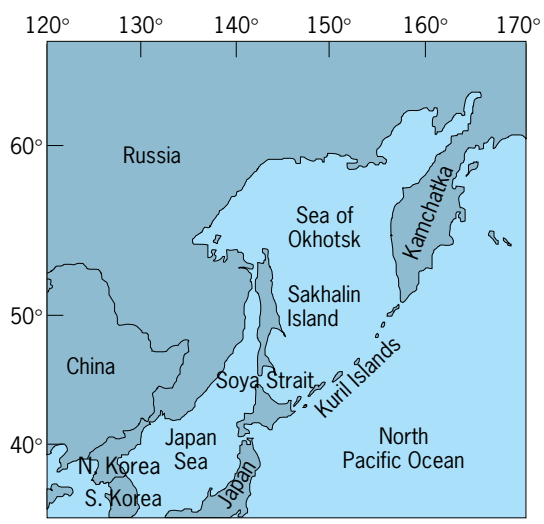

Sea of Okhotsk and adjacent landmasses.

Islands are shallower than about 900 ft (300 m), there is a limit to the amount of water that can be exchanged with the North Pacific. Below the sill depth, the water of the Sea of Okhotsk is trapped and can leave only via very slow vertical mixing with the surface waters.

The amount of water exchanged between the Sea of Okhotsk and the North Pacific Ocean is not well known, but it is thought that the waters of the Sea of Okhotsk that do enter the North Pacific may play an important role in the Pacific's large-scale circulation. The reason is the extreme winter conditions over the Sea of Okhotsk: its waters are generally colder and have a lower salinity than the waters at the same density in the North Pacific. It has been suggested that this flow from the Sea of Okhotsk into the North Pacific is the source of the large-scale subsurface tongue of low-salinity water at depths of 600–1800 ft (200–600 m) that can be seen over much of the subtropical North Pacific, from Asia to nearly the North American continent and south into the near-equatorial regions of the Pacific Ocean. *See* OCEAN CIRCULATION; SEAWATER.

The Sea of Okhotsk was generally off limits for scientific studies or commercial ventures from the 1940s through the mid-1990s. In recent years, the Russian government has relaxed restrictions on access to the sea. As a result, commercial fishing of salmon and pollack in the region by several nations is increasing rapidly, and offshore oil drilling has recently begun in the area just east of Sakhalin Island. A number of scientific studies of the Sea of Okhotsk have been undertaken that should help to ensure that these abundant natural resources are managed wisely. Since 1998, a joint Japanese, Russian, and United States observation program has examined some of the more important scientific questions pertaining to the interaction of ice, the atmosphere, and the ocean in the region. The new measurements suggest that tidal flows through the straits in the Kuril Islands are as strong as anywhere in the world, providing a source for the transport of the Sea of Okhotsk's properties into the North Pacific. It also appears that there is a great deal of year-to-year variability in ice production of in the Sea of Okhotsk that is closely related to regional changes in the state of the atmosphere. With limited exploration of the Sea of Okhotsk now permitted, it is expected that scientific knowledge of the region will continue to grow. *See* TIDE. Stephen C. Riser

Bibliography. S. Martin, R. Drucker, and K. Yamashita, The production of ice and dense shelf water in the Okhotsk Sea polynyas, *J. Geophys. Res.*, 103:27,771–27,782, 1998; I. Polyakov and S. Martin, Interaction of the Okhotsk Sea diurnal tides with the Kashevarov Bank polynya, *J. Geophys. Res.*, 105:3281–3294, 2000; J. L. Reid, Jr., *Northwest Pacific Ocean Waters in Winter*, Johns Hopkins University Press, 1973; L. D. Talley, Distribution and formation of North Pacific Intermediate Water, *J. Phys. Oceanog.*, 23:517–537, 1993; Y. You et al., Roles of the Okhotsk Sea and Gulf of Alaska in forming the North Pacific Intermediate Water, *J. Geophys. Res.*, 105:3253–3280, 2000.

Sea squirt

A sessile tunicate of the class Ascidiacea (phylum Chordata), so named because water is squirted from two openings when the animal is touched. Sea squirts may be solitary, colonial, or compound, but always have a permanent enclosing structure, the test or tunic, composed of polysaccharide material structurally similar to cellulose. In colonies, individuals are loosely attached together; in compound ascidians the individual zooids are embedded in a common test.

The individual is roughly cylindrical in shape and attached at the base (see **illus.**). The test is lined by the mantle or body wall, a structure containing blood vessels and muscle fibers. The mantle encloses a space or atrial cavity. Suspended within the atrial cavity is a pharynx or branchial sac perforated by numerous oval slits (stigmata), each one bearing a ring of marginal cilia. Constant beating of the cilia draws in water through the inhalant or oral aperture into the branchial sac. Water passes through the stigmata into the atrial cavity and is discharged at the exhalant or atrial aperture. In the branchial sac, food particles from the water are filtered on a mucous sheet which is then rolled up and passed to the esophagus. The anus is located in the atrial cavity below the exhalant aperture so that feces are ejected with the exhalant water current (as are renal and reproductive products).

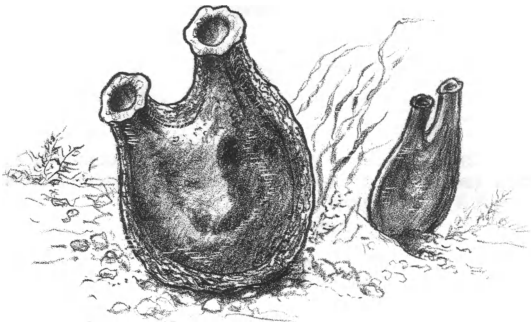

Sea squirts; these saclike tunicates serve as scavengers and as food for higher forms.

Sea squirts are dioecious, and fertilization may be internal or external. In some the egg develops into a tadpole which swims briefly before attaching by the anterior and metamorphosing into an adult; in others this free-swimming larval stage is suppressed.

Sea squirts occur in all seas and at all depths in the ocean. Most require a hard substratum on which to settle, but a few live in sand or mud. They are particularly common in shallow-water environments such as rocky shores, piers, and boat hulls, and in mangrove lagoons. They thrive best where there is a large amount of suspended particulate matter in the water to provide a source of food. *See* ASCIDIACEA; TUNICATA. Ivan Goodbody

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1988; H. H. Plough, *Sea Squirts of the Atlantic Continental Shelf from Maine to Texas*, 1978.

Sea urchin

A marine echinoderm of the class Echinoidea. These invertebrates are found commonly in shallow waters. The soft internal organs are enclosed in and protected by a test or shell (see **illus.**) consisting of a number of plates which fit closely together and are located under the skin. The oral surface is in contact with the substratum. Five teeth, located in the mouth, form part of Aristotle's lantern, a complex chewing structure. Average size for a sea urchin is about 2–3 in. (5–8 cm) in diameter; however, a specimen of *Sperosoma giganteum* from Japan measured 1 ft (30 cm) in diameter. Many species burrow into the sand, while others move into rock crevices to protect themselves from severe tidal action.

Like many other species of echinoderms, sea urchins use tube feet for locomotion and pincer-like structures called pedicellariae to keep the shell clean. In those species with long spines, tube feet serve a tactile function and the spines are used as locomotary structures. Sea urchins feed on most available animal and vegetable materials. The edible sea urchin from the Mediterranean (*Echinus esculentus*) is able to climb by using the tube feet as suction cups. The turban urchin (*Diadema antillarum*), a tropical species found in West Indian waters, has long, slender, and needle-sharp poisonous spines. These serve as defensive structures and can be moved in the direction of any threatening object "seen" by the numerous eye spots on the body surface. *See* EYE (INVERTEBRATE).

Fertilization is external. The larvae that develop are known as echinoplutei, bilaterally symmetrical organisms. The echinopluteus undergoes a transformation to radial symmetry during metamorphosis. When sea urchins are injured, regeneration occurs; this is especially true of the test. Their regenerative ability does not appear to be as great as that of starfishes. Severe injury to a mature animal containing reproductive cells has been advanced as a cause for what is termed a breeding epidemic. It is postulated that under such conditions a hormone is liberated which induces or stimulates adjacent animals to reproduce. Since sea urchins have the ability to produce large numbers of reproductive cells, they are used experimentally in developmental studies; the eggs have been sent into outer space to study the effects of that environment on development. *See* DEVELOPMENTAL BIOLOGY; ECHINODERMATA; ECHINOIDEA; INVERTEBRATE EMBRYOLOGY; REGENERATION (BIOLOGY). Charles B. Curtin

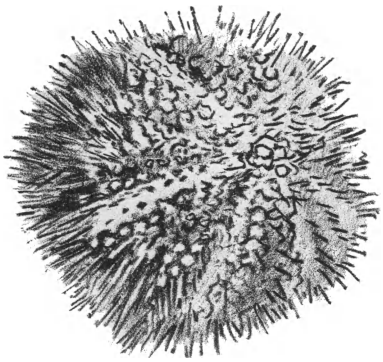

The sea urchin shell is covered with spines arranged in five broad areas that are separated by narrow unprotected areas.

Seaborgium

A chemical element, symbol Sg, atomic number 106. Seaborgium has chemical properties similar to tungsten. It was synthesized and identified in 1974. This discovery of seaborgium took place nearly simultaneously in two nuclear laboratories, the Lawrence Berkeley Laboratory at the University of California and the Joint Institute for Nuclear Research at Dubna in Russia.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

The Berkeley group, under the leadership of A. Ghiorso, used as its source of heavy ions the Super-Heavy Ion Linear Accelerator (SuperHILAC). The production and positive identification of the isotope of seaborgium with the mass number 263 decays with a half-life of 0.9 + 0.2 s by the emission of alpha particles of principal energy 9.06 + 0.04 MeV. This isotope is produced in the reaction in which four neutrons are emitted: $^{249}Cf(^{18}O,4n)$.

The Dubna group, under the leadership of G. N. Flerov and Y. T. Oganessian, produced its heavy ions with a heavy-ion cyclotron. They found a product that decays by the spontaneous fission mechanism with the very short half-life of 7 ms. They assigned it to the isotope ^{259}Sg, suggesting reactions in which two or three neutrons are emitted: $^{207}Pb(^{54}Cr,2n)$ and $^{208}Pb(Cr^{54},3n)$. *See* NOBELIUM; NUCLEAR CHEMISTRY; PERIODIC TABLE; TRANSURANIUM ELEMENTS. Glenn T. Seaborg

Bibliography. S. Hofmann, *On Beyond Uranium: Journey to the End of the Periodic Table*, 2002; G. T. Seaborg (ed.), *Transuranium Elements: Products of Modern Alchemy*, 1978; G. T. Seaborg and W. D. Loveland, *The Elements Beyond Uranium*, 1990.

Seamount and guyot

A seamount is a mountain that rises from the ocean floor; a submerged flat-topped seamount is termed a guyot. By arbitrary definition, seamounts must be at least 3000 ft (about 900 m) high, but in fact there is a continuum of smaller undersea mounts, down to heights of only about 300 ft (100 m). Some seamounts are high enough temporarily to form oceanic islands, which ultimately subside beneath sea level. There are on the order of 10,000 seamounts in the world ocean, arranged in chains (for example, the Hawaiian chain in the North Pacific) or as isolated features. In some chains, seamounts are packed closely to form ridges (for example, the Walvis Ridge in the South Atlantic). Very large oceanic volcanic constructions, hundreds of kilometers across, are called oceanic plateaus (for example, the Manihiki Plateau in the South Pacific). *See* OCEANIC ISLANDS.

Leaving aside small tectonic fragments of continents isolated in oceanic surroundings (for example, the Seychelles in the Indian Ocean), almost all seamounts are the result of submarine volcanism, and most are built within less than about 1 million years. Seamounts are made by extrusion of lavas piped upward in stages from sources within the Earth's mantle to vents on the seafloor. The lavas in most seamounts within the oceanic basins are composed mainly of the tholeiitic variety of basalt, but lavas are commonly of more alkalic varieties in later stages of eruption. The basaltic lavas generally erupt quietly, but the rapid escape of gases dissolved in the magmas may fragment the cooling lavas, forming pyroclastic (ashy and blocky) deposits. Pyroclastics become increasingly important as the seamount nears sea level, where confining hydrostatic pressures diminish. Seamounts closely associated with island arcs are commonly andesitic rather than basaltic, and these commonly are more explosive than basaltic volcanoes. *See* BASALT; MARINE GEOLOGY; VOLCANO.

Seamounts provide data on movements of tectonic plates on which they ride, and on the rheology of the underlying lithosphere. The trend of a seamount chain traces the direction of motion of the lithospheric plate over a more or less fixed heat source in the underlying asthenosphere part of the Earth's mantle. During cooling of the lavas, the magnetic minerals are magnetized in the direction of the Earth's magnetic field, and this direction is preserved in the rocks. From magnetic surveys or from oriented samples of seamount basalt, the paleolatitude of a seamount can be determined. The weight of a seamount flexes the underlying lithosphere, with the result that seamounts and seamount chains are generally flanked by a moatlike depression that may be in turn bordered by a broad arch. The form of the moat and arch depends on the rigidity of the lithosphere at the time of construction of the volcano: the older the lithosphere, the more rigid. *See* LITHOSPHERE; PLATE TECTONICS.

Many smaller seamounts are flat topped, possibly as a result of summit collapse, but larger seamounts tend to be more or less conical. As a seamount grows upward, some lavas, rather than being extruded from the summit of the volcano, are emplaced laterally along radial rift zones. Volcano slopes are built by lava flows, modified by undersea avalanching and sliding, which may alter the original conical form of a seamount to a starfishlike ground plan, with arms along the rift zones. Typical side slopes range about 15–30°.

Erosion of the subaerial summit region of an island seamount, or the accumulation of shallow-water reefal deposits of calcium carbonate atop a tropical seamount during subsidence, result in a nearly flat top. Termed guyots, some that were at sea level 100 million years ago have summits that are at present at water depths of about 1 mi (1500 m). *See* MARINE GEOLOGY; VOLCANO. Edward L. Winterer

Bibliography. H. Aoki and S. Iizuka (eds.), *Volcanoes and Tectonosphere*, 1976; E. L. Winterer, D. M. Hussong, and R. W. Decker (eds.), *The Eastern Pacific Ocean and Hawaii*, 1989.

Seaplane

An airplane capable of navigating on, taking off from, and alighting upon the surface of water. Seaplanes are grouped into two main types: flying boats and float planes. In the flying boat, the hull, which provides buoyancy and planing area, is an integral part of the airframe, a specially designed fuselage which supports the wings and tail surfaces and houses the crew, equipment, and cargo. Although multihull flying boats have been built, modern use is confined to single-hull boats with lateral stability on the water provided by small floats or pontoons attached to the wings (**Fig. 1**). The float plane is a standard landplane made capable of water operation by the addition of floats which are attached to the airframe by struts. In practice the twin float is used exclusively, lateral stability on the water being provided by the separation of the two identical floats (**Fig. 2**). A seaplane with retracting wheels which permit either land or water operation is known as an amphibian. *See* AIRFRAME; FUSELAGE.

Takeoff dynamics. At rest or at low water speeds, the seaplane is supported by buoyancy and acts as a displacement vessel. Directional control in this

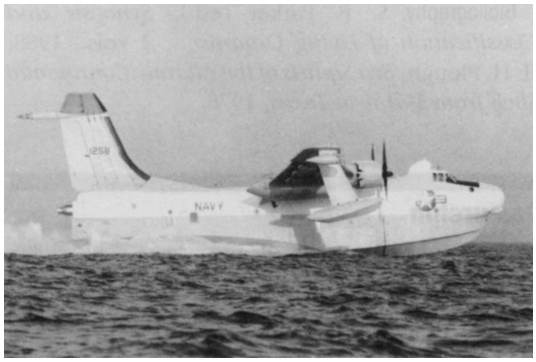

Fig. 1. Martin P5 Marlin twin-engine flying boat. (*Martin Marietta Corp.*)

Fig. 2. Single-engine float plane.

regime is achieved by water rudders or by asymmetric engine thrust in multiengine craft. As speed increases, hydrodynamic forces act on the bottom of the hull or floats, and the seaplane is supported by a combination of aerodynamic and hydrodynamic lift forces. The speed regime where the seaplane is supported by dynamic force rather than by buoyancy is known as the planing regime. While planing, hydrodynamic drag is a function of the apparent weight of the seaplane in relation to the water. Since aerodynamic lift supports more of the weight with increasing speed, the hydrodynamic drag decreases, allowing the seaplane to accelerate to flying speed. The water speed of maximum drag is known as the hump speed, and is the speed where the seaplane goes "on the step" or changes from a displacement to a planing body. Directional control in the planing regime is through the seaplane's aerodynamic control surfaces. *See* AERODYNAMICS; BUOYANCY; HYDRODYNAMICS.

Hull or float design. While the hull or float is designed to be as perfect a streamlined body as possible, certain deviations are necessary to ensure satisfactory planing action. These deviations are required because water exerts suction on convex surfaces. Although it is impossible to avoid the use of convex surfaces near the bow, the bow is in the water only at very low speeds, its curved surfaces rising above the surface even before full planing is developed. The lines formed by the intersection of the float bottom with vertical planes parallel to the float centerline are known as buttock lines. These buttock lines must be straight for most of the float forebody length in order to avoid suction and to allow dynamic planing forces to be generated efficiently. These lines must terminate at their aft ends in discontinuities called steps, since any upward change of direction of a continuous line would cause suction. The step is usually located near the longitudinal center of gravity of the seaplane, for it limits the aft travel of the hydrodynamic center of pressure. The unbalanced moment caused by hydrodynamic forces acting through a center of pressure not coincident with the center of gravity must be overcome by the seaplane's aerodynamic controls.

Although a flat surface is the most efficient planing bottom, most seaplanes have a vee bottom, with the sharpness of the vee expressed in dead-rise angle, measured from the horizontal to the nearest leg of the vee. A vee bottom lessens the impact forces experienced during landing and takeoff. The deadrise angle selected for a particular seaplane is a compromise between planing efficiency and impact reduction and is usually in the 20–30° range.

Spray control. Control of the spray thrown up by the floats or hull is an important consideration in seaplane design, for that spray can damage the propellers or structure and impair the pilot's vision. The most common spray-control technique is to install a vertical spray strip at the chine (where the bottom meets the side of the float). The spray strip alters the direction of the spray so that it does not touch vital areas.

Takeoff and landing conditions. Seaplanes are designed primarily for operation from fairly calm or sheltered waters. Seaplane landings or takeoffs from rough water are analogous to a landplane operating from very rough ground. Since impact forces on water contact are proportional to the square of the impact velocity, the only truly successful rough-water seaplane has been the Japanese Shin Meiwa PS-1, launched in 1967, which used high-lift devices to achieve touchdown and lift-off speeds of about one-half those of a conventional aircraft.

Development. The first seaplane was flown in France in 1910 by Henri Fabre. Rapid design progress led to large multiengined flying boats by World War I. Subsequently, long-distance and international flight was dominated by the seaplane, and culminated in regular passenger service across the Pacific and Atlantic by Martin and Boeing Clippers. During World War II, thousands of seaplanes were used by all combatants for reconnaissance, rescue, and logistic support missions. Wartime research led to the ultimate large seaplane, the four-jet, 600-mi/h (270-m/s) Martin P6M Seamaster (**Fig. 3**) launched in 1955.

Use. Seaplanes suffer a number of disadvantages in relation to landplanes: greater aerodynamic drag because of the float and hull shape; greater weight because of the additional structure required; susceptibility to corrosion when in salt water; inability to operate when the water is frozen or rough; and a requirement for a high level of pilot skill. The one main advantage of the seaplane, the ability to

Fig. 3. Martin P6M Seamaster four-jet flying boat. (*Martin Marietta Corp.*)

operate from an unprepared surface, has been negated by the worldwide availability of airports. Consequently, the large seaplane has all but disappeared from commercial and military service, and its rescue function has been assumed by the long-range helicopter. *See* HELICOPTER.

Most modern seaplanes are of the small single- or twin-engine type and are used for sport or recreational flying. They still serve a valuable logistic support rule in remote areas without airports but with many lakes and rivers. *See* AIRPLANE.

Richard A. Hoffman

Bibliography. M. Allward, *An Illustrated History of Seaplanes and Flying Boats*, 1982, reprint 1990; M. Faure, *Flying a Floatplane*, 3d ed., 1996; H. R. Palmer, *The Seaplanes*, 1980; J. Stroud, *The World's Civil Marine Aircraft*, 1979.

Seasons

The four divisions of the year based upon variations of sunlight intensity (solar energy per unit area at the Earth's surface) at local solar noon (noontime) and daylight period. The variations in noontime intensity and daylight period are the result of the Earth's rotational axis being tilted 23°.5 from the perpendicular to the plane of the Earth's orbit around the Sun. The direction of the Earth's axis with respect to the stars remains fixed as the Earth orbits the Sun. If the Earth's axis were not tilted from the perpendicular, there would be no variation in noontime sunlight intensity or daylight period and no seasons.

A common misconception is that the seasons are caused by variation of the Earth-Sun distance: many people think that the Earth is nearest to the Sun in summer. However, the Earth is farthest from the Sun during the first week of July (early summer in the Northern Hemisphere) and nearest to the Sun during the first week of January; also, winter and summer occur simultaneously on opposite sides of the Equator. The Earth-Sun distance does not influence the seasons because it varies only slightly, and it is overwhelmed by the effects of variations in sunlight intensity and daylight period due to the alignment of the Earth's axis.

The more hours the Sun is above the horizon, the greater the heating effect produced by the Sun. Daylight period varies for all locations except the Equator, which has 12 h of daylight every day. The poles have the widest variation in daylight period, experiencing 6 months of continuous daylight and 6 months of continuous night.

Sunlight intensity at a location depends upon the angle from the horizon to the Sun at local solar noon; this angle in turn depends upon the location's latitude and the position of the Earth in its orbit. At increased angles, a given amount of sunlight is spread over smaller surface areas, resulting in a greater concentration of solar energy, which produces increased surface heating (see **illus.**). Intensity is a more important factor than the number of daylight hours in determining the heating effect at the Earth's surface.

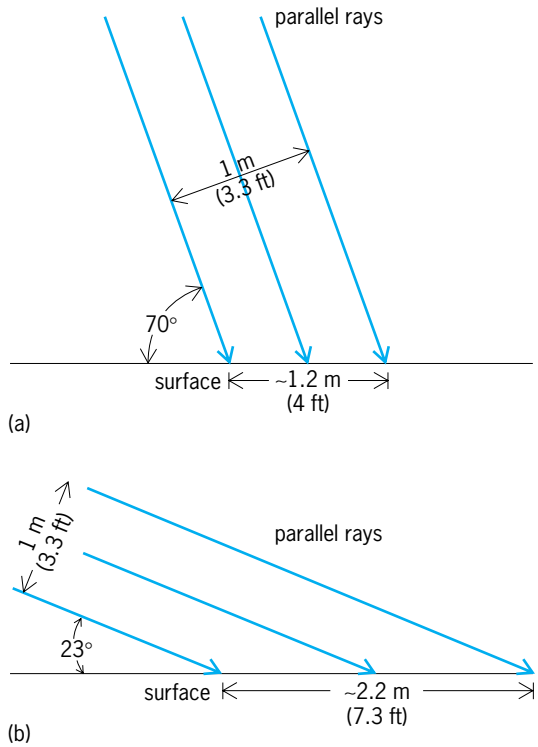

Relation of the angle of the Sun above the horizon to the area covered by a beam of sunlight 1 m (3.3 ft) on a side at 43° latitude. (a) The beam at solar noon on the summer solstice; the beam covers an area 1.2 m (4 ft) on a side, or about 1.4 m^2 (16 ft^2). (b) The beam at solar noon on the winter solstice; the beam covers an area 2.2 m (7.3 ft) on a side, or about 4.8 m^2 (53 ft^2). Since sunlight at noon on the winter solstice is spread over an area more than three times larger than that covered on the summer solstice, less solar heating occurs in winter than in summer.

For example, the poles receive 6 months of continuous daylight, a situation that would seem likely to produce significant heating. However, the maximum sunlight intensity at the poles is equal only to the noontime intensity on the first day of winter at 43° latitude (for example, the northern United States). The many hours of daylight at the poles provide relatively little heating effect because of the low sunlight intensity. By contrast, the maximum noontime intensity at 43° latitude equals the solar energy per unit area received at the Equator's surface in late November.

For locations on or north of the Tropic of Cancer (23°.5 north) and on or south of the Tropic of Capricorn (23°.5 south), maximum sunlight intensity and daylight period occur on the summer solstice; minimum intensity and period occur on the winter solstice. The solstices occur on about June 21 and December 21, the June solstice marking the start of summer in the Northern Hemisphere and the beginning of winter in the Southern Hemisphere. *See* EARTH ROTATION AND ORBITAL MOTION.

Harold P. Coyle

Bibliography. J. Bennett et al., *Cosmic Perspective*, 3d ed., Addison Wesley, 2004; E. Chaisson and S. McMillan, *Astronomy Today*, 5th ed., Prentice Hall, 2005; J. M. Pasachoff, *Science Explorer: Astronomy*, Prentice Hall, 2005; J. M. Pasachoff and

A. Filippenko, *The Cosmos: Astronomy in the New Millenium*, 3d ed., Brooks Cole, 2007.

Seawater

An aqueous solution of salts of a rather constant composition of elements whose presence determines the climate and makes life possible on the Earth and which constitutes the oceans, the mediterranean seas, and their embayments. The physical, chemical, biological, and geological events therein are the studies that are grouped as oceanography. Water is most often found in nature as seawater (about 98%). The rest is ice, water vapor, and fresh water. The basic properties of seawater, their distribution, the interchange of properties between sea and atmosphere or land, the transmission of energy within the sea, and the geochemical laws governing the composition of seawater and sediments are the fundamentals of oceanography. *See* HYDROSPHERE; OCEANOGRAPHY.

Physical Properties

Water is arguably the most important chemical substance in the terrestrial system, its large specific heat (the highest of almost any substance) is of extreme import to weather, climate, and the ocean. Its physical properties include a large dielectric constant, elevated boiling and freezing points, high osmotic pressure, and strong dissolving power—properties that are modified by the dissolved salts and the other materials found in the ocean.

The major chemical constituents of seawater are cations (positive ions) and anions (negative ions) [**Table 1**]. In addition, seawater contains the suspended solids, organic substances, and dissolved gases found in all natural waters. A standard salinity of 35 practical salinity units (psu; formerly parts per thousand, or ‰) has been assumed. While salinity does vary appreciably in oceanic waters, the fractional composition of salts is remarkably constant throughout the world's oceans. In addition to the dissolved salts, natural seawater contains particulates in the form of plankton and their detritus, sediments, and dissolved organic matter, all of which lend additional coloration beyond the blue coming from Rayleigh scattering by the water molecules. Almost every known natural substance is found in the ocean, mostly in minute concentrations. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

The density of seawater is an important quantity that enters sensitively into many dynamical and thermodynamical processes. An internationally accepted equation of state has been derived that allows the accurate computation of density, given measurements of salinity, temperature, and pressure. A high accuracy is required to compute the small density variations that set up horizontal pressure gradients which, together with the Earth's Coriolis force, balance large-scale geostrophic flows such as the Gulf Stream. Salinity is usually measured via electrical conductivity. *See* CORIOLIS ACCELERATION; GULF STREAM.

Seawater is slightly compressible, implying that sea level is approximately 30 m (100 ft) lower than it would be if it water were incompressible. A parameter known as the thermodynamic space is quite convenient for displaying measurements of temperature and salinity taken during vertical probes of the water column, and allows the identification of the origins of various water masses via a technique called isentropic analysis.

The density of sea ice drops to the range of 857–924 kg/m^3, depending on the amount of brine. These effects of salinity are related to the very large osmotic pressure of seawater.

Various thermodynamic quantities define other properties of seawater (**Table 2**). All of these coefficients are functions of salinity, temperature, and density, and some of their variations with pressure, for example, lead to subtle but profound effects in the sea.

Acoustics. Because of the complex temperature and salinity structure of the upper ocean, highly variable sound propagation paths occur there. The oceanic mixed layer often serves as a leaky waveguide for a sound source located in that layer; in the absence of a vertical temperature gradient, sound propagates along the mixed layer by a combination of upward refraction and surface reflection–forward scattering. However, for a source located beneath the thermocline, the colder water initially refracts propagation paths downward as the sound speed decreases. This is partially offset by the slow increase in speed due to pressure at depth, with the result that there develops a minimum in overall sound speed, the so-called deep sound channel, a secondary waveguide. In tropical and temperate seas, this minimum lies near 1200 m (4000 ft) depth with the result that detectable levels of sound may be heard across entire ocean basins. In a typical sound

TABLE 1. Major constituents of seawater (salinity 35 psu)*

Positive ions	Amount, g/kg	Negative ions	Amount, g/kg
Sodium (Na^+)	10.752	Chloride (Cl^-)	19.345
Magnesium (Mg^{2+})	1.295	Bromide (Br^-)	0.066
Potassium (K^+)	0.390	Fluoride (F^-)	0.0013
Calcium (Ca^{2+})	0.416	Sulfate (SO_4^-)	2.701
Strontium (Sr^{2+})	0.013	Bicarbonate (HCO_3^-)	0.145
		Boron hydroxide [$B(OH)_3^-$]	0.027

* Water, 965 psu: psu dissolved materials, 35 psu.

TABLE 2. Thermodynamic properties of seawater*

Quantity	Symbol	Value
Isothermal-isohaline compressibility	a_p	4.27×10^{-10}/Pa
Specific heat at constant volume and salinity	$C_{\alpha s}$	3939 J/kg-°C
Specific heat at constant pressure and salinity	C_{ps}	3994 J/kg-°C
Isobaric-isohaline thermal expansion coefficient	a_T	2.41×10^{-4}/°C
Isobaric-isothermal saline contraction coefficient	a_s	7.45×10^{-4}/psu
Speed of sound	c	1520 m/s (3280 mi/h)
Thermal conductivity	kq	0.596 W/m-°C
Latent heat of vaporization	L_v	2.453×10^6 J/kg
Latent heat of fusion	L_f	0.335×10^6 J/kg

*All quantities are evaluated at $s = 35$ psu and $T = 20$°C (68°F), except L_f, for which $T = 0$°C (32°F).

speed profile for a midlatitude ocean, the major variability occurs in the seasonal mixed layer.

At even greater depths, the pressure continues to increase the sound speed and to bend rays upward. Near-surface rays that are launched beneath the mixed layer at shallow angles will initially penetrate into the deep ocean while undergoing refraction, until near depths of 5000 m (16,500 ft) the rays will have been refracted through the horizontal; subsequently they will bend upward toward the surface again to intersect it at distances near 60–65 km (35–40 mi) from the source. Thus, there is a focusing of sound intensity into an annular region at the surface centered on the source, the so-called first convergence zone; sound levels there may be as much as 20 decibels above those expected from nominal spreading and attenuation. Within the annular region at the surface, reflection and scattering again launch acoustic energy downward, which then goes through the same processes once again; as a result, a series of high-intensity convergence zones are developed in the form of circular annuli surrounding the source, which have radii that are integral multiples of the basic 60–65 km (35–40 mi) and widths of perhaps 5 km (3 mi).

In addition to these intensity variations due to spreading and refraction losses, there exists attenuation from absorption and volume scattering in the seawater. While the attenuation coefficient nominally varies as the square of the frequency, there are significant modifications from dissolved chemicals and acoustic scattering.

Underwater noise levels limit the detection of sound in the sea; they vary enormously with frequency, wind, rain, and shipping traffic, ranging from approximately 40 dB at 10 Hz to −30 dB at 10 kHz. At higher frequencies, receiver thermal noise is often limiting. *See* ACOUSTICS; SOUND; UNDERWATER SOUND.

Dynamical properties. The most important dynamical characteristic of the sea are waves, ocean currents, and tides; however, this discussion is limited to air-sea interaction and diffusion. *See* OCEAN WAVES; TIDE.

Air-sea interaction. The sea absorbs roughly 98% of the visible energy incident on it at high sun angles, and converts it to heat, thereby warming the upper layers of the ocean. The solar spectral irradiance is attenuated as sunlight penetrates beneath the sea surface. Under the influences of wind wave stress, turbulence, and convective overturning of the upper layers, the thermalized solar energy that is deposited into the upper layers is mixed down to depths ranging perhaps 30–150 m (100–500 ft), depending on location and season. Because of the high specific heat of seawater, this mixed-layer storage represents by far the largest heat capacity on the face of the Earth. Subsequent exchanges across the interface are governed by wind stress and air-sea temperature differences. When the air is cooler than the ocean, as occurs during cold-air outbreaks or at night, significant evaporation sets in. This results in turbulence in the atmospheric boundary layer, which causes rougher seas than when the air-sea temperature difference is positive. Values of air-sea temperature differences in excess of roughly −1°C (−1.8°F) lead to strong interchanges during such unstable conditions. When the air and sea temperatures are nearly the same, the exchanges are greatly reduced. *See* WIND STRESS.

Evaporation of water from the sea results in rising moist air, cloud formation and, upon condensation into clouds at altitude, the release of the latent heat of evaporation to the atmosphere, warming the air. Because the elevated warm air is farther from the Earth's axis of rotation than the air was at sea level, its changed angular momentum is transformed into winds via the Coriolis force. Thus the combination of latent heat release and Coriolis effect initiates the global atmospheric circulation, the wind stress on the ocean, and ultimately both the wind-driven and density-driven currents in the sea. Starting in the tropics then, the general circulation of the atmosphere is initiated via air-sea interaction. *See* ATMOSPHERIC GENERAL CIRCULATION.

The great poleward-directed systems (such as the Gulf Stream) that are found in every ocean basin transport heat out of the tropics and into temperate and polar regions. During their journey, the warm currents surrender to the atmosphere large amounts of latent and sensible heat, moisture, and momentum via the processes of evaporation, turbulent fluctuations, radiation, and conduction, in decreasing order of importance. In midlatitudes, for example, this warming of the atmosphere is responsible for maintaining the western parts of Europe at significantly

higher average temperatures than in the same latitudes in North America. Over longer times, these large-scale air-sea exchanges constitute one of the most important processes governing weather and climate dynamics.

Turbulence in the planetary boundary layer and the oceanic mixed layer is largely responsible for many of the secondary processes conditioning the air-sea interface. Surface wave generation, mixing, and exchanges of heat, momentum, and moisture all rely on turbulent diffusion for their efficacy. *See* HEAT BALANCE, TERRESTRIAL ATMOSPHERIC; OCEAN CIRCULATION.

Diffusive properties. Beyond waves and advection by currents, properties such as heat, salt, and momentum are transported by diffusive effects in the sea. Molecular diffusion coefficients are quite small; but eddylike fluctuations lead to turbulence and eddy diffusion coefficients which are orders of magnitude greater than their molecular counterparts and which depend on the scale of the fluctuations. The coefficients also differ for horizontal diffusion as compared with vertical diffusion, because buoyancy greatly inhibits transport in the vertical. The energy going into waves, currents, and diffusion of properties is ultimately removed from the system at scales near the dissipation length, that is, at about 1 mm. This fluid friction, along with bottom friction, acts to slow down both the atmosphere and ocean, but their large-scale dynamics are maintained by a continuing but unsteady input of energy from the Sun and the atmosphere. *See* DIFFUSION.

Electromagnetics. Electromagnetic and optical properties of seawater include both surface characteristics and bulk attributes.

Conductivity. The ocean is a moderately good conductor of electricity. Electrical conductivity, like the other material properties, depends not only on salinity but on temperature and pressure as well, although the dependence on pressure is quite weak. In seawater, conductivity is conventionally measured, and an algorithm is then used to convert it to salinity; a method for computing the inverse relationship is available as well. Not only is conductivity important for determination of the oceanic density field, but along with the dielectric constant, it is relevant to problems involving remote sensing of the sea surface by microwave methods.

Scattering and emission. Microwave power reflection coefficients (or reflectivities) are dependent on polarization and incidence angle. Waves on the sea modify the reflection coefficients significantly. The customary measure of reflection-scattering from a rough surface is the normalized radar cross section, relative to 1 m^2 of ocean surface. This is a function of frequency, incidence, azimuth, and reflection angles; polarization; wind speed–sea state; and surfactant coverage. Emission of thermal energy at microwave frequencies is used to determine thermodynamic temperatures, wind speed and, at lower frequencies, salinity of the sea surface. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; MICROWAVE; POLARIZED LIGHT.

Optics. The optical characteristics of seawater derive from a combination of physical, chemical, biological, and geological properties. Central to the subject are the questions of light transmission in the sea (for studies of biological productivity and underwater signaling), the color of seawater, and the use of variations in water color as a means for flow visualization in optical images of the sea.

Surface reflectivities. Surface reflectivities are due to a combination of intrinsic Fresnel reflection and rough surface scattering. The smooth-surface reflectivity is polarization sensitive and is similar to that for microwave frequencies. Since direct sunlight is unpolarized, the variation of reflectivity with incidence angle is given by the average of vertical and horizontal values. As the wind speed increases, reflection falls off.

Index of refraction. Over the nominal range of variation of wavelength, temperature, salinity, and pressure, the index of refraction does not depart from 1.34 by more than 0.015.

Volume absorption and scattering. While seawater is a reasonable conductor at microwave frequencies and below, at optical frequencies clearest seawater is a remarkably low-loss dielectric with a broad transparency band in the visible. The spectral beam absorption coefficient is a measure of the fractional absorption of a collimated beam of light in traversing 1 m of water. The addition of photosynthetic biological material profoundly alters the adsorption coefficient; chlorophyll *a* and phaeophytin *a* are the major absorbers, along with so-called yellow substance, or dissolved tanninlike compounds. Beyond absorption, the process of scattering is important, and this is described by a beam-scattering coefficient which is large in the forward direction but much smaller in the backscatter direction. Thus, the total beam attenuation coefficient is due to absorption and scattering.

Remote sensing with imaging optical spectrometers on satellites has yielded the first global estimates of near-surface chlorophyll and sediment concentrations by utilizing the concepts of ocean optics in the interpretation of multispectral images from satellites and aircraft. *See* OPTICAL DETECTORS; REMOTE SENSING; SCIENTIFIC AND APPLICATIONS SATELLITES.

John R. Apel

Chemical Properties

The complex mixture of salts and other substances found dissolved in seawater is quantitatively dominated by chloride and sodium ions, with somewhat lesser amounts of sulfate, magnesium, calcium, and other ions. The total salt content of seawater is expressed by the term salinity. This measure of the salt concentration has been rigorously defined in several ways; since 1978 it has been defined and measured according to the ratio of the electrical conductivity of a seawater sample to that of a standard solution of potassium chloride. For both technical and historical reasons, the salinity is not exactly the total mass of salts in solution, but the two values differ by only about 0.5%.

TABLE 3. Major inorganic nongaseous constituents in seawater*

Substance	g/kg	mmol/kg
Sodium ion (Na^+)	10.781	468.96
Potassium ion (K^+)	0.399	10.21
Magnesium ion (Mg^{2+})	1.284	52.83
Calcium ion (Ca^{2+})	0.4119	10.28
Strontium ion (Sr^{2+})	0.00794	0.0906
Chloride ion (Cl^-)	19.353	545.88
Sulfate ion (SO_4^{2-})	2.712	28.23
Bicarbonate ion (HCO_3^-)	0.126	2.06
Bromide ion (Br^-)	0.0673	0.844
Boron hydroxide [$B(OH)_3^-$]	0.0263	0.425
Fluoride ion (F^-)	0.00130	0.086

*Total mass concentration of above constituents, 35.170; total concentration of all other dissolved substances, ~0.03; water, ~964.80.

The average salinity of the world ocean is 34.73‰ (parts per thousand by weight, or grams per kilogram of seawater) and is remarkably constant. Only about one-tenth of all the water in the ocean departs by more than 1% from this mean value, so measurement techniques of the most exquisite sensitivity must be used to discern the real, and important, differences in salinity from place to place within the oceans. The salinity departs from this relative constancy in estuaries and other near-coastal regions where seawater is diluted with river water, in polar regions where ice melts in the spring, and in coastal lagoons or small seas where evaporation may sometimes exceed the input of fresh water by rivers and rain. Since evaporation and rainfall generally change only the total salt content and not the relative proportions of the salts, it is conventional to express the concentrations of substances in the ocean on the basis of a standard ocean water with a salinity of 35‰ (**Tables 3** and **4**).

Major and minor elements. It is also conventional to consider that the major dissolved constituents of seawater consist of 11 substances (Table 3), all present in concentrations of greater than 1 mg/kg or 1 part per million (ppm). In addition, however, the dissolved gases nitrogen and oxygen, while variable in concentration, generally amount to a total of about 10–30 mg/kg, and dissolved organic matter is present in variable concentrations of about 1 to several milligrams per kilogram.

The total concentration of salts in seawater, about 35 g in a kilogram, is great enough that the chemical activity of each ion is strongly affected by the presence of the others, and in addition numerous ion pairs and other complexes are present. These effects must be taken into account in any quantitative evaluation of the behavior of substances in seawater.

TABLE 4. Concentrations of dissolved gases in seawater*

Gas	mg/kg	μmol/kg
Nitrogen (N_2)	12.04	428.5
Oxygen (O_2)	7.48	233.7
Helium (He)	0.0000067	0.0017
Neon (Ne)	0.00014	0.0069
Argon (Ar)	0.457	11.44
Krypton (Kr)	0.00021	0.0025
Xenon (Xe)	0.000043	0.00035

*At 18°C (64°F) and salinity of 35‰, and at pressure of one standard atmosphere.

With the exceptions of oxygen and helium, the dissolved gases in seawater (Table 4) are not subject to any processes that can significantly change their concentrations once these concentrations have been established by coming to equilibrium with the atmosphere. They will not change during the circulation of that water into deeper parts of the ocean, except by mixing with other parcels. The concept of attaining equilibrium with the atmosphere is, however, a little fuzzy. It is common to assume standard atmospheric pressure, but atmospheric pressure varies. During windy conditions, breaking waves cause the entrainment of bubbles to some depth in the water; as these dissolve, they increase the concentrations, sometimes up to several percent above saturation. Observation shows that all ocean water has been affected to a measurable extent by these effects, so that they have to be taken into account when making careful calculations. Once below the surface, however, the concentrations of nitrogen and the noble gases (with the exception of helium) are not subject to further change. Helium, especially the rare isotope helium-3, is increased in deep water by its release from the mantle in places where fresh hot basalt comes in contact with seawater. Nitrogen is used and produced by biological processes, but not in amounts that can easily be measured in most of the ocean. Oxygen is both produced and consumed in biological reactions, and measurements of the concentration of oxygen are thus an important tracer of oceanic processes.

The minor elements exhibit a great range of concentrations. Most are present in only tiny fractions of a part per billion, providing a great challenge to the analyst. In many cases the limitation has been the great difficulty of collecting seawater from ships without contamination being far greater than the concentrations in the water, as well as the difficulty of obtaining reagents and equipment that do not contain more of the trace elements than the water samples themselves. It is possible to measure such low concentrations by combining great attention to ultraclean sampling and processing with modern highly sensitive instrumentation.

Nutrients. Unlike the major elements and the conservative minor elements, the important plant nutrients (phosphate, nitrate, and silicate) are not uniformly distributed in the ocean according to the salinity of the water. Instead, they exhibit marked changes from place to place, both vertically and horizontally. Characteristically, the nutrients are depleted in surface waters and are enriched at depth (**Fig. 1**). Similar and related profiles from much of the ocean have provided the general understanding of the interaction of ocean circulation and biogeochemical processes.

Small plants (the mostly unicellular photosynthetic plankton) in the sunlit surface waters use these

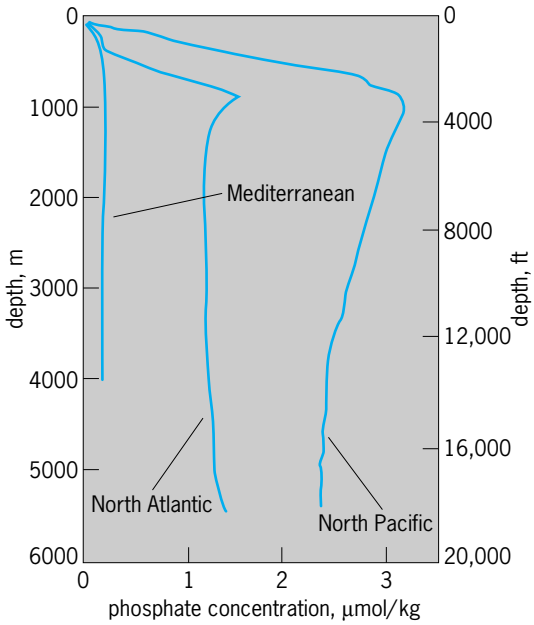

Fig. 1. Vertical distribution of phosphate (PO_4^{3-}) at three stations to show a range of characteristics found in major oceanic regions. North Atlantic is from a region in the central North Atlantic northeast of Bermuda, and North Pacific is from the central North Pacific northeast of Hawaii.

nutrients (along with carbon and smaller amounts of other nutrients) to form their body structures; these in turn provide the base of the food chain from zooplankton all the way up to the largest fish and whales. Particulate matter, including plant cells, feces, and other debris, sinks below the surface and is eaten by deeper-living plankton, fish, bacteria, and other organisms throughout the water column and on the bottom. The metabolism of these deeper living organisms results in the release and enrichment of the deeper water with carbon dioxide, phosphate, and the other nutrients, and the consumption of oxygen. *See* FOOD WEB; MARINE ECOLOGY; PHYTOPLANKTON.

The difference in the concentrations of the nutrients between the Mediterranean and the Atlantic is explained by the circulation of seawater. Surface water from the Atlantic flows into the Mediterranean, and is already depleted in nutrients. The return flow through the Strait of Gibraltar is somewhat deeper, underneath the surface inflow, so that the small accumulation of nutrients in the deeper water tends to be removed. The Pacific Ocean provides the reverse example. Deep water from the Atlantic flows into the Pacific, carrying its load of nutrients; in the Pacific it upwells, and the return flow to the Atlantic includes a considerable amount of surface water, depleted in nutrients through the activities of the biota that caused the transport of the nutrients back to the deep water in the Pacific. The nutrients therefore tend to be retained and to build up in the Pacific; the same is also true of the Indian Ocean. In many cases, more detailed analyses also provide analogous explanations of the structure seen in the vertical profiles. *See* ATLANTIC OCEAN; INDIAN OCEAN; MEDITERRANEAN SEA; PACIFIC OCEAN.

Surface water concentrations of nutrients vary considerably throughout the ocean, depending on the detailed physical processes that bring water from some depth to the surface. These upwelling regions are especially notable, for example, off the western shores of North and South America, off the southwestern shore of Africa, and in the complex current systems flowing along the Equator. Accordingly, these are regions of high biological productivity and also regions where there is a very strong transport of nutrients from the surface downward as the debris sinks. Shallow water regions tend also to be very productive, for two reasons. First, the interaction of ocean currents with these shallow regions causes turbulence that enhances the mixing of surface and deeper nutrient-rich water. Second, the sinking debris cannot go very deep and much is metabolized on the shallow bottom, so that the resulting nutrients are easily mixed into the surface water again. *See* SEAWATER FERTILITY; UPWELLING.

In general, an inverse relationship prevails between the concentration of oxygen (O_2) in the deep water of the ocean and the concentrations of phosphorus (P) as phosphate [and nitrogen (N) as nitrate] and carbon dioxide (CO_2). The reason is that, on average, the debris that sinks to depth has a generally uniform composition, and the nutrients are released in proportion to the oxygen used in metabolism of the debris. The quantitative relationship is expressed in the Redfield ratio: $175O_2$:122C:16N:1P; that is, for every 175 molecules of oxygen that are utilized, 122 atoms of carbon (as carbon dioxide), 16 atoms of nitrogen (as nitrate), and 1 atom of phosphorus (as phosphate) are released. Variations in this ratio from place to place appear to be only a few percent. Additional carbon dioxide, in the form of carbonate ion (CO_3^{2-}), is released by particles of calcium carbonate in places where these dissolve.

In addition to phosphorus, nitrogen, and carbon, organisms need a number of other substances in small amounts. For example, some metals are needed, such as iron, copper, molybdenum, zinc, and cobalt, with iron in the greatest amount. Iron is quite insoluble in seawater, however, and is much depleted in the surface water of the open ocean. There are a number of places in the ocean, generally far from land, where the other nutrients are available, but the growth of plankton is limited by the availability of iron. In these cases the surface depletions of phosphate and nitrate are less evident.

Many trace elements show vertical distributions that, to varying degrees, resemble the vertical distributions of phosphate. For those that are known to be plant nutrients, these distributions are easily explicable, as they must be carried down with the sinking debris from life at the surface. In other cases it appears that some elements may be simply adsorbed to the surfaces of sinking particles, carried down, and released at depth. The vertical profiles of concentration of some elements exhibit quite different shapes; in these cases other processes must be called upon to explain them.

Carbon dioxide. The chemistry of carbon dioxide in seawater is complicated and important, not least

because the oceans are taking up a considerable part of the carbon dioxide produced by the burning of fossil fuels and by other human activities. When carbon dioxide dissolves in seawater, it enters into a number of chemical equilibrium reactions. The equilibrium constants of all such reactions are affected by temperature and other parameters. Of particular importance are those reactions that are strongly affected by the temperature and salinity of the water and by hydrostatic pressure, and accurate knowledge of all these parameters is important in assessing the chemical state of the carbonate system. Because of the participation of the hydrogen ion (H^+) in some reactions, the concentration of this ion (expressed usually as the pH) is an important indicator of the state of the system. When the concentration of carbon dioxide increases, the pH drops; and when carbon dioxide decreases, the pH rises. Accordingly, as anthropogenic carbon enters the ocean it causes a small decrease in the pH. The effect is, so far, much smaller than the changes commonly caused by the growth of plants (increase in pH) and by the respiration of all living things (decrease in pH).

An additional complication is the precipitation of calcium carbonate by planktonic plants and animals (as well as by many types of organisms on coral reefs). The solubility of calcium carbonate is also strongly affected by temperature, salinity, and pressure. The shells and other bits of calcium carbonate formed in the plankton also sink, directly or in fecal pellets, out of surface layers and tend to dissolve at depth. In many places they dissolve completely because of the combined effects of pressure and the acidification of the deep water by respiratory carbon dioxide. The concentration of carbon dioxide in deep water is therefore increased, relative to the surface (**Fig. 2**). As with the nutrients, this concentration increases from the North Atlantic to the South Atlantic, through the South Pacific to the North Pacific. *See* CHEMICAL EQUILIBRIUM; HYDROGEN ION; PH.

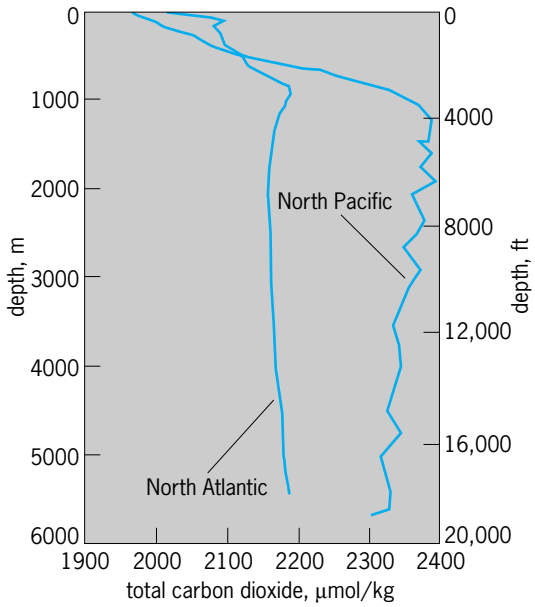

Fig. 2. Vertical distribution of total carbon dioxide concentrations at locations in the North Atlantic and North Pacific. Below depths of a few hundred meters, the distributions are generally characteristic of fairly large regions in the two oceans.

TABLE 5. Partial list of radioactive nuclides in seawater*

Isotope	Principal source
Potassium-40 (^{40}K)	Primordial
Rubidium-87 (^{87}Rb)	Primordial
Uranium-238 (^{238}U)†	Primordial
Uranium-235 (^{235}U)†	Primordial
Uranium-234 (^{234}U)†	Decay of ^{238}U
Strontium-90 (^{90}Sr)	Bomb explosions
Cesium-137 (^{137}Cs)	Bomb explosions
Carbon-14 (^{14}C)	Cosmic rays in atmosphere and bomb explosions
Hydrogen-3 [tritium] (^{3}H)	Cosmic rays in atmosphere and bomb explosions

*The total number detected is well over 40, but those not listed are present in very small concentrations.
†Nuclides of uranium that are the parents of long series of radioactive decay products.

Radioactive materials. Most radioactivity in seawater is due to nuclides present since the world began (**Table 5**). The distributions of many of the radioactive nuclides resulting from the decay of uranium isotopes and those produced by cosmic-ray-induced processes in the atmosphere, as well as those produced artificially and deposited into the ocean after nuclear explosions in the atmosphere, have been utilized to help trace mixing and transport processes in the ocean, and have helped provide quantitative understanding of the rates of these processes. *See* RADIOACTIVITY; RADIOISOTOPE; URANIUM.

Changes in composition. The overall chemical composition of the ocean results from the balance of numerous processes. These processes include (1) the weathering of exposed rocks on continents and islands, with the transport of the resulting materials to the oceans; (2) the uptake and release of various substances when seawater comes in contact with hot spots on the sea floor; (3) exchange with the atmosphere, mostly gases; (4) uptake and release of many substances by organisms; (5) sinking of particles from surface water into deep water; (6) chemical exchanges between seawater and sediments, and chemical reactions occurring within the sediments; and (7) precipitation of complex mixtures of salts in places where seawater can evaporate to dryness, a process that has varied greatly in magnitude over the course of geological time. *See* HOT SPOTS (GEOLOGY).

It should not be assumed that all these processes are constant over the span of geologic time. Known changes in climate, rates of weathering of the continents, rates of sea-floor spreading, and rearrangements of geological structures must have caused some changes in the chemistry of the ocean. Most of these changes would not likely be perceptible on the time scale of human history. However, the activities of human civilization since the beginning of the industrial revolution have resulted in detectable chemical signals throughout most of the ocean, especially in surface water.

Michael E. Q. Pilson

Distribution of Properties

The distribution of physical characteristics in the ocean is principally the result of radiation (of heat), exchange with the land (of water and salt), exchange with the atmosphere (water and heat), and mixing and stirring processes within the ocean.

Temperature. The ocean is heated by the sun, but very little of its radiation penetrates more than a few meters beneath the surface. Essentially all the heat lost by the ocean is across the air-sea interface by long-wave radiation, evaporation, and conduction. Although the heat losses balance the heat gains over an annual cycle, they seldom balance at any one location. The surface ocean gains more heat than it loses equatorward of about 35°, and it loses more heat than it gains poleward of 35°. As a consequence, the sea surface temperature is high (more than 28°C or 82°F) in equatorial regions and low (less than 1°C or 34°F) in polar regions. Below a shallow surface layer of a few hundred meters, the temperature structure of the ocean is a result of this continuing poleward transfer of the order of 10^{15} W of heat energy.

Salinity. Various dissolved solids have entered the sea from the land and have been so mixed that their relative amounts are everywhere nearly constant. For well over 95% of the ocean, the salt content is 3.43–3.51% by weight. Oceanographers routinely measure salinity to a few thousands of a percent, and salinity is reported in parts per thousand (‰). Considerable insight in oceanic processes can be achieved by examining changes in seawater salinity distributions of the order of ±0.02‰. In the middle latitudes the evaporation of water exceeds precipitation, and the surface salinity is high; in low and high latitudes precipitation exceeds evaporation, and dilution reduces the surface salinity. The Atlantic Ocean is more saline than the Pacific and Indian oceans.

Density. The density of seawater varies with temperature, salinity, and pressure. It can vary horizontally only in the presence of currents, and hence its distribution depends closely upon the structure of the flow.

Pycnocline, thermocline, and halocline. The ocean is stratified. The upper layer, in contact with the atmosphere, is warmed and cooled by exchange of heat through the surface and freshened and made more saline by rainfall and evaporation. Over the ocean this layer varies in depth from less than 100 m (330 ft) to more than 300 m (1000 ft). During summer heating, a thin cap of warmer water may be formed at the top, disappearing in fall and winter by cooling and mixing with the underlying water.

Beneath the upper layer, there is a strong increase in density, called the pycnocline. This increase may include a sharp drop in temperature (the thermocline) or a rise in salinity (the halocline). In most middle latitudes, both temperature and salinity decrease. In high latitudes, both temperature and salinity may increase downward if the vertical gradient of salinity is strong enough.

Characteristics beneath the surface. The oceans of the world are comprised of layers (**Figs. 3–5**). Because seawater is slightly compressible, it is made warmer with pressure. The potential temperature which is defined as the temperature that a parcel of water would have if it were acted upon by atmospheric pressure alone. Near 4000 m (14,000 ft), where pressures exceed 400 atm (41,000 kilopascals), the actual temperatures are a few tenths of a degree higher than the potential temperatures.

Abyssal waters. The densest waters of the open ocean are formed in high latitudes by cooling of some of the waters that have flowed poleward from the more saline regions. The principal source of the abyssal waters is the Weddell Sea, with some contribution from the Ross Sea, and some from the Norwegian-Greenland Sea passing over the shallow sill between Greenland and Iceland into the North Atlantic Ocean (Fig. 3). The dense waters formed in this fashion are cold and moderate in salinity. They flow equatorward along the western boundaries of the oceans and become progressively warmer, more saline, and less dense by mixing with the overlying water.

The abyssal waters formed in the Weddell and Ross seas are not formed by convection to the bottom in the open ocean. Though these seas are at freezing temperature most of the year, the salinity in the upper layer is so low (about 34.0 parts per thousand) that even at the freezing point the surface waters are not as dense as the underlying waters. The densest waters are formed on the continental shelves, where two processes raise the salinity. Higher-salinity waters from lower latitudes extend southward beneath the surface and raise the salinity on the continental shelves. Leaching out of brine from the ice adds significant amounts of salt to the shallower layer on the shelves. At these raised salinities and the freezing point, the shelf waters become dense enough to spill down the continental slopes to abyssal depths. These waters mix with the open-ocean waters as they pour down. As they reach the bottom, their temperature is near −0.25°C (31.6°F). Their salinity is 34.66–34.7 parts per thousand, far higher than any of the open-ocean salinity values of the Antarctic Ocean.

Overturn to the bottom. Overturn to the bottom does not occur in the open ocean, but takes place only in adjacent seas. The warm and saline waters of the open North Atlantic Ocean flow into the Norwegian-Greenland Sea, where the intense cooling increases their density and convection occurs. The passages from the Norwegian-Greenland Sea to the open ocean are narrow and shallow; and where the dense Norwegian-Greenland Sea waters spill back into the Atlantic, they mix with the less dense waters and their density is reduced. The waters that pour down from the Denmark Strait, between Greenland and Iceland, and still dense enough to reach the bottom in the North Atlantic Ocean, but their density is not as great as those at the bottom of the Weddell Sea. The mixture from the overflow between Scotland and Iceland is less dense and reaches only about 3000 m (10,000 ft). Their major contribution is to the middepth water.

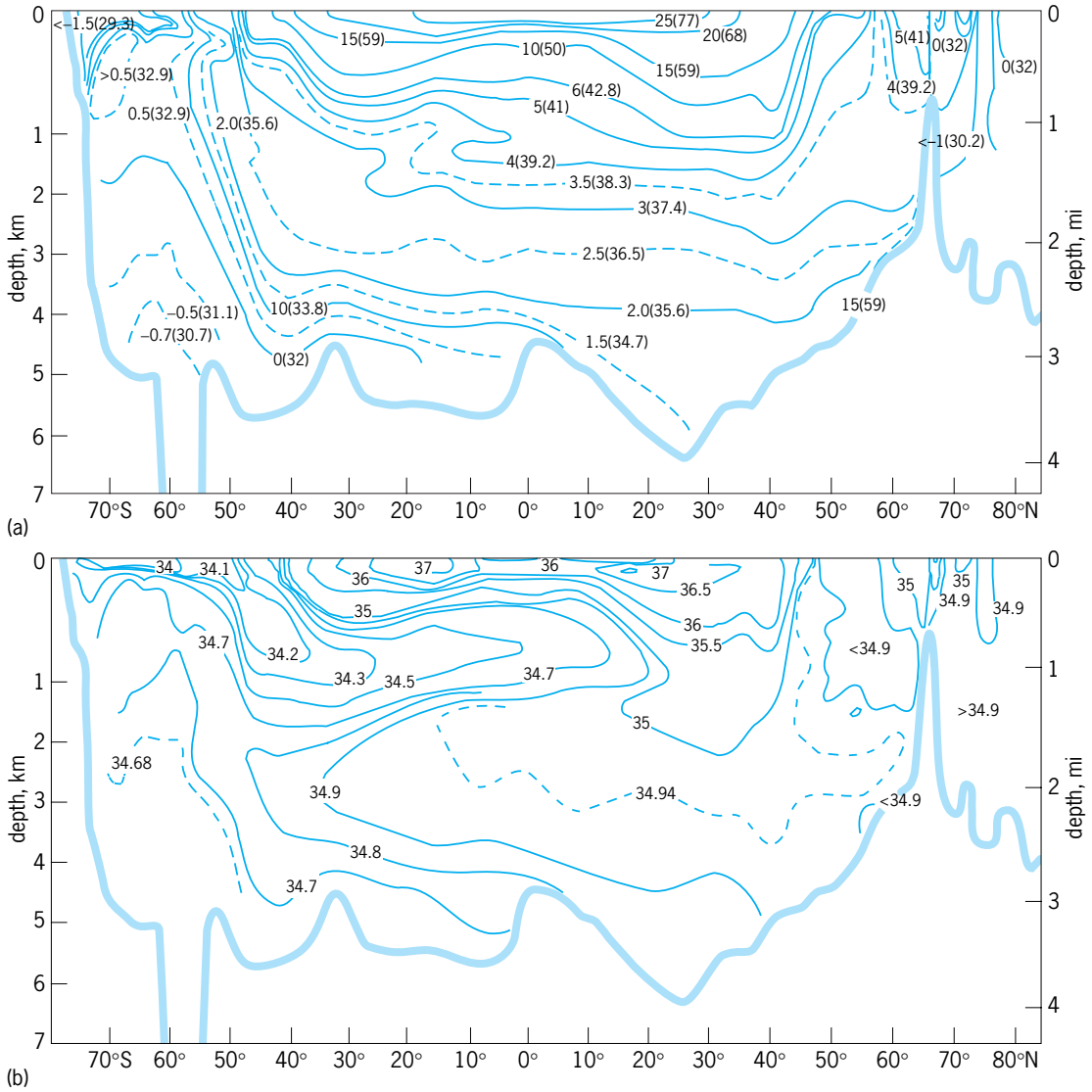

Fig. 3. Layers in seawater in a section through the Atlantic Ocean from the Greenland Sea (right) to the Weddell Sea (left). The thick line represents depth. (*a*) Potential temperature; contour values are in °C (°F). (*b*) Salinity; contours are in parts per thousand.

The Mediterranean Sea is an evaporation basin. Water from the Atlantic flows in at the surface to make up the loss to the atmosphere. Evaporation raises the salinity so high (more than 38 parts per thousand) that even at a temperature of 12°C (36°F) this water is denser than any of the Atlantic waters, and the Mediterranean Sea overturns to the bottom. Then the water pours out through the Strait of Gibraltar, beneath the inflowing Atlantic water, and spills down the slope. Rapid mixing with the Atlantic waters and the water's low compressibility (from its high temperature) keep it from penetrating to the bottom. It provides the very high salinity and temperature that characterize the middepth Atlantic Ocean.

Middepth water. Between the upper layer and the dense abyssal layer, the middepth waters are also supplied from the higher latitudes. Warm and extremely saline water from the Mediterranean Sea makes the North Atlantic the warmest and saltiest of the oceans at depths of 1000–2500 m (3300–8300 ft; Figs. 3–5). Between Iceland and Scotland, warm and saline water from the near-surface waters of the eastern North Atlantic meets the cold and extremely dense water from the Norwegian Sea. Their mixture produces a denser layer of warm and saline water. This water flows westward south of Iceland and Greenland and into the Labrador Sea, where it turns equatorward along the western boundary.

These three North Atlantic sources (Greenland Sea, Mediterranean Sea, and Norwegian Sea–Northeastern Atlantic) account for the characteristics of much of the middepth water in the world ocean. The high salinity in this layer extends southward through the Atlantic (Fig. 3) at depths of 1500–2500 m (5000–8300 ft). It rises with the Antarctic Circumpolar Current near 45 to 60°S latitude and turns westward into the South Indian Ocean (Fig. 4). It continues through the Pacific Ocean (Fig. 5), where part of it turns northward and part continues

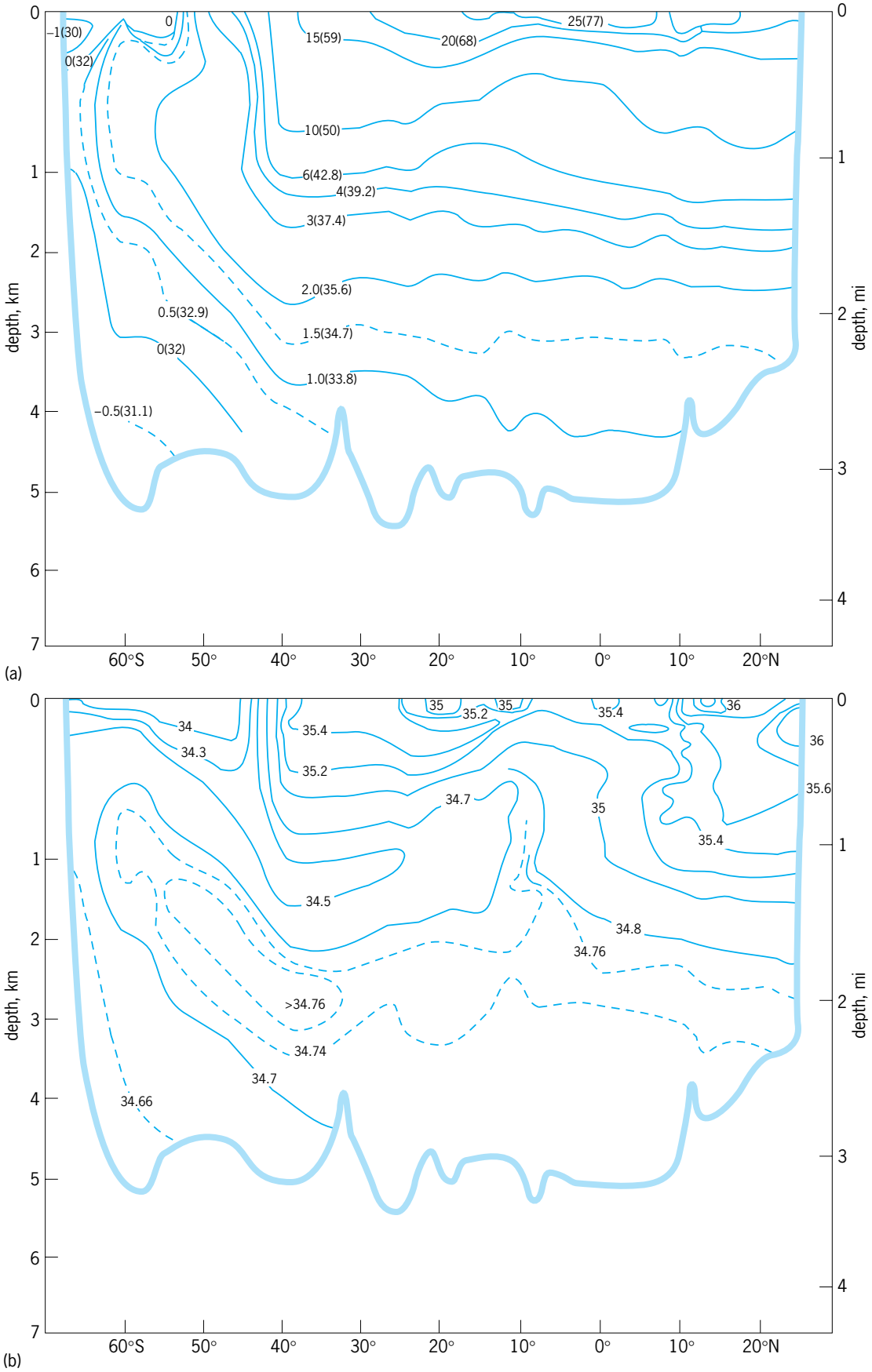

Fig. 4. Layers in seawater in a section through the western Indian Ocean from India (right) to Antarctica (left). The thick line represents depth. (*a*) Potential temperature; contour values are in °C (°F). (*b*) Salinity; contours are in parts per thousand.

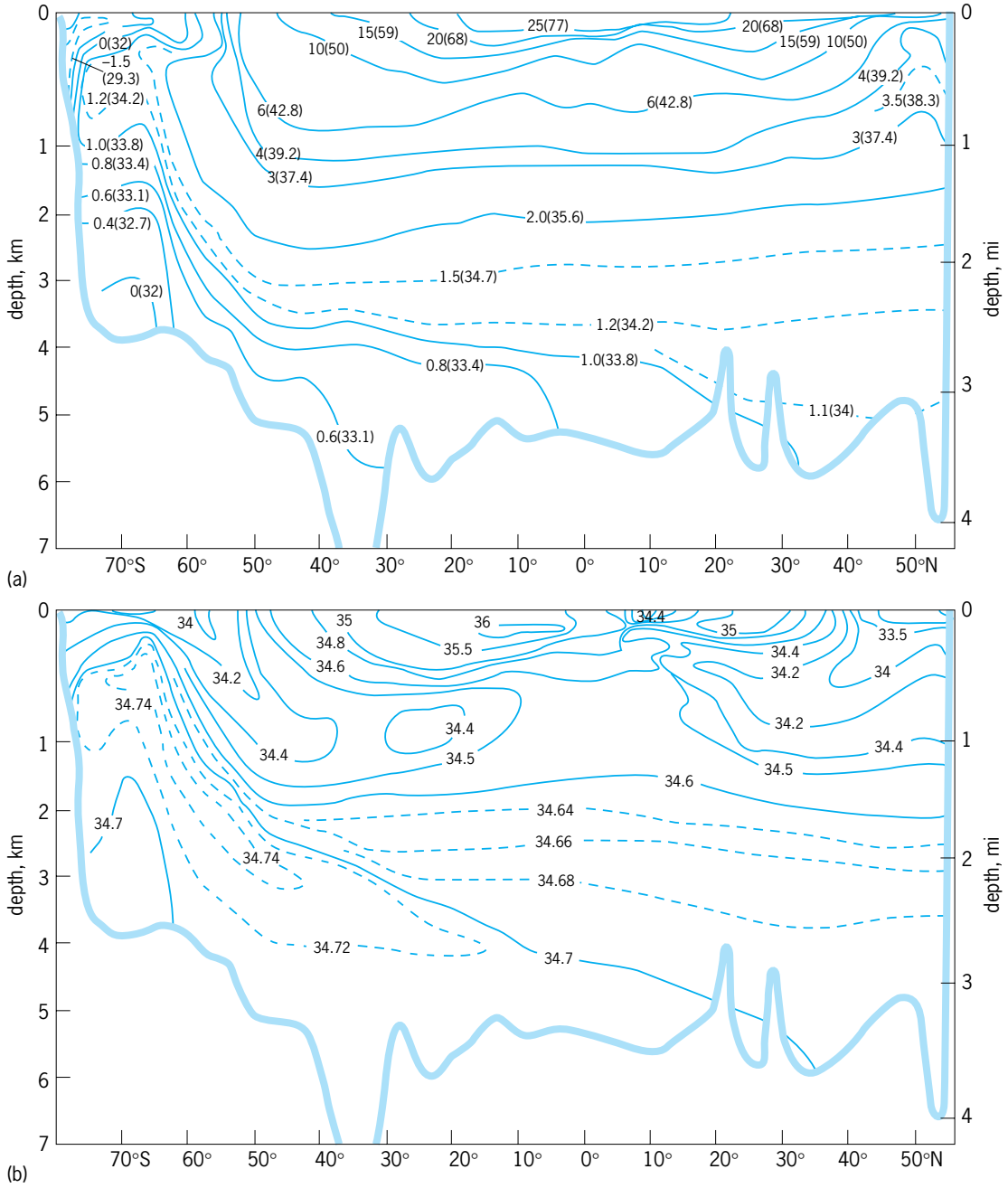

Fig. 5. Layers in seawater in a section through the Pacific Ocean from the Aleutian Islands (right) to the Ross Sea (left). The thick line represents depth. (*a*) Potential temperature; contour values are in °C (°F). (*b*) Salinity; contours are in parts per thousand.

eastward to return, in a diminished form, to the Atlantic Ocean.

Layer of low salinity. The surface waters in high latitudes are made colder and less saline and dense than the adjacent midlatitude waters by exchange of heat and water with the atmosphere. As they extend equatorward, they descend below the upper layer because of their higher density. They are recognized near 800–1000 m (2700–3300 ft) as layers of low salinity from the high southern latitudes of all three oceans, and also from the north in the Pacific.

Layers of high salinity. The highest salinity in the open ocean is found in the upper layer near the tropic circles, where evaporation exceeds precipitation. Vertical mixing through the pycnocline takes place only slowly, and the high surface salinity does not penetrate very well. But these high salinities do spread laterally beneath the less dense surface waters of lower latitudes. In all three oceans (Figs. 3–5), shallow layers of high salinity can be seen extending equatorward from the great high-salinity cells around the tropic circles.

Joseph L. Reid

Thermal Structure

There is a rich spectrum of temperature variations in the ocean. Those with lateral scales of thousands

of kilometers affect the horizontal density distribution and, hence, have a crucial role in determining the geostrophic currents. At the smallest scales (on the order of a centimeter), the temperature variations are usually so small that they have no effect on the density. In these cases the temperature behaves as a passive tracer that shows the stirring of water parcels. Between these extreme spatial scales, temperature can be a tracer of the motion or an active participant in the motion via its effect on the density.

Since temperature is relatively easy to measure, it has been a popular way to study the processes which stir and mix the oceans. Two general types of measurements predominate: (1) vertical profiles of temperature or its gradient, either single profiles or repeated profiles from a steaming or drifting vessel; and (2) horizontal sections of isotherm depth, derived from vertical arrays of many thermistors towed through the water. Operational problems are significant, and the data are always limited by the fact that the vertical profiles are spaced horizontally and the towed arrays have fixed and finite vertical spacing between the sensors.

Fine structure. Temperature generally decreases with depth in the ocean. This decrease is neither steady nor monotonic. The detailed fluctuations with vertical scales on the order of meters are called fine structure. Because of the stratification of the ocean, these features are relatively thin and wide; that is, they have a large aspect ratio. Fine-structure features extend horizontally 100–1000 times (or more) further than they extend vertically.

A parcel of water that is warmer or colder than the water found both above and below must have come into the water column laterally. Known as intrusions, they are quite common (and especially dramatic) in frontal regions where waters with substantially different salinity and temperature characteristics meet. Intrusions 20 m (66 ft) thick with horizontal scales of several kilometers are observed in coastal regions. The rate of change of temperature with depth varies greatly in magnitude and even changes sign. Intrusions are driven by the horizontal pressure gradients that arise from the horizontal density differences. It is possible to have intrusions without a temperature maximum or minimum. These cannot be recognized from the temperature alone, but require examination of the salinity, the density, or some other property such as oxygen.

Another form of temperature fine structure is created by vertical mixing. As the mixing events in the ocean are rarely thorough enough to produce a truly homogeneous patch of water, the result is a relatively low gradient in the patch and relatively high gradients at the edges. A special case of fine structure due to mixing is associated with double diffusive convection, and leads to very strong layering of almost isothermal water with a horizontal extent of many kilometers.

Fine structure is also created by internal waves. Waves produce differential motion with depth, producing alternating regions of horizontal convergence and divergence. Even in an ocean with a smooth temperature profile that decreases monotonically, internal waves can produce local regions of higher and lower gradient. As a result, a smooth temperature profile develops kinks but not inversions. The irregular fine-structure features in the profile can be observed to propagate vertically.

Fine structure due to internal wave motion appears and disappears with the arrival and departure of the wave energy. Hence, it is called reversible fine structure. Lateral intrusions and vertical mixing are not reversible processes; they generate irreversible fine structure. Irreversible fine structure is left behind after the processes that produced it have finished their work. Measurements indicate that in most regions free from intrusions the reversible fine structure dominates over the irreversible fine structure, which would arise from vertical mixing. In regions with intrusions, no general statement can be made.

Microstructure. At scales smaller than the fine structure, from less than 1 m (3.3 ft) down to a few millimeters, the temperature fluctuations are called microstructure. This microstructure is usually due to turbulent mixing processes, with temperature acting as a tracer of the motion. The aspect ratio of these features (ratio of horizontal to vertical scale) decrease with the vertical scale. It might be as large as 100 for 1-m vertical scales and shrink toward 1 at the centimeter scale. The magnitude of the temperature variations decreases with decreasing size.

The magnitude of the fluctuations is determined by the local temperature gradient as well as the size and intensity of the mixing event. Thus, there is no simple relation between the temperature variations and the motion in the water. Molecular thermal diffusivity smooths out the inhomogeneities at all scales, but this effect is most noticeable at scales of a few centimeters and smaller. At these smallest scales, the time for the temperature gradient to diffuse away is from a few seconds (1 mm) to 10 min (2 cm). Therefore, each patch of microstructure is transient, lasting only slightly longer than the processes which produce it.

The detailed features of the thermal structure portend similar features in other parameters such as salinity, density, oxygen, and nutrients. The physical processes that produce fine structure and microstructure also affect the distribution of phytoplankton, small herbivores, and carnivores. However, there are many other processes that affect the chemical and biological distributions in addition to the physical processes that dominate the thermal structure.

Thomas Osborn

Bibliography. J. R. Apel, *Principles of Ocean Physics*, 2d ed., 2001; A. E. Gill, *Atmosphere-Ocean Dynamics*, 1982; M. N. Hill (ed.), *The Sea*, vol. 1: *Physical Oceanography*, 1982, vol. 2: *Composition of Seawater*, 1982; J. A. Knauss, *Introduction to Physical Oceanography*, 2d ed., 1996; E. B. Kraus and J. Businger, *Atmosphere-Ocean Interaction*, 1994; S. K. Majumdar et al. (eds.), *The Oceans: Physical-Chemical Dynamics and Human Impact*, 1994; H. Medwin, *Sounds in the Sea: From Ocean*

Acoustics to Acoustical Oceanography, 2005; F. J. Millero and M. L. Sohn, *Chemical Oceanography*, 2d ed., 1996; M. E. Pilson, *Introduction to the Chemistry of the Sea*, 1998; R. W. Spinrad, K. L. Carder, and M. J. Perry, *Ocean Optics*, 1997; T. Teramoto (ed.), *Deep Ocean Circulation: Physical and Chemical Aspects*, 1993.

Seawater fertility

A measure of the potential ability of seawater to support life. Fertility is distinguished from productivity, which is the actual production of living material by various trophic levels of the food web. Fertility is a broader and more general description of the biological activity of a region of the sea, while primary production, secondary production, and so on, is a quantitative description of the biological growth at a specified time and place by a certain trophic level. Primary production that uses recently recycled nutrients such as ammonium, urea, or amino acids is called regenerated production to distinguish it from the new production that is dependent on nitrate being transported by mixing or circulation into the upper layer where primary production occurs. New production is organic matter, in the form of fish or sinking organic matter, that can be exported from the ecosystem without damaging the productive capacity of the system.

An example illustrates the difference between fertility and productivity. Waters around Antarctica during the austral summer have dense blooms of diatoms and large swarms of krill, the shrimplike zooplankton that feed on the diatoms and in turn are eaten by whales, seals, fish, and penguins. Krill are harvested in Antarctic waters, which contain the only unexploited fishery resource that can significantly increase the supply of protein from the sea (**Fig. 1**). In the austral summer, both primary production, the synthesis of organic matter by photosynthesis, and secondary production, the growth of krill, proceed at high rates. In contrast, during the dark and stormy Antarctic winter there is little sunlight for photosynthesis, and winter storms deeply mix the surface waters. Primary and secondary production proceed at undetectably low rates. While productivity of Antarctic waters varies throughout the year from very low to very high, it is correct to refer to the Antarctic as a fertile region because as a region it has the potential to support high levels of biological activity.

Another analogy that clearly illustrates the distinction between fertility and productivity deals with soil. The rich black soil of the Corn Belt in Illinois and Iowa is always fertile, but in winter the actual productivity of the soil is very low because of low temperatures, short day lengths, and snow cover. *See* BIOLOGICAL PRODUCTIVITY.

The potential of the sea to support growth of living organisms is determined by the fertilizer elements that marine plants need for growth. Fertilizers, or inorganic nutrients as they are called in oceanography, are required only by the first trophic level in the food web, the primary producers; but the supply of inorganic nutrients is a fertility-regulating process whose effect reaches throughout the food web. When there is an abundant supply to the surface layer of the ocean that is taken up by marine plants and converted into organic matter through photosynthesis, the entire food web is enriched, including zooplankton, fish, birds, whales, benthic invertebrates, protozoa, and bacteria. Since the level of seawater fertility is initially set by the primary producers' response to nutrient conditions, this article is limited to the geographic and temporal variations in nutrients and primary productivity. *See* DEEP-SEA FAUNA; FOOD WEB; MARINE FISHERIES.

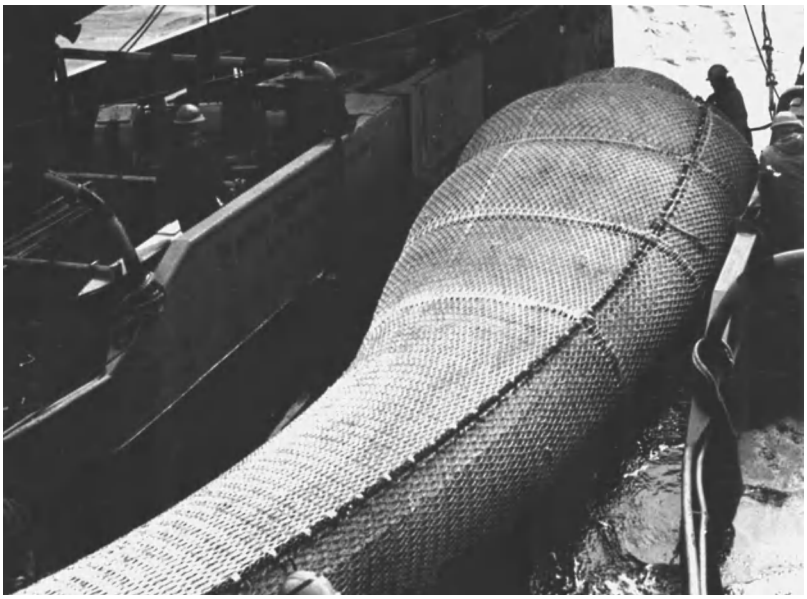

Fig. 1. A 17-ton (15-metric-ton) catch of krill harvested in Antarctic waters. (*Photograph by Dr. Gotthilf Hempel*)

Nutrients. The elements needed by marine plants for growth are divided into two categories depending on the quantities required: The major nutrient elements that appear to determine variations in ocean fertility are nitrogen, phosphorus, and silicon. The micronutrients are elements required in extremely small, or trace, quantities including essential metals such as iron, manganese, zinc, cobalt, magnesium, and copper, as well as vitamins and specific organic growth factors such as chelators. Knowledge of the fertility consequences of variations in the distribution of micronutrients is incomplete, but consensus among oceanographers is that the overall pattern of ocean fertility is set by the major fertilizer elements—nitrogen, phosphorus, and silicon—and not by micronutrients.

Nitrogen, phosphorus, and silicon are required for plants in relatively large quantities because these elements are building blocks of living material. While the other elements in proteins, carbohydrates, and fats such as carbon, oxygen, or sulfur are necessary to living organisms, seawater contains a large excess of them in proportion to the demands of marine plants in the synthesis of new organic matter. Hence the supply of carbon, sulfur, or oxygen does not

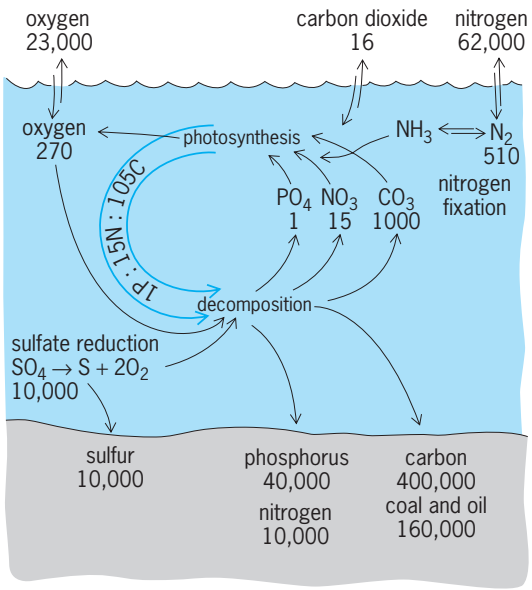

Fig. 2. Biochemical cycle of nutrient elements in the sea. Numbers represent the quantities of elements relative to phosphorus in the ocean and show the excess abundance of carbon and sulfur in comparison with phosphorus and nitrogen. (*After A. C. Redfield, The biological control of chemical factors in the environment, Amer. Sci., 46:205–221, 1958*)

determine variations in ocean fertility (**Fig. 2**). Understanding the pattern of fertility requires knowledge of the pattern of nitrogen, phosphorus, and silicon distribution in the sea.

The nutrient requirement for nitrogen is met in the ocean by forms of fixed inorganic nitrogen such as nitrate, nitrite, or ammonium and by dissolved organic nitrogen species such as urea or amino acids. The nitrate transported across the thermocline into the upper layer supports most of the new primary production in the ocean, while organic nitrogen recycled within the upper layer supports regenerated production. The requirement for phosphorus is met by the phosphate ion and that for silicon by ions of silicate and silicic acid.

Deep waters of the ocean beneath the permanent thermocline contain high concentrations of nitrate, phosphate, and silicate—dissolved inorganic nutrients maintained by a balance of processes. A steady but very small flux of sinking organic particles, especially zooplankton fecal pellets, into the subthermocline water exports material from the upper layer to deep water. In deep water, nitrogen, phosphorus, and silicon are released as the organic matter is decomposed by bacteria and protozoa; the released elements pass from particulate phase back to dissolved phase and are eventually remineralized to the stable terminal oxidation products of nitrate, phosphate, and silicate. In the deep ocean there is no light for photosynthesis, so the abundant inorganic nutrients cannot be taken up by plants; the fertilizer elements accumulate in solution until the nutrient-rich deep water returns to the surface by circulation or mixing. When and where subsurface water is transported to the surface layer, nutrients are brought together with light, and the stage is set for enhanced primary production and greatly increased fertility.

Oceanographers express the atom ratios of inorganic nutrients in seawater in terms of concentrations in microgram-atoms (of N or P) per liter of seawater or the more typical chemical units of micromoles (of nutrient ions) per kilogram of seawater. Since nitrate, phosphate, and silicate are monoatomic ions, the two conventions are quantitatively equal. Deep waters of the Pacific and Indian oceans have a nitrate concentration of about 40 μmol/kg, while Atlantic deep water has half that. Surface waters having a nitrate concentration of more than 10 μmol/kg are highly fertile; those having less than 0.1 μmol/kg are biological deserts. The relative abundances of phosphate and silicate vary in proportion to nitrate.

Primary producers. Two types of marine plants carry out primary production in the ocean: microscopic planktonic algae collectively called phytoplankton, and benthic algae and sea grasses attached to hard and soft substrates in shallow coastal waters.

The benthic and planktonic primary producers are a diverse assemblage of plants adapted to exploit a wide variety of marine niches; however, they have in common two basic requirements for the photosynthetic production of new organic matter: light energy and the essential elements of carbon, hydrogen, nitrogen, oxygen, phosphorus, sulfur, and silicon for the synthesis of new organic molecules. These two requirements are the first-order determinants of photosynthetic growth for all marine plants and, hence, for primary productivity everywhere in the ocean.

Phytoplankton. The term phytoplankton refers to a wide variety of planktonic plants, including diatoms, dinoflagellates, coccolithophorids, microflagellates, blue-green algae, and even ciliates that have endosymbiotic chloroplasts. Diatoms are abundant in the more fertile upwelling and coastal regions of the ocean, and are probably the major group responsible for the new production that provides material to the pelagic food web and to the bottom for consumption by the benthos (see **table**). Coccolithophorids and microflagellates maintain relatively constant but

Primary production and fish production in various provinces of the ocean

Province	Percent of ocean area	Area, mi^2 (km^2)	Average primary production, g C/(m^2)(yr)	Total primary production, tons/yr (metric tons/yr)	Number of trophic levels	Trophic efficiency, percent	Fish production, tons/yr (metric tons/yr)
Offshore ocean	90.0	125×10^6 (326×10^6)	50	17.9×10^9 (16.3×10^9)	5	10	18×10^5 (16×10^5)
Coastal zone	9.9	14×10^6 (36×10^6)	100	4.0×10^9 (3.6×10^9)	3	15	13×10^7 (12×10^7)
Upwelling areas	0.1	1.4×10^5 (3.6×10^5)	300	0.1×10^9 (0.1×10^9)	1.5	20	13×10^7 (12×10^7)

low levels of productivity in the less fertile offshore provinces of the ocean; in these areas microflagellates may have rapid specific growth rates. That is, the rate of synthesis of new cell material per unit of existing cellular material may be high, but since the biomass of phytoplankton in the offshore provinces is low, the absolute productivity is low when compared with more fertile areas.

Dinoflagellates are a functionally complex group and exploit a variety of marine niches. Some dinoflagellate species are slow-growing and persistent in tropical waters, while other species form concentrated but sporadic so-called red tide aggregations that may be poisonous to both marine life and humans; still others occur with diatoms, during nutrient-rich conditions. However, the niche most often associated with dinoflagellates is intermediate between that occupied by diatoms at the fertile end of the nutrient spectrum and that of coccolithophorids and microflagellates at the nutrient-poor end of the spectrum.

Highly fertile regions of the ocean have been studied frequently because important fisheries are concentrated there. The poisonous red tide dinoflagellate outbreaks have also received considerable study. While diatoms and dinoflagellates are major primary producers where the intensity of production is high, microflagellates and coccolithophorids are important in the large tropical and subtropical offshore province of the ocean where production is low (see table). The offshore province occupies 90% of the ocean area, so that the coccolithophorids and microflagellates make a significant contribution to the global carbon cycle even though they may not support important fisheries. *See* PHYTOPLANKTON.

Benthic plants. An advance in the understanding of ocean fertility has been recognition of the role played by benthic algae and sea grasses. These plants occur in dense stands in shallow waters and carry out primary production at extremely high rates throughout the year. Attached algae and grasses are efficient at nutrient scavenging since they remove dissolved nutrients from many volumes of water as it passes over them. Most of the production by attached algae and sea grasses is not grazed directly by herbivores but becomes detritus, and it is consumed in that form by a wide variety of deposit and suspension feeders. Examples of the important benthic primary producers are brown algae such as *Macrocystis*, *Nereocystis*, *Fucus*, or *Laminaria*, and the sea grasses *Zostera* and *Thallasia*.

Productivity measurement. Productivity is usually expressed as the increase in organic matter per unit area per unit time since light enters the marine environment as an areal function from the surface. This convention makes it possible to compare the productivity of the ocean directly with that of land. The units most frequently used are grams of carbon per square meter (g C/m^2). Primary production is measured over short time intervals of 1 day or less, producing data units of g C/(m^2) (day). For food web considerations these data are converted to annual production as g C/(m^2) (year), which introduces very large errors because oceanographers have not made daily measurements throughout annual cycles to determine the correct daily-to-annual conversion factors for different marine ecosystems.

Productivity conventionally is expressed in terms of carbon. However, there is no fundamental rationale for choosing carbon over nitrogen, phosphorus, or a cell property like protein. Carbon is convenient, however, because the most precise technique for measuring primary production uses the radioactive isotope of carbon. The excess inorganic carbon in seawater satisfies the requirement of the radioactive isotope of carbon. The excess inorganic carbon in seawater satisfies the requirement of the radiocarbon method that there be no detectable change in the precursor substrate during the measurement. With nitrogen, phosphorus, or silicon, the requirement that there be significant change in the substrate concentration is not often satisfied, so determination of the in-place rate of uptake of these elements is difficult or impossible with isotopic tracers. Conversely, production in terms of nitrogen, silicon, or phosphorus can be accurately measured by following the disappearance of these nutrients from seawater, a technique not possible with carbon.

In fertile areas primary production may be over 1.0 g C/(m^2)(day), but the sea grasses, kelp, or upwelling phytoplankton can exceed 10.0 g C/(m^2)(day). It is suspected that these high daily rates of production in fertile areas do not persist for long in the annual cycle, so oceanographers estimate that even highly fertile areas have an annual primary production of only about 300 g C/(m^2)(year). This crude estimate must suffice until time series studies are done. Provinces of the ocean that are permanently nutrient-poor due to strong thermal stratification have less than 0.1 g C/(m^2)(day) with little variation throughout the annual cycle. The yearly production is estimated to be 30–40 g C/(m^2)(year). While the coccolithophorids and microflagellates that are the dominant primary producers in this infertile area are important in the global flux of carbon, they are insignificant in the production of fish, as shown in the table. *See* ALGAE.

Patterns of fertility. The regions of the world's oceans differ dramatically in overall fertility. In the richest areas, the water is brown with diatom blooms, fish schools are abundant, birds darken the horizon, and the sediments are fine-grained black mud with a high organic content. In areas of low fertility, the water is blue and clear, fish are rare, and the bottom sediments are well-oxidized carbonate or clay. These extremes exist because the overall pattern of fertility is determined by the processes that transport nutrients to the sunlit upper layer of the ocean where there is energy for photosynthesis.

Upwelling. The fisheries of the world are concentrated in regions of enhanced fertility; it is not generally appreciated that the fertile food-producing areas of the ocean, like on land, are concentrated into a very small proportion of the total area. It has been estimated that 99% of the world's fish catch comes from

only 10% surface of the area of the ocean (see table), and of that, 50% comes from the 0.1% ocean area where upwelling carries nutrient-rich water to the surface. Upwelling on the west coasts of the Americas and Africa is a well-known process that enriches fertility, modifies the coastal climate, and has great economic impact because of the large upwelling-based fisheries. However, there are other processes that replenish nutrients to the depleted surface layer, but these processes are not nearly as well known as upwelling. *See* UPWELLING.

The annual cycle of heating and cooling of the Earth's surface in temperate and high latitudes drives the transport of nutrients into the upper layer of the ocean. When there is net negative heat flux to the ocean from the atmosphere, the surface water cools, becomes denser and unstable, and subsequently mixes with the underlying water. Mixing replaces nutrient-depleted surface water with nutrient-rich water and sets the stage for high productivity. When the heat flux becomes positive, the now nutrient-rich water is heated, becomes less dense, and forms a stable upper layer separated from the cooler deep water by a thermocline. In the stabilized upper layer, light for photosynthesis is available, and a rapid burst of phytoplankton growth occurs. This brief burst of high primary productivity by the diatom-dominated phytoplankton assemblages is the famous spring bloom of regions such as Long Island Sound, the Gulf of Maine, or the North Sea. The spring bloom in temperate zone seas depletes the upper layer of nitrate, phosphate, and silicate, causing a productivity decrease to a lower maintenance level until the next mixing event replenishes the nutrients.

Estuaries. Studies in lateral estuaries of the Chesapeake Bay and along the coast of Maine show that mixing during the strong tidal flows of the full and new moon generate short blooms in phase with the lunar cycle. In the waters around the British Isles, tidal mixing produces quasistationary blooms where the tidal currents hit subsurface banks or ledges and water mixes upward. Downstream from the wake-generating obstruction, heating restabilizes the upper layer, and blooms of phytoplankton exploit the available nutrients. Wake mixing over topographic features such as Georges Bank, Grand Banks, or Dogger Bank complements seasonal mixing and supports continuously high productivity in a manner analogous to the topographically fixed blooms of upwelling centers. An important portion of the world's fish catch comes from these banks. This yield is made possible by the physical transport of nutrients upward to replace the nitrogen, phosphorus, and silicon removed in the form of organic exports such as harvested fish. Physical processes drive both the seasonal heat-driven mixing and stabilization and the tidal mixing and stabilization; both processes are important to the continuing fertility of fishing areas such as Georges Bank.

Polar seas. In the polar seas a short burst of high production occurs in midsummer as increased sunlight becomes available for photosynthesis. At that time the water stabilizes slightly from solar heating, and the melting ice packs release fresh water that further stabilizes the surface layer. Phytoplankton blooms in both the Bering and Weddel Seas have some of the highest concentrations of phytoplankton ever recorded in the open ocean, but in both the Arctic and Antarctic the blooms are short-lived and very patchy in distribution. A brief burst of primary production supports a rich food web with large populations of higher-trophic-level organisms such as krill fish and marine mammals. The ecology of these highly fertile polar seas is not well understood. Because the valuable unexploited food resources of the polar areas merit careful management, it is important to have a basic understanding of these ecosystems.

A major difference in the productivity cycle of polar seas is that phytoplankton growth does not proceed to nutrient depletion, and productivity is regulated at the end of the brief summer by the lack of light, deep wind-driven mixing, and low temperatures. Polar and tropical waters are dramatic opposites in regulation of fertility. Polar waters always have an abundance of nutrients, so they are described as being continuously fertile; but there is light limitation of photosynthesis during most of the year. Tropical and subtropical offshore waters are permanently nutrient-depleted; light is never limiting in the upper layer, but the solar heating is so intense that thermal stratification is not broken down in the annual cycle. A biological desert is maintained because the subthermocline nutrient-rich water never mixes or circulates to the upper layer. (Dramatic exceptions to this are coastal and equatorial upwelling and the wake-stream mixing that occurs behind islands and atolls.) In the tropical and subtropical offshore waters that are permanently low in fertility, there is very efficient recycling of nutrients and high turnover rates. That is not true for the regenerated production of tropical and subtropical waters, which explains the fragile nature of tropical and subtropical ecosystems in comparison to the relatively resilient character of higher latitude marine ecosystems. Richard T. Barber

Size of populations and fluctuations. In temperate waters, fish often spawn at the same place at the same time of year. The larvae drift in a current from spawning ground to nursery ground, and the adults migrate from the feeding ground to the spawning ground. Adolescents leave the nursery to join the adult stock on the feeding ground. This migration circuit is based on the track of the larval drift, and the population is isolated by its unique time of spawning. The larvae suffer intense mortality, and numbers in the population are perhaps regulated naturally during the period of larval drift. The same principles apply to the upwelling areas in a more diffuse way. The tuna spawn nearly all year round across the North Pacific, and their larvae and juveniles drift in the major currents.

There are three basic methods of measuring the sizes of marine populations. The first is by census based on samples, that together constitute a known

fraction of the whole population, as in the case of sessile species such as shellfish, or of a particular age range, as in the case of fish that have pelagic eggs whose total abundance can be measured by fine-meshed nets hauled vertically through the water column. The second is by marking or tagging, in which a known number of marked individuals are mixed into the population and the ratio of marked to unmarked fishes is subsequently estimated from samples. This method may survive in lakes, but in the sea it is no longer used, because better estimates of stock may be obtained by the third method. From the catches in numbers at age in a fish population, sequential population analysis is used to obtain estimates of fishing mortality (or stock) by age. From the oldest age group in a year class, with educated guesses of fishing and natural mortality, the analysis proceeds back in the year class to the youngest, correcting the estimate as it proceeds.

The largest measured populations in numbers are of pelagic fish, particularly of the families of herring and sardines. The Atlantic herring (*Clupea harengus*) comprises a number of distinct stocks, the largest of which comprises 1 billion mature individuals. Each population ranges over hundreds or even thousands of kilometers in the northeast Atlantic.

Populations of bottom-living fish tend to fluctuate slowly over periods of 50–75 years. But those of pelagic fish, like the Atlantic herring, may fluctuate dramatically. Fish that spawn at a fixed season are obviously vulnerable to climatic change. During long periods there are shifts in wind strength and direction, with consequent delays or advancements in the timing of the production cycle. If the cycle is progressively delayed, the fish larvae, hatched at a fixed season, become short of food, and the populations decline with time. The match or mismatch of larval production to that of their food is probably the agent by which variation in the magnitudes of the recruiting year classes are determined. Climatic change affects the population during the period of the larval drift when isolation is maintained and when numbers are probably regulated. Recruitment is determined by the growth and mortality of the fish larvae and the juveniles, variations in which may be modified by a range of matches and mismatches in the production of larvae and juveniles to production of their food. *See* CLIMATE HISTORY. D. H. Cushing

Biological species and water masses. Biogeographical regions in the ocean are related to the distribution of water masses. Their physical individuality and ecological individuality are derived from partly closed patterns of circulation and from amounts of incident solar radiation characteristic of latitudinal belts. Each region may be described in terms of its temperature-salinity property and of the biological species which are adapted to all or part of the relatively homogeneous physical-chemical environment.

Cosmopolitan species. The discrete distributions of many species are circumscribed by the regions of oceanic convergence bounding principal water masses. Other distributions are limited to current systems. Cosmopolitan species are distributed across several of the temperature-salinity water masses or oceans; their wider specific tolerances reflect adaptations to broadly defined water types. No pelagic distribution is fully understood in terms of the ecology of the species.

A habitat is integrated and maintained by a current system: oceanic gyral, eddy, or current, with associated countercurrents. This precludes species extinction that could occur if a stock were swept downstream into an alien environment. The positions of distribution boundaries may vary locally with seasonal or short-term changes in temperature, available food, transparency of the water, or direction and intensity of currents.

Phytoplankton species are distributed according to temperature tolerances in thermal water masses, but micronutrients (for example, vitamin B_{12}) are essential for growth in certain species. The cells of phytoplankton reproduce asexually and sometimes persist in unfavorable regions as resistant resting spores. New populations may develop in prompt response to local change in temperature or in nutrient content of the water. Such species are less useful in tracing source of water than are longer-lived, sexually reproducing zooplankton species.

Indicator organisms. The indicator organism concept recognizes a distinction between typical and atypical distributions of a species. The orgin of atypical water is indicated by the presumed affinity of the transported organisms with their established centers of distribution.

Zooplankton groups best understood with respect to their oceanic geography are crustaceans such as copepods, and euphausiids, chaetognaths (arrow worms), polychaetous annelids, pteropod mollusks, pelagic tunicates, foraminiferids, and radiolarians. Of these the euphausiids are the strongest diurnal vertical migrants (60–200 ft or 200–700 m). The vertical dimension of euphausiid habitat agrees with the thickness of temperature-salinity water masses, and many species distributions correspond with the positions of the masses. In the Pacific different species, some of which are endemic to their specific waters, occupy the subarctic mass (such as *Thysanoessa longipes*), the transition zone, a mixed mass lying between subarctic and central water in midocean and between subarctic and equatorial water in the California Current (for example, *Nematoscelis difficilis*), the barren North Pacific central (such as *Euphausia hemigibba*) and South Pacific central masses (for example, *E. gibba*), the Pacific equatorial mass (such as *E. diomediae*), a southern transition zone analogous to that of the Northern Hemisphere (represented by *Nematoscelis megalops*), and a circumglobal subantarctic belt south of the subantarctic convergence (such as *E. lucens*).

Epipelagic fishes and other strongly swimming vertebrates are believed to be distributed according to temperature tolerances of the species and availability of food. However, distributions of certain bathypelagic fishes (such as *Chauliodus*) have been related to water mass. *See* SEAWATER. Edward Brinton

Bibliography. C. Lalli and T. Parson, *Biological Oceanography: An Introduction*, 2d ed., 1995; J. S. Levinton, *Marine Biology: Function, Biodiversity, Ecology*, 2d ed., 2001; K. H. Mann and J. R. N. Lazier, *Dynamics of Marine Ecosystems: Biological-Physical Interactions in the Oceans*, 2d ed., 1996; T. R. Parsons, M. Takahashi, and B. Hargrave, *Biological Oceanographic Processes*, 3d ed., 1984; A. C. Redfield, The biological control of chemical factors in the environment, *Amer. Sci.*, 46:205–221, 1958.

Sebaceous gland

A gland which produces and liberates sebum, a mixture composed of fat, cellular debris, and keratin. When the gland arises in association with a hair follicle, it forms a thickened outpushing from the side of the developing follicle near the epidermis. Central cells in these sebaceous glands form oil droplets within the cytoplasm. These cells disintegrate to liberate the sebaceous substance and are therefore of the holocrine type. The Meibomian or tarsal glands, within the tarsus or supporting plate at the edge of the eyelids, are sebaceous and complex tubuloacinous structures. The numerous separate glands open along the entire edge of the upper and lower lids. Retained secretions of the tarsal glands produce a cyst termed a chalozion or Meibomian cyst. *See* EPITHELIUM; GLAND.

Olin E. Nelsen

Second sound

A type of wave propagated in the superfluid phase of liquid helium (helium II) and in certain other substances under special conditions. Predicted independently by L. Tisza in 1938 and L. Landau in 1941, such waves were observed first by V. Peshkov in 1944 and have subsequently provided a rich source of information for the study and understanding of the superfluid state. The name is misleading since second sound is not in any sense a sound wave, but a temperature or entropy wave. In ordinary or first sound, pressure and density variations propagate with very small accompanying variations in temperature; in second sound, temperature variations propagate with no appreciable variation in density or pressure.

Two-fluid model. The two-fluid model of helium II provides further insight into the nature of second sound. In this model the liquid can be described as consisting of superfluid and normal components of densities ρ_s and ρ_n, respectively, such that the total density $\rho = \rho_s + \rho_n$. The superfluid component is frictionless and devoid of entropy; the normal component has a normal viscosity and contains the entropy and thermal energy of the system. As the temperature goes from 0 K to T_λ (the superfluid transition temperature), ρ_s goes from ρ to 0, and ρ_n from 0 to ρ. In a temperature or second-sound wave, the normal and superfluid flows are oppositely directed so that $\rho_s \mathbf{V}_s + \rho_n \mathbf{V}_n = 0$, where $\mathbf{V}_s$ and $\mathbf{V}_n$ are the superfluid and normal flow velocities. Thus a variation in relative densities of the two components, and hence a temperature fluctuation, propagates with no change in total density or pressure. In a first-sound wave, the two components move in phase, that is, $\mathbf{V}_n \cong \mathbf{V}_s$.

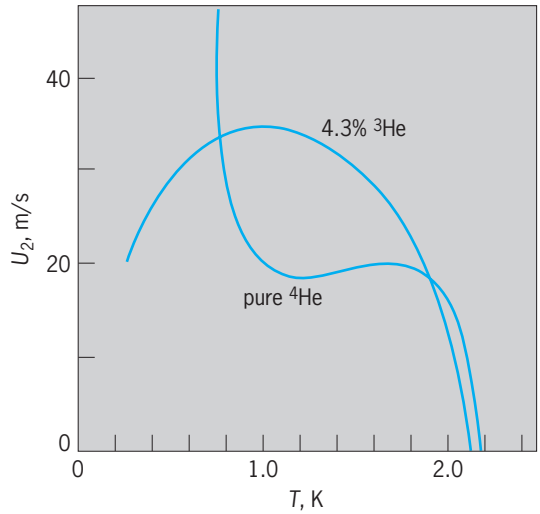

Velocity of second sound as a function of temperature for pure ^{4}He and a 4.3% ^{3}He mixture. (*After J. C. King and H. A. Fairbank, Second sound in ^{3}He-^{4}He mixtures below 1 K, Phys. Rev., 93:21, 1954*)

Second sound can be observed by methods similar to those used with first sound except that a heater and thermometer replace the sound transmitter and receiver. Thus an electrical resistor in the form of a thin metallic film or wire is the most common type of transducer for second sound.

Velocity. The velocity of second sound goes from 0 at T_λ to a value approaching that of first sound at 0 K as seen in the **illustration**. This is in excellent agreement with the theoretical relation in the equation shown below deduced from the two-fluid

$$u_2{}^2 = \frac{\rho_s}{\rho_n}\frac{T}{C}S^2$$

model, where T is the temperature, C the specific heat, and S the entropy of the liquid. At temperatures below about 0.5 K, Landau showed that the only excitations present in superfluid helium are longitudinal phonons (quantized sound waves), and hence heat pulses would be propagated as a fluctuation in the density of these phonons. As the temperature approaches 0, the equation shown for second-sound velocity reduces to $u_2 = u_1\sqrt{3}$, where u_1 is the phonon or first-sound velocity. Experimentally this value of velocity is approached at about 0.5 K, but another effect interferes—the interactions between phonons is very weak and the phonon mean free path goes rapidly toward infinity as the temperature and the density of phonons are reduced. Thus, below 0.5 K second sound is no longer observed in the true sense but the phonons propagate ballistically, interacting only with the walls of the container. This behavior is indeed found, as evidenced by large spreading of temperature pulses and the arrival of the leading edge with the phonon velocity.

In a variety of experiments second sound displays

the properties expected in wave phenomena such as diffraction, interference, and shock effects for large amplitude pulses.

^{3}He-^{4}He mixtures. Even a small addition of the lighter isotope helium-3 significantly affects many of the superfluid properties including second sound, especially in the very low temperature range. At the lowest temperatures where the ^{3}He excitations dominate the entropy and specific heat, the second-sound velocity is proportional to the square root of the temperature as expected for first sound in an ideal gas. Study of the second sound in this region provides an excellent means of measuring the details of the ^{3}He interactions.

Liquid ^{3}He. Liquid ^{3}He is a superfluid only at very low temperatures (below about 0.002 K). Second sound should exist in the principal phases (A and B) of this superfluid, but the expected velocity is quite small and the attenuation so large as to make direct detection unlikely. However, a new condensed form of second-sound wave and spin density wave can be propagated in a restricted narrow region of the ^{3}He phase diagram, the A_1 phase. This phase lies between the normal phase and the superfluid A phase and exists only in a strong magnetic field.

The superfluid component of the A_1 phase is apparently made up of pairs of ^{3}He atoms with their nuclear magnetic moments lined up parallel to the strong applied magnetic field. Counterflow of the normal and superfluid components, therefore, involves both the temperature and spin density changes in this normal second-sound wave/spin density wave combination. The velocity is directly proportional to the reduced temperature $(1 - T/T_{A1})$ and reaches a maximum value of about 1.4 m/s at the lower temperature limit of the A_1 phase. Here T_{A1} is the transition temperature between the A_1 phase and the normal phase. *See* LIQUID HELIUM; SUPERFLUIDITY.

Solid dielectric crystals. Theoretical predictions that second sound should exist in certain solid dielectric crystals under suitable conditions have been confirmed experimentally for solid helium single crystals at temperatures between 0.4 and 1.0 K. In a dielectric solid, heat pulses are carried only by phonons, as for liquid helium II, at very low temperatures. If resistive phonon scattering processes dominate, such as phonon-phonon umklapp processes, scattering by isotope or chemical impurities, or scattering by boundaries or imperfections in the crystal, the first-order diffusive flow equation would govern heat or temperature propagation. However, if the mean free path for such processes is large and the mean free path for nondissipative phonon-phonon normal processes is small, the second-order wave equation is obeyed, and heat pulses are propagated as second sound. That second sound has not yet been seen in most other dielectric crystals is probably due to the diffusive scattering of the phonons by impurities. However, evidence of second-sound propagation has also been found in pure sodium fluoride (NaF) crystals, and fully developed second sound has been found in pure semimetal bismuth crystals. Second sound has provided a good method of measuring directly the mean free path for phonon normal scattering processes. *See* LOW-TEMPERATURE ACOUSTICS.

Liquid crystals. Another quite different class of materials can exhibit second sound. In smectic A liquid crystals, when the wave vector is oblique with respect to the layers of these ordered structures, a modulation of the interlayer spacing can propagate at nearly constant density. It is likely that second sound will continue to be a valuable probe for studying the properties of interesting ordered materials.

Henry A. Fairbank

Bibliography. K. H. Bennemann and J. B. Ketterson, *The Physics of Liquid and Solid Helium*, 1976; J. Wilks and D. S. Betts, *An Introduction to Liquid Helium*, 2d ed., 1987; P. Wolf and P. Volhardt (eds.), *The Superfluid Phases of Helium-3*, 1990.

Secondary emission

The emission of electrons from the surface of a solid into vacuum caused by bombardment with charged particles, in particular, with electrons. The mechanism of secondary emission under ion bombardment is quite different from that under electron bombardment; the discussion here is limited to the latter case because it is in this sense that the term secondary emission is generally used.

The bombarding electrons and the emitted electrons are referred to, respectively, as primaries and secondaries. Secondary emission has important practical applications because the secondary yield, that is, the number of secondaries emitted per incident primary, may exceed unity. Thus, secondary emitters are used in electron multipliers, especially in photomultipliers, and in other electronic devices such as

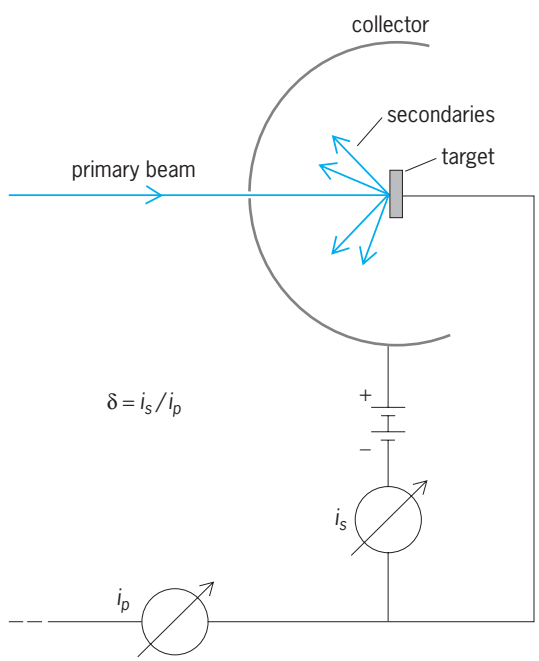

Fig. 1. Schematic circuit for measuring secondary yield; i_p and i_s represent the primary and secondary currents.

television pickup tubes, storage tubes for electronic computers, and so on.

Secondary yield. The most thoroughly investigated property of secondary emission is the yield as a function of the energy of the primaries. The yield may be measured by means of the circuit shown schematically in **Fig. 1**. A beam of primary electrons strikes a target with an energy determined by the potential difference between the target and the cathode. The primary beam passes through a hole in the collector, which has been made positive with respect to the target. The secondaries emitted by the target then flow to the collector, and the yield is obtained as the ratio of the secondary current i_s to the primary current i_p.

Mechanism of the process. The emission of secondary electrons can be described as the result of three processes: (1) excitation of electrons in the solid into high-energy states by the impact of high-energy primary electrons, (2) transport of these secondary electrons to the solid-vacuum interface, and (3) escape of the electrons over the surface barrier into the vacuum. The efficiency of each of these three processes, and hence the magnitude of the secondary emission yield δ, varies greatly for different materials.

Taking into account the material characteristics that, in addition to the value of the primary energy E_p determine the yield δ, one arrives at the equation

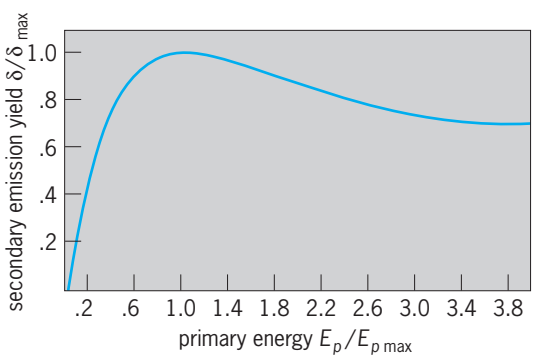

Fig. 2. Theoretical curve of secondary emission yield as a function of primary energy in normalized coordinates.

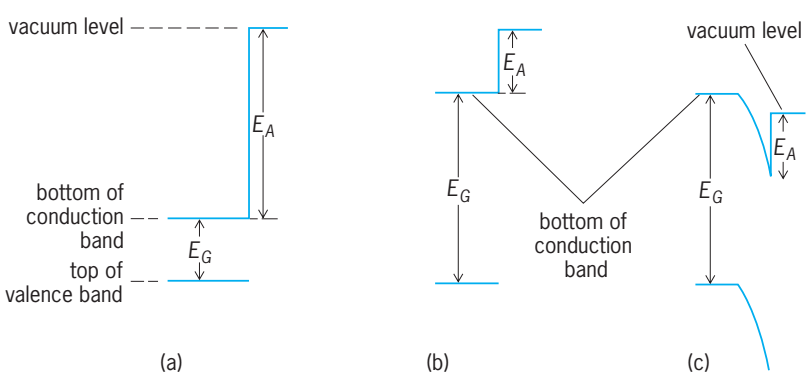

Fig. 3. Typical energy band models for semiconductors (or insulators). (*a*) $E_G \ll E_A$. (*b*) $E_G \gg E_A$. (*c*) Negative electron affinity.

below for the dependence of δ upon E_p, where ϵ

$$\delta = \frac{B_1 B_2 E_p}{\epsilon R}(1 - e^{-R/L})$$

is the energy required to produce a secondary electron; R is the range of primary electrons; B_1 is the coefficient, taking into account that only a fraction of the excited electrons diffuse toward the surface; L is the mean free path of the secondary electron; and B_2 is the probability that an electron reaching the solid-vacuum interface can escape over the surface barrier. Here ϵ and R are associated with process 1 above, B_1 and L with process 2, and B_2 with process 3. Without giving a detailed derivation of the equation above, it is qualitatively plausible that δ increases with increasing B_1, B_2, E_p, and L and decreases with increasing ϵ and R.

On the basis of the equation, a universal curve can be derived (**Fig. 2**) in which δ/δ_{max} is plotted versus $E_p/E_{p\max}$, where δ_{max} is the maximum yield and $E_{p\max}$ is the corresponding primary energy. Whereas curves for the absolute values of δ versus E_p vary over a wide range for different materials, experiments have generally confirmed the validity of the curve in Fig. 2. The peak in the curve can be interpreted as follows: With increasing E_p, the number of secondaries produced within the solid increases, but at the same time the primaries penetrate to a greater depth in the material. Because of energy loss processes, the "escape depth" of the secondaries has a finite value which is determined by some of the parameters entering the equation above. Thus the peak of the curve represents the point beyond which the number of secondaries produced at a depth greater than the escape depth exceeds the number of additional secondaries produced due to the higher primary energy. Because the escape depth is predominantly determined by L and B_2 in the equation, the variations in these two parameters are the main reasons why the δ values for different materials vary over such a wide range.

Experimental yield curves. In a discussion of measured δ and E_p curves, it is useful to consider metals and semiconductors (or insulators) separately. In metals the secondaries lose their energy rapidly by electron-electron scattering. As a result, L and B_2 in the equation are small, and the escape depth is of the order of, at most, nanometers. Hence, δ_{max}, and consequently $E_{p\max}$, have low values, typically well below 2.

For semiconductors and insulators, the situation is more complicated and is best understood in terms of energy-band models. Referring to **Fig. 3**, the highest δ values that can be obtained depend on the relative position of the top of the valence band (where the secondary electrons originate), the bottom of the conduction band, and the vacuum level. Three typical band models are shown in Fig. 3. The model shown in Fig. 3*a* is characterized by a small ratio of band-gap energy E_G to electron affinity E_G to electron affinity E_A. In the model shown in Fig. 3*b* the E_G to E_A ratio is large. In the model shown in Fig. 3*c* the bands are bent downward to such an extent that the

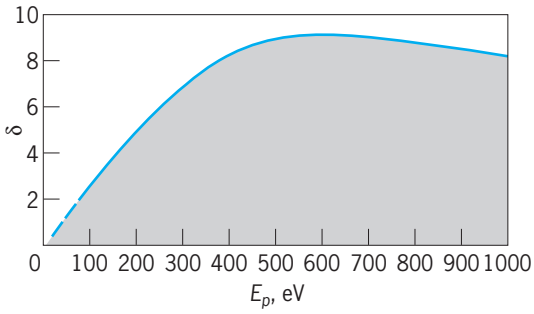

Fig. 4. Secondary emission yield versus primary energy for MgO.

vacuum level lies below the bottom of the conduction band in the bulk. A material with this characteristic is said to have negative effective electron affinity. This concept is also of great importance in photoelectric emission. The differences in secondary emission yields associated with each of the three band models can be qualitatively summarized as follows. *See* PHOTOEMISSION.

$E_G \ll E_A$ model. Secondary electrons excited from the valence band to levels above the vacuum level tend to lose their energy by exciting additional electrons from the valence band into the conduction band and thus arrive at the solid–vacuum interface with insufficient energy to overcome the surface barrier. In other words, the escape depth is very small, and the maximum δ values are below 2, similar to those of metals. Examples of this model are germanium and silicon.

$E_G \gg E_A$ model. Whereas secondary electrons excited from the valence band gradually lose energy by phonon-phonon scattering, an appreciable number of secondaries reach the solid–vacuum interface with sufficient energy to overcome the surface barrier. In other words, the escape depth is larger than in the case where $E_G \ll E_A$, of the order of tens of nanometers, and maximum δ values in the 8–15 range are typically obtained. Most of the materials used in practical devices fall into this category. Examples are MgO (see **Fig. 4**), BeO, Cs_3Sb (cesium antimonide), and KCl.

Negative effective electron affinity. Here the vacuum level is below the bottom of the conduction band. This case differs drastically from that shown in Fig. 3*b* because electrons that have dropped to the bottom of the conduction band as a result of phonon-phonon scattering still have enough energy to escape into the vacuum. Because the lifetime of electrons in the bottom of the conduction band is orders of magnitude longer than in states above this level, the escape depth of the secondaries is orders of magnitude greater than in the case represented in Fig. 3*b*. The most important material in this category is cesium-activated gallium phosphide, GaP(Cs). **Figure 5** shows the δ versus E_p curve for GaP(Cs) by comparison with MgO. Because of the much greater escape depth, δ values exceeding 100 are readily obtained. The curve for GaP(Cs) still follows quite closely the universal curve (Fig. 2), but the $E_{p\max}$ value is now in the 5–10-kV region compared with several hundred volts for materials represented in Fig. 3*a* or *b*.

Materials of the GaP(Cs) type represent a major breakthrough in the use of secondary emission for practical devices. In photomultipliers, for example, GaP(Cs) is superior to the conventional materials represented in Fig. 3*b* for a number of reasons, the greatest advantage being the improved signal-to-noise ratio. Negative effective electron affinity materials are not in universal use because of the more complex activation procedure and the associated higher cost. *See* BAND THEORY OF SOLIDS; SEMICONDUCTOR.

Dependence on angle of primary beam. The secondary yield for a given primary energy increases as the angle θ between the primary beam and the normal to the surface increases; the secondaries are then produced closer to the surface and consequently have a larger escape probability. At the same time,

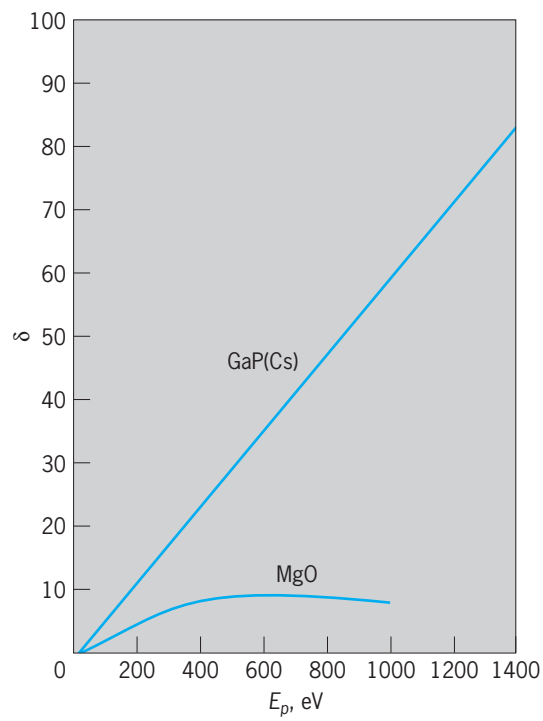

Fig. 5. Secondary emission yield versus primary energy for GaP(Cs). The curve for MgO is shown for comparison.

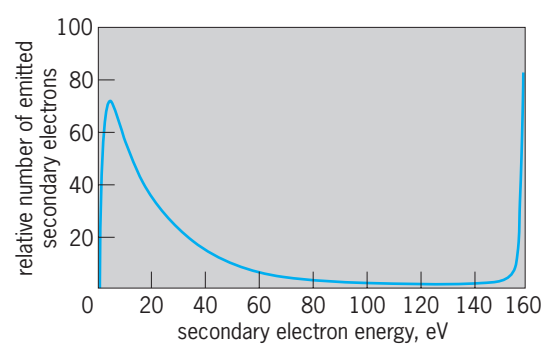

Fig. 6. Relative number of secondary electrons as a function of secondary energy for $E_p = 160$ V.

the energy for which the yield reaches its maximum value increases with increasing θ.

Secondary energies. A typical energy distribution of secondary electrons emitted by a silver target bombarded with primaries of 160-eV energy is given in **Fig. 6**. Note that most of the secondaries have relatively low energies. A small fraction of the emitted electrons have the same energy as the incident primaries and are called reflected primaries.

Alfred H. Sommer

Bibliography. R. L. Bell, *Negative Electron Affinity Devices*, 1973; A. J. Dekker, Secondary electron emission, in *Solid State Physics*, vol. 6, p. 251, 1958; A. Modinos, *Field, Thermionic, and Secondary Emission*, 1984.

Secondary ion mass spectrometry (SIMS)

An instrumental technique that measures the elemental and molecular composition of solid materials. Secondary ion mass spectrometry also provides methods of visualizing the two- and three-dimensional composition of solids at lateral resolutions approaching several hundred nanometers and depth resolutions of 1–10 nm. This technique employs an energetic ion beam to remove or sputter the atomic and molecular constituents from a surface in a very controlled manner. The sputtered products include atoms, molecules, and molecular fragments that are characteristic of the surface composition within each volume element sputtered by the ion beam. A small fraction of the sputtered atoms and molecules are ionized as either positive or negative ions, and a measurement by secondary ion mass spectrometry determines the mass and intensity of these secondary ions by using various mass analysis or mass spectrometry techniques. In this technique, the sputtering ions are referred to as the primary ions or the primary ion beam, while the ions produced in sputtering the solid are the secondary ions. Most elements in the periodic table produce secondary ions, and secondary ion mass spectrometry can quantitatively detect elemental concentrations in the part per million to part per billion range by using appropriate standards and specialized analytical conditions. Molecular detection limits by secondary ion mass spectrometry range between 0.1% and part-per-million concentration levels.

Dynamic and static SIMS. Secondary ion mass spectrometry analyses are divided into two broad categories known as dynamic and static. In the dynamic type, a relatively intense primary ion beam sputters the sample surface at sputter rates ranging 1–20 atomic layers/s (0.5–10 nm/s). This high sputter rate provides a very useful method for determining the in-depth concentration of different elements in a solid and is the most common method of secondary ion mass spectrometry analysis. Since the primary ion beam is eroding away the surface at a high rate, most molecular or chemical bonding information is rapidly destroyed, and hence the most common secondary ions detected in a dynamic secondary ion mass spectrometry analysis are elemental ions or clusters of elemental ions.

Static or molecular secondary ion mass spectrometry utilizes a very low intensity primary ion beam, and static analyses are typically completed before a single monolayer has been removed from the surface. Most static analyses are stopped before 1% of the top surface layer has been chemically damaged or eroded; under these conditions, molecular and molecular fragment ions characteristic of the chemical structure of the surface are often detected. Thus, static secondary ion mass spectrometry is best suited for near-surface analysis of molecular composition or chemical structure information, while dynamic secondary ion mass spectrometry provides the best technique for bulk and in-depth elemental analysis. *See* SPUTTERING.

Ion production. Secondary ions are formed by kinetic and chemical ionization processes in which sputtering is achieved by energy transfer from the primary ions to the solid surface. Typically, SIMS primary ions impact the surface with kinetic energies of 5–20 keV, and this energy is ultimately transferred to the sample atoms and molecules. The energy transfer initiates a collision cascade within the solid that ejects atoms and molecules at the solid surface-vacuum interface. A fraction of the ejected or sputtered species are in metastable or ionized states; however, most secondary ions formed within or near the surface are neutralized by unbound electrons in the solid before the ions can escape from the surface. However, metastable neutral atoms, that is, those neutral atoms having an internal energy above the ionization potential of the atom or molecule, can eject an electron above the sample surface and form a positive ion. Similarly, neutral atoms or molecules can abstract an electron from the solid, forming a negative secondary ion if the energy gained in the electron attachment exceeds the energy required to remove the electron from the surface. Kinetic sputtering processes are those in which the primary ion does not alter the electronic state of the solid. Kinetic sputtering is most commonly observed using primary ions of noble gases such as argon (Ar^+), krypton (Kr^+), or xenon (Xe^+). Chemical ionization processes occur in primary ion sputtering when the primary species alters the electronic state of the solid.

Two types of surface alterations are most commonly used in secondary ion mass spectrometry, and these modifications either reduce or increase the number of unbound electrons in the solid. For example, bombarding the surface with various oxygen ions (O_2^+, O^+, or O^-) reduces the number of conduction band electrons and favors the formation of positive ions. Similarly, bombarding a surface with cesium ions (Cs^+) increases the number of conduction band electrons and favors the formation of negative secondary ions. The use of either oxygen or cesium ion beams in dynamic secondary ion mass spectrometry is referred to a reactive-ion sputtering, because secondary ion yields are typically enhanced by one or more orders of magnitude over yields achieved with

unreactive primary species such as Ar^+. *See* CONDUCTION BAND; INERT GASES; ION.

Instrumentation. Instrument designs for secondary ion mass spectrometry range from custom-built, laboratory instruments to general-purpose, commercially available units. All instruments require a primary ion gun or column to generate and transport the primary ion beam to the sample surface, a sample chamber with sample mounting facilities, a mass spectrometer which performs mass-to-charge separation of the different secondary ions, and an ion-detection system. The complete instrument is typically housed in an ultrahigh vacuum chamber. *See* ION SOURCES; MASS SPECTROSCOPE.

Secondary ion mass spectrometry instrumentation is often distinguished by the method used to form secondary ion images. The ion microscope (see **illus.**) is used in a direct imaging technique that utilizes a relatively large primary ion beam ranging between 10 and 50 micrometers in diameter to sputter the sample surface. The secondary ions produced by this beam are extracted from the sample surface, mass analyzed, and detected. Secondary ion images mapping the lateral distribution of selected ions formed at the surface are acquired by transporting a two-dimensional image of the ions formed at the sample surface through the mass spectrometer and projecting a magnified image of these ions onto a two-dimensional ion detector. Typical image resolutions range from 1 to 5 μm from analysis of spots 100 μm in diameter. The lateral resolution of ion images formed by the ion microscope is determined by the quality of the secondary ion optics and the topography of the sample surface. Thus, smooth or flat samples typically produce higher-resolution images than rough surfaces.

The other mode of SIMS imaging is achieved by using an ion microprobe technique in which a microfocused primary ion beam is rastered about the sample surface, producing secondary ions at each raster position. Mass analysis of the secondary ions at each point in the raster, coupled with synchronization of the detection circuit with primary beam position, provides a method of generating secondary ion images at high spatial resolutions. The lateral resolutions of ion microprobe images are limited only by the spot size of the primary ion beam, and innovative source designs using liquid-metal ion guns routinely produce 50-nm spot sizes. Ion microprobe imaging thus provides the highest spatial resolution in secondary ion mass spectrometry analyses; however, ion microscope imaging is a faster imaging technique, since many pixels are irradiated per unit time.

A variety of mass spectrometers are used in secondary ion mass spectrometry instrumentation, and the most common include single- and double-focusing magnetic sector designs, quadrupole mass filters, and time-of-flight mass spectrometers. Magnetic sector and quadrupole mass filters are used most commonly in dynamic secondary ion mass spectrometry instruments, while time-of-flight spectrometers provide optimal performance for static secondary ion mass spectrometry applications. Both

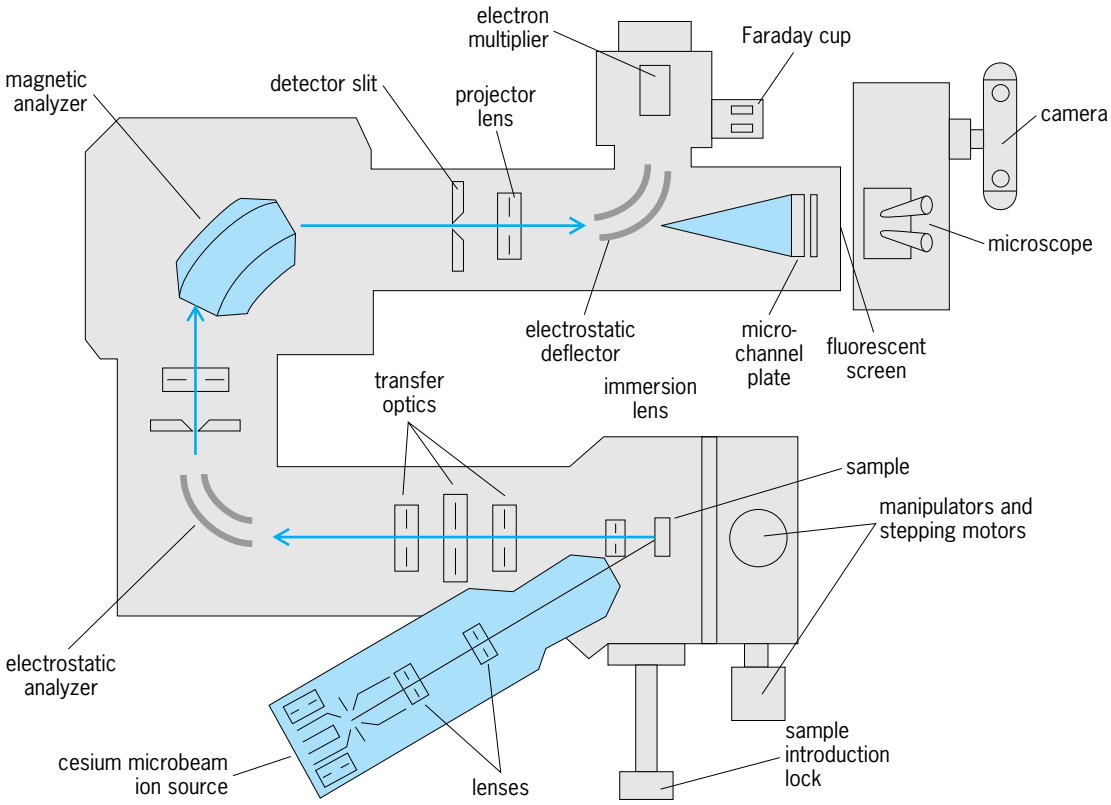

Diagram of an ion microscope.

magnetic sector and time-of-flight spectrometers can operate as either ion microscopes or ion microprobes, while instruments using quadrupole mass filters image almost exclusively with ion microprobe techniques. *See* TIME-OF-FLIGHT SPECTROMETERS.

Analytical applications. Both dynamic and static secondary ion mass spectrometry are used for analysis.

Dynamic. Dynamic secondary ion mass spectrometry has been successfully utilized in diverse applications, including the characterization of semiconductor materials, metals and metal alloys, catalysts, and geological and biological samples. The analytical issues are the detection and localization of specific elements on the surface or in the bulk of the materials. The most common application area of dynamic SIMS is in semiconductor analysis. Another important contribution of secondary ion mass spectrometry to materials science is the identification and localization of trace elements in metal grain boundaries, providing detailed insight into the chemistry of welds and alloys. Dynamic secondary ion mass spectrometry also has the unique ability to detect specific catalyst poisons on the surface and in the bulk of used or spent catalysts. In geological sciences, SIMS has been used to detect isotopic anomalies in the composition of various geological and meteoritic samples. Sensitive isotope ratio data have played a very important role in a more complete understanding of the age of the universe. Finally, dynamic secondary ion mass spectrometry has found extensive applications to the characterization of both hard and soft biological tissue.

The most important application of dynamic SIMS is the bulk and in-depth analysis of semiconductor materials. Specific analyses include detection of surface contaminants such as alkalis, alkaline-earth metals, and transition metals [titanium (Ti), chromium (Cr), iron (Fe), nickel (Ni), and zinc (Zn)] on silicon (Si) and gallium arsenide (GaAs) surfaces, high-sensitivity (bulk) analysis of boron (B) and oxygen in silicon, and in-depth analysis of semiconductor dopants such as B, phosphorus (P), and As. Essentially all of these studies employ reactive primary ion sputtering to achieve high detection sensitivity. Bulk analysis for specific impurities such as B or O in silicon are performed at high primary currents in order to maximize sputter rate and, hence, detection sensitivity. Bulk analyses generally do not provide useful image information since the erosion or sputter rates are high ($\sim$10 nm/s).

The most common dynamic SIMS application is depth profiling. In this technique the intensities of one or more secondary ions are determined as a function of sputter depth into the solid. Knowledge of the relative sensitivity factor of the implanted species and the total analysis depth permits converting the data from raw signal intensity versus sputter time into an elemental concentration versus depth. Depth profiling is used to determine the distribution and peak depths of ion implants into semiconductor solids. Ion implantation is a very common process for doping semiconductor materials with low levels of electrically active species such as B, P, and As. Doping via ion implantation is achieved by forming an ion of the dopant species, accelerating this ion to high energies (50–250 keV), and implanting the dopant ion into the solid at known doses (number of dopant atoms/cm^2). The most critical parameters in ion implantation are the peak depth, peak concentration, and shape of the implant profile. Since modern semiconductor devices require low implant peak concentrations ($<10^{19}$ dopant atoms/cm^3 or 0.1% concentration), dynamic secondary ion mass spectrometry is an ideal method for measuring and quantifying these ion implants.

Accurate depth profiling does not detect ions produced at the edges or walls of the sputter crater, since these signals could be generated over a range of depths. Crater wall discrimination is achieved either by mechanical or electronic aperturing of the secondary ion signal.

A typical example of such an analysis is determination of the C, Si, Al, and O distributions in a GaAs/Al-GaAs laser diode structure. A depth profile for this type of analysis can be achieved by using a Cs^+ primary ion beam and negative secondary ion detection. The raw data (secondary ion intensity versus sputter time) are converted to atomic concentration (atoms/cm^3) versus sputter depth (μm) by using ion implant standards and accurate crater-depth measurement. *See* SEMICONDUCTOR.

Static. Applications of static secondary ion mass spectrometry to analyses of the near-surface region of solids has become increasingly important as ever more sophisticated materials are developed. The near-surface region is generally defined as the top five monolayers (2–3 nm) of a solid. Examples of technology areas in which the chemistry of the near-surface region plays a critical role include high-performance glass coatings, liquid-crystal displays, manufacturing, semiconductor processing, biopolymer and biocompatible materials development, and polymer adhesives and coatings. The development of high-performance time-of-flight mass spectrometry techniques coupled with ion beam sputtering revolutionized the applications of static secondary ion mass spectrometry in many of these areas. The techniques combine the sensitivity of static secondary ion mass spectrometry with a high-transmission, high-mass-resolution instrument. In addition, time-of-flight secondary ion mass spectrometry performs high lateral resolution imaging by using either ion microscope or ion microprobe techniques. *See* MASS SPECTROMETRY; MATERIALS SCIENCE AND ENGINEERING; SEMICONDUCTOR DIODE.

Robert W. Odom

Bibliography. A. Benninghoven, F. G. Rudenaur, and H. W. Werner, *Secondary Ion Mass Spectrometry: Basic Concepts, Instrumental Aspects, Applications and Trends*, 1987; E. de Hoffmann and V. Stroobant, *Mass Spectrometry: Principles and Applications*, 2001; J. C. Vickerman, A. Brown, and N. M. Reed, *Secondary Ion Mass Spectrometry: Principles and Applications*, 1989.

Secretion

The export of proteins by cells. With few exceptions, in eukaryotic cells proteins are exported via the secretory pathway, which includes the endoplasmic reticulum and the Golgi apparatus. Secreted proteins are important in many physiological processes, from the transport of lipids and nutrients in the blood, to the digestion of food in the intestine, to the regulation of metabolic processes by hormones. *See* CELL (BIOLOGY); CELL ORGANIZATION; ENDOCRINE MECHANISMS.

Synthesis on endoplasmic reticulum. Proteins destined for export or for insertion into the plasma membrane are synthesized initially on ribosomes attached to the outside of the rough endoplasmic reticulum, a portion of the membrane-bounded endoplasmic reticulum that is specialized for the synthesis of secretory proteins and most of the cell's membrane proteins. These proteins have a unique address tag called a signal sequence, usually located at or near the amino terminus, that ensures their recognition by a complex of proteins in the cytoplasm, known as the signal recognition particle, during their synthesis on the ribosomes. The signal recognition particle mediates the delivery of the growing polypeptides to the translocation apparatus. The translocation apparatus in turn mediates the transfer of the proteins across the membrane of the endoplasmic reticulum and into its interior space (lumen) or, in the case of membrane proteins, into the endoplasmic reticulum membrane itself. Proteins in the lumen of the endoplasmic reticulum or its membrane assume their appropriate three-dimensional structure, a process known as folding that is accelerated by chaperone proteins in the endoplasmic reticulum. *See* ENDOPLASMIC RETICULUM; PROTEIN; RIBOSOMES.

Export to Golgi apparatus. After they are folded, proteins to be exported enter small vesicles for transport to the Golgi apparatus (see **illus.**). The vesicles are coated with coat protein complex II (COPII) on their cytoplasmic surface, and they emerge from unique sites of the endoplasmic reticulum. The assem bly of the coat is thought to deform the membrane and facilitate the process of vesicle budding. Although the proteins making up the coat have been identified, it is not understood how secretory and membrane protein cargoes are concentrated within the COPII transport vesicles. Some glycoproteins, including the coagulation factors V and VIII, bind to a specific receptor protein, ERGIC-53, which acts to promote their incorporation into transport vesicles. Genetic mutations in ERGIC-53 lead to hemophilia (poor blood coagulation following injury) due to deficient secretion of factors V and VIII. *See* GLYCOPROTEIN; GOLGI APPARATUS; HEMOPHILIA; MUTATION.

The COPII vesicles are transported toward the Golgi apparatus in a process involving movement along microtubular tracks. As they approach the Golgi apparatus, they fuse together to form the *cis*-Golgi network, a tubulated structure on the *cis* side of the Golgi stack (the side closest to the endoplasmic reticulum). In the *cis*-Golgi network, proteins destined for further transport are segregated from

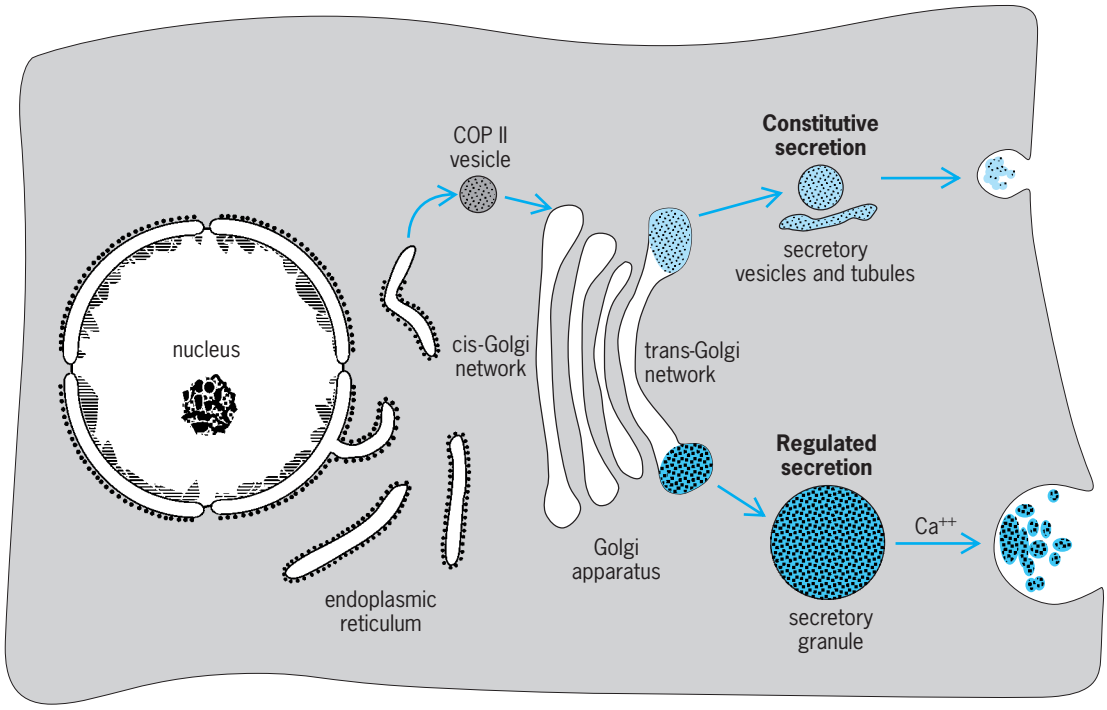

Secretory pathway of eukaryotic cells. Proteins are synthesized on ribosomes bound to the endoplasmic reticulum and subsequently transferred to the *cis*-Golgi network in COPII-coated vesicles. After traversing the Golgi apparatus, proteins are packaged into transport vesicles and tubules forming in the *trans*-Golgi network. In constitutive secretion, these vesicles and tubules subsequently fuse with the plasma membrane of the cell. In regulated secretion, newly synthesized proteins exit the *trans*-Golgi network and are routed to storage organelles called secretory granules. Secretory granule exocytosis occurs when an external signal is received by the cell and cellular calcium (Ca^{2+}) levels increase.

those whose predominant site of action is in the endoplasmic reticulum, such as ERGIC-53. ERGIC-53 and other endoplasmic reticulum proteins are recycled back to the endoplasmic reticulum by vesicular transport. The cargo is then transported across the cisternae of the Golgi stack, from the *cis* to the *trans* side, in a process that is dependent upon the formation of COPI-coated vesicles.

Sorting and transport to plasma membrane. When proteins reach the last cisterna of the Golgi, a highly tubulated region known as the *trans*-Golgi network, they are sorted and packaged again into transport vesicles, some of which are in the form of elongated tubules. From here, there are two pathways that proteins can take to the cell surface, depending on the cell type. Proteins can be transported directly to the plasma membrane (constitutive secretion) or to secretory granules (regulated secretion).

Constitutive secretion. In all cells, there exists a constitutive secretion pathway whereby vesicles and tubules emerging from the *trans*-Golgi network fuse rapidly with the plasma membrane (see illus.). The emerging vesicles and tubules can attach to microtubules, cytoskeletal elements emanating from the Golgi region, that accelerate their transport to the plasma membrane. In polarized epithelial cells, there can be two types of constitutive transport carriers that form in the *trans*-Golgi network.

Polarized epithelial cells have two plasma membrane domains, the apical and basolateral, that are separated by tight junctional complexes. The apical membrane domain faces the external milieu, lining the lumen of glands and the inside of the intestinal tract, for example. The basolateral surface is oriented toward the tissue interior and is in proximity to blood vessels. The transport carriers (vesicles and tubules), which contain distinct subsets of cargo molecules, will fuse with either the apical or basolateral plasma membrane domain. Signals in the cargo molecules themselves determine which type of transport vesicles, apical or basolateral, they enter. However, the receptors in the *trans*-Golgi network that sort the proteins into different vesicle populations have not been identified, nor have coat protein complexes been identified in the *trans*-Golgi network that might function as COPI and COPII do in the early secretory pathway. In some epithelial cell types, such as liver hepatocytes, nearly all of the protein traffic from the *trans*-Golgi network is routed to vesicles that fuse with the basolateral (sinusoidal) membrane. A subset of the membrane-bound proteins is subsequently redistributed to the apical (canilicular) plasma membrane in a process resembling endocytosis, known as transcytosis. *See* ABSORPTION (BIOLOGY); CELL MEMBRANES; ENDOCYTOSIS.

Regulated secretion. In cells that secrete large amounts of hormones, digestive enzymes, or enzymes that digest immune complexes and pathogens, most secretory and membrane proteins emerging from the *trans*-Golgi network are not immediately secreted, but are stored in membrane-bounded secretory granules (see illus.). Secretory granules forming in the *trans*-Golgi network contain large aggregates of content proteins. The aggregates are segregated away from proteins transported by constitutive transport vesicles. During the formation of secretory granules, hormone precursor proteins may be cleaved by enzymes to generate mature hormone molecules. Insulin, for example, is generated from its precursor, proinsulin, during the formation of secretory granules from the *trans*-Golgi network. Secretory granules release their contents into the extracellular space in a process known as exocytosis, when their membranes fuse with the plasma membrane. Exocytosis occurs only after the cell receives a signal, usually initiated by the binding of a hormone or neurotransmitter to a receptor on the cell surface. The receptor triggers a signal transduction cascade that results in increased concentrations of second messengers such as cyclic adenosine 3′,5′-monophosphate and phosphatidylinositol triphosphate. In most secretory cells, the second messengers or the hormone receptors themselves trigger the opening of calcium channels. Calcium ions stream into the cytoplasm through these channels, entering either from the extracellular milieu or from intracellular storage sites in the endoplasmic reticulum and secretory granules. Calcium initiates the docking of the secretory granules with the plasma membrane and the activation of the fusion apparatus (see illus.). *See* ENZYME; HORMONE; SIGNAL TRANSDUCTION.

Proteins destined for lysosomes also follow the secretory pathway (although they are not ultimately secreted). However, after reaching the *trans*-Golgi network, most are diverted into clathrin-coated vesicles. The clathrin-coated vesicles forming in the *trans*-Golgi network are similar to those formed during endocytosis at the cell surface, but contain a different set of the accessory proteins known as adaptors. Clathrin-coated vesicles fuse with endosomes after budding from the *trans*-Golgi network, and their cargo is subsequently delivered to lysosomes.

Fusion of transport vesicles with target membranes. Each step of transport along the secretory pathway requires that vesicles forming from one membrane-bounded organelle fuse with their correct target membranes. A specialized set of fusion proteins exists in the membrane of every vesicle population (v-SNAREs) and every target membrane (t-SNAREs). v-SNAREs form complexes with t-SNAREs, and these complexes mediate the fusion of the membranes of the vesicle and the target. Each type of vesicle has its own unique v-SNARE, and different target membranes contain different sets of t-SNAREs. The interaction between sets of v-SNAREs and t-SNAREs, a highly selective process, contributes to the specificity of vesicle fusion along the secretory pathway. The ability of the v-SNAREs and t-SNAREs to participate in membrane fusion events depends, in turn, on the activity of accessory proteins. One family of these proteins, which hydrolyze guanosine 5′-triphosphate, is known as the *rab* protein family. Different members of the *rab* family are responsible for activating SNARE interactions at each step during the transport of proteins from the endoplasmic reticulum to the plasma membrane. Thus, *rabs*

and other organelle-specific accessory proteins supplement the SNAREs in regulating the specificity of transport-vesicle targeting and fusion along the secretory pathway.

Unconventional secretion. In exceptional cases, proteins can be exported directly from the cytoplasm without using the secretory pathway. One such protein is fibroblast growth factor, a hormone involved in the growth and development of tissues such as bone and endothelium. Several interleukins, proteins that regulate the immune response, are also released via an unconventional route that may involve transport across the plasma membrane through channel proteins. These channels have adenosine 5′-triphosphatase (ATPase) enzyme activity and use the energy derived from the hydrolysis of ATP to catalyze transport. *See* ADENOSINE TRIPHOSPHATE (ATP); CELLULAR IMMUNOLOGY.

Secretion in prokaryotes. Bacteria also have the ability to export proteins, including toxins that cause illness in humans. Proteins exported via this route have signal sequences that are recognized by protein complexes in a manner analogous to recognition by signal recognition particle in eukaryotes. Proteins are directed during translation (or in some cases, post-translationally) for insertion into pores in the plasma membrane in a process resembling that occurring in the endoplasmic reticulum of eukaryotic cells. In gram-negative bacteria, such as *Escherichia coli*, the exported proteins are found primarily in the periplasmic space between the inner and outer membranes. A similar pathway is used in bacteria by membrane-bound proteins and proteins that form cell walls. *See* BACTERIA.

Michael Rindler

Bibliography. Y. A. Chen and R. H. Scheller, SNARE-mediated membrane fusion, *Nat. Rev. Mol. Cell Biol.*, 2:98–106, 2001; J. Klumperman, Transport between ER and Golgi, *Curr. Opin. Cell Biol.*, 12:445–449, 2000; H. Lodish et al., *Molecular Cell Biology*, W. H. Freeman, 2000; G. Palade, Intracellular aspects of the process of protein synthesis, *Science*, 189:347–358, 1975; M. J. Rindler, Biogenesis of storage granules and vesicles, *Curr. Opin. Cell Biol.*, 4:616–622, 1992.

Secretory structures (plant)

Cells or organizations of cells which produce a variety of secretions. The process of secretion is a separation of a substance from the protoplast of a cell. The secreted substance may remain deposited within the secretory cell itself or may be released from the cell. Substances may be secreted to the surface of the plant or into intercellular cavities or canals. Some of the many substances contained in the secretions are not further utilized by the plant (resins, rubber, tannins, and various crystals), while others take part in the functions of the plant (enzymes and hormones). Secretory structures range from single cells scattered among other kinds of cells to complex structures involving many cells; the latter are often called glands.

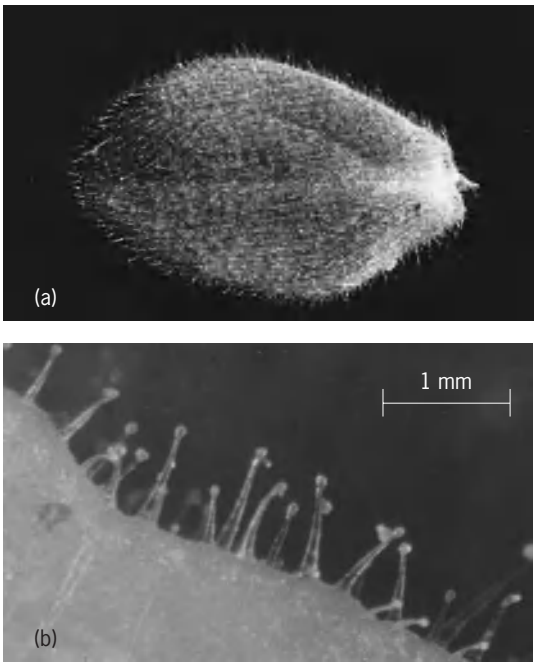

Fig. 1. ***Petunia***. **(*a*) Young leaf covered with glandular hairs. (*b*) Higher magnification of the hairs.**

Glandular hairs. Epidermal hairs of many plants are secretory or glandular. Such hairs commonly have a head composed of one or more secretory cells borne on a stalk (**Fig. 1***a* and *b*). The shaggy hairs of winter buds of many trees produce sticky secretions that permeate and cover the buds with a protective film. The hair of a stinging nettle is bulbous below and extends into a long, fine process above (**Fig. 2**). If one touches the hair, its tip breaks off, the sharp edge penetrates the skin, and the poisonous secretion is released.

Nectaries. Glands secreting a sugary liquid—the nectar—in flowers pollinated by insects are called nectaries (**Fig. 3**). Nectaries may occur on the floral stalk or on any floral organ: sepal, petal, stamen, or ovary. The nectary may be flat, depressed, or padlike. Its surface layer, one or more cells deep, consists of

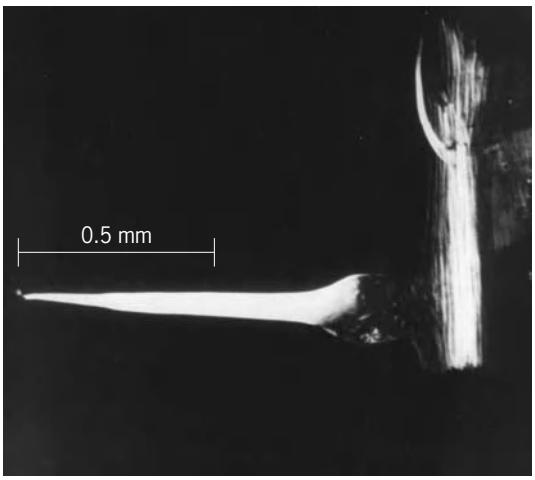

Fig. 2. Hair of stinging nettle.

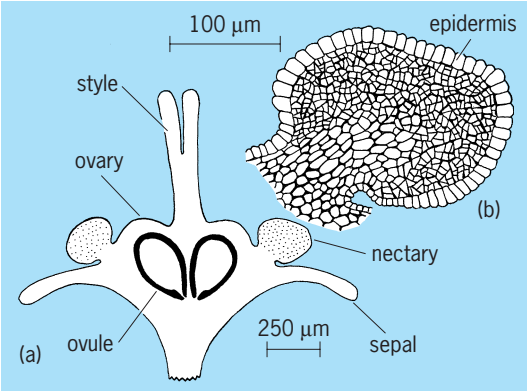

Fig. 3. ***Ceanothus*. (*a*) Longitudinal section of flower. The nectary is ringlike and surrounds the ovary. (*b*) Enlarged view of a section of the nectary. The epidermis and the small cells beneath it are secretory.**

secretory cells which usually have dense cytoplasm. Under the electron microscope these cells show few small vacuoles, a dense ground substance, numerous mitochondria, and a strongly developed endoplasmic reticulum, but the dictyosomes are not conspicuous. In the nectaries of monocotyledonous flowers, which occur in the partitions in the ovary (septal nectaries), the outer walls of the secretory cells develop numerous protuberances toward the interior of the cell. Since the plasmalemma lines these protuberances, its surface area is greatly increased. Thus a large secretory surface is provided. The vascular tissue, especially the sugar-conducting phloem, occurs close to the secretory tissue. The nectar may contain as much as 60% sugar.

Hydathodes. The hydathode structures discharge water—a phenomenon called guttation—through openings in margins or tips of leaves. The water flows through the xylem to its endings in the leaf and then through the intercellular spaces of the hydathode tissue toward the openings in the epidermis. Strictly speaking, such hydathodes are not glands because they are passive with regard to the flow of water. The pressure forcing the water to be discharged originates in the root. Some hydathodes, however, are like nectaries in having a layer of actively secreting cells.

Digestive glands. Some carnivorous plants have glands that produce secretions capable of digesting insects and small animals. These glands occur on leaf parts modified as insect-trapping structures. In the sundews (*Drosera*) the traps bear stalked glands, called tentacles. Each gland consists of four layers of cells, the innermost being in contact with the conducting tissue of the stalk. When an insect lights on the leaf, the tentacles bend down and cover the victim with a mucilaginous secretion, the enzymes of which digest the insect. In the Venus' flytrap (*Dionaea*) the insect is trapped by the sudden folding of the leaf. The upper surface of the leaf bears digestive glands and hairs acting as triggers for the folding mechanism. The traps in the pitcher plants are pitcher-shaped leaves that bear nectaries on the surface and inside the trap near its top; the digestive glands are deeper inside. Both consist of a layer of secretory columnar cells, a layer of rounded cells, and a layer of suberized cells. If an insect, attracted by the nectar, falls into the pitcher, it is prevented from escaping by rigid hairs pointing downward and is digested within the trap. The digested insects provide nitrogen that is essential for the plant's growth. Under the electron microscope the mucilage-producing cells are seen to contain large numbers of dictyosomes. The mucilage, a polysaccharide stainable with ruthenium red, occurs in the Golgi vesicles, which are derived from the dictyosomes. The vesicles separate from the dictyosomes and migrate toward the cell wall, frequently the anticlinal wall. The vesicle membrane and the plasmalemma open and join, and the contents of the vesicle are released into the space between the plasmalemma and the wall. The material then migrates along the wall toward the outside wall, where it is secreted by passing through the wall and cuticle. Dictyosomes do not appear to be involved in the production of digestive (proteolytic) enzymes. As judged by the scutellar cells of grass embryos, cells forming such enzymes show a strong development of the endoplasmic reticulum.

Resin ducts. Resin ducts are canals lined with secretory cells that release resins into the canal (**Fig. 4**). The canals are intercellular spaces that originate by separation of cells. Resin ducts are common in gymnosperms and occur in various tissues of roots, stems, leaves, and reproductive structures.

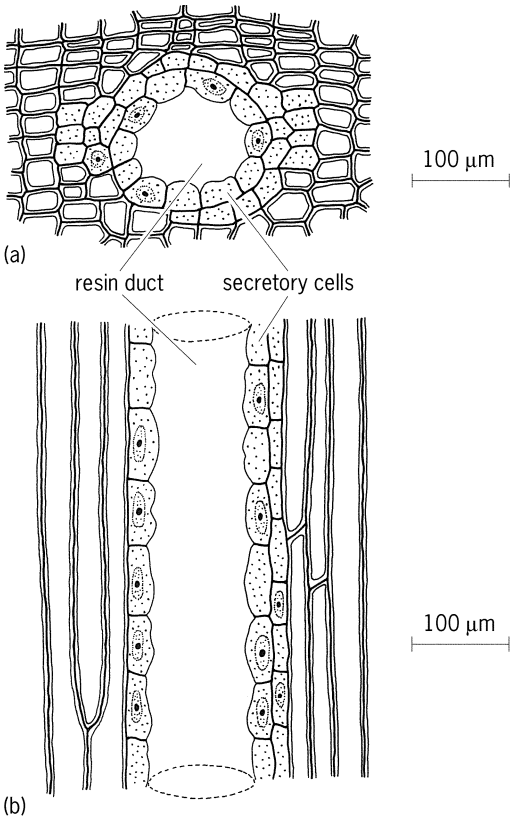

Fig. 4. Resin duct from pinewood in (*a*) transverse section and (*b*) longitudinal section.

They may arise during normal development or as a result of injury.

Gum ducts. These ducts are similar to resin ducts and may contain resins, oils, and gums. Usually, the term gum duct is used with reference to the dicotyledons, although gum ducts also may occur in the gymnosperms.

Oil ducts. Oil ducts are intercellular canals whose secretory cells produce oils or similar substances. Such ducts may be seen, for example, in various parts of the plant of the carrot family (Umbelliferae). In contrast to the oil ducts, the oil cavities of the kind found in the fruit rind and other plant parts of citrus result from a breakdown of oil-containing cells.

Laticifers. These are cells or systems of cells containing latex, a milky or clear, colored or colorless liquid. Latex occurs under pressure and exudes from the plant when the latter is cut. The single-cell laticifer is often much branched (**Fig. 5**); a few laticifer cells originating in the embryo may branch and invade all newly produced plant parts, forming the entire laticifer system of the mature plant. Laticifers derived from more than one cell result from breakdown of walls between adjacent cells; they are sometimes called laticiferous tubes. Both kinds of laticifers have multinucleate protoplasts. The latex is the product of this protoplast, but its relation to the cytoplasm seems to be more complex than that between the vacuole and cytoplasm in ordinary cells. The latex of some plants has a high content of rubber and is used as the source of natural commercial rubber. Examples of rubber-yielding plants are the Brazilian (*Hevea*) and Indian (*Ficus*) rubber trees. *See* INSECTIVOROUS PLANTS; PLANT METABOLISM; PLANT TISSUE SYSTEMS.

Katherine Esau

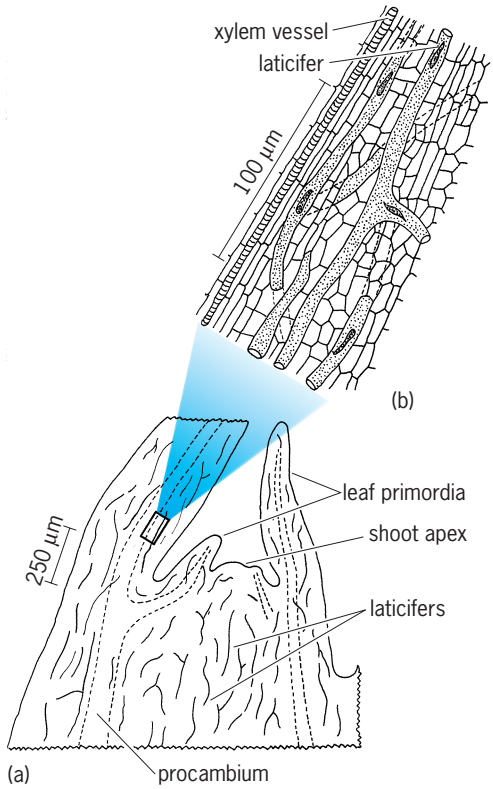

Fig. 5. Laticifers of *Nerium oleander*. (*a*) Longitudinal section of shoot tip. The wavy solid lines represent parts of laticiferous cells, which are arranged irregularly. (*b*) Enlarged view of the area in the rectangle in *a* showing parts of laticifers with cytoplasm and nuclei.

Sedative

A medication capable of producing a mild state of inhibition of the central nervous system (CNS) associated with reduced awareness of external stimuli. Numerous pharmacologic agents can induce different degrees of sedation, depending on the following variables: dosage; route of administration; absorption, metabolism, and excretion rates of the compound; specific receptor sites in the central nervous system that are affected by the agent; environmental setting; and state of the patient. *See* CENTRAL NERVOUS SYSTEM.

Types. Ethanol was probably the first sedative compound and was widely used for its analgesic and hypnotic properties, as well as for its ability to decrease inhibitory anxiety with resultant relaxation and occasional euphoria. Subsequent sedative compounds such as the barbiturates (for example, phenobarbital and secobarbital) and the benzodiazepines (for example, diazepam and alprazolam) also affect the same γ-aminobutyric acid (GABA) complex inhibitory receptor area in the central nervous system as ethanol does; however, they occupy different sites of this complex. Although these sedative compounds possess properties of tolerance and habituation, they vary in their addictive potential according to specific receptor sites and the particular type of patient. For example, in one controlled study, alprazolam was shown to be more likely to produce euphoric responses in alcoholics and their male offspring than a placebo or buspirone (an anxiety-reducing drug that is not a benzodiazepine). In other clinical studies, it has been noted that drug abusers and alcoholics appear to be more likely than other populations to abuse benzodiazepines. *See* ADDICTIVE DISORDERS; BARBITURATES.

In addition to those compounds that affect the GABA complex, other classes of chemical agents are used as sedatives. The antihistamines as well as some antidepressant drugs possess sedative side effects in addition to their primary pharmacologic properties. Since these classes of compounds are not addicting, they can be safely used as hypnotics. *See* TRANQUILIZER.

Effects. The sedative drugs that act on the GABA receptor sites, including GABA, are effective anticonvulsants and show cross-tolerance. Thus, they can be interchanged to treat their respective withdrawal syndromes. The benzodiazepines are useful in treating patients with the psychiatric diagnosis of generalized anxiety disorder. However, sedative side effects and possible interference with the ability to sustain attention and manual dexterity are possible. The duration of the side effects and withdrawal symptoms is related to the half-life of each of these sedative

compounds. Those sedative agents with short half-lives that are used to induce sleep can produce rebound insomnia when discontinued, and the long-half-life hypnotics are more likely to cause cognitive deficits on the day following administration. Cognitive impairment and drug dependence with subsequent withdrawal symptoms are more likely to occur if the dosage of the benzodiazepine is high, if use is of more than 4 months' duration, and if the patient is elderly or has had a prior history of alcohol or drug dependence. High-potency, short-half-life compounds appear to be more capable of producing pharmacologic dependence. *See* ANXIETY DISORDERS.

Donald M. Gallant

Bibliography. American Psychiatric Association Task Force Report on Benzodiazepine Dependency, Benzodiazepine dependence, toxicity, and abuse, *Amer. J. Psychiat.*, 1990; D. A. Ciraulo et al., Parental alcoholism as a risk factor in benzodiazepine abuse, *Amer. J. Psychiat.*, 146:1333–1335, 1989; H. Y. Meltzer (ed.), *Psychopharmacology: The Third Generation of Progress*, 1987.

Sedentaria

A group of 28 or more families of polychaete annelids in which the anterior, or cephalic, region is more or less completely concealed by overhanging peristomial structures, or the body is divided into an anterior thoracic and a posterior abdominal region; the pharynx or proboscis is usually soft and epithelial, lacking hard jaws or paragnaths. *See* POLYCHAETA.

Orbiniidae. This family has 13 genera and about 80 species. They are transitional between Errantia and Sedentaria in that the prostomium is exposed (**Fig. 1**); the thorax and abdomen are weakly separable; and the pharynx is muscular in the more primitive members and epithelial in the more developed members. There are two subfamilies: the Protoaricinae with 6 genera, chiefly from the Southern Hemisphere; and the Orbiniinae with 7 genera, including large, well-known species from worldwide areas in shallow to great ocean depths. The best known are the species in *Naineris*, *Orbinia*, *Scoloplos*, and *Phylo*. *See* ERRANTIA.

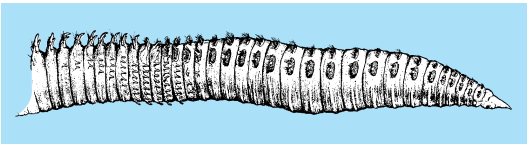

Fig. 1. ***Orbinia*** **(Orbiniidae) in right lateral view.**

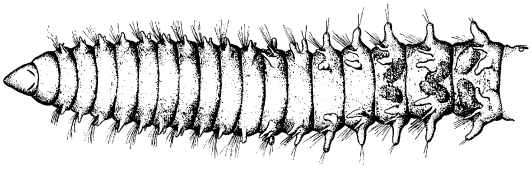

Fig. 2. ***Paraonis*** **(Paraonidae) in dorsal view.**

Paraonidae. This family has 7 genera and about 30 species, most members being small to minute and slender and inconspicuous (**Fig. 2**). They occur in all seas, often to great depths, and sometimes in tremendous numbers. Most have been named in recent years and come from deep seas, for example, the genera *Paraonis* and *Aricidea*.

Spioniform worms. The next 9 families make up the spioniform worms, characterized by the presence of a pair of short to long, grooved palpi near the mouth.

Apistobranchidae. This family has 1 genus and 2 species; both are rare and come from the North Atlantic Ocean and southern California.

Spionidae. This family has about 24 genera and more than 180 species. Some members are economically important; for example, *Polydora* drills in mollusk shells and is an oyster pest, and others foul bottoms of ships and ocean piers. The best-known genera are *Boccardia*, *Laonice*, *Nerine* (**Fig. 3**), *Polydora*, *Prionospio*, *Spiophanes*, and *Spio*. Most of these occur in intertidal or shallow seas and a few in abyssal depths.

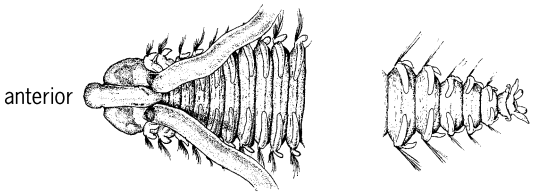

Fig. 3. ***Nerinides*** **(Spionidae) showing the anterior portion and the tail section, in dorsal view.**

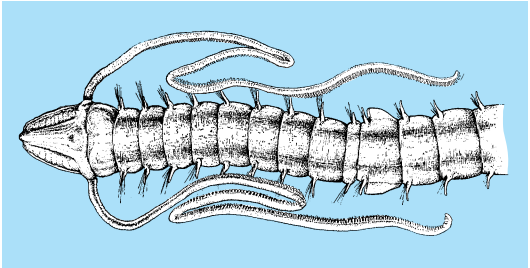

Fig. 4. ***Magelona*** **(Magelonidae) showing the thorax and three abdominal segments, in dorsal view.**

Magelonidae. This family has 1 genus (*Magelona*) and 14 species; they have a worldwide distribution, usually in shallow depths (**Fig. 4**).

Disomidae. This family has 2 genera and 11 species, occurring in cosmopolitan areas in the shelf and at great ocean depths; they include the *Disoma* and *Poecilochaetus*, the latter placed in a separate family by some authorities.

Heterospionidae. This family has 1 genus and 3 species, living in shallow and abyssal depths in the Atlantic and Pacific oceans.

Chaetopteridae. This family has 6 genera and about 31 species (**Fig. 5**). The best-known genus, *Chaetopterus*, is cosmopolitan, occurring in polar and tropical latitudes; its pelagic larvae survive long transport in oceanic currents. Other genera are *Telepsavus*, *Phyllochaetopterus*, and *Mesochaetopterus*.

Flabelligeridae. This family, the cage worms (**Fig. 6**), has at least 15 genera and 110 species. The anterior part of the body is often concealed by a setal cage, formed by setae of the first few segments. The body surface is papillated or covered with mud, and the blood is green. Common shallow-water genera are *Pherusa* and *Flabelligera*. Deep-water genera include *Ilyphagus* and *Brada*. *Flota* is pelagic throughout its life; it is covered by a thick mucoid sheath and has two kinds of setae. *Fauveliopsis* is an aberrant form which lacks the characteristic cephalic structures; it may therefore represent another family.

Poeobiidae. This family has a single, pelagic genus (*Poeobius*) and species from the North Pacific Ocean (**Fig. 7**). It lacks setae but in other respects resembles the flabelligerids.

Psammodrilidae. This family has 2 genera and 2 species, both occurring in the littoral sands in France. The length is 0.04–0.24 in. (1–6 mm) and the characters are aberrant in that the body resembles that of Protodrilidae, but the neuropodial hooks resemble those of Sedentaria. *See* ARCHIANNELIDA.

Fringe worms. This group includes 3 families: the Cirratulidae, Cossuridae, and Ctenodrilidae; they are characterized by long, tentacular structures along the side of the body. The Cirratulidae (**Fig. 8**) has 13 genera and about 120 species. They occur in all seas and at all depths and are important detritus feeders in coastal waters. *Dodecaceria* is associated with the construction of calcareous reefs or drilling in calcareous shells. Other genera are *Cirratulus*, *Chaetozone*, *Cirriformia*, and *Tharyx*. The Cossuridae (**Fig. 9**) have 1 genus (*Cossura*) and 8 species, occurring in worldwide areas. The Ctenodrilidae have 3 genera and 5 species; among the genera are *Ctenodrilus* and *Zeppelina*.

Limivorous (mud-swallowing) worms. Seven families make up this group, which is characterized by an eversible, pouchlike proboscis; some are tubicolous, and others occur in burrows or wander over the surface muds.

Scalibregmidae. This family has about 12 genera and 35 or more species; they are found chiefly in sublittoral and great depths, from slope to deep-ocean bottoms. Included in the genera are *Oncoscolex*, *Polyphysia*, *Scalibregma*, and *Sclerocheilus*.

Opheliidae. This family has 10 or more genera and more than 100 species; all are errantiate in habit or live in burrows without tubes. The most abundant and ubiquitous genus in intertidal zones is *Polyophthalmus*, which resembles a large nematode. Other well-known genera are *Ophelia*, the malodorous *Travisia*, *Armandia*, and *Ammotrypane*. An intertidal population of *Thoracophelia* in southern California was estimated to number about 158,000,000 individuals in a mile of beach sand and to total 7 tons (6.3 metric tons), wet weight; 14,600 tons (13,140 metric tons), of organic matter in sand has been cycled annually. Other limivorous annelids perform similar functions in recycling the sediments. *See* NEMATA.

Sternaspidae. This family has 1 genus, *Sternaspis*, sometimes called the gooseberry worm; it has several named species which are not clearly identifiable.

Capitellidae. Sometimes called bloodworms because of the dark-red color, this family has about 24 genera and more than 70 species (**Fig. 10**). They abound in intertidal mud flats and estuaries, and some are tolerant to varying salinities. Although chiefly mud-swallowers, some invade egg capsules of sea squids,

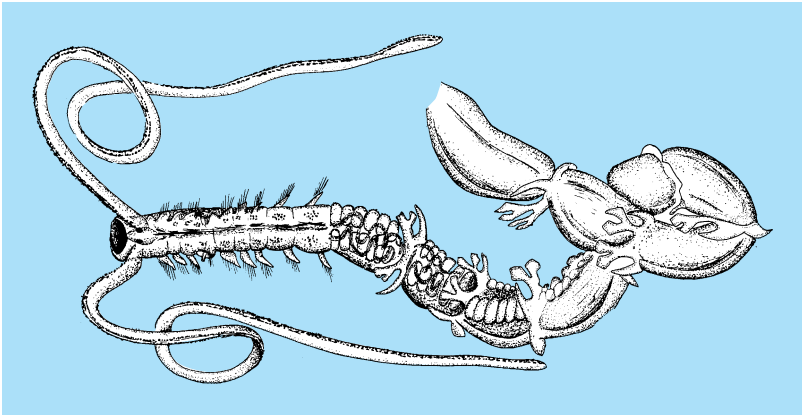

Fig. 5. *Telepsavus* (Chaetopteridae) in dorsal view.

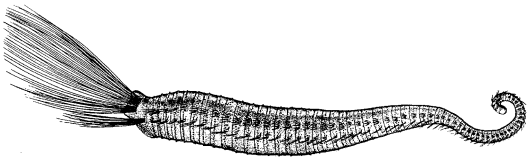

Fig. 6. *Pherusa* (Flabelligeridae) or cage worm, shown in right lateral view.

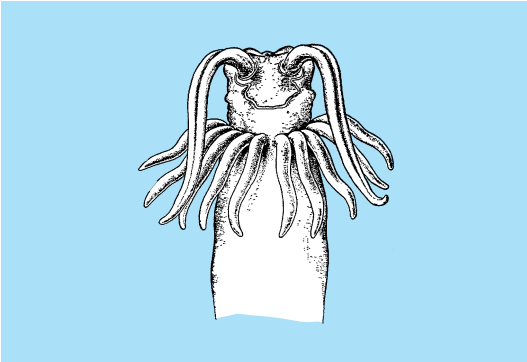

Fig. 7. Anterior end of *Poeobius* (Poeobiidae) in dorsal view.

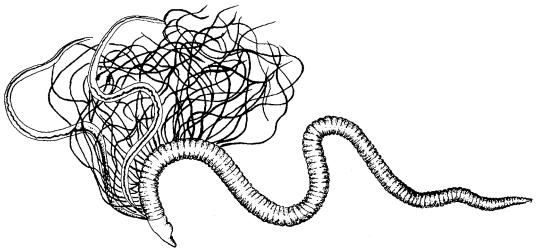

Fig. 8. *Chaetozone* (Cirratulidae) in left lateral view.

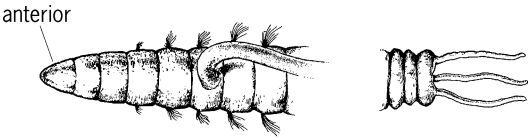

Fig. 9. ***Cossura*** **(Cossuridae) anterior and posterior ends in dorsal view.**

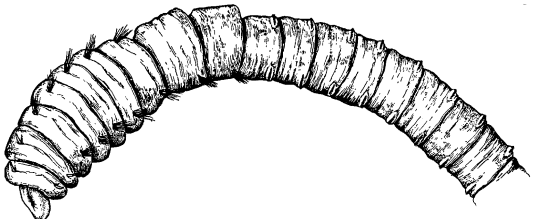

Fig. 10. ***Capitomastus*** **(Capitellidae) in left lateral view.**

and others occur in brackish and fresh water. One member is host to a monstrillid copepod. *See* COPEPODA.

Stypocapitellidae. This family takes its name from a monotypic genus for a small species from western Germany.

Arenicolidae. Known as the lugworms, this family has 4 genera and about 16 species. They are chiefly littoral, in sand and mud flats, where they form far-reaching colonies.

Maldanidae. Commonly called bamboo worms, this family has about 30 genera and 200 species, living in worldwide seas and in all oceanic depths (**Fig. 11**). They sometimes form beds of great extent in shelf and slope depths. Some are large and multisegmented; others have definite (often 19) setigerous segments and short bodies. Five subfamilies are generally recognized: Maldaninae have cephalic and anal plaques with the anal aperture located dorsally; common genera are *Asychis* and *Maldane*. Euclymeninae have well-developed plaques and an anal pore within the plaque; common genera are *Axiothella*, *Clymenella*, *Clymenopsis*, and *Clymenura*. The Lumbriclymeninae are less modified and primitive in some respects; they include *Lumbriclymene*, *Praxillura*, and *Notoproctus*. The Nicomachinae include *Nicomache*, and the Rhodininae have *Rhodine*, often with segmental flanges.

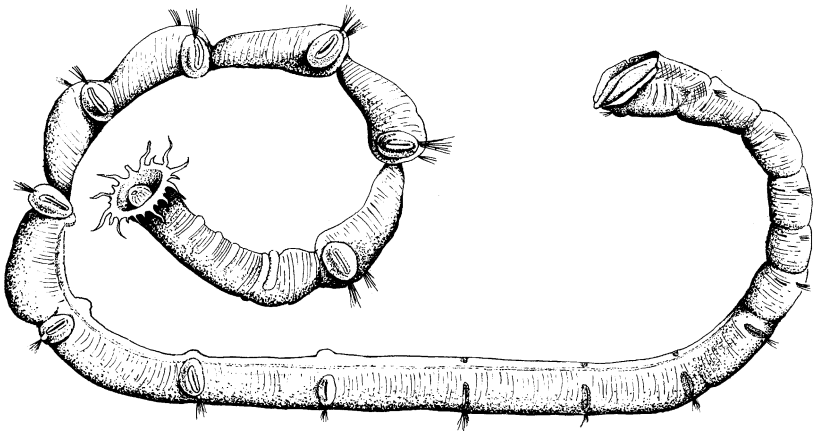

Fig. 11. ***Axiothella*** **(Maldanidae), in left lateral view, showing cephalic and anal plaques.**

Oweniidae. This family has 5 genera and about 20 species. *Owenia* is a shallow-water form, sometimes abundant in tropical inlets of sand. *Myriochele* occupies tapering tubes (**Fig. 12**) of characteristic structure, in which the animal is securely contained. *Myriowenia* (**Fig. 13**) is modified with a pair of long palpi. The remaining families are characterized by an opercular or tentacular crown which conceals head structures.

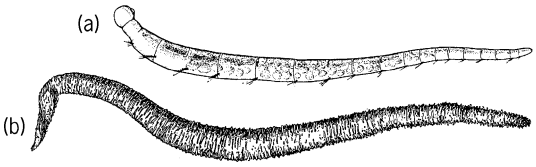

Fig. 12. ***Myriochele*** **(Oweniidae) with (*a*) entire ovigerous individual and (*b*) tube in right lateral view.**

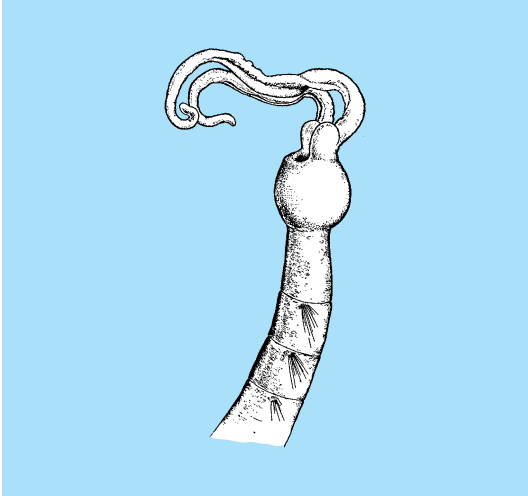

Fig. 13. ***Myriowenia*** **(Oweniidae) anterior end with long, grooved palpi, in left lateral view.**

Sabellariidae. Members of this family, called sand-cementing worms, have a compact operculum formed of setae of the first several segments (**Fig. 14**). The family has 7 genera and about 57 species, living at intertidal to abyssal depths in worldwide areas. Common genera are *Idanthyrsus*, chiefly tropical; *Phragmatopoma*, concentrated in the eastern Pacific; *Sabellaria*, well known from worldwide areas; and *Gunnarea* in South Africa. *Lygdamis* and *Phalacrostemma* are abyssal genera.

Pectinariidae. Commonly called cone worms, this family has 5 genera and about 46 species, occurring from intertidal sandy beaches to great oceanic depths. Common genera are *Cistenides* and *Pectinaria* from shallow bottoms and *Petta* from great depths.

Ampharetidae. This is a large, chiefly deep-water group, which includes more than 50 genera and at least 160 species. Some members occur in intertidal sands, and one group (*Hypania* and *Hypaniola*) occurs in the inland seas of Europe. They

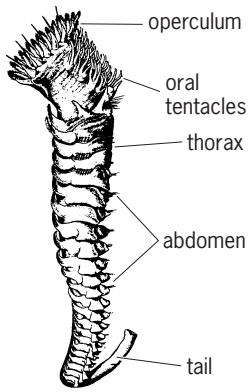

Fig. 14. ***Sabellaria*** **(Sabellariidae) in right lateral view.**

are characterized by retractile oral tentacles, unique cephalic structures, and highly modified parapodial parts. Most members construct tubes of specific design. The blood is green, and the body sometimes is highly colored. The family has 2 or 3 subfamilies: the Melinninae and Samythinae have a conspicuous dorsal membrane, with or without dorsal spines; and the Ampharetinae lack a dorsal membrane. Most members occur in slope and abyssal depths of the ocean; a few are intertidal, and *Hypania* and its allies are limited to the fresh waters of European inland seas. Included in the genera are *Amage*, *Ampharete*, and *Sabellides* in the Ampharetinae, and *Melinna* and *Melinnexis* in the Melinninae.

Terebellidae. This family has about 48 genera and more than 320 species. They are chiefly large, thick-bodied, tubicolous, rarely free-living worms, with the anterior end covered by a matted mass of tentacular cirri. Some harbor commensal polynoids in their tubes; others have commensal copepods. The family is usually divided into 4 subfamilies: the Amphitritinae with the genera *Amphitrite*, *Eupolymnia*, *Loima*, *Pista*, and *Terebella*; Polycirrinae with *Amaeana*, *Enoplobranchus*, and *Polycirrus*; Thelepinae with *Euthelepus*, *Thelepus*, and *Streblosoma*; and Artacaminae with *Artacama* and *Artacamella* (**Fig. 15**).

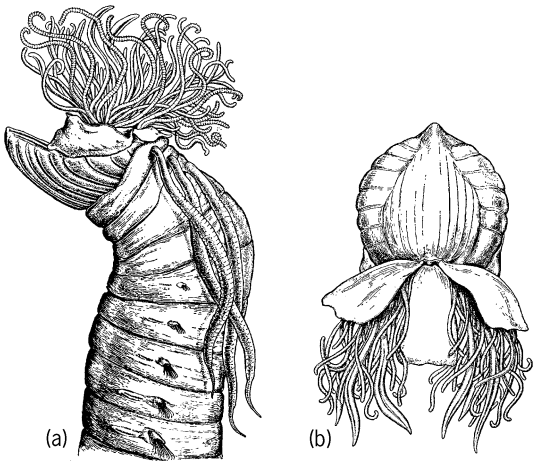

Fig. 15. ***Artacamella*** **(Terebellidae, Artacaminae). (*a*) Left lateral view. (*b*) Dorsal view.**

Trichobranchidae. This family has 5 genera and about 24 species, chiefly occurring at considerable ocean depths. *Terebellides* is the best known; all the other members of this family are rare and live at great ocean depths.

Sabellidae and Serpulidae. These 2 families make up the feather-duster worms. The Sabellidae have 31 genera and more than 257 species, often occurring in intertidal depths but descending to great abyssal depths. They are sometimes brilliantly colored, in conspicuous massed beds; some attain gigantic proportions, with a body length of 12–16 in. (30–40 cm); most are much smaller. The family has 3 subfamilies: The Sabellinae are the most numerous and the largest and include the genera *Eudistylia*, *Sabella* (**Fig. 16**), *Bispira*, *Hypsicomus*, and *Megalomma*. The Fabriciinae are chiefly small to minute, colonial, and intertidal or estuarine, but sometimes occur in deep water; common genera are *Chone*, *Euchone*, *Fabricia*, *Jasmineira*, and a fresh-water genus, *Manayunkia*, which is abundant in the Great Lakes district and its tributaries. The Myxicolinae are best known for *Myxicola*, which has been studied for its great neural fiber system.

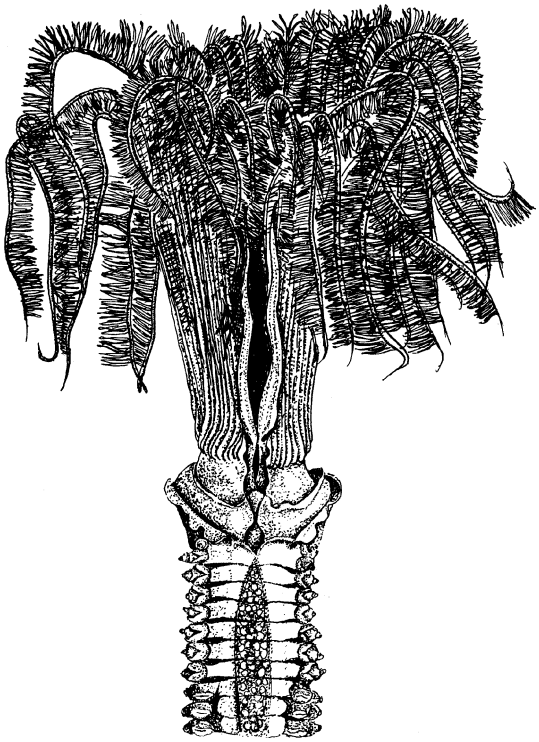

Fig. 16. ***Sabella*** **(Sabellidae) in dorsal view.**

The Serpulidae have more than 55 genera and 340 species, occurring in cosmopolitan areas and at all depths. They construct calcareous tubes, which may be so abundant as to clog drains and waterways. The peristomium is modified as a tentacular crown, with one filament modified as an operculum (**Fig. 17**). Common genera are *Crucigenia*, *Eupomatus*, *Hydroides* (a fouling form with cosmopolitan distribution), *Mercierella* (a euryhaline or estuarine form of

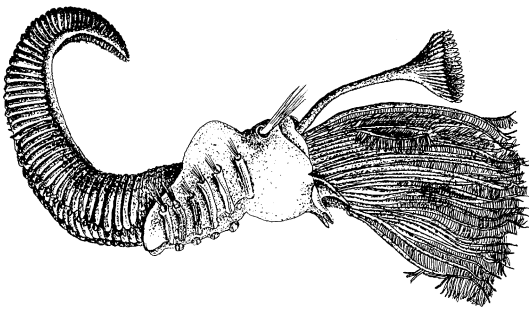

Fig. 17. *Serpula* (Serpulidae) in right lateral view.

economic importance), *Protula*, and *Serpula*. The Filograninae include *Filograna* and *Salmacina*. *See* ANNELIDA.

Olga Hartman

Bibliography. O. Hartman, *Catalogue and Index of the Polychaetous Annelids of the World*, Hancock Foundation, University of Southern California, 1959–1965; B. H. McConnaughey and L. F. Fox, The anatomy and biology of the marine polychaete *Thoracophelia mucronata* (Treadwell), Opheliidae, *Univ. Calif. (Berkeley) Publ. Zool.*, 47:319–328, 1949; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; D. P. Wilson, On the Mitraria larva of *Owenia fusiformis delle Chiaje*, *Phil. Trans. Roy. Soc. London*, ser. B, 221:231–334, 1932.

Sedimentary rocks

Rocks formed of any of the materials that accumulate at Earth's solid surface at ambient temperatures, regardless of their consistency. Together with lava flows, sedimentary rocks form a cover of stratified material, known as the stratisphere, that, lying on igneous and metamorphic rocks, covers over two-thirds of the lands and nearly all of the ocean floor. Generally measured in hundreds to a few thousand meters, it locally reaches up to 15 km (9 mi) in thickness. *See* EARTH; IGNEOUS ROCKS; METAMORPHIC ROCKS; ROCK.

Most sediments accumulate as sand or mud deposited by air or water under the influence of gravity. Because of changes in the material supplied and in the conditions of deposition, the resulting accumulations are generally layered (stratified) in a near-horizontal position, with each stratum representing an event or episode (**Fig. 1***a*).

Derivation of sediment. In theory, the Earth's surface forms an oblate spheroid (geoid). The oceanic surface comes close to this. However, the brittle shell (lithosphere) of the solid Earth, floating on a viscous asthenosphere, rises to different elevations, thus deviating from this ideal. Its surface includes the vast deep-sea floor with an average elevation of 3800 m (12,464 ft) below sea level and a continental platform with an average elevation of 840 m (2755 ft) above sea level. The plane of reference or base level to which erosion is reducing the lands is sea level. Most sediment is transported to the oceans, where it is deposited, along with lava flows, building up the bottoms toward this base level. Other forces—tectonics and terrestrial volcanism—keep erosion and deposition from achieving these ends. Plate tectonics is always renewing parts of the ocean floor and subducting others into the Earth's interior, and is always raising new mountain ranges. Two other factors complicate the process. Parts of the continents sink progressively over long periods, trapping sediments before they reach the sea. Also, the sea level oscillates, flooding continents to deposit shallow-water marine sediments, and then withdrawing and exposing them to erosion. *See* ASTHENOSPHERE; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; EROSION; LITHOSPHERE; PLATE TECTONICS; VOLCANO.

Weathering. The formation of sediment begins with weathering, the interaction of rock with the atmosphere. Some minerals are dissolved by rainwater. Others, such as the crystals of feldspar in granites, are decomposed by carbonic acid (H_2CO_3) in rainwater. These minerals break down into hydrous aluminum silicates (clay minerals), and release other metals and silicon dioxide (silica) to solution. This, along with heating and cooling, wetting, and drying, and frost action breaks up the rock. *See* CLAY MINERALS; GRANITE; SILICA MINERALS; WEATHERING PROCESSES.

This resulting residue of detritus is generally a mixture of clay and coarser fragments (clasts) subject to further weathering, and is known as the regolith. Organic compounds, such as wood and pollen grains, generally become incorporated. Clasts are also provided by surf tearing at sea cliffs, by the plucking and grinding of glaciers, and, rarely, by meteorite impact that blasts craters and ejects clasts, glass droplets, and vapor. *See* REGOLITH; SEDIMENTOLOGY.

Detrital transport and deposition. The regolith moves down slopes. At the same time, it becomes eroded by water, wind, and ice feeding dust to the air. Mostly the particles follow the gravipotential gradient to lower elevations. Some are trapped in subsiding continental basins, but most ultimately reach the ocean. Skeletal elements of organisms, added in transit and at depositional sites, may become incorporated and permanently preserved as fossils.

Some of this transport occurs en masse in mudflows or glaciers and in the bulk flow of aerial and aqueous suspensions (turbidity currents). However, most involves the movement of individual particles whose properties, above all, size, interact with fluid motion to sort them into separate populations. *See* TURBIDITY CURRENT.

Particles are classified primarily by their mean diameter (**Fig. 2**). Sand grains, pebbles, cobbles, and boulders are moved by traction—dragged, rolled, or bounced along the substrate—while the transport of clay is limited to suspension in the air, water, or ice. Deposition occurs as the wind or water slackens. Much of the clay begins in colloidal sizes, which coagulate into larger sizes. These coagulates tend to be filtered out by animals and generally are packaged with the animals' digestive waste into sand-grade pellets that, even in the deep ocean, reach the bottom in days or weeks. Dust particles in the air are caught by raindrops and delivered to the land or water surface.

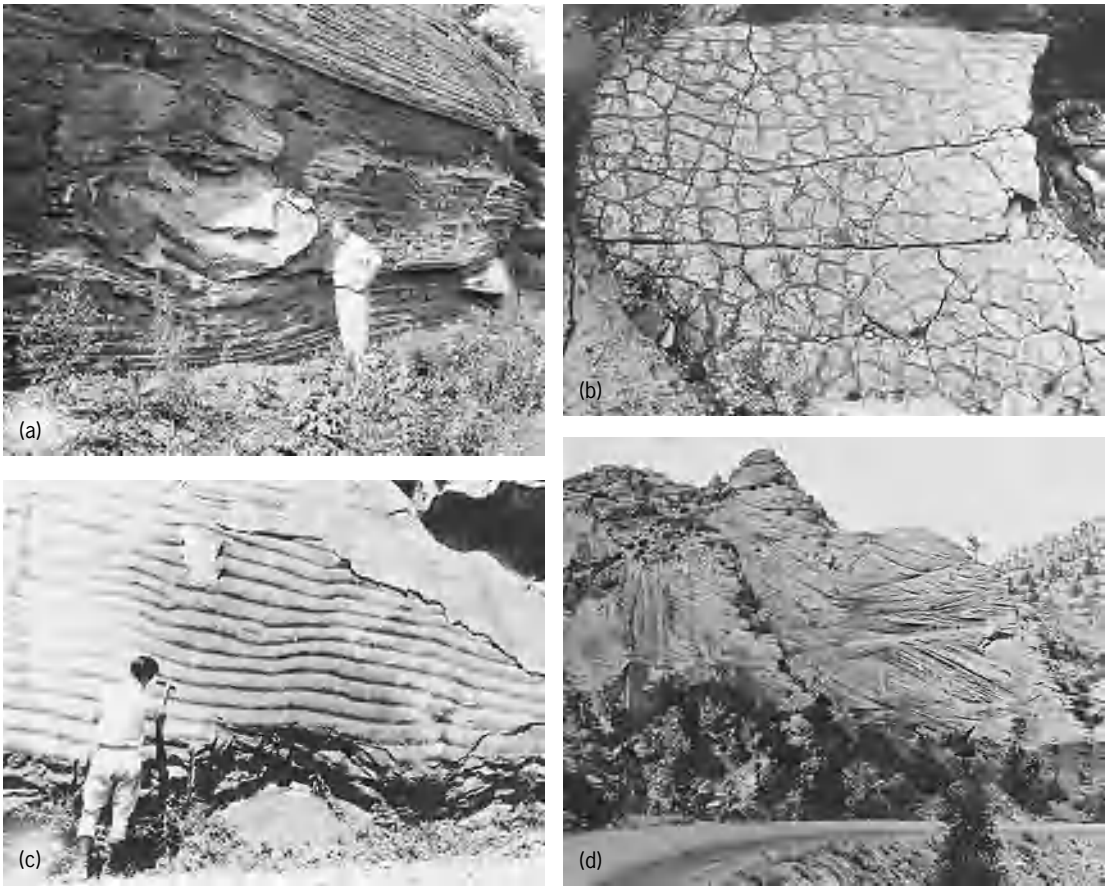

Fig. 1. Structures in sedimentary rocks. (*a*) Stratification in Carboniferous shales and sandstones near Montgomery, West Virginia, and a concretion. (*b*) Mud cracks in Lower Carboniferous limestone northeast of Bluefield, West Virginia. (*c*) Current ripples in Cambrian dolomite northeast of Bluefield, Virginia (*courtesy of Carl O. Dunbar*). (*d*) Dune cross-bedding in Jurassic Navajo Sandstone near entrance to Zion Canyon, Utah (*Virginia Division of Natural Resources*).

Hydrodynamics are not limited to the interaction between individual particles and the suspending fluid, but can involve bodies of water loaded with suspended sediment that travel en masse down subaqueous gradients as turbidity currents. These erode the sea floor as they gather speed and deposit sediment as they slow and spread out over deep-water plains. The resulting bed (turbidite) is texturally graded with the coarser particles, which settle out most rapidly, at the base and progressively finer particles above these. Millimeter-scale turbidites show a gradation from silt to clay, thus recording the settling-out of suspensions caused by a storm in a lake or marine basin. Centimeter-to-decimeter-scale beds, grading from sand to clay and traceable for thousands of km^2, record centurial or millennial land- or mud-stirring events such as earthquakes that triggered landslides or tsunamis. In addition, these beds characterize large bodies of sediment deposited in abyssal plains or trenches fronting continents or island arcs such as the flysch of the Northern Alps and similar deposits in most major mountain ranges. Events of truly extraordinary violence are recorded in meter-scale deposits, such as the Miocene turbidite with boulders in its base that has been traced for hundreds of kilometers in the Apennine Mountains, Italy. *See* MIOCENE; TURBIDITE.

Deposition may be temporary, with resumption of transport when the velocity of the transporting fluid increases. Thus detrital matter is moved in stages; for example, by streams into lakes and the sea, and then by waves and currents. Deposits along the way are temporary unless subsidence takes them below base level. Weathering continues throughout, and wear and tear reduces the clasts in size, rounding them and grinding up the softer minerals to gradually

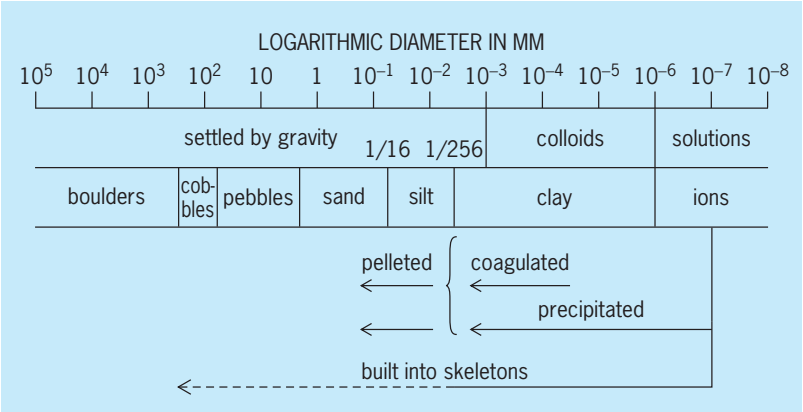

Fig. 2. Range of sedimentary particle sizes (Wentworth logarithmic scale). (*Adapted from W. Krumbein and L. L. Sloss*)

diminishing the less stable constituents in favor of quartz and clay minerals. *See* QUARTZ.

Transport and deposition of chemical sediments. Chemical sediments may derive from direct precipitation from sea water or from physical accumulation of skeletal matter precipitated by organisms. Solutes move with the water cycle. While dissolution occurs largely within the rock or regolith, the waters emerge in springs, flow as streams, and rest in lakes. Solutes are precipitated (Fig. 2) when concentrated to saturation to grow isolated crystals ranging in size from clay to sand or coarser grades, or to form aggregates of varying size, which may undergo mechanical transport or form solid crusts. In the last several hundred million years, most chemical sediments have derived from marine organisms, which extract calcium carbonate, silica, and calcium phosphate from water to build skeletons. After the organisms' death these skeletons may remain in place or become mechanically transported in the manner of detrital sediment.

Pyroclastic transport and deposition. Lumps of magma ejected from volcanoes freeze into ballistic volcanic "bombs" and help to build cones of sediment. Sand-grade crystal tuffs reach a wider distribution, and finer volcanic ash (microscopic shards) will travel over long atmospheric distances. Nuées ardentes, suspensions of glowing shards in hot gases, rush down the land surface to deposit a layer of consolidated ash (ignimbrite). Submarine eruptions yield the frozen glassy froth, called pumice, which contributes to deep-sea sediments and floats to beaches. *See* IGNIMBRITE; MAGMA; PUMICE; PYROCLASTIC ROCKS; TUFF.

Sedimentary structures. Bodies of sediment may be planar (sheetlike) as in Fig. 1*a* or pod-shaped as in dunes or reefs. Strata at the centimeter-to-meter scale are generally called beds, while those at the millimeter and micrometer level are distinguished as laminae. Typical dunes retain their pod shape, moving by transfer of sand from the windward to the lee side, where it accumulates in steeply inclined cross-laminae (Fig. 1*d*). Sand in water commonly forms ripples, which may grow into cross-laminated sheets (Fig. 1*c*). Most strata are formed by vertical accretion, resulting from a rain of suspended particles, and derive their stratification from changes in the particle size or composition of their components. *See* DUNE.

Polygonal shrinkage cracks (Fig. 1*b*) and large pores document the desiccation of ponds and intertidal flats. Animals leave tracks and trails on the surface of strata or may burrow for purposes of shelter or feeding in a process of bioturbation. In contrast to actual remains of organisms (body fossils), such structures are known as trace fossils. *See* FOSSIL.

Diagenesis. All sediments are overprinted by post-depositional (diagenetic) changes. Compaction results from the compression of the sediment under load, expelling pore water. It also tends to bring the platelike (platy) shapes of clay minerals into alignment with the bedding. Most mudstones show fissility parallel to bedding and are called shales. With burial to several kilometers, some clay minerals will give up some of their water, leading to further compaction. *See* DIAGENESIS; SHALE.

One of the first chemical overprints suffered by most sediments is chemical reduction. Weathering and transport generally occur in oxygen-rich environments, as does most deposition. But in most aqueous deposits (excepting the deep ocean floors where sediment supply is limited to wind-blown dust) the organic matter in the sediment exceeds the available oxidizing capacity. Bacteria use up the oxygen in the pore waters to decompose organic matter and decompose nitrate and sulfate to obtain more. The resulting oxygen deficit breaks down the ferric oxides, leading to the formation of sulfides and turning red and brown shades to green and black. *See* OXIDATION-REDUCTION.

Slow interaction between sediment and pore waters will gradually dissolve the less stable minerals, such as aragonite ($CaCO_3$) or magnesian calcite, to precipitate purer calcite or dolomite [$CaMg(CO_3)_2$] crystals. While these may replace the original grains, solution transfer may leave cavities to precipitate this solute to other sites, where the growth of crystals occludes pores and cements loose particles into solid rock. Solution transfer also dissolves load-bearing grains at points of contact (**Fig. 3**), precipitating on the walls of pores to convert porous and friable sandstone into hard, dense rock (solution welding). *See* ARAGONITE; CALCITE; DOLOMITE.

Local chemical gradients, such as caused by the decay of organisms, may induce the precipitation of calcite or siderite ($FeCO_3$) to form the dense nodules termed concretions (Fig. 1*a*). Large-scale chemical changes may dissolve disseminated fossils of opaline silica (sponge spicules, diatoms, radiolaria) to reprecipitate it as quartz in discrete nodules or beds of chert. *See* CHERT; CONCRETION.

In evaporite settings, magnesium ions may react with aragonitic calcium carbonate to form dolomite before burial. Large-scale conversion of calcite to dolomite occurs where evaporitic solutions penetrate buried limestone. In the Florida-Bahama platform, much of a Tertiary carbonate sequence up to 2 km (1.2 mi) thick has been converted to more coarsely crystalline dolomite. *See* LIMESTONE.

Of particular economic importance is the diagenesis of organic matter. Peat accumulated in bogs becomes compacted by burial and loses hydrogen in the form of methane to become first lignite, then bituminous coal, and, under high pressure, anthracite. *See* PEAT; COAL.

In lakes and the sea where phytoplankton are the main contributors of organic matter, bacteria in the sediment consume it with the available oxygen, including that obtained from nitrates and sulfates. Continuing anaerobically with release of methane, they convert the remaining organic matter into the waxy-appearing kerogen that is insoluble in most solvents. Burial to greater depths and higher temperatures results in the chemical cracking of kerogen, with the liberation of gases (such as methane, butane, and pentane) and liquid petroleum. *See* KEROGEN; PETROLEUM.

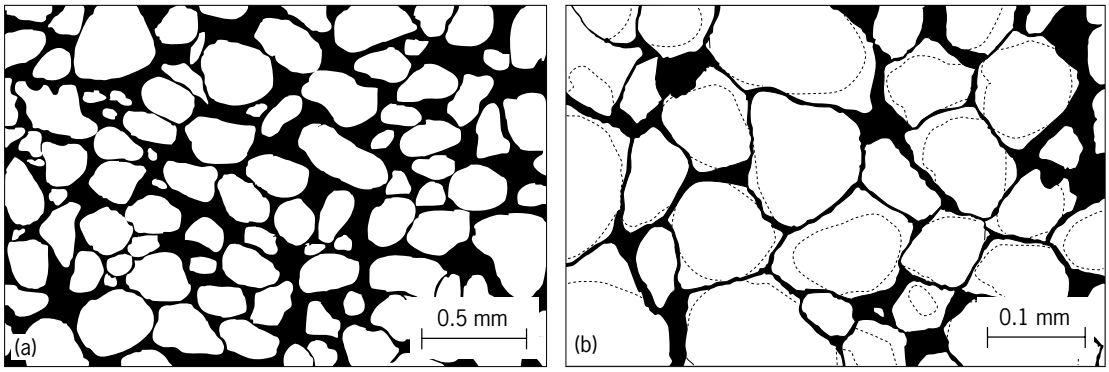

Fig. 3. Solution welding in sandstones. (*a*) Thin section of Jurassic Entrada Sandstone, Utah. The grains (white) are well sorted and rounded, and porosity (black) has been largely preserved. (*b*) Sandstone from Devonian Cedar Valley Formation, Iowa. Outlines of the original rounded grains, revealed by cathodoluminescence, are shown by thin lines. Dissolution of quartz at grain contacts has welded them together, while the silica, thus dissolved, was precipitated on the free surfaces in optical continuity with the sand grains to form a hard sandstone with reduced pore space. (*After P. A. Scholle, A Color Illustrated Guide to Constituents, Textures, Cements and Porosities of Sanstones and Associated Rocks, Amer. Assoc. Petrol. Geol., Mem. 28, 1979*)

Classification of sedimentary rocks. Sedimentary rocks are mixtures of various components, which in theory could be divided into detrital, chemical, and biogenic fractions. However, the breakup of skeletal matter and the diagenetic overprint make it difficult to assess the chemical versus biogenic components (making it necessary to lump them), while division of the detrital components into sand-silt and clay fractions remains apparent. Most sedimentary rocks are therefore assignable to a four-component system in which the end members are sand and silt (mostly quartz), carbonate, and nondetrital silica (chert). The mixtures of these can be visualized within a tetrahedron in which the great majority of rocks fall close to the sand-clay-carbonate triangle (**Fig. 4**). Individual strata that fall outside this scheme, mainly chemical precipitates, are dealt with as miscellaneous sediments.

Detrital sedimentary rocks. Rudites are detrital rocks dominated by clasts in the pebble-cobble-boulder range (Fig. 2). They include the loose gravels of stream or beach action, or their cemented equivalents (conglomerates). Rudites with sharply angular clasts, distinctive of short travel from their points of origin, are distinguished as sedimentary breccia, as distinct from breccias of tectonic origin that, cutting across strata, are not part of the stratisphere. *See* BRECCIA; CONGLOMERATE; GRAVEL.

Gravels are sought for road building and coarse aggregates in concrete. Deposited in settings of high water velocity, their finer accessories tend to be dense minerals, such as gold, silver, compounds of tin and uranium, and diamonds, which may form economically significant placer deposits.

Arenites or sandstones are abundant on the continental platforms. Here they may accumulate as dune sands, in streams, as beach sands, and in the shallows of lakes and seas. Turbidity currents carry sands to the depths of lakes and to the abyssal plains and trenches fronting the continents. *See* SANDSTONE.

Sandstones derived from immaturely weathered granites, retaining much feldspar, are termed arkose; those derived from metamorphic terranes, containing fragments of schist, slate, and greenstone are termed graywacke; while mature weathering preserving mostly the grains of quartz are called orthoquartzite. *See* ARKOSE; GRAYWACKE.

Texturally, sandstones may be divided into well-sorted grainstones whose open pores lend them porosity and permeability (not necessarily retained, as shown in Fig. 3*a*). In packstone, the presence of much fine material fills the interspaces between sand grains, while in wackestone the sand grains are floating in the finer matrix.

Porous and permeable sandstones are major sources of water and petroleum (gas, oil). Sandstone was widely used for building, and sand is an important aggregate for concrete and tarmac. Like rudites, sandstones serve as ores where they have concentrated silt-grade grains of heavy minerals such as gold and silver and minerals of tin and titanium.

Siltstone is a term generally reserved for silt-grade sedimentary rock that lack appreciable admixture

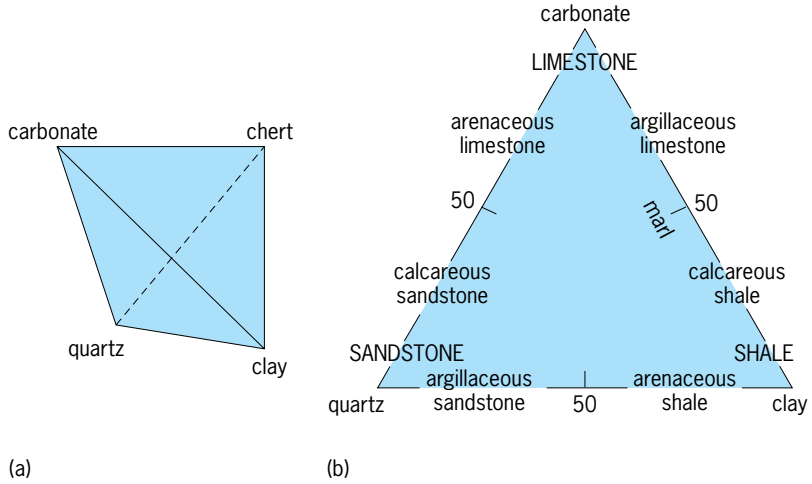

Fig. 4. Classification of the common sedimentary rocks. (*a*) Tetrahedron showing the main end members: detrital sand and silt (mainly quartz), clay fraction (mainly clay monerals), carbonates (generally a mixture of chemical and biogenic components), and siliceous chemical and biogenic components (chert). (*b*) Carbonate-quartz-clay face shown in detail. (*After W. Krumbein and L. L. Sloss*)

of clay. A notable example is loess, a deposit of wind-blown dust that accumulates on the lee side of deserts and peripheral to ice sheets and is prized for its agricultural properties. *See* LOESS.

Argillaceous or clay-bearing rocks (mudstones, shales) are the most abundant of all sedimentary rocks. They are initially plastic, but compaction turns them brittle and generally aligns their platy clay minerals to lend them a bed-parallel fissility, for which they are generally referred to as shales. Most are silty, and many are mixed with carbonate and then termed marls. Most extensive are the red clays of abyssal oceanic depths, consisting largely of the finest of wind-carried dust. *See* ARGILLACEOUS ROCKS; MARL.

Clays are an essential ingredient of portland cement, and are the raw materials for ceramics, from bricks to pottery. Some are mined as fuller's earth for filtering and other industrial purposes. Others, rich in zeolites, are mined for their base-exchange properties. Oil shales are rich in kerogen, and will yield petroleum on thermal cracking. Clays derived from regolith reduced to bauxite by intense tropical weathering are our source of aluminum. *See* BAUXITE; FULLER'S EARTH; OIL SHALE; ZEOLITE.

Carbonate sediments. Carbonates form more than 20% of sedimentary rocks. Rain contains dilute carbonic acid; hence waters becomes calcium-enriched in weathering. The air in soils, enriched in carbon dioxide (CO_2) due to organic respiration, adds to the calcium-carrying capacity of ground water. Where calcium carbonate–saturated waters reequilibrate with the atmosphere in caves, calcite is precipitated in hanging stalactites, standing stalagmites, and the gravity-defying helictites—all included in the term speleothems. In spring-fed water bodies, this reequilibration, aided by photosynthesis of plants, may precipitate calcite in porous but coherent (boundstone) aggregates. These calcareous tufas or travertines may dam streams and form waterfalls. Cyanobacteria and algae may coat pebbles with calcite and in lakes may build calcitic mounds (stromatolites). The photosynthesis of lacustrine phytoplankton may precipitate calcite mud (micrite or calcilutite) or mix with clay or marl. Skeletal elements of calcite include the egg capsules of the charophyte algae and the shells of mussels and snails. Buried whole or in fragments, these may accumulate as limestone, which may have been sorted into calcarenites divisible, like detrital sandstones, into carbonate grainstones, packstones, or wackestones. *See* CARBON DIOXIDE; STALACTITES AND STALAGMITES; STROMATOLITE; TRAVERTINE; TUFA.

Similar processes operate in the oceans. Due to the interference by magnesium, chemical precipitation at times, such as the present, yields the mineral aragonite rather than calcite. To depths of 4–5 km (2.5–3 mi), oceanic waters are generally carbonate-saturated. But as this saturation varies directly with pressure and inversely with temperature, water masses adjust their carbonate content by precipitation or dissolution. At greater depths (below the calcite compensation depth), the water is undersaturated and the carbonate constituents are dissolved. *See* MARINE SEDIMENTS.

Upwelling waters warmed on the Bahama Banks precipitate micritic aragonite muds, with or without aid from photosynthesis, or form round sand-size pellets (ooids) which, sorted by waves and currents, accumulate to form oolite, a rock found in all geologic ages. *See* OOLITE.

On tropical intertidal flats free of detrital matter, surficial mats of cyanobacteria and algae precipitate laminated and commonly crinkly laminae of micritic aragonite with characteristic large pores (fenestra), while in shallow water they produce mounds (stromatolites) in some restricted areas. Stromatolites were far more widespread in the geological past and dominated the carbonate deposits in the Precambrian seas. *See* PRECAMBRIAN.

Beginning in Cambrian time, some 550 million years ago, protists (for example, algae) and metazoans began to build skeletons out of calcite or aragonite and now dominate carbonate deposition. At present, the largest depositors are the prymnesiophyte algae (coccoliths) and the planktonic foraminifera of the oceanic plankton. *See* CAMBRIAN.

Deposition by organisms commonly produces a wide range of particle sizes. Storms in shallow waters commonly sort out shell beds (coquinas). *See* COQUINA.

Carbonate boundstones are chiefly produced in coral reefs and mainly held together by the colonies of red algae, corals, and other invertebrates. Surf and tsunamis break the reef rock into blocks that may accumulate as massive talus deposits of calcirudite. Mountain ranges, such as the Dolomites (in the southern Alps) or the Guadalupe Mountains of Texas-New Mexico, owe their massive cliff faces to ancient reefs bodies of boundstone and rudite, hundreds of meters thick. *See* TALUS.

Calcarenites include oolite, concentrations of sand-size skeletons such as the tests of foraminifera, the ossicles of echinoderms, or fragments of larger shells, and may form carbonate beach sands, as in the Florida Keys. *See* CALCARENITE.

Chalks are marine carbonate rocks of mudstone grade, consisting largely of coccoliths—complex plates of calcite secreted by prymnesiophyte algae—and of the tests of planktonic globigerinacean protozoans. Since the Late Jurassic, they have been the most widespread sedimentary rock on Earth, but have played a relatively small role on the continental shelves, except during Cretaceous times. *See* CHALK; CRETACEOUS; JURASSIC.

Carbonate rocks have great importance as reservoirs for water and petroleum. They are crushed for road gravel and smelting flux, and burnt for quicklime and portland cement. Many of the great buildings of England are made of the Jurassic limestones of Bath, and in the United States of the Carboniferous limestones of Indiana. Italian travertine is widely used as a building material. The Jurassic Solnhofen limestone of Germany provides lithographic stone. Much of the ornamental marble used worldwide is not true metamorphic marble but dense limestone, often fossiliferous such as the Devonian reef rocks of Belgium and the Jurassic Rosso Ammonitico of Italy. *See* MARBLE.

Siliceous chemical and biogenic sediments. Silica (SiO_2) is second only to carbonate in the solutes derived from weathering. In early times, it was presumably precipitated as silica gel and in silicates, but skeletal organisms since Cambrian time have extracted most of it to produce skeletal elements of opal, an amorphous and hydrated form of silica. These, in the form of sponge spicules and the microscopic tests of radiolarians and diatoms, cover large parts of the ocean floor with radiolarian and diatomic ooze. Diagenesis can transform these deposits of silt-grade opal into the lightly cemented diatomites, such as those of California's Miocene, with alteration to opal-CT (platy, dehydrated and crystalline), but ultimately converts them into dense, hard chert or flint, consisting of quartz. Diatomite is mined for filters, abrasives, and insulation, while flint was long vital to the lighting of fires and discharge of guns. *See* DIATOMACEOUS EARTH; OPAL.

Evaporites. In arid regions, evaporation of constricted water bodies raises solute contents to the level of precipitation. Evaporation of seawater sequentially yields a little carbonate, then much calcium sulfate (gypsum or anhydrite), then sodium chloride (halite), and then salts, mainly chlorides of potassium and magnesium. Gypsum is the source of plaster of paris, halite is the source of table salt and ice-clearing brines. Potash salts, such as mined in the Permian of Germany and New Mexico, are a major source of fertilizer. *See* ANHYDRITE; GYPSUM; HALITE; SALINE EVAPORITES.

Evaporites formed from continental salt lakes produce strata of different minerals such as the sodium carbonate mineral, trona, mined in the Eocene of Wyoming or the borax deposits in the Eocene of California.

Carbon-hydrogen sediments. In marshes, the remains of higher plants, such as Sphagnum moss and woody materials rich in cellulose and lignin, may accumulate where bacteria are inhibited by acidity. Diagenesis converts the resulting peat first into lignite, then into bituminous coal, and at high pressures into anthracite. These most abundant fossil fuels are mined in strata that range up to tens of meters thick, while the coal-associated methane is now produced as well.

While organic matter in marine settings rarely became a dominant rock component, the kerogen formed by diagenetic processes is sufficiently enriched in oil shales to constitute a potential source of petroleum for the future.

Miscellaneous chemical sediments. In phosphorites, apatite skeletons of certain brachiopods or vertebrate teeth and bones are generally associated with phosphatic pellets of sand to granule grade. They appear to be by-products of bacterial decomposition, which may enrich the phosphate content of pore waters to saturation. Such deposits form chiefly at sites of marine upwelling that localize organic productivity. Deposits of Late Cretaceous to early Tertiary times off North Africa or in Miocene times in Florida have given rise to large fertilizer industries. *See* APATITE.

Sedimentary iron ores are of several types. Bog iron ores of ochreous limonite form where reduced acid waters carrying dissolved ferrous iron meet the atmosphere, as in the bogs of New Jersey, and where bacteria make their living by catalyzing the oxidation reaction. Such ores were formerly widely smelted and served American colonists. *See* LIMONITE.

Oolitic iron ores are rocks of sand to granule-grade pellets of clay minerals of the chamosite group, variously altered to red hematite. Such ores have been widely mined for iron in the Silurian sedimentary rocks of North America, Jurassic sedimentary rocks of western Europe, the Cretaceous sedimentary rocks of North Africa, and the Oligocene sedimentary rocks of Colombia. *See* OLIGOCENE.

The largest sources of iron are the banded-iron formations of Precambrian age. At the time, the ocean and atmosphere lacked free oxygen, under which conditions much iron remained dissolved in the oceans. Photosynthesis by cyanobacteria provided an aerobic atmosphere, which then aerated the oceans and lead to precipitation of ferric iron oxides. Thick sequences of the banded-iron formation, consisting of alternating layers of hematite and chert, are now mined on all of the inhabited continents. *See* BANDED IRON FORMATION.

Manganese oxides occur in ores that parallel those of the oolitic and the banded iron ores, as well as in crusts and nodules formed on the deep-sea floor. *See* MANGANESE NODULES.

Related to the oolitic iron ores are the pelleted rocks composed of green glauconite. Normally accessory to marine sands, they may come to represent the bulk of the sedimentary rock, as in the Paleocene of New Jersey. Their texture and high potassium content have led to their use as a textural and fertilizing additive to soils. *See* GLAUCONITE; PALEOCENE.

Diamictites and intrusive sediments. Diamictites are doughlike mixtures of particles, ranging from clay to cobbles or boulders. As particles melt out of glaciers, they are kneaded under the moving ice to form deposits of glacial till. Diamictites can also be formed by landslides and in the ejecta blankets of large meteorite impacts. *See* GLACIAL GEOLOGY AND LANDFORMS.

The rapid burial of watery clays, especially if rich in hydrocarbons and thereby light and mobile, can lead to their penetration into the overlying sediments in domes that rise to the surface as mud diapirs. Extreme examples of these form in subduction zones, where sediments scraped off the underthrusting ocean plate become tectonically mixed with muds contributed from the continent to form melanges. These may erupt and flow out to form layers in the stratigraphic succession, layers that in Timor, Indonesia, can reach thicknesses of a kilometer and which represent diamictites containing exotic masses of stratigraphic successions of all sizes, up to hundreds of meters. *See* DIAPIR; SUBDUCTION ZONES.

The most common intrusive sediment bodies are salt domes. Having low density and yielding readily to pressure, halite can penetrate overlying rocks in the form of salt domes in cylindroid or tabular bodies with dimensions from hundreds to thousands of

meters. The updoming of the strata around salt domes has provided many highly productive oil pools. *See* SALT DOME.

Historical content. Each stratum represents an event or episode. Its composition and structure record physical and chemical happenings, and its fossils may reveal much more about the site of deposition. In the succession of strata, each is younger than the one it covers and older than the one that succeeds it, thus forming a link in a chain of history. These chains have recorded the evolution of life, the movements of the crustal plates, the numerous reversals in the Earth's magnetic field, changes in the fractionation of various isotopes in the oceanic reservoir, and climatic oscillations related to Earth's orbital variations. The layers of the stratisphere thus represent a vast archive of Earth history, the domain of stratigraphy. *See* GEOCHEMISTRY; GEOCHRONOMETRY; PALEONTOLOGY; ROCK MAGNETISM; STRATIGRAPHY. A. G. Fischer

Bibliography. H. Blatt, *Sedimentary Petrology*, 2d ed., 1992; S. Boggs, *Principles of Sedimentology & Stratigraphy*, 3d ed., 2000; W. Krumbein and L. L. Sloss, *Stratigraphy and Sedimentation*, 1963; G. Middleton (ed.), *Encyclopedia of Sediments and Sedimentary Rocks*, Kluwer, Dordrecht, 2004; F. J. Pettijohn, *Sedimentary Rocks*, 3d ed., 1983; F. J. Pettijohn and P. E. Potter, *Atlas and Glossary of Sedimentary Structures*, 1964; M. E. Tucker, *Sedimentary Rocks in the Field* (*Geological Field Guide*), 2003.

Sedimentation (industry)

The separation of a dilute suspension of solid particles into a supernatant liquid and a concentrated slurry. If the purpose of the process is to concentrate the solids, it is termed thickening; and if the goal is the removal of the solid particles to produce clear liquid, it is called clarification. Thickening is the common operation for separating fine solids from slurries. Examples are magnesia, alumina red mud, copper middlings and concentrates, china clay (kaolin), coal tailings, phosphate slimes, and pulp-mill and other industrial wastes. Clarification is prominent in the treatment of municipal water supplies.

The driving force for separation is the difference in density between the solid and the liquid. Ordinarily, sedimentation is effected by the force of gravity, and the liquid is water or an aqueous solution. For a given density difference, the solid settling process proceeds more rapidly for larger-sized particles. For fine particles or small density differences, gravity settling may be too slow to be practical; then centrifugal force rather than gravity can be used. Further, when centrifugal force is inadequate, the more positive method of filtration may be employed. All those methods of separating solids and liquids belong to the generic group of mechanical separations. *See* CENTRIFUGATION; CLARIFICATION; FILTRATION; THICKENING.

Settling of spheres through fluids. The basic laws of physics that apply to sedimentation describe the resistance exerted by a liquid to the motion of a solid particle flowing through it. When a particle is released in a fluid in a force field, such as gravity, initially it accelerates. But the motion of the particle generates a frictional resistance which acts to attenuate the acceleration. Quickly, a balance is reached between the force moving the particle and the resisting force; then by Newton's law of motion, the acceleration drops to zero, and the particle proceeds at a constant velocity as long as the active force can act. This steady velocity, called the terminal velocity, prevails in most sedimentation processes.

The law of resistance of a solid particle moving through a fluid is complicated by many variables: the particle diameter and shape, the fluid density and viscosity, the particle velocity (generally its acceleration has a negligible effect), and the proximity of the subject particle to other particles and to equipment surfaces. The ideal situation is considered to be described by a single, smooth, spherical particle, sufficiently far from other particles, and from equipment surfaces, that it is unaffected by their presence. **Figure 1** shows how, for this situation, the terminal velocity u_t is influenced by the acceleration of gravity g, the diameter D_p and density ρ_p of the particle, and the density ρ and viscosity μ of the fluid. These variables are related by dimensionless groups (1) and (2). Figure 1 presents a particular relation between

$$\text{Reynolds number} = \frac{D_p u_t \rho}{\mu} \qquad (1)$$

$$\text{Drag coefficient} = \frac{g D_p (\rho_p - \rho)}{\rho} \qquad (2)$$

these two dimensionless groups which permits the terminal velocity to be calculated directly. The use of such a relation requires that the units used to

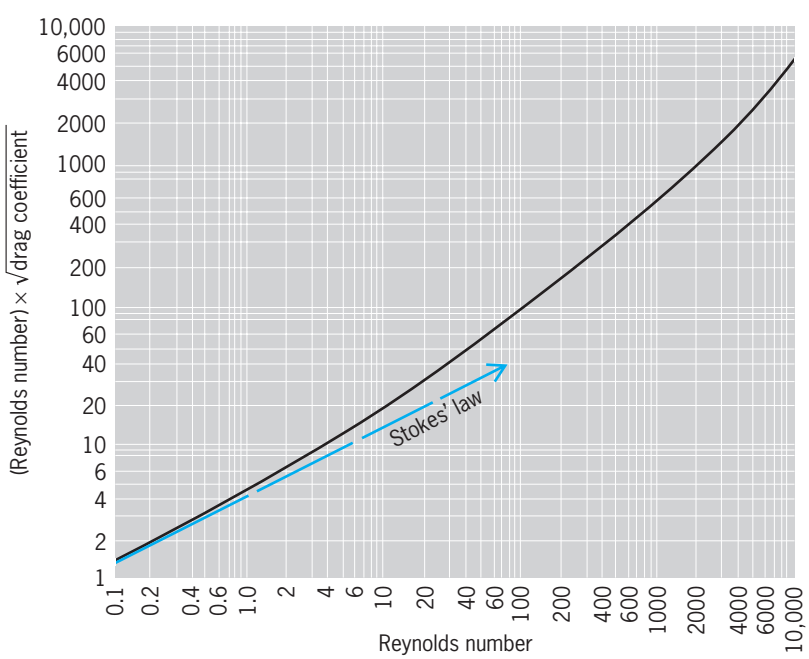

Fig. 1. Relation between the drag coefficient and the Reynolds number for spheres settling in fluids.

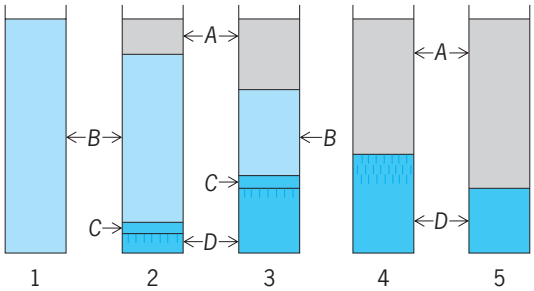

Fig. 2. Stages 1–5 in batch sedimentation of flocculated particles. Layers A–D are explained in the text. (*After W. L. McCabe, J. C. Smith, and P. Harriott, Unit Operations of Chemical Engineering, 4th ed., McGraw-Hill, 1985*)

describe each variable be self-consistent so that the two terms are indeed dimensionless. *See* DIMENSIONLESS GROUPS.

To use Fig. 1, for example, to predict a terminal velocity, the magnitude for the ordinate is calculated from known values of g, D_p, μ, ρ, and ρ_p. The corresponding value of the abscissa is read from the solid curve in Fig. 1, and the terminal velocity is calculated from this abscissa value.

Stokes' law. For small particles, or those moving in a very viscous liquid, the law of settling is governed by Stokes' law, which is Eq. (3). This equation

$$u_t = \frac{gD_p^2(\rho_p - \rho)}{18\mu} \tag{3}$$

is represented by the straight, broken line in Fig. 1; it applies for Reynolds numbers less than 0.3.

Sedimentation rates in practice. Particles too minute to settle at practical rates may form flocs by adding agents such as sodium silicate, alum, lime, and alumina. Because the agglomerated particles act like a single large particle, they settle at a feasible rate and leave a clear liquid behind.

A batch of flocculated pulp passes through several stages during sedimentation. **Figure 2** represents the original homogeneous flocculated pulp ready to settle. In stages 1–5 shown in Fig. 2, layer B is a uniform suspension of the same solid concentration as that of the original pulp, and layer A is clear supernatant liquid. Layer D consists of flocs resting lightly on one another, with liquid filling the voids between the flocs. Layer C is a transition layer, whose solid concentration varies continuously from that in layer B to that in layer D. As settling continues, layers C and B decrease and finally disappear. The end of this process is shown as stage 4 in Fig. 2. Subsequently, a new effect, called compression, begins. During compression, the weight of the deposit collapses the structure of the flocs, and some of the liquid in the flocs of layer D is expelled as small geysers. The thickness of layer D decreases to an equilibrium height called the ultimate height (Fig. 2, stage 5), and the process stops. *See* MECHANICAL SEPARATION TECHNIQUES; REYNOLDS NUMBER. Vincent W. Uhl

Bibliography. W. L. McCabe, J. C. Smith, and P. Harriott, *Unit Operations in Chemical Engineering*, 6th ed., 2000; R. H. Perry and D. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; J. F. Richardson, J. H. Harker, and J. Backhurst, *Chemical Engineering Volume 2 (Coulson & Richardson's Chemical Engineering)*, 5th ed., 2002.

Sedimentology

The study of natural sediments, both lithified (sedimentary rocks) and unlithified, and of the processes by which they are formed. Branches of this discipline have been known as sedimentation or sedimentary petrology, or they have been included in stratigraphy. Sedimentology includes all those processes that give rise to sediment or modify it after deposition: weathering, which breaks up or dissolves preexisting rocks so that sediment may form from them; mechanical transportation; deposition; and diagenesis, which modifies sediment after deposition and burial within a sedimentary basin and converts it into sedimentary rock. Sediments deposited by mechanical processes (gravels, sands, muds) are known as clastic sediments, and those deposited predominantly by chemical or biological processes (limestones, dolomites, rock salt, chert) are known as chemical sediments. **Figure 1** shows a diagram of the processes that are involved in the sedimentary cycle. *See* STRATIGRAPHY; WEATHERING PROCESSES.

Sedimentary cycle. The raw materials of sedimentation are the products of weathering of previously formed igneous, metamorphic, or sedimentary rocks. In the present geological era, 66% of the continents and almost all of the ocean basins are covered by sedimentary rocks. Therefore, most of the sediment now forming has been derived by recycling previously formed sediment. Identification of the oldest rocks in the Earth's crust, formed more than 3×10^9 years ago, has shown that this process has been going on at least since then. Old sedimentary rocks tend to be eroded away or converted into metamorphic rocks, so that very ancient sedimentary rocks are seen at only a few places on Earth. The area (or volume) of sedimentary rocks declines exponentially with age, and the half-life of a sedimentary rock is about 2×10^8 years. *See* EARTH CRUST; IGNEOUS ROCKS; METAMORPHIC ROCKS.

Major controls. The major controls on the sedimentary cycle are tectonics, climate, worldwide (eustatic) changes in sea level, the evolution of environments with geological time, and the effect of rare events.

Tectonics. These are the large-scale motions (both horizontal and vertical) of the Earth's crust. Tectonics are driven by forces within the interior of the Earth but have a large effect on sedimentation. These crustal movements largely determine which areas of the Earth's crust undergo uplift and erosion, thus acting as sources of sediment, and which areas undergo prolonged subsidence, thus acting as sedimentary basins. Rates of uplift may be very high (over 10 m or 33 ft per 1000 years) locally, but probably such rates prevail only for short periods of time. Over

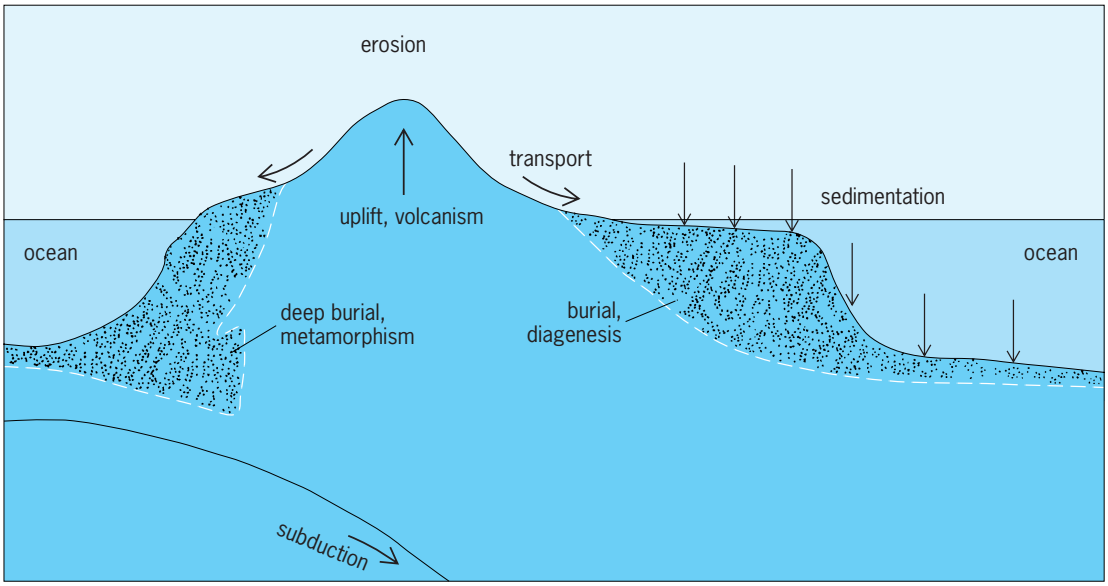

Fig. 1. Diagram showing processes that constitute the sedimentary cycle. Heavy arrows indicate mass fluxes of material. Light arrows indicate where deposition is taking place.

millions of years, uplift even in mountainous regions is about 1 m (3.3 ft) per 1000 years, and it is closely balanced by rates of erosion. Rates of erosion, estimated from measured rates of sediment transport in rivers and from various other techniques, range from a few meters per 1000 years in mountainous areas to a few millimeters per 1000 years averaged over entire continents. Though these rates are very slow by human standards, over periods of millions of years they produce immense quantities of sediment; all the continents would soon be reduced to sea level if it were not for the renewed uplift produced by tectonics. It is estimated that the total volume of all sediment and sedimentary rocks is over 6.5×10^{10} km^3 (1.55×10^8 mi^3), a volume that would cover the Earth to a depth of 1.3 km (0.82 mi) if spread evenly over its surface. Recent studies of the Moon and planets show that the Earth is very active tectonically compared with other bodies of similar size in the solar system. *See* BASIN; PLATE TECTONICS.

Climate. This plays a secondary but important role in controlling the rate of weathering and sediment production. The more humid the climate, the higher these rates are; evidence exists, however, that in some continental climates (for example, in the southwestern United States) the rate of sediment production reaches a maximum in semiarid regions (250–750 mm or 30–90 in. of rainfall per year). A combination of hot, humid climate and low relief permits extensive chemical weathering, so that a larger percentage of source rocks goes into solution, and the clastic sediment produced consists mainly of those minerals that are chemically inert (such as quartz) or that are produced by weathering itself (clays). Cold climates and high relief favor physical over chemical processes. Source rocks disintegrate to form sediment that includes little clay and many minerals that are chemically unstable in the Earth's surface environments. *See* CLIMATOLOGY.

Sea level. Tectonics and climate together control the relative level of the sea. In cold periods, water is stored as ice at the poles, which can produce a worldwide (eustatic) lowering of sea level by more than 100 m (330 ft). Changes in the rate of sea-floor spreading can produce large alterations in sea level by changing the average depth of the ocean basins. Widespread, but not worldwide, uplift or subsidence can also result in emergence or flooding of substantial areas of the continents. Changes in sea level, whether local or worldwide, strongly influence sedimentation in shallow seas and along coastlines; sea-level changes also affect sedimentation in rivers by changing the base level below which a stream cannot erode its bed.

Evolution of environments. One of the major conclusions from the study of ancient sediments has been that the general nature and rates of sedimentation have been essentially unchanged during the last billion years of geological history. However, this conclusion, uniformitarianism, must be qualified to take into account progressive changes in the Earth's environment through geological time, and the operation of rare but locally or even globally important catastrophic events. The most important progressive changes have been in tectonics and atmospheric chemistry early in Precambrian times, and in the nature of life on the Earth, particularly since the beginning of the Cambrian. *See* CAMBRIAN.

Rare events. Throughout geological time, events that are rare by human standards but common on a geological time scale, such as earthquakes, volcanic eruptions, and storms, produced widespread sediment deposits. There is increasing evidence for a few truly rare but significant events, such as the rapid drying up of large seas (parts of the Mediterranean) and collisions between the Earth and large meteoric or cometary bodies (bolides).

Sediment movement and deposition. Sediment is moved either by gravity acting on the sediment particles or by the motions of fluids (air, water, flowing ice), which are themselves produced by gravity. Deposition takes place when the rate of sediment movement decreases in the direction of sediment movement; deposition may be so abrupt that an entire moving mass of sediment and fluid comes to a halt (mass deposition, for example, by a debris flow), or so slow that the moving fluid (which may contain only a few parts per thousand of sediment) leaves only a few grains of sediment behind.

Fluids transport sediment mainly by traction (sliding, rolling, bouncing on the bed) and suspension caused by turbulence. Though the irregular, turbulent motion of fluids such as air and water are the most important way that sediment is held up within a mass of flowing fluid, other mechanisms such as buoyancy and impacts between the grains may also operate. *See* TURBULENT FLOW.

A fundamental dynamic property of sediment grains is their settling velocity. This property is the constant velocity of fall that a grain attains in a fluid at rest, when the force of gravity is balanced by the forces of buoyancy and fluid resistance. The settling velocity depends on the density and viscosity of the fluid, as well as on the size, shape, and density of the grains. As measured under standard conditions (in water at 20°C or 68°F), it can be used directly to characterize sediment or be converted into an equivalent diameter (the diameter of a quartz sphere with the same settling velocity as the measured value). Settling velocities in water vary from less than 0.1 mm (0.004 in.) per second for silt, to 5–10 cm (2–4 in.) per second for average sand, to meters per second for large boulders.

Transport and deposition by flowing water. The different ways in which sediment is transported by water flows can be illustrated by describing the changes that are observed in a laboratory channel (flume) as the discharge of water is gradually increased over a flat bed of sand. The flow strength just necessary to start movement of grains on the bed defines the competence of the flow for a particular type (size, shape, density) of sediment. The first grain motion is by sliding or rolling of the grains in contact with the bed, and this is soon followed by a leaping motion known as saltation. In saltation, grains rise up at a steep angle from the bed, are caught by the flow, and drop back to the bed; they follow a so-called ballistic trajectory. Sliding, rolling, and saltation generally are grouped together as traction. Sediment moved in this way is described as bed load. At low flow strengths, the grains do not rise far above the bed (only a few grain diameters in water) and they move much more slowly than the fluid. Rates of sediment movement (as measured by the sediment discharge, or mass rate of transport of sediment past a particular cross section of the channel) depend on a high power of flow strength. Bed load is supported by direct contact with the bed (and by buoyancy). Its motion responds to the fluctuating shear stress on the bed produced by turbulence in the flow rate,

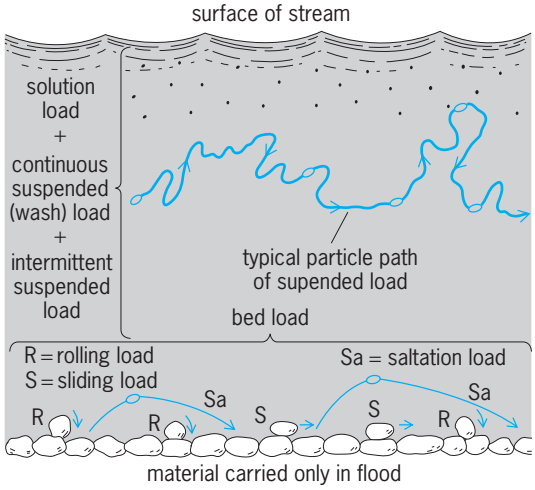

Fig. 2. Diagram of types of load carried by a stream. (*After R. M. Garrels, A Textbook of Geology, Harper and Row, 1951*)

but it is otherwise not much affected by turbulent fluid motions (**Fig. 2**).

As flow strength is increased, however, the upward velocity components of turbulent eddies become strong enough to overcome the settling velocity of the sediment, and the grains are taken into suspension. A part of the sediment load is still moved by traction, but a larger and larger part is moved by suspension (suspended load). Suspended sediment, at first concentrated near the bed, becomes more uniformly distributed in the flow as the flow strength increases. The total rate at which sediment can be moved (capacity) depends upon the availability of erodible sediment in the banks or bed and upon the sediment properties and flow strength. Sediment discharge continues to increase with flow strength, but not as rapidly as in the early stages of sediment transport. There is almost no limit to the amount of fine sediment that can be carried in suspension: concentrations as high as 30% by weight have been recorded. As concentration increases, so does the likelihood of collisions between the coarser grains. It has been suggested that such collisions become an important support mechanism (dispersive pressure) for sand and gravel at volume concentrations higher than about 10%, and such concentrations are certainly present close to the bed at high flow strengths.

As the rates of sediment transport increase, the nature of the bed also changes. Even if the bed is flat at first, it develops asymmetrical waves or bedforms migrating downstream (**Fig. 3**). Small current ripples (10–30 cm or 4–12 in. long) form soon after sediment movement begins, but at higher flow strengths they are replaced by large dunes on the bed; in deep flows, these may reach heights of several meters and lengths of tens of meters. At very high flow strengths, however, the dunes are washed out and the bed becomes almost flat again. If the flow becomes supercritical, nearly symmetrical bedforms (antidunes) migrating upstream may be

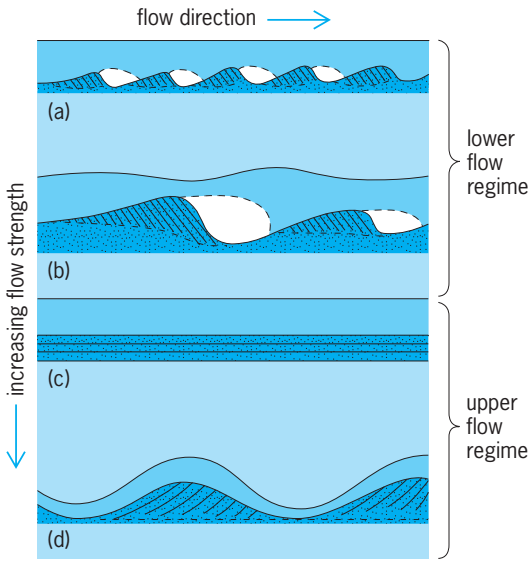

Fig. 3. Bedforms *a–d* produced by flow over a sand bed. (*After H. Blatt, G. V. Middleton, and R. Murray, Origin of Sedimentary Rocks, 2d ed., Prentice-Hall, 1980*)

formed. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; FROUDE NUMBER; STREAM TRANSPORT AND DEPOSITION.

Transport by waves. Water waves produce an oscillating motion of water close to the bed. Therefore, movement of sediment by waves differs in detail from that produced by unidirectional flows. Net sediment movement generally still results, however, because of the asymmetrical nature of the oscillating flow close to the bed. Waves produce stages of sediment transport similar to those resulting from unidirectional flows. Suspension replaces traction at high levels of wave activity. Bedforms show a transition from small to larger ripples, and a plane bed reappears at the highest wave intensities. Breaking of waves greatly enhances the ability of waves to carry sediment into suspension. After breaking, waves form a bore (swash) that washes sediment up the beach, and down again in the backwash. The approach of waves is generally oblique to the shore and produces a movement of sand along the beach that is known as longshore drift. *See* NEARSHORE PROCESSES.

Transport by wind. Sediment transport by wind is similar to transport by water, but different forms of transport predominate because of the low density of air compared with water, and the great thickness and high speed attained by moving air masses. Whereas sand is easily transported in suspension by water flowing at speeds of the order of 1 m (3.3 ft) per second, it can be suspended only rarely by winds. Most sand is moved by wind in saltation, which is a far more effective mechanism of sediment transport in air than in water. Saltating sand grains rise to heights of several centimeters and return to the bed with enough momentum to drive forward other grains, a process known as surface creep. Ripples produced by saltating grains are smaller and more regular than those produced by flow of water, but dunes develop to much larger sizes and show a wide variety of forms because of the thickness of the wind boundary layer and the variability of wind directions. *See* AIR MASS.

Only clay and silt are easily carried into suspension by the wind, but once suspended, these grains may be transported for hundreds of kilometers. Fine sediment (ash) is produced by explosive volcanic eruptions and carried high into the upper atmosphere, where it is transported great distances by the wind. Deposits of silt (loess) several meters thick were formed at the end of the last ice age by winds that blew sediment derived from rivers that drained the ice sheets. *See* DUNE; DUST STORM; LOESS; WIND.

Transport by ice. Glaciers and ice sheets incorporate sediment that falls onto the ice surface or is eroded by grinding or plucking from material beneath the ice. Such material is transported en masse to the point where the glacier melts. Sediment deposited by melting directly out of ice is almost completely unsorted and is known as till. Some ice sheets melt directly into lakes or the sea, producing debris that is modified as it settles to the bottom. Much of the sediment that is derived from glaciers is reworked by mass flow (flow till) or by meltwater (outwash). Some of the sediment eroded by glaciers consists of finely ground material known as rock flour. The abundance of this material in lakes fed by glaciers is at least partly responsible for their intense green color. *See* GLACIOLOGY; TILL.

Sediment gravity flows. These are flows whose motion results from gravity acting on the sediment grains rather than on the fluid in the flow (indeed, no fluid is necessary for certain types of gravity flow, and so these flows can operate in fluid-free environments, such as the Moon). Four classes of sediment gravity flow have been recognized, differentiated by the mechanisms that support sediment above the bed. If the mechanism is grain impacts, the flow is a grain flow; if it is turbulence, the flow is a turbidity current; if it is the upward escape of interstitial fluid, the flow is a liquefied flow; and if it is the strength of the matrix between the grains (for example, a mud matrix), the flow is a plastic debris flow. Such flows are generally short-lived, but they may be very large. Some recorded turbidity currents have transported as much as 100 km^3 (24 mi^3) of sediment and have flowed for several hundred kilometers over the ocean floor. *See* TURBIDITY CURRENT.

Modification by transport. Sediment is modified during transport by chemical action (continuation of the weathering process begun in the source) and by physical breakage, abrasion, and sorting. Physical impact is most effective on the larger grains, and it has little effect on grains small enough to be transported in suspension. Boulders and pebbles are rapidly rounded by abrasion of the corners and edges of the grain, and prolonged transport in rivers may lead to a nearly spherical shape. The more friable minerals found in sand (for example, feldspar) may be broken, but the main mineral, quartz, is very resistant to abrasion. Nevertheless, studies using the scanning electron microscope have shown evidence

of rounding and submicroscopic impact marks, and reduction in size by flaking of small chips from the surface. Transport by wind is far more effective than transport by water in rounding quartz grains, and reworking by waves on beaches is more effective than transport down rivers—even when the transport distance is hundreds of kilometers down major rivers.

Sediment weathered from an igneous or metamorphic source rock, or produced from it by mechanical breakage, tends to have a wide range of grain sizes that follow a statistical distribution known as Rosin's distribution. Though a similar distribution may be detected in some cases after several kilometers of river transport, generally it is rapidly modified by hydraulic sorting, which tends to segregate the variously sized fractions by their different mechanisms of transport. In a river, for example, the coarsest grains are confined to the deepest part of the channel; the sand is confined to bars within the channel or to the natural levees; and either the mud is transported in suspension rapidly through the system to the sea, or it is deposited during floods in overbank deposits on the floodplain. Because each of these deposits consists mainly on only one fraction of the original source material, they are better sorted and the size distribution is changed. Most naturally occurring size distributions appear to be composed of mixtures of more than one (lognormal) distribution, resulting from different sources or mechanisms of deposition. *See* FLUVIAL SEDIMENTS.

Flocculation. Gravel and sand grains are generally chemically inert and noncohesive. This is not true of the finest, clay-sized particles. Colloid-sized particles (finer than 0.2 micrometer) generally have electrically charged surfaces, even if composed of inert minerals. Many clay minerals, such as montmorillonite, readily react with dissolved ions. The surface charge depends in part on the nature and concentration of ions in the surrounding solution. Though the grains of clay minerals are very small (less than 1 μm), electrochemical effects may cause them to aggregate into large flocs whose settling velocity (though still relatively low, because of the low bulk density of the floc) may be many times higher than that of the individual grains. Flocculation is very common as clay is delivered by rivers into the sea; thus it is an important process in estuaries and deltas. *See* CLAY MINERALS; DELTA; ESTUARINE OCEANOGRAPHY; GRAVEL; SAND.

Chemical sedimentation. Chemical weathering dissolves rock materials and delivers ions in solution to lakes and the ocean. The concentrations of ions in river and ocean water are quite different, showing that some ions must be removed by sedimentation. Comparison of the modern rate of delivery of ions to the ocean, with their concentration in the oceans, shows that some are removed very rapidly (residence times of only a few thousand years) whereas others, such as chlorine and sodium, are removed very slowly (residence times of hundreds of millions of years). The main mechanisms by which ions are removed from solution are biochemical precipitation, evaporation, adsorption or exchange with clays and other rock materials (including glass in fine volcanic ash), and reaction with hot rock in hydrothermal systems associated with mid-oceanic ridges. The most important chemical sediments are calcareous (limestones and dolomites), followed by siliceous chert deposits and evaporites, but there are a very large variety of chemical sediments whose origins are not yet fully understood. Chemical dissolution and precipitation are not restricted to weathering and primary sedimentation from lake water or seawater, but are equally important in diagenesis, that is, in transforming sediment after deposition into sedimentary rock. *See* MARINE SEDIMENTS; MID-OCEANIC RIDGE.

Carbonate sediments. Organisms began to secrete skeletons composed of calcium carbonate during the Cambrian Period, and since then most carbonate sediment has been precipitated biochemically. Though the warm surface waters of the oceans are known to be supersaturated with respect to calcite (the thermodynamically stable form of calcium carbonate in most natural waters), direct chemical precipitation of calcite is known in only a few environments (for example, limestone caves). The common carbonate minerals in modern sediments are biochemically precipitated aragonite and calcite. Aragonite is metastable under surface conditions and is generally dissolved away within a few thousand years (rarely millions) to be replaced by diagenetic calcite or dolomite. *See* LIMESTONE.

The best-known biochemical accumulations of carbonate at present are the tropical coral-algal reefs, but a wide variety of organisms precipitate carbonate from seawater now or they have done so in the past. Besides reef organisms, pelecypods, algae, and foraminifera are important in cooler marine waters; and in the surface layers of the oceans, planktonic foraminifera, pteropods, and algae (coccolithophorids) are important today, and were so at other periods of the Cenozoic and Mesozoic. Cretaceous chalks, for example, are composed largely of microscopic coccoliths and foraminifera, not very different from modern deep-sea calcareous oozes (though most ancient chalks were deposited in water only a few hundred meters deep). In other geological periods, reefs were built by different types of corals, by specialized pelecypods (rudists), by algae, and by now-extinct groups such as stromatoporoids. *See* CENOZOIC; CHALK; CRETACEOUS; MESOZOIC; REEF.

The most important differences between ancient carbonate rocks and modern carbonate sediments, however, are due to diagenesis. Modern carbonate sediments are composed mainly of aragonite and a form of calcite containing several percent of magnesium. Ancient carbonates consist largely of calcite low in magnesium and of dolomite, a calcium-magnesium carbonate only rarely present in modern sediments. The mineralogical changes necessary to produce typical ancient carbonates take place rapidly, particularly under the action of meteoric water. Even the formation of dolomite, which generally requires the wholesale dissolution of aragonite and calcite and the exchange of large quantities of magnesium from water passing through the porous

rock, has been a common and frequently rapid process throughout geological history. *See* ARAGONITE; AUTHIGENIC MINERALS; CALCITE; CARBONATE MINERALS; DIAGENESIS; DOLOMITE; MARL.

Siliceous sediment. Sediments composed largely of fine-grained quartz (chert) are common in the geological record, extending back to the earliest Precambrian. Since the beginning of the Cambrian, most cherts have been formed by biochemical precipitation by organisms that secrete an opaline skeleton, such as radiolarians, diatoms, and some sponges, followed by diagenetic dissolution and reprecipitation as microcrystalline quartz and chalcedony. *See* CHALCEDONY; CHERT.

Evaporites. Evaporation of seawater in the laboratory gives rise to a sequence of different precipitates as evaporation proceeds. Small amounts of calcareous sediments are first precipitated, followed by gypsum, anhydrite, halite, and finally by complex salts containing sodium, potassium, magnesium, sulfate ions, and chloride ions. A close match with this experimentally determined sequence is known to exist in several ancient deposits, though the relative abundances of different minerals vary considerably. This indicates that these deposits were formed by evaporation of seawater, and that the composition of the ancient seawater was not very different from that of modern seawater. Evaporites can presently be observed forming in several shallow-water environments (lakes and marine lagoons), though none is comparable in magnitude to the evaporite basins inferred for some periods in the past. Formation of large evaporite deposits was probably a rare event, but once conditions were appropriate, large volumes could be formed in only a few tens of thousands of years. *See* SALINE EVAPORITES; SEAWATER.

Other chemical sediments. Other chemical sediments are rich in iron, phosphorus, manganese, and carbon. The carbonaceous sediments include marine muds rich in organic matter, which are important as the source rocks of petroleum, and nonmarine peats, some of which give rise to commercial coals. Many sedimentary rocks of unusual character are economically important; they serve as ores of iron, copper, manganese, phosphate, gold, diamonds, and so forth. Equally important economically are more common sedimentary rocks, which serve as sources of building stone, crushed rock for construction, limestone for smelting, pure quartz sands, and so forth. *See* COAL; ORE AND MINERAL DEPOSITS; PETROLEUM.

Biological effects. Many so-called chemical sediments are actually produced by biochemical action. Much is then reworked by waves and currents, so that the chemical sediment shows clastic textures and consists of grains rounded and sorted by transport. Depositional and diagenetic processes, however, are often strongly affected by organic action, no matter what the origin of the sediment. Plants in both terrestrial and marine environments tend to trap sediment, enhancing deposition and slowing erosion. Organisms burrow through the sediment, producing distinctive structures (bioturbation) and modifying or destroying original clastic structures and textures. Bacteria promote decay of organic particles, changing the chemical environment within the sediment, and producing gases such as carbon dioxide and methane. At deeper levels of burial within a sedimentary basin, the residual organic matter in the sediments (kerogen or sapropel) is broken down by heat to produce liquid and gaseous hydrocarbons and carbon dioxide, which may migrate many tens of kilometers upward along permeable strata within the basin. *See* KEROGEN; SAPROPEL.

Sedimentary environments and facies. Sedimentary rocks preserve the main direct evidence about the nature of the surface environments of the ancient Earth and the way they have changed through geological time. Thus, besides trying to understand the basic principles of sedimentation, sedimentologists have studied modern and ancient sediments as records of ancient environments. For this purpose, fossils and primary sedimentary structures are the best guide. These structures are those formed at the time of deposition, as opposed to those formed after deposition by diagenesis, or by deformation. In describing sequences of sedimentary rocks in the field (stratigraphic sections), sedimentologists recognize compositional, structural, and organic aspects of rocks that can be used to distinguish one unit of rocks from another. Such units are known as sedimentary facies, and they can generally be interpreted as having formed in different environments of deposition. Though there are a large number of different sedimentary environments, they can be classified in a number of general classes, and their characteristic facies are known from studies of modern environments. Furthermore, the transition from one environment to another is not a random process, because only those environments that originally existed side by side can succeed each other without a break (Walther's law). Thus, it is possible to produce plausible reconstructions of environments that existed many millions of years ago. Like all historical reconstructions based on fragmentary evidence, these depend upon the imaginative skill of the investigator, as well as upon reliable deductions from data. *See* FACIES (GEOLOGY); TRACE FOSSILS.

Basin analysis. This describes studies of the larger-scale aspects of sediment accumulation, such as the different types of sedimentary basins, and how they form, and the types of source areas and how they are linked to particular types of basins. Clastic sediments derived from preexisting rocks (terrigenous sediments) have textures and mineralogy that preserve clues about the nature of the source from which they were derived (their provenance). Sedimentary structures and textures can be used to indicate the direction of sediment transport at the place of deposition (paleocurrents). By carefully dating the deposits, or at least determining which deposits are the same age (stratigraphic correlation), by studying their provenance and paleocurrents, and by reconstructing their depositional environments, it is possible to reconstruct how a sedimentary basin developed and where the sediment came from. In

addition, the long-term rates of deposition can be determined, and by studying mineralogical changes produced by diagenesis it is possible to reconstruct the sequence of pressures and temperatures that have affected the sediments in the basin. Tectonic theories suggest models that simulate the patterns of deposition and diagenesis within sedimentary basins. *See* SEDIMENTARY ROCKS. G. V. Middleton

Bibliography. P. A. Allen, *Earth Surface Processes*, 1997; H. Blatt, G. V. Middleton, and R. Murray, *Origin of Sedimentary Rocks*, 2d ed., 1980; M. R. Leeder, *Sedimentology*: *Process and Product*, 1982; M. R. Leeder, *Sedimentology and Sedimentary Basins*: *From Turbulence to Tectonics*, 1999; R. K. Matthews, *Dynamic Stratigraphy*: *An Introduction to Sedimentation and Stratigraphy*, 2d ed., 1984; A. D. Miall, *Principles of Sedimentary Basin Analysis*, 2d ed., 1990.

Seebeck effect

The generation of a temperature-dependent electromotive force (emf) at the junction of two dissimilar metals. This phenomenon provides the physical basis for the thermocouple. In 1821, T. J. Seebeck discovered that near a closed circuit composed of two linear conductors of two different metals a magnetic needle would be deflected if, and only if, the two junctions were at different temperatures, and that if the temperatures of the two junctions were reversed the direction of deflection would also be reversed. He investigated 35 different metals and arranged them in a series such that at a hot junction, current flows from a metal earlier in the series to a later one. *See* ELECTROMOTIVE FORCE (EMF).

A thermocouple consists of a pair of wires of dissimilar metals, joined at the ends. One junction is kept at an accurately known cold temperature, usually that of melting ice, and the other is used for the measurement of an unknown temperature, by measuring the emf generated as a result of the Seebeck effect. *See* THERMOCOUPLE; THERMOELECTRICITY. A. Earle Bailey

Seed

A fertilized ovule containing an embryo which forms a new plant upon germination. Seed-bearing characterizes the higher plants—the gymnosperms (conifers and allies) and the angiosperms (flowering plants). Gymnosperm (naked) seeds arise on the surface of a structure, as on a seed scale of a pine cone. Angiosperm (covered) seeds develop within a fruit, as the peas in a pod. *See* FLOWER; FRUIT.

Structure. One or two tissue envelopes, or integuments, form the seed coat which encloses the seed except for a tiny pore, the micropyle (**Fig. 1**). The micropyle is near the funiculus (seed stalk) in angiosperm seeds. The hilum is the scar left when the seed is detached from the funiculus. Some seeds have a raphe, a ridge near the hilum opposite the micropyle, and a bulbous strophiole. Others such as nutmeg possess arils, outgrowths of the funiculus, or a fleshy caruncle developed from the seed coat near the hilum, as in the castor bean. The fleshy, edible aril of the Philippine kamanchile completely encloses the seed to form a fruitlike structure. The embryo consists of an axis and attached cotyledons (seed leaves). The part of the axis above the cotyledons is the epicotyl (plumule); that below, the hypocotyl, the lower end of which bears a more or less developed primordium of the root (radicle). The epicotyl, essentially a terminal bud, possesses an apical meristem (growing point) and, sometimes, leaf primordia. The seedling stem develops from the epicotyl. An apical meristem of the radicle produces the primary root of the seedling, and transition between root and stem occurs in the hypocotyl. *See* APICAL MERISTEM; ROOT (BOTANY); STEM.

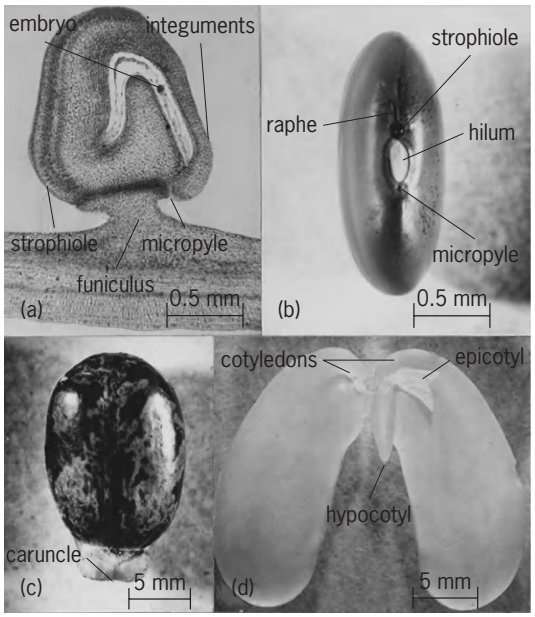

Fig. 1. Seed structures. (*a*) Median longitudinal section of pea ovule shortly after fertilization, showing attachment to pod tissues. (*b*) Mature kidney bean. (*c*) Mature castor bean. (*d*) Opened embryo of mature kidney bean.

Two to many cotyledons occur in different gymnosperms. The angiosperms are divided into two major groups according to number of cotyledons: the monocotyledons, including orchids, lilies, grasses, and sedges; and the dicotyledons, such as beans, roses, and sunflowers. Mature gymnosperm seeds contain an endosperm (albumen or nutritive tissue) which surrounds the embryo. In some mature dicotyledon seeds the endosperm persists, the cotyledons are flat and leaflike, and the epicotyl is simply an apical meristem (**Fig. 2**). In other seeds, such as the bean, the growing embryo absorbs the endosperm, and food reserve for germination is stored in fleshy cotyledons. The endosperm persists in common monocotyledons, for example, corn and wheat; and the cotyledon, known as the scutellum, functions as an absorbing organ during germination (**Fig. 3**). Grain embryos also possess a coleoptile and a coleorhiza sheathing the epicotyl and the radicle,

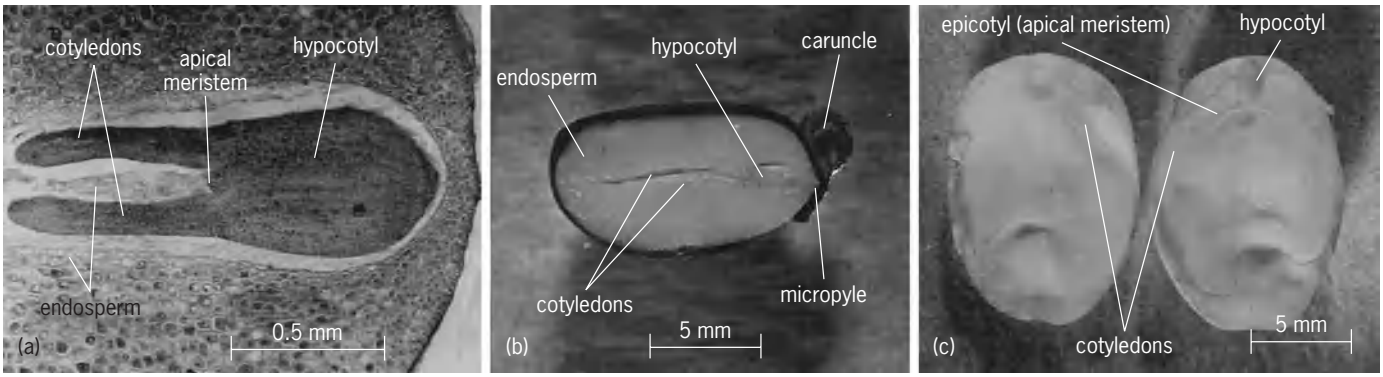

Fig. 2. Dicotyledon seeds. (*a*) Median longitudinal section of embryo of silk tassel bush (*Garrya elliptica*) embedded in endosperm which has been removed from mature seed. (*b*) Castor bean cut longitudinally. (*c*) Castor bean removed from seed coat and split longitudinally between cotyledons.

respectively. The apical meristems of lateral seed roots also may be differentiated in the embryonic axis near the scutellum of some grains.

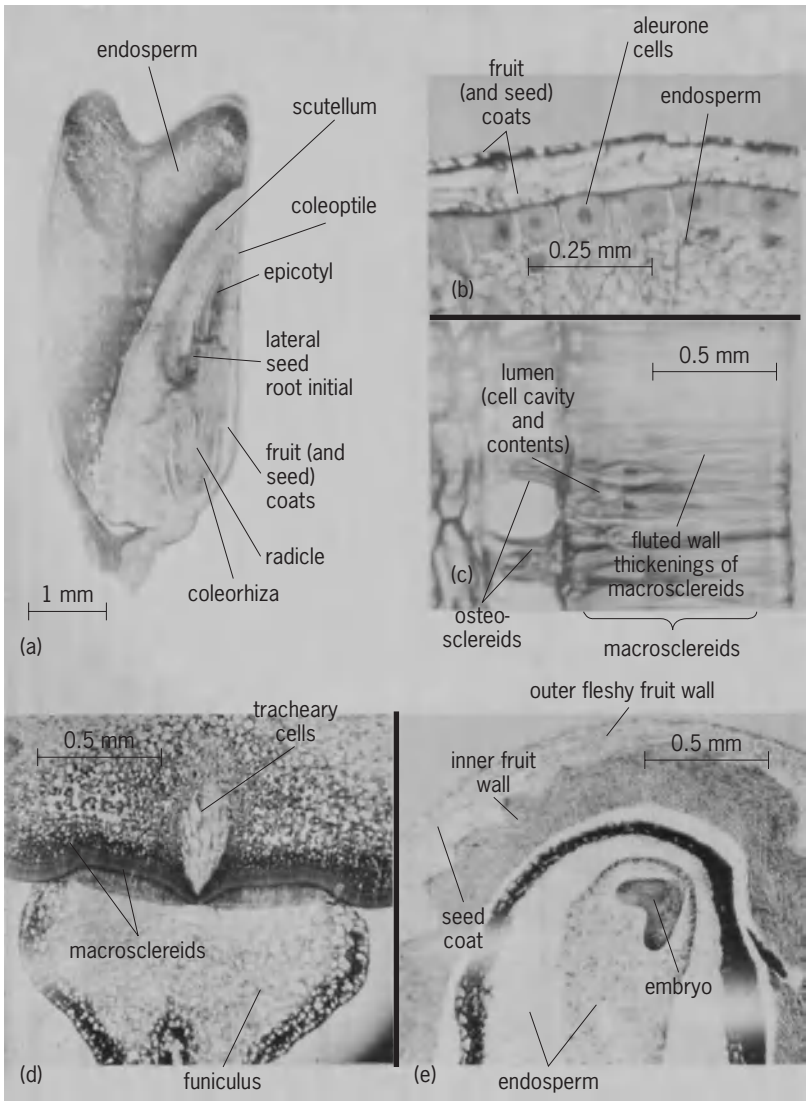

Fig. 3. Monocot seeds. (*a*) Section of corn kernel (*from M. J. Wolf et al., Cereal Chem., 29:321–382, 1952*). (*b*) Section of outer cell layers of wheat grain. (*c*) Section of outer cells of mature pea seed coat. (*d*) Section of hilum area of nearly mature pea. (*e*) Section through portion of young raspberry drupelet (segment of berry).

Monocotyledon and dicotyledon seeds also differ in seed coat structure. In grains, or caryopsis fruits, the mature fruit wall and seed coat may be fused and the outermost endosperm cells form an aleurone layer, rich in proteins. Flour brans consist of fragments of aleurone and fruit-seed coats. Most of the thiamine and riboflavin of grains occur in these tissues and near the scutellum.

Cellular structure of the seed coat varies in the dicotyledons. An outer, water-impervious layer of wax or cuticle is usually present. The outermost cells in bean and pea seed coats, known as macrosclereids (Malpighian cells), have fluted, cellulosic wall thickenings; underlying them is a layer of osteosclereids, short bone-shaped cells. Two layers of macrosclereids occur at the hilum in peas and beans. Underlying them is a group of tracheary cells with reticulate (netted) wall thickenings, and a surrounding spongy tissue of branched cells and large air spaces. This structure appears to be well suited for water absorption during germination.

Many so-called seeds consist of hardened parts of the fruit enclosing the true seed which has a thin, papery seed coat. Among these are the achenes, as in the sunflower, dandelion, and strawberry, and the pits of stone fruits such as the cherry, peach, and raspberry. Many common nuts also have this structure.

Dispersal. Mechanisms for seed dispersal include parts of both fruit and seed. Some dry fruits have membranous air sacs that aid in seed dispersal by water. In cockleburrs and beggarticks the seeds are enclosed in spiny, barbed fruits that are readily carried in the fur of animals. Seed dispersal by wind is aided by hair tufts of dandelion achenes, wings of elm and maple fruits and of conifer seeds, and by seed-coat hairs of willow, cottonwood, and milkweed seeds. Cotton fibers consist of greatly elongated epidermal hairs of the seed coat as do also the kapok fibers of the silk-cotton tree (*Ceiba pentandra*). *See* POPULATION DISPERSAL.

Economic importance. Propagation of plants by seed and technological use of seed and seed products are among the most important activities of modern society. Specializations of seed structure and composition provide rich sources for industrial

exploitation apart from direct use as food. Common products include starches and glutens from grains, hemicelluloses from guar and locust beans, and proteins and oils from soybeans and cotton seed. Drugs, enzymes, vitamins, spices, and condiments are obtained from embryos, endosperms, and entire seeds, often including the fruit coat. Most of the oils of palm, olive, and pine seeds are in the endosperm. Safflower seed oil is obtained mainly from the embryo, whereas both the seed coat and embryo of cotton seed are rich in oils. *See* FOOD; PLANT ANATOMY; REPRODUCTION (PLANT).

Roger M. Reeve

Physiology. Physical and biochemical processes of seed growth and germination are controlled by genetic and environmental factors. Conditions of light, temperature, moisture, and oxygen affect the timing and ability of a seed to mature and germinate. Seed development (embryogenesis) is concerned with the synthesis and storage of carbohydrate, protein, and oil to supply nutrients to the germinating seedling prior to soil emergence. Seed development occurs in several stages: rapid cell division, seed fill, and desiccation (**Fig. 4**). The timing of each stage is species-specific and environmentally influenced. In mature monocot seeds (for example, corn), starch in the endosperm is the main storage product, while small quantities of oil and protein are stored primarily in the embryo and aleurone (testa/pericarp), respectively. Major storage reserves in mature dicot seed cotyledons may be starch as in pea, or protein and oil as in soybean. In the case of soybean, seeds develop over a period of about 60 days. Protein and oil accumulation is confined to a period of carbon and nitrogen import from the vegetative plant, resulting in rapid dry-weight increases (seed fill). During late seed fill and desiccation, there is an accumulation of small amounts of the raffinose saccharides, raffinose and stachyose. Synthesis of raffinose and stachyose in soybean is correlated with increases in galactinol synthase activity, the enzyme that catalyzes the committing step to this biosynthetic pathway (**Fig. 5**). Total starch and soluble sugars account for only approximately 12% dry weight of a dry soybean seed. With desiccation, a viable seed will germinate if provided with suitable environmental conditions for germination and for overcoming any dormancy constraints.

Dormancy. Seed dormancy is the inability of a living seed to germinate under favorable conditions of temperature, moisture, and oxygen. Dormancy does not occur in all seeds, but typically occurs in plant species from temperate and colder habitats such as maple, pear, hazel, and pine. This process allows for a delay in seed germination until environmental conditions are adequate for seedling survival. At least three types of seed dormancy are recognized: primary, secondary (induced), and enforced. Primary dormancy occurs during seed maturation, and the seed does not germinate readily upon being shed. Internal factors such as chemical inhibitors, immaturity of the embryo, and impermeability of hard seed coat tissues to water and oxygen may be important in causing primary dormancy. Secondary and enforced dormancy occur after the seed is shed and may be caused by adverse environmental factors such as high or low temperature, absence of oxygen or light, low soil moisture, and presence of chemical inhibitors. Seeds with secondary dormancy will not germinate spontaneously when environmental conditions improve, and need additional environmental stimuli. Seeds with enforced dormancy germinate readily upon removal of the environmental limitation.

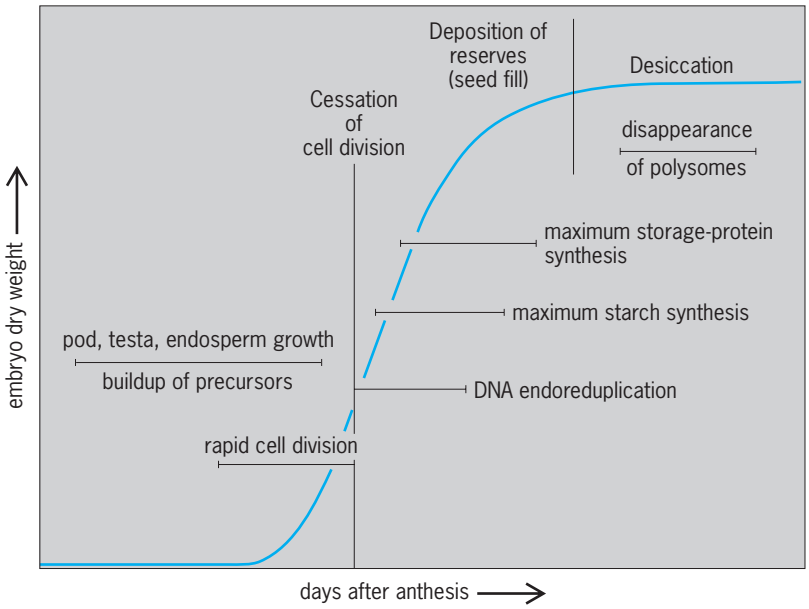

Fig. 4. Process of legume seed development (embryogenesis). (*After L. S. Dure, Seed formation, Annu. Rev. Plant Physiol., 26:259–278, 1975*)

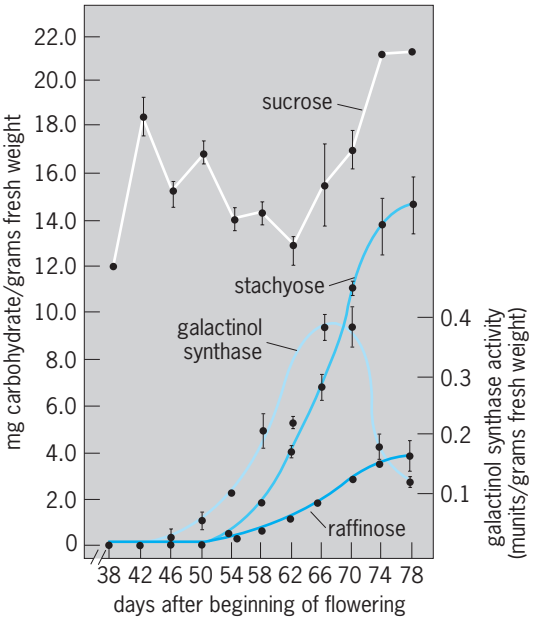

Fig. 5. Changes in levels of the sugars sucrose, raffinose, and stachyose, and concomitant increases in galactinol synthase activity in maturing soybean seeds. The term munits/grams fresh weight represents the amount of enzyme (galactinol synthase) necessary to form 1 nanomole of galactinol per minute at 30°C (86°F) per gram fresh weight. (*After L. W. Handley et al., Relationships between galactinol synthase activity and sugar composition of leaves and seeds of several crop species, J. Amer. Soc. Hort. Sci., 108:600–605, 1983*)

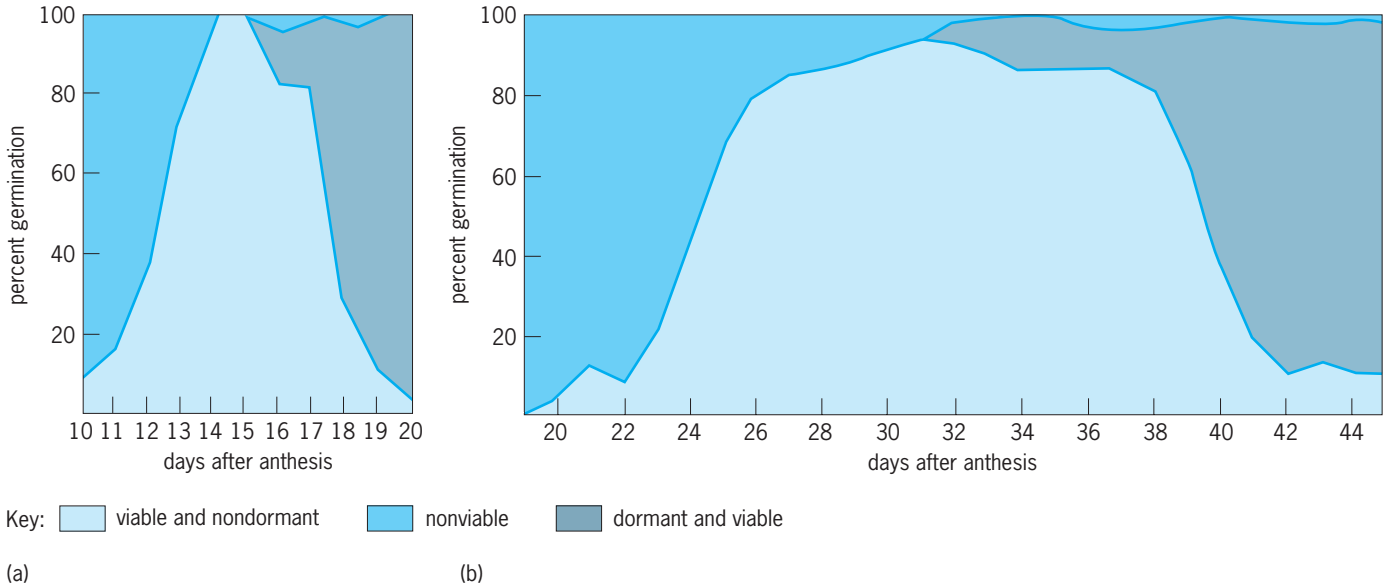

Fig. 6. Seed dormancy and viability during development under (*a*) field and (*b*) greenhouse conditions. Black medick (*Medicago lupulina*) seeds were detached from the plant and tested for germination. Three phases of germinability occurred depending on seed age; environmental conditions affect the length of each phase. (*After S. S. Sidhu and P. B. Cavers, Maturity-dormancy relationships in attached seeds of Medicago lupulina L. (black medick), Bot. Gaz., 138:174–182, 1977*)

Dormant seeds have all normal metabolic functions associated with nondormant seeds, except that the respiratory pentose phosphate pathway in dormant seeds may be inadequate to support germination. *See* DORMANCY; PLANT RESPIRATION.

Onset of dormancy. There is no one biochemical factor that has been determined to actually cause dormancy; however, factors affecting this process are genetic, environmental, and chemical (hormonal). Onset of seed desiccation during maturation appears to impart germability to seeds of barley and wheat, while low moisture content typical of a mature seed (about 15%) correlates with dormancy. Moisture loss specifically from seed coat tissues also is a factor in dormancy in common purslane (*Portulaca oleraceae*) and prickly sida (*Sida spinosa*). In seeds of black medick (*Medicago lupulina*), three phases of germinability/dormancy occur during development, the timing of which varies with the environment (**Fig. 6**). During early seed development, these seeds cannot germinate if detached from the plant. The seeds are subsequently viable and nondormant during seed fill, and can germinate on (precocious germination) or off the plant. These seeds finally become dormant during desiccation because of the development of an impermeable seed coat.

Both temperature and light may affect the onset of dormancy. Dormancy may be related to environmental temperatures during seed development in roses. The lower the daily average temperature the plants were grown at, the more pronounced the seed dormancy. In barley and wheat, embryos grown at 68–70°F (20–21°C) have greater incidence of dormancy than those grown at 82°F (28°C). In some cases such as seeds of *Syringa*, higher temperatures than normally encountered (64–75°F; 18–24°C) induce dormancy. Thermoperiod (different day/night temperatures) also may influence seed dormancy. In seeds of scarlet pimpernel (*Anagallis arvensis*), low dormancy occurs if seeds develop in a day/night temperature regime of 86°/77°F (30°/25°C), and extreme dormancy occurs in the seeds at 68°/59°F (20°/15°C) conditions.

Light-induced dormancy may be a photoperiodic (day/light time) or a phytochrome response. Long days appear to have greater effect than short days on inducing dormancy in developing seeds in species of *Chenopodium*, *Lactuca*, and *Portulaca*. The extent of this dormancy is related to the number of long days the developing seed receives. This time may or may not be correlated with seed coat thickness, depending on the species. Dormancy in a few species such as *Carrichtera anna*, however, is induced by short days. Nonphotoperiodic light effects on dormancy occur in mouse-ear cress (*Arabidopsis thaliana*) and cucumber, and they are controlled by the pigment phytochrome. Seeds from plants grown in far-red-enriched light (sunlight, incandescent) become dormant upon maturation, while those grown in red-enriched light (white fluorescent) are nondormant. *See* PHOTOPERIODISM; PHYTOCHROME.

Regulation of dormancy may be partly controlled by hormones. Maturing seeds, whether they undergo dormancy or not, accumulate abscisic acid. Induction of dormancy was correlated with abscisic acid in the embryo of mouse-ear cress. Abscisic acid also has been shown to inhibit translation by messenger ribonucleic acid (mRNA), an essential step in protein synthesis during early germination. *See* ABSCISIC ACID; RIBONUCLEIC ACID (RNA).

Release from dormancy. Dormancy is terminated in a large number of species when an imbibed seed is illuminated with white light. Biochemical control of this process is related to the functioning of a

single pigment, phytochrome, frequently located in the seed coat or embryonic axis. Phytochrome imparts to the seed the ability to interpret light quality, such as that under an existing vegetative canopy, and to distinguish light from dark with respect to its position in the soil. Phytochrome also is affected by temperature and is involved in the seasonal control of the ending of dormancy.

Light of three spectral regions interacts with phytochrome, blue (450 nanometers), red (650 nm), and far-red (750 nm); red light promotes, while far-red and blue light inhibit, seed germination. Phytochrome exists in two forms, P_r which absorbs red light and P_{fr} which absorbs far-red light. In a dormant seed in the dark, the P_r phytochrome form predominates. When light in the red range hits the seed, the phytochrome pigment converts from the P_r to P_{fr} form, and dormancy is terminated. The ratio of the two forms is critical to breaking dormancy. Germination is obtained when the P_{fr}/P_{total} ratio reaches a certain threshold value that is species-specific. Exposure to red light necessary to break dormancy may be a short burst, intermittent, or continual up to several days. The promotive effect of red light can be reversed by exposure to far-red light or long-term exposure to temperatures above 77°F (25°C). Blue light elicits a high-irradiance response in the seed. Short exposure to blue light reverses the effect of the red light and inhibits germination. The high-irradiance response, however, affects primarily elongation of the radicle after germination.

Temperature also can be a factor in ending seed dormancy, typically in imbibed seeds. Dormancy is broken via diurnal (daily) fluctuations or one-time shifts in temperature. In tobacco, diurnal alternations in temperature between 68 and 86°F (20 and 30°C) induce seed germination better than at either one temperature. Seed chilling or stratification at 34–59°F (1–15°C) for a few hours to days is a common requirement of temperate woody plants such as maple, apple, and hawthorn for germination. In many seeds with this requirement, stratification at 41°F (5°C) for several days prior to returning the seeds to ambient temperature is optimal for breaking dormancy in fully imbibed seeds.

Dormant seeds also may gradually overcome blocks to germination without imbibition (after-ripening). After-ripening is a time of physiological and anatomical maturation after the seed is shed from the plant that frequently occurs in cereals such as barley, rice, and wheat, and in grasses such as wild oat. After-ripening is stimulated by oxygen and high temperatures (86–104°F; 30–40°C).

Hormones that promote germination of dormant seeds include gibberellins, cytokinins, ethylene, and auxins. Exogenously applied gibberellins promote germination in dormant seeds, but internal concentrations of this hormone have not been shown to increase during dormancy. Cytokinin levels also have not been shown to change in dormant seeds, but cytokinins function in conjunction with gibberellins to induce germination when exogenously applied. Ethylene has stimulated germination of some species, and it has been linked to phytochrome and the cellular membrane state. Auxin does not promote germination when exogenously applied, but it does increase endogenously and is involved in loosening and expansion of cell walls, which is necessary for germination. Additional chemical compounds that stimulate seed germination when applied exogenously include plant or fungal chemical products (fusicoccin), respiratory inhibitors (cyanide, azide, hydroxylamine, hydrogen sulfide, carbon monoxide), oxidants, carbon dioxide nitrates or nitrites, sulfhydryl compounds, and certain organic compounds (ethanol, methanol). *See* AUXIN; CYTOKININS; ETHYLENE; GIBBERELLIN.

Germination. Germination is the process whereby a viable seed takes up water and the radicle (primary root) or hypocotyl emerges from the seed under species-specific conditions of moisture, oxygen, and temperature. Dormant seeds must undergo additional environmental stimuli to germinate. The germinating seed undergoes cell expansion, as well as increases in respiration, protein synthesis, and other metabolic activities prior to emergence of the growing seedling (**Fig. 7**).

As water, dissolved oxygen, and solutes from the soil enter the seed, there is a release of gas, unrelated to respiration, that was trapped in the seed following desiccation. Imbibition is accompanied by rapid leakage of substances of cytoplasmic origin such as sugars, organic acids, proteins, gibberellins, phenolics, phosphate, and potassium ions. Radicle expansion begins soon after the start of imbibition, and initial cell wall elongation proceeds slowly, lasting up to 46 h. This period is followed by rapid cell wall elongation, mobilization of metabolic reserves, and translocation of these materials to the growing radicle. Cell division may or may not be necessary for radicle protrusion. Cell expansion typically precedes cell division in seeds of corn, barley, broad bean, and pea.

Respiration of germinating seeds. Respiration in dry seeds is nearly nondetectable. In imbibed seeds, the three major respiratory pathways (glycolysis, pentose phosphate pathway, and Krebs cycle) are very active; they provide energy [adenosine triphosphate (ATP)], reducing power [β-nicotinamide adenine dinucleotide (NADH) and β-nicotinamide

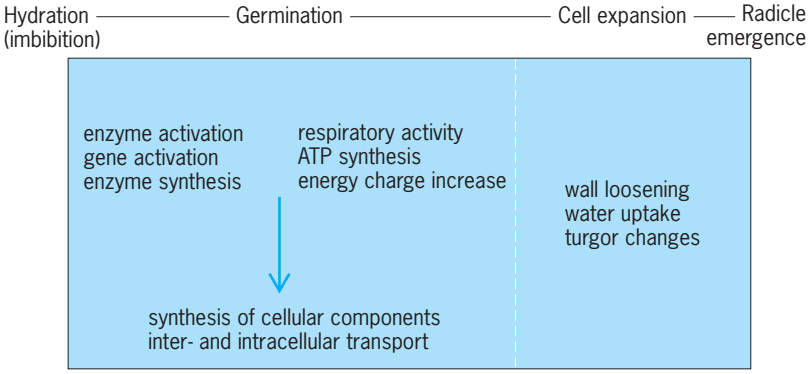

Fig. 7. Biochemical events during seed germination. (*After J. D. Bewley and M. Black, Physiology and Biochemistry of Seeds, vol. 2, Springer-Verlag, 1982*)

adenine dinucleotide phosphate (NADPH)], and key metabolic intermediates.

Respiration occurs in four phases as the seed imbibes; the length of each phase varies with species and environment. Aerobic respiration increases rapidly with imbibition and lasts approximately 10 h. This rapid respiration results from physical activation of preexisting cellular systems. A lag in respiration occurs 10–25 h after the start of imbibition. The seed is now completely hydrated, preexistent enzymes have been activated, and loosening and extension of the cell walls begin. Some of the respiration at this point may be anaerobic, and this lag may be due to a lack of oxygen. The seed may be held at this stage and even desiccated without loss of viability. Once rehydrated, the seed starts germination at the same point at which it was desiccated. The growing radicle subsequently breaks through seed coat tissues, and there is a second burst of rapid aerobic respiration. Oxygen is plentiful to the radicle with the breaking of the seed coat, and activity of newly synthesized mitochondria and respiratory enzymes in dividing cells increases dramatically. Finally, respiration decreases with the depletion of stored metabolic reserves. *See* ADENOSINE TRIPHOSPHATE (ATP); BIOLOGICAL OXIDATION; CITRIC ACID CYCLE.

Protein–nucleic acid synthesis. Protein synthesis is a prerequisite for radicle emergence, while synthesis of deoxyribonucleic acid (DNA) occurs only after germination and is associated with seedling growth. Protein synthesis begins when embryo cells are hydrated sufficiently to allow for the assembly of the necessary ribosomes and messenger RNA, within a few minutes to hours after the start of imbibition. In isolated embryos of rye, rice, and wheat, and dissected axes of lima bean and french bean, protein synthesis commences within 30–60 min after the start of inbibition. Protein synthesis results in formation of metabolic enzymes also found in the developing seed, as well as unique, so-called germination proteins such as certain proteolytic enzymes.

Mobilization of storage reserves. Carbon and nitrogen for seedling growth prior to soil emergence (heterotrophic growth) are derived from stored seed reserves in the form of carbohydrates, oils, and protein. Carbohydrate is the first storage compound to be used for energy by the growing embryo, beginning with imbibition. Starch stored in the endosperm of monocots such as barley and corn is degraded by enzymes synthesized in aleurone cell layers and secreted to the endosperm during imbibition. The resulting sugar, glucose, is absorbed by the scutellum, enzymatically converted into sucrose, and transferred to the growing embryonic axis. The scutellum also may synthesize and secrete hydrolytic enzymes. In dicots such as pea, where starch is also the major storage carbohydrate, this carbon source is mobilized directly in the cotyledons and transferred as sucrose to the growing axis. In the case where oligosaccharide sugars (sucrose, raffinose, stachyose) are the main storage carbohydrate as in cucumber and soybean (**Fig. 8**), there is a rapid enzymatic hydrolysis of raffinose and stachyose to sucrose with a concomitant formation of starch. This starch is degraded after germination to provide additional carbon energy for the seedling. Mobilization of raffinose saccharides is controlled by a balance of supply and demand between the cotyledons and growing radicle. Formation of starch may be a regulating mechanism for this process.

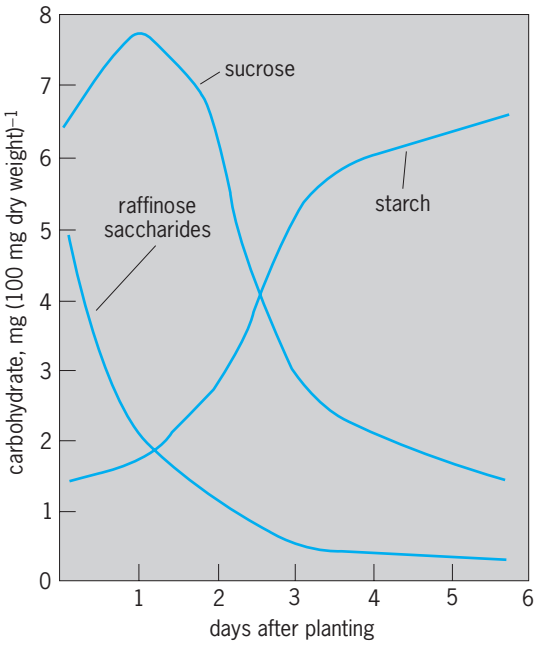

Fig. 8. Depletion of sucrose and raffinose saccharides (raffinose, stachyose) and accumulation of starch in etiolated soybean cotyledons when grown on moist germination paper. (*After C. S. Brown and S. C. Huber, Reserve mobilization and starch formation in soybean (Glycine max) cotyledons in relation to seedling growth, Physiol. Plant., 72:518–524, 1988*)

Genetic and hormonal control of carbohydrate mobilization in germinating seeds has been studied extensively in barley. Exogenously applied gibberellins and possibly those occurring endogenously appear to control the synthesis and release of hydrolytic enzymes for nutrient mobilization in the barley aleurone layer. For the starch-degrading enzyme α-amylase, gibberellins promote increases in an mRNA specific for this enzyme. Abscisic acid inhibits this gibberellin-induced synthesis of α-amylase and other hydrolytic enzymes.

Primarily as a postgerminative process, proteins and lipids are mobilized to supply nitrogen and carbohydrates, respectively, to the growing seedling. While protein is the main nitrogen source, RNA and amino acids also can contribute. Protein breakdown typically occurs in three stages. First, there is hydrolysis of proteins for the synthesis of additional hydrolytic enzymes for further protein degradation. Second is the hydrolysis of the main reserve proteins to provide amino acids for the growing seedling. Third, the seed storage tissue senesces and provides additional amino acids from cellular proteins and nuclear material for seedling growth. Oil-storing seeds typically are dicot species (for example, soybean, sunflower, and peanut), and mobilization of these reserves lags up to 2–3 days after imbibition. This

degradation then proceeds at a constant rate for 3–8 days. Mobilization of lipids is enzymatic, and cellular components (oil bodies, glyoxysomes, mitochondria, and cytosol) are involved in the conversion of lipids to carbohydrate for transport to the growing radicle. *See* CARBOHYDRATE; LIPID; PROTEIN. Cadance A. Lowell

Storage. Seed storage is necessary for food production and conservation of genetic resources. Seed companies and farmers need to keep seeds alive for short periods to ensure production of a high-quality crop. Conservationists need to keep seeds alive for decades to ensure that valuable genes will not be lost even if a plant species or population becomes extinct in the wild. In an effort to slow the continued erosion of genetic diversity in cultivated and noncultivated plant species, national and privately owned gene banks are storing seeds as germplasm for the future.

Seeds possess a remarkable attribute which allows storage until the next growing season or the next century. They survive drying. Drying slows the molecular motions needed for chemical reactions. When seeds dry, their metabolism stops and they go into "suspended animation." Many species of seeds survive complete desiccation, so they can be stored under standard conditions and their life spans can be predicted using models. This type of seed is termed orthodox. All major food crops (corn, wheat, and soybeans) produce orthodox seeds. Recalcitrant seeds represent another category: those which do not survive complete drying. Recalcitrant seeds must remain metabolically active and so cannot be stored for more than a year. Temperate forest trees and plants from tropical rainforests or aquatic ecosystems often produce recalcitrant seeds. Many tropical fruits, such as cacao, mangos, and avocados, produce recalcitrant seeds. The ability to survive complete desiccation is acquired during the final days of maturation in orthodox seeds. When scientists discover the details of these processes, they may be able to store recalcitrant seeds more effectively. *See* CACAO; CORN; PLANT METABOLISM; SOYBEAN.

The best way to preserve biological material is to keep it dry and cold. In the early 1960s, J. F. Harrington developed "thumb rules" to describe the relative longevity of orthodox seeds under various storage conditions. He stated that the longevity of a seed doubles for every 1% reduction in water content or 5°C (9°F) reduction in temperature. Seed gene banks use these principles to keep seeds alive for decades. The International Plant Genetic Resources Institute in Rome recommends drying seeds to water contents of 3–7%, packaging them in airtight bags, and then placing them in a freezer at −18°C (−0.4°F). Oily seeds such as peanuts store best at 3% water, and starchy seeds such as wheat store best at 7%. Equilibrating seeds to 20% relative humidity will achieve the optimum moisture content no matter what the chemical composition.

The U.S. Department of Agriculture's National Seed Storage Laboratory (NSSL) in Fort Collins, CO, is the largest seed gene bank in the world. The NSSL stores over 415,000 accessions or bags of seeds, representing over 8000 species of plants. Each bag contains more than 2000 seeds, so there are literally millions of individual plants being stored. In 1977, NSSL scientists began to store some seeds in liquid nitrogen vapor (−150°C or −238°F) with the hope that these seeds would survive centuries, if not millennia.

Longevity. While storing seeds under dry, cold conditions can extend life spans, the innate longevity of a seed is determined by the seed quality. Seed quality is a poorly understood feature that is a product of both the genetic background and conditions of growth and harvest. Seed species have characteristic longevities. Species from Malvaceae, Onagraceae, Polygonaceae, and Leguminosae tend to produce long-lived seeds, while species from Compositae, Papaveraceae, and Liliaceae tend to produce short-lived seeds. Usually there is tremendous variability in longevity among cultivars of a species and among harvest years of the same cultivar. Environmental conditions such as drought, temperature, and nutrient deficiencies and premature or late harvests can adversely affect seed longevity. Extremes in longevity are noted in seeds collected from wild populations since growing conditions are variable from site to site and year to year.

Records of seed longevity are often anecdotal but indicate remarkable potential for long-term survival (see **table**). In some surveys, seeds survived decades in unrefrigerated conditions. W. J. Beal buried seeds from 23 different species in inverted bottles in the autumn of 1879 with the intention of measuring how long they would remain dormant in the

Results of some key studies in seed longevity

Longevity, years	Species	Storage conditions	Researchers and year of publication
1700–110	*Chenopodium album* and other species	Soil	S. Odum, 1965
1288–271	*Nelumbo nucifera*	Soil	J. Shen-Miller et al., 1995
295–75	*Carex bigelowii*	Soil	J. B. Mcgraw et al., 1991
158–25	500 species	Museum	P. Becquerel, 1907
100	23 species	Soil	A. Kivilaan and R. S. Bandurski, 1981
50	600 species	Museum	A. J. Ewart, 1980
60–45	15 vegetable species	5 to −18°C	E. E. Roos and D. A. Davidson, 1992
50	98 species	Vacuum tubes, 15–25°C	F. W. Went, 1969 (unpublished data)
39	107 species	Soil	E. H. Toole and E. Brown, 1946
25	17 crucifer species	Vacuum tubes, −5°C	R. H. Ellis et al., 1993
20	9 species	−15°C	C. M. Rincker, 1983

soil. After 100 years, some seeds were still able to germinate. The sacred lotus (*Nelumbo*) is most often and most reliably reported as the longest-living seed. Seeds from herbarium specimens collected 250 years ago still germinated, and germinable seeds from the soil were carbon-dated at over 1200 years (see table). Archeological dating of soil samples collected along with germinable seeds of *Chenopodium album* suggested this sample may have lived for 1700 years. There also are claims that *Lupinus arcticus* survived for 10,000 years in the permafrost of Alaska, but the evidence to support this claim is weak. There is no scientific evidence for the popular myth of seeds surviving for thousands of years in Egyptian tombs.

Aging. Even when stored under optimum conditions, seeds eventually succumb with time. As they age, seeds lose the ability to tolerate stress, then they germinate slowly or produce abnormal seedlings, and finally they die. The cells in aging seeds change: membranes become leaky, enzymes lose catalytic activity, and chromosomes accumulate mutations. It was once proposed that seeds consume food reserves during storage, but the huge amount of reserves contained in seeds and the low level of metabolism measured in dried seeds challenge this hypothesis. Only a few chemical and physical reactions can occur when seeds are very dry, and most scientists believe that peroxidative reactions, in which free radicals arise from activated oxygen and attack cells, cause the deterioration in dry seeds. Dry seeds cannot repair the lesions to their cell machinery because metabolism has been shut off. It is ironic that the ability to stop metabolism allows seeds to survive for a very long time, but also leads to their eventual death.

Darren H. Touchell; Christina Walters; Eric E. Roos

Bibliography. J. D. Bewley and M. Black, *Physiology and Biochemistry of Seeds in Relation to Germination, Viability, Dormancy and Environment Control*, 1982; J. D. Bewley and M. Black, *Seeds: Physiology of Development and Germination*, 2d ed., 1994; L. S. Dure, Seed formation, *Annu. Rev. Plant Physiol.*, 25:259–278, 1975; O. L. Justice and L. N. Bass, Principles and Practices of Seed Storage, *Agri. Handb.*, no. 506, 1978; J. Kigel and G. Galili (eds.), *Seed Development and Germination*, 1995; A. M. Mayer and A. Poljakoff-Mayber, *The Germination of Seeds*, 4th ed., 1989; D. R. Murray (ed.), *Seed Physiology*, vol. 2: *Germination and Reserve Mobilization*, 1984; D. A. Priestly, *Seed Aging*, 1986; T. F. Stuessy and S. H. Sohmer (eds.), *Sampling the Green World: Innovative Conecepts of Collection, Preservation, and Storage of Plant Diversity*, 1996; J. R. Wilcox (ed.), Soybeans: Improvement, Production, and Uses, *Amer. Soc. Agron. Monog.*, no. 16, 1987.

Seiche

A short-period oscillation in an enclosed or semienclosed body of water, analogous to the free oscillation of water in a dish. The initial displacement of water from a level surface can arise from a variety of causes, and the restoring force is gravity, which always tends to maintain a level surface. Once formed, the oscillations are characteristic only of the geometry of the basin itself and may persist for many cycles before decaying under the influence of friction. The term "seiche" appears to have been first used to describe the rhythmic oscillation of the water surface in Lake Geneva, which occasionally exposed large areas of the lake bed that are normally submerged. *See* WAVE MOTION IN LIQUIDS.

Behavior. The most straightforward example of a surface seiche is that occurring in a long, narrow lake of uniform depth; in this case the water surface elevations have a simple physical analogy with the vibrations of a plucked string. Since the wave length is long relative to the water depth, seiche behavior conforms to the theory of shallow water waves. Thus the velocity is given by $\sqrt{gh}$, where g is the acceleration of gravity and h is the water depth. A standing wave can be represented by the sum of two waves traveling in opposite directions, each being reflected at the ends of the basin. The natural period of oscillation of a uniformly deep, completely enclosed rectangular basin is given by Merian's formula (1), where

$$T = \frac{2L}{n\sqrt{gh}} \tag{1}$$

L is the length of the basin and n is the number of nodal lines present, one for the fundamental or uninodal seiche, two for the binodal seiche, and so forth (**Fig. 1**). However, the formula is only approximate, and calculation of the higher modes of oscillation starts to depart seriously from that given by Eq. (1) due to the importance of nonuniformity of the basin shape.

The **table** gives examples of observed periods of surface seiches in different lakes. Figure 1 shows profiles of a standing wave in a closed basin of the fundamental, binodal, and trinodal seiche. Below the nodes there is maximum horizontal and no vertical motion; below the antinodes there is maximum vertical and no horizontal motion. **Figure 2** shows simultaneous records of water height near each end of Lake Vättern during a uninodal seiche. The two measurements show opposite movement of the water at each antinode. When several different seiches occur in a basin, their effects are superposed. Thus it is quite possible, for example, for a vessel to be located at the node of the fundamental seiche but at the antinode of the binodal seiche. The amplitudes of the different harmonics are determined by the initial disturbance.

It is also possible for transverse seiches to exist together with the longitudinal seiche. The appropriate generalization of Merian's formula for a deep, rectangular basin of length L and breadth B is given by Eq. (2), where $\alpha = L/B$ and m and n define the harmonic nodes in the transverse and longitudinal directions, respectively. However, the earlier comments

$$T_{mn} = \frac{2L}{\sqrt{gh}}\,(\alpha^2 m^2 + n^2)^{-1/2} \tag{2}$$

regarding nonuniformity of basin shape still apply, and Eq. (2) cannot be generally applicable to natural basins.

Seiches are also a common feature of semienclosed basins such as bays, gulfs, and harbors, the natural period for an open-end rectangular basin of uniform depth being twice that of the same-size basin closed at both ends. The node and antinode occur at the open and closed ends, respectively. A problem of great practical significance is the oscillation of water in harbors, which can generate strong and unpredictable currents affecting the safety of ships entering or leaving the harbor; under extreme conditions these can also cause damage to moored vessels.

Generation and decay. Seiches can be generated when the water is subject to changes in wind or atmospheric pressure gradients or, in the case of semienclosed basins, by the oscillation of adjacent connected water bodies having a periodicity close to that of the seiche or of one of its harmonics. Other, less frequent causes of seiches include heavy precipitation over a portion of the lake, flood discharge from rivers, seismic disturbances, submarine mudslides or slumps, and tides. The most dramatic seiches have been observed after earthquakes.

The amplitude and persistence of the resulting seiche depend not only on the magnitude of the energy source but also on the energy losses within the water body. Such losses include dissipative effects resulting from friction on the sides or bottom of the basin. For semienclosed basins, energy can also be lost by the radiation of waves away from the mouth. If the seiche is generated by an impulsive event such as a sudden change in atmospheric pressure gradient, the seiche amplitude is seen to decay by a nearly constant fraction with each succeeding period. The table shows, for example, that in Lake Geneva the decay has been estimated to be about 3% with each seiche period, whereas in Lake Erie it is 32%. In general, the rate of decay is greater for basins that are shallow or have narrow constrictions and complex topography.

In bays and gulfs a dominant source of energy is the tide. Since there is quite a rich array of tidal frequencies, it is sometimes possible to determine the natural seiche frequency of a bay simply by observing the amplification of different tidal components. If the tidal period is close to that of the seiche period for the bay, resonance increases the tidal amplitude.

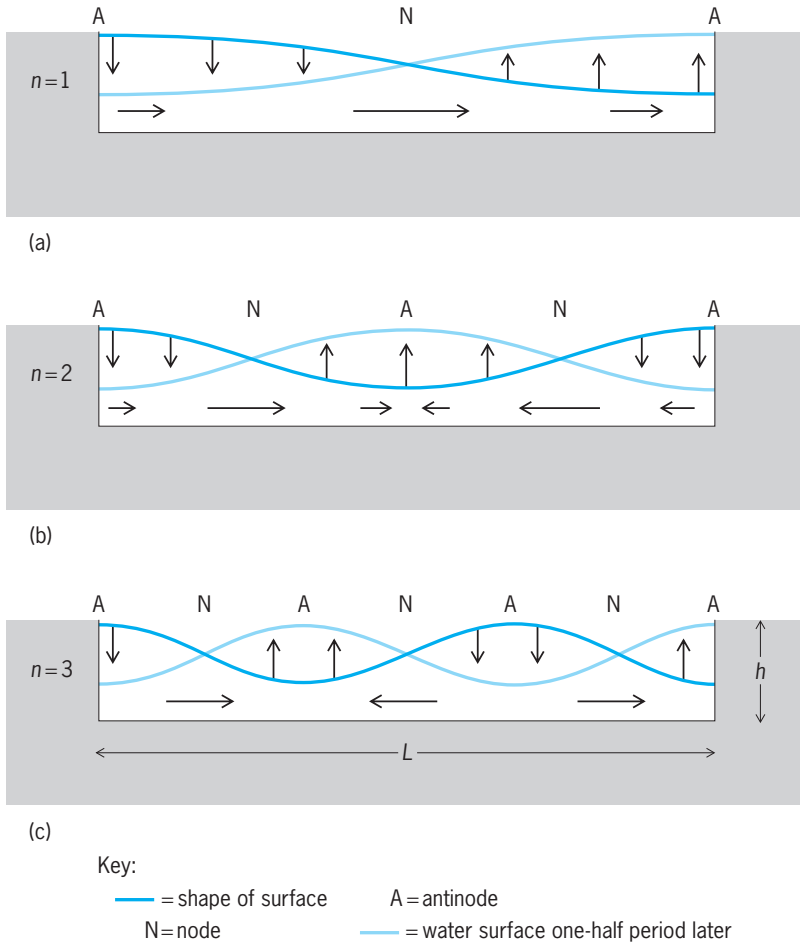

Fig. 1. Schematic diagram of a (*a*) uninodal, (*b*) binodal, and (*c*) trinodal seiche in a rectangular basin. Vertical and horizontal vectors show the direction of flow corresponding to the indicated change of water surface. Vertical movements are greatly exaggerated.

An example of this is the Bay of Fundy (Canada), where the tide can exceed 45 ft (15 m). The bay has a fundamental mode seiche period of about 13.3 h, close to the semidiurnal lunar tidal period of 12.4 h. A similar, near-resonance exists in the Gulf of Mexico, where the period of the fundamental seiche mode is close to 24 h, with resulting amplification of the diurnal tide. In smaller bays and harbors the generation of seiches has been attributed to surf beats. This effect is due to the coupling of radiation stress, in effect the thrust of the waves on the coast, with

Observed surface seiche periods in typical lakes and observed decay rates

Lake and location	Observed periods of oscillation, min						Fractional decrease in amplitude of free uninodal seiche over each successive period
	T_1	T_2	T_3	T_4	T_5	T_6	
Geneva (Switzerland–France)	74.0	35.5					.030
Garda (Italy)	42.9	28.6	21.8	15.0	12.1	9.9	.045
Loch Earn (Scotland)	14.5	8.1	6.0	4.0	3.5	2.9	
Erie (United States–Canada)	858.0	542.3	350.9	250.5			.322
Königssee (Germany)	11.6						.204
Vättern (Sweden)	179.0	97.5	80.7	57.9	48.1	42.6	.113
Yamanaka (Japan)	15.6	10.6	5.5				.099
Baikal (Russia)	278.2						

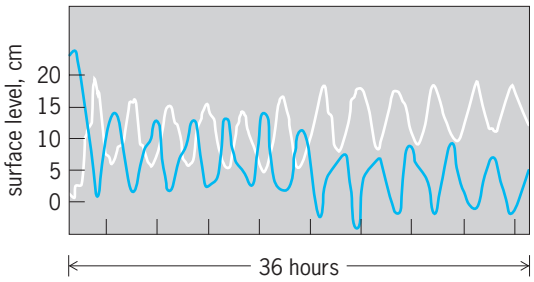

Fig. 2. Simultaneous recordings taken near each end of Lake Vättern during a uninodal seiche. The reading at one end is a mirror image of the observed displacement at the other end. 1 cm = 0.4 in. (*After A. Defant, Physical Oceanography, vol. 2, Pergamon, 1961*)

rhythmic changes in the height of the wind waves or swell. Since the swell can originate from distant storms, the existence of seiches is not necessarily closely correlated with local meteorological conditions. Coastal seiches can also be generated by traveling fronts or pressure disturbances, especially when the atmospheric disturbance travels at a speed close to that of the gravity wave ($\sqrt{gh}$). A less frequent but often important generation mechanism in harbors is the tsunami or seismically formed ocean wave. *See* TSUNAMI.

Effect of rotation. The response of large bodies of water is significantly influenced by the Coriolis effect produced by the rotation of the Earth. The effective period of the rotation decreases with latitude and is given by $^1/_2\ \tau/\sin\lambda$, where τ is the period of rotation of the Earth (24 h) and λ the latitude. The Coriolis force acts at right angles to the direction of motion of the water, to the right in the Northern Hemisphere and to the left in the Southern Hemisphere. The Coriolis force introduces a transverse slope to each of the traveling waves that combine to form a standing wave pattern, or seiche. The transverse motion, added to the motion along the channel, results in the movement of water in an elliptical path rotating once with each seiche period. The transverse motion has the effect of shrinking the nodal lines of zero vertical motion to small areas known as amphidromic points.

In the plan views of **Fig. 3***a* and *b*, arrows show surface currents at one-fourth and three-fourth cycles, respectively. After high water at the channel head, Coriolis deflection, indicated by broken arrows, causes an increase of water level in the right half of the current in both Fig. 3*a* and *b*, with a corresponding decrease to the left. The broken line is one of mean level with respect to the deflection. Water profiles along nodal lines are shown beneath Fig. 3*a* and *b*. Point O, the amphidromic point, is thus the only position of no water level change. Line O-N (O-N′) connects points having simultaneous high (low) water. Other such cotidal lines are shown in Fig. 3*c* for different phases of one seiche oscillation. The superposition of the transverse motion on the primary longitudinal motion causes this line to rotate counterclockwise (dotted arrow) for one complete oscillation in the Northern Hemisphere. *See* CORIOLIS ACCELERATION; TIDE.

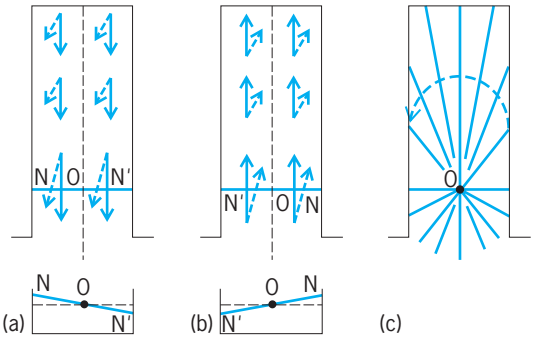

Fig. 3. Schematic representation of transverse oscillations in a bay in the Northern Hemisphere leading to the development of an amphidromic point. (*a*) Ebb current. (*b*) Flood current. (*c*) Cotidal lines.

Internal seiches. Internal seiches are standing waves that occur in water that has vertical variations in temperature or salinity resulting in a vertical density stratification. The waves produce vertical movements of water at depth, with very little vertical motion at the surface. They are typically generated by the piling up of water at one end of the lake by wind, although in semienclosed basins they may be forced by oscillations in the density structure at the mouth of the basin. **Figure 4** shows the measurements of two isotherm depths in Lake Alpnach (Switzerland) generated by a wind pulse as well as by diurnal wind forcing. The oscillations decay by one-third over about 1 day. As for surface seiches, the residence time of the energy varies, depending on the excitation (tens to hundreds of joules per square meter) and the depth of the basin, from less than 1 day up to several weeks.

Since the relative density differences within the water column are much smaller than the density difference at the air-water interface, the same amount of energy produces internal seiches of much greater amplitude than the corresponding surface seiche. However, the internal wave speed is much slower than the surface wave. The relative density difference in lakes during the summer is typically of the order of 0.002, so that wind effects, which might produce surface seiches of only a few centimeters, can produce internal seiches of meters to tens of meters (Fig. 4). In the Great Lakes of North America and the long and deep Rift valley lakes of Baikal and East Africa, amplitudes can reach almost 100 m with periods of many weeks to months.

In natural waters, the two-layer approximation, as depicted for surface seiches in Fig. 1, is a simplification, and in nature the density varies continuously with depth. Subsequently, the density structure allows several vertical modes, with characterizing vertical wavelengths, to exist. Such an example is given in Fig. 4 where, beside the fundamental seiche period of about 8 h, a second vertical mode with a 24-h periodicity is prominently displayed. The characteristic of the second vertical mode is the periodic widening and thinning of the middle layer (between 11 and 17°C in Fig. 4). The fluid of the lighter top layer (above 17°C) and the fluid in the heavier deep layer (below 11°C) move in phase, whereas the fluid

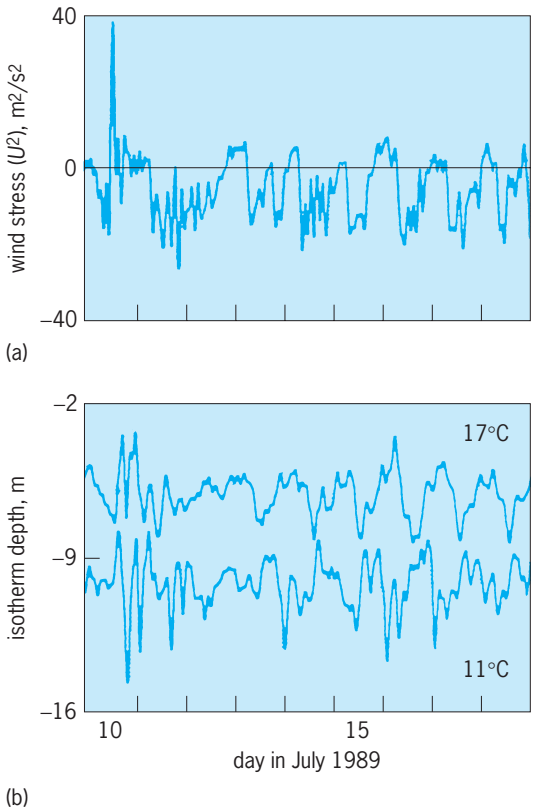

Fig. 4. Seiche measurements in Lake Alpnach, in Switzerland. (*a*) Time series of the lake-parallel component of the square of the wind speed (proportional to the wind stress) over the lake (maximum depth = 35 m). (*b*) Times series of the two selected isotherms of 11 and 17°C. The fluid layer between the two lines shows a strong second vertical mode characteristics of 24-h periodicity (V2H1). The fundamental mode (V1H1) of an 8 h period is also prominent. (*After M. Münnich, A. Wüest, and D. M. Imboden, Observations of the second vertical mode of the internal seiche in an alpine lake, Limnol. Oceanogr., 37:1705–1719, 1992*)

in the middle layer has the exactly opposite direction. Higher vertical resolution (by dividing the water column in more homogeneous layers) leads to even more vertical modes, generally expressed by *VpHq*, where *p* indicates the *p*-th vertical mode, and *q* the *q*-th horizontal mode. According to Eq. (2), the horizontal modes are combinations of the two horizontal coordinates (modes *m*, *n*). The Earth's rotation also alters the structure of internal seiches since the transverse slope of the internal interface induced by the Coriolis force is greatly increased for internal seiches.

Which modes are excited, and to what extent, in a particular water body depends on the forcing and the stratification, but is almost impossible to predict. As in the case of surface seiches, internal seiching is a mechanism that selectively extracts energy from the forcing (mostly wind) at a frequency close to that of the internal seiche. For example, as shown in Fig. 4, the second vertical mode in Lake Alpnach is resonantly forced at a 24-h period by the local diurnal wind.

The latter example demonstrates that the amplitude of internal seiches can reach a significant fraction of the basin depth. This enormous amplitude results in a distortion of the smooth sinusoidal perturbation of the interface predicted by linear theory, and the peak and trough of the wave travel at slightly different speeds. In addition, interaction with the topography and bottom friction leads to turbulence and the generation of short-period internal waves radiated into the water body. Especially in large basins and lakes, where resonant seiching is difficult to reach, most internal energy is contained in progressive short-period waves. Internal shear, bottom friction, and wave breaking leads to damping of the internal seiche energy. As a rule, the decay rate of the energy is about 1 day per 40 m of water depth (40 days in Lake Baikal). However, for very weak and almost linear seiching, damping may reduce the decay rate.

The enormous amplitude of internal seiches results in a distortion of the smooth sinusoidal perturbation of the interface predicted by linear theory. In effect, the peak and trough of the wave travel at slightly different speeds, resulting in steepening of a portion of the interface. Under the right conditions, the steepened interface can break down into a series of shorter waves, possibly accompanied by turbulence and mixing and usually referred to as an internal surge. Alfred Wüest; David M. Farmer

Bibliography. J. R. Apel, *Principles of Ocean Physics*, 1987; A. Defant, *Physical Oceanography*, vol. 2, 1961; J. Miles, Harbour seiching, *Annu. Rev. Fluid Mech.*, 6:17–36, 1974; M. Münnich, A. Wüest, and D. M. Imboden, Observations of the second vertical mode of the internal seiche in an alpine lake, *Limnol. Oceanogr.*, 37:1705–1719, 1992; B. W. Wilson, Seiches, *Advan. Hydrosci.*, 8:1–94, 1972.

Seismic exploration for oil and gas

Prospecting for oil and gas using exploration seismology, a geophysical method of determining geologic structure by means of prospector-induced elastic waves. By studying body waves such as compressional and shear waves propagating through the Earth's interior, the constituent and elastic properties of its solid and liquid core, its solid mantle, and its thin crust are defined. The major differences between earthquake seismology and petroleum exploration seismology are scales and knowledge of the location of seismic disturbances. Earthquake seismology studies naturally generated seismic waves, which have periods in minutes and resolution in kilometers. In exploration seismology, artificial sources are used that have periods of tenths of a second and tens of meters of resolution. Production seismology requires higher-frequency seismic waves and better resolution, often resolution in the order of a few meters. *See* EARTH; EARTH INTERIOR.

Seismic technology. Computer technology allows resolution of some of the theoretical complexities of elastic wave propagation so that deeper insight into the wave field phenomena can be obtained. The availability of a large number of channels in the recording instrument facilitates three-dimensional

and three-component acquisitions. Powerful supercomputers allow manipulation of larger and larger data sets, and they have facilitated display and interpretation of them as a single data unit through the use of advanced computer visualization techniques. *See* COMPUTER; SUPERCOMPUTER.

The availability of powerful workstations led to the wide use of interactive processing and interpretation. When such is coupled with technically advanced algorithms, the amount of information that the interpreter can obtain from the data increases significantly. Better quality control is provided, fine-tuning analysis is achieved more easily, and the data can be enhanced to meet specific objectives. As a result of technological advances, the working relationship among professionals of different disciplines has changed. Processing specialists, interpreters, geologists, and engineers use workstations concurrently to test new exploration concepts through computer simulations or modeling that can be validated through special processing of the seismic data. If there are discrepancies between the model and the real data, a hypothesis can be proposed based on information derived from the data. This process can be iterated until the Earth model derived is consistent with all available surface and subsurface geophysical, petrophysical, geological, and engineering data sets. *See* ALGORITHM; MODEL THEORY; SIMULATION.

Seismic method. The seismic method as applied to exploration of oil and gas involves field acquisition, data processing, and geologic interpretation.

Field acquisition. Seismic field acquisition requires placement of acoustic receivers (geophones) on the surface in the case of land exploration, or strings of hydrophones in the water in the case of marine exploration. The conventional geophone is a vertical-component velocity phone. A triaxial phone with two horizontal components may be used if shear-wave recording is desirable. A group of several receivers is generally used for each receiver point location. *See* HYDROPHONE.

After all instruments have been placed, an artificial earthquake is initiated by an impulsive or a vibratory source at the shot point. This generates elastic seismic waves that propagate through the subsurface. Commonly used sources are dynamite, vibrators, and air guns. These elastic seismic waves travel through different geologic strata at various speeds according to the acoustic properties of the composite rocks and minerals as well as the composite fluids in the reservoirs under various conditions. When these seismic waves arrive at the interfaces of acoustic impedance contrasts (impedance is a product of velocity and density), part of the energy is transmitted and part of the energy is reflected. The reflected waves eventually emerge back at the surface and are picked up by the receivers. The electric signal generated by the receivers is recorded digitally on magnetic tapes or cassettes in a multichannel seismic recorder, ready for preliminary processing in the field or further enhancement in processing centers.

Data processing. Seismic data processing is usually done in large computing centers with digital mainframe computers or a large number of processors in parallel configurations. Seismic data processing has several goals. One aspect of the processing is to compensate for wave propagation effects such as spherical spreading and attenuation, and to account for the time differences for the ray paths between different source-to-receiver distances. This includes seismic velocity estimation from the data. The processing also accounts for effects generated by different shots and the near-surface conditions such as the varying thickness and lithologic composition of the weathered layers. In addition, it ensures spatial and temporal consistency of the propagation wavelet. Seismic data processing can also improve the signal-to-noise ratio. Random noise is attenuated by the stacking of many redundant traces after normal moveout correction, which is the correction that has to be applied to seismic data to take into account the variation of reflection arrival time because of variation in the shotpoint-to-geophone distance. Organized noises such as surface waves, wind noise, repeated refraction, and multiples are filtered according to the principles of communication or information theory. Temporal and spatial resolution are achieved by using techniques known as deconvolution and migration, respectively. Deconvolution is an inverse filtering process that takes into account the characteristics of the shot signatures as well as some of the propagation effects of the wavelet as it propagates through the subsurface. Migration is a process that takes into account the true spatial position of reflectors. Migration ensures proper imaging of the reflective strata when it is performed correctly.

The end result of seismic data processing is the production of a subsurface profile similar to a geologic cross section. It is commonly plotted in a time scale, but it is also possible to plot it in depth. The accuracy of the depth plot, however, varies with the skill of the processor, the quality of data, and the knowledge of the prospecting area such as formation thickness and velocity information. These time or depth profiles are used for geologic interpretation.

Geologic interpretation. Geologic interpretaion of seismic data has two key components, structural and stratigraphic. Structural interpretation of seismic data involves mapping of the geologic relief of different subsurface strata by using seismic data as well as information from boreholes and outcrops. Stratigraphic interpretation looks at attributes within a common stratum and interprets changes to infer varying reservoir conditions such as lithology, porosity, and fluid content. These seismic attributes can be amplitude, velocity, travel time, absorption, frequency, phase, or amplitude-versus-offset, an interpretation technique that uses the difference in reflection strength as a function of offset (source-to-receiver separation) to infer reservoir properties. The interpretation of reservoir parameters from seismic data is an inversion process that is inherently nonunique. Information from other disciplines such as petrophysics, rock physics, and engineering must therefore be used as constraints for stratigraphic interpretations.

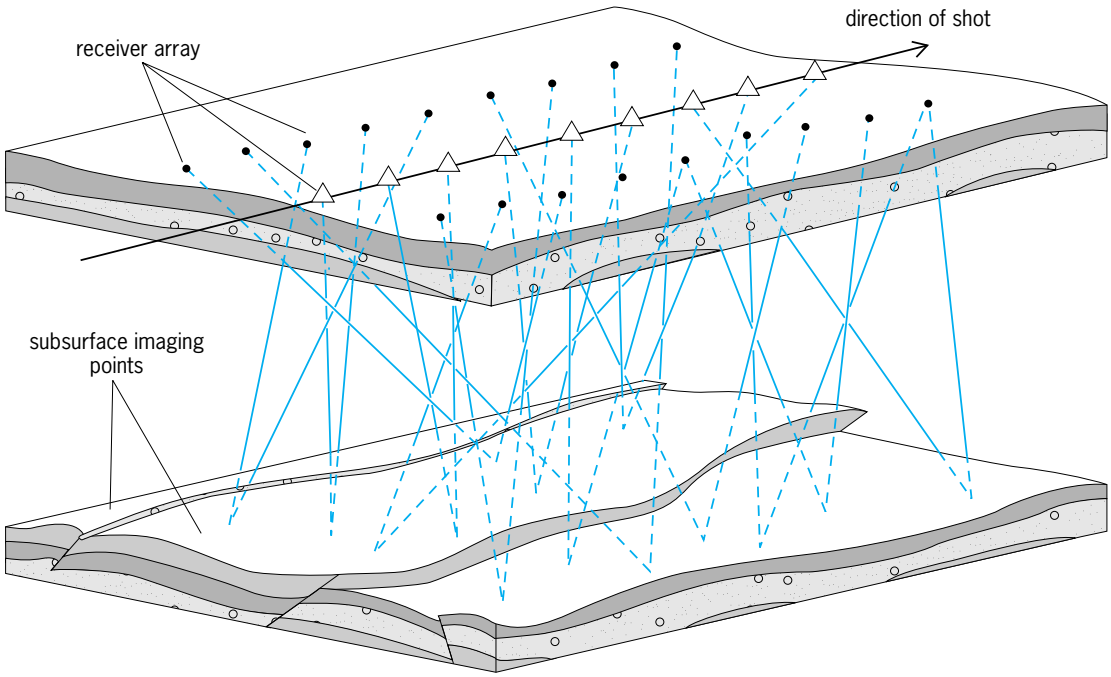

Fig. 1. Layout for a three-dimensional seismic survey.

Two- and three-dimensional seismic. Historically, surface seismic acquisition is done by placing sources and receivers along a straight line so that it can be assumed that all the reflection points fall in a two-dimensional plane formed between the line of traverse and the vertical. This is known as two-dimensional seismic.

Three-dimensional seismic is a method of acquiring surface seismic data by placing sources and receivers in an areal pattern (**Fig. 1**). One example of a simple three-dimensional layout is to place the receivers along a line and shoot into these receivers along a path perpendicular to this line. The common midpoints of a three-dimensional survey also form an areal surface. When all these traces or time series are displayed, they form a closely spaced data volume that represents an acoustic image of a three-dimensional subsurface. However, because of the large data set involved, display and processing of a three-dimensional seismic survey requires the use of workstations that include sophisticated computer graphics. *See* COMPUTER GRAPHICS.

A three-dimensional seismic survey provides a more accurate and detailed image of the subsurface. It offers significantly higher signal quality than the two-dimensional data commonly acquires. It also improves both spatial and temporal resolutions. The three-dimensional seismic technique is being applied to exploration and production of oil and gas, accounting for more than half of the seismic activity in the Gulf of Mexico and North Sea.

Three-component seismic. This refers to seismic surveys where either three-component receivers or three-component sources, or both, are being used.

In seismology, there are two modes of body-wave propagation in the subsurface. These waves are labeled primary (P), or compressional, and secondary (S), or shear. In the P wave, the volume of the rock changes, but not the shape, during the passage of the wave in response to alternating compressive and tensional stresses. The S wave, changes shape but not volume in response to the passage of shear stresses. The shear waves move perpendicular to the direction of propagation. The speed at which each body wave travels is a function of the elasticity of the rock.

Conventional seismic typically uses velocity phones in the vertical direction only. They are adequate for P-wave exploration. To record shear waves in three dimensions, however, three-component geophones are required. It is common practice to place the three-component phones orthogonally in the vertical direction and in the directions parallel and perpendicular to the direction of shooting.

If only P-wave sources are used in the survey, then only converted shear waves are recorded by using the three-component phones. Some surveys involve both P-wave and S-wave sources in three orthogonal directions, producing a total of nine seismic sections, one for each source-receiver direction. This is sometimes called elastic seismic or vector seismic. Such a survey contains azimuthal information related to additional constants of elasticity.

Production or reservoir seismology. Production seismology is the application of seismic techniques to problems related to the production and exploitation of petroleum reservoirs. The internal heterogeneity of a petroleum reservoir has a profound influence on its production performance. The internal structure of a reservoir is complex and generally poorly understood. Structural deformations, fractures, lithologic variations, and diagenetic alterations contribute to the creation or destruction of conduits and barriers to fluid flow. In addition, it may be important to

monitor changes in fluid flow or composition during the producing life of the field and in an enhanced oil recovery process. Since production geophysics is the only effective method available that can image the reservoirs under in-place conditions, it has become an active field of applied research aimed at improving descriptions and understanding of reservoirs and their fluid flow behaviors. *See* PETROLEUM ENHANCED RECOVERY; PETROLEUM RESERVOIR ENGINEERING.

The use of seismic waves to image detailed rock and reservoir properties is hampered by limited resolution. As the elastic seismic waves propagate through the Earth, they attenuate because of spherical divergence. Their higher-frequency components, however, are even more attenuated because of scattering and absorption, resulting in a drastic decrease in resolution. For this reason the majority of successful cases in production seismology have been limited to close-spacing imaging such as is found in shallow reservoirs or reservoirs very close to the wellbore.

By placing acoustic sources at appropriate distances from the receivers in one or more boreholes, higher-frequency seismic signals can be propagated further in the subsurface, resulting in a greater improvement in resolution. Borehole geophysics makes possible the mapping of reservoir heterogeneity, monitoring of secondary or tertiary recovery, and their use as proximity indicators to possible oil and gas accumulations.

Vertical seismic profiling. This is a measurement procedure in which a seismic signal generated at the surface of the Earth is recorded by geophones secured at various depths to the wall of a drilled well. When the source is directly at the top of the well, it is called zero-offset vertical seismic profiling; otherwise it is known as offset vertical seismic profiling (**Fig. 2***a*).

A primary result of the vertical seismic profiling survey is the time-to-depth relationship from which interval velocities as a function of depth can be determined. A zero-offset vertical seismic profiling survey provides a link between sonic log and surface seismic information in two important ways. First, it calibrates the integrated sonic log times to the vertical seismic profiling times, which are directly related to the seismic times. Synthetic seismograms generated by using calibrated sonic logs match the surface seismic data better. Second, the zero-offset vertical seismic profiling links seismic reflection character in time to the well information in depth. It helps to correlate the seismic response to the lithologic interfaces. Vertical seismic profiling also helps interpreters to understand and resolve some of the ambiguities and uncertainties associated with processing and interpretation of surface seismic data. *See* WELL LOGGING.

Offset vertical seismic profiling can also be used as a predictive tool for exploration. It can image the reservoirs ahead of a drill bit, investigate reservoir conditions several hundred meters around the wellbore, or monitor the propagation of a steam front in a thermal recovery process.

Reverse vertical seismic profiling. This is a reverse configuration where the source is placed at a given depth in a borehole and the receivers are deployed in different patterns at the surface (Fig. 2*b*). Although it is a more time-efficient and cost-effective method of downhole imaging. It requires a powerful source to be effective.

Cross-well tomography. Cross-well seismology is a method of seismic imaging where both the sources and receivers are placed in separate boreholes (Fig. 2*c*). By placing both sources and receivers at depths below the highly attenuating near-surface materials such as the weathered layers, seismic waves in frequencies up to several kilohertz can be propagated, resulting in a great improvement of resolving power. The limiting factor is the amount of acoustic energy that can be generated from downhole sources without risking the safety of the wellbore. This in turn limits the distance between the wells at which surveys can be safely conducted.

In a typical cross-well seismic survey between two wells, travel time and amplitude measurements for each source-receiver location across the wells are made. As each ray of transmissive seismic wave

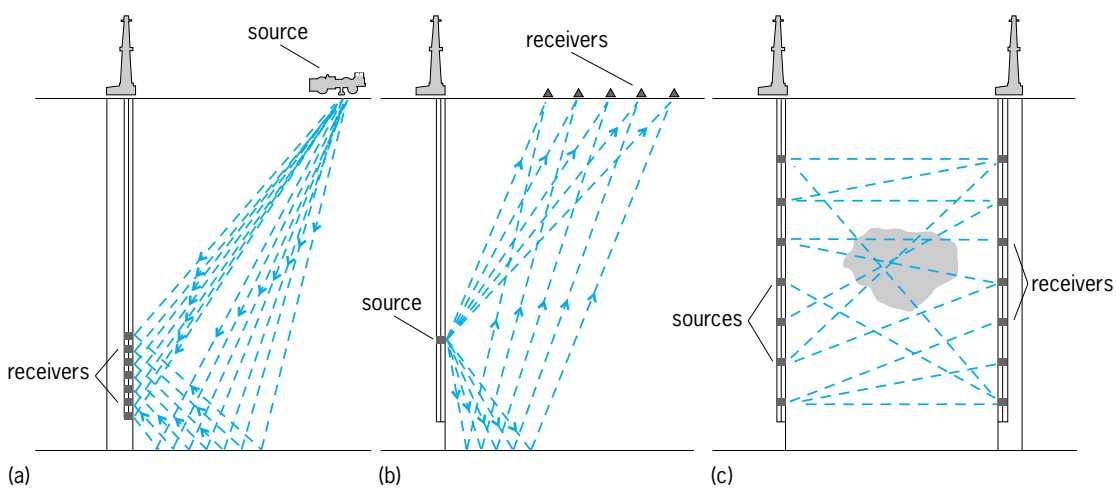

Fig. 2. Borehole seismic geometry. (*a*) Offset vertical seismic profiling. (*b*) Reverse vertical seismic profiling. (*c*) Cross-well tomography.

passes through the reservoir under investigation, it is modified according to the elastic properties of the reservoir. Tomography is an inverse processing technique that is used to reconstruct a two-dimensional image of these reservoir properties. Tomographic reconstruction techniques have been developed and applied in medical science for more than two decades. Conventional tomography assumes straight ray paths. When the effect of diffraction and scattering are considered in the mathematical model, it is called diffraction tomography. When reflections from different strata are considered, it is called reflection tomography. Cross-well seismic tomography has been successfully used to estimate velocities between wells, mostly for the purposes of monitoring enhanced oil recovery. *See* COMPUTERIZED TOMOGRAPHY.

Single-well seismic imaging. This is a technique where both seismic source and receiver are deployed in a single wellbore. The power of the source and the separation between source and receiver determine the depth of investigation.

Real-time seismic monitoring. This seismic technique is used to monitor the change in the conditions of a reservoir as it is producing through time by repeated surveys. It has been demonstrated to work most effectively in heavy-oil thermal recovery projects where steam front progressions have been successfully monitored by using repeated three-dimensional or cross-well seismic surveys. *See* GEOPHYSICAL EXPLORATION; PETROLEUM GEOLOGY; SEISMOLOGY.

Tai. P. Ng

Bibliography. J. K. Costain and C. Coruh, *Basic Theory in Reflection Seismology (Handbook of Geophysical Exploration Seismic Exploration)*, 2004; R. E. Sheriff (ed.), *Reservoir Geophysics*, 1992; M. A. Slawinski, *Seismic Waves and Rays in Elastic Media (Handbook of Geophysical Exploration Seismic Exploration)*, 2003; O. Yilmaz, *Seismic Data Processing*, 1988.

Seismic risk

The probability that social or economic consequences of earthquakes will equal or exceed specified values at a site, at several sites, or in an area, during a specified exposure time. Historically the term seismic risk has been used to describe an assortment of earthquake effects that range from ground shaking, surface faulting, and earthquake-induced landsliding to economic loss and casualties. As more quantitative methods for estimating the effects of earthquakes have been developed, terminology has become more precise. Although the term seismic risk is still sometimes used in a general sense to mean the potential for both the occurrence of natural phenomena and the economic and life loss associated with earthquakes, it is useful to differentiate between the concepts of seismic hazard and seismic risk. Seismic hazard may be defined as any physical phenomena (for example, ground shaking or ground failure) that are associated with an earthquake and that may produce adverse effects on human activities.

The exposure time is the time period of interest for seismic hazard or risk calculations. In practical applications, the exposure time may be considered to be the design lifetime of a building or the length of time over which it is of interest to estimate numbers of casualties. For example, the risk that a certain number of casualties will occur is the probability associated with all combinations of seismic hazards that may result from the earthquakes and all possible numbers of casualties resulting from those hazards. A general expression for the total risk is given by the equation below, in which $P(R_i)$ is the probability that the state

$$P(R_i) = \sum_{\text{all } j} P(R_i|S_j)P(S_j)$$

of the system is i, S_j means that the seismic hazard is level j, $P(S_j)$ is the probability that seismic hazard is level j, and $P(R_i|S_j)$ is the probability that the behavior state of the system will be R_i, given that seismic input, S_j, takes place. Thus, evaluation of seismic risk requires that the seismic hazard be specified.

Seismic hazard analysis. A seismic hazard can be any one of various physical phenomena that result either from surface faulting during shallow earthquakes or from the ground shaking resulting from an earthquake.

Physical phenomena. Surface faulting or rupturing of the Earth's surface usually occurs only in shallow earthquakes (12 mi or 20 km depth or less), but is not observed in all shallow earthquakes. The relative displacement of the two sides of a fault and the fault length both increase approximately exponentially with earthquake magnitude. The relative displacement between the two sides of a fault during great earthquakes may be as much as 40 to 90 ft (20 to 30 m), and the fault length in a single event may be 180 to 240 mi (300 to 400 km).

Landslides may be triggered by earthquake ground shaking, and under certain conditions can be catastrophic. The magnitude 7.4 earthquake that occurred off the coast of Peru in 1970 triggered a landslide above the village of Yungay in the Peruvian Andes, which resulted in the death of more than 15,000 people. Ground shaking also may cause a complex type of ground failure known as liquefaction, that is, the transformation of a granular material from a solid state into a liquefied state as a consequence of increased pore-water pressure. Liquefaction is known to cause widespread damage when concomitant to large earthquakes, as has been exemplified at Puerto Montt, Chile (1960), Niigata, Japan (1964), Anchorage, Alaska (1964), San Fernando, California (1971), and Caucete, Argentina (1977). Even moderate levels of ground shaking may result in liquefaction and subsequent ground failure. *See* LANDSLIDE.

Earthquakes may also result in tsunamis, which are water waves with periods of about 5 to 60 min that are believed to be caused by underwater tectonic displacements. Tsunamis may also be caused by landslides and subaqueous slides that may or may

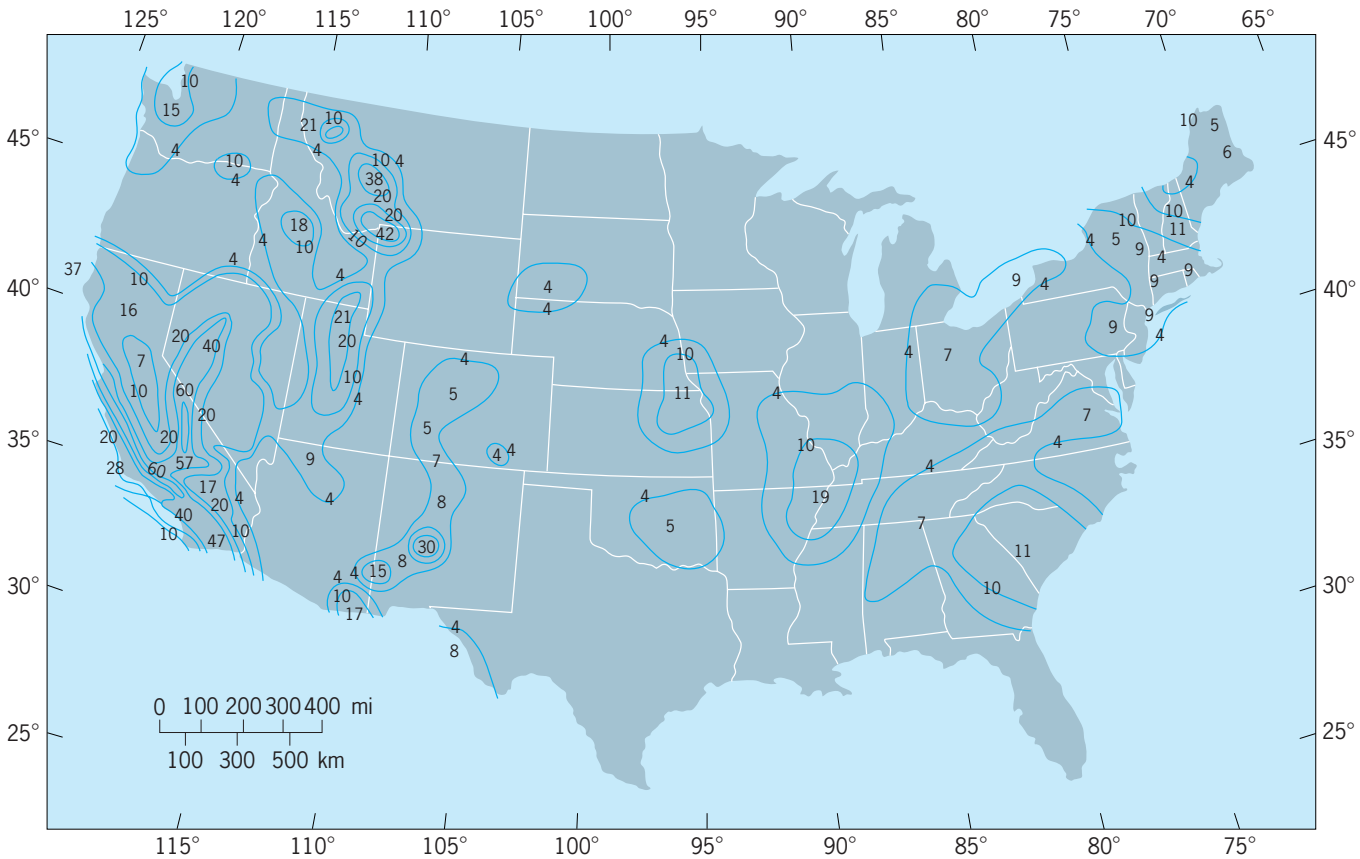

Acceleration hazard map of the United States. The numbers represent the horizontal acceleration in rock with a 90% probability of not being exceeded in 50 years. The acceleration is expressed as a percentage of the acceleration of gravity (≈10 m/s^2 or 32 ft/s^2). Therefore the numbers represent units of 10 cm/s^2 (4 in./s^2). Other numbers represent the highest acceleration attained within the contours. (*After S. T. Algermissen and D. M. Perkins, A Probabilistic Estimate of Maximum Acceleration in Rock in the Contiguous United States, U. S. Geol. Surv. Open-File Rep. 76–416, 1976*)

not have been triggered by an earthquake. Tsunamis have caused extensive property damage and life loss, particularly on Pacific islands and in coastal areas on the periphery of the Pacific Ocean, for example, those associated with the May 23, 1960, Chilean earthquake and the March 27, 1964, Alaskan earthquakes. *See* TSUNAMI.

Maps. Until about 1965, seismic hazards were estimated almost exclusively by use of deterministic methods. A number of seismic zoning maps have been prepared for the United States which provide estimates of earthquake effects. These have been compiled on the basis of a consideration of the regional geologic structure and its relationship to the size and spatial distribution of historic earthquake activity and of the resulting earthquake effects, described in degrees of intensity.

An intensity scale is a descriptive scale of earthquake effects arranged in degrees of increasing severity. The intensity scale most commonly used in the United States is the Modified Mercalli Intensity Scale (1931). This scale has 12 degrees, with I signifying earthquakes not felt or felt only rarely, under especially favorable circumstances, and XII signifying total destruction. Degrees I–V of the scale are based mostly on human reaction and perception of ground shaking; degrees VI–IX are based principally on the severity of damage to structures; and degrees X–XII are based mainly on geologic effects.

Deterministic hazard maps, which outline areas of potential fault rupture, landsliding, and liquefaction, have also been prepared.

Probabilistic models. Since 1965, interest in the probabilistic treatment of seismic hazard has increased. The most widely used (and perhaps the simplest) model assumes that earthquakes are exponentially distributed with regard to magnitude and that they follow a Poisson arrival process in time. The Poisson assumption is not in accord with concepts of the nature of earthquake occurrence, but it is suitable for the engineering evaluation of earthquake hazards. The occurrence of moderate to large earthquakes can closely approximate a Poisson process, and these earthquakes are the most important events in hazard and risk analyses. In addition, a suitable time-dependent model for large-earthquake occurrence is difficult to specify because of insufficient data. Bayesian statistics have been used with some success to model the earthquake process, but specification of prior probabilities will present difficulties until the earthquake mechanism is more clearly understood. A time history of ground motion resulting from an earthquake at any particular location of interest completely describes the ground-shaking hazard, but this

information is rarely available. Seismic intensity has been used frequently as a mapping parameter for seismic hazard, but because of its descriptive nature, it is somewhat unsuitable for formulating engineering specifications for the design of earthquake-resistant structures. The seismic ground-shaking hazard has been specified also by using peak ground acceleration, velocity, or displacement as a mapping parameter. Single-parameter descriptions of ground motion are not entirely satisfactory for the earthquake-resistant design of buildings, because the properties of ground motion that cause building damage cannot be described entirely by a single parameter. A probabilistic acceleration map of the United States is shown in the **illustration**. It has been used by the Applied Technology Council in the construction of two maps of the United States: one for effective peak acceleration and another for effective peak velocity. These two maps are used in constructing generalized spectra for ground motion on different kinds of superficial materials. These spectral shapes are considered to be the descriptions of ground motion for the determination of suitable earthquake-resistant designs for structures. The ground motions specified have an approximately 10% probability of being exceeded at least once in 50 years.

Risk studies. Seismic-risk analysis has been defined here as it involves probability, but risk studies can be deterministic or probabilistic, depending upon the requirements of the particular study. Practical techniques for probabilistic risk analysis in the earthquake-resistant design of structures have become highly developed. Seismic-risk estimates conducted on a regional or national basis for disaster preparedness and for economic studies have been largely deterministic in nature. Risk studies provide the framework for the evaluation of acceptable risks, those judged appropriate for determination of structure design or development of social or economic policy. Such analyses on a regional and national scale have yielded significant results. Studies have predicted that about 70% of the earthquake damage in the United States will occur in the western regions, even though this area contains only about 50% of the construction in the country.

While large earthquakes cause serious losses, the potential for cumulative losses from moderate earthquakes as least as great as those from individual large earthquakes exists because of their much more frequent occurrence. The Richter magnitudes for moderate earthquakes range from 5.5 to 6.5.

Earthquake losses in the United States are expected to become progressively larger with time because of the increase in population and the dispersion of the population. The difficulties in estimating seismic risk on a regional and national level are great because of the problem in estimating the construction characteristics, spatial distribution and value of structures, and their probable resistance to earthquake ground shaking. Seismic risk also is involved with the nature and importance of secondary economic losses resulting from an earthquake, such as those associated with suspension of business activities during building repair or restoration of utilities, increases in repair costs because of material shortages, and lack of skilled workers. The difficulty of evaluating total earthquake loss has prevented accurate and complete estimates of earthquake tolls. *See* EARTHQUAKE; RISK ASSESSMENT AND MANAGEMENT.

S. T. Algermissen

Bibliography. S. T. Algermissen and D. M. Perkins, *A Probabilistic Estimate of Maximum Acceleration in Rock in the Contiguous United States*, U.S. Geol. Surv. Open-File Rep. 76–416, 1976; S. Balassanian, A. Cisternas, and M. Melkumyan (eds.), *Earthquake Hazard and Seismic Risk Reduction: Advances in Natural and Technological Hazards Research*, 2000; C. A. Kircher and A. K. Chopra (eds.), *Seismic Engineering: Research and Practice*, 1989; F. Mulargia and R. J. Geller (eds.), *Earthquake Science and Seismic Risk Reduction,* NATO Science Series: IV: *Earth and Environmental Science*, 2004.

Seismic stratigraphy

Determination of the nature of sedimentary rocks and the fluids they contain from analysis of seismic data. Prior to the 1970s, seismic data were used only to map the structure of the Earth. Since then, improvements in seismic data quality have revealed reflection patterns that indicate the environment in which the sediments were deposited, and as a result the nature of the rocks. *See* SEQUENCE STRATIGRAPHY.

Further improvements in computers, data quality, data acquisition, and seismic theory have expanded the stratigraphic information that can be obtained from the patterns seen in seismic data. The better understanding of seismic stratigraphy has become important in both hydrocarbon exploration and production. *See* SEISMIC EXPLORATION FOR OIL AND GAS.

Sedimentation and seismic reflections. Seismic energy is partially reflected where changes occur in acoustic impedance, the product of density and velocity. At perpendicular incidence, reflectivity, R, is given by Eq. (1), where p and V are density and velocity, respectively, and the subscripts 1 and 2 indicate that these properties are measured above and below an interface where the values change. The fraction of the energy reflected is R^2. One expects density and velocity to change where rock type changes, and reflections to be associated with massive changes in rock type. However, stratification in sedimentary rocks is usually much smaller than a seismic wavelength, so that most reflections are the interference composites of many component reflections.

$$R = \frac{p_2V_2 - p_1V_1}{p_2V_2 + p_1V_1} \tag{1}$$

Deposition is usually episodic rather than smoothly continuous. On a macroscale, periods of land subsidence are interrupted by uplift events, or rivers shift course and begin depositing sediments in a new area, or similar events shift the locale of deposition. On a microscale, periods of quiet

water are interrupted by violent storms that rearrange sediments, or by floods which suddenly bring in large amounts of new material, or by similar events which rearrange sediments over a wide area. The microevents are associated with stratal surfaces; the macroevents often result in a geographic shift in the region of deposition, and are associated with unconformities. *See* ACOUSTIC IMPEDANCE.

Fairly large changes in density and velocity occur within the same rock type. Such changes are associated with changes in porosity or composition, such as changes in the sorting within a sand or changes in the type and degree of cementation. At least locally, such changes generally follow stratal surfaces which are relict depositional surfaces (isochronal surfaces); that is, they were the surface of the solid earth at some point in time. Stratal surfaces generally parallel each other; thus the component reflections do not change very much in short distances, do interfere in the same way on nearby seismic traces, and do produce coherent reflections. Consequently the reflections seen on seismic data generally show the attitude of isochronal surfaces.

The shifts in the geography of deposition associated with macroevents often result in the new depositional unit lying unconformably on older rocks. Usually the base of the unit is an unconformity associated with a hiatus in deposition, and also often is an erosional surface. The next macroevent which marks the end of the depositional unit may leave the top of the unit exposed to erosion and the next unit may be deposited elsewhere, so that the top is also apt to be an unconformity. The sediments above and below the unconformity are apt to differ, and erosion on the unconformities is apt to have changed the nature of the rock. Thus the acoustic impedance is apt to change significantly, and the unconformity will probably be a distinctive reflector. Often unconformities are responsible for the strongest reflections.

Thus seismic reflections can be associated with isochronal surfaces and unconformities (and fluid contact surfaces). The reflections associated with unconformities permit defining seismic-sequence units which are associated with depositional units, and those associated with isochronal surfaces permit determining the setting within the depositional unit, which may be indicative of the depositional environment. *See* SEDIMENTOLOGY; UNCONFORMITY.

Seismic-sequence analysis. The first step is to separate seismic-sequence units, also called seismic-facies units. This is usually done by mapping unconformities (**Fig. 1**) where they are shown by angularity. Angularity below an unconformity may be produced by erosion at an angle across the former bedding surfaces or by toplap (offlap), and angularity above an unconformity may be produced by onlap or downlap, the latter distinction being based on geometry. The unconformities are then followed along reflections from the points where they cannot be so identified, advantage being taken of the fact that the unconformity reflection is often relatively strong. The procedure often followed is to mark angularities in reflections by small arrows before drawing in the boundaries.

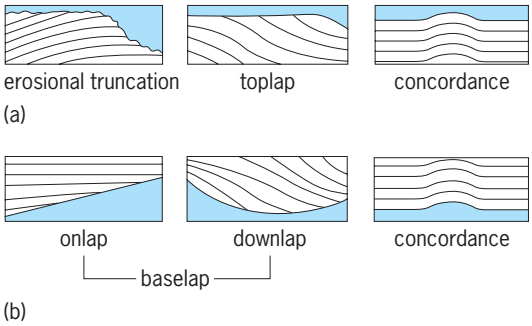

Fig. 1. Relation of strata to boundaries of depositional sequences. (*a*) At the upper boundary of a sequence, strata may terminate mainly as a result of erosion, toplap, or offlap (initially inclined strata may terminate mainly as a result of nondeposition) or top concordance (strata may be parallel to the boundary). (*b*) At the lower boundary, strata may terminate progressively against an initially inclined surface by onlap or downlap; strata may also parallel the lower boundary (base concordance). (*After C. E. Payton, ed., Seismic Stratigraphy: Applications to Hydrocarbon Exploration, Amer. Ass. Petrol. Geol. Mem. 26, 1977*)

Seismic-facies units are three-dimensional, and many of the conclusions from them are based on their three-dimensional shape. The appearance on seismic lines in the dip and strike directions is often very different. For example, a fan-shaped unit might show a progradational pattern in the dip direction and discontinuous, overlapping arcuate reflections in the strike direction (**Fig. 2**). *See* FACIES (GEOLOGY).

Reflection-character analysis. Often information from boreholes suggests that a particular interval may change nearby in a manner which increases its likelihood to contain hydrocarbon accumulations. Lateral changes in the wave shape of individual reflection events may suggest where the stratigraphic changes or hydrocarbon accumulations may be located. *See* DRILLING AND BORING, GEOTECHNICAL.

Where sufficient information is available to develop a reliable model, expected changes are postulated and their effects are calculated and compared to

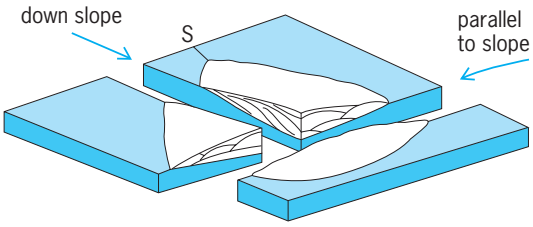

Fig. 2. Diagram of a fan-shaped depositional unit. The unit was fed from an entrenched source (S) and deposited on the continental slope or basin floor. The pattern in the downdip direction is a progradational one of successive units being deposited seaward on earlier units, whereas the pattern parallel to the slope shows the deposition jumping from one lobe to another to produce an overlapping cusp pattern.

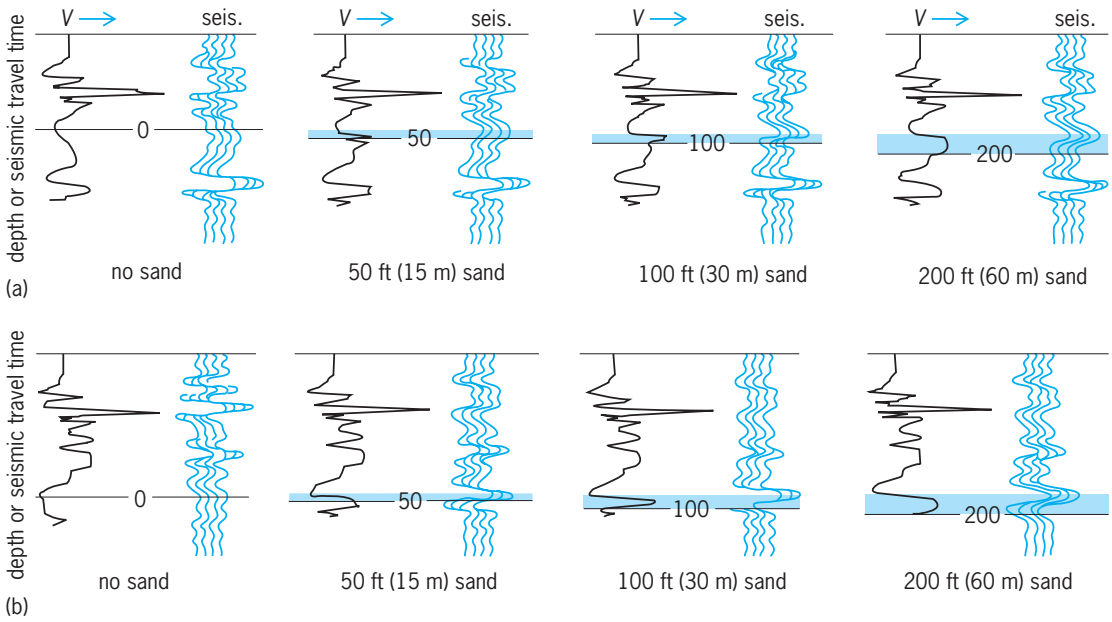

Fig. 3. Portions of sonic logs (V) and seismic traces (seis.) for four thicknesses of sand units; the sands are shaded. (*a*) Variable lower Wade. (*b*) Variable Medrano. The appearance of the sands produces a significant change in seismic wave shape, but the seismic wave shape does not change significantly as the sand thickens from 50 to 200 ft (15 to 60 m). (*After C. E. Payton, ed., Seismic Stratigraphy: Applications to Hydrocarbon Exploration, Amer. Ass. Petrol. Geol. Mem. 26, 1977*)

observed seismic data. The procedure is called synthetic seismogram manufacture; it usually involves calculating seismic data based on sonic and density logs from boreholes, sometimes based on a model derived in some other way. The sonic and density data are then changed in the manner expected for a postulated stratigraphic change, and if the synthetic seismogram matches the actual seismic data sufficiently well, it implies that the changes in earth layering are similar to those in the model. While the solution is nonunique, it gives an interpreter a feel for the relation between stratigraphic changes and their evidences in the seismic data, and is often very useful in actual practice. For example, interposing a possible reservoir sand into the middle of a shale (**Fig. 3**) may result in a significant reflection even where the sand thickness is only one-twentieth of a wavelength, but the reflection character may be insensitive to the thickness of the sand. As another example, the amplitude of a reflection may be markedly different where the pore spaces in the sand are filled with gas or liquid hydrocarbons rather than salt water, but only a minor amount of gas in salt water may produce equivalent effects.

Measurements made on seismic data such as amplitude, frequency, apparent polarity, and velocity are useful in suggesting where stratigraphic or structural features or hydrocarbon accumulations may be located. The general technique is to display attribute measurements, often as a color-coded overlay, along with the seismic section, and then examine the behavior of the attributes laterally along the reflection events. Constancy of attributes implies that there is no significant change in the rocks, a gradual change is generally interpreted as having a stratigraphic

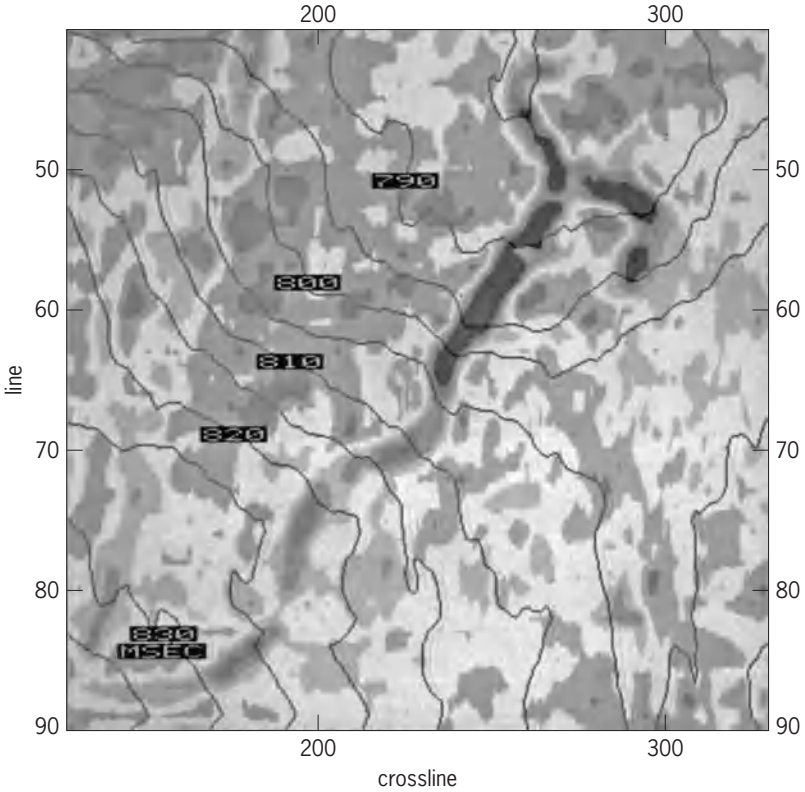

Fig. 4. Section through a 3D seismic volume in the Gulf of Mexico, showing a channel. The thin lines are arrival-time contours along a reflection event. They show the present-day relief on what was once the surface of the Earth. The different color intensities indicate different fluids filling the pore spaces. (*From A. Brown, 2004, courtesy of Chevron U.S.A. Inc.*)

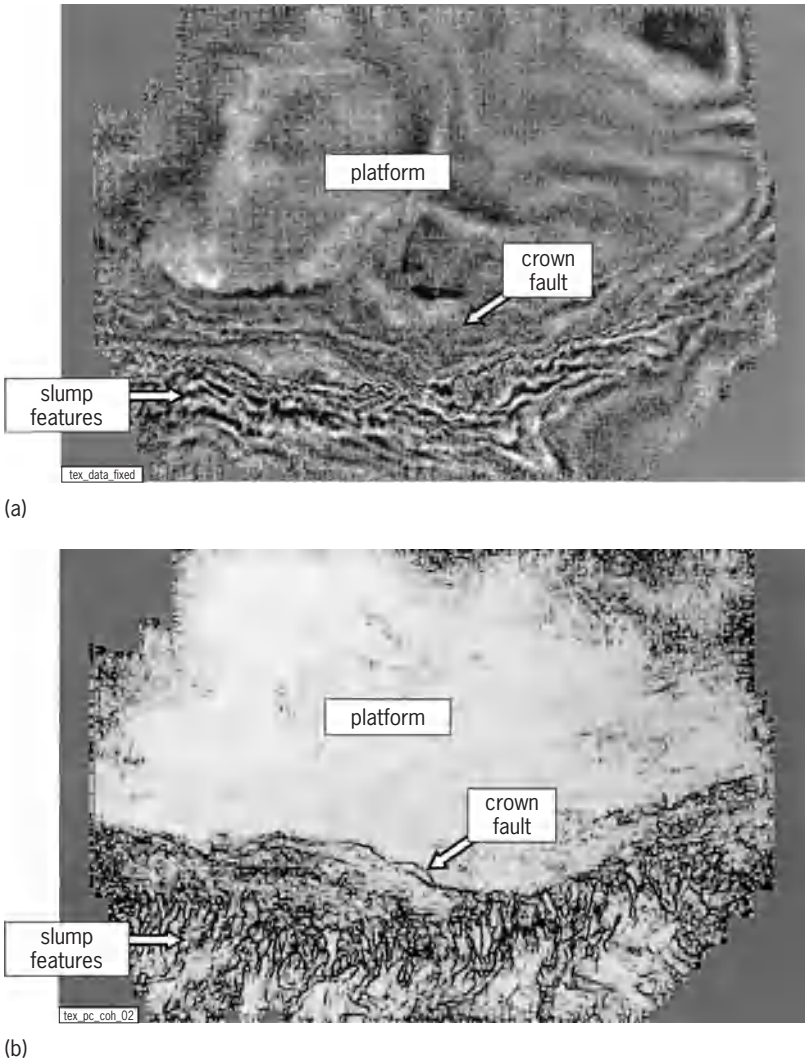

Fig. 5. Horizontal time slice at 900 ms through a 3D seismic volume in the Delaware Basin of New Mexico. The northern portion cuts through a relatively featureless carbonate platform. The edge of the platform has begun to break and slump as indicated by the crown fault. The southern part of the survey shows a suite of slumps from higher above the platform. Eolean sands in the slumping formation make excellent but compartmentalized reservoirs. (*a*) Seismic data. (*b*) Coherence volume. Coherence is an attribute that emphasizes incoherencies in the seismic data.

cause such as change in bed thickness or facies, and a sudden change may indicate a fault or a hydrocarbon accumulation. Attribute measurements and display may aid in defining the limits of hydrocarbon accumulations so that field development may be planned in a more economical and efficient manner.

The reflection coefficient equation (2) can be

$$p_2V_2 = P_1V_1\frac{1+R}{1-R} \qquad (2)$$

solved for the acoustic impedance. The acoustic impedance below an interface can thus be determined from that above the interface if noise can be removed from the seismic data sufficiently that the data represent the Earth's reflectivity. The result is a synthetic acoustic-impedance log (saile) or seismic log. The calculation process requires information not present in the seismic data, and additional velocity data usually provide the necessary supplemental information. For some purposes a seismic log display may be easier to interpret than the seismic data from which it is derived. The major problems with seismic logs usually involve the adequacy of isolating earth reflectivity effects from noise and equivalent wavelet shape, and there is controversy over their use. *See* GEOPHYSICAL EXPLORATION; SEISMOGRAPHIC INSTRUMENTATION; SEISMOLOGY; STRATIGRAPHY; WELL LOGGING.

Three-dimensional seismic methods. Whereas seismic stratigraphy was developed mainly by interpreting the vertical patterns that improved the vertical resolution revealed more clearly, the biggest single improvement in data quality has been in three-dimensional (3D) seismic methods that reveal horizontal (areal) patterns. The horizontal shape of features often reveals distinctive characteristics, such as the fan shape of delta lobes and the sinuosity that characterizes meandering streams (**Fig. 4**). Displays of quantities calculated from data measurements (attributes) often help us to see aspects that are otherwise difficult to see (**Figs. 5** and **6**). A multitude of different attributes are used to emphasize different properties. Having 3D data permits displaying attributes over slices through the data volume not only vertically and horizontally, but also along warped surfaces that follow seismic reflections. Some attributes require prior interpretation (such as picking of reflections, faults, and so on).

The nature of the fluid in a rock's pore spaces alters rock properties, especially velocity and density, which change the reflectivity at interfaces. Changes in reflectivity with the angle of incidence because of the conversion of pressure waves (P waves) into shear waves (S waves) are obscured by stacking. Prestack amplitude versus offset (AVO) analysis consists of examining the variation in amplitude with source-geophone distance (offset) and with incidence angle. Because the pore fluid affects P waves and S waves differently, AVO analysis often permits identifying the pore fluid, which has great economic value. Attribute displays that show where pore fluid changes occur help to locate where the hydrocarbons are and where stratigraphic features and faults affect fluid flow paths. Changes as hydrocarbons are produced (time-lapse or 4D seismic) show the movement of the fluids and have considerable economic value in managing production.

Multicomponent recording can provide S-wave data, in addition to P-wave data. Because P and S waves depend on different elastic properties, having both types of data permits a more definitive identification of the nature of the rocks and the fluids they contain. In areas where reservoir gas has permeated the overlying rocks so much that P-wave data is of very poor quality, the conversion of P waves to S waves upon reflection allows the mapping of the S waves, since they are relatively unaffected by the intruded gas. Multicomponent recording also helps determine the rock's anisotropic character, which helps reveal fractures that affect fluid flow. *See* PETROLEUM GEOLOGY; PETROLEUM RESERVOIR ENGINEERING.

Robert E. Sheriff

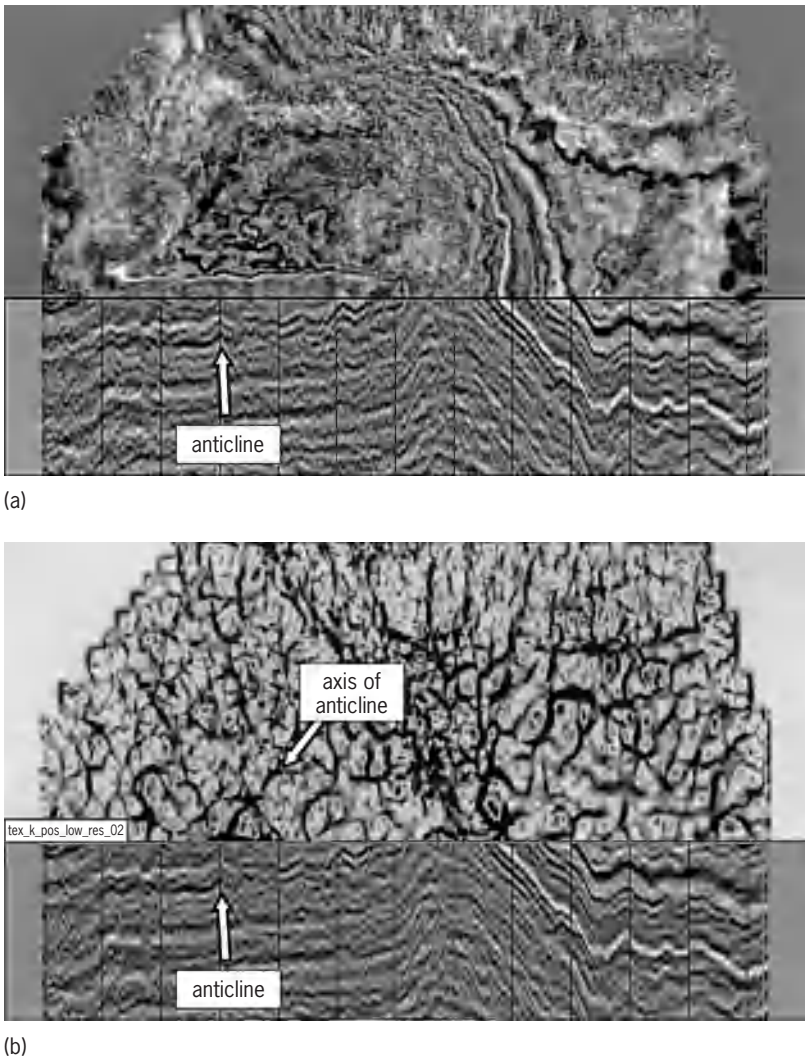

Fig. 6. Folded slice in the Delaware Basin, New Mexico. The northern part is a time slice at 1700 ms and the lower part is a vertical section at he southern edge of the time slice. (*a*) Seismic data. (*b*) Curvature attribute over the time-slice portion. Curvature is an attribute that shows complex folding. It is calculated for very small surfaces to fit the immediate surroundings of each point. Here it shows the orientation and extent of a small anticline. Curvature calculations do not require prior interpretation of the data.

Bibliography. A. R. Brown, *Interpretation of Three-Dimensional Seismic Data*, 6th ed., 2004; D. Emery and K. Myers (eds.), *Sequence Stratigraphy*, 1996; R. G. Loucks and J. F. Sarg, *Carbonate Sequence Stratigraphy*, 1993; R. E. Sheriff (ed.), *Reservoir Geophysics*, 1992; R. E. Sheriff and L. P. Geldart, *Exploration Seismology*, 2d ed., 1995; J. C. van Wagoner and G. T. Bertran (eds.), *Sequence Stratigraphy of Foreland Basin Deposits*, 1995; P. Weimer and H. W. Posamentier (eds.), *Siliciclastic Sequence Stratigraphy: Recent Developments and Applications*, 1994; P. Weimer and R. M. Slatt, *Petroleum Systems Of Deepwater Settings*, 2004.

Seismographic instrumentation

Various devices or systems of devices for measuring movement in the Earth. Ground motion is generally the result of passing seismic waves, gravitational tides, atmospheric processes, and tectonic processes. Seismographic instrumentation typically consists of a sensing element (seismometer), a signal-conditioning element or elements (galvanometer, mechanical or electronic amplifier, filters, analog-to-digital conversion circuitry, telemetry, and so on), and a recording element (analog visible or direct, frequency modulation, or digital magnetic tape or disk). Seismographs are used for earthquake studies, investigations of the Earth's gravity field, nuclear explosion monitoring, petroleum exploration, and industrial vibration measurement.

Design requirement. Seismographic instruments may be required to measure ground motions accurately over a range approaching 12 orders of magnitude, from as small as 10^{-11} m (the Earth's background noise level in the 2 mHz to 100 Hz band at very quiet sites) to as large as several meters (strong ground motion in the near-field of a very large earthquake). The instruments may be required to measure frequencies as low as $\sim 10^{-5}$ Hz (the semidiurnal gravitational tides), and even lower frequencies that are involved in tectonic strain monitoring, to as high as $\sim 10^{4}$ Hz (as observed from acoustic

emissions from rock failures in mines at distances of a few meters). Seismic waves from earthquakes are observed in the bandwidth of $\sim 3 \times 10^{-4}$ Hz (the gravest free oscillations of the Earth) to $\sim$200 Hz (a local earthquake recorded by a seismometer installed in a borehole 100+ meters deep). In exploration seismology the frequency range of interest is typically 10–1000 Hz. No single instrument can operate over such a large dynamic range and frequency bandwidth. Thus a variety of instrumental designs are seen in seismometry.

Seismometers. The seismometer is the basic sensing element in seismographic instruments, and there are two fundamentally different types: inertial and strain. The inertial seismometer generates an output signal that is proportional to the relative motion between its frame (usually attached to the ground or a point of interest) and an internal inertial reference mass. The strain seismometer (or linear extensometer) generates an output that is proportional to the distance between two points.

In the inertial seismometer (**Fig. 1**), the inertial mass is an element of a damped mechanical oscillator, in either a spring-mass or a pendulum configuration. The restoring force may be either an elastic force (including the inertial elasticity of a piezoelectric device), an electrical spring, or gravity. Damping may be achieved by the use of viscous fluids or by electrical feedback to provide a force that resists the relative velocity of the suspended mass.

Regardless of the design details, the behavior of inertial seismometers is described by the basic second-order differential equation for the harmonic oscillator, where x is the relative motion between the

$$\frac{d^2x}{dt^2} + 2\zeta\omega_n\frac{dx}{dt} + \omega_n^2 x = \frac{d^2z}{dt^2}$$

mass and the frame, ζ is the damping factor ($\zeta = 1$ is critical damping), ω_n is the natural frequency (in radians per second), and z is the frame (ground) motion. The relative displacement, x, or the relative velocity, dx/dt, is sensed by an appropriate transducer. Displacement transducers are typically optical, variable-capacitance, or linear vector differential transformer (LVDT) devices. Velocity transducers invariably used coil-magnet systems until recently.

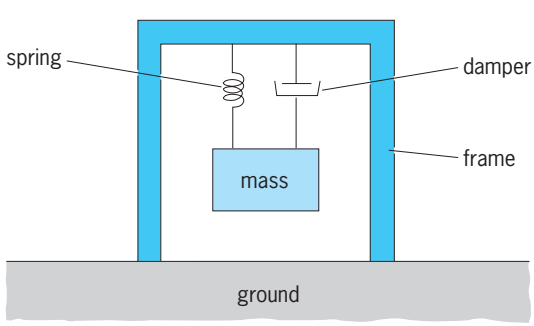

Fig. 1. Principle of the vertical-component inertial seismometer. Motion is measured between the mass and the frame.

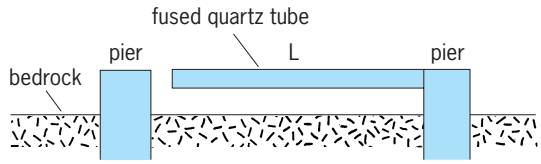

Fig. 2. Principle of the linear extensometer in a strain seismograph. Motion is measured between the left pier and the free end of the quartz tube.

However, a new type of velocity transducer has been developed which uses a molecular electronic transfer cell to detect the motion of an electrolytic fluid. In the case of piezoelectric (lead zirconate titanate, or PZT) seismometers, the voltage generated across the piezoelectric element is proportional to acceleration at frequencies below the resonant frequency (normally very high) of the piezoelectric device. *See* HARMONIC OSCILLATOR; PIEZOELECTRICITY; TRANSDUCER.

The output of a seismometer is a signal, usually electrical, that is proportional to some function of the ground motion, z. It is possible for this signal, over some frequency range, to be directly proportional to the ground displacement (z), the ground velocity (dz/dt), or the ground acceleration (d^2z/dt^2). Within such frequency ranges, seismographs are commonly referred to as displacement seismographs, velocity seismographs, or accelerographs, respectively. That is, the recorded amplitude represents the appropriate ground-motion quantity multiplied by a frequency-independent constant (the magnification, the velocity sensitivity, or the acceleration sensitivity of the seismograph) for motion in the particular frequency range.

A strain seismometer (or linear extensometer) [**Fig. 2**] is capable of providing a continuous precise measurement of the distance between two points, spaced typically 1 m to 1 km (3.3 ft to 0.6 mi) apart. A wide variety of designs have been implemented, using as reference lengths (L) from one point to the vicinity of the other point such standards as fused silica rods or light beams. Rod-based extensometers normally use a capacitance displacement transducer to sense variations in the position of the reference rod end with respect to the other reference point. Light-beam extensometers invariably incorporate some type of interferometer to detect changes in the length of the optical path between the two reference points. Longer paths require evacuated tubes and laser sources to achieve adequate sensitivity and stability in the length change measurement. The frequency range of interest is from zero to no more than $\sim$1 Hz in most systems. *See* INTERFEROMETRY.

The quantity measured in these instruments is the change ΔL in the reference length L. The ratio $\Delta L/L$ gives the horizontal component (typically) of the strain between the two points. With a displacement resolution of 10^{-9} m (3.3×10^{-9} ft), a 100-m (330-ft) extensometer can measure changes in strain of 1 part in 10^{-11}. However, practical limits in environmental controls (temperature, atmospheric pressure, and

mechanical stability) yield long-term stabilities no better than 10^{-7} per year for such instruments.

Secular (long-term) strains are also measured using very long baseline interferometry (VLBI) and geodetic Global Positioning System (GPS) techniques. VLBI uses phase differences in signals from extraterrestrial radio sources (quasars) observed simultaneously at points separated by continental dimensions. Accuracies of a few centimeters in a few thousand kilometers are observed, for a resolution of 10^{-9} in relative distance change between the two remote points. The GPS system is used to observe geodetic strain by absolute position measurement on the Earth's surface. The geodetic deformation is monitored with both continuous (permanent) stations and campaign measurements (temporary deployments at existing benchmarks) over a period of a few years. Accuracies of a few millimeters are attained. *See* GEODESY; SATELLITE NAVIGATION SYSTEMS.

Seismoscopes. A seismoscope is a device that indicates only the occurrence of relatively strong ground shaking and not its time of occurrence or duration. A typical seismoscope inscribes a hodograph of horizontal strong ground motion on a smoked watch glass.

Dilatometers. A dilatometer continuously and precisely measures volumetric strain. The quantity measured is the change ΔV in the reference volume V, and the ratio $\Delta V/V$ gives the volumetric strain. Dilatometers are typically installed in boreholes in competent rock (preferably granite) at a depth of 100–300 m (330–1000 ft).

Tiltmeters. A tiltmeter monitors the relative change in the elevation between two points, usually with respect to a liquid-level surface. The horizontal distance between the reference points may be as little as a few millimeters or as large as several hundred meters. A displacement transducer is typically employed to sense the vertical separation between the two liquid-level surfaces. Environmental effects limit the long-term stability to about 10^{-6} radian.

Tilt at a point can be measured by an inertial seismometer. Tilt of the horizontal surface is indistinguishable inertially from a horizontal acceleration of magnitude $g \sin \theta$, where θ is the angle and g is the acceleration due to gravity. Any seismometer with DC (direct current or zero frequency) response to acceleration (for example, a horizontal pendulum equipped with a displacement transducer) is therefore a tiltmeter with constant tilt sensitivity throughout the frequency range where its output is proportional to acceleration. *See* ACCELEROMETER.

Gravimeters. The gravity meter is just a vertical-component accelerometer, that is, a pendulum sensing ground motion and equipped with a displacement transducer, analogous to the inertial tiltmeter. In its most widely used form, the pendulum is small for portability and ease of thermal stabilization, with a natural frequency of $\sim$0.1 Hz. Transducer technology and environmental control methods allow the best gravimeters to have a repeatable accuracy of $\sim 10^{-8}\ g$. Gravimeters are widely used in geophysical exploration, in the study of earth tides, and in the recording of very low frequency (0.0003–0.01 Hz) seismic waves from earthquakes. *See* EARTH TIDES; EARTHQUAKE; GEOPHYSICAL EXPLORATION; SEISMIC EXPLORATION FOR OIL AND GAS; SEISMOLOGY.

Recording systems. The complete seismograph produces a record of the properly conditioned signal from the seismometer, along with appropriate timing information. The recording system may be a simple as a mechanical stylus scratching a line on a smoke-covered drum in a portable microearthquake seismograph, or as complex as a multichannel computer-controlled system handling 25,000 24-bit digital words per second in a modern seismic reflection survey for petroleum exploration. The range between these extremes includes many special-purpose seismographs, all designed to record ground motion in a particular application.

Deployment. Many methods are used for deploying seismographs, and they depend upon whether the site is temporary or permanent and whether the site is land-based or on the ocean bottom. Seismic station installations are susceptible to several types of noise that contaminate or even mask the desired signals, effectively reducing the operating sensitivity of the instrument. These noise sources are of several types: atmospheric, cultural, ocean current, microseismic, and ambient temperature variation. The atmosphere affects the seismometer through pressure fluctuations directly on the mechanical system and indirectly on the surrounding ground. Cultural and industrial vibrations from human activities can be measured near any population center. In the case of seismographs deployed on the sea floor, ocean currents are a large noise source. Microseisms, the natural background vibrations of the Earth, are generally larger near continental margins than inland—a result of wave action on the coastline. For feedback broadband seismometers, ambient temperature variation is also a significant source of noise.

In the conventional land-based permanent seismographic station, the seismometer is placed upon a stable (usually concrete) pier in an environment as noise-free as possible. Two deployment methods have been devised to mitigate against noise sources. One technique uses seismometers in closely spaced arrays, relying on the coherency of the signal of interest over the array dimensions, and the lack of correlation of the noise components in the seismic wave field. This method is applied successfully in seismic reflection surveying, and in the arrays deployed to monitor from large distance the seismic waves from detonation of underground nuclear explosions. A second noise-reduction method uses borehole or deep mine installations to take advantage of the natural attenuation with depth of much of the noise field due to shallow propagating surface waves, to attenuate the effects of the surface ground motion due to atmospheric pressure fluctuation, and to provide temperature stability.

When deploying a temporary array of portable seismographic stations, as required for aftershock studies or research investigations, the siting criteria also include obtaining permission from the land

owners and concerns about vandalism. The portable seismographs are typically housed in small enclosures that are placed on the surface or buried in the ground a few feet. The stations are generally battery- and solar-powered and have sufficient storage capacity to run unattended for weeks to months. Siting the seismometer on hard rock and away from cultural noise sources is preferable but not always possible. Instrument pools of portable seismographs exist for a wide variety of seismic studies. The Program for Array Seismic Studies of the Continental Lithosphere (PASSCAL) instrumentation center, for example, maintains a large stockpile of portable seismographs that are used by researchers.

In the deployment of an ocean bottom seismograph (OBS), the logistics and installation problems are formidable. However, there is considerable incentive to place seismographs on the sea floor since it represents the largest uninstrumented area of the Earth's surface. A number of specialized methods have been devised for deploying, servicing, and retrieving OBS instrumentation. Borehole installations in the sea floor produce the lowest noise levels. Most OBSs use retrievable instrument capsules with onboard recording and sufficient battery capacity for unattended operation from a few weeks to a year or more. Some permanent OBS stations have been installed along cables that link them to shore and provide telemetry and power. Transoceanic telephone cables, decommissioned as a result of increasing reliance on satellite communications, are being considered for the installation of permanent OBS stations on the sea floor. OBS instrument pools exist for a wide variety of seismic studies. For example, both the Ocean Bottom Seismograph Facility of the Scripps Institution of Oceanography and the Woods Hole Oceanographic Institution Marine Seismology Group maintain a pool of OBS instrumentation for use by researchers. Thomas V. McEvilly; Robert A. Uhrhammer

Bibliography. M. E. Ander, T. Summers, and M. E. Gruchalla, LaCoste & Romberg gravity meter: System analysis and instrumental errors, *Geophysics*, 64:1708–1719, Society of Exploration Geophysicists, 1999; G. Anderson, H. Staudigel, and F. K. Wyatt, A seafloor long-baseline tiltmeter, *J. Geophys. Res. B*, 102:20269–20285, 1997; C. R. Bradley, *Very Low Frequency Seismo-acoustic Noise below the Sea Floor*, Ph.D. thesis, MIT/WHOI Joint Program in Oceanography, 1994; C. Braitenberg, The hydrologic induced strain-tilt signal, a review: Analysis of environmental data for the interpretation of gravity measurements, *Marees Terrestres: Bulletin d'Informations*, 131:10171–10181, Observatoire Royal de Belgique, 1999; W. K. Cloud, Instruments for earthquake investigation: Earthquake investigations in the western United States, 1931–64, *U.S. Coast Geod. Surv. Publ.*, 41-2:5–20, 1965; J. Dewey and P. Byerly, The early history of seismometry (to 1900), *Bull. Seismol. Soc. Amer.*, 59:183–227, 1969; J. Fowler, PASSCAL: A facility for portable seismological instrumentation, *EOS: Trans. Amer. Geophys. Union*, 75:66, 1994; B. S. Melton, Earthquake seismograph development: A modern history, parts 1 and 2, *EOS: Trans. Amer. Geophys. Union*, 62:505–510, 545–548, 1981; G. D. Myren and M. J. S. Johnston, Borehole Dilatometer Installation, Operation and Maintenance at Sites along the San Andreas Fault, California, *USGS Open-File Rep.*, no. 89-0349, 1989; R. A. Stephen et al., The Seafloor Borehole Array Seismic Systems (SEABASS) and VLF ambient noise, *Mar. Geophys. Res.*, 16:243–286, 1996; R. A. Uhrhammer, B. Romanowicz, and W. Karavas, Broadband seismic station installation guidelines, *Seismol. Res. Lett.*, 69:15–26, 1998; E. Wielandt, The leaf-spring seismometer: Design and performance, *Bull. Seismol. Soc. Amer.*, 72:2349–2367, 1982.

Seismology

The study of the shaking of the Earth's interior caused by natural or artificial sources. The revolution in understanding the Earth brought about by the paradigm of plate tectonics in the early 1960s provided a rationale that has since guided an intense period of investigation into the structure of the Earth's crust, mantle, and core. Plate tectonics provided Earth scientists with a kinematic theory for global tectonics that gave context to many of their studies. Ancient regions of crustal deformation could, for instance, be seen as the result of past plate boundary interactions, and the global distribution of earthquakes could be understood as the result of present plate interactions. The extraordinary insights into the nature of the solid Earth provided by kinematic plate tectonics also indicated new avenues in which to pursue basic questions concerning the dynamic processes shaping the Earth. Throughout the period in which plate tectonics was advanced and its basic tenets tested and confirmed, and into the latest phase of inquiry into basic processes, seismology (and particularly seismic imaging) has provided critical observational evidence upon which discoveries have been made and theory has been advanced. The use of the term image to describe the result of seismic investigations is closely linked mathematically to inversion and wavefield extrapolation and applies both to structures within the Earth and to the nature of seismic sources. *See* PLATE TECTONICS.

Theoretical seismology. A seismic source is an energy conversion process that over a short time (generally less than a minute and usually less than 1–10 s) transforms stored potential energy into elastic kinetic energy. This energy then propagates in the form of seismic waves through the Earth until it is converted into heat by internal (molecular) friction. Most earthquakes, for example, release stored elastic energy, nuclear and chemical explosions release the energy stored in nuclear and molecular bonds, and airguns release the pressure differential between a container of compressed air and surrounding water. Large sources, that is, sources that release large amounts of potential energy, can be detected worldwide. Earthquakes above Richter magnitude 5 and explosions above 50 kilotons or so are large enough to be observed globally before the seismic

waves dissipate below modern levels of detection. The largest earthquakes, such as the 1960 event in Chile with an estimated magnitude of 9.6, cause the Earth to reverberate for months afterward until the energy falls below the observation threshold. Small charges of dynamite or small earthquakes are detectable at a distance of a few tens to a few hundreds of kilometers, depending on the type of rock between the explosion and the detector. The smallest of these may be detectable for only a few seconds. *See* EARTHQUAKE.

Seismic vibrations are recorded by instruments known as seismometers that sense the change in the position of the ground (or water pressure) as seismic waves pass underneath. The record of ground motion as a function of time is a seismogram, which may be in either analog or digital form. Advances in computer technology have made analog recording virtually obsolete: most seismograms are recorded digitally, which makes quantitative analysis much more feasible.

Equation of motion. The dissipation of seismic energy in the Earth generally is small enough that the response of the Earth to a seismic disturbance can be approximated by the equation of motion for a disturbance in a perfectly elastic body. This equation holds regardless of the type of source, and is closely related to the acoustic-wave equation governing the propagation of sound in a fluid. Deviations of the elastic body away from equilibrium or rest are resisted by the internal strength of the body. An elastic material can be thought of as a collection of masses connected by springs, with the spring constants governing the response of the material to an externally applied force or stress. More formally, the strain or distortion of an elastic body when subjected to an applied stress is described with a constitutive relation, which is just a generalization of Hooke's law. Instead of just one spring constant, however, a generalized elastic solid needs 21 constants to completely specify its constitutive relation. Fortunately, the Earth is very nearly isotropic; that is, its response to an applied stress is independent of the direction of that stress. This reduces the complete specification of the constitutive relation from 21 to just 2 constants, the Lamé parameters λ and μ. A well-known result is that the equation of motion for an isotropic perfectly elastic solid separates into two equations describing the propagation of purely dilatational (volume changing, curl-free) and purely rotational (no volume changing, divergence-free) disturbances. These propagate with wave speeds α and β [Eqs. (1) and (2)], respectively,

$$\alpha = \sqrt{\lambda + \frac{2\mu}{\rho}} \tag{1}$$

$$\beta = \sqrt{\frac{\mu}{\rho}} \tag{2}$$

where ρ is the density and the Lamé parameters have units of pressure (force/area).

Elastic-wave propagation. These velocities are also known as the compressional or primary (P) and shear or secondary (S) velocities, and the corresponding waves are called P and S waves. The compressional velocity is always faster than the shear velocity; in fact, it can be shown that mechanical stability requires that $\alpha^2 \geq 4/3\beta^2$. In a fluid, $\mu = 0$ and there are no shear waves. The wave equation then reduces to the common acoustic case, and the compressional wave is just the ordinary sound or pressure wave. In the Earth, α can range from a few hundred meters per second in unconsolidated sediments to more than 13.7 km/s (8.2 mi/s) just above the core–mantle boundary. (The fastest spot within the Earth is not its center, but just above the core–mantle boundary in the silicate mantle. The core, being predominantly iron, is in fact relatively slow.) Wave speed β ranges from zero in fluids (ocean, fluid outer core) to about 7.3 km/s (4.4 mi/s) at the core–mantle boundary. *See* HOOKE'S LAW; SPECIAL FUNCTIONS.

A P wave has no curl and thus only causes the material to undergo a volume change with no other distortion. An S wave has no divergence, thus causing no volume change, but right angles embedded in the material are distorted. Explosions are relatively efficient generators of compressional disturbances, but earthquakes generate both compressional and shear waves. Compressional waves, by virtue of the mechanical stability condition, always arrive before shear waves.

Compressional and shear waves can exist in an elastic body irrespective of its boundaries. For this reason, seismic waves traveling with speed α or β are known as body waves. A third type of wave motion is produced if the elastic material is bounded by a free surface. The free-surface boundary conditions help trap energy near the surface, resulting in a boundary or surface wave. This in turn can be of two types. A Rayleigh wave combines both compressional and shear motion and requires only the presence of a boundary to exist. A Love wave is a pure-shear disturbance that can propagate only in the presence of a change in the elastic properties with depth from the free surface. Both are slower than body waves.

Solutions of the elastic-wave equation in which a wave function of a particular shape propagates with a particular speed are known as traveling waves. An important property of traveling waves is their causality; that is, the wave function has no amplitude before the first predicted arrival of energy. The complete seismic wavefield can be constructed by summing up every possible traveling wave, usually a countably infinite set. A traveling-wave formalism is useful especially when it is desirable to concentrate on particular arrivals of energy. An alternative set of solutions to the wave equation can be constructed by considering standing-wave functions. Individual standing waves by definition do not propagate, but rather oscillate throughout the body with a fixed spatial pattern. However, the sum of all possible standing waves must be equivalent to the sum of all possible traveling waves, because they must represent the same total wavefield. Standing waves, also known as free oscillations or normal modes, are useful representations of the wavefield in the frequency domain.

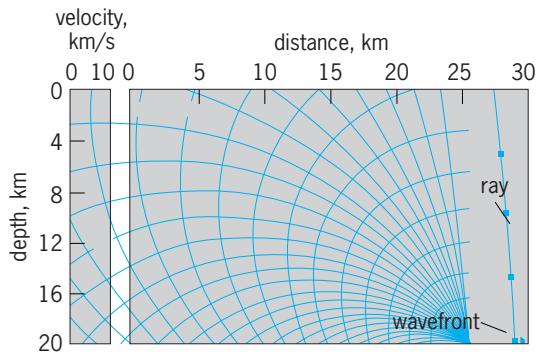

Fig. 1. Relationship between wavefronts and geometrical rays for a constant velocity gradient. 1 km = 0.6 mi.

Ray theory. Traveling-wave or full-wave theory provides the basis for a very useful theoretical abstraction of elastic-wave propagation in terms of the more common notions of wavefronts and their outwardly directed normals, called rays (**Fig. 1**). The validity of the ray treatment can be deduced from full-wave theory by taking infinite frequency approximations that treat the seismic disturbances as pure impulses having essentially no time duration and unit amplitude. Ray theory makes the prediction of certain kinematic quantities such as ray path, travel time, and distance by a simple geometric exercise involving what is essentially a generalization of Snell's law. Ray theory can be developed in the context of an Earth comprising flat-lying layers of uniform velocities; this is a very useful approximation for most problems in crustal seismology and can be extended to spherical geometry for global studies. Expressions will be developed for travel time and distance in both geometries.

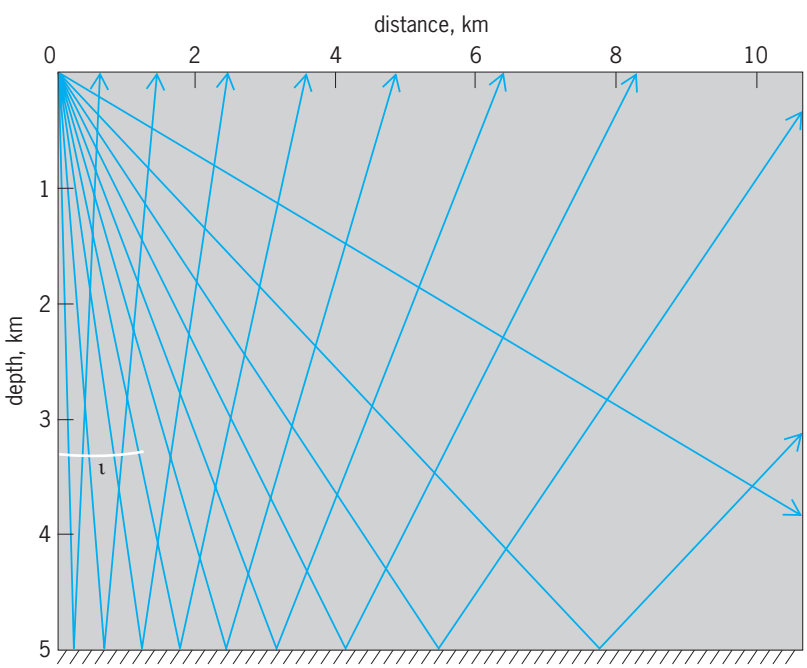

Fig. 2. Diagram showing seismic rays associated with a source at the Earth's surface bouncing off the base of a layer and being returned to the surface. The take-off angle ι for the ray that returns to the surface at a range of about 3.5 km (2.1 mi) is shown. 1 km = 0.6 mi.

Figure 2 is a diagram showing the rays of a seismic wavefront in a single layer expanding away from a source, a single layer. It is apparent from this simple ray diagram that the rays will intersect the base of the layer and be reflected to the surface. Each ray leaves the source at an angle ι, known as the take-off angle or angle of incidence. This angle can be used to characterize each of the rays by Eq. (3), where p is

$$p = \frac{\sin \iota}{\alpha} \tag{3}$$

the ray parameter. This term characterizes uniquely each ray in a system of rays. Its physical meaning is very important; if the term slowness is used to decribe the inverse of velocity, then p is the horizontal component of a ray's slowness. A vertical component q can also be defined, because unlike components of velocity the quantities p and q behave as vectors. This result makes it possible to cast the calculation of travel time and distance as geometric problems. Equation (4) gives the time $T(x)$ required to reach a given distance x from the source, and Eq. (5) gives

$$T^2(x) = T^2(0) + \left(\frac{x}{\alpha}\right)^2 \tag{4}$$

$$X(p) = \frac{2hp}{\sqrt{u^2 - p^2}} \tag{5}$$

the distance reached by a ray of given ray parameter, $X(p)$.

The first term on the right-hand side of Eq. (4) is the travel time of the vertical ray having a take-off angle of 0°, or $p = 0$. Equation (5) employs notation u to describe the slowness of the medium, rather than its velocity. Both Eqs. (4) and (5) are parametrized by layer thickness h and layer velocity α (or slowness u); and these are the unknowns needed to describe the Earth (α depends on the Lamé parameters and the density). Equation (4) is hyperbolic, which is sometimes observed in practice. This indicates that a simple representation of the Earth as a stack of uniform velocity layers might at times be reasonably close to reality.

Kinematic equations have been developed to describe what happens to rays as they impinge on the boundaries between layers. **Figure 3*a*** is a diagram of a single ray propagating in the stack of horizontal layers that define the model Earth. At each interface, part of the ray's energy is reflected, but a portion also passes through into the layer below. The transmitted portion of the ray is refracted; that is, it changes the angle at which it is propagating. The relationship between the incident angle and the refracted angle is exactly the same as that describing the refraction of light between two media of differing refractive index, as shown in Eq. (6). This is the seismic analog

$$\frac{\sin \iota_{\text{incident}}}{\alpha_{\text{upper layer}}} = \frac{\sin \iota_{\text{refracted}}}{\alpha_{\text{lower layer}}} \tag{6}$$

of Snell's law. The left-hand side of Eq. (6) is the ray parameter for the ray in the upper layer, and the right-hand side is the ray parameter in the lower layer. Thus Snell's law for seismic energy is equivalent to stating

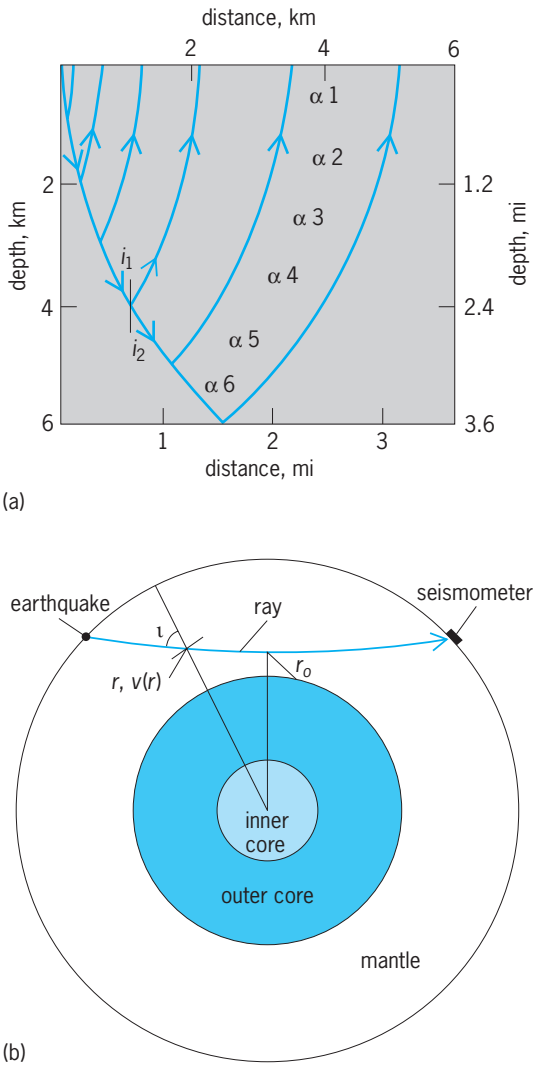

Fig. 3. Seismic ray paths. (*a*) A single ray passing through a multilayered Earth comprising a stack of uniform velocity layers will be reflected from each layer and also be refracted as it passes from one layer into the layer below in a manner that obeys Snell's law. Each ray therefore is considered to give rise to a new system of rays. (*b*) Ray diagram for a cross section of the spherical Earth. At the point labeled *v*(*r*), *r* = radial distance and *v* = velocity. r_0 = radial distance to the turning point. ι is the angle of incidence.

that the ray parameter or the horizontal slowness is conserved on refraction. *See* REFRACTION OF WAVES.

Equation (7) for multilayers describes the distance

$$X_n(p) = 2\sum_{k=1}^{n} \frac{p h_k}{\sqrt{u_k^2 - p^2}} \tag{7}$$

reached by a ray of given p after passing through n layers, each of which is associated with a slowness u_k and thickness h_k before being reflected to the surface. It is the sum of the individual contributions to $X(p)$ from each of the layers. The equation for $T(x)$ cannot be written as a simple sum of hyperbolic contributions, although doing so in some special situations can be a reasonable approximation to the real travel-time behavior. Because of this, complex seismograms are generally analyzed by using a "parametric" form such as $T(x) = T[X(p)]$.

The generalization of these kinematic equations to the case where the medium slowness varies continuously with depth is straightforward. More care must be taken, however, in defining what happens to the ray at its deepest point of penetration into the medium. In Eq. (5), the ray is reflected from the base of the deepest layer. More generally, information is sought for the ray traveling downward, then reversing direction and traveling upward. In a medium with continuously varying velocities, this happens when the ray is instantaneously horizontal, that is, $\iota = 90^\circ$ and the ray parameter or horizontal slowness p is equal to the medium slowness $1/\alpha(z_{\max})$; the term $z_{\max}$ is defined to be the turning point of the ray and is the deepest point in the medium sampled by the ray. The integral form of the distance equation is given by Eq. (8), where $u(z) = 1/\alpha(z)$ represents any

$$X(p) = 2\int_0^{z_{\max}} \frac{p\,dz}{\sqrt{u^2(z) - p^2}} \tag{8}$$

relationship of slowness versus depth.

The case of a spherical Earth is treated by replacing the flat-Earth ray parameter by Eq. (9), where r

$$p = \frac{r \sin \iota}{v(r)} \tag{9}$$

is the distance from the center of the Earth, $v(r)$ is the velocity at that radius, and ι is the angle of intersection between the ray and the radius vector (Fig. 3*b*). The velocity function $v(r)$ denotes either $\alpha(r)$ or $\beta(r)$, depending on the circumstance. The radius r_0 at which $\iota = 90^\circ$ is known as the turning radius of the ray and denotes the point at which the ray begins its journey back toward the surface. If $v(r)$ is a constant independent of radius, the ray path is straight but the turning point is still well defined. In a spherical Earth, the ray travels in the plane containing the source, the receiver, and the center of the Earth. Like its flat-Earth counterpart, the spherical-ray parameter must be conserved along the ray.

Seismogram synthesis. These simple geometric arguments can be extended to the computation of amplitudes provided that there are no sharp discontinuities in the velocity as a function of depth. More exact representations of the amplitudes and wave shapes that solve the full-wave equation to varying extents can be constructed with the aid of powerful computers; these methods are collectively known as seismogram synthesis, and the seismograms thus computed are known as synthetics. Synthetics can be computed for elastic or dissipative media that vary in one, two, or three dimensions.

The fundamental rationale for computing synthetics is rooted in the notion of the seismogram as an observable entity, and the distributions of Lamé parameters and density within the Earth as unknowns. In other words, it is the province of seismological imaging to estimate the three-dimensional variation

of elasticity and density from observations of seismic-wave propagation. Beyond elucidation of basic structure, the variation of elasticity and density is a proxy for the variation of the chemical and thermal properties within the Earth. These in turn provide constraints on the long-term evolution of the Earth as a physical system.

Seismogram synthesis is an example of a forward problem; given a mathematical representation of the Earth and a model of the seismic source, attempts are made to compute synthetic seismograms (or some observables like travel time). For every forward problem, there is a concomitant inverse problem. The general statement of the inverse problem is: Given a set of observations of some measurable seismic disturbance at the Earth's surface, what can be said about that part of the Earth through which the disturbance passed and the characteristics of the source of the disturbance? The source and structural inverse problems are dichotomous in that something must be known about the source before the seismogram can be inverted for structure, and vice versa. This is not a problem in crustal or exploration imaging, where the properties of the artificial source are for the most part known and controlled. It is a problem when natural sources like earthquakes are used and specific steps must be taken either to isolate source and structure inversions or to do them jointly.

The development of inversion methods is a major area of research in seismology. The kinematic predictions of ray theory provide a useful starting point for a discussion of seismogram interpretation and inversion.

Crustal imaging. In a typical experiment for crustal imaging, a source of seismic energy is discharged on the surface, and instruments record the disturbance at numerous locations. Many different types of sources have been devised, from simple explosives to mechanical vibrators and devices known as airguns that discharge a "shot" of compressed air. The details of the source–receiver geometry vary with the type of experiment and its objective, but the work always involves collecting a large number of recordings at increasing distance from the source. **Figure 4***a* shows a seismogram known as a $T(x)$ resulting from one such experiment conducted at sea where an airgun source was fired and recordings were made by devices towed several meters below the sea surface from a research vessel. The strength of an arriving disturbance can be judged approximately by how dark the record appears.

This seismogram is complex, exhibiting a number of distinct arrivals with a variety of shapes and having amplitudes that change with distance, though they do not simply lose amplitude with increasing distance. Although this seismogram clearly does not resemble the structure of the Earth in any sensible way and is therefore not what would normally be thought of as an image, it can be analyzed to recover estimates of those physical properties of the Earth that govern seismic-wave propagation. Some simple deductions can be made by inspection. For instance, in a flat Earth the travel time equation is hyperbolic in distance. In the distance ranges up to about 15 km (9 mi), the observed $T(x)$ seismogram in Fig. 4*a* indeed exhibits a number of segments which have a roughly hyperbolic shape. This can be taken immediately to indicate that the Earth here probably contains at least a few simple layers of uniform velocity. Some segments, particularly those developed at ranges greater than about 20 km (12 mi), do not show a hyperbolic form. In fact, their curvature is in the opposite sense to that expected for a hyperbolic reflection form. This indicates that the arrivals have propagated through a structure in which there is a continuous increase in velocity with depth, and are therefore turning or diving rays as discussed above.

The formal procedure of travel-time inversion involves deriving the turning-point depth for each ray parameter by using a formula for $z_{\max}$ that is obtained from Eq. (5) by an integral transform known as the Abel transform. This approach has been used extensively by seismologists, often very successfully. Methods have been developed that allow observational uncertainties to be included in the procedure so that the inversion calculates all of the possible $\upsilon(z)$ functions allowed by the data and their associated uncertainties, thereby giving an envelope within which the real Earth must lie. The approach has several difficulties principally associated with the nonlinear form of the $X(p)$ equation. Although the desired quantity, $\upsilon(z)$, is single-valued (the Earth has only one value of velocity at any given depth), $X(p)$ can be multivalued. $X(p)$ also has singularities that are caused when the quantity under the square root goes to zero. Thus, care must be taken not to use rays that do not turn but are reflected at interfaces. In practice, separating out only those arrivals needed for the inversion can be very difficult. Furthermore, $X(p)$ data are not actually observed but must be derived from observations. Methodologies that can overcome these problems attempt to linearize the equation with respect to small variations in $X(p)$ and obtain an inversion that includes solution bounds derived from data uncertainties by linear programming.

Much of the difficulties have been further overcome by a coordinate transformation as shown in Eq. (10), where $\tau(p) = T - pX$.

$$\tau(p) = 2\int_0^{z_{\max}} dz\sqrt{u(z)^2 - p^2} \qquad (10)$$

The term $\tau(p)$ is known as the intercept time, and it can be thought of as the time that a vertically traveling ray takes to propagate upward from the turning point of a ray with ray parameter p. Although Eq. (10) still represents a nonlinear relationship between data $\tau(p)$ and desired Earth model $u(z)$, the square root no longer presents a problem, and this leads to considerable stabilization of the inversion. The observed seismogram can be easily transformed to $\tau(p)$ by a process known as slant stacking, in which a computer search is made for all arrivals having a particular horizontal slowness; those data are

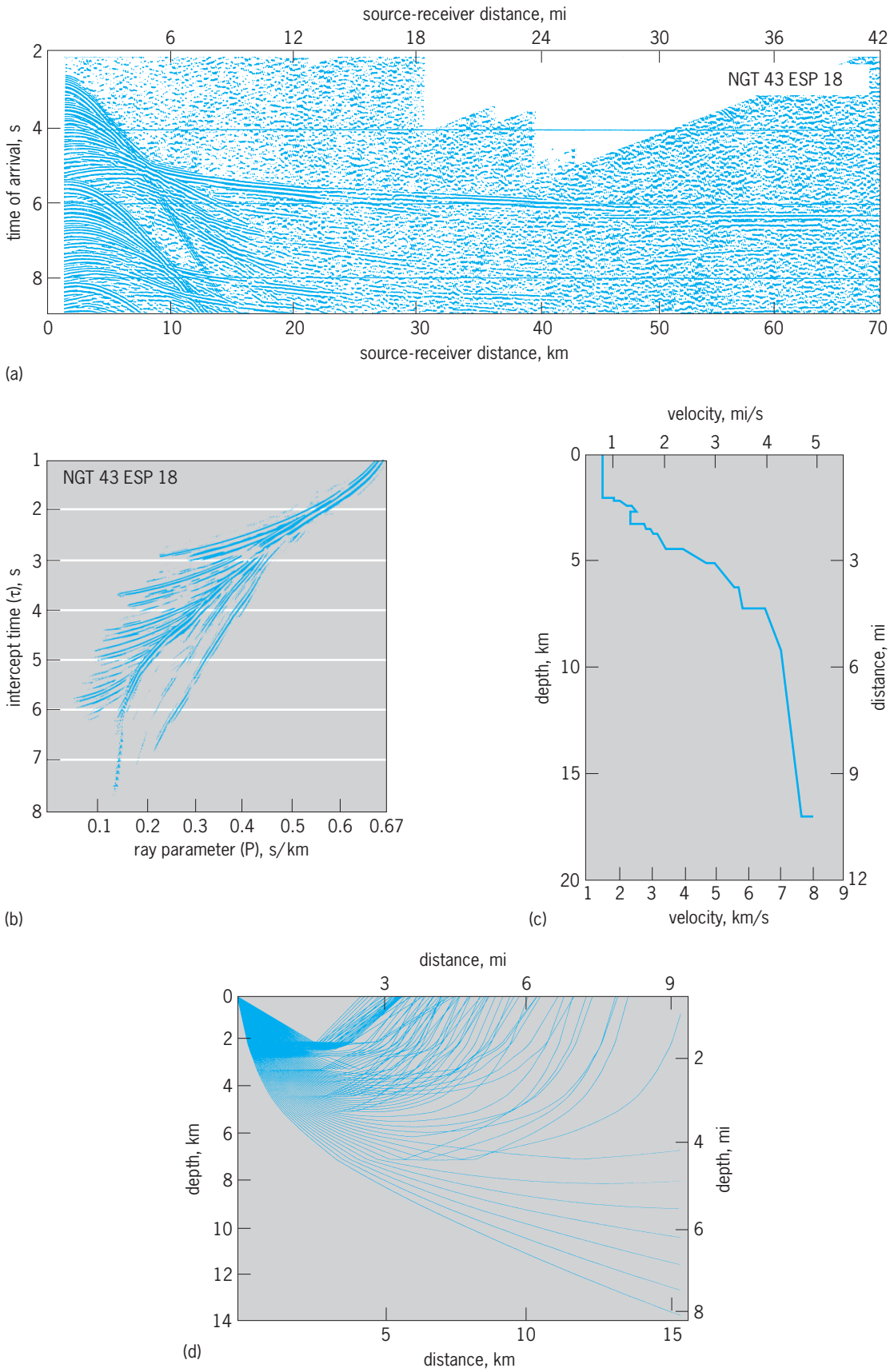

Fig. 4. Crustal imaging. (*a*) A *T*(*x*) seismogram recorded off the continental margin of East Greenland as part of a two-ship experiment designed to study the structure of the crust. (*b*) The *T*(*x*) seismogram data are transformed to $\tau(p)$ for analysis and for beginning the travel-time inversion. The part of the $\tau(p)$ data corresponding to the turning rays is organized into the strong band of energy in the seismogram on its lower right side. (*c*) By inversion and iterative forward modeling, an estimate of the change in compressional-wave velocity with depth is derived. (*d*) Rays that propagate through the structure derived from analysis of the seismogram take complex paths.

summed; and using the appropriate X and T data, $\tau(p)$ is computed (Fig. 4*b*). Most importantly, like the desired $\alpha(z)$ function these $\tau(p)$ data are single-valued, so that deriving the $\upsilon(z)$ function amounts to a point-by-point mapping from $\tau(p)$ to $\alpha(z)$.

Much of this theory has been known for quite some time, but the $\tau(p)$ approach to inversion and analysis of seismic data did not see extensive application until relatively recently. One reason is that the quality and sparseness of many early seismic experimental data did not allow them to be treated with $\tau(p)$ methods. Data often comprised just a dozen or so seismograms obtained by using explosive charges and a few recording stations, the records from which were made in analog form usually on paper records. The example seismogram (Fig. 4*a*) is typical of modern marine seismic data in comprising several hundred equally spaced traces. The seismic source consisted of an array of airguns. Seismic arrivals are recorded on another vessel towing a hydrophone array more than 2 km (1.2 mi) long containing several thousand individual hydrophones. Recording on this ship is synchronized to the shooting on the other vessel. Ship-to-ship ranges are measured electronically and written onto the same system that records the seismic data together with other information such as shot times. This experimental method leads to the dense sampling of the seismic wavefield that is

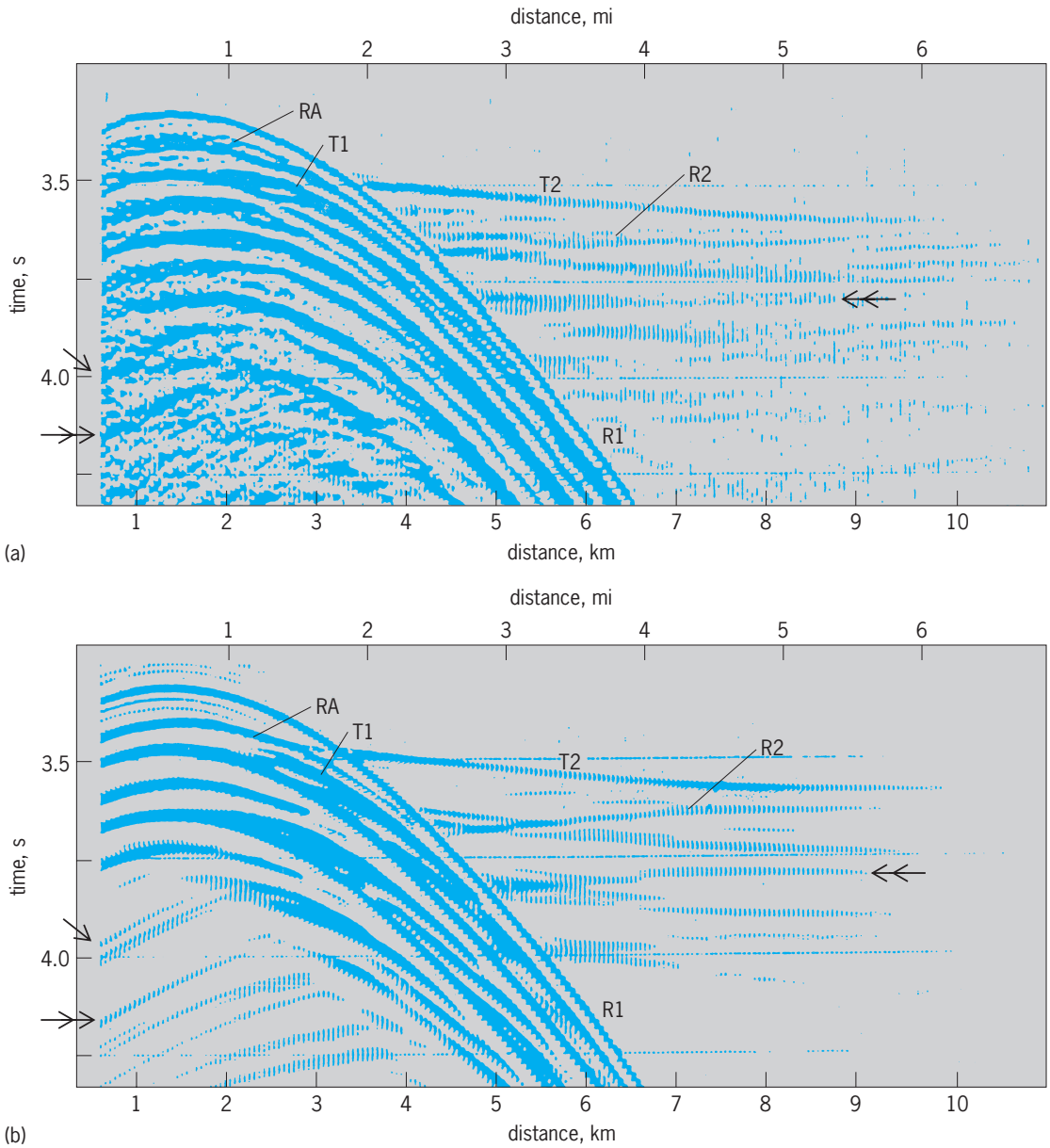

Fig. 5. Forward modeling; single and double arrows indicate the two types of reflected arrivals from the top of the axial magma chamber beneath the East Pacific Rise. (*a*) A *T*(*x*) seismogram obtained on the spreading center of the East Pacific Rise, used to obtain a velocity depth function by inversion and travel-time modeling. (*b*) A full synthetic seismogram, calculated and compared to the observed data as a way of resolving more detail in the structure than could be obtained from the travel times. In both parts, particular seismic phases [turning rays (T) and reflected rays (R)] are indicated; the vertical axis is a reduced time scale. (*From E. Vera et al., The structure of 0.0 to 0.2 m.y. old oceanic crust at 9°N on the East Pacific Rise from expanded spread profiles, J. Geophys. Res., 95:15529–15556, 1990*)

required for the correct computer transformation of the observed data into $\tau(p)$. The process is extremely demanding of computer time; it has come into common use only since relatively inexpensive, fast computers have made high computational power available to most researchers.

Figure 4*c* shows the result of analysis of the seismogram of Fig. 4*a*. It shows a profile of compressional-wave velocity against depth in the Earth at a location on the continental margin off East Greenland. The upper part of the crust where sedimentary layers are present is particularly complex, showing both uniform velocity layers, ones that show a gradational increase in velocity with depth, and one (at a depth of 3 km or 1.9 mi) in which the velocity is less than that of the layer above. This $\alpha(z)$ profile was constructed by the combination of several methodologies, including travel-time inversion from the $\tau(p)$ data in Fig. 4*b* and forward modeling of travel times and seismic amplitudes. Figure 4*d* depicts the rays that would propagate through such a structure; it is possible to recognize reflected arrivals that bounce off the interfaces between layers, and diving rays that smoothly turn in layers in which the velocity increases smoothly with depth.

Information from seismic waveforms. Real seismic disturbances have a finite time duration and a well-defined shape (see insert in Fig. 1). In passing through the Earth, any seismic disturbance changes shape in a variety of ways. The amount of energy reflected and transmitted at an interface depends on the ray incident angle (or equivalently on its ray parameter) and the ratio of physical properties across the boundary. It is sensitive both to the medium's compressional-wave and shear-wave velocities and to the ratio of densities. This is because, as well as splitting the arrival into reflected and refracted waves, boundaries act to convert waves from one mode of propagation to another. Thus, some of the compressional-wave energy incident on a boundary is always partitioned into shear energy.

Formulas for reflection and transmission coefficients are substantially more complex than those for the angles, and they are derived from considerations of traction across the boundary. At some incident angles the reflected energy becomes particularly strong. One point at which this happens is when the refracted wave travels horizontally below the interface, giving rise to a propagation mode known as a head wave, and essentially all the incident energy is reflected. This phenomenon is analogous to total internal reflection in optics. On the seismogram shown in Fig. 4*a*, there are several places where amplitudes become strong, and some can be related to critical-angle reflection, as this phenomenon is known in seismology.

Other strong regions correspond to places where energy is focused by the effect of changing subsurface gradients, which cause many diving rays to be returned to the surface near the same regions. Figure 4*d* was drawn by tracing a series of rays through the structure in Fig. 4*c*, with each ray incrementing in ray parameter by a fixed 0.005 s/km (0.008 s/mi). There are several regions where the density of rays returned to the surface is relatively sparse, while in other regions the density is very high (around 6 km or 3.7 mi, for instance). This pattern of amplitude variations can be used to refine models obtained by travel-time inversion and modeling to provide estimates of properties such as density, shear velocity, and attenuation (the amount of energy dissipated into heat by internal friction); however, it is often difficult to separate velocity and density uniquely, and estimates of acoustic impedance, the product of velocity and density, are more commonly derived.

Waveform information can be employed either by forward modeling or by using the $T(x)$ seismogram (Fig. 4*a*) directly in the inversion procedure. In the

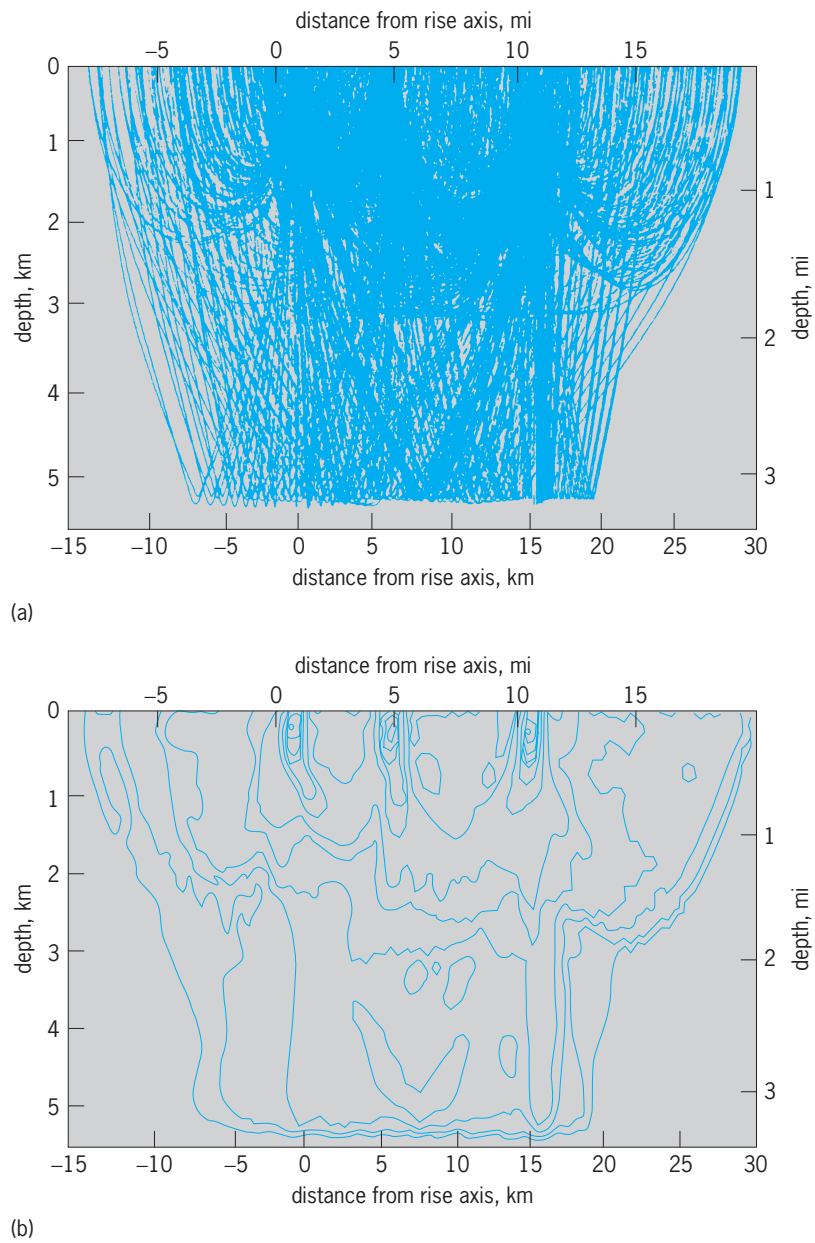

Fig. 6. Imaging by seismic tomography. (*a*) The type of ray pattern that can be used to conduct a tomographic inversion of travel-time data. (*b*) Travel-time partial derivatives that are closely related to the ray density and hence to the ability of the method to resolve structure.

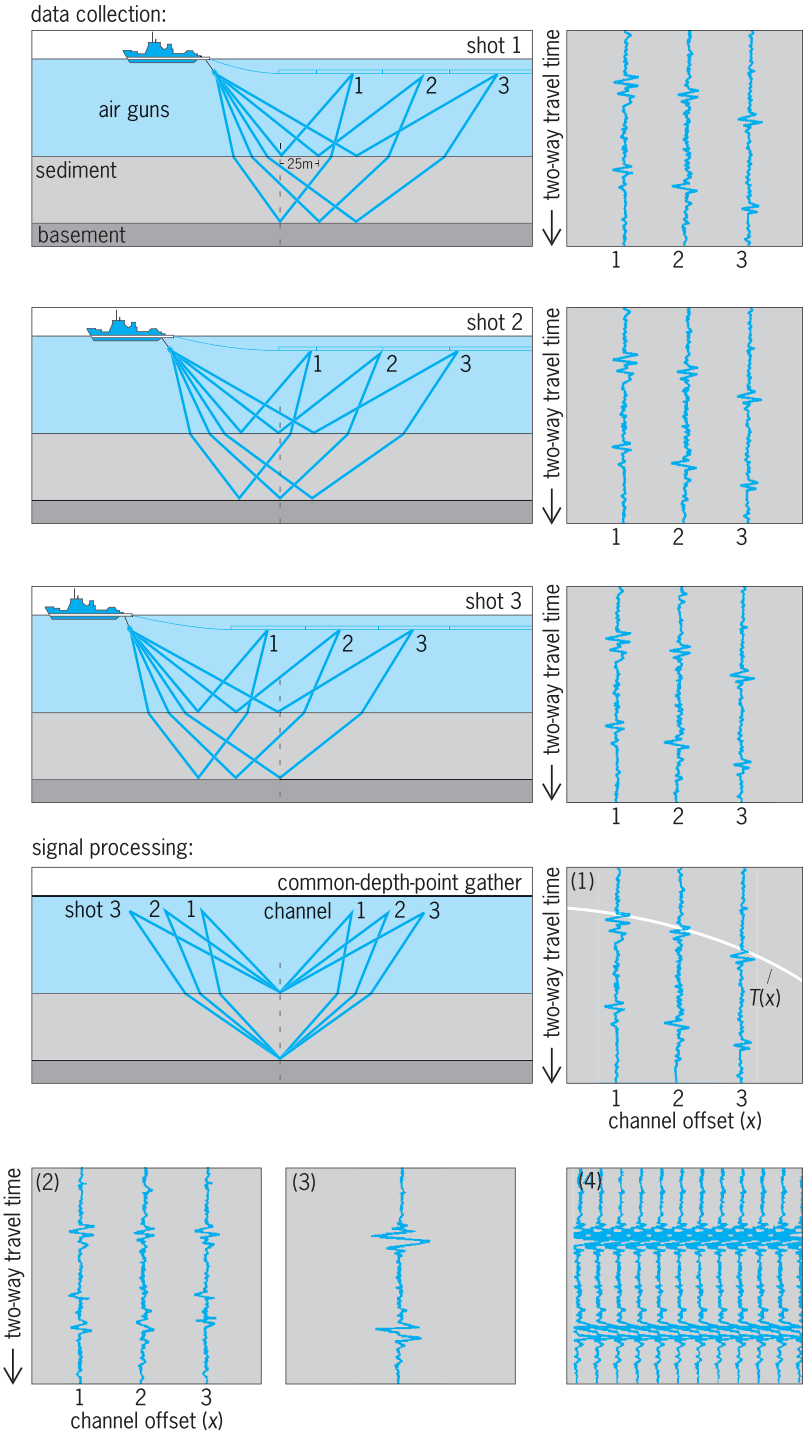

Fig. 7. Schematic representation of a standard multichannel seismic profiling experiment and the postcruise data processing. The first three rows show the ship towing an array of airguns and a recording streamer. The airguns fire at fixed intervals, sending acoustic energy (the heavy lines) into the water and the oceanic crust. The first (top) pulse represents the reflection of acoustic energy from the interface between the water and the sediment, and the second (bottom) pulse represents the reflection from the sediment–basement interface. The panels to the right show the *T*(*x*) seismograms recorded by each of the three hydrophones [channel offset (*x*) 1, 2, and 3] for shots 1, 2, and 3, with time increasing downward. The steps in the postcruise signal processing are denoted by (1)–(4) in the corner of the *T*(*x*) plots. Since the geometry of the shots and the recording streamer are maintained as the ship steams along, the *T*(*x*) seismograms of the three shots can be summed or "gathered" to reduce the effect of noise [plot (1)]. In (2), the common-depth-point gather is corrected for the effects of velocity moveout by fitting short sections of hyperbolas to the pulses. The results for each channel can then be gathered again to reduce noise further, as in (3). This is done continuously as the ship steams along, eventually producing an image of the interfaces, as in (4). (*After J. Mutter, Seismic images of plate boundaries, Sci. Amer., pp. 66–69, February 1986*)

forward methodology a synthetic seismogram is calculated on a computer; a model structure, usually derived from travel-time inversion, is used as a starting point, and the effect of propagation of a real seismic source through the structure is computed. The synthetic is then compared to the observed data, misfits noted, and adjustments made to the model until a satisfactory fit is obtained (**Fig. 5**). Computation of synthetic seismograms is enormously consuming of computer time, and the various methods in use all apply some computational shortcuts that make the problem tractable. Like inversion of travel-time data from $\tau(p)$, the advent of modern computers has made the computation of synthetics more realistic, though only supercomputers allow this to be done routinely.

These very fast machines may also allow waveform data to be incorporated more directly into the inversion of seismic data. Clearly the desired aim is to make use of all the information that the entire seismogram contains to yield estimates of all the physical properties that contribute to the observed waveform of the seismogram. The problem is a highly nonlinear one, and can be regarded as an optimization problem often addressed with a nonlinear least-squares procedure in which computed and observed data are systematically compared and adjusted until a fit is obtained that satisfies some optimization criterion. The computational demands of such a procedure are such that the use of supercomputers is essential. The automatic inversion of the complete seismogram to recover all the physical property information that affects the seismic waveform is one of the most challenging areas of research in crustal seismology.

Two- and three-dimensional imaging. A volume of the crust can be directly imaged by seismic tomography. In crustal tomography, active sources are used (explosives on land, airguns at sea) so that the source location and shape are already known. Experiments can be constructed in which sources and receivers are distributed in such a way that many rays pass through a particular volume and the tomographic inversion can produce relatively high-resolution images of velocity pertubations in the crust (**Fig. 6**). Crustal tomography uses transmitted rays like those that pass from a surface source through the crust to receivers that are also on the same surface.

Because most applications of tomography make use of rays that are refracted in their passage through a structure, they provide representations of the crust expressed in terms of smoothly varying contours of velocity, or velocity anomaly with respect to some reference. Many interfaces within the crust are associated with relatively small perturbations in velocity and occur on relatively small spatial scales, making their imaging by tomographic techniques essentially impossible. The finely layered strata of a sedimentary basin, for instance, cannot be imaged by such an approach.

The most successful approach that has been devised to address the detailed imaging of crustal structure involves the use of reflected arrivals and is obtained by a profiling technique. In the reflection

profile technique, energy sources and hydrophone (or geophone) arrays are the same as those used to create seismograms used for travel-time inversion; but in the profiling experiment, both are towed from the same vessel or moved along the ground together.

To produce an image from field recordings, they are first regrouped (gathered), corrected for the effect of varying source–receiver distance, and summed (stacked). The regrouping is made so that arrivals that are summed have come from the same reflection points and the time corrections are made assuming that all the arrivals obey hyperbolic travel times (**Fig. 7**) and Eq. (4). Although the latter assumption is not completely correct, it is an acceptable approximation for the relatively small source–receiver offsets (the distance between the shot position and the receiver) used in this type of experiment. After summation, the resultant seismogram has very high signal-to-noise characteristics and appears as if it were obtained in an experiment in which source and receiver were fixed at zero separation and moved along the profile line. Such a record is shown in **Fig. 8**.

Mathematically, if a recording of the wavefield at the surface has been obtained, and since it is known that propagation obeys the wave equation, then the equation can be used to move the wavefield back to its point of origin; that is, it can be extrapolated back down into the Earth to the place where it began to propagate upward. Having done this, and applying some condition that allows the strength of the reflection to be determined, it becomes possible to recover an essentially undistorted image of the structure. A conceptual similarity to travel-time inversion can be recognized in this methodology. In travel-time inversion, formulas are used for $X(p)$ or $\tau(p)$ to determine the turning-point depth of a ray. This could be restated by saying that the calculation extrapolates back down the ray from the surface to its turning point. Reflection imaging uses the wave equation to extrapolate the entire reflected wavefield back to its points of reflection, and can be considered the inversion methodology appropriate for the reflected wavefield.

Several wavefield extrapolation methods have been developed, all of which involve several stages of manipulation of the data, and all are very demanding of computer time. Figure 8 shows one example in which data obtained in shallow water in a sedimentary basin showing complex deformation structures have been imaged to reveal a variety of geologically interpretable structures. The velocity field in the structure must be known very well (and may not be) for the imaging to be successful. Most extrapolation methods involve some form of approximation to the full-wave equation to make the computations tractable. Most, for instance, allow only mild lateral variations in velocity. To do the job in a complete sense requires use of the full-wave equation and operating on prestack data.

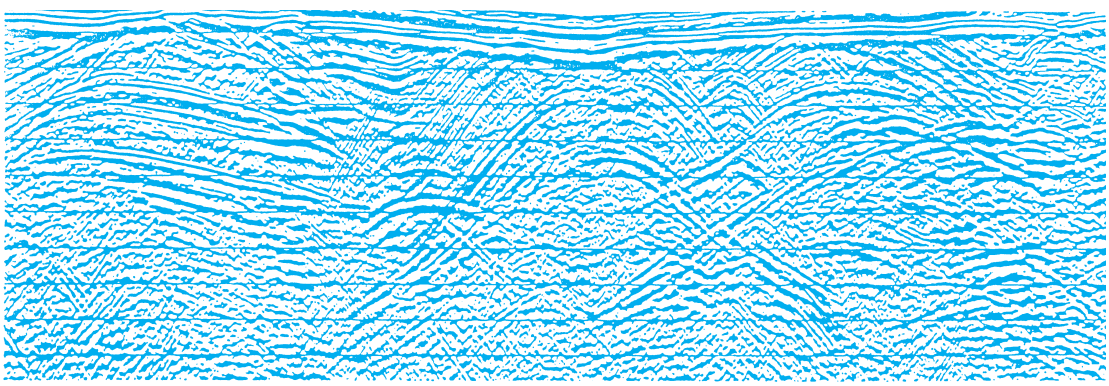

(a)

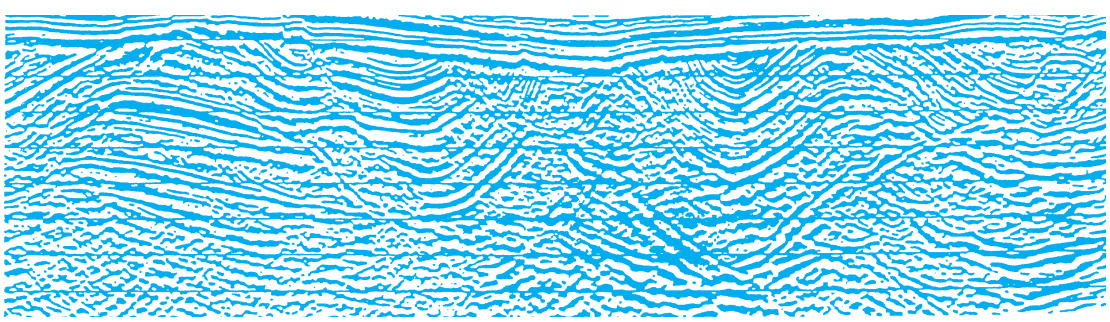

(b)

Fig. 8. Seismic profiling. (*a*) Stacked reflection profile obtained by using the common depth point (CDP) profiling method described in Fig. 7. The record represents the acoustic response of the Earth to near-vertical-incidence seismic energy along a traverse. It is analogous to the seismogram in Fig. 4*a*, which represents the response of the Earth at one location to seismic energy at a wide range of angles. (*b*) Reflection seismic image produced by migrating the data from *a*; migration of reflection seismograms can be treated as the equivalent of inversion of refraction seismograms.

Advances, typically led by industry requirements, have seen surveys conducted to provide three-dimensional images, usually for petroleum reservoir evaluation. Survey lines are conducted on orthogonal grids in which the line spacing is as little as 50 m (160 ft) in both directions. Large exploration vessels carry two long hydrophone arrays set apart by 50 m (160 ft) on large booms or deployed on paravanes. Correct imaging of structure from such recordings requires a fully three-dimensional wavefield extrapolation procedure, and this has been developed successfully by using the power of supercomputers. Once the migrated image has been made, it can be manipulated in a smaller computer to observe structure in ways that cannot be achieved by conventional means. Horizontal slices can be made to show buried surfaces, diagonal cuts can be made to examine fault structure, and the whole image can be rotated to view the volume from any chosen direction.

The correct three-dimensional imaging of the Earth, either the whole Earth by tomographic means or crustal structure by tomography and reflection seismic imaging, represents a frontier research area in seismology. *See* GEOPHYSICAL EXPLORATION; GROUP VELOCITY; PHASE VELOCITY; SEISMIC EXPLORATION FOR OIL AND GAS; WAVE EQUATION.

Discoveries. Although the theory of plate tectonics allowed Earth scientists to recognize that the global mid-ocean ridge system represents the locus of formation of oceanic lithosphere, it gave little direct insight into the processes operating at these spreading centers. At a few scattered locations throughout the world, small slivers of oceanic crust known as ophiolites have been thrust into exposure onto continents as a result of collisional tectonics. Their study by geologists led to the proposition that large, shallow, relatively steady-state magma chambers are responsible for producing the crust at all but the slowest-spreading ridges. Efforts directed toward seismic imaging of spreading centers since 1980 produced some unexpected results. *See* MID-OCEANIC RIDGE; OPHIOLITE.

Reflection imaging of the fast-spreading center at the East Pacific Rise combined with two ship refraction seismic measurements in 1985, and later tomographic imaging, showed that a magma body indeed exists beneath the axis but that it is small and discontinuous (**Fig. 9**). The magma body is so small, in fact, that the term chamber seems barely applicable. It is typically not more than 2–3 km (1.2–1.8 mi) wide, and the inferred region of melt (judged from the thickness of layers determined from the results of inversion and modeling of refraction seismic

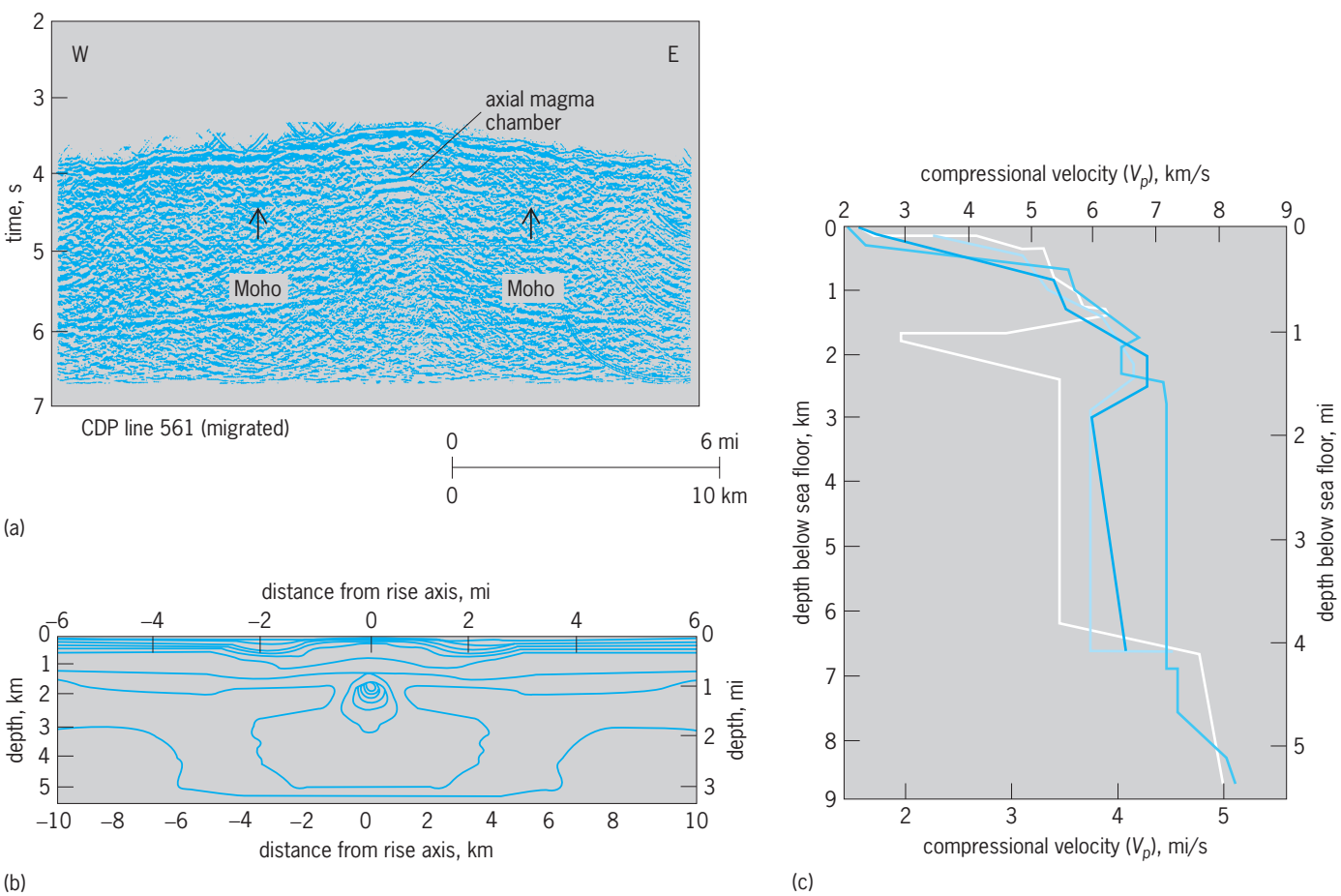

Fig. 9. Results of several different seismic investigations of the spreading center of the East Pacific Rise. (*a*) Reflection seismic image. (*b*) Velocity structure obtained by tomographic inversion. (*c*) Compilation of several velocity depth profiles obtained from inversion, travel-time, and synthetic seismogram modeling of refraction seismic data.

data) with greatly reduced velocities may be only a few hundred meters. This region is embedded in a broader zone of reduced velocities that may represent a region containing traces of melt but is more likely to be solid rock in which the temperature has been raised by proximity to the region of melt. The base of the crust, the Mohorovičić discontinuity (Moho), is a strong reflector formed very near the rise axis. *See* MOHO (MOHOROVIČIĆ DISCONTINUITY).

Images of continental crust and its underlying mantle have proven to be equally provocative (**Fig. 10**). Reflection profiling conducted on land in many regions of the world and on the shallow continental seas around Britain have shown that the deep crust is often characterized by an unreflective upper region and a band of strongly laminated reflections forming the lower third or so of the crust above a Moho that is quite variable in nature. The best available images show distinct events from within the mantle section itself. Mantle reflections may result from deep shear zones with reflectivity enhanced by deeply penetrating fluids. As exploration of the upper mantle has continued, it has become clear that structure in the mantle can be imaged by reflection methods to very great depths—perhaps even to the base of the lithosphere. It is possible that reflection methods will eventually be used to investigate the structure of the lithosphere as a whole, thereby complementing studies based on a different class of seismic methods that already have been developed. *See* LITHOSPHERE.

Global and regional seismograms. The basic unit of observation in global and regional seismology is a seismogram, but unlike their counterparts in crustal reflection and refraction, most seismometers used for larger-scale structural studies are geographically isolated from their neighbors. Thus the observational techniques and most methods of analysis used in global seismology have evolved quite differently from crustal seismology. Generally speaking, much more of each seismogram must be retained for analysis. This so-called spatial resolution gap is being closed slowly by the development of new portable instrumentation suitable for recording natural sources. Once these instruments are available in quantity, seismologists will be able to record seismic energy at closely spaced sites that will illuminate in much finer detail structures deep within the Earth.

Seismograms are classified according to the distance of the seismometer from the source epicenter. Those recorded within about 50–100 km (30–60 mi) of a large source are generally complex not only because of intervening structure but because the "dimensions" of the source are close to the propagation distance and different areas of stress release on the fault write essentially different seismograms. The term near-field is given to these seismograms to signify that the propagation distance for the seismic energy is less than a few source dimensions. Understandably, these seismograms are most useful for examining the details of the earthquake rupture. Beyond the near field to distances just past 1000 km (620 mi), the seismograms are dominated by energy propagating in the crust and uppermost mantle. These so-called regional seismograms are complicated, because the crust is an efficient propagator of high-frequency energy that is easily scattered; but there are still discernible arrivals. These seismograms are used to examine the velocity structure and other characteristics of relatively large blocks of crust. The domain beyond 1000 km (620 mi) is called teleseismic. Seismograms written at teleseismic distances are characterized by discrete and easily recognized body phases and surface-wave arrivals. These are relatively uncontaminated by crustal structure, and instead they are more sensitive to structure in the mantle and core.

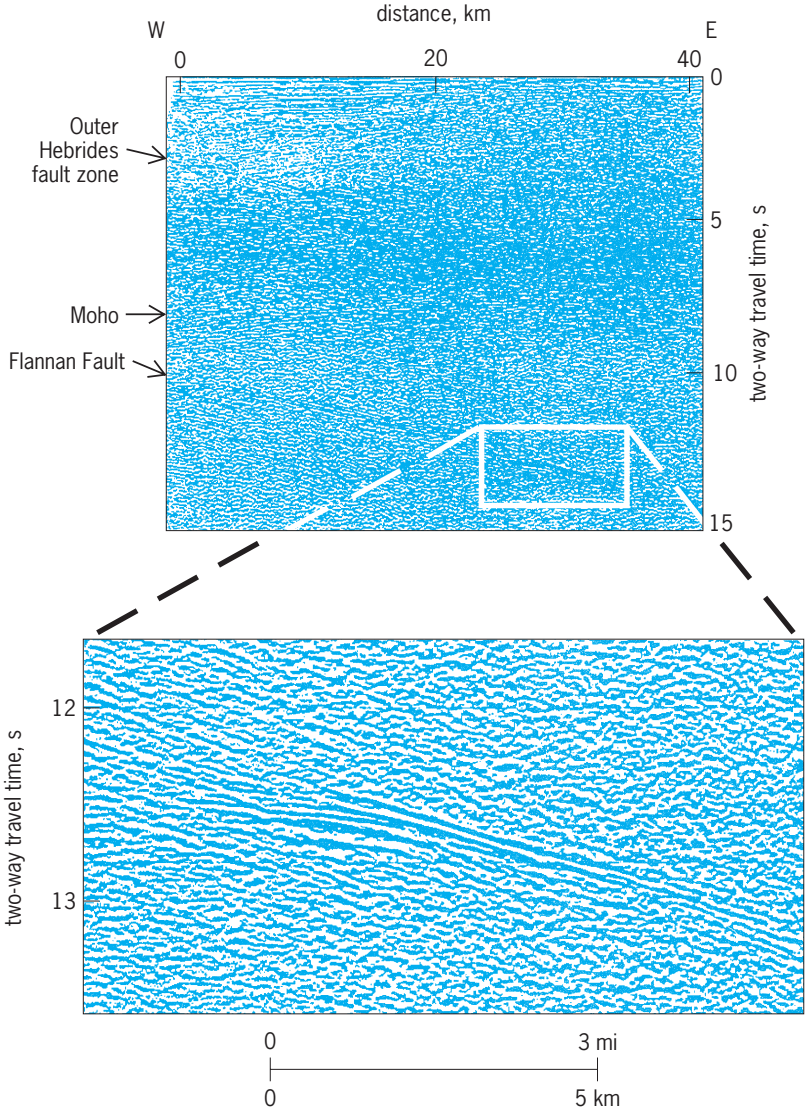

Fig. 10. Reflection seismic images of deep continental crust offshore Britain showing a variety of reflecting interfaces in and beneath the crust. (*From D. H. Mathews and M. J. Cheadle, Deep reflections from the Caledonides and Variscides west of Britain and comparison with the Himalayas, in M. Banazang and L. Brown, eds., Reflection Seismology: A Global Perspective, Geodynamics Ser., vol. 13, 1986*)

Figure 11 shows both regional and teleseismic seismograms; since seismic-wave motion is vector motion, three components (up-down, north-south, and east-west) are needed to completely record the

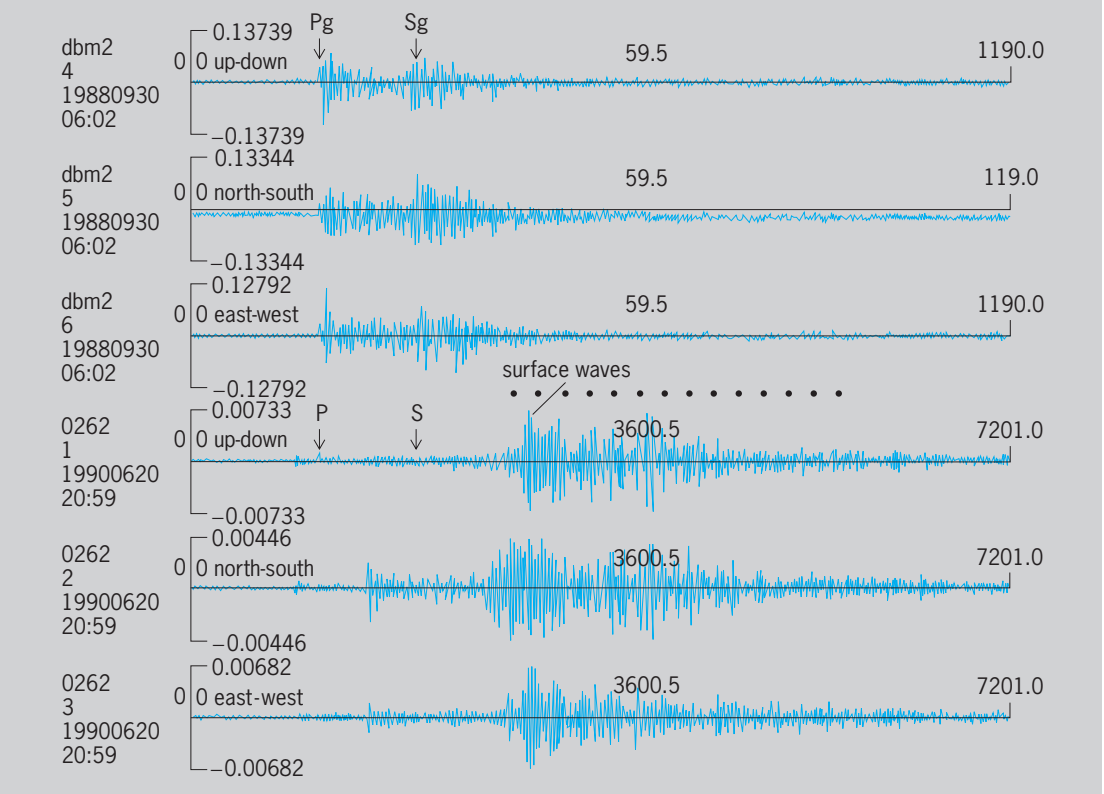

Fig. 11. Regional and teleseismic seismograms. The top three traces (seismometer installation dbm2) are the up–down, north–south, and east–west components of the regional seismogram from a 1-ton chemical explosion in the Adirondack Mountains of New York State (recorded at a distance of 107 km or 66 mi for 2 min). The regional phases Pg and Sg are indicated. These waves are guided by the crust over long distances and thus are very complicated. The bottom three traces (seismometer installation 0262) are the three components of the teleseismic seismogram produced by an earthquake in Iran (June 20, 1990; recorded at Palisades, New York, at a distance of 84°, or more than 9300 km or 5780 mi from the epicenter for 2 h). In addition to the P and S arrivals that have propagated through the Earth's mantle, the large-amplitude low-frequency arrivals later in the record are surface waves that have propagated in the Earth's outer layers. Detailed information is given in the text.

incoming wavefield. For the regional seismogram, the first arrival (Pg) is the direct P wave, which dives through the crust; the second arrival (Sg), beginning some 15 s after the Pg, is the crustal S phase. Even though explosions are inefficient generators of shear waves, it is common on regional seismograms to see substantial shear energy arising from near-source conversions of compressional to shear motion. The arrival time of the Pg phase can be picked easily, but the arrival time of Sg is somewhat more obscure. This is because some of the shear energy in the direct S is converted into compressional energy by scatterers in the crust near the receiver. This converted energy, traveling at compressional-wave speeds, must arrive before the direct S. The extended, very noisy wavetrains following the direct P and S are called codas; they represent energy scattered by small heterogeneities elsewhere in the crust. It is difficult to analyze these codas deterministically, and statistical procedures are often used. One result from the analysis of these and other seismograms is that the continental crust is strongly and three-dimensionally heterogeneous and scatters seismic energy very efficiently.

For the teleseismic seismogram, the teleseismic record is 60 times longer than the regional seismogram, and has been truncated only for plotting purposes. Plotted at this compressed time scale, it looks quite similar to the regional seismogram, but in fact there are important differences. The direct P and direct S waves are clearly evident; the S waves are larger on the horizontal components and larger than the P waves, a common feature for natural sources such as earthquakes. The very large-amplitude, long-period arrival starting some 20 min after the S are the surface waves. These waves are generally the most destructive in a major earthquake because of their large amplitude and extended duration. Between the direct P and the direct S are compressional body waves which have shallower take-off angles, and have bounced once (PP) or twice (PPP) from the surface before arriving at Palisades (the analogy here, though inverted, is like skipping a rock over water). The high level or energy between the S wave and the first surface wave is due to other multiply surface-reflected and -converted body waves, which are very sensitive to the structure of the upper mantle.

For purposes of analysis, the travel times of major phases have been inverted tomographically for large-scale structure. The International Seismological Centre (ISC) in England has been collecting arrival-time picks from station operators around the world since the turn of the century. Nearly one thousand

stations report arrival-time picks consistently enough so that their accuracy can be judged without viewing the original seismograms. More than 2 million of these P-wave arrival times have been inverted tomographically for the structure of the lower mantle, the core, and the core–mantle boundary. These inversions show that the core–mantle boundary and the mantle just above it are very heterogeneous. Tomographic inversions of special P phases that transit the boundary show that the boundary may have topography as well. Estimates of the amplitude of this topography range from a few hundred meters to several kilometers, but further work must be done.

Direct P phases dive deeply into the mantle and are not very sensitive to upper mantle structure unless the earthquakes are within 25° or so of the seismometer. Multiple bounce body phases are more sensitive to upper-mantle structure, but these must be picked carefully from the teleseismic records. This has been done for several thousand seismograms situated in the United States and Europe, and the resulting models give very finely detailed views of the structure of the mid to upper mantle. These methods are sensitive enough to resolve the remnants of old subducted slabs in the lower mantle and the large keels of high-velocity material which seem to underlie most old continents.

Going beyond tomographic methods requires analyzing more of the seismogram, for it is in the details of the wave shapes that information about major structures and the boundaries between them can be found. For example, surface waves disperse (different periods travel at different velocities) with wave speeds that depend on the fine details of the crust and upper mantle. An example of this dispersion can be seen in the Iranian seismogram (Fig. 11). Surface waves with longer periods arrive first and thus travel faster than surface waves at shorter periods. Measuring surface-wave dispersion is difficult but feasible, particularly for shallow earthquakes such as the Iranian event where the surface-wave arrivals dominate the seismogram. Seismologists began to measure dispersion in the late 1950s; some of the first applications of computers to seismological problems were the calculation of dispersion curves for oceanic and continental upper mantle and crust.

Measuring dispersion and other wave properties has evolved into a class of inversion procedures in which the entire seismogram is inverted directly for Earth structure. Such waveform inversion methods became feasible when computers became powerful enough to allow synthesis of most of the seismogram (**Fig. 12**). After many years of operation, even the sparse global digital arrays collect enough data to contemplate tomographic inversions. The most successful of these experiments have combined the seismogram-matching techniques of waveform inversion with a generalization of the tomographic approach to obtain models of the three-dimensional variation of upper mantle structure. The vertical integration shown in **Fig. 13** is a useful way of describing the geographic (two-dimensional) variation of a three-dimensional structure. Dark areas are fast relative to an average Earth, while light areas are slow. The darkest regions correspond to the oldest continental crust, known as cratons, while the lightest areas correspond to regions of active rifting either at mid-ocean ridges or at incipient oceans such as the Red Sea. *See* WAVEFORM.

Imaging the seismic source. The other half of the imaging problem in global seismology is constructing models of the seismic source. The simplest description of a so-called normal earthquake source requires two orthogonal force couples oriented in space. One force couple occurs on opposing sides of the rupture or fault surface and can be understood as the stress on either side of the fault induced by the earthquake rupture. The other couple is normal to the first and is required to conserve angular momentum. Implicit in this representation is the assumption that the earthquake is a point in both space and time (the point-source approximation), so that this type of source representation is known as a point-source

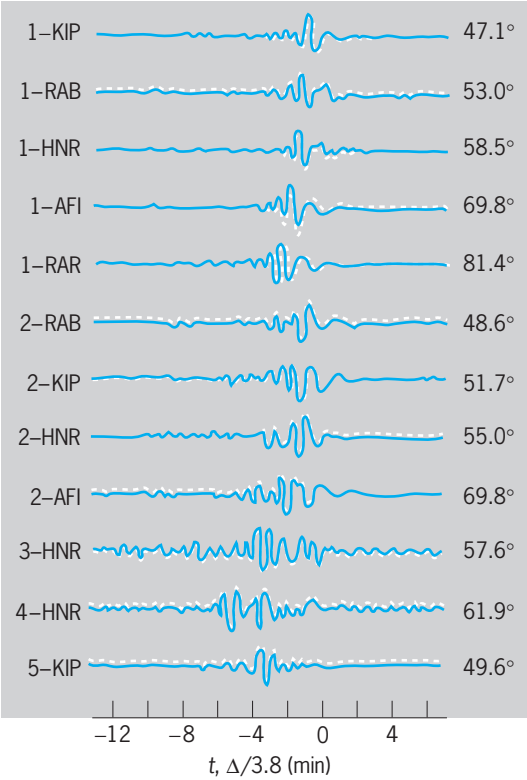

Fig. 12. Example of waveform inversion applied to fundamental and higher-mode surface waves crossing the western Pacific. The solid lines are data, and the broken lines represent synthetic seismograms computed for a model of the Earth obtained by waveform inversion. The numbers and letters to the left of each trace correspond respectively to earthquakes and stations used in the inversion. The distance between the epicenter and the recording station in degrees is given to the right of each trace. The seismograms are plotted as a function of reduced time, in which the time axis is translated by an amount proportional to the epicentral distance divided by a chosen velocity (Δ/3.8). This causes arrivals having the same velocity to line up vertically. The fits of the synthetic to the observed are very nearly "wiggle-for-wiggle" at nearly all distances, indicating the power of the waveform inversion technique.

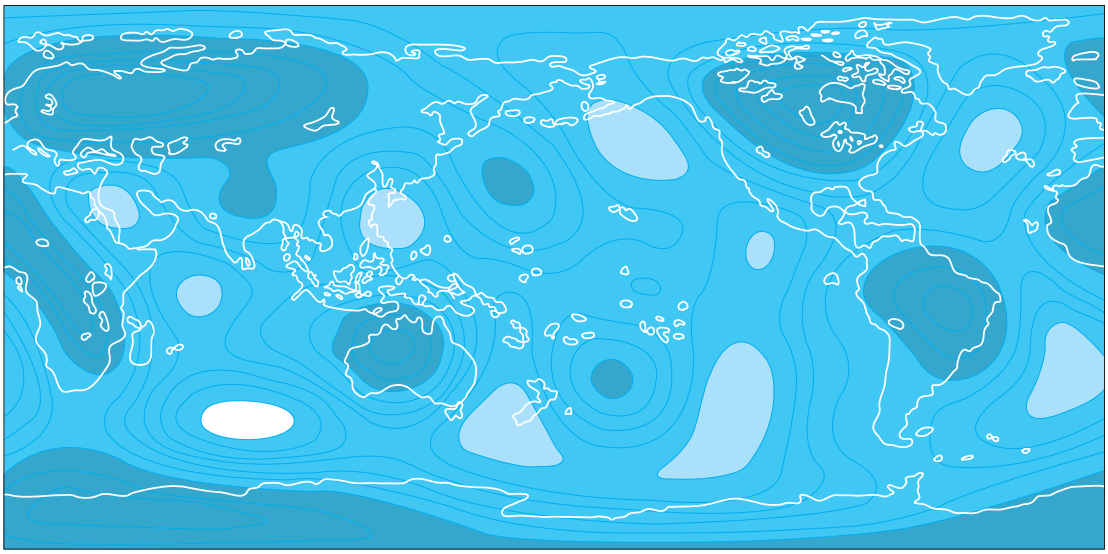

Fig. 13. Travel time anomaly from 400 to 40 km (240 to 25 mi), showing the deviation away from an average Earth of vertical shear-wave travel time derived by vertically integrating a three-dimensional model of shear–velocity variation. The model was developed by the inversion of several thousand long-period seismograms for upper-mantle velocity variations and then integrated. Dark areas are fast relative to the Earth, while light areas are slow. The darkest regions correspond to the oldest continental crust, known as cratons, while the lightest areas correspond to regions of active rifting either at mid-ocean ridges or incipient oceans such as the Red Sea.

double couple. Although there are some minor complications, the orientation of this double couple, and hence the orientation of the fault plane and the direction of rupture, can be inferred by analyzing the polarity of the very first P and S waves recorded worldwide. Thus this representation is known as a first-motion mechanism. It is not very difficult to make these observations provided that the instrument polarities are known, and the ensuing "beachball" diagrams are commonplace in the literature.

First-motion representations of seismic sources are the result of measurements made on the very first P waves or S waves arriving at an instrument; therefore they represent the very beginning of the rupture on the fault plane. This is not a problem if the rupture is approximately a point source, but this is true in practice only if the earthquake is quite small or exceptionally simple. An alternative is to examine only longer-period seismic phases, including surface waves, to obtain an estimate of the average point source that smooths over the space and time complexities of a large rupture. This so-called centroid-moment-tensor representation is an accurate description of the average properties of a source; it is routinely computed for events with magnitudes greater than about 5.5. Because an estimate for a centroid moment tensor is derived from much more of the seismogram than the first arrivals, it gives a better estimate of the energy content of the earthquake. This estimate, known as the seismic moment, represents the total stress reduction resulting from the earthquake; it is the basis for a new magnitude number M_W. This value is equivalent to the Richter body wave (m_b) or surface-wave magnitude (M_S) at low magnitudes, but it is much more accurate for magnitudes above about 7.5. The largest earthquake ever recorded, a 1960 event in Chile, is estimated to have had an M_W of about 9.5. For comparison, the 1906 earthquake in San Francisco had an estimated M_W of 7.9, and the 1964 Good Friday earthquake in Alaska had an M_W of 9.2.

While first-motion and centroid-moment-tensor representations are useful for general comparisons among earthquakes, there is still more information about the earthquake rupture process to be gleaned from seismograms. Types of analysis and instruments have been developed that demonstrate that some earthquakes, especially larger ones, are not adequately described by a first-motion or the averaged centroid-moment-tensor representations and require a more complex parametrization of the source process. These additional parameters may arise from relaxation of the point-source approximation in space or time or perhaps even of the double-couple restriction. In the former case, the earthquake is said to be finite, meaning that there is a spatial or temporal scale defining the rupture. The finiteness may also be manifested by multiple events occurring a few tens of seconds and a few tens of kilometers apart. In the latter case, the earthquake may have non-double-couple components that could be the result of explosive or implosive chemical phase changes, landslides, or volcanic eruptions. The size of the explosive non-double-couple component is one way of discriminating an earthquake from a nuclear explosion, for example.

One indication of source finiteness is an amplification of seismic signals in the direction of source rupture. Normally, the variation of seismic amplitudes with azimuth from a double-couple source varies in a predictable quadrupolar or bipolar pattern (the beach ball illustrates this best). For some earthquakes, however, the seismic waves that leave the source region at certain azimuths are strongly amplified. This has been ascribed to propagating ruptures

which "unzip" the fault along a fairly uniform direction. The best estimate is that most faults rupture at about two-thirds of the shear-wave velocity, but some faults may rupture even more slowly. Both unilateral and bilateral ruptures have been observed, and an important theoretical area in seismology is the examination of more complex ruptures and the prediction of their effects on observed seismic signals.

Another indication of source finiteness is the fact that some large events comprise smaller subevents distributed in space and time and contributing to the total rupture and seismic moment. The position and individual rupture characteristics of these subevents can be mapped with remarkable precision, given data of exceptional bandwidth and good geographical distribution. An outstanding problem is whether the location of these subevents is related to stress heterogeneities within the fault zone. These stress heterogeneities are known as barriers or asperities, depending on whether they stop or initiate rupture. The mappings of stress heterogeneities from seismological data is an active area of research in source seismology. *See* SEISMOGRAPHIC INSTRUMENTATION.

John Mutter; Art Lerner-Lam

Bibliography. J. Havskov and G. Alguacil, *Instrumentation in Earthquake Seismology (Modern Approaches in Geophysics: Vol. 22)*, 2004; C. H. Scholz, *The Mechanics of Earthquakes and Faulting*, 2002; S. Stein and M. Wysession, *An Introduction to Seismology, Earthquakes and Earth Structure*, 2002.

Seisonacea

A class of the phylum Rotifera which comprises a group of little-known marine animals. They form a single family with about seven species and are found only in Europe. The Seisonacea are epizoic or possibly ectoparasitic on crustacea, especially on members of the genus *Nebalia*. They have a very elongated jointed body with a small head; a long, slender neck region; a thicker, fusiform trunk; and an elongated foot, terminating in a perforated disk (see **illus.**). The Seisonacea differ from other rotifers in a number of characteristics. They have a very reduced corona, a quite aberrant type of mastax (fulcrate), paired ovaries but no vitellaria in females, and paired testes in the males. Furthermore, the sexes are of similar size and development; only one type of egg is produced, and it requires fertilization. The Seisonacea are also larger than other rotifers, attaining up to 0.12 in. (3 mm) in length. *See* ROTIFERA.

Elbert H. Ahlstrom

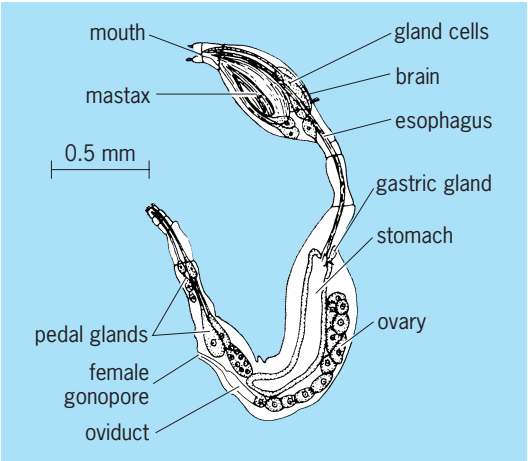

***Seison*, a rotifer.**

Seizure disorders

Conditions in which there are recurrent seizures. Such conditions are also known as epilepsy; the isolated occurrence of a seizure, however, is not designated as epilepsy. A seizure (ictus) is an event in which there is a sudden alteration in function of nerve cells, most commonly involving excessive electrical activity of the cells. This sudden change in nerve cell function is usually relatively brief, lasting seconds to minutes. Soon after a seizure, the brain may function quite normally. The manifestation of a seizure varies depending on which area of the brain is involved. *See* BRAIN; NERVOUS SYSTEM (VERTEBRATE).

Classification and symptomatology. A standard way to describe the seizures is outlined below. [The traditional nomenclature has been added in brackets.]

Nomenclature of the International Classification of Epileptic Seizures

I. Partial seizures (seizures beginning locally)
   A. Simple partial seizures (without impaired consciousness)
      1. With motor symptoms [focal motor seizures]
      2. With somatosensory or special sensory symptoms [focal sensory seizures]
      3. With autonomic symptoms
      4. With psychic symptoms [temporal lobe seizures]
   B. Complex partial seizures (with impaired consciousness)
      1. With impaired consciousness only [often confused with absence, or petit mal, seizures, from which it can be distinguished by an EEG]
      2. With automatisms [psychomotor seizures; temporal lobe seizures]
   C. Partial seizures secondary generalized [focal seizures progressing to grand mal (major motor) seizures, such as Jacksonian seizures]
II. Generalized seizures
   A. Generalized nonconvulsive seizures
      1. Absence seizures [petit mal seizures]
      2. Atonic seizures [atonic seizures; minor motor seizures; drop attacks]
   B. Generalized convulsive seizures
      1. Generalized tonic-clonic seizures [grand mal seizures; major motor seizures]

2. Myoclonic seizures [myoclonic seizures; minor motor seizures]
3. Tonic seizures [tonic seizures]
4. Clonic seizures [clonic seizures]

III. Unclassified epileptic seizures

Advances in the diagnostic evaluation of persons with epilepsy have also led to the classification of seizure disorders by epilepsy syndrome. The two major categories are considered as either generalized or localization-related (partial), with subclassifications of idiopathic, cryptogenic, symptomatic, or progressive for each syndrome.

The following are examples of the four common seizure types.

Focal motor epilepsy. This type of seizure, also known as a simple partial seizure with motor symptoms, is manifested by uncontrolled rhythmic jerking of the face, arm, or leg, caused by excessive abnormal discharges of nerve cells within the area of the brain, which under usual circumstances controls movement in that part of the body (**Fig. 1**). The abnormal nerve cell activity can be seen in the electroencephalogram recorded from electrodes on the scalp (**Fig. 2***b*). *See* ELECTROENCEPHALOGRAPHY.

Temporal lobe seizures. These seizures are known as simple partial seizures and may be manifested by a myriad of symptoms depending upon which part of the lobe is involved. Psychomotor, or complex partial, seizures are the most common type. Clinically, they are characterized by an alteration in the state of consciousness and performance of repetitive, patterned, non-goal-directed activity. Usually, the person is aware of having had the seizure but may not remember what happened during it. Occasionally, temporal lobe seizures produce visual, auditory, or olfactory hallucinations, or illusions,

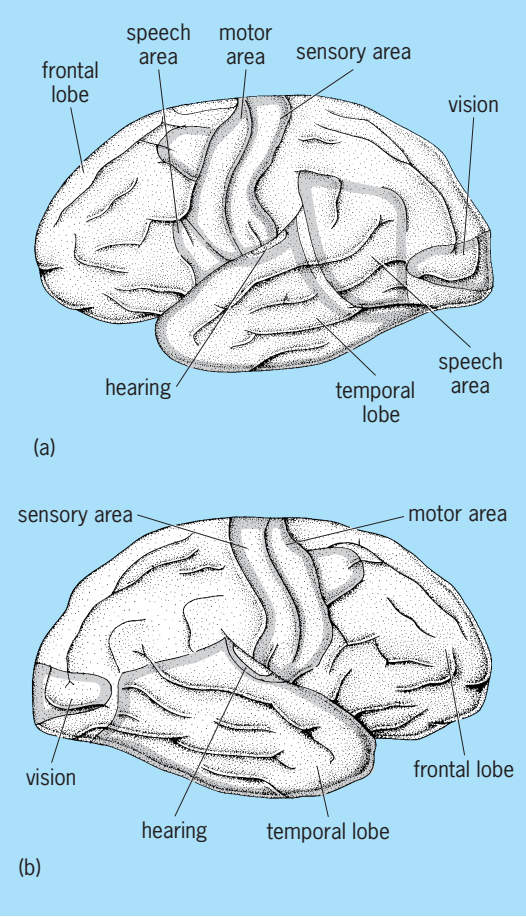

Fig. 1. Areas of the (*a*) left and (*b*) right brain.

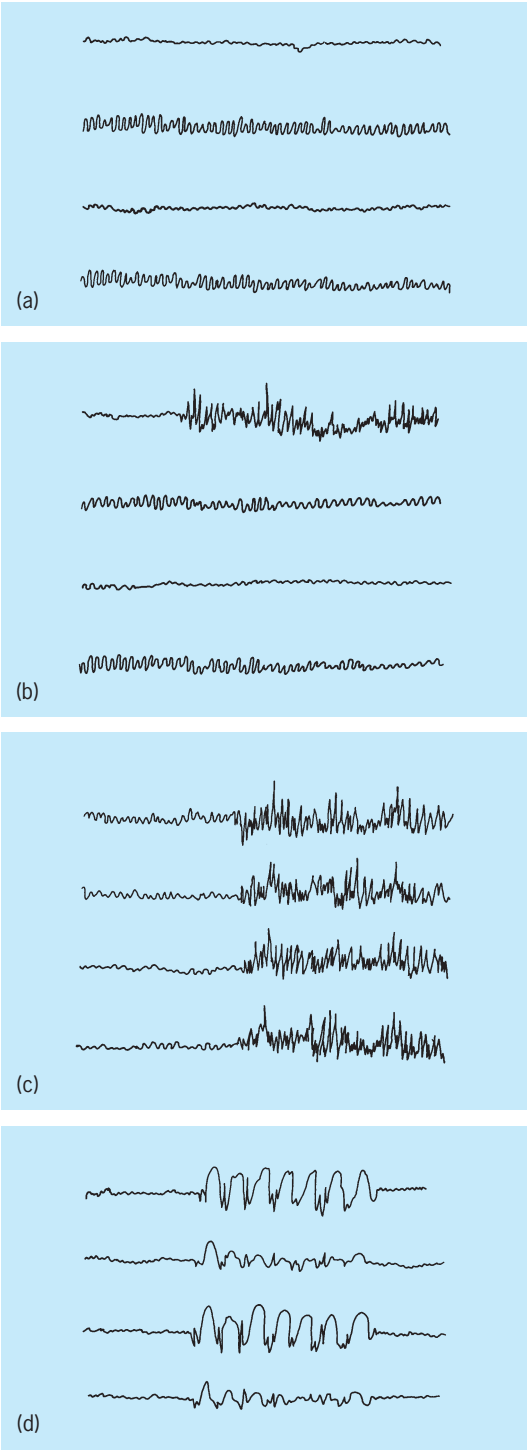

Fig. 2. Electroencephalograms: (*a*) Normal. (*b*) Partial (focal) seizure. (*c*) Generalized tonic-clonic (grand mal) seizure. (*d*) Absence (petit mal) seizure.

such as feelings of fear, familiarity, or distortions of size. These types of seizures are characterized electrically by abnormal discharges occurring within the temporal lobe for the duration of the seizure (Fig. 2*b*).

Grand mal. These seizures, also referred to as generalized tonic-clonic convulsive seizures, major motor seizures, or convulsions, occur if the abnormal discharges involve the entire brain all at once. In this condition there are forceful, generalized, symmetrical musculature contractions accompanied by loss of consciousness, and, at times, by urinary incontinence and tongue biting. In some instances, characteristic experiences of a sensory or visceral nature precede the more dramatic motor manifestations. Such usually brief subjective phenomena are termed auras. Actually, however, they represent the very beginning of an ictal episode. They are important for diagnostic purposes since they may indicate the portion of the brain in which the epileptogenic process is located.

A seizure comes to an end abruptly or gradually, and, depending on the pattern and intensity of the attack, the person may return immediately to normal preictal conditions or may continue to show various physical or mental deficits or both for some time. Amnesia is rather common, and after major seizures the subject often falls into a sleeplike state or goes through a period of confusion. *See* AMNESIA.

The relatively uncommon but serious life-threatening condition in which one convulsion occurs after another without recovery between them and lasts for hours is called status epilepticus.

Electroencephalographically, generalized tonic-clonic seizures are characterized by generalized rhythmic spike discharges in all areas of the brain at once which last several seconds and are followed by other generalized abnormalities for the duration of the seizure and for varying times thereafter (Fig. 2*c*).

Generalized nonconvulsive seizures. These may be atonic seizures, which are characterized by a sudden loss of muscle tone, or they may be absences (petit mal), which consist of brief periods of loss of consciousness and immediate recovery. The subjects with absences are often unaware of having had the episode. Electroencephalographically, they are characterized by generalized bilaterally synchronous spike and wave discharges occurring at three cycles per second (Fig. 2*d*). This type of seizure is more frequently seen in children than in adults.

Etiology. Epilepsy is not a disease in itself. It is a symptom of an underlying disease process. That disease process may be metabolic, such as uremia, decreased brain oxygen, or low calcium levels; or structural brain damage, such as from head trauma at birth, brain injuries, brain tumors, strokes, or congenital malformations, or previous encephalitis or meningitis. In many instances, however, a cause is not found. The disorder is then referred to as idiopathic or primary epilepsy. In some types of epilepsy there is apparently also a genetic component. The modes of genetic transmission are likely varied and may include the tendency to develop seizures with acquired brain damages as well as some types of "idiopathic" epilepsy. The role of genetic transmission in epilepsy must be evaluated on an individual basis.

Prognosis. The prognosis in epilepsy differs with the type of seizure. Some seizure types, once established, may persist. Other seizure types may have a high degree of spontaneous remission, especially those of childhood onset, such as febrile seizures and some absences (petit mal seizures). Seizures that are posttraumatic in origin may also remit spontaneously. In persons with a tendency to have epilepsy, events such as puberty, menses, alcohol intake, or emotional stress may precipitate seizures.

Diagnosis. Seizures should be distinguished from other conditions which have some clinical similarities. These include fainting from hypoglycemia, cardiovascular disorders, or hysteria.

The tools necessary to make this differential diagnosis are the history of the illness, the general physical exam, and a neurologic exam, all of which may be entirely normal or which may demonstrate signs of underlying disease processes. An electroencephalogram may demonstrate abnormal electrical discharges during a time the patient is seizure-free. These discharges may indicate that portion of the brain from which the seizures arise. Computerized tomography scans or magnetic resonance imaging (MRI) may show any gross structural abnormality. Examination of the blood may demonstrate abnormal circulating chemicals. At times an examination of cerebrospinal fluid may give information regarding the underlying cause of the epilepsy. In persons with medically refractory epilepsy, long-term EEG-video monitoring is employed to document seizure types, and the intracarotid sodium amytal (Wada) test is used to determine side of language dominance. Functional magnetic resonance imaging (fMRI) may assist in lateralization of language, and when used in conjunction with the EEG, in localization of epileptic discharges. *See* COMPUTERIZED TOMOGRAPHY; MEDICAL IMAGING.

Treatment. The ideal treatment of epilepsy is removal of the cause, such as excision of a tumor, or correction of a metabolic disorder, such as uremia or hypocalcemia. In many instances, however, the cause cannot be established or may not be amenable to direct treatment. When the cause cannot be removed, the symptoms (seizures) are treated.

It is important to control seizures for biologic and social reasons. Because of metabolic demands made by the abnormally discharging nerve cells, the seizures themselves may cause damage to nearby nerve cells and render the individual more susceptible to further seizures. The social importance of controlling seizures is obvious: a person who has frequent seizures cannot maintain a driver's license and may have difficulties with employment. The initial method of obtaining seizure control is antiepileptic drugs. About 80% of the people with epilepsy obtain good control or elimination of seizures with

medication. These medications are chemicals of varying structures and may be effective by a number of different brain mechanisms.

Medication. The following drugs are commonly used to control seizures. Phenytoin, carbamazepine, and phenobarbital have been used for decades for individuals with partial (focal) or secondarily generalized convulsive seizures while ethosuximide or valproic acid may be beneficial in generalized, nonconvulsive seizures (absence, petit mal). A new generation of drugs, including topirimate, lamotrigine, levetiracetam, zonisimide, and oxcarbazepine, are now available, and most of these may have broad spectrum efficacy against a variety of seizure types. In individuals whose seizures are more difficult to control, combinations of these drugs might be used, or other agents may be added. All antiepileptic drugs when used in high enough doses may produce toxic effects. The most commonly observed toxic effects from one or more of these agents include impairment of balance, double vision, or lethargy. It is possible to measure the amount of a circulating drug in the blood by the use of gas-liquid chromatography. In this way therapeutic ranges have been established that are a guide to the adjustment of medication. There are substantial individual differences in the way these drugs are absorbed and metabolized, so considerably different oral doses may be needed to achieve the same drug levels in blood in different patients. Some of these drugs may have an adverse effect on the fetus, and counseling with a neurologist and obstetrician is desirable before and during pregnancy.

Surgery. Brain surgery is a therapeutic potential for some of the 20% of the people with epilepsy who do not achieve control on medication. Surgery may be considered if the abnormally discharging nerve cells which cause the epilepsy are in a dispensable area of brain, that is, one of the frontal or temporal lobes (Fig. 1). One of these regions may be excised in one cerebral hemisphere without significant deficit. About 10% of the individuals refractory to medicine may be candidates for excision of the epileptic focus, and of these about 70% will achieve seizure control. Thus, epilepsy is a symptom, not a disease in itself, and when treated often has a good prognosis. However, about 15–20% of epileptics cannot achieve seizure control with the present modes of therapy.

Prevention. The occurrence of epilepsy can be decreased by measures to reduce head injury and birth defects. There are some indications that epilepsy can also be reduced by prophylactic administration of antiepileptic drugs in conditions which, if untreated, may lead to development of chronically recurring seizures. These include febrile convulsions of childhood, traumatic head injuries, and postoperative conditions of patients who have had brain surgery for other reasons.

Epilepsy was once considered a sacred disease and was surrounded by many myths and superstitions. Many rules concerning social interactions were imposed on people with seizures. In recent times, understanding of the entity as an organic disorder and recognition that it is not an all-encompassing disease have led to changes in legislation and discriminatory practices. Modern therapy has rendered many people free of seizures with little or no medication toxicity so that they may lead normal lives. Public education is another area which has been productive in determining early recognition and treatment of seizures and in prevention of epilepsy.

Linda Moretti Ojemann; Mark D. Holmes

Bibliography. F. E. Dreifuss, Proposal for revised clinical and electrographic classification of epileptic seizures, *Epilepsia*, 22:489–501, 1981; J. E. Engel and T. Pedley (eds)., *Epilepsy: A Comprehensive Textbook*, 3 vols., 1998; W. A. Hauser and D. C. Hesdorffer, *Epilepsy: Frequency, Causes and Consequences*, 1990; J. Laidlaw, A. Richens, and D. Chadwick (eds.), *A Textbook of Epilepsy*, 4th ed., 1993; G. A. Ojemann, Surgical therapy for medical intractable epilepsy, *J. Neurosurg.*, 66:489-499, 1987; C. P. Panayiotopoulos, *A Clinical Guide to Epileptic Syndromes and Their Treatment*, 2002; W. Penfield and H. Jasper, *Epilepsy and the Functional Anatomy of the Human Brain*, 1954; H. Weiser, Future aspects of epilepsy research, *Acta Neurochir. Suppl.*, 84:1–16, 2002.

Selachii

One of two subdivisions of the subclass Elasmobranchii, consisting of nine extant orders and 34 extant families, collectively known as sharks. The other subdivision, Batoidea, includes the sawfishes, guitarfishes, skates, and rays. In the classification system below, daggers indicate extinct orders. Each Recent order is covered by a separate article in the encyclopedia.

Class Chondrichthyes
- Subclass Elasmobranchii
  - Subdivision Selachii
    - †Order Cladoselachiformes
    - †Order Symmoriiformes
    - †Order Ctenacanthiformes
    - †Order Hybodontiformes
    - Order Heterodontiformes (bullhead sharks)
    - Orectolobiformes (carpet sharks)
    - Lamniformes (mackerel sharks)
    - Carcharhiniformes (ground sharks)
    - Hexanchiformes (frill shark and cow sharks)
    - Echinorhiniformes (bramble sharks)
    - Squaliformes (sleeper sharks and dogfish sharks)
    - Squatiniformes (angel sharks)
    - Pristiophoriformes (saw sharks)

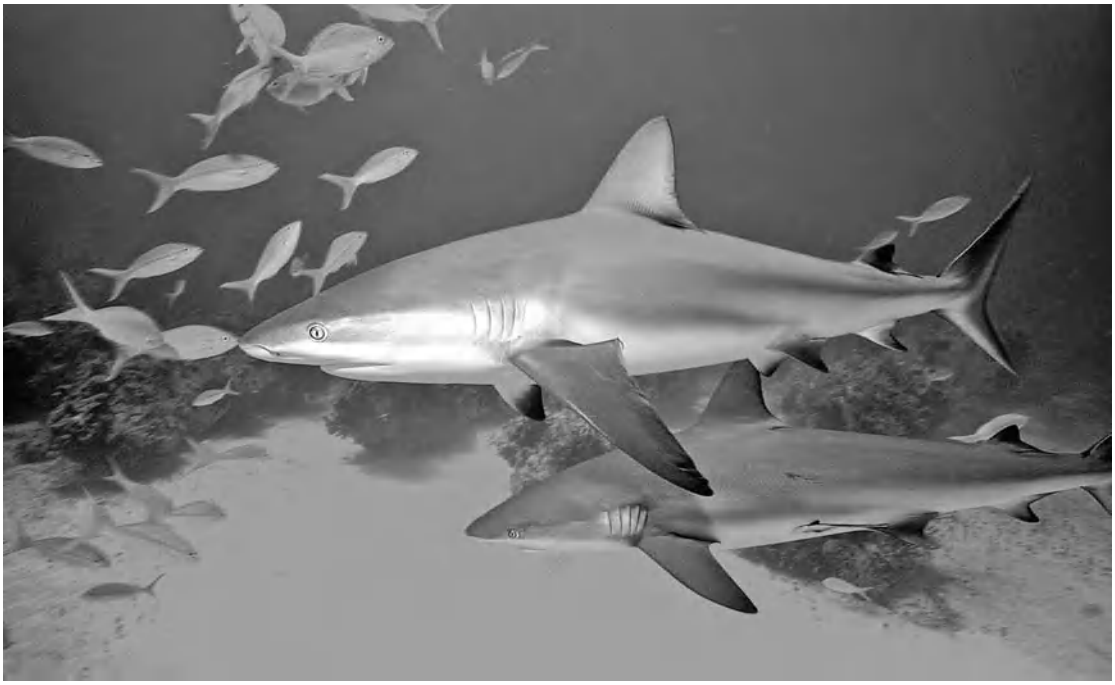

Caribbean reef shark (*Carcharhinus perezi*), a species of ground shark. (*Getty Images © Digital Vision*)

There are several characters common to all sharks that distinguish them unequivocally from all other elasmobranchs: the gill openings are at least partly lateral (that is, they can be seen in lateral view); the edges of the pectoral fins are not attached to the sides of the head anterior to the gill openings; and the upper margin of the orbit is free from the eyeball, that is, there is a free lid. *See* BATOIDEA; ELASMOBRANCHII.

Morphology. Sharks vary considerably in body form. On one end of the physiognomic spectrum are the mackerel and ground sharks (see **illustration**), with sleek fusiform bodies built for speed and power, while on the other end are angel sharks, with a batoid-shaped body adapted for cruising on or near the sea floor. Sharks vary greatly in size, ranging from only 0.4 m (1.3 ft) in adult smooth dogfishes to, reputedly, 18 m (59 ft) in whale sharks.

Ecology and development. Sharks are carnivorous predators for the most part, feeding on any flesh available, be it fresh or carrion; however, the largest of all sharks, whale sharks and basking sharks, feed on small schooling fishes, as well as tiny crustaceans. The poor eyesight of sharks is compensated by a very high order of olfaction. In most sharks the teeth vary in size and shape from front to back and in the upper and lower jaws. The teeth occur in series of rows, with replacement teeth always in reserve, waiting to rotate outward as teeth are lost. Development is oviparous in a few sharks, ovoviviparous in most, and viviparous in yet others.

Distribution. The great majority of sharks are found in relatively shallow waters around the world, primarily tropical and subtropical latitudes of both hemispheres. Their numbers decrease in temperate waters and drop precipitously in north temperate zones.

Herbert Boschung

Fossil selachii. All fossil and Recent sharks, with the possible exception of Devonian cladoselachians, may be included in the Selachii. The selachian endoskeleton is entirely cartilaginous, but it is frequently calcified superficially with calcium phosphate prisms. Although Recent sharks have a single layer of such tesserae, many Paleozoic forms had much thicker calcification consisting of many layers. The body is covered in minute scales, which in Recent sharks are simple and nongrowing (placoid), but which in earlier sharks often fused together and became enlarged during life (zonal scales) and may have been more regularly arranged over the body.

Most extant shark families are represented by the Late Cretaceous, apart from whale sharks (family Rhinodontidae), false cat sharks (Pseudotriakidae), and the frilled shark, *Chlamydoselachus*, all of which have a very poor fossil record. Cow sharks (Hexanchoidea), Port Jackson sharks (Heterodontidae), nurse sharks (Orectoloboidea), angel sharks (Squatinoidea), dogfish (Scyliorhinidae), and rhinobatoids had already developed before the close of the Jurassic. One family (Palaeospinacidae) appears in the Late Triassic (*Palaeospinax*) and ranges into the early Tertiary (*Synechodus*) before becoming extinct.

Because sharks periodically shed their teeth, sometimes replacing several thousand in a lifetime, their teeth are common in the fossil record. It is, therefore, rather surprising that the earliest records of sharks are from the Late Devonian, when they were already highly diversified. Knowledge of these early forms has improved greatly in recent years, particularly from the Pennsylvanian of North America.

Cladoselachians. Cladoselachians (late Paleozoic) are not known to possess male intromittent organs (claspers). Cladoselachian pectoral fins have a

single basal endoskeletal cartilage (metapterygium) and numerous long, unjointed radials, of which the anteriormost articulate directly with the shoulder girdle. The eyes are surrounded by a ring of sclerotic plates. A bladelike spine projects from the back, anterior to the first dorsal fin. *See* CLADOSELACHII.

Symmoriids. These mainly Carboniferous selachians have pectoral fins like cladoselachians, but the anterior radials are jointed to form a basal series. Radials of the other fins are also jointed once, and there is only one dorsal fin, behind which is a peculiar median cartilage. Most forms lack spines, but *Stethacanthus* has a single dorsal spine.

Ctenacanthids. The pectoral fin of these late Paleozoic sharks is tribasal, with a pro-, meso-, and metapterygium. In addition, there are two dorsal fins, each with a deeply inserted fin spine, and jointed radials in all fins. Since these characters also occur in the hybodontids, there is every possibility that ctenacanths are not a monophyletic taxon, but simply represent a primitive assemblage in need of further study. The aberrant xenacanths of the late Paleozoic may be closely allied to some ctenacanths. *See* XENACANTHIDA.

Hybodontids. These sharks primitively resemble ctenacanths and have long been considered intermediates between them and Recent sharks. However, hybodontids have distinctive cranial morphology, fin spines, head spines, and teeth, and calcified pleural ribs. Therefore, they are probably not ancestral to Recent sharks, particularly since they lack characters which unite these forms. John Maisey

Bibliography. H. B. Bigelow and W. C. Schroeder, Sharks, in *Fishes of the Western North Atlantic*, Mem. Sears Found. Mar. Res. Mem. 1 (pt. 1), pp. 29–556, 1948; L. J. V. Compagno, Interrelationships of living elasmobranchs, in P. H. Greenwood, R. S. Miles, and C. Patterson (eds.), *Interrelationships of Fishes*, pp. 15–61, Zool. J. Linn. Soc. 53, Suppl. 1, Academic Press, New York, 1973; J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 2006.

Selaginellales

An order of the class Lycopsida (club mosses), regarded as more advanced than the extinct Protolepidodendrales because they produce spores of two sizes, but less advanced than the Lepidodendrales and Isoetales because they lack wood and finite growth. All are perennial herbs, varying in growth habit from prostrate to climbing. *See* ISOETALES; LEPIDODENDRALES; PROTOLEPIDODENDRALES.

Although the approximately 700 extant species are traditionally assigned to a single genus, *Selaginella*, it is possible that this genus should include only the few species that possess undivided steles and microphyllous leaves of a single morphology. The bulk of the selaginellalean species are included in the genus *Stachygynandrum*. Fossil evidence suggests that the two lineages had already diverged by the early Carboniferous. Today, both genera are geographically widespread and occupy a wide range of habitats. Although selaginellaleans are rarely a significant element of the biomass in any of these communities, they dominate the ground flora in some tropical forests.

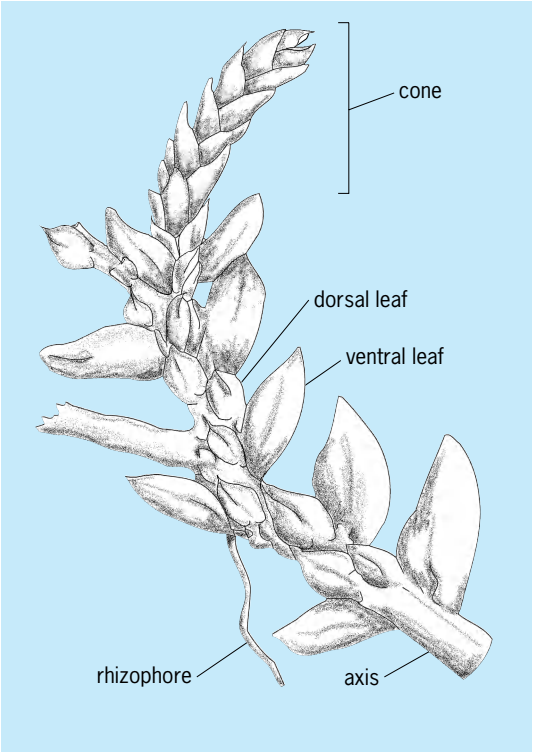

Characteristic features of a typical heterophyllous selaginellalean of the *Stachygynandrum* type. (*After T. R. Webster, Developmental problems in Selaginella in an evolutionary context, Ann. Mo. Bot. Gard., 79:633–647, 1992*)

Both selaginellalean genera show several remarkable biological features (see **illus.**). *Stachygynandrum* has leaves that are positioned both ventrally and dorsally. This character is regarded as an adaptation to efficient light capture; the orientation relative to light sources is precisely adjusted by differential growth, which is in turn controlled by auxins. Moreover, in many species the lower surface of the dorsal leaf and upper surface of the ventral leaf consist of more or less cubic cells that contain only one large chloroplast, whose orientation within the cell can be altered to maintain optimal levels of incident light for photosynthesis.

The origin of the rhizophores (specialized aerial rooting structures) is unclear.

Life history. Selaginellaleans have a free-sporing heterosporous life history. Sporangia are borne on sporophylls that are aggregated into cones. Each individual sporophyte is capable of producing both megasporangia, which on average contain four large megaspores, and microsporangia, which contain much larger numbers of much smaller microspores. Megasporangia generally develop in regions of the plant that experience the highest level of metabolic activity. Many species possess active

mechanisms for ejecting either microspores or entire microsporangia. *Selaginella selaginoides* disperses megaspores with a slingshot mechanism.

The megaspores and microspores germinate to produce independent, unisexual gametophytes. Both genders remain largely within the spore wall. The megagametophytes produce several archegonia that yield eggs, which chemically attract the swimming biflagellate spermatozoids released by the microspores. This requirement for an aquatic environment limits the habitats of sexually reproducing selaginellaleans.

Fossil record. Much of the fossil record of the Selaginellales is ambiguous. The more likely selaginellaleans are demonstrably both heterosporous and ligulate. *Selaginellites resimus*, from the early Carboniferous of England, is the earliest convincing record of the Selaginellales. Sporadic traces of possible selaginellaleans occur through the Mesozoic and Cenozoic. *See* LYCOPHYTA; LYCOPODIALES; LYCOPSIDA. Richard M. Bateman; William A. DiMichele

Bibliography. E. E. Karrfalt, The comparative and developmental morphology of the root system of *Selaginella selaginoides, Amer. J. Bot.*, 68:244–253, 1981; J. E. Skog and C. R. Hill, The Mesozoic herbaceous lycopsids, *Ann. Mo. Bot. Gard.*, 79:648–675, 1992; B. A. Thomas, Paleozoic herbaceous lycopsids and the beginnings of extant *Lycopodium* and *Selaginella, Ann. Mo. Bot. Gard.*, 79:623–631, 1992; T. R. Webster, Developmental problems in *Selaginella* in an evolutionary context, *Ann. Mo. Bot. Gard.*, 79:632–647, 1992.

Selection rules (physics)

General rules concerning the transitions which may occur between the states of a quantum-mechanical physical system. They derive in almost all cases from the symmetry properties of the states and of the interaction which gives rise to the transitions. The system may have a classical (nonquantum) counterpart, and in this case the selection rules may often be related to the classical conserved quantities. A first use of selection rules is in determining the symmetry classes of the states, but in a great variety of ways they may yield other information about the system and the conservation laws. *See* QUANTUM MECHANICS; SYMMETRY LAWS (PHYSICS).

Angular momentum and parity rules. For an isolated system the total angular momentum is a conserved quantity; this fact derives from a fundamental fact of nature, namely, that space is isotropic. Each state is then classifiable by angular momentum J and its z component M $(= -J, -J + 1, \ldots, +J)$. Angular momenta combine in a vectorial fashion. Thus, if the system makes a particle-emitting transition $J_1,M_1 \rightarrow J_2,M_2$, the emitted particles must carry away angular momentum (j, μ), where $\mathbf{j} = \mathbf{j}_1 - \mathbf{j}_2$. This implies that $\mu = M_1 - M_2$ and that j takes on values $J_1 - J_2, J_1 - J_2 + 1, \ldots, J_1 + J_2$. Thus in transitions $(J = 4 \leftrightarrow J = 2)$ the possible j values comprise only 2, 3, 4, 5, 6, and, if it is also specified that $M_1 - M_2 = \pm 4$, only 4, 5, 6. Observe that J_z is additive. *See* ANGULAR MOMENTUM; QUANTUM NUMBERS.

Another fundamental symmetry, the parity, which determines the behavior of a system (or of its description) under inversion of the coordinate axes, is conserved by the strong and electromagnetic interactions, and gives a classification of systems as even $(\pi = +1)$ or odd $(\pi = -1)$. Under combination the parity combines multiplicatively. Thus, if the transition above is $4^{\pm} \rightarrow 2^{\mp}$, it follows that $j^{\pi} = 2^-, \ldots, 6^-$, while $4^{\pm} \rightarrow 2^{\pm}$ would give $j^{\pi} = 2^+, \ldots, 6^+$. The angular momentum $\mathbf{j}$ may be a combination of intrinsic spin $\mathbf{s}$ and orbital angular momentum $\mathbf{l}$. Scalar, pseudoscalar, vector, and pseudovector particles are respectively characterized by $s^{\pi_s} = 0^+, 0^-, 1^-, 1^+$, where π_s is the "intrinsic" parity, while l always carries $\pi_l = (-1)^l$. *See* NONRELATIVISTIC QUANTUM THEORY; PARITY (QUANTUM MECHANICS); SPIN (QUANTUM MECHANICS).

Electromagnetic transitions. The photon is a transverse vector particle; this implies that its electric field vector (classically) or its intrinsic spin vector (quantum-mechanically) lies in the plane normal to the propagation direction. Thus, while an ordinary vector particle with total angular momentum j may, for $l > 0$, have $l = j, j - 1, j + 1$, corresponding to the three internal degrees of freedom, a transverse particle may have $l = j$ or else a particular linear combination of $l = j \pm 1$ (two internal degrees). An $l = j$ photon ("magnetic" type, or Mj) then carries parity $(-1)^{j+1}$, while the $l = j \pm 1$ combination (electric type, or Ej) carries $\pi = (-1)^j$; the magnetic-electric nomenclature is appropriate because, in the long-wavelength (LWL) limit (that is, when the photon wavelength is much greater than the radius of the source) the radiation is generated by a magnetic-type interaction (for example, $\mathbf{L} \cdot \mathbf{H}$ or $\mathbf{S} \cdot \mathbf{H}$ for M1, with $\mathbf{H}$ the magnetic field) or electric-type interaction ($\mathbf{P} \cdot \mathbf{E}$ for E1, with $\mathbf{P}$ the dipole operator and $\mathbf{E}$ the electric field). For $j = 0$, the only combination of s and l is $(s, l, j) = (1, 1, 0)$, which describes an oscillating E0 field, which however vanishes outside the source and hence does not radiate. Thus, for transitions of the form $J^{\pm}{}_1 \leftrightarrow J^{\pm}{}_2$, photons of type M1, E2, M3, ... may be emitted, and, for $J^{\pm}{}_1 \leftrightarrow J^{\mp}{}_2$, E1, M2, E3, ..., the j values being restricted as above. If M_1, M_2 are also specified, as in Zeeman transitions, there are further restrictions. For dipole transitions $(j = 1)$, only $\mu = 0$ (π lines) and $\mu = \pm 1$ (σ lines) occur, the former not being radiated parallel to the magnetic field because of the photon transversality. If $M_1 = M_2 = 0$, the transition is forbidden unless $J_1 + J_2 + j$ is even; this inversion rule, not restricted to electromagnetic transitions, arises from the behavior of states and operators under a rotation through π radians about an axis in the x-y plane. It is important in discussing molecular symmetries and isospin selection rules. *See* MULTIPOLE RADIATION.

Approximate rules. The "exact" selection rules are to be supplemented by approximate rules which depend on the internal structure of the system. To the extent that L, S, the total orbital and spin angular

momenta, are good quantum numbers, as happens in many atoms and a few nuclei, the angular momentum radiated in a long-wavelength E1 or E2 transition (whose transition operators are spin-independent) must come from the orbital structure; thus $(\Delta\mathbf{L}, \Delta\mathbf{S}) = (1, 0), (2, 0)$, respectively. For the M1 operator, which has the form $\alpha\mathbf{L} + \beta\mathbf{S}$, $(\Delta\mathbf{L}, \Delta\mathbf{S})$ can have the values (1, 0) or (0, 1) but not (0, 0) or (1, 1); however, beyond that, since **L** and **S** can generate no orbital, spin, or radial excitations, M1 transitions in *LS*-coupling occur only between two levels of the same term (although this is not true for *Mj* with $j > 1$).

Many processes, including photon emission and beta decay, occur via one-body operators [for example, by

$$\mathbf{P} = \sum_i e_i \mathbf{r}(i)$$

for E1, where the sum is over particles, with the *i*th particle having charge e_i and position operator $\mathbf{r}(i)$]. These generate only single-particle transitions, in which at most one particle changes orbit. (The number of such particles is $\delta n = 1$ for odd parity, as in

$$p^2 \xrightarrow{E1} sp,\ pd;$$

$\delta n = 0, 1$ for even parity, as in

$$d^2 \xrightarrow{E2} d^2,\ sd,\ dg;$$

$\delta n = 0$ for M1 in *LS*-coupling, but not in *jj*-coupling, since, for example,

$$f_{5/2} \xrightarrow{M1} f_{7/2}$$

is allowed.) Because of this "$\delta n < 1$" rule, the $(J\pi)$ rules apply to individual particles (for example, $\Delta l^{\pi} = 1^-$ for E1) and thus greatly facilitate study of the microscopic structures of the states involved. *See* ATOMIC STRUCTURE AND SPECTRA; NUCLEAR STRUCTURE.

Forbidden transitions. Forbidden transitions are those which are hindered by a significant selection rule or combination of them (the term is often applied in a more restricted sense in some domains, for example, in atomic physics, to all photon transitions except E1). Several examples will be used to illustrate the remarkable variety of information which they yield.

Parity violation. The alpha-decay process $A(J^{\pi}) \rightarrow B(0^+) + \alpha(0^+)$ is allowed only if $\pi = (-1)^J$ since the final angular momentum J is orbital only. The forbidden process, with A a $2-$ state in ^{16}O and B the ^{12}C ground state, has been observed, with a lifetime on the order of 10^{13} larger than that of a nearby $2+$ state in ^{16}O. Since neither of the decay products has a nearby 0^- companion, the decay can be ascribed to a small 2^+ admixture in the $2-$ state, the admixing "intensity" (square of the 2^+ amplitude) being then on the order of 10^{-13}. This gives a measure of a fundamental symmetry breaking, the violation of parity by the weak interaction. This example, like

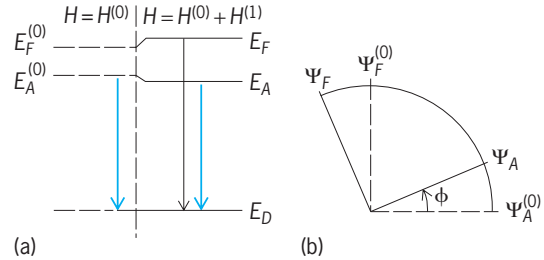

Symmetry admixing in a two-level substructure. (*a*) Energy-level diagram. (*b*) Rotation of states.

several of those following, involves symmetry breaking in a two-level substructure, the simple quantum mechanics of which is explained in the **illustration**.

Two close-lying states Ψ_A and Ψ_F are eigenfunctions of a hamiltonian $H = H^{(0)} + H^{(1)}$ with energies E_A and E_F (illus. *a*), where $H^{(0)}$ preserves a symmetry and the small $H^{(1)}$ breaks it. Transitions occur to a distant state Ψ_D (with energy E_D), in which the symmetry admixing is negligible. $\Psi_A^{(0)}$ and $\Psi_F^{(0)}$, eigenfunctions of $H^{(0)}$, have different symmetries. The selection rules are such that the transition $\Psi_A^{(0)} \rightarrow \Psi_D$ is allowed, while $\Psi_F^{(0)} \rightarrow \Psi_D$ is forbidden. The effect of $H^{(1)}$, which has no diagonal matrix elements in the $\Psi^{(0)}$ states, is to "rotate" the states through a small angle ϕ (illus. *b*), thereby admixing the symmetries, generating a weak transition for Ψ_F, and slightly shifting the energies. The ratio of the transition strengths S determines ϕ, which in turn fixes the admixing intensity and the level shifts. Diagonalizing the 2×2 matrix gives $S_F/S_A = \tan^2 \phi$; $(\Psi_F^{(0)} H^{(1)} \Psi_A^{(0)}) = (E_F - E_A) \sin \phi \cos \phi$; $(E_F - E_F^{(0)}) = -(E_A - E_A^0) = (E_F - E_A) \sin^2 \phi$. The perturbation solution (small ϕ) is often adequate. *See* PERTURBATION (QUANTUM MECHANICS).

Multiple-photon transitions. Because the M1 operator has no radial dependence in the long-wavelength limit, the single-photon transition between the $2S_{1/2}$ and $1S_{1/2}$ states in hydrogen is strongly inhibited; the lifetime should be on the order of 10^5 s. The very low-energy E1 transition to $2P_{1/2}^-$ (which lies below $2S_{1/2}^+$ by the Lamb shift energy), followed by E1 to ground, would take on the order of 10^8 s. The dominant process instead, with a lifetime on the order of 10^{-1} s, is the simultaneous emission of two photons, which share the energy to produce a continuous photon spectrum. In nuclei the two-photon $0^+ \rightarrow 0^+$ transition has also been observed, as well as E0 deexcitation by ejection of a penetrating atomic electron ("complete internal conversion") and by the closely related emission of an e^+-e^- pair. All of these processes are forbidden for $0^+ \leftrightarrow 0^-$, which has not been observed. *See* LASER SPECTROSCOPY; NONLINEAR OPTICS; RYDBERG CONSTANT.

LS-forbidden transitions. Early examples discovered in optical astronomy involve $6s6p\,^3P_0 \leftarrow 6s^2\ ^1S_0 \rightarrow 6s6p\,^3P_2$ in HgI. The first is absolutely forbidden, while the second would require that highly improbable M2, with wave-function corrections (since $s \rightarrow p$ would require E1). But the decoupling effect of the hyperfine interaction between valence electrons

and the nuclear magnetic moment generates a small $^3P_1$ amplitude in the $^3P_0$ state (the total angular momentum being $F = {}^1/_2$ in each case) and thereby, as in the illustration, an allowed but greatly retarded $^1S_0 \rightarrow {}^3P_1$ transition. A similar process occurs for the second case. Astronomical observation of such forbidden lines indicates that particle densities in the source are low; otherwise deexcitation would occur via collisions.

Spherical shell-model orbits in nuclei. The process of adding a neutron in orbit (l, j) to a nucleus, $A + n(l, j) \rightarrow B$, is realizable in (d, p) "direct" reactions, whose angular distribution determines l and whose magnitude gives the intrinsic neutron-capture probability. If one of the nuclear states involved is 0^+, the j value must equal the angular momentum of the other. For example, if A is describable as $(j^{2n})_{J=0^+}$, there should be large cross sections to $(j^{2n+1})_{J=j}$ and $((j^{2n})_0 j')_{J=j'}$, in which the particles in orbit j are themselves coupled to zero. Forbidden transitions to $(j^{2n+1})_{J=j'\neq j}$ are, however, often observed. They arise via admixtures (on the order of a few percent) of $((j^{2n})_0 j')$ in the final state which, by the analysis associated with the illustration, give measures of the orbital "purity." *See* NUCLEAR REACTION.

K selection rules in deformed nuclei. There is a J_z splitting of single-particle orbits in many heavy nuclei; this is produced by the large permanent quadrupole deformation of the nucleus which gives rise to a noncentral single-particle potential. K labels the angular-momentum component along the symmetry axis, whose rotation then generates from each intrinsic state a band of levels with $J \geq K$, all with the same parity. If two levels belong to bands with very different K values (specifically if $\Delta = |K_1 - K_2| - j$ is large, where the multipolarity j follows from the J^π values), electromagnetic transitions between them will be very highly forbidden. This is so formally because the j-multipole transition operator can only generate $|\Delta K| \leq j$, and more physically because the very different rotations involved in the two states generate large differences between all the corresponding orbits which play a role in the rotation. Empirically one encounters Δ values as high as 8 and finds retardations on the order of $10^{2\Delta}$, the process itself occurring via Coriolis-type admixtures. Similar multiparticle forbiddenness would obtain for transitions between states built on the two potential minima observed in fission studies for many nuclei. *See* NUCLEAR FISSION.

Molecular selection rules. The example of K selection rules in deformed nuclei involves both single-particle and rotational motion in these nuclei. Vibrational motion, especially quadrupole, can be found also. The classification of the motions derives from molecular physics and is most developed in that domain. To the extent that the molecular motions can be well separated, the angular momentum transferred in a transition must come from one or another of them. Along with the selection rules implied by this (the rule for harmonic vibrations is $n_\upsilon \rightarrow n_\upsilon \pm 1$), the various coupling possibilities and any special symmetries displayed by the intrinsic structure must be considered. Let N, Λ, Σ be the projection onto the symmetry axis (the line joining the two nuclei for the diatomic molecules to which the discussion is now restricted) of the collective rotational, the electronic orbital, and the electronic spin angular momenta. Λ and Σ may combine to Ω, the total intrinsic angular momentum along the symmetry axis; then the rotation gives $J = \Omega, \Omega + 1, \ldots$ (Hund's case A). Alternatively (case B), $\Lambda + N = K$, which combines with S to give $\mathbf{J} = \mathbf{S} + \mathbf{K}$, so that $|K - S| \leq J \leq K + S$. In both cases the supplementary rules for E1 (with natural extensions to other multipoles) that $\Delta\Lambda = 0, \pm 1$ and $\Delta S = 0$ are valid, while for case A, $\Delta\Sigma = 0$, $\Delta\Omega = 0 \pm 1$, and for case B, $\Delta\mathbf{K} = 1$, these being valid if the spin-orbit interactions are weak.

The special symmetries of the intrinsic structure refer to its behavior under a reflection through the symmetry center and through a plane containing the symmetry axis. The first operation, $\mathscr{P}$, defines the intrinsic parity, labeled as g (for $\pi = +1$) and u ($\pi = -1$). The second, $\mathscr{S}$, defines a symmetry $\pm$; in nuclear physics one commonly uses $\mathscr{R} = \mathscr{P}\mathscr{S}$, with eigenvalue $r = \pm 1$, which generates a rotation through π radians about an axis normal to the symmetry axis. For homonuclear molecules the interchange of the two nuclei defines states as symmetric (s) or antisymmetric (a). Then for E1 the rules $g \leftrightarrow u$, $+ \leftrightarrow -$ are valid, and with homonuclear molecules $s \rightarrow s$, $a \rightarrow a$. *See* MOLECULAR STRUCTURE AND SPECTRA.

Further symmetries. The isospin symmetry of the elementary particles is almost conserved, being broken by electromagnetic and weak interactions. It is described by the group SU(2), of unimodular unitary transformations in two dimensions. Since the SU(2) algebra is identical with that of the angular momentum SO(3), isospin behaves like angular momentum with its three generators $\mathbf{T}$ replacing $\mathbf{J}$. For a nucleon $t = {}^1/_2$, the values $t_z = {}^1/_2, -{}^1/_2$ may be assigned for protons (p) and neutrons (n), respectively. Then, for nuclear states, $T_z = (Z - N)/2$ (where Z is the atomic number and N is the neutron number) and $T\ (\geq |T_z|)$ is an almost good quantum number. The argument associated with the illustration can often be used to measure the isospin admixing; for example, with the pair of decays $^{15}\mathrm{N}\ (T = {}^1/_2, {}^3/_2) \rightarrow {}^{14}\mathrm{N}\ (T = 0 + n)$, the second of which is forbidden by the selection rule $\Delta T = {}^1/_2$, one finds about 4% admixing. Since the nucleon charge is $|e|\{{}^1/_2 + t_z\}$, the electromagnetic transition operators split into an isoscalar part (for which $\Delta\mathbf{T} = 0$) and an isovector ($\Delta\mathbf{T} = 1$). Since the isoscalar part of the (long-wavelength) E1 operator

$$|e| \left\{ {}^1/_2 \sum_i \mathbf{r}(i) + \sum_i t_z(i)\mathbf{r}(i) \right\}$$

is ineffective, being proportional to the center-of-mass vector, and the inversion rule described earlier forbids isovector transitions in self-conjugate nuclei ($T_z = 0$, $N = Z$), it follows that E1 transitions between states of the same T are forbidden in such nuclei.

The isospin group is a subgroup of SU(3) which defines a more complex fundamental symmetry of the elementary particles. Two of its eight generators commute, giving two additive quantum numbers, T_z and strangeness S' (or, equivalently, charge and hypercharge). The strangeness is conserved ($\Delta S' = 0$) for strong and electromagnetic, but not for weak, interactions. The selection rules and combination laws for SU(3) and its many extensions, and the quark-structure ideas underlying them, correlate an enormous amount of information and make many predictions about the elementary particles. *See* BARYON; ELEMENTARY PARTICLE; MESON; QUARKS; UNITARY SYMMETRY.

The fundamental beta-decay transition is $n \to p + e^- + \overline{\nu}$. For the Gamow-Teller decay mode (electron and antineutrino in a triplet state) and the Fermi mode (singlet), the transition operators are respectively components of

$$\sum_i \mathbf{s}(i)\ \mathbf{t}(i)$$

and of $\mathbf{T}$. When these decays are realized in a nucleus, in which also $p \to n + e^+ + \nu$ may occur, the rules $\Delta\pi = 0$, $\Delta\mathbf{T} = 1$, $\Delta T_z = \pm 1$ clearly hold, with $\Delta\mathbf{J} = 1, 0$ for Gamow-Teller and Fermi, respectively. But for Fermi decay there is also the very strong rule (which includes $\Delta T = 0$) that the decay takes place only to a single state, the isobaric analog, which differs from the parent only in T_z, the transition rate being then easily calculable. This special feature, analogous to that for M1 transitions in *LS* coupling, arises from the fact that the transition operator is a generator of an (almost) good symmetry. Besides superallowed Fermi transitions between $J = 0$ analog states (forbidden for Gamow-Teller), there are also forbidden Fermi decays between nonanalog states which, as in the illustration, arise from isospin admixing and have been used to give measures for it. Special rules, by no means as well satisfied, arise also for Gamow-Teller decays from the fact that the 15 components of the basic transition operators are the generators of an SU(4) group whose symmetry is fairly good in light nuclei. (For these same nuclei a different SU(3) realization than above gives a model for rotational bands.) Besides the allowed transitions, there is also a hierarchy of forbidden transitions in nuclei which, through recoil (retardation) effects, involve the orbital angular momentum and have therefore less restrictive selection rules. *See* NEUTRINO; RADIOACTIVITY.

A great variety of other groups have been introduced to define relevant symmetries for atoms, molecules, nuclei, and elementary particles. They all have their own selection rules, representing one aspect of the symmetries of nature. J. B. French

Bibliography. P. F. Bernath, *Spectra of Atoms and Molecules*, 1995; A. Bohr and B. R. Mottelson, *Nuclear Structure*, vol. 2, 1975; E. U. Condon and H. Odabasi, *Atomic Structure*, 1980; R. D. Cowan, *The Theory of Atomic Structure and Spectra*, 1981; K. S. Krane, *Introductory Nuclear Physics,* 1987.

Selectivity

The ability of a radio receiver to separate a desired signal frequency from other signal frequencies, some of which may differ only slightly from the desired value. Selectivity is achieved by using tuned circuits that are sharply peaked and by increasing the number of tuned circuits. With a sharply peaked circuit, the output voltage falls off rapidly for frequencies lower or higher than that to which the circuit is tuned. *See* Q (ELECTRICITY); RADIO RECEIVER; RESONANCE (ALTERNATING-CURRENT CIRCUITS). John Markus

Selenium

A chemical element, Se, atomic number 34, atomic weight 78.96. The properties of this element are similar to those of tellurium. *See* PERIODIC TABLE; TELLURIUM.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Selenium burns in air with a blue flame to give selenium dioxide, SeO_2. The element also reacts directly with a variety of metals and nonmetals, including hydrogen and the halogens. Nonoxidizing acids fail to react with selenium, but nitric acid, concentrated sulfuric acid, and strong alkali hydroxides dissolve the element. The only important compound of selenium with hydrogen is hydrogen selenide, H_2Se, a colorless flammable gas possessing a distinctly unpleasant odor, and a toxicity greater and a thermal stability less than that of hydrogen sulfide. Selenium oxyhalide, $SeOCl_2$, is a colorless liquid widely used as a nonaqueous solvent. The oxybromide, $SeOBr_2$, is an orange solid having chemical properties similar to those of $SeOCl_2$. The oxyflouride, $SeOF_2$, a colorless liquid with a pungent smell, reacts with water, glass, and silicon, and also forms additional compounds. Compounds in which C-Se bonds appear are numerous and vary from the simple selenols, RSeH, to molecules exhibiting biological activity such as selenoamino acids and selenopeptides. *See* ORGANOSELENIUM COMPOUND.

The abundance of this widely distributed element in the Earth's crust is estimated to be about 7×10^{-5}% by weight, occurring as the selenides of heavy elements and to a limited extent as the free element in association with elementary sulfur. Examples

of the variety of selenide minerals are berzelianite (Cu_2Se), eucairite (AgCuSe), and jermoite [$As(S,Se)_2$]. Selenium minerals do not occur in sufficient quantity to be useful as commercial sources of the element.

Major uses of selenium include the photocopying process of xerography, which depends on the light sensitivity of thin films of amorphous selenium, the decolorization of glasses tinted by the presence of iron compounds, and use as a pigment in plastics, paints, enamels, glass, ceramics, and inks. Selenium is also employed in photographic exposure meters and solar cells and as a metallurgical additive to improve the machinability of certain steels. Minor uses include application as a nutritional additive for numerous animal species, use in photographic toning, metal-finishing operations, metal plating, high-temperature lubricants, and as catalytic agents, particularly in the isomerization of certain petroleum products. John W. George

The biological importance of selenium is well established, as all classes of organisms metabolize selenium. In humans and other mammals, serious diseases arise from either excessive or insufficient dietary selenium. The toxic effects of selenium have long been known, particularly for grazing animals. In soils with high selenium content, some plants accumulate large amounts of selenium. Animals that ingest these selenium-accumulating plants develop severe toxic reactions. *See* SOIL CHEMISTRY.

Although toxic at high levels, selenium is an essential micronutrient for mammalian species. The accepted minimum daily requirement of selenium for adult humans is 70 micrograms. Many types of food provide selenium, particularly seafood, meats, grains, and the onion family. Mammals and birds require selenium for production of the enzyme glutathione peroxidase, which protects against oxidation-induced cancers. Other seleno-proteins of unknown function are found in mammalian blood, various tissues, and spermatozoa. *See* AMINO ACIDS.

Milton J. Axley; Thressa G. Stadtman

Bibliography. M. A. Brook, *Silicon in Organic, Organometallic and Polymer Chemistry*, Wiley-Interscience, 1999; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; P. Jutzi and U. Schubert (eds.), *Silicon Chemistry: From the Atom to Extended Systems*, Wiley, 2003; U. Schubert (ed.), *Silicon Chemistry*, Springer, 1999.

Seligeriales

An order of the true mosses (Bryidae), consisting of one family and five or six genera. The plants grow on rocks, and may be exceedingly small and gregarious to moderate in size and tufted. Though the gametophytes resemble the Dicranales, the double structure of the peristome is distinctive.

The stems are erect and simple or forked. The linear to lance-subulate leaves have a single costa which often fills the subula, and short, smooth cells. The operculate capsules are terminal, and the peristome is usually present; though apparently single, it consists of an endostome of 16 narrow, smooth or rarely papillose teeth often cleft at the tips, the outer peristome being reduced to a membrane adherent to the endostome. The calyptra is generally cucullate. Chromosome numbers of 11, 12, 13, and 14 are known. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA; DICRANALES.

Howard Crum

Semaeostomeae

An order of the class Scyphozoa including most of the common medusae, such as *Aurelia aurita*. The umbrella of these medusae is more flat than high and is usually domelike. The margin of the umbrella is divided into many lappets. Sensory organs are situated between the lappets. The tentacles arise between the lappets in *Pelagia* (see **illus.**), on the exumbrella in *Aurelia*, and on the subumbrella in *Cyanea*. They are generally well developed and very long except in a few forms such as *Aurelia*. The oral arms are well developed and either curtainlike, in *Cyanea*, or leaflike, in *Aurelia*. The stomach is cruciform, and several radial pockets or many radial canals which are sometimes connected with each other to form a network issue from it. The life history of this group shows the typical alternation of generations, with forms passing through the stages of planula, scyphopolyp, strobila, and ephyra. The Semaeostomeae are distributed mostly in temperate zones and are generally coastal forms with a few exceptions such as *Pelagia* and *Sanderia*. Some of them, such as *Dactylometra* and *Sanderia*, are known to have violent poison on their tentacles. This poison is not only harmful to the skin but also to the nerves, which are paralyzed temporarily. *Cyanea* is used as bait in fishing. Several fossils of this group were found in the strata of the Jurassic Period. *See* SCYPHOZOA. Tohru Uchida

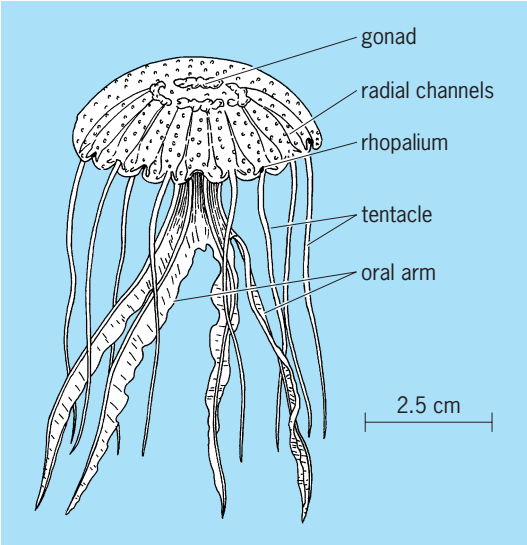

***Pelagia*, in the order Semaeostomeae. (*After L. Hyman, The Invertebrates, vol. 1, McGraw-Hill, 1940*)**

Semiconductor

A solid crystalline material whose electrical conductivity is intermediate between that of a metal and an insulator. Semiconductors exhibit conduction properties that may be temperature-dependent, permitting their use as thermistors (temperature-dependent resistors), or voltage-dependent, as in varistors. By making suitable contacts to a semiconductor or by making the material suitably inhomogeneous, electrical rectification and amplification can be obtained. Semiconductor devices, rectifiers, and transistors have replaced vacuum tubes almost completely in low-power electronics, making it possible to save volume and power consumption by orders of magnitude. In the form of integrated circuits, they are vital for complicated systems. The optical properties of a semiconductor are important for the understanding and the application of the material. Photodiodes, photoconductive detectors of radiation, injection lasers, light-emitting diodes, solar-energy conversion cells, and so forth are examples of the wide variety of optoelectronic devices. *See* INTEGRATED CIRCUITS; LASER; LIGHT-EMITTING DIODE; PHOTODIODE; PHOTOELECTRIC DEVICES; SEMICONDUCTOR DIODE; SEMICONDUCTOR RECTIFIER; THERMISTOR; TRANSISTOR; VARISTOR.

Conduction in Semiconductors

The electrical conductivity of semiconductors ranges from about 10^3 to 10^{-9} $\text{ohm}^{-1}\ \text{cm}^{-1}$, as compared with a maximum conductivity of 10^7 for good conductors and a minimum conductivity of 10^{-17} $\text{ohm}^{-1}\ \text{cm}^{-1}$ for good insulators. *See* ELECTRIC INSULATOR; ELECTRICAL CONDUCTIVITY OF METALS.

The electric current is usually due only to the motion of electrons, although under some conditions, such as very high temperatures, the motion of ions may be important. The basic distinction between conduction in metals and in semiconductors is made by considering the energy bands occupied by the conduction electrons. *See* IONIC CRYSTALS.

A crystalline solid consists of a large number of atoms brought together into a regular array called a crystal lattice. The electrons of an atom can each have certain energies, so-called energy levels, as predicted by quantum theory. Because the atoms of the crystal are in close proximity, the electron orbits around different atoms overlap to some extent, and the electrons interact with each other; consequently the sharp, well-separated energy levels of the individual electrons actually spread out into energy bands. Each energy band is a quasicontinuous group of closely spaced energy levels. *See* BAND THEORY OF SOLIDS.

At absolute zero temperature, the electrons occupy the lowest possible energy levels, with the restriction that at most two electrons with opposite spin may be in the same energy level. In semiconductors and insulators, there are just enough electrons to fill completely a number of energy bands, leaving the rest of the energy bands empty. The highest filled energy band is called the valence band. The next higher band, which is empty at absolute zero temperature, is called the conduction band. The conduction band is separated from the valence band by an energy gap which is an important characteristic of the semiconductor. In metals, the highest energy band that is occupied by the electrons is only partially filled. This condition exists either because the number of electrons is not just right to fill an integral number of energy bands or because the highest occupied energy band overlaps the next higher band without an intervening energy gap. The electrons in a partially filled band may acquire a small amount of energy from an applied electric field by going to the higher levels in the same band. The electrons are accelerated in a direction opposite to the field and thereby constitute an electric current. In semiconductors and insulators, the electrons are found only in completely filled bands, at low temperatures. In order to increase the energy of the electrons, it is necessary to raise electrons from the valence band to the conduction band across the energy gap. The electric fields normally encountered are not large enough to accomplish this with appreciable probability. At sufficiently high temperatures, depending on the magnitude of the energy gap, a significant number of valence electrons gain enough energy thermally to be raised to the conduction band. These electrons in an unfilled band can easily participate in conduction. Furthermore, there is now a corresponding number of vacancies in the electron population of the valence band. These vacancies, or holes as they are called, have the effect of carriers of positive charge, by means of which the valence band makes a contribution to the conduction of the crystal. *See* HOLE STATES IN SOLIDS.

The type of charge carrier, electron or hole, that is in largest concentration in a material is sometimes called the majority carrier and the type in smallest concentration the minority carrier. The majority carriers are primarily responsible for the conduction properties of the material. Although the minority carriers play a minor role in electrical conductivity, they can be important in rectification and transistor actions in a semiconductor.

Electron distribution. The probability f for an energy level E to be occupied by an electron is given by the Fermi-Dirac distribution function (1), where k

$$f = \left[1 + \exp\left(\frac{E - W}{kT}\right)\right]^{-1} \tag{1}$$

is the Boltzmann constant and T the absolute temperature. The parameter W is the Fermi energy level; an energy level at W has a probability of 1/2 to be occupied by an electron. The Fermi level is determined by the distribution of energy levels and the total number of electrons. *See* FERMI-DIRAC STATISTICS.

In a semiconductor, the number of conduction electrons is normally small compared with the number of energy levels in the conduction band, and the probability for any energy level to be occupied is small. Under such a condition, the concentration

of conduction electrons is given by Eq. (2), where

$$N_n = \frac{2}{h^3}(2\pi m_n kT)^{3/2} \exp\left(\frac{W - E_c}{kT}\right) \qquad (2)$$

h is Planck's constant, E_c is the lowest energy of the conduction band, and m_n is called the effective mass of conduction electrons. The effective mass is used in place of the actual mass to correct the coefficient in the equation and to bring the results in line with experimental observations. This correction is necessary because the theory leading to these equations is based upon electrons moving in a field free space, which is not the exact picture. The electrostatic Coulomb potential throughout the crystal is varying in a periodic manner, the variation being due to the electric fields around the atomic centers. The concentration of holes in the valence band is given by Eq. (3), where m_p is the effective mass of a hole

$$N_p = \frac{2}{h^3}(2\pi m_p kT)^{3/2} \exp\left(\frac{E_v - W}{kT}\right) \qquad (3)$$

and E_v is the highest energy of the valence band.

Mobility of carriers. The velocity acquired by charge carriers per unit strength of applied electric field is called the mobility of the carriers. The velocity in question is the so-called drift velocity in the direction of the force exerted on the carriers by the applied field. It is added to the random thermal velocity. In semiconductors the carrier mobility normally ranges from 10^2 to 10^5 $cm^2/(s)(V)$. A material's conductivity is the product of the charge, the mobility, and the carrier concentration.

Electrons in a perfectly periodic potential field can be accelerated freely. Impurities, physical defects in the structure, and thermal vibrations of the atoms disturb the periodicity of the potential field in the crystal, thereby scattering the moving carriers. It is the resistance produced by this scattering that limits the carriers to only a drift velocity under the steady force of an applied field.

Intrinsic semiconductors. A semiconductor in which the concentration of charge carriers is characteristic of the material itself rather than of the content of impurities and structural defects of the crystal is called an intrinsic semiconductor. Electrons in the conduction band and holes in the valence band are created by thermal excitation of electrons from the valence to the conduction band. Thus an intrinsic semiconductor has equal concentrations of electrons and holes. The intrinsic carrier concentration N_i is determined by Eq. (4), where E_g is the energy

$$N_i = \frac{2}{h^3}(2\pi kT)^{3/2}(m_n m_p)^{3/4} \exp\left(-\frac{E_g}{2kT}\right) \qquad (4)$$

gap. The carrier concentration, and hence the conductivity, is very sensitive to temperature and depends strongly on the energy gap. The energy gap ranges from a fraction of 1 eV to several electronvolts. A material must have a large energy gap to be an insulator.

Extrinsic semiconductors. Typical semiconductor crystals such as germanium and silicon are formed by an ordered bonding of the individual atoms to form the crystal structure. The bonding is attributed to the valence electrons which pair up with valence electrons of adjacent atoms to form so-called shared pair or covalent bonds. These materials are all of the quadrivalent type; that is, each atom contains four valence electrons, all of which are used in forming the crystal bonds. *See* CRYSTAL STRUCTURE.

Atoms having a valence of +3 or +5 can be added to a pure or intrinsic semiconductor material with the result that the +3 atoms will give rise to an unsatisfied bond with one of the valence electrons of the semiconductor atoms, and +5 atoms will result in an extra or free electron that is not required in the bond structure. Electrically, the +3 impurities add holes and the +5 impurities add electrons. They are called acceptor and donor impurities, respectively. Typical valence +3 impurities used are boron, aluminum, indium, and gallium. Valence +5 impurities used are arsenic, antimony, and phosphorus.

Semiconductor material "doped" or "poisoned" by valence +3 acceptor impurities is termed *p*-type, whereas material doped by valence +5 donor material is termed *n*-type. The names are derived from the fact that the holes introduced are considered to carry positive charges and the electrons negative charges. The number of electrons in the energy bands of the crystal is increased by the presence of donor impurities and decreased by the presence of acceptor impurities. Let N be the concentration of electrons in the conduction band and let P be the hole concentration in the valence band. For a given semiconductor, the relation $NP = N_i^2$ holds, independent of the presence of impurities. The effect of donor impurities tends to make N larger than P, since the extra electrons given by the donors will be found in the conduction band even in the absence of any holes in the valence band. Acceptor impurities have the opposite effect, making P larger than N. *See* ACCEPTOR ATOM; DONOR ATOM.

At sufficiently high temperatures, the intrinsic carrier concentration becomes so large that the effect of a fixed amount of impurity atoms in the crystal is comparatively small and the semiconductor becomes intrinsic. When the carrier concentration is predominantly determined by the impurity content, the conduction of the material is said to be extrinsic. There may be a range of temperature within which the impurity atoms in the material are practically all ionized; that is, they supply a maximum number of carriers. Within this temperature range, the so-called exhaustion range, the carrier concentration remains nearly constant. At sufficiently low temperatures, the electrons or holes that are supplied by the impurities become bound to the impurity atoms. The concentration of conduction carriers will then decrease rapidly with decreasing temperature, according to either $\exp(-E_i/kT)$ or $\exp(-E_i/2kT)$, where E_i is the ionization energy of the dominant impurity.

Physical defects in the crystal structure may have similar effects as donor or acceptor impurities. They can also give rise to extrinsic conductivity.

An isoelectronic impurity, that is, an atom which has the same number of valence electrons as the host atom, does not bind individual carriers as strongly as a donor or an acceptor impurity. However, it may show an appreciable binding for electron hole pairs, excitons, and thereby have important effects on the properties. An example is nitrogen substituting for phosphorus in gallium phosphide; the impurity affects the luminescence of the material.

Hall effect. Whether a given sample of semiconductor material is *n*- or *p*-type can be determined by observing the Hall effect. If an electric current is caused to flow through a sample of semiconductor material and a magnetic field is applied in a direction perpendicular to the current, the charge carriers are crowded to one side of the sample, giving rise to an electric field perpendicular to both the current and the magnetic field. This development of a transverse electric field is known as the Hall effect. The field is directed in one or the opposite direction depending on the sign of the charge of the carrier. *See* HALL EFFECT.

The magnitude of the Hall effect gives an estimate of the carrier concentration. The ratio of the transverse electric field strength to the product of the current and the magnetic field strength is called the Hall coefficient, and its magnitude is inversely proportional to the carrier concentration. The coefficient of proportionality involves a factor which depends on the energy distribution of the carriers and the way in which the carriers are scattered in their motion. However, the value of this factor normally does not differ from unity by more than a factor of 2. The situation is more complicated when more than one type of carrier is important for the conduction. The Hall coefficient then depends on the concentrations of the various types of carriers and their relative mobilities.

The product of the Hall coefficient and the conductivity is proportional to the mobility of the carriers when one type of carrier is dominant. The proportionality involves the same factor which is contained in the relationship between the Hall coefficient and the carrier concentration. The value obtained by taking this factor to be unity is referred to as the Hall mobility.

Materials and Their Preparation

The group of chemical elements which are semiconductors includes germanium, silicon, gray (crystalline) tin, selenium, tellurium, and boron.

Elemental semiconductors. Germanium, silicon, and gray tin belong to group 14 (group IV) of the periodic table and have crystal structures similar to that of diamond. Germanium and silicon are two of the best-known semiconductors. They are used extensively in devices such as rectifiers and transistors. Gray tin is a form of tin which is stable below 13°C (55°F). White tin, which is stable at higher temperatures, is metallic. Gray tin has a small energy gap and a rather large intrinsic conductivity, about 5×10^3 ohm^{-1} cm^{-1} at room temperature. The *n*-type and *p*-type gray tins can be obtained by adding aluminum and antimony, respectively.

Selenium and tellurium both have a similar structure, consisting of spiral chains located at the corners and centers of hexagons. The structure gives rise to anisotropy of the properties of single crystals; for example, the electrical resistivity of tellurium along the direction of the chains is about one-half the resistivity perpendicular to this direction. Selenium has been widely used in the manufacture of rectifiers and photocells.

Semiconducting compounds. A large number of compounds are known to be semiconductors. Copper(I) oxide (Cu_2O) and mercury(II) indium telluride ($HgIn_2Te_4$) are examples of binary and ternary compounds. The series zinc sulfide (ZnS), zinc selenide (ZnSe), and zinc telluride (ZnTe), and the series zinc selenide (ZnSe), cadmium selenide (CdSe), and mercury(II) selenide (HgSe) are examples of binary compounds consisting of a given element in combinations with various elements of another column in the periodic table. The series magnesium antimonide (Mg_2Sb_2), magnesium telluride (MgTe), and magnesium iodide (MgI_2) is an example of compounds formed by a given element with elements of various other columns in the periodic table. *See* PERIODIC TABLE.

A group of semiconducting compounds of the simple type AB consists of elements from columns symmetrically placed with respect to column 14 of the periodic table. Indium antimonide (InSb), cadmium telluride (CdTe), and silver iodide (AgI) are examples of III–V, II–IV, and I–VI compounds, respectively. The various III–V compounds are being studied extensively, and many practical applications have been found for these materials. Some of these compounds have the highest carrier mobilities known for semiconductors. The compounds have zincblende crystal structure which is geometrically similar to the diamond structure possessed by the elemental semiconductors, germanium and silicon, of column 14, except that the four nearest neighbors of each atom are atoms of the other kind. The II–VI compounds, zinc sulfide (ZnS) and cadmium sulfide (CdS), are used in photoconductive devices. Zinc sulfide is also used as a luminescent material. *See* LUMINESCENCE; PHOTOCONDUCTIVITY.

Binary compounds of the group lead sulfide (PbS), lead selenide (PbSe), and lead telluride (PbTe) are sensitive in photoconductivity and are used as detectors of infrared radiation. The compounds, bismuth telluride (Bi_2Te_3) and bismuth selenide (Bi_2Se_3), consisting of heavy atoms, are found to be good materials for thermocouples used for refrigeration or for conversion of heat to electrical energy. *See* THERMOELECTRICITY.

The metal oxides usually have large energy gaps. Thus pure oxides are usually insulators of high resistivity. However, it may be possible to introduce into some of the oxides impurities of low ionization energies and thus obtain relatively good extrinsic conduction. Copper(I) oxide (Cu_2O) was one of the first semiconductors used for rectifiers and photocells: extrinsic *p*-type conduction is obtained by producing

an excess of oxygen over the stoichiometric composition, that is, the 2-to-1 ratio of copper atoms to oxygen atoms. A number of oxide semiconductors can be obtained by replacing some of the normal metal atoms with metal atoms of one more or less valency. The method is called controlled valence. An example of such a semiconductor is nickel oxide containing lithium.

Some compounds with rare-earth or transition-metal ions in their composition, such as EuTe and NiS_2, are semiconductors with magnetic properties. Another interesting type of semiconductor is characterized by layered structures. The interaction within a layer is significantly stronger than that between layers. A number of semiconductors of this type are known, such as PbI_2, GaSe, and various transition-metal dichalcogenides such as $SnSe_2$ and MoS_2.

Preparation of materials. The properties of semiconductors are extremely sensitive to the presence of impurities. It is therefore desirable to start with the purest available materials and to introduce a controlled amount of the desired impurity. The zone-refining method is often used for further purification of obtainable materials. The floating zone technique can be used, if feasible, to prevent any contamination of molten material by contact with crucible. *See* ZONE REFINING.

For basic studies as well as for many practical applications, it is desirable to use single crystals. Various methods are used for growing crystals of different materials. For many semiconductors, including germanium, silicon, and the III–V compounds, the Czochralski method is commonly used. The method of condensation from the vapor phase is used to grow crystals of a number of semiconductors, for instance, selenium and zinc sulfide. For materials of high melting points, such as various metal oxides, the flame fusion or Vernonil method may be used. *See* CRYSTAL GROWTH.

The introduction of impurities, or doping, can be accomplished by simply adding the desired quantity to the melt from which the crystal is grown. Normally, the impurity has a small segregation coefficient, which is the ratio of equilibrium concentrations in the solid and the liquid phases of the material. In order to obtain a desired impurity content in the crystal, the amount added to the melt must give an appropriately larger concentration in the liquid. When the amount to be added is very small, a preliminary ingot is often made with a larger content of the doping agent; a small slice of the ingot is then used to dope the next melt accurately. Impurities which have large diffusion constants in the material can be introduced directly by holding the solid material at an elevated temperature while this material is in contact with the doping agent in the solid or the vapor phase.

A doping technique, ion implantation, has been developed and used extensively. The impurity is introduced into a layer of semiconductor by causing a controlled dose of highly accelerated impurity ions to impinge on the semiconductor. *See* ION IMPLANTATION.

An important subject of scientific and technological interest is amorphous semiconductors. In an amorphous substance the atomic arrangement has some short-range but no long-range order. The representative amorphous semiconductors are selenium, germanium, and silicon in their amorphous states, and arsenic and germanium chalcogenides, including such ternary systems as Ge-As-Te. Some amorphous semiconductors can be prepared by a suitable quenching procedure from the melt. Amorphous films can be obtained by vapor deposition.

Rectification in Semiconductors

In semiconductors, narrow layers can be produced which have abnormally high resistances. The resistance of such a layer is nonohmic; it may depend on the direction of current, thus giving rise to rectification. Rectification can also be obtained by putting a thin layer of semiconductor or insulator material between two conductors of different material.

Barrier layer. A narrow region in a semiconductor which has an abnormally high resistance is called a barrier layer. A barrier may exist at the contact of the semiconductor with another material, at a crystal boundary in the semiconductor, or at a free surface of the semiconductor. In the bulk of a semiconductor, even in a single crystal, barriers may be found as the result of a nonuniform distribution of impurities. The thickness of a barrier layer is small, usually 10^{-3} to 10^{-5} cm.

A barrier is usually associated with the existence of a space charge. In an intrinsic semiconductor, a region is electrically neutral if the concentration n of conduction electrons is equal to the concentration p of holes. Any deviation in the balance gives a space charge equal to $e(p - n)$, where e is the charge on an electron. In an extrinsic semiconductor, ionized donor atoms give a positive space charge and ionized acceptor atoms give a negative space charge. Let N_d and N_a be the concentrations of ionized donors and acceptors, respectively. The space charge is equal to $e(p - n + N_d - N_a)$.

A space charge is associated with a variation of potential. A drop in potential $-\Delta V$ increases the potential energy of an electron by $e\Delta V$; consequently every electronic energy level in the semiconductor is shifted by this amount. With a variation of potential, the electron concentration varies proportionately to $\exp(eV/kT)$ and the hole concentration varies as $\exp(-eV/kT)$. A space charge is obtained if the carriers, mainly the majority carriers, fail to balance the charge of the ionized impurities.

A conduction electron in a region where the potential is higher by ΔV must have an excess energy of $e\Delta V$ in order for it to have the minimum energy on reaching the low potential region. Electrons with less energy cannot pass over to the low potential region. Thus a potential variation presents a barrier to the flow of electrons from high to low potential regions. It also presents a barrier to the flow of holes from low to high potential regions.

Surface barrier. A thin layer of space charge and a resulting variation of potential may be produced

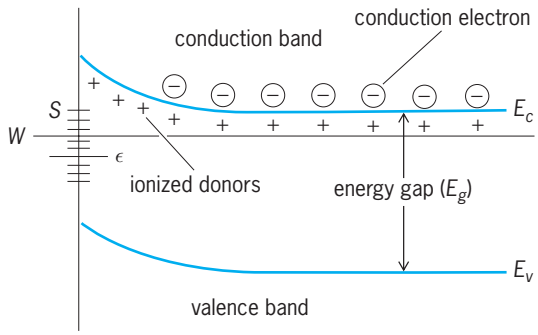

Fig. 1. Energy diagram of a surface barrier as employed in an *n*-type semiconductor.

at the surface of a semiconductor by the presence of surface states. Electrons in the surface states are bound to the vicinity of the surface, and the energy levels of surface states may lie within the energy gap. Surface states may arise from the adsorption of foreign atoms. Even a clean surface may introduce states which do not exist in the bulk material, simply by virtue of being the boundary of the crystal.

The surface is electrically neutral when the surface states are filled with electrons up to a certain energy level ϵ in the energy gap E_g, which is the energy difference between the bottom of the conduction band E_c and the top of the valence band E_v. If the Fermi level W in the bulk semiconductor lies higher in the energy gap, more surface states would be filled, giving the surface a negative charge. As a result the potential drops near the surface and the energy bands are raised for *n*-type material (**Fig. 1**). With the rise of the conduction band, the electron concentration is reduced and a positive space charge due to ionized donors is obtained. The amount of positive space charge is equal to the negative surface charge given by the electrons in the surface states between ϵ and the Fermi level.

Contact barrier. The difference between the potential energy E_o of an electron outside a material and the Fermi level in the material is called the work function of the material. **Figure 2** shows the energy diagram for a metal and a semiconductor, the work functions of which differ by eV. Upon connecting the two bodies electrically, charge is transferred between them so that the potential of the semiconductor is raised relative to that of the metal; that is, the electron energy levels in the semiconductor are lowered. Equilibrium is established when the Fermi level is the same in the two bodies. In this case, the metal is charged negatively and the semiconductor is charged positively. The negative charge on the metal is concentrated close to the surface, as is expected in good conductors. The positive charge on the semiconductor is divided between the increase of space charge in an extension of the barrier and the depopulation of some of the surface states. The charging of the semiconductor is brought about by a change of eV_2 in the barrier height ϕ. The sum of eV_2 and the potential energy variation eV_1 in the space between the two bodies is equal to the original difference eV between the work functions.

With decreasing separation between the two bodies, the division of eV will be in favor of eV_2. However, if there is a very large density of states, a small eV_2 gives a large surface charge on the semiconductor due to the depopulation of surface states.

It is possible that eV_2 is limited to a small value even at the smallest separation, of the order of an interatomic distance in solids. In such cases, the barrier height remains nearly equal to the value ϕ of the free surface, irrespective of the body in contact. This situation has been found in germanium and silicon rectifiers. Before the explanation was given by J. Bardeen, who postulated the existence of surface states, it had been assumed that the height of a contact barrier was equal to the difference of the work functions.

The understanding and the application of metal-semiconductor contacts have been extended to various kinds of contacts, such as that between different semiconductors, heterojunctions, and metal oxide semiconductor (MOS) junctions.

Single-carrier theory. The phenomenon of rectification at a crystal barrier can be described according to the role played by the carriers. Where the conduction property of the rectifying barrier is determined primarily by the majority carriers, the single-carrier theory is employed. Such cases are likely to be found in semiconductors with large energy gaps, for instance, oxide semiconductors. **Figure 3** shows the energy diagrams of metal-semiconductor contact rectifiers under conditions of equilibrium. The potential variation in the semiconductor is such as to reduce the majority carrier concentration near

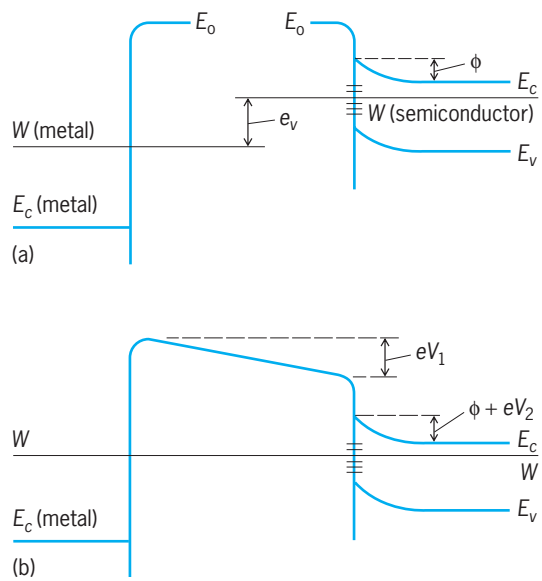

Fig. 2. Energy diagram for a metal (left) and an *n*-type semiconductor (right). E_0 is the potential energy of an electron outside the material, E_c is the energy at the bottom of the conduction band, and Ev is the energy at the top of the valence band. (*a*) Semiconductor and metal isolated. (*b*) Semiconductor and metal in electrical contact, $eV_1 + eV_2 = eV$.

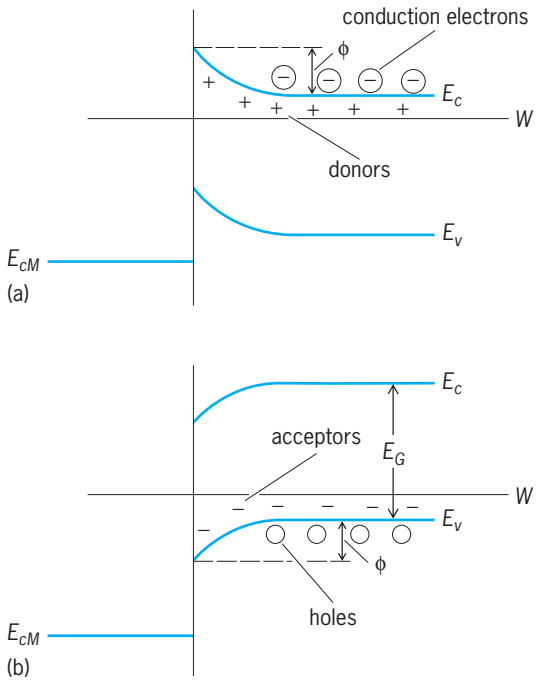

Fig. 3. Energy diagrams of a rectifying contact between a metal and a semiconductor: (*a*) *n*-type semiconductor; (*b*) *p*-type semiconductor.

the contact. If the energy bands were to fall in the case of an *n*-type semiconductor or to rise in the case of a *p*-type semiconductor, the majority carrier concentration would be enhanced near the contact, and the contact would not present a large and rectifying resistance. It is clear that in the cases shown in Fig. 3, the minority carrier concentration increases near the contact. However, if the energy gap is large, the minority carrier concentration is normally very small, and the role of minority carriers may be still negligible even if the concentration is increased.

Under equilibrium conditions, the number of carriers passing from one body to the other is balanced by the number of carriers crossing the contact in the opposite direction, and there is no net current. The carriers crossing the contact in either direction must have sufficient energies to pass over the peak of the barrier. The situations under applied voltages are shown in **Fig. 4** for the case of an *n*-type semiconductor. When the semiconductor is made positive, its energy bands are depressed and the height of the potential barrier is increased, as shown in Fig. 4*a*. Fewer electrons in the semiconductor will be able to cross over into the metal, whereas the flow of electrons across the contact from the metal side remains unchanged. Consequently, there is a net flow of electrons from the metal to the semiconductor. The flow of electrons from the metal side is the maximum net flow obtainable. With increasing voltage, the current saturates and the resistance becomes very high. Figure 4*b* shows the situation when the semiconductor is negative under the applied voltage. The energy bands in the semiconductor are raised. The flow of electrons from the semiconductor to the metal is increased, since electrons of lower energy are able to go over the peak of the barrier. The result is a net flow of electrons from the semiconductor to the metal. There is no limit to the flow in this case. In fact, the electron current increases faster than the applied voltage because there are increasingly more electrons at lower energies. The resistance decreases, therefore, with increasing voltage. The direction of current for which the resistance is low is called the forward direction, while the opposite direction is called the reverse or blocking direction. A general expression for the current can be written in the form of Eq. (5), where j is the current density, n is

$$j = enC\left(\exp\frac{-\phi}{kT}\right)\left[\exp\left(\frac{eV}{kT}\right) - 1\right] \qquad (5)$$

the carrier concentration in the bulk of the semiconductor, ϕ is the barrier height, and V is the applied voltage taken as positive in the forward direction. The factor C depends on the theory appropriate for the particular case.

Diffusion theory. When there is a variation of carrier concentration, a motion of the carriers is produced by diffusion in addition to the drift determined by the mobility and the electric field. The transport of carriers by diffusion is proportional to the carrier concentration gradient and the diffusion constant. The diffusion constant is related to the mobility, and both are determined by the scattering suffered by

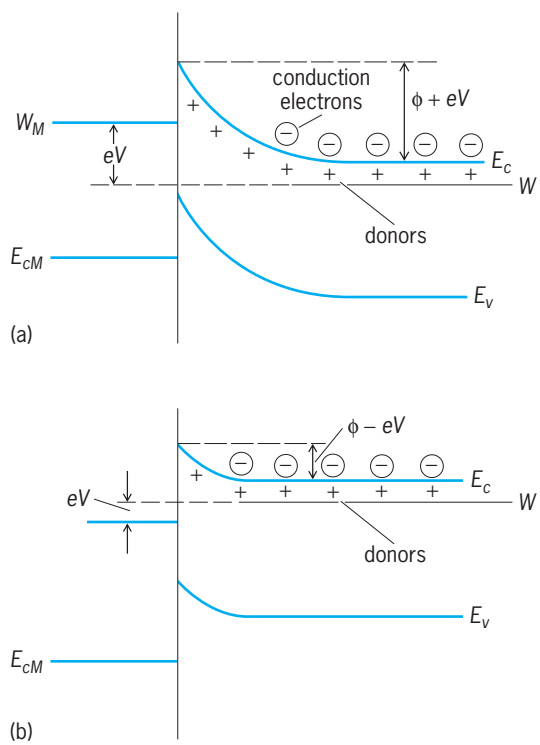

Fig. 4. Energy diagrams of a rectifying contact between a metal and an *n*-type semiconductor under an applied voltage *V*. (*a*) Positive semiconductor. There is a net flow of electrons from metal to semiconductor. (*b*) Negative semiconductor. There is a net flow of electrons from semiconductor to metal.

moving carriers. The average distance traveled by a carrier in its random thermal motion between collisions is called the mean free path. If barrier thickness is large compared to mean free path of carriers, motion of carriers in the barrier can be treated as drift and diffusion. This viewpoint is the basis of the diffusion theory of rectification. According to this theory, the factor C in Eq. (5) depends on the mobility and the electric field in the barrier.

Diode theory. When the barrier thickness is comparable to or smaller than the mean free path of the carriers, then the carriers cross the barrier without being scattered, much as in a vacuum tube diode.

According to this theory, the factor C in the rectifier equation is $v/4$, where v is the average thermal velocity of the carriers.

Tunneling theory. Instead of surmounting a potential barrier, carriers have a probability of penetrating through the barrier. The effect, called tunneling, becomes dominant if the barrier thickness is sufficiently small. This effect is important in many applications. *See* TUNNELING IN SOLIDS.

Two-carrier theory. Often the conduction through a rectifying barrier depends on both electron and hole carriers. An important case is the *pn* junction between *p*- and *n*-sections of a semiconductor material. Also, in metal-semiconductor rectifiers, the barrier presents an obstacle for the flow of majority carriers but not for the flow of minority carriers, and the latter may become equally or more important.

Rectification at pn junctions. A *pn* junction is the boundary between a *p*-type region and an *n*-type region of a semiconductor. When the impurity content varies, there is a variation of electron and hole concentrations. A variation of carrier concentrations is related to a shift of the energy bands relative to the constant Fermi level. This is brought about by a variation of the electrostatic potential which requires the existence of a space charge. If the impurity content changes greatly within a short distance, a large space charge is obtained within a narrow region. Such is the situation existing in a rectifying *pn* junction.

When a voltage is applied to make the *n*-region negative relative to the *p*-region, electrons flow from the *n*-region, where they are abundant, into the *p*-region. At the same time, holes flow from the *p*-region, where holes are abundant, into the *n*-region. The resistance is therefore relatively low. The direction of current in this case is forward. Clearly, the resistance will be high for current in the reverse direction.

With a current in the forward direction, electrons in the *n*-region and holes in the *p*-region flow toward the junction and there must be continuous hole-electron recombination in the neighborhood of the junction. The minority carrier concentration in each region is increased near the junction because of the influx of the carriers from the other region. This phenomenon is known as carrier injection. When there is a current in the reverse direction, there must be a continuous generation of holes and electrons in the neighborhood of the junction, from which electrons flow out into the *n*-region and holes flow out into the *p*-region. Thus current through a *pn* junction is controlled by the hole-electron recombination or generation in the vicinity of the junction.

The transistor consists of two closely spaced *pn* junctions in a semiconductor with an order *pnp* or *npn*.

Contact rectification. If the height of a rectifying contact barrier is high, only a very small fraction of majority carriers can pass over the barrier. The fraction may be so small as to be comparable with the concentration of the minority carriers, provided the energy gap is not too large. The current due to the minority carriers becomes appreciable if the barrier height above the Fermi level approaches the energy difference between the Fermi level and the top of the valence band (Fig. 3).

The concentration of minority carriers is higher at the contact than in the interior of the semiconductor. With a sufficiently high barrier, it is possible to obtain at the contact a minority carrier concentration higher than that of the majority carriers. The small region where this condition occurs is called the inversion layer.

As in the case of a *pn* junction, a forward current produces injection of minority carriers. With the presence of an inversion layer, the injection can be so strong as to increase appreciably the conductivity in the vicinity of the contact. Ordinarily, contact rectifiers consist of a semiconductor in contact with a metal whisker. For large forward currents, the barrier resistance is small, and the resistance of the rectifier is determined by the spreading resistance of the semiconductor for a contact of small area. By increasing the conductivity in the vicinity of the contact where the spreading resistance is concentrated, carrier-injection may reduce considerably the forward resistance of the rectifier.

Surface electronics. The surface of a semiconductor plays an important role technologically, for example, in field-effect transistors and charge-coupled devices. Also, it presents an interesting case of two-dimensional systems where the electric field in the surface layer is strong enough to produce a potential wall which is narrower than the wavelengths of charge carriers. In such a case, the electronic energy levels are grouped into subbands, each of which corresponds to a quantized motion normal to the surface, with a continuum for motion parallel to the surface. Consequently, various properties cannot be trivially deduced from those of the bulk semiconductor. *See* CHARGE-COUPLED DEVICES; SURFACE PHYSICS.

H. Y. Fan

Bibliography. M. L. Cohen and J. R. Chelikowsky, *Electronic Structure and Optical Properties of Semiconductors,* 2d ed., 1988; K. Seeger, *Semiconductor Physics*, 7th ed., 1999; M. Shur, *Physics of Semiconductor Devices*, 1996; J. Singh, *Semiconductor Devices: An Introduction*, 1994; S. Sze, *Physics of Semiconductor Devices*, 2d ed., 1981; S. Wang, *Fundamentals of Semiconductor Theory and Device Physics*, 1989; R. K. Willardson et al. (eds.), *Semiconductors and Semimetals*, vols. 1–42, 1966–1994; M. Zambuto, *Semiconductor Devices*, 1989.

Semiconductor diode

A two-terminal electronic device that utilizes the properties of the semiconductor from which it is constructed. In a semiconductor diode without a *pn* junction, the bulk properties of the semiconductor itself are used to make a device whose characteristics may be sensitive to light, temperature, or electric field. In a diode with a *pn* junction, the properties of the *pn* junction are used. The most important property of a *pn* junction is that, under ordinary conditions, it will allow electric current to flow in only one direction. Under the proper circumstances, however, a *pn* junction may also be used as a voltage-variable capacitance, a switch, a light source, a voltage regulator, or a means to convert light into electrical power. *See* SEMICONDUCTOR.

Silicon and germanium are the semiconductors most often used in diodes. However, other materials may be used for special purposes: cadmium sulfide and cadmium selenide in photoconductors; gallium phosphide, gallium arsenide–phosphide, and silicon carbide in light-emitting diodes; and gallium arsenide in microwave generators. *See* GALLIUM; GERMANIUM; SILICON.

Diodes without pn junctions. The band structure of each semiconductor contains a forbidden energy gap. That is, there is a range of energies, ΔE, in which there are no quantum states that electrons may occupy. When the semiconductor is in thermal equilibrium, its electrons are distributed among the allowed quantum states according to Fermi statistics: The free electrons tend to reside near the minimum in the conduction band, and the unoccupied states in the valence band, called free holes, tend to be concentrated near the maximum in the valence band. **Figure 1** shows a schematic of the band structure in gallium arsenide, GaAs. According to Fermi statistics, the holes are located mainly at $E = 0$, $p = 0$; and the electrons are concentrated near $E = E_1$, $p = 0$.

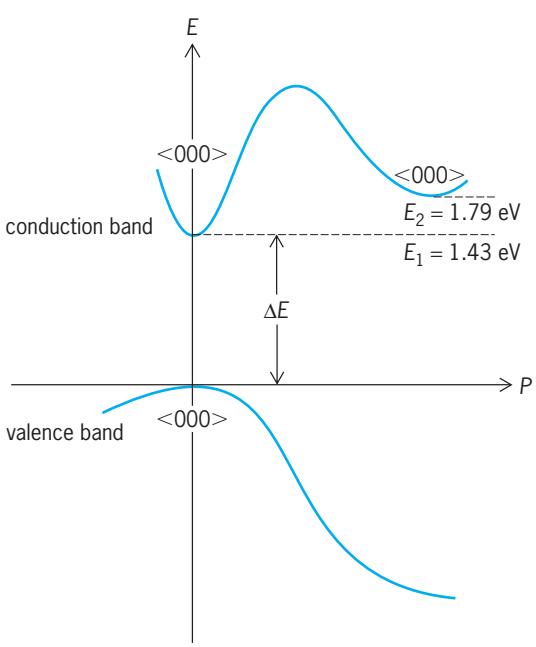

Fig. 1. Band structure of gallium arsenide; E = electron energy, P = momentum.

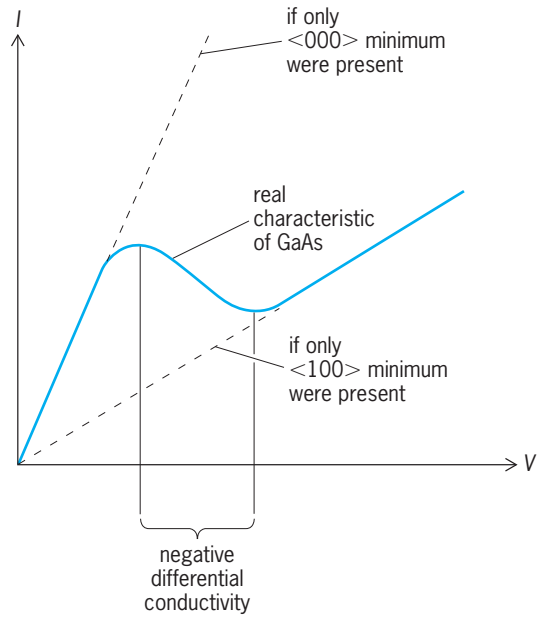

Fig. 2. Current-voltage characteristic bulk of GaAs.

The conductivity of a semiconductor is proportional to the number of electrical carriers (electrons and holes) it contains. In a temperature-compensating diode, or thermistor, the number of carriers changes with temperature. For example, in an intrinsic diode the number of carriers and the conductivity are proportional to $e^{-\Delta E/kT}$. As temperature T increases, the conductivity increases. This effect may be used to cancel out other conductivity changes in an electrical circuit so that the net change is zero. *See* THERMISTOR.

In a photoconductor the semiconductor is packaged so that it may be exposed to light. Light photons whose energies are greater than ΔE can excite electrons from the valence band to the conduction band, increasing the number of electrical carriers in the semiconductor. Thus the conductivity of a semiconductor is a measure of the light intensity striking the semiconductor. *See* PHOTOCONDUCTIVITY.

In some semiconductors the conduction band has more than one minimum. An example is gallium arsenide, as shown in Fig. 1. Here the ⟨000⟩ minimum is located 1.43 eV above the valence band, and the ⟨100⟩ minimum is located 0.36 eV above the ⟨000⟩ minimum. When an electric field is applied to GaAs, the electrons start out in the ⟨000⟩ minimum and the current-voltage relationship initially follows Ohm's law, as shown in **Fig. 2**. In this and other figures, I represents current and V represents voltage. When the field approaches 3200 V/cm, a significant number of electrons are scattered into the ⟨100⟩ minimum. Since carriers in the ⟨100⟩ minimum have lower mobilities than those in ⟨000⟩, the conductivity must decrease. This results in a region of negative differential conductivity, and a device operated in this

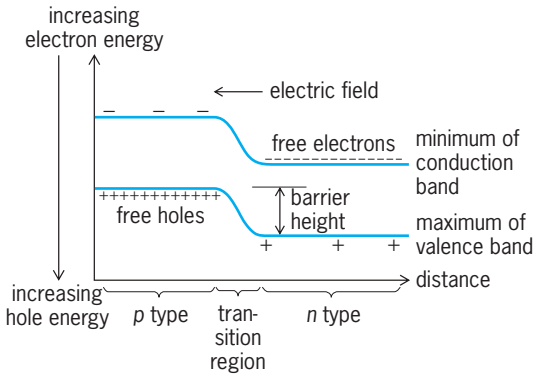

Fig. 3. Energy bands at the potential barrier for *pn* junction of a diode.

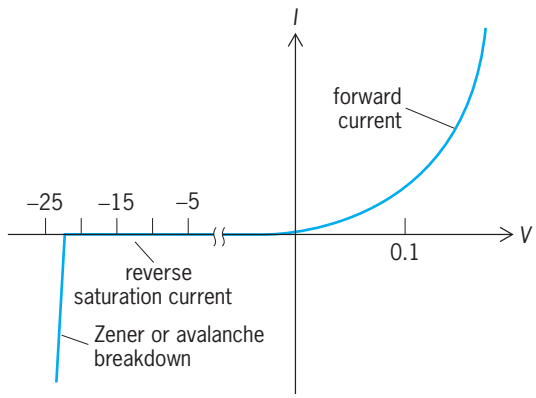

Fig. 4. Current-voltage characteristic of a *pn* junction.

region is unstable. The current pulsates at microwave frequencies, and the device, a Gunn diode, may be used as a microwave power source. *See* MICROWAVE SOLID-STATE DEVICES.

Diodes with pn junctions. A rectifying junction is formed whenever two materials of different conductivity types are brought into contact. Most commonly, the two materials are an *n*-type and a *p*-type semiconductor, and the device is called a junction diode. However, rectifying action also occurs at a boundary between a metal and a semiconductor of either type. If the metal contacts a large area of semiconductor, the device is known as a Schottky barrier diode; if the contact is a metal point, a point-contact diode is formed. *See* SCHOTTKY EFFECT.

The contact potential between the two materials in a diode creates a potential barrier which tends to keep electrons on the *n* side of the junction and holes on the *p* side. This barrier is shown in the energy band diagram of **Fig. 3**. When the *p* side is made positive with respect to the *n* side by an applied field, the barrier height is lowered and the diode is forward-biased. Majority electrons from the *n* side may flow easily to the *p* side, and majority holes from the *p* side may flow easily to the *n* side. When the *p* side is made negative, the barrier height is increased and the diode is reverse-biased. Then, only a small leakage current flows: Minority electrons from the *p* side flow into the *n* side, and minority holes from the *n* side flow into the *p* side. The current-voltage characteristic of a typical diode is shown in **Fig. 4**. Rectifying diodes can be made in a variety of sizes, and much practical use can be made of the fact that such a diode allows current to flow in essentially one direction only. At one extreme, single devices may be used to handle thousands of watts of power in changing alternating current to direct current. At the other extreme, a small diode may be used to detect an amplitude-modulated (AM) radio signal that has only microwatts of power.

The potential barrier at a *pn* junction consists of an electric dipole made up of charged impurity atoms, positively charged on the *n* side and negatively charged on the *p* side. Because the ends of the dipole are separated by a small distance, the junction acts like a capacitor. Moreover, the capacitance is voltage-variable since the barrier height is voltage-dependent. A diode which is specifically designed to utilize this capacitance characteristic is called a varactor.

When light photons with energies greater than ΔE strike a semiconductor near a *pn* junction, the potential barrier will sweep the conduction electrons that were generated by the light into the *n* side of the diode, and the newly created holes into the *p* side. If the diode, now called a solar cell, is connected to an external circuit, the diode will supply electrical power to the circuit as long as light strikes the *pn* junction. If, on the other hand, the diode is reverse-biased, the magnitude of the reverse saturation current can be used to measure light intensity. In this configuration the diode is called a photodiode. *See* PHOTOVOLTAIC EFFECT; SOLAR CELL.

When a junction diode is forward-biased, the number of electrons on the *p* side and the number of holes on the *n* side are increased far above their equilibrium values. This makes the recombination of electrons with holes highly probable, and the recombination energy may be released as photons of light.

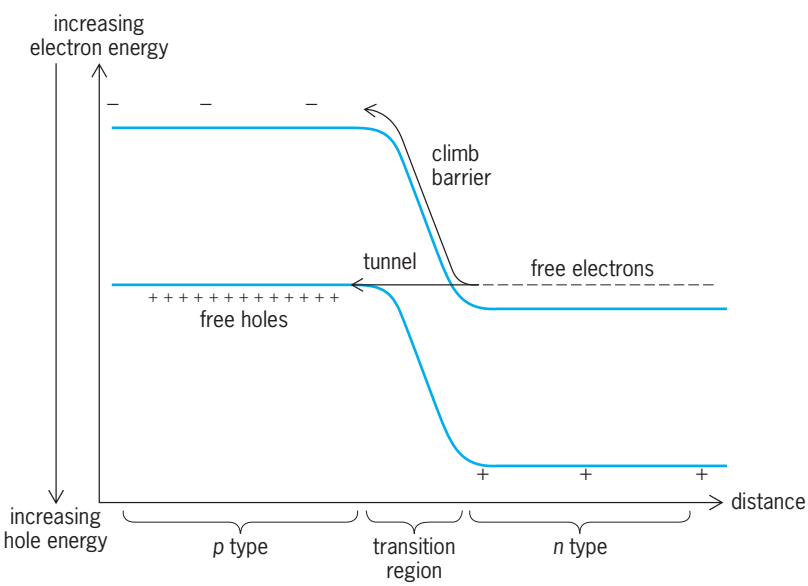

Fig. 5. Energy bands at *pn* junction in a tunnel diode.

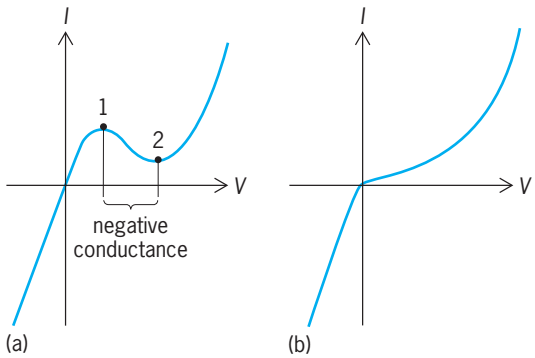

Fig. 6. Current-voltage characteristic of (*a*) a tunnel diode and (*b*) a backward diode.

In some materials, notably GaP for red or green light and GaAs for infrared light, the conversion of electrical energy directly to light can be made quite efficient. And in materials like GaAs, where the valence band maximum and the conduction band minimum are at the same value of crystal momentum, laser action may be achieved.

An Esaki tunnel diode is formed when both the n and p sides of a junction diode are very heavily doped. In this case, as shown in **Fig. 5**, an electron may move from the n side to the p side under forward bias in either of two ways: It may climb over the barrier, as in an ordinary diode; or it may quantum-mechanically tunnel through the barrier, since there are some available states on the p side that are at the same energy as states on the n side. Tunneling is highly probable at low and medium reverse voltages and at low forward voltages. As shown in **Fig. 6***a*, a tunnel diode has an unstable negative-resistance region. The diode may be operated either at point 1 or point 2, but may not be operated in between. This feature may be used to advantage in logic circuits, where the diode is made to switch from one point to the other. *See* TUNNEL DIODE.

A backward diode is similar to a tunnel diode except that it has no forward tunnel current (Fig. 6*b*). It is useful as a low-voltage rectifier.

Although simple diode theory predicts a constant reverse saturation current no matter how high the reverse voltage, all real diodes eventually break down either by tunneling or by avalanche (Fig. 4).

Avalanching occurs when the electrons or the holes gather so much energy that they can create new hole-electron pairs by collision. This process is cumulative and can result in a rapid buildup of current carriers. A diode that has a sharp breakdown provides a virtually constant voltage over a wide range of current. Popularly known as a Zener diode, this device may be used to provide a reference voltage or to regulate the voltage at some point in a circuit.

Stephen Nygren

Bibliography. J. Millman and A. Grabel, *Microelectronics*, 2d ed., 1987; G. W. Neudeck, *The PN Junction Diode*, 2d ed., 1989; J. Singh, *Semiconductor Devices: An Introduction*, 1994; E. S. Yang, *Microelectronic Devices*, 1988.

Semiconductor heterostructures

Structures consisting of two different semiconductor materials in junction contact, with unique electrical or electrooptical characteristics. A heterojunction is a junction in a single crystal between two dissimilar semiconductors. The most important differences between the two semiconductors are generally in the energy gap and the refractive index. In semiconductor heterostructures, differences in energy gap permit spatial confinement of injected electrons and holes, while the differences in refractive index can be used to form optical waveguides. Semiconductor heterostructures have been used for diode lasers, light-emitting diodes, optical detector diodes, and solar cells. In fact, heterostructures must be used to obtain continuous operation of diode lasers at room temperature. Heterostructures also exhibit other interesting properties such as the quantization of confined carrier motion in ultrathin heterostructures and enhanced carrier mobility in modulation-doped heterostructures. Structures of current interest utilize III–V and IV–VI compounds having similar crystal structures and closely matched lattice constants. *See* BAND THEORY OF SOLIDS; LASER; LIGHT-EMITTING DIODE; OPTICAL DETECTORS; REFRACTION OF WAVES; SOLAR CELL.

Carrier and optical field confinement. The most intensively studied and thoroughly documented materials for heterostructures are gallium arsenide (GaAs) and aluminum gallium arsenide ($Al_xGa_{1-x}As$). Several other III–V and IV–VI systems also are used for semiconductor heterostructures. The variation of the energy gap E_g and the refractive index $\overline{n}$ with aluminum arsenide (AlAs) mole fraction x are shown in **Fig. 1**. The lattice constant a_0 is also noted in Fig. 1 to emphasize that a_0 for GaAs and AlAs differs, but by an amount less than 0.14%. A close lattice match is necessary in heterostructures in order to obtain high-quality crystal layers by epitaxial growth and thereby to prevent excessive carrier recombination at the heterojunction interface. *See* CRYSTAL STRUCTURE.

A GaAs-$Al_{0.3}Ga_{0.7}As$ double heterostructure is illustrated in **Fig. 2**. The energy gap E_g versus x given in Fig. 1*a* shows that E_g for $Al_{0.3}Ga_{0.7}As$ is 0.37 eV greater than for GaAs. When a positive potential is connected to the p-type $Al_{0.3}Ga_{0.7}As$, electrons are injected into the GaAs layer from the wider-energy-gap n-type $Al_{0.3}Ga_{0.7}As$, and holes are injected into the GaAs from the wider-energy-gap p-type $Al_{0.3}Ga_{0.7}As$. The wider-energy-gap $Al_{0.3}Ga_{0.7}As$ layers also confine the injected electrons and holes to the GaAs layer, where they can recombine radiatively. This confinement is illustrated in Fig. 2*a*.

Figure 1*b* shows that the refractive index for the $Al_{0.3}Ga_{0.7}As$ layers will be less than for the GaAs layer. This refractive index step is shown in Fig. 2*b*. Solution of the reduced wave equation shows that the optical field will be confined to the larger-refractive-index GaAs layer. The resulting optical field distribution is shown in Fig. 2*c*. When this heterostructure is formed as a rectangular bar with parallel reflecting

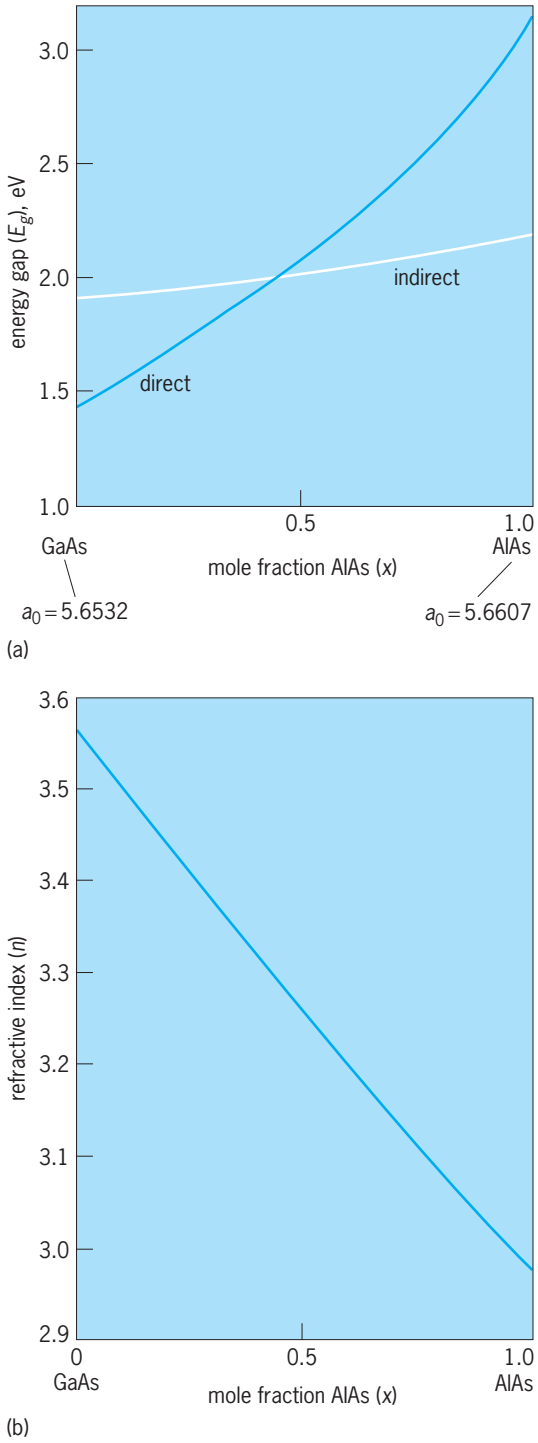

Fig. 1. Compositional dependence in $Al_xGa_{1-x}As$ of (*a*) the energy gap and (*b*) the refractive index at a photon energy of 1.38 eV. In both cases, temperature = 297 K (75°F). (*After S. E. Millet and A. G. Chynoweth, eds., Optical Fiber Telecommunications, Academic Press, 1979*)

surfaces, it can become a diode laser at current densities near 2×10^3 A/cm^2 at a forward voltage of 1.6 V.

Quantum well effects. When the narrow energy gap layer in heterostructures becomes a few tens of nanometers or less in thickness, new effects that are associated with the quantization of confined carriers are observed. These ultrathin heterostructures are referred to as superlattices, artificially layered structures, or quantum well structures, and they consist of alternating layers of GaAs and $Al_xGa_{1-x}As$ (**Fig. 3*a***). These structures are generally prepared by molecular-beam epitaxy. Each layer has a thickness in the range of 5 to 40 nanometers. The energy band diagram for the quantum well structure is shown in Fig. 3*b*. The abrupt steps in the energy gaps form potential wells in the conduction and valence bands. *See* ARTIFICIALLY LAYERED STRUCTURES.

In the GaAs layers, the motion of the carriers is restricted in the direction perpendicular to the heterojunction interfaces, while they are free to move in the other two directions. The carriers can therefore be considered as a two-dimensional gas. The Schrödinger wave equation shows that the carriers moving in the confining direction can have only discrete bound states. As the thickness of the quantum wells gets large, a continuum of states then results. The discrete states in the undoped GaAs quantum wells are illustrated in Fig. 3*b* and are accurately predicted by the well-known quantum-mechanical problem of the particle-in-a-box. *See* NONRELATIVISTIC QUANTUM THEORY.

The presence of the discrete quantum states may readily be observed by low-temperature optical absorption measurements. The observation

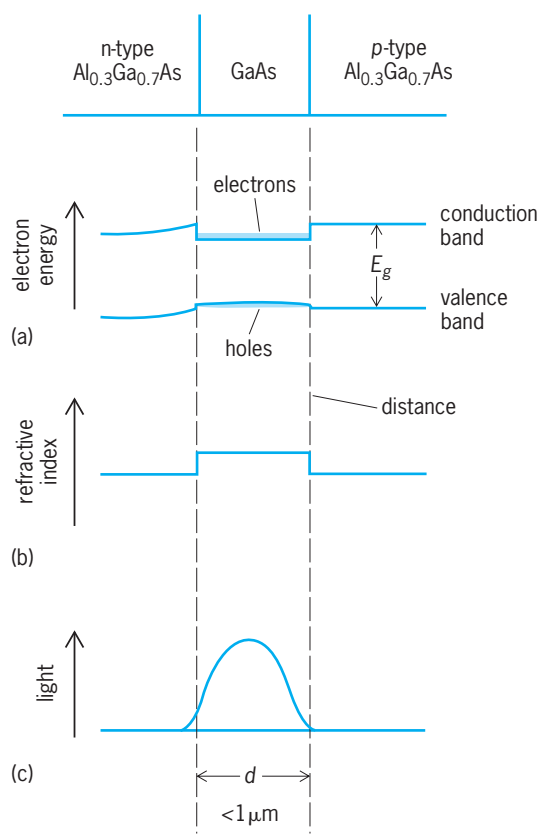

Fig. 2. Schematic representation of an $Al_{0.3}Ga_{0.7}As$-GaAs-$Al_{0.3}Ga_{0.7}As$ double heterostructure laser showing (*a*) energy band diagram at high forward bias, (*b*) refractive index profile, and (*c*) optical field distribution. (*After H. C. Casey, Jr., and M. B. Panish, Heterostructure Lasers, part A, Fundamental Principles, Academic Press, 1978*)

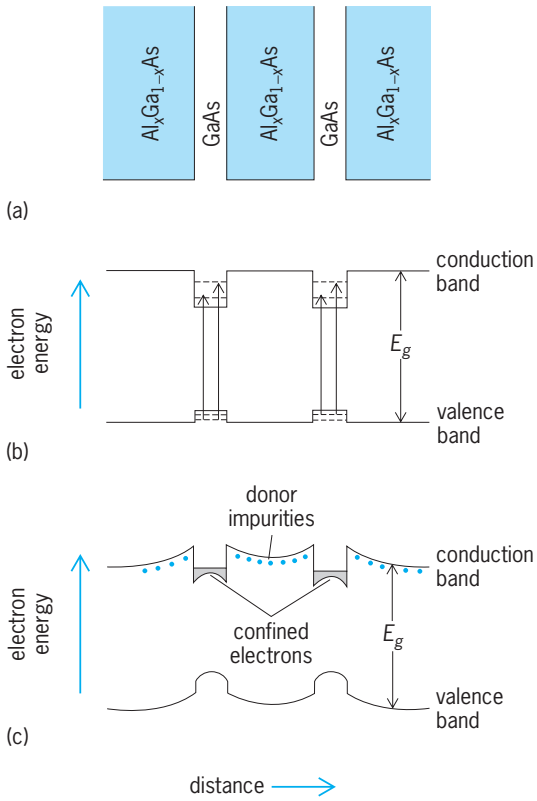

Fig. 3. GaAs-$Al_{0.3}Ga_{0.7}As$ superlattice. (*a*) Schematic representation. (*b*) Quantum wells, with the discrete conduction and valence band states for undoped GaAs. (*c*) Quantum wells with modulation doping.

wavelength is from 0.6 to 0.9 micrometer. These measurements are generally made at liquid helium temperatures with 20 or more GaAs quantum wells separated by $Al_xGa_{1-x}As$ layers. The absorption of photons takes place in the GaAs quantum wells by the transition of electrons in quantized states of the valence band to the quantized states of the conduction band. Sharp peaks in the absorption spectrum occur at the energies that separate the quantized states in the valence and conduction bands, and the photon energies of these absorption peaks vary with the thickness of the GaAs layer as predicted.

Modulation doping. Another property of semiconductor heterostructures is illustrated by a modulation doping technique that spatially separates conduction electrons and their parent donor impurity atoms. As illustrated in Fig. 3*c*, the donor impurities are incorporated in the wider-energy-gap $Al_xGa_{1-x}As$ layer but not in the GaAs layer. The GaAs conduction band edge is lower in energy than the donor states in the $Al_xGa_{1-x}As$, and therefore electrons from the donor impurities will move into the GaAs layers.

The useful feature of modulation doping is that the mobile carriers are the electrons in the GaAs layers and that these mobile carriers are spatially separated from their parent donor impurities in the $Al_xGa_{1-x}As$ layers. Since the carrier mobility in semiconductors is decreased by the presence of ionized and neutral impurities, the carrier mobility in the modulation-doped GaAs is larger than for a GaAs layer doped with impurities to give the same free electron concentration. Higher carrier mobilities should permit preparation of devices that operate at higher frequencies than are possible with doped layers.

Heteroepitaxy. In order to prepare optical sources and detectors that operate at a desired wavelength or to achieve high-mobility semiconductors, often it is necessary to use quaternary solid solutions. To obtain high-quality layers, the epitaxial layers must be grown on binary substrates that have the same lattice constant as the solid solutions. An example of considerable interest is the growth of $Ga_xIn_{1-x}P_yAs_{1-y}$ layers lattice-matched to InP. In this case, the heterojunction is used to obtain a similar lattice constant for dissimilar semiconductors at a desired bandgap. The heteroepitaxial growth techniques are chemical-vapor deposition, liquid-phase epitaxy, and molecular-beam epitaxy.

H. C. Casey, Jr.

Chemistry. The III–V and II–VI compounds and alloys commonly used in heterostructures are usually formed either by direct synthesis from the elements or by vapor-phase decomposition reactions using compounds containing hydrogen, halogen, or organic radicals. Substrates used for the growth of heterostructures are usually gotten from single-crystal boules synthesized directly from the elements. Most heterostructure materials can be doped either *n* or *p* type. However, in some cases one conductivity type can be difficult to obtain due to a "self-compensation" mechanism in which the formation of one type of vacancy is energetically favorable and compensates the opposite doping type. For example, the *p*-type beryllium (Be) doping of aluminum nitride (AlN) is compensated by the formation of *n*-type nitrogen vacancies. In general, variations in stoichiometry will greatly affect electrical and optical properties. Also, the properties of heterostructures greatly depend on the lattice constant match and chemical abruptness of the interface. A lattice mismatch of greater than 0.5% usually results in an interface with a significant density of both electron and hole traps. The chemical rule that determines the allowable alloy compositions is that the sum of the atom fraction of the group III (or group II) elements must equal the sum of the atom fractions of group V (or group VI) elements in the crystal, for example, $Ga_{1-x}Al_xAs$, $0 \leq x \leq 1$.

Fabrication. Semiconductor heterostructures are usually fabricated as single-crystal structures using thin-film epitaxial crystal growth techniques such as liquid-phase epitaxy (LPE), chemical vapor deposition (CVD), and molecular-beam epitaxy (MBE).

Liquid-phase epitaxy. The liquid-phase epitaxy method has been the most widely used technique for fabricating heterostructures. The principal advantage is the ease of fabrication of high-purity heterostructures with good electrical, optical, and interface properties as a result of impurity segregation from the growing layer into the melt during growth. The main disadvantage is the poor control of layer thickness and interface morphology and abruptness. Epitaxial growth by the liquid-phase

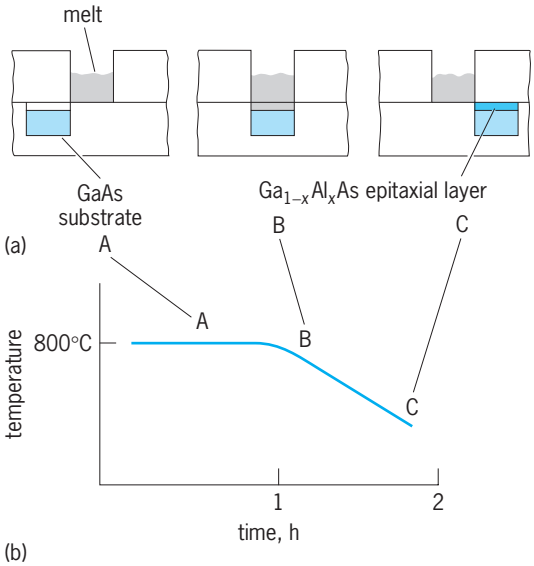

Fig. 4. Liquid-phase epitaxy method for GaAs-$Ga_{1-x}Al_xAs$ heterostructures. (*a*) Simplified diagrams of successive stages of the process. (*b*) Variation of temperature during the process. 800°C = 1500°F. (*After J. M. Woodall, III-V compounds and alloys: An update, Science, 208:908–915, 1980*)

epitaxy method occurs when a melt or solvent becomes supersaturated, usually by supercooling, with respect to a solid phase in presence of a single-crystal substrate. For the growth of III–V materials, the melt usually has a large concentration of one of the group III components in epitaxial layer. For example, epitaxial growth of $Ga_{1-x}Al_xAs$ for $0 \leq x \leq 1$ can be obtained from melts containing about 90% Ga and 10% Al plus As. A "typical" schedule for the liquid-phase epitaxy growth of a $Ga_{1-x}Al_xAs$ layer on GaAs is shown in **Fig. 4**. The important features are a "bottomless" bin to contain the melt and a fixture with a slightly recessed substrate which can be moved with respect to the melt bin. The first step is to thermally equilibrate the melt prior to contact with the substrate (A in Fig. 4). This means that the melt is either at solid-liquid equilibrium or slightly undersaturated. Next, the melt is cooled, usually at constant rate, while the substrate is in contact with it (B of Fig. 4). The desired layer thickness is obtained by programming either a change in temperature or elapsed time and then removing the substrate from the melt (C of Fig. 4). The growth of heterostructures can also be accomplished by using multimelt fixtures, moving the substrate between the various melts during cooling. The term liquid-phase epitaxy originally referred to growth from supercooled melts. The technique now includes growth from melts supersaturated by other techniques such as electroepitaxy, in which the solid–liquid interface region becomes supersaturated due to the flow of electric current across the interface and isothermal melt mixing techniques in which solid–liquid equilibrium melts of different compositions are mixed together and become supersaturated.

Chemical vapor deposition. The chemical vapor deposition method refers to the formation of thin solid films as the result of thermochemical vapor-phase reactions. When the films are epitaxial, the method is sometimes called vapor phase epitaxy (VPE). For III–V materials two different chemistries have been widely studied: the group III and group V halogen compounds and group V hydrogen compounds; and the group III metal-organic compounds and V—hydrogen compounds such as $Ga(CH_3)_3$ and AsH_3. The halogen transport reactions are of the "hot" to "cold" type in which the III-halogen is produced in a high-temperature zone by the reaction of the III element with HCl. The III-halogen then diffuses to the low-temperature zone, where it combines with the V species to form an epitaxial layer of III–V material. Vapor-phase epitaxy using metal-organic compounds (metal-organic chemical vapor deposition; MOCVD) occurs when the organic radical R of the III-R compound is "cracked" or pyrolyzed away at a "hot wall" or hot substrate in the presence of the V-H_3 compound. A schematic illustration of the metal-organic chemical vapor deposition technique applied to GaAs is shown in **Fig. 5**. Both the halogen and metal-organic chemistries have been successfully used to fabricate heterostructures. It is thought that the metal-organic chemical vapor deposition method may have long-term advantages such as better purity, composition, and thickness control and the ability to produce a wide range of III–V compounds and alloys.

Molecular-beam epitaxy. The molecular-beam epitaxy technique is relatively new compared with liquid-phase epitaxy and chemical vapor deposition. For heterostructure fabrication, it is capable of a layer thickness and interface abruptness control on a monoatomic scale. This represents almost two orders of magnitude improvement in structural resolution in the growth direction compared to liquid-phase epitaxy and chemical vapor deposition. The molecular-beam epitaxy technique is similar to

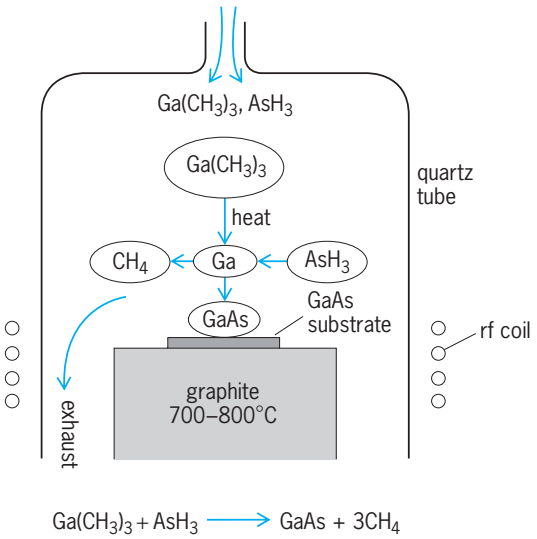

Fig. 5. Simplified diagram of the metal-organic method for chemical vapor deposition. 700–800°C = 1300–1500°F.

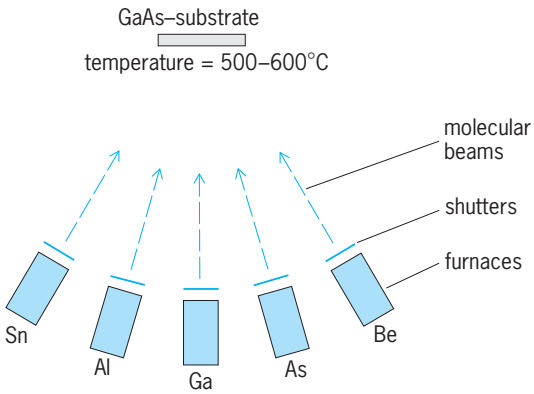

Fig. 6. Schematic illustration of a system configuration for growth of doped $Ga_{1-x}Al_xAs$ by molecular-beam epitaxy. 500–600°C = 900–1100°F. (*After P. E. Luscher, Crystal growth by molecular beam epitaxy, Solid State Technol., 20(12):43–52, 1977*)

evaporation techniques used to deposit thin metal films. The key features of the molecular-beam epitaxy method are the use of an ultrahigh-vacuum environment and epitaxial growth by the reaction of multiple molecular beams of differing flux and chemistry with a heated single-crystal substrate. A schematic illustration of molecular-beam epitaxy applied to doped $Ga_{1-x}Al_xAs$ is shown in **Fig. 6**. Each furnace heats a crucible which is charged with one of the constituent elements or compounds of the desired film. Tin (Sn) functions as an *n*-type dopant, and beryllium (Be) functions as a *p*-type dopant. The furnace temperature is chosen so that the vapor pressure is sufficient to produce the desired beam flux at the substrate surface. The furnaces are arranged so that the flux from each is maximum at the substrate position. The quality and composition of the epitaxial film are determined by the quality of the ultrahigh-vacuum system, the substrate temperature, and the furnace temperature. The shutters interposed between each furnace and the substrate allow the beams to be modulated. This feature, coupled with typical growth rates of a few tenths of a nanometer per second, facilitates the fabrication of very thin and very abrupt heterostructures. *See* CRYSTAL GROWTH; MOLECULAR BEAMS; SEMICONDUCTOR.

Jerry M. Woodall

Bibliography. M. G. Astles, *Liquid-Phase Epitaxial Growth of III–V Compound Semiconductor Materials and Their Device Applications*, 1990; N. G. Einspruch and W. R. Frensley (eds.), *Microstructure Science: Heterostructures and Quantum Devices*, 1994; M. A. Herman and H. Sitter, *Molecular Beam Epitaxy*, 2d ed., 1997; E. V. Ivchenko and P. E. Pikus, *Superlattices and Other Heterostructures: Symmetry and Optical Phenomena*, 2d ed., 1997; R. A. Kiehl and G. Sollner (eds.), *Semiconductors and Semimetals*, vol. 41: *High Speed Heterostructure Devices*, 1994; M. B. Panish and H. Temkin, *Gas Source Molecular Beam Epitaxy: Growth and Properties of Phosphorus-Containing III–V Heterostructures*, 1993; J. Singh, *Physics of Semicondutors and Their Heterostructures*, 1993; S. Sivaram, *Chemical Vapor Deposition*, 1995.

Semiconductor memories

Devices for storing digital information that are fabricated by using integrated circuit technology. Semiconductor memories are widely used to store programs and data in almost every digital system, and have replaced core memory as the main active computer memory.

Organization. Many different types of semiconductor memories are used in digital systems to perform various functions—bulk data storage, program storage, temporary storage, and cache (or intermediate) storage. Almost all of the memories are a form of random-access memory (RAM), in which any storage location can be accessed in the same amount of time.

The storage locations are set up as an array of rows and columns (**Fig. 1**). To the user, the storage array is organized as *M* words deep by *N* bits wide, even though the actual physical arrangement may be different. To access the data, the system supplies addresses—multiple-bit parallel words—that specify the locations to be accessed. Each address pinpoints the storage cell or group of cells (the memory bit or word, respectively), very much in the same way as a pair of grid coordinates specifies a location on a two-dimensional map. Various control lines then permit the system to either read data out from or store new data into the memory, or both, depending on the memory type.

There are also some specialized memory circuits that can be accessed only sequentially. These devices either are based on long strings of shift registers or use an internal RAM array but have a built-in address sequencer that automatically increments or decrements through all the memory locations when the circuit receives timing and control signals (**Fig. 2**). There is another type of serial memory that does permit random access. In this unique memory, the multibit address is loaded one bit at a time rather than in parallel, as with most other memories. After the address is loaded, a control signal tells the memory to access the data, and then the desired data are serially shifted in or out. Such serial memories are typically much slower to access since it is necessary to load in each address with many clock cycles, rather than a single cycle, as with the standard memory chips.

Read/write RAMs. Even though most semiconductor memories can be randomly accessed, they are not all referred to as RAMs. RAMs are memory chips that cannot retain data without power but permit data to be both read from or written into the memory chip's storage locations.

Dynamic and static types. Within the category of read/write RAMs, many subdivisions have been created to satisfy the performance and system architecture requirements of the various applications. Basically there are two types of read/write RAMs—dynamic and static (DRAMs and SRAMs). The terms "dynamic" or "static" refer to the structure of the actual storage circuit (the cell structure) used to hold each data bit within the memory chip. A dynamic memory uses a storage cell based on a transistor

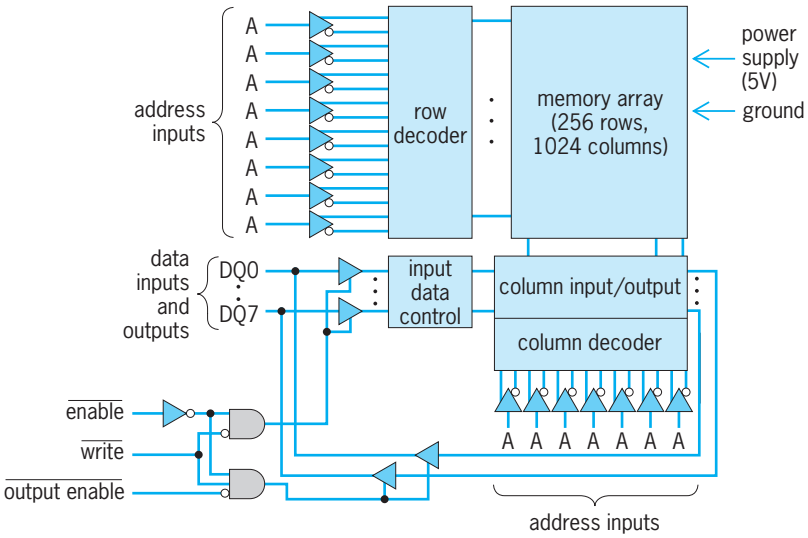

Fig. 1. Basic memory chip with array of memory cells. In this case, a 256-kbit static random-access memory is set up as 256 rows by 1024 columns, which are accessed when addresses are presented at the address inputs and are decoded to select the proper row and column. The 8-bit-wide data path, DQ0 to DQ7, is bidirectional, and just three active-low control lines—enable, write, and output enable—coordinate chip operation.

and capacitor combination, in which the digital information is represented by a charge stored on each of the capacitors in the memory array (**Fig. 3**). The memory gets the name "dynamic" from the fact that the capacitors are imperfect and will lose their charge unless the charge is repeatedly replenished (refreshed) on a regular basis (every few milliseconds) by externally supplied signals. Static memories, in contrast, do not use a charge-storage technique; instead, they use either four transistors and two resistors to form a passive-load flip-flop (**Fig. 4***a*), or six transistors to form a flip-flop with dynamic loads (Fig. 4*b*), for each cell in the array. Once data are loaded into the flip-flop storage elements, the flip-flops will indefinitely remain in that state until the information is intentionally changed or the power to the memory circuit is shut off.

The simple cell structure of dynamic RAMs (two elements—a transistor and a capacitor—that can be fabricated either side-by-side or one on top of the other) makes possible exceptionally small storage-cell areas. (As of 2001, the smallest memory cell area reported for a DRAM was just over 0.1 square micrometer. The small cell was fabricated using minimum feature dimensions of 0.1 μm.) Static RAMs, which require at least four or six transistors for each cell, consume almost four times the area to achieve the same storage capacity as dynamic RAMs. (In 2001, however, researchers crafted a loadless static memory cell that requires an area of just 0.99 μm^2 when fabricated using 0.13 μm design rules.) However, the continual refresh required by the dynamic RAMs demands that memory arrays be surrounded by support circuits (both on- and off-chip to provide the refresh timing signals), and that causes both the DRAM cells and refresh circuits to consume a few milliwatts of power even when the memories are not being accessed. That power drain must be taken into account when dynamic RAMs are used in battery-powered systems.

The interest in low-power system design that started in the early 1990s, coupled with the move downward in power supply voltage, has created the need for dynamic RAMs that operate with 3.3-V and lower power supplies. In the late 1990s, designers were working with even lower supply voltages. Many of the memories designed after 1998 use a 2.5-V power supply, further lowering the power consumption of the memory chips as densities go beyond 256 Mbits on a chip. Still lower power supply voltages are in demand—many designs are now targeted to operate from supply voltages as low as 1 V for cellular phones, digital personal assistants, and other portable systems.

Furthermore, the low-power dynamic RAMs must have special capabilities to retain the data during long standby periods with a minimal power drain. To accomplish this, memory designers have created special slow-refresh operating modes on low-power dynamic RAMs so that fewer refresh cycles are needed to retain the data in the dynamic memory cells, and thus operating power demands are reduced.

To keep dynamic RAM package pin counts as low as possible, the address information is divided into both a row address and a column address. The row and column addresses are time multiplexed with each other and sent to the memory over the same set of pins, thus doubling the amount of time required to supply the complete memory location address. As a result, the best access time for a traditional dynamic RAM can be five times slower than that of a

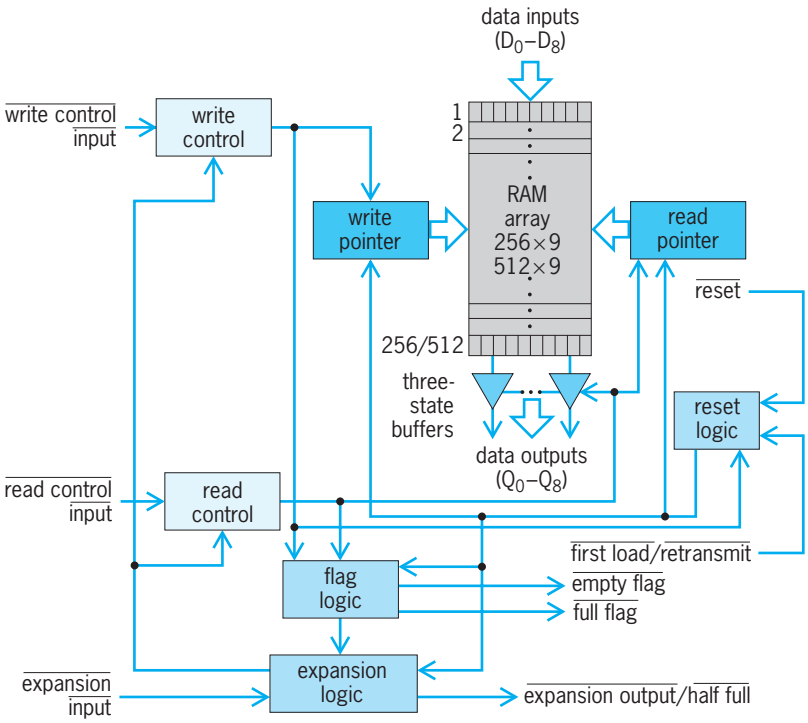

Fig. 2. A 9-bit-wide first-in/first-out (FIFO) memory chip. Inside the memory, read and write pointers provide sequential addresses that determine which memory location is being accessed. Separate 9-bit data paths allow data writes and data reads to take place simultaneously. Additional circuits on the chip provide information as to whether the memory array is empty, full, or half full.

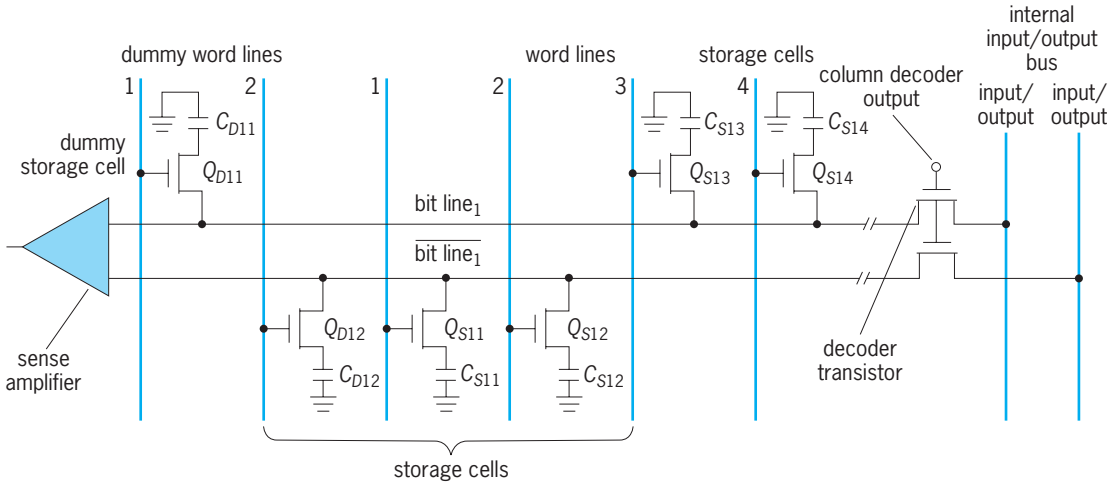

Fig. 3. Internal structure of a dynamic RAM. A typical bit line has many storage cells attached to it, with each cell consisting of a gating transistor Q_S controlled by the word line and a storage capacitor C_S connected to ground. When a particular cell is accessed on the bit line, a matching dummy cell (with gating transistor Q_D and capacitor C_D) is activated so that the sense amplifier receives a clean differential signal that represents the stored value.

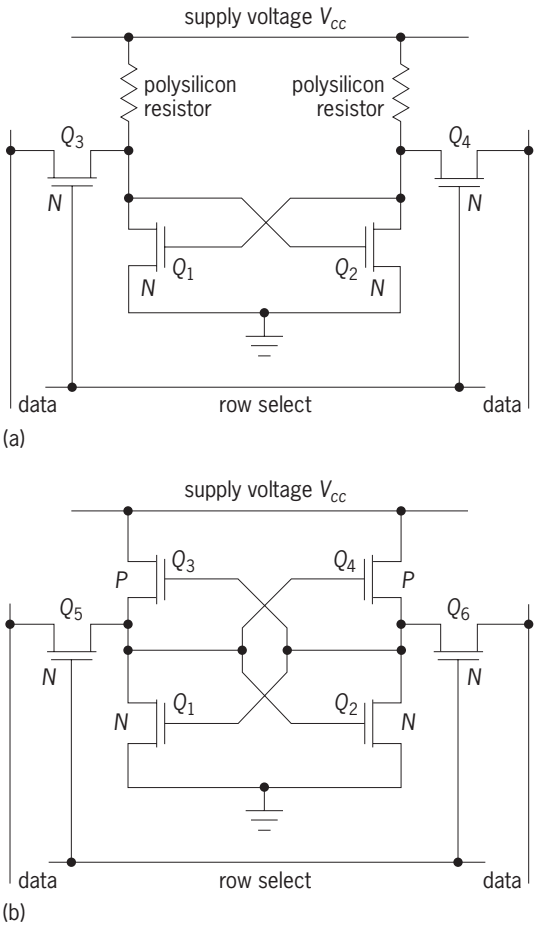

Fig. 4. Static RAM storage cells. (*a*) Cell with four NMOS transistors ($N = n$-channel), which form a simple cross-coupled flip-flop, and two large-value resistors, formed in an overlying polysilicon layer, that serve as the passive load elements. (*b*) Cell with PMOS transistors ($P = p$-channel) instead of resistors as the load elements. Leakage current in the cell is reduced since the off-resistance of the transistors is higher than that of polysilicon resistors.

high-performance static memory fabricated with the same process and design rules.

Alternative types. In addition to static and dynamic RAMs, there is an attempt to combine both technologies, thus merging the high storage density of dynamic memory cells with the simplicity of use of static RAMs. Referred to as pseudostatic or pseudodynamic RAMs, these memories include circuits on the chip to automatically provide the refresh signals needed by the dynamic cells in the memory array. Since the signals do not have to be supplied by the external system, the memory appears to function like a static RAM.

Dynamic RAMs and static RAMs continue to evolve, driven by the need for extremely fast access times and high memory data bandwidths to feed high-performance microprocessors. The standard DRAM has given way to devices with interfaces that incorporate fast-page, nibble, or extended-data-out (EDO) access modes, synchronous bus interfaces, and other more-dedicated interfaces based on standards such as Rambus (developed by Rambus Inc.) or Sync-Link (developed by the SLDRAM consortium), or novel devices such as the multibank DRAM architecture developed by Mosys Inc.

Most personal computers and workstations incorporate synchronous DRAMs (SDRAMs) for their main memory since the chips offer higher data transfer speeds (10 ns or less per transfer after the initial access) than previous memory types and thus allow higher-performance computer systems to be implemented. Fast-page, nibble mode, and EDO DRAMs are all fading in popularity since their performance cannot keep pace with the faster systems.

The synchronous dynamic RAM employs a standard dynamic RAM core, but the system interface timing is lock-stepped to the rest of the system through the use of a master clock input signal. Initial 16-Mbit and 64-Mbit synchronous memories were able to operate at system clock rates of 100 MHz, while the latest 256-Mbit SDRAMs can operate at system bus speeds of 133 MHz or faster. (1 kbit = 1024 bits;

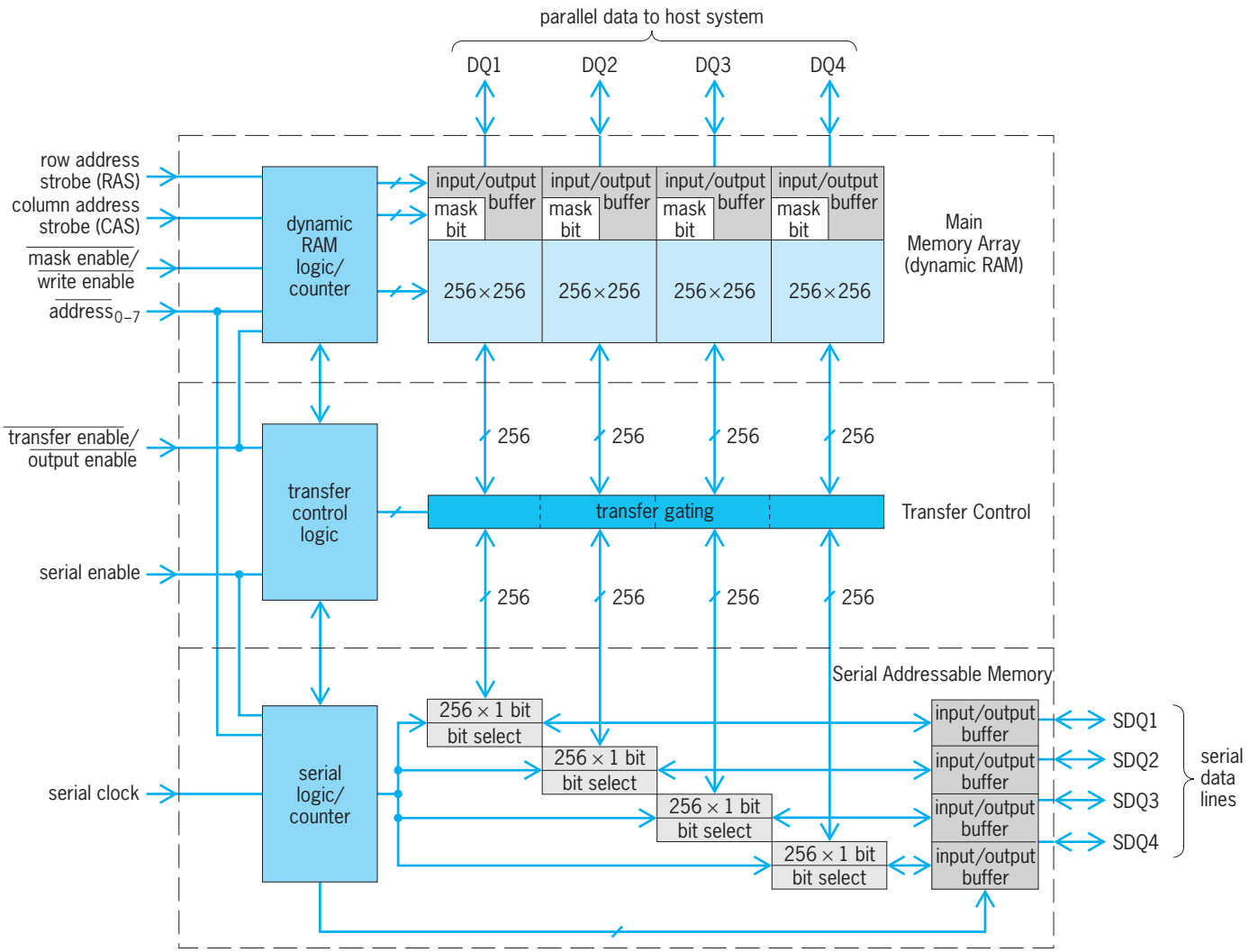

Fig. 5. Video memory organized as 64 kwords by 4 bits, with two data paths. On the top, a 4-bit-wide parallel data path to the 64-kword storage array is divided into four banks of 64 kbits each. The RAS and CAS signals, along with the eight address lines, permit any of the 64-k locations to be randomly accessed. From each of the four banks, a 256-bit string of bits can be transferred to separate bidirectional shift registers and serially shifted out at high speed over the four serial data lines. Conversely, data can be serially shifted into the four serial data lines and transferred into the chip's main memory array.

1 Mbit = 1024 × 1024 bits = 1,048,576 bits.)

New clocking techniques promise to rapidly reduce the transfer speeds of synchronous memories still further by using both the leading and trailing edges of the clock signal to transfer data. (Most synchronous memories prior to 1998 used only one edge of the clock signal to do the transfer.) This effective doubling of the transfer rate is called double-data-rate (DDR), and as of 2001 such DDR interfaces were available on SDRAM chips from just about every DRAM supplier. Although initially developed for use in SDRAMs, the technique can also be applied to synchronous static memories and can effectively double the data transfer speed of those memories as well.

Additionally, design automation tools, which make it much easier to design or modify a circuit, have opened up many new hardware options, especially for video and graphics. Specialized DRAMs that incorporate second and even third data ports provide the high bandwidths needed by high-performance graphics controllers and multimedia subsystems. The first of these variants was the video RAM which, in addition to the standard DRAM memory array, incorporated a high-speed 4- or 8-bit wide buffer memory that transfers a sequential string of data to the graphics controller (**Fig. 5**).

Several other memory types were developed for graphics applications during the 1990s, including devices such as the Rambus interface, Window RAMs, triple-port RAMs, and synchronous graphics RAMs (SGRAMs). For mass-market applications, such as the graphic-adapter cards used in most personal computers, the SGRAM became the most popular option. (The SGRAM is similar to the SDRAM, but has a few additional functions that better suit it for graphics applications.)

Computer systems in 2001 started to incorporate a variation of the original Rambus interface into main memory subsystems. The Direct Rambus interface, an enhanced memory-to-CPU (central processing unit) bus defined by Intel Corp. and Rambus Inc., allows data transfers of 800–1000 Mbytes/s and faster (**Fig. 6**). The Sync-Link DRAM (SLDRAM) interface, developed jointly by a consortium of companies to

serve as a standard, high-performance memory interconnect option, was dropped as an alternative to the SDRAM or Rambus DRAM (**Fig. 7**); however, parts of the SLDRAM definition will persist as designers finalize the definition of a high-performance second-generation DDR interface that incorporates some aspects of the SLDRAM interface.

When Direct Rambus DRAMs (RDRAMs) perform a read operation, a bank of memory in the DRAM core is activated (the core is actually split into multiple banks—16 in the case of the first Direct RDRAMs) and, based on the input address, transfers the requested data from the bank into the internal input/output buffer. That buffer, though, is 128 bits wide and actually receives data from all 16 banks so that after the first access, sequential bytes in the buffer can be accessed at the clock speed. Thus, when clocked at 333 MHz and both edges of the clock are used, the effective data transfer rate is 666 Mbytes/s for a Direct RDRAM device with an 8-bit interface, and over 1300 Mbytes/s for a device with a 16-bit data interface. If the desired data are not in the buffer, the latency encountered to reach the first piece of the data runs about 20 ns; however, subsequent accesses to other data in the buffer take place at the approximate 1.6 ns/clock transfer rate. Even faster versions that could clock at 400 MHz (800 Mtransfers/s) were available in 2000, and 500- and 600-MHz versions were sampled by suppliers in 2001.

Specialty dynamic RAMs with small blocks of static RAM, forming a unique type of cached dynamic RAM, were introduced in 1993. These unusual memories allow system designers to implement memory subsystems with the cache subsystem distributed across the dynamic RAM array. The advantage of the cache-dynamic RAM (CDRAM) integration is that in many smaller systems it can eliminate the need for a separate secondary cache and thus reduce the system cost or size.

In addition to the various DRAM options available for main memory, there are many specialized memory components that provide data buffering in video systems and data communications applications, or speed adjustments between two systems that operate at different bus speeds. First-in/first-out (FIFO) memories provide buffering for burst data transfers, while dual-port or multiport memories allow multiple subsystems, each operating at different clock rates, to read or write data to a central memory and thus provide rate buffering. Although most FIFO memories and multiport memories are fabricated with static RAM cells, a few of these specialty memories are based on arrays of DRAM cells surrounded by logic that turns the memory array into the desired memory type.

Static memories. Just as diversity has become the norm for DRAMs, the same is true for static memories. In addition to the basic asynchronous SRAM, the industry has crafted a variety of higher-performance SRAM approaches to meet the access time requirements of high-speed computers, telecommunications systems, and many other applications. Standard asynchronous SRAMs, which operate with the

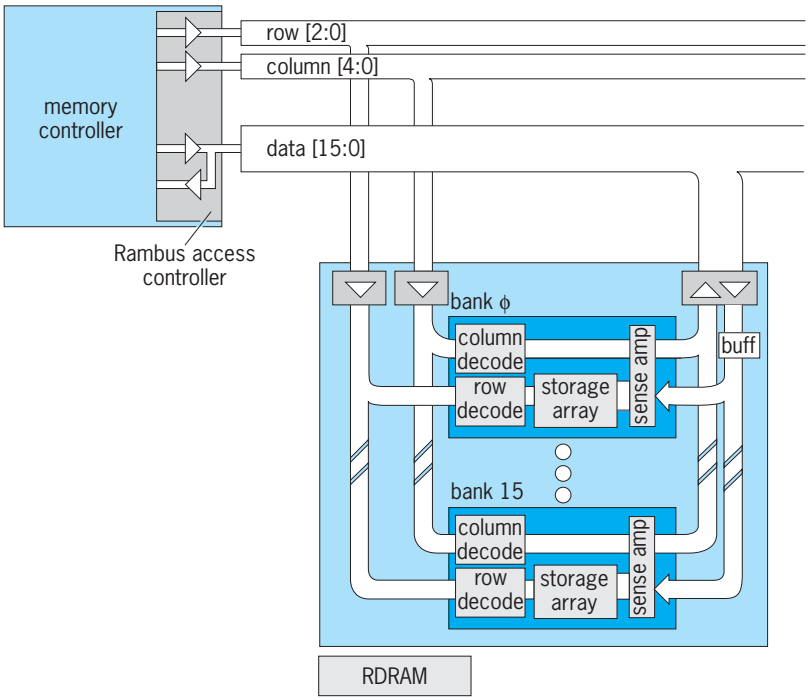

Fig. 6. Multiple memory banks are used in the Direct Rambus DRAM (RDRAM) to provide high-speed access to data. The 16-bit-wide interface and fast bank access allows the Direct RDRAMs to deliver a bandwidth of more than 4 Gbytes/s. (*Rambus Inc.*)

straightforward application of address inputs and a few simple control signals and data, are the most commonly used memory type and are found in many moderate-performance and low-power applications. Access times for asynchronous devices tend to be in the 60–200 ns range.

The huge growth in portable equipment, from cell phones to battery-powered medical instruments, has created a demand for low-power asynchronous memories. These devices consume less than a few hundred milliwatts of power when active, and microwatts of power when idling. To achieve that, especially for the higher-density devices (512 kbits and larger), memory designers are moving the memory cell structures from the commonly used four-transistor cell to the six-transistor cell to reduce leakage currents, simplify manufacturing, and improve the cell tolerance to a phenomenon called soft errors (random changes in bit values due to unpredictable impacts of alpha-particle radiation). *See* RADIATION HARDENING.

High-speed personal computer and workstation cache-memory subsystems typically employ synchronous devices. Some derivative versions of the synchronous memories include small counters on the chip that will let the memories deliver four words in a burst, such that the computer supplies the first address, and then upon each of three subsequent clock cycles, the counter increments and the next data word is sent to the computer, thus eliminating the need for the computer to send a new address for each word. These devices are called pipelined synchronous-burst SRAMs and are widely used with microprocessors such as the Intel Pentium and the IBM/Motorola PowerPC to implement level-2

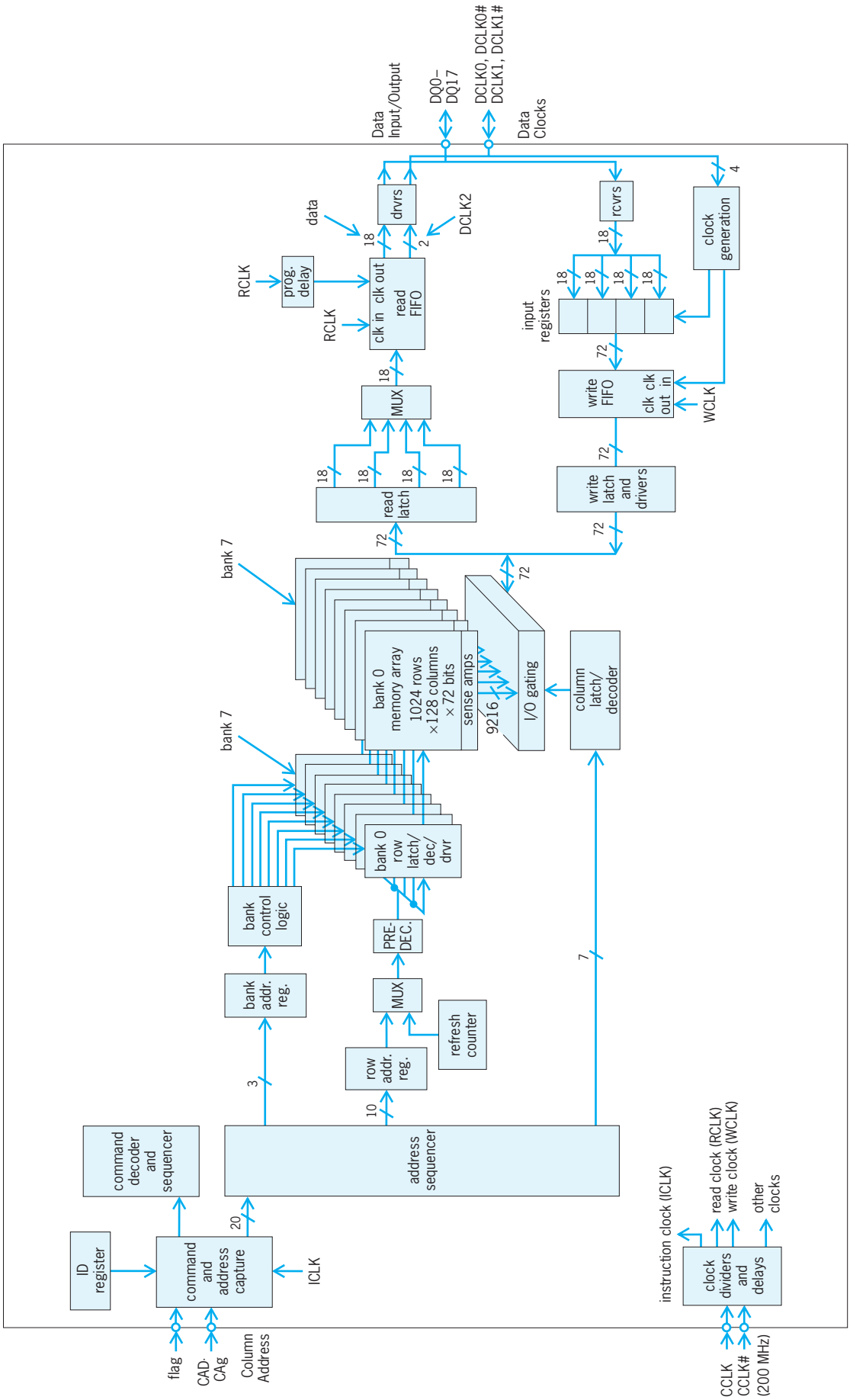

Fig. 7. Synchronous-link DRAM uses multiple memory banks and a packet-based transfer protocol. It decouples the address, control, and data paths so that multiple internal operations can take place concurrently, which speeds memory operation and allows data transfers of 800 Mbytes/s and faster. (*SLDRAM consortium*)

cache subsystems that supplement the on-chip level-1 caches of the processors. The on-chip counters can either be set to count in the special order required by x86 family processors or in standard linear sequence used by other processors.

Yet another variation to the synchronous interface, called zero-bus turnaround or no-bus latency, eliminates the short delay time required by the burst-type parts when the direction of data on the bus must be reversed. For example, when a read operation follows a write operation, typical burst memories must reset some internal circuits, and the adds a delay of a full memory cycle (several nanoseconds) to the response time.

With the no-latency devices, a read can follow a write or vice versa, with no performance penalty. Such memories provide a higher-performance alternative to equipment designers who must deal with reads and writes that occur in a totally unpredictable sequence rather than the typical sequential reads and writes in a desktop computer or workstation cache. Personal computer and workstation cache subsystems typically transfer four or more sequential words into internal processor caches. Those transfers are most often reads that are done by the central processing unit rather than writes to the memory, and thus the time it takes for the bus to reverse direction is of minimal consequence.

In 2001 a new high-performance SRAM interface was developed. Known as a quad-data-rate (QDR) interface, this scheme uses the standard DDR signals to transfer data. However, it also employs separate read and write memory ports. The read and write ports can be simultaneously active, thus doubling the maximum data transfer rate possible with standard single-port DDR memories. That doubling of the double-data-rate transfer speed results in a fourfold overall improvement over the standard single-data-rate memory.

Redundancy. As memory chips increase in storage capacity, manufacturing defects can often cause several storage cells to fail and thus render the chip useless. To counter that problem and thus improve production yield, many manufacturers have incorporated redundant storage cells that can be swapped into the memory array during unpackaged chip testing via the blowing of electrical fuses or the use of a laser to burn away microscopic fuses. Most static and dynamic RAMs now incorporate some degree of redundancy, and all memory types except the read-only memories typically include a small number of redundant cells as a hedge against defects.

Capacity and special features. In the early 1970s, dynamic memories with a capacity (density) of 1024 bits per chip were introduced. Since then, improvements in semiconductor processing and circuit design have made practical continual increases in density.

Traditional dynamic RAMs were introduced in organizations with single-bit word widths and depths that started at 1024 words. As the memory depths increased, the single-bit-wide memory chips did not provide the optimum increments (granularity) in all but the largest of systems. Thus, dynamic RAMs with 4-, 8-, 9- (the ninth bit is used for parity), 16-, and 32-bit-wide data paths have appeared to best match the architectures of diverse systems.

By the mid-1990s DRAMs with densities of 16 Mbits were commercially available, and although samples of 256-Mbit devices have been available since 1996, many manufacturers found it a little harder than expected to mass-produce such chips, with mass production starting in late 1999. In the interim, the 16- and 64-Mbit chips became the workhorse DRAMs, and in 1998 and 1999 a number of manufacturers introduced the first intermediate density, 128-Mbit DRAM. [In past generations manufacturers typically increased capacity by increments of 4X (1-k, 4-k, 16-k, 64-k, 256-k, 1-M, 4-M, 16-M, 64-M, 256-M, 1 Gbit, and so on); since one row and one column address signal must be added, the two extra address bits represent a fourfold increase in address range.]

Blocks of memory (notably DRAM and SRAM), functions such as FIFO buffers and multiport SRAMs, and flash EEPROMs are also available as building blocks in the design libraries of most manufacturers of application-specific integrated circuits (ASICs). Although SRAMs, FIFOs, and multiport SRAMs have been part of the design libraries since the 1980s, the mid-1990s saw the first commercial offerings of custom-designed logic chips that embedded large blocks of DRAM and large blocks of flash memory. Since then, DRAM blocks of up to 64 Mbits, SRAM blocks over 1 Mbit, flash memory blocks of up to about 4 Mbits, and blocks of FIFO and multiport memories have been available to designers of custom circuits.

The increased use of cache memories in the late 1980s spawned a new class of static memory chips with special features to ease their use in fast systems and eliminate the need for off-chip circuits to implement the functions. Some typical examples include individual bit-enable control lines on multibit outputs, parity bits carried through the memory, special cache-tag memories to carry the tag bits for the cache, special self-timed logic to permit extremely fast access times in pipelined systems, and very wide word sizes to reduce the chip count in very wide word but low word depth systems.

Access times for the fastest cache memory chips in the late 1990s were well below 5 ns, allowing them to keep pace with processors that operate at clock speeds of 300 MHz and higher. Still higher-speed processors introduced after 2000 actually eliminated the need for stand-alone cache memories since CPU designers incorporated both first- and second-level caches onto the same chip as the CPU. In cache subsystems, memory bandwidth is a key consideration, while memory depth for a second-level cache is typically limited to about 4 Mbytes. Thus memories with wide words are very popular for cache subsystems. In the cache-memory area, companies offer a range of memories with word widths of 8, 9, 16, 18, 32, 36, and 64 bits, and memory capacities of up to 8 Mbits (1 Mbyte) on a single chip. For small cache subsystems, a single cache-memory chip can provide all the desired memory.

Nonvolatile memories. There are many other forms of semiconductor memories in use—maskprogrammable read-only memories (ROMs), fuse-programmable read-only memories (PROMs), ultraviolet-erasable programmable read-only memories (UV EPROMs), electrically alterable read-only memories (EAROMs), electrically erasable programmable read-only memories (EEPROMs), flash EPROMs, nonvolatile static RAMs (NV RAMs), and ferroelectric memories. Most of these memory types are randomly accessible, but their main distinguishing feature is that once information has been loaded into the storage cells, the information stays there even if the power is shut off.

Some of these memories also come in specialty versions that were developed for a narrow range of applications. For instance, there are ROMs, UV EPROMs, and EEPROMs that are designed for serial interfaces to minimize chip pin count. There are also special UV EPROMs with small cachelike registers on the chip to deliver data at fast static-RAM-like speeds.

ROM. The ROM is programmed by the memory manufacturer during the actual device fabrication. Here, though, there are two types of ROMs: one is called late-mask or contact-mask programmable, and the other is often referred to as a ground-up design. In the late-mask type of ROM, the masks used in the final stages of the fabrication process determine the connections to the internal transistors. The connections, in turn, determine the data pattern that will be read out when the cell is accessed. The ground-up type of ROM is designed from the bottom up—all fabrication masks used in the multiple mask process are custom-generated—and that makes the chip slightly smaller and, in large volumes, less expensive than the late-mask version.

ROMs can be had with capacities of up to about 32 Mbits, with almost all chips available with an 8-bit word width. The increased use of 16-, 32-, and 64-bit microprocessors, however, has moved ROMs toward 16-bit-wide word organizations, and chips with such structures are available at the 1-Mbit level and above. Additionally, at such densities the cost of manufacturing the chip and the manufacturing losses due to imperfections that destroy stored data bits require that some method to correct bad bits be used. Thus, ROMs with built-in error detection and correction circuits have been put into limited production. The use of ROMs, however, has been declining since the late 1990s as the cost of flash memory devices and other nonvolatile storage options has decreased.

Fuse PROM. As an alternative to the mask-programmable memories, all the other nonvolatile memory types permit the users to program the memories themselves. The fuse PROM is a one-time programmable memory—once the information is programmed in, it cannot be altered. One type uses microscopic fuse links that are blown open to define a logic one or zero for each memory cell; the other type causes metal to short out base-emitter transistor junctions. The latter type that forms a short in the base-emitter junction connection is the predecessor to what today is referred to as antifuse technology. Antifuses are one-time programmable elements used to configure some field-programmable logic devices.

Most fuse PROMs were fabricated with bipolar transistors, and a few were implemented with complementary metal-oxide semiconductor (CMOS) technology. However, advances in CMOS technology and stored-charge-based nonvolatile memory structures have caused an almost total decline in the use of fuse-based memories. *See* TRANSISTOR.

Reprogrammable memories. The birth of the microprocessor in the early 1970s brought with it nonvolatile memory types that offered reusability. Information stored in the memory can be erased—in the case of the UV EPROM, by an ultraviolet light, and in the case of the EAROM, EEPROM, flash EPROM (often referred to as just a flash memory device), nonvolatile (NV) RAM, or ferroelectric memory, by an electrical signal. Then the circuit can be reprogrammed with new information that can be retained indefinitely. All of these memory types are starting to approach the ideal memory element for the computer, an element that combines the flexibility of the RAM with the permanence of the ROM when power is removed. *See* MICROPROCESSOR.

The ideal memory circuit should allow unlimited read and write operations without any unusual voltage levels or extra circuitry. At present, UV EPROMs require an external ultraviolet lamp to erase the stored information and a programming voltage of about 12.5 V for the newer versions, compared to a 5- or 3.3-V supply for the read-only mode. Most of the newer EAROMs and EEPROMs have been able to place control and timing circuitry on the same chip to boost the standard 5-V or 3.3-V supply voltage up to the level necessary to program or erase the cell, as well as provide the other necessary support functions. Low operating voltage is also important for nonvolatile memories used in portable systems, and many of the latest versions of these memories can operate from 3.3-V or lower power supplies.

An offshoot of the EEPROM and UV EPROM technologies is the flash EEPROM and flash EPROM. Unlike most EEPROMs, which allow the user to selectively change a single word, the flash memory permits the user to erase only the contents of the entire chip before storing new data into the chip. Some of the flash memories allow some degree of block erasure, typically one-half or one-sixteenth of the memory array. The limited electrical alterability makes possible significantly simpler circuit designs and chips that are higher in capacity, smaller in area, and thus less expensive.

Flash memories have evolved along two architectural paths. The first path consists of a standard RAM-like architecture that provides full random access to any data word within the memory chip. Such memories are based on a form of NOR-type storage cell (a few manufacturers employ an OR-type structure which is very similar) [**Fig. 8***a*]. The second architectural path consists of a disk-drive-like access scheme that incurs some system overhead to access the first byte (or word) of data specified by the address and then allows the remaining data in the desired block

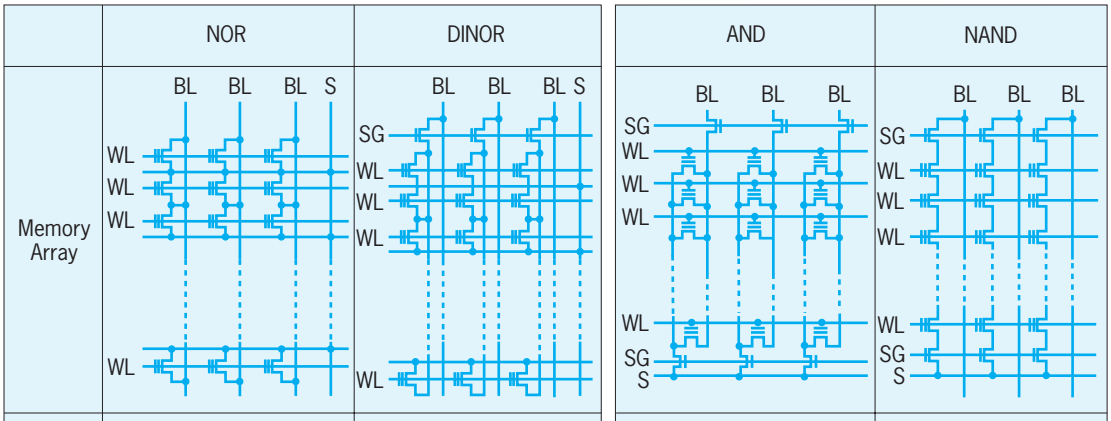

(a) (b)

Fig. 8. Flash memory architectures have taken two major directions: (*a*) Random access. A NOR-type (occasionally, OR) memory cell structure is used at the heart of a full random-access memory architecture. (*b*) Serial access. The byte-serial architecture is based on NAND-gate-like strings of series-connected MOS transistors. In some architectures an AND array is employed instead of NAND. BL = bit line; S = select; SG = select gate; WL = word line; FN = Fowler-Nordheim. (*Mitsubishi Electronics America Inc.*)

to be read out from or written into the memory in a byte (word) serial fashion using just timing and control signals. Since this second architectural approach does not employ a full random-access architecture, the internal memory array is much denser than the NOR/OR-type array. Part of that density is attributable to the very different memory cell approach used in the heart of the arrays. *See* LOGIC CIRCUITS.

Most of the byte-serial devices employ a NAND-based memory cell in which eight transistors are serially connected such that the source of one is connected to the drain of the next, and so on (Fig. 8*b*). [Alternatively, several companies have developed serial memories that are based on an AND-connected multitransistor cell structure but at the chip level perform much like the NAND-based memories.] The serial NAND (or AND) architecture is well suited for applications that might otherwise rely on electromechanical disk drives, but can operate 10–100 times faster than the disk drives. Memory cards using such chips are in widespread use in digital cameras, personal digital assistants, solid-state dictation systems, and many other applications.

Coming closest to the ideal specification is the NV RAM, a memory that combines a static RAM with a nonvolatile memory array, so that for every stored bit there are two memory cells, one of which is volatile and the other nonvolatile. During normal system operation, the NV RAM uses the volatile memory array, but when it receives a special store signal, information held in the RAM area is transferred into the nonvolatile section. Thus the RAM section provides unlimited read and write operations, while the nonvolatile section provides back-up when power is removed.

Another method employs a ferroelectric storage element deposited on top of a static RAM cell (or even a dynamic RAM cell). The potential advantages of the ferroelectric technology include a large number of storage cycles with little increase in chip area since the ferroelectric material can be deposited on top of the memory cells. However, the technology is still in the advanced experimental stage, with only a couple of companies offering devices with relatively low densities (16 kbits to 1 Mbit).

However, the EAROM, EEPROM, flash, NV RAM, and ferroelectric RAMs suffer one common failing that keeps them from reaching the ideal—they wear out. The electrical process used to store information in the nonvolatile array causes a steady deterioration in the ability of the memory to retain data for a guaranteed time period. Advances in materials technology are improving the number of store (write) cycles that the EAROM, EEPROM, flash, and ferroelectric memories can endure. Most commodity EEPROM and flash memories, for instance, can typically endure at least 100,000 write cycles, although a large number of suppliers can offer endurances of 1 million store cycles. Ferroelectric-based memories can offer still higher numbers of store cycles, often 10–100 times higher than flash or EEPROM-type devices.

Another new storage technology disclosed in 2000 is the magnetoresistive-RAM. The nonvolatile technology used to form the memory cells consists of two

ultrathin layers (tens of angstroms) of magnetically responsive materials separated by a thin dialectric layer. The memory cells can retain data indefinitely and have extremely high endurance levels since this technology does not have a wearout mechanism such as is found in flash memories.

To overcome the limited number of storage cycles possible with the true nonvolatile memory technologies, ultralow-power-consumption static RAMs (with a standby current of just tens of nanoamperes) have been combined with some power control circuitry and a button-type long-life battery in a single package. When built into a system, this form of NV RAM operates with the characteristics of a static RAM when the system supplies power—fast access times and unlimited reads and writes. However, when the power supply line drops below a preset level, the control circuitry in the memory package detects the drop, disables the memory's ability to store new data, and then switches over to the built-in battery to keep the stored data alive until main power returns.

Semiconductor technology. The semiconductor technology used to fabricate all the different memory types spans the entire range of available commercial processes (see **table**). All dynamic and static RAMs and most nonvolatile memories are fabricated only in CMOS (and in a few cases a merged bipolar-CMOS process). The use of CMOS

Summary of major memory types, performance levels, and technologies

Memory type	Maximum capacity for commercial chips, 2001–2002	Typical access speed/cycle time	Programmability	Technology
Dynamic RAM (standard, fast page, and extended data out)	No longer in widespread use	40 ns access, 60 ns cycle	Read/write, volatile	CMOS
Dynamic RAM (synchronous, single data rate and double data rate)	256 Mbit/512 Mbit (plus limited 1 Gbit production)	20–40 ns first access, subsequent sequential access at 4–8 ns/access, 125 MHz SDR, 250 MHz DDR	Read/write, volatile	CMOS
Dynamic RAMs (Rambus)	256 Mbit	1.6 ns access, 660 MHz clock	Read/write, volatile (8- or 16-bit sequential access)	CMOS
Dynamic RAMs (Sync Link)	64 Mbit (never commercialized)	2 ns access, 500 MHz clock	Read/write, volatile	CMOS
Cache DRAM	16 Mbit	5 ns access, 200 MHz clock	Read/write, volatile	CMOS
Multibank DRAM	64 Mbit	5 ns access, 200 MHz clock	Read/write, volatile	CMOS
Video RAM	16 Mbit (being phased out)	8 ns, serial port, 50 ns, random access port	Read/write, volatile	CMOS
Synchronous graphics RAM	32 Mbit (being phased out)	5 ns access, 200 MHz clock	Read/write, volatile	CMOS
Embedded DRAM	64 Mbits	5 ns access, 200 ns clock	Read/write, volatile	CMOS
Asynchronous static RAMs (no bus latency)	16 Mbits	20 ns access, 20 ns cycle	Read/write, volatile	CMOS
Synchronous static RAMs	16 Mbits	5 ns access, 8 ns cycle	Read/write, volatile	CMOS
Synchronous burst SRAM	8 Mbits	3 ns access, 5 ns cycle	Read/write, volatile	CMOS
FIFO buffers	4 Mbits	5 ns access, 5 ns cycle	Read/write, volatile	CMOS
Mulitport SRAMs	4 Mbits	5 ns access, 5 ns cycle	Read/write, volatile	CMOS
PROM	Being phased out and replaced with electrically reprogrammable memories	10 ns access	Electrically one-time programmable, nonvolatile	Bipolar, CMOS
UV EPROMs	32 Mbits (no new development)	60 ns access	UV erasable, electrically programmable cell	CMOS
EEPROM	1 Mbit	100 ns access, less than 10 μs/word for storage	Electrically programmable and erasable on a word-by-word basis, nonvolatile	CMOS
Flash EPROM (NOR type)	256 Mbits	50 ns access, less than 10 μs/word for storage	Bulk or block erasable, programmable on a word-by-word random basis, nonvolatile	CMOS
Flash EPROM (NAND type)	512 Mbits	60–100 ns for the first access, 20 ns for subsequent sequential accesses	Bulk or block erasable, programmable in word sequential order	CMOS

memory structures (the combination of both *p*- and *n*-channel devices to form the memory cells and supporting logic on the memory chip) permits a reduction of orders of magnitude in the standby power consumption, and often a reduction by a factor of 2–5 in the active power requirements of a memory chip, as compared to the older *n*-channel MOS (NMOS) equivalent.

The speed at which dynamic RAMs can be accessed has been shortening as each generation has been introduced. Static RAM access times have also improved over the years, and CMOS static RAMs with access times of as little as 10 ns are considered commodity products. Devices with access times as short as 6 ns are also readily available, with most of them targeted at applications such as cache subsystems and high-speed signal processing. Of course, not every application demands the minimum access time, and many speed grades are available, up to several hundred nanoseconds.

For even faster access times, merged bipolar-CMOS processes and full bipolar technologies such as emitter-coupled logic (ECL) have traditionally been used to offer access times of 3–5 ns for commercial memories. However, except for a few very specialized applications, CMOS technology has been able to catch up with the performance of most bipolar memories and to offer higher capacities as well. The use of gallium arsenide processes has led to the development of memory chips with access times of less than 3 ns at density levels of 1, 4, and 16 kbits.

Access times for the various nonvolatile memories range from less than 35 ns in their read modes to about 10 μs when they store data. The flash-style EEPROMs typically have access times in the 50–100-ns range, while memory capacities have reached levels of 256 Mbits. In 2001, prototypes of 512-Mbit flash memory chips were described at several technical symposia. In addition to the basic evenly sectored architecture, specialty flash versions referred to as boot-block flash memories and special serial-file flash memories have been developed.

The boot-block approach was developed specifically for use in personal computers to hold the basic input/output software control (BIOS) programs. The address space is divided into several unequal regions, so that part of the memory remains secured while other parts of the memory can be overwritten to update the control software.

Serial file memories were developed for use as a solid-state alternative to the electromechanical hard-disk drives. File memories access data words serially and are thus optimized for use with data that come in large blocks where all the words in the block can be transferred sequentially (such as when loading a program file or transferring a large table of data). Thus, rather than allowing random access to any stored data word, the file memories are designed to operate somewhat like a disk drive. The first access to a desired block of data contains several microseconds of overhead (much like the positioning of the read/write head over the right track in a disk drive), but then sequential accesses to the rest of the data block can be done at typicial random-access-like speeds (35–100 ns per transfer). When multiple chips are combined into an array, solid-state drive alternatives with capacities of 2 to more than 500 Mbytes can be implemented in a credit-card-sized format.

Two other innovative flash-based memory architectures first appeared at the 16-Mbit density level. The first employs a dynamic RAM-like interface so that it can be used in systems that would otherwise transfer code from slow ROMs into a faster dynamic or static RAM and then execute the code from the RAM to take advantage of the shorter access times of the RAMs. The second unique architecture takes the same base 16-Mbit flash memory array, but then integrates a synchronous interface on the chip so that it appears to function much like a clocked static RAM, thus reducing the access time to about 35 ns, about half that of the standard flash memory.

A novel storage scheme that was introduced in 1994 employs a charge-sensing scheme to double the number of bits that a chip can store. The scheme divides the charge held in the nonvolatile memory cell into four levels and, through the use of on-chip voltage comparators, will differentiate the charge values to allow two bits to be stored in each cell. Thus a 32-Mbit memory chip can be built on roughly the same size chip as a 16-Mbit memory. This double-density technique has been applied to produce flash memory chips with capacities up to 512 Mbits, and still higher-capacity chips can leverage this approach as process improvements allow the fabrication of chips with more than 256 million actual storage cells.

Additionally, a unique memory chip that combines UV EPROM, on-chip static RAM, and programmable logic was released in 1988. The programmable logic allows portions of the UV EPROM to be mapped into a different portion of the system's address space. The remapping can be used to improve addressing speed or to provide security by scrambling the address. *See* COMPUTER STORAGE TECHNOLOGY; INTEGRATED CIRCUITS.

David Bursky

Bibliography. C. Hu, *Nonvolatile Semiconductor Memories: Technologies, Design, and Applications*, 1991; Institute of Electrical and Electronics Engineers, *IEEE Standard Terminology of SSemiconductor Memories*, IEEE Stand. 662–1992, 1992; B. Prince, *Semiconductor Memories: A Handbook of Design, Manufacture, and Application*, 2d ed., 1991; A. K. Sharma, *Semiconductor Memories: Technology, Testing, and Reliability*, 1996; A. J. Van de Goor, *Testing Semiconductor Memories: Theory and Practice*, 1991.

Semiconductor rectifier

A semiconductor diode that is used in rectification and power control. The semiconductor diode conducts current preferentially in one direction and inhibits the flow of current in the opposite direction by utilizing the properties of a junction formed from two differently doped semiconductor materials (**Fig. 1**). Doped silicon is by far the most widely used semiconductor. Semiconductor diodes are intrinsic

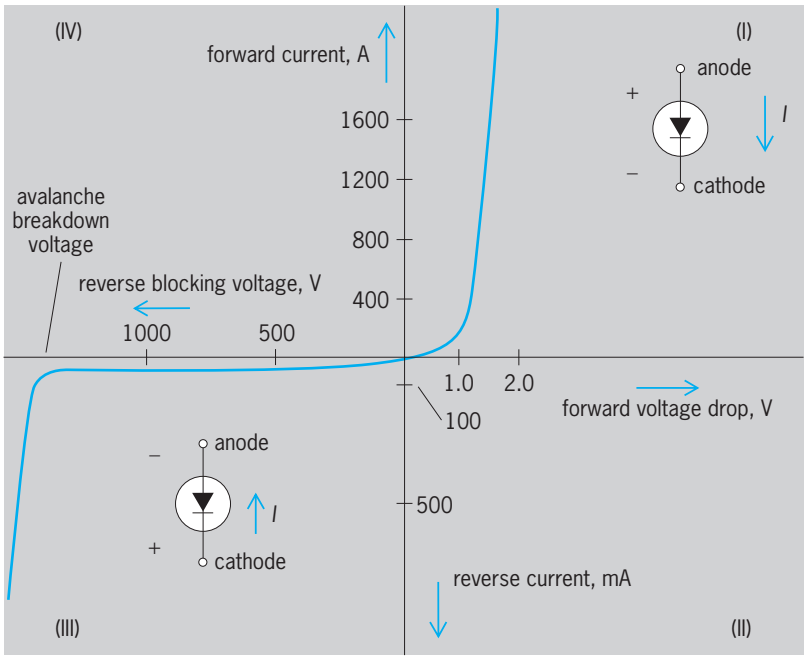

Fig. 1. Current-voltage characteristic of a commercially available 250-A silicon rectifier diode. Different voltage and current scales are used for positive and negative values.

to integrated circuits and discrete device technology and are used to perform a wide variety of isolation, switching, signal processing, level shifting, biasing, control, and alternating-current (ac) to direct-current (dc) conversion functions. This article focuses on using the semiconductor diode for ac-to-dc conversion, called rectification, as well as ac and dc power-control applications. *See* CONTROLLED RECTIFIER; RECTIFIER; SEMICONDUCTOR; SEMICONDUCTOR DIODE.

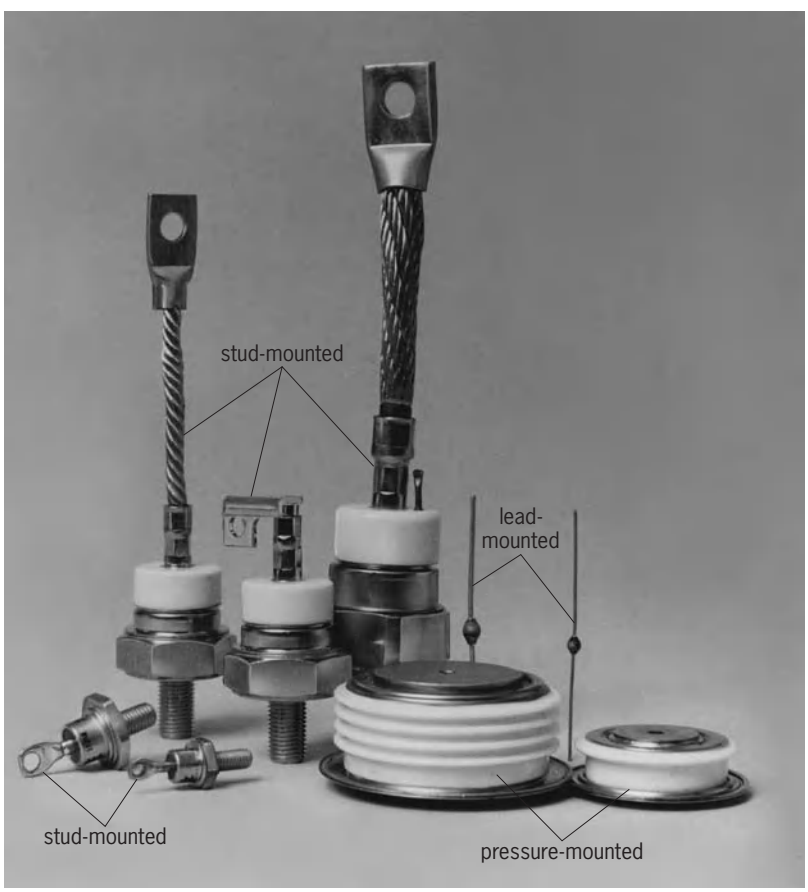

Fig. 2. Typical rectifier diode packages which allow heat dissipation and proper electrical contact. (*General Electric Co.*)

Either as a key element of an integrated circuit or as a discrete packaged part, the silicon rectifier diode is used in a plethora of applications from small power supplies for consumer electronics to very large power-rectification industrial installations. Many semiconductor diodes are used in non-power-conversion applications in signal processing and communications. These include avalanche or Zener diodes; diodes used for amplitude-modulation radio detection, mixing, and frequency translation; IMPATT, PIN, and step-recovery diodes, used at microwave frequencies; diodes fabricated from gallium arsenide and related compounds, used in optoelectronics; and light-emitting diodes (LEDs) and solid-state lasers. *See* AMPLITUDE-MODULATION DETECTOR; LASER; LIGHT-EMITTING DIODE; MICROWAVE SOLID-STATE DEVICES; MIXER; ZENER DIODE.

Silicon rectifier diodes. The electrical heart of the semiconductor diode is the junction between *p*-type and *n*-type doped silicon regions. The primary processing technology used to form a *pn* junction is high-temperature diffusion. In the diffusion process, a chemically very pure wafer of single-crystal silicon is exposed to a hot vapor of a dopant material. *See* JUNCTION DIODE.

The die, as a portion cleaved from a wafer, has a thickness on the order of 0.5 mm for a 100-mm-diameter (4-in.) wafer. For integrated circuit diodes, the current levels are low, in the range of a few microamperes, so that the junction dimensions are as small as a few micrometers on a side. For high-power applications rated in the thousands of amperes, the die could be as large as the entire wafer. The final die dimension is determined predominantly by current-handling requirements and packaging considerations. Advanced semiconductor processing technology utilizes 150-mm (6-in.) and even 200-mm (8-in.) wafers, although these are not used in whole-wafer form for semiconductor rectifiers because of economic factors. The die must be mounted in a suitable hermetically sealed protective package which allows the heat to be dissipated to the external environment and allows for proper electrical contact. Usually packages are lead-mounted for currents up to a few amperes, stud-mounted and flat-base mounted for currents up to a few hundred amperes, and external-pressure-mounted for the largest devices capable of conducting thousands of amperes (**Fig. 2**). The lead-mounted package dissipates the losses directly into the environment, while the other package styles are mounted on heat sinks for natural-convection-air, forced-air, or liquid cooling. *See* INTEGRATED CIRCUITS.

Discrete silicon diodes are commercially available with forward-current specifications from under 1 A to several thousands of amperes. Diodes may be connected in parallel for greater current capability as long as the design provides for the current being uniformly distributed between the parallel diodes.

This is usually done with a ballast resistor in series with each diode. *See* BALLAST RESISTOR.

Ideally, the current through a reverse-biased diode, called the saturation current (I_S) or reverse current (I_R), approaches zero (Fig. 1). Practically speaking, this current is several orders smaller than the forward current (I_F). The maximum value of the reverse blocking voltage is limited primarily by the structure and doping of the semiconductor layers. (The package design must be appropriate for the maximum design voltage.) This maximum voltage is referred to as the avalanche breakdown voltage, or the peak reverse voltage (PRV) or peak inverse voltage (PIV). It is a very important parameter for power supply and power conversion designs. Exceeding the peak inverse voltage is usually destructive unless the circuit design provides for limiting the avalanche current and resultant heating. In summary, at positive voltages and currents (quadrant I of the voltage-current characteristic; Fig. 1), the silicon rectifier diode shows the on-state conducting characteristic, with high current and low forward voltage drop; at negative voltages and currents (quadrant III), it shows the reverse-blocking or reverse-bias, off-state characteristic, with high blocking voltage and low (ideally zero) reverse blocking current. *See* ELECTRICAL BREAKDOWN.

Integrated-circuit diode-junction avalanche breakdown voltages are of the order of several tens of volts. Single silicon rectifier diodes designed for power conversion applications are available with ratings from a few hundred to a few thousand volts. Several diodes can be connected in series for greater voltage capability. Prepackaged series diode strings can be rated to tens of thousands of volts at several amperes. This series connection must ensure equal voltage division across each diode to guard against catastrophic failure of the entire series. Typically this is done by including a high-value equal-value resistor in parallel with each diode to obtain equal voltages, and a parallel capacitor to provide a low-impedance path for high-voltage transients that are often present in industrial environments.

Schottky silicon rectifier diodes. Unlike a silicon diode formed from a *pn* junction, the Schottky diode makes use of the rectification effect of a metal-to-silicon interface and the resultant barrier potential. The Schottky diode, sometimes called the Schottky-barrier diode, overcomes the major limitation of the *pn* junction diode; being a majority carrier device, it has a lower forward voltage drop (0.2–0.3 V, compared to 0.7–1.0 V) and faster switching speed than its minority-carrier *pn* junction counterpart. However, other factors confine its use to low-voltage power applications, chiefly the relatively small breakdown voltage, typically 45 V. Secondary shortcomings include a high reverse current and restricted temperature of operation, with commercial devices providing a maximum of 175°C (347°F) compared with 200°C (392°F) for *pn* junction diodes.

Integrated circuits used in computer and instrument systems commonly require voltages less than 15 V and as low as 3.3 V. Thus the advantage of low forward-voltage drop and faster switching favors the Schottky diode. This is particularly true for high-frequency switching voltage regulator power supply applications where voltages at 20–50 kHz must be rectified. The higher reverse current can be tolerated. However, cooling or heat sinking is more critical because of the higher reverse-current temperature coefficient and lower maximum operating temperature. *See* SCHOTTKY BARRIER DIODE.

Rectifier circuits. The greatest usage of rectifier diodes is the conversion of ac to dc. The single diode of a half-wave rectifier for a single-phase ac voltage (**Fig. 3***a,b*) conducts only on the positive half-cycle.

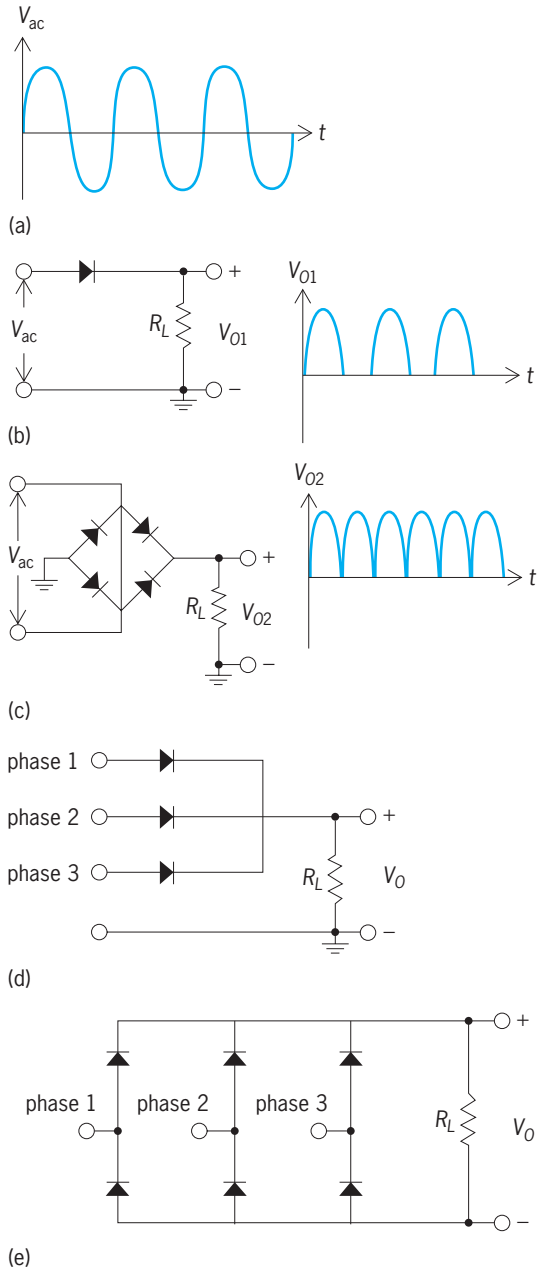

Fig. 3. Rectifier circuits. (*a*) Waveform of single-phase ac input voltage, V_{ac}, for both half-wave rectifier and full-wave rectifier circuits. (*b*) Half-wave rectifier circuit, showing rectification of input on the positive half-cycles to a pulsating dc output voltage, V_{O1}. R_L = load resistance. (*c*) Full-wave bridge rectifier circuit, showing rectification of input on both the negative and positive half-cycles to a pulsating dc output voltage, V_{O2}. (*d*) Half-wave and (*e*) full-wave rectifier circuits for three-phase ac inputs.

Because of this, the output voltage across the load resistance is unidirectional and has a nonzero average value. This output waveform is called a pulsating dc. Therefore the input ac voltage has been rectified to a dc voltage. For most applications, a filter, usually consisting of large electrolytic capacitors, must be employed at the output to smooth the ripple present on the pulsating dc voltage to come close to a constant dc voltage value. *See* CAPACITOR; ELECTRIC FILTER; ELECTRONIC POWER SUPPLY; RIPPLE VOLTAGE.

A widely used diode circuit for rectification of both the positive and negative half-cycles of the single-phase ac supply voltage is the full-wave bridge rectifier (Fig. 3*c*). It is commercially available as a single package. The output voltage is still a pulsating dc voltage, but the average value will be double that of the half-wave rectifier.

Many high-power industrial applications require the use of three-phase power (Fig. 3*d,e*). The output waveform exhibits less ripple voltage and is easier to filter. *See* ALTERNATING CURRENT.

In lower-power applications from a few watts to a few hundred watts, such as used in computers, television receivers, and laboratory instruments, a switching voltage regulator is commonly used to generate a 10-kHz–50-kHz ac signal from the high-ripple ac power supply voltage. The advantage is the ease and lower cost in filtering the ripple resulting from rectifying high-frequency ac as opposed to filtering low-frequency ac. *See* VOLTAGE REGULATOR.

Thyristors. Whereas the basic semiconductor rectifier has two terminals, an anode and cathode, a silicon controlled rectifier (SCR) has three terminals: an anode, cathode, and control electrode called the gate (**Fig. 4**). The silicon controlled rectifier is a four-layer device modeled as two interconnected *pnp* and *npn* transistors.

Normally, there is no current flow from the anode to the cathode. Both transistors are off; that is, they are blocking any current flow. By applying a relatively small trigger pulse control signal to the gate electrode, the *npn* transistor is switched on. When the *npn* transistor is switched on, the *pnp* transistor is also switched on. Consequently the silicon controlled rectifier is turned on and a current flows through the silicon controlled rectifier and external circuit. The resultant internal voltages keep both the *npn* and *pnp* transistors on even when the gate voltage is removed. The device is said to exhibit regenerative, positive-feedback, or latching-type switching action. There is a voltage drop of about 1 V across the on-state silicon controlled rectifier. The power dissipation rating required in specifying a silicon controlled rectifier is given by this 1-V drop multiplied by the peak current flowing through the device. *See* TRANSISTOR.

Current continues to flow even when the gate signal is reduced to zero (Fig. 4*d*). To reset the silicon-controlled rectifier, the external current must be reduced below a certain value. Thus, the thyristor can be switched into the on state (conducting condition) by applying a signal to the gate, but must be restored to the off state by circuit action. If the anode current momentarily drops below some holding current or if the anode voltage is reversed, the silicon controlled rectifier reverts to its blocking state and the gate terminal regains control. Typical silicon controlled rectifiers turn on in 1–5 microseconds and require 10–100 μs of momentary reverse voltage on the anode to regain their forward-blocking ability.

Other semiconductor diode topologies are also used for power control. A generic term for these power-control devices is the thyristor (**Fig. 5**).

The silicon controlled rectifier operates only in the first quadrant (forward current and voltage), that is, on either the positive or negative half-cycle of

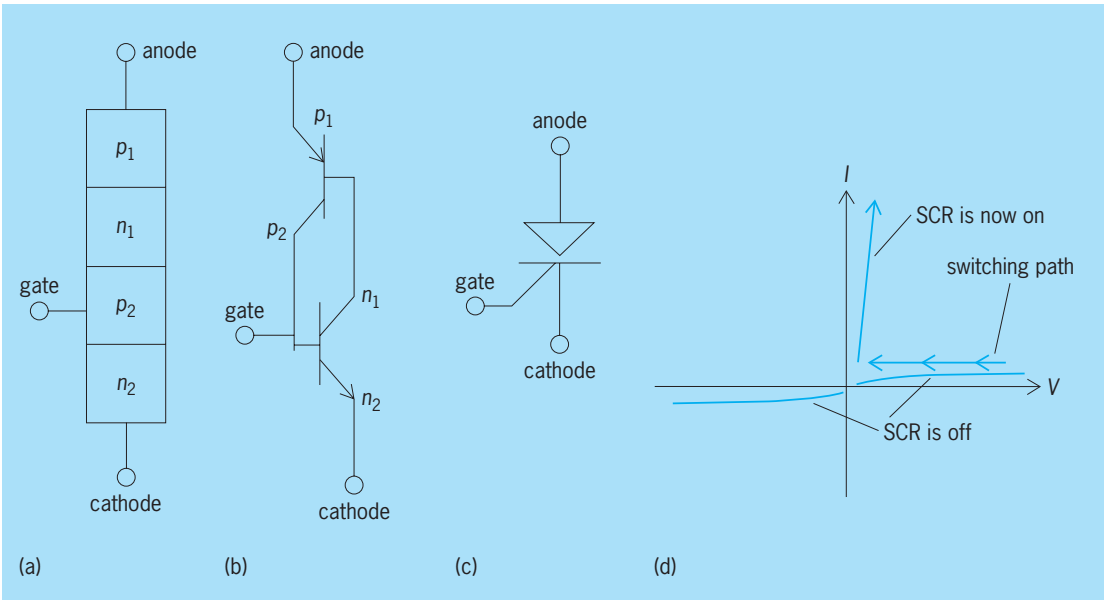

Fig. 4. Silicon controlled rectifier (SCR). (*a*) Four-layer structure. (*b*) Interconnected *npn-pnp* transistor model. (*c*) Preferred schematic symbol. (*d*) Current-voltage (*I-V*) characteristic. The switching path is followed when a voltage pulse is applied to the gate electrode of the *npn* transistor, switching the SCR to on.

Type of thyristor	Common circuit symbol	Current–voltage (I–V) characteristics
reverse-blocking triode thyristor (SCR)	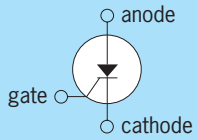	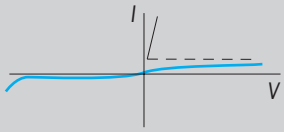
bidirectional triode thyristor (TRIAC)	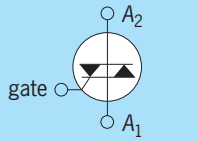	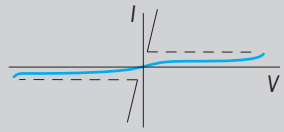
asymmetrical triode thyristor (ASCR)	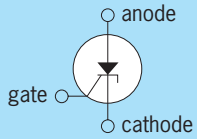	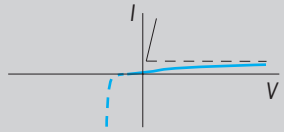
reverse-conducting triode thyristor (RCT)	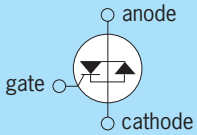	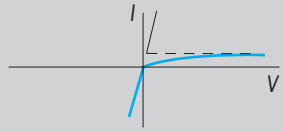
light-activated SCR (LASCR)	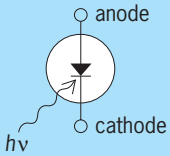	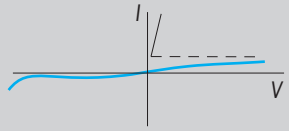
commutable reverse-blocking triode thyristor (GTO)	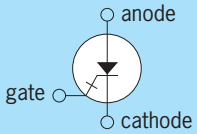	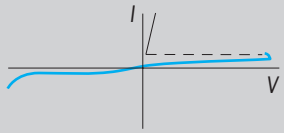
MOS controlled SCR (MOS SCR)	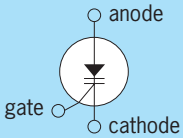	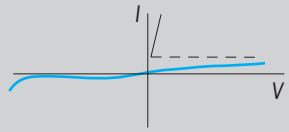
MOS controlled thyristor, TRIAC (MCT)	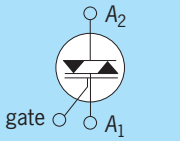	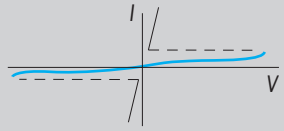

Fig. 5. Family of thyristor devices.

the ac voltage. For specialized ac-switching power control such as lamp dimmers and variable heating where bidirectional control is required, the bidirectional triode thyristor, called the TRIAC, is widely used. Its current-voltage characteristics (Fig. 5) are the same in the first and third quadrants as the first-quadrant silicon controlled rectifier characteristics. Functionally, the TRIAC behaves like two oppositely oriented silicon controlled rectifiers connected in parallel. The asymmetrical triode thyristor (ASCR), reverse conducting triode thyristor (RCT), and commutable reverse-blocking triode thyristor (GTO) are used in power circuits requiring operation from a dc source.

The asymmetrical triode thyristor has a much lower breakdown voltage (Fig. 5), which presents no problem in dc circuits because there is no need to block a large reverse voltage. The asymmetrical triode thyristor has a faster turn-off time, 5–25 μs, thus requiring a less costly auxiliary circuit to effect turn-off. The reverse conducting triode thyristor is the integrated equivalent of a discrete asymmetrical triode thyristor and includes a parallel discrete diode rectifier (Fig. 5). This provides a lower breakdown voltage. The reverse conducting triode thyristor also has a shorter turn-off time than the silicon-controlled rectifier. In general, both the asymmetrical triode thyristor and the reverse conducting triode thyristor have somewhat lower forward-voltage drops for comparable forward-voltage blocking ratings and silicon die area, thus decreasing the device power dissipation and increasing efficiency.

The gate trigger signal can also be generated optically. If a light pulse is allowed to shine on the second *pn* junction (p_2-n_2; Fig. 4*a*) of a light-activated silicon-controlled rectifier (LASCR), the device will be switched to the on state (Fig. 5). This type of device is used where electrical isolation between the gate-control circuit and output circuit is required as well as to detect the presence of opaque objects in a process assembly line. Its characteristics are similar to a standard silicon-controlled rectifier.

The gate turnoff device (GTO) is a symmetrical reverse-blocking triode thyristor which can block reverse voltages (Fig. 5). In addition, it can turn off anode current when a negative voltage is applied to the gate. Thus it does not require an auxiliary circuit to commutate it off as do the silicon controlled rectifier, asymmetrical triode thyristor, and reverse conducting triode thyristor devices.

All of the thyristors discussed so far are based upon switching a *pn* junction. The trigger circuit must provide adequate current to forward-bias the *pn* junction. A thyristor specification will include the gate turn-on voltage (V_{GT}) as well as the gate turn-on current (I_{GT}). In order to switch a thyristor from metal-oxide-semiconductor field-effect transistor (MOSFET) logic circuits or from low-power bipolar circuits, a circuit must be included in the gate trigger that can interface between the low-current drive capability of the switch-control circuitry and the higher-current requirements for switching on the thyristor. By including negative MOS (NMOS)

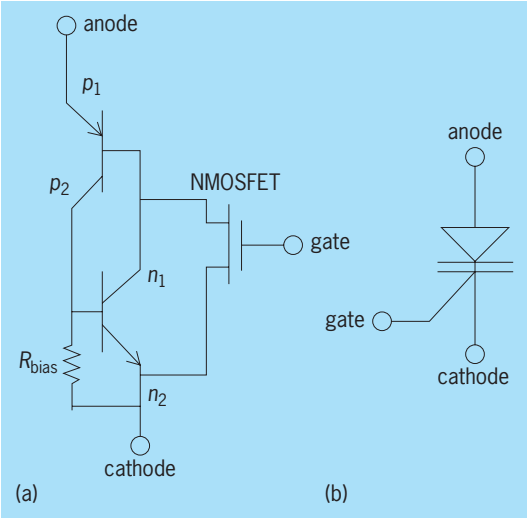

Fig. 6. MOS-controlled thyristor. (*a*) Model as interconnected *pnp* and *npn* transistors gated by an NMOSFET (negative metal-oxide-semiconductor field-effect transistor). (*b*) Schematic symbol.

technology in the thyristor fabrication (**Fig. 6**), the silicon-controlled rectifier can be switched from a voltage source that does not need to provide any significant current. The MOS controlled silicon controlled rectifier (MOS SCR) has high input impedance and requires a very low gate drive current (Fig. 5). Another type of MOS controlled thyristor (MCT) has a TRIAC-type characteristic (Fig. 5). MOS thyristors can be switched with a gate turn-on voltage of 15 V and control current densities of 1000 A/cm^2.

Forward current ratings for a single thyristor range from 1 to 5000 A, and peak reverse voltages may exceed 4000 V. Thyristor circuit topologies are found within many types of voltage-regulator integrated circuits and are used for short-circuit protection.

Thyristor circuits and applications. Thyristor applications fall into two general categories. The devices can be used from an ac supply, much like silicon rectifier diodes. However, unlike the rectifier diode, which conducts load current as soon as the anode voltage exceeds about 0.7 V, the thyristor will not conduct load current until it is triggered into conduction.

Therefore, the power delivered to the load can be controlled. This mode of operation is called ac phase control. It is extensively used in applications requiring conversion from ac to variable-voltage dc output, such as adjustable-speed dc motor drives, and in lighting and heating control. *See* DIRECT-CURRENT MOTOR.

The other category of applications is operation in dc circuits. This allows power conversion from a battery or rectified ac line to a load requiring either an alternating supply (dc-to-ac conversion) or a variable-voltage dc supply (dc-to-dc conversion). Since the rate of switching the thyristors in dc circuits can be varied by the control circuit, a thyristor inverter circuit can supply an ac load with a variable frequency. The fundamental approach in both cases is to

convert a dc voltage to a chopped voltage of controllable duty cycle. Changing the duty cycle either at a variable rate (frequency power modulation) or by varying the pulse width at a fixed frequency (pulse-width power modulation) effectively controls the power delivered to the load. *See* CHOPPING; PULSE MODULATION.

The major distinguishing feature of dc operation is the absence of any reversal of supply voltage polarity for turning off the thyristor, which would allow the gate electrode to regain control of the device. Either a gate turnoff device must be used, or an auxiliary circuit must be employed to effect turnoff of a silicon-controlled rectifier. One common technique to accomplish this is to switch a previously charged capacitor across the load-carrying silicon-controlled rectifier in such a manner that the voltage on the capacitor reverse-biases the silicon-controlled rectifier sufficiently to reduce the load current through it to zero, and then to allow the device a short time (about 10–50 μs) before reapplying the forward blocking voltage to the anode.

Important applications for dc-to-dc conversion, dc-to-ac power conversion at variable frequency, and dc-to-ac power conversion at fixed frequency are, respectively, control of battery-powered industrial vehicles such as forklift trucks and mining locomotives, adjustable-speed operation of ac synchronous and induction motors in industrial processing, and power transmission conversion. *See* ALTERNATING-CURRENT MOTOR; CONVERTER; INDUSTRIAL TRUCKS.

Stanley G. Burns

Bibliography. B. J. Baliga, *Power Semiconductor Devices*, 1996; *Electronic Engineer's Master Catalog (EEM)*, 4 vols., annually; J. T. Humphries and L. P. Sheets, *Industrial Electronics*, 4th ed., 1993; B. G. Streetman, *Solid State Electronic Devices*, 5th ed., 1999; A. Vladimirescu, *The SPICE Book*, 1993.

Semionotiformes

An order of actinopterygian fishes which appeared first in the upper Permian, reached maximum development in the Triassic and Jurassic, and persists in the Recent fauna as the gars (*Lepisosteus*). The gars are ranked by some as a separate order, called variously Lepisosteiformes, Lepisostei, and Ginglymodi. In the Semionotiformes the body is encased in a heavy armor of interlocking ganoid scales, which are thick, are more or less rhomboidal, and have an enamellike surface. The body form is diverse, and such Mesozoic forms as *Dapedius*, *Lepidotus*, and *Semionotus* were deep-bodied or stocky fishes with a blunt or moderate snout. Modern forms, however, are elongate and have bony jaws provided with enlarged conical teeth. The opposed dorsal and anal fins are short and placed far back on the body, and the caudal fin is abbreviated heterocercal. The vertebrae are distinctive in being opisthocoelous. There are paired prevomers and no gular plates, and the swim bladder is highly vascularized, thus permitting aerial respiration. The single Recent family, Lepisosteidae,

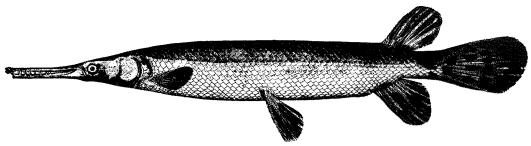

Spotted gar (*Lepisosteus oculatus*). (*After G. B. Goode, Fishery Industries of the United States, 1884*)

contains one genus, *Lepisosteus*, with seven species (see **illus.**). Although fossils are known from all continents, *Lepisosteus* is now restricted to lowland fresh and brackish waters of North and Central America.

Gars are despised by fishers because of their predation on valuable fishes. They are themselves caught with difficulty and are little esteemed as food. The roe is highly toxic when eaten by warm-blooded animals. *See* ACTINOPTERYGII; SWIM BLADDER.

Reeve M. Bailey

Bibliography. H. B. Bigelow (ed.), *Fishes of the Western North Atlantic*, Sears Found. Mar. Res. Memo. 1, pt. 3, 1968.

Sendai virus

A member of the viruses in the type species Parainfluenza 1, genus *Paramyxovirus*, family Paramyxoviridae; it is also called hemagglutinating virus of Japan (HVJ). Sendai virus was originally recovered in Sendai, Japan, from mice inoculated with autopsy specimens from newborns who died of fatal pneumonitis in an epidemic in 1952. Subsequent attempts to isolate this virus from humans were, however, mostly unsuccessful, although mice are commonly infected with Sendai virus along with rats, guinea pigs, hamsters, and pigs. It is believed that the natural host of Sendai virus is the mouse and that the virus is usually nonpathogenic for humans.

A virus serologically related to Sendai virus, named HA-2 virus, was isolated from children with respiratory illness in 1958. This virus is classified in the same type species as Sendai virus, and has been established as a common causative agent of croup in children.

Sendai virus has a remarkable ability to fuse cells. Cell fusion induced by ultraviolet-ray-inactivated Sendai virus is now widely used by investigators in many areas of cell biology, somatic cell genetics, virology, immunology, and oncology. It enables researchers to produce heterokaryons, hybrid cells, or macromolecule-introduced cells whose novel properties provide valuable insight into numerous aspects of cellular organization. *See* SOMATIC CELL GENETICS.

Under electron-microscopic observation, the shape of the Sendai virus particle is roughly spherical, but the diameter varies from 160 to 600 nanometers. The particle has an envelope of lipid bilayer, covered with several hundred spikes 12 nm in length, in which a helical ribonucleoprotein nucleocapsid of 18 nm diameter with a regular periodicity of 5 nm is enclosed.

The genome is a single-stranded ribonucleic acid

(RNA) molecule with a molecular weight of 50,000. The RNA is a negative strand; that is, it is not itself a messenger for viral protein synthesis, but functions as a template for the synthesis of six messenger RNA (mRNA) species in the cytoplasm of infected cells. Six viral structural proteins, as well as one nonstructural protein, are translated from the six mRNAs. *See* ANIMAL VIRUS.

Nakao Ishida

Bibliography. N. Ishida and M. Homma, Sendai virus, *Adv. Virus Res.*, 23:349–383, 1978; T. Shioda et al., Sequence of 3,687 nucleotides from the 3′ end of Sendai virus genome RNA and the predicted amino acid sequences of viral NP, P and C proteins, *Nucl. Acids Res.*, 11:7317–7330, 1983; T. B. Shows and A. Y. Sakaguchi, Gene transfer and gene mapping in mammalian cells in culture, *In Vitro*, 16:55–76, 1980; M. Yamaizumi et al., Rapid transfer of nonhistone chromosomal proteins to the nucleus of living cells, *Nature*, 273:782–784, 1978.

Sensation

The obtaining of information from the environment through the agency of the sense organs. Sensation commonly includes the interpretation of that information as a particular attribute of the world or self.

Sensation is the intermediary between reception and perception, but the boundaries are vague. Reception refers to the absorption of some form of energy (a stimulus) from the external or internal environment through a process called transduction. Perception includes the interpretation of that stimulus to generate a representation of the world. Within the nervous system, the interpretation is generated by multiple feedback loops, so there is no distinct point at which sensation ends and perception begins.

Sometimes there is sensation without perception. Brain damage can lead to a syndrome called blindsight, in which a person who is apparently blind can nevertheless indicate the general direction in which an object is moving. In one case, a patient who could not tell the position or orientation of a slot could nevertheless "post" a card into the slot with nearly normal accuracy. To some extent, normal individuals demonstrate sensory awareness without perception, as when making an abrupt catch. *See* BRAIN; COGNITION; PERCEPTION.

The senses. It is common to speak of the five senses (sight, hearing, taste, smell, and touch), but the senses also include hot and cold, pain, limb position, acceleration, and internal states such as hunger, thirst, and fatigue. Some senses have dedicated organs, such as the eyes, ears, tongue, or nasal mucosa, while others are more diffuse, with receptors throughout the skin or within other organs. Each sense has a specific pathway to the brain and primary areas within the brain where the signals are analyzed.

Sense receptors. Sensation is initiated by the stimulation of receptors. A receptor is a specialized cell that can absorb energy from its environment and signal its presence to other neurons. This is accomplished by a change in electrical potential of the receptor, resulting in a change in the amount of chemical transmitter released to secondary neurons. *See* NEUROBIOLOGY; NEURON.

The process of converting an external stimulus to a neural signal varies across receptor types, but each receptor has a preferred type of stimulus. The receptors in the eye are sensitive to light; they contain a chemical that absorbs photons, initiating a chain reaction that changes the electrical potential by closing the channels that allow ions to cross the cell membrane. Receptors in the ear respond to mechanical vibrations by bending their protruding cilia, opening channels that allow ions to cross the cell membrane. Cells in the tongue and nasal mucosa have receptor sites that bind specific molecules dissolved in the fluids surrounding them, leading to chain reactions that affect their electrical potential. Receptors in the skin and joints respond to mechanical stretch or pressure, or to hot or cold. Many of the channels that allow ions to enter and change the cell potential are members of a superfamily called transient receptor potential (TRP) channels. *See* BIOPOTENTIALS AND IONIC CURRENTS; SYNAPTIC TRANSMISSION.

Since each receptor type is tuned to a specific type of energy, the brain interprets a signal from any given receptor as arising from that form of energy. If a receptor is stimulated by an inappropriate stimulus, it will be misinterpreted. That is why a blow to the eye is seen as a flash of light, called a phosphene. Similarly, punctate pressure on the skin of the back of the hand at some positions causes a sensation of cold; this is where a cold receptor happens to be located.

Sensory receptors that transmit information about the body are found dispersed throughout the skin, muscles, and viscera. Receptors that sense the internal states of the body are generally in the viscera and are called interoceptors. These provide a general feeling of such states as hunger, nausea, distention, chill, or heat. Receptors that sense the position and motion of the body are called proprioceptors. They are generally found in the skin, muscles, and tendons. The receptors that transmit information about the external environment are called exteroceptors. Among these are the receptors for the sense of touch, known as proximoceptors; they provide information about what comes in contact with the body. Other exteroceptors transmit information about the world distant from the body, and are called teleceptors. *See* PROPRIOCEPTION.

Sense organs. Telereceptors are typically concentrated in special organs that channel the appropriate energy to the receptors. Often, the way these organs present the stimuli is a first essential step in the sensory process.

The organ of vision is the eye. The principal function of the cornea and lens of the eye is to focus an image of the visual scene upon the retina, the neural tissue at the back of the eye that includes the visual receptors. The retina signals the pattern of stimulation upon the receptor array. But the eyes also control what part of the scene is analyzed, as they are aimed at visual targets. The contractile iris of the eye also exerts some control over the amount of light presented to the retina, and the lids can shut out the image entirely. *See* EYE; VISION.

The organ of hearing is the ear. The ear also serves a focusing function; the mechanical properties of the inner ear cause different receptor cells to receive different frequencies of sound, as part of the process of analyzing tones. In many animals, the outer ear can be aimed at specific sound sources; humans turn their heads and cup a hand behind the ear to get this effect. *See* EAR; HEARING.

The sense of balance depends on the vestibular organ, an extension of the inner ear. The orientation of the semicircular canals of the labyrinth causes the fluids in them to circulate for particular rotational accelerations, allowing the receptors in each to be specific for those directions of motion. *See* POSTURAL EQUILIBRIUM; VESTIBULAR SYSTEM.

Common properties of the senses. Despite the wide variety of receptor types serving the various senses, and the very different functions each sense serves, there are common features that characterize sensation and sensory processing.

Multiple receptor types. Each sense is served by several different receptor types, often tuned to different aspects of the stimulus. In some cases, this allows coverage of a wider range of stimuli; in many cases, the differences in tuning are essential to be able to discriminate among stimuli of a particular type.

For example, in the vertebrate visual system, there are two receptor types: rods and cones. Rods are more sensitive than cones, being able to be stimulated by a single photon; cones cannot detect such dim light, but extend the range to far greater energies than the rods could handle. With the luxury of more light, cones can afford to be faster and more spatially selective, permitting faster and more accurate vision in brighter lights. In addition, there are three different types of cones in the normal human retina. Each type is most sensitive to a somewhat different range of the visible spectrum. Comparisons among the signals from these cone types is the basis for color vision.

The auditory system has two different types of receptor, the inner and outer hair cells. Like the rods and cones of vision, outer hair cells are more sensitive to weak stimuli than are the inner hair cells.

The olfactory (smell) and gustatory (taste) receptors come in a variety of types, with sensitivity to different chemicals. In the olfactory system, there are apparently thousands of different receptor sites with different distributions on individual receptor cells. *See* OLFACTION; TASTE.

Within the skin, there are nerve endings specific to sudden touch, and others that are sheathed in a fatty capsule so that they respond more slowly to sustained pressure. Pain receptors respond to the chemical products of cell damage, but interact with other receptors to indicate the kind of damage (for example, heat, pressure). Within the muscles, there are two different types of nerve endings on the receptive organ (a special kind of muscle fiber) that are specialized to signal either rapid or sustained lengthening of the muscle; another sensor in the tendons signals potentially damaging strong tension. *See* CUTANEOUS SENSATION; MECHANORECEPTORS; PAIN; PROPRIOCEPTION.

Adaptation. Most sensory systems (certainly all the exteroceptor systems) adapt to the general level of stimulation. By doing so, they concentrate on what is different from the general environment. The process of adaptation is slower than most responses, so one does not adapt to, and thus ignore, the stimuli of interest.

As an example, consider visual light and dark adaptation. If one steps into a darkened theater on a bright afternoon, it is impossible to see the way to an empty seat. After a few minutes, however, fellow theatergoers may be seen with ease. In fact, full dark adaptation in total darkness is not complete for nearly an hour. At the end of that time, sensitivity is such that one can see a single candle at 4 mi (6.4 km). The reverse occurs when one leaves the theater. The sunlight is blinding for a few moments, but then can be tolerated.

The other senses also adapt to the presence or absence of strong stimuli. A strong odor (if not too noxious) fades with time; you do not notice your own perfume or cologne. Gentle pressure is noticed only when first applied; you are not aware of the pressure of your clothing. The shock of jumping into cold water soon dissipates.

Processing to recognize changes. Much of early sensory processing involves comparisons. Localization of a stimulus depends on determining if there is more energy at a particular location than at others. Other aspects may be determined by a similar comparison, which may be thought of as localization in a dimension other than space.

In the visual system, each cell has a receptive field, an area within which light causes a response from that cell. Most receptive fields in early processing (within the retina and beyond) exhibit spatial antagonism; that is, light in a small area centered in the field has one effect, while light in the annular region surrounding that central region has the opposite effect. Thus, the greatest response occurs when the light is concentrated within the center; broader illumination decreases the effectiveness of the light in the center. In this way, the cells are tuned to the contrast of a stimulus—local light compared to the general illumination—and the greatest responses occur where there is the most contrast. The cells are thereby tuned to respond to edges and borders.

Another comparison is made between cones with different spectral properties. Cones responsive to long-wave light (red or orange) are somewhat responsive to all wavelengths, and so would not be able to signal whether a light was dim and red or bright and green (medium-wave light). But by differencing the responses from cones that prefer the green light (middle wavelengths), a positive response can be generated only when the long-wave-sensitive cones have a larger response than the middle-wave-sensitive cones, that is, for a long wavelength (red or orange) light. This processing narrows the response range of subsequent cells, making them more selective. Sensory processing further along in the visual pathway can then use these refined signals to distinguish colors. *See* COLOR VISION.

A similar comparison is made in the cochlear nucleus between cells responding to different locations in the inner ear. Since the different frequencies of sounds are maximally effective at different locations, this sharpens the tuning for specific frequencies.

Psychophysics. Psychophysics is the study of relations existing between energy in the environment and attributes of sensation. Changes of a particular type of sensation are often associated with changes of a particular aspect of the stimulus. For example, changing the spectral characteristics of a light affects the perceived color, while changing the overall energy is associated with a change of brightness. Similarly, changing the frequency of a sound changes the pitch, while changing the amplitude results in a change of loudness. However, these attributes are not independent. Changing the wavelength of a light also affects the brightness; changing the frequency of a sound also affects the loudness. Similarly, changing the energy in a light also affects the color; a more intense red light looks orange, while more intense green lights take a yellowish tinge (this is known as the Bezold-Brücke hue shift). Increasing the amplitude of a sound also affects the pitch, which is why stereo systems boost the bass for lower volumes. In addition, any extremely strong stimulus is interpreted as painful. *See* PSYCHOPHYSICAL METHODS.

Sensitivity and threshold. A primary task of psychophysics is to determine the minimum energy of a stimulus that can be detected, a level called the absolute threshold. Thresholds with different stimulus attributes can indicate the underlying mechanisms of a sensory system. Electrophysiologists who record the activity of neurons in a sensory pathway generally cannot define an absolute threshold, but prefer to find the energy level that produces a particular level of response. The inverse of that energy level is defined as sensitivity. Psychophysicists generally concur that absolute threshold is an arbitrary measure that depends on the subject's criterion; it is now common to determine instead detectability using signal detection theory. *See* SIGNAL DETECTION THEORY.

One application of determining sensitivity is to find different combinations of stimulus attributes that produce the same responses. For example, if a brief flash of light persists longer, it requires less intensity to have the same effect, up to some maximum duration (Bloch's law). That maximum time indicates the temporal summation properties of the system. If a stimulus spot is made larger, less intensity is required for the same effect, up to some maximum diameter. That maximum size indicates the spatial summation properties of the system (Ricco's law). These laws represent the static invariances.

Detection of differences. A revealing determination is the magnitude of the change in stimulus needed for a subject to detect that there is a difference. The requisite change generally is in direct proportion to the magnitude of the stimulus being compared. This is Weber's law, usually stated as $(\Delta I)/I_0 = k$, where ΔI is the change in stimulus energy, I_0 is the energy of the comparison, and k is Weber's fraction, which depends upon the particular sensory system and other attributes of the stimulus.

Scaling. Scaling asks a subject "how big?" is a stimulus. In general, the apparent strength of a stimulus does not grow in direct proportion to the physical energy in the stimulus. Rather, the sensation grows less rapidly for stronger stimuli, as one would expect from Weber's law. What this means is that all changes are relative.

Sensory pathways. Most sensory systems have complex pathways, involving multiple brain areas. Understanding how these systems work means determining the roles of these various brain areas.

Electrophysiologists examine the firing of cells in various brain areas when specific stimuli are presented. This work is generally done in animal models, on the assumption that the basic principles also apply to humans. These studies indicate the specific kinds of responses and cell types that participate in sensory processing.

Brain imaging has now made it possible to see which areas of the human brain are active during specific sensory tasks. There are two major kinds of imaging in use: positron emission tomography (PET) scans and functional magnetic resonance imaging (fMRI) scans. PET scans involve the injection of radioactive compounds and observation of the areas from which radioactivity is emitted to determine the specific chemicals involved in sensory processing pathways. MRI is a technique in which a strong magnetic field causes atoms to align. Subsequent stimulation with a high-frequency pulse temporarily disrupts the alignment, and the radio signals emitted by atomic realignment can then be detected and translated to produce anatomical images. In fMRI, the specific signals sought are from oxygenated hemoglobin. This blood oxygen level detection (BOLD) signal indicates the brain areas most active while the subject performs particular tasks. *See* MEDICAL IMAGING.

Michael W. Levine

Bibliography. R. Axel, The molecular logic of smell, *Sci. Amer.*, pp. 154–159, October 1995; H. B. Barlow and J. D. Mollon, *The Senses*, Cambridge University Press, 1982; R. Blake and R. Sekuler, *Perception*, McGraw-Hill, in press, 2006; D. Dobbs, Fact or phrenology?, *Sci. Amer. Mind*, pp. 24–31, April 2005; M. W. Levine, *Sensation and Perception*, 3d ed., Oxford University Press, 2000; V. S. Ramachandran and E. M. Hubbard, Hearing colors, tasting shapes, *Sci. Amer.*, pp. 53–59, May 2003; R. Shapley and V. H. Perry, *Trends Neurosci.*, pp. 229–235, May 1986; D. V. Smith and R. F. Margolskee, Making sense of taste, *Sci. Amer.*, pp. 32–39, March 2001; H. Wang and C. J. Woolf, *Neuron*, 46:9–12, 2005.

Sense amplifier

An electronic amplifier circuit used to sense and refresh the value of a bit stored in a memory cell of a dynamic random access memory (DRAM) integrated circuit.

In DRAM, bits are represented in memory by the presence or absence of an electric charge stored on tiny capacitors. The charged capacitor has a voltage equal to the charge stored on the capacitor measured in coulombs divided by the cell capacitance measured in farads. A binary one is represented by the presence of charge on the capacitor, while zero charge represents a binary zero. A DRAM chip may store 4 megabits or more on a small piece of silicon. Cell capacitance is directly proportional to cell area, and since the chip area devoted to each bit-cell must decrease with an increase in the memory capacity of a given size of chip, the resulting cell capacitance is tiny, typically on the order of femtofarads (1 femtofarad $= 10^{-15}$ farad). Since the voltage used to charge these capacitors is usually limited to 5 V or less, the charge stored in each memory cell also is quite small, on the order of femtocoulombs. *See* CAPACITANCE; CAPACITOR.

In order to read out the value of a given bit of a word in this type of memory, the bit-cell voltage, or equivalently the magnitude of its charge, needs to be sensed, and the results of this sense operation must be delivered to the rest of the circuit. Both the sense amplifier used to read the bit-cell voltage and the interconnect wiring (the bit line) between the bit-cell and the amplifier have inherent parasitic capacitances which, in total, are much larger than the bit-cell capacitance. If the sense amplifier is simply switched across the bit-cell, the charge on the bit-cell capacitor will redistribute across the parasitic capacitor and the bit-cell capacitor in accordance with Kirchhoff's law, which requires that the two capacitors have the same voltage, where the total charge must come from the original charge stored on the memory-cell capacitor. Thus the voltage seen by the amplifier is equal to the original bit-cell voltage multiplied by the ratio of the bit-cell capacitance to the sum of the two capacitances, and the charge remaining on the memory cell capacitor after sensing would be reduced from the original charge in the same proportion. If the parasitic capacitance is much larger than the cell capacitance, it follows that the voltage available at the sense amplifier input will be small and the voltage and charge remaining on the memory cell will be considerably reduced. More sophisticated approaches need to be taken to sense this tiny voltage and to refresh the memory cell during sense, so that the memory cell capacitor is left fully charged (or fully discharged) when sensing is complete. *See* KIRCHHOFF'S LAWS OF ELECTRIC CIRCUITS.

In practice, all bit-cells in the memory need to be sensed and refreshed periodically, whether they are read or not, since charge leaks from the bit-cell capacitor, even when it is undisturbed by external events. Consequently, a sense amplifier must operate reliably with charge and voltage levels considerably less than the maximum values mentioned above. Since refreshing the contents of the bit-cells takes time from the normal operation of the computer using this memory, maximizing the time between required refresh operations is highly desirable. The amplifier must be designed to function reliably with the smallest possible detectable charge on the bit-cell. Much effort has been expended in developing designs which will meet these stringent criteria.

The sense amplifier action is based on metal-oxide-semiconductor field-effect transistors (MOSFETs). These can be thought of as simple switches that are opened and closed in a predetermined sequence to carry out the read/refresh memory cycle. To assist the sense, a memory reference cell, containing a capacitor of approximately one-half the capacitance of the bit-cell capacitor, is added to the circuit. *See* SEMICONDUCTOR MEMORIES; TRANSISTOR.

Philip V. Lopresti

Bibliography. B. Prince, *Semiconductor Memories: A Handbook of Design, Manufacture, and Application*, 2d ed., 1996.

Sense organ

A structure which is a receptor for external or internal stimulation. A sense organ is often referred to as a receptor organ. External stimuli affect the sensory structures which make up the general cutaneous surface of the body, the exteroceptive area, and the tissues of the body wall or the proprioceptive area. These somatic area receptors are known under the general term of exteroceptors. Internal stimuli which originate in various visceral organs such as the intestinal tract or heart affect the visceral sense organs or interoceptors. A receptor structure is not necessarily an organ; in many unicellular animals it is a specialized structure within the organism. Receptors are named on the basis of the stimulus which affects them, permitting the organism to be sensitive to changes in its environment.

Photoreceptors. Those structures which are sensitive to light and in some instances are also capable of perceiving form, that is, of forming images, are called photoreceptors. Light-sensitive structures include the stigma of phytomonads, photoreceptor cells of some annelids, pigment cup ocelli and retinal cells in certain asteroids, the eyespot in many turbellarians, and the ocelli of arthropods. The compound eye of arthropods, mollusks, and chordates is capable of image formation and is also photosensitive. Dermal light reactions which may be observed in aquatic forms appear to be related to unidentified photosensitive materials which occur in epithelial cells. *See* PHOTORECEPTION.

Phonoreceptors. Structures which are capable of detecting vibratory motion or sound waves in the environment are phonoreceptors. The most common phonoreceptor is the ear, which in the vertebrates has other functions in addition to sound perception. Among the fishes it has been demonstrated that the air bladder and lateral line organs, which in certain species have receptors (neuromasts) in their canals, serve as sound receptors. Sound perception in snakes, which lack a middle ear and are sensitive to airborne sounds, is by bone conduction

which makes them sensitive to ground vibrations. Sensory hairs or hair sensilla and tympanal organs occur in insects. Other organs which may function as sound receptors are the statocysts of certain crustaceans, although they are primarily statoreceptors. *See* PHONORECEPTION.

Statoreceptors. Structures concerned primarily with equilibration, such as the statocysts found throughout the various phyla of invertebrates and the inner ear or membranous labyrinth filled with fluid, are statoreceptors. Halteres are unique structures found in the Diptera which aid in the orientation of these insects. *See* POSTURAL EQUILIBRIUM.

Olfactoreceptors. The sense of smell is dependent upon the presence of olfactory neurons in the olfactory epithelium of the nasal passages among the vertebrates. Olfactory hairs extend from each olfactoreceptor cell, giving it a brushlike appearance. Jacobson's organ in amphibians may play a minor role in olfaction. Little is known about the sensation of smell among the invertebrates. The auricle of planarians has olfactory receptors, and the rhinophores of certain land mollusks are regarded as smell receptors. Odor sensitivity has been observed among many of the invertebrates, but no specific structure has been implicated. The most widely studied invertebrate group is the insects, in which three types of structures have been described: the sensilla placodea, sensilla basiconica, and the sensilla coeloconica. *See* OLFACTION.

Gustatoreceptors. The sense of taste is mediated by the taste buds. In most vertebrates these taste buds occur in the oral cavity, on the tongue, pharynx, and lining of the mouth; however, among certain species of fish, the body surface is supplied with taste buds as are the barbels of the catfish. Again, among the invertebrates, the most studied group has been the insects, and most of the experimental work has been based on behavioral studies. However, because sensilla are known to occur on the antennae, mouthparts, tarsus, and tibia of many species, and on the ovipositor of a few forms, it has been postulated that these structures are involved. Contact stimulation appears to activate the receptors. *See* TASTE.

Cutaneous receptors. The surface skin of vertebrates contains numerous varied receptors associated with sensations of touch, pain, heat, and cold. Tangoreceptors are associated with the phenomena of touch and pressure. Because sensory endings of both tangoreceptors and algesiroreceptors (pain receptors) terminate in the skin, it is often difficult to distinguish between the two. Among the vertebrates, cutaneous receptors are Grandry's corpuscle, Herbst corpuscle, the bulb of Krause, Merkel's corpuscle, Meissner's corpuscle, the Pacinian corpuscle, and the end organ of Ruffini. The end bulb of Krause is a thermoreceptor and is stimulated by cold (frigidoreceptor), whereas the end organs of Ruffini are believed to be influenced by heat (calorereceptor). Sensory and labial pits of certain reptiles are temperature-sensitive, especially during developmental life cycles. Thermal perception among insects has been attributed to the antennae and mouthparts as well as the cerci and tarsi. The ampullae of Lorenzini found in elasmobranchs may have a thermosensory function according to recent studies. *See* CHEMICAL SENSES; PAIN; SENSATION. Charles B. Curtin

Sensitivity (engineering)

A property of a system, or part of a system, that indicates how the system reacts to stimuli. The stimuli can be external (that is, an input signal) or a change in an element in the system. Thus, sensitivity can be interpreted as a measure of the variation in some behavior characteristic of the system that is caused by some change in the original value of one or more of the elements of the system.

There exist many situations where the sensitivity measure indicates the ability of a system to meet certain design specifications. For example, in an electronic system the sensitivity of the output current with respect to the variation of the power-supply voltage can be very critical. In that case, a system with a minimum sensitivity of the output current with respect to the power-supply voltage must be designed. Another example is a high-fidelity audio amplifier whose sensitivity can be interpreted as the capacity of the amplifier to detect the minimum amplifiable signal.

A limiting factor in using the sensitivity of a system to characterize performance at low signal levels is the noise. Noise is a statistical description of a random process inherent in all elements in a physical system. The noise is related to the minimum signal that can be processed in a system as a function of physical variables such as pressure, visual brightness, audible tones, and temperature. *See* ELECTRICAL NOISE.

Sensitivity is commonly used as a figure of merit for characterizing system performance. As a figure of merit, the sensitivity is a numerical indicator of system performance that is useful for predicting system performance in the presence of elemental variations or comparing the relative performance of two or more systems that ideally have the same performance. In the latter case, the performance of the systems relative to some parameter of interest is rank-ordered by the numerical value of the corresponding sensitivity functions. If T is the performance characteristic and X is the element or a specified input level, then mathematically sensitivity is expressed as a normalized derivative of T with respect to X.

A widely used definition of sensitivity in systems is H. W. Bode's logarithmic sensitivity function. It is defined by Eq. (1), where the symbols ∂ and ln stand

$$S_X^T = \frac{\partial T}{\partial X}\frac{X}{T} = \frac{\partial(\ln T)}{\partial(\ln X)} \tag{1}$$

for partial derivative and natural logarithm, respectively.

When sensitivity is used as an indicator of system performance, it follows from Eq. (1) that the percent change in the performance characteristic of interest, $\Delta T/T$, due to a given percent change in the parameter X, $\Delta X/X$, can be expressed in terms of the sensitivity

function by Eq. (2). As an example, if the resonant frequency of a band-pass filter is given by Eq. (3),

$$\frac{\Delta T}{T} \simeq S_X^T \cdot \left[\frac{\Delta X}{X}\right] \tag{2}$$

$$\omega_0 = \frac{1}{\sqrt{R_1 R_2 C_1 C_2}} \tag{3}$$

then it follows from Eq. (1) that the sensitivity of ω_0 with respect to the resistor R_1 is $-{}^1/_2$ and from Eq. (2) that the percent change in ω_0 due to a 5% increase in R_1 is given by Eq. (4).

$$\frac{\Delta\omega_0}{\omega_0} \simeq S_{R1}^{\omega_0} \cdot \left[\frac{\Delta R_1}{R_1}\right] = (-{}^1/_2)(5\%) = -2.5\% \tag{4}$$

On the one hand, a relatively high sensitivity is desirable in a system whose performance characteristic must vary over a large range. An example could be a resonance filter with variable resonance frequency. More common, however, are situations where a minimum sensitivity is required; in that case, any variation in the component values should render small (or negligible) variation in the performance characteristics. An example could be a carburetor for a gasoline engine in which it is desired to maintain good performance (for example, fuel efficiency or horsepower) in spite of rather large manufacturing tolerances in the small mechanical components used in the construction of the carburetor and the wear of these components. Edgar Sánchez-Sinencio

Bibliography. H. W. Bode, *Network Analysis and Feedback Amplifier Design*, 1945; J. L. Bordewijk, Inter-reciprocity applied to electrical networks, *Appl. Sci. Res.*, 6:1–74, 1956; M. Eslami and R. S. Marleau, Theory of sensitivity of network: A tutorial, *IEEE Trans. Educ.*, 32(3):319–324, 1989; L. P. Huelsman, *Active and Passive Analog Filter Design: An Introduction*, 1993; J. D. Schoeffler, The synthesis of minimum sensitivity networks, *IEEE Trans. Circuit Theory*, CT–11(2):271–276, 1964; G. C. Temes and S. K. Mitra, *Modern Filter Theory and Design*, 1973.

Sensory cell regeneration

The replacement of receptor cells with sensory end organs (retina, cochlea, taste buds, olfactory epithelium), most commonly via addition of newly differentiated cells to the systems.

Humans and other organisms use a variety of sensory receptors to detect and interpret information from the surrounding environment. These receptors may be damaged by environmental toxins, injury, or overstimulation, thereby reducing the sensory input received by the organism and the animal's ability to respond appropriately to stimuli. In some systems, such as taste, new sensory receptors are produced on a regular basis. In others, including hearing in humans, the loss of sensory cells is permanent. This article briefly describes the major classes of sensory receptors in vertebrates and some mechanisms of receptor cell regeneration.

Receptor cells. Most sensory stimuli are grouped into three main categories: chemical, visual, and mechanical. Taste receptor cells in the mouth and olfactory receptor neurons in the nose detect chemical stimuli and provoke the senses of taste and smell. Photoreceptor cells in the eye detect visual stimuli. Hair cells in the inner ear detect mechanical (vibrational) stimuli as either sound or balance depending on the frequency of the vibration.

Sensory information from each type of peripheral receptor is transmitted to the brain along pathways that are distinct from one another, except within regions of the brain where sensory information is compared or integrated. This separation of sensory systems is fortunate, since damage to one will not significantly impair the function of another. On the other hand, even minor damage to a given pathway can result in substantial sensory deficits. The most peripheral components, the sensory receptor cells, are often the most vulnerable. *See* BRAIN; NERVOUS SYSTEM (VERTEBRATE).

This vulnerability of sensory receptor cells stems from their position within the organism and from the extreme sensitivity of the receptors to their respective stimuli. Receptor cells such as olfactory receptor neurons and taste receptors must come into direct contact with a chemical stimulus in order to generate a response. Direct exposure makes these cells susceptible to damage from environmental toxins such as harsh chemicals. Photoreceptors in the eye and sensory hair cells in the ear do not make direct contact with the external environment but these cells are easily damaged by intense light or sound. Photoreceptors are exquisitely sensitive, capable of responding to as little as one photon of light. Hair cells are sensitive to mechanical stimuli as minute as 1 nm, making them extremely sensitive to low-intensity sound and balance cues. Since the loss of sensory receptors makes an organism vulnerable to dangers in its environment, numerous mechanisms have evolved to protect sensory systems from debilitating permanent damage. In many sensory systems, one such protective mechanism is the ability to regenerate receptor cells. Researchers struggle to understand these mechanisms with the hope of one day applying their findings to human sensory systems such as the eye and ear. *See* CHEMICAL SENSES; HEARING (HUMAN); OLFACTION; SENSATION; TASTE; VISION.

Innate regenerative capabilities. Regeneration of sensory receptor cells occurs in three manners: repair of damaged receptors, direct conversion (transdifferentiation) of a local nonreceptor cell into a receptor, or proliferation of new receptor cells from a pool of progenitor cells. While all three regenerative processes may occur in sensory systems, proliferation of new cells is arguably the most important and will be discussed briefly here.

In most adult vertebrates, chemical receptor cells possess the greatest regenerative potential. Aging taste receptor cells and olfactory receptor neurons undergo programmed cell death (apoptosis) and are replaced by newly proliferated receptor cells that arise within the sensory epithelia. In the mouse, the average taste receptor cell lives for only nine days, demonstrating that continuous receptor turnover is

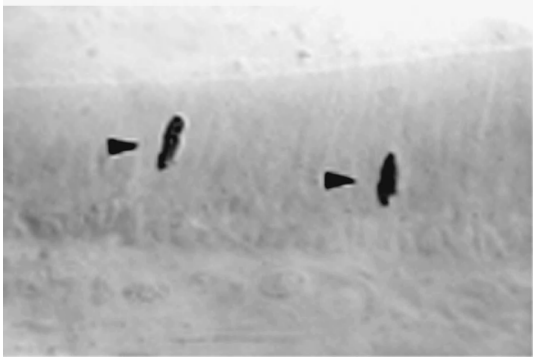

Fig. 1. Light micrograph of the inner ear of the goldfish, *Carassius auratus*. The inner ears of goldfish and other teleost fish continue to add sensory hair cells throughout the life of the animal. The nuclei of two progenitor cells are labeled with BrdU, a substance that marks dividing cells (arrowheads). These cells will go through mitosis and may eventually differentiate as hair cells in the sensory end organ.

normal in this system. There is a close link between apoptosis and proliferation in gustatory and olfactory systems such that a constant number of mature receptor cells exist in the system at any given time. Taste buds and olfactory epithelia can also increase production of receptor cells following end organ damage, leading to the restoration of epithelial structure and function. *See* APOPTOSIS; EPITHELIUM.

In contrast, regeneration of photoreceptors and sensory hair cells is dependent upon the species and age of the animal. In teleost fish and amphibians, the retina and inner ear epithelia grow throughout the life of the organism and new receptor cells are added to the growing sensory structure. Fishes and amphibians also possess increased regenerative capabilities in the eye and ear following sensory overstimulation or drug-induced damage to sensory receptors. This regeneration occurs mostly through an increase in proliferation of new receptor cells similar to that seen in olfactory and gustatory end organs (**Fig. 1** and **2**).

The situation in amniotic vertebrates (reptiles, birds, and mammals) is very different. It was originally thought that these vertebrates lost all regenerative capacity in the eye and ear. However, landmark studies in the past few decades show that birds and perhaps mammals do retain limited regenerative potential. In adult chickens, Müller glia cells in the retina respond to retinal damage by increasing proliferation and producing new retinal cells. Similarly, studies using a radioactive marker to identify proliferating cells in the ear demonstrate that sound or drug-induced damage to the auditory end organ in the bird ear causes dormant progenitor cells to proliferate, producing new sensory hair cells. The vestibular (balance) portions of the mammalian inner ear also possess the ability to produce small numbers of new hair cells. However, innate proliferative regeneration does not occur in the mammalian retina or cochlea. *See* CELL DIFFERENTIATION; EAR (VERTEBRATE); EYE (VERTEBRATE).

Therapeutic regenerative strategies. External intervention may be necessary to achieve regeneration in the mammalian retina and cochlea. Stem cell therapy provides one possible strategy. In vitro experiments with adult mammalian retinal cells show that proliferation is possible given the proper culture conditions. In some cases these newly proliferated cells do differentiate into specific retinal cell types. Transplantation of embryonic stem cells into damaged mammalian retina also shows promise, as some transplanted cells are incorporated into retinal cell layers. Introduction of neural stem cells into the mammalian cochlea has also met with some success, leading to differentiation of new cells that appear structurally similar to hair cells. However, functional recovery following stem cell transplantation has not yet been demonstrated in either sensory system. *See* STEM CELLS.

An alternative strategy is to harness the dormant proliferative potential of the mammalian eye and inner ear. Research on the molecular pathways involved in normal embryonic development of these sensory structures may reveal suites of molecular factors necessary for sensory receptor proliferation and differentiation. These factors could then be exogenously introduced in the damaged retina or cochlea to stimulate regeneration in the mature animal. The Notch signaling pathway, a genetic cellular patterning program, operates in normal retinal and cochlear development and may be involved in regeneration in nonmammalian vertebrates. *See* EMBRYONIC DIFFERENTIATION; EMBRYONIC INDUCTION; PATTERN FORMATION (BIOLOGY).

Notch signaling pathway. The Notch signaling pathway functions as a switching mechanism by

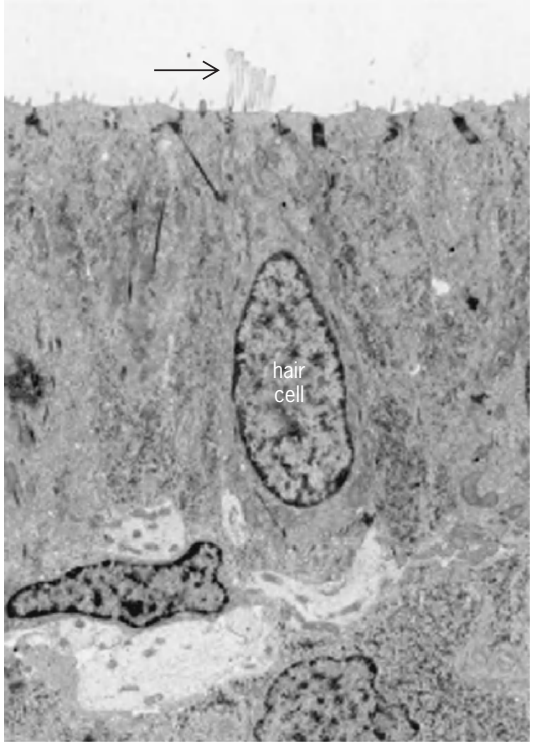

Fig. 2. Transmission electron micrograph of the goldfish inner ear. The inner ears of all vertebrates contain sensory hair cells, from which protrude specialized hairlike structures (arrow) to detect vibrations associated with the senses of hearing and balance.

which progenitor cells may be directed to exit the cell cycle and differentiate as one type of cell versus another, or to continue to divide. The pathway is activated when a progenitor cell bearing the receptor (Notch) comes into physical contact with a progenitor cell that bears one of the receptor ligands: Delta, Serrate, or Jagged1 (**Figs. 3** and **4**). Binding of Notch to the ligand results in an increase in the transcription of genes that encode inhibitory factors (*Enhancer of Split, HES*) [Fig. 4]. These inhibitory factors reduce the activity of cell-specific genes (for example, hair cell or retinal neuron-specific genes) that will enable the progenitor to differentiate as a particular cell type. Consequently, activation of the pathway in a given progenitor cell prevents that cell from differentiating and may result in further rounds of cell division. In this way, the Notch pathway acts as a developmental switching mechanism in which a single pool of progenitor cells may be divided into multiple mature cell types. The Notch signaling pathway plays an important role in retinal cell differentiation; in zebrafish, inhibition of this pathway has been shown to disrupt the spatial organization of retinal cell types during development. In the mammalian cochlea, activation of the Notch pathway apparently results in the presence of two basic cell types in the end organ: hair cells, which differentiate first, and nonsensory supporting cells, which differentiate after subsequent rounds of division. *See* CELL CYCLE.

Interestingly, the Notch pathway also appears to be involved in regeneration and postembryonic hair cell addition in nonmammalian vertebrates. For example, regions of the bird ear that produce large numbers of new hair cells as part of regeneration also express increased levels of Notch and other components of the Notch pathway. Notch expression continues in adult teleost fish, animals which generate postembryonic hair cells. Combined, these results indicate that Notch appears to play an important role in

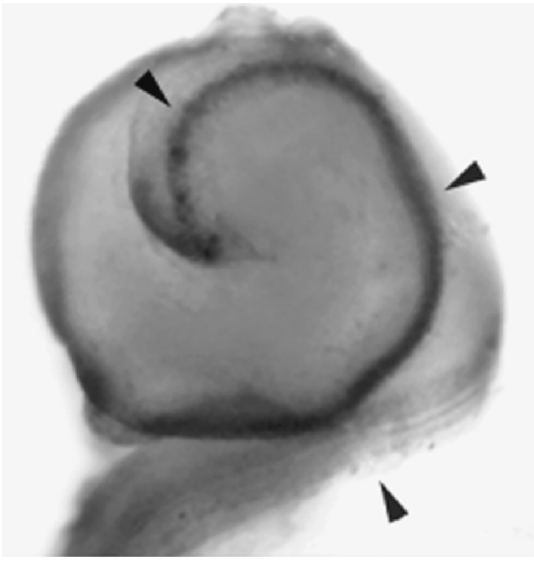

Fig. 3. Notch activity in the developing mouse cochlea. Gene expression in the cochlea (arrowheads) demonstrates the region of the cochlea in which one of the Notch ligands, Jagged1, is expressed.

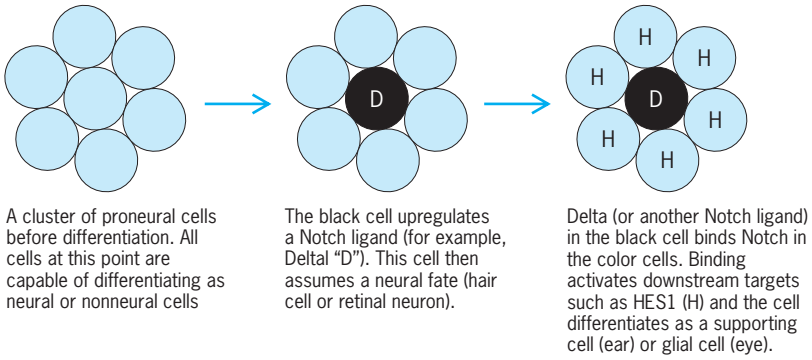

Fig. 4. Simplified model of Notch pathway activation during sensory system development. Differential activation of Notch and its ligands leads to patterning of neural and nonneural cells in the vertebrate inner ear and eye. Details are given under each panel. (*Adapted from J. E. Rooke and T. Xu, Positive and negative signals between interacting cells for astablishing neuronal fate, Bioessays, 20:209–214, 1998, and N. E. Baker, Notch signalling in the nervous system. Pieces still missing from the puzzle, BioEssays, 22:264–273, 2000*)

the postembryonic growth and regeneration of the inner ear in nonmammalian vertebrates. However, since Notch acts primarily as a molecular switch, the expression of Notch alone is not likely to produce new retinal neurons or hair cells. A second set of molecules, encoded by proneural genes, is required as well.

Proneural genes. Proneural genes produce a set of molecular factors that direct developing progenitor cells toward a neural fate. These genes are present in both vertebrates and invertebrates and promote differentiation of a variety of neural cell types depending upon the region of the nervous system in which they are expressed. Recent studies in zebrafish and mice have demonstrated that photoreceptor and hair cell progenitors express specific proneural genes, some of which are related to *atonal*, a gene first identified in the fly.

The mammalian retina expresses multiple proneural genes, including the atonal homologs *Math3* and *Math5*. Mutant mice lacking *Math3* and two other proneural genes, *Mash1* and *NeuroD*, show a severe decrease in the number of photoreceptors but an increase in ganglion cells (another retinal cell type), suggesting that multiple proneural genes act in a complementary fashion in the mouse retina to produce photoreceptor cells. Disruption of *Math5* expression, however, decreases retinal ganglion cells numbers and increases expression of other retinal proneural genes. Taken together, these results indicate that different proneural genes specify different subsets of retinal cell types and that loss of one or more cell types leads to overproduction of others.

In the mammalian inner ear, *Math1*, an atonal homolog, is both necessary and sufficient for hair cell development. In mutant mice that lack functional *Math1*, the inner ear develops normally but completely lacks hair cells. Conversely, introduction of *Math1* into the ears of postnatal guinea pigs leads to production of additional hair cells in regions of the ear that do not normally produce them. Finally, it has been shown that *Math1* is one of the cell-specific genes that is targeted by inhibitory factors of the Notch pathway and that decreased levels of Notch activity result in increased numbers of hair cells in

both mice and zebrafish. Collectively, these results indicate that progenitor cells in the inner ear are directed toward specific cell fates through a combination of proneural gene expression and Notch pathway activation.

In summary, many vertebrate sensory systems regenerate throughout the life of the animal, allowing continued sensitivity to sensory information in the environment. Regenerative capability is lost, however, in the postembryonic mammalian retina and cochlea. Research in nonmammalian model organisms such as birds and fishes that retain regenerative abilities as adults may lead to novel clinical therapies for human retinal degeneration and hearing impairment. *See* DEVELOPMENTAL GENETICS.

Pamela Lanford; Allison Coffin

Bibliography. T. Akagi et al., Requirement of multiple basic helix-loop-helix genes for retinal neuronal subtype specification, *J. Biol. Chem.*, 279:28592–28498, 2004; N. A. Bermingham et al., *Math1*: An essential gene for the generation of inner ear hair cells, *Science*, 284:1837–1842, 1999; R. L. Bernardos et al., Notch-Delta signaling is required for spatial patterning and Müller glia differentiation in the zebrafish retina, *Develop. Biol.*, 278:381–395, 2005; J. T. Corwin and D. A. Cotanche, Regeneration of sensory hair cells after acoustic trauma, *Science*, 240:1772–1774, 1988; C. Haddon et al., Delta-Notch signaling and the patterning of sensory cell differentiation in the zebrafish ear: Evidence from the *mind bomb* mutant, *Development*, 125:4637–4644, 1998; P. Hitchcock et al., Persistent and injury-induced neurogenesis in the vertebrate retina, *Prog. Ret. Eye Res.*, 23:183–194, 2004; R. C. Murray and A. L. Calof, Neuronal regeneration: Lessons from the olfactory system, *Cell Develop. Biol.*, 10:421–431, 1999; M. A. Parker and D. A. Cotanche, The potential use of stem cells for cochlear repair, *Audiol. Neurotol.*, 9:72–80, 2004; P. A. Tsonis and K. Del Rio-Tsonis, Lens and retina regeneration: Transdifferentiation, stem cells and clinical applications, *Exp. Eye Res.*, 78:161–172, 2004.

Sepioidea

An order of the class Cephalopoda (subclass Coleoidea) including the cuttlefishes (*Sepia*), the bobtail squids (*Sepiola*), and the ram's-horn squid (*Spirula*). The group is characterized by an internal shell that is calcareous and broad with closely packed laminate chambers (the cuttlebones of cuttlefishes); calcareous, cylindrical, and coiled with about 40 subspherical chambers (the ram's-horn shell of *Spirula*); chitinous and rod- or bladelike (the gladius of bobtail squids); or absent. The mouth is surrounded by ten appendages (eight arms and two longer tentacles) that bear suckers with chitinous rings. The tentacles are contractile and retractile into pockets at their bases (see **illus.**).

Cuttlefishes are common benthic or epibenthic (living on or just above the bottom, respectively) animals that occur in the warm and temperate waters of the nearshore and continental shelf zones of the Old World, but they are excluded from the Western Hemisphere (North and South America). They prey on shrimps, crabs, small fishes, and other cuttlefishes. The sexes are separate, and during mating, which follows a colorful ritualistic courting behavior, sperm is transferred to the female in cylindrical packets (spermatophores) by a modified arm (the hectocotylus) of the male. The eggs are relatively large, are often stained black with the female's ink, and are laid in grapelike clusters attached to algae, shells, or rocks. Young hatch in a few weeks to 2 months, depending on temperature, and animals live 1–2 years. The cuttlebones are used to control the buoyancy of the cuttlefishes as fluid is pumped into and out of the laminar chambers. Dried cuttlebones are a source of calcium for cage birds and are used for fine jewelry molds, dentifrices, and cosmetics. The cuttlefishes eject an attention-getting blob of brownish-black ink when threatened by predators, then change color, become transparent, and jet-swim away, leaving the predator to attack the false body (pseudomorph) of ink. Artists have used the ink, called sepia, for centuries. Cuttlefishes are important in world fisheries; about 200,000 metric tons are caught each year for human food. Sizes of sepioids range from the smallest known cephalopods of about 0.2–0.7 in. (6–17 mm) body length (*Idiosepius pygmaea*) to cuttlefishes of about 1.6 ft (50 cm) body length and 25 lb (12 kg). *See* CEPHALOPODA; COLEOIDEA.

Clyde F. E. Roper

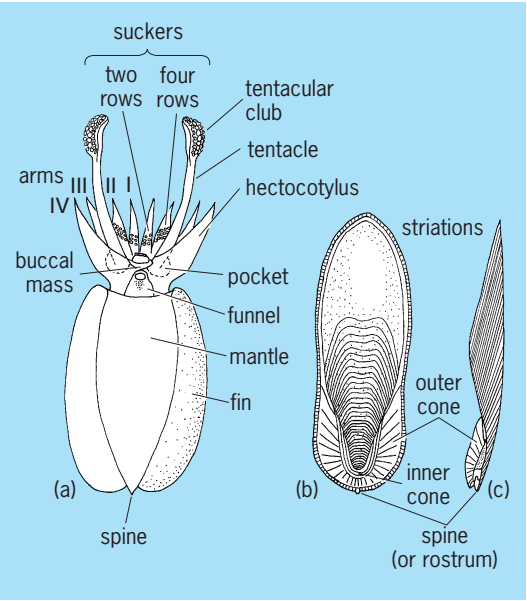

Diagram of some features of (*a*) a cuttlefish and of (*b*) a cuttlebone shown in ventral view and (*c*) in cross section.

Bibliography. P. R. Boyle (ed.), *Cephalopod Life Cycles*, vol. 1, 1983, vol. 2, 1987; F. W. Lane, *Kingdom of the Octopus*, 1960; K. N. Nesis, *Cephalopods of the World*, transl. from Russian, 1987; C. F. E. Roper, M. J. Sweeney, and C. E. Nauen, *Cephalopods of the World*, FAO Fisheries Synopsis 125, vol. 3, 1984; K. M. Wilbur et al. (eds.), *The Mollusca*, vol. 11: *Form and Function*, 1988; K. M. Wilbur and M. R. Clarke (eds.), *The Mollusca*, vol. 12: *Paleontology and Neontology of Cephalopods*, 1988.

Sepiolite

A complex hydrated magnesium silicate mineral named for its resemblance to cuttlefish bone, alternately named meerschaum (sea foam). The ideal composition, $Mg_8(H_2O)_4(OH)_4Si_{12}O_{30}$, is modified by some additional water of hydration, but is otherwise quite representative. The ideal density, 2.26 g/cm^3 (1.31 oz/in.^{3}), is never realized in aggregates. Interlaced disoriented fibers aggregate into a massive stone so porous that it floats on water. These stones are easily carved, take a high polish with wax, and harden when warmed. They have been valued from antiquity in the eastern Mediterranean region for ornaments and make attractive smoking pipes.

Two related types of fibers exist; either of these may occur as the massive interlacings, as paperlike mattings, as bedded clayey deposits, or as more pronounced asbestoslike fibers. They make up a group related in crystal structure to amphibole. Palygorskite and attapulgite are used as names when the amphibole chain relationship is in double units, and sepiolite when it is in triple units. *See* AMPHIBOLE.

Extensive bedded fibrous deposits of either type have been found useful as dispersoids in oil-well drilling fluids for operations that penetrate high-salt formation waters. They are also active as absorptive agents for decolorizing oils, as carriers for insecticides, and as fillers. The bedded palygorskite (attapulgite) variety is produced in quantity in Florida and Georgia in the United States, and in the Ukraine and the Caucasus in Russia; much of the sepiolite variety is produced in Spain, North Africa, and Arabia.

A third fiber, loughlinite, contains up to four sodium ions in place of two magnesium of the sepiolite formula but is changed to asbestoslike fibers of sepiolite by soaking in magnesium chloride. *See* CLAY MINERALS; SILICATE MINERALS. William F. Bradley

Septibranchia

A subclass of bivalve mollusks (class Bivalvia) that are unique in their possession of a muscular septum instead of a filamentous gill. The Septibranchia equate in great part to the superfamily Poromyacea, which includes the septibranch families Cuspidariidae and Poromyidae, and the Verticordiidae. The Verticordiidae have gills that are greatly reduced in size. Although there are a few cuspidariid species in shallow seas, the great majority of septibranchs are found at lower slope and abyssal depths which are deficient in food for filter-feeding bivalves. They live close to the surface in soft sediments. Most species are less than 20 mm maximum length. *See* BIVALVIA; MOLLUSCA.

The term septibranch remains extremely useful because it describes mollusks having a septum and other morphological specializations that relate to the septibranch's unique carnivorous habits. The septum is a muscular, horizontal partition dividing the mantle cavity (see **illus.**). It is derived from the enormous gill found in the mollusks of the subclass Lamellibranchia: The filaments are reduced in size and

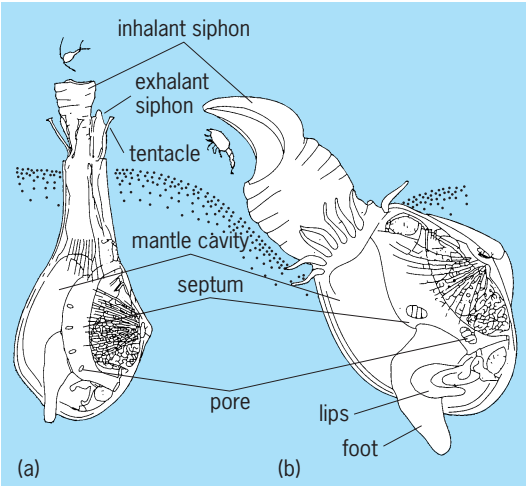

Morphology of two genera of septibranchs. (*a*) *Cuspidaria* (*after J. A. Allen and R. E. Morgan, The functional morphology of Atlantic deep water species of the families Cuspidariidae and Poromyidae (Bivalvia): An analysis of the evolution of the septibranch condition, Trans. Roy. Soc. Lond., 294:413–546, 1981*). (*b*) *Poromya* (*after B. Morton, Prey capture in the carnivorous septibranch Poromya granulata (Bivalvia: Anomalodesmata: Poromacea), Sarsia, 66:241–256, 1981*).

modified. The septum is perforated with cilia-lined pores on either side of the midline. Water circulates into the lower mantle cavity, and then through pores in the septum into the supramantle cavity and exits. This action is brought about by the raising and lowering of the septum and a related cycle of opening and closing the pores, in contrast to other bivalves in which this type of circulation is driven by the cilia of the gills.

The septum, as in all bivalve gills, is used in respiratory and feeding functions. In the cuspidariids the action of rapid raising of the septum draws water and epibenthic and planktonic animals into the mantle cavity. The capture of organisms is precise, the presence of prey being detected first by the siphonal tentacles, and then, via nerve pathways, which trigger a sharp inhalant response from the septum. In contrast, the carnivorous poromyids which have a less muscular septum capture their prey by means of a hood which is an extension of the inhalant siphon. The prey, in this case epibenthic organisms, is also located by sensory tentacles around the inhalant siphon. The closely related verticordiids have adhesive-coated, muscular, branched tentacles surrounding the inhalant aperture, to which planktonic organisms probably adhere. It is thought that the prey is passed into the mantle cavity by the tentacles' bending into the inhalant aperture, the prey being transferred to the mantle cavity possibly with the help of a valve to the inside of the aperture. Food then passes to enlarged manipulative lips of the mouth.

T 'ese prey-capturing movements require rapid redistribution of the blood, and there are various septal and mantle sinuses into which large quantities of blood are shunted in sequence. Other modifications include a muscular stomach for crushing the prey and high proteolytic activity of the gastric juice. The need for rapid muscular and hydraulic movements

can also be related to the nature of the hinge of the shell, the relatively small central ligament, and the development of a ventral part of the ligament (the lithodesma). The lithodesma provides a flexible hinge mechanism that is found only within the anomalodesmatan bivalves. Similarly, the ventral margins of the mantle are fused, assisting in preventing the escape of material from the mantle cavity. *See* FEEDING MECHANISMS (INVERTEBRATE). J. A. Allen

Bibliography. J. A. Allen, On the functional morphology of the family Verticordiidae (Bivalvia) with descriptions of new species from the abyssal Atlantic, *Trans. Roy. Soc. Lond.*, 268:401–536, 1974; J. A. Allen and R. E. Morgan, The functional morphology of Atlantic deep water species of the families Cuspidariidae and Poromyidae (Bivalvia): An analysis of the evolution of the septibranch condition, *Trans. Roy. Soc. Lond.*, 294:413–546, 1981; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Septic tank

A single-story, watertight, on-site treatment system for domestic sewage, consisting of one or more compartments, in which the sanitary flow is detained to permit concurrent sedimentation and sludge digestion. The septic tank is constructed of materials not subject to decay, corrosion, or decomposition, such as precast concrete, reinforced concrete, concrete block, or reinforced resin and fiberglass. The tank must be structurally capable of supporting imposed soil and liquid loads. Septic tanks are used primarily for individual residences, isolated institutions, and commercial complexes such as schools, prisons, malls, fairgrounds, summer theaters, parks, or recreational facilities. Septic tanks have limited use in urban areas where sewers and municipal treatment plants exist. *See* CONCRETE; REINFORCED CONCRETE; STRUCTURAL MATERIALS.

Design. Daily flow is based either on a per-capita flow or on gallons per bedroom. Minimum size of tank (rectangular or circular) for a private residence with one or two bedrooms should be 750 gal (3 m^3), while for a seven-bedroom house it should be 2000 gal (7.5 m^3). The tank should provide a minimum of 24-h detention, and up to 72 h is recommended.

The liquid depth, from the invert of the outlet line to the tank floor, should be at least 4 ft (1.2 m) but not greater than 6 ft (2 m). An air space of at least 12 in. (30 cm) should be provided above the stored

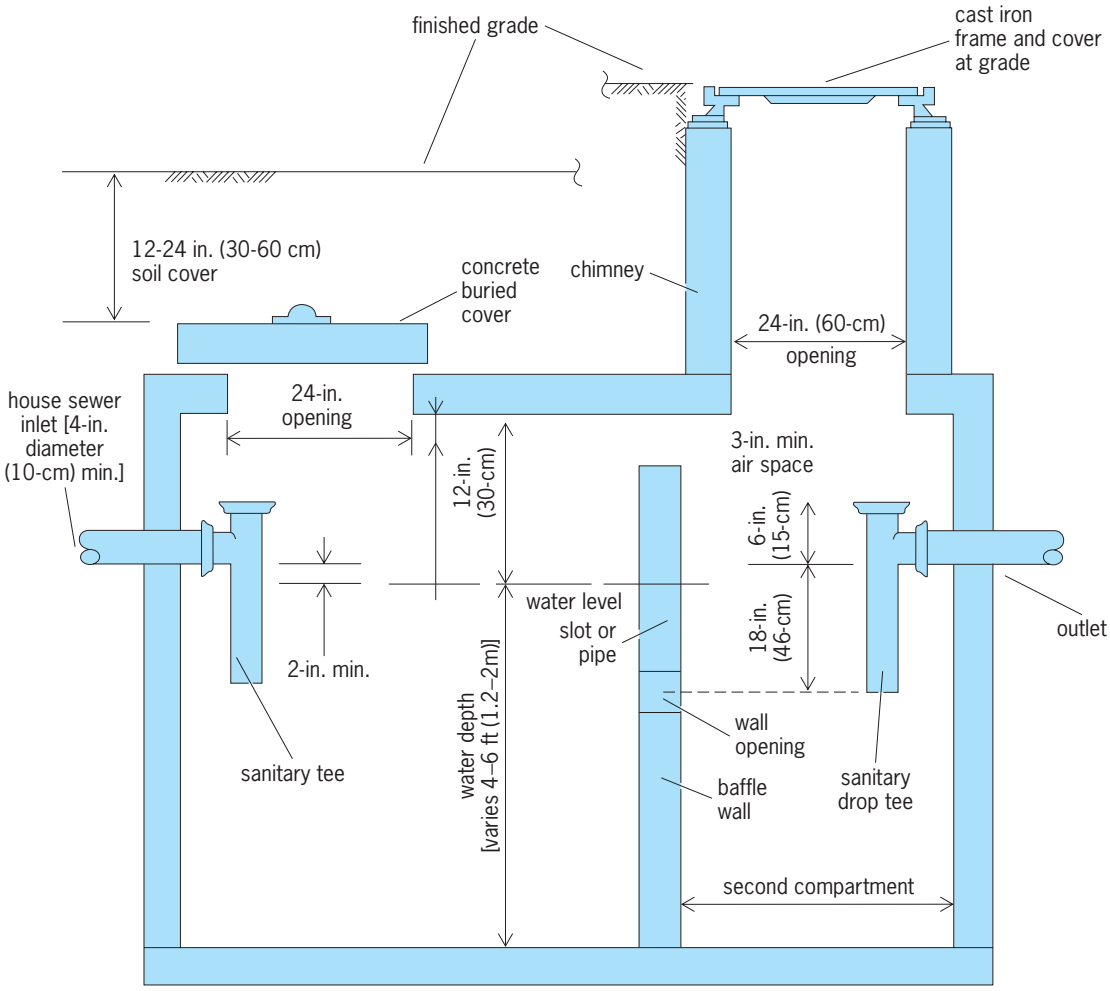

Section of a typical septic tank for on-site residential wastewater treatment.

liquid. The tank length, or diameter, should be not less than 5 ft (1.5 m), and for rectangular tanks should be at least three times the tank width. The inlet and outlet pipes should be at least 4 in. (10 cm) in diameter, with the inlet at least 2 in. (5 cm) above the outlet; and each pipe should be provided with vented tees or baffles to maintain effective retention of scum and sludge and to provide adequate venting. For better operation, the tank should be divided into compartments, with the first compartment containing one-half to two-thirds of the total volume. No compartment should be less than 24 in. (60 cm) in length. The compartment wall must have adequate liquid connections and not block the air-vent space. There should be adequate access to each compartment of the tank for inspection and sludge removal (see **illus.**).

Commercial and institutional septic tanks, with flows in excess of 1500 gal (6 m^3) per day, are designed to store the peak flow for a minimum of 12 h, in addition to sludge storage, although 24-h detention is recommended. If the septic tank is to receive ground garbage, the storage capacity should be increased by at least 30%, although 50% is recommended.

Operation. Septic tanks do not treat sewage; they merely remove some solids and condition the sanitary flow so that it can be safely disposed of to a subsurface facility such as a tile field, leaching pools, or buried sand filter. The septic tank, because of its surface area and detention time, permits heavier solids to settle and lighter particles to rise, thus clarifying the liquid to some extent. The organic solids retained in the tank undergo a process of liquefaction and anaerobic decomposition by bacterial organisms. The clarified septic tank effluent is highly odorous, contains finely divided solids, and may contain enteric pathogenic organisms. The detention and conditioning of the sanitary flow alters the raw sewage so that it may be applied to a properly sized, permeable subsurface soil stratum that will receive the flow with minimal plugging or clogging. The small amounts of gases produced by the anaerobic bacterial action are usually vented and dispersed to the atmosphere without noticeable odor or ill effects.

Maintenance. Septic tanks for private dwellings generally require yearly inspection, and removal of accumulated scum and sludge approximately every 3–5 years. If the inspection indicates that the sludge layer is within 12 in. (30 cm) of the bottom of the outlet tee, or the scum layer is either within 3 in. (7.5 cm) of the top of the outlet tee or more than 10 in. (25 cm) thick, the tank should be pumped out by an experienced scavenger and the contents disposed of in a safe and sanitary manner. *See* RURAL SANITATION; SEWAGE; SEWAGE TREATMENT.

Gerald Palevsky

Bibliography. G. Tchobanoglous, F. L. Burton, and H. D. Stensel (eds.), *Wastewater Engineering: Treatment and Reuse*, 4th ed., 2002; J. A. Salvato, *Environmental Engineering and Sanitation*, 4th ed., 1994.

Sequence stratigraphy

The study of stratigraphic sequences, defined as stratigraphic units bounded by unconformities. L. L. Sloss was the first to formally establish, in 1963, that Phanerozoic cratonic successions in North America could be subdivided into six broad sequences that spanned the entire interior of the continent. However, his work was based on widespread knowledge among North American petroleum geologists, going back to E. Blackwelder in 1909, that the stratigraphy of the continent could be subdivided into a limited number of unconformity-bounded sequences. In later work, Sloss and others demonstrated that these sequences could be correlated partially or entirely with similar successions within other continental interiors, in Russia and South America. With improvements in the acquisition and processing of reflection-seismic data by petroleum exploration companies, in the 1970s came the recognition that unconformity-bounded sequences could be recognized in most sedimentary basins. This was the beginning of an important development, seismic stratigraphy, which also included the use of seismic reflection character to make interpretations about large-scale depositional facies and architecture (**Fig. 1**). These breakthroughs were accompanied by the proposal by Peter Vail that the bounding unconformities are generated by erosion resulting from eustatic (worldwide) changes in sea level and they are therefore correlatable worldwide. These developments were rapidly accepted within the petroleum industry and brought about a revolution in stratigraphy that is still under way. *See* SEISMIC STRATIGRAPHY; UNCONFORMITY.

Correlation problem. If Vail's proposal is correct, it provides the basis for the development of a stratigraphic template for worldwide correlation. However, research during the 1980s and 1990s suggested that the global eustasy model is probably simplistic. There are several other important mechanisms that generate widespread unconformities and sequence boundaries, foremost among them tectonism—the uplift and subsidence of basins resulting from plate-tectonic movements and the effects of thermal convection in the mantle (dynamic topography). Such mechanisms operate regionally, from the scale of individual basins up to that of major continental areas; but tectonic mechanisms do not generate vertical motions of basins that are globally correlatable both in timing and in direction. This being the case, the validity of the global chart of stratigraphic cycles first published by Vail and his coworkers in 1977 (with a revised version appearing in 1987) has been questioned. Current methods for the dating of sequences (primarily biostratigraphy) are insufficiently accurate to provide independent confirmation of global correlations. *See* PLATE TECTONICS.

The controversy regarding global sequence correlation should not detract from the fact that techniques for the documentation, correlation, and analysis of stratigraphic successions have been revolutionized by the methods of sequence stratigraphy. This is because sequence concepts have led

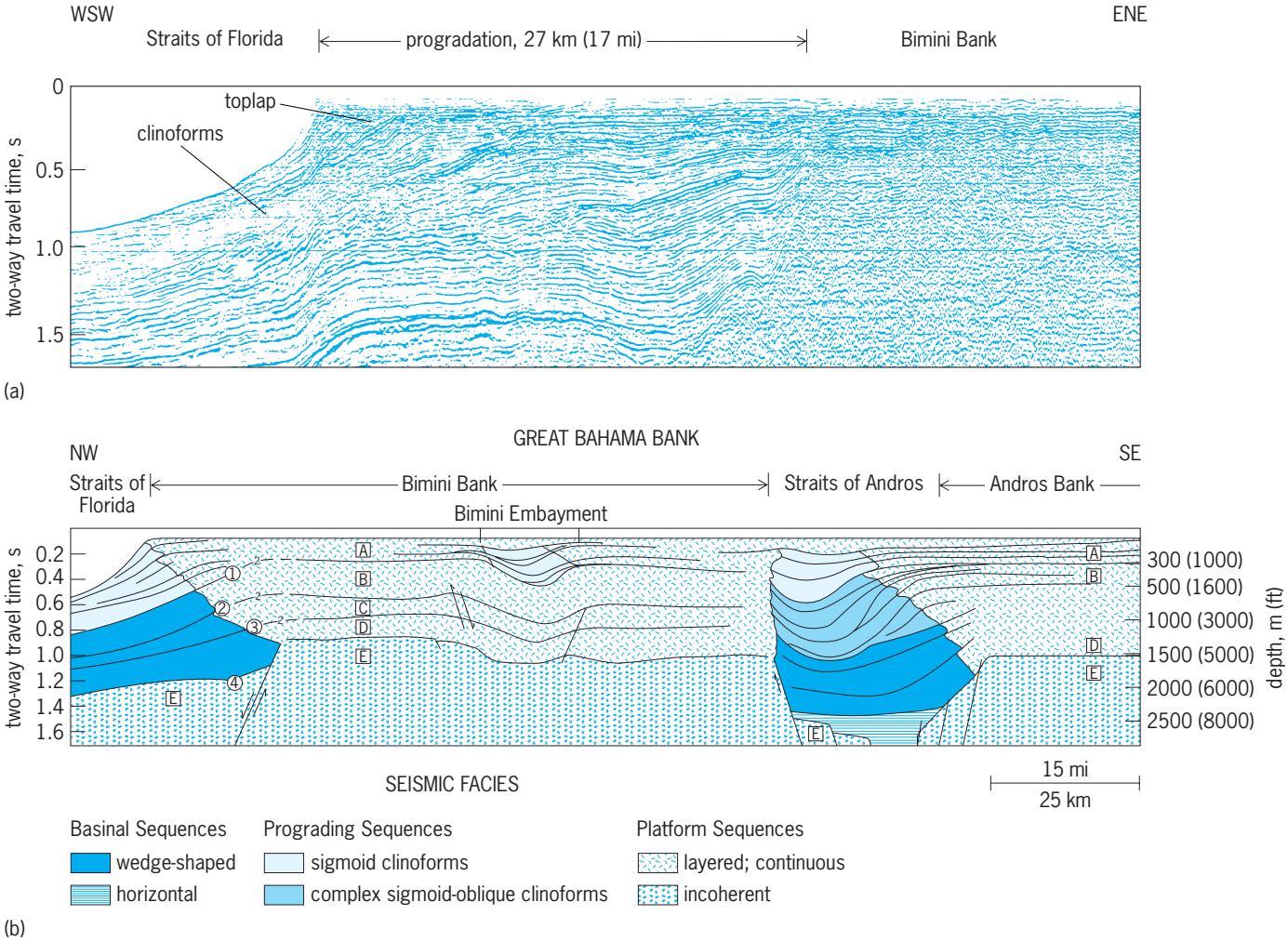

Fig. 1. Seismic reflection characterization. (*a*) Seismic transect across the western margin of Bimini Bank, Bahamas Platform, showing variations in seismic facies. Note the poorly reflective rocks at lower right, representing a sedimentary basement bounded by a normal fault, and well-defined layering in the carbonate-dominated succession above. (*b*) Interpretation of a seismic transect across the Bahamas Bank, showing subdivision of the succession into sequences (indicated by capital letters) bounded by regional unconformities (indicated by numbers). The oldest sequence is mid-Cretaceous. The youngest is still forming. (*After G. P. Eberli and R. N. Ginsburg, Soc. Sedimentary Geol. Spec. Publ., no. 44, 1989*)

to the development of predictive models that have dramatically improved methods for mapping and interpreting stratigraphic units. They are a vast improvement on the older methods of stratigraphic documentation and correlation, which were based on lithostratigraphy—the definition of units (formations, members, groups) based on similarities in facies. Lithostratigraphy embodies no predictive capability.

Sequence model. Underpinning sequence-stratigraphic methods are the following interrelated principles: (1) The volume of sediment accumulating in any part of a sedimentary basin is dependent on the space made available for sediment by changes in sea level or basin-floor elevation. This space is referred to as accommodation. (2) Changes in accommodation tend to be cyclic, and they are accompanied by corresponding changes in sedimentary environment and depositional facies. Thus, a rise in base level typically leads to an increase in accommodation, deepening of the water in the basin, with corresponding changes in facies, and a transgression, with a consequence landward shift in depositional environments and in depositional facies. A fall in base level may lead to exposure and erosion (negative accommodation), with the development of a widespread unconformity. (3) These predictable changes provide the basis for a model of the shape and internal arrangement or architecture of a sequence, including the organization and distribution of sedimentary facies and the internal bedding surfaces that link these facies together. *See* BASIN; DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; FACIES (GEOLOGY).

There are two major kinds of components of sequence models: systems tracts and bounding surfaces (**Fig. 2**). The details of these components vary with depositional and tectonic setting. Thus, the subsidence patterns and regional variations in generation of accommodation are very different in the different tectonic classes of sedimentary basin (rift basins, extensional-margin prisms, foreland basins, and so forth). Sequence architectures also vary

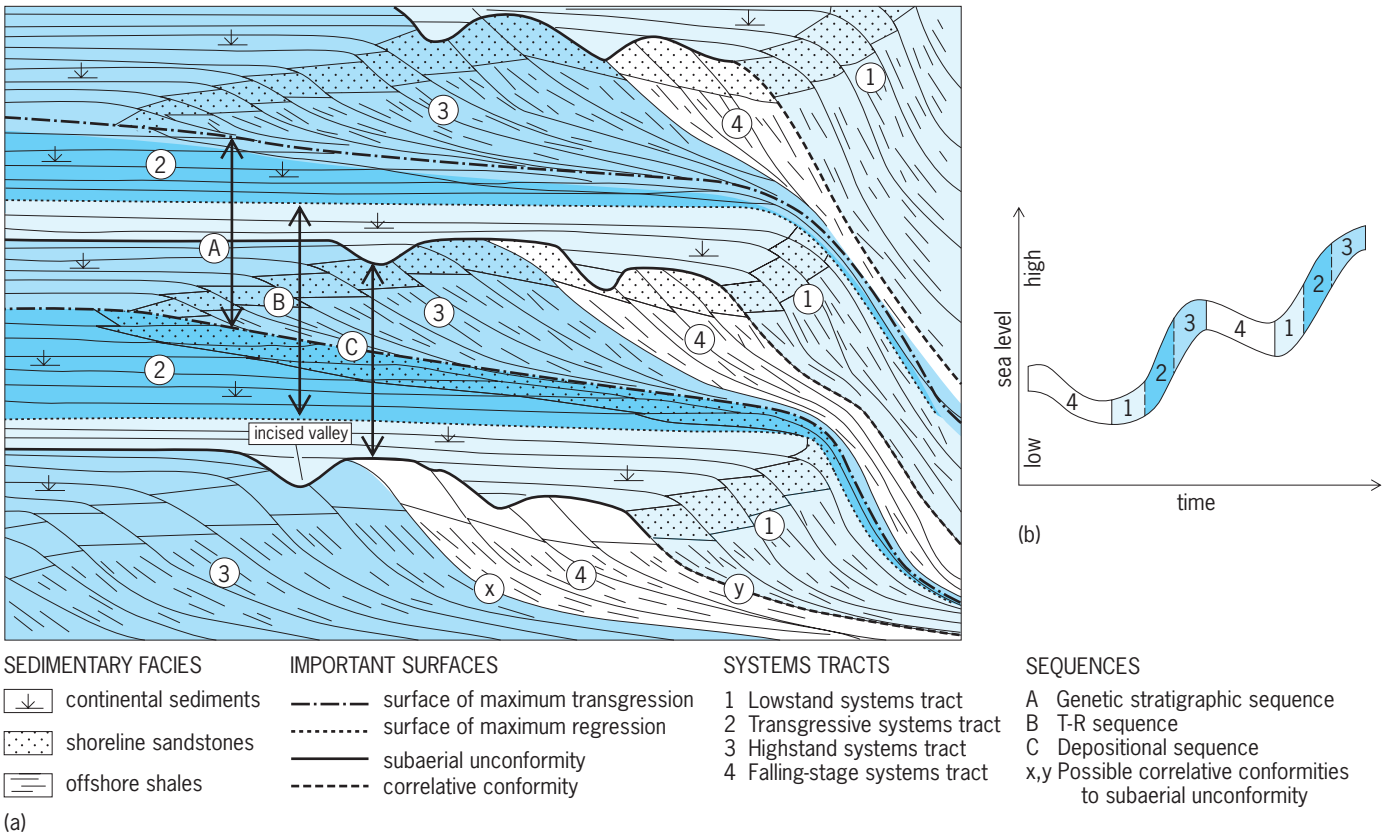

Fig. 2. Clastic-dominated sequence. (*a*) Model for clastic-dominated sequences on an extensional continental margin, showing the key stratigraphic surfaces and the four major systems tracts that develop within each sequence. Double-headed arrows show three ways in which sequences have been defined (see text). (*b*) Diagram showing how the systems tracts relate to changes in relative sea level with time. (*After W. Helland-Hansen and J. G. Gjelberg, Sedimentary Geol., 92:31–52, 1994*)

markedly with differences in the overall sedimentary style. For example, carbonate and clastic depositional systems respond very differently to changes in accommodation. The first sequence model was for clastic sequences on extensional continental margins. *See* CONTINENTAL MARGIN.

Clastic-dominated sequences. As originally defined, sequences are bounded by unconformities (the depositional sequences of Fig. 2). These surfaces (sequence boundaries) are typically well developed within coastal and shelf sediments, where they form as a result of subaerial exposure and erosion during falling sea level. In deeper-water settings, including the continental slope and base of slope, there may be no corresponding sedimentary break; and sequences may be mapped into such settings only if the unconformity can be correlated to the equivalent conformable surface (the correlative conformity). In some instances, the surface of marine transgression, which develops during the initial rise in sea level from a lowstand, forms a distinctive surface that is close in age to the subaerial unconformity and may be used as the sequence boundary. *See* MARINE GEOLOGY; MARINE SEDIMENTS.

The subaerial unconformity is commonly cut by valleys formed by lowstand incision, and may be characterized by erosional relief, making it a difficult surface to correlate and map in the subsurface. For this reason, some workers prefer to map sequences by correlating the deeper-water sediments, typically mudstones, formed at the time of maximum transgression (the top of which is commonly termed the maximum flooding surface). These tend to be widespread. Some workers map sequences using these mudstones as the sequence boundaries, and these sequences have been termed genetic stratigraphic sequences (Fig. 2).

The cycle of rise and fall of sea level may be divided into four segments: lowstand, transgressive, highstand, and falling stage. The deposits that form at each stage are distinctive, and are assigned to systems tracts named for each of these stages.

Falling stage. The sequence-bounding unconformity is formed by subaerial erosion during the falling stage. Incised river valleys form on the continental shelf, and they may preserve remnants of river terraces on the flanks of the valleys. The seaward ends of the valleys may form the heads of submarine canyons, down which sediment is funneled by sediment gravity flows to the base of the continental slope. The submarine-fan and other deep-marine clastic deposits in the deep ocean are generated primarily at this stage of the sea-level cycle because of the abundant sediment supply and connected transport system available to deliver the detritus. *See* SUBMARINE CANYON.

Sea-level lowstand. During the sea-level lowstand, incised valley systems may deepen and widen and

begin to fill with fluvial deposits. Deltas and other coastal depositional systems may build seaward at the margins of the continental shelf. On the coastal plain, mature soils develop on the upland areas between river systems.

Transgression. The transgressive systems tract forms as sea level rises, and accommodation generation is rapid. Typically, coastal regions undergo transgression, and coastal deposits backstep landward (retrogradational stacking of deposits such as shoreline and barrier facies). Incised valleys flood and become estuaries. Sedimentary facies in these valleys change from nonmarine to marine, commonly exhibiting the influence of waves or tides. Later, the transgression may drown the entire coastal plain, blanketing it with mud. Wave erosion may plane off coastal sediments as the shoreline retreats, forming a prominent time-transgressive erosion surface called a ravinement surface. Little sediment may reach deeper-water areas, which may accumulate organic and other detritus in a condensed section. Submarine fans are starved of sediment (except where they are fed by deep canyons at the mouths of some of the largest rivers, such as the Amazon) and are blanketed with a thin mud layer.

Highstand phase. The rate of sea-level rise slows as the cycle nears its end, and enters the highstand phase. The rate of sediment supply typically begins to outpace the rate at which new accommodation is generated, and coastal depositional systems change from retrogradational to aggradational (vertical stacking) and then typically to progradational (seaward-stepping), forming seaward-dipping deltaic and continental-slope depositional surfaces, known as clinoforms. These prograding deposits build back out over the mudstones and condensed sediments of the transgressive systems tract, generating a downlap architecture. During the main highstand phase, sediment supply outpaces the rate of generation of new accommodation, and much sediment bypasses the coastal plain and continental shelf. Slope clinoforms thin upward into these zones of bypass and nondeposition, a condition known as toplap (Fig. 1).

Seaward regression. In the early phases of the next fall in sea level, coastal environments regress seaward, and the shelf may be eroded by wave action, generating a regressive surface of marine erosion. Shoreline deposits may form thin sheets or isolated pods resting with sharp bases on this surface, but all these features may be removed by subaerial erosion as sea level continues to fall.

Carbonate-dominated sequences. Carbonate sedimentation is most active in warm, clear, shallow, shelf seas. During the sea-level cycle, these conditions tend to be met during the highstand phase. Sediment production may be so active, including that of reef development at the platform margin, that it outpaces accommodation generation, leading to deposition on the continental slope. Oversteepened sediment slopes there may be remobilized, triggering sediment gravity flows and transportation into the deep ocean. This process is called highstand shedding. During lowstand, continental platforms may be exposed, leading to erosion and karst development.

Sequences containing evaporites. Evaporites, such as gypsum, anhydrite, or halite, typically form during one of two phases of the sea-level cycle. During lowstand, some marine basins become partially or entirely cut off from marine circulation. Given a hot, arid climate, evaporation may outpace fluvial influx, and evaporite layers may form in environments ranging from shallow to deep. Evaporites may also form in shallow basins on the continental platform during transgression.

Scales and mechanisms of sequence development. The sequence cycle presented here has been described without reference to a time scale. This is because there are several processes of sequence generation that range from a few tens of thousands of years to hundreds of millions of years for the completion of a cycle of rise and fall of sea level. More than one such process may be in progress at any one time within a basin, with the production of a range of sequence styles nested within or overlayering each other.

High-frequency sequence generation is driven by orbital forcing of climate (the so-called Milankovitch effects), of which glacial eustasy is the best-known outcome (although subtle effects on oceanic temperatures and oxygenation levels can also generate significant sequence-stratigraphic effects). The effects of glacioeustasy have dominated continental-margin sedimentation since the freeze-up of Antarctica in the Oligocene. Major glacial episodes also occurred during the late Paleozoic, at the end of the Ordovician, and at several times during the Precambrian. Sequences with durations between about 20,000 and 100,000 years characterize all these periods of Earth history. *See* CLIMATE HISTORY; GEOLOGIC TIME SCALE; GLACIOLOGY; PALEOCLIMATOLOGY.

Regional tectonism—such as the process of thermal subsidence following rifting, and flexural loading in convergent plate settings—develops changes in basement elevation that drive changes in relative sea level. These cycles have durations of a few millions to a few tens of millions of years, and they are confined to individual basins or the flanks of major orogens or plate boundaries. Movement on individual thrust sheets or nappes may generate local cycles of loading and erosional unloading with durations of thousands to tens of thousands of years within limited areas of flanking foreland basins. These tectonically driven sequences commonly dominate the regional stratigraphy; but they are strictly local to regional in extent, and cannot be used in the construction of any global cycle chart.

Low-frequency cycles of eustatic sea-level change are driven by changes in the average rate of sea-floor spreading. Fast rates generate thermally elevated seafloor spreading centers that displace the ocean waters up onto the land (highstand phase); whereas reorganization of spreading centers tends to lead to thermal subsidence, deepening of the ocean basins, and eustatic lowstands. Regional changes in plate

motions occur every few millions to tens of millions of years, with corresponding consequences for global sea level. These changes tend to be additive during the gradual formation and breakup of supercontinents over periods of a few hundreds of millions of years. The highest global sea levels occur during supercontinental break-up, as during the fragmentation of Pangea and the formation of the Atlantic, Indian, and Southern oceans. During the Cretaceous, when this process was at its peak, global sea levels were 100–300 m (300–1000 ft) higher than at the present. Exceptionally low sea levels follow the assembly of a supercontinent, when spreading centers undergo reorganization. *See* PALEOGEOGRAPHY.

Applications. Sequence concepts enable petroleum exploration and development geologists to construct predictive sequence models for stratigraphic units of interest from the limited information typically available from basins undergoing petroleum exploration. These models can guide regional exploration, and can also assist in the construction of production models that reflect the expected partitioning of reservoir-quality facies within individual stratigraphic units.

Documentation of the sequence stratigraphy of sedimentary basins is an essential component of regional geological analysis, and may yield invaluable insights into the tectonic and climatic history of the area. *See* GEOPHYSICAL EXPLORATION; STRATIGRAPHY; WELL LOGGING. Andrew D. Miall

Bibliography. D. Emery and K. Myers (eds.), *Sequence Stratigraphy*, Blackwell Science, Oxford, 1996; A. D. Miall, *The Geology of Stratigraphic Sequences*, Springer-Verlag, Berlin, 1997; P. R. Vail et al., Seismic stratigraphy and global changes of sea-level, in C. E. Payton (ed.), Seismic Stratigraphy: Applications to Hydrocarbon Exploration, *Amer. Ass. Petrol. Geol. Mem.*, 26:49–212, 1977; J. C. Van Wagoner et al., *Siliciclastic Sequence Stratigraphy in Well Logs, Cores, and Outcrops*, American Association of Petroleum Geologists Methods in Exploration Series, no. 7, 1990.

Sequoia

The giant sequoia or big tree (*Sequoia gigantea*) occupies a limited area in California and is said to be the most massive of all living things (**illus.** *a*). The leaves are evergreen, scalelike, and overlapping on the branches (illus. *b*). In height sequoia is a close second to the redwood (300–330 ft or 90–100 m) but the trunk is more massive. Sequoia trees may be 27–30 ft (8–9 m) in diameter 10 ft (3 m) from the ground. The stump of one tree showed 3400 annual rings. The red-brown bark is 1–2 ft (0.3–0.6 m) thick and spongy. Vertical grooves in the trunk give it a fluted appearance. The heartwood is dull purplish-brown and lighter and more brittle than that of the redwood. Because the loss in felling the trees is so great, the logs so difficult to handle, and the wood so brittle, sequoia is little marketed. The wood and bark contain much tannin, which is probably the cause of

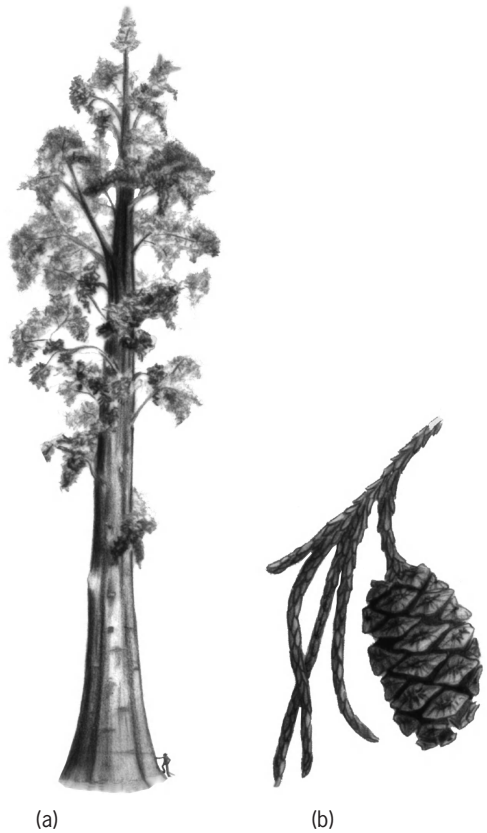

Giant sequoia (*Sequoia gigantea*). (*a*) Tree, showing relative size of man at base. (*b*) Branch and cone.

the great resistance to insect and fungus attack. The most magnificent trees are within the Kings Canyon and Sequoia National Parks, where there are many individual trees 25 ft (7.5 m) in diameter, containing 500,000 board ft (1180 m^3) and up, and probably 3000 years old. The largest specimen weighs about 6000 tons (5400 metric tons). Giant sequoia is sometimes grown in the eastern United States and responds well to cultivation in the British Isles and in central Europe. *See* FOREST AND FORESTRY; PINALES; REDWOOD; TREE. Arthur H. Graves; Kenneth P. Davis

Series

The indicated sum of a succession of numbers or terms. Series are used to obtain approximate values of infinite repeating decimals, to solve transcendental equations, to obtain values of logarithms or trigonometric functions, to evaluate integrals, and to solve boundary value problems.

For a finite series, with only a limited number of terms, the sum is found by addition. For an infinite series, with an unlimited number of terms, a sum or value can be assigned only by some limiting process. When the simplest such process yields a value, the infinite series is convergent. Many tests for convergence enable one to learn whether a sum can be found without actually finding it.

If each term of an infinite series involves a variable x and the series converges for each value of x in a

certain range, the sum will be a function of x. Often the sum is a given function of x, $f(x)$, for which a series having terms of some given form is desired. Thus the Taylor's series expansion

$$f(x) = \sum_{n=0}^{i} f^{(n)}(a)\frac{(x-a)^n}{n!}$$

can be found for a large class of functions, the analytic functions, and represents such functions for sufficiently small values of $|x - a|$. For a much less restricted type of function on the interval $-\pi < x < \pi$, a Fourier series expansion of the form

$$^1/_2 A_0 + \sum_{n=1}^{i}(A_n \cos nx + B_n \sin nx)$$

can be found.

Finite series. Here the problem of interest is to determine the sum of the first n terms,

$$S_n = u_0 + u_1 + u_2 + \cdots + u_{n-1}$$

when u_n is a given function of n. Examples are the arithmetic series, with $u_n = a + nd$ and $S_n = (n/2)[2a + (n-1)d]$, and the geometric series, with $u_n = ar^n$ and $S_n = a(1 - r^n)/(1 - r)$. *See* PROGRESSION (MATHEMATICS).

If v_n is any function such that $v_{n+1} - v_n = u_n$, then $S_n = v_n - v_0$. For methods of solving this difference equation $\Delta v_n = u_n$ *see* INTERPOLATION.

For example, if u_n is any polynomial of the nth degree, v_n will be a polynomial of the $(n+1)$st degree. In particular,

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

$$1^3 + 2^3 + 3^3 + \cdots + n^3 = \frac{n^2(n+1)^2}{4}$$

As an example with u_n a rational function of n, from

$$\frac{-1}{n+2} - \frac{-1}{n+1} = \frac{1}{(n+2)(n+1)}$$

it may be concluded that if

$$u_n = \frac{1}{(n+1)(n+2)}$$

then

$$S_n = 1 - \frac{1}{n+1} = \frac{n}{n+1}$$

Convergence and divergence. An infinite series is the indicated sum of an unlimited number of terms

$$u_0 + u_1 + u_2 + \cdots + u_n + \cdots$$

or more briefly

$$\sum_{n=0}^{\infty} u_n$$

or simply Σu_n, read "sigma of u_n." The sum S_n of the first n terms is known as the nth partial sum. Thus S_n is the finite sum

$$\sum_{k=0}^{n-1} u_k$$

If, as n increases indefinitely or becomes infinite, the partial sum S_n approaches a limit S, then the infinite series Σu_n is convergent. S denotes the sum or value of the series. For example, if $|r| < 1$,

$$S = \Sigma ar^n = \frac{a}{1-r} \quad \text{since} \quad S_n = a\frac{1-r^n}{1-r}$$

and

$$S = \Sigma \frac{1}{(n+1)(n+2)} = 1$$

since

$$S_n = 1 - \frac{1}{n+1}$$

If, as n becomes infinite, the partial sum S_n does not approach a finite limit, then the infinite series Σu_n is divergent. For example, $\Sigma 1$ diverges, since here $S_n = n$ becomes infinite with n. Also $\Sigma(-1)^n$ diverges, since here $S_n = {}^1/_2[1 - (-1)^{n+1}]$ which is alternately one and zero.

It follows from the definition of S_n that $S_{n+1} - S_n = u_n$. For a convergent series, one may take limits in this equality and so deduce that

$$\lim_{n\to\infty} u_n = S - S = 0$$

Thus, if as n becomes infinite, u_n either approaches a nonzero limit or fails to approach a limit, the series must diverge. This checks the earlier conclusion about $\Sigma 1$, with $\lim u_n = 1$, and also that about $\Sigma(-1)^n$, where $u_n = (-)^n$ does not approach a limit. A convergent series remains convergent if a finite number of terms is added, removed, or changed either at the beginning or distributed throughout the series. This changes all S_n after a certain point by the same finite constant. Thus the limit S is changed by this same constant, but there is no change in the fact of approach to a limit.

Positive series. These are series each of whose terms is a positive number or zero. For such series, the partial sum S_n increases as n increases. If for some fixed number A, no sum S_n ever exceeds A, the sums are bounded and admit A as an upper bound. In this case, S_n must approach a limit, and the series is convergent. If every fixed number is exceeded by some S_n, the sums are unbounded. In this case, S_n must become positively infinite and the series is divergent. The tests for convergence of positive series are tests for boundedness, and this is shown by a comparison of S_n with the partial sums of another series or with an integral.

Integral test. Let the function $f(x)$ be positive and always decrease as x increases for x greater than m, some fixed positive integer. Then the series $-f(n)$, with $u_n = f(n)$, converges if the integral

$$\int_m^{\infty} f(x)\,dx$$

converges. The series diverges if the integral diverges. The principal application of this test is that with $f(x) = 1/x^p$, where p is any positive constant. The result is that the series

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

converges if p is greater than 1 and diverges if p is less than or equal to 1. Such divergent series as

$$\sum_{n=1}^{\infty} \frac{1}{n} \quad \text{or} \quad \sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$$

illustrate that series may diverge and still have $\lim u_n = 0$. The slow divergence of

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

may be seen from the fact that over 10^{433} terms must be taken to make the partial sums exceed 1000.

Comparison tests. Let k be any positive constant, Σu_n a positive series to be tested, and Σc_n a positive series known to be convergent. Then if $u_n \leqq kc_n$ for all n greater than some fixed integer m, the series Σu_n converges.

If Σd_n is a positive series known to be divergent, and $u_n \geqq kd_n$ for all n greater than some fixed integer m, then the series Σu_n is divergent.

As corresponding tests involving limits, if

$$\lim_{n\to\infty} \frac{u_n}{c_n} = L$$

where L is a finite limit, then Σu_n converges. But if

$$\lim_{n\to\infty} \frac{u_n}{d_n} = L \quad (L > 0,\ L = +\infty)$$

then the series Σu_n is divergent.

Ratio test. For positive series, the simple ratio test is based on a consideration of

$$\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = t$$

If t is less than unity, the series converges. If t is greater than unity, or if the ratio becomes positively infinite, the series diverges.

If $t = 1$, no conclusions can be drawn directly. But in many such cases

$$\frac{u_{n+1}}{u_n} = 1 - \frac{b}{n} + \frac{c}{n^2} + \cdots$$

may be written. In this case the series converges if $b > 1$, and diverges if $b \leqq 1$.

Cauchy's test, which is related to the ratio test but depends on a single term, is as follows. If for a positive series

$$\sqrt[n]{u_n} \leq r < 1 \quad \text{for all } n \text{ greater than } m$$

the series converges. If $\sqrt{u_n} \geqq 1$ for an infinite number of values of n, the series diverges.

Alternating series. These are series whose terms are alternately positive and negative. For such a series, if each term is numerically less than the preceding term, and

$$\lim_{n\to\infty} u_n = 0$$

the series converges. An example is

$$\Sigma(-1)^n \frac{1}{n+1}$$

For such a series, the difference between S_n, the nth partial sum, and S, the sum of the series, is numerically less than the first unused term, u_n.

Absolute convergence. For any series Σu_n which may have both positive and negative terms, the series of absolute values, $\Sigma|u_n|$, is a positive series whose convergence may be proved by one of the tests for positive series. If $\Sigma|u_n|$ converges, then Σu_n necessarily converges and is said to converge absolutely. The sum of an absolutely convergent series is independent of the order of the terms.

Conditional convergence. A series which converges but which does not converge absolutely is said to be conditionally convergent. For such a series, a change in the order of the terms may change the sum or cause divergence. In fact, by a suitable rearrangement, any sum may be obtained. The series $1 - {}^1/_2 + {}^1/_3 - \cdots$ is conditionally convergent with sum $\log_e 2$. The rearrangement $1 + {}^1/_3 - {}^1/_2 + {}^1/_5 + {}^1/_7 - {}^1/_4 + \cdots$ obtained by taking blocks of two positive terms and then one negative term is conditionally convergent with sum ${}^3/_2 \log_e 2$.

Operations on series. Two convergent series, if added termwise, give a convergent series. If $\Sigma u_n = S$ and $\Sigma v_n = T$, then $\Sigma(u_n + v_n) = S + T$. If both series are absolutely convergent, then the double series $\Sigma\Sigma u_m v_n$, where m and n each run from 1 to infinity, converges to the product ST absolutely, and so does any rearrangement. In particular, this is true for the Cauchy product, with $u_0 v_n + u_1 v_{n-1} + \cdots + u_{u-1} v_1 + u_n v_0$ as its $(n + 1)$st term.

If Σu_n and Σv_n each converge, and if the Cauchy product series converges, its sum is ST. This will necessarily be the case if at least one of the series converges absolutely.

In any convergent series, parentheses may be inserted to form a new convergent series, with the same sum. But the removal of parentheses may convert a convergent series to a divergent one; for example, $\Sigma(1 - 1) = 0$ becomes $\Sigma(-1)^n$, which diverges.

Power series. These are series with $u_n = a_n x^n$. For such a series, it may happen that

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = A$$

If $A = 0$, the series converges for all values of x. If $A \neq 0$, the series converges for all x of the interval $-1/A < x < 1/A$. It will diverge for all x with $|x| > 1/A$. For any power series, the interval of convergence is related in this way to a number A, which, however, in the general case has to be given by the superior limit of

$$\sqrt[n]{|a_n|}$$

Similar remarks apply to the series with $u_n = a_n(x - c)^n$. Here the interval of convergence is $|x - c| < 1/A$, where A may be

$$\lim \left| \frac{a_{n+1}}{a_n} \right| \quad \text{or} \quad \lim \sqrt[n]{|a_n|}$$

if these limits happen to exist. In any case, A is equal to the superior limit

$$\overline{\lim} \sqrt[n]{|a_n|}$$

One of the most important power series is the binominal series:

$$1 + mx + \frac{m(m-1)}{1 \cdot 2} x^2 + \cdots + \frac{m(m-1)(m-2)\cdots(m-n+1)}{n!} x^n + \cdots$$

When m is a positive integer, this is a finite sum of $m + 1$ terms which equals $(1 + x)^m$ by the binomial theorem. When m is not a positive integer, the interval of convergence is $-1 < x < 1$, and for x in this interval, the sum of the series is $(1 + x)^m$. *See* BINOMIAL THEOREM.

The sum function of a power series is continuous inside the interval of convergence. If the series converges at either end of the interval, the function is continuous at this end. Inside the interval of convergence, a power series may be integrated termwise. At any point inside the interval, it may be differentiated termwise. For example, from the series

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots + x^n + \cdots \qquad \text{for } -1 < x < 1$$

differentiation gives

$$\frac{1}{(1-x)^2} = 1 + 2x + \cdots + nx^{n-1} + \cdots$$

Integration gives

$$\log_e(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots - \frac{x^{n+1}}{n+1} - \cdots$$

Since this converges when $x = 1$, $\log_e 2 = 1 - {}^1/_2 + {}^1/_3 - {}^1/_4 + \cdots$.

Taylor series. Let the power series $\Sigma a_n(x - c)^n$ have the sum function $f(x)$. Then $a_0 = f(c)$, $a_n = f^{(n)}(c)/n!$, and the series is the Taylor series of $f(x)$ at $x = c$. Thus, every power series whose interval of convergence has positive length can be put in the form

$$f(x) = f(c) + f'(c)\frac{x-c}{1!} + \cdots + f^{(n)}(c)\frac{(x-c)^n}{n!} + \cdots$$

where $f(x)$ is the sum function.

The Maclaurin series is the special case of Taylor series with $c = 0$

$$f(x) = f(0) + f'(0)\frac{x}{1!} + \cdots + f^{(n)}(0)\frac{x^n}{n!} + \cdots$$

Remainder term. For any function which is finite together with all of its derivatives for $x = c$, the difference between the function and the first $(n + 1)$ terms of its Taylor series is the remainder R_n. For a suitable value x_1 in the interval $c < x_1 < x$, which is

$$R_n = f^{(n+1)}(x_1)\frac{(x-c)^{n+1}}{(n+1)!}$$

Lagrange's form for R_n. This gives $x_1 = c + \theta h$, where $h = x - c$, and θ is a suitable value between 0 and 1. It is true that frequently, if $h \to 0$, $\theta \to 1/(n+2)$. Cauchy's form for R_n is

$$R_n = f^{(n+1)}(c + \theta h)(1-\theta)^n \frac{h^{n+1}}{n!}$$

This may be used to prove that the binomial series converges to $(1 + x)^m$, by showing that $R_n \to 0$ as $n \to \infty$. Lagrange's form for R_n may be used to show that

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

are each Maclaurin series valid for all values of x.

Operations on power series. Two power series $\Sigma a_n(x - c)^n$ and $\Sigma A_n(x - c)^n$ may be added or multiplied by the Cauchy product rule. The resultant series will converge in an interval at least as large as the smaller of the intervals of convergence of the two given series. The first series may be divided by the second, provided that the divisor series is not zero at $x = c$, to give a series for the quotient with some nonzero interval of convergence. A Taylor series with constant term c may be substituted in another series about $x = c$. A power series may be inverted; that is, if $y = \Sigma a_n(x - c)^n$, with $a_1 \neq 0$, there is an expansion $x = \Sigma b_n(y - a_0)^n$ where $b_0 = c$, and the other b_n may be found by substitution.

Complex series. The series Σz_n, in which the general term is the complex number $x_n + iy_n$, with $i^2 = -1$, is said to converge to $A + iB$ if $\Sigma x_n = A$ and $\Sigma y_n = B$. Thus the tests for real series may be made on Σx_n and Σy_n, or the positive series $\Sigma|x_n|$ may be tested. If this converges, the given complex series converges absolutely, and necessarily converges.

Complex power series. The series with $u_n = a_n z^n$, where the $a_n = p_n + iq_n$ are now complex numbers, and the complex variable $z = x + iy$, are of great importance in the theory of analytic function of a complex variable. More generally, power series occur in $(z - c)$, with $u_n = a_n(z - c)^n$, where c is any complex number. There is always a radius of convergence $R = 1/A$, where A may be

$$\lim \left| \frac{a_{n+1}}{a_n} \right| \quad \text{or} \quad \lim \sqrt[n]{|a_n|}$$

if these limits exist. In any case, A is equal to the superior limit

$$\overline{\lim} \sqrt[n]{|a_n|}$$

If A is finite and not zero, so is R, and there is a circle of convergence $|z - c| < R$ within which the series converges. If $A = 0$, the series converges for all values of z. If $A = \infty$, $R = 0$, the series converges for no value except $z = c$. For example, $\Sigma n! z^n$ converges for $z = 0$, but for no other values of z; $\Sigma(z - 3)^n/2^n$, with $A = 1/2$, $R = 2$, converges [to $2/(5 - z)$] for $|z - 3| < 2$; $\Sigma\, z^n/n!$, with $A = 0$, $R = \infty$, converges (to e^z) for all values of z.

Uniform convergence. Let each term of a series be a function of z, $u_n = g_n(z)$. Let S_n be the sum of the first n terms, and S the sum to which the series converges for a particular value of z. Then $R_n = S - S_n$ is the remainder after n terms, and for the particular value of z,

$$\lim_{n\to\infty} R_n$$

must equal zero. If, for a given range of z, it is possible to make $R_n(z)$ arbitrarily small for sufficiently large n without specifying which z in the range is under consideration, the series converges uniformly. The Weierstrass comparison test may be used to test for uniformity. If ΣU_n is a convergent series of positive constants, or a uniformly convergent series of positive terms, and $|u_n| \leqq U_n$ in the range considered, then Σu_n converges uniformly in the range. For a uniformly convergent series, if each $u_n(z)$ is continuous, the sum function $S(z)$ is continuous. A uniformly convergent series may be integrated termwise. If the differentiated series converges uniformly, the series may be differentiated termwise.

Analytic functions. A function $f(z)$ is analytic in a two-dimensional region if, at each point of the region, the derivative $f'(z)$ exists. A uniformly convergent series of analytic functions has a sum function which is analytic and, in particular, may be integrated or differentiated termwise. These results all apply to any power series whose radius of convergence R is not zero, since every power series $\Sigma a_n(z - c)^n$ converges uniformly in any circle $|z - c| < R_1$, where $R_1 < R$. Thus the sum function $f(z) = \Sigma a_n(z - c)^n$ is an analytic function of z for any z such that $|z - c| < R$. For such z, the function $f(z)$ possesses derivatives of every order, each expandable in a series obtained by termwise differentiation. Each such series has the same radius of convergence R. The power series is necessarily the Taylor series of the function $f(z)$, so that $a_n = f^{(n)}(c)(z - c)^n/n!$. If a function is single-valued and analytic in a two-dimensional region of the complex plane, and c is any point inside this region, then there is a Taylor expansion $f(z) = \Sigma f^{(n)}(c)(z - c)^n/n!$. Its radius of convergence R is at least as great as the largest circle with center at c, all of whose interior points are interior points of the given region of analyticity.

Fourier series. Let $f(x)$ be a periodic function of period T, so that $f(x+T) = f(x)$. Then the Fourier series for $f(x)$ is

$$A + \sum_{n=1}^{\infty}(A_n \cos n\omega x + B_n \sin n\omega x)$$

where $\omega = 2\pi/T$ and

$$A = \frac{1}{T}\int_a^{a+T} f(x)\,dx \quad A_n = \frac{2}{T}\int_a^{a+T} f(x)\cos n\omega x\,dx$$

$$B_n = \frac{2}{T}\int_a^{a+T} f(x)\sin n\omega x\,dx$$

That is, A is the average of $f(x)$, A_n is twice the average of $f(x)\cos n\omega x$, and B_n is twice the average of $f(x)\sin n\omega x$ over any interval of length T. Because of the periodicity, it is immaterial which interval, or value of a, is used. If, in each interval of period T, the graph of $f(x)$, which need not be continuous, is made up of arcs which collectively have finite length, then the Fourier series necessarily converges to $f(x)$ at each point where $f(x)$ is continuous. At each point of discontinuity the series converges to $\frac{1}{2}[f(x-) + f(x+)]$, the average of the right- and left-hand limits. If $f(x)$ is continuous at all points, and has a uniformly bounded derivative, the series will converge uniformly for all values of x.

A Fourier series may always be integrated termwise, but it may not be permissible to differentiate such a series termwise unless $f'(x)$ satisfies some condition for development in a Fourier series.

Fourier sine series. On the interval $0 < x < L$, the Fourier sine series

$$f(x) = \sum_{n=1}^{\infty} B_n \sin n\omega x$$

where

$$\omega = \frac{\pi}{L} \qquad B_n = \frac{2}{L}\int_0^L f(x)\sin n\omega x\,dx$$

may be considered as the Fourier series for the function $F(x)$ which is odd, $F(x) = -F(-x)$, of period $2L$, $F(x + 2L) = F(x)$, and which equals $f(x)$ on the interval $0 < x < L$. Thus if $f(x)$ was an odd function for $-L < x < L$, the sine series is valid inside this interval.

Fourier cosine series. On the interval $0 \leqq x \leqq L$, the Fourier cosine series

$$f(x) = A + \sum_{n=1}^{\infty} A_n \cos n\omega x$$

where

$$\omega = \frac{\pi}{L} \qquad A = \frac{1}{L}\int_0^L f(x)\,dx$$

$$A_n = \frac{2}{L}\int_0^L f(x)\cos n\omega x\,dx$$

may be considered as the Fourier series for the function $F(x)$ which is even, $F(x) = F(-x)$, of period $2L$, $F(x+2L) = F(x)$, and which equals $f(x)$ on the interval $0 < x < L$. Thus if $f(x)$ was an even function for $-L \leqq x \leqq L$, the cosine series is valid inside this interval. The sum of the sine or cosine series is always $\frac{1}{2}[F(x-) + F(x+)]$ for each value of x, so that the sum of $F(x)$ at each point of continuity.

Cesaro summability. Let S_n be the sum of the first n terms of a series Σu_n, and form the sequence

$$C_1 = \frac{S_1}{1},\ C_2 = {}^1\!/_2(S_1 + S_2), \ldots, C_n = \frac{1}{n}\sum_{k=1}^{n} S_k$$

If

$$\lim_{n\to\infty} C_n = L$$

exists, the series is said to be summable in the sense of Cesaro, or $C(1)$ to L. A convergent series with sum S is necessarily summable $C(1)$ to S, but a divergent oscillating series may be summable $C(1)$. For example, if $u_n = (-1)^n$, the series is $1 - 1 + 1 - \cdots$ which diverges. But $S_k = 1$ and 0 alternately, and $L = {}^1\!/_2$. Thus the series is summable $C(1)$ to ${}^1\!/_2$, a value also made plausible from consideration of the limit as $x \to 1$ of the identity

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \cdots$$

There are other more elaborate methods of summability by which sums can be assigned to certain divergent series.

Fejér theorem. There are continuous functions whose Fourier series do not converge. But, for any continuous periodic function, the theorem of Fejér asserts that the Fourier series is always summable $C(1)$ to the function for every value of x.

Convergence in the mean. Let $f(x)$ be periodic with T, and

$$T_n = a + \sum_{k=1}^{n=1}(a_k \cos k\omega x + b_k \sin k\omega x)$$

be any trigonometric sum of order $n - 1$. Then T_n may be regarded as an approximation to $f(x)$, and the degree of the approximation may be measured by the average of the square of the error,

$$E_n = \frac{1}{T}\int_a^{a+T} [f(x) - T_n(x)]^2\, dx$$

Then

$$E_n = \frac{1}{T}\int_a^{a+T} [f(x)]^2\, dx - A^2 - {}^1\!/_2 \sum_{k=1}^{n-1}(A_k{}^2 + B_k{}^2)$$
$$+ (A - a)^2 + {}^1\!/_2 \sum_{k=1}^{n-1}(A_k - a_k)^2 + (B_k - b_k)^2$$

This shows that, for all trigonometric sums of given order, S_n, the one formed with Fourier coefficients, $a = A$, $a_k = A_k$, $b_k = B_k$ makes the average error least. Moreover, for any function such that

$$\int_a^{a+T} [f(x)]^2\, dx$$

is finite, as n becomes infinite the limit of E_n, the average squared error for S_n, is zero. Thus

$$\lim_{n\to\infty} \int_a^{a+T} [f(x) - S_n(x)]^2\, dx = 0$$

This is the condition for the trigonometric sums S_n to converge in the mean to $f(x)$.

Integration. If the sequence $S_n(x)$ converges in the mean to $f(x)$, and $g(x)$ is any square-integrable fixed function, it is true that

$$\lim \int S_n(x)g(x)\, dx = \int f(x)g(x)\, dx$$

for the same limits on the integrals in both members. That is, the series with partial sums $S_n(x)$ may be integrated termwise, and the same is true of the series with partial sums $g(x)S_n(x)$. This fact may be used to derive the formulas for the Fourier coefficients in terms of integrals which were given above.

Again, if $T_n(x)$ converges in the mean to $g(x)$, then

$$\lim \int S_n(x)T_n(x)\, dx = \int f(x)g(x)\, dx$$

with the same limits on the integrals in both members. From this Parseval's identity may be deduced,

$$\int_a^{a+T} f(x)g(x)\, dx = AA' + {}^1\!/_2 \sum_{k=1}^{\infty}(A_kA_k{}' + B_kB_k{}')$$

where A', $A_k{}'$, $B_k{}'$ are Fourier coefficients for $g(x)$.

Examples of Fourier series. For $-\pi < x < \pi$, there is a series

$$\sin ax = \frac{2\sin a\pi}{\pi} \cdot \left(\frac{\sin x}{1^2 - a^2} - \frac{2\sin 2x}{2^2 - a^2} + \frac{3\sin 3x}{3^2 - a^2} - \cdots\right)$$

The analogous expansion for $\cos ax$, with $x = 0$ and $a = z/\pi$ leads to

$$\csc z = \frac{1}{z} - \frac{2z}{z^2 - \pi^2} + \frac{2z}{z^2 - 2^2\pi^2} - \frac{2z}{z^2 - 3^2\pi^2} + \cdots$$

The expansion for $\cos ax$ is valid for $x = \pi$. With this and $z = z/\pi$, it gives

$$\cot z = \frac{1}{z} + \frac{2z}{z^2 - \pi^2} + \frac{2z}{z^2 - 2^2\pi^2} + \frac{2z}{z^2 - 3^2\pi^2} + \cdots$$

The last two expansions hold for any complex z for which no denominator is zero.

Infinite product for the sine. A series for $\log_e [(\sin z)/z]$ may be found from the expansion of $\cot z - 1/z$ by integration. This leads to

$$\sin z = z\left(1 - \frac{z^2}{\pi^2}\right)\left(1 - \frac{z^2}{2^2\pi^2}\right)\left(1 - \frac{z^2}{3^2\pi^2}\right)\cdots$$

the infinite product for the sine. Putting $z = \pi/2$ leads to

$$\frac{\pi}{2} = \frac{2\cdot 2\cdot 4\cdot 4\cdot 6\cdot 6\cdots}{1\cdot 3\cdot 3\cdot 5\cdot 5\cdot 7\cdots}$$

which is Wallis's product. Equating the z^3 term on the right with the term $-z^3/3!$ in $\sin z$ gives

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots = \frac{\pi^2}{6}$$

This is a special case of

$$\sum_{n=1}^{\infty} \frac{1}{n^{2k}} = \frac{2^{2k-1}\pi^{2k}}{(-1)^{k-1}(2k)!} B_{2k}$$

where the B_{2k} are rational fractions, the Bernouilli numbers. $B_2 = {}^1/_6$, $B_4 = -{}^1/_{30}$, $B_6 = {}^1/_{42}$. For n odd, B_n is zero unless $n = 1$, and $B_1 = -{}^1/_2$. These B_n are the coefficients in

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!}$$

They occur in such expansions as

$$\tan x = \sum_{k=1}^{\infty} \frac{(2^{2k} - 1)2^{2k}(-1)^{k-1}}{(2k)!} B_{2k} x^{2k-1}$$

$$\cot x = \frac{1}{x} + \sum_{k=1}^{\infty} \frac{(-1)^k 2^{2k}}{(2k)!} B_{2k} x^{2k-1}$$

$$\csc x = \frac{1}{x} + \sum_{k=1}^{\infty} \frac{(2 - 2^{2k}(-1)^k}{(2k)!} B_{2k} x^{2k-1}$$

Stirling's formula for m!. The expansion

$$\log_e(m!) = \log_e \sqrt{2\pi} + (m + {}^1/_2)(\log_e m) - m + \sum_{r=1}^{n} \frac{B_{2r} m^{-2r+1}}{2r(2r-1)} + R_n$$

leads to a divergent series if R_n is omitted and $n \to \infty$. But the series is asymptotic in the sense that R_n is always numerically less than the last term in the sum which involves B_{2n}. Thus for large m, a few terms of the sum give a good approximation. This leads to Stirling's approximation to $m!$, $m! \sim \sqrt{2\pi m}\, m^m e^{-m}$. For $m \to \infty$, the absolute error becomes infinite, but the percentage error is of the order of $e^{1/12m} \sim {}^1/_{12}\, m$ which is small even for moderate m.

Applications of series. Series sometimes appear in disguised form in arithmetic. Thus the approximation of a rational number by an infinite repeating decimal is really a geometric series, ${}^1/_3 = 0.3$ being a series with $u_n = {}^1/_3 > (10^{n+1})$ with sum ${}^1/_3$.

It is possible to use $1 = 0.9$ to find the fraction represented. For example, $0.285714 = 285714/999999 = {}^2/_7$.

Roots are often conveniently found by the binomial series. For instance, for small x, $(1 + x)^{1/n} \sim 1 + (x/n)$, $\log_e(1 + x) \sim x$.

Solution of equations. Sometimes algebraic or transcendental equations are best solved by reverting series. By translations of x and y, $y = y_0 + Y$, $x = x_0 + X$, it is possible to make $Y = 0$ correspond to $X = 0$ near the points of interest. Also, a change of scale, $X = kX'$, makes the first coefficient one. Then with new notation, $y = x + bx^2 + cx^3 + dx^4 + ex^5 + \cdots$, with x and y small. The reverted series is $x = y - by^2 + (2b^2 - c)y^3 - (5b^3 - 5bc + d)y^4 + (14b^4 - 21b^2c + 6bd + 3c^2 - e)y^5 + \cdots$. For example, with $d = 0$, $e = 0$, $y = -q$, this gives a series solution of the cubic equation $cx^3 + bx^2 + x + q = 0$ for small values of q.

Tables. The values of logarithms for tables may be computed by judicious use of the series

$$\log_e \frac{1+x}{1-x} = 2\left(x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots\right)$$

The expansions

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

$$\sin^{-1} x = x + \frac{1}{2}\frac{x^3}{3} + \frac{1 \cdot 3}{2 \cdot 4}\frac{x^5}{5} + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\frac{x^7}{7} + \cdots$$

or

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$$

are useful in computing tables of trigonometric functions. The last series makes it possible to compute π easily by using relations like

$$\frac{\pi}{4} = 2\tan^{-1} {}^1/_3 + \tan^{-1} {}^1/_7$$

or

$$\frac{\pi}{4} = 4\tan^{-1} {}^1/_5 - \tan^{-1} {}^1/_{239}$$

Integrals may often be found from series. For example,

$$\int_0^x e^{-x^2}\, dx = x - \frac{x^3}{3} + \frac{x^5}{5 \cdot 2!} - \frac{x^7}{7 \cdot 3!} + \cdots$$

makes it possible to evaluate the probability integral. For large values of x it is easier to use the divergent but asymptotic expression

$$\int_0^x e^{-x^2}\, dx = \frac{\sqrt{\pi}}{2} - e^{-x^2}\left(\frac{1}{2x} - \frac{1}{2^2x^3} + \frac{3}{2^3x^5} - \frac{3 \cdot 5}{2^4x^7} + \cdots\right)$$

Ordinary differential equations. An ordinary point of a linear differential equation is a value of x at which all the coefficients are analytic, with the first coefficient not zero. For x near such a value, the complete solution may be expressed in terms of Taylor series. For example, $x = 0$ is an ordinary point of

$$(1 - x^2)\frac{d^2y}{dx^2} - 2x\frac{dy}{dx} + n(n+1)y = 0$$

which is Legendre's equation. For n zero or an integer, one solution is

$$P_n(x) = \sum_{k=0}^{\infty} (-1)^k \frac{(2n - 2k)!}{2^n k!(n-k)!(n-2k)!}$$

the Legendre polynomial of degree n.

Fourier-Legendre series. If, for $-1 < x < 1$, the function $f(x)$ satisfies the conditions for expansion in a Fourier series, there is an expansion

$$f(x) = \sum_{n=0}^{\infty} a_n P_n(x)$$

where

$$a_n = \frac{2n+1}{2} \int_{-1}^{1} f(x) P_n(x)\, dx$$

The sum to $(n + 1)$ terms gives the polynomial of the nth degree best approximating $f(x)$ in the sense of least square error.

Regular singular point. Let a differential equation have the form

$$A_2 \frac{d^2y}{dx^2} + A_1 \frac{dy}{dx} + A_0 y = 0$$

where $A_2(x)$ is analytic and not zero at x_0, $(x - x_0)A_1(x)$ and $(x - x_0)^2 A_0(x)$ are each analytic at x_0. Then if x_0 is not an ordinary point, it is a regular singular point. Near such a point, at least one solution may be found in the form $(x - x_0)^s$ times a Taylor series, where s is some real or complex value. For example, for n zero or any positive integer, the Bessel function of order n

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{n+2k}}{2^{n+2k} k!(n+k)!}$$

is a solution of Bessel's differential equation

$$\frac{d^2y}{dx^2} + \frac{1}{x}\frac{dy}{dx} + \left(1 - \frac{n^2}{x^2}\right) y = 0$$

Partial differential equations. Certain boundary-value problems in partial differential equations may be solved by the use of series. Thus, in two dimensions, Laplace's equation in polar coordinates is

$$\frac{\partial^2 U}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 U}{\partial \theta^2} + \frac{1}{r}\frac{\partial U}{\partial r} = 0$$

Let the values on the boundary of a circular region $r = a$ be given in the form $f(\theta)$, and let the Fourier expansion of this function of period 2π be

$$f(\theta) = A + \sum_{n=1}^{\infty} (A_n \cos n\theta + B_n \sin n\theta)$$

Then the solution of Laplace's equation with $U(a,\theta) = f(\theta)$ is

$$U(r, \theta) = A + \sum_{n=1}^{\infty} \left(\frac{r}{a}\right)^n (A_n \cos n\theta + B_n \sin n\theta)$$

Again, the solution of Laplace's equation

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} = 0$$

with $U(0,y) = 0$, $U(L,y) = 0$, $U(x, +\infty) = 0$, $U(x,0) = f(x)$, with Fourier sine expansion

$$f(x) = \sum_{n=1}^{\infty} B_n \sin \frac{n\pi x}{L}$$

is given by

$$U(x, y) = \sum_{n=1}^{\infty} B_n e^{-n\pi y/L} \sin \frac{n\pi x}{L}$$

See FOURIER SERIES AND TRANSFORMS.

Philip Franklin; Salomon Bochner

Bibliography. W. Kaplan, *Advanced Calculus*, 4th ed., 1991; I. S. Sokolnikoff and R. M. Redheffer, *Mathematics of Physics and Modern Engineering*, 2d ed., 1966; S. K. Stein, *Calculus and Analytical Geometry*, 5th ed., 1992.

Series circuit

An electric circuit in which the principal circuit elements have their terminals joined in sequence so that a common current flows through all the elements.

The circuit may consist of any number of passive and active elements, such as resistors, inductors, capacitors, electron tubes, and transistors.

The algebraic sum of the voltage drops across each of the circuit elements of the series circuit must equal the algebraic sum of the applied voltages. This rule is known as Kirchhoff's second law and is of fundamental importance in electric circuit theory. *See* KIRCHHOFF'S LAWS OF ELECTRIC CIRCUITS.

When time-varying voltages and currents are involved, it is necessary to employ differential or integral equations to express the summation of voltages about a series circuit. If the voltages and currents vary sinusoidally with time, functions of a complex variable are used in place of the calculus. *See* ALTERNATING-CURRENT CIRCUIT THEORY; CIRCUIT (ELECTRICITY); DIRECT-CURRENT CIRCUIT THEORY.

Robert Lee Ramey

Serology

The study of antibodies in the serum or other body fluids as well as the study of the antigens present on cells of the blood. It is the oldest branch of immunology. *See* ANTIBODY; ANTIGEN; ANTIGEN-ANTIBODY REACTION; IMMUNOLOGY.

Blood consists of red and white cells, platelets, and plasma (the fluid medium). After blood coagulates, serum (consisting of plasma minus the clotting proteins) separates from the clot. It is generally easier to prepare and more convenient to study in a laboratory than plasma. Even before antibodies were identified, it was noted that the serum of a person who had recovered from a bacterial infection could agglutinate, or clump, those bacteria, so a simple test became available to help identify the organisms with which a person was, or had been, infected. It is easier and more reliable to measure antibodies in serum than the microorganisms themselves, which often are present at low or undetectable numbers. *See* AGGLUTINATION REACTION; SERUM.

Serologic tests. In modern laboratories there are many techniques for identifying antibodies against

microorganisms. Agglutination is still used, as well as complement fixation, enzyme-linked immunosorbent assays (ELISA), virus neutralization assays, and immunofluorescence. Most serologic tests can be set up quickly in the face of an emerging infectious disease epidemic. *See* COMPLEMENT-FIXATION TEST; IMMUNOASSAY; IMMUNOFLUORESCENCE.

Medicine. Large reference laboratories may offer serologic tests for the detection of over 100 pathogens. Some of the familiar tests include the oldest, such as the serologic test for syphilis, which is required to obtain a marriage license in some jurisdictions. Similarly, many governments require serologic proof of immunity to rubella (German measles) before marriage, because rubella infection early in pregnancy can result in serious damage to the fetus. Serology is also of great importance in veterinary medicine to monitor the health of animal herds as well as individual companion animals. *See* RUBELLA; SYPHILIS; VETERINARY MEDICINE.

Severe acute respiratory syndrome. Within weeks of the first cases of severe acute respiratory syndrome (SARS), an ELISA was available. In this test, a synthetic peptide derived from the sequenced SARS-CoV virus genome is attached chemically to the bottom of a well in a plastic dish. The patient's serum is then added, and the well washed so that only antibody, if present, remains bound to the peptide. Next a second antibody (often from an animal) that is able to generically recognize human antibodies is added; it binds only if SARS antibody has previously bound. This second antibody comes tagged with an enzyme. A colorless substrate for this enzyme is then added, which becomes a colored product whose intensity depends on how much antibody is present in the patient's serum. This test is very sensitive, and it is safe for laboratory workers because they do not have to be exposed to a potentially dangerous pathogen. *See* ANIMAL VIRUS; VIRUS.

Human immunodeficiency virus. Human immunodeficiency virus (HIV) infection is confirmed by serology, using an ELISA test. Since the ELISA test for HIV can sometimes yield false-positive results, a more precise test called a Western blot is used to confirm positive results. HIV highlights the limitations of serology, because infected persons usually do not produce detectable antibodies for about 6–9 weeks, during which time they contain the virus and may be contagious. Although expensive, there are tests that can detect the virus itself and are applied when necessary. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS).

Seroconversion. Seroconversion is the process whereby antibodies to a pathogen become detectable over time due to infection or immunization; thus, initial serologic test results are negative, but subsequent tests yield positive results. In most diseases, concentrations of antibody in the serum (called titers) can remain elevated for months or years after recovery from the illness. This means that a single positive serologic result indicates that the patient has encountered the pathogen but does not mean the person is currently ill from it; thus, a second test is ordered, usually about 2 weeks later, to confirm the results. If the infection is quite recent, the second test may show a rise in the titer; if less recent, the titer may be lower; and if the infection was contracted a long time ago, there is likely to be no change in the titer level.

Forensic science. Serology plays a major role in forensic science. Police serologists can type the cells in various body fluids using similar methods to those used in blood banks. Using enough markers, with high probability they can either rule out or implicate suspects. Forensic serologic techniques can even distinguish between the blood of identical twins: Although the cell antigens and markers of the twins are the same, the pathogens they have been exposed to are likely to differ over a lifetime. By comparing antibody titers in the crime scene samples with those in blood taken from the twins, the stains can be definitively ascribed to one or the other. *See* FORENSIC BIOLOGY.

J. John Cohen

Bibliography. R. Burns (ed.), *Immunochemical Protocols* (*Methods in Molecular Biology*), 3d ed., Humana Press, 2005; A. P. Johnstone and M. W. Turner (eds.), *Immunochemistry: A Practical Approach.*, Oxford University Press, 1997; K. Landsteiner, *The Specificity of Serological Reactions*, Dover Publications, 1990 (reprint of original 1945 publication).

Serotonin

A neurotransmitter, also known as 5-hydroxytryptamine (5-HT), that is released throughout most of the brain and spinal cord by the terminals of neurons originating in the raphe nuclei of the brain stem. (These neurons also release serotonin from their dendrites, where it can act on nearby autoreceptors to control the frequency with which the neurons fire.) Serotonin is also released from platelets to promote blood clotting and from enterochromaffin cells in the intestines as a hormone; moreover, in the pineal gland serotonin is an intermediate in the synthesis of the hormone melatonin (5-methoxy *N*-acetyltryptamine).

Serotonin is made from tryptophan, an essential amino acid, and the concentration of tryptophan in the central nervous system (CNS) normally controls the rate at which serotonin is synthesized. Thus, eating carbohydrate foods that elevate brain tryptophan levels (by increasing the amino acid's transport across the blood-brain barrier) can increase serotonin's synthesis and its release into synapses. This unusual ability of serotonin-releasing neurons to couple the release of their neurotransmitter to the composition of each meal allows these neurons to participate in maintaining nutritional homeostasis. However, this ability also underlies the "carbohydrate craving" and weight gain associated with a number of mood disturbances, in which patients learn to use dietary carbohydrates as though they were antidepressant drugs. A large number of protein molecules—including enzymes, receptors,

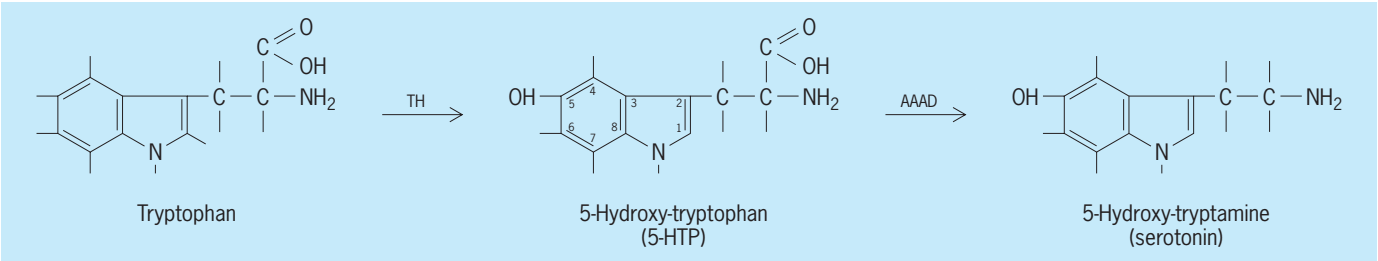

The synthesis of serotonin from tryptophan. AAAD, aromatic L-amino acid decarboxylase; TH, tryptophan hydroxylase.

and binding proteins—are involved in serotonin's actions, allowing an impressive array of relatively specific drugs to be developed for treating serotonin-related disorders.

Synthesis, and transport across blood-brain barrier. The initial and rate-limiting step in serotonin's synthesis is tryptophan's hydroxylation to the amino acid 5-hydroxytryptophan (5-HTP), a process catalyzed by the enzyme tryptophan hydroxylase. 5-HTP is then decarboxylated to serotonin by a ubiquitous and nonspecific enzyme, aromatic-L-amino acid decarboxylase (AAAD) [see **illustration**]. Tryptophan hydroxylase is a low-affinity enzyme, and the tryptophan levels found in the brain are normally too low to saturate it with substrate. Hence, any process that increases or decreases brain tryptophan levels can rapidly affect the synthesis of serotonin, as well as the amounts released into synapses each time the neuron fires. The main such process is eating: a meal that is rich in carbohydrates but contains little or no protein will cause brain tryptophan levels to rise; however, a high-protein meal or snack will fail to elevate brain tryptophan. These relationships are paradoxical, since carbohydrates contain no tryptophan and do not elevate blood tryptophan levels, while all proteins do contain tryptophan and thus do elevate blood tryptophan levels. These food-induced effects occur due to a characteristic of the transport protein, within the endothelial cells lining brain capillaries, which mediates the transport of tryptophan from the blood into the brain.

The transport protein carries numerous additional large neutral amino acids (LNAAs) besides tryptophan—particularly leucine, isoleucine, valine, tyrosine, and phenylalanine—across the blood-brain barrier, and these compounds all compete with tryptophan for access to the transport sites. Eating most simple and complex carbohydrates (but not fructose) causes insulin to be secreted, which then lowers blood levels of the competing LNAAs by allowing them to enter muscle cells. However, insulin has very little effect on blood tryptophan levels, so tryptophan's ability to compete for transport sites is enhanced, causing brain tryptophan levels to rise. In contrast, most proteins decrease tryptophan's ability to compete for transport sites because they contain 20 times (or more) of the other LNAAs than of tryptophan, so brain tryptophan levels fail to rise and may actually fall. (Eating the Thanksgiving turkey does not raise brain tryptophan and serotonin; however, eating the dessert probably does.) Brain tryptophan can also be raised by administering tryptophan itself; however, this product was withdrawn from the market in 1989, when an impure batch caused a new disease, eosinophilia-myalgia syndrome [characterized by high levels of eosinophils (circulating white blood cells) and debilitating muscle pain], that killed many people.

Metabolism. Serotonin is metabolized by the enzyme monoamine oxidase (MAO), which is present in virtually all cells, including the neurons that produce serotonin. Its action generates an intermediate aldehyde which is then either oxidized to an acid (5-hydroxyindole acetic acid; 5HIAA) or reduced to an alcohol. Although none of these metabolites retains serotonin's ability to transmit signals, the process by which "active" serotonin—that is, serotonin present within a synapse—is inactivated involves not its metabolism but its reuptake into the neuron that released it. Inhibition of MAO (for example, by the original antidepressant drugs), like consumption of a carbohydrate-rich meal or snack, elevates serotonin levels within the terminals of raphe neurons and, consequently, the amounts released into synapses when the neurons fire. *See* MONOAMINE OXIDASE; NEUROSECRETION.

Role in the CNS. Once released into synaptic clefts, serotonin (1) can interact with any one of at least 14 postsynaptic receptors (5-HT receptors) to affect second-messenger synthesis or, via the 5-HT_3 receptor, ion channels, (2) can be inactivated by reuptake, a process inhibited by selective serotonin reuptake inhibitors (SSRIs), such as the antidepressants fluoxetine (Prozac®), sertraline (Zoloft®), and Paroxetine (Paxil®), or (3) can act on autoreceptors on the terminals or dendrites of serotonin-releasing neurons to slow its firing frequency or diminish the amount of serotonin subsequently released per firing. (The time needed to desensitize these autoreceptors may explain why antidepressant drugs can require weeks to become clinically effective, even though they block serotonin's reuptake or inhibit MAO activity soon after they are administered.)

Serotonin in the CNS has numerous known physiological and behavioral roles. Drugs that increase its levels within brain synapses, or interact with its receptors, are used to relieve depression, anxiety, excessively aggressive behavior, and bulimia; to enhance satiety and suppress carbohydrate craving; to treat premenstrual syndrome; to suppress

vomiting (for example, in patients receiving cancer chemotherapy); and to treat migraines. Serotonin release within the brain also promotes sleep, and within the spinal cord, it lowers blood pressure and acts as a "gate," limiting pain-generated signals ascending to the brain. *See* BRAIN; NEUROBIOLOGY; PSYCHOPHARMACOLOGY; PSYCHOTOMIMETIC DRUG; SYNAPTIC TRANSMISSION.

Blood serotonin. Serotonin is present within the blood plasma and platelets. Plasma serotonin can be thought of as a hormone, synthesized by specialized cells, the enterochromaffin cells of the intestine, and secreted into the bloodstream. Plasma serotonin is the source of platelet serotonin, since platelets have a serotonin-uptake protein, such as the one in terminals of raphe neurons, but lack the enzyme (tryptophan hydroxylase) required to synthesize serotonin. When platelets coagulate, for example at the site of an injury, they release serotonin, causing local capillaries to constrict and decrease bleeding. *See* BLOOD.

Serotonin levels in plasma are normally quite low. If they are made to rise substantially—for example, in patients suffering from carcinoid tumors, which release very large quantities of serotonin into the bloodstream—this can cause flushing of the skin, diarrhea, wheezing, and sometimes severe damage to the heart valves. Two biochemical pathways exist for removing serotonin from the plasma: (1) uptake into platelets and (2) destruction by MAO in the liver, kidneys, and elsewhere. If both pathways are blocked by the concurrent administration of an SSRI and a MAO inhibitor, plasma serotonin levels will also rise. This mechanism may explain the heart valve lesions and pulmonary hypertension observed in some patients treated for obesity with fenfluramine plus phentermine ("fen-phen"), since fenfluramine suppressed serotonin uptake while phentermine inhibited MAO. Administration of either drug alone was very rarely associated with such side effects, and their incidence may have been no greater than that found in untreated, control populations. It is illegal to give patients both an SSRI and an MAO inhibitor, but unfortunately not all compounds with some MAO-inhibiting activity are labeled as such.

Melatonin. The mammalian pineal gland can also produce serotonin from circulating tryptophan; however, it further transforms the serotonin to the hormone melatonin. Most of the melatonin made during each 24-h period is synthesized during the hours of darkness. Because of its great lipid solubility, it rapidly diffuses across the cellular membrane to enter the cerebrospinal fluid and the bloodstream without being actively secreted. Hence, there are marked circadian rhythms in human plasma melatonin levels, rising from about 8–10 pg/ml during the day to 100–200 pg/ml at night. The mammalian pineal gland is informed about what time it is, and whether or not light is available, by a special neuronal network that runs from the retinas, through the brain and upper spinal cord, to sympathetic nerves, which stimulate melatonin synthesis by releasing norepinephrine onto the pineal cells. The nocturnal secretion of melatonin is an important factor in initiating and sustaining sleep: Many older people whose largely calcified pineal glands release subnormal quantities of melatonin at night exhibit a characteristic type of insomnia, causing them to awaken frequently during the latter half of the night and then experience difficulty falling back asleep. This type of insomnia can usually be treated by administration of oral melatonin, in the tiny doses needed to restore nighttime blood melatonin levels to what they probably were earlier in life. Melatonin is known to act on several types of receptors within the brain; however, the particular receptor which underlies its sleep-promoting effects (as opposed to its effects on other circadian rhythms) awaits discovery. *See* BIOLOGICAL CLOCKS; HORMONE; PINEAL GLAND.

Richard J. Wurtman

Bibliography. J. Cooper, F. Bloom, and R. Roth, *The Biochemical Basis of Neuropharmacology*, 8th ed., 2003; R. J. Wurtman et al., Effects of normal meals rich in carbohydrates or proteins on the plasma tryptophan ratio, *Am. J. Clin. Nutrition*, 77:127–132, 2003; R. J. Wurtman and J. J. Wurtman, Carbohydrates and depression, *Sci. Am.*, 260(1):68–75, 1989; A. F. Schatzberg and C. B. Nemeroff, *Textbook of Psychopharmacology*, 3d ed., American Psychiatric Publishing, Washington, D.C., 2004; I. V. Zhdanova et al., Melatonin treatment for age-related insomnia, *J. Clin. Endocrin. Metab.*, 86(10):4727–4730, 2001.

Serpentine

The name traditionally applied to three hydrated magnesium silicate minerals, antigorite, chrysotile, and lizardite. All have similar chemical compositions [ideally $Mg_3Si_2O_5(OH)_4$] but with three different but closely related layered crystal structures. Serpentine also has been used as a group name for minerals with the same layered structures but with a variety of compositions. The general formula is $M_3T_2O_5(OH)_4$, where M may be magnesium (Mg), ferrous iron (Fe^{2+}), ferric iron (Fe^{3+}), aluminum (Al), nickel (Ni), manganese (Mn), cobalt (Co), chromium (Cr), zinc (Zn), or lithium (Li); and T may be silicon (Si), Al, Fe^{3+}, or boron (B).

Basic crystal structure. The basic layered structure is made up of an octahedral sheet, ideally composed of divalent cations of the M group bonded in octahedral coordination to oxygens (O) and hydroxyls (OH), linked through shared oxygens to a tetrahedral sheet composed of cations of the T group, dominantly silicon atoms, bonded in tetrahedral coordination to oxygens (see **illus.**). Hydrogen bonds between the hydroxyls in the outer surface of the octahedral sheet and the basal oxygens of the overlying tetrahedral sheet bond successive structural layers. All the octahedral sites are occupied in serpentine, unlike structurally related kaolinite, in which only two-thirds of the sites are occupied. In the magnesium-serpentine minerals, this produces an octahedral sheet with slightly larger lateral dimensions than the tetrahedral sheet, and this

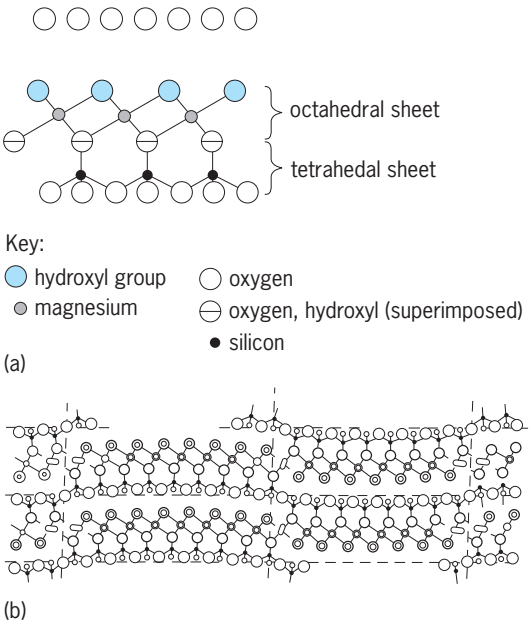

Serpentine structures. (*a*) Crystal structure drawing of lizardite showing the octahedral and tetrahedral sheets and the planar nature of the layered structure. (*b*) Crystal structure drawing of antigorite showing the alternating wave structure (*after G. Kunze, Die gewellte struktur des antigorites, I. Zeit., Krystallographica, 108:82–107, 1956*).

misfit produces the different serpentine structures. *See* HYDROGEN BOND; HYDROXYL.

Magnesium-serpentine minerals. Lizardite has a planar structure (illus. *a*), with the misfit accommodated by slight adjustments of the atomic positions within the layers. This structure is similar to that of other layer silicates and clay minerals, and it has a well-developed cleavage parallel to the layers. Curvature of the sheets also relieves the misfit along the direction of curvature, and there are two curved serpentine structures. Chrysotile has a cylindrical structure in which the layers are either concentrically or spirally rolled to produce fiber commonly ranging from 15 to 30 nanometers in diameter, and micrometers to centimeters in length. These fibers have great strength and flexibility and are the most abundant and commonly used form of asbestos. Successive layers in both lizardite and chrysotile can be offset by series of systematic shifts, so that each mineral has a series of slightly different structures, or polytypes. Antigorite has a modulated wave structure, with wavelengths generally varying between 3 and 5 nm (illus. *b*). At each point of inversion, magnesium and hydroxyl groups are omitted relative to the silicate sheet, so that the composition of antigorite is slightly and systematically different from lizardite and chrysotile. *See* ASBESTOS.

Serpentine minerals often form in an environment that does not reach equilibrium, and transmission electron microscope studies often show one structure type passing into another laterally along the layers. Thus, the fine details of structural relationships can be varied and complex. Polygonal serpentine, or polygonal fibers, is a common and interesting form of lizardite. It is composed of either 15 or 30 polygonal sectors of lizardite, often around a chrysotile fiber core. A fourth mechanism for overcoming the misfit has been recognized in the structure of the mineral carlosturanite. Although this is not a layer silicate in the strict sense, it has been identified in several serpentinites, and it may be a commonly occurring mineral. The structure is believed to be composed of a octahedral sheet linked to a tetrahedral sheet modified into strips of tetrahedra, with the breaks between strips occupied by hydroxyls and water (H_2O). This produces a composition with higher magnesium and H_2O and lower silicon contents than the ideal composition.

The magnesium-serpentine minerals usually contain some iron and aluminum substituting for magnesium, and aluminum or ferric iron for silicon, but they may also contain lesser amounts of nickel, manganese, chromium, and cobalt, reflecting the composition of the parent rock. Chrysotile has the smallest range of substitution and antigorite is intermediate, both limited by the curvature of the structure. Lizardite has the greatest range, and it is possible that a solid solution exists between lizardite and amesite $[Mg_2Al[SiAl]O_5(OH)_4]$. The coupled substitution of aluminum for silicon and magnesium reduces the misfit between the octahedral and tetrahedral sheets, making the former smaller than the latter and producing a fifth variation of the layered structure. *See* SERPENTINITE.

Other serpentine minerals. The serpentine group of minerals include, in addition to the magnesium-serpentines, nepouite and pecoraite $[Ni_3Si_2O_5(OH)_4]$, the nickel analogs of lizardite and chrysotile respectively; and brindleyite $[Ni_2Al[SiAl]O_5(OH)_4]$, berthierine $[Fe_2Al[SiAl]O_5(OH)_4]$, kellyite $[Mn_2Al[SiAl]O_5(OH)_4]$, fraipontite $[Zn_2Al[SiAl]O_5(OH)_4]$, and cronstedtite $[Fe_2^{2+}Fe^{3+}[SiFe^{3+}]O_5(OH)_4]$, which can be considered analogs of amesite. In greenalite $[Fe_3Si_2O_5(OH)_4]$ and caryopilite $[Mn_3Si_2O_5(OH)_4]$ the substitutions of iron and manganese produce misfits that are too large to be accommodated in the regular lizardite-type structure, and they have modulated structures. Manandonite $[Li_2Al_4[Si_2AlB]O_{10}(OH)_8]$ is a rare mineral that has come to be recognized as a serpentine mineral. *See* CHEMICAL BONDING; CLAY MINERALS; CRYSTAL STRUCTURE; KAOLINITE; SILICATE MINERALS.

Frederick J. Wicks

Bibliography. S. W. Bailey (ed.), *Hydrous Phyllosilicates (Exclusive of Micas)*, 1988; G. D. Guthrie and B. T. Mossman (eds.), *Health Effects of Mineral Dusts*, 1993.

Serpentinite

A common rock composed of serpentine minerals, serpentinite is usually formed through the hydration of ultramafic rocks, dunites, and peridotites in a process known as serpentinization. The result is the formation of hydrated magnesium-rich minerals, such as antigorite, chrysotile, or lizardite, commonly with magnetite or, less frequently, brucite. *See* DUNITE; PERIDOTITE.

Origin. Serpentinites can be distinguished by, and are named for, the dominant serpentine mineral in

the rock, that is, antigorite-serpentinite, chrysotile-serpentinite, and lizardite-serpentinite. Lizardite-serpentinites are the most abundant. They have been formed in retrograde terrains and are characterized by the pseudomorphic replacement of the original olivine, pyroxenes, amphiboles, and talc by lizardite with or without magnetite or brucite. Each host mineral produces an easily recognized characteristic pseudomorph in which relicts of the original minerals may remain if serpentinization is not complete. Thus, the textures of the preserpentinization rock are preserved by the pseudomorphs; so that a great deal of information on the original grain shapes, fabrics, deformation features, and distribution of the original minerals is retained.

Antigorite-serpentinites can form directly from minerals such as olivine, pyroxene, and so forth in retrograde terrains similar to lizardite, but at a high temperature. However, antigorite-serpentinites most commonly have been formed in prograde terrains, often by the recrystallization of lizardite-serpentinites. The recrystallization of the rock obliterates the previous minerals and textures, replacing them with easily recognized interpenetrating blades of antigorite.

Chrysotile-serpentinites usually occur only in chrysotile asbestos deposits. The complex sequence of serpentinization, recrystallization, folding, and faulting that produce chrysotile asbestos deposits produce serpentinites that are mineralogically and texturally complex. The typical sequence begins with serpentinization of an anhydrous peridotite or dunite to produce a lizardite-serpentinite. Subsequent recrystallization produces a second generation of lizardite, chrysotile, or antigorite, with all three often occurring in a single rock. Obviously, such rocks did not reach equilibrium conditions. *See* ASBESTOS.

Accessory minerals. Magnetite is the main accessory mineral, a by-product of the serpentinization of olivine, and it may outline the original grains, making them easier to recognize. However, it may not be formed or may be resorbed by a later recrystallization event. The serpentinization of dunite often produces brucite rather than magnetite. In lizardite-serpentinites the brucite is intimately associated with the lizardite on a submicrometer scale, but in antigorite and chrysotile-serpentinites it occurs as discrete easily recognized grains. In the reducing conditions that produced some serpentinites, nickel/iron alloys are a minor accessory mineral. *See* BRUCITE; MAGNETITE; OLIVINE.

Calcium in pyroxenes is released during serpentinization, but because it cannot be accommodated in the serpentine crystal structures, calcium-rich fluids are produced. Dikes and blocks of other rocks such as granite, diabase, and gabbro, or rocks at the contacts with serpentinites may be altered by these fluids to form calcium-rich rocks called rodingites. The strong chemical gradient between the magnesium-rich serpentinites and the silica-rich country rock can produce other reactions that result in nephrite jade, jadeite, chlorite/talc, or, if carbon dioxide is present, talc/carbonate reaction zones along the contacts. *See* METAMORPHISM.

The occurrence of serpentinites is widespread, particularly in greenstone belts, mountain chains, and mid-ocean ridges, where they have formed through the serpentinization of ultramafic rocks. Serpentinites also are formed infrequently during the intrusion of diabase dikes into dolomitic marbles. This type of serpentinite is unusual, because it contains no magnetite; and the serpentine minerals contain little or no iron, a reflection of the very low iron content of the original marble. *See* AMPHIBOLE; MID-OCEANIC RIDGE; PYROXENE; SERPENTINE.

Frederick J. Wicks

Bibliography. S. W. Bailey (ed.), *Hydrous Phyllosilicates (Exclusive of Micas)*, 1988; G. D. Guthrie and B. T. Mossman (eds.), *Health Effects of Mineral Dusts*, 1993; D. S. O'Hanley, *Serpentinites: Records of Petrologic and Tectonic History*, 1966.

Serum

The liquid portion that remains when blood is allowed to clot spontaneously and is then centrifuged to remove the blood cells and clotting elements. It is approximately of the same volume (55%) as plasma and differs from it only by the absence of fibrinogen which, after conversion into fibrin, is incorporated in the separated clot. *See* FIBRINOGEN.

Blood serum contains 6–8% solids, including macromolecules such as albumin, antibodies and other globulins, and enzymes; peptide and lipid-based hormones; and cytokines; as well as certain nutritive organic materials in small amounts, such as amino acids, glucose, and fats. Somewhat less than 1% of the serum consists of inorganic substances such as sodium, chloride, sulfates, phosphates, calcium, potassium, bicarbonate, magnesium, iodine, zinc, and iron. Small amounts of respiratory gases such as oxygen and carbon dioxide are dissolved in the serum, as is the gas nitric oxide, which serves as a chemical messenger and vasodilator. Small amounts of waste material such as urea, uric acid, xanthine, creatinine, creatine, bile pigments, and ammonia are also present in the serum. These substances, along with other small molecules which are not bound to blood proteins, are filtered out as blood flows through the kidney. *See* BLOOD; CLINICAL PATHOLOGY; KIDNEY.

Certain types of sera, both human and animals, are used in clinical medicine. Immune serum and hyperimmune serum either are developed by naturally occurring disease or are deliberately prepared by repeated injection of antigens to increase antibody titer for either diagnostic tests or the treatment of active disease. These sera are referred to as antisera, since they have a specific antagonistic action against specific antigens. This ability of antibodies to detect and/or immobilize specific antigenic molecules provides a very important tool for biomedical research. There are also homologous and heterologous sera, the former pertaining to serum prepared and used in animals of the same species, and the latter referring to sera prepared and used on animals of different

species. *See* ANTIBODY; ANTIGEN; BIOLOGICALS; IMMUNITY.

By custom, the clear portion of any liquid material of animal origin separated from its solid or cellular elements is also referred to as sera, for example, the fluid in a blister and the fluids present in certain natural cavities of the body, such as the peritoneal, pleural, or pericardial sacs. These fluids are more properly referred to, however, as effusions. *See* SEROLOGY.

Rueben Straus; Bruce A. Stanley

Servomechanism

A system for the automatic control of motion by means of feedback. The term servomechanism, or servo for short, is sometimes used interchangeably with feedback control system (servosystem). In a narrower sense, servomechanism refers to the feedback control of a single variable (feedback loop or servo loop). In the strictest sense, followed in this article, the term servomechanism is restricted to a feedback loop in which the controlled quantity or output is mechanical position or one of its derivatives (velocity and acceleration). *See* CONTROL SYSTEMS.

The purpose of a servomechanism is to provide one or more of the following objectives: (1) accurate control of motion without the need for human attendants (automatic control); (2) maintenance of accuracy with mechanical load variations, changes in the environment, power supply fluctuations, and aging and deterioration of components (regulation and self-calibration); (3) control of a high-power load from a low-power command signal (power amplification); (4) control of an output from a remotely located input, without the use of mechanical linkages (remote control, shaft repeater).

Many of these objectives could be attained by a nonfeedback system if accurately calibrated components were available. The salient advantage of the closed-loop feedback method of control is that it permits economical design and production of accurate systems using primarily inaccurate components. A typical servomechanism requires only a few of its components to have high accuracy.

Servomechanisms were first used in speed governing of engines, automatic steering of ships, automatic control of guns, and electromechanical analog computers. Today, servomechanisms are employed in almost every industrial field. Among the applications are cutting tools for discrete parts manufacturing, rollers in sheet and web processes (such as steel rolling, paper calendering, and high-speed printing), elevators, automobile and aircraft engines, robots, remote manipulators and teleoperators, telescopes, antennas, space vehicles, mechanical knee and arm prostheses, and tape, disk, and film drives. *See* COMPUTER STORAGE TECHNOLOGY; FLIGHT CONTROLS; GOVERNOR; MAGNETIC RECORDING; MEDICAL CONTROL SYSTEMS; REMOTE MANIPULATORS; ROBOTICS; VIDEO DISK.

Servo Loops

Figure 1 shows the basic elements of a servomechanism and their interconnections; in this type of block diagram the connection between elements is such that only a unidirectional cause-and-effect action takes place in the direction shown by the arrows. The arrows form a closed path or loop; hence this is a single-loop servomechanism or, simply, a servo loop. More complex servomechanisms may have two or more loops (multiloop servo), and a complete control system may contain many servomechanisms. *See* BLOCK DIAGRAM.

In Fig. 1 the objective of controlling the output variable c is accomplished if the feedback signal b can be made identically equal to the reference input signal r. If b and r are not equal, a deviation or actuating signal e is produced. This signal e, suitably amplified, is impressed on the servomotor which corrects the output variable c so as to bring b into coincidence with r. If the amplifier gain is large, the deviation e required for a given corrective action is quite small. Adequate control action can be achieved provided the following conditions are met: (1) The servomotor is large enough to drive the load and perform the required corrections. (2) The measuring devices and their connection to the shafts have adequate accuracy. (3) The summing device is accurate. (4) The amplifier gain is sufficiently large. *See* GAIN.

The accuracy of the forward controlling elements and of the load does not enter into these considerations. In other words, with a sufficiently

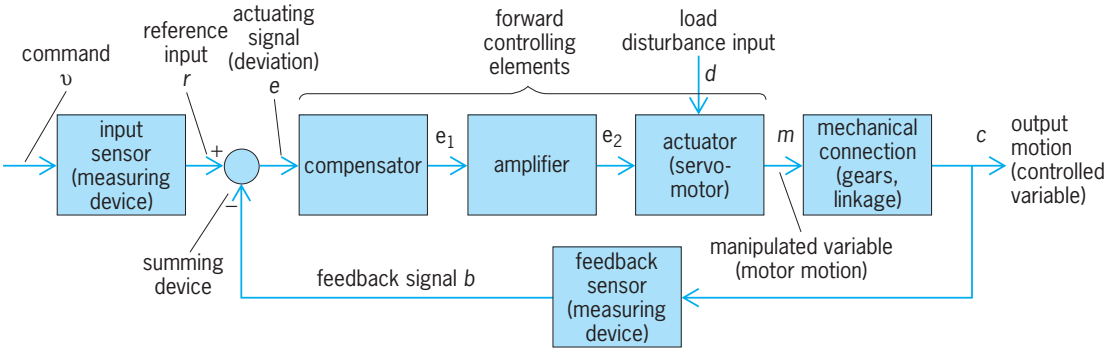

Fig. 1. Servo loop elements and their interconnections. Cause-and-effect action takes place in the directions of arrows. (*After American National Standards Institute, Terminology for Automatic Control, ANSI C85.1*)

large motor and sufficiently high amplifier gain, the accuracy of the system is dependent primarily on the accuracy of the measuring and summing devices and is independent of the characteristics of the high-power portion of the servo, amplifier, motor, and load. Variations in amplifier, motor, and load characteristics, either from unit to unit in mass production or in one unit over a period of time, have little effect on the servo accuracy. The devices which determine the overall accuracy, the measuring and summing elements, are low-power components, and they can be fabricated to a high order of accuracy. A high-grade control requires accurate means of measurement and instrumentation. The quality of the servomechanism can never exceed that of the instruments which measure input and output motion.

The use of high amplifier gain in feedback systems tends to cause overcorrection, and the resulting system may be poorly damped or unstable. The purpose of the compensator in Fig. 1 is to minimize or eliminate these instabilities and to provide accurate, rapid, and smooth response. The functions of the summing device and compensator are usually provided by a single component called the controller. *See* CONTROL SYSTEM STABILITY; DAMPING.

Servomechanisms may be classified in several ways. The following classifications show some of the many available variations:

1. Positional or velocity servo, depending on whether the variable to be controlled is the output position or the output velocity.
2. Electric, hydraulic, pneumatic, or mixed, depending on the form of the energy used in actuation, measurement, and control.
3. Analog, pulse, digital, or mixed, depending on the way the measurement, actuation, signal conditioning, and control functions are implemented.
4. Proportional or on-off (relay) servo, depending on the number of levels in the motor drive signal.
5. Translational or rotational, depending on the type of actuator and on the output motion.
6. Direct-drive or geared, depending on the manner of connection between actuator and load.
7. Direct-current (dc), alternating-current (ac), pulse-modulated, or mixed, depending on the electrical signal format.
8. Absolute value or incremental motion, depending on whether the input command and feedback sensor possess a datum or zero level.
9. Power or instrument servo, the latter referring to a small servo whose load consists primarily of displays and sensors.

A few of these combinations are described more fully below.

Positional servomechanism. The feedback sensor generates a signal representing a measured position of the output shaft. **Figure 2** is an example of a positional servo employing a dc motor as the actuator and a resistance potentiometer as a feedback sensor. The input is a shaft position which is measured by an identical potentiometer. A change in the input command causes a change in the reference voltage r, which produces a deviation e. This in turn causes a voltage to be applied to the motor armature, which then repositions the load so as to reduce the deviation to zero. This servo is called a shaft position repeater, because the output and input shaft positions are equal at equilibrium. This is an example of an analog dc servo with proportional actuation.

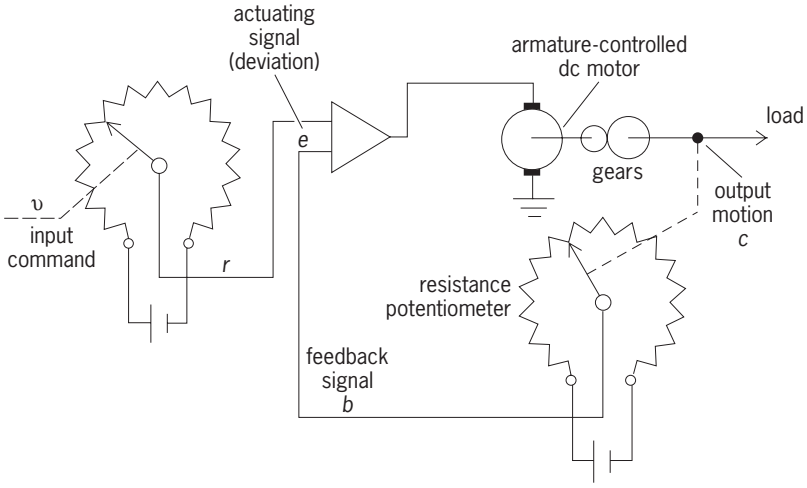

Fig. 2. Positional servomechanism, employing a dc motor as the actuator and a resistance potentiometer as a feedback sensor.

Velocity servomechanism. This is also called a rate servo. The feedback-measuring device generates a signal representing a measured value of the velocity of the output shaft. **Figure 3** is an example of a rate servo employing a dc motor as actuator and a dc generator as a feedback sensor. The speed is controlled by the electric reference voltage r. A transient disturbance such as a change in load causes a transient change in speed, after which the speed recovers to its original value. In the example shown, the servo does not have any positional reference; that is, the shaft speed returns to its original value after a transient disturbance, but the shaft may not be in the proper position. More sophisticated digital speed control systems are able to provide both velocity and position control (phase lock).

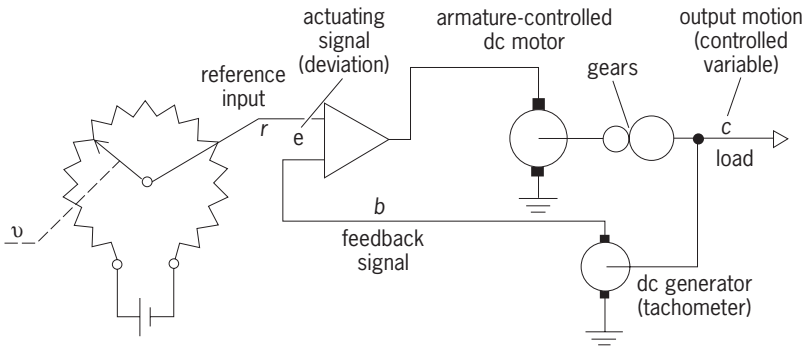

Fig. 3. Velocity or rate servomechanism, employing a dc motor as actuator and a dc generator as a feedback sensor.

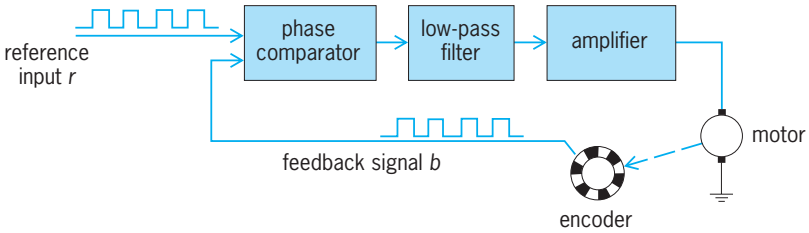

Fig. 4. Phase-locked servo loop.

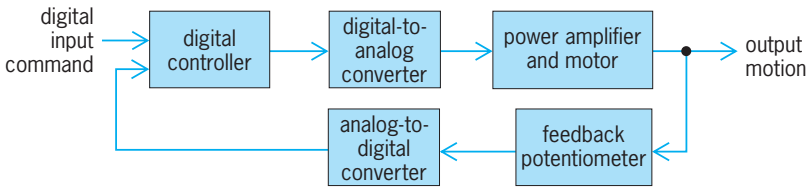

Fig. 5. Digitally controlled positional servomechanism.

The phase-locked servo of **Fig. 4** employs an encoder or pulse tachometer as the feedback sensor. The command signal and the feedback signal are square waves or sine waves of variable frequency. The servo is phase-locked when the frequencies and phases of the two signals become identical. As long as the servo is phase-locked, the motor velocity follows the frequency of the command independently of component parameter variations. The phase-locked servo is really a positional control system used for velocity control: in the event of a disturbance, such as a change in load torque, the motor will quickly recover to its original velocity and position.

Digital servomechanism. Digital control has many advantages over the analog control schemes of Figs. 2 and 3. A digitally controlled positional servo (**Fig. 5**) has digital-to-analog and analog-to-digital converters that interface the digital microcontroller with the motor and the feedback potentiometer, both of which are analog components. *See* ANALOG-TO-DIGITAL CONVERTER; DIGITAL CONTROL; DIGITAL-TO-ANALOG CONVERTER.

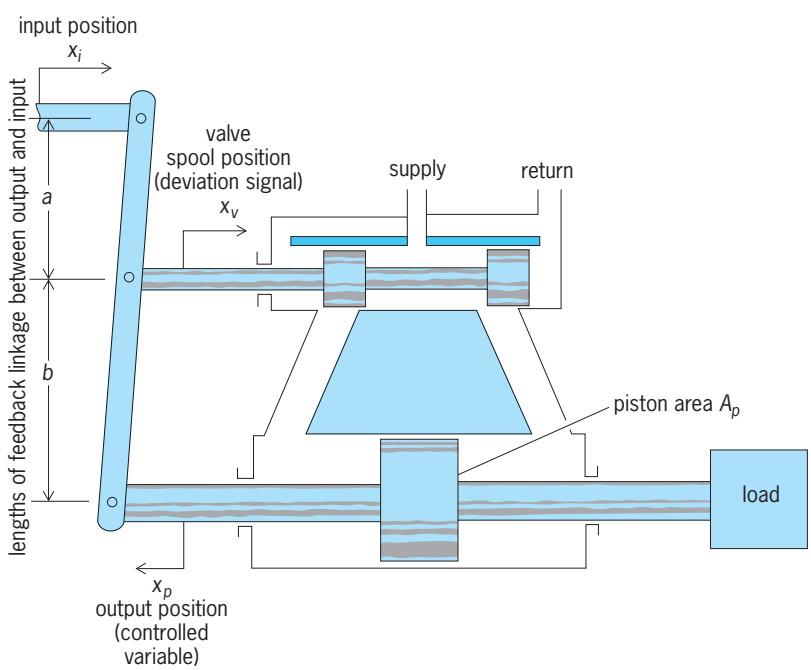

Fig. 6. Hydromechanical servomechanism, with two-stage servovalve. (*After H. E. Merritt, Hydraulic Control System, John Wiley and Sons, 1967*)

Stepping servomechanism. The stepping servo employs a special kind of motor, a stepping motor (also called a step motor or stepper). Any electric motor will move in discrete steps when it is energized from a sequence of discrete voltage pulses, but the size of these steps cannot be exactly controlled. The stepping motor turns by a precisely determined angle with each pulse. It is basically an ac synchronous motor with special construction features for stepping service. *See* STEPPING MOTOR.

Because of its precise response, positional control with a stepping motor is often done open-loop, except for an index mark that determines a starting point or zero positional reference. Starting at the index, the stepper turns by a predetermined angle for each pulse. An error due to a lost or spurious pulse corrects itself as soon as the stepper returns to the index position.

A closed-loop stepping servo is used when errors due to lost or spurious pulses are unacceptable, such as in speed control. Feedback makes it possible to drive the motor at a higher stepping rate, and permits rapid resumption of positional control after a power interruption. The feedback transducer can be an analog component, such as a resistance potentiometer or a dc tachometer generator, or it can be an optical encoder. The stepping servo with encoder feedback is sometimes called an all-digital servo, but the stepper and the incremental encoder are really pulse analog rather than digital devices. Pulse analog signals are much easier to interface with digital controllers than the customary voltage-amplitude analog signals.

Alternating-current servomechanism. It is sometimes desirable to use inductive and capacitive transducers, such as synchros and differential transformers, which require ac excitation. The mechanical input to such a transducer results in a suppressed-carrier modulated output voltage (double-sideband amplitude modulation without carrier). Currently a mixture of ac-dc analog signals or of ac analog with digital signals is used. A typical application is in 360° position control, for example, an antenna azimuth drive, which employs a synchro as a position sensor and a dc motor as an actuator. Synchronous detectors (phase-sensitive rectifier circuits) convert the synchro output ac voltage to dc. Synchro-to-digital and digital-to-synchro converters interface with digital controllers and computers. *See* MODULATION; SINGLE SIDEBAND.

Hydraulic and pneumatic servomechanisms. Hydraulic and pneumatic drives have a number of advantages over electric drives. The principal one is their ability to develop a force at zero speed without power dissipation; on the other hand, obtaining a standstill or holding force from an electric motor or solenoid entails power dissipation and temperature rise which severely limit the available force. This

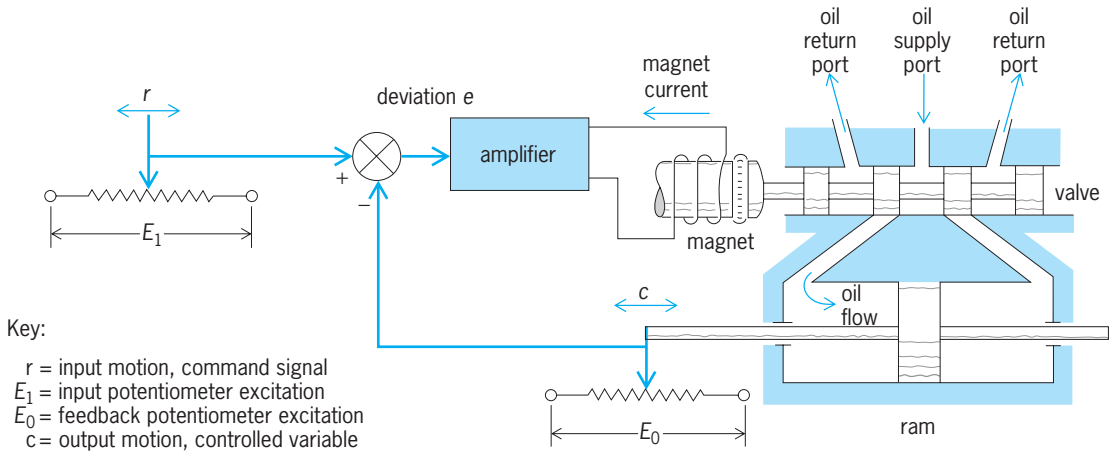

Fig. 7. Electrohydraulic servomechanism. (*After R. N. Clark, Introduction to Automatic Control Systems, John Wiley and Sons, 1962*)

advantage is a major consideration in positional servomechanisms where operation at standstill is often the rule. For a hydraulic drive, other advantages are smaller weight per output power and faster response time. *See* HYDRAULIC ACTUATOR.

Hydraulic power systems are either valve-controlled or pump-controlled. A valve-controlled system consists of a servovalve controlling the flow from a constant-delivery pump to the motor. A pump-controlled system has a variable-delivery pump supply fluid to the motor. Valve control is faster, but less efficient because servovalves are dissipative. Pump control is slow, bulky, but power-efficient, and may be preferred at very high power levels.

Figure 6 shows a valve-controlled hydraulic or hydromechanical servo. The input motion x_i opens and closes the ports of the servovalve, the upper part of the mechanism, which in turn controls the pressure applied to the sides of the piston, the lower portion. The piston motion is fed back and subtracted from the input motion so as to null the servovalve. A hydraulic pump (not shown) supplies the fluid pressure.

The electrohydraulic servo of **Fig. 7** uses resistance potentiometers for input and feedback motion sensing. The electric error signal is amplified and applied to the winding of an electromagnet, generating a force which positions the hydraulic valve spool.

An interesting variation of the above two schemes is an electrohydraulic servo with two-stage servovalve. This two-stage valve is precisely the hydromechanical servo of Fig. 6.

Figure 8 shows a pneumatic servo, called a valve positioner, which is used as a component in process control systems. The input signal, a variable air pressure from the process controller, is converted to motion of a linkage by the air bellows. This adjusts the opening of the pilot valve, which in turn controls the pressure applied to the spring-loaded diaphragm of the process valve or valve motor. The diaphragm converts this pressure change into a motion of the valve stem and thereby controls the flow of the process stream through the main valve body. The main valve stem motion is fed back through the lever and subtracted from the input motion so as to null the opening of the pilot valve.

The purpose of such a servo is to improve the operating characteristics of valves used in industrial process control. Such valves must have good seals between stem and body to isolate the process stream from the instrument air, and these seals exhibit high Coulomb friction or "stiction." Sticky valves are responsible for a phenomenon called mechanical hysteresis, which makes it difficult to control the performance of the system. The servomechanism, by enclosing the sticky valve in a feedback loop, reduces the hysteresis and improves smoothness of operation. *See* HYSTERESIS; PROCESS CONTROL.

The dynamic characteristics of hydraulic and pneumatic servos depend on the interaction of motor and load and on their interconnection. The effect of fluid compressibility cannot be ignored in hydraulic systems, particularly in the case of large masses and long fluid columns. Hydraulic liquids are

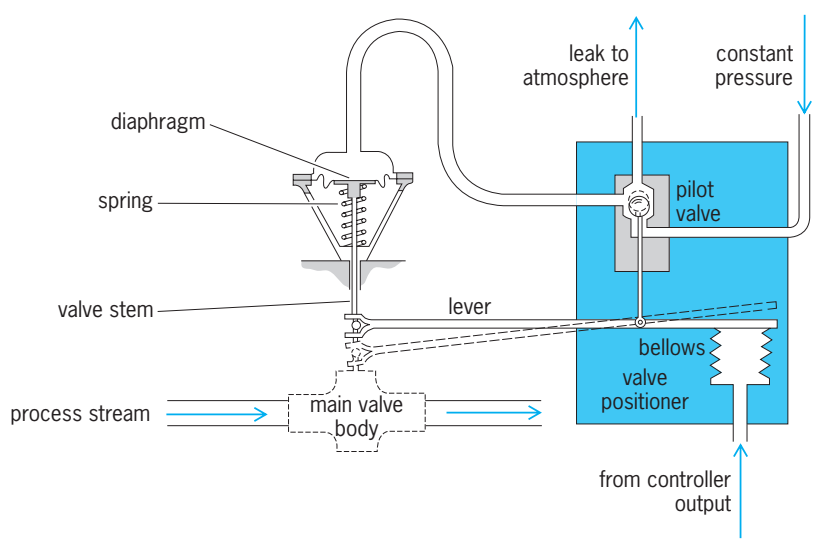

Fig. 8. Pneumatic servomechanism, or valve positioner. (*After W. A. Hadley and G. Longobardo, Automatic Process Control, Addison-Wesley, 1963*)

selected for low compressibility and must be kept free from entrained air. In pneumatic systems the compressibility is always one of the major factors in the system dynamics.

All-hydraulic and all-pneumatic servos such as the ones shown in Figs. 6 and 8 have rather limited applications. Mixed electric systems are the rule in fast-acting, high-accuracy servomechanisms. Electric motors (torquers) or electric servos are used to move the spool of the servovalve or the lever of the delivery pump. The power package (drive) is hydraulic or pneumatic, but all low-level components are electric.

Servomotors

The servomotor, actuator, or final control element is the "muscle" of the servomechanism. It must be large enough to move the load over the required range of velocities and accelerations. Load may mean inertia, friction, or external torque, or any combination thereof.

The motor is the middle part of a three-part power package consisting of (1) the power amplifier which raises the low-level controller output signal to the required power level, (2) the motor itself, and (3) the connection between motor and load, such as gears, cables, or, in some cases, merely a direct solid shaft connection. The combination of these elements is called a power drive. The drive itself often accounts for most of the weight and volume of the complete servomechanism. Servo dynamic performance, such as stability and response time, is determined primarily by the combined parameters of the power drive and load. All these elements interact, and they cannot be studied in isolation. In some servos the load is negligible and the servo dynamics depends primarily on the drive alone.

Electric motors employed in servomechanisms are usually standard motors developed for general (open-loop) use. This is particularly so in the case of large motors, 1 hp (0.75 kW) and more, where special modifications would be very costly. For small sizes, manufacturers have developed special motors for servomechanisms, originally as a matter of necessity, and subsequently to optimize designs. Three areas critical to servomechanisms are stickiness, response time, and shaft compliance.

Stickiness in electric servomotors has the same deleterious effect on system performance as described above for the pneumatic valve positioner. In the electric motor the stickiness is due not so much to mechanical friction but to an effect called magnetic cogging. The nonuniformity of the motor air gap due to the winding slots causes torque fluctuations (torque ripple) as the motor rotates; the motor has a tendency to lock itself into a position of minimum energy.

The response time of a motor is limited by the moment of inertia of its rotor, by the inductance of its winding, and by the maximum torque which can be generated without overheating. Inertia and inductance must be minimized, and torque per watt maximized.

Shaft compliance is reduced and backlash is minimized if as many moving components as possible can be assembled and mounted on a common shaft. One step in this direction is the integrally mounted motor-tachometer or the integrally mounted motor-encoder. Another approach is the gearless or direct-drive connection from motor to load.

Direct current servomotor. This type is favored in servo applications because its torque and speed can be readily controlled by varying the applied voltage. Small- and medium-size dc servomotors have permanent-magnet fields and are controlled by a voltage applied to the armature circuit (armature control). Large dc motors require field windings, and can be controlled by varying the applied field voltage (field control), by armature control or by a combination. A principal drawback of the dc motor is the complexity, limited life, unreliability, and electromagnetic interference associated with its commutator and brushes. *See* COMMUTATOR.

The amplifier that controls the dc servomotor from the available ac power source can have a number of forms. Linear amplifiers are employed at power levels up to about 100 W. Large dc motors are driven from switching amplifiers, thyristors (controlled rectifiers), or magnetic amplifiers. For very large drives, a scheme called the Ward-Leonard system uses a separately excited dc generator to supply the armature voltage of the dc servomotor. A small change in generator field voltage produces a large change in armature voltage and power; hence the term rotating amplifier. The shaft of the dc generator is mechanically driven by an ac motor, or perhaps by a diesel engine, which acts as a power source and runs at near-constant speed. *See* AMPLIFIER; DIRECT-CURRENT MOTOR; MOTOR; SEMICONDUCTOR RECTIFIER.

Besides the conventional dc servomotor, a number of specialized dc motors tailored to servomechanism applications are in use, such as the moving-coil motor, the disk-armature motor (printed motor); and the brushless dc motor.

Moving-coil motor. This type of motor superficially resembles a standard dc motor, with a permanent-magnet-field structure on the stator and an armature on the rotor, but the armature structure (**Fig. 9***a*) is hollow (cup or shell armature). Inside this rotating shell is a stationary iron core, which serves as a magnetic return path for the field flux. Thus, the magnetic path is stationary in its entirety, and torque ripple is virtually eliminated. Integrally mounting the armatures of a moving-coil motor and a dc tachometer-generator on a common shaft (Fig. 9*b*) eliminates the backlash between motor and tachometer and minimizes the compliance.

Brushless dc motor. In this type of motor, the mechanical commutator and brushes of the conventional dc motor are replaced by position sensors and electronic switches. A practical design (**Fig. 10**) has its armature winding carried on the stator and its permanent-magnet field on the rotor. This reversal of the usual dc motor arrangement eliminates all sliding contacts and improves heat transfer out of the motor.

Fig. 9. Cup armatures of a moving coil servomotor. (*a*) Open basket-type construction designed for high power dispersion when air-cooled. (*b*) Resin-molded moving coil winding and integrally coupled tachometer-generator armature. (*Robbins Myers/Electro-Craft Corp.*)

Alternating current servomotor. This type does not require a commutator, and is thus generally more reliable and economical than the dc motor. It is inherently fixed-speed, and its use as a control motor requires a variable-frequency power supply. The frequency synthesizer, a microprocessor-controlled electronic switching amplifier, converts the available fixed-frequency power into a controllable variable-frequency source. The two-phase servo induction motor is a special design with variable speed characteristic, but its inefficiency limits its size to a few watts. Even in these low-power applications, it has been largely superseded by the dc servomotor or by the stepping motor. As stated above, the stepping motor is essentially an ac synchronous motor, and its speed is determined by the applied pulse rate. *See* ALTERNATING-CURRENT MOTOR; INDUCTION MOTOR; SYNCHRONOUS MOTOR.

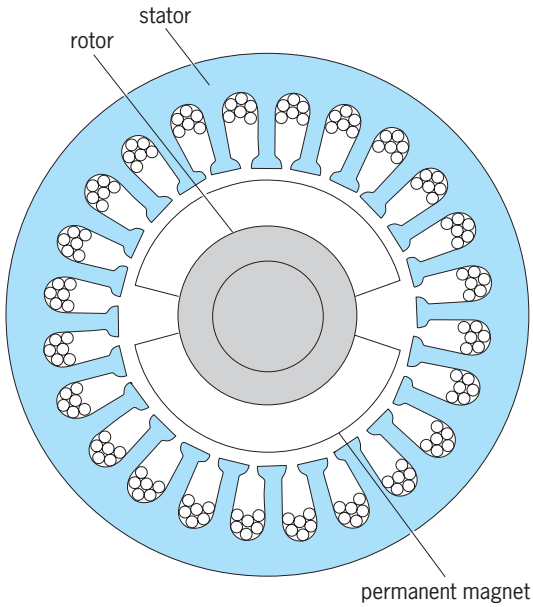

Fig. 10. Cutaway view of brushless dc motor assembly. (*After Electro-Craft Corp. Staff, Handbook of DC Motors, Speed Controls, and Servo Systems, 5th ed., Robbins & Myers/ Electro-Craft Corp., 1980*)

Unconventional motors. When the motion to be controlled is rotary but limited in angle, or when it is translational (rectilinear), an unconventional electric motor may be superior, particularly in small instruments such as cameras, magnetic-disk players, video-disk players, and gyroscopes.

A motor for limited rotational control is required to supply a specified torque, but it never really picks up full speed. Such a motor is called a torquer (or torque motor or torque generator). In direct-drive (gearless) designs, the torquer is often supplied in the form of separate rotor and stator pieces without bearings or housing, for direct installation in the user equipment.

Translational motors can take a variety of forms. The voice-coil motor resembles a conventional electrodynamic loudspeaker movement. A piezoelectric crystal is often employed as the actuator (bender, pusher, puller) for tiny motions, such as the positioning of lenses and prisms in optical instruments. Magnetostriction is also used in motors for fine motion control. The linear induction motor is a stretched-out, translational form of the ac induction motor that has been used in shuttle mechanisms and for proposed high-speed electric railroads. *See* LOUDSPEAKER; MAGNETOSTRICTION; PIEZOELECTRICITY.

Measuring Devices (Sensors)

The quantities measured in motion control are displacement, velocity, acceleration, and force, both in translation and in rotation. Many physical phenomena have been exploited in instrument design. Most sensors and transducers used in laboratory and industrial practice have been or can be adapted to servos. Electrical and optic devices predominate. A few of the most common transducers will be described. *See* PHYSICAL MEASUREMENT; TRANSDUCER.

Analog transducers. The resistance potentiometer is used primarily to measure limited-angle rotation, but can also be designed for translation. The servo potentiometer is a precision component with high-quality bearings. Resistance linearity may be as good as 0.05% of full scale. *See* POTENTIOMETER.

The differential transformer is frequently employed in the control of small translational motion. It is an inductive transducer and requires ac excitation. The generation of the ac input with the conversion of the output to dc is available packaged inside the transformer. *See* DIFFERENTIAL TRANSFORMER.

Synchros are angular position sensors that can handle the measurement of limitless 360° motion. They are inductive devices with ac output, and come in two kinds. Torque synchros, always operated in pairs, are induction motors with interconnected rotors which make them self-aligning or self-synchronous. They require no further electronic

amplifiers, but their accuracy is poor. They have been largely superseded by the family of control synchros: transmitters, differentials, control transformers, and resolvers. The control synchro is an outstandingly accurate component (typically 3 arc-minutes, equivalent to 0.014% of full travel).

The dc tachometer is the most common analog transducer for velocity (rate generator). Physically it is a small permanent-magnet dc generator, designed to obtain high linearity, high stability with temperature, and low output ripple. Linearity of 0.2% of full scale is achievable. Servo tachometers are supplied as individual bearing-mounted components, or as part of an integral motor-tachometer unit with a common shaft and, possibly, a common field magnet. *See* TACHOMETER.

A pulse tachometer can be constructed by depositing regularly spaced tick marks along the edge of the body whose speed is to be measured, be it in rotation or translation. The tick marks are detected optically or magnetically and used to generate pulses. The observed pulse rate or pulse repetition frequency is proportional to speed. When the pulse tachometer is supplied as a separate bearing-mounted unit, it is effectively an encoder, described below.

Gyroscopic and seismic instruments sense and measure absolute motion in space, and are employed in the control of systems mounted on moving vehicles and stabilized platforms, such as cameras, antennas, and guns. *See* ACCELERATION MEASUREMENT; ACCELEROMETER; GYROSCOPE.

Optical encoders. Encoders come in two kinds, absolute encoders and incremental encoders. The absolute encoder is a position transducer with output in the form of parallel binary digits; it is in effect an electromechanical analog-to-digital converter. The incremental encoder produces a single output pulse for a given displacement. It is in effect a pulse tachometer and can be used to measure either position or velocity. The encoder can be constructed to measure rotation (rotary encoder; formerly known as shaft encoder) or translation (linear encoder).

Absolute encoder. The heart of the absolute encoder is a patterned code disk with a number of tracks, one track per binary digit (**Fig. 11**) constructed of alternating transparent and opaque segments and read by means of a light source, focusing system, and photosensor array. Typical resolution is from 6 to 22 binary digits (bits) in 1- to 10-in. (2.5- to 25-cm) disk sizes. (Twenty-two bits represent a 0.3-arc-second angular resolution.) Because of the difficulty of fabricating small disks with many tracks, higher-resolution encoders are sometimes multiturn units, with two disks geared together.

It is not possible to construct an encoder with such mechanical perfection that all sensors will consistently operate in perfect synchronism. As a result, the natural binary code illustrated in Fig. 11 results in errors when there is a transition which requires a change in more than one digit, for example, in going from binary 011 to binary 100. All digital codes in which adjacent numbers can differ in more than one digit (polystrophic codes) suffer from the same

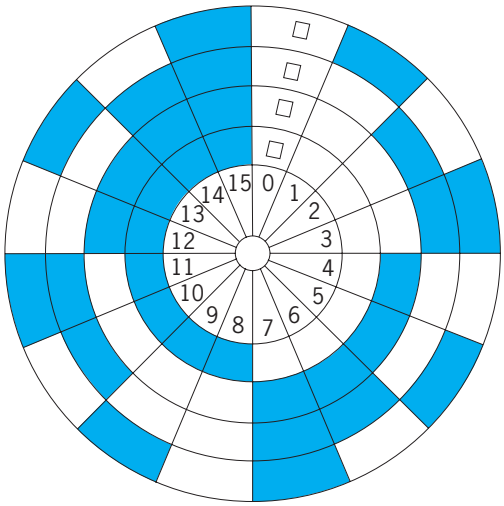

Fig. 11. Four-digit encoder disk. (*After E. O. Doebelin, Measurement Systems, 3d ed., McGraw-Hill, 1983*)

switching ambiguity. There are two solutions to this problem, the addition of extra sensors, or the employment of a special binary code which changes only one digit at a time (monostrophic code), such as a reflected or cyclic code, of which the Gray code is the most common. The **table** shows the Gray code for a three-bit disk; all transitions between adjacent numbers require only one bit change, including the transition from the last to the first number. The output of an encoder may actually be in any of the codes employed in computers, such as binary-coded-decimal (BCD) or excess-3. Logic circuits to convert from the disk code to the output code are then packaged inside the encoder housing. *See* INFORMATION THEORY; LOGIC CIRCUITS.

Older encoders were read by electric contacts, (brushes), magnetically, or capacitatively. A brush encoder disk has conductive and nonconductive segments read by a set of brushes riding on each track. A magnetic code disk has magnetized segments sensed by tiny toroidal coils serving as reading heads. To avert switching ambiguities, brush encoders have two brushes per track, arranged in various configurations known as U-scan and V-scan. Diode logic circuits select the proper sequence of brushes to provide ambiguity-free output.

Incremental encoder. In its simplest form, the incremental encoder has only one track with on-off segments. The number of pulses produced by the sensor is a measure of the displacement, while the pulse

Natural binary and Gray codes

Decimal	Natural binary	Gray
0	000	000
1	001	001
2	010	011
3	011	010
4	100	110
5	101	111
6	110	101
7	111	100

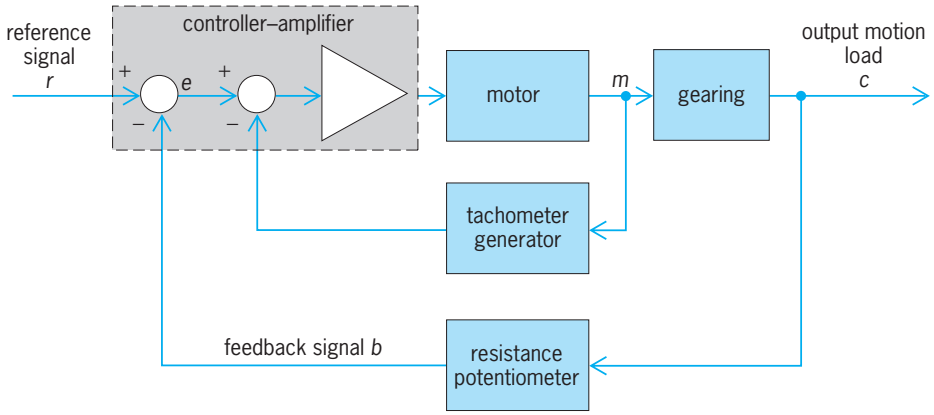

Fig. 12. Tachometer-damped positional servomechanism.

rate indicates the velocity. For absolute motion measurement, the incremental encoder is used with an electronic counter which totalizes the output pulses. Bidirectional sensing is accomplished by employing two identical tracks offset by one-half of a segment length. A third track which generates one pulse per revolution is sometimes provided as a zero marker or index. In the event of a signal interruption or power failure, the error due to lost pulses can be corrected when the index is sensed. The incremental encoder is also used as a unidirectional (one track) or bidirectional (two tracks) pulse tachometer.

Compensation (Control Algorithm)

A servomechanism must be stable and exhibit specified static and dynamic accuracy. These requirements are usually incompatible; the trade-off is that an increase in loop gain improves static accuracy but makes the servo less stable. The process of overcoming this conflict and achieving simultaneously the desired degree of accuracy and stability is called compensation (or stabilization) in analog technology; the term control algorithm has a similar meaning for digital systems.

Static accuracy depends on the zero-frequency (static) loop gain, which becomes infinite if the feedback loop contains an integration. This is indeed the case in the positional servo which uses a conventional motor with output speed proportional to input voltage. The desired output variable, motor displacement, is the integral of motor speed. Hence, the positional servo contains an integration and can have zero static error for a constant input. In Fig. 1, the actuating signal e that represents this error will be zero when the command v is constant and the disturbance d is zero. The velocity servo, on the other hand, does not include an inherent integration, and the loop error as exhibited by the signal e will not be zero. Feedback loops are classified as types 0, 1, 2, and so forth, according to the number of integrations enclosed in the loop. The higher the type number the better the static accuracy, but the poorer the stability.

Standard analog or digital controllers in the compensator block of Fig. 1 implement the proportional-integral-derivative (PID) function and can produce the control modes proportional control (P), proportional-plus-integral (PI), proportional-plus-derivative (PD), or combined PID control. Compensators can also be characterized in terms of their frequency response, the simplest functions being lag, lead, and lag-lead. To some extent, this is only a change in terminology: phase lead compensation is equivalent to PD control, phase lag is an approximation to PI control, and lag-lead approximates PID control. However, integral control adds an integration to the loop, while phase lag does not. This small difference may be significant in the positional servo because PI control converts it into a type-2 loop, which could be unstable. For large loads with structural resonances, more elaborate control algorithms such as notch filters (band-rejection filters) may be required. *See* ELECTRIC FILTER.

Some other methods of compensation do not fit into the simple framework of Fig. 1. For positional servos, rate feedback (**Fig. 12**) is often very effective. Here, a second, inner, loop is added that feeds back a signal proportional to output velocity; this scheme is a form of linear-state feedback. The tachometer for the auxiliary feedback need not be accurate but must be tightly coupled to the motor. In positional or velocity servos with large torque load, it is possible to improve the performance by feeding a signal proportional to torque into the controller. This is, in fact, a feed-forward control scheme. *See* LINEAR SYSTEM ANALYSIS.

Gerald Weiss

Bibliography. B. W. Andersen, *The Analysis and Design of Pneumatic Systems*, 1967, reprint 2001; G. Biernson, *Principles of Feedback Control*, 2 vols., 1988; E. B. Canfield, *Electromechanical Control Systems and Devices*, 1965, reprint 1977; Y. Dote and S. Kinoshita (eds.), *Brushless Servomotors: Fundamentals and Applications*, 1990; Electro-Craft Corp. (Eden Prairie, Minnesota) Staff, *Handbook of DC Motors, Speed Controls, and Servo Systems*, 5th ed., 1980; H. E. Merritt, *Hydraulic Control Systems*, 1967, reprint 1991; R. T. Pickett, *Feedback Control Systems for Technicians*, 1988; J. J. Pippenger, *Hydraulic Valves and Controls*, 1984; D. P. Sante, *Automatic Control System Technology*,

1980; R. B. Walters (ed.), *Hydraulic and Electro-Hydraulic Control Systems*, 1992.

Set theory

A mathematical term referring to the study of collections or sets. Consider a collection of objects (such as points, dishes, equations, chemicals, numbers, or curves). This set may be denoted by some symbol, such as X. It is useful to know properties that the set X has, irrespective of what the elements of X are. The cardinality of X is such a property.

Cardinality of sets. Two sets A and B are said to have the same cardinal, written $C(A) = C(B)$, provided there is a one-to-one correspondence between the elements of A and the elements of B. For finite sets this notion coincides with the phrase "A has the same number of elements as B." However, for infinite sets the above definition yields some interesting consequences. For example, let A denote the set of integers and B the set of odd integers. The function $f(n) = 2n - 1$ shows that $C(A) = C(B)$. Hence, an infinite set may have the same cardinal as a part or subset of itself.

Subset. A is called a subset of B if each element of A is an element of B, and it is expressed as $A \subset B$. The collection of odd integers is a subset of the reals. Also, each set is a subset of itself.

Continuum hypothesis. An infinite set is called uncountable if it cannot be put in a one-to-one correspondence with the positive integers. Here is one of the unsolved problems of set theory. It is of particular interest since so many mathematicians have tried unsuccessfully to solve it. If X is an uncountable subset of the reals R, is $C(X)$ equal to $C(R)$? The conjecture that the answer is in the affirmative is called the continuum hypothesis. It has been shown that the answer cannot be decided by using the ordinary axioms of set theory.

Comparing cardinality of sets. One says that $C(A) \leqq C(B)$ if there is a one-to-one correspondence between the elements of A and a subset of the elements of B. One useful theorem that can be proved states that any two sets A, B are comparable, that is, either $C(A) \leqq C(B)$ or $C(B) \leqq C(A)$ (possibly both). Another theorem states that if $C(A) \leqq C(B)$ and $C(B) \leqq C(A)$, then $C(A) = C(B)$. Each of these results may be proved by using well orderings of A and B.

Ordering. An ordering is one way of setting up a one-to-one correspondence between two sets of the same cardinality. A relation $<$ is an order relation for a set X if it satisfies the following conditions:

1. If x_1, x_2 are two elements of X, either $x_1 < x_2$ or $x_2 < x_1$ (any two elements are related).
2. $x_1 \not< x_1$ (no element is less than itself).
3. If $x_1 < x_2$ and $x_2 < x_3$, then $x_1 < x_3$ (the order relation is transitive).

The following are examples of ordered sets: a horizontal line where $<$ means "to the left of"; the reals where $<$ means "is less than"; and the collection of words in the dictionary where ordering is alphabetical.

Well ordering. An ordering of a set is called a well ordering if it satisfies the additional condition:

4. Each non-null subset Y of X has a first element; that is, there is an element y_o of Y such that if y' is another element of Y, $y_o < y'$.

The natural ordering of the positive integers is a well ordering, but the natural ordering of neither the integers nor the reals is a well ordering. A well ordering for the integers is $0, 1, 2, \ldots, -1, -2, \ldots$. Since a well ordering of the reals cannot be written down, one might guess that there is none. This guess is shown to be false by the theorem that states that any set X has a well ordering. In proving this, one considers the collection Z of all non-null subsets of X, selects a point $x_a = f(z_a)$ from each element z_a of Z, and well orders X so that if S_a is the set of all elements that precedes x_a, then $x_a = f(Z - S_a)$.

Some of the theorems proved by well ordering are so strange that their truths do not seem intuitively obvious. Well ordering is also used to construct pathological examples which serve as counterexamples to various conjectures. These counterexamples are useful since they show that the conjectures are false and it is useless to try to prove them.

Formation of sets. One approved method of forming a set is to consider a property P possessed by certain elements of a given set X. The set of elements of X having property P may be considered as a set Y. The expression $p \in X$ is used to denote the fact that p is an element of X. Then $Y = \{p | p \in X$ and p has property $P\}$. Another approved method is to consider the set Z of all subsets of a given set X. It may be shown in this case that $C(X) < C(Z)$.

Paradoxically, it is not permissible to regard the collection of all sets as a set. If such a collection X were called a set, and Z were used to denote the set of all subsets of X, one would arrive at the absurdity that $C(X) < C(Z)$.

Operations with sets. In set theory, one is interested not only in the properties of sets but also in operations involving sets: addition, subtraction, multiplication, and mapping.

Sum or union. The sum of A and B ($A + B$ or $A \cup B$) is the set of all elements in either A or B; that is, $A + B = \{p | p \in A$ or $p \in B\}$.

Intersection, product, or common part. The intersection of A and B ($A \cdot B$, $A \cap B$, or AB) is the set of all elements in both A and B; that is, $A \cdot B = \{p | p \in A$ and $p \in B\}$. If there is no element which is both A and B, one says that A does not intersect B and writes $A \cdot B = 0$.

Difference. The expression $A - B$ is used to denote the collection of elements of A that do not belong to B; that is, $A - B = \{p | p \in A$ and $p \notin B\}$. If $A \subset B$, it is expressed as $A - B = 0$.

An example. If a person were to squirt some black ink on a plane, the set A of points in the dark spot would be an example of a point set. Suppose a set B is determined by squirting some red ink on the plane. Then $A + B$ designates the set of points covered by ink, $A \cdot B$ designates the set covered by both kinds of ink, and $A - B$ designates those covered by black but not red ink (**Fig. 1**).

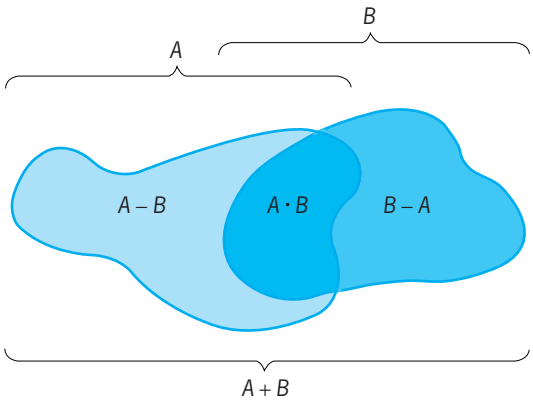

Fig. 1. Sets of points on a plane.

Boolean algebra. By using the previous notation, it follows that the sets of Fig. 1 satisfy some of the familiar laws of algebra as

$$A + B = B + A$$
$$A \cdot (B + C) = A \cdot B + A \cdot C$$

However, other identities are not so familiar:

$$X - (A + B) = (X - A) \cdot (X - B)$$
$$X - A \cdot B = (X - A) + (X - B)$$

See BOOLEAN ALGEBRA.

Transformations. A transformation of a set X into a set Y is a function that assigns a point of Y to each point of X. The transformation shown in **Fig. 2** is the vertical projection of a set X onto a segment Y. The point assigned to X under a transformation f is called the image of x and denoted by $f(x)$. Also the set of all points x sent into a particular point y of Y is called the inverse of y and denoted by $f^{-1}(y)$. The inverse of y shown in Fig. 2 is the sum of three segments. The equation $f(x) = x^2$ represents a transformation that takes each real number into its square, for example, 2 to 4, −5 to 25, and so on. Rotations, congruences, and similarities are examples of transformations from geometry. However, in general, a transformation may change both the size and shape of an object.

Topology. Topology is one of the branches of mathematics that makes extensive use of set theory. Not only does one have sets of points for consideration but also collections of interiors of spheres or neighborhoods. These neighborhoods enable one to study limit points and the continuity of transformations. *See* CONFORMAL MAPPING; RING THEORY; TOPOLOGY.

R. H. Bing

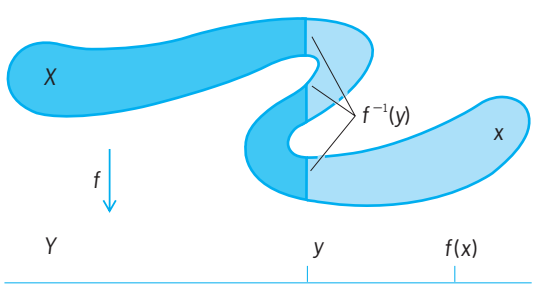

Fig. 2. Vertical projection of set *X* onto segment *Y*.

Bibliography. K. Devlin, *The Joy of Sets: Fundamentals of Contemporary Set Theory*, 2d ed., 1994; M. D. Potter, *Sets: An Introduction*, 1991; J. Roitman, *Introduction to Modern Set Theory*, 1990; R. L. Vaught, *Set Theory: An Introduction*, 2d ed., 1994.

Sewage

Water-carried wastes, in either solution or suspension, that flow away from a community. Also known as wastewater flows, sewage is the used water supply of the community. It is more than 99.9% pure water and is characterized by its volume or rate of flow, its physical condition, its chemical constituents, and the bacteriological organisms that it contains. Depending on their origin, wastewaters can be classed as sanitary, commercial, industrial, or surface runoff.

The spent water from residences and institutions, carrying body wastes, ablution water, food preparation wastes, laundry wastes, and other waste products of normal living, are classed as domestic or sanitary sewage. Liquid-carried wastes from stores and service establishments serving the immediate community, termed commercial wastes, are included in the sanitary or domestic sewage category if their characteristics are similar to household flows. Wastes that result from an industrial process or the production or manufacture of goods are classed as industrial wastes. Their flows and strengths are usually more varied, intense, and concentrated than those of sanitary sewage. Surface runoff, also known as storm flow or overland flow, is that portion of precipitation that runs rapidly over the ground surface to a defined channel. Precipitation absorbs gases and particulates from the atmosphere, dissolves and leaches materials from vegetation and soil, suspends matter from the land, washes spills and debris from urban streets and highways, and carries all these pollutants as wastes in its flow to a collection point. Discharges are classified as point-source when they emanate from a pipe outfall, or non-point-source when they are diffused and come from agriculture or unchanneled urban land drainage runoff. *See* HYDROLOGY; PRECIPITATION (METEOROLOGY).

Wastewaters from all of these sources may carry pathogenic organisms that can transmit disease to humans and other animals; contain organic matter that can cause odor and nuisance problems; hold nutrients that may cause eutrophication of receiving water bodies; and may contain hazardous or toxic materials. Proper collection and safe, nuisance-free disposal of the liquid wastes of a community are legally recognized as a necessity in an urbanized, industrialized society.

In urban areas, wastewaters are generally conveyed away through a series of buried pipes. The inlets, manholes, pipes, and other appurtenances constitute the sewerage system. A system carrying only domestic wastes is known as a sanitary sewer system; if only storm runoff is collected in the drains, it is a

storm drainage system; but if both storm water and wastewater are carried in the same sewerage system, the system is known as a combined system. In rural or suburban areas, sanitary sewage from individual buildings may be collected and disposed of on-site through the use of septic tanks and disposal fields. *See* SEPTIC TANK.

Sewage flows. The used water supply leaving a community is directly related to the quantity and timed use of water entering that community. To establish values for the engineered design of collection systems, pump stations, and treatment plants, the water consumed within a community must be measured; or it must be evaluated, taking into account climate, extent of industrialization, types of residential districts, socioeconomic status, water cost, water quality, and delivery pressure. Not all water delivered to a community ends up as flow into the sewerage system. Water is lost through public use, fire fighting, street washing, and park and home irrigation, and is consumed in commercial and industrial production and for steam generation and cooling. Per-capita water use tends to rise with increased community size (**Table 1**). *See* WATER SUPPLY ENGINEERING.

Some of the lost flow is replaced by additions from ground-water infiltration, from inflows and, at times, from private wells and industrial activities. Infiltration is ground-water leakage through poor joints, fittings, and connections into the sewerage system lying below the ground-water level. Inflow is surface water, or directed water entering the sewerage system through manhole covers, illegal roof leaders, foundation and area drains, cooling water discharges, and street washing. *See* GROUND-WATER HYDROLOGY.

A community's wastewater flow may be only 60–75% of the average water supply when new sewerage systems with tight joints and no illegal connections are installed, and infiltration/inflow is kept to a minimum. However, in time, infiltration/inflow as well as unauthorized connections increase; so engineers frequently assume the community's wastewater flow is equal to community water supply.

Human activities follow a diurnal pattern; thus fluctuations in water supply and consequently in sewage flows occur according to time of day, day of the week, weather conditions, holidays, and seasons (**Fig. 1**). Size of community, as well as the presence and type of industry, and institutions such as colleges, prisons, or hospitals all affect the quantity and fluctuation of flows. Within the same community, wastewater peaks are smaller and lag behind water supply peaks. Water supply is a pressure system, with demand anywhere in the service area being met instantaneously. Flow into the distribution network is measured at the point of supply to the total pressurized supply system. The sewerage collection system is essentially a gravity operating system whose flow is usually measured at a final collection point, the treatment plant. The wastewater, collected from all parts of the community, requires time (lag time) to flow from the user through the pipes to the treatment plant. The flow is further detained by storage within access holes and within the pipes themselves. On any given day the flow is lowest during the night, increasing to peaks at breakfast time and in the evening. In general, the smaller the community, the greater the daily variation from the average. Water use, and hence sewage flow, not only varies with time of day but also varies with time of year. During hot, dry months the flows are greater than during cold, wet periods. Infiltration/inflow varies with rainfall frequency and intensity.

For proper engineering design to safely control and contain all anticipated variations in flow, adequate capacity must be provided within the sewerage system, pumping stations, diversions, treatment plants, and outfalls. To develop practical design flow values, not only the diurnal flow variations but also peak flows must be considered. Peak flows are the maximum flows that may occur over a relatively short period of time, usually time of concentration for storm flows and up to 1 h for other wastewater flows. The total annual flow for a community, divided by 365 days, is termed the average daily flow, and this value is taken as unity. Ranges of ratios of anticipated flow fluctutations are frequently used in wastewater system design (**Table 2**).

Peaking factors, combined with infiltration/inflow values, must be analyzed for the total community, as well as for each major district within the community. The smaller and more homogeneous the study area, the more intense the peaking factor. Peaking factors are vital for proper design of pumping stations, treatment plants, and sewerage systems. Many regulatory agencies have set 100 gpd (378 lpcd) as the average daily flow for the design of sewers and plants, with peak flows into sewers at 400 gpd (1500 lpcd) for small populations, decreasing to 200 gpd (750 lpcd) for populations over 100,000.

Storm wastewater. This is that portion of precipitation that, in modern urban environments, must be safely collected and economically removed through a controlled system. Precipitation follows three paths at the Earth's surface. In any rainfall occurrence a portion of that storm's water is removed as it initially wets and cools the ground, is retained on foliage, is temporarily stored in ground depressions, is evaporated from the surface, or is transpired by vegetation. As more rain falls and overcomes this initial abstraction, a portion of the precipitation infiltrates the soil mantle and percolates downward to the ground-water table. If the rainfall exceeds the rate that can percolate into the ground, this excess water flows overland and is conveyed away as storm water runoff. *See* SURFACE WATER.

For the proper engineering design of storm water drainage systems, the volumes of storm wastewater to be handled must be determined. In the rational method, runoff is related to rainfall through the formula $Q = cIA$, where Q is the peak runoff rate from a drainage area in ft^3/s, c is a dimensionless ratio incorporating all abstractions and losses and is an indication of the portion of the rainfall that can be anticipated to actually run off from the subject drainage area, I is rainfall intensity in in./h based on the time of

TABLE 1. Average wastewater flow from various sources*

Source	Unit	Flow, gallons per unit per day† Range	Typical
Residential sources			
Apartment			
High-rise	Person	35–75	50
Low-rise	Person	50–80	65
Hotel	Guest	30–55	45
Individual residence			
Typical home	Person	45–90	70
Better home	Person	60–100	80
Luxury home	Person	75–150	95
Older home	Person	30–60	45
Summer cottage	Person	25–50	40
Motel			
With kitchen	Unit	90–180	100
Without kitchen	Unit	75–150	95
Trailer park	Person	30–50	40
Commercial sources			
Airport	Passenger	2–4	3
Automobile service station	Vehicle served	7–13	10
	Employee	9–15	12
Bar	Customer	1–5	3
	Employee	10–16	13
Departmental store	Toilet room	400–600	500
	Employee	8–12	10
Hotel	Guest	40–56	48
	Employee	7–13	10
Industrial building (sanitary waste only)	Employee	7–16	13
Laundry (self-service)	Machine	450–650	550
	Wash	45–55	50
Office	Employee	7–16	13
Restaurant	Meal	2–4	3
Shopping center	Employee	7–13	10
	Parking space	1–2	2
Institutional sources			
Hospital medical	Bed	125–240	165
	Employee	5–15	10
Hospital, mental	Bed	75–140	100
	Employee	5–15	10
Prison	Inmate	75–150	115
	Employee	5–15	10
Rest home	Resident	50–120	85
School, day			
With cafeteria, gym, and showers	Student	15–30	25
With cafeteria only	Student	10–20	15
Without cafeteria and gym	Student	5–17	11
School, boarding	Student	50–100	75
Recreational sources			
Apartment, resort	Person	50–70	60
Cabin, resort	Person	8–50	40
Cafeteria	Customer	1–3	2
	Employee	8–12	10
Campground (developed)	Person	20–40	30
Cocktail lounge	Seat	12–25	20
Coffee shop	Customer	4–8	6
	Employee	8–12	10
Country club	Member present	60–130	100
	Employee	10–15	13
Day camp (no meals)	Person	10–15	13
Dinning hall	Meal served	4–10	7
Dormitory, bunkhouse	Person	20–50	40
Hotel, resort	Person	40–60	50
Store, resort	Customer	1–4	3
	Employee	8–12	10
Swimming pool	Customer	5–12	10
	Employee	8–12	10
Theater	Seal	2–4	3
Visitor center	Visitor	4–8	5

*After Metcalf & Eddy, Inc., G. Tchobanoglous, and F. L. Burton, *Wastewater Engineering: Treatment, Disposal, and Reuse,* 3d ed., 1991.
†1 gal = 3.7854 liters.

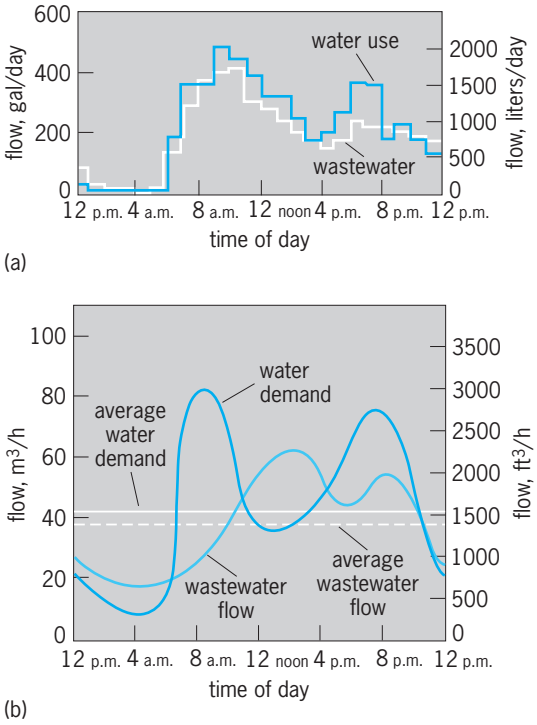

Fig. 1. Comparison of water use and wastewater flows (*a*) for a residential community (*after W. Viessman, Jr., and M. J. Hammer, Water Supply and Pollution Control, 4th ed., Harper & Row, 1985*) and (*b*) for a municipal system (*after S. R. Qasim, Wastewater Treatment Plants: Planning, Design and Operation, Holt, Rinehart and Winston, 1985*).

concentration t_c, and A is the drainage area in acres. If a rainfall with an intensity of 1 in./h falls on a drainage area of 1 acre that is totally impervious ($c = 1.0$), the peak runoff is 1 ft^3/s (28.317 liters/s). Since all the units in the U.S. Customary System work out without any conversions, this formula has been called the rational formula. The rational formula is used for urban areas of under 1 mi^2. The rational formula can be used in the International System (SI) with $Q = 2.78 \times cIA$, where Q is in liters/s, I is in mm/h, and A is in hectares. The runoff coefficient c can never be unity, because no surface is totally impervious and there are always abstractive losses. However, many regulatory agencies use a value of 1.0 for building roof areas and for concrete surfaces to simplify calculations and to incorporate a factor of safety. The wide range of runoff coefficients (**Table 3**) demands realistic evaluation of local conditions and use of engineering judgment.

Time of concentration t_c is the total time required for a drop of water to travel from the hydraulically farthest point of the tributary area to the point of collection. The term t_c is composed of two components, inlet flow time t_0 and pipe flow time t_p. Inlet (overland) flow time is the maximum time required for surface water to reach the inlet, and that time is dependent on ground slope, surface texture, and physical distance. Pipe flow time is established by dividing the length of sewer run, from inlet to point of concentration, by the velocity of the wastewater flow within the pipe. The time of concentration is equal to the sum of the two times ($t_0 + t_p$).

Rain gages at various locations in an area measure and record the quantity of precipitation collected per unit time. The rate and quantity of rainfall varies during any one storm and among the gages. Intensity–duration curves are developed by statistical analysis of rainfall data (**Fig. 2**). From accumulated gaged values, over extended time periods, a probability of frequency of occurrence can be developed. The return period is the anticipated average time interval in years between occurrences of a specified magnitude. Rainfall intensity values, correlated with the time of concentration t_c, are obtained from these intensity–duration–frequency curves. *See* PRECIPITATION MEASUREMENT.

Urban storm water volumes to be collected and carried away are evaluated on a cost-benefit basis. The greater the return period, the greater the probability of a longer time interval between flood

TABLE 2. Ranges of the ratios of flow fluctuation to average annual value

Design factor	Ratio
Maximum day in a year	(1.7–2.0) : 1
Maximum hour any day	(1.4–1.8) : 1
Peak hour in year (maximum)	(2.85–3.25) : 1
Minimum day in a year	(.50–.67) : 1
Minimum hour any day	(.35–.50) : 1
Minimum hour in year	(.20–.30) : 1
Daily average during maximum month (late summer)	(1.25–1.40) : 1
Daily average during minimum month (late winter)	(.65–.80) : 1
Daily average during maximum week (dry summer)	(1.30–1.65) : 1

TABLE 3. Recommended runoff coefficients for a specified type of area or surface*

Area or surface	Runoff coefficients†
Business	
Downtown	0.70–0.95
Neighborhood	0.50–0.70
Residential	
Single-family	0.30–0.50
Multiunits, detached	0.40–0.60
Multiunits, attached	0.60–0.75
Residential (suburban)	0.25–0.40
Apartment	0.50–0.70
Industrial	
Light	0.50–0.80
Heavy	0.60–0.90
Park, cemetery	0.10–0.25
Playground	0.20–0.35
Railroad yard	0.20–0.35
Unimproved	0.10–0.30
Pavement	
Asphalt and concrete	0.70–0.95
Brick	0.70–0.85
Roof	0.75–0.95
Lawn, sandy soil	
Flat, 2%	0.13–0.17
Average, 2–7%	0.18–0.22
Steep, 7%	0.25–0.35

* After American Iron and Steel Institute, *Modern Sewer Design*, 1980.
† The coefficients are applicable for storms of 5–10-year frequencies. Less-frequent, higher-intensity storms require the use of higher coefficients because infiltration and other losses have a proportionally smaller effect on runoff. The coefficients are based on the assumption that the design storm does not occur when the ground surface is frozen.

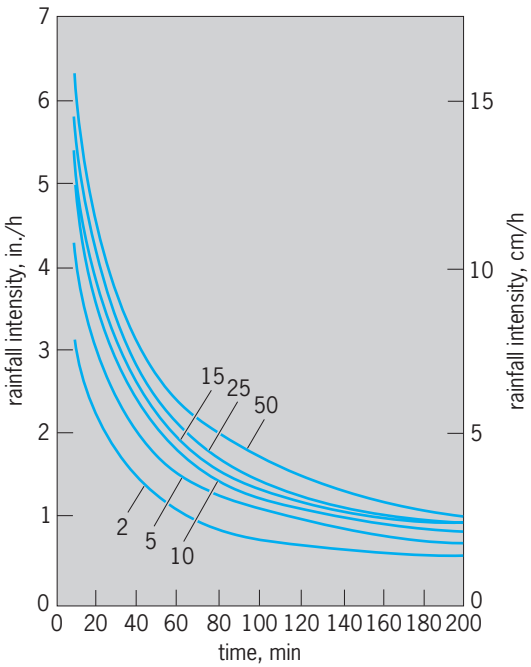

Fig. 2. Rainfall intensity-duration-frequency curves. Frequency (in years) is indicated by numbers on curves. (*After Design of Wastewater and Stormwater Pumping Stations, Manual of Practice no. FD-4, Water Pollution Control Federation, 1981*)

damage or hazardous conditions. A larger return period also means larger design flow volumes, greater required capacity of the installed system, and consequently greater construction costs. Engineering judgment must be used to balance construction cost against reduction in future inconvenience, traffic delays, and potential flooding damages. The more serious or costly the anticipated flooding, the greater the justification for the initial higher installed costs. In areas with small structural development, a 5-year return period might be acceptable, while in a highly developed area with an extensive infrastructure or a major transportation hub, a more conservative design with a 25-year, or even a 50-year, return might be reasonable.

Design period. This is the time span for which a proposed system or installed infrastructure should be able to function in a safe manner, that is, be capable of providing the desired service without structural failure. The design period is also known as project life expectancy. The U.S. Internal Revenue Service provides estimated life expectancies for many industrial components, with hydraulic items including piping among them. Many nontechnical issues such as legal or regulatory constraints, bonding limitations, politics, or other issues may influence the design periods. Sewerage systems are designed to handle the anticipated peak load that will develop under anticipated population growth within the subject area for a period of 40–50 years. Appurtenances and mechanical equipment are not expected to last that long, and 15–20-year design periods are usual.

Pumping sewage. Pumping facilities become necessary when it is uneconomical or physically impossible to convey sewage or storm water by gravity to the desired point of discharge. Pumps can be installed in deep basements of buildings to raise the wastes to the level of street sewers, or they can be used to overcome cost and construction problems of deep excavation or topographic impediments like crossing a ridge or bypassing a watercourse. Pumping facilities are sited after evaluating topography, subsurface conditions, flooding hazards, and economic feasibility.

Pumping facilities can be factory-assembled or built in place. They can be either wet-pit (**Fig. 3**) or wet-well/dry-well (**Fig. 4**) types. Wet-pit stations are generally employed for small flows up to 700 gal/min (0.044 m^3/s). They are usually factory-assembled. The wet-pit stations are automated for routine unattended operation, using pneumatic ejectors, suction lift pumps, vertical wet-pit pumps, or submersible pumps that are nonclogging or self-cleaning. Built-in-place stations, using a separate wet well to receive and store the incoming wastes, have centrifugal pumps installed in an adjacent dry well, and typically handle flows in excess of 1000 gal/min (63 m^3/s). Since a pump station is usually installed at a topographic low point, the design must incorporate flood protection and access under all weather conditions. Many states require flood protection for recurrence intervals (return periods) of 25–100 years. *See* PUMP; PUMPING MACHINERY.

Sewage pumps, even though they are fitted with open impellers designed to pass the large solids anticipated in the flow, must still be protected from clogging. Pumps are additionally protected by grit chambers, comminutors, or bar racks (screens), which must be accessible for service and maintenance. Because organic matter is present in wastewater and decomposes with time, malodorous, corrosive, toxic, and hazardous gases frequently develop in pump station wet wells. Proper ventilation, explosion-proof construction, and personnel safety procedures must be adhered to stringently. Pumping facilities are provided with two or more pumps of the size required to handle the peak anticipated inflow when one pumping unit is out of service. Pumps, with electric drive units set above flood level or otherwise protected, should have dual power sources, either from two separate and distinct electric supplies or through the use of a fuel-fed auxiliary system. Pump stations, particularly unattended ones, must be protected against freezing and should be provided with an alarm system to notify appropriate personnel in the event of any type of power interruption or high-water conditions at the facility. *See* SEWAGE TREATMENT.

Examination. Tests and analysis are employed to establish the characteristics of wastewaters such as composition; condition; and physical, chemical, and biological constituents. Water, as used within a community, undergoes only a slight delay in its passage along the hydrologic cycle. In the modern world, water separated from sewage may eventually be directed into a body of water that may be used as a source for a potable water supply. As such, knowledge about what is being returned to

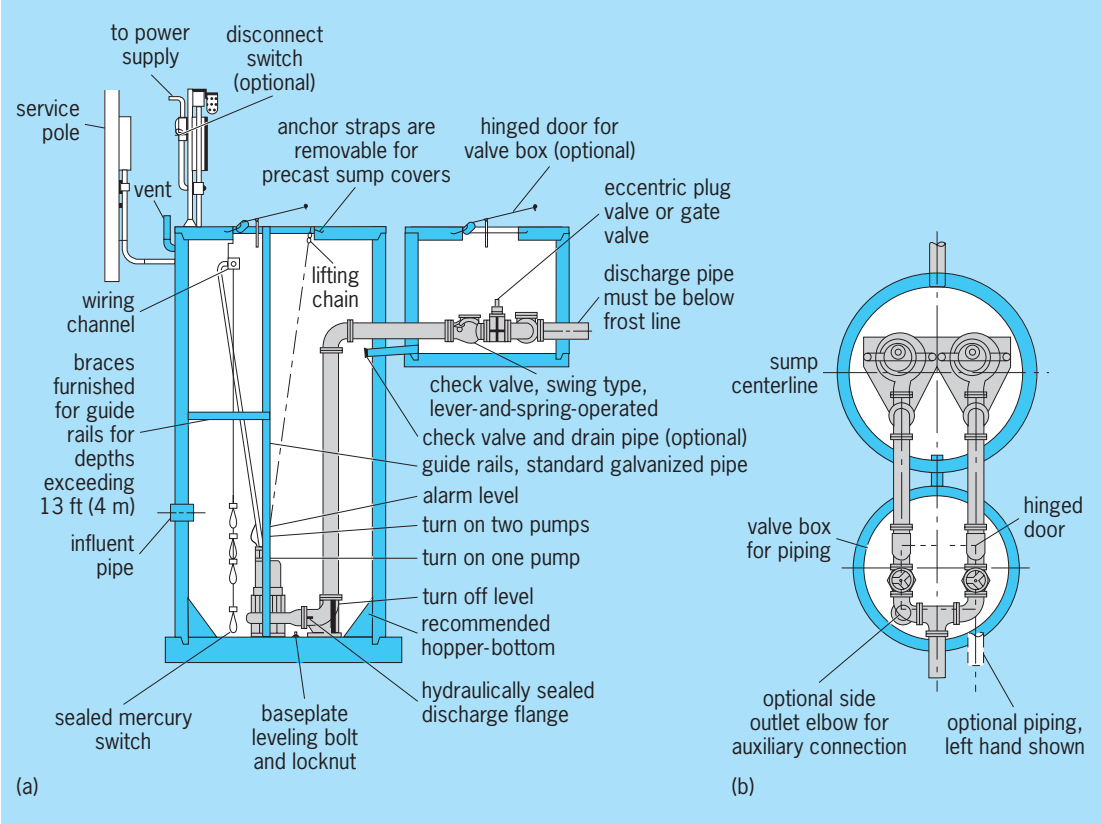

Fig. 3. Wet-pit pumping station with submersible pumps shown (*a*) in section and (*b*) in plan view. (*After Metcalf & Eddy, Inc., and G. Tchobanoglous, Wastewater Engineering: Collection and Pumping of Wastewater, McGraw-Hill, 1981*)

the watercourse and into the immediate environment is essential. Wastewaters contain a myriad of dissolved and suspended materials, including minerals, salts, organic compounds, heavy metals, bacteria, pathogenic organisms, sediments, and toxic substances. It is society's responsibility to return the spent waters to the water-use cycle with safe and acceptable constituents. Tests are used to determine treatability of the waste, amounts of solids and methods to handle these residues, economical amounts of necessary or required additives, and presence of pathogenic organisms, and to establish safe limits for constituents found in effluents that are to be returned to the environment. Technologies and mensuration methods now employed for testing wastes can establish quantities as low as 1 part per trillion (ppt). *See* ANALYTICAL CHEMISTRY; PUBLIC HEALTH.

Results of chemical laboratory tests are usually expressed in mg/liter or g/m^3. For much smaller values, micrograms/liter are used. Since wastewater is more than 99.9% water, mg/liter can be interchanged with parts per million (ppm) and μg/liter with parts per billion (ppb). For solids and sludges, where concentrations are greater than 1 kg/liter, percentages are commonly used. Flows are reported in liters/s, m^3/s, or 10^6 gal/day (mgd). In U.S. Customary units, lb/day can be evaluated from the flow rate in mgd $\times$ dose rate in ppm, or mg/liter $\times$ 8.33 lb/gal.

Since any examination is only as useful and as valid as the sample, planning and care must be exercised in selecting the sampling point, the type and frequency of collection, and the actual manner in which the sample is obtained. Sampling may be a single, random sample, known as a grab sample, or it can be a more representative sample that is made up of a repeated number of collections that are composited by either flow or time. All samples collected must be preserved so that there is no significant change in quality between time of sampling and actual testing.

Tests are characterized as physical, chemical, for solids, bioassay, or bacteriological. The physical tests made for turbidity, color, oils, suspended solids, dissolved oxygen, and temperature relate to the esthetic senses. Sewage may be categorized by its dissolved-oxygen content as fresh, stale, or septic. Fresh sewage has a dissolved oxygen content well in excess of 2 mg/liter, is grayish in color, usually is turbid containing visible solids, but has little or no odor. As organic matter undergoes bacterial decomposition, the dissolved oxygen content drops, color darkens, and a musty odor is noted. When all the dissolved oxygen is used up, sewage becomes dark, odors develop, and the waste is septic. Sewage in this last phase is esthetically unacceptable and also more difficult to treat. *See* STREAM POLLUTION; WATER POLLUTION.

Various tests are employed to determine the type and quantity of organics present. Aerobic organisms present in wastewater require dissolved-oxygen and organic matter for growth. As they multiply, they

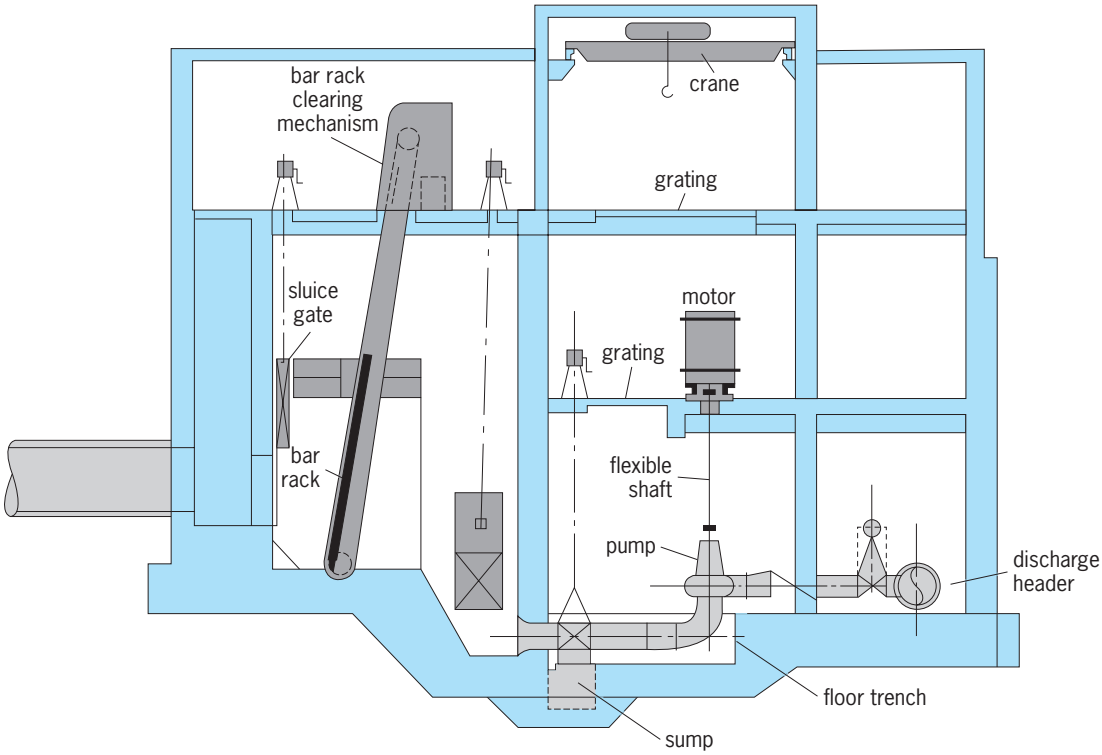

Fig. 4. Section view of conventional wet/dry-well pumping station with protective devices. (*After Metcalf & Eddy, Inc., and G. Tchobanoglous, Wastewater Engineering: Collection and Pumping of Wastewater, McGraw-Hill, 1981*)

use up the dissolved oxygen but stabilize the organic matter and reduce its putrescibility. By measuring the rate at which these bacteria deplete the dissolved-oxygen during 5 days at 68°F (20°C), the organic strength, reported as the biochemical oxygen demand, is determined. The chemical oxygen demand test uses a strong oxidizing agent instead of bacteria. This quicker method indicates that portion of the organic matter that is susceptible to chemical oxidation. Though both tests provide a measure of oxygen required by organics in wastes, they are different processes producing results that may differ widely and not provide good correlation.

In a small waste sample or where there is a low concentration of organic matter, the organics can be determined by the total organic carbon test. In this test, a known quantity of the sample is injected into a specially designed high-temperature retort. The resulting carbon dioxide produced is automatically read by an infrared analyzer. Still another test for organic content is the total oxygen demand test.

Wastewater solids provide important information about the strength, character, and the quantity of residue that will ultimately have to be disposed of. A measured quantity of a waste that is evaporated to dryness and weighed establishes the total residue. Passing another measured quantity of the same waste through a filter and weighing the trapped residue indicates the suspended solids. Subtracting the weight of the suspended solids from the total residue yields the weight of dissolved solids. Settleable solids are determined by placing a liter of the same waste in an Imhoff cone, letting it stay quiescent for an hour, and measuring what has settled out. Once dry weights of the tested waste have been determined, each sample is burned in a muffle furnace. Organic matter is volatilized, and inert matter or ash remains. The weight loss represents the organic component in each constituent portion of the solids found in that waste (**Table 4**). *See* SEWAGE SOLIDS.

Many tests are used to evaluate plants involved in pollution control operations and to define the residuals discharged in the effluent. Tests are conducted for pH, alkalinity, residual chlorine, nitrogen in its various forms, and phosphorus. Bioassay tests permit rapid, realistic interactions between minute quantities of difficult-to-detect toxic components on test organisms in their local environment. Bioassays, using fish as test organisms, establish maximum concentrations of toxics, and also provide the means of establishing required dilutions of effluents, so that they will not have significant impact on the local environment. *See* BIOASSAY; TOXICOLOGY.

The sources of most waterborne pathogenic diseases are human and animal fecal discharges. Sewage contains these discharges, but the pathogenic organisms are both few in number and difficult to test for. A human voids 2–3 $\times$ 10^{11} coliform organisms daily. Since these harmless organisms are hardy, plentiful, and relatively easy to test for, and emanate from the body together with any pathogens, they are used as the index organisms for fecal pollution. Consequently when coliforms are found, it is presumed pathogenic organisms are also present. There are two different bacteriological procedures to determine the number of coliform organisms in a sample. The first method, known as the most probable number, employs the fact that coliform organisms can

TABLE 4. Typical composition of untreated domestic wastewater*

Constituent	Concentration†		
	Strong	Medium	Weak
Solids, total	1200	720	350
Dissolved, total	850	500	250
Fixed	525	300	145
Volatile	325	200	105
Suspended, total	350	220	100
Fixed	75	55	20
Volatile	275	165	80
Settleable solids, ml/liter‡	20	10	5
Biochemical oxygen demand, 5-day, 68°F (20°C)	400	220	110
Total organic carbon (TOC)	290	160	80
Chemical oxygen demand (COD)	1000	500	250
Nitrogen (total as N)	85	40	20
Organic	35	15	8
Free ammonia	50	25	12
Nitrites	0	0	0
Nitrates	0	0	0
Phosphorus (total as P)	15	8	4
Organic	5	3	1
Inorganic	10	5	3
Chlorides§	100	50	30
Alkalinity (as $CaCO_3$)§	200	100	50
Grease	150	100	50

* After Metcalf & Eddy, Inc., and G. Tchobanoglous, *Wastewater Engineering: Treatment, Disposal, Reuse*, 2d ed., McGraw-Hill, 1979.
† All values except settleable solids are expressed in mg/liter; mg/liter = 0.001 oz/ft^3.
‡ ml/liter = 11.5 in.^{3}/ft.^{3}
§ Values should be increased by amount in domestic water supply.

ferment lactose broth and produce gas. The presence or absence of gas, following 24–48 h of incubation at 95°F (35°C) in multiple fermentation tubes of broth that contain inoculations of serial dilutions of a sample, is noted. A statistical estimate of the probable density of coliforms can be established from the data. Results are recorded as the most probable number of coliforms per 100 ml of sample. *See* WATER-BORNE DISEASE.

The second procedure, the membrane filter method, is a direct enumeration method; a known volume of sample is passed through a fine-pore membrane that traps the bacteria present on its surface. The filter is then placed in a special petri dish containing requisite nutrients and incubated at 95°F (35°C) for 24 h. Each trapped coliform organism develops into a colony large enough to be identified and counted directly.

Microscopic examinations, frequently performed as a part of evaluation of plant operations, seek particular organisms within the activated sludge floc or in the zoogloea (a gelatinous mass characteristic of bacteria in fixed films). Microscopic and biologic examinations of receiving waters are made for evaluation of water quality. *See* SEWAGE COLLECTION SYSTEMS.

Gerald Palevsky

Bibliography. American Public Health Association, American Water Works Association, Water Pollution Control Association Joint Committee, *Standard Methods for the Examination of Water and Wastewater*, 20th ed., 1999; American Water Works Association, *Water Quality and Treatment: A Handbook of Public Water Supplies*, 5th ed., 1999; R. A. Corbitt, *Standard Handbook of Environmental Engineering*, 2d ed., 1998; M. J. Hammer, *Water and Wastewater Technology*, 5th ed., 2003; S. R. Qasim, *Wastewater Treatment Plants: Planning, Design, and Operation*, 2d ed., 1998; G. Tchobanoglous, F. L. Burton, and H. D. Stensel (eds.), *Wastewater Engineering: Treatment and Reuse*, 4th ed., 2002; W. Viessman, Jr., and M. J. Hammer, *Water Supply and Pollution Control*, 7th ed., 2004.

Sewage collection systems

Configurations of inlets, catch basins, manholes, pipes, drains, mains, holding basins, pump stations, outfalls, controls, and special devices to move wastewaters from points of collection to discharge. Sewage collection systems must be designed and constructed to function without nuisance to the public, with a minimum of maintenance, with maximum safety for workers, without damage due to corrosion or erosion, with all components sized large enough to be capable of adequately serving the community through a justifiable design period with assured durability, and with an economically acceptable cost. The system of pipes and appurtenances is also known as the sewerage system. Wastewaters may be sanitary sewage, industrial wastes, storm runoff, or combined flows.

A sewer is a constructed ditch or channel designed to carry away liquid-conveyed wastes discharged by houses and towns. Modern sewer systems typically are gravity-flow pipelines installed below the ground surface in streets and following the ground slope. The depth of cover over pipelines is controlled by

factors such as the location of rock and ground water, the ability to receive flows from all buildings by gravity, depth to frost line, economics of maintaining gravity flow as compared with pumping, and location and elevation of other existing utilities and infrastructures.

Sewerage systems are designed to carry the liquid wastes smoothly, without deposition, with a minimum of wasted hydraulic energy, and at minimum costs for excavation and construction; they should provide maximum capacity for future populations and flows. Engineered construction, controlled by availability of time, material, personnel, and finances, affects the choice and use of individual components within sewerage systems.

Components. Sewage collection systems include pipes, joints, manholes, house sewers, street inlets, catch basins, inverted siphons, backwater gates, diversions and regulators, outlets, and pumps.

Pipes. Pipes, house sewers, laterals, trunks, mains, and culverts are constructed of various materials to meet diverse design conditions. Sewerage systems are composed of commercially manufactured pipe sections joined together; in special circumstances they are constructed in place by using concrete or masonry, or the pipe sections are assembled in place from fabricated component pieces. Pipelines are usually circular, but they can also be elliptical, horseshoe, or arched shapes, boxes, or specially fabricated shapes (**Fig. 1**).

Criteria for choice of system piping include friction factors, size and availability of materials, types of joints, crushing strength, ease of installation, resistance to degradation from corrosion and erosion, and total costs of construction. Vitrified clay, concrete, cast iron, fabricated steel, and plastic are the materials most commonly used for pipes. Standards for uniformity, material quality control, size tolerances, guaranteed strength, and jointing conditions are established by standards organizations and the federal government, as well as industrial, trade, and professional organizations.

Different materials are manufactured in specified diameters, in various standard lengths, and with defined types of ends or joints. The shorter the length of pipe, the more joints are necessary. This can lead to higher labor costs, increased possibility of misalignment, leakage, chance of failure, and lack of durability. Increasing pipe length increases section weight, making pieces heavy, unwieldy, and difficult to place and join in a trench, thus requiring special equipment and, often, more personnel. The choice of pipe frequently depends upon economics, and it may be controlled by construction conditions.

Vitrified clay, generally 4–36 in. (10–91 cm) in diameter, is nonabsorptive and resistant to corrosion, erosion, and scour. It is available in lengths of 1–8 ft (0.3–2.4 m), and it varies in strength with diameter and category (**Table 1**). Clay pipes with bell-and-spigot joints formerly were used in traditional house drains and street sewer lines. *See* CLAY, COMMERCIAL.

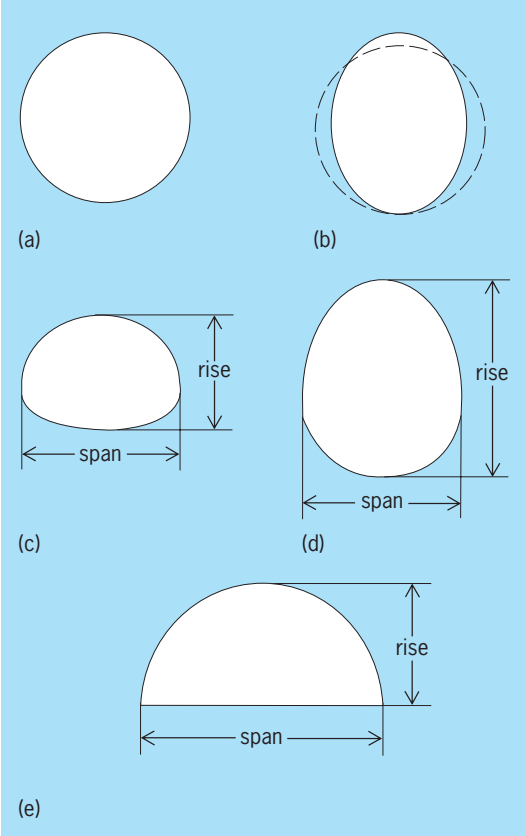

Fig. 1. Various pipe shapes. (*a*) Round. (*b*) Vertically elongated (ellipse). (*c*) Pipe-arch. (*d*) Underpass. (*e*) Arch. (*After American Iron and Steel Institute, Handbook of Steel Drainage and Highway Construction Products, 2d ed., 1971*)

Cast-iron soil pipe is a bell-and-spigot pipe of limited strength. It is used in plumbing work to carry sanitary wastes from water closets and other fixtures to the street sewer. Soil pipe is often used for house sewers. *See* CAST IRON.

Concrete pipe is manufactured in both circular and elliptical shapes that are 4 in. to 12 ft (10 cm to 3.6 m) in diameter, either plain or reinforced. Various methods, including spinning, casting, and packing, are used in its manufacture. The laying lengths and strength of concrete pipe can be varied and specifically controlled by placing different styles, configurations, and thicknesses of metal reinforcement, as well as by adjusting wall thickness (Table 1). Concrete is subject to corrosion from acids and hydrogen sulfide, and also possibly to scour or erosion where there is excessive velocity and heavy sediment load. *See* CONCRETE; REINFORCED CONCRETE.

Cast-iron pipe and ductile iron pipe 2–48 in. (5–122 cm) in diameter and with lengths to 20 ft (6 m) and various types of joints provide very tight connections, flexibility, and greater strength than concrete. These pipes, as well as those made of welded steel, are most often used for pipe crossings, in areas of high water, for special load conditions, and for pressure lines. Linings and wrappings are applied inside and out to protect these metal pipes from corrosion. *See* CORROSION.

TABLE 1. Crushing strength of clay and concrete pipe by the three-edge bearing method, lb/linear ft*

Internal diameter, in.†	Clay		Nonreinforced concrete			Reinforced concrete, ultimate strength				
	Standard	Extra strength	Class I	Class II	Class III	Class I	Class II	Class III	Class IV	Class V
3	—	2000								
4	1200	2000	1500	2000	2400					
6	1200	2000	1500	2000	2400					
8	1400	2200	1500	2000	2400					
10	1600	2400	1600	2000	2400					
12	1800	2600	1800	2250	2600	—	1500	2000	3000	3750
15	2000	2900	2000	2600	2900	—	1875	2500	3750	4690
18	2200	3300	2200	3000	3300	—	2250	3000	4500	5620
21	2400	3850	2400	3300	3850	—	2625	3500	5250	6560
24	2600	4400	2600	3600	4400	—	3000	4000	6000	7500
27	2800	4700	2800	3950	4600	—	3375	4500	6750	8440
30	3300	5000	3000	4300	4750	—	3750	5000	7500	9380
33	3600	5500	3150	4400	4875	—	4125	5500	8250	10,220
36	4000	6000	3300	4500	5000	—	4500	6000	9000	11,250
39	—	6600	—	—	—	—	4825	6500	9750	—
42	—	7000	—	—	—	—	5250	7000	10,500	13,120
48	—	—	—	—	—	—	6000	8000	12,000	15,000
54	—	—	—	—	—	—	6750	9000	13,500	16,880
60	—	—	—	—	—	6000	7500	10,000	15,000	18,750
66	—	—	—	—	—	6600	8250	11,000	16,500	20,620
72	—	—	—	—	—	7200	9000	12,000	18,000	22,500
—	—	—	—	—	—	—	—	—	—	—
—	—	—	—	—	—	—	—	—	—	—
108	—	—	—	—	—	10,800	13,500	18,000	27,000	33,750

*1 lb/linear ft = 0.0146 kN/m.
†1 in. = 2.54 cm.
SOURCE: *Annual Book of ASTM Standards*, American Society for Testing and Materials, 1977.

Fabricated steel pipe is constructed of lightweight corrugated flat sheets that provide strength and flexibility; they can be fabricated into a variety of cross-sectional shapes. Circular, elliptical, partial sections, arches, and pipe arches can be either fabricated in the shop or erected in the field in sizes up to 25 ft (7.5 m). As the corrugations increase in depth for increased strength and size opening, hydraulic coefficients become lower and flow characteristics change. Corrugated pipes are frequently coated and paved to fill the corrugations in order to increase flow efficiency and reduce erosion and corrosion problems.

Plastic pipe of poly(vinyl chloride) (PVC), polypropylene, and acrylonitrile-butadiene-styrene resin (ABS), plastic truss pipe, and other novel materials are being used in sewerage systems. These pipes are very light, are resistant to corrosion, and are available in extended lengths. *See* POLYMER.

Joints. The joints or junctions between two adjacent pipes or between a pipe and an appurtenance must be watertight, possess a degree of flexibility, and have a long life expectancy. Most joints utilize elastomeric O rings that compress between two adjacent pipes to form a tight, yet somewhat flexible, seal. Mortar or grout is used to fill openings or spaces in order to prevent roots from growing into the joints or to prevent soil constituents from interacting with the sealing rings. Corrugated pipe typically uses a band coupling that compresses the O rings set in the corrugated grooves as bolts draw the band tight to the outside of the pipe (**Fig. 2**).

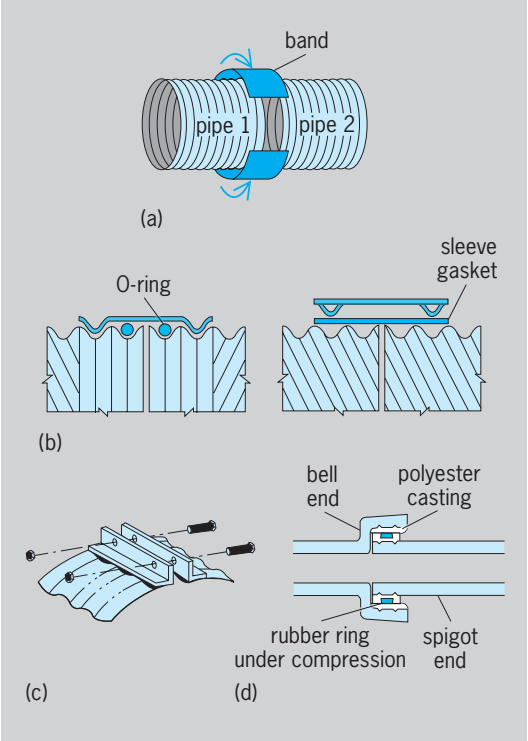

Fig. 2. Band-coupled pipe joints. (*a*) Typical band coupling; (*b*) standard corrugated steel pipe band connectors; (*c*) band angle connector (*after American Iron and Steel Institute, Modern Sewer Design, 1980*). (*d*) Details of an O-ring compression joint for clay, cast iron, and concrete pipes.

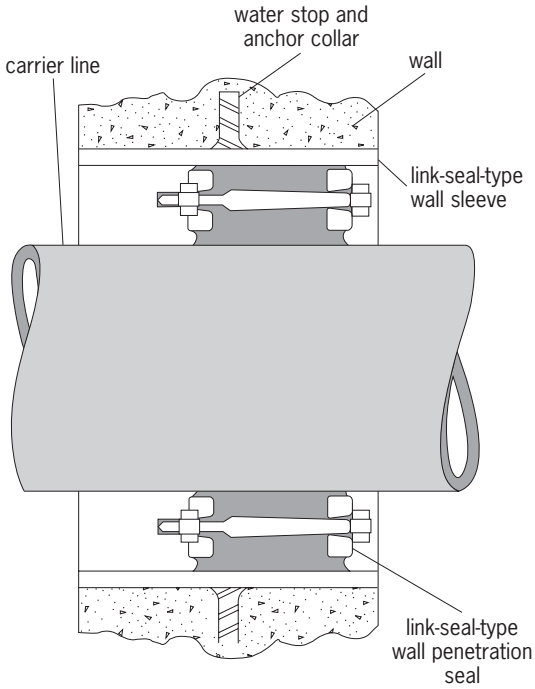

Fig. 3. Diagram of a mechanical compression water stop and seal. (*Link Seal, Thunderline Corp., Wayne, Michigan*)

Appurtenances and structures that permit connection to, entry into, or maintenance of sewerage system piping require seals at openings where pipes penetrate these structures. The openings can be grouted closed, or a mechanical compression type of closure may be used (**Fig. 3**).

Manholes. Also known as access holes, manholes are openings from the ground surface that are fitted with removable covers and permit personnel to enter, inspect, maintain, and repair the gravity flow system (**Fig. 4**). Manholes are placed at ends of lines and wherever a pipe must be changed in size, elevation, direction, or slope. On straight runs, for up to 24-in. (600-mm) diameter, they are placed no more than 350 ft (100 m) apart to permit mechanical cleaning of the intervening pipe. Spacing increases with diameter. For sewer lines 8–24 in. (20–61 cm) in diameter, manholes 4 ft (1.2 m) in diameter sit directly over the pipe, with the steps of the manhole set into the vertical barrel side. For pipes larger than 24 in. (61 cm), manholes may be offset to the side of the pipe; or larger structures known as junction chambers may be used.

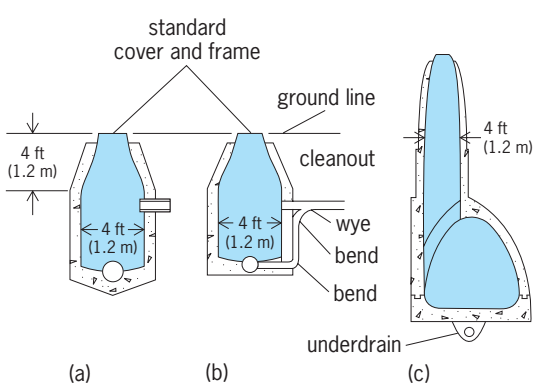

Fig. 4. Some typical sewer manholes with standard 24-in. (61-cm) cover and frame. (*a*) Inside drop. (*b*) Outside drop. (*c*) Unit for a large sewer. (*After R. L. Linsley and J. B. Franzini, Water Resources Engineering, 3d ed., McGraw-Hill, 1979*)

The channel within the bottom of the manhole should provide a smooth passage for the flow entering and leaving the structure. The paved invert in the manhole permits transitions from a smaller pipe entry to a larger pipe leaving, and it offers the means to change direction by developing a paved curve. Where the entering line is considerably above the existing pipe, a drop pipe is constructed to permit a smooth transition and yet maintain a clear straight line for ease of cleaning (**Fig. 5**).

Manhole covers, usually in roadway traffic lanes, are generally made of cast iron, set in fabricated frames, and circular in shape so that they cannot fall through the opening when they are lifted for service entry. Covers must lie flat with the pavement, seat well to prevent rattle or movement, be capable of supporting typical traffic loads without damage, and prevent storm water inflow. Covers are frequently provided with special locking devices to prevent removal, vandalism, unauthorized entry, or illegal dumping of wastes into the system.

House sewers. House or building sewers, connections, or service laterals are the small-diameter sewer lines that connect the individual residence or property with the street sewer or lateral (**Fig. 6**). House connections must be made with tight joints, and care must be taken to support the pipe adequately so that subsequent construction does not break the pipe, open joints, or develop pockets in the line. Unused stubs, the openings that have been provided in street sewers for future house sewer connections, must be tightly plugged to control infiltration. Special appurtenances such as flushing manholes, terminal cleanouts, relief overflows, lamp holes, and bypasses are installed in sewerage systems to overcome special topographic or design conditions.

Street inlets. Street inlets are structures that intercept, collect, and transfer storm runoff from street surfaces to the underground storm drain system. Street inlets are located along paved streets and are used to keep pedestrian walkways and crossways relatively free of flowing storm runoff. There are three types: gutter, curb or combined. A gutter inlet has a flat grate at the paved surface that allows flowing storm water to fall through to the inlet structure below. A curb inlet has a vertical opening in the curb and is depressed below the gutter line, allowing water to flow and drop laterally into the inlet structure. A combined inlet has both a gutter and curb opening with the two aligned over the structure. Special designs for areas of very heavy gutter flows, or where gutter flow velocities are very high, use multiple openings, deflector inlets with notches, grooves, or ridges. The outlet line from the inlet to the storm drainline manhole drains the entire structure (**Fig. 7**).

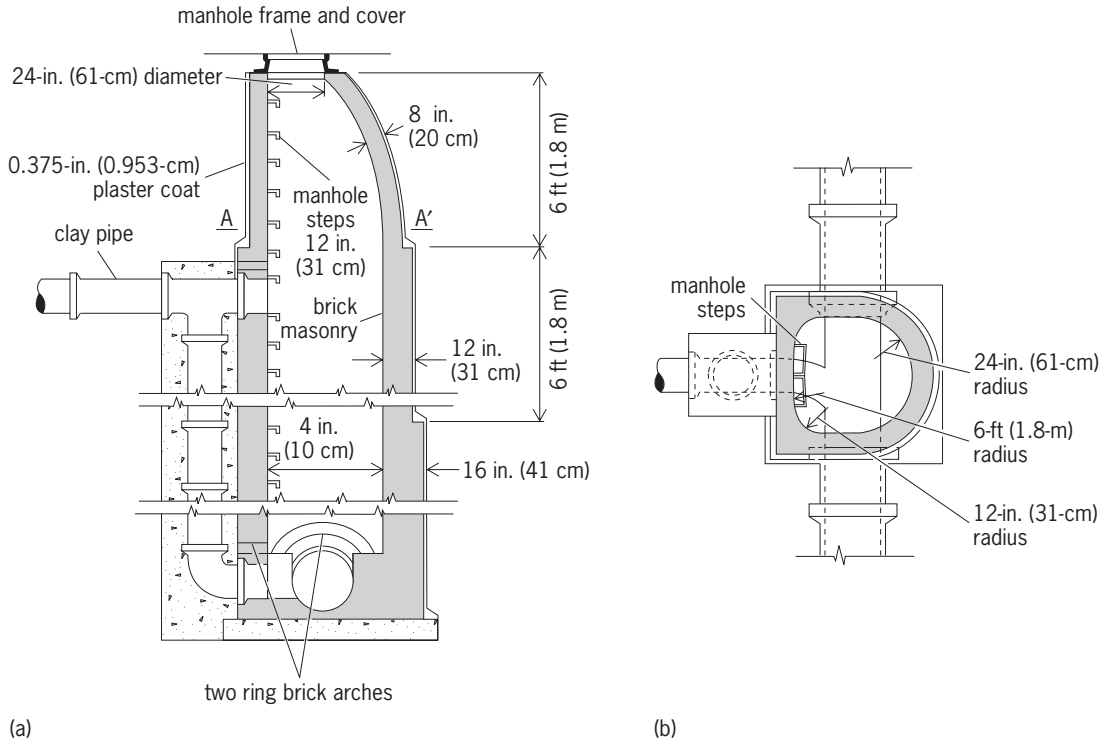

Fig. 5. Drop manhole detail. (*a*) Vertical section. (*b*) Horizontal section (A-A in part a). (*After National Clay Pipe Institute, Clay Pipe Engineering Manual, 1962*)

Catch basins. At one time, catch basins were an integral part of combined systems. They existed for storm drainage systems when streets were poorly paved or not maintained and drain lines were not constructed to maintain self-cleaning velocities. Catch basins are constructed with a sump below the outfall line to trap all grit, sand, street sweepings, and other debris within the structure. Catch basins have hatches, covers, or easily removable gutter grates to permit access for periodic maintenance and cleaning of accumulated solids. Special hoods or overflow designs are required on combined systems to prevent sewer gases from venting into the street (**Fig. 8**).

Inverted siphons. An inverted siphon is a depressed or suppressed sewer. The pipe enters the upper end chamber and usually leaves in a number of smaller lines. The multiple lines pass under the obstruction, structure, or stream with the outlet chamber being at a lower invert elevation so that inflow enters through the inlet and flows out the outlet structure. The section of pipe that is below the hydraulic grade line is always full of liquid, even when there is no flow into

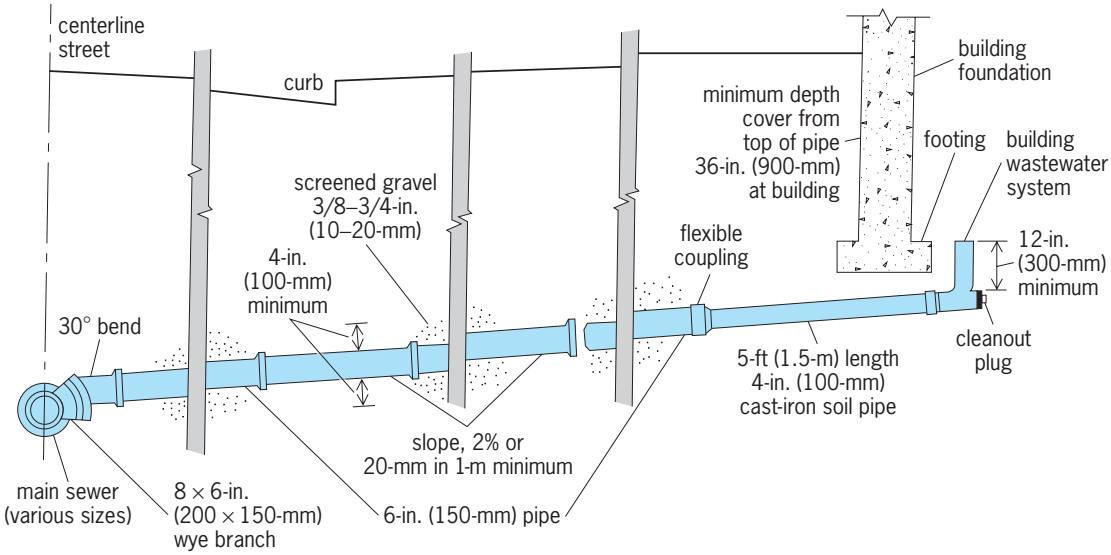

Fig. 6. Typical house sewer connection. (*After Metcalf & Eddy, Inc., and G. Tchobanoglous, Wastewater Engineering: Collection and Pumping of Wastewater, McGraw-Hill, 1981*)

the line. The multiple pipes of the inverted siphon are arranged so that as increased flow enters the inlet structure, successive lines are pressed into use, thus maintaining self-cleaning velocities above 3.5 ft/s (1.0 m/s) to prevent clogging.

Backwater gates. Backwater gates are devices that function like check valves to provide unidirectional flow and prevent the reversal of water flow in a line when the water elevation in the receiving water body rises above the pipeline's hydraulic gradient. They are also known as backflow valves, tide gates, or flap valves.

Diversions and regulators. These structures bypass, divert, or dump excess combined flow through use of weirs, of relief siphons, or by float, electrical, or pressure-operated valves and gates. Treatment plants are designed for flows up to three times the volumes of sanitary flows that occur in dry weather but they cannot cope with the huge volumes of water delivered by combined sewers during storms. Excess storm flow, above the design flow that is to continue to the treatment plant, is diverted by these structures to relief sewers, to overflows to nearby watercourses, or to holding facilities for later treatment.

Outlets. At the end of any sewerage system—be it from storm water drains, diversion or regulator outlet line, an overflow, or the end of the treatment facility—the treated or untreated flows must be discharged to the environment without causing nuisance, offense, or damage. Outlet structures vary from a single pipe retained by a straight headwall, to flared pipes, structures with special wingwalls, riprapped aprons, or specially constructed devices to dissipate hydraulic energy and prevent erosional damage as water discharges from an outlet and spills over the surface of the land. When the outlet flows are large and velocities high, special devices are used to dissipate energy. They include hydraulic jumps, hydraulic buckets, side channel discharges, and other structures designed specifically for energy dissipation.

Treatment plant outfalls usually discharge below the surface of the receiving water body. Discharges from outfall pipes are generally diffused through a series of manifolded lines fitted with multiple ports. Lines are placed deep enough to allow for good mixing and dilution, prevent so-called water boils, and maintain sufficiently low exit velocities to preclude erosion. They are anchored to prevent undermining or movement by currents or tides, and they are placed in locations where they will not interfere with normal use of the receiving body of water. *See* EROSION; HYDRAULIC JUMP; HYDRAULICS.

Pumps. Pumping facilities must be included in sewage collection systems when topography or soil conditions preclude typical piping layouts or construction, and where normal gravity flow cannot be employed.

Sewer system design. Sewer design is composed of several phases: initial, preliminary, and final. The initial conceptual phase defines the specific problem with generalized solutions, based on population, area, and financing. Conceptual design addresses broad issues of politics, public policy, probable costs, design period, and phasing of construction. An environmental assessment is necessary. In addition, evaluations should be made of bonded indebtedness, how bonds will be issued, availability of state and

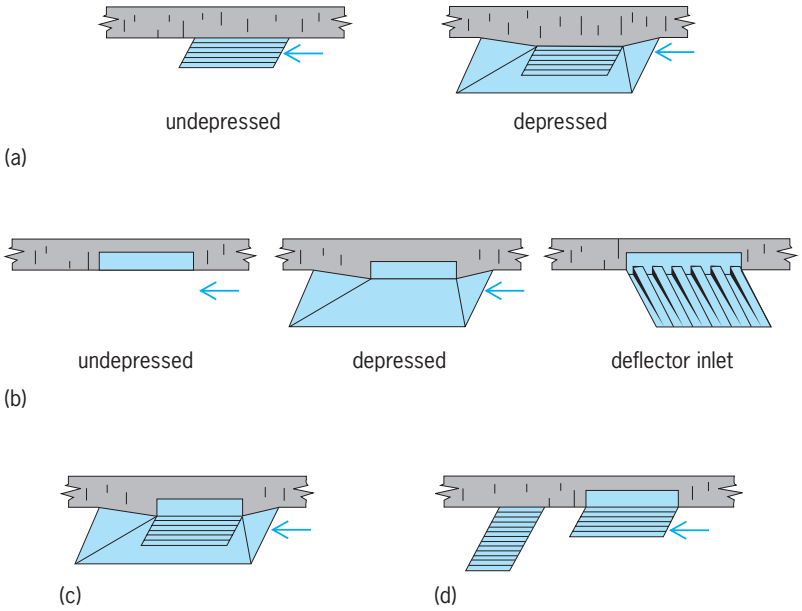

Fig. 7. Typical stormwater street inlets. (*a*) Gutter. (*b*) Curb. (*c*) Combination; grate placed directly in front of curb opening and depressed. (*d*) Multiple; undepressed. (*After Joint Committee of American Society of Civil Engineers and Water Pollution Control Federation, Design and Construction of Sanitary and Storm Sewers, ASCE MOP 37, WPCF MOP 9, 1969*)

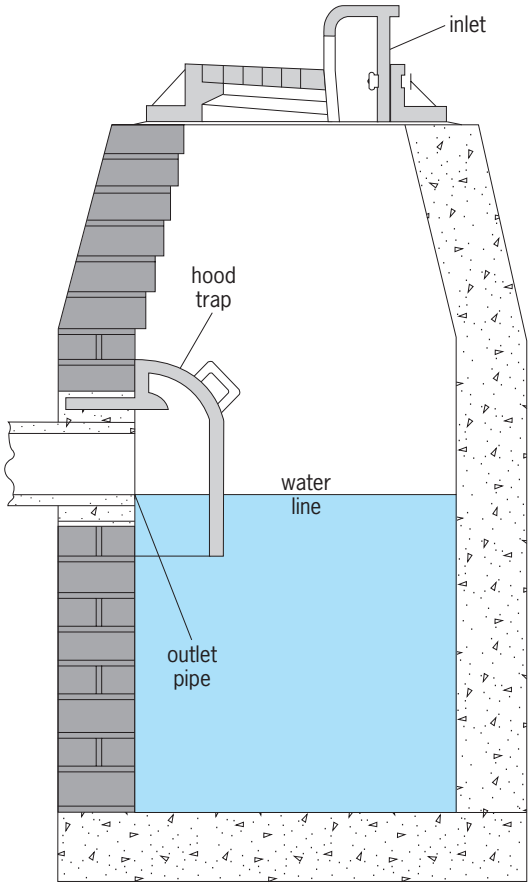

Fig. 8. Typical catch basin with removable hood trap for cleanout. (*Neenah Foundry Co., Neenah, Wisconsin*)

federal assistance programs, and preliminary determination regarding tax structures and rates which may be set by the public service commission.

Conceptual design is comprehensive and is based on physical, demographic, economic, and legal information. Projected population distribution and density, dependent upon zoning regulations and economic development within a political subdivision, must be investigated. Topography, rainfall, flood information, soil conditions, ground-water elevations, extent of existing road systems and paving, extent of existing and future infrastructure, and locations for collection, treatment, and outfall must be established. Environmental regulations, regional or multimunicipal systems, intra- and interstate compacts and legal requirements of flood control, flood plain insurance, and final disposal of liquid and solid effluents must all be resolved. *See* ENVIRONMENTAL ENGINEERING; LAND-USE PLANNING.

The preliminary design phase uses reliable data to permit development of realistic evaluations of the proposed project. The area to be served and the initial and future quantities of flow to be handled are established, and treatment and discharge points are determined. Tentative pipeline runs are prepared by using topographic maps and aerial photographs to locate major concentrations of buildings, roads, possible easements, streams, and special features. Preliminary sewer lines and grades are approximated, and appurtenances, diversions, pump stations, overflows, treatment plants, and outfalls spotted. Preliminary costs based on projected lengths of pipe, number of manholes, special structures, appurtenances, and approximate excavation are estimated. *See* AERIAL PHOTOGRAPH; TOPOGRAPHIC SURVEYING AND MAPPING.

Final design is based on refined information of the topography and soil that has been obtained by field surveys and borings. Design flows are established by detailed evaluation of areas and population projections based on demographics. Pipelines and appurtenances are located, sized, and drawn on contract plans. Specifications for construction are prepared, and cost estimates for the designed system are made. *See* SURVEYING.

Final construction plans for a sewerage system contain both plan and profile sheets (**Fig. 9**). Plan sheets provide graphical information about the location, dimensions, and clearances of all proposed pipes and appurtenances with respect to topography, legal grades, existing and proposed infrastructure, water mains, overhead power lines,

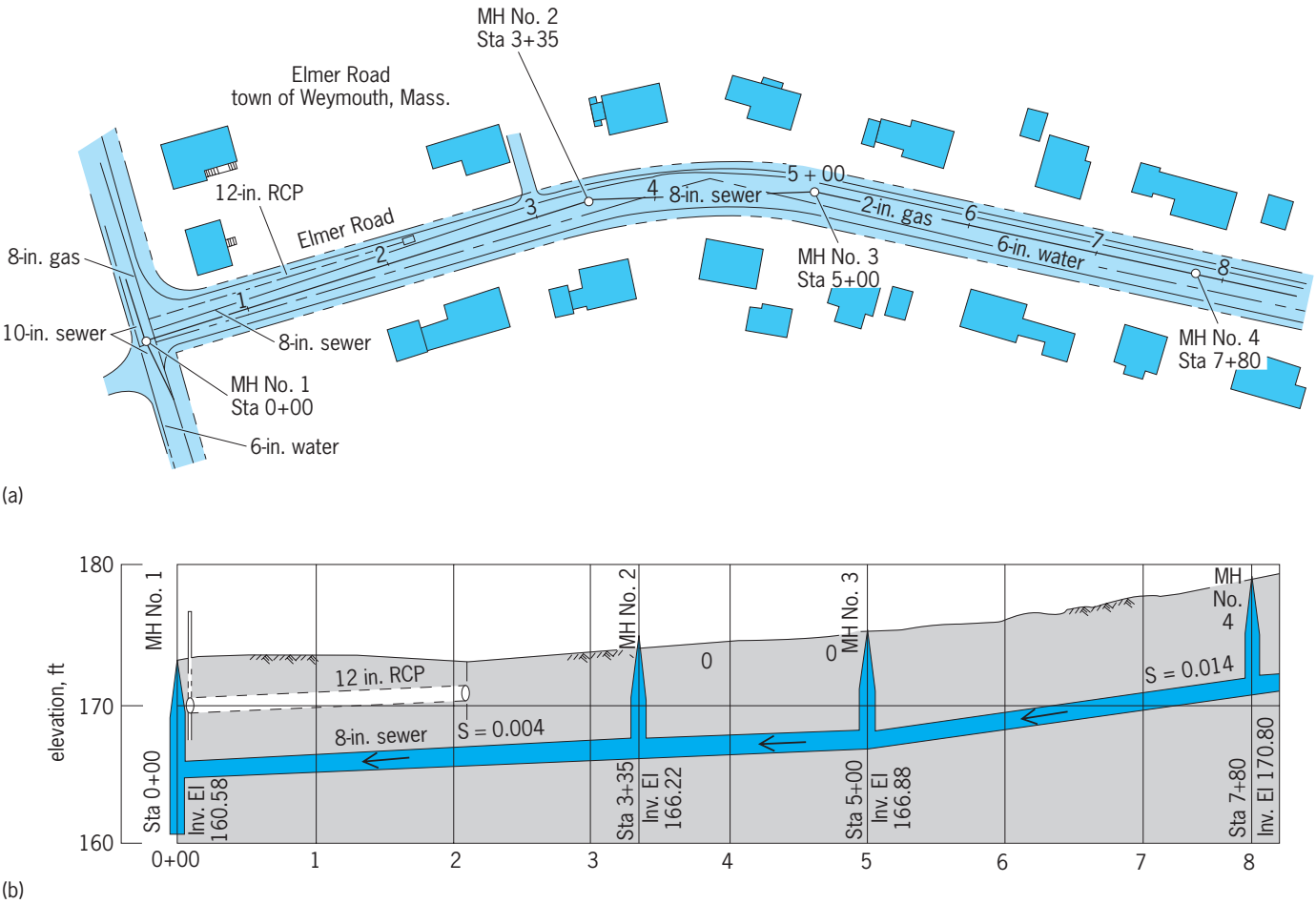

Fig. 9. Typical contract plan and profile sheet for sewer lateral. (*a*) Plan. (*b*) Profile; numbers on horizontal axis refer to sites with same numbers in part *a*. MH = manhole. Sta = Station. Inv. El. = Invert elevation. RCP = reinforced concrete pipe. S = slope in ft/ft or m/m. 1 ft = 0. 3 m. 1 in. = 2.54 cm. (*After Metcalf & Eddy, Inc., and G. Tchobanoglous, Wastewater Engineering: Collection and Pumping of Wastewater, McGraw-Hill, 1981*)

trees, and special situations. The profile sheet, a developmental plan along the centerline of the proposed sewerage system, with stationing and references to the plan sheet, clearly shows initial and final ground elevations; diameter, material, length, and slope of pipes; location of and connections to appurtenances; all control invert elevations; all known belowground infrastructure crossings; clearances and cover over proposed pipe; and special bedding conditions under the pipe.

Additional sheets showing typical details, standard castings, construction limits, trench conditions, and pipe bedding, together with designs of special structures and appurtenances for erosion and sediment control during construction, complete the contract drawings. Specifications defining the quality of material, workmanship of construction, esthetics of finished work, testing of system, and guarantees are an integral portion of the public works contract documents.

The flows established in the design of a sanitary sewer system are dependent upon population. Storm drainage systems are based on rainfall-runoff values coupled with risk evaluation of flood damage. Combined systems are more complex, since they are designed on values derived from both of the above. Combined systems must carry dry weather flows without deposition, must carry limited storm water flow without surcharging treatment facilities, and must provide for diversion and overflows when heavy storm flows develop.

Pipe flow. Liquid flows from a point of higher energy to a point of lower energy. Energy at a point is made up of position, pressure, and kinetics. Position is elevation above a measured datum; pressure is due to applied external force; and kinetic energy is developed from the velocity of the moving liquid. The difference in energy levels between an upstream and a downstream point is known as the available head.

Within the normal range of sizes used, a pipe section between manholes is a gravity sewer that is an open-channel system. The pipe section between manholes is of a constant size, placed on a continuous slope, with the liquid quantity between manholes assumed constant. Under these conditions, since there is no external pressure, uniform flow exists, and velocity is constant.

The difference in physical elevation between two points along the bottom of the pipe (difference in invert elevation), divided by the horizontal distance between the two points (reach), is the invert slope. Since quantity of flow and pipe size are constant, the depth of flow in the pipe (normal depth) is constant, and the water surface slope, or hydraulic gradient, is parallel to the invert slope. Because the liquid is moving, it develops kinetic energy that is known as the velocity head and is given by Eq. (1), where h_v is

$$h_v = \frac{V^2}{2g} \tag{1}$$

the velocity head in ft or m, V is velocity in ft/s or m/s, and g is the gravitational acceleration constant. The energy gradient is the theoretical line depicting the total energy in the fluid at every point along the pipe. Under uniform flow, the energy gradient, the hydraulic gradient, and the bottom of the pipeline are all parallel, and their slopes are identical. **Figure 10**

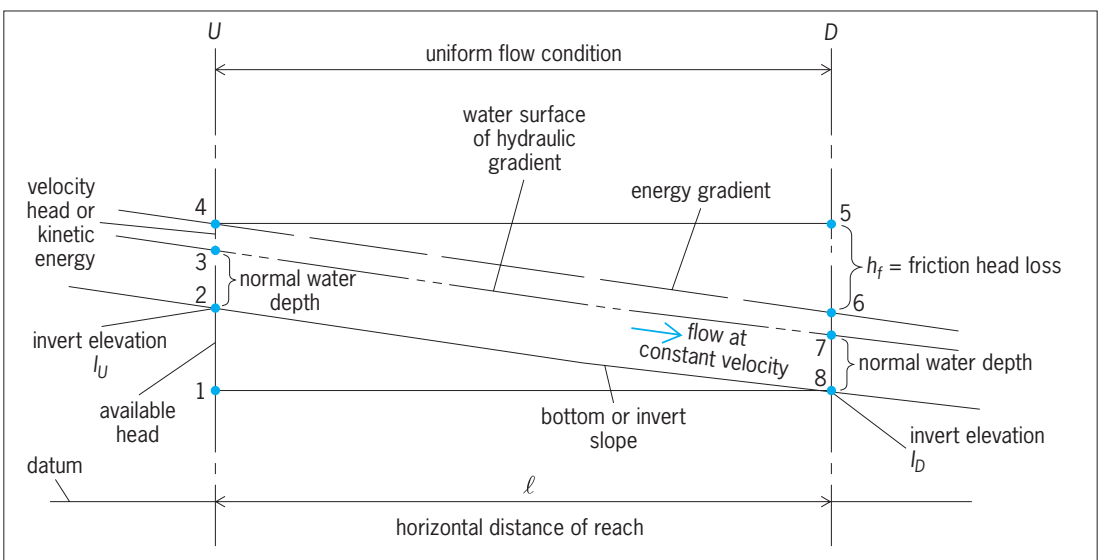

- 1-8 is parallel to 4-5 and is parallel to datum
- 1-2 = 5-6 available head at U = head loss at D
- invert slope = (1-2)/ℓ
- energy gradient = h_f/ℓ
- 3-4 = 6-7 = kinetic head = $V^2/2g$ = velocity head
- energy gradient is parallel to hydraulic gradient is parallel to invert slope
- total energy at U = datum to point 4; potential energy = datum to point 3

Fig. 10. Hydraulic profile for uniform flow condition, showing values at points *U* and *D*.

TABLE 2. Coefficients of roughness values of *n* for use in the Manning equation

Surface	Best	Good	Fair	Poor
Uncoated cast-iron pipe	0.012	0.013	0.014	0.015
Coated cast-iron pipe	0.011	0.012*	0.013*	
Commercial wrought-iron pipe, black	0.012	0.013	0.014	0.015
Commercial wrought-iron pipe, galvanized	0.013	0.014	0.015	0.017
Smooth brass and glass pipe	0.009	0.010	0.011	0.013
Smooth lockbar and welded O.D. pipe	0.010	0.011*	0.013*	
Riveted and spiral steel pipe	0.013	0.015*	0.017*	
Vitrified sewer pipe	{0.010, 0.011}	0.013*	0.015	0.017
Common clay drainage tile	0.011	0.012*	0.014*	0.017
Glazed brickwork	0.011	0.012	0.013*	0.015
Brick in cement mortar; brick sewers	0.012	0.013	0.015*	0.017
Neat cement surfaces	0.010	0.011	0.012	0.013
Cement mortar surfaces	0.011	0.012	0.013*	0.015
Concrete pipe	0.012	0.013	0.015*	0.016
Wood stave pipe	0.010	0.011	0.012	0.013
Plank flumes				
Planed	0.010	0.012*	0.013	0.014
Unplaned	0.011	0.013*	0.014	0.015
With battens	0.012	0.015*	0.016	
Concrete-lined channels	0.012	0.014*	0.016*	0.018
Cement-rubble surface	0.017	0.020	0.025	0.030
Dry-rubble surface	0.025	0.030	0.033	0.035
Dressed-ashlar surface	0.013	0.014	0.015	0.017
Semicircular metal flumes, smooth	0.011	0.012	0.013	0.015
Semicircular metal flumes, corrugated	0.0225	0.025	0.0275	0.030
Canals and ditches				
Earth, straight and uniform	0.017	0.020	0.0225*	0.025
Rock cuts, smooth and uniform	0.025	0.030	0.033*	0.035
Rocks cuts, jagged and irregular	0.035	0.040	0.045	
Winding sluggish canals	0.0225	0.025*	0.0275	0.030
Dredged-earth channels	0.025	0.0275*	0.030	0.033
Canals with rough stony beds, weeds on earth banks	0.025	0.030	0.035*	0.040
Earth bottom, rubble sides	0.028	0.030*	0.033*	0.035
Natural-stream channels				
1. Clean, straight bank, full stage, no rifts or deep pools	0.025	0.0275	0.030	0.033
2. Same as 1, but some weeds and stones	0.030	0.033	0.035	0.040
3. Winding, some pools and shoals, clean	0.033	0.035	0.040	0.045
4. Same as 3, lower stages, more ineffective slope and sections	0.040	0.045	0.050	0.055
5. Same as 3, some weeds and stones	0.035	0.040	0.045	0.050
6. Same as 4, stony sections	0.045	0.050	0.055	0.060
7. Sluggish river reaches, rather weedy or with very deep pools	0.050	0.060	0.070	0.080
8. Very weedy reaches	0.075	0.100	0.125	0.150

*Values commonly used in designing.
SOURCE: Metcalf & Eddy, Inc., and G. Tchobanoglous, *Wastewater Engineering: Collection and Pumping of Wastewater,* McGraw-Hill, 1981.

shows a typical hydraulic profile. *See* FLUID-FLOW PRINCIPLES; FLUID MECHANICS; OPEN CHANNEL.

Energy is required to move a fluid from one point to another. This energy is used to overcome the resistance or friction of the liquid along the surface of the conduit carrying the flow. The rougher the surface texture, and the more joints or imperfections in alignment, the greater will be the resistance to flow. Energy for flow requires available head, which is the change in elevation, or slope of the energy gradient from manhole to manhole.

An equation used to compute the velocity of uniform flow in an open channel is known as the Manning equation. It provides the formula that is used the most for sewerage system design. The Manning equation (2) relates velocity within a pipe or open

$$V = \frac{1.486}{n} \times R^{2/3} \times S^{1/2} \qquad (2a)$$

$$V = \frac{1}{n} \times R^{2/3} \times S^{1/2} \qquad (2b)$$

channel to slope, to a roughness coefficient that is a measure of the surface friction, and to the hydraulic radius. Here V is velocity in ft/s [Eq. (2*a*)] or m/s [Eq. (2*b*)]; n is the coefficient of roughness and is dependent upon the material and finish of the open channel (**Table 2**); R is the hydraulic radius, obtained by dividing the cross-sectional area of water by the wetted perimeter in the open-channel flow zone (R in pipes flowing full is pipe diameter ÷ 4); and S is the slope in ft/ft [Eq. (2*a*)] or m/m [Eq. (2*b*)].

Sewerage systems should be designed with a velocity high enough to prevent deposition of solids, but not so high that the flow causes erosion or excess thrust or movement of pipe or appurtenances (**Table 3**). The basic Manning equation is developed for pipes flowing full; design charts and nomographs are used for sizing pipes. Many regulatory agencies limit maximum flow within pipes to some proportion of full pipe capacity as a safety factor. In addition, inflow to sewerage systems is not a constant; it varies with time and season for volume, depth, and velocity. Since the hydraulic radius R is dependent on cross-sectional area and wetted perimeter, and these do not change in a constant or uniform manner, actual

TABLE 3. Minimum slope for gravity flow in sanitary sewers

Size		Slope, m/m*	
mm	in.	$n = 0.013$	$n = 0.015$
200	8	0.0033	0.0044
250	10	0.0025	0.0033
300	12	0.0019	0.0026
375	15	0.0014	0.0019
450	18	0.0011	0.0015
525	21	0.0009	0.0012
600	24	0.0008	0.0010
675	27	0.0007†	0.0009
750	30	0.0006†	0.0008†
900	36	0.0004†	0.0006†

*Based on Manning's equation with a minimum velocity of 0.6 m/s. Where practicable, steeper slopes should be used. 1 m = 3.2808 ft.
†The minimum practicable slope for construction is about 0.0008 m/m.
SOURCE: Metcalf & Eddy, Inc., and G. Tchobanoglous, *Wastewater Engineering: Collection and Pumping of Wastewater,* McGraw-Hill, 1981.

flows within the pipes must be determined for these different conditions. Various graphs for determining hydraulic elements for partial flow are available (**Fig. 11**); they are used to determine depths of flow and velocities in sections having different shapes.

Good design of a gravity sewer line attempts to keep the slope of the energy gradient as flat as possible in order to minimize excavation and prevent loss of available energy, which cannot be recovered. Proper calculations for manhole losses must be considered, since wherever pipes change slope, direction, or size within a manhole energy losses occur. A general rule of sewerage system construction has been that the soffit (inner top) of the outflow line from a manhole should be at least 0.1 ft (3 cm) lower than the soffits of all lines entering the manhole. This is done to prevent the energy gradient from sloping upward or against the direction of flow, a condition that produces deposition of solids.

When pipe slope is significantly flatter than the overlying ground, insufficient cover over pipes may develop. This is overcome by the use of drop manholes (Figs. 4 and 5). A localized energy loss is sustained as the energy gradient is made rapidly discontinuous. Special designs for inside and outside drop manholes may become necessary to absorb excess turbulence and energy, direct the flows, and prevent scour or erosion from the falling water. Where the pipe is installed on steep grades, with steep slopes and high velocities, special collars, anchorages, thrust blocks, and joint restraints are designed and provided to prevent movement or undermining of the pipe. *See* PIPE FLOW; PIPELINE.

Pipe installation. Considerations in installing pipe include trenches, bedding, maintenance, and safety.

Trenches. Sewerage system pipelines must be installed accurately, to proper line and grade. Construction methods for normal cut and cover, or open-trench excavation, are similar to other types of excavation for general construction. Since sewerage lines are most frequently installed under pavements within the rights of way for roadways, adjacent to existing buildings, with overhead and underground utilities, special care must be exercised to protect work and workers from traffic, cave-ins, flooding, and electrical shock. Professional construction practices and regulations of the Occupational Health and Safety Administration (OSHA) must be followed for shoring, bracing, and dewatering excavations. Spoil, the excavated material from the trench, should be placed on the uphill side to prevent surface water from flowing into the trench; yet it should be placed far enough away from the trench so that it does not add weight and cause a cave-in. If the spoil is of satisfactory material, it may be used for backfilling the trench after the pipe is bedded in place. Trench widths are kept to a minimum to reduce transmitted load to the pipe. If the trench is too narrow, workers cannot work within it, make proper joints, and backfill around the pipe safely. Sheeting, shoring, and bracing, used to maintain stability of the trench excavation and prevent cave-ins, are removed after the pipe is in place and backfilling is under way. Since piping is installed on dry, stable trench bases, dewatering is necessary where ground-water levels are high. Removed water must be disposed of in a safe manner. Pipe must be properly joined with watertight joints, so that a smooth, continuous invert, on

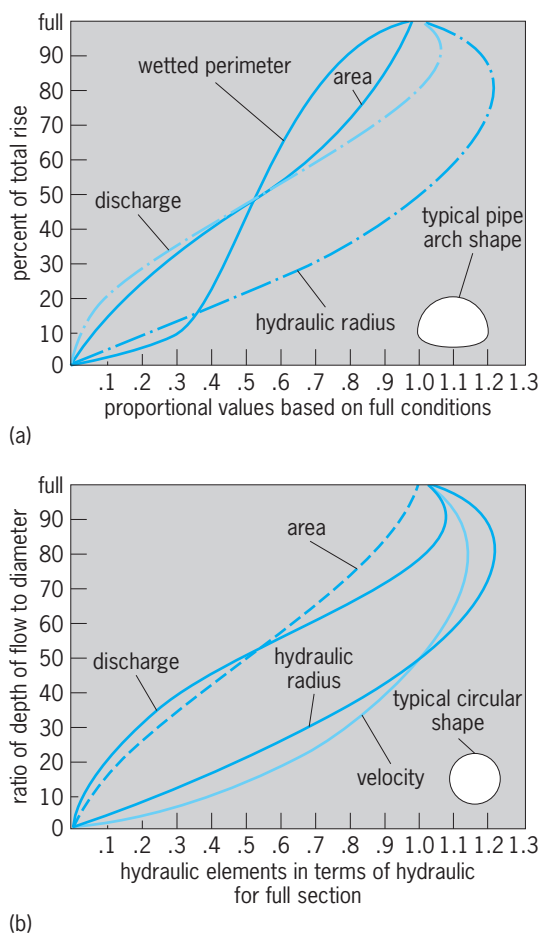

Fig. 11. Hydraulic element graphs for conditions of partial flow in terms of full section. (*a*) Pipe-arch section. (*b*) Circular section. (*After American Iron and Steel Institute, Modern Sewer Design, 1980*)

TABLE 4. Values of load coefficients C_d for trench condition

Fill material:	Sand and gravel	Saturated topsoil	Clay	Saturated clay
Specific weight, lb/ft^3 (kN/m^3):	100 (15.7)	100 (15.7)	120 (18. 9)	130 (20.4)
Ratio of cover depth to trench width				
1.0	0.84	0.86	0.88	0.90
2.0	1.45	1.50	1.55	1.62
3.0	1.90	2.00	2.10	2.20
4.0	2.22	2.33	2.49	2.65
5.0	2.45	2.60	2.80	3.03
6.0	2.60	2.78	3.04	3.33
7.0	2.75	2.95	3.23	3.57
8.0	2.80	3.03	3.37	3.76
9.0	2.88	3.11	3.48	3.92
10.0	2.92	3.17	3.56	4.04
12.0	2.97	3.24	3.68	4.22
14.0	3.00	3.28	3.75	4.34

SOURCE: R. K. Linsley and J. B. Franzini, *Water Resources Engineering,* 2d ed., McGraw-Hill, 1979.

the proper slope, with no pockets, depressions, or offsets is developed. Stubs for house sewer connections and manhole openings must be made at proper locations and elevations. Stubs or openings must be sealed to prevent infiltration. Pipe must be properly backfilled to finished grade; the excavation must be paved to provide a safe surface for carrying traffic.

Pipe tunneling and jacking are other construction techniques used for pipe installations. These are generally used to pass under highways, existing embankments, railroads, and streams, or for other special situations. New sewer lines are tested for watertightness, generally with low air pressure, to establish conformance with infiltration and exfiltration limits.

Bedding. Installed sewer lines must be able to withstand the external loads superimposed on them due to dead load from soil backfill and overburden, and from live loads of traffic and impact. The pipe's ability to withstand these loads is based on the crushing strength of the pipe material used, the method of installation, and the type and style of bedding provided. The standard for the crushing strength to which the pipe is manufactured is defined by the laboratory three-edge bearing test (Table 1).

The method of installation is defined as either trench or embankment condition. In the trench condition, pipe is installed in a relatively narrow trench, excavated into firm soil, with trench walls extending above the top of the pipe. Backfill is then placed around and over the pipe and carried up to the final ground surface. An embankment condition occurs when pipe is essentially placed on the existing ground surface or in a very wide trench, with backfill placed over the pipe and carried up to the final ground condition. Bedding, used to reduce stress concentrations and increase pipe support, is the prepared foundation between the bottom of the pipe and the firm trench or support ground surface. The safe support strength or load carrying ability of the pipe is increased with better bedding conditions.

The Marston formula [Eq. (3) for trench condition and Eq. (4) for embankment condition] provides a

$$W_d = C_d w B_d^2 \tag{3}$$

$$W_c = C_c w B_c^2 \tag{4}$$

value for trench load on a buried pipe. Here W_d or W_c is the vertical dead load of fill in lb/linear ft or kN/m on the buried pipe, w is the unit weight of fill material in lb/ft^3 or kN/m^3, B_d is the width of the trench in ft or m, measured 1 ft (0.3 m) above the top of the buried pipe, B_c is the outside pipe diameter in ft or m, and C_d or C_c are load coefficients for pipe (**Tables 4** and **5**). As the width of the trench increases, a point is reached where there is no longer a trench and the pipe is in the embankment condition. In this case the smaller value obtained for W_d or W_c controls. In addition to dead loads, vehicular and impact loads must be considered. As the live load is applied to the ground above the pipe, the load is distributed; its intensity decreases with increased depth of cover. Depth of cover is a major concern in sewerage systems to maintain pipeline integrity.

TABLE 5. Values of load coefficients C_c for embankment condition*

Ratio of cover depth to pipe diameter	C_c
1.0	1.2
2.0	2.8
3.0	4.7
4.0	6.7
6.0	11.0
8.0	16.0

* Rigid pipe, unyielding base, noncohesive backfill.
SOURCE: R. K. Linsley and J. B. Franzini, *Water Resources Engineering,* 2d ed., McGraw-Hill, 1979.

The sum of all loads, multiplied by a safety factor usually taken as 1.5, divided by the rated three-edge crushing strength of the selected pipe, establishes the load factor or type of bedding condition required (Table 1). The combination of pipe material, crushing strength, and bedding is an economic consideration in the design of sewerage systems. *See* LOADS, DYNAMIC; LOADS, TRANSVERSE; STRENGTH OF MATERIALS.

Maintenance. Clogging from accumulations of debris, tree roots, or illegally dumped materials that prevents proper flow of liquid through appurtenances and pipes is the major problem encountered in sewerage systems. Cleaning equipment to remove these obstructions includes cutting or boring tools to cut or rip through roots or balled-up debris and special tools known as pigs to move deposits of sand, grit, or other materials mechanically and hydraulically from the line to a manhole, where orange-peel buckets lift out the deposited materials.

In modern systems, breaks, cracks, spalling of pipe, poor joints, or other defects are inspected through the use of underwater television cameras. Specialized methods of concreting, patching, or even inserting a liner within the pipe have been developed to reduce infiltration, maintain the usefulness of pipes, and reduce the amount of surface excavation needed to make repairs. *See* UNDERWATER TELEVISION.

Safety. A sewerage system is a confined space where organic matter is undergoing organic decomposition. Hazardous and toxic gases generally exist; in addition, because these areas are belowground, there may be a lack of oxygen. Before personnel can safely enter a manhole, vault, or chamber, the space must be ventilated mechanically for a period of time and then tested for gases.

Entry into the sewerage system is through manholes that have narrow entry hatches or manholes with steps cast into the wall. Care must be exercised that at least two people work together, that the entering person wears a safety harness, and that the steps are safe. As a minimum, OSHA recommendations and regulations should be followed. *See* INDUSTRIAL HEALTH AND SAFETY; SEWAGE. Gerald Palevsky

Bibliography. American Concrete Pipe Association, *Concrete Pipe Design Manual*, 16th ed., 2004; American Iron and Steel Institute, *Modern Sewer Design*, 4th ed., 1999; M. J. Hammer, *Water and Wastewater Technology*, 5th ed., 2003; R. K. Linsley and J. B. Franzini, *Water Resources Engineering*, 4th ed., 1991; G. Tchobanoglous, F. L. Burton, and H. D. Stensel (eds.), *Wastewater Engineering: Treatment and Reuse*, 4th ed., 2002.

Sewage solids

The accumulated, semiliquid material consisting of suspended, colloidal, and dissolved organic and inorganic matter separated from wastewater during treatment. Sludges are developed as contained pollutants, and contaminants are separated by mechanical, hydraulic, biological, or chemical processes. The various classes of solids that are removed and collected must be disposed of in a safe, nuisance-free manner without hazard to health or the environment. Collection, handling, transporting, and disposal of removed solids are difficult and costly, since they are offensive and putrescible, with 92–99.5% water content. Sewage solids must be treated by thickening, chemical conditioning, mechanical dewatering, thermal action, biological stabilization, or digestion to convert putrescible organic matter to relatively inert end products, remove water, and reduce weight and volume.

Types. Sewage solids are classified as screenings, scum, grit, septage, or sewage sludges (see **table**).

Screenings. Large solids, carried by incoming wastewater, are captured mechanically on screens or racks with openings of various sizes. These protective units remove floating debris, including wood, clothing, cans, rags, paper, rubber and plastic goods, and stringy material that could damage equipment or create problems in plant maintenance and operation. Fine screens, or microstrainers, may be employed at plant outfalls to prevent discharge of esthetically objectionable floating solid matter. Modern, self-cleaning, inclined, rotary-drum or centrifugal screens are being installed to augment or replace grit chambers and primary settling tanks.

Screening volume varies with type of waste flow, season, weather, degree of urban development, and screen opening. Increased use of mechanical cutters and shredding devices has reduced volume of removed screenings. Coarse screenings range from 1.0 to 4.0 $ft^3/10^6$ gal (7.5 to 30 $m^3/10^6$ m^3).

Scum. This is defined as the floating fraction of sewage solids, with specific gravity under 1.0, that, under quiescent conditions, rises to the surface of the wastewater. Primary tank skimmings contain oils, fats, soaps, rubber and plastic hygienic products, cigarette filter tips, paper, and similar materials.

Grit. Heavy suspended solids consisting of sand, cinders, coffee grounds, seeds, small metal objects, and other generally inorganic particles carried in wastewater inflow are collectively known as grit. The amount of grit varies with type of sewer, season, weather, intensity of runoff, condition of streets and sewers, and use of household garbage disposal units. In new, separate sanitary sewers, with proper street sweeping, grit volumes in sanitary sewage are generally less than 4 $ft^3/10^6$ gal (30 $m^3/10^6$ m^3).

Septage. This consists of partially digested material pumped from on-site sanitary wastewater disposal systems. It contains a mixture of grit, scum, and suspended solids, adding to treatment plant sludge. *See* LEACHING; SEPTIC TANK.

Sewage sludge. Sedimentation, subsidence, or clarification of suspended matter from turbid water by gravity occurs when the velocity of that liquid has been reduced below its transport capacity. Transport capacity is controlled by varying flow-through velocity, which causes differential hydraulic subsidence. Thus heavy inorganic solids, grit, can be settled

Sources of solids and sludge from a conventional wastewater treatment facility*

Unit operation or process	Type of solids or sludge	Remarks
Screening	Coarse solids	Coarse solids are often comminuted and returned to the wastewater for removal in subsequent treatment facilities.
Grit removal	Grit and scum	Scum-removal facilities are often omitted in grit-removal facilities.
Preaeration	Scum	In some plants, scum-removal facilities are not provided in preaeration tanks.
Primary sedimentation	Primary sludge and scum	The quantities of both sludge and scum depend on the nature of the collection system and whether industrial wastes are discharged to the system.
Aeration tank	Suspended solids	Suspended solids are produced from the conversion of biological oxygen demand (BOD). If wasting is from the aeration tank, flotation thickening is normally used to thicken the waste-activated sludge.
Secondary sedimentation	Secondary sludge and scum	Provision for scum removal on secondary settling tanks is now a requirement of the U.S. Environmental Protection Agency.
Sludge-processing facilities	Sludge and ashes	The characteristics and moisture content of the sludge and ashes depend on the operations and processes that are used.

*After Metcalf & Eddy, Inc., G. Tchobanoglous, and F. Burton, *Wastewater Engineering: Treatment, Disposal, and Reuse,* 3d ed., McGraw-Hill, 1990.

in one chamber while light agglomerated organic particles, sludges, settle in others.

Sludge derives its name from the unit process from which it settles out. Primary sludge, or raw sludge, develops as solids in incoming wastewater settle hydraulically. Raw sludge, containing up to 5% solids by weight, is gray, greasy, viscous, unsightly, contains visible fecal solids and scraps of household wastes, and has a disagreeable odor. Use of household garbage grinders can increase primary sludge weight and volumes by 25–40%.

Chemical precipitation may be used to increase the removal of primary solids, to adjust pH for further treatment or to remove nutrients as phosphorus. Chemical dosage is high, producing large volumes of sludge. *See* PRECIPITATION (CHEMISTRY).

Secondary treatment uses aerobic biota to enmesh, coagulate, oxidize, nitrify, and mineralize the unsettled colloidal and dissolved organic matter that remains after primary treatment. Aerobic processes such as trickling filters, rotating biologic disks, or activated sludge utilize microorganisms to reduce organic and nutrient matter, and form sludges of settled biologic floc, which are light tan to dark brown in color, spongy, generally uniform in texture and form, having an earthy odor with solids content up to 1.5% by weight. Anaerobic decomposition will cause these sludges to turn septic rapidly, become dark to black and develop malodorous gases. *See* WATER TREATMENT.

Sludge thickening. This process removes water, increases the concentration of solids, reduces weight and volume, and prepares sludges for further treatment and handling. Flotation uses fine air bubbles to entrain and carry solids to the surface for skimming. Gravity thickening allows self-compaction of accumulated sludge, while gentle stirring causes water to separate. Use of a centrifuge compresses sludge against a porous surface that retains solids but allows liquid to pass through. Supernatant, the liquid drawn off from a sludge, is turbid, contains fine solid particles, has a high organic content, is odorous, and is returned to the inflow for retreatment.

Stabilization. When microorganisms multiply rapidly, anoxically, on organic and volatile substances contained within sludges, putrefaction and offensive odors develop. Sludges are stabilized biologically, chemically, or thermally to make them less putrescible and odorous, reduce pathogenic content, reduce weight and volume, and prepare them for final disposal. When stabilized biologically, either aerobic or anaerobic microorganisms reduce organics, volatile matter, and nutrients by converting these to gases, water, and cell tissue. Use of chemicals kills microorganisms, causing biota cells to shrivel and coagulate, releasing and separating water. Thermal processes sterilize (pasteurize) the biota, causing cell destruction, coagulation, and release of water. Chemical and thermal processes generally do not change organic or nutrient levels.

Aerobic stabilization. This allows aerobic microorganisms, during prolonged periods of aeration, to oxidize organic matter within sludge to carbon dioxide, water, ammonia, and nitrates. A humuslike, relatively odor-free and biologically stable mass that has some useful soil conditioning values is developed in some 20 days of controlled operation. The aerobic stabilization process is relatively easy to operate, and it produces a fairly clear supernatant liquor. Drawbacks are high power costs, difficulty with final settling, dewatering, and drying of the sludge, and absence of produced methane gas as a useful by-product.

Anaerobic digestion. Digestion, the most common method of handling sludges in municipal plants, occurs in closed tanks. Anaerobic digestion comprises a series of complex, interrelated biochemical reactions, performed by groups of microorganisms functioning in the absence of free oxygen. Under controlled temperature and pH, the anaerobic microorganisms produce methane gas as they decompose the sludge to a drainable, relatively inert, humic mass.

The digestion process can be envisioned as occurring in discrete, independent steps. In the first stage, acid fermentation, pH drops rapidly as colloidal solids, lipids, dissolved carbonaceous matter,

and sugars are rapidly converted to organic acids, with evolution of carbon dioxide and hydrogen sulfide. In the second stage, acid regression, pH rises slowly as organic acids formed earlier, and some proteins, are digested to acetate and ammonia compounds. In the third stage, alkaline fermentation or alkaline digestion, proteins, organic acids, and amino compounds are attacked. The pH rises to around neutral (pH 7), with large volumes of gases, predominantly methane, produced. In a properly functioning anaerobic digester the raw sludge is changed from a putrescible mass with over 70% volatile solids and high water content to a more mineralized, easily draining and drying humic mass having an inoffensive tarry odor with less than 50% volatile solids.

The sizing of digestion tanks is a function of raw sludge volumes added, digestion and water separation rates, and the volume of digested sludge drawoff. Sludge digestion occurs at temperatures from 42 to 140°F (5.6 to 60°C), with increased temperatures increasing bacterial action and reducing digestion time. Digestion time is expressed by the days necessary to attain 90% of ultimate gas production (see **illus.**). Thickening sludge prior to its being fed into the digester reduces input sludge volumes, assists in maintaining tank operation at optimum digestion temperatures, and saves space and energy.

The presence of methane gas, as it forms and leaves the mixed sludge liquor, indicates organic stabilization. Gas production at optimum mesophilic temperatures is 12–18 ft^3/lb (0.75–1.12 m^3/kg) of volatile matter destroyed. Digester gas, 60–70% methane, has a heating value of 600–700 Btu/ft^3 (22–26 kJ/m^3). The digester gas is generally used at treatment plants as an economical fuel for powering engines and heating digester tanks. *See* METHANE.

Chemical stabilization. This method uses large dosages of chlorine or lime to render sludges unsuitable for microorganisms to metabolize the organic substances, to grow or survive. The chemical stabilization process generates large volumes of sludge, often with problems of ultimate disposal.

Thermal processes. These use a variety of heating methods and sludge handling to destroy cell structure and reduce the quantity of moisture. The end solid is relatively dry, sterilized, generally free of offensive odors, and is frequently used as a soil amendment.

Disposal. Materials removed from the wastewater must be disposed of by ultimately returning them to the land, ocean, or atmosphere without damaging the receiving environment. Regulations have prohibited ocean dumping, and strict air emission rules have reduced the use of thermal systems. Solids are generally disposed of in landfills, buried, composted, or recycled as soil amendments.

Screenings, after shredding, may be returned to inflow for retreatment or are buried or incinerated. Washed grit is buried or used as fill or cover within landfills.

Digested liquid sludges may, in isolated areas, be spread or plowed under fields, but this is not a generally accepted practice. Wastewater solids may contain pathogenic organisms, but more importantly, higher concentrations of initial pollutants than were in the original wastewater. Disposal of removed solids must neither adversely affect the food chain nor endanger the environment. *See* FOOD WEB.

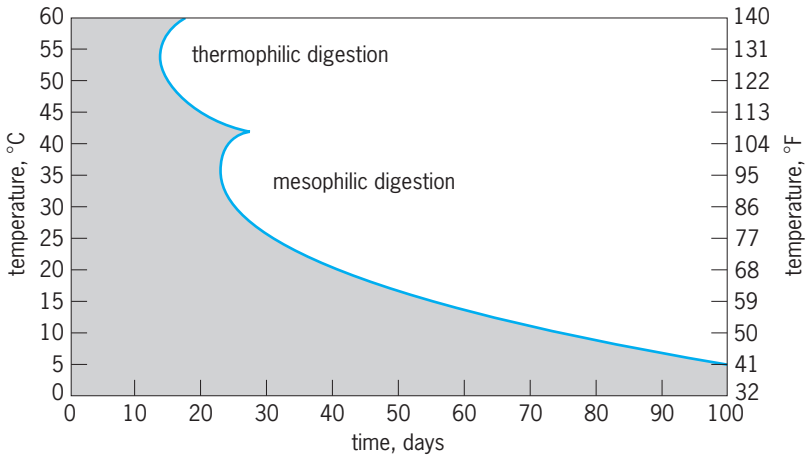

Time required for 90% digestion of sewage sludge at different temperatures, (*After K. Imhoff and G.M. Fair, Sewage Treatment, John Wiley & Sons, 1940, reprint 1946*)

Digested sludge as withdrawn is semiliquid; therefore, for ease of handling, storing, transporting, or landfilling and for increased efficiency in further thermal processing, it must be dewatered or dried. Where adequate area exists, open drying beds can be employed to reduce water content; otherwise centrifuges, belt presses, vacuum filters, and thermal or other processes are used. When added to soil, the dried humuslike digested sludge increases the water-holding capacity of the soil, improves texture and tilth, reduces erosion, and provides some nutrient benefit. Digested sludge, an incomplete fertilizer, requires additions of nitrogen, phosphorus, potassium, and even trace elements.

Composting, a natural thermophilic biological process, further reduces pathogens and volatile solids in digested sludge, forming a stable humuslike soil-conditioning material that can be recycled.

Heat drying uses external energy and reduced pressure to reduce moisture content to under 12%. Since drying by heat pasteurizes (sterilizes) the sludge and stops further bacterial action, dried digested sludges can easily be processed mechanically, stored, handled, and applied to land.

Thermal processes, including wet oxidation, pyrolysis, and incineration, require sludges with high content of organic solids that can produce significant heat value. Prestabilization of the sludge with reduction in volatile sludge content is not usually employed. In wet oxidation sludge is mixed with compressed air under high heat to oxidize organic matter. Pyrolysis, a thermal cracking process, converts sludge in reduced oxygen atmospheres to gases, water, and char. In both wet oxidation and pyrolysis, withdrawn supernatant liquors are characterized by high biochemical oxygen demand and solids and are retreated. Incineration is complete combustion and, as practiced in large plants, provides a reduction in volume. Environmental considerations require that

air pollution standards be met. Residual ash is usually disposed of in landfills. *See* AIR POLLUTION; HAZARDOUS WASTE; PYROLYSIS; SEWAGE; SEWAGE TREATMENT. Gerald Palevsky

Bibliography. Environmental Protection Agency, *Sludge Treatment and Disposal*, Process Des. Man. EPA 625/1-79-011, Center for Environmental Research Information, Cincinnati, 1979; A. F. Gaudy, Jr., and E. T. Gaudy, *Microbiology for Environmental Scientists and Engineers*, 1980; J. G. Henry and G. W. Heinke, *Environmental Science and Engineering*, 2d ed., 1996; Metcalf & Eddy, Inc., G. Tchobanoglous, and F. Burton, *Wastewater Engineering: Treatment, Disposal, and Reuse*, 3d ed., 1990; H. S. Peavy, D. R. Rowe and G. Tchobanoglous, *Environmental Engineering*, 1985; T. D. Reynolds, *Unit Operations and Processes in Environmental Engineering*, 2d ed., 1995; P. A. Vesilind, *Treatment and Disposal of Wastewater Sludges*, rev. ed., 1979.

Sewage treatment

Unit processes used to separate, modify, remove, and destroy objectionable, hazardous, and pathogenic substances carried by wastewater in solution or suspension in order to render the water fit and safe for intended uses. Treatment removes unwanted constituents without affecting or altering the water molecules themselves, so that wastewater containing contaminants can be converted to safe drinking water. Stringent water quality and effluent standards have been developed that require reduction of suspended solids (turbidity), biochemical oxygen demand (related to degradable organics), and coliform organisms (indicators of fecal pollution); control of pH as well as the concentration of certain organic chemicals and heavy metals; and use of bioassays to guarantee safety of treated discharges to the environment.

In all cases, the impurities, contaminants, and solids removed from all wastewater treatment processes must ultimately be collected, handled, and disposed of safely, without damage to humans or the environment.

Treatment processes are chosen on the basis of composition, characteristics, and concentration of materials present in solution or suspension. The processes are classified as pretreatment, preliminary, primary, secondary, or tertiary treatment, depending on type, sequence, and method of removal of the harmful and unacceptable constituents. Pretreatment processes equalize flows and loadings, and precondition wastewaters to neutralize or remove toxics and industrial wastes that could adversely affect sewers or inhibit operations of publicly owned treatment works. Preliminary treatment processes are employed to protect plant mechanical equipment; remove extraneous matter such as grit, trash, and debris; reduce odors; and render incoming sewage more amenable to subsequent treatment and handling. Primary treatment employs mechanical and physical unit processes to separate and remove floatables and suspended solids and to prepare wastewater for biological treatment. Secondary treatment utilizes aerobic microorganisms in biological reactors to feed on dissolved and colloidal organic matter. As these microorganisms reduce biochemical oxygen demand and turbidity (suspended solids), they grow, multiply, and form an organic floc, which must be captured and removed in final settling tanks (**Fig. 1**). Tertiary treatment, or advanced treatment, removes specific residual substances, trace organic materials, nutrients, and other constituents that are not removed by biological processes.

On-site sewage treatment for individual homes or small institutions uses septic tanks, which provide separation of solids in a closed, buried unit. Effluent is discharged to subsurface absorption systems. *See* SEPTIC TANK; STREAM POLLUTION; UNIT PROCESSES; WATER TREATMENT.

Pretreatment. Wastewater, depending on its source, varies in quantity, strength, and composition. Pretreatment, particularly of sewage containing industrial wastes, neutralizes, alters, or removes nonbiologically degradable, toxic, hazardous, and highly corrosive materials before they enter the municipal sewer system. Chemical neutralization brings pH values to the range 6.0–8.0 and provides buffering capacity. Coagulation, flotation, and surface skimming of oils, greases, or other light constituents are enhanced by adding air or chlorine.

In the equalization process, widely varying flows are stored and then discharged to the treatment facility in a more uniform manner over a longer time period. Mixing and storing different flow streams neutralizes, dilutes, and dampens both quantities and concentrations, reducing shock loads on receiving systems. Equalization tanks are aerated to mix the sewage, inhibit septicity, increase gas exchange, and prevent sedimentation.

Preliminary treatment. Wastewaters contain sand, cinders, ceramic shards, glass, and other granular inorganic matter that is abrasive, causes wear on pumps and equipment, clogs lines, and occupies tank space. By controlling velocities in grit chambers (**Fig. 2**) at about 1 ft/s (0.3 m/s), dense granular matter [larger than 0.008 in. (0.2 mm)] settles, while materials with low specific gravity remain in suspension within the flow. Settled solids, being removed from the bottom by grit elevators or detritors, are agitated within the incoming flow. Agitation abrades and separates clinging organic matter from the inert solids, and the flowing water carries these removed putrescible substances away, leaving a washed grit that is ready for landfill disposal. *See* SEWAGE SOLIDS.

Unwanted materials like plastic objects, throwaway sanitary hygiene products, diapers, and towels, and items ranging from string to bed sheets or toothpicks to tree limbs become part of the sewage flow. Screens are provided to protect equipment from impact damage, from stringy material winding around and fouling rotating shafts, or from plastics and rags that clog valves. Bar racks and screens remove such large solids and flotables. Bar racks composed of rigid

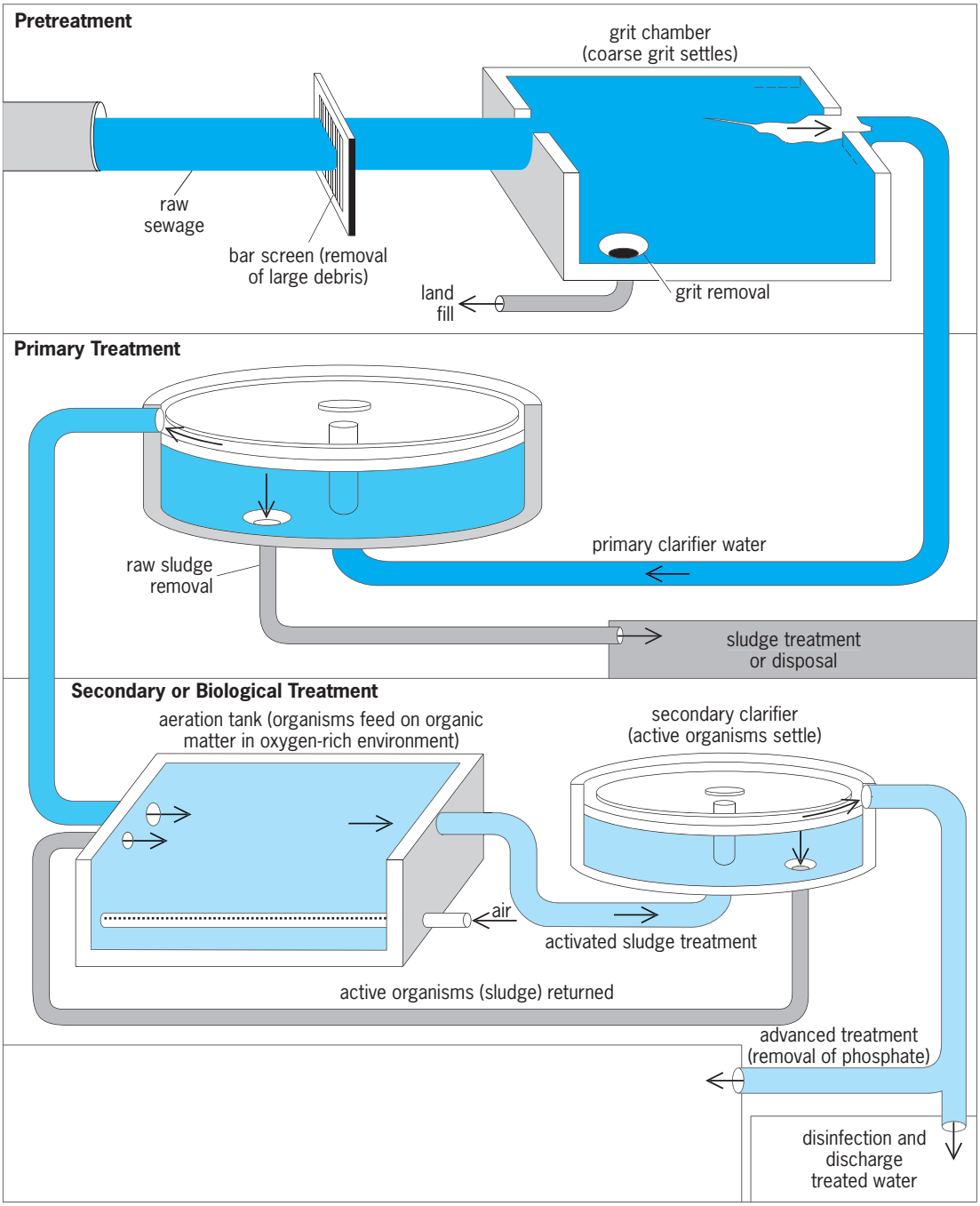

Fig. 1. Schematic diagram of liquid wastewater treatment through the secondary process. (*After B. J. Nebel, Environmental Science: The Way the World Works, 2d ed., Prentice-Hall, 1981*)

parallel bars, spaced 0.5–3 in. (12.5–75 mm) apart, held in place by support cross members, are installed within a controlled flow channel to catch this debris. Large, coarse material that is carried in the wastewater impacts the bar rack and is retained on the upstream face, from which it is mechanically removed and collected by traveling rakes.

Screens are formed of wire mesh, screen cloth, or perforated plates, with openings 0.5 in. (12.5 mm) or less. Coarse screens, with openings greater than 0.2 in. (5 mm) are used as pretreatment. Medium screens, 0.01–0.06 in. (0.8–1.5 mm), may be used in lieu of primary treatment. Fine screens, up to 0.003 in. (0.075 mm) openings, are used as final screening before plant discharge.

Disk, drum, band, and other styles of preliminary treatment screens are movable, with only a portion rotating within the flow. As the screen rotates, clean surfaces enter the wastewater, and previously immersed areas are lifted out and cleaned of solids caught on the surface.

Removal, handling, and disposal of screenings from bar racks and coarse screens are generally difficult maintenance problems, with associated odors and esthetically unpleasant conditions. To deal with this situation and to develop a smaller, more uniform

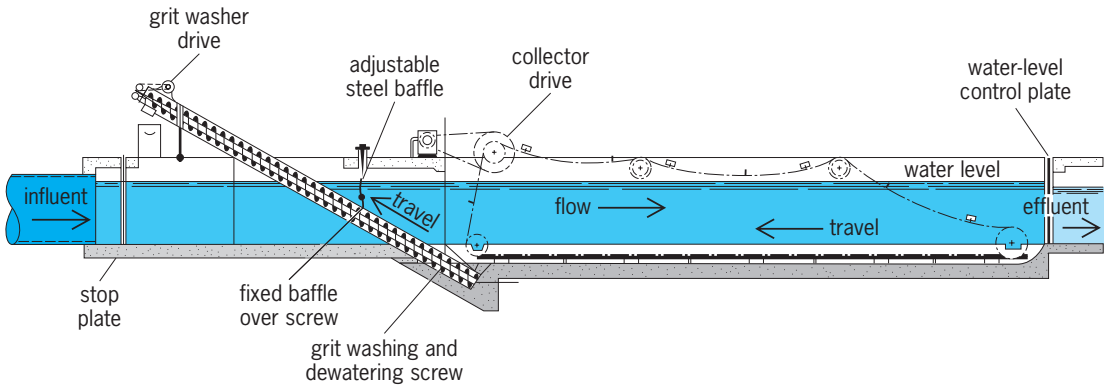

Fig. 2. Diagram showing components of a grit chamber. (*After Joint Committee of Water Pollution Control Federation and American Society of Civil Engineers, Wastewater Treatment Plant Design, MOP 8, 1977*)

sized solid, comminution (shredding) devices are used. Solids, trapped and held by the force and flow of liquid on a slotted plate or bars, are cut up or shredded by moving cutters or teeth. The solids, reduced in size, pass through the openings and are carried downstream to further treatment.

Primary treatment. Primary sedimentation, or clarification, is the first process where removal of substantial quantities of suspended solids and materials causing biochemical oxygen demand in wastewater flow occurs. Sedimentation tanks, rectangular or circular, operate on a continuous flow-through basis, yet they maintain relative quiescence to permit suspended solids to settle if the specific gravity is greater than that of water, or to float if the specific gravity is less. Mechanical devices remove the accumulated suspended solids (sludge) from the tank bottom, while floatable materials (skimmings) are taken off the surface. The clarified liquor, known as primary effluent, is discharged over the tank's effluent weirs.

Discrete particles of uniform size, shape, and specific gravity settle according to Stokes' law. Theoretically, if the velocity of a downward-settling particle is greater than the horizontal velocity of the carrying liquid, the particle will be captured and retained within the sedimentation tank. However, empirical values are used because sewage solids vary in shape, size, specific gravity, and concentration, and because the rate and path of settling are affected by temperature, age, density, and wind currents as well as eddies developed by mechanical scrapers and turbulence at influent and outlet structures.

Removal of suspended solids and biochemical oxygen demand is based on surface overflow rate and detention time. The surface overflow rate is expressed in gal/ft^2 (m^3/m^2) of tank surface area per day. Detention time varies from 30 min for minimal removal prior to some types of biological treatment, to 60–90 min (typical values within conventional plants), and longer times if chemical coagulation or nutrient removal is employed (**Fig. 3**).

A solids contact clarifier diverts the incoming wastewater through a layer or blanket of previously

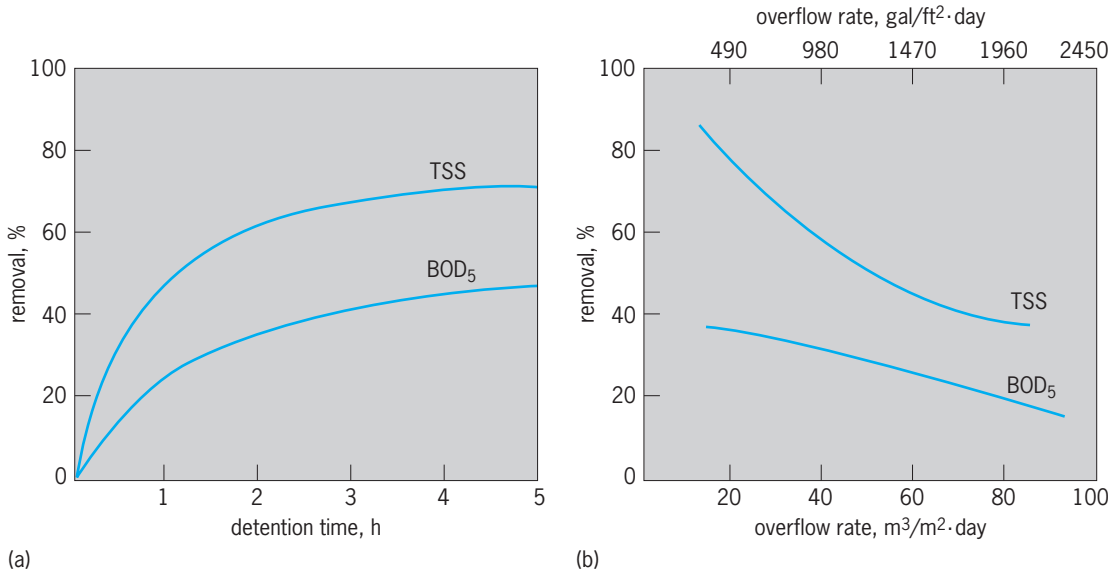

Fig. 3. Removal of suspended solids and biochemical oxygen demand in primary clarifiers at various (*a*) detention times and (*b*) overflow rates. (*After S. R. Qasim, Wastewater Treatment Plants, Planning, Design and Operations, CBS College Publishing, Holt, Rinehart and Winston, 1985*)

settled solids. Suspended solids within incoming sewage contacts the sludge blanket, is agglomerated, and remains trapped within the sludge blanket as the liquid rises and overflows the outlet weirs (**Fig. 4***a*). Although efficient for chemical and special suspensions, and frequently used in water treatment, solids contact clarifiers are used infrequently for biological floc, because extended detention times, particularly when it is warm, leads to septicity, gasification, and odors.

Based on discrete particles, suspended solids removal is related to available surface area and is independent of tank depth. The use of inclined tubes or plates (lamella) within sedimentation tanks increases the theoretical tank surface area and reduces both the settling time of the particles and the depth at which they come to rest. With highly inclined plates or tubes, there is increased settling efficiency, and solids contact benefits may also be derived, but some problems with biological decomposition remain (Fig. 4*b* and *c*).

Use of additives such as lime, alum, chlorine, clays, or polyelectrolytes increases suspended solids removal from wastewater, but produces more sludge that impacts subsequent handling of solids. The added costs for chemicals, feeders, and handling of solids make the use of chemicals less attractive.

Various configurations are employed to reduce velocities at inlet and outlet structures and prevent adverse mixing, currents, short circuiting, or scour. Weir overflow rates are limited to 15,000 gal per day per lineal foot of weir face (186 $m^3/m \cdot d$).

Static or inclined self-cleaning screens, centrifugal screens, and rotary-drum screens are being used to upgrade overloaded primary units and to replace preliminary and primary units. Removal of suspended solids and biochemical oxygen demand by screens is generally not as complete as by conventional units, but overall costs are lower. Subsequent plant units must be sized to compensate for these differences in removal. *See* PRECIPITATION (CHEMISTRY); SEDIMENTATION (INDUSTRY).

Secondary treatment. Biological organisms, predominantly aerobic bacteria, convert and metabolize dissolved and colloidal matter remaining in wastewater to new cellular material, carbon dioxide, and water. These biological processes use organisms that form either a fixed film (attached growth) or are motile (suspended growth); they convert degradable organics to different, larger forms without significantly reducing total biochemical oxygen demand. Safe discharge to the environment and compliance with effluent standards require that the biota and newly formed cellular material be removed. This is accomplished by provision of secondary settling tanks that function as an integral part of the biological process.

Given adequate detention time, dissolved oxygen, controlled pH, appropriate temperature, absence of any toxic materials or shock loads, these biological processes can reduce biochemical oxygen demand to under 15 mg/liter and suspended solids to under 20 mg/liter. Not all wastewater organics can be fully

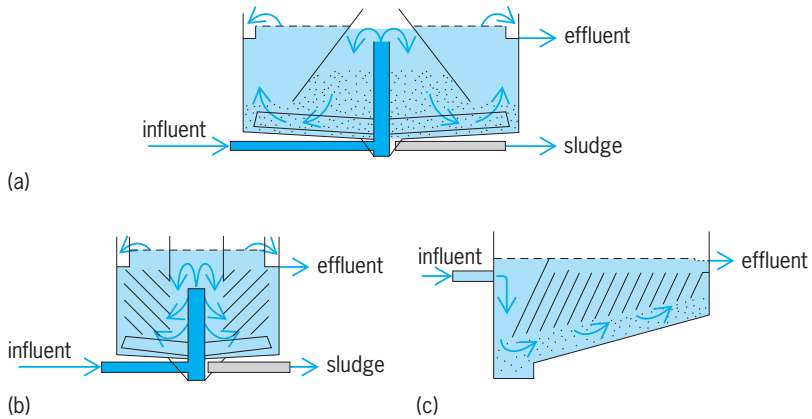

Fig. 4. Diagrams showing solids contact and inclined lamella settlers. (*a*) Circular settling tank using solids contact with sludge blanket. (*b*) Inclined lamella within a circular settling tank. (*c*) Inclined lamella within a rectangular sedimentation tank. (*After S. R. Qasim, Wastewater Treatment Plants, Planning, Design and Operation, CBS College Publishing, Holt, Rinehart and Winston, 1985*)

or adequately degraded in the time available and under the constraints found in publicly owned systems.

Fixed-film systems. Trickling filters (**Fig. 5**) contain beds of filter media—inert materials, with large surface area and voids, that support a slime growth, that is, an attached zoogleal mass composed of algae, fungi, protozoa, rotifera, nematoda, and most importantly, aerobic bacteria. Wastewater is applied to the surface of the filter media, which are made up of rock, slag, or synthetically formed shapes of manufactured materials, of different depths. The organic materials in this intermittently applied wastewater, flowing over and through the aerobic gelatinous matrix adhering to the support surfaces, undergo biologic coagulation, precipitation oxidation, and clarification as the flow moves downward to the underdrains. Organisms, in the presence of oxygen, convert the organic substances in the wastewater by various physical, chemical and enzymatic processes to energy for life, growth, and replication, and form carbon dioxide, nitrates, and sulfates as by-products.

This reactive zoogloeal mass converts organic matter and trapped particles to a humic mass as long

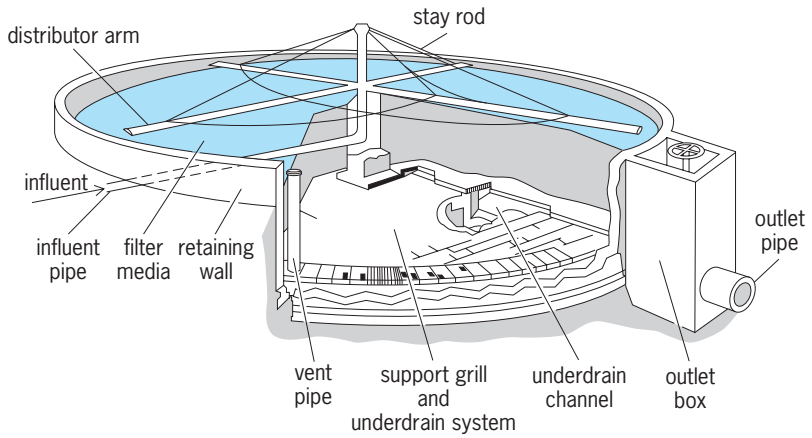

Fig. 5. Diagram of the components of a typical trickling filter. (*After Water Pollution Control Federation, Operation of Wastewater Treatment Plants, MOP 11, 1976*)

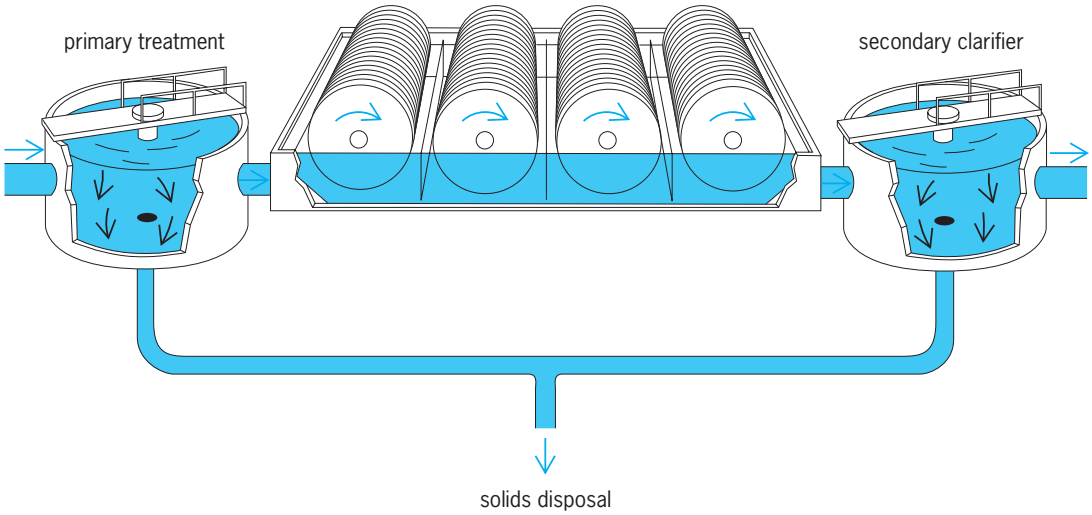

Fig. 6. Schematic diagram of a rotating biological contactor. (*After Water Pollution Control Federation, Operation of Wastewater Treatment Plants, MOP 11, 1976*)

as there is adequate oxygen transfer. If the film becomes too thick, or otherwise so impacted or compacted that oxygen cannot diffuse, portions become anoxic, and the film loses its capability to adhere to the support surface. Applied wastewater physically scours away nonattached portions, rejuvenating the filter by exposing areas for new colonization. Dislodged humic mass, known as the slough-off, is washed through the filter's large voids and conveyed through the underdrains to the secondary settling tank for ultimate capture and disposal.

Trickling filters are classified as standard or low-rate, intermediate, high-rate, and roughing filters according to hydraulic and organic loadings and rate of recirculation. Two-stage operation, the placing of two filters in series, is used for high organic loads. Rotating biological contactors (**Fig. 6**), also known as biodisks, use closely spaced plastic disks mounted on a rotating shaft. An aerobic zoogleal mass attaches itself and grows on the disks as they are intermittently exposed to organic materials and air as they rotate, with approximately 40% of their area immersed in the wastewater flow. The relative motion between the rotating disks and the flowing wastewater in the tank scours the attached film and carries the slough-off to the secondary settling tank.

Because daily wastewater flows into a publicly owned treatment works are highly variable both in amount of flow and in organic loading, they cause fluctuations in the liquid applications to the trickling filters. Rapid swings in biological density within the attached matrix develop, reducing efficiency. To provide flow and loading uniformity, a portion of the wastewater is recirculated after passing through the filter. This recirculation maintains continuous growth and sloughing, improves operation, increases efficiency, and reduces bed volume. Recirculation rate is the ratio of recycle flow to average daily plant inflow.

Total liquid flow, generally applied through rotating arms over the surface of the bed, intermittently doses each portion and is known as the hydraulic application rate. The average daily dry weight of new organic matter applied to the total bed volume is the organic loading rate. Typical application rates and design data are shown in **Table 1**.

Activated sludge. This is an aerobic suspended-growth process (**Fig. 7**) in which biodegradable organics in wastewater are intimately mixed with a concentrated mass of biota and oxygen within an aeration tank. New microorganisms grow and flocculate as the biotic mass adsorbs, oxidizes, and reduces the organic wastes. As the mixed liquor leaves the aeration tank following several hours of aeration, the biotic mass with the newly formed floc is separated within final settling tanks. A portion of the settled floc, the activated sludge, is returned to the aeration tank to maintain the required concentration of biota, while excess sludge is removed for solids handling and ultimate disposal. Air, required for utilization of dissolved oxygen in metabolism and respiration, is also provided to maintain mixing and prevent sedimentation in the aeration tank.

Modifications of the activated-sludge process relate to how, where, when, and for how long returned activated sludge and oxygen are introduced into the aeration tank and maintained in contact with the mixed liquor. These modifications of conventional activated sludge are known as tapered aeration, step aeration, extended aeration, contact stabilization, complete mix, and pure oxygen. Mechanical or compressed air systems are used to introduce and

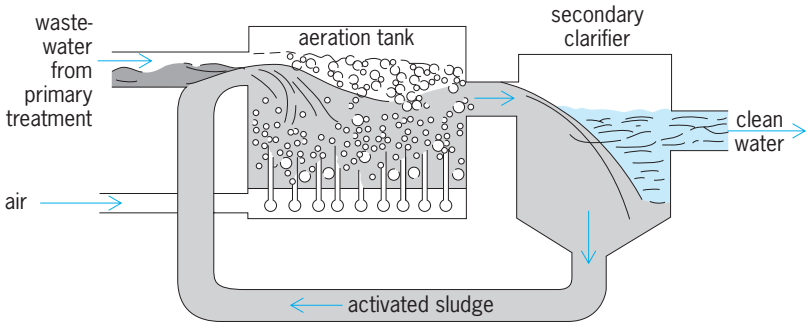

Fig. 7. Schematic diagram of activated sludge treatment. (*After B. J. Nebel, Environmental Science: The Way the World Works, 2d ed., Prentice-Hall, 1981*)

TABLE 1. Typical design information for trickling filters*

Item	Low rate	Intermediate rate	High rate	Superhigh rate	Roughing	Two-stage
Filter medium	Rock, slag	Rock, slag	Rock	Plastic	Plastic, redwood	Rock, plastic
Hydraulic loading, gal/ft^2 (liters/m^2) min	0.02–0.06 (1.2–3.6)	0.06–0.16(3.6–9.4)	0.16–0.64(9.438)	0.2–1.20 (12–70)	0.8–3.2 (47–187)	0.16–0.64 (9.4–38)
Mgal/acre (10^6 liters/hectare) day	1–4 (9–36)	4–10 (36–90)	10–40 (90–360)	15–90 (135–810)	50–200 (450–1800)	10–40 (90–360)
BOD$_5$† loading, lb/10^{3}ft^3 (kg/m^3) day	5–25 (0.08–0.40)	15–30 (0.24–0.48)	30–60 (0.48–0.96)	30–100 (0.24–1.6)	100–500 (1.6–8)	60–120 (0.96–1.9)
Depth, ft (m)	6–8 (1.8–2.4)	6–8 (1.8–2.4)	3–6 (0.9–1.8)	10–40 (3–12)	15–40 (4.5–12)	6–8 (1.8–2.4)
Recirculation ratio	0	0–1	1–2	1–2	1–4	0.5–2
Filter flies	Many	Some	Few	Few or none	Few or none	Few or none
Sloughing	Intermittent	Intermittent	Continuous	Continuous	Continuous	Continuous
BOD$_5$† removal efficiency, %	80–90	50–70	65–85	65–80	40–65	85–95
Effluent	Well nitrified	Partially nitrified	Little nitrification	Little nitrification	No nitrification	Well nitrified

*After Metcalf & Eddy, Inc., G. Tchobanoglous, and F. L. Burton, *Wastewater Engineering: Treatment, Disposal, and Reuse*, 3d ed., 1991, and S. R. Qasim, *Wastewater Plants: Planning, Design, and Operation*, 1985.
†BOD$_5$ = Measure of strength of waste in miligrams per liter.

maintain dissolved oxygen and mixing. Mechanical systems entrain atmospheric air through brushes, impellers, propellers, or turbines, while compressed air is introduced near the tank bottom through porous diffuser plates, spirally wound tubes, spargers, nozzles, or injection jets (**Fig. 8**).

Depending upon the process used, parameters such as empirical ratios of biodegradable organics (food) to concentration of microorganisms, aeration contact time, concentration of solids in the mixed liquor, and loading criteria for aeration tanks are established to achieve the desired degree of reduction of biochemical oxygen demand. Suspended growth systems, with relatively short detention times, are more sensitive to hydraulic, organic, and toxic shock loads and require greater operational skill and monitoring than fixed-film processes. However, suspended-growth systems occupy much less space and can produce effluents with lower soluble biochemical oxygen demand and suspended solids (**Table 2**).

For the suspended-growth system to operate optimally, the solids in the mixed liquor leaving the aeration tank must be completely separated, settled, and returned to the aeration tank quickly. Final settling tanks, similar in general configurations to primary tanks, generally have lower values of surface overflow rate and are deeper, since both separation and sludge densification of the light, fluffy biofloc must be accomplished.

For industrial wastes and small plants, aerated lagoons, oxidation ponds, and waste-stabilization lagoons are employed for secondary treatment or for final polishing after conventional treatment. Oxidation ponds, with large surface areas, depths less than 5 ft (1.5 m), with loading rates not to exceed 50 lb/acre (56 kg/ha) of biochemical oxygen demand, rely on natural reaeration, diffusion, and photosynthesis. Bacterial and algal actions reduce putrescible organics, converting them to stable algal cells. Although some solids settle and undergo anaerobic decomposition, oxidation ponds can support fish life. Evaporation, screened or filtered settled outflow, maintains liquid levels.

Aerated lagoons are facilities that are intermediate between oxidation ponds and activated sludge, since no sludge is recycled. They are frequently used to pretreat sewage with high organic loads or

TABLE 2. Common design parameters for activated sludge systems

Process modification	Sludge retention time (θ_c), d	F/M,[a] d^{-1}	Aerator loading,[b] kg/m^3·d	MLSS,[c] mg/liter	Aeration[d] period, h	Recirculation ratio (Q_r/Q)
Conventional	5–15	0.2–0.4	0.3–0.6	1500–3000	4–8	0.25–0.5
Tapered aeration	5–15	0.2–0.4	0.3–0.6	1500–3000	4–8	0.25–0.5
Step aeration	5–15	0.2–0.4	0.6–1.0	2000–3500	3–5	0.25–0.75
Complete mix aeration	5–15	0.2–0.6	0.8–2.0	3000–6000	3–5	0.25–1.00
Extended aeration	20–30	0.05–0.15	0.1–0.4	3000–6000	18–36	0.5–2.0
Contact stabilization	5–15	0.2–0.6	1.0–1.2	1000–4000[e] 4000–10000[f]	0.5–1.0[e] 3.0–6.0[f]	0.5–1.0
Pure oxygen	8–20	0.25–1.0	1.6–3.3	6000–8000	2–5	0.25–0.5

[a]The F/M (ratio of food to microorganisms) is kilograms of BOD$_5$ applied per day per kilogram of MLVSS in the aeration basin [lb BOD$_5$/(lb MLVSS · d)]. MLSS = mixed-liquor suspended solids in milligrams per liter. MLVSS = mixed-liquor volatile suspended solids, that is, that portion of the solids in the aeration tank that are organic and hence biodegradable. Q_r/Q = ratio of return sludge to average daily plant inflow.
[b]Aerator loading is kilograms of BOD$_5$ applied per day per cubic meter of aeration tank capacity [lb/(10^3 ft^3 · d)].
[c]Generally the ratio of MLVSS to MLSS is 0.75–0.85.
[d]Aeration period is the time in hours that the mixed liquor is in the aeration tank.
[e]Contact tank.
[f]Reaeration or stabilization tank.

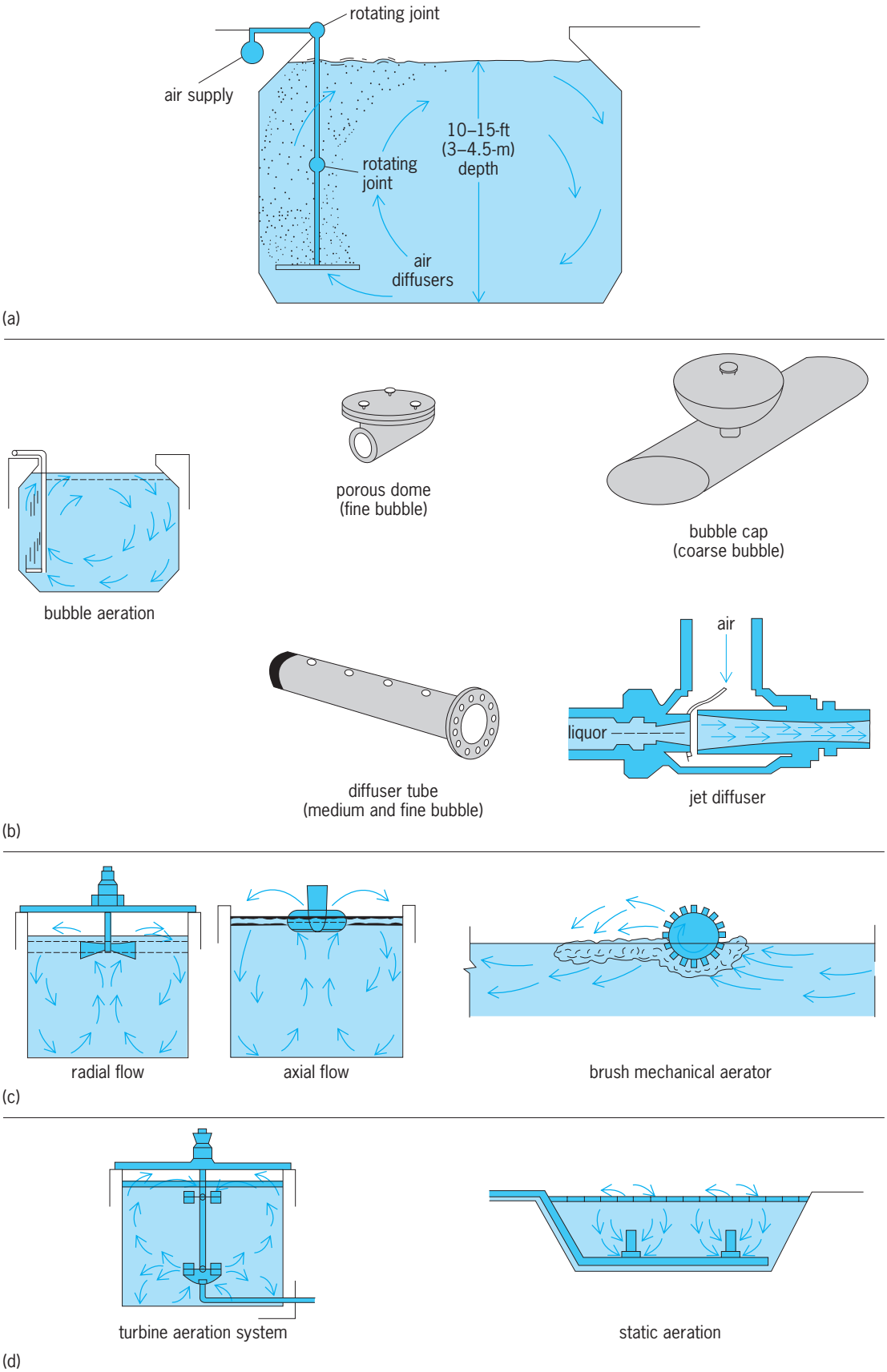

Fig. 8. Aeration systems and devices. (*a*) Cross section of a typical spiral-flow aeration tank showing the spiral-flow pattern created by aeration along one side (*after W. Viessman, Jr., and M. J. Hammer, Water Supply and Pollution Control, 4th ed., Harper and Row, 1985*). (*b*) Diffused aeration units. (*c*) Surface aeration systems. (*d*) Below-surface air release. (*b*, *c*, *d after W. W. Eckenfelder, Jr., Industrial Water Pollution Control, McGraw-Hill, 1989*)

industrial wastewaters prior to discharge to publicly owned treatment works. Basins are 8–18 ft (2.4–5.5 m) deep, and complete mixing is carried out by using mechanical devices, with detention times of several days. Some solids may settle and undergo anaerobic decomposition. When used as secondary treatment, basins are frequently followed by facultative lagoons, where liquid is retained for up to 20 or more days. The upper portion of a facultative lagoon is maintained aerobic, while the bottom, benthic zone, where solids accumulate, undergoes anaerobic decomposition. Released gases, rising particles, and bacteria are enmeshed, absorbed, destroyed, and converted to algal cells in the upper portion. Effluent may be subject to final mechanical screening, although frequently the effluent evaporates, percolates, or is used for land applications.

Final polishing. Rotating-drum microscreens, with stainless steel or plastic cloth having openings 30 micrometers in diameter, remove solids carried over from biological treatment and provide an effluent that is relatively free of turbidity. Filters are used for polishing of the final effluent as well as for removal of chemical precipitates from advanced treatment systems. Filtration, which is a complex process of screening, adsorption, and sedimentation, is defined and classified by the media, physical force, or application method employed. Flow may be induced by gravity, pressure, or vacuum; may be applied intermittently or continuously; and may be upflow, downward, or across sand, mixed media, granular carbon, membranes, or diatomaceous earth. Cleaning of filters, usually by backwashing, results in a new wastewater that must be captured, handled, and ultimately disposed of.

Disinfection. This process differs from sterilization in that it kills pathogenic organisms by physical or chemical means without total destruction of all the organisms present. Disinfecting agents should, at low concentrations, be highly selective and toxic to targeted pathogens. They must be stable, noncorrosive, economical, safe to humans and the environment, unaffected by pH or temperature, and long-lasting. None of the available agents or processes meets all or even most of these criteria. Disinfecting agents such as heat, phenols, alcohols, strong acids or alkalies, and halogens cause coagulation and denaturing of cell protein. Soaps and detergents injure cell membranes, change surface tension, or affect permeability. Oxidizing agents such as halogens, hydrogen peroxide, ozone, and heavy metals inhibit or denature enzymes and remove the free sulfhydryl (—SH) group. Direct sunlight, ultraviolet rays, or ionizing radiation destroys pathogens. Thermal processes and moist heat cause protein coagulation. Even though there are some problems with chlorine in the formation of trihalomethanes, it remains the agent most frequently used for waste water disinfection. Even with greater attention to control and dichlorination, it is relatively inexpensive, readily applied and controlled, and can be monitored by a simple and quick test. *See* ANTIMICROBIAL AGENTS; ENVIRONMENTAL ENGINEERING; STERILIZATION; ULTRAVIOLET RADIATION (BIOLOGY).

Advanced wastewater treatment. This is a treatment process (tertiary or polishing) that goes beyond the biological stage. Advanced wastewater treatment produces high-quality effluent, with reduced targeted toxic chemicals, nutrients, and trace organics. Such systems are a step in the cycle of transforming sewage to drinking water; they are closely related to processes that are used for industrial water treatment and for providing safe drinking water. Most advanced wastewater treatment systems include denitrification and ammonia stripping, carbon adsorption of trace organics, and chemical precipitation.

Nutrient removal of phosphorus and nitrogen is possible by chemical coagulation using lime or by biologic methods. The latter utilize systems having aerobic, anaerobic, and anoxic sequences to nitrify and then dentrify wastewaters. Ammonia can be removed by air stripping, breakpoint chlorination, or ion exchange. Carbon absorption traps refractory organics onto exposed carbon surfaces. Liquid can be passed over carbon in packed columns, or powdered carbon can be added to treated liquid and then filtered out on diatomaceous earth filters, on microstrainers, or by ultrafiltration. Evaporation, distillation, electrodialysis, ultrafiltration, reverse osmosis, freeze drying, freeze-thaw, floatation, and land application, with particular emphasis on the increased use of natural and constructed wetlands, are being studied and utilized as methods for advanced wastewater treatment to improve the quality of the treated discharged to reduce unwanted effects on the receiving environment. *See* ABSORPTION; DISTILLATION; EVAPORATION; SEWAGE; ULTRAFILTRATION; WETLANDS.

Gerald Palevsky

Bibliography. Metcalf & Eddy, Inc., G. Tchobanoglous, and F. L. Burton, *Wastewater Engineering: Treatment, Disposal, and Reuse*, 3d ed., 1991; S. R. Qasim, *Wastewater Treatment Plants, Planning, Design and Operation*, 1985; A. M. Thayer, Water treatment chemicals: Tighter rules drive demand, *Chem. Eng. News*, 68(13):17–34, 1990; W. Viessman and M. J. Hammer, *Water Supply and Pollution Control*, 4th ed., 1985; Water Environment Federation, *Design of Municipal Wastewater Treatment Plants*, Man. Prac. 8, 1993; Water Environment Federation, *Operation of Municipal Wastewater Treatment Plants*, Man. Prac. 11, 1991; S. R. Qasim, *Wastewater Treatment Plants, Planning, Design and Operation*, 2d ed., 1998; A. M. Thayer, Water treatment chemicals: Tighter rules drive demand, *Chem. Eng. News*, 68(13):17–34, 1990; W. Viessman and M. J. Hammer, *Water Supply and Pollution Control*, 6th ed., 1998; Water Environment Federation, *Design of Municipal Wastewater Treatment Plants*, Man. Prac. 8, 1993; Water Environment Federation, *Operation of Municipal Wastewater Treatment Plants*, Man. Prac. 11, 1991.

Sewing machine

A mechanism that stitches cloth, leather, book pages, and other material by means of a double-pointed needle or an eye-pointed needle. In ordinary

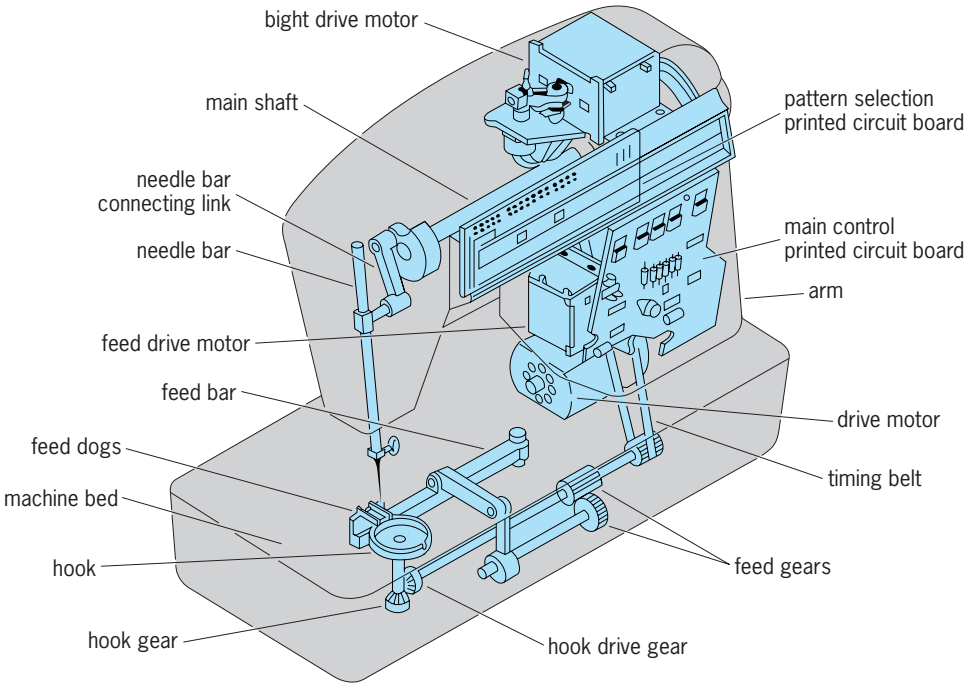

Components of a modern sewing machine. (*Singer Co.*)

two-threaded machines, a lock stitch is formed (see **illus**.). An eye-pointed reciprocating needle carries an upper thread through the layers of fabric, forming a loop beneath the material. A shuttle carrying a bobbin of under thread passes through the loop. Alternatively, a rotary hook takes the loop of upper thread and passes it around the bobbin of under thread. The needle withdraws, and a thread take-up lever pulls the stitch tight. The machine carries out these necessary motions and also feeds the material past the needle intermittently between each pass of the needle. A presser foot held against the material with a yielding spring adjusts itself automatically to variations in thickness of material and allows the operator to turn the material as it feeds through the machine. A cluster of cams, any one of which can be selected to guide the needle arm, makes possible a variety of stitch patterns. *See* CAM MECHANISM.

Frank H. Rockett

Sex determination

The genetic mechanisms by which sex is determined in all living organisms. The nature of the genetic basis of sex determination varies a great deal, however, among the various forms of life.

Forms of sexuality. There are two aspects of sexuality: the primary form involves the gametes, and the secondary aspect is gender.

Gametes. In its broadest usage the term "sex" refers to the processes that enable species to exchange materials between homologous chromosomes, that is, to effect recombination. Generally, recombination is essential to their mechanism for reproduction. For most organisms this involves, either exclusively or as one stage in the life cycle, the formation of special cells, known as gametes, by meiosis. *See* GAMETOGENESIS.

Most sexually reproducing species produce two different kinds of gametes (**Fig. 1**). The relatively large and sessile form, an ovum or egg, usually accumulates nutriments in its cytoplasm for the early development of the offspring. The relatively mobile form, a sperm (or pollen grain in many plants), contributes little beyond a haploid chromosome set. Thus the primary form of sex differentiation determines which kind of gamete will be produced. The formation of gametes usually involves the concomitant differentiation of specialized organs, the gonads, to produce each kind of gamete. The ova-producing gonad is usually known as an archegonium or ovary (in flowering plants it is part of a larger organ, the pistil or carpel); the gonad producing the more mobile gametes is usually known as a testis in animals and an antheridium or stamen in plants. *See* OVARY; OVUM; SPERM CELL; TESTIS.

Gender. In most animals and many plants, individuals become specialized to produce only one kind of gamete. These individuals usually differ not only in which kind of gonad they possess but also in a number of other morphological and physiological differences, or secondary sex characteristics. The latter may define a phenotypic sex when present, even if the typical gonad for that sex is absent or nonfunctional. The form that usually produces ova is known as female; the one that usually produces sperm or pollen is known as male. Since some sexual processes do not involve gametes, the more universal application of the term "gender" refers to any donor of genetic material as male and the recipient as female.

Humans. Human males and females tend to differ in the following secondary sex characteristics: cephalic

hairline, facial hair, larynx and voice, breasts, axillary hair, body hair, pubic hair, body configuration, phallic size, carrying angle of the arm, and knee relations. The female phenotype usually includes breast development and a characteristic disposition of fat over the hips and thighs. The pelvis tends to increase its lateral diameters, producing in adaptation a typical angulation of the forearm to the arm and the touching of the knees in the standard anatomical position. The pubic hair is kinky or curly and level at the upper margin, and the facial hair is inconspicuous. In contrast, the male pubic hair is straight and extends toward the umbilicus, while the facial hair increasingly becomes more conspicuous, as does that on the chest. After puberty the male's hairline tends to recede, the larynx to increase in size (the so-called Adam's apple), and the voice to deepen. The phallic organ (penis or clitoris) is usually much larger in the male, especially after puberty. Most of these characteristics are controlled by hormones, produced principally by the gonads (and therefore traceable to the primary sex determination). Their anomalous appearances in the two sexes are usually traceable to endocrine abnormalities. *See* ANDROGEN; ENDOCRINE MECHANISMS; ESTROGEN; HORMONE.

Plants. Sexual reproduction is not always accompanied by the kinds of differentiation described above. Thus, the major structures in the life cycles of some simple organisms such as the pond scum *Spirogyra*, a green alga, are filaments composed of haploid cells; these organisms sometimes reproduce sexually by passing the contents of a cell from one filament into a cell of an adjacent filament by means of a protoplasmic bridge or conjugation tube (**Fig. 2**). Since both conjugating cells are the same size and shape, they are known as isogametes. (The corresponding term for unequal-sized gametes is anisogamy; oogamy is anisogamy when the larger gamete is nonmotile.) Union of the two haploid isogametes produces a diploid zygospore that undergoes meiosis and germinates to form new filaments composed of haploid cells. The majority of plant species are monoecious, with both kinds of gonads on the same plant (**Fig. 3*a***). Plants such as the date palm, willow, poplar, and papaya that bear male and female gonads on separate plants are dioecious (Fig. 3*b*). They occur in about 60 of the 400 or so families of flowering plants, 20 of which are thought to contain exclusively dioecious species. *See* REPRODUCTION (PLANT).

Animals. Although the sexes are distinct in most animals, many species are hermaphroditic; that is, the same individual is capable of producing both eggs and sperm. This condition is particularly common among sessile or sluggish, slowly moving forms, including most flatworms, a large percentage of the worms that are parasitic in humans and other mammals, many members of the annelid phylum (such as earthworms and leeches), many mollusks (such as garden snails, oysters, scallops, and clams), and some fishes. Some hermaphroditic and monoecious species are homothallic; that is, the eggs and sperm of the same individual can combine successfully; but most are heterothallic, the gametes being capable only of cross-fertilization, often evolving special mechanisms to ensure its occurrence. In some cases the cross-fertilization can involve only certain members of the same species, a condition very similar to separation of the sexes even though the cross-fertilizing individuals are both hermaphroditic.

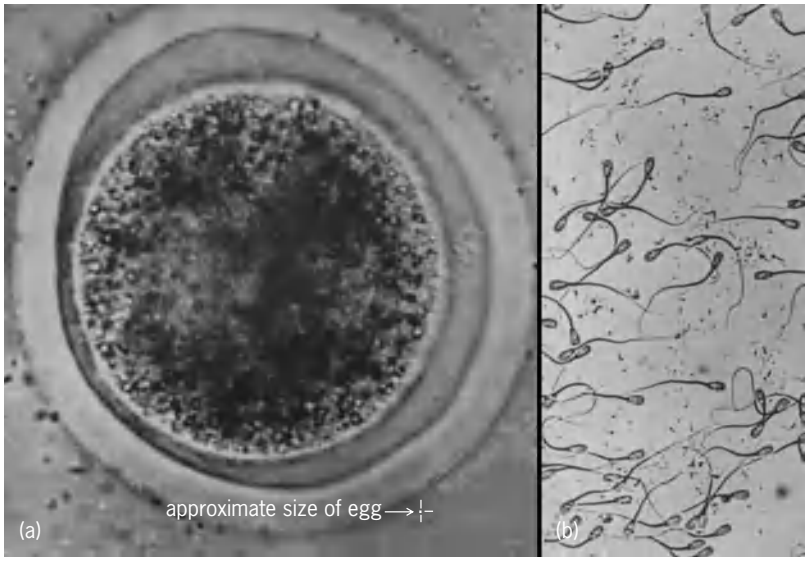

Fig. 1. Bovine gametes. (*a*) Egg. (*b*) Sperm. (*From R. C. Cook, A one-celled cow, J. Hered., 23(5):193–199, 1932*)

Chromosomal sex differentiation. Sex differentiations are often accompanied by consistent chromosomal dimorphisms, leading to the presumption that the chromosomal differences are related to, and possibly responsible for, the sex differences. Indeed, the chromosomes that are not alike in the two sexes were given the name sex chromosomes very soon after their discovery. Some workers use the term "heterosomes" to distinguish them from the autosomes,

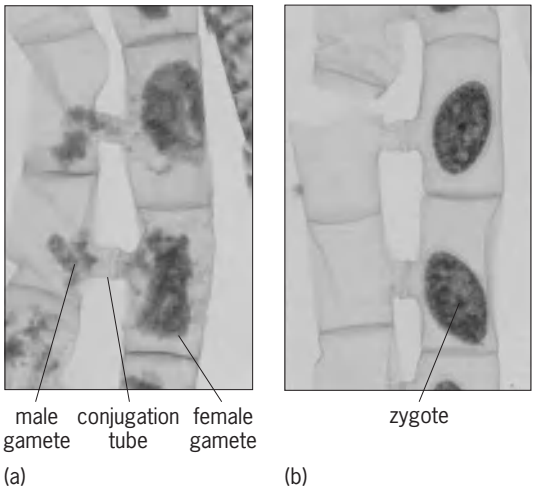

Fig. 2. *Spirogyra* (watersilk), showing the union of isogametes. (*a*) The condensed protoplasts in the cells on the left are functioning as male gametes that are migrating through the conjugation tubes to the stationary female gametes on the right. (*b*) Zygotes have been produced in the cells as a result of fusion of gametes. (*Adapted with permission from K. R. Stern, S. Jansky, and J. E. Bidlack, Introductory Plant Biology, 9th ed., McGraw-Hill, 2003*)

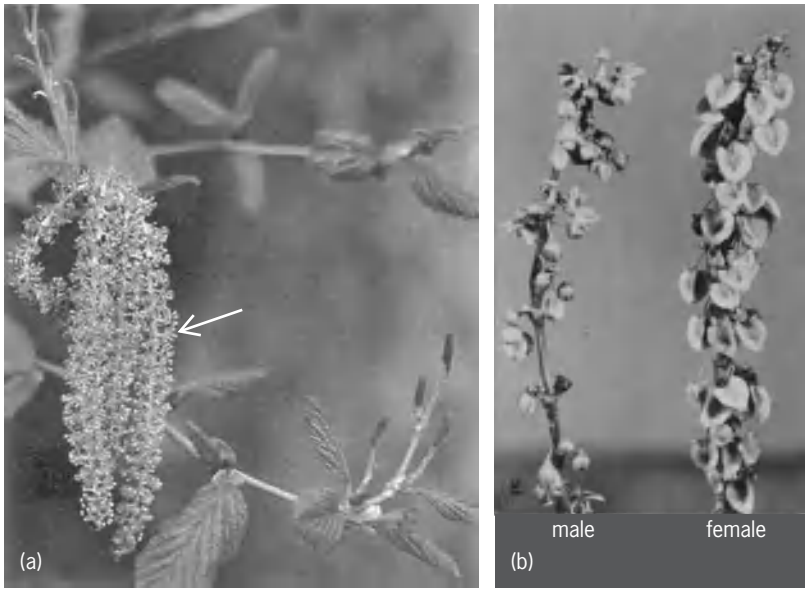

Fig. 3. Sexuality in plants. (*a*) Staminate and pistillate flowers of a birch (*Betula* sp.). Birches are monoecious; their staminate flowers hang down in long, yellowish tassels, while their pistillate flowers mature into clusters of small, brownish, conelike structures (*reprinted with permission from P. H. Raven and G. B. Johnson, Biology, 6th ed., McGraw-Hill, 2002*). (*b*) Dioecious plant *Rumex hastatulus*, with stamens and pistil located on separate plants (*courtesy of B. W. Smith*).

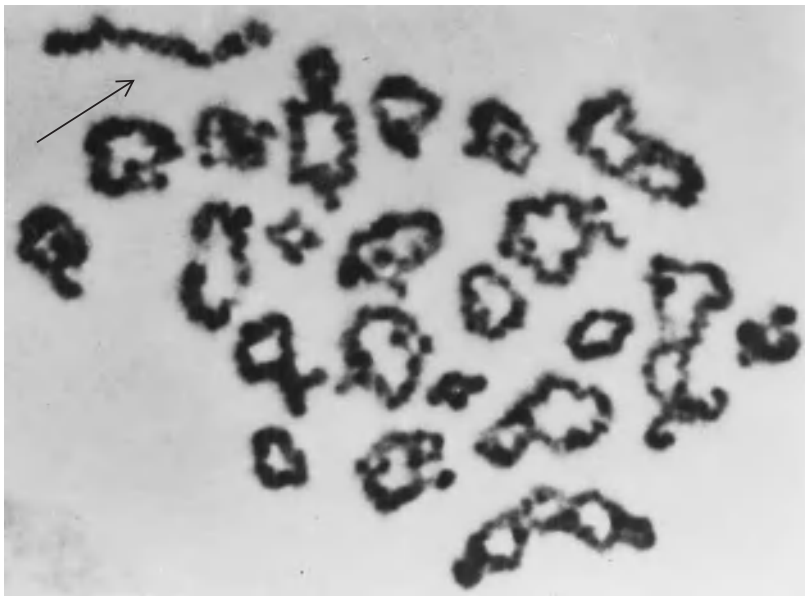

Fig. 4. Evidence of end-to-end pairing between X and Y chromosomes (arrow) during human spermatogenesis during diakinesis, the last stage of the first meiotic prophase in the male. Chromosome pairing is an indication of the presence of regions of homology. The figure shows the end-to-end pairing between X and Y chromosomes, reflecting the very short region of homology between them. (*From M. Levitan, Textbook of Human Genetics, 3d ed., Oxford University Press, 1988*)

which are the chromosome pairs that are morphologically identical to each other in both sexes.

In most species, one of the sex chromosomes, the X chromosome, normally occurs as a pair in one gender but only singly in the other. The gender with two X chromosomes is known as the homogametic sex, because each gamete normally receives an X chromosome after meiosis. The gender with only one X chromosome generally also has a morphologically different sex chromosome, the Y chromosome. The X and Y chromosomes usually pair to some extent at meiosis (**Fig. 4**), with the result that the XY is the heterogametic sex, with half its gametes containing an X and half containing a Y.

Although the earliest demonstrations of sex chromosomes were based on differences in gross morphology, geneticists soon noted that the fundamental dimorphism of X and Y chromosomes lay in their genic contents: X chromosomes of the species share homologous loci, just as do pairs of autosomes, whereas the Y chromosome usually has few, if any, loci that are also represented on the X. Thus X and Y chromosomes are sometimes very similar in shape or size but are almost always very different in genetic materials.

In the larger bisexual organisms, females are normally XX and males are XY. This is true in mammals, including humans (**Fig. 5**); in many insects, including most species in the orders Coleoptera and Diptera as well as many species of the orders Heteroptera, Dermaptera, and Neuroptera; and in many dioecious plants, including gingkos, poplars, and willows. Some organisms have a similar system of homogametic, XX, females and heterogametic males, but the males lack a Y chromosome, and their diploid number is odd, being one less than that of females. Such an XX:XO arrangement (O indicating the absence of a Y or X chromosome) occurs in many insects of the orders Orthoptera, Homoptera, and Heteroptera. Y chromosomes are commonly absent also in spiders and nematodes.

The reverse system (that is, the males are homogametic and the females heterogametic) occurs in

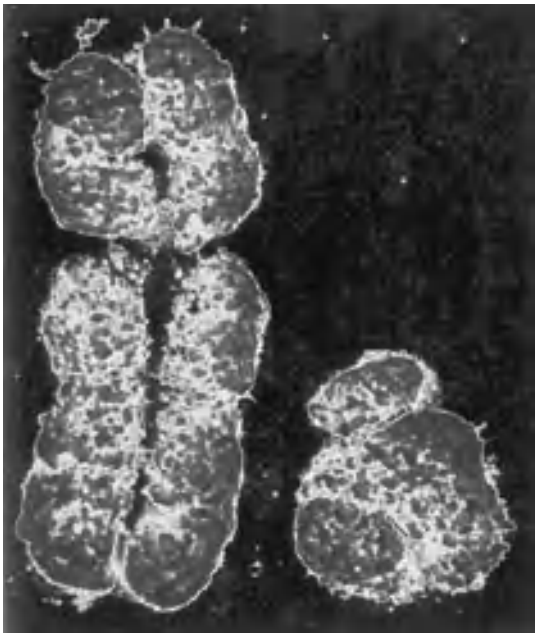

Fig. 5. The X and Y chromosomes. In humans, females are the homogametic sex (XX) and males are the heterogametic sex (XY). (*Reprinted with permission from R. Lewis, Human Genetics: Concepts and Applications, 5th ed., McGraw-Hill, 2003*)

the insect order Lepidoptera (moths and butterflies), all birds, the blood fluke (*Schistosoma* douthitti) and, among plants, the strawberry. Some moth females are XO instead of XY. The XX males and XY females are less exclusively found in the insect orders Diptera and Odonata (dragonflies), crustaceans, fishes, amphibians, and reptiles. Instead of XX and XY, the sex chromosomes in these organisms are sometimes termed ZZ and ZW, respectively. In some fishes, female heterogamety is the rule in "domesticated" forms, but wild members of the same species exhibit male heterogamety. Many more complex systems of sex chromosomes exist, particularly species with more than one kind of X or more than one kind of Y. These systems may be classified formally as X_1X_2Y, $X_1X_2X_3Y$, X_1X_2O, XY_1Y_2, and so forth, according to the number of nonhomologous (or only partly homologous) chromosomes in the heterogametic sex. The spiders and nematodes belong to this group. Since the species with these systems are usually closely related to species with a simple sex chromosomal difference (XX:XY or XX:XO), the more complex systems are thought to have evolved by means of chromosomal interchanges between sex chromosomes and autosomes that in effect converted autosomes, or parts of them, into sex chromosomes. *See* CHROMOSOME.

Genetic basis of sex differentiation. Even organisms having the same form of chromosomal dimorphism often differ in the genetic basis for sex determination. Thus, although both humans and *Drosophila melanogaster* have the XX female, XY male system, a fly with two or more X's and a Y is a female (usually fertile), whereas humans that are XXY, XXXY, XXXXY, and so on are phenotypic males (though usually infertile). In contrast, an XO *Drosophila* is a sterile male, whereas an XO human is a phenotypic female, but also sterile and with gonadal dysgenesis.

Y-dominant system. The major factor in sex differentiation in humans is a locus on the short arm of the Y chromosome designated SRY or SrY (for sex-determining region of the Y). This comparatively small gene contains no introns and encodes for a protein with only 204 amino acids. The protein appears to be a deoxyribonucleic acid (DNA)-binding type that causes somatic cells of the developing gonad to become Sertoli cells that secrete a hormone, Müllerian inhibiting substance (MIS), that eliminates the Müllerian duct system (the part that would produce major female reproductive organs). The gonad is now a testis, and certain cells in it become the Leydig cells that produce testosterone, which causes the primordial Wolffian duct system of the embryo to develop the major male reproductive organs (**Fig. 6**). If no MIS is produced, further development of the Müllerian duct structures occurs, and in the absence of testosterone the Wolffian ducts disappear, producing the normal female structures. Embryos lacking SRY or having mutated forms of it normally become females even if they are XY.

The presence of SRY does not ensure the male phenotype. XY's with testicular fertilization, for example, are phenotypic females, lacking ovaries and a uterus because a mutant X-chromosome gene results in absent or defective androgen receptors. However, phenotypic males that are apparently XX have been found to contain SRY that had been translocated elsewhere in their genome. In a number of instances, SRY came to reside in the X chromosome derived from the father as a result of crossing-over in the "pseudoautosomal" region in which the small amount of pairing between the X and Y chromosomes occurs.

This system of sex determination is called Y-dominant. It appears to be characteristic of almost all mammals, even marsupials, among others. In Y-dominant salamanders of genus *Axolotl*, the pertinent genes are female-determining, as the males are the homogametic sex. *See* MUTATION.

While SRY is the primary gene, many other genes, both autosomal and X-chromosomal, are involved in the course of developing the two sexes in mammals, as manifested by disorders in gender development or sex reversal when they mutate. Many genes remain to be fully elucidated, and more probably remain to be discovered. Some, referred to as *SOX* genes, are autosomal and X-linked loci that are SRY-related in that they a share with *SRY* a 79-amino-acid sequence

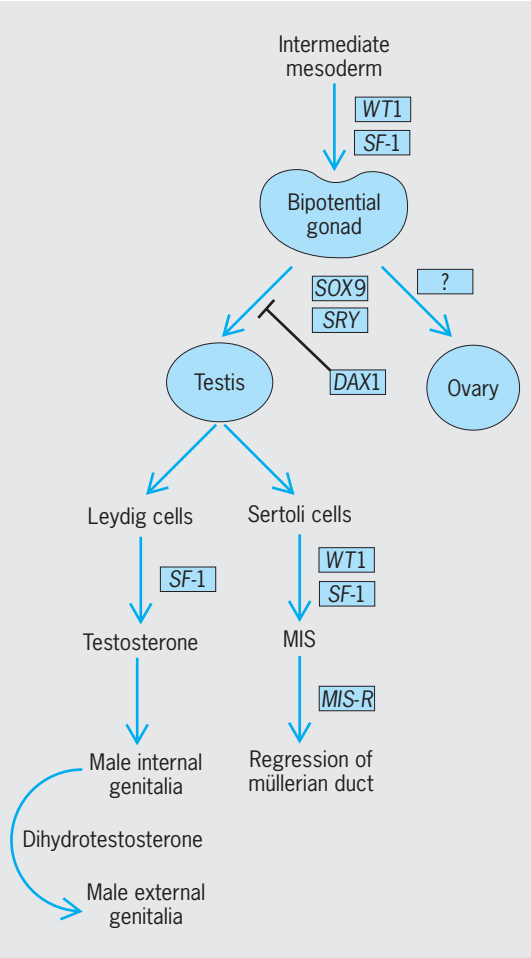

Fig. 6. *(Reprinted with permission from D. Federman, Three Facets of Sexual Differentiation, NEJM, 350:323–324, 2004. © 2004 Massachusetts Medical Society. All rights reserved)*

(the HMG box); they are believed to be involved in gonadal function, spermatogenesis, and development of the nervous system (Fig. 6). Another gene, *WT*, whose mutated form causes Wilms' tumor of the kidneys, is expressed in Sertoli cells and thought to be normally involved with the action of *SRY* (Fig. 6), as homozygotes for complete absence of the normal allele fail to form gonads in mice. An X-chromosome gene, *Dax1*, appears to be an antagonist of *SRY*; that is, it can interfere with testis formation when present in double dose in XY individuals, and it may normally be important in ovary development (Fig. 6). Female development also involves an autosomal gene called *Wnt-4*, one of a group of similar genes that have critical signaling actions during embryogenesis in animals as diverse as the worm *Caenorhabditis elegans, Drosophila*, mice, and humans.

Balance theory. The Y-dominant mechanism would not fit the *Drosophila* data, the XX/XO situations, or most of the more complex sex chromosome complements. These are best explained by the so-called balance theory. This hypothesis holds that the sexual phenotype in species with male heterogamety depends on a balance between many female-determining genes on the X-chromosome and male-determining genes on the autosomes, often referred to as the X:A ratio. It is essentially a ratio of the relative quantities of proteins determined by the X-chromosome and autosomal genes involved. When at least two X's are present together with a normal complement of autosomes, then X:A = 1, and the individual would develop as a female. In the presence of only a single X, X:A = 0.5; however, the autosomal genes prevail, and the result is a male.

In *Drosophila melanogaster* the so-called master switch in this mechanism is a gene called sex-lethal (*Sxl*). Very early in development, around the blastomere stage, if X:A = 1, *Sxl* is activated, and the result is a female; if X:A = 0.5, *Sxl* is silent, and the result is a male. Activated *Sxl* produces a protein which controls the degree of transcription of other genes, such as the normal alleles of *dsx* (doublesex) and *msl* (male-specific lethal), in development of the gonad and female characteristics, and in dosage compensation. Much remains to be worked out about the exact nature and action of all the genes involved.

When X:A balance determination occurs in female heterogamety, the female-determining genes would be the ones on the autosomes, and vice versa. The balance method is used also by another well studied species, the nematode worm, *Caenorhabditis elegans*. The master switch is again an X-linked gene, *xol-1*, but, unlike in *Drosophila*, it is activated when the *male* X:A ratio is 0.5, and repressed when the ratio X:A = 1.

Theoretically, the balance hypothesis would be compatible with a hypothesis of single genes on the X chromosomes or the autosomes being responsible. Indeed, a number of autosomal mutations are known in *Drosophila* that transform chromosomally female individuals into normal, albeit sterile, males or into intersexes. Experiments have shown, however, that the number of female-determining genes on the X chromosome is large and that they exist on all regions of the chromosome. The balance hypothesis implies that the Y chromosome of many organisms contributes little, if anything, to sex determination. However, in certain moth species all the female-determining genes are thought to be on the Y chromosome (moth females being XY) and the male-determining genes are on the X, but the primary mechanism is still one of balance. *See* GENE; GENE ACTION.

Because the sex determination methods of *D. melanogaster* and *C. elegans* have been the best studied at the genic level, it was assumed that the balance method was more widespread than has proved to be the case. Although *Sxl* has been found to be generally present even in Diptera outside the genus *Drosophila* (Muscids, for example), it has been determined that it does not act as a master switch in these organisms, leading to the realization that "Y-dominance" or some modification of it, rather than X:A balance, is probably the predominant form of sex determination.

Other systems. In many of the social insects, such as the honeybees, ants, and wasps, diploid individuals become females, while haploid ones, which develop parthenogenetically from unfertilized eggs, become males. Sex chromosomes are not identifiable as such, but one of the chromosomes contains a locus with multiple alleles that is critical to sex determination. Heterozygosity at this locus results in a viable female, whereas homozygosity results in males that usually (or in some species, always) die early or are sterile; only those with a single allele develop as viable males.

A single locus that determines sex has also been demonstrated in garden asparagus. Usually dioecious, occasionally viable seeds are produced by self-fertilization in staminate plants that also have rudimentary pistils. The offspring of these seeds produce staminate and pistillate plants in mendelian ratios that are typical for segregation in a heterozygote for a single pair of alleles; the gene for staminate is dominant to the one for pistillate.

In bacteria that can undergo recombination, the process resembles conjugation in the green alga. Meiosis does not occur, however, so that the exchange takes place directly between the transferred genetic material and the chromosome of the recipient cell. The capacity to donate genetic material in this process (maleness) derives from the presence in some cells of an F (fertility) factor, that is, a self-duplicating circular piece of deoxyribonucleic acid (DNA). The F factor is an episome: that is, it may exist free in a cell or be integrated into the chromosome. Strains where it is integrated are known as Hfr (high frequency of recombination) because they undergo conjugation about a thousand times as often as strains in which the F factor is free-standing. The presence of an F factor brings about a change of morphology not unlike a secondary sex characteristic; the resultant male produces appendages, called sex

pili or fimbriae, which attach to a female cell (one lacking an F factor) and provide the cytoplasmic bridge for transfer of DNA. *See* BACTERIAL GENETICS.

Environmental influence. In some organisms, sex determination seems to depend on the environment rather than on genes or chromosomes. In the sea worm *Dinophilus*, large eggs become females and small eggs become males. In another sea worm, *Bonnelia*, sex seems to depend on where the free-swimming larvae land: those that land on the sea bottom develop into females, each with a long proboscis; those that land on this proboscis develop into tiny males that live as parasites in the genital ducts of the female. In the horsetail plant, *Equisetum*, sex depends on growth conditions: if conditions are good, the plants become female, and if conditions are poor, they are male. Similarly, in many fishes and reptiles the temperature at which the eggs develop determines the resultant sex ratio. In alligators, for example, warm temperatures result in an excess of males, whereas cool temperatures produce an excess of females. In a number of turtle species the opposite is true, with females predominating in the warmer environment.

The internal environment also must play an important role in sex determination much the same way that such factors determine how some cells of the pluripotential zygote become nervous structures whereas others become the linings of the gut. This is particularly likely in hermaphroditic animals and monoecious plants, and the many organisms, such as the algae, where the male and female tissues seem to have the same genes and chromosomes.

X inactivation. Development of the Y-dominant pattern of sex determination in which the Y chromosome is almost entirely devoid of genes that are counterparts (homologs) of those on the X chromosome posed a serious developmental problem: Since the genome is effectively balanced when there is a single X chromosome with a diploid set of autosomes, as occurs in XY and XO individuals, one would expect that having a second X chromosome would be detrimental to the homogametic sex because of the resultant imbalance. In other words, it is a problem of dosage compensation.

The mammalian solution of the problem was discovered by the British geneticist Mary Lyon, and is referred to as the Lyon hypothesis. From her work with mice, she concluded that one of the X chromosomes in somatic tissues of the XX female is inactivated early in embryogenesis, so that in effect the genes on only one X chromosome are active, just as in the heterogametic sex. Both the active and inactive X's are faithfully reproduced in subsequent mitoses.

The inactivated X is the late-replicating one. Since inactivation is, in most cases, a random process, in some cell lineages the paternal, while in others the maternal, X chromosomes are the ones affected. Hence, mammalian females that are heterozygous for X-linked loci are usually mosaics for them. The exception is the locus for steroid sulfatase, which is never inactivated.

X inactivation appears to occur in all mammals, even the monotremes (such as Platypus), though there is some variability in the amount of X chromosome and the type of cells that escape inactivation. In humans with supernumerary X's (for example, XXXY males or XXX females), all X chromosomes in excess of one are usually inactivated, correlating with the number of heterochromatic sex chromatin (Barr) bodies seen in their somatic cells.

Inactivation begins at an inactivation center in band 13 on the long arm of the chromosome (Xq13) and spreads upstream and downstream from it, extending to normally autosomal cells that have been traslocated to the portion of the X chromosome bearing the inactivation center (but not to the X chromosome genes on the portion of the translocated X that lacks Xq13). The part of Xq13 involved has been narrowed down to a gene called Xist (for inactive specific transcript), which governs the production of an RNA (but no protein) that remains in the nucleus and binds to the inactive X. Apparently it acts to block transcription from the inactive X. The corresponding gene on the active X appears to be silenced by methylation. *See* HUMAN GENETICS. M. Levitan

Bibliography. J. S. Bogan and D. C. Page, Ovary? testis?: A mammalian dilemma, *Cell*, 76:603–607, 1994; J. J. Bull and E. L. Charnov, Enigmatic reptilian sex ratios, *Evolution*, 43:1561–1566, 1989; K. M. Cadigan and R. Nusse, *Wnt* signaling: A common theme in animal development, *Genes Develop.*, 11:3286–3305, 1997; T. W. Cline, Autoregulatory functioning of a *Drosophila* gene product that establishes and maintains the sexually determined state, *Genetics*, 107:231–277, 1984; T. W. Cline and B. J. Meyer, Vive la difference: Males vs. females in flies vs. worms, *Annu. Rev. Genet.*, 30:637–702, 1996; J. Godwin, J. A. Luckenbach, and R. J. Borski, Ecology meets endocrinology: environmental sex determination in fishes, *Evcl. Develop.*, 5:40–49, 2003; L. O. F. Penalva and L. Sanchez, RNA binding protein Sex-lethal (*Sxl*) and control of *Drosophila* sex determination and dosage compensation, *Microbiol. Mol. Biol. Rev.*, 67:343–359, 2003; G. Saccone, A. Pane, and L. C. Polito, Sex determination in flies, fruitflies, and butterflies, *Genetica*, 116:15–23, 2002; D. C. A. Shearman, The evolution of sex determination systems in dipteran insects other than *Drosophila, Genetica*, 116:25–43, 2002; J. Weiss et al., *Sox3* is required for gonadal function, but not sex determination, in males and females, *Mol. Cell. Biol.*, 23: 8084–8091, 2003.

Sex-linked inheritance

The inheritance of a trait (phenotype) that is determined by a gene located on one of the sex chromosomes. These traits, including many diseases, have characteristic patterns of familial transmission; for a century these patterns were used to localize the genes for these traits to a specific chromosome. As the genomes of many organisms have now been

completely sequenced, reliance on the specific pattern of inheritance to map genes on sex chromosomes has waned.

Sex determination. The expectations of sex-linked inheritance in any species depend on how the chromosomes determine sex. For example, in humans, males are heterogametic, having one X chromosome and one Y chromosome, whereas females are homogametic, having two X chromosomes. In human males, the entire X chromosome is active (although not all genes are active in every cell), whereas one of a female's X chromosomes is largely inactive. Random inactivation of one X chromosome occurs during the early stages of female embryogenesis, and every cell that descends from a particular embryonic cell has the same X chromosome inactivated. The result is dosage compensation for X-linked genes between the sexes (the Lyon hypothesis). A specific gene on the long arm of the X chromosome, called *XIST* at band q13, is a major controller of X inactivation. This pattern of sex determination occurs in most vertebrates, but in birds and many insects and fish the male is the homogametic sex. *See* SEX DETERMINATION.

Sex-linked traits. In general terms, traits determined by genes on sex chromosomes are not different from traits determined by autosomal genes. Enzymes, structural proteins, receptor proteins, and so forth are encoded by genes on both autosomes and sex chromosomes. Sex-linked traits are distinguishable by their mode of transmission through successive generations of a family. The term sex linkage was introduced by T. H. Morgan in 1914, based on his Nobel prize–winning investigations of *Drosophila*. His experiments with the X chromosome of fruit flies provided the first proof of linkage of any type, the first linkage maps, convincing evidence that heritable traits are encoded on chromosomes, and the concept that x-rays produce heritable changes (mutations) by disrupting the chromosome.

Sex-linked inheritance has been studied extensively in humans, for whom it is preferable to speak in terms of X-linked or Y-linked inheritance.

X-linked traits. Red-green color blindness was the first human trait proven to be due to a gene on a specific chromosome. This was accomplished by E. B. Wilson in 1911 when he correctly interpreted family studies showing X-linked recessive inheritance. The characteristics of this pattern of inheritance are readily evident (**Fig. 1***a*). Males are more noticeably or severely affected than females; in the case of red-green color blindness, women who have one copy of the mutant gene (that is, are heterozygous or carriers) are not at all affected. Among offspring of carrier mothers, on average one-half of their sons are affected, whereas one-half of their daughters are carriers. Affected fathers cannot pass their mutant X chromosome to their sons, but do pass it to all of their daughters, who thereby are carriers. A number of other well-known human conditions such as XY gonadal dysgenesis behave in this manner, including the two forms of hemophilia, Duchenne muscular dystrophy, the fragile X syndrome that is a common cause of mental retardation, and glucose-6-phosphate dehydrogenase deficiency, which predisposes to hemolytic anemia. *See* ANEMIA; COLOR VISION; HEMOPHILIA; MUSCULAR DYSTROPHY.

Human color vision is based on expression of blue, green, and red pigments in specific cells of the retina. The gene for the blue pigment is located on chromosome 7, and the genes for the red and green pigments are closely linked at the end of the long arm of the X chromosome (**Fig. 2**), exactly where classic family studies of red-green color blindness had predicted them to be. In people with normal color vision, there is one copy of the red pigment gene followed by a variable number of green pigment genes in tandem array. Various degrees of red-green color blindness result from disruption of one or more of these genes, primarily by the process of unequal recombination. *See* RECOMBINATION (GENETICS).

Y-linked traits. A number of Y-linked conditions (such as Sertoli-cell only syndrome) account for a substantial percentage of infertility in males. Y-linked traits never occur in females and are passed to every son of an affected father (Fig. 1*b*). Variations in DNA sequence along the Y chromosome, which are often completely common and harmless (polymorphisms), are also passed only from fathers to sons. Such polymorphisms have been extremely useful to anthropologists in studying migration patterns and male lineages.

Advances in mapping. Refined cytogenetic and molecular techniques have supplemented family

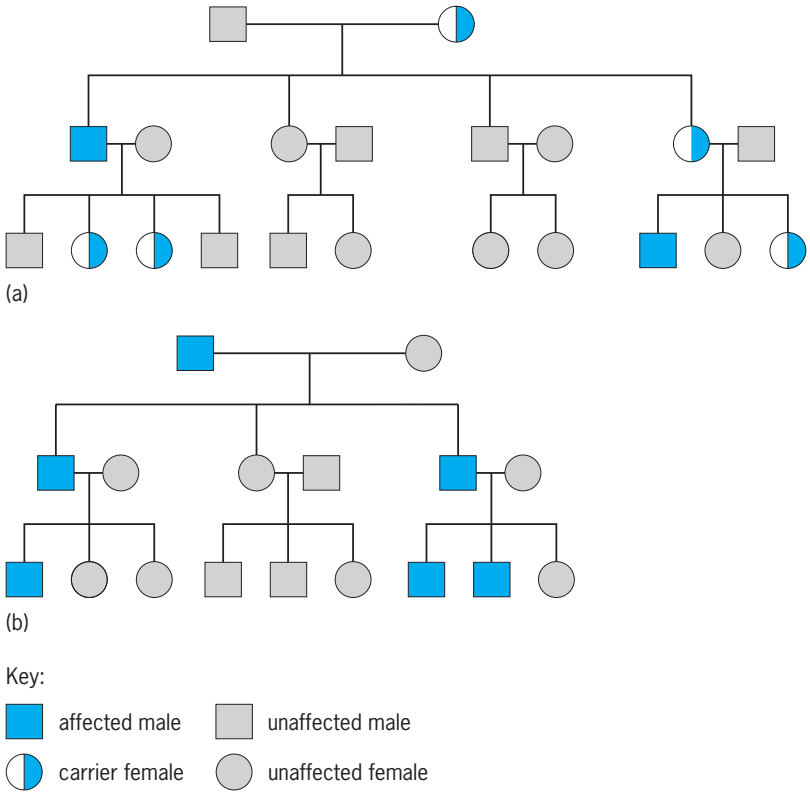

Fig. 1. Patterns of sex-linked inheritance in humans. (*a*) X-linked recessive inheritance. (*b*) Y-linked inheritance.

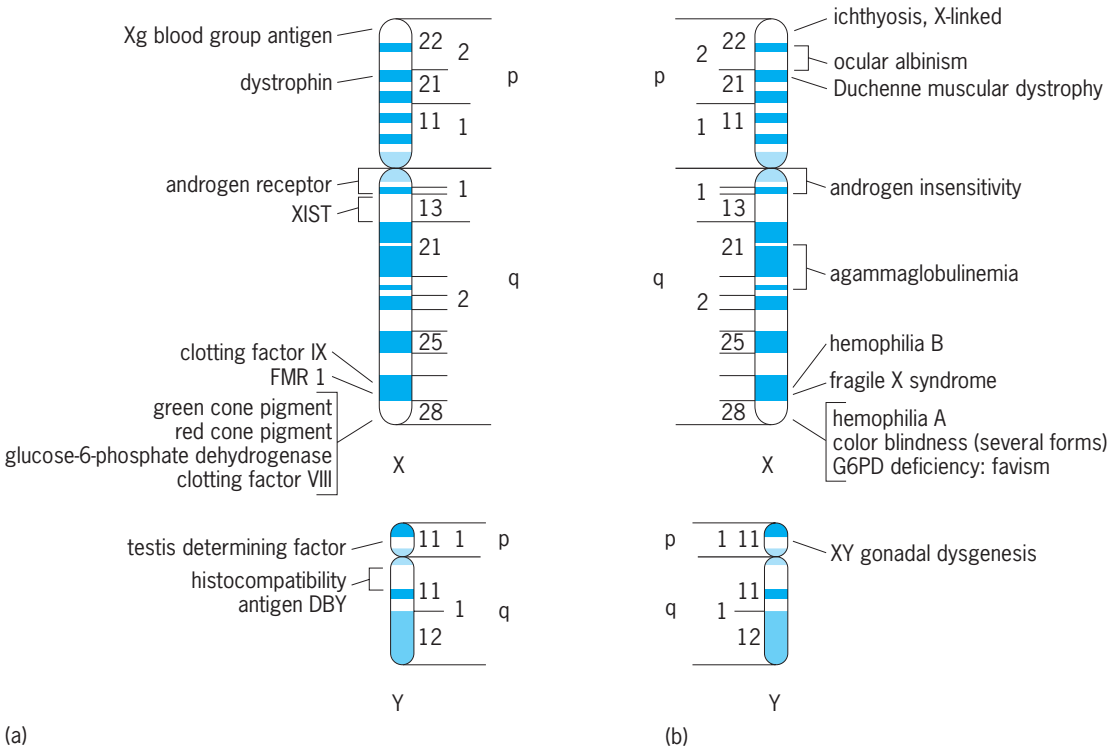

Fig. 2. Genetic maps of the human sex chromosomes. (*a*) Genes and (*b*) diseases mapped to specific chromosome regions.

studies as a method for characterizing sex-linked inheritance and for mapping genes to sex chromosomes in many species. Over 500 human traits and diseases seem to be encoded by genes on the X chromosome, and nearly 500 genes have been mapped. Among mammals, genes on the X chromosome are highly conserved. Thus, identifying an X-linked trait in mice is strong evidence that a similar trait, and underlying gene, exists on the human X chromosome. The human Y chromosome harbors about 50 genes. *See* GENETICS; HUMAN GENETICS.

Reed E. Pyeritz

Bibliography. J. Brennan and B. Capel, One tissue, two fates: Molecular genetic events that underlie testis versus ovary development, *Nat. Rev. Genet.*, 5:509–521, 2004; M. Kohn et al., Wide genome comparisons reveal the origins of the human X chromosome, *Trends Genet.*, 20:598–603, 2004; C. Krausz et al., Y chromosome polymorphisms in medicine, *Ann. Medi.*, 36:573–583, 2004; L. Ronfani and M. E. Bianchi, Molecular mechanisms in male determination and germ cell differentiation, *Cell Mol. Life Sci.*, 61:1907–25, 2004.

Sextant

A navigation instrument used for measuring angles, primarily altitudes of celestial bodies. Originally, the sextant had an arc of 60°, or one-sixth of a circle, from which the instrument derived its name. Because of the double-reflecting principle used, such an instrument could measure angles as large as 120°.

In modern practice, the name sextant is commonly applied to all instruments of this type regardless of the length of the arc, which is seldom exactly 60°. Occasionally, the terms octant, quintant, and quadrant are applied to instruments having arcs of 45°, 72°, and 90°, respectively.

The optical principles of the sextant are similar to those of the prismatic astrolabe. *See* PRISMATIC ASTROLABE.

Development. John Hadley, an Englishman, and Thomas Godfrey, an American, are credited with inventing the marine sextant independently. The sextant has remained virtually unchanged since each man designed similar instruments in 1730. Various, less accurate, more cumbersome angle-measuring devices preceded the sextant, including the common quadrant, astrolabe, cross-staff, backstaff or sea

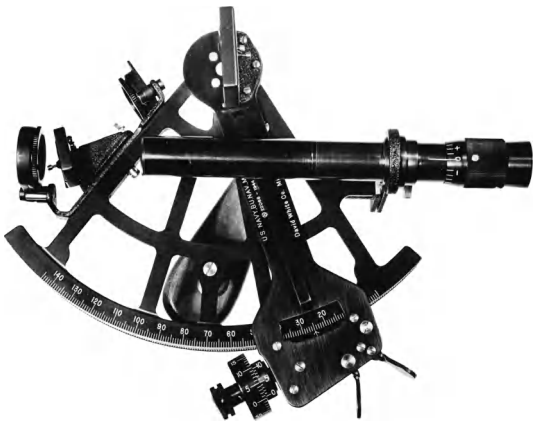

Fig. 1. Marine sextant.

Fig. 2. Periscopic sextant for use in aircraft.

quadrant, and nocturnal. Tycho Brahe (1546–1601) invented several instruments with arcs of 60°, which he called sextants. In 1700, Isaac Newton designed an instrument incorporating the double-reflecting principle, but his invention was not made public until after sextants had been constructed by Hadley and Godfrey.

Classes. Modern instruments may be grouped into two classes, marine sextants and air sextants.

Marine sextant. This is an instrument designed for use by mariners. It utilizes the visible sea horizon as the horizontal reference. It is equipped with a removable telescope to magnify the image of the horizon, and with shade glasses or filters to reduce glare (**Fig. 1**). The marine sextant has been widely used as a symbol of navigation.

Air sextant. An instrument designed for use in aircraft is called an air sextant. Such sextants vary widely in design, but all modern instruments have built-in artificial horizons. A bubble is most commonly used for this purpose, but pendulums and gyroscopes have also been used. Most modern air sextants are periscopic to permit observation of celestial bodies without need of an astrodome in the aircraft (**Fig. 2**).

Sextant altitude corrections. Altitudes measured with the sextant are subject to certain corrections. These may be classified as those related to (1) inaccuracies in reading, (2) inaccuracies in horizontal reference, (3) bending of the ray of light from the body, (4) adjustment to equivalent reading of the center of the body, and (5) adjustment to the equivalent reading at the center of the Earth. *See* CELESTIAL NAVIGATION.

Alton B Moody

Bibliography. N. Bowditch, *American Practical Navigator*, U.S. Defense Mapping Agency Hydrographic Topographic Center Publ. 9, 1977; E. S. Maloney, *Dutton's Navigation and Piloting*, 14th ed., 1985.

Sexual dimorphism

Any difference, morphological or behavioral, between males and females of the same species. In many animals, the sex of an individual can be determined at a glance. For example, roosters have bright plumage, a large comb, and an elaborate tail, all of which are lacking in hens. Sexual dimorphism arises as a result of the different reproductive functions of the two sexes and is a consequence of both natural selection and sexual selection. Primary differences, such as the structure of the reproductive organs, are driven by natural selection and are key to the individual's function as a mother or father. Other differences such as the peacock's (*Pavo cristatus*) enormous tail are driven by sexual selection and increase the individual's success in acquiring mates. *See* ORGANIC EVOLUTION.

Reproduction. A less obvious sexual dimorphism is the difference in size of male and female gametes. In animals, females produce eggs (ova) and males, sperm. In plants, females produce ovules and males, pollen. Fusion of egg and sperm is required for production of offspring. In nearly all cases, the sperm (or pollen) are substantially smaller and more numerous than the ova. Eggs are large because they contain nutrients essential for development of the embryo. However, the sole purpose of sperm is to fertilize the egg. Sperm do not contain any nutrients and can therefore be small. For the same investment of nutrients, a male can produce more sperm than a female can produce eggs. Human males, for example, produce about 300 million motile sperm per ejaculate, whereas females normally produce only one egg (30,000 times larger than a single sperm) per month. *See* GAMETOGENESIS.

In order to perform their sexual function, males have testes in which sperm are produced. In animals with internal fertilization, males often have an intromittent organ (penis) to transfer sperm to the female. Females often have a more complex reproductive system. As well as ovaries producing eggs, female mammals have a uterus, placenta, and mammary glands. Female insects have sperm storage organs and an ovipositor for laying eggs. *See* FERTILIZATION; INSECTA; REPRODUCTION (ANIMAL); REPRODUCTIVE SYSTEM.

Body size. In nearly all animal groups (apart from mammals and birds), females are larger than males because bigger females tend to produce more eggs. Therefore large female size is favored. For example, in the Atlantic silverside fish (*Menidia menidia*), large mothers lay more eggs than do small mothers, and the same is true for the Edith's checkerspot butterfly (*Euphydryas editha*). Males of many species, such as weevils, prefer to mate with bigger females. However, size does not appear to limit males' ability to produce sperm. The male blanket octopus (*Tremoctopus violaceus*) weighs only 0.25 g and is dwarfed by the female which may be up to 10,000 times heavier. *See* OCTOPUS.

In contrast, among mammals and birds males are generally the larger sex. For example, male great bustards (*Otis tarda*, a rare ground-living bird) are more than twice as heavy as females, and male elephant seals (*Mirounga angustirostris*) weigh eight times as much as the females. The reason that male elephant seals are so large is that they fight for access

to females, with only the largest male getting to mate with all the females on the beach. *See* PINNIPEDS.

In a few species, males may be either large or small but rarely of intermediate size. The males of the Pacific coho salmon (*Onchorhynchus kisutch*) come in two sizes. Large hook-nosed males fight over females, whereas inconspicuous satellite males sneak in to fertilize the eggs, undetected by their large rivals. *See* SALMONIFORMES.

Differences in body size and shape can be caused by factors other than reproductive success. Sexual dimorphism can arise as a consequence of competition between the sexes over resources, or because the sexes use different resources. For example, in many species of snake, males and females use different habitats and eat different food, which has led to differences in their head shape and feeding structures. Spider females often cannibalize their partners during mating. Being small is advantageous for the male, as he can sneak up and mate with the female without prior detection, and his small size may also make him less tempting as a meal.

Weapons. Some males use weapons to compete with each other for access to mates. Deer and antelope use their antlers to fight over females, as do stag beetles (*Lucanus cervus*) and narwhals (*Monodon monoceros*). Elephants fight with their tusks, and pheasants fight using leg spurs. Females of these species do not fight over mates and therefore lack these weapons.

Ornamentation in animals. Males and females also show dimorphic structures that are not involved in combat but are used in mate choice. Female reproductive capacity is limited by resources (both egg production and pregnancy are energetically costly), whereas male reproductive capacity depends largely on how many females he mates with. Males tend to compete with each other for access to as many females as possible, while females look for the best candidate among different males.

Female preference for showy males can lead to exaggerated male ornaments. For example, the peacock has an impressively large tail which he uses in an elaborate dance to impress the drab peahen. Bright plumage and showy tails, specifically long tails, are widespread ornaments in male birds. Females tend to prefer the male with the most elaborate tail because only males of high genetic quality are capable of producing showy tails. Therefore, by mating with a male with a showy tail, the female will have high-quality offspring. Tail length may be an honest indication of the male's ability to resist parasites which will be passed on to the offspring. For example, male barn swallows (*Hirundo rustica*) with long tails carry fewer parasitic mites than do short-tailed males. Females prefer these long-tailed males, and their offspring inherit the ability to resist mite infection.

Female choosiness can lead to increasingly exaggerated male ornaments, even when the ornamentation interferes with his survival. The longer the peacock's tail, the more it hampers his flight and maneuverability. Further exaggeration within the species is halted when the tail finally becomes too costly to the bearer, and the peacock with the overexaggerated tail does not survive long enough to mate and pass the genes encoding this trail to the next generation.

Females may also prefer bright males even when ornamentation conveys no information about quality. Female stickleback fish are drab silver, whereas males have a red belly that varies in brightness. Females prefer mates with a bright red belly, hence bright males mate more often than dull males. Sons inherit the red color from their fathers, and daughters inherit a preference for red-bellied mates from their mothers.

While sexual selection tends to favor bright, showy males, natural selection often promotes cryptic coloration in females. For example, the male mallard duck (*Anas platyrhynchos*) has iridescent plumage to woo females. In contrast, the female is drab and inconspicuous, which helps her to avoid predators while incubating the eggs.

Ornamentation in plants. Plants also differ in showiness. Many plants bear both male and female flowers (simultaneous hermaphrodites), but male flowers are sometimes larger and more conspicuous. For example, the female catkins of willow are dull gray compared with the bright yellow male catkins, because male flowers compete with each other to attract pollinators. In plant species with separate sexes (dioecious), males tend to produce more flowers than females. For example, males of the American holly (*Ilex opaca*) produce seven times as many flowers as females in order to increase their chances of pollen transfer to females. *See* FLOWER; POLLINATION.

Vocalization and chemical signals. Animals and plants show marked sexual dimorphism in other traits. Calling, singing, pheromones, and scent marking can all be explained by competition between males and by female mate choice. *See* ANIMAL COMMUNICATION.

Vocalization. Just as with visual ornaments, male vocal calls are used in competition with other males and to attract females. The male red-winged blackbird (*Agelaius phoeniceus*) sings to defend his territory from rival males as well as to attract a female. Male bullfrogs call only during the breeding season, and their impressive croaks reveal their size to females. The silent females prefer the call of large, high-quality males. Male crickets also call to attract mates, and the longer a male calls, the more females he will attract. Unfortunately for the male, calling also announces his presence to predators and parasitoids (such as bats and flies), exemplifying how mating success and individual survival can sometimes conflict. *See* TERRITORIALITY.

Chemical signals. Chemical signals are commonly used for attracting mates in insects, attracting pollinators in plants, and territorial marking in mammals. In many nocturnal insects, males use pheromones rather than visual ornaments to attract mates. Male and female moths produce different pheromones. Females emit long-range scents to attract males. At close range, males court the female using a

different scent and females choose the sweetest-smelling male. *See* PHEROMONE.

In mammals such as domestic cats and mice, males use pheromones to stake out their territory. Males can detect the status of a rival from his scent. When choosing between mates, female mice prefer the odor of dominant males.

There is also some evidence that females prefer males with different (complementary) major histocompatibility complex (MHC) genes and that this preference is based on scent. The MHC is a large cluster of genes essential to the immune system of vertebrates. MHC-based mate choice may maximize parasite resistance in offspring and/or minimize the risks of mating with kin. Even human females can discriminate between potential partners by their smell, preferring mates with a different MHC variant to their own. Remarkably, in humans the same odor receptor genes are present both on sperm and in the nose. *See* HISTOCOMPATIBILITY.

Plants also use scent to attract pollinators and reward them with nectar. In some species, only male flowers provide nectar. In the Central American tree *Jacaratia dolichaula*, female flowers smell just as sweet as male flowers but do not provide any nectar rewards.

Behavior. Associated with morphological sexual dimorphism are several behavioral differences between males and females. Many of these are related to locating a mate, competition between males, and female choosiness.

In most animal and plant species, males invest more time and energy into finding mates. Female angler fish live in total darkness at a depth of 2000 m (6600 ft) or more. The male finds his spouse using his huge, sensitive nose. He invests enormous resources into locating a female, having the largest nose relative to his size of any vertebrate. In many insects, only the males have wings so that they can fly in search of females. Males of many species are ready to mate before females. For example, many male insects reach adulthood before females, enabling them to monopolize and mate with females as soon as they become sexually mature. Males of most migrating birds species arrive and set up territories at the breeding ground before females. Male plants tend to flower earlier in the season and more often than females. This aids pollen flow toward the scarcer, less attractive female flowers. In the shrimplike *Gammarus*, males guard females by carrying them around and waiting for them to lay their eggs. Males expend a lot of energy in carrying females for several days. As a result, unusually for invertebrates, *Gammarus* males are larger than females. *See* GAMMARIDEA.

Males may form large aggregations (leks) where they display their sexual ornamentation to onlooking females. During the breeding season, sage grouse (*Centrocercus urophasianus*) form leks of 15–20 males in which they strut and display their plumage to females. Females often visit several males before selecting a mate. Frogs also form large leks, with several hundred males calling throughout the night. The more vocal the frog, the more attractive he is to females.

Male displays are costly to produce and depend on overall condition, which is determined by many genes. For example, in dung beetles (*Onthophagus taurus*) there is genetic variation between males in courtship rate. Females prefer males with a high courtship rate, as this reflects their overall condition.

Some males even vigorously display to females when other males are not present. The male bird of paradise performs an elaborate and complicated dance to impress a female, and the male bowerbird builds an elegant structure, the bower, which he adorns with feathers and flowers to impress the female. The male bowerbird may even destroy rival males' bowers and steal their decorations to enhance the beauty of his own creation. *See* BEHAVIORAL ECOLOGY; REPRODUCTIVE BEHAVIOR.

Parental care. Animals also show sexual dimorphism relating to their roles as parents. Many parents continue to provide for their young after birth, with the female performing the bulk of the care in most species. In marsupials such as kangaroos, females have a brood pouch in which the young develop and feed on the mother's milk. Similarly, female cichlid fish protect their growing young by brooding them in their mouths. Earwig females tend the nest, keeping their young free of fungi, and scorpion mothers carry their young on their back. Female mammals suckle their young, whereas males cannot because they lack mammary glands. However, some mammals (such as gibbons and prairie voles) and many birds share parental duties, with both males and females feeding and protecting the young. *See* MATERNAL BEHAVIOR.

Sex role reversal. In some species, males invest more time and resources in their offspring than females. This reversal in sex roles is reflected in reversed sexual dimorphism. Both male sticklebacks and South American daddy longlegs (*Zygopachylus albomarginis*) construct a nest for their young, and in pipefish, males instead of females brood the young. Female pipefish are the showy sex, and males prefer to mate with females with the brightest coloration. Seahorse (*Hippocampus*) males have a special brood pouch and even a placenta to provide the growing young with nutrients. The American jacana (*Jacana spinosa*) is rare among birds because the male exclusively cares for the young and females fight with each other over males. Consequently, in contrast with most birds, female jacanas have a larger body size than males, and even possess weapons in the form of wing spurs, which are used in female-female combat.

Many male insects provide food items (nuptial gifts) to the female, which she invests in egg production. In some katydid species, females fight for access to males and their gifts, and males prefer to mate with the largest female because she will lay more eggs. In some empidid dance flies, males seduce females with a nuptial gift and only females are

ornamented. Females deceive males by inflating air sacs on their abdomen to appear larger and, therefore, more attractive.

Sex determination and sex change. For many animals, sex is determined at conception by the sex chromosomes that they inherit from each parent. In humans and most other mammals, daughters inherit two X chromosomes, one from their mother and one from their father. Sons inherit an X chromosome from the mother and a Y chromosome from the father. Conversely, male birds and butterflies inherit a Z chromosome from their mother and a Z from their father, while females receive a Z from their father and a W from their mother. These sex-chromosome differences underlie all other aspects of sexual dimorphism between males and females.

One potential problem is that males and females have different numbers of X (or Z) chromosomes. Sex chromosome–linked genes will therefore be expressed at different levels in the two sexes (for example, twice as much in XX females), which can cause developmental disruption. This problem is solved through dosage compensation, a regulating mechanism ensuring that genes are equally expressed in males and females.

Many animal species have a default sex regulated by the sex chromosomes. For example, in humans the default sex is female, and male development is initiated by testosterone production during embryonic development. In contrast, the default sex in the fruit fly (*Drosophila melanogaster*) is male. The sexual phenotype is regulated by a genetic switch that controls other genes by turning them on or off; when this switch is turned on, the fly develops as a female.

More rarely, sex is determined by environmental conditions experienced during development (environmental sex determination). For example, in the Atlantic silverside fish, eggs laid early in the season become females, and eggs laid later, males. This is because individuals laid early have more time to grow, and large size is favored in females since they can produce more eggs. The peanut worm (*Bonellia viridis*) shows extreme sexual size dimorphism that is associated with environmental sex determination. Larvae that settle on the muddy bottom become females and burrow in the mud, feeding through a long proboscis. A larva that settles on or near a female becomes a dwarf male one-thousandth of her size and lives in her proboscis until he matures and makes a one-way journey to fertilize her eggs. Temperature even regulates sexual phenotype in some animals that have sex chromosomes, including alligators and brush turkeys (*Alectura lathami*).

Not all animals have fixed reproductive roles. Some start life as one sex and then change to the opposite sex later (sequential hermaphrodites). The clownfish (*Amphiprion akallopisos*), which lives in anemones on coral reefs, begins life as a small male. As the fish grows, it becomes a female. This sex change is favorable since size is unimportant for male mating success whereas bigger females lay more eggs. In contrast, the blue-headed wrasse (*Thalassoma bifasciatum*) male defends a territory, and only large males are successful territory holders. These fish start life as females and become males when they are large enough to defend a territory and compete successfully for mates. *See* SEX DETERMINATION.

Males and females tend to share most of their genes, yet many genes are sexually antagonistic as they benefit one sex while being costly for the other. For example, bright coloration is advantageous for male guppies in obtaining mates. In females, conspicuousness would be costly due to high predation risk, and does not confer any benefit in terms of mate attraction. One solution to the problem of sexually antagonistic genes is sex-limited expression, in which the genes are expressed only in the sex in which the trait is beneficial. One way of achieving this is to place such genes on the sex chromosomes. For example, in guppies most genes for body coloration are present only on the male sex chromosome (Y) and, therefore, are never expressed in females. *See* SEX-LINKED INHERITANCE.

Sexual dimorphism is seen throughout the animal and plant kingdoms, with males and females differing in traits ranging from gamete size to ornaments, weapons, and behavior. The driving force behind this dimorphism is the different investment made by males and females in mate attraction and offspring production. Alison M. Dunn; Nina Wedell

Bibliography. M. Andersson, *Sexual Selection*, Princeton University Press, 1994; J. J. Bull, *Evolution of Sex Determining Mechanisms*, Benjamin-Cummings, California, 1983; T. H. Clutton-Brock, *The Evolution of Parental Care*, Princeton University Press, 1991; J. R. Krebs and N. B. Davies, *An Introduction to Behavioural Ecology*, 3d ed., Blackwell Science, Oxford, 1993; R. H. Peters, *The Ecological Implications of Body Size*, Cambridge University Press, 1986.

Sexually transmitted diseases

Infections that are acquired and transmitted by sexual contact. Although virtually any infection may be transmitted during intimate contact, the term sexually transmitted disease is restricted to conditions that are largely dependent on sexual contact for their transmission and propagation in a population. The term venereal disease is literally synonymous with sexually transmitted disease, but traditionally is associated with only five long-recognized diseases (syphilis, gonorrhea, chancroid, lymphogranuloma venereum, and donovanosis) and sometimes is considered to have pejorative connotations. Sexually transmitted diseases occasionally are acquired nonsexually (for example, by newborn infants from their mothers, or by clinical or laboratory personnel handling pathogenic organisms or infected secretions), but in adults they are virtually

never acquired by contact with contaminated intermediaries such as towels, toilet seats, or bathing facilities. However, some sexually transmitted infections (such as human immunodeficiency virus infection, viral hepatitis, and cytomegalovirus infection) are transmitted primarily by sexual contact in some settings and by nonsexual means in others.

Incidence and epidemiology. The incidences of most sexually transmitted diseases increased in industrialized countries beginning in the early 1960s. During the same period, sexually transmitted infections and their complications also increased in most developing nations, where they constitute some of the major causes of morbidity and contribute importantly to socioeconomic deprivation. In the United States, the reported incidence of gonorrhea began to decline slowly after 1975, due in part to national control measures instituted early in the decade. Although this trend accelerated in the 1980s, the incidence fell more rapidly in some segments of the population than others. The incidence of syphilis also fell, but began to rise rapidly once again after 1985. The decline in syphilis in the early to mid-1980s was due largely to falling rates in homosexually active men, a population that underwent major changes in sexual behavior in response to the epidemic of acquired immunodeficiency syndrome (AIDS). The rapid rise in the late 1980s occurred primarily in urban heterosexual populations with high rates of substance abuse, poverty, and prostitution. The rise in reported rates of *Chlamydia trachomatis* infection since the mid-1980s is due primarily to increasing case detection and reporting. The true incidence of chlamydial infection probably was relatively stable in the 1980s and early 1990s, at incidence rates five- to tenfold higher than those of gonorrhea.

Several behavioral, demographic, and biologic factors have fueled the epidemic of sexually transmitted diseases in industrialized countries. The "sexual revolution" that began in the 1960s and extended into the 1980s was associated with an increasing frequency of sexual activity with larger numbers of partners, especially for young women, and with increasing reliance on the birth control pill for contraception, at the expense of barrier methods that provided partial protection against sexually transmitted diseases. The maturing of the postwar "baby boom" generation resulted in unprecedented numbers of persons aged 18 to 35 years, the time of maximal sexual activity. Changing patterns of marriage and divorce, increasing urbanization, and other social factors abetted these demographic and behavioral changes and contributed to the epidemic. In developing nations, the social disruption associated with conversion from tribal-rural to industrialized-urban societies, economic deprivation, high prevalences of prostitution, and poor diagnostic and treatment facilities have contributed to the epidemic.

Biological factors influencing the sexually transmitted disease epidemic are related both to the sexually transmitted pathogens themselves and to changes in the human hosts. For example, increasing antimicrobial resistance among some organisms, notably *Neisseria gonorrhoeae* and *Haemophilis ducreyi*, has reduced the reliability of traditional inexpensive treatments, and waxing and waning of specific strains of *N. gonorrhoeae* has resulted in differing clinical manifestations and patterns of transmission of gonorrhea. The average duration of infection is another biological factor (partly influenced by behavioral and societal factors, such as access to health care) that markedly influences the epidemiology of sexually transmitted diseases. Infections with most sexually transmitted viruses, such as herpes simplex virus or human papillomavirus, are incurable, persist indefinitely, and remain transmissible for many years, often for life. Therefore, low rates of sexual partner change are sufficient to sustain infection, and these infections are highly prevalent in all segments of the population. In contrast, the mean interval from acquisition to treatment or to spontaneous cure of gonorrhea or chancroid is measured in days to weeks, so that frequent new sexual partnerships are required to sustain infection in the community. Thus, the classical bacterial sexually transmitted diseases are largely limited to small subsets of the population with especially high rates of partner change. The clinical manifestations of *C. trachomatis* infections are subtle, so that the mean duration of infectivity is longer than that of gonorrhea or chancroid; however, it is shorter than that of the viral infections. Thus, intermediate rates of partner change are required to sustain the prevalence of infection, and chlamydial infections are more widely distributed in society than gonorrhea but less so than genital herpes or human papillomavirus infection.

Some populations may be more susceptible to some sexually transmitted diseases than in the past. For example, at one time most young adults had partially protective antibodies to the herpes simplex viruses, but the majority of persons in North America now lack such immunity when they reach sexual maturity. This reduced immunity may partly explain the emergence of genital herpes as one of the most common sexually transmitted infections. By contrast, genital herpes is less common in developing countries, where oral herpes infection in childhood is almost universal. A similar phenomenon may have contributed to the increasing incidence of sexually transmitted infections with cytomegalovirus, hepatitis B virus, human papillomavirus, and molluscum contagiosum.

Although sexually transmitted diseases are common in all populations, in North America and Europe they are more prevalent in urban than in rural areas; in blacks, Hispanics, and native Americans than in whites and persons of Asian ethnicity; and in lower than in higher socioeconomic and educational strata. In the United States, the incidence ratio of gonorrhea in Blacks compared with that in Whites rose from 10:1 in the early 1980s to 25:1 in 1989. A similar trend was observed for syphilis. Minority populations are especially likely to seek care

at public sexually transmitted disease clinics, where case ascertainment and reporting are greater than in the private health care sector. However, this probably explains only a small proportion of the difference in incidence between racial groups. These contrasts reflect cultural differences in sexual behavior, in real or perceived access to health care, and in the behavioral response to early or mild symptoms of infection. Such behaviors in turn are related to social disruption, poor education, anonymity, and family instability. In addition, the abuse of alcohol and illicit drugs has long influenced the risk of sexually transmitted infections by affecting the selection of sexual partners and because substance abuse is closely linked with prostitution.

Although the rates of sexually transmitted diseases are higher for men than women, for most infections women are more likely than men to suffer complications. Moreover, in their early stages many sexually transmitted diseases more commonly are asymptomatic or cause nonspecific symptoms and signs in women than in men, fostering delays in seeking medical care, underdiagnosis, and inadequate treatment. Sexually active adolescent girls have dramatically higher rates of most sexually transmitted diseases and their complications than older women (see **illus.**), due to enhanced behavioral risks and perhaps to age-related physiologic differences.

There are important links between AIDS and the other sexually transmitted diseases. The sexual transmission of the human immunodeficiency viruses is enhanced in the presence of genital ulcers, such as those caused by chancroid, syphilis, and genital herpes. Preliminary data suggest that chlamydial infection, gonorrhea, and trichomoniasis may have similar effects. It is probable that any inflammatory condition of the genitals enhances the efficiency of HIV transmission, and that the wide variation in the epidemiology of HIV infection in heterosexual population around the world is largely due to differences in the frequency of other sexually transmitted infections. The clinical course and response to treatment of some sexually transmitted diseases (for example, syphilis and genital herpes) are adversely affected by human immunodeficiency virus infection.

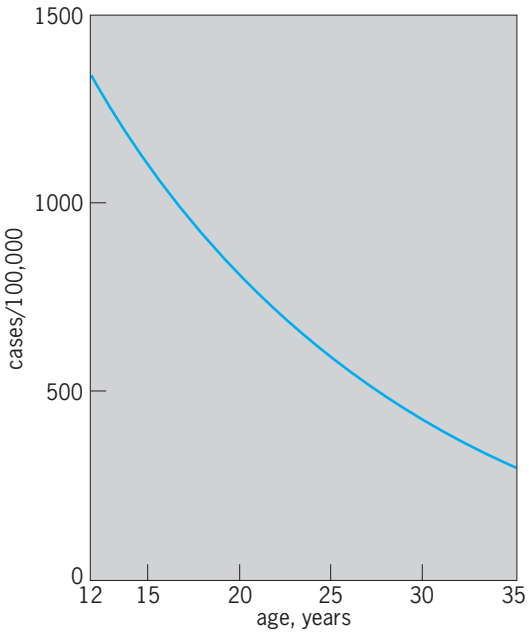

Age-related incidence of hospitalized cases of acute pelvic inflammatory disease in sexually experienced women in the United States. (*From K. K. Holmes et al., eds., Sexually Transmitted Diseases, McGraw-Hill, 1990*)

Sexually transmitted diseases are not uncommon in children. In newborns they are acquired before birth or from an infected birth canal during delivery. From the newborn period to 1 year of age, sexually transmitted infections usually are acquired by nonsexual contact with an infected mother or by sharing a bed with an infected adult, often in a setting of poor hygiene. From age 1 year to puberty, most cases are the result of sexual abuse.

Classification. The sexually transmitted diseases may be classified in the traditional fashion, according to the causative pathogenic organisms, as follows:

Bacteria
- *Chlamydia trachomatis*
- *Neisseria gonorrhoeae*
- *Treponema pallidum*
- *Mycoplasma genitalium*
- *Mycoplasma hominis*
- *Ureaplasma urealyticum*
- *Haemophilis ducreyi*
- *Calymmatobacterium granulomatis*
- *Salmonella* species
- *Shigella* species
- *Campylobacter* species

Viruses
- Human immunodeficiency viruses (types 1 and 2)
- Herpes simplex viruses (types 1 and 2)
- Hepatitis viruses B, C, D
- Cytomegalovirus
- Human papillomaviruses
- Molluscum contagiosum virus
- Kaposi sarcoma virus

Protozoa
- *Trichomonas vaginalis*
- *Entamoeba histolytica*
- *Giardia lamblia*
- *Cryptosporidium* and related species

Ectoparasites
- *Phthirus pubis* (pubic louse)
- *Sarcoptes scabiei* (scabies mite)

Sexually transmitted diseases may also be classified according to clinical syndromes and complications that are caused by one or more pathogens as follows:

1. Acquired immunodeficiency syndrome (AIDS) and related conditions

2. Pelvic inflammatory disease
3. Female infertility
4. Ectopic pregnancy
5. Fetal and neonatal infections (conjunctivitis, pneumonia, pharyngeal infection, encephalitis, cognitive impairment, physical anomalies, deafness, immunodeficiency, death)
6. Complications of pregnancy (spontaneous abortion, premature labor, premature rupture of fetal membranes, chorioamnionitis, postpartum endometritis)
7. Neoplasia (cervical dysplasia and carcinoma, Kaposi's sarcoma, hepatocellular carcinoma, squamous cell carcinomas of the anus, vulva, and penis)
8. Human papillomavirus and genital warts
9. Genital ulcer–inguinal lymphadenopathy syndromes
10. Lower genital tract infection in women (cervicitis, urethritis, vaginal infection)
11. Viral hepatitis and cirrhosis no12
12. Urethritis in men
13. Late syphilis
14. Epididymitis
15. Gastrointestinal infections (proctitis, enteritis, proctocolitis, enterocolitis)
16. Acute arthritis no17
17. Mononucleosis syndromes no18
18. Molluscum contagiosum
19. Ectoparasite infestation (scabies, pubic lice)

Most of these syndromes may be caused by more than one organism, often in conjunction with nonsexually transmitted pathogens. They are listed in the approximate order of their public health impact, and it is noteworthy that the first half of the list is dominated by conditions that predominantly affect women. The remainder of this article considers the most important sexually transmitted diseases according to both the etiologic and syndromic classifications.

Bacteria. Bacteria cause the best-known and most frequently recognized sexually transmitted diseases, as well as several other sexually transmitted infections.

Neisseria gonorrhoeae. *Neisseria gonorrhoeae*, or the gonococcus, is the cause of gonorrhea, one of the longest-recognized sexually transmitted diseases. The incidence of gonorrhea in most industrialized countries peaked in the early to mid-1970s, then began to decline. Nevertheless, the disease persists at epidemic levels in most areas; in the United States, an estimated 600,000–800,000 cases occur annually. Increasing resistance of *N. gonorrhoeae* to penicillin and other antibiotics has hampered control efforts, especially in developing countries, where the expense of alternative treatments limits their availability. Cephalosporin antibiotics (for example, ceftriaxone or cefixime) or fluoroquinolones (for example, ofloxacin or ciprofloxacin) usually are used to treat gonorrhea in industrialized countries. *See* DRUG RESISTANCE.

Treponema pallidum. *Treponema pallidum* is the cause of syphilis, which has been recognized since antiquity. An inability to cultivate *T. pallidum* is an important impediment to understanding the pathogenesis of syphilis and to its prevention. Syphilis is a protean disease known historically for its ability to mimic many other conditions. Penicillin is highly effective against *T. pallidum* and remains the treatment of choice in all stages of the disease. However, even penicillin does not cure all individuals, especially those with underlying immunodeficiency due to AIDS or other conditions. The changing epidemiology of syphilis is a prime example of the influence that demographic, social, and behavioral factors can have on an infectious disease despite the ready availability of effective antibiotic therapy. *See* SYPHILIS.

Chlamydia trachomatis. *Chlamydia trachomatis* is a small bacterium that, like viruses, requires cell culture techniques for isolation. Some serovariants (L_1, L_2, and L_3) cause the classical sexually transmitted disease lymphogranuloma venereum (LGV), now uncommon in North America and Europe. Infection with the non-LGV serovariants of *C. trachomatis* is the most common bacterial sexually transmitted disease in industrialized countries. During the early 1990s, an estimated 3–4 million cases occurred annually in the United States. *Chlamydia trachomatis* is the most common cause of pelvic inflammatory disease and female infertility in industralized countries.

Several factors have inhibited control of chlamydial infections, including their insidious nature and the subtlety of early signs and symptoms; the lack of widespread screening programs for high-risk individuals; the lack of effective single-dose antibiotic therapy (in contrast to gonorrhea and syphilis); and inadequate knowledge among physicians, health authorities, and the public about the nature, extent, and implications of the chlamydia epidemic. The necessity for cell culture testing for chlamydial infection, an expensive and technically demanding procedure, contributed to these problems. The development of diagnostic tests based on genetic amplification methodology, such as the polymerase chain reaction and the ligase chain reaction, is an important advance that has greatly enhanced the diagnosis, prevention, and control of chlamydial infections. Chlamydial infections respond to treatment with the macrolide or tetracycline classes of antibiotics, such as azithromycin and doxycycline, respectively; penicillin is not effective.

Genital mycoplasmas. The genital mycoplasmas, *Ureaplasma urealyticum* and *Mycoplasma hominis*, are cell wall–free bacteria that reside in the urethra or vagina of most sexually active men and women; over 50% of persons with five or more lifetime sexual partners are colonized with either or both organisms. Sexual contact may not be the sole means of transmission, since some virgin women harbor these organisms. Although most infections are not associated with clinical disease, 20–30% of nongonococcal

urethritis cases in men may be caused by *U. urealyticum*. This organism also has been linked with low birth weight and perhaps premature delivery of infants born to infected mothers. *Mycoplasma hominis* may contribute to some cases of acute pelvic inflammatory disease and bacterial vaginosis. *See* MYCOPLASMAS.

Haemophilus ducreyi. *Haemophilus ducreyi* is the cause of chancroid, or soft chancre, one of the five classical sexually transmitted diseases. It is characterized by painful genital ulcerations, often accompanied by tender enlarged regional lymph nodes. Chancroid is uncommon in the United States, but the incidence rose rapidly from the early 1980s, when several hundred cases were reported annually, to 1989, when 5191 cases were reported; the incidence then declined once again, to fewer than 1,000 cases per year during the early 1990s. By contrast, chancroid is among the most common sexually transmitted diseases in many developing countries. In both industrialized and developing countries, chancroid is more closely associated with prostitution and illicit drug use than any other sexually transmitted infections. Several antibiotics are effective in treatment, including azithromycin, erythromycin, ceftriaxone, and ciprofloxacin.

Calymmatobacterium granulomatis. *Calymmatobacterium granulomatis* causes donovanosis, or granuloma inguinale, one of the five classical sexually transmitted diseases. It causes morphologically typical ulcerations of the skin on or adjacent to the genitals. Donovanosis is extremely rare in industrialized nations and uncommon in most of the world, except for a few localized areas such as Papua New Guinea and the Indian subcontinent.

Enteric bacteria. Several species of the intestinal pathogens *Shigella, Salmonella*, and *Campylobacter* commonly are transmitted by sexual practices that foster fecal-oral contamination, especially among homosexually active men.

Viruses. Although viruses are among the most common and serious sexually transmitted pathogens, recognition of the importance and public health impact of most viral sexually transmitted diseases is relatively recent.

Human immunodeficiency viruses. The human immunodeficiency viruses types 1 and 2 (HIV-1 and -2) cause AIDS. These viruses are transmitted almost exclusively by sexual contact, by direct exposure of the bloodstream to the virus, or by perinatal exposure of the fetus or newborn to maternal infection. Sexual transmission is enhanced in the presence of sexually transmitted diseases, especially those that cause genital ulceration. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS).

Herpes simplex viruses. Genital herpes, caused by herpes simplex virus type 2 and less commonly by the type 1 virus, is among the most common and rapidly increasing sexually transmitted diseases. About 24% of the population of the United States becomes infected with herpes simplex virus type 2 by age 35. Most genital herpes infections are asymptomatic or cause trivial symptoms whose significance is misunderstood by the individual and physician alike. Nonetheless, genital herpes is a serious public health problem due to its potential to cause devastating infections in infants born to infected mothers and to the psychological stress that often accompanies infection. Acyclovir and related antiviral drugs can control the clinical manifestations of infection but are not curative. *See* HERPES.

Hepatitis viruses. Hepatitis B virus is transmitted by direct contact with infected blood or body secretions, as is hepatitus D virus (formerly delta agent), which causes disease only in the presence of hepatitis B infection. Hepatitis C virus, an important cause of viral hepatitis due to neither hepatitis A nor B, is bloodborne and probably is sexually transmitted as well, albeit less efficiently than hepatitis B virus. Routine screening of blood donors has markedly reduced the risk of transfusion-acquired hepatitis B, leaving sexual contact the most common mode of transmission. Sexual contact also plays a distinct but minor role in transmission of the hepatitis A virus. Immunization with hepatitis B vaccine is an effective but underutilized prevention measure. *See* HEPATITIS.

Cytomegalovirus. Cytomegalovirus is an extremely common herpes group virus that ultimately infects 50–90% of persons in industrialized nations and 90–100% of those in developing countries. Sexual contact is an important mode of transmission of cytomegalovirus in young adults in some industrialized societies, but nonsexual transmission is also common. Although cytomegalovirus infection usually is asymptomatic in healthy adults, it is nevertheless a major cause of congenital infection that may result in miscarriage, malformation, mental retardation, or deafness. The virus causes lifelong latent infections, and life-threatening disease due to reactivation of asymptomatic cytomegalovirus infection is common in persons with impaired immune function due to AIDS, cancer, and other conditions. *See* CYTOMEGALOVIRUS INFECTION; EPSTEIN-BARR VIRUS.

Human papillomaviruses. The human papillomaviruses, of which there are more than 50 distinct varieties, cause warts. Most types cause common warts of the hands, feet, or other parts of the body and are not sexually transmitted. Most people with genital human papillomavirus infection are asymptomatic and have no visible manifestation, but human papillomavirus infection may be the most common of all sexually transmitted infections. Human papillomavirus types 6 and 11 cause most cases of clinically apparent genital warts. By contrast, types 16, 18, and several others usually do not cause visible warts, but are closely linked with cancer of the cervix, anus, vulua, and penis. While tobacco use and other (largely unknown) factors influence the development of malignancy, the weight of evidence suggests that cervical cancer is a sexually transmitted disease of which human papillomavirus is the direct cause. Clinically apparent warts can be ablated by freezing or by chemical irritants, but no curative treatment exists.

Molluscum contagiosum. The molluscum contagiosum virus causes characteristic painless wartlike skin lesions that have no known serious sequelae. Seen most commonly in young children, who acquire the disease by nonsexual contact with other children, molluscum contagiosum in adults usually involves the genital region and is transmitted sexually.

Kaposi's sarcoma. The epidemiology of Kaposi's sarcoma, a malignancy that is frequent in homosexual men with AIDS but less common in other individuals with AIDS, suggested that it might be caused by a sexually transmitted pathogen other than human immunodeficiency virus.

Protozoa. *Trichomonas vaginalis* is a flagellated protozoon that causes vaginal inflammation and discharge, but asymptomatic vaginal colonization also is common. Sexual contact is the primary mode of acquisition. Most infections of the male genital tract are asymptomatic, but urethritis may occur. The infection is readily treated with metronidazole or related antibiotics. Complications of trichomoniasis are rare. *Giardia lamblia, Entamoeba histolytica*, and *Cryptosporidium* are intestinal protozoa that are most commonly acquired by ingestion of the organisms' cysts in contaminated water. As for intestinal bacteria, however, they commonly are sexually transmitted, especially among homosexually active men. *See* PROTOZOA.

Ectoparasites. *Phthirus pubis*, the pubic (crab) louse, is slowly mobile, survives only for a few hours away from a human host, and usually requires direct genital apposition for transmission. *Sarcoptes scabiei*, the itch mite, is also transmitted by direct skin-to-skin contact. Sex is an important but not an exclusive mechanism of spread of both intestations.

Clinical syndromes. Sexually transmitted diseases result in a variety of clinical syndromes. The major syndromes are reviewed below except for AIDS and syphilis, which are discussed in separate articles.

Pelvic inflammatory disease. Acute pelvic inflammatory disease, or salpingitis, is the result of infection of the Fallopian tubes (salpinges) with bacteria that ascend from the cervix or vagina, in the absence of recent childbirth or gynecological surgery. Salpingitis is often accompanied by uterine inflammation (endometritis), peritonitis, or—in its most advanced form—formation of abscesses involving the Fallopian tubes, ovaries, or peritoneal cavity. Pelvic inflammatory disease and its complications constitute one of the most important health problems of women. It is estimated that 85–90% of cases are sexually acquired. Up to 1,000,000 cases are estimated to occur annually in the United States. In North America and Europe, 10–20% of acute pelvic inflammatory disease episodes (depending on the bacterial etiology and clinical severity) result in infertility due to bilateral scarring and obstruction of the Fallopian tubes; the rate undoubtedly is higher in socioeconomically disadvantaged women who lack ready access to health care. In addition to infertility, the major complications of the disease are ectopic (tubal) pregnancy, chronic pelvic pain, and recurrence of salpingitis.

Among sexually experienced women, the risk of pelvic inflammatory disease is inversely correlated with age (see illus.). Use of certain intrauterine devices for contraception results in a two- to ninefold enhanced risk, and accounts for many nonsexually acquired cases of pelvic inflammatory disease. In contrast, the use of hormonal contraception reduces the risk by about 50%. A prior history of pelvic inflammatory disease markedly increases the risk. Moreover, recurrent episodes enhance the risk of long-term complications, and up to 75% of women with three or more episodes become infertile. Vaginal douching is a strong independent risk factor for pelvic inflammatory disease. As a result, douching is strongly discouraged for either so-called feminine hygiene or treatment of vaginal infection.

The most common causes of pelvic inflammatory disease are gonorrhea and chlamydial infection, but the relative contribution of each varies widely. Historically, *Neisseria gonorrhoeae* was long considered the primary cause. However, gonorrhea accounts for only 15–40% of cases in much of North America and Europe, where 20–50% are associated with chlamydial infection. Many women with the syndrome are infected with both of these organisms. Other bacteria commonly contribute to salpingitis, either alone or in combination with *Chlamydia trachomatis* or *N. gonorrhoeae;* these include *Mycoplasma hominis* and several aerobic and anaerobic species of the normal bacterial microflora of the vagina.

The clinical hallmarks of pelvic inflammatory disease are low abdominal pain and tenderness of the uterus, Fallopian tubes, and ovaries, elicited during bimanual pelvic examination. These commonly are accompanied by vaginal discharge or abnormal uterine bleeding. Nausea, vomiting, or fever commonly occur in severe cases. However, mild and even asymptomatic cases are common and cause many cases of infertility and ectopic pregnancy. Clinical management of pelvic inflammatory disease includes treatment with antibiotics, and occasionally surgery to drain or remove abscesses. Because of the broad spectrum of bacteria involved, two or more antibiotics typically are used in combination, such as ofloxacin with metronidazole or ceftriaxone with doxycycline. Correction of reversible predisposing factors also is important (such as removal of an intrauterine device and examination and treatment of sexual partners).

Infertility. The incidence of female infertility has risen in parallel with the sexually transmitted disease epidemic, and it is estimated that 100,000 or more women in the United States become infertile each year as the result of sexually transmitted infections. In localized parts of Africa, up to half of the women of childbearing age are estimated to be infertile because of prior gonococcal pelvic inflammatory disease. In the industrialized countries, about half of all female infertility is the result of tubal scarring or obstruction. About 75% of those

with tubal infertility have antibodies to *C. trachomatis*, indicating past chlamydial infection, compared with 0–30% of fertile women and those with other forms of infertility. Chlamydial infection may be the most common preventable cause of female infertility in industrialized countries. Most of the women with chlamydia-associated tubal obstruction lack past histories of overt pelvic inflammatory disease or unexplained abdominal pain, indicating that asymptomatic chlamydial infection can cause sufficient salpingitis to result in infertility. The treatments for tubal infertility (surgical reconstruction and in vitro fertilization) are expensive and largely unsatisfactory. Although infertility in men can result from testicular inflammation (epididymitis) due to *N. gonorrhoeae* or *C. trachomatis*, sexually transmitted disease apparently is not a common cause of male infertility.

Ectopic pregnancy. Ectopic (tubal) pregnancy occurs when an embryo implants in the Fallopian tube rather than the uterus; growth of the embryo and placenta ultimately causes tubal rupture and internal bleeding, a potentially fatal medical emergency. Ectopic pregnancy, like tubal infertility, usually is the result of Fallopian tube scarring. The incidence of ectopic pregnancy rose in parallel with the sexually transmitted disease epidemic, and the risk of ectopic pregnancy is increased about tenfold after the occurrence of pelvic inflammatory disease. *See* PREGNANCY.

Fetal and perinatal infections. Most sexually transmitted pathogens are transmissible either to the unborn fetus or to the newborn during or immediately after birth to an infected mother, with effects that range from trivial to devastating. The most dangerous sexually transmitted diseases for the newborn are human immunodeficiency virus infection, herpes, and syphilis. Neonatal herpes is most common when the mother experiences her first infection with herpes simplex virus immediately prior to delivery, and commonly results in encephalitis, hepatitis, and multiple organ failure. The mortality is about 50%, and most surviving infants suffer severe neurological and developmental impairment. It is estimated that several thousand cases occur annually in the United States.

Reported cases of congenital syphilis in infants under 1 year of age in the United States rose from 107 in 1980 to almost 700 in 1989 more than in any year since the 1940s. Neonatal human immunodeficiency virus infection rose similarly during this period. Both occur most frequently in socioeconomically disadvantaged populations and are especially common in some developing countries.

Cytomegalovirus is a common cause of congenital or perinatal infection. Although most cases are benign, several thousand clinically significant cases occur annually in the United States. In its most severe form, which usually follows primary maternal infection during pregnancy, congenital cytomegalovirus infection results in severe neurodevelopmental abnormalities.

Chlamydial infection, acquired during birth to an infected mother, is the most common neonatal sexually transmitted disease in industrialized countries. One to five percent of all newborns in the United States were estimated to be infected in the 1980s, primarily with asymptomatic nasopharyngeal infections or ocular infections (neonatal inclusion conjunctivitis), and less commonly with pneumonia. Undetected perinatal chlamydial infections may contribute to middle ear inflammation (otitis media) and persistent cough due to bronchiolitis in young children.

Gonococcal conjunctival infection (ophthalmia neonatorum) now is rare in industrialized countries but still is common in some locales; delayed treatment often leads to blindness. Neonatal hepatitis B virus infection is a major health problem in most tropical areas and in some populations in industrialized countries. Transmission of genital human papillomavirus infection to the newborn can result in warts of the genitals, perianal regions, or throat (pharyngeal papillomatosis); the latter condition is particularly difficult to treat.

Complications of pregnancy. In addition to infections of the fetus or newborn, many sexually transmitted pathogens adversely affect the course of pregnancy. Most surveys in North America and Europe have shown 5–10% of pregnant women to have chlamydial infection. Similar prevalences of gonorrhea exist in socioeconomically disadvantaged pregnant women in much of the world. Gonorrhea and probably chlamydial infection increase the risk of premature rupture of the fetal membranes, premature delivery, postpartum uterine infections (endometritis), and perhaps miscarriage and stillbirth. Infection of pregnant women with the genital mycoplasmas also may be responsible for some cases of prematurity, low birth weight, and postpartum endometritis. Bacterial vaginosis also is associated with early rupture of the fetal membranes, prematurity, and other complications of labor and delivery.

Neoplasia. For several decades, the risk for squamous cell cancer of the cervix, the most common of all malignancies in men and women, was known to be linked with certain sexual behaviors, including young age at first intercourse and childlessness. Both of these factors now are known to be linked with increased numbers of sexual partners and therefore with sexually transmitted disease risk. It also was shown that the risk of cervical cancer in permanently monogamous women is correlated with their husbands' lifetime sexual partners. Despite their relative youth, women attending sexually transmitted disease clinics have a higher prevalence of cervical cancer and its precursors than any other group of women that has been surveyed.

Because of such epidemiologic observations, at various times gonorrhea, trichomoniasis, and syphilis were postulated as causes of cervical cancer, but they clearly play no direct roles. There are conflicting data regarding herpes simplex virus, but this agent probably has no primary role. In contrast, a growing body of data suggests that some strains

Lower genital tract infections in women

Syndrome	Sites affected	Usual causes
Cervicitis	Uterine cervix	*Chlamydia trachomatis* *Neisseria gonorrhoeae* Herpes simplex virus Other or unknown
Vaginal infection	Vagina, labia	*Candida albicans* and other yeasts *Trichomonas vaginalis* Herpes simplex virus *Gardnerella vaginalis* with vaginal anaerobic bacteria Foreign body Other
Urethritis-cystitis	Urethra, bladder	*Escherichia coli* and other coliform bacteria *Staphylococcus saprophyticus* *Chlamydia trachomatis* *Neisseria gonorrhoeae* Herpes simplex virus

of human papillomavirus directly cause precancerous changes (dysplasia) and overt cancer. All women with past histories of sexually transmitted disease are particularly encouraged to undergo a cervical cancer screening examination (Papanicolaou test) annually.

Other malignancies also have been linked with sexually transmitted diseases, though generally less strongly than has cervical cancer. Homosexual men are unusually susceptible to anal cancer, probably due to human papillomavirus infection. Chronic carriers of hepatitis B virus have an enhanced risk of liver cancer (hepatocellular carcinoma). Cancer of the skin of the vulva and penis also is related to human papillomavirus. Kaposi's sarcoma, believed to be a malignancy of lymphatic epithelial cells, may be due to a sexually transmitted herpes group virus. *See* CANCER (MEDICINE).

Genital warts. Warts of the genitals, anus, or rectum (condyloma acuminatum) are a common and rapidly growing problem among sexually active individuals in North America and Europe. Although genital warts were long considered a benign (albeit unpleasant and inconvenient) health problem, the association of some strains of human papillomavirus with cancer has elevated the perceived public health impact of this infection. Most affected tissues are not visibly abnormal, and in most patients the true extent of infection is substantially greater than indicated by overt warts. Partly for this reason, eradication of genital warts often is difficult, and no treatment has been shown to eradicate the virus or to affect the risk of later malignancy.

Genital ulcer/inguinal lymphadenopathy syndromes. Several sexually transmitted infections cause ulcers or other lesions of the skin or mucous membranes of the genitals, tenderness and swelling of the lymph nodes of the groin (inguinal lymphadenopathy), or both. Primary syphilis is a common cause throughout the world, but the frequencies of the other diseases have great geographic or demographic variation. Genital herpes is by far the most common cause of genital ulcer disease in most of Europe and North America, where chancroid is rare; the reverse is true in most tropical settings. Lymphogranuloma venereum and donovanosis are the other major causes; both are rare in the United States.

Lower genital tract infections in women. Lower genital tract infection in women refers to cervical infection (cervicitis), vaginal infection, and infection of the bladder and urethra (cystitis-urethritis). These syndromes cause various combinations of vaginal discharge, vulvar pain or itching, and urinary symptoms such as painful, frequent, or urgent urination. Most lower genital tract infections are clinically mild. However, cervicitis and some vaginal infections predispose to pelvic inflammatory disease, and cystitis-urethritis can lead to serious infections of kidneys (pyelonephritis). Each syndrome may be caused by sexually transmitted infections or by nonsexually transmitted agents (see **table**).

Cervicitis is the female counterpart to urethritis in men; it usually results in discharge from the vagina or painful urination (dysuria), but commonly is asymptomatic. *Chlamydia trachomatis* and *Neisseria gonorrhoeae* are the most common causes; the proportion caused by each varies geographically and with other demographic factors, along the lines described above for pelvic inflammatory disease. A few percent of cases are due to herpes simplex virus, usually in conjunction with overt genital herpes. The etiology cannot be defined in up to one-half of the cases; the role of sexual transmission in these cases is uncertain.

There are three major causes of vaginal infection. Vulvovaginitis due to various yeast-type fungi (usually *Candida albicans*) affects the majority of women at one time or another, usually causing scant discharge and prominent vulvar irritation. It is not sexually transmitted.

Bacterial vaginosis usually causes a nonirritating white or gray vaginal discharge, often with a distinctive fishy odor. It results from overgrowth of various normal vaginal bacteria, especially anaerobic organisms (which cause the distinctive odor), *Gardnerella vaginalis*, the genital mycoplasmas, and *Mobiluncus curtisii*, and by depletion of *Lactobacillus* species, especially strains that produce hydrogen peroxide.

However, the precise cause of bacterial vaginosis is unknown. The syndrome occurs most frequently in women at risk for sexually transmitted diseases, but a sexually acquired pathogen has not been demonstrated. Long considered a benign condition, bacterial vaginosis is known to predispose individuals to potentially serious uterine infections (endometritis) after childbirth or gynecologic surgery, to early rupture of the fetal membranes and premature labor, and perhaps to pelvic inflammatory disease. Antibiotics with activity against anaerobic bacteria, such as clindamycin or metronidazole, are effective either by mouth or by intravaginal administration. No male counterpart to bacterial vaginosis has been defined, and treatment of the individual's sexual partners does not improve, cure, or prevent recurrence of the syndrome.

Trichomonas vaginalis causes discharge that may have a fishy odor. Trichomoniasis usually is acquired sexually, and the male sexual partner, who usually is asymptomatically infected, should be treated. However, a woman can develop symptoms after carrying the organism for several years, so cases are common in women not currently at risk for a sexually transmitted infection.

Numerous other conditions can result in increased vaginal discharge or irritation. Common causes include chemical irritation or allergy (for example, from commercial douche preparations or vaginal contraceptives); foreign body (such as a forgotten tampon); estrogen-lack ("atrophic") vaginitis following menopause or surgical removal of the ovaries; and increased normal discharge.

Cystitis-urethritis in young women may be due to sexually transmitted disease, especially gonorrhea or chlamydial infection. However, these are less common than nonsexually acquired urinary tract infections, the most common causes of which are *Escherichia coli* and *Staphylococcus saprophyticus*.

Viral hepatitis. These infections and their complications have been especially common in homosexual and bisexual men who have many sexual partners, persons who share equipment used to inject illicit drugs, and their sexual partners. Chronic hepatitis B and C are common causes of cirrhosis and liver failure, and chronic hepatitis B is the primary cause of hepatic carcinoma, one of the most common cancers in developing countries. *See* HEPATITIS.

Urethritis in men. Urethritis (urethral infection) is the most common of all sexually transmitted disease syndromes in men. Urethritis usually is associated with visible discharge of exudate from the penis and often with painful urination or a sense of irritation or itching within the urethra. These symptoms may be mild, and asymptomatic sexually transmitted urethral infections are common. Gonorrhea historically has been the most common cause of urethritis, and remains so in much of the world. However, nongonococcal urethritis, which refers to all urethritis that is not due to gonorrhea, is substantially more common in most populations in North America and Europe. About 30–40% of cases are due to *Chlamydia trachomatis*, 20–30% to *Ureaplasma urealyticum*, and under 5% each to herpes simplex virus and *Trichomonas vaginalis*. Some cases may be due to a newly recognized genital mycoplasma, *Mycoplasma genitalium*. The causes of up to 50% of cases are uncertain, although epidemiologic evidence indicates that these infections also are sexually acquired. About 10–20% of men with urethral gonorrhea in industrialized countries have simultaneous urethral chlamydial infections. Urethritis itself usually is a mild condition; its major significance to public health is as a source of sexually transmitted disease in women.

Epididymitis. Acute epididymitis is infection of the sperm-collecting ducts (epididymis) adjacent to the testis; the testis itself also may be involved (epididymo-orchitis). Epididymitis causes testicular pain and swelling, and results from ascending infection from the urethra via the vas deferens. It is the male counterpart to pelvic inflammatory disease, but is far less common than that syndrome. In young, sexually active men, *N. gonorrhoeae* and *C. trachomatis* are the major causes; in most settings in Europe and North America, chlamydial infection is a much more common cause than gonorrhea. Nonsexually transmitted bacteria, such as *Escherichia coli*, usually are responsible for epididymitis in men over age 35 and those who recently have undergone genitourinary surgery or instrumentation. Bilateral epididymitis can result in sterility. *See* URINARY TRACT DISORDERS.

Gastrointestinal infections. Sexually transmitted infections of the gastrointestinal tract are major health problems among homosexually active men; some of them also are common in heterosexuals, especially women. The most common rectal infections (proctitis) are gonorrhea, chlamydial infection, herpes, and syphilis. Proctitis results from direct inoculation of infectious material into the rectum, either via rectal intercourse or, in many women, through contamination of the anal area by vaginal discharge. Most other sexually transmitted gastrointestinal infections affect the small intestine (enteritis) or the colon (colitis) and are acquired during sexual activity that facilitates fecal-oral contamination. The protozoa *Entamoeba histolytica* and *Giardia lamblia* and the bacteria *Shigella, Salmonella*, and *Campylobacter* are the predominant causes. On a worldwide scale, however, the large majority of such infections are acquired by nonsexual routes.

Arthritis. Gonorrhea occasionally causes an acute arthritis that if untreated can result in permanent joint damage. Chlamydial infection is the most common trigger of Reiter's syndrome, a type of chronic arthritis that occasionally causes severe disability. These two diseases are among the most common causes of acute arthritis in sexually active young adults in the United States. *See* ARTHRITIS.

Mononucleosis. Mononucleosis is an illness characterized by fever, malaise, enlargement of lymph nodes, and often pharyngitis. Mononucleosis due to cytomegalovirus often is sexually acquired. The

more common infectious mononucleosis, due to Epstein-Barr virus, has been called a "kissing disease," reflecting carriage of the virus in saliva and the potential for sexual transmission. *See* INFECTIOUS MONONUCLEOSIS.

Ectoparasite infestation. Scabies and pubic louse infestation are common and cause pruritic skin rashes that usually are nuisances rather than serious threats to health.

Prevention and control. Prevention strategies for communicable diseases are categorized as primary prevention, the prevention of infection as such; secondary prevention, the detection of infection and early treatment, before complications ensue; and tertiary prevention, the treatment of complications and amelioration of long-term consequences. Primary prevention of sexually transmitted diseases includes educating those at risk to select sexual partners at low risk of infection, to use condoms or other barrier contraceptives, and to maintain permanent, monogamous sexual relationships. Secondary prevention emphasizes screening persons at risk to detect subclinical infection, antibiotic treatment, and notification and treatment of the sexual partners of infected persons. Examples of tertiary prevention for sexually transmitted infections include treatment of pelvic inflammatory disease or of infants with neonatal herpes; in vitro fertilization or fallopian tube reconstructive surgery for infertile women following pelvic inflammatory disease; and the provision of institutional care for infants with birth defects due to maternal sexually transmitted infections. Primary and secondary prevention are far more cost-effective than tertiary prevention and are emphasized in most control programs.

The only certain methods for prevention of sexually transmitted disease are sexual abstinence and permanent mutual monogamy. Less stringent limitations on sexuality confer varied levels of protection; examples include avoiding multiple simultaneous partners and foregoing sexual contact with persons suspected to be promiscuous, and with individuals known to have genital or other symptoms suggestive of sexually transmitted disease. The consistent and proper use of either male or female condoms provides a high degree of protection against pathogens transmitted between mucous membranes (for example, *C. trachomatis*, *N. gonorrhoeae*, and the human immunodeficiency viruses). Condoms confer significant but less complete protection against pathogens transmitted between skin surfaces (for example, herpes simplex virus and *T. pallidum*), because not all sexually exposed surfaces are covered. Latex or polyurethane condoms are believed to provide better protection than condoms made of biological membranes. No sexually transmitted pathogen passes through intact latex or polyurethane condoms. Diaphragms and vaginal spermicides may help to prevent some infections in the event of exposure. Intrauterine devices should be avoided by women at high risk for sexually transmitted disease.

Persons at high risk for sexually transmitted diseases should seek periodic examinations to detect asymptomatic infections. Prompt medical attention for appropriate symptoms usually prevents the development of complications. Preventive treatment with an appropriate antibiotic before or immediately after exposure may be effective, but is not practical in most circumstances.

Communitywide control measures include provision of adequate clinical services for diagnosis and treatment of sexually transmitted infections, which usually implies a combination of publicly supported clinics and appropriate training of health care providers in the community. Disease detection by routine screening of individuals at risk is central to control of gonorrhea, syphilis, chlamydial infection, and human immunodeficiency virus infection. Measures to confidentially locate and assure treatment and counseling of the sexual partners of infected persons are important to interrupt chains of transmission. For hepatitis B, and perhaps for other sexually transmitted diseases in the future (such as genital herpes), vaccination programs are valuable. Directed health education, with attention to the preventive measures outlined above, is important. However, educational programs that use scare tactics or whose primary goal is proscription or limitation of sexual activity ("moral prophylaxis") historically have been ineffective. Training of medical students, physicians, and other health care providers in the increasingly complex field of sexually transmitted disease is a crucial measure. The reporting of cases to health authorities is necessary to facilitate the targeting of resources for disease control. Finally, basic and applied research provides the growing knowledge on which all other control measures are based. *See* PUBLIC HEALTH.

H. Hunter Handsfield

Bibliography. A. M. Brandt, *No Magic Bullet: A Social History of Venereal Disease in the United States Since 1890*, 2d ed., 1990; H. H. Handsfield, *Color Atlas and Synopsis of Sexually Transmitted Diseases*, 2d ed., 2000; K. K. Holmes et al. (eds.), *Sexually Transmitted Diseases*, 2d ed., 1989; E. O. Laumann et al., *The Social Organization of Sexuality: Sexual Practices in the United States*, 2000.

Sferics

Electromagnetic radiations produced primarily by lightning strokes from thunderstorms. It is estimated that globally there occur about 2000 thunderstorms at any one time, and that these give rise to about 100 lightning strokes every second. The radiations are short impulses that usually last a few milliseconds, with a frequency content ranging from the low audio well into the gigahertz range. Sferics (short for atmospherics) are easily detected with an ordinary amplitude-modulation (AM) radio tuned to a region between radio stations, especially if there are thunderstorms within a few hundred miles. These

sounds or noises have been identified and characterized with specific names, for example, hiss, pop, click, whistler, and dawn chorus. They fall into what is generally known as radio noise. *See* ATMOSPHERIC ELECTRICITY; ELECTROMAGNETIC RADIATION; LIGHTNING; THUNDERSTORM.

Types of sferics. The various types of sferics include terrestrial, magnetospheric, or Earth-ionospheric.

Terrestrial. Occasionally lightning is associated with other natural phenomena. Violent volcanic eruptions accompanied by phreatic explosions when the molten lava comes into contact with seawater have been observed to produce lightning. The birth of the Icelandic island of Surtsey was accompanied by spectacular lightning displays, as was the eruption on the nearby island of Heimaey. Both eruptions provided a copious source of sferics, as do the somewhat regular eruptions of the Japanese volcano Sakurajima on Kyushu island. Dust storms and dust devils have also been observed to produce sferics. *See* DUST STORM.

Anthropogenic noise is also present in the sferics population: automobile ignition, motor brushes, coronas from high-voltage transmission lines, and various high-current switching devices are among human contributions to radio noise. Most of these noise sources, however, are too small to compete with the amplitude and power of the lightning-generated impulse except in close proximity to the source. One of these noises that does compete with the lightning sferic is the electromagnetic pulse that accompanies the detonation of a hydrogen (fusion-type) bomb. *See* ELECTRICAL NOISE.

Magnetospheric. Lightning-generated sferics are sometimes coupled into the magnetosphere, where they are trapped and guided by the Earth's magnetic field. In this mode, the impulse travels in an ionized region. As a result, the frequencies present in the original impulse are separated by dispersion (the higher frequencies travel faster than the lower) and produce the phenomena known as whistlers. Amplification may occur, and reflection at the ends of the geomagnetic ducts (magnetic North and South poles) result in the return of the sferic to the source. The increased length of the path effectively stretches the original impulse into a train of decreasing audio frequencies sometimes several seconds in duration. Sferics often undergo several reflections in the duct. Studies of whistlers have led to an increased understanding of the magnetosphere and regions of enhanced ionization. *See* MAGNETOSPHERE.

Earth-ionospheric. By far the dominant and most readily observed sferics are the lightning-produced impulses that travel in the spherical cavity formed by the ionosphere and the Earth's surface. Lightning currents produce strong radiation in the very low-frequency band, 3–30 kHz, and in the extremely low-frequency band, 6 Hz–3 kHz. For frequencies above about 3 kHz, the propagation in the Earth–ionosphere cavity for distances up to about 600 mi (1000 km) can be treated as a ground wave and a series of multiple reflections between the Earth and the ionosphere. The farther the observing station from the source, the greater the number of reflections that may reach the receiver. An example of a very low-frequency sferic followed by a series of reflections is given in **Fig. 1**. The initial downward deflection (about one-third the size of the maximum negative deflection) is the radiated electric field from a lightning stroke lowering negative charge to the Earth. This vestige of the ground wave has been strongly attenuated in traveling over the land. The first four ionospheric reflections show much less attenuation. More than 15 reflections are discernible as the sferic propagates in the Earth–ionosphere cavity. The train of high-frequency noise preceding the negative stroke impulse is produced by the downward-moving stepped leader that initiated the lightning flash.

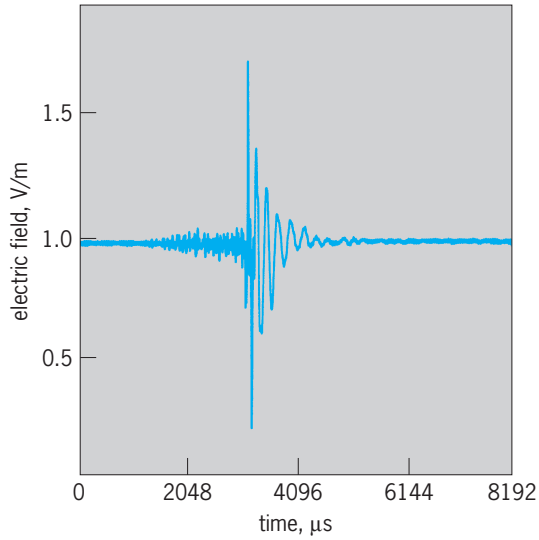

Fig. 1. Sferic that originated 1006 mi (1610 km) from the receiver.

An interesting variant of wave propagation in the Earth–ionosphere cavity is given for the mode $n = 0$. In this mode, known as the Earth–ionosphere waveguide mode, only frequencies less than about 2–3 kHz propagate, so that a lightning sferic with both a strong high- and low-frequency content reaches the receiver with two sets of signals: the normal very low-frequency component, $n = 1$ and greater, via a ground wave and ionospheric reflections, and a set with a slower group velocity (frequencies below 3 kHz) that propagate in the $n = 0$ waveguide mode. The sferic then exhibits a so-called slow tail (**Fig. 2**). The slow tail of this sferic indicates a significant frequency component below about 1 kHz in the source. It appears that strokes that lower positive charge to Earth exhibit a slow tail far more often than the common negative strokes. Positive strokes are a reasonably rare occurrence in summer thunderstorms: about 99% of all strokes to ground lower negative charge to Earth.

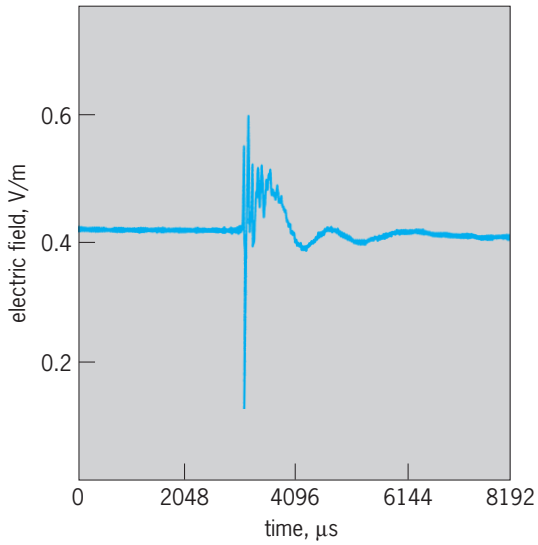

Fig. 2. Sferic exhibiting a slow tail; it originated in a positive lightning stroke 812 mi (1300 km) from the receiver.

The time between the first very low-frequency peak and the peak of the slow tail can be used to estimate the distance to the lightning flash. Similarly, the time between successive ionospheric reflections can also be used to estimate the distance to the source, provided the height of the ionosphere is known. Values for ionospheric height (which change with time of day) are approximately 40 mi (70 km) in daylight and 50 mi (90 km) at night. Although source distances can be estimated by the single-station methods described above, multiple-station measurements are usually preferred.

Lightning discharges are also observed to excite the Earth–ionosphere cavity at its fundamental and higher resonant frequencies, starting at about 8 Hz. These are known as the Schumann resonances. Studies of these resonances complement the studies of sferics at higher frequencies and have provided many interesting details related to specific properties of the ionosphere. *See* IONOSPHERE; RADIO-WAVE PROPAGATION.

Lightning detection. There are two principal methods used to locate lightning sources using sferics. One method uses a pair of loops or coils to detect a voltage induced by the magnetic field of the sferic. When properly calibrated, the coils produce equal voltages for a source located on the bisector of the angle between the loops. Usually, the coils are placed in vertical planes perpendicular to each other. The direction of the arriving signal is obtained from the ratio of the induced voltages in the coils. A value of the magnetic field strength related to the current impulse at the source is obtained by squaring and adding the voltages in the coils. The ambiguity in the direction of arrival of the sferic (from in front or from behind the coils) is resolved by making an independent measurement of the electric field of the sferic. Two or more stations provide two or more directions that intersect at the apparent source of the sferic. Accuracies of about 2° have been claimed for the crossed-loop systems.

The second system in use is based upon a measurement of the time of arrival of the sferic at three or more stations. This system requires excellent timing (microsecond accuracy) and also that there be specific characteristics of the sferic that are identifiable in the signal at each station, so that the arrival time of that particular feature can be estimated accurately. Both the time of arrival and the crossed-loop lightning-direction-finding systems are commercially available. They are used by the weather services and by private industry in many countries throughout the world. *See* STORM DETECTION. Marx Brook

Bibliography. J. Galejs, *Terrestrial Propagation of Long Electromagnetic Waves*, 1972; R. A. Helliwell, *Whistlers and Related Ionospheric Phenomena*, 1965; M. A. Uman, *The Lightning Discharge*, 1987, reprint 2001; H. Volland (ed.), *Handbook of Atmospherics*, 1982; J. R. Wait, *Electromagnetic Waves in Stratified Media*, 1962.

Shadow

A region of darkness caused by the presence of an opaque object interposed between such a region and a source of light. A shadow can be totally dark only in that part called the umbra, in which all parts of the source are screened off. With a point source, the entire shadow consists of an umbra, since there can be no region in which only part of the source is eclipsed. If the source has an appreciable extent, however, there exists a transition surrounding the umbra, called the penumbra, which is illuminated by only part of the source. Depending on what fraction of the source is exposed, the illumination in the penumbra varies from zero at the edge of the full shadow to the maximum where the entire source is exposed. The edge of the umbra is not perfectly sharp, even with an ideal point source, because of the wave character of light. *See* DIFFRACTION.

For example, the shadow of the Moon shown in the **illustration** has an umbra which barely touches the Earth during a total eclipse of the Sun. The eclipse is only partial for an observer situated in the penumbra. Astronauts circling the Earth or the Moon observe this shadow pattern during each orbit. *See* ECLIPSE.

The term shadow is also used with other types of radiation, such as sound or x-rays. In the case of sound, pronounced shadows are formed only for high frequencies. The high frequency of visible light waves makes possible shadows that are nearly black. X-ray photographs are shadowgrams in which the bones and tissues appear by virtue of their different

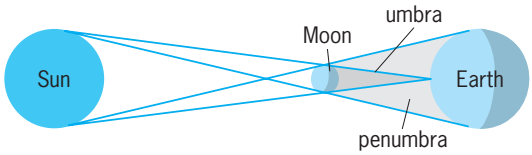

Shadow of the Moon at the time of an eclipse of the Sun. Relative sizes are not to scale.

opacities or degrees of absorption. The bending of light rays by reflection or refraction may also produce shadow patterns. An important practical example is the schlieren method of observing flow patterns in wind tunnels. *See* RADIOGRAPHY; SCHLIEREN PHOTOGRAPHY. Francis A. Jenkins; William W. Watson

Shadowgraph

An optical method of rendering fluid flow patterns visible by using index-of-refraction differences in the flow. As a simple example, sunlight passing through the air above a warm toaster may be seen to form darker and brighter streaks on a surface a meter or two from the toaster. The method relies on the fact that rays of light bend toward regions of higher refractive index while passing through a transparent material. The fluid is usually illuminated by a parallel beam of light. **Figure 1** depicts the method as it might be applied to a fluid sample undergoing thermal convection between two parallel plates, with the lower plate being kept warmer than the upper one. As illustrated, the rays bend toward the cooler down-flowing regions, where the refractive index is higher, and away from the warmer up-flowing ones. After they have passed through the fluid layer, the rays tend to focus above the cooler regions and defocus above the warmer regions. If an image of the light beam is recorded not too far from the sample, brighter areas of the image will lie above regions of down flow, where the rays have been concentrated, and darker areas will lie above regions of up flow. If an image is recorded at a distance greater than the distance for which the rays actually focus, then it is possible to have a situation in which darker, rather than brighter, areas lie above down-flowing regions. In practice this is not a problem, because the image can be observed beginning near the sample and gradually moving away, until a suitable situation is achieved. Because the light passes completely through the sample, the bending effect for each ray is averaged over the sample thickness. *See* CONVECTION (HEAT); REFRACTION OF WAVES.

In convection experiments the refractive index varies because of thermal expansion of the fluid, but the method is not restricted regarding the mechanism responsible for disturbing the refractive index. Thus the same method may be used to visualize denser and less dense regions in a gas flowing in a wind tunnel, including Mach waves and shock waves, where the denser regions have a higher-than-average refractive index. *See* SHOCK WAVE; SUPERSONIC FLOW; WIND TUNNEL.

In a modern research apparatus, images are usually recorded by means of a charge-coupled-device (CCD) camera, digitized, and stored in a computer. Such a digitized image consists of an array of numbers, each number being proportional to the brightness at a particular point in the image. The image points (pixels) form a closely spaced rectangular grid. A reference image may be taken in the absence of any fluid flow, and the reference image may be divided point by point into images taken with the fluid moving. The divided image is then almost free of artifacts associated with nonuniformity of optical elements, such as windows, as well as nonuniformity of the intensity of the original beam (**Fig. 2**). *See* CHARGE-COUPLED DEVICES. David S. Cannell

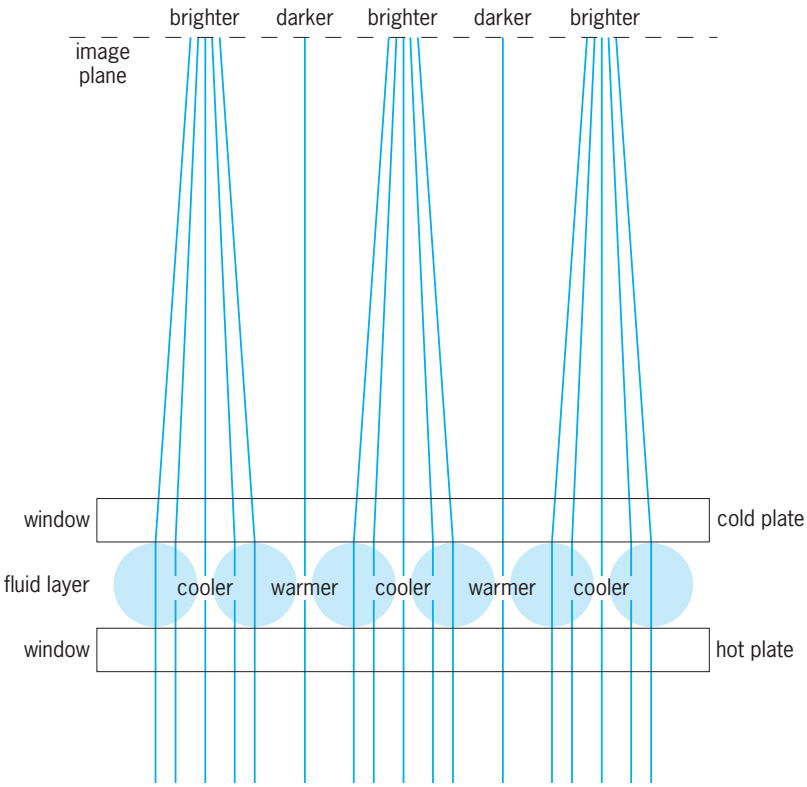

Fig. 1. Schematic of the shadowgraph method applied to reveal convection patterns in a fluid layer. A cross section of the apparatus perpendicular to the convection rolls is shown.

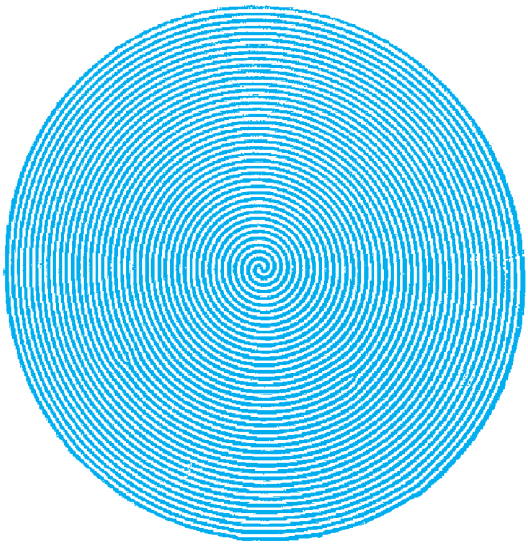

Fig. 2. Background-divided shadowgraph image of convection in high-pressure carbon dioxide gas. The convection rolls are in the form of a two-armed spiral lying within a set of concentric circular rolls. (*From E. Bodenschatz et al., Transitions between patterns in thermal convection, Phys. Rev. Lett., 67:3078–3081, 1991*)

Bibliography. J. R. de Bruyn et al., Apparatus for the study of Rayleigh-Bénard convection in gases under pressure, *Rev. Sci. Instrum.*, 67(6):2043–2067, 1996; R. J. Goldstein (ed.), *Fluid Mechanics Measurements*, 2d ed., Taylor & Francis, 1996; D. R. Jenkins, Interpretation of shadowgraph patterns in Rayleigh-Bénard convection, *J. Fluid Mech.*, 190:451–469, 1988; W. Merzkirch, *Flow Visualization*, 2d ed., Academic Press, Orlando, FL, 1987; G. S. Settles, *Schlieren and Shadowgraph Techniques*, Springer-Verlag, 2001.

Shaft balancing

The process (often referred to as rotor balancing) of redistributing the mass attached to a rotating body in order to reduce vibrations arising from centrifugal force.

Static versus dynamic balancing. A rotating shaft supported by coaxial bearings (for example, ball bearings) together with any attached mass (such as a turbine disk or motor armature) is called a rotor. If the center of mass of a rotor is not located exactly on the bearing axis, a centrifugal force will be transmitted via the bearings to the foundation. The horizontal and vertical components of this force are periodic shaking forces that can travel through the foundation to create serious vibration problems in neighboring components. The magnitude of the shaking force is $F = me\omega^2$, where m = rotor mass, ω = angular speed, and e = distance (called eccentricity) of the center of mass from the rotation axis. The product me, called the unbalance, depends only on the mass distribution, and is usually nonzero because of unavoidable manufacturing tolerances, thermal distortion, and so on. When the center of mass lies exactly on the axis of rotation ($e = 0$), no net shaking force occurs and the rotor is said to be in static balance. Static balance is achieved in practice by adding or subtracting balancing weights at any convenient radius until the rotor shows no tendency to turn about its axis, starting from any given initial orientation. Static balancing is adequate for relatively thin disks or short rotors (such as automobile wheels). However, a long rotor (for example, a turbogenerator) may be in static balance and still exert considerable forces upon individual bearing supports. Suppose that a long shaft is supported on two bearings, and that the unbalance m_1e_1 of the left-hand half of the shaft is equal in magnitude to the unbalance m_2e_2 of the right-hand side, but e_1 and e_2 have opposite signs. In this case, the rotor is in static balance; however, equal but oppositely directed centrifugal forces act on each of the two bearings, transmitting a so-called shaking couple to the foundation. This condition is called dynamic unbalance.

It may be shown that any rigid shaft may be dynamically balanced (that is, the net shaking force and the shaking couple can be simultaneously eliminated) by adding or subtracting a definite amount of mass at any convenient radius in each of two arbitrary transverse cross sections of the rotor. The so-called balancing planes selected for this purpose are usually located near the ends of the rotor, where suitable shoulders or balancing rings have been machined to permit the convenient addition of mass (lead weights, calibrated bolts, and so on) or the removal of mass (by drilling or grinding). Long rotors, running at high speeds, may undergo appreciable elastic deformations. For such flexible rotors it is necessary to utilize more than two balancing planes.

Balancing machines and procedures. The most common types of rotor balancing machines consist of bearings held in pedestals that support the rotor, which is spun at constant speed (as by a belt drive or a compressed air stream). Electromechanical transducers sense the unbalance forces (or associated vibrations) transmitted to the pedestals, and electric circuits automatically perform the calculations necessary to predict the location and amount of balancing weight to be added or subtracted in preselected balancing planes.

For very large rotors, or for rotors which drive several auxiliary devices, commercial balancing machines may not be convenient. The field balancing procedures for such installations may involve the use of accelerometers on the bearing housings, along with vibration meters and phase discriminators (possibly utilizing stroboscopy) to determine the proper location and amount of balance weight.

Committees of the International Standards Organization (ISO) and the American National Standards Institute (ANSI) have formulated recommendations for the allowable quality grade $G = e\omega$ for various classes of machines. A typical value of G (for fans, machine tools, and so on) is 0.25 in./s or 6.3 mm/s, but values as high as 63 in./s or 1600 mm/s (for crankshaft assemblies of large two-cycle engines) and as low as 0.02 in./s or 0.4 mm/s (for gyroscopes and precision grinders) have been recommended. *See* MECHANICAL VIBRATION; SHAFTING.

Burton Paul

Bibliography. J. P. Den Hartog, *Mechanical Vibrations*, 1956, reprint 1985; F. Fujisaw et al., Experimental investigation of multi-span rotor balancing using least squares method, *J. Mech. Des., Trans. ASME*, 102:589–596, 1980; C. M. Harris, *Shock and Vibration Handbook*, 4th ed., 1995; B. Paul, *Kinematics and Dynamics of Planar Machinery*, 1979; S. S. Rao, *Mechanical Vibrations*, 3d ed., 1995.

Shafting

The machine element that supports a roller and wheel so that they can perform their basic functions of rotation. Shafting, made from round metal bars of various lengths and machined to dimension the surface, is used in a great variety of shapes and applications. Because shafts carry loads and transmit power, they are subject to the stresses and strains of operating machine parts. Standardized procedures have been evolved for determining the material characteristics and size requirements for safe and economical construction and operation.

Types. Most shafting is rigid and carries bending loads without appreciable deflection. Some shafting is highly flexible; it is used to transmit motion around corners.

Solid shafting. The normal form of shafting is a solid bar. Solid shafting is obtainable commercially in round bar stock up to 6 in. (15 cm) in diameter; it is produced by hot-rolling and cold-drawing or by machine-finishing with diameters in increments of $^1/_4$ in. (6 mm) or less. For larger sizes, special rolling procedures are required, and for extremely large shafts, billets are forged to the proper shape. Particularly in solid shafting, the shaft is stepped to allow greater strength in the middle portion with minimum diameter on the ends at the bearings. The steps allow shoulders for positioning the various parts pressed onto the shaft during the rotor assembly.

Hollow shafting. To minimize weights, solid shafting is bored out or drilled, or hollow pipes and tubing are used. Hollow shafts also allow internal support or permit other shafting to operate through the interior. The main shaft between the air compressor and the gas turbine in a jet aircraft engine is hollow to permit an internal speed reduction shaft with the minimum requirement of space and weight. A hollow shaft, to have the same strength in bending and torsion, has a larger diameter than a solid shaft, but its weight is less. The center of large shafts made from ingots are often bored out to remove imperfections and also to allow visual inspection for forging cracks.

Functions. Shafts used in special ways are given specific names, although fundamentally all applications involve transmission of torque.

Axle. The primary shafting connection between a wheel and a housing is an axle. It may simply be the extension of a round member from each side of the rear of a wagon, and on the end of each the hub of a wagon wheel rotates. Similarly, railroad car axles are large, round bars of steel spanning between the car wheels, supporting the car frame with bearings on the axle outside the wheels. Axles normally carry only transverse loads, as in the examples above, but occasionally, as in rear automobile housings, they also transmit torsion loads. *See* WHEEL AND AXLE.

Spindle. A short shaft is a spindle. It may be slender or tapered. A spindle is capable of rotation or of having a body rotate upon it. It is similar to an arbor or a mandrel, and usage defines the small drive shaft of a lathe as a live spindle. The term originated from the round tapering stick on a spinning wheel on which the thread is twisted.

Head. A short stub shaft mounted as part of a motor or engine or extending directly therefrom is a head shaft. An example is the power takeoff shaft on a tractor.

Countershaft. A secondary shaft that is driven by a main shaft and from which power is supplied to a machine part is called a countershaft. Often it is driven by gears, and thus rotates counter to the direction of the main shaft. Countershafts are used in gear transmissions to obtain speed and torque changes in transmitting power from one shaft to another.

Jackshaft. A countershaft, especially when used as an auxiliary shaft between two other shafts, is termed a jackshaft.

Line shafting. One or more pieces of shafting joined by couplings is used to transmit power from, for example, an engine to a remotely located machine. A single engine can drive many lines of shafting which, in turn, connect in multiple fashion to process equipment machines. Belts operate on pulleys to transmit the torque from one line to another and from the shafting to the machines. Clutches and couplings control the transfer of power from the shafting. *See* CLUTCH; COUPLING.

The delivery of power to the machines in a shop has generally been converted from line shafting to individual electric motor drives for each machine. Thus, in a modern processing plant, line shafting is obsolete. *See* BELT DRIVE; PULLEY; SHAFT BALANCING.

James J. Ryan

Shale

A class of fine-grained clastic sedimentary rocks with a mean grain size of less than 0.0625 mm (0.0025 in.), including siltstone, mudstone, and claystone. One-half to two-thirds of all sedimentary rocks are shales. *See* SEDIMENTARY ROCKS.

Origin and deposition. Shale is deposited as mud, that is, small particles of silt and clay. These particles are produced by weathering and erosion on land and are transported as a hydraulic suspension. The particles are deposited when fluid turbulence caused by currents or waves is no longer adequate to counteract the force of gravity, or if the water evaporates or infiltrates into the ground. There is sufficient turbulence in most bodies of water to keep clay-sized particles suspended, but clay particles often form larger aggregates which settle from suspension more rapidly than individual particles.

Aggregates may be formed by the inorganic process of flocculation, or by filter-feeding organisms which excrete fecal material made up of particle aggregates. Aggregates formed by flocculation tend to be arranged edge-to-face at oblique angles, whereas the particles in aggregates formed by organisms may form stacks of face-to-face clay particles called domains.

Individual silt particles and clay aggregates are often deposited as thin layers less than 10 mm (0.4 in.) thick called laminae. The arrangement of alternate silt and clay laminae in many shales indicates that the silt laminae are traction deposits of weak currents, whereas clay laminae are usually deposited from suspension (see **table**). *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS.

Compaction and diagenesis. When deposited, the particle aggregates in clay laminae have highly disordered orientations, which results in a high percentage of void space (porosity) occupied by water. After deposition, the mud may be buried by deposition of more sediment. The weight of the overlying sediment causes compaction. During compaction, the

Sedimentary structures in shale

Type of structure	Interpretation	Arrangement
Massive beds	Bioturbation or very rapid sedimentation	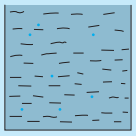
Graded beds	Deposition from decelerating muddy turbidity currents	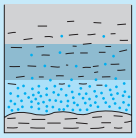
Horizontal lamination	Traction deposition by bottom currents	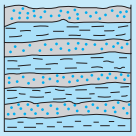
Ripple lamination	Ripple formation by by bottom currents	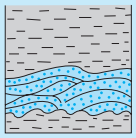
Convolute lamination	Loading of rapidly deposited silt laminae into muddy sediment	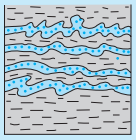
Bottom structures	Erosion or dewatering of mud due to silt deposition	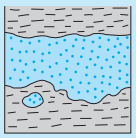
Slumps	Slumping during or after deposition	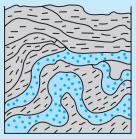
Burrows	Burrowing by organisms	

Key: silt clay 0.8 in. 20 mm

domains or floccules may be rearranged into a more nearly parallel alignment, which results in a loss of porosity, a reduction in average pore size, and the expulsion of water. The amount of porosity loss which occurs at a specific burial depth is dependent on a number of factors, including the original orientation of the particles, the amount of silt, and the rate of burial. The curve in **illus.** *a* shows the general trend of porosity loss for many shales; the most rapid loss of porosity occurs at shallow depths, and the rate of porosity loss decreases exponentially. *See* MARINE SEDIMENTS.

Diagenesis occurs when the elevated temperatures associated with deep burial cause some of the mineral and organic matter in shales to be transformed into more stable phases. Two chemical transformations in shales which have been studied in detail are the conversion of smectite to illite and the maturation of organic matter into hydrocarbons. The extent of diagenesis depends on the original composition of the sediment, the composition of the fluids which saturate the shale, the rate of deposition, and the thermal gradient. *See* DIAGENESIS.

Mineralogical and chemical composition. Mineralogically, most shales are made up of clay minerals, silt-sized quartz and feldspar grains, carbonate cements, accessory minerals such as pyrite and apatite, and amorphous material such as volcanic glass, iron and aluminum oxides, silica, and organic matter.

The most common clay minerals in shales are smectite, illite, kaolinite, and chlorite. The type of clay particles deposited is dependent on the mineralogy, climate, and tectonics of the source area. Paleozoic shales rarely contain smectite or kaolinite, which is at least partially due to the diagenetic conversion of smectite to illite. However, smectite and kaolinite are weathering products commonly formed in soils. The absence of smectite and kaolinite in most Paleozoic rocks may also be related to less intense chemical weathering before land plants evolved in the late Paleozoic. Clay minerals also differ in particle size and in tendency to flocculate; clay mineralogy is thus also dependent upon the distance from the source and the depositional environment. *See* CLAY MINERALS; CHLORITE; ILLITE; KAOLINITE; MONTMORILLONITE; PALEOZOIC.

The amount and size of silt-sized quartz and feldspar in shales are related to the distance of transport and the depositional mechanism. Carbonate cements, amorphous material, and accessory minerals such as pyrite often precipitate in open pores or replace other minerals during burial as a result of diagenesis.

Organic material is normally oxidized shortly after deposition. Shales that are rich in organic material are normally black or dark in color and indicate zones of high productivity, rapid deposition and burial, or lack of oxygen in the depositional environment. After burial, some types of organic matter mature into hydrocarbons; this process is the major source of petroleum and natural gas deposits. *See* NATURAL GAS; PETROLEUM.

Physical and mechanical properties. Most physical and mechanical properties of shale are highly dependent upon changes which occur during burial as a result of compaction and diagenesis. For example, the bulk density of shale is a function of the average grain density, the fluid density, and the porosity. However, fluid and grain density variations are usually insignificant relative to the large amount of porosity lost during burial. Thus bulk density of shale tends to have an inverse relation to the porosity trend (illus. *b*); organic-rich shales generally are less dense,

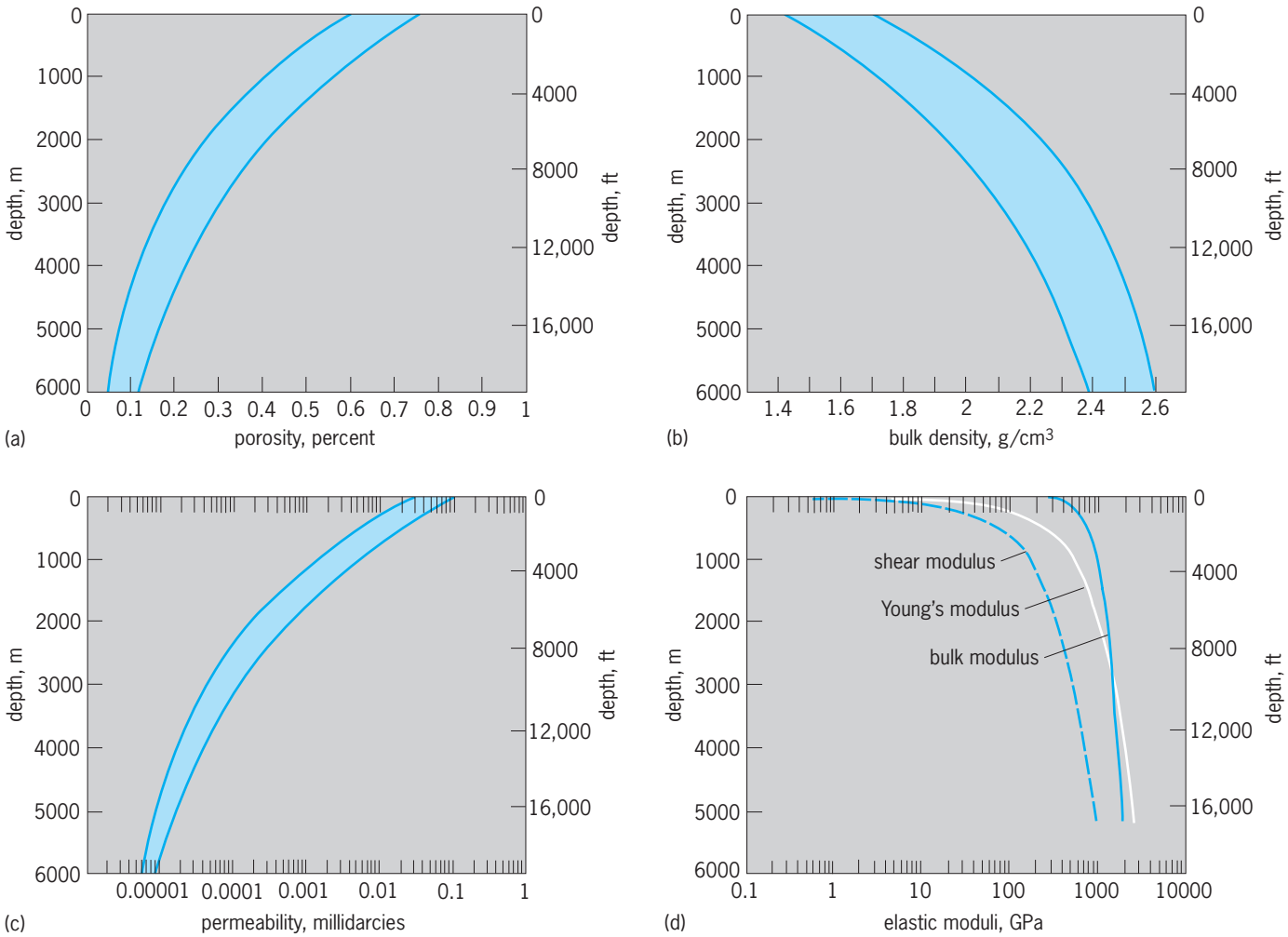

Depth trends of shale properties: (*a*) porosity, (*b*) bulk density, (*c*) permeability, and (*d*) elastic moduli. Shaded areas show the range of values for most shales.

and shales which contain carbonate cement are more dense than average.

The small size of pores in shale relative to those in sandstone causes shale permeability to be much lower than sand permeability. Permeability in shale decreases exponentially as a function of porosity (illus. *c*); the exponential decrease of porosity with depth causes permeability to decrease four to five orders of magnitude with burial to depths of about 3000 m (10,000 ft). Although fracturing due to compaction stresses or to tectonic movements can create high-permeability deviations from this general trend, shales often form permeability barriers to fluid movement; this has important bearing on the occurrence of subsurface water and hydrocarbons. Ground-water aquifers are commonly confined by an underlying low-permeability shale bed or aquiclude, which prevents further downward movement of the water. Hydrocarbon reservoirs are often capped by low-permeability shale which forms an effective seal to prevent hydrocarbons from escaping. *See* AQUIFER.

The high water content of mud before it is compacted generally results in much less mechanical strength than in uncompacted sand. The realignment of particles, rapid porosity loss, and decrease in pore size cause the mechanical strength to increase rapidly with compaction (illus. *d*). However, even after compaction, the strength of shale is usually less than for most other rocks.

Classification. Shales are usually classified or described according to the amount of silt, the presence and type of lamination, mineralogy, chemical composition, and color. Variations in these properties are related to the type of environment in which the shale was deposited and to postdepositional changes caused by diagenesis and compaction. Fissility, the tendency for shales to split along smooth surfaces parallel to bedding planes, is sometimes confused with lamination. However, fissility is a weathering phenomenon, and although it may be related to lamination, fissility is no longer recommended for use in shale classification.

Most of the individual particles or aggregates which make up shale are too small to be resolved by the human eye, but without using a microscope or other laboratory equipment, shales can be described by determining the presence of silt, presence and type of lamination, presence of carbonate cement, and color. More complete description

of shales requires optical and electron microscopy, x-radiography, x-ray diffraction, and chemical analyses to accurately determine the details of particle size distribution, particle orientation, mineralogy, and chemical composition. *See* PETROFABRIC ANALYSIS; PETROLOGY. Joseph R. Davis

Bibliography. K. Frye, *Mineral Science: An Introductory Survey*, 1992; C. Klein, *Manual of Mineralogy*, revised 22nd ed., 2001; D. Prothero and F. Schwab, *Sedimentary Geology: An Introduction to Sedimentary Rocks and Stratigraphy*, 1996; C. E. Weaver, *Clays, Muds, and Shales (Developments in Sedimentology)*, 1989.

Shape memory alloys

A group of metallic materials that can return to some previously defined shape or size when subjected to the appropriate thermal procedure. That is, shape memory alloys can be plastically deformed at some relatively low temperature and, upon exposure to some higher temperature, will return to their original shape. Materials that exhibit shape memory only upon heating are said to have a one-way shape memory, while those which also undergo a change in shape upon recooling have a two-way memory. Typical materials that exhibit the shape memory effect include a number of copper alloy systems and the alloys of gold-cadmium, nickel-aluminum, and iron-platinum.

A shape memory alloy may be further defined as one that yields a thermoelastic martensite, that is, a martensite phase that is crystallographically reversible. In this case, the alloy undergoes a martensitic transformation of a type that allows the alloy to be deformed by a twinning mechanism below the transformation temperature. A twinning mechanism is a herringbone structure exhibited by martensite during transformation. The deformation is then reversed when the twinned structure reverts upon heating to the parent phase.

Transformation characteristics. The martensitic transformation that occurs in shape memory alloys yields a thermoelastic martensite and develops from a high-temperature austenite phase with long-range order. The martensite typically occurs as alternately sheared platelets, which are seen as a herringbone structure when viewed metallographically. The transformation, although a first-order phase change, does not occur at a single temperature but over a range of temperatures that is characteristic for each alloy system.

There is a standard method of characterizing the transformation and naming each point in the cycle. Most of the transformation occurs over a relatively narrow temperature range, although the beginning and end of the transformation during heating or cooling actually extends over a much larger temperature range. The transformation also exhibits hysteresis in that the transformation on heating and on cooling does not overlap. This transformation hysteresis varies with the alloy system. *See* ALLOY.

Thermomechanical behavior and alloy properties. The mechanical properties of shape memory alloys vary greatly over the temperature range spanning their transformation. The only two alloy systems that have achieved any level of commercial exploitation are the nickel-titanium alloys and the copper-base alloys. Properties of the two systems are quite different. The nickel-titanium alloys have greater shape memory strain, tend to be much more thermally stable, have excellent corrosion resistance compared to the copper-base alloys, and have higher ductility. The copper-base alloys are much less expensive, can be melted and extruded in air with ease, and have a wider range of potential transformation temperatures. The two alloy systems thus have advantages and disadvantages that must be considered in a particular application.

Applications. The manufacturing process and the technology associated with both of the two commercial classes of shape memory alloys are quite different as are the performance characteristics. Therefore, in applications where a highly reliable product with a long fatigue life is desired, the nickel-titanium alloys are the exclusive materials of choice. Typical applications of this kind include electric switches and actuators. However, if high performance is not mandated and cost considerations are important, then the use of copper-zinc-aluminum shape memory alloys can be recommended. Typical applications of this kind include safety devices such as temperature fuses and fire alarms.

In order to maintain a satisfactory shape recovery during manufacturing, the deformation strain must be constrained from exceeding a critical value that is a function of a variety of processing parameters and the service environment. The critical strain depends upon the geometrical attributes of the product, the qualitative and quantitative nature of the loads, and the number of loading cycles that the product is designed to withstand.

Similarly, heating a shape memory alloy product to a temperature above some critical temperature is not recommended. The critical temperature for nickel-titanium is approximately 250°C (480°F), and for CuZnAl approximately 90°C (190°F). Extended exposure to thermal environments above these critical temperatures results in an impaired memory function regardless of the magnitude of the load.

Another technical consideration in the practical application of shape memory alloys pertains to fastening or joining these materials to conventional materials. This is a significant issue due to the fact that shape memory alloys undergo expansions and contractions not encountered in traditional materials. Therefore, if shape memory alloys are welded or soldered to other materials, they can easily fail at the joint when subjected to repeated loading. Alloys of nickel-titanium and copper-zinc-aluminum can also be brazed by using silver filler metals; however, the brazed region can fail because of cyclic loading. It is therefore desirable to devise some other mechanism for joining shape memory alloys to traditional

materials. Shape memory alloys cannot be plated or painted for similar reasons.

Free recovery. In this case, a component fabricated from a shape memory alloy is deformed while martensitic, and the only function required of the shape memory is that the component return to its previous shape upon heating. A prime application is the blood-clot filter in which a nickel-titanium wire is shaped to anchor itself in a vein and catch passing clots. The part is chilled so it can be collapsed and inserted into the vein; then body heat is sufficient to return the part to its functional shape.

Constrained recovery. The most successful example of this type of product is undoubtedly a hydraulic coupling. These fittings are manufactured as cylindrical sleeves slightly smaller than the metal tubing that they are to join. Their diameters are then expanded while martensitic and, upon warming to austenite, they shrink in diameter and strongly hold the tube ends. The tubes prevent the coupling from fully recovering its manufactured shape, and the stresses created as the coupling attempts to do so are great enough to create a joint that, in many ways, is superior to a weld.

Force actuators. In some applications, the shape memory component is designed to exert force over a considerable range of motion, often for many cycles. Such an application is the circuit-board edge connector. In this electrical connector system, the shape memory alloy component is used to force open a spring when the connector is heated. This allows force-free insertion or withdrawal of a circuit board in the connector. Upon cooling, the nickel-titanium actuator becomes weaker, and the spring easily deforms the actuator while it closes tightly on the circuit board and forms the connections.

An example based on the same principle is a fire safety valve, which incorporates a copper-zinc-aluminum actuator designed to shut off toxic or flammable gas flow when fire occurs.

Proportional control. It is possible to use only a part of the shape recovery to accurately position a mechanism by using only a selected position of the recovery, because the transformation occurs over a range of temperatures rather than at a single temperature. An example is a device for controlling the rate of fluid flow by carefully heating a shape memory alloy component just enough to close a valve the desired amount.

Other applications. A number of applications are based on the pseudoelastic (or superelastic) property of shape memory alloys. Some eyeglass frames use superelastic nickel-titanium alloy to absorb large deformations without damage. Guide wires for steering catheters into vessels in the body have been developed using wire fashioned of nickel-titanium alloy, which resists permanent deformation if bent severely. Arch wires for orthodontic correction also use this alloy.

Shape memory alloys have found application in the field of robotics. The two main types of actuators for robots using these alloys are biased and differential. Biasing uses a coil spring to generate the bias force that opposes the unidirectional force of the shape memory alloy. In the differential type, the spring is replaced with another shape memory alloy, and the opposing forces control the actuation. A microrobot was developed with five degrees of freedom corresponding to the capabilities of the human fingers, wrist, elbow, and shoulder. The variety of robotic maneuvers and operations are coordinated by activating the nickel-titanium coils in the fingers and the wrist in addition to contraction and expansion of straight nickel-titanium wires in the elbow and shoulders. Digital control techniques in which a current is modulated with pulse-width modulation are employed in all of the components in order to control their spatial positions and speeds of operation.

In medical applications, in addition to mechanical characteristics, highly reliable biological and chemical characteristics are very important. The material must not be vulnerable to degradation, decomposition, dissolution, or corrosion in the organism, and must be biocompatible.

Nickel-titanium shape memory alloys have also been employed in artificial joints such as in artificial hip joints. These alloys have also been used for bone plates, for marrow pins for healing bone fractures, and for connecting broken bones. *See* METAL, MECHANICAL PROPERTIES OF; PROSTHESIS. Mel Schwartz

Bibliography. T. W. Duerig, *Engineering Aspects of Shape Memory Alloys*, 1990; S. Eucken (ed.), *Progress in Shape Memory Alloys*, 1992; M. V. Gandi and B. S. Thompson, *Smart Materials and Structures*, 1992; D. E. Hodgson, *Using Shape Memory Alloys: Shape Memory Applications*, 1988.

Shear

A straining action wherein applied forces produce a sliding or skewing type of deformation. A shearing force acts parallel to a plane as distinguished from tensile or compressive forces, which act normal to a plane. Examples of force systems producing shearing action are forces transmitted from one plate to another by a rivet that tend to shear the rivet, forces in a beam that tend to displace adjacent segments by transverse shear, and forces acting on the cross section of a bar that tend to twist it by torsional shear (**Fig. 1**). Shear forces are usually accompanied by normal forces produced by tension, thrust, or bending.

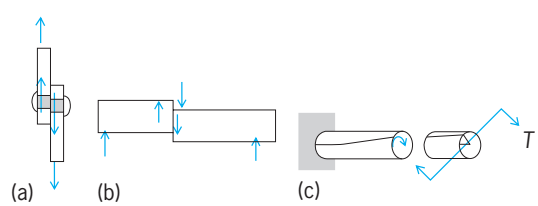

Fig. 1. Shearing actions. (*a*) Single shear on rivet. (*b*) Transverse shear in beam. (*c*) Torsion.

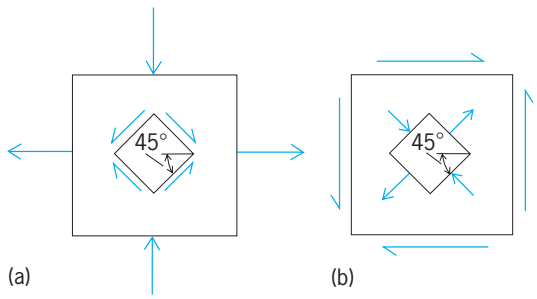

Fig. 2. State of pure shear. (*a*) Shear produces biaxial tension and compression. (*b*) Biaxial stresses induce shear.

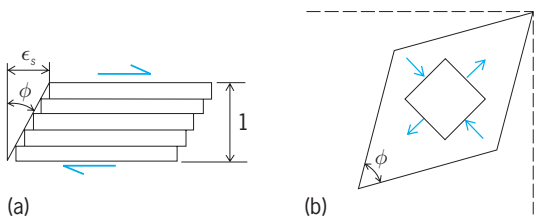

Fig. 3. Shear (*a*) strain and (*b*) distortion.

Shearing stress is the intensity of distributed force expressed as force per unit area. Shear stresses on mutually perpendicular planes at a point in a stressed body are equal. When no normal stresses exist, the state of stress is pure shear, which induces both normal and shear stresses on oblique planes. Elements oriented 45° to planes of pure shear are subjected to biaxial tension and compression equal to the shear stress (**Fig. 2***a*). Similarly, pure shear is induced by equal and opposite biaxial stresses (Fig. 2*b*). Pure shear is the maximum shear stress at the point. Under combined stress, the shear stress is found by principal stress analysis.

Shearing strain is the displacement ϵ_s of two parallel planes, unit distance apart, which accompanies shear stresses acting on these planes. Shearing distortion is visualized as sliding without separation or contraction of all planes parallel to the shear forces, like cards in a pack (**Fig. 3***a*). For planes unit distances apart, the relationship is shown in Eq. (1),

$$\epsilon_s = \tan\phi \approx \phi \text{ radians} \tag{1}$$

where ϕ is the small angle of distortion.

Modulus of rigidity, designated E_s or G, is the shearing modulus of elasticity, which according to Hooke's law is the constant of proportionality between shearing stress S_s and shearing strain S_s/ϕ during elastic behavior.

After shearing distortion, the angles of a rectangular element are altered by the shear strain ϕ (Fig. 3*b*). Changes in length of diagonals must be consistent with the biaxial strains, which leads to the relationship shown in Eq. (2), where E is Young's modulus

$$E_s = \frac{E}{2(1+\mu)} \tag{2}$$

and μ is Poisson's ratio. The value of μ can be found when E_s and E are experimentally determined. *See* STRESS AND STRAIN.

John B. Scalzi

Shear center

A point on a line parallel to the axis of a beam through which any transverse force must be applied to avoid twisting of the section. A beam section will rotate when the resultant of the internal shearing forces is not collinear with the externally applied force. The shear center may be determined by locating the line of action of the resultant of the internal shear forces. A rolled wide flange beam section has two axes of symmetry, and therefore the shear center coincides with the geometric center or centroid of the section. When such a beam member is loaded transversely in the plane of the axes, it will bend without twisting.

However, a channel section with one axis of symmetry is subject to twisting unless the externally applied load passes through the shear center of the section (**Fig. 1**). The shear center for a channel bending about the strong axis (x–x) is located a distance e from the center of the web and lies behind the web. An expression may be derived to determine this eccentricity for different sizes of channels of uniform thickness, such that $e = B^2h^2t/4I_x$, where B is the

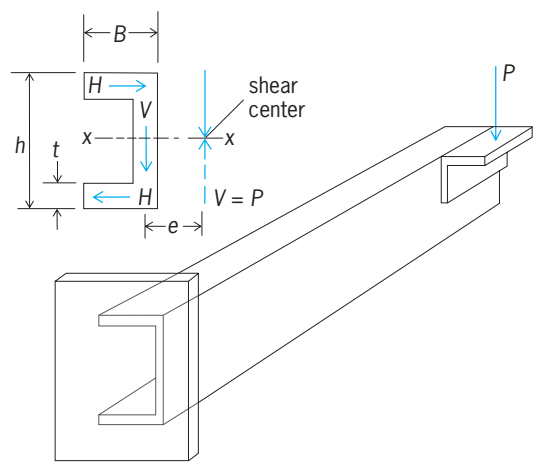

Fig. 1. Shear center under *P* at distance *e* from centerline web, on symmetrical axis.

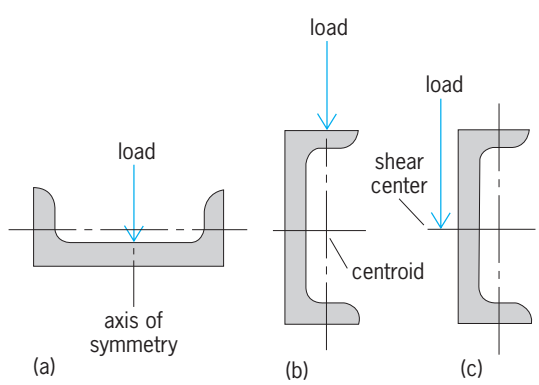

Fig. 2. Three conditions of channel section under load. Channel (*a*) resists bending, (*b*) twists, and (*c*) resists twisting.

flange width, b is the depth of section, t is the thickness of web and flange, and I_x is the moment of inertia about the x axis.

Other structural sections with one axis of symmetry will twist unless properly restrained or loaded through the shear center. **Figure 2***a* illustrates the loading of a channel through the shear center causing bending about the nonsymmetrical axis and no twist; Fig. 2*b* illustrates the condition which will cause twisting; and Fig. 2*c* indicates a condition which will resist twisting. *See* LOADS, TRANSVERSE.

John B. Scalzi

Bibliography. J. S. Iffland (ed.), *Steel Structures*, 1989.

Sheep

Sheep are members of the family Bovidae in the order Artiodactyla, the even-toed, hoofed mammals. Sheep were possibly the first animals to be domesticated about 11,000 years ago in southwestern Asia, a time when sheep were closely associated with the religious, civic, and domestic life of the people. They are referred to more often in the Bible than any other animal.

Wild sheep range over the mountainous areas of the Northern Hemisphere. They have long, curved, massive horns and mixed coats with a hairy outercoat and a woolly undercoat. Although wild sheep vary in size, they are usually larger than domestic sheep and have shorter tails.

Distribution. Over 1 billion domestic sheep extend over most of the land areas of the world. They are most heavily concentrated in the warmer parts of the temperate zones and are found in greatest numbers in the Southern Hemisphere in Australia, New Zealand, South Africa, and South America. Other countries where there are high sheep numbers include Russia, China, India, Iran, Turkey, and the United Kingdom. Sheep are more common in drier regions and at high elevations. Hair sheep are found in the wet tropics.

Domestic sheep were first brought to the Americas by Spanish explorers. Sheep were introduced into the English colonies almost as soon as they were settled. By 1840, New England was the center for sheep production in the United States; the center moved westward to Ohio by 1850. This westward movement of sheep continued, encouraged by the availability of year-round grazing, and by 1860 there were many sheep in California and Texas. The sheep population reached a peak in 1884 with over 50 million head. Since then numbers have declined to under 11 million in the 1980s. Sheep are common in the north-central states and from Missouri to Virginia, but the bulk are in the western states and Texas.

Terminology. Sheep are called lambs until about 12 months of age, from which time to 24 months they are referred to as yearlings, and thereafter as two-year-olds, and so on. The female sheep is called a ewe and the male, a ram or buck; the castrated male is called a wether. Sheep meat is called lamb or mutton depending on the age at slaughter: Young carcasses, where the forefeet are removed at the break joint or temporary cartilage just above the ankle, are sold as lamb; mutton carcasses are those of older animals where it is necessary to take the feet off at the ankle. Mutton is sometimes less desired, brings a lower price than lamb, and may be used in manufactured meats.

General description. A precise description is difficult since sheep are so variable. Horns, if present, tend to curl in a spiral. If horns occur in both sexes, they are usually larger for the ram. In some breeds only the rams have horns, but many breeds are hornless. The hairy outercoat common to wild sheep was eliminated by breeding in most domestic sheep. Wool may cover the entire sheep, or the face, head, legs, and part of the underside may be bare; or wool may be absent as in some breeds (bred for meat) with a hairy coat or a coat which constantly sheds. Although domestic sheep are mostly white, shades of brown, gray, and black occur, sometimes with spotting or patterns of color. The faces and legs are white in some white breeds while black or mottled in others. The shape of the face and head varies from a small dished face to one with a large prominent nose. The ears may be small and pointed; somewhat straight and moderate-sized or long and pendulous; and erect or drooping. Body form varies from the angular, uneven shape of wool-type breeds to the smooth, rounded, blocky appearance of meat-type breeds. The tail tends to be long in most domestic breeds, but is often removed or docked soon after birth. It may be short in some breeds, but may be bare or completely covered with wool in others. Some desert breeds have large fat deposits on the rump or the tail.

The shape of the body of a sheep is determined by the bony skeleton and the amount and distribution of muscle and fat. The thoracic and lumbar vertebrae have both lateral and vertical extensions which contribute to the shape of the back. The front legs have no direct connections to the body skeleton, but the shoulder blades are attached by muscles. Body weights for mature animals may vary from well under 100 to 400 lb (45 to 181 kg) or more. Body measurements vary widely with breed, sex, age, and degree of fatness.

As in other ruminants the stomach of the sheep has four compartments and is adapted to regurgitation and cud chewing. Prior to regurgitation and further chewing, food is stored in the rumen, the largest compartment where roughages are broken down by microorganisms. When food is swallowed a second time, it passes into the third stomach or omasum, which is made up of lengthwise leaves and is connected to the abomasum or true stomach. *See* DIGESTIVE SYSTEM.

The sheep has 20 lamb or temporary teeth, and 32 permanent teeth of which 8 are incisors; a cartilaginous pad is found instead of incisors in the upper jaw. The sheep bites by clamping herbage between the incisor teeth and the dental pad and then jerking or twisting the head.

Although sheep mature in 1–2 years, they may not reach maximum weights and production until 3–5 years of age. The average life-span is about 6–7 years with the productive life sometimes being shorter: survival beyond 10–11 years is rare. Replacement of the lamb incisors with a pair of larger permanent incisors is generally used to estimate age of sheep, although this practice is quite inexact. The first central pair of permanent incisors appears around 1 year of age, and an additional pair is added approximately yearly until a full mouth of four pairs is obtained. After 4 years, age may be estimated from the spread, wear, or loss of the teeth. However, the only sure way to determine age is through individual or group identification from birth throughout life.

Wool and skin. Wool fiber is made up largely of insoluble, sulfur-containing proteins or keratins. The crimps or waves of the wool fiber vary from only 1–2 up to 25 or more per inch (1 up to 10 or more per centimeter), but the waves of adjacent fibers are generally parallel so that blocks of fibers in locks of wool or even over the entire fleece tend to be uniformly crimped. Fiber diameter varies from 10 to 70 micrometers. The number of crimps per inch is greater in fine wool which has fibers of smaller diameters. Fiber length is generally recognized as length of staple, that is, length of wool as it grows on the sheep, and this may vary from about 1 to 6 in. (2.5 to 15 cm) or more for 1 year of growth. Length and diameter are associated to some extent and tend to vary with breed, nutrition, and possibly other environmental factors. Wool grows continually at a somewhat regular rate over the lifetime of the sheep even if it is not removed.

The skin of a sheep is about $^1/_{12}$ in. (2 mm) thick, may be pigmented to various degrees, and contains sweat and sebaceous glands which, together with the hair and wool follicles, are generally arranged in a typical pattern. Those follicles which develop prenatally are called primary follicles, and each is associated with a sweat gland, oil gland, and a smooth muscle. Secondary follicles develop later and have only an oil gland. Primary follicles usually occur in groups of three with a variable number of secondary fibers. There are great variations within breeds and on individual animals regarding the total number of fibers per square inch; totals range from less than 5000 to more than 60,000 per square inch (775 to more than 9300 per square centimeter).

Reproduction. Sheep may reach puberty as early as 100–150 days, although 5–10 months is more common. The first estrus in ewe lambs may occur somewhat later than the first sperm production in ram lambs. Sexual maturity may be more closely related to weight than to age and usually occurs when body weights reach 40–60% of mature weight; it may be delayed even beyond 1 year of age if nutrition is poor. Sexual maturity varies with breed, being earlier for fast-growing breeds such as the Suffolk than for slow-growing breeds such as the Merino, and is generally earlier for crossbred lambs than for purebreds. Ram lambs can be used successfully in breeding, but males are more commonly mated first at 18–20 months of age. Although ewes are often bred first to lamb at 2 years of age, it is not uncommon to breed ewe lambs. Ewes breeding first as lambs generally will have greater lifetime lamb production than ewes bred first as yearlings. Conception rates for ewe lambs may vary 10–60% or more.

The breeding season is related to day length and generally commences as days become shorter; most domestic breeds of sheep have a breeding season restricted to the fall and winter months. There is considerable variation in estrous periods among breeds and among individuals, varying from a few periods to almost year-round breeding. The presence of rams near the beginning of the breeding season may stimulate the onset of the first estrus. Rams do not show a restricted breeding season, but semen production and quality are generally higher in the fall and may decline in the spring and summer; summer sterility may occur under continuous high temperatures.

Estrus. In the ewe, estrus can only be recognized by willingness to stand and to allow the ram to mount. Sometimes the vulva is enlarged, and there is usually a liquefaction and flow of mucus from the cervix. Cornified epithelial cells may be observed in the vaginal smear following estrus. Estrus may vary in duration from a few hours to 3 days or more with an average of 1–2 days; it may be shorter when rams are with ewes continuously. Ovulation normally occurs near the end of estrus.

The normal estrous cycle in the ewe varies from 14 to 19 days with an average of 17 days. Ovulation without estrus generally occurs one cycle before the first estrus. A postpartum estrus may occur within the first day or two after parturition but is not accompanied by ovulation. Ovulation and estrus may occur during lactation but both are more likely to occur after the first 4–10 weeks. Giving progesterone for 12–14 days during the breeding season may synchronize estrus, with estrus and ovulation usually occurring 1–5 days after the treatment is stopped. There may be some reduction in fertility if ewes are bred at the first synchronized estrus. *See* ESTRUS.

Ovulation. The number of eggs shed per heat period (ovulation rate) limits the number of lambs produced and therefore is an important aspect of productivity. It varies with year, season, age, breed, and nutrition. The most common ovulation rate as judged from the number of lambs born is one, although two occur often, three occasionally, four sometimes, and five or six rarely; up to nine have been reported. Ovulation rate increases with age, reaching a maximum at 3–6 years of age. It tends to be high in the early part to the middle of the breeding system but decreases toward the end.

Pregnancy. The average duration of pregnancy varies among breeds from 144 to 155 days; normal pregnancies may range from 138 to 159 days in individuals. Early-mating breeds such as the Southdown and Shropshire have short gestation periods of about 145 days, as do high-fertility breeds such as Finnsheep or Romanov with periods of 142–146 days, while late-maturing breeds such as the

Merino or Rambouillet may have average periods as long as 151 days. Crossbred types such as the Columbia or Targhee tend to be intermediate. Pregnancy is shorter for twin lambs than for singles; it may be longer in older ewes.

The sex ratio is generally 49–50% males. Identical twins are rare. A few cases of freemartin lambs have been reported.

A number of reliable pregnancy tests may be used with sheep. Methods which are both rapid and accurate include palpation by laparotomy, observation with peritoneoscope, and ultrasonic detection of pregnancy either by echo or by the Doppler technique. This last method is highly accurate at about 60 days following conception or even earlier. Diagnosis with x-rays and with fetal electrocardiography is possible but of limited usefulness. Hormonal methods are coming into practical use with accurate diagnosis as early as 18 days after conception.

Parturition. The termination of pregnancy is determined in part by hormonal changes in which fetal adrenal activity initiates a series of hormonal trends culminating in parturition. The pelvic ligaments, vagina, and cervix tend to relax near the end of pregnancy. Parturition may be more common in the forenoon and late afternoon. The first signs of parturition are uneasiness, pawing and frequent turning, or lying down and standing up. The forefeet of the lamb generally appear first in front of the head; one or both front legs turned back are regarded as malpresentations. The lamb is licked dry by the mother following birth and sucks milk within an hour or two. The afterbirth is generally expelled 2–4 h after parturition. The uterus normally returns to nonpregnant size in 2 weeks and is completely involuted by 4 weeks. Synchronization of parturition is feasible with hormonal treatments.

Male physiology. The spermatogenic cycle in the ram is about 49 days. The average sperm production has been measured at 5.5 billion per day. High-quality ram semen is characterized by having at least 2 billion sperm per cubic centimeter, a rapid swirling movement, a pH about 7.0 or with more acidity, and with less than 25% abnormal sperm or less than 5% abnormal heads. Rams show wide variation in frequency of copulation from once in about 10 h to over once an hour. The number increases in direct proportion to the number of available ewes in heat.

Altering reproductive cycles. Artificial insemination techniques have been developed for sheep, but fresh semen gives higher fertility than frozen semen. Practical use of artificial insemination generally requires synchronization of estrus, involvement of genetic improvement, and repeated reproductive cycles to reduce the length of dry periods.

The reproductive rate of sheep may be increased by selection, choice of breed, crossing with high-fertility breeds, crossbreeding, control of reproduction, multiple lambing per year, single-gene high ovulation rate, immunizing techniques, and milk replacers. Combinations of these may be used, but selection is basic because it improves all aspects of weight of lambs weaned of ewes mated, a large part of efficiency, and also improves adaptability to the particular environment. Complete control of reproduction in sheep is now feasible through synchronization of estrus, artificial insemination, early pregnancy diagnosis, and synchronization of parturition, but is not in general use. *See* BREEDING (ANIMAL).

Adaptation. Although sheep are well adapted to cold climates, they have excellent tolerance to heat because of the excellent temperature regulation effected by their wool. Shearing may increase heat tolerance in summer and decrease it in cooler periods. Body temperature maintained at about 102°F (39°C) may vary from about 98 to 106°F (37 to 41°C). The most important source of heat loss is evaporation from the respiratory tract; however, sheep do sweat to a limited extent.

Sheep are probably most productive in temperate climates. Hot climates are unfavorable for lamb meat or market lamb production because of reduced fertility and reduced feed intake. Wool may usually be produced efficiently in hot, dry climates. Internal parasites become abundant and cause serious losses in warm, humid climates. Both lamb and wool are produced well in cold climates, but environmental protection may be necessary at lambing, after shearing, and under wet or stormy conditions.

Breeds. Hundreds of breeds of sheep of all types, sizes, and colors are found over the world; some of the more prominent breeds are shown in the **illustration**. Wool-type breeds, mostly of Merino origin, are important in the Southern Hemisphere, but both fine- and long-wool types are distributed all over the world. Sheep with fat tails or fat rumps are common in the desert areas of Africa and Asia. These usually produce carpet wool. Milk breeds are found mostly in central and southern Europe. Meat breeds from the British Isles are common over the world.

Over 40 breeds of sheep, about half of which were imported from the British Isles, are represented in the United States. The Suffolk and Hampshire breeds represent over half of the purebred registered sheep of the country and are used as sires of slaughter lambs. Other British breeds used for this purpose include Southdown, Shropshire, Dorset, Oxford, and Cheviot. The Dorset is favored for its ability to breed in the spring and fall. About 90% of registered Dorsets are polled.

In the United States, fine wool breeds of the Merino from Spain and the Rambouillet from France and Germany have formed the basis for much of the western range sheep which are often crossed with meat-type rams to produce slaughter lambs. These two European breeds and the Debouillet resulting from their cross are used in the Southwest where wool production is important.

Breeds from long-wool–fine-wool breed crosses have spread over the entire United States and are truly dual-purpose in type, although they are often crossed with meat-type rams for commercial lamb production. The Corriedale, imported from New Zealand, was first used in the West but is now more common in the Midwest; the Columbia and Targhee

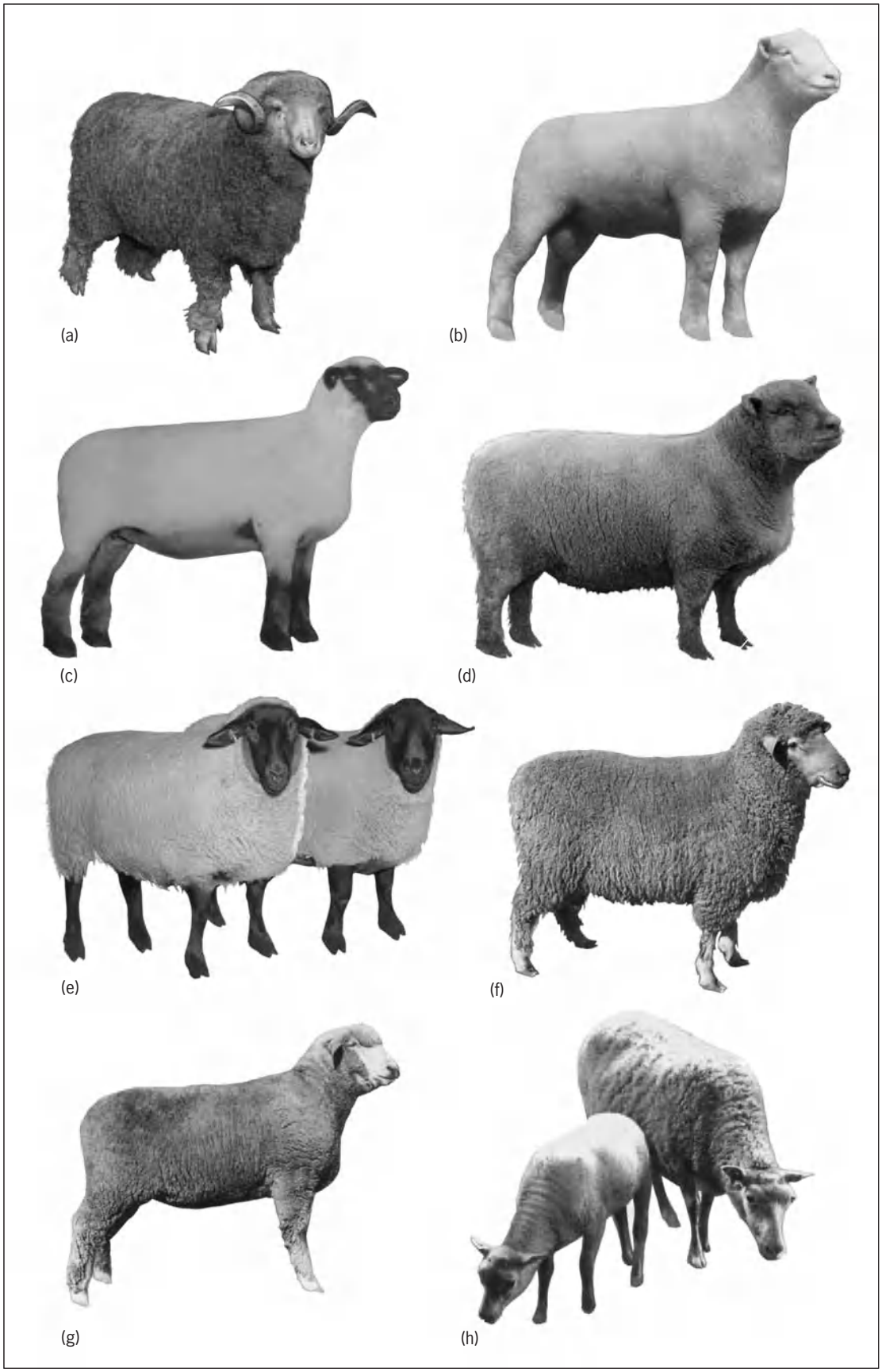

Examples of prominent breeds of sheep. (*a*) Rambouillet ram (*Sheep and Goat Raiser*). (*b*) Dorset ram (*Continental Dorset Club*). (*c*) Hampshire ram. (*d*) Southdown ram (*Southdown Association*). (*e*) Suffolk ewe and ram (*photograph by A. L. Henley*). (*f*) Columbia ram (*photograph by A. Sponagel*). (*g*) Targhee ram (*Cook and Gormley*). (*h*) Finnsheep ewe and lamb (*U.S. Department of Agriculture*).

breeds, developed by the Department of Agriculture, are large, rapidly growing, white-faced, polled sheep. The Columbia, which resulted from crossing Lincoln rams on Rambouillet ewes, produces heavy fleeces of 50–58s in spinning count while the Targhee, a three-quarter Rambouillet and one-quarter long-wool, produces fleeces of 60–62s in spinning count. Another crossbred type is the Montadale, from Columbia–Cheviot.

The Romney is the most numerous of the long-wool breeds, although some Lincolns are maintained primarily for crossing with fine-wool sheep. The Cotswold and Leicester breeds have largely disappeared from the United States. The Tunis and Karakul breeds are quite minor in importance.

Breed improvement. Performance and progeny testing of rams both in central stations and on farms has been carried on in many states. The proportion of purebred sheep selected on production records is small but increasing. The Rambouillet and Hampshire breeds have led in development of plans for certifying rams on the performance of their offspring.

The average flock size for purebred flocks tends to be as small as 25–30 head with only one or two rams used per season. This small flock size limits selection so that breed improvement results from decisions made in only a very few flocks. Wool-type breeds tend to be kept in larger flocks than the meat-type breeds.

Crossbreeding for slaughter lambs has been used extensively for a number of years. Crosses of meat-type rams on wool-type or dual-purpose ewes not only produce desirable slaughter lambs but also result in hybrid vigor which, from two-, three-, and four-breed crosses, has produced important increases in number and pounds of lambs weaned per ewe bred over comparable averages of the purebred parents.

Some exotic breeds might be used to advantage in the United States if disease restrictions on importation could be overcome. Finnsheep have shown their ability to increase lambing rates. Other desirable traits for improvement include adaptability to particular environments, less restricted breeding seasons, ability to produce more than one lamb crop per year, greater efficiency of feed use, and more desirable carcass traits.

Nutrition and feeding. Forages make up about 95% of sheep diets and supply most of the energy needs. Diets should contain 10–15% protein, generally from leguminous plants. Sheep prefer fine forages, but they eat a large variety of grasses, legumes, weeds, herbs, and shrubs. Legume hays are preferable, but corn or sorghum fodder, cereal or grass hays, and straws are often fed. Silages from corn, sorghum, cereal, and other plants may be used to replace part of the hay ration. Nonleguminous roughages may need protein and mineral supplements, but lack of energy may be the most common nutritional deficiency of sheep. Concentrates, including the common farm grains of oats, barley, corn, wheat, and sorghum, are fed sparingly, usually in winter, also prior to and following lambing, and to growing and fattening lambs. Common protein supplements include cake or oil meal from soybeans, flaxseed, or cottonseed. Feeding of by-products, crop residues, and wastes such as poultry or other manures is increasing. *See* ANIMAL FEEDS.

Minerals such as calcium, phosphorus, iodine, copper, and cobalt may need to be supplied. Sheep may do without salt, but if supplied they will normally consume about 1 lb (0.45 kg) per month. Vitamins A, D, and E are required and may have to supplement certain rations. Vitamins C and the B complex are synthesized by sheep.

Feed required for maintenance and growth or production varies with age, weight, phase of reproduction, and climatic conditions. Nonlactating ewes and those in early gestation require 1.6–2.2% of body weight of air-dry feed. Requirements of mature ewes increase in late pregnancy and are greatest in early lactation with a range of 3.2–4.8% of body weight. Lambs and yearlings require 2.2–4.5% of their body weight with greater requirements per unit of body weight at lighter weights.

Pelleting of feeds reduces handling, storage space, and waste. Pelleting of high-roughage diets generally results in increased feed consumption and eliminates selection of the more palatable portions of the roughage. With pellets there is greater feed intake resulting from a more rapid passage of the finely ground material through the digestive tract.

Antibiotics may improve performance of growing and fattening lambs, particularly under conditions of stress. Supplemental or creep feeding of lambs from 2 to 6 weeks of age up to the time of weaning is often advantageous, particularly on ordinary pastures.

Production systems. Sheep under farm conditions are generally kept in fenced pasture from the spring to fall months. Housing, or some shelter, and harvested feed are usually provided in winter, especially in the colder and more humid climates. Flock size is typically small, 25 head or less, but may range up to 150 and sometimes up to several hundred.

Range sheep of the western states are often herded in bands of 700–1000 head of ewes with lambs and up to 2000 head or more of dry ewes. Some rangeland is fenced, and range sheep are sometimes herded only part of the year or not at all. Flock size is generally large with individual holdings often of several thousand up to 15,000 or 20,000. Lambing is often done in sheds although lambing on the range is not uncommon.

Intensive production involving large numbers of sheep in concentrated areas, generally with confinement for part or all of the year, may be practical. This system often involves lambing in the fall and winter months and may include lambing more often than once per year.

Management. Management practices are similar for different production systems although wide variations do exist. Breeding normally occurs in the fall for a period of 45–90 days. In purebred matings individual rams are mated to groups of ewes, while in commercial flocks rams are generally group mated, with from 2 or 3 rams per 100 ewes.

Extra care is usually provided at spring lambing time to increase the survival rate of the lambs. Ewes are commonly shorn or crutched just prior to lambing. The newborn lamb may be assisted in suckling and is often given extra protection from cold or wet weather. Docking and castrating may be done with rubber bands at birth or soon after, or by cutting within 6–10 days or up to 1 month after birth. The umbilical cord of the newborn lamb is dipped in iodine at birth to prevent infection. Vaccination for tetanus, overeating disease, and sore mouth may be done on the young lamb or sometimes later. Purebred lambs may be individually identified with ear tags at birth. Lambs are often identified with their mothers with scourable paint brands.

Shearing to remove the annual growth of wool is usually done in the spring but may range from January to June. Fleeces are individually tied with paper string and are packed in bags holding about 20 to 40 fleeces. Average fleece weight is about 8 lb (3.6 kg).

Lambs may be weaned as early as 3 weeks but weaning under weights of 30–45 lb (13.6–20.4 kg) is not recommended. While weaning age may extend up to 5 months, most ewes decline in milk production as lactation progresses. Lambs are often practically weaned by their mothers by 4 months more or less. The artificial rearing of lambs with milk replacer from birth is desirable when the lamb would otherwise die.

Culling is usually done in the fall. Ewes which fail to reproduce, are unsound, have bad udders, have poor or missing teeth, or are losing weight with age should be eliminated. Ewes in their sixth year or older which fail to wean lambs should generally be culled. Ewes with low lamb production may be culled at any age. Ewes with light fleeces should generally be culled when young.

Disease and parasites. Losses from diseases are moderate to low, although sheep suffer from many maladies. Common diseases include abortion, pneumonia, and overeating disease. Foot rot can be eliminated or controlled by tedious trimming and treating of feet and isolation of infected animals. Vaccination is effective for some diseases, particularly blue tongue and sore mouth. Epididymitis may interfere with ram fertility.

External parasites such as lice or keds are troublesome and may be controlled by dipping, spraying, or dusting. Sheep scabies has been eradicated in the United States.

Internal parasites cause tremendous losses, particularly in warm humid areas. The stomach worm is probably the most common parasite. Weaning the lambs onto clean pasture from dry lot is generally effective in preventing or delaying the increase of parasites on summer pastures.

Products. Wool is an important source of income from sheep but it is less important than lamb meat. Wool is valued or graded chiefly on fineness or fiber diameter, staple length, clean yield, and freedom from contamination with foreign material. Color, strength, handle, and crimp also are often considered. Wool is a superior fiber, even though synthetic fibers are often cheaper, and it is especially useful for warmth, moisture absorbability, and resistance to fire. It is unique in extensibility, resilience or wrinkle resistance, felting ability, and durability. Fine and medium wools are generally used for apparels, and coarse wools for rugs and carpets.

Lamb meat is the most important product of sheep in the United States. High-quality carcasses are thick and blocky with a high proportion of lean, large loin-eye muscle, and with a thin covering of fat. The fat has a high melting point, and the meat has a characteristic flavor. The loin chops, rib chops, and legs are the favored cuts. Consumption of lamb is about 2 lb (0.9 kg) per person per year. Lamb is not well distributed in the United States and therefore not available to many.

Sheep are used for milk production as well as for meat and wool in many countries, particularly in southern Europe. Ewe milk averages about 6% fat and is superior for making cheese. Sheep may yield up to 130 gal (492 liters) of milk annually. Sheep manure is valued in some areas. Skins are used for a variety of lightweight leathers. Pelts may be tanned with the wool on for items such as coats, rugs, and slippers. Fur pelts are produced from young Karakul lambs. Sheep have been used rarely for work and sport. *See* AGRICULTURAL SCIENCE (ANIMAL). Clair E. Terrill

Bibliography. I. E. Coop (ed.), *Sheep and Goat Production*, 1982; M. E. Ensminger and R. O. Parker, *Sheep and Goat Science*, 1985; R. Jensen and B. L. Swift, *Diseases of Sheep*, 1982.

Sheet-metal forming

The shaping of thin sheets of metal (usually less than $^1/_4$ in. or 6 mm) by applying pressure through male or female dies or both. Parts formed of sheet metal have such diverse geometries that it is difficult to classify them. In all sheet-forming processes, excluding shearing, the metal is subjected to primarily tensile or compressive stresses or both. Sheet forming is accomplished basically by processes such as stretching, bending, deep drawing, embossing, bulging, flanging, roll forming, and spinning. In most of these operations there are no intentional major changes in the thickness of the sheet metal. *See* METAL FORMING.

There are certain basic considerations which are common to all sheet forming. Grain size of the metal is important in that too large a grain produces a rough appearance when formed, a condition known as orange peel. For general forming an ASTM no. 7 grain size (average grain diameter 0.00125 in. or 32 micrometers) is recommended. Another type of surface irregularity observed in materials such as low carbon steel is the phenomenon of yield-point elongation that results in stretcher strains or Lueder's bands, which are elongated depressions on the surface of the sheet. This is usually avoided by cold-rolling the original sheet with a reduction of only 1–2% (temper rolling). Since yield-point elongation reappears after

some time, because of aging, the material should be formed within this time limit. Another defect is season cracking (stress cracking, stress corrosion cracking), which occurs when the formed part is in a corrosive environment for some time. The susceptibility of metals to season cracking depends on factors such as type of metal, degree of deformation, magnitude of residual stresses in the formed part, and environment. *See* METAL, MECHANICAL PROPERTIES OF.

Anisotropy or directionality of the sheet metal is also important because the behavior of the material depends on the direction of deformation. Anisotropy is of two kinds: one in the direction of the sheet plane, and the other in the thickness direction. These aspects are important, particularly in deep drawing.

Formability of sheet metals is of great interest, even though it is difficult to define this term because of the large number of variables involved. Failure in sheet forming usually occurs by localized necking or buckling or both, such as wrinkling or folding. For a simple tension-test specimen the true (natural) necking strain is numerically equal to the strain-hardening exponent of the material: thus, for instance, commercially pure annealed aluminum or common 304 stainless steel stretches more than cold-worked steel before it begins to neck. However, because of the complex stress systems in most forming operations, the maximum strain before necking is difficult to determine, although some theoretical solutions are available for rather simple geometries. *See* ALUMINUM; STAINLESS STEEL.

Considerable effort has been expended to simulate sheet-forming operations by simple tests. In addition to bend or tear tests, cupping tests have also been commonly used, such as the Swift, Olsen, and Erichsen tests. Although these tests are practical to perform and give some indication of the formability of the sheet metal, they generally cannot reproduce the exact conditions to be encountered in actual forming operations.

Stretch forming. In this process the sheet metal is clamped between jaws and stretched over a form block. The process is used in the aerospace industry to form large panels with varying curvatures. Stretch forming has the advantages of low die cost, small residual stresses, and virtual elimination of wrinkles in the formed part.

Bending. This is one of the most common processes in sheet forming. The part may be bent not only along a straight line, but also along a curved path (stretching, flanging). The minimum bend radius, measured to the inside surface of the bend, is important and determines the limit at which the material cracks either on the outer surface of the bend or at the edges of the part. This radius, which is usually expressed in terms of multiples of the sheet thickness, depends on the ductility of the material, width of the part, and its edge conditions. Experimental studies have indicated that if bending is carried out in a pressurized environment, with pressures as high as 500,000 lb/in.^{2} (3.4 gigapascals), the minimum bend radius decreases substantially. This observation is important in forming brittle materials. Heating the workpiece is another method of improving bendability.

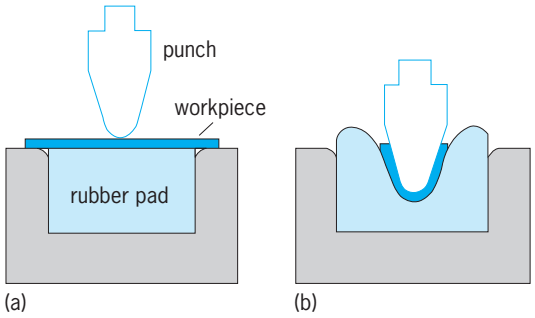

Fig. 1. Bending process with a rubber pad. (*a*) Before forming. (*b*) After forming.

Springback in bending and other sheet-forming operations is due to the elastic recovery of the metal after it is deformed. Determination of springback is usually done in actual tests. Compensation for springback in practice is generally accomplished by overbending the part; adjustable tools are sometimes used for this purpose.

In addition to male and female dies used in most bending operations, the female die can be replaced by a rubber pad (**Fig. 1**). In this way die cost is reduced and the bottom surface of the part is protected from scratches by a metal tool. The roll-forming process replaces the vertical motion of the dies by the rotary motion of rolls with various profiles. Each successive roll bends the strip a little further than the preceding roll. The process is economical for forming long sections in large quantities.

Rubber forming. While many sheet-forming processes are carried out in a press with male and female dies usually made of metal, there are four basic processes which utilize rubber to replace one of the dies. Rubber is a very effective material because of its flexibility and low compressibility. In addition, it is low in cost, is easy to fabricate into desired shapes, has a generally low wear rate, and also protects the workpiece surface from damage.

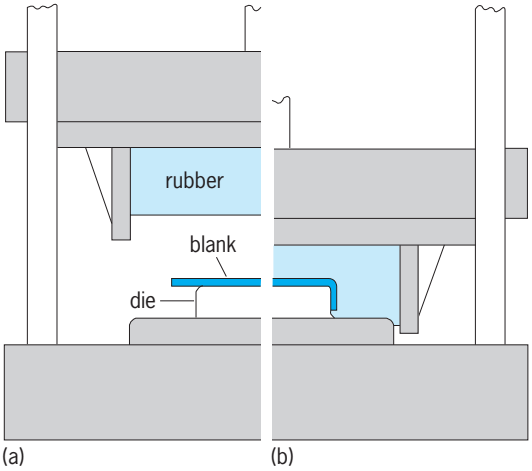

Fig. 2. Guerin process, the simplest rubber-forming process. (*a*) Before forming. (*b*) After forming.

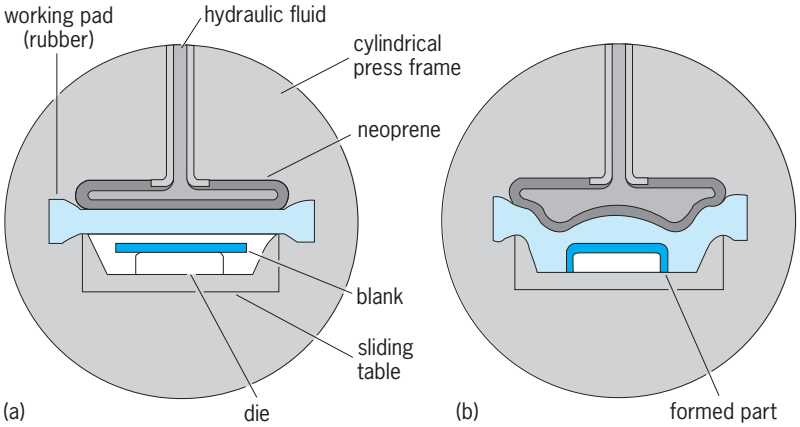

Fig. 3. Verson-Wheelon process. (*a*) Before forming. (*b*) After forming.

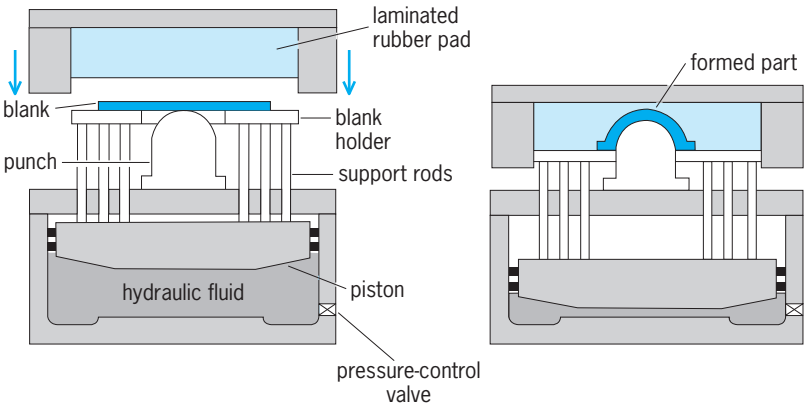

Fig. 4. Marform process.

The simplest of these processes is the Guerin process (**Fig. 2**). Auxiliary devices are also used in forming more complicated shapes. In the Verson-Wheelon process (**Fig. 3**) hydraulic pressure is confined in a rubber bag, the pressure being about five times greater than that in the Guerin process. For deeper draws the Marform process is used (**Fig. 4**). This equipment is a packaged unit that can be installed easily into a hydraulic press. In deep drawing of critical parts the Hydroform process (**Fig. 5**) is quite suitable, where pressure in the dome is as high as 15,000 lb/in.^{2} (100 megapascals). A particular advantage of this process is that the formed portions of the part travel with the punch, thus lowering tensile stresses which can eventually cause failure.

Bulging of tubular components, such as coffee pots, is also carried out with the use of a rubber pad placed inside the workpiece; the part is then expanded into a split female die for easy removal.

Deep drawing. The basic components of a deep-drawing operation are shown in **Fig. 6**. A great variety of parts are formed by this process, the successful operation of which requires a careful control of factors such as blank-holder pressure, lubrication, clearance, material properties, and die geometry. Depending on many factors, the maximum ratio of blank diameter to punch diameter ranges from about 1.6 to 2.3. *See* DRAWING OF METAL.

This process has been extensively studied, and the results show that two important material properties for deep drawability are the strain-hardening exponent and the strain ratio (anisotropy ratio) of the metal. The former becomes dominant when the material undergoes stretching, while the latter is more pertinent for pure radial drawing. The strain ratio is defined as the ratio of the true strain in the width direction to the true strain in the thickness direction of a strip of the sheet metal. The greater this ratio, the greater is the ability of the metal to undergo change in its width direction while resisting thinning.

Anisotropy in the sheet plane results in earing, the appearance of wavy edges on drawn cups. Clearance between the punch and the die is another factor in

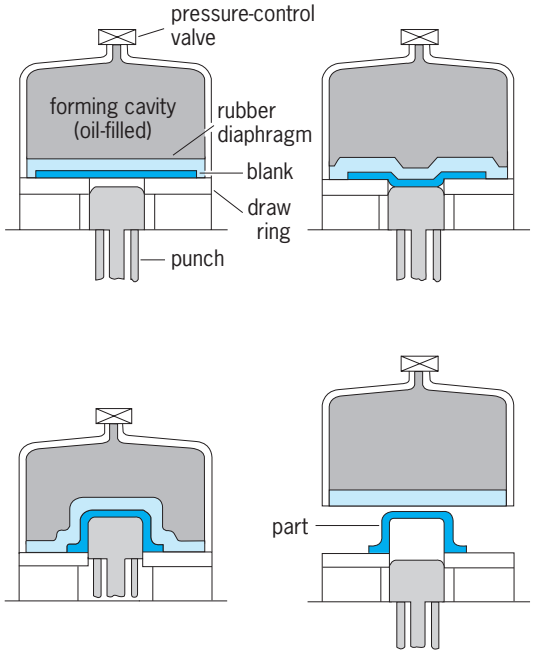

Fig. 5. Hydroform process.

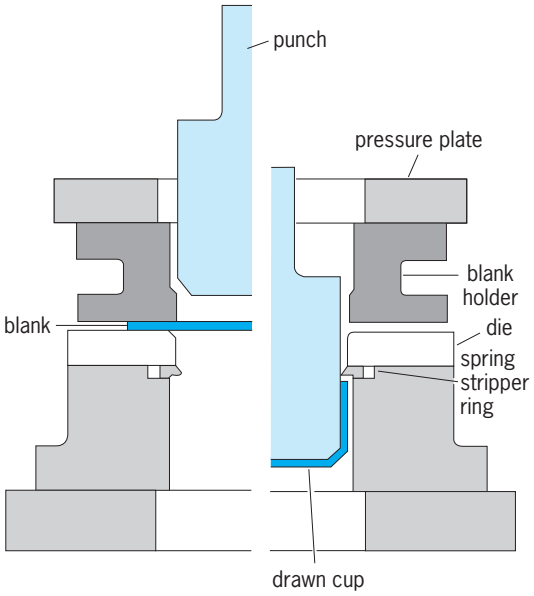

Fig. 6. Deep-drawing process.

this process; this is normally set at a value of not more than 1.4 times the thickness of the sheet. Too large a clearance produces a cup whose thickness increases toward the top, whereas correct clearance produces a cup of uniform thickness by ironing. Also, if the blank-holder pressure is too low, the flange wrinkles; if it is too high, the bottom of the cup will be punched out because of the increased frictional resistance of the flange. For relatively thick sheets it is possible to draw parts without a blank holder by special die designs.

Miscellaneous processes. Many parts require one or more additional processes; some of these are described briefly here. Embossing consists of forming a pattern on the sheet by shallow drawing. Coining consists of putting impressions on the surface by a process that is essentially forging, the best example being the two faces of a coin. Coining pressures are quite high, and control of lubrication is essential in order to bring out all the fine detail in a design. Shearing is separation of the material by the cutting action of a pair of sharp tools, similar to a pair of scissors. The clearance in shearing is important in order to obtain a clean cut. A variety of operations based on shearing are punching, blanking, perforating, slitting, notching, and trimming. *See* COINING.

Sheet-forming equipment. The most common equipment is a mechanical or hydraulic press. Presses may be single, double, or triple acting, depending on the number of slides to do a certain portion of the operation. Common mechanical presses are the eccentric, crank, knuckle joint, and toggle. The design of dies requires considerable experience because of the great number of problems that may arise in production. Tool and die materials for sheet forming are generally made of cast irons, cast alloys, die steels, and cemented carbides for high-production work. Nonmetallic materials such as rubber, plastics, and hardwood are also used as die materials. The selection of the proper lubricant depends on many factors, such as die and workpiece materials, and severity of the operation. A great variety of lubricants are commercially available, such as drawing compounds, fatty acids, mineral oils, and soap solutions. *See* HYDRAULIC PRESS.

Pressures in sheet-metal forming generally range between 1000 and 8000 lb/in.^{2} (7 and 55 MPa) (normal to the plane of the sheet); most parts require about 1500 lb/in.^{2} (10 MPa).

Spinning. This process forms parts with rotational symmetry over a mandrel with the use of a tool or roller. There are two basic types of spinning: conventional or manual spinning, and shear spinning (**Fig. 7**). The conventional spinning process forms the material over a rotating mandrel with little or no change in the thickness of the original blank. Parts can be as large as 20 ft (6 m) in diameter. The operation may be carried out at room temperature or higher for materials with low ductility or great thickness. Success in manual spinning depends largely on the skill of the operator. The process can be economically competitive with drawing; if a part can be made by both processes, spinning may be more economical than drawing for small quantities. *See* SPINNING (METALS).

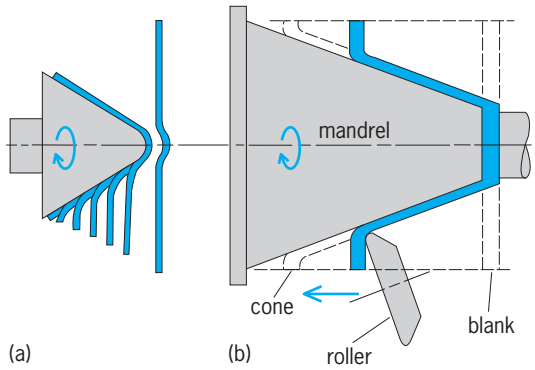

Fig. 7. Spinning processes of (*a*) conventional and (*b*) shear type.

In shear spinning (hydrospinning, floturning) the deformation is carried out with a roller in such a manner that the diameter of the original blank does not change but the thickness of the part decreases by an amount dependent on the mandrel angle. The spinnability of a metal is related to its tensile reduction of area. For metals with a reduction of area of 50% or greater, it is possible to spin a flat blank to a cone of an included angle of 30° in one operation. Shear spinning produces parts with various shapes (conical, curvilinear, and also tubular by tube spinning on a cylindrical mandrel) with good surface finish, close tolerances, and improved mechanical properties. *See* METAL COATINGS. Serope Kalpakjian

Bibliography. R. Baril, *Modern Machining Technology*, 1987; D. F. Eary and E. A. Reed, *Techniques of Pressworking Sheet Metal: An Engineering Approach to Die Design*, 2d ed., 1997; R. A. Lindberg, *Processes and Materials of Manufacture*, 4th ed., 1990.

Shellac

The lac resin (secreted by the lac insect) when used in flake (or shell) form. Shellac varnish is a solution of shellac in denatured alcohol.

Shellac varnish is used in wood finishing where a fast-drying, light-colored, hard finish is desired. Drying is by simple evaporation of the alcohol. Shellac varnish is not water-resistant and is not suitable for exterior coatings. When used as a finish, it has the distinct advantage that it remains soluble. When touch-up is required, it therefore merges completely with the original finish, and no scratches or worn spots show.

Natural shellac has a brown color and is often referred to as orange shellac. Bleaching with chlorine produces bleached or white shellac. Bleached shellac is slightly less durable and less hard than orange shellac but is lighter in color.

Shellac consists of a number of acids and esters and about 3% shellac wax, which is removed to make dewaxed shellac. Shellac varnishes are usually sold

on the basis of the number of pounds of shellac dissolved in a gallon of alcohol, as 3-, 4-, or 5-lb (1.4-, 1.8-, or 2.3-kg) cuts. Shellac varnish should be freshly prepared because it deteriorates on aging. Drying rates become slower as the number of pounds of shellac per gallon increases. *See* PAINT; SURFACE COATING; VARNISH. C. R. Martinson; C. W. Sisler

Bibliography. R. Lambourne, *Paints and Surface Coatings: Theory and Practice*, 1999; D. Stoye (ed.), *Paints, Coatings, and Solvents*, 2d ed., 1998.

Ship design

A process which translates a set of owner's requirements into the drawings, specifications, and other technical data necessary to actually build a ship. Naval architects lead the process, but engineers and designers with many other skills contribute. These other skills include marine engineering, structural design, and production engineering. The ship design process is iterative, and is subdivided into several phases during which the design is developed in increasing degrees of detail. Typically, the owner's requirements specify the mission that the new ship must perform and define such parameters as required speed, fuel endurance, and cargo weight and capacity. Generally, the cost to build and operate a ship is constrained by the prospective owner since the ship will be part of a transportation or industrial system, which must be profitable. The ship design process involves numerous trade-off studies in order to achieve the desired capability and, at the same time, stay within the established cost. *See* NAVAL ARCHITECTURE.

Mission requirements. Ships are designed and built to satisfy the mission requirements and constraints specified by the prospective owner. These requirements and constraints are unique to each ship being considered and form the basis for design development. For some ships, such as point-to-point cargo ships, the mission requirements can be simply stated; for example, "Transport 5000 20 ft ISO standard cargo containers at an average sea speed of 18 knots with 10,000 nautical miles between refuelings. On- and off-load the 5000 containers using shore-based cranes in less than XX hours." For other ship types, such as industrial ships performing missions at sea, the mission requirements are more complex. The requirements for a fisheries research vessel, for example, might specify the ability to catch fish using several different techniques, radiated noise limitations, required sonar performance, and several different aspects of maneuvering and seakeeping performance, such as low-speed stationkeeping and the ability to maintain a specified track over the sea floor in the face of cross currents, winds, and seas. *See* MERCHANT SHIP; OCEANOGRAPHIC VESSELS.

Generally, constraints as well as requirements are imposed on the designer by the prospective owner. Cost, both to design and build the ship and to operate it, is usually constrained. The two primary elements of operating cost are crew and fuel, so there is nearly always pressure on the designer to reduce crew size and fuel consumption. Physical constraints may also be imposed on the design related to construction, operational, or maintenance requirements. Weight or dimensional constraints may be imposed if the ship is to be built or maintained in a specific dry dock. Pier or harbor limitations may also impose dimensional constraints. Ship length may be limited by the requirement to tie up to a certain pier. Ship air draft (vertical distance from the water surface to the highest point on the ship) may be limited by the need to pass under bridges of a certain height. Ship navigational draft (vertical distance from the water surface to the lowest point on the ship) may be limited by the depth of a dredged channel in a particular harbor.

In addition to unique mission requirements and constraints, every ship must satisfy certain physical principles. The fundamental principles are that (1) the ship hull and superstructure must have adequate storage space, and (2) the ship must float at an acceptable waterline (draft neither too great nor too small) when it is fully loaded. A ship floats because the buoyant force acting vertically upward is equal to the total weight of the ship and its contents. Per Archimedes' principle, the buoyant force is equal to the weight of the water displaced by the hull and its appendages (propeller, rudder, and so on). This weight is termed the displacement. Another principle is that the ship must be statically stable; that is, when it is displaced from its equilibrium condition, it must tend to return to that condition. For example, when the ship is heeled to one side by a disturbing force such as a wind gust, it must tend to return to the vertical rather than continuing to roll and capsizing. The ship's hull must have sufficient strength to withstand the forces that will act upon it over a range of loading and sea conditions. The ship must possess sufficient propulsive power to achieve the desired speed even with a fouled bottom and in adverse sea conditions. In addition, it must generate sufficient electric power to satisfy the requirements of mission systems; ship machinery; heating, ventilation, and air conditioning (HVAC) systems; hotel; and other ship services. *See* ARCHIMEDES' PRINCIPLE; BUOYANCY; HYDROSTATICS.

Design approach. Ship design is an iterative process. The various synthesis and analysis tasks are repeated in increasing detail as the design is developed, as shown by the design spiral in **Fig. 1**. In the figure, each complete circuit around the center is intended to represent a complete set of the steps necessary to configure and assess a design concept. The steps are repeated in the second circuit, the third, and so on. Initially, assumptions are made as to the eventual size and configuration of the ship in order to get the process started. As the design proceeds, these initial assumptions are replaced by better data as more becomes known. The nature of the analysis work also changes. Initially, rough order-of-magnitude estimates are made using simplified procedures. Later, more detailed calculations are made using more complex procedures. The initial circuits around the spiral

are made quickly by one or two designers. Later in the process, many more persons are involved and the circuits are made more slowly. Figure 1 depicts an idealized design process. In fact, after the initial ship concept is developed, design work in the various design disciplines proceeds more or less in parallel, with periodic halts to integrate the results and assess overall progress.

In planning, managing, and describing the ship design process, a distinction is made between level I (total ship) design and level II (ship system) design. Level I design concerns the synthesis and analysis of total ship attributes such as hull form and general arrangement, and the assessment of total ship attributes such as weights and centers. Level II design concerns the synthesis and analysis of the major individual ship elements such as structure; propulsion machinery; electric power generation and distribution; ship control, navigation, and communication systems; auxiliary and mechanical systems, including piping and HVAC systems; and outfit and furnishings. Many outputs of level II design are inputs to level I design assessments. Examples of such outputs are data on weight, reliability, required crew size, cost, and risk.

Design integration is a major challenge. If the level II elements of a ship design are developed and optimized independently, the result will be a nonoptimum and disappointing level I total ship design. The elements of the design must be integrated so that they work well together. This will generally require that compromises be made in the level II system designs.

Design phases. The four traditional phases of ship design are concept, preliminary, contract, and detail design (Fig. 1). The phases correspond to milestones in the process used to develop and build a new ship; they permit periodic top-level reviews of the development status by the top management of both the design/build team and the owner (and enable the project to be halted if desired); and they reflect the fact that the design team changes size and composition during design development.

Concept design. The first design phase helps the owner and the designer to define the top-level requirements for the new design and to develop a ship concept design that reflects those requirements. Typically, the top-level design requirements are defined through an iterative dialog between the owner and the designer. The designer will develop numerous sketch designs, called ship feasibility studies, that explore parametric variations of the key requirements, which are major design drivers in terms of ship size, cost, and performance. Costs are estimated for each option. The prospective owner will review the sketch designs and their estimated costs and make tentative decisions to narrow the range of investigation. The designer will then develop a second set of design alternatives, representing a more focused set of options, and present them to the owner for review. By means of such an iterative dialog, the major requirements for the ship are defined, and a feasibility study, representing the agreed-to requirements, is

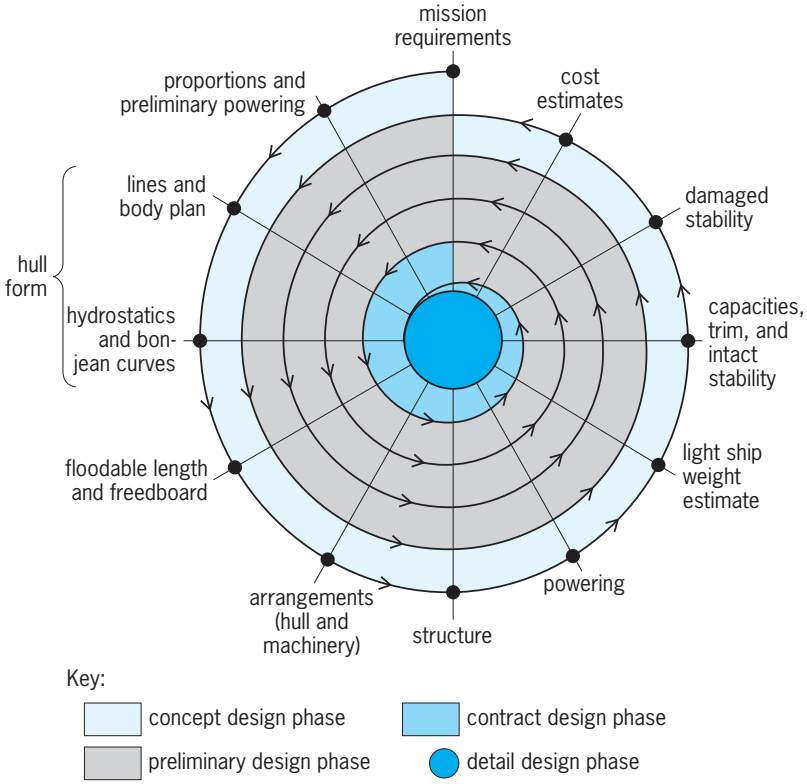

Fig. 1. Basic design spiral, showing the iterative ship design process.

developed. Subsequently, this brief study is further developed into a concept design.

Preliminary design. During the second design phase, the major ship characteristics which drive ship cost and performance are emphasized. For each characteristic, trade-off studies are performed in which a range of alternatives is studied. Issues typically addressed during this phase include hull shape (proportions and form coefficients), propulsion plant [prime movers: number and type(s), number of shafts], general arrangement options, topside configuration, structural configuration, hull material, and electric plant configuration. The initial strategy for building the ship is developed. At the conclusion of this phase, a budget-quality cost estimate is prepared. Major ship characteristics, such as the principal hull dimensions, installed propulsion power, and cargo capacity, are not expected to change upon completion of this phase. *See* PROPELLER (MARINE CRAFT).

Contract design. During this phase, the design is further developed and refined. Drawings and specifications are produced which form an integral part of the contract for ship detail design and construction. Hydrodynamic model tests are run, including resistance, propulsion, seakeeping, and maneuvering tests. Hull form and appendage shape definitions are completed. The primary hull structure is defined, including stiffener and plate scantlings, based on refined load estimates and finite element analyses. Distributive systems are defined in diagrammatic form. The equipment in ship operating and living spaces is selected and arranged. The ship weight and center-of-gravity estimate, which accounts for the weight

and location of each major item in the ship, is further refined. Intact and damage stability analyses for a range of ship loading conditions are refined. The ship specifications, a primary product of this phase, define quality standards for the hull and hull outfit as well as performance standards for each item of ship machinery and equipment. The specifications also define the tests and trials that must be performed by the shipbuilder as the ship nears completion. The specified test criteria must be satisfied for the ship to be considered acceptable. *See* FINITE ELEMENT METHOD.

Detail design. The fourth and final design phase is devoted to completing the ship design in every detail and to translating the completed design into easily understood work packages that can be used directly by the artisans who will build the ship without further translation or interpretation. The work packages are the fabrication and installation instructions to the ship fitters, welders, sheet metal workers, pipefitters, and electricians. In the detail design phase, especially in the development of work packages, great emphasis is given to production engineering, to make the ship construction process as speedy and efficient as possible. *See* ENGINEERING DESIGN; PRODUCTION ENGINEERING; SHIPBUILDING.

Level I (total ship) design elements. Level I design includes such topics as hull form and general arrangement, weights and centers, stability, hydrodynamic performance, crew size, cost, reliability, supportability, operability, and risk.

Hull form. Naval architects describe hull forms by their principal dimensions (length, L; beam, B; draft, T; and depth, D), proportions (ratios of length to beam, L/B; length to depth, L/D; and beam to draft, B/T), and by form coefficients, which characterize the hull shape. L is the hull length between perpendiculars at or near the ends of the ship and may also be the length on the waterline at which the ship is designed to float when fully loaded. This waterline is called the design waterline, or DWL. B is the hull beam, generally the beam at the DWL midship (at the mid-length of the ship). T, the ship draft, is measured amidship vertically down from the DWL to the lowest point on the hull. D, the hull depth, is also measured vertically at midship from the lowest point on the hull to the main or weather deck. Freeboard is the height of the hull above the water surface and is equal to the difference between D and T, or $D - T$.

The hull form for a new ship design must satisfy the ship's mission requirements as well as the physical principles mentioned above. The hull must provide adequate internal volume, buoyancy when floating at or near the DWL, and stability. In addition, the freeboard must be sufficient to provide adequate stability at large heel angles, and adequate reserve buoyancy in the event that the hull suffers damage and a portion of the hull is flooded. The freeboard must also be adequate to keep the weather deck reasonably dry in rough seas. Hydrodynamic performance is a major consideration in defining the hull form. Resistance and propulsion, seakeeping, and maneuvering behavior must all be considered. *See* SHIP POWERING, MANEUVERING, AND SEAKEEPING.

First among the considerations is resistance and propulsion. The hull proportions and form coefficients must enable the ship to be propelled efficiently at the desired top speed as well as at the specified economical cruising speed. These two speeds may be very different, as on a warship, or may be the same, as on a point-to-point cargo ship. The two most significant form coefficients from the standpoint of resistance and propulsion are the displacement-length ratio and the prismatic coefficient. The displacement-length ratio indicates the hull fatness or slenderness. It is the ratio of the hull's displacement volume divided by the hull length cubed. Traditionally, in English units, the displacement-length ratio is defined as the hull displacement in long tons of salt water divided by hull length in feet/100. In this form, the ratio varies from a low of about 10–15 for a racing shell used by college crews, up to about 500 for a harbor tug. The prismatic coefficient C_p indicates the bluntness or fineness of the underwater hull form and is defined as the displacement volume divided by the volume of a circumscribing prism with length L and cross section equal to the largest cross section of the ship's immersed form, at or near midlength. A very fine hull has a C_p value of about 0.55, while a very full hull such as that of an oil tanker might have a C_p of 0.90 or higher.

The speed-length ratio (SLR) is an important concept when discussing ship resistance and propulsion. SLR is defined as the ship speed in knots divided by the square root of the hull length L in feet. The two primary components of hull resistance to forward motion are frictional resistance and wavemaking resistance. Frictional resistance is due to the friction between the hull's wetted surface and the water passing over it. Wavemaking resistance is due to the waves generated by the moving hull; creating these waves requires energy input. Below an SLR of about 0.75, virtually all of the hull resistance is frictional resistance. As the SLR increases above 0.75, the importance of the wavemaking component or resistance steadily increases. At an SLR of about 1.0, the frictional and wavemaking components of resistance are about equal in importance. Above an SLR of 1.0, the wavemaking component dominates. At an SLR of about 1.30, a conventional displacement ship is moving at its "hull speed." At this speed, the ship's speed-power curve is very steep and it is very difficult to increase speed further, that is, a great deal of power is required to increase speed even a small amount.

Frictional resistance is a function of the hull's wetted surface and hull smoothness. The wavemaking portion of resistance is affected by the hull shape, and its importance thus increases as the SLR increases above about 0.75. For SLR values in the range of 0.8–1.1, the optimum C_p is low, in the 0.55–0.57 range. For SLR ratios approaching 1.30 and higher, the optimum C_p increases to about 0.64. At such high speeds, the effect of displacement-length

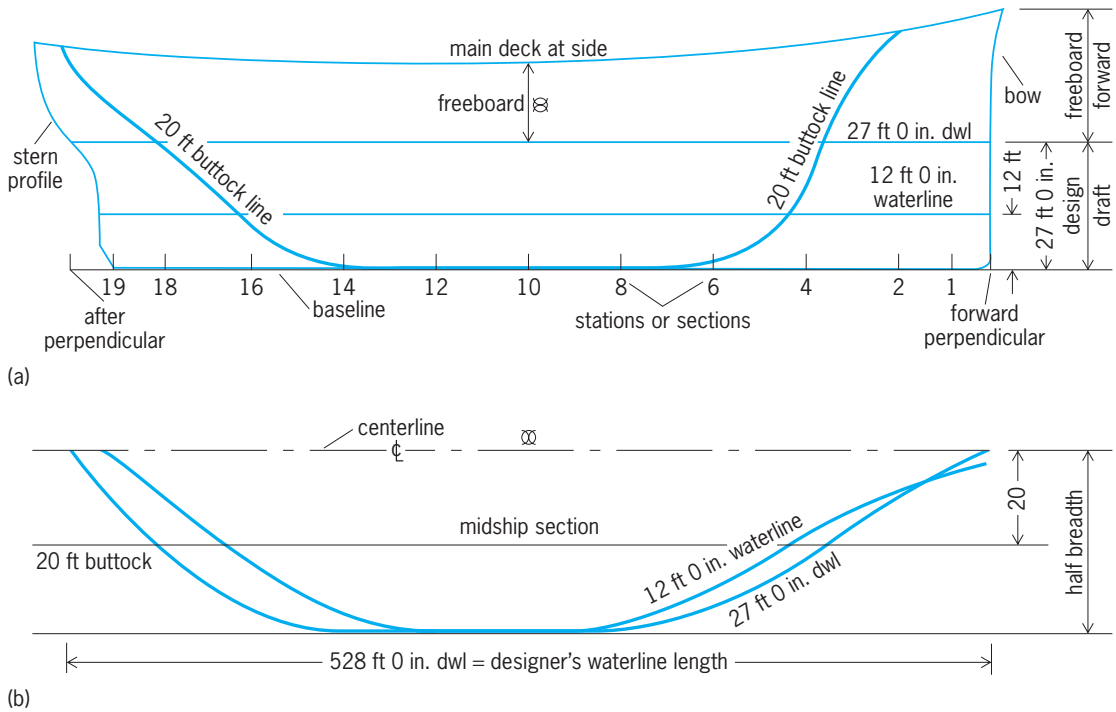

Fig. 2. Geometry of a ship. (*a*) Profile view. (*b*) Half-breadth plan. These views, along with the transverse view given in Fig. 3, make up the lines drawing, or lines of the vessel. 1 ft = 0.3 m; 1 in. = 2.5 cm.

ratio is dominant and low values are desired. *See* FRICTION.

For the naval architect, length is the most important hull dimension. It has a greater effect on construction cost than any other dimension. It also has a powerful effect on hull resistance characteristics through its effect on both the SLR and the displacement-length ratio. Trade-off studies are performed during the design process to find the optimum balance between reduced length to save construction cost and increased length to improve hull resistance characteristics. Hull beam is the second most costly dimension in terms of construction cost, and has the most important effect on transverse static stability. Thus, selection of the optimum beam is also a subject of study during design development. Hull depth is important because of its effect on freeboard and on the strength of the ship's hull girder.

The ship's hull form is delineated in a so-called lines drawing. This drawing depicts the hull form in three views: the elevation or profile plan, looking at the hull from the side; the half-breadth plan, looking down from above; and the body plan, looking at the hull from the bow or stern (**Figs. 2** and **3**). Each view depicts the lines of intersection of the hull with a series of parallel planes. In the profile view, these planes are parallel to the hull centerplane, the plane of symmetry of the hull, and the lines of intersection are called buttocks. In the half-breadth plan, the planes are parallel to the design waterplane, and the lines of intersection are called waterlines. In the body plan, the planes are parallel to the transverse midship plane of the hull and are perpendicular to the other two sets of planes. The lines of intersection are called sections.

Ordinarily, only one side of the hull is depicted in the lines drawing since most hulls are symmetrical about their longitudinal centerplane. Using the three projections of the lines drawing in conjunction with one another, it is possible to determine the position in space of any point on the hull surface. The lines drawing is used to perform many of the analyses during ship design development. Also, many other ship drawings are based on the hull geometry defined by the lines drawing.

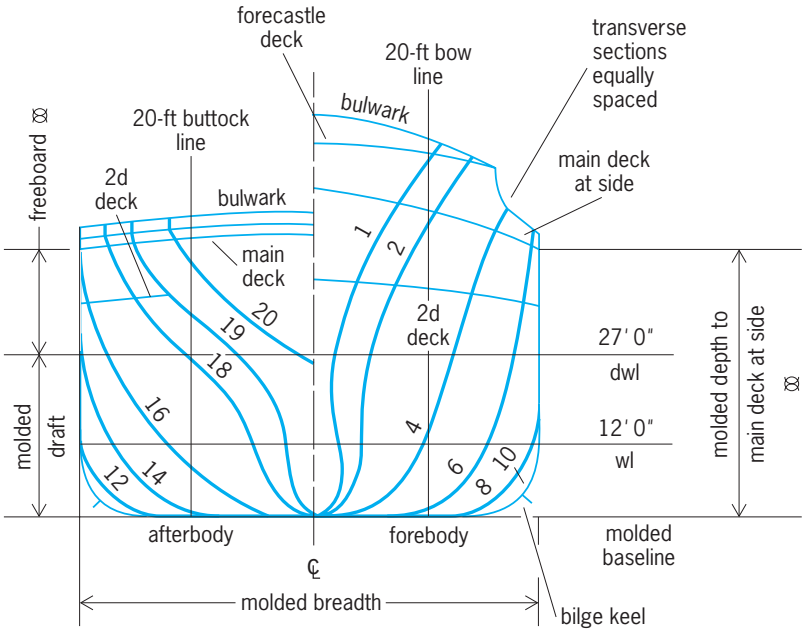

Fig. 3. Transverse view, or body plan, of ship. 1 ft = 0.3 m; 1 in. = 2.5 cm.

Subdivision. The internal hull volume is subdivided into compartments by horizontal surfaces called decks, and vertical surfaces called bulkheads. Bulkheads are either transverse or longitudinal, depending on whether they are perpendicular to or parallel to the ship's centerplane.

Decks permit the internal space to be efficiently used for living and working spaces. Decks may be omitted from main machinery spaces or cargo holds where large heights are required. Bulkheads serve several purposes. They add to the strength of the hull, restrict fire and flooding, support the decks, and separate compartments with dissimilar functions.

Longitudinal bulkheads are sometimes used to add structural strength or to bound vital spaces such as shaft alleys. Generally they are avoided, however, since in the event of hull damage a watertight longitudinal bulkhead restricts flooding to one side of the ship. This can result in a large heel angle after damage and may cause the ship to capsize.

The primary function of transverse bulkheads is to restrict the longitudinal extent of flooding in the case of hull damage. Transverse bulkheads are watertight up to the bulkhead deck. Ships are generally designed so that the flooding, resulting from a specified extent of damage, will not immerse the hull so deeply that the bulkheads restricting the flooding are immersed. If this occurs, seawater can flow over the tops of the bounding bulkheads into adjacent undamaged compartments, and the resulting progressive flooding can cause ship loss, as occurred with the *Titanic*.

The spacing of transverse bulkheads is a critical design decision. If the spacing is too great, the ship will not be able to survive flooding beyond the capability of the ship's pumps to dewater the flooded space. Too close a spacing may interfere with cargo handling and stowage, passenger accommodations, machinery arrangements, or other ship functions. This could cause the ship to be unsuccessful commercially. Naval architects develop floodable length curves to aid in locating transverse bulkheads (**Fig. 4**). The floodable length at any point along the hull is the length of flooding, centered at that point, which will cause the ship to sink down to a waterline 3 in. (7.6 cm) below the bulkhead deck. Compartment permeabilities are reflected in the floodable length curve, as noted in the figure. The permeability of a compartment μ is the volume of flooding water that can enter that compartment, expressed as a percentage of the compartment volume. An empty cargo hold or ballast tank would have a permeability of close to 100%, while a full ballast tank would have a permeability of zero. Most ship spaces other than tanks have permeabilities in the 60–80% range. There are international regulations that specify the subdivision requirements for various ship-type categories, such as passenger ships, oil tankers, and dry cargo vessels. The passenger ship regulations are driven by passenger safety concerns, while the oil tanker regulations are driven by pollution concerns.

General arrangement. Naval architects define the arrangement of external and internal ship features in a drawing termed the "general arrangement." This drawing typically shows the exterior of the ship in several views. For the ship's interior, the inboard profile (a profile view along the ship's centerplane) and plan views of each deck in the hull and superstructure are shown. The general arrangement shows all internal tanks and compartments, including passageways and vertical trunks provided for access. The development of the ship's general arrangement during the design process is one of the principal ship design tasks, and the drawing is used to help guide and control the entire design process. During arrangement development, top priority is given to the ship's mission requirements and the associated payload. After the hull form and principal subdivisions have been defined, blocks of space are assigned to accommodate the various ship functions. Alternatives are often considered and evaluated, such as propulsion machinery midships versus aft. Later, the principal access routes within the ship are outlined, and individual compartments are sized and located. Additional detail is developed as the design proceeds.

Capacities. Available cargo hold and tank capacities (fuel, fresh water, ballast) are calculated from the hull form and general arrangement. The results are

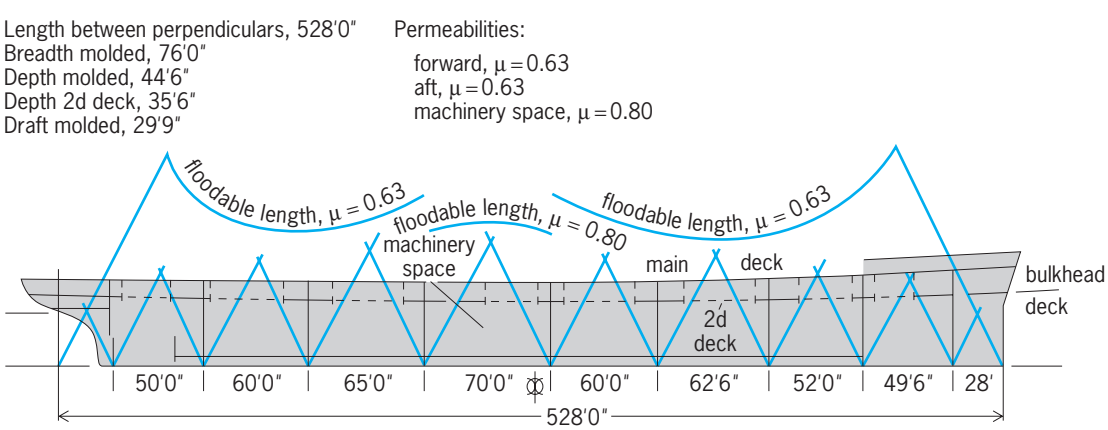

Fig. 4. Subdivision diagram and floodable length curve. 1 ft = 0.3 m; 1 in. = 2.5 cm.

compared with the required capacities, and adjustments are made if necessary.

Weights and centers. Careful analysis of ship weight and center of gravity is essential to a successful ship design. The fully loaded weight of a ship has two elements: the light ship weight, the weight of the ship itself; and the loads, called the deadweight. The light ship weight includes the weight of the hull structure; propulsion machinery; electric plant; ship navigation, control, and communications systems; piping and ventilation systems; cargo-handling equipment; outfit and furnishings; and all other equipment and fittings necessary to make the ship complete. The deadweight consists of all the variable loads that can be moved onto or off the ship, including cargo, fuel, crew and their personal effects, stores, and fresh water. The light ship weight plus the deadweight equals the full load weight or displacement, the weight which must be supported by the buoyancy of the hull. Early in the design process, weights and centers are estimated using similar "parent" ships or historical data, which are analyzed to create estimating relationships or algorithms. Margins are applied to the estimates to allow for unknowns and estimating errors. As the design proceeds, more and more of the total weight can be calculated based on equipment lists and drawings, such as the structural drawings. As the quality of the estimate improves, the margins applied can be reduced. The total weight and center-of-gravity position are calculated for a range of ship operational loading conditions, varying from light to full load. The results are used to calculate trim and list when the ship is at rest pierside and also to analyze stability and seakeeping characteristics. *See* CENTER OF GRAVITY.

Stability. A ship must be stable under all normal conditions of loading and performance at sea. This means that when it is inclined from its at-rest position by a disturbing force, such as a wind gust, it should return to the at-rest position when the force is removed. Stability is addressed in both the longitudinal and transverse directions; for most surface ships, transverse stability is a greater concern. Stability is an issue both when the hull is intact and after the hull has been damaged and partially flooded. Thus both intact and damage stability are addressed by the naval architect during the design of a ship.

Figure 5 shows the cross section of a ship that has been inclined from its normal upright position by a disturbing force. The weight of the ship W acts down at the ship's center of gravity, point G in the figure. When the ship is upright, the buoyant force, equal to W, acts up at the center of buoyancy, point B. After the ship is inclined, the center of buoyancy shifts toward the low side to point B'. The buoyant force is now offset from the weight force by the distance GZ. A couple is formed acting to right the ship. The righting moment is $W \times GZ$. The distance GZ is termed the righting arm. For a given center-of-gravity position G, the righting arm is a function of the ship's hull form. GZ can be computed for a range of heel angles and plotted versus the heel angle to yield the ship's righting arm curve. Generally, GZ will increase

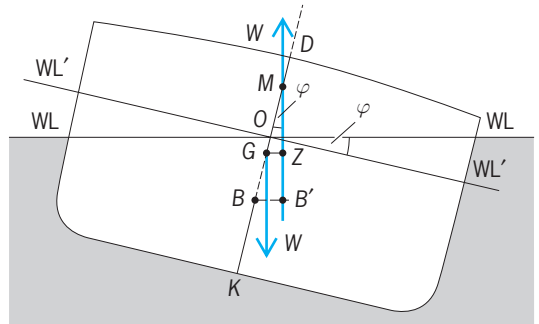

Fig. 5. Diagram of inclined ship, showing points, lines, and angles that are important in determining the ship's stability. W′L′ = inclined water line, WL = position of waterline on upright ship. Other symbols are explained in text.

with heel angle until a maximum is reached and then decrease as heel is further increased until it reaches zero. At this point the ship will capsize.

The range of stability refers to the range of heel angles over which the righting arm is positive. Increasing hull beam will increase the height of the righting arm curve. Increasing freeboard will increase both GZ (at large heel angles) and the range of stability. For a given hull form, GZ is a function of the height of point G, that is, the vertical position of the ship's center of gravity. As G is raised, GZ decreases. If GZ is negative, that is, if Z is to the left of G in Fig. 5, then the ship is unstable—the moment is tending to capsize the ship rather than return it to the upright. Intact stability is evaluated for several assumed hazards. These include a beam wind (perpendicular to the direction that the ship is sailing) combined with ship rolling due to beam seas, and the case of all the ship's crew and passengers crowding to one side of the ship.

The point M in Fig. 5 is called the metacenter. It is the point about which the center of buoyancy is rotating. For small heel angles, M lies on the ship's centerplane at the point where the line of action of the buoyant force intersects it. This fact permits naval architects to analyze stability at small heel angles, called initial stability, using simplified methods. Initial stability is checked very early in the design process. Later, large-angle intact stability and stability after damage are analyzed. For initial stability, the distance GM is called the transverse metacentric height. If M is above G, GM is positive and the ship is stable. If M is below G, then GM is negative and the ship is unstable. The height of M above the ship's keel, distance KM in the figure, is a function of the ship's hull form. KM has two parts: the distance KB, the height of the ship's center of buoyancy above the ship's keel, plus the distance BM, the height of the metacenter above the ship's initial center of buoyancy. Distance BM is called the metacentric radius. BM is equal to I/V, where I is the transverse moment of inertia of the ship's waterplane at the water surface and V is the immersed hull volume. Both KB and BM are functions of the ship's hull form and can be readily computed in early-stage design (or estimated using simplified relationships). I can be expressed

as a constant $\times$ LB^3, and V can be expressed as a constant $\times$ $L \times B \times T$ (B is the hull beam in these expressions). Thus I/V varies as beam squared. This is why a small beam increase has such a strong effect on GM and hence initial stability. Note in Fig. 5 that the righting arm GZ is equal to GM times the sine of the heel angle ϕ for small heel angles (when M is on the ship's centerplane). *See* MOMENT OF INERTIA.

A ship's natural roll period varies inversely with the square root of its transverse GM. Increasing GM increases the ship's stiffness and improves its initial stability but reduces its roll period. Shorter roll periods mean greater lateral accelerations and increased discomfort for shipboard personnel and passengers. Thus the naval achitect must find the best balance between stability and seakindliness when establishing GM in a new design. Generally, a GM equal to about 6% of the maximum beam on the waterline in the full load condition is appropriate for commercial vessels. Passenger ships and warships have higher GM values in order to meet more stringent damage stability requirements. Damage stability is analyzed after one or more adjacent compartments have been flooded. The presence of a beam wind combined with rolling is assumed. Naval architects use the terms one-, two-, or three-compartment ships, referring to the number of adjacent compartments that can be flooded while still satisfying the established stability criteria.

Freeboard and load lines. A load line is a mark on the side of a ship's hull which indicates the deepest draft to which it can be legally loaded in fresh or salt water in various operating regions and seasons of the year. International conventions have established regulations for establishing load lines. The regulations are based on considerations of reserve buoyancy, hull structural strength, and the provisions for making the hull watertight in extreme conditions, such as strong, tight hatch covers. Reserve buoyancy is the volume (potential buoyancy) of the watertight hull above the load waterline. Credit is also taken for weathertight superstructure. The regulations establish the minimum acceptable freeboards for the various operating conditions. The naval architect must ensure that the ship being designed will satisfy the load line regulations when carrying the desired payload. For passenger ships carrying more than 12 passengers, in addition to the above regulations, there is a limit termed the subdivision draft. The subdivision draft is a function of the spacing of internal watertight transverse bulkheads and hence the number of compartments that can be flooded in a damaged condition before reserve buoyancy and stability disappear. If the subdivision draft is less than the limiting draft per the load line regulations, the subdivision draft governs.

Tonnage. Tonnage measurement is the result of continuing efforts by government authorities to establish a uniform means of measuring ships' relative size and carrying capacities for the purposes of taxation. Since about 1860, tonnage has generally been a measure of internal volume or cubic capacity expressed in units (tons) of 100 ft^3 (3 m^3). Gross tonnage refers to the entire internal volume, less certain exempted spaces. Net tonnage is the tonnage remaining after the non-earning spaces, such as machinery and crew living spaces, have been deducted from the gross tonnage. Net tonnage is intended to be a measure of a ship's earning ability, and is the basis on which charges, such as canal transits and port entry, are determined. Until recently, different countries and canal authorities have had their own tonnage measurement regulations. In 1969, the International Convention on the Tonnage Measurement of Ships established a simplified method of determining tonnage. The Convention replaced the myriad of special tonnage systems previously used.

Hydrodynamic performance. Satisfactory hydrodynamic performance is critical to a new ship's success, both in terms of mission effectiveness and of economic success. Thus the evaluation of hydrodynamic performance is an important element of ship design. The three primary elements of hydrodynamic performance are calm water resistance and propulsion, maneuvering, and seakeeping. The latter includes ship motions as well as the effects of rough water on resistance and propulsion and the structural loads transmitted to the hull. Hull size and shape have strong effects on resistance and ship motions. These effects are primary considerations in early-stage design when hull size and gross shape are being determined. Propulsion, maneuvering, and other aspects of seakeeping get greater attention later in the design process. To date, naval architects have been unable to satisfactorily predict hydrodynamic performance by analytical means. The complexities of the phenomena involved have simply been too great. Thus, in early-stage design, rather crude analytical predictions of hydrodynamic performance have been made. Extensive model testing has traditionally been used in the middle phases of design to evaluate attractive design alternatives and make the best possible performance predictions. Advances in computational fluid dynamics (CFD), coupled with reductions in the cost of high-speed computing, are rapidly changing this picture. Today, model testing is generally used later in design only to confirm sophisticated analytical predictions of hydrodynamic performance; in the future, physical model tests may be phased out entirely. *See* COMPUTATIONAL FLUID DYNAMICS; HYDRODYNAMICS.

Propulsion. The shaft power (SHP) required to propel a ship at a specific speed in calm water is equal to the power required to tow the ship at that speed (EHP) divided by the quasi-propulsive coefficient, P. C. EHP equals the tow rope pull (total ship resistance) times the velocity converted to power. P. C. equals the open-water propeller efficiency times a factor called the hull efficiency, to account for the effect of the propeller on the ship resistance and the effect of the hull and appendages on the inflow velocity to the propeller. Thus, to estimate the SHP at a specific speed, the following must be

estimated: total resistance, open-water propeller efficiency, and hull efficiency. Total resistance consists of bare hull resistance plus appendage resistance plus still air drag (negligible for most ships). Bare hull resistance consists of frictional resistance plus residual resistance, which is the sum of wavemaking resistance plus separation and eddymaking resistance. Appendage drag and hull efficiency can be estimated from self-propelled model test results for similar ship designs. Frictional resistance can be calculated using rather simple relationships based on turbulent flow, knowing the immersed wetted area of the hull. Still air drag is significant only for high-speed ships, and can be estimated based on the projected frontal area of the ship topsides and superstructure and an assumed drag coefficient. Open-water propeller efficiency can be estimated using propeller model test data or lifting line theory. This leaves residual resistance, which can be estimated from the model test results for a similar hull form, the results of model tests on a standard series of hull forms, the results of regression analyses of model tests of a large number of hull forms, analytically computed wavemaking resistance multiplied by a correction factor based on experience, or on model tests of the new hull form. In early-stage design, speed-power estimates are based on combinations of analytical and empirical methods, using previous model tests of similar hull forms. Later in the design process, model tests are generally performed for the new design. These tests include bare hull and appended hull resistance and self-propulsion tests. *See* TOWING TANK.

Seakeeping. Ship motions are the direct effect of the seaway on a ship. In higher sea states, these motions may result in bottom slamming (when the ship's keel emerges from the water and then violently reenters) and deck wetness. Slamming and deck wetness lead to voluntary speed reductions to reduce their frequency and severity. Ship motions themselves result in accelerations which cause loss of crew effectiveness and passenger seasickness, and may even damage equipment. Excessive motions and accelerations may prevent certain missions from being carried out, such as helicopter or small boat operations.

The complex motions of a ship at sea are analyzed by decomposing them into six components, three linear and three rotational, with reference to orthogonal axes through the ship's center of gravity. The linear components are surge (fore and aft motion), sway (side-to-side motion), and heave (vertical motion). The rotational components are roll (rotation about the longitudinal axis), pitch (rotation about the transverse axis), and yaw (rotation about the vertical axis).

Three of the motion components—pitch, heave, and roll—have the dominant effects on seakeeping performance. Pitch and heave are strongly coupled; their effects are most severe in head seas (waves coming from directly ahead). They cause vertical accelerations, deck wetness, and bottom slamming. Once the overall hull size and shape have been established, little can be done to ameliorate pitch and heave, since the forces necessary to do this are very large for conventional hull forms.

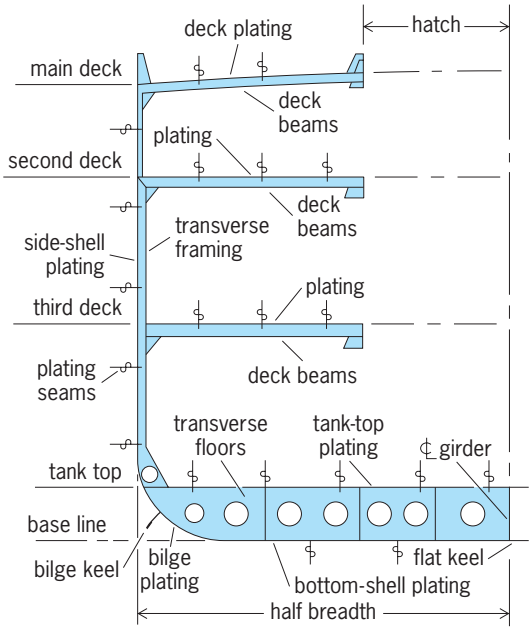

Fig. 6. Schematic midship section showing structural arrangement of plating and transverse framing within a steel merchant ship.

Roll motion, on the other hand, causes lateral accelerations and is caused, and can be reduced, by relatively small forces. For this reason, several practical means of reducing roll motion have been developed and employed. Bilge keels are long projecting structures, fitted at the turn of the bilge (**Fig. 6**), which are inexpensive and quite effective at reducing roll amplitudes over the entire speed range. Active or passive antiroll tanks can be effective at most speeds and headings, but they have significant space and weight penalties, are noisy, and can increase roll at some headings and ship loading conditions. Active antiroll fins are very effective at mid-to-high ship speeds but are costly and ineffective at low speeds.

Seakeeping behavior is assessed analytically in early-stage design and may also be evaluated by means of model tests later in design. Seakeeping assessments are an important part of the hull form trade-off process. Roll stabilization studies are often performed.

Maneuvering. The third primary element of ship hydrodynamic performance is maneuverability. Important aspects of maneuverability are directional stability, turning ability, and responsiveness to course changes, all at significant forward speeds, and maneuverability at low speeds, at pierside, for example. If a ship is directionally stable, it will tend to return to its original course after being forced off course by a minor disturbance such as a wind gust. Turning ability is expressed in terms of the turning diameter at various speeds and rudder angles. Maneuverability is influenced more by the ship's appendages, such as

the rudder and skeg, than it is by the basic hull proportions, although these do have a significant influence. For this reason, in sizing and shaping the hull in early-stage design, speed-power and seakeeping considerations generally play a greater role than maneuverability. Initially, appendages are often sized using empirical relationships or by ratiocination from similar successful designs. Later, analytical models are used to prepare refined performance predictions. Finally, physical model tests are generally performed to confirm the predicted performance. Low-speed maneuverability is often enhanced by installing thrusting devices of various sorts at the bow or stern which can be used to create strong side forces at low forward speeds.

Crew size. The total life cycle cost of a ship is the sum of its design and build (D&B) cost plus its operating cost over the ship's life. Crew cost is invariably the largest fraction of the total life cycle cost, greater even than the D&B cost. Thus ship owners and naval architects give great attention to minimizing crew size. Three scenarios must be considered in establishing crew size: normal ship operations, ship maintenance performed at sea, and damage control in the event of fire or flooding. Federal safety regulations and union rules are important factors for commercial ships. Automation of functions traditionally performed by humans is a technique used to reduce crew size. However, such automation is costly and adds maintenance and crew training requirements which must be considered. Thus crew size versus automation trade-off studies are a typical element of the design process.

Reliability. A ship and its systems must be highly reliable from the standpoints of both safety and good business. During design, reliability analyses are performed for critical ship systems in order to ensure that sufficient reliability is designed in. Techniques for improving reliability include paying more for more reliable system components or increasing redundancy to permit continued operation in the event of a failure.

Supportability. A ship must be properly supported to be capable of sustained, reliable operations. Properly trained crew members must be available when required to fill voids. The ship must receive timely replenishment of consumables, other stores, fresh water, fuel, other liquids, and repair parts. A routine for regular ship maintenance must be established and provisions made to accomplish unscheduled repairs when needed. Supportability considerations must be an integral part of the design process. Decisions must be made on when, where, and how frequently the ship will be resupplied. A ship maintenance plan must be developed which addresses how the ship will be maintained and by whom (for example, maintenance at-sea versus ashore, in-place repairs versus unit replacement, and maintenance by the ship's crew versus shore support personnel). The ship's storerooms and workshops must reflect the maintenance and supply support plans. Adequate internal accesses must be provided to permit stores handling and necessary maintenance actions to be performed. Crew training requirements must be defined and training materials developed for unique ship systems.

Cost. The prospective owner of a new ship will nearly always have cost limits which must not be exceeded. Thus throughout the design process, from the development of initial sketchy concepts through detail design, trade-offs are made in which projected costs are a major consideration. There are two major elements of the life cycle cost of a ship: the acquisition (or design and build) cost, and operating costs, that is, the costs of items such as fuel, crew salaries, supplies, maintenance, and insurance. D&B cost is further split into two categories: recurring and nonrecurring costs. Nonrecurring costs are the costs of design and development associated with building the first ship of a new class, while recurring costs are incurred for each ship of a multiship class. There is often a trade-off between acquisition and operating cost; for example, operating cost may be reduced by spending more initially. A more costly but more efficient propulsion plant might reduce fuel consumption and operating cost. The ship designer must know the owner's D&B and operating cost thresholds, as well as the owner's preference as to the relative weight to be given to D&B versus operating cost in performing design trade-offs.

Risk. Risk is an issue which must be addressed in each new ship design. Technical risk relates to performance, both performance at delivery and over the long term. Will the completed ship meet the owner's performance requirements? Program risk includes cost and schedule as major elements. Will the ship be delivered on time and at a cost at or below the owner's threshold? For the ship designer, there is a generally a trade-off between innovation and risk. Innovative design solutions may have the potential of improving performance or reducing cost, but they generally increase technical or program risks because they have not been fully developed or have not been proven previously in service. Generally ship owners are extremely conservative and reluctant to be the first to try new systems, equipment, or design approaches. A bold owner can gain a significant market advantage with an innovative design but runs the risk of disappointment and financial losses. The ship designer must be sure to understand the owner's willingness to accept risk before starting a design project. Methods have been developed to formally assess and manage risks during the design process so as to reduce the chances of rude surprises after ship completion. *See* RISK ASSESSMENT AND MANAGEMENT.

Level II (ship system) design elements. The total ship consists of major level II elements or systems which can be classified as follows: structure, propulsion plant, distributive (electric, piping, and HVAC) and mechanical systems, and outfit and furnishings. The ship design process includes the design of each of these level II elements as well as their functional and physical integration into a balanced whole.

For information on propulsion plant design *See* MARINE ENGINE; MARINE ENGINEERING; MARINE MACHINERY.

Structure. The ship's hull structure can be thought of as a hollow box beam. The main deck and the shell plating on the exterior of the hull comprise the outer surface of the beam, and the decks and bulkheads within the hull subdivide and strengthen the beam. The hull structure must withstand the overall loads acting on the hull, which tend to bend and twist the beam, as well as the local loads acting at any point. The overall hull loads result from differences between the local weight and buoyancy along the length of the hull in still water, and loads resulting from operation in rough seas (wave loads). Local loads result from sources such as water pressure, breaking seas, the weight of cargo, and loads due to components of the ship itself, such as heavy machinery, masts, and superstructure.

The hull of a typical steel merchant ship (Fig. 6) is a complex structure. The side shell and decks are supported locally by framing members called frames and beams, respectively. The bottom is supported by deep beams called floors. The purpose of the framing members is to maintain the adjacent plate membranes in their designed contours and positions relative to one another, as the plates are subjected to internal, external, and in-plane loads. The local framing members are, in turn, supported by more substantial structure. An important structural design decision is whether the deck and side shell frames should run longitudinally or transversely. Cost, weight, cargo handling, and maintenance considerations all enter into the decision. Stiffener spacing is an important variable. Wider stiffener spacings add weight but reduce cost by reducing the number of structural parts and the amount of welding required. Generally, ship hulls are built of steel, but other materials are sometimes used. In terms of steel grade, higher grades of steel are more costly to purchase and fabricate but may save weight, which is of great importance in some designs. *See* LOADS, TRANSVERSE; STRESS AND STRAIN; STRUCTURAL DEFLECTIONS.

The loads acting on a ship's hull are complex, and some cannot be accurately determined. Thus the structural designer must be guided by experience as well as theory. Classification societies (the American Bureau of Shipping in the United States) have developed rules for the structural design of commercial ships based on a combination of theory and experience. These rules can be used to define the structural dimensions used for a new design, as well as to gain an understanding of preferred construction practices. *See* STRUCTURAL DESIGN.

Distributive systems. The distributive systems in a ship are the electrical; piping; and heating, ventilation, and air conditioning (HVAC) systems. The design process for these systems consists of (1) analyzing the demands or loads which the system must support, (2) postulating one or more system concepts, as defined by a 1-line block diagram that depicts the system elements and the connections between elements, (3) sizing the major equipment elements of the system such as the electric generators, pumps, compressors, fans, and so on, (4) sizing the connections between system elements, such as the electric power cables, pipes, and vent ducts, and (5) developing the physical arrangement of the system within the ship. Historically, this was done using two-dimensional drawings, but increasingly, three-dimensional computer-aided design (CAD) models are employed. As with most other aspects of ship design, the process is an iterative one. For example, the required size of a pump will be dependent on the details of the input and output pipe sizes and runs; these are not defined until quite late in the design process. Thus, estimates are made of the effects of these details to enable an estimate of pump size to be made earlier in the process. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

Mechanical systems. A ship's mechanical systems include handling systems for anchors, boats, and cargo. For each such system, requirements must be defined, system concepts postulated, and analyses performed to size system components. Major equipment items are generally selected from vendor options.

Outfit and furnishings. Outfit and furnishings include items such as deck coverings, insulation, joiner bulkheads, ceilings, doors, lighting, shelving, port lights, and compartment furnishings. The ship's living and working spaces must be arranged and furnishings selected.

Warship design. Warship design is necessarily more complex than most commercial ship design. Warships carry and support sophisticated combat systems which include weapons systems and command, control, and communications systems. The design of these systems is a tremendous challenge. *See* NAVAL ARMAMENT.

Survivability is a key consideration in warship design. Low signatures, decoys, and active self-defense systems are employed to avoid being hit by an enemy mine or projectile. Armor and other passive protection features, as well as extensive subdivision and a high standard of damage stability, are built into the ship to contain damage and prevent ship loss if the ship is hit. Damage control features are provided to fight fires and control flooding after a hit, with the goal of restoring some mobility and combat capability.

The top speed of most warships is significantly higher than the speeds of most commercial ships. On the other hand, warships operate at top speed a small fraction of the time. Most of their at-sea time is spent at much lower transit or cruising speeds. Thus, unlike most merchant ships, warship propulsion plants must operate at two widely different power levels; this has a significant effect on the plant design. Weight must be controlled to achieve high speeds at practical power levels. Thus warship designs typically incorporate lighter-weight components, systems, and structure which are more costly than their commercial equivalents since they employ more expensive materials and are more difficult to

fabricate. One way to reduce structural weight, for example, is to reduce stiffener spacings, which increases the number of structural elements and hence fabrication costs.

Warship crews are typically much larger than merchant ship crews, and many naval ship crew members have little experience. These two factors have many ramifications for warship design, including more extensive living facilities, more trash and waste to handle while satisfying antipollution regulations, and greater challenges for crew training to properly operate and maintain the ship.

U.S. Navy warships are designed for long periods at sea. Thus they have large storerooms for supplies and spare parts, as well as extensive capabilities for self-maintenance. In addition, the ships are capable of being replenished with fuel, ordnance, and supplies while under way at sea. This is done by helicopters shuttling to and from the supply ship or by high-line transfers while connected to the supply ship, both ships proceeding on parallel courses. *See* NAVAL SURFACE SHIP.

Peter A. Gale

Bibliography. J. B. Caldwell and G. Ward (eds.), *Practical Design of Ships*, 2 vols., 1992; E. V. Lewis (ed.), *Principles of Naval Architecture*, 3 vols., 1988; H. Schneekluth and V. Bertram, *Ship Design for Efficiency and Economy*, 2d ed., 1998; R. Taggart (ed.), *Ship Design and Construction*, 3d ed., Society of Naval Architects and Marine Engineers, 1980.

Ship nuclear propulsion

Nuclear reactors for shipboard propulsion can be of any type used for the production of useful heat. Nuclear power is particularly suitable for vessels which need to be at sea for long periods without refueling or for powerful submarine propulsion. Only the pressurized water reactor (PWR) and the liquid metal reactor (LMR) have actually been applied to operating vessels. The pressurized water reactor has been most widely applied since it uses a readily available coolant and has a relatively simple cycle and control system and a large industrial and technical base. The supposed advantages of a liquid metal reactor (compactness, fast response, and higher propulsion plant efficiency) have not been proven in application, and liquid metal reactors are not now in marine service. For basic information on fission reactors *See* NUCLEAR POWER; NUCLEAR REACTOR; REACTOR PHYSICS.

In all the shipboard nuclear power plants that have been built, energy conversion is based on the steam turbine cycle, and that portion of the plant is more or less conventional (see **illus.**). There are two types in use: a steam turbine geared to a fixed-pitch propeller (called a geared turbine), and a steam turbine generator whose output drives an electric motor connected to a propeller (called a turboelectric unit). Any energy conversion process that converts heat into mechanical energy could be used to propel a ship. For example, a closed-cycle helium gas turbine has been studied, but none has been built for ship propulsion. *See* MARINE ENGINE; MARINE MACHINERY; PROPELLER (MARINE CRAFT); SHIP POWERING, MANEUVERING, AND SEAKEEPING.

Civil applications. Three early civilian applications, the *Savannah* (United States), the *Otto Hahn* (Germany), and the *Mutsu* (Japan), were technology demonstrations. They did establish that ships using nuclear power could not compete economically with ships using cheap fossil fuel for cargo/passenger applications on the oceans of the world.

In contrast, the use of nuclear propulsion has proven successful in the unique environment of the Russian Arctic. The long time between refueling and the energy required for ice breaking made this application comparatively practical. The icebreaker *Lenin* remained in service for 30 years from 1957, though a major accident is believed to have occurred in one or more of its reactors, which led to the replacement of the three reactors with two others in 1967–1970. The operational use of the *Lenin* led to larger icebreakers, the *Arktika* class of ships, the first of which was launched in 1975. These vessels have two pressurized water reactors and are used in deep Arctic waters to transport cargo (typically ores) and as passenger excursion vessels. (The *Arktika* was the first surface vessel to reach the North Pole.) *See* ICEBREAKER.

For use in shallow water such as estuaries and rivers, several shallow-draft *Taymyr*-class icebreakers were built in Finland and fitted with a single pressurized water reactor in Russia. Eight or nine icebreakers and one nuclear-powered container ship have operated. The nuclear-powered container ship *Sevmorput*, which began operation in 1988, had one pressurized water reactor; it was designed to meet International Maritime Organization (IMO) standards and has transported cargo to the Russian Far East and Cuba.

Military applications. Work on naval nuclear propulsion started in the late 1940s, and the first test reactor, a United States pressurized water reactor, started up in 1953. The first nuclear-powered submarine, the *Nautilus* of the United States, went to sea in 1955. The first Soviet nuclear-powered submarine went to sea 3 years later. *See* SUBMARINE.

Most naval reactors have been pressurized water reactors, which differ from commercial pressurized water reactors producing electricity in that they must have high power density and a small volume, and therefore run on highly enriched uranium (more than 20% uranium-235). They also have a long reactor core life, so that refueling is not needed for 10 years or more (depending on the amount of high-speed operation). This lifetime is extended by the relatively high enrichment of the uranium and by the incorporation of burnable poisons in the reactor assembly, which are progressively used up as the fuel is depleted and fission product nuclear poisons accumulate. By careful design these two effects cancel out, and long reactor life is attained. Unlike their power station counterparts, military reactors operate at varying power levels, which also influences the design of the reactor.

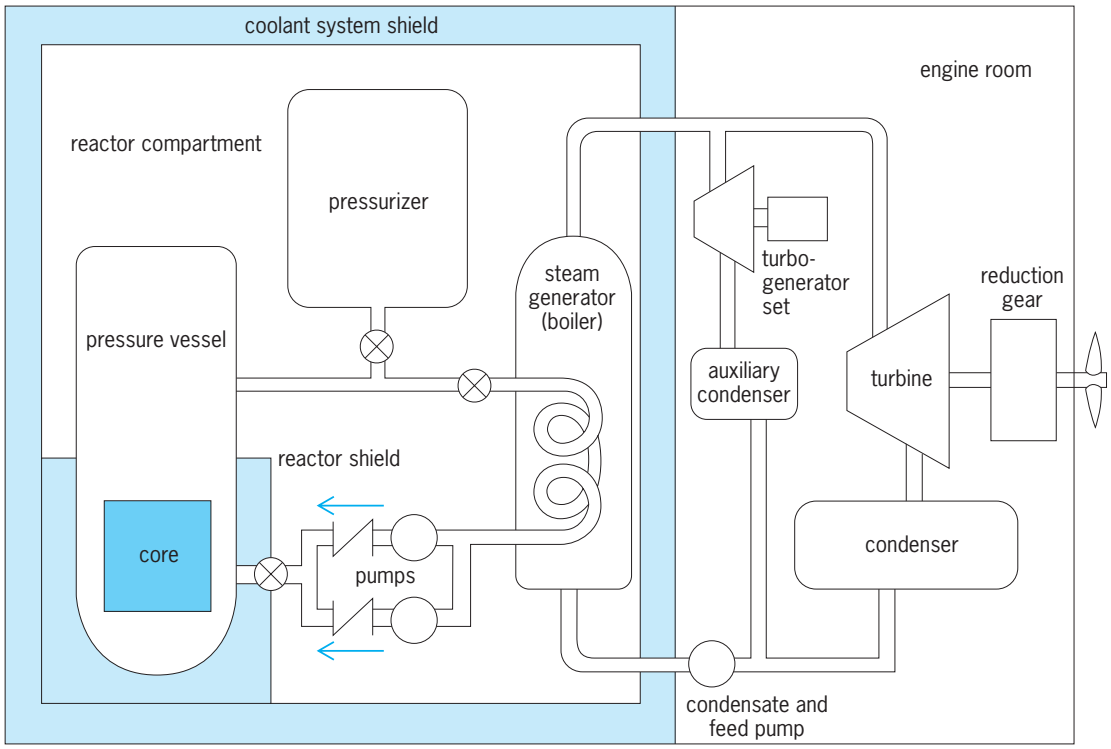

Schematic diagram of a typical pressurized nuclear propulsion plant (relative sizes of the various components are not indicated). (*Society of Naval Architects and Marine Engineers*)

United States. Most of the United States nuclear-powered ships have used a pressurized water reactor with a geared turbine. Several submarines have been turboelectric. All of the submarines were built as either ballistic-missile-carrying (SSBN) or hunter-killer (SSN) submarines. Many of the later SSNs carry cruise missiles. The second nuclear-powered submarine used a liquid metal reactor (sodium) which was replaced by a pressurized water reactor since the sodium design provided no operational advantages and had several disadvantages as a military vessel. The nuclear-powered surface ships, comprising carriers and cruisers, have used multiple pressurized water reactors of various sizes depending on the shaft horsepower required for the application.

The United States has built nearly 200 nuclear submarines. A small deep-diving nuclear-powered research vessel was built by the Navy for oceanographic work. The United States has also built nine nuclear-powered aircraft carriers, all of which (including the first, the *Enterprise*, commissioned in 1961) are still in operation. The nine nuclear cruisers have all been retired or will be soon. *See* NAVAL SURFACE SHIP.

Soviet Union. Most of the Soviet military nuclear-powered ships were pressurized water reactors with geared turbines. The Soviet Union built many classes of submarines (SSNs, SSBNs, and SSGNs) of differing designs depending on desired speed, diving depth, and size. Many of the early submarine classes used two reactors for reliability. The largest submarines in the world, the *Tyfun* class of SSBNs, used two reactors (each in its own pressure hull).

The Soviet Union built four nuclear-powered cruisers. The cruisers had an additional oil-fired steam system to provide superheated steam for a more efficient steam turbine plant. There was a small pressurized water reactor designed to be retrofitted into existing diesel submarines. There were also several deep-diving ocean research minisubmarines, each with a single small pressurized water reactor. Several classes of turbo-electric submarines used a liquid metal reactor with a lead-bismuth alloy as the primary coolant. This allowed a higher steam pressure and a more compact primary and secondary plant. This design had a number of operational and safety problems, and all units have all been removed from service. The Soviet Union built nearly 250 submarines. They also built a nuclear-powered communications ship (with two pressurized water reactors) which operated briefly in the Pacific (probably in support of the Soviet space program).

United Kingdom. The United Kingdom has built about 20 nuclear-powered submarines. All of them have used pressurized water reactors with geared turbines. Both SSN and SSBN types were built. The earliest application was an adaptation of a United States pressurized water reactor, but later submarines have used an indigenous reactor design.

France. All of the French military nuclear-powered submarines have used the pressurized water reactor with a turbo-electric drive. There are two types, SSBN and SSN. France has built more than 10 submarines and a nuclear-powered aircraft carrier using two reactors (based on the current SSBN pressurized

water reactor) driving propellers through geared turbines.

China. The Chinese nuclear-powered submarines, both SSBN and SSN, are believed to be based on early Soviet pressurized water reactor designs and are turbo-electric.

Other countries. The only other country that may be involved in submarine nuclear propulsion is India, which has operated a borrowed Russian nuclear-powered submarine of obsolete design and has announced development of an indigenous design.

Status. At the end of the cold war in 1989, there were more than 400 nuclear-powered submarines operational or being built. By 1994 the number was halved. More than 150 had been taken out of service, and many planned for construction were stopped. The size of the miliary fleets depends on a number of factors. The trend is toward a smaller, more modern fleet. The extra expense of operating and maintaining a nuclear-powered ship has placed pressure on military/national budgets. Underwater detection technology has made many of the early submarines obsolete before the end of their design lives, and the cruisers were not updated with the latest antimissile technology. Arms reduction treaties eliminated many of the SSBNs. *See* ANTISUBMARINE WARFARE.

Currently there are less than 200 nuclear-powered ships with one or two reactors each. Over 5000 reactor years of naval operation have been accumulated. Most are submarines, but the ship types range from icebreakers to aircraft carriers. All of the countries named above continue to build nuclear submarines of various types and, in the case of the United States, an aircraft carrier. There have been no announced plans for new nonnaval nuclear-powered ships. Alan R. Newhouse

Bibliography. *Jane's Fighting Ships*, 1997–1998 edition; *The United State Naval Nuclear Propulsion Program*, USDoE/USDoD, June 1996.

Ship powering, maneuvering, and seakeeping

Powering, maneuvering, and seakeeping constitute the three central areas of ship hydrodynamics. Although this article focuses on these areas in connection with displacement ships, that is, vessels supported entirely (or almost entirely) by hydrostatic forces, basic concepts of powering, maneuvering, and seakeeping are also critical to an understanding of high-speed craft, such as planing boats, hydrofoil craft, and air-cushion vehicles, even if the physical phenomena governing dynamically supported craft are somewhat different. *See* AIR-CUSHION VEHICLE; HYDROFOIL CRAFT; MERCHANT SHIP; NAVAL ARCHITECTURE; NAVAL SURFACE SHIP; SHIP DESIGN; SHIPBUILDING; SUBMARINE.

Powering

Even in antiquity a body of practical knowledge existed among shipbuilders regarding the relationship between hull proportions and speed. In the ancient world, with propulsion limited to human power acting through paddles and oars, the speed advantages of increased length, a high ratio of length to breadth (beam), and hull slenderness were well known, but without a formal theoretical basis. In naval architecture, the term "hull slenderness" has a special meaning; it is often used to describe the relationship between length and displacement. The volumetric coefficient of the hull, for example, is a measure of slenderness defined as displaced volume divided by the length cubed. Generally, lower values of volumetric coefficient are associated with low-resistance forms: long, slender hulls which are more easily driven through the water. In the era of human power, increased length also increased the space available for paddlers or oarsmen, thus making the effect on speed even greater.

With the introduction of sails, first to supplement oars or paddles, and culminating in pure sailing ships, the effects of a ship's stability on sail-carrying power tended to reduce the benefits of slenderness and high length-to-beam ratio. Consequently, even very fast sailing ships were by no means as slender as many earlier galleys or other oared ships. To some extent, the balance was shifted back again by the advent of steam propulsion and the connection between the powering of a ship and its economic performance. The power required to move a ship determined the size, power, and cost of its engines, the cost of fuel per voyage and, because of the weight and volume required for coal, the range and cargo capacity of the ship. By the middle of the nineteenth century, the need for an adequate theory of ship powering was apparent, and the work of William Froude, in particular, marked the beginning of the scientific study of powering. The field of powering is divided into two related issues: resistance, the study of forces opposed to the ship's forward speed, and propulsion, the study of the generation of forces to overcome resistance.

Resistance of ships. A body moving through a fluid experiences a drag, that is, a force in the direction opposite to its movement. In the specific context of a ship's hull, this force is more often called resistance. Resistance arises from a number of physical phenomena, all of which vary with speed, but in different ways. These phenomena are influenced by the size, shape, and condition of the hull, and other parts of the ship; and several of them are interrelated. Frictional resistance is produced by the viscosity of the fluid and the velocity of the ship. Frictional resistance is directly related to the ship's wetted surface area. Therefore, reducing wetted surface, other things being equal, is favorable to reduced resistance. Because of frictional forces, the pressure in a stream of fluid is also lower on the afterbody than on the forebody of the ship. Form drag is the result of this pressure drop. When the afterbody shape is particularly blunt, such as on an immersed transom stern, large-scale eddies are visible, and the form drag is very high. For this reason, form drag is sometimes also identified as eddy-making drag. Because both

frictional resistance and form drag are related to viscous losses, they are often grouped together as viscous resistance. The distinction is that skin friction would exist even on a flat thin plate moving edgewise through the fluid, while form drag exists because of the shape of the body. *See* VISCOSITY.

When a ship moves at or near the surface of the sea, forces are produced which are not the result of viscosity, but the result of surface waves created by pressure variations around the ship. Wavemaking resistance reflects the energy carried away from the ship by the wavetrain. The shape of the wave surface also affects the flow around the hull, and the actual wetted surface as well. Similarly, eddies and other viscous flow phenomena near the stern can affect the formation of the wavetrain. For these reasons, wavemaking and viscous resistance are not entirely independent, although it is convenient to consider them separately.

In addition to the resistance of the bare hull, a ship's appendages can produce significant amounts of drag. Typical appendages include rudders, exposed propeller shafts, shaft bearings and struts, bossings where shafts emerge from the hull, skegs, bilge keels and stabilizing fins, sonar domes, and auxiliary thrusters. Appendage resistance can be a significant part of the total resistance of the ship, often 3–5% for single-screw ships, and 15% for twin-screw ships with exposed shafts and struts.

Because the density of air is only about 1/800 that of seawater, the air resistance of the superstructure and above-water hull of the ship is usually not a very large component of the total. However, superstructures of most ships are quite bluff (broad and flattened) in shape, and the drag contribution may not be negligible. Typically, still air resistance, that is, air resistance due solely to the ship's speed, is a significant part of total resistance only for very fast ships. In addition to still air resistance, wind can add both a direct drag increase and indirect effects. For example, a wind not from dead ahead causes a ship to make leeway (that is, drift slightly sideways through the water), and may also require use of the rudder to hold course. Both of these effects cause additional resistance. Bare hull resistance plus appendage resistance, and generally with still air resistance included, is often referred to as resistance fully appended.

Viscous resistance. Basic to the understanding of viscous resistance is the notion of the boundary layer. This is a relatively thin layer of water (thin with respect to ship dimensions), adjacent to the hull surface, in which viscous forces dominate the flow. Osborne Reynolds, as a result of studies involving flow in pipes, introduced in 1883 a dimensionless ratio of inertial forces to viscous forces in a fluid system. The Reynolds number is often written as $\mathrm{Re} = VL/\nu$, where V is the velocity of flow, L is the length of the system or body under consideration, and ν is the kinematic viscosity of the fluid [given by $\nu = \mu/\rho$, where μ is the (dynamic or absolute) viscosity and ρ is the fluid mass density]. The Reynolds number largely determines the characteristics of the boundary layer, including its thickness, velocity and pressure gradients, turbulence, and friction coefficient. The frictional resistance of a body is given by $R_f = {}^1/_2\rho SV^2C_f$, where ρ is the mass density of the fluid, S is the wetted surface, V is the velocity, and C_f is the friction coefficient, which is a function of Reynolds number only. This function has been determined experimentally and has been given many empirical formulations, of which the present standard used in ship design is $C_f = 0.075/[\log_{10}(\mathrm{Re}) - 2]^2$. *See* BOUNDARY-LAYER FLOW; REYNOLDS NUMBER.

Because form drag is directly related to viscous phenomena, it is usually expressed as a multiplier of the friction coefficient, the form drag factor. There is a wide variation in form drag factor for hulls of differing proportions, but a typical range of values is from 1.05 for a slender, sharp-ended hull to 1.25 or even higher for bluff forms such as tankers.

Wavemaking resistance. Unlike viscous resistance, wavemaking resistance is characterized by behavior which changes dramatically with speed. Water waves, including those produced by a ship's forward motion, propagate with a speed related to their wavelength. At a very low speed, no significant waves are produced and the ship's entire resistance is viscous in origin. It can also be shown, although it is not as intuitively obvious, that wavemaking vanishes in the limit of very high speed, generally too high a speed to be of practical interest for displacement ships. Again, however, the ship's resistance is mainly viscous, although spray also becomes a drag contributor. At intermediate speeds, wavemaking resistance is extremely important. Depending on the hull form, wavemaking often determines the maximum speed that a displacement hull can reach.

Waves exist because a water surface, subject to an initial disturbance, cannot be in static equilibrium; gravity forces the surface to oscillate about its static level. The dimensionless number which characterizes wavemaking resistance is the Froude number, the ratio between inertial and gravitational forces in a fluid system. The Froude number may be written as $\mathrm{Fr} = V/(gL)^{1/2}$, where V is the velocity, g is the acceleration of gravity, and L is the length of the ship, generally taken at the waterline. The form of the Froude number is highly important; it implies that the wave systems generated by geometrically similar ships of various lengths are "in scale" when the ship speeds vary with the square root of length. (A quantity proportional to Froude number, but not dimensionless, appears in many English and American sources. This is the speed-length ratio, defined as $V_k/L^{1/2}$, where V_k is the speed in knots, and the length L is in feet.) *See* FROUDE NUMBER.

In his studies of ship resistance, about 1870, William Froude assumed that the total resistance of a ship, or of a model of the ship, can be decomposed into frictional and residuary resistance components. The frictional resistance of a ship or model was approximated by that of a flat plate of equal area and length. Residuary resistance, then, was obtained by deducting frictional resistance from total resistance.

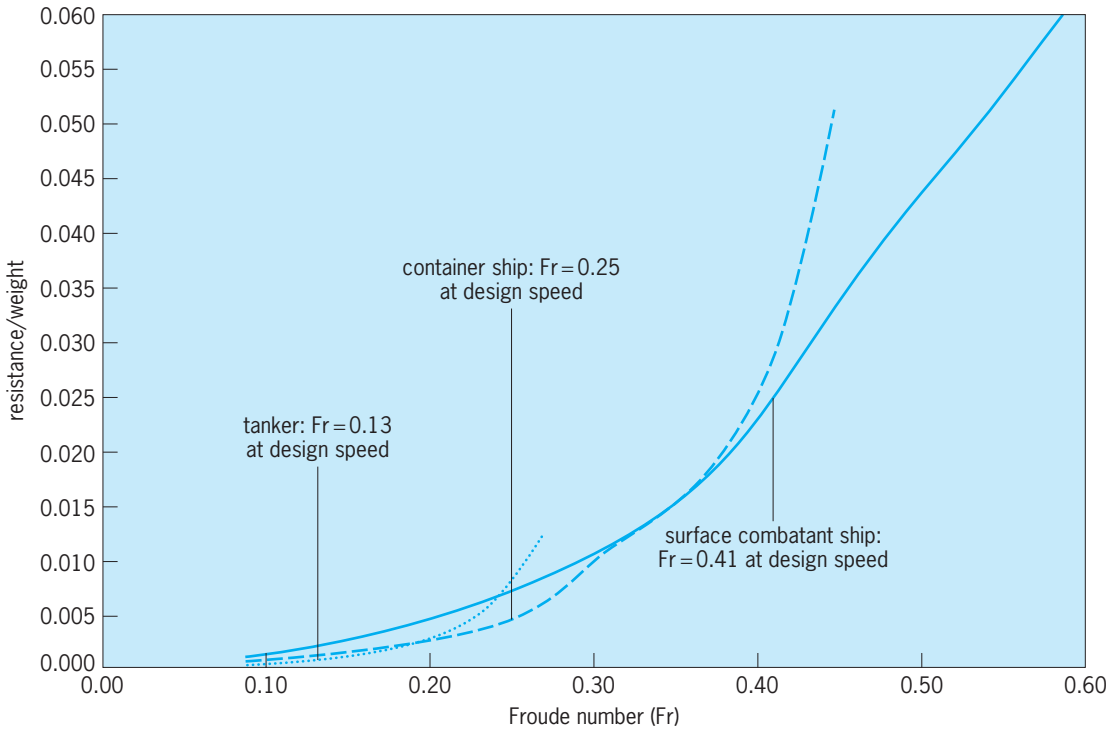

Fig. 1. Comparative resistance of three hull forms intended for different speed regimes. The tanker has displacement of 350,000 metric tons (344,000 long tons), dimensions of 366 × 50 × 22.2 m draft (1200 × 164 × 72.8 ft draft), and design speed of 15 knots or 7.7 m/s (Fr = 0.13). The container ship has displacement of 63,000 metric tons (62,000 long tons), dimensions of 274 × 32.2 × 10.7 m draft (900 × 105.6 × 35.1 ft draft), and design speed of 25 knots or 13 m/s (Fr = 0.25). Surface combatant ship has displacement of 11,300 metric tons (11,100 long tons), dimensions of 195 × 19.5 × 6.5 m draft (640 × 64.0 × 21.3 ft draft), and design speed of 35 knots or 18 m/s (Fr = 0.41).

The essence of Froude's contribution is that to capture the correct wavemaking behavior, a model of a ship must be tested at a corresponding speed to the full-scale ship, that is, at the same Froude number. At corresponding speeds, the model's wavemaking resistance can then be scaled up to the full-size ship simply by multiplying by the cube of the linear scale ratio, that is, $(L_{\text{ship}}/L_{\text{model}})^3$. Frictional resistance must then be recalculated at the ship's Reynolds number, and added to residuary to obtain full-scale ship bare hull resistance. This simple approach, barring the details of how form drag is to be accounted for, is the foundation of ship resistance model testing.

The influence of hull form on wavemaking is profound and rather complicated. Analytical predictions of wavemaking provide a conceptual understanding of the wavemaking phenomenon, but purely mathematical treatments have not yielded reliable design methods for all types of ship forms. Instead, in early stages of a ship design, estimates of wavemaking resistance are based on past data, using systematic series data (results from a series of model tests on a closely related family of hull forms), or regression methods (results derived from a statistical summary of tests for a range of ships, not necessarily systematically related, but generally similar to the hull being designed). As a hull design is developed, it is increasingly common practice to obtain more detailed wavemaking estimates using computational fluid dynamic methods, with final verification by model testing of the design hull. *See* COMPUTATIONAL FLUID DYNAMICS; TOWING TANK; WATER TUNNEL.

The two most important elements in determining a ship's wavemaking resistance are slenderness and whether the hull's volume is spread out into the ends or concentrated nearer amidships. Other form parameters, such as the sharpness of the bow and stern waterline endings, and the shape of the hull sections, also are known to affect both wavemaking resistance and form drag. Ships designed for different purposes and speed regimes have markedly dissimilar hull forms, and their resistance characteristics are quite different (**Fig. 1**).

Unlike viscous losses, the waves generated by a ship's hull may be partially canceled at a particular speed by incorporating hull features which produce waves of opposite phase. A widely used feature of this type is the bow bulb, often seen on ships of moderate to high Froude number, such as container ships, and some naval ships. Bulbous bows are also found on slower ships, operating at low Froude number, such as oil tankers; however, in these applications the effect of the bulb is related to moderating local wave-breaking phenomena around a bluff bow, rather than overall wave cancellation.

Hull roughness correction and correlation allowance. The friction coefficient discussed above applies to a hydrodynamically smooth surface, typically a freshly painted model surface. Consequently, when resistance predictions are based on model test extrapolation,

an additional correction is added to account for differences between the typical surface condition of a ship's hull and an ideal smooth surface. This is sometimes referred to as a hull-roughness correction, or δC_f. Often, however, a roughness correction is included as part of a more general correction, termed a correlation allowance, which encompasses all differences between model test and ship trial conditions not explicitly contained in the friction coefficient. Consequently, the correlation allowance involves not only the resistance of the ship but propulsion as well, and is not independent of specific model-test and trial practices.

In service, the frictional resistance of a ship is obviously affected by changes in surface roughness. In fact, excessive hull roughness and fouling by marine organisms can increase the powering requirements dramatically, increase fuel consumption, prevent efficient powerplant operation, and ultimately impair economic performance. Maintaining the smoothness of hull surfaces is a major concern in ship operation.

Propulsion. Many devices have been used to propel a ship. In approximate historical order they include paddles, oars, sails, draft animals (working on a canal towpath), paddlewheels, marine screw propellers, vertical-axis propellers, airscrews, and waterjets. A key distinction is whether or not propulsive forces are generated in the same body of fluid that accounts for the main sources of the ship's resistance, resulting in hull-propulsor interaction.

Any propulsor can be understood as a power conversion device. Delivered power for a rotating propulsor is the product of torque times rotational speed. The useful power output from the system is the product of ship resistance times ship speed, termed effective power. The efficiency of this power conversion, often termed propulsive efficiency, is of great importance. *See* BOAT PROPULSION.

Wake, thrust deduction, and hull efficiency. The efficiency of a propeller working in an undisturbed uniform inflow is called open-water efficiency. This is a function of propeller characteristics alone: diameter, pitch, blade geometry, revolutions per minute (rpm), and speed of advance, that is, the average velocity of flow into the propeller. For several reasons, however, when a propeller works behind a ship's hull the overall efficiency can depart significantly from the open-water efficiency. *See* PROPELLER (MARINE CRAFT).

Working near the hull, the propeller's speed of advance is different from the speed of the ship through the water. To a great extent this difference is due to the presence of the hull's viscous boundary layer, but the velocity is also influenced by the form of the stern, and in some cases, the presence of the ship's wave system. The ratio of speed of advance V_a to ship speed V is given by $V_a/V = (1 - w)$, where w is the wake fraction, representing the amount of fluid velocity change caused by the hull at the propeller's location. Wake fraction typically varies from values near zero (for twin-screw ships with large propellers extending well outside the hull boundary layer, such as naval combatants) to values of 0.3 or more (for ships with relatively full sterns and propellers deeply embedded in the hull's viscous wake, such as large single-screw tankers).

At the same time, the action of the propeller influences the flow over the after portions of the hull, compared with conditions without the propeller at work. Typically, the effect of a screw propeller near the stern is to reduce the pressure acting on the afterbody. Consequently, the propeller must generate a thrust somewhat in excess of the hull's resistance. The ratio of thrust to resistance is denoted by $T/R = 1/(1 - t)$, where t is the thrust deduction. Typically, thrust deduction is somewhat smaller than wake fraction for single-screw ships, and of the same order or slightly larger for twin-screw.

The ratio of effective power to delivered power involves both wake and thrust deduction. Effective power is defined as $P_e = RV$, where R is the total resistance and V is the ship's speed. The power transferred from propulsor to fluid, the so-called thrust power, is defined as $P_t = TV_a$. From the definitions of wake, thrust deduction, and effective power, this can be rewritten as $P_t = [R/(1 - t)][V(1 - w)]$.

The quantity $(1 - t)/(1 - w)$ is called hull efficiency. It is usually denoted by η_b, and can also be written in terms of power as $\eta_b = P_e/P_t$. It is noteworthy that for single-screw ships, η_b generally exceeds 1. For full-sterned hulls in particular, such as tankers, with propellers located deep in the wake, values of η_b as high as 1.25 to 1.3, or more, are by no means uncommon. Values of hull efficiency greater than 1 do not violate overall conservation of energy; they merely reflect the fact that the ship's hull has transferred energy into the fluid wake which can be partially recovered by the propeller.

Wake adaptation. In open water, the ratio between thrust power and delivered power is simply the open-water efficiency, $\eta_o = P_t/P_d$. However, the distribution of velocities in a ship's wake varies considerably at different radii from the shaft line, and consequently it is often possible to design a propeller to suit the wake distribution of a given hull significantly better than a propeller designed for open flow. The result is that such a wake-adapted propeller design may have a rather low value of η_o when tested in open-water conditions, but is actually more efficient in the wake for which it has been designed. The ratio between the propeller efficiencies in the "behind" and open-water conditions is a measure of the propeller's wake adaptation to a given hull, and is usually called relative rotative efficiency, η_r. Like hull efficiency, its value may exceed unity, but only slightly.

Propulsive efficiency is often defined as effective power divided by delivered power. This can be expressed as a product of the propeller's open-water efficiency, the hull efficiency, and the relative rotative efficiency, that is, $\eta_{\text{propulsive}} = P_e/P_d = \eta_o\eta_b\eta_r$. Mechanical losses between the propeller and the engine include frictional losses in shaft bearings, and in other portions of the transmission system, such

as couplings and gears. These losses are sometimes included as an additional factor in the efficiency, the transmission efficiency, η_t. In this case, the propulsive efficiency is defined as $\eta_{\text{propulsive}} = P_e/P_{\text{brake}} = \eta_o\eta_b\eta_r\eta_t$, where P_{brake} is the brake power of the propulsion engine or motor.

Propulsive effects of fouling. Increased surface roughness of propeller blades can impair efficient powerplant operation to an even greater extent than hull roughness. Blade roughness increases the torque required to turn the propeller at a given rpm, while at the same time blade lift forces and propeller thrust begin to degrade, even at a constant rpm. The result is a significant decrease in propulsive efficiency, even for a constant required thrust. However, when this occurs concurrently with excessive hull fouling, which increases the required thrust, the results are compounded.

Maneuvering

Maneuvering (more generally, ship controllability) includes consideration of turning, course-keeping, acceleration, deceleration, and backing performance. The field of maneuvering has also come to include more specialized problems of ship handling, for example, the production of sideways motion for docking or undocking, turning in place, and position-keeping using auxiliary thrusters or steerable propulsion units. In the case of submarines, maneuvering also includes depth-change maneuvers, either independently or in combination with turning.

Turning. When a ship's rudder is turned from its neutral position, it forces the stern toward the outside of the turn (**Fig. 2**). As a result, the hull is forced to operate at a drift angle, that is, with an angle between the ship's centerline and the path of its center of gravity. It is the hull's side force which actually moves the ship in its curved path, not the rudder. The rudder must provide sufficient moment to start the turn expeditiously and then to balance other moments which would tend to return the ship to a straight course (for example, moments generated by a fixed skeg or by the propellers).

Typically, a ship's tactical diameter is approximately three to five ship lengths. Smaller tactical diameter is generally associated with larger rudder area, larger maximum rudder angle, and higher rudder rate (the angular speed with which the rudder can be moved by the steering engine). Typically, maneuverable ships may have rudders sized as large as 2% or even 3% of the lateral plane area of the hull (approximately the product of length times draft). Many merchant ships, however, have rudders considerably less than 1% of lateral plane area. Normally, maximum rudder angle is limited to 35° port and starboard, although some ships have rudders designed for high maneuverability that can be deflected further. However, smaller tactical diameter is almost always associated with the hull assuming a larger drift angle. Because this produces extremely high drag as well as side forces, the speed loss in a sustained tight turn is very large. Further, at high speeds, extreme turns also can produce large angles of heel. For a displacement ship, the heel is toward the outride of the turn, and the angle in a high-speed turn may become large enough to produce a hazardous condition.

Directional stability. Depending on hull, skeg, and rudder design, a ship may be either directionally stable or unstable. A stable ship, if perturbed from a straight trajectory with the rudder remaining fixed in its neutral position, will tend to return to a straight trajectory. An unstable ship, by contrast, will tend to go into a turn, either to port or to starboard, and remain in a turn, even with the rudder neutral. Consequently, once a turn is established, merely returning the rudder to its neutral position may or may not result in recovery from the turn, even over a long period of time. Further, a directionally unstable ship can keep to a straight course only by moving the rudder slightly from one side to another, more or less continually, which produces an additional drag.

However, while an excessively unstable ship is difficult to control, some degree of directional instability is not necessarily unsafe or even undesirable. Most large tankers, for example, are directionally unstable. However, the inertia of large ships is so great, and their responses are consequently so slow, that tendencies to diverge from the intended course can be easily countered by the helmsman or an automatic controller or autopilot. By contrast, some excessively stable ships are also slower-turning.

Acceleration, deceleration, and backing. Compared with aircraft or land vehicles, even the most powerful

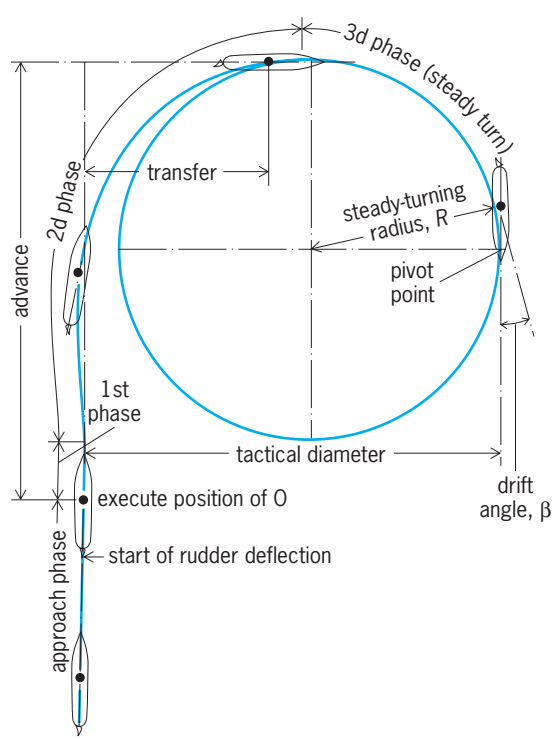

Fig. 2. Turning path of a ship. (*After E. V. Lewis, ed., Principles of Naval Architecture, Society of Naval Architects and Marine Engineers, 2d ed., 1988*)

ships have a very low power-to-weight ratio. Consequently, the accelerative performance of ships is generally a secondary design issue at best. However, acceleration, deceleration, and backing have important effects on controllability. The effectiveness of a rudder, whether in turning or course-keeping, is increased when the rudder is located directly behind a propeller, because the outwash from the propeller, when going ahead, increases the velocity of flow over the rudder, and consequently the available rudder side force. For this reason, when a propeller is producing additional ahead thrust for acceleration, the rudder is also able to act with more authority. Conversely, when a propeller is going astern, the flow over the rudder is reduced and the controllability of the ship may be compromised. Because of the geometry of a ship's stern, a propeller generally produces side force as well as an axial thrust. Consequently, when backing, it is not unusual for a single-screw ship to have markedly asymmetric turning behavior. Further, when attempting to decelerate hard (the so-called crash-back maneuver), single-screw ships tend to yaw in one direction, depending on the direction of propeller rotation, while the rudder may not be sufficiently effective to prevent this.

Lateral thrusters. Many ship-handling tasks require fine control of lateral position and speed. These include docking and undocking (large ships not fitted with lateral thrusters usually employ tug assistance); dynamic positioning (hovering over a position on the sea floor for oil exploration or drilling, for example), and precision track-keeping (for example, while cable laying or surveying).

Lateral thrusters of several different types are often fitted on merchant ships. The most common type is the tunnel thruster, a screw propeller mounted inside a transverse tunnel open at either side of the ship. Thrusters may be mounted at the bow only, or at both bow and stern, where the hull is relatively narrow. Tunnel thrusters are effective only at low speeds, typically below 3 knots (1.5 m/s). Other types, using steerable waterjets or steerable propellers, may remain effective at somewhat higher speeds.

Seakeeping

The modern term "seakeeping" is used to describe all aspects of a ship's performance in waves, affected primarily by its motions in six degrees of freedom (**Fig. 3**). Seakeeping issues are diverse, including the motions, accelerations, and structural loads caused by waves. Some are related to the comfort of passengers and crew, some to the operation of ship systems, and others to ship and personnel safety. Typical issues include the incidence of motion sickness, cargo shifting, loss of deck cargo, hull bending moments due to waves, slamming (water impact loads on sections of the hull), added powering in waves, and the frequency and severity of water on deck.

Even in the era of the wooden sailing ship, some elements of ship geometry were adapted to reduce deck wetness, and to avoid extremely high structural loads due to waves. High freeboard, sheer (the upward sweep of the decks toward bow and stern), elevated forecastles and sterncastles, well-rounded bows and sterns, and sections with higher deadrise (sloping rather than flat-bottom frames) were common features that were known to contribute to seaworthiness and seakindliness, a general and now largely obsolete term referring to deep but "easy" motions (in effect, low accelerations).

A more scientific approach to seakeeping improvement began with the consideration of the ship's roll, especially quick roll, as an impediment to accurate gunnery. It was known centuries ago that a ship with excessive initial stability also tended to roll quickly. More recently, the development of seakeeping theory and design features intended to improve seakeeping has been led by a desire for passenger comfort, and by naval requirements, in particular the operational need for helicopter operations from small or medium-sized surface warships.

Linear theory of ship motions. Beginning in the early 1970s, a simple but very useful theory of ship motions became widely used in ship design. Mathematically, the ship's oscillatory behavior in waves may be considered as that of a linear system consisting of springs, masses, and dampers, subjected to a harmonic excitation due to the forces and moments imposed by an incident wave. Six coupled linear equations of motion are involved, one for each degree of freedom, and the relevant characteristics of the ship may be expressed as coefficients for restoring force, inertia, and damping.

Computation of these coefficient matrices is required. The restoring force coefficients are computed from simple hydrostatics. The inertia matrix elements, however, must include not only the actual mass and moments of inertia of the ship itself, but also the component of the hydrodynamic force which acts in phase with the acceleration of the ship's harmonic motion. This hydrodynamic component is known as added inertia or added mass. Once the required coefficients are computed, the coupled equations may be solved. The resulting solution, at a given wave frequency, speed, and heading, gives the magnitudes and phases of the six ship motions in response to an assumed sinusoidal wave of unit amplitude. Repeating the computation for a range of

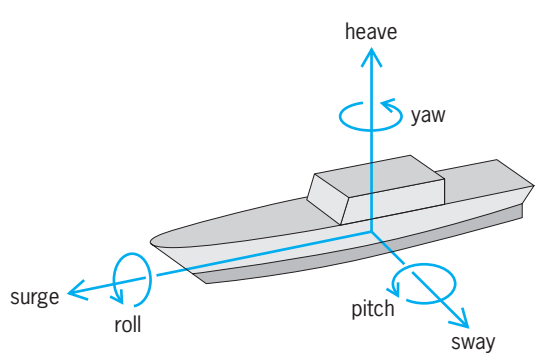

Fig. 3. Ship motion degrees of freedom.

wave frequencies yields six functions of frequency, each known as a response amplitude operator.

Actually, ocean waves are not regular sinusoids, but are random and irregular, characterized by sea spectra analogous to the spectra of other oscillatory natural phenomena such as sound or light. However, linear theory permits real sea conditions to be modeled as a superposition of sinusoidal components, which are linearly additive. Consequently, if a sea spectrum is specified, and the ship's response amplitude operators have been calculated, then spectra of all six motions of the ship can be computed as well. From the motion spectra, statistical measures of the ship's motions and accelerations, including predicted probabilities and extreme values of the motions (the largest values that can be expected to occur within a given time), can be derived.

In spite of the strong simplifying assumptions required to formulate a linear theory of ship motions, results for the vertical plane motions, heave and pitch, are surprisingly accurate, provided that certain limits are not exceeded. However, for roll and lateral motions coupled with roll (sway and yaw), the results of a purely linear treatment are less accurate. Largely this is because roll-restoring moments are often quite nonlinear even at moderate roll angles, while roll damping also includes both linear and viscous (strongly nonlinear) components. As a result, linear theory often tends to underestimate roll in mild conditions but overestimate it in severe conditions.

Added resistance in waves. When a ship operates among waves, even in relatively moderate sea conditions, its resistance and required power generally increase. Added resistance in waves arises from several sources. Waves generated by the ship's motions radiate away from the ship, and the radiated energy in these waves corresponds partly to work which must be done against an added drag force. Also, the incident wave is partially reflected by the presence of the ship, even if its motions are small, and this entails a transfer of momentum or force.

Typically, ship motions contribute most significantly to the added resistance of fast, fine-hulled ships, while reflection forces contribute more in the case of large, massive ships, such as tankers. Apart from direct wave-induced forces, a ship's wetted surface and underwater form are constantly changing due to waves and ship motions, and the result is often increased viscous drag. Finally, the cyclic use of the rudder to remain on course adds yet another drag.

The effects of waves and ship motions on propeller performance are less straightforward. The average thrust required tends to increase due to added resistance, and in general efficiency is reduced. In extreme seas, of course, where the propeller may move vertically a great deal, additional cavitation, air drawing, and even emergence of the propeller can occur, with dramatic losses of thrust. Consequently, even if a ship's speed is not reduced voluntarily in rough seas, a loss of speed often occurs due to power limitations.

Motion reduction. Because roll is relatively lightly damped and easily excited, it is often troublesome as a cause of motion sickness even in relatively benign sea conditions. The incidence of seasickness is particularly high for passengers, who are often less accustomed to the ship's motion than crew members. Most large ships are fitted with bilge keels, appendages which increase roll damping and added inertia. On passenger and cruise ships, among other types, roll stabilization systems are often fitted in addition to bilge keels. Dramatic roll reduction, occasionally approaching 90%, can be obtained with active antiroll fins that are controlled by signals from a gyroscope sensing unit. However, fins are only fully effective at relatively high speeds. Other antiroll systems include active and passive tanks, in which a fluid is either actively pumped or permitted to flow passively, and tuned in such a way as to produce rolling moments out of phase with the wave-induced roll, even if the ship speed is low.

Progress has been made in the design of hull forms with increased damping of pitch and heave motions, including modification of section shapes, and even the experimental use of antipitch fins. In general, these efforts have only been moderately successful, and forms with improved seakeeping performance are often compromised from the standpoint of powering. Another approach to improved seakeeping is reduced wave excitation. Excitation forces are largely generated by changes in the ship's buoyancy during a wave encounter. Consequently, a ship with a small waterplane area is inherently less affected by the motion of the sea surface. Developments such as the small-waterplane-area twin-hull (SWATH) concept have been successful in improving motions in adverse sea conditions (**Fig. 4**).

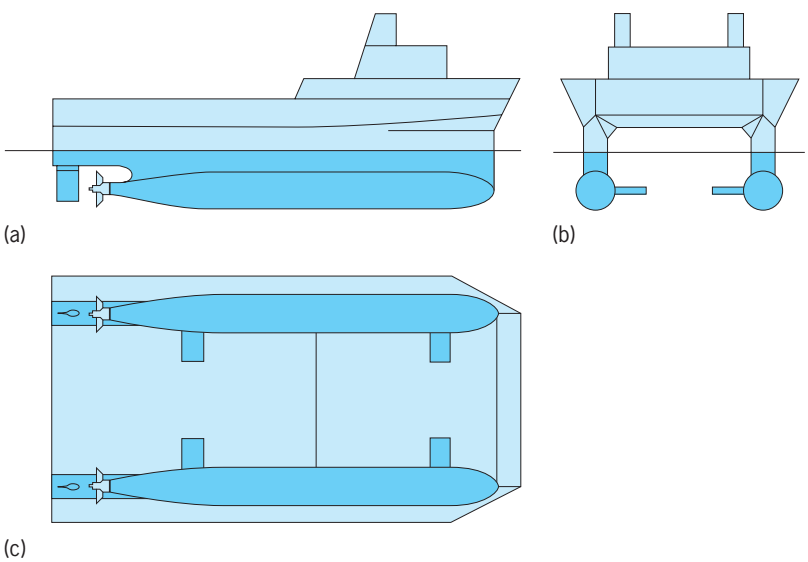

Fig. 4. Small-waterplane-area twin-hull (SWATH) vessel for reduced motions in waves. (*a*) Side view. (*b*) Front view. (*c*) Bottom view.

Role of Computational Fluid Dynamics

In the past, initial powering, maneuvering, and seakeeping predictions for new designs depended

almost entirely on design rules of thumb or, at best, applicable series or regression data from previous model tests, subsequently refined by additional model tests. With the increase in computing power available to the naval architect, computational fluid dynamic methods are now applied to some of these problems in various stages of the ship design process.

Numerical methods for computing the pressures and velocities in the flow around a ship's hull facilitate estimates of resistance and maneuvering forces. These methods include potential flow codes, which give approximate solutions in which viscous forces are neglected, and more complete viscous flow solutions, making certain modeling assumptions regarding the boundary layer. Computational fluid dynamic methods permit estimation of the resistance and of the geometry of the ship's wave system and wake (including predictions of propeller-hull interactions and propeller-induced vibrations), and are also useful in improving propulsor designs. These methods have been applied with some success in simulating surface ship and submarine maneuvers. Computational fluid dynamics (CFD) offers the potential for solving nonlinear seakeeping problems, especially important in extreme conditions where linear theory proves inadequate. While extremely useful in some areas, computational fluid dynamic methods have not been uniformly well developed for practical ship design, and many promising approaches remain in the area of underlying research. Robert M. Scher

Bibliography. R. L. Harrington (ed.), *Marine Engineering*, 1992; E. V. Lewis (ed.), *Principles of Naval Architecture*, 2d ed., 3 vols., 1998; H. E. Saunders, *Hydrodynamics in Ship Design*, 3 vols., 1956; R. Taggart (ed.), *Ship Design and Construction*, 1980.

Ship routing

The selection of the most favorable route for a ship's transit in the open ocean. Safety, economics, and environmental protection play a key role in determining this selection. In the late 1890s, the practice of following predetermined routes for shipping was adopted by shipping companies operating in the North Atlantic. As ships grew both in size and number, the ability to plan open-ocean transits based upon the environmental conditions expected during a voyage and the effect of these conditions upon the ship's performance became important factors in planning a ship's route. Ship routing is also referred to as weather routing, optimum track ship routing, and meteorological navigation.

Figure 1 illustrates the concept of ship routing. A ship is scheduled to depart the English Channel bound for Norfolk, Virginia, and the owners desire a least-time track. Routes can be computed to minimize transit time, avoid waves over a specific height, or, more typically, provide a compromise between travel time and the roughness of the seaway to be encountered. The shortest track is the modified great circle, but in this case the severe storm off Ireland would generate strong opposing winds and heavy head seas. In addition, the ship would obtain little benefit from the favorable, following winds that could be realized from the second storm as it proceeds eastward. The standard track avoids the head seas to some degree, but is a longer route and approaches the second storm so closely that the benefits of following winds are negated by the heavy seas and associated violent ship motions. An optimum route for reducing travel time avoids the head seas, yet takes advantage of the favorable conditions by paralleling the second storm just the right distance to the north. This example is oversimplified, since storms move erratically at times, change intensity, or become stationary. To compensate for this variability, routing offices compute the initial route based on the long-range weather forecast, but maintain a daily plot of both the ship and the storm positions. If conditions warrant, an amended route is sent to the ship master based on the latest weather changes. *See* OCEAN WAVES; STORM DETECTION; WEATHER FORECASTING AND PREDICTION.

Effect of environmental conditions. In addition to all the latest environmental predictions, the ship router must understand how these conditions will affect the specific ship under routing. For this, ship performance graphs are used, similar to **Fig. 2**, which relate the average speed that a ship of a particular class will make under increasing heights of head, beam, and following waves. For some ships, a relationship has been developed between ship speed and both wave height and length. This knowledge is useful, because for some wave lengths the ship's pitching motion becomes accentuated through resonance. The loss of ship's speed is due more to the master's reducing engine speed to ease the ship motions than to direct resistance of the seaway to the ship's passage. *See* SHIP DESIGN; SHIP POWERING, MANEUVERING, AND SEAKEEPING.

Although the effect of waves is the predominant factor considered, the ship router also utilizes knowledge of ocean currents, wind, and hazards to navigation such as ice fields or fog banks. On U.S. east coast runs, the currents are more important than the waves. Most major oil companies have their tankers follow the Gulf Stream on northbound runs and avoid it on the return trip. This practice is logical, since a ship can gain a 3–5-knot (1.5–2.5 m/s) push by riding the current. The National Oceanic and Atmospheric Administration (NOAA) issues the weekly *Gulf Stream Bulletin*, which provides ships with the latest position of the Gulf Stream as measured by meteorological satellites carrying infrared sensors that detect warm water. *See* GULF STREAM; METEOROLOGICAL SATELLITES.

Extent of application. Ship routing is available from a number of environmental consulting companies, and is utilized by many of the large shipping companies. The U.S. Navy operates its own routing service, which provides some 1200 transits per year to everything from seagoing tugs to aircraft carriers. Routing services are available from at least five countries other than the United States: the United

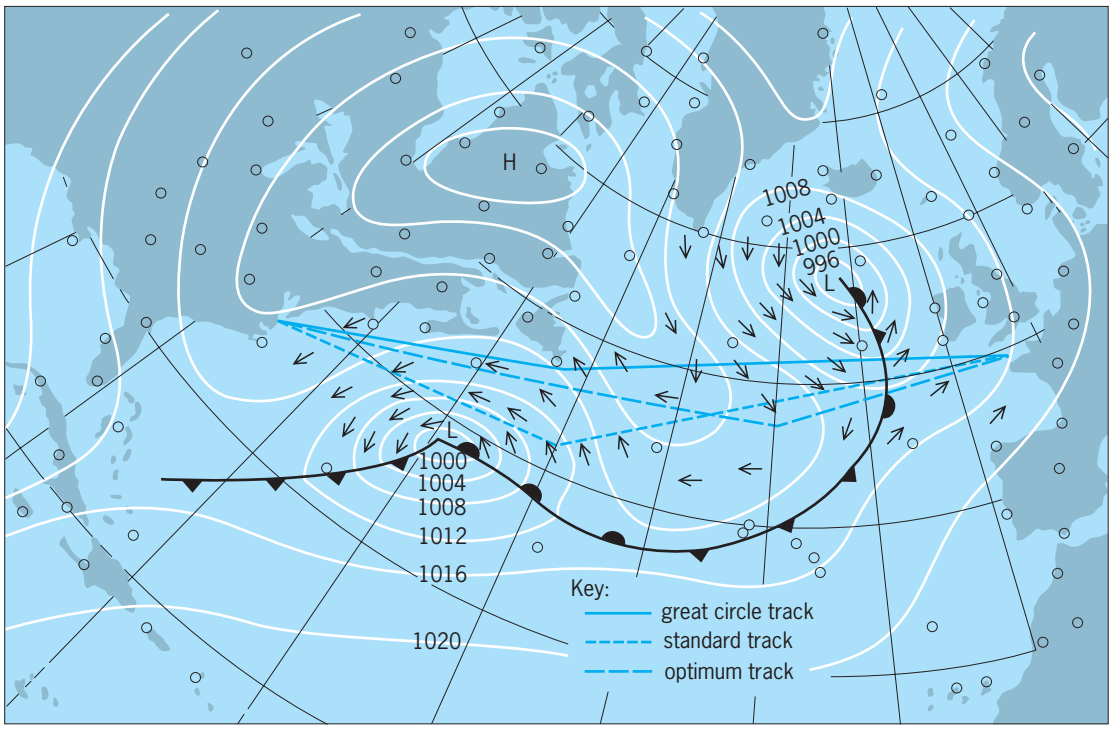

Fig. 1. Meteorological chart illustrating concept of ship routing. (*After R. W. James, The present status of ship routing, in Proceedings of Interocean '70, Düsseldorf, November 1970*)

Kingdom, Germany, Netherlands, Norway, and Russia. Additional information on weather routing services is available from the World Meteorological Organization.

Ship routing also benefits from improvements in all weather navigation capabilities such as those provided by the Global Positioning System (GPS). Additionally, new standards being implemented by the International Maritime Organization (IMO) provide for the automated exchange of ship information through a Universal Shipborne Automatic Identification System (AIS). This system employs automated transponders designed to operate in a ship-to-ship mode for collision avoidance and a ship-to-shore mode when a ship's route begins to approach coastal areas. Providing coastal authorities with information about a ship's speed, course, position, and cargo assists in the efficient management of ship traffic in congested shipping lanes and ports. *See* SATELLITE NAVIGATION SYSTEMS; VESSEL TRAFFIC SERVICE.

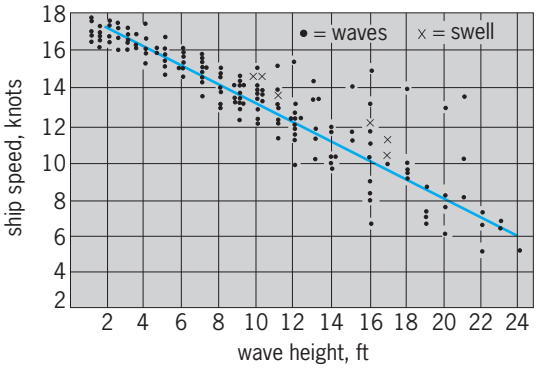

Fig. 2. Ship performance graph showing relationship between speed of ship and head waves. (1 ft = 0.3 m; 1 knot = 0.5 m/s.)

Benefits. Ship routing is very cost-effective. The charges for a route are low compared to the savings realized by reduced storm damage, less fuel consumption, and shorter travel times. For a large company, the reduction in transit time means that more cargo is carried in fewer ships than would be possible if weather routing were not used. Additional intangible benefits accrue to routed ships in terms of crew and passenger well-being, since the ship has a smoother crossing than nonrouted ships. When the savings of ship routing are coupled with improved navigation capabilities and automated ship reporting systems for coastal and port management, further cost saving are realized.

Ship routing has progressed from the simple concept of the sailing masters to "follow the winds" to modern practice which utilizes such technology as satellite data to track storms and measure currents, high-speed computer calculations, and modern communications to follow the ship's progress and provide route diversions where necessary for safety, economics, and protection of the environment.

Richard W. James; Douglas Taggart

Bibliography. N. Bowditch, National Imagery and Mapping Agency, *The American Practical Navigator, 2002 Bicentennial Edition*, Celestaire, Inc./Paradise Cay Publications, 2002; T. J. Cutler, *Dutton's Nautical Navigation*, 15th ed., Naval Institute Press, 2003.

Ship salvage

Voluntary response to a maritime peril by other than the ship's own crew. Salvage is usually time-critical. The property in danger can be any type or size of vessel or maritime cargo. Salvage is encouraged by compensating the salvor based on risk to the salvor and the salvor's equipment, conditions under which the salvage service was performed, the value of the vessel and cargo saved and, more recently, minimizing environmental damage.

Aside from international agreements requiring the saving of life at sea, salvage is voluntary, and is driven by economics. Given sufficient time and money, salvage masters and salvage engineers can accomplish virtually any salvage operation. However, the payoff may not be worth the risk and expense, and the ship and cargo owners or the underwriters may not want to spend more than the value of the ship and cargo in order to save them.

Contrary to popular belief, the salvor does not own or gain title to recovered property, but establishes a maritime lien. The salvor's award may be established by agreement between the salvor and the owner, master, or insurance underwriter, or it may be adjudicated either by the courts under admiralty law or by an arbitrator agreed to by the owner, master, or underwriter and the salvor. Changes in maritime law and fee structures, and public response to major pollution casualties, have focused the salvor's attention on preventing or minimizing oil discharge when attending a casualty.

The salvor must often work with limited resources, usually in a remote location in bad weather. Salvors are a combination of maritime ambulance, engineer, and adventurer. Critical decisions must be made and implemented quickly.

Ships, equipment, and personnel. Salvage services can be performed using either dedicated salvage vessels or vessels hired for the specific operation. Salvage ships are typically large seagoing tugs, usually capable of 5000–10,000 hp (3750–7500 kW), designed and outfitted to work at remote sites in all weather (**Fig. 1**). The largest tugs can be as powerful as 40,000 hp (30,000 kW) and exert a pull against a fixed object, called bollard pull, of up to 250 tons (227 metric tons) with their main tow wire. Wire used in salvage is made from multiple twisted steel wires and can exceed 2.5 in. (6.4 cm) in diameter. Salvage ships carry a variety of portable salvage equipment such as firefighting gear, pumps and patching material, electrical generators, compressors, diving equipment, wire and chain, beach gear, pollution control equipment, and the material to make whatever might be needed on-site.

Maritime countries sometimes maintain ships on "salvage station" near high-density shipping lanes or areas where problems historically occur. With improvements in navigation, communications, and engineering reliability, there has been a reduction in the number of casualties, although the potential for pollution and damage and award are generally larger per casualty. The net result has been a reduction in the number of dedicated salvage ships.

Some salvage companies maintain portable equipment at a central location and mobilize it quickly by truck, rail, or air to vessels of opportunity, usually a tug or barge, near the salvage site. Other salvors

Fig. 1. USS *Safeguard* (ARS50). (*U.S. Navy*)

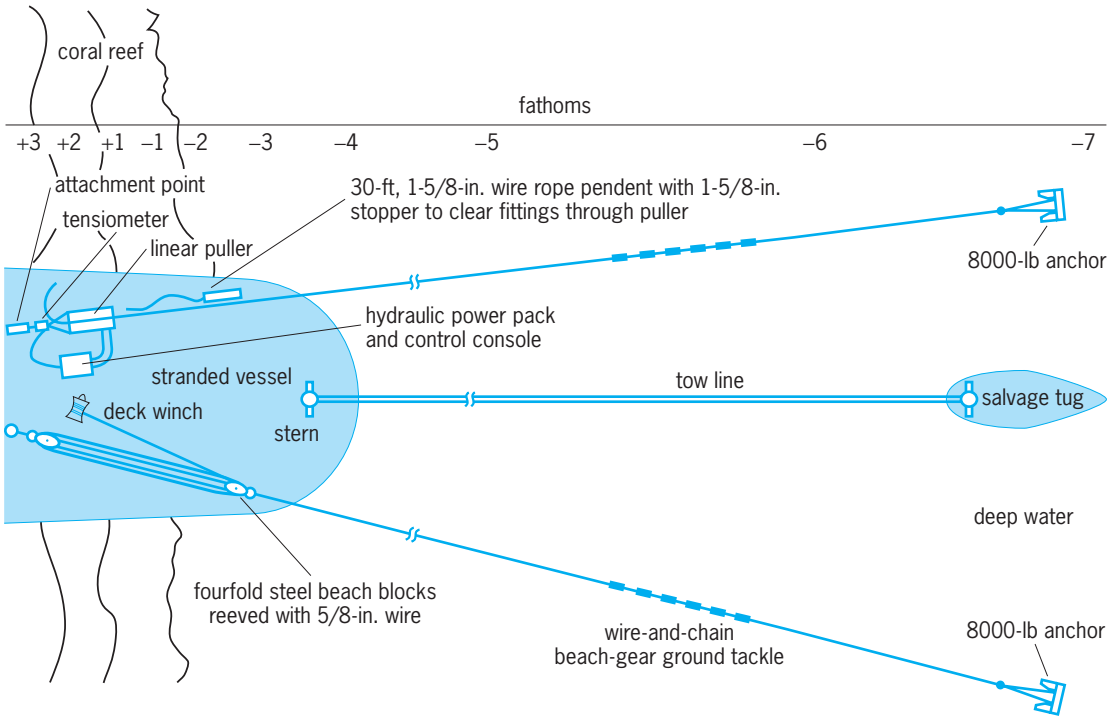

Fig. 2. Beach-gear deployment. Beach gear enables a stranded ship to assist in pulling itself into deep water. 1 fathom = 1.8 m; 8000 lb = 3600 kg; 30 ft = 9 m; 1–5/8 in. = 41 mm; 5/8 in. = 16 mm.

have established loosely allied groups of specialists and equipment providers that they call upon as required.

Trained salvors frequently have multiple skills, and can be riggers, welders, divers, crane operators, carpenters, and pump engineers. They are supervised by an experienced salvage master. A salvage engineer is often part of the team and can determine the hull strength and stability of a casualty. Together, the salvage engineer and the salvage master determine the best way to refloat a vessel, and develop a detailed salvage plan to accomplish the work.

Salvage situations. True salvage is time-critical, involving assets that are immediately available. Assistance is provided in situations involving collisions, firefighting, flooding and damage control, damage from hostile action, and breakdowns of propulsion or steering systems. The salvage team may fight fire from off-ship or board the casualty to fight fires. Salvors also may board the ship to control flooding and stabilize the casualty if required. Rescue towing often results from one or more of the above and the need to reach a safe haven or avoid going aground.

Stranding. This results when a ship drives, or is driven, aground through engine or steering casualty, error, problem with navigation or navigation aid, or bad weather, and cannot extricate itself. Time-criticality may be due to coming bad weather, tidal range, ongoing damage to the ship from the grounding, or other requirement such as clearing a blocked shipping channel. The salvor must assess the condition of the casualty and what is required to refloat it, the potential for further damage to the ship, and the potential for loss of cargo and damage to the environment. A wide range of other factors such as weather, tides and currents, the weight of the ship now supported by the ground (ground reaction), and the type of bottom (coral, mud, rock, sand, and so forth) also must be considered in developing a salvage plan. The situation will usually deteriorate quickly without immediate intervention by the salvor.

A lightly grounded vessel can frequently be freed by tugs pulling at high tide. This is often combined with moving fuel or cargo, especially liquids, within the ship to lighten the grounded portion of the hull, or jettisoning or removing fuel or cargo.

The tops of damaged, flooded compartments such as tanks or engine rooms can be temporarily sealed. Compressed air is then pumped into the compartments. The air displaces the liquid tank contents, usually a combination of flooding water and tank contents such as oil or water, and provides a bubble of air in the tank to restore some buoyancy to the ship.

Beach gear adds great pulling force with relatively little equipment (**Fig. 2**). Beach gear consists of anchors, wires and chain, and a pulling power source such as a winch or hydraulic puller. Anchors are placed to seaward and connected by chain and wire to the stranded vessel. The wire is led either to a hydraulic puller, which has clamps that "walk" a wire in using hydraulic rams, or to a multiple purchase tackle system leading to one of the ship's winches. Each leg of beach gear can exert a force of about 50–60 tons (45–54 metric tons) to help pull the ship off.

Environmental salvage. This is the protection of the environment from damage by pollutants that may result

from a maritime casualty. International agreements such as the International Convention for the Prevention of Pollution from Ships (MARPOL 73/78) and national laws such as the Oil Pollution Act of 1990 highlight the need for protecting the environment. This is most effectively accomplished by keeping pollutants in a ship during a casualty, minimizing dispersion, and rapidly recovering any spilled pollutants, usually cargo crude oil or petroleum products, or ship's fuel stored in bunkers. Salvage actions such as pumping air into a damaged tank to restore buoyancy may blow oil out through the damaged hull below the waterline. Salvors must be prepared to deal with pollutants when this occurs, and plan salvage operations to minimize oil release. Regulations and standard salvage agreements provide financial incentives for pollution reduction. Many salvage companies are capable of controlling pollution.

Wreck removal. This involves recovery of a stranded or sunken vessel, and is usually not time-critical. The Suez Canal clearance in 1974 is an example of a large wreck removal operation.

Many salvage techniques can be brought to bear. If the main deck of the stranded or sunken vessel is above water, a combination of patching and pumping may be sufficient to refloat the casualty. If the main deck is not submerged too deeply, a watertight wall, known as a cofferdam, may be built around the main deck to above the water surface. Patching and pumping may then refloat the casualty.

Floating cranes, some capable of lifting up to 14,000 tons (12,700 metric tons), can be used to lift vessels using wires attached to the hull by divers. The casualty is raised to the surface and pumped out, lifted onto a barge, or moved to a shallower location for patching and pumping. In 1982, a floating crane and barge-cradle combination was used to raise Henry VIII's flagship, the *Mary Rose*, sunk in 1545, and place it in dock in Portsmouth, England, for conservation.

Parbuckling is one of the techniques used to right a vessel that is on its side or inverted (**Fig. 3**). Wires are run from the edge of the main deck on the uppermost side of the hull, around the hull, across the keel, under the hull (passing the main deck on the opposite side), and then to a source of lift, usually a crane, on the surface. Applying a lifting force on the wires causes the vessel to be rolled to an upright position. Other techniques can then be applied to float the vessel.

Pontoons, attached to the sunken vessel by wires and then inflated, can provide lifting force. They are difficult to manage, and are generally used for lighter lifts or to supplement other forms of lift or buoyancy.

Heavy-lift craft can raise sunken ships in stages (**Fig. 4**). Wires are run under the casualty between the flooded-down heavy-lift craft. The heavy-lift craft are pumped out on a rising tide, and the casualty is moved to shallower water, where it rests on the bottom as the tide falls, and the heavy-lift craft are again flooded. This process is repeated until the casualty is in water shallow enough for patching and pumping efforts to begin.

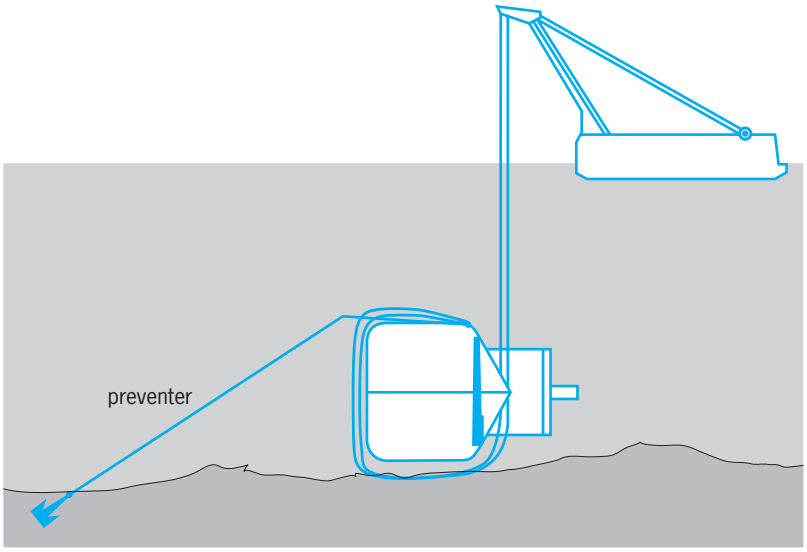

Fig. 3. Parbuckling technique, used to right a ship that is on its side or capsized. Wires are passed from the uppermost deck edge, around under the hull, and to a source of lift. Tensioning the wires rolls the ship upright, and it can then be raised by customary means. A preventer may be used to keep the ship from dragging.

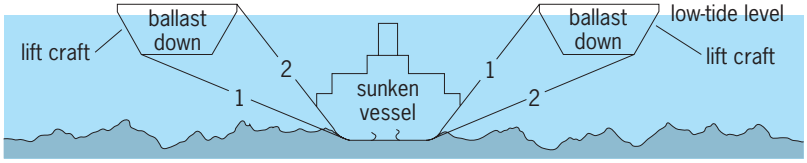

(a)

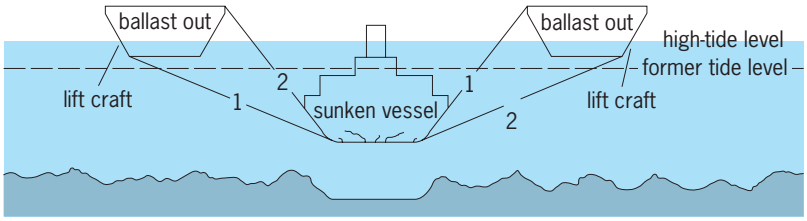

(b)

Fig. 4. Lift craft operation. (*a*) At low tide, the lift craft are ballasted down, and cables are passed beneath the sunken vessel and pulled tight. (*b*) At high tide, the lift craft are deballasted, and the sunken vessel is raised above the ocean bottom and towed away.

Wrecks can also be removed by breaking them into pieces small enough to be lifted using cranes on site. This can be accomplished with cutting torches, explosives, or something as simple as a large steel plate, sharpened on one edge, repeatedly raised by a crane, and dropped onto the same spot, cutting the wreck into liftable pieces.

Cargo. A ship's cargo may be of more value than the ship itself, and the salvor may recover the cargo without recovering the ship. The best-known type of cargo salvage is the recovery of jewels or precious metals. But even something as mundane as copper ingots may be considered worth recovering by both the salvor and the owner or underwriter.

Salvage contracts. The concept of "no cure/no pay" is unique to marine salvage contracts. With the exception of recovering expenses for environmental

protection efforts, the salvor who does not save at least some property for the owner is not entitled to compensation. Most time-critical salvage agreements incorporate these terms.

Perhaps the best-known agreement is the Lloyds Open Form (LOF) and its variants, originated in 1892 at the early center of maritime insurance, Lloyds of London, and updated at infrequent intervals. The LOF is an agreement to have an arbitrator at Lloyds evaluate the salvage service in light of the degree of danger, value of the property saved, salvor's skill and effort in preventing environmental damages, and salvor's expenses and long-term investment in salvage equipment. The amount of a salvage award cannot exceed the value of property saved.

A salvor who finds a ship abandoned at sea and tows it to safe haven without a contract has performed voluntary salvage. The salvor does not own the derelict, but has a claim against the ship for an award. If the claim cannot be resolved with the ship owner, the salvor may go to court to obtain an award.

Technology changes. Advances in technology have had a significant impact on ship salvage. Improvements in navigation and piloting equipment, training, engineering reliability, and ship safety have reduced the number of ship casualties requiring salvage assistance. On the other hand, the increased size of ships, particularly oil tankers, has increased the potential impact of each casualty. Ship owners are now required to have extensive casualty plans.

Larger salvage tugs and larger floating cranes have greatly increased the pulling and lifting forces that can be brought to bear on a casualty, and enabled salvage operations to take place in worse weather than they could before.

Salvage engineering advances have enabled salvors to more precisely calculate required lifts, pulls, and pumping. Computer programs enable modeling of ships and rapid calculation of casualty strength and stability, including the impact of damage on hull strength and of flooding on ship stability.

Deep-ocean search, observation, and recovery techniques, including towed sonars and cameras, remotely operated vehicles (ROVs), small crewed submersibles, and saturation diving have brought many otherwise unrecoverable and potentially abandoned wrecks within reach. In 1992, the U.S. Navy located and recovered a crashed helicopter in over 17,000 ft (5180 m) of water. Inaccessible for years, the submerged Titanic is now visited routinely. This access to deeper wrecks has resulted in much legal activity in admiralty law concerning ownership, abandonment, and salvage claims, particularly for high-value cargoes or vessels having significant archeological or historic value. Search, observation, and recovery techniques using towed sonars, ROVs, divers, and salvage ships and equipment are also now used routinely in aircraft salvage such as the crash of TWA Flight 800 off Long Island, New York, in 1996. *See* SONAR; UNDERWATER VEHICLE; UNDERWATER SOUND; UNDERWATER TELEVISION.

Richard P. Fiske

Bibliography. E. Ellsberg, *Under the Red Sea Sun*, 1946, paperback 1974; G. Forsberg, *Salvage from the Sea*, 1977; W. I. Milwee, Jr., *Modern Marine Salvage*, Cornell Maritime Press, 1996; F. Mowat, *The Grey Seas Under*, 1958 reprint, Lyons Press, 2001; National Research Council, *A Reassessment of the Marine Salvage Posture of the United States*, 1994; G. H. Reid, *Marine Salvage: A Guide for Boaters and Divers*, Sheridan House, 1996.

Shipbuilding

The construction of large vessels which travel over seas, lakes, or rivers. A ship must be capable of self-sufficiency in a hostile marine environment for extended periods of time while simultaneously providing the necessities and comforts expected by passengers, the safety and timely arrival of cargo, the offensive and defensive capabilities required by warships, and a variety of other particular functions. The bulk-cargo carrier, such as an oil tanker in which the cargo is voluminous and homogeneous with no premium on speed of transport, is the most efficient mode of freight transportation.

Many different approaches have been used in the construction of ships. Sometimes a ship must be custom-built to suit the particular requirements of a low-volume trade route with unique cargo characteristics. Such a ship may be one of a kind, and as a consequence there may be little opportunity to employ labor-saving techniques during construction due to the low incidence of repetitive processes. A cable-laying ship, used to lay communication lines across bodies of water, could be of this type. On the other hand, there are many instances where a significant number of similar ships are constructed, providing an opportunity to employ procedures which take advantage of repetitive processes. Examples of this construction procedure, known as series production, include the production of U-boats in Germany during World War II, oil tankers in Japan during the 1960s, and destroyers by the Ingalls Shipyard in Mississippi during the 1970s.

Shipyards can be laid out so as to be ideally suited for series production; however, such a layout may be entirely unsuitable for the construction of dissimilar ships if market conditions change. Obviously, the layout used to build U-boats would not be suitable for the construction of ocean liners. Consequently, shipyard facilities are designed to take advantage of labor-saving techniques which can be used in the construction of ships which have an immediate market, while, at the same time, avoiding inflexible facilities that cannot be easily modified to construct ships of different types.

Shipyard layout and facilities. A shipyard site must be on deep water accessible to the ocean (or, in the United States, to the Great Lakes) and should have railroad connections for the delivery of materials. **Figure 1** shows the layout of a large American shipyard. In such shipyards, where a variety of ships are built concurrently, sufficient space must be provided for the storage of vast amounts of material and equipment due to the impracticability of scheduling such items to be delivered to the shipyard

at precisely the time required. In addition to the space required for the storage of purchased steel plates, rolled steel sections, pipe, and other materials, space is required for the proper storage of steam boilers, turbines, reduction gears, pumps of all types, and the broad variety of specialized mechanical, electrical, and electronic equipment that is usually purchased by the shipbuilder. Some shipbuilders manufacture their own specialized equipment; others have an extensive manufacturing capability, but elect to purchase such equipment from other suppliers. The shipbuilder's decision is determined by shipyard capability, worker-power availability, and economics. Since equipment manufacturers serve many shipyards and limit their production to specific types of equipment, they are able to take advantage of labor-saving techniques in their area of specialization, and this often gives them an economic advantage. However, such is not always the case. The production of slow-speed diesel engines for ship propulsion does not lend itself to a small-scale operation. These diesel engines may be more than 42 ft (13 m) high and weigh 1375 short tons (1250 metric tons), with many of the engine parts weighing several metric tons. Consequently, the large, heavy, and expensive facilities required to manufacture and assemble these diesels are similar to those required for other shipbuilding processes. As a result, some shipbuilders manufacture their own diesel engines.

The handling and processing of steel through the production processes require the greatest amount of facilities and space in a shipyard. Not only must the steel be received, inspected, sorted, and stored, but it must also be blasted, primed, cut to shape, perhaps formed to the proper configuration, welded to make assemblies, and then stored again until needed to make larger erection units.

The building of a ship can be divided into seven phases: design, construction planning, work prior to keel laying, ship erection, launching, final outfitting, and sea trials. Ship design and sea trials are discussed in other articles; therefore, this article covers the remaining five phases and is generally applicable to the building of both merchant and naval ships. *See* SHIP DESIGN; SHIP POWERING, MANEUVERING, AND SEAKEEPING.

Nomenclature. The process of constructing the steel hull of a ship requires an understanding of terms which are not used in other industries. The type of structure used varies considerably for different types of ships, but the descriptive terms do not. For illustration purposes, consider the structure of a simple cargo vessel shown in **Fig. 2**. The keel consists of a flat keel and a vertical keel. The flat keel is the center strake (row of plates) of the bottom shell plating; and the vertical keel is a vertical plate on the bottom centerline, running the entire length of the ship. The bottom shell and side shell form the outer, watertight envelope of the ship.

Floors are vertical plates extending across the bottom of the ship, perpendicular to the centerline. Longitudinal girders are similar vertical plates, generally running parallel to the centerline. An inner bottom is several feet above the bottom shell, and provides protection in case of minor bottom damage and a smooth bottom surface for the holds and machinery spaces. Decks are horizontal surfaces corresponding to the

Fig. 1. Newport News shipyard. (*Newport News Shipbuilding*)

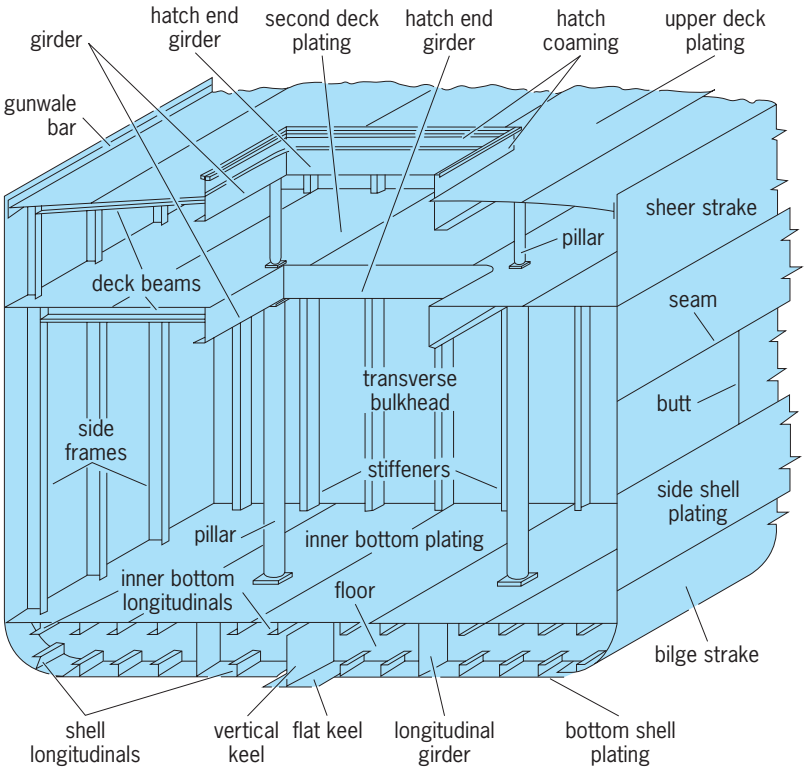

Fig. 2. Section through the structure of a simple cargo vessel.

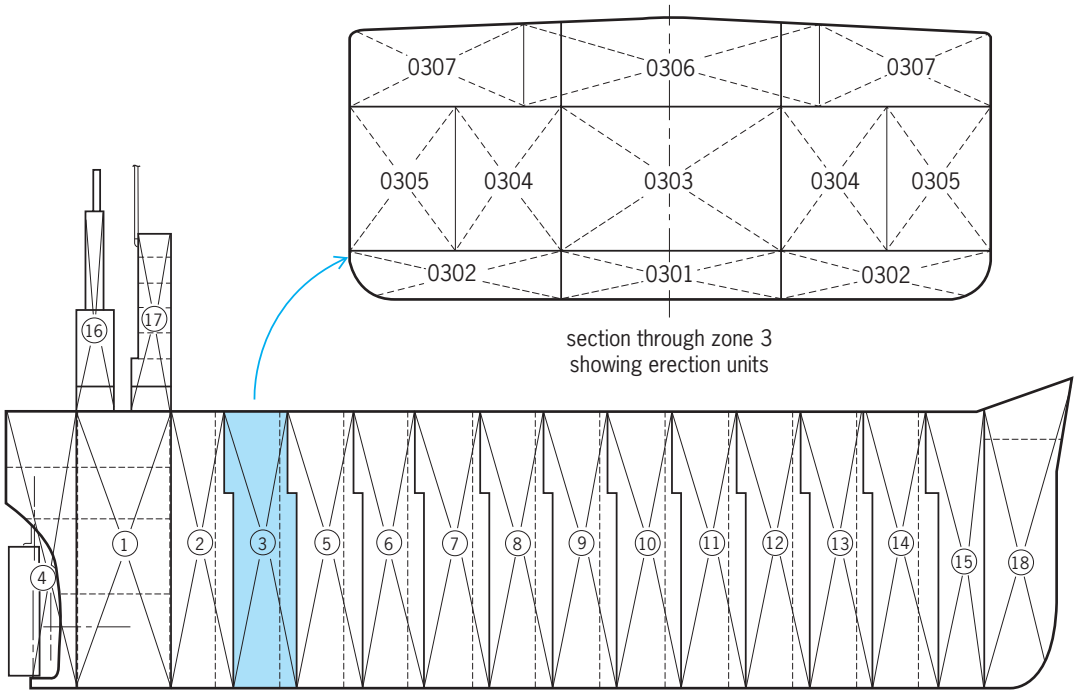

Fig. 3. Erection diagram for a tanker.

floors in a building; and bulkheads are vertical partitions, either longitudinal (fore and aft) or transverse (perpendicular to the centerline of the ship), that divide the ship into compartments. Frames, either longitudinal or transverse, are the structural members which hold the shell against the pressure of the sea. Deck beams support the deck and its loads, and in turn are supported by girders which rest on stanchions or pillars. Seams are the weld connections between the longer edges (usually fore and aft) of plates, and butts are weld connections between the shorter edges of plates. Seams connect strakes, and butts connect the plates within a strake.

Construction planning. The construction planning process is one of the most important aspects of shipbuilding because it establishes the construction techniques to be used and the schedules which all of the shipbuilding activities must follow. Construction planners generally start with an erection diagram (**Fig. 3**) on which the ship is shown broken down into erection zones and units. The first zone to be erected includes the machinery space which is beneath the deck houses; following that, zone 2 (which is just forward of the machinery space and includes the cargo pump room) is erected; and then zone 3 (the first of a series of cargo tanks) is placed in position. The typical midship section shown by Fig. 3 indicates a breakdown of the units which compose erection zone 3. To facilitate the fabrication of steel, insofar as possible, the erection units are designed to be identical. The size (or weight) of the erection units selected is usually limited by the amount of crane capacity available. Crane capacity at the various shipyards varies widely from less than 110 short tons (100 metric tons) to more than 1100 short tons (1000 metric tons), and it is the construction planner's job to plan the work to make optimum use of the facilities available.

Larger crane capacities do not necessarily equate to greater shipbuilding productivity, inasmuch as some shipyards with low crane capacities can be very competitive; however, the general trend of shipyards is toward higher crane capacities which provide the ability to assemble (erect) a ship by using large erection units which have been preoutfitted to the greatest practical extent. The practice of erecting a ship by using large erection lifts provides the advantage of minimizing the time that a dock is occupied by a ship being built, and permits large erection units to be constructed away from the dock under more favorable conditions. For example, with proper planning and suitable crane service, the steel structure can be turned during fabrication so that most of the welding is done in the more efficient downhand position. Also, the erection units can be transported to paint sheds where they can be painted or coated in the controlled environment required in some climates. In addition, the use of staging (temporary structure, such as ladders, for access) during ship construction is minimized. However, the use of large erection lifts has the disadvantage of requiring facilities for transporting and lifting erection units, as well as requiring facilities and space for the fabrication, processing, and storage of erection units. Of course, many ships have been erected essentially one plate at a time, and there is no reason why this practice should not continue.

Once the construction planners have established the manner in which the ship is to be erected and the sequence of construction, the schedules for construction can be developed. Working backward from the time an erection unit is required in the dock, with

allowances made for the many processes involved, a schedule of working plans is prepared. This is the schedule that the ship designers must follow in order to properly support the construction trades. Similarly, a schedule for the procurement of all purchased equipment must be developed. This requires that the construction planners decide at what point in the construction process the purchased equipment is going to be installed; that is, the equipment may be preoutfitted on an erection unit (thereby requiring an early delivery), it may be landed aboard ship just before the deck is installed over it (thereby permitting a later delivery), or it may be practicable to install the equipment after launch (thereby permitting the latest possible delivery).

Work prior to keel laying. Before the keel of a ship is laid (or when the first erection unit is placed in position) a great deal of work must have been accomplished for work to proceed efficiently. The working drawings prepared by ship designers completely define a ship, but often not in a manner that can be used by the construction trades people. Structural drawings prescribe the geometry of the steel plates used in construction, but they cannot be used, in the form prepared, to cut steel plates. Instead, the detailed structural drawings must be translated into cutting sketches, or numerical-control cutting tapes, which are used to fabricate steel. The trend is toward the use of computer-driven cutting torches which require the use of numerical-control cutting tapes. Several organizations have developed sophisticated computer programs which readily translate detailed structural drawings into machine-sensible tapes which can be used to drive cutting torches (**Fig. 4**). In addition to cutting the steel parts, many of the plates must be rolled into the desired configuration (**Fig. 5**). The parts must also be fabricated to form the erection units, and the erection units must be coated and preoutfitted. The preoutfitting (installation of piping, machinery, or other nonstructural items) of erection units may be required by the construction schedule and, if so, must be accomplished if work is to proceed efficiently.

Fig. 5. Steel plate being formed in a roll press. (*Newport News Shipbuilding*)

Fig. 4. Steel plate being cut to the desired configuration by a computer-driven cutting torch. (*Newport News Shipbuilding*)

Ship erection. If all of the preceding work has been accomplished properly and on schedule, the erection of a ship can proceed rapidly; however, short-term problem areas invariably arise. When erecting a ship one plate at a time, there are no serious fitting problems; but when 990-short-ton (900-metric-ton) erection units do not fit (or align) properly, there are serious problems which tend to offset some of the advantages previously mentioned for this practice.

Figure 6 shows a crude-oil tanker in the process of being erected. In Fig. 6*a* the centerline erection unit, which has been preoutfitted to contain the cargo-oil piping before being placed in the dock, can be seen to be substantially complete. Figure 6*b* shows a number of erection units being preoutfitted alongside the dock and aft of the ship. By referring to Fig. 3, the manner in which the ship is erected may become apparent. Figure 6*c* illustrates an advantage in the use of large erection lifts. The forward deckhouse (zone 17 in Fig. 3), which includes the living spaces, offices, and pilot house, was manufactured and outfitted complete with carpets, chairs, and curtains by a subcontractor and shipped by barge to the shipyard, where it is shown being placed aboard ship in a single 990-short-ton (900-metric-ton) lift.

To reduce the time required at the outfitting pier (the time period between launch and sea trials), the outfitting process is initiated as soon as feasible. For example, the final alignment of the propulsion machinery is accomplished as soon as all of the major structural welds are completed; this permits the propulsion machinery to be substantially ready to

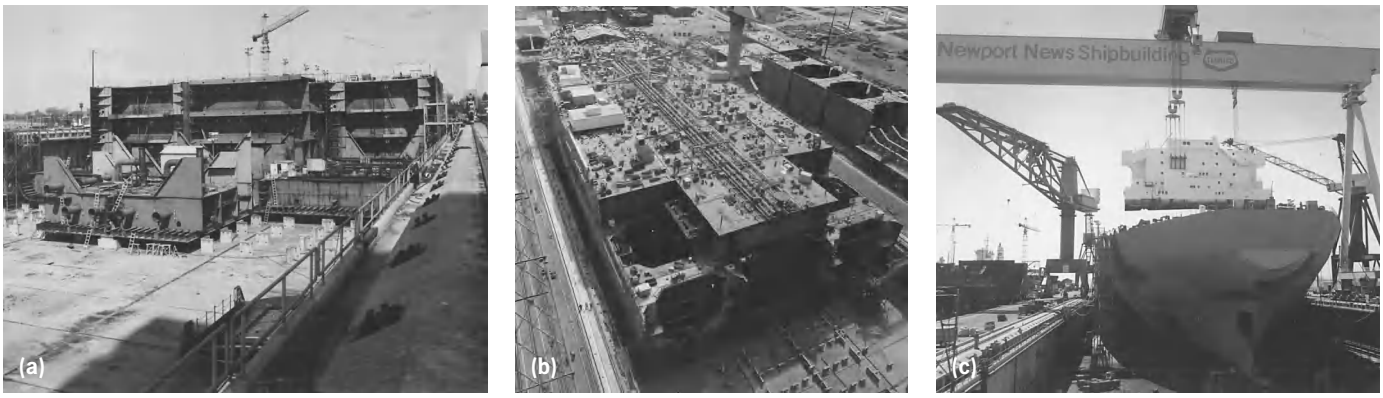

Fig. 6. Crude-oil tanker being constructed with preoutfitted erection units, showing (*a*) centerline erection unit, preoutfitted to contain the cargo-oil piping, (*b*) erection units being preoutfitted alongside the dock and aft of the ship, and (*c*) forward deckhouse being lowered into position. (*Newport News Shipbuilding*)

put to sea when the ship is launched. *See* MARINE MACHINERY.

Launching. The degree of completion that a ship has reached at the time of launching depends on many factors, such as the need for the shipway or dock being used for a subsequent ship, the relative ease of handling materials on the shipway or dock as compared with the outfitting pier, and the degree of completion of the boilers and other equipment which require cooling water for test purposes. However, in general, a ship is launched as soon as the hull structure is sufficiently complete to withstand the strain. Ships may be launched endwise, sidewise, or by in-place flotation. Endwise launchings were the most common method (and are still used) before the advent of side launchings (which are ideally suited for particular circumstances), and the more general use of flotation (for example, graving docks). The use of a graving dock requires a greater investment in facilities than either of the other two methods, but in some cases there may be an overall advantage due to the improved access to the ship and the simplified launch procedure. *See* DRYDOCKING.

Outfitting. The final outfitting of a ship is the construction phase during which checks are made to ensure that all of the previous work has been accomplished in a satisfactory manner; and last-minute details, such as deck coverings and the top coat of paint, are completed. It is considered good practice to subject as much of the ship as possible to an intensive series of tests while at the dock, where corrections and final adjustments are more easily made than when at sea. As a part of this test program, the main propulsion machinery is subjected to a dock trial, during which the ship is secured to the dock and the main propulsion machinery is operated up to the highest power level permissible. The full-power rating of the machinery cannot be developed during a dock trial because rated torque will not turn the propeller at the rated revolutions per minute with the ship restrained from forward motion. Also, for highly powered ships the strength of the pier may limit the power which can be developed. Still, many checks and tests can be made to ensure that the propulsion plant components and systems are correctly installed and function properly. For example, checks can be made to ensure that all relief valves are set at the correct pressure level and are functioning properly. Of particular concern are the extensive automation and remote control systems necessitated by the trend toward reduced crew levels.

When a comprehensive program of dockside tests has been completed, the only capabilities which have not been demonstrated are the operation of the steering gear during rated-power conditions and the operation of the main propulsion machinery at rated power; these capabilities must be demonstrated during trials at sea. However, the remaining equipment and systems can be operated at either actual or simulated rated conditions. For example, the anchor windlass cannot be operated with the anchor and the full scope of 165 fathoms (300 m) of chain extended at dockside; but the anchor and an added weight equivalent to the full scope of chain can be hoisted and lowered at dockside a sufficient number of times to demonstrate the operational readiness of the windlass.

After the ship has satisfactorily passed a comprehensive program of testing and has been determined to be seaworthy, it is ready to be taken on sea trials, followed by a delivery to the owners. *See* MERCHANT SHIP; NAVAL SURFACE SHIP; SUBMARINE.

R. L. Harrington

Bibliography. R. L. Harrington (ed.), *Marine Engineering*, 2d ed., 1992; E. V. Lewis (ed.), *Principles of Naval Architecture*, 3 vols., 2d ed., 1990; H. J. Pursey, *Merchant Ship Construction*, 1987; R. L. Storch et al., *Ship Production*, 2d ed., 1994; R. Taggart (ed.), *Ship Design and Construction*, 3d ed., 1980.

Shipping fever

A severe inflammation of the lungs (pneumonia) commonly seen in North American cattle after experiencing the stress of transport. This disease occurs mainly in 6–9-month-old beef calves transported to feedlots. Many infectious agents, both viruses and bacteria, cause symptoms of respiratory disease in

stressed cattle. However, the characteristic shipping fever pneumonia is caused primarily by the bacteria *Mannheimia* (formerly *Pasteurella*) *haemolytica* serotype A1 and, to a lesser extent, *Pasteurella multocida*; thus a synonym for shipping fever is bovine pneumonic pasteurellosis.

Both bacteria are normally present in low numbers in the nasal passages and tonsil area of unstressed cattle. However, with the stress of shipment and infection with respiratory viruses, these pathogenic bacteria replicate rapidly in the upper respiratory tract and are inhaled into the lungs, where pneumonia develops in the deepest region of the lower respiratory tract (pulmonary alveoli). Viruses can also concurrently damage pulmonary alveoli and enhance the bacterial pneumonia. *See* ANIMAL VIRUS; INFLAMMATION; PASTEURELLA.

Clinical signs. Symptoms of shipping fever occur 1–2 weeks after transport. Disease can occur in 50% or more of a group of cattle. Initially, cattle have reduced appetite, high fever, rapid and shallow respiration, depression, and a moist cough. During the later stages of disease, cattle lose weight and have labored breathing. A nasal and ocular discharge ranging from thin and watery to thick and yellow is usually present. Abnormal lung sounds are heard with a stethoscope. Without vigorous treatment, shipping fever can cause death in 5–30% of affected cattle.

The lesion causing these symptoms is a severe pneumonia accompanied by inflammation of the lining of the chest cavity, and is called a fibrinous pleuropneumonia. The chest cavity contains yellow, cloudy fluid, rich in plasma proteins (especially fibrin), and the lung surface is covered with yellow clots of fibrin. Large areas of the lungs become firm with red and gray mottling because pulmonary alveoli become filled with red blood cells, fibrin, and white blood cells. At autopsy, *Mannheimia haemolytica* can be cultured from affected lungs or pleural fluid. Other bacteria, notably *Mycoplasma bovis, Arcanobacterium* (formerly *Actinomyces*) *pyogenes*, and *Histophilus somni* (formerly *Haemophilus somnus*), are also isolated from cattle with shipping fever, especially in more chronic cases. *See* ACTINOBACILLUS; HAEMOPHILUS.

Virulence factors. *Mannheimia haemolytica* has at least three components, called virulence factors, that are responsible for causing pneumonia. The first is a bacterial cell wall–associated endotoxin that damages pulmonary alveolar capillaries, causing red blood cells, white blood cells, and plasma proteins (particularly fibrin) to escape from capillaries and fill the alveoli. The second is leukotoxin, released from the bacterium into alveoli. This toxin damages the cytoplasmic membranes of white blood cells that are normally present in the lung (pulmonary alveolar macrophages) and white blood cells that infiltrate the lung from the blood. Substances such as enzymes and altered oxygen molecules (free radicals) leak from damaged cells and injure alveolar lining cells (epithelium). Affected white blood cells eventually rupture. The third factor is a thick polysaccharide capsule that surrounds the bacterium and helps prevent the alveolar macrophages and infiltrating white blood cells from killing it. *Pasteurella multocida* also contains endotoxin and is surrounded by a thick capsule. *See* TOXIN.

Treatment and prevention. Treatment is aimed at eliminating bacteria, limiting the inflammatory reaction in the lung, and providing supportive care for cattle. To eliminate bacteria, diseased cattle are injected with antibacterial drugs, usually antibiotics. To reduce inflammation, nonsteroidal antiinflammatory drugs are given. These drugs reduce inflammation, pain, and fever but do not impair the immune response to the viruses and bacteria. Supportive care includes separating affected cattle from the stress of crowding, supplying adequate water, and feeding a ration high in roughage.

Vaccines can stimulate immunity to the viruses and bacteria associated with shipping fever. Although vaccination of cattle does not prevent shipping fever, it may reduce the number of cattle affected and the severity of the disease in individual animals. Viral vaccines should stimulate antibodies that neutralize viruses, by blocking their entry into cells and preventing their spread in the animal, thereby reducing susceptibility to bacterial infection. *Mannheimia haemolytica* vaccines should also stimulate production of two types of antibodies. The first prevents leukotoxin from damaging pulmonary alveolar macrophages and white blood cells, thereby decreasing damage to the lungs. The second attaches to the bacterial surface, allowing the bacteria to be more easily killed by pulmonary alveolar macrophages and white blood cells. *Pasteurella multocida* vaccines stimulate antibodies against the entire bacterium. Vaccines are often administered to cattle soon after shipment. This practice may be of some benefit; however, it does not allow enough time for cattle to develop an adequate antiviral and antibacterial antibody response before the onset of disease. As part of a disease control program called preconditioning, vaccines are given to cattle before shipment. This practice has the potential for better control of shipping fever. Those preconditioned cattle have antiviral and antibacterial antibodies before shipment, and they can better control viral and bacterial infections. *See* PNEUMONIA; RESPIRATORY SYSTEM DISORDERS; VACCINATION.

Anthony W. Confer

Bibliography. M. R. Ackermann and K. A. Brogden, Response of the ruminant respiratory tract to *Mannheimia (Pasteurella) haemolytica, Microbes Infect.*, 2:1079–1088, 2000; D. A. Mosier, Bacterial pneumonia, *Vet. Clin. North Amer. Food Anim. Prac.*, 13:483–493, 1997; B. P. Smith (ed.), *Large Animal Internal Medicine*, 3d. ed., 2001.

Shipworm

A wood-boring mollusk of the family Teredinidae. These borers, also called pileworms, are found around the world in tropical to boreal and marine to nearly fresh-water habitats, from the intertidal to depths of about 2000 ft (600 m). They arose in the

Cretaceous, and since then have been the most important organisms involved in recycling woody plant material entering the sea. It was not until humans became navigators that shipworms became pests. Before the advent of steel ships, they were greatly feared, and even today hundreds of millions of dollars are spent annually on prevention and repair due to their activity.

Teredinids are specialized to bore in woody plant material and utilize it as food. They are elongate and wormlike, with a small shell anteriorly and with siphons and pallets posteriorly (**Figs. 1** and **2**). Adult shipworms vary in length from about 0.8 in. to over 6.5 ft (2 cm to 2 m). There are 15 genera and about 70 species in this variable, specifically adapted, family. They bore by scraping filelike ridges on the anterior part of the shell against the anterior end of the burrow. The posterior or outer end of the burrow has a calcareous lining to which the muscles of the siphons and pallets are permanently attached. If conditions are adverse, the borer can withdraw the siphons and close the burrow entrance with the pallets. In this way they can survive for weeks under anaerobic conditions. *See* FEEDING MECHANISMS (INVERTEBRATE).

Fig. 2. Living *Lyrodus floridanus*.

Reproduction in the teredinids is similar to that in other bivalves. Fertilization of the egg may take place in the sea (oviparous species), or the eggs may be fertilized as they are deposited in the gills of the female, where they develop to the straight-hinge or first feeding stage (short-term brooders) or until they reach the pediveliger or settling stage (long-term brooders). The length of the free-swimming period is inversely proportional to the length of the brooding period. Tropical species may reproduce throughout the year, but most temperate and boreal species breed during the summer. At settlement the larvae are almost microscopic (200–350 micrometers); they rapidly penetrate the softened wood surface, complete metamorphosis, commence boring, and for the remainder of their life are confined to their burrow. *See* BIVALVIA. Ruth D. Turner

Bibliography. E. B. G. Jones and S. Eltringham (eds.), *Marine Borers, Fungi and Fouling Organisms of Wood*, Organization for Economic Cooperation and Development, Paris, 1971; N. B. Nair and M. Saraswathy, The biology of wood-boring teredinid molluscs, *Advan. Mar. Biol.*, 9:335–509, 1971; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; R. D. Turner, *A Survey and Illustrated Catalogue of the Teredinidae*, Spec. Publ. Mus. Compar. Zool., Harvard Univ., 1966.

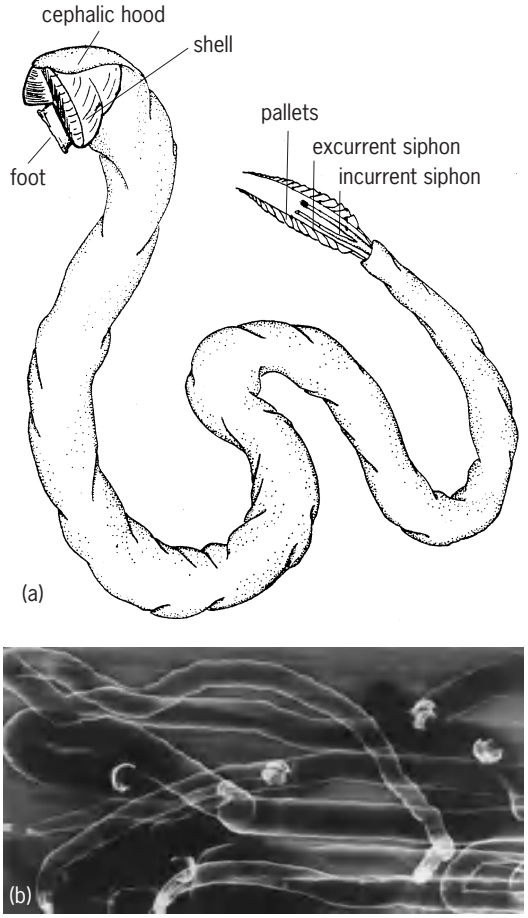

Fig. 1. *Bankia gouldi*. (*a*) Diagram showing position of shells, foot, pallets, and siphons (*after R. D. Turner, A Survey and Illustrated Catalogue of the Teredinidae, Spec. Publ. Mus. Comp. Zool., Harvard Univ., 1966*). (*b*) X-ray of collecting panel exposed for 6 months showing burrows with pallets and shells in place.

Shock absorber

Effectively a spring, a dashpot, or a combination of the two, arranged to minimize the acceleration of the mass of a mechanism or portion thereof with respect to its frame or support.

The spring type of shock absorber (**Fig. 1**) is generally used to protect delicate mechanisms, such as instruments, from direct impact or instantaneously applied loads. Such springs are often made of rubber or similar elastic material. Their design in relation to the natural frequency of the supported system and the forcing frequency of the applied load is most important. *See* SHOCK ISOLATION; SPRING (MACHINES).

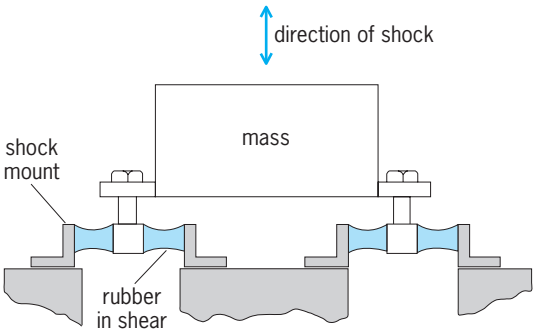

Fig. 1. Spring-type shock absorber.

The dashpot type of shock absorber is best illustrated by the direct-acting shock absorber in an automotive spring suspension system (**Fig. 2**). Here the device is used to dampen and control a spring movement. The energy of the mass in motion is converted to heat by forcing a fluid through a restriction, and the heat is dissipated by radiation and conduction from the shock absorber. *See* AUTOMOTIVE SUSPENSION; VIBRATION DAMPING.

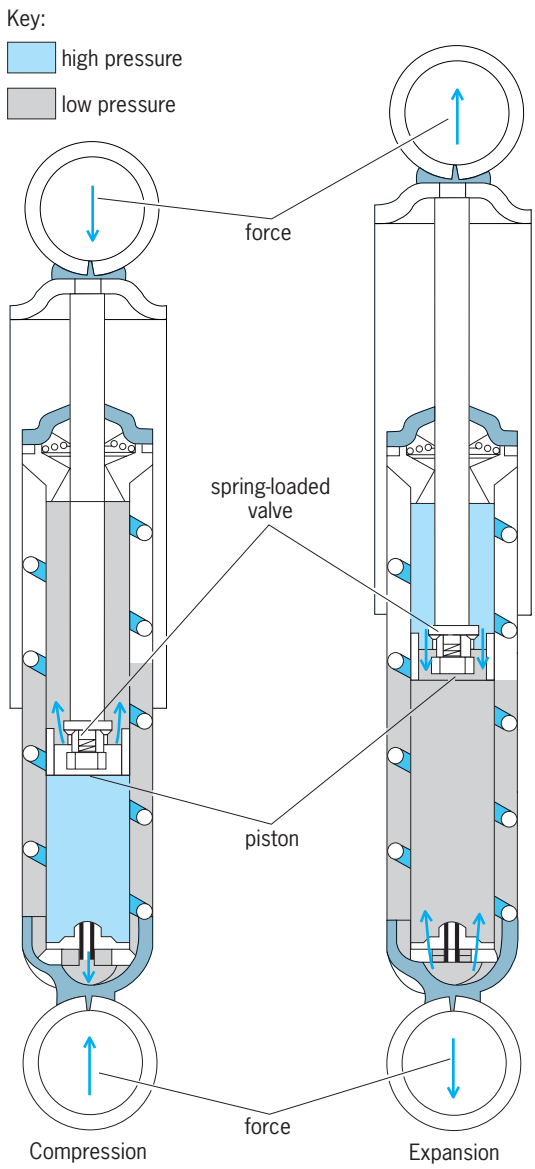

Fig. 2. Dashpot-type shock absorber. (*Plymouth Division, Chrysler Corp.*)

There are also devices available which combine springs and viscous damping (dashpots) in the same unit. They use elastic solids (such as rubber or metal), compressed gas (usually nitrogen), or both for the spring. A flat-viscosity hydraulic fluid is used for the viscous damping. L. Sigfred Linderoth, Jr.

Shock isolation

The application of isolators to alleviate the effects of shock on a mechanical device or system. Although the term shock has no universally accepted definition in engineering, it generally denotes suddenness, either in the application of a force or in the inception of a motion. *See* SHOCK WAVE.

Shock isolation is accomplished by storing energy in a resilient medium (isolator, cushion, and so on) and releasing it at a slower rate. The effectiveness of an isolator depends upon the duration of the shock impact. An isolator may be effective in one case where there is a high G loading with a short duration, 0.001 millisecond or less, but may magnify the shock where there is a lower G loading but a longer duration (0.001–0.015 ms). The quantity G is equal to the so-called limit acceleration a divided by the acceleration of gravity g and is discussed later. Most shock isolators, also known as shock mounts or shock absorbers, that are available commercially are effective for the 0.001-ms or less interval.

Rubber is the most common material used in commercial shock isolators. Rubber isolators are generally used where the shock forces are created through small displacements. For larger displacement shock forces, such as those experienced by shipping containers in rough handling conditions, thick cushions of felt, rubberized hair, sponge rubber, cork, or foam plastics are used. Shock isolation systems which use the various cushion materials are generally custom designed to the particular application and cannot be considered from the standpoint of standardized isolators, but rather from the standpoint of the basic principles involved. *See* RUBBER.

Absorption of shock. The shock load must be divided between the case, the shock cushion, and the equipment. The case, since it must withstand effects of rough handling such as sliding and dropping, is by necessity rigid. The more rigid the case the closer to a 1:1 ratio will be the transfer of the shock from outside to inside. The absorption of the shock is primarily between the cushion and the equipment.

The dissipation of the energy of a 1-ft (0.3-m) drop with a cushion having a linear spring rate would require a thickness of 2 ft (0.6 m) of cushion. Since cushions of such thickness are not feasible, the equipment itself must withstand part of the shock. The cushion that is needed to dissipate the energy

from various heights of drop with equipment sharing part of the load may be determined by consideration of the principles which are involved. *See* SHOCK ABSORBER.

Limit acceleration. When a body moving with velocity v has to be stopped to complete rest, a deceleration (negative acceleration) must be applied. In order to make the stopping process smooth, a maximum value for the acceleration a is prescribed as a limit. Usually the full amount of the limit acceleration cannot be attained for the entire duration of the stopping process. However, an ideal process can be imagined with constant limit acceleration a. Deviation from this ideal case will be considered later.

Such uniformly decelerated motion is exactly the reverse process of uniformly accelerated motion. The same formulas apply.

The total distance traveled S is given by Eq. (1),

$$S = \frac{v^2}{2a} \tag{1}$$

where v is the initial and final velocity. When the velocity to be stopped is produced by a free fall, the necessary height of drop H must be as given by Eq. (2), where g is the gravitational acceleration. From these formulas Eq. (3) is obtained. Since the

$$H = \frac{v^2}{2g} \tag{2}$$

$$v^2 = 2aS = 2gH \tag{3}$$

velocity at the end of the free fall is the same as it was at the beginning of thestopping process, Eq. (4)

$$2aS = 2gH \tag{4}$$

is obtained. It is customary to give the limit acceleration in the form of a G value, defined by $G = a/g$. With this notation, the stopping distance can be written in the form given in Eqs. (5).

$$S = H\frac{g}{a} \qquad S = \frac{H}{G} \tag{5}$$

Force-distance diagrams. It is useful to analyze the ideal case again from another aspect. The energy E stored in the falling body after fall from height H is $E = WH$, where W is the weight of the moving body. The same amount of energy must be consumed during the stopping period.Therefore, Eq. (6) is ob-

$$E = GWS \tag{6}$$

tained. The active force is GW in this stopping phase. Equating the last two equations for the energy yields Eq. (1). As represented in **Fig. 1**, the energy can be visualized as the area under the curve in a force-distance diagram. The curve is a horizontal line in the case of constant deceleration.

Whatever device is used for checking the velocity, it is hardly possible to obtain a constant deceleration

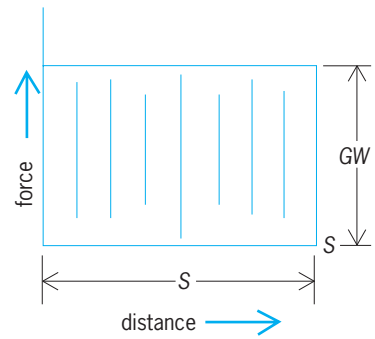

Fig. 1. Force-distance diagram for ideal case.

of exactly the limit value. Therefore, allowances have to be made for practical considerations.

If the opposing force is provided by an ordinary spring, no constant force is produced. The force F is built up gradually with the compression of the spring, as shown in **Fig. 2**. The area under the force-displacement curve is in this case a triangle, representing the energy given in Eq. (7). The kinetic

$$E = {}^1/_2 GWS_1 \tag{7}$$

energy from the fall is the same as before, $E = WH$. Equation (8), the stopping distance for the spring of

$$S_1 = \frac{2H}{G} \tag{8}$$

Fig. 2, is found by equating Eqs. (6) and (7). This means a distance twice as long as that for the ideal case of Fig. 1 is needed. This less favorable condition is caused by the fact that the permitted maximum force is used only at the end of the process. The diagram of Fig. 2 is not '"filled up" completely; it is only half-filled.

Preloaded springs have a better-filled diagram, as shown in **Fig. 3**. An inherent disadvantage of the

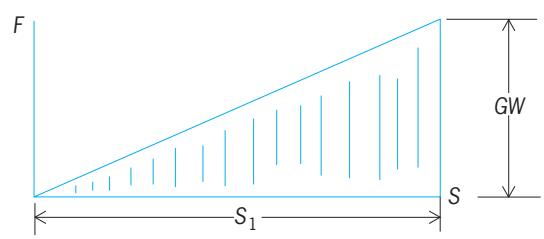

Fig. 2. Force-distance diagram for ordinary spring.

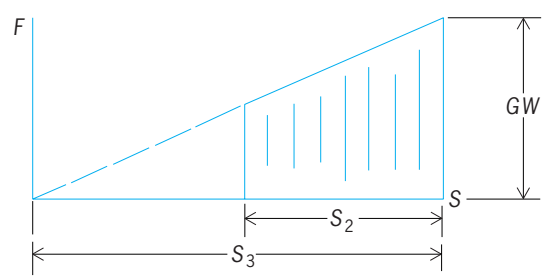

Fig. 3. Force-distance diagram for preloaded spring.

better-filled-up spring diagrams is the steep slope at the beginning, which means that the spring system must be designed for a predetermined shock. This design does not help in isolating small shocks. It has not been proved which shocks do the most damage—the large shock that happens occasionally or the small, repeated shock that occurs almost continuously. From the isolation standpoint, the spring characteristic of an isolator should start up with a moderate slope and continue with gradually sharper increase of force. This is why a system representing a completely filled diagram cannot be applied, and therefore a certain deviation from the ideal condition represented by Eq. (1) is unavoidable.

Damping forces are helpful for attaining longer deceleration force at the beginning of the stopping process. Pure viscous damping alone would result in an elliptical diagram like the one shown in **Fig. 4**.

Since the high velocity at the beginning produces large opposing forces, spring force and damping effects can be combined to make a well-filled diagram, as shown in **Fig. 5**.

Practical shock absorbers usually fall between the two cases of completely filled diagrams and half-filled diagrams. The stopping distance provided therefore lies between Eqs. (5) and (8). The value given by Eq. (8) is considered conservative and therefore is recommended, since a certain tolerance is necessary.

Using Eq. (8) and assuming that the equipment can withstand 30*G*, the operating deflection of the cushion under a 24-in. (60-cm) drop must be as given by Eq. (9). Since *S* is the operating distance, the total

$$S = \frac{2 \times 24}{30} = 1.6 \text{ in. } (4 \text{ cm}) \qquad (9)$$

thickness of cushion will depend on the compression ratio of the material.

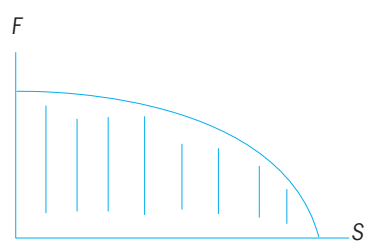

Fig. 4. Force-distance diagram for pure viscous damping.

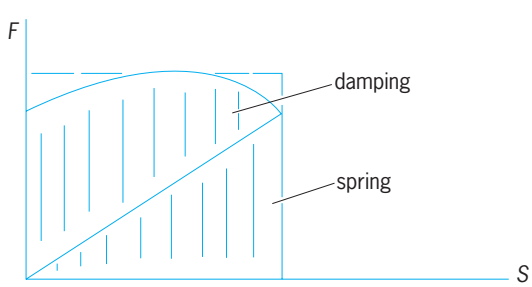

Fig. 5. Force-distance diagram for combined effects of spring force and damping.

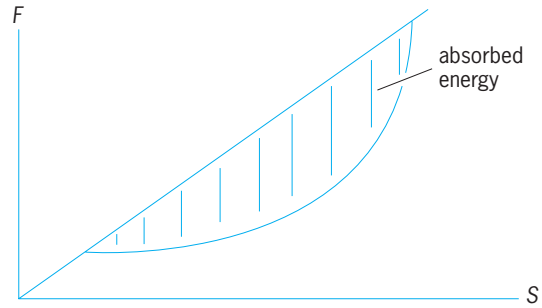

Fig. 6. Diagram of absorbed energy for ideal spring system with damping.

The preceding equation provides the operating thickness needed to bring a given mass to rest with a linear spring cushion, leaving the equipment to share 30*G* of the impact load. It must be remembered that cushions, springs, and so forth store energy and, depending upon their inherent friction, return this energy. Thus, in an ideal spring system, no energy is absorbed. However, by adding damping or by using a material with inherent damping, there is provided a resilient medium (**Fig. 6**).

To accomplish such a system requires engineering information on the cushioning materials under dynamic conditions. Once their behavior is known, the materials that have the desired force-displacement characteristics can be selected, and the damping forces may be added to provide a shock-absorbing system approaching the effectiveness of the one shown in Fig. 6. *See* DAMPING; SPRING (MACHINES); VIBRATION DAMPING. K. W. Johnson

Bibliography. C. M. Harris and C. E. Crede (eds.), *Shock and Vibration Control Handbook*, 4th ed., 1995.

Shock syndrome

A clinical condition caused by widespread impairment of blood flow resulting in an inadequate supply of oxygen to the body's tissues. Shock is a sudden and progressive condition that, if not reversed, can lead to death. Shock affects the entire body, as opposed to other conditions where blood flow is reduced to only a portion of a single organ, such as stroke or heart attack. It is typically, but not always, associated with a low blood pressure that can be felt as a weak pulse.

Shock can be categorized according to the initiating event: hypovolemic (loss of circulating blood volume), cardiogenic (loss of cardiac pumping activity), distributive (shunting of blood flow away from the tissues), or obstructive (a blockage of blood flow through the heart or major vessels).

Regardless of the initiating event, shock will incite responses from the body. These responses will often maintain blood flow to vital organs for a period of time but at the expense of metabolic stress and exhaustion in the tissues. The nervous system reacts to shock by sending signals to the blood vessels in

nonvital organs (such as the skin and bowel) that constrict these blood vessels, reduce blood flow to those organs, and allow blood to go preferentially to the vital organs (heart and brain). The nervous system also increases the heart rate in an attempt to pump more blood.

As shock progresses, the metabolically active tissues are starved for oxygen (a condition termed ischemia) and change their metabolic processes from those that use oxygen (aerobic) to those that do not require it (anaerobic). Anaerobic metabolism is a very inefficient process. Once cells are unable to maintain basal metabolic activities, the cell dies and releases harmful intracellular substances.

Advanced shock is generally associated with impaired cardiac contractility (even if this was not the initial cause) and slugging of blood flow in the microcapillaries. By this time, shock has progressed to the degree that, despite reversal of the inciting process, the impaired tissues do not recover and death results.

Hypovolemic shock. Hypovolemic shock occurs when the circulating blood volume (approximately 70 to 80 mL/kg of body weight or 5 to 6 L in an adult human) is reduced below a critical level. The most common cause of blood loss (termed hemorrhage) occurs when a hole develops in the wall of a major blood vessel and blood escapes either outside the body (external hemorrhage) or to spaces inside the body (internal hemorrhage). The amount of sudden blood loss necessary to produce shock is around 40% of the total blood volume, although this threshold varies according to the individual's physiologic condition and the rapidity of loss.

Blood is approximately 40% red blood cells and 60% plasma, a mixture of water, salt, and proteins. Shock is possible if large amounts of water are lost, as might occur from severe vomiting or diarrhea.

Hypovolemic shock is classically associated with cool, pale skin (due to reduced blood flow to the skin), sweating (due to stimulation of the nervous system), and a fast heart rate (an attempt to pump more blood).

The treatment of hypovolemic shock consists of rapid restoration of circulating blood volume, initially with an isotonic solution of water and salt (termed crystalloid), followed by replacement of lost red blood cells through blood transfusions.

Cardiogenic shock. Cardiogenic shock results when the pumping activity of the heart is suddenly reduced. The main pumping chamber of the heart is the left ventricle. Cardiogenic shock becomes common when approximately 40% of the left ventricular heart muscle loses its ability to contract. The most common reason for cardiogenic shock is a sudden blockage (occlusion) of blood flow in a major artery (coronary artery) suppling blood to the working heart muscle. Even though the heart pumps blood through its four chambers, the heart muscle is, like every other tissue, dependent on blood vessels that course within the muscle itself for its oxygen supply.

A sudden occlusion of a coronary artery (heart attack) is usually caused by a blood clot forming on the site of a long-standing narrowing due to the deposition of cholesterol underneath the inner lining of the artery. Cardiogenic shock is usually associated with pain from the heart attack as well as the previously mentioned signs of shock.

The treatment of cardiogenic shock involves restoration of blood flow through the occluded coronary artery (reperfusion). Reperfusion is possible using drugs which dissolve the occluding blood clot (termed thrombolytics), by mechanically breaking up the blood clot that occludes or narrows the artery using catheters advanced through the circulatory system, or by surgically establishing a route for blood flow around the occluding blood clot (termed coronary artery bypass grafting).

Distributive shock. Distributive shock results from widespread impairment of blood flow to the tissues despite adequate circulating blood volume and cardiac pumping activity. This may appear paradoxical, but in reality it is blood flow though the capillaries that is most important. Only in the capillaries is oxygen delivered to the metabolically active cells. If blood flow is shunted away from the capillaries, then oxygen does not reach the cells and shock may develop. Portions of the nervous system regulate the distribution of blood flow and can finely adjust blood flow to balance the metabolic needs of local tissues.

There are several conditions that lead to distributive shock. A major impairment of neurologic regulation, termed neurogenic shock, is one. Nerve signals pass from the brain, down the spinal cord, and out into the peripheral nerves. Damage to the spinal cord will impair the transmission of these regulatory signals, with the greatest effect if the damage is toward the upper portion of the spinal cord. Thus, patients who sustain spinal cord injury, usually in the neck region, are at risk for developing neurogenic shock.

Infections can produce distributive shock from the toxic substances elaborated by the infecting organisms and by the body's response to the infection. Bacteria produce some of the more severe infections and stimulate the body to produce several chemical substances that impair circulation. The body's response to infection is called sepsis, and if shock develops the process is called septic shock.

Severe allergic reactions can produce widespread improper distribution of blood flow, a process termed anaphylaxis or anaphylactic shock. Such severe allergic reactions may result from insect stings (such as by bees and wasps), foods (such as nuts and shellfish), or medications (for example, penicillin).

Common clinical characteristics of distributive shock are the tendency for skin perfusion to be maintained initially (warm, pink, even flushed) and an absence of sweating. Patients with neurogenic shock have impaired innervation to the heart, so the heart rate is typically low or normal. Patients in septic shock often have a fever. Patients in anaphylactic

shock may have other signs of allergic reaction (such as hives or wheezing).

The initial treatment of distributive shock is the expansion of circulating volume with isotonic crystalloid. The amounts required to restore perfusion may be large: 20 to 40 mL/kg or 2 to 4 L in a typical adult. Many patients will require the administration of drugs that stimulate the cardiovascular system to increase blood flow and improve tissue perfusion.

Obstructive shock. Obstructive shock results from a mechanical blockage of blood flow through the heart or major blood vessels in the chest. For example, a blood clot can form in a leg vein, break off, and travel with the blood from the leg to the heart. It will typically pass through the heart and its valves but become lodged in the smaller pulmonary artery that leads from the heart to the lungs. If the obstruction is large enough, blood flow from the heart is impaired and shock develops.

Other mechanical conditions that may occlude blood flow are the presence of a constriction around the heart that limits pumping action (pericardial tamponade) and blockage of blood flow out of the heart into the major artery of the body (aortic dissection).

Patients with obstructive shock have typical features of shock (cool pale skin, sweating, and fast heart rate) in addition to features specific to the underlying event. The treatment of distributive shock is removal of the mechanical blockage to blood flow, which usually requires a surgical procedure. *See* BLOOD; CARDIOVASCULAR SYSTEM; RESPIRATION.

J. Stephan Stapczynski

Bibliography. D. Annane, E. Bellissant, and J. M. Cavaillon, Septic shock, *Lancet*, 365(9453):63–78, January 1, 2005; C. A. Graham and T. R. J. Parke, Critical care in the emergency department: Shock and circulatory support, *Emerg. Med. J.*, 22(1):17–21, January 2005; F. A. Moore, B. A. McKinley, and E. E. Moore, The next generation in shock resuscitation, *Lancet*, 363(9425):1988–1996, June 12, 2004; M. Mullner et al., Vasopressors for shock, *Cochrane Database Syst. Rev.*, 3:CD003709, 2004; T. A. Sanborn and T. Feldman, Management strategies for cardiogenic shock, *Curr. Opin. Cardiol.*, 91(6):608–612, November 2004.

Shock tube

A laboratory device for rapidly raising confined samples of fluids (primarily gases) to preselected high temperatures and densities. This is accomplished by a shock wave, generated when a partition that separates a low-pressure (driven) section from a high-pressure (driver) section is rapidly removed. In most shock tubes the two sections are separated by a diaphragm that supports the pressure differential until it is punctured (**Fig. 1***a*). Shock tubes operated by rapidly opening valves also have been developed. The tubes can be circular, square, or rectangular in cross section, with diameters that range from 0.8 to 24 in. (2 to 60 cm). The lowest gas pressure to be used in the driven section is such that the corresponding molecular mean free path is less than several hundreds of the diameter of the driven tube; the length of that section should be 40–60 times its diameter. Driver lengths vary, depending on the designed test time and the desired ultimate pressure in the shock tube. Concurrently with propagation of the high-pressure (heating) wave from the partition toward one end wall (Fig. 1*b*), an expansion (cooling) wave propagates in the opposite direction. Although shock waves can be generated under laboratory conditions by means of a variety of devices, such as ballistic pistons, electrical discharges, distributed explosives, and the impact of hypersonic flows on stationary targets, shock tubes are the simplest to construct and to operate.

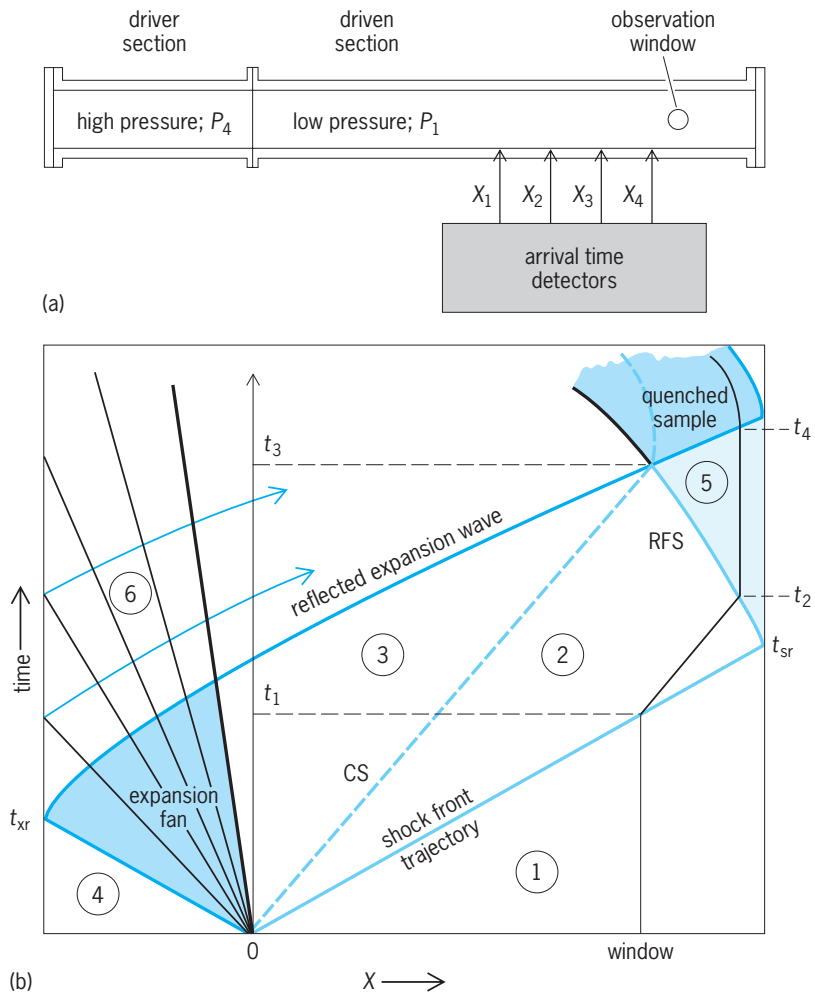

Fig. 1. Shock tube. (*a*) Schematic. (*b*) Flows in a shock tube. The special conditions in a single-pulse tube are shown, and the circled numbers identify the different regions.

In the driven section, gaseous samples can be heated to temperatures as high as 26,500°F (15,000 K) under strictly homogeneous conditions. The material in any tube cross section is thus rapidly and uniformly heated, except for a very thin layer adjacent to the walls, for a duration that depends

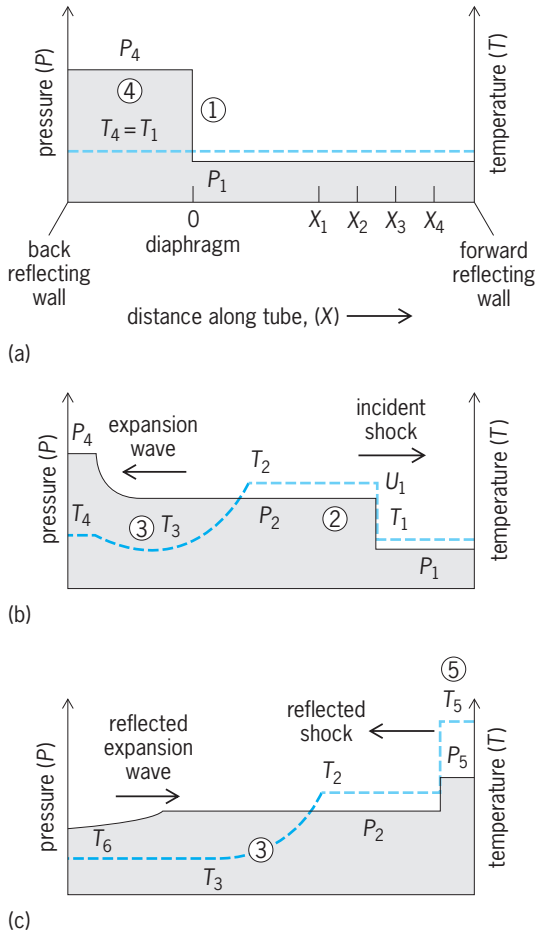

Fig. 2. Schematic profiles of pressure and temperatures along the shock tube at selected times (*t*). Circled numbers (regions) correspond to those shown in Fig. 1*b*. (*a*) At $t = 0$. (*b*) At $t = t_1$. (*c*) At $t \geq t_{sr}$.

on its location along the tube. At the shock front the heating rate is about 180°F/microsecond (100 K/μs), and the transition from the unshocked to the high-temperature condition is of short but of finite duration; with high-resolution diagnostic devices it is possible to separate, in time, sequences of relaxations. The incident shock wave is reflected at the tube terminus (t_{sr}), generating a slower-moving wave (RFS) that additionally heats the compressed gas. The elevated temperature can be maintained for times of the order of milliseconds. The expansion wave is also reflected at the (left) terminus (t_{xr}). Upon convergence of the two reflected waves (t_3), the processed material is quenched at cooling rates of 0.9–9.0°F/μs (0.5–5 K/μs). On the high-pressure side the expansion wave never attains a steep front (**Fig. 2**); the cooling rate is both time- and position-dependent. *See* SHOCK WAVE.

Operation. With the diaphragm in place, both sections of the shock tube are evacuated; in particular, the driven section must be well pumped to minimize contamination of the sample by residual gases and material desorbed from the walls. That section is then filled to a modest pressure [$P \approx 0.01$–1.0 atm or 10^3–10^5 pascals at initial temperature (T_1)] with the mixture under investigation. Most often this mixture consists of 90–99.9% argon (or a heavier inert gas), plus the gas of interest. The driver section is filled to a high pressure ($P_4 \approx 10$–100 atm or 10^6–10^7 Pa) with a low-molecular-weight gas: helium or hydrogen. When the diaphragm is rapidly ruptured, expansion of the high-pressure gas acts as a low-mass piston that generates a steepening pressure front, which moves ahead of the boundary between the driver and the test gases (trajectory CS, Fig. 1*b*). A planar shock develops in 6–10 tube diameters: it moves at a hypersonic speed (U_1), determined by the pressure ratio (P_4/P_1) and the heat capacity ratio [$\gamma_{\text{test}}/\gamma_{\text{driver}}$: γ is the ratio of the specific heat at constant pressure (C_p) to the specific heat at constant volume (C_v)]. The shock speed is measured by recording its arrival times at several precisely located stations ($X_1, X_2, \ldots$); the transition from P_1 to P_2 occurs over a distance equivalent to about five mean free paths, measured in the unshocked gas: 10^{-8} to 10^{-7} cm for $P_1 \approx 10$ torr. The test sample is raised to temperature $= T_2$, density $= \rho_2$, and pressure $= P_2$, thereby initiating vibrational or electronic excitation, isomerization, bond scission, ionization, or bimolecular reactions.

Shock tubes are unique reactors: on the one hand, they may be used to investigate the gas dynamics of shocks; on the other hand, they may be used for the preparation of test samples, for equilibrium or kinetic study of the conversion of materials at elevated temperatures (or at low temperatures in the driver section). Multifold diagnostic devices are in use for tracking the changes of state that occur in the processed materials with microsecond resolution. Characteristic spectral line sources may be used to chart the evolution of specific absorbing species on a state-by-state basis. Rapid response detectors, photovoltaic and photoconducting elements, as well as photoelectron multipliers, coupled to fast digitizers and computers, permit the concurrent recording of vast amounts of data for each shock. Appropriate spectrographs and detectors permit measurement of intensities and wavelengths of the emitted light. With selected frequency sources, particularly lasers, specific absorption spectra and changes in refractive index can be recorded. Spectrometers equipped with parallel-array detectors permit the acqusition of time-dependent spectra at numerous wavelengths concurrently. In some experiments the shock-heated gas is expanded, through a tiny aperture at the terminus of the test section, into a rapid-scan mass spectrometer for gas analysis. In a single-pulse shock tube a ballast tank is attached to the tube, close to the diaphragm on the low-pressure side, with an aperture comparable to the tube diameter. This greatly attenuates subsequent reflected shock waves and permits single-pulse processing of the driven gas from region 5 (Fig. 1*b*). A sample of the heated and subsequently quenched material may then be extracted for extended chemical analysis. *See* MASS SPECTROMETRY.

Thermodynamic processes. The shock wave carries energy and momentum. Hence, when it reaches the end plate it is reflected as another shock wave, now propagating into the previously heated gas. Additional heating (to T_5) and compression (to ρ_5) occur. At the observation window, the incident shock front is seen first at t_1, followed by samples that were exposed for various times to T_2. The sample of gas initially at the observation point is swept toward the end plate. It is exposed to T_2 for the period $t_1 \rightarrow t_2$; then to T_5 for the period $t_2 \rightarrow t_4$, when it is quenched by the rarefaction wave that met the reflected shock wave at t_3. In a properly designed shock tube, such samples may be extracted for detailed chemical analysis, since they have been exposed to a single high-temperature pulse of known magnitude (T_5) and duration ($\tau = t_4 - t_2$). The production of a large number of chemical species at high temperatures, integrated over τ, can thus be ascertained. Figure 2 shows schematic profiles of pressures and temperatures along the shock tube at selected times.

The thermodynamic state of the shock-heated sample is determined by the equation of state of the test gas (over the full temperature range); the conservation conditions of mass, momentum, and energy for the material that passes through the shock front; and the velocity of the shock wave, which is equal but opposite to the mass velocity with which molecules enter the front. The measured shock speed is usually expressed as a dimensionless Mach number ($M_1 = U_1/\alpha_1$); $\alpha_1 = (\gamma\ P_1/\rho_1)^{1/2}$ is the velocity of sound in the unshocked test gas. The pressure jump across the shock front (P_2/P_1) is proportional to M_1, as in Eq. (1). For ideal gases, the ratios for density (ρ) and temperature (T) are given by Eqs. (2) and (3). The higher the value of γ, the larger these ratios.

$$\frac{P_2}{P_1} = \frac{2\gamma M_1^2 - (\gamma - 1)}{(\gamma + 1)} \tag{1}$$

$$\frac{\rho_2}{\rho_1} = \frac{(\gamma + 1)M_1^2}{2 + (\gamma - 1)M_1^2} \tag{2}$$

$$\frac{T_2}{T_1} = \frac{\left[2\gamma M_1^2 - (\gamma - 1)\right]\left[2 + (\gamma - 1)M_1^2\right]}{(\gamma + 1)^2 M_1^2} \tag{3}$$

The Mach number M_1 increases with the initial pressure ratio across the diaphragm (P_4/P_1) and with the sound speed of the driver gas (α_4). Thus, $M_1 = 2.8$ for argon driving into argon, when $P_4/P_1 = 10^3$. The same ratio develops $M_1 = 6$ for helium into argon, and $M_1 = 8.3$ when hydrogen drives argon. Since the enthalpy and heat capacity of the material of interest progressively change during relaxation to equilibrium at T_2, argon, which is an inert diluent and has a high γ (1.67), is used as the major component in the test gas. Then, the drift in the parameters of state due to the presence of 0.1–1.0% of the component of interest is minimized. *See* ENTHALPY; FLUID FLOW.

Applications. The many gas dynamic problems that can be investigated in shock tubes include thermal boundary-layer growth, shock bifurcation, and shock-wave focusing and reflection from plane and wedge-shaped surfaces. However, shock tubes have their greatest utility as devices for preparing gases for study at very low or very high temperatures without thermal contact with extraneous surfaces. For example, if condensable vapors are commixed with the driver gas, upon expansion (Figs. 1 and 2) they attain various degrees of supersaturation prior to homogeneous condensation. Nonequilibrium condensation of methanol vapor has thus been measured at the shock-tube end wall behind reflected expansion waves.

The optical and spectroscopic parameters of samples that attain thermodynamic equilibrium after the passage through the shock front can be measured: for example, line intensities in emission and absorption, effects of collisional broadening and the presence of ions on spectral line contours, indices of refraction, and microwave attenuation. Conversely, when the indices of refraction or the spectroscopic parameters of the various species suspected of being present in the shock-heated gases are known (or can be reliably estimated), their measurement provides information on the composition and hence the thermodynamic state of the gas.

Shock-tube techniques are particularly useful for studying the sequence of states through which a sample passes after its temperature is precipitously raised from T_1 to T_2 (or from T_2 to T_5). The evolution in time of the post-shock-front samples can be followed to address a variety of questions, ranging from simple ones (such as how long does it take for the initial high-translational, low-vibrational energy partition to equilibrate to a uniform temperature) to very complex problems [such as what is the kinetic mechanism that determines ignition delays in a combustible mixture of fuel and oxidizer; specifically, the effect of functional groups on ignition delays in H_3CX/O_2 mixtures (X = CH_3, CH_2, NO, CN, Br) was found to correlate with the C-X bond dissociation energy].

The gas density in the vicinity of the shock front, $\rho(X)$, can be scanned by recording, interferometrically, the profile of the index of refraction. A more sensitive measurement may be obtained from the gradient of the index of refraction. Then, from the measured shock speed (U_1) the time-dependent change in the density immediately past the shock front can be derived. This reflects the change in the local temperature, and consequently the shuffling of enthalpy from the translational to the internal molecular degrees of freedom.

One of the first significant applications of shock tubes to chemical kinetics was the determination of diatom dissociation rates; this was followed by studies of small, then larger, polyatomic molecules. Such processes could be investigated only with difficulty in conventional high-temperature reactors, either static or flow type, because of unavoidable

concurrent complex reactions that take place at hot reactor walls. Isomerization reactions constitute another simple type of conversion; for example, cyclopropane → propene; *cis*-butene-2 → *trans*-butene-2. For these the measured rates fit well with values extrapolated from kinetic parameters obtained at lower temperatures. *See* MOLECULAR ISOMERISM; MOLECULAR STRUCTURE AND SPECTRA.

It is possible to generate mixtures of highly supersaturated metal vapors by shock-heating volatile precursors diluted with argon (Ar) [iron from $Fe(CO)_5$; bismuth from $Bi(CH_3)_3$; silicon from SiH_4]. Initial precursor pressures are selected such that complete decomposition occurs rapidly in the postshock region (T_5), thus generating a highly supersaturated mixture of metal/Ar species. Then condensation occurs to produce minute metallic clusters (up to 10–30-nanometer radius). The kinetics of such processes can be followed by recording the intensity of scattered laser light and sample turbidity. The results of these experiments proved to be inconsistent with classical nucleation theory.

The use of shock tubes as a research tool for investigations of gas dynamics, determination of parameters of state, spectroscopy, and chemical kinetics—at high temperatures—is detailed in an ongoing series, Proceedings of Shock Tube Symposia, held biannually. Aspects that focus on chemical kinetics have been summarized in two reviews (1990 and 1993). Therein are listed over 350 specific reactions for which temperature-dependent rate constants were measured over 800–2500 K; 600 references to the literature are included. In one review, by W. Tsang and A. Lifshitz (1990), chemical kinetics studies were classified into five categories, based on the diagnostic devices that were used. In one of the major techniques the rates of production (or destruction) of atomic species were measured with microsecond resolution, by recording their atomic resonance absorptions, which is both highly specific and very sensitive; for example, notation (4).

$$\begin{gathered} \underline{H} + CH_4 \longrightarrow H_2 + CH_3 \\ k = 10^{14.9} \exp(-7577/T) \\ CH_3 + Ar \longrightarrow CH_2 + \underline{H} \\ k = 10^{16.3} \exp(-46{,}100/T) \end{gathered} \tag{4}$$

Broad-band absorptions and emissions have also been used to follow the evolution (or decay) of molecules or radicals, such as notation (5).;

$$\begin{gathered} \underline{CH_3I} \rightarrow CF_3 + I \\ k = 10^{14.67} \exp(-27{,}200/T) \end{gathered} \tag{5}$$

The "single pulse" configuration that develops a rapid quench of the shock-heated gas (Fig. 1*b*) has been extensively exploited for investigations of fragmentation sequences of large molecules at high temperatures. Free-radical precursors have been pyrolized to generate initially the desired free radicals, and their reactions with commixed reagents followed either spectroscopically or by analysis of the distribution of products in the quenched samples. Kinetic parameters for many reactions, derived via three additional diagnostic procedures, were also listed: use of laser-schlieren (high-resolution laser diagnostic technique) for measuring density gradients at shock fronts (to determine enthalpy flow among the heated gas components), mass-spectrometric sampling of the shock-processed gas, and determination of induction periods for the onset of combustion when mixtures of fuel and oxidizer were shock-heated.

Simon H. Bauer

Bibliography. E. F. Greene and J. P. Toennies, *Chemical Reactions in Shock Waves*, 2d ed., 1964; J. V. Michael and K. P. Lim, *Annu. Rev. Phys. Chem.*, 44:429 1993; W. Tsang and A. Lifshitz, *Annu. Rev. Phys. Chem.*, 41:559, 1990.

Shock wave

A mechanical wave of large amplitude, propagating at supersonic velocity, across which pressure or stress, density, particle velocity, temperature, and related properties change in a nearly discontinuous manner. Unlike acoustic waves, shock waves are characterized by an amplitude-dependent wave velocity. Shock waves arise from sharp and violent disturbances generated from a lightning stroke, bomb blast, or other form of intense explosion, and from various types of supersonic flow over stationary objects. Extraterrestrial examples include supernovae expanding against neighboring intergalactic gas clouds and solar wind flowing over magnetized planets and satellites. This article discusses shock waves in gases and in condensed materials.

Gases

The abrupt nature of a shock wave in a gas can best be visualized from a schlieren photograph or shadow graph of supersonic flow over objects. Such photographs show well-defined surfaces in the flow field across which the density changes rapidly, in contrast to waves within the range of linear dynamic behavior of the fluid. Measurements of fluid density, pressure, and temperature across the surfaces show that these quantities always increase along the direction of flow, and that the rates of change are usually so rapid as to be beyond the spatial resolution of most instruments except when the shock wave is very weak or at very low gas densities. These surfaces of abrupt change in fluid properties are called shock waves or shock fronts. *See* SCHLIEREN PHOTOGRAPHY; SHADOWGRAPH; WAVE MOTION IN FLUIDS.

Shock waves in supersonic flow may be classified as normal or oblique according to whether the orientation of the surface of abrupt change is perpendicular or at an angle to the direction of flow. A schlieren photograph of a supersonic flow over a blunt object is shown in **Fig. 1**. Although this photograph was obtained from a supersonic flow over a stationary model in a shock tube, the general shape of the shock wave around the object is quite typical of those observed in a supersonic wind tunnel, or of similar objects (or projectiles) flying at supersonic speeds in a stationary atmosphere. The shock wave in this case assumes an approximately parabolic shape and

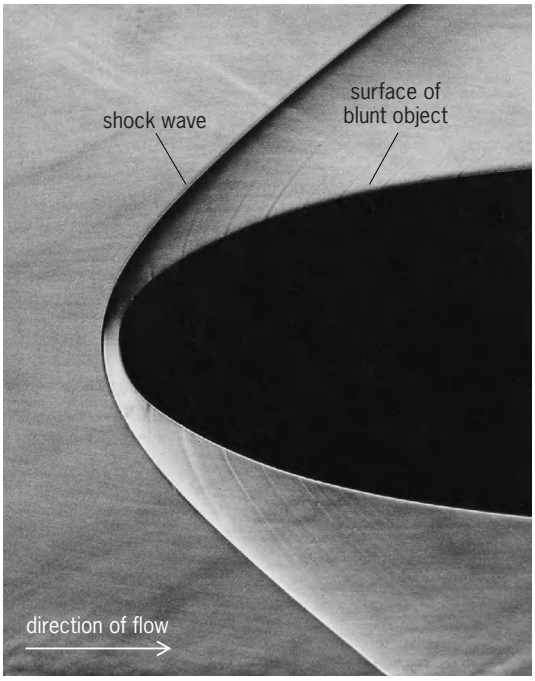

Fig. 1. Schlieren photograph of supersonic flow over blunt object. Shock wave is approximately parabolic, and detached from object. (*Avco Everett Research Laboratory, Inc.*)

is clearly detached from the blunt object. The central part of the wave, just in front of the object, may be considered an approximate model of the normal shock; the outer part of the wave is an oblique shock wave of gradually changing obliqueness and strength. The stand-off distance Δ (that is, the closest distance between the shock wave and the object) is governed by the density ratio across the normal shock, ρ_2/ρ_1, which increases with the upstream

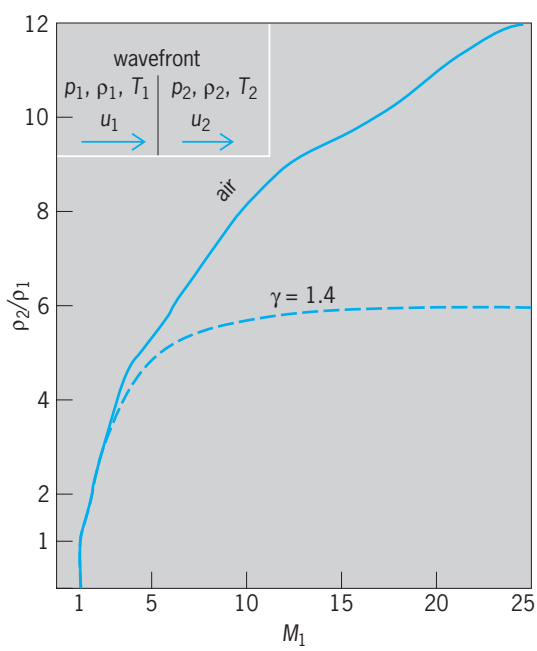

Fig. 2. Density ratio ρ_2/ρ_1 across normal shock wave in air at standard atmospheric density as a function of Mach number M_1. Inset shows variables in case of steady flow across a stationary wavefront.

Mach number, M_1, and varies with gas composition (**Fig. 2**). At high Mach numbers, the value of Δ becomes a small fraction of the nose radius of the object, R_N, such that Δ/R_N is approximately equal to ρ_1/ρ_2. For supersonic flow over sharp objects, such as slender cones and thin wedges, Δ approaches zero as the nose radius R_N vanishes. The shock waves then become oblique shock waves attached to the tip of the sharp object. For a discussion of experimental techniques for supersonic flow *see* BALLISTIC RANGE; WIND TUNNEL. *See also* SONIC BOOM; SUPERSONIC FLIGHT.

Curved shock waves of the type shown in Fig. 1 can also be found in quasi-steady supersonic flows over blunt objects with flexible skins. A good example is the solar wind flow over the Earth's magnetopause. The stand-off distance of the bow wave from the self-adjusting curved surface of the magnetopause also varies with Mach number in the same way as in terrestrial supersonic flow, provided that the Mach number is suitably redefined as the ratio between solar wind velocity and the highest wave speed for propagating small pressure and momentum flux disturbances in the upstream direction, which may or may not be the ordinary sound speed. *See* MAGNETOSPHERE; SOLAR WIND.

Normal shock wave. The changes in thermodynamic variables and flow velocity across the shock wave are governed by the laws of conservation of mass, momentum, and energy, and also by the equation of state of the fluid. For the case of normal shock, the inset in Fig. 2 illustrates a steady flow across a stationary wavefront (as in the case of a stationary model in a shock tube). The mass flow and momentum equations are the same as for an acoustic wave. However, in a shock wave, changes in pressure p and density ρ across the wavefront can no longer be considered small. As a consequence, the velocity of propagation of the shock wave relative to the undisturbed fluid is givenby Eq. (1), where the initial state

$$u_1^2 = \frac{\rho_2(p_2 - p_1)}{\rho_1(\rho_2 - \rho_1)} \tag{1}$$

of the fluid is denoted by subscript 1 and variables behind the shock front are denoted by subscript 2. In addition, conservation of thermal and kinetic energy across the shock front requires the validity of Eq. (2),

$$h_1 + {}^1/_2 u_1^2 = h_2 + {}^1/_2 u_2^2 \tag{2}$$

where h is the specific enthalpy (or total heat per unit mass) of the fluid, and u_1 and u_2 are fluid velocities relative to the shock wave. By eliminating u_2 and u_1 with the aid of Eq. (1) and the law of conservation of mass, $\rho_1 u_1 = \rho_2 u_2$, the energy equation becomes Eq. (3). If the thermodynamic properties of the fluid

$$h_2 - h_1 = {}^1/_2 \left(\frac{1}{\rho_1} + \frac{1}{\rho_2}\right)(p_2 - p_1) \tag{3}$$

are known, specific enthalpy h can be expressed as a function of pressure and density, or of any other pair of thermodynamic variables. Equations (1) and (3),

together with the appropriate equation of state of the fluid, are known as the Rankine-Hugoniot equations for normal shock waves. From this set of equations, all thermodynamic variables behind the shock front (denoted by subscript 2) can be expressed as functions of the propagation velocity of the shock wave and the known initial state of the fluid (denoted by subscript 1). For example, if the fluid is a perfect gas of constant specific heats, enthalpy h can be written as Eq. (4), where C_p is the specific heat at constant

$$h = C_p T = \frac{\gamma}{\gamma - 1}\frac{p}{\rho} = \frac{a^2}{\gamma - 1} \tag{4}$$

pressure, T is the thermodynamic temperature, γ is the ratio of specific heats, and a is the adiabatic speed of sound given by $(\gamma RT)^{1/2}$, where R is the gas constant. For this case, the pressure and density ratios across the shock front are given by Eqs. (5) and (6).

$$\frac{p_2}{p_1} = \frac{2\gamma M_1^2 - (\gamma - 1)}{\gamma + 1} \tag{5}$$

$$\frac{\rho_2}{\rho_1} = \frac{u_1}{u_2} = \frac{(\gamma + 1)M_1^2}{2 + (\gamma - 1)M_1^2} \tag{6}$$

The temperature ratio, deduced from the perfect gas law, is given by Eq. (7). These expressions show that

$$\frac{T_2}{T_1} = \frac{\left[2\gamma M_1^2 - (\gamma - 1)\right]\left[2 + (\gamma - 1)M_1^2\right]}{(\gamma + 1)^2 M_1^2} \tag{7}$$

the ratios of all thermodynamic variables and flow velocities across the shock depend on only one parameter for a given gas, which is the Mach number M_1 of the flow relative to the shock front, where $M_1 = u_1/a_1$, or in other words, M_1 equals the velocity of the shock wave divided by the speed of sound for the gas into which the shock propagates. Because of this, the magnitude of M_1 is often used as a measure of the strength of the shock wave. For comparison with the amplitude of acoustic waves, sound waves correspond to values of $M_1 \cong 1.001$ or less. *See* ENTHALPY; GAS; HEAT CAPACITY; MACH NUMBER.

The results of Eqs. (5), (6), and (7) have been derived for gases of constant specific heats. From molecular and atomic physics, it is well known that, when a gas is heated to high temperatures, vibrational excitation, dissociation, and ionization take place, with accompanying changes in heat capacities of the gas. Therefore, for strong shock waves, the appropriate expression for the specific enthalpy h, and the equation of state, which takes into account these phenomena, must be used in place of Eq. (4) to obtain the shock wave solution from Eqs. (1), (2), and (3). The ratios ρ_2/ρ_1, T_2/T_1, and p_2/p_1 for normal shock waves in air at typical atmospheric pressure $p_1 = 1$ bar (75 cm Hg or 100 kPa) and temperature $T_1 = 288$ K (59°F) are plotted in Figs. 2, **3**, and **4**. The approximate solutions, as given by Eqs. (6) and (7), hold only for the weaker shock waves ($M_1 < 6$), even though the pressure ratio is relatively insensitive to the changes in heat capacities of the gas.

As the upstream Mach number M_1 increases, contributions from vibrational excitation, molecular dissociation, and ionization to the specific heats and enthalpy of the shock-heated gas mixture become important. The resultant effect is to lower the temperature T_2 and to increase the compressibility. The increased compressibility then leads to a higher value of the mass density ρ_2.

The relative insensitivity of the pressure ratio p_2/p_1 to changes in M_1, on the other hand, is due to the fact that at high Mach numbers conservation of momentum across the normal shock depends primarily on a balance between the downstream pressure p_2 and the upstream momentum flux $\rho_1 u_1^2$. As M_1 further increases, the other terms, p_1 and $\rho_2 u_2^2$, become even less important, whatever the thermal properties of the shock-heated gas.

Because the Rankine-Hugoniot equations do not impose any limit on the value of that is M_1, there remains the question of whether a shock wave can propagate into an undisturbed gas at a speed lower than the speed of sound of this gas, that is, $M_1 < 1$. Although this question cannot be answered by the first law of thermodynamics, an examination of the change in specific entropy, using Eq. (8), may provide the answer. Thus Eq. (8) holds, which, for gases of constant specific heats, becomes Eq. (9), showing

$$\Delta S_{12} = S_2 - S_1 = \int_1^2 \frac{dE + p\,dv}{T} \tag{8}$$

$$\frac{\Delta S_{12}}{R} = \frac{1}{\gamma - 1}\ln\frac{\left[2\gamma M_1^2 - (\gamma - 1)\right]\left[2 + (\gamma - 1)M_1^2\right]^{\gamma}}{(\gamma + 1)^{\gamma+1} M_1^{2\gamma}} \tag{9}$$

that entropy change ΔS_{12} assumes a negative value when M_1 is noticeably less than unity. This violates the second law of thermodynamics, which states that the entropy accompanying any naturally occurring processes always tends to increase. Therefore, it follows that shock waves always travel at supersonic speeds relative to the fluids into which they propagate. *See* ENTROPY; THERMODYNAMIC PRINCIPLES; THERMODYNAMIC PROCESSES.

The Rankine-Hugoniot equations relating the thermodynamic variables p, ρ, and T on the two sides of a normal shock wave are based on the fundamental laws of mechanics, that is, conservation of mass, momentum, and energy during any adiabatic change of state in an isolated system. Since the laws of mechanics are invariant to the choice of inertial frame of reference, the computed values of ρ_2/ρ_1, T_2/T_1, and p_2/p_1 shown in Figs. 2, 3, and 4 for a stationary shock wave in supersonic flow, are equally applicable to a propagating shock wave generated by a violent explosion in an otherwise quiescent atmosphere.

Oblique shock wave. The changes in flow variables across an oblique shock wave are also governed by the laws of conservation of mass, momentum, and

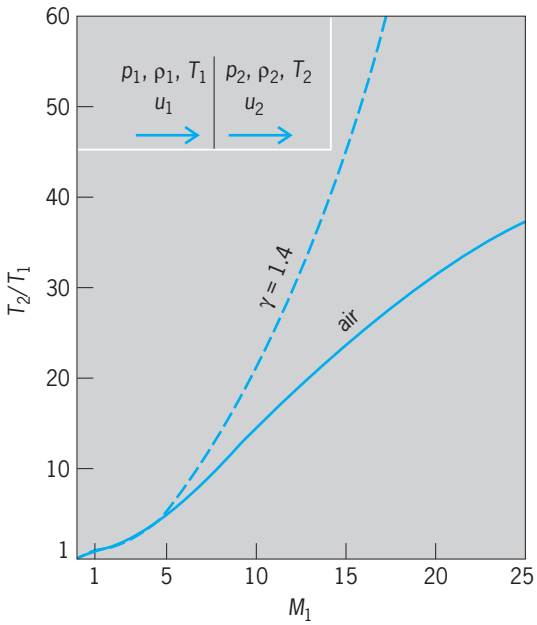

Fig. 3. Temperature ratio T_2/T_1 across normal shock wave in air at standard atmospheric density as a function of Mach number M_1. Inset shows variables in case of steady flow across a stationary wavefront.

energy in a coordinate system which is stationary with respect to the shock front. In this case, the problem is slightly complicated by the fact that the flow velocity will experience a sudden change of direction as well as magnitude in crossing the shock front. Thus, if β_1 and β_2 denote the acute angles between the initial and final flow velocity vectors and the shock surface (**Fig. 5**), then in crossing the oblique shock, the flow will be deflected by a finite amount $\theta = \beta_1 - \beta_2$.

The oblique shock solution can be obtained directly from the complete set of conservation equations. However, the solution already obtained for normal shock waves provides the following simplifying information.

The rate of mass flow per unit area across the shock wave is determined by the normal component of the flow velocity (Fig. 5). Thus, for conservation of mass across the shock, Eq. (10) holds. On

$$\rho_1 u_1 \sin\beta_1 = \rho_2 u_2 \sin\beta_2 \tag{10}$$

the other hand, conservation of the parallel component of momentum across the shock front requires that Eq. (11) hold. Equations (10) and (11) show that Eq. (12) holds. Itis equivalent to the statement

$$\rho_1 u_1^2 \sin\beta_1 \cos\beta_1 = \rho_2 u_2^2 \sin\beta_2 \cos\beta_2 \tag{11}$$

$$u_1 \cos\beta_1 = u_2 \cos\beta_2 \tag{12}$$

that the tangential component of the flow velocity must remain unchanged in crossing the oblique shock wave. Therefore, the resultant flow across the oblique shock shown in Fig. 5 will be identical to what would be seen by an observer who moved at a uniform velocity $u_1 \cos\beta_1$ along the surface of a normal shock wave propagating at a velocity $u_1 \sin\beta_1$. Such a translation of the frame of reference in the direction parallel to the shock front should not change the strength of the shock wave; thus, the changes in thermodynamic variables across the shock should depend only on the velocity component normal to the shock wave. The substitution of M_1 in Eqs. (5), (6), (7), and (9) with $M_1 \sin\beta_1$ gives the corresponding expressions for oblique shock waves in gases of constant specific heats. Again, thermodynamic considerations show that the normal component of the flow velocity into the oblique shock wave must be at least sonic. Therefore, in a supersonic stream of Mach number $M_1 > 1$, the value of β_1 must lie within the range given by expression (13).

$$\sin^{-1}\frac{1}{M_1} \le \beta_1 \le \frac{\pi}{2} \tag{13}$$

The lower limit corresponds to the Mach angle of acoustic waves, while the upper limit corresponds to that of the normal shock wave. At both limits, the flow deflection angle $\theta = \beta_1 - \beta_2$ will be zero. For any intermediate value of β_1, the value of β_2 can

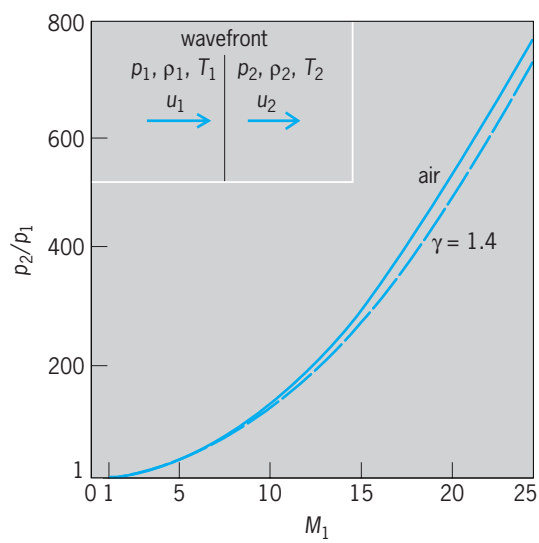

Fig. 4. Pressure ratio p_2/p_1 across normal shock wave in air at standard atmospheric density as a function of Mach number M_1. Inset shows variables in case of steady flow across a stationary wavefront.

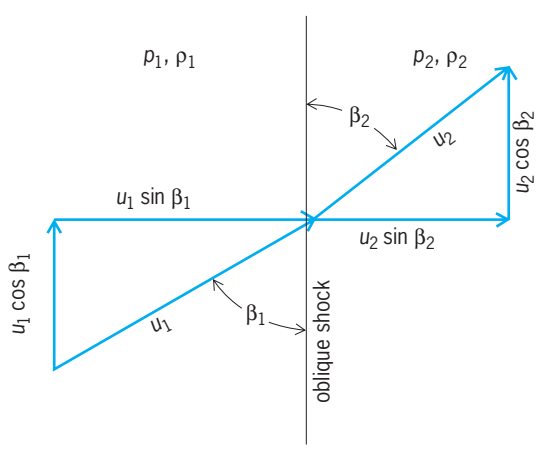

Fig. 5. Flow across an oblique shock wave.

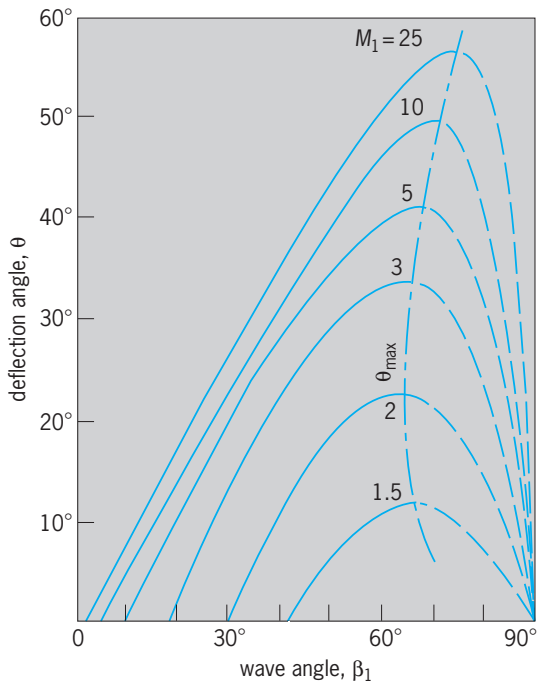

Fig. 6. Flow-detection angle for oblique shock waves in air at standard atmospheric density.

be obtained from Eqs. (10) and (12). Thus Eq. (14)

$$\beta_2 = \tan^{-1}\left(\frac{\rho_1}{\rho_2}\tan\beta_1\right) \tag{14}$$

holds. The flow deflection angle θ for oblique shock waves in air at normal density is plotted in **Fig. 6** as a function of M_1 and β_1. For a given Mach number M_1, the flow deflection angle first increases with the wave angle β_1, reaches a maximum value θ_{max}, and then decreases toward zero again as the wave angle approaches that of the normal shock. Conversely, for any given flow deflection angle $\theta < \theta_{max}$ such as would be produced by sudden introduction of a wedge or an inclined plane surface into the initially uniform supersonic stream, there exist two possible values of β_1. The higher value of β_1 corresponds to the stronger shock. If the wedge angle or the inclination of the plane surface so introduced exceeds the value of θ_{max} for the given M_1, the shock wave either will become detached from the obstructing object or will form a more complicated pattern. The question of exactly what shock-wave pattern to expect from a given situation is complicated by interaction of the resultant shock wave and flow pattern on the overall boundary condition as well as on the local state of the flow.

Bomb blast. When energy is suddenly released into a fluid in a concentrated form, such as by a chemical or a nuclear explosion, the local temperature and pressure may rise instantly to such high values that the fluid tends to expand at supersonic speed. When this occurs, a blast wave forms and propagates the excess energy from the point of explosion to distant parts of the fluid. If the point of explosion is far from any fluid boundary, the blast wave assumes the form of an expanding spherical shock wave followed by a radially expanding fluid originating from the point of detonation. The changes in thermodynamic variables across the spherical shock are the same as those for a normal shock propagating at the same instantaneous velocity. However, because of the continuous expansion and the finite amount of energy available from the explosion, both the strength of the shock and the specific energy of the expanding fluid must decay with time. The decay of a blast wave goes through three principal stages. A strong shock period begins immediately after the formation of the blast wave, during which the shock strength decays rapidly with distance from the point of detonation. During this stage, the shock velocity decays with the inverse $^3/_2$ power of the distance, and the overpressure behind the shock decays with the inverse cube of the distance. The second stage is a transition period, during which the strong spherical shock gradually changes into an acoustic wave. During the last stage or residual acoustic decay period, the acoustic wave carries the sound of explosion great distances from the point of detonation. As characteristic of sound propagation in three dimensions, the overpressure carried by the spherical acoustic wave decays inversely with distance, and velocity of propagation is constant. *See* EXPLOSIVE.

Extraterrestrial shock waves. The formation and propagation of shock waves are natural processes that cause massive stars to end their lives in supernova explosions and spread their remnants to provide heavy elements that are essential for sustaining life on Earth. While prominent supernovae, such as the one observed by Chinese astronomers in 1054 and Kepler's supernova in 1604, are rare, in the sense that only roughly 25 (most of them obscured by dust) are believed to occur in the Milky Way Galaxy each 1000 years, this slow pace still adds up to tens of millions of supernova explosions over the cosmic time scale of billions of years before the birth of the solar system. *See* SUPERNOVA.

A supernova explosion occurs after a massive star has converted most on its available fusion energy into radiation and its fusion ash, iron, has accumulated at its core to a mass significantly greater than the solar mass, and to a size comparable to that of our Sun. As the fusion energy conversion rate slows and the iron core continues to cool and shrink, the gravitational pressure at its center will eventually overwhelm the degeneracy pressure of the atomic electrons, allowing them to be suddenly captured by the nuclear protons to form neutrons. In the process, neutrinos are released to carry away the excess energy, the specific volume of iron is suddenly reduced to that of nuclear matter, and the iron core implodes gravitationally into a proto-neutron star (p-n star). *See* GRAVITATIONAL COLLAPSE; NEUTRON STAR.

Before the p-n star stabilizes into a neutron star at the end of the core-collapse process (assuming that the star was not massive enough to collapse into a black hole), it is overcompressed momentarily by the momentum flux of the imploding stream and recoils. At that moment, a shock wave begins to form at the

p-n star interface to bounce the incoming stream outward. The formation of this shock wave and its propagation through the dense stellar layers just above the p-n star are affected by p-n star reverberation, gravitational deceleration, energy transfer from neutrinos to the dense stellar plasma behind the shock front, heavy-element production, and so forth, and hence are quite complex. The shock wave is strong enough to superheat the stellar plasma to temperatures far beyond the local precollapse values (about 10^{11} K near the core and 10^4 K in the photosphere).

As the shock wave propagates out of the stellar envelope and into the near vacuum of interstellar space, the superheated stellar plasma explodes into an extremely bright supernova. During the explosion, the stellar plasma expands rapidly in the manner of an unrestrained expansion wave. The shock wave ahead of it continues to propagate into the low-density interstellar gas. As the supernova plasma cools, it eventually condenses into a luminous cloud that becomes the remnant of the supernova. The Crab Nebula is believed to be such a remnant from the 1054 supernova. *See* CRAB NEBULA.

The supernova shock wave is generally invisible. However, it can be detected when it propagates into a gas cloud of substantial density. A good example is Supernova 1987A, whose shock wave was invisible but was detected 11 years later in 1998 by the *Hubble Space Telescope* and the *Chandra X-ray Observatory*. *See* CHANDRA X-RAY OBSERVATORY; HUBBLE SPACE TELESCOPE.

Another example of an extraterrestrial shock wave is the solar wind shock wave just outside the Earth's magnetopause, where the solar wind momentum flux is just balanced by magnetic field pressure so that there is no net flow across the equipotential surface. Space probes launched by the National Aeronautics and Space Administration (NASA) beginning in 1958 suggested the existence of such a shock wave. Plasma probe measurements onboard the *Imp 1* (*Explorer 18*) spacecraft in 1963–1964 confirmed the existence of a quasisteady solar wind of relatively high velocity (300–800 km/s or 200–500 mi/s) and low density (less than 10 ions/cm^3) outside the magnetopause transition region. Magnetometer signals also indicated the presence of plasma turbulence and magnetohydrodynamic wave structure inside the transition region.

The mean profiles of the outer and inner boundaries of the magnetopause transition region mapped out by the first 30 orbits of the *Imp 1* spacecraft seemed to resemble the typical bow shock wave shown in Fig. 1 for steady supersonic flows over blunt objects. In 2005, computer simulations using real-time data from NASA's *ACE* (*Advanced Composition Explorer*) solar wind "weather report" showed that both the shock wave and magnetopause surfaces actually change shape and position rapidly in response to the rapidly changing upstream solar wind conditions during periods of strong solar activities.

Structure, thickness, and stability. Shock wave structure and shock wave stability are topics of considerable interest in plasma physics and astrophysics, as well as in high-temperature gas dynamics. In certain engineering problems, such as the feasibility and efficiency of laser-driven fusion, the question of shock wave stability is crucial.

In gas dynamics, shock wave structure usually refers to spatial distributions of thermodynamic variables such as ρ, P, and T along the flow direction across a normal shock wave. In diatomic and polyatomic gas mixtures, it may include spatial variations of chemical composition and molecular energy level population. In plasma physics and astrophysics, it may also include spatial variations of magnetic field pressure, plasma wave spectra, energy transfer rate, and so forth. From such distributions, one may define an overall shock wave thickness as the spatial extent of the transition zone through which the working fluid changes from the initial equilibrium state 1 to the final equilibrium state 2. However, the term "shock wave thickness" or "shock front thickness" is commonly used to refer only to the length scale over which the sudden jump in gas pressure or density occurs across a normal shock wave. *See* GAS DYNAMICS.

Kinetic theory and low-density shock tube experiments show that the shock front thickness for sudden pressure or density change is only about the length l_1 of one mean free path for molecular collisions at the upstream gas density and temperature over a very wide range of values of shock strength and initial chemical composition. Thus, for a normal shock in air at normal sea level atmospheric pressure and temperature, l_1 is about 0.06 μm, which is the optical wavelength in the deep ultraviolet. This explains why the shock front thickness is difficult to measure. On the other hand, in the upper stratosphere at 50 km (30 mi) altitude, where $p_1 = 0.5$ torr (67 Pa) and $T_1 = 355$ K (179°F), the molecular mean free path l_1 is about 0.1 mm and so would be the thickness of the shock front.

The reason for such a thin shock front is that diatomic gas molecules need only one close (but not head-on) collision to exchange translational momentum and rotational energy. In contrast, only head-on or nearly head-on collisions at sufficiently high kinetic energies are effective for vibrational excitation, dissociation, and ionization, so completion of these processes takes place at much greater distances behind the shock front.

Shock wave stability refers to the ability of the shock wave to maintain its geometrical configuration in the presence of extraneous disturbances, or its ability to recover its original configuration after such disturbances. In telecommunications, knowledge about solar wind shock wave stability would be valuable as it pertains to the recovery time of the Earth's magnetosphere after a severe magnetic storm.

In laser-driven fusion, a form of inertial confinement nuclear fusion which has been proposed as an alternative to magnetic confinement fusion for solving the world's long-term energy supply problem (for a million years or more according to some estimates based on the world's current total energy consumption rate and the total amount of deuterium

and tritium in ocean water), the ability of a converging shock wave to maintain its spherical symmetry in a compressed deuterium–tritium (D–T) mixture would be crucial in the efficient conversion of the fusion reaction energy into electricity. In this case, the shock wave, driven by the reactive force (as in rocket propulsion) of a rapidly heated and outwardly expanding D–T vapor plasma generated by intense focused multiple laser beams, can be expected to behave like an ordinary shock wave in condensed matter. *See* NUCLEAR FUSION.

In early shock tube experiments, converging cylindrical shock waves appeared to be relatively stable up to certain shock strengths, but beyond that, the stability of the shock wave appeared to be at best neutral; that is, initial spatial nonuniformity was maintained but not smoothed out as the converging process continued. This, perhaps, was not too surprising since momentum transfer along the transverse direction of a curved ordinary shock wave cannot be expected to extend much beyond one shock front thickness, or one mean free path for molecular collisions in gaseous media. In plasmas and in condensed matter, the length scale for effective momentum transfer can be quite different. Nevertheless, the early two-dimensional cylindrical shock wave experiments already pointed to the need to maintain an extremely spherically symmetric distribution of focused laser beam intensity in laser-driven fusion in order to achieve efficient energy conversion.

Shao-chi Lin

Condensed Materials

Some common examples of shock waves in condensed materials are encountered in the study of underground or underwater explosions, meteorite impacts, and ballistics problems. The field of shock waves in condensed materials (solids and liquids) has grown into an important interdisciplinary area of research involving condensed matter physics, geophysics, materials science, applied mechanics, and chemical physics. The nonlinear aspect of shock waves is an important area of applied mathematics.

The discussion below will be restricted to one-dimensional plane-wave propagation, in which material properties vary with respect to only one spatial coordinate and time. This situation is the simplest to treat theoretically and is the most frequently studied experimentally. Problems involving one-dimensional cylindrical- or spherical-wave propagation have certain advantages for the study of solids and for many applications, but no satisfactory measurement techniques have been developed to investigate them.

Formation and behavior. The production of one-dimensional plane waves requires rapid, simultaneous deformation of a large flat surface. As the shock wave propagates through the sample thickness, it compresses each macroscopic (or continuum) volume element along the wave propagation direction. Due to inertial confinement, the volume element cannot be deformed laterally, and a state of one-dimensional or uniaxial strain results in the medium. The volume elements at the lateral edge of a finite sample are unconfined, and their expansion results in a rarefaction (or decompression) wave that destroys the uniaxial strain condition and also lowers the pressure in the sample. Experimental measurements are typically made in the central region of the sample before the arrival of the edge rarefactions. Hence, the macroscopic condition of uniaxial strain is satisfied exactly during this time, and the shock wave appears to be extended infinitely in the lateral direction for times of interest.

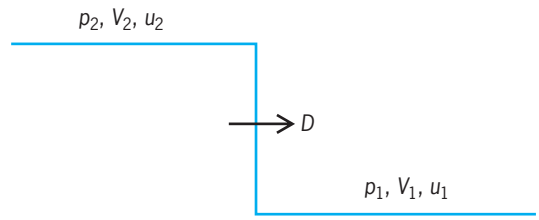

Fig. 7. Diagram of discontinuous shock front moving with a velocity *D* and connecting two uniform states. Here *p* is pressure or equilibrium stress, *V* is specific volume, *u* is particle or mass velocity, and subscripts 1 and 2 represent the initial and final states.

The formation of shock waves is a consequence of the nonlinear response of condensed materials and can be qualitatively understood by examining the effect of a piston moving into a compressible material. The piston motion and the resulting compression can be considered as a series of smaller steps. The propagation velocity for each step is given by $c + u$, where c represents the acoustic velocity corresponding to the compression for that step, and u is the particle velocity at the end of the previous step. For most materials, the propagation velocity increases with each succeeding step, and the compressive wave motion steepens with propagation distance. Eventually, a nearly discontinuous shock front propagating at a supersonic velocity is formed (**Fig. 7**). Although true mathematical discontinuities do not occur in real materials, the idealization depicted in Fig. 7 is a reasonable approximation because of the very short rise times observed in experiments. The above explanation can be reversed to show that a rarefaction or unloading wave will spread out with propagation distance. *See* NONLINEAR ACOUSTICS.

Experimentally, shock waves are produced by rapidly imparting momentum over a large flat surface. This can be accomplished in many different ways: rapid deposition of radiation using electron or photon beams (lasers or x-rays), detonation of a high explosive in contact with the material, or high-speed impact of a plate on the sample surface. The impacting plate itself can be accelerated by using explosives, electrical discharge, underground nuclear explosions, and compressed gases. The use of compressed gas to accelerate projectiles with appropriate flyer plates provides the highest precision and control as well as convenience in laboratory experiments.

By using gas guns, stresses (or pressures) ranging from 50 megapascals (0.5 kilobar) to 1 terapascal (10 megabars) have been achieved. By using lasers and underground nuclear explosions, pressures of the order of 10 TPa (100 Mbar) have been reported (the pressure at the center of the Earth is 350 gigapascals or 3.5 Mbar). Time durations range from a nanosecond to several microseconds; the upper limit is determined by sample size. Because of the short time durations, the shock process can be considered as adiabatic. The irreversible nature of shock compression can result in a large temperature rise in the sample. At the very highest pressures, temperature changes are of the order of 10^5 K, and the electronic contribution to pressure becomes significant. *See* ADIABATIC PROCESS.

Two other important characteristics of shock compression are the fast rate of loading (possibly subnanosecond) and the production of large shear stresses in solids. The rapid loading rates make it possible to examine the kinetics of atomic and molecular rearrangement due to large compressions. Although the term pressure is commonly used in shock-wave studies, the use of stress is more appropriate because the state of stress is nonhydrostatic in general. The generation of large shear stresses makes possible the investigation of various deformation and yield mechanisms in solids that could not be studied otherwise. *See* STRESS AND STRAIN.

Large-amplitude one-dimensional compression and shear waves have been studied in solids. In these experiments, a macroscopic volume element is subjected to both a compression and shear deformation. The combined deformation state is produced by impacting two parallel flyer plates that are inclined at an angle to the direction of the plate motion. Momentum conservation coupled with different wave velocities for compression and shear waves leads to a separation of these waves in the sample interior. An example of a shear wave is shown in **Fig. 8**. These experiments provide direct information about the shear response of shocked solids, and subject samples to more general loading states than the uniaxial strain state.

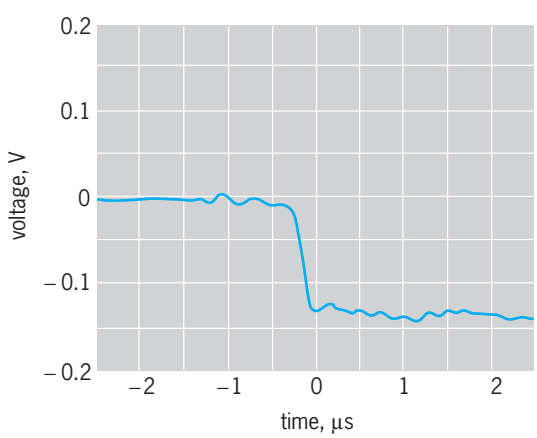

Fig. 8. Typical large-amplitude shear-wave profile. The particle velocity has been converted to a voltage signal. The negative signal is of no consequence. The small perturbation prior to the shear wave is caused by the large compression wave and imperfections in the experimental apparatus.

Theory. Equations (15)–(17) govern continuous

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} = 0 \tag{15}$$

$$\rho \frac{\partial u}{\partial t} = \frac{\partial \sigma}{\partial x} = 0 \tag{16}$$

$$\frac{dE}{dt} + \sigma \frac{\partial V}{\partial t} = 0 \tag{17}$$

one-dimensional planar compression in the absence of heat conduction. Here ρ is density, u is particle or mass velocity; σ is the compressive stress in the x direction and includes viscous and other time-dependent contributions; E is the specific internal energy; and $V (= 1/\rho)$ is the specific volume. Equations (15)–(17), written in the spatial (or eulerian) representation, are valid for both solids and fluids. In analyzing experimental data, it is often more convenient to write these equations in the material (or lagrangian) representation. In the more general case of combined compression and shear waves, the governing equations are generalized to include shear deformation.

For discontinuous jumps or steady shocks, there is no change in the wave shape, and Eqs. (15)–(17) can be integrated to give the Rankine-Hugoniot jump conditions, Eqs. (18)–(20), where D is the shock

$$\rho_1(D - u_1) = \rho_2(D - u_2) \tag{18}$$

$$\sigma_2 - \sigma_1 = \rho_1(D - u_1)(u_2 - u_1) \tag{19}$$

$$E_2 - E_1 = {}^1/_2(\sigma_2 + \sigma_1)(V_1 - V_2) \tag{20}$$

velocity. Equations (18)–(20) relate the initial (subscript 1) and final (subscript 2) states, assumed to be equilibrium states, on either side of the shock front. Hence, σ appearing in the jump conditions represents equilibrium stress (or pressure for fluids) unlike the σ in Eqs. (15)–(17). Within the shock front, deviations from equilibrium can be appreciable.

When the jump conditions are applicable, the locus of end states achieved through a shock transition is referred to as the Rankine-Hugoniot (R–H) curve (**Fig. 9**) or merely the Hugoniot. Unlike curves commonly used in quasistatic experiments, the Rankine-Hugoniot curve in Fig. 9 does not represent the pressure-volume path followed in a given experiment. The actual states that a material element follows in a shock transition lie along the chord connecting the end states and are referred to as the Rayleigh line. The slope of this chord is a measure of the shock velocity and increases with the shock amplitude.

The discussion of thermodynamics of shock compression will be restricted to fluids to permit simplification. The third jump condition, Eq. (20), refers only to thermodynamic quantities and can therefore be combined with the equation of state for a material to generate the Rankine-Hugoniot curve in the

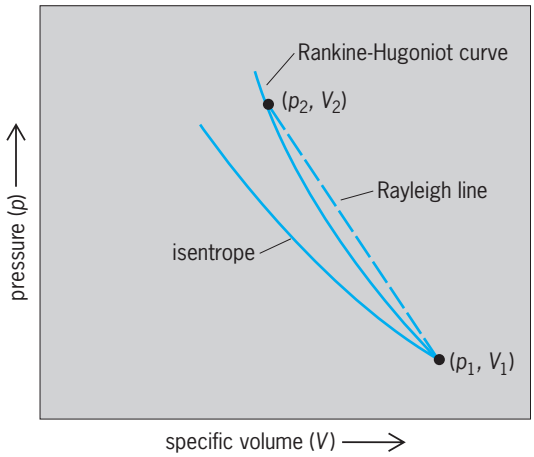

Fig. 9. Rankine-Hugoniot curve and the Rayleigh line for an initial state with pressure and volume (p_1, V_1). Different initial states will lead to different Rankine-Hugoniot curves. An isentrope through the initial state is also shown. For a solid, pressure should be replaced by equilibrium stress.

pressure-volume plane. This procedure, though simple in principle, is difficult to implement in practice because the equation of state is often not known over the pressure-temperature range of the shock experiments. Hence, it is the converse problem that is most useful: how to obtain information about the equation of state from shock-wave data. Measurement of any two quantities appearing in the jump conditions permits a complete solution to Eqs. (18)–(20). Typically, the shock velocity (D) and either pressure or particle velocity are measured in experiments. These data are then used to calculate the values of pressure, specific volume, and specific internal energy along the Hugoniot curve. Calculation of temperature along the Hugoniot and thermodynamic quantities away from the Hugoniot curve require additional assumptions. For solids, there is the additional complexity of correcting for material strength in the stress measurements.

The irreversible nature of the shock transition gives rise to entropy production, which for weak shocks (approximately 15% compression) can be written by expanding the entropy S in a Taylor's series, Eq. (21), where p is the pressure. This equation

$$S_2 - S_1 = \frac{1}{12T_1}\left(\frac{\partial^2 p}{\partial V^2}\right)_{S_1,V_1}(V_1 - V_2)^3 + \cdots \quad (21)$$

shows that the entropy increases as the third order in compression and is positive when $(\partial^2 p/\partial V^2)_S$ is positive. The pressure-volume relation for most normal materials is indeed concave upward as shown in Fig. 9. In addition, the isentrope and the Hugoniot curve have a second-order contact (a point of contact at which the second derivatives are different) at the initial state because of the result in Eq. (21). The unloading curve from a shocked state is along an isentrope and is to the right of the Hugoniot. If the sample is unloaded to zero pressure in a shock-wave compression-release experiment, the final volume will be higher than the initial volume because of thermal expansion due to a higher temperature. *See* ISENTROPIC PROCESS.

The thickness of the shock front is also closely related to the question of thermal equilibrium in these experiments. The shock front thickness is a consequence of the balance between nonlinear compressibility and dissipative mechanisms such as viscosity, heat conduction, and plastic deformation in solids. Because all real materials have some dissipation, the hypothesis of a shock discontinuity in Fig. 7 is an idealization. *See* CONDUCTION (HEAT); PLASTIC DEFORMATION OF METAL; VISCOSITY.

The Hugoniot curve shown in Fig. 9 is applicable for most so-called normal materials. However, deviations from this behavior are of considerable scientific interest and result in shock instability or multiple waves. The general theory of shock stability and its relationship to thermodynamics is complicated and is an area of ongoing research. For nonreactive materials, the stability condition can be written as inequality (22), where c_2 is the sound speed in the

$$D - u_2 < c_2 \quad (22)$$

shocked state. This condition implies that the shock front is subsonic with respect to material behind the shock front, and a disturbance from behind can catch up with the shock front.

Condensed matter properties. Shock waves subject matter to unusual conditions and therefore provide a good test of understanding of fundamental processes. The majority of the studies on condensed materials have concentrated on mechanical and thermodynamic properties. These are obtained from measurements of shock velocity, stress, and particle velocity in well-controlled experiments. Advanced techniques using electromagnetic gages, laser interferometry, piezoelectric gages, and piezoresistance gages have given continuous, time-resolved measurements at different sample thicknesses. These data have permitted improved resolution of the shock front and states behind the front. Time resolution in the 500 picosecond to 1 nanosecond range has been achieved. A large body of high-pressure data has been gathered for most elements and a large number of other materials. These experiments have achieved pressure and temperature ranges well beyond those of static high-pressure experiments, and have provided data for comparison with theoretical predictions based on quantum-mechanical calculations of the equation of state.

Two particular areas of interest in studying mechanical and thermodynamic response are dynamic yielding (plastic deformation or yielding under dynamic loading) and polymorphic phase transitions. Both of these responses lead to an unstable shock, and each can result in a two-wave structure. If both phenomena are present in the experimental range, as for iron, a three-wave structure is observed (**Fig. 10**). The first wave travels with an elastic-wave velocity and has an amplitude that is related to the dynamic strength of the material. The second wave travels with the plastic-wave velocity (close to bulk sound

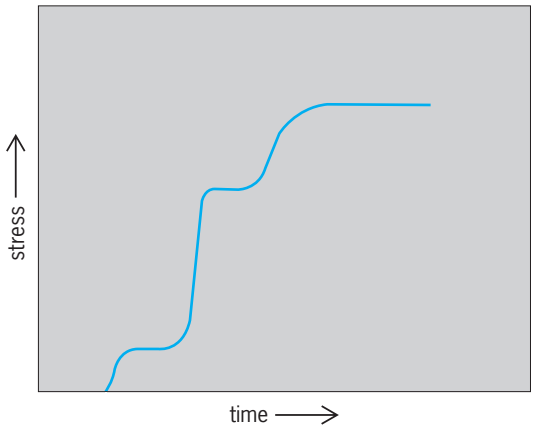

Fig. 10. Representative stress-time profile at a particular sample depth, showing the three-wave structure observed in iron. The first wave is the elastic precursor, the second-wave amplitude (13 GPa or 130 kilobars) marks the onset of the phase change, and the third wave is a measure of the peak stress for the experiment.

speed) and has an amplitude that depends on the peak stress or the phase-transition stress (if the material can undergo a phase change). In the event of a phase change, the amplitude and velocity of the third wave depend on the peak stress. Although a large number of studies have been carried out to investigate shock-induced yielding and phase transitions, the atomic mechanisms controlling these effects are not well understood. Some important unresolved questions concern the applicability of conventional dislocation theory to shock-induced yielding; the mechanisms for shear-strength loss in brittle materials; the precise mechanisms for going from one crystal structure to another in shock-induced phase changes; what controls the fast kinetics of phase changes; the role of material strength in shock-induced phase changes; and the process by which materials melt under shock loading. *See* CRYSTAL DEFECTS; CRYSTAL STRUCTURE; PHASE TRANSITIONS.

Two experimental issues fundamental to thermodynamic and mechanical studies are the complete determination of the stress state under shock loading, and the measurement of temperature (when applicable). Some progress has been made on the first issue, and a complete characterization of the stress state is expected in regions where the stress is nearly constant (the flat portions of the stress-time profile). The second problem is a major challenge for shock-wave experimentalists. Attempts to infer temperature from time-resolved emission from shocked fluids appear promising, but there are two fundamental problems with temperature measurements: the question of thermodynamic equilibrium and an independent check on the inferred temperature. Resolution of the latter problem requires comparison of results from different methods, and temperature measurements have been attempted in a range below that of the emission data. Another approach has involved the propagation of high-pressure ramp waves to approximate isentropic compression in the megabar (10^{11} Pa) range. This development should aid in ascertaining temperature effects and result in a better understanding of thermodynamic states away from the Hugoniot.

Electric, magnetic, and optical properties have not been extensively studied in shock-wave experiments, due to the difficulty in acquiring and interpreting these data. Some work has focused on shock-induced conductivity changes, demagnetization, and piezoelectric and ferroelectric studies, but a detailed understanding of these phenomena does not exist.

Technological developments in lasers and time-resolved optical diagnostics have stimulated considerable interest in spectroscopic measurements in condensed materials. The most extensive work has involved electronic spectroscopy of shocked carbon disulfide (CS_2); a typical absorption spectrum from a multiple shock experiment is shown in **Fig. 11**. Light passes through a thin layer of liquid carbon disulfide (1–300 micrometers) contained between two sapphire plates, and the intensity as a function of wavelength and time is recorded. The latter variable correlates directly with increasing pressure and temperature in the sample. The steps seen in the spectrum correspond to shock reverberation steps. The shift in absorption edge (the bottom part of the band) has been correlated with pressures and temperatures calculated from the carbon disulfide equation of state.

A clearer picture of the molecular processes governing the spectroscopic results will require numerous experiments with systematic variation of parameters and careful analysis. Despite the many challenges in both experiments and analysis, this area of research along with electrical measurements is essential to understanding the shock response of condensed materials on a molecular level.

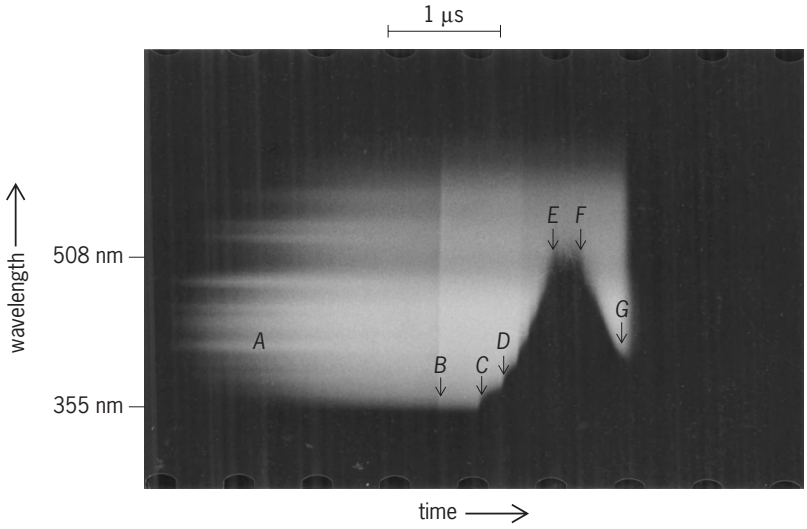

Fig. 11. Time-resolved ultraviolet-visible spectrum of liquid carbon disulfide (CS_2) compressed by shock-reverberation between two sapphire plates. Initial thickness of carbon disulfide = 0.320 mm. Peak pressure = 7.67 GPa = 76.7 kilobars. Symbols indicate events: A = xenon lamp turns on; B = impact; C = shock enters carbon disulfide; D = shock reflects from rear sapphire; E = peak pressure is reached; F = rarefaction wave decreases pressure; G = cell breaks up.

The study of residual effects, that is, the postshock examination of samples subjected to a known pulse amplitude and duration, is of considerable importance to materials science and metallurgy. The conversion of graphite to diamonds is noteworthy. Other effects that have been observed are microstructural changes, enhanced chemical activity, changes in material hardness and strength, and changes in electrical and magnetic properties. The generation of shock-induced lattice defects is thought to be important for explaining these changes in material properties. There has been growing interest in using shock methods for material synthesis and powder compaction. *See* DIAMOND; HIGH-PRESSURE CHEMISTRY; HIGH-PRESSURE MINERAL SYNTHESIS; HIGH-PRESSURE PHYSICS. Y. M. Gupta

Bibliography. J. D. Anderson, *Modern Compressible Flow: With Historical Perspective*, 3d ed., McGraw-Hill, 2003; F. N. Frenkiel and W. R. Sears (eds.), Proceedings of the International Symposium on Magneto-Fluid Dynamics, *Rev. Mod. Phys.*, 32(4):695–1033, 1960; I. I. Glass and J. P. Sislian, *Nonstationary Flows and Shock Waves*, Oxford University Press, 1994; W. D. Hayes and R. F. Probstein, *Hypersonic Flow Theory*, Academic Press, New York, 1959; W. B. Kunkel (ed.), *Plasma Physics in Theory and Application*, McGraw-Hill, 1966; M. A. Lieberman and A. L. Velikovich, *Physics of Shock Waves in Gases and Plasmas*, Springer-Verlag, 1986; M. A. Saad, *Compressible Fluid Flow*, 2d ed., Prentice Hall, 1998; A. Sawaoka (ed.), *Shock Waves in Materials Science*, Springer-Verlag, 1993; J. Smoller, *Shock Waves and Reaction-Diffusion Equations,* 2d ed., Springer-Verlag, 1994; Ya. B. Zeldovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Phenomena*, vols. 1 and 2, translated and edited by W. D. Hayes and R. F. Probstein, Academic Press, 1967.

Short circuit

An abnormal condition (including an arc) of relatively low impedance, whether made accidentally or intentionally, between two points of different potential in an electric network or system. *See* CIRCUIT (ELECTRICITY); ELECTRICAL IMPEDANCE.

Common usage of the term implies an undesirable condition arising from failure of electrical insulation, from natural causes (lightning, wind, and so forth), or from human causes (accidents, intrusion, and so forth). From an analytical viewpoint, however, short circuits represent a severe condition that the circuit designer must consider in designing an electric system that must withstand all possible operating conditions. The short circuit thus is important in dictating circuit design parameters (wire size and so on) as well as protective systems that are intended to isolate the shorted element. *See* ELECTRIC PROTECTIVE DEVICES; ELECTRICAL INSULATION; LIGHTNING AND SURGE PROTECTION.

In circuit theory the short-circuit condition represents a basic condition that is used analytically to derive important information concerning the network behavior and operating capability. Thus, along with the open-circuit voltage, the short-circuit current provides important basic information about the network at a given point. *See* ALTERNATING-CURRENT CIRCUIT THEORY; DIRECT-CURRENT CIRCUIT THEORY; NETWORK THEORY; THÉVENIN'S THEOREM (ELECTRIC NETWORKS).

The short-circuit condition is also used in network theory to describe a general condition of zero voltage. Thus the term short-circuit admittance (or impedance) is used to describe a network condition in which certain terminals have had their voltage reduced to zero, for the purpose of analysis. This leads to the terms short-circuit driving point admittance, short-circuit transfer admittance, and similar references to the zero voltage condition. *See* ADMITTANCE.

Short-circuit protection is a separate discipline dedicated to the study, analysis, application, and design of protective apparatus that are intended to minimize the effect of unintentional short circuits in power supply systems. For these analyses the short circuit is an important limiting (worst) case, and is used to compute the coordination of fuses, circuit reclosers, circuit breakers, and other devices designed to recognize and isolate short circuits. The short circuit is also an important parameter in the specification of these protective devices, which must have adequate capability for interrupting the high short-circuit current. *See* CIRCUIT BREAKER; FUSE (ELECTRICITY).

Short circuits are also important on high-frequency transmission lines where shorted stub lines, one-quarter wavelength long and shorted at the remote end, are used to design matching sections of the transmission lines which also act as tuning elements. *See* TRANSMISSION LINES. Paul M. Anderson

Bibliography. P. M. Anderson, *Analysis of Faulted Power Systems*, 1973; *IEEE Standard Dictionary of Electrical and Electronics Terms*, 5th ed., IEEE Std. 100, 1992.

Short takeoff and landing (STOL)

The term applied to heavier-than-air craft that cannot take off and land vertically but can operate within areas substantially more confined than those normally required by aircraft of the same size. A pure STOL aircraft is a fixed-wing vehicle that derives lift primarily from free-stream airflow over the wing and its high lift system, sometimes with significant augmentation from the propulsion system. Although all vertical takeoff and landing (VTOL) machines, including helicopters, can lift greater loads by developing forward speed on the ground before liftoff, they are still regarded as VTOL (or V/STOL craft), operating in the STOL mode. *See* VERTICAL TAKEOFF AND LANDING (VTOL).

Runway length. It has been customary to define STOL capability in terms of the runway length required to take off or land over a 50-ft (15-m)

obstacle, the concept of "short" length being variously defined as from 500 to 2000 ft (150 to 600 m), depending on the high-lift concept employed and on the mission of the aircraft.

In addition to being able to operate from short runways, STOL aircraft are usually expected to be able to maneuver in confined airspace so as to minimize the required size of the terminal area. Such aircraft must therefore have unusually good slow-flight stability and control characteristics, especially in turbulence and under instrument flight conditions.

Development and uses. The term STOL came into general use during the 1950s, following many efforts during the prior 20 years to develop this capability in aircraft. In these early concepts, STOL was primarily pursued by developing and applying numerous geometric means to improve the high-lift characteristics of the wing. Aircraft of this type were used in off-airport operations in Alaska, in the Canadian north, and in developing countries.

Following these early activities, propulsion system power was more fully explored in a number of schemes to augment wing lift for STOL (**Fig. 1**). Two concepts predominated in the Air Force advanced medium STOL transport competitions. The externally blown flap concept (Fig. 1*e*), used in the YC-15, employs the engine exhaust blowing through, and being turned down by, the flap system to augment the lift. The upper-surface-blowing concept (Fig. 1*f*), used in the YC-14, employs engines mounted above the wing with their exhaust flattened and spread so that it enhances flap effectiveness and is turned downward by the flap. Both of these concepts were flight-tested and demonstrated significant reduction in field length. An augmentor wing aircraft (Fig. 1*d*) was successfully flight-tested, further demonstrating the attributes of power to provide STOL capability.

Subsequently, STOL technology has been applied to explore further improvements in transport takeoff and landing performance as well as potential applications to fighter aircraft. The upper-surface-blowing concept has been employed in a number of flight-test and operational aircraft. This concept has a noise advantage because placing the engine exhaust above the wing shields people on the ground from one of the primary noise sources. The quiet short-haul research aircraft (QSRA) has been used in acoustic, aerodynamic, stability and control, and terminal area investigations. Flight tests of the ASKA (**Fig. 2**) began in Japan in 1987; the Russian Antonov An 74 (derived from the An 72) is operational.

The externally blown flap propulsive-lift concept is used in the U.S. Air Force C-17 transport aircraft. The C-17 has the capability of airlifting large payloads over long distances and delivering them into simple, rough fields. Using technology proven in the YC-15 program, the aircraft has a wing span of 165.0 ft (50.3 m) and a maximum design takeoff gross weight of 572,000 lb (259,457 kg). *See* MILITARY AIRCRAFT.

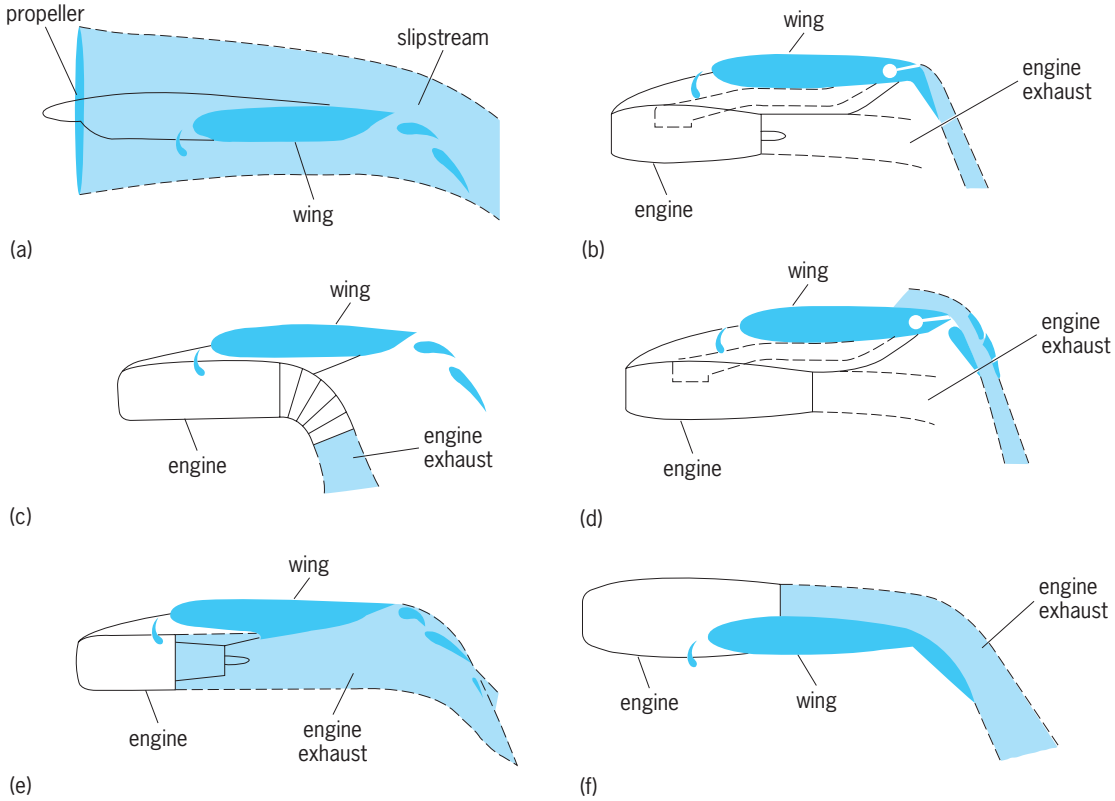

Fig. 1. Propulsive-lift concepts for STOL transports. (*a*) Propeller slipstream. (*b*) Internally blown flap. (*c*) Vectored thrust. (*d*) Augmentor wing. (*e*) Externally blown flap. (*f*) Upper-surface blowing.

Fig. 2. Japanese ASKA upper-surface-blowing transport aircraft.

Under a U.S. Air Force program, an F-15 fighter aircraft was modified to include movable canards, thrust-vectoring/thrust-reversing two-dimensional nozzles, and an advanced, four-channel digital, fly-by-wire flight/propulsion control system. Following 43 flights with standard circular engine nozzles, the aircraft first flew with the two-dimensional nozzles in 1989. The program is to flight-test technologies to provide fighter aircraft with the ability to land on wet, short, bomb-damaged runways. *See* AIRPLANE.

W. P. Nelms

Bibliography. J. A. Cochran et al., *Quiet Short-Haul Research Aircraft: The First Three Years of Flight Research*, Amer. Inst. Aeron. Astronaut. Pap. 81-2625, 1981; D. D. Few, *A Perspective on 15 Years of Proof-of-Concept Aircraft Development and Flight Research at Ames-Moffett by the Rotocraft and Powered-Lift Flight Projects Division, 1970–1985*, NASA Ref. Pub. 1187, August 1987; *Jane's All the World's Aircraft*, published periodically; *Proceedings of the International Powered Lift Conference*, Santa Clara, California, December 7–10, 1987; D. W. Riddle et al., *Quiet Short-Haul Research Aircraft: A Summary of Flight Research since 1981*, Soc. Automotive Eng. Pap. 872315, 1987.

Shrew

A mammal of the family Soricidae (previously Order Insectivora, now listed as Soricomorpha by some). Shrews are found in Asia, Africa, Europe, North America, and northern South America. The family includes 376 species in 26 genera grouped into three subfamilies: the red-toothed shrews (Soricinae), the white-toothed shrews (Crocidurinae), and the African white-toothed shrews (Myosoricini).

Subfamilies. The red-toothed shrews (Soricini) include 14 genera and 148 species and occur throughout much of the major part of the world. The largest genus is *Sorex*, with 77 species. All of the shrews of North America (five genera) are in the Soricini (see **illustration**). Three of the genera (*Sorex*, *Notiosorex*, and *Megasorex*) have long tails, and two (*Blarina* and *Cryptotis*) have short tails that are a little longer that the hind foot. Some of the better known species of red-toothed shrews are the short-tailed shrew (*Blarina brevicauda*), least shrew (*Cryptotis parva*), masked shrew (*Sorex cinereus*), water shrew (*S. palustris*), and trowbridge shrew (*S. trowbridgei*) of the New World, and the European common shrew (*S. araneus*) and European pygmy shrew (*S. minutus*) of the Old World.

White-toothed shrews (Crocidurinae) include 9 genera and 210 species. They are found only in the Old World and are the dominant Old World shrews. The two largest and best known genera are *Crocidura* (172 species) and *Suncus* (18 species).

The African white-toothed shrews (Myosoricine) include 3 genera and 18 species, with the largest genus being Myosorex (14 species).

Morphology and predation. At least two shrews are among the smallest mammals in the world; these are the European pygmy shrew (*Suncus etruscus*) and southern Indiana populations of the American pygmy shrew (*Sorex hoyi*), in which the individuals in some populations weigh about two grams. Another contender for the world's smallest mammal is the bumblebee bat of Asia. The largest shrew is the Asian house shrew, *Suncus murinus*. It weighs about 100 g.

The dental formula for most shrews is I 3/1, C1/1, P3/1, and M3/3 × 2, for a total of 32 teeth, except for *Cryptotis*, which has one less upper premolar (2/1), thus 30 teeth. Shrews have the first upper and lower incisors enlarged, thereby effectively forming a set of tweezers with which they can capture prey and extract them from small places. In addition, the teeth behind the first incisors are small and not well differentiated. There are four or five of these teeth (called unicuspids), which are helpful in shrew classification and identification. The milk teeth are resorbed prior to birth; thus shrews are born with their adult teeth. *See* DENTITION.

Shrews are small secretive animals that constantly run about searching for food. The eyes are minute,

Dusky shrew (*Sorex obscura*), native to western North America. (*Dr. Lloyd Glenn Ingles © California Academy of Sciences*)

the snout is sharp-pointed, and the ears are very small (see illustration). Musk glands on the flank region exude an odor, possibly having a protective function. However, the musk glands do not prevent shrews from being preyed upon by owls, reptiles, and various mammals, including house cats. (When house cats capture the shrews they often do not eat them, presumably because of the bad smell and taste.) The main function of the scent glands in some shrews, at least, is to mark tunnels for reproductive purposes. The glands are much more developed in males than females. Males searching for mates will enter burrows with little odor (that is, those of females) but will avoid those with male odor.

Mating, reproduction, and development. Shrews are solitary and establish distinct territories. Males and females are found together only during mating, which occurs when they are about 1 year old. The female builds a nest, usually in an underground burrow. The nest is lined with leaves and moss. She deposits her litter of about 4–10 young, born after a gestation period of approximately 3 weeks, with usually 3 or 4 litters produced during the season. The young are nearly full size when they leave the nest, mate in their second year, and die during the fall after the reproductive season, or at around 15 months. The fall and winter population is mostly of juveniles and subadults.

Diet and eating habits. Because of their small size and the fact that they lose heat rapidly, these animals have a very high metabolism. They alternate between short (3 h) stretches of feeding and sleeping throughout the 24 h, and eat approximately their own weight in food per day. During winter they do not hibernate, but continue to actively search for food. Shrews are, or were, classed in the order Insectivora, but they do not feed entirely on insects. Although they are often said to feed on whatever is available, in fact they exhibit quite a bit of selectivity. Major foods of some of the larger shrews are earthworms, snails, and centipedes, whereas small ones often feed on small insect larvae. In addition, many shrews feed on hypogenus fungi such as *Endogone*. Some shrews (such as *Blarina* and *Sorex*) have poisonous saliva which is used to overcome prey. Some shrews store food items, such as earthworms and snails. These remain alive but are paralyzed, and thus remain fresh and at the site until eaten.

Novel traits. A few shrews, such as the water shrew, *Sorex palustris*, European water shrew, *Neomys fodiens*, and Tibetian water shrew (*Nectogale elegans*), are aquatic. The latter species is the only shrew with webbed feet. Other aquatic shrews have toes lined with a fringe of hairs. At least some shrews have echolocation, which they use to negotiate their burrow systems, and many also have various high pitched squeaks and twitterings, probably used in communication. Another interesting behavior of at least some shrews is caravanning, in which the young follow the mother by each shrew grasping the one ahead by the base of the tail. Some shrews (*Sorex cinereus*, *Crocidura*) practice coprophagy—or "fecal feeding"—in which partially digested food and fat globules are extruded from the rectum and eaten; it is thought that shrews obtain nutrients from this process that otherwise would not be digested. Shrews are generally solitary and antagonistic to one another, yet the least shrew, *Cryptotis parva*, has been found in groups, with one of the groups numbering 31 individuals.

Another interesting trait of shrews is that they tend to partition their habitat such that any one habitat contains an assemblages of different sized species, which feed on different sized prey and thereby reduce competition. Overall, shrews are a poorly known, but very interesting group of animals deserving of much further study. *See* MAMMALIA.

John O. Whitaker, Jr.

Bibliography. S. Churchfield, *A Natural History of Shrews*, Cornell University Press, Ithaca, 1990; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Facts on File Publications, New York, 1984; R. M. Nowak, *Walker's Mammals of the World* (3 vols.), Johns Hopkins University Press, Baltimore, 1999; J. O. Whitaker, Jr., and W. J. Hamilton, Jr., *Mammals of the Eastern United States*, Cornell University Press, Ithaca, 1998; D. E. Wilson and D. M. Reeder, *Mammal Species of the World. A Taxonomic and Geographic Reference*, 3d ed., in press; D. E. Wilson and S. Ruff, *The Smithsonian Book of North American Mammals*, American Society of Mammalogists, 1999.

Shrink fit

A fit that has considerable negative allowance so that the diameter of a hole is less than the diameter of a shaft that is to pass through the hole, also called a heavy force fit. Shrink fits are used for permanent assembly of steel external members, as on locomotive wheels. The difference between a shrink fit and a force fit is in method of assembly. Locomotive tires, for instance, would be difficult to assemble by force, whereas a shaft and hub assembly would be convenient for force fit by a hydraulic press. In shrink fits, the outer member is heated, or the inner part is cooled, or both, as required. The parts are then assembled and returned to the same temperature. *See* ALLOWANCE; FORCE FIT.

Paul H. Black

Shunting

The act of connecting an electrical element in parallel with (across) another element. The shunting connection is shown in **illus.** *a*.

An example of shunting involves a measuring instrument whose movement coil is designed to carry only a small current for a full-scale deflection of the meter. To protect this coil from an excessive current that would destroy it when measuring currents that exceed its rating, a shunt resistor carries the excess current.

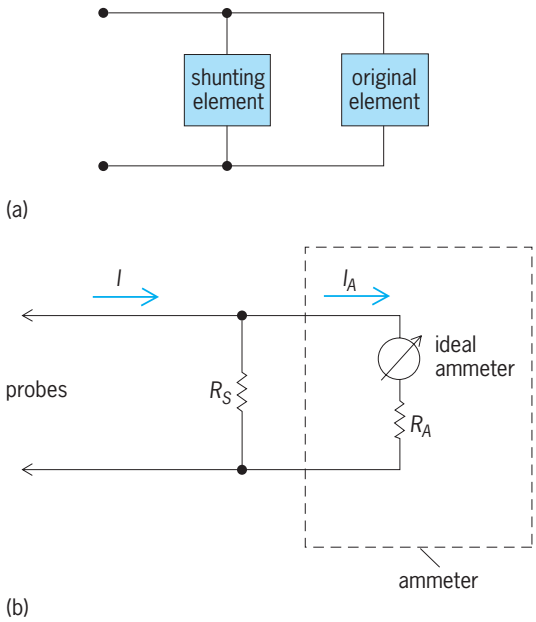

Shunting. (*a*) Shunting connection. (*b*) Ammeter shunted by resistor R_S.

Illustration *b* shows an ammeter (a current-measuring instrument) with internal resistance R_A. It is shunted by a resistor R_S. The current through the movement coil is a fraction of the measured current, and is given by the equation below. With different

$$I_A = \frac{R_S}{R_A + R_S} I$$

choices of R_S, the measuring range for the current I can be changed. *See* AMMETER; CURRENT MEASUREMENT.

Similar connections and calculations are used in a shunt ohmmeter to measure electrical resistance. Shunt capacitors are often used for voltage correction in power transmission lines. A shunt capacitor may be used for the correction of the power factor of a load. In direct current shunt motors, the excitation (field) winding is connected in parallel with the armature. *See* DIRECT-CURRENT MOTOR; OHMMETER; POWER FACTOR; RESISTANCE MEASUREMENT.

In electronic applications, a shunt regulator is used to divert an excessive current around a particular circuit. In broadband electronic amplifiers, several techniques may be used to extend the bandwidth. For high-frequency extension, a shunt compensation is used where, typically, a capacitor is shunted across an appropriate part of the circuit. Shunt capacitors (or more complicated circuits) are often used to stabilize and prevent undesired oscillations in amplifier and feedback circuits. *See* ALTERNATING-CURRENT CIRCUIT THEORY; AMPLIFIER; DIRECT-CURRENT CIRCUIT THEORY; FEEDBACK CIRCUIT.

Shlomo Karni

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; M. Kaufman and A. H. Seidman, *Handbook for Electronics Engineering Technicians*, 2d ed., 1985; J. Markus and C. Weston, *Essential Circuits Reference Guide*, 1988.

Sickle cell disease

An inherited disorder of red blood cells characterized by lifelong anemia and recurrent painful episodes. The sickle cell mutation is caused by a single nucleotide effecting a change in the β-globin gene, resulting in the substitution of valine for glutamic acid as the sixth amino acid of β-globin. The short circulatory survival of red blood cells that contain sickle cell hemoglobin S results in anemia, and their abnormal rigidity contributes to painful obstruction of small blood vessels. *See* ANEMIA; GENETIC CODE; HEMOGLOBIN.

Genetics. The sickle cell gene is found most commonly among individuals of African ancestry, but also has a significant incidence in Mediterranean, Middle Eastern, and Asian Indian populations. Its evolutionary persistence in these populations is related to its ability to partially protect heterozygous individuals from death due to malaria. The geographic dispersion of this gene is related to its having arisen as a spontaneous mutation at least four different times and to its transmission along ancient trade routes.

Inheritance of one sickle gene and one normal β-globin allele results in a simple heterozygous condition known as sickle cell trait. This benign carrier condition is associated with a normal life expectancy, and it does not cause either anemia or recurrent pain. However, carriers may occasionally experience blood in their urine and be more subject to the risk of sudden death when engaging in rigorous prolonged physical conditioning. The large amounts of hemoglobin A within sickle-cell-trait red blood cells protect against the deleterious effects of hemoglobin S. The inheritance of homozygous sickle cell anemia results in sufficiently high intracellular concentration of sickle cell hemoglobin S to cause clinical disease. *See* HUMAN GENETICS.

Pathophysiology. The pathophysiology of sickle cell disease has three aspects: molecular, cellular, and rheological. The property of sickle cell hemoglobin S responsible for clinical disease is its insolubility when deoxygenated. Oxygenated sickle cell hemoglobin S is as soluble as oxygenated normal hemoglobin, but when it is deoxygenated it aggregates and forms an insoluble polymer.

The effects of sickle cell hemoglobin on sickle red blood cells is catastrophic in terms of their circulatory survival and deformability. While normal red blood cells circulate for 4 months, sickle cells survive in the circulation for only about 12–20 days. Polymerization of sickle cell hemoglobin within deoxygenated sickle cells reversibly reduces cellular deformability and distorts cells to the sickle shape (see **illus.**). Sickle cells usually return from the venous circulation to the arterial, where the hemoglobin is reoxygenated and the cells unsickle. The so-called vicious cycle of erythrostasis that has been proposed to account for the episodic vascular

occlusion of sickle cell disease suggests that transient stasis of blood flow results in deoxygenation of sickle cells and the generation of an acidic milieu, both of which make possible the polymerization and sickling responsible for recurrent pain and organ dysfunction. Initiation of vascular occlusion, however, appears to be the result of abnormally sticky young sickle cells adhering to the vascular endothelium.

Persistent cycles of sickling and unsickling result in the generation of dehydrated, very dense sickle cells; these are irreversibly sickled cells that are incapable of resuming a normal shape when reoxygenated. Their generation is the result of abnormal cell losses of cations and water via potassium-chloride cotransport and calcium-sensitive potassium efflux, called the Gardos pathway.

As a result of the poor deformability of individual sickle red blood cells, sickle cell blood has high viscosity. The impaired rheologic properties of sickle blood are compounded by abnormal adherence of sickle red cells to endothelial cells lining the blood vessels.

Clinical manifestations. The short-lived nature of sickle red blood cells results in lifelong chronic hemolytic anemia with which accelerated red blood cell production cannot keep pace. The increased turnover of red blood cells results in elevated levels of hemoglobin degradation and bilirubin production by the liver and in very frequent formation of gallstones. Individuals with sickle cell disease are also afflicted with episodes of increased severity of anemia, usually due to cessation of erythropoiesis. Vasoocclusive complications of sickle cell disease include the episodic painful crises and both chronic and acute organ dysfunction, for example, strokes, painful necrosis of bone, loss of vision, diminished function of the liver and kidneys, blood in the urine, and poorly healing ulcers of the legs. Older individuals frequently have a degree of heart failure that may be related to both anemia and vascular occlusion. Individuals with sickle cell disease are generally regarded as having severely shortened life expectancy, but survival has improved to the extent that average life expectancy is in the fifth decade. One disease manifestation that is particularly problematic in young children is susceptibility to infections. Prior to the use of prophylactic antibiotics in infants and toddlers, septicemia was the leading cause of death in this age group.

Diagnosis. The standard method of diagnosing sickle cell syndromes is hemoglobin electrophoresis. The replacement of negatively charged glutamic acid by neutrally charged valine results in sickle cell hemoglobin S having a different electrophoretic mobility from normal hemoglobin. On electrophoresis the hemoglobin of subjects with homozygous sickle cell anemia is virtually all sickle cell hemoglobin, while that of individuals with sickle cell trait includes both normal and sickle cell hemoglobin. Another method of detecting sickle cell hemoglobin is solubility testing in a high-ionic-strength solution, a test that will be positive in sickle trait as well as in sickle cell disease. A simple diagnostic tool is to review the peripheral blood smear microscopically for the presence of irreversibly sickled cells. These cells are almost always found in individuals with sickle cell disease but are absent in those with sickle trait. *See* ELECTROPHORESIS.

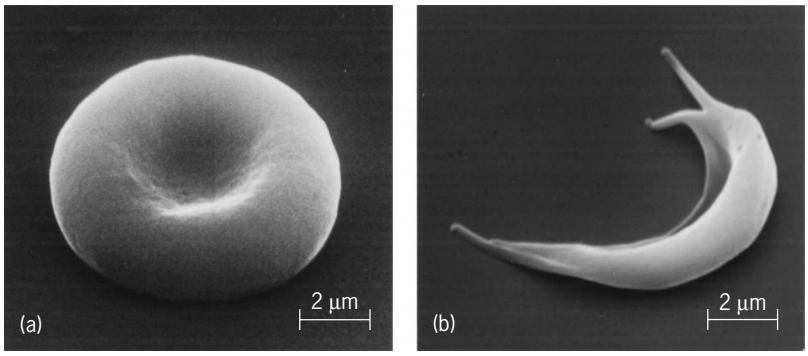

Scanning electron micrographs of (*a*) an oxygenated sickle red blood cell and (*b*) a deoxygenated sickle red blood cell. The appearance of the oxygenated sickle cell is indistinguishable from a normal red blood cell. (*Electron micrographs by Dr. James White*)

There are a variety of diagnostic tests based on deoxyribonucleic acid (DNA). These tests depend on the ability to detect substitution of the single nucleotide responsible for the sickle gene in native genomic DNA or in DNA that has been amplified enzymatically from genomic DNA using the polymerase chain reaction. These DNA-based diagnostic methods are particularly useful for prenatal diagnosis of sickle cell disease, where often only small amounts of DNA for diagnosis are obtained by amniocentesis or chorionic villus sampling. *See* DEOXYRIBONUCLEIC ACID (DNA); PRENATAL DIAGNOSIS.

Therapy. Despite the profound understanding of sickle cell disease, treatment of painful episodes often consists of only symptomatic therapy, including analgesics for pain, antibiotics for infections, and transfusions for episodes of severe anemia. Previous attempts at inhibiting polymerization of sickle cell hemoglobin have been unsuccessful because of intolerable side effects or lack of beneficial effect. Genetic counseling and prenatal diagnosis remain important therapeutic approaches. *See* BLOOD; GENETIC ENGINEERING.

Stephen H. Embury

Bibliography. H. F. Bunn and B. G. Forget, *Hemoglobin: Molecular, Genetic and Clinical Aspects*, 1986; W. A. Eaton and J. Hofrichter, Sickle cell hemoglobin polymerization. *Adv. Prot. Chem.*, 40:63–279, 1990; S. H. Embury et al., *Sickle Cell Anemia: Scientific Principles and Clinical Practice*, 1994; R. P. Hebbel, Beyond hemoglobin polymerization: The red blood cell membrane sickle cell disease pathophysiology, *Blood*, 77:214–237, 1991; G. R. Serjeant, *Sickle Cell Disease*, 3d ed., 2001.

Sideband

The range of the electromagnetic spectrum located either above (the upper sideband) or below (the lower sideband) the frequency of a sinusoidal

carrier signal $c(t)$. The sidebands are produced by modulating the carrier signal in amplitude, frequency, or phase in accordance with a modulating signal $m(t)$ to produce the modulated signal $s(t)$. The resulting distribution of power in the sidebands of the modulated signal depends on the modulating signal and the particular form of modulation employed. *See* AMPLITUDE MODULATION; FREQUENCY MODULATION; PHASE MODULATION.

Amplitude-modulation systems. In conventional analog amplitude modulation (AM), changing the amplitude of a carrier signal of frequency f hertz in accordance with the amplitude of a modulating signal whose power spectrum extends from 0 to B hertz produces an upper sideband extending from f to $f + B$ hertz and a lower sideband extending down from f to $f - B$ hertz. Except for this shift in frequency, the distribution of power in the two sidebands equals that of the modulating signal, with one sideband being the mirror image of the other about the carrier frequency. The total power of the modulated signal is the sum of the power in the carrier signal at f hertz and the power in the modulating signal, the latter being evenly divided over the two sidebands.

Modifications. Several modifications of conventional analog amplitude modulation are used in practice. Since the carrier signal conveys no information about the modulating signal, it can be eliminated. This so-called double sideband–suppressed carrier (DSB-SC) modulation reduces the total amount of transmitted power in the modulated signal. In addition, either sideband contains all the information about the modulating signal, so that it suffices to transmit only one of them. This single-sideband (SSB) modulation therefore reduces the required channel bandwidth by a factor of 2. In a modification of single sideband, known as vestigial sideband (VSB), the transmitted signal consists of one complete sideband, the carrier signal, and a small fraction of the other sideband. The latter can be used to simplify the demodulation process relative to the complexity of single-sideband demodulation. *See* SINGLE SIDEBAND.

Digital AM systems. In practical terms the power spectra of digital amplitude-modulation schemes such as on-off keying (OOK) or amplitude-shift keying (ASK), in which the carrier signal is shifted in amplitude between a discrete and finite number of values, are similar to those for analog amplitude modulation. The upper and lower sidebands are usually taken to extend over the range from f to $f + R$ and from f to $f - R$, respectively, where R is the rate at which the amplitude of the carrier signal is shifted.

Angle-modulation systems. In analog angle-modulation systems the power spectral distribution of the modulated signal generally does not exhibit the same symmetry about the carrier frequency as for amplitude modulation. However, if the modulating signal is periodic, the power in the modulated signal exists at discrete and evenly spaced frequencies, with equal amounts of power at pairs of frequencies symmetrically disposed about the carrier frequency. *See* ANGLE MODULATION.

Optimization of power spectrum. In general, the power spectral distribution in the sidebands is a function of the type of modulation employed and the power spectral distribution of the modulating signal. Some modulation schemes, in particular offset quadrature phase-shift keying and minimum-shift keying, offer a distinct advantage over other forms in that most of the power is concentrated in a relatively narrow portion of the sidebands, resulting in reduced transmission bandwidth requirements. *See* MODULATION.

Hermann J. Helgert

Bibliography. B. P. Lathi, *Modern Digital and Analog Communication Systems*, 3d ed., Oxford University Press, 1998; M. Schwartz, *Information Transmission, Modulation, and Noise*, 4th ed., McGraw-Hill, 1990; D. R. Smith, *Digital Transmission Systems*, 2d ed., Van Nostrand Reinhold, 1993.

Sidereal time

One of several kinds of time scales used in astronomy, whose primary application is as part of the coordinate system to locate objects in the sky. It is also the basis for determining the solar time commonly used in everyday living.

Time measurement. Time is measured with respect to some repetitious event. The measure of time is a parameter whose scale and units are adopted for specific purposes at hand. The most common measurements of time are based on the motions of the Earth that most affect everyday life; rotation on its axis, and revolution in orbit around the Sun. Objects in the sky reflect these motions and appear to move westward, crossing the meridian each day. A particular object or point is chosen as a marker, and the interval between its successive crossings of the local meridian is defined to be a day, divided into 24 equal parts called hours. The actual length of the day for comparison between systems depends on the reference object chosen. The time of day is reckoned by the angular distance around the sky that the reference object has moved westward since it last crossed the meridian. In fact, the angular distance west of the meridian is called the hour angle. *See* MERIDIAN.

Sidereal day and year. The reference point for marking sidereal time is the vernal equinox, one of the two points where the planes of the Earth's Equator and orbit appear to intersect on the celestial sphere. The sidereal day is the interval of time required for the hour angle of the equinox to increase by 360°. One rotation of the Earth with respect to the Sun is a little longer, because the Earth has moved in its orbit as it rotates and hence must turn approximately 361° to complete a solar day. A sidereal year is the time required for the mean longitude of the Sun to increase 360°, or for the Sun to make one circuit around the sky with respect to a fixed reference point. *See* EQUINOX.

Astronomical coordinate system. Sidereal time is essential in astronomy and navigation because it is part of the most common coordinate system for

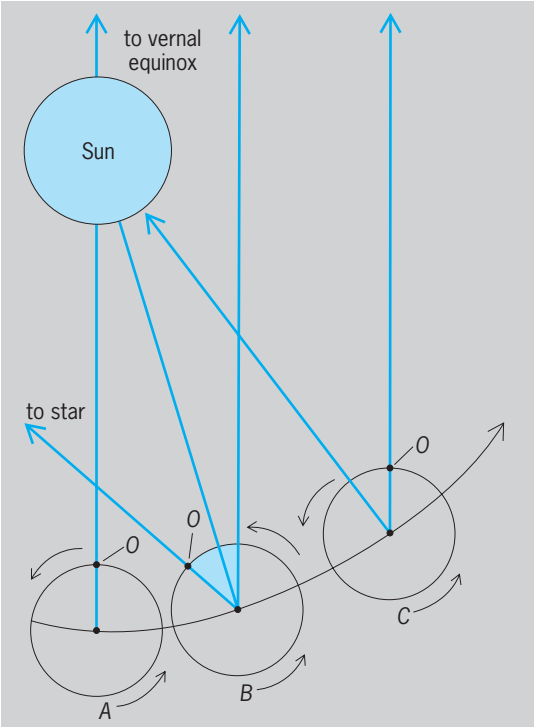

Diagram of the Earth's motions, illustrating the concepts of sidereal time, hour angle, and right ascension. Symbols are explained in text.

describing the position of objects in the sky. These coordinates are called right ascension and declination, analogous to longitude and latitude on the surface of the Earth, with the vernal equinox playing the same role as the Greenwich meridian for the zero point. The local sidereal time is the hour angle of the vernal equinox past (westward of) the local meridian, and also the right ascension of any object on the meridian. By a sidereal clock, a star rises and crosses the local meridian at the same times every day. Greenwich sidereal time is local sidereal time plus the longitude of the observer west of Greenwich. *See* ASTRONOMICAL COORDINATE SYSTEMS.

These concepts are further described in the **illustration**, in which the reader is looking down on the North Pole as the Earth moves counterclockwise in orbit around the Sun. An observer at point *O* is considered. When the Earth is at position *A*, the Sun is taken to be in line with the vernal equinox and overhead for the observer. At position *B*, the Earth has rotated counterclockwise through the shaded angle, which represents the hour angle of the vernal equinox, or the local sidereal time for the observer and also the right ascension of a star overhead for the observer. At position *C*, the Earth has made a complete rotation with respect to the vernal equinox but not with respect to the Sun. The right ascension of the Sun has increased by approximately 1°.

Determination. For a measure of time to be of practical use, determination of it by observation must be possible. Sidereal time, defined by the daily motion of the vernal equinox, is determined in practice from observations of the daily motions of stars. Solar time, which is used in the conduct of everyday life, is not directly observable. It is determined from the sidereal time by a known relationship.

Effect of motion of Earth's axis. Upon more sophisticated examination, the sidereal day is not exactly the period of rotation of the Earth, being shorter by about 0.0084 s. The Earth's axis of rotation moves very slowly but regularly in space because of the gravitational effects of other bodies in the solar system, and the Earth's crust also shifts very slowly and irregularly about the axis of rotation. As a result, neither the equinox nor the local meridian plane is absolutely fixed. Consequently, the relation of sidereal time to the measure of time defined by the rotational motion of the Earth alone is highly complex. *See* EARTH ROTATION AND ORBITAL MOTION; PRECESSION OF EQUINOXES; TIME.

Alan D. Fiala

Bibliography. J. Kaler, *The Ever-Changing Sky*, Cambridge University Press, 1996; D. Morrison, S. Wolff, and A. Franknoi, *Abell's Exploration of the Universe*, 7th ed., Saunders College Publishing, 1995; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, University Science Books, 1992; W. M. Smart and R. M. Green, *Textbook on Spherical Astronomy*, 6th ed., Cambridge University Press, 1977, reprint 1999.

Siderite

A mineral ($FeCO_3$) with the same space group and hexagonal crystal system as calcite ($CaCO_3$). Siderite has a gray, tan, brown, dark brown, or red color, has rhombohedral cleavages, and occasionally may show rhombohedral crystal terminations. It may display curved crystal faces like dolomite ($CaMg[CO_3]_2$), but more commonly is found as massive, compact, or earthy masses. *See* CARBONATE MINERALS; CRYSTAL STRUCTURE; DOLOMITE.

Siderite forms extensive solid solutions with magnesite ($MgCO_3$), rhodochrosite ($MnCO_3$), and less commonly with smithsonite ($ZnCO_3$); with calcite and ankerite ($Ca[Fe,Mg][CO_3]_2$), it has wide miscibility gaps that vary with temperature. Its physical properties are variable, depending on the amounts of magnesium (Mg), manganese (Mn), and zinc (Zn) solid solutions; pure iron(II) carbonate ($FeCO_3$) has a high specific gravity of 3.94, a medium hardness of 3.5–4, and a high index of refraction, 1.88. It is only weakly soluble in dilute hydrochloric acid (HCl) but is very soluble in hot hydrochloric acid and other hot strong acids. It may resemble some sphalerite ([Zn,Fe]S) but yields carbon dioxide (CO_2) and not hydrogen sulfide (H_2S) upon reaction with acids. *See* ANKERITE; CALCITE; MAGNESITE; RHODOCHROSITE; SOLID SOLUTION.

Siderite, a widespread mineral in near-surface sediments and ore deposits, occurs in hydrothermal veins, lead-silver ore deposits, sedimentary concretions formed in limestones and sandstones, and Precambrian banded iron formations that precipitated under acidic conditions. Upon metamorphism it may be oxidized or decarbonated to form the iron

oxide magnetite (Fe_3O_4) at 300–500°C (570–930°F). In the presence of quartz (SiO_2) it reacts to form various ferrous hydroxyl silicates upon metamorphism, including minnesotaite ($Fe_3Si_4O_{10}[OH]_2$), at higher metamorphic grades (250–400°C or 480–750°F), or greenalite ($Fe_3Si_2O_5(OH)_4$) at lower metamorphic temperatures (<250°C or 480°F). The exact metamorphic reaction involving the decomposition of siderite depends on the relative pressures of carbon dioxide, water, and oxygen as well as temperature. Siderite may also alter to various iron oxyhydroxides such as goethite during weathering or low-temperature alteration.

Famous localities for siderite are found in Styria (Austria), Westphalia (Germany), Cornwall (Britain), Wawa (Northern Ontario, Canada), Minas Gerais (Brazil), and Llallagua and Potosi (Bolivia). These and other occurrences have provided locally significant quantities of siderite as iron ore. *See* IRON; MAGNETITE; METAMORPHISM; ORE AND MINERAL DEPOSITS.

Eric J. Essene

Bibliography. L. L. Y. Chang, R. A. Howie, J. Zussman, *Rock-Forming Minerals,* vol. 5B: *Non-Silicates: Sulphates, Carbonates, Phosphates, Halides*, 1995; C. Klein and C. S. Hurlbut, Jr., *Manual of Mineralogy*, 21st ed., 1993, rev. 1999; R. J. Reeder (ed.), *Carbonates: Mineralogy and Chemistry*, Rev. Mineral. 11, 1983.

Siderophores

Low-molecular-mass (generally 500–1000 daltons) molecules that have a high specificity for chelating or binding iron. Siderophores are produced by many microorganisms, including bacteria, yeast, and fungi, to obtain iron from the environment. More than 500 different siderophores have been identified from microorganisms. Some bacteria produce more than one type of siderophore. *See* BACTERIA; BIOINORGANIC CHEMISTRY; CHELATION; FUNGI; IRON; YEAST.

Iron is required by aerobic bacteria and other living organisms for a variety of biochemical reactions in the cell, including production of adenosine triphosphate (ATP) for energy and for synthesis of deoxyribonucleic acid (DNA). Although iron is the fourth most abundant element in the Earth's crust, it is not readily available to bacteria. Iron is found in nature mostly as insoluble precipitates that are part of hydroxide polymers. Bacteria living in the soil or water must have a mechanism to solubilize iron from these precipitates in order to assimilate iron from the environment. Iron is also not freely available in humans and other mammals. Most iron is found intracellularly in heme proteins and ferritin, an iron storage compound. Iron outside cells is tightly bound to proteins such as transferrin in serum and lactoferrin on mucosal surfaces. Therefore, bacteria that grow in humans or other animals and cause infections must have a mechanism to remove iron from these proteins and use it for their own energy and growth needs. Siderophores have a very high affinity for iron and are able to solubilize and transport ferric iron (Fe^{3+}) in the environment and also to compete for iron with mammalian proteins such as transferrin and lactoferrin. The majority of bacteria and fungi use siderophores to solubilize and transport iron. Microorganisms can use either siderophores produced by themselves or siderophores produced by other microorganisms. *See* ADENOSINE TRIPHOSPHATE (ATP); DEOXYRIBONUCLEIC ACID (DNA); IRON METABOLISM.

Structure. The many different types of siderophores can generally be classified into two structural groups, hydroxamates and catecholate compounds. Despite their structural differences, all form an octahedral complex with six binding coordinates for Fe^{3+} (**Fig. 1**). Siderophores can be detected in bacterial cultures using chemical assays, bioassays, and functional assays. There are specific chemical assays for hydroxamate and catecholate groups. However, these assays do not work for all siderophores because some siderophores do not have the appropriate functional group. Bioassays use selected strains of bacteria that require addition of siderophores for growth in a low-iron medium. Cultures or extracellular fluids of siderophore-producing bacteria can often promote growth of these strains. The bioassay strains, however, do not detect all siderophores. The chrome azurol S assay is a functional assay that is often used to detect siderophores. Chrome azurol S is a blue compound that binds iron but changes to an orange color when iron is removed from it by siderophores. The chrome azurol S assay is independent of siderophore structure and has made it relatively easy to detect siderophores in many microorganisms. *See* BIOASSAY; CULTURE.

Uptake by bacteria. Gram-negative bacteria have a cell envelope consisting of an outer membrane and an inner or cytoplasmic membrane; these membranes are separated by the periplasm. The outer membrane contains proteins called porins that form

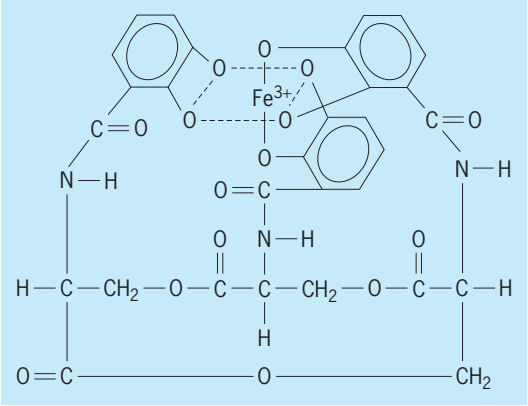

Fig. 1. Structure of ferric enterobactin, a catechol-type siderophore produced by many enteric bacteria. Three molecules of enterobactin bind each molecule of Fe^{3+} to provide the six binding coordinates. (*Adapted from J. Biol. Chem., 270:26723–26726, 1995*)

water-filled channels. Small substrates with masses of up to 600 daltons, including sugars, phosphates, and nucleosides, can diffuse through these channels in the membrane. However, iron-siderophore complexes are too large to diffuse through these channels. Therefore, gram-negative bacteria produce specific proteins in the outer membrane that function as receptors for iron-siderophore complexes. These receptor proteins have a very high affinity for specific siderophores and concentrate the iron-siderophore complex at the cell surface. These complexes are then translocated into the periplasm of the cell by an active transport mechanism. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; CELL (BIOLOGY); CELL MEMBRANES; CELL PERMEABILITY; PROTEIN.

Siderophore receptors, proteins ranging in mass from 70,000 to 90,000 daltons, are produced by bacteria when iron is not readily available. Bacteria may produce receptors for several different siderophores, including those produced by themselves and those produced by other bacteria that might be present in their environment. It is not unusual for bacteria to produce six or more different siderophore receptors. These receptors enable them to compete with other microorganisms for iron and provide a competitive advantage that allows these bacteria to grow in nature or in a mammalian host. Bacteria even have receptors for siderophores produced by fungi. For example, *Escherichia coli, Pseudomonas aeruginosa*, and *Yersinia enterocolitica* are bacteria that cause diseases in humans. These bacteria all have receptors for ferrichrome siderophores produced by fungi. *See* ESCHERICHIA; PSEUDOMONAS; YERSINIA.

In addition to specific siderophore receptors, other proteins are involved in translocating iron from the surface of the bacteria into the cytoplasm, where it is needed for various biochemical functions (**Fig. 2**). Proteins known as TonB, ExbB, and ExbD form a complex in the inner membrane. TonB extends from the inner membrane through the periplasm and along with ExbB and ExbD provides energy to the siderophore receptor, located in the outer membrane, and facilitates transport of the iron-siderophore complex into the periplasm. Periplasmic iron-binding proteins carry the iron through the periplasm to the inner membrane. ATP-binding proteins and permeases then transport the iron into the cytoplasm. In the cytoplasm, ferric iron, Fe^{3+}, is reduced to ferrous iron, Fe^{2+}, for use by the cell.

Although gram-positive bacteria do not have an outer membrane or specific siderophore receptor proteins, they do have proteins that are similar to the periplasm iron-binding proteins that directly bind the iron-siderophore complex. Gram-positive bacteria also have a TonB, ExbB and ExbD complex, and the inner-membrane permease and ATP-binding proteins involved in iron transport.

Iron toxicity. Although iron is very important for bacterial metabolism, too much iron can be toxic to cells. Free Fe^{2+} in bacteria and other cells can act as a catalyst to generate hydroxyl radicals (·OH) from hydrogen peroxide (H_2O_2). Generation of hydroxyl radicals from superoxide ($O_2{\cdot}^-$) and hydrogen peroxide is known as the Haber-Weiss cycle. Free hydroxyl radicals can cause damage to the cell, including damage to DNA. Therefore, bacteria tightly control the amount of iron taken into the cells to ensure that there is sufficient iron for cell growth and replication, but not an excess amount which would be toxic and cause cell death. Bacteria control iron levels in the cell by regulating the amount of siderophores and receptors produced. *See* ENZYME; HYDROXYL.

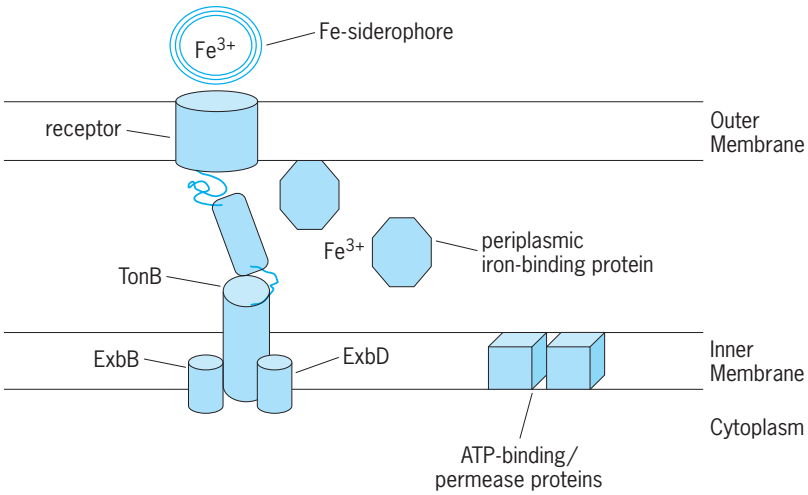

Fig. 2. Siderophore-iron complexes bind to a specific receptor in the outer membrane of gram-negative bacteria. The TonB and Exb proteins (ExbB and ExbD) form a complex that provides energy to the receptor to facilitate transport of Fe^{3+} into the periplasm. Periplasmic iron-binding proteins carry iron through the periplasm to the cytoplasmic membrane. ATP-binding proteins transfer Fe^{3+} into the cytoplasm, where it is needed for bacterial growth and metabolism.

Regulation of siderophore uptake. Siderophores and their specific receptors are produced only when bacteria sense that they are in an iron-deprived environment. Gram-negative bacteria control expression of the genes for siderophore synthesis and receptor expression with a protein called Fur (standing for ferric uptake regulation). When iron levels in the cell are sufficient for the needs of the bacteria, Fe^{2+} binds to Fur. The Fe-Fur complex binds to DNA upstream of the siderophore and receptor genes, and acts as a repressor of transcription so that these genes stop being expressed. Gram-positive bacteria also have proteins functionally similar to Fur called ideR (iron-dependent regulation) that repress siderophore production. Careful control of siderophore uptake systems guards against iron toxicity. *See* BACTERIAL GENETICS.

Some bacteria also have additional mechanisms that turn on production of specific siderophore receptors only when that siderophore is present. These regulatory processes consist of complex sensing mechanisms and transcriptional activator genes. These systems help conserve energy in bacteria that are capable of producing several different siderophore receptors, preventing the unnecessary

synthesis of receptors for siderophores not present in their environment.

Bacterial disease. Bacteria that infect humans and animals and cause disease symptoms must have a mechanism to remove iron from host iron proteins such as transferrin, lactoferrin, or heme. There are two mechanisms that bacteria use to get iron. Some bacteria, for example *Neisseria meningitidis*, bind transferrin, lactoferrin, or heme to specific receptors on their outer membrane. These receptors are similar to siderophore receptors. Iron is removed from the bound host proteins and internalized by a process that also requires the TonB complex. These bacteria cannot grow outside humans or animals because there is no transferrin or lactoferrin in soil or natural environments. However, most pathogenic bacteria do not directly bind host iron proteins, but produce siderophores and siderophore receptors that can remove iron from transferrin, lactoferrin, or precipitates in the environment. These bacteria can also live in soil and water, where host proteins are not present. *See* INFECTIOUS DISEASE; MENINGITIS.

There are many studies performed on culture media that demonstrate that siderophores promote growth of pathogenic bacteria in iron-starvation conditions. Mutant strains of bacteria that do not produce siderophores, siderophore receptors, or components of the uptake process such as TonB are not able to grow in these iron-restricted conditions. In addition, animal model systems have shown that siderophore production and uptake is required for virulence in some bacteria. *See* BACTERIAL GROWTH; BACTERIOLOGY.

Pseudomonas aeruginosa and *Burkholderia cepacia* cause chronic respiratory infections, and *P. aeruginosa* also causes burn-wound infections. In experimental animal infection studies involving these bacteria, siderophores have been shown to be required for virulence. *Yersinia pestis*, which causes bubonic plague, has been shown to require production of the siderophore yersiniabactin for virulence in mice. Other types of *Yersinia* that cause diarrhea and intestinal disease also produce this siderophore. The siderophore aerobactin has been shown to promote growth and virulence of *E. coli* and several other gram-negative pathogens involved in human and animal infections. Most of the invasive strains of *E. coli* that cause septicemia produce aerobactin. *Vibrio* species cause a variety of diseases in humans and fish and produce related siderophores. *Vibrio cholerae* causes the intestinal disease cholera and *V. vulnificus* causes septicemia and wound infections, respectively. *Vibrio anguillarium* causes fatal septicemia in salmon and is a serious problem in the fish industry. Siderophores are required for the virulence of *V. vulnificus* and *V. anguillarium*, and probably *V. cholerae*. Not as much is known about the role of siderophores in bacterial plant infections, but the siderophore chrysobactin contributes to the virulence of *Erwinia chrysanthemi*, a pathogen that causes disease in a wide range of plants. *See* CHOLERA; CLINICAL MICROBIOLOGY; FOOD POISONING; MEDICAL BACTERIOLOGY; PLAGUE; VIRULENCE; YERSINIA.

Biocontrol. Some soil and water bacteria have been found to naturally promote growth of plants. These bacteria aggressively colonize the roots of plants and prevent the growth and inhibit the virulence of many species of fungi as well as other bacterial species, including *Erwinia* and *Fusarium*. These pathogens destroy root hairs and root tips, resulting in root rot, wilting, and stunted growth. Production of siderophores by beneficial bacteria that bind iron and make it unavailable to the pathogenic bacteria or fungi is one mechanism by which bacteria can promote the growth of plants. The growth-promoting bacteria produce specific siderophores that cannot be used by the pathogenic organisms because they lack the appropriate siderophore receptor. The pathogens are therefore not able to compete for necessary iron in the environment of the plant root system. Bacterial species that promote growth of plants include *Pseudomonas*, *Rhizobium*, *Burkholderia*, *Alcaligenes*, and *Agrobacterium*. There is considerable interest in the agriculture industry in using bacteria as natural biocontrol agents instead of chemicals to inhibit growth of plant pathogens. *See* MICROBIAL ECOLOGY; PLANT PATHOLOGY; RHIZOSPHERE; ROOT (BOTANY).

Clinical uses. Siderophores have potential applications in the treatment of some human diseases and infections. Some siderophores are used therapeutically to treat chronic or acute iron overload conditions in order to prevent iron toxicity in humans. Individuals who have defects in blood cell production or who receive multiple transfusions can sometimes have too much free iron in the body. However, in order to prevent infection during treatment for iron overload, it is important to use siderophores that cannot be used by bacterial pathogens.

A second clinical application of siderophores is in antibiotic delivery to bacteria. Some gram-negative bacteria are resistant to antibiotics because they are too big to diffuse through the outer-membrane porins. However, siderophore-antibiotic combination compounds have been synthesized that can be transported into the cell using the siderophore receptor. The antibiotic is delivered into the cell, where it can act on its target to either inhibit growth or kill the bacterial pathogen. This approach to development of new antibiotics is under investigation. *See* ANTIBIOTIC; DRUG RESISTANCE.

Pamela A. Sokol

Bibliography. V. Braun and H. Killmann, Bacterial solutions to the iron-supply problem, *Trends Biochem. Sci.*, 24:104–109, 1999; J. J. Bullen and E. Griffiths (eds.), *Iron and Infection: Molecular, Physiological and Clinical Aspects*, Wiley, New York, 1999; R. B. Byers and J. E. L. Arceneaux, Microbial iron transport: Iron acquisition by pathogenic microorganisms, *Metal Ions Biol. Sys.*, 35:37–66, 1998; D. J. O'Sullivan and F. O'Gara, Traits of *Pseudomonas* spp. involved in suppression of plant root pathogens, *Microbiol. Rev.*, 56:662–676, 1992; C. Ratledge and L. G. Dover, Iron metabolism in pathogenic bacteria, *Annu. Rev. Microbiol.*, 54:881–941, 2000.

Sigma-delta converter

A class of electronic systems containing both analog and digital subsystems whose most common application is to the conversion of analog signals to digital form and vice versa. The device is also known as a delta-sigma converter. The main advantage of the sigma-delta approach to signal conversion is its minimal reliance on the quality of the analog components required. To achieve this end, the system uses pulse density modulation to create a high-rate stream of single-amplitude pulses. For analog-to-digital conversion, the rate at which the pulses are generated depends on the amplitude of the analog voltage being sensed. For digital-to-analog conversion, the pulse density depends on the numeric digital quantity applied at the converter input. *See* PULSE MODULATION.

The simplest implementation of a sigma-delta analog-to-digital converter uses an analog circuit to generate the single-valued pulse stream from an analog source, and a digital system to repeatedly sum the number of these pulses over a fixed number of pulse intervals. The summing operation converts the pulses to a numeric value, achieving analog-to-digital conversion. Conversely, in a digital-to-analog converter, a digital circuit is used to convert numeric values from a digital processor to a pulse stream, and these pulses then are low-pass-filtered by a relatively simple analog system to produce an analog waveform. This low-pass filtering effectively sums the uniform analog pulse amplitudes over a fixed interval. The circuit—analog or digital—used to generate the pulse stream is called a sigma-delta (or delta-sigma) modulator. *See* ELECTRIC FILTER.

In each case, analog information is contained not in the pulse amplitude but in the number of pulses that occur during the conversion interval. This distribution of the analog information makes the conversion process essentially independent of the amplitude of the pulses and greatly simplifies the design and fabrication of the analog portion of the converter. It does, however, require that the sampling process be rapid, since the resolution of the conversion depends on the number of pulses that can exist in the conversion interval. For the simplest sigma-delta converter implementation, for example, if a maximum of n pulses can exist in the conversion interval, only $n + 1$ (from zero to n inclusive) different numeric values can be generated. The high-rate pulse production process, where a multitude of pulses are generated in the conversion interval, is called oversampling and usually is described as the ratio of the number of high-rate pulse intervals to the number of conversion intervals in which they exist. This oversampling ratio determines the maximum resolution of the converter. Converters are characterized by their maximum resolution, and the number of pulses in the conversion interval is usually chosen to be a power of 2. Thus, a 2^n converter is called an n-bit analog-to-digital (or digital-to-analog) converter. An eight-bit converter, for example, has a resolution of one part in two to the eighth power, or 256 levels. Therefore this simplest version of an eight-bit sigma-delta converter needs 256 pulse periods in its basic conversion interval.

The reciprocal of the duration of the conversion interval over which the pulses are counted is called the sampling frequency. This sampling frequency sets the highest rate at which a signal can change and still be represented adequately by a digital data set. The most rapid signal permitted must change from high to low and back to high in exactly two sampling intervals. This highest frequency, exactly one-half the conversion rate, is called the Nyquist frequency. A digital processor in a digital-to-analog converter cannot generate signals higher than this rate. Conversely, if an analog-to-digital converter tries to sample signals higher than this rate, they will look like frequencies lower than the desired one to the converter and will be digitized as such. This undesired frequency translation effect is called aliasing. *See* ELECTRICAL COMMUNICATIONS; INFORMATION THEORY; PULSE MODULATOR.

Sigma-delta modulator. Sigma-delta modulators used to generate a pulse-density-modulated data stream can take a variety of forms, depending on their application. The simplest version, a single-loop sigma-delta modulator (**Fig. 1**), has three essential components: (1) a differencing circuit, which takes the difference between the value of the quantity being converted and a feedback signal from a quantizer; (2) a summing circuit, that adds these

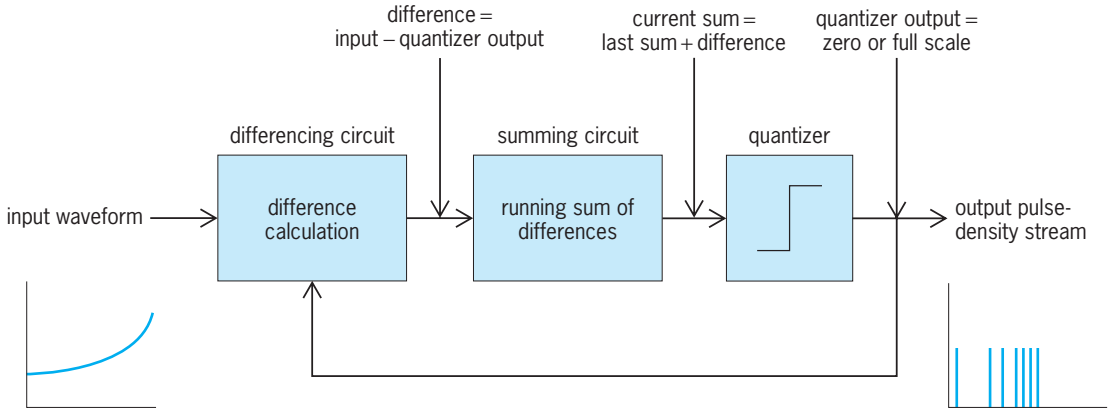

Fig. 1. First-order sigma-delta modulator block diagram.

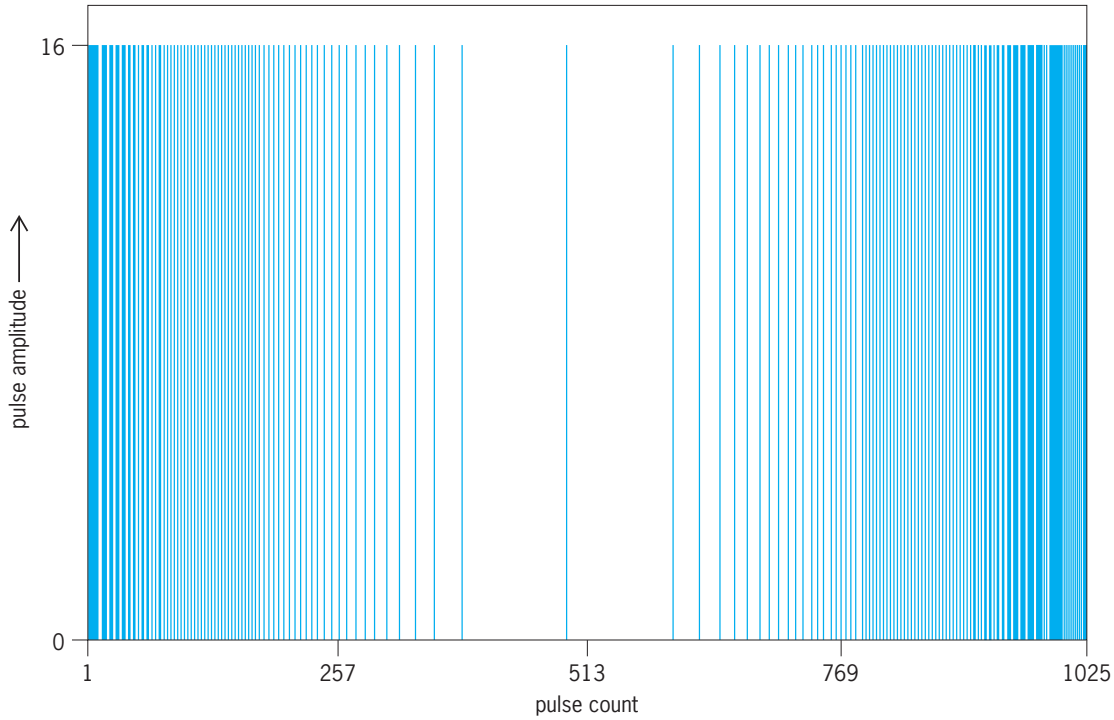

Fig. 2. Pulse density produced for a half-cycle of a periodic input of length 2048.

differences to previously calculated ones; and (3) a quantizer, which produces a signal equal to either the maximum allowable input value or zero. The quantizer output is maximum if the output of the summing circuit is greater than zero, and is zero if the summing output is negative. The input to this sigma-delta model can be considered to be either digital or analog. The modulator's name comes from the conventional mathematical symbols used to represent the two operations that it performs: differencing (delta) and summing (sigma). The 1024-pulse density pattern produced at the quantizer output for one-half cycle of a periodic input waveform is shown in **Fig. 2**.

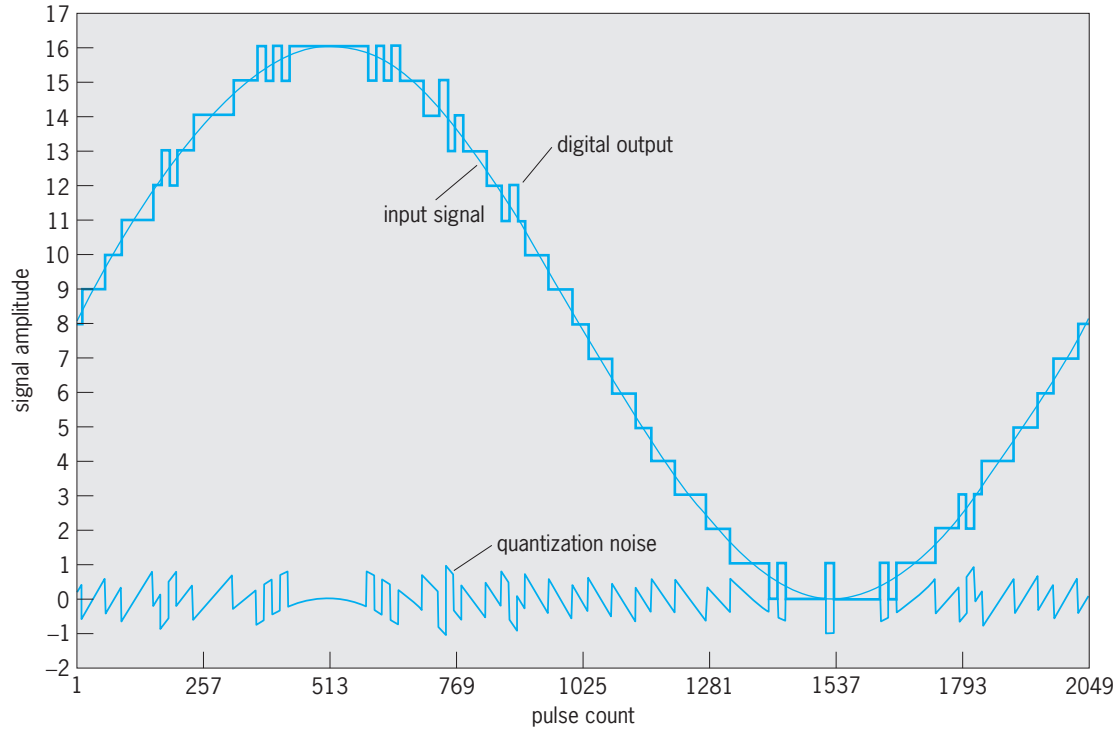

Fig. 3. Digital output produced with a 16-pulse decimation interval.

Low-pass filter, or decimator. To recover the quantized amplitude of the input signal from the pulse-density waveform, the pulses can be averaged over exactly one conversion period. For an analog-to-digital converter, this means counting the number of pulses over the conversion interval and dividing the result by the maximum number of pulses allowed in the interval. The digital filter that performs these operations automatically produces a binary number equal to the pulse count at the end of each conversion interval. The converter produces an output once every conversion interval, or once every 2^n pulses. This process, which produces data at a rate lower than the original pulse rate, is called decimation.

The output produced by a simple averaging decimation filter operating on the output of a single-loop sigma-delta modulator is shown in **Fig. 3**. The oversampling rate for this example is 16. Therefore, with simple averaging decimation, the output signal has just 17 levels (from 0 to 16) with which to represent the input. The digitized input signal is included on the plot, along with a third waveform, the quantization noise introduced by the digitizing process. Practical applications employ much higher resolution than the four-bit resolution of this example, and the quantization noise correspondingly is much lower.

Sigma-delta digital-to-analog conversion. Digital-to-analog conversion based on the pulse density concept is performed in a manner substantially identical to analog-to-digital conversion. In this case, the sigma-delta modulator is implemented as a digital process. A digital quantizer value is subtracted from the numeric value to be converted, and the resulting difference is added to preceding differences in a digital summing, or integrator, circuit. If the integrator produces a negative output, the quantizer output level becomes zero. A positive integrator output, on the other hand, sends the quantizer to its high level. The pulses produced by the modulator in this case are treated as analog signals. A suitable analog low-pass filter then processes these pulses. The net effect of this analog filtering is effectively to add the pulses over an interval that corresponds to the conversion interval used in analog-to-digital conversion.

Many variations on this basic sigma-delta configuration have been investigated and have been used in a multitude of practical applications. The one most commonly encountered in everyday use is the ubiquitous "one-bit digital-to-analog converter" found in audio compact disk players. *See* ANALOG-TO-DIGITAL CONVERTER; COMPACT DISK; DIGITAL-TO-ANALOG CONVERTER.

Philip V. Lopresti

Bibliography. J. C. Candy and G. C. Temes (eds.), *Oversampling Delta-Sigma Data Converters: Theory, Design, and Simulation*, IEEE, New York, 1992; J. A. Cherry and W. M. Snelgrove, *Continuous-Time Delta-Sigma Modulators for High-Speed A/D Conversion: Theory, Practice and Fundamental Performance Limits*, Kluwer, New York, 1999; S. R. Norsworthy, R. Schreier, and G. C. Temes (eds.), *Delta-Sigma Data Converters: Theory, Design, and Simulation*, IEEE, New York, 1996.

Signal detection theory

A theory in psychology which characterizes not only the acuity of an individual's discrimination but also the psychological factors that bias the individual's judgments. Failure to separate these two aspects of discrimination had tempered the success of theories based upon the classical concept of a sensory threshold. The theory provides a modern and more complete account of the process whereby an individual makes fine discriminations.

Modern detection theory was developed by communications engineers to help distinguish weak radar signals from the static in their background. Quite naturally, the first applications in psychology were to discriminations that taxed sensory acuity. The theory has been applied as well to cognitive discriminations, made under limitations imposed, for example, by memory or available information.

Parameters and methods. Quantitative investigations of sensory acuity fall under the heading of psychophysics, the study of relations existing between energy changes in the environment and the attributes of sensation. These attributes include intensity, quality, duration, and extent. Energy changes commonly used as signals in laboratory experiments include flashes of light, bursts of tones, gases directed to the smell receptors, solutions applied to taste buds, and vibrations applied to the skin. *See* PSYCHOPHYSICAL METHODS.

In cognitive psychology signals richer in meaning are used, for example, brief presentations of printed or spoken words. A subject might be asked whether or not the word on a card was on a list of words the subject examined earlier, whether a given statement of opinion most likely appeared in magazine A or magazine B, or whether a particular journal article, represented by its abstract, is relevant or irrelevant to a particular need for scientific information.

When the discrimination called for is between the presence and absence of a specified signal—is the light present? is a word present?—the process is called detection. Distinguishing a particular signal from among a set of possible signals—is the light red or green? which word is being flashed?—is called recognition or identification. In every instance the individual observes a stimulus, which may or may not be a signal, and attempts to choose the response alternative associated with that stimulus. The discrimination is difficult to make because the signal is weak or blurred by some random background interference (noise) or masked by another stimulus or similar to other signals that must be identified.

Components of the theory. The theory of signal detection has two parts of quite different origins. The first comes from mathematical statistics and is a translation of the theory of statistical decisions. The major contribution of this part of the theory is that it permits a determination of the individual's

discriminative capacity, or sensitivity, that is independent of the judgmental bias or decision criterion he may have had when he made the discrimination. The second part of the theory comes from the study of electronic communications. It provides a means of calculating for simple signals, such as tones and lights, the best discrimination that can be attained. The prediction is based upon physical measurements of the signals and their interfering noise.

This opportunity to compare the sensitivity of human observers with the sensitivity of an "ideal observer" for a variety of signals is of considerable usefulness, and of growing interest, in sensory psychology. However, the ability afforded by the first part of the theory to analyze discrimination decisions into two components, thereby separating the influences of discriminative capacity and decision criteria, is of more general relevance to psychology and is the subject of the remainder of this article.

Analysis of discrimination decisions. A simple example will illustrate the main points of the theory. To begin, it may be generally acknowledged that children tend to detect the bell of an ice cream vendor at a greater distance than do their parents. Do children hear better than their parents? That is quite possible; there is a defect of hearing associated with advancing age termed presbycousis. On the other hand, it is also possible that children are more likely to interpret any faint tinkling sound as having been produced by the vendor's bell. That is, children may have a greater bias in favor of making a response like "The ice cream truck!" than their parents do; they may have, in the language of decision theory, a more lenient criterion for making a positive judgment about an ambiguous stimulus of this kind. This is only common sense: One perceives what one wants or expects to perceive. In decision theory lingo, the value of detecting the bell when it is there and the cost of failing to detect the bell when it is there are higher for the child than for the adult. The probabilities also favor a more lenient decision criterion for the child; the child is actively expecting the bell to ring any minute, while the parent is thinking about something else. Then again, the decision criterion of an individual may vary from time to time, in a way independent of changes in his or her acuity. For the child in this example, a cool day, with a lower value of ice cream and a lower probability of the bell being in the neighborhood, may produce a more stringent criterion for a positive response.

Measurement and evolution. It is clear, according to the theory of signal detection, that recording only the proportion of times a signal is detected when it is present, called the proportion of "hits," fails to give an accurate measure of an individual's ability to discriminate between signal presence and absence. This proportion of hits will be higher or lower not solely as a result of differences in acuity, but also as a result of factors unrelated to acuity, such as expectancy and motivation. The theory indicates the need to obtain a second proportion: the proportion of times a signal is detected when it is, in fact, absent. This proportion is called the proportion of "false alarms" and tells a good deal about the individual's response bias or decision criterion. A lenient criterion, of necessity, leads to more false alarms, as well as to more hits, than a strict criterion. This is true in any discrimination problem where the discrimination is less than perfect. *See* MOTIVATION.

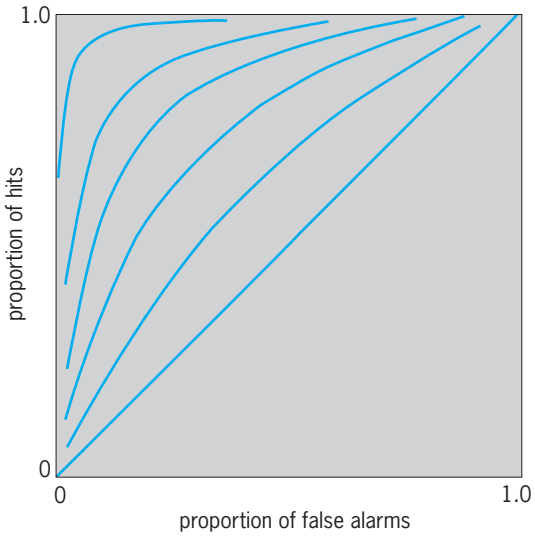

Receiver operating characteristics.

The procedures and analysis suggested by the theory for a simple detection problem are briefly as follows: Many trials are defined; on some a specified signal is present while on others, randomly selected, the signal is absent. From one group of trials to another the subject is encouraged to shift his or her criterion for a positive response, perhaps by changes in the values and costs assigned to the various possible decision outcomes (the payoff matrix) or by changes in the probability of a signal being present.

Plotting the proportion of hits against the proportion of false alarms, for each of several decision criteria, will lead to a curve similar to one of those shown in the **illustration**. By varying physical characteristics (for example, the energy level) of the signal or noise, a number of such curves is obtained. Each curve, called a receiver operating characteristic, shows the trading relation between the two types of response outcome for discrimination of a given degree. The degree of acuity corresponding to each curve can be represented by a single number, which gives essentially the distance of the curve from the diagonal line (no discrimination) toward the upper left-hand corner (perfect discrimination). This number, it can be seen, is independent of the decision criterion. If one wishes to represent any criterion the subject may have adopted, then one can abstract a second number from the curve, such as the slope of the curve at the point yielded by that criterion.

Other experimental procedures used in the context of detection theory typically yield measures of discrimination that agree well with those obtained from the procedure exemplified—the so-called yes-no procedure. These include the confidence-rating

procedure, in which the subject is asked to relate each observation to several criteria held simultaneously, and the forced-choice procedure, in which the task is to state in which one of a specified set of intervals, of time or space, the signal was presented. The latter procedure largely bypasses the issue of decision criteria, and it is often used in studies directed solely at questions of sensory function.

Comparison with threshold theories. The theory of signal detection is an alternative to theories based on the classical concept of a threshold. As indicated, detection theory sees the information obtained from an observation as ranging along a continuum of more or less certainty about the stimulus that gave rise to it and sees the observer as able to place a decision criterion at any point on the continuum. On the other hand, threshold theories view the sensory information available to the observer as "all or none" or, at most, as falling into one of very few categories of relative certainty, with the categories demarcated by unalterable criteria. The term threshold connotes a barrier that must be overcome: The notion is that when the observation yields enough information to exceed the barrier, then a signal is clearly seen or heard or felt, as the case may be, and that when the information does not exceed this fixed quantity, then absolutely no information about the signal exists. The older threshold theories are incompatible with much of the present evidence. The measures of discrimination they suggest are not only confounded by response biases, but the values of these measure depend too heavily upon which experimental procedure is used to determine them. *See* PSYCHOPHYSICAL METHODS.

Applications. Signal detection theory has been applied to several topics in experimental psychology in which separation of intrinsic discriminability from decision factors is desirable. Included are attention, imagery, learning, conceptual judgment, personality, reaction time, manual control, and speech. Also included are studies of animal behavior, with rats, pigeons, and monkeys producing neat curves of the sort shown in the illustration. These curves are called ROC curves, for receiver (or relative) operating characteristic.

The analytical apparatus of the theory, primarily the ROC curve, has been of value in the evaluation of the performance of systems that make decisions based on uncertain information. Such systems may involve only people, or people and machines together, or only machines. Examples come from medical diagnosis, where clinicians may base diagnostic decisions on a physical examination, or on an x-ray image, or where machines make diagnoses, perhaps by counting blood cells of various types. Further, manual and automatic systems for retrieving articles in a library have been compared for accuracy relative to cost. And one may attempt to determine the validity of systems designed for forensic purposes, for example, for identifying an unknown speaker on the basis of a recording, or for detecting lies by means of a polygraph test. *See* HEARING (HUMAN); LIE DETECTOR; PSYCHOLOGY; VISION. John A. Swets

Bibliography. J. P. Egan, *Signal Detection and ROC-Analysis*, 1975; D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics*, 1966, reprint 1989; J. A. Swets, The relative operating characteristic in psychology, *Science*, 182:990–1000, 1973; J. A. Swets, *Signal Detection and Recognition by Human Observers*, 1964, reprint 1989.

Signal generator

A piece of electronic test equipment that delivers a sinusoidal output of accurately calibrated frequency. The frequency may be anywhere from audio to microwave, depending upon the intended use of the instrument. The frequency and the amplitude are adjustable over a wide range. The oscillator must have excellent frequency stability, and its amplitude must remain constant over the tuning range.

The Wien-bridge oscillator is commonly used for frequencies up to about 200 kHz. For a radio-frequency signal generator up to about 200 MHz, a resonant circuit oscillator is used (such as a tuned-plate tuned-grid, Hartley, or Colpitts). Beyond this range, very high frequency and microwave oscillators are used.

Many signal generators contain circuitry that allows the output to be either amplitude- or frequency-modulated. The most common forms of amplitude modulation are sinusoidal, square-wave, and pulse. The frequency is kept constant, is sinusoidal-modulated, or is swept linearly across a band of frequencies. For example, for testing broadcast receivers, it is important to sweep the generator frequency over a range of ± 10 kHz at a low rate, say 60 times a second. *See* OSCILLATOR; PULSE GENERATOR; WAVE-SHAPING CIRCUITS. Jacob Millman

Bibliography. A. D. Helfrick and W. D. Cooper, *Modern Electronic Instrumentation and Measurement Techniques*, 1990; J. Markus and C. Weston, *Essential Circuits Reference Guide*, 1988; D. L. Schilling et al., *Electronic Circuits: Discrete and Integrated*, 3d ed., 1989.

Signal processing

Operations that modify or create signals. A simple representation of a signal processing operation has an input signal *A* that is modified to produce an output signal *B* (**Fig. 1**).

For a signal to be useful, it must have one or more characteristics that vary, usually over time or distance, to represent information. Amplitude, frequency, and phase are common signal characteristics

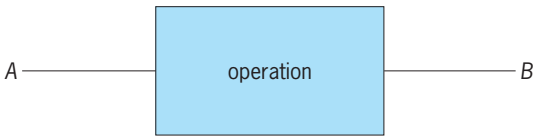

Fig. 1. Simple signal processing representation.

that can carry information. Time-varying signals are written as $A(t)$. *See* AMPLITUDE (WAVE MOTION); FREQUENCY (WAVE MOTION); PHASE (PERIODIC PHENOMENA).

A simple signal processing operation is scaling, which multiplies the input signal by a constant value to produce the output signal. Another simple operation is the addition of a constant value to the input to produce the output.

A signal processing operation may have more than one input. For example, two input signals may be added together to produce an output signal. Addition is a linear mathematical operation. An operation may also be nonlinear, as is the case where an input signal is squared or when two input signals are multiplied together.

Logical operations may also be employed in signal processing to make decisions and route signals. For example, a switch may be used to select between two signals based on the value of a third signal. *See* LOGIC CIRCUITS.

An important signal processing operation is time delay. The output signal is a delayed copy of the input signal.

Simple mathematical, logical, and time-delay operations are combined into signal processing systems that perform more complex processing operations. The simple operations, or blocks, are connected together in an appropriate manner to implement the desired complex operation. **Figure 2** shows the block diagram of a system that performs a time-averaging operation. The input signal and a delayed version of the input signal are added together and then scaled by 0.5 to produce the averaged output.

Signal processing is performed in analog and digital circuits. Analog signal processing has been used since the beginning of audio and radio electronics. Over the past few decades, digital signal processing has become an increasingly common method, especially for more complex systems. In many cases, the same operation can be performed equivalently in both the analog and digital domains, although one may have advantages depending on various factors, such as the physical domain, complexity, and required adjustability of the operation.

Because computers use numbers and perform calculations on numbers, they can be programmed to perform digital signal processing. By approximating analog signals and operations in digital form, analog signal processing can also be performed on computers using simulation programs. *See* COMPUTER; DIGITAL COMPUTER; SIMULATION.

Characteristics of operations and systems. Signal processing operations and systems can be described in different ways. The choice depends on the physical process being described, the approach to designing the processing system, the desired outcomes of the design, the modeling of the design, and the capabilities of the designer and design tools.

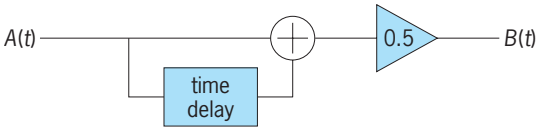

Fig. 2. Signal averager.

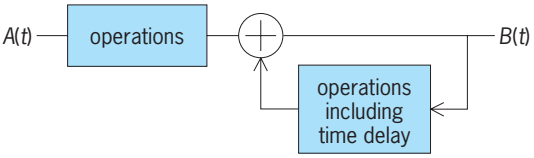

Fig. 3. System containing feeback.

Linear and nonlinear signal processing. One characteristic already discussed is whether an operation is linear or nonlinear. An operation with input x and output y is linear if Eq. (1) holds true, where m and b are

$$y(t) = m \cdot x(t) + b \tag{1}$$

constants. A nonlinear operation does not meet this requirement.

If an operation does not vary over time, as in Fig. 2 or in Eq. (1), it is said to be time-invariant. If instead the value of the scaling in Fig. 2 changes over time, the scaling operation and thus the whole system is considered time-variant. Time-invariant systems are generally easier to design, model, and analyze.

Linear time-invariant systems are by far the signal processing systems most commonly designed, analyzed, and implemented. They can be represented by simple mathematical equations called transfer functions, which describe completely the input-output relationship of the system for any input signal. *See* LINEAR SYSTEM ANALYSIS; LINEARITY.

Feedback systems. A system may contain feedback, which means at least one signal within the system returns to a point that indirectly feeds into the same signal, creating a loop. Delay must be present in the loop. The system in Fig. 2 does not have feedback. A general representation of a system with feedback is shown in **Fig. 3**. The output signal $B(t)$ is fed back through operations that include time delay, with the delayed result added to the processed input to create the output. A key attribute of a feedback system is whether or not it is stable. An unstable system will produce an output that is uncontrolled and unpredictable, making the system unusable. Methods that analyze the operations in the loop can be used to determine if a particular feedback system is stable. *See* CONTROL SYSTEM STABILITY; CONTROL SYSTEMS.

Real and complex signal processing. The representation of signals as having complex values is a very useful tool in signal processing. Here, complex does not refer to the complexity of a signal but rather to a signal having real and imaginary components. A signal that is not complex may be described as real or scalar. Operations on complex signals are also complex and have real and imaginary parts. Complex signals and operations allow certain types of signal processing tasks, like those in digital radio systems, to be performed more efficiently than if scalar signals and operations are used. They also allow signal processing implementations to follow more directly the mathematical models on which they are based.

To denote the imaginary part of a complex signal, the symbol i is used. The complex signal $C(t)$ is represented with $A(t)$ as the real part and $i{\cdot}B(t)$ as the imaginary part, as in Eq. (2). The symbol i is generally used in mathematical equations, but in signal processing the equivalent symbol j is more common. The quantities $A(t)$ and $B(t)$ are themselves real. In a signal processing system these can be signals that could for example be carried on two wires. The wires taken together are treated as a real-imaginary pair forming a complex signal. Thus, even though the complex signal has a so-called imaginary component, the signal can exist in the real world and complex operations on the signal can be performed.

$$C(t) = A(t) + i{\cdot}B(t) \tag{2}$$

The symbol j actually has the value of $\sqrt{-1}$. Although this expression cannot be calculated, it can be used in calculations describing complex operations. For example, the product $P(t)$ of two complex signals $C(t)$ and $D(t)$ is calculated in Eq. (3). In block diagram form, the product can be represented as a complex operation on complex signals or as real operations on real signals (**Fig. 4**).

$$\begin{aligned} P(t) = C(t)\cdot D(t) &= [A(t) + j\cdot B(t)]\cdot[E(t) + j\cdot F(t)] \\ &= A(t)\cdot E(t) + j\cdot B(t)\cdot j\cdot F(t) \\ &\quad + A(t)\cdot j\cdot F(t) + j\cdot B(t)\cdot E(t) \\ &= A(t)\cdot E(t) + j\cdot j\cdot B(t)\cdot F(t) \\ &\quad + j\cdot A(t)\cdot F(t) + j\cdot B(t)\cdot E(t) \\ &= A(t)\cdot E(t) - B(t)\cdot F(t) \\ &\quad + j\cdot[A(t)\cdot F(t) + B(t)\cdot E(t)] \\ &= V(t) + j\cdot W(t) \end{aligned} \tag{3}$$

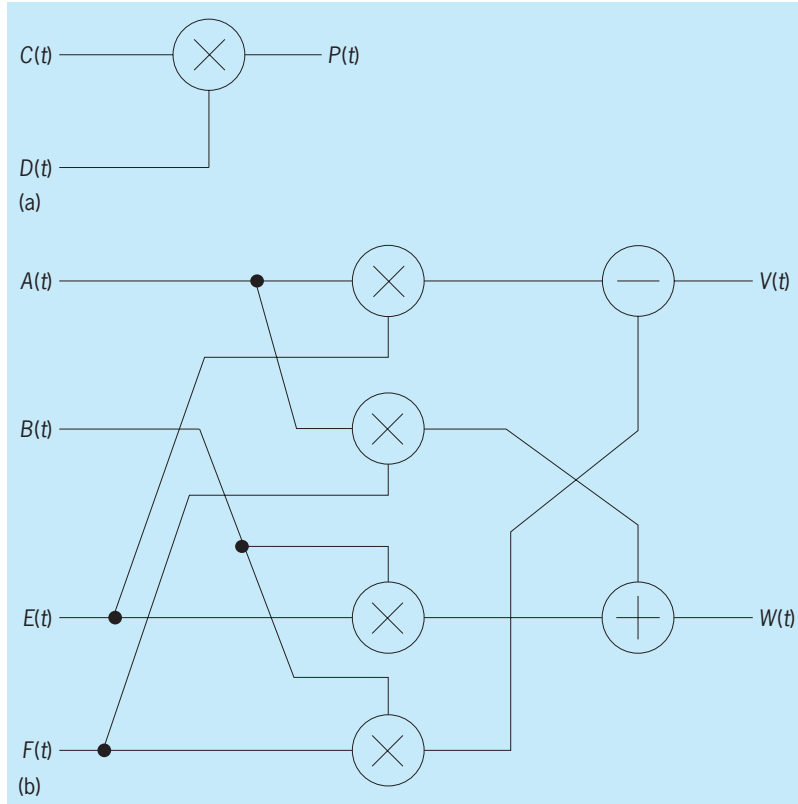

Fig. 4. Complex product operation (*a*) represented as a complex operation on complex signals, (*b*) represented as real operations on real signals.

The result is also complex. If Eq. (3) were to be implemented in an electronic circuit, all of the real signals $A(t)$ to $F(t)$ would exist as electronic signals. Similarly, if the complex operation were to be implemented in a computer program, these signals would be represented by program variables. Operations that process scalars usually have equivalent complex versions.

A complex signal $A(t) + j\cdot B(t)$ can also be represented in the form $M(t)\cdot e^{j\cdot\varphi(t)}$, where $M(t)$ and $\varphi(t)$ are the magnitude and phase of the signal. The magnitude and phase can be calculated from the real and imaginary parts, as in Eqs. (4).

$$M(t) = \sqrt{A^2(t) + B^2(t)} \tag{4a}$$

$$\varphi(t) = \arctan[B(t)/A(t)] \tag{4b}$$

Signal processing in multiple dimensions. Signal processing may be performed across a one-dimensional domain, like time, or in two dimensions. Time-based processing has already been discussed. A common application of two-dimensional processing is the generation and manipulation of visual images. The domain is space instead of time, and the space is typically measured along a horizontal and vertical axis. The image is represented as a two-dimensional array of pixels, and processing of the image involves operations on individual pixels or groups of pixels. For example, a simple way to soften an image is by averaging the intensities of its pixels. This reduces the intensity variations from pixel to pixel, and is performed along both axes, averaging short groups of adjacent pixels first in each horizontal line of the image, and then in each vertical line. Other operations can be performed to sharpen images, remove noise, change coloration, extract features, and combine images. Image-processing operations can have the various characteristics already discussed, such as nonlinearity. *See* IMAGE PROCESSING.

Processing in higher dimensions and in combinations of different domains (like time and space) are also possible. This depends only on how the signal processing designer chooses to represent a signal processing system and on the design tools available.

Creating signals. In addition to operating on input signals, signal processing is also used to create signals. For example, a sine wave oscillator is a signal source employed in radio and music synthesis systems. A signal source may have control inputs to set its output characteristics, like an oscillator's frequency and amplitude. *See* OSCILLATOR.

Analog signal processing. Whether a signal processing operation or system is implemented in analog or digital form is one of its most important characteristics. This is because the methods for implementing each form are very different. Analog processing is performed in traditional analog circuits, which are built from analog components. The signals in these circuits carry information according to

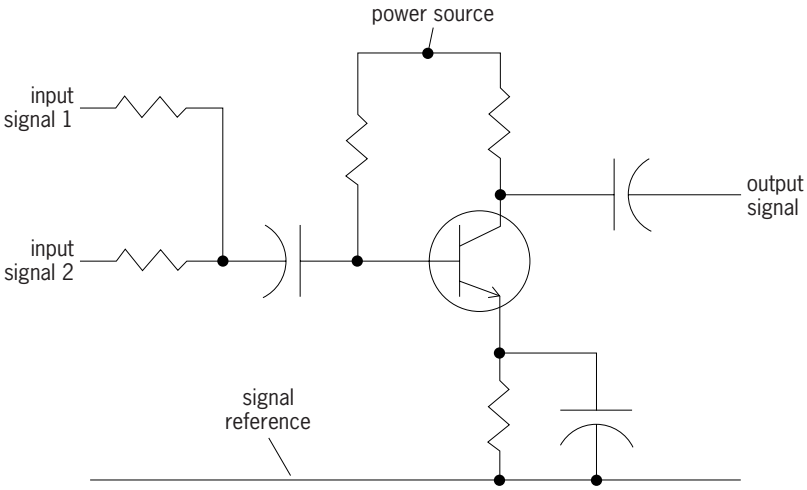

Fig. 5. Amplifier using a transistor, resistors, and capacitors.

the time-varying values of their voltage and current. Analog signals and operations are continuous in value and in time.

Analog processing components. Two kinds of analog components are used: passive and active. Resistors, capacitors, inductors, and transformers are passive components, and the operation of each can be described by a simple relationship between the voltage across the terminals of the component and the current passing through the component. Passive components do not require an external power source to operate. *See* CAPACITOR; INDUCTOR; RESISTOR; TRANSFORMER.

Active components are based on semiconductors or in older or high-power circuits on electron tubes. Transistors and diodes are common types of semiconductor components. Operations that require some type of amplification or nonlinear operation must use active components, and therefore a power source is required. To form useful signal processing operations, semiconductor components must be combined with passive components to create circuits, such as an amplifier that sums two input signals (**Fig. 5**). *See* AMPLIFIER; DIODE; ELECTRON TUBE; TRANSISTOR.

Integrated analog processing. A very useful analog device is the operational amplifier, or op amp. It is a circuit consisting of active and passive components that have been packaged, or integrated, into a single, higher-level active component. Op amps were created to make analog signal processing design easier than designing circuits based directly on semiconductors and supporting passive components. This in turn allows more complex circuits to be designed, using multiple op amps mixed with other active and passive components. *See* OPERATIONAL AMPLIFIER.

Very complex and high-performance analog signal processing systems can be integrated into a single component, or integrated circuit. This allows, for example, most of the signal processing for a radio or television to be contained in a package the size of a quarter, helping to reduce the cost to manufacture these products. *See* INTEGRATED CIRCUITS.

Analog processing circuit applications and design. Despite the now widespread use of digital signal processing, analog signal processing remains very important. Analog circuits must be used for operations on signals that vary at high frequencies, approximately 100 MHz and higher. Certain types of processing, for example in the input stages of radio and television receivers, require analog circuits combining high-frequency signal handling with the ability to input signals at very weak and very strong levels. High-power processing, such as that of a radio transmitter, must also be performed in analog. Even in digital signal processing systems, analog processing must be employed to convert between the digital domain and the mostly analog domain of the physical world. *See* RADIO RECEIVER; RADIO TRANSMITTER; TELEVISION RECEIVER.

Digital signal processing. Up until the past few decades, most electronic products, such as telephones, radios, televisions, and sound processors (like audio mixing consoles), were implemented completely with analog processing circuits. Slowly, more of the processing in many products has been changed to digital implementations. Two important aspects of all-digital signal processing systems separate them from analog systems: sampling and quantization.

Sampling. In digital signal processing (DSP) systems, signals exist only at discrete points, usually in time or space. Such signals are called sampled signals and have a value at each of these points. Thus, a sampled signal A can also be written as $A(n)$, where n is an index of the values, or samples (**Fig. 6**). Usually the samples are arranged in time or space so that the time or distance between any two adjacent samples is the same, or uniform. The time or distance between adjacent uniform samples is the sample period, and the inverse of the sample period is the sample frequency, or sample rate. For sample period T, $A(n)$ is the sequence of samples of signal A at the discrete points in time $n \cdot T$. It is convenient to view $A(n)$ as having been created by sampling a continuous-time signal $\hat{A}(t)$, although a sampled signal does not need to originate from a continuous-time signal.

How a signal varies over time or space can be described by how fast it changes, that is, by its frequency content. A sampled signal can contain frequencies no higher than one-half its sample rate. And, for a signal with bandwidth B to be accurately represented as a sampled signal, it must be sampled

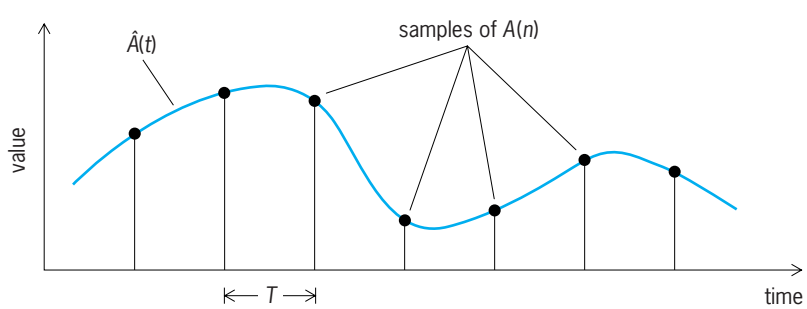

Fig. 6. Time-sampled signal.

at a rate $2 \cdot B$ or higher. This minimum rate is called the Nyquist rate, and half this rate is the Nyquist frequency. In practice, the sample rate must be somewhat higher than the Nyquist rate. *See* INFORMATION THEORY.

Quantization and signal representations. In digital signal processing systems, signals are represented by binary numbers, and as such are quantized, meaning they can take on only a limited set of values. For example, an 8-bit form of a signal can take on values from 0 to 255, or −128 to 127, depending on what the signal represents or how it is processed. The quantization of a signal means that for each sample, its value must be approximated to one of the values, or quantization steps, allowed by the binary form representing the signal.

This approximation process adds quantization noise to the signal, because each sample has some amount of error compared to the exact value of the signal. With the proper design of a digital signal processing system, quantization noise can be controlled and will have a minimal adverse effect on the system's operation. *See* PULSE MODULATION.

Two forms of numbers are used to represent signals. The simpler form is fixed point, meaning the binary point, similar to a decimal point, is in a fixed position. In the second form the binary point is allowed to move, or float, to allow a wider range of values to be represented, and so this form is called floating point. Some of the bits of the number indicates the position of the binary point.

Because digital signal processing uses numbers for signals, processing operations can be represented directly by mathematical and logical equations. The general signal processing operation of the addition of two signals, for example, is performed by the actual addition of the binary values representing the signals. A delay operation is easily implemented by delaying a signal by a number of sample periods using a digital delay line. Multiplication and division are performed in the manner of long-hand arithmetic.

Digital signal processors and other implementations. As with analog signal processors, complex digital signal processing systems may be combined into single, highly integrated circuits. These are commonly called digital signal processors, and are commercially available from many sources and for many purposes. Their processing is controlled by internally or externally stored programs consisting of sequences of simple program steps, or instructions, that execute mathematical and logical operations. These processors can be programmed by designers for many different purposes, and for added versatility may include various types of interfaces for signal input and output.

A typical digital signal processor contains an arithmetic logic unit (ALU), a multiplier, memories to store signals (data) and coefficients, and a section that controls the operation of these blocks and the routing of signals between blocks and into and out of the processor (**Fig. 7**). At each program step, the ALU performs an addition, subtraction, or logic operation, according to the mode selected by the controller. Coefficients are values that do not change as often as signals, such as the gain of a gain block. Higher-performing processors may contain multiple ALUs and multipliers, and other blocks like barrel shifters and hardware accelerators, which are small nonprogrammable sections of the processor executing dedicated operations for higher speed and efficiency.

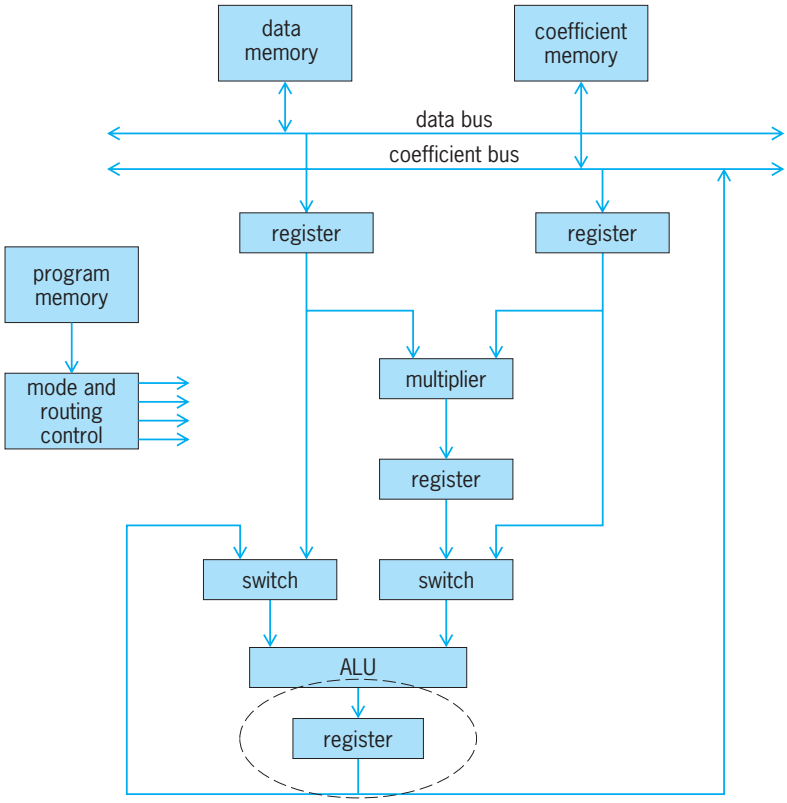

Fig. 7. Programmable digital signal processor.

Generally, a fixed sequence of program instructions is executed during each sample period. More complex programs may alter the performed sequences depending on inputs to the processor or results from previous processing steps. The sequence of executed instructions is varied using special branch instructions that cause the controller to redirect the instruction flow.

Digital signal processing can also be performed in circuits that are not programmable, but rather consist of simple processing operations hardwired together to perform a single processing task. This type of processing is sometimes more efficient to use for signal processing at higher sample rates, approximately 100 MHz and higher.

Another highly integrated device used for digital signal processing is the field-programmable gate array (FPGA). In terms of programmability and speed, it falls between programmable digital signal processors and hardwired processing. In FPGAs, logic gates are connected to create simple signal processing operations, like adders, which in turn can be connected together inside the FPGA to create more complex operations. The configuration of the connections is programmed into the FPGA device.

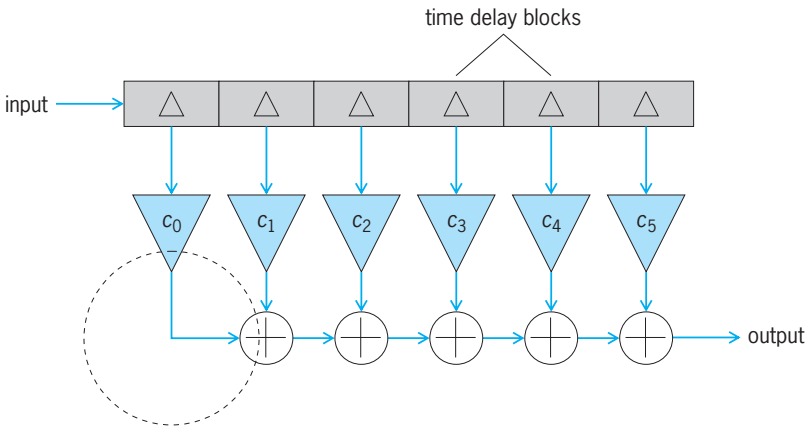

Fig. 8. Finite impulse response filter.

Hardwired, programmable, FPGA, and even analog processing may be integrated into very complex processors. These are developed for very cost sensitive products that are manufactured in high volumes, or that require a very small physical implementation, such as cellular telephones.

Analog-to-digital and digital-to-analog converters. Many applications of digital signal processors involve processing analog signals, such as the signal from a microphone. Analog-to-digital (A/D) and digital-to-analog (D/A) converters provide the interfaces between analog signals and the digital signals inside processors. Converters must sample at the Nyquist rate of the signal being converted.

Analog-to-digital converters must have sufficiently fine quantization to avoid adding excessive noise, and their input signals must usually be analog filtered to avoid a condition called aliasing. If an analog input signal has undesired frequency content above the Nyquist frequency, this content must be removed by analog filtering prior to sampling by the converter. If this antialiasing filtering is not performed, the undesired signal content will alias into the same region of frequencies as the desired content, combining with it in the converter output. When this occurs, it is impossible to separate the desired and undesired content. *See* ANALOG-TO-DIGITAL CONVERTER.

Digital-to-analog converters must usually be followed by an analog filter, called a reconstruction filter. The filter smoothes the quantization steps of the signal. In the case of an audio signal these steps would make the signal sound noisy if filtering was not used. *See* DIGITAL-TO-ANALOG CONVERTER.

Types of signal processing blocks. Signal processing systems are built from various types of signal processing blocks. Most can be implemented in analog or digital form, although the digital form is usually more directly related to a block's mathematical model.

The blocks that will be described are examples of the many types of operations used in signal processing systems. They also show how more complex blocks can be constructed from simpler ones.

Linear processing blocks and filters. Addition, scaling, and delay blocks are the most basic linear blocks, and the simplest to implement, especially using digital signal processing. From these, other linear blocks, like filters, can be created.

There are two basic types of linear filters. The first, which does not contain feedback, is the finite impulse response (FIR) filter, also called the transversal filter. Multiple delayed versions of the input signal are scaled and then added together to form the output. The FIR filter in **Fig. 8** adds six versions of the signal, and is therefore a six-tap filter. The input signal is delayed by a series of connected delay blocks, forming a delay line with the needed six taps. The scaling gains c_0 to c_5 are the filter's coefficients. The averager in Fig. 2 is an FIR filter.

The second basic filter type is the infinite impulse response (IIR) filter, which is similar to the FIR filter but contains feedback (**Fig. 9**).

Most analog signal processing filters are IIR and are formed from networks of resistors, capacitors, and inductors. Digital filters are formed directly according to the structures shown in Figs. 8 and 9.

The coefficients of an FIR or IIR filter determine the filter's frequency response, and whether the response is lowpass, highpass, bandpass, bandstop, or has a more complex shape. A filter's coefficients may also be designed to create differentiators and integrators.

If the coefficients of FIR and IIR filters are constant with time, the filters are linear time-invariant, and so can be represented by transfer functions. The transfer function H of the IIR filter in Fig 9, if implemented in digital form, is given in Eq. (5). The

$$H(z) = \frac{b_0 \cdot z^2 + b_1 \cdot z + b_2}{z^2 - a_1 \cdot z - a_2} \tag{5}$$

symbol z is an operator related to frequency, and allows the response of the filter to be conveniently evaluated for any input signal. In the analog domain, similar transfer functions are used, but with the Laplace operator s. *See* ELECTRIC FILTER; LAPLACE TRANSFORM; Z TRANSFORM.

Nonlinear processing blocks. Nonlinear blocks for multiplication, division, squares and higher powers, square roots, and logarithms can all be performed in signal processing. Analog implementations require the use of active components. The digital forms of some of these operations, like multiplication and division, are usually executed in the same way long-hand arithmetic is performed. Others, like

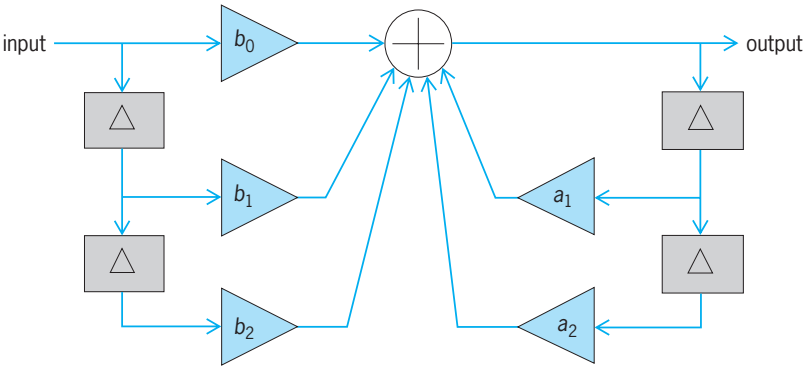

Fig. 9. Infinite impulse response filter.

logarithms, use approximations based on polynomial equations (themselves consisting of multiplication and addition operations). Look-up tables may also be used to calculate logarithms, just as they were calculated from tables in books before the advent of slide-rules and calculators. Actually, any nonlinear operation for a block may be approximated using a look-up table.

Modulator blocks process audio, image, and data signals so they can be transmitted from one point to another, as in a radio broadcast, or stored into physical media, like a compact disk. An amplitude modulator of a radio transmitter varies the amplitude of a high-frequency carrier in proportion to the level of the relatively low-frequency audio signal to be transmitted (**Fig. 10**). A demodulator performs the

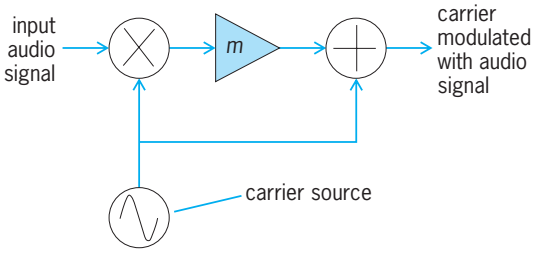

Fig. 10. Amplitude modulator.

opposite operation of its corresponding modulator to recover the original modulated signal. A related block is a frequency translator, which shifts a carrier signal from one frequency to another while preserving the modulation on the carrier. Combining a frequency translator block and a filter block allows signals to be selected, or tuned, from within a frequency space containing many signals, such as all of the radio stations in a city. *See* AMPLITUDE MODULATOR; DEMODULATOR; MODULATION; MODULATOR.

A level detector block outputs an estimate of the level of its input signal. A common type of detector measures the root-mean-square (RMS) level of a signal (**Fig. 11**). It can be used with other blocks to construct automatic gain controllers (AGC) and level compressors. *See* AUTOMATIC GAIN CONTROL (AGC).

Logical blocks. Logic signals can be created in signal processing systems and used to control other signals. For example, in a comparator block an input signal is compared to a threshold to produce a logical 0 or 1 output. A switch block outputs one of its two inputs according to a logic selection input to the block. Powerful signal processing systems can be created from combining logic-based blocks with other types of blocks. An example of this is a neural network, which is capable of complex tasks like pattern recognition in images. *See* NEURAL NETWORK.

Transform blocks. Mathematical transforms can be implemented in signal processing systems. The Fourier transform is very useful in communication and signal analysis systems, and is usually implemented in an efficient digital form called the fast Fourier transform (FFT). *See* FOURIER SERIES AND TRANSFORMS.

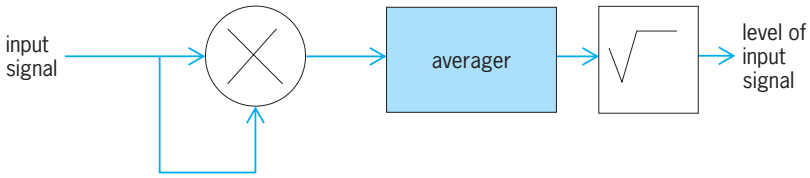

Fig. 11. RMS level detector.

Signal source blocks. Signals may be created for various purposes, such as for the carrier signal in Fig. 10. A sine wave is a common and useful signal, and may be created by designing a block in the form of an IIR filter so that its internal feedback is on the verge of instability, causing the output to vary periodically, or oscillate, in a controlled manner to produce the sine wave. Another method to create a sine wave, and other signals of any arbitrary waveform shape, is a look-up table. Samples of the waveform are precalculated and stored in the table. A counter creates an index to select the table entries in sequence, thus creating the desired waveform (**Fig. 12**). *See* SIGNAL GENERATOR; WAVEFORM GENERATOR.

Examples of signal processing systems. Radio signal processing is a key element of many electronic systems, including cellular telephones, radar, point-to-point radio communication, and transmitters and receivers for AM, FM, television, and more recently satellite broadcast systems. These systems usually employ a combination of analog and digital processing, and consist of many processing blocks. The main blocks in radio systems are frequency translators, filters, modulators, demodulators, AGC and amplifiers, A/D and D/A converters, and other blocks for manipulating data. These blocks appear in the diagram of a cellular telephone in **Fig. 13**. *See* MOBILE RADIO.

Point-to-point communication is used for voice and data signals in military and space applications and for many other purposes. Signal processing on space probes allows large amounts of telemetry data, like images of other planets, to be transmitted to radio receivers on Earth, where more processing converts the data back into pictures and other usable forms. Military radios may include processing to encrypt and decrypt signals. *See* TELEMETERING.

Devices for producing and reproducing music make heavy use of signal processing, especially digital processing. Electronic keyboards contain processing to create multiple sound sources capable

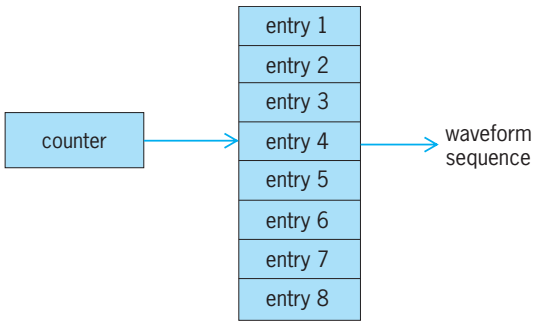

Fig. 12. Signal source using a look-up table.

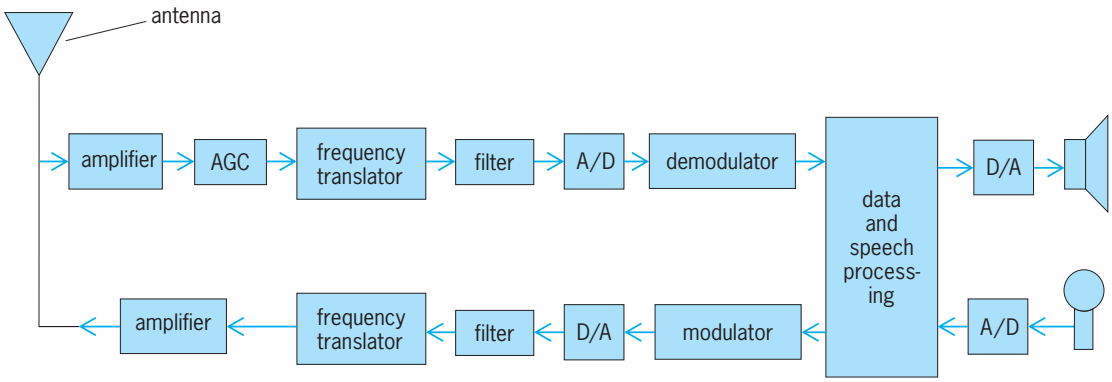

Fig. 13. Cellular telephone block diagram.

of producing many timbres, and they employ additional processing to manipulate and combine these sources into audio outputs. An example of sound manipulation is reverberation, which uses delay and gain blocks arranged into feedback loops to create echoes. Portable MP3 music players employ signal processing to convert highly compressed stored sound data into music output. *See* MUSICAL INSTRUMENTS; SOUND-REINFORCEMENT SYSTEM.

Automated systems, including industrial robots, are commonly used in factories to control machinery and position materials. The physical parts of this type of system, such as a robot arm, are moved by actuators controlled by a signal processing program, with sensors in the system providing feedback to the program regarding the positions and movement of parts of the system. The program, actuators, physical parts, and sensors form a feedback loop that allows the program to finely control the system's movements. *See* ROBOTICS.

A few more examples of the many systems that employ signal processing are medical-imaging devices, sonar, DVD recorders, engine controllers in cars, multichannel sound systems, night-vision goggles, graphics design programs, power plants, and guidance systems for rockets and missiles.

J. William Whikehart

Bibliography. J. J. Carr, *Secrets of RF Circuit Design*, 3d ed., McGraw-Hill, 2001; R. J. Higgins, *Digital Signal Processing in VLSI*, Prentice-Hall, 1990; J. H. McClellan, R. W. Schafer, and M. A. Toder, *DSP First: A Multimedia Approach*, Prentice-Hall, 1997; S. W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*, California Technical Publishing, 1997.

Signal-to-noise ratio

The quantity that measures the relationship between the strength of an information-carrying signal in an electrical communications system and the random fluctuations in amplitude, phase, and frequency superimposed on that signal and collectively referred to as noise. For analog signals, the ratio, denoted S/N, is usually stated in terms of the relative amounts of electrical power contained in the signal and noise. For digital signals the ratio is defined as the amount of energy in the signal per bit of information carried by the signal, relative to the amount of noise power per hertz of signal bandwidth (the noise power spectral density), and is denoted E_b/N_0. Since both signal and noise fluctuate randomly with time, S/N and E_b/N_0 are specified in terms of statistical or time averages of these quantities.

An information signal usually accumulates noise at all stages of its generation, transmission, and reception. The noise contributed by the transmitters and receivers tends to have a power spectral density that is constant over a wide range of frequencies and is thus referred to as white noise, in analogy to the spectral distribution of white light. Furthermore, its level is proportional to the temperature of the various electrical components such as oscillators, amplifiers, filters, and detectors that constitute the signal generation and detection circuitry. Noise of this type, referred to as thermal noise, is caused in large part by the random motion of thermally excited electrons in conducting media. Its power spectral density is given by $N_0 = kT/2$ watts per hertz, where T is the medium's absolute temperature in kelvins and $k = 1.38 \times 10^{-23}$ joule/kelvin is Boltzmann's constant. In contrast, noise accumulated in the transmission path typically derives from external electromagnetic sources such as microwave systems, or radio and television stations. This type of noise is known as colored noise by virtue of the fact that its power spectral density varies with frequency. In most cases the total noise carried by the signal is in part derived from external sources and in part from the transmitter and receiver hardware. *See* BOLTZMANN CONSTANT.

The magnitude of the signal-to-noise ratio in a communications systems is an important factor in how well a receiver can recover the information-carrying signal from its corrupted version and hence how reliably information can be communicated. Generally speaking, for a given value of S/N the performance depends on how the information quantities are encoded into the signal parameters and on the method of recovering them from the received signal. The more complex encoding methods such as phase-shift keying or quadrature amplitude-shift

keying usually result in better performance than simpler schemes such as amplitude- or frequency-shift keying. As an example, a digital communication system operating at a bit error rate of 10^{-5} requires as much as 7 dB less for E_b/N_0 when employing binary phase-shift keying as when using binary amplitude-shift keying. *See* ELECTRICAL COMMUNICATIONS; ELECTRICAL NOISE; INFORMATION THEORY; MODULATION. Hermann J. Helgert

Bibliography. B. P. Lathi, *Modern Digital and Analog Communication Systems*, 3d ed., Oxford University Press, 1998; M. Schwartz, *Information Transmission, Modulation, and Noise*, 4th ed., McGraw-Hill, 1990.

Signal transduction

The transmission of molecular signals from a cell's exterior to its interior. Molecular signals are transmitted between cells by the secretion of hormones and other chemical factors, which are then picked up by different cells. Sensory signals are also received from the environment, in the form of light, taste, sound, smell, and touch. The ability of an organism to function normally is dependent on all the cells of its different organs communicating effectively with their surroundings. Once a cell picks up a hormonal or sensory signal, it must transmit this information from the surface to the interior parts of the cell—for example, to the nucleus. The cell then uses this information to respond in an appropriate manner.

Information is transmitted from the cell surface to the interior via signal transduction pathways that are very specific, both in their activation and in their downstream actions. Thus, the various organs in the body respond in an appropriate manner and only to relevant signals. *See* CELL (BIOLOGY).

Cell signals. All signals received by cells first interact with specialized proteins in the cells, called receptors, which are very specific to the signals that they receive. Signals received by the cell can be in various forms. The most common are chemical signals, which include all the hormones and neurotransmitters secreted within the body as well as the sensory (external) signals of taste and smell. The internal hormonal signals include steroid and peptide hormones, neurotransmitters, and biogenic amines, all of which are released from specialized cells within the various organs. The external signals of smell, which enter the nasal compartment as gaseous chemicals, are dissolved in liquid and then picked up by specialized receptors. Other external stimuli are first received by specialized receptors (for example, light receptors in the eye and touch receptors in the skin), which then convert the environmental signals into chemical ones, which are then passed on to the brain.

Receptors. Receptors are specialized proteins that perform unique functions, the most important of which are to recognize external signals and convert them into biochemical or physiological functions. Once a receptor has received a signal, it must transmit this information effectively into the cell. This is accomplished either by a series of biochemical changes within the cell or by modifying membrane voltage potential by the movement of ions into or out of the cell. Receptors that initiate biochemical changes can do so either directly via intrinsic enzymatic activities within the receptor or by activating intracellular messenger molecules. Receptors may be broadly classified into four classes: G-protein–coupled receptors, receptor tyrosine kinases, ion channels, and steroid receptors. These types of receptors differ in their mode of action and in the molecules that activate them (**Fig. 1**).

G-protein–coupled receptors. The largest family of receptors is the G-protein–coupled receptors (GPCRs), dependent on guanosine triphosphate (GTP) for their function. Many neurotransmitters, hormones, and small molecules bind to and activate specific G-protein–coupled receptors. The G-protein–coupled receptor is located at the cell membrane; it is made up of an extracellular ligand-binding domain, seven

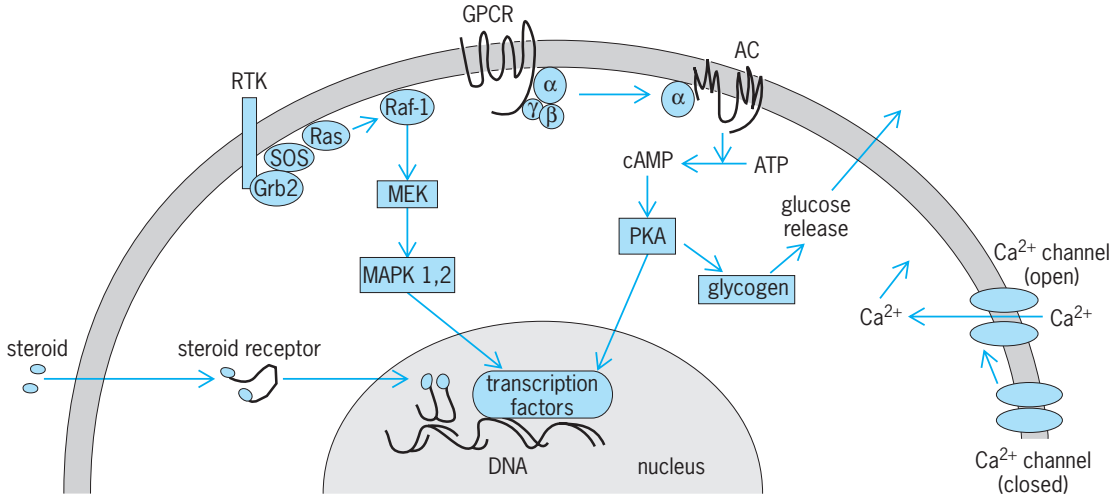

Fig. 1. Different receptors and their mechanism of action: G-protein–coupled receptor (GPCR), receptor tyrosine kinase (RTK), adenylyl cyclase (AC).

membrane-spanning domains, and an intracellular domain. In the nonstimulated state the receptor is bound within the cell to a complex of three different proteins (a heterotrimer), named alpha (α), beta (β), and gamma (γ), which together constitute the G-protein. The α subunit is also bound to guanosine diphosphate (GDP) in the inactive stage. Upon activation of the receptor, the α subunit binds GTP, the heterotrimer disassociates from the receptor, and the α subunit dissociates from the β and γ subunits. The activated α and, $\beta\gamma$ subunits (which are bound together) then bind to and regulate their effector molecules. Once the signal has been passed on, the α subunit hydrolyzes the bound GTP to GDP, which makes the α subunit inactive, and it then binds the $\beta\gamma$ subunits. At this point, the subunit is back to its resting state and capable of being activated again by the receptor.

Receptor tyrosine kinases. A second family of membrane-bound receptors is the receptor tyrosine kinases (RTKs). They function by phosphorylating themselves and recruiting downstream signaling components. Upon activation the receptor becomes phosphorylated on tyrosine residues, allowing the receptor to engage intracellular pathways.

Ion channels. These proteins form channels that open upon activation, thereby allowing the passage of ions across the membrane. Ion channels are responsive to either ligands or to voltage changes across the membrane, depending on the type of channel. The movement of ions changes the membrane potential, which in turn changes cellular function.

Steroid receptors. Steroid receptors are not located at the cell membrane but are found within the cell. They bind cell-permeable molecules such as steroids, thyroid hormone, and vitamin D, all of which are lipophilic and can thus pass through the cell membrane into the cell. Once these receptors are activated by ligand, they translocate to the nucleus, where they bind specific deoxyribonucleic acid (DNA) sequences to modulate gene expression. *See* STEROID.

Signal flow within cells. Once the receptor has received its signal, it must effectively propagate this message to the cell. The intracellular component of signal propagation, also known as signal transduction, is receptor-specific. A given receptor will activate only very specific sets of downstream signaling components, thereby maintaining the specificity of the incoming signal inside the cell. In addition to maintaining specificity, signal transduction pathways amplify the incoming signal. Signal amplification is achieved by a signaling cascade, in which molecule A activates several molecule B's, which in turn activate several molecule C's. In this way the activation of a few receptors eventually leads to the activation of numerous signaling molecules, resulting in an appropriate physiological response by the cell.

Signaling by receptors. The signaling cascade is initiated at the level of the receptor. The various receptors activate different signaling molecules and cascades. Activation of G-protein–coupled receptors results in Gα and G$\beta\gamma$ subunit activation. Gα subunits can activate or inhibit signaling molecules, depending on the Gα subtype that is bound to the G-protein–coupled receptor. Gαs stimulates adenylyl cyclases (AC) to produce cyclic adenosine monophosphate (cAMP), whereas Gαi inhibits adenylyl cyclases and the production of cAMP. Thus, an additional level of connection specificity is introduced in that different G-protein–coupled receptors bind to specific Gα subunits. Once the Gα subunit is ON (activated), it uses adenylyl cyclases to amplify its signal. The Gα subunit has a built-in OFF (inactivated) mechanism (the hydrolysis of bound GTP to GDP) which autoregulates the duration of the signal.

Initiation of the signaling cascade by receptors that have intrinsic enzyme activities, such as receptor tyrosine kinases, proceeds in a different manner. Upon binding hormone, the receptors become enzymatically competent and activate downstream signaling components. These activated messenger components then activate signaling components farther downstream, bringing into play a signaling cascade. The basic design here is that of a "bucket brigade" (**Figs. 2** and **3**). Ion channels function in yet another way to initiate signaling cascades. The receptors themselves are channels that can open and close to permit the movement of ions across the cell membrane. Once the ion channels are activated they undergo a change in their structural conformation, allowing ions to pass into the cell. When activated the channel opens, allowing an ion such as calcium (Ca^{2+}) into the cell. These ions can then trigger the activation of signaling pathways.

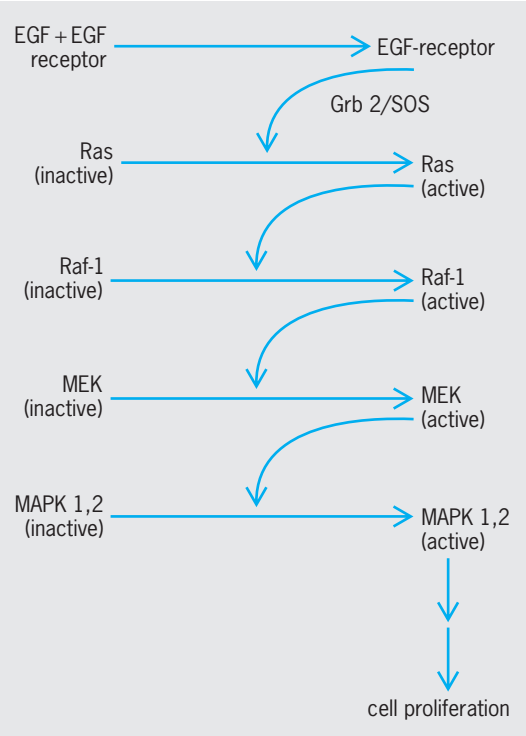

Fig. 2. Steps in the activation of the MAPK pathway leading to cell proliferation.

Signaling by small molecules. Several small molecules within the cell act as intracellular messengers. These include cAMP, cyclic guanosine monophosphate (cGMP), nitric oxide (NO), and Ca^{2+} ions. Increased levels of Ca^{2+} in the cell can trigger several changes, including activation of signaling pathways, changes in cell contraction and motility, or secretion of hormones or other factors, depending on the cell type. Increased levels of nitric oxide cause relaxation of smooth muscle cells and vasodilation by increasing cGMP levels within the cell. Increasing cAMP levels can modulate signaling pathways by activating the enzyme protein kinase A (PKA).

Signal transduction cascades. Signals often flow within the cell as cascades. As the signal is transmitted downstream from one signaling molecule to another, it is amplified. Two of the most well-defined cascades are the mitogen-activated protein kinase (MAPK) pathway leading to cell proliferation, and the glucose 6-phosphate (G6P) pathway leading to the metabolism of glycogen (Fig. 2).

In the case of the MAPK pathway, extracellular growth factors such as epidermal growth factor (EGF) bind to and activate the EGF receptor. Activation of the receptor causes autophosphorylation of the receptor and recruitment of signaling components Grb2 and SOS. Once these components are recruited to the membrane, they activate Ras, a G protein, by promoting the exchange of GDP to GTP. Ras-GTP binds and activates Raf-1, which then phosphorylates and activates mitogen-activated protein or extracellular signal-regulated kinase kinase (MEK). Activated MEK in turn phosphorylates and activates MAPK 1,2. Activated MAPK 1,2 is capable of phosphorylating and activating a number of proteins that are involved in transcriptional and translational functions. Several of the transcription factors activated by MAPK modulate genes that are important in regulating the cell cycle, thereby promoting cell division and proliferation.

The glucose 6-phosphate pathway works in a manner similar to the MAPK pathway except that one of the signaling components is the molecule cAMP. In the case of the glucose 6-phosphate pathway, glucagon binds to and activates the glucagon receptor. The activated glucagon receptor, which is coupled to a G protein, activates its G protein. The GTP-bound activated Gα subunit binds and activates adenylyl cyclase, which promotes the production of cAMP from ATP. Increased levels of cAMP activate cAMP-dependent protein kinase (PKA). Activated PKA phosphorylates and activates the enzyme phosphorylase kinase, which activates the enzyme glycogen-phosphorylase A. Activated phosphorylase A phosphorylates glycogen, leading to its breakdown to glucose, which is then released from the cell. In addition to this metabolic function, PKA can regulate gene expression (Fig. 1).

Regulation of physiological function. One of the most important functions of cell signaling is to control and maintain normal physiological balance within the body. Activation of different signaling pathways leads to diverse physiological responses (Fig. 2). For example, activation of the MAPK pathway results in cell proliferation, whereas activation of the cAMP pathway results in the release of glucose. In this manner all the normal physiological functions such as cell proliferation, death, differentiation, and metabolism are maintained. Signaling pathways in cells play more than the role of a simple transmitter. Several of the signaling pathways interact with each other and can serve as signal integrators. This added level of complexity is necessary in order to maintain normal physiology. An average cell in the body is constantly exposed to several different factors and hormones; therefore the final physiological outcome is dependent on the signal integration capabilities of the cell. Signal integration can occur in a number of ways: negative and positive feedback loops in pathways can modulate signals within a pathway; positive interactions between two signaling pathways can increase duration of signals; and negative interactions between pathways can block signals. Therefore, signaling pathways in cells play a vital role in maintaining physiological function. *See* CELL NUCLEUS; CELL ORGANIZATION; ENDOCRINE SYSTEM (VERTEBRATE); NORADRENERGIC SYSTEM.

Prahlad T. Ram; Ravi Iyengar

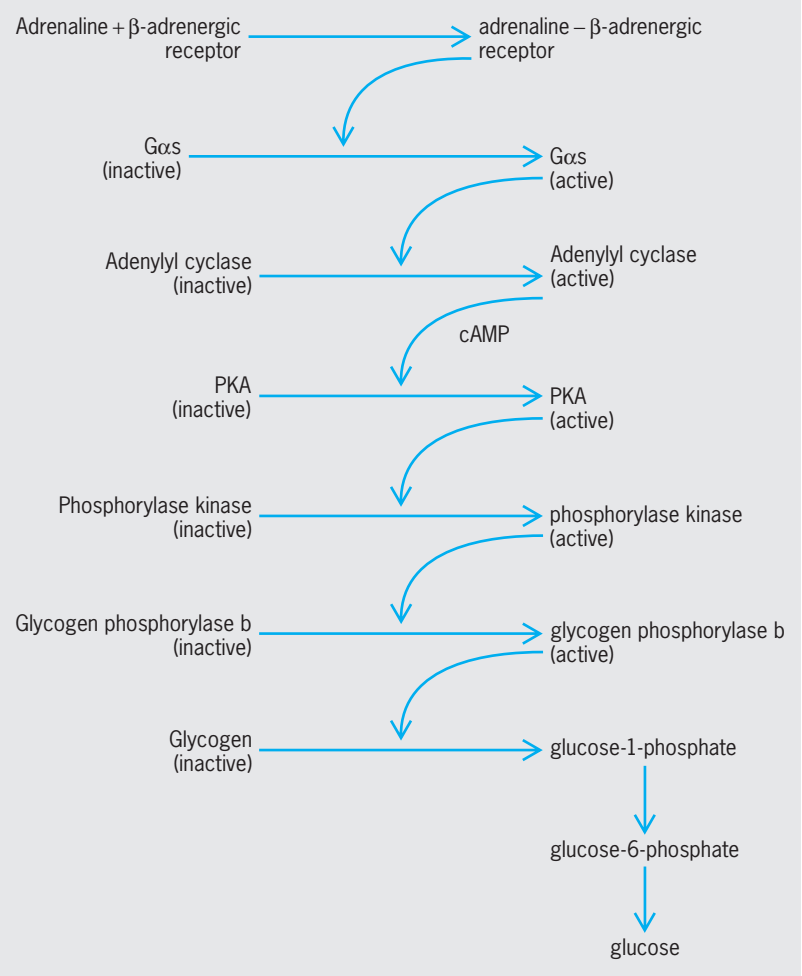

Fig. 3. Steps in the glucagon-stimulated release of glucose stores from cells.

Bibliography. H. E. Hamm, The many faces of G protein signaling, *J. Biol. Chem.*, 273:669–672, 1998; T. Hunter, Signaling: 2000 and beyond, *Cell*, 100:113–127, 2000; R. Iyengar and L. Birnbaumer (eds.), *G Proteins*, Academic Press, New York, 1989; J. D. Jordan and R. Iyengar, Modes of interactions between signaling pathways, *Biochem. Pharmacol.*, 55:1347–1352, 1998.

Significant figures

Digits that show the number of units in a measurement expressed in decimal notation.

Scientific notation is useful in showing which digits are significant. In scientific notation, a number is expressed as the product of a number 1 to 10 and a power of 10, or the product of 1 and a power of 10. Thus, the number 123,000 is 1.23×10^5 (see **table**).

The precision of a measurement is based on the size of the unit of measurement. The smaller the unit, the more precise is the measurement.

All the numbers expressing measurements are approximate. Care must be exercised not to increase the precision of a measurement simply by affixing zeros. For example, a measurement of 14.2 g, indicating precision to the nearest tenth, should not be written as 14.20 because this indicates precision to hundredths.

The absolute error in a measurement is one-half the unit. For example, the error in 480 m, measured to the nearest 10 m, is one-half of 10 m, or 5 m. The relative error, usually expressed as a percent, is the ratio of the absolute error to the measurement. The relative error for 480 m, measured to the nearest 10 m, is 5/480, or approximately 1%. *See* PERCENT.

Computations cannot improve the precision of the measurement. To add measures, they should all be rounded to the unit of the least precise measurement. The sum 8.6 cm + 0.14 cm + 2.75 cm is found by rounding each to tenths: 8.6 cm + 0.1 cm + 2.8 cm = 11.5 cm. Even by doing this, the absolute error might be as large as 0.05 + 0.005 + 0.005 or 0.06, and affect the result by as much as 0.1.

In multiplying and dividing approximate numbers, the product or quotient is rounded to the number of significant digits in the number with the fewest significant digits. For example, 6.2 m × 8.75 m by computation is 54.25. The product needs to be rounded to 54 m^2 so as to show two significant digits. *See* NUMBERING SYSTEMS; NUMERICAL ANALYSIS; STATISTICS.

Joseph N. Payne

Examples of significant digits and scientific notation

Measurement	Unit of measurement	Significant digits	Scientific notation
480 m	10 m	4 and 8	4.8×10^2
37.53 m	0.01 m	3, 7, 5, and 3	3.753×10^1
14.2 g	0.1 g	1, 4, and 2	1.42×10^1
193,000,000 mi	1,000,000 mi	1, 9, and 3	1.93×10^8
270.0 kg	0.1 kg	2, 7, 0, and 0	2.700×10^2

Bibliography. L. Elden, *Numerical Analysis: An Introduction,* 1990.

Silica minerals

Silica (SiO_2) occurs naturally in at least nine different varieties (polymorphs), which include tridymite (high-, middle-, and low-temperature forms), cristobalite (high- and low-temperature forms), coesite, and stishovite, in addition to high (β) and low (α) quartz. These forms have distinctive crystallography, optical characteristics, physical properties, pressure-temperature stability ranges, and occurrences (see **table**).

The transformation between the various forms are of two types. Displacive transformations, such as inversions between high-temperature (β) and low-temperature (α) forms, result in a displacement or change in bond direction but involve no breakage of existing bonds between silicon and oxygen atoms. These transformations take place rapidly over a small temperature interval and are reversible. Reconstructive transformations, in contrast, involve disruption of existing bonds and subsequent formation of new ones. These changes are sluggish, thereby permitting a species to exist metastably outside its defined pressure-temperature stability field. Two examples

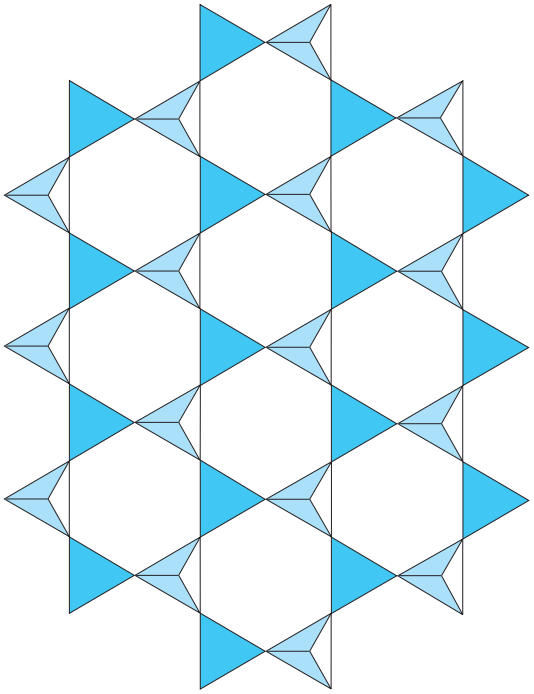

Fig. 1. Portion of an idealized sheet of tetrahedrally coordinated silicon atoms similar to that found in tridymite and cristobalite. Sharing of apical oxygens (which point in alternate directions) between silicons in adjacent sheets generates a continuous framework. (*After J. J. Papike and M. Cameron, Crystal chemistry of silicate minerals of geophysical interest, Rev. Geophys. Space Phys., 14:37–80, 1976*)

Optical and physical properties of silica minerals

Mineral	Specific gravity	Optic sign	Mean index	Nα or Nω*	N*	Nγ or N*	Cleavage	Color	Hardness (Mohs scale)	Twinning
Stishovite	4.35	Uniaxial (+)	1.81	1.799		1.826	N.R.†	Pale gray	N.R.	N.R.
Coesite	3.01	Biaxial (+)	1.59	1.594	1.596	1.597	None	Colorless	8	{021}
Low (α) quartz	2.65	Uniaxial (+)	1.55	1.544		1.553	{1011}, {0111}, {1010}‡	Colorless, white to variable	7	*c* axis {1120}, {1122}, and others
Low (α) cristobalite	2.32	Uniaxial (−)	1.48	1.487		1.484	None	Colorless, white, or yellowish	$6^1/_2$	{111}
Low (α) tridymite	2.26	Biaxial (+)	1.47	1.471–1.479	1.472–1.48	1.474–1.483	None to poor prismatic	Colorless, white	7	{110}

*Nα, Nβ, and Nγ are the three principal refractive indices for biaxial crystals; Nω and Nε are the principal refractive indices for uniaxial crystals.
†N.R. = not reported.
‡Cleavages not usually observed so that quartz is frequently described as having no cleavage.

of reconstructive transformations are tridymite ⇌ quartz and quartz ⇌ stishovite.

Crystal structure. The crystal structures of all silica polymorphs except stishovite contain silicon atoms surrounded by four oxygens, thus producing tetrahedral coordination polyhedra. Each oxygen is bonded to two silicons, creating an electrically neutral framework. Stishovite differs from the other silica minerals in having silicon atoms surrounded by six oxygens (octahedral coordination). *See* CRYSTAL STRUCTURE.

Ideal high tridymite is composed of sheets of SiO_4 tetrahedra oriented perpendicular to the *c* crystallographic axis (**Fig. 1**) with adjacent tetrahedra in these sheets pointing in opposite directions. The apical oxygens of the tetrahedra are bonded to silicons in neighboring sheets, thus generating a continuous framework with hexagonal (P_{6_3}/mmc) symmetry. Naturally occurring meteoritic high tridymite deviates somewhat from the ideal structure because adjacent sheets are slightly offset, producing orthorhombic (C_{222_1}) symmetry (**Fig. 2**). Low tridymite has a similar structure, but displacive transformations alter the geometry of the sheets, producing a lower symmetry. Terrestrial low tridymite contains three types of sheets in a complex stacking sequence resulting in triclinic (pseudoorthorhombic) symmetry. Meteoritic low tridymite has a monoclinic (Cc) cell caused by a distortion of the hexagonal rings shown in Fig. 1. Little is known about the structure of middle tridymite, but it is believed to be hexagonal.

High cristobalite, like tridymite, is composed of parallel sheets of SiO_4 tetrahedra with neighboring tetrahedra pointing in opposite directions. However, the hexagonal rings are distorted and adjacent sheets are rotated 60° with respect to one another, resulting in the geometry shown in **Fig. 3**. The structure is cubic (Fd3m), with the layers of tetrahedra parallel to (111). Further distortion of these sheets at low temperature causes an inversion to tetragonal ($P_{4_1 2_1 2}$ or $P_{4_3 2_1 2}$) low crystobalite.

Coesite also contains silicon atoms tetrahedrally coordinated by oxygen. These polyhedra share corners to form chains composed of four-membered rings. Bonding between chains creates a continuous framework of monoclinic symmetry (C2/c) in which each oxygen is shared between two silicons. The structure, part of which is shown in **Fig. 4**, is significantly more dense than tridymite and cristobalite.

Silicon in stishovite is octahedrally coordinated by oxygen (**Fig. 5**). These coordination polyhedra share edges and corners to form chains of octahedra parallel to the *c* crystallographic axis. The resultant structure is tetragonal (P_{4_2}/mnm) and is similar to the structure of rutile (TiO_2).

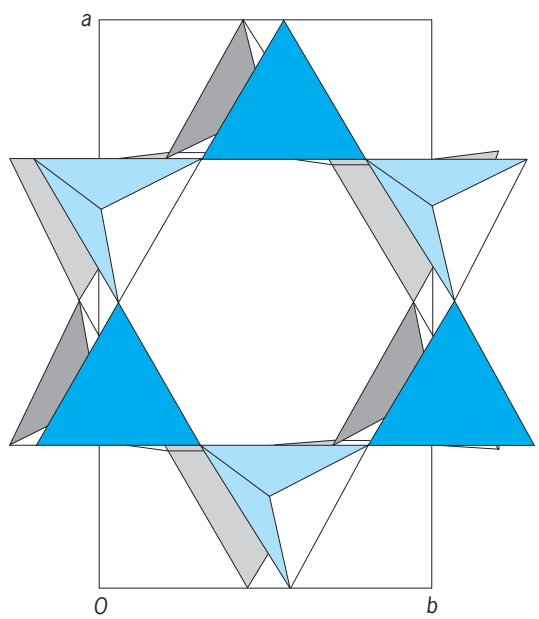

Fig. 2. Portion of the orthorhombic high tridymite structure viewed down the *c* axis showing the slight displacement between adjacent sheets. In ideal hexagonal high tridymite, neighboring sheets would be superimposed. (*After J. J. Papike and M. Cameron, Crystal chemistry of silicate minerals of geophysical interest, Rev. Geophys. Space Phys., 14:37–80, 1976*)

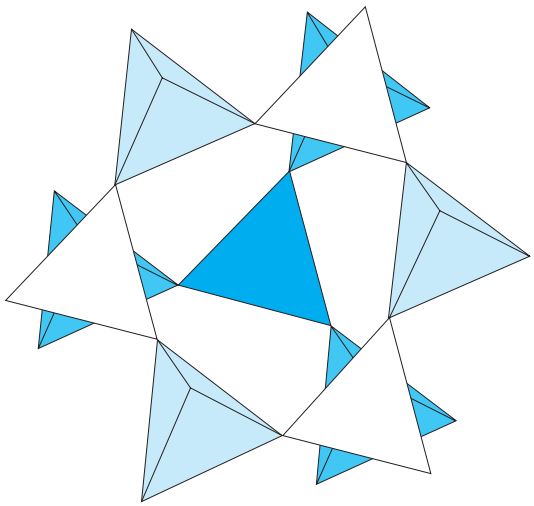

Fig. 3. Portion of cubic high cristobalite illustrating the distortion of tetrahedral sheets, which are oriented parallel to (111), and the 607 rotation of adjacent sheets. (*After J. J. Papike and M. Cameron, Crystal chemistry of silicate minerals of geophysical interest, Rev. Geophys. Space Phys., 14:37–80, 1976*)

Properties. The physical, optical, and chemical characteristics of the silica polymorphs vary in relation to their crystal structures. Stishovite, which has the most dense packing of constituent atoms, has the highest specific gravity ($G = 4.35$), whereas tridymite and cristobalite, which have relatively open structures, are the silica polymorphs with the lowest specific gravities ($G_{Tr} = 2.26$, $G_{Cr} = 2.32$). In all instances the β forms of the polymorphs have less densely packed structures, hence lower specific gravities, than α counterparts. The refractive indices of the polymorphs are also related to crystal structure, and the mean refractive index is proportional to specific gravity, ranging from 1.47 for tridymite to 1.81 for stishovite. Specific optical and physical properties for the various silica minerals are given in the table. Twinning is common in most of these minerals and, as in the case of quartz, frequently results from the inversion of a high-temperature, high-symmetry form to a low-temperature, low-symmetry form.

Chemically, all silica polymorphs are ideally 100% SiO_2. However, unlike quartz which commonly contains few impurities, the compositions of tridymite and cristobalite generally deviate significantly from pure silica. This usually occurs because of a coupled substitution in which a trivalent ion such as Al^{3+} or Fe^{3+} substitutes for Si^{4+}, with electrical neutrality being maintained by monovalent or divalent cations occupying interstices in the relatively open structures of these two minerals. Such substitutions may decrease the SiO_2 content to as little as 95 wt %.

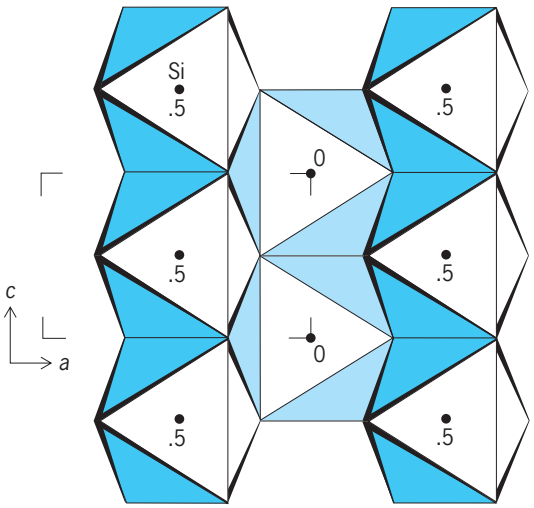

Fig. 5. Octahedral oxygen coordination polyhedra surrounding central silicons in the stishovite structure. The numbers (0, 0.5) indicate the relative height of the silicon atoms on the *a* crystallographic axis. The chains of octahedra formed by sharing polyhedral edges run parallel to the *c* axis. Chains at different levels are interconnected by silicon-oxygen bonds. (*After J. J. Papike and M. Cameron, Crystal chemistry of silicate minerals of geophysical interest, Rev. Geophys. Space Phys., 14:37–80, 1976*)

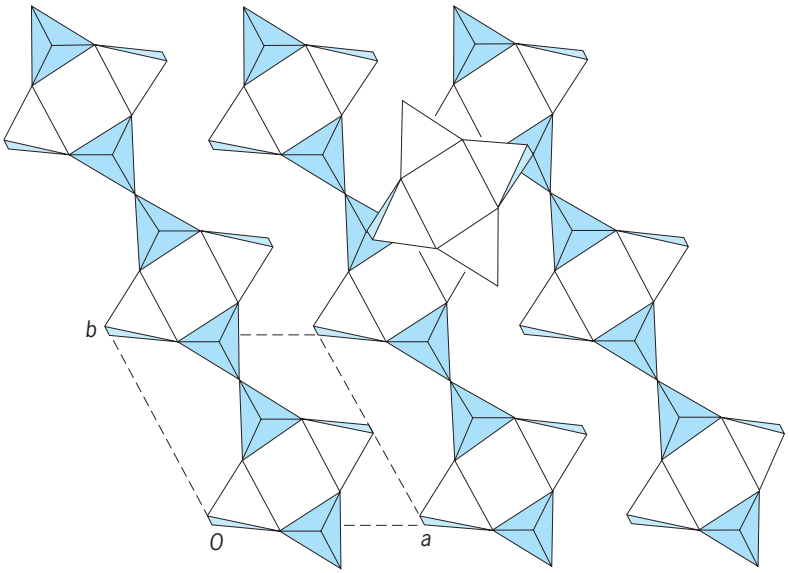

Fig. 4. Four-membered tetrahedral rings in coesite as seen in the plane perpendicular to the twofold axis of symmetry. Bonding between rings in this plane and adjacent planes (as indicated by the single ring at an elevated level) produces the dense coesite structure. (*After J. J. Papike and M. Cameron, Crystal chemistry of silicate minerals of geophysical interest, Rev. Geophys. Space Phys., 14:37–80, 1976*)

Phase relationships and occurrences. The pressure-temperature stability fields for the silica minerals are shown in **Fig. 6**. Low quartz is the stable form at low temperatures and pressures. At 1 atm (10^2 kilopascals) high quartz is stable from 573 to 870°C (1063 to 1600°F), at which temperature tridymite becomes stable. The stability field for cristobalite ranges from 1470 to 1720°C (2680 to 3130°F), at which temperature SiO_2 melts. Coesite and stishovite are stable only at exceedingly high pressures. However, because of the sluggishness with which reconstructive changes take place, various polymorphs can exist and occasionally crystallize metastably outside their defined limits of stability. Thus high tridymite (β_2) can persist to 163°C (325°F), at which temperature it transforms displacively to middle tridymite (β_1), which exists metastably to 117°C (243°F), at which

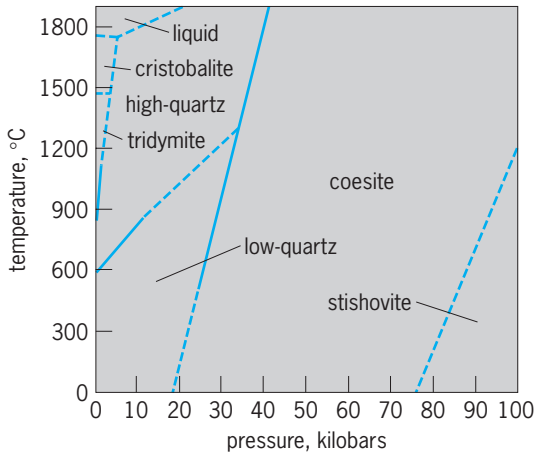

Fig. 6. Stability field of the silica polymorphs. °F = (°C × 1.8) + 32.1 kilobar = 10^2 MPa. (*After C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*)

low (α) tridymite forms. Impurities or disordering of the stacking sequence can cause these transformation temperatures to vary and possibly merge into a single high/low transformation. High cristobalite can exist metastably below 1450 to 268°C (or 2640 to 514°F; or lower), at which temperature it transforms displacively to low cristobalite. Variation in the temperature of inversion is caused by compositional variability and stacking disorder.

Tridymite and cristobalite crystallize as primary minerals in siliceous volcanic rocks such as rhyolites, trachytes, and andesites in which both minerals may occur in vesicles (cavities) or in the groundmass. Less frequently they are found in basaltic igneous rocks. They have been identified in siliceous sedimentary rocks, such as sandstones and arkoses, subjected to high temperatures during contact thermal metamorphism. Both minerals also occur in the silicate portions of meteorites and in some lunar igneous rocks. Cristobalite, and to a much lesser extent tridymite, is a common intermediate product in the transformation of amorphous biogenic silica (opal A) in marine sediments to quartz during the process of diagenesis (the lithification and alteration of unconsolidated sediments at low temperatures and pressures). The diagenetic sequence is as follows: opal A → opal CT (= porcellanite = cristobalite containing nonessential water in which the stacking sequence of silica sheets is disordered) → chalcedony (microcrystalline quartz) → quartz.

Naturally occurring coesite and stishovite were both first discovered in the Coconino sandstone adjacent to Meteor Crater, Arizona. These minerals are considered mineralogical indicators of the high pressures associated with the shock wave that resulted from the crater-forming meteorite impact. Since their discovery at Meteor Crater, they have both been found in the highly shocked rocks surrounding Ries Crater in Germany. Coesite has also been identified in rocks associated with several other impact or suspected impact craters, in tektites, in material ejected from human-caused explosion craters, and as inclusions in diamonds from kimberlite originating at great depth (hence at great pressure) within the Earth. Stishovite and coesite, therefore, are of great geological significance as indicators of very high pressure due either to depth of formation within the Earth or to shock processes such as those occurring during meteorite impact. *See* COESITE; QUARTZ; STISHOVITE. John C. Drake

Bibliography. W. H. Blackburn and W. H. Dennen, *Principles of Mineralogy*, 2d ed., 1993; W. A. Deer, R. A. Howie, and J. Zussman, *Rock Forming Minerals*, vol. 4: *Framework Silicates*, 1963; C. Frondel, *Dana's The System of Mineralogy*, 7th ed., vol. 3: *Silica Minerals*, 1962; K. Frye (ed.), *The Encyclopedia of Mineralogy*, 1982; J. J. Papike and M. Cameron, Crystal chemistry of silicate minerals of geophysical interest, *Rev. Geophys. Space Phys.*, 14:37–80, 1976; D. Stöffler, Coesite and stishovite in shocked crystalline rocks, *J. Geophys. Res.*, 76:5474–5488, 1971.

Silicate minerals

All silicates are built of a fundamental structural unit, the so-called SiO_4 tetrahedron. The crystal structure may be based on isolated SiO_4 groups or, since each of the four oxygen ions can bond to either one or two silicon (Si) ions, on SiO_4 groups shared to form complex isolated groups or indefinitely extending chains, sheets, or three-dimensional networks. Mixed structures in which more than one type of shared tetrahedra are present are also known. *See* SILICON.

Classification. Silicates are classified according to the nature of the sharing mechanism, as revealed by x-ray diffraction study. An abbreviated form of such a classification is given in the **table**. The sharing mechanism gives rise to a characteristic ratio of Si to O, but it is possible for oxygen ions that are not bonded to Si to be present in the structure, and sometimes some or all of any aluminum present must be counted as equivalent to Si. The constitution and classification of the silicates were controversial before the advent of x-ray structure analysis methods in 1912. The silicates were then usually considered as salts of silicic acids, many of them hypothetical, and a chemical classification, such as orthosilicates, metasilicates, and the like, was applied.

Structure. A dominant feature of the crystal chemistry of the silicates that in large part determines the chemical complexity of these species is the dual role played by aluminum in the crystal structure. The radius ratio of this ion with oxygen is near the critical value between four-coordination and six-coordination, and the aluminum (Al) ion can occur in one or the other or both roles simultaneously. When in four-coordination, the trivalent Al ion substitutes for the quadrivalent Si ion, introducing a valence deficiency of one unit and requiring a concomitant substitution of another cation elsewhere in the crystal structure to provide valence compensation. This mechanism usually involves the coupled substitution of a divalent for a monovalent cation, as of calcium (Ca) for sodium (Na), less frequently of a trivalent

Silicate structures and their characteristic Si/O ratios

Type	Nature of Si-O linkage	Si/O ratio	Examples
Nesosilicates	Isolated SiO_4 groups	1:4	Olivine, garnet
Sorosilicates	Isolated compound groups, Si_2O_7, Si_6O_{18}, and so on	2:7, 6:18, and so on	Thortveitite, beryl
Inosilicates	One-dimensional extended chains and bands	1:3, 4:11, and so on	Amphiboles, pyroxenes
Phyllosilicates	Two-dimensional extended sheets	2:5	Mica, clays, talc, chlorite
Tectosilicates	Three-dimensional networks	1:2	Feldspars, feldspathoids, zeolites

for a divalent ion. Other less common mechanisms involve coupled omissions or substitutions among the anionic units of structure. In some silicates, such as those with a silicon-oxygen framework based on a polymorph of silica, the serial substitution of Al for Si is compensated by the entrance of a cation, such as Na, into vacant interstices of the crystal structure.

The detailed crystallographic and physical properties of the various silicates are broadly related to the type of silicate framework that they possess (see **illus.**). Thus, the phyllosilicates as a group typically have a platy crystal habit, with a cleavage parallel to the plane of layering of the structure, and are optically negative with rather high birefringence. The inosilicates, based on an extended one-dimensional rather than two-dimensional linkage of the SiO_4 tetrahedra, generally form crystals of prismatic habit; if cleavage is present, it will be parallel to the direction of elongation. The tectosilicates commonly are equant in habit, without marked preference for cleavage direction, and tend to have a relatively low birefringence.

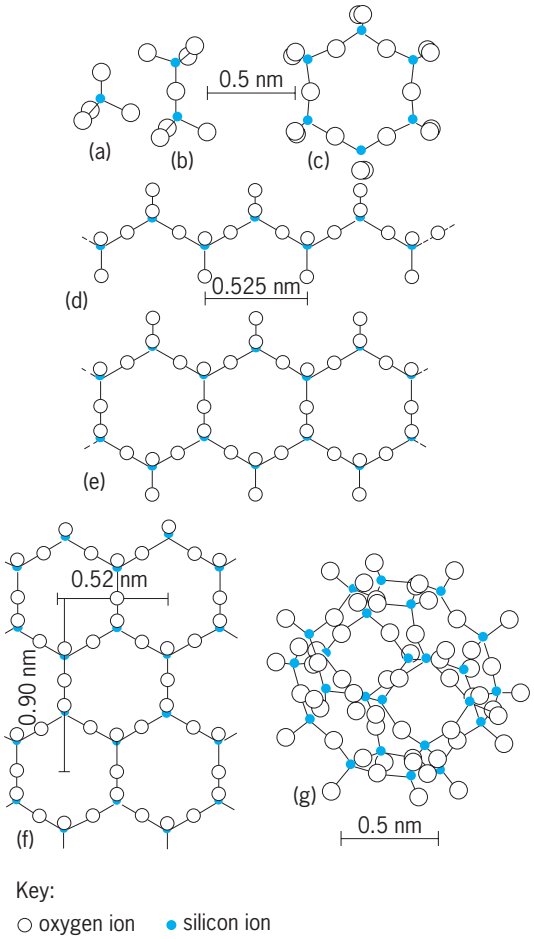

Forms of silicon-oxygen linkage. (*a*) Nesosilicate $(SiO_4)^{4-}$. (*b*) Sorosilicate $(Si_2O_7)^{6-}$. (*c*) Sorosilicate, or cyclosilicate $(Si_6O_{18})^{12-}$. (*d*) Inosilicate $(SiO_3)^{2-}$, showing chain structure of pyroxenes. (*e*) Inosilicate $(Si_4O_{11})^{6-}$, showing band structure of amphiboles. (*f*) Phyllosilicate $(Si_2O_5)^{2-}$, showing extended sheets. (*g*) Tectosilicate, showing three-dimensional structure of lazurite. (*After W. L. Bragg, Atomic Structure of Minerals, Cornell, 1937*)

Important minerals. Silicate minerals make up the bulk of the outer crust of the Earth and form in a wide range of geologic environments. Many silicates are of economic importance. Among the clays, feldspars, and refractory minerals, andalusite and wollastonite are used in the ceramic industries, mica as an electrical insulating agent, asbestos and exfoliated vermiculite as thermal-insulating agents, and garnet as an abrasive. Talc is a constituent of facial powder. Other silicates are important as ore minerals, beryllium being obtained from beryl, zirconium and hafnium from zircon, and thorium from thorite. Some silicates such as jadeite and nephrite are prized as ornamental materials, and peridot, garnet, tourmaline, and aquamarine are well-known gemstones. *See* CLAY, COMMERCIAL; CLAY MINERALS; GLASS; SILICATE PHASE EQUILIBRIA.

For discussions of certain silicate mineral groups *See* AMPHIBOLE; ANDALUSITE; CHLORITE; CHLORITOID; EPIDOTE; FELDSPAR; FELDSPATHOID; GARNET; HUMITE; MICA; OLIVINE; PYROXENE; SCAPOLITE; SERPENTINE; ZEOLITE.

Clifford Frondel

Silicate phase equilibria

Silicate phase equilibria studies define the conditions of temperature, composition, and pressure at which silicates can stably coexist. Silicate phase equilibria relations are used by geologists, ceramists, and cement manufacturers to explain the variation of composition of silica-bearing minerals, as well as their number and order of appearance in rocks, slags, glasses, and cements. Silicate phase equilibria relations are also useful to interpret the chemistry of refractories, boiler scale deposits, and welding fluxes.

Silica itself makes up nearly 60% by weight of the Earth's crust. The next most abundant oxides, in decreasing order, are Al_2O_3, CaO, Na_2O, FeO, MgO, K_2O, and Fe_2O_3; all of these occur principally combined with silica as silicates. Free silica and the hundreds of silicate minerals make up nearly 97% of the Earth's crust. The study of silicate phase equilibria was initiated by geologists seeking to apply the phase rule of J. Willard Gibbs to these abundant natural substances. Much of the work since 1907 has been done at the Geophysical Laboratory of the Carnegie Institution of Washington. *See* PHASE EQUILIBRIUM; SILICATE MINERALS.

Silicate phase equilibria are determined in systems generally specified by naming the components involved. Components are the smallest number of independently variable chemical constituents necessary to express the composition of each phase present. A binary system has two components, which may be simple oxides such as Na_2O and SiO_2, or complex compounds such as Na_2SiO_3 and $Na_2Si_2O_5$. In either case, the preferred order of writing the components or oxides follows the rule that the oxides are grouped according to increasing valence of the cation, and then in alphabetical order, as K_2O–Na_2O–CaO–Al_2O_3–SiO_2 for a quinary system and $K_2O \cdot Al_2O_3 \cdot 6SiO_2$ or $KAlSi_3O_8$ for a complex compound.

Equilibrium relations between molten silicate liquids and crystalline silicates, called liquidus equilibria, are most often determined. Equilibria between crystalline silicates in the absence of any liquid silicate melt, called subsolidus equilibria, are determined only infrequently.

Historically, research has been from binary or ternary systems toward more complicated systems. More than 500 systems have been studied to date. The most complicated studies involve limited portions of a system with six components. Specific portions of multicomponent systems are determined because of their importance. For example, in 1913, the first complete silicate phase equilibrium study, the binary system $NaAlSi_3O_8$ [albite (Ab)]–$CaAl_2Si_2O_8$ [anorhite (An)], representing the very abundant rock-forming plagioclase feldspars, was determined by N. L. Bowen, one of the most famous pioneering experimental petrologists. However, this binary system was merely a subsystem from the very complicated quaternary system Na_2O–CaO–Al_2O_3–SiO_2, which has not yet been completely determined.

Silicate crystal structures in general require at least 30% by weight of silica; hence the less siliceous portions of systems are often not determined. The limited number of possible silicate crystal structures determines the possible ratios of oxides to silica in silicates. Certain elements substitute for each other in silicate crystal structures; this reduces the possible varieties of silicates so that those making up the overwhelming bulk of rocks fall into less than a dozen major groups, called rock-forming silicates. Typical of such groups are the feldspars and feldspathoids in which alkali ions substitute for one another and aluminum and ferric iron ions substitute in part for silicon ions, and the olivine and pyroxene groups in which divalent iron, magnesium, manganese, and calcium ions substitute for each other. Most experiments on silicate phase equilibria have been and continue to be on the relations within and between these important groups of minerals. *See* FELDSPAR; FELDSPATHOID; OLIVINE; PYROXENE.

Dynamic and static methods. Both methods are used to determine equilibrium in silicate systems. Dynamic methods like those used for metals and alloys, such as measuring the heat effects of phase changes on heating or cooling, are not particularly satisfactory with silicates. These methods require large samples of silicates that are difficult to prepare, and require also that equilibrium be reached quickly. Silicates in general are slow to react, and supercooling or superheating of hundreds of degrees before reaction occurs is common. Many silicates react sluggishly at temperatures of 1000°C (1830°F) or higher. Crystals in liquid may persist metastably for weeks, nucleation of new phases may require months, and certain crystallographic processes, such as the ordering of aluminum and silicon atoms into different structural sites, may take a lifetime. Subsolidus reactions are extremely sluggish and may require geologic time before equilibrium is attained or even approached. It is this fact that permits the geologist and petrologist to determine the past thermal history of a rock; the higher-temperature products are preserved for study, even through very long cooling intervals, by the slow rate of attaining equilibrium.

Most silicate phase equilibria are determined by the static method of holding a sample under controlled conditions until equilibrium is attained, then quenching the sample for examination. Charges of 10–100 mg of pure components or glass, gel, or crystals of known composition are wrapped in platinum foil and held in air in furnaces in which the temperature is carefully regulated. Controlled partial pressures of oxygen or inert gases are required to study systems with components, such as FeO, MnO, Fe_2O_3, and TiO_2 that are oxidized or reduced by contact with air, or that alloy with platinum. The samples are quenched rapidly to room temperature by dropping them into water or mercury. Most silicate liquids with a ratio of oxygen to aluminum plus silicon of 2:1 or 3:1 are viscous and quench to glass. Crystals grown from liquid during the quench are usually fine-grained and fibrous and are easily recognized. Certain displacive polymorphic transitions in silicate crystals cannot be quenched and are studied by special techniques, but in general the crystals stable at high temperature can be quenched and studied at room temperature. The composition, amount, and number of glassy and crystalline phases are determined by optical microscopy, x-ray diffraction, and thermal analysis.

Establishment of equilibrium. Equilibrium is established when the products obtained by heating a sample to a given temperature are identical with the products of cooling a sample to that temperature; no requirement is made regarding the texture, shape, or grain size of the products. Another criterion used

in recognizing equilibrium is that no change of the sample can be observed after holding the charge at a given temperature for very long periods of time. Particularly useful for this purpose are thin, sharp, crystal fragments that sinter and become rounded on melting, or splinters of glass that devitrify to crystals before larger fragments are affected. When the same products are obtained regardless of starting materials, equilibrium is thought to be attained. At equilibrium the products are usually uniformly distributed in the sample, and there is consistency between the equilibrium results of the system studied and those portions common to other equilibrium systems. Finally, the results of equilibrium are consistent with the trends shown by natural occurrences of the same minerals.

Diagrammatic representation. The use of diagrams to present silicate phase equilibria is customary, since such diagrams express quantitatively the amount and composition of each phase present at any bulk composition in the system at any temperature. Even when molten, silicates have very low vapor pressures, and the vapor pressure and composition of the vapor are practically unaffected by large variations of temperature. Therefore the vapor phase is not considered, and the phase diagrams are drawn isobarically at a total pressure of 1 atm (10^2 kilopascals). However, diagrams involving only a single component are usually drawn with temperature plotted against total pressure.

Binary systems. Binary systems are plotted with composition (usually in percentage by weight) against temperature at constant pressure. An example of such a binary system with complete solid solution without a maximum or minimum is the system $NaAlSi_3O_8$–$CaAl_2Si_2O_8$ (**Fig. 1**). The region labeled liquid plus feldspar is a two-phase field where feldspar, whose composition lies along the lower curve, coexists with liquid, whose composition lies on the upper curve. A line connecting the liquid and feldspar that coexist at a given temperature is called a tie line. The curve separating the liquid field from the liquid-plus-feldspar field is called the liquidus. The

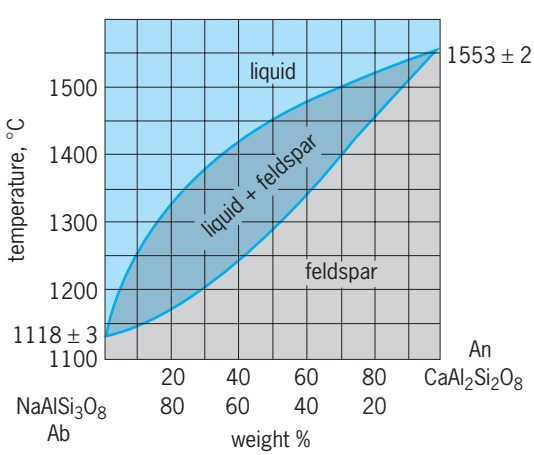

Fig. 1. System $NaAlSi_3O_8$–$CaAl_2Si_2O_8$. °F = (°C × 1.8) + 32. (*After N. L. Bowen, Amer. J. Sci., ser. 4, 35:577–599, 1913*)

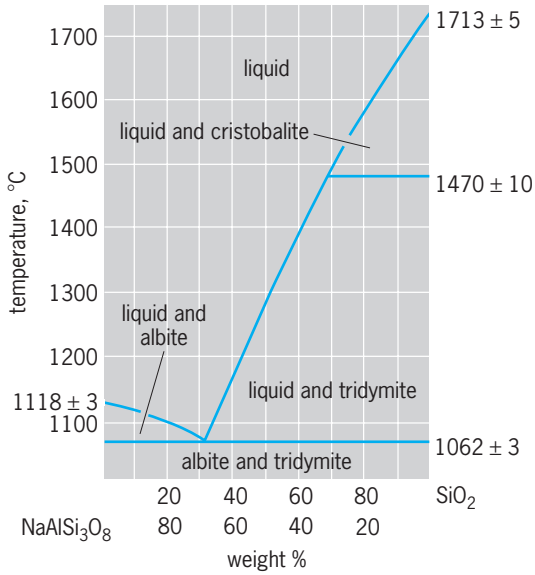

Fig. 2. System $NaAlSi_3O_8$–SiO_2. (*After J. F. Schairer and N. L. Bowen, Amer. J. Sci. 254:161, 1956*)

curve separating the liquid-plus-feldspar field from the feldspar field is called the solidus.

A composition such as 50% by weight $NaAlSi_3O_8$ and 50% by weight $CaAl_2Si_2O_8$, or more simply $Ab_{50}An_{50}$, is all liquid at temperatures above 1450°C (2640°F). The first crystals to appear in it on cooling would be of composition $Ab_{18}An_{82}$ on reaching the liquidus at 1450°C (2640°F). If equilibrium is maintained during further cooling, these crystals will react continuously with the liquid and change in composition toward $Ab_{50}An_{50}$ along the solidus and at the same time increase in amount. Simultaneously the composition of the liquid will change along the liquidus toward $Ab_{86}An_{14}$. The ratio of liquid to crystals can be found at any temperature by the lever rule, which states that the amounts of the two phases are inversely proportional to the distances of their compositions from the bulk composition. At 1285°C (2345°F), the last liquid of composition $Ab_{86}An_{14}$ will react with crystals and the entire mass will consist of feldspar of $Ab_{50}An_{50}$ composition.

Fractional crystallization occurs when crystals are removed from contact with the liquid for any reason, such as by crystal settling or by protection by armoring of the core by a rim of different composition. The magnitude of the change produced in the remaining liquid can be determined from the diagram if the composition and amount of crystals isolated are known. Thus, the equilibrium diagram can be used to determine the consequences of disequilibrium.

The diagram for the binary system $NaAlSi_3O_8$–SiO_2 shows no solid solution in the crystalline phases (**Fig. 2**). In most systems of this type a solid solution in the crystalline phases does exist, but it is too limited to be observed by ordinary techniques or to be shown on the scale of the diagram.

In the two-phase region of this system, a crystal of composition of the appropriate components exists with liquid until, upon heating, it dissolves in the

liquid at the temperature appropriate for the particular bulk composition chosen. When cooled to the temperature of the eutectic, the first crystal of the composition of the second component appears, and the temperature remains constant until all the liquid has crystallized. The bulk composition is then represented as an aggregate of crystals of both components. Only at the eutectic composition and when the components are pure do crystals melt directly to a liquid without an intervening stage of crystals and liquid. Contrary to the behavior of systems of metals, eutectic mixtures in silicate systems rarely form intimately intergrown aggregates with so-called eutectoid texture. Discrete crystals of both phases appear instead.

Many other types of binary systems are known, and one additional type is described below.

Ternary systems. Ternary systems are plotted in triangular coordinates with the three components at the apexes of an equilateral triangle; temperatures on some surface such as the liquidus are indicated by contours projected onto the diagram, as in the system $KAlSiO_4$–$NaAlSiO_4$–SiO_2 (**Fig. 3**). The fields of primary crystallization of the various solid phases are indicated, as are the compositions and melting temperatures of the end-members of solid solutions, intermediate compounds, and pure components of the system. Binary systems appear as the lines on ternary diagrams; the system $NaAlSi_3O_8$–SiO_2 discussed above appears as a segment of the boundary $NaAlSiO_4$–SiO_2.

Isothermal sections. Ternary systems may also be graphically represented by the use of isothermal sections. Phase relations and compositions are shown in detail at a specified temperature. An isothermal section at 600°C (1110°F) for the system $KAlSiO_4$–$NaAlSiO_4$–SiO_2 is a typical example (**Fig. 4**). More complex systems are represented by tetrahedrons or by various projections.

Pseudobinary system. A pseudobinary system is shown in the temperature-composition diagram for $KAlSi_3O_8$–$NaAlSi_3O_8$ (**Fig. 5**). These components make up the abundant alkali feldspars. The system is pseudobinary; the equilibria represented by the heavy lines are truly binary, but the equilibria represented by the light lines are ternary and indicate a section through regions in which the compositions of the phases do not lie in the plane of the diagram. The ternary nature of this system results from the incongruent melting of solid solutions at and near $KAlSi_3O_8$ composition to leucite, $KAlSi_2O_6$, and a liquid whose composition is indicated by the intersection of the liquidus surface and the prolongation of the line drawn from leucite composition through the bulk composition of the feldspar under consideration. These relations are apparent in the illustration of the system $KAlSiO_4$–$NaAlSiO_4$–SiO_2 (Fig. 3).

The truly binary portion of the system $KAlSi_3O_8$–$NaAlSi_3O_8$ is an example of another major group of binary systems with a minimum in the liquidus and solidus, or with a maximum. The one crystal in the solid solution series at the composition of the minimum melts directly to a liquid without an intervening region of liquid and crystals.

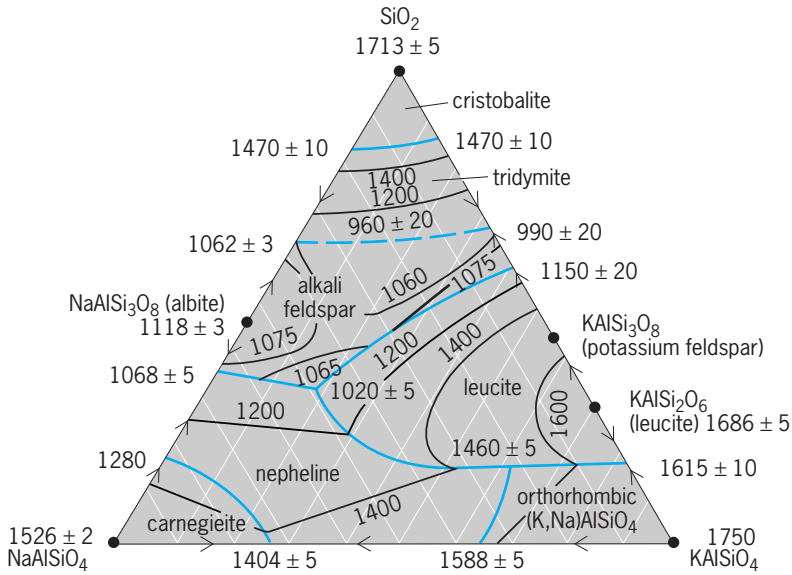

Fig. 3. Ternary system $KAlSiO_4$–$NaAlSiO_4$–SiO_2. The primary phases are indicated, and their primary fields of crystallization are outlined. Arrows indicate direction of falling temperature. Temperatures are given in °C [°F = (°C × 1.8) + 32]. (*After J. F. Schairer, J. Geol., 58:514, 1950*)

The approximate limits of a two-phase subsolidus region are shown, within which a homogeneous alkali feldspar of intermediate composition will separate into two alkali feldspar solid solutions. The curve defining this region is the locus of points representing pairs of compositions that coexist; it is called a solvus.

A subsolidus, nonquenchable polymorphic transition from monoclinic to triclinic symmetry is also shown.

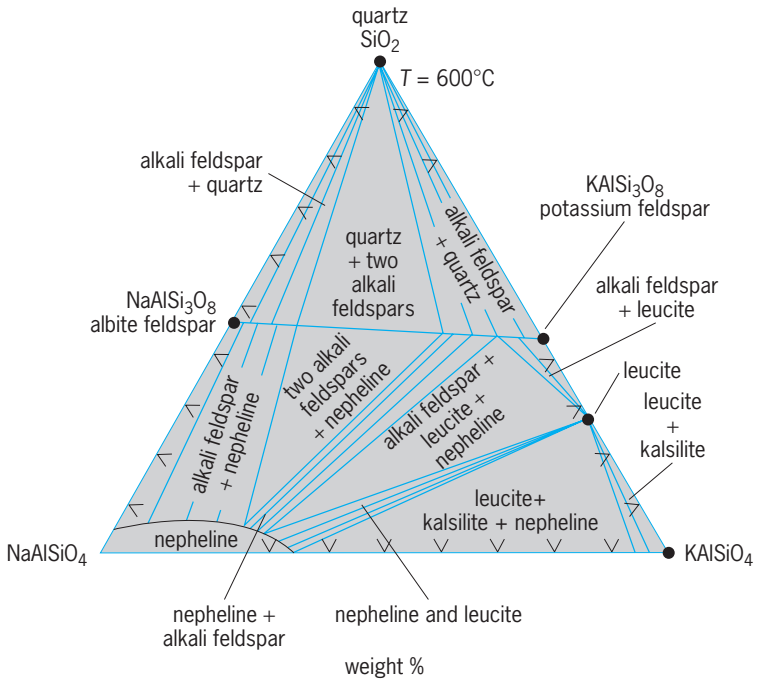

Fig. 4. Isothermal section, in part schematic, showing system, $KAlSiO_4$–$NaAlSiO_4$–SiO_2, at 600°C (1110°F). (*After O. F. Tuttle and J. V. Smith, Amer. J. Sci., 256:587, 1958*)

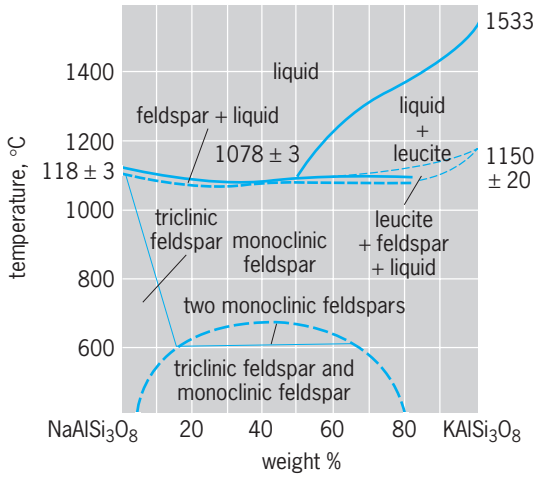

Fig. 5. System $KAlSi_3O_8$–$NaAlSi_3O_8$. Heavy solid or broken curves refer to binary equilibrium. Light solid or broken curves refer to ternary equilibrium. °F = (°C × 1.8) + 32. (*After J. F. Schairer, J. Geol., 58:515, 1950; N. L. Bowen and O. F. Tuttle, J. Geol., 58:497, 1950; W. S. MacKenzie, Amer. J. Sci., Bowen vol., p. 331, 1952*)

Applications. Silicate phase diagrams permit the user to trace the paths of crystallization of a given composition both when equilibrium is maintained and when disequilibrium occurs. These paths place limits on possible relations and thereby restrict interpretation of natural occurrences.

There are several important generalizations which result from the determination of hundreds of phase diagrams for silicates. The most important of these is the generalization that solid solutions are very prevalent among silicates, and wide variations of composition are possible within a given mineral group. An important consequence is that the composition of the crystal or crystals growing from silicate melt changes as the temperature in the system varies. The liquids also change composition during this process and, if liquids are separated from crystals, fractionation and differentiation of silicate magma can proceed very extensively. Furthermore, careful study of the compositions of coexisting mineral phases can, after comparison with experimental results, indicate details of the temperatures and pressures at which the minerals were formed.

The extent of solid solution in a mineral group varies with the temperature and pressure; the limits are indicated by the appearance of another crystalline phase. High temperature favors greater solid solution in general, whereas high pressure decreases the amount of solid solution. The extent of solid solution may be used, in certain cases where the pressure of formation is known or can be approximated, to determine the temperature of formation. The method requires that two minerals with a common component coexist at equilibrium. If they are cooled so that they do not react, then their compositions, by comparison with the appropriate solvus, will indicate the temperature at which the assemblage formed. Solvi are known in the pyroxene, feldspar, nepheline, and other abundant mineral groups, so this method has wide geologic application. *See* GEOLOGIC THERMOMETRY.

Since 1950 a few experimental studies of the distribution of trace elements between coexisting silicates have been made. The results suggest that the minor elements may also be useful to evaluate the temperature and pressure of formation of mineral assemblages.

Another generalization is that there are frequently sharp changes in the number and relative amounts of the phases present as a given composition is cooled from an all-liquid to an all-crystalline state. Other sharp changes occur even in the crystalline state. A great many points have been established, most of which are peritectic (reaction) points rather than eutectic points. A decrease in temperature may cause an assemblage of minerals and liquid to transform completely into other minerals plus liquid, to several new minerals, or to a single mineral of intermediate composition. For example, siliceous liquid and leucite, $KAlSi_2O_6$, react with each other on cooling to form potassium feldspar, $KAlSi_3O_8$, at 1150°C (2100°F).

Phase equilibria studies of aluminosilicates have shown that, when incongruent melting occurs, the liquid formed is richer in silica than the newly formed crystal in most cases. This generalization is of considerable significance in petrogenesis, as it indicates a reason for the progressive enrichment in silica of late magmatic differentiates. *See* MAGMA.

Phase equilibria studies of silicate systems have also shown that the hypothesis of the immiscibility of granitic and basaltic magmas is not useful geologically. Liquid immiscibility has been shown to occur experimentally, but only in compositions that are not found in nature.

The available phase diagrams show that the silicate liquids that crystallize last and at the lower temperatures, either by equilibrium or disequilibrium crystallization, tend to be enriched in ferrous iron, silica, and alkali aluminosilicates with a molecular ratio of alkalies to alumina of unity. Further crystallization tends to remove the ferrous iron, and the last liquid remaining very closely approaches the system $KAlSiO_4$–$NaAlSiO_4$–SiO_2 shown above. This is known as petrogeny's residua system. Two major types of rocks are represented by this system. The most abundant are the granites, represented by the subsystem $KAlSi_3O_8$–$NaAlSi_3O_8$–SiO_2. Analyses of hundreds of granites plot in or close to the low-temperature "valley" on the liquidus of this system, running from the $KAlSi_3O_8$–$NaAlSi_3O_8$ binary toward SiO_2. This constitutes a major reason for believing that most granites originate from silicate magmas.

The second major type of rocks represented are the silica-deficient feldspathoidal syenites of the subsystem $KAlSiO_4$–$NaAlSiO_4$–$KAlSi_3O_8$–$NaAlSi_3O_8$. More experimental detail is required to understand all of the geologic consequences of this system, but there is excellent agreement between synthetic and natural assemblages of phases. A major problem remains in finding the reason why certain magmas differentiate toward the silica-deficient subsystem. The

solution to this problem undoubtedly involves many additional components. It is apparent, however, that the join $KAlSi_3O_8$–$NaAlSi_3O_8$, where binary, is a barrier or hump across which liquids do not move as a consequence of equilibrium crystallization.

The concept of incompatible mineral assemblages is another major consequence of studies in silicate phase equilibria. Nepheline, $NaAlSiO_4$, is incompatible at equilibrium with quartz, SiO_2, as they react to form albite feldspar, $NaAlSi_3O_8$. Similarly leucite, $KAlSi_2O_6$, and quartz are incompatible and react to form potassium feldspar, $KAlSi_3O_8$. Tens of incompatible assemblages are now known to petrologists. The recognition of any of these assemblages in the field indicates unusual conditions in which equilibrium was not attained.

Since 1948, silicate phase equilibria in systems with oxidizable components have been intensively studied to determine the effect of variations of oxygen pressure upon the phase relations. The results show that, when the oxygen partial pressure is changed, the temperature at which minerals containing an oxidizable component can exist is changed. The magnitude of the changes is quite remarkable; a small fraction of an atmosphere of oxygen may change the stability limits of an iron-bearing silicate by as much as 300°C (570°F). This principle, of course, is fundamental to the operation of blast furnaces. In nature, changes of partial pressure of oxygen are expected as gases are expelled from crystallizing magma, and these changes are now believed to affect profoundly the course of differentiation of the magma. The trend in silicate research is toward more detailed study of this effect, particularly as it affects the silica content of the residual liquid.

Much geological field evidence indicates that the temperatures required to obtain silicate liquids in experimental work are hundreds of degrees higher than those determined by various methods to have existed in magmatic assemblages or observed in volcanic eruptions. Silicate phase equilibria studies of increased complexity indicate that, as the number of components is increased, the temperatures required for liquid to form are lowered. The magnitude of this effect is inadequate to explain the geological discrepancy. Geologists have suggested that volatile components, notably H_2O, cause the great lowering of liquidus temperatures. This has indeed been shown to be the case, but experimentally the volatile components must be contained by high external pressure. Pressures of H_2O of 15,000 lb/in.^{2} (103 megapascals) lower the temperatures of the liquidus as much as 700°C (1290°F); simultaneously, several percent of H_2O dissolve in the molten silicate. Reactions are greatly accelerated, and important groups of hydrous minerals become stable. So spectacular have been the results that most current research in silicate phase equilibria for geological applications is conducted under high pressures of volatile components. *See* HIGH-PRESSURE MINERAL SYNTHESIS.

Thermodynamic theory is sufficiently developed to permit silicate phase equilibria to be calculated when values of the necessary physical-chemical parameters are available and ideal solution occurs. Accurate measurements of all these parameters are difficult, and the results of most phase equilibria studies are not precise enough to allow calculation of thermodynamic parameters. Some simple subsolidus and liquidus relations have been rigorously calculated. In general, however, the quality of calculations of complex phenomena is poorer than rough experimental work, the usual difficulties of establishing equilibrium notwithstanding. *See* CEMENT; CERAMICS; GLASS; LITHOSPHERE; PETROLOGY; PORCELAIN; REFRACTORY.

David B. Stewart

Bibliography. F. A. Hummel, *Introduction to Phase Equilibria in Ceramic Systems*, 1984; Y.-H. Li, *A Compendium of Geochemistry*, 2000; B. Mysen and P. Richet, *Silicate Glasses and Melts: Properties and Structure: Developments in Geochemistry*, 2005; R. S. Roth (ed.), *Phase Equilibria Diagrams: Phase Diagrams for Ceramists: Oxides*, 1995; J. V. Walther, *Essentials of Geochemistry*, 2005; S. L. Webb, *Silicate Melts: With 33 Figures: Lecture Notes in Earth Science*, 1997.

Silicate systems

Assemblages of silicate minerals. Studies of the nature of these minerals provide information on their origin and their role in rock formation. This article discusses kinetic processes in silicate systems, oxidation-reduction equilibria in high-temperature systems, and volatile equilibria in systems at high temperatures and pressures. *See* SILICATE MINERALS.

Kinetic Processes

The minerals that form by crystallization in a magma (molten rock) or by recrystallization in a solid rock are determined by phase equilibria and kinetics. Phase equilibria are a measure of which minerals are the most stable for a given temperature, pressure, and bulk composition. In general, minerals with lower free energies form at the expense of minerals with higher free energies. The mineral assemblage with the lowest free energy for a given set of conditions is the most stable and is known as the equilibrium mineral assemblage. Kinetics determines the rate at which chemical reactions occur; it includes such factors as nucleation rates, crystal growth rates, and diffusion rates. The rate at which a rock forms, relative to the kinetics associated with particular reactions, will determine whether the equilibrium mineral assemblage for a given set of conditions is formed. *See* PHASE EQUILIBRIUM.

Metastable equilibrium. An equilibrium mineral assemblage formed at one set of conditions of temperature, pressure, and composition may no longer be the lowest free-energy assemblage if the conditions change. Minerals that persist at conditions where they are not part of the equilibrium assemblage are said to exist metastably. Such metastable assemblages are common in igneous and metamorphic rocks, and they may persist almost indefinitely if the kinetic factors associated with the nucleation and growth

of the stable assemblage are extremely slow. Examples of metastablity include zoned minerals, volcanic glasses, and high-grade metamorphic rocks at room temperature and pressure. Minerals may also form metastably if the nucleation and growth rates of the equilibrium assemblage are slow relative to the nucleation and growth rates of a metastable assemblage. In such situations, minerals will nucleate and grow spontaneously if the total free energy of the products of the reaction is less than that of the reactants; it is not necessary that the products have the lowest possible free energy. An example of metastable crystallization occurs in volcanic rocks when cristobalite or tridymite forms under conditions where quartz has a lower free energy. Although all three minerals have the identical composition, silicon dioxide (SiO_2), tridymite, and cristobalite nucleate much more readily than does quartz because their structures are less complex.

Nucleation. The first step in the formation of a new mineral is the nucleation of seed crystals or nuclei of that mineral. Nucleation of a new crystal may occur on the surface of another crystal or on a plane of disruption within a crystal or melt, in which case the process is known as heterogeneous nucleation. Nucleation that occurs spontaneously within a liquid or solid, independent of local inhomogeneities, is known as homogeneous nucleation. The formation of individual nuclei is controlled by the free-energy change (ΔG) associated with the growth of small particles of that crystal, as in Eq. (1), where

$$\Delta G_{\text{total}} = \Delta G_V + \Delta G_S \qquad (1)$$

ΔG_V = volume free-energy change and ΔG_S = surface free-energy change. Thus this relationship includes the free energy associated with transforming a unit volume of material from the reactants to the products (volume free-energy change) and the free energy associated with the formation of a new crystal interface (surface free-energy change). The surface free-energy term for the growth of a new particle is always positive and is directly proportional to the surface area. The volume free-energy term may be either positive or negative: if it is positive, no nuclei will form; if it is negative, nuclei will grow spontaneously when the decrease in ΔG_V with increasing size exceeds the increase in ΔG_S, so that ΔG_{total} decreases as the nucleus grows. For very small particles, ΔG_{total} increases with increasing radius and, as a result, small nuclei are rapidly resorbed. Particles having a radius smaller than some critical radius r_c can grow only by random collisions. If ΔG_V is negative, however, there is a critical size beyond which ΔG_{total} decreases with increasing radius and nuclei grow spontaneously (**Fig. 1**). The critical radius for heterogeneous nucleation is less than that for homogeneous nucleation, because the ΔG_S term in heterogeneous nucleation is partially offset by the free energy associated with local inhomogeneities.

Systems at equilibrium are said to be saturated in the minerals associated with that equilibrium reaction, but no nucleation will occur at saturation because ΔG_V is 0. As systems get further out of equilibrium and become increasingly supersaturated in a mineral, ΔG_V increases and the critical radius decreases. The nucleation rate is a measure of the number of nuclei per unit time that exceed the critical radius. Experimental studies show that, in general, nucleation rates increase with increasing supersaturation (**Fig. 2**). *See* FREE ENERGY; NUCLEATION.

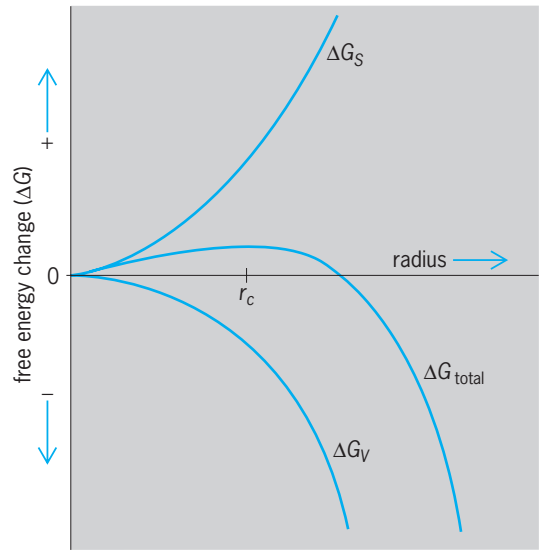

Fig. 1. For growth of spherical nuclei, surface free-energy change (ΔG_S) increases as a function of $4\pi r^2$ (the area of a sphere), while volume free-energy change (ΔG_V) decreases as a function of $^4/_3\pi r^3$ (the volume of a sphere). Nuclei that exceed the critical radius (r_c) will grow spontaneously; nuclei smaller than the critical radius will spontaneously disaggregate.

Crystal growth. The rate at which crystals grow, once stable nuclei have formed, is dependent on the rate at which chemical components attach to the surface of a crystal and the rate at which included components diffuse toward, and excluded components diffuse away from, a growing crystal. The crystal morphology that develops is dependent upon the

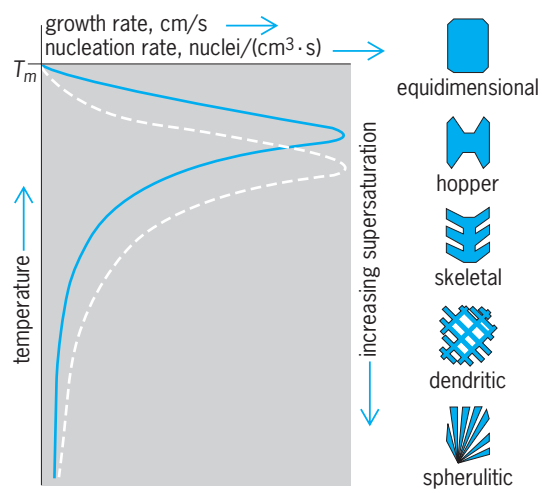

Fig. 2. Growth rate (solid curve) and nucleation rate (broken curve) as a function of temperature for a hypothetical silicate melt and resulting crystal morphology. T_m is the melting temperature of the crystal.

degree of supersaturation when nucleation first occurs. Equidimensional crystals result from interface-controlled growth in which the rate-limiting factor is the attachment of new components at the crystal interface. Cellular morphologies (hopper, skeletal, dendritic, or spherulitic crystal forms) result from diffusion-controlled growth in which the rate-controlling process is the diffusion of components to and from the crystal interface. Experimental studies indicate that cellular morphologies result when the growth rate (cm/s) exceeds the pertinent diffusion coefficients (cm^2/s). Studies also show that after an initial increase, the growth rate reaches a maximum and then decreases with increasing supersaturation. The fall-off in the nucleation and growth rates is the result of a decrease in molecular motion and diffusion rates at lower temperatures. Like nucleation rates, growth rates are equal to 0 at saturation and increase with increasing supersaturation (Fig. 2). In igneous rocks, the relative crystal morphologies are controlled by the rate and order in which minerals grow. In metamorphic rocks, the relative crystal morphologies are controlled by the surface free energies of the competing minerals; the best-developed crystal faces occur in minerals with the highest surface free energies. *See* CRYSTAL GROWTH.

Diffusion. The rate at which chemical components move through solids and liquids is known as the diffusion rate, and is an important control on the rates of crystal nucleation and growth. Diffusion coefficients are a measure of the rate at which components move across a plane perpendicular to the direction of diffusion, and they can be used to compare relative diffusion rates. Diffusion without a compositional gradient is known as intrinsic diffusion and results in no net flux of material; diffusion along a compositional gradient is known as chemical diffusion and results in a net flux of material from areas of high concentration to areas of low concentration. In general, diffusion rates in magmas increase with increasing temperature, increasing water content, and decreasing silica content. Diffusion rates are several orders of magnitude lower in solids than in melts, but can be greatly increased if a fluid phase is present along grain boundaries. *See* CHEMICAL EQUILIBRIUM; DIFFUSION; MAGMA; METAMORPHISM; MINERAL. H. R. Naslund

Redox Equilibria in High-Temperature Systems

The term "oxidation-reduction" (redox) refers to the gain and loss of electrons by atoms or ions. A redox equilibrium is one in which both processes occur.

Iron. Iron (Fe) is one of the most abundant of the elements that form terrestrial planets; it has the property of ready adoption of altervalent states (Fe^0, Fe^{2+}, Fe^{3+}), making redox equilibria particularly important in terms of geochemical processes.

For example, it is believed that all of the terrestrial planets possess a central metallic core, which on cosmochemical grounds most likely is predominantly metallic iron (Fe^0). The formation and gravitational segregation of Fe^0 in a differentiating planet has important consequences for the thermal history (including melting), volatile degassing and evolution, and initial redox state of the overlying silicate mantle of the planet. *See* COSMOCHEMISTRY.

Iron is also generally the most abundant altervalent element presently in the Earth's crust and mantle. Consequently, redox equilibria such as that shown in reaction (2), are important in terms of the min-

$$2FeO + {}^1/_2O_2 \rightleftharpoons Fe_2O_3 \qquad (2)$$

eral assemblages developed, the nature of the volatile species present in multicomponent geologic fluids through equilibria such as shown in reaction (3), and

$$8Fe_3O_4 + CO_2 + 2H_2O \rightleftharpoons 12Fe_2O_3 + CH_4 \qquad (3)$$

the character of silicate melts produced during genesis of magma.

Other redox controls. Within the range of redox states commonly found within the Earth, ranging from equilibria involving Fe^0 to air, there are other less abundant altervalent elements that are important in high-temperature silicate systems. These include the couples of nickel (Ni^0-Ni^{2+}), europium (Eu^{2+}-Eu^{3+}), chromium (Cr^{2+}-Cr^{3+}), titanium (Ti^{3+}-Ti^{4+}), and manganese (Mn^{3+}-Mn^{4+}). Furthermore, in the absence of metal redox equilibria involving silicates, carbon polymorphs and carbonates may be locally the most important redox controls defining the oxygen fugacity (f_{O_2}) in the Earth's upper

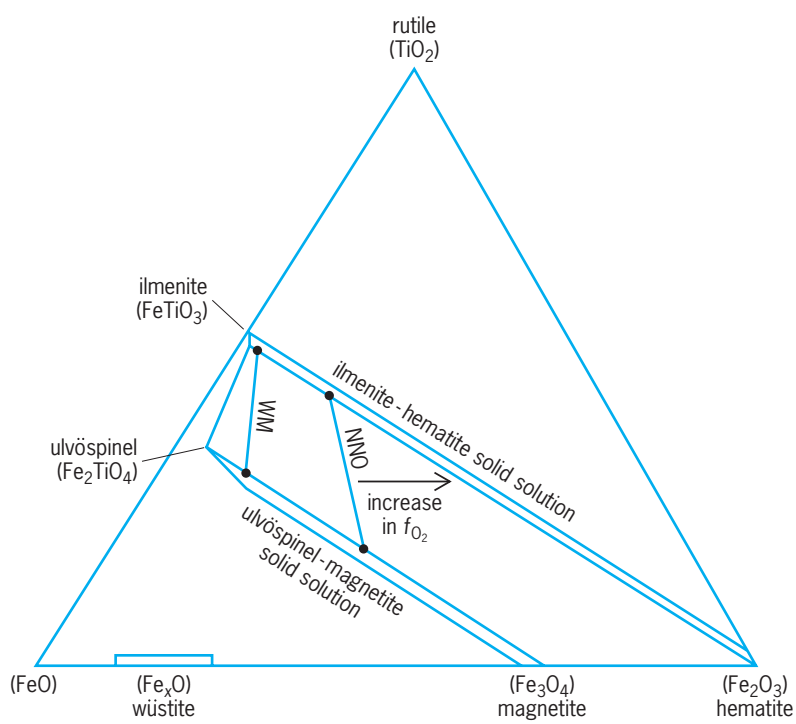

Fig. 3. A portion of the system Fe-Ti-O at 1000°C (1800°F) and 1 bar (10^5) pascals. The double lines between ilmenite-hematite and ulvöspinel-magnetite represent continuous solid solutions, and the width of these zones reflects the extent of Fe-Ti nonstoichiometry. Lines WM and NNO link coexisting phases of solid solutions of ilmenite-hematite and ulvöspinel that are stable under the conditions of the wustite-magnetite and nickel–nickel oxide oxygen buffers, respectively, at the temperature and pressure indicated.

mantle, as shown in reaction (4).

$$\underset{\text{Enstatite}}{MgSiO_3} + \underset{\text{Magnesite}}{MgCO_3} \rightleftharpoons \underset{\text{Forsterite}}{Mg_2SiO_4} + \underset{\text{Diamond or graphite}}{C} + O_2 \qquad (4)$$

Experimental measurement of oxygen fugacity. Determination of the oxygen fugacity prevailing in high-temperature silicate systems has been pursued with a variety of methods. These are based on thermodynamic calculations involving mineral assemblages at equilibrium, electrochemical cell equilibria, and comparison of the measured ratio Fe^{2+}/Fe^{3+} of magmas with laboratory calibrations of the temperature–oxygen fugacity–compositional dependence of this ratio.

Thermodynamic calculations. As an example of the thermodynamic approach, the coexistence of different phases that are sensitive to temperature and oxygen fugacity in the system Fe-Ti-O has proved particularly important (**Fig. 3**). Precise compositions of coexisting Fe-Ti oxides are monitors of the temperature-redox conditions under which a given rock formed, and these two "tie lines" are examples of the many coexisting compositions that can be found within the natural range of redox states.

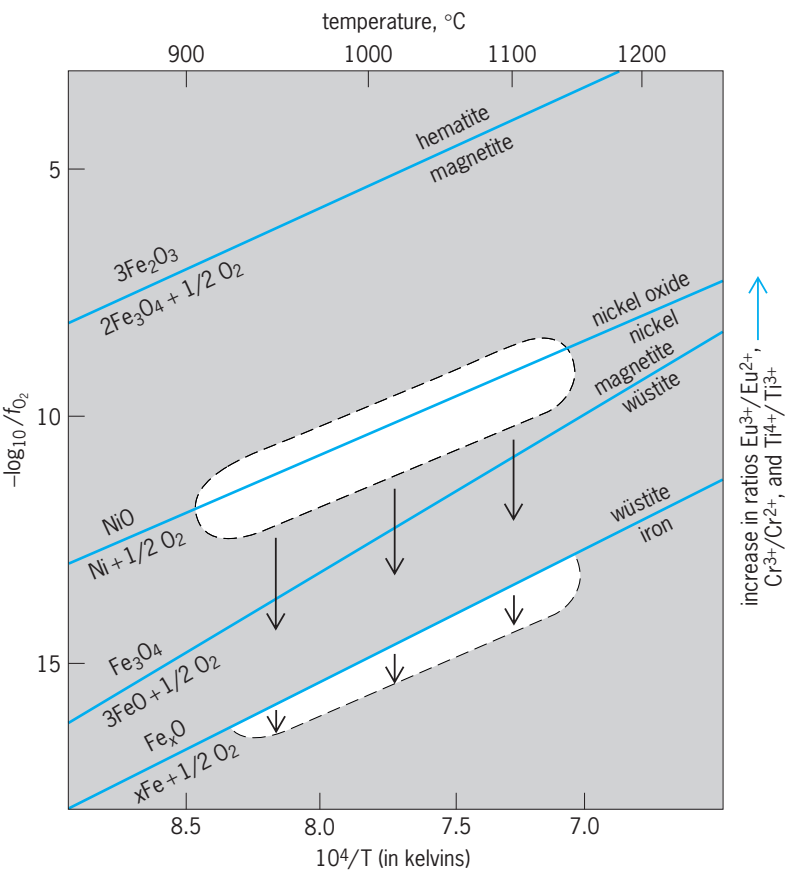

Fig. 4. Range of redox states present within most terrestrial igneous and metamorphic rocks represented in terms of oxygen fugacity (f_{O_2}) as a function of temperature. Most rocks have formed between the oxygen buffers represented by the iron-wustite and magnetite-hematite redox equilibria. In fact, the great majority of erupted rocks lies close to the nickel–nickel oxide redox equilibrium. The arrows indicate a gradation between the areas enclosed in broken lines from (top) most erupted igneous rocks to (bottom) some high-pressure terrestrial, lunar, and meteoritic systems. °F = (°C × 1.8) + 32°.

The precise compositions of coexisting solid solutions of ilmenite-hematite (i-h.ss) and ulvöspinel-magnetite (u-m.ss) found in a wide variety of igneous and metamorphic rocks are a function of the absolute temperature (T) and f_{O_2} of equilibration. The f_{O_2}–T dependence may be described by reaction (5), from

$$\underset{\text{in u-m.ss}}{4Fe_3O_4} + O_2 \rightleftharpoons \underset{\text{in i-h.ss}}{6Fe_2O_3} \qquad (5)$$

which Eq. (6) can be derived, where ΔG° = stan-

$$\ln f_{O_2} = \frac{\ln\left(a^{\text{i-h.ss}}_{\text{hematite }(Fe_2O_3)}\right)^{6-}}{\ln\left(a^{\text{u-m.ss}}_{\text{magnetite }(Fe_3O_4)}\right)^{4}} + \frac{\Delta G^\circ}{RT} \qquad (6)$$

dard Gibbs free energy of formation or reaction, T = absolute temperature, R = gas constant, and the notation a^{b}_{c} means the activity of component c in phase b. Thermodynamic data (enthalpies, entropies, and cell volumes) are available for the pure end-member components of these solid solution series for derivation of ΔG°. Adoption of appropriate solution models for the activity terms in Eq. (6) allows calculation of the variation of f_{O_2} as a function of T. *See* ACTIVITY (THERMODYNAMICS); SOLID SOLUTION.

Electrochemical cell equilibria. In the second experimental approach, an electrochemical cell is constructed with a high-temperature, O^{2-}-specific solid electrolyte [typically stabilized zirconium oxide (ZrO_2)]. One electrode is formed from the geological material of interest containing at least one redox pair (for example, Fe^{2+}-Fe^{3+}) in one or more phases. The other electrode is a reference material of known oxygen potential. At a given temperature, the electromotive force (emf; E) measured across the cell with a high-impedance voltmeter is a function of the difference in oxygen potential between the two electrodes, and is given by the Nernst equation (7), which is the ex-

$$E = \frac{RT}{4F}\ln\frac{(f_{O_2\text{sample}})}{(f_{O_2\text{reference}})} \qquad (7)$$

pression for the activity dependence of emf; F is the Faraday constant.

Knowing the variation of the oxygen fugacity of the reference material as a function of T and measuring E allows calculation of the oxygen fugacity of the sample. *See* ELECTROMOTIVE FORCE (EMF); ENTHALPY; ENTROPY.

Comparison of ratios with laboratory calibrations. The theoretical basis for the third approach to experimental measurements is founded in reaction (2), which can be developed as Eq. (8). Empirical calibrations

$$\ln (f_{O_2})^{1/2} = \frac{\Delta G^\circ}{RT} + \frac{\ln\left(a^{\text{magma}}_{Fe_2O_3}\right)}{\left(a^{\text{magma}}_{FeO}\right)^2} \qquad (8)$$

of this equilibrium with a wide range of natural and synthetic silicate melts have been achieved.

Interpretation. Application of the three experimental approaches either singly or in combination has shown that the majority of igneous rocks that have

erupted at the Earth's surface lie within the f_{O_2} range defined by the iron-wustite and magnetite-hematite equilibria (**Fig. 4**). At the more reduced end of this redox range, a readily measurable fraction of the Eu, Ti, and Cr in igneous and metamorphic rocks occurs in the lower of the valence states listed earlier.

In the early years of quantitative experimental studies of silicate systems, the control of oxygen fugacity, iron redox state, and the loss of iron from the systems under investigation through alloying with sample containers was difficult. With a few exceptions, the problems were mostly avoided by studies of iron-free, synthetic analogs. However, it was particularly through the efforts of E. F. Osborn and colleagues that methods of atmospheric control in studies of silicate systems were developed during the 1950s and onward. Initially, these studies were pursued with support from the steel-making industry because of the relevance of optimizing the performance of iron-bearing silicate slags. In addition, it became apparent that many of the processes involved in the heterogeneous phase equilibria of iron-bearing silicates, oxides, melts, and vapor were relevant to natural magmatic systems.

For example, it has long been recognized that within the subalkaline spectrum of igneous rock compositions, two major lineages are present that are characterized by systematic early increases of the FeO/MgO ratio during increases in concentration of SiO_2 (tholeiitic trend), and a relatively constant FeO/MgO value during equivalent increases in SiO_2 (calcalkalic trend), respectively. Although it is recognized that a number of alternative processes may be involved to account for the development of these lineages at different times and places, the timing of saturation and fractional crystallization of a spinel rich in the magnetite (Fe_3O_4) component can be critical. The stability of spinel in a given magma is a function of pressure, temperature, bulk composition, and oxygen fugacity.

Much experimental effort has been expended since 1970 in efforts to control the intensive parameters and retention of iron in synthetic and natural systems during experimental studies. Rewards for this effort include a thorough understanding of the partitioning of FeO and MgO between olivine and silicate melt, and the acquisition of detailed knowledge of the phase behavior of subsolidus and supersolidus relations in the upper mantle.

The nature of solid-melt-fluid speciation in the system C-O-H-S is also well understood under conditions relevant to the deep crust and upper mantle. The predominance of CH_4-H_2O in fluids close to iron-wustite buffer conditions and CO_2-H_2O in more oxidized settings has been determined, and this is clearly relevant to the nature of gas species that could have formed a primitive terrestrial atmosphere during the early evolutionary stages of the Earth.

Finally, the possibility of secular changes in the redox state of the Earth's upper mantle has been examined. It is clear that the return of relatively oxidized oceanic lithosphere to the interior of the planet is a prime agent of local oxidation of the mantle, and indeed analysis of magmas and fragments of the mantle associated with subduction zones indicates that these regions are relatively oxidized compared with other zones of magma generation, such as at mid-oceanic ridges.

Research is focused on the mechanisms of redox control within the Earth's mantle and possible causes for spatial and temporal variations. Comparative studies of the redox evolution of the Earth, Moon, and differentiated meteorites are also of interest with respect to core separation events, metal-silicate equilibria, volatile loss, and recycling. *See* FUGACITY; OXIDATION-REDUCTION. Richard Arculus

Volatile Equilibria at High Temperatures and Pressures

Volatiles are compounds that tend to form gases or fluids in high-temperature and relatively low-pressure geologic environments. They occur as a gas phase, as essential constituents of some minerals, and as dissolved constituents in magmas (molten rock) and fluids. Thermodynamic modeling and experimental research on volatiles determine the conditions of pressure, temperature, and composition at which volatile-bearing rocks form. By definition, volatile equilibria are equilibrium relationships established between volatile-bearing minerals, magmas, and fluids. Volatile equilibria control many high-temperature and high-pressure geologic processes in igneous and metamorphic mineral assemblages, including volcanic eruptions and ore deposition from fluids.

Volatile species. Analyses of volcanic gases, fluid inclusions (small gas or fluid-filled bubbles within rocks), and volatiles in rocks and minerals indicate that the dominant volatiles include water (H_2O) and compounds of carbon (C), sulfur (S), nitrogen (N), fluorine (F), and chlorine (Cl)—that is, carbon dioxide (CO_2), carbon monoxide (CO), methane (CH_4), sulfur dioxide (SO_2), hydrogen sulfide (H_2S), ammonia (NH_4), hydrogen fluoride (HF), and hydrogen chloride (HCl). The volatile species that occur in a given igneous or metamorphic environment vary as a function of pressure, temperature, and system composition. The stability of some of the important species is defined by the volatile equilibria shown in reactions (9)–(13). The dominant volatile species re-

$$C + 2H_2 \rightleftharpoons CH_4 \quad (9)$$

$$H_2 + {}^1/_2O_2 \rightleftharpoons H_2O \quad (10)$$

$$C + O_2 \rightleftharpoons CO_2 \quad (11)$$

$$O_2 + {}^1/_2S_2 \rightleftharpoons SO_2 \quad (12)$$

$$H_2 + {}^1/_2S_2 \rightleftharpoons H_2S \quad (13)$$

leased during eruptions of basaltic volcanoes include H_2O, CO_2, and SO_2, and the dominant species generated during thermal metamorphism include H_2O and CO_2 with or without CO, CH_4, and H_2S. The extent to which these volatiles alter the stability of minerals, magmas, and fluids is largely proportional to

Maximum concentrations of volatiles in silicate minerals and magmas*

	Components, wt %			
	Water	Carbon dioxide	Chlorine	Fluorine
Igneous minerals	7	4	2	21
Granitic magmas†	10	0.5	1	5
Metamorphic minerals	14	4	1.5	9

*Concentrations in minerals are for mica, amphibole, chlorite, serpentine, topaz, humite, and scapolite.
†Concentrations are for pressures of ≤150,000 lb/in.^{2} (1030 megapascals); the solubilities of H_2O and CO_2 are strongly pressure-dependent.

their concentration in the system (see **table**). *See* VOLCANO.

A fluid phase is not always stable in volatile-bearing systems. Each component in an assemblage exerts a pressure on the system, and the highest relative pressures are exerted by the more volatile components. A fluid phase is unstable if the sum of the individual partial pressures of the volatiles is less than the confining rock pressure. The relationship between partial pressure and total gas pressure is given by Eq. (14), where P_i is the partial pressure of

$$P_i = P_{\text{total}} X_i \tag{14}$$

species i, P_{total} the total fluid pressure in the system, and X_i the concentration of i expressed as the mole fraction. Alternatively, at pressures or temperatures lower than that of a liquid's critical point, a single fluid phase may be replaced by a gas coexisting with a liquid.

Influences. Water is the most significant volatile to influence geologic processes in many environments, primarily because of its greater abundance relative to most other volatiles. Water fluxes the melting of rocks. For example, the presence of only 5 wt % H_2O in granitic and basaltic magmas reduces the liquidus temperatures by 200–300°C (392–572°F). Furthermore, during melting of mica- and amphibole-bearing rocks, the water in these hydrous minerals increases the relative stability of silicate melt by reducing magmatic solidus temperatures [reaction (15)]. An increasing abundance of H_2O in silicate

$$\begin{aligned} SiO_2 + KMg_3(AlSi_3O_{10})(OH)_2 &\rightleftharpoons \\ KAlSi_3O_8 + MgSiO_3 &+ \text{hydrous melt} \end{aligned} \tag{15}$$

magmas also profoundly disrupts the internal structure of silicate magma, reduces magmatic viscosities, enhances rates of diffusion of species though magma, increases or decreases crystallization rates, and increases the explosive nature of volcanic eruptions.

Increasing the extent of thermal metamorphism decreases the stability of hydrous mineral assemblages and increases the relative stability of H_2O-bearing fluids, such as in reaction (16). As the

$$\begin{aligned} KAl_2(AlSi_3O_{10})(OH)_2 + SiO_2 &\rightleftharpoons \\ KAlSi_3O_8 + Al_2SiO_5 &+ H_2O \text{ fluid} \end{aligned} \tag{16}$$

hydrous assemblage, muscovite plus quartz, is subjected to increasing thermal metamorphism, it is increasingly unstable and gradually dehydrates to an anhydrous assemblage of potassium feldspar and, depending on temperature and pressure, one of the aluminosilicate minerals (kyanite/sillimanite/andalusite) and a water-rich fluid.

An increasing abundance of carbon dioxide in a fluid decreases the stability of hydrous mineral assemblages and variably alters the melting temperatures of hydrous and anhydrous magmas. Because of its relatively low solubility in most magmas, carbon dioxide has a fairly small effect on magmatic solidus and liquidus temperatures, in contrast to water. The solubility of carbon dioxide in silicate magmas has been found experimentally to vary strongly with changes in the composition of the magma. Devolatilization resulting from increasing thermal metamorphism generates H_2O- and CO_2-bearing fluids via dehydration and decarbonation reactions, respectively. For example, fluids rich in CO_2 are generated by equilibria such as that in reaction (17).

$$\underset{\text{Calcium carbonate}}{CaCO_3} + SiO_2 \rightleftharpoons \underset{\text{Calcium silicate}}{CaSiO_3} + CO_2 \text{ fluid} \tag{17}$$

The release of sulfur dioxide (SO_2) from magma to the atmosphere during explosive volcanic eruptions can produce short-term climate variations. Sulfur dioxide reacts with water vapor in air to form small droplets of sulfuric acid (H_2SO_4), and these aerosol droplets have been shown to reflect sunlight away from Earth and hence to reduce the extent of solar heating of the Earth's surface for short periods.

The volatile elements fluorine, chlorine, and sulfur and their compounds are important in the transport of ore and non-ore minerals in CO_2- and H_2O-bearing fluids. Under conditions of high temperature and pressure, they enhance the solubilities of minerals in such fluids. Fluorine is similar to water in that it shows relatively large solubility in silicate magmas and, as a result, high concentrations of fluorine in silicate magmas can significantly reduce solidus and liquidus temperatures of magmas, enhance diffusivities, and disrupt magma structures. Sulfur compounds are also important for the deposition of many ore metals in the form of sulfide minerals. *See* ORE AND MINERAL DEPOSITS.

Experimental studies. The relative stability of a given silicate assemblage can be predicted through thermodynamically based computations, but specific values of pressure, temperature, and composition at which an assemblage is stable are best fixed by experimental research. In experimental simulations, volatiles and natural minerals or synthetic analogs of

minerals are heated and pressurized in sealed capsules of precious metals. Depending upon the kinetics of the reactions, the establishment of equilibrium requires durations of experimental runs ranging from minutes to months. Examination and chemical analysis of the materials at the end of an experiment determine the identity and composition of the minerals and the melt or fluid that were stable during the experiment.

In mineral-, melt-, and fluid-bearing systems at equilibrium, the most stable assemblage is the one that exhibits the lowest total free energy (G_{total}). In reaction (16), for example, muscovite and quartz are the stable phases at those temperatures where G_{total} of muscovite and quartz is lower than that of the fluid and anhydrous mineral assemblage.

If natural mineral or melt comparisons are not homogeneous or if a mineral assemblage contains minerals that are not mutually compatible, then chemical equilibrium has not been established within the assemblage and it is metastable. *See* HIGH-PRESSURE MINERAL SYNTHESIS; IGNEOUS ROCKS; METAMORPHIC ROCKS; SILICATE PHASE EQUILIBRIA.

James D. Webster

Bibliography. M. Carroll and J. R. Holloway, *Volatiles in Magmas*, Mineralogical Society of America, Reviews in Mineralogy Series, vol. 30, 1994; J. Crank, *The Mathematics of Diffusion*, 1975; J. M. Ferry and D. M. Burt, *Characterization of Metamorphism through Mineral Equilibria*, Mineral Society of America, Reviews in Mineralogy Series, vol. 10, pp. 207–262, 1982; P. Henderson, *Inorganic Geochemistry*, 1982; J. R. Holloway, *Advances in Physical Chemistry 1*, 1981; R. J. Kirkpatrick, Crystal growth from the melt: A review, *Amer. Mineralog.*, 60:798–814, 1975; R. J. Kirkpatrick, Theory of nucleation in silicate melts, *Amer. Mineralog.*, 68:66–77, 1983; G. E. Lofgren, An experimental study of plagioclase crystal morphology: Isothermal crystallization, *Amer. J. Sci.*, 274:243–273, 1974; D. K. Nordstrom and J. L. Munoz, *Geochemical Thermodynamics*, 1987; R. C. Reid, J. M. Prausnitz, and B. E. Roling, *The Properties of Gases and Liquids*, 4th ed., 1987; H. Sigurdsson, *Encyclopedia of Volcanoes*, pp. 149–170, 1999; G. C. Ulmer, *Research Techniques for High Pressure and High Temperature*, 1971.

Siliceous sediment

Fine-grained sediment and sedimentary rock dominantly composed of the microscopic remains of the unicellular, silica-secreting plankton diatoms and radiolarians. Minor constituents include extremely small shards of sponge spicules and other microorganisms such as silicoflagellates. Siliceous sedimentary rock sequences are often highly porous and can form excellent petroleum source and reservoir rocks. *See* SEDIMENTARY ROCKS.

In the modern oceans, the distribution of siliceous sediments closely resembles the distribution of nutrients in the uppermost oceanwaters, because diatoms and radiolarians are most abundant in oceanic regions of high productivity. Some of the greatest concentrations of siliceous sediment are adjacent to Antarctica; in the equatorial Pacific Ocean; and along the continental shelves of Peru, the western United States, and west Africa—all of which are oceanographic regions with elevated biologic productivity. The continued productivity results in siliceous sediment being deposited on the sea floor in these regions. Sometimes large-scale die-offs of siliceous planktonic communities can result in the deposition of mats of intertwined diatom assemblages on the sea floor. *See* SILICA MINERALS.

Siliceous sediment undergoes a complicated chemical and physical process during burial on the sea floor (see **illus.**). Diatoms and radiolarians secrete a form of silica that is an amorphous mineral called opal-A. Because seawater is undersaturated with respect to this silica mineral, most of the opal-A microskeletons dissolve on or within the uppermost few centimeters of the sea floor. With continued burial to a few hundreds of meters, the small proportion of opal-A that survives initial burial continues to dissolve, and eventually (with continued sedimentation burying the mud to higher temperatures and pressures) a new siliceous mineral is precipitated, called opal-CT. This opaline mineral consists of disordered cristobalite and tridymite, structures which are also made of silica. With still more burial to many hundreds of meters below the sea floor, the opal-CT in turn becomes chemically unstable, and a new mineral called biogenic quartz is formed. Biogenic quartz is the dominant mineral in the siliceous rock chert (also known as flint or, more rarely, jasper). This mineral is chemically stable, and although it is chemically similar to the quartz found on sandy beaches, its process of formation is entirely different. Biogenic quartz owes its existence to the diatoms and radiolarians that initially thrived in the water column above the sea floor, whereas the quartz found on typical beaches results from the erosion of igneous and metamorphic rocks. Chert and flint are hard, dense rocks, and their fine grain size results in maintenance of sharp edges, a characteristic capitalized upon by early humans who used chert for arrowheads, other weapons, and jewelry. *See* OPAL; QUARTZ.

Given their biologic composition, siliceous sediments provide some of the best geologic records of the ancient oceans. Diatoms did not evolve until the late Mesozoic; thus the majority of siliceous rocks older than approximately 150 million years are formed by radiolarians. Geologists map the distribution of ancient siliceous sediments now pushed up onto land by plate tectonic processes, and can thus determine which portions of the ancient seas were biologically productive; this knowledge in turn can give great insight into regions of the Earth's crust that may be economically productive (for example oil-containing regions). The vast oil reserves of coastal California are predominantly found in the Monterey Formation, a highly porous diatomaceous

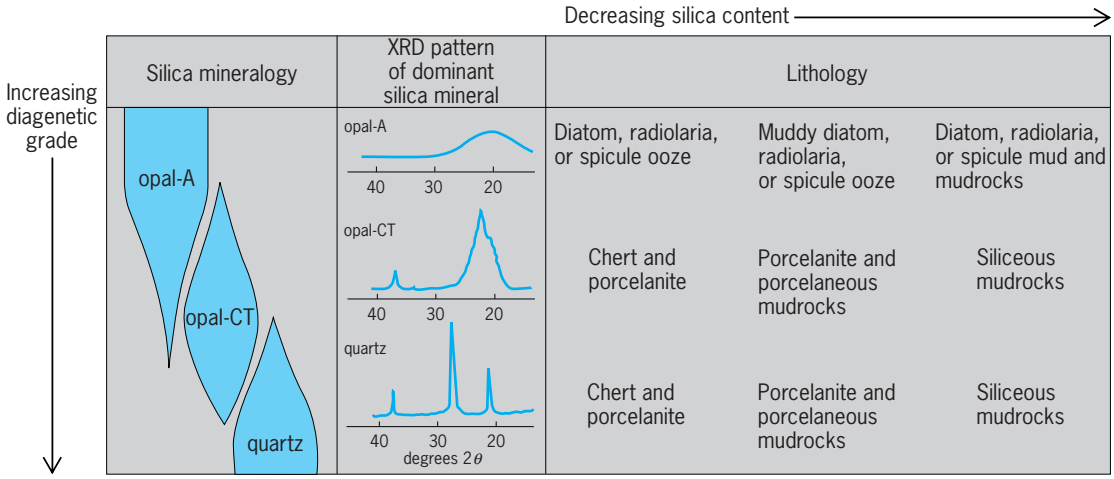

Generalized relationship between silica mineralogy and rock types. (*After K. A. Pisciotto, Distribution, thermal histories, isotopic compositions, and reflection characteristics of siliceous rocks recovered by the Deep Sea Drilling Project, in The Deep Sea Drilling Project: A Decade of Progress, vol. 32, ed. by J. E. Warme, R. G. Douglas, and E. L. Winterer, Society of Economic Paleontologists and Mineralogists, Tulsa, 1981*)

siliceous sedimentary sequence distributed along the western seaboard of the United States. These sediments and rocks are particularly well exposed along the beaches of Santa Barbara, where some of the outcrops are so petroliferous that they can be lit with a common household match and can burn slowly for several hours. *See* CHALK; CHERT; LIMESTONE.

R. W. Murray

Bibliography. G. Bohrmann et al., Pure siliceous ooze a diagenetic environment for early chert formation, *Geology*, 22:207–210, 1994; J. R. Hein, *Siliceous Sedimentary Rock-Hosted Ores and Petroleum*, Van Nostrand Reinhold, New York, 1987; A. E. S. Kemp and J. G. Baldauf, Vast Neogene laminated diatom mat deposits from the eastern equatorial Pacific Ocean, *Nature*, 362:141–143, 1993; M. Leinen et al., Distribution of biogenic silica and quartz in recent deep-sea sediments, *Geology*, 14:199–203, 1986; R. W. Murray, Chemical criteria to identify the depositional environment of chert: General principles and applications, *Sed. Geol.*, 90:213–232, 1994; K. A. Pisciotto, Distribution, thermal histories, isotopic compositions, and reflection characteristics of siliceous rocks recovered by the Deep Sea Drilling Project, in *The Deep Sea Drilling Project: A Decade of Progress*, vol. 32, 1981.

Siliceous sinter

A porous silica deposit formed around hot springs. It is white to light gray and sometimes friable. Geyserite is a variety of siliceous sinter formed around geysers. The siliceous sinters are deposited as the hot subterranean waters cool after issuing at the surface and become supersaturated with silica that was picked up at depth. The sinters are frequently deposited on algae that live in the pools around the hot springs. *See* GEYSER; SPRING (HYDROLOGY).

Raymond Siever

Silicoflagellata

A class of marine planktonic chromists. Their skeletons are composed of hollow, siliceous rods and resemble those of the Radiolaria, with which they have been grouped. They are usually subpyramidal or hemispherical in shape, delicately filigreed, and range in size from 10 to 150 micrometers (see **illus.**).

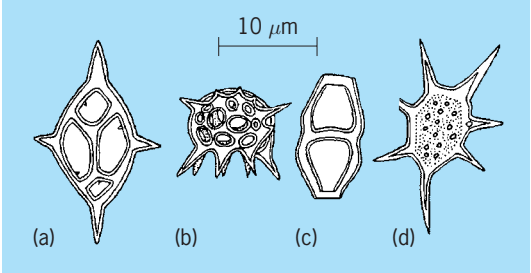

Examples of fossil and modern Silicoflagellata. (*a*) *Dictyocha*, Cretaceous to Recent; (*b*) *Cannopilus*, Miocene; (*c*) *Naviculopsis*, Eocene to Miocene; and (*d*) *Vallacerta*, Upper Cretaceous.

Two families and 11 genera of silicoflagellates have been described from siliceous sedimentary rocks ranging in age from the early Cretaceous to Recent, in association with abundant diatoms and siliceous sponge spicules; however, this may change pending systematic revision. One genus, *Dictyocha*, lives in the ocean today. *See* CHRYSOPHYCEAE; MICROPALEONTOLOGY; PHYTOPLANKTON.

Daniel J. Jones; Sally Walker

Silicoflagellida

An order of the phylum Protozoa, subphylum Sarcomastigophora, superclass Mastigophora, class Phytamastigophorea. These organisms are marine flagellates which have an internal, siliceous, tubular skeleton; numerous small, discoid, yellow

chromatophores; and a single flagellum. At times the organisms also put forth, from the ends of their skeletal tubes, long, rather threadlike pseudopodia. The skeleton forms a basket within which the moiety of protoplasm lies, but skeletal elements always have at least a thin covering. Nuclei are not conspicuous. Reserve material consists of leucosin.

There is now a single genus, with *Dictyocha fibula* being the most abundant species, although three other species have been described. The organism is not uncommon in Atlantic, Pacific, and Gulf of Mexico waters. However, it is normally not abundant enough to play an important role in ecology.

In past ages a large number of species were abundant, judging from fossil remains. The skeletons vary from quadrate through pentagonal, hexagonal, and octagonal types. *See* PHYTAMASTIGOPHOREA; PROTOZOA.

James B. Lackey

Silicon

A chemical element, Si, atomic number 14, and atomic weight 28.086. Silicon is the most abundant electropositive element in the Earth's crust. The element is a metalloid with a decided metallic luster; it is quite brittle. It has a specific gravity of 2.42 at 20°C (68°F), melts at 1420°C (2588°F), and boils at 3280°C (5936°F). The element is usually tetravalent in its compounds, although sometimes divalent, and is decidedly electropositive in its chemical behavior. In addition, pentacoordinate and hexacoordinate compounds of silicon are known. *See* METALLOID; PERIODIC TABLE.

Crude elementary silicon and its intermetallic compounds are used in alloying constituents to strengthen aluminum, magnesium, copper, and other metals. Metallurgical silicon of 98–99% purity is used as the starting material for manufacturing organosilicon compounds and silicone resins, elastomers, and oils. Silicon chips are used in integrated circuits. Photovoltaic cells for direct conversion of solar energy to electricity use wafers sliced from single crystals of electronic-grade silicon. Silicon dioxide is used as the raw material for making elementary silicon and for silicon carbide. Sizable crystals of it are used for piezoelectric crystals. Fused quartz sand becomes silica glass, used in chemical laboratories and plants as well as an electrical insulator. A colloidal dispersion of silica in water is used as a coating agent and as an ingredient in certain polishes.

1																	18
1												13	14	15	16	17	2
H	2																He
3	4											5	6	7	8	9	10
Li	Be											B	C	N	O	F	Ne
11	12											13	14	15	16	17	18
Na	Mg	3	4	5	6	7	8	9	10	11	12	Al	Si	P	S	Cl	Ar
19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
55	56	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86
Cs	Ba	Lu	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
87	88	103	104	105	106	107	108	109	110	111	112	113					
Fr	Ra	Lr	Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg							

lanthanide series	57	58	59	60	61	62	63	64	65	66	67	68	69	70
	La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb

actinide series	89	90	91	92	93	94	95	96	97	98	99	100	101	102
	Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No

Naturally occurring silicon contains 92.2% of the isotope of mass number 28, 4.7% of silicon-29, and 3.1% of silicon-30. In addition to these stable, natural isotopes, several artificially radioactive isotopes are known. Elementary silicon has the physical properties of a metalloid, resembling germanium below it in group 14 of the periodic table. In very pure form silicon is an intrinsic semiconductor, although the extent of its semiconduction is greatly increased by the introduction of minute amounts of impurities. Silicon resembles the metals in its chemical behavior. It is about as electropositive as tin, and decidedly more positive than germanium or lead. In keeping with this rather metallic character, silicon forms tetrapositive ions and a variety of covalent compounds; it appears as a negative ion in only a few silicides and as a positive constituent of oxy acid or complex anions.

Several series of hydrides are formed, a variety of halides (some of which contain silicon-to-silicon bonds), and also many series of oxygen-containing compounds which may be either ironic or covalent in their properties.

Silicon occurs in many forms of the dioxide and as almost numberless variations of the natural silicates. For a discussion of the structures and compositions of the representative classes, *see* SILICATE MINERALS.

In abundance, silicon exceeds by far every other element except oxygen. It constitutes 27.72% of the solid crust of the Earth, whereas oxygen constitutes 46.6%, and the next element after silicon, aluminum, accounts for 8.13%.

Silicon is reported to form compounds with 64 of the 96 stable elements, and it probably forms silicides with 18 other elements. Besides the metal silicides, used in large quantities in metallurgy, silicon forms useful and important compounds with hydrogen, carbon, the halogen elements, nitrogen, oxygen, and sulfur. In addition, useful organosilicon derivatives and silicon-containing polymers have been prepared.

Edwin P. Plueddemann

Bibliography. M. A. Brook, *Silicon in Organic, Organometallic and Polymer Chemistry*, Wiley-Interscience, 1999; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; P. Jutzi and U. Schubert (eds.), *Silicon Chemistry: From the Atom to Extended Systems*, John Wiley & Sons, 2003; U. Schubert (ed.), *Silicon Chemistry*, Springer, 1999.

Silicone resins

Soluble or meltable materials based on highly branched polymers composed of alternating silicon and oxygen atoms and having organic (R) substituents attached to the silicon atoms. On curing, these polymers are converted into highly cross-linked, insoluble materials. They may contain four types of units, usually symbolized by the letters M, D, T, Q, as follows:

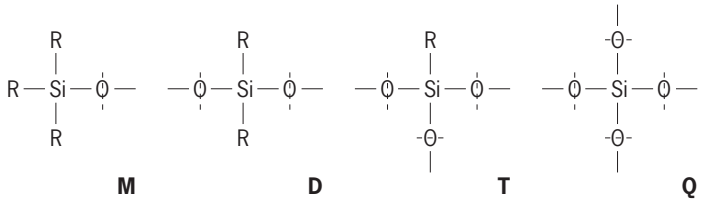

Each of the oxygen atoms is shared by neighboring units or by the unit and a functional OR group (R = H or alkyl). Therefore, the formulas of units are M—$R_3SiO_{1/2}$, D—R_2SiO, T—$RSiO_{3/2}$, Q—SiO_2.

Silicone resins are combinations of these units, the proportions of which are varied over a broad range. They are characterized by the R/Si ratio, which usually is in the range 1–1.7. Silicone resins constitute one of the three main classes of silicone products. The other two are silicone rubbers (R/Si = 2) and silicone fluids (R/Si > 2). The structure and properties of silicone resins are strongly dependent on the R/Si value. The lower the R/Si value, the less miscible the resin with organic polymers, the easier it undergoes gelation, and the harder it is after curing. Conversely, resins with higher R/Si values (containing more organic groups) are more flexible and compatible with organic resins. The basic component of silicone resins is usually the T unit. Polymers composed exclusively of T units, having the $(RSiO_{3/2})_n$ structure, are called polysilsesquioxanes and may contain fragments of a ladder or a cage structure. T resins having R/Si = 1 are known, but more common are DT resins. The D units give the resin softness and flexibility, which is particularly important in paints and varnishes. M and Q units also are often used. Resins composed exclusively of M and Q units, called MQ resins, are very important and have recently found numerous applications, such as pressure-sensitive adhesives and flexible silicone coatings, as well as high-performance defoamers, release agents, and additives for improving the mechanical properties of silicone rubber. *See* INORGANIC AND ORGANOMETALLIC POLYMERS; ORGANOSILICON COMPOUND; POLYMER; RUBBER.

The organic group is most often methyl or phenyl. Commonly used phenylmethyl silicone resins contain both phenyl and methyl in various proportions. Phenyl groups give the material superior thermal resistance. Cured resins containing a large proportion of phenyl groups are less brittle and more readily miscible with organic polymers. Other organic groups are introduced to give silicone resins special properties. For example, enhanced water repellency is attained by introducing fluoroalkyl groups or long-chain hydrocarbon groups, such as dodecyl ($C_{12}H_{25}$). Silicone resins contain a considerable number of hydroxy (HO—) or alkoxy (for example, CH_3O—) groups which, in the curing process, cross-link the polymer by condensation. These reactive groups may also form hybrids with organic resins. Vinyl groups or SiH groups are sometimes introduced in silicone resins to give them the ability to cure or form hybrids by the addition to olefinic bonds in organic polymers.

The presence of reactive silanol (SiOH) groups in the resins may make them unstable, particularly if the R/Si ratio is low and the molecular mass is relatively high. Their stability may be increased by the partial replacement of hydroxy groups by alkoxy groups. The stability of T resins (R/Si = 1) used for masonry is attained by conversion of the ≡SiOH to ≡SiONa or ≡SiOK groups. [Here ≡ means a triply bonded Si atom; that is, in ≡SiOH only SiOH group reacts and the other three (undefined) do not.]

Silicone resins are usually sold as solutions of the highly branched polymer with a molecular mass of 1000–60,000 daltons in an inert solvent such as an aromatic hydrocarbon, chlorinated hydrocarbon, or alkyl ester. Commercial solvent-free resins are either viscous liquids or powders which melt at 50–80°C (120–180°F). Some silicone resins are available as emulsions or dispersions in water.

Preparation. Of the basic monomers used for the synthesis of silicone resins, CH_3SiCl_3 and $C_6H_5SiCl_3$ are precursors of T units. For the generation of D units, $(CH_3)_2SiCl_2$, $C_6H_5MeSiCl_2$, and $(C_6H_5)_2SiCl_2$ are used. The precursor of the M unit is usually $(CH_3)_3SiCl$. The precursors of Q units are esters of orthosilicic acid $(C_2H_5O)_4Si$ (TEOS) or $(CH_3O)_4Si$ (TMOS), their polymers (for example, ethyl polysilicate 40 which contains 40% SiO_2), inorganic silicates, and $SiCl_4$. Vinyl groups may be introduced using $H_2C{=}CHSiCl_3$ or $H_2C{=}CHSi(CH_3)Cl_2$, while $(CH_3)_2SiHCl$ or CH_3SiHCl_2 is used to introduce SiH groups.

Methyl-substituted monomers CH_3SiCl_3, $(CH_3)_2SiCl_2$, $(CH_3)_3SiCl$ are synthesized using exclusively the Rochow-Müller direct synthesis from elemental silicon and methyl chloride (CH_3Cl) at 250–350°C (480–660°F), catalyzed with copper. $(CH_3)_2SiHCl$ and CH_3SiHCl_2 may be obtained by the same process by supplying HCl to the reactor. Phenyl-substituted monomers are manufactured by direct synthesis from Si and C_6H_5Cl at 500°C (930°F), catalyzed with Ag, or by the Grignard reaction from C_6H_5Cl and $SiCl_4$. *See* GRIGNARD REACTION.

Formerly, TEOS and TMOS were obtained by alcoholysis of $SiCl_4$ produced by the chlorination of silicon. However, direct synthesis from silicon and alcohol has become important. Inorganic silicates are products of reaction of sodium or potassium hydroxide with silica.

Synthesis of silicone resins is based on the hydrolytic polycondensation of a mixture of the organochlorosilane monomers. Reactions (1)–(5)

$$\equiv SiCl + H_2O \rightleftharpoons \equiv SiOH + HCl \quad (1)$$

$$\equiv SiCl + ROH \rightleftharpoons \equiv SiOR + HCl \quad (2)$$

$$\equiv SiCl + HOSi\equiv \rightleftharpoons \equiv SiOSi\equiv + HCl \quad (3)$$

$$\equiv SiOH + HOSi\equiv \rightleftharpoons \equiv SiOSi\equiv + H_2O \quad (4)$$

$$\equiv SiOR + HOSi\equiv \rightleftharpoons \equiv SiOSi\equiv + ROH \quad (5)$$

occur in the polycondensation system; however, part of the ≡SiOH and ≡SiOR groups must remain unreacted.

This process may be performed in various ways. Direct hydrolysis-condensation by the addition of the mixture of monomers to an excess of water is difficult to control as considerable differences in hydrolysis rates and in condensation rates of various components lead to a nonuniform product. Better control is gained when water is added slowly to the mixture of monomers and HCl is removed. A mixture of water with alcohol moderates the rate of the process and introduces alkoxy groups to the resin. The presence of alkoxy in place of ≡SiOH makes the resin more stable, and the cured material with remaining alkoxy is more flexible. Using a solvent for the monomers and resins that is immiscible with water slows down the process, diminishes the possibility of gelation, and decreases the polymer viscosity. The process must be stopped well before the gelation point when the desired viscosity is achieved. The process is terminated by cooling, washing, neutralization, and in some cases the addition of solvent.

MQ resins are obtained by two routes. The older process uses sodium silicate and trimethylchlorosilane as precursors. The more recent, easier to control route exploits the hydrolytic condensation reaction between hexamethyldisiloxane [$(CH_3)_3SiOSi(CH_3)_3$] and TEOS, TMOS, or ethyl silicate 40.

Some properties of silicone resins may be improved and the price of material may be lower if a hybrid of a silicone resin and an organic resin, such as an alkyd, polyester, epoxy, or acrylic resin, is used. These hybrids are usually obtained by condensation of silanol or alkoxy groups on a silicone resin with hydroxyl groups on an organic resin. If poor resin compatibility is a problem, silicone intermediates (reactive additives), such as (**1**) and (**2**), can be used to increase compatibility.

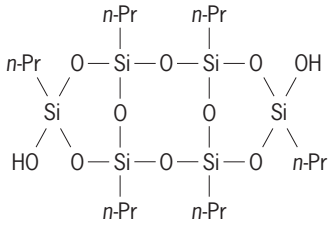

Dow Corning Z-6018

(**1**)

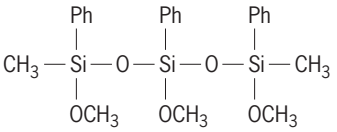

Dow Corning Z-6188

(**2**)

Curing. Curing of silicone resins is the process in which the three-dimensional network is formed. It is most often accomplished by the condensation reaction of the silanol and alkoxysilane groups remaining in the resin, according to Eqs. (**4**) and (**5**). The process is run at high temperature using catalysts such as acid and bases, inorganic salts, titanium, and organotin reagents. Curing of silicone coatings by air oxidation involves the cleavage of methyl groups with siloxane-bridge (≡SiOSi≡) formation. This process takes place at temperatures above 150°C (300°F). Resins containing vinyl groups may be cured by free-radical coupling, initiated by peroxides. Curing of resins containing both the SiH and $SiCH{=}CH_2$ groups occurs under mild conditions but needs to be catalyzed by metal complexes. The advantage of this method is that no volatile by-products are formed. *See* FREE RADICAL; ORGANIC SYNTHESIS; REACTIVE INTERMEDIATES.

Properties and applications. Silicone resins have outstanding thermal stability and weathering resistance. They change little on exposure to humidity, oxygen, heat, and sunlight. They are also physiologically inert, retain their gloss for a long time, and have good electrical resistance. Their mechanical and electrical properties are not very dependent on temperature. Thus, although some organic resins may have a higher hardness and strength at room temperature, silicone resins are superior to other resins at extreme temperature conditions. Silicone resins have very interesting surface properties, including very low surface energy. They are water-repellent and foreign bodies do not stick to them. They also show high permeability for gases.

Silicone resins are particularly useful in applications where the material is exposed to extreme temperatures and adverse weather conditions, with many types of silicone resins produced for specific applications.

A variety of silicone resins are used for paints. Heat-resistant paints and varnishes are commonly used on parts for furnaces and ovens, chimneys, pipelines for transporting heating media, exhaust pipes, engine and turbine parts, as well as machinery and apparatus operating at high temperatures. These paints contain a considerable amount of phenyl groups and are often used together with aluminum pigments. They withstand temperatures up to 450°C (840°F) and may work for a long time at 250°C (480°F).

Silicone paints based on methyl-substituted polymers are used on the exterior walls of buildings. They resist weathering, preserve gloss, and repel water, dust, and dirt. These water-repellent silicone-resin coatings make the surfaces hydrophobic, protecting them against humidity. They are most often used for protecting masonry. There are two general types of these materials: water-soluble sodium or potassium polymethylsiliconate resin, and solutions of methyl-substituted silicone resin in an organic solvent, most often a hydrocarbon. These clear, colorless coatings preserve the original appearance of the masonry and leave the masonry pores open, which permits the wall to "breathe."

Among the silicone impregnating varnishes and resins, particularly important are those used for the protecting electrical machinery and devices. These silicone resins provide high thermal stability and excellent electrical resistance and are used for

coating motor stator coils, transformer windings, and generator coils.

Silicone resins also are used as release coatings to prevent other bodies from sticking to surfaces. The most important are glazes for baking pans, trays, and molds. The surface properties of the resins are most critical in these applications, and high-modulus (stiff) methylsilicone resins of a low R/Si ratio are used. In contrast, paper coatings require a resin of lower modulus (more flexible). Silicone resins for pressure-sensitive adhesives are based on MQ resins with an average molecular mass of 1000–10,000 daltons. Silicone DT laminating resins for composites that are produced for the electrical industry contain a large number of silanol groups. Silicones for molding compounds are often solvent-free DT resins.

The range of use of silicone resins is expanding, with many new applications such as antifouling coatings, sensor coatings, resins for microelectronics, materials for photopattering, adhesives for porous ceramic materials, abrasion-resistant coatings, materials for optical communication, and biomedical materials. Particularly important are various silicone-organic hybrid resin materials.

Julian Chojnowski

Bibliography. B. Arkles, *Organosilicon Chemistry IV*, ed. by N. Auner and J. Weis, pp. 606–612, Wiley, Weinheim, 2000; M. A. Brook, *Silicon in Organic, Organometallic and Polymer Chemistry*, Wiley, 2000; B. Hardman and A. Torkelson, *Encyclopedia of Polymer Science and Engineering*, ed. by H. F. Mark et al., Wiley-Interscience, vol. 15, pp. 265-270, 1989; W. Huang, Y. Huang, and Y. Yu, Synthesis of MQ silicone resins through hydrolytic condensation of ethyl polysilicate and hexamethyldisiloxane, *J. Appl. Polym. Sci.*, 70(9):1753–1757, 1998; S. Paul, *Surface Coatings Science and Technology*, pp. 200–216, Wiley, Chichester, 1985; A. Tomanek, *Silicones and Industry: A Compendium for Practical Use, Instruction and Reference*, pp. 71–75, Wacker Chemie, Munich, 1991.

Silk

The lustrous fiber produced by the larvae of silkworms; also the thread or cloth made from such fiber. The United States is the greatest importer and consumer of silk. Silk will probably always be prized by the consumer even though certain manufactured fabrics now have some qualities that were formerly possessed only by silk.

Japan, the first country to use scientific methods in cultivating the silkworm, has always ranked highest in the production of fine silk, although satisfactory types are made in other silk-producing countries. The domestication and cultivation of the silkworm, which began in Japan about 3000 B.C., requires extreme care and close supervision, and the reeling of the filament from the cocoons can be done successfully only by skilled operators whose training is the result of generations of experience.

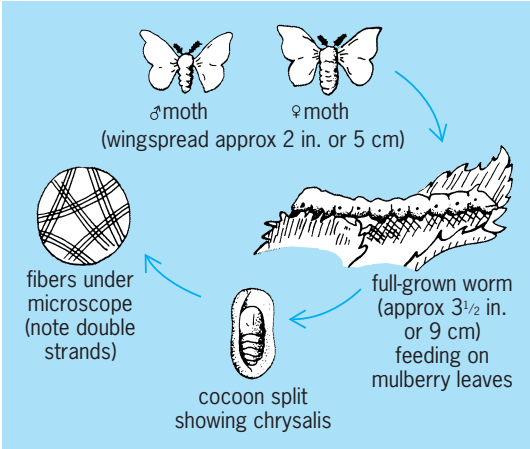

Life cycle of the silkworm.

Life cycle of the silkworm. Since the discovery that the filament composing the cocoon of the silkworm can be unwound and constructed into a beautiful and durable fabric, silkworms have been bred for the sole purpose of producing raw silk. The production of cocoons for their filament is called sericulture. The cocoon of the silkworm is the second stage of development of the life cycle of *Bombyx mori*, a species which spins a thread of high quality. In sericulture, all four stages of the life cycle of this moth (see **illus.**) are important, because some of the better cocoons must be set aside to permit full development, thus supplying eggs for another hatching. By scientific breeding, silkworms may be hatched three times a year; under natural conditions, breeding occurs only once a year. The life cycle includes (1) the egg, which develops into the larva or caterpillar, the silkworm; (2) the silkworm, which spins its cocoon for protection and to permit development into the pupa or chrysalis; (3) the chrysalis, which emerges from the cocoon as the moth; and (4) the moth, the female of which lays eggs thus continuing the life cycle.

Within 3 days after emerging from the cocoons, the moths mate and the female lays 350–400 eggs. The adults seldom fly, do not eat, and live only a few days. Each healthy egg hatches into a grub, or larva, about $^1/_8$ in. (3 mm) long. The larva requires careful nurturing for approximately 20–32 days. During this period, the tiny worm has a voracious appetite, requiring five daily feedings of chopped mulberry leaves. After four changes of skin, or moltings, the worm reaches full growth in the form of a smooth, grayish-white caterpillar about $3^1/_2$ in. (9 cm) long. After about 6 weeks its interest in food ceases. It shrinks somewhat in size and acquires a pinkish hue, becoming nearly transparent. A constant restless rearing movement of the head indicates that the worm is ready to spin its cocoon. Clusters of twigs or straw are provided for this purpose.

The silkworm begins to secrete a proteinlike substance through its spinneret, a small opening under its jaws. With a bending motion a filament is spun around the worm in the form of the figure eight.

The silkworm is hidden from view within 24 h; in 3 days the cocoon is completed. It is about the size and shape of a peanut shell. The filament is in the form of a double strand or fibroin, which is held together by a gummy substance called sericin, or silk gum. Chemically, the silk fibroin and sericin are composed of approximately 95% protein and 5% wax, fats, salts, and ash. The liquid substance hardens immediately on exposure to the air. If left undisturbed, the chrysalis inside the cocoon develops into a moth within 2 weeks. To emerge, the moth must break through the top of the cocoon by excreting an alkaline liquid that dissolves the filament. As this cutting through damages the cocoon so that the filament cannot be unwound in one long thread, the growers terminate the life cycle at this point by a process known as stoving or stifling. The cocoons are heated to suffocate the chrysalis, but the delicate silk filament is not harmed. *See* LEPIDOPTERA.

Filature operations. The cocoons are delivered to a factory, called a filature, where the silk is unwound from the cocoons and the strands are collected into skeins. Some of the cocoons are produced scientifically in such factories. They are sorted according to color, size, shape, and texture, as all these affect the final quality of the silk.

After the cocoons have been sorted, they are put through a series of hot and cold immersions, because the sericin must be softened to permit the unwinding of the filament in one continuous thread. Raw silk consists of about 80% fibroin and 20% sericin. In this step only about 1% of the sericin is removed because this silk gum is a needed protection during the further handling of the delicate filament.

The process of unwinding the filament from the cocoon is called reeling. The care and skill used in the reeling operation prevent defects in the raw silk. As the filament of a single cocoon is too fine for commercial use, 3–10 strands are usually reeled at a time to produce the desired diameter of raw silk thread. The cocoons float in water, bobbing up and down as the filaments are drawn upward through porcelain eyelets and rapidly wound on wheels or drums while the operator watches to detect flaws. As the reeling of the filament from each cocoon nears completion, the operator attaches a new filament to the moving thread. Skilled operators have an uncanny ability to blend the filaments, always retaining the same diameter of the rapidly moving silk strand. The sericin acts as an adhesive, aiding in holding the several filaments together while they are combined to form the single thread.

The usable length of the reeled filament is from 1000 to 2000 ft (300 to 600 m). The remaining part of the filament is valuable raw material for the manufacture of spun silk.

The term reeled silk is applied to the raw silk strand that is formed by combining several filaments from separate cocoons. It is reeled into skeins, which are packed in small bundles called books, weighing 5–10 lb (2.3–4.6 kg). These are put into bales, ranging in weight from 135 to 145 lb (61 to 66 kg). In this form the raw silk is shipped to all parts of the world.

From the filature, the books of reeled silk go to the throwster where they are transformed into silk yarn, also called silk thread, by a process known as throwing. Persons engaged in this work are called throwsters. Silk throwing is analogous to the spinning process that changes cotton, linen, or wool fibers into yarn. Unlike the manufacture of these fibers, however, the manufacture of silk yarn does not include carding, combing, and drawing out, the usual processes for producing a continuous yarn. The raw silk skeins are sorted according to size, color, and length or quantity. Then they are soaked in warm water with soap or oil. This softening of the sericin aids in handling the thread. After being mechanically dried, the skeins are placed on light reels, from which the silk is wound on bobbins.

During this winding operation single strands may be given any desired amount of twist. If two or more yarns are to be doubled, they are twisted again in the same or in a reverse direction, depending on the kind of thread to be made. To equalize the diameter, the thread is run through rollers. It is then inspected and packaged ready for shipment to manufacturers for construction into fabric.

Wild silk. Wild or tussah silk may be distinguished from cultivated silk by its coarse, thick form, which appears flattened. Cultivated silk is a narrow fiber with no markings. Wild silk is a broader fiber with fine, wavy longitudinal lines running along its surface, giving it a dark hue under a microscope. *See* NATURAL FIBER.

M. David Potter

Sillimanite

A nesosilicate mineral of composition $Al_2O[SiO_4]$, crystallizing in the orthorhombic system, space group Pbnm, $a = 0.744$, $b = 0.760$, and $c = 0.575$ nanometers, with four formula units in the cell. Sillimanite commonly occurs in slender crystals or parallel groups, and is frequently fibrous, hence the synonym fibrolite (see **illus.**). There is one perfect

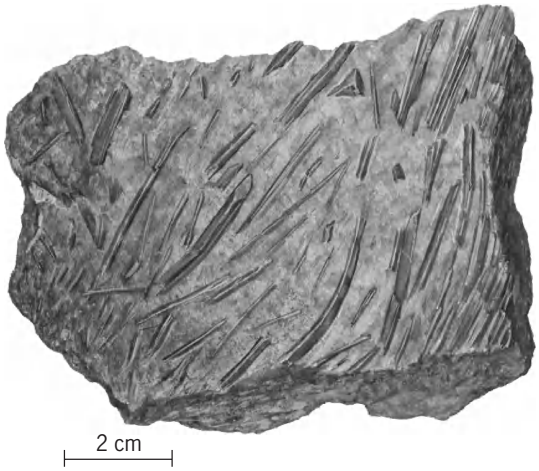

Sillimanite crystals from Chester, Connecticut. (*Specimen from Department of Geology, Bryn Mawr College*)

cleavage, luster is vitreous, color is brown, pale green, or white, hardness is 6–7 on Mohs scale, and the specific gravity is 3.23.

Sillimanite, andalusite, and kyanite are polymorphs of $Al_2O[SiO_4]$. In sillimanite one Al atom is coordinated by four and the other by six oxygens; in andalusite the Al atoms are five- and six-coordinate; and in kyanite all Al atoms are six-coordinate. The triple point where all three polymorphs coexist has been a source of much debate, but is probably close to 3.76 kilobars (376 megapascals) and 501°C (934°F), with kyanite the high-pressure phase. The three $Al_2O[SiO_4]$ polymorphs are important in assessing the metamorphic grade of the rocks in which they crystallized. Sillimanite occurs in high-temperature metamorphic argillaceous rocks. Common associated minerals are cordierite, biotite (a mica), and almandine (a garnet). In rocks which are poor in silica, sillimanite can coexist with corundum. *See* ANDALUSITE; KYANITE; SILICATE MINERALS.

Paul B. Moore

Bibliography. D. M. Kerrick, *Reviews in Mineralogy*, vol. 22: *The Al_2SiO_5 Polymorphs*, 1990.

Silurian

The third oldest period of the Paleozoic Era, spanning an interval from about 412 to 438 million years before the present. The Silurian system includes all sedimentary rocks deposited and all igneous and metamorphic rocks formed in the Silurian Period. Both the base and top of the Silurian have been designated by international agreement at the first appearances of certain graptolite species in rock sequences at easily examined and well-studied outcrops. *See* GEOLOGIC TIME SCALE.

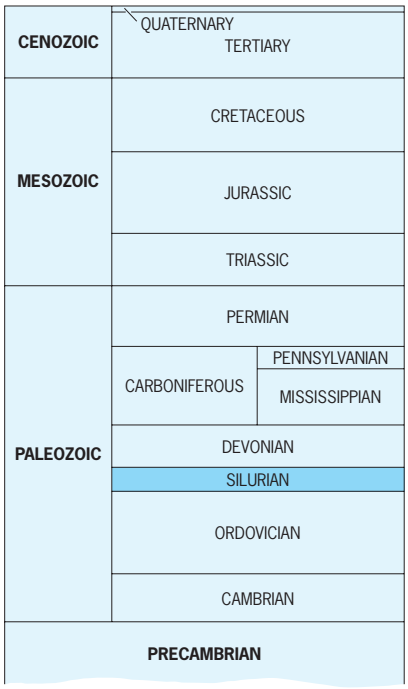

Divisions. The type area of the Silurian System is in adjoining parts of Wales and England. Based on studies of rocks and their contained faunas in that area, R. I. Murchison divided the Silurian System into three series. Each series is distinguished by a characteristic faunal association. Epochs are time intervals during which rocks of the corresponding series accumulated. From oldest to youngest, the three widely recognized divisions of the Silurian are Llandovery, Wenlock, and Ludlow. Analyses of faunas obtained from Silurian rocks led to recognition of a fourth and youngest series and epoch within the Silurian, the Pridoli. The Pridoli type area is in Bohemia, an area close to Prague, where marine faunas younger than those of the Ludlow and older than those of the Devonian have been obtained from outcrops. The stratotype or type section for the base of the Devonian is in the same area as rocks bearing typical Pridoli faunas.

Intervals with durations shorter than those of the epochs have been recognized in Britain. Three divisions of the Llandovery, two of the Wenlock, and two of the Ludlow are recognized on the basis of shelly (primarily brachiopod) and graptolite faunas. The divisions are termed stages. Ages are the time intervals during which rocks of the corresponding stages accumulated. Graptolites, the fossil remains of colonial, marine, planktic (floating) organisms, have been used to divide the Silurian in many parts of the world into intervals even shorter than those of the stages. These intervals are zones. Graptolite zonal successions have been developed in dark, graptolite-bearing shales of Silurian age in many parts of the world. Twenty-eight generally recognized graptolite zones have been proposed as divisions of the Silurian that had shorter durations than stages. Most of these zones may have had durations of about a million years, making separation of evolutionary and biogeographic developments within Silurian graptolite-bearing strata remarkably precise. Fifteen conodont zones are recognized as zonal divisions of the Silurian. They are most useful in carbonate successions. Organic-walled microfossil (chitinozoans and acritarchs) zonations have been developed that are especially useful in the high-latitude siliclastic successions. *See* GRAPTOLITHINA.

Paleogeography and lithofacies. Plate positions and plate motions during the Silurian significantly influenced the depositional environments, climates, and life of the period. Integration of data from remanent magnetism, distributions of reefs and other prominent successions of carbonate rocks, positions of shorelines, and positions of glacial deposits may be used to indicate certain features of Silurian paleogeographies and changes in them through the Silurian (see **illus.**). *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; FACIES (GEOLOGY); PALEOGEOGRAPHY; PALEOMAGNETISM; PLATE TECTONICS.

The most prominent feature of Silurian paleogeography was the immense Gondwana plate. It included much of present-day South America, Africa, the Middle East, Antarctica, Australia, and the Indian subcontinent. Numerous small plates lay near its margins.

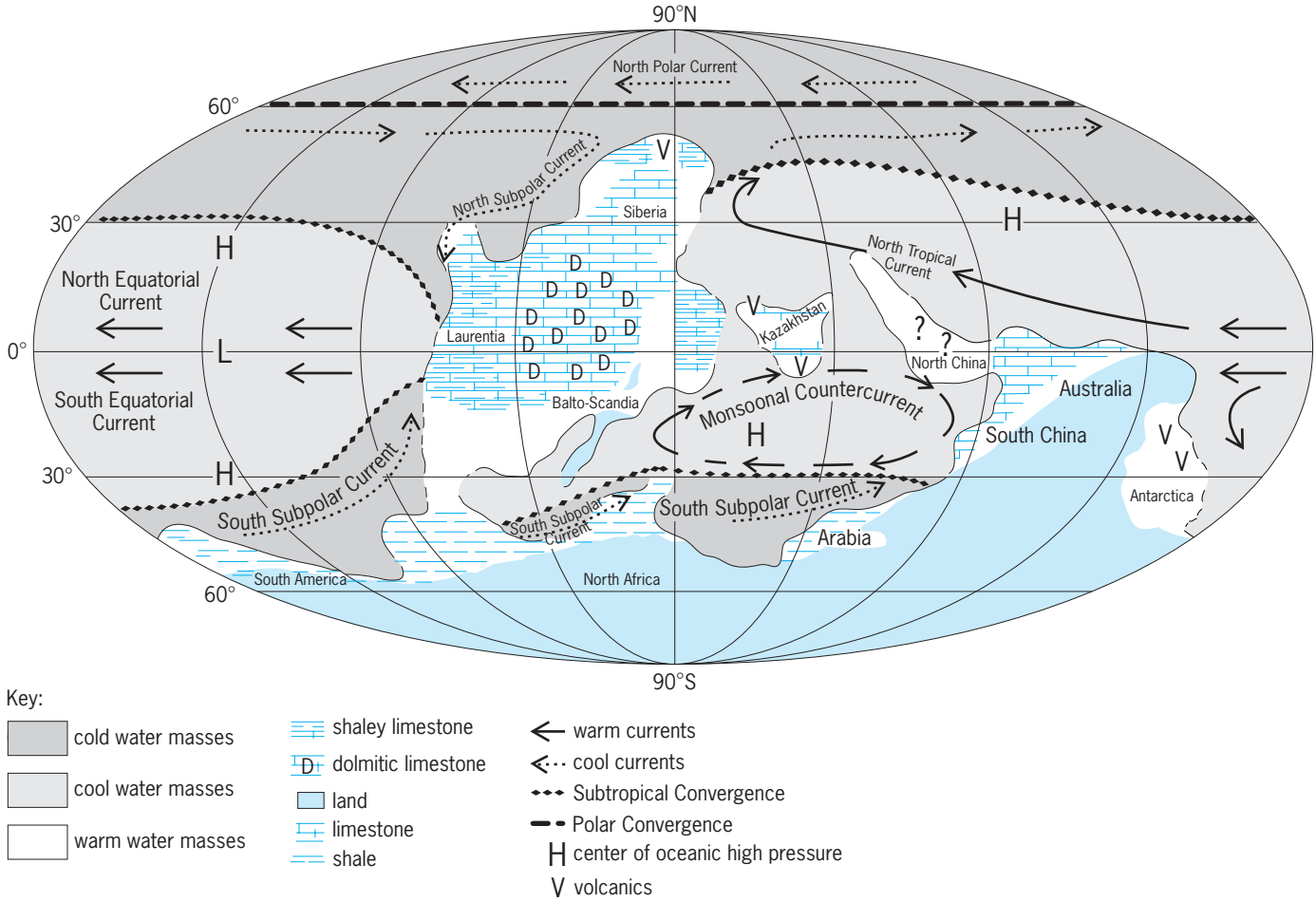

Lithofacies of the Upper Silurian (Ludlow), modified from A. J. Boucot et al. (1968). Paleogeographical reconstructions after Scotese (1986). (*After Spec. Pap. Paleontol. 44, 1991*)

Some of the small plates close to northern or equatorial Gondwana included certain areas of North China, Tibet, Southeast Asia, Asiatic China, New Guinea, and New Zealand. Small plates that were in mid to low latitudes included the plates that today make up southern Europe and Florida, as well as small plates now found as fragments in the Alps, South China, and Tarim (in Asiatic China). The modern South American and African portions of Gondwana lay in high latitudes about the South Pole. That pole probably was positioned approximately in eastern South America. A large salient (landform projection) that was composed of areas of the Middle East, India-Pakistan, Antarctica, and Australia extended north across the Equator from modern eastern Africa into the Northern Hemisphere. The small plates that constitute modern Southeast Asia-Malaysia were within the Northern Hemisphere tropics.

During the Silurian, many plates continued the relative northward motion that had commenced during the mid-Ordovician. As a consequence of these motions, the Avalon plate (South Wales, southern England, and nearby continental Europe; and the Avalon Peninsula of Newfoundland and parts of maritime Canada) collided with the eastern margin of Laurentia (much of modern North America, Greenland, Scotland, and portions of northwestern Ireland and western Norway) in the latest Ordovician. As that collision developed, a large mountainous area formed. It was the source of siliclastics that comprise the Queenston deltaic deposits in Appalachian rock sequences. It was also the source of siliclastic sediments that accumulated in Avalon plate sites. Baltica (the Baltic states, Scandinavia, and eastern Europe east to the Urals) moved north across the Equator and collided with the Avalon and Laurentian plates during the early part of the Silurian. Late Silurian deposits of east-central Laurentia include significant thicknesses of salt and other evaporites. This stratigraphic record suggests that by the latter part of the Silurian, the eastern side of Laurentia was close to 30° south latitude. The plate fragments that today constitute much of southern Europe moved northward to approach, and perhaps begin to collide with, each other. Most of these plates moved into the tropics during the latter part of the Silurian. The core of one such plate, Perunica (modern Czech Republic), was the site of a large volcano for much of the Silurian. Potentially, that plate resembled a modern atoll because carbonates flank the volcanics. The Kazakh plate (modern Kazakhstan) may have been a set of atolls in the Silurian. Volcanic rocks

of different types are common there, as are carbonates. The plate seems to have included one or more tropical atolls. The Siberian plate (much of modern Asiatic Russia, Mongolia, and areas in northwestern China) moved northward through the Northern Hemisphere tropics. Portions of it may have entered the relatively temperate conditions north of 30° north latiude by the end of the Silurian. The northern part of modern Africa also moved northward during the Silurian, coming close to or even entering the tropics by the end of the period. *See* CONTINENTS, EVOLUTION OF.

Plate positions and plate motions as well as topographic features of the plates controlled depositional environments and lithofacies. These, in turn, significantly influenced organismal development and distributions.

Much of high-latitude Gondwana was the scene of continental glaciation in the latest Ordovician. Initially, the ice melted rapidly, leading to a rapid sea-level rise that continued from the latest Ordovician into the early Silurian. Central areas of Gondwana rose after the ice melted, presumably as a result of isostatic rebound. That rebound was followed by tectonic doming. The consequence of a rising central Gondwana is reflected in the steady northward and westward spread of continental and nearshore marine deposits away from land areas and across continental shelves during the Silurian. The Gondwanan land was the source for much of the highly organic-rich mudstone that characterizes North African and Middle Eastern Silurian strata. Isotopic ratios of strontium-87 to strontium-86 may be obtained from brachiopod shells and conodonts. In such analyses, a relatively higher proportion of strontium-87 to strontium-86 suggests that an old granitic source land is being eroded. Studies of the ratio of the strontium-87 to strontium-86 for Silurian brachiopods and conodonts are consistent with continued erosion of a relatively old granitic source during the Silurian. As the southern European and modern Alpine plates moved northward to enter the tropics, they became sites of, initially, cool-water carbonate deposition. That was followed by reef formation and extensive bahamian-type carbonate deposition. *See* CHEMOSTRATIGRAPHY; GLACIAL EPOCH; ISOTOPE; STRONTIUM.

Three brief glacial episodes have been recognized in the Llandovery stratigraphic record of the Amazon Basin. Each of these episodes took place essentially at the time of each of the boundaries between the stage divisions of the Llandovery Series. The South American stratigraphic record also suggests that a brief glaciation took place in the late Wenlock, at about the height of impact of the Baltic plate against Avalon-Laurentia. The South American Silurian record of glaciation suggests that the modern eastern part of South America lay close to or across the South Pole at that time.

Plates that were within the tropics were sites of extensive carbonate deposition. Reefs developed along and near the margins of most plates that were within the tropics.

Silurian Northern Hemisphere plates, other than a portion of Siberia, are not known north of the Northern Hemisphere tropics. Presumably, nearly all of the Northern Hemisphere north of the tropics was ocean throughout the Silurian.

Ocean circulation. Absence of plates bearing continental or shallow shelf marine environments north of about 45° north latitude indicates that ocean circulation in most of the Silurian Northern Hemisphere was zonal. Ocean currents were relatively strong and flowed from east to west north of 60° north. Major ocean currents between 30 and 60° north flowed from west to east. Surface currents would have been deflected south along the west side of Siberia, and upwelling conditions could have formed in that area.

Ocean surface currents in the tropics would have been influenced strongly by the prevailing westerlies. Primary surface circulation from the Equator to 30° north and south of it would have been from east to west. The large peninsulalike salient of Gondwana would have deflected currents northerly along it in the Northern Hemisphere and southerly along it in the Southern Hemisphere. North China's eastern shores would have been sites of upwelling. The presence of richly fossiliferous coralline limestones and volcanics there is consistent with surface-water turbulence and upwelling conditions along the margins of volcanic islands. Much the same conditions appear to have prevailed along the margins of the Kazakh plate.

The large size of the Gondwana plate and the presence of land over much of it would have led to development of seasonal monsoon conditions. Monsoonal conditions would have led to seasonal reversal of major surface circulation adjacent to the land, as is seen in modern India. In the austral summer–Northern Hemisphere winter, winds would have blown offshore from the warm land and resulted in an east to west current flow near 30° south and a west to east surface current near the Equator. In the Northern Hemisphere summer and austral winter, surface water flow west of the lands on the Gondwana plate would have reversed. These seasonal reversals in surface current directions could have created long-term, year-long upwelling in a pattern similar to that observed in modern oceans off the Somali coast. Consequently, intense upwelling across northern Africa could have been generated during the Silurian in a pattern similar to that in the modern Arabian Ocean. That upwelling generated vast quantities of organic matter which became organic-rich shales that characterize the North African Silurian. In general, surface circulation west of tropical Gondwana would have been east to west. However, surface circulation between the several plates that drifted into the tropics as well as those that moved closer to each other than they had been previously would have created many small east-to-west flowing gyres in the tropics. Modest-to-strong upwelling along many plate margins was likely as a consequence. The occurrence of numerous long-studied reefs, such as those on the island of Gotland off the Swedish coast, close to or along

tropical plate margins is consistent with plate margin upwelling.

Surface circulation south of 30° south would have hit the western side of Laurentia and flowed generally northward along it to about 30°, at which latitude the surface currents probably turned to flow east to west between 30° south and the Equator. *See* CORIOLIS ACCELERATION; PALEOCEANOGRAPHY; UPWELLING.

Climate and hydrology. Collision of the Avalonian and Laurentian plates in the latest Ordovician coincides with development of the Southern Hemisphere continental glaciation. Erosion of the land area formed at the Avalon-Laurentian plate collision generated a large volume of coarse to fine-grained siliclastic materials. Such extensive erosion could have reduced the atmospheric concentration of carbon dioxide, creating climate conditions cool enough to allow glaciation to develop. Although atmospheric carbon dioxide concentration probably was significantly greater during the Silurian than it is today, solar heat coming to the Earth was about 4–5% less than today. Thus, depression of the atmospheric carbon dioxide content could have been enough to allow glaciation. Reduction in the extent and height of the land being eroded could have resulted in increased atmospheric carbon dioxide concentration, ending the conditions suited to generating glaciers by the latter part of the Wenlock.

That part of South America (modern eastern South America) near the South Pole for the early part of the Silurian was not just cold, but also the site of as many as four brief glacial episodes. These episodes took place during the time that the Avalon and Baltica plates collided with each other and with the Laurentian plate.

Rainfall would have been significant on the mountainous lands formed along the Avalon-Laurentian plate collision boundary because that land lay within the prevailing winds. *See* PALEOCLIMATOLOGY.

Atmosphere. Analyses of major rock suites have led to the suggestion that the Silurian global air temperature was about 9°F (5°C) warmer than today. The postulated global air temperature is consistent with an estimated carbon dioxide content of three to four times that of the modern atmosphere. Oxygen content of the Silurian atmosphere has been estimated to have been about two-thirds to three-fourths the present content. A lesser oxygen concentration in the atmosphere would have resulted in less oxygen available to be stirred into the oceans. A lesser oxygen concentration in the surface oceans coupled with a warmer global temperature meant that less oxygen was available to marine organisms in the Silurian seas than is available to modern oceanic organisms. Silurian marine organisms, especially those living at depths significantly below the surface, must have survived on significantly less oxygen than do modern organisms. Furthermore, the ocean oxygen minimum zone would have attained shallower depths than it does in modern oceans. Oxygen content in ocean waters may have declined to near zero by 330 ft (100 m) in tropical waters. Atmospheric conditions would have had a significant influence on the distributions of marine benthic dwellers. *See* ATMOSPHERE, EVOLUTION OF.

Land life. Both nonvascular and vascular plants continued to develop in land environments following their originations in the early mid-Ordovician. Many of these Silurian plants were mosslike and bryophytelike. Plants with vascular tissues had developed in the mid-Ordovician. These plants continued their spread in terrestrial environments during the Silurian. Psilophytes assigned to the genus *Cooksonia* were relatively widespread in Late Silurian terrestrial environments. The probable lycopod (club moss) *Baragwanathia* apparently lived in nearshore settings in modern Australia during the latter part of the Silurian.

Silurian land life also included probable arthopods and annelid worms. Fecal pellets of wormlike activity have been found as well as remains of centipede-, millepede-, and spiderlike arthropods. *See* PALEOBOTANY; PALEOECOLOGY.

Marine life. Shallow marine environments in the tropics were scenes of rich growths of algae, mat-forming cyanobacteria, spongelike organisms, sponges, brachiopods, bryozoans, corals, crinoids, and ostracodes. Nearshore marine siliclastic strata bear ostracodes, small clams, and snails and trilobites. Certain nearshore strata bear the remains of horseshoe-crab-like eurypterids. Some of them may have been significant predators. *See* ALGAE; BRACHIOPODA; BRYOZOA; CRINOIDEA; CYANOBACTERIA; OSTRACODA; TRILOBITA.

Fish are prominent in a number of Silurian nearshore and some offshore marine environments. Jawless armored fish of several kinds occur in Silurian strata. These fish include many species of thelodonts that had bodies covered with minute bony scales, heterostracans, and galeaspids that had relatively heavily armored head shields, and anaspids that possessed body armor consisting of scales and small plates. Jawed fish were relatively rare in the Silurian. They were primarily spiny sharks or acanthodians. As well, there are remains of true sharklike fish and fish with interior bony skeletons (osteichthyes) in Late Silurian rocks. *See* ANASPIDA; HETEROSTRACI; OSTEICHTHYES; THELODONTIDA.

The acanthodians appear to have been relatively common in the latter part of the Silurian. The oldest known placoderms were found in Silurian strata in South China. Fish began to diversify in the latter part of the Silurian in many shelf sea environments. The planktic colonial marine graptolites are the prominent organism found in rocks that formed under anoxic or near-anoxic waters. Their remains are most plentiful in rocks that accumulated on the outer parts of shelves and in basins of the Silurian. *See* ANOXIC ZONES.

The extinct microfossil group, the conodonts, were relatively common in many carbonates deposited in shelf seas. Small, slender shells of squidlike cephalopods occur in many shelf-sea rock suites, including certain of the black, organic-rich graptolite-bearing sequences. These cephalopods appear to

have been nektic in habit. *See* CEPHALOPODA; CONODONT.

Biogeography. Both land plants and marine animals were distributed in patterns reflective of latitudinal temperatures during the Silurian. Those organisms living along the margins of the Gondwana plate in the general position of 30 to 45° south latitude constituted one floral and faunal realm, the Malvinokaffric, which persisted throughout the Silurian. Organisms that typify it were adapted to cool climates.

Land plants and marine organisms living during the early part of the Silurian (Llandovery into Wenlock) in warm temperate to tropical conditions outside of the cool temperate to polar Malvinokaffric Realm were essentially cosmopolitan. This distribution reflected sparsity of tropical marine life in the tropics after the major Late Ordovician into earliest Silurian extinctions among marine organisms. Both land plants and marine organisms were increasingly provincial during the latter part of the Silurian. Marine bottom-dwelling invertebrates living on shelves of plates in the tropics were noticeably provincial by the close of the Silurian. The northern part of the Siberian plate was characterized during the latter part of the Silurian by a unique association of shelled marine invertebrates of which the brachiopod *Tuvaella* is characteristic. Laurentia-Baltoscanian shallow marine environments were inhabited by marine bottom-dwelling faunas, of which the brachiopod *Salopina* has been selected to give its name. The Kazakh plate shelf seas had a unique, endemic marine invertebrate fauna.

Nearshore marine fish had distributions similar to those of the marine bottom-dwelling invertebrates. Silurian fish from South China are an exception. They constituted a distinct faunal province.

Among Late Silurian land plants, the truly tracheophytic *Baragwanathia* known from eastern Australia was distinct from coeval land plants known in other parts of the tropics. Although sparse, tropical land plants appear to have had distributions similar to those of the marine organisms. This evidence suggests that plate positions and the ocean and wind circulation patterns they influenced were major factors in the distribution of life during the latter part of the Silurian. *See* BIOGEOGRAPHY.

Economic resources. Silurian dark, organic-rich rocks form one of the six prominent suites of petroleum source rocks known. The six suites have provided more than 90% of the world's known oil and gas reserves. Silurian source rocks have generated approximately 9% of the world's reserves. Most of the Silurian source beds for petroleum and gas are those that lay under the upwelling waters along the northern and northwesterly margins of the Gondwana plate. Other known prominent Silurian source rocks are those that accumulated under upwelling waters on the eastern side of the Baltoscanian-Avalonian plates and those that accumulated under upwelling waters along the westerly side of the Laurentian plate.

Sedimentary iron ores accumulated in tidal wetlands on the Avalonian, especially the West Avalon, plates during the Llandovery. At that time, the Avalonian plates lay south of the tropics.

Silurian carbonate rocks have been quarried in many parts of the world for building stone and for the raw material for cement. Quarries in the area of the type Wenlock, at Wenlock Edge in Britain, have yielded richly fossiliferous limestones used in construction. Late Silurian orthoceroid cephalopod-bearing limestones interbedded with the black, graptolite-bearing shales in southern Europe have been quarried for a spectrum of ornamental uses, including tables and decorative paneling. Silurian salt deposits have been mined extensively in the eastern United States. These salts were a major economic resource in the 1800s. Silurian carbonate rocks in Nevada bear large quantities of gold, which occurs as tiny flakes. Today the amount of gold recovered from the Nevada sites has made this area one of the two or three largest gold producers in the world. Silurian rocks have played significant roles in a number of local and regional economies. *See* FOSSIL; PALEOZOIC; PETROLEUM GEOLOGY.

William B. N. Berry

Bibliography. M. G. Bassett, P. D. Lane, and D. Edwards (eds.), *The Murchison Symposium: Proceedings of an International Conference on the Silurian System*, Spec. Pap. Palaeontol. 44, Palaeontological Association, London, 1991; W. B. N. Berry and A. J. Boucot, *Correlation of the African Silurian Rocks*, GSA Spec. Publ. 147, 1973; W. B. N. Berry and A. J. Boucot, *Correlation of the North American Silurian Rocks*, GSA Spec. Publ. 102, 1970; J. Gray and W. Shear, Early life on land, *Amer. Sci.*, 80:444–456, 1992; J. C. Gutierrez-Marco and Rabano (eds.), *Proceedings of the 6th International Graptolite Conference of the GWPG (IPA) and the 1998 Field Meeting of the International Subcommission on Silurian Stratigraphy (ICS-IUGS)*, Instituto Tecnologico Geominero de Espana, Temas Geologico-Mineros, 23, 1998; C. H. Holland and M. G. Bassett (eds.), *A Global Standard for the Silurian System*, Nat. Mus. Wales Geol. Ser. 10, 1989; E. Landing and M. Johnson (eds.), *Silurian Cycles: Linkages of Dynamic Stratigraphy with Atmospheric, Oceanic, and Tectonic Changes*, New York State Mus. Bull. 491, 1998; G. T. Moore et al., A paleoclimate simulation of the Wenlockian (Late Early Silurian) world using a general circulation model with implications for early land plant paleoecology, *Palaeogeog. Palaeoclimatol. Palaeoecol.*, 110:115–144, 1994.

Siluriformes

The catfishes, a large and highly diverse order of the Ostariophysi. Another ordinal name commonly used is Nematogathi, from the Greek *nematos*, thread, and *gnathos*, jaw, a descriptive name referring to the long barbels around the mouth of most catfishes. Catfishes differ from other members of the Ostariophysi in having a more complex Weberian apparatus and in lacking intermuscular bones, as well as lacking subopercle, symplectic, or intercalar bones in the head. They are further characterized by either

lacking scales (in which case they are covered with tough skin) or being covered with overlapping bony plates; having vomer, pterygoid, and palatine bones that are usually toothed; having maxillae that are rudimentary and toothless; lacking a parietal bone; usually having one to four pairs of barbels (some members of the family Loricariidae lack visible barbels); having (usually) an adipose fin and having pectoral and dorsal fins that are each preceded by a strong spine in most species.

The order includes about 2800 species in 35 families and is represented on all continents, albeit only by fossils in Antarctica. Catfishes are chiefly freshwater species; Ariidae (tropical waters of both hemispheres) and Plotosidae (Indo-Pacific) are considered marine families, but they too include freshwater species. Aspredinidae (tropical South America), considered a freshwater family, has some brackish species. Twelve families of catfishes are known only in Central and South America; three families are restricted to Africa; five are endemic to eastern Asia; three are shared by Africa and Asia; two are Indonesian, one Eurasian; one ranges from India to southeast Asia;, and one, the Ictaluridae, is endemic to North America.

The following families were selected to represent the world of catfishes and give the reader some insight into their diversity.

Ictaluridae. Ictaluridae is the largest family of fishes indigenous to North America, where their native range is eastward from the Continental Divide to the Atlantic and Gulf drainages and from southern Canada southward to Guatemala. Characters that distinguish the ictalurids are four pairs of barbels, one pair nasal, one pair of maxillary, and two on the chin; naked skin; usually six soft dorsal fin rays; dorsal and pectoral fins with spines (dorsal spine absent in *Prietella*); and an adipose fin free or adnate.

Three of its seven genera (*Prietella, Satan*, and *Trogloglanis*) are represented by blind, subterranean species. The bullheads (*Ameiurus*) with seven species, the catfish (*Ictalurus*) [see **illustration**] with nine species, and the flathead catfish (*Pylodictis*) are large-sized members of the family. Pond-reared channel catfish (*I. punctatus*) support a large food industry in the southeastern United States. The 25 species in the genus *Noturus* are pygmies among the catfishes, the largest hardly exceeding 150 mm (6 in.) and most being far smaller than that. Madtoms spend rather secretive lives in small streams, where they hide by day and become active at night. They have a poison gland at the base of the pectoral fin spine enabling them to deliver a painful sting.

Clariidae. These are the so-called air-breathing catfishes or labyrinth catfishes of Africa, India, southeastern Asia, western Indonesia, and the Philippines. They are distinguished by having a very long dorsal fin not preceded by a spine; having four pairs of barbels; and lacking pectoral and pelvic fins in some species. Most interesting are the labyrinthic organs in the upper branchial cavities, which greatly increase the surface area for exchange of respiratory gases, thus allowing the fish to breathe atmospheric air. Not all clariids are so equipped, but those that are can move short distances over land, thus the "walking catfish." The ability to move over land enhances dispersal, as demonstrated by *Clarias batrachus*, a species accidentally introduced in Florida.

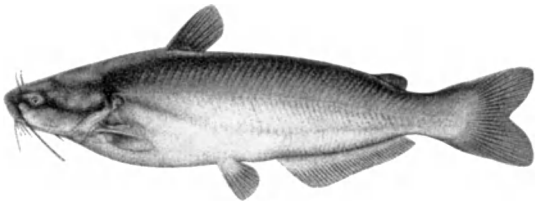

Blue catfish (*Ictalurus furcatus*). (*Drawing © Joseph R. Tomelleri*)

Ariidae. Ariid catfishes are in general appearance similar to the genus *Ictalurus* (Ictaluridae), but easily distinguished from them by the absence of nasal barbels. They are represented in all tropical and subtropical seas of the world, and some species enter freshwater on a regular basis, while others occupy freshwater throughout their lives. Most species of the family practice oral incubation (eggs are carried in the mouth until hatching). The male carries the eggs throughout their development, and the young are not released until they can fend for themselves. The mature eggs are the largest known among teleost fishes, measuring 14–18 mm (0.5–0.7 in.) in diameter. During the incubation period the oral cavity is enlarged by depressing the hyoid bones, making room for several dozen 40-mm (1.6 in.) young.

Doradidae. Doradids are called the thorny catfishes. They are freshwater species of South America, primarily in Guiana, Brazil, and Peru. The body, with few exceptions, has a lateral row of bony plates, each usually armed with a spine; there are three pairs of barbels, nasal barbels being absent; the mandibular barbel may be branched; pectoral and dorsal fin spines are stout and serrated; the skull extends as a bony plate to the base of the dorsal fin. These are the "talking catfishes." By movements of the pectoral spines or by vibrations of the swim bladder, the fishes make a growling sound. The smaller species of the family, along with the similar families Callichthyidae and Loricariidae, also from South America, are popular aquarium fishes.

Trichomycteridae. Members of this family occur in Costa Rica, Panama, and throughout South America. They are identified by the following features: elongate body lacking scales; chin barbels usually absent, nasal barbels usually present; dorsal and anal fins short; pelvic fins absent in some species; usually no adipose fin; and posteriorly directed spines on the opercle. These characteristics adapt the fish for a parasitic life. Indeed, they do enter the gill cavity of larger fish, where they nip the gills and feed on the ensuing blood. The infamous candiru even parasitize humans, by entering the urethra in a rheotaxic response to a stream of urine. Once the candiru enters the urethra, the opercular spines prevent its removal and surgery is required.

Not only are catfishes diverse in morphology,

biology, and life histories, but also they vary tremendously in size. Adults of the little madtoms (*Noturus*) of North America may weigh only a few grams, while the Mekong giant catfish (*Pangasianodon gigas*, family Pangasidae) has been reported by the Thai Inland Fisheries Department to weigh up to 293 kg (646 lb). *See* ACTINOPTERYGII; CYPRINIFORMES; OSTEICHTHYES. Herbert Boschung

Bibliography. T. M. Berra, *An Atlas of Distribution of Freshwater Fish Families of the World*, University of Nebraska Press, Lincoln, 1981; J. G. Lundberg, The comparative anatomy of the toothless blindcat, *Trogloglanis pattersoni* Eigenmann, with a phylogenetic analysis of the ictalurid catfishes, *Misc. Publ. Mus. Zool. Univ. Michigan*, 163:1–85, 1982; J. G. Lundberg, The phylogeny of ictalurid catfish: A synthesis of recent work, pp. 392–420 *in* R. L. Mayden, (ed.), *Systematics, Historical Ecology, and North American Freshwater Fishes*, Stanford University Press, Stanford, California, 1992; J. S. Nelson, *Fishes of the World*, 3d ed. Wiley, New York, 1994; W. R. Taylor, A revision of the catfish genus *Noturus* Rafinesque with an analysis of higher groups in the Ictaluridae, *Bull. U.S. Nat. Mus.*, 282:1–315, 1969.

Silver

A chemical element, Ag, atomic number 47, atomic mass 107.868. It is a gray-white, lustrous metal. Chemically it is one of the heavy metals and one of the noble metals; commercially it is a precious metal. There are 25 isotopes of silver with atomic masses ranging from 102 to 117. Ordinary silver is made up of the isotopes of masses 107 (52% of natural silver) and 109 (48%). *See* PERIODIC TABLE.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Although silver is the most active chemically of the noble metals, it is not very active in comparison with most other elements. It does not oxidize as iron does when it rusts, but it reacts with sulfur or hydrogen sulfide to form the familiar silver tarnish. Electroplating silver with rhodium prevents this discoloration. Silver itself does not react with dilute nonoxidizing acids (hydrochloric or sulfuric acids) or strong bases (sodium hydroxide). However, oxidizing acids (nitric or concentrated sulfuric acids) dissolve it by reaction to form the unipositive silver ion, Ag^+. The Ag^+ ion is colorless, but a number of silver compounds are colored because of the influence of their other constituents.

Silver is almost always monovalent in its compounds, but an oxide, fluoride, and sulfide of divalent silver are known. Some coordination compounds of silver, also called silver complexes, contain divalent and trivalent silver. Although silver does not oxidize when heated, it can be oxidized chemically or electrolytically to form silver oxide or peroxide, a strong oxidizing agent. Because of this activity, silver finds considerable use as an oxidation catalyst in the production of certain organic materials.

Soluble silver salts, especially $AgNO_3$, have proved lethal in doses as small as 2 g. Silver compounds may be slowly absorbed by the body tissues, with a resulting bluish or blackish pigmentation of the skin (argyria).

Silver is a rather rare element, ranking 63rd in order of abundance. Sometimes it occurs in nature as the free element (native silver) or alloyed with other metals. For the most part, however, silver is found in ores containing silver compounds. The principal silver ores are argentite, Ag_2S, cerargyrite or horn silver, AgCl, and several minerals in which silver sulfide is combined with sulfides of other metals; stephanite, $5Ag_2S \cdot Sb_2S_5$; polybasite, $9(Cu_2S, Ag_2S) \cdot (Sb_2S_3, As_2S_3)$; proustite, $3Ag_2S \cdot As_2S_3$; and pyragyrite, $3Ag_2S \cdot Sb_2S_3$. About three-fourths of the silver produced is a by-product of the extraction of other metals, copper and lead in particular. *See* SILVER METALLURGY.

Pure silver is a white, moderately soft metal (2.5–3 on Mohs hardness scale), somewhat harder than gold. When polished, it has a brilliant luster and reflects 95% of the light falling on it. Silver is second to gold in malleability and ductility. Its density is 10.5 times that of water. The quality of silver, its fineness, is expressed as parts of pure silver per 1000 parts of total metal. Commercial silver is usually 999 fine. Silver is available commercially as sterling silver (7.5% copper) and in ingots, plate, moss, sheets, wire, castings, tubes, and powder.

Silver, with the highest thermal and electrical conductivities of all the metals, is used for electrical and electronic contact points and sometimes for special wiring. Silver has well-known uses in jewelry and silverware. Silver compounds are used in many photographic materials. In most of its uses, silver is alloyed with one or more other metals. Alloys in which silver is an ingredient include dental amalgam and metals for engine pistons and bearings. *See* PHOTOGRAPHIC MATERIALS; SILVER ALLOYS. William E. Cooley

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; J. R. Davis (ed.), *Metals Handbook: Desk Edition*, 2d ed., ASM International, 1998; D. R. Lide, *CRC Handbook of Chemistry and Physics*, 85th ed., CRC Press, 2004.

Silver alloys

Combinations of silver with one or more other metals. Pure silver is very soft and ductile but can be hardened by alloying. Copper is the favorite hardener and normally is employed in the production of sterling silver, which must contain a minimum of 92.5% silver, and also in the production of coin silver.

In the United States coin silver contained 90% silver, the balance being copper. To reduce the consumption of silver in the coinage, the 10-cent and 25-cent pieces were replaced by a composite material with cupro-nickel layers outside a core of copper. The density and electrical conductivity of the composite are such that the coins are accepted by coin-operated mechanisms. The 50-cent piece was made of outer layers of a silver-copper alloy with a core of copper silver. This peculiar combination must be ascribed to political metallurgy.

Pure silver melts at 961.93°C (1763.5°F) on the International Practical Temperature Scale of 1968, but this is lowered by the addition of copper to a minimum of 779°C (1434°F) at 28% copper. This silver-copper eutectic and modifications containing other elements such as zinc, tin, cadmium, phosphorus, or lithium are widely used for brazing purposes, where strong joints having relatively good corrosion resistance are required. Where higher strengths at elevated temperature are required, silver-copper-palladium alloys and other silver-palladium alloys are suitable. The addition of a small amount of silver to copper raises the recrystallizing temperature without adverse effect upon the electrical conductivity. For this reason, copper containing 10–20 oz (280–560 g) of silver per ton is desirable for commuter bars and armature windings in certain dc machinery.

An unusual use for silver is the 25% silver—75% palladium alloy used as a diffusion septum transmitting only hydrogen. The high electrical and thermal conductivities of pure silver, slightly exceeding those of copper, along with its resistance to oxidation, have led to the large use of silver for electrical contacts. Because of the tendency of silver to form sulfide films, it is desirable to employ voltages in excess of about 12 V and reasonable pressures in using such contacts. Frequently, the silver for contact is alloyed with 10% copper or with a small amount of cadmium, or, better still, with cadmium oxide, which improves the behavior of the contact material under many conditions. Cadmium alloyed with silver raises electrical resistivity. This effect, undesirable for electrical contacts, is reduced if the cadmium is present as the oxide. Oxide content materials may be made by powder metallurgy methods or by heating a silver cadmium alloy in air for some hours during which oxygen diffuses into the alloy and converts the cadmium to the oxide.

Noble-metal alloys. Silver may be alloyed with gold or palladium in any ratio, producing soft and ductile alloys; certain of these intermediate alloys are useful for electrical contacts, where resistance to sulfide formation must be achieved. This requires about 40–50% palladium or slightly more gold. Sterling and coin silver cost slightly less. In an early effort to reduce the cost of silver articles, the technique of bonding silver to copper and rolling the composite material into sheets was developed in England, particularly in the vicinity of Sheffield; the product was therefore known as Sheffield plate. However, with the development of the method of electrodepositing silver from a double cyanide solution by G. R. and H. Elkington in 1844, the production of Sheffield plate almost disappeared and was replaced by electroplated silver. This was the first commercial use of electroplating, and basically the same solution is still employed. The Elkingtons found that the most satisfactory base metal upon which silver could be plated was the ancient Chinese alloy known as paktong, which was found to contain nickel, copper, and zinc. It is still made under the general name of German silver or nickel-silver and still constitutes the most satisfactory base material. *See* ELECTROPLATING OF METALS.

Technical alloys. Silver is a useful component for high-duty bearings in aircraft engines, where it may be overlaid with a thin layer of lead and finally with a minute coating of indium. *See* METAL COATINGS.

The fact that silver is readily oxidized and reduced is a limitation to its use for certain electrical purposes, particularly where silver conductors are in contact with vulcanized fiber or similar materials and are subjected to a reasonable voltage gradient. Under these circumstances, the silver tends to migrate and develop conducting paths within the insulation which are difficult to detect.

Specially developed alloys of silver with tin, plus small percentages of copper and zinc in the form of moderately fine powder, can be mixed with mercury to yield a mass which is plastic for a time and then hardens, developing relatively high strength despite the fact that it contains about 50% mercury. This material was developed specifically for dental use and is generally known as amalgam, although the term amalgam actually includes all the alloys of mercury with other metals. Dental amalgam is widely used as a tooth-filling material and is a useful dental material. It possesses some limitations in strength and also discolors in the mouth, but silver has some germicidal effect which probably is helpful in this instance in preventing further decay, which is particularly important in children's teeth. *See* AMALGAM.

Silver is a component in many of the colored gold alloys used not only for jewelry but also for dentistry; when it is added to the gold-copper alloys or the palladium-copper alloys, a large improvement results in the strengths obtainable by heat treatment. The addition of platinum or palladium to some of these alloys further augments their strength or hardenability and also renders them more suitable for rubbing electrical contacts. The softness of silver, its insolubility in iron, and its freedom from oxidation at high temperatures make silver powder useful in preventing the threads in high-temperature bolting from sticking at temperatures up to about 650°C (1200°F). In the absence of silver, such bolts may bond together after a short time so that they cannot be unscrewed.

Where silver is to be used at high temperatures, the pure metal should be employed; it has been so used for electrical windings in motors operating as high as 500°C (930°F). At moderate temperatures, silver coatings are frequently useful in electrical devices to ensure good electrical contact in bolted joints and to provide a low-resistivity coating in equipment which will be operated at high frequency where the skin conduction is of primary importance. Thin silver or gold electroplate also may be applied to copper and copper parts to facilitate soldering, particularly if the parts are to be stored for some time and may become oxidized and difficult to solder.

Because silver does not oxidize on heating, it is used in applying electrically conductive coatings to ceramics. A paste containing finely divided silver plus certain additives is applied to the ceramic. Upon heating to redness, the silver coating becomes firmly bonded to the ceramic. With care, the product can be soldered; the method is used in making connections to ceramic capacitors and similar devices. Similar coatings are used for decorating glassware. These may be coated with additional silver by electrodeposition and finally with rhodium to prevent tarnishing.

Because of its resistance to acetic acid and many other organic materials, as well as to alkalies, silver-lined equipment finds use in the chemical industry—in autoclaves, piping, and similar pieces of equipment, some of it very large. In high-temperature applications, the high solubility of oxygen in silver and the rapid diffusion of oxygen through it must be recognized. Silver, alloyed with a base metal, will suffer internal oxidation when heated in air, and a silver coating applied to a base metal, such as iron, affords no useful protection at high temperatures because the oxygen will go through the silver, oxidizing the iron at the interface. Furthermore, silver will be weakened along grain boundaries if heated in an oxidizing atmosphere and then in a reducing atmosphere, particularly if the latter contains hydrogen. *See* SILVER; SILVER METALLURGY. Edmund M. Wise

Bibliography. A. Butts (ed.), *Silver, Economics, Metallurgy and Use*, 1967, reprint 1975; R. P. Elliot (ed.), *Constitution of Binary Alloys: First Supplement*, 1965, reprint 1989; M. Hansen, *Constitution of Binary Alloys*, 2d ed., 1988; W. Hume-Rothery and R. E. Smallman, *The Structure of Metals and Alloys*, 1988; F. A. Shunk (ed.), *Constitution of Binary Alloys: Second Supplement*, 1969, reprint 1986.

Silver metallurgy

The art and science of extracting silver metal economically from various ores, and the reclamation of silver from the myriad types of industrial processes or scrap produced therefrom. It includes all processes of separating silver from its ores, alloys, and solutions, as well as the smelting, refining, and working of the metal and its alloys and compounds. It deals with the technical application of the chemical and physical properties of silver, to its concentration, extraction, purification, alloying, working, and compounding to meet the requirements of ever-expanding technical needs.

Silver is widely distributed in nature and occurs in both native and combined forms. Although one of the most important ores of silver, tetrahedrite ($Cu_8Sb_2S_7$ with silver sulfide in chemical combination), is hypogene in origin, native silver, argentite (Ag_2S), and cerargyrite (AgCl) are distinctly supergene; and pyrargyrite (Ag_3SbS), stephanite (Ag_5SbS_4), polybasite (Ag_9SbS_6), and proustite (Ag_3AsS) are probably hypogene, having developed as a result of supergene sulfide enrichment. In a few cases, ores are treated for silver only, or for combined gold and silver values. However, at least three-fourths of the world silver production is obtained as a by-product of base-metal ores, mainly those of copper, lead, and zinc. *See* ACANTHITE; CERARGYRITE; PYRARGYRITE; TETRAHEDRITE.

Extractive processes. Silver extraction has been effected by a number of processes. A few of the important historical processes, as well as modern-day practice, are discussed below.

Amalgamation. This is a process of separating metal from ore by alloying the metal with mercury. The Patio process was introduced in Pachuca, Mexico, in 1557. The silver ore was finely ground, mixed with water to form a thin mud, and then treated with salt, copper sulfate, and mercury. The entire mass was then thoroughly mulled by walking horses or mules through the mixture. The resulting silver amalgam was retorted to separate the silver and mercury. Large amounts of silver were produced by this method from the time of its inception until the end of the nineteenth century, when it was superseded by other methods. Other amalgamation processes of importance are the Tintin, Cazo, Kröhnke, Washoe, Combination, Boss, Barrel or European, Pan or Reese-River, and Franke-Tina, all of which were derived from the basic Patio process.

Chloridization. Chloridization is a method of producing silver by precipitating it from aqueous solutions. In chloridization methods the silver contained in ores or metallurgical products is first converted by means of a chloridizing roast into a compound which is soluble in water or in certain aqueous solutions. The silver is then precipitated as an insoluble compound by suitable reagents and the precipitate worked for the metal. The soluble silver compounds are in the form of chloride, sodium thiosulfate solution, sulfate, or the double cyanide of silver and potassium. Principal chloridization processes are the Augustin process, which uses brine as a solvent; the Patera process, in which sodium thiosulfate is used; the Kiss process, in which calcium thiosulfate is used; the Russell process, in which silver metal or sulfide is brought into the solution with sodium-copper thiosulfate; and the Ziervogel process, in which the silver is converted to the sulfate and dissolved in hot water.

Cyanidation. This process differs only in minor details, such as strength of the solution and time of treatment, from the cyanide process used in the recovery of gold. Further, it differs from the chloridization process described above in that no preliminary metallurgical treatment other than fine grinding is required, as the cyanide is capable of dissolving

metallic silver and, to a greater or lesser extent, all the silver compounds occurring in the various ores. *See* GOLD METALLURGY.

Pattinson lead process. Although the Pattinson lead process has been superseded by the Parkes process described below, it is still occasionally used. The Pattinson process is based on the fact that, if molten argentiferous lead is slowly cooled, crystals of silver-poor lead separate from the still molten portion which is correspondingly enriched in silver. If the richer liquid portion is separated, it can again be separated into a poorer solid portion and a still richer liquid alloy. This operation can be repeated until the silver is concentrated to approximately 750 oz/ton (24,000 g/metric ton), at which point no further separation is possible. Similarly, the portions poor in silver may be successfully treated as above until, by repetition, lead very low in silver is produced. For economic reasons this process is no longer in extensive use.

Parkes process. This process is based on the greater affinity of silver for zinc than for lead. Slab zinc is added to an argentiferous lead bath, the temperature of which has been raised higher than the melting point of the zinc. When the zinc has been melted and thoroughly mixed into the lead bath, the bath temperature is lowered and a silver-zinc alloy separates and floats on the top of the kettle. This zinc crust, as it is now called, is pressed off to remove the excess lead. The crust is broken up, cooled, and later retorted to recover the major part of the zinc, which is then reused. The resultant high silver-lead retort bullion is then cupelled to recover the silver and gold as doré and the lead as litharge. *See* LEAD METALLURGY.

Cupellation. This is the oldest and most widely known method of separating and recovering gold and silver from lead. It is used extensively to recover doré metal from the silver-lead end product of the Parkes process. The silver-lead bullion resulting from retorting of the zinc crust of the Parkes process is charged to a reverberatory-type cupellation furnace. After the charge has melted, air is blown across the top of the molten bath, causing oxidation of the lead and other impurities which separate as an impure litharge slag, leaving behind the silver and gold as a doré alloy. The doré is cast into anodes and sent to the refinery, where it is electrolytically separated into pure gold and silver.

Slimes smelting. By far the most important method of silver production results from the smelting and subsequent treatment of the silver-rich slimes resulting from the electrolytic refining of copper. Before smelting, the raw slimes are subjected to treatment with a sulfuric acid and air leach to remove excess copper and to combine the silver with the selenium and tellurium that are present. The resultant leached slimes are filtered through plate and frame presses to remove the excess solution and charged to a reverberatory-type furnace with appropriate fluxes. The melt slag resulting from this initial melting contains most of the base-metal impurities. Upon completion of this step, the material remaining in the furnace consists of gold, silver, selenium, tellurium, and residual base metals, and is called matte. At this point, treatment of the matte begins first with sodium carbonate to purge residual sulfur, if any, and then with niter (potassium nitrate) and air, forming a water-soluble slag containing the selenium and tellurium. In some cases caustic soda is used as a collector in conjunction with the niter and air treatment. Upon completion of the selenium and tellurium removal, the white metal or impure bullion is subjected to continued niter treatments, sometimes with the addition of lead, to remove the residual base metals. The resulting doré is then cast into anodes for subsequent electrolytic refining.

Refining. Refined metal from the doré furnaces (or from any previously mentioned production method) is treated electrolytically in either the Möbius or Balbach-Thum cell or variants thereof. The Möbius cell is more conventional, since because of the vertical suspension of the anodes and the cathodes, it is similar to the arrangement of a copper refinery. In the Balbach-Thum cell the cathode is placed on the cell bottom and the anodes are suspended in a basket parallel to the cathode. The major difference in the two methods is the fact that the Möbius cell requires less space than the Balbach-Thum cell in relation to the quantity of silver produced but produces 15–25% scrap which must be reprocessed, whereas the Balbach-Thum cell produces no scrap. The silver crystals produced by both methods are collected, washed, and cast into fine silver bars which are then ready for industrial or artistic use.

Recovery from secondary sources. Of importance to the world silver supply is the recovery of silver from scrap generated by the arts and industry. The more important sources of secondary silver are sterling and silver plate production; electrical components such as batteries, relays, and contact points; solders and brazing and other alloys; chemicals; mirrors, lamps, and jewelry; catalysts; and photographic and related processes. Because of the complexity of reclamation and the value of the above items, the sources of secondary silver may enter the smelting and refining process at the initial blast or reverberatory furnace smelting stage, the copper or lead refining stage, or, if high enough in silver purity, the doré production stage. In all cases, however, the final product is refined silver. *See* SILVER; SILVER ALLOYS.

R. D. Mushlitz

Bibliography. A. Butts (ed.), *Silver: Economics, Metallurgy, and Use*, 1967, reprint 1975; L. Coudurier, D. W. Hopkins, and I. Wilkomirsky, *Fundamentals of Metallurgical Processes*, 1985; M. A. Meyers and K. K. Chanla, *Mechanical Metallurgy: Principles and Applications*, 1984; H. W. Pollack, *Materials Science and Metallurgy*, 4th ed., 1988.

Silviculture

The theory and practice of controlling the establishment, composition, and growth of stands of trees for any of the goods (including timber, pulp, energy, fruits, and fodder) and benefits (water, wildlife habitat, microclimate amelioration, and carbon

sequestration) that they may be called upon to produce. In practicing silviculture, the forester draws upon knowledge of all natural factors that affect trees growing upon a particular site, and guides the development of the vegetation, which is either essentially natural or only slightly domesticated, to best meet the demands of society in general and ownership in particular. Based on the principles of forest ecology and ecosystem management, silviculture is more the imitation of natural processes of forest growth and development than a substitution for them.

Origins. Silviculture is the oldest deliberate application of ecology, recognized as a practice long before the term "ecology" was coined. Some of the most complex silvicultural techniques have been practiced for centuries in the tropics within tree garden systems, but on very small scales. However, in most parts of the world the dominant historical influence has been the clearing of forests for farmland either permanently or through more traditional cultivation of swidden systems (slash-and-burn agriculture). In Europe and the Middle East, much of this clearance occurred in pre-Roman times. Up until the late Middle Ages, the main use of woodland was in the provision of small and medium-scale materials for fuelwood, building poles, fencing, and wattle. These products were supplied mostly by a variety of short rotation coppice systems (thickset of small trees). Demand for timber was satisfied by standard trees left to continue to grow among the coppice or to form isolated trees in wood pastures where animals grazed the undergrowth or grass beneath. In the early 1600s, wood scarcity and the decline in land productivity became an important resource issue across much of Europe. Modern silviculture was therefore born out of necessity to develop restoration techniques for sustaining the productivity of woodlands and forests; even in the twenty-first century, exploitation of forests and depletion of resources usually precede silvicultural restoration. In the eighteenth and nineteenth centuries, silviculture began to break away from its empirical basis to find a foundation in the underlying sciences that were developing.

Treatments. The techniques proceed on the assumption that the natural vegetation of any site normally tends to extend itself to occupy all available growing space, making the fullest possible use of the available growth factors, such as light and moisture. The forester attempts, usually by cutting while harvesting useful wood products, to create vacancies in the forest vegetation that will provide environments favorable either to the establishment of new, desirable trees (reproduction treatments) or to the enhanced growth of remaining trees (intermediate operations).

Intermediate operations. These are made during the life of a particular crop of trees to alter species composition, to provide income, to increase the amount or value of the goods and services produced, and to alter forest structure in order to improve wildlife habitat. Under intensive management, intermediate operations may include various kinds of release cuttings to free desirable trees not past the sapling stage

Fig. 1. Release treatment for timber tree saplings in a mixed dipterocarp rainforest in Malaysia. (*Yale University School of Forestry and Environmental Studies*)

from the competition of taller or faster-growing trees that either are undesirable species or are defective specimens of otherwise acceptable species (**Fig. 1**). If the trees to be removed in such cuttings are not to be harvested, they may be killed either by girdling or by applying herbicides. *See* HERBICIDE.

If trees large enough to have rough bark are to be killed in place, the girdling is usually done by cutting away or injecting herbicides into a band of bark and cambium. This halts the downward movement of carbohydrates from the leaves to the roots; the starvation of roots ultimately kills the whole tree. Smaller trees or clumps of sprouts with thin bark also can be killed by basal spraying in which oil solutions of herbicides are applied entirely around the bases of the stems to girdle them and also to kill any dormant buds beneath the bark. It is possible to kill whole stands of undesirable woody plants by spraying foliage from the ground or air; this technique depends on essentially normal movement of materials through the phloem to carry the herbicides from the leaves to the roots at times when the roots will be killed. Under proper conditions, foliage spraying can be used to kill broad-leaved trees without significant harm to narrow-leaved conifers beneath them. *See* PLANT ANATOMY; PLANT TRANSPORT OF SOLUTES.

Thinnings are made in immature stands to stimulate the growth of the trees that remain and to increase the total production (**Fig. 2**). Thinning has its basis in that new stands start with many hundreds or thousands of trees per acre and ultimately decline to several hundreds or merely dozens as a result of severe competition. The crowns and roots

of individual trees expand, and the individuals that grow fastest in height overtop and ultimately kill the laggards.

Artificial thinning represents a judicious acceleration and guidance of this process. As long as the trees of a stand are sufficiently numerous to be capable of expanding to occupy the available growing space completely, the gross volume of raw wood produced by a stand of given composition and age on a particular site remains nearly independent of the number of trees per acre. Because a given volume of wood has greater value and utility if produced in a small number of large trees rather than a large number of small trees, it is generally desirable to reduce the stocking periodically to the smallest number of good trees that will fully occupy the site. The process can enhance the growth of the remaining trees in diameter but not in height. A little production is lost because some space is temporarily unoccupied by foliage.

Thinning also increases the actual yield of a stand through the harvest of surplus marketable trees that might otherwise be killed by competition and lost to decay. The practice often imitates the natural process by eliminating the smaller trees in favor of the larger. However, if the smaller trees happen to be more desirable than the larger trees, and if they are of sufficient vigor, it is possible to alter the pattern of removals to rescue them. Thinning enables the forester to guide an existing stand along lines of optimum development and to secure current income from an otherwise immature crop.

Thinning can be used to reallocate growing space for the purpose of altering wildlife habitat or the visual esthetics in a stand. For example, removing some of the larger trees will increase the vertical structural heterogeneity of the forest by releasing smaller trees that will grow at different rates among each other. Removing all the smaller trees in a stand can increase the depth and viewshed for a person walking along a trail.

Pruning is the removal of branches from the stems of living trees, ordinarily for the purpose of increasing the volume of clear (knot-free) wood. The practice is best confined to high-quality trees destined to be favored by thinning as part of the final crop.

Reproduction treatments. These are made when the prospective growth of an old stand is deemed less valuable or desirable than that of a new stand. Reliance is often on natural seeding or upon the release of "advance regeneration" already present beneath the old stand. Some species, mostly angiosperms, can renew themselves simply by sending up new sprouts from stumps or roots after their tops are killed by cutting or fire. Under intensive practice, it is common to depend on artificial (or direct) seeding or, more commonly, the planting of nursery-grown seedlings. These forms of artificial regeneration are used when natural regeneration is regarded as too slow, where there is no source of natural seed, when selected superior strains are to be introduced, or simply when there is reason to use logging methods that are incompatible with natural regeneration. *See* HERBICIDE.

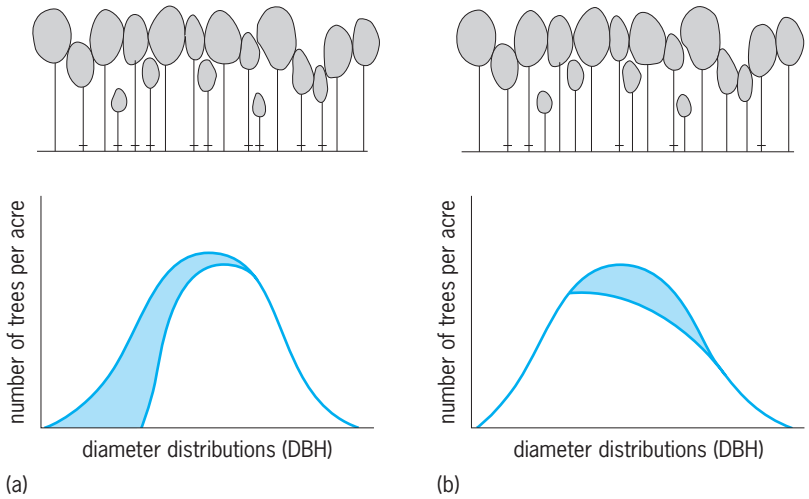

Fig. 2. Types of thinning. (*a*) Previously untreated even-aged stand marked (by horizontal lines at the base of the stems) for a low thinning whereby trees removed are the smallest trees in the stand. (*b*) Crown thinning whereby trees removed are the direct competitors of the remaining dominant canopy. In the diameter distributions (DBH), the shading depicts that part of the original stand that is removed in the thinning.

Successful regeneration from seed, whether naturally or artificially applied, demands basic knowledge of the ecological requirements for germination and survival of the desired species. An abundant source of seed is essential. The condition of the seed bed and the arrangement of the remaining trees also must be adjusted to provide an adequate number of spots where the environment is more favorable to establishment of the desired species than to any others. The native flora of any forest region includes species adapted to colonize almost any kind of vacancy that might be created in the forest by natural disturbances, great or small. The forester attempts to simulate the kind of disturbance that will lead to establishment of the particular species desired. *See* DENDROLOGY.

Stand patterns and processes. The spatial patterns in which old trees are removed and the species that replace them determine the structure and developmental processes of the new stands. If all the trees are replaced at once with a single species, the result is so-called pure stand or monoculture in which all of the trees form a single canopy of foliage that is lifted ever higher as the stand develops. If several species start together from seeds, from small trees already present, or from sprouts, as a single cohort, the different species usually tend to grow at different rates in height. Some are adapted to develop in the shade of their sun-loving neighbors. The result is a stratified mixture (**Fig. 3**). Such stands grow best on soils or in climates, such as in tropical moist forests, where water is not a limiting factor and the vegetation collectively uses most of the photosynthetically active light. If trees are replaced in patches or strips, the result is an uneven-aged stand which may be of one species or many.

These different spatial and temporal patterns of stand structure are created by different methods of reproduction. The simplest is the clear-cutting

Fig. 3. Stand patterns. (*a*) Stratified mixture of red oak and eastern hemlock in a temperate moist forest in northeastern Connecticut. (*b*) Stratified mixture of mahogany with several subcanopy species of palm and an understory shrub layer in a moist subtropical rainforest in Puerto Rico. (*Yale University School of Forestry and Environmental Studies*)

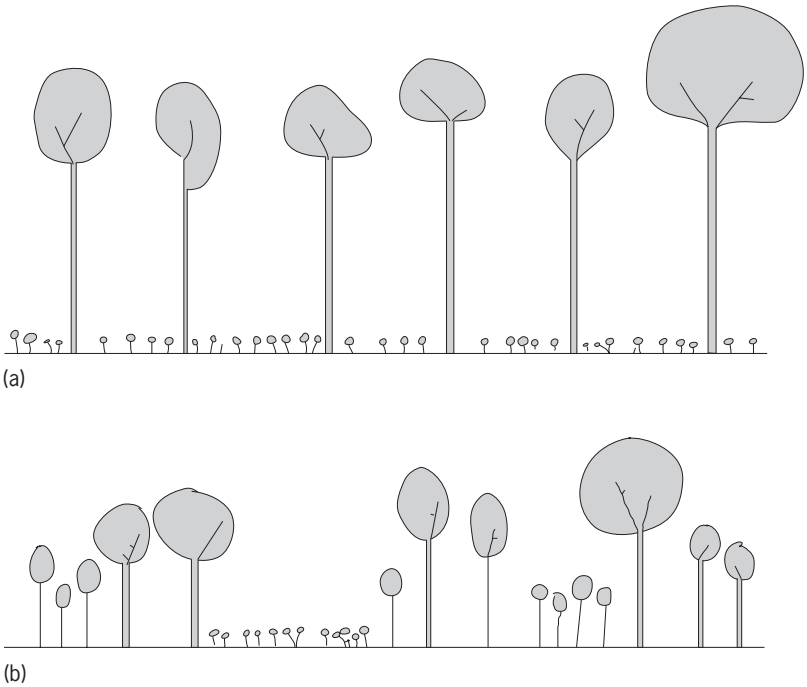

Fig. 4. Reproduction treatments. (*a*) Even-aged stand after uniform shelterwood cuttings to establish advance regeneration. The uniform shelterwood method involves the removal of the entire stand in a series of partial cuttings that extend over a fraction of the rotation. (*b*) Uneven-aged stand after single tree selection treatments. The selection method involves the removal of scattered individuals or strips or groups of trees at relatively short intervals; such cuttings are repeated indefinitely.

method, in which virtually all of the vegetation is removed. Although it is sometimes possible to rely on adjacent uncut stands as sources of seed, it is usually necessary to reestablish the new stand by artificial seeding or planting after clear-cutting. The seed tree method differs only in that a limited number of trees are temporarily left on the area to provide seed.

In the shelterwood method, enough trees are left on the cutting area to reduce the degree of exposure significantly and to provide a substantial source of seed (**Fig.** 4*a* and **Fig. 5***a, b*). In this method the growth of a major portion of the preexisting crop continues, and the old trees are not entirely removed until the new stand is well established. The three methods just described lead to the creation of essentially even-aged stands.

The various modifications of the selection method, however, lead to the creation of uneven-aged stands characterized by the presence of three or more distinct age classes (Fig. 4*b* and **Fig. 6**). Reproduction treatments under the selection method occur in the same stand at more or less equal intervals of time and are not concentrated near the end of the life of the crop as in the even-aged stands. Both the shelterwood and selection methods are subject to very large amounts of variation and modification; both are applied in one form or another to the management of almost all species and forest types, except the few that actually need exposure to become established. In reality, selection and shelterwood methods have a high degree of overlap. This overlap can be

Fig. 5. Shelterwood methods. (*a*) Uniform shelterwood cutting in an 100-year-old stand consisting of oak and other hardwood trees in Connecticut. (*b*) Group shelterwood where regeneration has been established in distinct canopy openings throughout a beech stand in the Czech Republic. (*Yale University School of Forestry and Environmental Studies*)

perceived as dependent upon the size of the openings created within a single entry for a selection treatment (for example, the amount of trees removed) as compared to the arrangement, age-class distribution, and amount of reserve trees actually left standing until the next rotation for a shelterwood.

The so-called coppice methods involve establishment of new stands from sprouts arising from the stumps or roots of cut trees. They are used mainly for the production of quick crops of such species as the poplars for fuel and pulp. However, some species, such as oaks and coast redwoods, can grow to very large sizes from sprouts.

Clear-cutting is not necessarily crude silviculture; it may even represent the final harvesting operation following a series of frequent intermediate cuttings. Conversely, light periodic partial cuttings do not invariably represent highly efficient practice, but may involve "high-grading," in which the few good trees of an otherwise untended stand are cut to leave it dominated by undesirable trees.

The application of partial cutting often involves the concept of financial maturity as a guide in determining which trees to cut and which to reserve for additional growth. According to this concept, the market value of each standing tree is regarded as a capital investment, and an attempt is made to predict the increase in value of the tree between the cutting immediately in prospect and the next cutting. If this increase in value does not represent an acceptable rate of compound interest on the investment represented by the present value of the tree, the tree is marked for immediate cutting. Under this concept, fast-growing trees of good quality are reserved, while poor unthrifty trees are likely to be cut.

Ecological factors. The choice of methods of regeneration cutting depends on the ecological status of the species and stands desired. Species that characterize the early stages of succession, the so-called pioneers, will endure, and usually require the kind of exposure to sunlight resulting from heavy cutting or, in nature, severe fire, catastrophic windstorms, floods, and landslides. These pioneer species usually grow rapidly in youth but are short-lived and seldom attain large size. They are exemplified by slash, lodgepole, and jack pines as well as by aspens, paper birch, and red alder. The species that attain greatest age and largest size are ordinarily those which are intermediate in successional position and tolerance of shade. They are moderately resistant to exposure when young; while they can usually endure the shade of light-foliaged pioneer trees, shrubs, or herbs, they seldom persist beneath a closed canopy of their own species. Some of them will become established after the severe exposure of clear-cutting, but they often start best with light initial shade such as that created by shelterwood cutting. This group is exemplified by Douglas-fir, loblolly pine, the five-needled pines (white pines), sweet gum, tulip poplar, and the ashes. Their longevity and large size result from the fact that they are naturally adapted to reproduce after disturbances occurring at relatively long intervals. *See* ECOLOGICAL SUCCESSION; FOREST ECOSYSTEM.

The shade-tolerant species representing late or climax stages in the succession are adapted to reestablish themselves in their own shade. They usually require such protection when young; they can survive as stunted advance regeneration for many years and suddenly start rapid growth when released. These species represent natural adaptations to the kinds of fatal disturbance caused by insects, disease, and atmospheric agencies rather than to the more complete disturbance caused by fire. Since they can reestablish themselves almost continuously, longevity is not necessary for survival of the species. If these species are already established as advanced regeneration, new stands can be created simply by removing the overstory in any pattern that is otherwise desirable, provided that the small trees are

Fig. 6. Single-tree selection method of reproduction in a Norway spruce–silver fir stand in Bavaria, Germany. (*Yale University School of Forestry and Environmental Studies*)

not destroyed in the process. If advanced growth is not established when the time comes to replace the stand, comparatively light partial cutting (shelterwood or selection) is necessary to accelerate the regeneration process. Among the species of this group are the true firs, spruces, hemlocks, maples, beeches, and oaks. Some of the limitations that ecological factors place on the germination and early survival of seedlings can be evaded by planting ones that have survived the crucial early stages in the protected environment of a nursery.

Mixed and pure stands. Much silvicultural practice is aimed at the creation and maintenance of pure, even-aged stands of single species. This approach is analogous to that of agriculture and simplifies administration, harvesting, and other operations. The analogy is often carried to the extent of clear-cutting and planting of large tracts with intensive site preparation, especially with species representative of the early or intermediate stages of succession.

Mixed stands are more difficult to handle from the operational standpoint but are more resistant to injury from insects, fungi, and other damaging agencies which usually tend to attack single species. They are also more attractive than pure stands and make more complete use of the light, water, and nutrients available on the site. They usually do not develop unless these site factors are comparatively favorable; where soil moisture or some other factor is limiting, it may be possible for only a single well-adapted species to grow. If the site is highly favorable, as in tropical rainforest, river floodplains, or moist ravines, it is so difficult to maintain pure stands that mixed stands are inevitable.

In mixed stands, the different species have a tendency to occupy different strata in the crown canopy, depending on their rate of height growth or position in succession. If all start in the same year or within a few years of one another, any pioneer species that are present quickly occupy the upper stratum even though they may succumb after a few decades. Thereafter the ascendancy passes to species that are intermediate in succession, and in nature they may dominate for the remainder of the next century. When they begin to dwindle in number from natural attrition, the long-suppressed species of the late successional stages begin to take over more of the upper canopy levels. In the silvicultural treatment of such stands, it is often appropriate to anticipate, imitate, accelerate, and guide the sequence of events. The modifications usually take the form of cuttings in which individual trees or strata are removed when they are mature. It is possible to proceed in such ways that either the even- or uneven-aged structure is created or maintained; in other words, mixed stands can be either even- or uneven-aged.

The various strata representing different sets of species and successional stages can also appear or be caused to develop at successive intervals separated by many decades. If otherwise desirable, the harvesting of the successive overstories can proceed in a manner that even more nearly duplicates the events of natural succession than where all species become established nearly simultaneously.

Accessory practices. The application of silviculture involves a number of accessory practices other than cutting. Fire, like the ax, is an agent that can be used wisely to build the forest or unwisely in its witless destruction. In localities of high fire risks, it may be desirable to burn the slash (logging debris) after cutting. Not only does this reduce the potential fuel, but it may also help the establishment of seedlings by baring the mineral soil or reducing the physical barrier represented by the slash. Slash disposal is most often necessary where the cutting has been very heavy or where the climate is so cold or dry that decay is slow.

Deliberate prescribed burning of the litter beneath existing stands of fire-resistant species is sometimes carried out even in the absence of cutting to reduce the fuel for wild fires, to kill undesirable understory species (including those representing stages in plant succession later than those desired), to enhance the production of forage for wild and domestic animals, and to improve seedbed conditions. Prudence demands that such burning be prescribed only by experts under carefully chosen weather conditions and with ample provision for controlling the fires. The practice has been used mainly in stands of fire-resistant two- and three-needled pines, primarily in the southern United States, where it closely imitates the effect of natural fires. *See* FOREST FIRE.

Heavy disk plows, rolling brush cutters, modified bulldozers, and similar equipment also find use in the elimination of undesirable vegetation and preparation of the soil for the establishment of new crops by planting or natural seeding. The growth of a stand is sometimes improved by chemical fertilization or by drainage of swamps and bogs.

Silvicultural practices may be conducted not only for the growth of wood products but also for development of forest stands of form and composition consistent with wise management of water, wildlife, scenery, and other forest resources. Almost all forests provide sources of water. Earthworms and other soil organisms that feed on fallen leaf litter generally maintain soils of high porosity. Water that sinks into porous soils does not erode or move sediments into streams but can emerge from clear-water springs after it has been filtered through the soil. Almost any type of forest vegetation will maintain suitable soil porosity. The cutting of trees does not cause erosion, floods, or reduced rainfall, but may actually reduce transpiration losses of water enough to cause small increases of stream runoff, usually during dry periods. However, turbid water can flow off roads and trails, so it is important that water is diverted from them as soon as possible and, preferably, into areas where the sediment will filter out before they reach streams. In regions where winter snows are melted mainly by the sun rather than rain, water can be stored in snow drifts; these can be induced to form by cutting narrow openings in the forests.

Silviculture that is aimed at wildlife management (such as that for timber management) works by controlling the structure and species composition of stands. However, it also involves the creation of more kinds of stands. Since animals are mobile, it is much more difficult to learn their life histories than is the case with trees, and to determine the kinds of stands to create. The objective is to build the kinds of habitats that will call forth the desired animal population.

If maximum diversity of plant and animal species is desired, forests are required to have a wide range of age classes, stand structures, and species. Some wildlife species require conditions (such as hollow trees suitable for nests) that go with stands or trees older than those that would be grown just for timber production. Others, including many game animals, may require very young stands or even the elimination of trees in favor of shrubs, grasses, or herbaceous plants. If grazing of cattle is to be done beneath the trees, the establishment of grasses is promoted by heavy thinning, elimination of understory vegetation, or prescribed burning.

Managing cutting areas for esthetic purposes may include a variety of measures. All cutting areas look bad, so it is best to screen them from public view by leaving trees and removing debris along routes of travel. Log loading areas should also be kept out of sight. Where cutting areas will be visible for long distances, it is best to avoid straight boundaries and to effect irregular shapes that simulate natural openings. *See* CONSERVATION OF RESOURCES.

Ecosystem management. Integrated schedules of treatment for stands are called silvicultural systems. They cover both intermediate and reproduction treatments but are classified and named in terms of the general method of reproduction cutting contemplated. Such programs are evolved for particular situations and kinds of stands with due regard for all the significant biological and economic considerations involved. These considerations include the desired uses of the land, kinds of products and services sought, prospective costs and returns of the enterprise presented by management of the stand, funds available for long-term investment in stand treatments, harvesting techniques and equipment employed, reduction of losses from damaging agencies, and the natural requirements that must be met in reproducing the stand and fostering its growth.

It is also necessary to coordinate silvicultural systems for the treatment of individual stands with plans for aggregation of stands that are called forests in an administrative sense. It is generally desirable to plan the length of rotations and the arrangement of different age classes of trees or stands (**Fig. 7**) so that the forest as a whole will be the source of continuous, sustained yield of products as well as other benefits that forests can provide indefinitely (water, recreation). For a discussion of the planning, organization, and the administration of whole forests for these objectives *see* FOREST MANAGEMENT.

Identifying the natural pattern within a landscape is the best starting point for defining appropriate scales of stands for forest management. Physiographical attributes of soil nutrition, texture, and soil moisture availability to plants, along with coarser descriptors such as aspect, slope, and topographic position, when integrated together, are the best determinants

Fig. 7. Two age classes of ponderosa pine developed from natural seeding. The stand in the background is about 60 years old and that in the foreground is 35 years old. (*Yale University School of Forestry and Environmental Studies*)

of site productivity. The ability to identify how site productivity changes across different temporal and spatial scales provides the base guideline for defining thresholds of management practice from one site to another. Conceptually this seems logical, but in reality threshold limits in site productivity usually change across landscapes as continua rather than within any clearly identifiable abrupt transition. Such an example might be a ridge subject to drought, with a thin to bedrock soil that obviously will have a lower site productivity with a narrower set of management options and ecological constraints, as compared to a fertile, moist, and deep soil in the adjacent valley. Economic values demand that the two sites be managed differently, which means that somewhere between the two sites the management on the more productive site will be delineated from that on the less productive. This will likely accentuate abruptness between ridge and valley ecosystems.

Therefore, though ecological processes are without boundaries, ecosystem management is only feasible when silvicultural operations can be considered at certain discrete scales. The three most important scales for accommodating ecosystems processes include forest stands, forest watersheds, and physiographic regions. The different scales are appropriate for different kinds of forest management and should be integrated as a whole. Stand scales are best for managing forests where direct manipulation for the purposes of yielding a product is necessary (including timber, fruits, and fuelwood). Other service-related values must be considered at watershed (landscape) scale (such as water management, recreation, and wildlife habitat). The whole physiographic region is a useful scale for conceiving management operations for sequestering carbon or preserving biodiversity. David M. Smith; Mark S. Ashton

Bibliography. J. W. Barrett (ed.), *Regional Silviculture of the United States*, 3d ed., 1994; T. W. Daniel, J. A. Helms, and F. S. Baker, *Principles of Silviculture*, 2d ed., 1979; R. D. Nyland, *Silviculture: Concepts and Applications*, 1996; C. D. Oliver and B. C. Larson, *Stand Dynamics*, 1996; D. M. Smith et al., *Practice of Silviculture: Applied Forest Ecology*, 9th ed., 1997; U.S. Forest Service, *Silvics of North America*, 1990; U.S. Forest Service, Silvicultural Systems for the Major Forest Types of the United States, *Agr. Handb.*, no. 445, 1983.

Simple machine

Any of several elementary machines, one or more of which is found in practically every machine. The group of simple machines usually includes only the lever, wheel and axle, pulley (or block and tackle), inclined plane, wedge, and screw. However, the gear drive and hydraulic press may also be considered as simple machines. The principles of operation and typical applications of simple machines depend on several closely related concepts. *See* BLOCK AND TACKLE; HYDRAULIC PRESS; INCLINED PLANE; LEVER; SCREW; WEDGE; WHEEL AND AXLE.

Two conditions for static equilibrium are used in analyzing the action of a simple machine. The first condition is that the sum of forces in any direction through their common point of action is zero. The second condition is that the summation of torques about a common axis of rotation is zero. Corresponding to these two conditions are two ways of measuring work. In machines with translation, work is the product of force and distance. In machines with rotation, work is the product of torque and angle of rotation.

Work is the product of a force and the distance through which it moves. For example, the work done in raising a 10-lb object 15 ft is 150 ft-lb. In this example the work done on the weight goes into increasing the potential energy of the object. Work and energy, both potential and kinetic, have the same units, and in general the purpose of a machine is to convert energy into work. *See* EFFICIENCY; MECHANICAL ADVANTAGE; WORK.

For rotating machines, it is more convenient to consider torque and angular displacements than force and distance. Work is then expressed as the product of the torque and the angle (in radians) through which the object rotates while acted on by the torque. Torque, in turn, is the force exerted at a given radius from an axis of rotation. Thus, a 10-lb force at the end of a 15-ft crank exerts a torque of 150 lb-ft. *See* TORQUE.

Power is the rate of doing work. For example, one horsepower is arbitrarily defined as 550 ft−lb per second, or 33,000 ft-lb per minute. *See* MACHINE; POWER.

Richard M. Phelan

Simulation

A broad collection of methods used to study and analyze the behavior and performance of actual or theoretical systems. Simulation studies are performed, not on the real-world system, but on a (usually computer-based) model of the system created for the purpose of studying certain system dynamics and characteristics. Simulation is both an experimental science and an art or craft in the development of the precise, "valid" model that captures the necessary elements of the system, and the use of the model to analyze system behavior under various scenarios.

Simulation is used in many disciplines to model different types of systems. For example, simulation can be used in task or situational training areas in order to allow humans to anticipate certain situations and be able to react properly; decision-making environments to test and select alternatives based on some criteria; scientific research contexts to analyze and interpret data; and understanding and behavior prediction of natural systems, such as in studies of stellar evolution or atmospheric conditions.

The focus of this article is the use of simulation in decision making. Simulation helps provide information to aid the decision maker in selecting the proper alternatives. Simulation then becomes a decision-aiding tool, and with simulation a decision maker can try out new designs, layouts, software programs, and systems before committing resources to their acquisition or implementation; test why certain phenomena occur in the operations of the system under consideration; compress and expand time; gain insight about which variables are most important to performance and how these variables interact; identify bottlenecks in material, information, and product flow; better understand how the system really operates (as opposed to how everyone thinks it operates); and compare alternatives and reduce the risks of decisions.

Advantages. Simulation is a powerful and flexible tool because a simulation study can be tailored to provide answers to many different questions. The goal of each simulation study varies; however, there are some fundamental reasons why people use simulation to analyze real-world systems:

1. Simulation can be used to predict outcomes. If a bank wants to reconfigure its operations, simulation can be used to test various layouts and scenarios of operation.

2. Most real-world systems are stochastic (changes in the system are determined by chance) and dynamic (the system changes over time). In this case, the development of an analytical solution for studying the systems is either not feasible or impractical. *See* STOCHASTIC PROCESS.

3. Simulation can account for the effects of variance. Multiple runs of the model can be made under different conditions, and not only the average outcomes but also the variability of outcomes can be examined.

4. Simulation analysis promotes what is known as total solutions. A simulation model is designed and built to take into account all aspects that directly impact the system; indirect impacts exist in the system environment. Since the total system is reflected in the model, any outcomes suggested by the analysis of the model represent total solutions to the system behavior.

5. Simulation can bring together expertise, knowledge, and information about a system and its operations. The process of model development should incorporate any outside human expertise and knowledge about the system. In addition, information on the current status of inputs and processes is incorporated directly into the model and its analysis. Finally, simulation can be a very cost effective tool, especially in the design of new systems. Many different scenarios can be examined and analyzed. Thus, even new and creative ideas for system modifications can be examined without harm to the current operations.

Simulation is essentially a four-step process to formulate (build) a valid model of a real world system, develop operational scenarios of interest (from the real world) that exercise the model, analyze the output of these scenarios, and translate and project the model outputs back to the real-world system.

Systems. The word "system" refers to a set of elements (objects) so interconnected (interdependent, interrelated) as to aid in driving toward a desired goal. This definition has two connotations: First, a

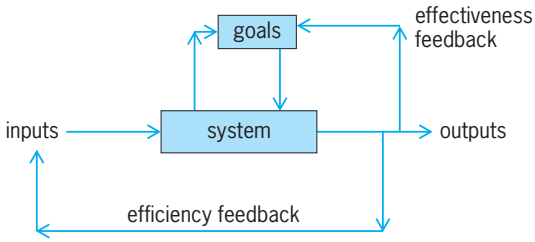

System feedback.

system is made of parts (elements) that have relationships between them (or processes that link them together). These relationships or processes can range from relatively simple to extremely complex. One of the necessary requirements for creating a "valid" model of a system is to capture, in as much detail as possible, the nature of these interrelationships. Second, a system constantly seeks to be improved. Feedback (output) from the system (*see* **illus.**) must be used to measure the performance of the system against its desired goal. Both of these elements are important in simulation. *See* SYSTEMS ENGINEERING.

System environment and system state. A system is often affected by changes that occur outside the system. These changes are said to occur in the system environment. In the modeling of systems, it is important to decide on the boundary between the system and its environment. Changes in the system can be reflected (or modeled) by changes in the system state. The state of the system is defined to be the collection of variables necessary to describe the system at any time.

For example, in the study of a bank operation, the elements in the system include the tellers, the waiting area (how many people are allowed to wait at one time), and the services (processes) provided by the tellers to the customers. The environment of this system includes the town in which the bank is located. If it is a payday for a large company in town, the bank may see more customers on that particular day. The state of the bank system could be defined as the number of tellers who are working, the number of tellers who are currently serving customers, and the numbers of customers waiting. Each time one of these variables changes, the state of the bank system changes.

Classification of systems. Systems can be classified in three major ways. The classification of systems (*see* **table**) revolves around the types of elements that exist in the system, the nature of time and how the system state changes in relation to time, and whether or not the system changes over time at all. This categorization affects the type of modeling that is done and the types of simulation tools that are used.

Classification of systems and models

Classifiers	Systems
Elements	Deterministic versus stochastic
Time	Discrete-event versus continuous
Change	Static versus dynamic

A system that has elements or interrelationships betwen the elements known with complete certainty is called a deterministic system. In deterministic systems, every time a particular input is placed into the system a known output always results. Deterministic systems are easier to model than stochastic systems. Stochastic systems have at least one element or interrelationship that is determined by chance and therefore must be modeled with probabilities. The bank is an example of a system containing stochastic elements because the arrival of customers and the length of time to service a customer are both stochastic.

Time is an important element is many real-world systems. The nature of time in a system and how time affects the system state is an important classifier. A system that changes in a continuous manner is a continuous-time system. Differential calculus, specifically differential equations, are often used to describe the nature of the system state through a model. In the bank example, events, such as the arrival of a customer or the completion of service of a customer, happen at discrete instants in time; these systems are known as discrete-event systems.

The nature of change in the system also acts as a classifier. Systems that do not change over time are called static, whereas systems that do change over time are called dynamic. Most real-world systems are both dynamic and stochastic, and simulations models are usually developed for such systems. Different tools exist to model systems that are continuous-time or discrete-event.

Modeling. The purpose of any model is to enable its users to draw conclusions about the real system by studying and analyzing the model. The major reasons for developing a model, as opposed to analyzing the real system, include economics, unavailability of a "real" system, and the goal of achieving a deeper understanding of the relationships between the elements of the system.

The model should include only the elements of the system that are within the boundary of the system relative to the objectives of the study. While the system environment does have an effect on the model outcome, the environment does not have to be modeled explicitly.

Models of actual or theoretical physical systems can take many forms. Models, like the systems they represent, can be static or dynamic, discrete or continuous, and deterministic or stochastic. Simulation models are composed of mathematical and logical relations that are analyzed by numerical methods rather than analytical methods. Numerical methods employ computational procedures to run the model and generate an artificial history of the system. Observations from the model runs are collected, analyzed, and used to estimate the true system performance measures. *See* MODEL THEORY.

Simulation studies. There is no single prescribed methodology in which simulation studies are conducted. However, most simulation studies cover each

of the steps below to some level of detail depending on the study objectives. Most simulation studies proceed around four major areas: the problem, the system and model, the experiment, and the analysis.

Problem formulation. In the problem area, the major activity is in the formulation of the problem. In formulating the problem, a key step in any problem-solving endeavor, the analyst must strive to understand and define the problem in a precise and systematic form. The analyst must be able to state a clear understanding for the entities in the system and the interrelationships between the various entities. In this stage a clear and measurable set of objectives of the simulation analysis must be stated, and with those objectives a well-defined performance metric (or index) must be determined. In the problem definition, all assumptions as to the performance of the model must be clearly and explicitly stated. The scenarios that drive the simulation must be developed so that key inputs to the system (and therefore the model) are also stated. Finally, any restrictions on the system solutions (physical, political, economic, and social) must be clearly identified.

The output of this step is a detailed system description. In this description a baseline system is developed to show the movement of entities through the system; this can be done using a flow diagram or schematic depiction of the baseline system. The diagram should include entry points and exits to and from the system, any processing stations, and decision points where an entity might take a different path throughout the remainder of the system.

Once a baseline system is detailed, the analyst can start to design a new system diagram that would show possible changes to the baseline system that will be studied to improve overall system performance. The new system diagram can entail small changes to the baseline or significant deviations from the processing of entities from the baseline. This new system diagram is used to generate a list of potential questions that are to be answered with the simulation efforts. These questions must be answerable using the existing performance criteria (metrics that were defined in the earlier part of the problem formulation step).

The activities that are undertaken in this step help the simulation analyst to prepare for the next step in the problem analysis phase of the simulation study: defining and developing the model.

Model development. The development of a system model requires both science and art. A great detail of experience in developing models is required to obtain a valid and meaningful model of the system under study. The analyst must determine to what level of detail the system needs to be modeled and what are the essential features of the model, which are critical to the study objectives. In addition, the analyst must make an effort to involve the users of the model in the building process. If the users of the model are not satisfied that the model solves their problems, the modeling effort will be wasted.

The building of a system model is a top-down and often incremental process. The analyst will begin by developing a high-level model that captures the flow of entities through the system and leaves many of the finer details for future development. The analyst continues to refine the model in an effort to ensure that the model is a true representation of the system.

There is a constant interplay between the development of the model and the collection of needed input data. Before building the model (and as the model changes), the analyst must collect data from the real-world system and represent that data within the processes of the model. In a stochastic system, most of the real-world data that are collected are transformed and represented by probability distributions. Each of the processes within the model must attempt to mimic the real-world processes as closely as possible. The transformation of real-world data into a probability distribution is a well-documented statistical procedure, and the analyst must be able to use judgment and experiences to choose the right distribution for the situation and then follow the statistical procedure to ensure that the parameters of the distribution are accurate. *See* DISTRIBUTION (PROBABILITY); STATISTICS.

Model verification is used to check the correctness of the model functionality. In order to check the correctness, multiple runs of the simulation model are performed to ascertain that the correct outputs are generated for the particular inputs. Model validation is a much more difficult endeavor. It requires a check of the degree to which the model is representative of the real system; this usually requires an independent credibility check. Users and outside experts may be called on to review the structure of the model, the level of detail of the model, and the output of the model in comparison to the real-world system.

Another important activity in understanding the system and developing a valid model is the use of sensitivity analysis. Sensitivity analysis is a group of statistical techniques that help the analyst (and users) to understand the effect of input variables on the output. Sensitivity analysis is undertaken to examine the input variables and ascertain which variables have a high degree of "sensitivity." It is important in the understanding of the system to assess how small changes in the input of the system can have large effects on the output. Sensitivity analysis is extremely important in the design of the simulation experiments that are undertaken with the model. It can also lead to further refinement (more detail) in the data collection process or the study of a particular input parameter. *See* SENSITIVITY (ENGINEERING).

Running the model. Simulation is an experimental science, so care must be taken to design experiments with the simulation model to obtain results that address the objectives of the study. Simulation is not designed to find the "best" or optimal answer to the problem at hand, but the nature of simulation makes it possible to compare many different alternatives. Solutions found using simulation are only as good as the alternatives being investigated. In other words, simulation is not a discovery science that allows a

creative and innovative solution to be uncovered unless that particular solution is being tested directly.

In designing simulation experiments, it is largely the job of the analyst to determine the proper inputs that need to be changed in order to meet the study's objectives. By controlling the change in these inputs over multiple runs of the model, the analyst can minimize the effect of variability that is introduced to the output of the mode. This in turn makes the results of the study more meaningful.

Many systems that simulation deals with are both dynamic and stochastic. Since a simulation model can be viewed as an input-output transformation, and since many of the inputs are random variables, the outputs of a simulation model are also random. Since these outputs are random, a single run of the model is insufficient in making a judgment of the result of the model. Multiple runs (or replications) of a stochastic model must be made, and statistical inference techniques on these replications must be used to make credible recommendations for system changes.

Analyzing the model output. Output analysis is the analysis of data generated by a simulation. Output analysis is used to predict system performance or compare the performance of two or more system designs. In the problem-formulation stage of the study, the analyst developed a problem description that included the specific metrics that will be used to measure the performance of the system. These metrics will differ from study to study depending on the objectives of the study. However, regardless of the metric to be used, each run of the simulation will yield a single value for the metric and this value will differ from run to run. The job of the analyst is to determine the long-term behavior of the system; this is accomplished by taking multiple runs of the simulation model and averaging these runs.

When analyzing simulation output data, a distinction is made with regard to the type of system being modeled. A terminating (or transient) system is one that starts at time 0, under a prespecified set of initial conditions, and proceeds until some specified time, at which time the system "closes." The objectives and the metrics for a terminating system may differ from that of a nonterminating or steady-state simulation. In analyzing data from a steady-state simulation, the goal is to make a valid statistical inference about the long-term average behavior of the system based on the sample average from a specified number of replicate simulation runs. The tool most often used to make this determination is confidence intervals. A confidence interval is an interval estimate for a parameter that specifies a level of confidence. A confidence interval provides a way of quantifying imprecision. The goal of a confidence interval procedure is to form an interval, with end points determined by the sample, that will contain, or "cover," the target parameter with a prespecified (high) probability called the confidence level.

Statistical inference methods also allow the comparison of various competing system designs or alternatives. For example, estimation and hypothesis testing make it possible to discuss the outputs of the simulation and compare the system metrics.

Applications. Simulation is a versatile tool that can be applied to many different systems operating (or proposed) in society. Many of the applications of simulation are in the area of manufacturing and material handling systems. Manufacturing systems are the focus of many simulation textbooks, and simulation is taught in many engineering and business curricula with the focus of the applications also being on manufacturing systems. The characteristics of these systems, such as physical layout, labor and resource utilization, equipment usage, products, and supplies, are extremely amenable to the simulation modeling methods described above. *See* COMPUTER-INTEGRATED MANUFACTURING; FLEXIBLE MANUFACTURING SYSTEM.

There are many other areas where simulation is prominent. Transportation systems, service systems (such as hospitals and banks), communication and network systems, environmental systems, business process reengineering, and social and economic systems are all candidate fields for the use of simulation.

Simulation began in the late 1950s and 1960s as a very expensive and specialized tool that was generally used by large corporations, primarily steel and aerospace companies. With the advent of personal computers and the relative richness of current simulation software packages, simulation has become a mature tool that many smaller firms can use with little capital investment. *See* COMPUTER-AIDED ENGINEERING; MATHEMATICAL SOFTWARE; MICROCOMPUTER; SOFTWARE.

Current applications of simulation are concerned primarily with the design or redesign of complex systems. In the future, simulation will concentrate its efforts on the actual control of the system. This requires that models be kept current, but also allows for easy testing of the new systems over time. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING; CONTROL SYSTEMS; ENGINEERING DESIGN.

James Pomykalski

Bibliography. J. Banks et al., *Discrete-Event System Simulation*, 3d ed., Prentice Hall, Upper Saddle River, NJ, 2000; T. J. Gogg and J. R. Mott, *Improve Quality and Productivity with Simulation*, 3d ed., Consulting Group, Palos Verdes, CA, JMI, 1996; B. Hannon and R. Matthias, *Dynamic Modeling*, 2d ed., Springer-Verlag, New York, 2000; A. M. Law and W. D. Kelton, *Simulation Modeling and Analysis*, 3d ed., McGraw-Hill, New York, 2000; W. D. Kelton, R. P. Sadowski, and D. A. Sadowski, *Simulation with Arena*, McGraw-Hill, Boston, 1998; A. B. Pritsker, *Introduction to Simulation and SLAM II*, 4th ed., Halsted Press, New York, 1995.

Sine wave

A wave having a form which, if plotted, would be the same as that of a trigonometric sine or cosine function. It generally results from the solution of a problem having a one-dimensional space coordinate, such as the transverse vibrations of a string, longitudinal vibrations of a bar, or the propagation of plane waves of electromagnetic radiation or sound. *See*

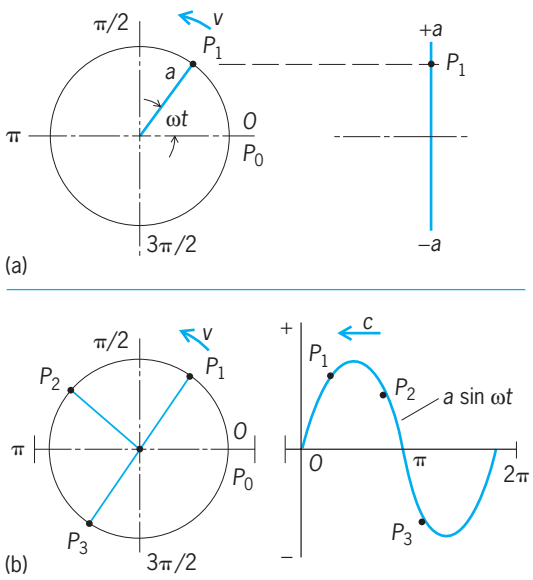

Simple harmonic motion and sine wave. (*a*) Point *P* moves around the circle at constant tangential speed *v*; its projection moves up and down between limits $+a$ and $-a$. (*b*) Same conditions as in preceding, except that the plane on which the projection is plotted is traveling to the left at speed *c*, thereby generating a sine wave. Angular speed is ω rad/s; the total angle traced out from zero position equals ωt rad.

ELECTROMAGNETIC RADIATION; SOUND; VIBRATION; WAVE EQUATION; WAVE MOTION; WAVE MOTION IN FLUIDS.

The sine wave may be thought of as the projection on a plane of the path of a point moving around a circle at uniform speed. For example, in **illus.** *a*, assume that the moving point travels around the circle at constant speed v. The projection onto a plane would trace back and forth on the line indicated as the point went around the circle. If the plane is now moved to the left at constant speed c, as in illus. *b*, the resulting trace has the form of the graph of the sine function.

The sine wave trace of illus. *b* describes a body moving in simple harmonic motion. This motion changes in magnitude and time so that it repeats itself exactly, as long as uniform speed is maintained. It is characteristic of one-dimensional vibrations and one-dimensional waves having no dissipation. *See* HARMONIC MOTION.

The sine wave is the basic function employed in harmonic analysis. It can be shown that any complex motion in a one-dimensional system can be described as the superposition of sine waves having certain amplitude and phase relationships. The technique for determining these relationships is known as Fourier analysis. *See* FOURIER SERIES AND TRANSFORMS; WAVEFORM.

William J. Galloway

Single crystal

In crystalline solids the atoms or molecules are stacked in a regular manner, forming a three-dimensional pattern which may be obtained by a three-dimensional repetition of a certain pattern unit called a unit cell. When the periodicity of the pattern extends throughout a certain piece of material, one speaks of a single crystal. A single crystal is formed by the growth of a crystal nucleus without secondary nucleation or impingement on other crystals. *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Growth techniques. Among the most common methods of growing single crystals are those of P. Bridgman and J. Czochralski. In the Bridgman method the material is melted in a vertical cylindrical vessel which tapers conically to a point at the bottom. The vessel then is lowered slowly into a cold zone. Crystallization begins in the tip and continues usually by growth from the first formed nucleus. In the Czochralski method a small single crystal (seed) is introduced into the surface of the melt and then drawn slowly upward into a cold zone. Single crystals of ultrahigh purity have been grown by zone melting. Single crystals are also often grown by bathing a seed with a supersaturated solution, the supersaturation being kept lower than is necessary for sensible nucleation.

When grown from a melt, single crystals usually take the form of their container. Crystals grown from solution (gas, liquid, or solid) often have a well-defined form which reflects the symmetry of the unit cell. For example, rock salt or ammonium chloride crystals often grow from solutions in the form of cubes with faces parallel to the (100) planes of the crystal, or in the form of octahedrons with faces parallel to the (111) planes. The growth form of crystals is usually dictated by kinetic factors and does not correspond necessarily to the equilibrium form. *See* CRYSTAL GROWTH; CRYSTALLIZATION; ZONE REFINING.

Physical properties. Ideally, single crystals are free from internal boundaries. They give rise to a characteristic x-ray diffraction pattern. For example, the Laue pattern of a single crystal consists of a single characteristic set of sharp intensity maxima. *See* X-RAY DIFFRACTION.

Many types of single crystal exhibit anisotropy, that is, a variation of some of their physical properties according to the direction along which they are measured. For example, the electrical resistivity of a randomly oriented aggregate of graphite crystallites is the same in all directions. The resistivity of a graphite single crystal is different, however, when measured along different crystal axes. This anisotropy exists both for structure-sensitive properties, which are strongly affected by crystal imperfections (such as cleavage and crystal growth rate), and for structure-insensitive properties, which are not affected by imperfections (such as elastic coefficients).

Anisotropy of a structure-insensitive property is described by a characteristic set of coefficients which can be combined to give the macroscopic property along any particular direction in the crystal. The number of necessary coefficients can often be reduced substantially by consideration of the crystal symmetry; whether anisotropy, with respect to a given property, exists depends on crystal symmetry.

The structure-sensitive properties of crystals (for example, strength and diffusion coefficients) seem

governed by internal defects, often on an atomic scale. *See* CRYSTAL DEFECTS. David Turnbull

Bibliography. I. V. Markov, *Crystal Growth for Beginners*, 1995; B. R. Pamplin, *Crystal Growth*, 2d ed., 1980; J. P. Van der Eerden, *Fundamentals of Crystal Growth*, 1993.

Single sideband

An electronic signal-processing technique in which a spectrum of intelligence is translated from a zero reference frequency to a higher frequency without a change of frequency relationships within the translated spectrum. Single-sideband (SSB) signals have no appreciable carrier. After translation, only the single-sideband energy remains. This form of intelligence transmission requires distortion-free amplification of the SSB signal prior to transmission and occupies only the spectrum bandwidth of that intelligence. These advantages allow SSB to be selected for communication of voice and digital data wherever there is a premium on having a minimum of transmitted power and available frequency spectrum.

Amplitude-modulated (AM) signals have identical upper and lower sidebands symmetrically located on each side of the translation frequency, which is often called the carrier. The SSB spectrum differs from the AM spectrum in having little or no carrier and only one sideband. (A double-sideband signal is one in which only the carrier is suppressed, leaving both sidebands. This technique is not commonly used.) *See* AMPLITUDE MODULATION.

Generation. In the SSB signal-processing action, the intelligence spectrum to be translated is applied to the signal input port of a balanced modulator. A higher-frequency sinusoidal signal, often called a carrier, is applied to the other input port of this circuit. Its function is to translate the zero reference spectrum to the carrier frequency and to produce the upper and lower sidebands, which are symmetrically located on each side of the carrier. The carrier frequency power is suppressed to a negligible value by the balanced operation of the modulator and does not appear at the output. Generally, the balanced modulator operates at an intermediate frequency which is lower than the frequency of transmission. Following the balanced modulator is a sideband filter which is designed to remove the unwanted sideband signal power and to allow only the desired intelligence spectrum to pass. A block diagram and drawings of the signal characteristics at each major circuit interface are shown in the **illustration**. *See* AMPLITUDE MODULATOR; MODULATOR.

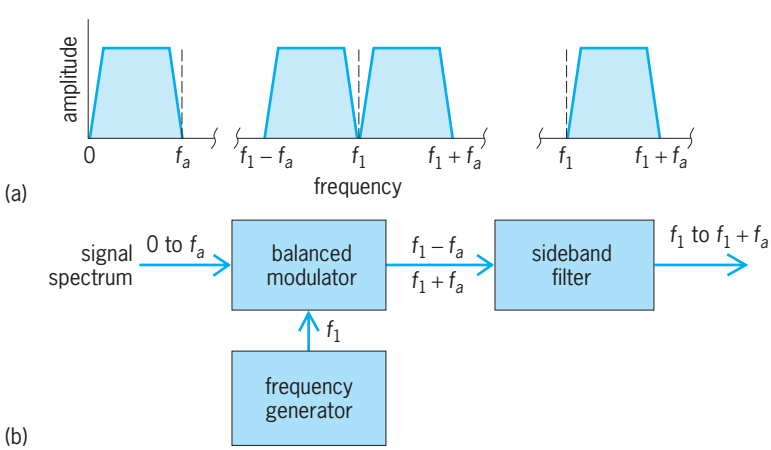

Generation of a single-sideband signal. (*a*) Spectrum. (*b*) Block diagram.

The outputs of several sideband filters having adjacent frequency passbands may be added to produce a succession of intelligence channel groups, each with differing information content. These may be translated to higher frequency and amplified for transmission; amplified and transmitted directly; or used as an input to a frequency, phase, or pulse code modulation system. *See* FREQUENCY MODULATION; MULTIPLEXING AND MULTIPLE ACCESS; PHASE MODULATION; PULSE MODULATION.

The SSB signal-handling processes are carefully designed to avoid distortion of the frequency relationships within the sideband spectrum. Amplifier circuits and components are chosen to provide a highly linear signal-handling capability. Linearity is measured by determining the level of the third-order tone and the fifth-order tone produced when two tones (f_1, f_2) occupy the intelligence channel. Typically linearity requirements allow a third-order tone level less than 1/1000 (30 decibels below) the desired tone's power level and a fifth tone almost 1/1,000,000 (60 decibels below) the desired tone's power level. Filters are designed for equal amplitude response and uniform time delay of the frequency components in the signal passband.

A technique is under development and in early commercial use which compresses both the frequency range and the amplitude range of the intelligence spectrum before translation. This technique is referred to as amplitude compandered sideband (ACSB). ACSB processing improves the signal-to-noise performance for a given transmitted power and furthermore decreases the required transmission bandwidth.

Reception. Reception of SSB signals is essentially a reverse process from signal generation. The signal energy derived from the antenna is selected for a desired frequency range by radio-frequency filtering, is amplified to a desired level, and is translated down to an intermediate frequency. The signal-processing circuitry required to derive the desired intelligence channel spectrum from the intermediate frequency requires a demodulator which is similar to the modulator of the SSB generator. ACSB reception requires an inverse signal processing of the intelligence spectrum to that used in transmission in which the original frequency and amplitude range are precisely restored.

Frequency stability requirements. Frequency stability of a high order is required of the signal generators that provide the translation frequencies in both SSB generation and reception. In the precise recovery of the intelligence spectrum there is a direct dependence upon total system frequency

accuracy in returning the received intelligence sideband to zero reference frequency. Voice communication intelligibility suffers little with a system frequency error of up to 50 Hz, although digital data may be disrupted severely with only a few hertz error. Many SSB systems use a frequency synthesizer which makes all translation frequencies dependent upon a single piezoelectric frequency generator which can be designed for frequency accuracy on the order of 1 part in 10^9 under severe environmental conditions. Atomic standards for frequency generation can provide reference frequency accuracy on the order of 1 part in 10^{12}.

The Doppler shift, caused by motion of high-velocity vehicles carrying SSB equipment, for example, supersonic aircraft or satellites, necessitates transmitting some reference frequency power (suppressed carrier) to provide at the receiver a source for frequency or phase locking of the frequency standard. In ACSB transmission and reception there is need for a high relative stability of the reference (carrier) frequency. Thus a pilot tone is transmitted, and upon reception it provides a stable frequency reference to which the translator frequency is phase-locked. *See* AUTOMATIC FREQUENCY CONTROL (AFC); DOPPLER EFFECT.

Communications. There are many advantages in the use of SSB techniques for communication systems. The two primary advantages are the reduction of transmission bandwidth and transmission power. The bandwidth required is not greater than the intelligence bandwidth and is one-half that used by amplitude modulation. The output power required to give equal energy in the intelligence bandwidth is one-sixth that of amplitude modulation.

Propagation of radio energy via ionospheric refraction and when the transmitted signal energy is reflected from terrain objects or from buildings (as in land mobile communications) provides the possibility for multiple paths of differing path length which cause a selective cancellation of frequency components at regular frequency spacings. This produces in AM a severe distortion of the intelligence because of the critically dependent carrier-to-sideband amplitude and phase relationships. Frequency-modulated (FM) multipath signals in conjunction with the motion of a vehicular receiver cause a rapid fading to occur so that when the median received signal is near the required noise quieting threshold level the receiver will emit bursts of noise.

SSB is much less affected under these conditions. SSB does not have an abrupt noise quieting threshold as does FM, since with SSB there is a nearly linear increase of noise as the signal level decreases. SSB has proved to be a highly satisfactory communication signaling method where signal-to-noise ratio levels of under 10:1 are present, and SSB can produce intelligible voice communication at less than unity signal-to-noise ratio.

SSB and its close relative ACSB are inherently less affected by interference than are either AM or FM due to the narrow bandwidth occupied by SSB/ACSB, and since the reference frequency is adjusted for clarity of the desired signal, the undesired signal appears noiselike. Thus in SSB the signal-to-noise ratio of the desired signal is reduced by interference. In FM the interfering signal either can capture the demodulator when the interference is stronger than the desired signal or can severely distort the desired signal when the interference is less than the desired signal. In AM a weaker interfering signal will appear as an intelligible signal at a lower level than the desired signal. SSB techniques are finding wide use in wire line, coaxial cable, long-range high-frequency communications, microwave multiplex, ground-to-air–air-to-ground voice communication and land mobile (ASCB) communications, and are being used in experimental satellite-to-Earth channels. High-channel-density microwave systems which formerly employed FM multiplex techniques are now being installed using SSB-modulation techniques to achieve more than twice the number of channels in the same spectrum bandwidth. SSB microwave has equal or improved performance when compared to that of FM multiplexed microwave. *See* RADIO-WAVE PROPAGATION.

History. The discovery of SSB can be credited to John R. Carson in 1915. He found from the mathematical equations that there was a preservation of the premodulation spectrum in the remaining sideband after the removal of the lower sideband and the carrier of an AM signal. He experimentally proved this by using the filtering capability of a frequency-selective antenna. His Patent No. 1,449,382, filed in 1915 and granted in 1923, recognized this. Transoceanic experiments were also in progress by 1923, and were soon followed by commercial radio telephone channels. Wider utility followed developments in stable frequency generators and improved filter technology and amplifier components.

David M. Hodgin

Bibliography. American Radio Relay League, *Radio Amateur's Handbook*, revised periodically; American Radio Relay League, *Single Sideband for the Radio Amateur*, revised periodically; J. M. Gale, *Marine SSB Operation*, 1990; National Association of Broadcasters Staff, *NAB Engineering Handbook*, 1985; W. E. Sabin and E. O. Schoenike, *Single Sideband Systems and Circuits*, 1995; M. Schwartz, *Information Transmission, Modulation, and Noise*, 4th ed., 1990.

Single-site catalysts (polymer science)

Metallocenes and related compounds which have a single catalytic site and are used to produce uniform polymers and copolymers with tailored structures and physical properties. Metallocenes are "sandwich compounds" in which a π-bonded metal atom is situated between two aromatic ring systems (**Fig. 1**). *See* CATALYSIS; METALLOCENES; POLYMER.

Classical catalysts. In 1953, Karl Ziegler succeeded in polymerizing ethylene into high-density polyethylene at standard pressure and room temperature, based on the use of the catalyst titanium

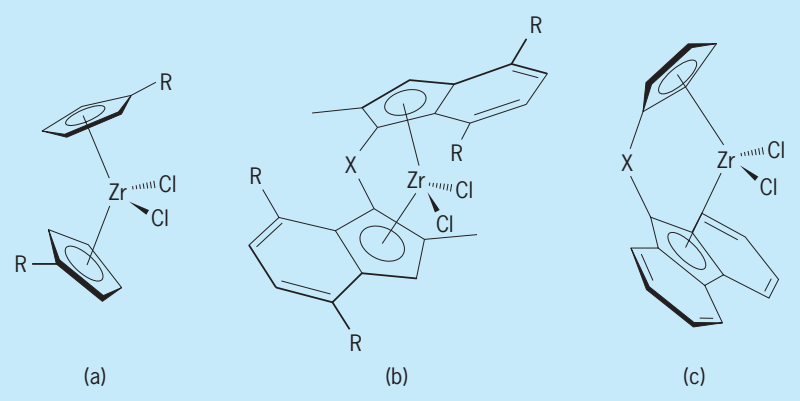

Fig. 1. Structures of metallocenes (in this case, zirconocenes). (*a*) Nonchiral structure produces atactic polypropylene. (*b*) Racemic mixture of this chiral structure produces isotactic polypropylene. (*c*) C_s-symmetric structure produces syndiotactic polypropylene. R = alkyl substitution [H, CH_3, C_2H_5, C_6H_5]; X = bridge [—CH_2—CH_2—, $(CH_3)_2Si$, $(CH_3)_2C$].

tetrachloride combined with diethylaluminum chloride. Some months later, Giulio Natta demonstrated a catalyst system that was capable of polymerizing propylene into semicrystalline polypropylene. For their work, Ziegler and Natta shared the Nobel Prize for Chemistry in 1963. Modern Ziegler-Natta catalysts are mixtures of solid and liquid compounds, often containing $MgCl_2/TiCl_4/Al(C_2H_5)_3$ and, for propylene polymerization, different internal and external donors such as ethylbenzoate, silanes, or ethers to increase stereoregularity (tacticity). An important catalyst for ethylene polymerization is the Phillips catalyst, that is, chromium trioxide on a silica support. This catalyst is used in the gas-phase polymerization of ethylene, but is unable to produce polypropylene with methyl groups always on the same side of the chain (isotactic). As Ziegler-Natta catalysts are heterogeneous and complex systems with different active sites, the resulting polymer structure can be controlled only to a limited degree. *See* HETEROGENEOUS CATALYSIS; POLYMERIZATION; POLYMER STEREOCHEMISTRY AND PROPERTIES; POLYOLEFIN RESINS.

Metallocene catalysts. In contrast to Ziegler systems, metallocene catalysts have only one active site (single-site catalysts) and produce polymers with a narrow molecular weight distribution $M_w/M_n = 2$ (M_w and M_n are the weight average and number average molecular weights). Their catalytic activity is 10–100 times higher than that of the classical Ziegler-Natta systems. In addition, metallocene catalysts are soluble in hydrocarbons or liquid propylene, and their structure can be easily changed. These properties allow one to predict accurately the properties of the resulting polyolefins by knowing the structure of the catalyst used during their manufacture, as well as to control the resulting molecular weight and distribution, comonomer content, and tacticity by careful selection of the appropriate reactor conditions. *See* MOLECULAR WEIGHT.

Metallocenes, in combination with the conventional aluminumalkyl cocatalysts used in Ziegler systems, are capable of polymerizing ethylene, but only at a very low activity. In 1977 in Hamburg, it was discovered that methylalumoxane (MAO) greatly improved catalysis with metallocenes, enhancing their activity, surprisingly, by a factor of 10,000. Methylaluminoxane

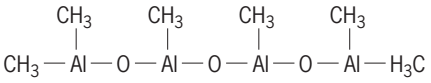

is a compound in which aluminum and oxygen atoms are arranged alternately and free valences are saturated by methyl substituents. It is prepared by the careful partial hydrolysis of trimethylaluminum and consists mainly of units of the basic structure $Al_4O_3Me_6$. As the aluminum atoms in this structure are coordinatively unsaturated, the basic units join, forming cluster compounds. These have molecular weights from 1200 to 1600 and are soluble in hydrocarbons. Metallocene catalysts, especially zirconocenes, treated with MAO can polymerize up to 100 tons of ethylene per gram of zirconium (Fig. 1). Such highly active catalysts can remain in the product. The insertion time (for the insertion of one molecule of ethylene into the growing chain) is only 10^{-5} s. A comparison to enzymes is not far-fetched. It is generally assumed that the function of MAO is to first undergo a fast ligand exchange reaction with the metallocene dichloride, thus producing the metallocene methyl and dimethylaluminum compounds. In the next step, either Cl^- or CH_3^- is abstracted from the metallocene compound by an aluminum center in MAO, thus forming a metallocene cation and an MAO anion. The alkylated metallocene cation represents the active center. Meanwhile, other weakly coordinating cocatalysts, such as tetra(perfluorophenyl)borate anions, $[(C_6F_5)_4B]^-$, have been successfully used to activate metallocenes and other single-site catalysts. Polyolefins with different microstructures and characteristics can be custom-made by varying the ligands of the metallocene (Fig. 1). By combining different olefins and cycloolefins with one another, the range of characteristics can be further broadened. Conventional heterogeneous catalysts have not yet produced polyolefins with tailored microstructures or chemically uniform copolymers. Using metallocene catalysts, it was possible for the first time to produce polyethylenes, polypropylenes, and copolymers with narrow molecular-weight distribution, syndiotactic (alternating stereoregular structure) polypropylene, syndiotactic polystyrene, cyclopolymers of 1,5-hexadiene, and cycloolefin copolymers (COCs) with high catalytic activity. *See* COPOLYMER; POLYSTYRENE RESIN.

Other single-site catalysts. "Half sandwich" metallocenes are highly active for the copolymerization of ethylene and 1-octene (**Fig. 2**). These catalysts produce long-chain branched polyolefins by incorporating oligomers with vinyl endgroups, which are formed during polymerization by β-hydrogen transfer. Long-branched copolymers of ethylene with 1-octene show elastic properties as long as the

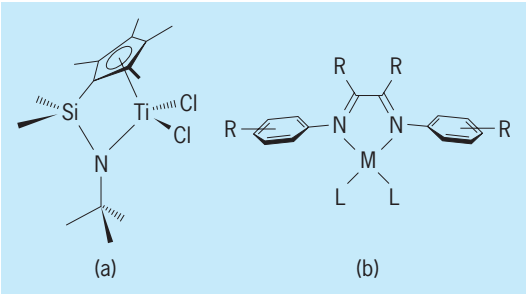

Fig. 2. Structures of (*a*) "half-sandwich" metallocenes, dimethylsilylamido(cyclopentadienyl)titanium compound for the production of linear low-density polyethylene (LLDPE), and (*b*) diimin complexes with M = Ni, Pd; R = H, alkyl; L = halogen.

comonomer content is more than 20%. Other elastomers with different microstructures can be synthesized from dienes.

New palladium, nickel, iron, and cobalt complexes activated by MAO or borate cocatalysts are also active for the polymerization of olefins. Cationic Pd(II) and Ni(II) complexes are able to polymerize ethylene and other α-olefins when sterically bulky α-diimine ligands and noncoordinating counterions are used. The lower oxophilicity and presumed greater functional group tolerance of late-transition metals relative to early-transition metals make them likely targets for the development of catalysts for the copolymerization of ethylene with polar monomers under mild conditions.

Polypropylene. Propylene and long-chain olefins can be assembled in a stereoregular manner, such that the building blocks have a defined and recurring arrangement in the resulting polymer chain. The structural arrangement of polypropylene has a considerable influence on its functional properties. A random arrangement leads to amorphous (atactic) polypropylene which flows at room temperature, whereas stereoregular polypropylene (isotactic and syndiotactic) is crystalline having a melting point of 165°C (330°F). In addition to isotactic and atactic polypropylene, syndiotactic polypropylene as well as isoblock and stereoblock polypropylene have been made for the first time in large quantities and in high purity (**Fig. 3**). The block lengths of equal tacticity in isoblock polypropylene (meaning the position of the methyl groups in one direction) can be varied in a broad range on the nanoscale (4–100 units). This leads to microcrystalline polypropylene which is suitable for making transparent films. The development of applications for elastic stereoblock polypropylene has just begun. Polypropylenes made using metallocenes exhibit distinct differences to conventionally produced polypropylenes, such as narrow molar mass distributions, higher stiffness, and greater tensile strength (see **table**). This is caused not only by the uniform structure but also by the extremely low fractions of low-molecular-weight (oligomeric) products. These fractions amount to less than 0.1%, compared to 2–4% in Ziegler-Natta polypropylene.

Elastomers. One of the biggest impacts of single-site catalysts promises to be in the manufacture of elastomers with elastic properties similar to rubber. The copolymers of ethylene and propylene with a molar ratio of 1:0.5 up to 1:2 are of great industrial interest. These ethylene-propylene copolymers (EPM) show elastic properties. In addition, ethylene-propylene-diene terpolymers (EPDM), made by including 2–5 wt % of a diene as a third monomer, are elastomers. The diene, 5-ethylidene-2-norbornene (ENB) [norbornene = bicyclo[2.2.1]-hept-2-ene], is mainly used. Since there are no double bonds in the backbone of this polymer, it is stable to oxidation and photodegradation reactions. *See* RUBBER.

Until recently, most processes for the production of EPM and EPDM rubber used soluble or highly dispersed vanadium catalysts. The elastomers produced by vanadium catalysts are pale yellow in color because of traces of remaining vanadium catalysts. Similar, less colored elastomers can be obtained with the single-site catalysts because the concentration of the single-site catalysts used for the polymerization is about a factor of 100 lower than the vanadium catalysts.

Cycloolefin copolymers. Single-site catalysts can be used to polymerize and copolymerize strained cyclic olefins such as cyclopentene or norbornene. The melting points for these homopolymers are extremely high, such as 400°C (750°F) for

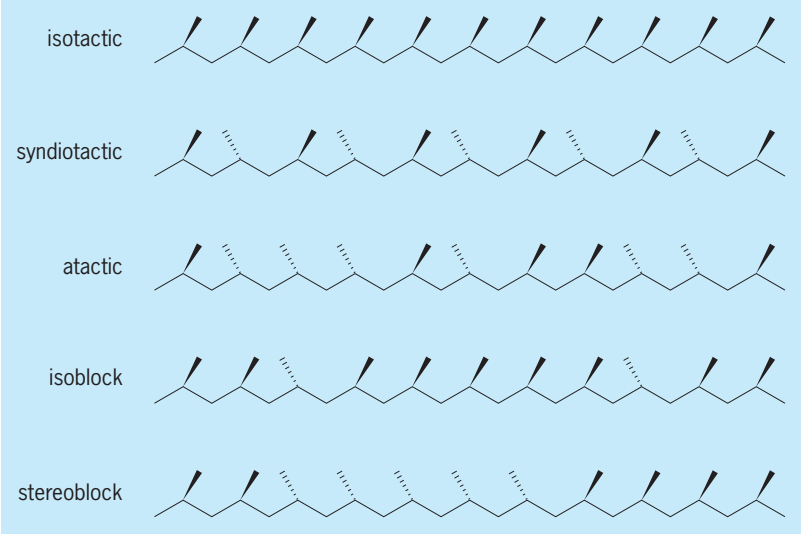

Fig. 3. Microstructures of polypropylenes.

Comparison of some properties of polypropylene made by Ziegler-Natta and metallocene catalysts

	Polypropylene from metallocene catalyst	Polypropylene from Zigler-Natta catalyst
Melting point (°C)	161	162
M_w/M_n	2.5	5.8
Stress tensile (N mm^{-2})	1620	1190
Hardness (N mm^{-2})	86	76
Extractable fraction (%)	0.1	2–4

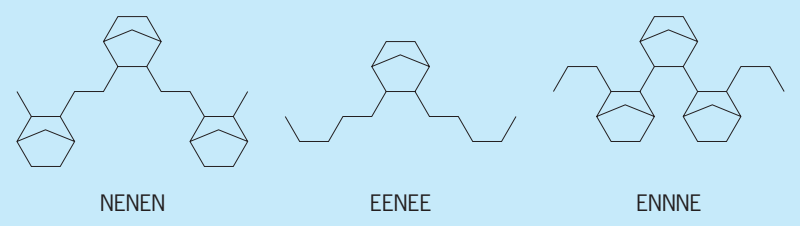

Fig. 4. Structures of norbornene (N)/ethylene (E) copolymers. Alternating blocks can also occur.

polycyclopentene, and as a result they may be used as engineering plastics. Copolymers of cyclic olefins with ethylene (COC) represent a new class of thermoplastic amorphous materials (**Fig. 4**). COCs are characterized by excellent transparency and very long life at high service temperatures. In addition they are chemically resistant and can be melt-processed. Due to their high carbon/hydrogen ratio, these polymers feature a high refractive index, for example, 1.53 for the ethene-norbornene copolymer at 50 mol % norbornene incorporation. Their stability against hydrolysis and chemical degradation, in combination with their stiffness, makes them desirable materials for optical applications such as compact discs, lenses, optical fibers, and films. W. Kaminsky

Bibliography. S. D. Ittel, L. K. Johnson, and M. Brookhart, Late-metal catalysts for ethylene homo- and copolymerization, *Chem. Rev.*, 100:1169–1203, 2000; W. Kaminsky, Highly active metallocene catalysts for olefin polymerization, *J. Chem. Soc., Dalton Trans.*, pp. 1413–1418, 1998; J. Scheirs and W. Kaminsky, *Metallocene Based Polyolefins*, vols. 1 and 2, Wiley, Chichester, 1999; K. B. Sinclair, Future trends in polyolefin materials, *Macromol. Symp.*, 173:237–261, 2001.

Sintering

The welding together and growth of contact area between two or more initially distinct particles at temperatures below the melting point, but above one-half of the melting point in kelvins. Since the sintering rate is greater with smaller than with larger particles, the process is most important with powders, as in powder metallurgy and in firing of ceramic oxides.

Powder. Although sintering does occur in loose powders, it is greatly enhanced by compacting the powder, and most commercial sintering is done on compacts. Compacting is generally done at room temperature, and the resulting compact is subsequently sintered at elevated temperature without application of pressure. For special applications, the powders may be compacted at elevated temperatures and therefore simultaneously pressed and sintered. This is called hot pressing or sintering under pressure.

Sintering is observed as an increase in mechanical properties (strength, ductility, and so on) and in many physical properties (for example, electrical and thermal conductivity). In many, but by no means all, sintering processes, the density of the compact increases, that is, the dimensions of the compacts become smaller (the compacts "shrink"), but the shape of the compact is generally preserved. The final density of the sintered compact depends upon the pressure with which the powder is compacted—the higher the pressure, the greater the density—and upon the shrinkage of the compact during sintering. Compacts from a single component powder, for example, a powder of a pure metal or a pure oxide, must be sintered below the melting point of the component, that is, without a liquid phase. Certain compacts from a mixture of different component powders may be sintered under conditions where a limited amount of liquid, generally less than 25 vol%, is formed at the sintering temperature. This is called liquid-phase sintering, important in certain powder-metallurgy and ceramic applications. *See* CERAMICS.

Mechanism. The driving force in sintering is surface energy, which decreases because the total surface area decreases as sintering proceeds. When two spheres originally in tangential contact sinter, the area of contact between the spheres increases and a neck is formed. In amorphous materials such as glasses, the process causing neck formation is viscous flow (as when two drops of water coalesce to form one). This type of viscous flow is not possible in crystalline solids. Here the most important material transport process is self-diffusion. Because the material in the neck surface has a highly convex curvature, the number of defects in its crystal structure (vacant lattice sites) is considerably higher than on a flat or concave surface. These defects move by self-diffusion from the convex neck surface to the adjacent flat surface, which means material moves in the opposite direction, that is, the neck grows (**Fig. 1***a*). Instead of coming from the adjacent flat surface, the material forming the neck may also move by self-diffusion from the grain boundary between the two spheres to the convex neck surface (Fig. 1*b*). This latter type of movement explains why compacts shrink, because in this case the centers of the spheres approach each other during sintering.

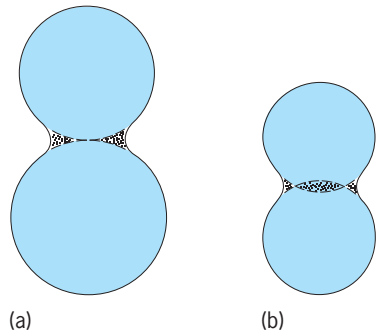

Fig. 1. Neck formation. (*a*) Neck growth through movement of defects from the neck surface to the adjacent flat surface. (*b*) Movement of neck-forming material from the grain boundary to the neck surface. (*After G. H. Gessinger et al., Continuous observation of the sintering of silver particles in the electron microscope, ASM Trans. Quart., 61(3):598–604, 1968*)

Densification (shrinkage of compacts) is inhibited by rapid grain growth during sintering. During grain growth the grain boundaries, which are the sources of the material which is transported into the convex neck surface, that is, into the pores of the compact, are swept out of the compact.

Liquid. The mechanism of liquid-phase sintering is more complicated. It involves viscous flow of the liquid phase, but also solution of the solid phase in the liquid phase and its reprecipitation in such a way as to make the compact more dense. As in solid-phase sintering, the driving force in liquid-phase sintering is a decrease in surface energy. *See* CERMET; POWDER METALLURGY. F. V. Lenel

Boundary migration. Powder compacts for sintering are usually made of mixtures of elemental particles, which form alloys as composition change occurs during the sintering treatment. When solute atoms diffuse into the grain boundaries, either present in the original particles or formed in the early stage of the sintering treatment, the grain boundaries may migrate, forming alloyed regions behind them. Such grain boundary migration has been observed in a number of metallic and ceramic materials, and is usually called diffusion-induced grain boundary migration (DIGM) [**Fig. 2***a*]. Solute atoms are supplied from vapor, melt, or solid. During liquid-phase sintering, thin liquid films form between the grains and may migrate if the grains are not chemically in equilibrium with the liquid (Fig. 2*b*). This phenomenon, called chemically induced liquid-film migration or simply liquid-film migration (LFM), shows characteristics similar to DIGM. Liquid-film migration can be induced either by adding solute atoms to the liquid phase or by changing the temperature. If the amount of liquid in the grain-liquid alloy is large at high temperatures, the solid grains may become corrugated by reprecipitation on certain surface regions and dissolution of adjacent regions. When the composition change by solute addition is large, new grains often form and grow, resulting in diffusion-induced recrystallization (DIR) [Fig. 2*c*].

Mechanism. Both DIGM and LFM occur when there is depletion of solute atoms from an alloy, such as by the evaporation of zinc from the surface of a polycrystalline iron-zinc alloy, and both occur at relatively low temperatures, where lattice diffusion of the solute atoms is not rapid. Often the boundaries move against their curvatures. As for any reaction to equilibrium, the driving force for DIGM (and LFM) arises from a decrease in the chemical free energy associated with the composition change. It has been proposed and verified in some experiments that solute atoms diffuse into the grains in front of the moving boundaries and produce coherency strains, which become the driving force for migration. The migration velocity then depends on the size and concentration change of the solute atoms. If the boundaries do not migrate, lattice diffusion is the dominant mode of composition change. The mechanism of forming new grains in DIR is not well understood, but once the grains form, their boundaries move by the same mechanism as for DIGM and LFM.

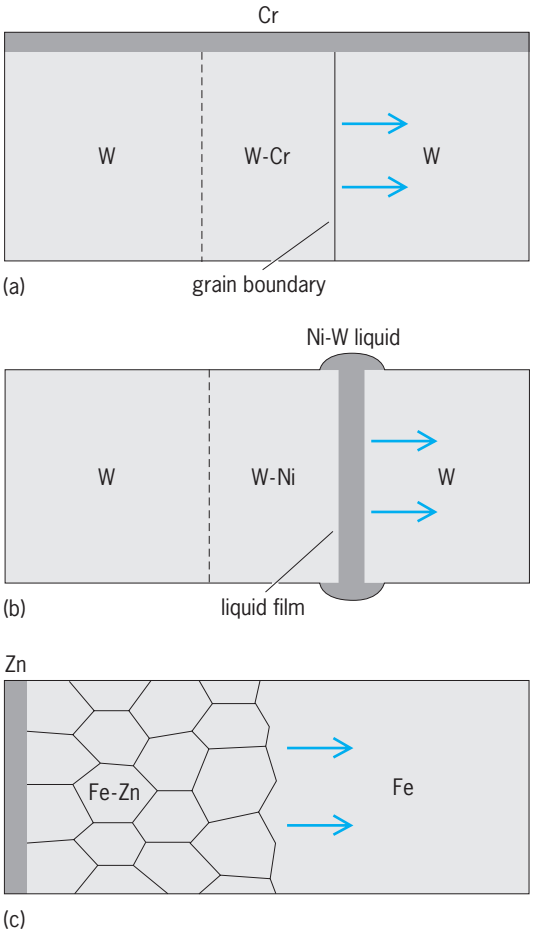

Fig. 2. Boundary migration. (*a*) Diffusion-induced grain boundary migration in tungsten-chromium (W-Cr). (*b*) Liquid-film migration in tungsten-nickel (W-Ni). (*c*) Diffusion-induced recrystallization in iron-zinc (Fe-Zn). The initial positions of the grain boundary and liquid film are shown by broken lines and the migration directions by arrows.

Applications. Both DIGM and LFM and possibly DIR are found in a variety of practical processes, such as heat treatment, welding, oxidation, sputtering, and melting. In the alloys prepared by liquid-phase sintering, rapid microstructural and compositional changes occur by DIGM and LFM during cooling from the sintering temperature. In electronic ceramics, the boundaries migrate during infiltration with a liquid phase. Such microstructural and composition changes produce significant effects on the resulting properties. The overall composition changes by DIGM, LFM, and DIR in polycrystals, including those produced by sintering, are much more rapid than that expected by diffusion only along stationary boundaries and in the lattice. *See* ALLOY; DIFFUSION. W. Gust; D. Y. Yoon

Bibliography. G. Flight, *Introduction to Ceramics*, 1990; W. D. Kingery, *Introduction to Ceramics*, 2d ed., 1976; B. Kubicki, *Sintered Machine Elements*, 1994; G. C. Kuczynski (ed.), *Sintering*, 1980; G. C. Kuczynski (ed.), *Sintering Processes*, 1980; E. Rabkin, C. Y. Ma, and W. Gust, Diffusion-induced grain boundary phenomena in metals and oxide ceramics, in J. Novotny (ed.), *Science of Ceramic*

Interfaces II, pp. 353–369, Elsevier, Amsterdam, 1994; L. H. Yaverbaum (ed.), *Technology of Metal Powders: Recent Developments*, 1980; D. Y. Yoon, Theories and observations of chemically induced interface migration, *Int. Metall. Rev.*, 40:149–179, 1995.

Sinus

Any space in an organ, tissue, or bone, but usually referring to the paranasal sinuses of the face. In humans, four such sinuses (see **illus.**), lined with ciliated, mucus-producing epithelium, communicate with each nasal passage through small apertures. The ethmoid and sphenoid sinuses are located centrally between and behind the eyes. The frontal sinuses lie above the nasal bridge, and the maxillary sinuses are contained in the upper jaw beneath the orbits. The mastoid portion of the temporal bone contains air cells lined with similar epithelium.

These sinuses vary in size and shape with the individual and with age. They are minute or absent in the infant, and usually develop to adult size as the skull bones enlarge after eruption of the permanent teeth.

Fluid or inflammatory products may accumulate in the sinuses, altering the voice resonance and producing headache or pain. Such accumulations are frequently the result of acute or chronic infections, or sinusitis. In this type of inflammation, abundant mucus is produced, which cannot drain into the nose because the narrow ducts are obstructed by the swollen mucous membrane. Pressure rises in the sinus, causing pain in the form of localized headaches. The localization varies according to the sinus involved. Extension of the infection to the neighboring structures becomes dangerous because of the close relationship to the brain. In chronic sinusitis, polyps can develop that protrude into the nose.

The membranes of the sinus of humans are well supplied with blood, lymphatics, and nerves, particularly nerves sensitive to pain and pressure. Erect posture decreases natural drainage of the human paranasal sinuses.

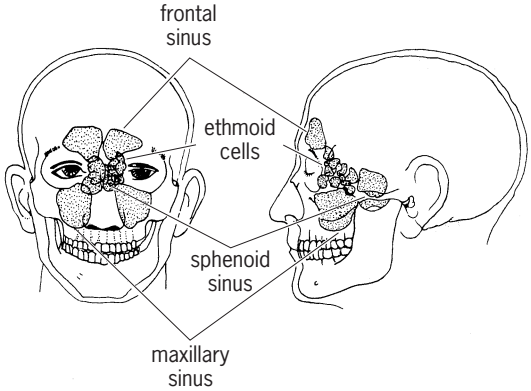

Frontal and side view of face, showing sinuses. (*After N. L. Hoerr and A. Osol, eds., Gould Medical Dictionary, 2d ed., Blakiston–McGraw-Hill, 1956*)

Similar sinuses are found in most mammals and may be very large, as in elephants, in which they serve to lighten the head, but they have no known function in humans. Most submammalians lack them, but crocodilians have comparable cavities emptying into the nasal cavities. Thomas S. Parsons

Siphonaptera

An order of insects commonly known as fleas. These insects are bloodsucking ectoparasites of animals and humans and may transmit serious diseases.

Adult fleas are easily recognized. They are small, wingless, about $^1/_8$ in. (3 mm) in length, dark brown or black, and laterally flattened. They have three pairs of legs, with the third pair modified for jumping. The body of a flea is characteristically oval-shaped and armed with combs, setae, and spines, which prevent the flea from being dislodged from the host. The head bears mouthparts adapted for bloodsucking (see **illustration**).

Life history. Fleas undergo complete metamorphosis (holometabolous); that is, their life history involves four distinct phases: the egg, larva, pupa, and adult. The only exception is the neosomic *Tunga monositus* (sand flea), which has only two stages. Female adult fleas may lay large numbers of eggs (400 or more). Flea eggs are white, ovoid, and nonadhesive. They are laid in the bedding or nest of the host, where they eventually hatch into the larva. Fleas have three larval stages (instars), which are characterized by differences in size (1.5 – 10 mm). The first larval stage has an egg tooth, which is used to open the eggshell. Flea larvae are eyeless and legless, and sparsely covered with bristles. They usually bear an anal comb of bristles on the tenth segment. Only the head and mouthparts are sclerotized. Larvae are seldom seen because they usually hide in the nesting material or in the surroundings of the host. They feed on organic material, mainly droplets of dried, digested blood from the adults. When fully developed, the larva spins a silken cocoon with incorporated debris, to form the pupal stage. This is a nonfeeding, resting period, which may last from a few days to several months, often depending on host availability. The adult flea emerges from the pupa. Under favorable conditions, the entire life cycle may be completed in as little as 2 weeks.

Adult fleas feed exclusively on the blood of the host. They have specialized sucking mouthparts with three piercing blades (stylets) see illustration. Adult fleas have three-segmented antennae, and their eyes are rudimentary. Most species possess combs on their head or the first thoracical segment. The thorax bears three pairs of legs. The third pair, in combination with an elastic structure (pleural arch) within the thorax, enables the flea to jump. The flea abdomen is divided into tergites (dorsal plates) and sternites (ventral plates), which are connected through flexible tissue that enables the abdomen to expand during a blood meal.

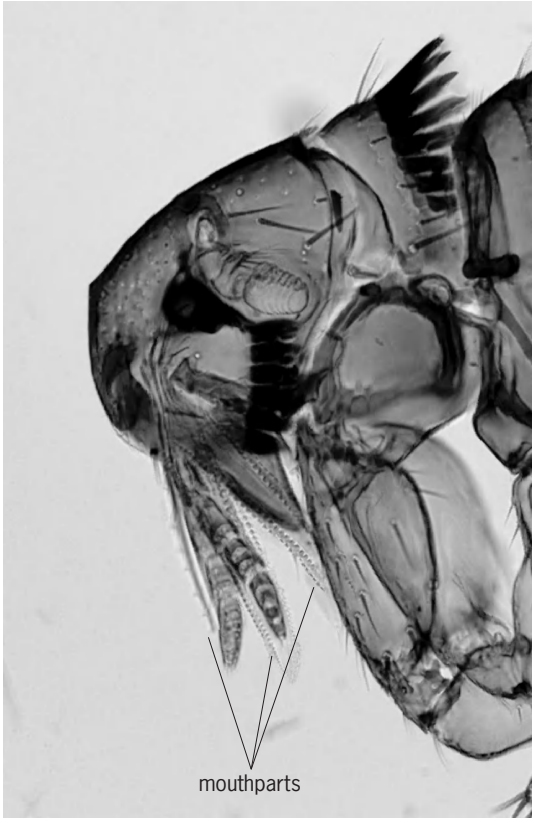

Head of *Spilopsyllus cuniculi* (rabbit flea), with extended mouthparts.

The last abdominal segment has a sensory plate, called the sensillium, containing a variety of sensory pits, whose function is mainly unknown. Adult fleas spend considerable time off the host, and may be found in or around the bedding or nest of the host. They only visit the host to feed, and may feed multiple times during one blood meal.

Medical importance. During the three decades following World War II, the study of fleas increased, and the world flea fauna became very well known; specialists now recognize 2575 species, belonging to 244 genera (see **table**). The great majority of fleas occur on mammals (95%); the remainder live on birds. The host specificity of fleas, as well as their seasonal and geographical distribution, is largely dependent on specific requirements of the larvae. The least host specificity is shown by those taxa whose immature stages tolerate wide ranges of temperature, humidity, and nutritional requirements (for example, cat fleas). Adult fleas may occasionally attack humans if given the opportunity of contact. Since many fleas are carriers and transmitters (vectors) of disease, humans can contract infections through their bite.

Plague, which is responsible for millions of deaths throughout the history of humankind, is a disease of rats It is transmitted between rats and from rats to humans by fleas, particularly the oriental rat flea (*Xenopsylla cheopis*). This is a cosmopolitan species occurring on rats throughout the world. Plague occurs naturally in the wild (sylvatic plague) and usually affects rodents, such as ground squirrels, chipmunks, and prairie dogs of the western United States (Colorado, New Mexico, Arizona, Wyoming). Although many species of fleas are capable of transmitting sylvatic plague, human cases are rare, perhaps because humans do not often come in contact with the animal and insect reservoirs.

Murine or endemic typhus fever, another disease transmitted from domestic rats to humans by rat fleas, was of great public health significance in the southeastern United States during the period 1944–1950, when thousands of cases occurred. Tularemia, or rabbit fever, a disease of small wild animals, particularly of the western and midwestern United States, may be transmitted to humans by fleas. However, this disease, which is not common, is more often contracted while skinning infected animals. Flea allergy is the term applied by physicians to the severe reactions which sometimes result from flea bites. This disease is regarded as a major dermatological problem in California. Often the urticarial papular eruption resulting from flea bites in allergic persons is mistakenly diagnosed as hives and attributed to some food. The fleas involved are those species that most often occur as household pests, the dog flea (*Ctenocephalides canis*), the cat flea (*C. felis*), and the human flea (*Pulex irritans*). In tropical America a flea known as the chigoe, nigua, or sand flea (*Tunga penetrans*) can penetrate into the skin between the toes of humans and cause festering sores which may lead to tetanus and gangrene. The sticktight, or chicken flea (*Echidnophaga gallinacea*), is similar to the nigua but attacks the heads of poultry, cats, and dogs and sometimes annoys

Important species of Siphonaptera

Scientific name	Common name	Importance
Ctenocephalides canis	Dog flea	Intermediate host of tapeworm, readily attacks humans, cosmopolitan, rare flea
Ctenocephalides felis	Cat flea	Common household pest, intermediate host of tapeworm, readily attacks humans, very abundant on cats and dogs, cosmopolitan
Leptopsylla segnis	Mouse flea	Common on mice and rats, may transmit plague and tularemia among rodents, cosmopolitan
Oropsylla montana		Transmits sylvatic plague and tularemia among rodents, common on prairie dogs
Pulex irritans	Human flea	Probably originated with caviomorph rodents, attacks humans, not as common as cat and dog fleas, transmits plague and murine typhus
Xenopsylla cheopis		Main vector of the plague and endemic typhus, cosmopolitan

humans. Fleas are also intermediate hosts of cat and dog tapeworms, and children may become infected by accidentally swallowing dog or cat fleas. *See* PLAGUE; TULAREMIA.

Control of fleas. Bubonic plague and murine typhus fever are controlled by measures directed against rodents, such as poisoning, trapping, and rodent-proofing buildings, and by the application of carbaryl or bendiocarb dust to rodent runs and harborages to kill rodent fleas. Dog and cat fleas can also be controlled through application of the same pesticides. To kill all stages of dog and cat fleas, cellars, outbuildings, and living quarters should be sprayed with bendiocarb (wettable powder), carbaryl (wettable powder), chloropyrifos (emulsifiable concentrate), diazinone (emulsifiable concentrate), and pyrethrum (aerosols, spray). Simultaneously, pets should be treated by veterinarians to control fleas on their bodies. Flea collars impregnated with carbamate or organophosphate insecticides can be used to repel fleas. In humans, shoes and overlapping light-colored clothing should be worn as prophylactic measures along with application of repellent spray containing diethyltoluamide (DEET) on clothing and exposed skin.

Katharina Dittmar; Irving Fox

Bibliography. R. Lewis, Resume of the Siphonaptera (Insecta) of the world, *J. Med. Entomol.*, 35:377–389, 1998; A. Mallis, S. A. Hedges, and D. Moreland, *Handbook of Pest Control*, 9th ed., 2004; D. K. Mardon, *Pygiopsyllidae: An Illustrated Catalogue of the Rothschild Collection of Fleas (Siphonaptera) in the British Museum (Natural History)*, British Museum, vol. VI, 1981; F. G. A. M. Smith, *Malacopsyllidae and Rhopalopsyllidae: An Illustrated Catalogue of the Rothschild Collection of Fleas (Siphonaptera) in the British Museum (Natural History)*, Brithish Museum, vol. VII, 1987; R. Traub, M. Rothschild, and J. Haddow, *The Rothschild Collection of Fleas: The Ceratophyllidae—Key to the Genera and Host Relationships*, Cambridge University Press, 1983.

Siphonophora

An order of the class Hydrozoa of the phylum Coelenterata, characterized by an extremely complex organization of components of several different types, some having the basic structure of a jellyfish, others of a polyp. The components may be connected by a stemlike region or may be more closely united into a compact organism. The medusalike components include swimming bells, a gas-filled float, bracts which are tough gelatinous bodies presumed to be protective, and gonophores which produce the eggs and sperms. Polyplike components include gastrozooids with mouth and one long hollow tentacle; dactylozooids without mouth and often with one tentacle; and gonozooids which bear the gonophores.

Most siphonophores possess a float and are animals of the open seas. Best known is the Portuguese man-of-war, *Physalia*, with a float as much as 16 in. (40 cm) long and tentacles which extend downward for many meters. These animals may be swept shoreward and may make swimming not only unpleasant but dangerous. Some siphonophores propel themselves by jets of water produced by the swimming bells. *See* COELENTERATA; HYDROZOA.

Sears Crowell

Siphonostomatoida

An order of Copepoda; all members are parasites of marine and fresh-water fishes or a variety of invertebrate hosts. In fact, it is estimated that 67% of copepod parasites of fishes belong to the Siphonostomatoida. Their obvious success in the parasitic mode of life appears at least in part to be a result of two morphologic adaptations. The first is the modification of the buccal apparatus (mouth and oral appendages). The siphonostomatoid mouth forms a tube or siphon through the partial or complete fusion of the labrum ("upper lip") and labium ("lower lip") around the esophageal opening. The resulting mouth cone has a small opening near the base through which the mandibles are free to enter. The mandibles themselves are styliform (see **illus.**) and may or may not be armed with denticles. The maxilliae are subchelate or brachiform, superficially resembling a human arm, and serve as the appendage for attachment to the host. The second adaptation is the development of a frontal filament, which is a larval organ of attachment. In those siphonostomatoids possessing a frontal filament, the brachiform second maxilla may be used to manipulate the frontal organ, and in some species fusion of those two structures forms the attachment structure.

Although siphonostomatoids all share the highly specialized character of a tube- or siphonlike mouth that is well adapted to a parasitic mode of life, other morphologic aspects are quite diverse, and so it has been speculated that the order is not in fact monophyletic. Diversity makes it impossible to give a generalized account of the order. Most familiar are the ectoparasites of fishes belonging to the family Caligidae and the mesoparasites of the Pennellidae that embed in the tissues of marine mammals (see illus.).

Those siphonostomatoids parasitic on invertebrates typically resemble most closely the cyclopoids in having primitive podoplean body segmentation and reduction of the fifth thoracic legs. The remaining siphonostomatoids, all parasites of vertebrates, are more highly modified, often bearing little resemblance to Copepoda at all. Most specializations have involved adaptations of the cephalic appendages to serve as holdfasts and body streamlining that would enhance adhesion to the body surface of the host.

It has been hypothesized that the Siphonostomatoida evolved from the podoplean line much earlier than the other major order of parasitic Copepoda,

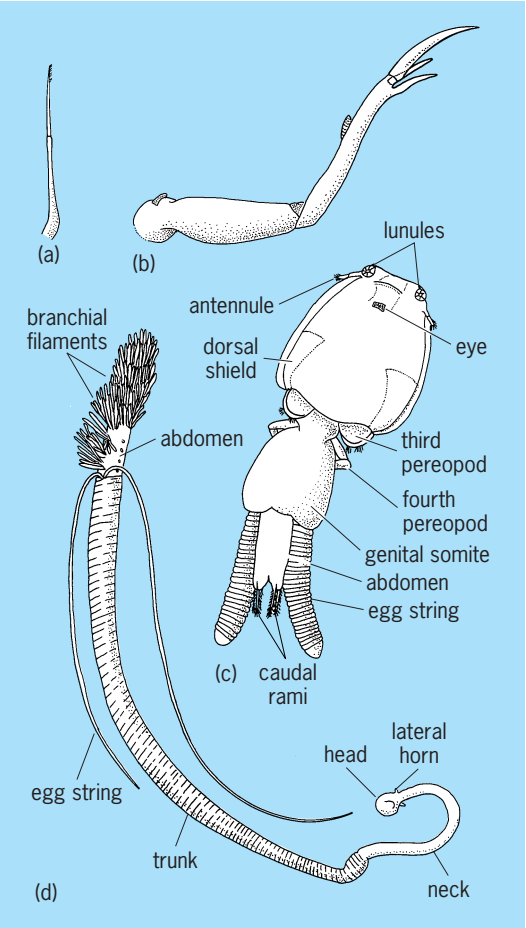

Siphonostomatoida. (*a*) Representative siphonostomoid mandible; (*b*) brachiform maxilla (*after Z. Kabata, Parasitic Copepoda of British Fishes, Ray Society, London, vol. 152, 1979*). (*c*) Representative female caligid; (*d*) female Pennella (*after P. A. McLaughlin, Comparative Morphology of Recent Crustacea, W. H. Freeman, 1980*).

the Poecilostomatoida. Their pathway presumably involved an epibenthic, epiphytic, or epizoic ancestor of the Cyclopoida, in which the siphonlike mouth cone developed first in invertebrate parasites where the hosts were at least temporarily stationary and in contact with a firm substrate. *See* COPEPODA; CRUSTACEA; POECILOSTOMATOIDA. Patsy A. McLaughlin

Bibliography. Z. Kabata, *Parasitic Copepoda of British Fishes*, Ray Society, London, vol. 152, 1979; B. M. Marcotte, Evolution within the Crustacea, pt. 2: Copepoda, in L. G. Abele (ed.), *The Biology of Crustacea*, vol. 1: *Systematics, the Fossil Record, and Biogeography*, 1982.

Sipuncula

A phylum of sedentary marine vermiform coelomates that are unsegmented, but possibly distantly related to the annelids; they are commonly called peanut worms. There are 17 genera and approximately 150 species living in a wide variety of oceanic habitats within the sediment or inside any protective shelter such as a discarded mollusk shell, foraminiferan test, or crevice in rock or coral.

Anatomy. Sipunculans have two body regions: the trunk and a retractable introvert (see **illustration**). The mouth is at the tip of the introvert and in one class is surrounded by tentacles of varying degrees of development and complexity. Behind the tentacular region is a zone which may bear posteriorly directed hooks either scattered or arranged in regular rings. The introvert varies from less than half the trunk length in some species to many times its length in others. Adult sipunculans range in trunk length from 2 to over 500 mm (0.08 to over 20 in.). The shape of the body ranges from almost spherical to a slender cylinder. Sipunculans have a variety of epidermal structures (papillae, hooks, or shields). Many species lack color, but shades of yellow or brown may be present.

Internal anatomy is relatively simple. The digestive tract has a straight esophagus and a double-coiled intestine extending toward the posterior end of the body and back (see illustration) terminating in a rectum, sometimes bearing a small cecum. The dorsal anus is located at the anterior end of the trunk, except in one genus (*Onchnesoma*) in which it is on the introvert. Near the anus a threadlike spindle muscle originates from the body wall and extends down the center of the gut coil, increasing its stability (see illustration). This extends to the posterior end of the trunk in many genera. A pair (only one in two genera) of simple, saclike metanephridia open ventrolaterally at the anterior end of the trunk. The longitudinal muscles on the inside of the body wall form a smooth, uniform layer in some genera, but are gathered into bundles in most. Four retractor muscles (sometimes fused to form two or rarely one muscle) control the introvert. A ventral nerve cord (see illustration) with lateral nerves and circumenteric connectives to the pair of cerebral ganglia are present. Two or four pigmented eyespots may be present on the cerebral ganglia, and a chemoreceptor (nuchal organ) is usually present.

Reproduction. The reproductive biology of sipunculans has been examined in representatives of all families. The timing and length of the reproductive period varies depending on species and geographic region. Small gametes are produced from a transient strip of tissue at the base of the ventral retractor muscles during the breeding season. They are released into the spacious coelom at an early stage, where they undergo the remainder of their growth and differentiation. One species is known to be monoecious (containing both male and female gametes), one is known to be parthenogenic (eggs develop without sperm), and two are capable of reproducing by asexual, transverse fission. Most species of sipunculans are dioecious (have two sexes) that are similar that is, they lack any sexual dimorphism. The sipunculan gametes are retained in the coelom until mature, then released through the nephridiopores, fertilization occurring externally. Free-swimming trochophore larvae generally occur, but some brooding of early stages within the maternal "home" has been

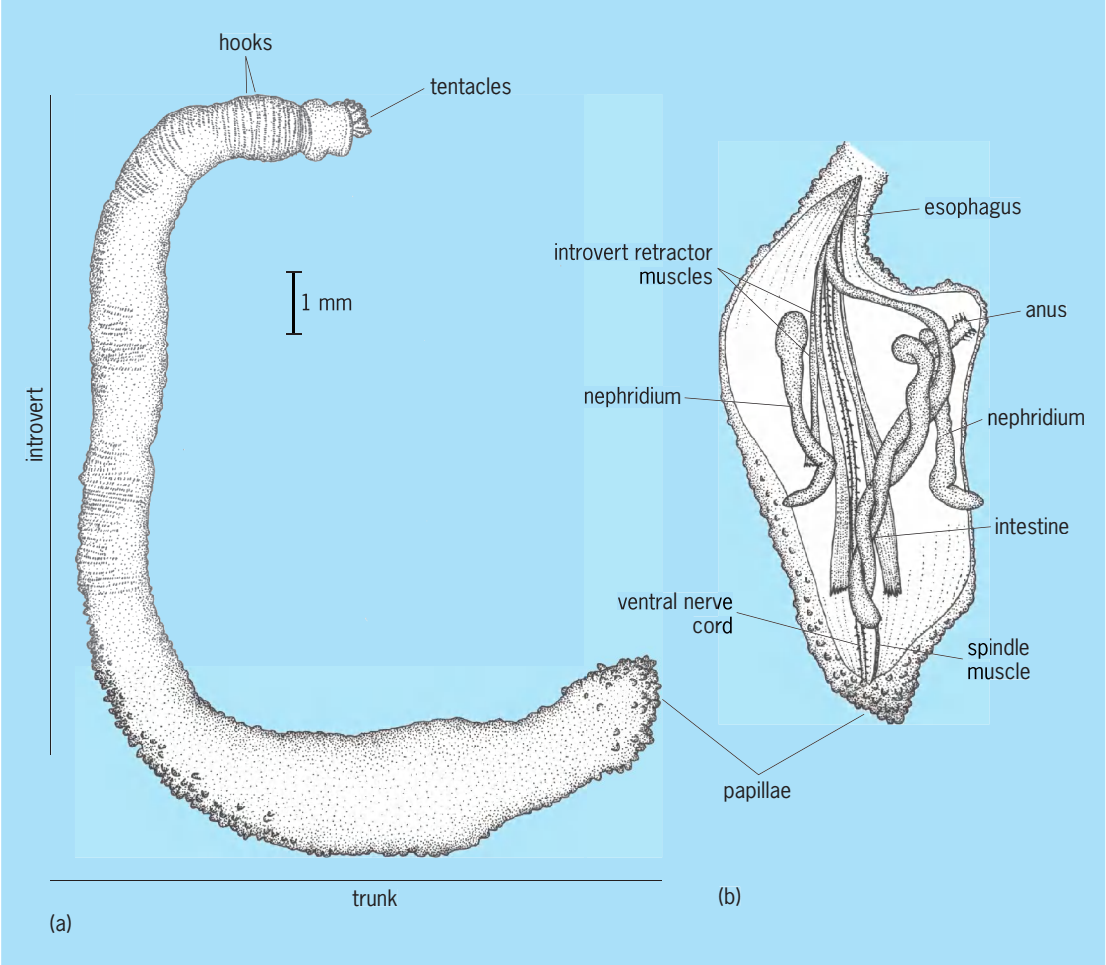

Sipunculan body regions. (*a*) External morphology and (*b*) internal anatomy of *Phascolosoma nigrescens*. *(Illustration by Anju Schulze)*

observed in *Nephasoma minuta* and may occur in other cold- or deep-water forms. The unique long-lived pelageosphera larvae are produced by some genera, allowing for transoceanic dispersal. After metamorphosis, the juveniles settle onto a suitable substratum and probably remain in that immediate vicinity throughout their lives, although there is no knowledge of growth rate or longevity.

Ecology. Most sipunculans are deposit feeders, but a few (sucnas *Themiste*) are filter feeders with elaborate tentacular crowns. Therefore, these worms play a part in the recycling of detritus and probably consume smaller invertebrates in the process. They are in turn preyed on by fishes and probably other predators (including humans in China). *See* FEEDING MECHANISMS (INVERTEBRATE).

Taxonomy. In recent years the phylogenetic relationships of the members of the phylum Sipuncula have begun to be investigated by analysis of deoxyribonucleic acid (DNA) sequence data from several different genes. While this work is currently in progress, preliminary results do support the monophyly of the phylum. However, support for all of the existing higher taxa (orders, families, and genera) is not universal, and it is likely that in the near future some changes will be proposed to the currently accepted system, which uses morphological characters to define two classes: Sipunculidea and Phascolosomatidea. Sipunculidea, with peripheral tentacles around the mouth, and simple scattered hooks when present, contains two orders (Golfingiiformes and Sipunculiformes), four families, and 11 genera. Phascolosomatidea, without peripheral tentacles, have short tentacles around the dorsal nuchal organ, and complex hooks in rings when present. It contains two orders (Phascolosomatiformes and Aspidosiphoniformes), two families, and six genera. Edward B. Cutler; A. Schulze

Bibliography. E. B. Cutler, *The Sipuncula: Their Systematics, Biology, and Evolution*, Cornell University Press, Ithaca, New York, 1994; P. E. Gibbs and E. B. Cutler, A classification of the phylum Sipuncula, *Bull. Brit. Mus. Nat. Hist. (Zoology)*, 52(1):43–58, 1987; A. B. Maxmen et al., Evolutionary relationships within the protostome phylum Sipuncula: A molecular analysis of ribosomal genes and histone H3 sequence data, *Mol. Phylogen. Evol.*, 27(3):489–503, 2003; M. E. Rice, Sipuncula, in K. G. Adiyodi and R. G. Adiyodi (eds.), *Reproductive Biology of Invertebrates*, pp. 263–280, Oxford and IBH, Kerala,

1988; M. E. Rice, Sipuncula, in F. W. Harrison and M. E. Rice (eds), *Microscopic Anatomy of Invertebrates*, vol. 12: *Onychophora, Chilopoda, and Lesser Protostomata*, pp. 238–325, Wiley-Liss, New York, 1993.

Siren (acoustics)

A sound source that is based on the regular interruption of a stream of fluid (usually air) by a perforated rotating disk or cylinder. The components of a siren are a source of air, a rotor containing a number of ports which interrupt the airflow at the desired frequency, and ports in a stator through which the air escapes. The air is supplied by a compressor, and a motor drives the rotor. The frequency of the sound wave produced by the siren is the product of the speed of rotation and the number of ports in the rotor. The shape of the rotor and stator ports determines the wave shape at the entrance of the stator port. The stator ports feed into a horn in order to improve radiation. Siren performance parameters are sound power output, acoustic pressure, and efficiency, that is, the ratio of acoustic power output to compressor power. *See* SOUND PRESSURE.

The first sirens were developed in the first half of the nineteenth century. They provided a sound source with a known and variable frequency and were used to study pitch and beats. Sirens were first used as foghorns and air warning devices in the latter half of the nineteenth century.

Sirens are used as high-intensity sound sources for sonic aerosol applications. Examples are acoustic agglomeration (the coagulation of aerosol particles in the presence of an acoustic field), combustion enhancement (for example, that of coal particles due to increased heat and mass transfer caused by the acoustic wave), and acceleration of gaseous reaction rates. Other applications of sirens include acoustic levitation (the use of radiation pressure to levitate small objects), broadband underwater sound projectors, and sonic fatigue (fatigue life and failure of structures subjected to fluctuating pressures generated by acoustic waves). *See* ACOUSTIC LEVITATION.

Electromechanical sirens use an electric motor instead of a compressed air supply to generate the acoustic signal. A second motor spins the rotor. The stator and horn increase the sound power output and efficiency. Electromechanical sirens are widely used as warning devices. *See* SOUND. Bart Lipkens

Bibliography. R. T. Beyer, *Sounds of Our Times: Two Hundred Years of Acoustics*, Springer-Verlag, New York, 1999; T. F. Hueter and R. H. Bolt, *Sonics*, Acoustical Society of America, 2000.

Sirenia

An order of herbivorous aquatic placental mammals, commonly known as sea cows, that includes the living manatees and dugongs and the recently exterminated Steller's sea cow. The order has an extensive fossil record dating from the early Eocene Epoch, some 50 million years ago.

West Indian manatee (*Trichechus manatus*) and calf. (*U.S. Fish & Wildlife Service/Galen Rathbun*)

The earliest known sirenians were quadrupedal and capable of locomotion on land. Fossils clearly document the evolutionary transition from these amphibious forms to the modern, fully aquatic species, which have lost the hindlimbs and transformed the forelimbs into paddlelike flippers. The living species have streamlined, fusiform bodies with short necks and horizontal tail fins like those of cetaceans, but no dorsal fins (see **illustration**). The skin is thick and nearly hairless. The nostrils are separate, and the ears lack external pinnae. The apex of the heart is bifid. The thorax and lungs are elongate and the diaphragm is nearly horizontal so that the buoyant lungs lie wholly dorsal to the abdominal viscera. The skeleton typically shows both an increase in the bulk of individual bones (pachyostosis) and an increase in the proportion of compact bony tissue (osteosclerosis). These conditions provide ballast to neutralize the buoyancy of the lungs, which do not collapse during submergence in these shallow-diving animals. The manus has five digits, and bears small nails in two of the living manatee species. The organs of feeding include large, fleshy lips; numerous tactile and prehensile vibrissae; and horny pads covering the toothless anterior parts of the palate and mandibular symphysis (area of union of bones). The cheek teeth are primitively low-crowned, with two rows of rounded cusps, but the dentitions of the Recent species are divergently specialized from this primitive pattern. Where incisors are present, the first upper incisor is the largest (and in later species the only one), and it often forms an enlarged tusk. Sirenians typically feed on aquatic angiosperms, especially seagrasses (Hydrocharitaceae and Potamogetonaceae), but in ecologically marginal situations they also eat algae and even some animal material. They are normally found in tropical or subtropical marine waters, but some have become adapted to fresh water or colder latitudes. Body sizes have ranged from less than 3 m up to 9–10 m (10 to 30–33 ft). Sirenians mate and give birth in the water, bearing a single calf (occasionally twins) after about 13–14 months of gestation and then nursing it from one pair of axillary

mammae. The closest relatives of sirenians among living mammals are the Proboscidea (elephants). *See* PROBOSCIDEA.

A classification scheme is given below.

Order Sirenia
- Family: Prorastomidae
  - Protosirenidae
  - Trichechidae
  - Subfamily: Miosireninae
    - Trichechinae
- Family Dugongidae
  - Subfamily: Halitheriinae
    - Hydrodamalinae
    - Dugonginae

Primitive sirenians. Sirenians are presumed to have originated in the Old World, along the shores of the former Tethys seaway. However, the most primitive fossils of the order have turned up by chance in Jamaica (*Prorastomus* and related forms). These were pig-sized amphibious creatures with long bodies, short legs, and narrow snouts that were not deflected from the palatal plane. They were apparently selective browsers on assorted aquatic plants. The dentition of these and other Eocene sirenians is noteworthy in having five premolars like the most primitive placental mammals. Whether this condition is really primitive or a sirenian specialization is unresolved. Some prorastomids already show weakening of the connection between the spinal column and the hindlimbs. Protosirenids, known from the Eocene of North Africa, Asia, Europe, and possibly North America, show further reduction of strength of the hindlimbs and (like all later sirenians) significant downturning of the snout (an adaptation for bottom feeding). *See* DENTITION.

Dugongids. Known from the middle Eocene to the present, dugongids have been the most diverse and successful sirenian family. The Eocene fossils of this group show a steady reduction of the hindlimb and the pelvic bones until the body takes on, by the end of the Eocene, the outward form seen in modern sea cows. Fossil dugongids have been found in tropical or formerly tropical areas around the globe, and seem always to have inhabited marine waters. Their anatomically most conservative members are placed in the subfamily Halitheriinae, which gave rise to the Dugonginae in the Oligocene Epoch and the Hydrodamalinae in the Miocene Epoch.

Dugongines apparently arose in the West Atlantic and the Caribbean regions but quickly spread to the Old World. Many of them evolved large tusks that were probably used for digging up seagrass rhizomes. They are represented today by the single species *Dugong dugon*, found in the Indian and southwest Pacific oceans. It is notable for its reduced, peglike cheek dentition: the enamel crowns of the molars wear off quickly and the chewing surface is then formed only by the roots, which in the case of the last two molars are ever growing.

Hydrodamalines evolved in the North Pacific Ocean and remained confined to that region. Over the course of the Miocene and Pliocene epochs, they adapted to cooler climates and higher-energy habitats. The lineage culminated in Steller's sea cow (*Hydrodamalis gigas*), the largest of the sirenians, which tolerated temperate and cold waters and fed mainly on kelp (large brown and red algae). It was completely toothless, and its highly modified, hook-like flippers lacked phalanges. It was exterminated by human hunting around 1768.

Trichechids. Apart from the extinct and poorly known miosirenines, which inhabited northern Europe and may have been shellfish eaters, the trichechids comprise the sirenians known as manatees. Trichechids may have arisen from primitive dugongids, in the late Eocene Epoch. The fossil record of manatees is poor, but the group seems to have evolved mainly in South America before spreading to North America, and even crossing the Atlantic to West Africa, late in the Tertiary Period. While still in South America, manatees adapted to a diet composed largely of abrasive aquatic grasses (Gramineae) by evolving the ability to replace their cheek teeth without limit. As the teeth become worn, they are shed at the front of the toothrow while replacement molars erupt at the rear, and the whole row shifts forward like a conveyor belt. This form of horizontal tooth replacement has been misleadingly likened to that of elephants, but the only mammal to have evolved a truly similar replacement process is a species of Australian wallaby. Today, manatees are represented by three species: *Trichechus manatus*, the West Indian manatee, and *T. senegalensis*, the West African manatee, both of which inhabit salt water as well as fresh water; and *T. inunguis*, the Amazonian manatee, which lives only in fresh water. Unlike dugongs, which have dolphinlike flukes, manatees have rounded, paddle-shaped tail fins, and only six cervical vertebrae instead of the normal mammalian complement of seven. The vestigial pelvis is even more reduced than in dugongids. *See* MAMMALIA.

Daryl P. Domning

Bibliography. D. P. Domning, Bibliography and Index of the Sirenia and Desmostylia, *Smithson. Contrib. Paleobiol.*, 80:1–611, 1996; H. Marsh, Our tropical siren, *Austral. Geog.*, 21:42–57, 1991; R. M. Nowak (ed.), *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; J. E. Reynolds III and D. K. Odell, *Manatees and Dugongs*, Facts on File, 1991; W. Zeiller, *Introducing the Manatee*, University Press of Florida, 1992.

Sirius

The star α Canis Majoris, also referred to as the Dog Star, the brightest of all the stars in the night sky (apparent magnitude -1.47). Sirius owes its apparent brightness both to its close distance to Earth, only 2.64 parsecs (8.14×10^{13} km or 5.06×10^{13} mi), and to its intrinsic luminosity, which is more than

20 times that of the Sun. It is a main-sequence star of spectral type A1 with an effective temperature of about 9400 K (16,500°F). *See* SPECTRAL TYPE.

Sirius is a very interesting binary system in which the companion is a degenerate star (α CMa B). It is the brightest among the known white dwarfs (apparent magnitude 8.4) and the closest to the Sun. It orbits around the bright star once every 50 years at a mean angular separation of 7.5″, corresponding to about 20 astronomical units. The masses of α CMa A and B as computed from the orbit are 2.14 and 1.03 times the solar mass, respectively. The luminosity of the white dwarf is 400 times smaller than that of the Sun, which makes it extremely difficult to see because of the overpowering brightness of the nearby primary. Its temperature is estimated to be 24,790 K (44,654°F), much hotter than the primary. The spectrum shows hydrogen lines broadened by extreme pressure, a direct consequence of the fact that the white dwarf's mass is comparable to the Sun, yet is concentrated in a body the size of the Earth. The mean density of the material in α CMa B is about 2×10^6 times that of water, or 2 metric tons/cm^3 (nearly 40 tons/in.^{3}). *See* BINARY STAR; WHITE DWARF STAR.

Sirius has been found to have chemical anomalies. Certain elements such as barium, lanthanum, and cerium are overabundant by a factor of about 100 compared to the Sun. These anomalies may be related to the fact that Sirius is in a binary system with a white-dwarf companion. Other stars with similar peculiarities in binary systems with a white dwarf are believed to have been polluted by material shed by the companion, which was originally the more massive star in the system and evolved more rapidly. *See* STAR; STELLAR EVOLUTION. David W. Latham

Sirocco

A southerly or southeasterly wind current from the Sahara or from the deserts of Saudia Arabia which occurs in advance of cyclones moving eastward through the Mediterranean Sea. The sirocco is most pronounced in the spring, when the deserts are hot and the Mediterranean cyclones are vigorous. It is observed along the southern and eastern coasts of the Mediterranean Sea from Morocco to Syria as a hot, dry wind capable of carrying sand and dust great distances from the desert source. The sirocco is cooled and moistened in crossing the Mediterranean and produces an oppressive, muggy atmosphere when it extends to the southern coast of Europe. Rain that falls in this air is often discolored by the dust or sand which is precipitated along with the waterdrops. Under sunny conditions, the sirocco can produce temperatures in excess of 100°F (38°C) in southern Europe. Various other names are used to denote the sirocco in specific localities, such as khamsin in Egypt. *See* AIR MASS; WIND.

Frederick Sanders

Sisal

A fiber obtained from the leaves of *Agave sisalana*, produced in Brazil, Haiti, and several African countries, including Tanzania, Kenya, Angola, and Mozambique.

A well-developed sisal plant has a stem or trunk about 12–16 in. (30–40 cm) in diameter and 3 ft (1 m) in height at maturity. The leaves, 30–60 in. (75–150 cm) in length and 4–6 in. (10—15 cm) wide, are bright to dark green, thick, succulent, and smooth, with hardened margins that are smooth with occasional small, curved spines and a sharp terminal spike 0.4–1.2 in. (1–3 cm) long. When the plant is 6–10 years old, a flower stalk grows up from the top to a height of 10–23 ft (3–7 m). The stalk branches and produces flowers and then bulbils; seed is seldom produced.

Cultivation. Sisal grows best in the tropics where rainfall is low to moderate and the soil is calcareous with a texture that allows aeration of the roots. However, the plant does well on a fairly wide range of nonacid soils.

Land for planting should be plowed and cultivated to kill weeds. Sisal plants are propagated by bulbils or by suckers which grow from the roots of established plants. Bulbils, which tend to produce better plants, are planted in a nursery for 1 to 2 years; suckers usually are moved directly to the field location. A widely used planting pattern is a series of double rows about 3 ft (1 m) apart with a 11.5–13 ft ($3^1/_2$–4 m) avenue between each pair of rows. Plants are spaced about 3 ft (1 m) apart in the row.

Harvesting and processing. The first harvest is ready in about 3 years after field planting from the nursery stage. About 20–25 leaves having an angle of 40° or more from vertical are cut at intervals of 6–12 months, depending on growing conditions.

The fibers are found in the fleshy tissue of the leaf, mostly in association with the vascular bundles. The freshly cut leaves are brought to a central location for the defibering operation, frequently called decortication, which involves beating and scraping. The basic defibering units are wheels or drums 3–5 ft (100–150 cm) in diameter and about 12 in. (30 cm) wide. Blunt beater blades 2 in. (5 cm) high and spaced 10 in. (25 cm) apart, are attached to the face of the drum at right angles to the direction of rotation. An adjustable breastplate one-third the circumference of the drum is set so that the beater blades clear it by about the thickness of a fiber at its narrowest part, the trailing edge. The space is wider where the leaf enters at the leading edge. As the leaf passes between the breastplate and the drum, the pulp is beaten out and the fibers are scraped clean. Jets of water, directed onto the fiber as it passes through each defibering unit, wash the fiber and carry away the waste. A carrying and gripping device moves a layer of leaves sideways into one end of the machine through the first defibering unit, where the lower half of the leaves are cleaned. Then the grip is transferred to the

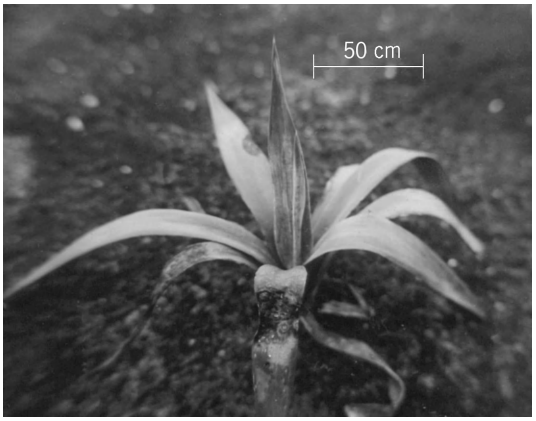

Leaf spot of sisal caused by *Colletotrichum agaves*.

cleaned fiber, and the layer of partly cleaned leaves continues through the second defibering unit, where cleaning is completed. The fiber is now ready to be dried on poles in the sun or in an apron conveyor dryer.

A smaller, lighter defibering machine with the same action is also used. The top of the leaf is handheld while the base end is inserted against the revolving drum to about the center of the leaf and withdrawn. The clean fiber is then grasped, and the other end is cleaned. This machine cleans fiber with less loss, but production per unit of labor is very low.

Uses. Sisal is used mainly for twine and rope, but some of the lower grades are used for upholstery padding and paper. The greatest quantity goes into farm twines, followed by industrial tying twine and rope. Most sisal-fiber ropes made in the United States are small to medium in size, intended for light duty. Sisal is sometimes used for marine cordage in Europe. *See* NATURAL FIBER.

Elton G. Nelson

Diseases. Sisal is subject to fewer diseases than most crops, with many disease symptoms being caused by nutrient deficiencies. Fungus diseases includes two bolt (stem) rots. *Aspergillus niger* has been associated with one which is a soft wet rot. The other, a dry rot, has been attributed to *Fusarium solani*. The fungus enters through the exposed base of harvested leaves and spreads inwardly.

Leaf spots are caused by the fungi *Microdiplodia agaves, Marssonina agaves,* and *Colletotrichum agaves* (see **illus.**).

Nutritional diseases include leaf banding, chlorotic leaf spot, and tip withering and mottling.

The only major insect pest of sisal is the weevil *Scyphophorus interstitialis*. It can be reduced by field sanitation and by spraying with aldrin or dieldrin.

The cause of the destructive "red rot" disease is suspected to be a virus. The fiber of the closely related plants *A. fourcroydes* and *A. cantala* is sometimes called sisal fiber; these species are affected by many of the above diseases. *See* LILIALES; PLANT PATHOLOGY; PLANT VIRUSES AND VIROIDS.

Thomas E. Summers

Skarn

A broad range of rock types made up of calc-silicate minerals such as garnet, regardless of their association with ores, that originate by replacement of precursor rocks. It was a term originally coined by miners in reference to rock consisting of coarse-grained, calc-silicate minerals associated with iron ores in central Sweden. Ore deposits that contain skarn an gangue are termed skarn deposits; such deposits are the world's premier sources of tungsten. They are also important sources of copper, iron, molybdenum, zinc, and other metals. Skarns also serve as sources of industrial minerals such as graphite, asbestos, and magnesite. *See* ORE AND MINERAL DEPOSITS; SILICATE MINERALS.

Characteristic features and origin. Skarns are composed largely of calc-silicate minerals (for example, garnet, pyroxenoid, pyroxene, olivine, and amphibole). They are relatively coarse-grained (for example, the garnets up to 1 cm (0.4 in.) in diameter and the pyroxenes up to 10 cm (4 in.) in length). Skarns are characterized by few minerals for the number of chemical components; and they are commonly zoned (consisting of zones of different minerals arranged systematically around the locus of replacement). Skarns form in geological environments characterized by igneous intrusion, metamorphism, and circulation of high-temperature (800–400°C or 1470–750°F) aqueous fluids at relatively shallow crustal levels (1–10 km or 0.6–6 mi). Skarn formation involves complex reactions between aqueous fluids and rocks (or protoliths). These reactions result in changes in the chemical composition, mineralogy, and texture of the protoliths. The most common protoliths are calcareous (for example, limestone or dolostone), although noncalcareous rocks (such as igneous rocks) also are replaced by skarn. The most common sources of skarn-forming fluids are metamorphic water released when hydrous minerals such as clays are subjected to increasing temperature; and magmatic water released during final crystallization of granitic magmas. Some magmatic waters are enriched in metals and volatile components, and these may form skarn deposits where they encounter calcareous rocks. *See* METAMORPHISM.

Classification. Based on mineralogy, three idealized types of skarn are recognized: calcic skarn characterized by calcium- and iron-rich silicates (andradite, hedenbergite, wollastonite) most commonly replaces limestone; magnesian skarn characterized by calcium- and magnesium-rich silicates (forsterite, diopside, serpentine) most commonly replaces dolostone; and aluminous skarn characterized by aluminum- and magnesium-rich calc-silicates (grossularite, vesuvianite, epidote) replaces either limestone or igneous rocks. Based on dominant process, three idealized skarn types are recognized: metamorphic skarn from isochemical recrystallization of metal-rich, calcareous sedimentary rocks; reaction skarn by local exchange of components along contacts between reactive rocks; and metasomatic skarn as the result of transfer of components into the

local environment from an external source (such as a magma). The vast majority of skarn deposits contain metasomatic skarn. *See* DOLOMITE; LIMESTONE; METAMORPHIC ROCKS.

Classification of deposits. Skarn deposits, are generally classified on the basis of dominant economic metal association. Because different metal associations reflect different sources of hydrothermal fluids or different local geological and geochemical environments, the metal classification of skarn deposits also results in a separation of deposit types on the basis of geological and mineralogical features. The most abundant of the skarn deposit types include iron deposits (Sarbai in Siberia, Russia; Eagle Mountain in California); tungsten deposits (MacMillan Pass in Yukon, Canada; King Island in Tasmania, Australia); copper deposits (Bingham, Utah; Ok Tedi, Papua New Guinea); and zinc-lead deposits (Yeonhwa-Ulchin, South Korea; EI Mochito, Honduras). *See* COPPER; IRON; LEAD; TUNGSTEN; ZINC.

M. T. Einaudi

Bibliography. M. T. Einaudi, L. D. Meinert, and R. J. Newberry, Skarn deposits, *Economy Geology 75th Anniversary Volume*, pp. 317–391, 1981; V. A. Zharikov, Skarns, *Int. Geol. Rev.*, 12:541–559, 619–647, 706–775, 1970.

Skeletal system

The supporting tissues of animals which often serve to protect the body, or parts of it, and play an important role in the animal's physiology.

Skeletons can be divided into two main types based on the relative position of the skeletal tissues. When these tissues are located external to the soft parts, the animal is said to have an exoskeleton. If they occur deep within the body, they form an endoskeleton. All vertebrate animals possess an endoskeleton, but most also have components that are exoskeletal in origin. Invertebrate skeletons, however, show far more variation in position, morphology, and materials used to construct them.

Exoskeletons. Many of the invertebrate phyla contain species that have a hard exoskeleton, for example, corals (Cnidaria); limpets, snails, and *Nautilus* (Mollusca); and scorpions, crabs, insects, and millipedes (Arthropoda). However, these exoskeletons have different physical properties and morphologies. The form that each skeletal system takes presumably represents the optimal configuration for survival. *See* ARTHROPODA; MOLLUSCA; ZOOPLANKTON.

Calcium carbonate is the commonly found inorganic material in invertebrate hard exoskeletons. The stony corals have exoskeletons made entirely of calcium carbonate, which protect the polyps from the effects of the physical environment and the attention of most predators. Calcium carbonate also provides a substrate for attachment, allowing the coral colony to grow. However, it is unusual to find calcium carbonate as the sole component of the skeleton. It normally occurs in conjunction with organic material, in the form of tanned proteins, as in the hard shell material characteristic of many mollusks. Cuticular exoskeletons are widely distributed among the invertebrates. They are formed by proteins that have been stiffened by the chemical action of phenols (the process is called tanning). Mineral salts may be incorporated within the cuticle for additional strength and stiffness.

The major advantage of a hard exoskeleton is the high degree of protection afforded to the body organs against mechanical damage and desiccation. Some disadvantages are the relatively large mass of the skeletal structure in proportion to the size of the soft tissues; the inability of some animals to remodel the skeleton and repair damage; and problems for a growing animal bounded by rigid skeletal tissues. Arthropods have solved this last problem by periodically shedding the old skeleton (molting) and replacing it with a new, larger one. Immediately after molting, the new skeleton is soft and compliant, allowing it to stretch to accommodate the increased size of the animal. During this period the animal is more vulnerable to predation. *See* MOLTING (ARTHROPODA).

Endoskeletons. Internal hard skeletons are less common among the invertebrates but are a feature of some mollusks and echinoderms. There are various degrees of skeletal reduction seen in these groups, as evidenced by the urchins which have a complete rigid skeletal test formed of calcareous ossicles sutured together, and by sea cucumbers in which the skeleton is reduced to microscopic ossicles. The vertebrate endoskeleton is usually constructed of bone and cartilage; only certain fishes have skeletons that lack bone. In addition to an endoskeleton, many species possess distinct exoskeletal structures made of bone or horny materials. This dermal skeleton provides support and protection at the body surface. *See* ECHINODERMATA.

Human Skeleton

Various structural components make up the human skeleton, including collagen, three different types of cartilage, and a variety of bone types.

Collagen. Collagen is the most abundant protein in the body. It is of fundamental importance in all organ systems and is found in all parts of the musculoskeletal system. Tropocollagen, a structure composed of three polypeptide chains, forms the basic building unit of collagen. The rodlike tropocollagen molecules polymerize into large collagen fibrils. Cross-links between adjacent molecules and also between fibrils give collagen its tensile strength. In tendons and ligaments, these fibrils are bundled together to form larger fibers. Tendons are formed from parallel bundles of fibers, an arrangement which allows tendons to support high uniaxial tensile loads to which they are subjected during activity. Although most ligament fibers are parallel, some have an oblique orientation. This reflects the more complex loading patterns applied to ligaments during movement, where small loads may be applied in a variety of directions. Collagen is stiff and strong only in tension. *See* COLLAGEN.

Cartilage. The formation of cartilage (chondrogenesis) from mesenchyme occurs in many areas of the embryo, such as the skull, limbs, and vertebral column. Most embryonic cartilage is replaced by true bone in endochondral bone formation, but in many regions the cartilage remains throughout life. The tissue consists of cartilage cells (chondrocytes) which manufacture, secrete, and maintain the extracellular organic matrix that surrounds them. This matrix contains a dense network of collagen fibrils within a concentrated solution of protein-polysaccharide molecules called proteoglycans. The three main types of cartilage are fibrocartilage, elastic cartilage, and hyaline cartilage.

Hyaline cartilage, the most abundant form of cartilage in the body, ossifies during development to become bone. In the adult it covers the articular surfaces of bones, supports the trachea and bronchi, forms the costal cartilages linking the first ten ribs to the sternum, and reinforces the nose. This type of cartilage contains extremely fine collagen fibrils that can be seen only with the electron microscope. Hyaline articular cartilage is a special form found within synovial joints. Its material characteristics are perfectly adapted to this mechanically demanding environment.

Fibrocartilage is found predominantly in the pubic symphysis, the menisci of the knee, and the intervertebral discs. This type of cartilage is very durable and can withstand large tensile and compressive forces. It also represents a transitional material found at tendon and ligament insertions into bone and at the margins of some joint cavities. Histologically it can be characterized by large collagen fibers running through the matrix.

Elastic cartilage is similar to hyaline cartilage but contains abundant elastin fibers which give it a yellowish appearance and make it very flexible, while still maintaining its strength. It is found in the outer ear, the auditory tube, and the larynx. *See* CONNECTIVE TISSUE.

Bone. This specialized rigid connective tissue comprises cells which produce an organic extracellular matrix of collagen fibers and ground substance, and inorganic materials in the form of mineral salts. In vertebrates, the mineral portion of bone consists primarily of calcium and phosphate in the form of hydroxyapatite crystals. The organic phase gives bone its resilience and flexibility, while the inorganic phase makes the bone hard and rigid.

The functions of bone are numerous, relating to the maintenance of mineral (mainly calcium) homeostasis, formation of blood cells, and mechanical requirements.

Mineral (mainly calcium) homeostasis. Plasma calcium ion (Ca^{2+}) concentrations are maintained at about 5 mEq/L. This concentration is effectively regulated by calcitonin and parathyroid hormone and is required for normal blood clotting and nerve and muscle function. Low concentrations of calcium in the extracellular fluid that bathes the parathyroid glands elicits parathyroid hormone release. This stimulates osteoclastic breakdown of bone, releasing calcium and phosphorus into the extracellular fluid. Simultaneously, calcium absorption from the gut is increased, calcium loss via the kidneys is decreased, and urinary phosphate excretion is elevated. Calcitonin, released from the thyroid gland in response to elevated calcium concentrations, inhibits calcium removal from bone and increases urinary calcium excretion.

Formation of blood cells. During embyonic development, cellular elements of the blood (red and white blood cells and platelets) are produced by a process known as hematopoiesis. Early production occurs in the vessels of the yolk sac, but following the development of other organ systems, major blood cell production occurs in the liver, spleen, thymus, and bone marrow. At about the sixth developmental month, red bone marrow becomes the major site for the production of red and white blood cells and platelets. In adults, bone is the primary site for white cell production and the only site for red cell production, with portions of the sternum, ribs, vertebrae, skull, pelvis, scapulae, and proximal femoral and humeral heads of particular importance.

Mechanical requirements (support, protection, and leverage). Bone has a protective function, particularly with regard to the central nervous system. The skeleton is the supporting framework for the body and provides stiff levers on which muscles act to generate movement. *See* CALCIUM METABOLISM; HEMATOPOIESIS; HOMEOSTASIS.

There are a number of easily recognized bone types whose microstructure can be related to the mechanical functions required of the bone. The only type of bone that does not require a preexisting surface or structure for its formation is woven bone. Woven bone is formed during early stages of development, and in adults it is encountered only in pathological conditions, including bone fracture repair. It consists of randomly oriented, small-diameter, highly mineralized collagen fibers. It is not a strong form of bone because of its unordered structure and, particularly, because much of its volume is unmineralized. Its value lies in its ability to rapidly bridge the gap between broken ends of bone to act as temporary scaffolding during the process of repair. It is later removed by bone remodeling events.

Lamellar bone is the dense hard material that constitutes most of the skeleton. In a typical bone, such as the femur, lamellar bone has two distinct types of organization. In the shaft (diaphysis) the material is deposited in layers to form compact or cortical bone. A transverse section through the diaphysis would show a tree-ring-like arrangement of bone. Bone cells (osteocytes) occupy spaces between adjacent lamellae. They interconnect with vascular spaces and one another by fine cellular extensions running in narrow channels (canaliculi). These cells maintain the integrity and normal functioning of bone. Two other main cell types important in normal bone are the bone-forming cells (osteoblasts), and the bone-destroying or -resorbing cells (osteoclasts).

Trabecular (spongy) bone is another form of lamellar bone found particularly in the expanded ends (epiphyses) of long bones and in the vertebrae.

Although it is made of the same material as cortical bone, the mechanical characteristics of trabecular bone differ as a result of the honeycomb arrangement of bone into interlacing struts (trabeculae). This spongy bone is less dense and stiff than cortical bone, and has a large surface area which makes this tissue important for the exchange of calcium between the skeleton and the bloodstream.

Plexiform or fibrolamellar bone is a medium-density, relatively strong bone. It is found at sites of rapid growth where strength is also needed, such as in the limbs of large, fast-growing animals, which include humans during the growth phase in puberty.

Haversian systems or secondary osteons are structural units of bone that are formed secondarily within preexisting bone. This intracortical remodeling involves the formation of a tunnel through bone, achieved by the action of osteoclasts generating a resorption cavity. Following a reversal phase, in which a cement lining is deposited around the perimeter of the cavity, the tunnel is refilled with concentric rings of new bone by osteoblastic activity. A small-diameter central canal containing blood vessels is left unfilled. A completed secondary osteon is a branched structure. Secondary osteons form throughout life and are implicated in skeletal adaptation, repair processes, and mineral exchange between blood and bone (**Fig. 1**). *See* BONE.

Development of the skeleton. All the components of the skeleton are derived from mesenchyme of either mesodermal or neural crest ectodermal origin. The majority of bones in the human body develop by the process of endochondral ossification; that is, the mesenchyme first forms a cartilaginous model which is subsequently replaced by true bone. The facial bones and certain bones of the cranium are formed by intramembranous ossification, in which mesenchyme is converted into skeletal elements by forming bone directly, without need of a cartilage stage. Sesamoid bones are specialized intramembranous bones that form within tendons.

Ossification. The development of bone from the embryo to the adult depends on the orderly processes of mitotic division, growth, and remodeling. These are determined largely by genetics, but are strongly influenced by hormonal action and nutrition. In endochondral ossification, the cartilaginous model is gradually calcified, resulting in cartilage cell (chondrocyte) death. Osteoblasts, together with a blood supply, invade the model and begin to secrete osteoid, which subsequently mineralizes and forms a primary ossification site. This process is repeated in the epiphyses of long bones to form secondary ossification sites. It is normal for the primary ossification site to form earlier, grow more rapidly, and cease ossification later than the secondary sites. The locations and ages at which ossification sites are active in the developing human skeleton are well documented. Hand and wrist x-radiographs are used to determine the skeletal age of individuals because of the orderly progression of ossification in these regions. Such information is important for checking whether an individual's growth rate is abnormal and whether hormone treatment is indicated. Applications also exist in the forensic field, where age at death can be accurately predicted from skeletal remains.

Two parts of developing bones remain as cartilage: epiphyseal plates and articular cartilage. Epiphyseal plates are situated between the diaphysis and the epiphysis. Longitudinal bone growth occurs by chondrocyte proliferation on the diaphyseal sides. The plates are finally obliterated by an extension of diaphyseal ossification into these regions, thus preventing further growth in the length of the bone. Articular cartilage never ossifies except in pathological situations such as osteoarthritis. *See* ARTHRITIS.

Changes in the radial dimensions of long bones occur by new material being deposited beneath the periosteum, a connective tissue membrane surrounding the bone. The endosteal membrane is the equivalent structure lining the internal surfaces of tubular bones. Bone tissue can be added or removed

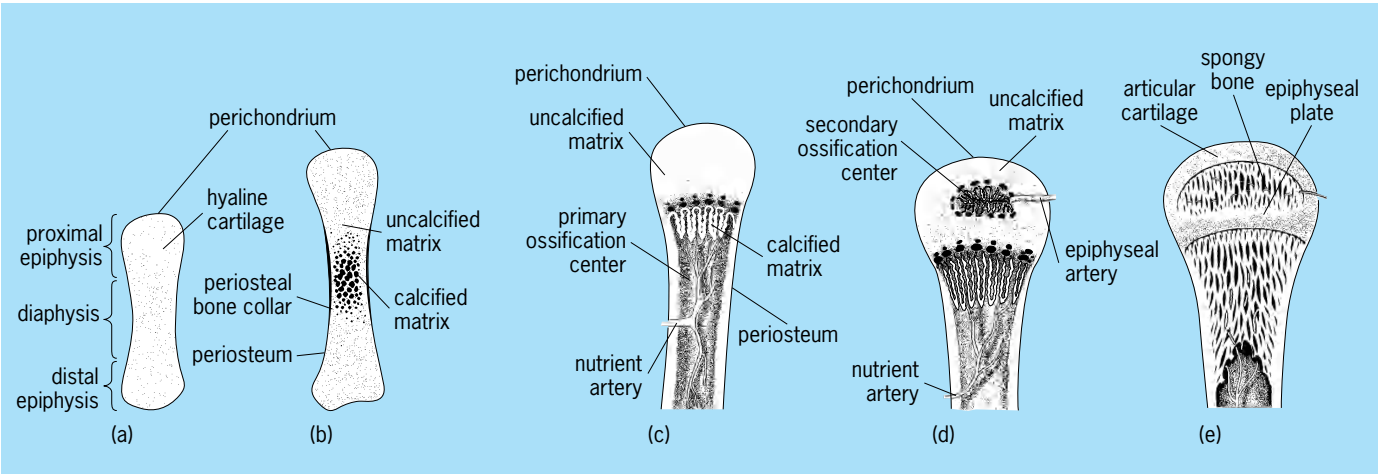

Fig. 1. Development of endochondral ossification in a long bone. (*a*) Mesenchymal cells differentiate into chondroblasts which form the hyaline cartilage model. (*b*) The cartilage model grows, and chondrocytes in the midregion calcify the matrix. (*c*) The primary ossification center and the medullary cavity form. (*d*) Postnatal development of the secondary ossification center occurs. (*e*) Remnants of hyaline cartilage as articular cartilage and epiphyseal plate persist. (*After G. J. Tortora, Principles of Human Anatomy, 6th ed., Harper Collins, 1992*)

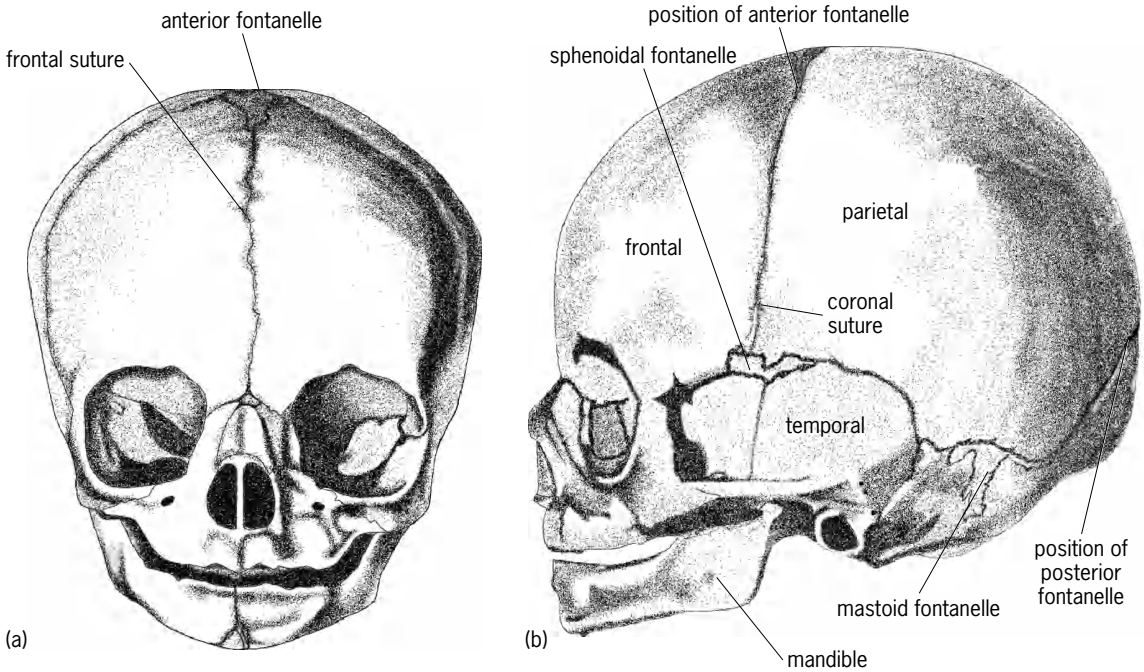

Fig. 2. Skull in a human newborn infant. Cranial sutures and areas of temporary cartilage (fontanelles) form articulations between bones. (*a*) Frontal view. (*b*) Lateral view. (*After G. J. Romanes, Cunningham's Textbook of Anatomy, 10th ed., Oxford University Press, 1964*)

at either site by the action of osteoblasts or osteoclasts, respectively. During growth and aging, there is a tendency for net bone deposition periosteally, with net resorption endosteally.

Bone growth and fracture repair can be increased by mechanical and electrical stimulation. Mechanical deformation of bone causes fluid flow through microscopic channels in bone, resulting in fluid shear stresses on cell membranes and/or production of strain-generated electrical potentials. Both effects have been shown to have osteogenic effects, with osteocytes responding to fluid flow changes, and osteoblast-mediated bone deposition occurring subsequent to mechanical loading in which strain-generated electrical potentials are recorded. In addition, electrical stimulation alone can promote osteogenesis under a wide range of conditions, including nonunion bone fractures and the maintenance of bone mass even in the absence of mechanical function. The mechanisms appear to involve insulinlike growth factors and/or transforming growth factor-β. Elevated levels of both accompany electrical stimulation, and insulinlike growth factors are capable of stimulating proliferation and differentiation of osteoprogenitor cells. Transforming growth factor-β is known to have important effects on bone formation, osteoblast proliferation and differentiation, and matrix synthesis. Furthermore, in vertebrates there are transforming growth factor-β related signaling proteins (bone morphogenetic proteins and activin) that effectively regulate animal development, can induce cartilage and bone formation, and play critical roles in modulating mesenchymal differentiation.

Joints. In the human body there are fibrous, cartilaginous, and synovial joints. Functionally, these joints can be considered as immovable (synarthrosis), partly movable (amphiarthrosis), and freely movable (diarthrosis) joints, respectively.

Immovable joints are represented by cranial sutures and epiphyseal plates prior to their ossification. Little or no movement is available at these joints. Instead, their primary function is to allow bone growth at their margin (**Fig. 2**).

Partly movable joints include the pubic symphysis that joins the two halves of the pelvis anteriorly and the fibrocartilaginous intervertebral discs. While the amount of movement between adjacent vertebrae is quite limited by disc stiffness, the vertebral column as a whole is quite flexible. Flexibility is achieved by combining each of the small movements permitted by individual discs.

Freely movable joints are the most complex and varied of the three types of joint, with their sizes and shapes matched to the functional requirements of the skeletal system at each location. At each joint, the surfaces of opposing bones are covered with a layer of articular cartilage that is a few millimeters thick. The joint is enclosed within a flexible joint capsule, the internal surface of which is lined by the synovial membrane that secretes the lubricating synovial fluid into the joint space. The main functions of articular cartilage are to distribute compressive loads over a wide area in order to reduce contact stresses, and to allow relative movement of opposing joint surfaces with minimum wear and friction. The combination of articular cartilage and synovial fluid gives these joints these remarkable properties. *See* JOINT (ANATOMY).

Articular cartilage. Collagen fibrils, enmeshed within a proteoglycan solution, form the superficial part of articular cartilage. Beneath the relatively soft

articular cartilage lies an interface (the tidemark) that indicates the start of the calcified cartilage layer. Beneath the cartilage lies a layer of subchondral bone which itself lies over the trabecular bone of the epiphysis. A gradient of increasing stiffness exists from the articular surface to the trabecular bone that helps protect the cartilage from splitting under loading conditions.

Synovial joints are subjected to an enormous range of forces. For example, the joints in the lower limb of humans have to support transient forces between two to five times body weight during running and jumping, and moderate but prolonged loading during standing. The nearly frictionless operation of joints is thought to be a function of lubricating films of synovial fluid between, and an adsorbed boundary lubricant on, the articular cartilage surfaces. The ability of cartilage to resist deformation when subjected to high stresses resides in complicated interactions between collagen fibrils, water, and proteoglycans, particularly keratin sulfate, chondroitin sulfate, and hyaluronic acid.

Damage to cartilage is thought to be a function of both the magnitude of applied forces and the speed at which forces are applied (with impact loading causing more damage). Due to its avascular nature, articular cartilage has a limited capacity for repair of such damage.

Vertebrate Skeletal Anatomy

The vertebrate skeleton consists of the axial skeleton (skull, vertebral column, and associated structures) and the appendicular skeleton (limbs or appendages). The basic plan for vertebrates is similar, although large variations occur in relation to functional demands placed on the skeleton.

Axial skeleton. The axial skeleton supports and protects the organs of the head, neck, and torso, and in humans it comprises the skull, ear ossicles, hyoid bone, vertebral column, and rib cage.

Skull. The adult human skull consists of eight bones which form the cranium, or braincase, and 13 facial bones that support the eyes, nose, and jaws. There are also three small, paired ear ossicles—the malleus, incus, and stapes—within a cavity in the temporal bone. The total of 27 bones represents a large reduction in skull elements during the course of vertebrate evolution. The three components of the skull are the neurocranium, dermatocranium, and visceral cranium. *See* EAR (VERTEBRATE).

The brain and certain sense organs are protected by the neurocranium. All vertebrate neurocrania develop similarly, starting as ethmoid and basal cartilages beneath the brain, and as capsules partially enclosing the tissues that eventually form the olfactory, otic, and optic sense organs. The basal and ethmoid plates expand to meet the olfactory and otic capsules to form a floor on which the brain rests. The optic capsule becomes the fibrous sclerotic coat of the eyeball and remains free to move independently of the skull. Further development produces cartilaginous walls around the brain. Passages (foramina) through the cartilages are left open for cranial nerves and blood vessels. The largest opening is the foramen magnum, through which the spinal cord passes. Endochondral ossification from four major centers follows in all vertebrates, except the cartilaginous fishes.

It is likely that the bones of the dermatocranium derive from the dermal armor (bony scales) of early fishes. In modern vertebrates, these bones form via intramembranous ossification from subdermal mesenchyme.

The basic structure from which the tetrapod dermatocranium probably evolved is seen in the primitive crossopterygian fishes and early amphibians. Four kinds of bones are involved. (1) Roofing bones form above and on each side of the brain and neurocranium. These comprise paired nasal, frontal, parietal, and postparietal bones that cover the dorsal parasagittal area, while prefrontal, postfrontal, postorbital, infraorbital, and lacrimal bones form a ring around the orbit. The posterolateral part of the skull is formed from the paired intertemporal, tabular, supratemporal, squamosal, and quadratojugal bones. (2) Dermal, tooth-bearing bones (premaxillae and maxillae) form the margins of the upper jaw. (3) Bones of the primary palate form the roof of the mouth in lower tetrapods and the oropharynx of fishes, and are the paired vomers, palatines, pterygoids, and ectopterygoids and the unpaired parasphenoid. (4) Opercular bones extend posteriorly from the hyoid arch to cover the gill slits. These are paired bones, represented by large opercular bones and smaller subopercular, preopercular, and interopercular bones (**Fig. 3**). *See* TETRAPODA.

Visceral skeleton. This skeleton of the pharyngeal arches is demonstrated in a general form by the elasmobranch fishes, where all the elements are cartilaginous and support the jaws and the gills. Each pharyngeal arch is typically composed of a number of cartilage elements, most of which support gills, but the first and second arches are modified to function in feeding. The mandibular (first) arch consists of two elements on each side of the body: the palatoquadrates dorsally, which form the upper jaw, and Meckel's cartilages, which join ventrally to form the lower jaw. The hyoid (second) arch has paired dorsal hyomandibular cartilages and lateral, gill-bearing ceratohyals. This jaw mechanism attaches to the neurocranium for support. Variations in this articulation occur between species and effectively determine jaw movement, and hence, the feeding abilities of fishes.

Evolution of neurocranium and dermatocranium. Modern amphibian skulls differ considerably from the primitive forms. Most changes involve the loss of roofing bones, particularly those from around the orbit, and massive reduction of the primary palate. These changes allow the eyes, which normally protrude from the head for good vision, to retract until they bulge into the oral cavity.

Early reptiles maintained the primitive pattern of the skull roof, with loss of only the intertemporal bone. Later reptiles exhibited large losses of bony elements, mostly from the temporal region and the back margin of the skull. Major changes in the roof

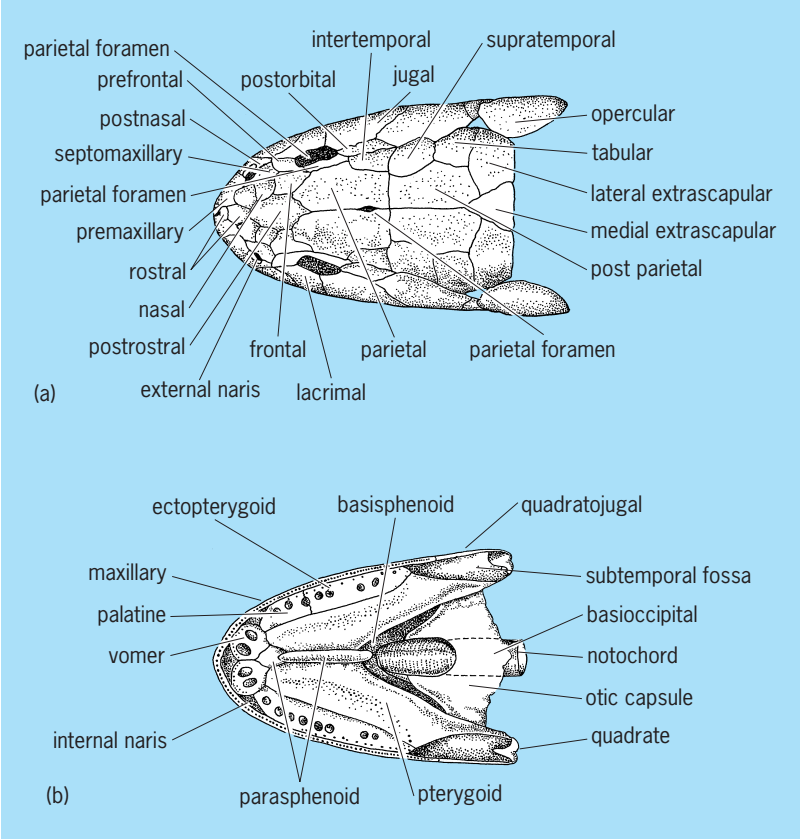

Fig. 3. Skull of *Eusthenopteron*, a generalized fossil crossopterygian fish from Devonian time. (*a*) Dorsal aspect. (*b*) Ventral aspect. (*After A. S. Romer, The Vertebrate Body, 3d ed., Saunders, 1962*)

of the skull are associated with the development of temporal fenestrae, resulting in four recognized design types: anaspids exhibit a primitive condition with a solid skull roof (turtles and stem reptiles); diaspids have two openings on each side of the skull (archosaurs and the ancestors of lizards and snakes); euryapsids have a single, dorsal opening on each side (ichthyosaurs and plesiosaurs); and synapsids have a single, lateral opening (pelycosaurs and therapsids). Birds derive from archosaurian reptiles and have, essentially, a diaspid skull design. The pelycosaur stock, with a synapsid skull, gave rise to the mammals.

A secondary palate appears for the first time in the reptiles as a horizontal shelf of bone dividing the oral cavity into separate nasal and oral passages. The remaining bones of the more primitive primary palate are located in the nasal passageway. Only the crocodilians, among the reptiles, have a complete secondary palate formed from medial extensions of the maxillae, premaxillae, pterygoid, and palatine bones. *See* REPTILIA.

The avian skull has been modified from the diaspid reptilian type in line with increased brain size and changes in life-style, particularly feeding and flight behavior. Premaxillae and dentary bones form the majority of the upper and lower beak, respectively. Some birds can raise the upper portion of the beak relative to the rest of the skull by a hinge mechanism between the frontal bone and the premaxillae, maxillae, and nasal bones. This is one of the many forms of cranial kinesis, where motion (usually of the upper jaw) is possible, partially independent of other parts of the skull. Most vertebrates other than mammals are capable of cranial kinesis, which is a feeding adaptation. *See* AVES.

The therapsids are mammallike reptiles that evolved large temporal fenestrae in the dermal bones of the cheek region. Other notable changes involved the loss, fusion, or reduction of roofing bones and a large variation in bone proportions and sizes. A secondary palate formed from the premaxillae, maxillae, and palatines separates the nasal and oral cavities in a manner that was similar to the crocodilians. This palate continues posteriorly as a fold of skin, the soft palate. The value of a divided airway is that it gives the animal the ability to breathe while the mouth is full. *See* THERAPSIDA.

Notable increases in the size of the brain during mammalian evolution have resulted in considerable changes to the neurocranium. The zygomatic arches are the true remnants of the original lateral walls of the skull. The temporal bone, as seen in humans, is a single unit, but is really a composite of a large number of fused skeletal elements of endochondral and membranous origin. The occipital bones are usually fused, and the single occipital condyle, present in reptiles and birds, divides to become a paired structure as in modern amphibians. Large changes occur in the organization of the otic capsule, with the inclusion of a new skull bone (the entotympanic) and dermal bone (derived from the angular bone of the reptilian jaw) in a protective role for the middle ear. *See* MAMMALIA.

Evolution of visceral cranium. In all jawed vertebrates except mammals, an articulation between the posterior ends of the palatoquadrate and Meckel's cartilages (which may be ossified or ensheathed in bone) occurs between the upper and lower jaws (**Fig. 4**). The bony fishes have elaborated on the primitive condition, where the upper jaw was fused to the skull and the lower jaw or mandible could move only in the manner of a simple hinge. Teleosts are able to protrude the upper and lower jaws. This motion, coupled with expansion of the buccal cavity, enables these fishes to generate the large suction forces used to draw food into the mouth.

In the course of mammalian evolution, the dentary of the lower jaw enlarged and a ramus expanded upward in the temporal fossa. This eventually formed an articulation with the squamosal of the skull. With the freeing of the articular bone (seen as an ossified posterior end of Meckel's cartilage in teleosts) and the quadrate from their function in jaw articulation, they became ear ossicles in conjunction with the columella, that is, a skeletal rod that formed the first ear ossicle, arising in the amphibia to conduct sound waves from the eardrum to the otic capsule.

The remaining visceral skeleton has evolved from jaw and gill structures in the fishes to become an attachment site for tongue muscles and to support the vocal cords in tetrapods.

Vertebral column, ribs, and sternum. The vertebral column is an endoskeletal segmented rod of mesodermal origin. It provides protection to the spinal

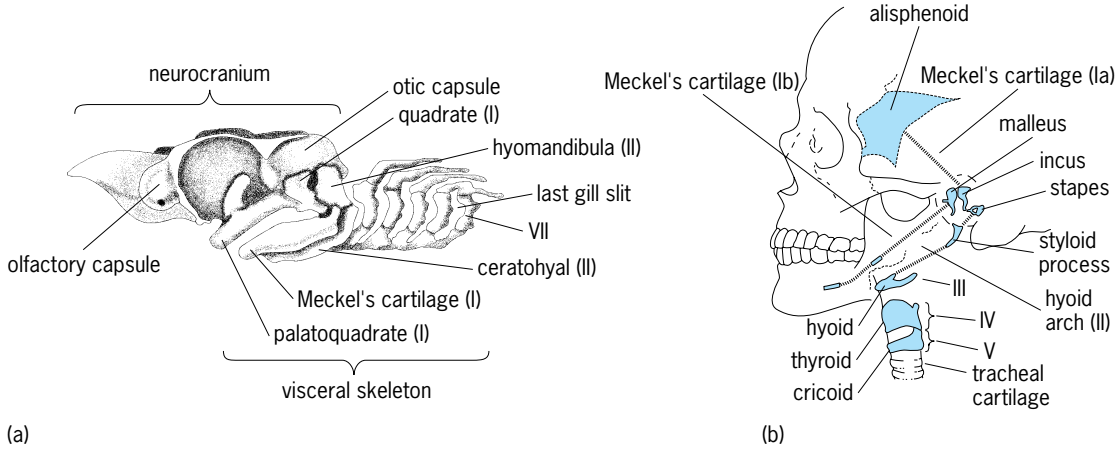

Fig. 4. Skull and visceral skeleton. (*a*) Shark. The skeleton of the first, second, and seventh pharyngeal arches is indicated. (b) Human. Shaded regions show the components of the human skull and neck that are derived from the visceral skeleton. (*After G. C. Kent, Comparative Anatomy of the Vertebrates, 7th ed., Mosby-Year Book, 1992*)

cord, sites for muscle attachment, flexibility, and support, particularly in land-based tetrapods where it has to support the weight of the body (**Fig. 5**).

1. Vertebrae. Hard, spool-shaped bony vertebrae alternate with tough but pliable intervertebral discs. Each typical vertebral body (centrum) has a bony neural arch extending dorsally. The spinal cord runs through these arches, and spinal nerves emerge through spaces. Bony processes and spines project from the vertebrae for the attachment of muscles and ligaments. Synovial articulations between adjacent vertebrae (zygapophyseal joints) effectively limit and define the range of vertebral motion.

Vertebral morphology differs along the length of the column. There are two recognized regions in fishes (trunk and caudal) and five in mammals (cervical, thoracic, lumbar, sacral, and caudal), reflecting regional specializations linked to function. Humans have seven cervical, twelve thoracic, five lumbar, five (fused) sacral, and four coccygeal vertebrae. Most amphibians, reptiles, and mammals have seven cervical vertebrae regardless of neck length (giraffes have only seven), whereas the number is variable in birds. Specific modification to the first two cervical vertebrae in most reptiles, birds, and mammals gives the head extra mobility. In humans, the occipital condyles articulate with the atlas, which in turn articulates with the axis, all via synovial joints. The atlas is a ringlike vertebra whose centrum is represented as a peg of bone fused to the axis. This odontoid process provides a pivot about which the atlas, and hence the head, swivels. The nodding movement is provided by rotational motion at the atlanto-occipital joint.

The presence of large ribs in the thoracic region often limits spinal flexibility. Birds utilize fusion of thoracic vertebrae to provide a rigid structure, adapted for the demands of powered flight. Additionally, fusion of the lumbar, sacral, and proximal caudal vertebrae produces a rigid back, the synsacrum.

In typical tetrapods, the sacral region is usually modified for support of the pelvic girdle, while the number of caudal vertebrae varies greatly (from 0 to 50) between and within animal groups. Birds have a specialized end cap to the caudal vertebrae (pygostyle) for the attachment of feathers. *See* VERTEBRA.

2. Sternum and ribs. Jawed fishes have ribs that help maintain the rigidity and support of the coelomic cavity. These ribs typically follow the connective tissue septa that divide successive muscle groups. In the caudal region, they are often small paired ventral ribs, fused on the midline to form the haemal arches. Ancestral tetrapods had ribs on all vertebrae, and their lengths varied between the vertebral regions. Modern amphibia (frogs and toads) have few thoracic ribs, and these are much reduced and never meet ventrally. Reptiles have varied rib arrangements, ranging from snakes with ribs on each vertebra (important for locomotor requirements) to turtles with only eight ribs which are fused to the inside of the carapace. Flying birds and penguins have a greatly enlarged sternum that links the ribs ventrally. This provides a large surface area for the origin of the powerful pectoralis muscles that are used to power the downstroke of the wing. In humans there are twelve pairs of ribs which form a strong but movable cage encompassing the heart and lungs.

Appendicular skeleton. This section of the skeletal system comprises the pectoral and pelvic limb girdles and bones of the free appendages. The girdles provide a supporting base onto which the usually mobile limbs attach.

Pectoral girdle. The pectoral girdle has both endoskeletal and dermal components. The dermal components are derived from postopercular dermal armor of primitive fishes, and are represented by the clavicles and interclavicles in modern vertebrates, except where they are secondarily lost. Endochondral bone forms the scapula. In fishes, the main component of the girdle (the cleithrum) is anchored to the skull by other bony elements. Increased mobility of the girdle is seen in amphibia as it becomes independent of the skull. Further development and skeletal reduction have resulted in a wide range of morphologies, culminating in the paired clavicles and scapulae of mammals.

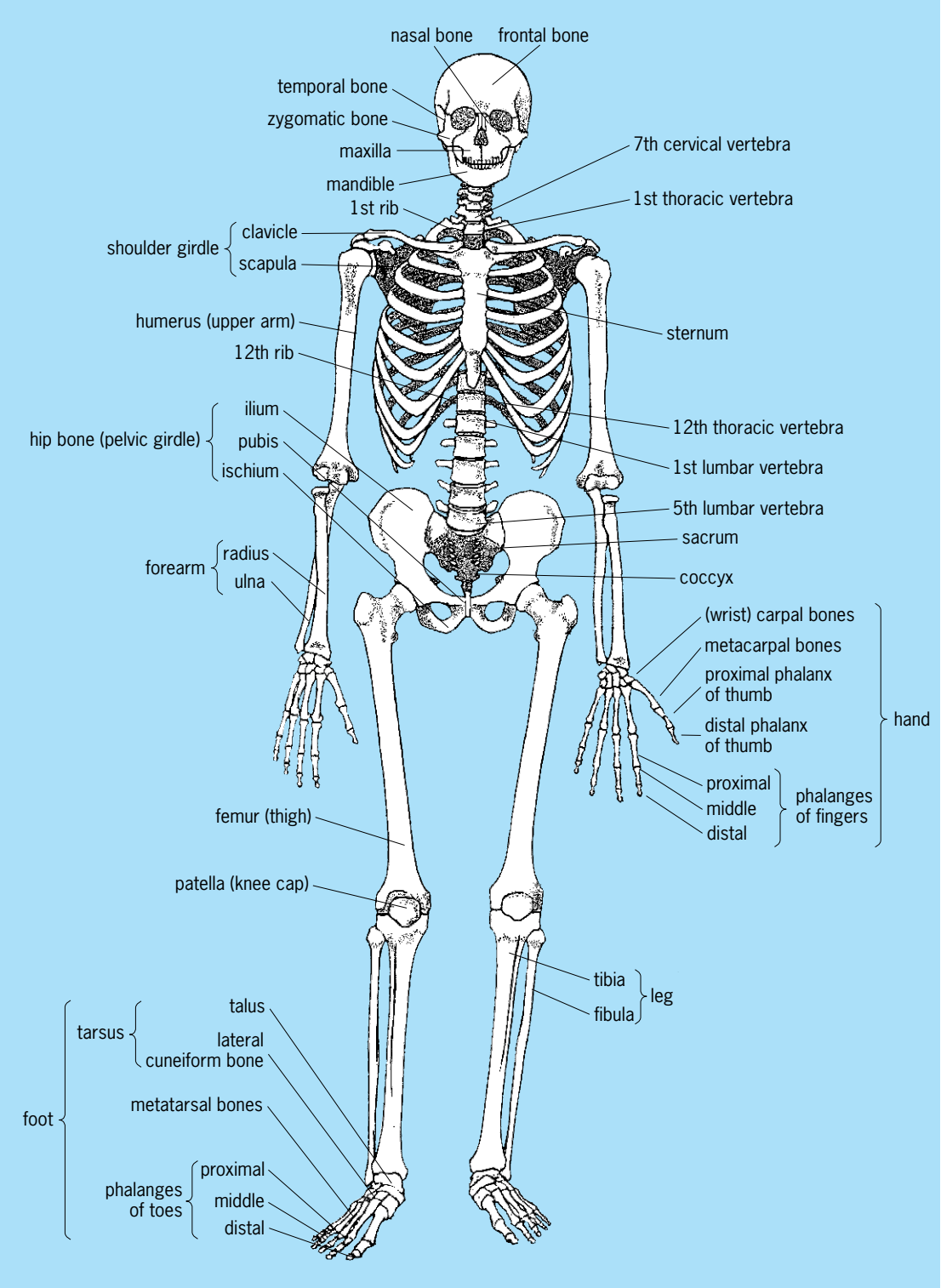

Fig. 5. Human skeleton (anterior view). (***After G. J. Romanes, ed., Cunningham's Textbook of Anatomy, 10th ed., Oxford University Press, 1964***)

Birds have fused their paired clavicles and single interclavicle to form the wishbone or furcula. This provides an increased site for the origin of the pectoral muscles, both directly and from the fascial sheet spanning the two arms. Bending of the furcula during flight is thought to assist breathing by compressing or expanding underlying air sacs. It also has a role in storing, and subsequently returning, elastic strain energy at functionally useful parts of the wing beat cycle. Clavicles have disappeared in certain groups of bounding mammals to allow greater movement of the scapula. Although humans, and most other mammals, have a coracoid process on the scapula, other tetrapods typically have a separate coracoid bracing

the scapula against the sternum and forming part of the glenoid fossa.

Pelvic girdle. The pelvic girdle forms by endochondral ossification, that is, the conversion of cartilage into bone. In the fishes, it is a small structure embedded in the body wall musculature just anterior to the cloaca. Each half of the girdle provides an anchor and articulation point for the pelvic fins. In tetrapods, the girdle attaches to the vertebral column to increase its stability and assist in the support of body weight and locomotor forces. Humans, like all other tetrapods, have a bilaterally symmetrical pelvic girdle, each half of which is formed from three fused bones: the ischium, ilium, and pubis. A part of each of these elements forms the acetabulum, the socket-shaped component of the hip joint, that articulates with the femoral head. The human pelvis is bowl-shaped, helping to support the viscera in upright stance. The large ilium (a feature shared with certain heavy-bodied animals such as cattle and horses) is the site of origin of hip extensor muscles, the gluteals. They prevent the body from bending sharply at the hip and also assist powerful thigh muscle extension during the propulsive stage of running.

All urogenital and digestive products have to pass through the pelvic outlet. This accounts for the pelvic sexual dimorphism seen in most mammals, where the pelvic opening is broader in females, because of the physical demands of pregnancy and parturition. In birds (with the exception of the ostrich and the rhea), both sexes have an open pelvic girdle, a condition also found in female megachiropteran bats (flying foxes), gophers, and mole-rats.

Paired fins and tetrapod limbs. Paired fins in fishes come in different forms, but all are involved in locomotion. In the simplest form they are fairly rigid and extend from the body, functioning as stabilizers, but they are also capable of acting like a wing to produce lift as in sharks. In many fishes, the pectoral fins have narrow bases and are highly maneuverable as steering fins for low-speed locomotion. In addition, some fishes (such as the Australian lungfish) use their pectoral and pelvic fins to walk on the river bed, while others have greatly enlarged pectoral fins that take over as the main propulsive structures (for example, rays, flatfish).

The basic mammalian pectoral limb consists of the humerus, radius, ulna, carpals, five metacarpals, and fourteen phalanges (arranged as 2-3-3-3-3, starting at the thumb; reptiles tend to have a 2-3-4-5-3 arrangement); and the pelvic limb consists of the femur, tibia, fibula, tarsal, five metatarsals, and fourteen phalanges (mammals have various digital arrangements; most reptiles have a 2-3-4-5-4 arrangement). Variation in the shape and number of elements occurs primarily in the hand (manus) and foot (pes). A typical bird pelvic limb consists of a femur, tibiotarsus (formed by fusion of the tibia with the proximal row of tarsal bones), fibula, and tarsometatarsus (formed by fusion of metatarsals II–IV), metatarsal I, and four digits (each consisting of two to five phalanges).

Locomotion and skeletal adaptations. Throughout evolution, the skeletal system has adapted to the needs of many different types of organisms. Such adaptations have been made for walking and running, speed, power, digging and burrowing, locomotion without limbs, and aerial and aquatic locomotion.

Walking and running. The undulatory side-to-side flexion of the vertebral column seen in most fishes continued to play a role in early tetrapod locomotion. It still exists in the urodeles (for example, newts) and many reptiles where limbs project laterally, with the proximal element swinging in the horizontal plane. The feet provide points of firm contact with the ground in this sprawling locomotor style, with forward progression powered by lateral trunk flexion. Most terrestrial tetrapods have raised the trunk off the ground, with the limbs projecting ventrally to support the body. Changes in joint alignment resulted in the limbs and trunk flexing and extending in the sagittal plane.

Adaptations for speed. Animals use various recognizable forms of locomotion, or gaits, for traveling at different speeds. The voluntary selection of a particular gait appears to be linked to reducing locomotor forces on the limb skeleton and minimizing the energetic cost of travel.

Obvious modifications to the basic tetrapod skeleton have accompanied the acquisition of high-speed locomotion. The lengthening of limb segments allows for longer strides which, coupled with stride frequency, determine running speed. Limb elongation has been accompanied by a reduction in the number of skeletal elements in running specialists (**Fig. 6**).

The normal pentadactyl (five-digit) limb is seen in plantigrade animals, such as primates and insectivores, where the ankles, wrists, and all bones distal to these joints contact the ground. Humans have a bipedal plantigrade posture. A plantigrade posture typically provides stability (large foot-ground contact area) and good manual or pedal dexterity (useful for climbing, and holding and manipulating objects), but does not provide particularly good running performance (**Fig. 7**).

Many of the Carnivora, Rodentia, and Lagomorpha have either lost or reduced the first digit. They are digitigrade; that is, they walk and run on four toes, with the metacarpals and tarsals raised off the ground. The functional increase in limb length contributes to their fleetness of foot, useful for catching prey or avoiding capture; and the reduction, particularly of the metacarpals, metatarsals, and digits, can be viewed as an adaptation for speed. *See* CARNIVORA; LAGOMORPHA; RODENTIA.

Hoofed (unguligrade) animals have further reduced their digits. Even-toed ungulates (sheep, cattle, camel) and odd-toed ungulates (horse, rhinoceros) walk on the tips of the remaining digits. Claws or nails are often expanded to form protective structures such as the horse's hoof. The remaining metatarsals and metacarpals are often very elongate. Digital reduction is extreme in horses, where a single digit (digit III, equivalent to the human middle finger or toe) remains. All limb reduction results in loss

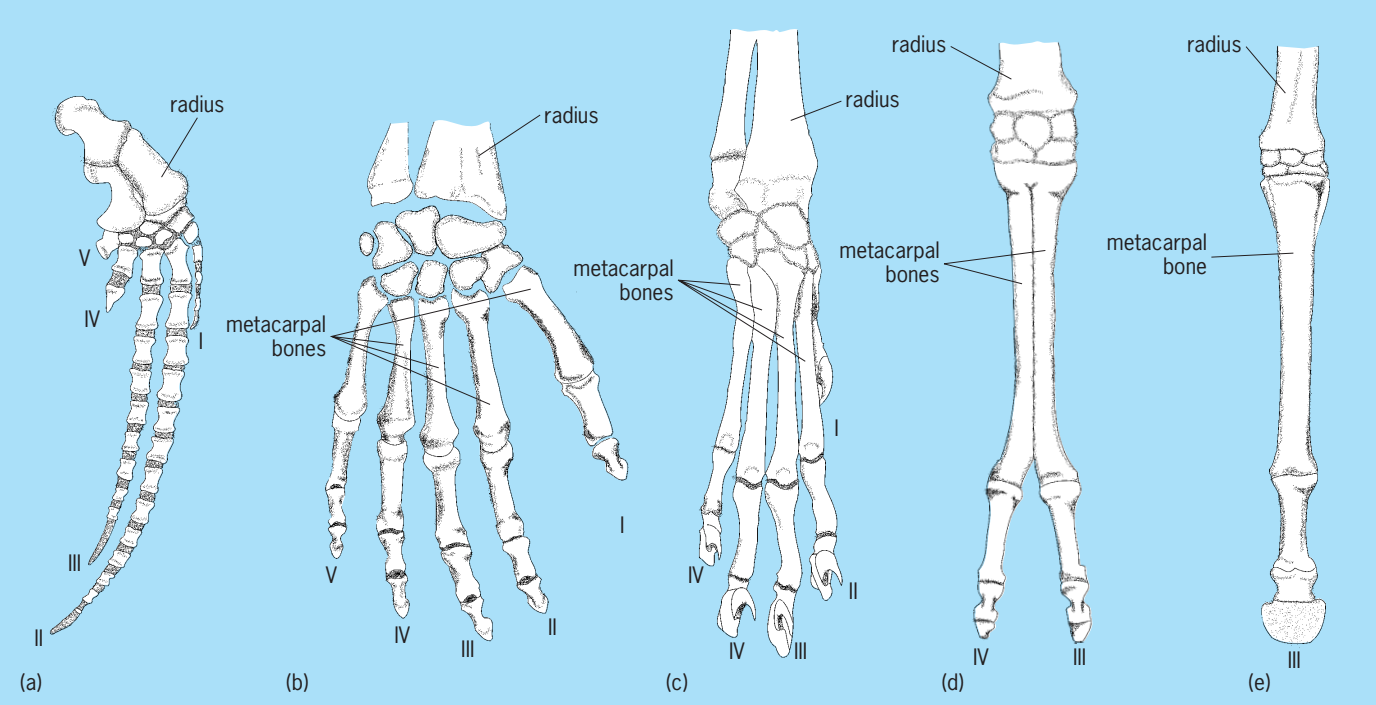

Fig. 6. Changes in the tetrapod distal limbs in relation to different locomotor requirements. (*a*) Whale (*after A. S. Romer and T. S. Parsons, The Vertebrate Body, 6th ed., CBS College Publishing, 1986*). (*b*) Human; (*c*) cat; (*d*) camel; (*e*) horse (*after G. C. Kent, Comparative Anatomy of the Vertebrates, 7th ed., Mosby-Year Book, 1992*)

of dexterity, and the horse, as an extreme example, cannot use its limbs for any function other than locomotion. However, digital reduction is extremely important for economical locomotion: multiple bones weigh more than a single element of equal strength, and thus limb mass can be reduced by reducing the number of bones without compromising limb strength.

There is a similar trend in birds. The ostrich is the only bird that has two toes (didactyl); rheas, cassowaries, emus, and others have three (tridactyl), whereas most species have four. However, variation in toe number, size, and orientation reflects more than just locomotor behavior because there are three-toed species (kingfishers, parrots) that are not runners.

Other skeletal adaptations for speed include the loss of clavicle in cats, allowing greater scapular movement which results in greater reach (greater stride length when running) and helps reduce impact forces to the axial skeleton when running; and increased flexibility of the vertebral column in the sagittal plane—another way to increase stride length.

Speed versus power. The speed of movement of limb bones is of fundamental importance for running animals. This depends on the sizes and properties of muscles and where they attach relative to the joints. A muscle that is attatched close to a joint has a small moment arm (the perpendicular distance from the line of action of the force to the joint center) and will be able to move the distal end of the bone rapidly when the muscle contracts. The larger its moment arm about the joint, the slower the movement. However, the opposite is true for exertion of force at the distal end of the bone. Thus, skeletal morphologies differ, depending on whether predominantly powerful or rapid movements are required.

Digging and burrowing. Most changes in hole-dwelling animals are related to muscular insertions for increased power in their movements. However, skeletal adaptations occur in the teeth, pectoral and pelvic limbs, and vertebral column. Major modifications include limbs becoming shorter and more robust. The

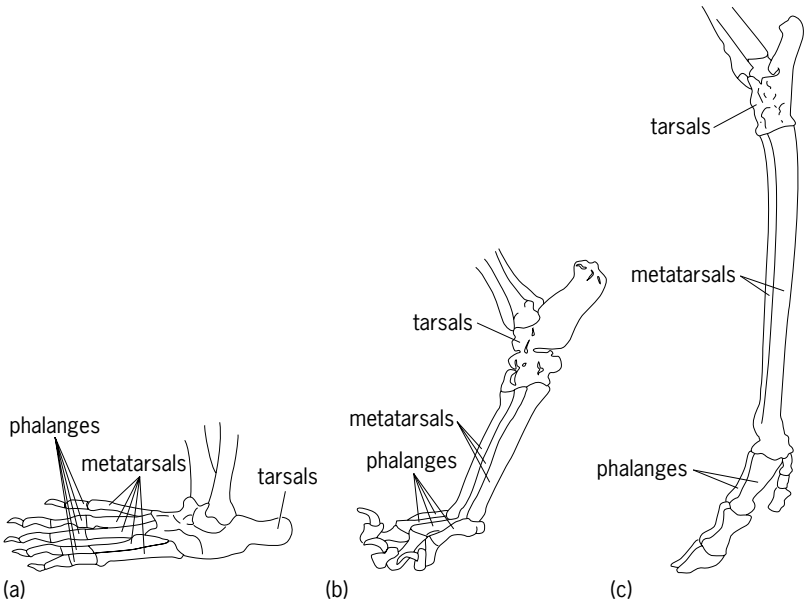

Fig. 7. Limb postures. (*a*) Plantigrade. (*b*) Digitigrade. (*c*) Unguligrade. (*After G. C. Kent, Comparative Anatomy of the Vertebrates, 7th ed., Mosby-Year Book, 1992*)

medial humeral epicondyle enlarges to accommodate an increased muscle bulk of carpal and digital flexor muscles and forearm pronators. Claws are either sharp to break hard soil or flattened to help shift soft soil, and often have large extensor tubercles on the dorsal surfaces. These act as passive bony stops to prevent overextension of the joints. Many joints are restricted to planar motion (flexion-extension) to increase their stability and make them less likely to dislocate.

Limbless locomotion. Frogs, birds, and terrestrial mammals have never produced limbless forms. Snakes (Ophidia) are the most diverse limbless group, but there are lizards (Sauria) and amphibia (Apoda) that show a reduction or loss of one or both limb pairs and girdles. The main adaptation accompanying limb loss is elongation of the body. More than 500 vertebrae have been documented in a snake. The overlapping bony scales used for protection and to gain purchase for movement are formed secondarily in the skin.

Aerial locomotion. Birds and bats (and possibly the extinct pterosaurs) are the only examples of vertebrates capable of continuous muscle-powered flight. Many others use the more passive gliding or parachuting. Flying fishes (for example, Exocoetidae, Hemirhamphidae) have enlarged pectoral fins enabling them to glide out of water to escape predators. A number of reptiles (for example, *Draco* spp.) glide with the aid of flight surfaces formed by elongated ribs.

Birds have a highly modified forelimb complex (**Fig. 8**). The humerus, ulna, and radius have surface features linked to flight, but the real adaptation of these (and other bird and pterosaur bones) is their air-filled or pneumatic character. This reduces their mass and, therefore, the muscular work involved in moving the wings up and down. In addition, the major wing bones of bats, birds, and pterosaurs have relatively large diameters and are relatively thin-walled. These traits maximize torsional strength for any given mass of bone material, which is important for flying animals in which twisting forces on the wing skeleton are commonly encountered. Support of the primary flight feathers is provided by the bones of the carpus, carpometacarpus (a fusion of carpals and metacarpals unique to birds), and manus, which in adult birds has only three digits (recently identified as digits II-III-IV). Wing length is effectively determined by the lengths of the arm, forearm, and hand skeleton, but wing shape (breadth and sharpness of wingtips) is determined by feather morphology. Secondary adaptations to flight are seen in the pelvic girdle, with its extensive connection to the synsacrum presumed to be linked to absorbing the shock of landing.

Bats have adopted a different skeletal strategy. They use highly elongated bones (humerus, radius, metacarpals, phalanges) to help support the whole wing, not only the leading edge as seen in birds. These bones are not pneumatic, but contain marrow. Weight savings do occur in distal elements due to a reduction in mineralization. Large pectoral muscles attach to ribs, and there is no massive, keeled sternum as in birds. Pelvic girdle reduction, thinner

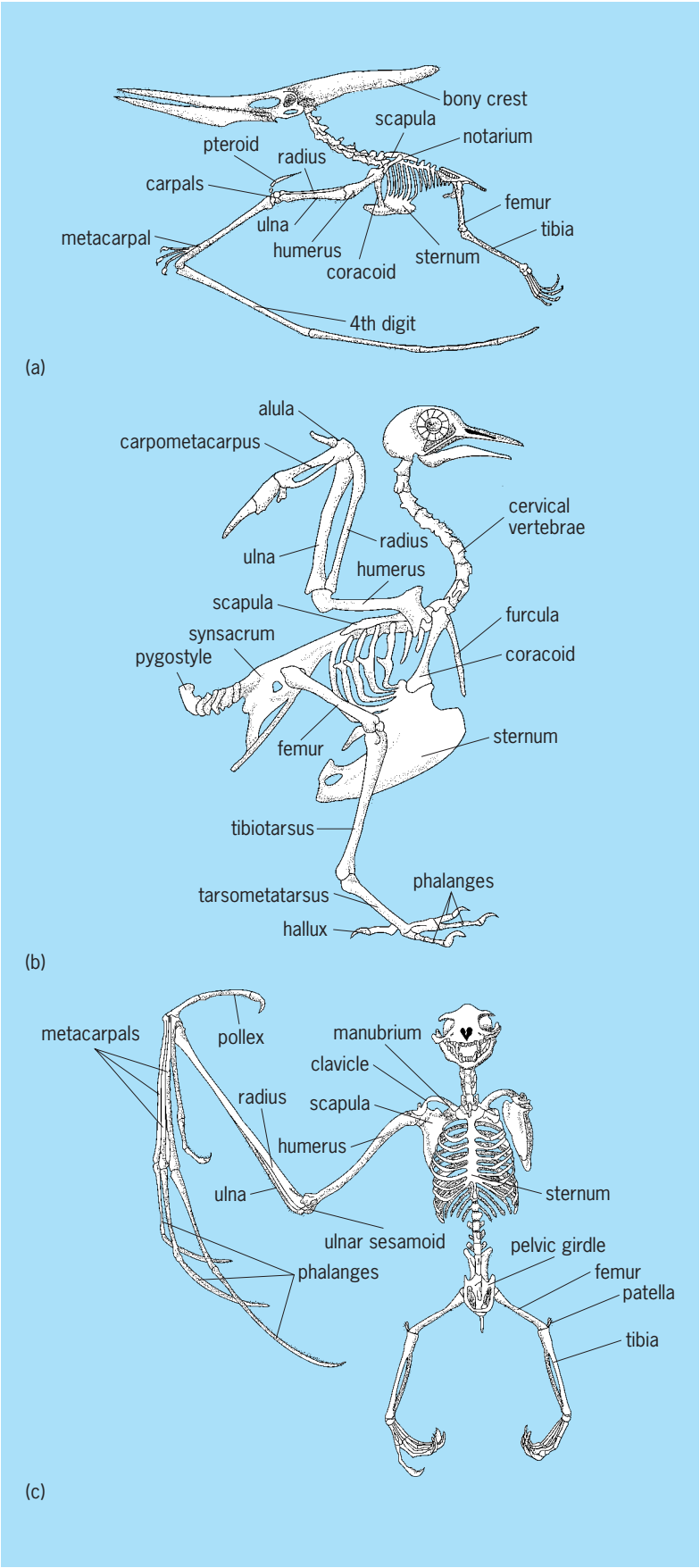

Fig. 8. Skeletons of flying vertebrates. (*a*) Extinct pterosaur (*Pteranodon*). (*b*) Typical bird skeleton. (*c*) Skeleton of a fruit bat or flying fox. (*After L. B. Halstead, Vertebrate Hard Tissues, Wykeham Publications, London, 1974*)

and straighter limb bones (they no longer have to resist large bending forces), rotation of the limb so the patella faces posteriorly, partial loss of the fibula, and possession of large claws are all skeletal adaptations for a hanging life-style. *See* FLIGHT.

Aquatic locomotion. In contrast to terrestrial locomotion, the limbs of aquatic organisms are hydrofoils or paddles. The skeletal elements are usually the same as typical tetrapods, with altered dimensions. Some specialist swimmers (such as porpoises and whales) have increased the number of phalanges in the hand. The manus is fairly stiff, the bones being bound together with connective tissue. Reduction or loss of the pelvic girdle and limbs has occurred in cetaceans and sirenians, and thrust is supplied by dorsal-ventral flexion of the vertebral column. Seals and walruses retain the pelvic structures, swimming by lateral flexion of the vertebral column.

Skeletal strength. Bones have to be stiff in order to act as levers for effective muscle action, and strong to resist failure. Although the form of the skeleton is genetically determined, the basic plan can be modified within limits by hormone action and the mechanical environment.

The limb skeleton of terrestrial tetrapods is subjected to large dynamic forces during rigorous exercise. For example, the feet of human runners generate loads of two to three times body weight. Even larger forces are applied to the skeleton by muscle action (for example, to the heel bone via the Achilles tendon). Direct measurement of bone deformation under such conditions shows that the limb bones in most animals undergo similar levels of strain (that is, change in specimen length per original length) and are typically three to five times as strong as they need to be to resist failure.

Long bones. Lamellar or cortical bone has the same mechanical properties whether taken from a mouse or an elephant. The strength of whole bones is therefore a function of the amount of bone present and its geometric arrangement. Long bones in tetrapod limbs are typically tubular, even though the marrow contained within the hollow center appears to have little functional use. Why the bones are not solid can be explained by referring to principles of biomechanics. As most bones have a degree of longitudinal curvature, they tend to bend in a predictable direction when loaded, for example, during running. One surface of the bone will tend to be under compression (squashed), and the opposite surface will be under tension (pulled). As there is a gradient across the bone running from compression to tension, there is a position between the surfaces where the forces cancel out (the neutral axis). Bone material along and close to this axis does not contribute much strength to the bone in bending. Such redundant material can be removed (forming the marrow cavity) without compromising bone strength.

Size and the skeleton. The bones of humans and elephants are proportionately more robust than those of mice. This occurs because a doubling in linear dimensions results in a fourfold increase in area and an eightfold increase in mass. If animals simply increased in size, it would result in the bones undergoing larger stresses in bigger animals. To keep bone stresses at the same level as in small animals requires that the bones be made thicker.

Another obvious feature is that small animals adopt crouching (noncursorial) postures, while larger ones are typically cursorial, standing with straighter limbs. A more upright limb posture results in a better alignment of the forces generated during locomotion with the long axes of the bones. This reduces both the forces that muscles need to generate to keep limbs extended, and the bending moments imposed on the bones. Cursorial postures lower stresses on the limb bones but reduce agility. *See* BIOPHYSICS; VERTEBRATA.

Mike Bennett

Bibliography. R. McN. Alexander, *Animal Mechanics*, 2d ed., 1983; J. D. Currev, *The Mechanical Adaptations of Bones*, 1984; M. Hildebrand et al. (eds.), *Functional Vertebrate Morphology*, Belknap Press, Cambridge, MA, 1985; A. S. King and J. McLelland, *Form and Function in Birds*, vol. 3, Academic Press, London, 1985; J. Klein-Nulend et al., Sensitivity of osteocytes to biomechanical stress in vitro, *FASEB J.*, 9:441–445, 1995; R. B. Martin and D. B. Burr, *Structure. Function and Adaptation of Compact Bone*, Raven Press, New York, 1989; F. H. Martini, *Fundamentals of Anatomy and Physiology*, 4th ed., Prentice Hall, Englewood Cliffs, NJ, 1998; M. Nordin and V. H. Frankel, *Basic Biomechanics of the Musculoskeletal System*, 2d ed., Lea and Febiger, Philadelphia, 1989; H. Zhuang et al., Electrical stimulation induces the level of TGF-β1 mRNA in osteoblastic cells by a mechanism involving calcium/calmodulin pathway, *Biochem. Biophys, Res. Comm.*, 237:225–229, 1997.

Skin

The entire outer surface of the body and the principal boundary between the external environment and the body's internal environment of cells and fluids. Skin serves as the primary barrier against the intrusion of foreign elements and organisms into the body, and also as a large and complex sense organ through which animals explore and learn about the external world. In addition, skin functions to maintain the homeostasis of the body's constituents, acting as a barrier to the loss of various ions and nutrients by diffusion. For terrestrial animals, it also serves as an effective barrier to water loss, without which most land animals would rapidly become desiccated and die.

Structure of mammalian skin. The skin of humans and other mammals can be divided into two distinct regions, the epidermis and the dermis (**Fig. 1**). The latter merges on its lower side with a looser connective tissue region called the hypodermis, which anchors the skin to deeper layers of connective tissue and is the principal site of fat deposition under the skin.

Epidermis. The outermost layer of the skin is the epidermis. It varies in thickness from 0.1 mm in

most of the protected areas of the skin to approximately 1 mm in those regions exposed to considerable friction, such as the soles of the feet and palms of the hands. The epidermis consists of a great many horizontally oriented layers of cells. The innermost layers, known as the stratum germinativum, consist of living and rapidly dividing cells. As cell division occurs in this region, the older cells are pushed toward the surface of the skin. In the process of this vertical migration, they become progressively flattened and lose their nuclei and most of their intracellular material. In effect, they die and become empty shells packed with the protein keratin. The outermost layer of the epidermis, the stratum corneum, consists of many layers of this packed cellular debris, forming an effective barrier to water loss from lower layers of the skin. Normally, its water content is exceptionally low, approximately 10% by weight. After soaking, however, the stratum corneum imbibes some external water, resulting in a wrinkled and shrunken (actually swollen) appearance. The outermost layer of the stratum corneum is continually sloughed off as small flakes. The greater the external wear and removal, the greater is the rate of cell division in the stratum germinativum and replacement of lost cells. It takes approximately 1 month for a cell to traverse the epidermis and be sloughed from its surface.

The lowest levels of stratum germinativum constitute the portion of the skin that contains melanocytes, cells that produce the dark pigment melanin. Different levels of melanin secretion are responsible for the large range of pigmentation observed among humans. Melanocytes respond with increased melanin secretion to a variety of stimuli, in particular, to ultraviolet radiation damage, resulting in tanning of the skin.

Dermis. The dermis plays a supportive and nutritive role for the epidermis. The epidermis has no blood supply of its own; that is, it does not contain any capillaries. However, nutrients and oxygen are apparently provided by diffusion from the blood supply of the underlying dermis, which extends in small, fingerlike projections (papillae) into the lower margin of the epidermis (Fig. 1). The dermis contains relatively few cells and numerous collagenous and elastic fibers, which give strength, cohesion, and resilience to the skin. The average thickness of the dermis is 1–3 mm. It is in this layer that the sebaceous and sweat glands are located and in which the hair follicles originate. The hair follicles are formed from downward invaginations of the stratum germinativum of the epidermis around dermal papillae of connective tissue. The surrounding dermis provides structural support and nutrition for the growing hair root. Associated with all hair follicles are oil-producing (sebaceous) glands, which empty their fatty contents into the hair follicle and lubricate the hair shaft. Two types of sweat glands, apocrine and eccrine, also originate in the dermis. Apocrine glands are confined to the axillary and anal-genital areas and produce viscous secretions; subsequent bacterial growth on these secretions is responsible for body odor. Eccrine glands are widely distributed over the human body and produce sweat, which has a thermoregulatory role in humans during exercise or heat stress. Both types of sweat glands have a duct running directly from the main part of the gland through the epidermis and emptying directly on the skin surface. The products of all these sets of glands are derived from the rich blood supply of the dermis. Hair, sweat glands, and mammary glands (which are modified sweat glands) are skin inclusions unique to mammals. *See* HAIR; SWEAT GLAND.

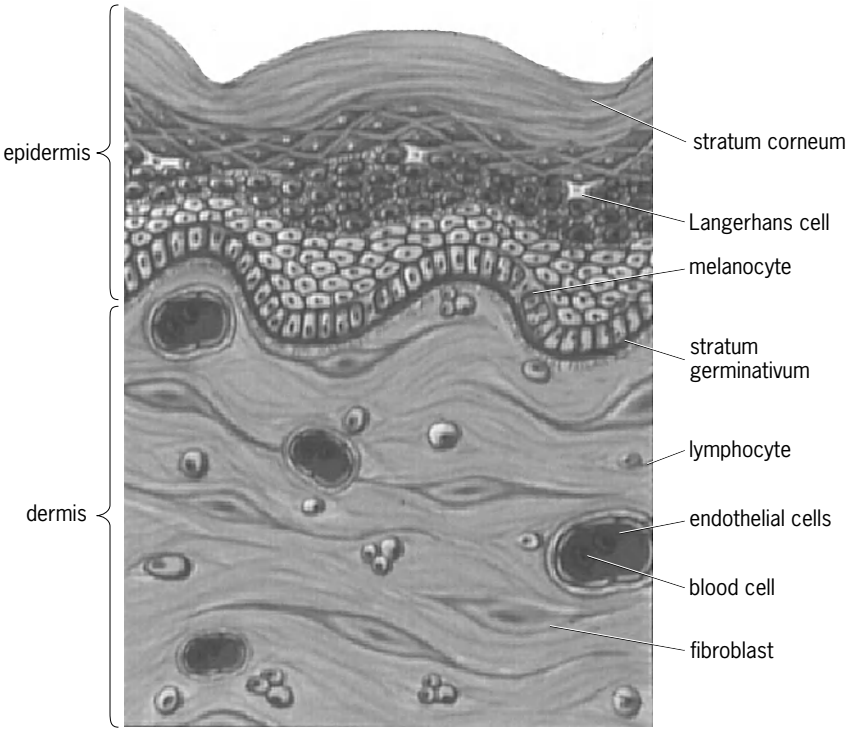

Fig. 1. Cross section of human skin. (*After Organogenesis, Inc.*).

Thermoregulation in mammals. Mammals are endothermic homeotherms; that is, their principal source of body heat is derived from their internal metabolic processes, and they maintain a constant body temperature. The skin plays a primary and essential role in thermoregulation in mammals, because it is across the skin that most heat is lost to or gained from the environment. Changes in skin insulation alone are very effective in adjusting rates of heat loss over a broad range of environmental temperatures, so that no extra metabolic heat has to be produced for thermoregulation. In this thermal neutral zone, insulation is altered by two mechanisms: piloerection and adjustment of dermal blood flow. To accomplish piloerection, a small smooth muscle fiber attaching the hair follicle to the skin contracts, raising the hair vertically. Erect hair traps a still layer of air next to the skin. This trapped air is a very efficient insulator, partially isolating an animal from its external thermal environment. The thicker and longer the hair, the more air is trapped and the more effective the insulation. Human hair on the body is too sparse to form an effective insulatory barrier, but the effects of piloerection can be seen in the

goose bumps that are produced on the skin when exposed to cold temperatures. Heat loss is also controlled by varying dermal blood flow. In warmer environments, the skin becomes flushed with blood, and metabolically produced heat can be dissipated into the environment. In cold environments, blood flow to the skin is restricted, effectively increasing the insulatory barrier between the internal source of heat production and the environment. These insulatory adjustments within the zone of thermal neutrality can be very effective at restricting heat loss. In the arctic fox, for example, an increase in metabolic rate due to shivering does not begin until environmental temperatures fall below −4°F (−20°C). *See* THERMOREGULATION.

In warm environments or during intense physical activity, which produces large amounts of metabolic heat, humans sweat. Sweat is a fluid similar in salt composition to that of blood plasma, from which it is derived. Sweating places large amounts of water directly on the surface of the skin. Some of this water evaporates, that is, is converted from a liquid to a gaseous state. Thus, metabolically produced heat may be brought to the skin through increased blood flow to the dermis, and this heat may then be dissipated to the environment through evaporation. Note, however, that the sweat must actually evaporate. If it simply drips off the body in its liquid form, no evaporative cooling is achieved. In very hot and dry environments, such as deserts, evaporation may be so rapid that humans might not be conscious of sweating, but under these conditions sweating proceeds at high rates and body dehydration can become a serious problem unless fluids are replaced regularly.

Sweating is the primary means of heat dissipation in humans, but it plays a substantially smaller role in most other mammals. While some mammals, such as horses, have widespread sweat glands for thermoregulation, many others have a very restricted distribution. In dogs, for instance, sweat glands are confined to the pads of the feet and probably serve principally to adjust pad friction during running. In mammals without a general sweat gland distribution, evaporative heat loss is undertaken by panting or by spreading saliva over the surface of the skin. In the latter case, the effect is the same in dissipating heat as if the fluid had been provided by sweating.

Evolution of vertebrate skin and scales. The most primitive living vertebrates are the agnathans, lampreys, and hagfish. While they are very different from other vertebrates in many aspects of anatomy and physiology, the basic organization of their skin is strikingly similar to that of mammals. They have an avascular, multicellular epidermis (although it contains mucus-secreting glands) underlain by a dermis containing elastic and collagen fibers and relatively few cells. There is a rich layer of melanocytes at the junction between the two layers. This basic bilayer organization of vertebrate skin must have evolved very early, or perhaps was a heritage from prevertebrate ancestors. *See* SCALE (ZOOLOGY).

Dermal structures. The first vertebrate group of which there is fossil evidence is the ostracoderms. These jawless, fishlike animals lived as bottom foragers in the seas a half billion years ago. Much of their body (especially their head and thorax) was covered with dermal bony armor. This material contained bone and denticles, the latter being composed of dentine and enamel, similar to that found in teeth in later vertebrates. This bony material was located in the dermis of the skin and was used as a defensive armor against predatory attack. A variety of similar hard dermal structures are found in some modern vertebrates. Sharks, for instance, have denticles that originate in the dermis and protrude through the epidermis, giving a rough, sandpaperlike feel to their skin. Teeth are thought to have evolved from such denticles located in the jaw region. Many fish and even some reptiles (for example, crocodiles) have bony scales (osteoderms) located in the dermis. These structures help to provide partially protective armor against external attack. Such protective elements are usually absent in mammals, although they did evolve convergently in a few groups, such as armadillos.

Scales appeared in the very first fossils of the next major vertebrate group, the placoderms. Scales are composite units of the skin, with a dermal core and overlying epidermal covering. They come in a wide variety of sizes and forms, sometimes containing bony or other supportive material, sometimes having only a connective tissue core. They are found in all descendant classes of vertebrates. They are particularly well developed in fish and reptiles. Birds, however, have well-developed scales on their legs, and their feathers were evolutionarily derived from scales. Certain groups of mammals also have scales: for example, rates and opossums have them on the tails. Most modern amphibians are scaleless; however, ancestral amphibians were almost certainly scaled, and some caecilians (apodans), an order of tropical wormlike amphibians, retain ancestral scales buried in the skin. Scales are apparently very convenient functional units in the skin. They permit a degree of protection for the underlying tissues and yet retain flexibility in the entire integument, permitting bending of the body or appendage by stretching or compressing the flexible areas between the scales. *See* FEATHER.

Functional role of skin. The functional role of the skin changed with the transition of vertebrates from aquatic to terrestrial environments. Even in aquatic environments, the skin serves as partial barrier to water entry or loss. Osmotic concentration of body fluids in most teleost fish (either marine or freshwater) is about 30% that of sea water. Thus, there is a tendency for fish to dehydrate (that is, to lose water from their body fluids) in ocean water and to become waterlogged in fresh water. Water permeability across the skin of fish, with its thick, scaly layers, is low, greatly lowering the potential water exchange with the surrounding fluid. In both marine and freshwater environments, however, water is abundant and freely available. Upon emergence into the terrestrial environment and air, primitive tetrapods found themselves in a very different hydric environment. Water was not always present and, even more significantly,

there was a continual loss of water from their bodies to the air by evaporation. Thus, any terrestrial environment in which the relative humidity is less than 100% is necessarily a desiccating environment. Much of the evolution of the tetrapod vertebrates was conditioned by the necessity of restricting water loss from the body. This was accomplished either by selecting wet and cool environments or by making a variety of physiological and morphological adaptations to retain internal water stores.

Among the tetrapods, there has been a general association between possession of a scaly skin and resistance to water loss across the integument. This association probably results from comparisons of scaleless frogs and salamanders, which have very high rates of cutaneous water loss, and scaled reptiles, such as lizards and snakes, which have the lowest cutaneous water loss rates among the vertebrates.

Water loss from amphibian skins. The problems of desiccation in the terrestrial environment are seen in sharpest relief among the amphibians, frogs, toads, and salamanders. Of all the terrestrial vertebrates, these organisms have the most permeable skins and greatest water permeabilities. Although some species have specialized adaptations to restrict cutaneous water loss, the vast majority lose water across their skins as if their bodies consisted of nothing more than an unprotected free surface of water. A skinned frog loses water at exactly the same rate as one with an intact skin, so the latter forms virtually no barrier to water loss. Most terrestrial species of amphibians have water permeabilities that are very similar to those of fully aquatic species; the adoption of terrestrial existence within this group has not been achieved by a reduction in the water permeability of the skin. Many terrestrial species would dehydrate and die in less than a day if they were placed in air and given no access to water. They are able to occupy terrestrial environments only because they behaviorally seek out and rehydrate in cool and wet areas. Thus, the entire life cycle of most amphibians, terrestrial as well as aquatic, is closely tied to the proximity to water.

Burrowing. In spite of their inherently high water permeabilities, some frogs and toads live in exceptionally hot and dry environments. For example, the spadefoot toad, *Scaphiopus couchii*, lives far from any permanent source of water in the arid desert of southeastern California, an exceedingly inhospitable place for an amphibian. It is able to survive by burrowing underground for most of its life and emerging onto the ground surface only after summer rainstorms. These rainstorms are infrequent and unpredictable, and are sometimes separated by several years in any particular local area. When a sufficient amount of rain falls to create temporary pools in the desert, the spadefoot toads emerge and rehydrate, not by drinking but by sitting in pools of rainwater and absorbing water across their skin. This process is under hormonal control, and rehydration can be exceptionally fast, with rates of water uptake of as much as 30% of their body weight per hour in small toads. The toads feast on the abundant insects available at that time, primarily termites which emerge with the rain. The toads mate and lay eggs in the temporary pools, and then dig down into the soil again. The tadpoles in the temporary rain pools develop exceptionally fast, for they are in a race to metamorphose into toadlets before the pool dries up. The entire period of surface activity for the toads may span only a couple of weeks out of the year, and in some years the toads may not emerge at all. Under the soil, the toads form cocoons of shed skin and enter a reduced metabolic state called estivation, in which their metabolic rates may fall to only one-fifth normal values. During estivation, they do not breathe: all gas exchange occurs across their skin. Their major source of fuel for the burrowing period is fat accumulated in abdominal fat bodies and throughout other organs. In addition, the toads may utilize 20–30% of their body's protein as a source of fuel. As a result of this protein breakdown, they also accumulate high concentrations of the nitrogen excretory product urea in their bodies. These high levels of urea aid water retention in the burrowed toads by reducing the tendency of water to diffuse across the skin into the surrounding soil. Further, less water is lost via urination by the storage of urea rather than its excretion.

Many species of frogs and toads use burrowing to escape from seasonally inhospitable environments. In many aspects of this response, their skin plays a crucial role, particularly during rehydration, underground respiration, and the formation of a cocoon. The cocoons are particularly important in reducing water loss to the surrounding dry soils. In some species, these consist of many layers of the shed portion of the stratum corneum. The cocoon in composed of dozens of layers of keratin and lowers the overall rate of water loss across the skin to almost immeasurable levels. The cocoons are shed one cell layer at a time, sometimes as often as once a day.

Mucus secretion. Some arboreal frogs are also exposed to very drying environments. They perch on tall grasses or shrubs, and are sometimes even exposed to direct sunlight. These frogs cannot escape dehydration by burrowing. Instead, they have developed various means to restrict water loss across their skins. Many treefrogs (for example, spring peepers, genus *Hyla*) have rates of water loss that are only one-third to one-half these of nonarboreal frogs of the same size. Part of this reduction in water loss appears to be the result of the accumulation of dried mucus on the skin. These treefrogs release mucus onto their backs when exposed to dry air, and as this evaporates, the residual protein is left on the skin surface. Water loss is further restricted when the frogs adopt a water-conserving posture of folding their limbs beneath them and pressing their bodies against the substrate. This behavior minimizes the amount of skin exposed to dry air from which evaporation could occur.

Waterproofing. In some African and South American treefrogs, the reduction in water loss across the skin has reached spectacular levels. These frogs are

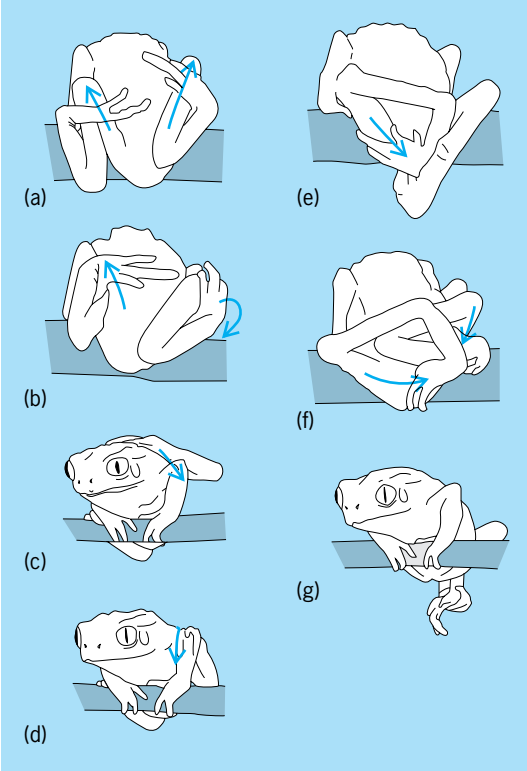

Fig. 2. The treefrog *Phyllomedusa sauvagei* wipes a waxy secretion from its skin all over its body with its feet in sequence *a–g*. Arrows show the direction of movement. After drying, this coating forms a very effective barrier to water loss from the skin. (*After L. A. Blaylock et al., Skin structure and wiping behavior of phyllomedusine frogs, Copeia, 1976:283–295, 1976*)

effectively waterproofed, reducing water loss in dry environments to levels fully equal to those of reptiles. Members of the genus *Chiromantis* live in very hot semiarid regions of southern Africa, perched on shrub branches, sometimes sitting directly in the sun. Thus, their ability to conserve water is crucial for survival. The functional basis of the waterproofing in *Chiromantis* is not known, although the stratum corneum on the backs (but not on the abdomens) of these frogs contains a unique layer immediately below the outer surface. This layer is likely involved in their low water permeability. Treefrogs of the genus *Phyllomedusa* also have extremely low rates of water loss across their skins; one species lives on exposed tree limbs in the arid Chaco region of South America and has been found to have body temperatures that are the highest ever recorded for an amphibian. These frogs are waterproofed by a unique lipid gland in the stratum corneum of their skin. The gland produces a waxy secretion that the frog wipes all over its body with its feet in a highly stereotyped behavioral sequence (**Fig. 2**). After this covering dries, the frog appears shiny, and droplets of water placed on its back bead up as on a waxed surface. This covering forms an extremely effective barrier to water loss. When body temperatures reach levels near 104°F (40°C), the treefrog remains cool by secreting water onto its back from numerous mucus glands located next to the lipid glands in its skin. The rate of this water secretion is carefully controlled so that heat gain from the environment is exactly balanced by evaporative heat loss, and body temperature remains constant. *See* INTEGUMENTARY PATTERNS; LIPID.

Albert F. Bennett

Bibliography. M. E. Feder and W. W. Burggren (eds.), *Environmental Physiology of the Amphib- ians*, 1992; W. N. McFarland et al., *Vertebrate Life*, 3d ed., 1989.

Skin disorders

The skin is subject to localized and generalized disorders, as well as those of primary occurrence in the skin and those secondary to involvement of other tissue. Diseases and disorders may affect any of the structures of the skin. They may be caused by external agents, either infectious or noninfectious, or by the abnormal accumulation of normal or abnormal skin elements, either inborn or acquired.

Infectious diseases. These are classified by the type of infectious agent—bacterial, parasitic, fungal, or viral.

Bacteria. Bacterial infections are distinguished clinically by the skin layer or appendage affected, but treatment is based upon the organism causing the infection. Impetigo and cellulitis are the most common infections of skin. Impetigo more commonly affects children and young adults, and it is usually more superficial and localized. Cellulitis is more common in older individuals and affects deeper layers in and larger areas of the skin, most often on the lower legs. Both infections may be caused by streptococcus, including group B hemolytic streptococcus, and staphylococcus, often species resistant to penicillin. Erysipelas is a rare streptococcal infection of the dermis, often of the face, characterized by high fever; it is potentially fatal if not promptly diagnosed and treated. Treatment is guided by culturing and determining the antibiotic sensitivity of the organism. Depending upon the results of the culture, treatment with antibiotics that are effective against organisms resistant to penicillin is recommended. *See* ANTIBIOTIC; DRUG RESISTANCE; STAPHYLOCOCCUS; STREPTOCOCCUS.

Skin appendages such as hair follicles may be similarly infected; the condition is known as folliculitis when individual follicles are involved, and furunculosis and carbunculosis when several or many adjacent follicles are involved in a large nodule or boil. Treatment is with antibiotics determined by culture but incision and drainage of the furuncles and carbuncles speeds treatment time. Acute infections of other skin appendages are less common, but sebaceous glands may be chronically infected in a triad of diseases including severe cystic acne, hidradenitis suppurativa, and dissecting cellulitis of the scalp. Treatment includes long-term antibiotic therapy, selective surgical intervention, and, in certain cases, retinoid drugs.

Parasites. Parasitic skin conditions are most often seen in epidemics among individuals who are in close contact, or where hygiene is poor, such as in

schools, mental institutions and other chronic care facilities, and shelters. Head, body, and pubic lice, and scabies are the most common. Pubic lice and scabies are also often transmitted by sexual contact. Scabies is also common in the general population, particularly in children and young adults. The mites of scabies tunnel within the epidermis, whereas all forms of lice live on the surface of the skin or in clothing. Treatment consists of topical pesticides, most often lindane. Treatment of allergic reactions to the insect parts or bites, particularly in scabies, is often necessary. *See* MEDICAL PARASITOLOGY; PESTICIDE.

Fungi. Fungal infections are extremely common. In the United States, most fungal infections in humans are with species incapable of infecting tissue other than keratinized epidermis. These organisms are known as dermatophytes; they cause tinea pedis (athlete's foot), tinea cruris (some forms of jock itch), and tinea capitis (a scalp condition responsible for some forms of hair loss). Regardless of the specific dermatophyte organism involved, all such infections are treated with topical or oral antifungal antibiotics that work by interfering with formations of the fungal wall. The length of treatment varies from as little as a few days of oral treatment for tinea capitis to a lengthy treatment that may require as much as 2 years for fungal toenail disease. *See* MEDICAL MYCOLOGY.

Viruses. Viral diseases often involve the skin, including conditions such as measles, chickenpox, and many others. This discussion, however, will be limited to viral diseases that affect only the skin. Viral warts (verruca vulgaris) and molluscum contagiosum are the primary examples. Both are characterized by single or multiple, somewhat contagious, skin tumors that usually are small but can in rare instances exceed 0.4 in. (1 cm) in diameter. They respond to various treatments, and they may disappear spontaneously. Both diseases can affect genital skin, and they are then much more likely to be spread by sexual contact. *See* ANIMAL VIRUS; CHICKENPOX AND SHINGLES; HERPES; MEASLES; SMALLPOX.

Inflammatory disease. Most itchy rashes are due to inflammation of the skin; they are usually known as eczema or dermatitis. Inflammatory skin diseases can occur as allergic reactions to specific chemicals in the environment, such as poison ivy; or they are caused by direct damage to the skin, such as by excessive exposure to soaps, detergents, or chemicals. In the acute stage, eczematous dermatitis is characterized by a vesicular, oozing condition. If the condition becomes chronic, the skin becomes dry, scaling, and leathery. When a specific cause can be identified, it must be avoided. Topical corticosteroids are the mainstay of treatment.

Seborrheic dermatitis is a common eczematous condition affecting primarily the areas of skin that bear sebaceous glands, that is, scalp, central face, chest, axilla, and groin. This type of dermatitis also usually responds to topical corticosteroids. Evidence has implicated the organism *Pityosporum ovale*, which explains the response to oral and topical ketoconizole (an antibiotic that kills the organism) in some cases, and the very high incidence in patients with acquired immune deficiency syndrome (AIDS), who are susceptible to many infectious diseases.

Hereditary disease. Atopic dermatitis is the skin manifestation of atopy, a clinically apparent hypersensitivity. This condition may also be associated with asthma and pollen allergies. It is hereditary and first manifests itself in infancy or early childhood. Some children outgrow it. However, it can persist into adulthood, and sometimes it does not appear until later in life. Atopic dermatitis is more common in blacks and people from Asia, particularly the Philippines. The main treatment is topical corticosteroids.

Psoriasis is a common disease of unknown etiology. It has been estimated to occur in up to 3% of the population of the United States. The disease is hereditary, but not in a simple mendelian pattern. Psoriasis is usually seen in young and middle-aged adults and it is rare in young children. Psoriasis is usually chronic for years or for the rest of the individual's lifetime. There is, however, a poststreptococcal form of the disease that is seen in adolescents and young adults; this type responds to antibiotics, and it may be cured completely, without recurrences. The typical psoriasis consists of well-defined patches and plaques of red skin with a silvery scale that often results in pinpoint bleeding when removed. The most common and persistent sites are the elbows, knees, and scalp, but any area of skin may be involved, including the palms and soles; there may be significant morbidity and disability. Nails may be involved, even in the absence of skin involvement. Psoriatic arthritis, which is similar to but distinct from rheumatoid arthritis, occurs in a small number of individuals; it may be severe enough to cause disability. There are a number of treatment modes for psoriasis, both local and systemic. Each patient requires an individual course of treatment, based on the concept that it should be adequate and not worse than the disease in terms of time, expense, discomfort, unsightliness, or potential medical complications.

Other conditions. Acne vulgaris is an extremely common skin disorder, affecting 80–90% of young adults, usually during adolescence. The condition varies widely in severity, distribution, and duration between individuals and over time in the same individual. The pilosebaceous organ of the skin is the primary target, particularly the sebaceous gland, its duct, and the infundibulum of the hair follicle. This organ is responsive to hormones and is variously stimulated by the different steroid hormones, including the sex hormones; this is the cause of the prevalence of the disease in the years immediately after puberty. In addition to increased production of sebum, bacteria contribute to the development of acne lesions. These lesions consist of open and closed comedones (blackheads), papules, pustules, and cysts. Scarring may result, particularly in cystic acne or in lesions that are physically manipulated. A large variety of treatments exist, directed at the bacterial, sebum production, and structural aspects of acne. Often several agents are used together to treat several aspects of the problem simultaneously. Each

individual must be evaluated, and a treatment regimen selected that can be modified as the condition responds or changes in severity over time.

Reactions to ingested materials, such as food or medications, often appear in the skin, usually either as a red, itchy, measleslike rash or as urticaria (hives). These reactions indicate allergy to the material, which must then be avoided. Reexposure may result in a more severe, potentially fatal, anaphylactic reaction. Treatment for these reactions consists of antihistamines and corticosteroids, both topically and, in severe cases, systemically. Under rare circumstances, usually where allergy to a potentially life-saving antibiotic or other drug exists, the agent may be given in a carefully controlled situation in the hospital or similar facility where immediate care for a severe allergic reaction can be provided. *See* ALLERGY.

Neoplasms. Skin neoplasms may be benign or malignant, congenital or acquired, and they may arise from any component of the skin. Almost all skin neoplasms are benign and acquired. The common mole (melanocytic nevus) is a neoplasm of benign melanocytes; usually it is acquired, but it may be present at birth, when it is often known as a birthmark. With the exception of giant congenital melanocytic nevi, which have approximately a 10% chance of becoming malignant and thus should be removed, nevi do not seem to have any greater risk of becoming malignant than any other part of the skin. Other common congenital nevi or birthmarks are of vascular origin, including strawberry and cavernous hemangiomas and port wine stains. With rare exception, these are of cosmetic significance only.

Malignant neoplasms may arise from cellular elements of the epidermis or dermis, or by infiltration of the skin by malignant cells arising from other tissues. By far the most common are basal cell and squamous cell cancers, which arise from basal and squamous keratinocytes of the epidermis, respectively. These constitute the most common form of cancer in the United States, with over 600,000 new cases per year. They are caused by the cumulative effects of ultraviolet radiation on the skin. Luckily, they are locally invasive in almost all cases, rarely metastasize or cause death, and are readily recognized and treated with standard outpatient procedures. They usually are characterized by a nonhealing sore, persistent red scaling or crusting patch, or a slowly growing pearly nodule on skin that has been exposed to the sun; they occur mostly on the head, neck, hands, and arms.

Malignant melanoma arises from the pigment-forming melanocyte, and thus it is usually pigmented. It is a metastasizing cancer that is often fatal if not removed surgically in the early stage. Once the cancer is advanced, chemotherapy and other interventions are rarely more than palliative. It can be recognized as a pigmented lesion, often thought to be a benign nevus initially, that increases in size, changes color, particularly with admixtures of black, blue, red, or white along with the usual tans or browns; and it becomes irregular in size and shape. Other signs and symptoms, such as bleeding, pain, and itching, are less frequent, often occurring as late manifestations. Malignant melanoma is one of the most rapidly increasing cancers—probably related to increased ultraviolet radiation exposure. *See* CANCER (MEDICINE).

Two unusual multicentric primary skin malignancies are mycosis fungoides and Kaposi's sarcoma. Mycosis fungoides is a lymphoma of the skin, usually present in several sites when first diagnosed, that may remain confined to the skin for 10 or more years before eventually spreading to internal organs and causing death. It may be extremely difficult to diagnose, both clinically, when it can appear only as eczematous patches, and histologically, for months or years. Often, multiple biopsies over time are required to make the diagnosis. It is not clear whether early aggressive chemotherapy can result in a cure or increased life-span, or whether such therapy may even shorten life. Kaposi's sarcoma occurs in two forms, the classic form seen on the legs of elderly Mediterranean men, and a form associated with HIV-1 infection and AIDS that may occur on any part of the body. It is derived from skin blood vessels, is multicentric, and usually appears as red to violet patches, plaques, or nodules. It is usually not fatal, although it may eventually spread to internal organs and may cause significant morbidity. It is treated palliatively by ionizing radiation and local surgery. When it is extensive and internal, the treatment includes chemotherapy. There are numerous other primary skin cancers, but they are rare. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); SKIN.

Alan N. Moshell

Bibliography. T. B. Fitzpatrick et al., *Dermatology in General Medicine*, 1987; L. A. Goldsmith (ed.), *Biochemistry and Physiology of the Skin*, 1983; W. F. Lever and G. Schaumburg-Lever, *Lever's Histopathology of the Skin*, 8th ed., 1997; W. Montagna and P. F. Parakkal, *The Structure and Function of Skin*, 1974; S. L. Moschella and H. J. Hurley, *Dermatology*, 3d ed., 1992.

Skin effect (electricity)

The tendency for an alternating current to concentrate near the outer part or "skin" of a conductor. For a steady unidirectional current through a homogeneous conductor, the current distribution is uniform over the cross section; that is, the current density is the same at all points in the cross section. With an alternating current, the current is displaced more and more to the surface as the frequency increases. The conductor's effective cross section is therefore reduced so the resistance and energy dissipation are increased compared with the values for a uniformly distributed current. The effective resistance of a wire rises significantly with frequency; for example, for a copper wire of 1-mm (0.04-in.) diameter, the resistance at a frequency of 1 MHz is almost four times the dc value. *See* ALTERNATING CURRENT; ELECTRICAL RESISTANCE.

A mathematical description of skin effect can be

derived from Maxwell's equations, but exact solutions have been obtained for only several simple shapes, including cylindrical, tubular, and flat conductors. For a plane conductor carrying a sinusoidal alternating current, the current density is a maximum at the surface and its magnitude decreases exponentially with distance into the conductor. A skin depth or penetration depth δ is frequently used in assessing the results of skin effect; it is the depth below the conductor surface at which the current density has decreased to $1/e$ (approximately 37%) of its value at the surface and is given by $\sqrt{1/(\pi f \mu \sigma)}$, where f is the frequency and σ and μ are the conductivity and permeability of the material respectively. This concept applies strictly only to plane solids, but can be extended to other shapes provided the radius of curvature of the conductor surface is appreciably greater than δ. *See* MAXWELL'S EQUATIONS.

At a frequency of 60 Hz the penetration depth in copper ($\sigma = 5.8 \times 10^7$ S m^{-1}, $\mu = 4\pi \times 10^{-7}$ H m^{-1}) is found from the above expression to be 8.5 mm (0.33 in.); at 10 GHz it is only 6.6×10^{-7} m. Waveguide and resonant cavity internal surfaces for use at microwave frequencies are therefore frequently plated with a high-conductivity material, such as silver, to reduce the energy losses since nearly all the current is concentrated at the surface. Provided the plating material is thick compared to δ, the conductor is as good as a solid conductor of the coating material. *See* CAVITY RESONATOR; MICROWAVE; WAVEGUIDE.

In the case of a cylindrical conductor carrying a sinusoidal alternating current, the solution to the skin-effect equation is determined in terms of the zero-order Bessel functions of the first kind. When the frequency is high, the exact formulation reduces to the exponential solution for a plane conductor. A useful approximation to the ratio of ac-to-dc resistance of a round conductor of radius r is $r/2\delta$, if $\delta \leq 0.1r$. For low frequencies or conductors of small radii, the exact ratio in terms of the Bessel functions must, however, be employed. At high frequencies the central part of a cylindrical conductor carries so little current that it is often removed to give a tube which saves on weight and can also be convenient for passing cooling fluid. The skin-effect equation for tubular conductors is solved in terms of Bessel functions of the first and second kind. *See* BESSEL FUNCTIONS.

Frank A. Benson

Bibliography. B. I. Bleaney and B. Bleaney, *Electricity and Magnetism*, 3d ed., 1976, reprint 1989; W. H. Hayt, *Engineering Electromagnetics*, 6th ed., 2001; M. A. Plonus, *Applied Electromagnetics*, 1978; S. Ramo, J. R. Whinnery, and T. van Duzer, *Fields and Waves in Communication Electronics*, 3d ed., 1994.

Skink

A widely dispersed reptile belonging to the family Scincidae. The family contains approximately 1200 species and is represented on every continent except Antarctica. There are over 150 species of skink in Australia alone. Characteristics include elongate bodies, moderate to long tails, and proportionately short, reduced, or absent limbs. Some scales on the head and body have cores of bone. In many species the eyelids are transparent. Skinks exhibit pleurodont dentition; that is, the teeth are attached on the side of the jaw. *See* DENTITION.

Five-lined skink (*Eumeces fasciatus*). (*Photo by George W. Robinson; © California Academy of Sciences*)

Skinks vary widely in size and longevity (some have been known to live to 20 years). They are carnivorous, eating invertebrates and, in the case of larger species, even small mammals, reptiles, and birds. Three reproductive modes can be found among skink species: oviparity (egg-laying), viviparity (bearing live young), and ovoviviparity (producing eggs that develop internally and hatch before or soon after extrusion). Three genera, *Neoseps*, *Scincella*, and *Eumeces*, occur in the United States. They are represented by species such as the federally threatened sand skink (*N. reynoldsi*) of Florida, which has reduced limbs and ear openings covered by scales that are specifically adapted for burrowing; the ground skink (*S. lateralis*), the only *Scincella* species found throughout the eastern regions; and about 15 species of the common genus *Eumeces*, the striped skinks (see **illustration**).

A number of species of the genus *Tiliqua*, represented by the largest skinks in terms of body size, inhabit Malaysia and Australia. *Tiliqua gigas* reaches a length of 2 ft (0.6 m). The blue-tongued skink (*T. scincoides*), also native to Australia, can survive a wide temperature range. Conversely, the Egyptian skink (*Scincus officianalis*), a common species of the Sahara region, is lethargic unless the temperature is near 100°F (38°C). It has a flattened snout that is used for burrowing into the sand. Its wide, thin toes resemble feathers and are well adapted for walking on sand. Another Sahara species, *Chalcides sepoides*, has a cylindrical body and four small limbs that fold against the body when the animal moves through sand, permitting a snakelike body movement. A similar species, the three-toed sand skink (*C. striatus*), common in France and the Iberian Peninsula, also folds its legs when crawling, but unlike the previous species it has depressions in the

side of the body to accommodate the folded limbs. These highly adapted species are but a few examples of the diversity of skinks worldwide. *See* REPTILIA; SQUAMATA. W. Ben Cash

Bibliography. F. H. Pough et al., *Herpetology*, 3d ed., Prentice Hall, Upper Saddle River, New Jersey, 2004; G. R. Zug, L. J. Vitt, and J. P. Caldwell, *Herpetology* Academic Press, San Diego, 2001.

Skunk

A carnivore long classified in the subfamily Mephitinae in the family Mustelidae (otters, weasels, mink, badgers, and their relatives), but since 1997 placed in the independent family Mephitidae along with the Oriental stink badger *Mydaus*. This new placement was based on convincing DNA evidence linking skunks (genera *Mephitis*, *Spilogale*, and *Conepatus*) and *Mydaus* genetically.

Skunks are found only in the New World, ranging from Canada to Chile. These mammals have short limbs with five clawed toes on each foot. The claws are curved and nonretractile. A pair of well-developed anal scent glands, used for defense against predators, are present. A short rostrum is present on the elongate and flattened skull. Carnassial teeth (used to shear flesh and bone) are well developed.

Striped and hooded skunks (Mephitis). These two species range from southern Canada to Costa Rica. Adults are about the size of a house cat, standing approximately 18 cm (7 in.) at the shoulder, with adults usually attaining a head and body length between 550 and 750 mm (21 and 29 in.), including a tail 175–300 mm (7–12 in.) long. Adults usually weigh between 1.3 and 4.5 kg (3 and 10 lb). The dental formula is I 3/3, C 1/1, Pm 3/3, M 1/2 × 2, for a total of 34 teeth. *See* DENTITION.

The relatively small head possesses small, short, rounded ears, small eyes, and a pointed muzzle. The feet are semiplantigrade with long claws on the forefeet for digging. The pelage (the coat, including fur and hairs) is long, coarse, and oily. The body is black with a narrow white stripe running up the middle of the forehead and a broad white area on the nape of the neck that usually divides into a V at about the shoulders (see **illustration**). The resulting two white stripes may continue to the base of the bushy tail. There is, however, much variation in the length and width of the stripes; some skunks have very broad, well-defined stripes, whereas in others the stripes are almost or completely lacking.

Striped skunks prefer bushy and sparsely wooded areas and are most abundant in agricultural regions where the forest gives way to open lands and farm crops. Hooded skunks occur at low elevations in desert areas. Both species are mainly nocturnal. Daylight hours are usually spent resting or sleeping in a hollow log, in wood piles, beneath an abandoned building, along a fence row, or in a burrow in the ground. Striped skunks may become dormant for prolonged periods during severe winter weather, but

Striped skunk (*Mephitis mephitis*). (*Photo © Creatas/Punch Stock*)

they do not hibernate. Striped and hooded skunks are not good climbers. When frightened or annoyed, the skunk contracts the muscles surrounding the scent glands, thus forcing the scent out in a fine, almost invisible spray. At least seven major volatile components have been identified in the secretion of the striped skunk. The fluid has a most distasteful and nauseating stench and produces intense smarting and burning if it comes in contact with the membranes of the eyes, nose, or mouth. Incongruous as it may seem, the oily yellow fluid produced by the skunk is used as a base for perfumes.

Striped and hooded skunks are omnivorous and feed on insects, crayfish, fish, frogs, turtle eggs, snakes, small birds, and small mammals. During fall and winter, fruit, small mammals, carrion, grains and nuts, grasses, leaves, and buds are important.

Striped skunks are polygamous and usually mate from January through March. Most females normally produce a single annual litter. Following a gestation of 59–77 days, a litter of 3–10 blind young are born in a nest of dry leaves and grasses. They have a fine covering of hair in which the adult pattern is evident. The eyes open at 22–35 days of age, and by 7–8 weeks they are out following their mother in single file on her nightly foraging trips. They are weaned at 8 weeks. Striped skunks may live 5 to 6 years in the wild. The main predators are the great horned owl and the barred owl.

Spotted skunk (Spilogale). Spotted skunks are the smallest and most weasel like in appearance of the three genera of skunks. This genus consists of three species which range from southern British Columbia to Oaxaca, Mexico, and eastward to Pennsylvania and the southeastern United States. Spotted skunks

may be distinguished from the other two genera by their smaller size, forehead patch, and pattern of stripes and spots. The basic color pattern consists of six white stripes extending along the back and sides; these are broken into smaller stripes and spots on the rump. There is a triangular patch in the middle of the forehead, and the tail is usually tipped with white. Variations, however, are infinite. A pair of scent glands is present under the base of the tail. The dental formula is I 3/3, C 1/1, PM 3/3, M 1/2 × 2, for a total of 34 teeth. Adults are between 185 and 565 mm (7 and 22 in.) in total length, including an 80–220 mm (3–8 in.) tail. They weigh 200–1000 g (7–35 oz).

Spotted skunks inhabit a variety of brushy, rocky, and wooded habitats but avoid dense forests and wetlands. In the states, they inhabit riparian woodlands and areas of vegetation along fences. The omnivorous diet consists mainly of insects and vegetation in the summer and rodents and other small animals in the winter. Unlike striped and hooded skunks, spotted skunks are agile climbers.

Breeding is not the same in all parts of the range, with some populations mating during March and April and others mating in September and October. Gestation may range from 50 to 65 days in spring breeders to 230 to 250 days in fall breeders. The period of delayed implantation may range from 2 weeks in spring breeders to as long as 6 months in fall breeding populations. Litters average 3 to 6 young. The eyes open at 32 days, and they are weaned at about 54 days of age. Longevity in the wild is unknown, but a captive lived 9 years 10 months.

Hog-nosed skunks (Conepatus). Five species of hog-nosed skunks range from southern Colorado and eastern Texas to southern Chile and Argentina. These skunks have the coarsest fur of all skunks. This genus can be distinguished from the other two by the absence of a white head stripe and by its long head and naked muzzle, somewhat like a hog's snout. This adaptation allows the skunk to root in the ground for insects, which make up a good part of its diet. The color pattern is variable. In some, the top of the head, the back, and the tail are white, and the remainder of the animal is black. In others, the pattern resembles that of *Mephitis*, with the pelage being black except for two white stripes and a mostly white tail. Adults are 460–900 mm (18–35 in.) long, including a 160–410 mm (6–16 in.) tail. They weigh 2.3–4.5 kg (5–10 lb). Little is known concerning reproduction and longevity. Donald W. Linzey

Bibliography. J. W. Dragoo and R. L. Honeycutt, Systematics of mustelid-like carnivores, *J. Mammal*, 78:426–443, 1997; D. W. Linzey, *The Mammals of Virginia*, McDonald & Woodward Publishing, 1998; R. N. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999; J. O. Whitaker, Jr., and W. J. Hamilton, Jr., *Mammals of the Eastern United States*, 3d ed., Comstock Publishing Associates, 1998; D. E. Wilson and S. Ruff (eds.), *The Smithsonian Book of North American Mammals*, Smithsonian Institution Press, 1999.

Slate

Any of the deformed fine-grained, mica-rich rocks that are derived primarily from mudstones and shales, containing a well-developed, penetrative foliation that is called slaty cleavage (see **illus.**). Slaty cleavage is a secondary fabric element that forms under low-temperature conditions (less than 540°F or 300°C), and imparts to the rock a tendency to split along planes. It is a type of penetrative fabric; that is, the rock can be split into smaller and smaller pieces, down to the size of the individual grains. If there is an obvious spacing between fabric elements (practically, greater than 1 mm), the fabric is called spaced. Slates typically contain clay minerals (for example, smectite), muscovite/illite, chlorite, quartz, and a variety of accessory phases (such as epidote or iron oxides). Under increasing temperature conditions, slate grades into phyllite and schist. *See* ARGILLACEOUS ROCKS; CLAY MINERALS; PHYLLITE; SCHIST.

Cleavage categories. There are four main categories of cleavage based on morphological characteristics: disjunctive cleavage, pencil cleavage, slaty cleavage, and crenulation cleavage.

Disjunctive cleavage. This type of cleavage is defined by an array of subparallel fabric elements, called cleavage domains, in which the original rock fabric

Slaty cleavage in folded mica-rich rock near Walland, Blount County, Tennessee. (*USGS*)

and composition have been markedly changed by the process of pressure solution. Domains are separated from one another by intervals, called microlithons, in which the original rock fabric and composition are more or less preserved. The cleavage domains cut across a preexisting fabric (usually bedding) that is preserved in the microlithons. In a horizontal bed of argillaceous (clay-rich) limestone or sandstone that is subjected to a compressive stress, dissolved ions created by pressure solution diffuse away from the site of dissolution through a fluid film that adheres to the grain surfaces. The ions precipitate at crystal faces where compressive stress is less, precipitate nearby in pressure shadows adjacent to rigid grains, or enter the pore fluid system and are carried out of the local rock environment entirely. Where it occurs, the process preferentially removes more soluble grains. Thus, in an argillaceous limestone, calcite is removed and clay and quartz are progressively concentrated. In an argillaceous sandstone, the process is effectively the same, except that quartz is the mineral that preferentially dissolves and clay alone is concentrated. Eventually, a discrete domain develops in which there is a selvage composed of mostly clay with some relict corroded calcite or quartz grains. If deformation continues, the domain further thickens by pressure solution along its edges. As a result, compositional contrasts between cleavage domains and microlithons become so pronounced that a new layering is defined in the rock.

Pencil cleavage. When a fine-grained sedimentary rock (shale or mudstone) breaks into elongate pencillike shards, it is said to have a pencil cleavage. Typically, pencils are 5–10 cm (2–4 in.) long and 0.5–1 cm (0.2–0.4 in.) in diameter. In outcrop, pencil cleavage looks like it results from the interaction of two fracture sets, but the parting reflects the internal alignment of clay grains in the rock. Because of their strong shape anisotropy, clay flakes attain a preferred orientation parallel to bedding when they settle out of water and are compacted together. This preferred orientation imparts the tendency for clay to break on bedding planes that is displayed by shale. If the shale is subjected to layer-parallel shortening, rotation and microfolding of preexisting grains and growth of new grains occur. The plane of new flakes or parts of deformed flakes is roughly perpendicular to the shortening direction (that is, at a high angle to bedding). At an early stage during this process, the new deformation fabric is comparable in degree of development to the initial bedding fabric, creating the pencils.

Slaty cleavage. Pencil cleavage may be considered a snapshot of an early stage in the process by which slaty cleavage develops. If shortening perpendicular to cleavage planes continues, clay throughout the rock displays a preferred orientation at an angle to the primary sedimentary fabric. Eventually, the deformation fabric dominates the primary fabric. Formation of this secondary fabric occurs by much the same process as the formation of disjunctive cleavage; however, the resulting domains are so closely spaced that effectively there are no uncleared microlithons, and the clay in the entire rock displays the tectonically induced preferred orientation. When a rock has this type of fabric, it is said to have slaty cleavage. Thus, slaty cleavage is defined by a strong dimensional preferred orientation of clay in a very clay rich, low-grade metamorphic rock, and the resulting rock is a slate. Slaty cleavage tends to be smooth and planar. Coupled with the penetrative nature of slaty cleavage, these characteristics enable slates to split into very thin sheets. This and the durability of the rock are reasons why slates are used in the roofing industry, in the tile industry, and in the construction of pool tables.

Crenulation cleavage. Penetrative microfolding of the original bedding foliation or a secondary foliation (such as slaty cleavage) produces the fourth morphological cleavage type, called crenulation cleavage. Crenulation cleavage is characterized by the presence of microfolds with near-parallel limbs that define the cleavage surface. *See* ROCK CLEAVAGE.

Cleavage and strain. In order to determine the relationship of cleavage to strain, it is necessary to consider strain markers in a cleaved rock, such as reduction spots. Reduction spots are small regions where the iron in a red slate is reduced, and therefore look greenish. If these spots start out as spheres, they make ideal strain markers; and geologists find that the deformed spots are flattened ellipsoids with the plane of flattening almost parallel to the slaty cleavage domains. However, cleavage is probably not perfectly parallel to the flattening plane of strain, given the evidence for shear parallel to the cleavage. Estimates of the total shortening associated with cleavage formation are in excess of 60%. There is a question whether strain resulting from cleavage formation is a volume-constant strain or a volume-loss strain. If ions of soluble minerals enter the pore water system, they can be carried out of the local rock system by movement of ground water. In some rocks the volume may have decreased by up to 50% during cleavage formation. *See* SHALE. Ben A. van der Pluijm

Bibliography. B. A. van der Pluijm and S. Marshak, Fabrics, ch. 11 in *Earth Structure: An Introduction to Structural Geology and Tectonics*, McGraw-Hill, Dubuque, 1997.

Sleep and dreaming

Sleep is generally defined as an easily reversible, temporary, periodic state of suspended behavioral activity, unresponsiveness, and perceptual disengagement from the environment. Compared with other states of temporary unresponsiveness such as syncope or coma, sleep is easily reversible with strong or meaningful sensory stimuli (for example, the roar of a nearby tiger, or a voice speaking the sleeper's name). Sleep should not be considered a state of general unconsciousness. The sleeper is normally unconscious (but not always) of the nature of events in the surrounding environment; this is the meaning of perceptual disengagement: a lack of conscious perception and meaningful responsiveness to environmental

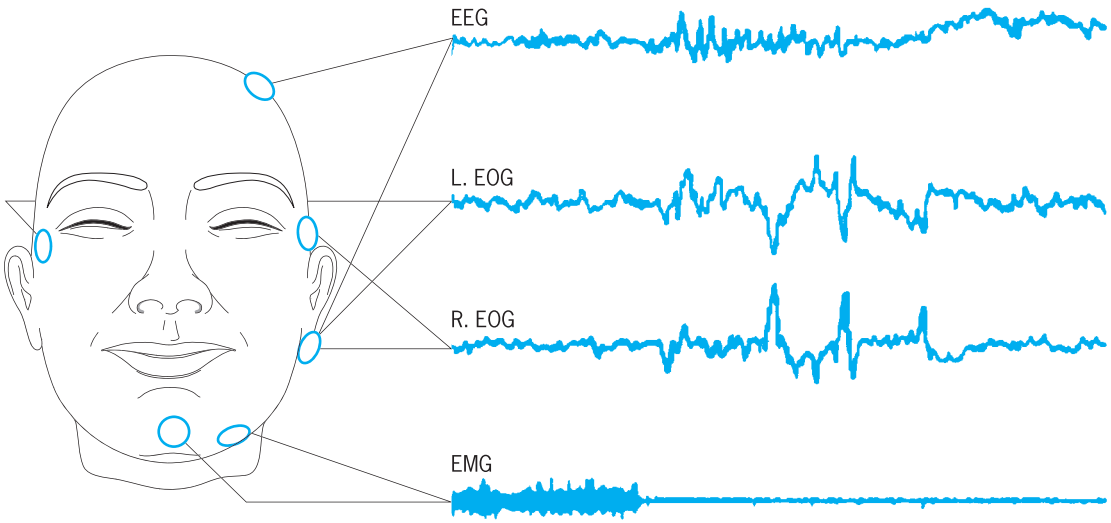

Fig. 1. Standard electrode placements for sleep research. Electroencephalogram (EEG) waves are recorded from left (C3) or right (C4) central locations, which give the best discrimination of sleep stages. Electrooculogram (EOG) electrodes are typically sited to allow both vertical and horizontal eye movements to be recorded. The standard chin (submental) electromyogram (EMG) site is chosen to show reliable suppression of muscle tone during onset of REM (rapid eye movement) sleep.

stimuli. However, the sleeper's attention may be fully engaged in experiencing a dream. And if reportability is accepted as a sufficient condition for conscious mental processes, any dream that can be recalled must be considered conscious. Dreaming, then, can be simply defined as the world-modeling constructive process through which people have experiences during sleep, and a dream is just whatever the dreamer experienced while sleeping.

Nature of sleep. In general, biological organisms do not remain long in states of either rest or activity. For example, if a cat's blood sugar level drops below a certain point (the cat's reference level for hunger-related action), the cat is motivated by hunger to venture from its den in search of a meal. It will continue to hunt until it finds food (or starves). After satisfying the urge to eat, the cat is no longer motivated to expend energy tracking down uncooperative prey; now its biochemical state motivates a return to its den, to digest in peace, conserve energy, and generally engage in restful, regenerative activities, including sleep. This example tracks a cat through one cycle of its basic rest-activity cycle (BRAC). Such cyclic processes are ubiquitous among living systems.

Sleep and wakefulness are complementary phases of the most salient aspects of the brain's endogenous circadian rhythm, or biological clock. Temporal isolation studies have determined the biological clock in humans to be slightly longer than 24 hours. Several features of sleep are regulated by the circadian system, including sleep onset and offset, depth of sleep, and rapid eye movement (REM) sleep intensity and propensity. In the presence of adequate temporal cues (for example, sunlight, noise, social interactions, and alarm clocks), the internal clock keeps good time, regulating a host of physiological and behavioral processes. *See* BIOLOGICAL CLOCKS; NORADRENERGIC SYSTEM.

Sleep is not a uniform state of passive withdrawal from the world, as scientists thought until the twentieth century. In fact, a version of the BRAC continues during sleep, showing a periodicity of approximately 90 minutes. There are two distinct kinds of sleep: a quiet phase (also known as quiet sleep or QS, slow-wave sleep) and an active phase (also known as active sleep or AS, REM sleep, paradoxical sleep), which are distinguished by many differences in biochemistry, physiology, psychology, and behavior. Recordings of electrical activity changes of the brain (electroencephalogram or EEG), eye movements (electrooculogram or EOG), and chin muscle tone (electromyogram or EMG) are used to define the various stages and substages of sleep (**Fig. 1**). *See* ELECTROENCEPHALOGRAPHY.

Sleep cycle. If sleepy enough, most people can fall asleep under almost any condition even while sitting up or being subjected to loud noise, bright lights, and painful electric shocks. However, humans usually lie down and decrease environmental stimulation when preparing to go to sleep. After lying in bed for a few minutes in a quiet, dark room, drowsiness usually sets in. The subjective sensation of drowsiness can be objectively indexed by a corresponding change in brain waves (EEG activity): formerly continuous alpha rhythms (**Fig. 2***a*) gradually break up into progressively shorter trains of regular alpha waves and are replaced by low-voltage mixed-frequency EEG activity. When less than half of an epoch [usually the staging epoch is the 20–30 seconds it takes to fill one page of polygraph (sleep recording) paper] is occupied by continuous alpha rhythm, sleep onset is considered to have occurred and stage 1 sleep is scored (Fig. 2*b*). At this stage, the EOG usually reveals slowly drifting eye movements (SEMs) and muscle tone might or might not decrease. Awakenings at this point frequently yield reports of hypnagogic (leading into sleep) imagery, which can often be extremely vivid and bizarre.

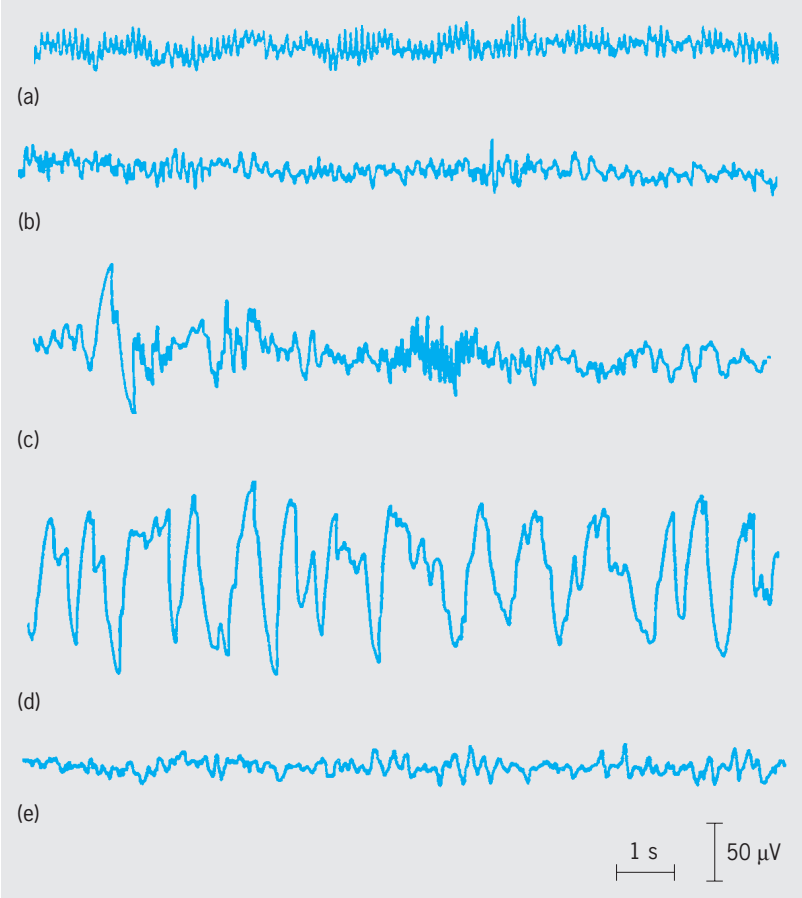

Fig. 2. Human EEG associated with different stages of sleep and wakefulness. (*a*) Relaxed wakefulness (eyes shut) shows rhythmic 8-12-Hz alpha waves. (*b*) Stage 1 non-REM sleep shows mixed frequencies, especially 3-7-Hz theta waves. (*c*) Stage 2 non-REM sleep shows 12–14-Hz sleep spindles and K-complexes. (*d*) Delta sleep shows large-amplitude (>75 μV) 0.5-2-Hz delta waves. (*e*) REM sleep shows low-amplitude, mixed frequencies with sawtooth waves.

Stage 1 is a very light stage of sleep described by most subjects as "drowsing" or "drifting off to sleep." Normally, it lasts only a few minutes before further EEG changes occur, defining another sleep stage. It is at this point that startlelike muscle jerks known as hypnic myoclonias or hypnic jerks occasionally briefly interrupt sleep. As the subject descends deeper into sleep, the EEG of stage 2 sleep is marked by the appearance of relatively high-amplitude slow waves called K-complexes as well as 12–14-Hz rhythms called sleep spindles (Fig. 2*c*). The EOG would generally indicate little eye movement activity, and the EMG would show somewhat decreased muscle tone. Reports of mental activity from this stage of sleep are likely to be less bizarre and more realistic than those from stage 1. However, light sleepers sometimes report lengthy and vivid dreams upon awakening from stage 2 sleep, especially late in the sleep cycle.

After several minutes in stage 2, high-amplitude slow waves (delta waves) gradually begin to appear in the EEG. When at least 20% of an epoch is occupied by these (1-2-Hz) delta waves, stage 3 is defined. Usually this slow-wave activity continues to increase until it completely dominates the appearance of the EEG. When the proportion of delta EEG activity exceeds 50% of an epoch, the criterion for the deepest stage of sleep, stage 4, is met. During stages 3 and 4, often collectively referred to as delta sleep (Fig. 2*d*), the EOG shows few genuine eye movements but is obscured by the high-amplitude delta waves. Muscle tone is normally low, although it can be remarkably high, as when sleepwalking or sleep-talking occurs. Recall of mental activity on arousal from delta sleep is generally very poor and fragmentary and is more thoughtlike than dreamlike. It should be noted that cognitive functioning immediately after abrupt wakening from sleep is likely to carry over some of the characteristics of the preceding sleep state. This phenomenon, known as sleep inertia, can be used as an experimental tool for studying (by inference) cognition during different stages of sleep. For example, studies have shown that stories told in response to thematic apperception test cards (a psychological test used to uncover the dynamics of personality) immediately after awakening from REM sleep are more creative than those told right after non-REM sleep. The combination of confusion, disorientation, cognitive and motor impairment, and amnesia sometimes following abrupt awakening from delta sleep is justifiably termed sleep drunkenness and can be genuinely life-threatening for people such as navy fighter pilots on alert who must fly their highly complicated supersonic jets off an aircraft carrier in the middle of the night within 5 minutes of being awakened.

After about 90 minutes, the progression of sleep stages is reversed, back through stage 3 and stage 2 to stage 1 again. But now the EMG shows virtually no activity at all (Fig. 2*e*), indicating that muscle tone has reached its lowest possible level, and the EOG discloses the occurrence of rapid eye movements—at first only a few at a time, but later in dramatic profusion. This is REM (or active) sleep. Breathing rate and heart rate become more rapid and irregular, and both males and females show signs of sexual arousal (for example, erections and increased vaginal blood flow). Brain metabolic rate increases to levels that typically exceed the waking state average. This state of intense brain activation is normally accompanied by experiences that seem vividly real while they last, but often vanish within seconds of waking. When people are abruptly awakened from REM sleep, 80–90% of the time they recall vivid and sometimes extremely detailed dreams.

While all this activity is happening in the brain, the body remains almost completely still (except for small twitches), because it is temporarily paralyzed during REM sleep to prevent dreams from being acted out. The brainstem system that causes the paralysis of REM sleep does not always inactivate immediately upon awakening. The resulting experience, known as sleep paralysis, can be terrifying, but it is quite harmless and normal if it occurs only occasionally. However, frequent sleep paralysis can be a symptom of a disorder of REM sleep called narcolepsy. *See* SLEEP DISORDERS.

After a REM period lasting perhaps 5 to 15 minutes, a young adult will typically go back through

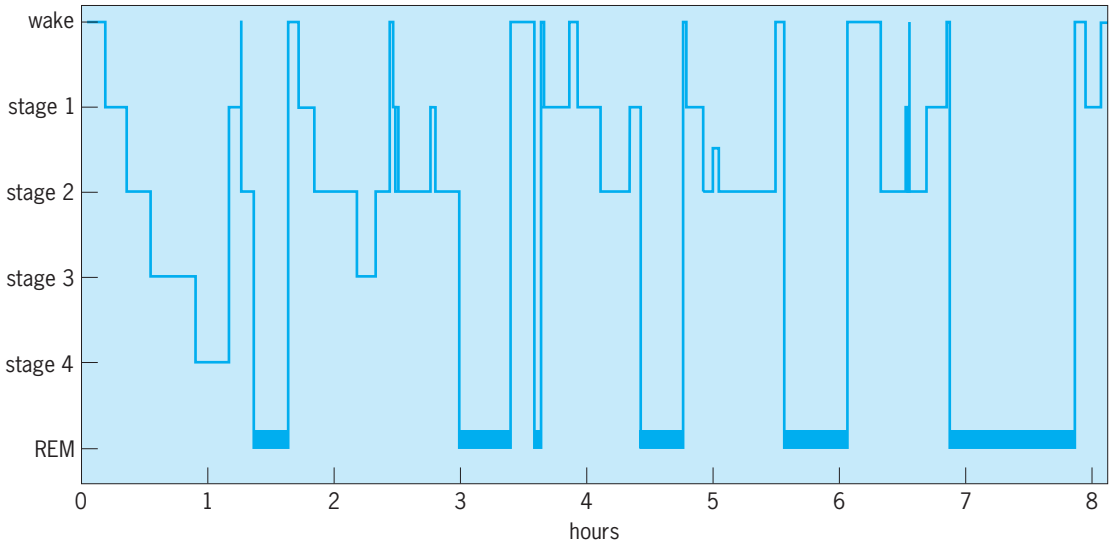

Fig. 3. Sleep histogram showing sleep cycles during a typical night. Note the increasing amount of REM sleep as the night progresses.

the preceding cycle stages, dreaming vividly three or four more times during the remainder of the night (**Fig. 3**), with two major modifications. First, decreasing amounts of slow-wave EEG activity (stages 3 and 4 or delta sleep) occur in each successive cycle. Later in the night, after perhaps the second or third REM period, no delta sleep appears on the EEG at all, only non-REM, stage-2, and REM sleep. Second, as the night proceeds, successive REM periods tend to increase in length, up to a point. While the first REM period commonly lasts less than 10 minutes, later REM periods often last 30 to 40 minutes, and even an hour or more is not exceptionally uncommon late in the sleep cycle. At the same time that the REM periods are getting longer, the intervals between them tend to decrease in length from the approximately 90 minutes characteristic of the first part of the night to as little as 20 or 30 minutes in the late morning. The fact that for humans most REM occurs in the last portion of the sleep cycle as dawn approaches suggests that REM serves a function related to preparation for waking activity.

Finally, after four or five periods of dreaming sleep, the sleeper wakes up (for perhaps the tenth time during the night) and gets up for the day. It may be difficult to believe that brief awakenings occur this frequently during an average night; however, they are promptly forgotten. This retrograde amnesia is a normal feature of sleep: information in short-term memory at sleep onset is usually not transferred into more permanent storage.

Measuring sleepiness. Probably the most obvious relationship between sleep and wakefulness is the fact that variations in the amount of sleep received predict subsequent variations in alertness the following day or days. Too little sleep one night is likely to result in increased sleepiness and decreased alertness during the following day. Variations in alertness/sleepiness form a continuum of states ranging from hyperexcitation and seizure at one extreme to coma and death at the other. Abnormal or pathological states of excessive activation are accompanied by restless hypervigilance, distractibility, agitation, insomnia, and mania. Optimal levels of activation are accompanied by feelings of energy, interest, and motivation to act. Suboptimal activation is accompanied by decreases in such feelings and motivations, along with varying levels of sleepiness, a specific drive motivating rest, inaction, and sleep. Whether sleepiness is experienced as pleasant or unpleasant depends on whether conditions allow consummation of the desire to sleep.

There are various behavioral indications of sleepiness (for example, yawning, head nodding, eye closure, and performance deficits), as well as subjective self-evaluation scales. For example, the Stanford Sleepiness Scale consists of statements that describe seven states of sleepiness/alertness varying from "wide awake" (1), through "somewhat foggy, let down" (4), to "no longer fighting sleep, sleep onset soon, having dreamlike thoughts" (7). (Of course, "asleep" is off the scale.) Subjective scales are quick, cost-effective, and easy to administer but offer relatively low reliability compared with objective measures.

The "gold standard" for quantifying sleepiness is the Multiple Sleep Latency Test (MSLT). It is based on the assumption that degrees of sleepiness are best measured by the speed with which a person falls asleep under standard conditions. Typically, the MSLT is carried out in a sleep laboratory with subjects wired for sleep recordings (EEG, EOG, and EMG). The MSLT is typically performed in five sessions, starting at 10 a.m. and after every 2 hours thereafter until 6 p.m. Care is taken to ensure that the subject is fully comfortable during each session. After taking care of any requirements for comfort (such as a drink of water and going to the bathroom) in time for the scheduled session, the subject enters a dark, quiet bedroom and the electrodes placed on the

subject's body are plugged into the polysomnograph. The subject lies on the bed, and at the scheduled start time is told to close his or her eyes and try to go sleep. Usually, prominent alpha waves are visible in the EEG at first. Sleep onset is scored at the abrupt disappearance of alpha waves, and sleep latency is measured as the time from eyes-shut to sleep onset. Immediately at sleep onset, the subject is promptly awakened and required to get out of bed, preventing the accumulation of more than a few seconds of sleep per trial. If the subject fails to fall asleep within 20 minutes, the trial ends and the sleep latency is scored as the maximum, 20 minutes.

The results from the multiple tests are averaged to provide a single measure of daytime alertness. MSLT scores greater than 15 minutes merit the label "Good Alertness"; scores more than 10 but less than 15 minutes, "Can Get By"; greater than 5 but less than 10 minutes, "Borderline"; and scores less than 5 minutes, "Twilight Zone" because memory and clarity of thinking are usually substantially impaired with this degree of sleepiness.

The quickest method of rapidly increasing sleepiness is total sleep deprivation. An MSLT study showed that young adults were in the Twilight Zone after a single night of total sleep deprivation. After two 9-hour recovery nights of sleep following two nights of total sleep deprivation, they were out of the Twilight Zone in only half the day's MSLT sessions. The fact that two full nights of sleep were insufficient to return alertness to full wakefulness suggests that a sleep debt has been incurred by the total sleep deprivation. But total sleep deprivation is not the usual source of increased daytime sleep tendency, because even partial sleep deprivation has cumulative effects. For example, after three baseline nights with 9 hours in bed per night, subjects were restricted to 5 hours in bed per night for seven consecutive nights, and the MSLT was administered on each subsequent day. Subjects showed a progressive increase in sleepiness for every night of the study. Subjects' scores dropped from Good Alertness to Can Get By after only one night of partial sleep deprivation. By night 5 of partial sleep deprivation, they reached Borderline and would have been in the Twilight Zone by night 10 if the study had continued.

The MSLT also reveals that it is typical that sleepiness reaches a peak in midafternoon. A person can be very sleepy at 4 p.m. but by 6 or 8 p.m. feel quite alert–without having received any sleep in the meantime. This is due to another factor determining alertness/sleepiness in addition to immediately prior sleep or wakefulness, namely a circadian rhythm. According to the opponent process model, there are two independent processes involved which together produce the circadian process of sleep and wakefulness. One process, whose function is to induce and maintain sleep, is called sleep homeostasis. The other process, whose function is to induce and maintain alertness, is called clock-dependent alerting.

A variation on the sleep deprivation technique is selective sleep deprivation. By waking subjects whenever they show a particular sleep stage, it is theoretically possible to determine the effects of that particular sleep stage, usually REM sleep. REM deprivation was pioneered by W. C. Dement, and the early studies produced dramatic demonstrations of the need for REM sleep. Later, more carefully controlled studies found rather more modest effects of REM deprivation on personality, cognitive function, and motivation. Probably the most consistent effect is an intensification of drives and a lowering of thresholds for cortical stimulation. What manifests as hypersexuality and hyperphagia in cats shows itself in humans as increased irritability and perhaps some tie-loosening. In fact, it is probably the threshold-lowering effect of REM deprivation that causes a single night of REM deprivation to significantly improve mood in depressed patients. *See* AFFECTIVE DISORDERS.

Both total and selective quiet-sleep deprivation produces corresponding increases of sleep-drive and intensification of delta wave activity (Fig. 2*d*). After a night of total sleep deprivation, the first few hours of subsequent sleep will be almost exclusively non-REM sleep, suggesting that quiet sleep may be even more functionally important than REM sleep.

Evolution and function of sleep. What selective forces were responsible for the evolution of sleep? The major source of daily change in local conditions on Earth derives from day and night. It is a fact of life that most organisms are adapted to either the dark or light phase of the cycle. Therefore, a biological process that limits activity to the phase of the cycle to which the organism is adapted will enhance survival. Most likely sleep developed out of the rest phase of the BRAC, allowing organisms to minimize interactions with the world during the less favorable phase, while engaging in a variety of "off-line" internal maintenance operations, including growth, repair, synthesis of biochemicals consumed during waking activity, and energy conservation, as well as memory consolidation. The fact that different species have many differences in sleep structure, process, and function fits with the idea that sleep serves the specific adaptive needs of each species. For example, quiet sleep and active sleep have different evolutionary histories and serve different biological functions. Quiet sleep appears to be an older form of sleep with simpler and more universal functions related to energy conservation, growth, and restoration. Active sleep is a mammalian invention with functions that appear to be related to specifically mammalian needs such as live birth. The portion of total sleep composed of REM is at its highest level perinatally: newborn humans spend 8 hours per day in REM sleep, with even more time during the last 6 weeks before birth. The time of maximal REM corresponds to the time of maximal growth of the brain. A number of theorists, including H. Roffwarg and M. Jouvet, suggest that this is not a coincidence but points to the main evolutionary function of REM: to provide a source of endogenous stimulation supporting the unfolding of genetic programming and self- organization of the brain. REM time decreases continuously until puberty, when it reaches the 2 hours or less it maintains throughout

the remainder of life. The fact that REM time does not decrease to zero after the full development of the nervous system suggests secondary adaptive advantages afforded by REM during adulthood, which may include facilitation of difficult learning and preparation of the brain for optimal functioning on arousal. This latter function would account for the unequal distribution of REM throughout the night, with most REM closest to the end of the night. The amount of time spent asleep by different species, and especially the amount of time spent in the deep stage of REM, varies widely as a function of lifestyle. For example, predators sleep deeply and long, while prey species sleep little with still less REM.

Both phases of sleep appear to serve a variety of distinct functions related to learning and memory. For example, quiet sleep has been shown to enhance consolidation of declarative memory, while active sleep enhances consolidation of procedural memory.

As mentioned above, the delta activity of NREM sleep is homeostatically regulated, increasing with wakefulness and decreasing during sleep. In 2004, R. Huber, G. Tononi, and colleagues found local enhancement of delta activity in specific brain areas associated with a particular motor-learning task, providing strong support for a cellular need for sleep, as well as a specific role in learning consolidation.

A number of studies have recently established that performance improvements on a variety of perceptual, motor, and cognitive tasks following a night of sleep are specifically due to sleep rather than, say, the passage of time.

Although the modern understanding of sleep and dreaming has had the benefit of half a century of scientific research, there are still no simple and conclusive answers to the questions of why we sleep. The more we learn, the more complex sleep seems.

Dreams. From the biological perspective, the basic task of the brain is to predict and control the organism's actions and regulate those actions to achieve optimal outcomes (in terms of survival and reproduction). To accomplish this task, the brain in some sense internally "models" the world. The waking brain bases the features of its world model primarily on the current information received from the senses and secondarily on expectations derived from past experience.

In contrast, the sleeping brain acquires little information from the senses. Therefore in sleep, the primary sources of information available to the brain are the current motivational state of the organism (for example, current concerns, fears, hunger, and desires) and past experience (for example, both species-specific instincts and personal memories of recent and associated past experiences of potential relevance). According to this theory, dreams result from brains using internal information to create a simulation of the external world, in a manner directly parallel to the process of waking perception, minus most sensory input. *See* PERCEPTION.

Dreaming experience is commonly viewed as qualitatively distinct from waking experience. Dreams are often believed to be characterized by lack of reflection and an inability to act deliberately and with intention. However, this view has not been based on equivalent measurements of waking and dreaming state experiences. To achieve equivalence, it is necessary to evaluate waking experience retrospectively, in the same way that dreams are evaluated. In 1995, S. LaBerge et al. published the results of a study examining cognition in dreaming and waking by sampling recall of recent experiences from both conditions and collecting responses to questions pertaining to thought processes during the period recalled. Each of 167 subjects contributed detailed reports and questionnaires from a randomly selected sample of waking experience and a corresponding sample of dreaming experience. After recording each experience, the subjects answered a series of questions about the experience described, assessing whether or not they had engaged in a number of cognitive and other activities. These included deliberate choice between alternatives, reflection, sudden distraction of attention, focused intention, self-consciousness, and emotion. Differences appeared between waking and dreaming samples for several variables: 49% of subjects reported deliberate choice from dreaming versus 74% from waking; 41% of subjects reported public self-consciousness from dreams versus 30% from waking; and 86% reported emotion from dreams versus 74% from waking. It is notable that significant differences between dreaming and waking were not evident for other cognitive activities, and none of the measured cognitive functions were absent or rare in dreams. In terms of the typical, waking and dreaming experiences have much in common, with dreams being characterized by somewhat more emotionality and somewhat less choice.

However, less typical examples reveal a wider range of qualities of consciousness in dreaming than in the usual waking state. For example, at one end of the spectrum, dreams at times exhibit cognitive and perceptual errors similar to those produced by brain damage. However, dream bizarreness also shows the creative recombinatory potential of the dreaming brain. Moreover, as illustrated by lucid dreams, dreaming consciousness can also be as volitional and rational as waking consciousness. *See* CONSCIOUSNESS.

Cognizant or lucid dreaming. Typically, people are not explicitly aware of the fact that they are dreaming while they are dreaming; however, at times a remarkable exception occurs. During such lucid dreams it is possible to freely remember the circumstances of waking life, to think clearly, and to act deliberately upon reflection or in accordance with plans decided upon before sleep, all while experiencing a dream world that seems vividly real.

Although accounts of lucid dreaming go at least as far back as Aristotle, until recently dream reports of this sort were received with considerable skepticism. In the absence of objective proof, sleep researchers doubted that the dreaming brain was capable of such a high degree of mental functioning and consciousness. A new technique involving eye movement signals, developed independently by

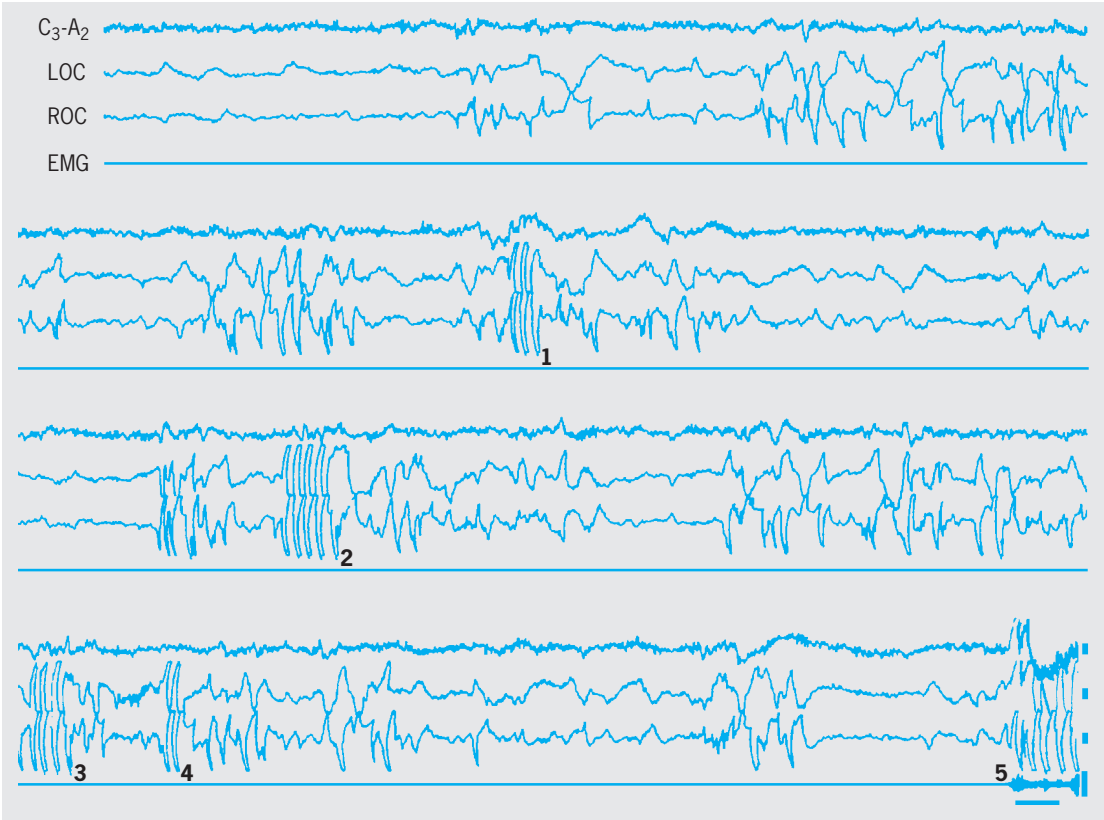

Fig. 4. Voluntary eye movement signaling marks lucid dreaming during uninterrupted REM sleep. Four channels of physiological data [central EEG (C3-A2), left and right eye movements (LOC and ROC), and chin muscle tone (EMG)] from the last 8 minutes of a 30-minute REM period are shown. Upon awakening, the subject reported having made five eye movement signals (labeled 1–5). The first signal (1, LRLR) marked the onset of lucidity. Emotional sweating artifacts can be observed in the EEG at this point. During the following 90 seconds the subject "flew about," exploring his dream world until he believed he had awakened, at which point he made the signal for awakening (2, LRLRLRLR). After another 90 seconds, the subject realized that he was still dreaming and signaled (3) with three pairs of eye movements. Realizing that this was too many, he correctly signaled with two pairs (4). Finally, upon awakening 100 seconds later, he signaled appropriately (5, LRLRLRLR). Calibrations are 50 μV and 5 seconds.

researchers in the United States and in England in the late 1970s, proved the reality of lucid dreaming. The technique was based on earlier studies which found that the directions of eye movements recorded during REM sleep sometimes exactly corresponded to the directions in which subjects reported they had been looking in their dreams. It was reasoned that if lucid dreamers can in fact act volitionally, they should be able to prove it by making a prearranged eye movement signal marking the exact time they became lucid. Using this approach, researchers verified reports of lucid dreams from five subjects by eye movement signals. All of the signals, and therefore lucid dreams, had occurred during uninterrupted REM sleep (**Fig. 4**). This result has been replicated in a number of independent laboratories around the world.

Lucid dreaming as a method for studying consciousness. The eye movement signaling experiments illustrate an important approach to the study of dreaming consciousness. The attempt to apply rigorous scientific methodology to the study of such phenomena as mental imagery, hallucinations, dreaming, and in general, conscious processes faces a major challenge: The most direct account available of the private events occurring in a person's mind is that individual's subjective report. However, subjective reports are difficult to verify objectively, and introspection is far from being an unbiased or direct process of observation. There are two strategies likely to increase the reliability of subjective reports: (1) the use of highly trained subjects (in the context of dream research, this is best achieved with lucid dreamers) who are skillful reporters of consciousness, and (2) the use of a psychophysiological approach to provide validation of subjective reports by correlating them with physiological measures. *See* HALLUCINATION.

Using these approaches in a series of studies, researchers found that various dreamed experiences (including time estimation, breathing, singing, counting, and sexual activity) produce effects on the dreamer's brain (and to a lesser extent, body) remarkably similar to the physiological effects that are produced by actual experiences of the corresponding events while awake. The results of these studies support the following picture: During REM dreaming, the events that a person consciously experiences (or seems to experience) are the results of patterns of central nervous system activity that produce, in turn, effects on the autonomic nervous system and that are body-isomorphic (similar in appearance) to the effects that would occur if a person actually experienced the corresponding events while awake. This explains, in part, why people regularly mistake

dreams for reality. To the functional systems of neuronal activity that construct the experiential world model of consciousness, dreaming of perceiving or doing something is equivalent to actually perceiving or doing it.

Dream recall. The average person remembers a dream only once or twice a week. Given the fact that every person dreams each night, that leaves at least 95% of most dreams forgotten (in REM alone; assuming five REM periods per night with perhaps two dreams per REM period). A variety of theories have been put forward suggesting often fanciful explanations of why dreams are so easily forgotten, ranging from Sigmund Freud's belief that dreams are repressed because they contain so much taboo dream thought, to F. Crick's view that the content of dreams is what the brain is trying to unlearn, and therefore ought not to be remembered. Of course, standard memory theory explains much of what is remembered of dreams and what is not, which aspects of dream content are more readily recalled and which aspects are not, and so on. *See* MEMORY; PSYCHOANALYSIS.

But why are dreams so much more difficult to recall than waking experiences? Evolution may provide the answer. As S. LaBerge reasoned in 1985, humans learn that dreams are distinct from other experiences by talking with other humans. Nonspeaking animals, however, have no way to tell each other how to distinguish dreams from reality. Explicit dream recall would thus be disorienting and maladaptive for all the nonlinguistic mammals having REM sleep during the 140 million years since REM evolved. The fact that explicit dream recall obviously is maladaptive for cats, dogs, bats, whales, and all of the rest of the mammalian dreamers except humans very likely explains why dreams are difficult for humans to recall. The dreams may be so, according to this view, because of natural selection. Like all mammals, humans and their ancestors might have been protected from dangerous confusion by the evolution of mechanisms that made forgetting dreams the normal course of affairs. *See* ORGANIC EVOLUTION. Stephen LaBerge

Bibliography. K. M. Fenn, H. C. Nusbaum, and D. Margoliash, Consolidation during sleep of perceptual learning of spoken language, *Nature*, 425:614–616, 2003; P. Fenwick et al., Lucid dreaming: Correspondence between dreamed and actual events in one subject during REM sleep, *Biol. Psychol.*, 18:243–252, 1984; S. Fischer et al., Sleep forms memory for finger skills. *Proc. Natl. Acad. Sci. USA*, 99:11987–11991, 2002; R. Huber et al., Local sleep and learning, *Nature*, 430:78–81, 2004; S. LaBerge, *Lucid Dreaming*, J. P. Tarcher, Los Angeles, 1985; S. LaBerge, Lucid dreaming: Psychophysiological studies of consciousness during REM sleep, in R. R. Bootsen, J. F. Kihlstrom, and D. L. Schacter (eds.), *Sleep and Cognition*, pp. 109–126, APA Press, Washington, DC, 1990; S. LaBerge, T. Kahan, and L. Levitan, Cognition in dreaming and waking, *Sleep Res.*, 24A:239, 1995; N. Malcolm, *Dreaming*, Routledge, London, 1959; S. C. Mednick et al., The restorative effect of naps on perceptual deterioration, *Nat. Neurosci.*, 5:677–681, 2002; A. Rechtschaffen, The single-mindedness and isolation of dreams, *Sleep*, 1:97–109, 1978; H. Roffwarg et al., Dream imagery: Relationship to rapid eye movements of sleep, *Arch. Gen. Psychiat.*, 7:235–238, 1962; C. Smith, Sleep states and memory processes, *Behav. Brain Res.*, 69:137–145, 1995; R. Stickgold, L. James, and J. A. Hobson, Visual discrimination learning requires sleep after training. *Nat. Neurosci.*, 3:1237–1238, 2000.

Sleep disorders

A reversible state during which the individual's voluntary functions are suspended but the involuntary functions, such as circulation and respiration, are uninterrupted; the sleeping subject assumes a characteristic posture with relative immobility and decreased responses to external stimuli.

The sleep state can be divided into nonrapid eye movement (NREM) and rapid eye movement (REM) sleep; NREM sleep, in turn, is subdivided into four stages based on electroencephalographic criteria. The first REM in a normal adult occurs 60–90 min after sleep onset, and there are usually four or five NREM-REM cycles, each lasting for 90–110 min. Most dreams occur during REM sleep. The sleep-wake pattern for humans follows a circadian rhythm. An average adult needs approximately 7–8 h of sleep, but in elderly people there are frequent awakenings.

Clinical disorders. The latest International Classification of Sleep Disorders (ICSD) divides sleep disorders into eight broad categories: (1) insomnias; (2) sleep-related breathing disorders, including obstructive sleep apnea syndrome; (3) hypersomnias of central origin (including narcolepsy) that are not secondary to sleep-related breathing disorders, circadian rhythm sleep disorders, or other causes of disturbed night sleep; (4) circadian rhythm sleep disorders; (5) parasomnias (abnormal movements or behavior intruding into sleep, such as sleepwalking, sleep terrors, confusional arousals, REM behavior disorders, nightmare, sleep enuresis); (6) sleep-related movement disorders, such as restless legs syndrome; (7) isolated symptoms (for example, long sleeper, short sleeper, snoring, sleep talking, etc.; these symptoms belong to the borderline between normal and abnormal sleep or are part of the continuum of normal to abnormal events in sleep), including apparently normal variants and unresolved issues due to insufficient information; and (8) other sleep disorders, which do not conform to the listing outlined above or may have features overlapping some other conditions listed in other categories. In addition, ICSD includes Appendix A, which includes among other conditions sleep-related epilepsy; and Appendix B, which includes psychiatric and behavioral disorders causing sleep dysfunction.

Most common complaints. Excessive daytime sleepiness, insomnia (sleeplessness), and abnormal movements and behavior during sleep are the major sleep complaints. Individuals with hypersomnia complain of excessive daytime sleepiness. An individual with insomnia may have difficulty initiating or maintaining

sleep; repeated awakenings or early morning awakenings; or daytime fatigue and impairment of performance. The commonest cause of insomnia is psychiatric or psychophysiologic disorder (for example, depression, anxiety, or stress), but other causes include medical disorders or pain. Circadian-rhythm sleep disorders may be associated with either insomnia or hypersomnia.

Obstructive sleep apnea syndrome. There are two forms of sleep apnea. In central apnea, both the airflow at the mouth and nose and the effort by the diaphragm decrease. In obstructive apnea, the airflow stops but the effort by the diaphragm continues. Obstructive sleep apnea syndrome is common in middle-aged and elderly obese men. The individual stops breathing for many seconds during sleep and wakes up repeatedly with a loud snore. Repeated apneas associated with reduction of oxygenation in the blood may cause serious complications, such as high blood pressure and irregularities and/or failure of the heart. Obstructive sleep apnea syndrome results from excessive relaxation of the upper airway muscles during sleep, coupled with an unknown dysfunction of the respiratory neurons.

Narcolepsy. Narcoleptic sleep attacks usually begin in individuals between the ages of 15 and 25. Narcolepsy is characterized by an irresistible desire to sleep, and the attacks may last 15–30 min. Most individuals also have cataplexy, during which there is transient loss of muscle tone that causes the individual to fall to the ground or slump the head forward for a few seconds. These attacks are often triggered by an emotional outburst. Other symptoms may include fearful dreams or feeling of loss of power at sleep onset or offset or nocturnal sleep disturbance. Narcolepsy cannot be cured.

Restless legs syndrome. This syndrome is another lifelong condition occurring during middle age. It is characterized by the urge to move the legs, usually accompanied by intensely disagreeable feelings in the legs at rest and repose, predominantly in the evening, with partial or total relief from these symptoms sometimes occurring. Most individuals with this problem also have periodic limb movements in sleep.

Diagnosis and treatment. Overnight polysomnography and multiple sleep latency tests are both important for sleep disorders. Overnight polysomnography involves recording of brain waves, muscle activities, eye movements, heart activity, airflow at the nose and mouth, respiratory effort, and oxygen saturation. Polysomnography is needed for individuals with excessive daytime sleepiness, narcolepsy, parasomnias, restless legs–periodic limb movements, and nocturnal seizures. The multiple sleep latency test is important in documenting pathologic sleepiness and diagnosing narcolepsy. The test records brain waves, muscle activities, and eye movements.

Treatment of obstructive sleep apnea syndrome consists of avoidance of sedatives or alcohol consumption, weight loss, and use of continuous positive airway pressure to open the upper airway passage so that obstructive apneas can be eliminated. The optimal continuous positive airway pressure is determined during overnight polysomnography. Narcolepsy is treated by stimulants and short naps, while cataplexy is treated with tricyclic antidepressants, selective serotonin reuptake inhibitors, and sodium oxybate. Chronic insomnia can best be treated by sleep hygiene and behavioral modification. A short-acting nonbenzodiazepine (for example, zolpidem, zaleplon, eszopiclone) or a short: or intermediate-acting benzodiazepine hypnotic may be used for transient insomnia. Most parasomnias do not require special treatment. However, psychotherapy may be helpful in some cases. Circadian rhythm disorders may be treated by bright light and chronotherapy. *See* BIOLOGICAL CLOCKS; SLEEP AND DREAMING.

Sudhansu Chokroverty

Bibliography. S. Chokroverty (ed.), *Sleep Disorders Medicine: Basic Science, Technical Considerations and Clinical Aspects*, 2d ed., 1999; *The International Classification of Sleep Disorders, Diagnosis, and Coding Manual*, 2d ed., American Academy of Sleep Medicine, 2005; M. Kryger, T. Roth, and W. Dement (eds.), *Principles and Practice of Sleep Medicine*, 4th ed., 2005.

Slider-crank mechanism

A four-bar linkage with output crank and ground member of infinite length. A slider crank (**Fig. 1**) is most widely used to convert reciprocating to rotary motion (as in an engine) or to convert rotary to reciprocating motion (as in pumps), but it has numerous other applications. Positions at which slider motion reverses are called dead centers. When crank and connecting rod are extended in a straight line and the slider is at its maximum distance from the axis of the crankshaft, the position is top dead center (TDC); when the slider is at its minimum distance from the axis of the crankshaft, the position is bottom dead center (BDC). *See* FOUR-BAR LINKAGE.

For a given crank throw, or radius, the action of the slider approaches simple harmonic motion as the length of the connecting rod is increased. Maximum accelerations occur at reversal of the slider. For constant angular velocity of the crank, slider acceleration at TDC is somewhat greater, and at BDC somewhat less, than accelerations that would occur if the motion were simple harmonic.

Although the idea of combining a crank with a connecting rod is quite old (fifteenth century), the crank was not successfully applied to a steam engine

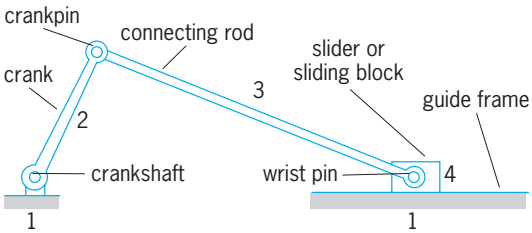

Fig. 1. Principal parts of slider-crank mechanism.

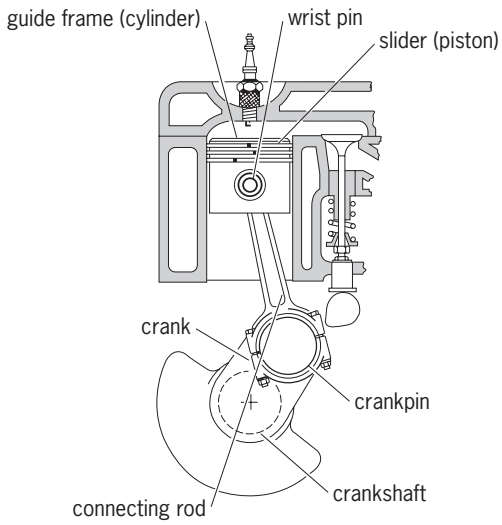

Fig. 2. Slider crank applied to internal combustion engine.

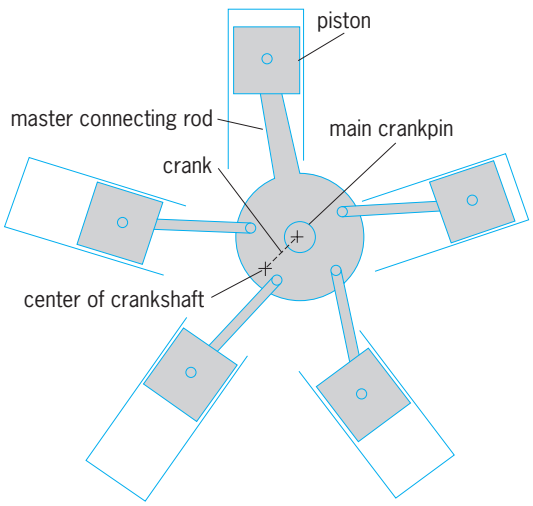

Fig. 3. Multiple-crank mechanism in radial engine.

until 1780; the completion of the linkage to include a slider had to wait for satisfactory lubrication.

Many attempts were made between 1780 and 1830 to produce rotary motion directly and thus eliminate the need for the slider-crank mechanism in prime movers. Hundreds of different rotary engine designs (many of which employed slider-crank linkage in a form not recognized by the inventor) were proposed but, for a prime mover that depends for its operation upon such a fluid as steam or air within a chamber, no arrangement has been found superior to the conventional one described here.

Internal combustion engine. The conventional internal combustion engine employs a piston arrangement in which the piston becomes the slider of the slider-crank mechanism (**Fig. 2**). Although satisfactory for engines of moderate life-span, the reversal of side thrust on the piston twice during each revolution complicates the problem of keeping the piston tight enough in the cylinder to contain the working medium in the combustion space. Because of angularity of the connecting rod, most wear occurs at the lower end of the cylinder. In some large, low-speed engines, a crosshead, differing in arrangement but similar in principle to the crosshead of a steam engine, is used to reduce cylinder wear.

Radial engines for aircraft employ a single master connecting rod to reduce the length of the crankshaft (**Fig. 3**). The master rod, which is connected to the wrist pin in a piston, is part of a conventional slider-crank mechanism. The other pistons are joined by their connecting rods to pins on the master connecting rod.

Reciprocating compressors. To convert rotary motion into reciprocating motion, the slider crank is part of a wide range of machines, typically pumps and compressors. The crankshaft may be driven through belting or by an electric motor. Portable, reciprocating compressors use pistons; stationary compressors generally employ crossheads and guides. *See* COMPRESSOR.

Other applications. Another use of the slider crank is in toggle mechanisms, also called knuckle joints. The driving force is applied at the crankpin (Fig. 1) so that, at TDC, a much larger force is developed at the slider. This is limited by practicality to the forces at which the links and pins yield.

Other forms that the slider crank can take can be studied by inversion (**Fig. 4**). Compared to a

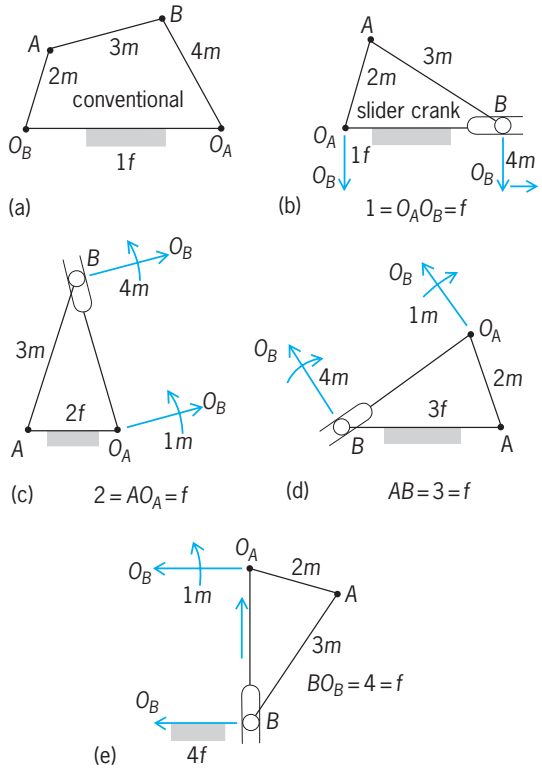

Fig. 4. Inversion of slider-crank mechanism.
(*a*) Conventional four-bar linkage of which slider crank is a special case. (*b*) Link 4 made infinitely long, and link 1, which carries the slider guide, held fixed, produce the usual slider-crank mechanism. (*c*) Intermediate link 2 held fixed. (*d*) Link 3, carrying the slider head, held fixed. (*e*) Virtual link 4, to which slider *B* is pinned, held fixed. Links marked *f* are fixed; those marked *m* move.

conventional four-bar linkage (Fig. 4*a*) where link 1 is the fixed member, the slider crank replaces finite moving link 4 with the slider (Fig. 4*b*); in this form link 1 remains the fixed member. Although moving member 4 loses its separate identity, comparison of angles ABO_B in Fig. 4*a* and *b*, and comparison of angles BO_BO_A in the two diagrams, shows the actions of the two mechanisms to involve the same number of links and turning pairs within their closed loops. By successively holding each link fixed, the alternative forms of Fig. 4*c*–*e* are obtained. Douglas P. Adams

Bibliography. H. H. Mabie and C. F. Reinholtz, *Mechanisms and Dynamics of Machinery*, 4th ed., 1987; G. H. Ryder and M. D. Bennett, *Mechanics of Machines*, 2d ed., 1990; J. E. Shigley and J. J. Uiker, *Theory of Machines and Mechanisms*, 2d ed., 1994.

Slip (electricity)

A numerical value used in describing the performance of electrical couplings and induction machines. In an electrical coupling, slip is defined simply as the difference between the speeds of the two rotating members. In an induction motor, slip is a measure of the difference between synchronous speed and shaft speed.

When the stator windings of an induction motor are connected to a suitable alternating voltage supply, they set up a rotating magnetic field within the motor. The speed of rotation of this field is called synchronous speed, and is given by Eq. (1) or Eq. (2),

$$\omega_s = \frac{4\pi f}{p} \qquad \text{rad/s} \tag{1}$$

$$n_s = 120\frac{f}{p} \qquad \text{rev/min} \tag{2}$$

where f is the line frequency and p is the number of magnetic poles of the field. The number of poles is determined by the design of the windings. In accord with Faraday's voltage law, a magnetic field can induce voltage in a coil only when the flux linking the coil varies with time. If the rotor were to turn at the same speed as the stator field, the flux linkage with the rotor would be constant. No voltages would be induced in the rotor windings, no rotor current would flow, and no torque would be developed. For motor action it is necessary that the rotor windings move backward relative to the magnetic field so that Faraday's law voltages may be induced in them. That is, there must be slip between the rotor and the field. *See* ELECTROMAGNETIC INDUCTION; INDUCTION MOTOR.

The amount of slip may be expressed as the difference between the field and rotor speeds in revolutions per minute or radians per second. However, the slip of an induction motor is most commonly defined as a decimal fraction of synchronous speed, as in Eq. (3) or Eq. (4). Here n is the motor speed in

$$s = \frac{n_s - n}{n_s} \tag{3}$$

$$s = \frac{\omega_s - \omega}{\omega_s} \tag{4}$$

revolutions per minute, ω is its speed in radians per second, and s is the slip, or more properly the per unit slip. Typical full-load values of slip for an induction motor range from 0.02 to 0.15, depending on rotor design. Slip is sometimes expressed in percent of synchronous speed, rather than per unit. If an induction machine is driven faster than synchronous speed, the slip becomes negative, and the machine acts as a generator, forcing energy back into the electrical supply line. *See* ELECTRIC ROTATING MACHINERY. George McPherson, Jr.

Bibliography. Institute of Electrical and Electronics Engineers, *The Authoritative Dictionary of of IEEE Standards Terms (IEEE 100)*, 7th ed., 2000; G. McPherson and R. Laramore, *An Introduction to Electrical Machines and Transformers*, 2d ed., Wiley, 1990.

Slip rings

Electromechanical components which, in combination with brushes, provide a continuous electrical connection between rotating and stationary conductors. Typical applications of slip rings are in electric rotating machinery, synchros, and gyroscopes. Slip rings are also employed in large assemblies where a number of circuits must be established between a rotating device, such as a radar antenna, and stationary equipment.

Electric rotating machines. Slip rings are used in wound-rotor induction motors, synchronous motors and alternators, and rotary converters to connect the rotor to stationary external circuits. These slip rings are usually constructed of steel with the cylindrical outer surface concentric with the axis of rotation. Insulated mountings insulate the rings from the shaft and from each other. Conducting brushes are arranged about the circumference of the slip rings and held in contact with the surface of the rings by spring tension. A typical assembly is shown in **Fig. 1**. Other arrangements, such as concentric slip rings mounted on the face of an insulating disk, may be employed in special cases. Alternating-current

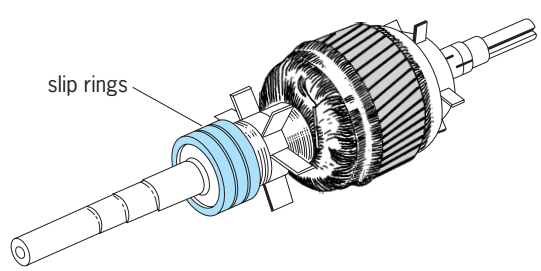

Fig. 1. Rotor of electric rotating machine.

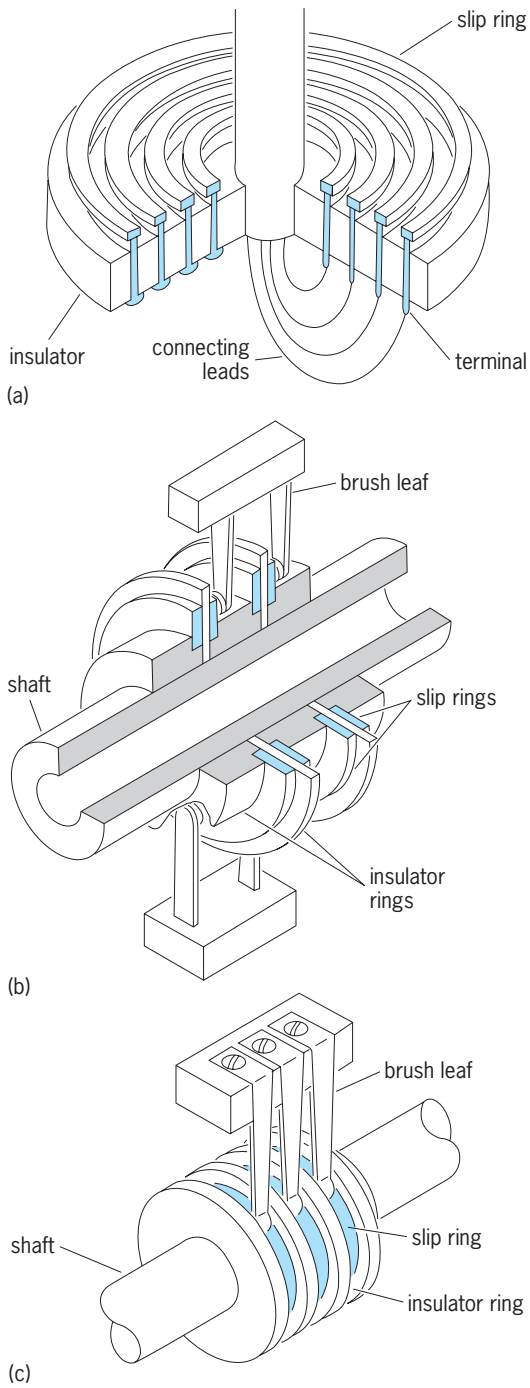

Fig. 2. Slip-ring assembly configurations. (*a*) Concentric-ring. (*b*) Back-to-back. (*c*) Drum.

windings normally require one slip ring per phase except that a single-phase winding requires two slip rings. Two slip rings are required for rotating dc field windings. *See* ELECTRIC ROTATING MACHINERY; GENERATOR; MOTOR.

Arthur B. Eckels

Slip-ring assemblies. These integral mechanical structures contain a plurality of slip rings which, in combination with self-contained brushes, provide continuous electrical connection between electric and electronic equipment mounted on stationary and rotating platforms.

Slip-ring assemblies are designed for a wide range of electric circuits. The same assembly may have circuits for power up to several hundred kilowatts; high voltage to 50 kV; power pulses for radar transmitters, and data signals, including radio frequencies to 100 MHz; strain-gage signals and thermocouple signals. Ring surface speeds range from a few feet per minute to over 15,000 ft/min (75 m/s)

Slip rings for slip-ring assemblies are made usually of coin silver, stamped from hard rolled sheet or cut from drawn tubing, or made of strip silver overlay on copper, formed into a ring, and silver brazed. Fine silver rings may be electroformed onto the insulating material. Surface finish is machined or mill rolled from 4 to 16 microinches (100 to 400 nanometers). Brushes for slip-ring assemblies are graphite combined with copper or silver in proportions suitable for the application, and may be welded or brazed to a spring-temper leaf. Leaf brush pressure is from 1.5 to 3 oz (43 to 85 g).

Large assemblies, such as those used with radar antennas, are fabricated from individual insulators and conducting materials arranged in either the concentric-ring configuration (**Fig. 2***a*) or the back-to-back ring configuration (Fig. 2*b*). Small assemblies, such as those used to transmit signals through gimbals or high-speed turbines, may be fabricated in the back-to-back ring configuration or the drum configuration (Fig. 2*c*).

Other manufacturing methods employ casting the individual rings into filled epoxy resins, electroforming fine silver into grooves machined on filled epoxy resin tubes, or molding the individual rings with electrical grades of thermosetting resins. The casting, electroforming, and molding methods have the advantage of low tolerance buildup, which is important for the synchro sizes.

Background noise for strain-gage signals should not be greater than a few microvolts; intercircuit interference (crosstalk) should not be greater than 70 dB down at 30 MHz, insertion loss at 30 MHz no greater than 0.5 dB, and brush contact resistance approximately 0.005 ohm.

Synchro slip rings are made from fine silver or gold alloy. Brushes are also made from precious-metal alloys, usually in the form of hard-drawn, spring-temper wires. Surface speeds are usually low.

Willard F. Mason

Sloan Digital Sky Survey

A large-scale imaging and redshift survey which, when complete, will map in detail over one quarter of the sky, determining the positions and brightness of over 100 million celestial objects and the redshifts of over a million galaxies and quasars. A primary objective of the survey is to obtain a map of the three-dimensional distribution of galaxies in the universe

Many decades of observational and theoretical research have converged on a unified picture of the nature of the universe. Instrumental in these

advances have been detailed measurements of the expansion rate of the universe by scientists using the Hubble Space Telescope, of the radiation from the cosmic microwave background radiation by the Wilkinson Microwave Anisotropy Probe, and of the distribution of galaxies by several redshift surveys, of which the largest to date have been the Two Degree Field Survey and the Sloan Digital Sky Survey (SDSS). *See* HUBBLE SPACE TELESCOPE; WILKINSON MICROWAVE ANISOTROPY PROBE.

These last projects probe the universe by measuring the redshifts of galaxies, that is, the shift in wavelength of spectral features from their known rest wavelengths to longer wavelengths. Redshifts are caused by the expansion of the universe. Because of this expansion, the farther away from us a galaxy lies the more rapidly it is receding from us. This relation can be turned around, for once the proportionality between distance and redshift (the Hubble "constant" or parameter) has been measured, it can be used together with a galaxy's redshift to give its distance. Thus observations of the positions on the sky and of the redshifts of a sample of galaxies can be used to measure the three-dimensional distribution of galaxies. *See* HUBBLE CONSTANT; REDSHIFT.

Three-dimensional galaxy distribution. Why is this of interest? Galaxies turn out not to be randomly distributed in space, but their distribution is structured in gravitationally bound clusters, sheets, and filaments. The first three-dimensional maps of significant volumes of the universe were made in the 1980s, and incorporated observations of 1000–2000 galaxies. These maps revealed a frothy, filamentary distribution of galaxies, with the largest feature being a "Great Wall" (in a map made by Margaret Geller, John Huchra, and Valerie de Lapparent) of galaxies more than 600 million light years across and extending to the boundaries of the survey volume, so that it could have been even larger.

Encoded in these maps is the result of the initial mass fluctuations in the early stages of the universe and subsequent evolution due to the expansion of the universe, galaxy formation, and gravity: Material falls into dense regions of space, which thereby accrete more mass as the universe evolves. The exciting results discussed above pointed out the need for surveys of much larger volumes of space, so that a representative sample of the universe could be studied. Current theories of the universe say that, in a statistical sense, it is uniform at all locations and in all directions. The requirement on a redshift survey that can fairly be said to sample the universe is, then, that the volume is large enough to contain many tens of examples of the largest possible structures. It was clear from the early results that to do this would involve measuring redshifts for many hundreds of thousands of galaxies. Galaxies are faint, and accumulating enough light to measure a single redshift can take of order an hour of telescope time. The possibility of conducting surveys large enough to map the local universe in the necessary detail developed with the advent of technology that could measure thousands of redshifts in a single night, and of fast, powerful computers to acquire and reduce the data automatically.

Observing strategy. The Sloan Digital Sky Survey is one of several large surveys recently finished or under way. It is unique in several important ways: It is the largest redshift survey, with more than 400,000 galaxy redshifts measured as of June 2004 and plans to measure a million; and it acquires the imaging data from which the positions of the galaxies can be measured during the same time period as it measures the redshifts.

The reason for this strategy is that the Sloan Digital Sky Survey acquires a much better image of the sky than has previously been available for more than very small areas of sky, so it is both an imaging and a redshift survey. This improved imaging is necessary to understand better the intrinsic properties of each observed galaxy (brightness, dominant stellar population, and so forth). The Sloan Digital Sky Survey images the sky through five filters essentially simultaneously, measuring five brightnesses for each object. The filters cover the ultraviolet, green, red, far-red, and near-infrared regions of the optical spectrum. A star that is very hot, and therefore blue, will be much brighter in the ultraviolet filter than in the far-red filter, and vice versa for a cool, red star. Thus objects can be sorted and selected by their colors automatically. The filters are designed to separate quasars from stars by color, for the Sloan Digital Sky Survey is also a redshift survey of 100,000 quasars. Galaxies can be separated from stars and quasars because they appear to be fuzzy rather than pointlike on images (see **illustration**). The data are acquired by electronic imagers called charge-coupled devices (CCDs), similar to those used in digital cameras. The word digital applies both to cameras and the Sloan Survey because the imager converts light to electrons, which can be directly measured and counted as integers by computers. *See* ASTRONOMICAL IMAGING; CHARGE-COUPLED DEVICES; QUASAR.

The normal way in which redshift surveys are done is to image the sky (and before Sloan Digital Sky Survey, all large-area imaging of the sky had been done by photographic plates), select the objects for which redshifts will be measured, and then measure the redshifts. To do this with the Sloan Digital Sky Survey would have been prohibitively time-consuming and expensive. All optical observing needs clear weather, but imaging needs more: pristine clarity that does not change over the night, and a steady atmosphere, so that the images of stars are as close to pointlike as possible. Even the best sites have such conditions only a small fraction of the time. They do, however, have good atmospheric conditions much more frequently. Such conditions are perfectly adequate for obtaining the spectra that give redshifts. Accordingly, the Sloan Digital Sky Survey follows a dual observing strategy: (1) imaging is done in the best weather; (2) objects are selected from the imaging according to carefully defined criteria by powerful computer software; and their spectra are observed during the less-than-perfect weather. In this way the Sloan Digital Sky Survey efficiently uses all the available

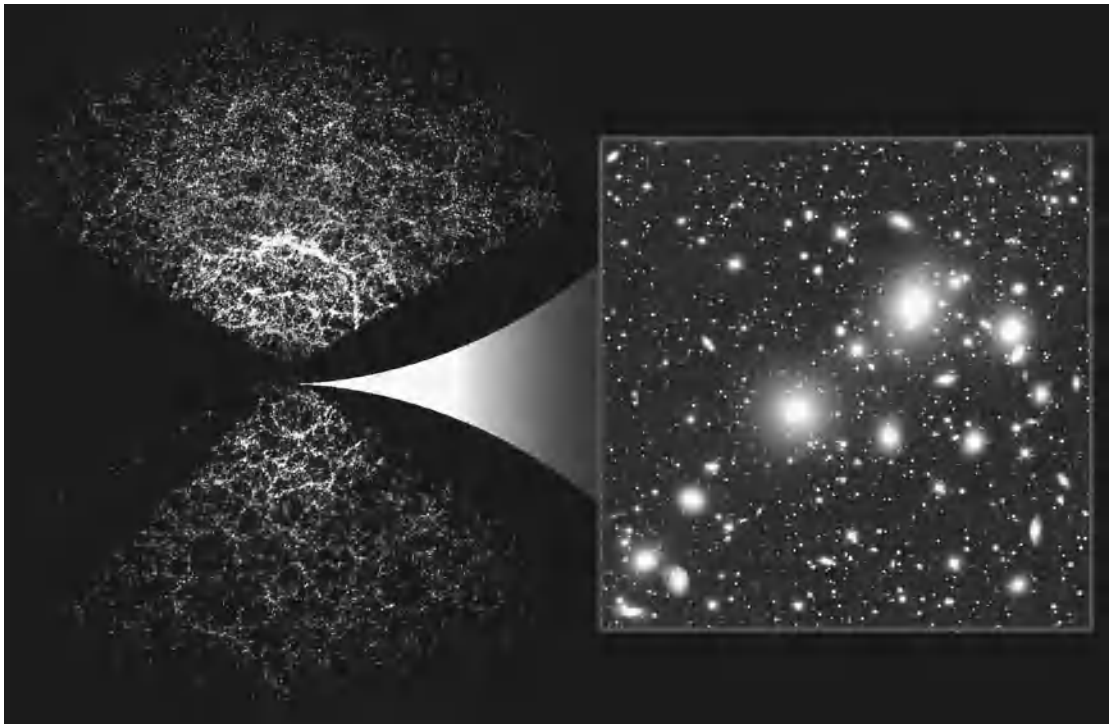

Distribution of galaxies from the Sloan Digital Sky Survey (SDSS). The image on the left shows the redshift-space distribution for galaxies in the SDSS. The dark cones to the left and right are areas of the universe which cannot be observed because they lie behind the obscuring dust of the Milky Way Galaxy. Each point represents one galaxy. The image on the right is a small piece of sky constructed from the SDSS multicolor imaging data. The extended fuzzy objects are galaxies, and the small pointlike objects are stars in the Milky Way Galaxy. (*Courtesy of M. Tegmark and R. Lupton*)

observing time and can be completed in a reasonable time (5–7 years). The spectra are observed with devices that can obtain 640 spectra simultaneously; 7 or 8 sets, or several thousand spectra, can be observed in a single night. To do this, the Sloan Digital Sky Survey uses a dedicated 2.5-m (98-in.) telescope at the Apache Point Observatory, New Mexico. The telescope is equipped with a multi-charge-coupled-device camera and two fiber-fed spectrographs, and observes every possible clear, dark night year-round except for a short maintenance period in the summer. *See* SPECTROGRAPH; TELESCOPE.

Results. The Sloan Digital Sky Survey has obtained several fundamental results. A "Great Wall" of galaxies, 1.4×10^9 light-years long and 80% longer than the "Great Wall" found by Geller, Huchra, and Lapparent, is the largest structure in the universe observed so far. The distribution of galaxies is in excellent agreement with the current cosmological model, in which only 4% of the universe is normal matter, with the rest being dark matter and dark energy. The signature of the formation of galaxies in lumps of dark matter is seen, and comparison of the distribution of the galaxies with that of the cosmic microwave background confirms the presence of dark energy. Cosmology with the Sloan Digital Sky Survey is only just beginning to yield significant information, and many more results, measuring the universe in great detail and with great accuracy, will be forthcoming in the next few years. The current map of the universe, together with a Sloan Digital Sky Survey image of a cluster of nearby galaxies, is shown in the illustration.

The Sloan Digital Sky Survey quasar work has been enormously successful. Quasars are extremely high luminosity objects, thought to be due to the infall of gas into massive black holes. Because they are so luminous, quasars can be seen to great distances, and we observe distant objects as they were a long time ago because of the light travel time. The most distant quasars are seen when the universe was about 5% of its present age. The distant quasars have allowed the discovery of the Gunn-Peterson effect, predicted in the 1960s, whereby the light from distant objects is absorbed by hydrogen if that hydrogen is atomic. This would prevent our seeing the distant universe unless the intervening hydrogen were ionized, that is, the electrons separated from the protons, in which case the atoms could no longer absorb light. The detection of the Gunn-Peterson effect therefore locates the epoch at which the dark ages of the universe ended and the hydrogen was ionized by light from newly formed stars.

A survey as powerful as the Sloan Digital Sky Survey produces results in many other areas. The Sloan Digital Sky Survey has been instrumental in discovering that the Milky Way Galaxy is surrounded by streams of stars from the tidal destruction of small satellite galaxies that used to orbit the Galaxy. The Sloan Digital Sky Survey has discovered many gravitational lenses, in which light from distant quasars is bent by the warping of space by lumps of mass (the lenses) between the Earth and the quasar. This effect is predicted by Einstein's theory of general relativity. The Sloan Digital Sky Survey has discovered the

most massive lens yet found. All masses will distort the light passing them, and this effect has allowed the Sloan Digital Sky Survey to measure the masses of galaxies directly; the more massive the galaxy, the more the light path is bent. The Sloan Digital Sky Survey has discovered the lowest-luminosity galaxy ever found, a satellite of the nearby Andromeda Galaxy. It has discovered large numbers of brown dwarfs, objects intermediate in mass between planets and stars. It has produced the largest sample of small bodies in the solar system with accurate colors, allowing their compositions to be deduced. It can confidently expected that these findings are only the beginning of decades of discovery from the Sloan Digital Sky Survey data. *See* BROWN DWARF; GRAVITATIONAL LENS; MILKY WAY GALAXY.

The data from the Sloan Digital Sky Survey are made freely available to all, with a regular schedule of data releases. As well as being available to professional astronomers in support of their research, they are available for educational purposes and come with a large set of educational tools. *See* COSMOLOGY; GALAXY, EXTERNAL; UNIVERSE. Gillian R. Knapp

Bibliography. B. S. Ryden, *Introduction to Cosmology*, Addison-Wesley, 2003; M. A. Strauss, Reading the blueprints of creation, *Sci. Amer.*, 290(2):42, February 2004; M. Tegmark et al., The three-dimensional power spectrum of galaxies from the Sloan Digital Sky Survey, *Astrophys. J.*, 606:702–740, 2004; D. G. York et al., The Sloan Digital Sky Survey: Technical Summary, *Astron. J.*, 120:1579–1587, 2000.

Slope

The trigonometric tangent of the angle α that a line makes with the x axis as in **Fig. 1**. In **Fig. 2** the slope of a plane curve C at a point P of C is the slope of the line that is tangent to C at P. If $y = f(x)$ is an equation in rectangular coordinates of curve C, the slope of C at $P(x_0, y_0)$ is the value of the derivative $dy/dx = f'(x)$ at P, denoted by $f'(x_0)$, and hence an equation of the nonvertical tangent to C at P is $y - y_0 = f'(x_0)(x - x_0)$. *See* ANALYTIC GEOMETRY; CALCULUS. Leonard M. Blumenthal

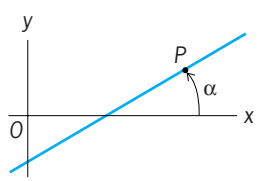

Fig. 1. Slope of a line.

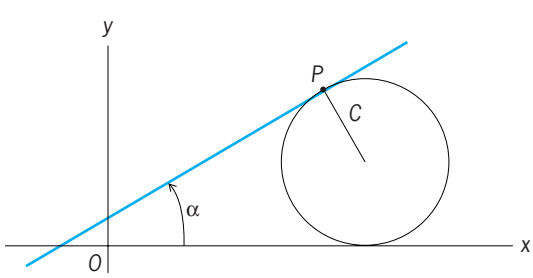

Fig. 2. Slope of a curve.

Sloth

A mammal classified in the order Xenarthra along with anteaters, tamanduas, and armadillos. The sloth differs from all other mammals by having additional articulations (xenarthrales) between their lumbar vertebrae (called xenarthrous vertebrae). They also have a double rather than a single inferior vena cava, the vein that returns blood from the posterior portion of the body to the heart. The dental formula is I 0/0, C 0/0, PM 0/0, M 5/4 × 2, for a total of 18 teeth which grow throughout life. Females have a primitive, divided uterus and a common urinary and genital tract; males have internal testes. *See* DENTITION.

Two extant families of tree sloths (Bradypodidae and Megalonychidae) inhabit the lowland and upland tropical forests of Central and South America.

Bradypodidae. The family Bradypodidae includes a single genus, the three-toed tree sloth (*Bradypus*) [see **illustration**]. This is a slender-bodied animal with a head and body length of 413–700 mm (16–27 in.) and a tail length of 20–90 mm (0.75–3.5 in.) that looks heavier because of its long, thick, coarse hair. The three-toed sloth weighs 2.25–6.20 kg (5–13.5 lb). It has three toes on each front and rear foot. The head is small and round, the tail is short and stumpy, and the eyes and ears are reduced—the animal can neither hear nor see well but depends on its senses of smell and touch. The three species in this genus have 30 ribs and long necks containing nine vertebrae (two more than most mammals). These sloths are drab brown or gray with patches of white. Males have a bright yellow spot, about the size of a silver dollar, on their back. In this spot, the hair is short and glossy.

Megalonychidae. The family Megalonychidae (two-toed tree sloths) contains a single genus (*Choloepus*) with two species. Two-toed tree sloths, or unau, with two toes on their front feet and three on their hind feet, are somewhat larger and more active than their three-toed cousins. Adults range 540–740 mm (21–28 in.) in total length and weigh 4.0–8.5 kg (9–18.5 lb). The coloration is grayish brown with a paler face. The large hooklike claws are 8–10 cm (3–4 in.) long. These sloths possess 48 ribs, six to eight vertebrae, and their tail is either absent or vestigial.

Physiology and ecology. Tree sloths are solitary and spend most of their lives hanging upside down from the upper branches of a tree. They hang by the long, hooked claws at the ends of their toes and slowly move along, hand over hand in the inverted position. They spend their waking hours slowly picking and eating leaves, especially the leaves of the Cecropia

Three-toed tree sloth (*Bradypus tridactylus*). (*Photo by Dr. Lloyd Glenn Ingles; © California Academy of Sciences*)

tree. Their stomachs contain cellulose-digesting bacteria and may retain food for as long as a month before passing it on to the small intestine. Sloths are so lethargic that two blue-green algae, which look like moss, actually grow in the grooves of their coarse hair. In the rainy season, their fur often has a distinct greenish tinge due to the luxuriant growth of algae; in times of drought, the fur turns yellowish. The greenish color serves to camouflage the animals in the tree canopy.

Sloths have a low body temperature of 30–34°C (86–93°F) which fluctuates both daily and seasonally. They regulate their temperature by shifting their position in and out of the sun. Sloths make no attempt to construct a den or shelter in the trees. They sleep and give birth while hanging upside down. They normally sleep about 18 hours a day on a vertical branch, keeping a firm hold with all four feet with their head hanging down and resting between the front limbs. Predators include jaguars and eagles, although habitat destruction and excessive hunting may prove to be the most detrimental factors affecting these species. One species (*B. torquatus*) is classified as endangered by the International Union for Conservation of Nature and Natural Resources (IUCN) and the U.S. Department of the Interior.

Reproduction and development. Most sloths are believed to breed throughout the year. A single young is born after a gestation ranging from 3.5 to 11.5 months. During its first few months, the young sloth clings flat to the mother's belly. They cease nursing between 4 and 5 weeks of age. Female two-toed sloths reach sexual maturity at 3 years of age, whereas males are sexually mature at 4–5 years of age. Life expectancy in the wild is up to about 12 years, although they have lived up to 31 years in captivity. Donald W. Linzey

Bibliography. D. Macdonald (ed.), *The Encyclopedia of Mammals*. Andromeda Oxford Limited, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Slow neutron spectroscopy

The use of beams of slow neutrons, from nuclear reactors or nuclear accelerators, in studies of the structure or structural dynamics of solid, liquid, or gaseous matter. Studies of the chemical or magnetic structure of substances are usually referred to under the term neutron diffraction, while studies of atomic and magnetic dynamics go under the terms slow neutron spectroscopy, inelastic neutron scattering, or simply neutron spectroscopy. The results obtained are to a considerable extent complementary to those obtained by use of optical spectroscopy and x-ray diffraction. *See* NEUTRON DIFFRACTION; NUCLEAR REACTOR; PARTICLE ACCELERATOR; SPECTROSCOPY; X-RAY DIFFRACTION.

Experiments. In a neutron spectroscopy experiment, a beam of neutrons is scattered by a specimen and the scattered neutrons are detected at various angles to the initial beam. From these measurements, the linear momenta of the incoming and outgoing neutrons (and the vector momentum changes experienced by individual neutrons) can be computed. Experiments are often carried out in a mode in which the vector momentum change $\mathbf{Q}$ is held constant and the energy transfer ϵ is employed as the independent variable for the experiment.

In general, just those neutrons which have been scattered once only by the specimen are useful for analysis; the specimen must be "thin" with respect to neutron scattering power as well as to neutron absorption. In practice, the experiments are usually intensity-limited, since even the most powerful reactors or accelerators are sources of weak luminosity when, as here, individual slow neutrons are to be considered as quanta of radiation.

Neutron spectroscopy requires slow neutrons, with energies of the order of neutrons in equilibrium with matter at room temperature, or approximately 0.025 eV. The corresponding de Broglie wavelengths are approximately 0.2 nanometer, of the order of interatomic spacings in solids or liquids. The fast neutrons emitted in nuclear or slow fission reactions can be slowed down to thermal velocities in matter which is transparent to neutrons and which contains light elements, such as hydrogen, carbon, and beryllium, by a process of diffusion and elastic (billiard-ball) scattering known as neutron moderation. By selection of those diffusing neutrons which travel in a certain restricted range of directions

(collimation), a beam of thermal and near-thermal neutrons can be obtained. The beam of slow neutrons so produced can be selected as to energy by means of their velocities, by using an electronic or mechanical velocity selector, or by means of their wavelengths, through diffraction from a specially selected and treated single-crystal monochromator. As theoretically expected, the two methods are found to be physically equivalent, but have important differences for the strategy of experiments. *See* NEUTRON; QUANTUM MECHANICS; THERMAL NEUTRONS.

Interpretation. The bulk of the observations can be accounted for in terms of scattering of semiclassical neutron waves by massive, moving-point scatterers in the forms of atomic nuclei and their bound electron clouds. The spatial structure of the scatterers, time-averaged over the duration of the experiment, gives rise to the elastic scattering from the specimen that is studied in neutron diffraction; the spatial motions of the scatterers give rise to the Doppler-shifted inelastic scattering involved in slow neutron spectroscopy.

Quasiparticles. Analysis of the scattering patterns is particularly straightforward when the neutrons can be envisaged as creating or annihilating quantum excitations called quasiparticles in the course of the scattering event. The single-quantum component then includes peaks in the scattering pattern—either sharp peaks, whose energy width reflects the energy resolution of the experiment and instrument, or broadened peaks, whose energy width is interpreted in terms of natural lifetimes for the excitations and of interactions between them. Each peak then yields, through the assumption of conservation of energy and momentum between neutron and quasiparticle, an energy ϵ and a momentum $\mathbf{Q}$ for the (created or annihilated) single quasiparticle involved. Repetition under different conditions permits deduction of the dispersion relation, $\epsilon_q = \epsilon_q(\mathbf{Q})$, of the quasiparticle concerned. Consideration of the intensity of the neutron line allows assignment of the polarization and symmetry type of the excitation concerned—for example, phonon or magnon, longitudinal or transverse, acoustic or optical. *See* DISPERSION RELATIONS; LATTICE VIBRATIONS; LIQUID HELIUM; MAGNON; PHONON.

Likewise, through analyses of their widths and shapes, the broadened peaks yield estimates of the lifetimes $\tau_q(\mathbf{Q})$ of the quasiparticle and information on its modes of interaction in the specimen. Inverse lifetimes ($1/\tau$) or frequencies (ϵ/h, where h is Planck's constant) in the range from 10^{10} to 10^{13} Hz are accessible with this method.

Diffusion phenomena. Broadened peaks may occur at zero energy transfer. This quasielastic scattering is usually the result of slow diffusionlike motions of scattering entities through the specimen. What is unusual is that distances involved are of atomic dimensions and times correspondingly short. Thus it is possible to observe self-diffusion in classical liquids on an atomic scale. The phenomena are best considered through the correlation function approach discussed below. *See* DIFFUSION.

Continuous spectra. For continuous spectra, interpretation is more difficult and indirect. Always evocative in principle (and at times useful in practice) are the Van Hove space-time correlation functions $G(\mathbf{r},t)$. These are natural extensions to space-time of the Patterson pair correlation functions of crystallography and the Zernike-Prins pair distribution functions in liquids. For a monatomic liquid, the self-correlation function $G_s(\mathbf{r},t)$ has a mathematical definition which can be approximately expressed as follows: Given an atom at position $\mathbf{r} = 0$ at time $t = 0$, then $G_s(\mathbf{r},t)$ is the probability that the same atom is at position $\mathbf{r}$ at time t. Similarly, the pair correlation function $G_p(\mathbf{r},t)$ of a monatomic liquid can be approximately defined by the statement: Given an atom at position $\mathbf{r} = 0$ at time $t = 0$, then $G_p(\mathbf{r},t)$ is the number density of other atoms at position $\mathbf{r}$ at time t. In simple cases of continuous spectra, these functions provide a conceptually simple and powerful method of direct analysis, similar to the role played by quasiparticles in the study of line spectra. This is also the case for magnetic inelastic scattering when there is only one species of magnetic atom involved.

When more than one type (j) of atom is involved, the scattering pattern is related to a weighted superposition of correlation functions (G) between pair types (j,j'). In this case, isotopic substitution (for example, deuterium for hydrogen, or ^{60}Ni or ^{62}Ni for natural nickel) can assist in sorting out the contributions. (The weights are proportional to the products of the neutron scattering amplitudes of the pair types.) When the scattering is magnetic, the three spatial components of the vector magnetic moments of the atoms can also be involved through their various self and pairwise correlations.

Applications. Just as (slow) neutron diffraction is the most powerful available scientific tool for study of the magnetic structure of matter on an atomic scale, so slow neutron spectroscopy is the most powerful tool for study of the atomic magnetic and nuclear dynamics of matter in all its phases. The direct nature of the analysis has in some cases added considerable support to the conceptual structure of solid-state and liquid-state physics and thus to the confidence with which the physics is applied. For example, neutron spectroscopy has confirmed the existence of phonons, magnons, and the quasiparticles (rotons) of liquid helium II. Detailed information has been obtained on the lattice vibrations of most of the crystalline elements and numerous simple compounds, on the atomic dynamics of many simple liquids, on the dynamics of liquid helium in different phases, and on the atomic magnetic dynamics of a great variety of ferromagnetic, ferrimagnetic, antiferromagnetic, and modulated magnetic substances.

Bertram N. Brockhouse

Bibliography. E. Balcar and S. W. Lovesey, *Theory of Magnetic Neutron and Photon Scattering*, 1989; H. Glyde, *Excitations in Liquid and Solid Helium*, 1995; J. S. Higgins and H. C. Benoit, *Polymers and Neutron Scattering*, 1994, reprint 1996; S. W. Lovesey, *Theory of Neutron Scattering from Condensed Matter*, 2 vols., 1986.

Slug (zoology)

A terrestrial pulmonate mollusk in which the shell is absent or reduced to a small internal or external rudiment. The slug form has evolved independently several times. The incorporation into the muscular foot region of the body organs (which are contained within the shell in other mollusks) results in a streamlined body shape (see **illus.**), enabling the animal to enter small holes or crevices.

***Limax maximus*, up to 8 in. (20 cm) in length. There are two pairs of tentacles on the head, and the opening to the lung is visible.**

The skin is soft and moist, and readily loses water by evaporation or gains it by absorption. Maximum water loss occurs when crawling, due to laying down the slime trail. Water loss is minimized by behavioral mechanisms; thus most slugs are abundant on the surface only on windless wet nights.

Lung respiration occurs as in other pulmonates, but skin respiration is probably at least as important.

Slugs are simultaneous hermaphrodites. Copulation follows an often elaborate courtship, and is frequently reciprocal. In many species, sperm are packaged in a spermatophore. Cross-fertilization is normal, but self-fertilization is possible and may indeed be common. Eggs are laid in moist crevices as they have little resistance to desiccation.

A few slugs are carnivorous, such as *Testacella* which eats earthworms, but the majority are catholic herbivores and may become serious horticultural and agricultural pests. *See* MOLLUSCA; PULMONATA.

Norman Runham

Smallpox

An acute infectious viral disease characterized by severe systemic involvement and a single crop of skin lesions that proceeds through macular, papular, vesicular, and pustular stages. Smallpox is caused by variola virus, a brick-shaped, deoxyribonucleic acid–containing member of the Poxviridae family. Strains of variola virus are indistinguishable antigenically, but have differed in the clinical severity of the disease caused. From at least 1157 B.C., when the disease probably killed the Egyptian Pharaoh Ramses V, hundreds of millions of cases and millions of deaths due to smallpox occurred throughout the world. On October 26, 1977, Ali Maow Maalin, a cook from Merca, Somalia, had onset of the smallpox rash; he was the last patient to contract the disease through natural transmission, although a laboratory outbreak of smallpox occurred in the United Kingdom in 1978. Following a 13-year worldwide campaign coordinated by the World Health Organization (WHO), smallpox was declared eradicated by the World Health Assembly in May 1980. Smallpox is the first human disease to be eradicated.

Epidemiology. Humans were the only reservoir and vector of smallpox. The disease was spread by transfer of the virus in respiratory droplets during face-to-face contact. Before vaccination, persons of all ages were susceptible. As the eradication program progressed, attack rates were highest in young children and older adults. Children had low vaccination coverage due to high birth rates and poor health infrastructure in developing countries; older persons, often exposed to smallpox while caring for sick family members, lost immunity if not revaccinated since youth. In 1967, when the intensified global eradication program began, 33 nations reported endemic disease, and 14 other countries reported importations of cases. Over 131,000 cases were recorded by WHO in 1967, but special surveys indicated that less than 1% of the cases were reported. *See* VACCINATION.

It is estimated that 10–15 million cases and 1–3 million deaths due to smallpox occurred yearly in the mid-1960s. All countries with endemic smallpox were in the tropics; in 1967, about 65% of the recorded cases were from India, while 78% of the smallpox-endemic countries were in Africa.

Smallpox was a winter-spring disease; there was a peak incidence in the drier spring months in the Southern Hemisphere and in the winter months in temperate climates. The spread of smallpox was relatively slow. The incubation period was an average of 10–12 days, with a range of 7–17 days. Fifteen to forty percent of susceptible persons in close contact with an infected individual developed the disease.

Manifestations. Smallpox was once one of the most feared diseases because it caused widespread skin lesions, often leaving disfiguring scars on the face. Lesions began as reddened or dark spots (macules) and evolved through papules, vesicles, and pustules over 7–14 days. The lesions were round, approximately 0.5 cm in diameter, and firmly fixed to the basal layers of the skin. Over the next 7–14 days, the pustules would dry and form crusts that would fall off, leaving pitted areas with reduced pigmentation, particularly over skin areas with a heavy concentration of sebaceous glands. In time, these areas would flatten and develop excess pigmentation. The eruption was distributed peripherally, differing from the rash of chickenpox (varicella), which has a typical distribution over the torso but was often confused with smallpox.

There were two main clinically distinct forms of smallpox, variola major and variola minor. Variola major, prominent in Asia and west Africa, was the more severe form, with widespread lesions and case fatality rates of 15–25% in unvaccinated persons, exceeding 40% in children under 1 year. From the early 1960s to 1977, variola minor was prevalent in

South America and south and east Africa; manifestations were milder, with a case fatality rate of less than 1%.

Treatment. There is no specific treatment for the diseases caused by poxviruses. Supportive care for smallpox often included the systemic use of penicillins to minimize secondary bacterial infection of the skin. When lesions occurred on the cornea, an antiviral agent (idoxuridine) was advised.

Prevention. Although inoculation with material from cowpox lesions was known in England in the mid-1700s, Edward Jenner, a British general medical practitioner who used cowpox to prevent smallpox in 1796, is credited with the discovery of smallpox vaccine (vaccinia virus). Over the next two centuries, major efforts were made to keep the fragile vaccine potent and stable in warm climates. These problems were overcome with the application of the freeze-dry technique to the manufacture of smallpox vaccines in the 1960s and with the strict quality control by WHO of vaccine used in the field. However, the intensified global smallpox eradication program did not rely only on vaccination. Although the strategy for eradication first followed a mass vaccination approach, experience in the west and central African campaign during 1968 and 1969 showed that intensive efforts to identify areas of epidemiologic importance, to detect outbreaks and cases, and to contain them would have the greatest effect on interrupting transmission.

Maintenance of eradication. Escape of variola virus from laboratories in the United Kingdom resulted in three cases of smallpox in 1973 and two cases in 1978, but these outbreaks were not due to natural transmission. WHO has since established requirements to assure strict containment of variola virus in laboratories.

In 1978, WHO established an International Commission to confirm the absence of smallpox worldwide. The 19 recommendations made by the commission included reducing the number of laboratories retaining variola virus, pursuing research on selected orthopoxviruses (such as monkeypox, an animal virus that causes a smallpoxlike eruption but only rarely infects humans and has little potential for interhuman spread), and abandoning routine vaccination except for laboratory workers at special risk. By 1990, only the Soviet Union and the United States were maintaining variola virus, which is kept under high security. Scientists and health officials in these countries are considering destroying the last remaining strains of variola virus after the viral genome has been completely mapped. *See* ANIMAL VIRUS.

Joel Breman

Bibliography. J. G. Breman and I. Arita, The confirmation and maintenance of smallpox eradication, *N. Engl. J. Med.*, 303:1263–1273, 1980; A. Deria et al., The world's last endemic case of smallpox: Surveillance and containment measures, *Bull. WHO*, 58:279–283, 1980; F. Fenner et al., *Smallpox and Its Eradication*, World Health Organization, 1988; D. A. Henderson, The eradication of smallpox, *Sci. Amer.*, 235:25–33, 1976.

Smart card

A plastic card, the size of a credit card, that contains an embedded silicon computer chip. Its primary purpose is the portable storage and retrieval of data used to authorize various types of electronic transactions. Compared to their predecessors, magnetic stripe cards (for example, credit cards), smart cards can store data relating to more than one institution, facilitate more than one type of use, carry substantially larger volumes of data, process transactions at higher rates, and provide logical and physical security to the card's data. Smart cards work by interacting with a card reader—an interface between the smart card and an institution.

History. Smart cards were developed in the late 1960s to 1970s, when inventors in France, Germany, and Japan formed the idea of cards holding microchips (**Table 1**). Throughout the 1980s, smart card implementation was mainly confined to Europe and Japan. The United States favored wired telecommunication, which supported on-line cardholder/transaction fraud detection, reducing the need to adopt stringent forms of identify verification associated with batch transaction processing. As a result, smart card developments were limited to trials by a few banks, the Department of Defense, and a few universities. Since the September 11, 2001, terrorist attacks, there has been a race to implement stringent identification procedures involving smart cards, primarily in high-liability operating environments.

Applications. Smart cards support a wide range of transactions based on their ability to identify and store information about the cardholder, card issuer, authorized accounts and their activities, authorized merchants, and authorized transactions. The validation of a cardholder's identity may be knowledge-based, where the individual knows something, such as a personal identification number (PIN). It may also be possession-based, where the presence of a "token" (for example, an ID card or a bank card) validates the person's identity. In addition, smart cards

TABLE 1. Invention and deployment of the smart card

Year	Event
1968	Jorgen Dethloff and Helmut Grotrupp (Germany) patent the idea of combining computer chips with plastic cards.
1970	Kunitaka Arimura (Japan) receives Japanese patent related to the IC card.
1974	Roland Moreno (France), journalist, receives patent in France for chip cards. Considered father of smart cards, he founded Innovatron, which issued first global licenses.
1976	Honeywell Bull (France) receives licenses as a result of DGT initiative.
1979	Flonic Schlumberger (France) and Honeywell Bull & Philips (Netherlands) receive first Innovatron licenses.
1980	GIE sponsors first smart card trials in three French cities.
1982	First USA trials held in North Dakota and New Jersey.
1996	First USA trials held on university campus.

can carry biometric information (for example, fingerprints or iris scans), providing greater security for the identification process.

Some common smart card applications include credit transactions, debit transactions, stored-value transactions, information management applications, loyalty/affinity transactions, and multiple applications.

Credit transactions. Smart cards reduce fraud by securing merchant transactions. Smart credit cards give cardholders financial purchasing flexibility in the form of preapproved cash advances and loans (initiated at the point of sale), which are associated with specific electronic "purses" contained on the card.

Debit transactions. Smart cards provide electronic direct-debit at automated teller machines and at specific points of sale. This improves the cardholder's access to specific goods and services in locations where the cardholder's identity and credit history are unknown.

Stored-value transactions. Institutions can create smart cards with a fixed value initially encoded in the card chip's memory. This stored value may be disposable (debit "read-only") or reloaded (debit/credit "rewrite"). From this sum, purchases can be deducted. These transactions do not require access to the cardholder's personal identity, so no PIN is required. Examples include telephone cards and retail merchant gift cards.

Information management applications. Smart cards can be used for the portable storage and use of cardholder's personal information such as bank credit/debit accounts, insurance information, medical history, emergency contact information, and travel documentation.

Loyalty/affinity transactions. Smart cards can apply vendor incentives (such as points, credits, discounts, or direct delivery of products or services) at the point of sale, which are tied to the purchase of goods or services.

Multiple applications. Smart cards are able to combine two or more applications on a single card (some of which are able to share PIN numbers). At present, this form of combined identity management requires the capabilities of an integrated circuit.

Physical specifications. The size of a smart card is the same as that of a typical magnetic stripe credit, debit, or ID card. The International Standards Organization specification, ISO 7810, established the size of these cards, called ID-1 cards. ISO 7810 specifies that cards be 85.7 mm (3.4 in.) wide, 55.2 mm (2.1 in.) high, and 0.76 mm (0.03 in.) deep (**Table 2**). Smart cards adhere to this standard to promote backward compatibility with the existing transaction processing infrastructure.

TABLE 2. Smart card standards

Standard	Description
ISO 7810	ID cards—physical characteristics
ISO 7811	Identification cards, magnetic stripe, embossing
ISO 7816	Design and use of contact ICC cards
ISO/IEC 10536	Contactless ICC cards
ISO/IEC 14443	Proximity cards
ISO/IEC 15693	Vicinity cards

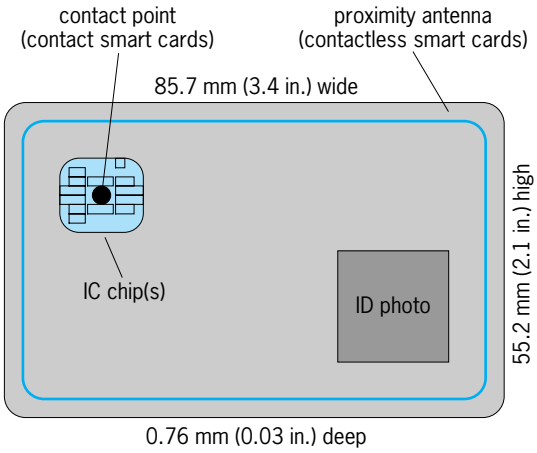

Fig. 1. Smart card showing both the contact point and the proximity antenna.

Contact versus contactless smart cards. Contact cards require that the card remain in physical contact with the reader. ISO 7816 specifies the requirements for contact cards. In general, they have a gold plate on one side of the card, covering contact points between the card's chip and the reader (**Fig. 1**). A contact smart card must be physically inserted into the reader and remain in contact with it during the session. The reader verifies that the card is properly situated and that the power supplies of both card and reader are compatible. It then supplies the card with the power necessary to complete the transaction.

The contactless card (ISO 10536, 14443, and 15693) transmits information to and from the card reader via an antenna embedded in the card (Fig. 1). Contactless smart cards rely on electrical coupling for their power. At present, contactless cards tend to be slower, less reliable, and more expensive than contact cards. However, this technology is rapidly improving and will eventually replace contact smart cards.

Integrated circuits. The computer chips embedded in smart cards can be either memory chips or microprocessors. At present, memory chips have 1–5 kilobits of memory. Less expensive than microprocessors, they depend on the card reader for processing and security. *See* INTEGRATED CIRCUITS; MICROPROCESSOR; SEMICONDUCTOR MEMORIES.

Smart cards with microprocessors are "true" smart cards. In many ways, the microprocessor makes them a miniature computer, complete with input/output ports, an operating system, and a hard drive. Microprocessor chips are available with 8-, 16-, and 32-bit architectures. They contain both short- and long-term memory cells, ranging from a few hundred bytes of read-only memory (ROM) to several megabytes of random-access memory (RAM). Usually, the operating system is loaded in the ROM, while the RAM provides a "scratch pad" for processing information.

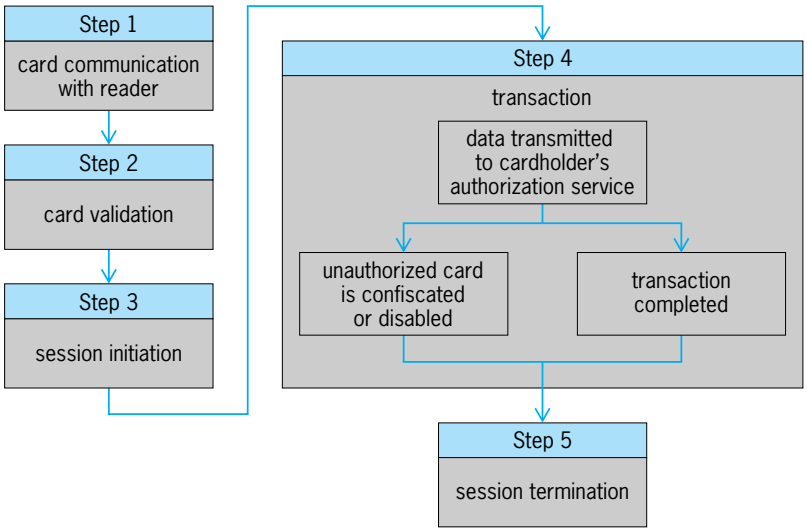

Fig. 2. Typical smart card transaction.

While smart cards are programmed using a variety of languages, the most popular is Java, which can be compiled directly on the card. Current innovations center on the development of minidatabases that can be stored on the card. As a result, smart card applications are increasingly important enablers of identify-based transactions. *See* PROGRAMMING LANGUAGES.

Transaction. A typical smart card transaction session consists of communication, validation, initiation, transaction, and termination (**Fig. 2**).

Communication. Communication with the reader may be direct or in the vicinity of the reader, depending on the type of card (contact or contactless) and the capability of the reader.

Validation. The card must be validated in order to establish a session. If the card is valid, the system captures an identification number to establish an audit trail and processes the transaction using the authorized "value" that is stored on the chip (for example, removing cash from an electronic purse). This activity is handled by a secure access module (SAM) on a chip housed in the reader. The SAM has the electronic keys necessary to establish communication with the reader's operating system.

Initiation. A reset command is issued to establish communication between the card and the reader, and the clock speed is established to control the session. In the case of a both contact-and-contactless smart card, the reader obtains the required data from the card and initiates the requested transaction.

Transaction. Data are exchanged between the card and the reader through the defined contact (input/output) points. A record of the transaction is stored on both card and reader.

Termination. For contact smart cards, the contacts are set to a stable level, the supply voltage is terminated, and the card is ejected from the terminal. For contactless cards, a termination protocol ends the session and resets the reader.

Security. There are two levels of smart card security: physical and data.

Physical security. This involves protecting the smart card against damage from cardholders, damage from the environment, and physical tampering. Bending a card, scratching its contacts, or breaking chip connections will destroy the card. New manufacturing technologies and protective sleeves are being designed to protect cards.

A range of environmental conditions can damage or destroy smart cards. Examples include electric voltage, frequency of use, humidity, light, temperature, and radiation exposure.

Various methods of tampering with smart cards have been suggested or tried. Smart card manufacturers have protected the layout of the IC chip by hiding its data pathways. Cards that show physical traces of tampering can be identified and their use blocked by readers. Manufacturers also use other technologies including holograms, hidden characters, and fingerprints to make counterfeiting unprofitable for criminals.

Data security. Smart card systems are susceptible to attempts to invade systems to steal or destroy data. This may be done by visually capturing the keystrokes of cardholders in order to learn the PIN associated with the card. A more sophisticated threat is infiltration of the system, with the goal of finding and stealing valuable data. In the case of contactless smart cards, data that are wirelessly transmitted between cards and readers potentially may be intercepted.

In response to these threats, smart cards use key-based cryptography for data encryption. There are various methods for this based on the logic of the "one-time pad," a randomly generated single-use ciphertext (encrypted text) that is decrypted by both the reader and the card and required to complete a transaction. *See* CRYPTOGRAPHY.

In a symmetric-key-encrypted transaction, the card accesses a public-key directory (stored into the reader), called a certificate authority (CA), to encrypt the public key with a "hash" (a document digest of

the key). The information used to create this hash includes version number, serial number, signature algorithm, card issuer's name, and so on. Both the card and the CA receive a decrypted copy of the hash. Next, the reader accesses the CA to get the decrypted copy of the hash. This validates and authorizes the identity of both parties, allowing the transaction to take place. The transaction recipient accesses the CA for a final hash, which is encrypted and acts as a response to the card, identifying and authorizing the reader and notifying the card of the close of the transaction.

Asymmetric encryption makes use of both public and private keys to provide additional security for sensitive transactions.

Institutions using smart cards have developed several measures to deal with insecure cards. The most popular among these is "hot listing" of insecure cards. Early on, when only contact smart cards were available, the reader was programmed to retain problematic cards. This proved problematic when readers "ate" damaged but secure cards. Hot listing of contactless cards usually involves notifying the readers of insecure cards and programming the system to refuse transactions. *See* COMPUTER SECURITY.

Katherine M. Shelfer; Kevin Meadows

Bibliography. *Open Smart Card Infrastructure for Europe V 2*, Vol. 6: *Contactless Technology*, pt. 1: *White Paper on Requirements for the Interoperability of Contactless Cards*, eESC TB 8 Contactless Smart Cards, March 2003; K.M.Shelfer et al., Smart cards, *Advances in Computer Science*, vol. 60, 2003.

Smith chart

A graphical procedure for converting impedances into reflection coefficients and vice versa. At high-frequency circuit operation, voltages and currents behave like traveling waves propagating over finite-length components. Copper traces on printed circuit boards, coaxial cables, and even simple wires become transmission lines. Changes in the length or operating frequency of these transmission lines result in periodic impedance behaviors not encountered in low-frequency circuits. Frequently, the impedance is replaced by the reflection coefficient, a more convenient way to quantify the transmitted and reflected voltage-current waves. To show how the impedance can be converted into a reflection coefficient and vice versa, P. H. Smith developed an ingenious graphical procedure based on conformal mapping principles. His approach permits an easy and intuitive display of the reflection coefficient as well as the complex impedance in a single graph. Although such a graphical procedure, nowadays known as the Smith chart, was developed in the 1930s prior to the computer age, it has retained its popularity and can be found in every data book describing passive and active high-frequency components and systems. Almost all computer-aided design programs utilize the Smith chart for the analysis of circuit impedances, design of matching networks,

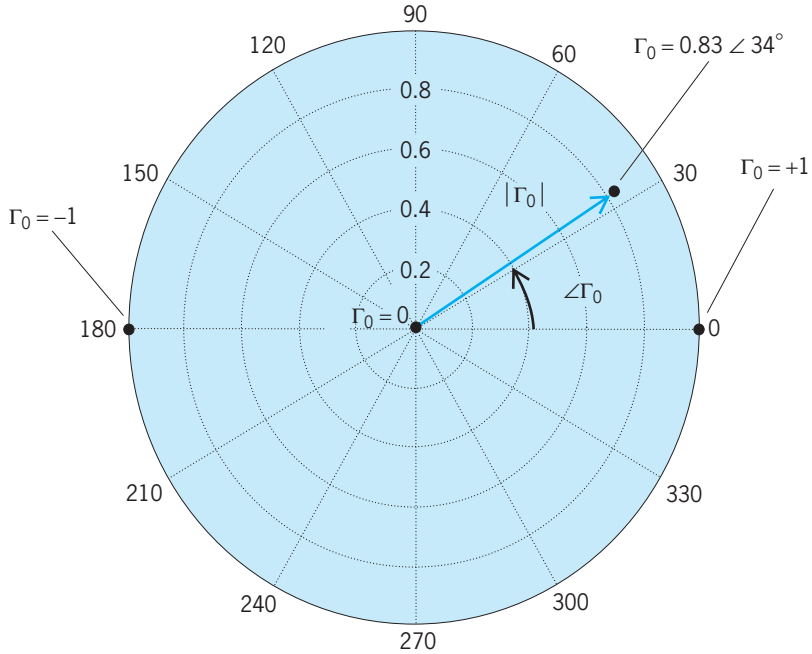

Fig. 1. The reflection coefficient Γ_0 as a complex phasor in the Γ-plane. Open- and short-circuit conditions are located at the right and left sides, while impedance matching to the characteristic impedance is achieved at the center.

and computations of noise figures, gain, and stability circles. Even instruments such as the ubiquitous network analyzer have the option to represent certain measurement data in a Smith chart format. *See* CONFORMAL MAPPING; ELECTRICAL IMPEDANCE; REFLECTION AND TRANSMISSION COEFFICIENTS; TRANSMISSION LINES.

Reflection coefficient. At its core the Smith chart displays the complex reflection coefficient in magnitude and phase. If we consider the end of a transmission line, we can define a reflection coefficient Γ_0 in the form given by the equation below, where

$$\Gamma_0 = |\Gamma_0|\angle\Gamma_0 = \frac{Z_L - Z_0}{Z_L + Z_0}$$

Z_L is a generic complex termination impedance and Z_0 is the reference impedance, typically the characteristic impedance of the transmission line. **Figure 1** depicts the reflection coefficient as a complex vector in polar form with magnitude and phase. In this example a load impedance of $Z_L = (50 + j150)\ \Omega$ is attached to a $Z_0 = 50\ \Omega$ transmission line. The reflection coefficient $\Gamma_0 = 0.83\angle 34°$ can then be displayed as a phasor in the complex plane. It is interesting to note that open-circuit ($Z_L \to \infty$) and short-circuit ($Z_L \to 0$) conditions result in the real axis points $\Gamma_0 = +1$ and $\Gamma_0 = -1$. In other words, all values of the reflection coefficient reside within the unit circle. *See* ALTERNATING-CURRENT CIRCUIT THEORY; COMPLEX NUMBERS AND COMPLEX VARIABLES.

Impedance and reflection coefficient representation. Smith's contribution was the recognition that a generic complex impedance $Z = R + jX$ can be expressed in terms of circles that are superimposed over the reflection coefficient representation. The

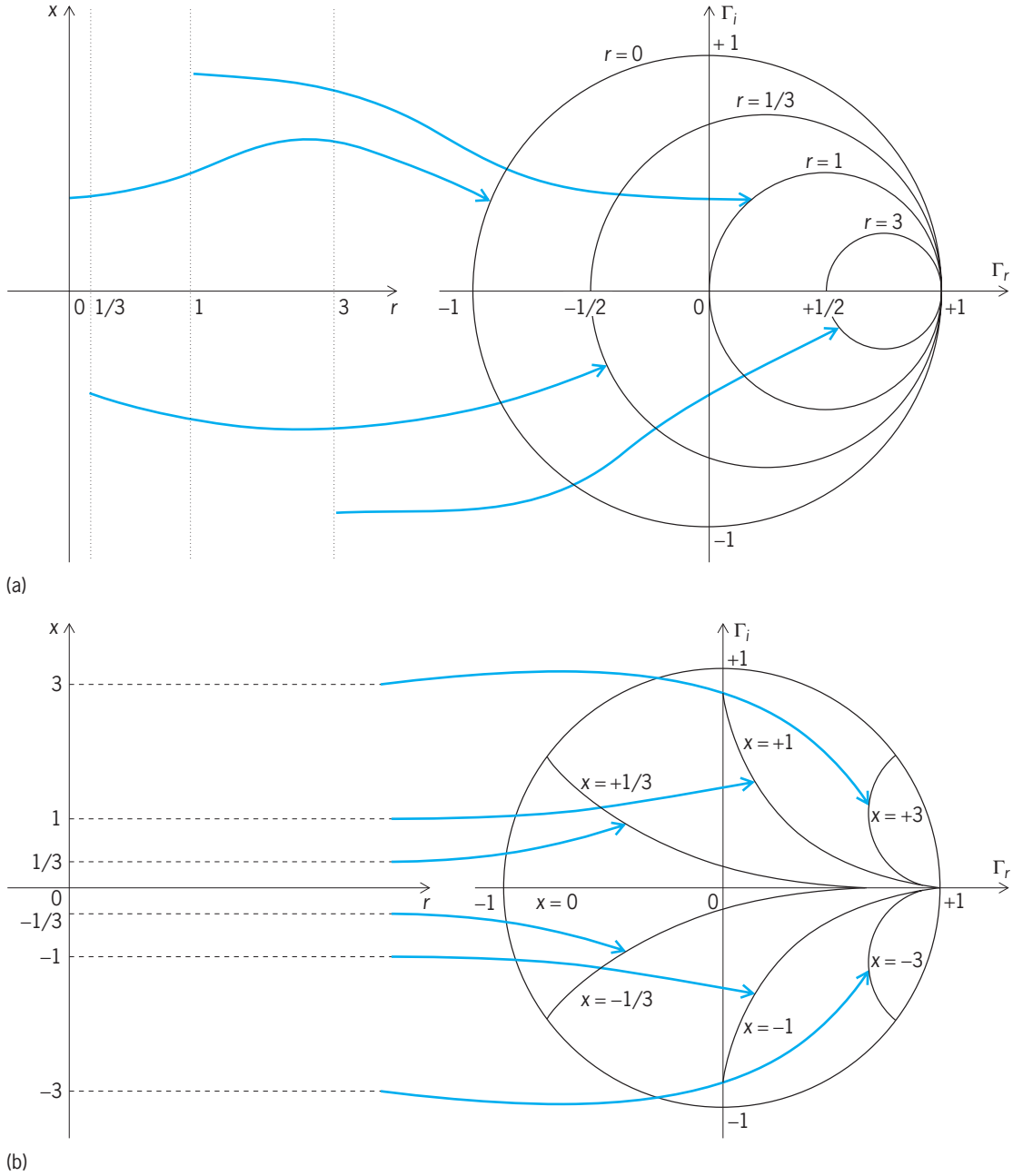

Fig. 2. Mapping of normalized resistance and reactance values into the complex reflection coefficient plane. (*a*) Mapping of normalized resistance values $r = 0$, 1/3, 1, 3 of the *z*-plane to corresponding *r*-circles in the complex Γ-plane. (*b*) Mapping of normalized reactance values $x = 0, \pm 1/3, \pm 1, \pm 3$ of the *z*-plane to corresponding *x*-circles in the complex Γ-plane.

mapping from the impedance plane into the reflection coefficient plane is done by first normalizing the impedance with respect to the characteristic impedance Z_0, $z = Z/Z_0 = r + jx$, with $r = R/Z_0$ and $x = X/Z_0$. This allows the display of all impedances within the unit circle. **Figure 2***a* depicts the mapping of *r*-values into shifted circles along the real axis of the reflection coefficient plane. As $r \to \infty$, the shifted circles get smaller and begin to collapse to a single point at the right side. Figure 2*b* shows a similar pattern for the reactance. However, unlike the resistance circles, the shifted reactance circles reside in the upper or lower plane, consistent with the fact that a circuit can exhibit inductive or capacitive behavior. *See* REACTANCE.

The trick is now to superimpose both the resistance and reactance circles over the complex unit circle plane, resulting in the final form of the Smith chart, shown in **Fig. 3**. In this figure it is evident how various complex normalized impedances can be interpreted as corresponding reflection coefficients. For instance, the four impedances of $(50 + j150)\ \Omega$, $(150 + j50)\ \Omega$, $(16.66 - j16.66)\ \Omega$, and $(150 - j150)\ \Omega$ are normalized with respect $Z_0 = 50\ \Omega$, yielding the normalized impedances of $(1 + j3)$, $(3 + j1)$, $(1/3 - j1/3)$, and $(3 - j3)$. These four points in the z plane are mapped into corresponding resistance and reactance circles in the complex Γ plane, as seen in Fig. 3. The intersecting resistance and reactance circles are identified as the

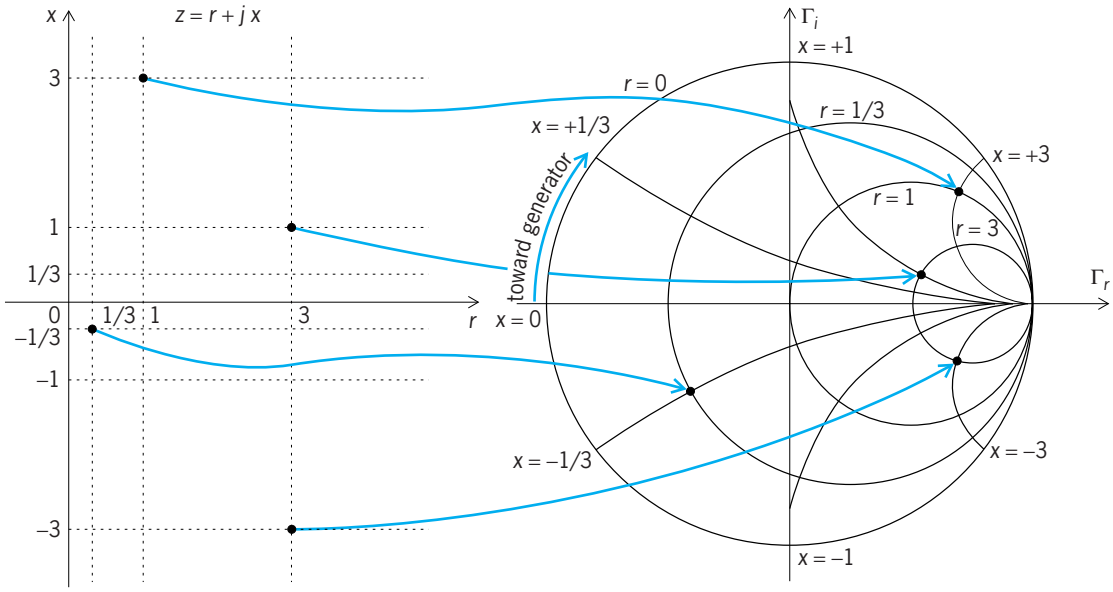

Fig. 3. The Smith chart is simply an overlay of resistance and reactance circles in the complex Γ-plane.

endpoints of vectors which are equivalent reflection coefficient phasors.

An important location is the normalized impedance value of $z = 1$; this spot maps into the center of the Smith chart, that is, $\Gamma = 0$. In the context of the above equation for the reflection coefficient, at this point the impedance is matched to the characteristic line impedance, $Z_0 = 50\ \Omega$. Matching to 50 Ω is often a key design requirement when dealing with conditioning the input and output ports of circuits like amplifiers, mixers, and oscillators.

The identification of an impedance point in the Smith chart and its corresponding reflection coefficient is but one consideration. Conversely, a reflection coefficient with its magnitude and phase can also be identified in terms of the normalized impedance value. Most radio-frequency (RF) devices are recorded in data sheets based on their scattering or S-parameters. For instance, to obtain the so-called input and output reflection coefficients S_{11} and S_{22}, it is sufficient to simply identify the intersecting resistance and reactance circles in order to obtain the equivalent normalized input and output impedances of the device.

Computations with the Smith chart. A typical example of how the Smith chart is employed to find the input impedance of a variable length transmission line (coaxial cable) is shown in **Fig. 4**. Here a frequency-independent load impedance of $Z_L = (30 + j60)\ \Omega$ is attached to the transmission line with characteristic impedance of $Z_0 = 50\ \Omega$. At a fixed operating frequency f of 2 GHz and based on a given phase velocity of $v_p = 0.5c$ (that is, 50% of the speed of light c), the coaxial cable is varied in length d. As the length is increased from 0 to 3 cm, the input impedance Z_{in} varies from $(30 + j60)\ \Omega$ to $(12.4 + j15.5)\ \Omega$. This is due to the fact that the reflection coefficient $\Gamma_0 = 0.63\angle 72^\circ$ begins to rotate in the mathematical negative sense from $\Gamma_{in} = \Gamma_0$ to the new location $\Gamma_{in} = \Gamma_0\angle(-2\beta d) = 0.63\angle -216^\circ$. Here $\beta d = 2\pi f d/v_p = 2.51$ rad (or 144°) is also known as the electric length of the coaxial cable. The rotation involves twice the electric length owing to the fact the voltage-current waves have to travel from the input of the coaxial cable to the load and back again. Reading off the normalized resistance and reactance circles, it is found that $r = 0.248$ and $x = 0.31$. Multiplying these values by $Z_0 = 50\ \Omega$ provides the final result of $Z_{in} = (12.4 + j\,15.5)\ \Omega$.

Additional applications. Besides impedance transformations, the Smith chart finds widespread use for a number of RF applications; it is routinely utilized for noise figure, gain, and voltage standing-wave ratio (VSWR) computations. As an example, for the design of a low-noise amplifier (LNA) the optimal reflection coefficient for the minimum noise figure can typically be found in the data sheets of the

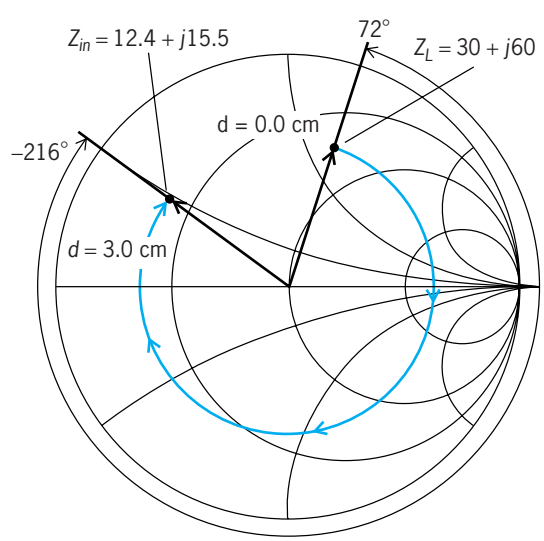

Fig. 4. Input impedance transformation of a coaxial cable from Z_L (at length $d = 0$) to Z_{in} (at $d = 3.0$ cm) at 2 GHz. The reflection coefficient of fixed magnitude is rotated from its original angular location of 72° to −216° in clockwise (or mathematical negative) direction toward the generator.

RF device. From the knowledge of the corresponding impedance value, the design engineer can then noise-match the device by designing an appropriate network whose impedance is matched to the conjugate complex input impedance of the device. Besides being useful for the design and simulation of RF systems, the Smith chart is frequently found in a number of measurement devices, notably the network analyzer. Here, the RF system under test is attached to the input and output ports of the network amplifier; a frequency sweep then allows the display of the input and output reflection coefficients as traces in the Smith chart. *See* MICROWAVE IMPEDANCE MEASUREMENT; RADIO-FREQUENCY AMPLIFIER.

Reinhold Ludwig

Bibliography. K.-C. Chan and A. Harter, Impedance matching and the Smith chart—The fundamentals, *RF Design Mag.*, pp. 52–66, July 2000; R. Ludwig and P. Bretchko, *RF Circuit Design, Theory and Applications*, Prentice Hall, 2000; B. Schweber, The Smith chart: More vital after all these years, *EDN Mag.*, pp. 50–52, March 19, 1999; P. H. Smith, *Electronic Applications of the Smith Chart*, 2d ed., Crestone Technical Books/Noble Publishing, 1995; P. H. Smith, Transmission-line calculator, *Electronics*, 12:29–31, 1939.

Smithsonite

A naturally occurring rhombohedral zinc carbonate ($ZnCO_3$), with a crystal structure similar to that of calcite ($CaCO_3$). The structure consists of alternating layers of zinc ions (Zn^{2+}) and carbonate ions (CO_3^{2-}) oriented perpendicular to the *c* crystallographic axis. Smithsonite has a hardness on the Mohs scale of $4^1/_2$, has a specific gravity of 4.30–4.45, and exhibits perfect rhombohedral cleavage. Optically, smithsonite is uniaxial negative, which means that it has a single optic axis and two principal refractive indices referred to as the ordinary refractive index and the extraordinary refractive index. Smithsonite is classified as optically negative because the ordinary refractive index is greater than the extraordinary index. ($N_o > N_e$). *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY; HARDNESS SCALES.

Smithsonite most commonly forms as an alteration product of the mineral sphalerite (ZnS) during supergene enrichment of zinc ores in arid or semiarid environments. During this process, primary zinc minerals are oxidized and dissolved near the surface of the Earth. Although zinc ion is significantly more soluble than many other metals present in sulfide ore deposits (for example, copper, iron, lead, and silver), if solutions rich in zinc react with carbonate-rich surrounding rocks such as limestone, smithsonite may precipitate. Smithsonite thus deposited seldom occurs as well developed crystals but forms colliform, or botryoidal, encrusting masses, usually brownish or grayish white in color. It is seldom pure zinc carbonate, commonly containing other divalent metal ions such as manganese (Mn^{2+}), ferrous iron (Fe^{2+}), magnesium (Mg^{2+}), calcium (Ca^{2+}), cadmium (Cd^{2+}), cobalt (Co^{2+}), or lead (Pb^{2+}) substituting for zinc ion. Substitution of other elements for zinc may result in different colors such as blue-green (copper), yellow (cadmium), and pink (cobalt). Minerals frequently associated with smithsonite are hemimorphite [$Zn_4Si_2O_7(OH)_2 \cdot H_2O$], sphalerite (ZnS), galena (PbS), calcite ($CaCO_3$), and other secondary, supergene minerals.

Smithsonite has been mined as an ore of zinc and has also been used as an ornamental stone. Notable localities include Laurium, Greece; Sardinia, Italy; the Broken Hill Mine, Zambia; Tsumeb, Namibia; Broken Hill, New South Wales, Australia; and Kelly, New Mexico. *See* CALCITE; CARBONATE MINERALS; GALENA; HEMIMORPHITE; SPHALERITE; ZINC.

John C. Drake

Bibliography. L. L. Y. Chang, R. A. Howie, J. Zussman, *Rock-Forming Minerals*, vol. 5B: *Non-Silicates: Sulphates, Carbonates, Phosphates, Halides*, 1995; C. Klein and C. S. Hurlbut, Jr., *Manual of Mineralogy*, 21st ed., 1993; C. Palache, H. Berman, and C. Frondel, *The System of Mineralogy*, 7th ed., vol. 2, 1951; R. J. Reeder (ed.), *Reviews in Mineralogy*, vol. 11: *Carbonates: Mineralogy and Chemistry*, 1983.

Smog

The noxious mixture of gases and particles commonly associated with air pollution in urban areas. Harold Antoine des Voeux is credited with coining the term in 1905 to describe the air pollution in British towns. *See* AIR POLLUTION.

The constituents of smog affect the human cardiorespiratory system and pose a health threat. Individuals exposed to smog can experience acute symptoms ranging from eye irritation and shortness of breath to serious asthmatic attacks. Under extreme conditions, smog can cause mortality, especially in the case of the infirm and elderly. Smog can also harm vegetation and likely leads to significant losses in the yields from forests and agricultural crops in affected areas.

The only characteristic of smog that is readily apparent to the unaided observer is the low visibility or haziness that it produces, due to tiny particles suspended within the smog. Observation of the more insidious properties of smog—the concentrations of toxic constituents—requires sensitive analytical instrumentation. Technological advances in these types of instruments, along with the advent of high-speed computers to simulate smog formation, have led to an increasing understanding of smog and its causes. In much of the industrially developed world, this understanding has been used to design and implement pollution control measures to mitigate smog's harmful effects. The smoggiest cities today are mostly found in the developing world (for example, Chengdu, China; Mexico City, Mexico; and Santiago, Chile) where pollution controls are generally given a lower priority than economic growth. However, even in the developed world smog remains a significant environmental problem. For example, in the United States about 70 million people are

estimated to live in areas where smog poses a threat to human health. Moreover, while smog is generally most severe in and around urban centers, it is by no means only an urban problem. Under the appropriate conditions, smog can extend over thousands of kilometers and encompass agricultural and forested areas as well as urban areas. Some data even suggest that smog affects the composition of the atmosphere over the entire Northern Hemisphere.

Role of meteorology. Smog is an episodic phenomenon because specific meteorological conditions are required for it to accumulate near the ground. These conditions include calm or stagnant winds which limit the horizontal transport of the pollutants from their sources, and a temperature inversion which prevents vertical mixing of the pollutants from the boundary layer into the free troposphere. (The boundary layer refers to the lowest 1–2 km, or 0.6–1.2 mi, of the atmosphere, where winds are slowed by frictional resistance with the Earth's surface. The free troposphere refers to that part of the atmosphere lying between the boundary layer and the stratosphere.) *See* STRATOSPHERE; TROPOSPHERE.

The importance of meteorology also explains why some areas experience especially severe and frequent smog episodes and others do not. In cities such as Chengdu, Los Angeles, Vancouver, and Mexico City, surrounding mountain ranges restrict wind flow and prevent pollutant ventilation. In coastal cities such as Houston, the combination of nighttime offshore land breezes and daytime onshore sea breezes recycles urban pollutants and helps maintain smoggy conditions. *See* METEOROLOGY.

Smog can be classified into three types: London smog, photochemical smog, and smog from biomass burning.

London smog. Becoming prevalent in the midnineteenth century, London smog arises from the by-products of coal burning. These by-products include soot particles and sulfur oxides. During cool damp periods (often in the winter), the soot and sulfur oxides can combine with fog droplets to form a dark acidic fog.

The generation of London smog is not unique to modern times. For example, Theophrastus, in the third century B.C., wrote of the noxious quality of air exposed to coal burning. During the industrial revolution, the growing use of coal coupled with urbanization fostered more frequent and severe smog episodes in many European cities and especially in London. The first known scientific treatise on air pollution was produced by John Evelyn in 1661 and focused specifically on London smog. Probably the worst smog episode occurred in London in 1952 with an estimated 4000 deaths. Other documented killer smogs include episodes occurring in Glasgow in 1909, over the Meuse Valley in Belgium in 1930, in Donora, Pennsylvania, in 1948, in New York City in 1953, and in London in 1962. *See* ACID RAIN; COAL.

Photochemical smog. As nations switch from coal to cleaner-burning fossil fuels such as oil and gas as well as alternate energy sources such as hydroelectric and nuclear, London smogs cease. Unfortunately, photochemical smog then tends to appear. Photochemical smog is more of a haze than a fog and is produced by chemical reactions in the atmosphere that are triggered by sunlight. Because photochemical smog requires high solar intensities, its occurrence is rare when the dark foggy conditions of a London-type smog are present. *See* FOSSIL FUEL; PHOTOCHEMISTRY.

The phenomenon of photochemical smog was discovered and studied in the 1950s in Los Angeles when a growing urban population coupled with an increasing use of the automobile correlated with a discernible increase in hazy skies and an increase in the concentrations of compounds known as oxidants. At the same time, there appeared to be an increasing incidence of damage to crops in the areas surrounding Los Angeles.

A. J. Hagen-Smit first unraveled the chemical mechanism that produces photochemical smog. He irradiated mixtures of volatile organic compounds (VOC) and nitrogen oxides (NO_x) in a reaction chamber. (Note that both VOC and NO_x are emitted into the atmosphere by automobiles and other industrial processes, and NO_x represents the sum of nitric oxide or NO and nitrogen dioxide or NO_2.) After a few hours, Hagen-Smit observed the appearance of cracks in rubber bands stretched across the chamber. Knowing that ozone (O_3) can harden and crack rubber, Hagen-Smit correctly reasoned that photochemical smog was caused by photochemical reactions involving VOC and NO_x, and that one of the major oxidants produced in this smog was O_3. *See* NITROGEN OXIDES; OZONE.

Since Hagen-Smit's work, scientists have used more sophisticated reaction chambers, called smog chambers, along with advanced chemical monitors and computer simulation models to more fully elucidate the chemical reactions that form smog. A typical reaction sequence might include (1)–(8). Here, RH

$$\mathrm{RH + OH \longrightarrow R + H_2O} \tag{1}$$

$$\mathrm{R + O_2 + M \longrightarrow RO_2 + M} \tag{2}$$

$$\mathrm{RO_2 + NO \longrightarrow RO + NO_2} \tag{3}$$

$$\mathrm{RO + O_2 \longrightarrow HO_2 + carbonyl} \tag{4}$$

$$\mathrm{HO_2 + NO \longrightarrow OH + NO_2} \tag{5}$$

$$\mathrm{2 \times (NO_2 + h\nu \longrightarrow NO + O)} \tag{6}$$

$$\mathrm{2 \times (O + O_2 + M \longrightarrow O_3 + M)} \tag{7}$$

$$\mathrm{RH + 4O_2 + 2h\nu \longrightarrow 2O_3 + carbonyl + H_2O} \tag{8}$$

represents a hydrocarbon (that is, a type of VOC), R is a hydrocarbon radical or compound fragment (for example, CH_3CH_2 for RH = ethane), $h\nu$ represents a photon from the Sun, and M is a nonreactive, energy-absorbing molecule referred to by chemists as a third body. In the atmosphere, this third body is most often molecular nitrogen (N_2) or molecular oxygen (O_2). Carbon monoxide (CO) can also act as

a VOC in photochemical smog. In this case the reaction sequence is initiated by (9) and (10) and then followed by (5)–(7).

$$CO + OH \longrightarrow CO_2 + H \qquad (9)$$

$$H + O_2 + M \longrightarrow HO_2 + M \qquad (10)$$

In these reaction sequences, VOC and CO are oxidized and consumed, while NO_x acts as a catalyst. Moreover, the by-product labeled "carbonyl" is itself a VOC and can react and produce more O_3. The OH radical, which triggers the sequence by reacting with VOC or CO, is initially generated in the atmosphere from O_3 as in reactions (11) and (12). Here O^* is an

$$O_3 + h\nu \longrightarrow O^* + O_2 \qquad (11)$$

$$O^* + H_2O \longrightarrow 2OH \qquad (12)$$

electronically excited atom of oxygen. Because the OH consumed by reaction (1) or (9) is regenerated by reaction (5), and because more OH can be generated during the breakdown of the carbonyl produced by reaction (4), the reaction sequences above are net sources of O_3. *See* ATMOSPHERIC CHEMISTRY.

Noxious constituents. While generally not as dangerous as London smog, photochemical smog contains a number of noxious constituents. Ozone, a strong oxidant that can react with living tissue, is one of these noxious compounds. Another is peroxyacetyl nitrate (PAN), an eye irritant that is produced by reactions between NO_2 and the breakdown products of carbonyls. Particulate matter having diameters of about 10 micrometers or less is another harmful smog component. Particles in this size range are of concern because they can penetrate into the human respiratory tract during breathing and have been implicated in a variety of respiratory ailments.

Mitigating photochemical smog. Because VOC and NO_x are required to generate O_3, they are referred to as the precursors of O_3. Human-induced NO_x is produced in high-temperature combustion, and human-induced VOC arises from the incomplete combustion of fossil fuels and the evaporation of solvents. Volatile organic carbons are also emitted into the atmosphere naturally by trees and other vegetation. Because the two precursors have different sources, controls that limit VOC emissions usually do not reduce NO_x emissions, and vice versa. Thus, before initiating a program to mitigate photochemical smog and O_3 pollution, it must first be determined which O_3 precursors to target for emission reductions.

Since both VOC and NO_x are required to produce O_3, it might be concluded that smog can be reduced by any one of three strategies, based on lowering (1) the emissions of VOC, (2) NO_x emissions, or (3) both. However, this conclusion would be incorrect. The photochemical smog mechanism is complex, and the relative effectiveness of pollution control strategies based on lowering VOC or NO_x emissions can vary depending upon local conditions. In various locales, VOC controls might be effective and NO_x controls ineffective, and vice versa. Ozone mitigation can be further confounded by natural emissions of VOC. In sylvan cities, such as Atlanta, Georgia, large emissions of hydrocarbons from trees (primarily as isoprene) render pollution control strategies based on lowering human-induced VOC emissions relatively ineffective and necessitate a greater focus on lowering NO_x emissions. Another complication arises from the fact that, under some circumstances, lowering NO_x emissions can actually lead to a temporary O_3 increase and occurs because of high concentration NO_x scavenger OH radicals, thereby preventing initiation of O_3 production via reactions (1) and (9). This is referred to as a NO_x disbenefit. It is also the case that strategies that effectively lower O_3 concentrations may not reduce the concentrations of other noxious smog components such as particulate matter. For these reasons, determining the best pollution control strategy for a given area has proven to be a difficult task, requiring the application of sophisticated air quality simulation models.

Tropospheric and stratospheric ozone. In addition to being a component of photochemical smog, O_3 is a natural constituent of the atmosphere (see **illus.**). In the stratosphere, where it is most abundant, O_3 actually protects life by filtering out potentially harmful ultraviolet radiation from the Sun. In the troposphere (the lowest 10–12 km or 6–7.5 mi of the atmosphere), O_3 arises naturally from the downward transport of stratospheric O_3 as well as photochemical production. Interestingly, the same photochemical

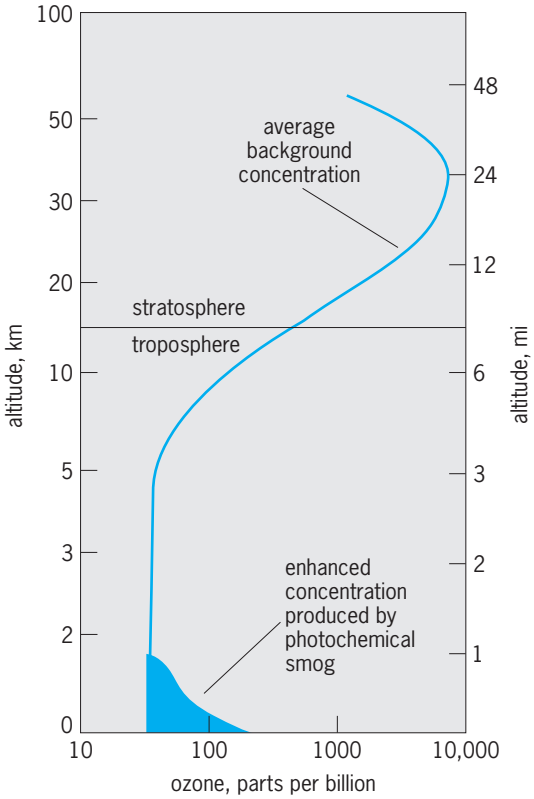

Average background concentration of ozone in the atmosphere as a function of altitude. Note the enhancement of O_3 near the Earth's surface that can occur in polluted areas as a result of photochemical smog.

reactions that produce O_3 in photochemical smog also produce tropospheric O_3. In the latter case, methane (CH_4) and carbon monoxide (CO) act as VOC precursors, while NO_x is primarily supplied by lightning. It is likely that pollution has accelerated the rate of production of tropospheric O_3 on a hemispheric scale by adding to the total atmospheric VOC, CO, and NO_x burdens. Some data suggest that present-day O_3 concentrations throughout much of the Northern Hemisphere are a factor of 2 or more larger than they were a hundred years ago.

Smog from biomass burning. Probably the oldest type of smog known to humankind is produced from the burning of biomass or wood. It combines aspects of both London smog and photochemical smog since the burning of biomass can produce copious quantities of smoke as well as VOC and NO_x.

Biomass-burning smog was most likely ubiquitous in preindustrial times. For example, in 1542 Juan Rodriguez Cabrillo, upon first arriving at the Los Angeles harbor, noted the heavy smoke from Indian campfires and named the place Bahia de los Humos (Bay of Smokes).

Even in modern times, biomass-burning smog can pose a serious environmental threat. In the 1980s, Nobel prize winner Paul Crutzen and his colleagues documented the large-scale degradation in air quality over South America from agricultural burning. During the summer of 1997, forest fires and agricultural burning in Indonesia produced a thick blanket of smoke-laden smog over much of Southeast Asia. The smog caused widespread economic disruptions and agricultural losses. The human toll was also high, with tens of thousands falling ill and many deaths. Indeed, appreciation of the dangers of biomass-burning smog led Crutzen and John Birks in 1982 to warn that a nuclear exchange and the resulting fires could cause a "nuclear winter" with catastrophic environmental consequences for the planet. William L. Chameides

Bibliography. J. Fishman and R. Kalish, *Global Alert: The Ozone Pollution Crisis*, Plenum, New York, 1990; National Research Council, *Rethinking the Ozone Problem in Urban and Regional Air Pollution*, National Academy Press, Washington, DC, 1991; R. P. Turco, *Earth Under Siege: From Air Pollution to Global Change*, Oxford University Press, New York, 1997.

Smoke

A dispersion of small, solid particles and liquid droplets suspended in a gaseous medium, usually produced by a combustion process. Smoke particles start as groups of carbon atoms or small fragments of the burning material which grow in size by coagulation and condensation as they are carried from the combustion zone by the buoyancy of the hot gases. These particles are usually found to have diameters in the range 0.01–5.0 micrometers. The solid particles tend to be somewhat irregular in shape and are often seen as clusters or long chains of dark material. Smokes from flaming combustion consist mainly of solid particles. *See* COMBUSTION.

Liquid particles are always spherical and consist of water and high-molecular-weight compounds which have condensed on a small, solid nucleus. These smokes usually appear gray or white but may also appear to be lightly colored by the condensed compounds. Smokes produced by pyrolysis or smoldering combustion are largely made up of liquid particles.

Smokes can also be produced by chemical reactions or by mechanical generators such as atomizers or nebulizers. Chemical smokes use such compounds as phosphorus pentoxide (P_2O_5), ammonium chloride (NH_4Cl), titanium tetrachloride ($TiCl_4$), or aluminum chloride ($AlCl_3$). Mechanically generated smokes can be produced from various oils such as dioctylpthalate, corn oil, or silicones, or by atomizing salt solutions (for example, sodium chloride or potassium chloride) or sugar solutions (glucose or sucrose), leaving solid particles of a fairly uniform size after evaporation of the surrounding liquid.

These chemical smokes are used for such things as airflow visualization, theatrical effects, vision obscurants for crowd control or by the military, and daylight signaling devices. Mechanically generated smokes are often used for measuring the efficiency of filters because their sizes are fairly well defined by the material and production technique employed. *See* AIR FILTER; SCREENING SMOKE. Richard W. Bukowski

Bibliography. J. K. Beddow (ed.), *Particle Characterization in Technology*, 2 vols., 1984; R. Clift, *Particulate Technology*, 1991; H. K. Hasegawa (ed.), *Characterization and Toxicity of Smoke*, 1990.

Snail

Any of the approximately 74,000 species in the class Gastropoda of the phylum Mollusca or, alternatively, any of the 12 or so species of land pulmonate gastropods used as human food.

General characteristics. The shell of snails is in one piece and typically turbinate (see **illus.**), but may be planospiral or limpet-shaped, or may be secondarily lost (as in land slugs and marine nudibranchs). In development, gastropods have undergone torsion (the visceral mass and the mantle-shell covering it have become twisted through 180° in relation to the head and foot) so that the mantle cavity is placed anteriorly above the head.

A few primitive snails retain two aspidobranch (featherlike) gills (the characteristic ctenidia which are structurally and functionally homologous throughout the Mollusca). Most gastropods show considerable asymmetry and simplification with reduction to a single aspidobranch ctenidium or, more often, to a pectinibranch ctenidium (or comblike "half-gill"). This is accompanied by corresponding changes that reduce paired internal organs to a single kidney, a single gonad, and a single cardiac auricle. True ctenidia are lost in most of the shellless marine nudibranchs, and in the fresh-water and

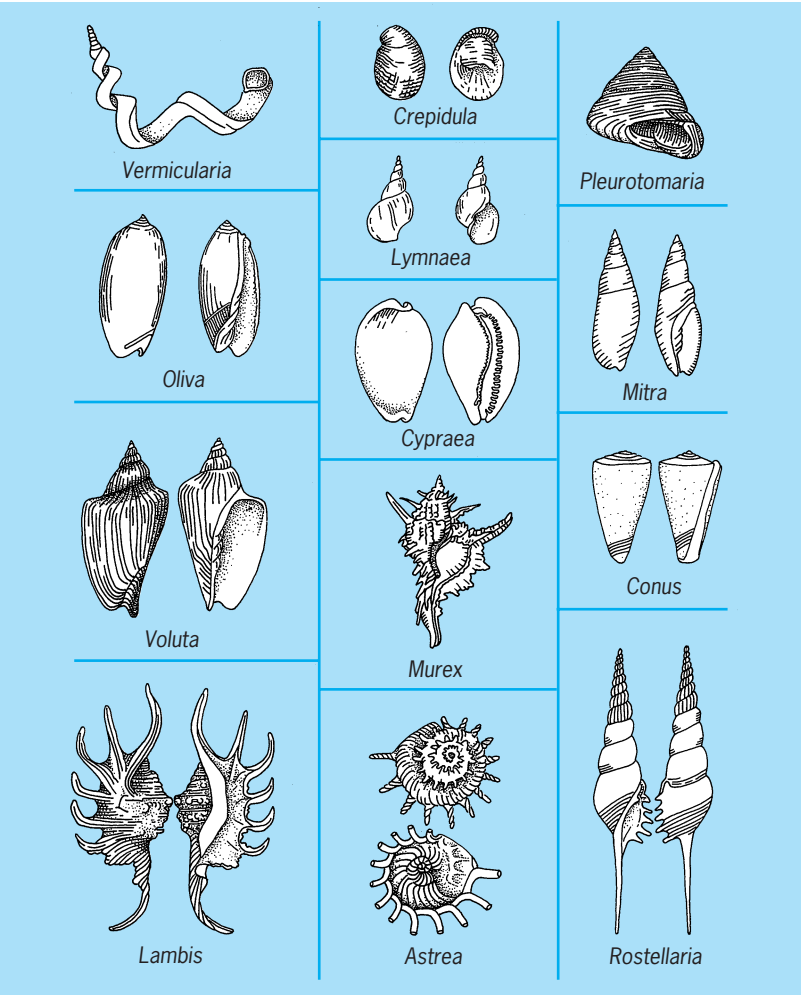

Diversity of snail shells. Except for *Lymnaea* (a fresh-water pulmonate), all genera represented are marine prosobranchs. (*After R. R. Shrock and W. H. Twenhofel, Principles of Invertebrate Paleontology, 2d ed., McGraw-Hill, 1953*)

terrestrial snails of the subclass Pulmonata, in which the mantle cavity is an air-breathing lung.

In feeding, all snails use a characteristic rasping tongue or radula. This is a chitinous ribbon bearing teeth which is moved over a supporting protrusible "tongue" (the odontophore) with a to-and-fro action. The radular apparatus has a twofold function: it serves both for rasping off food material (mechanically like an inverted version of the upper incisor teeth of a beaver) and for transporting the food back into the gut like a conveyor belt. Different patterns of radular tooth rows are found in the major gastropod groups; in general, carnivorous snails have fewer (but more differentiated) teeth per row.

The class Gastropoda is divided into three somewhat unequal subclasses, the Prosobranchia, Opisthobranchia, and Pulmonata. The Prosobranchia, made up of marine gastropods with a few nonmarine forms, are the largest and most diverse. The Opisthobranchia, entirely marine, consist largely of shell-less sea slugs and the highly colorful nudibranchs. The Pulmonata are basically air-breathers that use the mantle cavity as a lung, and include the most highly evolved gastropods, with representatives on land, in fresh waters, and (by secondary reversion) on the seashore. *See* OPISTHOBRANCHIA; PROSOBRANCHIA; PULMONATA.

Both fresh-water and terrestrial snail species serve as vectors (first or sole intermediate hosts) in the transmission of flukes (Trematoda) infecting humans or domestic animals. For further discussion of these parasitic diseases *see* DIGENEA; SCHISTOSOMIASIS.

Edible snails. Although many marine gastropods are eaten by humans, including periwinkles (*Littorina*), moon snails (*Polinices*), and "conchs" of various genera, the "snails" or *Schnecken* or *escargots* of gastronomy belong to a very few species of larger pulmonate land snails in the family Helicidae. Certain species belonging to genera such as *Cepaea*, *Otala*, and *Helix* are widespread in many continents as a result of human cultural migrations. All are capable of long periods of dormancy (with the shell aperture sealed by a series of membranes, the epiphragm, and metabolism reduced over twentyfold), and this facilitates human transport.

Two such European land snails, *C. hortensis* and *C. nemoralis*, are widespread and abundant. Recent work on their population genetics includes some of the most convincing demonstrations of natural selection in action. Both species are markedly polymorphic in shell colors and patterns. The basic shell color may be bright yellow, dull brown, or any shade from very pale fawn through pink and orange to red. Up to five longitudinal dark bands may be present, and all possible combinations of presence or absence of them, and of fusion between adjacent bands, have been described. All of these variables are genetically determined.

Most natural populations of *Cepaea* consist of two or more different varieties living together. The proportions of varieties in each colony correlate with the general class of background in the habitat where the snails are living. Predators include songbirds and small mammals, and in some cases it is possible to have direct quantification of selection pressures by comparing the proportions of each variety eaten to the proportions present in the snail colony. Considerable efforts of population geneticists are being directed to the elucidation of other factors involved in maintenance of this complex polymorphism in natural populations of *Cepaea*. *See* GASTROPODA; MOLLUSCA; POLYMORPHISM (GENETICS).

W. D. Russell-Hunter

Snap ring

A form of spring used principally as a fastener. Piston rings are a form of snap ring used as seals. The ring is elastically deformed, put in place, and allowed to snap back toward its unstressed position into a groove or recess. The snap ring may be used externally to provide a shoulder that retains a wheel or a bearing race on a shaft, or it may be used internally to provide a construction that confines a bearing race in the bore of a machine frame (see **illus.**). The size of the ring and its recess determines its strength under

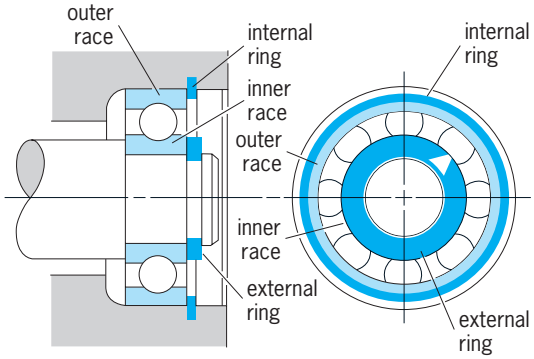

Snap rings hold ball-bearing race in place. Internal ring supports axial thrust of axle. External ring aligns inner race against shoulder of shaft.

load. Sufficient clearance and play is needed in the machine so that the ring can be inserted and seated. *See* ANTIFRICTION BEARING; INTERNAL COMBUSTION ENGINE; SPRING (MACHINES). L. Sigfred Linderoth, Jr.

Bibliography. H. Carlson, *Spring Designers Handbook*, 1978.

Snow

Frozen precipitation resulting from the growth of ice crystals from water vapor in the Earth's atmosphere.

Formation and growth. As ice particles fall out in the atmosphere, they melt to raindrops when the air temperature is a few degrees above 32°F (0°C), or accumulate on the ground at colder temperatures. At temperatures above −40°F (−40°C), individual crystals begin growth on icelike aerosols (often clay particles 0.1 micrometer in diameter), or grow from cloud droplets (10 μm in diameter) frozen by similar particles. At lower temperatures, snow crystals grow on cloud droplets frozen by random molecular motion. At temperatures near 25°F (−4°C), crystals sometimes grow on ice fragments produced during soft hail (graupel) growth. Snow crystals often grow in the supersaturated environment provided by a cloud of supercooled droplets; this is known as the Bergeron-Findeisen process for formation of precipitation. When crystals are present in high concentrations (100 particles per liter) they grow in supersaturations lowered by mutual competition for available vapor. **Figure 1** shows the vapor pressure of water and ice and vapor pressure excess of supercooled water over ice at different temperatures. The difference has a maximum near 9.5°F (−12.5°C).

Ice crystals growing under most atmospheric conditions (air pressure down to 0.2 atm or 20 kilopascals and temperatures 32 to −58°F or 0 to −50°C) have a hexagonal crystal structure, consistent with the arrangement of water molecules in the ice lattice, which leads to striking hexagonal shapes during vapor growth. The crystal habit (ratio of growth along and perpendicular to the hexagonal axis) changes dramatically with temperature. Both field and laboratory studies of crystals grown under known or controlled conditions show that the crystals are platelike above 27°F (−3°C) and between 18 and −13°F (−8 and −25°C), and columnlike between 27 and 18°F (−3 and −8°C) and below −13°F (−25°C). Aspect ratios can be large—hexagonal plates being 50 μm thick and several millimeters across, and hexagonal columns being 100 μm across and several millimeters long. Individual crystals fall in the atmosphere at velocity up to 0.5 m s^{-1} (1.6 ft s^{-1}). As crystals grow, they fall at higher velocity, which leads, in combination with the high moisture availability in a supercooled droplet cloud, to sprouting of the corners to form needle or dendrite skeletal crystals. These are more evident for growth

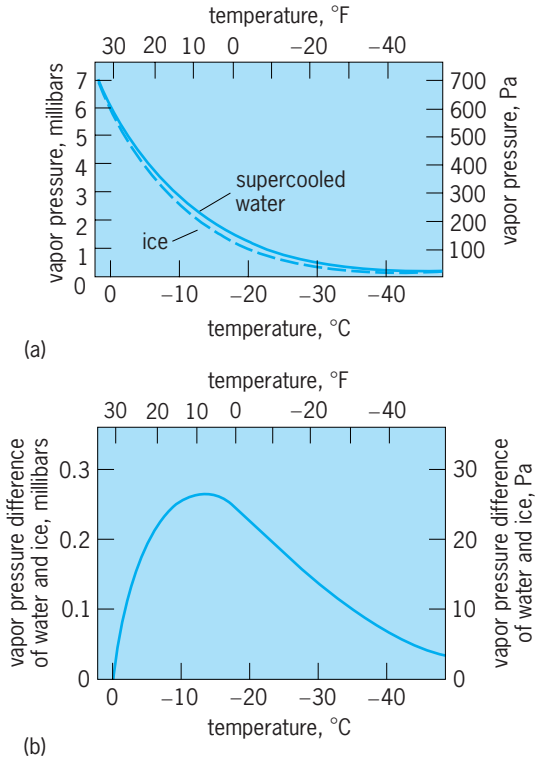

Fig. 1. Graphs showing variation of vapor pressure with temperature. (*a*) Vapor of supercooled water and ice. (*b*) Vapor pressure difference between supercooled water and ice. Maximum occurs at 9.6/12.5°F (−12.5°C).

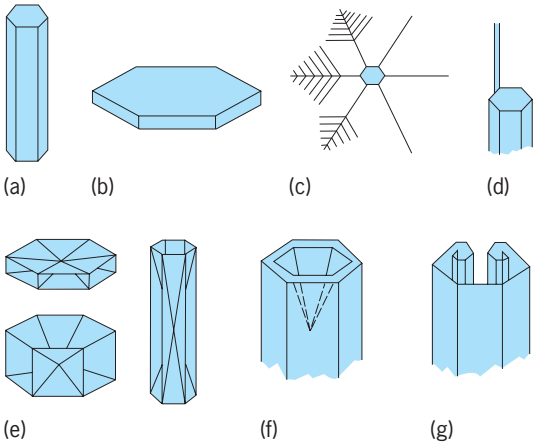

Fig. 2. Different forms of snow crystals. (*a*) Prism. (*b*) Plate. (*c*) Dendrite. (*d*) Needle. (*e*) Hopper-type crystals. (*f*) Hollow prism. (*g*) Scroll.

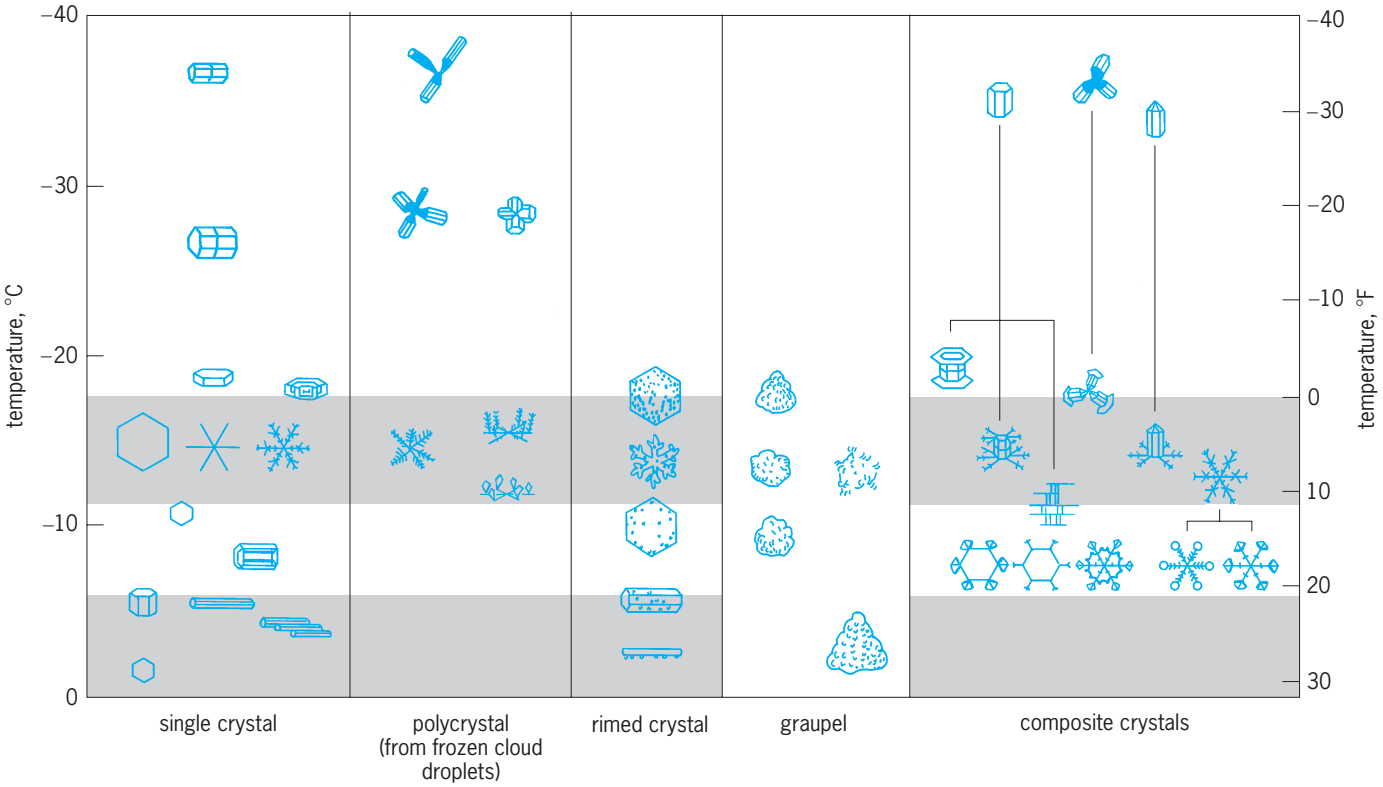

Fig. 3. Diagram showing forms of snow crystals and the temperature zones in which they occur (shaded areas). Lines indicate paths in which one type of composite crystal grows into another. (*After J. Hallett, How snow crystals grow, Amer. Sci., 72:582–589, November–December 1984*)

between 25 and 23°F (−4 and −5°C) and 10 and 3°F (−12 and −16°C) [**Fig. 2**].

As crystals fall through the atmosphere into levels of different temperature and moisture availability, their growth form changes with the environmental conditions; crystals of composite growth form (plates or dendrites with columns at the corners, a column with plates or dendrites on the end) demonstrate the subtleties of these changes. When crystals grow on cloud droplets frozen at low temperature (below 14 to −4°F or −10 to −20°C), crystals of several orientations grow radially outward to give a radial array of columns, plates, or dendrites (**Fig. 3**).

Through its shape, a snow crystal represents a record of the temperature–moisture–fall velocity history of its growth (**Fig. 4**). Two crystals that appear visually identical would have to have an identical origin and to follow an identical path through the atmosphere to reproduce temperature, moisture, and fall velocity. In view of the turbulent nature of the atmosphere, the likelihood of actually observing two crystals that followed an identical path is quite remote.

Recent concerns about global warming resulting from carbon dioxide (CO_2) production by the combustion of fossil fuel have emphasized the importance of high-level ice clouds (cirrus) forming at some 10–15 km (6–9 mi) in the atmosphere, both for their role in scattering back sunlight to space and thermal radiation back to the lower atmosphere and the ground. Such clouds are composed of complex

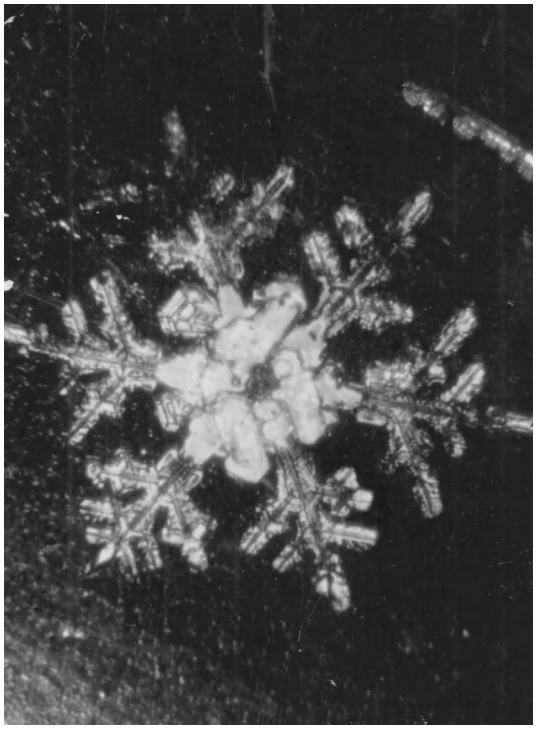

Fig. 4. Dendrite snow crystal, showing that its initial growth (center) began as a plate.

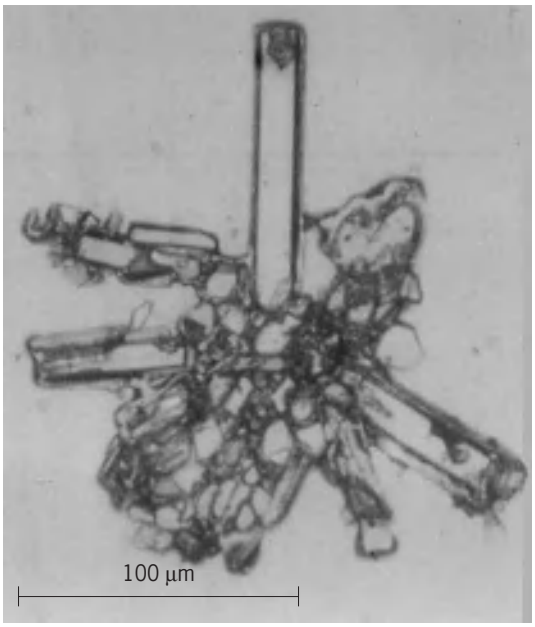

Fig. 5. Replica of an ice crystal collected by an aircraft during flight through a cirrus cloud. It is in the form of a rosette of individual columns, suggesting that the nucleating droplet froze as a polycrystal.

column-type crystals togather with plates (**Fig. 5**). The relative proportion depends on the nucleation processes (**Fig. 6**; cloud droplets may freeze as a single crystal or polycrystal) and the supersaturation as determined by the release of water vapor by local updraft and the removal by the growing crystals—which in turn depends on the crystal concentration. The effect on radiation depends either on crystal thickness or on crystal area, which depends in turn on the absorption at different wavelengths and the size and shape which influences scattering. Satellite measurements of temperatures of these clouds, used as a forecast tool, require interpretation in terms of the properties of high-level crystal cirrus shields. *See* GREENHOUSE EFFECT; WEATHER MODIFICATION.

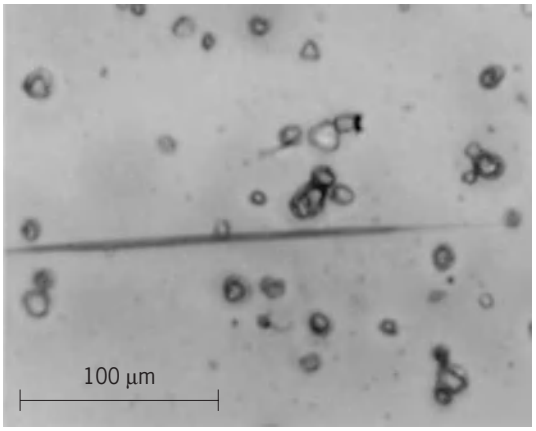

Fig. 6. Replica of crystals collected in high-level cirrus at about −50°C (−122°F) showing the simultaneous presence of hexagonal and triangular plates, a long column, and also more complex but equiaxed crystals. The habit variability suggests sequential nucleation events and growth under varying conditions. As such crystals fall to lower levels where vapor is available for growth, they form the core for larger snow crystals.

Frost crystals. Frost crystals differ from snow crystals in that the latter grow on a fixed surface and are ventilated by ambient wind. The frost crystals usually fail to have the exquisite symmetry of falling snow crystals since shielding of one side of the growing crystal often occurs at its point of attachment.

Crystal orientation. Small crystals tens of micrometers across are usually randomly oriented during their fall, but as they grow they become oriented by asymmetric flow around the falling particle, ultimately with their longest dimension horizontal by lee eddies formed by the airflow around the crystal. Larger crystals, greater than a few millimeters across (depending on shape and density), oscillate as they fall along a helical path as these eddies are shed. These orientations lead to complicated optical effects, often colored, such as haloes and arcs in sunlight and sometimes moonlight. *See* HALO.

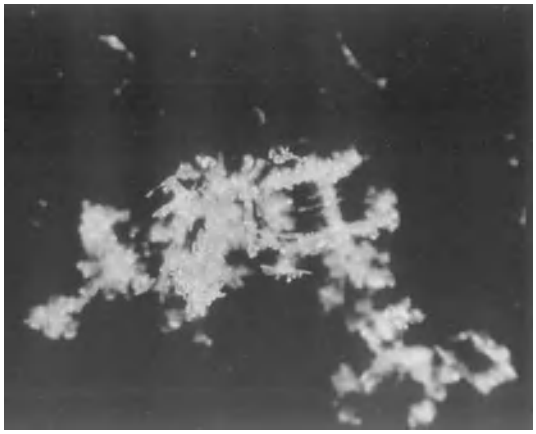

Fig. 7. Snowflake made up of several hundred individual crystals.

Snowflakes. Under some conditions crystals aggregate to give snowflakes. This happens for the dendritic crystals that grow near 5°F (−15°C), which readily interlock if they collide with each other, and for all crystals near 32°F (0°C). Snowflakes typically contain several hundred individual crystals (**Fig. 7**) and fall at 1–2 m s^{-1} (3–6 ft s^{-1}). Casual observation of falling snowflakes often reveals a complex path, the flake oscillating and spinning, which results from the shedding of eddies as the flake falls. The shape of the flake tends to be planar (rather than spherical), which represents molding by aerodynamic forces as the flake falls through the air.

Snowpack. When snow reaches the ground, changes take place in the crystals leading to properties of the snowpack that are important for skiing conditions and avalanche formation. At temperatures near 32°F (0°C) the crystals rapidly lose the delicate structure acquired during growth, sharp edges evaporate, and the crystals take on a rounded shape, some 1–2 mm (0.04–0.08 in.) in diameter. These grains sinter together at their contact points to give snow some

structural rigidity. The specific gravity varies from ~0.05 for freshly fallen "powder" snow to ~0.4 for an old snowpack. At lower temperatures, crystals retain their shape and the snowpack retains a low density, unless compressed at depth. Strong vertical gradients of temperature in the snowpack caused by low air temperatures above the pack lead to an evaporation-growth process in the pack itself. Crystals grow in a skeletal form as hollow columns (hopper crystals; Fig. 2*f*). These forms lead to low density and planes of weakness in the pack, which lead to avalanche release under high loading rates. For this to occur most readily, a strong vertical gradient of temperature (~2.3°F/in. or 0.5°C/cm) must coincide with the 25°F (−4°C) isotherm in the snowpack. A similar effect occurs when frost crystals have grown on the snowpack from an overlying fog near this temperature and are subsequently buried by a snowfall. *See* AVALANCHE.

Snowfall rates tend to be maximum in association with adequate supplies of moisture from a warm winter storm (surface temperature near 32°F or 0°C), and are enhanced when associated with mountain ranges where air rising over the mountain makes additional moisture available. These combinations give typical snowfall rates of 0.5 cm h^{-1} (0.2 in. h^{-1}) water equivalent. Higher rates of ice precipitation usually are associated with soft hail formed by riming of falling snow particles as they collide with these supercooled cloud droplets during their fall. *See* CLOUD PHYSICS; HAIL; PRECIPITATION (METEOROLOGY).

Growth mechanism and crystal shape. Crystals grow in different shapes as a result of interplay of two processes, namely heat-moisture transport between the growing crystal and its environment, and the molecular processes that take place at the crystal-air interface.

At low supersaturations (lower than usually occur for extended periods in the atmosphere) crystalline defects (departures from regular molecular packing) may be necessary for growth of specific crystal faces of high molecular density. By contrast, under most atmospheric conditions growth of faces proceeds by nucleation of layers at corners of a growing crystal which grow back toward the face center. In the case of a plate, as the layer grows out a new one is nucleated and the process repeats; in the case of a dendrite, growing in the presence of higher vapor content or at higher ventilation, layers grow only a short distance before a new one is nucleated so complete faces do not form. The reason for columns or plates is related to the properties of water molecules absorbed on the different faces, giving rise to different rates of nucleation and layer growth. Confirmation of these differences awaits development of laboratory techniques such as ellipsometry for studying the properties of a surface with a high vapor pressure, for such studies cannot make use of high vacuum techniques utilized in studies of similar crystals of technological importance. *See* CRYSTAL GROWTH; INTERFACE OF PHASES.

John Hallett

Bibliography. W. A. Bentley and W. J. Humphreys, *Snow Crystals*, 1931, reprint 1962; J. R. Eagleman, *Meteorology: The Atmosphere in Action*, 2d ed., 1992; J. Hallett, How snow crystals grow, *Amer. Sci.*, 72:582–589, November–December 1984; P. V. Hobbs, *Ice Physics*, 1975; V. Keller and J. Hallett, Influence of air velocity on the habit of ice crystal growth from the vapor, *J. Crys. Growth*, 60(11):91–106, 1982; C. Magono and C. Lee, Meteorological classification of natural snow crystals, *J. Fac. Sci.*, Hokkaido Univ., ser. 7, no. 2, pp. 321–335, 1966; U. Nakaya, *Snow Crystals*, 1954.

Snow line

A term generally used to refer to the elevation of the lower edge of a snow field. In mountainous areas, it is not truly a line but rather an irregular, commonly patchy border zone, the position of which in any one sector has been determined by the amount of snowfall and ablation. These factors may vary considerably from one part to another. In regions where valley glaciers descend to relatively low elevations, the summer snow line on intervening rock ridges and peaks is often much higher than the snow line on the glaciers, and in most instances it is more irregular and indefinite. If by the end of summer it has not disappeared completely from the bedrock surfaces, the lowest limit of retained snow is termed the orographical snow line, because it is primarily controlled by local conditions and topography. On glacier surfaces it is sometimes referred to as the glacier snow line or névé line (the outer limit of retained winter snow cover on a glacier).

Year-to-year variation in the position of the orographical snow line is great. The mean position over many decades, however, is important as a factor in the development of nivation hollows and protalus ramparts in deglaciated cirque beds. The average regional level of the orographical snow line in any one year is the regional snow line. Since the regional snow line is controlled entirely by climate and not influenced by local conditions, glaciologists sometimes call it the climatological snow line. The average position of the climatological snow line over a decade or more may be termed the mean climatological snow line. On broad ice sheets and glaciers, the climatological snow line is the same as the névé line. The average position of this, over a decade or more of observation, may be termed the mean névé line. *See* GLACIOLOGY; SNOWFIELD AND NÉVÉ.

Maynard M. Miller

Snow surveying

A technique for providing an inventory of the total amount of snow covering a drainage basin or a given region. Most of the usable water in western North America originates as mountain snowfall that accumulates during the winter and spring and appears several months later as streamflow. Snow surveys were established to provide an estimate of the snow water equivalent (that is, the depth of water

produced from melting the snow) for use in predicting the volume of spring runoff. They are also extremely useful for flood forecasting, reservoir regulation, determining hydropower requirements, municipal and irrigation water supplies, agricultural productivity, wildlife survival, and building design, and for assessing transportation and recreation conditions.

Conventional snow surveys are made at designated sites, known as snow courses, at regular intervals each year throughout the winter period. A snow course consists of a series of 5–10 sampling points, spaced 50–100 ft (15–30 m) apart, along a permanently marked and mapped route. A snow sampler is used to measure the snow depth and water equivalent at each point. Average depth and water equivalent are calculated for each snow course. There are about 3000 snow courses operated throughout North America.

Accurate estimates of snow water equivalent over a basin can be obtained only after allowances are made for instrument, observation, or siting biases. Snow samplers in common use overmeasure by up to 10%. Accurate basin snow cover estimates of depth, density, and water equivalent are best obtained from a snow course network or survey sampling scheme specifically designed to represent local vegetation and terrain variations. This type of snow survey could differ from a conventional snow course in the number and spacing of sample points.

A viable alternative snow survey procedure is the airborne gamma-ray technique. Gamma rays are emitted by the natural radioactive decay of uranium, thorium, and potassium in the ground. Since water, in either liquid or solid phase, attenuates gamma radiation, snow pack water equivalent along a flight line can be calculated based on the attenuation of the radiation signal by the snowpack compared to the condition during no snow cover (see **illus**.). Corrections for soil moisture and for any contribution of atmospheric radon gas and cosmic rays to the measured signal are required. Typically, the aircraft flies along a 12-mi (20-km) flight line at 500-ft (150-m) altitude measuring the radiation signal over a 1000-ft-wide (300-m) swath. Each line provides a measured areal average snow water equivalent over a 2.3-mi^2 (6-km^2) area. The airborne gamma method has been used operationally in North America, Scandinavia, and Russia, mainly in shallow snowpack regions.

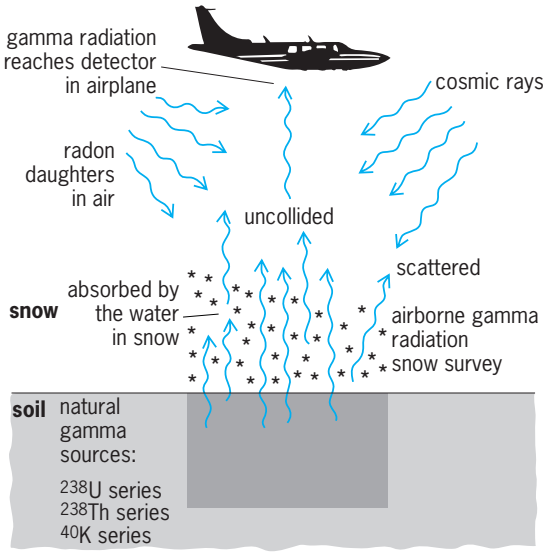

Snow survey using airborne gamma-radiation method. (*Courtesy of T. R. Carroll*)

Satellite remote sensing and data relay are technologies used to obtain information on snow cover in more remote regions. Satellite imagery is used to map snow-covered areas; in the future, passive microwave sensors may provide information on snow depth and water equivalent. Other communications technologies (satellite data retransmission, meteor burst) permit real-time access of measurements by automatic devices, such as snow pillows or snow depth sensors, reducing the need to travel to more remote sites to conduct snow surveys. *See* GAMMA-RAY DETECTORS; REMOTE SENSING; SURFACE WATER.

Barry E. Goodison

Bibliography. T. R. Carroll, J. E. Glynn, and B. E. Goodison, A comparison of U.S. and Canadian airborne snow water equivalent measurements, *Proceedings of the Western Snow Conference*, 51st Annual Meeting, Vancouver, Washington, pp. 27–37, 1983; B. E. Goodison (ed.), Hydrological applications of remote sensing, *Proceedings of the Hamburg Symposium, August 1983*, IAHS Publ. 145, 1985; D. M. Gray and D. H. Male (eds.), *Handbook of Snow*: *Principles, Processes, Management and Use*, 1981; E. L. Peck, T. R. Carroll, and S. C. Van Demark, Operational aerial snow surveying in the United States, *Hydrol. Sci. Bull.*, 25:51–62, 1980.

Snowfield and névé

The term snowfield is usually applied to mountain and glacial regions to refer to an area of snow-covered terrain with definable geographic margins. Where the connotation is very general and without regard to geographical limits, the term snow cover is more appropriate; but glaciology requires more precise terms with respect to snowfield areas. These terms differentiate according to the physical character and age of the snow cover. Technically, a snowfield can embrace only new or old snow (material from the current accumulation year). Anything older is categorized as firn or ice. The word firn is a derivative of the German adjective *fern*, meaning "of last year," and hence refers to hardened snow not yet metamorphosed to ice which has been retained from the preceding year or years. Thus, by definition, a snowfield composed of firn can be called a firn field. Another term familiar to glaciologists is névé, from the French word for a mass of hardened snow in a mountain or glacier environment. In English, rather than a specific word for the material itself, a descriptive phrase is used such as: "consolidated granular snow not yet changed to glacier ice." Because of

the need for simple terms, however, it is acceptable to use the French term *névé* when specifically referring to a geographical area of snowfields on mountain slopes or glaciers (that is, an area covered with perennial "snow" and embracing the entire zone of annually retained accumulation). For reference to the compacted remnant of the snowpack itself, that is, the material which is retained at the end of an annual melting period on a snowfield or névé, it is appropriate to use the derivative German term firn. *See* GLACIATED TERRAIN; GLACIOLOGY.

Maynard M. Miller

Bibliography. D. I. Benn, *Glaciers & Glaciation*, 1998; S. C. Colbeck, *Dynamics of Snow and Ice Masses*, 1980; W. D. Kingery, *Ice and Snow: Properties, Processes and Applications*, 1963.

Soap

A cleansing agent described chemically as an alkali metal salt of a long-carbon-chain monocarboxylic acid, represented for example by sodium myristate ($NaCOOC_{13}H_{27}$), whose structure is shown below.

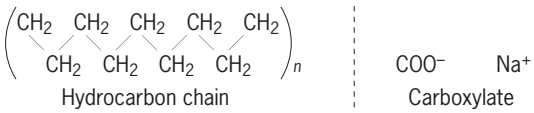

The hydrocarbon portion is hydrophobic and the carboxylate portion is hydrophilic.

For detergency purposes, the most useful hydrophobic portion contains 12–18 carbon atoms. Sodium soaps are not useful for soil removal with hydrocarbon chains of less than 12 carbons. When the chain length exceeds 18 carbons, they become insoluble in water.

Manufacturing process. Historically, soap was made from animal fats and wood ashes (which contain the alkali potassium carbonate). The Romans were doing this 2500 years ago, and a similar process was used in colonial America. In 1791, the Leblanc process for preparing sodium hydroxide was developed, making possible a soap-making process based on an alkaline hydrolysis reaction, saponification, according to the reaction below, where R represents the hydrocarbon chain.

$$\underset{\text{Fat}}{C_3H_5(OOCR)_3} + \underset{\text{Caustic soda}}{3NaOH} \rightarrow \underset{\text{Soap}}{3NaOOCR} + \underset{\text{Glycerin}}{C_3H_5(OH)_3}$$

Until the late 1930s, soap was made in huge kettles into which fats, oils, and caustic were piped and heated until boiling vigorously. After cooling for several days, tons of salt were added, causing the mixture to separate into two layers with the neat soap on top and the lye on the bottom. One kettle could produce carload quantities of soap. The product was pumped into giant mixers called crutchers where builders, perfumes, and other desired ingredients were added. (Builders are alkaline compounds which improve the cleaning performance of the detergent.) Finally, the soap was rolled into flakes, cast or milled into bars, or spray-dried into soap powder.

Most soap is now made by a continuous process. Fats and oils may be converted directly to soap by the reaction with caustic, and the neat soap separated by a series of centrifuges or by countercurrent washing. The saponification may be carried out with heat, pressure, and recirculation. Other important modern processes hydrolyze the fats directly with water and catalysts at high temperatures. This permits fractionation of the fatty acids, which are neutralized to soap in a continuous process. Advantages for this process are close control of the soap concentration, the preparation of soaps of certain chain lengths for specific purposes, and easy recovery of the by-product glycerin. After the soap is recovered, it is pumped to the crutcher and treated just as the product from the kettle process. Tallow and coconut oil are the most common fatty materials used for soapmaking.

Characteristics. Because of the marked imbalance of polarity within the molecules of synthetic detergents and soaps, they have unusual surface and solubility characteristics. The feature of their molecular structure which causes these properties is the location of the hydrophilic function at or near the end of a long hydrocarbon (hydrophobic) chain. One part of the molecule is water-seeking while the other portion is oil-seeking. Thus these molecules concentrate and orient themselves at interfaces, such as an oil-solution interface where the hydrophobic portion enters the oil and the hydrophilic portion stays in the water. Consequently, the interfacial tension is lowered and emulsification can result. At an agitated air-solution interface, the excess detergent concentration leads to sudsing.

Considerable research has also been done on the aqueous solutions of soap. Many of the phase characteristics now known to occur in most surfactant-water systems were first discovered in soap systems. These include micelle formation where, at a critical concentration, soap molecules in solution form clusters (micelles). The hydrocarbon chains associate with each other in the interior, and the polar groups are on the outside. Since the interior of the micelle resembles a hydrocarbon, materials such as oil-soluble dyes, which are insoluble in water, will frequently dissolve in a detergent solution containing micelles. At higher concentrations of soap, two different liquid crystalline phases form. Around 50% water–50% soap "middle" phase occurs, and is believed to form when soap molecules coalesce into long chains or rods in a hexagonal array. It is a very sticky, viscous material which must be avoided in the manufacture of soap. At still higher concentrations (30% water–70% soap) neat phase occurs where the soap molecules form extensive bilayer sheets. Although viscous, neat soap can be pumped and is the phase in which soap is recovered after the saponification. *See* MICELLE; SURFACTANT.

There is one serious disadvantage to the general use of soap which has caused its replacement by synthetic detergents for home laundering. The problem is that carboxylate ions of the soap react

with the calcium and magnesium ions in natural hard water to form insoluble materials called lime soap. In areas having a hard water supply, this is the floating curd which appears in the washbowl and as bathtub ring. In laundry products, attempts were made to overcome the problem by adding materials such as sodium carbonate to "build" the action of soap by precipitating calcium and magnesium carbonates. However, both the carbonates and the soap curd deposited on the fabrics being washed, and gradually caused them to become dull, gray, and sometimes stiff.

Uses. Bar soap for personal bathing is the greatest remaining market for soap in the United States, Canada, and western Europe. However, in the developing countries, the detergent industry is still dominated by soap. Some commercial laundries having soft water continue to use soap powders. Metallic soaps are alkaline-earth or heavy-metal long-chain carboxylates which are insoluble in water but soluble in nonaqueous solvents. They are used as additives to lubricating oils, in greases, as rust inhibitors, and in jellied fuels. *See* DETERGENT.

Roy C. Mast

Bibliography. *Kirk-Othmer Encyclopedia of Chemical Technology*, 3d ed., vol. 21, 1983; P. L. Layman, Brisk detergent activity changes picture for chemical suppliers, *Chem. Eng. News.*, 62(4):40, 1984; W. C. Preston, Some correlating principles of detergent action, *J. Phys. Colloid. Chem.*, 52:84, 1948; E. G. Thomssen and J. W. McCutcheon, *Soaps and Detergents*, 1949.

Soapstone

A soft talc-rich rock. Soapstones are rocks composed of serpentine, talc, and carbonates (magnesite, dolomite, or calcite). They represent original peridotites which were altered at low temperatures by hydrothermal solutions containing silicon dioxide, SiO_2; carbon dioxide, CO_2; and other dissolved materials (products of low-grade metasomatism). Among the rock products thus formed are antigorite schists, actinolite-talc schists, and talc-carbonate rocks. To the last belongs the true soapstone, but the whole group of rocks may loosely be referred to as soapstones because of their soft, soapy consistency. Such rocks were selected by prehistoric humans for making primitive vessels and pots, and for making rough carvings for ornamental purposes. *See* TALC.

T. F. W. Barth

Social hierarchy

A fundamental aspect of social organization that is established by fighting or display behavior and results in a ranking of the animals in a group. Social, or dominance, hierarchies are observed in many different animals, including insects, crustaceans, mammals, and birds. In many species, size, age, or sex determines dominance rank. Dominance hierarchies often determine first or best access to food, social interactions, or mating within animal groups.

Dominance and agonistic behavior. When two animals fight, several different behavioral patterns can be observed. Aggressive acts (such as biting, kicking, or stabbing with horns or antlers) and submissive acts (such as a dog rolling on its back and exposing its stomach) are both parts of a fight. Aggression and submission, together, are known as agonistic behavior.

Agonistic encounters often result in a winner, or dominant animal, and a loser. The losing individual may simply retreat or leave the area, so that it is not part of the winner's social group. If the loser stays in the vicinity of the winner, it usually displays submissive behavior when the winner approaches in order to prevent further attacks. The submissive individual is subordinate in the relationship.

An agonistic relationship in which one animal is dominant and the other is submissive is the simplest type of dominance hierarchy. In nature, most hierarchies involve more than two animals and are composed of paired dominant-subordinate relationships. For example, if a young female bighorn sheep enters a herd of other females, she will have agonistic encounters with other females and will establish dominant-subordinate relationships with the other herd members that determine her rank in the group.

The simplest dominance hierarchies are linear and are known as pecking orders. In such a hierarchy the top individual (alpha) dominates all others. The second-ranked individual (beta) is submissive to the dominant alpha but dominates the remaining animals. The third animal (gamma) is submissive to alpha and beta but dominates all others. This pattern is repeated down to the lowest animal in the hierarchy, which cannot dominate any other group member. Linear hierarchies can be observed in chickens.

Other types of hierarchies result from variations in these patterns. If alpha dominates beta, beta dominates gamma, but gamma dominates alpha, a dominance loop is formed. In some species a single individual dominates all members of the social group, but no consistent relationships are formed among the other animals. In newly formed hierarchies, loops or other nonlinear relationships are common, but these are often resolved over time so that a stable linear hierarchy is eventually observed.

Hierarchies in social insects. In some social wasps and bees the queen has agonistic relationships with her workers, and a linear dominance hierarchy is established. In paper wasps (*Polistes* sp.) the queen lunges at her workers and opens her mandibles in a threatening manner; in extreme cases, biting, grappling, and stinging occur. The workers submit by crouching down to the comb. The dominance hierarchy assures that the queen lays most of the eggs in the colony, while the workers build comb and bring food to the queen and her offspring. Males are not a part of the societies of stinging social insects (ants, bees, and wasps) and do not participate in these hierarchies.

In some of the more advanced social insects, such as the honeybee, the queen uses chemical signals to dominate her workers, rather than aggressive behavior. If a honeybee queen dies or is removed from the colony, the workers are able to lay eggs; fights and agonistic behavior over dominance then occur among the workers. *See* CHEMICAL ECOLOGY; SOCIAL INSECTS.

Dominance and mating. Males often fight over access to females and to mating with them. Male dominance hierarchies are seen in many hooved mammals (ungulates). Horns or antlers are usually adaptations used in agonistic encounters among males, and body size is often the determinant of the dominant male. Males may come together in bachelor herds during the nonmating season; when the females are ready to mate, agonistic behavior among the males intensifies and the dominant male is determined. *See* REPRODUCTIVE BEHAVIOR.

Dominance and feeding. Herds of females use dominance hierarchies to determine access to food. A female occupying a favored feeding site will yield to a dominant female but will displace a subordinate female. Agonistic interactions among females are often not as overtly aggressive as those among males, but the effects of the dominance hierarchy can easily be observed by recording which female is able to displace which others. In female dairy cattle, the order of entry into the milking barn is determined by dominance hierarchy, with the alpha female entering first.

Dominance hierarchies occur among very young siblings in many species; dominance determines which nipple is available for suckling. Piglets prefer the foremost nipples, which produce the most milk. Piglets relegated to the hindmost nipples receive less food and grow more slowly. Similar nipple-order dominance appears among domestic kittens.

Genetic factors. Because dominant animals may have advantages in activities such as feeding and mating, they will have more offspring than subordinate animals. If this is the case, then natural selection will favor genes for enhanced fighting ability; a striking example is seen in males with armament used in fighting for access to mates. Large horns or antlers will be be favored by natural selection, as will ferocity in fighting. In some cases, such selection results in seemingly bizarre or disproportionate weapons. However, practicalities, such as ability to move in the environment or to evade predators, may place a limit on such selection. Heightened aggressive behavior may be counterselected by the necessity for amicable social interactions in certain circumstances, such as parent-offspring relationships.

Artificial selection has been used by humans to raise (pit bull terriers, fighting cocks) or to lower (most breeds of dogs, horses, cattle) fighting tendencies and dominance behavior. Domesticated animals are generally easier to handle and manipulate if they are less likely to fight and injure one another or human handlers. They may also be kept under more crowded conditions if fighting tendencies are reduced. Selection for passivity toward humans usually results in relaxation of dominance within animal social groups. *See* BREEDING (ANIMAL).

Complex social hierarchies. Many higher primates live in large groups of mixed sex and exhibit complex social hierarchies. In these groups, intra- and intersexual dominance relationships determine many aspects of group life, including feeding, grooming, sleeping sites, and mating. In rhesus macaques the rank of an individual's mother is very important in determining rank. The cultural and genetic history of a macaque may be more important in determining social rank than its body size. Macaque, baboon, and chimpanzee societies are characterized by cooperative alliances among individuals that are more important than individual fighting ability in maintaining rank.

Significance of hierarchies. Social hierarchies provide a means by which animals can live in groups and exploit resources in an orderly manner. In particular, food can be distributed among group members with little ongoing conflict. Although social hierarchies can be viewed as a group characteristic, they have evolved as a result of the effects of natural selection on individual animals. The dominant animals gain first access to food, mates, or other important environmental features without needing to drive all other members of the species away; the selective benefits to the dominant are obvious. Subordinate animals reduce their risk of injury in combat by submitting, but by doing so they also give up priority of access to resources. However, subordinates are close to resources that may be taken if there is a surplus and, perhaps most important, if the alpha animal dies they may compete for its position.

Another motivation for group living is mutual defense. Even though subordinates receive less food or have fewer opportunities to mate, they may have greatly increased chances of escaping predation. This is particularly striking in herds of ungulates, such as antelopes, in which a few animals are always in a vigilant posture (head up, looking for predators) while the others feed. Thus social hierarchies make group living orderly, allocate resources, and encourage mutual defense; the peaceful coexistence of competing individuals is facilitated by such hierarchies. *See* BEHAVIORAL ECOLOGY; POPULATION ECOLOGY; SOCIAL MAMMALS; TERRITORIALITY. Michael D. Breed

Bibliography. J. Archer, *The Behavioural Biology of Aggression*, 1988; I. Eibl-Eibesfeldt, *Human Ethology*, 1989; J. L. Gould, *Ethology*, 1982; D. H. Morse, *Behavioral Mechanisms in Ecology*, 1990; E. O. Wilson, *Sociobiology: The New Synthesis*, 1990.

Social insects

Insects that share resources and reproduce cooperatively. The shared resources are shelter, defense, and food (collection or production). After a period of population growth, the insects reproduce in several ways.

As social insect groups grow, they evolve more differentiation between members but reintegrate into a

more closely organized system known as eusocial. These are the most advanced societies with individual polymorphism, and they contain insects of various ages, sizes, and shapes. All the eusocial insects are included in the orders Isoptera (termites) and Hymenoptera (wasps, bees, and ants). *See* HYMENOPTERA; INSECTA; ISOPTERA; POLYMORPHISM (GENETICS).

Polymorphism. The social insects have evolved in various patterns.

Termites. The most primitive termites live in small patches of dead wood. After incubation of the fertilized clutch of eggs, the larvae that hatch are fed on glandular food at first and later, after several molts, are given semidigested food from the anus of the adults. In this way they obtain important intestinal protozoa that help them digest wood. Generally, the young do not develop wings and gonads, and most become workers in the nest. A few may molt twice (**Fig. 1**) and form a wingless adult with a large head that defends the others; these are known as soldiers. For some time after, only workers are formed, as the reproductive pair inhibit the formation of their own kind, and so do the soldiers. This control is achieved because development inhibitors, specific for each of the castes inhibited, are secreted into the common food stream. The inhibitors are hormonal in nature and are sex- and soldier-specific.

The most advanced termites have evolved distinct lines of development from egg stage: the reproductive line can be identified from the existence of individuals with wing rudiments (nymphs), but their metamorphosis into sexual adults depends on the season as well as on whether the pair of reproductives withhold the development inhibitors. The asexual line consists of termites of both sexes in various sizes; they can molt into either workers or soldiers that are adult only in that they are developmentally fixed. Usually, soldiers are formed from one sex, which is specific for the species. The ratio of soldiers to workers is also a characteristic of the species; one family of termites no longer produces soldiers at all. Polymorphism in higher termites is thus a complex process that is responsive to both social and environmental conditions.

In the Hymenoptera, the society is composed of only females; males are produced periodically for their sperm. They usually congregate and attract females, or they visit colonies with virgin females and copulate there. In the Hymenoptera, sex is determined largely by whether the individual has one or two sets of chromosomes. Thus the queen has the power to determine the sex of her offspring: if she lets any of her stored sperm reach the egg, a female is produced; if not, a male results. *See* SEX DETERMINATION.

Bees and wasps. In the more primitive bees and wasps, social role (caste) is influenced by interaction with like but not necessarily related individuals. The female that can dominate the others assumes the role of queen, even if only temporarily. Domination is achieved by aggression, real or feigned, or merely by a ritual that is followed by some form of salutation by the subordinates. This inhibits the yolk-stimulating glands and prevents the subordinates from contributing to egg production; if it fails to work, the queen tries to destroy any eggs that are laid. Subordinate females take on more and more of the work of the group for as long as the queen is present and well. At first, all the eggs are fertilized and females develop, with the result that virgin females inhabit the nest for the first batches. They are often undernourished, and this, together with their infertility, reduces their urge to leave the nest and start another one. Such workers are said to be produced by maternal manipulation.

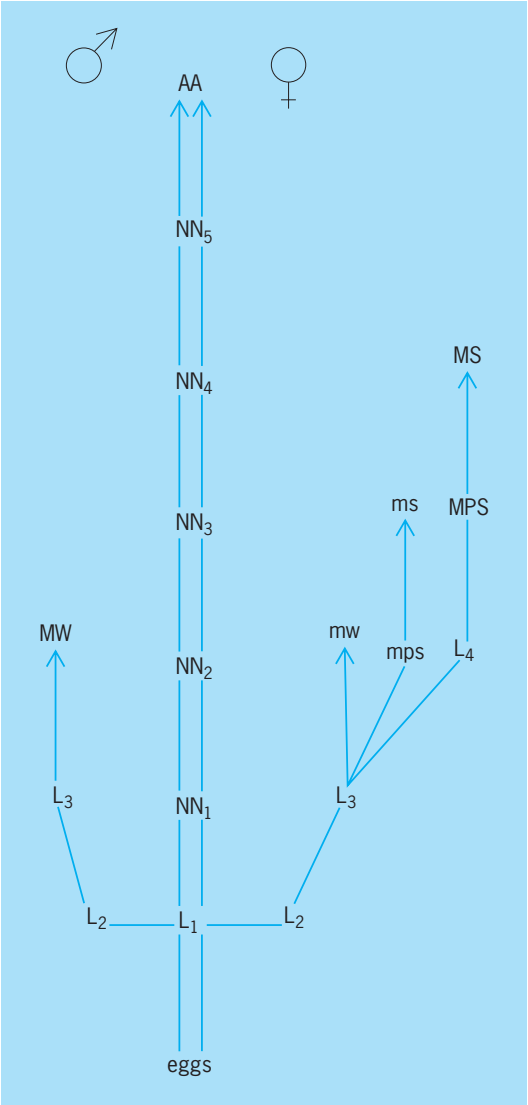

Fig. 1. Scheme of postembryonic development in mature colonies of *Macrotermes*, an advanced African termite that lives in savanna. Larvae (stages L_1–L_4) develop into large (MW) and small (mw) workers, and large (MPS) and small (mps) presoldiers, which become large (MS) and small (ms) soldiers, respectively; soldiers are female only. Sexual nymphs (N_1–N_5) and sexual adults (AA) develop only from larval stage L_1. (*After B. M. Okot-Kotber, Instars and polymorphism of castes Macrotermes michaelseni, Insectes Sociaux, 28(3):233–246, 1981*)

Though many more advanced societies of wasps are often perennial, the common wasps of temperate

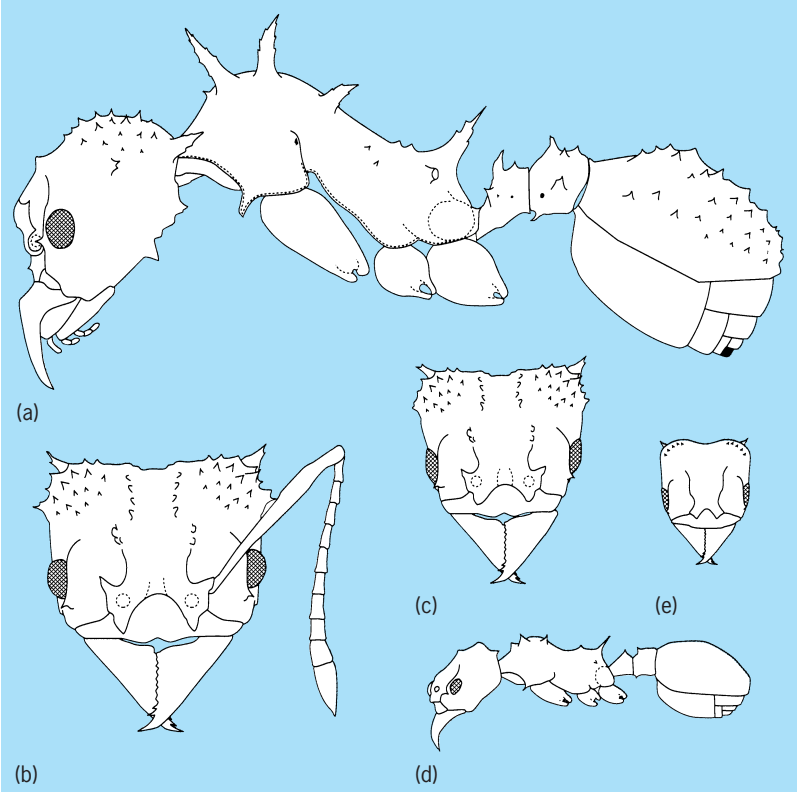

Fig. 2. Heads of the polymorphic *Acromyrmex octospinosus*, an ant that cultures fungus in Trinidad: (*a*) profile; and (*b*) front view of major worker; (*c*) front view of media worker; (*d*) profile; and (*e*) front view of minor worker. (*After W. M. Wheeler, Mosaics and Other Anomalies among Ants, Harvard University Press, 1937*)

regions have a single queen that is distinctly larger than her workers, having been reared in larger cells and given more and better food. If such a queen is removed, the society becomes disorganized, and the workers fight and lay their own unfertilized eggs, giving rise to males. Thus the current queen is not replaced, and the society dies out.

Stingless bees also have queens and workers of unequal size. The queens are fed more than the workers, and when these virgin queens mature they are held in their cells and fed until the current queen fails in some way, probably as an egg layer or a pheromone producer. When this time comes, a virgin is let out onto the comb and assessed in relation to the mother queen. With this type of system, there is little change in the society if the queen is removed. With the honeybee, however, removal of the queen sets up a series of reactions leading to her replacement by a daughter.

This replacement activity is a special case of the normal replacement that takes place in spring. It starts with the formation of queen cells on the lower rims of the central combs in which the mother queen lays fertile eggs. The larva that hatches from this type of egg is fed throughout its life on a jelly from the mouth glands of the workers (worker-cell larvae get jelly for only 3 days). Not only does this special diet yield an adult queen that is larger than the worker, and with a bigger ovary, but she also differs considerably in behavior and lives years rather than months. Queens leave the nest soon after emerging from their cell and search for males; they mate with 10 or more males. After 2 or more years they leave the nest in a swarm of thousands of workers, which amounts to about two-thirds of the total population, to establish a new colony. Workers generally stay in the nest to clean, make cells, and feed larvae until they are strong enough to forage for nectar and pollen. The queen normally stops the reproductive activities of workers by means of pheromones that affect worker behavior and are passed from one to the other by contact, not in the food stream. *See* PHEROMONE.

Ants. Reproductive ants, like termites, engage in a massive nuptial flight, after which the females, replete with sperm, go off to start a new nest. At some stage after the nuptials, the reproductives break off their wings, which have no further use. Workers, however, never have wings because they develop quickly and pass right through the wing-forming stages; their ovaries and genitalia are also reduced. Ant queens can prevent the formation of more queens; as with the honeybee, they do this behaviorally by using pheromones. They also force the workers to feed all larvae the same diet. To this trophogenic caste control is added a blastogenic control; eggs that are laid have a developmental bias toward one caste or another. This is not genetic; bias is affected by the age of the queen and the season: more worker-biased eggs are laid by young queens and by queens in spring. In some ants, workers mature in various sizes (**Fig. 2**). Since they have dis-

Fig. 3. Thorns of a species of *Acacia*, showing the cavity in which the ant *Pseudomyrmex* nests after entering by a hole near the tip. (*After A. J. Beattie, The Evolutionary Ecology of Ant-Plant Mutualisms, Cambridge University Press, 1985*)

proportionately large heads, the biggest workers are used mainly for defense; they also help with jobs that call for strength, like cutting vegetation or cracking nuts.

Nest construction. Social insects make remarkable nests that protect the brood as well as regulate the microclimate. The simplest nests are cavities dug in soil or soft wood, with walls smoothed and plastered with feces that set hard. Chambers at different levels in the soil are frequently connected by vertical shafts so that the inhabitants can choose the chamber with the best microclimate.

Termites and ants also make many different types of arboreal nests. These nests are usually made of fecal material, but one species of ant (*Oecophylla*) binds leaves together with silk produced in the salivary glands of their larvae that the workers hold in their jaws and spin across leaves. A whole group of ants (for example, *Pseudomyrmex*) inhabit the pith of plants (**Fig. 3**).

Termite mounds such as those of *Macrotermes* (**Fig. 4**) that form conspicuous features in savanna all start as a hole in the ground excavated by a male and a female. They grow to include various types of chambers for adults, juveniles, and the symbiotic fungus, as well as ventilation shafts and cellars that go down to water level. The large mounds of dry leaves made by the wood ant (*Formica*) in temperate forests function to shed water and collect solar radiation. Galleries may go deep into the subsoil and are used by the workers for hibernation.

Each species of wasp cuts a characteristic type of wood and chews it into a soft pulp that it thins out to the consistency of paper and applies gradually to form the nest. Combs of great strength and precise design are constructed in this way, arranged in tiers

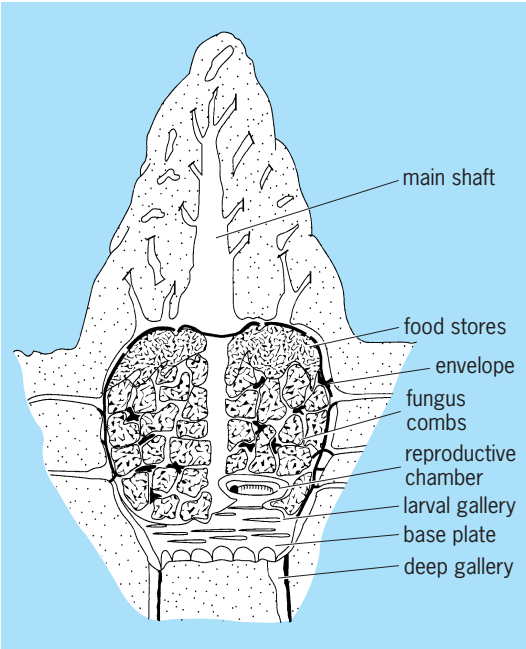

Fig. 4. Section of a *Macrotermes* nest in Nigeria about 18 ft (5.5 m) high and 9–12 ft (2.7–3.7 m) in basal diameter. (*After N. M. Collins, The nests of Macrotermes bellicosus, from Mockwa, Nigeria, Insectes Sociaux, 26(3):240–246, 1979*)

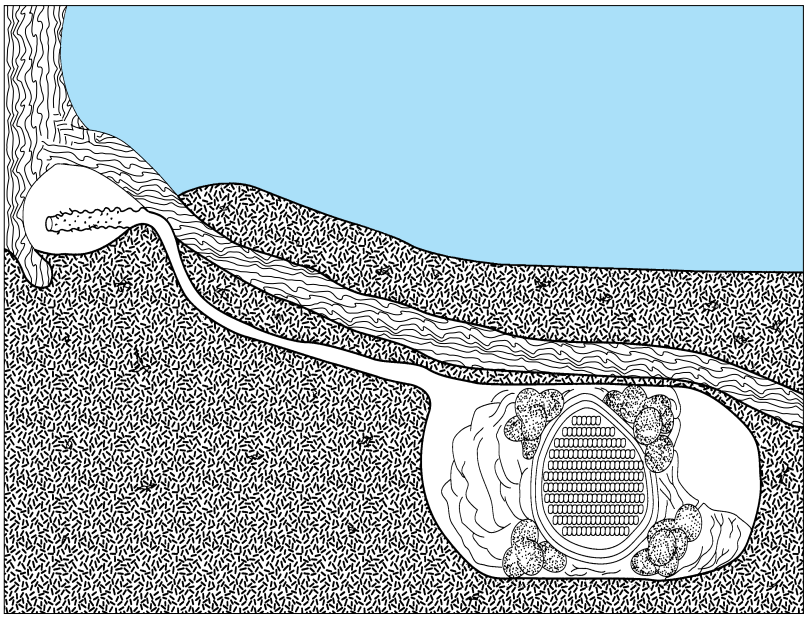

Fig. 5. Section of the nest of a stingless bee (*Trigona dorsalis*) in soil under a tree root. Twelve combs of brood cells are surrounded by a wax envelope, outside of which are many large storage cells supported by struts; the tube to the outside opens under the tree. (*After A. Wille and C. D. Michener, The nest architecture of stingless bees with special reference to those of Costa Rica, Revista de Biologia Tropical, 21(suplemento 1): 1–278, Universidad de Costa Rica, 1973*)

and surrounded by a paper envelope that is heat-refractory as well as impermeable to other insects. The cells are packed tightly in a hexagonal formation and may, in advanced wasps like *Vespula*, be large for reproductives and small for workers. Such a strong, light nest can be suspended from a branch by struts that are smeared with a chemical to repel ants.

Social bees use wax secreted by their cuticular glands and frequently blended with gums from tree exudates for their nest construction. Cells are made cooperatively by a curtain of young bees that scrape wax from their abdomen, chew it with saliva, and mold it into the correct shape; later it is planed and polished. Some primitive species, such as *Bombus*, make combs with spherical cells that are used first for brooding and later for storage of pollen and honey. Stingless bees (for example, *Trigona*) often stratify groups of cells and link cell layers with struts. Such a brood zone may be surrounded with wax envelopes, and these in turn with large globular storage cells (**Fig. 5**). The whole nest is usually in a cavity and is connected to the open air by a long tube. With honeybees the hexagonal comb reaches perfection as a set of back-to-back cells, each sloping slightly upward to prevent honey from running out. The same cells are used repeatedly for brood and for storage; or they may be made a size larger for rearing males. Only the queen cell is pendant, with a circular cross section and an opening below.

Symbiosis with other organisms. The ubiquity and ecological power of social insects depend as much on their ability to evolve mutualistic relations with other organisms as on the coherence of their social organization. Wood and the cellulose it contains is normally available as a source of energy only to

bacteria and fungi. However, it is used as a basic resource by both termites and ants that have evolved a technique of culturing these organisms. Though lower termites have unusual protozoa as intestinal symbionts, higher termites have bacteria in pouches in their hind gut. Many have a fungus that they culture in special chambers in their nests (Fig. 4). The termites feed on woody debris, leaves, and grass cuttings; the fungus digests these materials with the aid of termite feces and produces soft protein-rich bodies that the termites share with their juveniles and reproductives, neither of which are able to feed themselves. Protected from the weather, the fungus can remain active throughout the dry seasons—an inestimable advantage in the subtropics.

The ant–fungus symbiosis is similar but has evolved only in the central American region. The ants carry leaves to the nest to feed their fungus. Leaves from a great many different species are cut, and a distinct preference is shown. In the nest, the waxes are first rubbed off, and the oils, extracted by chewing, are used as a source of energy for the workers. The paste that results is treated with ant feces and enzymes from the fungus to degrade the material and convert it into an easily digested protein that is fed to the larvae and queen. Apart from producing a digestible, energy-rich material for the ant, the fungus extracts and destroys many toxic substances that are unpalatable to ants.

Many ants collect and store seeds that they mash and feed directly to their larvae. Provided they are collected when dry and stored in well-ventilated chambers, these seeds can remain viable and edible for an entire season. The plants benefit because not all the seeds are eaten; some that start to germinate are thrown out with the rubbish of the colony—in effect a way of planting them. Others are left behind by the ants when they change nests, as they often do when they suffer shading by encroaching vegetation. In this way, grass seeds can extend their range into dry areas that they could not reach alone. Some plants produce a hard shiny seed that the ants cannot grip and, to make it attractive, they add an aril, which is a growth that contains attractive foods. The ants take the seed back to their nest and eat the aril without harming the seed, which they later throw away or just leave buried in the soil of their nest. In woodland, most ants choose to live in sunny spots that suit the plant as well.

The dispersal of plant pollen by bees is a well-known symbiosis, and it has led to the evolution of many strange shapes, colors, and scents in flowers. Quite specific flower–bee relationships may exist in which one plant may use very few species of bees for the transfer of pollen. However, it is possible that this relationship may be breaking down, since many species of bee can find the nectar lure directly and cut through the plant tissue without pollinating the plant at all. Also, many bees collect pollen from flowers that do not have nectar; they want the pollen itself, which is a very good food. Highly successful species of stingless bees and honeybees exhibit behavior that attracts other bees of the same nest to newly discovered open flowers. *See* POLLINATION; POPULATION ECOLOGY.

M. V. Brian

Bibliography. M. D. Breed (ed.), *The Biology of Social Insects*, 1982; M. V. Brian, *Social Insects: Ecology and Behavioral Biology*, 1983; W. Engels (ed.), *Social Insects*, 1990; B. Holldobler and E. O. Wilson, *The Ants*, 1990; C. D. Michener, *The Social Behaviour of the Bees*, 1990; T. D. Seeley, *Honeybee Ecology: A Study of Adaptation in Social Life*, 1985; E. O. Wilson, *The Insect Societies*, 1990.

Social mammals

Mammals that exhibit social behavior. This may be defined as any behavior stimulated by or acting upon another animal of the same species. In this broad sense, almost any animal which is capable of behavior is to some degree social. Even those animals which are completely sedentary, such as adult sponges and sea squirts, have a tendency to live in colonies and are social to that extent. Social reactions are occasionally given by species other than the animals' own; an example would be the relations between domestic animals and humans.

Mammals are warm-blooded vertebrates which typically live on the land surface of the globe. Several orders, however, have semiaquatic forms, such as muskrats and beavers among rodents, and otters and seals among carnivores. One entire order, the whales, has become completely water-living, and another, the bats, has developed the power of active flight and has taken to the air.

A common characteristic of the entire group is the mammary gland, used for the feeding of the newborn young. Because this is developed only in females, there is a general tendency for a differentiation of labor in which the females care for the young and the males specialize in fighting behavior. In the subclass Marsupialia, the nipples are enclosed in a brood pouch, in which the young spend most of their early life. The early ingestive behavior of mammals, or nursing, thus has a high degree of social significance.

In addition to their great variety of habitats, mammals eat a great variety of foods, and occupy many different ecological niches. There are almost as many mammals which are nocturnal as are active chiefly during the day.

Behavioral capacities. Depending on its habitat, a given species may use one or more of the sense organs more than the others. Some nocturnal rodents such as mice and rats use the tactile and auditory senses a great deal more than the eyes, whereas a diurnal tree-living animal such as a squirrel may chiefly use its eyes. All major sense organs are well developed in each species.

Mammals show a great variety of motor capacities. The limbs can be highly developed for digging as they are in moles and badgers, or used almost entirely for locomotion as they are in the hoofed mammals. An arboreal primate such as the chimpanzee has all four limbs developed for manipulation and prehension.

Social behavior and relationships in mammals

Relationship	Type of social behavior	Typical combinations formed
Care-dependency	Care giving Care soliciting Ingestive (especially nursing) Shelter seeking Eliminative (in some species)	Adult-young (especially female-young)
Sexual	Sexual Investigatory Eliminative (in some species)	Male-female
Dominance-subordination	Agonistic	Male-male (but also in all other combinations)
Leader-follower	Allelomimetic	Female-young (usually weaker in other combinations)
Mutual care	Care giving (grooming)	Adult-adult
Mutual defense	Agonistic	Male-male

Even the tail can be used for prehension by some of the monkeys. At the other extreme, the hindlimbs of whales are externally invisible and the forelimbs are used only as flippers.

Mammals have relatively large and complex brains, with a large development of the cerebral cortex. There is also a considerable capacity for behavioral adaptation through learning. Mammals exhibit relatively few instances of elaborate patterns of behavior which are completely fixed by heredity. In general, mammals are highly adaptable, both from the evolutionary viewpoint and in their capacity to adjust to daily changes in the environment (see **table**).

Social behavior of the American elk. This is one of the largest living species of the deer family, standing as high as a horse. The eyes, ears, and nose are all well developed. The males grow large antlers which are used chiefly in the rutting season and are shed in the early spring. Elk were originally plains-living animals, but are now restricted by hunting to the open areas in the Rocky Mountains. They are highly social animals and live in large herds.

Social life is closely related to the seasons. In the autumn the male herds break up. Each male attempts to round up a herd of females with whom he mates as each comes into estrus. If another male appears, the two begin a furious pushing contest, ending when one is exhausted and driven off. *See* ESTRUS.

As winter comes on, the males and females separate into different herds and migrate to lower altitudes. Hundreds of individuals may move together, producing a migration trail which looks almost as if it were a roadway.

In the spring there is a return migration to the higher altitudes. Each band separates from the large winter herd and goes back to the locality from which it came. Males go higher than females, which stay together and bear their calves. As is true with many members of the deer family, the young calves will "freeze," or stay quiet while the mothers leave to graze or browse. Usually at least one female stays in close proximity to the young animals. The calves at first obtain most of their nourishment by nursing from their mothers. When alarmed, they call to the mother. They grow rapidly during the summer and follow their mothers throughout the first year. The seasonal round of behavior begins again with the rutting season in the autumn.

Types of behavior. The most highly developed type of social behavior is allelomimetic, in which each animal does the same things as those nearby, all responding to each other. This results in the formation and continued existence of a herd. Also of importance in these herbivorous animals is ingestive behavior. Adults spend a great deal of time browsing and grazing. A more highly social sort of ingestive behavior is the nursing of the calves. Related to this is the caregiving or epimeletic behavior of the mothers and the care-soliciting behavior of the young. Agonistic behavior, consisting of fighting and escape behavior, is common in the males during the rutting season. As is true of most herd animals, fighting is individual and groups never combine against one animal. Sexual behavior occurs at the same time, but only for a brief period with each female. Investigative behavior is common, and the animals frequently raise their heads to look around while engaged in grazing or other activities. There is no special pattern of eliminative behavior and relatively little social shelter seeking.

Social relationships. Social behavior results in the formation of social relationships. This is partly based on biological differentiation into three types, males, females, and young, each with its characteristic behavior. There are six possible combinations between them, and each kind of social behavior may result in the formation of a social relationship within a particular combination.

The sexual relationship developed between males and females is a brief one, lasting only as long as the male is able to maintain his dominant position in relation to a herd of females. It has no importance at other times of the year, although it may play some part in holding the male group together.

Both male-female and female-female relationships can involve dominance and subordination. However, as is the case in many grazing animals, whose food is scattered in such a way as to prevent competition, dominance has little importance except in the breeding season. Even here a male may be able to hold a dominant position only for a portion of the season.

In female-young associations the care-dependency relationship is important and long-lasting. As is the case in other herd animals, females will allow only their own young to nurse, and a close association

is built up throughout the nursing period. This relationship lasts into later life and forms the foundation of leadership within the herds. The young born in the same year and within the same herd also form a lasting relationship with each other. This group is particularly important for the males when they separate from the females as adults. In the female herds a leader-follower relationship develops, first between mother and offspring and eventually between older and younger females. This basically allelomimetic relationship is one of the two most important social relationships developed in an elk society.

Unlike the situation among many birds and fishes, the migration trails are apparently learned from one generation to the next and have been altered within historical times. The animals become localized, that is, they become attached to the area in which they were brought up, but there is no defense of territory. Strange herds mingle with each other without fighting, and agonistic behavior is important only in the rutting season.

Other ungulates. The social behavior of other even-toed ungulates follows a pattern similar to that of the elk. The red deer of Scotland are quite similar to them, and even more distantly related animals such as goats, sheep, and bison have the same general characteristics. There is much emphasis on ingestive behavior, because enormous quantities of vegetation must be eaten. A typical grazing animal follows a regular cycle of grazing and stops to chew its cud every few hours. Among the more social species this is always done in concert with others, so that allelomimetic behavior is also highly important. Under natural conditions the result is the enormous herds of ungulates which were once seen on the plains of North America and on the plains of South Africa.

The predominant social relationship of the even-toed ungulates is the mother-young relationship. Its formation has been studied in detail in sheep. The mother sheep will allow only its own lamb to nurse and will drive all others away. The behavior of the mother toward its own lamb is fixed within the first 4 h after birth. The result of this close and definite relationship is that the leadership is developed by females, and social inheritance tends to be matrilineal.

Social behavior of wolves. These carnivorous animals are especially interesting because of their wide distribution and their close association with humans in fact and legend. The first domestic animal, the dog, was derived from them and has almost identical patterns of social behavior.

Under ordinary conditions the food supply of a carnivore is not regularly available in any one area. Consequently the chief activity of wolves is hunting, a form of investigatory behavior. Much of this is done at night. Having spent the day in or near the den, the wolf pack assembles toward evening and hunts far away, usually returning by morning. The hunting range may be 20–50 mi (30–80 km) across. In their hunting activities, they show a great deal of allelomimetic behavior, both in searching for and pulling down game, but there is little evidence of consistent leadership by one individual.

Mating and care of young. Mating behavior takes place in January or February. The entire estrous cycle of a female extends for 6 weeks or more, during which time there is long preliminary courtship as well as actual mating. A peculiarity of sexual behavior is the sexual tie, in which the male and female remain coupled for as long as $^1/_2$ h. It is not known whether the female mates with more than one male, but in related animals which do not run in packs, such as foxes and coyotes, a female typically mates with one male, and the two form a lasting association.

The cubs are born in early spring and are constantly attended by the mother during the first few days of life. As they grow older the mother must leave them to feed and hunt. There is usually a good supply of food at the den, either carried back or vomited by the adults. The cubs begin to eat freshly vomited food at about 3 weeks of age and are completely weaned at 6 or 7 weeks. However, they are still unable to hunt and remain dependent on the adults until approximately 6 months of age.

Thus the litter is in constant association from birth, but parents and other adults leave them for long periods each day. The closest social relationships are made with litter mates, and this is the foundation for the adult pack. Some young wolves are known to remain in the pack in which they were raised, but there is no clear evidence as to how new packs are formed. It is possible that in some cases a group of litter mates may leave an old pack to form one of their own. Most wolf packs consist of five or six adults plus any young animals running with the pack but not old enough to leave it.

Dominance. Agonistic behavior within the pack is organized into a definite dominance order. Because wolves occupy such large ranges, there is no possibility of guarding the boundaries as territories, although some packs regularly hunt over circular runways. Wolves set up scent posts which are regularly visited and at which the animals urinate and defecate. It is possible that in this way a wolf can tell whether a range is occupied and so avoid it. The den area is a definite territory; a strange wolf coming to the den is attacked and driven off.

Wolves are relatively long-lived animals, and social relationships, including the sexual one, are quite persistent. The allelomimetic relationship between litter mates is a most important one, and also important is the dominance-subordination relationship. The relationships between parents and offspring are relatively unimportant.

Social behavior of cat family. By contrast, members of the cat family tend to emphasize the mother-offspring relationship. This is associated with a different method of hunting: lying in wait for the prey and pursuing it for a short distance rather than ranging widely. In addition, cats are less omnivorous and more dependent on meat than members of the dog family. The young develop rapidly and are assisted in

their first hunting by the mother. With the exception of lions and cheetahs, cats tend to be solitary in their hunting activities and show very little allelomimetic behavior.

Social behavior of seals. Alaskan seals show highly organized social behavior upon their breeding grounds. The males take up territories along the shore in the late spring and defend them against males and other mammals. As females arrive, the males guard them as a harem. A few days after arrival the females produce their pups and shortly come into estrus, mating only once. Females leave their pups in groups while they swim off to feed, but males do not feed in this period. Thus two social relationships are prominent on the breeding grounds: the dominance relationship of the harem males, and the mother-offspring relationship. The sexual relationship has little social importance. Females do not breed until 4 or 5 years of age, and males are not able to maintain harems till about 10 years. Large groups of immature males stay outside the territories of the harem males. After the young are mature enough to leave the breeding grounds, the seal herds break up and become almost entirely aquatic until the new breeding season.

Social behavior of carnivores. In general, the social behavior of carnivores is related to their food habits. Investigatory behavior is always prominent. Possibly because of this, carnivores have a strong reputation for intelligence. Some are strongly allelomimetic, but many are almost solitary in their activities. The care of the young is greatly emphasized, because successful hunting depends upon size and experience. Agonistic behavior is also of great importance, as might be expected in a hunting animal. The parent-offspring relationship is important in early life because of the difficult manner of obtaining food. In some species, but not all, the sexual relationship is an important one, and there is a tendency toward the development of lasting bonds between males and females. *See* CARNIVORA.

Social behavior of rodents. Among mammals, rodents are good builders and diggers, and this is reflected in their social life. Perhaps the highest development of rodent society is found in the prairie dog, a large ground squirrel of the western plains in the United States. These animals dig numerous burrows and pile the dirt in mounds, which are the outward signs of a prairie dog town. A town is subdivided into territories, each inhabited by a coterie of prairie dogs. A coterie includes one or two male and two or three female adults, which produce numerous young and guard the territory of perhaps half an acre. Mating takes place in February and March, at which time there is a good deal of fighting between males and some reorganization of territories. Each female in the territory retires to a particular burrow and guards it as long as the young are belowground. Males must stay in other holes. When the young come out in April the females relax and all is again peaceful within the coterie.

The young learn the territorial boundaries by being challenged as they cross them. There is some social organization throughout the entire town, and any prairie dog that sees danger gives a warning signal which alerts all the rest. In June and July the adults leave the territory to construct new burrows on the outskirts of the colony. Thus there is a regular cultural inheritance. The young learn the territory from the parents and inherit the burrows. This system provides for colonization and also means that the experienced adult animals live on the outskirts of the colony, the part most exposed to danger.

In behavior typical of most other rodents, the principal type of social behavior is the care of the young in the nest, a relationship which tends to break up as the young leave. The male-female relationship is more important than in ungulates but is relatively unstable. The young are bound together chiefly by the territorial system. There is almost no allelomimetic behavior and consequently very little coordinated group activity. In the less social rodents such as mice, an entire society may consist of a mother and her young. *See* RODENTIA.

Social behavior of primates. Primates form a large and variable order consisting of nearly 200 species. With the exception of humans and a few species that have penetrated into mild temperate regions, primates live in tropical climates, perhaps because of their highly developed fingers, which become stiff and unusable in cold weather. However, they have evolved in widely different ways as they have become adapted to different habitats. The plains-living baboons with their highly developed teeth and habits of running on all fours are similar in many ways to dogs, while the tree-living galagos have a general appearance similar to that of domestic cats. Gorillas have become entirely herbivorous and in their large size and huge paunches resemble the larger ungulates. There is even a slothlike primate, the slow loris, which lives entirely in the tops of trees and moves deliberately, although not hanging upside down continually as do true sloths.

Unlike most other mammals, primates have a highly developed sense of color vision, and some species, such as mandrills, have evolved great differences in color and form between the sexes and between adults and young, similar to the plumage differences in certain birds. Again as in birds, hearing is well developed and olfaction is relatively unimportant.

Primates have the best-developed capacities for manipulation and prehension of any group of mammals. An opposable thumb is found in many species, and the hindlimbs are often well developed for manipulation. In some arboreal forms the tail is an important grasping organ. The characteristic mode of reproduction involves the production of a single infant. Unlike the young of the ungulates, those of the primates are usually born in a relatively immature state and require constant care and considerable maternal manipulation over long periods.

Baboons. These plains-living primates are in many aspects of their social life more similar to humans

than are the more closely related anthropoid apes. A savannah baboon society consists of several males of various ages, a larger group of females, and assorted young of both sexes. The whole group occupies an area of approximately 2 mi^2 (5 km^2), wandering over it in search of food in the daytime and returning to certain trees for sleeping at night. The males are organized into a definite dominance order. The most dominant male usually stays in the center of the group with other males spaced out around him. This results in the younger and presumably more active and alert males being on the edge of the group. When a baboon sights a predator, it gives an alarm call, and all the males quickly gather on the side of the troop toward the predator and threaten or attack it. The combined force is usually enough to scare off most predators except lions.

The females and their infants are likely to be found close to the dominant male, not because he is attractive, but because this is the safest place in the group. Females in estrus may mate with the dominant male or may prefer another. There is no evidence of permanent consortship, and females freely move from one male to another. There is no fighting between males over females, and almost all displays of agonistic behavior from whatever cause are confined to threats.

The dominance order is developed largely through playful activities of the young. When an infant is hurt by another baboon or when a young animal playfully attacks an older one, the aggressor is chased and threatened by one of the older males. In this way the young baboon soon learns his place in the hierarchy. There is no special relationship with the biological father and his offspring, but successive offspring of the same female retain a special relationship with her and each other. This is usually shown by contact and mutual grooming. Baboons thus develop a cohesive and cooperative society which maintains itself over long periods under favorable conditions.

Related primates such as rhesus monkeys and other macaques develop a similar social organization although there is considerable variation with respect to the amount of overt agonistic behavior and the amount of direct contact between individuals. Cultural transmission of information has been observed in the Japanese macaque. One troop was fed grain on a sandy shore. After one member learned to separate sand and grain by flotation, the habit spread through the group, the oldest members adopting it last.

Langurs. The Indian langur monkey is much more arboreal, but a troop will frequently descend to the ground to feed. They live chiefly on leaves. The adults have a striking appearance, with dark faces surrounded by white hair rising to a crest on top. As in most other primates, there are almost always more males than females in the usual troop. In addition, small all-male groups are common. Compared to macaques, langur groups are more relaxed and show less aggressive behavior. The adult males in a troop are organized into a definite dominance order and are always dominant over females. Dominance relationships among females are less definite. Males determine the direction of movement of the group. When an infant is born, all adult females are intensely interested and the infant is passed from hand to hand.

Group activities are integrated by whooping calls from males, which draw the group together in forested areas and also communicate the troop's location to nearby troops. Interactions of an aggressive nature are common between groups. Attacks are usually initiated by the low-ranking males, and adjacent troops show indications of a dominance order with respect to each other. As in many other primates, there is a core area in which the troop spends most of its time and a larger area over which the troop wanders. Ranges overlap, and there is no defense of territorial boundaries as such. Intertroop conflicts increase as the home ranges are compressed by population pressure.

Howler monkeys. These New World monkeys are entirely arboreal. They exhibit very little agonistic behavior of any kind, except that when troops draw near to each other the males roar loudly, and the two groups thus avoid closer contact. All feeding is done in the tops of trees. As a troop moves from place to place, males usually take the lead, although there is no single leader of the group. As with baboons, females in estrus move from male to male without conflict. The care-dependency relationship between mothers and offspring is the most prominent.

Anthropoids. Of the three major genera, orangutans are virtually unknown with respect to social organization in the wild, and much information remains to be gathered about the other two. Gorillas are the largest anthropoids. They live in small groups consisting of one or two males plus a few females and their offspring, and perhaps a few solitary males. The sex organs are inconspicuous, and sexual behavior is rarely observed. Little fighting takes place within the group. The animals respond to potentially dangerous predators by beating their chests, by threatening, and occasionally by attacking. Mature males are organized into a dominance order, but there is no hierarchy among females. Home ranges overlap, and when two troops meet they may exhibit a few threats, or mingle together and travel peacefully for a time. Both gorillas and chimpanzees live in dense forests and frequently build temporary nests in trees.

Social organization among chimpanzees is much less definite than in many other primates. The animals in a given area rarely form large groups but associate with each other in small, temporary bands consisting of females with young, all-male groups, or groups of males and females without young. There are few indications of dominance. Fighting is rare but there is some threatening. Much of the social behavior is composed of mutual grooming, as in other primates. Most food is fruit, but chimpanzees occasionally kill small animals and eat them. They also use crude tools in the form of twigs and rolled-up

leaves to extract termites from their nests, and these habits are apparently learned from mothers by infants, indicating the existence of some degree of culture.

Thus while chimpanzees are, as individuals, more similar to humans than any other primates, their unstable social organization is quite dissimilar. The size and composition of primate societies vary from species to species. In the gibbons, in which the two sexes are much alike in physique and both are highly aggressive, the typical group includes only one adult male and female. The forest-living great apes, such as chimpanzees and gorillas, live in small groups of eight or nine, mostly young, with one or two each of adult males and females. In contrast, the macaques may form groups as large as 150. Thus the ground-living primates form larger groups than those which live in trees. However, there is a great variety of social organization and behavior in primates, and no one species can be taken as an exact model for primitive human behavior. *See* APES.

Development of social organization. The postnatal development of each species is closely related to the social organization typical of the adults. Every highly social animal has a short period early in life when it readily forms attachments to any animal with which it has prolonged contact. This critical period for primary socialization can be demonstrated by removing young from their natural parents and rearing them by hand. Experience before this period has no effect; during the period the young animal transfers all its native social behavior to its human contact; thereafter this process becomes increasingly difficult, although captivity and hand feeding have some effect even on adults.

Process of socialization. The process of socialization begins almost immediately after birth in ungulates like the sheep, and the primary relationship is formed with the mother. In dogs and wolves the process does not begin until about 3 weeks of age, at a time when the mother is beginning to leave the pups. Consequently the strongest relationships are formed with litter mates, thus forming the foundation of a pack. Many rodents stay in the nest long after birth; primary relationships are therefore formed with nest mates. Young primates are typically surrounded by a group of their own kind, but because they are carried for long periods the first strong relationship tends to be with the mother.

Mammalian society. In summary, mammals may develop all types of social behavior to a high degree, but not necessarily in every species. The above examples include four highly developed societies from representative orders; in other species certain types of social behavior may be completely absent. Mammals have great capacities for learning and adaptation, which means that social relationships are often highly developed on the basis of learning and habit formation as well as on the basis of heredity and biological differences. The resulting societies tend to be malleable and variable within the same species and to show considerable evidence of cultural inheritance from one generation to the next. Mammalian societies have been completely described in relatively few forms, and new discoveries will probably reveal the existence of a greater variety of social organization.

Basic human social organization and behavior obviously differs from that of all other primates, although relatedness does exist. At the same time, the range of variability of human societies as seen in the nuclear family does not approach that in mammals as a whole. Human societies are characterized by the presence of all fundamental types of social behavior and social relationships rather than by extreme specialization. *See* ETHOLOGY; HUMAN ECOLOGY; REPRODUCTIVE BEHAVIOR; SOCIOBIOLOGY. John P. Scott

Bibliography. W. C. Allee, *The Social Life of Animals*, 1938, reprint 1976; S. A. Altmann (ed.), *Social Communication among Primates*, 1967, reprint 1982; C. A. Bramblett, *Patterns of Primate Behavior*, 2d ed., 1993; C. R. Carpenter, *Behavioral Regulators of Behavior in Primates*, 1974; W. E. Etkin, *Social Behavior and Organization among Vertebrates*, 1964; J. P. Scott, *The Evolution of Social Systems*, vol. 3 of *Monographs in Psychology*, 1989; N. Tinbergen, *Social Behaviour in Animals*, 2d ed., 1990.

Sociobiology

A scientific discipline that applies principles of evolutionary biology to the study of animal and human social behavior. It is a synthesis of ethology, ecology, and evolutionary theory in which social behavior is viewed as a product of natural selection and other biological processes. *See* ECOLOGY; ETHOLOGY; ORGANIC EVOLUTION.

Underlying principles. Sociobiology predicts that individuals will behave in ways that maximize their fitness (their success at projecting copies of their genetic material into succeeding generations) and argues that such behaviors can arise through the same evolutionary processes that operate on other trait systems. The central principle underlying sociobiology is that an individual's behavior is shaped, in part, by its genes, and thus is heritable and subject to natural selection. Natural selection is simply the result of the differential survival and reproduction of individuals who show heritable variation in a trait. The variants of a trait that allow individuals to survive better or reproduce better—the variants that convey greater fitness—will increase in frequency in a population over time. *See* BEHAVIORAL ECOLOGY.

As an example of natural selection, consider a population made up of asexually reproducing, haploid individuals carrying one of two alternative forms of a gene, allele A or allele a, initially at equal frequency. In each generation, individuals carrying one of these alleles have more offspring than individuals carrying the other. Over time, the frequency of the allele associated with greater fertility will increase in the

population, while that associated with lower fertility will decrease. Applied to behavioral traits, the logic is similar. Individuals carrying allele *A* are more likely to perform behavior *X*, whereas individuals carrying allele *a* are more likely to perform alternative behavior *x*. If the different behaviors, *X* and *x*, result in differential survival or reproduction, those individuals performing the more successful behavioral option will pass on more copies of the corresponding allele to the next generation. Eventually, the population will tend to be composed of individuals carrying the more successful allele and performing the more fit behavior. For diploid, sexually reproducing animals, this process is somewhat more complex (because of the reassortment of alleles at meiosis), but the outcome is fundamentally equivalent. *See* POPULATION GENETICS.

The distinctive sociobiological perspective on behavior views behaviors as strategies that have evolved through natural selection to maximize an individual's fitness. Sociobiologists employ this perspective in an inductive way, predicting that, all else being equal, animals should behave in ways that lead to greater fitness. Note that under this view there is absolutely no need to assume that maximizing fitness is in any way a conscious goal underlying an animal's behavior—successful behavior is simply an emergent result of the evolutionary process and not the force driving that process.

Some critics of the sociobiological perspective argue strongly against the notion of any genetic determinism for behavior, instead viewing environment and experience as the most important causal variables shaping behavior. This resurrects the ancient debate over nature versus nurture. Environment and experience are undoubtedly important, and the stance of most sociobiologists is that behavior—like all other aspects of an organism's phenotype—is the result of an interaction between genes and environment. Neither nature (genes) nor nurture (environment and experience) alone determines behavior. Rather, all behavior is influenced by both factors to varying degrees. It is the interaction between genotypes and environments that determines the direction and degree to which behavioral variants are associated with differential reproduction and survival—thus, the nature versus nurture debate is largely misguided. *See* BEHAVIOR GENETICS.

Altruism. Animals sometimes behave in ways that seem to reduce their own personal fitness while increasing that of other individuals. For example, in many species of social mammals and in numerous species of birds, individuals commonly give alarm vocalizations that alert others to the presence of a predator while seemingly rendering themselves more conspicuous to attack. In many species of social insects, large numbers of nonreproductive workers often labor on behalf of a very small number of reproductive individuals. Evolutionary biologists since Charles Darwin have noted that these seemingly altruistic acts pose a challenge to the notion of an animal's behavior being a strategy to maximize individual fitness. *See* SOCIAL INSECTS; SOCIAL MAMMALS.

Inclusive fitness. One resolution to this paradox comes from taking a "gene's eye view" of evolution, which recognizes that Darwinian fitness (an individual's success in producing offspring) is simply a specific case of the more general concept of inclusive fitness. Because an individual shares genes with its relatives, there are actually two different routes by which it can pass on copies of those genes to the next generation: first, through personal or direct reproduction; and second, by helping relatives to reproduce. Inclusive fitness is a composite measure of an individual's genetic contribution to the next generation that considers both of these routes.

Kin selection. Animals who behave altruistically toward relatives may, in fact, be behaving selfishly in the genetic sense if their behavior sufficiently enhances their inclusive fitness through its effects on the survival and reproduction of relatives. This phenomenon is referred to as kin selection. Because close relatives share more genes, on average, with an individual than do distant relatives, such seemingly altruistic behavior tends to be directed toward closer kin.

Numerous studies have demonstrated the importance of kin selection on structuring individual social behavior in a wide array of animal taxa. For example, female ground squirrels are especially likely to give alarm calls when their relatives are nearby, but they call far less often when close relatives are absent. Similarly, in cooperatively breeding birds and callitrichid primates (marmosets and tamarins), individuals that assist a breeding pair in rearing young are typically older offspring of the pair and thus are helping to rear their own younger siblings. Kin selection has also been implicated in the evolution of complex caste societies in some species of social insects. Because of their peculiar genetic system, workers in many species of social wasps and bees are more closely related to their sisters (other workers) than they would be to their own offspring if they were to reproduce. By helping the queen (their mother) to reproduce, they actually do more for their own inclusive fitness than if they were to attempt to produce offspring on their own. *See* SOCIAL HIERARCHY.

Reciprocity. Another solution to the altruism paradox—one that can even operate among nonrelatives under certain conditions—is reciprocity. If the recipient of some altruistic behavior is likely to repay a donor in the future, it may be beneficial for the donor to perform the behavior even at some immediate cost to its own fitness. This is especially true if the cost is low and the anticipated benefit is high. Examples include food sharing among unrelated vampire bats and alliance behavior in some primates.

Under some conditions, it is possible for altruism to result from natural selection operating at the level of groups, rather than at the level of individuals or genes (for example, if groups containing altruists were at a sufficient selective advantage compared with groups containing only selfish individuals). However, within such groups, selection acting

on individuals would still tend to favor selfishness over altruism. In any event, much of the cogency of sociobiology has resulted from the recognition that selection seems to be most powerful when it acts at lower levels—notably, genes and individuals rather than groups or species.

Other social strategies. Over the past three decades, the sociobiological view of behavior as an evolved strategy for maximizing inclusive fitness has proven to be a powerful explanatory paradigm for investigating many other aspects of animal social behavior. The phenomenon of sociality, for example, has been extensively explored from this perspective. Living in a social group is a behavior that involves numerous potential benefits (including improved access to food, more effective deterrence of or defense against predators, and enhanced opportunities for nepotism) as well as numerous potential costs (such as having to share food resources and competition over mates). Each species—and individuals within each species—can be investigated as to the relative costs and benefits associated with the decision of whether or not to be social.

Similar considerations apply to many other strategic (though typically unconscious) decisions that animals make during their lives, such as whether to reproduce sexually, when in life to begin reproducing (with regard to season and age at maturation), how many partners to mate with and who those partners should be, whether to compete with other individuals over access to partners, and whether to bestow parental care and the level of care to give. For example, consider the case of mating and parenting strategies for males versus females. Biologists define males as the sex that produces a large number of small, relatively mobile gametes (sperm), and females as the sex that produces fewer, larger, and relatively immobile gametes (eggs). In many species, this initial inequality between the sexes in parental investment is compounded by additional forms of care provided primarily by females (for example, gestation and lactation in mammals). Fundamental differences between males and females in the level of investment that each provides result in marked differences between the sexes in maximum potential lifetime fitness. In general, individual males of most species are potentially able to produce many more offspring than even the most fertile females since the male contribution to an offspring can be as little as a single sperm. Male reproductive success is thus limited, in many cases, by access to fertile females and is often enhanced by achieving a maximum number of successful copulations. For a female, fitness is much more limited by access to resources than by access to mates because of her much lower lifetime reproductive potential and relatively greater investment in each offspring. In general, the sex providing less parental investment (typically males) is expected to compete among themselves for access to the sex providing greater investment (typically females). Males, accordingly, are likely to maximize their fitness by being aggressive and competitive with one another and are expected to be less discriminating about mate choices, whereas females are likely to be more selective in their mate choices. Significantly, species in which the pattern of parental investment is reversed (as among certain species of birds, fish, and frogs in which males provide substantially more prenatal or postnatal care than females) also show a reversal in the pattern of competitive behavior, with females tending to be larger, showier, and more aggressive than males. *See* MATERNAL BEHAVIOR; REPRODUCTIVE BEHAVIOR; SEXUAL DIMORPHISM.

The sociobiological concept of behaviors as fitness-enhancing evolutionary strategies has led to many other insights concerning the expression of animal social behaviors, including predictions regarding parent-offspring conflict, sibling rivalry, birth order effects, development of dominance hierarchies, alliance and coalition formation, intentional deception, status striving, and cooperative hunting. Sometimes, distinct behavioral strategies for dealing with the same set of environmental challenges may exist in a population simultaneously—for example, some individuals might be likely to escalate fights while others back away. Sociobiologists sometimes interpret this as an example of an evolutionarily stable strategy set in which the frequency of one behavioral type may depend on the frequency of others maintained in the population.

Human behavior. Although most of the research in sociobiology has focused on understanding the behavior of nonhumans, sociobiological explanations have been used to interpret patterns of human behavior as well. This has generated much controversy within both the scientific and lay communities. Many people have objected to what they see as the political and ethical implications of applying concepts from evolutionary biology to the human species. Much of this concern is well founded, given the past abuses of Social Darwinism and various other racist and eugenic programs. However, an important distinction must be made between trying to understand the biological underpinnings of human behavior and morally condoning or condemning that behavior. Sociobiology may well shed considerable light on why humans do the things they do—for example, why men generally tend to be more aggressive and competitive than women, or why people tend to favor their relatives over nonkin—but it does not render judgment. In fact, a sociobiological perspective on human behavior emphasizes the existence of cross-cultural universals—features of behavior that all human beings share by virtue of their biology that transcend notions of race, class, and ethnicity. As such, sociobiological studies could constitute a potent antidote to racist ideologies.

David P. Barash; Anthony Di Fiore

Bibliography. D. P. Barash, *Sociobiology and Behavior*, 2d ed., Elsevier, 1982; L. L. Betzig (ed.), *Human Nature: A Critical Reader*, Oxford University Press, 1996; L. Cronk, N. Chagnon, and W. Irons (eds.), *Adaptation and Human Behavior: An Anthropological Perspective*, Aldine De Gruyter, 2000; R. Dawkins, *The Selfish Gene*, Oxford University Press, 1990; J. R. Krebs and N. B. Davies (eds.),

Behavioural Ecology: An Evolutionary Approach, 4th ed., Blackwell Science, 1997; V. Standen and R. A. Foley (eds.), *Comparative Socioecology: The Behavioural Ecology of Humans and Other Mammals*, Blackwell Science, 1989; E. O. Wilson, *Sociobiology: The New Synthesis*, 25th Anniversary Edition, Harvard University Press, 2000.

Soda niter

A nitrate mineral having chemical composition $NaNO_3$ (sodium nitrate). Soda niter, otherwise known as nitratite, is by far the most abundant of the nitrate minerals. It crystallizes in the rhombohedral division of the hexagonal system. It sometimes occurs as simple rhombohedral crystals but is usually massive granular. The mineral has a perfect rhombohedral cleavage, conchoidal fracture, and is rather sectile. Its hardness is 1.5 to 2 on Mohs scale, and its specific gravity is 2.266. It has a vitreous luster and is transparent. It is colorless to white, but when tinted by impurities, it is reddish brown, gray, or lemon yellow.

Soda niter is a water-soluble salt found principally as a surface efflorescence in arid regions, or in sheltered places in wetter climates. It is usually associated with niter, nitrocalcite, gypsum, epsomite, mirabilite, and halite.

The only large-scale commercial deposits of soda niter in the world occur in a belt roughly 450 mi (725 km) long and 10–50 mi (16–80 km) wide along the eastern slope of the coast ranges in the Atacama, Tarapaca, and Antofagasta deserts of northern Chile. The deposits consist of a thin bed of nitrates and associated minerals, varying from a few inches to a few feet in thickness, overlain by a shallow overburden of sand and gravel. The crude soda niter, known as caliche, is about one-fourth sodium nitrate, admixed with other salts, notably bloedite, anhydrite, gypsum, polyhalite, halite, glauberite, and darapskite, together with minor amounts of various iodates, chromates, and borates. Because of the presence of the iodate minerals lautarite and dietzeite, these deposits yield most of the world's supply of iodine.

The origin of these deposits is controversial. It is generally agreed that the nitrates were transported by ground water and deposited by evaporation. The source of the nitrate has been attributed to (1) guano; (2) nitrogen fixation by electrical storms; (3) the bacterial fixation of nitrogen from vegetable matter; and (4) a volcanic source in nearby Triassic and Cretaceous rocks. The last source seems the most probable. *See* NITROGEN CYCLE.

Chilean nitrate had a monopoly of the world's fertilizer market for many years, but now occupies a subordinate position owing to the development of synthetic processes for nitrogen fixation which permit the production of nitrogen from the air. This led to the commercial production of artificial nitrates and reduced the need to import Chilean nitrates for use in the manufacture of fertilizers and explosives. *See* CALICHE; EXPLOSIVE; FERTILIZER.

Small deposits of soda niter similar to those in Chile are found in Bolivia, Peru, North Africa, Egypt, Russia, India, and the western United States. *See* NITER; NITRATE MINERALS.

George Switzer

Sodalite

A mineral of the feldspathoid group, crystallizing in the isometric system, with chemical composition $Na_4Al_3Si_3O_{12}Cl$. Its tektosilicate framework, in which half the tetrahedrally coordinated Si atoms are replaced by Al, is arranged in four- and six-membered rings. The rings define sets of channels which intersect to form cavities. Na atoms reside in the channels and Cl in the cavities. Crystals are rare, usually dodecahedrons. Sodalite is usually massive or granular with poor cleavage. The Mohs hardness is 5.5–6.0, and the density 2.3. The luster is vitreous, and the color is usually blue but may also be white, gray, or green.

Sodalite is common in nephelite syenites and associated rock types. It also occurs in metasomatized carbonates at contacts with alkaline igneous rocks. Notable occurrences are at Mount Vesuvius, Italy; at Bancroft, Ontario, Canada; and on the Kola Peninsula, Russia. In carbonaceous chondrite meteorites, it occurs as a gas-phase alteration product of refractory condensate inclusions. *See* FELDSPATHOID; SILICATE MINERALS.

Lawrence Grossman

Sodium

A chemical element, Na, atomic number 11, and atomic weight 22.9898. Sodium is between lithium and potassium in the periodic table. The element is a soft, reactive, low-melting metal with a specific gravity of 0.97 at 20°C (68°F). Sodium is commercially the most important alkali metal. The physical properties of metallic sodium are summarized in the **table** below. *See* PERIODIC TABLE.

Sodium ranks sixth in abundance among all the elements in the Earth's crust, which contains 2.83% sodium in combined form. Only oxygen, silicon,

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Physical properties of sodium metal

Property	Temperature °C	Temperature °F	Metric (scientific) units	British (engineering) units
Density	0	32	0.972 g/cm^3	60.8 lb/ft^3
	100	212	0.928 g/cm^3	58.0 lb/ft^3
	800	1472	0.757 g/cm^3	47.3 lb/ft^3
Melting point	97.5	207.5		
Boiling point	883	1621		
Heat of fusion	97.5	207.5	27.2 cal/g	48.96 Btu/lb
Heat of vaporization	883	1621	1005 cal/g	1809 Btu/lb
Viscosity	250	482	3.81 millipoises	4.3 kinetic units
	400	752	2.69 millipoises	3.1 kinetic units
Vapor pressure	440	824	1 mm	0.019 lb/in.^{2}
	815	1499	400 mm	7.75 lb/in.^{2}
Thermal conductivity	21.2	70.2	0.317 cal/(s)(cm)(°C)	76 Btu/(h)(ft)(°F)
	200	392	0.193 cal/(s)(cm)(°C)	46.7 Btu(h)(ft)(°F)
Heat capacity	20	68	0.30 cal/(g)(°C)	0.30 Btu/(lb)(°F)
	200	392	0.32 cal/(g)(°C)	0.32 Btu/(lb)(°F)
Electrical resistivity	100	212	965 microhm-cm	
Surface tension	100	212	206.4 dynes/cm	
	250	482	199.5 dynes/cm	

aluminum, iron, and calcium are more abundant. Sodium is, after chlorine, the second most abundant element in solution in seawater. The important sodium salts found in nature include sodium chloride (rock salt), sodium carbonate (soda and trona), sodium borate (borax), sodium nitrate (Chile saltpeter), and sodium sulfate. Sodium salts are found in seawater, salt lakes, alkaline lakes, and mineral springs. *See* ALKALI METALS.

Sodium reacts rapidly with water, and even with snow and ice, to give sodium hydroxide and hydrogen. The reaction liberates sufficient heat to melt the sodium and ignite the hydrogen. When exposed to air, freshly cut sodium metal loses its silvery appearance and becomes dull gray because of the formation of a coating of sodium oxide.

Sodium does not react with nitrogen. Sodium and hydrogen react above about 200°C (390°F) to form sodium hydride. Sodium reacts with ammonia, forming sodium amide. Sodium also reacts with ammonia in the presence of coke to form sodium cyanide.

Sodium does not react with paraffin hydrocarbons but does form addition compounds with naphthalene and other polycyclic aromatic compounds and with arylated alkenes. The reaction of sodium with alcohols is similar to, but less rapid than, the reaction of sodium with water. Sodium reacts with organic halides in two general ways. One of these involves condensation of two organic, halogen-bearing compounds by removal of the halogen, allowing the two organic radicals to join directly. The second type of reaction involves replacement of the halogen by sodium, giving an organosodium compound. *See* ORGANOMETALLIC COMPOUND.

Sodium chloride, or common salt, NaCl, is not only the form in which sodium is found in nature but (in purified form) is the most important sodium compound in commerce as well. Sodium hydroxide, NaOH, is also commonly known as caustic soda. Sodium carbonate, Na_2CO_3, is best known under the name soda ash.

A major use of sodium is in the reduction of animal and vegetable oils to long-chain fatty alcohols; these alcohols are raw materials for detergent manufacture. Sodium is used to reduce titanium and zirconium halides to their respective metals. Sodium chloride is used in curing fish, meat packing, curing hides, making freezing mixtures, and food preparation (including canning and preserving). Sodium hydroxide is used in the manufacture of chemicals, cellulose film, rayon soap pulp, and paper. Sodium carbonate is used in the glass industry and in the manufacture of soap, detergents, various cleansers, paper and textiles, nonferrous metals, and petroleum products. Sodium sulfate (salt cake) is used in the pulp industry and in the manufacture of flat glass. *See* GLASS; HALITE; PAPER. Marshall Settig

The sodium ion (Na^+) is the main positive ion present in extracellular fluids and is essential for maintenance of the osmotic pressure and of the water and electrolyte balances of body fluids. Hydrolysis of adenosine triphosphate (ATP) is mediated by the membrane-bound enzyme Na^+, K^+-ATPase (this enzyme is also called sodium pump). The potential difference associated with the transmembrane sodium and potassium ion gradients is important for nerve transmission and muscle contraction. Sodium ion gradients are also responsible for various sodium ion–dependent transport processes, including sodium-proton exchange in the heart, sugar transport in the intestine, and sodium-lithium exchange and amino acid transport in red blood cells. *See* POTASSIUM. Duarte Mota de Freitas

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; C. A. Pasternak (ed.), *Monovalent Cations in Biological Systems*, 1990.

Sodium-vapor lamp

An arc discharge lamp with sodium vapor as the emitting species. Low-pressure and high-pressure types are in common commercial usage.

Low-pressure lamps. The low-pressure sodium lamp has remarkably high luminous efficiency, or efficacy, producing as much as 200 lumens per watt of input power. The radiation is nearly monochromatic yellow in color, and is used where color rendition is unimportant such as on piers, in harbors, for low-energy-cost security lighting, on bridges, and on roadways. The very high luminous efficacy is due to very stringent control of heat losses from the arc. The key to the heat conservation is the indium oxide film (**Fig. 1**) used to reflect infrared energy back to the arc tube and the evacuated outer jacket that minimizes the conducted thermal losses. Lamps at 18 and 10 W of power have been developed for very low-energy-cost uses. The 10-W lamp produces light at 100 lm/W. *See* ILLUMINATION; LUMINOUS EFFICACY.

The low-pressure sodium lamp has an operating sodium pressure of about 10^{-5} atm (1 pascal), and 10^{-3} to 10^{-2} atm (100 to 1000 pascals) of inert gas (typically neon) for arc starting and operating purposes. The arc tube wall temperature is about 520°F (270°C).

The chemistry problems of lamps containing elemental sodium are formidable. Ordinary glasses react at temperatures near 480°F (250°C), and low-pressure sodium lamps are specially coated with a sodium-resistant glass for protection to 520°F (270°C).

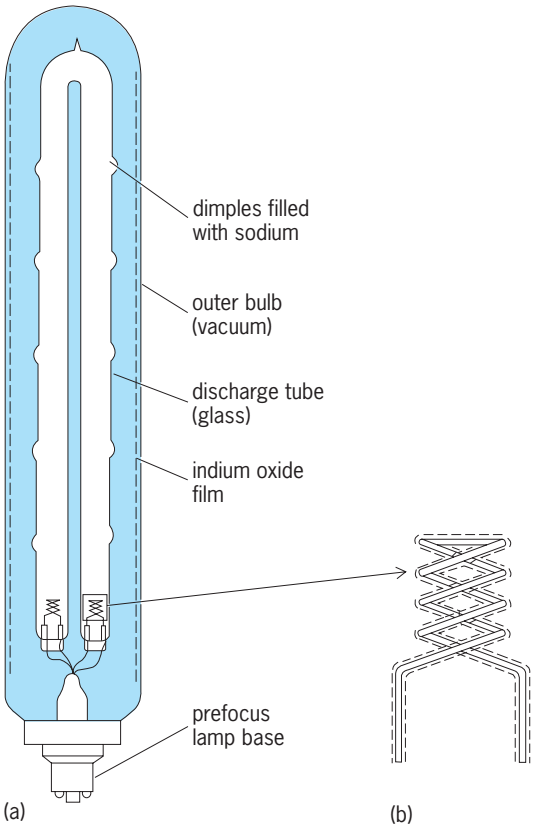

Fig. 1. Low-pressure sodium lamp. (*a*) Lamp structure. (*b*) Structure of electrode consisting of coiled tungsten wire coated with oxides of barium, strontium, and calcium.

High-pressure lamps. When the sodium pressure is increased by a higher liquid-sodium-pool temperature, the yellow sodium D-line resonance radiation reaches a condition where pressure broadening and self-absorption occur, producing a lamp no longer monochromatic yellow but golden-white in color. The high-pressure sodium lamp has a maximum arc tube temperature typically above 2000°F (1100°C) with a sodium amalgam reservoir temperature about 1300°F (700°C). The term high-pressure is used merely to distinguish this light source from the low-pressure sodium lamp. The plasma arc column of the high-pressure sodium lamp has a total pressure of sodium, mercury, and inert gas of typically slightly less than 1 atm (10^5 Pa).

The severity of reactions involving elemental sodium increases with the higher operating temperature and pressure of the high-pressure sodium lamp, and special materials must be used for the arc containment vessel and for the electrical feedthrough seals. Indeed, the success of the high-pressure sodium lamp is due to the invention of an arc containment material resistant to sodium vapor at high temperature. This material is a very high-purity aluminum oxide ceramic (**Fig. 2**) with the ability to pass light with little attenuation. (Previously all discharge lamps used either glass or fused quartz as the containment material.) In order to conserve losses, the high-pressure sodium lamp uses low-thermal-conductivity buffers of mercury and xenon in the sodium discharge. The mercury and xenon make almost no contribution to the lamp spectra.

Properties and uses. In comparison with other arc discharge lamps such as high-pressure mercury, metal halide, and fluorescent, the sodium-vapor lamps produce very low levels of ultraviolet radiation: about 0.5% of the input lamp power results in ultraviolet with the major amount at 360 nanometers. The ultraviolet radiation in the erythema-sensitive regions near 300 nm is virtually nonexistent. Thus tanning of the skin is not possible with sodium-vapor lamps. *See* FLUORESCENT LAMP; MERCURY-VAPOR LAMP; METAL HALIDE LAMP; ULTRAVIOLET LAMP; ULTRAVIOLET RADIATION.

Low-pressure sodium lamps are used commercially in lamp powers from 10 to 180 W at efficacies from 100 to 200 lm/W. High-pressure sodium lamps are offered in powers ranging from 35 to 1000 W and from 60 to 150 lm/W. The maximum efficacies of sodium lamps are considerably greater than those of other light sources such as fluorescent lamps (approximately 90 lm/W), high-pressure mercury lamps (60 lm/W), tungsten halogen lamps (25 lm/W), and incandescent lamps (15 lm/W). Both types of sodium lamps convert 30 to 35% of input power to useful visible light.

Because of the more compact source size, the light from a high-pressure sodium lamp can be

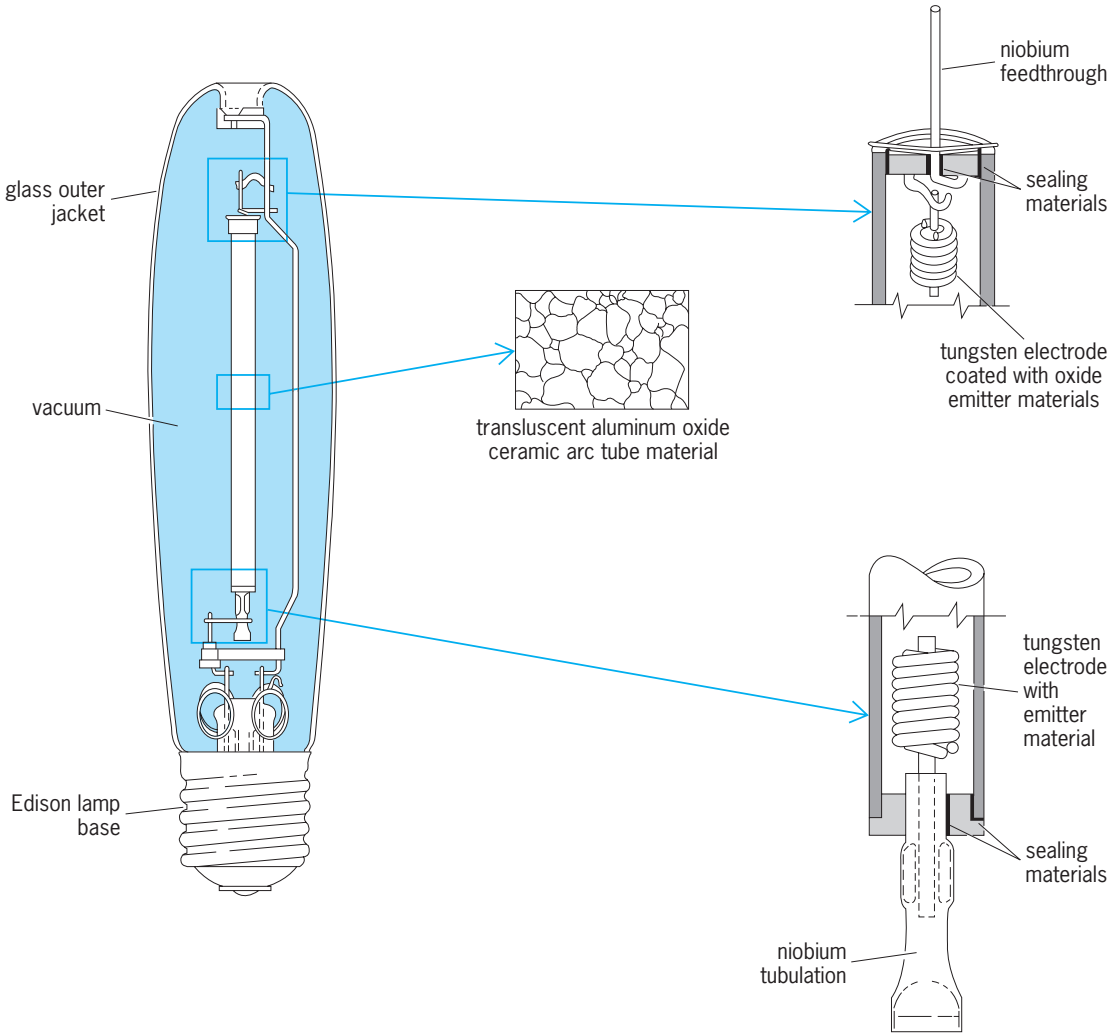

Fig. 2. High-pressure sodium lamp and components.

more effectively directed by reflectors and refractors than that from the low-pressure sodium lamp. Therefore the higher efficacy of the low-pressure sodium lamp is offset by its reduced luminous efficiency. Sodium-vapor lamps are replacing high-pressure mercury lamps in many floodlight and roadway applications. More sodium-vapor lamps are manufactured annually than the high-pressure mercury and metal halide lamps combined. *See* LAMP; VAPOR LAMP.

Charles I. McVey

Bibliography. M. A. Cayless and A. M. Marsden, *Lamps and Lighting*, 3d ed., 1983; J. J. DeGroot and A. J. Van Vliet (eds.), *The High Pressure Sodium Lamp*, 1986; Special issue on light source technology, *IEE Proceedings*, vol. 127A, no. 3, 1980.

Sofar

The acronym for sound fixing and ranging. The concept is based on finding the geographic location of sound that is created by an explosive or other impulsive event, such as a missile impact, by using the time of arrival of the signal at three or more receiving stations. This is known as hyperbolic fixing. This concept was used to locate aircraft that crashed at sea. Although other techniques have replaced this application, the term sofar channel has remained in use, referring to what is also known as the deep sound channel. *See* HYPERBOLIC NAVIGATION SYSTEM.

Sofar channel. Sound signals that are created in the deep oceans are known to travel very long distances. This is due to two important properties of sound in seawater. First, the speed of sound, which is approximately 1500 meters per second (5000 feet per second), is dependent on temperature and pressure, increasing with both as they increase. Second, the absorption of sound is very small, especially at lower frequencies. For most deep seas the water temperature is highest at or near the surface and decreases with depth down to a few hundred meters, where it reaches a nearly constant value, which in most oceans is about 4°C (39°F). Pressure, of course, increases with depth. Sound speed will, therefore, decrease with depth due to temperature, and increase with depth due to pressure. Typically this leads to a minimum value of sound speed at a depth of around

1000 m (3300 ft). This depth is usually referred to as the channel axis.

The sound radiating from an explosive source, for example, can be depicted as a set of rays that are emitted in all directions. Each ray can be traced using a simple rule that is based on Snell's law. From Snell's law it can be shown that a ray will always bend in a direction away from a region of higher sound speed. Thus there will be a set of rays that starts upward, and then bends toward the axis without reaching the sea surface. They will pass through the channel axis and then, since the sound speed increases with depth below the axis, they will then bend upward toward the axis where they will then repeat the same excursion, and so on. There will be another set that starts downward and then bends upward toward the channel axis. Thus part of the sound is trapped in the channel and, due to the very low loss due to absorption, intensity is essentially decreased only due to the cylindrical spreading of the rays. Without the channel sound would be spread spherically and thereby lose much more energy. *See* REFRACTION OF WAVES; UNDERWATER SOUND.

Shape of received signal. Another important feature of the sofar channel is its effect on the shape of the received signal. As an example, consider a signal that originates on the channel axis and is received at a range that is also on the axis. Each ray, having its own vertical angle of emission, will follow its own serpentine path. Using Snell's law, each path can be traced and, knowing the sound speed as a function of depth, the time of arrival can be determined. At each point along the channel axis there will be a unique set of rays that arrive. However, it will be found that they do not have the same arrival time. Rays that make the widest paths from the axis will arrive the earliest, with the ray that stays on the axis arriving last. The signal will then be the sum of a family of individual signals. It is the arrival time of the last of these that most accurately determines range.

Ralph R. Goodman

Software

A set of instructions that cause a computer to perform one or more tasks. The set of instructions is often called a program or, if the set is particularly large and complex, a system. Computers cannot do any useful work without instructions from software; thus a combination of software and hardware (the computer) is necessary to do any computerized work. A program must tell the computer each of a set of minuscule tasks to perform, in a framework of logic, such that the computer knows exactly what to do and when to do it. *See* COMPUTER PROGRAMMING.

Programming languages. Programs are written in programming languages especially designed to facilitate the creation of software. In the 1950s, programming languages were numerical languages easily understood by computer hardware; often, programmers said they were writing such programs in machine language. The computers understood certain number sequences to mean that a particular computer operation was to be performed, and programmers would write down such operations, one instruction per line, telling the computer both the number of the operation and the numerical address of one or more pieces of data on which the operation was to be performed (**Fig. 1**). [Data were stored in the computer at a location identified by a numerical address.]

Example of relationship between assembler (mnemonic) and machine-languge code*

Mnemonic	Code	Instruction†
A	21	Add
AM	11	Add (I)
B	49	Branch
BB	42	Branch Back
BD	43	Branch on Digit
BI	46	Branch No Indicator
BNF	44	Branch No Flag
BNI	47	Branch on Indicator
BNR	45	Branch No Record Mark
BT	27	Branch and Transmit
BTM	17	Branch and Transmit (I)
C	24	Compare
CF	33	Clear Flag
CM	14	Compare (I)
DN	35	Dump Numerically
H	48	Halt
K	34	Control
M	23	Multiply
MM	13	Multiply (I)
NOP	41	No OPeration
RA	37	Read Alphamerically
RN	36	Read Numerically
S	22	Subtract
SF	32	Set Flag
SM	12	Subtract (I)
TD	25	Transmit Digit
TDM	15	Transmit Digit (I)
TF	26	Transmit Field
TFM	16	Transmit Field (I)
TR	31	Transmit Record
WA	39	Write Alphamerically
WN	38	Write Numerically

*From R. C. Houghton, Growing up with software tools, in R. L. Glass, *In the Beginning: Recollections of Software Pioneers*, pp. 272–291, IEEE Computer Society Press, 1998; © 1998 IEEE.
†I means Immediate.

Machine language was cumbersome, error-prone, and hard to change. In the latter 1950s, assembler (or assembly) language was invented. Assembler language was nearly the same as machine language, except that symbolic (instead of numerical) operations (see **table**) and symbolic addresses were used, making the code considerably easier to change, although it remained cumbersome and error-prone. Some assembler code is still written, usually for applications where extremely rapid code performance is needed, or where the code must interact at a fairly intimate level with the computer hardware.

The programmable aspects of computer hardware have not changed much since the 1950s. Computers still have numerical operations, and numerical addresses by which data may be accessed. However, programming languages have changed considerably. In the late 1950s, they began to move away from the machine view and toward a view of the problem being solved. Now programmers use

high-level languages, which look much more like English than a string of numbers or operation codes (**Fig. 2**).

Well-known programming languages include Basic, Java, and C. Basic is a fairly simple language, easily taught yet general in application. It has been modified into Visual Basic, a language useful for writing the portion of a program that the user "talks to." (This portion of a program is called the user interface. Since modern interfaces tend to be pictorial or "graphical" for ease of use, most such program portions are now called graphical user interfaces or GUIs.) Java is especially useful for creating software that runs on a network of computers, and is designed so that a program created to be used on one computer can be easily moved to a different kind of computer and run there with minimal change. (This so-called porting of software from one machine to another is normally extremely complicated and error-prone.) C and C^{++} are powerful but complex languages for writing such software as systems software and games. *See* HUMAN-COMPUTER INTERACTION; LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

COBOL (Common Business-Oriented Language) and Fortran (for formula translation) are heavily used in spite of their origins in the 1950s. They are useful for solving, respectively, business problems or scientific and engineering problems involving heavy use of mathematics. Both languages benefit from ongoing modernization programs; although their modern versions resemble those of the 1950s, the languages have been updated to reflect some (but not all) of the best of modern-day thinking regarding programming languages. *See* PROGRAMMING LANGUAGES.

Programs have never been written with 1's and 0's (binary numbers). They are represented in binary inside the computer, but the programmer never sees that representation. *See* NUMBERING SYSTEMS; NUMERICAL REPRESENTATION (COMPUTERS).

Types of software. There have been enormous changes in the software field. Perhaps the change of greatest significance to the general public is the development of already-built software that may be purchased. In the early days of software, when none had been created before, all software was built from scratch. A large amount of such custom-built software is still constructed for the specific needs of its users, since new tasks continue to be invented for computers and software, and since the intricate details of software often require that a new variant be custom-built to perform a task similar to one done before.

Some portions of custom-built software may be reused from previous products, although the ability to reuse components eludes the plug-together-software vision of those who advocate massive-scale reuse. Frequently, software can be modified from older versions rather than built from scratch to perform a task similar to those done before. This task is known as software maintenance. (Traditionally, in other fields, the word "maintenance" means to repair faults. Fault repair in software is also known as maintenance, but that activity is less common than enhancing old software to do new tasks.) Old software that continues to be modified to do new things is often referred to as legacy software.

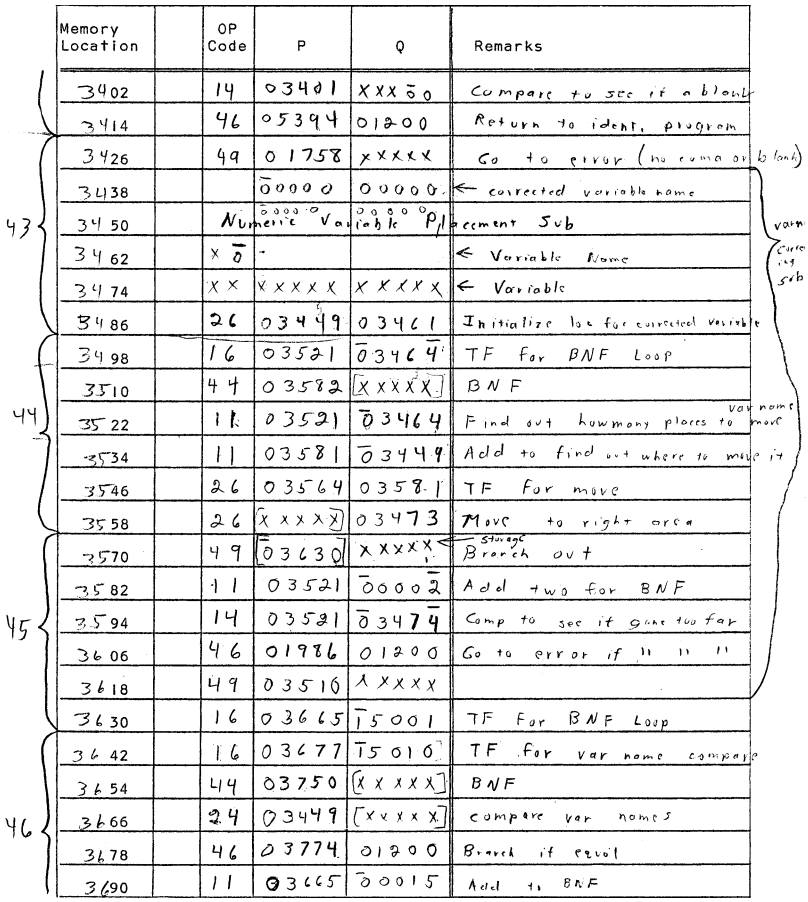

	Memory Location		OP Code	P	Q	Remarks
	3402		14	03401	XXX$\bar{0}$0	Compare to see if a blank
	3414		46	05394	01200	Return to ident. program
	3426		49	01758	XXXXX	Go to error (no comma or blank)
	3438			$\bar{0}$0000	00000	← corrected variable name
43	3450			Numeric Variable Placement Sub		
	3462		X $\bar{0}$	-		← Variable Name
	3474		XX	XXXXX	XXXXX	← Variable
	3486		26	03449	03461	Initialize loc for corrected variable
	3498		16	03521	$\bar{0}$3464	TF for BNF Loop
	3510		44	03582	[XXXXX]	BNF
44	3522		11	03521	$\bar{0}$3464	Find out how many places to move
	3534		11	03581	$\bar{0}$3449	Add to find out where to move it
	3546		26	03564	03581	TF for move
	3558		26	[XXXX]	03473	Move to right area
	3570		49	[$\bar{0}$3630]	XXXXX	Branch out
	3582		11	03521	$\bar{0}$0002	Add two for BNF
45	3594		14	03521	$\bar{0}$3474	Comp to see if gone too far
	3606		46	01986	01200	Go to error if " " "
	3618		49	03516	XXXXX	
	3630		16	03665	$\bar{1}$5001	TF for BNF Loop
	3642		16	03677	$\bar{1}$5010	TF for var name compare
	3654		44	03750	[XXXX]	BNF
46	3666		24	03449	[XXXX]	compare var names
	3678		46	03774	01200	Branch if equal
	3690		11	03665	$\bar{0}$0015	Add to BNF

Fig. 1. Example of machine-language code. (*From R. C. Houghton, Growing up with software tools, in R. L. Glass, In the Beginning: Recollections of Software Pioneers, pp. 272–291, IEEE Computer Society Press, 1998; © 1998 IEEE*)

Packaged software such as word processors, spreadsheets, graphics and drawing tools, email systems, and games are widely available and used. Some software packages are enormous; for example, enterprise resource planning (ERP) software can be used by companies to perform almost all of their so-called backoffice software work. Several vendors make ERP backoffice software so generalized that a company can use it for the majority of its accounting, marketing, personnel, manufacturing, and other information needs. *See* COMPUTER GRAPHICS; ELECTRONIC MAIL; VIDEO GAMES; WORD PROCESSING.

Because software is such a versatile product, it can be used for an enormously varied set of tasks. The number and variety of those tasks are so huge that it is difficult to even categorize them. One generally accepted categorization, however, identifies the following major areas.

Systems software. This is software necessary to support the running of an application program. Operating systems are needed to link the machine-dependent needs of a program with the capabilities of the machine on which it runs. Compilers translate programs from high-level languages into machine

```
procedure RecordErrorMessage
  (
     ErrorCode: ErrorCode_t;
     var Status: Status_t
     );
{ This routine outputs an error message based on an error code
  supplied by the calling routine. The way it outputs the message
  depends on the current processing state, which it retrieves
  on its own. It returns a variable indicating success or failure . }

var
     ProcessingMethod: ProcessingMethod_t;
     ErrorMessage:     Message_t;
     FileStatus:       Status_t;

begin
    { set the default status }
    Status := Failure;

     { look up the message based on the error code }
     LookupErrorMessage( ErrorCode, ErrorMessage );

     { if the error code is valid }
     if ( ErrorMessage.ValidCode ) then begin

          { determine the processing method }
          ProcessingMethod := CurrentProcessingMethod;

          { if doing interactive  processing }
          if ( ProcessingMethod = Interactive ) then begin

               { print the error message interactively and declare success }
               PrintInteractiveMessage( ErrorMessage. Text );
               Status := Success
          end

          { else doing batch processing }
          else if ( ProcessingMethod = Batch ) then begin

               { if the batch message file opens properly }
               FileStatus := OpenMessageFile;
               if ( FileStatus = Success ) then begin

                    { log the error message to the batch file, close the file,
                      and declare success }
                    LogBatchMessage( ErrorMessage.Text );
                    CloseMessageFile;
                    Status := Success
               end { if }
          end { else }
     end

     { else the message code is not valid }
     else begin

          { notify the user that an internal error has been detected }
          PrintInteractiveMessage( 'Internal Error: Invalid error code',
               'in RecordErrorMessage()' )
     end
```

Fig. 2. Example of high-level (Pascal) code. (*From S. McConnell, Code Complete: A Practical Handbook of Software Construction, Microsoft Press, 1993; reproduced by permission of Microsoft Press; all rights reserved*)

languages. Database programs keep track of where and how data are stored on the various storage facilities of a typical computer, and simplify the task of entering data into those facilities or retrieving the data. Networking software provides the support necessary for computers to interact with each other, and with data storage facilities, in a situation where multiple computers are necessary to perform a task, or when software is running on a network of computers (such as the Internet or the World Wide Web). *See* DATABASE MANAGEMENT SYSTEM; INTERNET; OPERATING SYSTEM; TELEPROCESSING; WORLD WIDE WEB.

Business applications software. This software processes transactions, produces paychecks, and does the myriad of other tasks that are essential to running any business. ERP systems are cutting down on (but not eliminating) the need to build customized business systems. Roughly two-thirds of software applications are in the business area.

Scientific and engineering software. This software satisfies the needs of a scientific or engineering user to perform enterprise-specific tasks. For example, a rocket propulsion company might develop software to evaluate the thrust that a particular combination of propellants will produce when they are mixed in a propulsion chamber. Such software is often highly mathematical in nature. Because scientific and engineering tasks tend to be very enterprise-specific, there has been no generalization of this application area analogous to that of the ERP for backoffice business systems. The scientific-engineering application usually is considered to be in second place only to business software in terms of number of software products built.

Edutainment software. This software instructs (educates) or plays games with (entertains) the user. Such software often employs elaborate graphics and complex logic. This is one of the most rapidly growing software application areas, and includes software to produce special effects for movies and television programs.

Real-time software. This software operates in a time-compressed, real-world environment. Although most software is in some sense real-time, since the users of modern software are usually interacting with it via a GUI, real-time software typically has much shorter time constraints. For example, software that controls a nuclear reactor, or a bomb dropping from an aircraft, must make decisions and react to its environment in minuscule fractions of a second, not the 1–5-second response time typically needed by a human being. Software with such tight timing requirements often must be produced to even tighter constraints than other software; performance analysis of such software is essential to make sure that the software can perform its tasks in the available time.

The preceding areas are often known as application areas or application domains, since they involve applying the computer to a certain kind of problem's solution.

Internal functioning. A look at how software functions internally reveals several matters of interest.

Program listing. Software is an ethereal concept since it is a set of instructions, not a physical entity, and it has no weight and occupies no space. To the programmer, a program is embodied in its listing, the textual representation of the set of instructions, written in the language of the programmer's choosing. A listing may be printed on a ream of paper (the typical program contains many thousands, perhaps many millions, of programming language statements) or scrolled across a computer screen. To change, or add to, a program, the programmer uses a compiler to manipulate the statements of the programming language in much the same fashion as a word processor manipulates written text (with the big difference that, when the program is complete, the compiler translates the program into machine code).

Client-server and three-tier systems. The average program has at least three major parts. As software becomes more complex, those parts have come to represent three separate program portions.

One portion is the GUI. Expectations of users for their interaction with a program are far higher than in the past, and the GUI has, therefore, increased enormously in complexity. For significant software systems, it is not unusual for the GUI not only to be a separate piece of software but also to run on a separate computer. This GUI computer is often known as the client computer, because the client or user sits at that computer when using the system.

Another portion of the program interacts with the environment in which the system runs. Often, especially for business applications, that will involve manipulating data in a database. Because database access (or any other kind of environment interaction) is a very different problem from, for example, interacting with the user, this portion of the program also often runs on its own computer. This computer is called the server computer, since it serves up information from the environment to the remainder of the program. Often, especially for business applications, the software on the server is a database package especially built for convenient access to masses of data.

The third portion of a typical contemporary program is called the application portion. It contains those portions of the program that provide the solution to the user's specific problem. A decision must be made as to where the application portion should run. Running it on the client computer along with the GUI results in a so-called fat client (fat because the client computer contains more software than it would have without that decision). Running it on the server computer is sometimes called a fat server approach. But more often, the application program runs in its own computer. A three-computer solution is known as a three-tier system, and the application computer then becomes known as the middle-tier computer. For programs and systems of any complexity, the three-tier approach is becoming more common. *See* CLIENT-SERVER SYSTEM; DISTRIBUTED SYSTEMS (COMPUTERS).

Computing environment. The idea of a two- or three-tiered approach to computing hardware, and the resulting two- or three-way distribution of linked pieces of software, is relatively recent. Before such client-server or three-tiered solutions were employed, most software ran on a single computer. If the program in question was quite large, most often it would run on a mainframe computer. Mainframe computers are still used for the majority of legacy software and for some new custom software development. Mainframes are typically very large, powerful, and costly. Because mainframes have been around for many years, they provide software support often not present in the more recent computing approaches. For example, system security software, to prevent unauthorized users from accessing material intended to be secure, has matured in the mainframe environment, and is still considerably weaker in the more recent types of systems. *See* COMPUTER SECURITY.

Other kinds of computers for which application software is commonly written are minicomputers (mid-level, mid-speed, mid-price machines, typically used for mid-sized applications or as one element of a multitiered system) and microcomputers (now more commonly called personal computers or PCs), which have come to be ubiquitous and range in power and speed from modest to quite high. *See* MICROCOMPUTER.

Choice of programming language. With the advent of multiple program portions, software development has become considerably more complicated. The choice of a programming language, for example, is much more difficult. Whereas it was formerly considered sensible to develop all of a software system in the same programming language, now the different portions are often developed in entirely different languages.

The relatively complex GUI, for example, can most conveniently be developed in one of the so-called visual languages, since those languages (such as Visual Basic) contain powerful facilities for creating it. Such languages are said to be event-driven, helping the programmer write software that reacts to user-initiated events on the client computer screen.

The server software, on the other hand, will likely be built using a database package and the database language SQL (a Structured Query Language, for inquiring into the contents of a database). If the server software is also responsible for interacting with a network such as the Internet, it may also be coded in a network-support language such as Java. An object-oriented approach (which focuses on the objects to be manipulated by the program, rather than the functions to be performed or the data on which functions are to be performed) may be adopted in its development, since the software will need to manipulate objects on the Internet.

The best choice for the middle tier will probably be a language that facilitates writing software for the application domain in question. Even though COBOL and Fortran are quite old, their application focus makes them reasonable choices in this environment, if the application falls clearly into their domain. Alternatively, C and C^{++} are often used for writing the

middle tier of software. *See* COMPUTER PROGRAMMING; OBJECT-ORIENTED PROGRAMMING; SOFTWARE ENGINEERING. Robert L. Glass

Bibliography. F. P. Brooks, Jr., *The Mythical Man-Month: Essays on Software Engineering*, Addison-Wesley, 1995; T. DeMarco and T. Lister, *Peopleware: Productive Projects and Teams*, Dorset House, 1999; W. S. Humphrey, *Managing Technical People: Innovation, Teamwork, and the Software Process*, Addison-Wesley, 1997; S. McConnell, *Code Complete: A Practical Handbook of Software Construction*, Microsoft Press, 1993; M. C. Paulk et al., *The Capability Maturity Model: Guidelines for Improving the Software Process*, Addison-Wesley, 1995.

Software engineering

The process of manufacturing software systems. A software system consists of executable computer code and the supporting documents needed to manufacture, use, and maintain the code. For example, a word processing system consists of an executable program (the word processor), user manuals, and the documents, such as requirements and designs, needed to produce the executable program and manuals.

Software engineering is ever more important as larger, more complex, and life-critical software systems proliferate. The rapid decline in the costs of computer hardware means that the software in a typical system often costs more than the hardware it runs on. Large software systems may be the most complex things ever built. This places great demands on the software engineering process, which must be disciplined and controlled.

To meet this challenge, software engineers have adapted many techniques from older engineering fields, as well as developing new ones. For example, divide and conquer, a well-known technique for handling complex problems, is used in many ways in software engineering. The software engineering process itself, for example, is usually divided into phases. The definition of these phases, their ordering, and the interactions between the phases specify a software life-cycle model. The best-known life-cycle model is the waterfall model (see **illus.**) consisting of a requirements definition phase, a design phase, a coding phase, a testing phase, and a maintenance phase. The output of each phase serves as the input to the next. *See* SYSTEMS ENGINEERING.

Requirements. The purpose of the requirements phase is to define what a system should do and the constraints under which it must operate. This information is recorded in a requirements document. A typical requirements document might include a product overview; a specification of the development, operating, and maintenance environment for the product; a high-level conceptual model of the system; a specification of the user interface; specification of functional requirements; specification of nonfunctional requirements; specification of interfaces to systems outside the system under development; specification of how errors will be handled; a listing of possible changes and enhancements to the system; and other auxiliary information such as a glossary and index. Each requirement, usually numbered for reference, must be testable. The requirements definition process is difficult for several reasons.

Developing requirements requires that the customers of a system be identified and that their needs be understood. There may well be communication problems between customers who know their business lines but may be unfamiliar with computer technology, and software engineers who know computer technology but do not understand customers' business lines. Since there is a time lag between the requirements definition and the delivery of the system, the requirements of the customer may change while

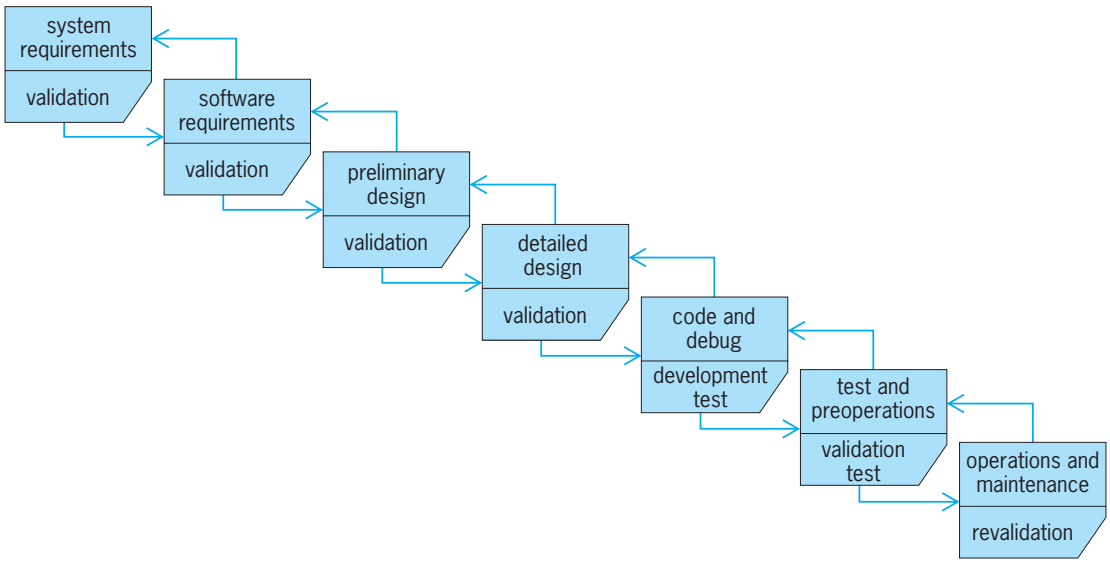

Waterfall software life-cycle model. (*After B. W. Boehm, Software engineering, IEEE Trans. Comp., C-25:1226–1241, 1975;* © *IEEE*)

the system is being built. This problem is known as requirements volatility.

Another problem stems from the language that is used to express the requirements. Requirements stated in a natural language, such as English, are expressive and understandable by users of the system but may be imprecise. Formal languages such as predicate calculus, set-theoretic notations, and finite-state machines are precise but may lack expressiveness and may be difficult or impossible for end users of the system to understand. In practice, most requirements documents are written in natural language, but formal languages are sometimes used for parts of the system. Finite-state machines, for example, are often used to specify telecommunication system activities such as call handling.

The traceability problem refers to the difficulty of cleanly mapping a given requirement to its implementation in design and code. This makes it difficult to determine which parts of a system implement a given requirement and whether a given requirement has been completely and accurately implemented.

To deal with these problems, many approaches to requirements definition have been tried. One of the more successful is the JAD (joint application design) method, which uses a focused multiday meeting of customers and developers to produce a requirements document. JAD uses rapid prototyping to help specify user interfaces. Data collected from JAD sessions indicate that significant improvements in requirements productivity and quality can be achieved using JAD.

A prototype is a partial implementation of a software system. Throwaway prototypes are built to clarify or validate requirements or to explore feasibility and then are discarded. A breadboard prototype is built as the initial part of a system and is not discarded. A key to prototyping is to use a language tool that allows rapid prototype creation. For example, interface generators allow user interfaces to be drawn, and then the system generates the code for the interface.

Design. In the design phase, a plan is developed for how the system will implement the requirements. The plan is expressed using a design method and notation. Many methods and notations for software design have been developed. Each method focuses on certain aspects of a system and ignores or minimizes others. This is similar to viewing a building with an architectural drawing, a plumbing diagram, an electrical wiring diagram, and so forth.

Principles. Principles of software design have also emerged. Abstraction, the suppression of details, is used throughout the design process to identify key design objects and operations. Separation of concerns guides the decomposition of a system into smaller parts that can be solved separately. Reusable patterns and components are used to improve system quality and reliability. Design tools support design tasks, though they primarily support the drawing of diagrams using a particular design notation and database functions, rather than guiding the design process per se.

Researchers have identified several common architectural styles for software systems. These include the pipeline architecture where the system is viewed in terms of data flowing through a pipe and transformed by data filters, data abstraction where each subsystem hides a type representation, a repository architecture where computational processes write to and read from a common data store, and a layered architecture where each architectural layer provides certain services and hides lower levels of abstraction.

A key method for handling the complexity of large software systems is modularity. A system is modular if it is composed of simple independent parts that communicate through well-defined interfaces. The conceptual simplicity of the modular system parts allows them to be more easily documented and tested and to be better maintained.

Another key method for handling design complexity is information hiding, a decomposition method where each module hides internal processing details from other modules. Each module thus has two parts; a public part with information needed to use the component, and private information such as implementation details that a user does not need to know.

Two other important design principles are coupling and cohesion. Cohesion refers to the degree to which parts of a program unit, for example functions in a module, are related. A highly cohesive module performs a single task and can usually be described by a specific verb and object like "print a file" or "sort a list." The best type of cohesion is among elements that jointly implement an abstract data type. More cohesive modules tend to be more modular and to be easier to understand and document. Coupling refers to the strength of the linkage between two modules, based on the amount and kinds of information they exchange. Minimal coupling is a design goal. The worst type of coupling is among modules that share global data (data that are defined outside the scope of individual modules). The best is coupling where modules manipulate the same data type.

Design methods. The most common design methods are structured design, data structure design, dataflow design, and object-oriented design. Each of these methods has a defined process to guide design and a set of notations for recording the design. One of the earliest methods of design focused on the flow of control in a software system based on program branching and functional activities. This information was recorded in flowcharts which provided graphical symbols for procedures, decisions, starting and stopping, and connecting pages of the diagram. This method has fallen into disfavor because flow of control is thought to be a low-level implementation detail that should be specified at the end, rather than the beginning, of the design process.

Structured design, sometimes called top-down design, begins with a high-level functional view of a system and progressively decomposes it into lower and lower levels of detail. Structure charts are one notation for recording top-down designs. This

notation provides symbols for system units at various levels of decomposition, and connectors that show how data are passed among units. Structured design is widely used and is fairly easy to understand, but fails to provide detailed guidance for system decomposition.

Dataflow design focuses on the data that flows through the system and on the processes needed to transform it. Flow of control issues are deferred until the end of the design process. Dataflow notation provides symbols for data flowing through the system, data transforms, logical alternation, and the beginning and end of data processes. Data dictionaries provide precise definitions of the data in the system, and pseudo-code is used in the final steps of the design process to record flow of control sequences. Dataflow design provides a complete method, and dataflow notation is useful in communicating system development issues to end users. The method is less appropriate for systems that are algorithmically intense, such as mathematical software. The method also provides little opportunity to incorporate abstract data types. *See* DATAFLOW SYSTEMS.

A data type is a set of data values and allowed operations on those values. For example, an integer data type has integers as its data values and the operations addition, subtraction, multiplication, and division. An abstract data type is one with implementation details abstracted away, providing an implementation-neutral mathematical abstraction. Formally defining an abstract data type involves developing a name for it, defining the sets needed for its description, defining the syntax of its operations, and defining the semantics of the operations. *See* ABSTRACT DATA TYPE; DATA STRUCTURE.

Data abstraction is a key idea in object-oriented design, which views a software system as a collection of objects which communicate via message passing. An object is an encapsulated collection of data and operations, called methods. A message consists of the name of an object, the name of one of its methods, and the parameters needed to use the method. A class is a specification of the shared features of related objects. Objects belonging to a class are said to be instances of the class. Inheritance is the process by which one class acquires the variables and operations of another. *See* OBJECT-ORIENTED PROGRAMMING.

Coding. The coding phase of the software lifecycle is concerned with the development of code that will implement the design. This code is written is a formal language called a programming language. Programming languages have evolved over time from sequences of ones and zeros directly interpretable by a computer, through symbolic machine code, assembly languages, and finally to higher-level languages that are more understandable to humans. *See* PROGRAMMING LANGUAGES.

Most coding today is done in one of the higher-level languages. When code is written in a higher-level language, it is translated into assembly code, and eventually machine code, by a compiler. Many higher-level languages have been developed, and they can be categorized as functional languages, declarative languages, and imperative languages.

Functional languages are modeled on mathematical function theory. Computations in these languages, such as Lisp and ML, are specified using the application of functions.

Declarative languages are modeled on formal logic. Computations in these languages, such as Prolog, which is based on a subset of predicate calculus, are specified by stating facts and using logical implication.

Imperative (algorithmic) languages, such as Fortran, Cobol, and C, use a precise series of instructions to specify computations. Object-oriented languages, such as C^{++} and Java, are a special kind of imperative language that adds constructs to support object-oriented programming. These constructs include class declarations and inheritance. Almost all engineering of large software systems is done using an imperative language.

Following the principle of modularity, code on large systems is separated into modules, and the modules are assigned to individual programmers. A programmer typically writes the code using a text editor. Sometimes a syntax-directed editor that "knows" about a given programming language and can provide programming templates and check code for syntax errors is used. Various other tools may be used by a programmer, including a debugger that helps find errors in the code, a profiler that shows which parts of a module spend most time executing, and optimizers that make the code run faster.

Testing. Testing is the process of examining a software product to find errors. This is necessary not just for code but for all life-cycle products and all documents in support of the software such as user manuals.

The software testing process is often divided into phases. The first phase is unit testing of software developed by a single programmer. The second phase is integration testing where units are combined and tested as a group. System testing is done on the entire system, usually with test cases developed from the system requirements. Acceptance testing of the system is done by its intended users.

Testing can be done either top-down of bottom-up. In top-down testing, high-level routines are implemented and tested and then used as a testing environment, called a test harness, for lower-level routines. Bottom-up testing proceeds by first developing low-level routines, testing them, and then progressively combining and testing them as parts of larger and larger program units. In practice, both methods are used together in a process called sandwich testing.

The basic unit of testing is the test case. A test case consists of a test case type, which is the aspect of the system that the test case is supposed to exercise; test conditions, which consist of the input values for the test; the environmental state of the system to be used in the test; and the expected behavior of the system given the inputs and environmental factors.

Test cases can be either black-box or white-box.

Black-box test cases are developed without reference to the internal structure of the code under test—that is, the code is treated as a black box. In white-box testing, the internals of the system are examined to help develop test cases.

Ideally, one would like to test all possible input-output-environment combinations, but this is impossible because such a set is typically very large, or even infinite. To address this problem, a method called equivalence partitioning is used. In equivalence partitioning, the test input and environment space are divided into classes based on common attributes. A few members of each class are then selected as representative test cases for the class.

Since it is impossible to test software exhaustively, various heuristics are used to allocate test cases to different portions of the system. One heuristic is to focus testing on the parts of the system that are most crucial. Another is to allocate testing resources on estimates of where the most errors are likely to be found, based on the size or complexity of system parts or on the parts that consume the most processing time.

Another practical problem is knowing when enough testing has been done and the product is ready to be released. This can be addressed using a cumulative fault plot, a bivariate graph of the cumulative faults found via testing versus time. Such a plotted curve typically displays an S shape since at first few faults are found as the testing process gears up, then the curve rises steeply as most of the errors are found, and finally the curve flattens again since the remaining errors are more difficult to detect. A system is considered ready to ship when the final part of the curve remains consistently flat.

In white-box testing, the aim is to ensure that all of the code has been tested. This measurement can be in terms of lines of code, branch points, or execution paths through the program. Tools to support white-box testing instrument (add tracking code to) assembly or source code in order to keep track of which parts of the code have been exercised by test cases and which have not.

The coverage analysis process uses this information to design test cases that cumulatively test the code to the desired percentage. The process is to first examine the code to see which parts have not been exercised, design test cases to exercise the untested code, run the test cases, and then run a coverage analysis tool. If the desired level of test coverage has been reached, the process ends; otherwise, the process is repeated.

When software is changed to fix a bug or add an enhancement, a serious error is often introduced. To ensure that this does not happen, all test cases must be rerun after each change. The process of rerunning test cases to ensure that no error has been introduced is called regression testing. Various tools, including Unix shell scripts, can be used to support regression testing. *See* SOFTWARE TESTING AND INSPECTION.

Walkthroughs and inspections. Walkthroughs and inspections are used to improve the quality of the software development process. Consequently, the software products created by the process are improved. A quality system is a collection of techniques whose application results in continuous improvement in the quality of the development process. Elements of the quality system include reviews, inspections, and process audits.

A quality system should address two different aspects of quality: process and product quality. Process quality is the fitness of the development process for producing a product, while product quality is the fitness for use of the product resulting from some process. Quality systems should answer two basic questions about process and product quality: Are we building the right product (assessing product quality)? Are we building the product right (assessing process quality)?

Many of the things that must be tested in the software engineering process are not amenable to the kinds of techniques discussed for software. For the many textual documents that are typically produced for a software product and for software engineering processes, the best available testing methods are walkthroughs and inspections. Since code is also a kind of document, it can be tested with walkthroughs and inspections. Several types of walkthroughs and inspections are used.

A software review is an objective evaluation of a software engineering product or process. Though informal reviews, lacking systematic procedures and structures, may be adequate for small simple projects, larger projects will benefit from formal reviews. According to data collected on many projects, inspections can detect errors earlier in the life-cycle. This is important because the cost to repair errors rises rapidly as the software life-cycle progresses. Earlier detection via inspections can give up to 25% improvements in productivity. Other data indicate that defects found by inspections can reduce costs tenfold.

Code inspections are group readings of code by its designer, programmer, and tester. A moderator manages the inspection. During inspection, the programmer reads the code aloud and describes the intent of each statement. Errors are documented and categorized, but no attempt is made to fix the error during the inspection; this is done later. Average inspection rates for code are about 90 noncommentary source lines per hour. The inspection process can be guided by an error checklist of kinds of errors that might be encountered such as data declaration errors, computation errors, and control flow errors.

Another kind of inspection is a process audit whose purpose is to improve quality and productivity by improving the software production process. The best-known process evaluation for software is the SEI (Software Engineering Institute) process capability maturity model (CMM). This model is based on Philip Crosby's model for general process quality improvement. Crosby's model is based, in turn, on one of the basic principles of modern quality theory, which is that the best way to assure the quality of a product is to control the process which produces it. *See* QUALITY CONTROL.

The CMM is an evaluation model that classifies a process as being at one of five levels: (5) Optimizing — constant improvement based on measurement and feedback. (4) Managed—measured process and quality and productivity. (3) Defined—process defined and institutionalized. (2) Repeatable—process dependent on individuals, but basic project controls established. (1) Initial—no formal procedures or management control. The highest level, 5, includes continual process improvement, another hallmark of modern quality theory. Experience has shown that most software organizations today are at level 1, and only a very few are at level 5.

Maintenance. Large software systems are not static; rather, they change frequently both during development and after deployment. Maintenance is the phase of the software life-cycle after deployment. The maintenance phase may cost more than all of the others combined and is thus of primary concern to software organizations. The Y2K problem was, for example, a maintenance problem.

Activities. Maintenance consists of three activities: adaptation, correction, and enhancement. Enhancement is the process of adding new functionality to a system. This is usually done at the request of system users. This activity requires a full life-cycle of its own. That is, enhancements demand requirements, design, implementation, and test. Studies have shown that about half of maintenance effort is spent on enhancements.

Adaptive maintenance is the process of changing a system to adapt it to a new operating environment, for example, moving a system from the Windows operating system to the Linux operating system. Adaptive maintenance has been found to account for about a quarter of total maintenance effort. Corrective maintenance is the process of fixing errors in a system after release. Corrective maintenance takes about 20% of maintenance effort.

Software configuration management. Since software systems change frequently over time, an important activity is software configuration management. This consists of tracking versions of life-cycle objects, controlling changes to them, and monitoring relationships among them. Configuration management activities include version control, which involves keeping track of versions of life-cycle objects; change control, an orderly process of handling change requests to a system; and build control, the tracking of which versions of work products go together to form a given version of a software product.

Version control is supported by version control tools such as the Revision Control System (RCS) and Source Code Control System (SCCS). These tools support the storage and retrieval of versions of text files. To use them, a file is put into the change control system. After a version of a file is retrieved and modified, it is resubmitted to the system, which calculates changes between the old and new versions and stores only the changes, rather than both complete versions. This greatly reduces storage overhead. The system assigns unique numbers to versions and variations, allowing an orderly tracking process. A version replaces the original file, and a variation is an alternative to the original file.

Change control must be done in an orderly way on a large project. When a change is requested to some life-cycle work product, a form, sometimes called a modification request (MR), is created and logged into a change control database. A review board decides the severity of the problem described in the modification request and how the modification request should be handled, and then assigns personnel to address the handling of it.

Build control is concerned with how to put various versions of life-cycle work products together to build a system or part of a system. The make tool on UNIX, for example, allows the specification of which versions of software modules are needed to build an executable program, and the commands to actually do the build. In the C language, for example, an executable file depends on the object and library files needed to generate it. These, in turn, depend on source code files written in the C language, which are used by the compiler to create the object and library files.

Software reuse. The methods described so far, while effective in many cases, may not be adequate for building systems of the size and complexity sometimes required. One method that has been proposed to address this problem is software reuse, that is, the use of knowledge or artifacts from previous systems to build new ones. While software reuse has always been done in some way, recent research has attempted to get large productivity and quality gains by focusing on reuse that is systematic and domain-based.

A key idea in systematic domain-based reuse is that organizations seldom build completely new software systems. Rather, they build similar systems with variants to address different customer needs. These similar systems are said to be in a domain, which may be thought of informally as a business line. For example, a company's domain may be word processing software. The company will produce new versions with added functionality over time and may well provide variants for different operating systems.

Part of the impetus for more software reuse has come from considering other engineering disciplines. In electrical engineering, for example, systems are constructed from standardized components such as integrated circuits, resistors, and capacitors. A study of the history of hardware manufacturing has shown that that field made the transition from craft-based manufacturing to manufacturing based on interchangeable parts around 1840. Software production is in this transition process now.

Software reuse can be either horizontal or vertical. Horizontal reuse is reuse across domains. Much reuse is horizontal. For example, standard function libraries that support activities like searching and sorting are of this kind. Vertical reuse is reuse within a domain. For example, a component library might be

developed specifically for word processing systems that includes modules to handle text buffers, font management, and so on.

Actions that can produce more reuse fall into four categories: economic, legal, technical, and managerial. Clearly reuse improvement is undertaken to produce some benefit. The bottom-line benefit for a company will be greater profits. More immediate benefits, presumed to improve profits, include higher productivity, better quality, faster time to market, and improved bid and late life-cycle estimation. There is evidence that reuse can indeed produce these results under the right circumstances. These circumstances will be different for different organizations, requiring the proper blend of managerial, technical, legal, and economic factors.

Managerial factors include a rewards program for producers and consumers of reusable assets. Making an asset reusable will require more effort than making an equivalent one-use asset. This additional cost must be amortized across the multiple uses of the component. Incentives for producers of components must be established. Some companies, for example, provide cash bonuses for reusable assets accepted by the company reuse library. Since most software engineers are not currently trained to do systematic reuse and domain engineering, management must also provide education and motivation.

Legal factors may become the greatest barrier to systematic reuse. Current laws regarding software copyrights, patents, and liabilities are unclear. For example, if a company sells an asset and it is integrated into an application that fails because of the asset, causing loss of life or property, does the purchasing organization have the right to sue for damages? Another problem is that many software algorithms are now being patented each year. This may, in the future, provide a disincentive to reuse.

Technical factors include the issues of how to create reusable assets, how to design new applications to take advantage of them, and what kind of software engineering environment is needed to support reuse. Other technical factors concern what kind of software development process is needed and how to create and maintain a reuse library.

Economic factors include how to do a cost benefit analysis of reuse, performance measurement of a reuse program, methods for system costing, product pricing criteria and strategies to take reuse into account, and how to finance support costs for a reuse program.

A reuse library consists of a database for storing reusable components and a search interface that allows users to find and retrieve items in the database. Components for the database can be purchased, built, or reengineered from existing software. Before being added to the library, the components must be certified to assure that they meet quality standards, and must be classified.

The most common method in classifying reusable components is enumerated classification, where the subject area is broken into mutually exclusive, usually hierarchical classes. Other classification methods are faceted classification where the subject area is analyzed into basic terms which are organized as attribute-value pairs, and free-text keyword where index terms are extracted directly for the reusable component description. Reuse library tools should ideally be integrated with tools for configuration management, since the reusable assets in the database will change as error corrections and enhancements are made.

Domain engineering. Domain engineering is the process of analyzing a domain, recording important information about the domain in domain models, creating reusable assets, and using the domain models and assets to create new systems in the domain. It has two phases: domain analysis and domain implementation.

Domain analysis is the process of finding common and variable parts of systems in a domain and constructing domain models. Code, documents, and expert information are used as inputs to the domain analysis process. A domain vocabulary is developed which is used to describe systems in the domain. A key model that must be developed is a generic architecture that describes all of the systems in the domain. Reusable components will also often be identified in the domain analysis process.

These models are used to develop tools to speed the process of developing systems in the domain. One tool is domain-specific components that can be integrated into systems in the domain. Domain-specific programming languages can also be developed. These languages can be either large complete languages that include domain-specific operators and operands, or little languages that are small domain-specific languages developed for a specific purpose. Another approach, if sufficient formal information about the domain is known, is to develop application generators.

Application generators are programs that produce all or most of a software system when given a high-level specification. Application generators can be particularly powerful tools for reuse, giving great quality and productivity improvements. To develop an application generator, a language must be developed that allows the desired system to be specified. A specification in this language is then passed to the generator that produces the code for the application in the desired target language. Lex, for example, is an application generator for lexical analyzers. The specifications for a desired lexical analyzer are written in the lex language. The lex generator then produces the C code to implement the lexical analyzer. *See* SOFTWARE.

William B. Frakes

Bibliography. B. I. Blum, *Software Engineering: A Holistic Approach*, Oxford University Press, 1992; E. J. Braude, *Software Engineering: An Object-Oriented Perspective*, Wiley, 2000; M. Dorfman and R. H. Thayer (eds.), *Software Engineering*, IEEE Computer Society, 1996; R. S. Pressman, *Software Engineering: A Beginner's Guide*, McGraw-Hill, 1988; R. S. Pressman, *Software Engineering: A*

Practitioner's Approach, 5th ed., McGraw-Hill, 2000; I. Sommerville, *Software Engineering*, 6th ed., Addison-Wesley, 2000.

Software metric

A rule for quantifying some characteristic or attribute of a computer software entity. For example, a simple one is the FileSize metric, which is the total number of characters in the source files of a program. The FileSize metric can be used to determine the measure of a particular program, such as 3K bytes. It provides a concrete measure of the abstract attribute of program size. Of course, there are better size metrics. The FileSize metric is used for source programs; other metrics can be used for software entities such as requirements documents, design object models, or database structure models. Metrics for requirements and design documents can be used to guide decisions about development and as a basis for predictions, such as for cost and effort. Metrics for programs can be used to support decisions about testing and maintenance and as a basis for comparing different versions of programs. Ideally, metrics for the development cost of software and for the quality of the resultant program are desirable; but, lacking those, metrics for size and complexity are commonly used.

The fundamental property that must be satisfied by all software metrics is the representation condition. It requires that if one entity is less than another entity in terms of a selected attribute, then any metric for that attribute must associate a smaller number to the first entity than it associates with the second entity. For example, the FileSize metric is not truly a metric since it does not meet the representation condition. It is possible to take a "small" program and replace all variables with very long names, so that the measure of FileSize of this "small" program is larger than that of some "larger" program that uses short variable names.

Size metrics. The earliest size metric for programs was Lines of Code (LOC). Unfortunately, LOC, although commonly used, is not a good predictor of cost of development, likelihood of errors, and so forth. Additionally, there is no standard way to measure LOC; variations include counting or not counting blank lines, comments, declarations, and extra lines used for formatting.

Several metrics that better define program size have been introduced. The number of unique operators, η_1 (operators are symbols such as + and function and procedure names), and unique operands, η_2 (generally operands are the constants and identifiers in a program), can be used to predict software size according to the equation below. Similar metrics

$$N = \eta_1 \log_2 \eta_1 + \eta_2 \log_2 \eta_2$$

that relate to complexity have been defined, but they have not been widely used.

Complexity metrics. These metrics are intended to provide a quantitative measure of the abstract attribute of complexity. Intuitively, if one software entity is more complex than another software entity, then it follows that the more complex entity may have more errors, may be harder to understand, or may be more difficult to maintain. Unfortunately, complexity metrics provide only an imperfect measurement of the attribute of complexity.

A commonly used complexity metric for programs is based on the cyclomatic number used in graph theory. The cyclomatic number on the control flow graph of a program or design pseudocode has been shown to be equal to one plus the number of decisions in the control graph. A good rule of thumb is that program modules with cyclomatic number greater than ten should be avoided; that is, these programs need to be rewritten.

Whereas the Cyclomatic Number metric is intended to measure complexity of single program units (such as a function, procedure, or method), a number of metrics have been introduced to measure the complexity of object models and programs. In object-oriented software, objects can be inherited from other objects, and one metric is the length of these inheritance lists. Objects can also call functions in other objects, and the number of these calls can be measured. Since functions within an object use data local to the object, lack of cohesion can be measured by noting how many pairs of functions do not use any common data. *See* OBJECT-ORIENTED PROGRAMMING.

Software requirements metric. The Function Points (FP) metric has been developed for application to software requirements. It is used to predict the size and cost of development of a proposed software entity. The developer identifies and counts all inputs, outputs, inquiries, files, and interfaces and rates each as simple, average, or complex. Additional characteristics of the project are ranked on a scale of 1 to 5.

Types of scales. The area of measure theory specifies five possible types of measurement scales: nominal, ordinal, interval, rational, and absolute. Every software metric must be one of the five types of scales. The type of scale determines the types of statistics that can be applied to the metric. A metric must be at least an interval scale to take averages and must be at least a rational scale to consider ratios between values. For example, the cyclomatic number is an interval scale, and thus it would be correct to calculate the average cyclomatic number of a set of modules. However, it could not be correctly said that one module has twice the cyclomatic number of another module. For example, if a new module was created from two copies of a module, the new module would not have twice the cyclomatic number of the original module. *See* MEASURE THEORY.

Goal-Question metric (GQM). This is a method for developing metrics that will be more acceptable to developers. The first step is to select goals that should be maximized or minimized. For example, a goal

might be customer satisfaction. Next, questions are developed that are related to the goals—for example, "Is the customer finding too many errors?" Finally, metrics are proposed that quantify answers to the questions. For example, a metric could be the number of customer-reported errors. Since the metrics are closely related to the goals, use of the metrics should contribute to achievement of the goal. *See* COMPUTER PROGRAMMING; SOFTWARE ENGINEERING.

David A. Gustafson; William Hankley

Bibliography. N. Fenton, *Software Metrics*, Chapman and Hall, 1991; A. Melton (ed.), *Software Measurement*, International Thomson Computer Press, 1995; H. Zuse, *Software Complexity Measures and Methods*, W. de Gruyter, 1991.

Software testing and inspection

When software does not operate as it is intended to do, a software failure is said to occur. Software failures are caused by one or more sections of the software program being incorrect. Each of these incorrect sections is called a software fault. The fault could be as simple as a wrong value. A fault could also be the omission of a decision in the program. Faults have many causes, including misunderstanding of requirements, overlooking special cases, using the wrong variable, misunderstanding of the algorithm, and even typing mistakes. Software that can cause serious problems if it fails is called safety-critical software. Many applications in aircraft, medicine, nuclear power plants, and transportation involve such software.

Software testing is the execution of the software with the purpose of detecting faults. Software inspection is a manual process of analyzing the source code to detect faults. Many of the same techniques are used in both procedures. Other techniques can also be used to minimize the possibility of faults in the software. These techniques include the use of formal specifications, formal proofs of correctness, and model checking. However, even with the use of these techniques, it is still important to execute software with test cases to detect possible faults.

Software testing involves selecting test cases, determining the correct output of the software, executing the software with each test case, and comparing the actual output with the expected output. More testing is better, but costs time and effort. The value of additional testing must be balanced against the additional cost of effort and the delay in delivering the software. Another consideration is the potential cost of failure of the software. Safety-critical software is usually tested much more thoroughly than any other software.

The number of potential test cases is huge. For example, in the case of a simple program that multiplies two integer numbers, if each integer is internally represented as a 32-bit number (a common size for the internal representation), then there are 2^{32} possible values for each number. Thus, the total number of possible input combinations is 2^{64}, which is more than 10^{19}. If a test case can be done each microsecond (10^{-6} second), then it will take thousands of years to try all of the possible test cases. Trying all possible test cases is called exhaustive testing and is usually not a reasonable approach because of the size of the task, as this example indicates.

A test case should include the values of the inputs and the expected values of all outputs. In the following example, a test case includes three numbers that represent the inputs and a string that represents the value of the output. For example, (3 3 3 "equilateral") is a test case where the inputs are 3, 3, and 3, and the expected output is "equilateral."

The triangle problem, introduced by Glenford Myers, will be used to illustrate software testing. The problem is a simple one: Given the lengths of three sides of a triangle, tell whether the triangle is scalene, isosceles, or equilateral. The C++ code for one solution is as follows:

```
float a,b,c;
cin ≪ a ≪ b ≪ c;
char* type;
type = "scalene";
if (a==b || a==c || b==c)
{
        type = "isosceles";
}
if (a==b && a==c)
{
        type = "equilateral";
}
if (a>=b+c || b>=a+c || c>=a+b)
{
        type = "not a triangle";
}
if (a<=0 || b<=0 || c<=0)
{
        type = "bad inputs";
}
cout<<type;
```

In this solution, the three triangle types are possible outputs, but there are also two error conditions: The inputs may fail to be positive numbers, and the values may fail to be values of the sides of a legitimate triangle. The code has the property that the last "if" statement that is true for a particular test case is the one that determines the answer. For example, all of the "if" statements are true for the inputs 0 0 0, therefore the answer is "bad inputs."

Most of the testing criteria are easier to visualize using the control flow graph shown in the **illustration**. Each circle on the graph represents a block of code in the program. Each of these blocks is a section of code without any decisions inside the block. Thus, node *a* represents the code up to the logical expression of the first "if" statement. Node *b* represents the assignment of "equilateral" to the variable type. The arcs between the nodes represent the

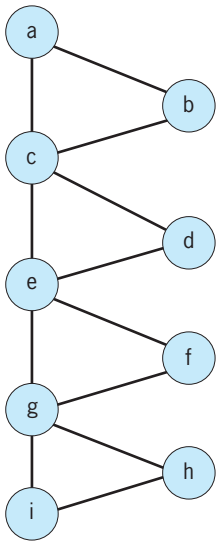

Control flow diagram for the triangle problem.

possible flow of control. The arrow from node *a* to node *b* represents the flow of control if the first "if" statement is true. The arrow from node *a* to node *c* represents the flow of control if the first "if" statement is false.

Structural testing. One approach to software testing is to find test cases so that all statements in a program are executed. For example, the test case (3 4 5 "scalene") will execute all the statements except the assignment of "isosceles," "equilateral," "not a triangle," and "bad inputs" (nodes *a*, *c*, *e*, *g*, and *i* are covered). Additional test cases will be required to execute those assignments. This approach is a type of structural testing and is called "every statement coverage." Structural test methods use the code to select test cases. In this case the criterion is to execute all statements. "Every statement coverage" is often considered the minimal acceptable coverage for software testing. This coverage does not guarantee that every fault that exists in the program will be detected.

A more extensive criterion for test selection is "every branch coverage." This means that each branch coming out of every decision is tested. On the control flow graph "every branch coverage" is satisfied by executing every arc. A minimal test set that achieves this coverage would be (0 0 0 "bad inputs") and (3 4 5 "scalene").

Instead of just requiring the whole decision to be true or false, the "multiple condition coverage" criterion requires all combinations of truth values for each simple comparison in a decision to be covered. For example, the first decision statement is true if *a* equals *b* or *a* equals *c* or *b* equals *c*. The test case (4 3 3 "isosceles") makes the first comparison true (TFF). Because of how compilers handle multiple conditions, if a condition of a set of "or-ed" conditions is true, the conditions to the right will not be checked. Thus, this test case covers TFF, TTF, TFT, and TTT. Test case (3 4 3 "isosceles") covers FTF and FTT; (4 3 3 "isosceles") covers FFT; and (3 4 5 "scalene") covers FFF. (If two of the comparisons are true, then the third comparison is also true.)

Another approach is called dataflow coverage. The basis for coverage is the execution paths between the statement where a variable is assigned a value (a def or definition) and a statement where that value is used. These paths must be free of other definitions of the variable of interest. The basic criterion is to cover all def-use pairs. There are a number of variations. In the triangle problem, the variable of interest is type. There are five definitions of the variable type and one use of the variable type (the cout statement). Thus, test cases must be found for the paths *acegi*, *bcegi*, *degi*, *fgi*, and *hi*.

Functional testing. Testing compares the actual behavior of the software with the expected behavior. That expected behavior is usually described in a specification. In fact, software without a specification of what it should do could be considered correct because it does what it does. In functional testing, the test case selection is based on information found in the specification. In the triangle example, the specification says that software is supposed to identify the type of triangle. A very simple but effective approach is to ensure that all possible outputs are generated by the test set. This would require that the test set include an equilateral, an isosceles, and a scalene test case.

More involved functional test case selection involves analyzing the conditions inherent in the task. For the triangle problem, these conditions would be the definition of equilateral, isosceles, scalene, and being a triangle. For example, being a triangle requires that each side is less than the sum of the other two sides. Each valid combination of all these conditions would be considered to be a subdomain that should be tested. The test criterion would be to include at least one test case from each subdomain.

Boundary testing. Another approach to test selection concentrates on the boundaries between the subdomains. This approach recognizes that many faults are related to the boundary conditions. Test cases are chosen to check whether the boundary is correct. Test cases on the boundary and test cases just off the boundary are chosen. An example in the triangle problem is not-a-triangle decision. One boundary is $a = b + c$. Test cases such as (6 4 2 "not a triangle") and (100 50 50 "not a triangle") might be chosen on the boundary. An off-the-boundary test case might be (51 25 25 "isosceles").

Prospects. None of these approaches guarantees that faults will be detected. Current research is aimed at measuring the effectiveness of these testing approaches and investigating ways to improve the fault detection ability. Work is also being done on quantifying the effectiveness of testing in terms of predicting the reliability of software that has been successfully tested with a specific technique. *See* SOFTWARE; SOFTWARE ENGINEERING. David A. Gustafson

Bibliography. B. Beizer, *Software Testing Techniques*, 2d ed., Van Nostrand Reinhold, 1990; D. Gustafson, *Schaum's Outline of Theory and*

Problems of Software Engineering, McGraw-Hill, 2001; B. Marick, *The Craft of Software Testing*, Prentice Hall, 1995; G. Myers, *The Art of Software Testing*, Wiley, 1979.

Soil

Finely divided rock-derived material containing an admixture of organic matter and capable of supporting vegetation. Soils are independent natural bodies, each with a unique morphology resulting from a particular combination of climate, living plants and animals, parent rock materials, relief, the ground waters, and age. Soils support plants, occupy large portions of the Earth's surface, and have shape, area, breadth, width, and depth. (However, the term "soil" as used by engineers means unconsolidated rock material.) *See* PEDOLOGY.

Origin and Classification

Soil covers most of the land surface as a continuum. Each soil grades into the rock material below and into other soils at its margins, where changes occur in relief, ground water, vegetation, kinds of rock, or other factors which influence the development of soils. Soils have horizons, or layers, more or less parallel to the surface and differing from those above and below in one or more properties, such as color, texture, structure, consistency, porosity, and reaction (**Fig. 1**). The horizons may be thick or thin. They may be prominent or so weak that they can be detected only in the laboratory. The succession of horizons is called the soil profile. In general, the boundary of soils with the underlying rock or rock material occurs at depths ranging from 1 to 6 ft (0.3 to 1.8 m), though the extremes lie outside this range.

Origin. Soil formation proceeds in stages, but these stages may grade indistinctly from one into another. The first stage is the accumulation of unconsolidated rock fragments, the parent material. Parent material may be accumulated by deposition of rock fragments moved by glaciers, wind, gravity, or water, or it may accumulate more or less in place from physical and chemical weathering of hard rocks. *See* WEATHERING PROCESSES.

The second stage is the formation of horizons. This stage may follow or go on simultaneously with the accumulation of parent material. Soil horizons are a result of dominance of one or more processes over others, producing a layer which differs from the layers above and below.

Major processes. The major processes in soils which promote horizon differentiation are gains, losses, transfers, and transformations of organic matter, soluble salts, carbonates, silicate clay minerals, sesquioxides, and silica. Gains consist normally of additions of organic matter, and of oxygen and water through oxidation and hydration, but in some sites slow continuous additions of new mineral materials take place at the surface or soluble materials are deposited from ground water. Losses are chiefly of materials dissolved or suspended in water percolating through the profile or running off the surface. Transfers of both mineral and organic materials are common in soils. Water moving through the soil picks up materials in solution or suspension. These materials may be deposited in another horizon if the water is withdrawn by plant roots or evaporation, or if the materials are precipitated as a result of differences in pH (degree of acidity), salt concentration, or other conditions in deeper horizons.

Other processes tend to offset those that promote horizon differentiation. Mixing of the soil occurs as the result of burrowing by rodents and earthworms, overturning of trees, churning of the soil by frost, or shrinking and swelling. On steep slopes the soil may creep or slide downhill with attendant mixing. Plants may withdraw calcium or other ions from deep horizons and return them to the surface in the leaf litter.

Saturation of a horizon with water for long periods makes the iron oxides soluble by reduction from ferric to ferrous forms. The soluble iron can move by diffusion to form hard concretions or splotches of red or brown in a gray matrix. Or if the iron remains, the soil will have shades of blue or green. This process is called gleying, and can be superimposed on any of the others. *See* DIFFUSION.

The kinds of horizons present and the degree of their differentiation, both in composition and in structure, depend on the relative strengths of the processes. In turn, these relative strengths are determined by the way humans use the soil as well as by the natural factors of climate, plants

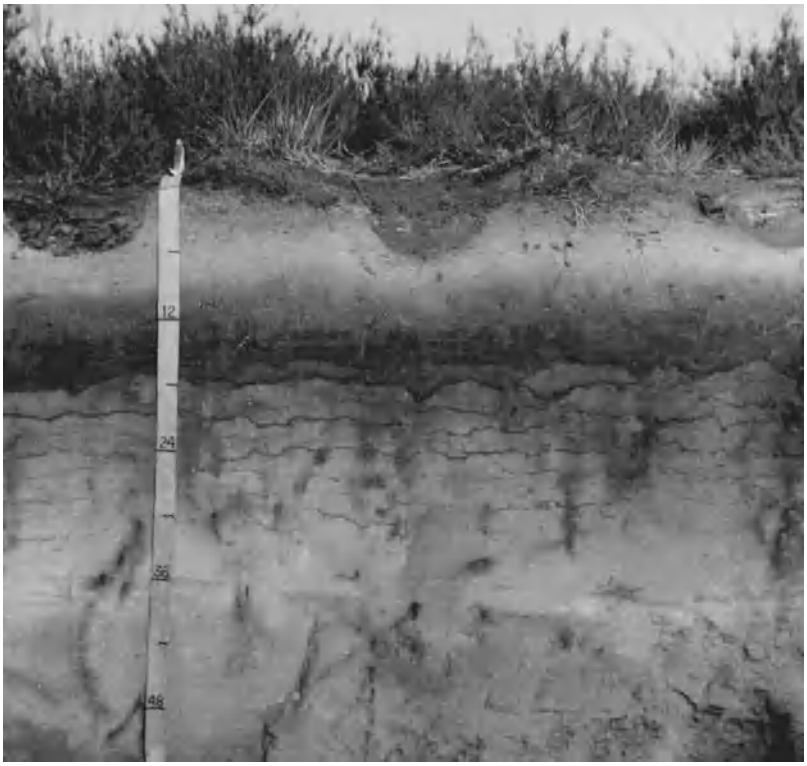

Fig. 1. Photograph of a soil profile showing horizons. The dark crescent-shaped spots at the soil surface are the result of plowing. The dark horizon lying 9–18 in. (23–45 cm) below the surface is the principal horizon of accumulation of organic matter that has been washed down from the surface. The thin wavy lines were formed in the same manner.

and animals, relief and ground water, and the period of time during which the processes have been operating.

Composition. In the drier climates where precipitation is appreciably less than the potential for evaporation and transpiration, horizons of soluble salts, including calcium carbonate and gypsum, are often found at the average depth of water penetration.

In humid climates, some materials normally considered insoluble may be gradually removed from the soil or at least from the surface horizons. A part of the removal may be in suspension. The movement of silicate clay minerals is an example. The movement of iron oxides is accelerated by the formation of chelates with the soil organic matter. Silica is removed in appreciable amounts in solution or suspension, though quartz sand is relatively unaffected. In warm humid climates, free iron and aluminum oxides and low-activity silicate clays accumulate in soils, apparently because of low solubility relative to other minerals. *See* CHELATION.

In cool humid climates, solution losses are evident in such minerals as feldspars. Free sesquioxides tend to be removed from the surface horizons and to accumulate in a lower horizon, but mixing by animals and falling trees may counterbalance the downward movement.

Structure. Concurrently with the other processes, distinctive structures are formed in the different horizons. In the surface horizons, where there is a maximum of biotic activity, small animals, roots, and frost action keep mixing the soil material. Aggregates of varying sizes are formed and bound by organic matter, microorganisms, and colloidal material. The aggregates in the immediate surface tend to be loosely packed with many large pores among them. Below this horizon of high biotic activity, the structure is formed chiefly by volume changes due to wetting, drying, freezing, thawing, or shaking of the soil by roots of trees swaying with the wind. Consequently, the sides of any one aggregate, or ped, conform in shape to the sides of adjacent peds.

Water moving through the soil usually follows root channels, wormholes, and ped surfaces. Accordingly, materials that are deposited in a horizon commonly coat the peds. In the horizons that have received clay from an overlying horizon, the peds usually have a coating or varnish of clay making the exterior unlike the interior in appearance. Peds formed by moisture or temperature changes normally have the shapes of plates, prisms, or blocks.

Horizons. Pedologists have developed sets of symbols to identify the various kinds of horizons commonly found in soils. The nomenclature originated in Russia, where the letters A, B, and C were applied to the main horizons of the black soils of the steppes. The letter A designated the dark surface horizon of maximum organic matter accumulation, C the unaltered parent material, and B the intermediate horizon. The usage of the letters A, B, and C spread to western Europe, where the intermediate or B horizon was a horizon of accumulation of free sesquioxides or silicate clays or both. Thus the idea developed that a B horizon is a horizon of accumulation. Some, however, define a B horizon by position between A and C. Subdivisions of the major horizons have been shown by either numbers or letters, for example, Bt or B2. No internationally accepted set of horizon symbols has been developed. In the United States the designations (**Fig. 2**) have been widely used since about 1935, with minor modifications made in 1962. Lowercase letters were added to numbers in B horizons to indicate the nature of the material that had accumulated. Generally, "h" is used to indicate translocated humus, "t" for translocated clay, and "ir" for translocated iron oxides. Thus, B2t indicates the main horizon of clay accumulation.

Classification. Systems of soil classification are influenced by concepts prevalent at the time a system is developed. Since ancient times, soil has been considered as the natural medium for plant growth. Under this concept, the earliest classifications were based on relative suitability for different crops, such as rice soils, wheat soils, and vineyard soils.

Early American agriculturists thought of soil chiefly as disintegrated rock, and the first comprehensive American classification was based primarily on the nature of the underlying rock.

In the latter part of the nineteenth century, some Russian students noted relations between the steppe and black soils and the forest and gray soils. They developed the concept of soils as independent natural bodies formed by the influence of environmental factors operating on parent materials over time. The early Russian classifications grouped soils at the highest level, according to the degree to which they reflected the climate and vegetation. They had classes of Normal, Abnormal, and Transitional soils, which later became known as Zonal, Intrazonal, and Azonal. Within the Normal or Zonal soils, the Russians distinguished climatic and vegetative zones in which the soils had distinctive colors and other properties in common. These formed classes that were called soil types. Because some soils with similar colors had very different properties that were associated with differences in the vegetation, the nature of the vegetation was sometimes considered in addition to the color to form the soil type name, for example, Gray Forest soil and Gray Desert soil. The Russian concepts of soil types were accepted in other countries as quickly as they became known. In the United States, however, the term "soil type" had been used for some decades to indicate differences in soil texture, chiefly texture of the surface horizons; so the Russian soil type was called a Great Soil Group.

Many systems of classification have been attempted but none has been found markedly superior; most systems have been modifications of those used in Russia. Two bases for classification have been tried. One basis has been the presumed genesis of the soil; climate and native vegetation were given major emphasis. The other basis has been the observable or measurable properties of the soil. To a

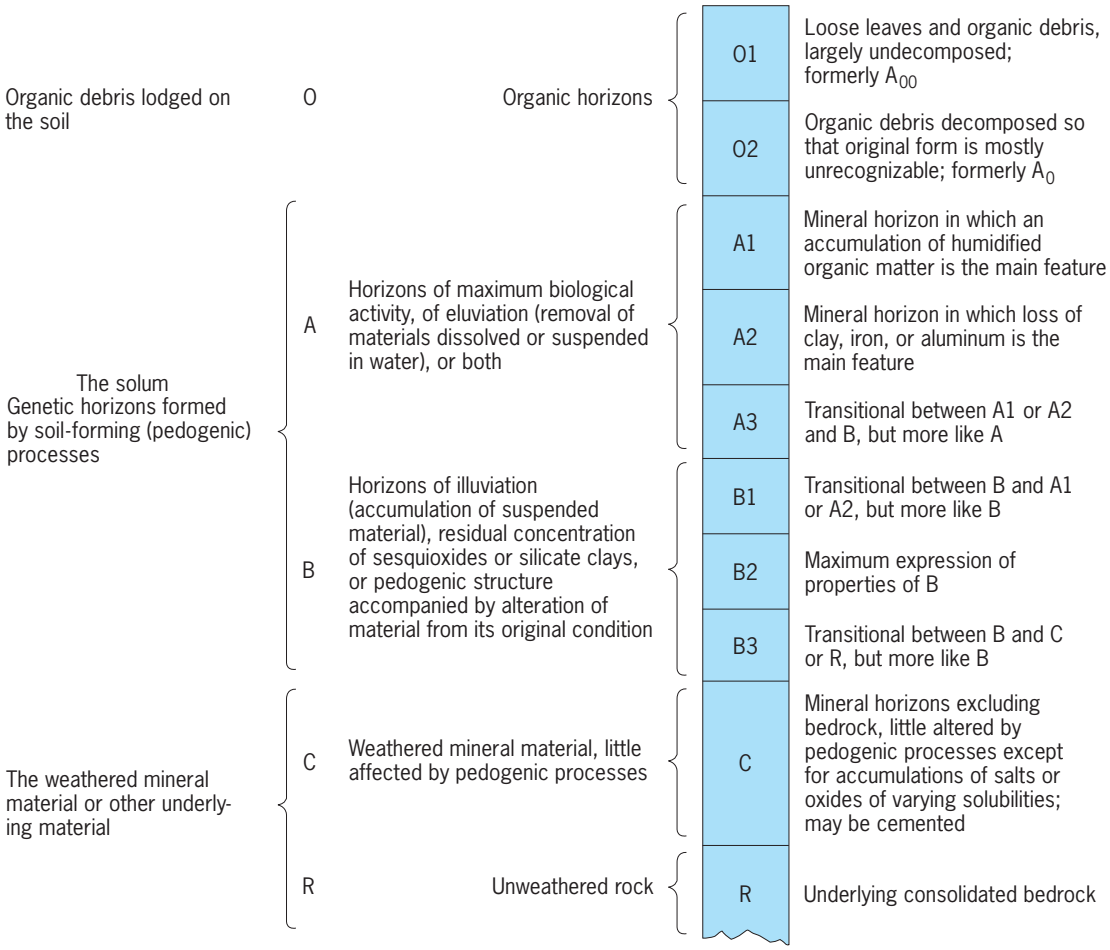

Fig. 2. Hypothetical soil profile having all principal horizons. Other symbols are used to indicate features subordinate to those indicated by capital letters and numbers. The more important of these are as follows: ca, as in Cca, accumulations of carbonates; cs, accumulations of calcium sulfate; cn, concretions; g, strong gleying (reduction of iron in presence of ground water); h, illuvial humus; ir, illuvial iron; m, strong cementation; p, plowing; sa, accumulations of very soluble salts; si, cementation by silica; t, illuvial clay; x, fragipan (a compact zone which is impenetrable by roots).

considerable extent, of course, these are used in the genetic system to define the great soil groups. The morphologic systems, however, have not used soil genesis as such, but have attempted to use properties that are acquired through soil development.

The principal problem in the morphologic systems has been the selection of the properties to be used. Grouping by color, tried in the earliest systems, produces soil groups of unlike genesis.

The Soil Survey staff of the U.S. Department of Agriculture (USDA) and the land-grant colleges adopted a different classification scheme in 1965. The system differs from earlier systems in that it may be applied to either cultivated or virgin soils. Previous systems have been based on virgin profiles, and cultivated soils were classified on the presumed characteristics or genesis of the virgin soils. The system has six categories, based on both physical and chemical properties. These categories are the order, suborder, great group, subgroup, family, and series, in decreasing rank.

Nomenclature. The names of the taxa or classes in each category are derived from the classic languages in such a manner that the name itself indicates the place of the taxa in the system and usually indicates something of the differentiating properties. The names of the highest category, the order, end in the suffix "sol," preceded by formative elements that suggest the nature of the order. Thus, Aridisol is the name of an order of soils that is characterized by being dry (Latin *arudys*, dry, plus *sol*, soil). A formative element is taken from each order name as the final syllable in the names of all taxa of suborders, great groups, and subgroups in the order. This is the syllable beginning with the vowel that precedes the connecting vowel with "sol." Thus, for Aridisols, the names of the taxa of lower classes end with the syllable "id," as in Argid and Orthid (**Table 1**).

Suborder names have two syllables, the first suggesting something of the nature of the suborder and the last identifying the order. The formative element "arg" in Argid (Latin *argillus*, clay) suggests the horizon of accumulation of clay that defines the suborder.

Great group names have one or more syllables to suggest the nature of the horizons and have the suborder name as an ending. Thus great group names

TABLE 1. Soil orders

Order	Formative element in name	General nature
Alfisols	alf	Soils with gray to brown surface horizons, medium to high base supply, with horizons of clay accumulation; usually moist, but may be dry during summer
Aridisols	id	Soils with pedogenic horizons, low in organic matter, and usually dry
Entisols	ent	Soils without pedogenic horizons
Histosols	ist	Organic soils (peats and mucks)
Inceptisols	ept	Soils that are usually moist, with pedogenic horizons of alteration of parent materials but not of illuviation
Mollisols	oll	Soils with nearly black, organic-rich surface horizons and high base supply
Oxisols	ox	Soils with residual accumulations of inactive clays, free oxides, kaolin, and quartz; mostly tropical
Spodosols	od	Soils with accumulations of amorphous materials in subsurface horizons
Ultisols	ult	Soils that are usually moist, with horizons of clay accumulation and a low supply of bases
Vertisols	ert	Soils with high content of swelling clays and wide deep cracks during some seasons

have three or more syllables but can be distinguished from order names because they do not end in "sol." Among the Argids, great groups are Natrargids (Latin *natrium*, sodium) for soils that have high contents of sodium, and Durargids (Latin *durus*, hard) for Argids with a hardpan cemented by silica and called a duripan.

Subgroup names are binomial. The great group name is preceded by an adjective such as "typic," which suggests the type or central concept of the great group, or the name of another great group, suborder, or order converted to an adjective to suggest that the soils are transitional between the two taxa.

Family names consist of several adjectives that describe the texture (sandy, silty, clayey, and so on), the mineralogy (siliceous, carbonatic, and so on), the temperature regime of the soil (thermic, mesic, frigid, and so on), and occasional other properties that are relevant to the use of the soil.

Series names are abstract names, taken from towns or places near where the soil was first identified. Cecil, Tama, and Walla Walla are names of soil series.

Order. In the highest category, 10 orders are recognized. These are distinguished chiefly by differences in kinds and amount of organic matter in the surface horizons, kinds of B horizons resulting from the dominance of various specific processes, evidences of churning through shrinking and swelling, base saturation, and lengths of periods during which the soil is without available moisture. The properties selected to distinguish the orders are reflections of the degree of horizon development and the kinds of horizons present.

Suborder. This category narrows the ranges in soil moisture and temperature regimes, kinds of horizons, and composition, according to which of these is most important. Moisture or temperature or soil properties associated with them are used to define suborders of Alfisols, Mollisols, Oxisols, Ultisols, and Vertisols. Kinds of horizons are used for Aridisols, compositions for Histosols and Spodosols, and combinations for Entisols and Inceptisols.

Great group. The taxa (classes) in this category group soils that have the same kinds of horizons in the same sequence and have similar moisture and temperature regimes. Exceptions to horizon sequences are made for horizons so near the surface that they are apt to be mixed by plowing or lost rapidly by erosion if plowed.

Subgroup. The great groups are subdivided into subgroups that show the central properties of the great group, intergrade subgroups that show properties of more than one great group, and other subgroups for soils with atypical properties that are not characteristic of any great group.

Family. The families are defined largely on the basis of physical and mineralogic properties of importance to plant growth.

Series. The soil series is a group of soils having horizons similar in differentiating characteristics and arrangement in the soil profile, except for texture of the surface portion, and developed in a particular type of parent material.

Type. This category of earlier systems of classification has been dropped but is mentioned here because it was used for almost 70 years. The soil types within a series differed primarily in the texture of the plow layer or equivalent horizons in unplowed soils. Cecil clay and Cecil fine sandy loam were types within the Cecil series. The texture of the plow layer is still indicated in published soil surveys if it is relevant to the use of the soil, but it is now considered as one kind of soil phase.

Classifications of soils have been developed in several countries based on other differentia. The principal classifications have been those of Russia, Germany, France, Canada, Australia, New Zealand, and the United States. Other countries have modified one or the other of these to fit their own conditions. Soil classifications have usually been developed to fit the needs of a government that is concerned with the use of its soils. In this respect soil classification has differed from classifications of other natural objects, such as plants and animals, and there is no international agreement on the subject.

Many practical classifications have been developed on the basis of interpretations of the usefulness of soils for specific purposes. An example is the capability classification, which groups soils according to the number of safe alternative uses, risks of damage, and kinds of problems that are encountered under use.

Surveys. Soil surveys include those researches necessary (1) to determine the important characteristics of soils, (2) to classify them into defined series and other units, (3) to establish and map the boundaries between kinds of soil, and (4) to correlate and predict adaptability of soils to various crops, grasses, and trees; behavior and productivity of soils under different management systems; and yields of adapted crops on soils under defined sets of management practices. Although the primary purpose of soil surveys has been to aid in agricultural interpretations, many other purposes have become important, ranging from suburban planning, rural zoning, and highway location, to tax assessment and location of pipelines and radio transmitters. This has happened because the soil properties important to the growth of plants are also important to its engineering uses.

Soil surveys were first used in the United States in 1898. Over the years the scale of soil maps has been increased from $^1/_2$ or 1 in. to the mile (8 or 16 mm to the kilometer) to 3 or 4 in. to the mile (47 to 63 mm to the kilometer) for mapping humid farming regions, and up to 8 in. to the mile (126 mm to the kilometer) for maps in irrigated areas. After the advent of aerial photography, planimetric maps were largely discontinued in favor of aerial photographic mosaics. The United States system has been used, with modifications, in many other countries. *See* AERIAL PHOTOGRAPH.

Two kinds of soil maps are made. The common map is a detailed soil map, on which soil boundaries are plotted from direct observations throughout the surveyed area. Reconnaissance soil maps are made by plotting soil boundaries from observations which are made at intervals. The maps show soil and other differences that are of significance for present or foreseeable uses.

The units shown on soil maps usually are phases of soil series. The phase is not a category of the classification system. It may be a subdivision of any class of the system according to some feature that is of significance for use and management of the soil, but not in relation to the natural landscape. The presence of loose boulders on the surface of the soil makes little difference in the growth of a forest, but is highly

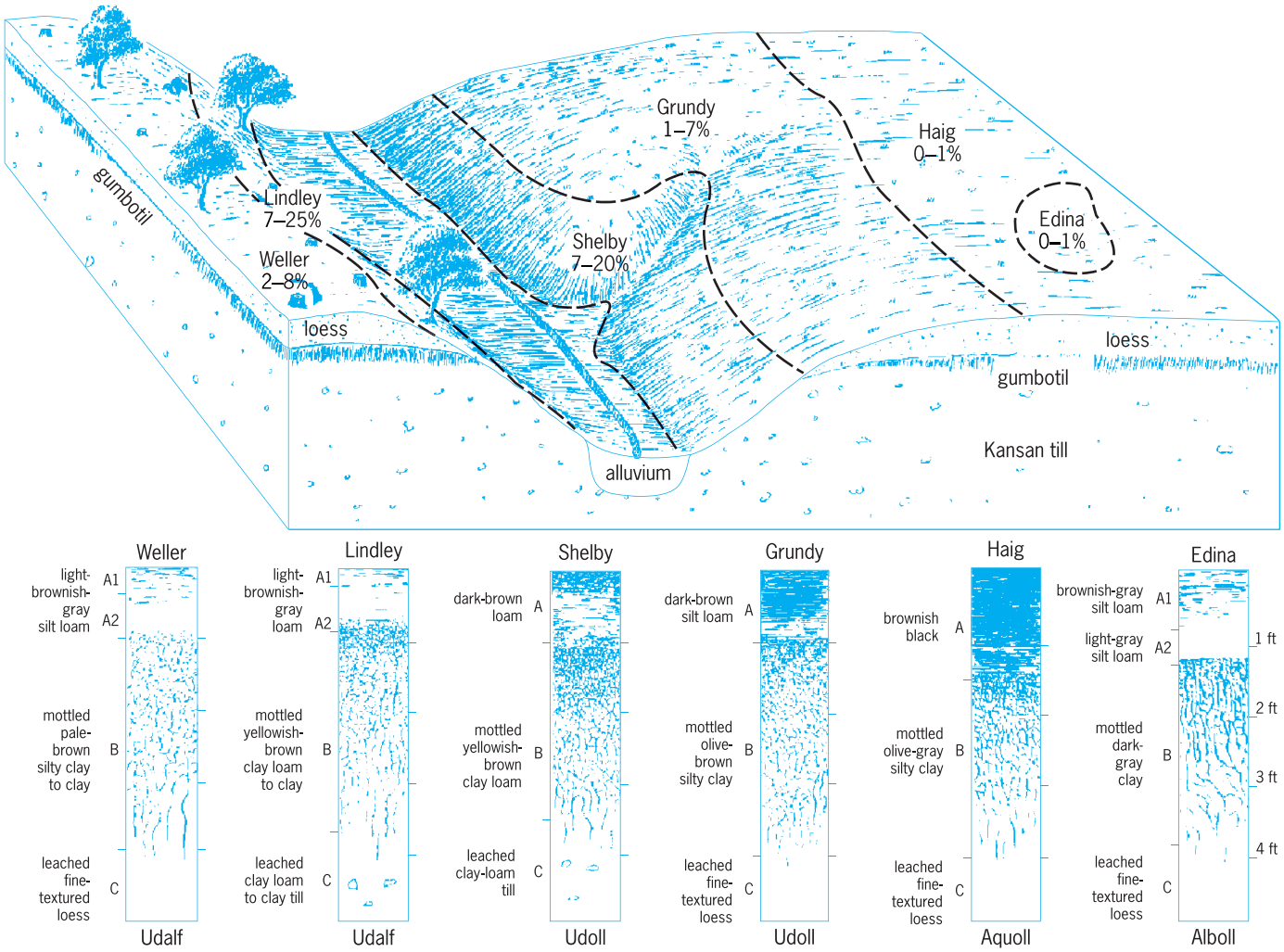

Fig. 3. Relation of the soil pattern to relief, parent material, and native vegetation on a farm in south-central Iowa. The soil slope gradient is expressed as a percentage. 1 ft = 30 cm. (*After R. W. Simonson et al., Understanding Iowa Soils, Brown, 1952*)

significant if the soil is to be plowed. Phases are most commonly based on slope, erosion, presence of stone or rock, or differences in the rock material below the soil itself. If a legend identifies a phase of a soil series, the soils so designated on a soil map are presumed to lie within the defined range of that phase in the major part of the area involved. Thus, the inclusion of lesser areas of soils having other characteristics is tolerated in the mapping if their presence does not appreciably affect the use of the soil. If there are other soils that do affect the use, inclusions up to 15% of the area are tolerated without being indicated in the name of the soil.

If the pattern of occurrence of two or more series is so intricate that it is impossible to show them separately, a soil complex is mapped, and the legend includes the word "complex," or the names of the series are connected by a hyphen and followed by a textural class name. Thus the phrase Fayette-Dubuque silt loam indicates that the two series occur in one area and that each represents more than 15% of the total area.

In places the significance of the difference between series is so slight that the expense of separating them is unwarranted. In such a case the names of the series are connected by a conjunction, for example, Fayette and Downs silt loam. In this kind of mapping unit, the soils may or may not be associated geographically.

It is possible to make accurate soil maps only because the nature of the soil changes with alterations in climatic and biotic factors, in relief, and in ground waters, all acting on parent materials over long periods of time. Boundaries between kinds of soil are made where such changes become apparent. On a given farm the kinds of soil usually form a repeating pattern related to the relief (**Fig. 3**).

Because concepts of soil have changed over the years, maps made 50 or more years ago may use the same soil type names as modern maps, but with different meanings.

Guy D. Smith

Soil Suborders

Soil suborders are broad classes at one level in the soil classification system adopted in the United States in 1965. A total of 47 suborders form the full set of classes in the second highest category (each category is a set of classes of parallel rank) in the system.

The number of local kinds of soils in a large country is also large. For example, 11,500 soil series have been recognized in soil surveys made in the United States through 1979. On the average, a series consists of six phases, which are the local kinds of soils. This means that approximately 66,000 local kinds have been defined up to the present time in a single large country, though all parts of that country have not been studied.

Despite the myriads of local kinds over the land surface of the Earth, all soils share some characteristics. They can all be related to one another in some way. The relationships are close for some pairs of local kinds and distant for others. The similarities and differences among the thousands of local kinds permit their grouping into sets of progressively broader classes in order to show degrees of kinship.

Though it is impossible for a single mind to retain concepts of 11,500 series or 66,000 phases, the salient features of a few dozen broad classes can be remembered. Consequently, the nature of the 47 suborders is described in this article. The purpose is to provide a general picture of the kinds of soils in the United States and the world. The 47 suborders grouped into the 10 orders are:

- Alfisols
  - Aqualfs
  - Boralfs
  - Udalfs
  - Ustalfs
  - Xeralfs
- Aridisols
  - Argids
  - Orthids
- Entisols
  - Aquents
  - Arents
  - Fluvents
  - Orthents
  - Psamments
- Histosols
  - Fibrists
  - Folists
  - Hemists
  - Saprists
- Inceptisols
  - Andepts
  - Aquepts
  - Ochrepts
  - Plaggepts
  - Tropepts
  - Umbrepts
- Mollisols
  - Albolls
  - Aquolls
  - Borolls
  - Rendolls
  - Udolls
  - Ustolls
  - Xerolls
- Oxisols
  - Aquox
  - Humox
  - Orthox
  - Torrox
  - Ustox
- Spodosols
  - Aquods
  - Ferrods
  - Humods
  - Orthods
- Ultisols
  - Aquults
  - Humults
  - Udults
  - Ustults
  - Xerults
- Vertisols
  - Torrerts
  - Uderts
  - Usterts
  - Xererts

The brief individual descriptions of the suborders are arranged in the same sequence as the list.

Broad regional distribution of soils for the world has been recognized (**Fig. 4**). Each region outlined on the map is identified by the name of the most extensive suborder among the component soils. In every region, suborders other than the most extensive one are important.

Alfisols. These soils have A and E horizons that are mostly pale in color and that have lost silicate clay, sesquioxides, and bases such as calcium and magnesium.

The soils have B horizons with accumulations of silicate clay and with moderate to high levels of exchangeable calcium and magnesium. The C horizons are usually lighter in color and lower in clay than the B horizons.

Alfisols are most extensive in humid, temperate regions but range from the edges of the tundra and the desert into the tropics. Mostly, the soils were under forest or savanna vegetation, though some were under prairie. All have been formed on land surfaces that are neither old nor yet among the youngest in the world.

Occurring as they do in many parts of the world, these soils are used for a wide variety of crops. Some remain in forest, and those under drier climates are used chiefly for grazing.

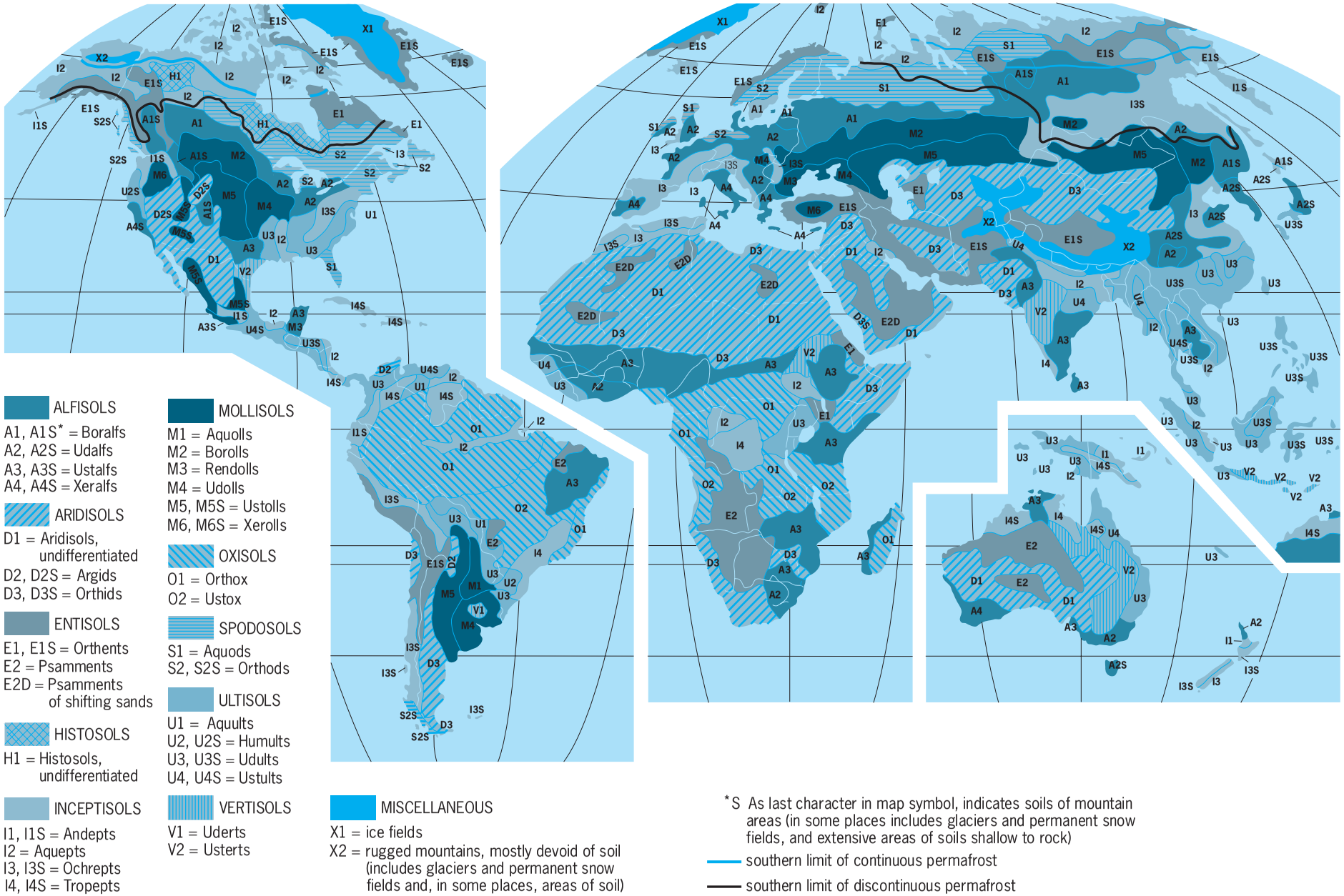

Fig. 4. General soil map of the world. Each region is identified by the name of the most extensive suborder. Other suborders are present in every region and are important in most of them. (*Soil Conservation Service, USDA*)

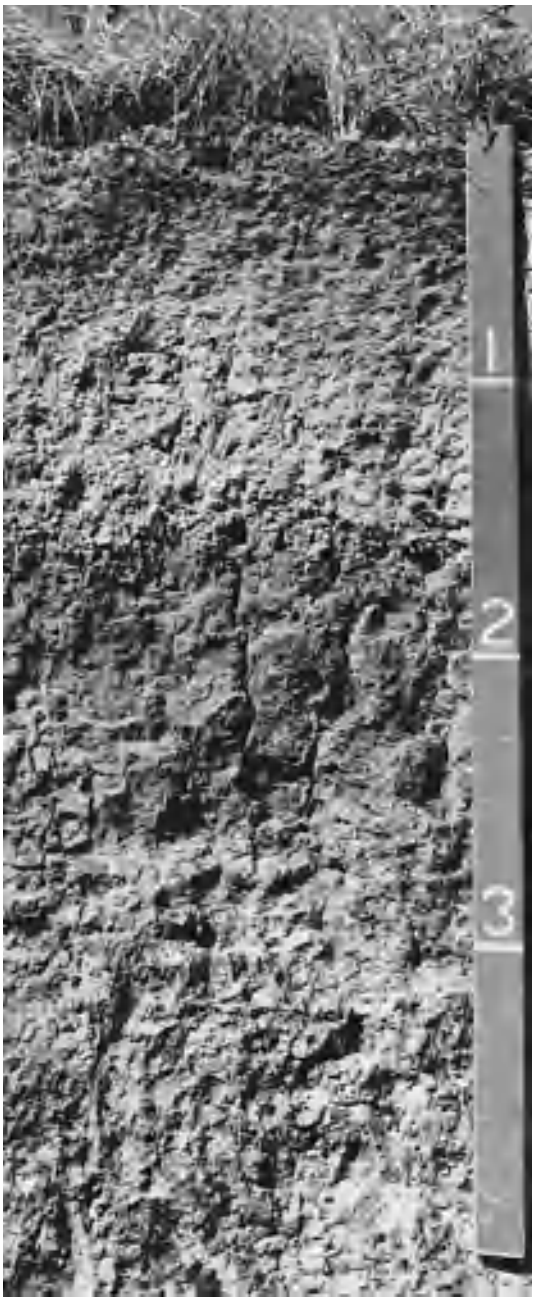

Fig. 5. Profile of an Aqualf with pale A and E horizons about 18 in. (45 cm) thick resting on B horizon high in clay which grades into C horizon at a depth of about 4 ft (1.2 m); numbers on scale indicate feet. 1 ft = 30 cm. (*Photograph by R. W. Simonson*)

Aqualfs. These are the seasonally wet Alfisols. They generally occur in depressions or on rather wide flats in local landscapes. In addition to the general morphology and composition shared with other soils of the order, Aqualfs are marked by gray or mottled colors reflecting their wetness (**Fig. 5**).

Boralfs. These are the well-drained Alfisols of cool or cold regions, such as west-central Canada and Russia. The soils occur either at high altitudes or in high latitudes, including some frigid zones.

In their morphology and composition, the soils are much like the Udalfs, though colors are more dull on the whole and the surplus of calcium and magnesium a little higher.

Udalfs. These are the well-drained Alfisols of humid, temperate climates. The soils are important in the north-central part of the United States, in western Europe, and in eastern Asia. Udalfs differ from Aqualfs in that B horizons are characteristically brown or yellowish brown and lack marks of wetness. These have higher mean annual temperatures than do the Boralfs and are moist for higher proportions of the year than the Ustalfs and Xeralfs (**Fig. 6**).

Ustalfs. These are well-drained Alfisols occurring in somewhat drier and mostly warmer regions than Udalfs. On the whole, the soils have more reddish B horizons and are a little higher in calcium and magnesium than Udalfs. These soils are intermittently dry during the growing season.

Xeralfs. These are well-drained Alfisols found in regions with rainy winters and dry summers, in what are called mediterranean climates. Like the Ustalfs in nature of B horizons, the soils have A horizons that tend to become massive and hard during the dry season. Some of the soils have duripans [cemented

Fig. 6. Profile of a Udalf with A and E horizons 12 in. (30 cm) thick over darker B horizon with blocky structure grading into C horizon at a depth of about 4 ft (1.2 m); larger numbers on scale indicate feet. 1 ft = 30 cm. (*Photograph by R. W. Simonson*)

layers at depth of 2 or 3 ft (0.6 or 0.9 m)] that interfere with root growth.

Aridisols. These are major soils of the world's deserts, which form about one-fourth of the land surface. Soils of other orders, especially the Entisols, are also present but less extensive in the deserts.

Formed under low rainfall, Aridisols have been leached little and are therefore high in calcium, magnesium, and other more soluble elements. The low rainfall has also limited growth of plants, mostly shrubs and similar species, so that the soils are low in organic matter and nitrogen. The combined A and B horizons are rarely more and usually less than 1 ft (30 cm) thick. The A horizons are light-colored and usually calcareous. All horizons are neutral or mildly alkaline in reaction.

Most Aridisols in use provide some grazing for nomadic herds. On the other hand, if water and other resources, including adequate skills, are available and climate is favorable, some Aridisols will support a large variety and produce high yields of crops.

Argids. These well-drained Aridisols have B horizons of silicate clay accumulation. The B horizons are characteristically brown or reddish in color and grade into lighter colored C horizons marked by carbonate accumulation. On the whole, these soils occupy the older land surfaces in desert regions (**Fig. 7**).

Orthids. These Aridisols lack B horizons of clay accumulation. Many are free of carbonates in the A and upper B horizons; most are well drained. Common colors are gray or brownish gray with little change from top to bottom of the profile. A few Orthids are fairly high in soluble salts such as sodium sulfate and sodium chloride, whereas others are high in calcium carbonate throughout. More extensive than the Argids, generally, Orthids occupy younger but not the youngest land surfaces in deserts.

Entisols. These soils have few and faint horizons, with reasons for the limited horizonation differing among suborders. Reasons for the practical absence of horizons are indicated for individual suborders.

Entisols occur in all parts of the world and may be found under a wide variety of vegetation. Most, though not all, are on young land surfaces, distributed from the tundra through the tropics and from the deserts to the rainiest climates. Entisols have a wide range in usefulness. Some are highly productive and others are not.

Aquents. These Entisols have been under water until very recent times at the margins of oceans, lakes, or seas. The wetness is reflected in the bluish-gray or greenish-gray colors. Examples are the soils in recently reclaimed polders of the Netherlands. The total extent of Aquents in the world is very small.

Arents. These are Entisols because of severe disturbance of soils formerly classifiable in other orders. The sequence of horizons has been disrupted completely, and remnants of those horizons can be found randomly distributed in the profiles of Arents.

Fluvents. These well-drained Entisols are in recently deposited alluvium. They occur along streams or in fans where the rate of sediment deposition is high. Marks of sedimentation are still evident, and identifiable horizons are lacking, except for slightly darkened surface layers or A horizons. Small bodies of these soils are scattered over all parts of the world.

Fig. 7. Profile of an Argid with pale silty A and E horizons, darker B horizon higher in clay, and calcareous C horizon; profile shown in 20 in. (50 cm) deep; numbers on scale indicate inches. 1 in. = 2.5 cm. (*Photograph by R. W. Simonson*)

Orthents. These well-drained Entisols are of medium or fine texture, mostly on strong slopes. The grass strips serve to remove sediment from overland flow. The soils may have A horizons or slightly darkened surface layers an inch or so thick but otherwise lack evidence of horizonation. Many of the soils are shallow to bedrock.

Psamments. These Entisols are of sandy texture. Like the Orthents, these may have thin A horizons, which grade into thick C horizons. The sandiness is the distinctive character of the suborder (**Fig. 8**).

Histosols. These are wet soils consisting mostly of organic matter, popularly known as peats and mucks. Most have restricted drainage and are saturated with

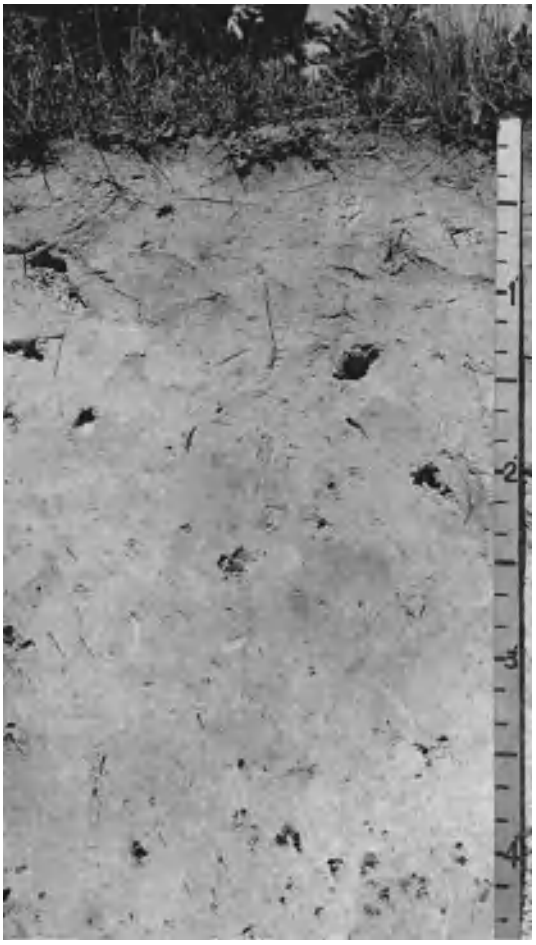

Fig. 8. Profile of a Psamment lacking evident horizons and consisting of sand throughout; numbers on scale indicate feet. 1 ft = 30 cm. (*Photograph by R. W. Simonson*)

water much of the time. A few are wet but not fully saturated. Widely distributed over the world, these soils may occur in small or large bodies, with the latter occurring chiefly at high latitudes. A large proportion of the total area is idle. Where the climate is favorable, some of the soils have been drained and are producing vegetables and other crops. *See* BOG; PEAT.

Fibrists. These Histosols consist mainly of recognizable plant residues or sphagnum moss. They are saturated with water most of the year unless drained (**Fig. 9**).

Folists. These Histosols consist of forest litter resting on rock or rubble. Drainage is not restricted, but a combination of rainfall, fog, and low temperatures keeps the litter wet.

Hemists. These Histosols consist of partially decayed plant residues. Plant structures have largely been destroyed but an appreciable share of the mass remains as fibers when rubbed vigorously. The soils are saturated with water much of the time unless drained.

Saprists. These Histosols consist of residues in which plant structures have been largely obliterated by decay. A very small part of the mass remains as fibers after vigorous rubbing. The soils are saturated with water much of the time unless drained. Most Saprists in the United States are known as muck.

Inceptisols. These soils have faint to moderate horizonation but lack horizons of accumulation of translocated substances other than carbonates and silica. Two of the suborders have distinct dark A horizons, and most have B horizons formed by losses and transformations without corresponding gains in substances. Thus, the Inceptisols are in some ways intermediate in horizonation between the Entisols and Vertisols on the one hand, and the Alfisols, Mollisols, Spodosols, and Ultisols on the other.

Inceptisols are widely distributed, ranging from the arctic through the tropics and from the margins of the desert into regions of heavy rainfall. They may consequently be found under a wide variety of vegetation. Usefulness of the soils has as wide a range as does their distribution. Some are highly productive and others are of little or no value.

Andepts. These Inceptisols are formed chiefly in volcanic ash or in regoliths with high components of ash. Mostly, the soils tend to be fluffy. They have thick dark A horizons, rather high levels of acidity, and poorly crystalline clay minerals. The soils are

Fig. 9. Profile of a Fibrist with little or no change from the surface to a depth of 5 ft (1.5 m); soil consists of partly decayed plant residues; numbers on scale indicate feet. 1 ft = 30 cm. (*Photograph by R. W. Simonson*)

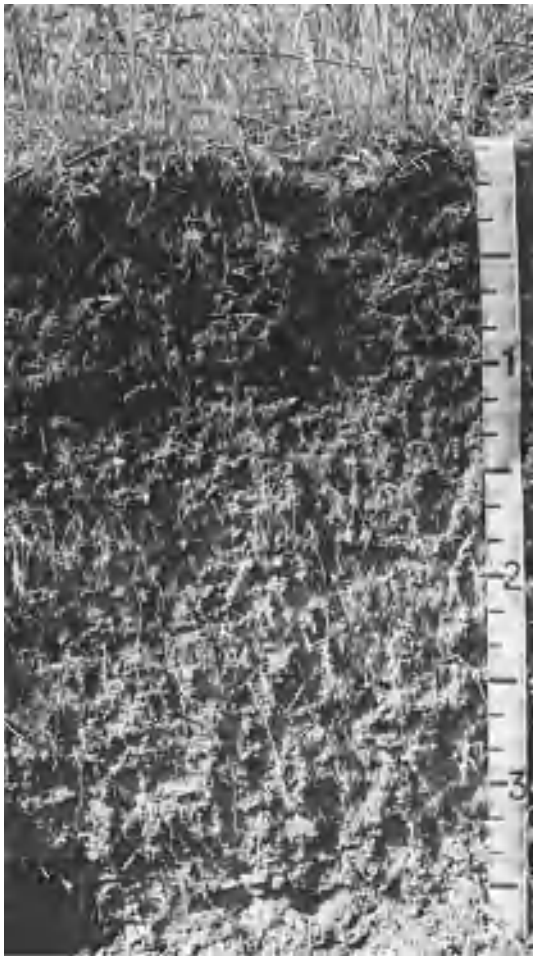

Fig. 10. Profile of an Andept with thick, dark A horizon, faint B horizon, and lighter C horizon; fine plant roots are numerous; numbers on scale indicate feet. 1 ft = 30 cm. (*Photograph by R. W. Simonson*)

widely distributed but seem to be restricted to regions of fairly recent volcanic activity (**Fig. 10**).

Aquepts. These Inceptisols are wet or have been drained. Like the Aqualfs, the soils have gray or mottled B and C horizons, but they lack silicate clay accumulation in their profiles. The A horizons may be dark and fairly thick or they may be thin, as they are in many of the soils (**Fig. 11**).

Ochrepts. These Inceptisols have pale A horizons, darker B horizons, and lighter colored C horizons. The B horizons lack accumulations of translocated clay, sesquioxides, or humus. The soils are widely distributed, occurring from the margins of the tundra region through the temperate zone but not in the tropics. Ochrepts also occur in the fairly dry regions though not in deserts (**Fig. 12**).

Plaggepts. These Inceptisols have very thick surface horizons of mixed mineral and organic materials added as manure or as human wastes over long periods of time. For the world as a whole, such soils are of negligible extent, but they are conspicuous where found.

Tropepts. These Inceptisols have moderately dark A horizons with modest additions of organic matter, B horizons with brown or reddish colors, and slightly paler C horizons. The soils are less strongly weathered than the geographically associated Ultisols and Oxisols. In general appearance, the profiles are much like that of the Orthox (**Fig. 13**). Tropepts are restricted to tropical regions, largely to those of moderate and high rainfall.

Umbrepts. These Inceptisols have dark A horizons more than 10 in. (25 cm) thick, brown B horizons, and slightly paler C horizons. The soils are strongly acid, and the silicate clay minerals are crystalline rather than amorphous as in the Andepts. The Umbrepts occur under cool or temperate climates, are widely distributed, and are of modest extent.

Mollisols. These soils have dark or very dark, friable, thick A horizons high in humus and bases such as calcium and magnesium. Most have lighter colored or browner B horizons that are less friable and about as thick as the A horizons. All but a few have paler C horizons, many of which are calcareous.

Major areas of Mollisols occur in subhumid or semiarid cool and temperate regions. They meet the desert along their drier margins and meet soils such

Fig. 11. Profile of an Aquept with thin, dark A horizons, fairly thick, light-gray B horizon, and stone in C horizon beside tape; numbers on scale indicate feet. 1 ft = 30 cm. (*Photograph by R. W. Simonson*)

Fig. 12. Profile of an Ochrept with litter on the surface, dark A horizon 4 in. (10 cm) thick, thin B horizon, and pale C horizon; deeper profile is marked by plant roots and traces of former roots; numbers on scale indicate inches. 1 in. = 2.5 cm. (*Photograph by R. W. Simonson*)

as the Alfisols at their more humid margins. Mollisols were formed under vegetation consisting chiefly of grasses and are thus the major ones of former prairies and steppes. The soils occupy rather young land surfaces.

Though they produce a variety of crops, Mollisols are largely used for cereals. These soils now produce a major share of the world's output of corn and wheat. Topography is generally favorable for the operation of large machinery, and many Mollisols are therefore in large farms. Yields have a wide range, depending on climatic conditions. Wide fluctuations in yield with wet and dry years are normal for the Mollisols marginal to arid regions. On the other hand, yields are consistently high for those under more humid climates.

Albolls. These are Mollisols with dark A horizons, pale E horizons, distinct B horizons marked by clay accumulation, and paler C horizons. The soils are set, especially in the upper part, for some part of the year. Mostly, the soils occur on upland flats and in shallow depressions.

Aquolls. These are wet Mollisols unless they have been drained. Because they were formed under wet conditions, the soils have thick or very thick, nearly black A horizons over gray or mottled B and C horizons. If they have not been drained, the soils may be under water for part of the year, but they are seasonally rather than continually wet (**Fig. 14**).

Borolls. These are Mollisols of cool and cold regions. Most areas are in moderately high latitudes or at high altitudes. The soils have fairly thick, nearly black A horizons, dark grayish-brown B horizons, and paler C horizons that are commonly calcareous. The B horizons of some soils have accumulations of clay. These soils are extensive in western Canada and Russia.

Rendolls. These are the Mollisols formed in highly calcareous parent materials, regoliths with more than 40% calcium carbonate. The soils may be calcareous to the surface and must have high levels of carbonates within a depth of 20 in. (50 cm). Rendolls do not have horizons of carbonate accumulation. The profiles consist of dark or very dark A horizons grading into pale C horizons. For the most part, Rendolls are restricted to humid, temperate regions.

Udolls. These are Mollisols of humid, temperate and warm regions where maximum rainfall comes during the growing season. The soils have thick, very dark A horizons, brown B horizons, and paler C horizons.

Fig. 13. Profile of an Orthox with slightly darkened A horizon about 1 ft (30 cm) thick, little further change with depth, and deep penetration by fine plant roots; scale in feet. 1 ft = 30 cm. (*Photograph by R. W. Simonson*)

Fig. 14. Profile of an Aquoll with very thick, dark A horizon, signs of mixing and burrowing by animals at depths between 3 and 3.5 ft (0.9 and 1.1 m), and lighter C horizon at the bottom; the numbers on the scale to the left indicate feet. 1 ft = 30 cm. (*Photograph by R. W. Simonson*)

Throughout the profile these soils are browner than the Borolls and are not as cold. Udolls lack horizons of accumulation of powdery carbonates. Some of the soils have B horizons of clay accumulation and others do not. These soils are major ones of the Corn Belt of the United States (**Fig. 15**).

Ustolls. These are the Mollisols of temperate and warm climates with lower rainfall than the Udolls. The soils are therefore dry for an appreciable part of each year, usually more than 90 cumulative days. Horizons and their sequence are much the same as for Udolls except that many Ustolls have accumulations of powdery carbonates at depths of 40 in. (100 cm) or less.

Xerolls. These are Mollisols of regions with rainy winters and dry summers. The nature and sequence of horizons are much like those of the Ustolls. The soils are completely dry for a long period during the summer of each year.

Oxisols. These soils have faint horizonation, though formed in strongly weathered regoliths. The surface layers or A horizons are usually darkened and moderately thick, but there is little evidence of change in the remainder of the profile. Because of the intense or long weathering, the soils consist of resistant minerals such as kaolinite, forms of sesquioxides, and quartz. Weatherable minerals such as feldspars have largely disappeared. Moreover, the clay fraction has limited capacity to retain bases such as calcium and magnesium.

The soils are porous and readily penetrated by water and plant roots. A distinctive feature of Oxisols is the common occurrence of tubular pores about the diameter of ordinary pins extending to depths of 6 ft (1.8 m) or more.

Oxisols are largely restricted to low altitudes in humid portions of the tropics. Any occurring elsewhere seem to be relicts of earlier geologic ages. All occupy old land surfaces. Most were formed under forest, with some under savanna vegetation. Regions with Oxisols as major soils are extensive, ranking second in total area only to the Aridisols of deserts.

Most Oxisols remain in forest or savanna and produce little food and fiber. Many regions are sparsely inhabited, with natives depending on shifting cultivation for much of their food. A small proportion of the total area is cultivated with modern technology and is highly productive. Even so, management to

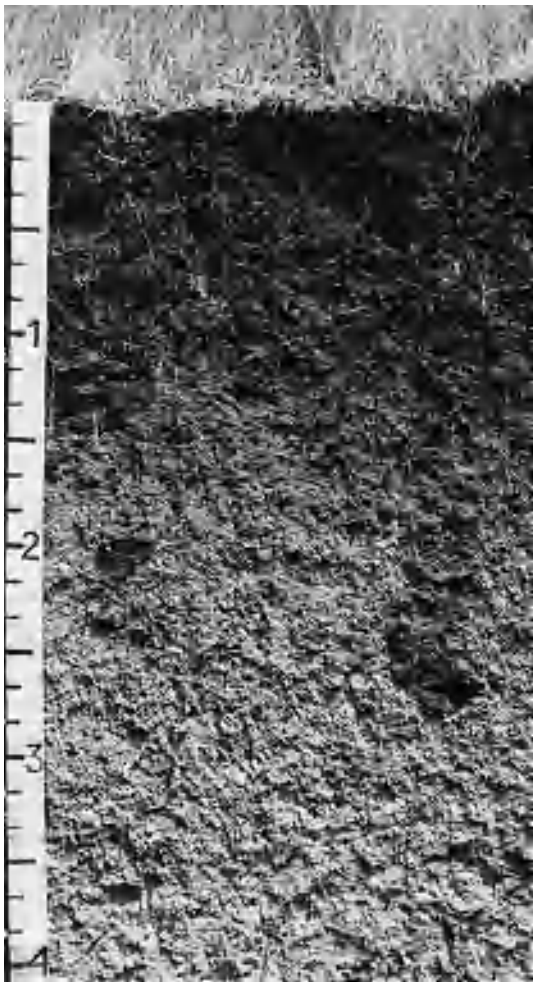

Fig. 15. Profile of a Udoll with thick, dark A horizon, B horizon gradational in color, and rather pale C horizon; filled former animal burrows in B horizon; numbers on scale indicate feet. 1 ft = 30 cm. (*Photograph by R. W. Simonson*)

ensure sustained high yields is still to be developed for the more strongly weathered Oxisols.

Aquox. These are seasonally wet Oxisols found chiefly in shallow depressions. Because of their wetness, deeper profiles are dominantly gray, with or without mottles and nodules or sheets of iron and aluminum oxides. Total extent is extremely small.

Humox. These are well-drained Oxisols high in organic matter and moist all or nearly all year. Profiles have dark A horizons 1 ft (0.3 m) or so thick over generally reddish B and C horizons. The high amounts of organic matter distinguish Humox from other suborders of Oxisols, and the soils are also moist much more of the year than are Torrox and Ustox. Humox are believed to be of limited extent, restricted to relatively cool climates and high altitudes for Oxisols.

Orthox. These are well-drained Oxisols moderate to low in organic matter and moist all or nearly all year. Orthox are much like Humox in general appearance, but their profiles are lower in organic matter. They are moist more of each year than are Torrox and Ustox. Although good data on extent are lacking, Orthox are believed to be extensive at low altitudes in the heart of the humid tropics (Fig. 13).

Torrox. These are well-drained Oxisols low in organic matter and dry most of the year. Profiles resemble those of Orthox except that A horizons are more poorly expressed. The soils are believed to have been formed under more rainy climates of past eras. Total extent of the Torrox seems to be extremely small.

Ustox. These are well-drained Oxisols low to moderate in organic matter and dry for periods of at least 90 cumulative days each year. Profiles resemble those of Orthox, on the whole. Ustox are lower in organic matter than Humox, dry for longer periods each year than Humox and Orthox, and moist for longer periods than Torrox. Good data on extent are lacking, but Ustox are believed to be extensive.

Spodosols. These soils have B horizons with accumulations of one or both of organic matter and compounds of aluminum and iron. The accumulated substances are amorphous in nature. They impart red, brown, or black colors to the B horizons, which may have irregular lower boundaries with tongues extending downward a foot or more. If the soil has not been disturbed, the surface layer consists of both fresh and partly decayed litter. This rests on a very pale, leached E horizon overlying a highly contrasting B horizon. The Spodosols formed from sands under boreal coniferous forests have some of the most striking profiles in the world. Mostly, the soils are strongly acid because of the small supplies of bases.

Spodosols are most extensive in humid, cool climates, but some occur at low elevations under tropical and subtropical climates. The soils were largely under forest. The bulk of the Spodosols have been formed in sandy regoliths, with others in loamy regoliths. Land surfaces are fairly young.

Most Spodosols remain in forest, but some are cultivated in both cool and tropical regions. The variety of crops produced is large because of the wide climatic range under which the soils occur. The range in yields is also wide, being dependent on the combination of climatic conditions and prevailing level of technology. Production is modest for most Spodosols under cool climates and simple management. Production is high from some soils cultivated with complex management in tropical climates.

Aquods. These are seasonally wet Spodosols. The soils may be wet most of each year but not all of the time. The B horizons are black or dark brown in color and some are cemented. Aquods occupy depressional areas or wide flats from which water cannot escape easily (**Fig. 16**).

Ferrods. These are well-drained Spodosols having B horizons of iron accumulation with little organic matter. Appearance of the profile is much like that of Orthods.

Humods. These are well-drained Spodosols having B horizons of humus accumulation, usually black or dark brown in color. Aluminum usually accumulates with the humus but iron is lacking, especially from the upper part of B horizons. Where formed in white or nearly white sands, the soils have striking profiles, as in parts of western Europe.

Orthods. These are well-drained Spodosols having B horizons of humus, aluminum, and iron accumulation. The B horizons are mostly red or reddish in color and are friable. They grade downward into

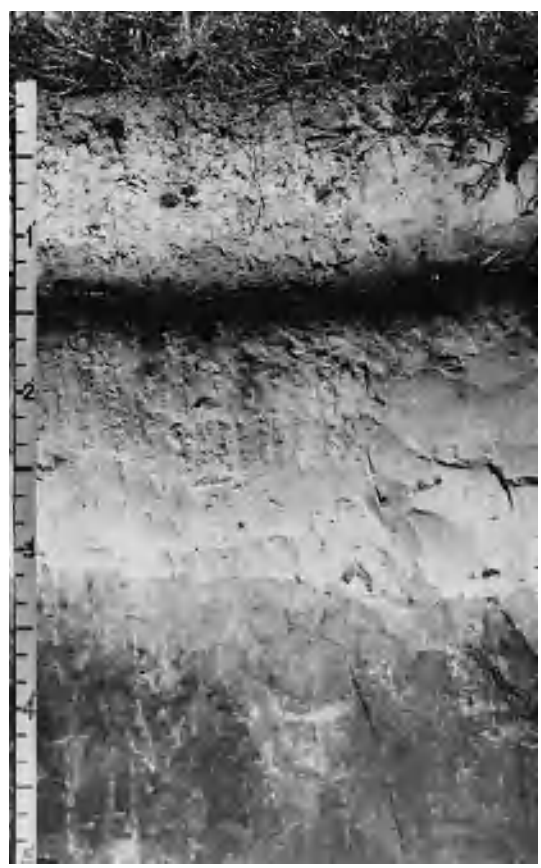

Fig. 16. Profile of an Aquod with distinct E horizon, dark B horizon at depth of 18 in. (45 cm), pale C horizon below, and part of buried profile below 4 ft (1.2 m); soil consists of sand; numbers on scale indicate feet. 1 ft = 30 cm. (*Photograph by R. W. Simonson*)

lighter-colored C horizons which are commonly less friable and may be very firm. Orthods form the most extensive suborder among the Spodosols, being widespread in Canada and Russia.

Ultisols. Like Alfisols, these soils have A and E horizons that have lost silicate clays, sesquioxides, and bases. Most A and E horizons are pale, though not all are. The B horizons have accumulations of silicate clays and low levels of exchangeable calcium and magnesium. The C horizons are usually lighter in color and lower in clay than the B horizons. Combined thickness of the A, E, and B horizons is greater, on the average, for Ultisols than for Alfisols. Ultisols are strongly acid throughout their profiles, reflecting the low levels of exchangeable bases.

Ultisols are most extensive under humid, warm-temperate climates but extend through the tropics. They are not found in cold regions. The largest bodies of the soils are in southeastern Asia, nearby islands, and the southeastern United States. The soils were usually under forest but some were covered by savanna vegetation. All were formed in strongly weathered regoliths on old land surfaces.

Many Ultisols remain in forest. Among those producing crops, a majority are used under some method of shifting cultivation. Production is limited in such circumstances. On the other hand, the variety of crops and yields obtained can be large if cultivators are in a position to apply complex technology to Ultisols.

Aquults. These are seasonally wet Ultisols, saturated with water for an important part of the year unless drained. Usually the soils have thin, dark A horizons, but they may be as thick as 20 in. (50 cm). Deeper profiles are gray, with or without red mottles. Aquults occur in depressions or on wide upland flats from which water moves very slowly.

Humults. These are well-drained Ultisols formed under rather high rainfall distributed evenly over the year. The soils are high in organic matter throughout their profiles, and most have darkened A horizons of moderate thickness. Deeper profiles tend to be brown, reddish-brown, or yellowish-brown in color. Humults are common in southeastern Brazil.

Udults. These are well-drained Ultisols of humid, warm-temperate and tropical regions. The soils are low or relatively low in humus and typically have thin, darkened A horizons. The B horizons are yellowish-red, red, brown, or yellowish-brown in color and are fairly thick. Rainfall is high enough and distributed evenly enough over the year so that soils are dry for only short periods. Udults are major soils in the southeastern parts of the United States and Asia (**Fig. 17**), and their total extent is large.

Ustults. These are well-drained Ultisols of warm-temperate and tropical climates with moderate or low rainfall. The soils are like the Udults in general appearance but are dry for appreciable periods each year. Examples of Ustults may be found in northeastern Australia. Total extent of the suborder is appreciably less than that of Udults.

Xerults. These are well-drained Ultisols of regions with warm, dry summers and cool, rainy winters. The soils are like Ustults in appearance but become and remain dry for longer periods in summers. Total extent is small.

Fig. 17. Profile of a Udult with pale A horizon 16 in. (40 cm) thick, darker B horizon higher in clay and iron oxides, and C horizon near bottom; numbers on scale indicate feet. 1 ft = 30 cm. (*Photograph by R. W. Simonson*)

Vertisols. These soils have faint horizonation for two main reasons. In the first place, Vertisols are formed in regoliths that are high in clay and therefore resistant to change. In the second place, the clay fraction in the soils has high levels of activity. The soils are therefore subject to marked swelling and shrinking as they wet and dry. Cracks formed as the soils become dry may extend to depths of several feet. Because of the shrinking and swelling of the soils, materials from deeper profiles are forced upward in places so that entire soils are slowly but continually overturned and mixed. Vertisols have therefore been called "self-swallowing" and "soils that plow themselves."

Some Vertisols have darkened A horizons, whereas others do not. All are low in organic matter and high in bases. Many are calcareous in deeper profiles. Most are neutral or mildly alkaline in reaction because of goodly supplies of bases.

Vertisols occur in warm-temperate and tropical climates with one or more dry seasons. The soils were under savanna vegetation for the most part with a

few in forest. Land surfaces are old or fairly old. Large bodies of Vertisols are found on the Deccan Plateau of India, in the Gezira of Sudan, and in Australia.

Because they are high in active clays, Vertisols are hard to cultivate. The soils therefore remain in savanna in many places, and the savannas are used for grazing or left alone. Large areas of the soils are cultivated, some with simple, bullock-drawn implements and others with large machinery. A wide variety of crops are produced, but yields are generally modest, especially for cultivators dependent on simple technology. The soils will, however, produce crops indefinitely under simple management.

Torrerts. These are Vertisols of arid regions, the driest soils of the order. Because the soils are dry most of the time, cracks that form tend to remain open. The soils do become wet enough at rare intervals to permit cracks to close. Torrerts are of limited extent.

Uderts. These are Vertisols of humid regions; each profile is moist in some part most of each year. The soils do dry out enough to permit formation of cracks once every year, as a rule. Uderts are moist more of each year than other Vertisols. The soils are of moderate extent.

Usterts. These are Vertisols of subhumid and semiarid regions, chiefly under climates with two rainy and two dry seasons each year. Cracks formed during dry seasons are open for at least 90 cumulative days per year. The Usterts are thus intermediate in moisture regimes between Uderts on the wet side and Torrets on the dry side. Usterts are dry for shorter periods during summer than Xererts. Usterts are extensive, represented by large bodies in Australia and India (**Fig. 18**).

Fig. 18. Profile of a Ustert with thick, dark A horizon, which is due in part to mixing and churning of soil mass and which grades into the lighter-colored C horizon; scale in feet. 1 ft = 30 cm. (*Photograph by E. H. Templin*)

Xererts. These are Vertisols of regions with warm, dry summers and cool, rainy winters—the mediterranean climates. Cracks formed as the soils dry out each summer remain open for at least 60 consecutive days. Xererts are higher in moisture than Torrerts and lower than Uderts. They have longer dry periods during warm seasons than Usterts. The soils are of limited extent.

Roy W. Simonson

Zonality

Many soils that are geographically associated on plains have common properties that are the result of formation in similar climates with similar vegetation. Because climate determines the natural vegetation to a large extent and because climate changes gradually with distance on plains, there are vast zones of uplands on which most soils have many common properties. This was first observed in Russia toward the end of the nineteenth century by V. V. Dokuchaev, the father of modern soil science. He also observed that on floodplains and steep slopes and in wet places the soils commonly lacked some or most of the properties of the upland soils. In mountainous areas, climate and vegetation tend to vary with altitude, and here the Russian students observed that many soils at the same altitude had many common properties. This they called vertical zonality in contrast with the lateral zonality of the soils of plains.

Zonal classification. These observations led N. M. Sivirtsev to propose in about 1900 that major kinds of soil could be classified as Zonal if their properties reflected the influence of climate and vegetation, as Azonal if they lacked well-defined horizons, and as Intrazonal if their properties resulted from some local factor such as a shallow ground water or unusual parent material.

This concept was not accepted for long in Russia. It was adopted in the United States in 1938 as a basis for classifying soil but was dropped in 1965. This was because the Zonal soils as a class could not be defined in terms of their properties and because they had no common properties that were not shared by some Intrazonal and Azonal soils. It was also learned that many of the properties that had been thought to reflect climate were actually the result of differences in age of the soils and of past climates that differed greatly from those of the present.

Zonality of soil distribution is important to students of geography in understanding differences in farming, grazing, and forestry practices in different parts of the world. To a very large extent, zonality is reflected but is not used directly in the soil classification used in the United States. The Entisols include most soils formerly called Azonal. Most of the soils formerly called Intrazonal are included in the orders of Vertisols, Inceptisols, and Histosols and in the aquic suborders such as Aquolls and Aqualfs. Zonal soils are mainly included in the other suborders in this classification.

The soil orders and suborders have been defined largely in terms of the common properties that result from soil formation in similar climates with

similar vegetation. Because these properties are important to the native vegetation, they have continuing importance to farming, ranching, and forestry. Also, because the properties are common to most of the soils of a given area, it is possible to make small-scale maps that show the distribution of soil orders and suborders with high accuracy.

Zonal properties. A few examples of zonal properties of soils and their relation to soil use follow. The Mollisols, formerly called Chernozemic soils, are rich in plant nutrients. Their natural ability to supply plant nutrients is the highest of any group of soils, but lack of moisture often limits plant growth. Among the Mollisols, the Udolls are associated with a humid climate and are used largely for corn (maize) and soybean production. Borolls have a cool climate and are used for spring wheat, flax, and other early maturing crops. Ustolls have a dry, warm climate and are used largely for winter wheat and sorghum without irrigation. Yields are erratic on these soils. They are moderately high in moist years, but crop failures are common in dry years. The drier Ustolls are used largely for grazing. Xerolls have a rainless summer, and crops must mature on moisture stored in the cool seasons. Xerolls are used largely for wheat and produce consistent yields.

The Alfisols, formerly a part of the Podzolic soil group, are lower in plant nutrients than Mollisols, particularly nitrogen and calcium, but supported a permanent agriculture before the development of fertilizers. With the use of modern fertilizers, yields of crops are comparable to those obtained on Mollisols. The Udalfs are largely in intensive cultivation and produce high yields of a wide variety of crops. Boralfs, like Borolls, have short growing seasons but have humid climates. They are used largely for small grains or forestry. Ustalfs are warm and dry for long periods. In the United States they are used for grazing, small grains, and irrigated crops. On other continents they are mostly intensively cultivated during the rainy season. Population density on Ustalfs in Africa is very high except in the areas of the tsetse fly. Xeralfs are used largely for wheat production or grazing because of their dry summers.

Ultisols, formerly called Latosolic soils, are warm, intensely leached, and very low in supplies of plant nutrients. Before the use of fertilizers, Ultisols could be farmed for only a few years after clearing and then had to revert to forest for a much longer period to permit the trees to concentrate plant nutrients at the surface in the leaf litter. With the use of fertilizers, Udults produce high yields of cotton, tobacco, maize, and forage. Ustults are dry for long periods but have good moisture supplies during a rainy season, typically during monsoon rains. Forests are deciduous, and cultivation is mostly shifting unless fertilizers are available.

Aridisols, formerly called Desertic soils, are high in some plant nutrients, particularly calcium and potassium, but are too dry to cultivate without irrigation. They are used for grazing to some extent, but large areas are idle. Under irrigation some Aridisols are highly productive, but large areas are unsuited to irrigation or lack sources of water. Guy D. Smith

Physical Properties

Physical properties of soil have critical importance to growth of plants and to the stability of cultural structures such as roads and buildings. Such properties commonly are considered to be (1) size and size distribution of primary particles and of secondary particles, or aggregates; (2) the consequent size, distribution, quantity, and continuity of pores; (3) the relative stability of the soil matrix against disruptive forces, both natural and cultural; (4) color and textural properties, which affect absorption and radiation of energy; and (5) the conductivity of the soil for water, gases, and heat. These usually would be considered as fixed properties of the soil matrix, but actually some are not fixed because of influence of water content. The additional property, water content—and its inverse, gas content—ordinarily is transient and is not thought of as a property in the same way as the others. However, water is an important constituent, despite its transient nature, and the degree to which it occupies the pore space generally dominates the dynamic properties of soil. Additionally, the properties listed above suggest a macroscopic homogeneity for soil which it may not necessarily have. In a broad sense, a soil may consist of layers or horizons, each consisting of roughly uniform soil materials of various types, which considered together may affect both mechanical properties and movement of water in a soil profile.

From a physical point of view it is primarily the dynamic properties of soil which affect plant growth and the strength of soil beneath roads and buildings. While these depend upon the chemical and mineralogical properties of particles, particle coatings, and other factors discussed above, water content usually is the dominant factor. Water content depends upon flow and retention properties, so that the relationship between water content and retentive forces associated with the matrix becomes a key physical property of a soil.

Structure and pore space. Although the mineralogical nature of the clay fraction and the textural classification of the soil are of basic importance to soil physical properties, other factors often have even greater importance. Decomposed organic matter, or humus, and certain inorganic compounds, such as iron and aluminum oxides, form coatings on soil particles, and these, along with some clays, bind soil particles together into granules or aggregates. The size and distribution of such aggregates are major factors in determining pore size and distribution in soil and are important in the physical behavior of soils. Additionally, fragmented organic materials, undecomposed or only partially so, if present in soil in significant quantity, also can play a major role in determining physical properties. The manner in which mineral soil particles are assembled and maintained in aggregate form, together with quantity, size, and distribution of fragments of partially decomposed

organic materials, is referred to as soil structure. Soil structure may develop as a consequence of natural processes such as wetting and drying, freezing and thawing, transport of minerals in moving water and their deposition, or mechanical forces exerted by plant roots or by other biological agents. Natural structure also may be modified by plowing, cultivation, and mechanical forces associated with cultural activities.

There are two aspects to the evaluation of soil structure which have major practical importance: measurements of pore size and configuration, and assessments of the stability of these characteristics against various natural and cultural forces which act to change the physical arrangement of primary particles and secondary aggregates. Pore size and configuration determine water retention and water and air transport properties of a soil and, in some cases, the ease with which plant roots penetrate and living organisms can move. The equivalent pore size of the largest water-filled pore in a soil sample (that is, the radius of a cylindrical pore which would behave with respect to water retention similarly to the largest water-filled soil pore) may be measured by using the hanging water column apparatus (**Fig. 19**), and the equation for capillary rise, $r = 2s/Dgh$, where s is the surface tension of water, D its density, g the acceleration of gravity, and h the height of rise of water in a capillary tube of radius r.

The porous plate apparatus for relating water-filled pore size to the force of removal exerted by the hanging water column (Fig. 19*a*) is based upon a capillary tube model where soil pores are presumed to be equivalent to small sections of capillary tubes (Fig. 19*b*). Pores in the porous plate are all considerably smaller than the size of the capillary tube (Fig. 19*a*), so that they remain water-filled as h is varied. More typically, the shape of air-water interfaces (Fig. 19*c*) has a radius of curvature, R, given as a combination of a positive radius r_1 measured axially and a negative radius r_2 at right angles, or $1/R = (1/r_1) + (1/r_2)$. A plot of water content of a soil sample in the hanging column apparatus against the height of the hanging water column characterizes the pore size distribution of the sample. A large change in water content with a small change in hanging water column length indicates a large volume of pores of similar size, whereas a small change in water content with a large change in hanging water column length indicates a relatively small volume associated with pores over a wide size range. More direct evidence of pore size and arrangement may be obtained by microscopic examination.

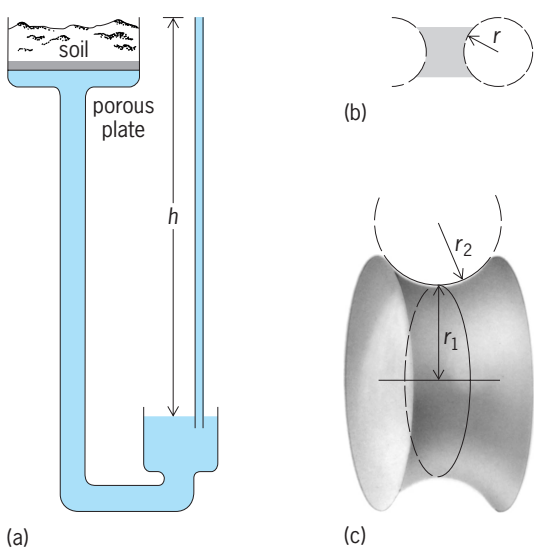

Fig. 19. Measurement of equivalent pore size. (*a*) Porous plate apparatus. (*b*) Schematic representations of the air-water interfaces and (*c*) the particle-water interfaces.

The second factor which is basic to determining the structural state of soil has to do with soil stability against disruptive forces. The existence of binding forces between clay particles themselves and between clay and other mineral surfaces, together with colloidal organic decomposition products, polysaccharides and polyuronides, and inorganic cementing substances, determines how well a particular geometrical organization of soil particles resists change against disruptive forces. Hence, structural evaluation usually involves both geometrical properties and some indication of how stable a particular arrangement of particles will be against the disruptive forces of a plow, falling raindrops, the weight of an animal, or some other force. Although the nature of the pore space, by implication at least, is a part of any characterization of soil structure, often the measurement involves application of a particular disruptive force and noting the degree to which an existing arrangement is destroyed.

Bulk density. An important property which can be used to characterize the structural state of a soil is the bulk density. Where the particle density of soil materials is uniform (often taken as 2.65 g/cm^3), it is possible to determine the total pore space fraction of a soil from measurement of the bulk density (mass of dry soil per unit bulk volume) and use of the formula, pore space fraction $= 1 - (D_b/D_p)$, where D_b and D_p are the bulk and particle densities. The presence of organic materials having particle densities differing appreciably from those of the mineral particles introduces some error into such computations. Agricultural soils in the plow depth ordinarily have bulk densities ranging from 1.0 to 1.5 g/cm^3 and a pore-space fraction from 0.4 to 0.6, or roughly half particles and half pore space.

Soil horizonation. From some perspectives the physical properties of soil may be presented adequately by a description of the properties of a sample of surface soil. However, horizonation, whether produced by natural or cultural processes, may profoundly affect water flow and retention near the surface or within the rooting zone of plants. Such influence is particularly noticeable at boundaries between materials having different pore sizes.

Horizonation may be the result of the method of original deposition, for example, alluvial or loess derived profiles, or it may develop over time as a consequence of weathering processes which differ near

the soil surface from those at depths beneath. Root growth and development and organic matter decomposition near the surface may promote both the development of aggregates and their stabilization. Also, fine soil particles, silts and clays, may be displaced downward through a profile by moving water, accumulating in layers below. Solutes carried in water also may dissolve minerals, carrying them in the direction of water flow and depositing them as precipitates or leaving them behind as water evaporates.

Horizonation also may occur in soil as a consequence of plowing and cultivation operations. It is not unusual for a plow pan to develop in cultivated soils (**Fig. 20**) as a consequence of the smearing action of an implement surface as it slides through the soil. Decreased porosity at an interface between disturbed and undisturbed soil can profoundly affect water flow and bring about both filtering of fine particles, which may be carried in moving water, and deposition of substances by precipitation. Also, soil may become severely compacted in the surface by vehicle and animal traffic, with a consequent decrease in porosity and increase in soil hardness.

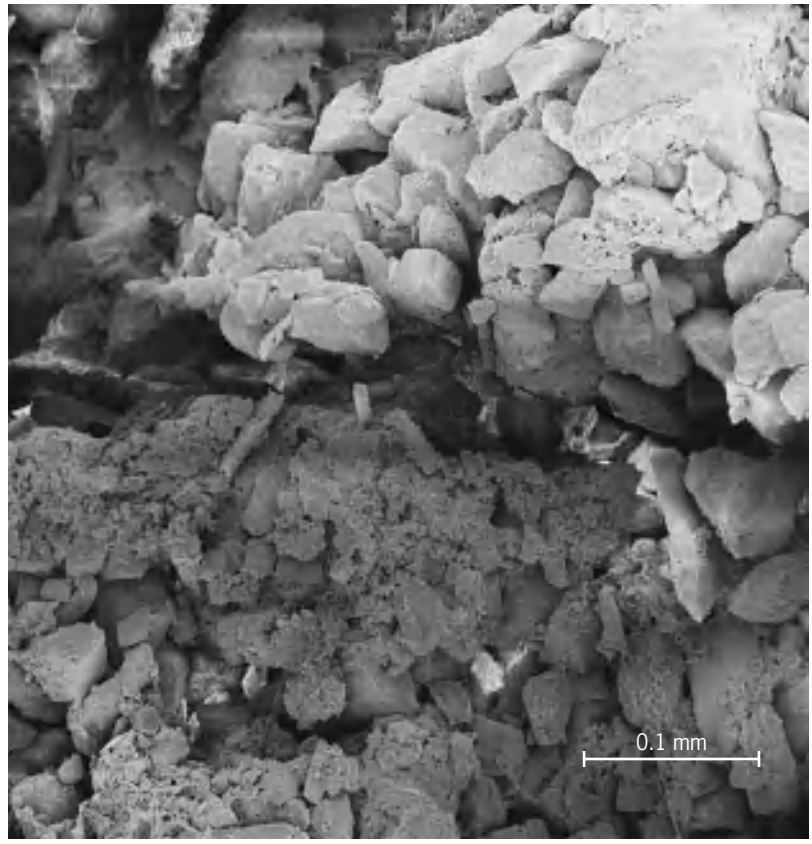

Fig. 20. Interface between disturbed soil and compact soil produced by the smearing action of a tillage implement as shown by a scanning electron micrograph. Water flow across such an interface is greatly reduced.

Forces acting on soil water. Flow of water in soil depends upon forces existing within the soil matrix and upon forces associated with gravity acting upon the water itself. In saturated soil the force of gravity acting upon water within the soil and by standing water upon the soil surface constitutes the moving force, and it can cause flow in any direction depending upon the geometry of the system. In unsaturated soil, gravity still acts upon the water but, except in very wet soil, this force may be small compared to forces due to the attraction of solid surfaces for water. These are the same forces that cause water to rise in a capillary tube, and they are sometimes referred to as capillary forces. In soil, such forces are called matric forces. In addition to matric forces, three other forces may be present in a soil system: osmotic forces, relating to the attraction between solutes and water; local gas pressures, which may exert a force upon air-water interfaces within soil pores; and overburden forces, which may arise where substances dissolved or suspended in water add to its density and increase its weight. Overburden forces may exist also in saturated soil, but usually such forces are neglected under both saturated and unsaturated conditions.

The nature of the various forces acting on water in soil differ appreciably, and the direction in which each acts may be highly complicated. Hence, it is difficult to apply them quantitatively to soil water. To circumvent this difficulty, the contribution of the various forces to the energy state of soil water is considered, thus permitting the addition of the contributions of each type of force. This involves measuring the work required to remove a unit quantity of water, and is known as the potential. Such work depends upon the degree of wetness, so that the energy state is not a linear function of water content. The units of measurement are energy content per unit of liquid volume of water. The commonly used units for potential are the bar, the millibar, and centimeter of water. Above the water table, work is required to remove water, and by convention, potential is negative. The potential at a flat or free air–water interface is taken to be zero.

Practical problems of water flow in unsaturated soil involve mostly the matric and gravity potentials. For flow into roots, because they are semipermeable membranes, additional work must be done by a plant to remove water from those ions which do not readily pass through root membranes. Thus, in dealing with flow of water in soil-plant systems, osmotic potentials also must be considered. However, gravity potential often is small in unsaturated soil compared to matric and osmotic potentials, and sometimes may be neglected. Matric potential is a function of water content, pore size, and pore size distribution in the soil. As with capillary tubes, small pores hold water tightly, and considerable work must be done to remove it. The capillary rise equation may be used to relate the size of the largest water-filled pores (the size of a cylindrical pore with similar water-retentive properties) to the force required to remove water from wet soil, $h = 2s/dgr$ as described earlier. Such air-water interfaces in small pores are essentially hemispherical so that this radius also approximates the spherical equivalent of the irregularly shaped pore.

The capillary rise equation can be applied to a porous soil system in the wet range (Fig. 19). Pores in the porous plate are all small enough to remain

water-filled at the elevation of the hanging water column (smaller than the radius of a capillary tube which would raise water to this elevation). The radii of the air-water interfaces in the soil are given by the capillary rise equation. If the elevation h is decreased, larger pores become water-filled; if h is increased, only smaller pores remain water-filled. The elevation is an index of the matric potential, and the potential on a mass basis may be obtained by multiplying by the acceleration of gravity: hg (or if potential is on a volume basis, the value is hDg, where D is the density of water). The pressure in the water just below the interface as given by the equation $p = 2s/r$, and if the radius is taken as negative (the radius of a raindrop is positive and the pressure inside also is positive), the pressure may be seen to be negative or less than the pressure of the atmosphere. The use of the hanging water column apparatus (Fig. 19) is confined to the very wet range because of the limitation in the length of a hanging water column due to possible cavitation. However, positive pressures applied in a pressure chamber over a porous plate upon which samples are placed achieve the same result without any serious limitation to the range of measurements which can be made.

Matric potential in wet soil equals the work done per unit mass or volume to remove water from the air-water interface or against surface tension forces. As the soil becomes drier, larger and larger proportions of the water are associated with particle surfaces, and more work per unit quantity must be done to remove this water, which is tightly held by adsorptive forces in the particle surface. Hence, matric potential may be seen to involve both surface tension and the attraction of particle surfaces for water. Curves showing this relationship over a wide range of water content are obtained experimentally and used to characterize different soils (**Fig. 21**). However, since it is the porosity (a characteristic dependent not only upon textural classification but also upon the degree and kind of aggregation) that defines the relationship, different curves can exist for the same basic soil material treated differently.

The water potential–water content curves (Fig. 21) are desorption curves obtained in the process of removing water from saturated soil. A slightly different curve would exist for the wetting cycle because filling and emptying of large pores is controlled by the size of entryways rather than the size of the pore itself. Air is easily trapped in large pores during filling, and water is retained in large pores during drying, until entryways are emptied. This hysteresis phenomenon sometimes complicates use of water potential–water content curves, since it is impossible to determine one from the other without knowledge of the wetting history.

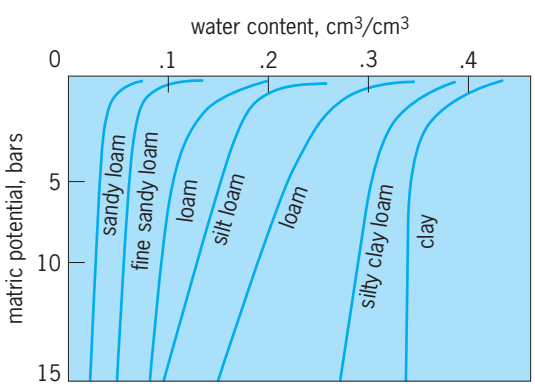

Fig. 21. Relationship between water potential (index of retentive forces) and water content for different soil materials. Wide variations often exist, even for materials in the same textural class. 1 bar = 10^2 kPa; 1 cm^3 = 0.6 $in.^3$

Water movement and retention. Water moves in soil in response to the net force acting. Above the water table in the region of a soil profile primarily involved with growing plants, matric forces are the major forces involved. Only near the water table where soil is very wet do gravitational forces play an important role. Where matric forces dominate, the water flow (flux) is the product of a conductivity factor and the moving force, which is the gradient of the matric potential. The conductivity factor is expressed as the product of a dimensionless number between 0 and 1 which expresses the degree to which the channel is filled with water and the value of the conductivity for saturated flow. This relationship describes saturated flow when the channels are full and where the matric potential is replaced by an equivalent potential derived from gravity and positive pressure, which is associated with saturation.

The channel-filling factor falls off rapidly as large pores empty, so that the cross section available for liquid flow is greatly reduced. As the soil desaturates, the pressure potential disappears and is replaced by a negative matric potential term. Matric forces are "pulling" rather than "pushing" forces, and water is pulled from regions of high potential into regions of low potential. Liquid water flows through water-filled interstitial space and along surfaces. The air-filled pores contribute nothing to liquid flow. Under such circumstances it may be seen that coarse materials, such as gravels, which would have high conductivity when saturated, would have extremely low conductivity at low water content. Thus water movement in unsaturated soil may be retarded in regions of large pores as well as in regions with fine pores. Aside from the influence of air-filled pores in reducing conductivity, conductivity also is reduced, beyond the amount expected from reduction of flow cross section, as pore size is reduced. The conductivity of capillary tubes varies with the fourth power of the capillary radius (Hagen-Poiseuille equation). On a unit area basis, the variation is with the square of the radius. Hence, with the force term held constant, reducing pore radius by a factor of 2 would reduce flow by a factor of 2^2, or 4.

Movement of water into soil, infiltration, and redistribution of water in soil following water addition slow down with time as a consequence of both changing gradients and reduction of unsaturated hydraulic conductivity as water moves from wet zones into drier zones. Slow redistribution is of particular

importance to consideration of water retention in soil. Water is nearly always moving, downward into dry soil or toward a water table, and upward in response to evaporative forces at the soil surface. Hence, soil has no unique retentive capacity. Water-retentive capacity is a dynamic property and must be defined in the context of change with time. A practical field capacity may be specified as the quantity of water in a defined depth of soil when rate of downward flow, or loss from a designated part of the profile, reduces to a value beyond which any further loss may be regarded as negligible. For agricultural purposes, and depending upon the nature of the profile, this rate would be reached in a period of time ranging from a fraction of a day to 10 or more days. The presence of soil horizons, involving either coarse or fine layers, reduces liquid water flow rates and increases water retention, generally reducing the time required to reach a negligible loss value.

An important application of unsaturated flow of water in soil concerns contamination of soil by radioactive materials or by other chemical or biological substances in spills. The seriousness of contamination depends upon the amount of water that is applied to soil as precipitation or through artificial means. From an agricultural point of view, slow movement of liquid water deep in the profile below the normal depth of roots may be of little consequence. However, when soil contains contaminants, particularly long-half-life isotopes, such flow could become important. Under some circumstances, the presence of layers of sand or gravel might constitute an effective barrier to the spread of some contaminants, but not when long periods might allow diffusion of substances in extremely thin films, which might under some conditions be present on particle surfaces or in vapor form.

Water vapor is present in soil air with relative humidity ranging from 98.9 to 100%, existing in soil wet enough to support plant growth. At uniform temperature the sum of matric and osmotic potentials, which affect evaporation, may be equated to relative humidity by the equation: matric + osmotic potential = $(RT/M) \ln (p/p_0)$, where R is the universal gas constant, T the Kelvin temperature, M the molecular weight of water, and p/p_0 the vapor pressure divided by the vapor pressure of saturated air at the same temperature, or the relative humidity. Where gradients of matric plus osmotic potential exist in soil, vapor pressure gradients also exist, and vapor flow can take place, even when liquid flow is negligibly small. If temperature gradients also exist, appreciable vapor flow can occur because of large vapor density gradients. However, in the absence of temperature gradients, vapor flow over large distances is small in dry soil. Because of low hydraulic conductivity values in dry soil, evaporative water loss from below depths of 25–50 cm (10–20 in.), in the absence of plants, generally is negligibly small, and moist soil, even in deserts, may be found below such depths. Flow of water in dry soil is primarily in vapor form, and the insulation properties of dry soil are high, so that high temperature gradients required for rapid diffusion of vapor do not exist. Where deep-rooted plants, particularly trees, are present without a water source, soil can be dried to great depths through the process of plant water uptake and evapotranspiration. Plant uptake has been used to remove unwanted chemical substances from soil and, in some coastal reclamation projects, to dry soil.

Air composition and flow. Gaseous composition of soil air, apart from water vapor discussed above, tends toward the composition of the atmosphere, with somewhat higher concentrations of carbon dioxide and lower concentrations of oxygen due to metabolic processes which utilize oxygen and give off carbon dioxide. Exchange with the atmosphere is most rapid in dry soil having high porosity and least rapid in dense or wet soil. The composition of the soil air affects plant growth and the growth of microorganisms and insects which inhabit the soil. Hence, limited air exchange with the atmosphere often has a marked effect on plant disease and plant growth generally. Oxygen requirements of plant roots vary with different plants, so that plant composition often is dependent upon aeration characteristics of soils.

Solute movement. Physical conditions of soils affect solute movement. Water is necessary in the vicinity of roots to solubilize nutrients contained in clays and other minerals and to form liquid paths through which the nutrients may diffuse and become available for absorption by roots. Also, solutes may be carried in water taken in through roots to form the transpirational stream. *See* PLANT MINERAL NUTRITION.

Solutes may be carried with moving water downward out of the rooting zone of plants and upward in moving water due to evaporation at the soil surface. In the latter case, transport to the surface and deposition through evaporation of pure water constitute a concentration process which may lead to high salt concentration in and on the soil surface. *See* ROOT (BOTANY).

Mechanical strength. Single-grain soil materials like dry sands have only a limited capability to withstand stress or compressive forces when unconfined. With small amounts of water added, this capability is increased because of surface tension forces in air–water interfaces which tend to hold particles together. Elimination of air–water interfaces by further additions of water, together with lubrication of particle surfaces, reduces the ability to withstand stress. Aggregated soil behaves somewhat in the same fashion except that stabilizing materials such as humus tend to bind particles into aggregates. Soil, particularly when moist, has some tensile strength. However, tensile strength is limited, as may be seen with swelling soils which may shrink and form large cracks upon drying. Compressive strength varies with degree of compaction brought about by animal or vehicle traffic, particularly at critical water contents somewhere between wet and dry. The water content for maximum compaction depends upon textural characteristics and types of clay minerals present. In a like manner an optimum water content generally exists for maximum effectiveness

of tillage operations designed to fragment the soil and reduce hardness.

Mechanical strength is an important consideration in agriculture as well as construction engineering. The agriculturalist needs soil which is strong enough to support small plants and trees, yet soft enough to permit easy movement of roots. Additionally, the soil should be sufficiently resistant to the mechanical forces of rain and running water to hold its position. The engineer needs soil with minimum shrink and swell properties and maximum strength over a wide range of water contents. *See* SOIL MECHANICS.

Soil temperature. Soil temperature depends upon absorption of solar radiation, reradiation from the surface, conductive exchange with the air, heat flow within, and the heat capacity of the soil. Soil color and surface texture influence both absorption and reradiation. Smooth, light-colored materials generally reflect light energy and are poor radiators, while rough, dark materials absorb or reradiate energy best. Thus, rough, dark soils tend to warm faster than smooth, light-colored materials. Organic residues on the soil surface play a major role in determining soil temperature through interception of both incoming and outgoing radiation and reduction in the velocity of air movement at the soil surface. Water content is the major factor involved in both heat transfer and heat retention, increasing both thermal conductivity and thermal capacity. Change of state of soil water—freezing, thawing, and evaporating—involves significant quantities of energy as latent heat. Soil temperature at a given depth and time depends on both heat conductance and storage, and complicated mathematical models are required for its prediction.

Plant growth and biological activity. Growth of plants, microorganisms, and insects in soil all involve establishment of an optimum physical environment. In turn, the presence in soil of organic materials and their decomposition products profoundly affects these physical properties. Broken or partially decomposed organic materials behave somewhat the same as mineral particles, and the decomposed material or humus acts as a binding substance imparting stability to mineral particle arrangements of soil aggregates. From a soil physical point of view, the major factors affecting biological activity are water, temperature, and gas composition. With some exceptions, soil supporting growing plants must have a matric-plus-osmotic potential ranging from about −15 bars (−1.5 megapascals) up to a small fraction of −1 bar (−0.1 MPa). If the soil is too wet, aeration becomes limiting. But, because aeration depends upon the quantity of air–filled pores, which will be different in different soils, it is difficult to specify a definite water potential value which will be limiting. Also, numerous agricultural plants grow best when the water potential is above −1 bar, even though plants might survive at potentials even below −15 bars.

Soils having mostly fine pores tend to remain wet in the spring and to warm slowly, so that many plants are delayed in growing because of low temperature. Optimum temperature for plants varies widely with species, and temperature thresholds exist for many plants; for example, corn and tomatoes grow poorly or not at all until day-time temperatures are well above 59°F (15°C), and growth is reduced if soil temperature greatly exceeds 86°F (30°C). However, many plants do well in cold soil.

Mechanical properties of soil can restrict root growth. Hard pans or plow pans formed by tillage operations often limit the growth of all but the hardiest of plants. Such pans may reduce the depth of the rooting zone. The hardness of the soil also is a factor in plowing and tillage operations. Since soil water content affects hardness, certain tillage operations are timed so as to take advantage of a water content that produces softer soil. If too wet, however, a soil may puddle badly when tilled. Thus an optimum water content for such operations exists. Appreciable attention is being given to no-till or minimum tillage to avoid creating hard soil by excessive traffic compaction and to maintain the protection of plant cover against erosion. Weed control, one of the important functions of tillage, is done chemically when these practices are followed. *See* HERBICIDE; PLANT GROWTH.

Walter H. Gardner

Spatial Variability

Soil properties vary vertically at a given site, horizontally from place to place, and both vertically and horizontally with time. Determination of the spatial variability of soil can be accomplished by kriging, a geostatistical technique that was developed to aid mining engineers make better estimates of ore deposits in unsampled locations. Geostatistics is the statistics of spatially correlated data; since soil properties are generally correlated spatially, geostatistics and kriging have become popular methods for describing and dealing with this complex aspect of soils. Kriging is a technique of making optimal, unbiased estimates of soil properties at unsampled locations.

Describing spatial variability of soil is important only insofar as it enables users of land resources to make better predictions of soil behavior and performance. The precision with which soil behavior and performance can be predicted depends on the uniformity of the land in question. While it may be possible to use an average value to estimate soil behavior or performance, it is often not very useful and sometimes dangerous to depend on averages, especially if different parts of a land parcel behave very differently. For example, a home owner who plans to install a septic tank needs to know the hydraulic characteristics of the plot and not the average value for the neighborhood. Since the performance of each plot as a septic tank field depends on the hydraulic behavior of the field, there is economic value in any method that enables a user to estimate the local value with greater precision than the regional mean. Kriging achieves this aim by using observed values in the neighborhood to estimate values at unsampled locations.

If all observed values are identical, the average of the observed values would be an excellent estimator of the same property at any point in the region. The

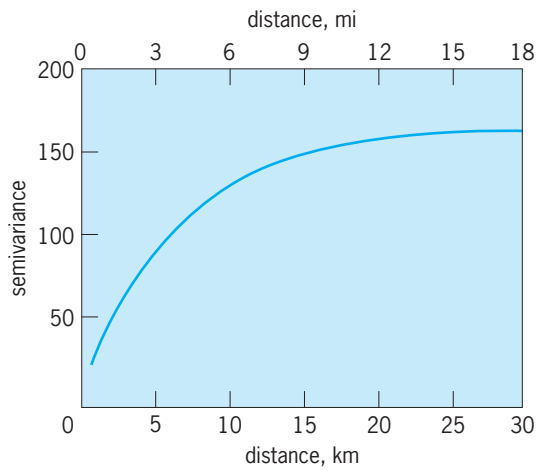

Fig. 22. Semivariogram of sand content showing a range of spatial depedence of about 9 mi (15 km). Data points represent the experimental values, and the curve is an exponential model.

problem arises when the variance in the observed data is high, and the user is unwilling to accept the mean as an estimator of values at unsampled locations. In such cases, kriging can provide better results, provided certain conditions are met.

Kriging depends on the fact that closely spaced samples tend to be more alike than samples separated by large distances. The difference in clay content between soil samples taken 1 m apart, for example, more likely will be smaller than that of the samples taken 1 km apart. In kriging, the clay content is expected to vary from point to point, but the difference in clay content between two points is expected to vary only with separation distance. Very often, however, the difference varies not only with separation distance but with orientation as well.

Considering the simpler case, in which the difference varies only with separation distance, that information can be captured and condensed in a statistical representation known as a semivariogram: each point on such a curve (**Fig. 22**) represents the sum of squared differences for a particular separation distance divided by twice the number of pairs. The semivariogram is calculated from Eq. (1), where

$$\gamma(h) = \frac{1}{2n} \sum [z(x) - z(x+h)]^2 \qquad (1)$$

$\gamma(h)$ is the semivariogram [$2\gamma(h)$ is the variogram], n is the number of pairs, z is the soil variable at location x, and h is the separation distance. In the "ideal" semivariogram, $\gamma(h) = 0$ when $h = 0$; that is, samples separated by zero distance are identical. In practice, the curve frequently intersects the vertical axis at a value above zero. The point of intersection is called the nugget variance, a term that links geostatistics to its early beginning in the gold mining industry. Another characteristic of the semivariogram is that it levels off to a value called the sill. It turns out that in the ideal case the sill corresponds to the ordinary sample variance. The sill, therefore, corresponds to variances associated with samples that are spatially independent. The maximum distance at which sample pairs are correlated is called the range. For sample distances shorter than the range, the sample property is said to be spatially dependent. Note that if by chance the sampling distance exceeds the range, one obtains a semivariogram in which the nugget variance and sill are the same. For purposes of kriging, separation distances within the range of spatial dependence are required. Choosing the minimum sampling distance so that it falls well within the range requires knowledge of the study area. Trends in the data and anisotropy are frequently encountered. Trends are regular changes in a value, such as soil temperature decreasing with elevation. Special techniques are required to eliminate the trend from the data. Anisotropy is encountered when the difference between sample values depends not only on separation distance but on orientation as well.

In order to use the structure in the variance for kriging, the experimental semivariogram is fitted to a model semivariogram. The three most commonly used models are the spherical, exponential, and linear models: the exponential model (**Fig. 23**) is fitted to the experimental semivariogram.

In kriging, local values are estimated from the weighted average of observed values within the range of the point to be estimated. The estimated value of the variable z at location x_0 is given by Eq. (2), where n is the number of observed values

$$z(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i) \qquad (2)$$

$z(x_i)$, and λ_i are weights applied to each $z(x_i)$. The weights are calculated so that the estimate is unbiased and the estimation variance is minimized. The estimation involves solving a set of $n + 1$ linear equations with $n + 1$ unknowns that describe the expected autocorrelation between observed values and the value to be estimated.

For purposes of mapping, the points to be estimated can be located in a square grid (Fig. 23). The

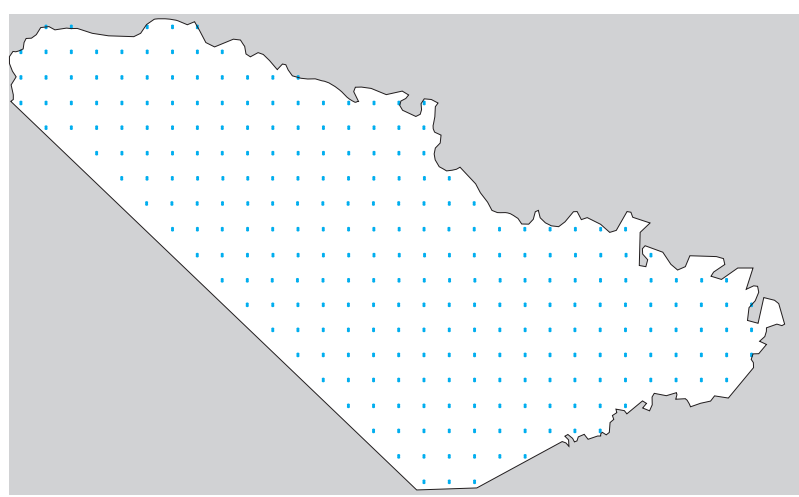

Fig. 23. Locations of 268 points for kriging sand content. The points are 1 km (0.6 mi) apart.

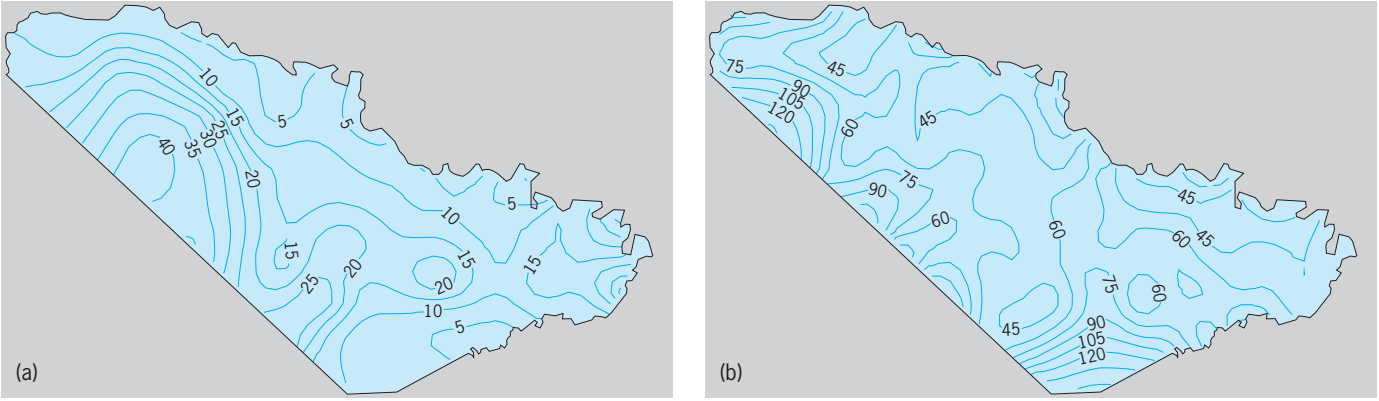

Fig. 24. Isarithm maps. (*a*) Map of actual sand content; values at the isarithms are in percent. (*b*) Map of the estimation variance of percent sand content; values at the isarithms are in percent squared.

estimated values and their corresponding estimation variance can then be used to produce isarithm maps (**Fig. 24**). Areas with low sampling densities will show high estimation variances, so that the estimation variance map can be used to guide subsequent sampling to improve map quality. If the variable in question is difficult or costly to measure, it is easier and cheaper to measure another variable that correlates highly with the variable of interest, and then compute linear estimates through the technique of cokriging. Local estimates of undersampled variables can also be improved through cokriging. Cokriging is, therefore, an interpolation technique that increases the accuracy of an estimated property by using additional information provided by the structure in the variance of a covarying property. Cokriging can be especially useful in soil science because most soil properties are spatially correlated and intercorrelated.

It is often more practical to estimate the local value of a block of land than a point in the landscape. This procedure, known as block kriging, provides a way to estimate an average value for an area of land. The area may be the size of a experimental plot, septic tank field, or a corn field. The difference between point kriging and block kriging is in the calculation of the weighing coefficients. In block kriging, the semivariances between observed points and the estimated values of point kriging are replaced by the average semivariances between the observed points and all points in the block. As in point kriging, the weights are calculated so that the estimation variance is minimized.

The ability to determine the property of a soil at any point or block in the landscape provides users of land resources with information that permits sounder economic and safer environmental land-use decisions to be made. Spatial variability can be costly to farmers and other users of land resources. Farmers who apply agricultural chemicals uniformly in a field with spatially varying soils will add too much to places where the requirement for the chemical is low, and too little where the requirement is high. Farmers who cultivate large tracts of land are finding it profitable to invest in farm machinery that can match application rates to spatially varying soil and crop requirements with a high degree of precision. While economic considerations serve as the main incentive, stricter environmental regulations may compel farmers to adopt precision farming, and kriging offers a way to predict optimum application rates of agricultural chemicals. Goro Uehara

Management

Soil management can be defined simply as the manner in which people use soils to produce food, fiber, and forages. Soil management includes determination and use of many factors and practices, such as land survey maps, cropping systems, organic matter and tilth, soil fertility, salinity, and irrigation. Most often, it is not a single soil that is managed; management is carried out on a field or a portion of a landscape composed of a number of physical and biological features, like climate, vegetation, topography, and drainage. In a true sense, applying soil management strategies involves an integration of a number of factors. Such integration involves a difficult synthesis of many individual characteristics, both measured and observed, to arrive at a meaningful interpretation of how a soil responds to management. In fact, the value of any agricultural soil is determined by how well that soil responds to proper management.

Standards. Meaningful assessment of how soils are responding to a particular management system can be judged only if standards are established. There are two primary standards for judging quality of management: prevention of soil deterioration or degradation, and improving the soil system in ways that result in increased plant production. Examples of deterioration that must be prevented are excessive erosion, fertility depletion, and accumulations of salt within the rooting zone of plants. Examples relating to improving the soil system are conservation of precipitation, irrigation development, reclaiming salt-damaged land, and increased water use efficiency by proper fertilization. *See* AGRICULTURAL SOIL AND CROP PRACTICES.

The two goals are very broad and therefore require the integration of many aspects of science. Incorporation of the basic areas of chemistry, physics, and microbiology into management planning is essential. The sciences of ecology, genetics, and various types of engineering are also essential to modern soil management. Computer modeling will play an increasingly important role in soil management by making simulation models directly available to farm and ranch operators. *See* SIMULATION.

Land capability classification. The first step in planning a management system is assessment of the soil and environmental resources. This is accomplished through land classification—more specifically, through land capability classification, which groups land according to properties essential for identifying the opportunities or constraints the land offers for various uses.

Interpretations of land capability classifications are most often based on physical, chemical, and economic considerations. Classification provides assessment of land suitability for agricultural and other uses, identifies adapted crops, estimates yields of crops under defined systems of management, indicates presence of specific soil management problems, and delineates opportunities and limitations of various management practices. Through land capability classification, there usually are identified several alternative management strategies for a given type of land. From this information, analyses can be made to determine which land use or management strategy will be most desirable for both economic and wise land utilization.

The basic requirements for developing a good classification system are soil-survey and soil chemical and physical data; climatic, topographic, and hydrologic data; field and laboratory research data; long-time land-use records; and experience and observations of land response to various uses.

Soil surveys are basic in that they identify soil properties important for determining land-use capability. Surveys provide maps showing the location and boundaries of soils. Soil depth, texture, kinds of minerals, salinity and kinds of salts, and acidity are some properties utilized in determining land capability. The particular set of properties needed for determining land capability is dependent on the particular use.

Land classifications cannot be static because they depend upon interpretation. As technology changes and economic and social conditions change, interpretations change. With basic soil, climatic, and topographic data and maps, however, these interpretations can be revised easily. Combinations of alternative land-use and management strategies, as well as social and economic conditions, do change, but the physical and chemical factors of well-managed land do not change very much.

Land classifications have been developed for specific purposes, for example, tax assessment, sales and credit, soil and water conservation, irrigation potential and management, wildlife suitability, watershed management, recreational uses, industrial uses, or highway construction.

Effective extrapolation of management or use strategies to the land depends on an adequate inventory of basic data and how well an existing classification is chosen or a classification developed that fits the specific need at a given time and place. With appropriate facts and maps, through land classification, predictions can be made about the results of using a specific type of land in a particular way. Then, planning of land use includes the practical combinations of management practices required, and the effects of management on the quality of land resources. Without land classification and appropriate maps, it is difficult to extrapolate experience and research results to the land.

Land capability classification systems have been utilized for many years by the Soil Conservation Service, USDA, and other land resource planning and management agencies. Their utilization is being enhanced greatly through the use of computer data storage and management systems. *See* LAND-USE CLASSES.

Long-term impacts. Modern soil management practices have a cumulative effect on the soil's future productivity. The foremost management question for the soil manager is whether a given practice or system is causing soil degradation. Many current agricultural problems have developed because managers did not realize the long-term impact of their management techniques. Some practices may appear to be sustainable for periods of time equal to a human lifetime, and yet they could be causing slow soil degradation that will ruin the soil for long-term use. Therefore, long-term plans that identify causes of soil deterioration and avoid it successfully are essential to a nation's future agricultural productivity. For example, proper use of fertilizers and manures can enhance long-term productivity and maintain environmental quality. Poor management choices including fertilizers and manures can actually ruin a soil and damage the environment, almost irreparably. Civilizations have often missed the "clues" that a particular system was not sustainable. This has led to loss of their food supply, and with their being conquered eventually or ceasing to exist.

Soil water. Water conservation and efficient use of stored water are important issues within soil management. Since plant growth is limited by water supply in almost every climate situation at some point in the crop cycle, having adequate stored soil water is critical. Irrigation becomes important in climates where the natural precipitation will not meet plant demands or in areas where rainfall patterns are erratic.

Water conservation principles are the same in either rain-fed or irrigated situations, and can be divided into two phases: water capture in soil, and water retention in the soil. Water capture is maximized by maintaining open pores at the soil surface and protecting these surface pores from raindrop impact which would otherwise destroy surface

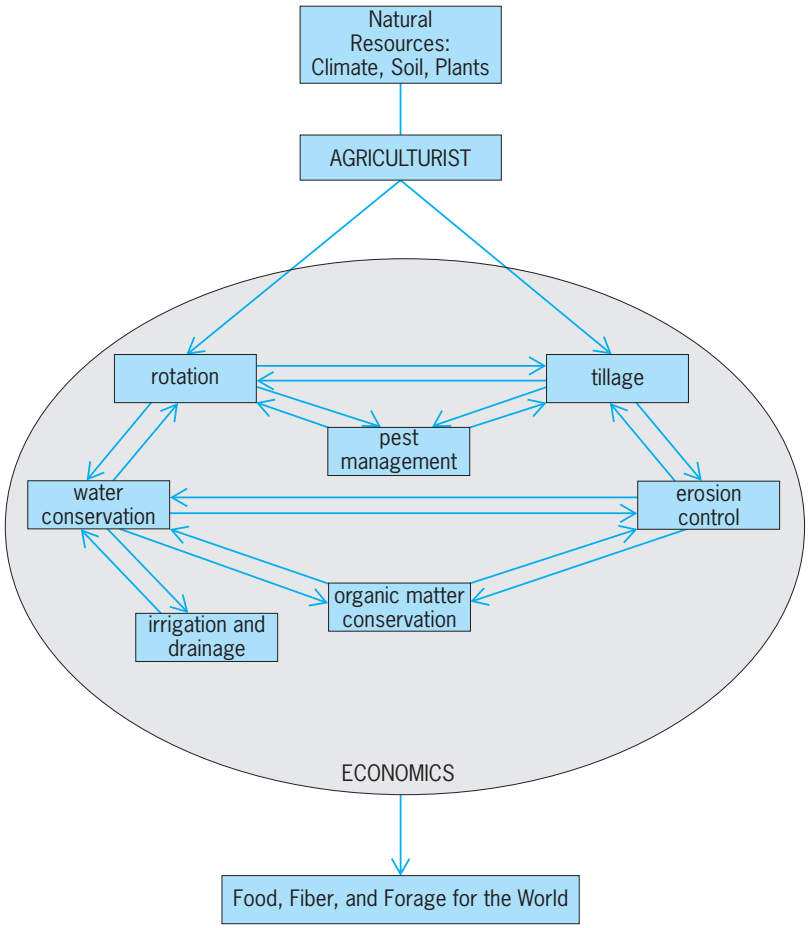

Fig. 25. Diagram showing an integrated system of soil management.

soil structure and ultimately plug the pores. Crop-residue mulches are effective protectors of structure, and these are best maintained by reducing tillage. Residues also slow runoff water and increase the opportunity for infiltration into the soil. The second phase of water conservation, retention of captured water, requires prevention of weed growth that would waste the captured water, and reduction of evaporation. Protecting the soil surface from exposure to sunlight and wind by maintaining crop residues at the surface is an important management tool for evaporation control. Residues serve as reflectors of light energy, as insulators to keep the soil surface cool, and as wind deflectors. Reducing and even eliminating tillage maximizes the crop residue on the soil surface and consequently maximizes water retention.

Irrigation is not always an option. Many locations in the world do not have adequate water supplies to allow irrigation, even though the crops could benefit from additional water. In some cases, underground water supplies are present but are so deep that the additional crop growth with irrigation would not offset the cost of pumping the water. Soil factors such as texture, depth, and salt content also dictate the feasibility of irrigation. Some very fine textured soils will not absorb water readily and are not good candidates for irrigation; others are so shallow that the storage capacity is too little. If soils have too many soluble salts, they are not suitable for irrigation. There are many combinations of climate, soils, water supply, and economic parameters that must be considered when determining if irrigation will be feasible or profitable.

All irrigation waters, whether they are diverted from rivers or pumped from the underground supply, contain some dissolved salts. Thus, when soils are irrigated they receive a dose of these salts. If the irrigation manager does not plan for these salts, the soil can become saline and lose productivity. *See* WATER CONSERVATION.

Integrated systems. Soil management involves integration (**Fig. 25**). The resource block includes climate, soil, and plants; these are the natural-resource inputs. Modifications can and will be made on them, but for the most part they are the constants of the system. The agriculturist enters the scheme at this point to impose a management system, which has as its goal the production of food, fiber, or forages. Figure 25 also shows that the cropping system imposed is directly related to the soil, climate, and plant resources. The manager implements a cropping system through choice of crop rotation and tillage system. Choice of tillage system is often conditioned by the rotation chosen.

Erosion control, organic matter conservation, and water conservation are the classical concerns of soil management; they are all highly linked. A change in one or more of these three factors creates feedbacks to the others. For example, control of erosion leads to direct savings in soil organic matter and soil water and thus soil conservation. Furthermore, the control of erosion, which relates to the first goal of preventing soil deterioration, is linked directly to the second goal, namely, improvement of the soil system. Erosion prevention techniques also improve soil water storage by reducing water runoff. Conservation of organic matter may increase water infiltration, which decreases runoff and thus decreases erosion. The other blocks in Fig. 25 represent pest management and irrigation and drainage. They are equally important, but have not been studied to as great an extent as the first three factors. In fact, pest and pesticide management are relatively new concepts in soil management. Soil tillage methods and rotations are intimately involved in pest and pesticide management.

Finally, basic economic principles are involved in all sound soil management systems, as indicated in Fig. 25 by the enclosure of all other factors by economics. This is true in even the most primitive civilizations. If the people cannot live on the food or income produced, the system cannot survive. In more developed nations, quality of management has even influenced the potential value of the land. The ability to cope with the problems of a particular soil eventually determines its economic value to society.

Forces external to the system, such as changes in government policy, can greatly impact the overall management system. For example, the government has funded land set-aside programs for multiple

reasons, ranging from attempts to decrease production of various crops to efforts to slow soil erosion. These interventions cause multiple, and often unexpected, interactions within the system. The programs encourage producers to take land from production of annual crops, such as corn, cotton, sorghum, and wheat, and to replace them with perennial grasses. During the contract period, the producer cannot harvest or graze the grass, and it is left as a cover crop on the soil.

Quite obviously these land set-aside programs have an economic impact on both the producer and local community. The producer experiences less risk because inputs to the system are low and the income is stable, no matter what the weather is. At the same time, the community merchants and the local labor market are often negatively impacted because the merchant sells fewer inputs and there is less need for seasonal labor. In contrast to the mixed economic outcome, the set-aside programs have nothing but positive influences on the soil itself. Soil erosion is slowed, relative to the conditions prior to the set-aside program, and soil organic matter contents stabilize or rise. However, these potentially positive soil changes may not be long-lasting because government policies change on a short-term basis (5–10 years) while soil changes require multiple decades to be of long-term value to the nation. Depending on the management scheme adopted by the producer, the benefits of 10 years in perennial grass can be lost in 1–2 years.

High-quality, efficient management of the natural resources of climate, water, and soil is needed to provide the food and fiber that will sustain life on Earth. Soil management decisions will have an important effect on the future food supply.

G. A. Peterson; W. O. Willis

Salinity

Soluble salt and exchangeable cation concentrations play major roles in determining the pH, physical characteristics, and chemical composition of soils. When a salt dissolves in water, it dissociates or separates into cations and anions. The predominant cations in salt-affected soils are calcium (Ca^{2+}), magnesium (Mg^{2+}), sodium (Na^{+}), and potassium (K^{+}); the predominant anions are chloride (Cl^{-}), sulfate (SO_4^{2-}), carbonate (CO_3^{2-}), and bicarbonate (HCO_3^{-}).

Clays and organic matter contain negative electrically charged sites. In salt-affected soils, this charge is satisfied by calcium, magnesium, sodium, and potassium ions. The cations, bound to the exchange sites by electrical charge, are known as exchangeable cations because they can be removed from the charged surface only by replacement with another cation from the soil solution. *See* HUMUS; SOIL CHEMISTRY.

Classification. Each soil can be classified as normal, saline, sodic, or saline-sodic, based on its salt content and exchangeable cation ratios.

Normal soils. These soils do not contain sufficient soluble salts or exchangeable sodium to adversely affect plant growth or soil physical properties. Normal soils have saturation paste pH values less than 8.3 and saturation paste extract electrical conductivities of less than 4 deciSiemens per meter (dS/m) and an exchangeable sodium percentage less than 15. A saturation paste is made by mixing just enough distilled water with a soil sample to fill the voids without having excess water standing on the surface of a well-mixed sample after 4–16 h. These electrical conductivities are defined upper limits, but if salt-sensitive plants are grown on soils with an electrical conductivity of 3.5 dS/m, a significant yield reduction will result. Likewise, using a high-volume sprinkler system to irrigate a soil with an exchangeable sodium percentage of 12 could produce serious runoff rates because of low infiltration rates. *See* HYDROLYSIS; PH.

Saline soils. These soils contain sufficient soluble salts (electrical conductivity greater than 4 dS/m) in the upper root zone to reduce yields of most cultivated crops and ornamental plants. The exchangeable sodium percentage is less than 15, the sodium absorption ratio is less than 13, and the pH is less than 8.3. The predominant cations are calcium, magnesium, and, in a few cases, potassium. The predominant anions are chloride and sulfate. Water entry and movement through these soils is not inhibited by high exchangeable sodium concentrations. Osmotic effects and chloride toxicity are the predominant causes of plant growth reduction. *See* OSMOSIS.

Sodic soils. These soils are lower in soluble salts than saline soils (electrical conductivity less than 4 dS/m). The exchangeable sodium percentage is greater than 15 and the sodium absorption ratio of the saturation paste extract is greater than 13. The pH of the saturation paste is greater than 8.5. Bicarbonate, carbonate, and hydroxide (OH^{-}) ions are the predominant anions in these soils; these anions cause calcium to precipitate from solution as calcium carbonate ($CaCO_3$; lime). The combination of high exchangeable sodium percentage and pH, and low electrical conductivity and exchangeable calcium causes the clay and organic matter to disperse, which in turn destroys the soil structure or tilth, causing so-called slick spots. These spots have extremely low rates of water and air exchange. They often have a black, greasy, or oily-looking surface due to the dispersed organic matter. Vegetation may be absent because of low water infiltration and insufficient plant-available water.

Saline-sodic. These soils are similar to saline soils in that the electrical conductivity is greater than 4 dS/m and the pH is below 8.3. Saline-sodic soils differ from saline soils in that more than 15% of the exchangeable cations are sodium and the saturation-paste sodium absorption ratio is greater than 13. The anions are a mix of bicarbonate, chloride, and sulfate. As long as the electrical conductivity remains above 4 dS/m, infiltration rates and hydraulic conductivities are similar to those of normal or saline soils. Irrigating saline-sodic soils with water having low salt concentrations will convert them into sodic soils if they do not contain gypsum (a calcium sulfate mineral).

This happens as the electrical conductivity decreases without a decrease in the exchangeable sodium percentage, causing the undesirable properties of sodic soils to be expressed. It is not uncommon to have a mix of two or more classes of salt-affected soils within a field. Salt-affected soils tend to be highly variable from one part of a field to another.

Sources. Most soluble salts and exchangeable cations are derived from rock and mineral weathering of the soil parent materials. In high-rainfall, humid, and tropical areas, rain and melting snow leach the salts from the soil as they form. In arid and semiarid areas, the annual evapotranspiration potential is greater than the total annual precipitation, and the salts are not always leached from the soil as they are released. With time, they may accumulate in the root zone at concentration levels that affect plant growth.

Salts can also accumulate above shallow water tables as water moves to the soil surface by capillary rise (wicking) and evaporates, leaving the salts on or near the surface. Shallow water tables may occur naturally, induced by irrigating poorly drained areas, by irrigating upslope from lowlands, or by construction activity that blocks natural subsurface lateral drainage.

All natural waters contain dissolved salt. In many arid and semiarid areas, good-quality irrigation water (low in salts and low in sodium) is not available; consequently, irrigation water is used that contains more salts or sodium than is desirable. If sufficient water does not move through the soil and leach the salts below the root zone, salts or sodium will accumulate in the soil. It is often stated that under irrigation "hard water makes soft soils and soft water makes hard soils." This implies that irrigation waters containing predominantly calcium and magnesium salts (sodium absorption ratio less than 3 or 4) tend to promote a more friable soil condition than do waters with high sodium concentrations.

Four conditions must be satisfied in order to remove soluble salts and excess sodium from soils: (1) less salt must be added to the soil than is removed; (2) salts must be leached downward through the soil; (3) water moving upward from shallow water tables must be removed or intercepted to avoid additional salts moving back to the soil surface; and (4) in sodic and saline-sodic soils the exchangeable sodium must be replaced with another cation, preferably calcium, and the sodium leached out. Applications of soil amendments (gypsum, iron sulfate, sulfur, or sulfuric acid) are beneficial only on sodic soils when leaching also occurs and on leaching of saline-sodic soils that do not contain natural gypsum.

Saline and sodic soils are found primarily in arid and semiarid areas of the world. Exceptions are recently drained coastal areas, salt marshes, and soils formed in depressions from marine deposits where the weathering products are not leached from the soil. Sodic and saline-sodic soils can also form where mists are carried from ocean waters by the wind and deposited on soil surfaces in arid and semiarid coastal areas. Aridisols and Entisols include most salt-affected soils. Low rainfall and unweathered soil materials result in insufficient salts leaching from the root zone of these soils. Mollisols, Alfisols, and Vertisols also contain considerable saline and sodic soil areas.

Human activities such as spills or intentional dumping of salts or solutions from oil well drilling-mud ponds, mines, food-processing plants, municipal sewage water, power-plant cooling-tower water, or heavy applications of wood ash can induce saline and sodic conditions in any soil when soluble salts are applied faster than they are leached from the soil.

C. W. Robbins

Identification and mapping. Because of the highly variable nature of salt-affected soils and the expense of reclamation, field mapping is prerequisite to effective reclamation. Visual observation mapping of plant growth and soil appearance is the quickest and least expensive method for mapping low-, medium-, and high-salinity impacted areas of the landscape. Native plant species and stunted growth or leaf burn in cultivated crops are effective identifiers of saline and sodic areas. Saline soil areas will also have white salt crusts, while sodic areas will be barren with dark-to-black, oily-looking slick shining surfaces bordered by water-stressed plants. Remote-sensing procedures, such as aerial photography, videography, infrared thermometry and imaging, multispectrial scanners, microwave sensors, and time domain reflectometry, also are used for salinity mapping. Once the productive and problem areas are defined, deep sample holes can be bored to determine if a high water table is the salt source. Water samples from high water tables should be analyzed for salt concentration and type. Soil samples should also be analyzed to determine salt type and concentration. Exchangeable cation concentrations are also needed to determine the degree of the sodium problem. Irrigation-water salt concentration and cation ratios are also important factors in salinity management. Once the salt sources, concentrations, and cation ratios are determined, a reclamation plan can be developed; or if reclamation is not practical, crops tolerant to the conditions can be selected.

Reclamation. Four conditions must be satisfied in order to reclaim salt-affected soils by removing soluble salts and excess sodium: (1) less salt must be added to the soil than is removed; (2) salts must be leached downward through the soil; (3) water moving upward from shallow water tables must be removed or intercepted to prevent additional salts from moving back to the soil surface; and (4) in sodic and saline-sodic soils the exchangeable sodium must be replaced with another cation, preferably calcium, and the sodium leached out. Applications of soil amendments (gypsum, iron sulfate, sulfur, or sulfuric acid) are beneficial only on sodic soils when leaching also occurs and on leaching of saline-sodic soils that do not contain natural gypsum. Adding chemical amendments such as gypsum to saline soils only adds more salts and is not needed unless the water has a sodium adsorption ratio of 10 or greater.

C. W. Robbins

Effects on plants. Many ions are essential to plant growth as major or minor nutrients. However, when ion concentrations become too high, plant growth is adversely affected by either the toxic effect of a specific ion or the general effects of high ion concentrations. Salinity decreases plant growth through a combination of nonspecific ionic effects and by causing a decrease in the water potential of the soil, which is principally an osmotic effect.

The point at which salinity limits plant growth varies because plants have adapted to a wide range of salinity environments. The ocean, which has salt concentrations in excess of 35 parts per thousand (ppt), is abundant in plant life and contains over half of the Earth's plant biomass. Concentrations at which specific ions become harmful to plant growth also vary over several orders of magnitude. For example, boron is toxic to some plants at soil-water concentrations as low as 0.05 mol/m^3, whereas some plants can tolerate chloride concentrations as high as 20 mol/m^3. Sodium and chloride are the most abundant salinity-inducing ions in soils and water; however, significant amounts of calcium, magnesium, sulfate, carbonate, and bicarbonate ions are also found in nature. The proportions of these ions vary with respect to soil types and local geology.

Terrestrial plants that are tolerant of high concentrations of soluble salts in their root zones (the area around the root from which a plant extracts water) are known as halophytes. Halophytes can survive and complete their life cycles at optimum salt concentrations of 1.2–30 ppt in their root zone. Most terrestrial plant species are not adapted to high salt concentrations and may be considered glycophytes.

In agriculture, salt is a serious hazard in irrigated areas if growers do not leach their soils properly during irrigation, fail to provide adequate drainage for their crops, or allow their water tables to rise too near the surface. High concentrations of salts in the water used for irrigation may also damage crops or reduce yields. High-salinity waters may include ground and surface water and water recycled from municipal and industrial uses. Salinity acts as an environmental plant stress that may cause leaf damage of reduced growth or, at high concentrations, may be lethal. Plant responses to excess salts in their root zones or on leaf surfaces (from ocean sprays or irrigation) are quantitatively dependent on salt concentration, composition, and time of exposure. Plant sensitivity to salt also varies according to growth stage. Generally, plants are salt-tolerant during germination, are most sensitive during the seedling stage, and become more tolerant with maturity. Plants also may show increased salt sensitivity during their reproductive stage, as seen by decreased pollen viability and seed setting ability.

Crop salt tolerance. Salt tolerance is the capacity of a plant to endure excess salt in its environment. This characteristic is quantitative and influenced by many soil, climate, and cultural factors. Tolerance assessment may be based on the ability of the plant to survive, to produce high yields, or to withstand adverse growth reductions. In nature, the measure of tolerance may be the ability to survive, reproduce, and compete with other species; whereas in sustenance agriculture, tolerance may be related to both survival and productive yield. In commercial agriculture, the ability of the crop to withstand salt effects without reducing yields below the profit margin is the most important consideration. Thus crop salt tolerance is usefully described by two parameters: the threshold salinity (T) at which yield reduction is significantly measurable, and the rate (R) at which yield decreases with increasing salinity beyond the threshold (**Fig. 26**). The rate of yield decrease for most crops can be described simply as the slope of a straight line. Salinity concentrations are described as an index of the electrical conductivity of a soil-saturated paste in units of dS/m. Salt tolerance parameters are useful for predicting how one crop may compare with another under similar conditions. However, such assessments are general and relate to crop growth after germination and seedling establishment.

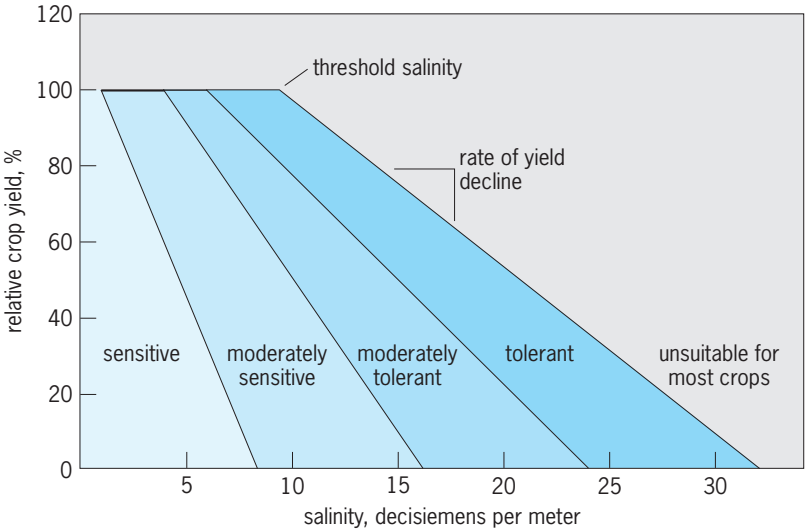

Fig. 26. Typical classifications for salt tolerance of crops based on their relative yields under nonsaline conditions in contrast to yields under increasing saline conditions.

Climate and agricultural management practices may reduce or increase the effects of salinity upon plants. Irrigation and management practices that leach salts away from, or maintain lower concentrations of salt in, the root zone during growth will reduce salt effects. Seed beds should be sloped or maintained in a manner that allows the irrigation water to move salts past the root zone. If excess salts in the seed bed are not kept low, the resulting reduction in plant stand will decrease yields far more than is predicted by salt tolerance parameters. Flood, furrow, drip, and sprinkler irrigations should also be applied at times and in ways that reduce salt accumulation on plant parts and within root zones. Climate factors such as high temperature, low humidity, and high wind speed will increase salt damage, whereas factors that reduce transpiration demand will reduce salt damage. Soil type is also an important factor. Sandy soils will not accumulate salts as readily and are easier to leach than clay soils. *See* IRRIGATION (AGRICULTURE).

Genetic variability. Many crops, such as sugarbeet, asparagus, date palm, cotton, and barley, are salt-tolerant and are standard crops in saline areas. Different types of beans and berries, as well as avocado and many fruit trees, are very sensitive to salt. In salt-sensitive species, sensitivity is often associated with the accumulation of a specific ion in leaves, usually chloride or sodium. Saline soil environments with high proportions of sulfate compared to chloride may have less severe salinity effects on chloride-sensitive plants. The ability to exclude specific ions at the level of the root or shoot is one of the major causes for genetic variability among plant varieties and ecotypes. Rice, although salt-sensitive, is grown on saline lands for reclamation because it has a shallow root zone and can be grown on flooded fields if water of good quality is available. Salt tolerance also varies less between cultivars and ecotypes of the same species.

Conventional breeding efforts to improve salt tolerance of crops include selection for more tolerant cultivars through hybridization among varieties of a species; hybridization of a cultivated species with related, wild salt-tolerant species to increase genetic variability prior to selection; and exploitation of the useful agronomic potential of wild halophytes. The ability of the grower to control the effects of salt in the field is of more greater consequence than the variability in salt tolerance among cultivars.

Morphological and physiological effects. Salinity reduces plant growth through both osmotic and ionic influences. The osmotic effects are a result of increased solute concentrations at the root-soil water interface, which create lower water potentials. Growth suppression is the result of total electrolyte concentration, soil water content, and matrix effects, and is manifested in reduced cell enlargement and metabolism. The plant suffers water stress for a short period until it can make some type of osmotic adjustment. Plants make this adjustment by accumulating more salt within their tissues (a halophytic response) or by the synthesis of organic solutes, which increases the osmotic potential of the cytoplasm so that water will flow into the root and plant tissues.

Ionic effects may be both general and specific. General ionic effects are the result of the increased ionic strength of the soil water. Ionic effects may interfere with the normal processes by which plants take up nutrients by changing the surface chemistry at the cell wall and plasma membrane. Specific ions may disrupt normal metabolic processes or upset nutrient balances. For instance, high sodium concentrations relative to other salts can disrupt root permeability to ions by displacing calcium in the plasma membrane. Upsetting calcium metabolism and nutrition within the cell may cause additional effects. At higher sodium-to-calcium ratios, soil structure, tilth, and permeability of the soil to water may be reduced (sodicity).

The specific physiological cause of growth reduction due to salt stress is undoubtedly complex in both the metabolic and genetic sense. Salt stress reduces plant growth primarily because it increases the metabolic energy needed to acquire water from the root zone and to make the biochemical and morphological adjustments necessary to maintain growth in a higher ionic environment. *See* PLANT-WATER RELATIONS; PLANTS, SALINE ENVIRONMENTS OF; SOIL FERTILITY.

Michael Shannon

Erosion

Soil erosion results from the detachment and transport of soil materials by water. Geologic erosion and erosion from human activities are the principal types. Long-term geologic erosion creates topographic features such as canyons, stream channels, and valleys. Removal of natural vegetation by human activities, such as farming, ranching, forestry, and construction, may also cause erosion.

Excessive erosion could threaten the world supply of agricultural and forest products. The efficiency of water conveyance and storage structures may be significantly impacted by sedimentation resulting from soil erosion. Excessive amounts of sediment in streams and rivers can reduce their suitability as a biological habitat and create water supply difficulties. Understanding the various types of erosion and the factors affecting erosion is necessary in identifying appropriate control practices.

Types. Erosion by water occurs when soil particles are detached from the soil surface and then transported by runoff. As runoff rate increases, small channels, called rills, begin to form. The region between rills is defined as the interrill area. When concentrated runoff is sufficiently large to cut deep channels, gully erosion occurs. Stream channel erosion may develop within a water course that has nearly continuous flow. Interrill erosion, rill erosion, gully erosion, and stream channel erosion each have distinct characteristics.

Interrill erosion. On interrill areas, raindrops impacting the soil surface serve to detach soil particles. Some of the soil particles are transported by thin interrill flow into rills. Residue materials from the previous crop that are left as a surface mulch are very effective in reducing interrill erosion. Raindrop energy is absorbed and dissipated by the residue mulch, thus protecting the soil surface.

Rill erosion. Soil materials removed by raindrop impact on interrill areas may eventually be delivered to rills where they are transported down a hillslope by rill flow. Rills are small enough to be removed by normal tillage operations. Substantial erosion may occur once rills have formed. If conservation measures are not employed to reduce rill erosion, rapid loss of soil productivity may result. Rill erosion can usually be controlled by contouring, strip cropping, and conservation tillage.

Gully erosion. Gullies are deep channels larger than rills that cannot be removed by tillage. Gully erosion generally occurs near the upper end of intermittent streams. Once they are formed, gullies become a permanent part of the landscape, and they may expand rapidly. Gully formation often develops where there is a water overfall causing the gully to move

upslope. Water moving through the gully may cause the channel to deepen. In addition, sections of the exposed banks may be undercut and slide into the gully, where they are later removed during large runoff events. Control measures, such as terraces or vegetated waterways, may be required to prevent gully erosion.

Stream channel erosion. Stream channel erosion results from the removal of soil from stream banks or beds. Runoff flowing over the side of the stream bank or scouring below the water surface can cause stream channels to erode, especially during severe floods. The major cause of erosion along stream banks is meandering. Stream channel erosion may increase when sediment delivery from upland areas is reduced through control practices or when upstream sediments are caught in water storage facilities.

Factors affecting erosion. The principal elements affecting soil erosion are rainfall characteristics, soil factors, topography, climate, and land use.

Rainfall characteristics. Runoff is rainfall that is neither absorbed by the soil nor accumulated on the surface but moves downslope. Rainfall rate and duration are important variables influencing runoff and erosion. Runoff occurs only when rainfall intensity exceeds soil infiltration rate, which decreases with time. Thus, no runoff may occur from a storm of short duration, while substantial runoff may result from a storm of the same intensity but of longer duration.

Both the rate and volume of runoff are influenced by rainfall intensity. Infiltration capacity is exceeded by a greater margin during a high-intensity storm than a less intense rainfall event. As a result, the high-intensity storm may produce a greater volume of runoff even though total precipitation was similar for the two events. The infiltration rate may also be substantially reduced by the destructive action of the storm on the soil surface.

Irrigation is used on some agricultural areas. Runoff may result from both irrigation and natural precipitation events. The runoff potential may be compounded on irrigated areas because of the increased quantities of water introduced through irrigation.

Soil factors. The physical, chemical and mineralogical characteristics of soils vary greatly, as does their susceptibility to erosion. Soil erodibility is influenced significantly by the size of primary soil particles, organic matter content, soil structure, and permeability. These soil characteristics affect the susceptibility of soil particles and aggregates to detachment.

For erosion to occur, runoff must be present. In general, as runoff rates become greater, erosion also increases. One of the most effective means of reducing erosion is to maintain high infiltration rates. Keeping crop residue materials on the soil surface to reduce sealing caused by raindrop impact helps to preserve high infiltration rates.

Topography. The degree and length of slope, and the size and shape of the watershed influence erosion. As slope gradient increases, the velocity of flowing water becomes greater. The ability of moving water to detach and transport soil particles increases substantially with larger flow velocity.

Rill erosion becomes greater on longer slopes because of an increased accumulation of overland flow. Concave slopes, with a smaller slope gradient at the bottom of the hillslope, are less erosive than convex slopes. Deposition frequently occurs at the bottom of concave slopes because of reduced transport capacity of flow.

Climate. The quantity of erosion that occurs from a given region is influenced by the total amount and intensity of rainfall. The dense vegetation found on areas that receive substantial rainfall reduces erosion potential. Regions with low rainfall and limited vegetation are often susceptible to erosion during high-intensity rain storms.

Runoff from melted snow and ice can cause serious erosion problems in colder climates. During several months of the year, frozen soil is not subject to erosion. However, if the snow cover melts rapidly and infiltration does not take place, substantial runoff may result. The rapid melting that may occur when rain falls upon a snow-covered surface can also produce significant runoff. Substantial erosion may occur as water moves over a thin layer of freshly thawed soil. In many colder climates, more erosion results from snowmelt than from rainfall.

Land use. Areas having complete ground cover throughout the year are least susceptible to erosion. Erosion from undisturbed forests is usually minimal, because a constant vegetative cover is maintained on the soil surface. On croplands, the amount of surface cover is influenced by the cropping and management conditions employed. One of the most critical periods exists after planting when residue cover is at a minimum and high-intensity rains frequently occur.

A study conducted in the southeastern United States demonstrated the effects of selected land use on runoff and soil loss (**Table 2**). The results showed that cultivated land left fallow with no vegetative cover is particularly vulnerable to erosion. Row crops such as corn and cotton grown continually on areas with steep slopes may also result in significant erosion. Planting row crops in rotation with grasses and legumes that maintain a dense surface cover substantially reduces erosion.

TABLE 2. Mean annual runoff and erosion under different land uses with mean annual rainfall of 1400 mm (55 in.)*

Land use	Runoff, mm (in.)	Soil loss, Mg/ha† (ton acre)
Rotation; coastal bermuda grass	70 (2.8)	5 (2)
and crimson clover after corn	70 (2.8)	5 (2)
Corn grown continuously	180 (7.1)	27 (12)
Cotton grown continuously	250 (9.8)	49 (22)
Fallow; cultivated with no vegetative cover	470 (19)	135 (60)

*After J. R. Carreker et al., *Soil and Water Management Systems for Sloping Land*, USDA, ARS-S-160, 1978.
†1 Mg = 10^6 grams.

Interseeding row crops with a legume can also be an effective conservation measure on areas that receive sufficient precipitation. The legume provides a protective surface cover during the critical planting period and also serves as a supplemental nitrogen source for the following cropping season. A herbicide is usually applied to kill the legume before the row crop is planted.

The dense sod found in pastures grown in humid areas is very effective in reducing erosion. Erosion is also minimal on natural rangelands where adequate surface cover is maintained. In regions with limited rainfall where bunch grasses are found, severe erosion may occur during intense storms from the exposed soil located between the bunches of grass. Reduction of vegetative cover through excessive grazing may also result in serious erosion.

On some rangeland areas, gravel and cobble materials are found throughout the soil profile. Since they are not easily transported by overland flow, gravel and cobble materials remain on the surface of eroded soils. This creates an armoring process in which the gravel and cobble materials reduce further erosion from rangeland soils.

Erosion is greatly diminished on forest lands because of the overhead canopy of trees and the surface layer of decaying organic matter. In an undisturbed forest, almost all erosion occurs from channel banks or adjacent steep slopes. Erosion rates may increase substantially on forest areas that are disturbed by timber harvesting, road construction, or fires.

Control. Contouring, strip cropping, conservation tillage, terraces, buffer strips, and stream channel erosion control measures have been used effectively to reduce the damage caused by soil erosion.

Contouring. Planting crops and performing tillage along the contours of the land can be an effective conservation measure. Surface runoff is confined in small depressions, thus reducing rill development. Ridge tillage systems are used to significantly increase the storage capacity of furrows. To maintain furrow storage capacity, row crops are planted on the top of the same furrow each year. As the slope gradient increases, the effectiveness of ridges in trapping runoff and reducing soil loss decreases.

Strip cropping. Strip cropping occurs when alternate strips of different crops are grown in the same field. The strips with the greatest vegetative cover reduce runoff velocity and capture soil eroded from upslope areas. The strip widths selected allow for the convenient use of farm equipment. For erosion control, the strips are usually planted on the contour in a rotation that shifts crops annually from one strip to the next.

Conservation tillage. Leaving a residue mulch from the previous crop on the soil surface greatly reduces erosion (**Fig. 27**). The reduction is related to the percent of residue cover left on the soil surface. Even small amounts of residue cover can cause substantial reductions in erosion. The amount of erosion protection provided with a given percent of surface cover is influenced by the type of residue material.

Conservation tillage has been defined as any tillage or planting system that leaves at least 30% of the soil surface covered with residue after planting (Fig. 27). When tillage is performed, implements are used that cause only minimal disturbance to the soil surface, thus maintaining existing crop residues. For some row crops, such as soybeans, no tillage is used before planting to leave sufficient residue cover to control erosion.

Terraces. Terraces are broad channels built perpendicular to the slope of steep land. The gentle grades used in terraces allow runoff to be carried around a hill at relatively low velocities, causing sediment to settle from the runoff water. Terraces usually empty onto grassed waterways or into underground pipes, thus preventing the formation of gullies.

Crops are usually planted parallel to the terrace channel, requiring the use of contour farming. Conservation tillage is also frequently used in conjunction with terracing. A significant investment is required to construct terraces, and farming operations are more difficult on terraced hillslopes. As a result, terraces are used only when other control measures cannot provide adequate erosion protection.

Buffer strips. Buffer strips are areas of land maintained with permanent vegetation, designed to intercept runoff. They are most effective when used in combination with other erosion control practices. Buffer strips are located at various positions along the landscape as part of a planned conservation system. The types of vegetation used in buffer strips is influenced by local conditions. Periodic maintenance may be required for sustained buffer strip performance. The types of buffer strips frequently used are contour buffer strips, filter strips, and grassed waterways.

Contour buffer strips containing perennial grasses are planted along steep slopes. The grass strips

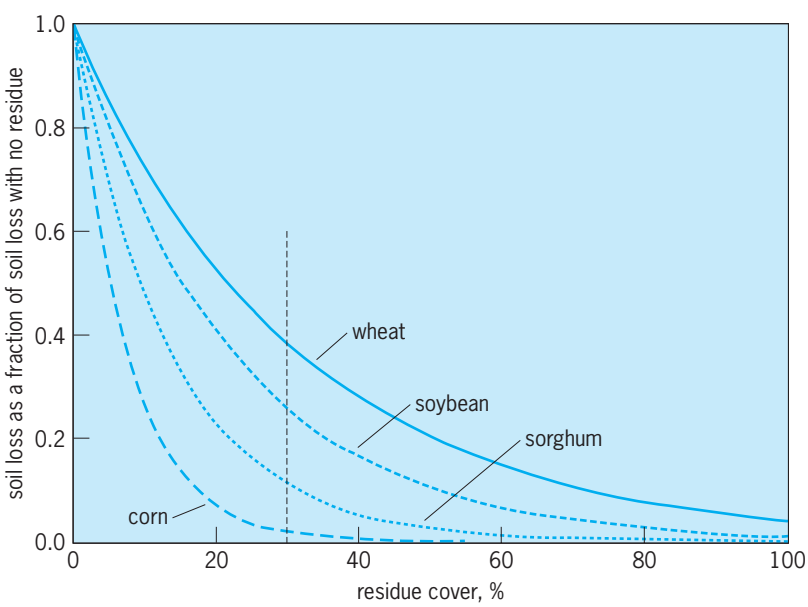

Fig. 27. Ratio of soil loss for given residue covers to soil loss with no cover. The vertical broken line indicates the residue cover necessary to be defined as conservation tillage. (*After T. S. Colvin and J. E. Gilley, Crop residue: Soil erosion combatant, Crops and Soils, 39(7):7–9, 1987*)

remove sediment from overland flow. The species of perennial grasses used and the spacing of the grass strips are tailored to local conditions. A narrow terrace may eventually form along the grass strip as a result of sedimentation. The expense of establishing contour buffer strips is substantially less than the cost for constructing terraces.

Filter strips can also be used to remove sediment from overland flow and provide increased infiltration. They are usually located on the edge of fields or adjacent to streams, ponds, or wetlands, and thus do not interfere significantly with normal farming operations. Filter strips are best suited for areas with gentle slopes where rilling is not a problem.

Grassed waterways can be used to convey runoff from terraces or other concentrated flow areas, and prevent channel erosion and gully formation. Sediment transported by overland flow is deposited in grassed waterways, thus reducing costly downstream sedimentation. A stable outlet below the grassed waterway is provided to reduce runoff velocity and disperse the flow before it enters a vegetated filter.

Stream channel erosion control measures. Vegetative, mechanical, or combined vegetative-mechanical means have been developed to reduce stream channel erosion. Depending on the size of the upstream drainage area, grading of the stream bank to a less severe slope may be necessary. Grass, shrubs, and trees have been successfully used to stabilize stream channels.

Dikes made of loose stone or rock piles are placed within the stream channel to divert the faster-flowing water away from the bank. Mechanical covers of stones, rocks, or other protective material may also be placed over the erodible bank. Areas with the greatest erosion hazard, such as the bottom of a stream bank, may be protected with a mechanical cover, while the upper portion of the stream banks is usually stabilized with vegetation. John E. Gilley

Wind erosion. Soil erosion by wind is a dynamic process. The results of wind erosion are evident when soil particles are dislodged from the soil surface, injected into the wind stream, and in some cases transported around the world as "dust" before being deposited. In the process of being dislodged by wind, soil particles are sorted according to size, like the winnowing of grain. Wind erosion is the geomorphological process responsible for the tremendous loess (wind-borne) deposits of highly fertile soil around the world. While wind erosion has always been an active process, human activities tend to accelerate it. As humans disturb large areas, loose soil particles may result. Large, coarse particles may move a few feet or be deposited in fence rows or road ditches at the edge of a field. When carried by strong winds, these coarse particles damage plants, abrade paint, break down soil crusts, and accelerate the wind erosion process. The loss of fine soil particles from a field reduces the capacity of the soil to produce crops; moreover, these particles reduce visibility and degrade air quality. *See* DUST STORM; LOESS.

When susceptible soils are exposed to erosive winds, ominous dust clouds occur. To effectively control the wind erosion that produces these dust clouds, all available resources, including the climate, soil, crop, and management systems, must be utilized. No single erosion control system will be equally effective for all wind erosion problem areas. For example, in regions that normally produce high-residue (vegetation) crops, the most effective wind erosion control systems maintain those residues either erect or on the soil surface. For semiarid regions growing crops that produce little residue, the most effective wind erosion control systems include a rough and cloddy soil surface and vegetative wind barriers. Options for humid or subhumid regions should include cover crops and residue management.

For all regions, if the erosive winds blow from the same direction, the benefits of residue management, soil roughness, or cover crops could be supplemented with strip cropping, annual or perennial crop wind barriers, and tree shelter belts. Farmers should be aware of all the wind erosion control practices that are available and then use whatever combination will be most effective for their conditions.

It is possible to estimate soil erosion with computer modes. Wind erosion models utilize weather, soil, crop, and management information in the calculation of predicted soil erosion losses. The losses can be estimated daily or for the entire crop-growing period.

Each erosion control strategy must consider the potential for rainwater to degrade soil surface roughness, temperature and rainfall requirements for growing cover crops, and the rate at which crop residues decompose.

Wind erosion damage can never be completely eliminated, but with careful planning and wise use of available resources (climate, soil, crop, and management) the impact of wind erosion on plants, soils, the atmosphere, and humans can be minimized. *See* EROSION; SOIL CONSERVATION. Donald Fryrear

Bibliography. P. W. Birkeland, *Soils and Geomorphology*, 1984; A. Blum, *Plant Breeding for Stress Environments*, 1988; N. C. Brady and R. R. Weil, *The Nature and Properties of Soils*, 11th ed., 1995; E. Bresler, B. L. McNeal, and D. L. Carter, *Saline and Sodic Soils*, 1982; H. D. Foth, *Fundamentals of Soil Science*, 8th ed., 1990; H. D. Foth, *Soil Fertility*, 1988; W. H. Gardner and J. C. Hsieh, *Water Movement in Soil*, 1960 (27-min time-lapsed video cassette, available in Soil Conservation Service and university audio-visual libraries, from Crop and Soil Sciences Club, Washington State University, P.O. Box 646420, Pullman, WA 99164-6420); C. T. Haan et al., *Design Hydrology and Sedimentology for Small Catchments*, Academic Press, San Diego, 1994; C. T. Haan, H. P. Johnson, and D. L. Brakensiek (eds.), *Hydrologic Modeling of Small Watersheds*, Amer. Soc. Agr. Eng. Monogr. 5, 1982; D. Hillel, *Environmental Soil Physics*, Academic Press, San Diego, 1998; N. Hudson, *Soil Conservation*, 3d ed., Iowa State University Press, Ames, 1995; E. H. Isaaks, *An Introduction to Applied Geostatistics*, 1990; H. Jenny,

The Soil Resource: Origin and Behavior, Ecological Studies, vol. 37, 1980; W. A. Jury, W. R. Gardner, and W. H. Gardner, *Soil Physics*, 5th ed., 1991; A. Klute (ed.), *Methods of Soil Analysis*, pt. 1: *Physical and Mineralogical Methods*, 2d ed., 1986; R. Lal, *Soil Erosion and Its Control*, 1988; A. Lieth and A. M. Massoun, *Towards Rational Use of High Salinity Tolerant Plants: Agriculture and Forestry under Marginal Soil Water Conditions*, 1993; L. Lyles, L. J. Hagan, and E. L. Skidmore, *Soil Conservation: Principles of Erosion by Wind*, Amer. Soc. Agron. Monogr. 23, 1983; National Academy of Sciences, *Soils of the Humid Tropics*, 1972; National Association of Review Appraisers Staff, *Land Classification for Land Uses Management and Valuation*, 1983; M. A. Nearing et al., A process-based soil erosion model for USDA: Water erosion prediction project technology, *Trans. Amer. Soc. Agr. Eng.*, 32(5):587–1593, 1989; F. J. Pierce and W. W. Frye (eds.), *Advances in Soil and Water Conservation*, Sleeping Bear Press, Chelsea, MI, 1998; K. G. Renard and G. R. Foster, *Soil Conservation: Principles of Erosion by Water*, Amer. Soc. Agron. Monogr. 23, 1983; G. Richardson and P. Mueller-Beilschmidt, *Winning with Water: Soil Moisture Monitoring for Efficient Irrigation*, 1988; G. O. Schwab et al., *Soil and Water Conservation Engineering*, 4th ed., John Wiley, Singapore, 1993; I. Shainberg and J. Shalhevet, *Soil Salinity under Irrigation: Processes and Management*, 1984; M. A. Sprague and G. B. Triplett (eds.), *No-Tillage and Surface Tillage Agriculture: The Tillage Revolution*, 1986; S. J. Stavarek and D. W. Rains, The development of tolerance to mineral stress, *HortScience*, 19:377–382, 1984; K. K. Tanji (ed.), *Agricultural Salinity Assessment and Management*, 1990; F. R. Troeh et al., *Soil and Water Conservation: Productivity and Environmental Protection*, 3d ed., Prentice Hall, Upper Saddle River, NJ, 1999; U.S. Department of Agriculture, *Soil: Yearbook of Agriculture*, 1957; U.S. Department of Agriculture Soil Conservation Service, *Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys*, 1988; W. H. Wischmeier and D. D. Smith, *Predicting Rainfall Erosion Losses: A Guide to Conservation Planning*, USDA Agr. Handb. 537, 1978.

Soil chemistry

The study of the composition and chemical properties of soil. Soil chemistry involves the detailed investigation of the nature of the solid matter from which soil is constituted and of the chemical processes that occur as a result of the action of hydrological, geological, and biological agents on the solid matter. Because of the broad diversity among soil components and the complexity of soil chemical processes, the application of concepts and methods employed in the chemistry of aqueous solutions, of amorphous and crystalline solids, and of solid surfaces is required. For a general discussion of the origin and classification of soils. *See* SOIL.

Elemental composition. The elemental composition of soil varies over a wide range, permitting only a few general statements to be made. Those soils that contain less than 12–20% organic carbon are termed mineral. (The exact percentage to consider in a specific case depends on drainage characteristics and clay content of the soil.) All other soils are termed organic. Carbon, oxygen, hydrogen, nitrogen, phosphorus, and sulfur are the most important constituents of organic soils and of soil organic matter in general. Carbon, oxygen, and hydrogen are most abundant; the content of nitrogen is often about one-tenth that of carbon, while the content of phosphorus or sulfur is usually less than one-fifth that of nitrogen. The number of organic compounds into which these elements are incorporated in soil is very large, and the elucidation of the chemistry of soil organic matter remains a challenging problem. *See* HUMUS.

Besides oxygen, the most abundant elements found in mineral soils are silicon, aluminum, and iron (**Table 1**). The distribution of chemical elements will vary considerably from soil to soil and, in general, will be different in a specific soil from the distribution of elements in the crustal rocks of the Earth. Often this difference may be understood in terms of pedogenic weathering processes and the chemical reactions that accompany them. Some examples are the accumulation of aluminum and iron in the Oxisols and of calcium in the Mollisols. The most important micro or trace elements in soil are boron, copper, manganese, molybdenum, and zinc, since these elements are essential in the nutrition of green plants. Also important are cobalt, which is essential in animal nutrition, and selenium, cadmium, and lead, which may accumulate to toxic levels in soil. The average natural distribution of trace elements in soil is not greatly different from that in crustal rocks (**Fig. 1**). This indicates that the total content of a trace element in soil usually reflects the content of that element in the soil parent material and, generally, that the trace element content of soil often is not affected directly by pedochemical processes.

The elemental composition of soil varies with depth below the land surface because of percolating water, chemical processes, and biological activity. The principal chemical processes are (1) hydrolysis (reaction with water to form products containing hydroxide ions), complexation (reaction of a metal with a ligand to form a product containing both), oxidation-reduction (changes in the oxidation state of an element), ion exchange (replacement of one ion by another on a solid surface), hydration (reaction with water to form a product containing water molecules), and flocculation-dispersion (settling-resuspension of solid particles in water). This last process affects soil particle removal by erosion or by translocation downward along soil pores. The principal effect of these processes is the appearance of illuvial horizons in which compounds such as aluminum and iron hydrous oxides, aluminosilicates, or calcium carbonate have been precipitated from

TABLE 1. Average percentages of the major elements and some microelements in subsurface soil clays and crustal rocks

	Soil order						
Element	Alfisol	Inceptisol	Mollisol	Oxisol	Spodosol	Ultisol	Crustal rocks
Silicon	19.20	24.69	23.01	12.43	5.79	16.02	27.72
Aluminum	12.38	19.61	10.29	19.33	15.86	17.49	8.20
Iron	8.04	3.81	6.83	10.83	3.29	11.96	4.10
Calcium	0.69	0.00	3.59	0.10	0.29	0.15	4.10
Magnesium	1.26	0.40	1.62	0.46	0.15	0.08	2.30
Sodium	0.18	2.52	0.04	0.00	0.27	0.06	2.30
Potassium	3.63	n.d.*	1.20	0.07	0.40	0.22	2.10
Titanium	0.40	0.28	0.44	1.32	0.16	0.50	0.56
Manganese	0.06	n.d.	0.06	0.08	0.06	0.05	0.10
Phosphorus	0.14	n.d.	0.14	0.27	0.17	0.12	0.10

*n.d. = no data.

solution or deposited from suspension. *See* WEATHERING PROCESSES.

Minerals. The minerals in soils are the products of physical, geochemical, and pedochemical weathering. Soil minerals may be either amorphous or crystalline. They may be classified further, approximately, as primary or secondary minerals, depending on whether they are inherited from parent rock or are produced by chemical weathering, respectively.

Primary minerals in soil. The bulk of the primary minerals that occur in soil are found in the silicate minerals, such as the olivines, garnets, pyroxenes, amphiboles, micas, feldspars, and quartz. The feldspars, micas, amphiboles, and pyroxenes commonly are hosts for trace elements that may be released slowly into the soil solution as weathering of these minerals continues. Chemical weathering of the silicate minerals is responsible for producing the most important secondary minerals in soil. *See* SILICATE MINERALS.

Secondary minerals in soil. The important secondary minerals that occur in soil are found in the clay fraction, sometimes in the form of coatings on other minerals. These include aluminum and iron hydrous oxides, carbonates, and aluminosilicates. The term allophane is applied to x-ray amorphous, hydrous aluminosilicates that are characterized by variable composition and a defect-riddled kaolinite structure containing aluminum in both tetrahedral and octahedral coordination. The significant crystalline aluminosilicates possess a layer structure; they are chlorite, halloysite, kaolinite, illite, smectite, and vermiculite. These clay minerals are identified in soil by means of the characteristic x-ray diffraction patterns they produce after certain pretreatments, although their positive identification may be difficult if two or more of the minerals are present at once. *See* CLAY MINERALS; X-RAY DIFFRACTION.

The distribution of secondary minerals varies among different soils and changes with depth below the surface of a given soil. However, under a leaching, well-oxidized environment, soil minerals do possess a differential susceptibility to decomposition, transformation, and disappearance from a soil profile. This has made possible the arrangement of the clay-sized soil minerals in the order of increasing resistance to chemical weathering known as the Jackson-Sherman weathering stages (**Table 2**). Minerals ranked near the top of the Jackson-Sherman list are present, therefore, in the clay fractions of slightly weathered soils (Entisols, Inceptisols, Aridisols, Gelisols), whereas minerals near the bottom of the list predominate in extensively weathered soils (Ultisols, Oxisols).

In zonal soils of humid-cool to subhumid-temperate regions, illite is the predominant clay mineral. Mixtures of kaolinite, vermiculite, and interstratified clay minerals are found in humid-temperate regions. In humid-warm regions, kaolinite, halloysite, allophane, gibbsite, and goethite are found. The mineralogical composition of the highly weathered and leached soils of the humid tropics is a subject of

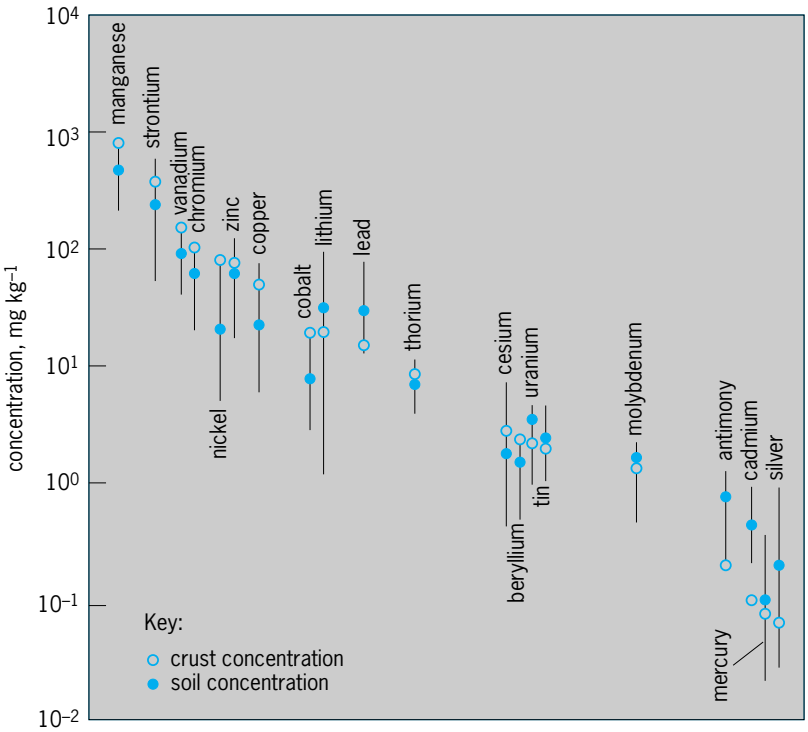

Fig. 1. Average concentrations and ranges of concentration of trace metals in surface soils compared with average crustal concentrations. (*After G. Sposito, Distribution of potentially hazardous trace metals, Metal Ions Biol. Sys., 20:1–20, 1986*)

TABLE 2. Jackson-Sherman weathering stages*

Characteristic minerals in soil clay fraction	Characteristic soil chemical and physical conditions
	Early Stage
Gypsum	Very low content of water and organic matter, very limited leaching
Carbonates	Reducing environments
Olivine/pyroxene/amphibole	Limited amount of time for weathering
Iron (II)-bearing micas	
Feldspars	
	Intermediate Stage
Quartz	Retention of sodium, potassium, calcium, magnesium, iron(II), and silica
Illite	Ineffective leaching and alkalinity
Vermiculite/chlorite	Igneous rock rich in calcium, magnesium, and iron(II), but no iron(II) oxides
Smectites	Silicates easily hydrolyzed
	Flocculation of silica
	Transport of silica into the weathering zone
	Advanced Stage
Kaolinite	Removal of sodium, potassium, calcium, and silica
Gibbsite	Effective leaching
Iron oxides (goethite, hematite)	Fresh-water oxidation of iron(II)
	Acidic compounds, low pH
Titanium oxides (anatase, rutile, ilmenite)	Dispersion of silica
	Aluminum-hydroxy polymers

*After M. L. Jackson and G. D. Sherman, Chemical weathering of minerals in soils, *Adv. Agron.,* 5:291–381, 1953, and M. L. Jackson, Clay transformations in soil genesis during the Quaternary, *Soil Sci.,* 99:15–22, 1965.

active investigation, in part because these soils (the Oxisols and Ultisols) are rapidly being converted to agricultural use by deforestation, with serious global change consequences. The soil minerals are dominated by iron and aluminum hydrous oxides, kaolinite, halloysite, and quartz. Weathering residues also are found in thin coatings on clay particle surfaces. Vermiculite and smectite with interlayer aluminum hydroxy polymers are common.

The chemical conditions favoring the genesis of kaolinite are the removal of the basic cations and Fe^{2+} by leaching, the addition of hydrogen ion (H^+) in fresh water, and a high aluminum-silicon (Al-Si) molar ratio. Smectite is favored by the retention of basic cations (arid conditions or poor drainage) and of silica. *See* GOETHITE; HALLOYSITE; ILLITE; KAOLINITE; MONTMORILLONITE; VERMICULITE.

Cation exchange. A portion of the chemical elements in soil is in the form of cations that are not components of inorganic salts but that can be replaced reversibly by the cations of leaching salt solutions. These cations are said to be readily exchangeable, and their total quantity, usually expressed in units of centimoles of proton charge [cmol(p^+)] per kilogram of dry soil, is termed the cation-exchange capacity (CEC) of the soil. The cation exchange capacity ordinarily is measured by leaching a known amount of soil with a salt solution, such as sodium acetate at pH 8.2, followed by an additional leaching with alcohol to remove residual salt, then determining the quantity of replacing cation [such as sodium (Na^+)] in the soil. However, the quantity of cations remaining after such treatment does not have a unique value characteristic of the soil alone, but depends as well on the concentration, the ionic composition, and the pH of the leaching solution. The cation exchange capacity of a soil generally will vary directly with the amounts of clay and organic matter present and with the distribution of clay minerals. *See* PH.

Soils that are less weathered because of recent origin, low precipitation, or temperate to cold climate have as exchangeable cations largely calcium (Ca^{2+}) and magnesium (Mg^{2+}). Some soils of dry areas contain significant amounts of exchangeable sodium. Extensively weathered soils, unless formed from basic parent material, have 20–95% of their exchangeable cations as aluminum. Prolonged leaching with fresh water supplies hydrogen (H^+) ions that eventually penetrate and disrupt the structures of soil aluminosilicates, thereby releasing aluminum cations, some of which remain in exchangeable form (**Table 3**).

The usual meaning of the term ion-exchange reaction in soil chemistry is the replacement of one adsorbed, readily exchangeable ion by another. On the molecular level, this means that ion exchange is a surface phenomenon involving charged species in outer-sphere complexes or in the diffuse-ion swarm. In practice, this conceptualization is adhered to only approximately. Cation-exchange reaction on soil humus, for example, include protons that may be adsorbed in inner-sphere surface complexes. Common extracting solutions for soil cation-exchange capacity measurements may displace metal cations from inner-sphere surface complexes as well as readily exchangeable metal cations. Careful experimental methods that will quantitate only readily

TABLE 3. Typical distributions of exchangeable cations in some soil orders*

Soil order (pH value)	CEC, cmol(p^+) · kg^{-1}	Percentages				
		Calcium	Magnesium	Sodium	Potassium	Hydrogen + aluminum
Alfisol (5.4)	38	23	11	0	2	64
Aridisol (6.3)	4	52	26	5	17	0
Entisol (6.4)	5	37	23	<1	6	33
Inceptisol (5.3)	104	6	1	<1	<1	91
Mollisol (6.7)	25	67	12	<1	2	18
Spodosol (3.6)	106	5	6	<1	1	87
Oxisol (4.9)	27	31	8	<1	2	58
Ultisol (3.5)	24	11	3	<1	<1	84

*After R. L. Donahue, R. W. Miller, and J. C. Shickluna, *Soils,* Prentice-Hall, 1983.

exchangeable ions or will partition adsorbed ions accurately into readily exchangeable and specifically adsorbed species remain the focus of research. *See* CHEMICAL EQUILIBRIUM; ION EXCHANGE.

Isomorphic substitution in clay minerals. One of the most important sources of cation exchange capacity in soils is the negative charge that occurs on 2:1 clay mineral surfaces because of isomorphic substitution. The replacement of the silicon cation (Si^{4+}) in the tetrahedral sheet by aluminum cation (Al^{3+}) or of Al^{3+} in the octahedral sheet by magnesium cation (Mg^{2+}) and iron(II) cation (Fe^{2+}) results in a permanent, negative surface charge on the clay mineral surface that is more or less localized in hexagonal cavities formed by rings of oxygen atoms. This negative charge is balanced either by the formation of surface complexes between the cavities and cations taken from the soil solution or by swarms of cations remaining free in the soil solution. Extensive isomorphic substitution occurs in the tetrahedral sheet of vermiculite, with the result that potassium cation (K^+) can form inner-sphere surface complexes stabilized by the fact that this cation does not hydrate easily and fits almost perfectly into the hexagonal cavity in the clay mineral surface. On the other hand, isomorphic substitution occurs principally in the octahedral sheet of the smectite, montmorillonite, and the resultant electrostatic force binding a cation is weaker than for vermiculite because it must act through a greater distance. Cations such as calcium (Ca^{2+}) which hydrate strongly tend to form outer-sphere surface complexes with the hexagonal cavities in a 2:1 clay mineral surface.

Ionizable hydroxyl groups on clay minerals. At the edges of the structural layers in the crystalline clay minerals and on the surfaces of the amorphous clay minerals, hydroxyl (OH) groups may be found bonded to exposed Al^{3+} cations. These hydroxyl groups can act as Brönsted acids, ionizing in aqueous solution when the pH is sufficiently high. Thus a pH-dependent, negative surface charge can develop that will contribute to the cation exchange capacity. The mineral surfaces for which this pH dependence occurs can form surface complexes with cations. When this occurs, the complexed cation (for example, H^+, Na^+, or Ca^{2+}) is termed a potential-determining ion for the clay mineral surface, and its degree of surface complexation is governed by its electrochemical potential. The importance of pH-dependent charge versus permanent charge is greater in 1:1 layer silicates than in 2:1 layer silicates. For montmorillonite the pH-dependent cation exchange capacity may be 20% of the total cation exchange capacity; for illite it may be 40%; while for kaolinite and allophane it is essentially 100%.

Ionizable functional groups in organic matter. The organic matter in soil contains two types of functional groups that contribute importantly to the cation exchange capacity: aromatic and aliphatic carboxyls and phenolic hydroxyls. When the pH is greater than 4, these acidic groups ionize in aqueous solution and provide a source of negative charge for metal cation adsorption. The cations of metals in group 1 and 2 (for example, Na^+ and Ca^{2+}) of the periodic table are readily exchangeable after adsorption by soil organic matter, whereas those of the transition metals [for example, iron(III) (Fe^{3+}) and copper (Cu^{2+})] often are not. The contribution to the cation exchange capacity from organic functional groups can be large at high pH values.

Anion exchange. The stoichiometric exchange of the anions in soil for those in a leaching salt solution is a phenomenon limited to chloride and nitrate in the general scheme of anion reactions with soils. Under acid conditions (pH < 5) exposed hydroxyl groups at the edges of the structural sheets or on the surfaces of clay-sized particles become protonated and thereby acquire a positive charge. The degree of protonation is a sensitive function of pH, the ionic strength of the leaching solution, and the nature of the clay-sized particle. The magnitude of the anion-exchange capacity (AEC) usually varies from near 0 at pH 9 for any soil colloid to as much as 50 cmol(−)/kg of allophanic clay at about pH 4. [cmol(−)/kg means centimoles of negative charge per kilogram.] Smectite and other clay minerals with high, pH-invariant cation exchange capacity values do not usually adsorb exchangeable anions to any degree. Anion-exchange capacity may be measured conveniently by shaking a sample of soil for 1 h in sodium chloride (NaCl) solution at a chosen pH, filtering, and displacing the adsorbed chloride ion (Cl^-). Anion-exchange

capacity is then the number of cmoles of Cl displaced per kilogram of soil.

Negative anion adsorption. Soils whose cation exchange capacity is large often display a significant negative adsorption of anions: the concentration of anions in a solution contacting a soil suspension is larger than that of the anions in the liquid phase of the suspension. This phenomenon may be understood simply on the basis of the presence of negative charge on the surface of the solid-phase particles in the suspension. This surface charge attracts cations and repels anions. The principal effects of this repulsion are to reduce the anion exchange capacity and to increase the ease with which anions may be leached from a soil. Negative cation adsorption occurs in soils whose anion-exchange capacity is significant.

Specific anion adsorption and precipitation. Anion exchange in the classic sense applies primarily to halide and nitrate ions. For other ions, in particular, borate, molybdate, selenite, and orthophosphate, the reaction with the solid matter in soil involves specific adsorption or precipitation. The selenite ion (SeO_3^{2-}), for example, can react with the accessible aluminum hydroxy ions of clay minerals and with hydrous aluminum and iron oxides in soil to form inner-sphere surface complexes by ligand exchange with OH groups. The orthophosphate ion can react the same way or form x-ray amorphous analogs of the known crystalline aluminum and iron phosphate minerals by precipitation. The nature of the reaction is strongly dependent on pH, the metal cations in solution, the acidity of the added phosphate compound, and the structure of the solid phase with which the phosphate ion reacts. Under conditions of a relatively high pH, low acidity of the added phosphate, or high degree of crystallinity of the solid phase reactant, phosphate ion (PO_4^{3-}) will tend to be "specifically adsorbed" by Al or Fe ions at the surface of the solid through ligand exchange. If these conditions are reversed, the combination of PO_4^{3-} with the solid phase may result in a nearly complete destruction of the reactant solid and the formation of an amorphous aluminum or iron phosphate. This is an area of active research in soil chemistry.

Soil solution. The solution in the pore space of soil acquires its chemical properties through time-varying inputs and outputs of matter and energy that are mediated by the several parts of the hydrologic cycle and by processes originating in the biosphere (**Fig. 2**). The soil solution thus is a dynamic and open natural water system whose composition reflects the many reactions that can occur simultaneously between an aqueous solution and an assembly of mineral and organic solid phases that varies with both time and space. This type of complexity is not matched normally in any chemical laboratory experiment, but nonetheless must be amenable to analysis in terms of chemical principles. An understanding of the soil solution in terms of chemical properties has proven to be essential to progress in the maintenance of soil fertility and the quality of runoff and drainage waters.

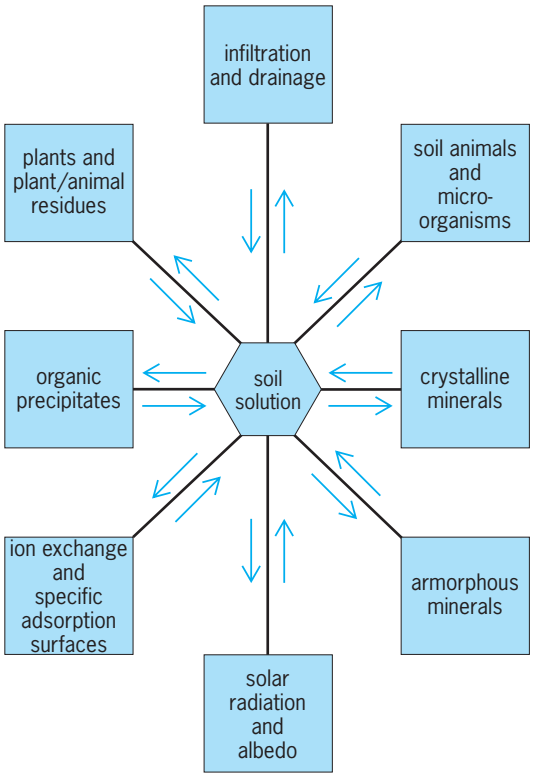

Fig. 2. Factors influencing the chemistry of a soil solution. (*After J. F. Hodson, Chemistry of the micronutrients in soils, Adv. Agron., 15:119–159, 1963*)

Chemical speciation of macrosolutes. The macrosolute composition of a soil solution will vary depending on pH, pE (negative common logarithm of the electron activity), organic matter content, input of chemical elements from the biosphere (including humans), and effectiveness of leaching. Under conditions of near-neutral pH, high pE, low organic matter content, no solute input from agriculture, and good but not excessive drainage, the expected macrosolutes are Ca^{2+}, K^+, Mg^{2+}, Na^+, Cl^-, biocarbonate ion (HCO_3^-), silicic acid [$Si(OH)_4$], and sulfate ion (SO_4^{2-}). If the pH is low, H^+ and Al^{3+} should be added to this list; if it is high, carbonate ion (CO_3^{2-}) should be added. If the soil has been fertilized, nitrate ion (NO_3^-) and hydrogen phosphate ion ($H_2PO_4^-$) become important. If the drainage is excessive, Al may be abundant and one or more of the solutes in the original list may be insignificant. If the drainage is poor and, therefore, the pE is low, SO_4^{2-} will be replaced by S^{2-} and CO_3^{2-} should be added. If the organic matter content is high, organic solutes become important. Combinations of these different environmental conditions will change the original list of solutes in still other ways (for example, low pE and nitrogen fertilizer addition could add NH_4^+ to the list).

The chemical speciation of macrosolutes (that is, their distribution among the free-ionic, complexed, precipitated, and adsorbed forms) depends on the nature of the solid matter in the soil, the composition of metals and ligands in solution and their concentration, and the pE value. Clearly the macrosolutes

themselves are interdependent in determining their speciation, and even the three factors just mentioned cannot be considered in complete isolation from one another. Nevertheless, it is possible to make some very broad statements about the macrochemical species to be expected by employing the principle of hard and soft acids and bases. *See* ACID AND BASE.

The macrosolute metal cations are hard acids. This means that they generally tend to form chemical bonds of a simple electrostatic type and, therefore, that their reactivities with any ligand (including those on a mineral surface) should be predictable on the basis of ionic charge and radius along with their common property of low polarizability (polarizability refers to the tendency of an ion to distort its electron cloud in the presence of an electric field). In a soil solution, the macrosolute metal cations will tend to (1) form complexes and sparingly soluble precipitates primarily with oxygen-containing ligands (hard bases), such as H_2O, OH^-, HCO_3^-, CO_3^{2-}, $H_2PO_4^-$, PO_4^{3-}, SO_4^{2-} and nonhydroxylated mineral surfaces: the stabilities of these complexes will tend to increase with the ratio of ionic charge to ionic radius of the metal; usually no reaction will occur with soft or nearly soft bases such as S^{2-} and Cl^-; (2) form complexes with carboxyl groups, but not with organic ligands containing only nitrogen and sulfur electron donors.

These generalizations make it possible to enumerate the probable complexes and precipitates of the macrosolute metals in a soil solution of known composition (including adsorbing surfaces). The exact speciation of the metals then can be calculated if stability constants are available for the expected chemical reactions. Usually a large set of nonlinear algebraic equations must be solved and ionic strength corrections must be performed.

Chemical speciation of microsolutes. The important microsolutes in soil include trace metals, such as iron, copper, and zinc, and trace element oxyanions, such as those formed by arsenic, boron, molybdenum, and selenium. The tableau of microsolutes in a given soil solution is more dependent on inputs from the lithosphere and the biosphere (particularly humans) and less on proton or electron activity and hydrologic factors than is the composition of macrosolutes. The trace metals present, for example, usually are derived from the chemical weathering of specific parent rocks, from the application of fertilizers, pesticides, and urban wastes, and from air pollution.

The most general features of the chemical speciation of the microsolutes also may be predicted on the basis of the principle of hard and soft acids and bases. The trace metal cations are soft or nearly soft acids, and the trace element anions are hard or nearly hard bases. The exception to this statement is Fe^{3+}, which is a hard acid. This means that the trace metal cations, except for Fe^{3+}, will tend to form chemical bonds of a covalent type whose strength will depend much more on detailed electronic structural considerations than on cationic size or charge. For the anions, the implication is that they will combine strongly with hard-acid metal cations, just as do other oxygen-containing ligands. Trace metal speciation thus presents a more complicated problem than does that of the macrosolute metals. Generally, the trace metal cations in a soil solution will (1) form complexes and insoluble precipitates more readily with the inorganic ligands Cl^- and S^{2-} and with sites on mineral surfaces that can bind covalently, than with oxyanions; stronger complexes also will form more readily with organic functional groups containing S, P, and N donors than with carboxyl groups; (2) tend to follow the Irving-Williams order in regard to the stabilities of strong complexes: $Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+} > Cd^{2+}$.

These broad predictions imply that trace-metal speciation in soil solutions will depend sensitively on the content and type of organic matter, the percentage of kaolinitic and amorphous hydrous oxide minerals, the pH, the pE, and the ionic strength. For example, a low solubility of the micronutrients Cu and Zn should occur for soils high in immobile organic matter (but not too low in pH) or for soil solutions with low pE values. Moreover, the solubilities of the trace metals should increase significantly with an increase in chloride concentration or in organic solutes. Since the complexes formed in these cases would reduce the ionic charge on the soluble trace metal species, a decrease in the affinity of a negative charge adsorbing surface (for example, that of montmorillonite) should also occur. Research has indicated that complex formation also decreases the bioavailability of a trace metal to a plant grown in soil. The uptake of a trace metal by a plant appears to be strongly correlated with the thermodynamic activity of the free metal cation species (**Fig. 3**).

Clay-organic complexes. The clay minerals in soils often are observed to be intimately associated with carbonaceous materials. These materials may be residues from plant or animal decomposition, herbicides or other pesticides, organic polymers and polyelectrolytes, surfactant compounds, or microbial metabolites. The complexes which they form

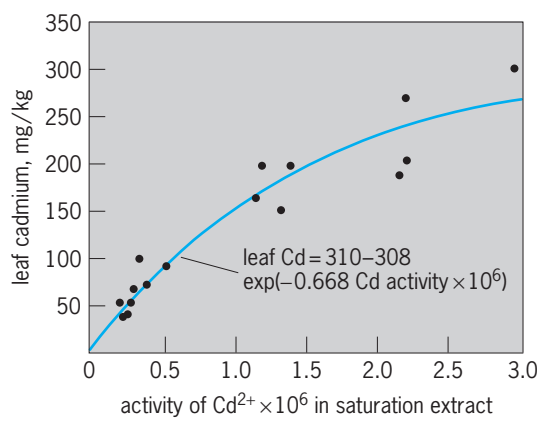

Fig. 3. Relationship between leaf concentration of a trace metal and its thermodynamic activity in the soil solution. The example shows Cd^{2+} uptake by Swiss chard, with the curve through the data represented by the equation. (*After F. T. Bingham, J. E. Strong, and G. Sposito, Influence of chloride salinity on cadmium uptake by Swiss chard, Soil Sci., 136:160–165, 1983*)

with clay minerals bear importantly on soil fertility [particularly in the rhizosphere (the near-root zone)], soil structure, soil moisture and aeration characteristics, the biological activity of organic compounds applied or disposed on soil, and the degradation of solid and liquid organic wastes in the soil environment. Generally, the organic component of a naturally occurring clay-organic complex will be of a very complicated nature that defies conclusive structural determination. Therefore, in order to obtain fundamental information about the mechanisms of bonding between clay minerals and organic matter in soil, a major line of research has involved the study of the reactions of known organic compounds with single types of clay minerals. On the basis of these studies, some important bonding mechanisms have been identified. They are expected to apply to the associations between clay minerals and organic matter in nature whenever the appropriate mineral species and organic functional groups are present.

Bonding mechanisms. The principal mechanisms through which organic compounds may bind to clay minerals have been elucidated largely by spectroscopic studies. They may be classified as follows:

1. Organic cation adsorption can occur, through protonated amine or carbonyl groups, onto any constant-charge clay mineral surface. The protonation of the functional groups may be either a pH effect or an acceptance of a proton that was formerly occupying an exchange site, was associated with a water molecule hydrating a metal cation, or was bound on another adsorbed organic cation. The affinity of an organic cation for a permanent-charge surface depends on molecular size, the nature of the functional groups present, and molecular configuration. Steric effects can be particularly significant because of the localized character of exchange sites and the quasirigid hydration envelope built up on a clay mineral surface. The stability of the water structure on a smectite surface is, in fact, great enough to require interstratified layers of either adsorbed metal cations or adsorbed organic cations when the clay mineral surfaces are only partially saturated with the organic compound. A mixture of the two types of cation in a single interlayer region disrupts the water structure too much to be stable.

2. Polar organic functional groups can bind to adsorbed cations through simple ion-dipole forces or complex formation involving covalent bonds (inner-sphere surface complex). The ion-dipole mechanism is to be expected, of course, for hard-acid metal cations, such as Ca^{2+} and Al^{3+}, whereas the formation of covalent bonds is to be expected for soft-acid metal cations, such as Cu^{2+}. As the organic functional groups often have soft-base character, the strength of binding by this mechanism should be greatest for exchangeable transition-metal cations. A sharp exception to this rule could occur with exposed Al^{3+} (or Fe^{3+}) in amorphous aluminosilicates that bind organic matter containing large numbers of carboxyl groups.

3. Large organic molecules can bind effectively to a clay mineral surface through hydrogen bonding. This bonding can involve a water bridge from a hydrated exchangeable cation to an oxygen-containing functional group (outer-sphere surface complex), a hydrogen bond from a more acidic functional group adsorbed directly on the clay mineral surface to a less acidic free one containing oxygen, or a direct hydrogen bond to a surface oxygen or hydroxyl plane in the clay mineral. If the exchangeable metal cation is a hard acid, the first type of bond will by far dominate the third type in importance. Direct hydrogen bonds to a plane of surface atoms would be accompanied by weaker dipole-dipole (that is, van der Waals) interactions, in general. This type of binding should be most important when very large organic molecules associate with a clay mineral surface containing relatively few exchange sites.

Catalysis reactions. Clay minerals have been shown often to catalyze reactions involving organic compounds. This catalytic function appears to be connected intimately with the presence of exchangeable metal cations and may be separated into two distinct types. The first type relates to the fact that the water molecules hydrating the exchangeable cations tend to dissociate protons very readily and, therefore, to endow the clay mineral surface with a pronounced acidity that increases markedly with desiccation. This enhanced proton-donating capability of the clay mineral, which will be greater the harder an acid the exchangeable cation is, serves a catalytic function in, for example, the surface protonation of amines and amino acids or in the abiotic degradation of pesticides.

A second type of catalytic function derives from the formation of inner-sphere organic complexes with exchangeable cations. This mechanism, which should be more significant the softer an acid the adsorbed metal cation is (again excepting Al^{3+} and Fe^{3+}), appears to play a basic role in, for example, the stabilization of humic compounds against degradation. These and other reactions catalyzed by clay minerals may prove to be very important in understanding how soil organic matter forms and how molecules of biological significance can be synthesized abiotically. *See* CATALYSIS; CHEMICAL BONDING; COORDINATION CHEMISTRY; COORDINATION COMPLEXES.

Garrison Sposito

Plant Nutrition

Soil chemistry controls the availability of plant nutrients within the soil and thus influences plant growth, yield, and nutritional value for human or animal consumption. The nutritional deficiencies or toxicities inherent to the morphology of the soil are known as mineral stress; they often represent a serious constraint for crop production and development. Globally, mineral stress occurs on about 23% of the world's soils. Plant nutritional problems are especially severe on highly weathered soils in tropical or subtropical geographic areas. This occurs because phosphate is made unavailable by iron and aluminum

oxides and because the clay minerals often have low cation exchange capacities.

The predominant soil chemical problems on well-aerated soils include acidity, salinity (high salt), and sodicity (high sodium). Important soil chemical processes include dissolution and precipitation of readily soluble minerals. The primary soil chemical problems in waterlogged soils include rapid disappearance of molecular oxygen, denitrification, and increased toxic concentrations of manganese and iron.

Nutrient movement. Soil chemical reactions occur continuously to maintain equilibria between available nutrients in soil solution and those existing as minerals within the soil or on exchange sites located on organic matter or clay minerals. Diffusion gradients (**Fig. 4***a*) and mass flow (Fig. 4*b*) are two primary mechanisms for nutrient movement to plant roots. Phosphorus and potassium move primarily by diffusion, but calcium and magnesium are usually supplied by mass flow. Although root interception (Fig. 4*c*) occurs, this process does not significantly contribute to plant nutrient accumulation.

Nutrient-specific effects. Carbon, hydrogen, and oxygen are supplied to plants by water and carbon dioxide (CO_2). Provided photosynthesis and transpiration are not limited, soil chemistry has very little effect on these nutrients.

Nitrogen (N) is assimilated from soil solution as ammonium (NH_4^+) or nitrate (NO_3^-) ions. Legumes also assimilate nitrogen through symbiotic fixation. The nitrogen cycle is biologically complex, but provided soils are not anaerobic and causing denitrification, soil chemistry has less affect on nitrogen than on a nutrient such as phosphorus. One exception is that the balance between NH_4^+ and NO_3^- forms can influence plant availability of nutrients such as magnesium. If NH_4^+ ions interfere with magnesium accumulation in plants, then availability of magnesium to grazing animals may be inadequate, leading to hypomagnesemia. *See* NITROGEN; NITROGEN CYCLE; NITROGEN FIXATION.

Phosphorus concentrations in soil solution are very low when compared with amounts present in organic matter and with amounts in distinct phosphate compounds such as monocalcium phosphate, dicalcium phosphate, octacalcium phosphate, hydroxyapatite, and fluorapatite. Phosphorus also exists as surface films coating soil aggregates and individual particles of sand, silt, and clay.

Phosphorus nutrition of plants is strongly affected by soil chemical conditions, especially soil reaction or pH (**Fig. 5**). Acid soils (pH < 5.0) have reduced availability of phosphorus because relatively insoluble iron and aluminum phosphate compounds are formed. In alkaline soils (pH > 7.0), insoluble calcium phosphate compounds, including octacalcium and hydroxyapatite, can substantially reduce available phosphorus. Primary geographic regions for alkaline phosphorus problems are the semiarid and arid regions, where free calcium carbonate ($CaCO_3$) is often found in the plant root zone. Phosphorus deficiencies associated with acid soil conditions are most prevalent in areas with high rainfall, including the southern United States as well as tropical or subtropical areas.

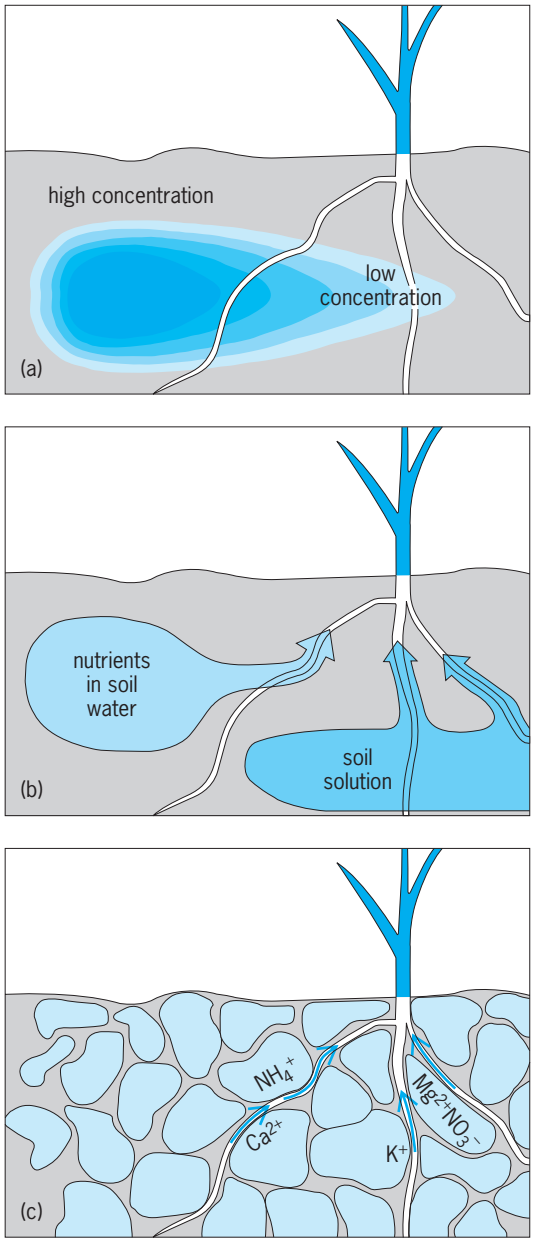

Fig. 4. Mechanisms of nutrient movement. (*a*) Diffusion; gradient movement. Fertilizer moves from high to low concentration. (*b*) Mass flow; water use controlled. (*c*) Root interception; contact required. NH_4^+ = ammonium ion; Ca^{2+} = calcium ion; K^+ = potassium ion; Mg^{2+} = magnesium ion; NO_3^- = nitrate ion.

Potassium is usually the predominant cation in plants, even though calcium is usually the predominant cation in soil. Deficiencies occur on all soils but are most prevalent on sandy soils that are low in organic matter, chalky soils, and soils derived from peat. Excess potassium can exist because of either natural mineral deposits or excessive additions of fertilizer or manure. Maintaining a balance among potassium, calcium, and magnesium is critical,

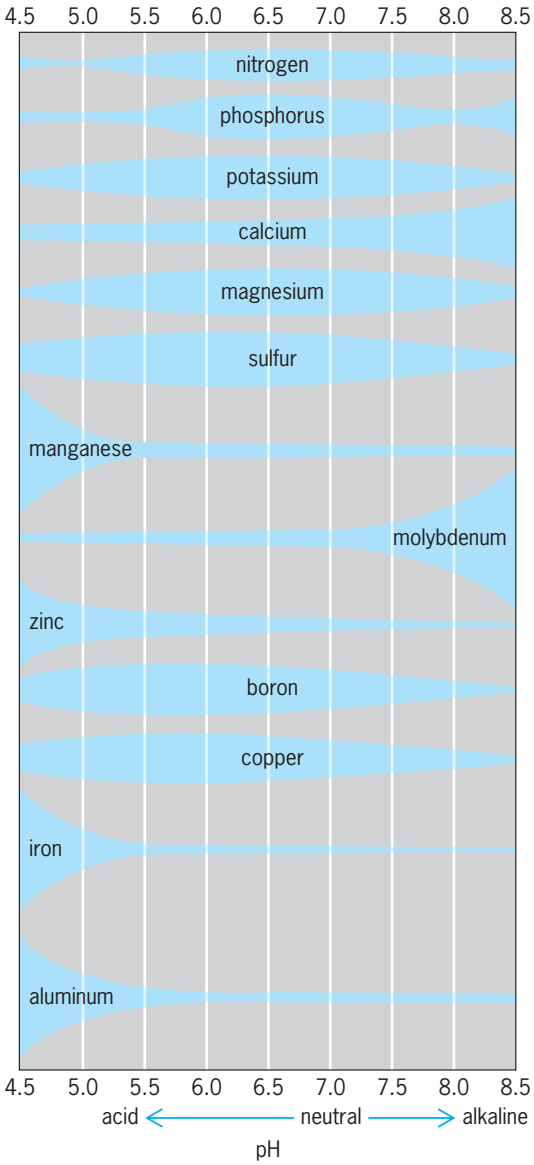

Fig. 5. Relationship of soil pH and plant nutrient availability, which is increased for the major plant nutrients when the soil pH is held between 6.5 and 7.0. Height of the tinted regions represents availability at a given pH level.

because imbalances can affect plant growth and development, and also can induce hypomagnesemia or grass tetany in cattle grazing forages on those soils.

Unless soils are extremely acidic, they can supply sufficient calcium for optimum plant growth. An exception is soils formed from serpentine minerals. Those soils often have a chemical imbalance between calcium and magnesium, and they have true calcium deficiencies. However, serpentine soils have many other chemical problems, including toxic levels of nickel and other heavy metals, as well as deficiencies of molybdenum.

Soil chemical processes generally supply magnesium in quantities sufficient to meet plant demand. Exceptions and areas where fertilizer response to magnesium may be found include sandy soils with low concentrations of magnesium, acid soils that have been treated with only calcitic limestone, and soils that have very high concentrations of potassium.

Saline and sodic soils. Saline, saline-sodic, and sodic soils are found in arid and semiarid regions where water loss through evaporation exceeds loss through leaching. This results in an accumulation of soluble salts, precipitated in the surface horizon and attached to cation exchange sites.

Saline soils contain enough chlorides and sulfates of sodium, calcium, and magnesium to interfere with growth of most plants. These are known as white-alkali soils, because the salts often give the soil light color. Saline-sodic soils also contain high amounts of neutral soluble salts and enough sodium to affect most plants seriously. These soils are most difficult to reclaim, because if they are leached with pure water, sodium ions will cause the clays to disperse and completely seal the soil surface. Reclamation must be done by using water with high concentrations of calcium or magnesium. Sodic soils do not contain large amounts of neutral soluble salts but have concentrations of sodium. They often have a very poor physical condition, because sodium disperses the clay particles. They are often known as black-alkali soils, because high concentrations of sodium carbonate (Na_2CO_3) disperse organic matter near the soil surface.

Micronutrients. Plant micronutrients are essential elements found in very low concentrations within plant tissues. These include boron, copper, iron, manganese, molybdenum, and zinc. Soil chemical processes that affect micronutrient availability include formation or dissolution of precipitates that contain these nutrients. Micronutrient availability is significantly influenced by soil pH (Fig. 5). Soil organic matter or microbial and root exudations also affect availability of micronutrients through formation of soluble coordination or chelation compounds. Any soil chemical change that influences formation of these compounds significantly affects micronutrient availability. *See* MYCORRHIZAE; RHIZOSPHERE.

Greenhouse culture. The principles of soil chemistry also apply to greenhouse growth media and potting mixes for home use. However, those media have physical and chemical properties that are distinctly different from natural soils. Since the mid-1990s, greenhouse operators have changed from using media containing soil to peat- or bark-based mixes containing manufactured materials such as perlite, vermiculite, or expanded polystyrene beads. These soilless growth media have good water-holding and aeration properties but limited nutrient holding capacities. As a result, fertility and plant nutrient management are more important than formerly.

Nutrient concentrations in the soil solution around the root are critical in such soilless growth media and depend upon the water-holding capacity of the medium. As water content increases, the nutrient concentration in solution decreases, because at complete saturation the water content is approximately four times the level at the permanent wilting point

and about two times that held against gravity. Total soluble nutrient concentration at saturation is one-fourth that at the permanent wilting point and one-half that at water held against gravity.

Developing a single set of fertilization guidelines is difficult because soilless growth media vary widely in bulk density (weight per unit volume). However, by analyzing a saturated media extract for pH, soluble salt content, and available nutrient levels, a balanced fertility program can be developed.

As with natural soils, growth-media pH influences the availability and uptake of all essential plant nutrients (Fig. 5). For peat-based media, however, the most desirable pH is 5.6–5.8 for most plants. If values exceed 6.5, micronutrient deficiencies may occur, and if the values are below 5.3, calcium or magnesium deficiencies or manganese toxicity may occur. All soluble ions or nutrients such as nitrate, ammonium, potassium, calcium, magnesium, chloride, and sulfate contribute to the soluble salt content of a growth media or water. The accumulation of these materials to toxic levels can be minimized by adding enough water to cause some leaching.

Whether in natural soil or soilless media, plants that have insufficient nutrients to maintain proper growth and development present several symptoms that can be used to help diagnose the primary problem. Beginning with the older leaves, plants deficient in nitrogen become light green in color. This occurs because the nitrate form of nitrogen is soluble and therefore mobile in the growth medium, so with watering, some may leach out. A fairly constant level

TABLE 4. Typical quantities of plant nutrients accumulated by harvested and nonharvested portions of various feed, food, and tree crops

Crop	Quantity, kg/ha*					
	Yield†	N‡	P	K	Mg	S
Corn grain	12,555	168	43	53	20	17
Stem, leaves, etc.		130	13	194	52	20
Wheat grain	5,380	100	22	25	13	6
Straw		50	5	126	13	17
Oat grain	3,590	90	12	19	6	9
Straw		40	7	116	17	12
Barley grain	5,380	123	20	32	9	11
Straw		45	7	107	10	11
Canola grain	2,350	74	16	15	—	13
Straw		44	7	62	—	10
Rice grain	7,850	86	23	26	9	6
Straw		39	7	112	7	8
Flax grain	1,255	45	8	14	—	3
Straw		16	2	28	—	3
Sorghum grain	8,970	134	30	28	16	25
Stem, leaves, etc.		146	15	158	34	18
Soybean grain	4,035	269§	24	78	19	13
Stem, leaves, etc.		94	8	54	11	14
Cotton lint + seed	1,680	105	19	41	12	8
Stalks, leaves, etc.		96	12	76	27	26
Sugarbeet roots	67,260	140	7	232	30	11
Tops		146	12	279	59	39
Sugarcane stalks	224,200	179	44	312	45	60
Tops and trash		224	32	255	67	36
Peanuts (nuts)	4,485	157§	11	32	6	11
Vines		112	8	140	22	12
Potatoes (tubers)	56,050	194	36	261	16	17
Vines		108	8	246	40	8
Sunflower seeds	3,925	140	30	36	13	7
Stems, leaves, etc.		57	5	83	34	11
Apples (fruit)	28,025	22	4	46	2	—
Leaves, new wood, etc.		90	19	121	25	—
Peaches (fruit)	33,630	39	5	60	13	—
Leaves, new wood, etc.		67	15	51	11	—
Oranges (fruit)	60,535	100	11	151	11	8
Leaves, new wood, etc.		196	16	152	31	24
Tomatoes (fruit)	67,260	112	11	187	9	24
Vines		90	12	112	22	22
Fababean (grain)	6,000	240	30	70	4	9
Field pea (grain)	6,000	187	29	75	8	16
Chickpea (grain)	1,500	91	6	50	11	9
Lentil (grain)	2,000	100	12	65	—	—
Alfalfa	17,935	504§	39	446	45	45
Tall fescue	7,850	170	32	172	14	—

*1 hectare = 10,000 square meters. Dividing by 10 gives values in grams per square meter.
†Yields are for the harvested portion of each crop (that is, grain for corn, wheat, oats, barley, canola, rice, flax, sorghum, soybean; lint + seed for cotton; roots for sugarbeet; stalks for sugarcane; leaves for tobacco; nuts for peanut; tubers for potatoes; seeds for sunflower; fruit for apples, peaches, oranges, tomatoes; and whole plant for alfalfa and tall fescue).
‡Accumulations given are total amounts taken up by the plants in both the harvested and the aboveground unharvested portions.
§ Legumes such as soybean, peanut, and alfalfa can get most of their nitrogen from the air.

of available nitrogen can be maintained by injecting additional nitrogen into the watering system. An adequate phosphorus supply is important for root system development, rapid growth, and flower quality in floral plants. Deficient plants exhibit slow root and top growth.

The nutrient most often limiting in greenhouse fertility programs is potassium. The lower or oldest leaves of plants deficient in potassium show marginal yellowing or chlorosis. Many greenhouse plants have a potassium requirements equal to or greater than their nitrogen requirement.

Calcium availability for plant uptake is dependent on pH of the growth media and levels of other cations present, especially potassium and magnesium. Calcium deficiency results in abnormal growth or death of the growing tip.

Nutrient balance is just as important in soilless mixtures as in natural soils. The balance among potassium, calcium, and magnesium is especially important since these nutrients compete for similar uptake sites at plant root surfaces. Maintaining a balanced level of nutrients also enables plants to tolerate stresses such as low fertility or high levels of soluble salts. *See* PLANT MINERAL NUTRITION.

Douglas Karlen

Nutrient accumulation and removal. The relationship between soil chemistry and plant nutrition is evident when the elemental composition of plants is determined and the amount of each nutrient within the vegetative material (biomass) is computed [concentration of a specific element (mass per unit dry matter) times the amount of biomass (dry matter per unit area or container for greenhouse conditions)]. These determinations, made for plants grown in either soil or greenhouse potting media, are used in several ways. By measuring nutrient accumulation several times throughout a plant's life cycle, rates of nutrient accumulation can be computed for either the whole plant or for its parts (leaves, stalks, seeds). Nutrient accumulation measurements are used to determine how nutrient availability (supply), movement, and other factors (water, temperature, light) are affecting physiological processes within plants grown in a specific type of soil, potting media, or solution culture (hydroponics). Computing the quantity of nutrients accumulated by plants (**Table 4**) is useful for determining fertilizer needs, estimating nitrogen fixation, and predicting food or feed value. It is also useful for understanding nutrient cycling between plants, soil micro- and macrofauna (mites, collembola, ants, beatles, earthworms), soil flora (algae, fungi, bacteria), and the mineral fraction (sand, silt, and clay) of a soil or greenhouse media (which may contain inert materials as a portion of the mineral fraction). *See* BIOMASS.

Differences in the quantity of nutrients accumulated within vegetative (stems, stalks, straw, leaves, vines, or new wood) or reproductive (grain, seeds, or fruit) plant parts, especially nitrogen, phosphorus, and potassium, illustrate the various physiological roles that each element has in supporting plant growth and development. Quantities of nitrogen and phosphorus tend to be greater in the grain or reproductive parts where protein and energy reserves are stored, while potassium accumulation is generally much greater in the vegetative material where it plays a vital role in cellular physiology and stem strength. Total nitrogen accumulation differs in legumes (soybean, peanuts, alfalfa), corn (or sorghum), and grass crops (wheat, oat, barley, rice). Also, it is the relative quantity or amount of each nutrient accumulated by plants (Table 4), and not their physiological importance, that provides the rationale for classifying nitrogen, phosphorus, and potassium as macronutrients; magnesium, sulfur, and calcium as secondary nutrients; and boron, copper, iron, manganese, and zinc as micronutrients by agronomists, soil scientists, and other plant nutrition professionals. *See* BORON.

Predicting the feed or food value of plants for animal and human consumption is another reason for measuring nutrient accumulation. However, with regard to nutritional value, the quantity of nutrients within a plant available to the animal (including humans) consuming that plant will usually be substantially lower than the total nutrients accumulated by the plant. In addition to soil chemistry effects, postharvest handling can significantly change the bioavailability of nutrients, minerals, amino acids, and vitamins. Soil chemistry, plant chemistry, and external factors (water, heat, light stress) are complex interactions. The simple statement that healthy soils produce healthy plants which sustain healthy animals and people is very difficult to demonstrate quantitatively.

Worldwide, nutritional deficiencies (of calories and protein) are still among the most crucial human problems. Thus research and technology efforts focusing on the judicious use of fertilizers to increase food, feed, and fiber production, without negatively impacting water or air resources, remain important. *See* MALNUTRITION.

Douglas Karlen

Animal Nutrition

Mineral elements that are essential for ruminant animals, such as cattle and sheep, are commonly arranged into two groups: macronutrients and micronutrients (**Table 5**). The groups are distinguished by relative dietary requirements and amounts of mineral elements in animal tissues. For example, the macronutrients are required in the diet at concentrations greater than 100 micrograms per gram (parts per million), and the micronutrients, or trace elements, are required at concentrations less than 100 ppm. The elements calcium, phosphorus, magnesium, potassium, sodium, sulfur, and chlorine are considered macronutrients because tissue concentrations and dietary requirements of these essential elements are relatively high. In contrast, elements that generally occur at relatively low concentrations in living tissues and in diets are designated trace elements. The trace elements are chromium, cobalt, copper, iodine, iron, manganese, molybdenum, selenium, and zinc. Additionally, since about 1970, other trace elements have been shown to be required

TABLE 5. Critical concentrations of mineral elements in forages and dietary amounts necessary to meet requirements for normal growth and development of ruminant animals

Element	Critical concentration*		Dietary requirement†		Potentially toxic dietary concentration	
	Grass	Legume	Cattle	Sheep	Cattle	Sheep
Macronutrients						
	g/kg dry weight (parts per thousand)‡					
Calcium	2	18	4	2	20	20
Chlorine	0.3§	0.3§	2	1	4	4
Magnesium	1	3	2	1	6	6
Phosphorus	2.5	2.5	2	2	10	6
Potassium	25	25	8	3	30	30
Sodium	NR¶	NR¶	1.5	1	3	3
Sulfur	2	2.5	1.5	1.5	4	4
Micronutrients						
	μg/g dry weight (parts per million)					
Copper	4	11	8	6	100	25
Iodine	NR¶	NR¶	0.1	0.1	50	50
Iron	50	30	40	40	1000	1000
Manganese	20	25	25	40	1000	1000
Molybdenum	0.1	0.5	0.1	0.1	5	6
Selenium	NR¶	NR¶	0.1	0.1	5	3
Zinc	12	20	25	30	500	1000

* Concentration of nutrient required to be nonlimiting to growth.
† Dietary requirement for normal growth and development.
‡ 1 g/kg = 1 part per thousand = 0.1%.
§ Concentration not adequately defined.
¶ Not required for most higher plant species.

or beneficial to some species of animals. Trace elements in this group are arsenic, boron, lead, lithium, nickel, silicon, tin, and vanadium. The beneficial effects of many of the elements in this third group have been established mainly in laboratory animals under highly controlled environmental conditions. At present, there is no information to indicate that deficiencies have occurred in grazing ruminant animals under practical conditions. *See* AGRICULTURAL SCIENCE (ANIMAL).

Table 5 shows the mineral elements considered by most authorities to be essential to the nutritional health of cattle and sheep. Mineral requirements of animals are often expressed as amounts needed per day or as a portion (percentage or ppm) of the dry matter of the diet. In general, mineral requirements are affected by numerous animal factors, including species, breed, sex, age, weight, pregnancy, and stage of lactation. Within the body, mineral elements have three broad functions: (1) some elements, particularly trace elements, either are integral constituents of enzymes or activate specific enzymes; (2) some elements are important constituents of the structural components of organs and tissues; and (3) some elements are important components of fluids and tissues where they are concerned with maintaining osmotic pressure, pH balance, impulse conduction, and membrane permeability. **Table 6** shows some generalized concentrations of macrominerals in selected tissues and fluids of grazing animals. Total amounts and concentrations of mineral elements in animal tissues are variable and are influenced by many factors, including age, sex, hormones, nutrient intake, and nutrient interactions. Consequently, amounts of minerals reported for tissues may not represent all cases or all ages of animals.

Grasses, legumes, and forbs are the main sources of mineral nutrients consumed by grazing ruminants, such as cattle, sheep, and goats. Also, these animals may obtain some minerals from the water they drink and the soil they ingest, either passively (because of soil deposition on foliage) or actively (licking or eating the soil directly). Plants obtain most of their mineral nutrients through absorption from the soil, but some elements may be absorbed through the leaves after atmospheric deposition. Complex factors regulate the movement of mineral elements from rocks to soil, uptake of mineral nutrients by plants from the soil, transport of mineral nutrients from roots to other parts of the plant, and deposition of mineral nutrients in various components of the plant. Consequently, concentrations of mineral nutrients in plant foods and forages are highly variable and difficult to predict (**Fig. 6**). *See* SOIL FERTILITY.

The amount of an element required in plant tissues to prevent growth or yield retardation in the plant is called the critical concentration. Critical concentrations of elements may deviate between plant species and the values may vary, depending on the stage of growth and anatomical part of the plant that is sampled. Therefore, the values presented in Table 5 represent approximate amounts for numerous species and should not be used to

TABLE 6. Concentrations of mineral elements in selected tissues and fluids of animals*

Element	Tissue or fluid: Blood plasma	Muscle	Milk	Wool
Macrominerals				
	mg/100 ml	*mg/g dry wt*	*g/L*	
Calcium	9–12	0.2–0.6	1.2	
Chlorine	365	3–4	1	
Phosphorus	4–9	9–13	0.95	
Magnesium	1.8–3.2	1–1.5	0.12	
Potassium	15–16	15–17	1.4	
Sodium	325	3–5	0.5	
Micronutrients				
	μg/100 ml	*μg/g dry wt*	*μg/L*	*μg/g*
Cobalt	0.04–0.08	0.1	0.5	
Copper	90–120	4	40–200	4–140
Iodine	0.08–60	< 1	20–150	
Iron	90–270	200	20–600	50
Manganese	0.5	3	20–50	18
Molybdenum	1–5	0.1–5	18–120	1–15
Selenium	2–30	0.04–0.2	10	0.1
Zinc	70–120	80–240	3000–5000	80–220

*Concentrations of mineral elements in animal tissues are highly variable and may be affected by numerous factors, such as sex, age, hormones, dietary intake, and nutrient interactions.

compare individual plant species. Many of the minerals essential to animals occur at sufficient concentrations in practical feedstuffs and forages to meet the needs of cattle and sheep. However, in some circumstances sufficient concentrations of mineral elements may accumulate in grasses to support optimal growth of the plants, but still not provide sufficient amounts of some minerals to prevent potential deficiency problems in grazing animals. Moreover, some plants may accumulate quantities of mineral elements that may be toxic when consumed by grazing animals.

In order to establish the minimum dietary mineral concentrations necessary to meet the requirements of animals, it is necessary to have information on the bioavailability of mineral nutrients in various feeds and forages. Bioavailability refers to that proportion of the total amount of a mineral element present in

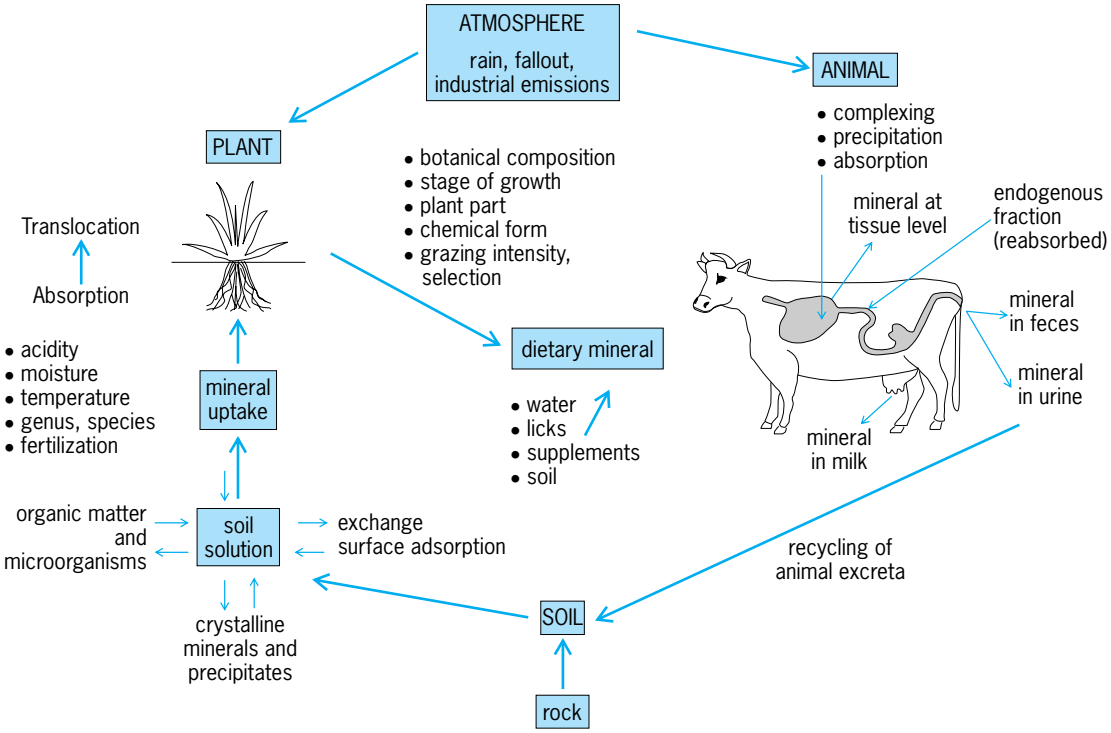

Fig. 6. Soil-plant-animal mineral cycle. Heavy arrows represent major pathways or shifts. Light arrows represent relatively minor pathways.

feeds or forages that is potentially absorbable in a metabolically active form. Moreover, as generally assessed, bioavailability is not a fundamental property of a feed or forage, but represents the response of a test animal to the diet or food. It would be very time-consuming and expensive to determine the bioavailability of mineral elements in all forages. Therefore, general estimates of mineral bioavailability, together with estimates of the amount of food consumed, have been used to approximate the minimal dietary concentrations of minerals necessary to meet the physiological needs of animals. Table 5 provides information on the amount of minerals required in the diet to meet the needs of growing animals. *See* ANIMAL GROWTH; NUTRITION.

Macronutrients. Macronutrients are the mineral elements required in ruminant animals' diet at concentrations greater than 100 μg/g.

Calcium. Calcium, the most plentiful mineral element in the body, is involved in numerous functions, such as blood clotting, bone structure, membrane permeability, muscle contraction, nerve impulse transmission, stability and activity of some enzymes, and secretion of specific hormones. Most of the calcium in the body (about 98%) occurs as part of the structural components of bone and teeth, and the remainder is found in extracellular fluids and soft tissues. The amount of calcium consumed from forage is affected by exchangeable calcium levels in soil, climatic conditions, plant species eaten, stage of plant growth, and portion of the plant consumed. *See* CALCIUM.

Phosphorus. Usually, phosphorus requirements and metabolism are considered with calcium because the two elements function together in bone. About 80–85% of the phosphorus in the body occurs in bone and teeth with the remainder in soft tissues. In addition to its role in bone structure, phosphorus is involved in several critical functions, including cell growth and differentiation, energy use, membrane structure, and maintenance of osmotic and pH balance. Phosphorus absorption is influenced by the amount and source of phosphorus consumed, the age of the animal, and dietary levels of fat, macrominerals (particularly calcium, magnesium, and potassium), and trace minerals (especially iron and manganese). In grazing animals, phosphorus deficiency may result in reduced growth, depraved or abnormal appetite, lowered reproduction, and impaired bone mineralization. A phosphorus deficiency may occur if the forage level is below 0.2%, a condition that may be found in tropical regions. *See* PHOSPHORUS.

Magnesium. Magnesium is involved in some important physiological functions, including maintenance of electrical potentials and transmission of neuromuscular impulses, structural integrity of bones and teeth, and activation of numerous enzyme systems. Also, as an ATP-magnesium complex, magnesium participates in all biosynthetic process. In cattle and sheep, about 60–70% of total body magnesium is in bone, 15% in muscle, 15% in other soft tissues, and 1% in extracellular fluids. In young animals, as much as 30% of the magnesium in bone can be mobilized during dietary magnesium deficiencies, but the value in adult animals is only about 2%. Grasses may have magnesium concentrations that are optimal for plant growth but lower than the requirements for grazing animals. *See* MAGNESIUM.

Potassium. Potassium, located mainly in intracellular fluids, affects osmotic pressure and pH balance within cells. Potassium also is involved in enzymatic reactions, nerve impulse transmission, muscle contraction, and transport of oxygen and carbon dioxide in blood. Secretion of potassium in milk increases the need for potassium by lactating animals, compared to growing animals. Forages usually contain sufficient potassium to meet the needs of grazing animals, but feed concentrates, such as corn silage, may not contain enough potassium to meet the needs of ruminant animals. Additionally, because potassium concentrations decline with increased maturity of forages, marginal potassium deficiencies, characterized by decreased food intake and reduced weight gain, may occur in animals grazing mature forages in winter or in drought periods. Severe potassium deficiencies are not likely to occur in grazing animals. *See* POTASSIUM.

Sodium and chlorine. In extracellular fluids, sodium is the major positively charged ion (cation) and chloride is the major negatively charged ion (anion). These elements function in controlling body fluid balance, regulating osmotic pressure, and maintaining pH balance. Also, sodium is involved in nerve impulse transmission, muscle contraction, and various transport processes. Chloride is necessary for the production of hydrochloric acid in gastric juice produced by the stomach and for the activation of the digestive enzyme amylase. Grazing beef cattle and sheep generally require less sodium than do lactating dairy cows because the cows secrete substantial amounts of sodium in milk. Sodium is not considered an essential nutrient for most forage species, and the sodium content of forages varies markedly. If sodium intake meets requirements, it is generally assumed that chlorine intake is sufficient because chlorine is usually associated with sodium. *See* SODIUM.

Sulfur. Sulfur is an integral component of several amino acids, particularly methionine, cysteine, and cystine, and it is part of the vitamins thiamin and biotin. Moreover, sulfur occurs in many compounds in the form of sulfate. One such compound, chondroitin sulfate, is an important part of blood vessels, bone, cartilage, and tendons. Moderate sulfur deficiency reduces food intake and lowers rates of protein synthesis by microorganisms in the rumen. When provided sufficient amounts of sulfur, the ruminal microorganisms can synthesize all of the sulfur-containing compounds required by the animal. Mature forages and forages grown in areas with sulfur-deficient soils may not provide enough sulfur to meet requirements. Dietary sulfur requirements are difficult to establish because of complex quantitative interactions between sulfur, copper, and molybdenum. Increasing the sulfur content of the diet lowers the absorption of copper. Additionally, sulfur

interacts with molybdenum to reduce copper retention. *See* SULFUR.

Micronutrients (trace elements). Micronutrients are the mineral elements required in ruminant animals' diet at concentrations less than 100 ppm.

Cobalt. Although cobalt is an indispensable micromineral, it is unique in that the requirement by most mammals is not for an ionic form of the element but for a preformed cobalt compound, vitamin B_{12}, also called cobalamin. Lack of this vitamin is the cause of signs of cobalt deficiency. Ruminant animals are unique in that vitamin B_{12} is synthesized by the microorganisms in the rumen. However, ruminal microorganisms are not able to synthesize sufficient quantities of vitamin B_{12} to meet the needs of the animal when dietary cobalt intake is inadequate, such as in areas of the world where soils are cobalt-deficient. In the United States, the sandy soils of the southeastern Coastal Plain and the leached soils derived from granite in New England are generally cobalt-deficient, and forages growing on these soils contain low amounts of cobalt. In lambs and calves, cobalt deficiency is likely to occur when forage concentrations of cobalt are below 0.1 ppm (dry-weight basis). Fertilizing pastures with cobalt or providing supplemental dietary cobalt are methods used to prevent cobalt deficiency in grazing ruminant animals. Establishing that certain pathological signs in grazing animals were attributable to cobalt deficiency is a well-documented example of a direct relationship between soil properties and animal health. *See* COBALT; VITAMIN.

Copper. Copper is an essential component of many enzymes. The amount of dietary copper required by grazing ruminants varies and is highly influenced by the amount of sulfur and molybdenum in the diet. Sulfur in dietary compounds may be reduced to sulfides by rumen microorganisms, and the sulfides may react with copper to form copper sulfide, a relatively insoluble compound that is not absorbed by the animal. Additionally, molybdenum in the diet may interact with sulfides in the rumen to form compounds called thiomolybdates. Then, copper may react with the thiomolybdates to form molybdenum-copper-sulfur compounds that are not absorbed. Moreover, some of the thiomolybdates may be absorbed, and further adversely affect copper metabolism in the body. The copper content of forages depends on the plant species and amount of available copper in the soil; grasses usually contain less copper than do legumes, such as clover. *See* COPPER.

Iron. Iron is an essential component of hemoglobin and myoglobin, the proteins involved in oxygen transport in blood and muscles, respectively. Within cells, iron is a component of compounds (cytochromes and proteins) that are involved in the use of oxygen to produce energy. Also, some enzymes contain iron or are activated by iron. For grazing ruminants, the dietary requirement for iron is about 50 ppm. Most forages and feedstuffs contain more than enough iron to meet the needs of grazing animals. Additionally, grazing animals may obtain iron from their drinking water or from eating soil. Excessively high intake of iron may depress the absorption of dietary copper. *See* IRON.

Manganese. Manganese is an essential component of some enzymes and an activator of other enzymes. Some of the metal activations are nonspecific; that is, for some enzyme activations, manganese can be replaced by other mineral ions, particularly magnesium. Manganese deficiency results in skeletal abnormalities in young animals, and affects reproduction in mature animals. The amount of manganese required for normal skeletal development is greater than the amount needed for maximum growth, but is less than the amount needed for optimal reproductive performance. Thus, the dietary manganese requirement for growing animals is about 20 ppm but that for breeding animals is about 40 ppm. Forage concentrations of manganese vary with plant species and are affected by soil acidity, soil moisture, and soil manganese content. Generally, forages contain sufficient amounts of manganese to meet the needs of grazing ruminants. *See* MANGANESE.

Molybdenum. Molybdenum is an essential component of several oxidase enzymes, but the dietary requirements for this element are not well established. Molybdenum requirements are low, and under practical grazing conditions there is no evidence that molybdenum deficiency occurs in grazing ruminant animals. However, in some cases supplemental molybdenum may increase ruminal microbial activity and thereby increase cellulose digestibility. Concentrations of molybdenum in forages vary greatly. Forages growing on moist (poorly drained) soils with high organic matter and neutral or alkaline pH are likely to have high concentrations of molybdenum. *See* MOLYBDENUM.

Selenium. Selenium is an indispensable part of an enzyme that protects cells from damage caused by some reactions that occur within cells. Additionally, selenium is an integral component of an enzyme responsible for converting thyroxine, a hormone that contains iodine, to its more active form, triiodothyronine. In grazing ruminant animals, selenium deficiency generally does not occur when the selenium concentration in forage is at least 0.1 ppm. However, there are regions in the United States and Canada where forages on range lands or in pastures contain less than adequate amounts of selenium because the selenium content of the soil is low. In contrast, there are other regions where soil selenium levels are high and plants growing on these seleniferous soils may accumulate very high amounts of selenium. Excessive consumption of these selenium accumulator plants may cause selenium toxicity. Maps produced by the U.S. Department of Agriculture show the regions where selenium deficiency and selenium toxicity are likely to affect the health of grazing ruminant animals. *See* SELENIUM.

Zinc. Zinc is an essential trace element because it is a component of a number of enzymes and because it is necessary for the activation of other enzymes. Also, zinc is important for the proper functioning of the immune system. The dietary zinc requirement for grazing ruminant animals has been set at 30 ppm.

However, the amount of zinc required by grazing ruminants is not as well defined as it is for other animals, partly because zinc status is difficult to assess. The amount of zinc in forages is affected by soil zinc content, soil acidity, stage of plant growth, and plant species. Grasses generally contain less zinc than do legumes, such as clover. Generally, the requirement for zinc in animal diets is higher than forage zinc concentrations necessary for normal plant growth. Forages in some areas, particularly those growing on calcareous soils with alkaline pH may not contain enough zinc to fully meet the needs of all grazing animals, and marginal or subclinical zinc deficiencies may occur. *See* ZINC. William A. House

Bibliography. C. B. Ammerman, D. H. Baker and A. J. Lewis (eds.), *Bioavailability of Nutrients for Animals*, 1995; K. C. Beeson and G. Matrone, *The Soil Factor in Nutrition: Animal and Human*, 1976; C. A. Black, *Soil Fertility Evaluation and Control*, 1993; J. J. R. Frausto da Silva and R. J. P. Williams, *The Biological Chemistry of the Elements*, 1991; M. B. McBride, *Environmental Soil Chemistry*, 1994; H. Marschner, *Mineral Nutrition of Higher Plants*, 2d ed., 1995; L. R. McDowell, *Minerals in Animal and Human Nutrition*, 1992; K. Mengel and E. A. Kirkby, *Principles of Plant Nutrition*, 4th ed., 1987; F. M. M. Morel and J. G. Hering, *Principles and Applications of Aquatic Chemistry*, 1993; J. J. Mortvedt et al. (eds.), *Micronutrients in Agriculture*, 2d ed., Soil Science Society of America, Madison, WI, 1991; National Research Council (NRC), *Nutrient Requirements of Beef Cattle*, 7th ed., 1996; NRC, *Nutrient Requirements of Sheep*, 6th ed., 1985; NRC, *Nutrient Requirements of Dairy Cattle*, 6th ed., 1988; E. W. Russell, *Russell's Soil Conditions and Plant Growth*, 11th ed., 1988; D. L. Sparks, *Environmental Soil Chemistry*, 1995; G. Sposito, *The Chemistry of Soils*, 1989; W. Stumm, *Chemistry of the Soil-Water Interface*, 1992; R. L. Westerman (ed.), *Soil Testing and Plant Analysis*, 3d ed., Soil Science Society of America, Madison, WI, 1990.

Soil conservation

The practice of arresting or minimizing artificially accelerated soil deterioration. Its importance has grown because cultivation of soils for agricultural production, deforestation and forest cutting, grazing of natural range, and other disturbances of the natural cover and position of the soil have increased greatly since the middle of the nineteenth century in response to the growth in world population and human technical capacity. Accelerated soil deterioration has been the unfortunate consequence.

Erosion extent and intensity. Accelerated erosion has been known throughout history wherever people have tilled or grazed slopes or semiarid soils. Much evidence of accelerated erosion exists in the eastern and central Mediterranean basin, Iraq, China, and many other areas. Wherever the balance of nature is a delicate one, as on steep slopes in regions of intense rainstorms, such as those along the Caribbean coast of Venezuela, or in semiarid regions of high rainfall variability, grazing and cultivation eventually have had to contend with serious or disabling erosion. Irrigation works of the Tigris and Euphrates valleys are thought to have suffered from the sedimentation caused by quickened erosion on the rangelands of upstream areas in ancient times. Hill sections of Israel, Jordan, Syria, Greece, and central and southern Italy experienced serious and often virtually total soil loss down to solid bedrock from almost total deforestation, overgrazing by sheep and goats, and other land-use mismanagement many centuries ago. Accelerated water erosion on the hills of southern China and wind erosion in northwestern China also date far back. The economic and social effects of these soil movements have been much debated, but their impact appears to have been serious in many cases, such as that of classical Rome and that of China several centuries ago.

The exact extent of accelerated soil erosion in the world today is not known, particularly as far as the rate of soil movement is concerned. However, it may be safely said that nearly every semiarid area with cultivation or long-continued grazing, every hill land with moderate to dense settlement in humid middle latitude and subtropical climates, and all cultivated or grazed hill lands in the Mediterranean climate areas suffer to some degree from such erosion. These recognized and often growing problems of erosion occur in such culturally diverse areas as southern China, the Indian plateau, South Africa, Russia and Ukraine, and Spain.

Within the United States the most critical areas have been the hill lands of the interior Southeast, the Great Plains, the Palouse area hills of the Pacific Northwest, southern California hills, and slope lands of the Midwest. The high-intensity rainstorms of the Southeast and the cyclical droughts of the Plains have predisposed these two larger areas to erosion. The light-textured A horizon formed under the Plains grass cover was particularly susceptible to wind removal, while the high clay content of many southeastern soils predisposed them to water movement. These natural susceptibilities were repeatedly brought into play by agricultural systems which stressed corn, cotton, and tobacco in the Southeast, corn in the Midwest, and intensive grazing and small grains on the Plains, the Palouse, and in California. The bare soil surface left by the traditional intertilled cultivation of the Southeast furnished almost ideal conditions for water erosion and at the same time caused heavy nutrient depletion of soils thus cropped. The bare fields during seasons between crops have also been susceptible to soil depletion. Bare fields have been especially disastrous to maintenance of soil cover during droughts of the Plains. Soil mismanagement thus has been a common occurrence in parts of the United States where stability of soil cover hangs in delicate balance. *See* EROSION.

Types of soil deterioration. Soil may deteriorate either by physical movement of soil particles from a given site or by depletion of the water-soluble

Fig. 1. Erosion of sandstone caused by strong wind and occasional hard rain in an arid region. (*USDA*)

elements in the soil which contribute to the nourishment of crop plants, grasses, trees, and other economically usable vegetation. The physical movement generally is referred to as erosion. Wind, water, glacial ice, animals, and tools in use may be agents of erosion. For purposes of soil conservation, the two most important agents of erosion are wind and water, especially as their effects are intensified by the disturbance of natural cover or soil position. Water erosion always implies the movement of soil downgrade from its original site. Eroded sediments may be deposited relatively close to their original location, or they may be moved all the way to a final resting place on the ocean floor. Wind erosion, on the other hand, may move sediments in any direction, depositing them quite without regard to surface configuration. Both processes, along with erosion by glacial ice, are part of the normal physiographic (or geologic) processes which are continuously acting upon the surface of the Earth. The action of both wind and water is vividly illustrated in the scenery of arid regions (**Fig. 1**). Soil conservation is not so much concerned with these normal processes as with the new force given to them by human land-use practices. *See* LAND-USE PLANNING.

Accelerated erosion may be induced by any land-use practice which denudes the soil surfaces of vegetative cover (**Fig. 2**). If the soil is to be moved by water, it must be on a slope. The cultivation of a corn or a cotton field is a clear example of such a practice. Corn and cotton are row crops; cultivation of any row crop on a slope without soil-conserving practices is an invitation to accelerated erosion. Cultivation of other crops, like the small grains, also may induce accelerated erosion, especially where fields

Fig. 2. Improper land use. Corn rows planted up and down the slope rather than on the contour. Note better growth of plants in bottom (deeper) soil in foreground as compared to stunted growth of plants on slope. (*USDA*)

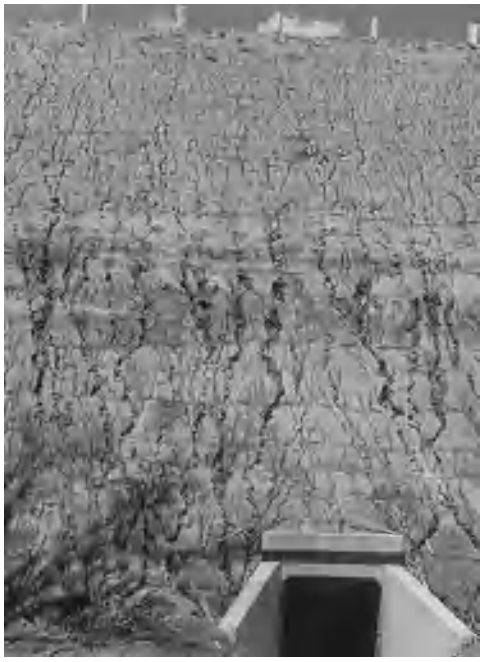

Fig. 3. Rill erosion on highway fill. The slopes have been seeded (horizontal lines) with annual lespedeza to bind and stabilize soil. (*USDA*)

are kept bare between crops to store moisture. Forest cutting, overgrazing, grading for highway use, urban land use, or preparation for other large-scale engineering works also may speed the natural erosion of soil (**Fig. 3**).

Where and when the soil surface is denuded, the movement of soil particles may proceed through splash erosion, sheet erosion, rill erosion, gullying, and wind movement (**Fig. 4**). Splash erosion is the minute displacement of surface particles caused by the impact of falling rain. Sheet erosion is the gradual downslope migration of surface particles, partly with the aid of splash, but not in any defined rill or channel. Rills are tiny channels formed where small amounts of water concentrate in flow. Gullies are V- or U-shaped channels of varying depths and sizes. A gully is formed where water concentrates in a rivulet or larger stream during periods of storm. It may be linear or dendritic (branched) in pattern, and with the right slope and soil conditions may reach depths of 50 ft (15 m) or more. Gullying is the most serious form of water erosion because of the sharp physical change it causes in the contour of the land, and because of its nearly complete removal of the soil cover in all horizons. On the edges of the more permanent stream channels, bank erosion is another form of soil movement.

Depletion of soil nutrients obviously is a part of soil erosion. However, such depletion may take place in the absence of any noticeable amount of erosion. The disappearance of naturally stored nitrogen, potash, phosphate, and some trace elements from the soil also affects the usability of the soil for human purposes. The natural fertility of virgin soils always is depleted over time as cultivation continues, but the rate of depletion is highly dependent on management practices. *See* PLANT MINERAL NUTRITION; SOIL; SOIL FERTILITY.

Causes of soil mismanagement. One of the chief causes of erosion-inducing agricultural practices in the United States has been ignorance of their consequences. The cultivation methods of the settlers of western European stock who set the pattern of land use in the United States came from physical environments which, because of the mild nature of their rainstorms and their prevailing soil textures, were in general far less susceptible to erosion than were the environments into which they migrated in North America. The principal European-sown grains, moreover, were wheat, barley, oats, and rye which covered the soil surface completely as they grew. Corn, cotton, and tobacco, however, were crops unfamiliar to European agriculture. In eastern North America the combination of European cultivation methods and American intertilled crops resulted in generations of soil mismanagement. In later years the Plains environment, with its alternation of drought

Fig. 4. Two most serious types of erosion. (*a*) Sheet erosion as a result of downhill straight-row cultivation. Note onions washed completely out of ground. (*b*) Gully erosion destroying rich farmland and threatening highway. (*USDA*)

Fig. 5. Wind erosion. The accumulation of topsoil was blown from the bare field on right. (*USDA*)

and plentiful moisture, was also an unfamiliar one to settlers from western Europe.

Conservational methods of land use were slow to develop, and mismanagement was tolerated because of the abundance of land in the eighteenth and nineteenth centuries. One of the cheapest methods of obtaining soil nutrients for crops was to move on to another farm or to another region. Until the twentieth century, land in the United States was cheap, and for a period it could be obtained by merely giving assurance of settlement and cultivation. With low capital investments, many farmers had little stimulus to look upon their land as a vehicle for permanent production. Following the Civil War, tenant cultivators and sharecroppers presented another type of situation in the Southeast, where stimulus toward conservational soil management was lacking. Management of millions of acres of farmland in the Southeast was left in the hands of people who had no security in their occupancy, who often were illiterate, and whose terms of tenancy and meager training forced them to concentrate on corn, cotton, and tobacco as crops.

On the Plains and in other susceptible western areas, small-grain monoculture, particularly of wheat, encouraged the exposure of the uncovered soil surface so much of the time that water and wind inevitably took their toll (**Fig. 5**). On rangelands, the high percentage of public range (for whose management little individual responsibility could be felt), lack of knowledge as to the precipitation cycle and range capacity, and the desire to maximize profits every year contributed to a slower but equally sure denudation of cover. *See* AGRICULTURAL SOIL AND CROP PRACTICES.

Finally, the United States has experienced extensive erosion in mountain areas because of forest mismanagement. Clearcutting of steep slopes, forest burning for grazing purposes, inadequate fire protection, and shifting cultivation of forest lands have allowed vast quantities of soil to wash out of the slope sites where they could have produced timber and other forest values indefinitely. In the United States the central and southern Appalachian area and the southern part of California have suffered severely in this respect, but all hill or mountain forest areas except the Pacific Northwest have had such losses. *See* FOREST MANAGEMENT.

Economic and social consequences. Where the geographical incidence of soil erosion has been extensive, the damages have been of the deepest social consequence. Advanced stages of erosion may remove all soil and therefore all capacity for production. More frequently it removes the most productive layers of the soil—those having the highest capacity for retention of moisture, the highest soil nutrient content, and the most ready response to artificial fertilization. Where gullying or dune formation takes place, erosion may make cultivation physically difficult or impossible. Thus, depending on extent, accelerated erosion may affect productivity over a wide area. At its worst, it may cause the total disappearance of productivity, as on the now bare limestone slopes of many Mediterranean mountains. At the other extreme may be the slight depression of crop yields which may follow the progress of sheet erosion over short periods. In the case of forest soil losses, except where the entire soil cover disappears, the effects may not be felt for decades, corresponding to the growth cycle of given tree species. Agriculturally, however, losses are apt to be felt within a matter of a few years.

Moderate to slight erosion cannot be regarded as having serious social consequences, except over many decades. As an income depressant, however, it does prevent a community from reaching full productive potentiality. More severe erosion has led to very damaging social dislocation. For those who choose to remain in an eroding area or who do not have the capacity to move, or for whom migration may be politically impossible, the course of events is fateful. Declining income leads to less means to

cope with farming problems, to poor nutrition and poor health, and finally to family existence at the subsistence level. Communities made up of a high proportion of such families do not have the capacity to support public services, even elementary education. Unless the cycle is broken by outside financial and technical assistance or by the discovery of other resources, the end is a subsistence community whose numbers decline as the capacity of the land is further reduced under the impact of subsistence cultivation. This has been illustrated in the hill and mountain lands of the southeastern United States, in Italy, Greece, Mexico, China, and elsewhere for many millions of farming people. Illiteracy, short lifespans, nutritional and other disease prevalence, poor communications, and isolation from the rest of the world have been the marks of such communities. Where they are politically related to weak national governments, indefinite stagnation and decline may be forecast. Where they are part of a vigorous political system, their rehabilitation can be accomplished only through extensive investment contributed by the nation at large. In the absence of rehabilitation, these communities may constitute a continued financial drain on the nation for social services such as education, public health, roads, and other public needs.

Effects on other resources. Accelerated erosion may have consequences which reach far beyond the lands on which the erosion takes place and the communities associated with them. During periods of heavy wind erosion, for example, the dust fall may be of economic importance over a wide area beyond that from which the soil cover has been removed. The most pervasive and widespread effects, however, are those associated with water erosion. Removal of upstream cover changes the regimen of streams below the eroding area. Low flows are likely to be lower and their period longer where upper watersheds are denuded than where normal vegetative cover exists. Whereas flood crests are not necessarily higher in eroding areas, damages may be heightened in the valleys below eroding watersheds because of the increased deposition of sediment of different sizes, the rapid lifting of channels above floodplains, and the choking of irrigation canals.

A long chain of other effects also ensues. Because of the extremes of low water in denuded areas during dry seasons, water transportation is made difficult or impossible without regulation, fish and wildlife support is endangered or disappears, the capacity of streams to carry sewage and other wastes safely may be seriously reduced, recreational values are destroyed, and run-of-the-river hydroelectric generation reaches a very low level. Artificial storage becomes necessary to derive the services from water which are economically possible and needed. But even the possibilities of storage eventually may disappear when erosion of upper watersheds continues. Reservoirs may be filled with the moving sediment and lose their capacity to reduce flood crests, store flood waters, and augment low flows. A dramatic example of such ongoing storage reduction with ominous implications for the future is Lake Nasser, one of the largest artificial lakes in the world, impounded by the great dam on the Nile River at Aswan, Egypt. For this reason, plans for permanent water regulation in a given river basin must always include watershed treatment where eroding lands are in evidence. *See* WATER CONSERVATION.

Conservation measures and technology. Measures of soil management designed to reduce the effects of accelerated erosion have been known in both the Western world and in the Far East since long before the beginning of the Christian Era. The value of forests for watershed protection was known in China at least 10 centuries ago. The most important of the ancient measures on agricultural lands was terrace construction, although actual physical restoration of soil to original sites also has been practiced. Terrace construction in the Mediterranean countries, China, Japan, and the Philippines represents the most impressive remaking of the face of the Earth before the days of modern earth-moving equipment (**Fig. 6**). Certain land management practices that were soil-conserving have been a part of western European agriculture for centuries, principally those centering on livestock husbandry and crop rotation. Conservational management of the soil was known in colonial Virginia and by Thomas Jefferson and others during the early years of the United States. However, it is principally since 1920 that techniques for soil conservation have been developed for many types of environment in terms of an integrated approach. The measures include farm, range, and forest management practices, and the building of engineered structures on land and in stream channels.

A first and most important step in conservational management is the determination of land capability—the type of land use and economic production to which a plot is suited by slope, soil type,

Fig. 6. Ifugao rice terraces, Philippines. (*Philippine Embassy, Washington, DC*)

Fig. 7. Land capability classes. Suitable for cultivation: 1, requires good soil management practices only; 2, moderate conservation practices necessary; 3, intensive conservation practices necessary; 4, perennial vegetation—infrequent cultivation. Unsuitable for cultivation (pasture, hay, woodland, and wildlife): 5, no restrictions in use; 6, moderate restrictions in use; 7, severe restrictions in use; 8, best suited for wildlife and recreation. (*USDA*)

drainage, precipitation, wind exposure, and other natural attributes. The objective of such determination is to achieve permanent productive use as nearly as possible. The U.S. Natural Resources Conservation Service has developed a widely employed system for such determination (**Fig. 7**). In this system, eight classes of land are recognized within United States territory. Four classes represent land suited to cultivation, from the class 1 flat or nearly flat land suited to unrestricted cultivation, to the steeper or eroded class 4 land that must be cultivated only infrequently and with extreme care. Three additional classes are grazing or forestry land, with varying degrees of restriction on use. The eighth class is suited only to watershed, recreation, or wildlife support. The aim in the United States has been to map all lands from field study of their capabilities, and to adjust land use to the indicated capability as it becomes economically possible for the farm, range, or forest operator to put conservational use into force.

Once the capability of land has been determined, specific measures of management come into play. For class 1 land, few special practices are necessary. After the natural soil nutrient minerals begin to decline under cultivation, the addition of organic or inorganic fertilizers becomes necessary. The return of organic wastes, such as manure, to the soil is also required to maintain favorable texture and optimum moisture-holding capacity. Beyond these measures, little need be added to the normal operation of cultivation.

On class 2, 3, and 4 lands, artificial fertilization will be required, but special measures of conservational management must be added. The physical conservation ideal is the maintenance of such land under cover for as much of the time as possible. This can be done where pasture and forage crops are suited to the farm economy. However, continuous cover often is neither economically desirable nor possible. Consequently, a variety of devices have been invented to minimize the erosional results from tillage and small grain or row crop growth. Tillage itself has become an increasingly important conservational measure since it can affect the relative degree of moisture infiltration and soil grain aggregation, and therefore erosion. Moreover, forms of conservation tillage that disturb the soil surface as little as possible have become very significant. Tillage machinery that can work the subsurface soil without turning over the topsoil is employed, as are planters that inject seed, fertilizer, and herbicides directly into the subsurface through small holes. Where wind erosion is the danger, straw mulches or row or basin listing (furrowing) may be employed and alternating strips of grass and open-field crops planted. Fields in danger of water erosion are plowed on the contour (not up- and downslope), and if cultivation by lister (a double-moldboard plow) is also employed, water storage capacity of the furrows will be increased. Strip cropping, in which alternate strips of different crops are planted on the contour, may also be employed. Crop rotations that provide for strips of

closely planted legumes and perennial grasses alternating with grains such as wheat and barley and with intertilled strips are particularly effective in reducing soil erosion. Fields may also be terraced, and the terraces strip-cropped.

Design of conservational cultivation also includes provision for grass-covered waterways to collect drainage from terraces and carry it into stream courses without erosion. Where suitable conditions of slope and soil permeability are found, shallow retention structures may also be constructed to promote water infiltration. These are of special value where insufficient soil moisture is a problem at times. Additional moisture always encourages more vigorous cover growth.

The measures just described for class 2, 3, and 4 lands may be considered preventive, but there are also rehabilitation measures where fields have already suffered from erosion but offer possibilities of restoration. Grading with mechanical equipment and the construction of small check dams across former gullies are examples.

For the remaining four classes of land, whose principal uses depend on the continuous maintenance of cover, management is more important than physical conditioning. In some cases, however, water retention structures, check dams, and other physical devices for retarding erosion may be applied on forest lands and rangelands. In the United States such structures are not often found in forest lands, although they have been commonly employed in Japanese forests. In forestry the conservational management objective is one of maximum production of wood and other services while maintaining continuous soil cover. The same is true for grass and other forage plants on managed grazing lands. For rangelands, adjustment of use is particularly difficult, because grazing must be tolerated only to the extent that the range plants still retain sufficient vitality to withstand a period of drought which may arrive at any time.

A last set of erosion-control measures is directed toward minimizing stream bank erosion, which may be large over the length of a long stream. This may be done through revetments, retaining walls, and jetties, which slow down current undercutting banks and hold sand and silt in which soil-binding willows, kudzu, and other vegetation may become established. Sediment detention reservoirs also reduce the erosive power of the current, and catchment basins or flood control storage helps reduce high flows. *See* AGRICULTURE; FOREST AND FORESTRY; LAND-USE CLASSES; RETAINING WALL; REVETMENT; RIVER ENGINEERING.

Irrigated agriculture. The impact of soil management on soil productivity is extremely dramatic when soils are irrigated. Flood irrigation has been utilized in many parts of the world, with stream water being diverted through canals into the furrows of individual plots of land. The enhanced population-support capacity in such regions helped give rise to ancient culture hearths such as those in the Nile and Tigris–Euphrates valleys and to the growth of large population numbers in the wet monsoon floodplains of India and China. Until the twentieth century, flood irrigation was the predominant method of managing water on soils, but with the explosion of technology came another form of water management—center-pivot irrigation—in which water is pumped to the surface from a central well and into a large sprinkler on wheels which is made to rotate slowly around the wellhead. Other, high-technology methods of irrigation have also appeared, such as fertigation in the Negev Desert of southern Israel, where computer-controlled desert water is mixed with fertilizers and then piped directly onto the root system of each crop plant. The twentieth century saw a startling worldwide increase in the total area under irrigation. World Watch Institute in Washington, DC, estimates that the Earth's irrigated acreage at the end of the twentieth century reached 630 million acres (255 million hectares), a global increase of some 35% in just 20 years.

Highly significant is the fact that this increase in irrigated soils is being counterbalanced in part by salinization, where plant growth is somewhat inhibited by accumulating salts both on and below the soil surface. Salinization occurs because all irrigation waters contain at least some dissolved salts that are left behind as the water either transpires or evaporates. Salinization is a serious problem in soil management. Soil conservation on irrigated soils can be effected through the control and even the diminution of salinization. Salinity control and desalinization require that additional amounts of water be applied to soils in order to leach accumulating salts downward beyond the root zone of crops. This process depends on the placing of drainage pipes some 10 ft (3 m) or more below the surface. As desalinization is expensive and poses the risk of soil waterlogging, it is less commonly implemented than most of the other conservation measures discussed. *See* DESALINATION; FLOODPLAIN. Donald J. Patton

Bibliography. J. L. Dodd, Desertification and degradation in sub-Saharan Africa, *BioScience*, 44:28–34, January 1994; G. Gardner, Shrinking fields: Cropland loss in a world of eight billion, *WorldWatch Paper*, no. 131, WorldWatch Institute, Washington, DC, July 1996; I. W. Heathcote, *Environmental Problem-Solving: A Case Study Approach*, McGraw-Hill, New York, 1997; R. L. McConnell and D. C. Abel, *Environmental Issues: Measuring Analyzing and Evaluating*, Prentice Hall, Upper Saddle River, NJ, 1999; D. Pimental et al., Environmental and economic costs of soil erosion and conservation benefits, *Science*, 267:1117–1123, February 24, 1995; J. D. Soule and J. Piper, *Farming in Nature's Image*, Island Press, Washington, DC, 1992.

Soil degradation

Loss in the quality or productivity of soil that is often the result of human activities, such as agriculture practices, deforestation, mining, waste disposal, and chemical spills. Degradation is attributed to changes in soil nutrient status, biota, loss of organic matter,

deterioration of soil structure, and toxicity due to accumulations of naturally occurring or anthropogenic (human-made) materials. The effects of soil degradation include loss of agricultural productivity, negative impacts on the environment and economic stability, and exploitation of marginally suitable or virgin land. Opinions as to the present threat from soil degradation range widely, with estimates of up to 17% of the Earth's 11.5 billion hectares (28.4 billion acres) of vegetated land considered by some to be degraded. Though the actual extent of the problem is unclear, it is certain that soil degradation remains a significant issue, particularly for impoverished regions. *See* SOIL.

Soil biology. Healthy soil is rich with life, particularly microorganisms. Because soil is highly compartmentalized, it offers diverse habitats that can accommodate many microbial lifestyles. Based on deoxyribonucleic acid (DNA) analysis, it appears that some 10,000 species of microorganisms can be present in a single gram (0.04 ounce) of soil. Bacteria are the most numerous microorganisms in soil, whereas fungi represent the largest fraction of the microbial biomass (living mass). Other biota comprising the soil community include algae, protozoa, nematodes, and a variety of larger organisms, such as arthropods and earthworms. Globally, it has been estimated that the approximately 26×10^{28} (260 million billion trillion) prokaryotic microbial cells present in terrestrial environments contain 26, 6.2, and 0.65 trillion kilograms (57, 14, and 1.4 trillion pounds) of the world's carbon, nitrogen, and phosphorus, respectively, rivaling the importance of plants in the global nitrogen and phosphorus budget.

The soil microbial community is essential to the cycling of nutrients and the decomposition of wastes; thus hindrance of microbial activities may have serious ecological results. Fortunately, soil microbial communities are usually resilient and able to recover from most environmental insults, such as deposition of toxic chemicals. However, some terrestrial ecosystems, such as tundra or soils low in nutrient-holding capacity, are more fragile than others and thus recover more slowly from perturbations. Soil microbial activity is generally closely related to the amount of soil organic matter; thus degradative forces such as erosion that deplete organic matter also result in long-term losses of microbial activity.

While microorganisms are the driving force in maintaining soil quality and productivity, agricultural practices may stimulate soil microorganisms that have detrimental effects on humans. Some fertilizer components, such as urea, rapidly undergo microbial oxidation to form nitrate, which is mobile in the environment and causes health concerns, particularly for infants. This problem can be exacerbated by application of fertilizers in the fall, when there is no crop present to take up the nutrients. Administration of feeding antibiotics to livestock as a prophylaxis or to achieve modest gains in animal growth may promote the occurrence of antibiotic resistance among terrestrial organisms, which may serve as an environmental reservoir for resistance genes that could be transferred to microbial pathogens. Excessive concentration of animals may also directly introduce microbial pathogens into surface water or shallow ground water, frequently resulting in soil and water contamination with nutrients, a form of chemical degradation. *See* SOIL ECOLOGY; SOIL MICROBIOLOGY.

Chemical degradation. Failing to replace nutrients removed with crop harvests depletes the soil over the long term. Excessive leaching of the basic cations that buffer soil pH may result in soil acidification, changing the solubility, and thus availability, of certain nutrients to plants. Soil testing is meant to determine the appropriate rate of nutrient addition; however, procedures used in soil testing vary from region to region and may not accurately predict fertilizer needs. In sub-Saharan Africa, soils are becoming depleted due to insufficient fertilization, whereas in some parts of North America excessive fertilization results in contamination of water and perhaps changes in soil function. Even prudent use of fertilizers and liming materials to maintain or enhance productivity may mask the effects of soil degradation resulting from other agricultural practices. It is also possible for land management practices to result in concentrations of chemicals that are toxic to plants. For example, in arid lands salts may accumulate at the surface due to irrigation practices. Other land management practices, such as clear-cutting of forests, may result in long-term changes in nutrient cycling. *See* SOIL CHEMISTRY; SOIL FERTILITY.

Mining. Historically the mining of coal and various metal ores has been among the most devastating forms of land use. Strip (surface) mining can expose large amounts of reduced (decreased oxidation state) minerals in the form of mine tailings and rubble. In the case of coal mining, pyrite (FeS_2) is the source of the problem. A combination of chemical oxidation and microbial oxidation of iron and sulfur in the pyrite releases sulfuric acid (H_2SO_4) and ferric hydroxide ($Fe[OH]_3$) that drain away from the site in storm water. The acidic mine tailings are extremely difficult to revegetate, and without intervention remain exposed, continuing to produce acidic drainage until oxidizable material is depleted. Left untreated, strip mines may require 50 to 150 years for recovery. Reclamation of strip mines involves establishment of a rooting medium, typically by the addition of topsoil. Vegetation is selected for tolerance to acidic conditions. Most developed nations now require complete reclamation of strip mines by responsible parties. *See* AUTOXIDATION; LAND RECLAMATION; SURFACE MINING.

Toxic chemical spills. Hundreds of thousands of different organic compounds are synthesized commercially for a variety of purposes. Many of these compounds eventually find their way into soil ecosystems, often without incident. In some cases, such as petroleum spills, the compounds are toxic to soil organisms when present at high enough concentrations, and thus produce large shifts in microbial community structure, especially in fragile tundra soils.

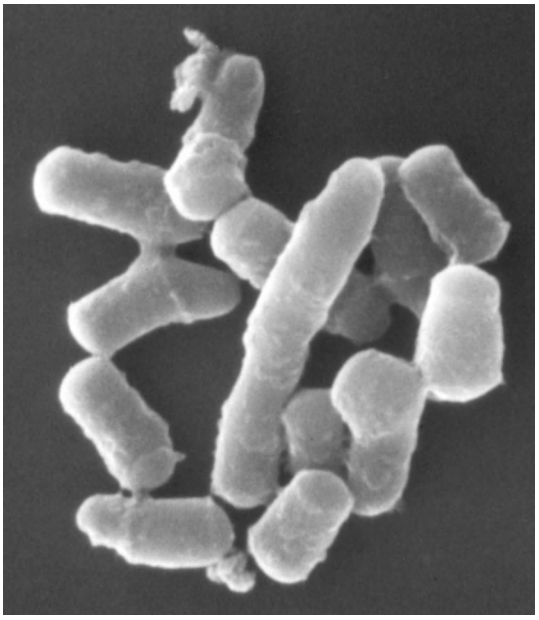

Scanning electron micrograph of a *Rhodococcus* species (bacterium) isolated from an environmental sample. This organism can degrade a diverse array of aromatic pollutants, including quinoline, pyridine, catechol, benzoate, and protocatechuic acid.

Spillage of other toxic chemicals can cause similar long-term effects on microbiology. Eventually, microorganisms present in the soil degrade most organic contaminants, thus alleviating toxic effects. The time required for microorganisms to degrade contaminants depends on both the chemicals and the organisms involved. In some cases, organisms originating from other sites can be used as inocula to enhance degradation of contaminants. For example, a species of the bacterium *Rhodococcus* (see **illustration**) is capable of degrading quinoline, a toxic organic compound associated with synfuel waste water.

Unlike organic contaminants, which are usually degraded to nontoxic forms, toxic metals become essentially permanent features of soils once introduced. Fortunately, adsorption of metals to soil colloids decreases the availability of the metals for movement in the environment or uptake by plants, animals, and microorganisms. The presence of toxic metals provides genetic pressure for the selection of metal-resistant microorganisms, which have been detected at metal-contaminated sites. Metal contamination may be either natural or anthropogenic.

Pesticide use. Pesticides are chemicals designed specifically to kill or suppress targeted groups of organisms, such as weeds, insects, or in some cases microorganisms. As a result, there has been concern that pesticide use will eventually result in changes in the functioning of the soil microbial community. Numerous investigations have been completed with very little evidence for such effects; however, it is possible that some organisms are affected with no measurable impact on the overall function of the microbial community. Until recently, it was difficult to examine the impacts of pesticides on soil microorganisms, because most cannot as yet be cultivated. Molecular biology tools now available allow collection of considerable information about microbial communities without the necessity of cultivating the organisms.

Effects of pesticides on soil microbial communities are likely to be transient. Though most pesticides are xenobiotic (foreign to life) and thus did not exist in nature until recently, there are microorganisms able to degrade most of these substances, sometimes with no prior exposure to them. *See* AGRICULTURAL CHEMISTRY; PESTICIDE.

Physical degradation. Physical effects of soil degradation, such as erosion, are more obvious than other forms of degradation and thus are the most widely documented. Soils vary widely in the relative proportions and organization of the solid and pore spaces (structure). Soil structure, in part, determines the suitability of land for crop production or its vulnerability to erosion. Misuse of soil often results in compaction or densification (loss of pore space). Compaction affects the flow of water through soil and impedes growth of plant roots. Compacted soils generate greater runoff during storms and tend to be sensitive to erosion. Erosion results not only in lost soil productivity but also contamination of lakes and streams with sediment and associated agrochemicals. Based on river sedimentation rates, it appears that human-induced erosion on average roughly equals natural erosion, around 11.2 billion metric tons (11 billion tons) per year on a global basis, though the relative rates may very considerably from place to place. *See* EROSION.

Soil degradation and climate change. Wind erosion of unprotected soil can induce the development of desertlike conditions (desertification). It is apparent that this process was involved in the disappearance of several ancient civilizations, including the Harappan (western India), Mesopotamian (western Asia), and Mayan (Central America) cultures. Desertification continues to be a problem in Africa, Kazakhstan, Uzbekistan, and northern China. Soil degradation can affect not only regional climate, but also global climate. In turn, global climate change, particularly global warming, may promote soil degradation and worsen existing soil problems. Higher temperatures promote microbial decomposition of organic matter, leading to decreased soil productivity and increased susceptibility to erosion. Organic matter decomposition releases carbon dioxide into the atmosphere, exacerbating warming. Changes in rainfall patterns associated with global warming may also promote soil erosion by lengthening dry periods and increasing the intensity of rainy periods. It is expected that deserts will expand because the interiors of continents will likely be drier. *See* CLIMATE MODIFICATION; DESERTIFICATION; GLOBAL CLIMATE CHANGE.

Soil conservation. Certain agricultural practices, such as no-till farming, reduce erosion losses; in the United States, conservation tillage has decreased water erosion by as much as 40%. Terracing and the use of cover crops also promote soil conservation. In

wealthy nations, addition of crop nutrients based on sound and suitable soil test results is the most promising way to reduce the problem of excess fertilization, whereas more affordable fertility practices, such as crop rotations and the use of nitrogen-fixing trees have been successful in poor regions. *See* AGRICULTURAL AND SOIL PRACTICES; COVER CROPS.

Sometimes, practices used to treat soil degradation may actually damage soils in other ways. For example, remediation of soil contamination by incineration or pyrolysis essentially destroys the soil along with the contaminants. Less invasive approaches, such as bioremediation (use of microorganisms or plants to clean up contamination), depend on natural processes to degrade contaminants and generally do not further degrade soil properties. Increased understanding of soil degradation processes has made it technically feasible to avoid many of the practices that damage soil; however, cultural and economic factors continue to promote degradative land-use choices. *See* BIODEGRADATION; SOIL CONSERVATION.

Gerald K. Sims

Bibliography. J. Kaiser, Wounding Earth's fragile skin, *Science*, 304(5677):1616–1618, 2004; R. Lal and B. A. Stewart, *Soil Management Experimental Basis for Sustainability and Environmental Quality*, CRC Press, Boca Raton, FL, 1995; G. W. Olson, Archaeology: Lessons on future soil use, *J. Soil Water Conserv.*, 36:261–64, 1981; C. A. Seybold, J. E. Herrick, and J. J. Brejda, Soil resilience: A fundamental component of soil quality, *Soil Sci.*, 164:224–234, 1999; G. K. Sims, Biological degradation of soil, *Adv. Soil Sci.*, 11:289–330, 1990; W. B. Whitman, D. C. Coleman, and W. J. Wiebe, Prokaryotes: The unseen majority, *Proc. Nat. Acad. Sci. USA*, 95:6578–6583, 1998.

Soil ecology

The study of the interactions among soil organisms, and between biotic and abiotic aspects of the soil environment. Soil is made up of a multitude of physical, chemical, and biological entities, with many interactions occurring among them. Soil is a variable mixture of broken and weathered minerals and decaying organic matter. Together with the proper amounts of air and water, it supplies, in part, sustenance for plants as well as mechanical support.

Soil is predominantly a sand-silt-clay matrix, containing living materials (biomass) including microorganisms, roots, and animals, plus dead organic matter, with varying amounts of gases and liquids. Soils, in addition to the three geometric dimensions, are greatly influenced by the fourth dimension, time.

Soils are the result of several factors, including climate, organisms, parent material, and topography (relief), acting through time (**Fig. 1**). These factors affect major processes, such as primary production, decomposition, and nutrient cycling, which lead to the development of ecosystem properties unique to a soil type. Such characteristics as cation-exchange capacity, texture, structure, and organic matter status are the outcome of these major processes operating as constrained by the controlling factors.

Soil profile. Abiotic and biotic factors lead to certain chemical changes in the top few decimeters (8–10 in.) of soil. In many soils, particularly in more mesic or moist regions of the world, there is a certain amount of leaching and redeposition of nutrients, often accompanied by a distinct color change. Thus, the air-litter surface is at the top of the profile, below are the zones of litter, fermentation, and humification (termed 00, 01, and 02, respectively), and below these is the mineral soil surface, which contains most of the organic matter (A horizon). The B horizon is next, containing deeper-dwelling organisms and somewhat weathered material. This is followed by the C horizon, the unconsolidated mineral material above bedrock.

The work of the soil ecologist is made easier by the fact that the surface 10–15 cm (4–6 in.) of the A horizon has the majority of plant roots, microorganisms, and fauna. A majority of the biological-chemical activities occur in this surface layer; indeed, most of the microorganism- and alga-feeding fauna, such as protozoa and rotifers and tardigrades, are in the top 1–2 cm (0.4–0.8 in.). The continual input of leaf, twig, and root materials, and the algal and cyanobacterial production and turnover make this region prime for biological activity.

The clay fraction in soil, so important to microorganism life and plant activity because of its nutrient content, comprises particles less than 2 micrometers in diameter. Unlike the sand-silt minerals, clays are weathered forms of primary minerals; they are therefore referred to as secondary minerals. Coarse clay particles (0.5 μm) often are derived from quartz and mica; finer clays (0.1 μm) are clay minerals or weathered products of these (such as hydrated ferric, aluminum, titanium, and manganese oxides). The effects of coarse and fine clays on organic matter dynamics have been studied intensively. Labile (that is, rapid turnover) constituents of organic matter may be preferentially adsorbed onto fine clay particles, and thus may be a prime source of energy for the soil microorganisms.

Input of organic matter to soil is one of the major agents of soil structure. The organic matter comes from both living and dead sources (roots, leaves, microorganisms, and fauna). Various physical processes, such as deformation and compression by roots and soil fauna and by freezing-thawing and wetting-drying also have significant influences on soil structure. Plant roots and resistant organic breakdown products (humus) are important in the formation of aggregates. Bacteria, fungi, and their metabolic products are equally important in promoting granulation. The interaction of organic matter and mineral components of soils has a profound influence on cation adsorption capabilities. This property is important to both agrosystem and natural ecosystem management. *See* AGROECOSYSTEM; HUMUS; SOIL CHEMISTRY.

Biological aspects of soil. The biological aspects of soil consist of major organic inputs, decomposition

by primary decomposers (bacteria, fungi, and actinomycetes), and interactions between microorganisms and fauna (secondary decomposers) which feed on them (Fig. 1).

The detritus decomposition pathway occurs on or within the soil after plant materials (litter, roots, sloughed cells, and soluble compounds) become available through death or senescence. Plant products are used by microorganisms (primary decomposers). These are eaten by the fauna, which thus affect flows of nutrients, particularly nitrogen, phosphorus, and sulfur. The immobilization of nutrients into plants or microorganisms and their subsequent mineralization are critical pathways. These processes are diagrammed in **Fig. 2**, showing flows out from the microorganisms and fauna via mineralization or direct losses into organic pools. The labile inorganic pool is the principal one that permits subsequent microorganism and plant existence. Scarcity of some nutrient often limits production. Most importantly, it is the rates of flux into and out of these labile inorganic pools which enable ecosystems to function successfully. *See* ECOSYSTEM; SYSTEMS ECOLOGY.

Model of soil interactions. Conceptual models are useful for envisioning ecosystem processes. The array of microorganisms, fauna, and roots all provide organic material belowground, which may move through the underground ecosystem in defined ways (**Fig. 3**). Pathways of organic matter utilization by primary decomposers include bacteria, actinomycetes, or saprophytic fungi. There is considerable grazing by such bacterial feeders as protozoa, nematodes, tardigrades, and enchytraeids. The final major primary decomposer group is the fungi, a wide array of saprophytic forms and vesicular-arbuscular mycorrhizae which take up nutrients and are mutualistic with the plant roots. The fungal grazers include stylet-bearing nematodes, collembolans, mites, and certain protozoa that are spore grazers. There is a general group of predators such as large predatory mites, nematodes, or centipedes which prey on bacteria-feeders and fungus-feeders and presumably some top predators as well.

While fauna in general account for less than 5% of the total detritus-decomposer respiration and even less biomass (see **table**), their indirect, catalytic role in decomposition is considerably greater in many ecosystems. Such indirect roles include feeding and its effects on microorganism (prey) populations, translocation of nutrients and microorganisms to different locations in the soil profile, and even, in some cases, immobilization of nitrogen and phosphorus in feces or in nests, such as those made by termites.

Earthworms play an important role in mixing and comminuting soil. They ingest large amounts of soil; a single worm such as *Lumbricus terrestris* can ingest several times its own body mass per day in a grassland or pasture. Hence, earthworms have a significant impact on nutrient transformation and availability. In addition to direct ingestion of leaf litter and other organic debris pulled from the soil surface, the earthworms will ingest and generally triturate any

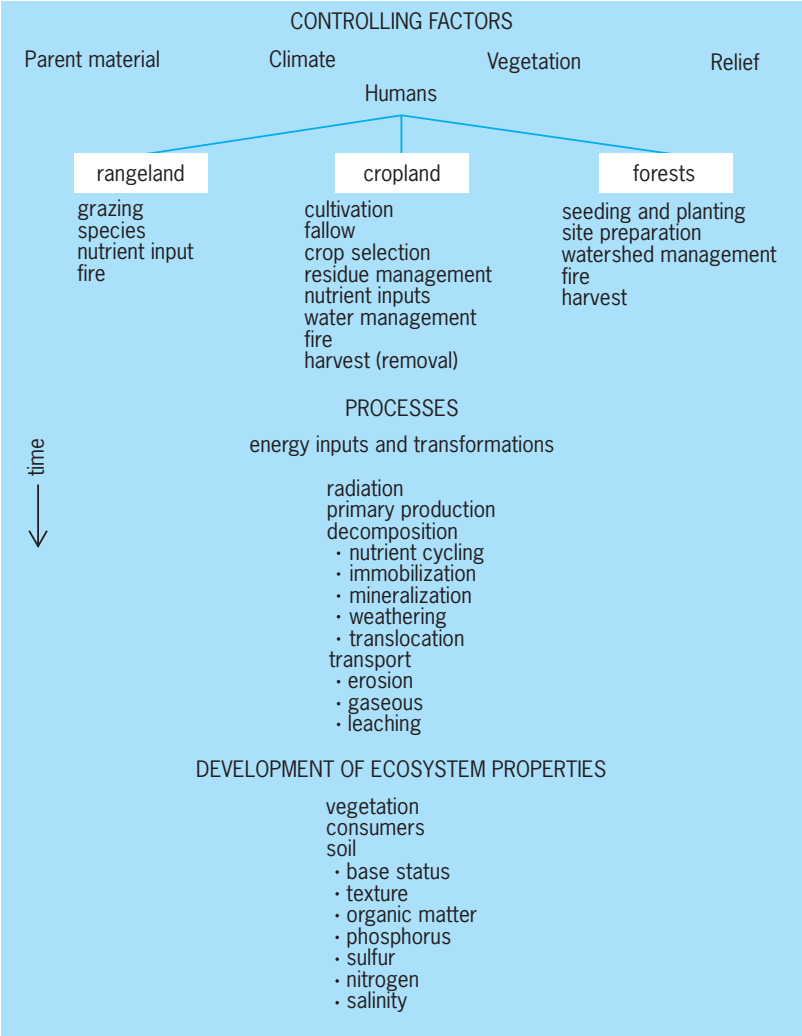

Fig. 1. Factors influencing soil development. Controlling factors affecting processes, over time, influence ecosystem properties. (*From D. C. Coleman, C. P. P. Reid, and C. V. Cole, Biological strategies of nutrient cycling in soil systems, Adv. Ecol. Res., 13:1–55, 1983*)

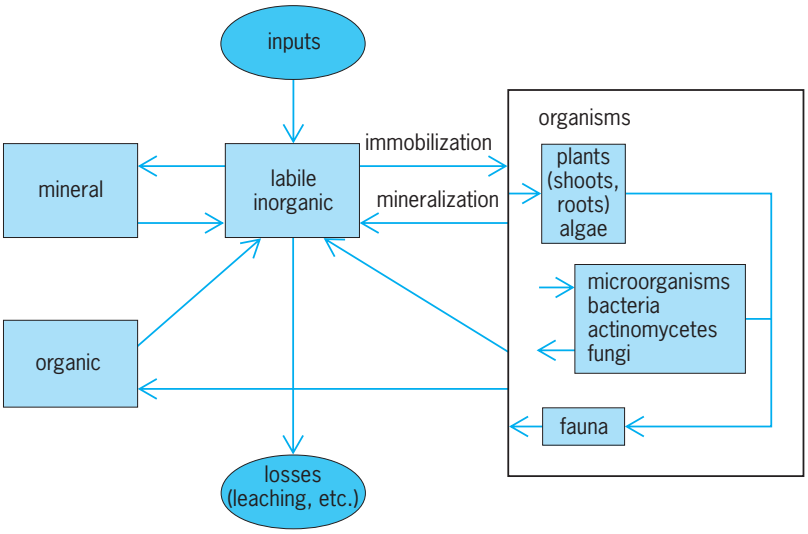

Fig. 2. Generalized nutrient-cycling scheme in soil. The biological activities of immobilization and mineralization are of major importance in ecosystem function. (*From D. C. Coleman, C. P. P. Reid, and C. V. Cole, Biological strategies of nutrient cycling in soil systems, Adv. Ecol. Res., 13:1–55, 1983*)

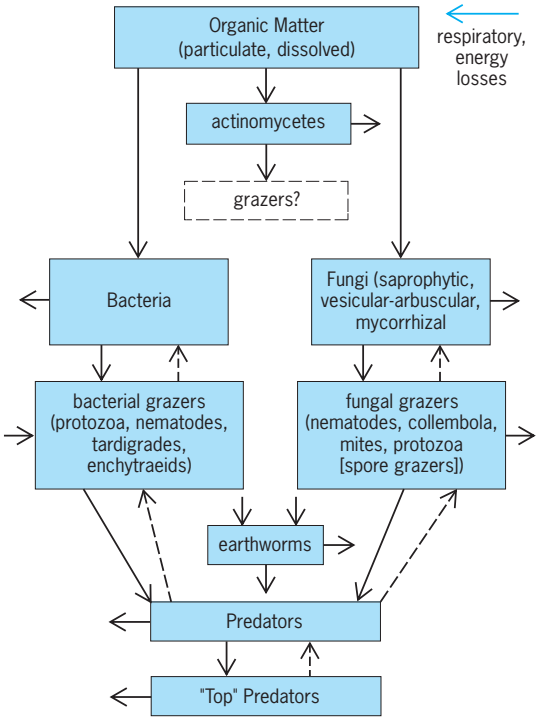

Fig. 3. Conceptual model of breakdown of particulate and soluble organic matter in agroecosystems. Vertical arrows show material flows; horizontal arrows show respiratory energy losses; and broken arrows show information feedback from lower trophic levels.

small bits of organic matter, including protozoa and nematodes. *See* OLIGOCHAETA.

Microorganism–fauna interactions. A number of studies have demonstrated the marked effects of protozoa on both bacterial numbers and nutrient dynamics. Protozoa, among the smallest and most abundant soil fauna, feed upon, and promote turnover of, microorganism populations. This has been shown in various natural ecosystems (forest, grassland, marine) as well as cropped-land ecosystems.

Soil incubations (microcosms) have been used to examine the role of some of the soil fauna in decomposition. The microfauna (protozoa) may be as abundant as 10,000–100,000 per gram of soil. Soil mesofauna, such as nematodes, are larger (about 100 times larger than protozoa), and still numerous (100–1000 per gram of soil). Both groups of organisms are voracious feeders on bacteria and fungi, and account for significant effects on nitrogen and phosphorus mineralization in soil (Figs. 2 and 3). Impacts on the primary decomposer populations vary, as noted below.

Several investigators have cultured fungus-feeding nematodes on various species of fungi and have monitored the fungal and nematode populations. Typically, they observed some inhibition of mycelial growth or respiration in cultures with nematodes versus those without nematodes. Grazing at very low nematode population levels actually stimulates the growth of fungi slightly, which is called a grazing-optimization response. The degree of such a response in natural soil detritus food chains is not known. In several studies, there have been complete shifts of fungal species abundances due to preferential feeding by Collembola (an order of primitive wingless insects) on certain fungal species in aspen woodland and in pine forests. *See* COLLEMBOLA.

Spatially, there are three regions where increased populations of microorganisms and fauna occur: roots, litter, and aggregates. These regions (or microsites) exist at interfaces, or zones where substrates accumulate. Roots produce exudates such as carbohydrates, amino acids, exfoliates (sloughed cells), and mucigel (in conjunction with microorganism activity) as the roots grow through soil. Microorganisms are more abundant in the rhizosphere (the volume 0.08–0.16 in. or 2–4 mm out from the root surface), and microorganism grazer populations (protozoa, nematodes) also are larger in the rhizosphere region. *See* RHIZOSPHERE; ROOT (BOTANY).

Several research groups in North America and Europe have investigated the effects of feeding by detrital food-chain organisms on plant growth. By using either protozoa or nematodes in experimental microcosms with growing plant seedlings, significant increases were observed in both plant shoot growth and nitrogen content when bacteria-feeding

Numbers of soil organisms in various ecosystems at four different sites*

Trophic group	Aspen poplar, Alberta, Canada	Lodgepole pine, Alberta, Canada	Dryland wheat, Nebraska	Mixed-grass prairies, Nebraska
Bacteria per gram soil	—	—	1.3×10^9	1.2×10^9
Fungi, meters per gram soil†	—	—	34	78
Protozoa per gram soil	—	—	37,000	43,000
Microbial biomass carbon, micrograms per gram litter or soil	21,000 (litter)	16,500 (litter)	966 (soil)	1400 (soil)
Nematodes (microbivorous‡) per gram soil	—	—	52	74
Collembola (fungal-feeding) per square meter	71,000	22,000	1300	2000
Mites (fungal-feeding and predatory) per square meter	158,000	135,000	900	1200

*1 g = 0.035 oz. 1 m^2 = 10.7 ft^2. 1 m = 3.28 ft.
†That is, hyphal length is measured in meters.
‡Bacterial and fungal feeding.

fauna were present. The fungus–fungus-feeder food chain, however, showed no significant increases in plant growth when compared with the appropriate control treatments, probably because the fungi are very efficient mineralizers of nitrogen.

Selective removal or inhibition of certain target organisms such as fungi, nematodes, and arthropods has been done by using biocides in the field to determine microorganism–fauna interactions. Fungicides, insecticides, and nematicides have been used singly or in combination in field plots in various field sites, and then decomposition has been measured. Compensatory responses of certain decomposers have been observed; with an inhibition of fungi, bacteria become more abundant and active. The bacteria-feeding fauna then multiply, and predators upon them in turn become more active and enhance mineralization processes. *See* SOIL MICROBIOLOGY.

Mutualistic interactions in soils. There are several root–microorganism interactions in soil which are mutualistic, that is, beneficial for both organisms. Prominent among these are the mycorrhizal associations between certain fungi and plant roots. Unlike the saprophytic fungi mentioned earlier, these fungi are dependent on the host plant for energy-containing carbon compounds. When the thread-like filaments (hyphae) of these beneficial fungi encounter the growing roots, they penetrate the cortical cells, or the spaces between the cells, and establish metabolic connections with the plant. The fungus receives energy compounds from the host, and there is a beneficial exchange of mineral nutrients. The hyphae grow through the soil, outside the rhizosphere zone that has been depleted of nutrients (primarily inorganic phosphorus and nitrogen), and transport them into the host plant. This mutualism is so beneficial to the plant that mycorrhizal plants can sometimes grow 20 or 30 times more rapidly than otherwise comparable nonmycorrhizal plants. In addition, they may be more resistant to drought and better able to repel disease organisms than nonmycorrhizal plants. *See* MYCORRHIZAE.

Within certain groups of plants, such as the legume family, including various species of beans, clover, and alfalfa, there is another important group of symbionts. When a legume root grows through soil near the symbiotic bacteria *Rhizobium*, the two future partners initiate production of several compounds, including complex carbohydrate recognition chemicals, called lectins, that allow the root to recognize the bacteria as beneficial rather than pathogenic. The bacteria then invade the root tissue and form a nodule, a swollen structure on the root surface that houses an active population of bacteria, in which nitrogen fixation takes place. By using photosynthetic energy from the plant and a specialized array of enzymes (nitrogenases), the *Rhizobium* bacteria can incorporate (fix) gaseous molecular nitrogen, N_2, breaking its triple covalent bonds and incorporating it into organic compounds, which are then mineralized and made available to the plant. This biological source of nitrogen frees legumes from a dependence on soil-borne sources and allows them to grow vigorously where many other kinds of plants could not thrive without fertilization. *See* BIOGEOCHEMISTRY; LECTINS; NITROGEN CYCLE.

Agroecosystems versus natural ecosystems. There have been several studies of minimum tillage or no-tillage in agricultural systems and how it affects microorganism populations and associated nutrient dynamics.

Soil microorganism and biochemical changes associated with reduced tillage in soybeans and corn have been studied. Differences were found in microorganism populations which convert nitrogenous compounds to nitrate, or produce gaseous nitrous oxide (N_2O) or N_2, related principally to changes in soil water, organic carbon and nitrogen, and pH when the reduced tillage regimes were imposed. There was generally a higher carbon, nitrogen, and water content in the surface soil (0–7.5 cm or 0–3 in.) under no-till, reflected by higher microorganism populations and enzyme activities. These relations were reversed from 7.5–15 cm (3–6 in.) depth in no-till versus conventional tillage, probably because plowing places crop residues at lower depths.

Interestingly, increased numbers of microarthropods and earthworms were found in zero-tilled versus conventional-tilled fields in several locations, such as Georgia and Kentucky, and in Europe. These increases in some faunal groups are probably due more to the effects of tilling on abiotic physical parameters than to increased food availability in no-till.

Systems studies. It seems important to use some integrative properties of ecosystems, particularly those relating to functional groups or guilds, so that groups of organisms that feed on similar types of substrates can be considered. Nematodes and protozoa feeding on bacteria may have a more related role in ecosystem function than various fungus-feeding nematodes and fungus-feeding mites and Collembola, which, in turn, form their own guilds (Fig. 3). A functional understanding of feeding guilds will enable better understanding of the diverse processes that function in the general framework of decomposition, immobilization, and mineralization. These processes are key focal points in determining dynamics of terrestrial ecosystems. *See* AGRICULTURAL SOIL AND CROP PRACTICES; ECOLOGY; GUILD; POPULATION ECOLOGY; SOIL.

David C. Coleman

Bibliography. D. C. Coleman et al., Interactions of organisms at root/soil and litter/soil interfaces in terrestrial ecosystems, *Agric. Ecosys. Environ.*, 24:117–134, 1988; D. C. Coleman, C. P. P. Reid, and C. V. Cole, Biological strategies of nutrient cycling in soil systems, *Adv. Ecol. Res.*, 13:1–55, 1983; E. T. Elliott and D. C. Coleman, Let the soil work for us, *Ecol. Bull.*, 39:23–32, 1988; E. T. Elliott, H. W. Hunt, and D. E. Walter, Detrital foodweb interactions in North American grassland ecosystems, *Agric. Ecosys. Environ.*, 24:41–56, 1988; K. Killham, *Soil Ecology*, 1994; D. Parkinson, Linkages between resource availability, microorganisms and soil invertebrates, *Agric. Ecosys. Environ.*, 24:21–32, 1988; R. E. Phillips, *No-Tillage Agriculture Principles*, 1984.

Soil fertility

The ability of a soil to supply plant nutrients. Sixteen chemical elements are required for the growth of all plants: carbon, oxygen, and hydrogen (these three are obtained from carbon dioxide and water), plus nitrogen, phosphorus, potassium, calcium, magnesium, sulfur, iron, manganese, zinc, copper, boron, molybdenum, and chlorine. Some plant species also require one or more of the elements cobalt, sodium, vanadium, and silicon. *See* PLANT GROWTH.

Nutrient forms in soil. While carbon and oxygen are supplied to plants from carbon dioxide in the air, the other essential elements are supplied primarily by the soil. Of the latter, all except hydrogen from water are called mineral nutrients.

Only part of the 13 essential mineral nutrients in soil are in a chemical form that can be immediately used by plants. The unusable (unavailable) parts, which eventually do become available to plants, are of two kinds: they may be in organic combination (such as nitrogen in soil humus) or in solid inorganic soil particles (such as potassium in soil clays). The time for complete decomposition and dissolution of these compounds varies widely, from days to hundreds of years.

The quantity of the immediately available nutrient is often small. For example, it is common for a soil to contain only 30 lb of available nitrate nitrogen per acre (33 kg/ha) in the top foot (1 ft = 0.3 m) of soil. The same quantity of soil might contain 4000 lb (1800 kg) of unavailable nitrogen in organic compounds (soil humus). As another example, a soil might contain 25,000 lb of total potassium per acre (28,000 kg/ha), with only 300 lb/acre (336 kg/ha) immediately available to crops.

Limits to plant growth. Soils exhibit a variable ability to supply the mineral nutrients needed by plants. This characteristic allows soils to be classified according to their level of fertility. Fertility levels can vary from a deficiency to a sufficiency, or even toxicity (too much), of one or more nutrients. A serious deficiency of only one essential nutrient can still greatly reduce crop yields. For example, a soil considered fertile in all other ways could have levels of available potassium too low for optimum plant growth. Control of crop growth by the deficient nutrient is known as Liebig's law of the minimum. This law states that the plant growth rate is controlled by that mineral nutrient present in the most limiting quantity in the soil (**Fig. 1**). Modern chemical fertilizers applied to the soil reduce or eliminate nutrient deficiency and ensure that fertility does not limit a soil's potential to produce crops. *See* FERTILIZER.

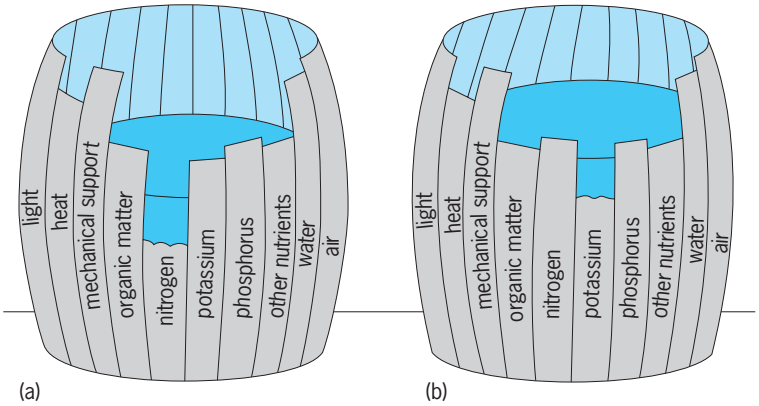

Fig. 1. An illustration of the principle of limiting factors. The level of water in the barrels above represents the level of crop production. (*a*) Nitrogen is represented as being the factor that is most limiting. Even though the other elements are present in more adequate amounts, crop production can be no higher than that allowed by the nitrogen. (*b*) When nitrogen is added, the level of crop production is raised until it is controlled by the next most limiting factor, in this case, potassium. (*After N. C. Brady, The Nature and Properties of Soils, 9th ed., Macmillan, 1984*)

Although fertilizers have dramatically increased the soil's potential to produce crops, other variables often control crop production. Both climatic and other soil factors can affect crop yields, as can crop variety, diseases, and insects. Elements of the climate such as insufficient water, temperature stresses (either too hot or too cold), and natural disasters such as wind and hail storms can all lower crop production regardless of a soil's inherent or cultivated fertility. Other undesirable soil conditions such as compacted soil layers, soil salinity, or acid soil layers that restrict root growth can also reduce crop yield.

Influence of crop type. The type of crop grown on a particular soil determines the required soil fertility level. Crops differ greatly in their need for nutrients and their ability to extract them from soils. Also, even if two crops extract a nutrient equally, one crop may use that nutrient more efficiently than the other.

Perhaps the most extreme example of the crop-type influence would be to compare a soil's nitrogen fertility for a legume such as alfalfa with its nitrogen fertility for a grass such as maize. Since alfalfa, through its symbiotic association with *Rhizobium* bacteria, obtains nitrogen from the dinitrogen (N_2) in the air, a severely nitrogen-deficient soil would not limit the growth of alfalfa. The growth of maize would be slowed by insufficient nitrogen, however, since maize depends directly on the uptake of available nitrogen in the soil. *See* NITROGEN CYCLE; SOIL MICROBIOLOGY.

Cation exchange and inherent fertility. Several soil properties are important in determining a soil's inherent fertility. One property is the adsorption and storage of nutrients on the surfaces of soil particles. Such adsorption of a number of nutrients is caused by an attraction of positively charged nutrients to negatively charged soil particles (similar to the attraction of the positive and negative poles of two magnets). This adsorption is called cation exchange (adsorbed cations can be exchanged with other cations in solution), and the quantity of nutrient cations a soil can adsorb is called its cation-exchange capacity. A cation is a single atom or a group of two or more atoms combined that has lost one or more of its electrons. Since atoms are made of equal numbers of protons (each has one positive charge) and electrons (each with one negative charge), they have zero charge. Some atoms and molecules lose one or more electrons, resulting in an excess of protons, thereby causing a

positive charge. The plant nutrient calcium is a cation with two more protons than electrons, giving it two positive charges (Ca^{2+}). Atoms or groups of atoms having more electrons than protons are negatively charged ions known as anions. *See* ADSORPTION; ION EXCHANGE; SOIL CHEMISTRY.

The negative charge in soils is associated with clay particles. This negative charge results largely from an imbalance of positive and negative charges from cations and anions within their crystalline structure. The negatively charged clay particle adsorbs enough cations to maintain an overall zero charge balance. The cations adsorbed by clay particles are largely the plant nutrients calcium (Ca^{2+}), magnesium (Mg^{2+}), and potassium (K^{+}) with lesser amounts of the essential micronutrients manganese (Mn^{2+}), zinc (Zn^{2+}), iron (Fe^{2+} and Fe^{3+}), and copper (Cu^{2+}). Sodium (Na^{+}) can also be adsorbed in significant quantities, especially in soils of arid and semiarid regions. The cations Ca^{2+}, Mg^{2+}, K^{+}, and Na^{+} are often referred to as a soil's exchangeable bases. The adsorbed cations are in equilibrium with the same cations in the soil water (soil solution). Plant roots absorb most nutrient ions directly from the soil solution. As nutrient cations are taken from the soil solution, exchange reactions resupply the solution with that ion from the soil clays to maintain a nearly constant concentration.

In addition to the exchangeable bases and other nutrient cations, acid cations such as trivalent aluminum, hydroxy-aluminum, and the hydrogen ion are also adsorbed by the soil's negative charges. *See* CLAY MINERALS.

Organic matter and cation exchange. Some of the soil's cation-exchange capacity may also arise from organic matter (humus) in the soil. Negative charges of organic matter arise largely from carboxylic and phenolic acid functional groups. Since these functional groups are weak acids, the negative charge from organic matter increases as the soil pH increases, as shown below in equilibrium for car-

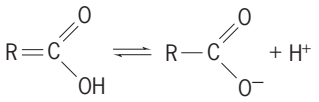

boxylic acid. As soil pH increases, hydrogen ion removal from the right side of the equilibrium causes the formation of additional negatively charged carboxyl groups. *See* CHEMICAL EQUILIBRIUM; HUMUS.

The negative charges of soil organic matter can adsorb the same cations as described for the soil clays. The proportion of cation-exchange capacity arising from mineral clays and from organic matter depends on the proportions of each in the soil and on the kinds of clays. In most mineral soils, the soil clays constitute the greater proportion of cation-exchange capacity. Within the class of mineral soils, those soils with more clay and less sand and silt have the greatest cation-exchange capacity.

The soil cation-exchange sites provide the immediate store of available nutrient cations that can be used by plants. Because of this store of available nutrients, those soils with greater cation-exchange capacities generally provide a greater supply of available nutrients for crop production. However, the fertility level of each plant nutrient must be balanced to the needs of the crop to achieve high productivity.

Soil acidity. The amounts and kinds of acid cations adsorbed on the cation-exchange sites can have a substantial influence on a soil's perceived fertility. Soils with a pH of 7 have their exchange sites nearly 100% saturated with the bases Ca^{2+}, Mg^{2+}, K^{+}, and Na^{+}. As soil pH's drop below 7, the proportion of the exchange complex saturated with acid cations begins to increase. Between pH 7 and 5.5, the predominant acid cation may be hydroxy-aluminum, which is tightly adsorbed to the clays and causes no reduction in plant growth. At soil pH's around 5.5 and below, the acid cation trivalent aluminum (Al^{3+}), which is toxic to plants, begins to appear on the cation-exchange sites. It and excess manganese (Mn^{+2}) together are the greatest hazards to the growth of healthy crops at low soil pH's. In some cases of extremely low soil pH, insufficient calcium may also limit crop growth, but such cases are rare. *See* PH.

Two factors cause soils to become acid: when crops are harvested, exchangeable bases are removed as part of the crop; and exchangeable bases move with drainage water below the crop's root zone (leaching). Leaching of bases is enhanced when acids are released to the soil from plant roots, from microbial respiration, and from other microbially mediated processes, especially nitrification, the conversion of ammonium nitrogen to nitrate nitrogen. For each ammonium ion converted to nitrate nitrogen, two hydrogen ions are released. Hydrogen ions or aluminum cations resulting from acid decomposition of clays can exchange some of the bases from cation-exchange sites to the soil solution to allow their loss by leaching. Since much of the nitrogen fertilizer supplied to crops contains ammonium nitrogen (this is true of both manure and chemical fertilizers), the addition of high rates of nitrogen enhances soil acidification. Finely ground limestone (calcium carbonate) can be applied to acid soils when needed to neutralize soil acids and to add Ca^{2+} and Mg^{2+}, thereby raising soil pH back to the desirable range for optimum crop production. *See* ACID AND BASE.

Nutrients in soil minerals. While the supply of some plant nutrient cations (especially calcium, magnesium, and potassium) is related to their amounts on a soil's cation-exchange sites, there is often a greater supply of these nutrients within the soil minerals themselves. However, nutrients from soil minerals become available to crops only after the minerals partially dissolve and release the individual ions to the soil solution. Most soil clays contain some potassium and magnesium within their crystalline structures. Feldspars in some soils contain potassium, while other minerals contain supplies of magnesium and calcium. Calcite in some high-pH soils constitutes a large supply of calcium. Other minerals supply many of the micronutrients in soil.

Minerals containing some micronutrients and, especially, the major nutrient phosphorus furnish

plant-available forms by slow dissolution. As plant roots absorb these nutrients, more of the soil mineral dissolves in the soil solution. Many micronutrients also have some immediately available forms adsorbed onto the cation-exchange sites, but phosphorus does not since its plant-available forms ($H_2PO_4^-$ and HPO_4^{2-}) are both anions.

Organic matter and nutrient supply. Most of the nitrogen and sulfur in soils are a part of soil organic matter. Soil organic matter can also contain phosphorus and micronutrients, although the major portions of these nutrients are components of soil minerals. Soils in the midwestern United States typically contain around 3000 lb of organic nitrogen per acre (3300 kg/ha) in their plow layer. During 1 year, less than 0.5% of the organic nitrogen might be converted to the plant-available ammonium and nitrate forms by soil microorganisms. The amount of total organic sulfur in soil and the amount converted to plant-available sulfur would both be approximately 10% of the nitrogen values. The amounts of mineralization (conversion from organic to available mineral forms) per year are only approximate, since mineralization depends on many factors, including soil properties, soil temperature, soil moisture, and the quality and quantity of decomposing crop residue. These factors can vary considerably from one location to another and from year to year at a given location.

Soil organic matter, especially the small proportion of soluble organic matter, can increase the supply of micronutrients available to plants by acting as a natural complexing agent. Complexing agents are large organic molecules that can encapsulate some micronutrients. This encapsulation maintains higher levels of that nutrient in the soil solution, increasing its availability for plant uptake. *See* COORDINATION COMPLEXES; MOLECULAR RECOGNITION.

Mineralization versus fertilizer use. Those nutrients that are stored primarily in soil organic matter (nitrogen and sulfur) may be mineralized in insufficient quantity for maximum production of organically grown crops. The amounts of nitrogen and sulfur required by crops are proportional to the yield of crop. For example, the adequate nutrition of corn generally requires 1.6 lb (0.7 kg) of nitrogen in the entire corn crop (roots, stalks, and grain) for each bushel of corn produced. A high-yielding corn crop of 200 bu/acre (80 bu/ha) would therefore require 320 lb (145 kg) of nitrogen per acre (200 bu/acre × 1.6 lb N/bu). Wheat requires about 2.5 lb (1.1 kg) of nitrogen in the entire crop for each bushel produced. A wheat crop of 70 bu/acre (28 bu/ha) would therefore require 175 lb N/acre (193 kg/ha). Other crops have their own specific nitrogen requirements that are proportional to their expected yield. The crop requirements for sulfur can be calculated similarly, except that the amounts required are about 6% of those for nitrogen. The expected yield would depend on other plant growth factors such as the anticipated rainfall or irrigation, the availability of other soil nutrients, or expected losses from pests.

The amount of nitrogen mineralized from soil organic matter to supply the crop's requirements for plant-available nitrogen depends on soil properties such as organic matter and clay concentrations and on environmental conditions such as temperature and the availability of water. The amounts of available nitrogen released from soil organic matter during the cropping season under typical agronomic practice of continuous grain crop production have been carefully measured in a limited number of studies and have been found to vary from 15 to 130 lb/acre (16.5 to 143 kg/ha), which is insufficient for high yields of corn or wheat. Following a nitrogen-fixing legume crop such as alfalfa, or immature legume or grass cover crops (often called green manure crops) that provide crop residues rich in nitrogen, the total amounts of nitrogen mineralized may be in the range of 115 to 230 lb/acre (126.5 to 253 kg/ha). Crop residues from green manure crops that are added to soil will decompose, and some of the green manure crop's nitrogen will mineralize, adding to the available nitrogen released from the soil organic matter alone. In these cases, no additional nitrogen fertilizer is usually needed when corn is grown following a good stand of alfalfa or a rich green manure crop, whereas nitrogen fertilizer is needed when corn follows another crop of corn. Therefore, when using organic crop production practices, it may be beneficial to include legumes or other green manure crops in rotation, or to add animal manures to enhance nitrogen availability and to ensure maximum crop production.

David E. Kissel

Soil testing. Soil testing is the use of rapid chemical analyses to determine the fertility status of a soil. It is used to identify those nutrients or substances that are present in either insufficient or excessive quantities for optimum plant growth. Analyses are also used to monitor increases or decreases in soil fertility over time. Soil test data are critical for determining rates of commercial fertilizer, animal wastes, and other nutrient sources that maximize profit to farmers that are environmentally responsible. Soil testing is also used by commercial greenhouse operators and by private individuals for lawn and garden management.

The first step for the user is collecting a representative sample from the field or area of interest. This process is of vital importance for the resulting soil tests to be meaningful. Attention must be given to the number of cores collected, the core locations, depth of sampling, time of sampling, and how the samples are handled after collection. The samples are usually submitted to a soil testing laboratory that will perform the tests and normally will provide an interpretation of the results, including fertilizer recommendations.

Another approach to soil sampling is zone or grid sampling. The objective is to generate a map of the levels of the soil parameter of interest rather than just an average for the entire field. Computerized fertilizer spreaders that obtain positioning information from satellites and can vary rates as they move across the field are used to apply fertilizer according to the mapped soil test levels.

Most of the laboratory procedures used for evaluating soil fertility involve use of specific chemical solutions to extract a portion of one or more nutrients from the soil. The solutions and procedures are

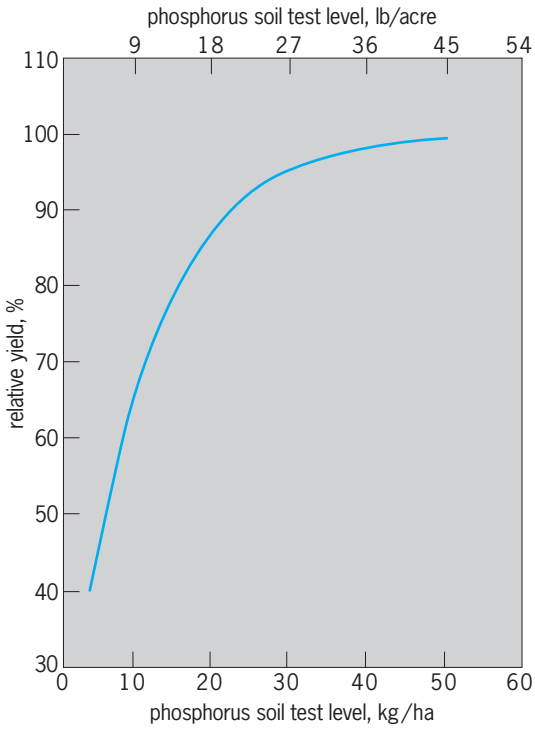

Fig. 2. Relationship between phosphorus soil test level and average relative corn yield or percent sufficiency in Illinois. (*After T. R. Peck et al., eds., Soil Testing: Correlating and Interpreting the Analytical Result, Amer. Soc. Agron. Spec. Publ. 29, 1977*)

selected from a database of correlation studies that have been conducted previously. These studies determine the correlation of the chemically extracted nutrient with the soil's ability to supply the nutrient to plants. The amount of the element present in the soil extract is determined by analytical instruments, which may be as basic as a colorimeter or specific ion electrode or as sophisticated as a plasma spectrophotometer capable of simultaneous multielement analysis.

Due to great diversity in soil chemical properties, several different laboratory procedures have been developed for each nutrient. The procedure must be appropriate for the soils in question or test results will be erroneous. For example, acid-based phosphorus extractants such as weak mixtures of hydrochloric acid and ammonium fluoride work well on acid or neutral soils. However, when used on alkaline soils containing calcium carbonate, the acid is often neutralized before it can extract any phosphorus, creating a false low soil test reading. Alkaline extractants, such as those containing sodium bicarbonate, were developed specifically for soils containing calcium carbonate. Soil testing laboratories determine which procedure to run, based on the properties of the soils being analyzed.

Indices of the soil nutrient availability are developed from the soil test values obtained from these determinations; they have practical significance only when the interpretation is based on extensive field calibration. In soil test calibration, researchers conduct many experiments across the geographic region of interest and across a range of soil test values. Each experiment includes rates of the nutrient in question such that the optimum rate for the site can be determined. Once a sufficient number of experiments have been performed, the probability of getting a response to application of a nutrient and the average percent sufficiency at a given soil test level can be determined (**Fig. 2**). The same studies are used to estimate the amount of a particular nutrient to apply to give optimum yield at a given soil test level.

The most subjective part of soil testing is the process of deriving specific fertilizer recommendations from soil test values. A multitude of factors can influence this process, including land tenure, supply of capital, fertilizer and crop prices, crop quality effects, crop rotation, and fertilizer placement. Therefore, the recommendations developed are dependent on the assumptions made with respect to these factors. Ideally they are fine-tuned to fit each individual situation. *See* AGRICULTURAL SOIL AND CROP PRACTICES; ANALYTICAL CHEMISTRY; PLANT MINERAL NUTRITION; SOIL.

Paul Fixen

Bibliography. C. A. Black, *Soil Fertility Evaluation and Control*, CRC Press, Boca Raton, FL, 1993; N. C. Brady, *The Nature and Properties of Soils*, 10th ed., Macmillan, New York, 1990; J. R. Brown (ed.), *Recommended Chemical Soil Test Procedures for the North Central Region*, Missouri Agricultural Experiment Station, Columbia, 1998; M. R. Carter (ed.), *Soil Sampling and Methods of Analysis*, CRC Press, Boca Raton, FL, 1993; R. H. Follett, L. S. Murphy, and R. L. Donahue, *Fertilizers and Soil Amendments*, Prentice Hall, Englewood Cliffs, NJ, 1981; J. L. Havlin et al., *Soil Fertility and Fertilizers: An Introduction to Nutrient Management*, Prentice Hall, Upper Saddle River, NJ, 1998; J. L. Havlin and J. S. Jacobsen (eds.), *Soil Testing: Prospects for Improving Nutrient Recommendations*, Soil Science Society of America, Madison, WI, 1994; T. R. Peck (ed.), *Soil Testing: Correlating and Interpreting the Analytical Results*, American Society of Agronomy, Madison, WI, 1977; R. L. Westerman (ed.), *Soil Testing and Plant Analysis*, rev. ed., Soil Science Society of America, Madison, WI, 1990.

Soil mechanics

The science that applies laws and principles of physics, especially continuum and particulate mechanics and hydraulics of porous media, to solve problems encountered in civil engineering that are related to force-deformation behavior of soil and the flow of fluids through soil. Civil engineers use the term soil for a naturally occurring multiphase material consisting of gas or liquid and solid particles, whether or not the material supports plant life. The term geotechnics or geotechnical engineering refers to a subspecialty of civil engineering that combines the disciplines traditionally known as soil mechanics, soil dynamics, rock mechanics, earthwork engineering, and foundation engineering for the design or analysis of natural or constructed facilities of, in, or on rock or soil. *See* ROCK MECHANICS.

Fundamental principles. Soil mechanics rests on five fundamental relationships. The effective stress

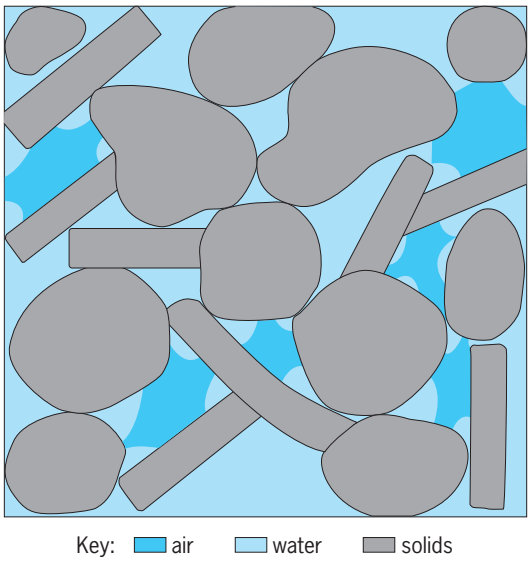

Fig. 1. Diagram showing the three phases in a thin section of a soil element.

principle formulated by K. Terzaghi in the early 1920s in relation to work on the consolidation of clays, is given by Eq. (1), where σ = total normal

$$\sigma = \overline{\sigma} + u \tag{1}$$

stress, $\overline{\sigma}$ = effective normal stress, and u = pore water pressure. The effective stress principle unifies and extends the applicability of other fundamental relationships such as Darcy's law, Eq. (2),

$$Q = kiA \tag{2}$$

where Q is flow rate, k permeability, i hydraulic gradient, and A cross-sectional area; Coulomb friction, Eq. (3), where τ is shear stress, σ_n normal stress,

$$\tau = \sigma_n \tan \phi \tag{3}$$

and ϕ friction angle; Newton's second law of motion, Eq. (4), where F is force, m mass, and a acceleration; and Hooke's law, Eq. (5), where σ is stress,

$$F = ma \tag{4}$$

$$\sigma = E\epsilon \tag{5}$$

E modulus of elasticity, and ϵ strain. Most of the advances that have been made in soil mechanics use these fundamental principles as their starting point. *See* FRICTION; HOOKE'S LAW; NEWTON'S LAWS OF MOTION.

Nature of soil. Soil consists of a multiphase assemblage of mineral or organic particles, as shown in the diagram in **Fig. 1**. The voids within this solid soil skeleton contain gas (generally air), liquid (generally water), or both. A saturated soil has its voids completely filled with water. Various mathematical equations are used to determine the relationships that exist among the three phases within a soil element, an arbitrarily small volume within a soil mass.

When an element of dry soil experiences a load (force F in **Fig. 2**), normal (N) and tangential (T) forces develop immediately at the contact points of the soil particles. The soil mass deforms as the air in the voids compresses or escapes. If the soil voids contain water (Fig. 2*b*), an essentially incompressible liquid, the spring–dashpot analogy helps visualize the load transfer process. Upon application of the load, the pore fluid (water in chamber) cannot compress, and the small size of the soil pores (valve orifice) restricts the escape of water. The pressure in the pore water increases to counteract the applied force F. Immediately after time $t = 0$, water slowly flows out through the tortuous path formed by the microscopic pores. The soil skeleton (spring) begins to feel an increased load, deforming as a load transfer from the pore water to the soil skeleton occurs. After sufficient time has elapsed, pressure within the pore water returns to zero, flow through the soil pores ceases, and the system reaches an equilibrium position, with a compressed soil skeleton resisting the applied force. The applied force imposes a stress F/A (force F per unit area A) on the system, and a strain d/L in the vertical direction (deformation d per unit

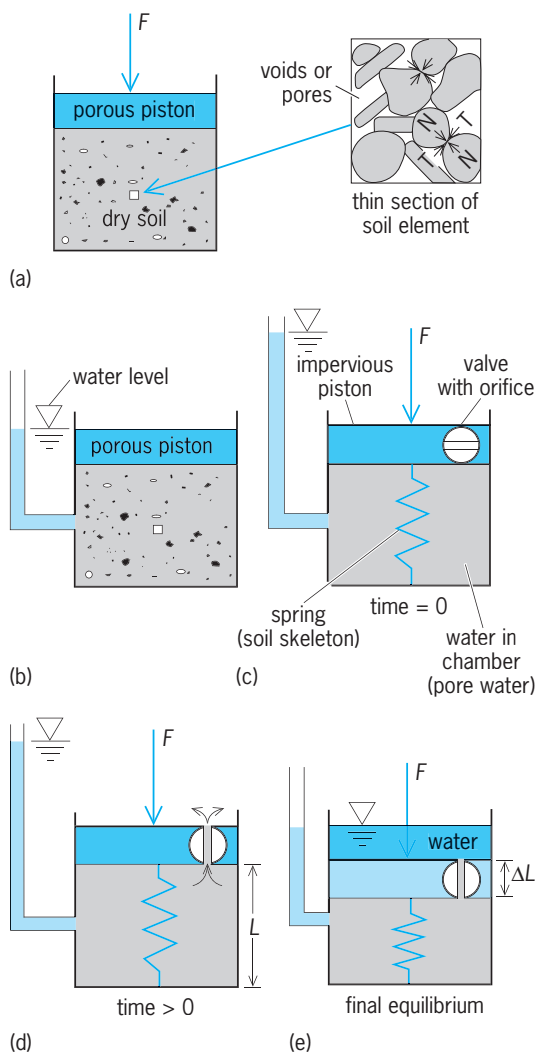

Fig. 2. Load transfer and consolidation; the effects of the application of a load (F) to soil. (*a*) Dry soil (N = normal force; T = tangential force). (*b–e*) Soil containing water and the corresponding spring and dashpot analogy from time = 0 until establishment of equilibrium.

length L in a given direction) in the equilibrium position. Geotechnical engineers consider stresses and strains positive when compressive. A system with large pores reaches equilibrium quickly, while a system with small pores requires a long time to reach equilibrium.

The spring–dashpot system exhibits linearly elastic behavior. Doubling the force F doubles the deformation. Upon removing the force F from the spring–dashpot system, the compressed spring pushes the piston back to its original position as water flows back into the chamber. In soils, crushing and slippage of particles results in nonlinear behavior and unrecoverable deformations that prevent a total rebound of the system to its original position. This nonlinear inelastic behavior means that soils do not obey Hooke's law relating stress to strain.

Since the early days of modern soil mechanics, Hooke's law served as a benchmark to evaluate how far soil behavior deviated from its simple generalization. Complex nonlinear stress–strain relationships attempt to model more closely the deformation characteristics of soil, but linear elasticity continues to provide a useful and sufficiently accurate simplification for the solution of many engineering problems.

Soil classification and index properties. The discipline of soil mechanics relies on the quantification of observed soil behavior. The geotechnical engineer utilizes soil classification to organize soils that exhibit similar behavior into groups. The **table** shows a simplified soil classification based on the widely accepted Unified Soil Classification System (USCS). Simple index tests such as grain size and plasticity help classify soils.

Among the most common index tests, the Atterberg limits that use the water content at the boundary between states of consistency to correlate soils. The water content of a remolded soil at which a $^1/_2$-in. (13-mm) section of a $^5/_{64}$-in.-wide (2-mm) groove closes from 25 successive shocks in a standard device corresponds to the liquid limit. In effect, the test measures the water content required for an arbitrary but constant shear strength of the remolded soil sample. Therefore, all soils have approximately the same shear strength at the liquid limit. The plastic limit corresponds to the water content at which a thread of soil with a diameter of $^1/_8$ in. (3.2 mm) crumbles when pressed together and rerolled. The shrinkage limit corresponds to the minimum water content required to completely fill all the voids of a dried soil pat (or similarly, the water content at which further drying will not cause additional shrinkage of the sample).

Correlations between fundamental soil properties (such as strength, permeability, compressibility, and expansivity) and index tests help the engineer obtain approximate solutions. The actual soil properties measured in laboratory or field tests can vary greatly from the values suggested by the index property correlations.

Pore pressures and heads. Water in the soil pores exists at a measurable pressure—positive, zero, or negative. Zero pore pressure corresponds to atmospheric pressure in the pore water. **Figure 3** shows the relationship between pore pressure and head. An instrument known as a piezometer measures total head near its sensor. The elevation of the piezometer sensor equals the elevation head, h_e. The dimensions of a piezometer sensor generally do not exceed 4 in. (10 cm) in diameter and 3.3 ft (1 m) in length. The total head, h_t, equals the elevation to which water rises in a piezometer. The height of water in the piezometer equals the pressure head, h_p. The pore pressure, u, is given by Eq. (6), where γ_w equals the

$$u = h_p \cdot \gamma_w \tag{6}$$

unit weight of water. Wells can have long collection

A simplified soil classification*

Material	Symbol	Grain size	USCA symbol	Name
Coarse-grained				
Boulder	None	12 in. + (30 cm +)		
Cobbles	None	3–12 in. (7.6–30 cm)		
Gravel	G		GW	Well-graded gravel
Coarse		$^3/_4$–3 in. (1.9–7.6 cm)	GP	Poorly graded gravel
Fine		No. 4, $^3/_4$ in. (1.9 cm)	GM	Silty gravel
			GC	Clayey gravel
Sand	S		SW	Well-graded sand
Coarse		No. 10–No. 4	SP	Poorly graded sand
Medium		No. 40–No. 10	SM	Silty sand
Fine		No. 200–No. 40	SC	Clayey sand
Fine-grained		Passing No. 200		
Silt	M		ML	Silt
			MH	Elastic silt
Clay	C		CL	Low-plasticity clay
			CH	High-plasticity clay
Organic	O		OL	Organic silt
			OH	Organic clay
			PT	Peat

*U.S. Standard Sieve Sizes (partial listing): 12 in. (300 mm); 3 in. (75 mm); No. 4: 0.2 in. (4.75 mm); No. 10: 0.08 in. (2.00 mm); No. 40: 0.02 in. (0.425 mm); No. 200: 0.003 in. (0.075 mm).

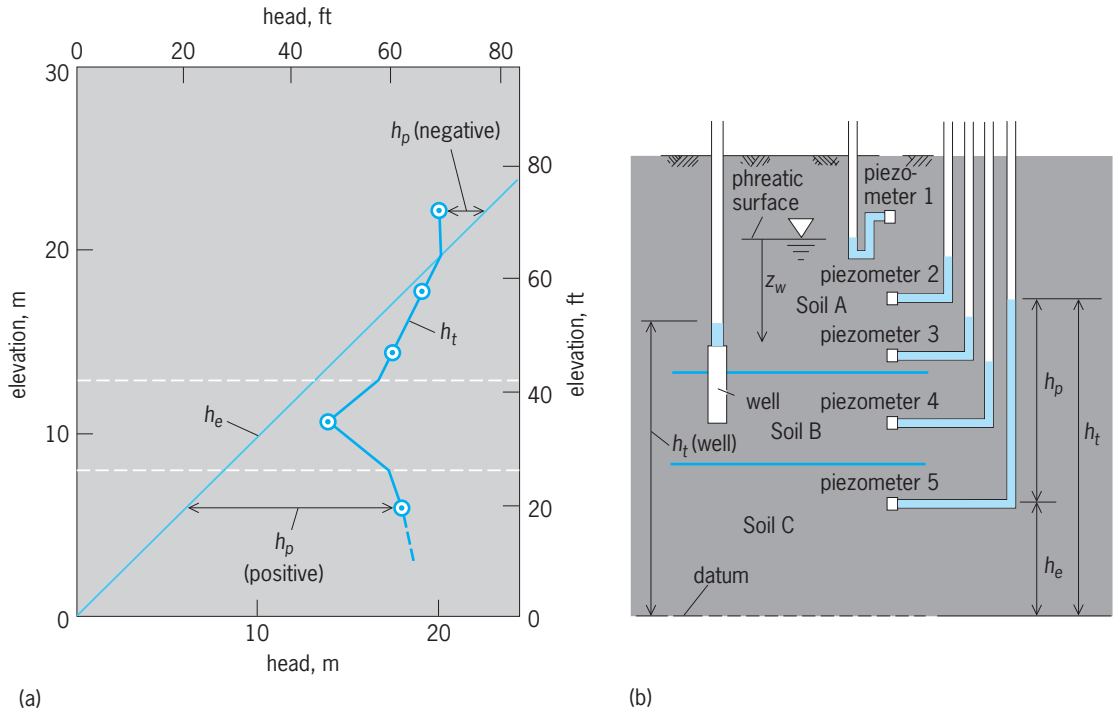

Fig. 3. Relationship between pore water and heads. (*a*) Relationship between heads in multilayered soil. (*b*) Placement of the piezometers and well [h_e = elevation head; h_p = pressure head (pore pressure); h_t = total head; z_w = depth below the phreatic surface (water table)].

zones that penetrate different soil layers. This characteristic makes total head data from wells difficult to interpret and use, as shown in Fig. 3*b*.

Positive pore pressures decrease effective stresses, which results in increased deformations and decreased soil shearing resistance. In general, the higher the pore pressure, the higher the possibility of encountering unsatisfactory performance.

For hydrostatic conditions, the total head remains constant everywhere, no flow occurs and the pore pressure, u, is given by Eq. (7), where z_w equals

$$u = z_w \cdot \gamma_w \tag{7}$$

the depth below the phreatic surface, the imaginary line where $h_p = 0$. Most soil deposits do not exist under hydrostatic conditions. A steady flow of water through the soil results from a constant pattern of pore pressures, known as the steady-state pore pressures or heads. Pore pressures different from steady-state values, known as excess pore pressures, result from loading, unloading, rainfall, increase in river or reservoir levels, leaks from pipelines or tanks, and drying of the soil, among other causes.

Flow. Total heads determine the direction of flow in soil. Water flows from a zone of higher total head toward one of lower total head. Darcy's law [Eq. (2)] describes accurately the flow of water through the soil pores. **Figure** 4*a* illustrates flow through soil under conditions similar to those used by Darcy for his experiments. Rewriting Darcy's law as Eq. (8),

$$Q = kiA = k\frac{h_{t_1} - h_{t_2}}{L} \tag{8}$$

yields a flow rate equal to 8.5×10^{-3} gal/s (3.2×10^{-5} m^3/s). [Here Q is flow rate, k permeability or hydraulic conductivity, i hydraulic gradient, h_{t_1} total head at entrance end of soil, h_{t_2} total head at exit end of soil, L length of soil sample, and A cross-sectional area of sample.] Darcy's coefficient of permeability varies more than any other engineering soil property, as much as 10 orders of magnitude (or 10^9 times) between a clean gravel and a plastic clay. Furthermore, discontinuities such as cracks or very pervious thin layers can greatly affect the permeability of a soil deposit.

A flow net provides a graphical representation of flow through soil, particularly useful for evaluating complex conditions. For two-dimensional flow through isotropic soil, the flow net consists of a set of flow lines and orthogonal equipotential lines that divide the soil into curvilinear squares. A flow line traces the flow path through the soil for a drop of

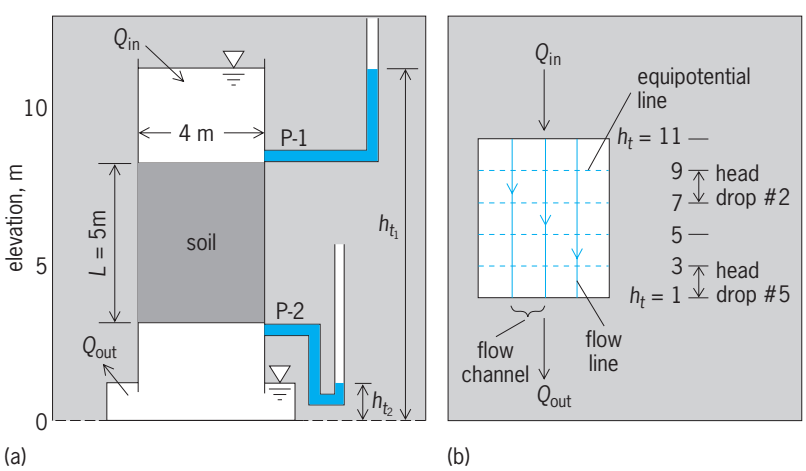

Fig. 4. Diagrams of flow through soil. (*a*) Flow through a permeameter. (*b*) Simple flow net.

water. While an infinite number of flow lines exist, a flow net utilizes an arbitrary selection to divide the soil into a convenient number of flow channels, n_f. Equipotential lines, drawn through points of equal total head, divide the flow net into a number of uniform head drops, n_d. The flow per unit length, l, perpendicular to the page is given by Eq. (9),

$$\frac{Q}{l} = k \cdot \Delta h \cdot \oint \tag{9}$$

where Δh = total head drop = $h_{t_1} - h_{t_2}$, and

$$\oint = \text{shape factor} = n_f / n_d$$

The flow net for the Darcy permeameter (Fig. 4*b*) has a shape factor equal to 4/5, which yields a flow equal to that computed by using Darcy's equation. The gradient between two points in the net equals the head drop divided by the distance between them. Since the net in Fig. 4 has equally spaced equipotentials, the gradient in the direction of flow remains constant.

Figure 5*a* shows a flow net for an earth dam with the following characteristics: Δh_t = 88 ft (27 m), $n_f = 9.9$, $n_d = 18$,

$$\oint = 0.55$$

and Q/l = 0.014 ft^2/day (0.0013 m^2/day) for $k = 4 \times 10^{-8}$ in./s (1×10^{-7} cm/s). The gradient in the direction of flow varies throughout the flow region, increasing as the flow approaches the drain. (Note that the distance between equipotentials decreases in Fig. 5*a*.) Areas of uncontrolled high gradients can experience instability problems from low effective stress and internal erosion. The graded filter shown on the upstream side of the drain ensures that the fine particles of the compacted fill do not get washed into the drain. Filters generally consist of layers of selected soil with particle sizes chosen to prevent migration of fines. Geosynthetic filters, rapidly growing in acceptance, consist of one or more layers of fabric engineered to retain soil fines. *See* DAM; GEOSYNTHETIC.

A flow net provides information on the pore pressure throughout the flow region. Rather than using the entire flow net for some analyses, engineers can obtain the information applicable to a surface of interest (for example, a potential shear surface) and portray the heads obtained from the net in terms of a piezometric line. Figure 5*b* shows that the piezometric line does not necessarily coincide with the phreatic surface. Under conditions of horizontal flow or no flow, the piezometric line and phreatic surface coincide.

Figure 5*c* shows a simplification that permits evaluation of the flow into a fully penetrating well using Darcy's law. The Dupuit assumption to Darcy's law states that the hydraulic gradient remains constant from top to bottom in the pervious formation and

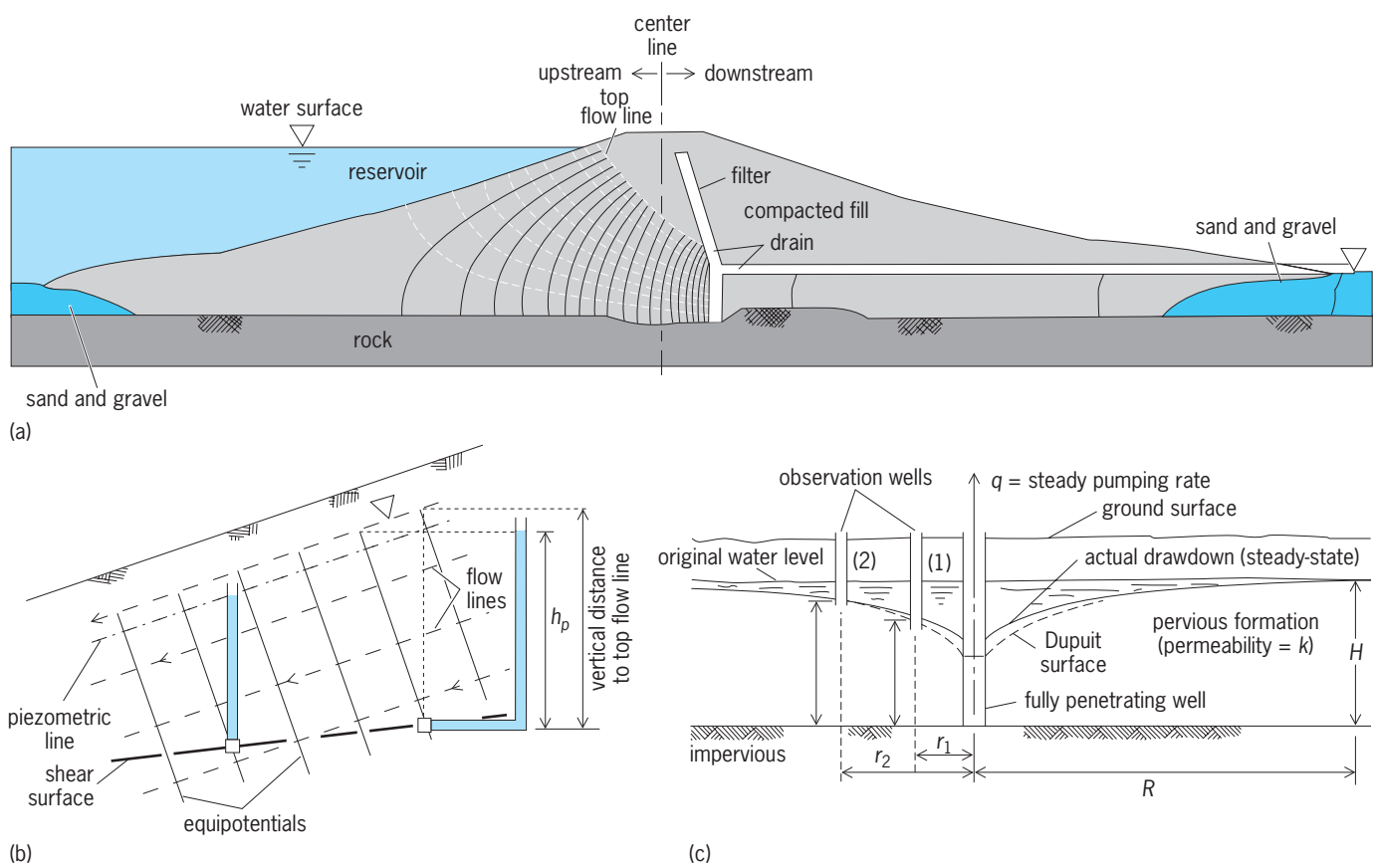

Fig. 5. Analysis of flow through soil of an earth dam. (*a*) Profile of flow through the dam. (*b*) Flow parallel to the slope. Piezometric line is a line located at a vertical distance h_p from a surface of interest; uppermost flow line represents the phreatic surface (water table). (*c*) Flow toward a fully penetrating well.

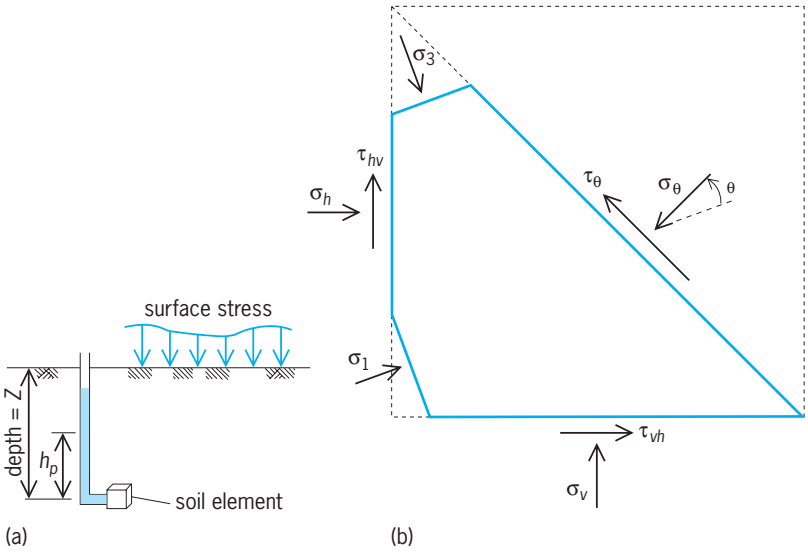

Fig. 6. Stress in soil represented by (*a*) soil element beneath loaded surface and (*b*) free-body diagram with stress vectors. τ = shear stress. Friction angle (θ) is measured counterclockwise from direction of normal stress (σ_1).

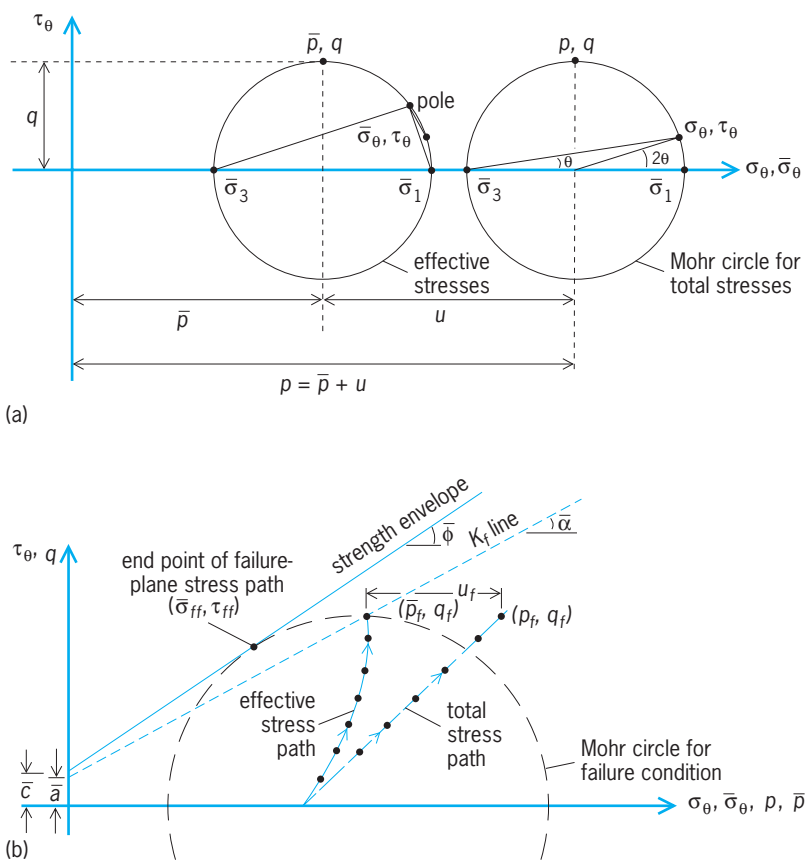

Fig. 7. Mohr's circle (*a*) for state of stress at a point and (*b*) stress path to failure for the soil element shown in Fig. 6.

equal to the slope of the water surface, as in Eq. (10).

$$i = dh/dr \tag{10}$$

Then, the flow rate, q, out of the well using any set of consistent units is given by formula (11), where q

$$q = \frac{\pi k(h_2^2 - h_1^2)}{\ln(r_2/r_1)} \tag{11}$$

is water quantity per unit time (Q/t). The values for h_1, h_2, r_1, and r_2 are measured (Fig. 5*c*) at observation wells 1 and 2, respectively. Pumped wells have many uses in projects involving civil engineering, including water supply, reduction of pore pressures to increase effective stresses and thus strength, and control of contaminants in the groundwater.

Stress in soil. Figure 6 illustrates the concept of stress in soil. Geotechnical engineers frequently work with normal stresses acting perpendicular to a surface of interest, shear stresses acting parallel to a surface of interest, and pore pressures (sometimes known as neutral stresses) acting equally in all directions, expressed as stress vectors (Fig. 6*b*). *See* SHEAR.

The effective stress principle [Eq. (1)] states that the total normal stress experienced by an element of soil consists of the effective normal stress, which controls strength and deformation, and the pore water pressure. Since water can not sustain shear, pore water pressures do not affect shear stresses. A bar over the symbol identifies a stress as effective or a property as related to effective stresses (for example $\bar{\sigma}$, $\bar{\phi}$, $\bar{c}$.

The stress history of a soil greatly influences its behavior. Two powerful mathematical tools used to understand and make use of stress history are Mohr's circle of stress and stress path (**Fig. 7**). Mohr's circle provides a graphical representation of the stresses in any direction acting at a point. A distance equal to the magnitude of the pore pressure separates the circle for total stresses from that representing effective stresses. Each point along the circle represents the normal and shear stresses acting on a particular plane. The pole of the Mohr circle relates the stresses to the physical planes on which they act. A line through the pole and point σ_θ, τ_θ will lie parallel to the plane on which σ_θ and τ_θ act.

Two particular points prove very helpful in portraying stress history—the point for stresses on the failure plane, often taken as the point of maximum obliquity (maximum value of $\tau/\bar{\sigma}$), and that for maximum shear stress τ (identified by coordinates p and q). The total stress path (TSP) and effective stress path (ESP) provide a locus of stress points representing successive states of stress in the history of the soil. For clarity, Mohr circles usually do not appear in stress path plots, although one could reconstruct the circle at any point along the stress path. Failure occurs when the effective failure-plane stress path reaches the strength envelope (also known as yield surface) or when the $\bar{p}$-q stress path reaches the K_f line as in Fig. 7.

The stress path can reach the strength envelope because of a loading, an unloading, an increase in pore pressure, or a combination of these. Stress paths help engineers identify mechanisms and select methods and soil parameters for solving a problem. *See* STRESS AND STRAIN.

Strength. Soil behaves as a frictional material; that is, it resists shear stresses by the frictional action between particles. **Figure 8** shows a strength test where a sample under a constant normal force gets sheared along a predetermined horizontal plane. The displacement-versus-shear stress plot (Fig. 8*b*)

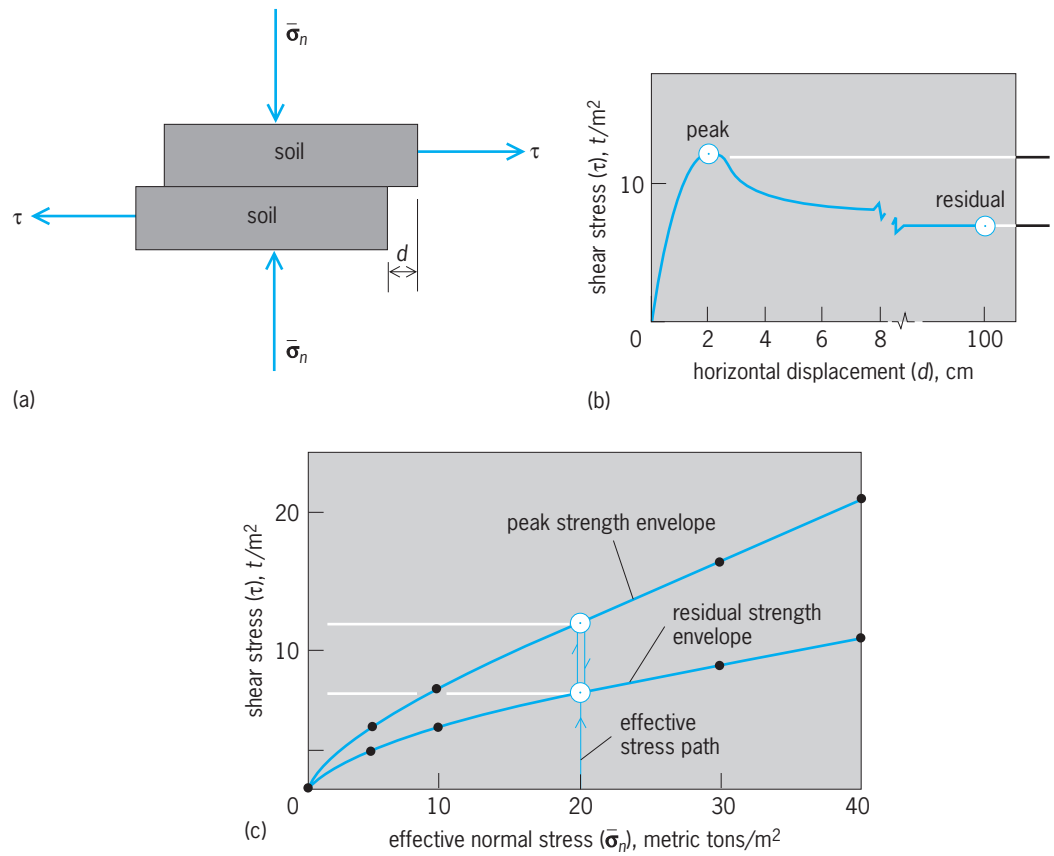

Fig. 8. Determination of soil strength. (*a*) Schematic diagram of a direct shear test. (*b*) Shear stress as a function of horizontal displacement. (*c*) Effective failure-plane stress path and strength envelopes. τ = shear stress; σ_n = effective normal stress. 1 m^2 = 10.8 ft^2; 1 cm = 0.4 in.; 1 metric ton = 1.1 ton.

shows two important characteristics—a peak shear strength and a residual value obtained after large deformations. A series of these tests at different values of normal stress yields the peak and residual strength envelopes shown. These strength envelopes, unique for practical purposes, remain valid for a wide range of effective stress paths to failure. Some soils exhibit a monotonically ascending displacement–versus–shear stress curve where the peak strength equals that at large deformations.

Soils generally do not have true cohesion, that is, when $\bar{\sigma} = 0$, strength = 0. Since strength envelopes can be curved, use of the linear Coulomb friction equation (3) requires care in the selection of parameters that are valid for the stress range of interest.

Consolidation and swelling. Consolidation refers to a process during which a saturated soil experiences a decrease in water content due to a decrease in the volume of voids. The opposite process, swelling (heave) results in an increase in water content due to an increase in the volume of voids. Consolidation and heave involve flow of water into or out of the soil, and a corresponding change in volume. Changed pore pressures (for example, from a change in load or heads) establish a hydraulic gradient, since the pore pressure near the drainage boundary remains essentially unchanged. In soils with low permeability, the water flows slowly, and a time lag exists between the event that changed the pore pressures and the completion of consolidation. The compression of a partially saturated soil involves compression, solution, or expulsion of air until the soil becomes nearly saturated. The spring–dashpot example in Fig. 2 illustrates the mechanism of consolidation.

Figure 9*a* shows a layer of clay with double drainage subjected to a uniform load. After application of the load, pore pressures near the center of the layer increase by an amount equal to the magnitude of the load. The high permeability of the sand permits rapid dissipation of excess pore pressures, and deformations occur almost instantaneously. The total head difference between the center of the clay and the sand layers creates a hydraulic gradient, which produces a flow from the center outward in two directions. Figure 9*b* shows a consolidation test (known as an oedometer test) that models the one-dimensional compression experienced by the clay layer. The expression for final settlement shown in Fig. 9*d* remains valid for fine- as well as coarse-grained soils.

The oedometer tests provide information to predict both the magnitude and rate of deformation. In coarse-grained soils, deformations occur very rapidly. In fine-grained soils, the theory of consolidation helps predict the rate of deformation using the coefficient of consolidation, a parameter influenced mostly by the permeability of the soil. The theory of consolidation permits computation of more detailed results, including distribution of pore pressures and deformations throughout the deposit at any time.

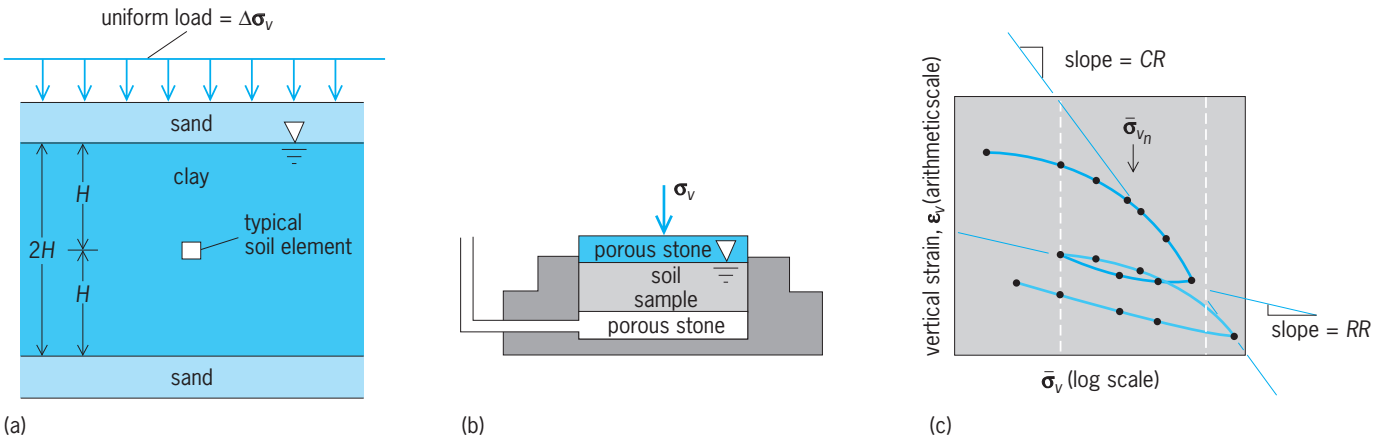

Fig. 9. Consolidation settlement. (*a*) Clay layer with double drainage, no lateral strains. (*b*) Oedometer test to obtain soil parameters. (*c*) Oedometer test results; each data point represents the end of a load increment; and left broken line represents 10 metric tons/m^2 and right broken line represents 100 metric tons/m^2. 1 metric ton = 1.1 ton; 1 m^2 = 10.8 ft^2.

Soil improvement. For many centuries, builders have tried to make soil stronger, denser, lighter, less compressible, less pervious, more pervious, or more resistant to erosion. One of the most common soil improvement techniques compacts the soil with a heavy roller or tamper. Compaction of fine-grained soils generally takes place in layers no more than 12 in. (30 cm) thick. Compaction densifies the soil, and it can increase its strength and decrease its compressibility and permeability. The key to obtaining the desired dry density lies in achieving a preselected water content in the soil undergoing compaction. Compaction at the optimum water content yields the highest dry density for a given compactive effort. Soil compacted at or near the optimum water content has negative pore pressures.

Other techniques use to improve engineering properties of soil include preloading; dewatering; stabilization with chemical admixtures; and reinforcement with rods, strips, grids, and fabrics made of natural of synthetic materials.

Slope stability. In earth slopes, the force of gravity pulls the soil mass downward. The internal friction between soil particles resists this force. As long as the resisting forces exceed the driving forces, the slope remains standing. A stability assessment of a slope involves a comparison of the stresses acting on the sliding mass and the strength of the soil along which sliding occurs. This assessment requires evaluation of the geometry and pore pressures, which together determine the stresses, and the strength of soils involved in a potential slide.

Geotechnical engineers use the term factor of safety (FS) to indicate the level of stability of a slope. Using forces, FS is the ratio of the shear resistance on a shear surface to the net shear force along a shear surface; or, using stresses, FS is the ratio of the shear strength along a shear surface to the average shear stress on a shear surface. When FS decreases to 1, failure occurs. **Figure 10** shows the forces acting on a wedge of soil. To increase FS would require an increase in shear resistance, S, or a decrease in shear force, T.

The following actions would increase S: (1) Increase W. (2) Increase $\bar{c}$ (densify, replace soil chemicals, temperature). (3) Increase $\bar{\phi}$. (4) Decrease U_s (drain).

The following actions would decrease T: (1) Decrease U_L (drain). (2) Decrease W (excavate, replace with lighter soil). (3) Decrease P_L (unload above slide). (4) Increase U_R (raise reservoir level). (5) Increase R (install tiebacks, piles, berm retaining wall).

Slope stability analyses generally use the principles of force and moment equilibrium, a technique known as limit equilibrium. In general, a free body bounded by the slope surface and a potential failure surface is isolated, and the stresses are determined by solving the static equilibrium equations. The ratio of the determined stresses and the shear strength of the soil yields the factor of safety. Repetition of the analysis with other possible failure surfaces yields different factors of safety because both the driving and resisting forces can vary from surface to surface. The most critical surface has the lowest factor of safety. Computer programs allow the rapid investigation of hundreds of potential failure surfaces along cross sections with different soil layers. *See* STATICS.

More sophisticated numerical techniques consider the stress–strain properties of the soil, and allow the analysis of slopes composed of soils with complicated stress–strain–strength properties. Finite-element stress–deformation analyses can

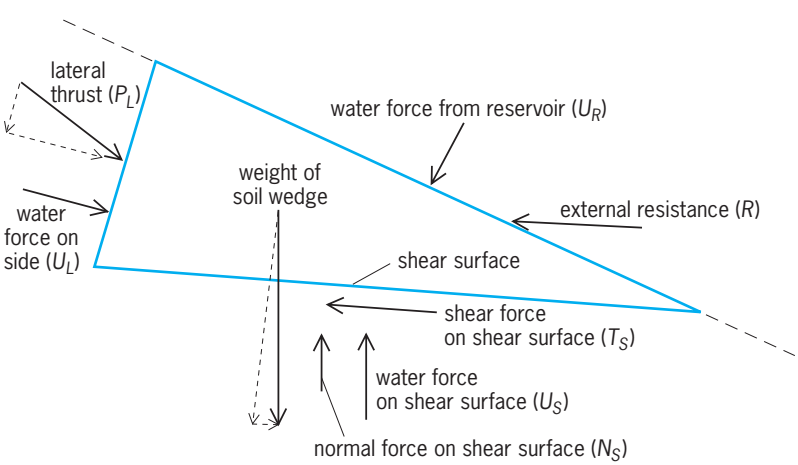

Fig. 10. Free-body diagram showing forces on a slope wedge.

provide a detailed description of the stresses acting throughout the slope. The engineer calculates the factor of safety by integrating the acting shear stress along a potential failure surface and comparing the result to the available shear strength along the surface. *See* FINITE ELEMENT METHOD.

The principles of stability analysis outlined above apply for natural slopes as well as for earth structures such as earth dams and highway embankments. With constructed facilities, the engineer has more detailed knowledge and some degree of control over the material properties. An earth structure must satisfy all the criteria needed to perform its intended function. These may include overall stability, limited deformation, weight limitations, specific flow characteristics, or particular dimensions. *See* HIGHWAY ENGINEERING.

Foundations. Nearly every large structure derives its support from contact with the earth. The foundation, whether of a building, tower, highway, retaining wall, bridge, or storage tank, transfers live and dead loads from the structure to the ground. Besides providing support, or adequate bearing capacity, the foundation must keep settlements within an acceptable range for the structure. Foundations vary from the simple preparation of a patch of ground to elaborate arrays of long steel piles driven to great depths. Other foundation types include rafts, footings, caissons, many types of piles, stone columns, grouted soil, frozen soil, mechanically reinforced soil, and hybrid combinations of these.

The bearing capacity of a foundation element depends on its ability to resist a shear failure, that is, a loss of stability. The principles of static equilibrium that govern stability of slopes also control foundation stability. The sum of the resisting forces must exceed the sum of the driving forces. *See* FOUNDATIONS.

Shallow. In practice, engineers design most shallow foundations by using semiempirical formulas that combine the concepts embodied in the theories with the experience gained by measuring the performance of actual foundations. Most shallow foundations have a large factor of safety against a general shear failure (3 or above) because restrictions on allowable settlement usually prove much more stringent and result in low allowable bearing stresses.

For either drained or undrained conditions, elastic theory predicts the stress change at any depth below a loaded area. The change in stress produces a strain in accordance with Hooke's law. The sum of the individual strains from the footing base to some depth where the stress change becomes negligible equals the settlement of the footing. In the case of shallow foundations on fine-grained soils, additional settlement will occur because of time-dependent consolidation under the applied load. Difficulties arise in selecting the appropriate elastic parameters, Young's modulus and Poisson's ratio, for the anticipated drainage and stress conditions. The stress path method, which attempts to simulate the anticipated changes in the state of stress with soil samples in the laboratory, can aid the engineer in selecting correct parameters. *See* ELASTICITY; YOUNG'S MODULUS.

Deep. Where soils of adequate strength do not lie sufficiently close to the ground surface, deep foundations can transfer the loads to stronger soil strata or distribute the loads throughout a greater mass of soil. Deep foundation elements, such as caissons or piles, can develop support along their entire length as well as at their base or tip. *See* CAISSON FOUNDATION.

Friction piles develop resistance primarily through shear resistance (skin friction) along their length. End-bearing piles develop resistance primarily at their tip by bearing on firmer soil or rock. Most piles combine the characteristics of the two pile types. Pile equipment has evolved that can drive the piles into the ground with powerful hammers, vibrate the piles through soil, or place the piles in predrilled holes. Pile widths or diameters range from about 4 in. (10 cm) for minipiles to several feet for offshore oil platforms, with lengths up to several hundred feet. Common pile materials include steel, concrete, and wood.

Caissons or drilled piers usually have greater diameters than piles and rely on end-bearing support. Their construction involves excavation instead of driving, and it often allows visual inspection of the ultimate bearing surface.

Tests and properties. Field and laboratory tests are used to determine soil properties. Those conducted in the field measure density, permeability, shear strength, penetration resistance, compressibility, and deposit boundaries. Those done in the laboratory measure water content, expansion index, grain size distribution, permeability, shear strength, dynamic stress, resonant frequency, and mineral composition.

The broad range of geotechnical instrumentation includes piezometers for measuring ground-water pressure, earth pressure cells for measuring total stress in soil, settlement plates for measuring vertical displacements, extensometers for determining deformation, tiltmeters for measuring inclination, inclinometers for measuring lateral deformation versus depth, piezoelectric transducers for measuring acoustic emissions, load cells for measuring load in a structural member, strain gages for measuring strain in a structural member, and various devices for measuring temperature.

Other soil engineering problems. A naturally occurring material such as soil presents geotechnical engineers with several particularly difficult problems. In addition to the complexities of working with a nonhomogeneous material with nonlinear anisotropic stress–strain properties, the strength varies from point to point, may vary with time or deformation, and remains difficult to measure precisely. The simple examples that have been discussed illustrate the fundamentals of soil mechanics. Engineers routinely extend and apply these fundamentals to the complicated conditions encountered in the field. Some of these applications include two- and three-dimensional problems; dynamic loading of soil; construction of earthquake-resistant facilities; earth pressures and retaining walls, tunnels, and excavations; subsidence; problems involving expansive soils; field exploration; probability and reliability;

and the use of synthetic materials such as fabrics, membranes, and superlight fills. *See* CLAY; CONSTRUCTION METHODS; ENGINEERING GEOLOGY; RETAINING WALL; SAND; SOIL; TUNNEL.

Francisco Silva-Tulla

Bibliography. American Society for Testing and Materials, *Annual Book of ASTM Standards*, sect. 4: *Construction*, vol. 04.08: *Soil and Rock; Building Stones, Geotextiles*, 1989; H. R. Cedergren, *Seepage, Drainage and Flow Nets*, 3d ed., 1997; R. F. Craig, *Soil Mechanics*, 6th ed., 1997; B. M. Das, *Fundamentals of Soil Dynamics*, 1983; J. Dunnicliff, *Geotechnical Instrumentation for Monitoring Field Performance*, 1988; T. W. Lambe and R. V. Whitman, *Soil Mechanics*, 1969; J. K. Mitchell, *Fundamentals of Soil Behavior*, 2d ed., 1993; K. Terzaghi and R. B. Peck, *Soil Mechanics in Engineering Practice*, 3d ed., 1996; U.S. Navy, *Soil Mechanics*, Design Man. 7.01, 1986.

Soil microbiology

The study of biota that inhabit the soil and the processes that they mediate. The soil is a complex environment colonized by an immense diversity of microorganisms. Soil microbiology focuses on the soil viruses, bacteria, actinomycetes, fungi, and protozoa, but it has traditionally also included investigations of the soil animals such as the nematodes, mites, and other microarthropods. These organisms, collectively referred to as the soil biota, function in a belowground ecosystem based on plant roots and litter as food sources.

Historically, soil microbiology has been concerned with the role of microorganisms in maintaining the soil's productivity for plant growth. Plant growth is often limited by the availability of nitrogen, and many of the early investigations focused on microorganisms that transform nitrogen into forms usable by plants. S. Winogradsky first isolated the bacteria responsible for the conversion of ammonium (NH_4^+) to nitrate (NO_3^-) in 1890. Since Winogradsky also advanced the understanding of microbial transformations of sulfur and iron and developed important concepts in classifying soil organisms, he is regarded as the founder of soil microbiology. Nitrogen fixation is a microbial process whereby atmospheric nitrogen (N_2) is converted to a form available for plants; it is the key to the nitrogen supply of leguminous plants (that is, clover, alfalfa, and soybean) and some woody species (such as alder and *Ceanothus*). Soil microorganisms have also been studied because of their abilities to produce medically useful antibiotics and to degrade potentially toxic pollutants. Modern soil microbiology represents an integration of microbiology with the concepts of soil science, chemistry, and ecology to understand the functions of microorganisms in the soil environment.

Microorganisms. The surface layers of soil contain the highest numbers and variety of microorganisms, because these layers receive the largest amounts of potential food sources from plants and animals. The soil biota form a belowground system based on the energy and nutrients that they receive from the decomposition of plant and animal tissues. The primary decomposers are the bacteria and fungi. Bacteria are single-celled prokaryotes and are the most numerous of the microorganisms in soil. Typically, there are about 1 billion bacteria representing thousands of different bacterial types in a single gram of soil. Bacteria are classified according to their sources of energy and carbon. Bacteria that decompose organic materials for carbon and energy are known as organotrophs or heterotrophs. Chemolithotrophic bacteria use alternative chemical sources of energy and inorganic carbon, and they are important in the transformations of nutrients and metals.

Fungi are the eukaryotic organisms commonly known as molds, mildew, yeasts, and mushrooms. Soil fungi exist primarily as thin, long filaments called hyphae. Many forest soils contain more than a kilometer of fungal hyphae in a single gram of soil. The mushrooms and fungi visible on the surface are the reproductive structures produced by the vast network of hyphae in the soil or woody material. The protozoa are single-celled, mobile eukaryotes such as the ciliate paramecia and the amebas. Nematodes (round worms) and the microarthropods are representatives of the soil animals. The protozoa and the soil animals graze or prey on the smaller soil organisms or on organic materials. Algae and lichens exist near the soil surface because they require light for photosynthesis. The distribution of soil organisms often depends on the size of the pore spaces between soil particles and the availability of water. *See* FOREST SOIL.

Ecological roles. Microorganisms, especially algae and lichen, are pioneering colonizers of barren rock surfaces. Colonization by these organisms begins the process of soil formation necessary for the growth of higher plants. After plants have been established, decomposition by microorganisms recycles the energy, carbon, and nutrients in dead plant and animal tissues into forms usable by plants. Therefore, microorganisms have a key role in the processing of materials that maintain life on the Earth. The transformations of elements between forms are described conceptually as the elemental cycles.

In the carbon cycle, microorganisms transform plant and animal residues into carbon dioxide and the soil organic matter known as humus. Humus improves the water-holding capacity of soil, supplies plant nutrients, and contributes to soil aggregation. Microorganisms may also directly affect soil aggregation. For example, fungal hyphae can physically bind microaggregates together, while bacteria produce polysaccharides that can bind soil particles into microaggregates. The extent of soil aggregation determines the workability or tilth of the soil. A soil with good tilth is suitable for plant growth because it is permeable to water, air, and roots. *See* HUMUS.

Soil microorganisms play key roles in the nitrogen cycle. The atmosphere is approximately 80% nitrogen gas (N_2), a form of nitrogen that is available to plants only when it is transformed to ammonia (NH_3) by either soil bacteria (N_2 fixation) or by

humans (manufacture of fertilizers). The N_2-fixing symbiotic bacteria, *Rhizobium*, form root nodules on legumes such as soybean, peas, beans, and alfalfa. Nitrogen-fixing crops are a key component of agricultural productivity, and their use decreases reliance on chemical fertilizers. Certain chemolithotrophic bacteria transform NH_4^+ to NO_3^- in the process called nitrification. Nitrate is very mobile in the soil and is readily available for plant assimilation. However, NO_3^- may also be lost from the soil through the processes of denitrification and leaching. Soil bacteria also mediate denitrification, which returns nitrogen to the atmosphere by transforming NO_3^- to N_2 or nitrous oxide (N_2O) gas. Microorganisms are crucial to the cycling of sulfur, phosphorus, iron, and many micronutrient trace elements.

In addition to the elemental cycles, there are several interactions between plants and microbes which are detrimental or beneficial to plant growth. Some soil microorganisms are pathogenic to plants and cause plant diseases such as root rots and wilts. Many plants form symbiotic relationships with fungi called mycorrhizae (literally fungus-root). Mycorrhizae increase the ability of plants to take up nutrients and water. The region of soil surrounding plant roots, the rhizosphere, may contain beneficial microorganisms which protect the plant root from pathogens or supply stimulating growth factors. The interactions between plant roots and soil microorganisms is an area of active research in soil microbiology. *See* BIOGEOCHEMISTRY; MYCORRHIZAE; NITROGEN CYCLE; NITROGEN FIXATION; NITROGEN OXIDES; RHIZOSPHERE.

Biotechnology and soil microorganisms. The incredible diversity of soil microorganisms is a vast reserve of potentially useful organisms. Antibiotics are compounds produced by microorganisms which can kill or inhibit the growth of other microorganisms. Many of the medically important antibiotics are produced by filamentous bacteria known as actinomycetes. The genus *Streptomyces* alone produces more than 3000 antibiotics. Fungi such as *Penicillum* also produce a wide variety of antimicrobial compounds. The soil is the largest reservoir of these medically important microorganisms. *See* ANTIBIOTIC.

The numerous natural substances that are used by microorganisms indicate that soil microorganisms have diverse mechanisms for degrading a variety of compounds. Human activity has polluted the environment with a wide variety of synthetic or processed compounds. Many of these hazardous or toxic substances can be degraded by soil microorganisms. This is the basis for the treatment of contaminated soils by bioremediation, the use of microorganisms or microbial processes to detoxify and degrade environmental contaminants. Often the degradation of contaminants by microorganisms can be stimulated by manipulating the environmental conditions, for example, by aeration. In some cases, adding specific degrading microorganisms can accelerate the breakdown of the pollutants. Recalcitrant compounds are extremely resistant to microbial degradation and may take many decades to degrade into harmless byproducts. Soil microbiologists study the microorganisms, the metabolic pathways, and the controlling environmental conditions that can be used to eliminate pollutants from the soil environment. *See* HAZARDOUS WASTE.

Methods for study. Microbiologists traditionally isolate pure strains of microorganisms by using culture methods. However, generally less than 10% of the soil microorganisms are recoverable by culture methods, and many soil microorganisms remain poorly characterized or unknown. Methods that do not rely on culturing microorganisms include microscopic observation and biochemical or genetic analysis of specific cell constituents. Microscopic methods generally use fluorescent stains so that microbial cells are visible against the opaque soil particles. Alternatively, cells can be separated from soil particles and then viewed. Fluorescent stains can be linked to antibodies that react only to a specific soil microorganism, a technique known as immunofluorescence, to obtain counts of a specific microorganisms without culturing.

Specific cell constituents, such as the fatty acids present in the cell membranes, can be extracted for biochemical analysis. Sequences of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) can also be used to identify specific microorganisms or groups of microorganisms. A known DNA or RNA sequence can be used as a probe to determine the amount of a specific gene in a soil population. The gene probe is labeled chemically or radioactively and incubated with the target DNA or RNA extracted from the soil. After the gene probe binds to identical sequences, genes can be detected and quantified. These techniques have recently been used to characterize the diversity of soil microorganisms that are not cultured with standard methods. *See* DEOXYRIBONUCLEIC ACID (DNA); RIBONUCLEIC ACID (RNA).

The rates or controlling factors for microbial processes are studied by using methods from chemistry, biology, and ecology. Typically, these studies involve measuring the rate of production and consumption of a compound of interest. Many investigators use stable or radioactive isotopes to trace the transformations of the compound. The results of these studies are commonly analyzed by using mathematical models. Models allow the information from one system to be generalized for different environmental conditions. *See* FLUORESCENCE MICROSCOPE; IMMUNOFLUORESCENCE; MICROBIOLOGY; MOLECULAR BIOLOGY; SOIL; SOIL CHEMISTRY; SOIL ECOLOGY.

Jeanette M. Norton

Bibliography. M. Alexander, *Soil Microbiology*, 2d ed., 1977; R. M. Atlas and R. Bartha, *Microbial Ecology*, 4th ed., 1997; D. P. Labeda, *Isolation of Biotechnological Organisms from Nature*, 1990; F. B. Metting, *Soil Microbial Ecology*, 1993; E. A. Paul and F. E. Clark, *Soil Microbiology and Biochemistry*, 2d ed., 1996; R. L. Tate, *Soil Microbiology*, 2d ed., 2000.

Soil sterilization

A chemical or physical process that results in the death of soil organisms. This control method affects many organisms, even though the elimination of only

specific weeds, fungi, bacteria, viruses, nematodes, or pests is desirable. Even if complete sterilization is achieved, it is short lived since organisms will recolonize this biological vacuum quite rapidly. Soil sterilization can be achieved through both physical and chemical means. Physical control measures include steam and solar energy. Chemical control methods include herbicides and fumigants. Dielectric heating and gamma irradiation are used less frequently as soil sterilization methods. Composting can be used to sterilize organic materials mixed with soil, but it is not used for the sterilization of soil alone. Soil sterilization is used in greenhouse operations, the production of high-value or specialty crops, and the control of weeds.

Steam. Heating the soil with steam to 180°F (82°C) for 30 min will kill most organisms except a few heat-tolerant viruses and weed seeds. These heat-tolerant organisms can be difficult to eliminate even at 212°F (100°C). However, heating soil to temperatures greater than 176°F (80°C) has disadvantages. These high temperatures require more energy, create a biological vacuum, and leave the soil toxic to plant growth because of excessive exchangeable manganese, nitrates, and ammonia. These toxic materials must be leached from soil with water, which can result in environmental pollution problems. An alternative to sterilization is pasteurization. The pasteurization of soil with aerated steam at 158°F (70°C) kills most plant pathogens without leaving a biological vacuum. The saprophytic microflora that survive pasteurization leave the soil more suppressive to plant pathogens. Pasteurization also avoids toxicity problems. Steam can be applied to soil by releasing it directly under plastic tarpaulins supported by nylon netting, or through perforated pipes buried 18–23 in. (45–60 cm) deep in the soil. A common problem in steaming a large ground bed area is using only a single steam inlet. An area no longer than 66 ft (20 m) and no wider than 16 ft (5 m) should be sterilized with one inlet.

Negative-pressure steaming is an alternative to conventional steaming and requires less energy. In this method a fan is used to set up a negative pressure within buried pipes so that steam supplied at the soil surface is actively moved downward. This arrangement provides more uniform and deeper distribution of heat than sheet steaming. *See* SOIL CHEMISTRY.

Solarization. Solarization is a simple technique that captures radiant heat energy from the Sun. This energy causes physical, chemical, and biological changes in the soil. These changes pasteurize the soil, leading to control or suppression of soil-borne plant pathogens such as fungi, bacteria, nematodes, and pests, along with weed seed and seedlings. Solarization is accomplished by placing a clear plastic tarpaulin over the soil surface and anchoring the edges by burying them in a trench around the area to be treated. The Sun's radiant energy then heats the soil beneath the plastic. Solarization is most effective when done during June and July in western states such as California, when the day lengths are long, air temperatures are high, and skies are clear. Depending on geographic location, solarization can also be effective in May, August, and September. The plastic should be left in place for 4–6 weeks to allow heating to reach greater depths. The clear plastic tarpaulin can be construction-grade plastic and vary in thickness from 1 to 6 mils (0.001 + 0.006 in. or 0.0025 to 0.015 cm). The thinner plastic allows better soil heating, since it reflects less solar energy. For some applications, ultraviolet-resistant plastic should be considered. This plastic is more expensive but it can be used more than once.

Efficacy can be further improved by combining solarization with allelopathy, which involves incorporating green plant material from crucifers into the soil. The gases released from this plant material are trapped under the plastic, increasing the efficacy of the solarization treatment.

Solarization can be done on a small or large scale. Large areas can be solarized by connecting a series of tarpaulins with long-lasting, heat-resistant glue. This technique can even be applied postplant to orchards with heat-tolerant plants such as pistachio. When the treatment is complete, removal of the plastic can be troublesome, especially on large acreage; if the plastic is left in place too long, it becomes brittle and begins to tear. Solarization is a safe and effective means to control soil-borne plant pathogens and is especially well suited for use by home gardeners and organic growers. On a large scale, this technique is economical only on high-value crops.

Herbicides. In the early 1900s, scientists began testing materials for the purpose of killing perennial weeds and sterilizing the soil. Salt and copper sulfate were tested, but the arsenicals proved the most satisfactory. The organic compounds introduced in the mid-1940s brought an end to the use of nonorganic arsenicals, since they were safer to use, had less potential for long-term harm to the environment, and provided a broader spectrum of weed control.

As agriculture became more mechanized, chemical weed control began to play an integral part in large-scale crop production. The herbicides used in crop production tend to be selective herbicides (temporary soil sterilants) and not long-term broad-spectrum sterilants. The broad-spectrum sterilants tend to be used more for weed control along railroads, highways, driveways, irrigation ditches, pipelines, power lines, and industrial sites. *See* HERBICIDE.

Fumigation. Sterilization by fumigation depends on infiltrating the soil with the vapor of a volatile chemical. Some of the chemicals have a broad spectrum of activity against soil microflora, fauna, and weeds, while others are toxic to fewer organisms. Like pasteurization with steam, fumigation usually leaves a fumigant-tolerant microflora, particularly at depth and at the edges of the treated area. Soil in a confined area (a garbage can or a plastic-covered area on a cement pad) can be fumigated much more thoroughly. Most fumigants require 2–3 days to kill organisms in the soil, depending on soil type and temperature. In order to keep the fumigants from

escaping, they are injected or incorporated into the soil, which is immediately covered with plastic or sealed by rolling. Some fumigants such as metham sodium can be applied in irrigation water through a sprinkler system. After 7–14 days the soil can be worked to allow the gas to escape. Some fumigants require a waiting period of several weeks before planting or seeding, to ensure that the fumigant has escaped or degraded. Nitrogen fertilization may also have to be adjusted since nitrifying bacteria are readily killed by fumigation. Leaching of the soil may also be required with some fumigants such as methyl bromide.

Fumigants are available as a gas, gas compressed to liquid, liquids alone or in mixtures, or granules. The combination of methyl bromide plus chloropicrin has strong activity against bacteria, fungi, nematodes, insects, and weeds. The other fumigants have a more limited range of activity. Some fumigants such as methyl bromide may move 20–32 in. (50–80 cm) in the soil from the point of injection, while other fumigants such as metham sodium may move only 4 in. (10 cm). The distance that vapors move in the soil is determined not only by the chemical applied but also by the size and continuity of soil pores, soil temperature, soil moisture, and soil type. The use of fumigants is expensive and generally limited to use on high-value crops such as potatoes, strawberries, bedding plants, and tobacco. Some fumigants are extremely toxic to humans and animals and must be handled with great caution. *See* FUMIGANT.

Soil sterilants and other pesticides have been closely regulated in the United States by the Environmental Protection Agency (EPA) under the Federal Insecticide, Fungicide, and Rodenticide Act and the Federal Food, Drug, and Cosmetic Act. More recently, the Montreal Protocol and the Food Quality Protection Act (FQPA) have created considerable concern for pesticide users since they will potentially eliminate many pesticides, including some soil sterilants. The pesticide methyl bromide has been included in the Montreal Protocol, an international environmental treaty, as an ozone depleter. Thus the EPA, based on the Clean Air Act, has frozen production levels of methyl bromide and set a schedule to phase it out of production by the year 2005. The loss of methyl bromide will be disruptive since the production, storage, shipping, or processing of nearly 100 crops depends on this pesticide. Methyl bromide is an effective herbicide, nematicide, insecticide, and fungicide and was used commercially in the United States for soil fumigation and quarantine purposes for most of the twentieth century. The pending elimination of methyl bromide has stimulated creative research that should lead to viable alternatives. However, finding a single alternative pesticide to replace the many uses of methyl bromide is very unlikely. Research will probably pursue ecologically based alternatives such as organic crop production, improved cultivars, organic soil amendments, biological controls, better crop rotations, soil solarization, better pesticide application methods, and the judicious use of low-risk chemical alternatives. Similar ecologically based alternatives will be needed for other pesticides when the EPA finishes implementing the FQPA and evaluating commercial products, particularly those classified as carbamates, organophosphates, and group B2 carcinogens (for example, iprodione). *See* INSECTICIDE; NEMATICIDE; PESTICIDE.

General considerations. The soil sterilization methods are improved by the presence of good soil moisture (not saturated soil). Good moisture enhances the probability that microorganisms will be in a metabolically active state, making them more susceptible to control measures. Sterilization methods are also improved by good soil tilth and the elimination of large pieces of plant debris. Some chemicals used for sterilization have soil temperature requirements, especially fumigants.

Soil sterilization through the use of chemicals is frequently the most practical method of weed control. However, soil sterilization for the control of soil-borne plant pathogens is an expensive control measure that affects many nontarget organisms. For control of soil-borne plant pathogens, consideration should be given to using safer and more cost-effective measures such as resistant plants, certified disease-free planting material, crop rotation, sanitation, seed treatments, fungicide soil drenches, and cultural practices (proper tillage, plant spacing, fertilization, and weed control). Prior to initiating control, a trained individual such as a county agent or university specialist should be consulted, to identify the target organism, recommend the safest and most cost-effective control measure, and help optimize effectiveness. *See* AGRICULTURAL CHEMISTRY; AGRICULTURAL SCIENCE (PLANT); FUNGISTAT AND FUNGICIDE.

Carl A. Strausbaugh

Bibliography. A. S. Crafts, *Modern Weed Control*, 1975; W. R. Jarvis, *Managing Diseases in Greenhouse Crops*, 1992; J. B. Ristaino and W. Thomas, Agriculture, methyl bromide, and the ozone hole: Can we fill the gaps?, *Plant Dis.*, 81:964–977, 1997.

Sol-gel process

A chemical synthesis technique for preparing gels, glasses, and ceramic powders. The synthesis of materials by the sol-gel process generally involves the use of metal alkoxides, which undergo hydrolysis and condensation polymerization reactions to give gels (see **illustration**).

The production of glasses by the sol-gel method is an area that has important scientific and technological implications. For example, the sol-gel approach permits preparation of glasses at far lower temperatures than is possible by using conventional melting. It also makes possible synthesis of compositions that are difficult to obtain by conventional means because of problems associated with volatilization, high melting temperatures, or crystallization. In addition, the sol-gel approach is a high-purity process that leads to excellent homogeneity. The sol-gel approach is adaptable to producing films and fibers as well as bulk pieces, that is, monoliths (solid

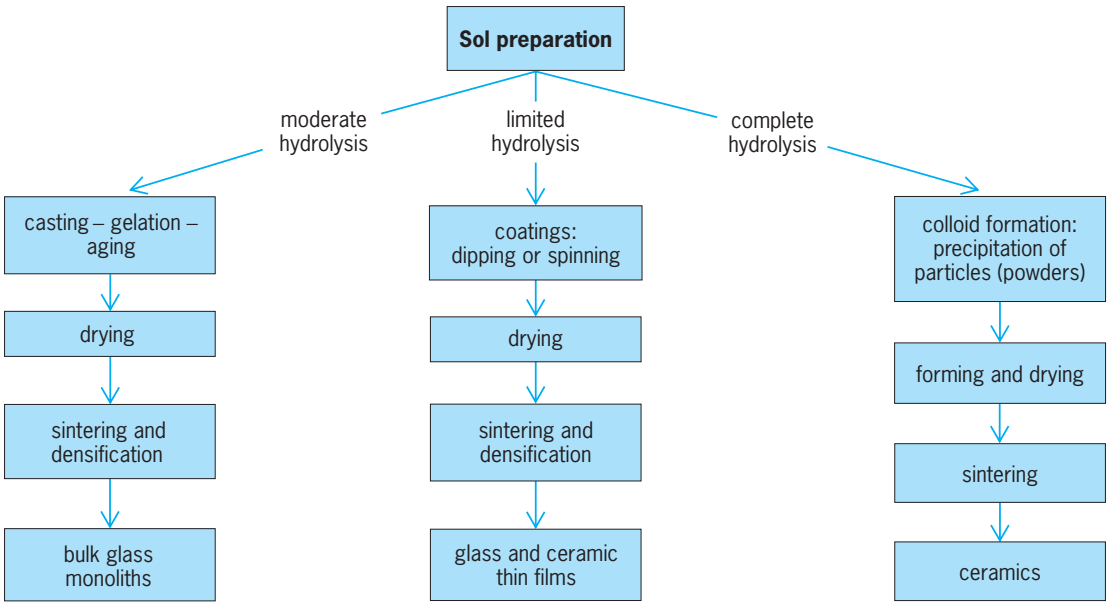

Flow chart showing the sol-gel methods used to prepare bulk glasses, thin films of glass or ceramics, and powders; the powders are then sintered to produce ceramics.

materials of macroscopic dimensions, at least a few millimeters on a side).

Glass formation. Formation of silica-based materials is the most widely studied system. An enormous range of multicomponent silicate or nonsilicate compositions have also been prepared.

The sol-gel process can ordinarily be divided into the following steps: forming a solution, gelation, drying, and densification. The preparation of a silica glass begins with an appropriate alkoxide such as tetraethylorthosilicate [$Si(OC_2H_5)_4$; TEOS], which is mixed with water (H_2O) and a mutual solvent, such as ethanol (C_2H_5OH), to form a solution. Tetramethylorthosilicate [$Si(OCH_3)_4$; TMOS] is also widely used as a precursor alkoxide. Hydrolysis leads to the formation of silanol groups (Si—OH). These species are only intermediates. Subsequent condensation reactions produce siloxane bonds (Si—O—Si) plus the by-product alcohol (ROH) or water. An example of the formation of a silica gel by the organometallic route is shown in reactions (1)–(3) [R is an alkyl group: $R = C_xH_{2x+1}$].

Hydrolysis and the formation of silanol groups:

$$\equiv Si-OR + H_2O \rightarrow \equiv Si-OH + ROH \qquad (1)$$

Alcohol condensation and the formation of siloxane groups:

$$\equiv Si-OR + HO-Si\equiv \rightarrow \equiv Si-O-Si\equiv + ROH \qquad (2)$$

Water condensation and the formation of siloxane groups:

$$\equiv Si-OH + HO-Si\equiv \rightarrow \equiv Si-O-Si\equiv + H_2O \qquad (3)$$

The silica gel formed by this process leads to a rigid, interconnected three-dimensional network consisting of submicrometer pores and polymeric chains. During the drying process (at ambient pressure), the solvent liquid is removed and substantial shrinkage occurs. The resulting material is known as a xerogel. When solvent removal occurs under hypercritical (supercritical) conditions, the network does not shrink and a highly porous, low-density material known as an aerogel is produced. Heat treatment of a xerogel at elevated temperature produces viscous sintering (shrinkage of the xerogel due to a small amount of viscous flow) and effectively transforms the porous gel into a dense glass.

Although the above discussion is based on silica, the sol-gel process is widely applicable to other inorganic oxides. By using various metal alkoxides $M(OR)_n$, where M is Al, Ti, V, Cr, Mo, W, Zr, and so on and OR is an alkoxy functionality as described previously, different oxides can be prepared. The resulting xerogels are either amorphous or nanocrystalline. Heat treatment of these materials leads to crystallization rather than the formation of a dense glass as occurs in the case of silica.

Hydrolysis and condensation. In general, the processes of hydrolysis and condensation polymerization are difficult to separate. The hydrolysis of the alkoxide need not be complete before condensation starts; and in partially condensed silica, hydrolysis can still occur at unhydrolyzed sites. Several parameters have been shown to influence the hydrolysis and condensation polymerization reactions, including the temperature, the solution pH, the particular alkoxide precursor, the solvent, and the relative concentrations of each constituent. In addition, acids and bases catalyze the hydrolysis and condensation polymerization reactions; therefore, they are added to help control the rate and extent of these reactions.

It is generally agreed that hydrolysis occurs by bimolecular nucleophilic displacement reactions. Under acidic conditions, an alkoxide group is protonated rapidly. As electron density is withdrawn from the silicon (Si), the silicon becomes more electrophilic and thus susceptible to attack by water.

Under basic conditions, water dissociates to rapidly form nucleophilic hydroxyl anions (OH^-). The hydroxyl anion then attacks the silicon atom. In one proposed mechanism for base hydrolysis, OH^- replaces OR^-, where OR^- is an alkoxide ion.

A number of factors influence hydrolysis. Under acidic conditions, such as in the presence of hydrochloric acid (HCl), the hydrolysis rate increases with alkyl substitution (electron-providing groups), while the reverse trend occurs with basic conditions, such as in the presence of ammonia (NH_3). The effect of the H_2O/Si ratio is to promote hydrolysis [reaction (1)] when the ratio is increased, while the alcohol-producing condensation mechanism [reaction (2)] is favored at less than stoichiometric amounts of water ($H_2O/Si < 4$). It should also be noted that the hydrolysis reaction may proceed in the reverse direction; in this reesterification, an alcohol displaces a hydroxyl group to produce water.

Condensation polymerization reactions also involve acid- or base-catalyzed bimolecular nucleophilic substitution reactions. The acid-catalyzed condensation mechanism involves protonated silanols, while the base-catalyzed mechanism involves deprotonated silanols. Electron-providing alkyl groups will reduce the acidity of the corresponding silanol and thus influence the pH dependence of the condensation mechanism. The nature of the solvent may also influence condensation. The ability of the solvent to encourage depolymerization has a significant effect on the morphology of the polymer structures produced. In addition, hydrogen bonding of the solvent (protic or aprotic) to the different silanols can either promote or retard condensation. *See* HYDROGEN BOND; HYDROLYSIS; POLYMERIZATION.

Microstructural development. The conditions under which hydrolysis and condensation occur have a profound effect on gel growth and morphology. These structural conditions greatly influence the processing of sol-gel glasses into various forms. It is well established that acid-catalyzed solutions with low water content (that is, conditions that produce linear polymers) offer the best type of solution for producing fibers. In the case of sol-gel films, control of precursor structure (that is, weakly branched to highly branched) is of vital importance in controlling the microstructure and the resulting properties of the film. Silicate gels prepared at low pH (<3) and low water content (<4 mol water per mol alkoxide) produce primarily linear polymers with low crosslink density. Additional crosslinks form during gelation, and the polymer chains become increasingly entangled. Silicate gels prepared under more basic conditions (pH 5–7) or higher water contents produce more highly branched clusters, which behave as discrete species. Gelation occurs by linking clusters together. At still higher pH and excess water content, colloidal silica is formed. *See* CATALYSIS; COLLOID; GEL.

Gelation and aging. As the hydrolysis and condensation polymerization reactions continue, viscosity increases until the solution ceases to flow. The time required for gelation to occur is an important characteristic that is sensitive to the chemistry of the solution and the nature of the polymeric species. This sol-to-gel transition is irreversible, and there is little if any change in volume. At this stage, the one-phase liquid is transformed into a two-phase system. The gels formed in alcohol are known as alcogels, while those formed in aqueous solutions are hydrogels. After gelation, gels are generally subjected to an aging process for a period of hours to days. The gels are kept in sealed containers, and very little solvent loss or shrinkage occurs. Syneresis (spontaneous separation of liquid from a gel) is frequently observed at this stage. During aging of silica gels, the conden\-sation polymerization reactions continue, increasing the degree of crosslinking in the network and producing an increase in gel viscosity. The mechanical strength and rigidity of the gel thus increase during aging.

Drying. The drying process involves the removal of the liquid phase; the gel transforms from an alcogel to a xerogel. Low-temperature evaporation is frequently employed, and there is considerable weight loss and shrinkage. The drying stage is a critical part of the sol-gel process. As evaporation occurs, drying stresses arise that can cause catastrophic cracking of bulk materials. The ability to dry without cracking serves to limit the sizes of monolithic pieces produced by the sol-gel method.

Much is known about the drying process, as it has been characterized by physical measurements, mechanical properties, and chemical methods. A comprehensive analytical model has been developed that has identified the various details related to the generation of drying stresses in terms of their origin, their spatial distribution, and how stresses lead to cracking and warping in gels. This model indicates that as the gel network becomes stiffer, shrinkage stops and drying stresses arise because of spatial variability in strain rate; that is, gel shrinkage is faster at the exterior surface than in the interior. It is at this point that maximum stress is developed at the surface of the gel and cracking is prone to occur. The analysis indicates that stresses may be minimized by a low rate of solvent evaporation or the use of additives to improve solvent permeability.

Densification. The final stage of the sol-gel process is densification, where the gel-to-glass conversion occurs and the gel achieves the properties of the glass. As the temperature increases, several processes occur, including elimination of residual water and organic substances, relaxation of the gel structure, and, ultimately, densification. At moderate temperatures (above 150°C or 302°F but below the glass transition temperature T_g), skeletal densification occurs from both condensation polymerization processes and structural relaxation. These processes account for considerable shrinkage and reduction in surface area, but the resulting material is still quite porous (>25% porosity). At higher temperatures (starting at temperatures close to T_g, viscous sintering becomes the predominant factor in removing porosity. In this way, fully dense sol-gel glasses can be produced without melting, that is, merely by heating to temperatures close to the glass transition temperature in order to obtain viscous sintering. The precise temperatures involved for producing dense

sol-gel glasses depend upon the specific system. In the case of silica materials, temperatures corresponding to a viscosity of 10^{13} poises (≈1200°C or 2200°F) are generally required in order to obtain dense sol-gel silica glass. *See* GLASS; SINTERING.

Applications. The sol-gel process offers advantages for a broad spectrum of materials applications. The types of materials go well beyond silica and include inorganic compositions that possess specific properties such as ferroelectricity, electrochromism, or superconductivity. The most successful applications utilize the composition control, microstructure control, purity, and uniformity of the method combined with the ability to form various shapes at low temperatures. Films and coatings were the first commercial applications of the sol-gel process. The development of sol-gel-based optical materials has also been quite successful, and applications include monoliths (lenses, prisms, lasers), fibers (waveguides), and a wide variety of optical films.

One application of sol-gel materials that has emerged in recent years is that of bioactive glasses for biosensors. Based on the guest-host concept, a wide range of biomolecules, including globular and membrane-bound proteins and enzymes, have been immobilized in the pores of sol-gel glasses, resulting in a material that is not only bioactive but mechanically, thermally, and chemically stable. Moreover, silica-based sol-gel materials have excellent optical transparency, making optical sensors possible. Owing to the nature of the sol-gel process, the amorphous solid network forms around the biomolecules, confining them within the pores of the solid-state matrix. Although the relatively large biomolecules are immobilized within the inorganic network, small ions and molecules are able to diffuse into and out of the network. The sol-gel encapsulated biomolecules retain their native biological activity and spectroscopic properties, and they can respond to chemical changes in their environment. These materials are being widely explored for long-term biomolecular storage and for a new generation of biomedical and chemical sensors. *See* BIOSENSOR; MICROSENSOR.

Aerogels are a class of sol-gel materials that possess very low density and high surface area. Aerogels are mesoporous solids in which nanometer-scale solid domains are networked with a high volume of continuous pores. The synthesis of aerogels is based on traditional sol-gel chemistry and the resulting solids offer unusual properties of very large surface areas (on the order of 1000 m^2/g) and high pore volumes (greater than 95%). Aerogel materials have been known since the 1930s and widely used as catalysts, catalyst supports, and thermal insulation materials. In recent times, the applications of aerogels have expanded to electrochemical devices, low dielectric constant materials for integrated circuits, and the capture of cosmic dust particles.

A key step in the synthesis of aerogels is the process of removing the liquid phase from the pores of the wet gel. If the pore liquid is allowed to evaporate, the volume of the material shrinks significantly due to the capillary pressure and accompanying collapse of the gel network, resulting in a xerogel. On the other hand, if the pore liquid is removed using methods where capillary stresses do not develop or are minimized, the porosity of the wet gel can be maintained with relatively little volume shrinkage, and an aerogel is produced. Traditionally, aerogels are prepared through supercritical drying; however, ambient-pressure drying using nonpolar solvents with low surface tension has also been widely used, with the resulting materials called ambigels.

Bruce Dunn; Esther H. Lan

Bibliography. C. J. Brinker, D. E. Clark, and D. R. Ulrich (eds.), *Better Ceramics Through Chemistry I-IV*, vol. 32, 1984, vol. 73, 1986, vol. 121, 1988, vol. 180, 1990; C. J. Brinker and G. Scherer, *Sol-Gel Science: The Physics and Chemistry of Sol-Gel Processing*, 1990; L. C. Klein (ed.), *Sol-Gel Optics: Processing and Applications*, 1994; L. C. Klein (ed.), *Sol-Gel Technology for Thin Films, Preforms, Electronics and Specialty Shapes*, 1988; J. D. Mackenzie and D. R. Ulrich, *Ultrastructure Processing of Advanced Ceramics*, 1988; W. Yuming and A. Wagendristel, *An Introduction to the Physics and Technology of Thin Films*, 1994.

Solanales

An order of flowering plants, division Magnoliophyta, in the euasterid I group of the asterid eudicotyledons. The order consists of two large and three small families, of approximately 4275 species. Solanaceae and Convolvulaceae account for all but 25 of the species. Solanales are generally characterized by sympetalous flowers (in early developmental stages only in *Montinia*) with a superior ovary, and alternate leaves. *See* ASTERIDAE; FLOWER; FRUIT; LEAF; MAGNOLIOPHYTA; MAGNOLIOPSIDA.

Solanaceae, approximately 2600 species, are cosmopolitan herbs, shrubs, lianas, and trees, with branched hairs and often spines and alkaloids. The family is of great economic significance, yielding potatoes, tomatoes, and eggplant (*Solanum*), peppers (*Capsicum*), tobacco (*Nicotiana*) and many ornamentals. Deadly nightshade (*Atropa*), jimson weed (*Datura*), and henbane (*Hyoscyamus*) are well-known poisonous members of the family. *See* BELLADONNA; EGGPLANT; ORNAMENTAL PLANTS; PEPPER; POTATO, IRISH; SOLANALES; TOBACCO.

Convolvulaceae, approximately 1650 species, include herbaceous and woody members that are often climbers. The main economic crop is sweet potato (*Ipomoea*), but the family also includes ornamentals (morning glory, *Ipomoea*; and *Convolvulus*), noxious weeds (*Calystegia* and *Convolvulus*), and parasites (*Cuscuta*). *See* POTATO, SWEET; WEEDS.

Michael F. Fay

Solar cell

A semiconductor electrical junction device which directly and efficiently absorbs the radiant energy of sunlight and converts it into electrical energy. Solar cells may be used individually as light

detectors, for example in cameras, or connected in series and parallel to obtain the required values of current and voltage for electric power generation.

Most solar cells are made from single-crystal silicon and have been very expensive for generating electricity, but have found application in space satellites and remote areas where low-cost conventional power sources have been unavailable. Research has emphasized lowering solar cell cost by improving performance and by reducing materials and manufacturing costs. One approach is to use optical concentrators such as mirrors or Fresnel lenses to focus the sunlight onto smaller-area solar cells. Other approaches replace the high-cost single-crystal silicon with thin films of amorphous or polycrystalline silicon, gallium arsenide, cadmium sulfide, or other compounds.

Solar radiation. The intensity and quality of sunlight are dramatically different outside the Earth's atmosphere from that on the surface of the Earth. The number of photons at each energy is reduced upon entering the Earth's atmosphere due to reflection, to scattering, or to absorption by water vapor and other gases. Thus, while the solar energy at normal incidence outside the Earth's atmosphere is 1.36 kW/m^2 (the solar constant), on the surface of the Earth at noontime on a clear day the intensity is about 1 kW/m^2.

On clear days the direct radiation is about 10 times greater than the diffuse radiation, but on overcast days the sunshine is entirely diffuse. The mean annual solar energy falling on the Earth's surface varies greatly from one location to another. The sunniest regions of the globe receive about 2500 kWh/m^2 per year of total sunshine on a horizontal surface. The Earth receives about 10^{18} kWh of solar energy each year. The worldwide annual energy consumption is about 80×10^{12} kWh, so that from a purely technical viewpoint, the world energy consumption corresponds to the sunlight received on about 0.008% of the surface of the Earth. *See* SOLAR RADIATION.

Principles of operation. The conversion of sunlight into electrical energy in a solar cell involves three major processes: absorption of the sunlight in the semiconductor material; generation and separation of free positive and negative charges to different regions of the solar cell, creating a voltage in the solar cell; and transfer of these separated charges through electrical terminals to the outside application in the form of electric current.

In the first step the absorption of sunlight by a solar cell depends on the intensity and quality of the sunlight, the amount of light reflected from the front surface of the solar cell, the semiconductor band-gap energy which is the minimum light (photon) energy the material absorbs, and the layer thickness. Some materials such as silicon require tens of micrometers' thickness to absorb most of the sunlight, while others such as gallium arsenide, cadmium telluride, and copper sulfide require only a few micrometers.

When light is absorbed in the semiconductor, a negatively charged electron and positively charged hole are created. The heart of the solar cell is the electrical junction which separates these electrons and holes from one another after they are created by the light. An electrical junction may be formed by the contact of: a metal to a semiconductor (this junction is called a Schottky barrier); a liquid to a semiconductor to form a photoelectrochemical cell; or two semiconductor regions (called a *pn* junction).

The fundamental principles of the electrical junction can be illustrated with the silicon *pn* junction. Pure silicon to which a trace amount of a fifth-column element such as phosphorus has been added is an *n*-type semiconductor, where electric current is carried by free electrons. Each phosphorus atom contributes one free electron, leaving behind the phosphorus atom bound to the crystal structure with a unit positive charge. Similarly, pure silicon to which a trace amount of a column-three element such as boron has been added is a *p*-type semiconductor, where the electric current is carried by free holes. Each boron atom contributes one hole, leaving behind the boron atom with a unit negative charge. The interface between the *p*- and *n*-type silicon is called the *pn* junction. The fixed charges at the interface due to the bound boron and phosphorus atoms create a permanent dipole charge layer with a high

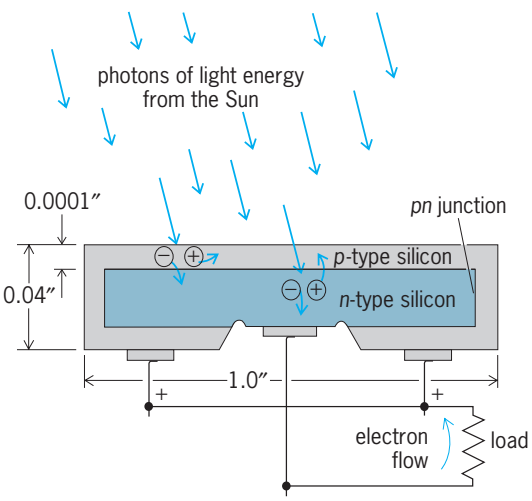

Fig. 1. Cross-sectional view of a silicon *pn* junction solar cell, illustrating the creation of electron pairs by photons of light energy from the Sun. 0.0001″ = 2.5 μm; 0.04″ = 1 mm; 1.0″ = 25 mm.

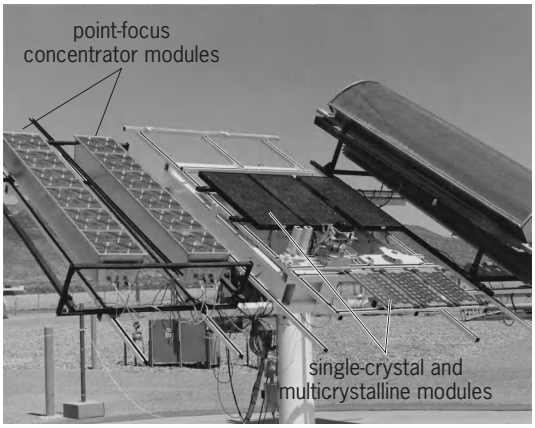

Fig. 2. Silicon solar cells assembled into modules and arrays to obtain higher voltage and current output. (*Sandia National Laboratories*)

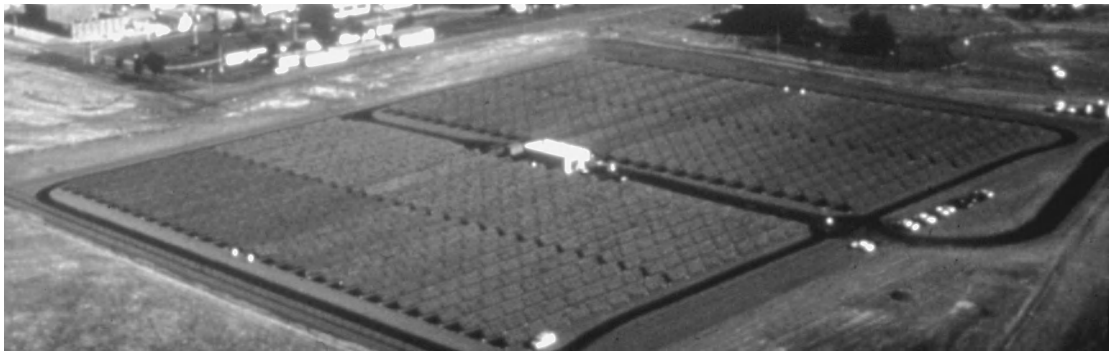

Fig. 3. A 2-MW photovoltaic array. (*Sacramento Municipal Utility District, California*)

electric field. When photons of light energy from the Sun produce electron-hole pairs near the junction, the built-in electric field forces the holes to the *p* side and the electrons to the *n* side (**Fig. 1**). This displacement of free charges results in a voltage difference between the two regions of the crystal, the *p* region being plus and the *n* region minus. When a load is connected at the terminals, electron current flows in the direction of the arrow, and electrical power is available at the load. *See* PHOTOVOLTAIC CELL; PHOTOVOLTAIC EFFECT; SEMICONDUCTOR; SEMICONDUCTOR DIODE.

Characteristics. The short-circuit current of a typical silicon *pn*-junction solar cell is directly proportional to light intensity and amounts to 28 mA/cm^2 at full sunlight (1000 W/m^2). The open-circuit voltage rises sharply under weak light and saturates at about 0.6 V for radiation between 200 and 1000 W/m^2. The maximum power output from the solar cell irradiated by full sunlight is about 11 mW/cm^2 at an output voltage of 0.45 V.

Under these operating conditions the overall conversion efficiency from solar to electrical energy is 11%. The output power as well as the output current is of course proportional to the irradiated surface area, whereas the output voltage can be increased by connecting cells in series just as in an ordinary chemical storage battery. Experimental samples of silicon solar cells have been produced which operate at efficiencies up to 18%, but commercial cell efficiency is around 10–12% under normal operating conditions.

Using optical concentration to intensify the light incident on the solar cell, efficiencies above 20% have been achieved with silicon cells and above 25% with gallium arsenide cells. The concept of splitting the solar spectrum and illuminating two optimized solar cells of different band gaps has been used to achieve efficiencies above 28%, with expected efficiencies of 35%. Thin-film solar cells have achieved between 4 and 9% efficiency and are expected in low-cost arrays to be above 10%. K. W. Mitchell

Arrays. Individual silicon solar cells or photovoltaic cells are about 100 cm^2 (15.5 in.^{2}) of surface area in size. At a 15% conversion efficiency, such a cell can deliver about 1.4 W at 0.5 V when in full sunlight. To obtain higher power and higher voltage, a number of cells must be assembled in panels or arrays (**Fig. 2**). Modules are of two basic types: flat plates (crystalline silicon or thin film) and concentrators. Cells may be connected in series to multiply their output voltage and in parallel to multiply their output current. Cells operated in series must be closely matched in short-circuit current since the overall performance of a solar cell array is limited by the cells having the lowest current.

Applications. Although the photovoltaic effect was discovered by A. C. Becquerel in 1839, practical solar cells made of silicon crystals were not developed until 1955. Beginning with *Vanguard 1*, launched in 1958, silicon solar cell arrays have become the almost exclusive power source for satellites. *See* SPACE POWER SYSTEMS.

Solar cell arrays have been used primarily to power small remote electrical loads that would otherwise be impractical or uneconomical to power by conventional means such as storage batteries or motor-generator sets. Solar cell arrays are sold worldwide to power such equipment as remote radio repeaters, navigational aids, consumer products, railroad signals, cathodic protection devices, and water pumps. Since most of the aforementioned uses require powering the load at times even when the Sun is not shining, electrical storage batteries are typically used in conjunction with solar cell arrays to provide reliable, continuous power availability.

Terrestrial uses take advantage of the ease of installation of photovoltaic arrays and the ease of adding or removing modules to meet changing electrical demands. These uses include agricultural, residential, commercial, and industrial applications (**Fig. 3**). When powering loads which require ac voltage, a static inverter is used to convert the dc voltage from the solar cell array into usable ac power. *See* SOLAR ENERGY. Margorie L. Tatro

Bibliography. M. A. Green, *Solar Cells*, 1982; T. Markvart (ed.), *Solar Electricity*, 2d ed., 2000; R. C. Neville, *Solar Energy Conversion: The Solar Cell*, 2d ed., 1995; L. D. Partain (ed.), *Solar Cells and Their Applications*, 1995; H. S. Rauschenbach, *Solar Cell Array Design Handbook*, 1980.

Solar constant

The total solar radiant energy flux (total irradiance) incident upon the top of the Earth's atmosphere at a standard distance (1 astronomical unit,

1.496×10^8 km or 9.3×10^7 mi) from the Sun. In 1980 it was discovered that the so-called solar constant actually varies with time, though only by small amounts, around a value of about 1366 $W \cdot m^{-2}$ (1.96 $cal \cdot cm^{-2} \cdot min^{-1}$). The solar radiant energy flux received at the Earth is more than 10^{10} times that of the next brightest star, and almost 10^6 times that of the full moon. This solar energy, averaged over its apparent variation (almost 7%) due to the Earth's orbital motions, maintains the climate on Earth. The solar constant represents a good measure of the solar luminosity (its total radiant energy), although anisotropic components (for example, a dark sunspot), which produce directional solar-constant variations, must be taken into account. *See* CLIMATOLOGY; EARTH ROTATION AND ORBITAL MOTION; PHOTOSPHERE; STAR.

An important test of the validity of any theory of stellar structure and evolution is the accuracy with which it predicts the steady luminosity of the Sun. However, the main astrophysical interest in variations of the solar constant concerns the new window it gives on transient processes in the solar interior. This new knowledge is welcome in view of the theoretical difficulties involved in calculating the solar neutrino flux successfully. *See* HELIOSEISMOLOGY; SOLAR NEUTRINOS; STELLAR EVOLUTION.

Causes of variability. Both expected and unexpected items contribute to the variability of the solar constant. Sunspots can produce deficits of up to a few tenths of 1% of the solar constant, on typical time scales of 1 week. Other surface manifestations of solar magnetic activity, faculae, contribute excesses rather than deficits. These features produce modulations at a roughly 27-day period due to solar rotation. Similar types of variability (due, for example, to so-called starspots) occur on other stars. These can have larger relative amplitudes. Convective motions and global oscillations of the solar interior, analogous to seismic waves on the Earth, produce variations of a few parts per million on time scales of a few hundred seconds; these waves are used to study solar interior structure and dynamics. Finally, and unexpectedly, there is an apparent 11-year sunspot cycle variation amounting to an approximately 0.1% increase of the solar constant during the sunspot maxima (see **illus.**). This long-term effect has the opposite dependence from that found for individual sunspots, which block the solar radiant energy and cause decreases rather than increases. If such small changes in solar energy flux affect the Earth's climate directly or indirectly, then this solar-cycle variation is consistent with an association noted some time ago: the so-called Maunder minimum period of greatly reduced sunspot activity in the seventeenth century corresponded with decades of cool climate in Europe, known as the Little Ice Age. *See* CLIMATE HISTORY; SUNSPOT.

Measurement. To measure the variability of the solar constant requires the deployment of a stable radiometer in space, above the Earth's atmosphere. The original high-quality observations came from

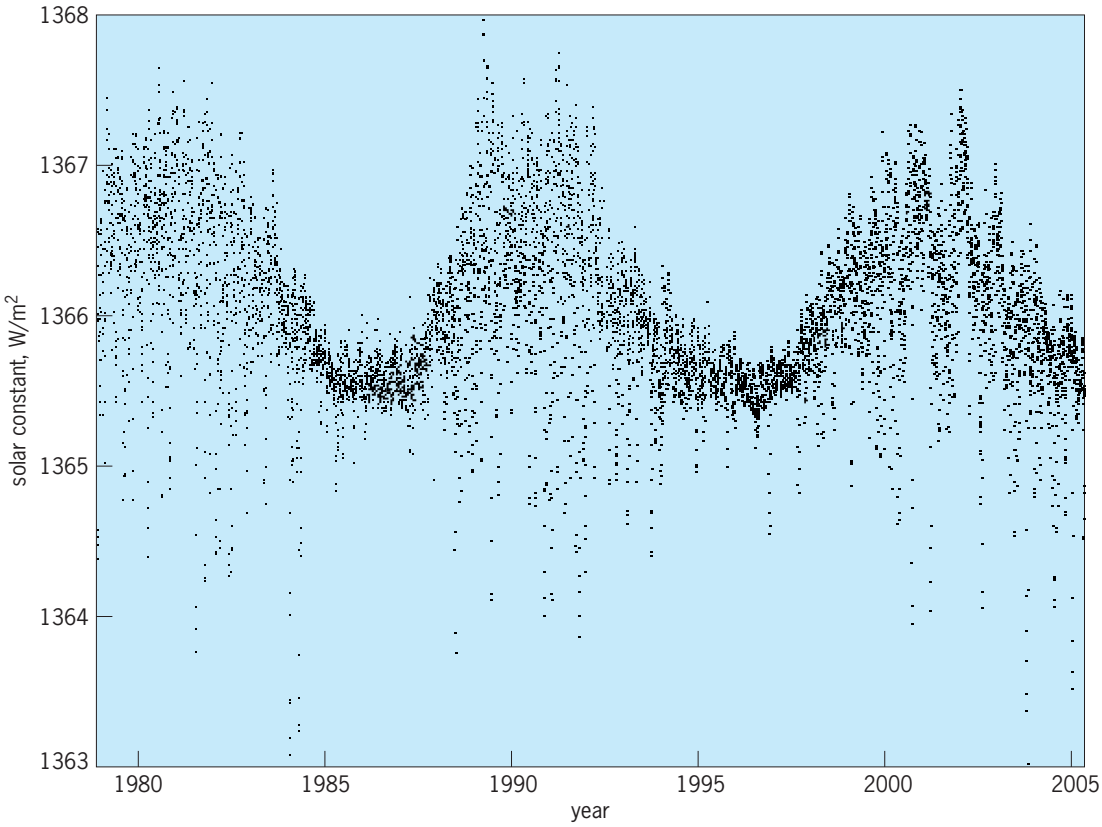

Composite time series of daily values of the solar constant as obtained from radiometers in space. (*Physikalisch-Meteorologisches Observatorium Davos/World Radiation Center; adapted from http://www.pmodwrc.ch/pmod.php?topic=tsi/composite/SolarConstant*)

NASA's *Solar Maximum Mission* (1980–1989), and several other spacecraft have since contributed to the composite time series shown in the illustration. To measure the total solar irradiance (solar constant) typically requires an active cavity radiometer making an absolute, not relative, measurement. Because of the difficulty of this kind of measurement, it is not presently possible to infer the presence of any gradual variation extending to time scales exceeding a single sunspot cycle. On shorter time scales the measurements are very precise and can reveal the parts-per-million variations resulting from solar global oscillations. *See* RADIOMETRY; SUN. Hugh S. Hudson

Bibliography. J. A. Eddy, R. L. Gilliland, and D. V. Hoyt, Changes in the solar constant and climatic effects, *Nature*, 300:689–693, 1982; C. Fröhlich, Long-term behaviour of space radiometers, *Metrologia*, 40:60–64, 2003; J. Pap and P. Fox (eds.), Solar variability and its effects on climate, *Geophys. Monogr. Series*, 141, 2004.

Solar corona

The outer atmosphere of the Sun. The corona is characterized by extremely high temperatures of several million kelvins, which cause it to extend far above the denser surface regions of the Sun. Coronal gas is constrained to follow the magnetic field of the Sun, forming it into the shapes seen during a solar eclipse, with a coronagraph, or in x-rays (see **illustration**). These shapes include long streamers that penetrate interplanetary space, loop-like tubes over the strongest fields, and vast regions of very low density called coronal holes. Near the north and south poles the coronal polar rays resemble the pattern formed by iron filings near a bar magnet, but at midlatitudes the more complicated fields of active regions dominate. The corona is hot enough to emit x-rays, and x-ray telescopes in space can form images of the corona. Such images display the magnetic loops connecting bright regions in the lower corona, while visible-light observations can detect the corona above the edge of the solar disk by carefully blocking out the strong photospheric light with an occulting disk (see illustration). *See* CORONAGRAPH; ECLIPSE; SOLAR MAGNETIC FIELD; X-RAY ASTRONOMY.

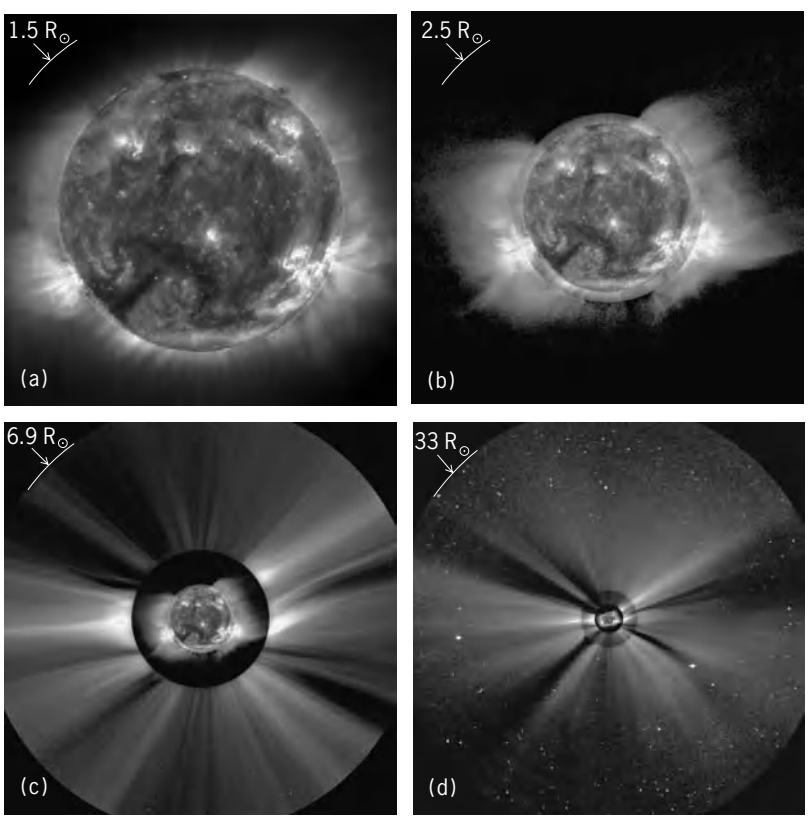

Global view of the solar corona, from observation by instruments on the *Solar and Heliospheric Observatory* (*SOHO*) on April 17, 1998. (*a*) On-disk images taken by the Extreme-Ultraviolet Imaging Telescope (EIT). (*b*,*c*,*d*) A set of nested visible-light images taken by the Large-Angle Spectroscopic Coronagraph (LASCO), showing the off-limb corona out to 30 solar radii ($R_\odot$). The active-region corona is shown in the on-disk images, while the streamer structure of the outer corona far above the solar surface is seen in the coronagraph images. (*Courtesy LASCO/EIT Consortium*)

Coronal holes are the source of solar matter streaming into interplanetary space. These streams often interact with the Earth's magnetic field to cause auroras and other geophysical phenomena. The interaction between the hot coronal plasma and the evolving magnetic fields in the corona sometimes leads to instabilities which eject large amounts of matter into space. These coronal mass ejections (CMEs) can have geophysical effects similar to those of the streams and ejecta. *See* AURORA; MAGNETOSPHERE; SOLAR WIND.

From a minimum in the lower atmosphere (approximately 4000 K or 7000°F), the temperature rises to values in the corona of $1–5 \times 10^6$ K ($2–9 \times 10^6$ °F) caused by dissipation of mechanical or magnetic energy produced by turbulent flows in the lower atmosphere. The mechanisms by which the corona is heated are uncertain, although results from the Transition Region and Coronal Explorer (TRACE) satellite indicate that the complex rearrangement of magnetic fields in the corona is largely responsible for coronal dynamics, via the rapid release of energy stored in the hot magnetized plasma. *See* SUN. Leon Golub

Bibliography. A. H. Gabriel (ed.), *The Solar Corona*, Elsevier Science, 1994; L. Golub and J. M. Pasachoff, *Nearest Star*, Harvard University Press, 2001; L. Golub and J. M. Pasachoff, *The Solar Corona*, Cambridge University Press, 1997; International Astronomical Union, *IAU Colloquium on Coronal Structures*, 1993; A. I. Poland (ed.), *Coronal and Prominence Plasmas*, NASA Conf. Pub. 2442, 1986; J. B. Zirker, *Journey from the Center of the Sun*, Princeton University Press, 2001.

Solar energy

The energy transmitted from the Sun. The upper atmosphere of Earth receives about 1.5×10^{21} watt-hours (thermal) of solar radiation annually. This vast

amount of energy is more than 23,000 times that used by the human population of this planet, but it is only about one two-billionth of the Sun's massive outpouring—about 3.9×10^{20} MW. *See* SUN.

The power density of solar radiation measured just outside Earth's atmosphere and over the entire solar spectrum is called the solar constant. According to the World Meteorological Organization, the most reliable (1981) value for the solar constant is 1370 ± 6 W/m^2. Of this power, 8% is in the ultraviolet wavelengths, 47% in the visible spectrum, and 45% in the infrared region. The solar constant actually is not a true constant, but is subject to a small continuous variation due to the shape of Earth's orbit, amounting to −3.3% from average about July 5, when Earth is at its greatest distance from the Sun, and +3.4% about January 3, when Earth is closest to the Sun. *See* SOLAR CONSTANT.

Solar radiation is attenuated before reaching Earth's surface by an atmosphere that removes or alters part of the incident energy by reflection, scattering, and absorption. In particular, nearly all ultraviolet radiation and certain wavelengths in the infrared region are removed. However, the solar radiation striking Earth's surface each year is still more than 10,000 times the world's energy use. Radiation scattered by striking gas molecules, water vapor, or dust particles is known as diffuse radiation. Clouds are a particularly important scattering and reflecting agent, capable of reducing direct radiation by as much as 80 to 90%. Because cloud distributions and types are highly variable, these reductions are quite unpredictable. The radiation arriving at the ground directly from the Sun is called direct or beam radiation. Global radiation is all solar radiation incident on the surface, including direct and diffuse. **Figure 1** illustrates the events that affect solar radiation as it passes through Earth's atmosphere. *See* SOLAR RADIATION.

The amount of atmospheric absorption and scattering of solar radiation is a function of the effective distance (depending on atmospheric thickness and content) through which the radiation travels. Spectral irradiance is a function of air mass. The term air mass zero is used to represent zero mass of air through which the Sun's rays must pass (the solar intensity outside Earth's atmosphere). Radiation arriving at the surface of Earth can be measured against that reaching the fringes of the atmosphere, where there is no air. The light of the Sun directly overhead at a 90° solar altitude, or zenith, at sea level is said to pass through an air mass of 1, providing an average peak intensity of 1 kW/m^2. The solar declination, the angle between the Earth-Sun line and Earth's equatorial plane, varies as Earth circles the Sun, and this affects the solar radiation (insolation) reaching the ground. Intensity weakens at Sun angles approaching the horizon since the rays have more atmosphere to penetrate. The amount of solar energy incident on a horizontal surface at sea level ranges up to 2400 Btu/ft^2-day (7 kWh/m^2-day). The sites between latitudes 35°N and 35°S receive between 2000 and 3500 h of sunshine per year; higher latitudes receive less.

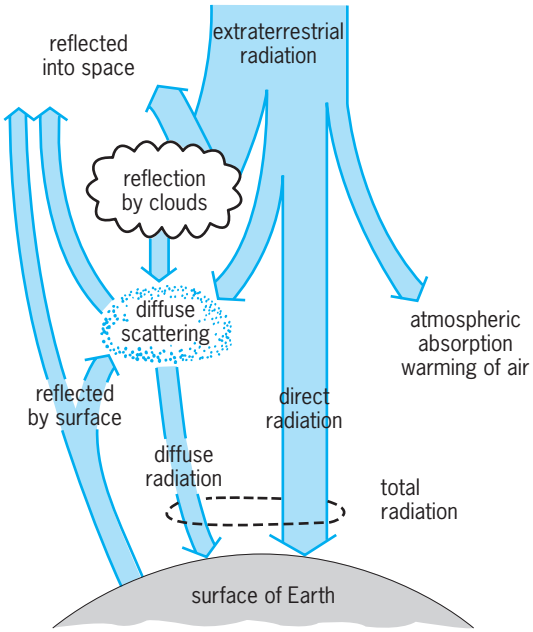

Fig. 1. Direct, diffuse, and total solar radiation. (*After W. C. Dickinson and P. N. Cheremisinoff, eds., Solar Energy Technology Handbook, pt. A: Engineering Fundamentals, Marcel Dekker, 1980*)

See ABSORPTION OF ELECTROMAGNETIC RADIATION; INSOLATION; SCATTERING OF ELECTROMAGNETIC RADIATION.

Solar research and technology development aim at finding the most efficient ways of capturing low-density solar energy and developing systems to convert captured energy to useful purposes. Also of significant potential as power sources are the indirect forms of solar energy: wind, biomass, hydropower, and the tropical ocean surfaces. With the exception of hydropower, these energy resources remain largely untapped. *See* ENERGY SOURCES.

Five major technologies using solar energy are being developed. (1) The heat content of solar radiation is used to provide moderate-temperature heat for space comfort conditioning of buildings, moderate- and high-temperature heat for industrial processes, and high-temperature heat for generating electricity. (2) Photovoltaics convert solar energy directly into electricity. (3) Biomass technologies exploit the chemical energy produced through photosynthesis (a reaction energized by solar radiation) to produce energy-rich fuels and chemicals and to provide direct heat for many uses. (4) Wind energy systems generate mechanical energy, primarily for conversion to electric power. (5) Finally, a number of ocean energy applications are being pursued; the most advanced is ocean thermal energy conversion, which uses temperature differences between warm ocean surface water and cooler deep water to produce electricity.

Solar Energy Systems in Buildings

Systems employing solar energy are classified as passive or active systems.

Passive systems. Passive solar energy systems use the building itself in combination with solar energy

as an energy-saving system. A passive building uses its environment, the features of its site, its structure, and its materials so that it requires much less fuel energy. Passive lighting, usually called daylighting, illuminates the interior of a building as a replacement for or supplement to electric lighting.

Heating systems. Most solar thermal systems consist of three basic parts: a collector, a storage medium, and a distribution system. In a passive solar heating system, the collector is usually the building itself. In direct systems, the Sun's rays enter the building, usually through windows, and heat the room or space in direct sunlight. Excess heat can be vented off or stored in the mass (walls and floor) of the building for later use. The entering solar energy is distributed through a combination of reradiation from the mass and natural convection of warm air between rooms. Indirect heating systems use one or more rooms as a collector, which is combined with a storage mass that separates that room from the rest of the building. The Sun's energy is reradiated and convected into the building after passing through the storage mass, which can store heat during the day for use at night. Popular passive heating systems include greenhouses and sunspaces.

Cooling systems. One type of passive solar cooling occurs when the cycle established in a passive heating system is reversed, allowing the building to radiate heat to the sky at night. As the mass is cooled, it lowers the temperature inside the house. Then, during the daytime, the cool mass helps maintain comfortable conditions in summer without using air conditioning. Sometimes, the radiant cooling process is assisted by placing a storage mass, such as bags of water, on top of the building. The storage mass is covered to prevent heating during the day but is exposed to the night sky.

Another form of passive cooling is natural ventilation, such as is caused by air currents entering through open windows or doors. Ventilation also occurs as heated air rises to the ceiling. If openings are placed near the ceiling of a room and at floor level, warm air will rise and exit through the top openings, drawing cooler air in through the lower openings.

Natural cooling also occurs when a building is shaded during the summer. If shade trees are deciduous (shed their leaves during the winter), they will not interfere with a passive heating system. The house can also be self-shading through the use of overhangs and extended walls.

Daylighting systems. In commercial, institutional, and industrial buildings, the largest energy use is often electric lighting. Daylighting can be used to replace or supplement electric lighting in buildings occupied primarily during the daytime. Windows and other types of apertures allow daylight to enter the building, and special control systems regulate the electric lighting to turn it off when there is sufficient daylight in the building. Unlike other passive systems, daylighting can occur on overcast days because the daylight in the building does not depend on the direct use of the Sun.

It is possible to use a variety of materials to effect passive solar performance, such as antireflective window coatings, insulating window shades and curtains, or modular water walls. Some window coatings carry an electric current that can change window transmittance in response to weather conditions. In addition, there are several chemical compounds that can be incorporated either into window glass or between panes to increase a window's storage or insulating capacity.

Active space conditioning and hot water systems. Solar active space conditioning and hot water systems use mechanical means to collect, store, and distribute solar energy to heat and cool buildings and to heat water for domestic or commercial use. In general, water heating, the simplest system, is characterized by the smallest collector array, the simplest design, and the lowest cost; retrofit can be accomplished easily. Space heating requires a large collector array, has the largest heat storage requirement, and is relatively expensive. Space heating and cooling is the most complex of the solar active systems, requiring the largest collector array, the highest temperatures, and a complicated mechanical process.

Active systems generally encompass four parts: collection (including storage), conversion, distribution, and control (**Fig. 2**).

Collection. Collectors can be flat-plate, evacuated-tube, or concentrating. In most active systems, flat-plate collectors are used in one or more modules. The Sun's radiation is absorbed by the collector and is converted to heat in a liquid (water or glycol) or gas (air) heat-transfer medium. The heat is transported by pumps or fans to the conversion and distribution systems for immediate use or is stored for later use. The heat-transfer medium in liquid systems usually is water, which must be protected from freezing in cold climates. An antifreeze can be added to the water, or the collectors can be drained when the outdoor temperature approaches freezing. If an antifreeze is added, a heat exchanger is used to isolate the water in the collectors from the internal building water.

Conversion. For both space heating and cooling and domestic hot water, a conversion system is required to convert the thermal energy to useful heat. The thermal energy may also be converted to mechanical work that can subsequently drive a conventional vapor compression cooling unit. Thermal energy can also be used directly, through absorption refrigeration, for space cooling.

Desiccant cooling for air conditioning uses recirculated, dehumidified air as well as evaporative cooling to produce cool, comfortable air for the building (**Fig. 3**). The desiccant material is regenerated and reused when water is driven off by solar-heated air.

Distribution. The distribution system transports the hot or cold fluid from the conversion system to the point of use in the building. Generally, one distribution system for both the solar system and auxiliary (backup) system is used to reduce duplication.

Control. The control system collects temperature information from inside and outside the building, processes the information, and sends commands as

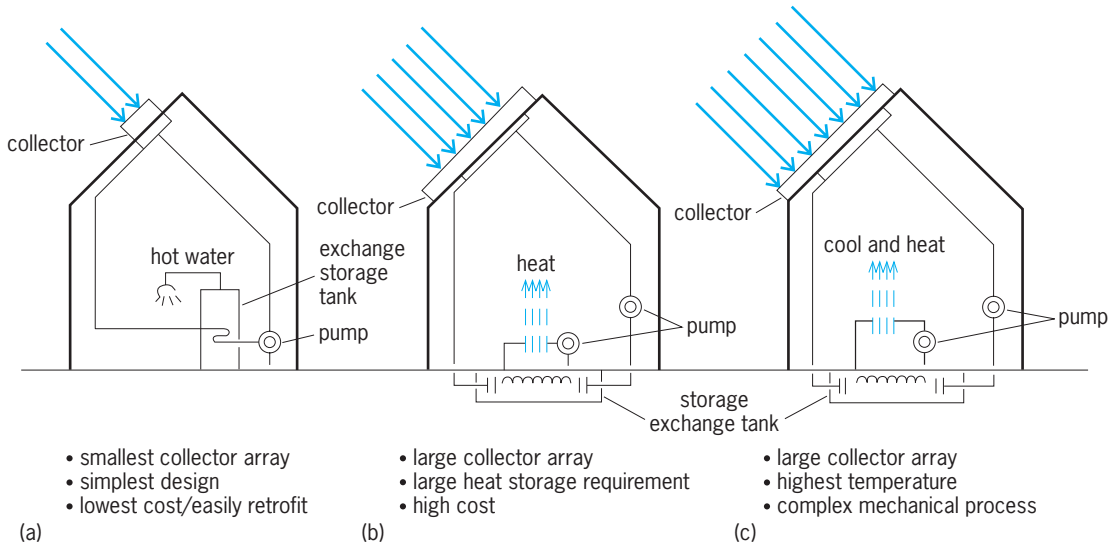

Fig. 2. Building space conditioning and hot water systems. (*a*) Water heating. (*b*) Space heating. (*c*) Space heating and cooling. (*After DOE 5-Year Research Plan for National Active Solar Heating and Cooling Program, U.S. Department of Energy, 1983*)

conditions dictate. Temperature sensors provide data to a network of relays or a microprocessor, which processes the data and sends commands to operate the collection, conversion, or distribution system.

Research has included efforts to improve the cost of collectors by testing lightweight flat-plate collectors using thin-film plastics rather than conventional metal-glass materials, identifying desiccants and other cooling system materials, and collecting data from a network of building sites on performance, reliability, and durability of active solar systems. *See* SOLAR HEATING AND COOLING.

One type of solar collector that could have wide applicability is the salt-gradient solar pond (**Fig. 5**). A thin layer of fresh or low-salinity water covers a deeper layer of water in which a salinity gradient is created (that is, salt concentration increases with depth). The bottom layer is usually at or near salt saturation. Sunlight passing through the water is absorbed and heats the bottom layer. The salt concentration in the deeper water serves to increase the liquid density and thereby prevents the natural convection that would normally mix both the warm lower layers and the cool layers above. The heat

Solar Thermal Technology

Solar energy can be converted to useful work or heat by using a collector to absorb solar radiation, allowing much of the Sun's radiant energy to be converted to heat. This heat can be used directly in residential, industrial, and agricultural operations; converted to mechanical or electrical power; or applied in chemical reactions for production of fuels and chemicals.

Collector systems. The collector system contains a concentrator and a receiver. The concentrator redirects and focuses sunlight on the receiver by using mirrors or lenses, and the receiver absorbs solar radiation and converts it to heat. Solar collectors are of two basic classes: nonfocusing and focusing. They are further distinguished on the basis of their concentrator optical properties (**Fig. 4**), and the operating temperature that can be obtained at the receiver (**Table 1**).

Nonfocusing collectors. Nonfocusing collectors are generally kept fixed (nontracking), yielding low-temperature outputs. Two types of this class of collectors are the solar pond and the evacuated-tube collector.

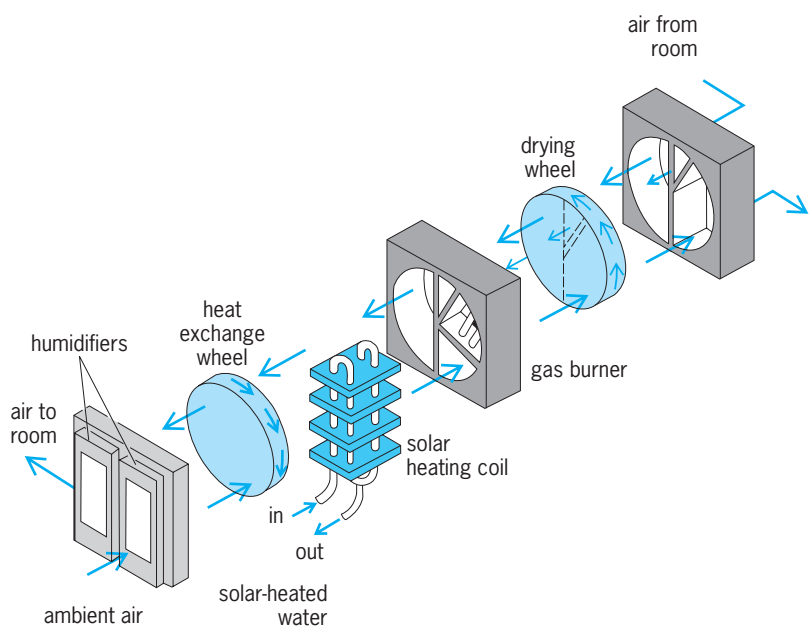

Fig. 3. Desiccant solar cooling system with gas backup in recirculation mode. (*After R. S. Barlow, Analysis of the Adsorption Process and of Desiccant Cooling Systems: A Pseudo-Steady Model for Coupled Heat and Mass Transfer, SERI/TR-631-1330, Solar Energy Research Institute, 1982*)

TABLE 1. Typical operating temperatures for solar thermal collectors

Type of collector	Operating temperature range, °F (°C)
Central receiver	400–2500 (200–1400)
Point focus (parabolic dish and Fresnel lens)	550–2500 (300–1400)
Line focus (parabolic trough and Fresnel lens, also multiple reflector)	160–600 (70–320)
Evacuated tube	120–350 (49–177)
Flat-plate	50–160 (10–71)
Solar pond	50–180 (10–82)

is thus trapped and stored at the bottom and can be extracted for use. Care must be taken to prevent the extraction process from upsetting the stability of the layers. Wind effects on the surface tend to cause mixing that is detrimental to pond performance, and floating or suspended debris will reduce the pond's transparency to sunlight.

Several solar ponds have been constructed and are operating around the world. The bottom hot brine, with temperatures near 180°F (82°C), can be used directly or via a heat exchanger to provide low-temperature heat for many industrial applications, for district space heating, or for electricity generation.

The evacuated-tube collector, sometimes incorporated with a reflective backing to provide greater solar concentration, can provide temperatures of up to 300°F (150°C), the temperature required to produce low-grade, industrial process steam.

Focusing collectors. This major class of collectors includes the distributed receiver system and the point-focus central receiver system.

Distributed receiver systems, such as line-focus (for example, parabolic trough) and point-focus (for example, parabolic dish), are commonly considered for remote community-power systems, military applications, individual factory or commercial building systems, or agricultural applications. These collectors must always point toward the Sun and cannot make use of diffuse and reflected light, which can account for a 10 to 20°F (5 to 10°C) temperature rise in flat-plate collectors.

1. Line-focus. The line-focus collector allows the concentrating, mirrored, reflector surface to follow the Sun by rotating about one axis (either east-west or north-south). Radiation is reflected onto an absorber tube (the receiver). The parabolic trough collector (**Fig. 6**) is the most common type of line-focus collector. It offers a wide range of operating temperatures from approximately 160 to 600°F (70 to 320°C). Therefore, it has been used for water heating below 212°F (100°C), steam production at 212 to 600°F (100 to 320°C), and power production at up to 600°F (320°C) using heat-transfer coils.

2. Point-focus. This major class of collectors offers the highest optical and thermal performance of any collector type because of its ability to track the Sun over its full range of diurnal motion and because of the relatively small area of the absorber used in the receiver unit. Parabolic dish–shaped collectors concentrate direct solar radiation to a point at the focus of the parabola. Each dish is a complete power-producing unit (**Fig. 7**), which can function either as an independent system or as part of a group of modules linked to form a larger system. A single parabolic-dish module can achieve fluid temperatures from 550 to 2500°F (300 to 1400°C) and can efficiently produce up to 25 kW of electricity.

The largest subsystem of a parabolic-dish module is the concentrator, a shallow dish with a reflective surface that tracks the Sun to focus sunlight on a receiver. Tracking along two axes ensures maximum solar energy collection during the day. Because the receiver absorbs concentrated solar radiation at very high temperatures (up to 2500°F or 1400°C), the system can be used for high-temperature chemical reactions or the production of steam, electric power, or fuels and chemicals. In many system designs, the engine for thermal-to-mechanical energy conversion is mounted with the receiver at the focus.

3. Central receiver system. The central receiver system concept consists of flat or slightly curved mirrors (heliostats) that rotate about two axes to reflect direct solar radiation onto a tower-mounted receiver. The absorbing surface can be a cylinder (external receiver) or a planar surface contained within a cavity (cavity receiver). Temperatures in excess of 2500°F (1400°C) and more and pressures of 1015 lb/in.^{2} absolute (7 megapascals) can be generated at the receiver. The central receiver–collector concept appears to hold promise for producing solar thermal electric power in capacities from approximately 1 to 100 MWe. **Figure 8** shows the mode of operation of a typical point-focus central receiver system for power generation using conventional steam technology.

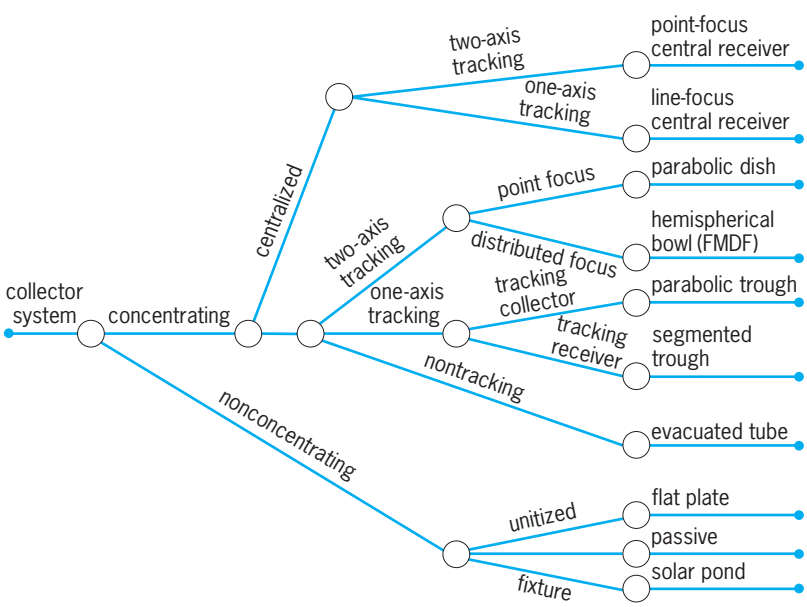

Fig. 4. Solar collector classification based on optical properties. (*After J. P. Thornton, A Comparative Ranking of 0.1-10 MWE Solar Thermal Electric Power Systems, SERI/TR-351-461, Solar Energy Research Institute, 1980*)

Conversion systems. Solar thermal heat can be converted directly to electricity, converted to mechanical and then electrical power, or used in the production of chemicals and fuels via an appropriate thermodynamic engine cycle. Three thermodynamic cycles have been considered for thermal-to-mechanical energy conversion: the Stirling cycle engine with closed-cycle operation; the Brayton cycle engine with either open- or closed-cycle operation; and the conventional Rankine cycle. Key concerns are the achievement of high engine performance and reliability and identification of operating requirements at the interface of the engine and receiver. Thermal energy conversion using parabolic-dish technology can be made at the point of heat generation, or at some distant location by piping heated fluid elsewhere for heat extraction and conversion. *See* BRAYTON CYCLE; RANKINE CYCLE; STIRLING ENGINE.

Storage. A solar energy system is normally designed to be able to deliver useful heat for 6 to 10 h a day, depending on the season and weather. Storage capacity in the solar thermal system is one way to increase a plant's operating capacity.

There are four primary ways to store solar thermal energy: (1) sensible-heat-storage systems, which store thermal energy in materials with good heat-retention qualities; (2) latent-heat-storage systems, which store solar thermal energy in the latent heat of fusion or vaporization of certain materials undergoing a change of phase; (3) chemical energy storage, which uses reversible reactions (for example, the dissociation-association reaction of sulfuric acid and water); and (4) electrical or mechanical storage, particularly through the use of storage batteries (electrical) or compressed air (mechanical). *See* ENERGY STORAGE.

Applications. Solar thermal systems that are used in electric power applications include the small, distributed power systems (typically 1000 kWe or less) and the large, central power systems of 10,000 kWe or more. Small-power uses include those for remote communities; military applications; and individual factory, commercial building, or agricultural applications. For these categories, the distributed receiver systems are most common. For generating electric power to distribute through a utility grid or to produce fuels and chemicals, the central receiver power and parabolic-dish systems are the most appropriate.

Industrial process heat is the thermal energy used directly in the preparation and treatment of materials and goods manufactured by industry. Hot water at temperatures between 120 and 200°F (50 and 100°C) can be supplied by directly heating water in the absorber tubes of evacuated-tube-type collectors or linear concentrating collectors. Hot air for industrial drying operations below about 350°F (180°C) can be supplied by collector systems designed to handle air as a circulating fluid or to circulate fluid pumped through an air-liquid heat exchanger. Concentrating solar thermal collector systems, both distributed and centralized, are commonly considered for steam generation at temperatures between 300 and 750°F (150 and 400°C) for use as direct heat in various manufacturing industries. Potential uses also include the production of ammonia in large agricultural fertilizer plants and enhanced oil recovery.

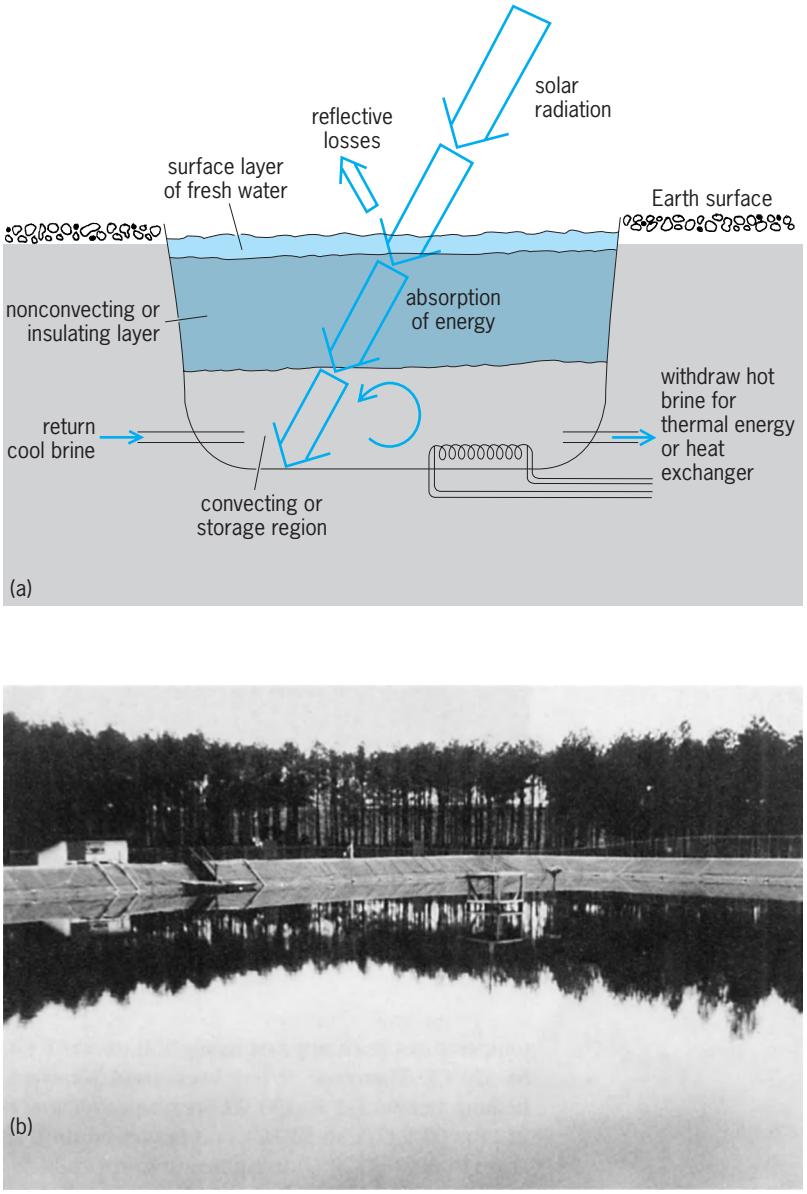

Fig. 5. Salt-gradient solar pond. (*a*) Schematic of the collection process (*after G. Franta et al., Solar Design Workbook, SERI/TR-62-308, Solar Energy Research Institute, 1981*). (*b*) Experimental pond (*Solar Energy Research Institute*).

Photovoltaics

Photovoltaic systems convert light energy directly to electrical energy. Using one of the most versatile solar technologies, photovoltaic systems can, because of their modularity, be designed for power needs ranging from milliwatts to megawatts. They can be used to provide power for applications as small as a wristwatch to as large as an entire community. They can be used in centralized systems, such as a generator in a power plant, or in dispersed applications, such as in remote areas not readily accessible to utility grid lines.

Historically, photovoltaic systems have been used

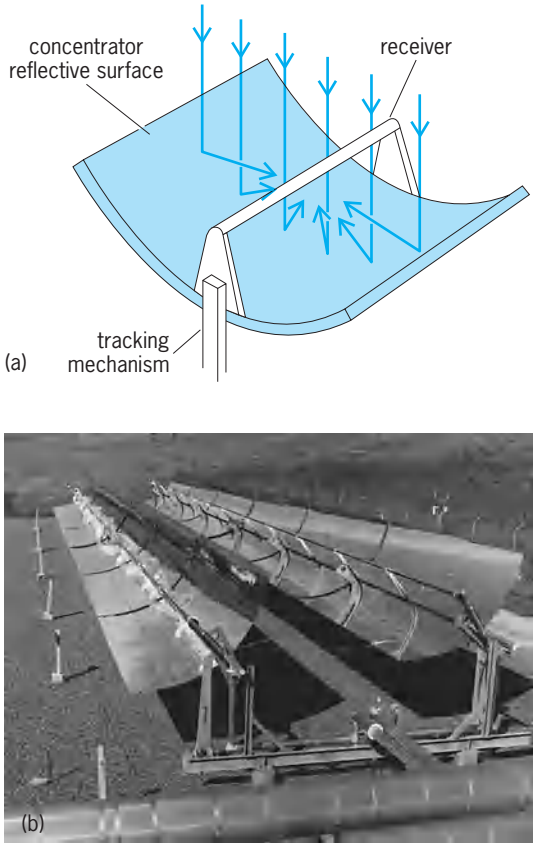

Fig. 6. Parabolic trough system. (*a*) Schematic of the components (*after Jet Propulsion Laboratory, Fact Sheet on Parabolic Trough Systems: A Solar Thermal Concentrating Collector Concept, 5106-22/3, 1982*). (*b*) Typical system (*Solar Energy Research Institute*).

where traditional electric power was unavailable or too expensive, providing power for satellites, remote cabins and houses, communication stations, observatories, water pumps, mobile military applications, and even for entire villages. Electricity has been brought to some areas of the world and has replaced fuel- and maintenance-reliant diesel generators. Photovoltaic systems are also used as power sources for hand-held calculators, watches, and portable television sets.

System components. The solar cell, composed principally of semiconductor material, is the basic component of the photovoltaic system. Though many materials and structures are being investigated, most commercial cells use a pure single-crystal silicon wafer that is doped *p*-type (that is, a small amount of an appropriate impurity is added, inducing extra unfilled bonds, or holes, in the wafer). Another impurity is diffused into a shallow region on the top of the wafer, promoting extra unbonded electrons, causing that portion to become *n*-type. A potential barrier, or electric field, is set up at the junction of the electrically different semiconductor layers. When photons strike the cell, they create ion pairs (negatively charged electrons and positively charged holes). These charges move around the silicon lattice structure until they recombine or cross the junction to be used in the photovoltaic process. The potential barrier segregates the ion pairs to opposite ends of the cell. The separated charges set up a potential (voltage), which is used to drive a current (composed of the light-generated electrons) through electrical contacts attached to both faces of the cell and through an external circuit (**Fig. 9**). *See* PHOTOVOLTAIC CELL; SEMICONDUCTOR.

The typical crystalline silicon cell produces a characteristic open-circuit voltage of about 0.5 V direct current (dc). The current generated depends on cell efficiency, cell area, and incident sunlight. A typical cell 4 in. (10 cm) in diameter can produce about 1 W of power under conditions where the Sun is directly overhead in a clear sky (one-sun conditions).

For larger voltages or currents, photovoltaic modules are formed by mounting groups of cells together on a rigid plate and interconnecting them in series or parallel. Stringing them in series increases the voltage; connecting them in parallel increases the current (**Fig. 10**). To further increase voltage or current output, modules may be grouped in parallel or strung in series to form photovoltaic arrays. The arrays, in turn, may be interconnected to form array fields. As

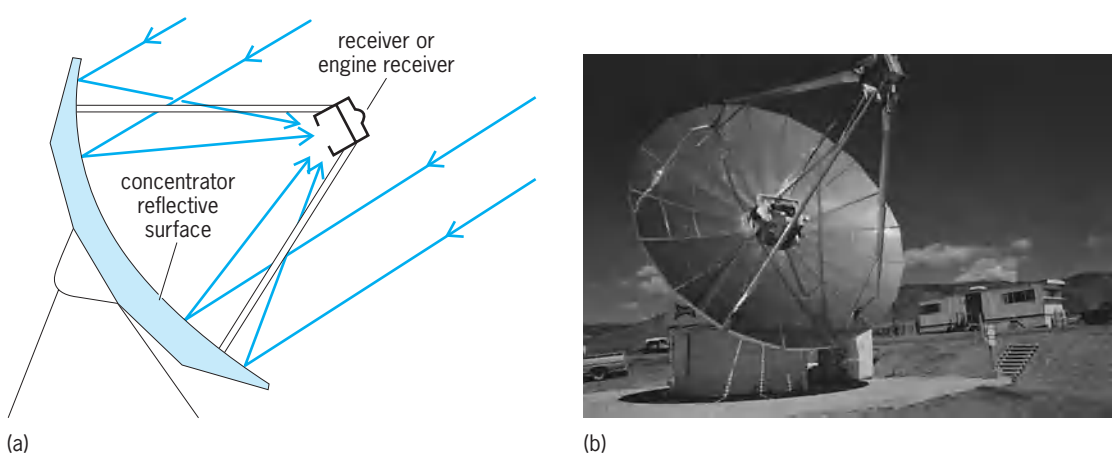

Fig. 7. Point-focus collector. (*a*) Schematic of collection process (*after Jet Propulsion Laboratory, Fact Sheet on Parabolic Dish Systems: A Solar Thermal Concentrating Collector Concept, 5106-22/2, 1982*). (*b*) Typical dish-receiver and engine module (*Jet Propulsion Laboratory*).

with cells, series combinations are voltage-additive, and parallel arrangements are current-additive.

Most commercial modules and arrays are based upon flat-plate collectors (which look like windows with cells behind them) that are able to use either direct or diffuse sunlight effectively. They can be mounted on a stationary surface, such as a roof top, and faced in a southerly direction (for systems in the Northern Hemisphere) at an appropriate angle. Concentrating systems, which use lenses or mirrors to focus sunlight on each cell, are also being used. Concentration increases the power produced by each cell. Since such systems require direct sunlight, they must use mechanisms to track the Sun. Also there must be provision for removing heat, since high temperatures can reduce cell output power. *See* PHOTOELECTRIC DEVICES.

Many electric power applications require alternating—current (ac) electricity. To obtain ac, a power-conditioning system is needed to convert the photovoltaic-generated dc to ac. A typical power-conditioning system consists of an inverter for converting the dc output to ac; filtering circuits to remove harmful or unwanted signals; a logic subsystem to control the on-off cycling of the photovoltaic system; a dc contactor to disconnect the dc source from ac lines during outages; and an isolation transformer to isolate the dc source voltage from the ac line voltage (**Fig. 11**).

When the Sun is not shining, at night or because of inclement weather, no power is generated by photovoltaic collectors. For this reason, and because the photovoltaic system must be able to start up and shut down smoothly and be able to satisfy user needs, a backup or storage system is often required. Systems installed on private residences usually can tie into the local utility grid, using it as an energy receptacle or resource, as conditions warrant. Other options include battery storage, pumped hydroelectric storage, and flywheel energy storage. A diesel engine-generator also can be used. *See* FLYWHEEL; GENERATOR.

Efficiencies. Under ideal conditions the theoretical maximum efficiency of a single-crystal silicon cell is about 25%, while the practical maximum efficiency hovers around 22%. Of the factors that combine to limit the efficiency, the most important are due to the inherent physics of the situation, the interaction between the incident photons and the silicon solar cell. A great percentage of the incident photon radiation is not energetic enough to be absorbed by the cell to create electron-hole pairs. An even greater percentage of the radiation is too energetic—the excess photon energy is not used in the conversion process.

By 1984, the efficiency of a typical commercial solar cell made from single-crystal silicon had increased by about a factor of two. Commercial cells became available that could convert from 13 to 15% of the incident radiant energy to electrical energy (at 77°F or 25°C and 1 kW/m^2 incident radiation). Single-crystal silicon cells have been produced in the laboratory with efficiencies approaching 19% under one-sun conditions, and further improvements are anticipated. Strides have also been made to increase the efficiencies of modules and arrays by decreasing intercell connection resistances and by employing protective circuitry that uses little power. This has resulted in array efficiencies of 12%, nearing the efficiency figures that some feel are necessary in order for photovoltaic systems to become competitive with conventional energy sources in the central utility market.

Advances have been made in power conditioning subsystems, another source of power loss for a photovoltaic system. Though commercially available converters in the early 1980s were only about 80% efficient over their broad range of operation, experimental systems have now reached 95% efficiency.

Other photovoltaic materials have energies that more suitably match the Sun's spectrum and thus make more efficient use of it. One of them, single-crystal gallium arsenide, has high efficiency, stability, and the ability to withstand higher operating

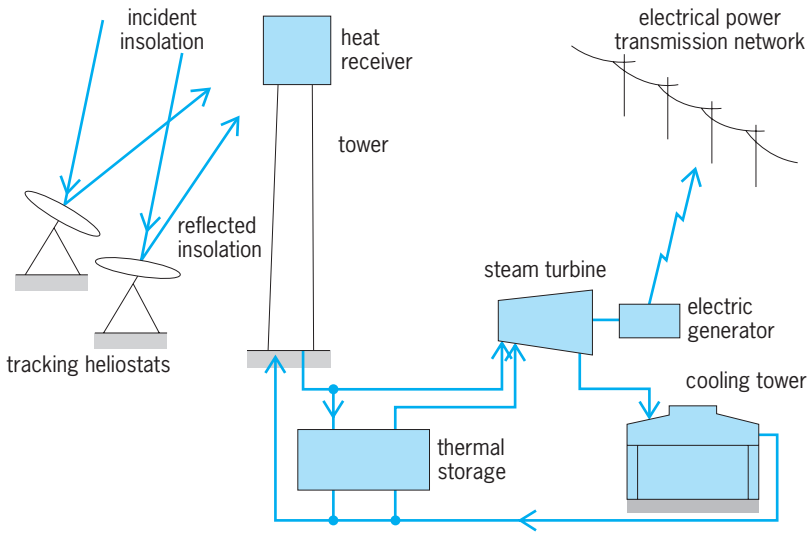

Fig. 8. Schematic of a typical central receiver collector for power generation using conventional steam technology. (*After F. Krawiec, J. Thorton, and M. Edesess, An Investigation of Learning and Experience Curves, SERI/TR-353-459, Solar Energy Research Institute, 1980*)

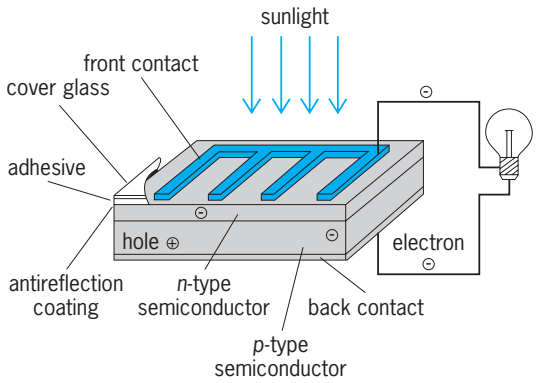

Fig. 9. Typical single-crystal silicon solar cell. (*After G. Cook, SERI Photovoltaic Advanced Research and Development: An Overview, SERI/SP-281-2235, Solar Energy Research Institute, 1984*)

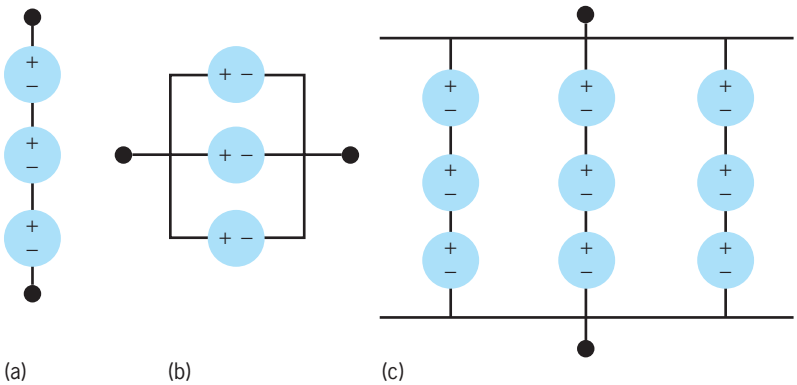

Fig. 10. Photovoltaic cell arrangements for building voltage and current. (*a*) String of photovoltaic cells (series). (*b*) Groups of cells (parallel). (*c*) Group of strings (combination). (*After Renewable Energy Technology Handbook for Military Engineers, SERI/SP-200-1413, Solar Energy Research Institute, 1982*)

temperatures than single-crystal silicon. For one-sun conditions, the highest theoretical efficiency of a single-crystal gallium arsenide cell is greater than 27%, and laboratory cells have reached efficiencies as high as 22%. Under sunlight concentrated 150 times, laboratory gallium arsenide cells have reached efficiencies nearing 24%.

For even higher efficiency, there is the multijunction (or stacked) device, a complex structure that uses several different semiconductor materials. Each layer has a junction sensitive to a different portion of the solar radiation spectrum, to make optimal use of all incident radiation. For example, an arrangement using three cells and three junctions—gallium aluminum arsenide, gallium arsenide, and silicon—has a theoretical maximum efficiency greater than 40% (**Fig. 12**).

Biomass

Biomass is plant and animal material. Biomass energy is solar energy stored in plant and animal matter. Through photosynthesis in plants, energy from the Sun transforms simple elements from air, water, and soil into complex carbohydrates. These carbohydrates can be used directly as fuel (for example, burning wood) or processed into liquids and gases (for example, ethanol or methane). Sources of biomass that can be converted to useful energy include agricultural crops, wastes, and residues; forest wood, waste, and residues; animal wastes; some municipal wastes; aquatic and desert plants; algae; and bacteria. Biomass is a renewable energy resource because it can be harvested periodically and converted to fuel. *See* BIOMASS; CARBOHYDRATE; PHOTOSYNTHESIS.

Feedstock production. Feedstock for biomass conversion can consist of wood, herbaceous plants, or aquatic plants.

Wood. Wood has been used as a fuel for tens of thousands of years. Woody biomass comes from a variety of sources, ranging from standing forests to wood processing operations. It includes mill and logging residues, timber thinnings, and harvests from well-managed natural stands or from short-rotation, fast-growing plantations. Mill residues are the by-products of processing operations that include the production of lumber, plywood, pulp and paper, and furniture. Dry mill residues from these operations are the most readily available of all tree biomass. Logging residues (residues left after tree harvest) include tree tops and branches, cull logs, standing dead trees, and stumps.

Wood has a heating value of about 8500 Btu/lb (20 megajoules/kg) when oven-dried compared to gasoline at about 21,000 Btu/lb (48 MJ/kg) and coal at about 13,500 Btu/lb (31 MJ/kg). Moisture dramatically affects the value of wood as a fuel. With a 30% moisture content, wood has a heating value of 5950 Btu/lb (14 MJ/kg); at 50% moisture the value is 4250 Btu/lb (10 MJ/kg). There are several methods of reducing moisture in wood to improve its performance as a fuel. Perhaps the simplest is drying, but some faster methods have been developed that apply residual combustion heat to dry the feedstock.

Improved silvicultural practices on harvested forests include overstocking (for example, 1000 stems per acre instead of 600) and improving regrowth from stumps following harvest. Another

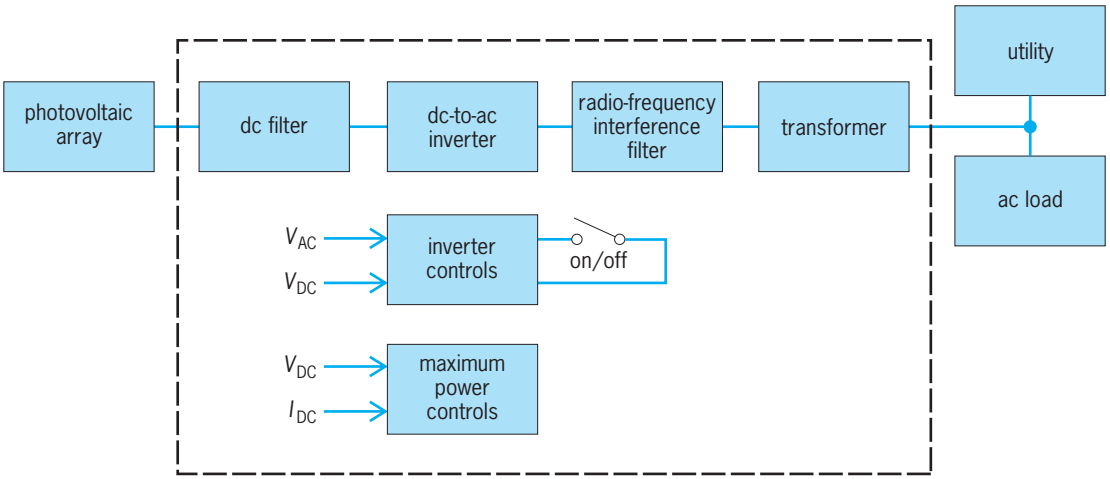

Fig. 11. Transformer isolation of dc from ac. (*After Renewable Energy Technology Handbook for Military Engineers, SERI/SP-200-1413, Solar Energy Research Institute, 1982*)

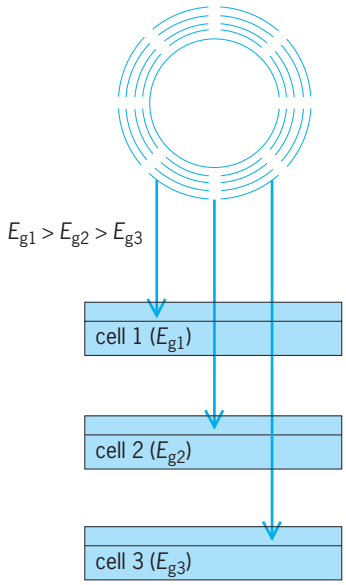

Fig. 12. Multiple-junction cell for highest efficiency. Gallium aluminum arsenide cell band gap (E_{g1}) = 1.82 eV. Gallium arsenide cell band gap (E_{g2}) = 1.45 eV. Silicon cell band gap (E_{g3}) = 1.1 eV. (*After G. Cook, SERI Photovoltaic Advanced Research and Development: An Overview, SERI/SP-281-2235, Solar Energy Research Institute, 1984*)

forestry practice improves production by cultivating fast-growing trees for energy conversion. These trees are selected for their rapid growth, ease of establishment, stump regeneration, and pest and disease resistance. Short-rotation plantations may require intensive cultivation practices similar to those used in agriculture such as irrigation, fertilization, and harvesting operations. *See* SILVICULTURE.

Herbaceous plants. Nonwoody (herbaceous) biomass feedstocks such as grasses and grains show promise as sources of fuel. Some herbaceous plants have both the potential for high yield and the ability to grow on arid and marginal land (**Fig. 13**). Herbaceous feedstock types can include plants intended as fuel for direct combustion to energy and plants grown for their carbohydrate or hydrocarbon content for conversion to liquid and gaseous fuel.

Grasses and legumes can be pressed into pellets for direct burning or gasification, or they can be converted via fermentation to alcohol. Starch and sugar crops also can be converted to energy via fermentation to alcohol. Agricultural residues not used for other purposes such as food or forage can be burned for heat or converted to alcohol.

Numerous plant and seed species are possible candidate crops for oil production. For example, about 2000 species of *Euphorbia* (for example, poinsettia, cassava, or the rubber tree) produce a latex rich in hydrocarbons that can be a substitute for petroleum in almost all its uses. Liquid hydrocarbons that are extracted from *Euphorbia* spp. may be processed into gasoline and similar fuels. *See* EUPHORBIALES; RUBBER TREE.

Aquatic plants. Certain aquatic plants can be cultivated on nonarable or nonproductive land. They have exceptional rates of biomass production and provide yields that surpass those of land plants by a factor of four or more. Microalgae have been studied as producers of gaseous (hydrogen) and liquid (oils) fuels. Macroalgae, such as the giant kelp, grow naturally in the open ocean and can be cultivated offshore or onshore in ponds. These species have high growth rates and produce significant amounts of

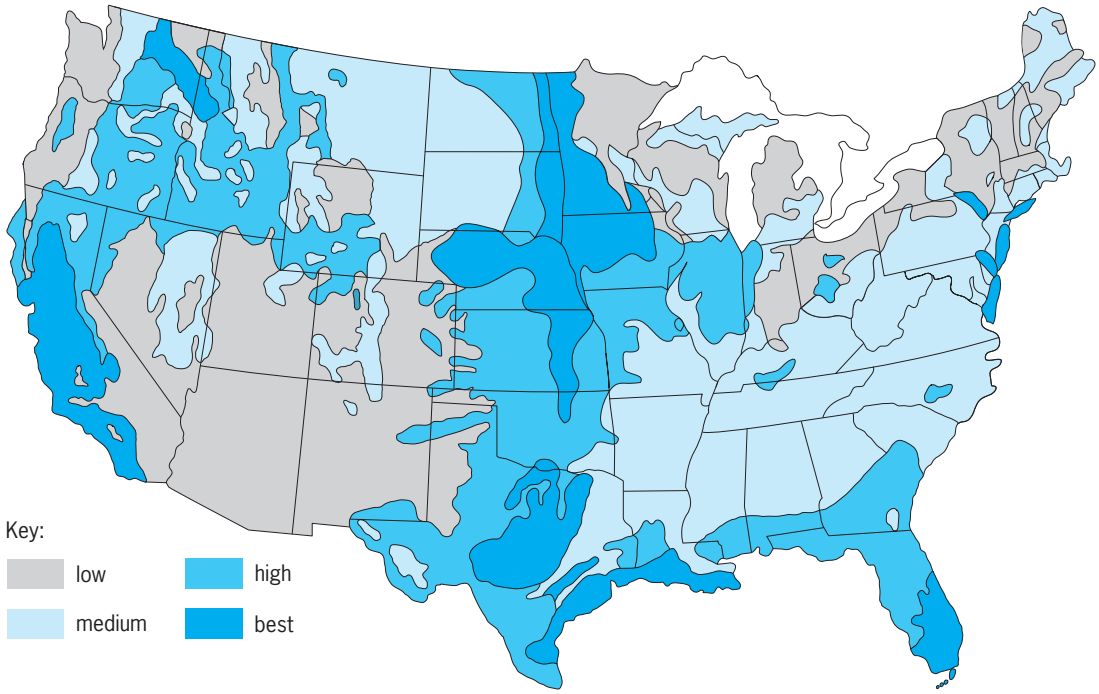

Fig. 13. Potential large-scale productivity of herbaceous biomass with all arable land considered. (*After Solar Energy Research Institute, FY 1982 Biomass Energy Technology Division Annual Technical Progress Report, vol. 1: Executive Summary, SERI/SP-281-2154, Solar Energy Research Institute, 1983*)

TABLE 2. Projected biomass resources for thermochemical conversion

Resource	Available quantity, 10^6 dry tons/year (10^6 dry metric tons/year)					
	1985		1990		2000	
	Maximum	Probable	Maximum	Probable	Maximum	Probable
Wood	464 (421)	167 (152)	429 (389)	154 (140)	549* (498)	71 (64)
Surplus agricultural crops†	45 (41)	34 (31)	36 (33)	31 (28)	33 (30)	39 (35)
Agricultural residues	220 (200)	115 (104)	240 (218)	123 (112)	278 (252)	110 (100)
Municipal solid waste	92 (83)	69 (63)	99 (90)	77 (70)	116 (105)	89 (81)
Other energy crops	8 (7)	104 (94)	69 (63)	203 (184)	172 (156)	247 (224)
Total	829 (752)	489 (444)	873 (792)	588 (533)	1148 (1041)	556 (504)

*Assumes wood from silvicultural energy farms starting in 1985.
†Assumes an aggressive program to establish sorghum as a cash crop that would divert land from corn in 1990 and 2000; data include surplus grain rejected for food processing.

carbohydrates that can be anaerobically digested to methane or fermented to alcohol. Emergent plants, such as cattails and bulrushes, and floating plants, such as water hyacinths, are also highly productive and represent potentially significant feedstock resources for conversion to methane or alcohol. *See* ALGAE.

Conversion to heat and fuel. Liquid and gaseous fuels usually can be transported more easily and efficiently than raw biomass. To be converted to these forms, most biomass must be treated either thermally, chemically, or biologically (**Table 2**). The oldest and simplest method for converting biomass is direct combustion. However, wood is 25 to 50% less efficient as a fuel source than fuel oil, natural gas, or coal. Even dried wood produces lower flame temperatures and requires larger fireboxes and heat transfer surfaces. Converting biomass to heat energy in large boilers is about 70 to 75% efficient, whereas wood can be burned in stoves in homes with efficiency ranges of 50 to 70%.

Gasification. Gasifiers can convert biomass as varied as hardwood chips or animal manure to combustible gases. Most gasifiers are supplied with air to produce a nitrogen-rich low-Btu gas. The oxygen in the air reacts with part of the biomass to produce heat. This heat is then used to produce gas from the remaining biomass. The low-Btu gas is usually used in nearby burners or engines. Conversion efficiencies are generally good, with 60 to 80% of the energy content of the dry biomass transferred to the gas. Use of biomass gasifiers offers flexibility in industrial plant design and provides a way of coupling biomass into existing plants with interchangeability between biomass fuel and conventional fuel (**Table 3**).

Another gasification process uses pure oxygen instead of air to produce synthesis gas (a mixture of carbon monoxide and hydrogen). Methanol, a liquid fuel, can be produced from the synthesis gas by one of several conventional processes. Efficiencies for the overall conversion to methanol are in the range of 40 to 50%. A process for the conversion of methanol to gasoline also is available. *See* ALCOHOL; METHANOL.

Pyrolysis. Producing charcoal by pyrolysis (breaking complex molecules into simpler ones by heat) has been done for centuries. Wood pyrolysis in general leaves a char that constitutes about 30% (by weight) of the wood and contains about 42% of its heat content. It also yields an oil that carries about 32% of the wood's heat content. The pyrolysis gases that are produced are usually burned to supply the processing plant's energy. *See* PYROLYSIS.

Fermentation. The conversion of sugars and starch by fermentation to ethanol is a well-established technology. Sugar from molasses and cheese whey, and starch from grains or tuberous plants are the most commonly used starting materials. However, most of the recent interest has been in fuel ethanol production involving the use of cornstarch, which is converted to glucose, which is then fermented to ethanol. Ethanol fermentation has traditionally been done with yeast in an 18–48-h batch process that produces 8 to 10% ethanol by volume. With modern

TABLE 3. Evaluation of biomass-gasification technologies

Technology	Advantages	Disadvantages
Pyrolysis	Low operating temperature; simple design; no other reactants required	Can get complex char, tar, and gas mixture; tar removal difficult; heat transfer rate may be limited in some cases
Air gasification	Simple design, only one reactor required; less tar than pyrolysis	Gas diluted with nitrogen (N_2) cannot be used for synthesis; low-Btu gas (LBG) may have a reduced combustion efficiency
Oxygen gasification	Medium-Btu gas (MBG) undiluted with N_2 is suitable for synthesis; can operate at high temperature and pressure if desired	Oxygen plant required, greatly increasing the capital cost at small scales
Steam gasification	MBG undiluted with N_2; no oxygen plant required; catalysts can adjust gas composition in gasifier	MBG is methane-rich; must be steam-reformed for methanol synthesis but is an advantage for synthetic natural gas (SNG) synthesis

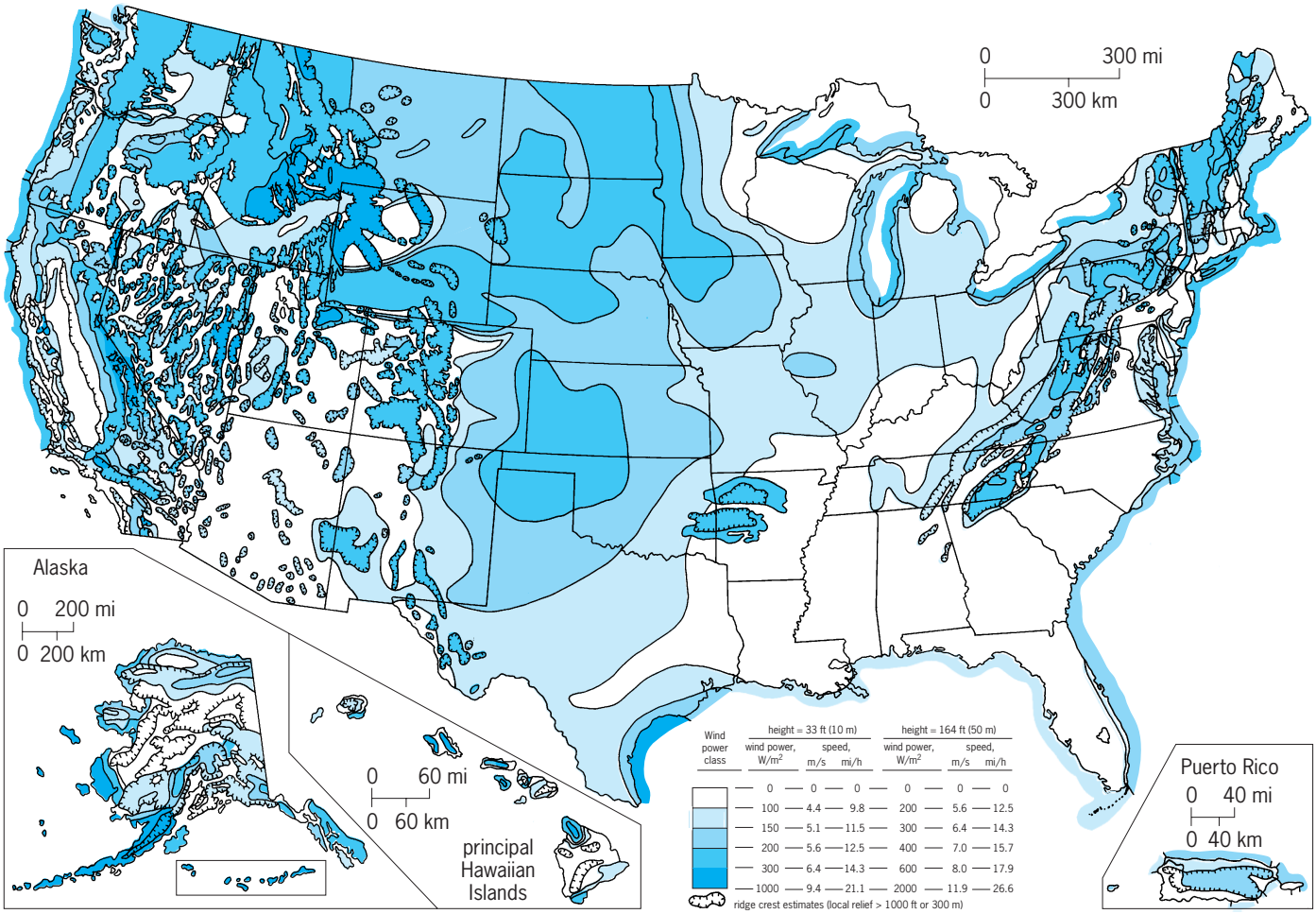

Fig. 14. Wind power in the United States (annual average). (*After Solar Energy Research Institute*)

techniques, the yield of ethanol is about 79 gallons per bushel of corn (8300 liters/m^3); 81% of the manufacturing cost of ethanol is the cost of the starting material. *See* ALCOHOL FUEL; ETHYL ALCOHOL.

Biomass with a high fiber content (woody biomass) is the largest renewable resource available for energy conversion. It is more complex in composition and structure than is starch, and is therefore more difficult to process into potential fuels. Theoretically, the carbohydrates in fibrous biomass can be converted to 100 gallons of ethanol per ton (420 liters/metric ton) of biomass at a 40% net energy conversion compared with 120 gallons/ton (500 liters/metric ton) for corn. One component of woody biomass, lignin, could be chemically removed and used since it has a heating value of about 15,000 Btu/lb (35 MJ/kg) to provide part of the heat needed to convert the carbohydrates to ethanol. *See* FERMENTATION; LIGNIN.

Anaerobic digestion. Anaerobic digestion by bacteria to make methane can use almost any biomass. The technology can be used in a digester of any size, from a small backyard fermenter that uses human waste to the large fermenters that use feedlot waste. Manure is particularly amenable to digestion because of its abundant nutrients, high moisture content, and lack of inhibitors of methane-producing bacteria. The gas produced during digestion is 50 to 70% methane and can be purified, burned as it is, or fed to a natural gas pipeline.

Photosynthetic microorganisms. Some types of microalgae and bacteria have the ability to manufacture energy-rich substances photosynthetically. For example, under appropriate conditions some microalgae will produce increased amounts of oils and lipids that can be used directly as fuel, and some bacteria will evolve hydrogen that also can be used directly as a fuel. Research concentrates on genetic engineering, artificially maintained production systems, and the basic chemistry and physics of photosynthesis.

Wind Energy Conversion Systems

Energy from the wind has been used for centuries to propel ships, to grind grain, and to lift water. Wind turbines extract energy from the wind to perform mechanical work or to generate electricity. Applications to generate electricity are recent and date from about 1930. This discussion concentrates on the production of electricity from wind. *See* WIND.

Wind resource. Wind is a source of energy derived primarily from unequal heating of Earth's surface by

the Sun. Wind speeds are altered by local terrain and vary significantly by season and day.

Figure 14 shows wind resources as annual average wind power classes based on average wind power density. Average wind power density rather than wind speed is used because it incorporates the combined effect of wind speed, wind speed distribution, and air density.

Since the average wind power density is proportional to the cube of wind speed, location or siting is very important. Wind power of at least 200 W/m^2 or wind speed 12 mi/h (19 km/h) is considered prerequisite for additional site evaluation and data gathering. The annual average wind speed is only one consideration. The distribution of wind speed by month, day, and hour is also important, especially in matching power output to power requirements.

The variation of wind speed with height is another important consideration. Generally, wind speed is greater above ground level than it is at the surface due to the frictional effects of Earth's surface.

The output power from a wind turbine can be defined by a power coefficient, C_p, which is a measure of the efficiency with which the rotor extracts power from the wind. The maximum theoretical efficiency, or Betz limit, for a propeller-type rotor is 0.593. Wind turbines available today generally have a C_p between 0.30 and 0.45. *See* WIND POWER.

Wind turbine system. The main components of a wind turbine system (**Fig. 15**) are the rotor, generator, tower, and storage devices (or utility interconnection). The rotor consists of blades and a hub that connects the rotor to the turbine system. The blades can spin around a horizontal axis parallel to the ground or around a vertical axis perpendicular to the ground. Both designs are used in commercial and residential applications. Most commercial models are horizontal-axis wind turbines (HAWTs). They range in output from less than 5 kW to 100 kW. The U.S. Department of Energy has tested a large, 2500-kW prototype HAWT, called the MOD-2, for utility applications; and design of a 3200-kW machine, the MOD-5B, has begun. A wind farm is a system in which many turbines are clustered at especially windy sites to generate megawatts of electricity to sell to utilities.

There are three common configurations of vertical-axis wind turbines (VAWTs); the most developed is the Darrieus design (**Fig. 16**). They, unlike HAWTs, are omnidirectional—they need not be turned as the wind direction changes. Rotors on HAWTs have many blade configurations (**Fig. 17**). In the familiar farm windmill, the blades constitute most of the area swept by the rotor as it spins about its axis. The farm windmill and the Dutch windmill combine drag and lift for inducing rotation. The rotors on these machines have low speed, but have high torque and are well suited to lifting water or grinding grain.

The blades of high-speed turbines are designed as airfoils. In these, the tip-speed ratio—the ratio of the speed of the blade tips to the wind speed—can be as high as 15. A high ratio improves efficiency and reduces the size of drive train components. This is important in electricity generation, whereas torque is important to lifting water or grinding grain. *See* AIRFOIL.

Control of rotor velocity is important to protect the turbine against overspeed in high winds. All wind

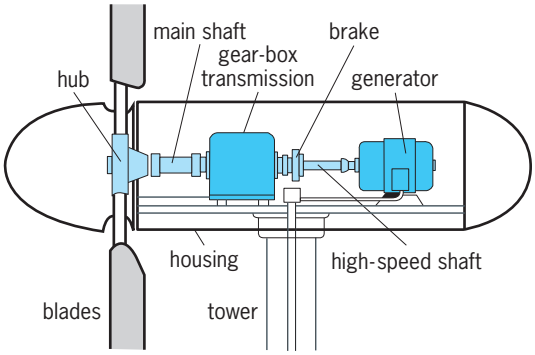

Fig. 15. Main components of a wind turbine. (*After B. H. Bailey, New York State Wind Energy Handbook, New York State Wind Energy Office, 1982*)

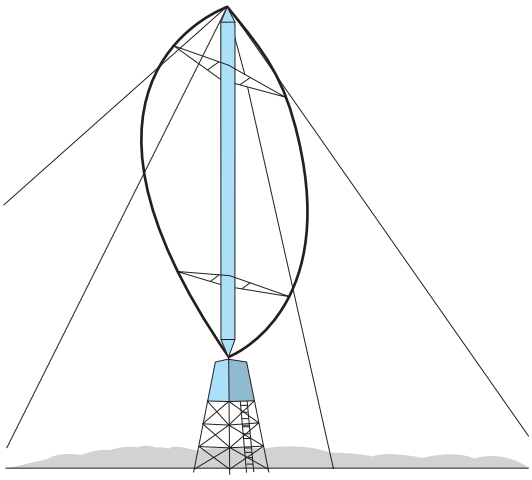

Fig. 16. Darrieus vertical-axis wind turbine. (*After B. H. Bailey, New York State Wind Energy Handbook, New York State Wind Energy Office, 1982*)

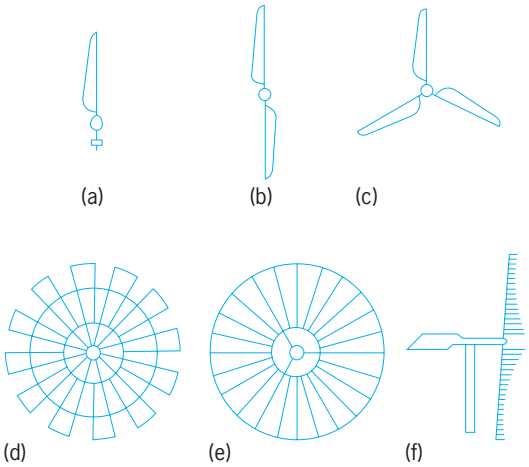

Fig. 17. Some horizontal-axis wind turbine configurations. (*a*) Single-bladed. (*b*) Double-bladed. (*c*) Three-bladed. (*d*) United States farm windmill multibladed. (*e*) Bicycle multibladed. (*f*) Sail wing. (*After Renewable Energy Technology Handbook for Military Engineers, SERI/SP-200-1413, Solar Energy Research Institute, 1982*)

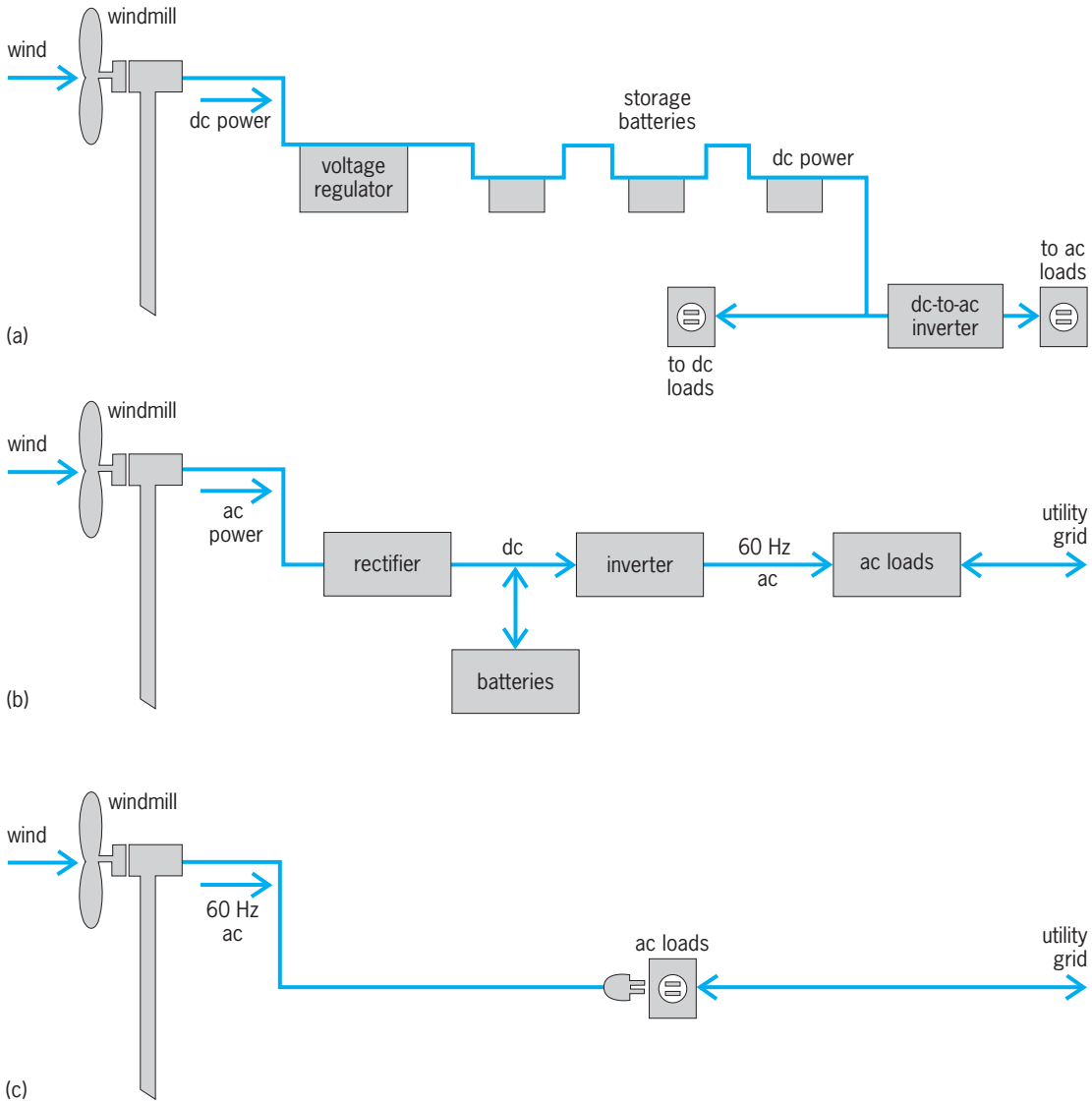

Fig. 18. Possible wind energy system configurations. (*a*) A dc generator windmill. (*b*) Variable-frequency ac generator windmill. (*c*) Induction generator windmill. (*After New York State Wind Energy Handbook*, *New York State Wind Energy Office*, *1982*)

turbines are designed to start at a certain minimum wind speed and stop when the maximum design speed is reached. Various mechanical, hydraulic, and aerodynamic devices are used to brake the rotor or to turn it out of the wind.

The wind energy captured by the rotor is transmitted by shafts and gears to the generator. Generators can produce either direct current or alternating current; the configurations of wind energy systems using different types of generators are shown in **Fig. 18**.

Wind turbines for residential applications are placed on towers usually at least 80 ft (25 m) tall. Towers are either free-standing or guyed, and either type can be a pole or lattice (truss) structure. Installation can cost as much as 30% of the total price of a small wind system. The cost of a tower can be partially offset by the increased output from higher wind speeds at the higher elevation.

Storage of wind energy can be by battery, pumped-hydro equipment, and other devices. Normally, wind turbines are connected to a utility, which provides backup electricity when needed. Excess electricity can, in turn, be fed into the grid and sold to the utility at a negotiated price. *See* TURBINE.

Ocean Thermal Energy Conversion

Ocean thermal energy conversion uses the temperature difference between surface water heated by the Sun and deep cold water pumped from depths of 2000 to 3000 ft (600 to 900 m). This temperature difference makes it possible to produce electricity from the heat engine concept. Since the ocean acts as an enormous solar energy storage facility with little fluctuation of temperature over time, ocean thermal energy conversion, unlike most other renewable energy technologies, can provide electricity 24 h a day. In addition, ocean thermal energy conversion plants can be located onshore, mounted on the continental shelf, or placed on moored or floating platforms. Various shapes have been suggested for floating platforms (**Fig. 19**).

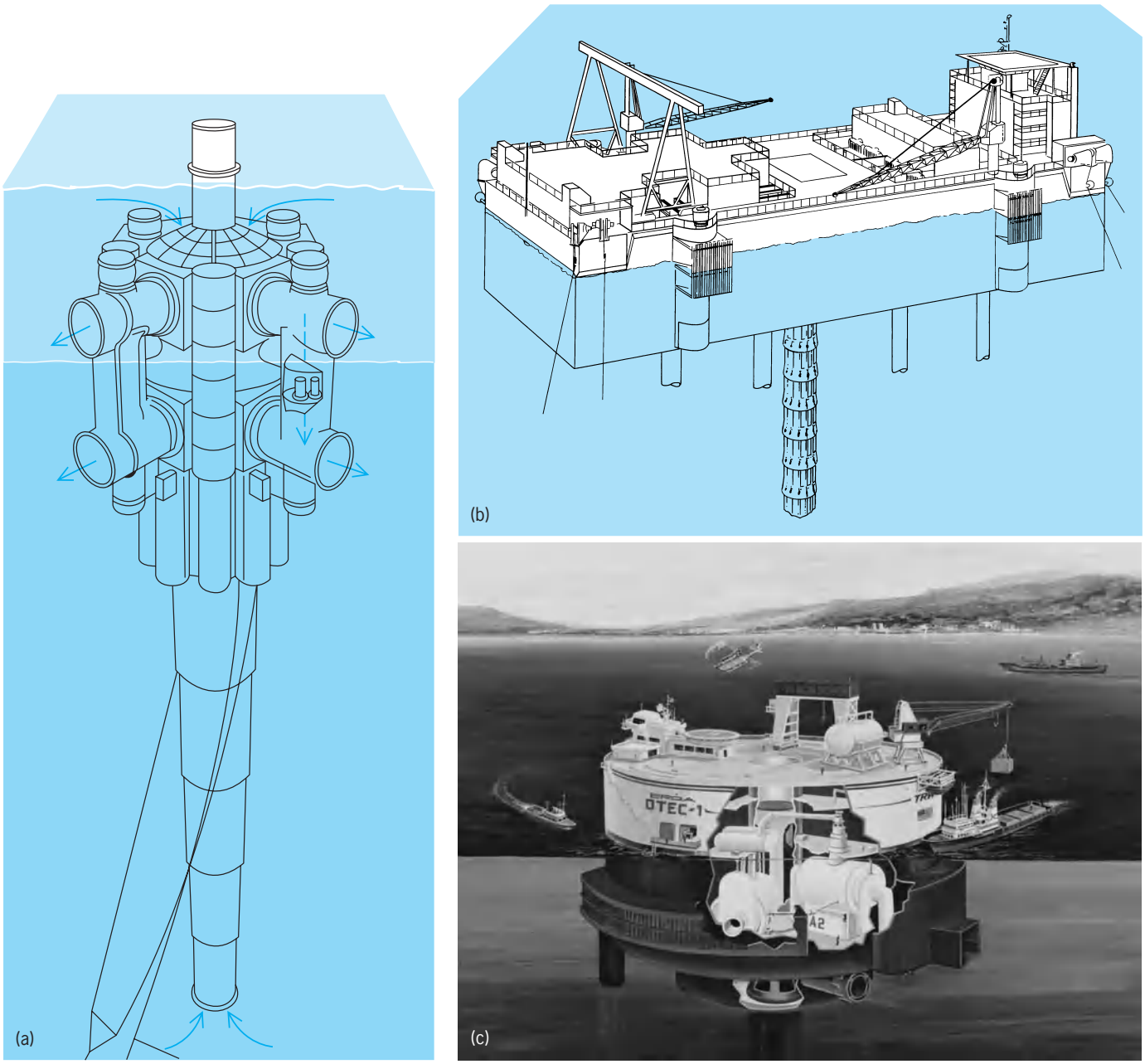

Fig. 19. Platform concepts for ocean thermal energy conversion (OTEC) systems. (*a*) Spar buoy. (*b*) Rectangular barge. (*c*) Cylindrical platform. (*After R. A. Meyers, ed., Handbook of Energy Technology and Economics, John Wiley and Sons, 1983*)

Energy content and electricity. A rough estimate of the thermal energy content of the ice-free, mixed layer of the ocean with a temperature difference of 18°F (10°C) is 70×10^{20} Btu (75×10^{23} joules). Thus, the ocean thermal energy resource is vast, but that suited for ocean thermal energy conversion facilities is only a fraction of this total, although still an enormous amount. Suitable sites are generally limited to an area within 30° north and 25° south of the Equator. With available technology, an annual average temperature difference of 36°F (20°C) or greater between water at the surface and that at a depth not exceeding 4700 ft (1500 m) is required for continuous and efficient operation of an ocean thermal energy conversion plant. Furthermore, the mean temperature difference during the coldest month should exceed 31°F (17°C). If system efficiency is improved, the required temperature difference will be less and a greater ocean area will be suitable for ocean thermal energy conversion use.

The possibility of generating electricity from solar thermal energy collected and stored by the ocean near the Equator was first suggested by A. d'Arsonval in 1881. In 1930, G. Claude built a power plant in Mantazas Bay, Cuba, and converted ocean thermal energy to electricity, using a temperature difference of 25°F (14°C) between the surface and a depth of 2200 ft (700 m) to generate 22 kWe. In 1979, a small closed-cycle test plant produced 50 kWe and a net output of 12 kWe by using ammonia as the working

fluid. In 1980, a 1000-kW test bed plant was mounted in a converted U.S. Navy tanker to test heat exchangers. A 100-kWe, shore-based plant was tested in 1981 by the Japanese for the island republic of Nauru. The Kyushu Electric Company in early 1983 began operating a 50-kWe test plant in Tokunoshima that uses waste heat from a diesel engine to boost the temperature of the warm seawater.

Conversion systems. There are three principal conversion systems for ocean thermal energy: closed Rankine cycle, open Rankine cycle (Claude cycle), and mist or foam-lift cycle. These systems are listed in order of technological development and in inverse order of theoretical energy conversion efficiency. In other words, the least technologically developed system, the lift cycle, theoretically has the highest potential efficiency. *See* RANKINE CYCLE.

Closed Rankine cycle. In the closed Rankine cycle, warm seawater is pumped through a heat exchanger to transfer stored thermal energy to a working fluid such as ammonia, which has relatively high vapor pressures at low temperatures (**Fig. 20**). The working fluid is vaporized, and the expanding vapor is used to drive a turbine generator. The vapor is condensed by cold seawater in another heat exchanger to complete the cycle. The major advantage of the closed cycle is that the process of heat exchange and energy conversion is established and familiar. The components can be manufactured with existing technology. The major disadvantage is that in adapting an energy conversion process designed for high temperatures and pressures to a low-temperature heat source, inefficiencies in the conversion process become magnified. The heat exchangers must be large, and a large volume of water must be pumped through them. About half of the total available temperature difference is required to transfer heat in the evaporator and condenser. Since the efficiency is a direct function of the temperature difference available after the transfer, the small but inherent drop in temperature through the heat exchange process reduces an already low temperature potential. *See* HEAT EXCHANGER.

Large heat exchanger surfaces and low working temperatures also make the control of slime formation (biofouling) and corrosion very critical. The design, construction, and operation of large closed-Rankine-cycle ocean thermal energy conversion facilities present complex engineering problems even though no technological breakthroughs are needed for any major structural or operating components.

Open Rankine cycle (Claude cycle). Claude showed that the heat exchange process for both evaporation and condensation could be performed by using warm seawater itself as the working fluid. Direct-contact heat exchange eliminates the need for large, costly heat exchangers. The evaporator and direct-contact condenser possibly can be made with inexpensive materials such as poly(vinyl chloride) and concrete.

In the open-cycle process, steam is generated from warm seawater by flash (very rapid) evaporation in an evacuated chamber. The steam turns a turbine and

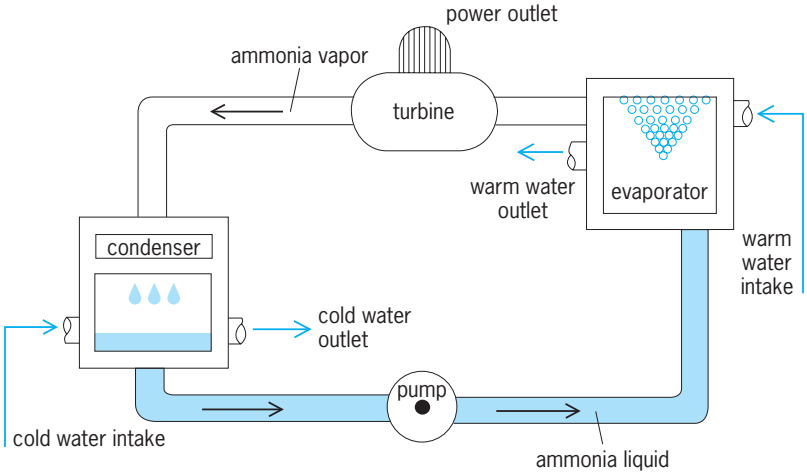

Fig. 20. Closed Rankine (ammonia) cycle. (*After The Ocean Option, SERI/SP-732-334, Solar Energy Research Institute, 1980*)

is condensed by direct contact with cold seawater to complete the cycle. The absence of any barrier between the working fluid and the heating and cooling source makes possible a higher overall system efficiency and lower seawater flow requirement than in the closed-cycle process. Using seawater rather than a substance like ammonia as the working fluid also lessens safety problems. The need for mechanical or chemical measures to control biofouling is also reduced.

The open-cycle process is similar to certain desalination processes, and if a typical closed-cycle surface condenser is used, fresh water can be produced as a by-product at some sacrifice in efficiency. The use of ocean thermal energy conversion facilities to produce fresh water as well as electricity may prove attractive in some locations. *See* WATER DESALINATION.

The elimination of volatile working fluids and heat exchangers means that the open-cycle process will function at very low steam pressures. The steam turbine, in turn, must be very large and housed in a correspondingly large vacuum vessel. Present technology is limited to relatively small plants.

Lift cycles. There are several other approaches to the open cycle that involve the conversion of the thermal gradient into a hydraulic head through the lift of seawater in vertical two-phase flow. The water column created can drive a turbine, as in a hydroelectric dam. An advantage of this approach is that the hydraulic turbine is more compact and less costly than the steam turbine required in the Claude system.

Warm seawater is introduced as a mist that is then lifted against gravity by the flow of steam from a low-pressure region to a lower-pressure region. This action is similar to the naturally occurring cycle that converts water evaporated by solar energy to hydropower. In ocean thermal energy conversion lift cycles, an artificial hydrological cycle is created.

Robert L. San Martin

Bibliography. B. Anderson, *Solar Building Architecture*, 1990; J. A. Duffie and W. A. Beckman, *Solar Engineering of Thermal Processes*, 2d ed.,

1991; M. A. Green, *Third Generation Photovoltaics: Advanced Solar Energy Conversion*, 2003; *Kirk-Othmer Encyclopedia of Chemical Technology*, vol. 17, 4th ed., 1995; Institute of Gas Technology, *Energy from Biomass and Wastes Symposium,* 1989; G. N. Tiwarik, *Solar Energy: Fundamentals, Design, Modelling and Applications*, 2002; S. Wieder, *An Introduction to Solar Energy for Scientists and Engineers,* 1982, reprint 1990.

Solar heating and cooling

The use of solar energy to produce heating or cooling for technological purposes. When the Sun's short-wave radiation impinges upon a blackened surface, much of the incoming radiant energy can be absorbed and converted into heat. The temperature that results is determined by: the intensity of the solar irradiance; the ability of the surface to absorb the incident radiation; and the rate at which the resulting heat is removed. By covering the absorbing surface with a material such as glass, which is highly transparent to the Sun's short-wave radiation but is opaque to the long-wave radiation emitted by the Sun-warmed surface, the effectiveness of the collection process can be greatly enhanced. The energy which is collected can be put to beneficial use at many different temperature levels to accomplish: distillation of seawater to produce salt or potable water; heating of swimming pools; space heating; heating of water for domestic, commercial, and industrial purposes; cooling by absorption or compression refrigeration; cooking; and power generation by thermal or photovoltaic means. *See* SOLAR ENERGY.

Distillation

The oldest of these applications dates back to a 50,000-ft^2 (4600-m^2) installation built in Chile in 1872 to distill saline water and make it potable. **Figure 1** shows in cross section the type of glass-roofed solar still used more than a century ago in Chile and used again in modern times on the Greek islands in the eastern Mediterranean and in central Australia. When the Sun shines through the glass cover into the salt or brackish water contained within the concrete channel, the water is warmed and some evaporates to be condensed on the underside of the glazing. The condensate runs down into the scuppers and then into a suitable container. More than 6000 gal (23,000 liters) of pure drinking water were produced on each sunny day by the Chilean still, and the larger version now operating at Coober Pedy, Australia, produces nearly twice as much.

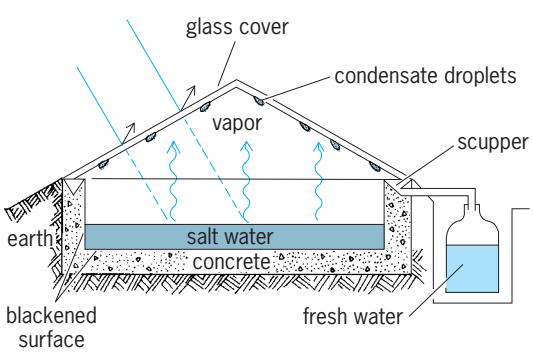

Fig. 1. Roof-type solar still.

Production of salt from the sea has been accomplished for hundreds of years by trapping ocean water in shallow ponds at high tide and simply allowing the water to evaporate under the influence of the Sun. The residue contains all of the compounds that were present in the seawater, and it is sufficiently pure for use in many industrial applications. *See* DISTILLATION; WATER DESALINATION.

Swimming Pool Heating

Swimming pool heating is a moderate-temperature application which, under suitable weather conditions, can be accomplished with a simple unglazed and uninsulated collector similar to the black polyethylene extrusion shown in **Fig. 2**. When the water to be heated is at almost the same temperature as the surrounding air, little or no heat will be lost from the absorber and so there is no need for either glazing or insulation. Such collectors are usually designed to empty themselves when the circulating pump is shut off so they can avoid the danger of freezing at night. For applications where a significant temperature difference exists between the fluid within the collector passages and the ambient air, both glazing and insulation are essential.

Space Heating

Space heating can be carried out by active systems which use separate collection, distribution, and storage subsystems (**Figs. 3** and **4**), or by passive designs which use components of a building to admit, store, and distribute the heat resulting from absorbing the incoming solar radiation within the building itself.

Passive systems. Passive systems can be classified as direct-gain when they admit solar radiant energy directly into the structure through large south-facing windows (**Fig. 5**), or as indirect-gain when a wall (**Fig. 6**) or a roof (**Fig. 7**) absorbs the solar radiation, stores the resulting heat, and then transfers it into the building. If the absorber and the storage components are not a part of the building fabric but are separate subsystems which operate by the natural circulation of warmed or cooled air, the term isolated-gain is generally used (**Fig. 8**).

Passive systems are generally effective where the number of hours of sunshine during the winter months is relatively high, where moderate indoor temperature fluctuations can be tolerated, and where the need for summer cooling and dehumidification is moderate or nonexistent.

Direct-gain systems. Most passive systems make use of the fact that, in winter, whenever the Sun is above the horizon it is in the southern part of the sky, and its altitude above the horizontal plane at noon is relatively low compared to the much higher position which it will attain in summer. This means that in winter

when heat is needed, vertical south-facing windows admit solar radiation freely as long as the Sun is shining. The use of double glazing reduces the transmission of the incoming solar radiation by about 16%, but it can halve the thermal loss due to the indoor-outdoor temperature difference. The use of movable insulation which can be placed over the windows at night and removed during the sunlit hours of the day can also greatly improve the performance of direct-gain systems. Window area on surfaces other than those facing south should be kept to the minimum permissible under local building codes. The building components (walls and floors) on which the solar radiation falls should possess as much mass as possible so that the excess solar heat gained during the day can be stored for use at night.

Indirect-gain systems. The indirect-gain concepts shown by Figs. 6 and 7 interpose a thermal mass between the incoming solar radiation and the space to be heated. The system shown in Fig. 6 uses a south-facing glazed wall of concrete or masonry with an air space between the wall's outer surface and the single or double glazing. This is known as a Trombe wall. Vents with dampers are provided near the floor and at the ceiling level so that cool air can be drawn by chimney action from the room into the space between the glass and the wall. The air is heated by contact with the wall, and rises to reenter the room at the ceiling level. At night, the chimney action stops and the dampers are closed to prevent the downflow of cooled air back into the room.

As the Sun's rays warm the outer surface of the concrete, a wave of heat begins to move slowly (at about 2 in. or 5 cm per hour) through the wall. The thickness of the wall is chosen to delay the arrival of the wave of warmth at the indoor surface until after sunset, when the long-wave radiation emitted by the wall will be welcome. Heat also flows outward from the wall to the glazing and thence to the outdoor environment, but this can be minimized by the use of movable insulation within the air space. Windows may be introduced into the Trombe wall, and the building need not be restricted to one story.

Another indirect-gain system, shown in Fig. 7, uses enclosed bags of water, called thermoponds, which are supported by a heat-conducting roof-ceiling and covered by horizontally movable insulating panels. This system collects solar energy during winter days by rolling the insulation away to a storage area, thus admitting Sun's rays into the thermoponds. At night, the insulating panels are rolled back to provide the insulation needed to retain the collected heat.

In summer, the operation is reversed, and the insulating panels are rolled back at night to expose the thermoponds to the sky and thus to enable them to dissipate heat that has been absorbed from the building during the day. Convection to cool night air, when it is available, and radiation to the sky on clear nights are two of the natural processes by which heat can be rejected. The third process, and the most potent, is evaporation, which can occur when the ponds are provided with exposed water surfaces by flooding or spraying them. *See* HEAT TRANSFER.

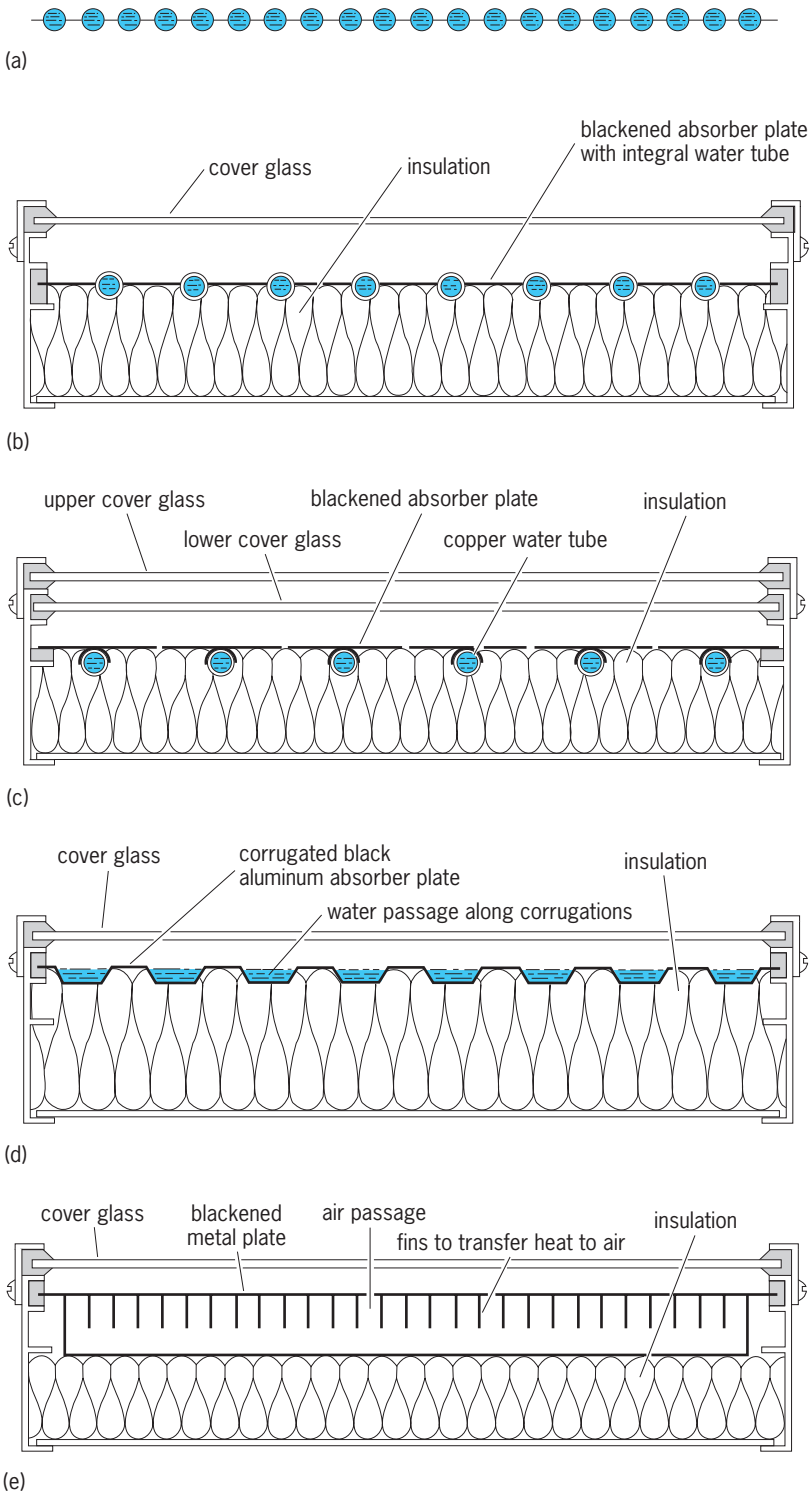

Fig. 2. Flat-plate collector types. (*a*) Extruded black plastic swimming pool heater. (*b*) Single-glazed collector with tube-in-sheet. (*c*) Double-glazed collector with extruded aluminum absorber element and copper water tube. (*d*) Open-flow collector used in the Thomason Solaris system. (*e*) Air heater using finned aluminum heat absorber.

Another indirect-gain system makes use of a greenhouse attached to the south side of a building to gather heat during the day. The wall between the glazed space and the building is warmed by the Sun's rays and by contact with the warm air in the

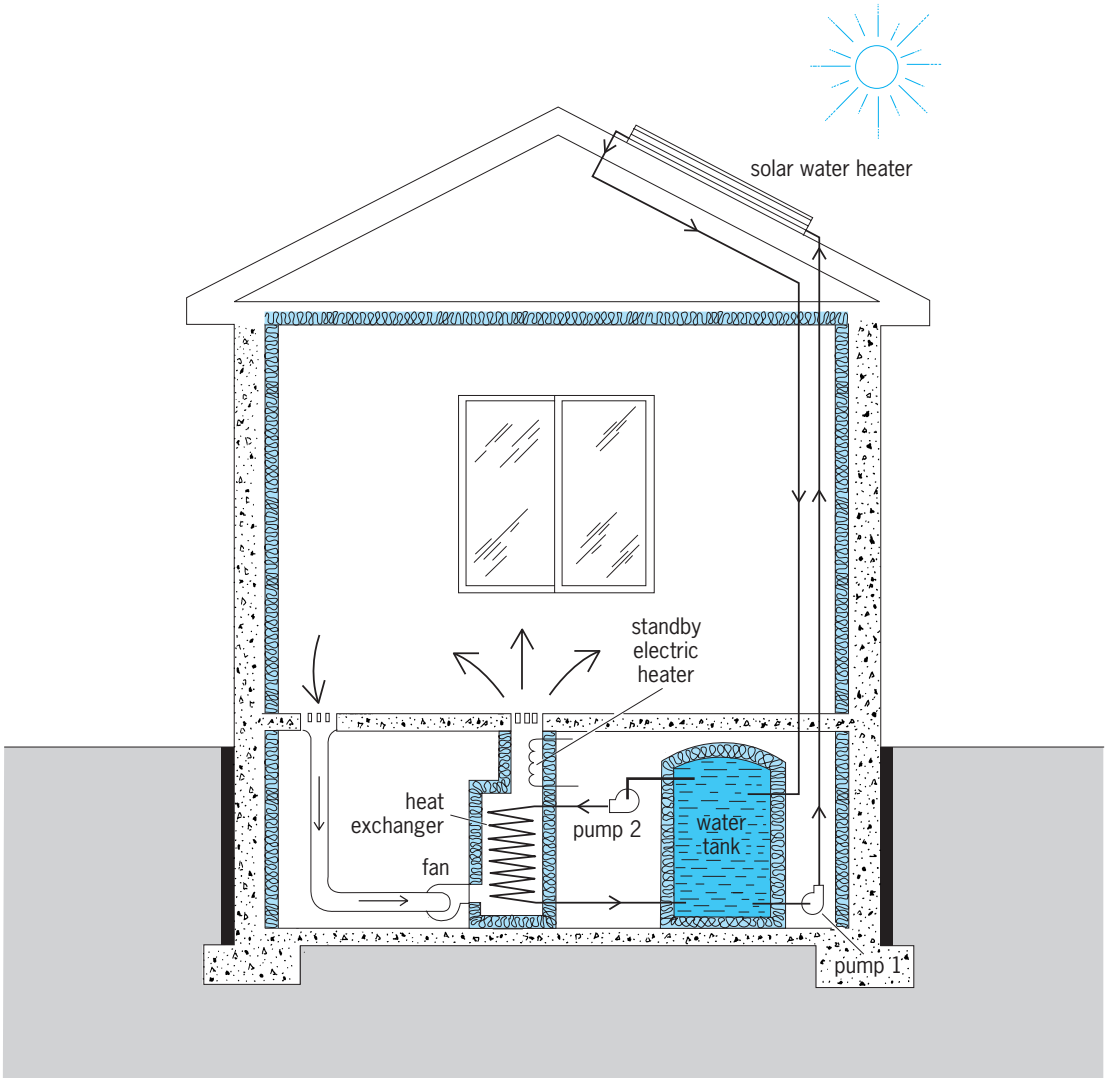

Fig. 3. Active system using roof-mounted solar water heater with storage tank in basement.

greenhouse. Additional stored heat can be transferred into the house by opening windows or vents. The warmth which is gathered even during cold winter days can thus be used to aid in heating the residence as well as to provide an environment in which plants can grow. *See* CONVECTION (HEAT); HEAT RADIATION.

Isolated-gain systems. The isolated-gain system shown in Fig. 8 uses a thermosyphon heater to warm air, which rises as it becomes hot, until it encounters the entrance to the rock-bed thermal storage component. The rocks abstract heat from the air which moves downward as it becomes cooler. The air eventually descends to the bottom of the rock bed and thence to the inlet of the heater. The cycle continues as long as the Sun shines on the collector, and heat is stored throughout most of the day. At night, when the occupants of the house sense the need for heat, the damper settings are changed and warm air from the top of the rock bed is admitted to the living space. At the same time, cool air from the north wall of the house is allowed to flow downward through ducts to enter the base of the rock bed. There the air is warmed, and the heat which has been stored during the day is used to warm the house at night.

Active systems. Active systems may use either water (Fig. 3) or air (Fig. 4) to transport heat from roof-mounted south-facing collectors to storage in rock beds or water tanks. The stored heat may be withdrawn and used directly when air is the transfer fluid. When the heat is collected and stored as hot water, fan-coil units are generally used to transfer the heat to air which is then circulated through the warmed space. Each system has both advantages and disadvantages.

Air cannot freeze or cause corrosion, and leakage is not a serious problem. Water requires relatively small pipes compared with the ducts needed to transport the same amount of heat in the form of warm air. Water tanks can store more than three times as much heat as rocks in a given volume per degree of temperature change, but rock beds are considerably lower in cost than water tanks, and rocks can

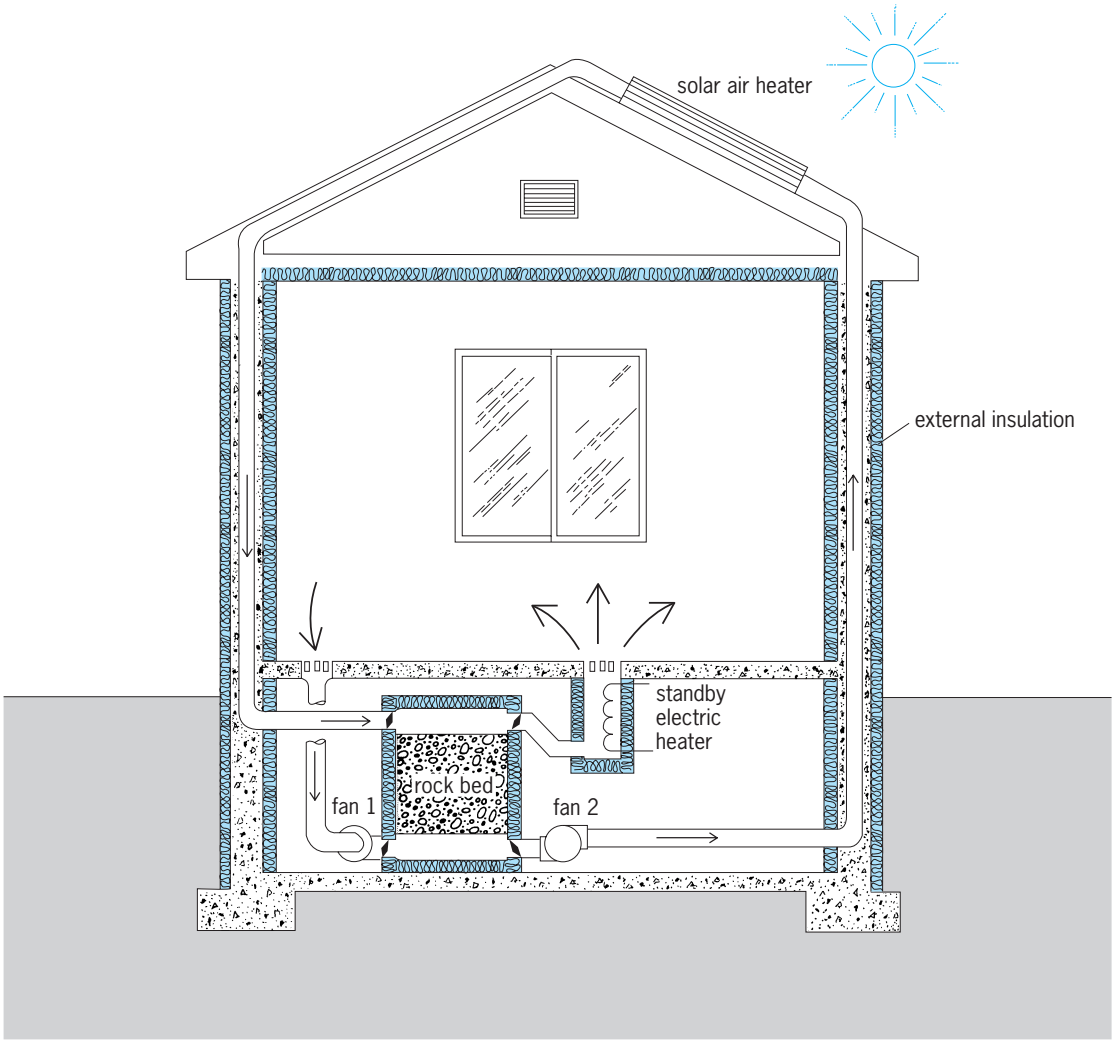

Fig. 4. Active system using air heater with rock-bed heat storage in basement.

tolerate virtually any temperature while water will boil at 212°F (100°C) unless its pressure is raised above atmospheric. The choice between water and air systems must be made carefully, taking into account the many features of each.

Heat storage. Heat storage can be accomplished with specific-heat materials such as water or rocks, which can store and discharge heat by simply undergoing a change in temperature. A different process is involved in the heat-of-fusion materials, which, like water at 32°F (0°C), can freeze when heat is removed and melt again when the same amount of heat is returned to them. Ice has the unique property of increasing in volume by 7% compared with the water from which it is formed, but virtually all other substances become denser when they solidify and so their solid components sink instead of floating. A number of materials are available which, like the well-known Glauber's salt, change their state from liquid to solid and back again at temperatures which are useful for either heating or cooling. None has yet been found, however, which possesses all of the attributes required for successful application to solar heating and cooling systems.

Standby electrical sources. Standby electrical sources are shown in Figs. 3 and 4, since some method of providing warmth must be included for use when

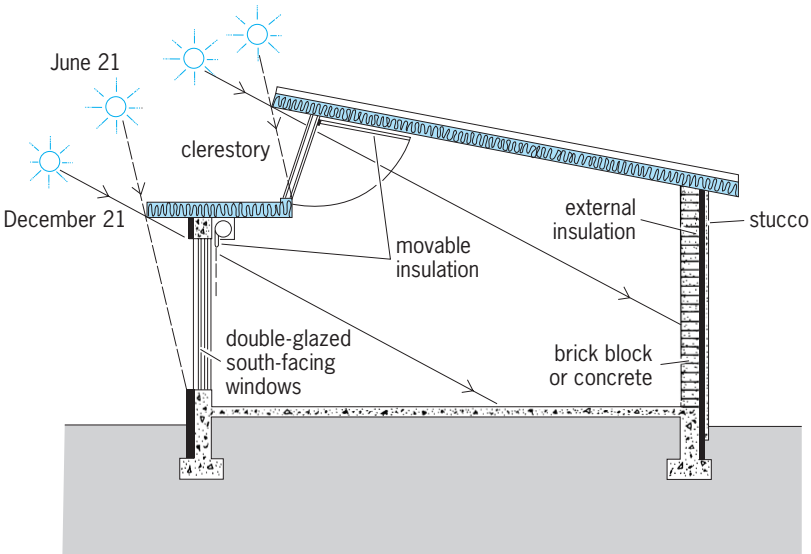

Fig. 5. Direct-gain passive system using large south-facing windows.

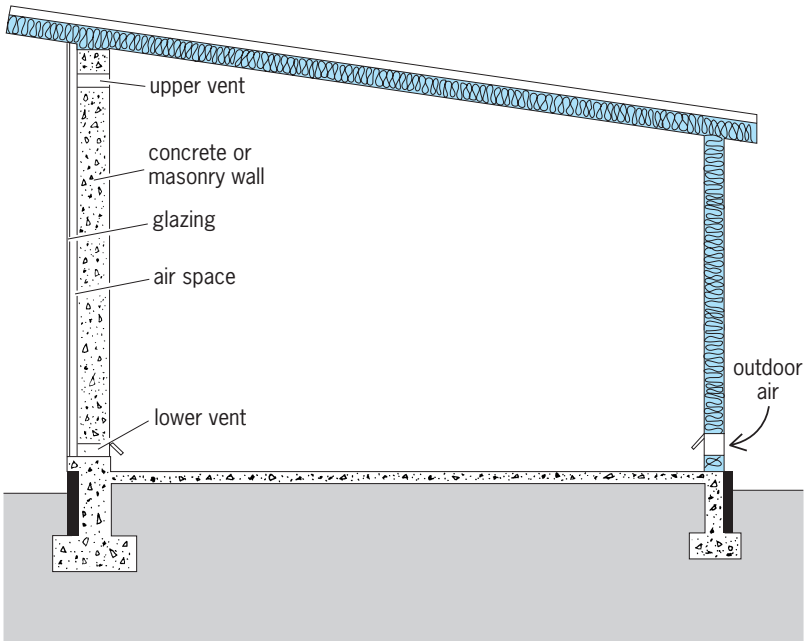

Fig. 6. Indirect-gain passive system using Trombe wall.

the Sun's radiant energy is inadequate for long periods of time. The standby heater may be something as simple as a wood-burning stove or fireplace, or as complex as an electrically powered heat pump. Simple electrical resistance heaters are frequently chosen for this service because of their low first cost, but their operating cost can become excessive in applications where they must be used for long periods of time. The rapidly escalating cost of electricity must be borne in mind when the standby energy source is selected. *See* COMFORT HEATING; HEAT PUMP.

Service Water Heating

Solar water heating for domestic, commercial, or industrial purposes is one of the oldest and most successful applications of solar-thermal technology. The most widely used water heater, and one that is suitable for use in relatively warm climates where freezing is a minor problem, is the thermosyphon type (**Fig. 9**). A flat-plate collector of one of the types shown in Fig. 2 is generally used with a storage tank which is mounted above the collector. A source of water is connected near the bottom of the tank, and the hot water outlet is connected to its top. A downcomer pipe leads from the bottom of the tank to the inlet of the collector, and an insulated return line runs from the top of the collector to the upper part of the storage tank which is also insulated.

The system is filled with water, and when the Sun shines on the collector, the water in the tubes is heated. It then becomes less dense than the water in the downcomer, and the heated water rises by thermosyphon action into the storage tank. It is replaced by cool water from the bottom of the tank, and this action continues as long as the Sun shines on the collector with adequate intensity. When the Sun moves away, the thermosyphon action stops, but the glazing on the collector minimizes heat loss from this component and the insulation around the storage tank enables it to retain the collected heat until needed. The system must be prevented from freezing, and the components must be strong enough to withstand operating water pressure.

For applications where the elevated storage tank is undesirable or where very large quantities of hot water are needed, the tank is placed at ground level. A small pump circulates the water in response to a signal from a controller which senses the temperatures of the collector and the water near the bottom of the tank. Heat exchangers may also be used with water at operating pressure within the tubes of the exchanger and the collector water outside to eliminate the necessity of using high-pressure collectors. Antifreeze substances may be added to the collector water to eliminate the freezing problem, but care must be taken to prevent any possibility of cross-flow between the collector fluid and the potable water. Many plumbing codes require special heat exchangers which interpose two metallic walls between the two fluid circuits. *See* HOT-WATER HEATING SYSTEM.

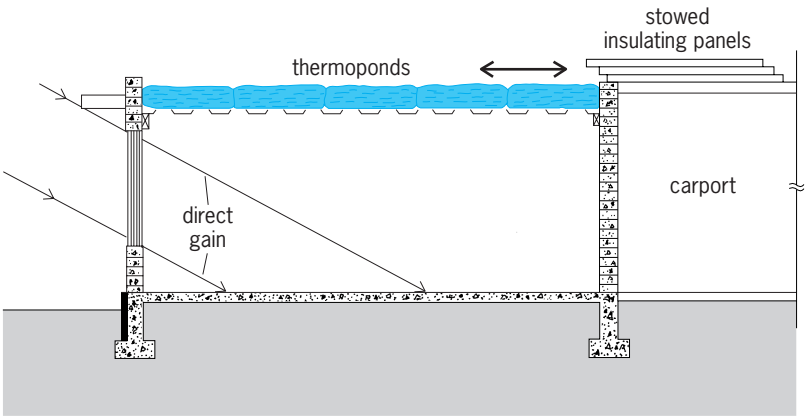

Fig. 7. Indirect-gain Skytherm system with movable insulation.

Cooling

Cooling can be provided by both active and passive systems.

Active cooling systems. The two feasible types of active cooling systems are Rankine cycle and absorption. The Rankine cycle system uses solar collectors to produce a vapor (steam or one of the fluorocarbons generally known as Freon) to drive an engine or turbine. A condenser must be used to condense the spent vapor so it can be pumped back through the vaporizer. The engine or turbine drives a conventional refrigeration compressor which produces cooling in the usual manner. *See* RANKINE CYCLE; REFRIGERATION.

The absorption system uses heat at relatively high temperature (180 to 200°F or 82 to 93°C) and employs a hygroscopic solution of lithium bromide to absorb water vapor at a very low pressure. The evaporating water is cooled to a temperature low enough (about 45°F or 7°C) to provide the chilled water needed to produce air conditioning. The heat introduced into this cycle by the high-temperature

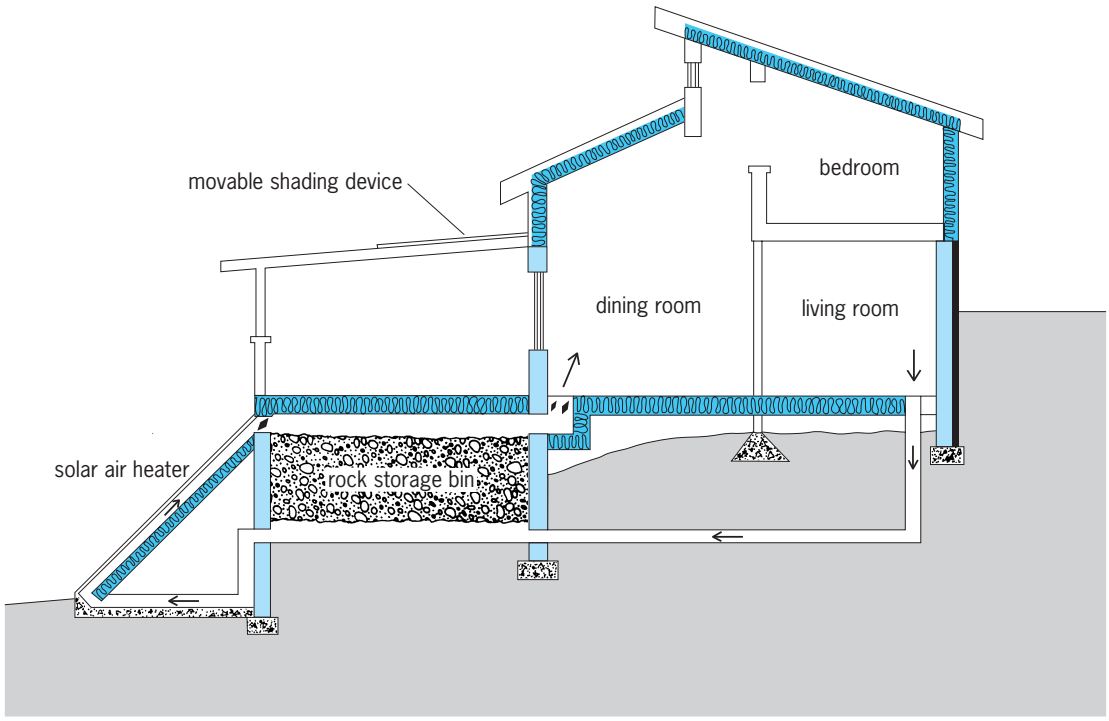

Fig. 8. Isolated-gain system with thermosyphon air heater and rock bed.

activating water and the low-temperature chilling water is removed by a third stream of water which has been cooled to below the outdoor air temperature by passing through an evaporative cooling tower. The cycle, simple in principle but much more complex in practice, is in wide use in large buildings where the necessary hot activating water is available either from solar collectors or as waste heat from other processes. Use of either Rankine cycle or absorption solar cooling systems for residences is restricted to a few experimental installations, but both hold promise for future development.

Passive cooling systems. Passive cooling systems make use of three natural processes: convection cooling with night air; radiative cooling by heat rejection to the sky on clear nights; and evaporative cooling from water surfaces exposed to the atmosphere. The effectiveness of each of these processes depends

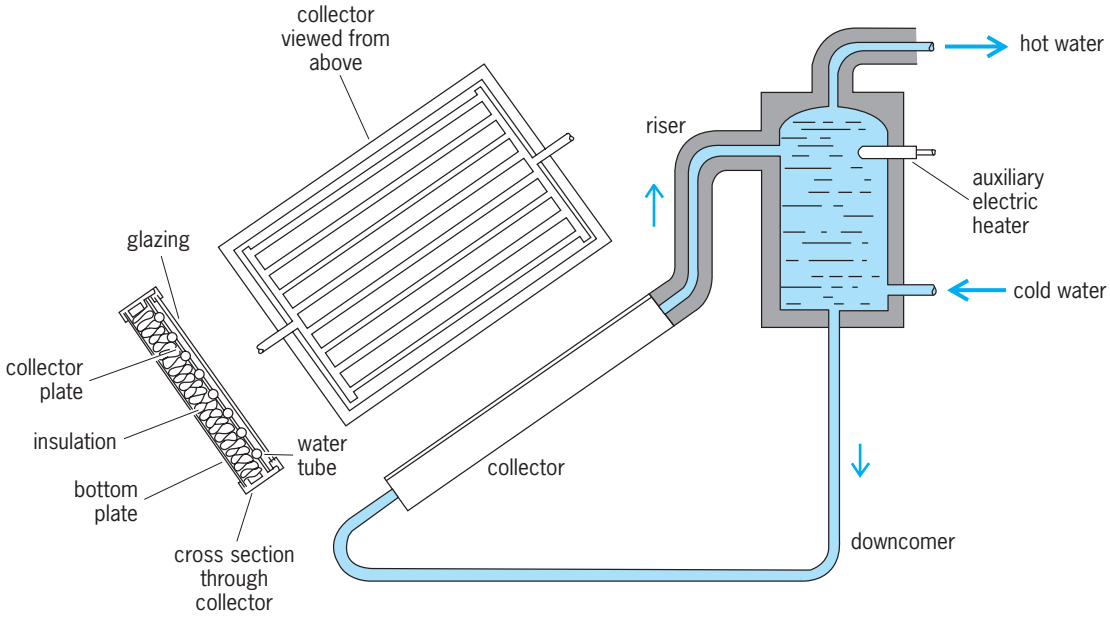

Fig. 9. Passive domestic water heater using thermosyphon system.

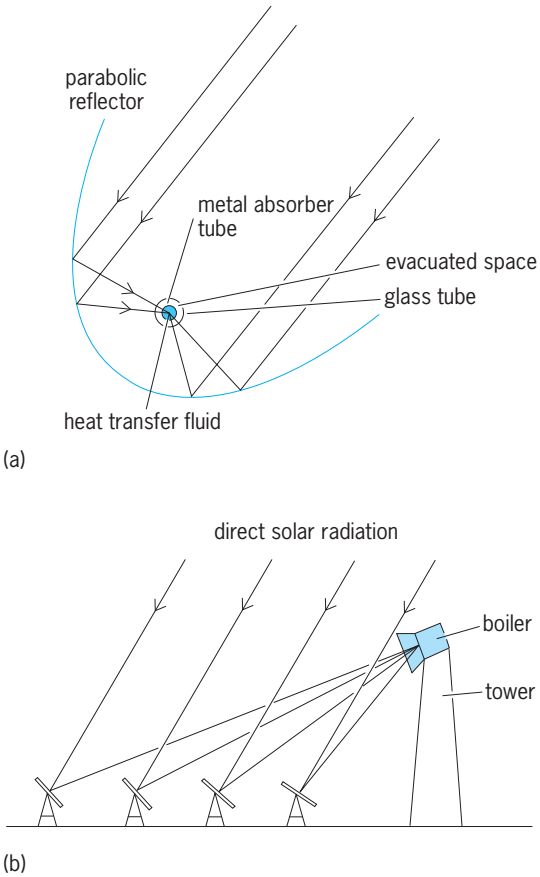

Fig. 10. Concentrating collectors. (*a*) Parabolic trough, for use in installations of moderate size. (*b*) Power tower, for large utility-type plants.

upon local climatic conditions. In the hot and dry deserts of the southwestern United States, radiation and evaporation are the two most useful processes. In hot and humid regions, other processes must be employed which involve more complex equipment that can first dehumidify and then cool the air.

Power Generation

Electric power can be generated from solar energy by two processes, the first of which uses the Rankine cycle described previously. In order to attain vapor temperatures high enough to make the cycle operate efficiently, concentrating collectors must be used. For installations of moderate size, parabolic troughs (**Fig. 10***a*) are adequate, but for very large utility-type plants the "power tower" concept (Fig. 10*b*) is preferred. Concentrators can use only the radiation which comes directly from the solar disk, so some means of tracking must be provided and energy must be stored during the daylight hours to ensure continuity of operation during intermittent cloudy periods. Storage systems for night-long operation are not yet available. Concentrating systems cannot make effective use of the diffuse radiation that comes from the sky, and so their use at the present stage of development of solar technology is limited to areas where bright sunshine is generally available.

Electric power can also be produced from solar radiation, without need for a thermal cycle, by photovoltaic cells which convert radiant energy directly into electricity. Originally used primarily in the United States and Soviet space programs, silicon solar batteries can now be produced at a cost low enough to justify their use in special circumstances for terrestrial applications. One development uses concentrators similar to those in Fig. 10*a* to enable a single solar cell to do the work of ten; the heat which must be removed to make the cells function efficiently can be used for many purposes requiring hot water or hot air. Solar cells produce direct current which can be stored in ordinary automobile-type batteries; the direct current can be converted to alternating current at the standard frequency, 60 Hz in the United States, and at any necessary voltage. One of the goals of the Department of Energy is the reduction of the cost of photovoltaic cells to the point where residential rooftop units become economically feasible. *See* SOLAR CELL.

Conclusion

Heating of space and service hot water is currently practical and economical. Passive heating and cooling are coming into wide use as the cost of alternative fuels continues to rise. Active cooling and power generation are too expensive for general use, but it is expected that the cost of these systems will diminish in the future. The major unsolved problem is how to store large quantities of energy for long periods of time. *See* ENERGY STORAGE. John I. Yellott

Bibliography. American Society of Heating, Refrigerating and Air Conditioning Engineers, *Active Solar Heating Systems Design Manual,* 1988; B. Anderson and M. Wells, *Passive Solar Energy: The Homeowner's Guide to Natural Heating and Cooling*, 2d ed., 1993; D. Halacy, *Understanding Passive Cooling Systems,* 1987; B. Norton, *Solar Energy Technology*, 1991.

Solar magnetic field

The magnetic field rooted in the Sun and extending out past the planets into the solar system. The field at the Sun's surface is detected remotely by its effect (the Zeeman effect) on atoms whose radiation is observed from Earth. This technique was first applied at the Mount Wilson Observatory in 1908 by G. E. Hale to detect the fields in sunspots. In 1952, H. D. Babcock and his son H. W. Babcock used a scanning technique to make the first magnetic maps of the entire visible disk of the Sun. Their daily "magnetograms" soon revealed a variety of magnetic features, including bipolar regions associated with sunspot groups, and unipolar regions whose fields extend far from the Sun and are responsible for recurrent geomagnetic activity at Earth. The field strengths range from a few gauss (a few hundred microtesla) in quiet areas to 3500 gauss (0.35 tesla) in sunspots. In 1962, the *Mariner 2* spacecraft, en route to Venus, made the first on-site sampling of the extended solar field in space. The average field strength was only 50 microgauss (5 nanotesla), reflecting the rapid (inverse

square) fall-off of field strength with distance from the Sun. *See* ZEEMAN EFFECT.

Bipolar regions and sunspots. Solar magnetic fields are related to the 11-year variation in the occurrence of sunspots. As the new sunspot cycle begins, concentrations of bipolar flux break through the Sun's surface in each hemisphere, beginning at about 40° latitude and gradually progressing toward the solar equator over the next few years. The emergence typically takes a few days, depending on the size of the region, but the subsequent spreading and decay may take several months so that a bipolar region may be visible from Earth again during the next 27-day rotation of the Sun on its axis. Thus, as the sunspot cycle advances and the eruption rate exceeds several per day, the sunspot belts become filled with the remnants of old bipolar magnetic regions.

Unipolar regions and polar fields. If the eruptions of bipolar flux were randomly oriented, the resulting patterns might be a complicated mixture of positive and negative polarities. However, the orientations are not random. The bipolar regions are oriented approximately east-west with their leading polarities (in the sense of solar rotation) all positive in the northern hemisphere and negative in the southern hemisphere during a given sunspot cycle (Hale's law). During the next 11-year cycle, the polarity scheme is reversed, with the leading polarities negative in the north and positive in the south. This systematic property of the sources of flux leads to systematic properties of the evolving magnetic patterns, and occasionally produces a large unipolar region whose opposite-polarity companion is far away.

In fact, the emerging bipolar magnetic regions are not oriented exactly east-west, but are tilted slightly so that their trailing polarities are closer to the Sun's poles and the leading polarities are closer to the equator. This small but systematic effect ultimately leads to the formation of unipolar regions at the Sun's poles, positive in one hemisphere and negative in the other, with an approximate 5-year lag between the peak in the eruption rate and the subsequent peaks in the strengths of the polar fields. When bipolar regions erupt with the opposite polarities during the next sunspot cycle, they will cause the polar fields to become reversed a few years later. This is why the solar magnetic cycle is referred to as a 22-year cycle rather than as an 11-year cycle like that of sunspots.

Mixed-polarity regions. Improvements in spatial resolution and sensitivity have revealed a mixed-polarity component of the solar magnetic field. This component acts like a weak background noise in which small-scale fields are continuously emerging, bumping into their counterparts of opposite polarity, and disappearing again. At sunspot minimum, this process involves more flux eruption than occurs in the larger, but less frequent, bipolar magnetic regions. However, because of its random occurrence, the mixed-polarity component seems to have little net effect on the large-scale solar magnetic field and its interplanetary extension.

Outward extension of the field. The occasional formation of large unipolar magnetic regions, isolated from their opposite-polarity counterparts by great distances, causes their magnetic field lines to extend far from the Sun without looping back to the surface. In these locations, material is able to stream outward from the Sun at relatively high speeds of 750 km/s (470 mi/h). When such open-field regions are located at low latitudes, their high-speed streams repeatedly sweep past Earth as the Sun rotates, inducing auroras and recurrent geomagnetic disturbances. *See* AURORA; GEOMAGNETIC VARIATIONS; MAGNETOSPHERE; SOLAR WIND.

The continual readjustment of the evolving patterns of magnetic fields leads to a variety of transient plasma processes, including solar flares, eruptive prominences, and coronal mass ejections. Without the solar magnetic field, none of these explosive phenomena would occur, and the Sun would be a much more placid late-type star.

Solar variability and other stars. As magnetic fields from bipolar regions spread out on the surface of the Sun, they heat the overlying chromosphere, causing an increase in ultraviolet radiation. The integrated (Sun-as-a-star) component of chromospheric radiation has been measured since 1976 and shows a clear 11-year variation, suggesting that the sunspot cycle would be visible to someone observing the Sun from elsewhere in the Milky Way Galaxy, and conversely that this technique could be used to see solar magnetic cycles in other stars. This idea motivated O. C. Wilson to begin such a survey at the Mount Wilson Observatory in 1966, and the continuation of this work has led to the identification of sunspot cycles in several late-type stars. *See* STELLAR MAGNETIC FIELD.

Origin of bipolar fields. It is presently believed that convection and differential rotation are responsible for the eruption of bipolar magnetic regions. The Sun's equator rotates faster than its polar caps and ought to wind up meridional field lines (like those that may join the Sun's poles), producing strong azimuthal fields below the sunspot belts. These strong fields would become evacuated to balance magnetic pressure against the external plasma pressure, and the lighter flux tube would rise to the surface (magnetic buoyancy) to form bipolar regions. The observed tilts of the bipolar regions are believed to result from cyclonal convection acting on the rising flux tubes. As increasing amounts of magnetic flux are packed more closely together, the very strong fields are believed to inhibit convection and produce the dark sunspots. *See* MAGNETISM; MAGNETOHYDRODYNAMICS; PLASMA (PHYSICS); SUN.

Neil R. Sheeley, Jr.

Bibliography. P. V. Foukal, *Solar Astrophysics*, 2d ed., Wiley, 2004; L. Golub and J. M. Pasachoff, *Nearest Star: The Surprising Science of our Sun*, Harvard University Press, 2001; E. J. Smith, Solar wind magnetic fields, in J. R. Jokipii (ed.), *Cosmic Winds and the Heliosphere*, University of Arizona Press, 1997; J. O. Stenflo, *Solar Magnetic Fields*, Kluwer Academic Publishers, 1994; M. Stix,

The Sun, 2d ed., Springer, 2004; H. Zirin, *Astrophysics of the Sun*, Cambridge University Press, 1988.

Solar neutrinos

Neutrinos produced in nuclear reactions inside the Sun. Neutrinos are produced as well in laboratory nuclear reactions. The first direct tests of how the Sun produces its luminosity (observed most conspicuously on Earth as sunlight) have been carried out by observing solar neutrinos. The results of these experiments confirm the theory of how the Sun shines and stars evolve. Moreover, the results show that neutrinos behave differently than is predicted by the standard model of particle physics.

Nuclear fusion in the Sun. The Sun shines because of fusion reactions similar to those envisioned for terrestrial fusion reactors. The basic solar process, called the proton-proton chain, is the fusion of four protons to form an alpha particle, two positrons (e^+), and two neutrinos (ν); that is, $4p \rightarrow \alpha + 2e^+ + 2\nu_e$ (see **table**). The rate for the initiating proton-proton (PP) reaction is predicted accurately ($\pm 1\%$) by the standard solar model. Unfortunately, the neutrinos from this reaction are below the energy thresholds for the first two experiments to detect solar neutrinos, with chlorine-37 (^{37}Cl) and with ultrapure water (H_2O). However, two experiments, one in Russia and one in Italy, which use gallium-71 (^{71}Ga), have successfully detected neutrinos from the PP reaction.

The proton-electron-proton (PEP) reaction, which is the same as the PP reaction except for having the electron in the initial state, is detectable in the ^{37}Cl experiment. The ratio of PEP to PP neutrinos is approximately independent of which solar model is used for the solar predictions. Two other reactions in the proton-proton chain are of special interest. The capture of electrons by ^{7}Be produces detectable neutrinos in the ^{37}Cl experiment. The ^{8}B decay is expected to be the main source of neutrinos for the ^{37}Cl experiment because of their relatively high maximum energy (14 MeV), although it is a rare reaction in the Sun. The ^{8}B neutrinos are the only substantial contributions to the experiments carried out with ultrapure water. There are also some less important neutrino-producing reactions from the carbon-nitrogen-oxygen (CNO) cycle, which will not be discussed here in detail since this cycle is believed to play a rather small role in the energy-production budget of the Sun. *See* CARBON-NITROGEN-OXYGEN CYCLES; NUCLEAR FUSION; NUCLEAR REACTION; PROTON-PROTON CHAIN.

Experimental tests. The first solar neutrino detector was based on the reaction $\nu_{\text{solar}} + {}^{37}\text{Cl} \rightarrow {}^{37}\text{Ar} + e^-$, which is the inverse of the electron-capture decay of argon-37 (^{37}Ar). This reaction was chosen for the first experiment because of its unique combination of physical and chemical characteristics, which are favorable for building a large-scale solar neutrino detector. Neutrino capture to form ^{37}Ar in the ground state also has a relatively low energy threshold (0.81 MeV) and a high sensitivity, nuclear properties that are important for observing neutrinos from ^{7}Be, ^{13}N, and ^{15}O decay and the PEP reaction.

Like all solar neutrino experiments, the ^{37}Cl detector was built deep underground to avoid cosmic-ray interactions in the detector that simulate the effects of neutrinos. The final detector system consists of an approximately 400,000-liter (100,000-gallon) tank of perchloroethylene, a pair of pumps to circulate helium through the liquid, and a small building to house the extraction equipment, all in the deep Homestake gold mine in South Dakota, 1500 m (5000 ft) below the surface. Sets of experimental runs were carried out in the ^{37}Cl experiment during the 1970s and 1980s, and continued into the 1990s. They have shown that the ^{37}Ar production rate in the tank is 0.48 ± 0.04 ^{37}Ar atoms per day. Even though the tank is nearly a mile (1.6 km) underground, a small amount of ^{37}Ar is produced by cosmic rays. The background rate has been estimated to be of order 0.1 ^{37}Ar atom per day.

Subtracting the cosmic-ray background, a positive signal of (2.6 ± 0.2) solar neutrino units (SNU) is inferred; 1 SNU $= 10^{-36}$ capture per target particle per second.

Principal reactions of the proton-proton chain in the Sun

Number	Reaction	Percentage of solar terminations*	Maximum neutrino energy, MeV
1†	$p + p \rightarrow {}^2\text{H} + e^+ + \nu$ or	99.75	0.420
2‡	$p + e^- + p \rightarrow {}^2\text{H} + \nu$	0.3	1.44 (monoenergetic)
3	${}^2\text{H} + p \rightarrow {}^3\text{He} + \nu$	100	
4	${}^3\text{He} + {}^3\text{He} \rightarrow {}^4\text{He} + 2p$ or	85	
5	${}^3\text{He} + {}^4\text{He} \rightarrow {}^7\text{Be} + \nu$	15	
6	${}^7\text{Be} + e^- \rightarrow {}^7\text{Li} + \nu$	15	0.861 (90%), 0.383 (10%) (both monoenergetic)
7	${}^7\text{Li} + p \rightarrow 2\,{}^4\text{He}$ or	15	
8	${}^7\text{Be} + p \rightarrow {}^8\text{B} + \nu$	0.02	
9	${}^8\text{Be} \rightarrow {}^8\text{Be}^* + e^+ + \nu$	0.02	14
10	${}^8\text{Be}^* \rightarrow 2\,{}^4\text{He}$	0.02	

*Percentage of solar terminations of the proton-proton chain in which this reaction occurs.
†PP reaction.
‡PEP reaction.

The predicted capture rates for the ^{37}Cl experiment (in SNU) for the standard solar model is as follows: PP reaction: 0; ^{8}B beta decay: 5.8; PEP reaction: 0.2; ^{7}Be electron capture: 1.2; ^{13}N decay: 0.1; and ^{15}O decay: 0.3. The total theoretical prediction is about 7.5 SNU. Many investigations have been undertaken of the best values to use for various input parameters for the calculation of the standard solar model. The calculated event rate is relatively insensitive to the input parameters, although the total predicted rate may well differ from 7.5 SNU by 1 SNU.

Solar neutrinos were also observed with the aid of electrons that the neutrinos scatter in a 700-metric-ton (770-short-ton) detector of pure water located in the Kamioka metal mine about 300 km (200 mi) west of Tokyo in the Japanese Alps. These results provided direct evidence that the observed neutrinos originate in the Sun; the electrons are scattered by neutrinos in the forward direction between the Earth and the Sun. Moreover, the neutrino events are registered at the exact time they are detected, making possible a sensitive search for possible time dependencies. The measurement of the Kamiokande collaboration yielded the result given by Eq. (1). The

$$\frac{\text{Observed}}{\text{Predicted}} = 0.55 \pm 0.08 \qquad (1)$$

indicated error is a 1-standard-deviation combined statistical and systematic uncertainty. The predicted event rate in the equation stands for the best estimate of the number of events calculated to occur according to the standard model of the Sun and the standard model of how neutrinos behave. The Kamioka result is sensitive only to high-energy neutrinos from ^{8}B decay.

The Super-Kamiokande detector is a much larger version of the original Kamiokande detector and is located in the same mine. The active detector in the Super-Kamiokande experiment contains 32,000 metric tons (35,000 short tons) of pure water and 11,200 specially designed tubes for observing electrons scattered by neutrinos. This more powerful superdetector finds a rate given by Eq. (2). The new rate is in

$$\frac{\text{Observed}_{\text{SuperK}}}{\text{Predicted}} = 0.46 \pm 0.01 \qquad (2)$$

good agreement with the previous measurement but is more precise.

Since the Japanese water experiments (Kamiokande and Super-Kamiokande) and the chlorine experiment are both primarily sensitive to the same ^{8}B neutrinos, it was difficult to understand why the discrepancy between observation and calculation is greater for the chlorine than for the water experiments. The favored interpretation is that neutrinos do not behave as predicted by the standard model of particle physics; neutrinos change their type as they travel from the center of the Sun to detectors on Earth. The chlorine detector is sensitive only to neutrinos associated with electrons, electron-type neutrinos, while the water detectors have some sensitivity to all types of neutrinos. This interpretation has been strengthened by the fact that the Super-Kamiokande experiment finds strong evidence that neutrinos produced by cosmic rays in the atmosphere change their type when they travel distances comparable to the Earth's radius.

Experiments have also been performed with gallium detectors located underneath a mountain in the Caucus region of Russia [60 metric tons (66 short tons) of gallium, SAGE experiment] and in an underground laboratory about an hour's drive north of Rome [30 metric tons (33 short tons) of gallium, GALLEX/GNO experiment]. These experiments are similar to the chlorine experiment. The two gallium experiments and the chlorine experiment all use radiochemical techniques to extract a small number of radioactive nuclei produced by neutrinos from huge tanks containing enormous numbers of atoms of chlorine or gallium. Unlike the Kamiokande and Super-Kamiokande experiments, the gallium and chlorine detectors do not determine specific energies for the neutrinos they detect. Instead, the radiochemical gallium and chlorine detectors specify only a minimum energy: 0.23 MeV for gallium and 0.81 MeV for chlorine.

The results from the gallium experiments reinforce the conclusion that new physics—adjustments or changes to the current understanding of the standard model of particle physics—is required. The observed rate is 69 ± 4 SNU, which is much less than the standard-model predicted rate of about 130 SNU. Almost all of the observed rate in the gallium experiments could be attributed to the fundamental low-energy neutrinos from the PP and PEP reactions. The rate expected from these neutrinos is largely independent of astronomical uncertainties. If the PP and PEP neutrinos are being detected at the expected rate, there is practically no room in the experimental results to accommodate the neutrinos from ^{7}Be, whose number is also believed to be reliably calculated with the solar models, and the ^{8}B neutrinos, which are actually observed in the Kamiokande and Super-Kamiokande experiments.

Possible explanations. Many explanations were advanced for the discrepancy between the observed and the predicted event rates in solar neutrino experiments. These explanations can be divided into three classes: (1) the standard solar model must be significantly modified; (2) something is seriously wrong with the experiments; (3) the standard model of how neutrinos behave must be significantly modified.

No one has succeeded in modifying a solar model so that it is consistent with all of the experimental data on solar neutrinos, as well as the astronomical data on the Sun and the laboratory measurements of the parameters used in constructing the solar models. Moreover, precise measurements of the thousands of frequencies with which the Sun pulsates on its surface (characteristic period of order 5 min) have confirmed to an accuracy of 0.1% the predictions of the standard solar model for these pulsation frequencies. This agreement convinced researchers that the standard solar model is an accurate description of the Sun.

All of the solar neutrino experiments have been examined carefully by many different researchers. A variety of checks have been made to test whether there was a significant error or a large uncertainty in one of the experiments that might explain the difference between prediction and observation. No significant previously unknown errors or uncertainties have been found. Moreover, the gallium experiments have been tested in the most direct fashion possible. Intense laboratory sources of neutrinos have been placed near the gallium detectors, and the expected number of events has been observed from these artificial sources. The consensus among scientists in the field is that the solar neutrino experiments are yielding a valid although surprising result.

The only remaining possibility is that the theory of how the neutrino behaves must be changed. Indeed, in March 2000 the results of a decisive experiment were published that showed unequivocally that solar neutrinos change their type on their way to Earth from the center of the Sun.

SNO experiment and combined results. An experiment called SNO (Sudbury Neutrino Observatory) uses 1000 metric tons (1100 short tons) of heavy water in a nickel mine in Sudbury, Ontario, Canada. This experiment began operating in the spring of 1999. It measures both the number of neutrinos of the electron type (associated with electrons) and the number of neutrinos of all types (associated with electrons, muons, and tau particles). In March 2000, the SNO collaboration combined their measurements of just electron-type neutrinos with the measurements from the Super-Kamiokande ordinary-water experiment (some sensitivity to neutrinos of all types) to obtain the total number of solar neutrinos from ^{8}B that reach the Earth. The SNO and Super-Kamiokande result announced in 2001 is given by Eq. (3), where the error shown reflects only

$$\frac{\phi}{\phi_{\text{predicted}}} = 1.08 \pm 0.20 \qquad (3)$$

the measurement uncertainty. Unlike the previous measurement made by Super-Kamiokande alone (for which the ratio of measured to predicted neutrino rate was 0.55), the new SNO plus Super-Kamiokande determination includes all of the ^{8}B solar neutrinos with equal sensitivity. The Super-Kamiokande experiment is mostly sensitive to electron-type neutrinos. By a series of improvements in the experimental measurements and the use of a more powerful analysis technique, the comparison between theory and measurement was refined, with the result, announced in September 2003, given by Eq. (4).

$$\frac{\phi}{\phi_{\text{predicted}}} = 1.01 \pm 0.04 \qquad (4)$$

An empirical value for the number of PP neutrinos has also been extracted from an analysis that includes all the solar neutrino experiments and relevant experiments with reactor neutrinos. The measured number of PP neutrinos is given by Eq. (5).

$$\frac{\phi\,(PP)}{\phi(PP)_{\text{predicted}}} = 1.02 \pm 0.02 \qquad (5)$$

All of the results from solar neutrino experiments are consistent with the conclusion that the standard solar model predicts accurately the number of neutrinos of different energies that are emitted by the Sun but that some of the neutrinos change their type on the way from the center of the Sun to the detectors on Earth. *See* NEUTRINO; STANDARD MODEL; STELLAR EVOLUTION; SUN.

John N. Bahcall

Bibliography. N. Abdurashitov et al., Solar neutrino flux measurements by the Soviet-American Gallium Experiment (SAGE) for half the 22-year solar cycle, *J. Exp. Theor. Phys.*, 95:181–193, 2002; Q. R. Ahmad et al., Measurement of the rate of $\nu_e + d \rightarrow p + p + e^-$ interactions produced by ^{8}B solar neutrinos at the Sudbury Neutrino Observatory, *Phys. Rev. Lett.*, 87:071301-1 to -6, 2001; M. Altmann et al., GNO solar neutrino observations: Results for GNO I, *Phys. Lett. B*, 490:16–26, 2000; J. N. Bahcall and C. Peña-Garay, A road map to solar neutrino fluxes, neutrino oscillation parameters, and tests for new physics, *J. High Energy Phys.*, 11004-0 to 11004-47, 2003; B. T. Cleveland et al., Measurement of the solar electron neutrino flux with the Homestake chlorine detector, *Astrophys. J.*, 496:505–526, 1998; S. Fukuda et al., Solar ^{8}B and hep neutrino measurements from 1258 days of Super-Kamiokande data, *Phys. Rev. Lett.*, 86:5651–5655, 2001; Y. Fukuda, Final solar neutrino results at Kamiokande, *Phys. Rev. Lett.*, 77:1683–1686, 1996.

Solar radiation

The electromagnetic radiation and particles (electrons, protons, alpha particles, and rarer heavy atomic nuclei) emitted by the Sun. The electromagnetic radiation covers a wavelength range from x-rays to radio waves, that is, from about 0.01 nanometer to 30 km. The annual mean irradiance at Earth, integrated over the whole spectrum, amounts to $1365\ \text{W} \cdot \text{m}^{-2}$, and 99% of its energy is carried by radiation with wavelengths between 278 and 4600 nm, with the maximum at 472 nm. The division into two energetically equal parts is at a wavelength of 731 nm. Although the total solar irradiance is often called the solar constant, it varies with time, for example, with the solar activity cycle with a period of about 11 years. These variations have been monitored since 1978 by radiometers on space platforms and amount to about 0.1%, with the radiation higher during the maximum of solar activity. There are also shorter-term variations such as decreases due to the passage of dark sunspots or increases due to faculae. *See* ELECTROMAGNETIC RADIATION; SOLAR CONSTANT.

The Sun also emits a continuous stream of particles, the solar wind, which originates in coronal holes and the upper corona. The density and speed of solar wind particles, and thus the solar wind's

influence on the Earth environment, such as the ionosphere, vary with the solar cycle. This wind also carries magnetic fields from the surface of the Sun into interplanetary space, and these fields in turn shield the Earth from the highly energetic cosmic rays coming from outside the solar system. Since these fields are stronger during periods of high solar activity, the shielding becomes more efficient and less cosmic rays reach the Earth at these times. On the other hand, high activity means also increased occurrence of explosive events on the Sun: the solar flares and coronal mass ejections. Solar flares are produced by the most powerful explosions, releasing energies of up to 10^{25} joules in 100–1000 s and high-speed electrons that emit intense radiation at radio and x-ray wavelengths. They also produce nuclear reactions in the solar atmosphere with the emission of gamma rays and of neutrons that move nearly at the speed of light. Coronal mass ejections expand away from the Sun at speeds of hundreds of kilometers per second, becoming larger than the Sun and removing up to 5×10^{13} kg of coronal material. Both events are believed to be ignited by the reconnection of magnetic fields, and the emitted particles are much more energetic and numerous than those of the solar wind. If these particles reach the Earth, they give rise to the aurora at high latitudes, and they can damage satellites, endanger humans in space, and on the Earth disturb telecommunications and even disrupt power systems. Such influences on the Earth environment are due to "space weather." Its prediction, and thus the continuous observation of the Sun, is of major importance for the functioning of civilization. *See* AURORA; COSMIC RAYS; IONOSPHERE; MAGNETOSPHERE; SOLAR WIND; SUN. Claus Fröhlich

Bibliography. J. Burch, The fury of space storms, *Sci. Amer.*, 284(4):72–81, April 2001; L. Golub and J. M. Pasachoff, *Nearest Star: The Surprising Story of Our Sun*, Harvard University Press, 2001; K. R. Lang, *SOHO* reveals the secrets of the Sun, *Sci. Amer.*, 276(3):32–47, March 1997; C. J. Schrijver and A. M. Title, Today's science of the Sun, *Sky Telesc.*, 101(2):35–40, February 2001, and 101(3):35–39, March 2001.

Solar system

The Sun and the bodies moving in orbit around it. The most massive body in the solar system is the Sun, a typical single star that is itself in orbit about the center of the Milky Way Galaxy. Nearly all of the other bodies in the solar system—the terrestrial planets, outer planets, asteroids, and comets—revolve on orbits about the Sun. Various types of satellites revolve around the planets; in addition, the giant planets have orbiting rings. The orbits for the planets appear to be fairly stable over long time periods and hence have undergone little change since the formation of the solar system. It is thought that some 4.56×10^9 years ago a rotating cloud of gas and dust collapsed to form a flattened disk (the solar nebula) in which the Sun and other bodies formed. The bulk of the gas in the solar nebula moved inward to form the Sun, while the remaining gas and dust are thought to have formed all the other solar system bodies by accumulation proceeding through collisions of intermediate-sized bodies called planetesimals. Planetary systems are believed to exist around many other stars in the Milky Way Galaxy. Solid evidence for the existence of Jupiter-mass planets around nearby solarlike stars now exists. *See* PLANET.

Composition. The Sun is a gaseous sphere with a radius of about 7×10^5 km (4×10^5 mi), composed primarily of hydrogen and helium and small amounts of the other elements. The Sun is just one of about 2×10^{11} stars in the Milky Way Galaxy. The solar system lies in the disk of the Galaxy and moves around the galactic center about once every 2×10^8 years on a circular orbit with a radius close to 3×10^4 light-years (about 3×10^{17} km or 2×10^{17} mi). The Sun is somewhat peculiar in that it is a single star, unaccompanied by another star; most stars in the disk of the Galaxy are in double or multiple systems where two or more stars orbit about their common center of mass. The Sun's mass of 2×10^{30} kg (4.4×10^{30} lb), however, is quite typical; stellar masses range from about 0.08 to about 60 times the Sun's mass, with the great majority of stars being similar in mass (and hence in other characteristics) to the Sun. *See* BINARY STAR; MILKY WAY GALAXY; STAR; SUN.

The terrestrial planets (Mercury, Venus, Earth, and Mars) are the closest to the Sun, with orbital radii ranging from 0.39 AU for Mercury to 1.5 AU for Mars. One astronomical unit (AU), the distance from Earth to Sun, equals 1.5×10^8 km or 9.3×10^7 mi. The terrestrial planets are composed primarily of silicate rock (mantles) and iron (cores). The Earth is the largest terrestrial planet, with an equatorial radius of 6378 km (3963 mi) and a mass of 5.972×10^{24} kg (1.317×10^{25} lb); Mercury is the smallest, with a mass of 0.055 times that of Earth. *See* EARTH; MARS; MERCURY (PLANET); PLANETARY PHYSICS; VENUS.

The outer planet region begins with Jupiter at a distance of 5.2 AU from the Sun. The outer planets are subdivided into the gas giant or Jovian planets (Jupiter and Saturn), the ice giant planets (Uranus and Neptune), and Pluto. By far the largest planet is Jupiter, with a mass 318 times that of the Earth, while the other giant planets are more massive by a factor of 15 or more than Earth. Jupiter and Saturn are composed primarily of hydrogen and helium gas, like the Sun, but with rock and ices, such as frozen water, methane, and ammonia, concentrated in their cores. Uranus and Neptune also have rock and ice cores surrounded by envelopes with smaller amounts of hydrogen and helium. Pluto, slightly smaller than the Earth's Moon, is probably composed primarily of rock and ice. *See* JUPITER; NEPTUNE; PLUTO; SATURN; URANUS.

The region between Mars and Jupiter is populated by a large number of rocky bodies called asteroids. The asteroids are smaller than the terrestrial planets, with most known asteroids being about 1 km (0.6 mi) in radius, though a few have radii of

hundreds of kilometers. Some asteroids have orbits that take them within the orbits of Earth and the other terrestrial planets, leading to the possibility of impacts with these planets; asteroids in the main belt beyond Mars are occasionally perturbed into these Earth-crossing orbits. Small fragments of asteroids (or comets) that impact the Earth first appear as meteors in the sky; any meteoric material that survives the passage through the Earth's atmosphere and reaches the surface is called a meteorite. The Allende meteorite, which fell in Mexico in 1969, contains inclusions that have been radiometrically dated at 4.56×10^9 years; this age, the oldest found so far, is presumably the age of the solar system. Along with the lunar rocks returned by the Apollo missions, Allende and other meteorites are samples of other bodies in the solar system. In addition, primitive components of meteorites such as Allende give important information about conditions in the very early solar system. A few meteorites are thought to have come from Mars or the Moon. *See* ASTEROID; METEOR; METEORITE.

Comets are icy bodies (so-called dirty snowballs) with diameters on the order of 10 km (6 mi). In contrast to the orbits of most planets, cometary orbits often are highly elliptical and have large inclinations that take them far above and below the plane where the planets orbit. The region well beyond Pluto's orbit is populated with a very large number (perhaps 10^{12}) of comets, out to a limiting distance of about 10^5 AU, at which point external forces (due to nearby stars, passing molecular clouds, and the Milky Way Galaxy itself) disturb any body trying to stay in a stable orbit around the Sun. The distribution of comets within this huge volume, the Oort Cloud, is uncertain; there may be an inner cloud extending outward to about 2×10^4 AU, and an outer cloud beyond that. In addition, comets have been detected orbiting in the plane of the solar system at distances of 30 to 55 AU; this flattened distribution is called the Edgeworth-Kuiper Belt, named for the scientists who predicted its existence in 1949 and 1951. The largest objects in the Edgeworth-Kuiper Belt can now be observed from Earth, with diameters in the range from 100 to 1000 km (60 to 600 mi); Pluto is one the largest bodies in this family of icy objects. Comets that enter the inner solar system are eventually evaporated by solar radiation, leaving behind interplanetary dust particles and perhaps rocky cores similar to certain asteroids. *See* COMET; KUIPER BELT.

Satellites orbit around all of the planets except Mercury and Venus. The Earth and its rocky Moon are unusual in that the Moon is about one-fourth the size of the Earth; nearly all other satellites are much smaller than their planet, with the exception of Pluto and its satellite Charon, which is half the size of Pluto. While Mars has only two tiny, rocky satellites, the Voyager missions, followed by ground-based observations, have shown that the giant planets all have large systems of satellites (Jupiter has at least 60), ranging in size from 0.5 to 2700 km (0.3 to 1700 mi) in radius and composed of varying amounts of rock and ice. Each giant planet also has a number of rings of particles whose orbits are interspersed with and controlled by the planet's innermost satellites. *See* MOON; SATELLITE (ASTRONOMY).

Origin. A number of imaginative theories of the origin of the solar system have been presented in the last several hundred years, but nearly everyone presently working on solar system cosmogony agrees on one fundamental concept, advanced in 1796 by P. S. de Laplace. This concept, the nebular hypothesis, holds that the Sun and the rest of the bodies in the solar system formed from the same rotating, flattened cloud of gas and dust, now called the solar nebula. The nebular hypothesis at once explains the gross orbital properties of the solar system: all planets orbit (and most rotate) in the same sense as the Sun rotates, with their nearly circular orbits being confined largely to a single plane almost perpendicular to the Sun's rotation axis.

While the grand experiment that led to the formation of the solar system can never be repeated, present-day regions of star formation in the Milky Way Galaxy can be observed to gain insight into the processes involved in the formation of stars similar in mass to the Sun. Assuming that present-day star formation proceeds much as it did 4.56×10^9 years ago, this approach should yield many constraints on the general process. These observations confirm the stellar implications of Laplace's nebular hypothesis: very young stars (protostars) are indeed found embedded in dense clouds of gas and dust that often show evidence for flattening and rotation. While understanding of solar system formation is almost certain to undergo further evolution, the following summarizes the basic framework in which cosmogonists are working.

Collapse phase. The solar nebula was produced by the collapse of a dense interstellar cloud (**illus.** *a*) composed primarily of molecular hydrogen (about 77% by mass) and helium gas (about 21%), with about 2% of the cloud mass being in the form of the other elements, mostly frozen into dust grains. Elements heavier than hydrogen and helium were formed by nucleosynthesis in previous generations of stars and then ejected into interstellar clouds through events such as supernovae. *See* NUCLEOSYNTHESIS; SUPERNOVA.

Radio telescopes have shown that dense interstellar clouds exist with masses comparable to that of the Sun, containing about 10^4–10^6 molecules/cm^3, and with quite low temperatures, about 10 K (-263°C or -441°F). Because of their low temperatures, gas pressure generally is not sufficient to support these clouds against their own self-gravity, which tries to pull the cloud into a smaller configuration. Many clouds appear to be supported primarily by magnetic fields; however, magnetic field support requires a continual contraction of the largely neutral cloud past the field lines, and eventually magnetic forces can no longer support the increasingly dense cloud. At this point, dense clouds enter the collapse phase, where supersonic inward motions develop that lead to the formation of a stellar-sized core at the center of

the cloud in about 10^5–10^6 years. Some of the gas in the cloud continues to fall onto this protostellar core, where it is suddenly stopped and its kinetic energy is converted into thermal energy. This heat leaves the cloud as infrared radiation; infrared telescopes have shown that many dense cloud cores contain deeply embedded young stars. *See* INFRARED ASTRONOMY; INTERSTELLAR MATTER; RADIO ASTRONOMY; STELLAR EVOLUTION.

Solar nebula. In a rotating cloud, not all of the infalling gas and dust falls directly onto the central protostar, because of the conservation of angular momentum. Instead, a disklike solar nebula forms, where the inward pull of gravity is expended in maintaining the rotational motion of the nebula. If too much angular momentum is present in the initial cloud, the collapse process is likely to result in the formation of a double or multiple star system instead of a single star and a planetary system; if extremely little angular momentum is present, only a single star forms.

Because the Sun contains 99.9% of the mass of the solar system but only about 2% of the angular momentum, the planetary system has considerably more angular momentum per unit mass than the Sun. This discrepancy is more than can be accounted for by forming the Sun out of the lowest-angular-momentum portions of a dense cloud, or by allowing for the fact that the Sun has lost substantial angular momentum during the age of the solar system through magnetic braking associated with the solar wind. *See* ANGULAR MOMENTUM; SOLAR WIND.

The solution to this angular momentum problem appears to lie in substantial evolution of the solar nebula after its formation. In order to have enough angular momentum to lead to planetary system formation, much of the mass of the newly formed solar nebula must reside in the disk rather than in the early Sun. The disk must then evolve in such a way as to transfer mass inward to feed the growing Sun, while transporting outward the excess angular momentum undesired by the Sun but required for the planets. While this sort of evolution may appear to be contrived if not miraculous, it is actually to be expected on very general grounds for any viscous disk that is undergoing a loss of energy, as the solar nebula will, through radiation to space. Several different physical mechanisms have been identified that potentially are capable of driving solar nebula evolution: turbulence caused by convective instability or velocity shear, gravitational torques between a bar-shaped protosun and spiral arms in the nebula, and magnetic torques associated with remanent magnetic fields or with a nebula magnetic field generated by the dynamo mechanism. These mechanisms are likely to lead to evolution of the bulk of the nebula gas inward onto the protosun (illus. *b*), on a time scale on the order of 10^5–10^7 years. Magnetic fields are the currently favored driver of nebula evolution.

Planetesimal formation. The portion of the nebula that is to form the planets must decouple from the gaseous nebula to avoid being swallowed by the Sun. This occurs by the process of coagulation of dust grains through mutual collisions; when solid bodies become large enough (roughly kilometer-sized), they will no longer be tied to the nebula through brownian motion (as is the case with dust grains) or gas drag (as happens with smaller bodies).

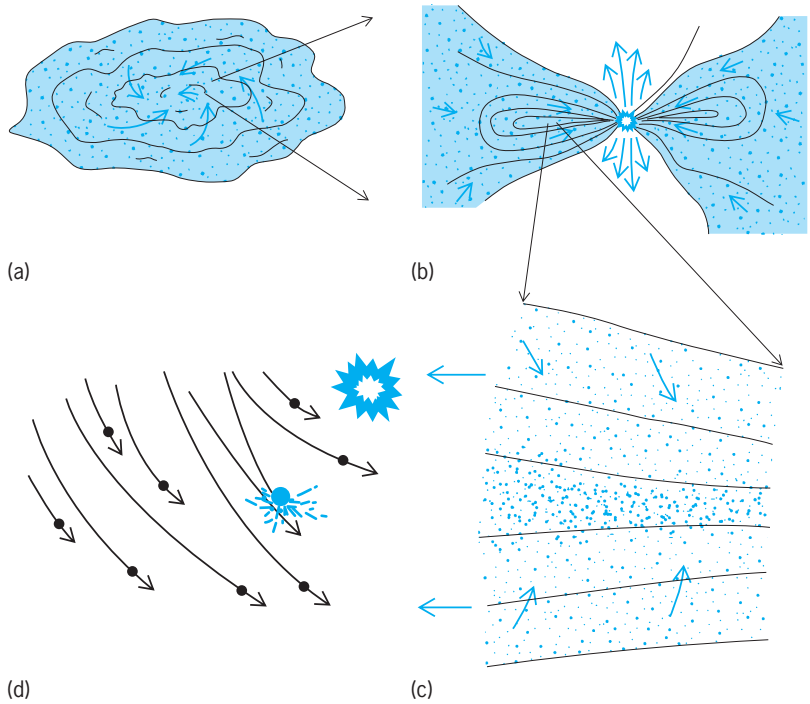

Schematic diagram of four phases of the formation of the solar system. (*a*) A dense, rotating interstellar cloud collapsed to form the solar nebula, a flattened cloud of gas and dust. (*b*) Most of the gas in the solar nebula flowed onto the growing protosun, whose energetic wind first flowed outward in a bipolar pattern and later removed the last vestiges of the nebula. (*c*) Dust grains sedimented to the nebula midplane and coagulated into kilometer-sized planetesimals. (*d*) Gravitational forces between the planetesimals resulted in collisions and their accumulation into planetary-sized bodies.

The interstellar dust grains that eventually form the solid bodies in the solar system start out with sizes on the order of 0.1 micrometer. These minuscule particles must stick together in order to begin the planet formation process; sticking can be caused by van der Waals forces. Coagulation is enhanced by the increase in the spatial density of dust grains, which occurs as the dust grains sediment down through the gaseous nebula to form a dense subdisk composed largely of solid particles (illus. *c*). Both dust grain sedimentation and coagulation are inhibited by the vigorous stirring expected in a highly turbulent nebula (the latter because grains may hit each other at relative velocities that are high enough to produce fragmentation rather than coagulation), but once the turbulence dies down, sedimentation and coagulation can produce centimeter- to meter-sized bodies in about 10^3–10^4 years. *See* INTERMOLECULAR FORCES.

The next phase of growth has long been thought to involve a collective gravitational instability of the dust subdisk that would rapidly produce bodies of kilometer size or larger, termed planetesimals. However, this instability may be prevented by turbulence induced by the difference in orbital speed between the gaseous nebula and the dust subdisk. (The gas orbits the Sun somewhat more slowly than the solid

bodies because the gaseous disk is partially supported against the Sun's gravity by gas pressure.) If the gravitational instability is prevented, then growth to planetesimal size may occur through further collisional coagulation; gas drag preferentially slows down smaller bodies, producing relative motions between larger and smaller bodies that can lead to collisions. In this case, growth to kilometer-sized planetesimals is expected to occur in about 10^4 years. There is currently much debate as to whether or not a dust disk instability was actually involved in growth to this size.

Planetary accumulation. About 10^{12} kilometer–sized planetesimals are needed to form just the terrestrial planets; significantly greater numbers of similarly sized bodies would be needed to form the giant planets. These plantesimals are already roughly the size of many asteroids and comets, suggesting that many of these bodies are simply leftovers from intermediate phases of the planet formation process.

Planetesimals are massive enough for gravitational forces to determine their collision probabilities; self-gravity is also needed to prevent debris from escaping following collisions. The subsequent growth of the planetesimals through gravitational accumulation (illus. *d*) is in two distinct phases. In the first phase, planetesimals grew by accumulation of other planetesimals at essentially the same distance from the Sun. Once the nearby planetesimals were all swept up, this phase ended. The first phase lasted about 10^4 years in the inner solar system and produced planetesimals about 500 km (300 mi) in size. This phase is likely to have been characterized by the runaway growth of relatively few bodies, because once one body becomes larger than its neighbors (as must happen after any collision occurs), its increased gravitational pull increases the chances that another body will impact it, and so on.

In the second phase, accumulation requires bodies at significantly different distances from the Sun to collide. This can happen only if the planetesimal orbits become highly elliptical and hence intersecting. Gravitational forces between the orbiting planetesimals can produce elliptical orbits, but only over relatively long time periods; about 10^7–10^8 years is required for accumulation to proceed by this means in the inner solar system. The second phase is currently thought to have proceeded differently from the first phase, with collisions occurring primarily between more or less equal-sized bodies, rather than the highly unequal-sized collisions that occur in a runaway accumulation process. The second phase may then have involved violent collisions between planetary-sized bodies, a spectacular finale to the entire planet formation process. A glancing collision between a Mars-sized and an Earth-sized body appears to be the best explanation for the formation of the Earth-Moon system; debris from the giant impact would end up in orbit around the Earth and later form the Moon.

Astronomical observations of young stars similar to the Sun imply that the gaseous portion of the solar nebula was removed some time between 10^5 and 10^7 years after the Sun began to form. While most of the gas presumably was added to the Sun through nebula evolution, the residual nebula gas and dust were probably removed from the solar system by the early solar wind, which had a mass loss rate roughly 10^6 times larger than at present (illus. *b*). The final formation of the terrestrial planets, lasting 10^7–10^8 years, then occurred in the absence of appreciable nebula gas, which explains the absence of significant hydrogen and helium in the inner solar system. However, formation of the giant planets must have occurred prior to the loss of the nebula gas; otherwise their rock and ice cores would not have been able to capture the gas needed to account for their present compositions. Formation of the giant planets within about 10^6 years appears to require that the second phase of planetesimal accumulation proceeded through runaway accretion all the way to bodies about 10 earth masses in size in the outer solar system. The satellite systems of the giant planets were largely formed in minisolar nebulae orbiting around each protoplanet, through processes that are similar to those described for planet formation.

Forming gas giant planets by the two-step process requires about 10^7 years for a 10-earth-mass core to form and then accrete a massive gaseous envelope, which may be longer than the life of typical gaseous disks. The alternative means for forming the gas giant planets is much more rapid, requiring only about 10^3 years for a gravitational instability of the gaseous nebula to produce a massive clump of gas and dust. The dust will settle to form a core at the center of the clump on a similar time scale. If gas giant planets form by the latter mechanism, then even the youngest stars will show evidence for Jupiter-mass companions; while if the two-step mechanism predominates, most young stars will not be old enough to have such companions. The formation of ice giant planets could occur in the disk instability mechanism if the solar nebula was born in a region of high-mass stars. Their radiation would photoevaporate the disk gas beyond Saturn's orbit as well as the gaseous envelopes of the protoplanets orbiting there.

Extrasolar planetary systems. Confirmation of the implications of the nebular hypothesis for planetary formation requires the detection of planetary systems orbiting around other stars; if the basic theory of solar system formation is correct, planetary systems should not be rare in the Milky Way Galaxy.

A new era in the study of planetary origins emerged in 1995, when Michel Mayor and Didier Queloz found evidence for a Jupiter-mass planet in orbit around the Sun-like star 51 Pegasi. Since then, over 130 other Jupiter-mass planets have been discovered, mostly by Geoffrey Marcy and Paul Butler, with orbital distances ranging from a surprising 0.04 AU to about 6 AU; the latter was to be expected based on the solar system, but not the former. These discoveries appear to show that some planets can migrate inward toward their stars following their formation. Other surprises undoubtedly await as additional extrasolar planets are discovered and characterized.

Alan P. Boss

Bibliography. A. P. Boss, *The Race to Find New Solar Systems*, Wiley, 1998; J. J. Lissauer, Planet formation, *Annu. Rev. Astron. Astrophys.*, 31:129–174, 1993; V. Mannings, A. P. Boss, and S. S. Russell (eds.), *Protostars and Planets IV*, 2000; V. S. Safronov, *Evolution of the Protoplanetary Nebula and Formation of the Earth and the Planets*, 1969, transl., NASA TTF-677, 1972; S. R. Taylor, *Solar System Evolution*, 2d ed., 2001; G. W. Wetherill, Formation of the Earth, *Annu. Rev. Earth Planet. Sci.*, 18:205–256, 1990.

Solar wind

A flow of completely ionized material away from the Sun. As this plasma moves through space, it creates a bubble known as the heliosphere. The processes that heat and accelerate the solar wind are not well understood at present. At heights of a few solar radii above the Sun's visible surface, the flow becomes supersonic, reaching speeds of 350–750 km/s (220–470 mi/s) in interplanetary space. The solar wind finally terminates at a distance of perhaps 150 astronomical units (1 AU is the distance from the Sun to Earth), where it interacts with the local interstellar medium.

This flow can be conveniently categorized as either ambient or time-dependent. The combination of solar rotation together with the ejection of material of different speeds from the Sun leads to a dynamical interaction between neighboring parcels of plasma and the formation of interaction regions and shocks that bound them. This forms the basis of the ambient solar wind. These structures appear stationary in a frame of reference rotating with the Sun and are thus known as corotating interaction regions (CIRs). Coronal mass ejections (CMEs) represent the most spectacular form of time-dependent phenomena.

During the course of the 11-year solar cycle, the structure of the solar wind changes at a fundamental level. At the minimum of the solar activity cycle a simple flow pattern exists consisting of (1) a slow wind about the magnetic equator (tilted or warped about the rotational equator), and (2) a fast wind emanating from large polar coronal holes. In contrast, at solar maximum the flow is considerably more complex and time-dependent. The rate, launch site, and properties of coronal mass ejections also vary with the phase of the solar cycle. The interaction of the solar wind with the Earth's geomagnetic environment represents an important and practical application of space physics known as space weather, which seeks to describe the conditions in space that affect both the Earth and its technological systems.

The study of the solar wind is a relatively new topic. As recently as the 1950s, its mere existence was speculation. Now, in part through the successful launch of a number of innovative spacecraft, considerably more is known. The first clue that the Sun was emitting a continuous wind can be traced back to observations of comet tails, which were found to always point away from the Sun no matter whether the comet was moving toward or away from the Sun. It turns out that comets often have two tails: a dust tail, produced by radiation pressure (that is, sunlight), and an ion tail. Differences in the properties of these two tails led L. Biermann in 1957 to propose the presence of a continuous stream of particles propagating away from the Sun that pushed the ions making up the ion tail. The following year, E. N. Parker derived a theory for the solar corona that required there to be a flow of plasma away from it. (Previously it had been thought that the extended atmosphere of the Sun was static and gravitationally bound to the Sun.) He named this flow the solar wind. In 1959, the Soviet spacecraft *Lunik 2* made the first direct measurements of the particles making up the solar wind. In the intervening years a number of interplanetary spacecraft have uncovered a broad spectrum of phenomena. *See* COMET; SPACE PROBE.

Coronal heating and solar wind acceleration. The visible surface of the Sun is known as the photosphere. Beyond this lies the chromosphere, and above this lies the corona. The corona extends many solar radii away from the Sun, eventually becoming the solar wind.

The temperature of the photosphere is a modest 10,000 K, while the solar corona is more than 10^6 K. Since the only source of energy that can heat the outer layers of the Sun comes from below, a major unsolved problem in space physics has been: How can the corona be hotter than the photosphere? Thermal conduction cannot transfer heat from something that is cold to something that is hot. Radiation can be ruled out too because the solar corona is transparent. What is known is that the continuous, agitating motion of the convective zone just below the photosphere provides the energy to heat the corona. However, how it is transported into the corona is still not well understood. It was once thought that magnetic or acoustic waves generated by the convective zone motions propagate up into the corona and dissipate their energy there. Unfortunately, it appears that acoustic waves are reflected back to the Sun at low altitudes and magnetic waves do not appear to dissipate sufficient energy into the corona.

A more promising explanation for coronal heating comes from a process known as magnetic reconnection. Coronal heating is strongest in regions of intense magnetic fields. In these so-called active regions, magnetic fields are constantly merging with one another and liberating large amounts of energy. This process manifests itself by the presence of jets, x-ray bright points, and many low-intensity flares. **Figure 1** illustrates how fountains of plasma lead to the formation of giant loops hundreds of thousands of kilometers above the surface of the Sun. The plasma on these loops can reach temperatures of several million kelvin and may be the energy source that heats the solar corona. *See* SOLAR CORONA.

Ambient solar wind. To a first approximation, two types of ambient solar wind can be distinguished: slow and fast (see **table**). The slow flow, which is associated with coronal streamers (features of the white light corona that have an archlike base

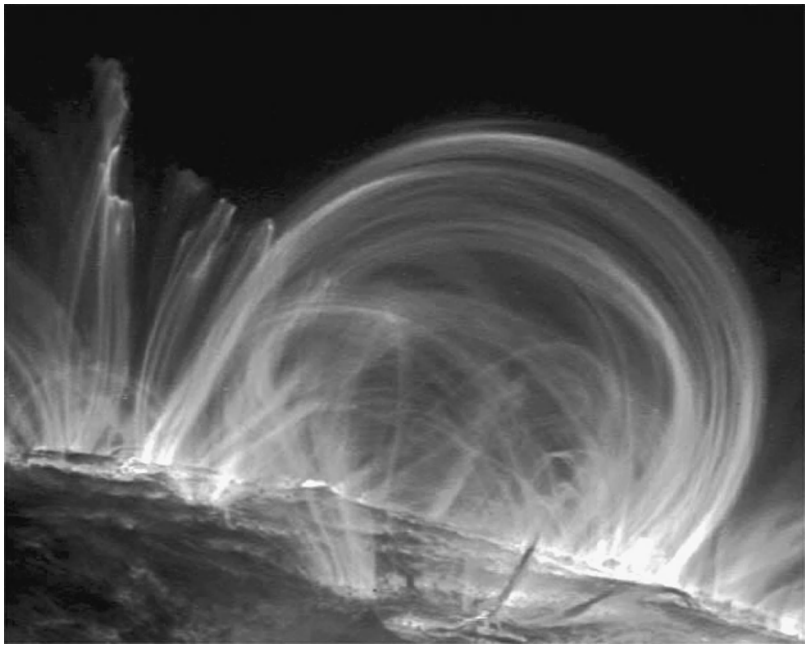

Fig. 1. Hot, dense plasma, showing the location of magnetic loops in the solar corona. This image was taken by the *Transition Region and Coronal Explorer (TRACE)* spacecraft, launched in 1998. (*Courtesy of the TRACE Consortium, Stanford-Lockheed Institute for Space Research, and NASA*)

and rays extending away from the Sun out to about one solar radius), travels away from the Sun at approximately 350 km/s (220 mi/s). It is denser and colder than the fast wind. The fast wind travels at more than twice the speed of the slow wind. It is also considerably steadier and emanates from coronal holes. The helium-to-proton abundance (the ratio of the number density of doubly ionized helium to protons) acts as a unique tracer of the origin of the solar wind in which the spacecraft is immersed. More sophisticated composition instruments on board *Ulysses* and *ACE* spacecraft now make it possible to measure the ratios of other minor species. Scientists are beginning to use these measurements to uncover the conditions and processes in the corona where the wind is accelerated as well as identifying the source regions of the solar wind in the corona.

This basic two-state flow pattern of the solar wind is illustrated in **Fig. 2**, which shows the speed measured at *Ulysses* (between 1.3 and 5.4 AU) as it completed its first orbit around the Sun. These measurements were taken during predominantly solar minimum conditions. At low latitudes, near the ecliptic, *Ulysses* observed slower, more variable solar wind. Comparison with the white-light image of the Sun shows that this flow is associated with coronal streamers. At higher latitudes, on the other hand, the flow is faster and significantly less variable. This flow is associated with the coronal holes (seen as dark regions in the ultraviolet disk image). The speed has been coded according to the magnetic polarity of the interplanetary magnetic field and shows that during this time interval the magnetic field was largely outward in the northern hemisphere and inward in the southern hemisphere, reflecting the dipole character of the solar field at this time.

Interplanetary magnetic field. The electrical conductivity of the solar wind plasma is sufficiently high that the magnetic field is effectively frozen to it. Thus, as the solar wind propagates away from the Sun, it drags the solar field with it. Because the field lines remain attached to the Sun at one end, solar rotation causes them to become wrapped into spirals. An initially radial field line at the Sun will have a pitch of approximately 45° at 1 AU and will lie almost perpendicular to the radial direction by 10 AU. Parker first provided a mathematical description of this pattern, known as the Parker spiral. *See* MAGNETOHYDRODYNAMICS; PLASMA (PHYSICS).

Spacecraft observations over several decades have confirmed Parker's basic model. However, there are some notable departures. For example, turbulence and waves produce large deviations of the magnetic field. It has also been suggested that motion of the foot points of the field lines on the surface of the Sun could lead to deviations from the Parker spiral at large distances from the Sun.

The heliospheric current sheet (HCS) is a surface that separates regions of opposite magnetic polarity. It is a fundamental feature of the heliosphere that is intimately related to the global dynamical flow of the solar wind. While it is the largest coherent structure within the heliosphere, it is also completely passive, responding to the dynamics of the solar wind flow pattern. The shape of the heliospheric current sheet also plays an important role in the modulation of galactic cosmic rays, which drift along it. *See* COSMIC RAYS.

Dynamic processes. As material flows away from the Sun, the combined effects of solar rotation together with different intrinsic speeds of the flow lead to slower material being caught by faster material, and faster material outrunning slower material. In the former case, the fast material plows into the slower material compressing and heating it. This compressed region is known as an interaction region (IR), and when the flow pattern back at the Sun does not change much with time, it is further known as a corotating interaction region. The interaction region or corotating interaction region is

Average properties of the slow and fast solar wind as observed by interplanetary spacecraft

Parameter	Slow wind	Fast wind
Proton speed	350 km/s (220 mi/s)	750 km/s (470 mi/s)
Proton density	11 cm^{-3}	3 cm^{-3}
Proton temperature	55,000 K	300,000 K
Helium to proton abundance	2.5%	4.3%

initially bound by waves that propagate away from it. As the entire structure propagates away from the Sun, these waves steepen nonlinearly, and by about 2–3 AU they have become shocks. In the case of faster material outrunning slower material, a region of rarefied plasma, known as a rarefaction region (or expansion fan), is created. *See* SHOCK WAVE.

Figure 3 summarizes many of these concepts. These are results from a global numerical simulation of the inner heliosphere during solar minimum conditions. The heliospheric current sheet is essentially flat with a single warp in it. The solar wind speed at high latitudes (over the poles of the Sun) is fast and steady. About the equator the flow is slower and more variable. The heliospheric current sheet folds back on itself by 4 AU. This is due to the acceleration of some material and the deceleration of other material. Magnetic field lines, which were initially radial near the Sun, are wound up into Parker spirals.

Time-dependent effects. Coronal mass ejections are spectacular events that involve the expulsion of massive amounts of solar material into the heliosphere. To initiate such an eruption requires a large amount of energy. It is thought that this energy is supplied by the magnetic field in the corona; however, how magnetic energy is converted into the kinetic energy needed to drive the eruption is not well understood. It is believed that a magnetic reconnection plays a central role.

Knowledge of coronal mass ejections is important for a number of reasons. First, coronal mass ejections appear to play a fundamental role in the way the corona responds to changes in the solar magnetic field associated with the 11-year solar activity cycle. The coronal mass ejections effectively provide a mechanism for the Sun to shed magnetic flux, which is thought to be essential to the cyclic renewal of the solar dynamo. Second, coronal mass ejections contribute a small, but significant amount to the overall solar wind flow at low latitudes. Third, fast coronal mass ejections have been identified as the leading cause of strong geomagnetic storms.

Figure 4 shows the evolution of a coronal mass ejection as it erupts from the Sun. Observations such as these have led to the suggestion that coronal mass ejections are often composed of helical magnetic fields, or flux ropes. Measurements of the magnetic field by interplanetary spacecraft subsequently added support to this picture.

At high heliographic latitudes, observations by the solar polar-orbiting spacecraft *Ulysses* have led to the identification of a new class of coronal mass ejections. These so-called overexpanding coronal mass ejections are probably ejected with internal pressures considerably higher than that of the surrounding solar wind, but with comparable speeds. This "overpressure" causes the coronal mass ejections to expand relative to the ambient solar wind, eventually producing an expansion shock that surrounds the coronal mass ejections and a region of low pressure within their central portions. While it is possible that these events are a distinct class from the low-latitude coronal mass ejections that have been studied for

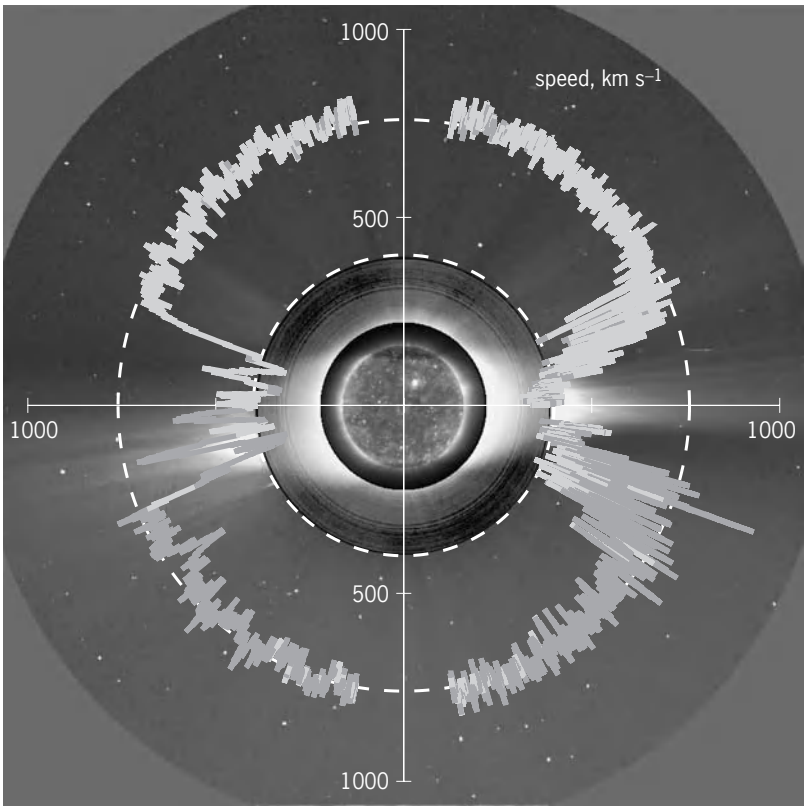

Fig. 2. Circular graph of the speed of the solar wind as measured by the *Ulysses* spacecraft, superposed on an image of the Sun and solar corona. The image is a composite made of an ultraviolet image of the solar disk from the EIT instrument on the *Solar and Heliospheric Observatory (SOHO)* and two white-light coronagraph images, one from the ground-based Mauna Loa MK3 and the other from *SOHO's* LASCO C2 instrument. The speed has been shaded to reflect the magnetic polarity of the field measured at *Ulysses*, light shading indicating outward (away from the Sun) polarity and dark indicating inward polarity.

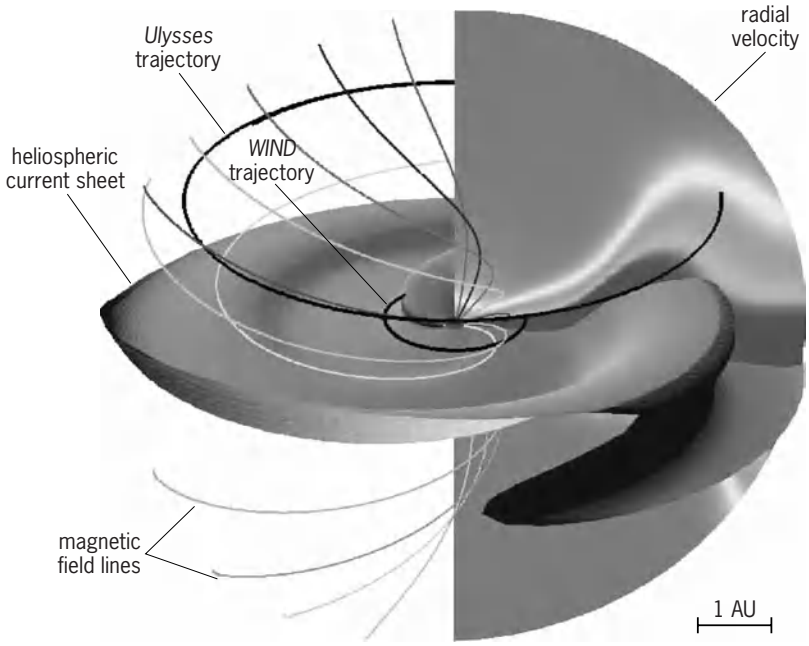

Fig. 3. Model of the heliosphere during solar activity minimum. The heliospheric current sheet is shown out to 5 AU. A slice of the radial velocity is shown at an arbitrary longitude. The darkest shading corresponds to slowest speeds (350 km/s or 220 mi/s) while the lightest shading corresponds to fastest speeds (750 km/s or 470 mi/s). Superimposed is a selection of interplanetary magnetic field lines originating from different latitudes. The trajectories of the *WIND* and *Ulysses* spacecraft are also marked.

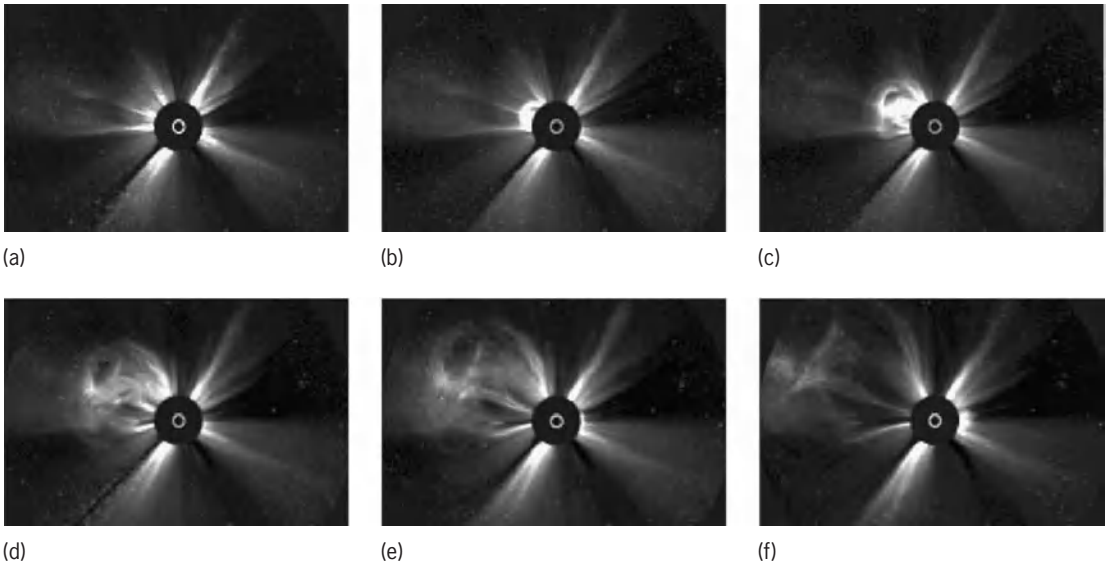

Fig. 4. Evolution of a coronal mass ejection (CME) observed over approximately 8 h on August 5–6, 1999. The white circle marks the location of the Sun, while the dark occulting disk in the center blocks the Sun. Images were taken by the LASCO C3 Coronagraph on board the *SOHO* spacecraft. (*a*) 18:18 Universal Time, August 5. (*b*) 18:42. (*c*) 21:18. (*d*) 23:18. (*e*) 00:42, August 6. (*f*) 2:42. (*Courtesy of the SOHO LASCO consortium. SOHO is a project of international collaboration between ESA and NASA.*)

decades, it is more likely that they represent a high-latitude extension of them.

Solar cycle variations. Over the course of approximately 11 years, the number of sunspots observed on the surface of the Sun rises from almost zero (solar minimum) to several hundred (solar maximum) and then falls back to zero. Other solar and solar wind parameters exhibit similar variations. For example, the pattern of coronal holes, which are the source of the fast solar wind, varies systematically with the solar cycle. At solar minimum, there are large, well-developed coronal holes over the northern and southern polar regions. By solar maximum, they have disappeared, being replaced by smaller, more transient coronal holes at essentially all heliographic latitudes. In turn, the fundamental character of the solar wind changes with the solar cycle. Whereas at solar minimum the solar wind consists of a simple two-state wind—slow, variable flow at low latitudes and fast, quiescent flow at higher latitudes—at solar maximum the flow pattern is much more complex. Fast flow can be found in equatorial regions as often as slow flow exists over the poles of the Sun.

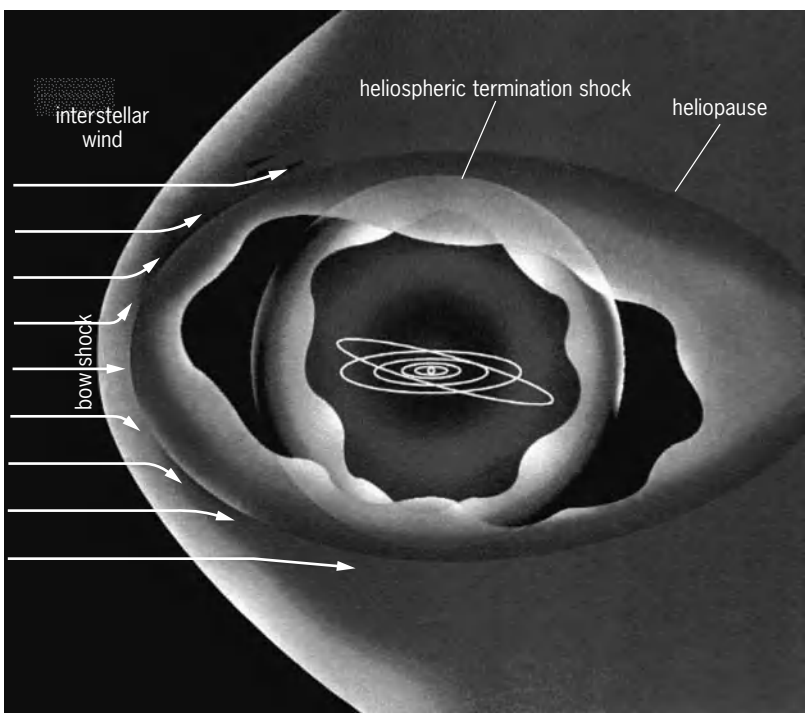

Fig. 5. Interaction of the heliosphere with the local interstellar medium. The heliopause marks the boundary between the heliosphere and the interstellar gas outside. Solar wind approaching this boundary slows suddenly, forming a shock wave. The orbits of the planets are shown by the closed curves. (*Jet Propulsion Laboratory*)

Outer heliosphere. Much of our knowledge of the heliosphere comes from interplanetary spacecraft, which measure the local properties of the surrounding plasma. As such, early knowledge of the solar wind was restricted to the environment surrounding the Earth, where spacecraft were most easily placed. As more ambitious missions were developed, explorations began in the region between the Sun and Earth (for example, *Helios* and *Pioneer Venus Orbiter*), high latitudes (*Ulysses*), and the outer heliosphere (*Pioneer 10* and *11* and *Voyager 1* and *2*).

Farther away from the Sun, the general characteristics of the solar wind change. High-frequency variations are damped out, leaving only large-scale structures. In much the same way that corotating interaction regions are formed from fast streams overtaking slower streams in the inner heliosphere, at larger distances from the Sun, more complex interactions can arise when one corotating interaction

region (or coronal mass ejection) overtakes another. The structures produced by such interactions are termed merged interaction regions (MIRs). They may be localized in latitude and longitude (such as when two coronal mass ejections interact) or global—essentially smearing into a thin spherical shell spanning 360° in longitude and up to high latitudes. Such structures play an important role in the modulation of galactic cosmic rays.

The solar system is traveling through the local interstellar medium at a speed of approximately 26 km/s (16 m/s). As the solar wind blows away from the Sun it carves out a cavity around it, called the heliosphere (**Fig. 5**). Upstream a bow shock is believed to exist, since the heliosphere is moving through the local interstellar medium supersonically. The surface where the solar wind and local interstellar medium meet is known as the heliopause. However, before reaching this point, the solar wind must be slowed down. This is accomplished by the presence of a shock wave, known as the heliospheric termination shock. Currently both of the *Voyager* spacecraft are moving away from the Sun and toward the termination shock. It is likely that at least one of them will encounter it before 2014.

Kinetic effects. In the preceding discussion, it has been implicitly assumed that the solar wind plasma behaves like a fluid. For environments like the Earth's atmosphere and oceans, this is a reasonable approximation: The collision rate between particles is very high, and the distribution of speeds within the gas closely approximates a Maxwellian distribution. In the solar wind, however, collisions are very rare. Several effects, however, conspire to give the solar wind a fluidlike behavior. The magnetic field, for example, constrains particles to orbit tightly about field lines so that they do not travel large distances between collisions. In addition, plasma instabilities tend to occur whenever there are large deviations from a Maxwellian distribution.

Kinetic effects, however, can be important. The proton temperature parallel to the magnetic field is typically 40% larger than the temperature perpendicular to it. And, while the bulk of the particles may define a Maxwellian-like distribution, secondary proton and alpha peaks in the velocity distribution imply the presence of a beam of particles streaming along the magnetic field. It has been suggested that these secondary beams are produced by reconnection in the lower corona and thus may contribute to the heating and acceleration of the solar wind. Solar wind electrons also display a secondary population, known as the halo. These halo electrons stream along field lines away from the Sun. Sometimes streams are seen in both directions, suggesting that the field lines are connected to the Sun at both ends. This usually coincides with the passage of a coronal mass ejection.

Space weather. The field of space weather is a relatively new discipline that aims to understand the conditions on the Sun, solar wind, magnetosphere, ionosphere, and thermosphere that can (1) affect the performance and reliability of space-borne and ground-based technological systems, and (2) endanger human life or health. An important component of space weather is the interaction of the solar wind with the Earth's magnetosphere. The boundary between these two regions is known as the magnetopause. Under quiet conditions the magnetopause provides a barrier to the flow of the solar wind, allowing only a fraction of the energy carried by the solar wind to enter the magnetosphere and be stored there. However, when solar activity increases, such as in a fast coronal mass ejection, the system is sufficiently perturbed that this stored energy is released suddenly, leading to a geomagnetic storm. Ultimately, it is desirable to be able to predict these events before they occur. For example, although no astronaut has yet been caught outside the shuttle during one of these storms, if such a situation were to occur, severe radiation sickness might result. *See* AEROSPACE MEDICINE; GEOMAGNETIC VARIATIONS; MAGNETOSPHERE; SPACE BIOLOGY; SUN.

Pete Riley

Bibliography. M. J. Carlowicz and R. E. Lopez, *Storms from the Sun: The Emerging Science of Space Weather*, Joseph Henry Press, 2002; K. R. Lang, *The Sun from Space*, Springer-Verlag, 2000; J. B. Zirker, *Journey from the Center of the Sun*, Princeton University Press, 2002.

Soldering

A low-temperature metallurgical joining method in which the solder (joining material) has a much lower melting point than the surfaces to be joined (substrates). Because of its lower melting point, solder can be melted and brought into contact with the substrates without melting them. During the soldering process, molten solder wets the substrate surfaces (spreads over them) and solidifies on cooling to form a solid joint.

The most important technological applications of solders are in the assembly of electronic devices, where they are used to make metallic joints between conducting wires, films, or contacts. They are also used for the routine low-temperature joining of copper plumbing fixtures and other devices. In addition, solder is used in the fusible joints of fire safety devices and other high-temperature detectors; the solder joint liquefies if the ambient temperature exceeds the solder's melting point, releasing a sprinkler head or triggering some other protective operation.

Solder alloys. To be effective, a solder must have a low melting point and form a good wetting bond to the metals it joins. A number of solder compositions are listed in the **table**, along with their melting points. Note that most of these are eutectic (lowest melting point alloy) compositions, and contain significant amounts of tin (Sn) or indium (In). Alloys of metals that are not completely miscible usually have an intermediate, eutectic composition at which the melting point has its minimum value and the molten

Compositions and melting points of simple solders

Alloy composition, wt %	Melting temperature, °C
51In-49Sn (eutectic)	120
57Bi-43 Sn (eutectic)	139
97In-3 Ag (eutectic)	141
62Sn-38 Pb (eutectic)	183
96.5Sn-3.5 Ag (eutectic)	221
95Sn-5 Sb	240
80Au-20 Sn (eutectic)	280
97Pb-3 Sn	320
87.5Au-12.5 Ge (eutectic)	361

alloy solidifies directly into a mixture of two solid solutions. Tin or indium content is included in solder to facilitate bonding to the metals that are most commonly soldered, such as copper (Cu), nickel (Ni), and gold (Au). Tin and indium form stable intermetallic compounds with copper and nickel, and indium also forms intermetallics with gold. The intermetallic reaction at the solder-substrate interface creates a strong, stable bond. *See* ALLOY; EUTECTICS; INTERMETALLIC COMPOUNDS.

Eutectic tin-lead (Sn-Pb) has been, by far, the most widely used solder, with its melting point conveniently located at 183°C (361°F). Its high tin content is suitable for rapid, efficient bonding to copper, nickel, and gold. In addition, tin-lead solders are easy to manufacture, widely available, and low in cost. This situation is, however, changing rapidly because of legislation intended to minimize the release of lead into the environment. Lead-rich solders are already prohibited in most plumbing fixtures and may be restricted in electronics. *See* LEAD ALLOYS; TIN ALLOYS.

The solders in the table are only a few of the many compositions that are commercially available. Different solders are used because of the need for specific melting-point ranges, for hierarchies of melting points, and for environmentally compatible compositions.

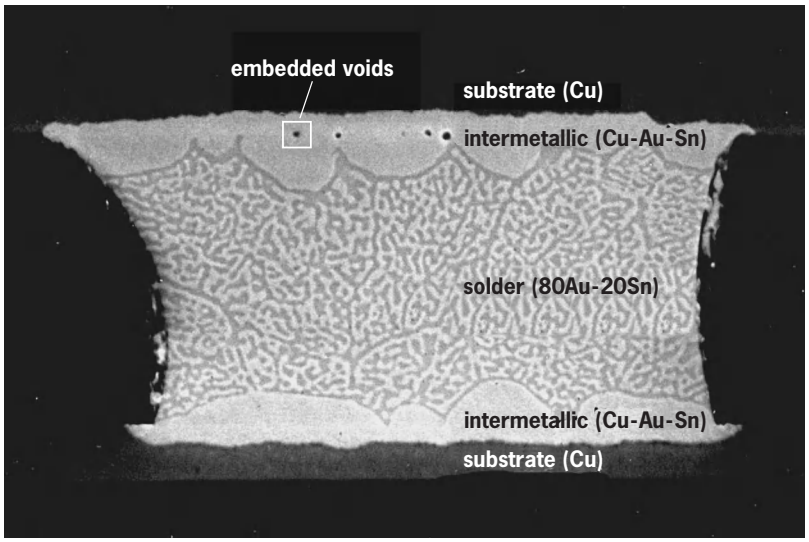

Polished and etched cross section of a eutectic Au-Sn solder joint between copper pads. The distinct phases are shown, as well as the fine voids present at the interface. (*Mark McCormack, Fujitsu Computer Packaging Systems*)

The melting point of the solder must be chosen to satisfy both manufacturing and service issues. For example, low-melting solders are needed for many microelectronic devices, particularly optoelectronic devices containing semiconductors that can be damaged by exposure to even moderate temperature. High-melting solders are required for "underhood" applications in automotive electronics, since the electronic components installed near the engine are exposed to relatively high temperatures.

A hierarchy of solder melting points is required in manufacturing operations where soldered devices are assembled in stages. In manufacturing a microelectronic device, for example, chips may be soldered to chip carriers, which are then soldered to circuit boards. It is desirable that subsequent stages of the operation use a solder with a sufficiently low melting point that the joints fabricated in previous operations are not melted or damaged. To accomplish this, chips may be mounted with 97Pb-3Sn and chip carriers mounted with 62Sn-38Pb. *See* ELECTRONIC PACKAGING; ELECTRONICS; INTEGRATED CIRCUITS; PRINTED CIRCUIT.

Environmental compatibility includes both chemical integrity in the device and environmental tolerance when the device is ultimately discarded. Solders may be metallurgically incompatible with one another [for example, a small amount of lead can cause a serious deterioration in the mechanical properties of eutectic bismuth-tin (Bi-Sn)]; they can exacerbate corrosion problems within the device; or they can be environmentally harmful on disposal, as are tin-lead alloys. *See* CORROSION; METAL, MECHANICAL PROPERTIES OF.

Circuit assembly. The manufacture of a sound solder joint requires a clean substrate to facilitate wetting and intermetallic formation. In particular, the oxide coatings that form spontaneously on copper and nickel must be removed or prevented. The oxide coatings are ordinarily removed by washing with appropriate cleaning agents called fluxes. The flux may be applied to the surface prior to soldering. Alternatively, the flux may be incorporated into the solder. This is usually done by making a solder paste, a mixture of fine solder particles with an appropriate flux. Since the most efficient fluxes are acidic, and pose environmental and safety problems, substrates are often cleaned during their original manufacture and coated for protection. The most common coatings are the solder iteself (in which case the substrate is said to be pre-tinned) and gold, which does not oxidize. Concepts such as flux-free soldering are under development, and include methods such as sputter cleaning of the substrate prior to soldering.

In typical electronic manufacturing, a solder joint is fabricated in two steps. The substrates to be joined are first coated with solder. They are then fitted together and heated to remelt or "reflow" the solder and make the joint. Solder paste is selectively applied to the substrate by stencil printing. A stencil is a mask with open areas that match the substrates that are to be coated, and the solder paste is pressed through the stencil openings using a squeegee. When molten

solder is deposited, a common method is wave soldering, in which the part or board containing the substrates is passed through a continuous stream of molten solder (solder wave). Because of its wetting characteristics, the molten solder adheres only to the substrates.

Solder joint metallurgy. In the **illustration**, a finished solder joint comprises copper substrates joined with eutectic Au-Sn solder. The solder is bonded to the substrates through an intermediate layer of intermetallic which, in this case, is a Cu-Au-Sn ternary compound. The thickness of the intermetallic layer is largely determined by the time of contact with the molten solder, and, hence, increases with the number of reflow operations that must be performed. In extreme cases, the intermetallic may bridge the joint. Since the intermetallic phase is ordinarily brittle, a thick intermetallic layer is undesirable.

The microstructure of the bulk solder in the as-solidified joint depends on its composition and on the rate at which it is solidified. The solidification rate is largely determined by the joint size; the smaller the joint, the higher its surface-to-volume ratio and the more rapidly it cools. Increasing the cooling rate refines the effective grain size of the solder and, generally, improves its mechanical properties. The bulk solder in the illustration has a classic eutectic microstructure, with alternating lamellae of gold- and tin-rich phases. At very high cooling rates the eutectic microstructure tends to be replaced by a fine mixture of small, equiaxed (having equal dimensional in all directions) grains of the two phases. Solder joints in microelectronic devices experience temperatures near their melting points during operation. These high homologous temperatures cause the microstructure to coarsen gradually as the device is used. *See* CRYSTAL STRUCTURE; GRAIN BOUNDARIES; METALLOGRAPHY; METALLURGY.

Another significant feature of the joint cross section shown in the illustration is the array of small voids along the interface. Such voids are common and form during solidification due to entrapped flux or outgassing from the substrate.

Reliability. The mechanical integrity of solder joints in microelectronic devices is a particular concern, since a joint fracture can break electrical contact, causing failure of the device. There are two principal sources of mechanical failure: overload fracture during handling, and thermal fatigue fracture in service. Overload fractures tend to happen at the interface and involve the brittle intermetallic layer. The strength deteriorates as the intermetallic layer thickens, and also decreases with the areal fraction of interfacial voids.

Thermal fatigue failures are caused by cyclic stresses that result from differential thermal expansion and differential heating, mainly when the device is turned on and off. These failures tend to occur in shear, by gradual crack propagation through the body of the solder. Thermal fatigue in solder is complicated by the high homologous temperature at which the joint operates. The deformation mode is dominated by high-temperature creep, whose rate depends strongly on both the temperature and the local microstructure. Moreover, the microstructure is not fixed, but evolves in a pattern that can be strongly affected by the deformation as well as the temperature. Given the difficulty of predicting thermal fatigue life in service, solder joints in critical locations tend to be designed with large margins of safety. *See* CREEP (MATERIALS); SHEAR. J. W. Morris, Jr.

Bibliography. D. R. Frear et al. (eds.), *The Mechanics of Solder Alloy Interconnects*, Van Nostrand Reinhold, New York, 1994; J. S. Hwang, Modern solder technology for competitive electronics manufacturing, McGraw-Hill, New York, 1996; J. H. Lau (ed.), *Solder Joint Reliability: Theory and Applications*, Van Nostrand Reinhold, New York, 1991; R. R. Tummala, E. J. Rymaszewski, and A. G. Klopfenstein (eds.), *Microelectronics Packaging Handbook*, 2d ed., Chapman & Hall, New York, 1997.

Solenodon

A rare insectivorous mammal classified in the family Solenodontidae. Solenodons are among the largest living insectivores. There are only two extant species. The Hispaniolan, or Haitian, solenodon (*Solenodon paradoxus*) is restricted to the remote, wet, densely vegetated central highland regions on the island of Hispaniola (Dominican Republic and Haiti). The coarse pelage (coat) varies from blackish to reddish brown. Most individuals possess a small, square, whitish area on the nape of the neck. Adults are 28–33 cm (11–13 in.) in body length, have a 17.5–26-cm-long (7–10.4-in.) tail, and weigh 600–1000 g (1.2–2.2 lb). The smaller Cuban solenodon, or almique (*S. cubanus*), has longer and finer fur which is blackish brown with white or buff. It is now restricted to Oriente Province in Cuba. Adults are 20–30 cm (10–12 in.) in body length, have a 26–30 cm-long (10.5–12 in.) tail, and weigh 600–700 g (1.3–1.5 lb). Two additional species are known only from skeletal remains in Haiti and Cuba.

Morphology. Solenodons resemble shrews in general appearance except that they are much larger and more stoutly built (see **illustration**). They have relatively large heads, tiny eyes, and large, partially naked ears. The head tapers to a long, flexible, and mobile proboscis that extends well beyond the lower jaw. A unique ball-and-socket joint attaches the proboscis to the skull in the Hispaniolan solenodon and provides even greater mobility. The nostrils are located at the sides of the tip of the proboscis. The senses of touch,

Solenodon, of the family Solenodontidae.

smell, and hearing are highly developed. The front legs are longer than the hind legs. All limbs possess five toes, with the toes on the forelimbs possessing long, stout, sharp claws. The tail is thick, scaly, and almost hairless. Scent glands in the armpit and the groin produce a fatty, musky, greenish secretion that has a goatlike odor. The mammary glands are located in the inguinal region. The dental formula is I 3/3, C 1/1, PM 3/3, M 3/3 × 2, for a total of 40 teeth. The second lower incisor has a goove on the lingual side (solenodon means "animal with grooved teeth"). The submaxillary salivary glands produce a toxic saliva. The grooves on the second lower incisor act as channels for carrying the venom.

Ecology. Solenodons are nocturnal and inhabit forests and brushy edges. Although able to climb, solenodons spend most of their time on the ground. They forage and nest in extensive burrow systems that they excavate in deep layers of humus. Although adult Cuban solenodons are considered to be solitary except when a mother is accompanied by her young, adult Hispaniolan solenodons may live in family groups in caves and natural hollows, with up to eight having been found in the same shelter. Vocalizations consist of chirps, squeaks, clicks, snorts, and grunts. A high-frequency click may function in echolocation. Food consists of a variety of invertebrates, including millipedes, ground beetles and other insects, insect larvae, earthworms, and snails, as well as amphibians and small reptiles and birds. The front claws are used to expose prey beneath rocks, bark, and in the soil.

The frequency and timing of reproduction in the wild in unknown. Mating is believed to be irregular and independent of the seasons. Normally a single young is born. At birth, young solenodons are naked except for long vibrissae on the snout and a small amount of fine, transparent body hair. They weigh about 40–55 g (1.4–1.9 oz). The first solid food is consumed at about 13 weeks of age. A young solenodon will remain with its mother for several months. Longevity in the wild is unknown. In captivity, a Hispaniolan solenodon has lived for more than 11 years, while a Cuban solenodon survived for 6.5 years.

Status. Both species are rare and are listed as highly endangered in the Red Data Book of the International Union for Conservation of Nature and Natural Resources (IUCN) and by the U. S. Department of the Interior. Their decline is due to their low reproductive rate, habitat destruction (deforestation), increasing human activity, and predation by introduced carnivores (dogs, feral cats, and mongooses). Mongooses were introduced to Hispaniola in 1872 for snake control, but they have also killed many other forms of wildlife, including solenodons. Prior to human settlement and the introduction of other carnivores, solenodons were the only natural terrestrial predators on the two islands on which they live. *See* INSECTIVORA. Donald W. Linzey

Bibliography. D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Solenoid (electricity)

An electrically energized coil of insulated wire which produces a magnetic field within the coil. If the magnetic field produced by the coil is used to magnetize and thus attract a plunger or armature to a position within the coil, the device may be considered to be a special form of electromagnet and in this sense the words solenoid and electromagnet are synonymous. In a wider scientific sense the solenoid may be used to produce a uniform magnetic field for various investigations. So long as the length of the coil is much greater than its diameter (20 or more times), the magnetic field at the center of the coil is sensibly uniform, and the field intensity is almost exactly that given by the equation for a solenoid of infinite length.

When used as an electromagnet of the plunger type, the solenoid usually has an iron or steel casing. The casing increases the mechanical force on the plunger and also serves to constrain the magnetic field. The addition of a butt or stop at one end of the solenoid greatly increases the force on the plunger when the distance between the plunger and the stop is small. The **illustration** shows a steel-clad solenoid

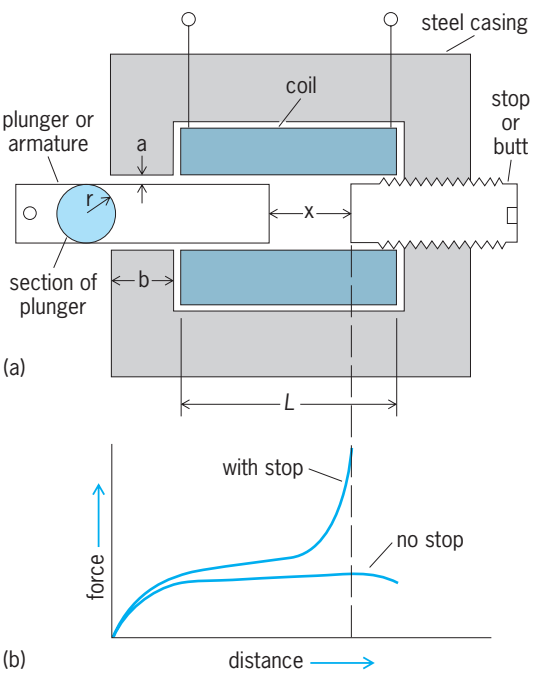

Steel-clad solenoid. (*a*) Cross-sectional view. (*b*) Relation of the force acting on the armature to the displacement of the armature.

with plunger and plunger stop. The relation of force versus distance with and without the stop is also shown.

The force for the solenoid rapidly increases as the plunger enters the coil because of the rapid rate of change of the reluctance for the magnetic path. For the solenoid shown in the illustration, the force on the plunger is given approximately by the equation below, where I is the coil current (in amperes), N is

$$F = ANI\left(\frac{\mu_0 NI}{2(c+x)^2} + \frac{1}{CL}\right) \text{ (newtons)}$$

the number of turns, $A = \pi r^2$ is the cross-sectional area of the plunger (in square meters; 1 ft^2 = 0.0929 m^2), μ_0 is the permeability of free space ($4\pi \times 10^{-7}$ henry/m), r, a, b, x, and L are as shown in the illustration (in meters; 1 ft = 0. 3048 m), and $c = ra/2b$. The value of the constant C depends upon the dimensions of the solenoid and the degree of plunger magnetization. For normal designs, C has a value in the range 0.6 to 2.9. For a conservative design, one may use the largest value for C or neglect the second term entirely. This term is due to leakage fluxes and becomes negligible as the plunger closes the air gap x.

In principle a solenoid works with either ac or dc excitation. In the dc solenoid the flux is always at its maximum value for a fixed plunger position. In an ac solenoid the force varies at twice the frequency of the supply voltage. The variation, which is caused by the variation of the magnetic flux, gives rise to excessive chattering or vibration unless a shading coil is embedded in the face of the plunger stop. The shading coil acts to smooth the variation of flux and thereby smoothes the attractive force.

When direct current is used, only the resistance of the coil wire limits the final value of current, while with ac excitation the inductance of the coil must also be considered. It is difficult to calculate the proper value of inductance, because it is a function of the position of the plunger. As the plunger moves into the solenoid, the inductance, or flux per ampere, becomes much greater because the reluctance of the flux path is much less. The current drawn by the coil for constant applied ac voltage becomes smaller as the plunger moves into the solenoid. It is common practice to laminate the plunger, stop, and casing in ac service to reduce eddy current losses. *See* ELECTROMAGNET. Jerome Meisel

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; G. R. Jones, M. A. Laughton, and M. G. Say (eds.), *Electrical Engineer's Reference Book*, 15th ed., 1993; D. B. Montgomery, *Solenoid Magnet Design*, 1969, reprint 1980; Y. M. Pulyer, *Electromagnetic Devices for Motion Control and Signal Processing*, 1992.

Solenoid (meteorology)

In meteorological usage, solenoids are hypothetical tubes formed in space by the intersection of a set of surfaces of constant pressure (isobaric surfaces) and a set of surfaces of constant specific volume of air (isosteric surfaces) or density (isopycnic surfaces). The isobaric and isosteric surfaces are such that the values of pressure and specific volume, respectively, change by one unit from one surface to the next. The state of the atmosphere is said to be barotropic when there are no solenoids, that is, when isobaric and isosteric surfaces coincide. The number of solenoids cut by any plane surface element of unit area is a measure of the torque exerted by the pressure gradient force, tending to accelerate the circulation of air around the boundary of the area. *See* BAROCLINIC FIELD; BAROTROPIC FIELD; ISOPYCNIC.

Frederick Sanders; Howard B. Bluestein

Bibliography. S. L. Hess, *Introduction to Theoretical Meteorology*, 1959, reprint 1979; H. G. Houghton, *Physical Meteorology*, 1985.

Solenopora

A genus of extinct calcareous red algae (Rhodophyta). *Solenopora* is characterized by encrusting, rounded, nodular growth forms ranging in size from several millimeters to a few centimeters. Internally, this alga consists of calcified rows or filaments of cells, commonly polygonal in shape, ranging from 20 to 100 micrometers in diameter. Although the cellular tissue is similar to modern coralline algae, the cells in *Solenopora* are generally larger and are not differentiated into more than one kind of cellular arrangement.

Solenopora and related fossil genera have been referred to invertebrate groups, but now there is general agreement that these fossils constitute a coherent group of extinct red algae separate from the Corallinaceae. Solenoporids probably provided an ancestral stock to the true coralline algae.

Solenopora appeared first in Cambrian time and became extinct in the early part of the Cenozoic. The genus flourished during the Ordovician, Silurian, and Jurassic periods. The environmental distribution of *Solenopora* was comparable to some modern coralline algae. *See* ALGAE. John L. Wray

Bibliography. J. L. Wray, *Calcareous Algae*, 1977.

Solid (geometry)

A three-dimensional geometric figure consisting of points continuously connected, and separated from the rest of space by a surface called the boundary of the solid. Points on the boundary are usually considered part of the solid. A solid is bounded if there exists a sphere having finite radius that could enclose the solid. A solid is convex if all points of any line segment having end points on the boundary also are points of the solid.

Volume. A measure of a solid is its volume: the amount of space enclosed by the boundary surface. A unit cube is a cube having an edge that is one unit long, where the unit normally is a standard unit of length. The volume of a unit cube is defined to be 1 cubic unit, or 1 $unit^3$. The volume of a solid is a number mulitiplied by the volume of a unit cube, with the number telling how many of these unit cubes would be required to occupy the same amount of space as does the solid. The volume of a bounded solid is finite; an unbounded solid might have finite or infinite volume. *See* UNITS OF MEASUREMENT.

Many solids have volumes that can be determined from formulas that involve one or more dimensions associated with the solid. A dimension is any measure associated with a geometric figure that has units of

length, such as height, width, or perimeter. Angles are not dimensions (they do not have units of length) but appear in some formulas for volumes.

Cavalieri's principle. The volumes of a large number of solids can be found with Cavalieri's principle: If a solid can be oriented so that all horizontal plane sections have the same cross-sectional area A, and if h is the distance between the lowest and highest plane sections, then the volume V is given by Eq. (1).

$$V = Ah \tag{1}$$

Prismoidal formula. Some solids have volumes given by the prismoidal formula: If the solid can be oriented so that the area A of any horizontal plane section satisfies $A = f(d)$, where d is the distance of that section from the lowest point of the solid, and f is a polynomial function of degree three or less, then the volume V is given by Eq. (2), where L, U,

$$V = \tfrac{1}{6}h(L + 4M + U) \tag{2}$$

and M are, respectively, the areas of the lowest plane section, the highest plane section, and a plane section midway between the lowest and highest (one or more of these areas might be zero); and h is the height of the solid (the distance between the highest and lowest plane sections). Prismatoids, spheres, ellipsoids, pyramids, frustrums, cones, cylinders, and many other solids satisfy this property. However, since the area of the midsection is usually not easily determined, the prismoidal formula is seldom a practical method for determining a volume.

Integral calculus techniques. Integral calculus supplies general techniques for determining the volumes of solids whose boundaries can be described with equations. These techniques can be classified as one-, two-, or three-variable. *See* CALCULUS; INTEGRATION.

1. One-variable: If the solid can be oriented so that cross sections perpendicular to an axis (say, the x axis) have areas $A(x)$ [the area is a function of x], then the volume V is given by Eq. (3), where a and b define the limits of the x axis for which this functional relationship holds. Some solids might have volumes that can be found by partitioning the solid, applying the technique to each part, and summing the volumes thus obtained.

$$V = \int_a^b A(x)\,dx \tag{3}$$

A solid of revolution is the part of space that is said to be swept out when a region of a plane is revolved about a line in that plane that does not contain an interior point of the region. A solid of revolution also can be defined as a solid whose boundary consists only of surfaces of revolution having a common axis. The volume of a solid of revolution might be found by the so-called cylindrical shell method, in which the solid is positioned so that the axis of revolution is the y axis. If the region that is being revolved can be defined by $y = U(x)$ and $y = L(x)$ [the upper and lower curves that are boundaries of the region], and $x = a$ and $x = b$ (the left and right straight-line boundaries of the region), then the volume V is given by Eq. (4).

$$V = \int_a^b 2\pi x[U(x) - L(x)]\,dx \tag{4}$$

See SURFACE (GEOMETRY).

2. Two-variable: If the solid can be oriented so that the up direction is the positive z axis, the upper surface of the solid is defined by $z = U(x,y)$, the lower surface is defined by $z = L(x,y)$, and the remaining boundary is a cylindrical surface whose generator is parallel to the z axis, then the volume V is given by Eq. (5). Here A represents the region of the xy plane

$$V = \iint_A U(x,y) - L(x,y)\,dA \tag{5}$$

enclosed by the cylindrical surface—the region over which the double-integral is to be evaluated.

3. Three-variable: If $\mathscr{V}\mathscr{V}$ represents the part of space occupied by a solid, then the volume V is given by Eq. (6).

$$V = \iiint_{VV} dV \tag{6}$$

The evaluations of the double and triple integrals above are normally attempted by replacing them with iterated integrals. The proper choice of coordinate system (rectangular, cylindrical, or spherical) and the order of integration can have an effect on the ease (or even the possibility) of obtaining the volume. If an expression for the volume cannot be obtained analytically, then numerical techniques can be used to obtain an approximate volume.

Volume theorem of Pappus. Some solids of revolution have volumes that are best found by using the volume theorem of Pappus: If a plane region is revolved about a line in that plane that does not contain an interior point of the region, then the volume of the solid of revolution generated is equal to the area of the region multiplied by the distance that the centroid of the region moves during one revolution. (If a thin material, such as sheet metal, were cut in the shape of the region, then the centroid of the region is the point where the region would balance.) This theorem is illustrated by a solid torus, the solid of revolution generated when a circular region is revolved about a line that does not intersect the circle. Clearly, the centroid of this region is the center of the circle. If R is the distance between this center and the axis of revolution, and r is the radius of the circle, then the area of the circle is πr^2 and the distance traveled by the centroid during one revolution is $2\pi R$. The volume V of a solid torus, therefore, is given by Eq. (7).

$$V = 2\pi^2 Rr^2 \tag{7}$$

See TORUS.

Important solids. A number of solids are frequently encountered in pure and applied mathematics, in addition to those already discussed.

Solids often receive the name of their boundary surface. In formal mathematical usage, "sphere,"

"cylinder," and "cone" are surfaces, but often these words refer to the solids having such surfaces as part or all of their boundary. Usually the distinction can be made from the context. The use of the word "solid" can remove any ambiguity.

Polyhedron. A polyhedron is any of a large class of bounded solids whose boundary surface consists of plane polygonal regions. Examples include cubes, pyramids, prisms, and many other familiar solids. *See* CUBE; POLYHEDRON.

Sphere. A solid sphere has a boundary that is a sphere. The size of a sphere is completely determined by either its radius or its diameter, which refer to the length of a radius (a line segment that joins the center with a point on the sphere) or the length of a diameter (a line segment containing the center whose end points lie on the sphere). The volume V of a solid sphere having radius R is given by Eq. (8).

$$V = {}^{4}\!/_{3}\pi R^3 \tag{8}$$

Two parallel planes that intersect a sphere, together with the part of the spherical surface between those planes, define a solid: a spherical segment (shaped like a tomato slice). The intersection of each plane with the sphere is a circle (a tangent plane intersects at only one point, which can be considered a degenerate circle with radius 0), and if the radii of these two circles are R_1 and R_2, and h is the distance between the planes, then the volume V of the spherical segment is given by Eq. (9).

$$V = {}^{1}\!/_{6}\pi h(3R_1^2 + 3R_2^2 + h^2) \tag{9}$$

A plane that contains the center of a sphere divides the sphere (both the surface sphere and the solid sphere) into two hemispheres (yet another word that can refer to either a surface or a solid). Another plane that contains the center will cut a solid hemisphere into two lunar segments (the smaller will be shaped like an orange segment). If θ is the radian measure of the dihedral angle formed by the planes that define the lunar segment, and R is the radius of the sphere, then the volume V of the lunar segment is given by Eq. (10).

$$V = {}^{2}\!/_{3}\theta R^3 \tag{10}$$

See SPHERE.

Cylinder. A cylindrical solid is formed when a cylindrical surface is intersected by two parallel planes (**Fig. 1**). The cylindrical surface between the planes is the lateral surface of the solid; the remaining two congruent regions are the bases of the solid. If each base has finite area A, and the distance between the planes (the height of the cylindrical solid) is h, then the solid's volume V will be given by Cavalieri's principle, Eq. (1). A cylindrical solid often is named by giving information about the shape of a base, and how the plane intersects a generator of the cylinder. The word "right" indicates that the plane is perpendicular to the generator. A prism is a special case of a cylindrical solid, in which the lateral surface consists of plane regions. *See* PRISM.

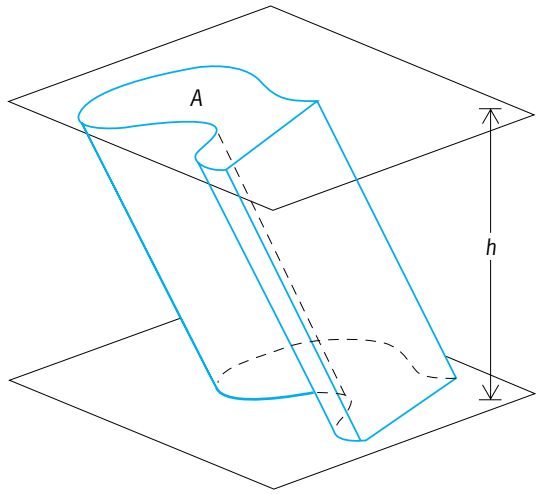

Fig. 1. Cylindrical solid. The height of the solid is *h*, and the area of the bases is *A*.

When referring to a solid, the word "cylinder," unaccompanied by any modifying words, invariably means a right circular cylinder. The axis of revolution is the axis of the cylinder. The bases of the cylinder are the circular regions defined by the intersection of the planes and the cylindrical surface. The radius of the cylinder is the radius of a base; the height (or altitude) of the cylinder is the distance between the bases. If r and h denote the radius and height, then the volume V of the cylinder is given by Eq. (11).

$$V = \pi r^2 h \tag{11}$$

See CYLINDER.

Cone. A conical solid is bounded by one nappe of a cone and a plane that intersects that nappe (**Fig. 2**).

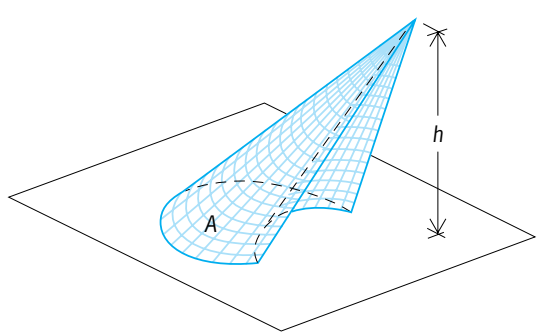

Fig. 2. Conical solid. The height of the solid is *h*, and the area of the base is *A*.

The plane region defined by this intersection is the base of the conical solid; the part of the nappe between the plane and the vertex is the lateral surface of the conical solid. The vertex of the solid is the vertex of the cone; the distance between the vertex and the plane is the height of the conical solid. If the base has area A and the height is h, then the volume

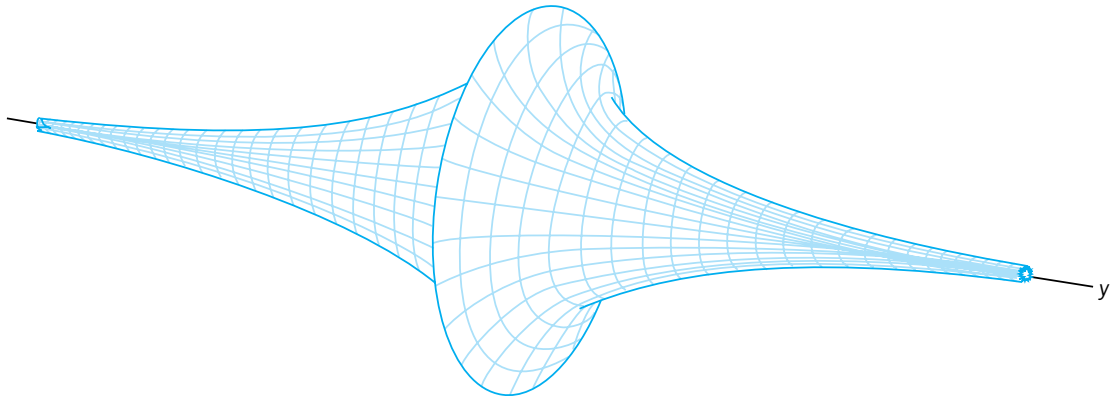

Fig. 3. Pseudosphere formed by revolving a tractrix about the *y* axis.

V of the conical solid is given by Eq. (12). A pyramid

$$V = {}^{1}/_{3}Ah \tag{12}$$

is a special case of a conical solid, in which the lateral surface consists of plane regions.

When used as the name of a solid, the single word "cone" invariably means a right circular cone: a solid whose boundary consists of one nappe of a circular cone and a plane perpendicular to the axis of revolution, which also is the axis of the cone. If r and h denote the radius and height, then the volume V of the cone is given by Eq. (13).

$$V = {}^{1}/_{3}\pi r^2 h \tag{13}$$

See CONE.

Frustrum. A frustrum (or frustum) of a conical solid is the solid remaining if a plane parallel to the base intersects the solid, and the part between that plane and the vertex is discarded. The intersecting plane defines a second (smaller) base of the frustrum. If B and b are the areas of the bases of the frustrum, and H is the height of the frustrum (the distance between the bases), then the volume V of a frustrum of a conical solid is given by Eq. (14).

$$V = {}^{1}/_{3}H(B + \sqrt{Bb} + b) \tag{14}$$

Ellipsoid. A solid ellipsoid is the solid whose boundary is an ellipsoid. A diameter of an ellipsoid is a line segment containing the center that has end points on the surface. The shortest and longest of all the diameters of the ellipsoid will be perpendicular, and a third diameter will be perpendicular to both. If these three diameters have lengths $2a$, $2b$, and $2c$, then the volume V of the solid ellipsoid is given by Eq. (15). If two of the three diameters are of equal

$$V = {}^{4}/_{3}\pi abc \tag{15}$$

length, then a solid spheroid (a solid ellipsoid of revolution) is generated. The spheroid will be oblate or prolate, according to whether the third diameter is respectively shorter than or longer than the diameters having equal length.

Pseudosphere. A solid pseudosphere is an unbounded solid of revolution formed when the region between a certain tractrix and a certain line is revolved about that line. If an object is tied to a cord of constant length, and the other end of the cord follows some curve, then a tractrix is the path followed by the object as it is dragged slowly through a resisting medium. A pseudosphere can be formed by revolving this tractrix about the y axis (**Fig. 3**): The end of the cord moves from the origin along the positive y axis, the initial position of the dragged object is the point $(a,0)$, and the length of the cord is a. (The entire tractrix consists of this first-quadrant curve and its reflection about the x axis.) A solid pseudosphere is formed by revolving the region lying between this tractrix and the y axis. The solid thus formed is not bound (since the tractrix is asymptotic to the y axis) but has a finite volume V given by Eq. (16).

$$V = {}^{2}/_{3}\pi a^3 \tag{16}$$

See TRACTRIX.

Harry L. Baldwin, Jr.

Bibliography. A. A. Albert, *Solid Analytic Geometry*, 1966; J. S. Frame, *Solid Geometry*, 1948; D. Hilbert and S. Cohn-Vossen, *Geometry and the Imagination*, 1932, reprint 1990; P. Pearce and S. Pearce, *Polyhedra Primer*, 1978.

Solid solution

Compositional variation of a crystalline substance due to substitution or omission of various atomic constituents within a crystal structure. Solid solutions can be classified as substitutional, interstitial, or omissional. They may also be categorized by the nature of their thermodynamic properties, such as enthalpy, entropy, and free energy (for example ideal, nonideal, and regular solid solutions).

Characteristics. The concept of solid solution can be understood by considering a specific mineral group such as olivine. Although olivine-group minerals may exhibit a range of compositions they all have a similar crystal structure consisting of isolated silicate $(SiO_4)^{4-}$ tetrahedra cross-linked by cations

in two specific types of sites, both of which are octahedrally (eightfold) coordinated to oxygen atoms. Thus all of the minerals in the olivine group are isostructural (having a similar crystal structure). Yet within this structural framework, compositions vary considerably. Such chemical variation can be described by defining the nature and extent of the atomic substitutions involved or by describing intermediate compositions in terms of limiting end-member compositions. For example, the composition of the mineral olivine ranges between the end members forsterite (magnesium silicate; Mg_2SiO_4) and fayalite (ferrous silicate; Fe_2SiO_4), representing varying amounts of iron (Fe^{2+}) and magnesium (Mg^{2+}) ion substitution. A complete solid solution series involving iron/magnesium substitution also exists between the olivine group minerals monticellite ($CaMgSiO_4$) and kirschteinite ($CaFeSiO_4$). There is, however, little solid solution between the monticellite series and the olivine series because the larger calcium ion (Ca^{2+}) does not readily substitute for either the magnesium ion or the ferrous iron ion. Thus in some instances solid solution between end members may be complete, whereas in other cases it may be limited. In some minerals, atomic substitution must be coupled in order to maintain electrical charge balance. In plagioclase feldspar, defined by the end-member compositions albite ($NaAlSi_3O_8$) and anorthite ($CaAl_2Si_2O_8$), compositional substitution of sodium (Na^{1+}) ion for calcium ion is coupled with the substitution of the silicon (Si^{4+}) ion for the aluminum (Al^{3+}) ion. *See* FELDSPAR; OLIVINE.

In addition to substitution within the crystal structure, compositional variation may take place by interstitial substitution or omission solid solution. Interstitial solid solution occurs when ions or atoms occupy a position in a crystal structure that is usually vacant. For example, this is relatively common in the zeolite mineral group and in beryl. When charged ions are thus incorporated in the crystal, additional coupled substitution elsewhere in the structure is necessary in order to maintain charge balance. Omission solid solution occurs when an atom or ion is missing from a specific crystallographic position. For example, the composition of the mineral pyrrhotite deviates from pure ferrous sulfide (FeS) by the omission of ferrous ion from the structure. In order to maintain charge balance, as ferrous ion is omitted, some of the iron present in pyrrhotite most likely exists as ferric (Fe^{3+}) ion. *See* BERYL; ZEOLITE.

Exsolution. Exsolution refers to the separation of a mineral solid solution that was initially stable as a single homogeneous substance into two or more substances with different compositions. This process generally occurs in response to decreasing temperature, with the aggregate composition of the new mixture remaining the same as the initial, single, homogeneous solid solution. Because this separation takes place in a solid substance, nucleation of the newly formed material commonly occurs on specific crystallographic planes or directions in the original mineral, producing exsolution lamellae.

Occurrence. Solid solution is widespread among minerals. In fact, very few naturally occurring minerals exist as pure end-member substances but exhibit trace to extensive solid solution. The extent of solid solution depends upon the relative sizes of the atoms or ions involved, the charges of the ions, the coexistence and composition of other minerals or liquid (for example, magma), and the temperature and pressure conditions of formation (with temperature having a more pronounced effect). Generally, solid solution among ions that differ in size by less than 15% is significant, whereas solid solution among ions that differ in size by more than 30% is very limited. If ions differ in charge, a coupled solid solution is required to maintain charge balance. In addition, the extent of solid solution reflects the abundance of various elements at the time of mineral formation. Although a complete solid-solution series between calcium carbonate ($CaCO_3$) and cadmium carbonate ($CdCO_3$) has been experimentally verified, it has not been observed in naturally occurring specimens because of generally low Cd^{2+} concentrations.

Temperature affects solid-solution limits in two ways. In general, increased temperature expands the limits of solid solution, because increased thermal vibration within a crystal structure reduces the geometrical size constraints on a specific structural site. Second, temperature (and pressure) variation affects the chemical equilibria among minerals and between minerals and fluids (such as magma). This affects the distribution of a given element among coexisting substances (phases) which, in turn, controls the composition of individual minerals. Because solid-solution limits are sensitive to temperature and pressure, mineral compositions and exsolution textures have been widely used to determine conditions of formation and subsequent cooling histories of many igneous rocks.

Compositional variation among metamorphic minerals (such as garnets and amphiboles) and exsolution phenomena have been used to determine temperature and pressure relationships during metamorphism. In addition, many of the optical properties of gemstones and minerals are the direct result of compositional substitutions. For example, the red color of ruby is due to chromium ion (Cr^{3+}) substitution in the mineral corundum (Al_2O_3), whereas titanium ion (Ti^{4+}) substitution in the same mineral produces blue sapphire. Similarly, beryl ($Be_3Al_2Si_6O_{18}$) exhibits a range of color (blue-green: aquamarine; green: emerald; pink: morganite) caused by trace atomic substitutions in the crystal structure. Alloys represent metallurgical solid solutions in which composition is controlled to produce specific physical, chemical, and electrical characteristics. *See* ALLOY STRUCTURES; AMPHIBOLE; CRYSTAL STRUCTURE; CRYSTALLOGRAPHY; GARNET; GEM; METAMORPHISM; MINERALOGY.

John C. Drake

Bibliography. I. S. E. Carmichael and H. P. Eugster (eds.), *Thermodynamic Modeling of Geological Materials: Minerals, Fluids, and Melts*, 1987; C. Klein and C. S. Hurlbut, Jr., *Manual of*

Mineralogy, 21st ed., 1993, rev. 1999; A. Putnis, *Introduction to Mineral Science*; 1992.

Solid-state battery

A battery in which both electrodes and the electrolyte are solids. Solid electrolytes are a class of materials also known as superionic conductors and fast ion conductors, and their study belongs to an area of science known as solid-state ionics. As a group, these materials are very good conductors of ions but are essentially insulating toward electrons, properties that are prerequisites for any electrolyte. The high ionic conductivity minimizes the internal resistance of the battery, thus permitting high power densities, while the high electronic resistance minimizes its self-discharge rate, thus enhancing its shelf life. Examples of such materials include Ag_4RbI_5 for Ag^+ conduction, LiI/Al_2O_3 mixtures for Li^+ conduction, and the clay and β-alumina group of compounds ($NaAl_{11}O_{17}$) for Na^+ and other mono- and divalent ions. At room temperature the ionic conductivity of a single crystal of sodium β-alumina is 0.035 S/cm, comparable to the conductivity of a 0.1 *M* HCl solution. This conductivity, however, is reduced in a battery by a factor of 2–5, because of the use of powdered or ceramic material rather than single crystals. Of much interest are glassy and polymeric materials that can be readily made in thin-film form, thus enhancing the rate capability of the overall system. *See* ELECTROLYTE; ELECTROLYTIC CONDUCTANCE; IONIC CRYSTALS.

In some cases, solid-state batteries also make use of fast ion conductors for the electrodes. These materials are, however, good conductors for both ions and electrons. Examples include the layered-structure disulfides of titanium and vanadium, TiS_2 and VS_2, which can be used as the cathode or sink for lithium, and aluminum-lithium alloys, which can be used as the anode or source of lithium in lithium batteries. These materials can sustain high reaction rates because of their unique crystalline structures which allow the incorporation of ions into their crystalline lattices without destruction of those lattices (**Fig. 1**). An example of such a cell is given by notation (1), having anode reaction (2) and cathode reaction (3). At the anode the face-centered cubic lattice of aluminum expands from 0.405 to 0.638 nanometer, whereas at the cathode the titanium disulfide sandwich structure expands only 10% on incorporation of the lithium.

$$\text{Li-Al}//\text{Li glass}//\text{TiS}_2 \quad (1)$$

$$\text{LiAl} \rightleftharpoons \text{Al} + \text{Li}^+ + e^- \quad (2)$$

$$\text{TiS}_2 + \text{Li}^+ + e^- \rightleftharpoons \text{LiTiS}_2 \quad (3)$$

See INTERCALATION COMPOUNDS; SOLID-STATE CHEMISTRY.

Use. Potentially the largest use of solid electrolytes in the battery field is in the high-energy-density sodium/sodium-β-alumina/sulfur battery. This system is expected to find application for powering of automotive vehicles and for energy storage at central power plants, such as solar plants, provided the lifetime of the ceramic electrolyte can be improved. However, this battery operates at about 570–750°F (300–4000°C), so that both electrodes are in the liquid state. *See* ENERGY STORAGE.

Solid-state batteries generally fall into the low-power-density and high-energy-density category. The former limitation arises because of the difficulty of getting high currents across solid–solid interfaces. However, these batteries do have certain advantages that outweigh this disadvantage: They are easy to miniaturize (for example, they can be constructed in thin-film form), and there is no problem with electrolyte leakage. They tend to have very long shelf lives, and usually do not have any abrupt changes in performance with temperature, such as might be associated with electrolyte freezing or boiling. Being low-power devices, they are also inherently safer. The major applications of these batteries are in electronic devices such as cardiac pacemakers, cameras, electrochromic displays, watches, and calculators.

Configuration. In a solid-state cell (**Fig. 2**), the polycrystalline pressed electrolyte is interspaced between a metallic anode and the solid cathode material. The electrodes are applied to the electrolyte by mechanically pressing the materials together, or in some cases the electrolyte is formed in place by reaction between the two electrodes. These cells are then stacked together to form a battery of the required voltage. A carbon current collector is often used on the cathode side, and this is frequently admixed with the cathode material. In some cases, where the

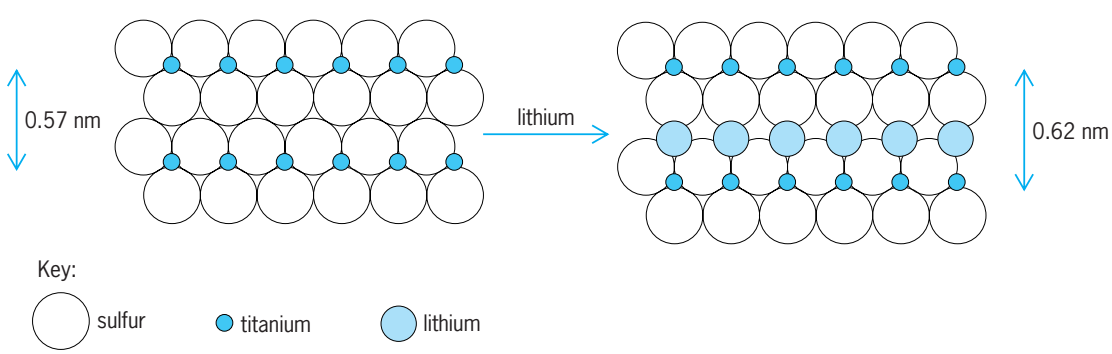

Fig. 1. Schematic illustration of an intercalation cathode reaction.

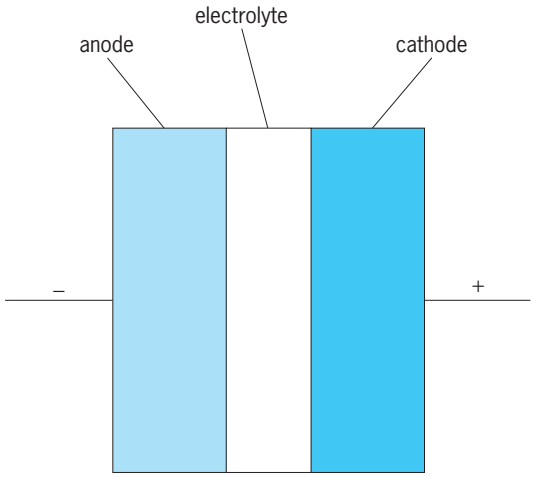

Fig. 2. Schematic diagram of solid-state battery.

cathode material is sufficiently conductive, for example titanium disulfide, no carbon conductor is needed. Although some of the earliest solid-state batteries were based on silver, these have not found any significant applications because of their low energy densities, their high cost, and competition from lithium batteries.

Lithium battery. The lithium-iodine battery employed in many cardiac pacemakers is constructed with a lithium anode and an iodine–polyvinyl pyridine cathode. This cathode has a tarlike consistency in the fresh battery, then solidifies gradually as the battery is discharged. In this cell the solid electrolyte, lithium iodide, is formed in place by the reaction of the lithium and iodine. Because the electrolyte continues to form as the battery discharges, the overall resistance of the cell continually increases with discharge. This results in a drop in cell voltage for a given current drain. The initial cell voltage is around 2.8 V.

The most important characteristics of the lithium battery for this application are its high reliability and very low rate of self-discharge. A typical implantable battery has a volume of 0.4 in.^{3} (6 cm^3), weighs 0.8 oz (23 g), and has a 2-amperehour capacity. The lifetime of this battery is 5–10 years since pacemakers typically draw only 15–30 microamperes. The battery is considered to be discharged when the output voltage drops to 1.8 V. This may be detected by monitoring the patient's pulse rate. *See* MEDICAL CONTROL SYSTEMS.

A reversible lithium solid-state battery uses a glass as the electrolyte, pure lithium as the anode, and an intercalation compound as the cathode. These compounds, such as titanium disulfide, react reversibly with lithium, as discussed above. The overall resistance, unlike that of lithium-iodine batteries, does not increase with discharge, nor is a conductive element such as carbon necessary in the cathode. The emf of such a cell is around 2 V and decreases slightly but continuously with loss of capacity; this allows the cell voltage to be used as a fuel gage, unlike most conventional batteries where there is an abrupt loss of voltage without warning when the cell is depleted.

An advanced lithium battery uses a lithium carbon compound as anode and lithium cobalt oxide ($LiCoO_2$) as cathode. The anode incorporates lithium between the layers of graphitic coke, thereby containing the lithium safely and making cell rechargeability much more feasible. The lithium cobalt oxide cathode has the same structure as lithium titanium disulfide (Fig. 1) and behaves in the same manner. However, it is much more oxidizing, leading to a cell emf of about 3.5 V. Since this is almost exactly three times that of Ni/Cd or Ni/hydride batteries, this battery can easily be used in their place. Moreover, since a single cell will suffice for such applications as portable computers, battery management is much easier.

Polymeric electrolytes. Solid-state batteries based upon polymeric electrolytes are being actively investigated, particularly for use with lithium and sodium anodes. Energy densities of 5–10 times that of lead-acid cells have been calculated based upon very thin cells combined with high-voltage arrays. These cell designs could eventually lead to rechargeable solid-state batteries with energy densities of 100–200 watthours per kilogram and comparable power levels. They are most likely to be used with lithium anodes and conductive cathodes such as titanium disulfide. The polymer most under study is polyethylene oxide mixed with a lithium salt such as $LiCF_3SO_3$. *See* BATTERY.

M. Stanley Whittingham

Bibliography. B. V. R. Chowdari and S. Radhakrishna (eds.), *Solid State Ionic Devices*, 1988; J. P. Gabano, *Lithium Batteries*, 1983; G. Pistoia (ed.), *Lithium Batteries,* 1994; J-M. Tarascon, Recent advances in rechargeable Li batteries, *Solid State Ionics,* 69:171–368, August 1994; C. A. Vincent, *Modern Batteries*, 2d ed., 1998; M. S. Whittingham and A. J. Jacobson (eds.), *Intercalation Chemistry*, 1982.

Solid-state chemistry

The science of the elementary, atomic compositions of solids and the transformations that occur in and between solids and between solids and other phases to produce solids. Solid-state chemistry deals primarily with those microscopic features which are uniquely characteristic of solids and which are the causes for the macroscopic chemical properties and the chemical reactions of solids. As with other branches of the physical sciences, solid-state chemistry also includes related areas that furnish concepts and explanations of those phenomena which are more characteristic of the subject itself.

The overlap of solid-state chemistry and solid-state physics is extensive. However, the perspectives of the two are sufficiently different that they can easily be identified. In general, solid-state physics treats properties, such as energy and entropy, which are continuously variable in the solid, whereas solid-state chemistry concerns those properties which are discontinuous because of chemical reactions; in a sense, by definition the properties are chemical because they are discontinuous. In another perspective,

solid-state chemistry tends to be based on structure in configuration space, whereas solid-state physics tends to be based on momentum space. *See* INORGANIC CHEMISTRY; SOLID-STATE PHYSICS.

Solid-state chemistry has no single unifying theoretical base and tends to be largely an experimental science supported by several theoretical bases. Consequently, its separation into topics is not well established. Those which are considered to be aspects of solid-state chemistry include, among others, chemical binding, crystal defects, crystal structures, crystal field theory, diffusion in solids, ionic crystals, lattice vibrations, and nonstoichiometry.

As with some other subdivisions of chemistry, the syntheses of new, unusual compounds and phases is one of the primary activities. In solid-state chemistry these compounds and phases have complex and varied structures with large unit cells; they are frequently composed of several different atoms and large integral numbers of atoms, and they tend to be nonstoichiometric phases of variable composition. One of the longer-ranged objectives is to classify these structures and to quantify the multiplicities of chemical bonds involved.

The techniques and procedures used to synthesize the compounds and phases are varied. They include (1) direct combination of the elements to produce an isolated compound or combination of compounds at high temperatures, high pressures, or both; (2) preparation from a solution or crystallization from molten compounds of the same composition or through the use of a molten flux; (3) chemical vapor transport from the reactants at one temperature to a product at another temperature; (4) growth of a single crystal at the interface of two gels, each containing one of the reactants or mixtures of such; (5) arc melting of a mechanical mixture that is sufficiently conductive to maintain an arc; (6) intercalation or deintercalation to produce metastable phases; and (7) implosive, thermomechanical impaction with reaction to produce a dense product. *See* VAPOR DEPOSITION.

Structures

Studies of structures provide a basis for understanding the chemical bonding in solids, their properties, reactions among them, and their variabilities of composition.

Spatial geometries. The spatial arrangement to form a stable compound of atoms or ions in a solid is determined by a minimum in the potential energy of the possible assemblages. For the most part, the factors involved are the atomic or ionic radii, the relative strengths of the bondings, and the directionality of the bonds. One of the objectives of solid-state chemistry is to develop a basis through which the minimum in the potential energy can be predicted for a given collection of atoms or ions. However, the problem is so complex that it has not been solved. Hence, the immediate objective is to obtain sufficient information so that more complete theories can be developed.

The placement of atoms in space is describable in terms of a unit cell. This cell is the smallest part of the crystal that contains all the symmetries of the structure. These are defined at four levels:

1. There are seven possible crystal systems defined by the shape of the unit cell or for a perfect crystal by its geometric shape. They are cubic, tetragonal, orthorhombic, monoclinic, hexagonal, triclinic, and trigonal.

2. To describe the lattices in these shapes, subdivisions of some of these are required. There are three cubic lattices, two tetragonal, four orthorhombic, two monoclinic, and one each of the hexagonal, triclinic, and trigonal.

3. A further subdivision is defined in terms of the rational and reflective invariance about a given lattice position. These consist of the 32 point groups.

4. A final classification is achieved by combining translation with rotation and reflection. These symmetry operations yield 230 space groups.

These idealized structures exist, strictly stated, only at 0 K (absolute zero). At any finite temperatures, these structures are imperfect or distorted in at least two general respects: the atoms or ions vibrate with amplitudes which increase with increasing temperature; and the lattice sites of the space groups become vacant or interstitial positions become occupied, or both occur, with the valencies of the cations or anions becoming altered. *See* ABSOLUTE ZERO; CRYSTAL DEFECTS; CRYSTAL STRUCTURE; LATTICE VIBRATIONS; VALENCE.

The concentration and spatial distribution of the defects in the solid produce various kinds of distortions of the idealized structures. When the defects interact only weakly and do not aggregate, the distortions are mostly localized near the point defects, which may be either randomly distributed or ordered on a long range. For this distribution, the x-ray or neutron diffraction pattern is essentially that of the idealized structures with a charged internuclear distance or lattice parameter and with additional diffraction maxima caused by the long-range order, that is, superstructure. This configuration of point defects occurs especially in several nonstoichiometric phases. In some cases the defects appear to be aggregated, with the result that a structured distortion occurs. Such a situation exists in cases known as shear structures that occur in tungsten trioxide (WO_3) and molybdenum trioxide (MoO_3), where extensive regions appear as an idealized structure connected by a plane in the lattice in which a displacement equivalent to a shear translation exists. *See* NONSTOICHIOMETRIC COMPOUNDS.

Layered structures. Numerous binary compounds or phases have layered structures that in themselves or in combination with other elements can be considered as extensions of the descriptions given above. The layers are displaced relative to each other in such a way that the structures can be classified in the space groups, but when other elements or compounds are incorporated between the layers, the structure is highly distorted. The most widely recognized material to possess the layered structure and the associated solid-state chemistry is graphite. Carbon atoms in this structure are arranged in hexagons connected along the edges, and thus form a plane throughout

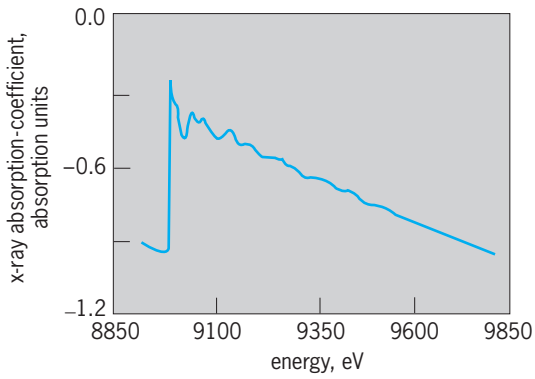

Fig. 1. X-ray absorption spectrum at 20 K (−420°F) for copper in $Ba_4Ca_3Sr_3Cu_4O_x$ obtained with x-radiation near the CuK x-ray edge. The extended x-ray absorption fine structure (EXAFS) occurs after the sharp increase. (*After P. P. Lottici, G. Antonioli, and F. Licci, T-dependence of EXAFS in the Bi-Ca-Sr-Cu-O system, Physics C, 152:468–474, 1988*)

the crystal. These planes are stacked in layers, with each layer displaced with respect to its adjacent layers by a distance such that the structure still belongs to one of the space groups. The bonding within a layer is covalent; between the layers it is metallic. Consequently, graphite is highly anisotropic. When it is exposed to bromine, alkali metals, and other chemical substances, these substances, at high temperatures, enter the spaces between the layers to form intercalation compounds; the idealized structure thus becomes highly distorted. A large number of compounds with layered structures are known to form a variety of intercalation compounds. Among them are titanium disulfide (TiS_2) and molybdenum disulfide (MoS_2). Two others, which are similar in that additional elements or compounds are located between lattice planes, are the tungsten bronzes and the β-aluminas. *See* GRAPHITE; INTERCALATION COMPOUNDS.

Multicomponent phases. For three or more components, the structural situation is generally complex. The number of possible compounds and phases is extremely large, and so is the variety of structures and distortions. Classification of the distorted space groups is difficult and is still being developed. The actual structures are generally described as distortions of some well-recognized structures, which in themselves are complex. Because many of these complex structures were found in minerals, the distorted structures are frequently classified as being derived from these minerals. Examples of such are spinels, perovskites, bixbyites, feldspars, and garnets.

Glasses. All of the structures described above are based on a lattice with defects but with sufficient long-range order to produce well-defined diffraction patterns. There exists another group or class of solids in which the long-range order is insufficient to yield such patterns. These are the glasses or amorphous solids that generally exist only in metastable conditions. The most readily known among these are the silicate glasses, but there are many others, including some even in metallic systems. *See* GLASS; METALLIC GLASSES.

Molecular solids. In all the cases described, there occur no decisive tendencies toward molecular formation in the solids. However, when the intramolecular forces are great enough, as in organic molecules, the solids are composed of molecular crystals. These solids in general have high volatilities and insufficient defects to cause reactions associated with the transfer of mass in the solids, but they have electronic structures which participate in photochemical reactions in organic solids.

Local and extended structures. All the structural information described above is represented by the location of atoms on the lattice network. The structures are determined by diffraction of x-rays, neutrons, or electrons. Two other features of the structure of solids are the local structure about a given atom and the extended structure on a more global scale. The first is determined by the absorption of x-radiation having a wavelength near that of the characteristic x-rays of a given atom (resonance absorption). The spectrum on the high-energy side of the absorption edge (that is, the sudden increase in **Fig. 1**) has fine structure extending beyond the edge and determined by the distance and number of neighbors in the successive shells surrounding the given, central atom [extended x-ray fine structure (EXAFS)]. The global structure is determined by high-resolution electron microscopy (HREM) that produces lattice imaging (**Fig. 2**). Regions of different lattice structures can frequently be identified, such as the contiguous intergrowth of microdomains in a matrix of the primary lattice, a situation that apparently exists in a nonstoichiometric phase that evolves as the composition is varied between two diphasic regions. *See* EXTENDED X-RAY ABSORPTION FINE STRUCTURE (EXAFS); ELECTRON MICROSCOPE.

Electronic structures. Aside from combinations near the center of the periodic system where covalency is dominant, the elements exist in solids as ions. For the regular elements, the outer, valence

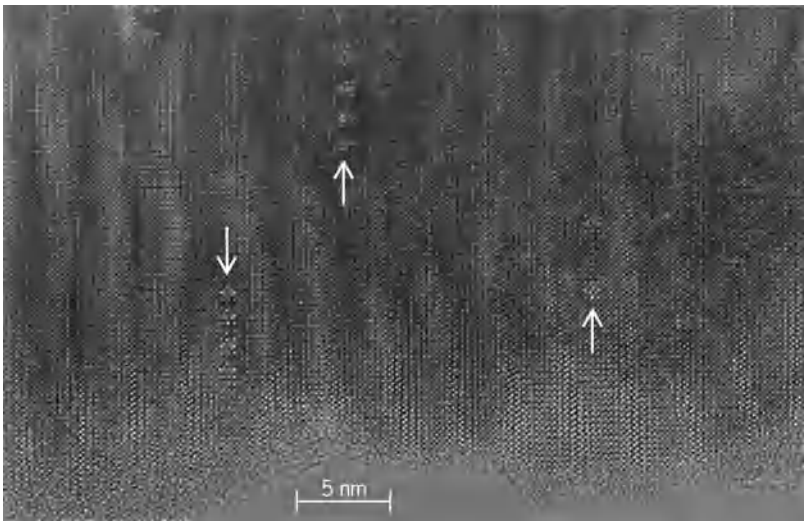

Fig. 2. High-resolution electron microscopic (HREM) lattice image of $Bi_{15}Ti_9Fe_{11}O_{57}$, an Aurivillius phase. Small defect clusters (microdomains) in the center of the perovskite slab are indicated by arrows. (*From M. Lundberg, J. L. Hutchison, and D. J. Smith, The structure of $Bi_{15}Ti_9Fe_{11}O_{57}$ and related compounds derived by high-resolution electron microscopy, J. Sol. State Chem., 80:178–188, 1989*)

orbitals of the ions are closed, with the inert gas shells having *s*-orbital configurations on the cations and hybridized *p*-orbitals on the anion. In both cases, the ions have filled orbitals and tend to be spherical; they are not highly directional. These ions have no low-lying states, and the crystal fields are insufficient to decouple the spins. Consequently, the electronic structures of solids formed with these ions are simple, and they contribute a basis for solid-state chemistry in terms of ionicity and normal polarization. These solids, which are essentially those formed from groups 1, 2, 14, and 17 in the periodic system, are transparent in the visible and near-infrared. In addition, some of the cations of the transition, lanthanide, and actinide groups have closed *p*-orbitals with no low-lying electronic states. These are the scandium and titanium subgroups (thorium is included in the latter). All the other transition, lanthanide, and actinide ions in solids display electronic transitions in their infrared and visible spectra. A large number of the observed transitions are attributable to the fact that the d^n or f^n degenerate states are separated by the electrostatic field of the crystal. In the cases of the lanthanide and actinide ions, some of the observed maxima are also attributable to *f*-to-*d* transitions. The presence of all of these states contributes to anomalous polarization or ionicity and a complexity of the chemistry. *See* ELECTRON CONFIGURATION; PERIODIC TABLE.

Photoacoustic spectroscopy with powdered materials can be used to study the spectra of such solids. Such studies furnish microscopic information for the solid-state chemistry associated with charge transfer, polarization, and valence states. *See* PHOTOACOUSTIC SPECTROSCOPY.

A feature of the electronic structure of solids, which is particularly essential to solid-state chemistry, is the valence state of the ion and the energy required to change the valency in the solid. For the transition, lanthanide, and actinide ions with open *d*- or *f*-orbitals, these energies are sufficiently small that transitions from one valency to another can be promoted by thermal (phonon) activation or the oxidation potentials of anions. The relative concentrations of two valencies in solids can be determined chemically through the stoichiometry or through chemical reactions which presumably do not change the relative concentrations. The known stoichiometries in such solids as the iron oxides Fe_3O_4 and Fe_2O_3 and the uranium oxide U_3O_8 imply mixed valence states. If these phases are dissolved under nonoxidizing conditions, then the initial ratio of the two valence states can be determined. The question of the existence of two distinct valence states in the solid has been debated for decades. Some have argued that a resonance condition exists so that two well-defined oxidation states cannot be identified. Most of these arguments are based on the results of x-ray diffraction observations, which give no evidence of two different cation sites. Two techniques have been developed that enable direct observation of the valence states in the solid, photoelectron spectroscopy and Mössbauer spectroscopy. *See* ELECTRON SPECTROSCOPY; MÖSSBAUER EFFECT; VALENCE; X-RAY DIFFRACTION.

Chemical Bonding

The structure of a solid is the result of the operation of interatomic or interionic forces and the size and shape of the atoms or ions. Hence, logically, bonding should be described first and structure second. However, the detailed role of the electrons in interionic forces is so complex, and the quantitative aspects of the problem of minimizing the potential energy with respect to all the possible configurations is so difficult, that structures cannot be derived. Rather, it is necessary to derive some information about bonding from structures, cohesive energies, refractive indices, electron binding energies, polarizabilities, and other properties through the use of models. Because of the number of ions and electrons involved, the wave-mechanical formulation of bonding in solids is extremely difficult to solve. Fortunately, some rather simple, classically based models modified by quantum-mechanical concepts have been and continue to be quite useful. These can be classified generally as ionic, covalent, and metallic bonding and combinations of the three. *See* CHEMICAL BONDING; STRUCTURAL CHEMISTRY.

Ionic bonding. In the ionic model, the solid is composed of positive (cations) and negative (anions) charges located on the respective lattice sites. These imagined point charges create through their collective coulombic interactions a net attractive force. The self-potential of the interionic forces can be calculated with high precision by summing the electrostatic energies. For crystals with high symmetries, this potential is the product of the ratio of a constant, the Madelung constant, to a common internuclear distance and the sum, characteristic for each structure, of terms whose numerators contain the point charges and whose denominators contain some multiple of the common internuclear distance. For crystals of lower symmetries, more sophisticated summations involving Patterson functions are required. The repulsive forces to balance the attractive forces at equilibrium consist essentially of the overlap of the valence electron orbitals. Additional attractive forces are frequently included in the model. These are occasionally referred to as the dispersive forces because they are the ones related to the dipole and multipole moments in ionic crystals and their indices of refraction. Solids having the most ionic character (ionicity) are those formed from combinations of the most electropositive with the most electronegative elements. In general, the ionicity increases with increasing cationic radius and with decreasing anionic radius. *See* IONIC CRYSTALS; X-RAY CRYSTALLOGRAPHY.

Covalent bonding. In the covalent model, the electrons in the bond are shared equally between the atomic cores, and the electrons involved are in bound orbitals, frequently referred to as molecular orbitals. In essence, the sources of the attractive and repulsive forces are electrostatic, but they are formulated wave-mechanically. In purely covalent

bonding, the atomic cores are identical, as in solids of the inert gases, halogens, oxygen, and sulfur.

Metallic bonding. In the metallic model, the solid is composed of cations on lattice sites surrounded by a uniform negative charge of the conduction or free electrons.

Combination bonding. In all real solids composed of different elements, the bonding consists of admixtures of these models. The bonding in ioniclike crystals composed of the regular metallic elements and the more electronegative elements can be rationalized as admixtures of ionic and covalent bonding. The degree of ionicity or covalency can be evaluated through optical dispersion theories of crystals, wherein the ionicity is evaluated from the index of refraction, the density of valence electrons, and some measure of the separation of the bonding and antibonding orbitals. In these formalisms, the group 14 elements carbon, silicon, and so on, are used as the reference for covalency. Thus covalency increases as the pairs of elements in the solid become more nearly the same as group 14 elements. In the case of the transition, lanthanide, and actinide elements in valence states having open shells, rationalization of the bonding on the basis of this description is inadequate. The crystal field of the solid removes the degeneracy of the orbitals, and the electrons in the bond can occupy states which increase the strength of the chemical bond. Thus there occurs a crystal field stabilization and the attendant anomalous dispersion accompanying the absorption in the infrared and visible spectra. Low-lying *f*-to-*d* transitions may also contribute to the increased bonding. In a sense, the role of these crystal field splittings and *d* states is to increase the polarizabilities of the cations.

In metals formed by higher-valent cations, starting at least with 3+, the anionic-forming elements can react to produce compounds and phases which still contain free electrons. Thus solids such as subhalides of scandium, the monosulfides and carbides of several elements, and some oxides such as titanium oxide (TiO) or vanadium oxide (VO) contain cations, anions, and conduction electrons or a bandgap sufficiently small to be *n*-type semiconductors. The roles of the valence states of the cations and the oxidation of the free electrons are illustrated through the variation of the conductivities and lattice parameters of the lanthanide monosulfides and monoselenides. All of these are conductors except those of europium, ytterbium, and samarium. The rationalization of this behavior is that in the metals the cations are in the 3+ valency in all cases except europium, ytterbium, and samarium, in which the valency is 2+. Thus the anionic elements, sulfur and selenium, are reduced by the free electrons to the 2− state so that the three cases with 2+ cations have no free electrons, whereas each of the others has approximately one free electron. The cationic radii and the oxidation-reduction potentials are consistent with this description.

Measurement. One of the more direct ways of measuring the bonding in a solid, and particularly one that measures it at the cationic and anionic sites separately, is by measuring the electron binding energies by photoelectron spectroscopy. Thus, through the ejection of electrons from valence orbitals with x-rays or ultraviolet radiation $h\nu$ and the measurement of the electrons' kinetic energy, their binding energies are determined by reaction (1). In this re-

$$M^{q+}cs \xrightarrow{h\nu} M^{(q+1)+}cs^{*} + e(g) \qquad (1)$$

action *cs* represents a cation site; the asterisk indicates that the cation site is unrelaxed after the ionization; and the electron is in the free gaseous (*g*) state. Solid-state physicists generally refer binding energies to electrons in the Fermi level. In reaction (1) the reference is the free electron; the two differ by the instrumental work function. This technique permits determination of the valence state, $q+$, and the lattice self-potential from which measures of polarization, ionicity, and so on are derived. The lattice self-potential is derived through a comparison of the binding energy with the ionization potential of the gaseous cation.

Chemical Composition

In a broad sense, the chemical composition of a solid can be classified into one of two groups. The first class comprises compositions in which the ratios of atoms are integers, albeit large in some cases; the structures of these are described in terms of the space groups. This situation is essentially the law of definite proportions, and the compounds are frequently referred to as daltonides. The second class comprises compositions that are continuously variable throughout a phase, because either one atom or ion can be substituted for another of comparable radius (usually known as solid solutions) or atoms or ions can be accommodated through the occupation of interstitial or created vacant sites. A change in valency of either the cation or the anion or both occurs as the composition is varied. This group is frequently referred to as berthollides. *See* DEFINITE COMPOSITION, LAW OF; SOLID SOLUTION.

At finite temperatures, and particularly at high temperatures, the partial vapor pressures of the components in a solid are different. Consequently, in general there occurs a preferential loss of one component so that any solid at equilibrium tends to contain lattice defects and to become nonstoichiometric. However, in a large number of cases the deviations from stoichiometry are not detectable, so that the number of nonstoichiometric phases that can be studied in solid-state chemistry is not extremely large. Among those which have been studied extensively are $Fe_{1-x}O$, CeO_{2-x}, UO_{2+x}, $YBa_2Cu_3O_{7-\delta}$, and various metal hydrides, as well as the other oxides, carbides, and hydrides.

Whenever the structure of a solid is such that an interstitial position of one of the sublattices can be occupied, both interstitials and vacancies occur. However, over the compositional range of nonstoichiometry, one or the other of the defects is usually in the higher concentration. Thus UO_2 has a fluorite structure such that the position at the center of the

unit cell can accommodate an oxygen ion with an attendant shift in the oxidation state of the uranium ions. In UO_{2+x} the defects are predominantly oxygen ions on interstitial sites and U^{5+} ions on some of the cation sites. In $UO_{2-x'}$ which exists at temperatures near 2000°C (3600°F), vacancies on the oxygen sublattice are at higher concentration. In a phase such as $Th_yU_{1-y}O_{2+x'}$ thorium ions are substitutional, and in $Pu_yU_{1-x}O_{2\pm x}$ plutonium is substitutional. In the last case, because plutonium has the 3+ and 4+ oxidation states accessible and uranium has the 4+ and 5+ states, the anion composition has an extensive range on both sides of stoichiometry. In metallic carbides, the carbon is often interstitial, and in UC_{1+x} the carbon tends to be incorporated in the UC lattice as a C_2 unit.

The role of the compositional variable in oxidation at the electronic level in the solid is illustrated in the x-ray photoelectron spectra of the valence band region of uranium at four stages of oxidation from metal to UO_3. For uranium metal and dioxide, the intensities of the 5*f* orbital electrons are the same; no change occurs in the number of 5*f* electrons. Hence, the oxidation of the metal to UO_2 involves only the conversion of the free electrons to bound electrons in the hybridized 2*p* orbital on the oxygen ion. When UO_2 is oxidized to U_4O_9, however, the intensity of the 5*f* part decreases and that of the 2*p* increases. With further oxidation to U_3O_8, the 5*f* intensities decrease further and the 2*p* intensity increases. Finally, in UO_3, the 5*f* electrons are completely oxidized to 2*p* electrons on the oxygen ion.

One of the measurements, which furnishes information needed to understand the compositional variable in solid-state chemistry, is the observation of partial vapor pressures, and their dependence on temperatures is determined. The chemical potentials and the partial molar enthalpies and entropies and their variation with composition are derived from these measurements. For a binary system, the chemical potential increases monotonically in the single-phase region, and is constant in the diphasic region. The construction and mathematical evaluation of statistical models to describe the chemical potential throughout both regions from the interionic forces, valence states, and defect energies contain conceptual problems associated with discontinuities at the diphasic region. Models can be constructed and evaluated to describe the homogeneous regions. One model is described by Eq. (2), in which θ_i is

$$\mu = \mu_0 + RT\left[\ln\left(\frac{\theta_i}{1-\theta_i}\right) - E_i - 2\theta_i E_{ii}\right] \qquad (2)$$

the fraction of interstitial sites occupied, E_i is the energy required to remove an ion from the interstitial site, and E_{ii} is the energy of interaction between two occupied interstitial sites. Above some critical temperature ($T_c = E_{ii}/2R$), this function has a sigmoidal shape with no maximum or minimum. Thus for $T > T_c$, the variation of μ with θ_i (composition) (**Fig. 3**) represents a monophasic region. At the critical temperature, the isotherm has a zero slope. Below this critical temperature, the function has a maximum, a minimum, and an inflection between them (as shown by the isotherm for $T < T_c$ in Fig. 3). Because the slope of μ versus composition in a real system cannot change sign, this behavior is unrealistic. The artifice that is introduced to excuse these van der Waals loops is to construct a horizontal line through the inflection point and use this to represent the diphasic region. A realistic model that describes the evolution to a diphasic region would be one in which the horizontal line is contained in the mathematics. In one technique that has been suggested, the chemical potential is described through the complex (mathematical) variable. In some cases (praesodymium oxides), the compositional isotherm displays hysteresis loops because of the defect complexes that form differently in the two directions of μ versus composition.

When a substantial nonstoichiometric phase is formed between an insulator and a metallic conductor, the electrical conductivity and concentration of free electrons change continuously. However, in some cases and perhaps in all cases, the change is so rapid over a small compositional range that the change can be viewed as an insulator-to-metal transition. One system that displays this behavior is samarium neodymium selenide ($Sm_{1-x}Nd_xSe$). SmSe

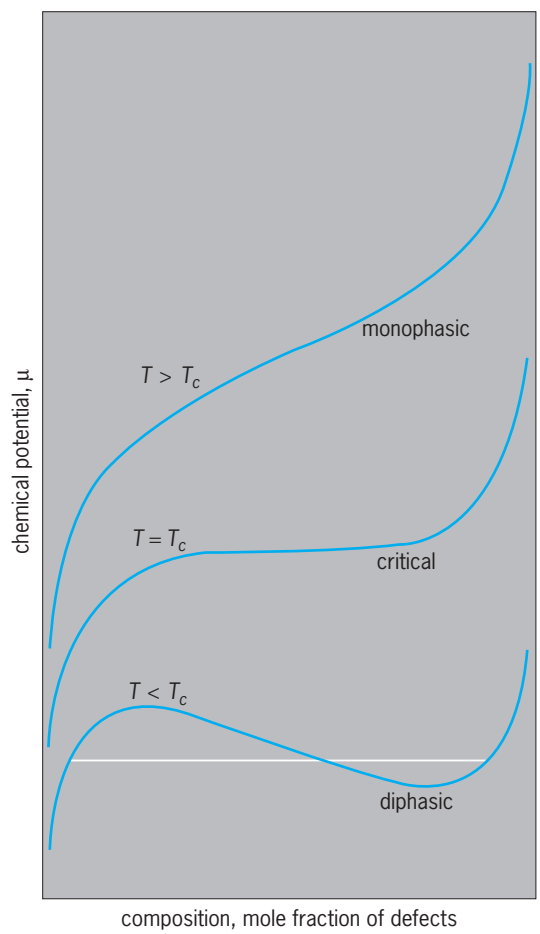

Fig. 3. Three isotherms for the variations of chemical potential with composition of defects.

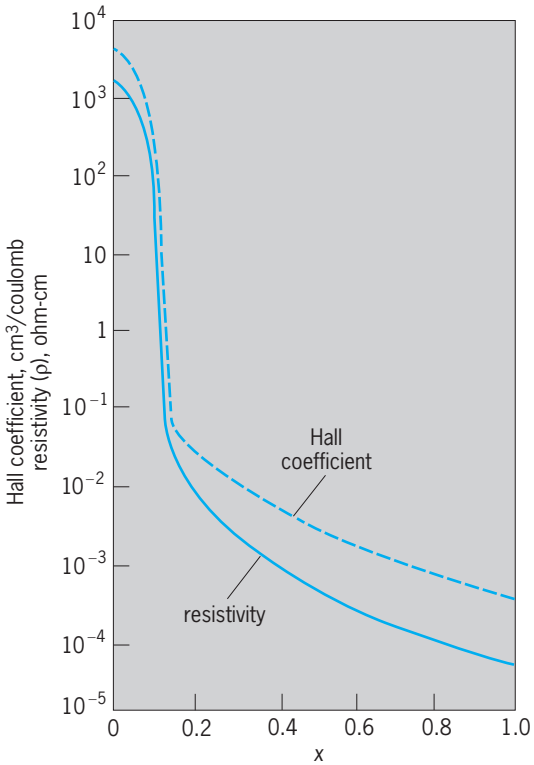

Fig. 4. Electrical properties versus composition in *n*-type $Sm_{1-x}Nd_xSe$, where *x* represents the variation in stoichiometry. The value of the Hall coefficient at $x = 1$ is calculated with the assumption of one electron per neodymium atom. Temperature = 300 K (80°F). (*After F. J. Ried et al., Electrical conduction in rare-earth monoselenides and monotellurides and their alloys, J. Phys. Chem. Sol., 25:969–976, 1964*)

is an insulator with Sm^{2+} and S^{2-}, and NdSe is a conductor with Nd^{3+}, S^{2-}, and one free electron per unit. As the composition is varied, the conductivity and concentration of free electrons change significantly, but at $x \cong 0.1$ the two change by five orders of magnitude (**Fig. 4**).

One of the essential features of the superconducting oxides is nonstoichiometry. The extent varies, but all appear to be phases of variable composition. One of the early ones to be discovered is the barium-lead-bismuth oxide $BaPb_{1-x}Bi_xO_3$, which is a single phase from $BaPbO_3$, a metalliclike conductor to $BaBiO_3$, a semiconductor. Both lead and bismuth have dual valences. The oxide that stimulated great interest in superconducting oxides is $La_{2-x}Ba_xO_{4-\delta}$ (at 35 K or −397°F), $0 \leq x \leq 0.15$ and δ small. The one having the largest range of oxygen composition is $YBa_2Cu_3O_{7-\delta}$ (at 92 K or −294°F), where Y = yttrium, $6 \leq \delta \leq 7$, and the formal oxidation states of copper (Cu) are from 1+ to 3+. Several others were found in a relatively short time. Intense research on these superconducting oxides awakened significant interest in solid-state chemistry.

The oxide $YBa_2Cu_3O_{7-\delta}$ has a perovskite structure (**Fig. 5**). The central part of this structure, which can be written as $(YCu_3O_6Ba_2)^-$, appears to play a rather minor role in determining the essential properties. The yttrium can be replaced by several lanthanide elements, especially in the second half of the series where the structures resemble those for yttrium. Substitution of holmium or gadolinium, with large magnetic moments, for yttrium has little if any effect. Barium that is adjacent to the basal plane where the chemical entity can be written as $(CuO_{1-\delta})^+$ appears to play an indirect role; its replacement by strontium does affect the properties; it decreases the temperature, but the primary variabilities are attributable to the oxygen content in the basal plane. For $\delta = 0$, the structure is orthorhombic; for $\delta = 1$, it is tetragonal. The transition from orthorhombic to tetragonal is second-order or order-disorder. The diamagnetic susceptibility varies sigmoidally with decreasing temperature during the transition to the superconducting state from a value near zero to a value near $-1/(4\pi)$ for the optimum situation (**Fig. 6**). As the oxygen content decreases from a value near 6.9, the asymptotically attained value of the susceptibility increases. *See* SUPERCONDUCTIVITY.

Chemical Reactions

The mechanisms of chemical reactions within and between solids are through lattice vibrations, lattice defects, and changes in valence states. These are the structural features through which migration of mass, charge, and energy occur. Consequently, diffusion and conductivity are integral, basic parts of solid-state chemistry, even though their quantitative roles in the totality of chemical reactions have not been developed. So long as the solid phase produced during reaction has a density nearly the same as those for the reactants, the microscopic description of the

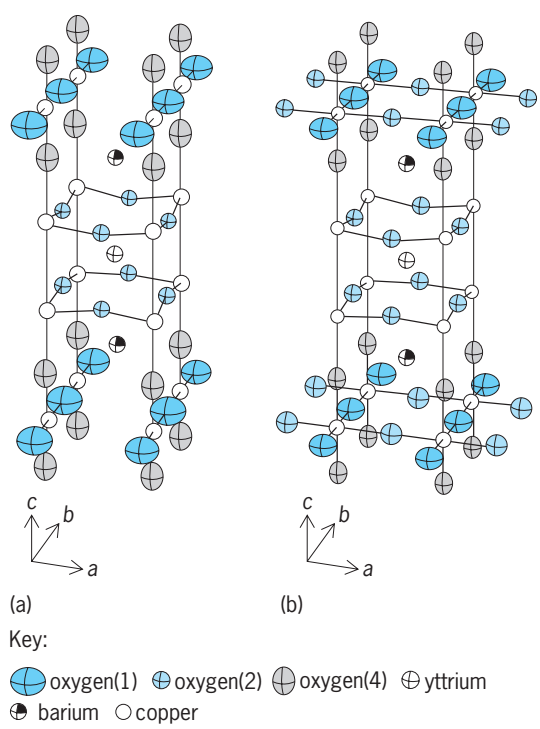

Fig. 5. Perovskite structures $YBa_2Cu_3O_{7-\delta}$ of (*a*) orthorhombic and (*b*) tetragonal. In the tetragonal structure the ellipsoidal symbols for the O(1) site is used to indicate that this site is not fully occupied. (*After J. D. Jorgensen et al., Structural properties of oxygen-deficient $YBa_2Cu_3O_{7-\delta}$, Phys. Rev. B, 41:1863–1877, 1990*)

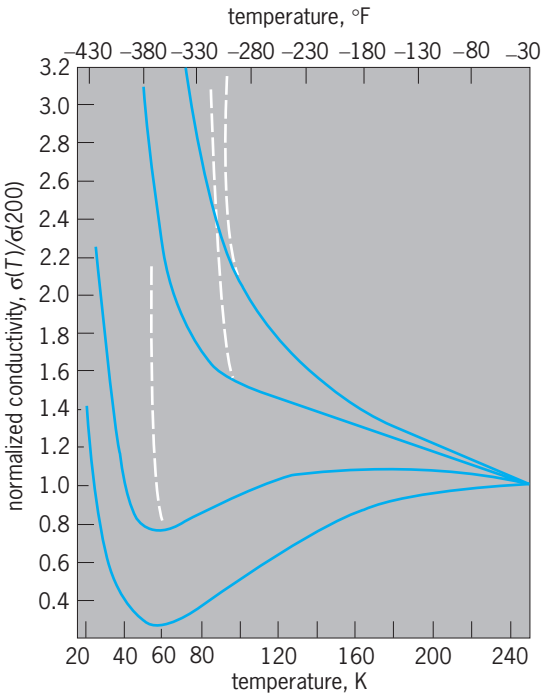

Fig. 6. Normalized electrical conductivity versus temperature for various concentrations of oxygen in $YBa_2Cu_3O_{7-\delta}$. Oxygen content decreases from the topmost curve to the lowest. (*After R. J. Thorn and C. E. Thorn, Semiconductive and semimetallic components in superconducting oxides, J. Phys. Chem. Sol., 50:153–161, 1989*)

reaction in terms of mass and charge transfer is feasible. However, in a number of reactions the molar volumes differ sufficiently to destroy the integrity of the product, so that much of the reaction then proceeds at interfaces, microcracks, and fissures. Thus the total net mechanism becomes a composite process so complex that a comprehensive theory is lacking. The theories that have been developed are generally discussed under topics such as corrosion. Thus, no formal classification of reactions in solids is universally recognized, or, in fact even cited, in the literature. *See* DIFFUSION; ELECTRICAL CONDUCTIVITY OF METALS.

Condensation. Although it is well recognized that condensation to form solids occurs when the translational energy of the gaseous reactants is dissipated into vibrational and electronic states, the processes involved are sufficiently complex that they are difficult to study and to describe in detail. In general, two kinds of condensation need to be described. In a supersaturated vapor, condensation is a stepwise process in which molecules form in successively larger clusters if the rate of growth of the clusters exceeds the rate of dissociation. At some critical size, the process results in condensation, because the translational energy can be dissipated into the cluster. This process occurs in shock tubes and jets which produce super-saturated vapors and in hydrodynamically flowing gaseous beams in which adiabatic expansion occurs. Another kind of condensation occurs if atoms or molecules in the gas are sufficiently accommodated thermally on a colder substrate. In this case, the growth occurs primarily through the flow of translational energy into the substrate. However, the atoms or molecules bound to the surface can be mobile on the surface so that growth can occur through a stepwise process therein. Usually the process is nonequilibrious, and an amorphous, metastable product is formed. For instance, gaseous silicon monoxide can be condensed on a cold substrate to produce solid, amorphous, metastable silicon monoxide. This is the only known way of producing it. Condensation to produce epitaxial layers on substrates is used to produce various kinds of solid-state devices.

An elementary condensation which can be used to illustrate the processes involved and which represents what could be called the first step in a solid-state chemistry is the formation of a metal. Because the process is one in which the neutral gaseous atom is converted to a collection of cations and free electrons, it involves internal oxidation and reduction, represented by reaction (3). Thus, in this reaction

$$M(---np^6) \rightarrow M^{q+}(---np^{(6-q)})e_q \qquad (3)$$

the gaseous atom with the electronic configuration of the core electrons (- - -) and the outer six *np* electrons condenses to form something resembling a cation of charge $q+$ associated with q free electrons. In the supersaturated vapor, the critical step probably is the one in which the cluster is large enough to have free or conduction electrons to serve to dissipate the translation energy. When gaseous metal-forming atoms impinge on a metal substrate, the free electrons in the latter serve to transfer the energy, so that the accommodation and condensation coefficients of metal on metal are generally unity. On nonmetallic insulators, however, the coefficients are not unity, and condensation occurs through nucleation on the surface. The significance of recognizing that the condensation to form a metal is a redox process is contained in what is involved upon further oxidation. When the solid metal undergoes further reaction with an anion-forming element, it is the free electrons which are used first, and the cation is subsequently oxidized. *See* CRYSTAL GROWTH; OXIDATION-REDUCTION.

Internal processes. One of the simplest reactions that occurs in solids is a phase transition. These transitions may be either first- or second-order. In the first case, a change occurs in the crystal structure, sometimes with the higher-temperature form having an apparently higher symmetry. Second-order transitions are usually order-disorder transitions such as the λ-type transitions which occur on only one of the sublattices. *See* PHASE TRANSITIONS.

The primary internal reactions which involve a chemical reaction are those associated with site occupancy, changes in valency, and clustering. Because it has been extensively studied, uranium dioxide (UO_2) serves as a useful illustration. In this phase, the interstitial position (*is*) of the fluorite lattice can accommodate an oxygen ion. Thus, an equilibrium exists between the occupancy of this site and the

regular anion sites (*as*), as shown in reaction (4).

$$O^{2-}as + is \rightarrow O^{2-}is + as \quad (4)$$

Whenever uranium dioxide is oxidized, the valence of uranium is increased from 4+ to 5+ so that reaction (5) is the one occurring primarily. Reactions

$$2U^{4+}cs + O(g) + is \rightarrow 2U^{5+}cs + O^{2-}is \quad (5)$$

similar to these are responsible for the existence of nonstoichiometric phases. The dual valency on the cation also is involved in hole conductivity in *p*-type semiconductors.

Another nonstoichiometric oxide that has been extensively studied is $YBa_2Cu_7O_{7-\delta}$ which, for δ approximately equal to 0.1, is a superconductor. In this phase the internal equilibrium is Eq. (6). In

$$Cu^{3+}cs + O^{2-}as = Cu^{2+} + O^{1-}as \quad (6)$$

the range of homogeneity $O \leq \delta \leq 1$, the structure changes from orthorhombic to tetragonal as oxygen is added in the basal plane (Fig. 5). For intermediate compositions, the neutron diffraction patterns are analyzed in terms of fractional occupancies of the sites in the basal plane, so the real unit cells are much larger. Electron diffraction reveals superstructure and, consistently, high-resolution electron microscopy (HREM) reveals that the phase can be represented as microdomains (such as in Fig. 2) in a matrix, so that the phase can be represented as $(YBa_2Cu_3O_7)_{1-\delta}$ $(YBa_2Cu_3O_6)_\delta$. Photoelectron spectroscopy reveals that the energy for this reaction is small; it occurs at low temperature. *See* CHEMICAL EQUILIBRIUM.

A comparison of the energies associated with these reactions in solids and in gases illustrates the differences in the two cases. In the gas phase, the O^{2-} ion is unstable; in the solids, it is stabilized by the lattice self-potential. The energy associated with reaction (6) in the gas phase equals the difference between the two ionization potentials. Usually, this difference is of the order of tens of electronvolts. However, in the solid, this difference is compensated by the difference in the lattice self-potentials. Consequently, the energy associated with reaction (6), in the solid phase, is only a few electronvolts. *See* HOLE STATES IN SOLIDS; SEMICONDUCTOR.

Interphase reactions. Reactions between phases accompanied by product formation with integrity maintained occur via chemical diffusion. This situation occurs when the reactants and the product have the same structure and comparable ionic radii. But even in this case, there generally is a volume change, the concentration of defects changes, and the initial interface between the reactant moves. A classical example is the reaction (that is, diffusion) between copper and copper-zinc alloy. The zinc diffuses more rapidly out of the alloy than the copper diffuses in, and the original interface moves toward the alloy. This phenomenon is known as the Kirkendall effect; it is one of the evidences that diffusion occurs via vacancies and interstitials rather than through positional exchange. *See* ALLOY STRUCTURES; INTERFACE OF PHASES; PHASE EQUILIBRIUM.

Two well-studied solid-state chemical reactions are sintering and corrosion. In the first reaction, mass is transferred between particles in a solid so that densification and frequently plastic deformation or creep occurs. The details of the transfer between particles is not well understood, but self-diffusion is generally involved. In corrosion, diffusion through the product layer generally controls the rate of reaction, although electromigration may also be significant. *See* CREEP (MATERIALS); PLASTIC DEFORMATION OF METAL.

In both cases, the mechanism can involve either the migration of a reacting component from the substrate through the layer or the migration of a corroding component from the external surface. When the layer becomes sufficiently thick and its density is sufficiently different, the layer develops fissures and cracks. Then the mechanism occurs along these, but the primary chemical driving force may still be the same as before. *See* CORROSION; SINTERING.

The oxides and other phases of the transition and lanthanide elements are catalysts for heterogeneous reactions at a gas–solid interface. It seems plausible that the lattice defects or the *d* and *f* electronic configurations serve as sites for adsorption of a gaseous reactant so that an energy barrier for the reaction is reduced. *See* CATALYSIS; HETEROGENEOUS CATALYSIS.

Electrochemical reactions. Electrolytic cells with solid electrodes and electrolytes have been used to determine chemical potentials, especially in nonstoichiometric phases. An example is a cell composed of nonstoichiometric oxides, UO_{2+x} and $Fe\text{–}Fe_{1-y}O$, as electrodes and stabilized zirconia as the electrolyte: $Fe\text{–}Fe_{1-y}O|ZrO_2|UO_{2+x}$. The electrolyte is an oxygen ion conductor, so that the voltage developed is produced by the difference in the chemical potential of oxygen in the two solid electrodes. Rechargeable batteries in which either the electrolyte or the electrodes are solids have been investigated. *See* ELECTROCHEMISTRY.

Photochemical reactions. Many of the photoinduced processes which occur in solids can be imagined to be solid-state chemical reactions. The classical ones, of course, are those involved in photographic plates and films. *See* PHOTOCHEMISTRY.

Sublimation. At sufficiently high temperatures, the rate of sublimation of all solids can be measured. The partial vapor pressures of most of the inorganic fluorides, chlorides, and oxides and several of the carbides and sulfides have been determined. In a number of cases, the vapor contains complex molecular species, even though the temperature required is high. The presence of these molecules is, in a sense, a consequence of the relatively high vapor pressure and not the high temperature, which of course at the constant pressure is a degradative factor. The chemical potentials and the partial molar enthalpies and entropies of sublimation are derived from the partial vapor pressures and their dependencies on temperature. These are direct measures of the bonding in

the solid. *See* HIGH-TEMPERATURE CHEMISTRY; SUBLIMATION. R. J. Thorn

Bibliography. D. M. Adams, *Inorganic Solids: An Introduction to Concepts in Solid State Structural Chemistry*, 1974; P. A. Cox, *The Electronic Structure and Chemistry of Solids*, 1987; W. Hayes and A. M. Stoneham, *Defects and Defect Processes in Nonmettalic Solids*, 1985; S. Holt et al. (eds.), *Solid State Chemistry: A Contemporary Overview*, 1980; W. Jost (ed.), *Physical Chemistry*, vol. 10: *Solid State*, 1970; P. Kofstad, *Nonstoichiometry, Diffusion, and Electrical Conductivity in Binary Metal Oxides*, 1972, reprint 1983; G. G. Libowitz, *The Solid State Chemistry of Binary Metal Hydrides*, 1965; *J. Phys. Chem. Solids*, monthly; *J. Solid State Chem.*, monthly; D. L. Nelson and T. F. George (eds.), *Chemistry of High Temperature Superconductors II*, 1988; D. L. Nelson, M. S. Wittingham, and T. F. George (eds.), *Chemistry of High Temperature Superconductors*, 1987; H. Schmalzried, *Solid State Reactions*, 2d ed., 1981; O. T. Sorensen, *Nonstoichiometric Oxides*, 1981; A. R. West, *Solid State Chemistry and Its Applications*, 1987.

Solid-state physics

The study of the physical properties of solids, such as electrical, dielectric, elastic, and thermal properties, and their understanding in terms of fundamental physical laws. Most problems in solid-state physics would be called solid-state chemistry if studied by scientists with chemical training, and vice versa. Solid-state physics emphasizes the properties common to large classes of compounds rather than the dependence of properties upon compositions, the latter receiving greater emphasis in solid-state chemistry. In addition, solid-state chemistry tends to be more descriptive, while solid-state physics focuses upon quantitative relationships between properties and the underlying electronic structure. *See* SOLID-STATE CHEMISTRY.

Many of the scientists who study the physics of liquids identify with solid-state physics, and the term "condensed-matter physics" is increasingly replacing "solid-state physics" as a division of physics. It includes noncrystalline solids such as glass as well as crystalline solids. *See* AMORPHOUS SOLID; GLASS.

Electronic structure of solids. In solid-state physics it is generally assumed that the electronic states can be described as wavelike. The individual electronic states, called Bloch states, have energies which depend on the wave number (a vector equal to the momentum divided by $\hbar$, where $\hbar$ is Planck's constant divided by 2π), and the wave number is restricted to a domain called the Brillouin zone. This energy given as a function of wave number is called the band structure. It is shown for two lines (Γ to X and Γ to L) in the Brillouin zone for metallic copper in the **illustration**. There are several curves, called bands, for each line. *See* BRILLOUIN ZONE.

These bands can be calculated by using quantum theory; the results of such a calculation are given as the continuous lines in the illustration. It is also possible to measure them directly by using the photoelectric effect, as indicated by the points in the illustration. In any solid the electrons occupy the lowest-energy electronic states, one electron of each of the two allowed spin states in each electronic state, or two electrons per atom in each band in a simple structure. In copper there are 11 electrons per atom to occupy the bands shown, and they fill the states up to the Fermi energy E_F, indicated in the illustration. One of the bands is partly occupied (that is, it crosses the energy E_F), so that this solid is a metal; under an applied field or temperature gradient, the electrons shift from state to state, causing currents to flow and making copper a good conductor of electricity and heat and making it shiny. Light at the blue end of the spectrum can also lift electrons from the nearly horizontal bands to the empty states; the absorption of the blue light gives the reddish color to copper. *See* CRYSTAL ABSORPTION SPECTRA; FREE-ELECTRON THEORY OF METALS; METAL; NONRELATIVISTIC QUANTUM THEORY; PHOTOEMISSION.

If copper had an additional electron per atom (as does zinc), there would be just enough electrons to fill the lower bands (the bottom five curves). The illustration suggests that there would then be no partially filled bands, and thus an energy gap separating the full bands from the empty ones. Actually, in copper and in zinc some of the upper bands dip below the peak energy of the lower bands, as would be the

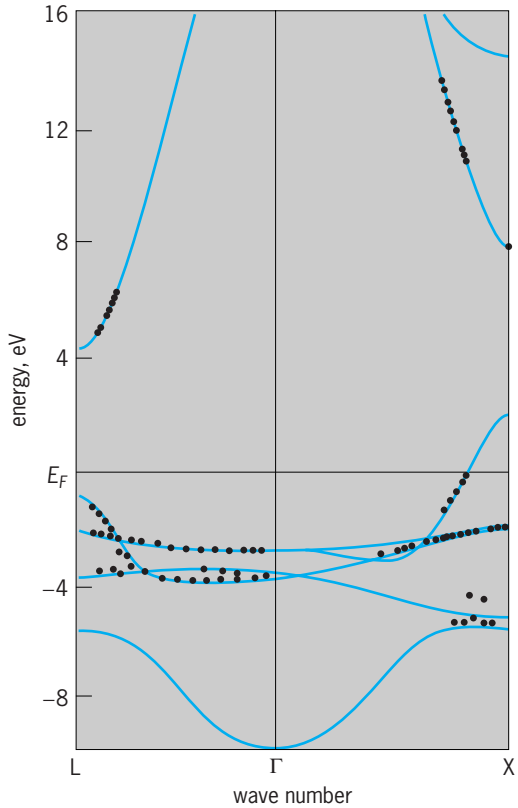

Energy bands of copper. (*After J. A. Knapp et al., Experimental energy band dispersions and lifetimes for valence and conduction bands of copper using angle-resolved photoemission, Phys. Rev., B19:4952–4964, 1979*)

case if the upper band at L dropped below 2 eV in the illustration. Thus copper would still be a metal, as is zinc, even if it had an additional electron.

When such energy gaps between occupied and empty bands do occur, as in silicon, the solid is no longer metallic. Applied fields only shift electrons from state to state within individual bands, and since the same states remain occupied, there is no current generated by, and no heat flow due to, the electrons. If the gaps are small, as is the 1.13-eV gap of silicon, a small number of electrons are thermally shifted at room temperature to the empty bands, called conduction bands, and electrical conductivity is introduced; it can also be accomplished by substituting impurities, called dopants. Phosphorus, for example, has one more electron than silicon and contributes that electron to the conduction band. Such systems in which conductivity can be introduced are called semiconductors. When the energy gaps are large, as is the 8.5-eV gap for rock salt, there are ordinarily no empty states in the lower bands and no electrons in the conduction bands, and the crystals are insulators. *See* BAND THEORY OF SOLIDS; ELECTRIC INSULATOR; SEMICONDUCTOR.

Cooperative phenomena. Each state in the copper bands shown in the illustration can accommodate two electrons of opposite spin. The two spin states have the same energy, so that they are both filled if the energy is less than E_F, or both empty if the energy is greater than E_F; there is no net electronic spin in the system and no magnetism. In nickel, less than copper in atomic number by one, the energies of the two spin states shift from each other. Electrons with spin of one sign become more numerous, and the metal becomes ferromagnetic; each of the bands shown in the illustration has become split into two. Such a transition is called a cooperative phenomenon since it is caused by the mutual lowering of each other's energies by electrons of the same spin. The theoretical and experimental study of such cooperative transitions is an active area of solid-state physics research. *See* FERROMAGNETISM; PHASE TRANSITIONS.

Superconductivity. Another cooperative phenomenon is superconductivity, in which pairs of electrons of opposite spin and moving in opposite directions form Cooper pairs that condense into a cooperative electronic state. There is again an energy gap between the occupied Cooper pair states and the empty states from broken pairs, but in this case the entire cooperative state can drift to carry a current without exciting individual electrons from the cooperative state. Such a flow is called supercurrent since it flows without any electrical resistance whatsoever; it occurs only in some metals and until 1986 it was observed only at very low temperatures.

In 1986 J. G. Bednorz and K. A. Müller discovered a copper oxide compound ($Ba_xLa_{2-x}CuO_4$, containing a small ratio x of barium atoms to copper atoms) that was superconducting at 30 K (−406°F). Soon, other copper oxide compounds were found ($YBa_2Cu_3O_{7-x}$) that were superconducting at 90 K (−298°F), well above the temperature of liquid nitrogen. The development of these so-called high-temperature superconductors may prove very important since it is much easier to cool materials with nitrogen than with the helium required for other superconductors. In spite of the enormous experimental and theoretical study of these materials, the origin of this new kind of superconductivity is not yet understood. *See* SUPERCONDUCTIVITY.

Small supercurrents can even flow through very thin insulators by quantum-mechanical tunneling, an effect predicted by B. D. Josephson. It was predicted and observed after the discovery by I. Giaever that individual electrons can tunnel through thin oxides; Giaever used this tunneling to make a direct measurement of the energy gap arising from the Cooper pairs in superconductors. *See* JOSEPHSON EFFECT; TUNNELING IN SOLIDS.

Crystal structure and lattice vibrations. The total energy of a solid includes a sum of the energies of the occupied electronic states. Since the energy bands depend upon the positions of the atoms, so does the total energy, and the stable crystal structure is that which minimizes this energy. The theory has proven adequate to predict the crystal structure of various solids. For some years it has been possible to predict the changes in energy under various distortions of the lattice. There are in fact three times as many independent distortions, called normal modes, as there are atoms in the solid. Each has a wave number, and the frequencies of the normal vibrational modes, as a function of wave number in the Brillouin zone, form vibrational bands in direct analogy with the electronic energy bands shown in the illustration. These can be calculated directly from quantum theory or measured by using neutron or x-ray diffraction. *See* CRYSTAL STRUCTURE; LATTICE VIBRATIONS; NEUTRON DIFFRACTION; X-RAY DIFFRACTION.

In quantum theory any normal mode can be excited only to discrete quantized vibrational energies. In solids a mode in the nth state of excitation is said to contain n phonons. These phonons interact with the electrons and cause the reduction in conductivity with increasing temperature in metals. This same electron-phonon interaction also provides the attraction that holds electrons together in Cooper pairs to form the cooperative superconducting state. *See* ELECTRICAL CONDUCTIVITY OF METALS; ELECTRICAL RESISTIVITY; PHONON.

Intrinsic and extrinsic properties. Solid-state physics includes the study of all of these properties of ideal crystals, called intrinsic properties. It also includes the study of defects in the structure, impurities, and surfaces, all of which are called extrinsic properties. The study of extrinsic properties can be made with very much the same experimental tools, and they can be understood by using quantum theory with very much the same concepts which are useful for studying intrinsic properties. *See* CRYSTAL DEFECTS.

Walter A. Harrison

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 2d ed., 2000; H. Ehrenreich et al. (eds.), *Solid State Physics*, vols. 1–48, 1955–1994; W. A. Harrison, *Electronic Structure and the Properties of Solids*, 1980; C. Kittel, *Introduction to Solid*

State Physics, 7th ed., 1996; S. A. Wolf and V. Z. Kresin (eds.), *Novel Superconductivity*, 1987.

Solids pump

A device used to move solids upward through a chamber or conduit. It is able to overcome the large dynamic forces at the base of a solids bed and cause the entire bed to move upward.

Solids pumps are used to cause motion of solids in process-type equipment in which treatment of solids under special conditions of temperature, oxidation, and reduction can be combined with upward motion and discharge of the spent solids overhead from the reacting vessel.

Solids pumps are inherently of the positive displacement type. One practical method uses a reciprocating piston mounted on a trunnion, permitting it to swing into an inclined position for filling and then to swing back into vertical position for discharge. **Figure 1** shows a mechanically driven solids pump in four positions through its cycle of operation. **Figure 2** shows a large hydraulically operated pump used in units of capacity exceeding 1000 tons/day (900 metric tons/day). Hydraulic activation permits very precise control of the feeder mechanism and good efficiency in operation.

The solids pump has found its principal application in the operation of oil-shale retorts. Here it is used to feed crushed shale into the bottom of a conical vessel, and as the shale moves upward through this vessel, air is drawn downward countercurrent.

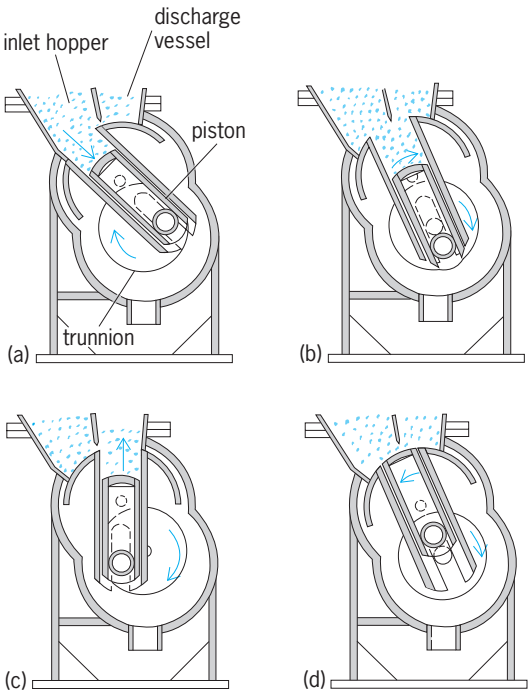

Fig. 1. Operation cycle of mechanically driven solids pump. (*a*) Filling with solids from inlet hopper. (*b*) Piston rotating on a trunnion toward its discharge position. (*c*) Discharge position, with piston pushing charge of solids upward. (*d*) Piston rotating back toward original filling position.

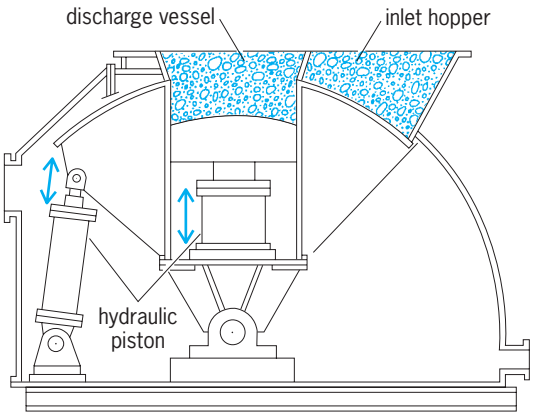

Fig. 2. Hydraulically operated solids pump.

At the top of the retort, the air burns the residual carbon on the shale ash. The hot flue gas so produced contacts shale in the midpoint of the reactor, educting the shale oil. These vapors together with the flue gas are cooled, and the oil condensed on the shale at the bottom of the retort. The oil flows out the bottom countercurrent to the upgoing bed of shale. *See* BULK-HANDLING MACHINES; PUMP; UNIT OPERATIONS.

Clyde Berg

Solifugae

An order of nonvenomous, spiderlike, predatory arachnids found chiefly in arid and semiarid, tropical, and warm-temperate regions worldwide. Solifugids can range in size from a few centimeters to almost 3 in. (7.5 cm). Most of the 1100 known species are nocturnal. They are commonly known as sun spiders, camel spiders, or wind scorpions (see **illustration**). The last name stems from the fact that they exhibit the most rapid sprint speeds for any known group of terrestrial arthropods.

Solifugids possess relatively large anterior appendages, known as chelicerae, that are used for grasping and crushing prey. Ounce for ounce, these jaws produce one of the most powerful bites in the animal kingdom. The second pair of appendages, the leglike and nonchelate palpi, are tactile in

Wind scorpion. (*Glenn and Martha Vargas © California Academy of Sciences*)

function and end in a structure (palpal sucker) that is adhesive in function and helps in procuring prey and mating. The first pair of legs is tactile, whereas the other three are ambulatory.

Most species have a life span ranging from 1 to 3 years. Life cycle stages include the egg, several nymphal instars, and adult. After laying her eggs, the female guards them within the nest until they hatch, and remains with the newly hatched nymphs for several days. Depending on their size, solifugids feed on a variety of arthropods and small vertebrates. These arachnids are most frequently found within burrows or rock crevices, or under rocks, logs, or surface debris. Many are excellent climbers, and some tropical species are arboreal. *See* ARACHNIDA; ARTHROPODA. Fred Punzo; C. Clayton Hoff

Bibliography. M. Moffett, Wind scorpions, *Nat. Geog.*, 206:94–101, July 2004; F. Punzo, *The Biology of Camel-Spiders (Arachnida, Solifugae)*, Kluwer Academic, Norwell, MA, 1998; F. Punzo, *Desert Arthropods: Life History Variations*, Springer, New York, 2000.

Soliton

An isolated wave that propagates without dispersing its energy over larger and larger regions of space. In most of the scientific literature, the requirement that two solitons emerge unchanged from a collision is also added to the definition; otherwise the disturbance is termed a solitary wave.

There are many equations of mathematical physics which have solutions of the soliton type. Correspondingly, the phenomena which they describe, be it the motion of waves in shallow water or in an ionized plasma, exhibit solitons. The first observation of this kind of wave was made in 1834 by John Scott Russell, who followed on horseback a soliton propagating in the windings of a channel. Scott Russell's account was published in 1845; 50 years later D. J. Korteweg and H. de Vries proposed an equation for the motion of waves in shallow waters which possesses soliton solutions, and thus established a mathematical basis for the study of the phenomenon. Interest in the subject, however, lay dormant for many years, and the major body of investigations began only in the 1950s. Researches done by analytical methods and by numerical methods made possible with the advent of computers gradually led to a complete understanding of solitons. *See* DIFFERENTIAL EQUATION; WAVE MOTION; WAVE MOTION IN LIQUIDS.

Eventually, the fact that solitons exhibit particle-like properties, because the energy is at any instant confined to a limited region of space, received attention, and solitons were proposed as models for elementary particles. However, it is difficult to account for all of the properties of known particles in terms of solitons. More recently it has been realized that some of the quantum fields which are used to describe particles and their interactions also have solutions of the soliton type. The solitons would then appear as additional particles, and may have escaped experimental detection because their masses are much larger than those of known particles. In this context the requirement that solitons emerge unchanged from a collision has been found too restrictive, and particle theorists have used the term soliton where traditionally the term solitary wave would be used. *See* ELEMENTARY PARTICLE; QUANTUM FIELD THEORY.

Sin ϕ or sine-Gordon model. The sin ϕ or sine-Gordon model is formulated in terms of a field $\phi(x,t)$ depending on one spatial coordinate x and the time variable t, and describes one-dimensional propagation of waves. The energy density, defined by Eq. (1), leads to Eq. (2) as the equation of motion,

$$\mathscr{E} = \frac{\alpha}{2}\left(\frac{\partial\phi}{\partial t}\right)^2 + \frac{\beta}{2}\left(\frac{\partial\phi}{\partial x}\right)^2 + \frac{\gamma}{2}\left(1 - \cos\frac{2\pi\phi}{\phi_0}\right) \tag{1}$$

$$\alpha\frac{\partial^2\phi}{\partial t^2} - \beta\frac{\partial^2\phi}{\partial x^2} = -\frac{\pi y}{\phi_0}\sin\frac{2\pi\phi}{\phi_0} \tag{2}$$

which is one of the simplest equations having soliton solutions. Here α, β, γ, and ϕ_0 are positive constant parameters. The first term in the energy density is sensitive to the rate of variation in time of the field at a definite location and becomes larger as this rate increases. The second term depends on the rate of variation of the field along the x-coordinate axis and also becomes larger as this rate of variation increases. Thus, the first two terms in the energy will be minimal (and indeed will equal zero) if the field is constant in space and time. The third term depends on the actual value of the field. When $\phi = 0, \pm\phi_0, \pm 2\phi_0$, or any integer multiple of ϕ_0, $\cos(2\pi\phi/\phi_0)$ [argument expressed in radians; otherwise 2π is replaced by 360°] becomes equal to 1 and the third term in the energy vanishes. When $\phi = \pm(\phi_0/2), \pm(3\phi_0/2), \pm(5\phi_0/2), \ldots$, the third term in the energy is maximal and equal to γ, and for all other values of ϕ it varies between 0 and γ. The state of minimum energy, or state of rest or vacuum state, is therefore achieved when the field takes a value constant in space and time and equal to $0, \pm\phi_0, \pm 2\phi_0, \ldots$. Thus there is not a single vacuum, but many states of minimal energy (multiple vacua).

Solitons occur when the field approaches different vacuum values in the two opposite directions of the x-coordinate axis. Simple and intuitive reasoning shows that if different vacua are reached as $x \to +\infty$ and $x \to -\infty$, a region of transition must exist where the energy density is nonvanishing. If these two vacuum values differ precisely by ϕ_0, the system exhibits a single soliton. The field profile, constant in time, which minimizes the energy gives rise to the static soliton solution (**Fig. 1**). The energy is thus confined to a definite region of space.

A more gradual passage of ϕ between the two vacuum states (which would spread out the energy) would lead to an increase of the contribution of the last term to the energy not compensated by the decrease in the second term. A more abrupt transition

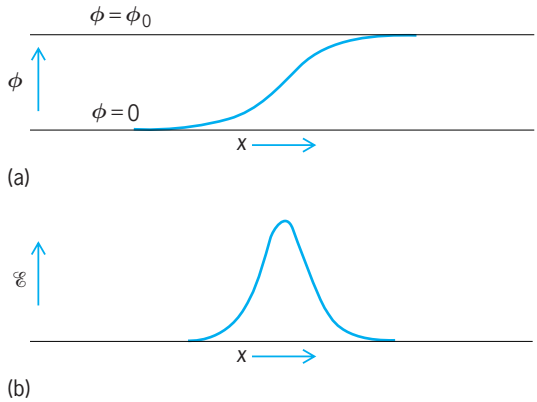

Fig. 1. Static soliton solution of the sin ϕ mode. (*a*) Field ϕ. (*b*) Energy density $\mathscr{E}$.

would also increase the total energy. Configurations in which the field undergoes more than one transition between neighboring vacuum states exhibit multiple solitons.

Solutions with moving solitons also exist. In an encounter of two solitons moving in opposite directions (**Fig. 2**), the two waves emerge from the collision with shape and speed unchanged. This remarkable feature of the solitary waves results from special properties of the sin ϕ equation (and of other equations having soliton solutions) and, as has already been mentioned, is often considered part of the definition of a soliton.

Other systems with solitons. Many equations describing one-dimensional propagation of waves give origin to soliton solutions. The equation of Korteweg and de Vries, the nonlinear Schrödinger equation, and the Boussinesq equation belong to this category. The preservation of shape and velocity of the waves after a collision has been related to the existence of conservation laws. Any isolated dynamical system obeys conservation laws: the total energy of the system and its total momentum, for instance, do not change in time. In a perfectly elastic frontal collision of two billiard balls of the same mass, these two conservation laws are sufficient to predict the outcome of the collision: the two balls just exchange their speeds. The systems exhibiting solitons are characterized by obeying a very large number of conservation laws, well beyond what would be expected of an ordinary dynamical system. The existence of these conservation laws implies, and in turn follows from, the fact that the system is integrable; by rather elaborate analytical techniques, a new set of variables may be found for the description of the motion which evolve in time independently and according to very simple rules. *See* CONSERVATION LAWS (PHYSICS); CONSERVATION OF ENERGY; CONSERVATION OF MOMENTUM; SYMMETRY LAWS (PHYSICS).

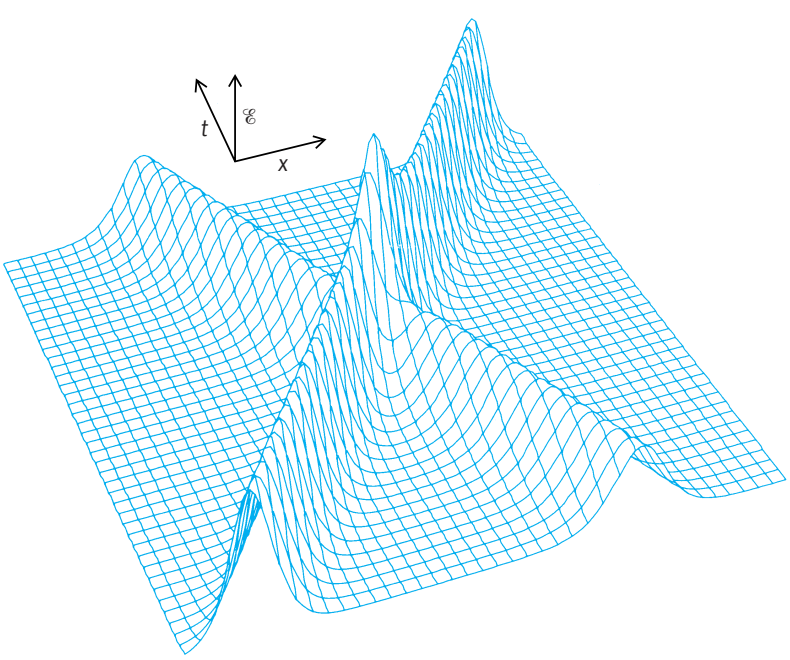

Fig. 2. Space and time variation of the energy density in a collision between two solitons moving in opposite directions. The wave with the higher peak energy is moving faster. The two waves emerge from the collision with shape and speed unchanged.

Derivation of solutions. Equations with soliton solutions have been widely studied; various mathematical methods have been devised to investigate their properties and to derive their solutions. In the method of inverse scattering transforms, an auxiliary problem is set up, which consists of studying the diffusion of a set of plane waves by a potential well. The shape of the potential well varies with time and is related to the profile of the solitons. From the initial soliton configuration, the scattering data, that is, the parameters that totally characterize the diffusion process, can be determined. These scattering data evolve in time according to some simple, soluble equations. Thus, to obtain the soliton profile at a later time, the evolution of the scattering data is followed, and eventually the shape of the diffusing potential is reconstructed from them. It is this last step, where the normal process of deriving the scattering data from a potential is inverted to determine the latter from the former, that gives the name to the method. Other mathematical techniques rely on special sets of transformations, called Bäcklund transformations, which convert given soliton solutions into other soliton solutions, frequently exhibiting a different number of solitons; or rely on embedding the original equations into larger sets of partial differential equations (prolongation schemes), which possess special algebraic symmetries. *See* INVERSE SCATTERING THEORY.

Two- and three-dimensional systems. While one-dimensional solitons are very interesting for a variety of applications, an isolated wave must be three-dimensional to be interpreted as a particle. The existence of systems exhibiting solitary waves in two and three spatial dimensions has been demonstrated. These confined waves are not expected to preserve their dynamical state in a collision (although at present it cannot be excluded that some have this property), but particle theorists have called them solitons as well. In some cases, the confinement of the wave can be related to the existence of multiple states of minimum energy (multiple vacua), as for the sin ϕ model; such waves are called topological solitons because of the special geometrical

properties of the arrangement of vacua around the wave. The topological features are not universal, however, and nontopological solitons exist as well.

Quantum phenomena. A quantum-mechanical formulation of the dynamics of the solitons has been developed, and it has been discovered that topological solitons are often associated with formerly unexpected, remarkable quantum phenomena. Some of the topological solitons are magnetic monopoles; that is, they carry an isolated pole of magnetic charge. Magnetic monopoles cannot be introduced in the ordinary theory of electromagnetism without allowing for some type of singularity in the fields, but no singularity is present in the soliton-monopole solutions, which exhibit everywhere regular field configurations. Topological solitons may violate the normal relation between the statistics of the fields and the spin of the particles. In quantum field theories, fields are classified as bosonic or fermionic, according to the algebraic properties of their quantization formulas. The elementary quantum excitations of bosonic fields are then particles with integer spin (in units of the fundamental constant $\hbar$), such as the photon and π-meson, whereas quantum excitations of fermionic fields are particles with half-integer spin, such as the electron, proton, and neutron. Yet, theories can be devised where all the fields are bosonic but where solitons behave as particles of half-integer spin. Finally, it has also been demonstrated that in some theories with topological solitons a phenomenon of fermion fractionalization occurs. Fermionic particles, which are normally indivisible units, can find themselves split in the presence of a soliton-antisoliton pair, and each soliton itself or each antisoliton carries a fraction of a fermion. In the domain of particle physics, such phenomena have only a theoretical basis, since solitons, as particles, have not been experimentally observed. However, certain solitons have found experimental verification in condensed matter systems. There is, for instance, evidence of fermion fractionalization in connection with the solitons that occur in long, dimerized molecular chains. *See* MAGNETIC MONOPOLES; QUANTUM STATISTICS; SPIN (QUANTUM MECHANICS).

Massive particles. Many field theories used to describe particles have solutions of the soliton type. These solutions, properly quantized, should not be identified with the known atomic, nuclear, or subnuclear particles, which are rather associated with elementary quantum-mechanical excitations of the fields or bound states of such excitations. The solitons, if indeed realized as particles in nature, would have much larger masses, to the point of making their production in present or conceivable future experiments impossible. They might have been formed, however, in the extreme conditions of temperature and energy density in the very early stages of the evolution of the universe, and their experimental detection would have remarkable implications for particle physics and cosmology. *See* COSMOLOGY.

Claudio Rebbi

Hydrodynamic solitons. A hydrodynamic soliton is simply described by the equation of Korteweg and de Vries, which includes a dispersive term and a term to represent nonlinear effects. Easily observed in a wave tank, a bell-shaped solution of this equation balances the effects of dispersion and nonlinearity, and it is this balance that is the essential feature of the soliton phenomenon.

Among the largest and most energetic solitons are those that appear in hydrodynamic systems. Tidal waves in the Firth of Forth were found by Scott Russell to be solitons, as are internal ocean waves and tsunamis. At an even greater level of energy, it has been suggested that the Great Red Spot of the planet Jupiter is a hydrodynamic soliton. *See* JUPITER; OCEAN WAVES; TSUNAMI.

Optical solitons. The most significant technical application of the soliton is as a carrier of digital information along an optical fiber. First observed in 1980, the optical soliton is governed by the nonlinear Schrödinger equation, and again expresses a balance between the effects of optical dispersion and nonlinearity that is due to electric field dependence of the refractive index in the fiber core. If the power is too low, nonlinear effects become negligible, and the information spreads (or disperses) over an ever increasing length of the fiber. At a pulse power level of about 5 milliwatts, however, a robustly stable soliton appears and maintains its size and shape in the presence of disturbing influences.

Because of unavoidable optical losses, it is necessary to amplify the pulse every 25 mi (40 km) or so. This requirement leads to the most important advantage of solitonic data transmission: the soliton retains its standard shape under optical amplification (using distributed Raman amplifiers). A linear pulse, however, cannot be optically reshaped. It must first be converted into an electronic signal for reashaping and amplification and then reconverted into a light pulse for transmission over the next 25-mi (40-km) leg of its journey, greatly reducing the bandwidth of the system. Present designs for data transmission systems based on the optical soliton have a total length of 4500 mi (7300 km) and a data rate of 4×10^9 bits per second at an error rate of 10^{-9}, which is the telecommunications industry standard. *See* OPTICAL COMMUNICATIONS.

Fluxons on long Josephson junctions. A carefully studied soliton system is the transverse electromagnetic (TEM) wave that travels between two strips of superconducting metal separated by an insulating layer thin enough (about 2.5 nanometers) to permit transverse Josephson tunneling (**Fig. 3**). In this case the relevant nonlinear wave equation is the sine-Gordon model given in Eq. (2) with added terms to represent bias current and electrical losses. Since each soliton carries one quantum of magnetic flux, it is also called a fluxon if the magnetic flux points in one direction, and an antifluxon if the flux points in the opposite direction.

A bias current flowing through the Josephson junction (Fig. 3) forces a fluxon to move transversely in one direction and an antifluxon in the other direction. Because a moving fluxon is reflected from the end of a junction waveguide as an antifluxon, steady bias current accelerates the fluxon (or antifluxon) until its speed becomes limited by electrical losses.

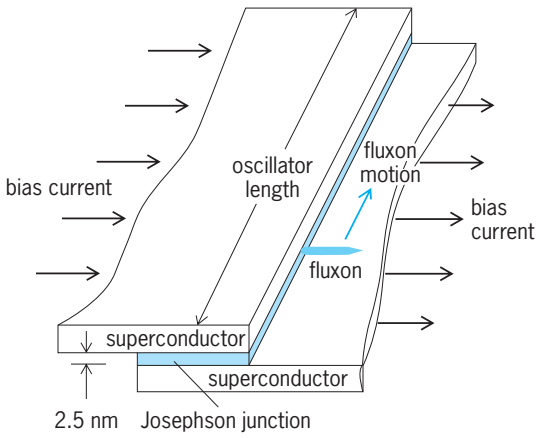

Fig. 3. Fluxon oscillator. Steady bias current flows down through the long Josephson junction and forces a fluxon to move in one direction and an antifluxon to move in the other direction. The fluxon is reflected from the end of the structure as an antifluxon, so the fluxon-antifluxon oscillates back and forth in the long direction.

(The limiting velocity of electromagnetic waves on the waveguide is about $1/\sqrt{20}$ the speed of light.) At this point, the system is a fluxon oscillator with its frequency under the control of the bias current. Since the oscillator length can be less than 100 micrometers, and since the maximum frequency is the limiting velocity divided by twice this length, such oscillators reach into the submillimeter wave region of the electromagnetic spectrum (frequencies greater than 10^{11} Hz). *See* JOSEPHSON EFFECT; WAVEGUIDE.

Biological solitary waves. The all-or-nothing action potential or nerve impulse that carries a bit of biological information along the axon of a nerve cell shares many properties with the soliton. Both are solutions of nonlinear equations that travel with fixed shape at constant speed, but the soliton conserves energy, while the nerve impulse balances the rate at which electrostatic energy is released from the nerve membrane to the rate at which it is consumed by the dissipative effects of circulating ionic currents. The nerve process is much like the flame of a candle. *See* BIOPOTENTIALS AND IONIC CURRENTS.

Theories for true (energy-conserving) solitons have been proposed for a variety of elongated biomolecular structures. These structures include the alpha-helix regions of protein molecules, where solitons may participate in the phenomenon of muscular contraction, carry electronic and protonic charge by polaronic mechanisms, or transport ions across semipermeable membranes; deoxyribonucleic acid (DNA), where solitons may help to open and close the double helix during transcription of the genetic code; and protein crystals of the cytoskeleton, where the solitons may serve as a basis for information processing in single-celled organisms such as the ameba, paramecium, and vorticella. Each of these biological solitons is closely related to one of the lattice solitons that have been observed in condensed matter, and describe the propagation of dislocations, ferroelectric and ferromagnetic domain walls, polarons, charge-density waves, and orientations of chemical bonds in a variety of crystalline structures. *See* CHARGE-DENSITY WAVE; CRYSTAL DEFECTS; CYTOSKELETON; DEOXYRIBONUCLEIC ACID (DNA); DOMAIN (ELECTRICITY AND MAGNETISM); POLARON; PROTEIN. Alwyn Scott

Bibliography. A. Barone and G. Paterno, *Physics and Applications of the Josephson Effect*, 1982; G. Chaohao (ed.), *Soliton Theory and Its Applications*, 1995; A. S. Davydov, *Solitons in Molecular Systems*, 2d ed., 1991; G. L. Lamb, Jr., *Elements of Soliton Theory*, 1980, reprint 1994; A. C. Newell and J. V. Moloney, *Nonlinear Optics*, 1991; C. Rebbi, Solitons, *Sci. Amer.*, 240(2):92–116, 1979; C. Rebbi and G. Soliani (eds.), *Solitons and Particles*, 1984; M. Remoissenet, *Waves Called Solitons: Concepts and Experiments*, 3d ed., 1999.

Solstice

The two days during the year when the Earth is so located in its orbit that the inclination (about $23^1/_2{}^\circ$, or 23.45°) of the polar axis is toward the Sun. This occurs on June 21, called the summer solstice, when the North Pole is tilted toward the Sun; and on December 22, called the winter solstice, when the South Pole is tilted toward the Sun (see **illus.**). The adjectives summer and winter, used above, refer to the Northern Hemisphere; seasons are reversed in the Southern Hemisphere.

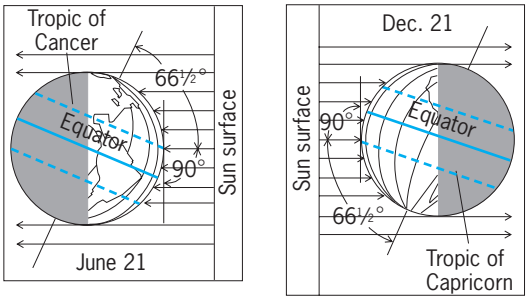

The Earth at the time of the summer and winter solstices. The dates may vary because of the extra one-fourth day in the year.

At the time of the summer solstice the Sun's rays are vertical overhead at the Tropic of Cancer, $23^1/_2{}^\circ$ north. At the North Pole the Sun will then circle $23^1/_2{}^\circ$ above the horizon; and at the Arctic Circle, $66^1/_2{}^\circ$ north, the noon Sun will be 47° above the horizon and the setting Sun will touch the horizon to the north. Thus, on this day every place north of the Arctic Circle will have 24 h of sunlight and the length of day at all places north of the Equator will be more than 12 h, increasing in length with increasing latitude.

Identical conditions are found in the Southern Hemisphere at the time of the Northern Hemisphere's winter solstice when the Sun is vertical above the Tropic of Capricorn, $23^1/_2{}^\circ$ south, and the South Pole is tilted toward the Sun. *See* MATHEMATICAL GEOGRAPHY. Van H. English

Solubility product constant

A special type of simplified equilibrium constant (symbol K_{sp}) defined for, and useful for, equilibria between solids (*s*) and their respective ions in solution, for example, reaction (1). For this relatively

$$AgCl(s) \rightleftharpoons Ag^+ + Cl^- \quad (1)$$

simple equilibrium, Eqs. (2) and (3) apply.

$$[Ag^+][Cl^-] \cong K_{sp} \quad (2)$$

$$(Ag^+)(Cl^-)/(AgCl) = K_{sp} \quad (3)$$

It can be demonstrated experimentally that a small increase in the molar concentration of chloride ion $[Cl^-]$ (produced, for example, by the introduction of NaCl, HCl, or other soluble chloride) causes a reduction in the concentration of silver present as Ag^+. Similarly, an increase in $[Ag^+]$ reduces $[Cl^-]$. The product of the two concentrations is approximately constant as indicated by Eq. (2) and equal to the K_{sp} of Eq. (3). Equation (3) is exact since the variables are activities instead of concentrations. In accordance with the choice of standard state usually made for a solid, the activity of solid AgCl is unity; hence Eq. (4) holds.

$$(Ag^+)(Cl^-) = K_{sp} = 1.8 \times 10^{-10} \text{ mole}^2 \text{ liter}^{-2} \quad (4)$$

In practice, various complications arise: addition of too much of either ion produces more complicated ions and hence actually increases the apparent concentration of the other ion. Addition of a salt without a common ion (that is, a salt supplying neither Ag^+ nor Cl^-) either may react with Ag^+ or Cl^- or may merely increase the concentration of both ions by a lowering of the mean ionic activity coefficient. (Sodium nitrate at a concentration 0.01 molar increases each concentration by about 10% and the product by 20%.)

It is usually assumed that an aqueous solution saturated with silver chloride contains only Ag^+, Cl^-, and the solvent. Some recent work indicates, however, that about 2.5% of the solute is present as undissociated AgCl. In practice, such effects are usually neglected. Equation (2) is especially useful in the explanation of analytical procedures in which it is desired to add a sufficient quantity of one ion to ensure (virtually) complete precipitation of the other.

An example of a salt of a different charge type is lead iodate $Pb(IO_3)_2$. The solubility equilibrium is represented by reaction (5).

$$Pb(IO_3)_2(s) \rightleftharpoons Pb^{2-} + 2IO_3^- \quad (5)$$

At about 77°F (25°C), the solubility of this salt is 0.024 g/liter. The mass of 1 gram-formula weight of the salt is 557 g. In the saturated solution, therefore, the concentrations are as in notation (6). Because

$$\begin{aligned} [Pb^{2+}] &= \frac{0.024}{557} = 4.3 \times 10^{-5} \text{ mole/liter} \\ [IO_3^-] &= \frac{2(0.024)}{557} = 8.6 \times 10^{-5} \text{ mole/liter} \end{aligned} \quad (6)$$

the solid is in its standard state, its activity is unity. Hence Eq. (7) applies.

$$[Pb^{2+}][IO_3^-]^2 = K_{sp} = 3.2 \times 10^{-13} \text{ mole}^3/\text{liter}^3 \quad (7)$$

It is important to note that the concentration, 8.6×10^{-5} mole/liter, is the concentration of IO_3^- ion in the solution; it is not twice the concentration of the IO_3^- ion, although it happens to be twice the concentration of Pb^{2+}. It is also important to observe that the square of 8.6×10^{-5} mole/liter enters this product because the coefficient of IO_3^- in reaction (5) is 2. *See* GRAVIMETRIC ANALYSIS; IONIC EQUILIBRIUM; PRECIPITATION (CHEMISTRY). Thomas F. Young

Solubilizing of samples

The process by which difficultly soluble samples are converted into different chemical compounds which are soluble. The sample may be heated in air to evolve volatile components or to oxidize a component to a volatile higher oxidation state with the formation of an acid-soluble form, as in the roasting of a sulfide to form the oxide and sulfur dioxide. Most frequently, the sample is treated with a solvent which reacts with one or more constituents of the sample. The dissolution of organic materials often follows the generalizations suggested by the solubility-based schemes of qualitative organic analysis.

Solvents. The choice of solvent is determined by the chemical reactions which are required. Reactions used include solvation, neutralization, complex formation, metathesis, displacement, oxidation-reduction, or combinations of these. Most water-soluble salts dissolve by solvation. Basic oxides such as ferric oxide dissolve in aqueous hydrochloric acid by neutralization followed by complex formation. This is shown by reaction (1). Amphoteric oxides

$$Fe_2O_3 + 8H^+ + 8Cl^- \rightarrow 2FeCl_4^- + 2H^+ + 3H_2O \quad (1)$$

dissolve in either acids or bases. Reaction (2) rep-

$$Al_2O_3 + 2Na^+ + 2OH^- \rightarrow 2AlO_2^- + 2Na^+ + H_2O \quad (2)$$

resents the dissolving in a base. Aqueous ammonia converts insoluble silver chloride into a soluble complex as shown by reaction (3). Carbonates are solubi-

$$AgCl + 2NH_4OH \rightarrow Ag(NH_3)_2^+ + Cl^- + 2H_2O \quad (3)$$

lized by treatment with hydrochloric acid to displace carbon dioxide and to form soluble chlorides. Reaction (4) represents this reaction. Many substances are

$$CaCO_3 + 2H^+ + 2Cl^- \rightarrow Ca^{2+} + 2Cl^- + H_2O + CO_2 \quad (4)$$

converted to easily dissolved mixtures by metathesis. An example is the boiling of barium sulfate with aqueous sodium carbonate to form barium carbonate and sodium sulfate. The insoluble barium carbonate dissolves easily in hydrochloric acid.

Insoluble metal salts have been transposed to soluble ethylenediaminetetraacetic acid complexes.

Insoluble phosphates, sulfates, and even silicates have also been transposed by batch digestions with cationic ion-exchange resins as shown by reaction (5).

$$MSO_4 + 2HR \rightarrow MR_2 + H_2SO_4 \qquad (5)$$

The metal which is bound to the resin can then be eluted or displaced into solution by treatment with an excess of an indifferent salt solution or with an acid whose anion is compatible with the cation present.

Dissolution of metals and alloys. Metals above hydrogen in the electrochemical series will dissolve in a nonoxidizing acid by reduction of hydrogen ion. Other metals such as copper and lead require an oxidizing acid, usually nitric acid. The treatment of alloys is determined by the constituents present. All of the components of brass and bronze are usually dissolved by nitric acid except tin, arsenic, and antimony, which precipitate as hydrated oxides. Alloy steels are usually dissolved by combinations of hydrochloric, nitric, phosphoric, and hydrofluoric acids, depending on the elements present. Aluminum-base alloys are treated with sodium hydroxide solution and any residues are dissolved in acid. Either total or component-selective anodic dissolution of metals and alloys may be achieved by applying a suitable anodic potential to the massive sample, weighed before and after attack, which is immersed in an electrolyte bath.

Fusions. Many substances do not dissolve at temperatures obtainable in the presence of liquid water. However, many refractory silicates, strongly ignited oxides of beryllium, cerium, plutonium, and so on, are completely attacked by aqueous hydrochloric acid at 300°C (570°F). A sealed quartz-tube procedure was developed which permits this widely useful type of sample attack.

However, fused salt reactions employing temperatures of 400–1100°C (750–2000°F) are necessary for the attack and decomposition of many types of samples. The material used as the solvent is called a flux, and the process of melting the mixture of dry, solid flux with the sample is called a fusion. Fusion is done in a crucible (usually platinum or nickel) which is not attacked by the flux or the sample constituents. The same types of reactions are used as with aqueous solvents. Sodium carbonate is used to attack acid materials such as silicates and for metathesis reactions with sulfates, as in reaction (6). Potassium bisulfate

$$CaSiO_3 + Na_2CO_3 \rightarrow CaCO_3 + Na_2SiO_3 \qquad (6)$$

on fusion yields potassium pyrosulfate, an acid flux, which attacks basic oxides such as alumina and ferric oxide and metals such as chromium and tungsten. This flux must be used in porcelain crucibles. Oxidizing fluxes such as sodium peroxide and mixtures of sodium carbonate with potassium nitrate are used with sulfides, chromium, and tin ores and some silicon samples. These usually require nickel crucibles. Calcium carbonate plus ammonium chloride is used to free the alkali metals from silicates. *See* ANALYTICAL CHEMISTRY.

Charles L. Rulfs

Bibliography. W. F. Hillebrand and E. F. Lundell, *Applied Inorganic Analysis*: *With Special Reference to the Analysis of Metals*, *Minerals, and Rocks*, 1980; D. A. Skoog, *Analytical Chemistry: An Introduction*, 7th ed., 2001.

Solution

A homogeneous mixture of two or more components whose properties vary continuously with varying proportions of the components. A liquid solution can be distinguished experimentally from a pure liquid by the fact that during transfers into other single phases at equilibrium (freezing and vaporizing at constant pressure) the temperature and other properties vary continuously, whereas those of a pure liquid remain constant. For an apparent exception *see* AZEOTROPIC MIXTURE; SOLVENT.

Gases, unless highly compressed, are mutually soluble in all proportions.

A solid solution is, similarly, a single phase whose composition and other properties vary continuously with changing composition of the liquid phase with which it is in equilibrium. *See* SOLID SOLUTION.

Types of intermolecular force. The extent to which substances can form solutions depends upon the kind and strength of the attractive forces between the several molecular species involved. It is necessary to consider the attractive forces exerted by molecules of the following types: (1) nonpolar molecules; (2) polar molecules, that is, those containing electric dipoles; (3) ions; and (4) metallic atoms.

London forces. The theory of attraction between nonpolar molecules, developed by F. London in 1930, is based upon the quantum-mechanical interaction between pairs of electron systems. For two molecules with electrons having frequencies ν_1 and ν_2, polarizabilities α_1 and α_2, and separated by the distance r between centers, the attraction potential is shown as Eq. (1), where h is the Planck constant. For molecules of the same species, this reduces to Eq. (2). The frequency ν_0 is that corresponding to $h\nu_0$, the zero-point energy, of the molecule in its unperturbed state. The perturbation by another molecule is related to its perturbation by light of varying frequencies, as seen in the variation of refractive index n with the frequency of light, that is, the dispersion. For this reason London designated these forces as dispersion forces. It is equally appropriate to speak of London forces by analogy with the nearly equivalent term, van der Waals forces. In the case of gases the dispersion n_ν is related to the frequency as in

$$\epsilon_{12} = -\frac{3\alpha_1\alpha_2 h}{2r^6} \cdot \frac{\nu_{0,1}\nu_{0,2}}{\nu_{0,1} + \nu_{0,2}} \qquad (1)$$

$$\epsilon_{11} = -\frac{3\alpha_1^2}{4r^6} h\nu_0 \qquad (2)$$

Eq. (3), where C is a constant. The polarizability α

$$n_\nu - 1 = \frac{C}{\nu_0^2 - \nu^2} \tag{3}$$

can be determined from the refractive index with the aid of the Lorentz-Lorenz formula. As a substitute for zero-point energy, London proposed the ionization potential.

The model upon which these relations are based is much simpler than the polyatomic molecules in most solutions of interest. In these the potential field is not central and radial; the interaction is between the electrons in the peripheral bonds. A striking example is octamethylcyclotetrasiloxane, whose core of alternating silicon and oxygen atoms is so buried within the eight methyl groups that it behaves toward other molecules essentially as an aliphatic hydrocarbon. The normal paraffins themselves are not symmetrical spherically. Moreover, the electrons in the molecules have many frequencies.

Although attempts have been made to extend London's basic concept to take account of such complexities, only the general implications of the concept as to the characteristics of the London forces are necessary here. These forces are (1) of very short range; (2) additive and nonspecific; (3) temperature independent; (4) operative between molecules of all types, whether nonpolar, polar, or metallic; (5) dependent in magnitude upon the number and "looseness" of the electrons; and (6) ordinarily less than average between molecules of different species. This last property can be seen by comparing ϵ_{12} with ϵ_{11} and ϵ_{22} as given by Eqs. (1) and (2). Eliminating the α terms one obtains Eq. (4). Most component pairs

$$\epsilon_{12} = \frac{(\nu_1\nu_2)^{1/2}}{(\nu_1 + \nu_2)/2} \cdot (\epsilon_{11}\epsilon_{22})^{1/2} \tag{4}$$

differ much less in ionization potential than in polarizability, and the factor representing frequencies in this expression is not far from unity. Thus, relation (5) can be written. This means that the interaction

$$\epsilon_{12} \approx (\epsilon_{11}\epsilon_{22})^{1/2} \tag{5}$$

potential between unlike molecules is less than the arithmetic mean of the like potentials.

The pair potentials can be integrated over all the molecules in the pure components as well as the solution to obtain approximate attraction constants, a's, corresponding to the attraction constant of the van der Waals equation. M. P. E. Berthelot proposed the relation $a_{12} = (a_{11}a_{22})^{1/2}$ between the attraction constants of like and unlike species. J. Hildebrand and H. M. Carter found the Berthelot relation to be valid within 1% for seven liquid pairs, for example, $a = 31.21$ for liquid CCl_4 and 64.79 for $SnBr_4$. The calculated geometric mean is 46.46, and a_{12}, observed, is 46.86.

The consequence of this geometric mean relation is that the cohesion in a mixture of two liquids having different cohesion is less than their average. This results in expansion in volume, absorption of heat, and vapor pressures greater than additive upon mixing.

This geometric mean relation is usually adhered to rather well in cases of unlike molecules whose outer electrons are of similar types, such as (1) "N-electrons," nonbonded, as in halogenated paraffins and halogens; (2) π electrons, as in olefins and aromatics; (3) bonding electrons only, as in H_2, CH_4, and other aliphatics; and (4) fluorochemicals. But deviations are found between molecules whose outer electrons are of different types. Illustrations will be found below in the sections on regular solutions and on solubility of gases.

Dipole interaction. This is the attraction between molecules containing permanent electric dipoles; it includes both the London forces and an electrostatic interaction of the dipoles. The latter depends upon the dipole moments of the molecules; it is temperature dependent because thermal agitation opposes the antiparallel orientation in which the interaction is greatest. Its magnitude depends also upon the geometry of the molecules because it is related to the distance of approach of the dipoles, not the molecular centers; the dipoles of some molecules are buried more deeply than those of others. This is the case with chloroform, which has solvent properties similar to those of carbon tetrachloride, except in a few specific cases. J. G. Kirkwood has expressed the degree of interaction between the dipoles of pure liquids by a g factor. For pyridine, the dipole moment μ is 2.20×10^{-18} cgs units, the dielectric constant ϵ is 12.5, and the dipole interaction g is 0.9. For water, $\mu = 1.84 \times 10^{18}$, $\epsilon = 78.5$, and $g = 2.7$; for ethyl alcohol, $\mu = 2.80 \times 10^{-18}$, $\epsilon = 24.6$, and $g = 3.0$.

It is the g factor, not the dipole moment or the dielectric constant, that is most significant for understanding solubility relations. Furthermore, in the case of molecules having more than one polar bond, it is the separate polar bonds, not their vector sum of the overall dipole moment, that determine solubility relations. The three isomeric dinitrobenzenes all affect the vapor pressure of benzene virtually to the same extent, even though their dipole moments are quite different.

The substances with the largest g factors are those that form hydrogen bonds. These have exceptionally high boiling points and are poor solvents for nonpolar substances. These liquids resist penetration by nonpolar molecules. The best-known pairs of incompletely miscible liquids are composed of a nonpolar liquid and water.

Electron donor-acceptor interaction. In the modern theory of generalized acids and bases, initiated by G. N. Lewis, a base is a substance having electrons that may be "accepted" into the vacant orbitals of other molecules, termed acids. This acceptance of electrons takes place reversibly and with little or no activation energy. Typical bases or donors are pyridine, acetone, ether, alkyl bromides, alkyl iodides, alkyl sulfides, iodide ion, thiocyanate ion, and aromatic hydrocarbons. Typical acids are the pure and mixed halogens, sulfur dioxide and trioxide, boron trichloride and trifluoride, aluminum halides, and stannic chloride. R. S. Mulliken and his co-workers have pointed out the close relationship between base

strength and ionization potential and elaborated a theory of charge transfer complexes. H. A. Benesi and Hildebrand discovered the strong absorption in the ultraviolet characteristic of such complexes. They found that the basic strength increases in the order benzene < toluene < xylene < mesitylene. R. L. Scott determined that acid strength increases in the order $Cl_2 < Br_2 < I_2 < BrI < BrCl < ICl$.

This type of interaction is specific and saturating, and it reduces the escaping tendencies of the components. It corresponds to $\epsilon_{12} > (\epsilon_{11}\epsilon_{22})^{1/2}$. In cases where it is weak it may reduce but not overcome the opposite effect of unequal London forces.

Ion-ion interaction. The ions in a solid or liquid salt attract and repel electrostatically according to Coulomb's law, but there is also a London force component, and large ions are polarized by smaller ones. This last effect is illustrated by solid silver bromide, which is colored although its ions in aqueous solutions are colorless and whose crystals have the sodium chloride structure. The evidence is that in both the solid and the fused salt the electron cloud of the bromide ion is distorted equally by each of its six neighboring silver ions.

Ion-dipole interaction. In order to dissolve a solid salt, its lattice energy must be supplanted by the ion-ion action of another salt already in the liquid state or by the predominantly electrostatic attraction of a polar solvent or by the specific chemical interaction represented by complex ions.

Ideal solution. It is profitable to deal with actual solutions in terms of their departure from a simple idealized model—a mixture of components having the same attractive fields, which mix without change in volume or heat content. This is analogous to an ideal gas mixture, which is formed with no heat of mixing and in which the total pressure is the sum of the partial pressures. In such a solution the escaping tendency of the individual molecules is the same, whether they are surrounded by similar or by different molecules. Therefore, the combined escaping tendency of all the molecules of species 1, f_1, is given by Eq. (6), where x_1 is the mole fraction of species 1 and f_1^0 is the escaping tendency of the molecules from the pure liquid. For a binary mixture, $x_1 + x_2 = 1$. If the gas imperfections of vapors are disregarded, vapor pressures may be substituted for fugacities to give Raoult's law (1886) in its usual form, $p_1 = p_1^0 x_1$. The total pressure is $P = p_1 + p_2 = p_1^0 x_1 + p_2^0 x_2$.

$$f_1 = f_1^0 x_1 \tag{6}$$

A more sophisticated derivation than the foregoing requires one to postulate molecules of the same size and shape. The gross structure of the solution containing n_1 plus n_2 molecules of two components is identical with that of the pure liquid components. The number of configurations of the components in the mixture within this structure is $(n_1 + n_2)!/n_1!n_2!$, and the configurational entropy of mixing is, by Stirling's formula, Eq. (7). The substitution of a number of moles of each, N_1 and N_2, gives Eq. (8). The partial derivative $(\partial \Delta S/\partial N_1)_{N_2}$ represents the partial molal entropy of transfer of component 1 from pure liquid into an ideal solution of mole fraction x_1, Eq. (9). Because the model postulates no change in enthalpy,Eq. (10) applies, and $f_1/f_1^0 = x_1$, which is Raoult's law, where f is fugacity. *See* FUGACITY.

$$\Delta S = k\left[n_1 \ln \frac{n_1 + n_2}{n_1} + n_2 \ln \frac{n_1 + n_2}{n_2}\right] \tag{7}$$

$$\Delta S = R\left[N_1 \ln \frac{N_1 + N_2}{N_1} + N_2 \ln \frac{N_1 + N_2}{N_2}\right] \tag{8}$$

$$\overline{S}_1 - S_1^0 = -R \ln x_1 \tag{9}$$

$$T(\overline{S}_1 - S_1^0) = -RT \ln \frac{f_1}{f_1^0} \tag{10}$$

But to arrive at this conclusion one must assume identical structures in the solution and the pure liquid components. This is very far from the case in solutions of high polymers in ordinary solvents, even though, as with polystyrene in benzene, the heat of mixing is practically zero.

Moderate difference in molal volume between components of high symmetry has little effect, as might be expected from the fact that the radius varies only with the cube root of the volume.

As a foundation for dealing with actual solutions in terms of the deviations of their properties from those of the model, it is necessary to derive other equivalent thermodynamic relationships.

Solubility of a crystalline solid. The fugacity of a crystalline substance f^s at temperature T is less than that of its supercooled liquid f_0 to an extent depending upon its melting point T_m and heat of fusion ΔH^F, as given by Eq. (11). If ΔH^F is assumed constant, this gives upon integration Eq. (12). If the heat capacities of the solid and liquid forms are known, the variation of ΔH^F with temperature can be taken into account. This is hardly necessary for the present purpose because the deviations from ideal solubility involve factors that are more uncertain than this.

$$\frac{d \ln(f^s/f^0)}{dT} = \frac{\Delta H^F}{RT^2} \tag{11}$$

$$\ln \frac{f^s}{f^0} = -\frac{\Delta H^F}{R}\left(\frac{1}{T} - \frac{1}{T_m}\right) \tag{12}$$

Molecular weight measurements. If a solid dissolves to form an ideal solution, its heat of solution is the same as its heat of fusion ΔH^F and $f_1^s/f_1^0 = x_1$. Therefore, Eq. (13) is formed. This is the approximate equation for solubility of a solid that forms an ideal solution. It can be transformed into one much used for determining the molal weight of a solute by the depression of the freezing point of the solvent, here

$$-\ln x_1 = \frac{\Delta H^F}{R}\left(\frac{T_m - T}{T_m T}\right) \tag{13}$$

component 1. For $-\ln x_1$, one can write $\ln(1 + N_2/N_1)$. Expanding in powers of N_2/N_1 gives Eq. (14).

$$\ln\left(1+\frac{N_2}{N_1}\right) = \frac{N_2}{N_1}\left[1-\frac{1}{2}\frac{N_2}{N_1}+\frac{1}{3}\left(\frac{N_2}{N_1}\right)^2-\cdots\right] \quad (14)$$

When $N_2 \ll N_1$, the higher powers may be neglected to give Eq. (15), where $\Delta T = T_m - T$. By measuring ΔT for a known weight of solute in N_1 moles of solvent, the molal weight of the solute can be calculated. Because of the approximations made and the fact that even good solvents for the solid in question are seldom ideal, the resulting molal weights are not very exact unless extrapolated to $x_2 = 0$ from a series of values.

$$\frac{N_2}{N_1} \approx \frac{\Delta H_1^F}{RT_m^2}\Delta \quad (15)$$

The lowering of the vapor pressure of a solvent upon the addition of a nonvolatile solute (component 2) may be offset by raising the temperature to restore the pressure of the solvent. These changes are related as in Eq. (16). This relation is far less useful than that for the freezing point depression because the heat of vaporization is much greater than the heat of fusion; therefore, the elevation of the boiling point is much smaller than the depression of the freezing point and also is harder to determine.

$$x_2 = \frac{\Delta H_1^{\text{vap}}}{RT_b^2}\Delta T \quad (16)$$

Osmotic pressure. One mole of a solvent can be removed from a large quantity of a solution in which its mole fraction is x_1 in two reversible, and hence equivalent, ways. If it is distilled from the solution into pure liquid, the gain in (Gibbs) free energy is $\Delta G_1 = RT \ln (f_1^0/f_1)$. If it is pressed out through a semipermeable membrane against the hydrostatic pressure difference, osmotic pressure ΔP, the gain in free energy is $\Delta P\overline{V}_1$, where $\overline{V}_1$ is the partial molal volume of the solvent. In an ideal solution this is the molal volume. Equating the free energy of these processes gives Eq. (17). Expanding as before and neglecting the higher powers gives $\Delta P\overline{V}_1 \approx (N_2/N_1)RT$ or $PV = RT$, where $V = N_2\overline{V}_1/N_1$, the volume containing 1 mole of solute.

$$\Delta P\overline{V}_1 = RT\ln\frac{f_1^0}{f_1} = RT\ln\frac{N_1+N_2}{N_1} \quad (17)$$

This is the van't Hoff law for osmotic pressure, put forth in 1887. The theoretical basis of Raoult's law, discovered at almost the same time, was not yet appreciated. The formal correspondence between the van't Hoff law and the perfect gas law seemed to lend unique significance to osmotic pressure and elevated the van't Hoff law to the status of an ideal solution law. It is a limiting law only and not valid at high concentrations; it neglects the specific nature of the solvent. The solvent is regarded as furnishing space for a quasi-gas solute. Thus the law cannot cover molecular states in solutions of finite concentration.

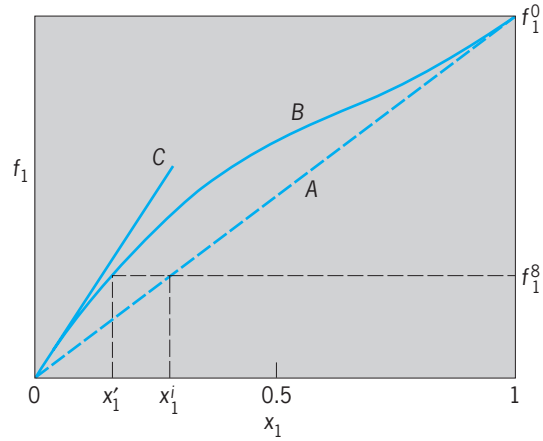

Fugacity and mole fraction. Line *A*, ideal. Line *B*, typical nonideal. Line *C*, Henry's law.

The determination of osmotic pressure offers a valuable means for determining molal weights of high polymers in solution, where high weight concentrations correspond to mole fractions so low as to have only minute effects upon the vapor pressure and the freezing point of the solvent. For example, consider a solution of 0.001 mole of solute in 1 mole of benzene; $T_m = 279$ K and $\Delta H^F{}_1 = 2370$ cal/mole, ΔT by Eq. (15) would be only 0.065°, but ΔP by Eq. (17) would be 194 mm. The latter is large enough to be measured with some precision.

Nonideal solutions. Unlikeness of the components of a binary mixture leads, as explained earlier, to fugacities in excess of ideal values. The excess is largest when the molecules of one species are surrounded mostly by those of the other, as shown in the **illustration**. They approach Raoult's law at the upper end and Henry's law, $p_1 = kx_1$ (or $f_1 = k_1x_1$), at the lower end, where k is an experimentally determined constant.

An important relation between the two components is given by the Gibbs-Duhem equation (18).

$$\left(\frac{\partial \ln f_1}{\partial \ln x_1}\right)_T = \left(\frac{\partial \ln f_2}{\partial \ln x_2}\right)_T \quad (18)$$

If Raoult's law holds in the limit when $x_1 = 1$, since $(\partial \ln f_1)/(\partial \ln x_1) = 1$, then also $(\partial \ln f_2)/(\partial \ln x_2) = 1$. Integrating gives $\ln f_2 = \ln x_2 + \ln k_2$ or $f_2 = kx_2$, where k is a constant of integration that cannot be evaluated unless Raoult's law holds for component 1 over the whole range. In that case it also holds for component 2.

The activity in the case of nonelectrolytes is defined as $a_1 = f_1/f_1^0$ and so on. In an ideal solution, $a_1 = x_1$ and $a_2 = x_2$. The activity coefficient is $\gamma_1 = a_1/x_1$ and so on. Alternate, equivalent forms of the Gibbs-Duhem equation include Eq. (19), and

$$\frac{\partial \ln a_1}{\partial \ln x_1} = \frac{\partial \ln a_2}{\partial \ln x_2} \quad (19)$$

$N_1 d\overline{G}_1 + N_2 d\overline{G}_2 = 0$, where $\overline{G}$ denotes partial molal Gibbs free energies.

If one component is a crystalline solid, its activity a_s is less than that of the liquid, which is 1, as given by Eq. (12); its maximum solubilities would be x^i_1 in the illustration, if in an ideal solvent, and x'_1, if in a real solution.

Regular solutions. There are many mixtures of nonpolar components in which, except in the immediate neighborhood of a critical mixing temperature, thermal agitation suffices to neutralize tendencies to segregate and yields virtually complete randomness of mixing, with a close approach to ideal entropy of mixing, Eqs. (8) and (9).

The enthalpy of mixing can be calculated as the difference between the potential energy of the mixture and the sum of the potential energies of the liquid components. The potential energy of a mole of liquid may be related to the potential between a pair of molecules $\epsilon(r)$. The lattice energy of a crystal is obtained by summation of $\epsilon(r)$ over all of the lattice distances; that of a liquid is obtained by integration over the continuous distribution function $\rho(r)$. The expression for a pure liquid is Eq. (20).

$$\Delta E_{\text{vap}} = -\frac{2\pi N_{Av}^2}{V} \int \rho(r)\epsilon(r) r^2 dr \qquad (20)$$

Here N_{Av} is the Avogadro number and V is the molal volume of the liquid. The corresponding expression for the potential of the mixture of N_1 and N_2 moles of the pure components involves their relative sizes. With certain simplifying assumptions, including the geometric mean for ϵ_{12} (r), Eq. (21*a*) was

$$\Delta E_M = \frac{N_1 V_1 N_2 V_2}{N_1 V_1 + N_2 V_2} \times \left[\left(\frac{\Delta E_1}{V_1} \right)^{1/2} - \left(\frac{\Delta E_2}{V_2} \right)^{1/2} \right]^2 \qquad (21a)$$

$$\overline{E} - E_2^0 = V_2 \phi_1^2 (\delta_1 - \delta_2)^2 \qquad (21b)$$

obtained for the energy of mixing N_1 and N_2 moles of two nonpolar liquids. The corresponding partial molal energy of transferring pure liquid to solution, for component 2, is Eq. (21*b*). Here ϕ_1 denotes volume fraction, neglecting expansion on mixing, and the δ's are $(\Delta E_{\text{vap}}/V)^{1/2}$, designated solubility parameters. Because energy and enthalpy are virtually identical for liquids, Eq. (21*b*) may be combined with the entropy of transfer as given by Eq. (9) to give the free energy of transfer, as in Eqs. (22*a*) and (22*b*).

$$\begin{aligned} \overline{G}_2 - G_2^0 &= \overline{H}_2 - H_2^0 - T(\overline{S}_2 - S_2^0) \\ RT \ln a_2 &= V_2 \phi_1^2 (\delta_1 - \delta_2)^2 + RT \ln x_2 \end{aligned} \qquad (22a)$$

$$\text{or} \quad RT \ln \gamma_2 = V_2 \phi_1^2 (\delta_1 - \delta_2)^2 \qquad (22b)$$

The quantity in square brackets in Eq. (21*a*) is the excess of the arithmetic mean, $^1/_2[(\Delta E_1/V_1) + (\Delta E_2/V_2)]$, over the geometric mean, $[(\Delta E_1/V_1) \times (\Delta E_2/V_2)]^{1/2}$. As mentioned earlier, the geometric mean assumption is amazingly valid for many species of similar electronic structure but fails for others.

Solubility parameters and molal volumes, 77°F (25°C)

Liquid	Formula	Molal volume, ml	Solubility parameter
Perfluoroheptane	C_7F_{16}	225	5.9
Perfluorotributylamine	$(C_4F_9)_3N$	360	6.0
Perfluoromethylcyclohexane	c-$C_6F_{11}CF_3$	196	6.1
n-Heptane	n-C_7H_{16}	147	7.4
Silicon tetrachloride	$SiCl_4$	115	7.6
Cyclohexane	C_4H_{12}	109	8.2
Carbon tetrachloride	CCl_4	97	8.6
Chloroform	$CHCl_3$	81	9.3
Benzene	C_6H_6	89	9.2
Carbon disulfide	CS_2	60	10.0
Bromine	Br_2	52	11.5
Iodine	I_2	59	14.1

Representative values of solubility parameters at 77°F (25°C) are given in the **table**. Parameters for substances solid at 77°F (25°C) have been calculated by Eq. (14*a*) from solubilities of these substances in solvents whose parameters are well determined. With paraffins the solubility data give concordant δ-values a little greater than $(\Delta E_{\text{vap}}/V)^{1/2}$.

Joel H. Hildebrand

Bibliography. *Annual Review of Physical Chemistry*, annually; E. A. Boucher, *Properties of Liquids and Solutions*, 2d ed., 1994; R. G. Mortimer, *Physical Chemistry*, 2d ed., 2000; J. H. Noggle, *Physical Chemistry*, 3d ed., 1997; O. Popovych and R. P. Tompkins, *Nonaqueous Solution Chemistry*, 1981, reprint 1990.

Solution mining

The extraction of valuable components from a mineral deposit using an aqueous leaching solution. In its original sense, solution mining refers to evaporite mining, the dissolution of soluble rock material such as salt by using borehole wells to pump water into the deposit and remove the resulting saturated brine. In its current usage, solution mining also includes ore leaching, the in-place (in-situ) leaching of valuable metal components from an orebody, and the mine-site procedures of heap leaching and dump leaching. Often included is the Frasch process of using superheated water to melt sulfur in its deep deposits and recover the molten sulfur in borehole wells. *See* ORE AND MINERAL DEPOSITS.

Evaporite Mining

Evaporites represent a broad class of water-soluble minerals (salts). Commercially important evaporites include halite (sodium chloride), sylvite and silvinite (potash), bischofite (magnesium chloride), nahcolite (sodium bicarbonate), trona (raw soda ash), and langbeinite and carnallite [both sources of potash and magnesia (magnesium oxide)]. Most of these

evaporite deposits are being mined in North America and Europe. *See* SALINE EVAPORITES.

The geologic setting (shape, attitude, and extent) and nature of each evaporite resource affect the mining method used. Sometimes conventional mechanical mining is more economical. In the Wyoming trona operations, either underground room-and-pillar or solution mining techniques are used, depending on localized conditions. Alternatively, natural brines may be processed directly for their salts; for example, soda ash is recovered from subsurface brines at Searles Lake, California.

Solution mining involves injecting a solvent into the pay zone of the deposit through a cased borehole. For evaporites, the solvent is hot water, which forms brine as the soluble minerals dissolve. The brine is brought to the surface via the casing system in the same or another borehole and sent to a processing facility for recovery by the controlled crystallization of the desired product, followed by dewatering and drying. Some minerals may require additional purification steps, such as the flotation of potash crystals. The depleted brine is chemically reconditioned and injected back into the deposit. Often a bleed stream, typically only 5-10% of the total brine flow, is withdrawn and treated to control the buildup of deleterious impurities in the solvent. Any impurity products, usually small in volume, woud be handled in the appropriate manner. Thus, solution mining creates minimal surface disturbance and little waste, compared to conventional mining. *See* WELL.

Boreholes are drilled to form a grid across the mining zone, and the injection pumping systems are located on the surface. Normally the injection pressure is sufficient to lift brine to the surface without submersible pumps in the recovery bores. To provide constant feed to the process plant, new wells are drilled periodically and old ones taken off-line and plugged as areas are leached out, until the entire evaporite bed is depleted.

Although solution mining is simple in principle, there are several key issues for successful operation. One is to maintain close control over the solvent parameters, such as pH, oxidation potential, temperature, and pressure. For example, a pressure loss would allow dissolved gas to escape, shifting the solubilities of the remaining constituents. *See* ELECTROCHEMICAL SERIES.

Another issue is isolation of the solution mining zone from the surrounding geologic structures. This requires effective well completion techniques that are compatible with the brine. Once the well is drilled and cased, the casing must be cemented into the formation. This prevents brine from migrating along the annular space outside the casing and

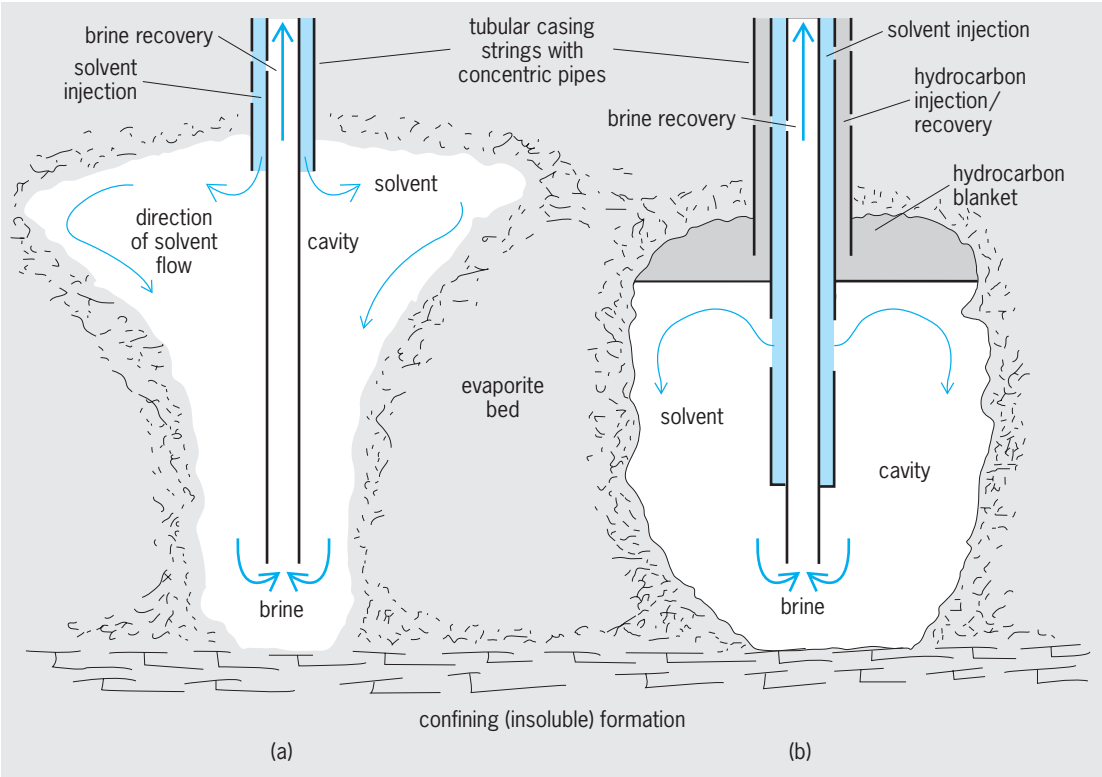

Fig. 1. Solution mining of a subsurface evaporite deposit. Mining is normally initiated at the bottom of the deposit and retreats vertically, leaving a mined-out cavity. Saturated brine is heavier (denser), and is recovered from the bottom of the cavity. (*a*) Early single-well solution mining system with uncontrolled solvent flow. Unsaturated solution (low-density) is injected at the top of the cavity, where most of the dissolution occurs, giving rise to a "morning glory"–shaped cavity. Roof span will continue to increase until it becomes unstable and collapses under the load of the overlying material. (*b*) Single-well system with hydrocarbon blanket. The blanket prevents excessive dissolution of salt at the top of the cavity. The hydrocarbon can be withdrawn periodically to allow for the controlled increase in cavity height.

contaminating adjacent aquifers. *See* OIL AND GAS WELL COMPLETION.

Proper brine flow patterns within the cavity control the shape of the cavity so that it is stable and not subject to excessive stresses, which might cause a ceiling collapse. An inert material (oil or gas) is often injected to form a blanket at the top of the cavity, retarding dissolution of the ceiling (**Fig. 1**).

Controlling the stability of solution mined cavities is particularly important in salt domes along the Gulf Coast. If the mined-out cavities are stable, they have great value for underground storage, as salt has very low porosity and permeability and is geotechnically stable. Cavities below 1000 ft (300 m) can withstand great pressures and are used for storing crude oil, liquefied petroleum products, and gases. Shallower cavities are used for climate-controlled storage, including isolation of waste materials. *See* SALT DOME.

W. Joseph Schlitt

Ore Leaching

Leaching large stockpiles and mine waste dumps (ore heaps) and, somewhat less important, modified in-situ leaching of caved underground orebodies annually accounts for about one-third of both new copper and gold production in the United States, the world's second largest source of both commodities. Ore leaching is an important contributor to the total world production of these commodities, along with silver. Ore leaching's very low unit processing cost, combined with low-cost earthmoving, allows profitable treatment of huge tonnages of low-grade material that could not otherwise be exploited.

Ore containing rock must first be fragmented to allow chemical leaching solutions to percolate through it. The metal-bearing minerals are liberated using an aqueous wetting fluid that penetrates and saturates the microfractures within the ore particles by capillary action. The ore minerals, originally deposited geochemically from hydrothermal solutions, are typically located in these micropores. Once dissolved, the metals are removed from the ore particles by diffusion through the solution-filled micropores and swept out of the ore heap by the flowing leachate. In effect, nature's ore deposition process has been reversed at a quicker pace. Nevertheless, commercial leaching times are measured in weeks, months, and even years. Leachates, or "pregnant liquors," are processed in surface plants to recover the metal and regenerate the leachant for reuse in the ore heap in a closed-loop system. *See* HYDROMETALLURGY; LEACHING.

In the case of gold and silver ores, the leachant is a basic sodium cyanide solution which forms soluble complex anions [$Au(CN)_2^-$]. Gold is recovered from solution either by adsorption on granular activated carbon or by reductive precipitation (cementation) with zinc dust. Ores amenable to heap leaching are usually fine-grained sedimentary rocks with highly disseminated, submicrometer, free gold grains. While low-grade, these deposits are typically massive and account for nearly all of the present primary gold production in the United States and several other countries. Coarse gold lode deposits and placer deposits are not good candidates for heap leaching because the gold dissolution rate is slow and unacceptably long times are required to dissolve coarse gold grains. Sealed rock microfractures will prevent extraction. A fairly common example of sealing is silicification, occurring after ore deposition. *See* ADSORPTION; GOLD METALLURGY; PRECIPITATION (CHEMISTRY); SEDIMENTARY ROCKS.

Oxidized copper ores contain minerals, such as chrysocolla and malachite, that are readily acid-soluble. Sulfuric acid is the universal leachant for these ores, with the copper dissolved as sulfate salts. However, in the more prevalent primary ores, chalcopyrite ($CuFeS_2$) is the dominant copper mineral, accompanied by relatively large amounts of pyrite (FeS_2). Metals are released from these and other sulfide minerals only upon biooxidation. In the case of pyrite, biooxidation generates sulfuric acid and large quantities of soluble iron within the ore heap. Consequently, primary copper ores generally do not require an external source of acid. *See* BIOLEACHING; COPPER METALLURGY.

Autotrophic bacteria, principally *Thiobacillus ferrooxidans*, are ubiquitous in sulfide ore leaching systems exposed to air and water. The primary role of these bacteria in sulfide leaching systems is to catalyze the oxidation of ferrous ions to ferric ions in the leachant while it is exposed to air inside the ore heap. At the necessarily modest (near-ambient) ore leaching temperatures, uncatalyzed iron oxidation is much too slow. Soluble ferric ions generated by biooxidation outside the ore particles diffuse under a concentration gradient through microfractures into the ore particles, causing sulfide mineral oxidation, and also causing generation and counterdiffusion of both ferrous ions and cupric ions out of the ore particle in concentration gradients. For air (oxygen) to be present, as required for sulfide leaching to occur, the ore heap must not be saturated with percolating leachant. However, if the ore fragmentation is too fine or if excessive amounts of clay are present, heap permeability will be inadequate and saturation may occur, thereby excluding air. *See* INDUSTRIAL MICROBIOLOGY; IRON; OXIDATION-REDUCTION.

Soluble iron and acid released by pyrite oxidation are eventually limited to steady-state concentrations by other consuming chemical processes. Acid attack of nonore minerals usually limits the pH to a range from 2.3 to about 2.9 depending on the individual ore mineral assemblage. Iron is precipitated by hydrolysis reactions, most commonly in ore heaps by generation of potassium jarosite. The concentrations of soluble iron species depend on solution oxidation potential and pH.

From process engineering and economical perspectives, the ore leaching time required to achieve a desirable extent of metal extraction is paramount. Leaching rates must be considered at different spatial scales. The microscale includes the heterogeneous dissolution kinetics of the ore minerals and the

behavior of bacteria. Transport of reactant and product species by diffusion in rock micropores depends on rock size, with the time to achieve equivalent fractional extraction increasing with the ore particle size squared (r_p^2). Heap leaching may involve ore fragmented only by blasting and excavation at the mine face, or it may involve one or more stages of crushing to reduce the rock size and substantially lower the required leaching time when rock micropore diffusion is rate-limiting. Often, crushed ore will be heap-leached for a fixed duration and removed to a spent-ore discard pile, followed by restocking the leaching pad with fresh ore (the on/off method). Alternatively, the leached ore may be covered with a new layer (a new ore lift). Generally, this approach is less expensive, and it is more common when the ore is not crushed, but it is not as well controlled as the on/off method. In both cases, a variety of leachant distribution methods using portable pipe and equipment are used to spread the leachant on the ore heap. These are similar to agricultural irrigation methods and include drip irrigation. Lightweight plastic materials are used for portability and avoidance of corrosion.

Many factors affect leaching extractions over time, including mineral grain size distribution and accessibility to open microfractures, rock microporosity (typically 1–6% in disseminated gold and copper ores), ore particle size and size distribution, and reactant concentrations. Because of the complexities and uncertainties of ore leaching, thorough amenability testing of each ore must be conducted. This is done by leaching representative ore samples in columns that mimic vertical sections through the planned ore heap.

In large ore heaps requiring oxygen for biooxidation of sulfide minerals, inadequate air flow into the heap usually limits the extraction rate. Heat from exothermic sulfide mineral oxidation reactions, and changes in the composition of oxygen and water vapor in the air cause some buoyancy, but this natural advection is very weak and usually inadequate for economic metal extraction rates. Forced-air ventilation with low-pressure fans using perforated air pipes, either underlying ore heaps or bored vertically into preexisting ore heaps, is usually affordable. This approach is shown in **Fig. 2**; a crushed ore with gold encapsulated in sulfide minerals (pyrite and arsenopyrite) is being treated by heap biooxidation prior to conventional cyanide heap leaching, which is conducted after removal of the ore and neutralization of its self-generated acid.

True in-situ leaching of permeable ore deposits, now the primary method of uranium production in the United States, is similar to production of brines from evaporite deposits, involving a field of alternate injection wells and production wells. However, the deposit permeability and porosity are generally very much lower than for evaporites, and a chemical leachant must be used. For uranium, the leachant is commonly an ammonium carbonate solution. True in-situ leaching of deep copper ore deposits (without fragmentation) has been attempted in major field tests but not yet successfully commercialized because of low formation permeability and poor sweep efficiency, coupled with high capital costs. *See* URANIUM METALLURGY.

Fig. 2. Forced-air ventilation of sulfide ore heaps during biooxidation.

Careful environmental planning, permitting, and control during ore leaching are essential. Containment to prevent spills and ground-water contamination, and adequate treatment of all discharged solutions to meet safe drinking water standards must be achieved. At a minimum, this involves geotechnically engineered, impermeable membrane leach pads and solution storage ponds. Ore leaching heaps are exposed to weather and cannot be turned off during maintenance shutdowns as conventional chemical plants are. Large sealed storage and overflow ponds must be provided to collect spills and handle variations in water balance through the year, including floods.

Reclamation of spent ore heaps, discard piles, and in-situ leached ore deposits is primarily directed at the soluble contaminants derived from both the ore and the chemicals introduced during leaching. Acid mine drainage, formed when rain water percolates through the spent ores, containing remnant sulfide minerals that continue to slowly oxidize, is a major problem that has required water treatment plants to be operated in perpetuity at closed mines. Biochemical reduction of sulfates using soluble organic media within the spent ore heap is a promising solution to this problem. As with leaching, environmental site reclamation is often a slow process, requiring months and years of treatment. *See* ENVIRONMENTAL ENGINEERING; LAND RECLAMATION; MINING.

Robert W. Bartlett

Frasch Process

Modern sulfur mining dates from the invention of the Frasch process in the late nineteenth century. H. Frasch accomplished in-place sulfur mining by using superheated water. The process has come down to the present with only minor modifications, and it is

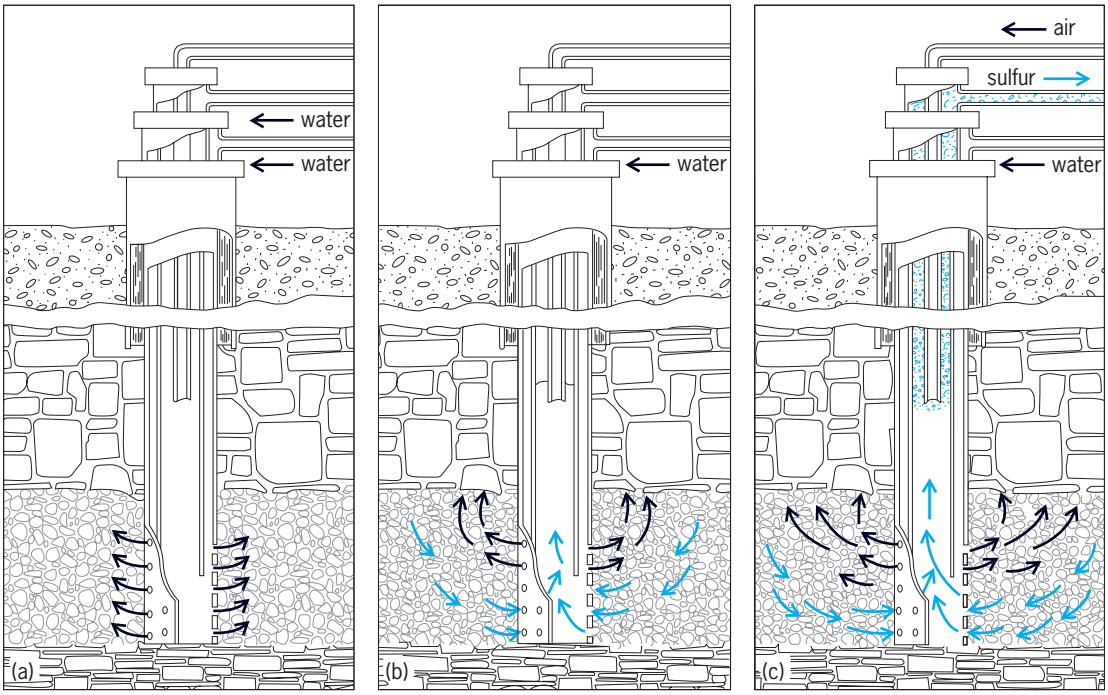

Fig. 3. Sulfur production well in a Gulf Coast salt dome deposit. (*a*) Superheated water is pumped down the inner and outer casing strings. (*b*) Molten sulfur enters inner casing string. (*c*) Sulfur is jetted to the surface as a result of compressed air being pumped down the inner line.

used for mining sulfur deposits in West Texas, offshore Louisiana, Iraq, Poland, and Russia.

Frasch mining begins by drilling wells into the sulfur deposit. Steel tubing (casing) is run into the drill hole to case off the barren overlying formation, and it is cemented in at the top of the sulfur deposit. Within this casing, three concentric strings of pipe are set within the sulfur deposit. The outermost string of pipe, 8 in. (20 cm) in diameter, is set to the bottom of the sulfur deposit. The lower 20 ft (6 m) of this pipe is perforated, providing access for drillhole fluids to the sulfur deposit. Within this string of pipe is another, 4 in. (10 cm) in diameter, with its lower 10 ft (3 m) perforated. In the annulus between the two strings of pipe, a seal (or packer) is placed about 10 ft (3 m) from the bottom of the two casing strings. Thus, superheated water at 325°F (163°C) can be pumped down the annulus between the two strings of pipe to leave the casing string through the perforations above the packer, and to circulate in the sulfur deposit (**Fig. 3***a*). Native sulfur melts at 275°F (135°C) in a cone-shaped area of influence around the well, and because liquid sulfur is denser than water, it settles to the bottom of the well.

The pressure from the heated water forces the molten sulfur into the lower set of perforations and into the inner string of casing (Fig. 3*b*). The molten sulfur rises to its hydrostatic head. This point is determined by practice, and a 0.75-in. (1.9-cm) titanium air line is set at this depth inside the 4-in. (10-cm) casing. Compressed air pumped through this inner tubing expands when it leaves the tubing, thus "jetting" the sulfur to the well collar, where it is collected in a tank (sulfur pan) [Fig. 3*c*]. The procedure to initiate sulfur production involves pumping hot water through inner and outer casing strings, followed by changing to the production setup with sulfur and air traveling through the inner casing strings. The water used to mine the sulfur cools and is removed from the sulfur-bearing zone by bleed-water wells.

Alan F. Edwards

Bibliography. R. W. Bartlett, Metal extraction from ores by heap leaching, *Metals Mater. Trans. B*, 28B:529–545, 1997; R. W. Bartlett, *Solution Mining: Leaching and Fluid Recovery of Materials*, 2d ed., Gordon & Breach, 1998; *Dictionary of Mining, Mineral and Related Terms*, 2d ed., American Geological Institute, Alexandria, VA, 1997; A. F. Edwards et. al., *The Geology of the Main Pass Sulphur and Oil Deposit, Amer. Inst. Mining Preprint*, no. 92–254, 1992; I. A. Given (ed.), *Mining Engineering Handbook*, vol. 2, sec. 21, SME-AIME, 1973; H. L. Hartman (ed.), *SME Mining Engineering Handbook*, 2d ed., sec. 15.2 and 15.3, SME, Littleton, CO, 1992; *In-Situ Leaching Mining, U.S. Bur. Mines Inform. Circ.*, no. 9216, March 1989; D. D. Rabb, Solution mining, *Min. Eng.*, 24(2):62–64, February 1972; D. D. Rabb, *Solution Mining: A Review*, SME-AIME, September 1986; W. J. Schlitt (ed.), *Salts & Brines '85*, SME-AIME, New York, 1985.

Solvation

The association or combination of a solute unit (ionic, molecular, or particulate) with solvent molecules. This association may involve chemical

or physical forces, or both, and may vary in degree from a loose, indefinite complex to the formation of a distinct chemical compound. Such a compound contains a definite number of solvent molecules per solute molecule.

Solvation occurring in aqueous solutions is referred to as hydration. In aqueous ionic solutions, the highly polar water molecules become oriented about the ions, forming spheres of hydration. As a result, the mobility of an ion under an applied voltage gradient is decreased. The extent of hydration depends upon the size and charge of the ion.

In certain colloidal suspensions, solvation is, to a large extent, responsible for the stability of the sol. Particles in lyophilic sols strongly adsorb on their surfaces one or more layers of solvent molecules. This protective layer prevents the particles from approaching so closely as to adhere. In water, starch and many proteins form suspensions which are highly solvated. Such systems exhibit a viscosity which is markedly higher than that of the pure solvent. *See* COLLOID; ELECTROLYTIC CONDUCTANCE; HYDRATION; SOLUTION.

Francis J. Johnston

Solvent

By convention, the component present in the greatest proportion in a homogeneous mixture of pure substances (solutions). Components of mixtures present in minor proportions are called solutes. Thus, technically, homogeneous mixtures are possible with liquids, solids, or gases dissolved in liquids; solids in solids; and gases in gases. In common practice this terminology is applied mostly to liquid mixtures for which the solvent is a liquid and the solute can be a liquid, solid, or gas. *See* SOLUTION.

Three broad classes of solvents are recognized—aqueous, nonaqueous, and organic. Formalistically, the nonaqueous and organic classifications are both not aqueous, but the term organic solvents is generally applied to a large body of carbon-based compounds that find use industrially and as media for chemical synthesis. Organic solvents are generally classified by the functional groups that are present in the molecule, for example, alcohols, halogenated hydrocarbons, or hydrocarbons; such groups give an indication of the types of physical or chemical interactions that can occur between solute and solvent. Nonaqueous solvents are generally taken to be inorganic substances and a few of the lower-molecular-weight, carbon-containing substances such as acetic acid, methanol, and dimethylsulfoxide. Nonaqueous solvents can be solids (for example, fused lithium iodide), liquids [sulfuric acid (H_2SO_4)], or gases [ammonia (NH_3)] at ambient conditions; the solvent properties of the first-named substances are manifested in the molten state, whereas the last-named substances must be liquefied to act as solvents.

Nonaqueous Solvents

The classification, often arbitrary, of nonaqueous solvents can be made on the basis of a variety of factors. Excluding utilitarian considerations and classifications based on chemical character, that is, the presence of distinctive groups such as carbonyl in the molecule, useful classification schemes involving nonaqueous solvents inevitably involve considerations of acid-base phenomena. Of the two major theories, the Brönsted-Lowry protonic concept has been the most useful because the early solvents of interest were invariably potential proton donors. However, the Lewis theory of acidity has become very useful for understanding solution phenomena in aprotic systems such as sulfur dioxide (SO_2). *See* ACID AND BASE.

The ability of a solvent to form a so-called onium species is an important factor in defining the nature of the solution phenomena it will support. Onium species can be formed by reaction with a potential proton donor or by self-ionization of the solvent. Four classes of solvents are generally recognized according to their ability to coordinate with the proton—basic, acidic, aprotic, and amphiprotic solvents. *See* COORDINATION CHEMISTRY.

Basic solvents. Such solvents form the onium species most readily and are generally derivatives of ammonia (amines, hydrazines, pyridine, and so on) or water (alcohols and ethers). Hydrogen-containing species in this classification can undergo self-ionization to form the onium species, for example, $2NH_3 \rightleftharpoons NH_4^+ + NH_2^-$, but the other substances in this class require the presence of a proton donor to do so, $R_2O + HX \rightleftharpoons R_2OH^+ + X^-$ (R = organic group; X = an anion).

Acidic solvents. These solvents exhibit a greater tendency to release protons than do basic solvents, and they form onium species with great reluctance. However, it is possible for acidic solvents to undergo self-ionization, for example, acetic acid, $2CH_3CO_2H \rightleftharpoons CH_3CO_2H_2^+ + CH_3CO_2^-$, to form the corresponding onium species. Even though a solvent may be classified as acidic, for example, CH_3CO_2H, it is possible to protonate such molecules with a more strongly acidic substance. Thus, the strong mineral acids such as hydrochloric acid (HCl) ionize in anhydrous acetic acid to form the onium species, $CH_3CO_2H + HCl \rightleftharpoons CH_3CO_2H_2^+ + Cl^-$.

Aprotic solvents. Solvents such as SO_2 and CO_2, also commonly called inert, have very little affinity for protons, and they are incapable of dissociating to give protons. Such aprotic solvents are also called indifferent, nondissociating, or nonionizing. *See* SUPERCRITICAL FLUID.

Amphiprotic solvents. These solvents are capable of either adding or donating protons. Thus, ammonia is amphiprotic because it can lose or accept a proton as shown in the following reactions:

$$(C_6H_5)_3C^- + NH_3 \rightarrow (C_6H_5)_3CH + NH_2^-$$

$$NH_3 + HX \rightarrow NH_4^+ + X^-$$

The classification of a substance as amphiprotic adds little to understanding solvent phenomena, because this designation depends only on the relative strengths of the acids or bases involved. It has always

been possible to produce a medium sufficiently acidic to protonate even acidic solvents which, of course, also would be expected to be good proton donors to bases. From this point of view, anhydrous acetic acid is considered amphiprotic. Whether a substance is an acid or base depends upon the character of the solvent in which it is dissolved. For example, urea, which is very weakly basic in water, is a weak acid in the more basic solvent ammonia and a strong base in the acidic solvent CH_3CO_2H.

Amphiprotic solvents exhibit a well-defined self-dissociation, of hydrogen sulfide, $2HS \rightleftharpoons H_2S^+ + S^-$ ($K_S = a_{H_2}S^+ \cdot a_{S^-}$), for which an autoprotolysis constant K_S can be measured (**Table 1**). It is generally difficult to measure K_S for nonaqueous solvents because apparent self-ionization can arise from the presence of traces of water or other amphiprotic contaminants. *See* IONIC EQUILIBRIUM.

Chemical behavior. Solvents intervene in chemical processes by producing species from solutes that are more reactive than if the solvent were not present. If the solute consists of ions, the energy for the dissolution process is supplied almost entirely from the solvation of ions; this is primarily an electrostatic process. However, two processes may occur when a covalent solute is dissolved: the solvation of molecules or the formation of ions. The solvation of molecular species usually involves dipolar interactions, specific interactions such as hydrogen-bond formation, and formation of covalent bonded species via coordinate covalent bond formation. In many instances it is possible for such intermediate species to undergo ionization. The formation of ions from covalent molecules is a measure of the ionizing power of the solvent, and attempts have been made to correlate this with the dielectric constant of the solvent. Unfortunately, there is a sufficient number of contradictions to make such relationships unreliable.

The extent to which a substance is ionized by neutral donor (solvent) molecules should increase with increasing stability of the cation resulting from nucleophilic attack, $S: + X\text{-}Y \rightarrow S\text{-}X^+ + Y^-$, compared with that of the unionized solute. The strength of the coordinate covalent bond found in the species SX^+ is related to the donor ability of S, the acceptor ability of X in the species XY, steric effects, and the magnitude of specific solvent-solute interactions such as hydrogen bonding. The donor strengths, expressed as donicity (D_n; **Table 2**) of solvent molecules have been defined relative to the reference acceptor antimony pentachloride ($SbCl_5$); the enthalpy of the reaction between $SbCl_5$ and a series of donors in an inert medium is taken as a measure of donicity. The donicity of a solvent has been interpreted as a measure of its donor intensity, nucleophilicity, or Lewis base strength and can be a useful guide in assessing the ionizing power of a solvent. There is no necessary correlation between the donicity of the solvent and its dielectric constant (ϵ). J. J. Lagowski

TABLE 1. Autoprotolysis constants (K_s) of some common solvents at 25°C

Solvent	log (K_s)	Dielectric constant (ϵ)
H_2O	14.0	78.4
H_2O_2	13	84.2
CH_3OH	16.9	32.6
H_2S	34.5 (–78°C; –61°F)	8.99 (–78°C; –61°F)
H_2SO_4	3.6	101
HCO_2H	6.2	58.5 (16°C; 61°F)
$H_2NC_2H_4OH$	5.1 (20°C; 68°F)	37.7
$(CH_3)_2SO$	32	46.6
$HCN(CH_3)_2$ (O ‖)	>21	36.7
CH_3CN	28.5	36.0
C_2H_5OH	19.1	24.3
$H_2NC_2H_4NH_2$	15.3	14.2 (20°C; 68°F)
CH_3CO_2H	14.5	6.1 (20°C; 68°F)

TABLE 2. Donicity (D_n) and dielectric constant (ϵ) of some common solvents

Solvent	D_n	ϵ
1,2-Dichloroethane	0.0	10.1
Sulfurylchloride	0.1	10.5
Tetrachlorethylene carbonate	0.2	9.2
Thionyl chloride	0.4	9.2
Acetyl chloride	0.7	15.8
Benzoyl chloride	2.3	23.0
Nitromethane	2.7	35.9
Dichloroethylene carbonate	3.2	31.8
Nitrobenzene	4.4	34.8
Acetic anhydride	10.5	20.7
Phosphorus oxide chloride	11.7	14.0
Benzonitrile	11.9	25.2
Selenium oxide chloride	12.2	46.0
Monochloroethylene carbonate	12.7	62.0
Acetonitrile	14.1	38.0
Sulfolane	14.8	42.0
Propylene carbonate	15.1	69.0
Benzyl cyanide	15.1	18.4
Ethylene sulfite	15.3	41.0
Isobutyronitrile	15.4	20.4
Propionitrile	16.1	27.7
Ethylene carbonate	16.4	89.1
Phenylphosphorus oxide difluoride	16.4	27.9
Methyl acetate	16.5	6.7
Butyronitrile	16.6	20.3
Acetone	17.0	20.7
Ethyl acetate	17.1	6.0
Water	18.0	81.0
Phenylphosphorus oxide dichloride	18.5	26.0
Diethyl ether	19.2	4.3
Tetrahydrofuran	20.0	7.6
Diphenylphosphorus oxide chloride	22.4	—
Trimethyl phosphate	23.7	6.8
Tributyl phosphate	23.7	6.8
Dimethylformamide	26.6	36.1
Dimethylacetamide	27.8	28.9
Dimethyl sulfoxide	29.8	45.0
Diethylformamide	30.9	—
Dimethylacetamide	32.2	—
Pyridine	33.1	12.3
Hexamethylphosphoramide	(38.8)	30.0

Organic Solvents

Organic solvents are liquids with atmospheric boiling points lying in the approximate range 40–200°C (100–390°F). They are relatively low cost products of the petrochemical industry, with small quantities

TABLE 3. Properties of commonly used solvents

Solvent	Boiling point, °C (°F)	Specific gravity	Polarity (water = 1)	Solubility parameter*	Evaporation time (ether = 1)
n-Pentane	36.1 (97.0)	0.626	0.009	7.0	1.0
n-Hexane	69 (156)	0.659	0.009	7.3	1.4
n-Heptane	98.4 (209)	0.684	0.012	7.5	3.0
Cyclohexane	80.7 (177)	0.778	0.006	8.2	3.4
Toluene	110.6 (231.1)	0.867	0.099	8.9	6.1
Xylenes	139 (282)	0.871	0.074	8.8	13.5
Ethylbenzene	136 (277)	0.867		8.3	8.8
Methanol	64 (147)	0.792	0.762	14.5	6.3
Ethanol	78 (172)	0.789	0.654	13.4	8.3
i-Propanol	82 (180)	0.786	0.546	11.5	11
n-Butanol	118 (244)	0.810	0.602	11.4	33
Ethylene glycol	198 (388)	1.1135	0.790	17.1	1550
Ethyl glycol	136 (277)	0.929	0.627	9.9	43
Butyl glycol	171 (340)	0.901	0.602	9.5	119
Acetone	56 (130)	0.790	0.355	10.0	2.0
Me Et ketone	80 (180)	0.805	0.327	9.3	3.3
Me iBu ketone	116 (241)	0.801	0.269	8.4	7.0
Dichloromethane	40 (100)	1.326	0.309	9.7	1.4
Trichloroethylene	87 (190)	1.464	0.160	8.0	3.6
Perchloroethylene	121 (250)	1.622		4.5	11.0
Tetrahydrofuran	66 (150)	0.888	0.207	9.9	2.2
Dioxane	101 (214)	1.0336	0.164	10.0	7.3
Ethyl ether	34.5 (94.1)	0.715	0.117	7.4	1.0
Me tBu ether	55 (130)	0.740	0.148	7.4	1.6
Methyl acetate	57 (135)	0.927	0.287	9.0	2.1
Ethyl acetate	77 (170)	0.902	0.228	9.1	3.0
n-Butyl acetate	126 (259)	0.883	0.241	8.6	12
Dimethylformamide	153 (307)	0.945	0.404	12.1	120
Dimethylacetamide	166 (330)	0.959	0.657		172
Acetonitrile	81.6 (179)	0.782	0.460	11.9	2.04
Pyridine	115 (239)	0.983	0.302	10.7	12.7

*Hildebrand solubility parameter, $(\text{calories/cm}^3)^{1/2}$.

also made from coke oven by-products, from turpentine and by fermentation.

Organic solvents are used in a wide variety of domestic and industrial products and processes falling into three main categories. (1) paints, printing ink, adhesives, toileteries, and other coatings where the solvent allows the solute to be put into position and to adhere to the surface when the solvent evaporates; (2) cleaners of materials ranging from fabrics to metal surfaces and electronic parts, where the solvent carries away unwanted contaminants; and (3) reaction media, largely in the chemical and pharmaceutical industries, where the solvent is removed as a mother liquor or by evaporation, or is incorporated in the product as a means of delivery (for example, in agrochemicals).

For dissolving resins and for cleaning when no chemical change takes place, the polarity of a solvent, its solubility parameter, and its evaporation rate are useful ways of predicting a solvent's performance (**Table 3**).

Since many of the ways of recapturing solvents from the air result in the recovered solvent being water-saturated, its water solubility and the ease with which, if necessary, it can be dried are also important.

Solvents used in chemical reactions can be chosen to modify the course of a reaction and to remove, by forming an azeotrope or by a phase separation, water or other unwanted products of reaction.

Hazardous properties. Apart from the operational performance of a solvent, its choice for any application will be made with its various hazardous properties in mind.

Toxicity by inhalation. This is an important consideration. Many of the solvents that were once widely used are being rejected because of the toxicity of their vapor. Benzene and chloroform are examples of solvents that are so toxic that they are difficult to handle in a safe way. *See* BENZENE.

Other organic solvents are comparatively harmless and can be used in domestic products without serious risk (**Table 4**). A parameter known as the threshold limit value (TLV) is defined as the concentration in air considered safe for people to work in on a regular basis. These limits vary slightly country to country.

Only a few solvents have strong enough odors at their threshold limit values to be reliably detected by smell, and even then prolonged exposure tends to reduce an individual's sensitivity.

Much higher concentrations of solvent in air are those giving immediate danger to life and health (IDLH), at which 30 min as a single exposure might have a permanently harmful effect; and even this is lower than the level to which solvent abusers expose themselves, which is close to the vapor pressure of the pure solvent (Table 4).

Toxicity by absorption. Some solvents, such as dimethyl sulfoxide and dimethylacetamide, are easily

TABLE 4. Health and fire hazards of commonly used solvents

Solvent	Threshold limit value, ppm	Immediate danger to life and health, ppm	Odor threshold*, ppm	Saturated vapor at 21°C (70°F), ppm	Flash point °C (°F)	Auto-ignition temperature, °C (°F)
n-Pentane	600	5,000	400	720,000	−40 (−40)	260 (500)
n-Hexane	50	5,000	130	200,000	−22 (−7.6)	225 (437)
n-Heptane	400	4,250	220	51,800	−4 (25)	230 (446)
Cyclohexane	300	10,000	25	155,700	−7 (1.4)	260 (500)
Toluene	100	2,000	0.2	31,000	4 (39)	480 (896)
Xylenes	100	10,000	1.1	9,180	23 (73)	480 (896)
Ethylbenzene	100	2,000	3	9,960	21 (70)	460 (860)
Methanol	200	25,000	100	133,000	15 (59)	470 (878)
Ethanol	1,000	20,000	84	60,000	13 (55)	423 (793)
i-Propanol	400	20,000	90	46,000	12 (54)	456 (853)
n-Butanol	50	8,000	3	6,300	35 (95)	367 (693)
Ethylene glycol	50			53	116 (241)	400 (752)
Ethyl glycol	5	6,000	200	5,400	40 (104)	238 (460)
Butyl glycol	25	700	13	860	60 (140)	244 (471)
Acetone	750	20,000	100	342,800	−18 (−0.4)	465 (869)
Me Et ketone	200	3,000	10	112,000	−6 (21)	515 (961)
Me iBu ketone	100		25	21,700	13 (55)	459 (858)
Dichloromethane	100	5,000	250	500,000	None	605 (1121)
Trichloroethylene	50	500	30	80,000	None	420 (788)
Perchloroethylene	50	400	5	23,000	None	None
Tetrahydrofuran	200		30	230,000	−15 (5)	321 (607)
Dioxane	25	200	170	38,000	12 (54)	180 (356)
Ethyl ether	400	19,000	10	19,000	−45 (−49)	160 (320)
Methyl acetate	200	10,000	50	290,000	−10 (14)	500 (932)
Ethyl acetate	400	10,000	4	100,000	−4 (25)	484 (903)
n-Butyl acetate	150	10,000	10	14,200	22 (72)	421 (790)
Dimethylformamide	10	3,500	100	3,700	62 (144)	445 (833)
Dimethylacetamide	10	400	50		70 (158)	490 (914)
Acetonitrile	40	4,000	40	96,000	6 (43)	524 (975)
Pyridine	5	3,600	0.01	22,000	20 (68)	522 (972)

*Individuals' abilties to detect solvent vapors by odor vary widely.

absorbed through the skin. While this property can sometimes be turned to good use, great care is needed to avoid exposure of unprotected skin to them. *See* DIMETHYL SULFOXIDE.

Fire safety. Of the widely used organic solvents with an assured future, only the chlorohydrocarbons (methylene chloride, trichloroethylene, and perchloroethylene) are not flammable. Some organic solvents produce explosive vapors at ambient temperatures. This characteristic is measured by the solvent's flashpoint (Table 4).

The potential of generating an explosive atmosphere will usually accompany a high rate of evaporation and call for precautions in storage and handling, such as blanketing storage tanks with nitrogen and using specially protected electrical equipment. With the sole exception of methanol, all organic solvents have a vapor with a density substantially higher than air, and ventilation should be planned with this in mind.

Some solvents, mostly ethers and glycol ethers, have low autoignition temperatures (Table 4), meaning that their vapors mixed with air may explode without a spark or naked flame if brought into contact with a modestly hot surface.

Many solvents, particularly hydrocarbons, combine a high electrical resistance with immisciblity with water, and they generate static electricity when a mixture of solvent and water flows through a pipe. To combat the danger of a static spark causing an explosion in the head space in a tank, careful electrical grounding of the entire installation is necessary.

Environmental pollution. All used solvent finally ends its life in the atmosphere, in water, or in incineration. For many applications, particularly in domestic and workshop use, incineration is impractical. Thus for all used solvents consideration must be given to their eventual environmental impact.

Solvents have three major impacts on the atmosphere: damage to the ozone layer, contribution to the production of low-level ozone, and smell.

The use of chlorofluorocarbons (CFC 11 and CFC 131) as cleaning solvents is being rapidly phased out, and the less harmful but widely used 1,1,1-trichloroethane is also being withdrawn. *See* HALOGENATED HYDROCARBON.

Local air pollution from ozone formation in urban areas has as a vital component the presence of volatile organic compounds. Organic solvents form up to 40% of all volatile organic compounds, and it is desirable to use solvents whose photochemical ozone creation potential (POCP) is low (**Table 5**).

Some people with sensitive noses find solvent odors at well below the hazardous level unpleasant (Table 4). *See* AIR POLLUTION; STRATOSPHERIC OZONE.

Discharge of solvents into rivers or ground water is potentially harmful. The biological degradation of the solvents in wastewater plants requires special treatment and may give rise to offensive odors.

TABLE 5. Photochemical ozone creation potential (POCP) of commonly used organic solvents

Solvent	POCP
n-Pentane	
n-Hexane	40
Cyclohexane	25
Toluene	55
Xylenes	90
Ethylbenzene	—
Methanol	10
Ethanol	27
i-Propanol	15
n-Butanol	55
Ethylene glycol	—
Ethyl glycol	75
Butyl glycol	75
Acetone	20
Me Et ketone	42
Dichloromethane	0.9
Trichloroethylene	7
Perchloroethylene	1
Tetrahydrofuran	70
Dioxane	—
Ethyl ether	60
Methyl acetate	3
Ethyl acetate	20
n-Butyl acetate	45
Dimethylacetamide	—
Acetonitrile	—
Pyridine	—

Low concentrations of chlorine-containing solvents can make water unfit for drinking. *See* WATER POLLUTION.

Usage. Global restrictions are causing a change in use of organic solvents away from the less effective and more polluting hydrocarbon solvents toward the oxygenated ones. Since the latter cost more and have a lower heating value, there is more incentive to recover rather than to incinerate solvents used in industry. *See* INDUSTRIAL HEALTH AND SAFETY; TOXICOLOGY.

Ian McN. Smallwood

Bibliography. H. Ohtaki and Y. Yamatera, *Structure and Dynamics of Solutions*, 1992; P. Politzer and J. S. Murray, *Quantitative Treatments of Solute Solvent Interactions*, 1994; C. Reichardt, *Solvents and Solvent Effects in Organic Chemistry*, 2d ed. 1988; D. J. De Renzo, *Solvents Safety Handbook*, 1986.

Solvent extraction

A technique, also called liquid extraction, for separating the components of a liquid solution. This technique depends upon the selective dissolving of one or more constituents of the solution into a suitable immiscible liquid solvent. It is particularly useful industrially for separation of the constituents of a mixture according to chemical type, especially when methods that depend upon different physical properties, such as the separation by distillation of substances of different vapor pressures, either fail entirely or become too expensive.

Industrial plants using solvent extraction require equipment for carrying out the extraction itself (extractor) and for essentially complete recovery of the solvent for reuse, usually by distillation. *See* DISTILLATION; EVAPORATION; STRIPPING (CHEMICAL ENGINEERING).

Applications. The petroleum refining industry is the largest user of extraction. In refining virtually all automobile lubricating oil, the undesirable constituents such as aromatic hydrocarbons, which have poor chemical and viscosity-temperature characteristics, are extracted from the more desirable paraffinic and naphthenic hydrocarbons. The principal solvents used are furfural, phenol, and a combination of phenol with propane and cresylic acid; nitrobenzene and 2,2′-dichloroethyl ether are used in minor amounts. Liquid propane is also used preferentially to extract the desirable constituents from unwanted asphaltic compounds.

By suitable catalytic treatment of lower boiling distillates, naphthas rich in aromatic hydrocarbons such as benzene, toluene, and the xylenes may be produced. The latter are separated from paraffinic hydrocarbons with such solvents as liquid sulfur dioxide, furfural, and ethylene glycol to produce high-purity aromatic hydrocarbons and high-octane gasoline.

Gasoline is "sweetened," or freed of its sulfur-containing compounds, by extraction with aqueous caustic solutions containing various naphthenic and aromatic acids, or methanol, to modify the solvent characteristics. Aqueous copper ammonium acetate is used to extract butadiene from other 4-carbon hydrocarbons in synthetic rubber production.

Vegetable oils are separated into relatively saturated and unsaturated glyceride esters with furfural or liquid propane as solvents. The former are edible products; the latter are drying oils used in paints. Fish oils are similarly treated and yield a high-vitamin fraction as well.

In by-product coke-oven plants, phenols and other tar acids are recovered from the ammoniacal liquors with benzene, tricresyl phosphate, butyl acetate, and other solvents in large installations. The pharmaceutical industry uses extraction to separate natural impurities or unwanted chemical by-products from products such as synthetic vitamins, penicillin, Aureomycin antihistamines, reserpine, and a host of others.

All uranium for atomic energy purposes is freed of its impurities in aqueous solution by extraction into diethyl ether, tributyl phosphate, and other solvents. The reprocessing of atomic energy fuels for the recovery of plutonium, and the separation of many of the other fission products such as the rare-earth metals, utilize solvent extraction extensively. The otherwise hard-to-separate metal pairs, zirconium-hafnium and niobium-tantalum, are separated in quantity with comparative ease by these methods.

Equipment. Extractors bring about direct contact of the feed (solution to be separated) and extracting solvent in order to permit diffusional transfer of the constituents from the feed to the solvent. The rate of transfer depends upon the contact area of the two liquids and the degree of turbulence developed within them. The extractor disperses one of the

liquids in the other to produce large surface area, and relative motion to produce turbulence. The extractor must also provide for the subsequent mechanical separation of the dispersion, based upon the different densities of the liquids, to permit withdrawal of the two effluent products, the extract (solvent containing the extracted constituents) and the raffinate (unextracted residue).

Mixer-settlers (**Fig. 1**) provide for these requirements in separate vessels. The feed and solvent flow continuously through the mixer, in which the rotating agitator disperses one of the liquids into small droplets immersed in the other. The size of this vessel must provide sufficient residence time for the liquids that the desired diffusional transfer occurs. The degree of agitation must be intense without, however, producing so fine a dispersion that subsequent settling is difficult. The dispersion flows to the settler, most simply a drum, in which low velocity and lack of agitation promote gravity settling and coalescence of the drops to provide clear effluents.

Since in such single-stage apparatus the extractable substance approaches a concentration equilibrium in the effluents, nearly complete extraction requires a multiplicity of stages. An arrangement for countercurrent interstage flow of the liquids reduces the amount of solvent needed (**Fig. 2**). The compact modification of **Fig. 3** has found particular favor in extraction of radioactive metals from aqueous solutions in processes associated with atomic energy operations.

To reduce the floor space and pump requirements for multistage extractors, a variety of vertical towers are also used. These involve countercurrent vertical flow, under gravity, of one of the liquids in dispersed form through a continuum of the other by virtue of the different liquid densities. A packed tower

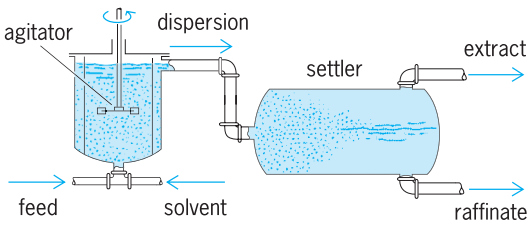

Fig. 1. Single-stage mixer-settler extractor.

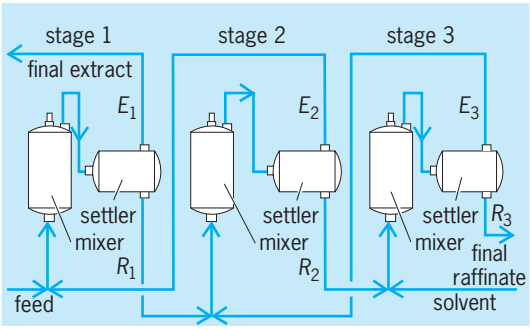

Fig. 2. Diagram of a three-stage countercurrent mixer-settler extractor. (*After R. E. Treybal, Mass Transfer Operations, 2d ed., McGraw-Hill, 1968*)

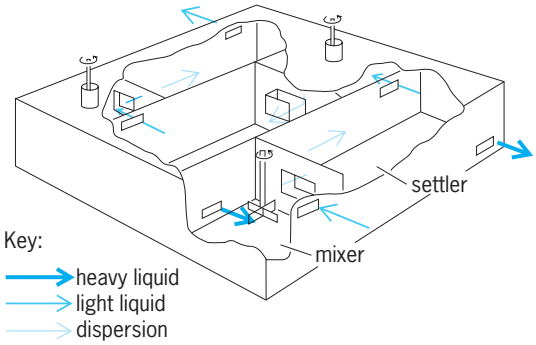

Fig. 3. Three-stage, box-type, mixer-settler extractor.

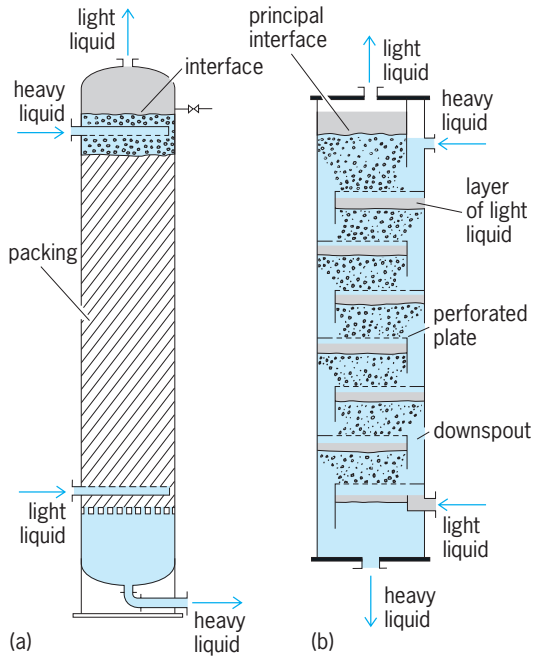

Fig. 4. Vertical tower extractors. (*a*) Packed-tower extractor. (*b*) Perforated-tray extractor. (*After R. E. Treybal, Mass Transfer Operations, 2d ed., McGraw-Hill, 1968*)

(**Fig.** 4*a*) is a cylindrical shell, the bulk of which is filled with manufactured packing, such as rings or saddles, randomly arranged. The more dense liquid, introduced at the top, flows downward as a continuum. The less dense liquid enters at the bottom through small nozzles. The resulting small droplets rise through the heavy liquid, during which time extraction occurs, and then coalesce into a bulk and leave at the top. The packing serves to maintain the dispersion and provide moderate turbulence. The dispersed liquid may be either feed or solvent, light or heavy. If it is heavy, the droplets settle downward. Although the liquids are not repeatedly dispersed and settled as in the multistage mixer-settler, nevertheless multistage effects are obtained. Spray towers contain no packing and are not as effective. *See* GAS ABSORPTION OPERATIONS.

In perforated-tray towers (Fig. 4*b*) the light liquid collects in a layer under each tray and is dispersed into droplets by the small perforations. The drops rise through the heavy liquid, which flows across

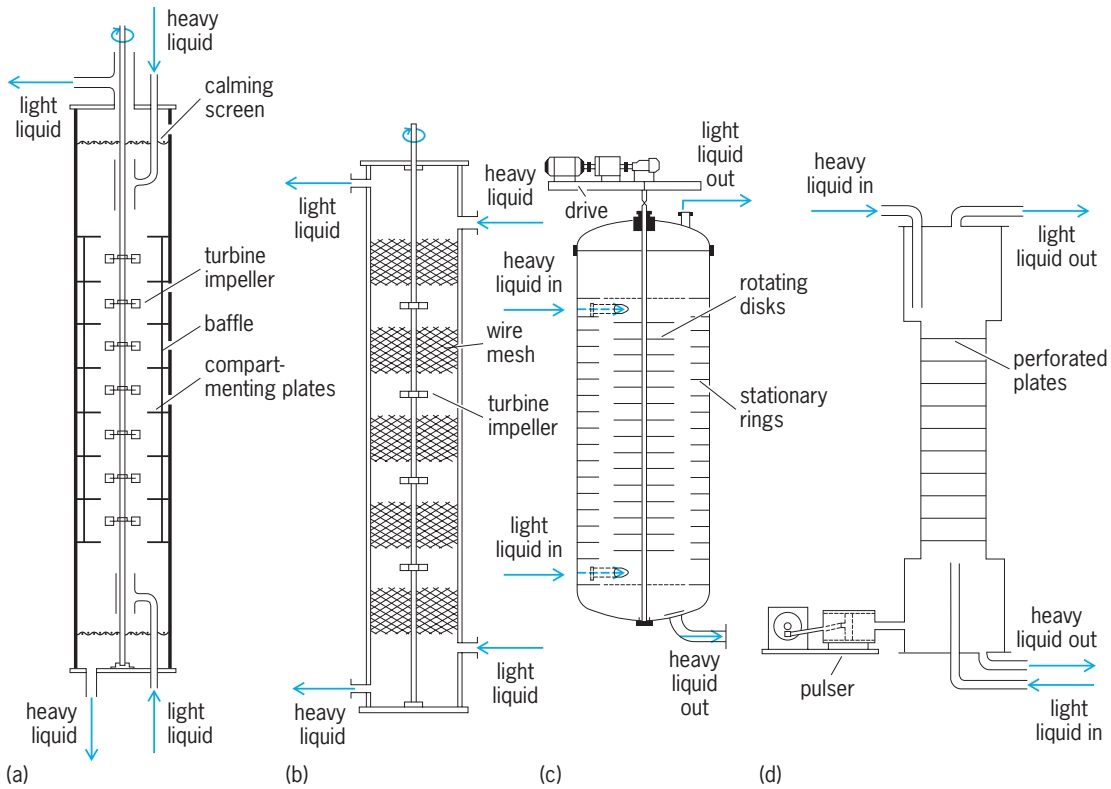

Fig. 5. Types of extractors with mechanical agitation. (*a*) Oldshue-Rushton extractor; (*b*) Scheibel-York extractor (*after R. E. Treybal, Mass Transfer Operations, 2d ed., McGraw-Hill, 1968*). (*c*) Rotating-disk extractor (*after G. H. Reman and R. B. Olney, Chem. Eng. Prog., 51:141–146, 1955*). (*d*) Pulsed extractor (*after T. B. Drew and J. W. Hoopes, eds., Advances in Chemical Engineering, vol. 1, Academic, 1956*).

each tray and through the downspouts. The frequent redispersion achieved makes these towers very effective. Alternatively, by turning the tower upside down, the heavy liquid may be dispersed.

Mechanical agitation, provided by rotating impellers as in the towers of **Fig. 5***a*, *b*, and *c*, is used to obtain finer dispersions and increased turbulence. The pulsed tower (Fig. 5*d*) provides the mechanical agitation by rapid (20–100 cycles/min), small-amplitude (0.25–2 in. or 0.64–5.0 cm), reciprocating motion superimposed upon the natural flow of liquids as they alternately pass through small perforations in the plates. This is particularly useful for handling radioactive liquids, since moving parts may be located in a place of safety.

In all these designs, the tower diameter is governed by the quantity of liquids to be handled, the height by the number of stages of extraction required. Towers up to 15 ft (4.8 m) in diameter and 125 ft (38 m) tall have been built. Auxiliary equipment may include pumps for movement of the liquids, motor drives for agitators, valves and flowmeters for control of flow rates, and liquid-level control instruments.

The centrifugal extractor (**Fig. 6**) consists of a series of perforated, concentric rings in a cylindrical drum, the whole rapidly rotated (2000–5000 rpm) on the horizontal shaft. Liquids enter and leave through the shaft; they flow radially and countercurrently in the rotating drum because the effects of density differences are increased by centrifugal force. The particular virtue of this machine is the low residence time of the liquids, which has made it especially useful in the extraction of antibiotic pharmaceuticals from fermentation broths. *See* COUNTERCURRENT TRANSFER OPERATIONS; EXTRACTION; MIXING; UNIT OPERATIONS.

Laboratory applications. Solvent extraction is carried out regularly in the laboratory by the chemist as a commonplace purification procedure in organic

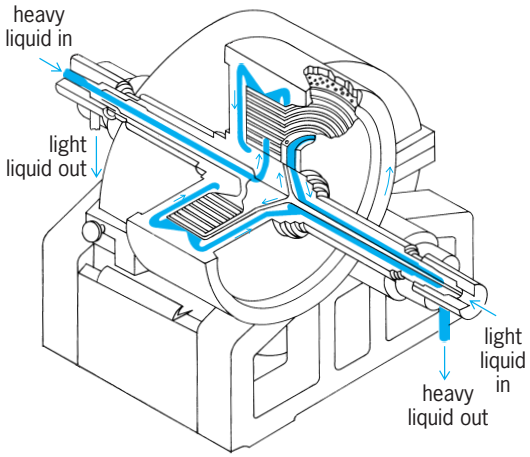

Fig. 6. Podbielniak centrifugal extractor, featuring low residence time of liquid. (*Podbielniak, Inc.*)

synthesis, and in analytical separations in which the extraordinary ability of certain solvents preferentially to remove one or more constituents from a solution quantitatively is exploited. Batch extractions of this sort, on a small scale, are usually done in separatory funnels, where the mechanical agitation is supplied by handshaking of the funnel.

Natural biological products, such as hormones, serums, vaccines, vitamins, and plant extracts, usually consist of many individual chemical compounds, frequently so similar and complex as to defy ordinary analytical procedures to separate them. Such substances, sometimes initially not even suspected of being mixtures, have been successfully separated into their constituents by the very ingenious laboratory device of L. C. Craig, which is capable of performing tens of thousands of extractions automatically.

Mechanism of extraction. Consider the single-stage extraction of a typical solute from a feed solution with a suitable immiscible solvent, as carried out in the device of Fig. 1. **Figure 7**, which is characteristic of the chemical nature of the system, shows the equilibrium distribution of solute between the two liquids as the curve *AGHB*. If two solutions whose respective solute concentrations plot as a point on this curve are brought into contact, they will remain at their initial concentrations indefinitely, and no solute transfer will occur. On the other hand, if their concentrations are such that they provide a point removed from the curve, a spontaneous transfer of solute will occur from one liquid to the other, in an effort to bring the system to equilibrium. If the feed solution of Fig. 1, flowing at rate F, is of concentration C_F (point D in Fig. 7), and the solvent, flowing at rate S, contains no solute, transfer of solute will occur from the feed to the solvent. The rate of extraction at any time is proportional to the degree of departure from equilibrium, the proportionality involving the interfacial area A and rate constants (mass-transfer coefficients k_E and k_R) characteristic of the circumstances. As the liquids enter the vessel, the initial rate of extractionis given by the equation below, where

$$\text{Rate} = k_E A \Delta C_E = k_R A \Delta C_R$$

ΔC_E and ΔC_R are the initial departures from equilibrium at G in the two liquids, as shown in Fig. 7. Because of the necessity of satisfying a material balance (amount of solute removed from the feed equals that taken up by the solvent), the concentrations within the liquids move, not toward G, but instead toward H, whose location is fixed by the relative rates of flow of liquids, as shown. As extraction proceeds, so that the liquids reach point E, the concentration departure from equilibrium is lessened ($\Delta C'_E$ and $\Delta C'_R$) and the rate of extraction is reduced. Theoretically, equilibrium for the effluents at H will never be reached since the rate will then have fallen to zero, but in practice a point J is ultimately reached which can be as close to equilibrium at H as desired, depending upon the time of residence (volume of vessel/volume rate of flow) in the vessel. The degree of approach to equilibrium actually attained is the stage efficiency of the extractor, which may be expressed as the ratio of line lengths DJ/DH.

The interfacial area A is the surface of all the droplets of dispersed liquid in the vessel. Since for reasons of cost it is desired that the rate of extraction be large and the time of residence and vessel volume small, A is made large by creating small droplets through vigorous agitation. In well-agitated vessels, A may be 1000 ft^2 or more for each cubic foot (or 3280 m^2 for each cubic meter) of liquid contents. Too vigorous agitation, which may produce such small droplets that permanent, unsettleable emulsions form, must be avoided, however. The droplets are smallest in the vicinity of the agitator owing to locally high turbulence intensity. They tend to coalesce and become larger in regions remote from the agitator.

The mass-transfer coefficients k_E and k_R are characteristic of the physical properties and the nature of the motion of the extract and raffinate liquids, respectively. If the liquids are entirely quiet, they are related to the molecular diffusion coefficient of

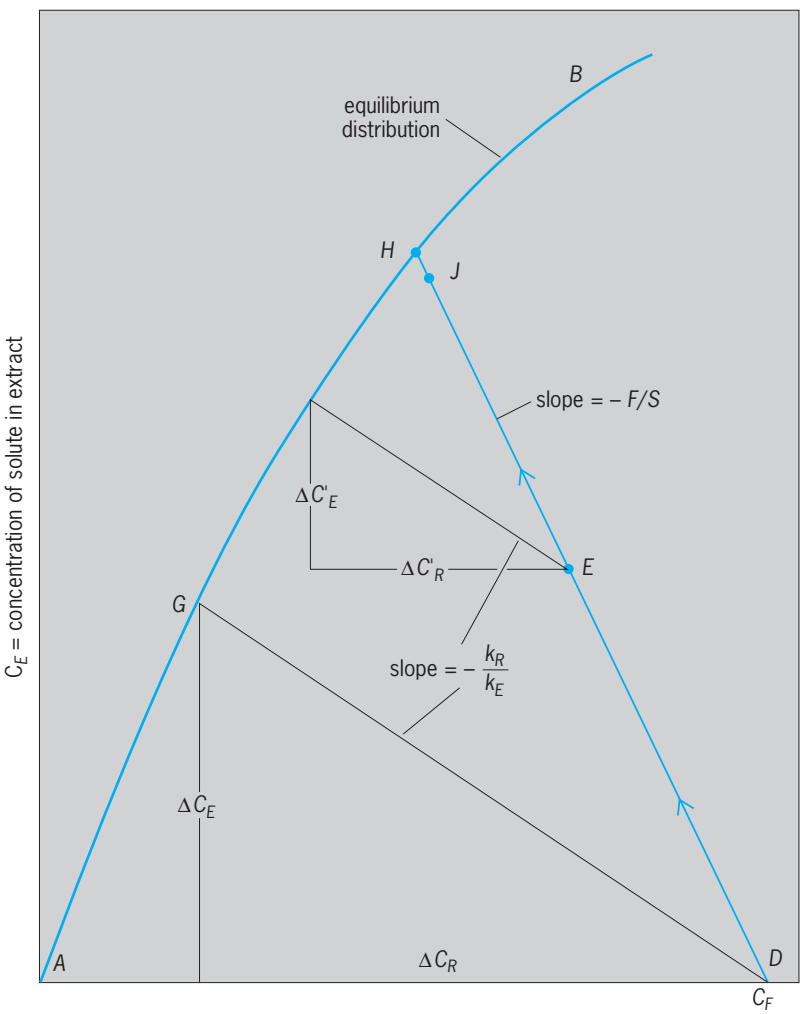

Fig. 7. Single-stage extraction. C_R = concentration of solute in raffinate.

the solute through the liquids. The transfer of solute molecules from the bulk of the raffinate, through the interface surface, and into the bulk of the extract, then depends upon the random, thermal motion of the molecules, and their passage is greatly hindered by their frequent collision with the more abundant solvent molecules. The resulting mass-transfer coefficients are then very small.

For liquids in motion, however, the coefficients are very much larger, to an extent depending upon the intensity and scale of turbulence within each liquid. Under these conditions, it is customary to ascribe the transport of solute from the bulk of the raffinate liquid to the interface surface, and ultimately into the extract, to eddies, or relatively large chunks of liquid which move rapidly, carrying with them their contained solute. The influence of the diffusion coefficients is then substantially less.

For a given degree of turbulence as produced by the agitator, a number of additional factors also have influence on the rate of extraction. These are discussed briefly below.

1. The relative motion of liquid droplets and the continuum in which they are immersed result in a transfer of momentum across the interface, and this in turn results in internal circulation of the droplet liquid, leading to relatively large mass-transfer coefficients for the dispersed liquid. The larger this internal circulation the larger the ratio of viscosities of continuous to dispersed liquid. In many systems of industrial importance, however, there are present trace concentrations of substances which tend to adsorb at the interface, producing a rigidity of the interface and a marked reduction in the internal circulation within the droplets, with consequent lowering of the mass-transfer coefficients. In one case, 6×10^{-5} g surface active agent/100 ml liquid reduced the expected rate coefficient by 68%. Other mechanisms have also been proposed to explain such a reduction. Blocking of portions of the interface by the adsorbed molecules and chemical interaction of the adsorbed substance and the extracting solute are two possible explanations.

2. The relative motion of the liquid droplets and the continuum result in unequal mass-transfer rates, and hence unequal solute concentrations, about the drop periphery. Since interfacial tension depends upon concentration, an interfacial tension gradient along the drop surface may result, which brings about a relatively violent surface movement (Marangoni effect). This in turn enhances the mass-transfer coefficient substantially.

3. When two droplets coalesce, the reduction of surface results in a substantial release of energy. This produces an increased relative motion of the droplets and the continuum, in turn leading to an increase in the mass-transfer coefficients.

4. Particularly in the extraction of metals from solution, extraction is frequently dependent upon a chemical reaction between the solute metal and a substance in the solvent, to make the metal soluble in the organic solvent. Such reactions may be relatively slow, and the rate of extraction will then be largely governed by the reaction rate.

While the mass-transfer coefficients and the stage efficiency of extraction may be computed with reasonable accuracy for ordinary circumstances, when any of the four conditions described above exist it is presently impossible to predict the rate of extraction with assurance. Resolution of these problems is the object of much current research. Robert E. Treybal

Bibliography. J. Marinsky (ed.), *Ion Exchange and Solvent Extraction*, 1995; T. Ritcey and A. W. Ashbrook, *Solvent Extraction*, pt. 1: *Principles and Applications to Process Metallurgy*, 1984; T. Sekine, *Solvent Extraction Chemistry*, 1977; R. E. Treybal, *Mass-Transfer Operations*: *Chemical Engineering*, 3d ed., 1979; A. Weissberger (ed.), *Techniques of Organic Chemistry*, 2d ed., vol. 3, 1957; J. A. Wesselingh and R. Krishna, *Mass Transfer*, 1991.

Somatic cell genetics

The study of mechanisms of inheritance in animals and plants by using cells in culture. In such cells, chromosomes and genes can be reshuffled by parasexual means rather than having to depend upon the genetic recombination and chromosome segregation that occur during the meiotic cell divisions preceding gamete formation and sexual reproduction. Genetics is concerned with the role of genes and chromosomes and their environmental interactions in the development and function of individuals and the evolution of species. Genetic analysis of complex multicellular organisms classically requires multigeneration families. Because fairly large numbers of progeny of defined matings have to be scored, genetic analysis of animals and plants with long generation times, small families, or lack of controlled matings is difficult and slow.

Somatic cell genetics circumvents many of these limitations. It has enhanced the scope and speed of genetic analysis in higher plants and animals, especially when combined with the powerful techniques of molecular biology. Cells in culture can be used to generate fertile plants and animals from single cells, including stem cells. With these methods, virtually every gene in any species of interest can be identified, its functions clarified, and the structural and functional changes it has undergone in particular phylogenetic lineages determined.

Cells in culture. Plant meristem cells can be grown readily in culture. They can be totipotent (capable of differentiating into all cell types) and can even grow into a mature plant. In mammals, the only totipotent cells, in addition to fertilized eggs, are the embryonic stem cells of early preimplantation embryos and some embryonal carcinoma cells. Even adults contain some stem cells, but these appear to be only pluripotent or even lineage-specific, able to differentiate into only a limited number of cell types. Totipotent cells appear to be immortal in culture. Other diploid cells senesce, dying after a limited

number of cell divisions. This is sufficient for chromosome studies, and hundreds of thousands of these are performed each year on normal and cancerous human cells. In addition, cells in early passage cultures of many species can be frozen and maintained in a viable state for years (cryopreservation). The cells can be thawed at any time and grown in culture. Thus, a single biopsy can provide cultured diploid cells for many years, including cells from endangered species. *See* STEM CELLS.

Normal cells in culture can be genetically and chromosomally stable, but genetic changes, or mutations, can be readily induced in them. Mutations provide the markers essential for genetic analysis. Mutations occur spontaneously, but they can be induced at much higher frequencies by a variety of physical, chemical, and living agents. Normal cells can be transformed into immortal cells (which can divide indefinitely) by exposure to certain chemicals or oncogenic viruses, such as SV40 or papillomavirus, and the transformed cells can form tumors in a suitable host. Transformed cells, like most cancer cells, are immortal. They are unstable, with both structural and numerical changes in the chromosomes. Somatic cell genetic methods have shown that many cancers are associated with specific chromosome rearrangements. *See* ANIMAL VIRUS; CHROMOSOME; CHROMOSOME ABERRATION; MUTAGENS AND CARCINOGENS; MUTATION; ONCOGENES; RECOMBINATION (GENETICS); TUMOR.

Somatic cell hybrids. Cells in culture can be induced to fuse when their cell membrane is modified by certain viruses (such as parainfluenza or measles), polyethylene glycol, or an electric shock (electroporation). After fusion of two cells, the nuclei are in a common cytoplasm. The fused cell is called a heterokaryon if the parental cells are of different origins. Heterokaryons are useful in complementation analysis, which is based on fusing cells from two individuals who have the same disease phenotype. If the two individuals have a mutation in the same gene, the heterokaryons will lack the gene product and will still show the abnormal phenotype. However, if the two individuals have mutations of different genes, each parental cell will provide the normal gene product missing in the other parent, and the heterokaryons will have a normal phenotype. Complementation analysis has shown that the human disease Fanconi's anemia, a potentially fatal blood disorder, is caused by a mutation in any one of at least six different genes.

The nuclei in a heterokaryon tend to become synchronized and enter cell division (mitosis) together. When the two nuclear membranes break down at metaphase of mitosis, a single mitotic spindle is formed, and each of the two daughter nuclei gets a complete set of chromosomes from each parental nucleus and is a true somatic cell hybrid.

Hybridomas are a special class of somatic cell hybrids. B lymphocytes, the antibody-secreting cells of the immune system, senesce very quickly in culture, but if they are hybridized with immortal malignant myeloma cells of the same species, the resultant hybridoma cells are also immortal, and they produce antibody molecules with the same unique specificity as the lymphocyte parent. These monoclonal antibodies are widely used in diagnostic studies and research. *See* ANTIBODY; ANTIGEN; MITOSIS; MONOCLONAL ANTIBODIES.

Interspecific cell hybrids, especially those involving distantly related species, tend to lose chromosomes during mitosis. Rodent–human hybrid cells tend to lose human chromosomes. In hybrids with a much reduced number of human chromosomes, any human gene still present must reside on one of the remaining human chromosomes. The analysis of rodent–human somatic cell hybrids has led to the mapping of thousands of genes and anonymous deoxyribonucleic acid (DNA) fragments. Interspecific somatic cell hybrids have been widely used to map genes in other mammalian species as well, including the owl monkey, mouse, cow, pig, and fox. More precise mapping has been achieved by in-situ hybridization of labeled DNA fragments or genes to chromosomes in normal cultured cells of the species being studied. *See* CHROMOSOME MAPPING.

Cell hybridization can lead to the extinction of phenotypic characteristics of one of the parental types. When cancer cells are fused with nonmalignant diploid cells, the resultant hybrid cells are nonmalignant. However, the cells regain their malignancy if they lose both copies of a particular chromosome derived from the normal parent. This suggests that the product of a particular gene on this chromosome is a tumor suppressor. Dozens of human tumor suppressor genes have been identified; mutations in them are important causes of cancers in many different tissues.

Extinction of one parental phenotype in somatic cell hybrids allows genetic dissection of processes such as cell differentiation and senescence. For example, rat hepatoma (liver tumor) cells express the liver-specific gene for albumin production, but when these cells are hybridized with mouse fibroblasts, albumin production is extinguished. Loss of both copies of mouse chromosome 1 from the hybrids leads to reexpression of albumin, indicating that a gene on this chromosome is responsible for blocking the expression of this liver-specific gene, and perhaps others. Fusion of mortal diploid cells with immortal chromosomally unstable cells produces hybrid cells that are mortal (senesce), perhaps because the immortal parental cells had lost both copies of a chromosome carrying a mortality gene. In fact, a gene has been identified by somatic cell genetic methods that reverses the immortal phenotype of cells into which it is inserted. *See* ALBUMIN; LIVER.

A wide variety of plant somatic cell hybrids have been made; a number of these, such as tomato-potato and carrot-parsley, are able to differentiate and form mature plants, although they have lost a few chromosomes and are sterile. Cell hybrids between more distantly related organisms, such as soybean and tobacco, lose the chromosomes of one parent quickly,

but may retain some genes derived from this parent (introgression).

Gene transfer. A gene of interest can be inserted into cells in culture by various techniques.

Viral vectors. One method is to insert the gene into an infectious virus and transfect this into cells. Cells of Fanconi's anemia complementation type E were transfected with a genomic complementary, or copy, DNA (cDNA) library in which each of the thousands of different cDNAs was present in a different copy of the viral vector. When the cells were grown in a selective medium in which survival required a functional copy of the gene, all the surviving cells had incorporated the same cDNA, thus identifying the gene. Such complementation cloning facilitates quick identification of important genes.

When the transfection vector is a retrovirus or an adenovirus, its DNA integrates into a random chromosomal site in the host cell and becomes a permanent part of the genome. If the gene of interest is expressed, it can provide long-lasting correction of a deficiency of that gene product in the host. Adenoviral vectors can enter nonproliferating cells by endocytosis (introduction of extracellular material into the cytoplasm). Because they remain extrachromosomal and thus do not replicate, they can provide gene therapy for only a limited period of time.

Nonviral methods. Genes can also be transferred into cultured cells by nonviral methods, such as precipitation of the DNA onto a monolayer of cells, fusion of DNA-containing lipid droplets with the cell membrane, insertion of an artificial chromosome containing the gene, or microinjection of the gene directly into the nucleus. When genes are introduced into plant cells from other plants, bacteria, yeast, or humans, they continue to function, reflecting the unitary origin of all living organisms and their virtually identical molecular machinery. *See* ANIMAL VIRUS; BACTERIAL GENETICS; PLANT VIRUSES AND VIROIDS.

Transgenic plants and animals. When a foreign gene (transgene) is introduced into a cell nucleus by most of the mechanisms previously described, it inserts at random into some chromosomal site, sometimes disrupting an important gene and causing deleterious effects. Fortunately, methods have been developed that target the transgene to the gene's normal chromosomal position. This has made it possible to replace one of the two copies of a gene in diploid cells with a normal or engineered copy of the gene from an external source.

Plants. Transgenic methods have tremendous potential for the development of plants resistant to insects, microorganisms, or environmental stress. Corn and cotton containing a transgene for a toxin produced by the bacterium *Bacillus thuringiensis* have greatly increased resistance to specific insect pests and led to higher crop yields. Tobacco or tomato callus cells containing a transgene for the antiviral protein of pokeweed (*Phytolacca americana*) have grown into mature transgenic plants that are resistant to a broad spectrum of viruses. Transgenic tobacco plants containing the gene for viral hepatitis B surface antigen produce large amounts of this protein; such transgenic plants may be useful for low-cost vaccine production. *See* IMMUNOGENETICS.

Animals. Transgenic animals are generated in ever-increasing numbers. Transgenic livestock can produce human proteins that may find broad use as medical drugs. Gene transfer techniques are being used to speed the growth of farmed tilapia fish and to reduce the threat to wild populations of tilapia by generating sterile variants. Increasing numbers of transgenic mice are being produced to study gene function, cell differentiation, and morphogenesis and to create animal models of human diseases whose study may speed the development of treatments.

Some transgenic animals are generated by microinjection into the early embryo, but a growing number are produced by targeted gene transfer into cultured embryonic stem cells. The role of a specific gene in development can be studied by replacing one copy of the normal gene with a nonfunctional copy by using gene targeting in embryonic stem cells, and introducing the resultant heterozygous cells into the blastocyst cavity of a developing embryo of the same species. Upon implantation into a pseudopregnant uterus, the preimplantation embryo can develop into a chimeric adult containing both normal host cells and heterozygous transgenic cells. If transgenic cells are present in the germ line, they can produce haploid sperm or ova containing only the nonfunctioning transgene, and progeny produced by these will be heterozygous for the transgene. One fourth of the zygotes produced by a mating of two individuals heterozygous for the transgene should be homozygotes with two copies of the nonfunctional transgene. The effect of such gene knockouts can be assessed at any stage of embryogenesis or at birth, or even later, if the homozygotes are viable. *See* CHIMERA.

Transgenic mice, including gene knockouts, have been produced in an attempt to understand the roles of many different kinds of genes in normal and abnormal development, including cancer. For example, the connexin proteins are important structural components of gap junctions, the channels between some adjacent cells that facilitate the passage of important low-molecular-weight molecules. Knockout mice missing one of these connexins die soon after birth because of the unexpected occurrence of a congenital heart malformation. Transgenic studies have shown that gap junctions are essential for the body's immunological defense against viral and other infections.

Another use of transgenic animals is to introduce a gene modified to express a reporter molecule that is easily detected by its color or fluorescence. This enables the identification and mapping of the cells that express the gene at specific stages in embryonic development or in the adult organism. Transgenic mice with a single copy of a rat heat shock protein gene that was engineered to be highly expressed where the actin gene is expressed normally produce a high level of the heat shock protein in heart muscle and

are protected against heart damage due to ischemia (reduced blood flow).

Implications. Somatic cell genetics has revolutionized genetic analysis in animals and plants. It is providing fundamental insights into developmental biology, and is increasingly important in the identification and characterization of certain genes in key species.

The complete nucleotide sequence of the DNA of a number of animals and plants has been determined in the last few years, and the number is growing rapidly. Knowledge of the complete DNA sequence of several organisms has led to considerably improved estimates of the total number of genes in each [for example, roughly 14,000 genes in the fruit fly, 18,000 genes in the silkworm (*Bombyx mori*), and 22,000 genes in humans]. Somatic cell genetic methods will play a major role in determining the functions of the enormous number of hitherto unknown genes revealed by sequencing complete genomes. In clinical medicine, this will lead to better methods of diagnosis and treatment of genetic diseases. The implications for agriculture and animal husbandry are equally exciting. *See* GENE; GENE ACTION; GENETIC ENGINEERING; GENETICS; HUMAN GENOME PROJECT.

Orlando J. Miller

Bibliography. M. J. Bertram et al., Identification of a gene that reverses the immortal phenotype of a subset of cells and is a member of a novel family of transcription factor-like genes, *Mol. Cell. Biol.*, 19:1479–1485, 1999; M. R. Capecchi, Targeted gene replacement, *Sci. Amer.*, 270(3):52–59, 1994; N. McLean et al., Transgenic tilapia and the tilapia genome, *Gene*, 295:265–277, 2002; J. E. Meija et al., Functional complementation of a genetic deficiency with human artificial chromosomes, *Amer. J. Human Genet.*, 69:315–326, 2001; M. Qaim and D. Zilberman, Yield effects of genetically modified crops in developing countries, *Science*, 299:900–902, 2003.

Somatostatin

A naturally occurring regulatory peptide that carries out numerous functions in the human body, including the inhibition of growth hormone secretion from the anterior pituitary gland. Somatostatin consists of 14 amino acids (see **illus.**); the two cysteine residues are joined by a disulfide bond so that the peptide forms a ring structure. A larger variant of this peptide, called somatostatin-28, is produced in some cells and has an additional 14 amino acids attached at the amino-terminal end of normal somatostatin (somatostatin-14).

Somatostatin acts primarily as a negative regulator of a variety of different cell types, blocking processes such as cell secretion, cell growth, and smooth muscle contraction. This peptide was first discovered in the hypothalamus, a region of the brain known to control the release of hormones from the anterior pituitary gland. Somatostatin is secreted from the hypothalamus into the portal circulation and travels to the anterior pituitary gland, where it inhibits the production and release of both growth hormone and thyroid-stimulating hormone.

Many tissues other than the hypothalamus contain somatostatin, suggesting that this peptide has numerous roles. Somatostatin is present in widespread areas of the nervous system, including the cerebral cortex, basal ganglia, brainstem, spinal cord, and spinal sensory ganglia. There is evidence that it acts as a neurotransmitter or neuromodulator, providing chemical communication between nerve cells involved in the perception of sensory stimuli. Somatostatin is also present in high concentrations in the δ cells of the pancreas and has the ability to inhibit the release of both pancreatic insulin and glucagon. Somatostatin is widely distributed in the gastrointestinal tract, where it inhibits the secretion of several gut hormones, gastric acid, and pancreatic electrolytes and digestive enzymes. It inhibits the contraction of the stomach, intestine, and gallbladder, and reduces absorption of nutrients. Somatostatin also has a role in regulating the immune system, where it inhibits inflammatory processes.

Each of the functions of somatostatin is initiated by the binding of the peptide to one or more of five different cell-surface receptor proteins. These five receptor proteins have similar structures; each passes through the cell membrane seven times. Binding of somatostatin to these receptors activates one or more intracellular G-proteins and initiates biochemical signaling pathways within the cell. *See* SIGNAL TRANSDUCTION.

Analogs of somatostatin have been synthesized that are smaller, more potent, longer-lasting, and more specific in their biological effects than natural somatostatin. Some of these analogs, for example octreotide and lanreotide, have become useful as drugs. They are used in treating acromegaly because they can inhibit the excessive secretion of growth hormone by pituitary tumors (which causes this disease) and may even shrink the tumors in some patients. Somatostatin analogs are also used to treat malignant carcinoid syndrome, vasoactive intestinal

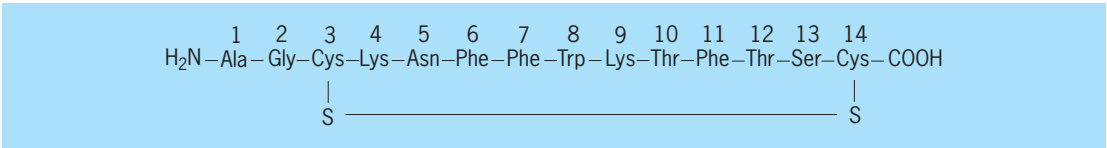

Amino acid sequence of somatostatin, including alanine (Ala), glycine (Gly), cysteine (Cys), lysine (Lys), asparagine (Asn), phenylalanine (Phe), tryptophan (Trp), threonine (Thr), and serine (Ser).

peptide (VIP)-secreting tumors, and other endocrine and neuroendocrine tumors that oversecrete hormones. Radiolabeled somatostatin analogs can be used to both localize and destroy tumors that produce an excessive number of somatostatin receptors. Newer analogs that specifically bind to just one of the five different somatostatin receptors should have more selective therapeutic responses with reduced side effects. *See* ENDOCRINE SYSTEM (VERTEBRATE); HORMONE; NEUROSECRETION; PITUITARY GLAND. William H. Simmons

Bibliography. P. Chanson, Somatostatin analogs in the treatment of acromegaly: The choice is now possible, *Eur. J. Endocrinol.*, 143:573–575, 2000; J. Epelbaum et al., The neurobiology of somatostatin, *Crit. Rev. Neurobiol.*, 8:25–44, 1994; E. P. Krenning et al., Radiolabelled somatostatin analogue(s) for peptide receptor scintigraphy and radionuclide therapy, *Ann. Oncol.*, 10(suppl. 2):S23–S29, 1999; Y. C. Patel, Somatostatin and its receptor family, *Frontiers Neuroendocrinol.*, 20:157–198, 1999; Z. Tulassay, Somatostatin and the gastrointestinal tract, *Scand. J. Gastroenterol.*, 33(suppl. 228):115–121, 1998.

Somesthesis

A general term for the somatic sensibilities aroused by stimulation of bodily tissues such as the skin, muscles, tendons, joints, and the viscera. Six primary qualities of somatic sensation are commonly recognized: touch-pressure (including temporal variations such as vibration), warmth, coolness, pain, itch, and the position and movement of the joints.

These basic sensory qualities exist because each is served by a different set of sensory receptors (the sensory endings of certain peripheral nerve fibers) which differ not only in their sensitivities to different types of stimuli, but also in their connections to structures within the central nervous system. Information is transmitted in the form of nerve impulses from the receptor endings to the central nervous system over the peripheral nerve fibers—the latter converging as spinal nerves and entering the central nervous system via the dorsal roots. The cell bodies of these peripheral nerve fibers are in dorsal root ganglia, located just outside the central nervous system. Fibers from regions in the limb or trunk enter the spinal cord, while those from the head and face enter the brainstem. The area of skin and deep tissues supplied by fibers from a single dorsal root is called a dermatome (**Fig. 1**). There is a pair of left and right dorsal roots for each cord segment. There is considerable overlap between cutaneous areas supplied by adjacent dermatomes, since a peripheral nerve projects into several roots. Thus the sensory loss resulting from transection of a peripheral nerve may be clearly outlined and complete in the tissue supplied by the nerve, but the loss following transection of a dorsal root is usually less noticeable. Usually three or more contiguous roots must be cut before there is a region of complete anesthesia. *See* CENTRAL NERVOUS SYSTEM.

Somatic pathways. The somatic sensory pathways are dual in nature. One major part, the lemniscal system, receives input from large-diameter myelinated peripheral nerve fibers (for example, those serving the sense of touch-pressure). Those originating in the trunk or limbs have branches that terminate on cells located within the dorsal horn of the spinal cord, and also branches that ascend within the dorsal columns to terminate on cells in the dorsal column nuclei (**Fig. 2**). Cells originating in these nuclei send fibers (axons) across the midline to the other side to project via the medial lemniscus to the ventrobasal nuclear complex in the thalamus. Thalamic cells send their axons to the primary somatosensory cortex, located in the postcentral gyrus and in the depths of the central sulcus. Axons of certain cells in the primary somatosensory cortex project to other cortical areas, including the secondary somatic sensory cortex, located in the upper bank of the lateral sulcus. Peripheral nerve fibers serving touch-pressure in the face and head terminate in the main sensory nucleus in the brainstem. Cells there project across the midline to ascend with the medial lemniscus to the thalamus.

The second major somatic pathway is called the anterolateral system. It receives input from small-diameter myelinated and unmyelinated peripheral nerve fibers carrying pain and temperature information. Those fibers from the limb and trunk terminate on cells within the dorsal horn of the spinal cord. Cells of origin there send their axons across the midline to ascend within the ventral (anterior) lateral part of the spinal cord called the anterolateral tract. These axons project upon areas in the brainstem, midbrain, and thalamus. Peripheral nerve fibers from the head and face serving pain and temperature terminate in the nucleus caudalis in the brainstem. Cells of origin there send their axons across the midline to project to areas in the brainstem, midbrain, and thalamus.

There are several differences between the anterolateral and lemniscal systems besides differences in anatomical projections and the sensory modalities they serve. These are revealed by electrophysiologically recording nerve impulses from single cells or by electrically stimulating local regions in the brain and spinal cords of animals or humans. The results of such studies indicate, first, that there is a precise topographical representation of the body form at each anatomical stage within the lemniscal system (for example, for the primary somatosensory cortex see **Fig. 3**), whereas that for the anterolateral system is less well organized. Second, the peripheral area of tissue projecting to each cell is small and continuous for the lemniscal system, but often large and discontinuous for the anterolateral system, indicating a better degree of spatial resolution in the lemniscal system. Third, the capacities of cells to respond to temporal variations in stimulation applied to the body (for example, the dynamic responses of certain cells to

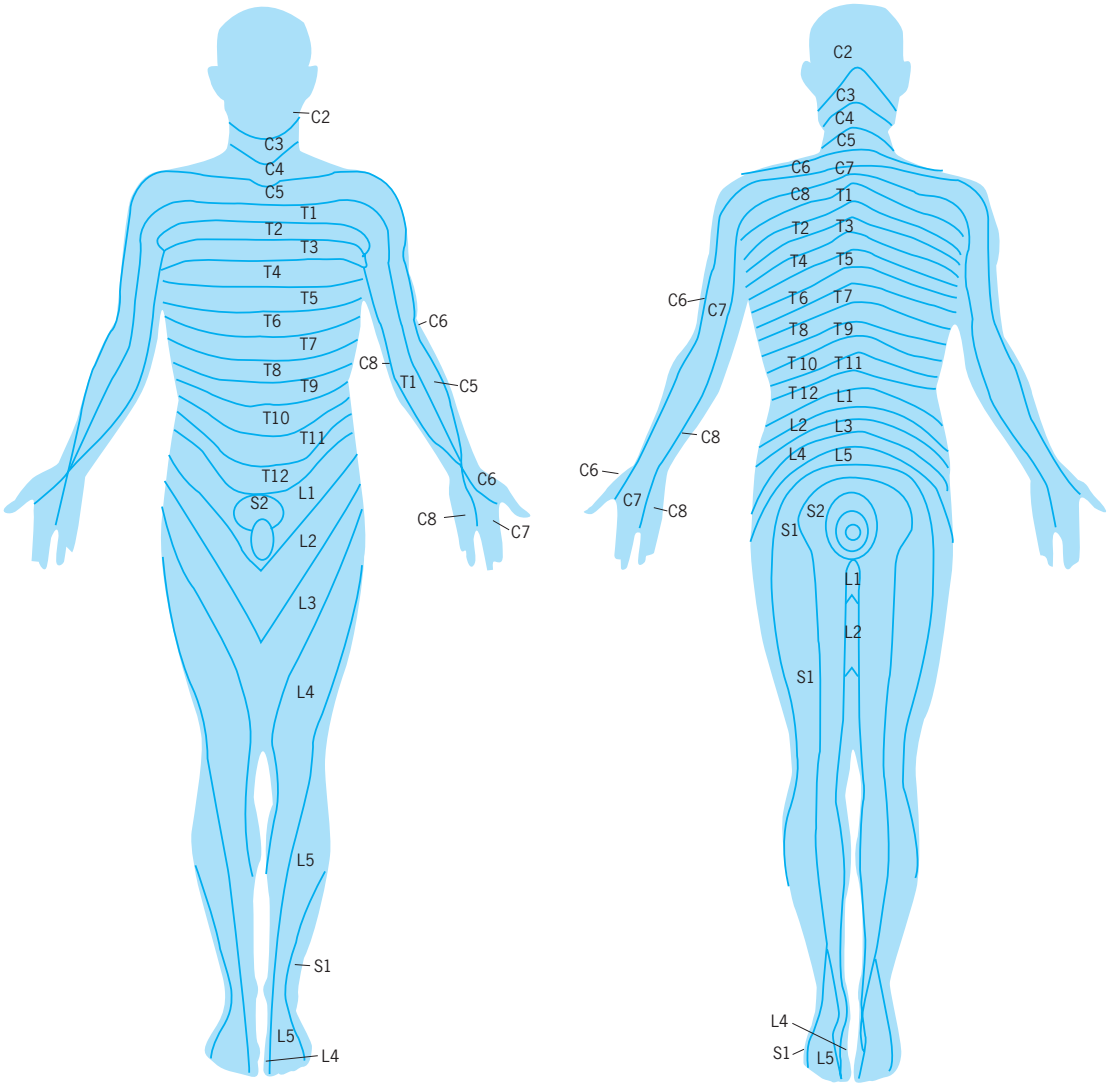

Fig. 1. Human dermatomes. The letters indicate the regions of the spinal cord: C, cervical; T, thoracic; L, lumbar; S, sacral. The numbers refer to segments of the spinal cord. (*After F. H. Netter, The CIBA Collection of Medical Illustrations, vol. 1: Nervous System, CIBA, 1967*)

a vibratory stimulus) are much more precise in the lemniscal system. These properties may reflect the greater requirement for temporal and spatial resolution in the touch-pressure sense, such as that needed to feel the spatial pattern of a finely textured cloth, than in the pain or temperature senses.

Somatosensory cortex. The functional organization of the somatosensory cortex occurs not only in the form of a representation of body geometry along its surface (Fig. 3), but also in the form of a vertical or columnar organization. All neurons encountered in a recording microelectrode penetration perpendicular to the cortical surface respond to stimulation of the same restricted area of the body. Further, all cells within a given column respond to the same type of stimulus, such as steady pressure applied to the skin or to movement of a joint or to warming the skin. Thus each column serves to process information of a particular stimulus modality arising from a specific location within or on the body.

Studies of the somatic sensory cortex have revealed that the somatotopic maps of the body are not static but can change as a result of sensory deprivation or stimulation. For example, after the amputation of the middle finger of one hand, the cells in the primary somatosensory cortex that had originally responded only to tactile stimulation of this digit become responsive to stimulation of the adjacent digits. Conversely, if an intact middle finger receives greater tactile stimulation than the adjacent fingers over an extended period of time, the cortical representation of this finger expands while that of each adjacent digit becomes smaller. Such mechanisms might contribute to the recovery of sensory function after a stroke (loss of blood supply to a local area resulting in cell death).

The reorganization of the sensory cortical map after sensory deprivation provides one possible explanation of why sensations from a body part, particularly a limb, may occur even in the absence of

the part. Sensations of the presence of a phantom limb in the place of a limb that was amputated, was traumatically deprived of its nerve supply, or was missing at birth are not uncommon. The phantom limb can appear quite lifelike and actually behave in a normal way, for example moving in coordination with an intact limb. Possibly, cells in primary and secondary somatosensory cortex that are denied the sensory input from the missing limb become endogenously active and are activated by input from adjacent, intact regions of the body, thereby giving rise to the phantom. It is not uncommon after the amputation of a limb for the phantom limb to be chronically painful. This medical condition is difficult to treat.

Lesions. Lesions (tissue destruction) of the anterolateral system may result in defects in pain and temperature sensibilities, with only mild losses in touch-pressure. A full surgical transection of the anterolateral column on one side of the spinal cord produces

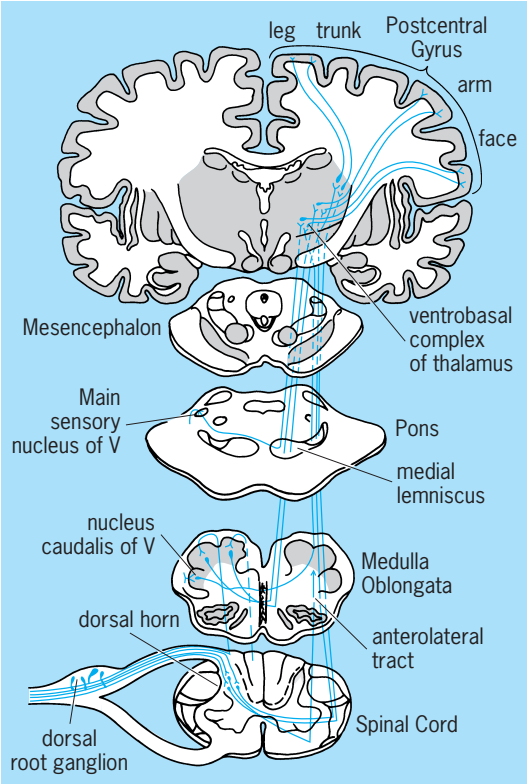

Fig. 2. Somatic afferent pathways. (*After A. Brodal, Neurological Anatomy in Relation to Clinical Medicine, Oxford University Press, 1969*)

a loss in sensitivity to pain and temperature stimuli on the opposite side of the body (whereas injury to the dorsal columns does not produce any such loss). Lesions of the lemniscal system, such as removal of SI, result in sensory defects in tactile discrimination, tactile localization, stereognosis (the perception of the shapes and sizes of objects), and the sense of

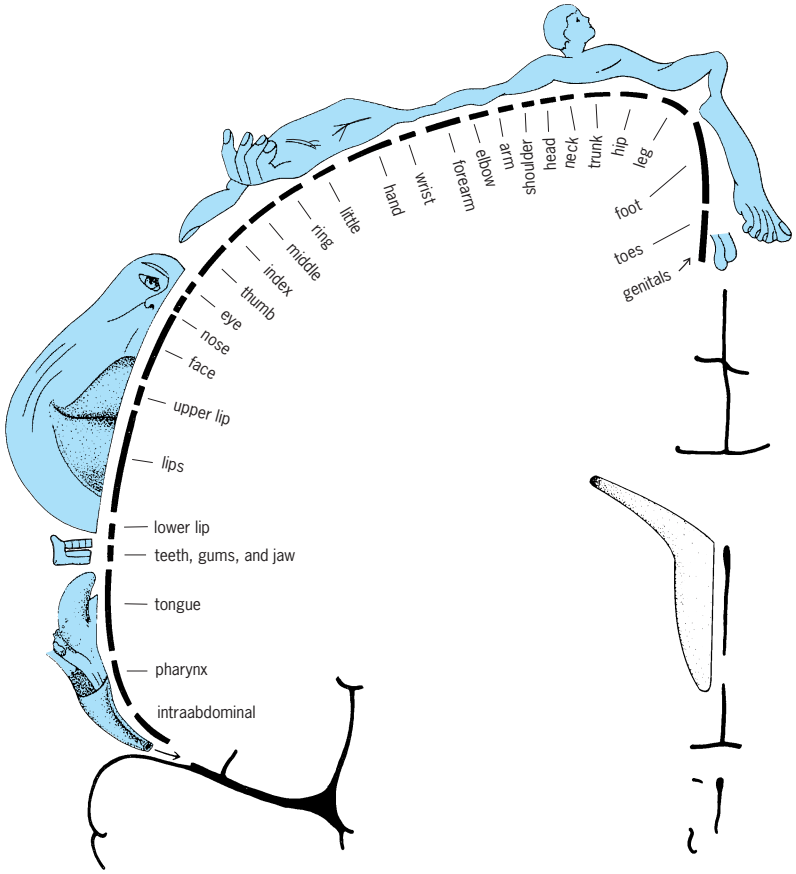

Fig. 3. Map of somesthetic areas in the human postcentral gyrus as determined by electrical stimulation (without anesthesia) and recording the subject's verbal description of his or her sensory experiences. The size of each part of the figurine is drawn approximately in proportion to the extent of sensory cortex devoted to it. (*After W. Penfield and T. Rasmussen, The Cerebral Cortex of Man, Hafner, 1968*)

position and movement of the limbs. *See* CUTANEOUS SENSATION; PAIN; PROPRIOCEPTION; PAIN.

Robert La Motte

Bibliography. E. R. Kandel, J. H. Schwartz, and T. M. Jessell, *Principles of Neural Science*, 4th ed., 2000; M. M. Merzenich et al., Somatosensory cortical map changes following digit amputation in adult monkeys, *J. Comp. Neurol.*, 224:591–605, 1994; R. F. Schmidt and G. Thews (eds.), *Human Physiology*, 2d ed., 1989.

Sonar

A remote sensing technique or device that uses sound waves to detect, locate, and sometimes identify objects in water. The term is an acronym for sound navigation and ranging. There are many applications, using a wide variety of equipment. Naval uses include detection of submarines, sea mines, torpedoes, and swimmers; torpedo guidance; acoustic mines; and navigation. Civilian uses include determining water depth; finding fish; mapping the ocean floor; locating various objects in the ocean, such as pipelines, wellheads, wrecks, and obstacles to navigation; measuring water current profiles; and

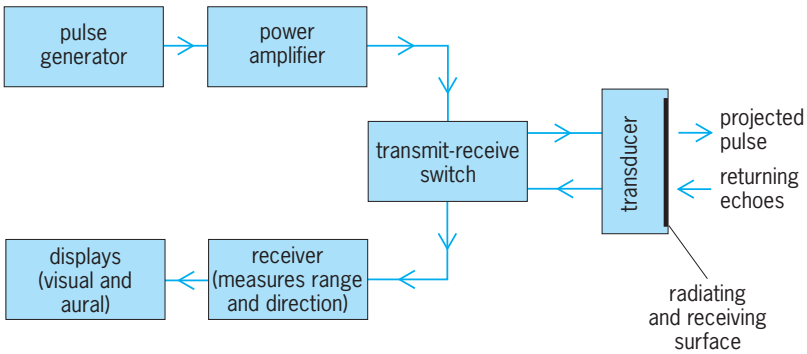

Fig. 1. Components of an elementary single-beam pulse sonar.

determining characteristics of ocean bottom sediments. Sound waves rather than electromagnetic waves (for example, radar and light) are used in these applications because their attenuation in seawater is much less. Some marine mammals use sound waves to find food and to navigate. *See* ACOUSTIC MINE; ACOUSTIC TORPEDO; ANTISUBMARINE WARFARE; ECHOLOCATION; MARINE GEOLOGY; MARINE NAVIGATION; UNDERWATER NAVIGATION; UNDERWATER SOUND.

There are two generic types of sonar: active (echolocation) and passive. An active sonar projects a signal (typically a short pulse of sound) into the water in a narrow beam, which propagates at a speed of about 1500 m/s (5000 ft/s). If there is an object (target) in the beam, it reflects a fraction of the sound energy to the sonar, which detects the echo. By measuring the elapsed time between projection and reception, the range to the target can be computed (range = sound speed × travel time ÷ 2). Direction to the target is determined from the orientation of the sound beam at the time of reception. Passive sonar does not radiate sound but depends on detecting sounds radiated by targets such as submarines and ships. Passive sonar determines direction to a target in the same manner as active sonar, but range determination is more difficult.

Active sonar. In an elementary active pulse sonar (**Fig. 1**), a pulse signal of certain frequency and duration is generated, amplified, and sent to an electroacoustic transducer, which converts the electrical signal into a sound signal, which then radiates into the water. If the transducer is reciprocal in character (typically the case), it also can be used to sense (detect) the returning echoes. The receiver amplifies the weak echoes and measures the range to each target, as well as the orientation of the receiving beam at the time of reception. This information is displayed in some form of range-direction plot.

Most active sonar transducers are mounted on the hulls of submarines or near the keels of surface ships. Sometimes, transducers are towed at a water depth that provides better operation. There are three basic transducer orientations. In the conventional depth sounder, the sound beam is directed downward. Echoes are reflected from the ocean bottom (and from fish that may be in the beam), and the depth of the ocean beneath the sonar can be determined. In the side-scan sonar configuration, the beam is oriented to the side of the ship (normal to the direction of travel) and (usually) slightly downward (**Fig. 2**). As the ship moves forward, a volume of water to the side of the ship is searched. Generally, two sonars are used, one searching to the right and one to the left. Side-scan sonars are well suited to search at a constant speed and along straight lines, such as in mapping the ocean bottom and in general searches of an area. The third, and most popular, sonar configuration involves rotating the sound beam about the vertical axis to search (scan) a sector of the water centered on the sonar platform. *See* ECHO SOUNDER.

Active sonar parameters. Sonar parameters such as operating frequency, bandwidth, beam widths, pulse duration, projected power level, and receiver sensitivity are interdependent. Ideally, the design goals are to maximize search rate, detection probability, accuracy of target location, and correctness of target classification. However, since these are conflicting goals, the parameters that optimize overall sonar performance for a specific application must be selected.

A restraint on all sonars is the speed of sound in the water. For example, if an elementary single-beam pulse sonar, known as a searchlight sonar, has a maximum search range of 6000 m (20,000 ft) and a beam width of 5°, then it takes 8 s to search a 5° sector (one beam width), and the beam cannot be rotated to search an adjacent volume until that

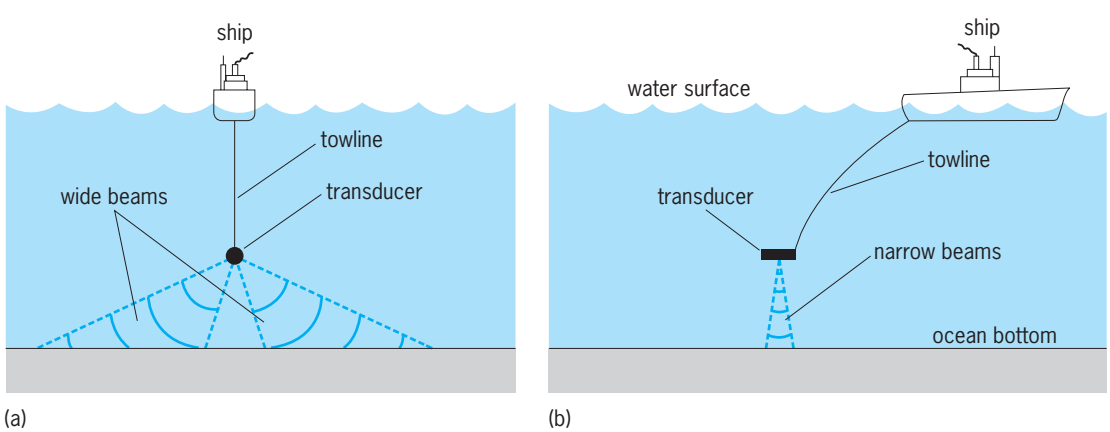

Fig. 2. Configuration for a two-beam side-scan sonar for bottom mapping. (*a*) End-on view. (*b*) Side view.

time has elapsed. About 9.6 min is required for a full 360° search in the horizontal plane. Most early active sonars were of the single-beam type, and many current small sonars still are. However, during and since World War II, much work has been done to increase search rates. The most popular technique is to project sound into a wide sector and to cover this sector with a fan of adjacent narrow receiving beams. In the example, 72 beams would be required. Rather than have 72 separate receivers, the output of each receiver beam is sampled sequentially by an electronic scanning switch and passed to a single receiver. In order not to leave gaps in the searched area, all 72 beams must be sampled during the pulse duration. If the sonar frequency is 15 kHz, the pulse needs to be at least 5 milliseconds or 75 cycles in duration in order to have at least one full cycle for each sample. An equivalent method is to electronically rotate (using phase-shift modulation) a single receiving beam through the search sector. Modern sonars, including advanced side-scan and depth sounders, employ large numbers of beams. *See* PHASE MODULATION.

Transducers. For radiation the transducer is called a projector (equivalent to an underwater loudspeaker), and for reception it is called a hydrophone (equivalent to an underwater microphone). Most electroacoustic transducers are reciprocal, meaning that they can be used for either function. The most popular type of electroacoustic transducer consists of a piezoelectric ceramic, such as barium titanate. For projection an alternating electric voltage is applied to electrodes on the ceramic, causing it to vibrate at the same frequency. For reception the alternating pressure of the echo causes the ceramic to generate an alternating electrical signal. Most practical transducers are formed from a number of small piezoelectric elements rather than a single large element and are called transducer arrays. By using many small elements, very large arrays can be fabricated in a variety of shapes (planes, cylinders, and spheres). Equally important, the elements can be connected in a variety of ways to form multiple beams simultaneously, some for projection and some for reception. *See* PIEZOELECTRICITY; TRANSDUCER.

The size and shape of the transducer, together with the sonar frequency or wavelength, determine the size and shape of the sound beam, both for projection and reception. For a given frequency the beam width decreases as the transducer size increases, and for a given transducer size the beam width decreases as the frequency increases. If the sonar is designed to search simultaneously in all directions in the horizontal plane, a suitable transducer might be a cylindrical array in which all elements are used to project, while groups of elements are connected to form multiple narrow beams for reception. Inhomogeneities in the ocean cause fluctuations in the speed of sound, which result in distortion of both the outgoing and returning sound waves and place an ultimate lower limit on the beam width; however, sonars with beams as narrow as 0.1° are generally feasible.

Frequency. The primary factor in selecting an operating sonar frequency is absorption of the sound energy by the water medium. This effect increases rapidly with increasing frequency (approximately as the frequency squared). Sonars used to locate submarines operate from about 3.5 to 35 kHz in order to achieve detection ranges of the order of 10 km (6 mi), and sometimes much greater depending on propagation conditions. Mine-hunting sonars must locate much smaller targets and operate at frequencies from about 100 kHz to 1.0 MHz, with ranges of a few hundred meters or less. Most other active sonars operate at frequencies within these extremes. *See* SOUND ABSORPTION.

Range. The range at which a target can be detected depends on the strength of the projected signal (source level), propagation losses to the target and return, reflection characteristics of the target (target strength), and sensitivity of the receiver. Also, the target echo must be stronger than various masking signals (noise and reverberation), which also are received by the sonar.

The ocean medium places an upper limit to the power that can be projected by a transducer. At low frequencies and shallow water depths, the limit likely will be due to cavitation bubbles, which essentially destroy the sound beam through absorption and scattering near the transducer. At high frequencies, the limit likely will be set by nonlinear characteristics of the water, which serve to convert any increase in power level into harmonics and heat. There is also a limit to the power that a transducer can handle without overheating or fracturing. *See* CAVITATION; NONLINEAR ACOUSTICS.

Propagation losses include absorption (discussed above) and scattering of the sound out of the beam because of inhomogeneities in the water. Finally, the sound beam spreads spherically as it travels from the projector and on the return path. This loss in intensity increases with the square of range for each path or as the fourth power of range for the round trip. For some portions of some propagation paths, the spreading is cylindrical rather than spherical, resulting in a loss proportional to the square of round-trip range. *See* SOFAR; SOUND.

Only a small fraction of the sound energy that strikes a target is reflected to the sonar. The rest is scattered in other directions or absorbed by the target. The ratio of reflected to incident intensity, known as target strength, generally increases with increasing target size, but also depends on target shape and composition, and is usually strongly dependent on target orientation, sometimes varying by a factor of 1000 depending on viewing angle.

The larger the receiving transducer, the more energy it intercepts from the returning wave, and thus the higher the sensitivity. However, the transducer also receives unwanted signals that tend to mask the desired ones. Ambient noise is a composite due to breaking surface waves, marine animals, seismic activity, and shipping. In general, ambient noise decreases with increasing frequency until thermal molecular noise dominates. Self-noise is

ship-generated noise, including turbulence as water flows past the transducer and ship.

Finally, the often dominant source of masking is the composite of unwanted echoes, known as reverberation, from inhomogeneities in the ocean volume such as animals, plants, and thermoclines; from the sea surface; and from the ocean floor. Generally, reverberation increases with frequency, but there are important exceptions. A reduction in receiver beam width results in a reduction in ambient noise and reverberation and an increase in receiver sensitivity, while a reduction in projector beam width results in an increase in source level.

Resolution. An important sonar performance characteristic is its ability to locate a target accurately and determine whether an echo is from a single target or from several targets close together. The uncertainty in the direction to the target is approximately the width of the receiving beam. Also, two targets at the same range can be resolved only if the receive beam width is less than the angle between the targets. Range accuracy and range resolution depend on the duration of the projected pulse. Two targets in the same direction at different ranges can be resolved if the separation is greater than one-half of the range extent of the pulse. There is a lower limit on pulse duration because the pulse must contain one or more full cycles of the sonar frequency (and many cycles for electronic scanning sonars). Also, the bandwidths of the receiver must equal the reciprocal of the pulse duration in order to receive the echoes without distortion. A wider bandwidth results in greater ambient noise. *See* BANDWIDTH REQUIREMENTS (COMMUNICATIONS).

Both range resolution and angular resolution are constant with range, but because of spherical spreading the beam width (in linear measure) increases linearly with range. At most operational ranges, range resolution is much better than cross-range resolution. If the resolution cell (range and cross-range) is small compared to the target, it is feasible to display a crude image of the target. Further, a target resting on or near the ocean floor blocks the sound beam and produces a reverberation-free acoustic shadow in back of the target. Some high-resolution sonars use target and shadow imaging for target classification.

Signal processing. There are many techniques to improve target detection against the masking of noise and reverberation, to increase detection ranges, and to improve target classification or identification. These include using complex pulses (for example, with shape and frequency modulation) and a wide variety of spatial and temporal frequency filters. Computer-aided detection, tracking, and target classification systems are often used. An important technique uses the Doppler phenomenon, whereby the echo from a moving target is distinguished from reverberation by its frequency shift relative to the incident frequency, and the approximate speed of target approach or recession can be determined. *See* ACOUSTIC SIGNAL PROCESSING; DOPPLER EFFECT.

Passive sonar. Passive sonars are used primarily to detect submarines and, to a lesser extent, surface ships. Because passive sonar does not radiate any sound that would reveal its location, it is the primary sensor used by submarines. The major weakness is that it cannot directly measure range to a target. To determine target location, the sonar must take bearings on a target from different locations. At short ranges (less than d^2/λ, where d is the length of the transducer and λ is the wavelength), the approximate range can be determined by measuring the curvature of the received sound wave.

Passive sonars depend on detecting noise radiated by targets, a mixture of sounds generated by propellers and hull vibrations (caused by motors, engines, pumps, and hydrodynamic forces). The noise has a continuous spectrum and discrete tones related to rotational speeds of propellers, engines, and so forth. By analysis of the received signals the sonar often can identify the type of target. Most of the radiated energy is in the audible frequency band and decreases in intensity with increasing frequency.

Most passive sonars use large receiving transducer arrays in order to achieve high sensitivity, discriminate against ambient noise, and determine precisely the direction to a target. In submarines, these arrays may be recessed into the structure or mounted on the hull. Submarines also may tow long slender line arrays. Typically, many beams are formed simultaneously to increase the search rate. Normally, surface combat ships make little use of passive sonars because of high self-noise levels. However, some special surface ships tow very long line arrays that maintain surveillance over large ocean areas. A number of very large fixed receiving arrays, placed on the ocean floor with cables running to shore stations, constantly observe strategically important ocean areas. Detection ranges for large passive sonars vary from hundreds of kilometers against noisy targets under good conditions to a very few kilometers against quiet targets.

The sonobuoy consists of a small surface buoy with a hydrophone array suspended beneath the water. Sounds received by the array are telemetered by radio link to an aircraft overhead. (There are also active sonobuoys.) Sonobuoys have relatively short detection ranges and are used primarily in tactical situations. *See* SONOBUOY; ULTRASONICS.

Chester M. McKinney

Bibliography. P. Blondel, *The Handbook of Sidescan Sonar*, Singer Praxis Books, 2004; H. E. Edgerton, *Sonar Images*, Prentice Hall, 1986; M. P. Hayes and P. T. Gough, Synthetic aperature sonar: a maturing discipline, in *Proceedings of the Seventh European Conference on Underwater Acoustics*, Delft, The Netherlands, pp. 1101-1104, 2004; H. T. Loeser, *Sonar Engineering Handbook*, Peninsula Publishing, 1984, reprint 1993; R. O. Nielsen, *Sonar Signal Processing*, Artech House, 1991; R. J. Urick, *Principles of Underwater Sound*, 3d ed., Peninsula Publishing, 1983, reprint 1996; O. B. Wilson, *Introduction to the Theory and Design of Sonar Transducers*, Peninsula Publishing, 1985, reprint 1989.

Sonic boom

An audible sound wave generated by an object that moves faster than the speed of sound (supersonic object). The sonic boom forms because the air is pushed away faster than the air molecules can move. The displaced air becomes highly compressed and creates a very strong sound wave, referred to as a compressional head shock or bow shock. At the back of the supersonic object the air has to fill the void left as the object moves forward; in this case, the gas becomes rarefied and a rarefractional tail shock develops. These shock waves are the main components of a sonic boom, and they are generated the entire time that an object flies faster than the speed of sound, not just when it breaks the sonic barrier. *See* SHOCK WAVE.

Sonic booms may be natural or generated by human activity. A natural sonic boom is thunder, created when lightning ionizes air, which expands supersonically. Meteors can create sonic booms if they enter the atmosphere at supersonic speeds. Human sources of sonic booms include aircraft, rockets, the space shuttle during reentry, and bullets. There are no reports of supersonic travel underwater; water induces a large drag that makes the concept impractical. However, a class of torpedoes is under development which may be able to travel supersonically; the torpedo creates a supercavitation bubble around itself which significantly reduces drag. Explosions can create shock waves; however, these shock waves spread spherically, like ordinary sound waves, and are not considered to be sonic booms. *See* METEOR; NAVAL ARMAMENT; THUNDER.

Supersonic aircraft. Sonic booms are commonly associated with supersonic aircraft. The most important parameter is the Mach number of the airplane, named after the Austrian physicist Ernst Mach (1838–1916). The Mach number is defined as $M = c_{plane}/c_0$, where c_{plane} is the speed of the airplane and c_0 the speed of sound of the air. At ground level, $c_0 =$ 343 m/s (1230 km/h or 770 mi/h). The first human to fly supersonically was Chuck Yeager, in a Bell X1 aircraft on October 14, 1947. The sonic barrier was broken on land by a British jet car, ThrustSSC, on October 13, 1997, just 50 years later. Two supersonic passenger airplanes have been built, the Anglo-French Concorde and the Soviet Tu-144, which is currently used only for experimental purposes. Both aircraft fly at Mach 2, that is, twice the speed of sound. The world's fastest airplane is the Lockheed SR-71 Blackbird, which was built in the early 1960s and can fly faster than Mach 3 (3220 km/h or 2000 mi/h) at an altitude over 26,000 m (85,000 ft). [The speed of sound decreases with altitude, principally because the upper atmosphere is colder. For example, the sound speed is 290 m/s (1050 km/h or 650 mi/h) at an altitude of 10,000 m (33,000 ft). Therefore, if an SR-71 were flying at 3220 km/h at sea level, it would attain only Mach 2.6.] *See* MACH NUMBER.

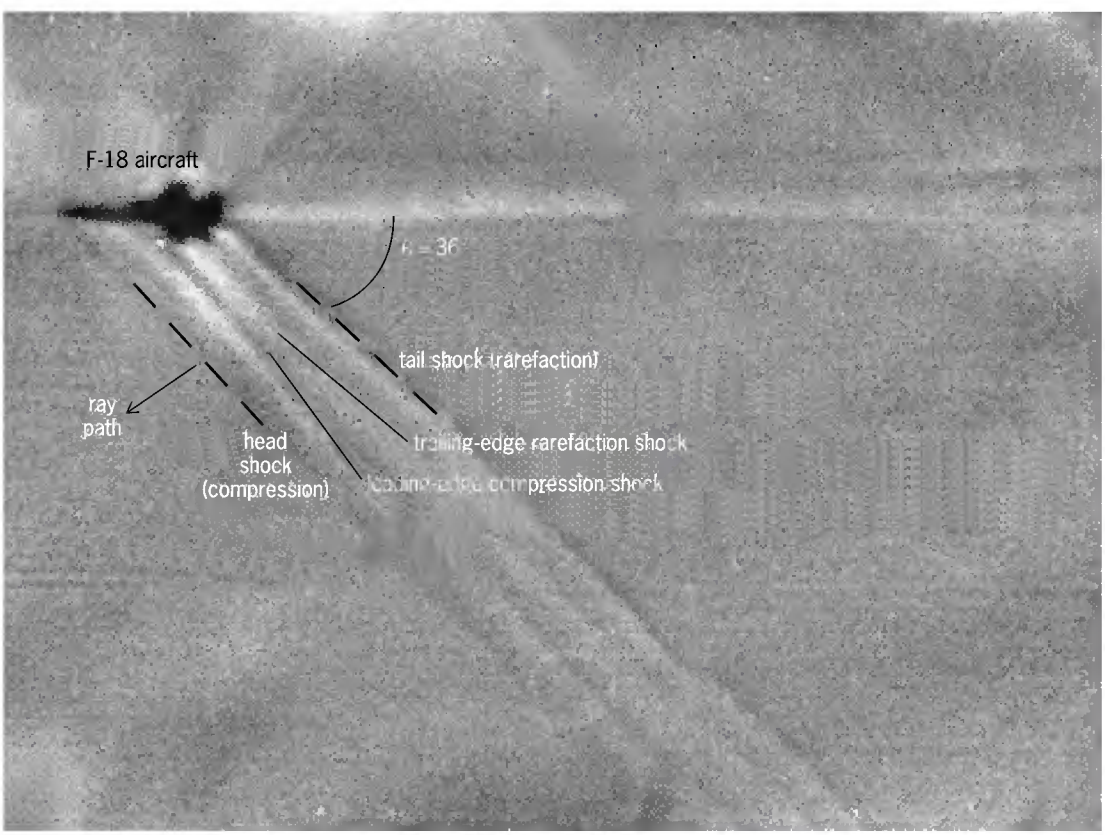

Fig. 1. Schlieren image of the shock waves generated by an F-18 aircraft flying at Mach 1.4. The Mach cone generated by the head and tail shocks can be seen as well as the shocks generated by the leading and trailing edges of the wings. (*NASA Dryden Flight Research Center*)

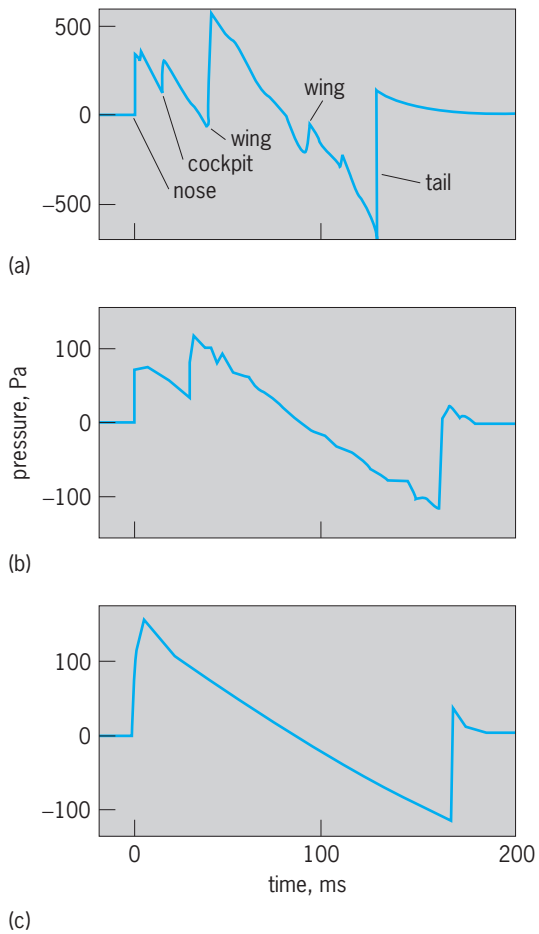

Fig. 2. Measured sonic boom waveforms from an SR-71 flying at Mach 2.4. (*a*) 165 m (540 ft) below the aircraft. Shocks associated with the nose, cockpit, leading edge of the wing, trailing edge of the wing, and the tail can be identified. (*b*) 1575 m (5075 ft) below the aircraft. Nonlinear distortion has resulted in the coalescence of some of the shocks. (*c*) At the ground, 9100 m (30,000 ft) below the aircraft, a classic N-wave signature with two shocks. (*NASA Dryden Flight Research Center*)

Mach cone. The shock waves associated with sonic booms propagate away from the aircraft in a unique fashion. These waves form a cone, called the Mach cone, that is dragged behind the aircraft. **Figure 1** shows the outline of the Mach cone generated by an F-18 fighter aircraft flying at Mach 1.4. The schlieren photographic technique was used to display the sonic boom, which is normally invisible. The half-angle of the cone is determined solely by the Mach number of the aircraft, $\theta = \arctan(1/M)$, 36° for $M = 1.4$. The shock waves travel along rays that are perpendicular to the Mach cone (Fig. 1). As the Mach number increases, θ becomes smaller and the sound travels almost directly downward. *See* SCHLIEREN PHOTOGRAPHY; SUPERSONIC FLIGHT.

Generation. The flow of air is highly complex near the aircraft body. Not only do the nose and tail of the aircraft create shock waves, but each leading edge, for example, the front of the wings, generates a compressive shock wave, and each trailing edge generates a rarefraction shock wave. The shock waves generated by the wings of the F-18 are indicated in Fig. 1. **Figure 2*a*** shows a waveform measured close to an SR-71. The compressional shocks generated by the nose, cockpit, and wing can be identified, as well as the rarefraction shocks from the trailing edge of the wing and the tail. The amplitudes of the shock waves are generally larger for faster aircraft and for larger aircraft, although the cross section of the aircraft is also important in determining the shock wave's strength; long, slender planes generate weaker shock waves than short, stubby planes.

Nonlinear distortion. As the sonic boom propagates along the rays away from the aircraft, the wave changes shape because the compression and rarefraction shocks are so strong that they change the speed of sound. The change in sound speed is small, less than 1%, but it is enough that the compression shocks generated by leading edges move faster than the head shock and can catch up to it. Conversely, the rarefraction shocks travel a little slower than the rest of the wave and eventually fall back into the tail shock. This nonlinear distortion is described by the theory of nonlinear acoustics. The nonlinear distortion results in the fine structure of the sonic boom being removed because the shocks coalesce as the boom propagates away from the aircraft (Fig. 2*b*). Eventually the sonic boom consists of a single head shock followed by a linear reduction in pressure to the single tail shock (Fig. 2*c*), a waveform referred to as an N-wave because of its shape. Near the aircraft the time interval between the head shock and the tail shock is $\Delta t = L_{\text{plane}}/c_0$, where L_{plane} is the length of the plane. The presence of nonlinear distortion can result in Δt increasing by as much as 25% by the time that the boom reaches the ground. For longer aircraft, flying at high altitudes (for example, the Concorde), Δt is greater than 0.3 s and the human ear can distinguish the head and tail shocks; that is, the sonic boom sounds like two very loud cracks, or like rifle shots in quick succession. For shorter aircraft (for example, the F-18) flying close to the ground, Δt is less than 0.1 s and the sonic boom sounds like a single bang. *See* NONLINEAR ACOUSTICS.

Atmospheric propagation. It can take up to 60 s for a sonic boom to travel from the plane to the ground (**Fig. 3**). As the sonic boom propagates away from the aircraft, it undergoes geometrical spreading and decreases in amplitude. If the sonic boom follows straight ray paths, the conical nature of the wavefront means that the spreading is cylindrical and the amplitude will decreases as $1/\sqrt{r}$, where r is the distance from the aircraft. In contrast, sound from a subsonic aircraft spreads spherically and the amplitude decreases as $1/r$, a much more rapid reduction. However, for sonic booms, nonlinear distortion adds extra losses, due to the presence of the shocks, and the amplitude reduction becomes $r^{-3/4}$, about half way between cylindrical and spherical spreading.

Refraction. The propagation of real sonic booms is more complex than a cylindrical spreading wave because the acoustic rays do not follow straight lines in the atmosphere. The rays undergo refraction due to the change in sound speed with altitude. Refraction in acoustics is directly analogous to that in optics, and

acoustic waves also obey Snell's law. Refraction acts to bend rays away from regions of high sound speed and toward regions of low sound speed. Therefore, as the sonic boom travels toward the ground, where the sound speed is highest, the rays are bent away. Refraction in the atmosphere is strong enough that some rays never reach the ground; they are refracted back up into the atmosphere (a form of total internal reflection). The sonic boom is therefore heard only in a narrow strip below the aircraft, referred to as the sonic boom carpet. The width of the sonic boom carpet is about 1 km for each 200 m of altitude (1 mi for each 1000 ft). For example, a supersonic aircraft flying at 16,000 m (50,000 ft) will be heard only along a sonic boom carpet about 80 km (50 mi) wide. The loudness of the sonic boom varies across the carpet; it is loudest directly beneath the aircraft and decreases with lateral distance from the flight path. *See* REFRACTION OF WAVES.

Effects of wind and turbulence. Wind also affects the propagation of the sonic boom. In the upper atmosphere the winds are typically steady and at fixed altitude; the jet stream is the strongest high-altitude wind. Wind changes the local sound speed and therefore can both affect the width of the sonic boom carpet and move this strip in the lateral direction. Close to the ground, typically within 1 km (0.6 mi), a turbulent region exists referred to as the atmospheric boundary layer. In this region, the wind and the ground interact to make the air move in a chaotic eddying motion. These turbulent eddies can strongly perturb the sonic boom waveforms. When the turbulence is strong, the N-wave structure can be broken up to produce very ragged looking waveforms. On occasion the turbulence can even focus sonic booms, resulting in intensified sound. *See* ATMOSPHERIC ACOUSTICS; BOUNDARY LAYER FLOW; TURBULENT FLOW.

Super booms. The nature of the sonic boom can be affected by aircraft maneuvers. When a supersonic aircraft accelerates, climbs, or turns, the sonic boom no longer forms a conical wavefront, but generates a curved wavefront. Such a wavefront can lead to focusing, resulting in what are termed super booms, with amplitudes up to twice that of booms generated by steady flight. These super booms can occur outside the sonic boom carpet; for example, when the Concorde decelerates over the Atlantic as it approaches Paris, a focused sonic boom can be heard in Sweden.

Impacts on the ground. The typical peak pressure amplitude (or overpressure) of a sonic boom on the ground is about 50–100 pascals (Fig. 2*c*). A sonic boom with 50 Pa (1 lbf/ft^2 or 0.007 psi) overpressure will produce no damage to buildings. Booms in the range of 75–100 Pa are considered disturbing by some people. Occasionally there is minor damage to buildings from sonic booms in the range of 100–250 Pa; however, buildings in good condition will be undamaged by overpressures up to 550 Pa. Very low flying aircraft (30 m or 100 ft) can produce sonic booms of 1000–7000 Pa. These pressures are still about five times less than that needed to injure the human ear, but can lead to damage to buildings, such as the breaking of glass windows and the cracking of plaster.

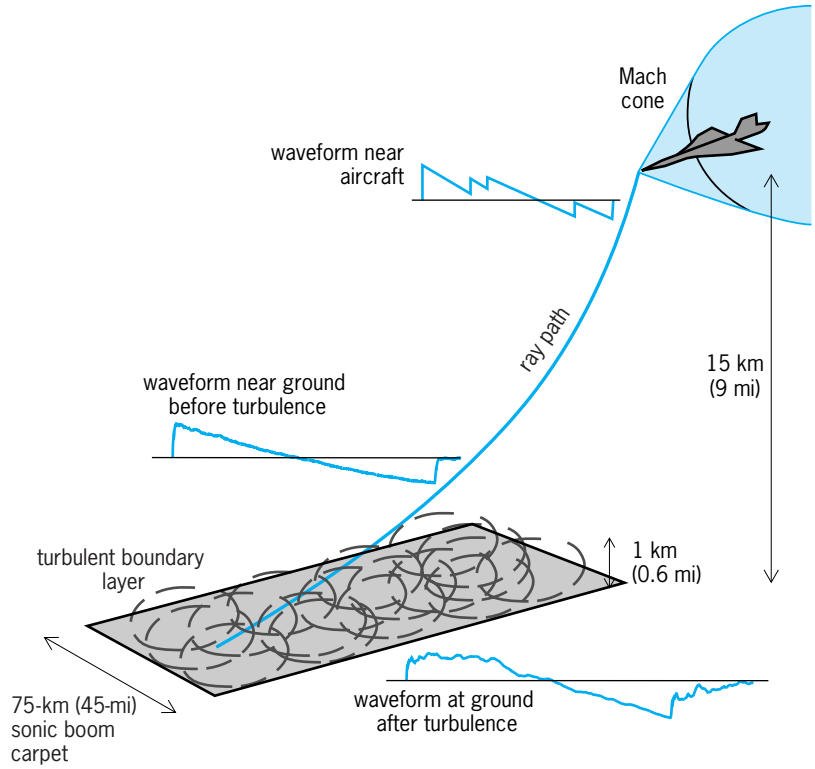

Fig. 3. Propagation of a sonic boom through the atmosphere from an aircraft flying at 15,000 m (50,000 ft).

Although sonic booms are not dangerous, they can evoke a strong startle response in people and animals. A consequence of the nonlinear distortion is that the shocks at the front and back of the sonic boom have a high-frequency content and sound like a gunshot. The high-frequency components are considered very annoying by people who are outdoors. People within houses are normally disturbed by the rattling effect of the sonic boom on the building. Experiments have shown that regular impacts by sonic booms generated by current aircraft are not acceptable to the general population, and so far it has not been possible to design aircraft that can fly supersonically and produce a boom that is acceptable to people underneath the flight path. For this reason, civilian aircraft are not permitted to generate sonic booms overland; therefore, the Concorde's flight path over the Atlantic is chosen to be longer than that for regular subsonic flights, to ensure that the plane is over water for the longest time and therefore able to remain in supersonic flight. Most military aircraft carry out supersonic flights over open water at altitudes greater than 3000 m (10,000 ft) and at least 23 km (15 mi) from shore. Supersonic operations over land are conducted above 10,000 m (30,000 ft) except for some specially designated areas, for example, near Fallon, Nevada, where low-level supersonic flights are permitted. Robin Cleveland

Bibliography. C. M. Darden et al., *Status of Sonic Boom Methodology and Understanding*, NASA CP

3027, 1989; A. D. Pierce and D. J. Maglieri, Effects of atmospheric irregularities on sonic-boom propagation, *J. Acous. Soc. Amer.*, 51:702–721, 1972; K. J. Plotkin, Review of sonic boom theory, *AIAA Pap.*, no 89-1105, April 1989; H. E. von Gierke and C. W. Nixon, Human response to sonic boom in the laboratory and the community, *J. Acous. Soc. Amer.*, 51:766–782, 1972.

Sonobuoy

An expendable device that enables aircraft to detect underwater objects, such as submarines, acoustically. Acoustics is the preferred energy form for use in salt water, because it tends to be the least attenuated by the medium.

A sonobuoy consists of an electronic radio link and antenna connected to a miniature sonar system. It contains the means for its launch from the aircraft, its entry into the water, separation of a floating antenna from the underwater transducer and sonar, and activation of a seawater battery upon entry, as well as a scuttling means for final sinking upon completion of its intended useful life (see **illus.**). The short life requires that each component be highly reliable and effective at low cost. The package must also be small and lightweight, since large numbers of packages are to be carried on an aircraft. *See* SONAR.

The aircraft is an extremely mobile platform that can rapidly follow up on an initial detection and classification of a submarine. Successive sonobuoy deployments permit tracking the submarine's course and taking or directing offensive action against a submarine target. *See* ANTISUBMARINE WARFARE; UNDERWATER SOUND.

The simplest sonobuoy is a passive sonar that senses sound with its hydrophone, amplifies and converts it to a radio signal, and transmits the signal from its antenna for analysis, evaluation, and storage in the aircraft. The sonobuoy system may be independently activated by an underwater sound source, usually an explosive device dropped from the aircraft. Two or more simple sonobuoys may be deployed to permit processing of directional information, passively or actively.

A sonobuoy that can transmit sound into the water and perform echo ranging on a submarine usually requires one of the larger standard-size configurations and larger launching tubes on the aircraft. Also, a command feature aboard the aircraft and a radio receiver in the sonobuoy are required. Directional transmission may be achieved using elementary superdirective beams such as dipoles and cardioids. Multiple sonobuoys and superdirective beamforming may be combined for sound reception. *See* DIRECTIVITY.

An Englishman, P. M. S. Blackett, proposed the concept of the sonobuoy in 1941. The original development was transferred to the National Defense and Research Committee of the United States, because British research gave higher priority to other tasks. By the end of World War II, substantial numbers of

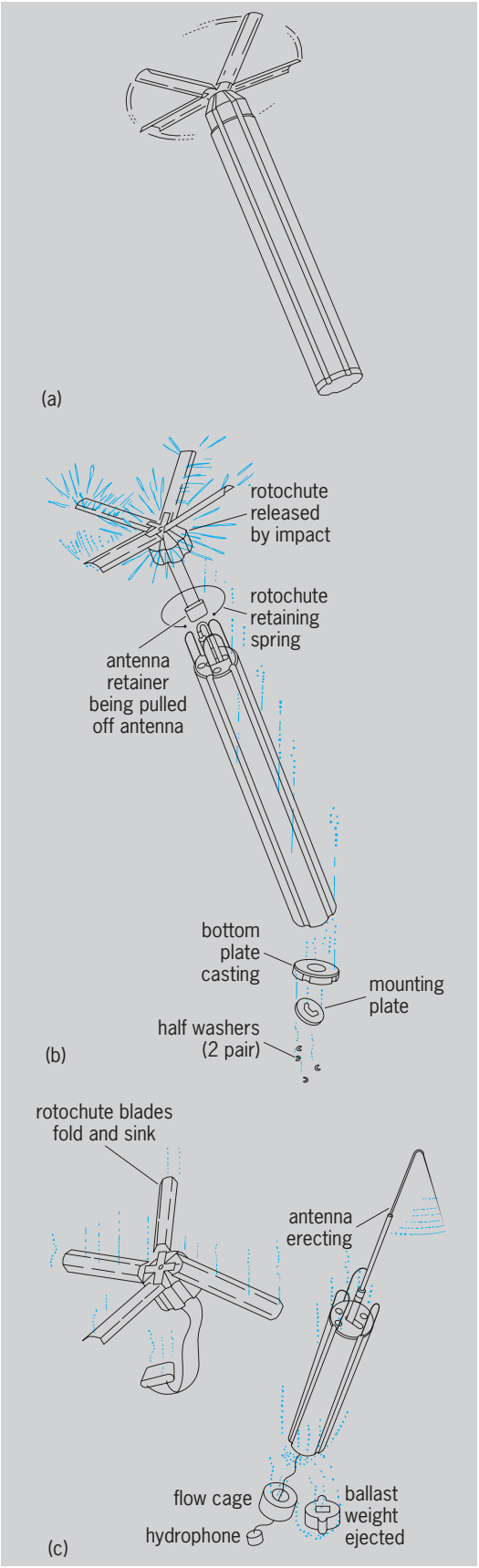

Launch of a sonobuoy from an antisubmarine warfare (ASW) aircraft. (*a*) In flight. (*b*) Buoy strikes the water in a nearly vertical attitude. (*c*) Antenna and hydrophone are positioned automatically.

sonobuoys had been designed, manufactured, and used in antisubmarine warfare (ASW).

Buoys of various types are used as sensors for oceanographic data such as sound speed, geophysical data such as earthquakes, bioacoustic data such as snapping shrimp, and other signal and noise sources. Some have been designed for long-term deployment on station, with provision for storing data until interrogated from an aircraft or a ship. Sonobuoys providing some of these functions have also been designed and deployed, and new applications can be anticipated. *See* BUOY. Stanley L. Ehrlich

Bibliography. T. A. Stefanick, *Strategic Antisubmarine Warfare and Naval Strategy*, 1987; R. J. Urick, *Principles of Underwater Sound*, 3d ed., 1983, reprint 1996.

Sonochemistry

The study of the chemical changes that occur in the presence of sound or ultrasound. Industrial applications of ultrasound include many physical and chemical effects, for example, cleaning, soldering, welding, dispersion, emulsification, disinfection, pasteurization, extraction, flotation of minerals, degassing of liquids, defoaming, and production of gas-liquid sols. *See* SOUND; ULTRASONICS.

Surprisingly, when liquids are exposed to intense ultrasound, high-energy chemical reactions occur, often accompanied by the emission of light. There are three classes of such reactions: homogeneous sonochemistry of liquids, heterogeneous sonochemistry of liquid-liquid or liquid-solid systems, and sonocatalysis (which overlaps the first two). In some cases, ultrasonic irradiation can increase reactivity by nearly a millionfold. Especially for liquid-solid reactions, the rate enhancements via ultrasound have proved extremely useful for the synthesis of organic and organometallic compounds. Because cavitation occurs only in liquids, chemical reactions are not generally seen in the ultrasonic irradiation of solids or solid-gas systems.

Mechanisms. Ultrasound spans the frequencies of roughly 20 kHz to 10 MHz (human hearing has an upper limit of less than 18 kHz). Since the velocity of sound in liquids is approximately 5000 ft/s (1500 m/s), ultrasound has acoustic wavelengths of roughly 7.5–0.015 cm. Because the wavelength of ultrasound are so much larger than molecular dimensions, sound does not cause such chemistry in a direct interaction. Instead, the chemical effects of ultrasound derive from several different physical mechanisms, depending on the nature of the system. For both sonochemistry and sonoluminescence, the most important of these mechanisms is acoustic cavitation: the formation, growth, and implosive collapse of bubbles in liquids irradiated with high-intensity sound (**Fig. 1**). During the final stages of cavitation, compression of the gas inside the bubbles produces enormous local heating and high pressures (but with very short lifetimes, in otherwise cold liquids). In clouds of cavitating bubbles, these hot spots have

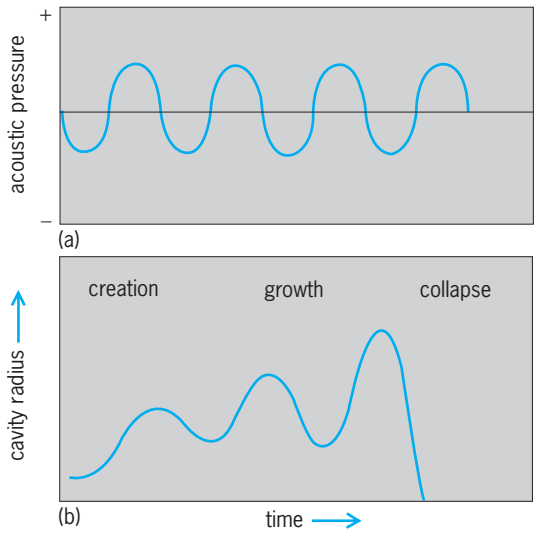

Fig. 1. Idealized representation of (*a*) expansion and compression waves and (*b*) the formation, growth, and collapse of a bubble in the process of acoustic cavitation. (*After K. S. Suslick, in F. G. A. Stone and R. West, eds., Advances in Organometallic Chemistry, vol. 25, 1986*)

equivalent temperatures nearly as high as the surface of the Sun (roughly 5000 K or 8540°F), pressures as large as at the bottom of the ocean (about 1000 atmospheres or 10^8 pascals), and lifetimes much less than a millionth of second. The dynamics of cavity growth and collapse are strikingly dependent on local environment, and cavitation in a homogeneous liquid should be considered separately from cavitation near an interface. The symbol $\xrightarrow{)))}$ will be used in this article to indicate ultrasonic irradiation (sonication or insonation). *See* CAVITATION.

When a liquid-solid interface is subjected to ultrasound, cavitation still occurs, but with major changes in the nature of the bubble collapse. No longer do cavities implode spherically. Instead, a markedly asymmetric collapse occurs, which generates a jet of liquid directed at the surface (**Fig. 2**). The jet velocities are greater than 330 ft/s (100 m/s). The origin of this jet formation is essentially a shaped-charge effect. The impingement of this jet can create a localized erosion (and even melting),

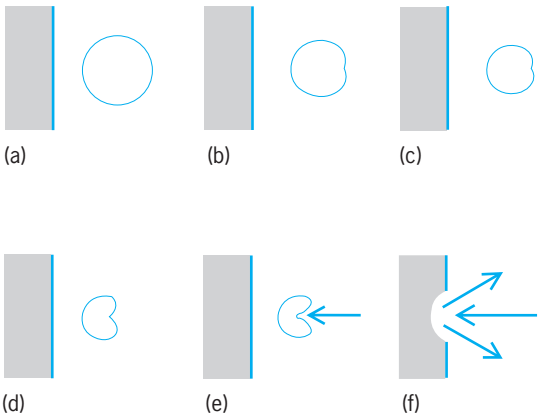

Fig. 2. Sequence *a–f* of a single bubble collapsing during cavitation near a surface.

responsible for surface pitting and ultrasonic cleaning. Enhanced chemical reactivity of solid surfaces is associated with these processes. The cavitational erosion generates unpassivated, highly reactive surfaces; causes short-lived high temperatures and pressures at the surface; produces surface defects and deformations; forms fines and increases the surface area of friable solid supports; and ejects material in unknown form into solution. The local turbulent flow associated with acoustic streaming improves mass transport between the liquid phase and the surface, thus increasing observed reaction rates.

Ultrasonic irradiation of liquid-powder suspensions produces another effect: high-velocity interparticle collisions. Cavitation and the shock waves that it creates in a slurry can accelerate solid particles to high velocities. The resultant collisions are capable of inducing dramatic changes in surface morphology, composition, and reactivity.

Homogeneous systems. A variety of novel reactivity patterns are distinct from either normal thermal or photochemical activation. Most of the reported reactions are stoichiometric in terms of a consumed reagent, but a few examples of true sonocatalysis have also appeared. The predominant reactions of homogeneous sonochemistry are bond breaking and radical formation. The effect of ultrasound on aqueous solutions has been studied for many years. The primary products are H_2 and H_2O_2; there is strong evidence for various high-energy intermediates, including HO_2, $H^\bullet$, $OH^\bullet$, and perhaps $e^-_{(aq)}$. Most recently there has been strong interest shown in the use of ultrasound to remediate low levels of organic contamination of water. The $OH^\bullet$ radicals produced from the sonolysis of water are able to attack essentially all organic compounds (including halocarbons, pesticides, and nitroaromatics) and, through a series of reactions, oxidize them fully. The desirability of sonolysis for such remediation lies in its low maintenance requirements and the low energy efficiency of alternative methods (for example, ozonolysis, ultraviolet photolysis). *See* HOMOGENEOUS CATALYSIS.

Sonochemistry also occurs in nonaqueous solutions. For example, the sonolysis of volalite organometallic complexes, whose thermal and photochemical reactivities were well characterized, gave quite unusual reactivities, including controlled multiple-ligand substitution and clusterification; examples are shown in reactions (1) and (2).

$$Fe(CO)_5 \xrightarrow{)))} Fe_3(CO)_{12} \quad (1)$$

$$Fe(CO)_5 + L \xrightarrow{)))} Fe(CO)_{5-n}L_n \quad (2)$$

$$(n = 1, 2, 3;\ L = \text{Lewis base})$$

See ORGANOMETALLIC COMPOUND.

Sonoluminescence. In addition to the initiation or enhancement of chemical reactions, irradiation of liquids with high-intensity ultrasound generates the emission of visible light. The production of such luminescence is a consequence of the localized hot spot created by the implosive collapse of gas- and vapor-filled bubbles during acoustic cavitation. In general, sonoluminescence may be considered a special case of homogeneous sonochemistry, and recent discoveries in this field have heightened interest in the phenomenon. In clouds of collapsing bubbles, emission from the excited states of atoms and molecules (very much as in flames) can be observed. For example, high-resolution sonoluminescence spectra were obtained from cavitation clouds in silicone oil under argon. The observed emission comes from excited-state diatomic carbon (C_2), and was modeled as a function of rotational and vibrational temperatures to find an effective cavitation temperature of 5050 ± 150 K (8630 ± 270°F). Recent work on metal atom sonoluminescence from volatile organometallics, such as iron pentacarbonyl [$Fe(CO)_5$], accurately confirms this temperature. Under conditions where an isolated, single bubble undergoes cavitation, recent studies on the duration of the sonoluminescence flash suggest that a shock wave may be created within the collapsing bubble, with the capacity to generate truly enormous temperatures and pressures within the gas—much higher than those seen with interacting clouds of bubbles. *See* CHEMILUMINESCENCE.

In a sense, chemistry is just the interaction of energy and matter. All chemical reactions require energy in one form or another to proceed. To a large degree, the properties of a specific energy source determine the course of a chemical reaction. Ultrasonic irradiation differs from traditional energy sources (such as heat, light, or ionizing radiation) in duration, pressure, and energy per molecule. The immense local temperatures and pressures and the extraordinary heating and cooling rates generated by cavitation bubble collapse mean that ultrasound provides an unusual mechanism for generating high-energy chemistry.

Heterogeneous systems. A major industrial application of ultrasound is emulsification. The first reported and most studied liquid-liquid heterogeneous systems have involved ultrasonically dispersed mercury. The use of such emulsions for chemical purposes has been delineated by extensive investigations. The effect of the ultrasound in this system appears to be due to the large surface area of mercury generated in the emulsion. *See* EMULSION.

The effects of ultrasound on liquid-solid heterogeneous organometallic reactions have been a matter of intense investigation. Various research groups have dealt with extremely reactive metals, such as lithium (Li), magnesium (Mg), or zinc (Zn), as stoichiometric reagents for a variety of common transformations. Examples are shown in reactions (3)–(6), where R represents an organic functional group.

$$RBr + R'R''CO \xrightarrow[Li]{)))} RR'R''COH \quad (3)$$

$$RBr + (CH_3)_2NCHO \xrightarrow[Li]{)))} RCHO \quad (4)$$

$$R_2SiCl_2 \xrightarrow[Li]{)))} \text{cyclo-}(R_2Si)_3 \quad (5)$$

$$RR'CO + BrCH_2CO_2R'' \xrightarrow[Li]{)))} RR'C(OH)CH_2CO_2R'' \quad (6)$$

Ultrasonic cleaning of the reactive metal surface to remove passivating impurities (for example, water, hydroxide, metal halide, or organolithium) is likely to be important in the origin of these effects.

The activation of less reactive metals continues to attract major efforts in heterogeneous catalysis, metal-vapor chemistry, and synthetic organometallic efforts. Given the extreme conditions generated by acoustic cavitation at surfaces, analogies to autoclave conditions or to metal-vapor reactors may be appropriate. It has been found that the use of ultrasonic irradiation facilitates the reduction of a variety of transition-metal salts to an extremely active form.

The effects of ultrasound on heterogeneous systems are quite general, and ultrasonic rate enhancements for many nonmetallic insoluble reagents also occur. Many organic reactions have been studied by various groups, including reductions with lithium aluminum hydride ($LiAlH_4$) or sodium hydride (NaH), oxidations with potassium permanganate ($KMnO_4$), and thioamidation with phosphorus pentasulfide (P_4S_{10}). Ultrasound has been used to enhance the rates of mass transport near electrode surfaces and thus to enhance rates of electrolysis. This has had some useful synthetic applications for the production of both organic and inorganic chalcogenides.

Applications of ultrasound to electrochemistry have also seen substantial recent progress. Beneficial effects of ultrasound on electroplating and on organic synthetic applications of organic electrochemistry have been known for quite some time. More recent studies have focused on the underlying physical theory of enhanced mass transport near electrode surfaces. Another important application for sonoelectrochemistry has been developed for the electroreductive synthesis of submicrometer powders of transition metals.

Applications to materials chemistry. Of recent interest is the development of sonochemistry as a synthetic tool for the creation of unusual inorganic materials (**Fig. 3**). As one example, the discovery of a simple sonochemical synthesis of amorphous iron helped settle the long-standing controversy over its magnetic properties. More generally, ultrasound has proved extremely useful in the synthesis of a wide range of nanostructured materials, including high-surface-area transition metals, alloys, carbides, oxides, and colloids. Sonochemical decomposition of volatile organometallic precursors in high-boiling solvents produces nanostructured materials (that is, built from nanometer-sized particles) in various forms. Nanometer colloids, nanoporous high-surface-area aggregates, and nanostructured oxide-supported catalysts can be prepared by this general route. The sonochemically prepared materials generally have very large surface areas, with surfaces that are extremely active as heterogeneous catalysts for a variety of important industrial reactions, including hydrogenations and dehydrogenations, hydrodesulfurization, and carbon monoxide reductions. Heterogeneous catalysts often require rare and expensive metals. The use of ultrasound offers some hope of activating less reactive, but also less costly, metals.

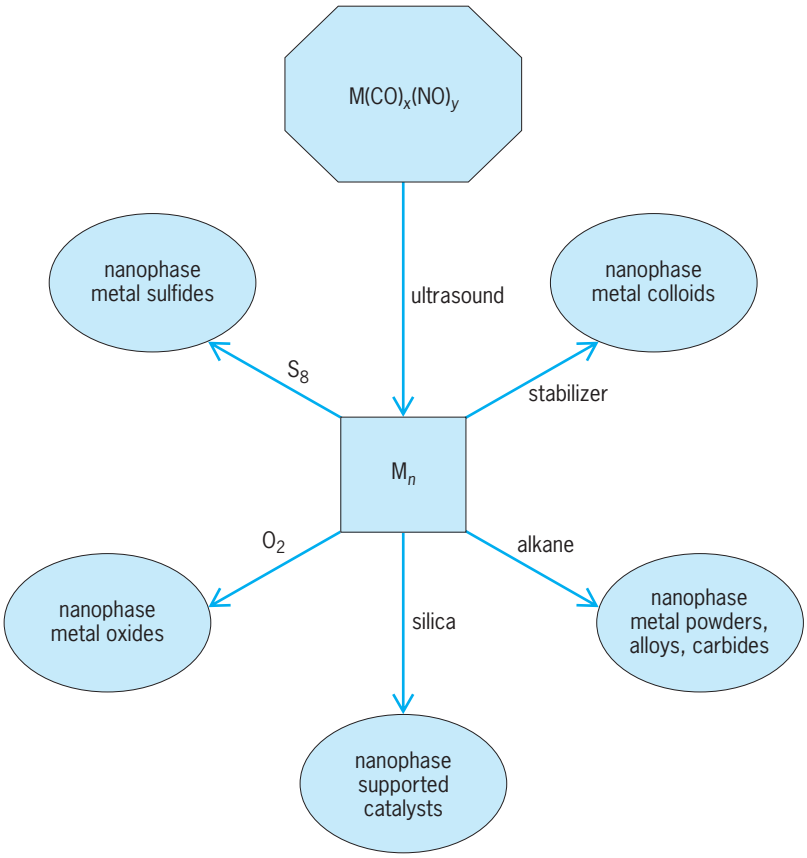

Fig. 3. Sonochemical synthesis of various forms of nanostructured materials. Starting with volatile precursors, such as $M(CO)_x(NO)_y$ where M can be any of several metals, ultrasound can be used to break the precursor apart and liberate metal atoms that form small metal clusters of 100–1000 atoms. These clusters are then used to form various nanostructured materials.

Such effects can occur during the formation of supported catalysts, by activation of preformed catalysts, or through enhancement of catalytic behavior during catalysis. *See* NANOCHEMISTRY; NANOSTRUCTURE.

Polymers. Sonochemistry is also proving to have important applications with polymeric materials. Substantial work has been accomplished in the sonochemical initiation of polymerization and in the

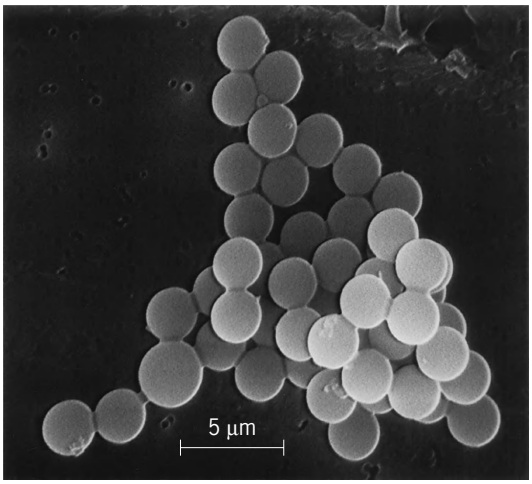

Fig. 4. Scanning electron micrograph of sonochemically synthesized hemoglobin microspheres. (*Courtesy of M. Wong, K. J. Kolbeck, and K. S. Suslick*)

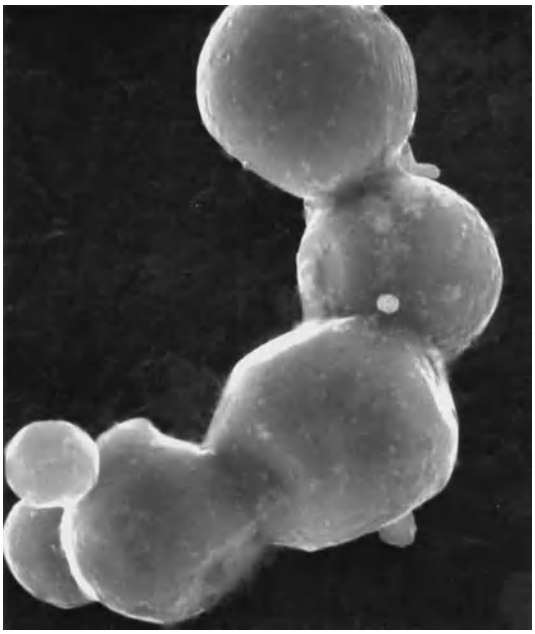

Fig. 5. Scanning electron micrograph showing particle agglomeration after ultrasonic irradiation of slurries of zinc powder (average diameter 5 μm). (*Courtesy of S. J. Doktycz and K. S. Suslick*)

modification of polymers after synthesis. The use of sonolysis to create radicals that function as radical initiators has been well explored. Similarly the use of sonochemically prepared radicals and other reactive species to modify the surface properties of polymers is actively being developed. Other effects of ultrasound on long-chain polymers tend to be mechanical cleavage, which produces relatively uniform size distributions of shorter chain lengths. *See* POLYMER.

Protein microspheres. Sonochemistry has found another recent application in the preparation of unusual biomaterials, notably protein microspheres (**Fig. 4**). Using high-intensity ultrasound and simple protein solutions, a remarkably easy method to make both air-filled microbubbles and nonaqueous liquid-filled microcapsules has been developed. Sonochemically prepared microspheres are stable for months, and being slightly smaller than red blood cells (erythrocytes), they can be intravenously injected and pass unimpeded through the circulatory system. The mechanism responsible for microsphere formation is a combination of two acoustic phenomena: emulsification and cavitation. Ultrasonic emulsification creates the microscopic dispersion of the protein solution necessary to form the proteinaceous microspheres. However, by itself, emulsification is insufficient to produce long-lived microspheres. The long life of these microspheres comes from a sonochemical cross-linking of the protein shell. Protein cysteine residues are oxidized during microsphere formation by sonochemically produced superoxide to create disulfide links between proteins (these same interprotein cross-links are responsible for the mechanical strength of hair, horn, and nails). These protein microspheres have a wide range of biomedical applications, including their use as echo contrast agents for sonography, magnetic resonance imaging contrast enhancement, and drug delivery. *See* PROTEIN.

Applications to heterogeneous catalysis. Ultrasonic irradiation can alter the reactivity observed during the heterogeneous catalysis of a variety of reactions. Sonication has shown such behavior by altering the formation of heterogeneous catalysts, by perturbing the properties of previously formed catalysts, or by affecting the reactivity during catalysis. However, ultrasonic rate enhancements of heterogeneous catalysis have usually been relatively modest (less than tenfold). The effect derived from irradiating operating catalysts is often due simply to improved mass transport. In addition, increased dispersion during the formation of catalysts under ultrasound will enhance reactivity, as will the fracture of friable solids (for example, noble metals on carbon or silica). It has been discovered that hydrogenation of alkenes by nickel powder is dramatically enhanced (by nearly a millionfold) by ultrasonic irradiation. Ultrasonic irradiation of liquid-solid slurries radically alters the surface composition of metal powders. The passivating oxide layer, which normally coats nickel and most metals, is rapidly removed by ultrasound. *See* CATALYSIS; HETEROGENEOUS CATALYSIS.

Scanning electron micrographs of metal powders after ultrasonic irradiation of slurries reveal the importance of interparticle collisions (**Fig. 5**). Such collisions are driven by the turbulent flow and shock waves that are created by high-intensity ultrasound. It appears likely that the observed changes in surface morphology, in the degree of aggregation, and in the surface's elemental composition originate from such interparticle interactions. The chemical effects derive in large part from the surface depassivation that results.

Kenneth S. Suslick

Bibliography. I. Hua, R. H. Hochemer, and M. R. Hoffmann, *Environ. Sci. Technol.*, 29:2790, 1995; T. G. Leighton, *The Acoustic Bubble*, Academic Press, London, 1994; K. S. Suslick, *MRS Bull.*, no. 20, 29, 1995; K. S. Suslick, Sonocatalysis, in G. Ertl et al. (eds.), *Handbook of Heterogeneous Catalysis*, vol. 3, Wiley-VCH, Weinheim, 1997; K. S. Suslick and L. A. Crum, Sonochemistry and sonoluminescence, in M. J. Crocker (ed.), *Handbook of Acoustics*, Wiley-Interscience, New York, 1998.

Sonoluminescence

The transformation of acoustic energy into light through collapsing bubbles. In particular, single-bubble sonoluminescence (SBSL) is the periodic light emission of a single acoustically trapped and periodically driven gas bubble. It occurs when the bubble collapses so violently that the energy focusing during collapse is intense enough for partial ionization of the gas inside the bubble to occur, leading to subsequent light emission. The phenomenon of single-bubble sonoluminescence was discovered in 1989 by Felipe Gaitan, then a graduate student, working with Lawrence A. Crum. A

micrometer-sized bubble is acoustically trapped in a water-filled flask at resonance. The driving pressure is typically $P_a = 1.2$–1.4 bar (1 bar $= 10^5$ Pa $\cong$ 1 atm), the driving frequency is 20–40 kHz, and the air saturation in the water is typically 20–40%. Once per cycle, at bubble collapse, the bubble emits a short pulse of light, that typically lasts 100–300 picoseconds. The origin of the light is thermal bremsstrahlung: During the adiabatic collapse, the gas inside the bubble is heated, presumably to about 15,000 K. Consequently, the gas partly ionizes. The ions and electrons interact, decelerate, and finally recombine, and the deceleration is connected with light emission. *See* BREMSSTRAHLUNG.

Though the energy-focusing power at the bubble collapse is about 12 orders of magnitude and the light emission is rather spectacular, it is negligible from an energy-balance point of view. The majority of the incoming acoustic energy is emitted again as sound (during the violent bubble collapse, and therefore at much higher frequencies), is converted into heat, or is eaten up by chemical reactions. Therefore single-bubble sonoluminescence can be understood as "illuminated cavitating bubble dynamics."

Bubble dynamics. The backbone of the theoretical understanding of single-bubble sonoluminescence is the Rayleigh-Plesset equation, given below. It describes the dynamics of a spherical bubble, with radius $R(t)$, in a liquid. Here σ denotes the surface tension, η the dynamic viscosity, P_0 the ambient pressure (usually 1 atm), $p_g(R(t))$ the gas pressure inside the bubble, and $P(t)$ the time-dependent external driving pressure, which in the case of single-bubble sonoluminescence is simply $P(t) = P_a \sin(2\pi f t)$, with the driving amplitude P_a and the driving frequency f.

$$R\ddot{R} + \frac{3}{2}\dot{R}^2 = \frac{1}{\rho}\left(p_g - P_0 - P(t) - 4\eta\frac{\dot{R}}{R} - \frac{2\sigma}{R}\right)$$

The typical behavior of the bubble radius in the sonoluminescent regime is shown in **Fig. 1**, together with the driving pressure: On pressure reduction, the bubble strongly expands to typically 10 times its equilibrium radius R_0. Once the pressure increases, the bubble violently collapses, followed by a "ringing" of the bubble at its eigenfrequency. As seen from Fig. 1, the Rayleigh-Plesset dynamics nicely describes the experimental data for the bubble radius.

The energy-focusing property of the Rayleigh-Plesset dynamics can be seen by neglecting all terms on the right side of Eq. (1), that is, by considering only the inertial terms, to obtain the simpler equation $R\ddot{R} + \frac{3}{2}\dot{R}^2 = 0$. This equation can be integrated to obtain the solution $R(t) = R_0[(t_* - t)/t_*]^{2/5}$, with the remarkable feature of a divergent bubble-wall velocity as t approaches the time t_* of the void collapse. It is this finite time singularity which leads to the adiabatic heating of the gas inside the bubble. The divergence is eventually cut off by the pressure building up inside the bubble and by the sound emission at bubble collapse.

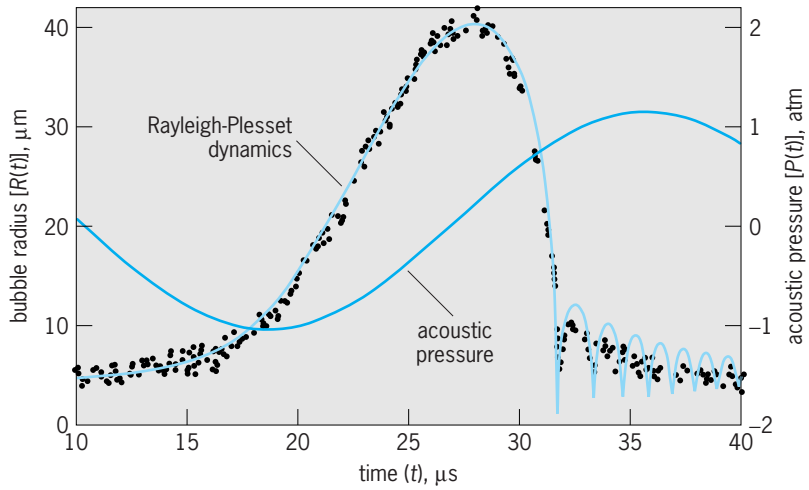

Fig. 1. Radius of a sonoluminescing bubble [$R(t)$, measured on the left axis] and acoustic pressure [$P(t)$, right axis] as functions of time. Rayleigh-Plesset dynamics (solid line) well describes the experimental data (circles) for the bubble radius.

The bubble dynamics also directly or indirectly determines the conditions under which stable single-bubble sonoluminescence can occur. These conditions are (1) that the threshold in the driving pressure for the Rayleigh collapse to occur be exceeded; (2) the shape stability of the bubbles (that is, the bubble must remain spherical and pinch-offs of microbubbles must not occur); (3) its diffusive stability (that is, the average number of gas molecules inside the bubble must not shrink or grow); and (4) the chemical stability of its ingredients. By applying these criteria, the phase diagrams of sonoluminescence can be calculated quantitatively (**Fig. 2**).

Chemical activity. The chemical activity inside the bubble is a consequence of the high temperatures

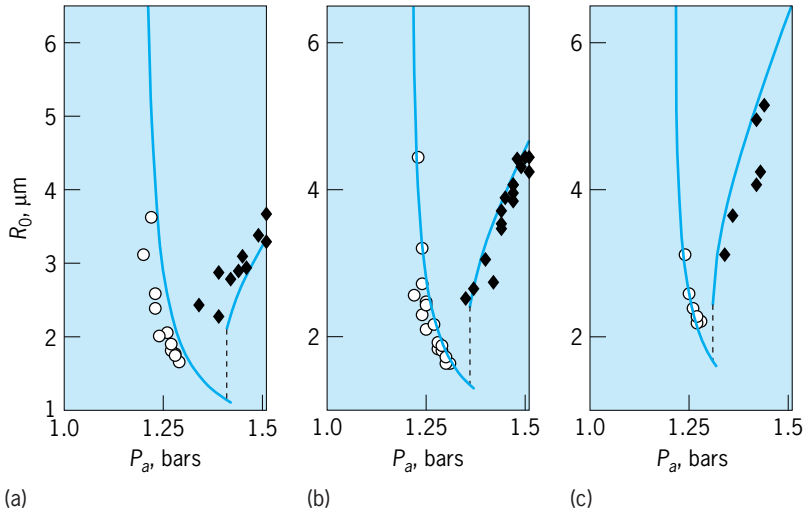

Fig. 2. Comparison of parameter-free model results (solid lines) and experimental data for air bubbles at (*a*) 10%, (*b*) 20%, and (*c*) 40% saturation in the ambient radius (R_0) versus forcing pressure (P_a) parameter space. Closed symbols refer to stable sonoluminescing bubbles, open symbols to stable bubbles that do not emit light. The ambient radius (that is, the radius of the bubble under normal pressure and temperature) is self-adjusting through diffusive processes like growth or shrinkage; there is no free parameter that the experimenter could choose to help the model results and the data agree. The driving frequency and the water temperature are $f = 33.4$ kHz, $T_0 = 293.15$ K $= 20^\circ$C $= 68^\circ$F.

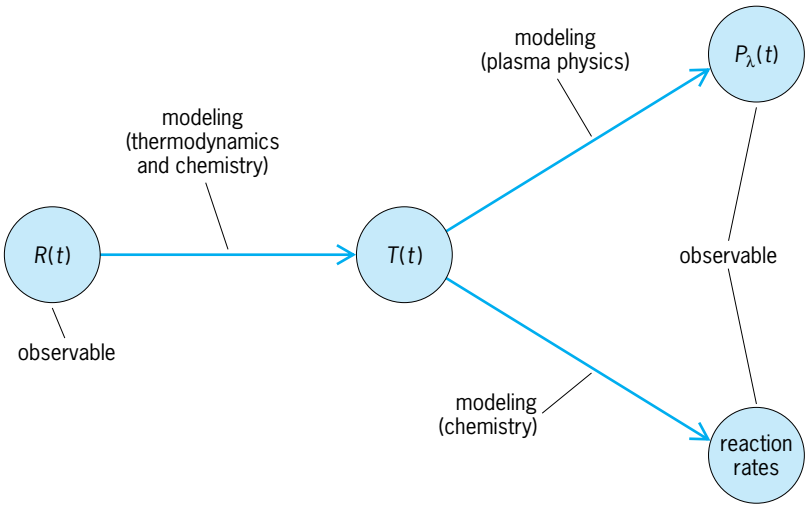

Fig. 3. Modeling steps that are involved in connecting the temperature $T(t)$ in the sonoluminescing bubble to observables. $T(t)$ is obtained from the radius dynamics $R(t)$, and in turn the spectral radiance $P_\lambda(t)$ and chemical reaction rates are deduced from the temperature. In 2005, the temperature of single-bubble sonoluminescence was directly measured and found to be 15,000 K.

achieved. The sonoluminescing bubble can be viewed as a high-temperature, high-pressure microlaboratory or reaction chamber, which can be controlled through the external parameters such as forcing pressure, frequency, or dissolved gas concentration. When the bubble is expanding, gas dissolved in the liquid and liquid vapor enter the bubble. During the adiabatic collapse, these gases are partly trapped inside the hot bubble and react. For example, nitrogen molecules will first dissociate to form nitrogen radicals and later react to form NH, NO, and so forth, which all dissolve readily in water when the bubble cools down and reexpands. Subsequent, reaction cycles, burn off further molecular gases. Eventually, all that remains in the bubble are the inert gases. In the case of air, this means argon. In this sense, stable sonoluminescing bubbles rectify argon.

Bubble temperature. The ultimate question in cavitation physics has always been how hot does the bubble get. That is, what limits the temperature increase, or in other words, what regularizes the singularity? The chief limiting processes are the sound emission from the collapsing bubble and the dissociation and eventually ionization of the gas inside the bubble, all of which consume energy. *See* SONOCHEMISTRY.

Direct measurement of the temperature in the bubble is difficult; measurements of most observable quantities enable only indirect deductions of the temperature, and modeling assumptions are necessary for its calculation. The situation is sketched in **Fig. 3**.

For the case of transient caviation, in which bubbles are destroyed at collapse and reform elsewhere (multibubble sonoluminescence or MBSL), Kenneth S. Suslick and coworkers succeeded in extracting the value of the temperature from observations of spectral lines in the early 1990s. Their result was 5000 K. However, MBSL light emission is rather weak compared to SBSL light emission, and so there was hope that much higher temperatures could be achieved in SBSL, nourished by the absence of any observed lines in SBSL spectra.

However, in 2005 Suslick and coworkers finally reported observing spectral lines in SBSL, enabling a direct temperature determination in that case as well. To achieve particularly strong SBSL, they used xenon and argon bubbles in sulfuric acid. This system combines three advantages: (1) high fluid viscosity, to allow the formation of large shape-stable bubbles; (2) low vapor pressure; and (3) monoatomic gases, so that the focused energy is not eaten up by the internal degrees of freedom of the gas–vapor mixture in the gas bubble. In this way, 1000 times more photons are emitted than with xenon or argon bubbles in water, allowing for good photon statistics and detailed spectral observations. From the atomic emission lines of argon, a temperature of 15,000 K could be deduced. Suslick and coworkers also observed strongly excited emissive states that are inconsistent with any thermal process, that is, light emission must originate from collisions of high-energy ions or electrons. This is the first direct experimental evidence that the light emission process of SBSL is thermal bremsstrahlung by an optically opaque plasma. Both the light-emitting process and the typical SBSL temperatures had been theoretically predicted, giving scientists confidence in the standard model of single-bubble sonoluminescence.

The reason for the much higher temperature in stable SBSL bubbles as compared to transient MBSL bubbles is that argon accumulates over hundreds of cycles in the SBSL case (which evidently is not possible if the bubble dies at collapse). Argon has an adiabatic exponent [equal to $(N + 2)/N$, where N is the number of degrees of freedom of the respective gas molecule] of 5/3, in contrast to nitrogen where it is only 7/5; water vapor has an even smaller adiabatic exponent. So, if the bubble is mainly filled with nitrogen, a large fraction of the energy from the collapse goes into rotational, vibrational, and dissociation energy, rather than leading to additional heating. Therefore, typically a temperature of only 5000 K can be achieved for transient cavitation. It is the bubble's stability in the SBSL case that allows for argon accumulation and therefore for higher temperatures of up to 15,000 K. Eventually, of course, further energy consuming processes take over, such as ionization (in addition to the rotation, vibration, and dissociation of the water vapor molecules, which are unavoidable). *See* ADIABATIC PROCESS; ISENTROPIC PROCESS.

Single-bubble sonoluminescence can be viewed as the hydrogen atom of cavitation physics: Single spherical bubble cavitation is the simplest building block of a sound-driven bubbly fluid, just as the hydrogen atom is the building block for more complicated atoms, molecules, or condensed matter. It is remarkable how many subdisciplines of physics and chemistry have been necessary to understand this conceptally simple building block: a single bubble oscillating in a sound field. These

subdisciplines range from acoustics to fluid dynamics, plasma physics, thermodynamics, atomic physics, spectroscopy, chemistry, dynamical system theory, and applied mathematics in general.

Detlef Lohse

Bibliography. B. P. Barber et al., Defining the unknowns of sonoluminescence, *Phys. Rep.*, 281:65, 1997; M. P. Brenner, S. Hilgenfeldt, and D. Lohse, Single bubble sonoluminescence, *Rev. Mod. Phys.*, 74:425, 2002; L. A. Crum, Sonoluminescence, *Phys. Today*, 47:22, 1994; D. J. Flannigan and K. S. Suslick, Plasma formation and temperature measurement during single-bubble cavitation, *Nature*, 434:52, 2005; S. Hilgenfeldt, S. Grossmann, and D. Lohse, A simple explanation of light emission in sonoluminescence, *Nature*, 398:402, 1999; W. C. Moss, D. B. Clarke, and D. A. Young, Calculated pulse widths and spectra of a single sonoluminescing bubble, *Science*, 276:1398, 1997; R. Toegel and D. Lohse, Phase diagrams for sonoluminescing bubbles: A comparison between experiment and theory, *J. Chem. Phys.*, 118:1863, 2003.

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

17 SOR-SUP

www.MHEST.com

Sorghum — Supramolecular chemistry

Sorghum

Sorghum includes many widely cultivated grasses having a variety of names. Sorghum is known as guinea corn in West Africa, Kafir corn in South Africa, mtama in East Africa, and durra in the Sudan. In India sorghums are called juar, jowar, or cholam; in China, kaoliang; and in the United States, milo. Cultivated sorghums in the United States are classified as a single species, *Sorghum bicolor*, although there are many varieties and hybrids. The two major types of sorghum are the grain, or nonsaccharine, type, cultivated for grain production and to a lesser extent for forage, and the sweet, or saccharine, type, used for forage production and for making syrup and sugar.

Grain Sorghum

Grain sorghum is grown in the United States chiefly in the Southwest and the Great Plains. It is a warm-season crop which withstands heat and moisture stress better than most other crops, but extremely high temperatures and extended drought may reduce yields. Sorghum responds well to optimum growing conditions, fertility, and management to produce large grain yields. It is extensively grown in Texas, Kansas, Nebraska, Oklahoma, Missouri, Colorado, and South Dakota. This grain production is fed to cattle, poultry, swine, and sheep primarily, with less than 2% going into nonagricultural markets such as starch, dextrins, flour, and industrial preparations. Sorghum is considered nearly equal to corn in feed value.

Origins and description. Sorghums originated in the northeastern quadrant of Africa. They have been grown in Africa and Asia for more than 2000 years. Introduction of a sorghum called chicken corn was made on the southern Atlantic coast in Colonial American times, but it was not successfully cultivated. The variety escaped and became a weed. Practically all grain sorghums of importance, until recent years, introduced into the United States were tall, late maturing, and generally unadapted. Since its introduction into the United States, the crop has been altered in many ways, these changes coming as a result of naturally occurring genetic mutations combined with hybridization and selection work of plant breeders. The rapid expansion in acreage came with the development of widely adapted varieties and later higher-yielding hybrids. The fact that hybrid grain sorghums with high yield potential could be produced with stems that are short enough for harvesting mechanically (**Fig. 1**) made the crop appealing to many farmers.

Grain sorghum is difficult to distinguish from corn in its early growth stages, but at later stages it becomes strikingly different. Sorghum plants may tiller (put out new shoots), producing several head-bearing culms from the basal nodes. Secondary culms may also develop from nodal buds along the

Fig. 1. Dwarf grain sorghum hybrids are grown throughout the sorghum belt of the United States because their short stems make them adaptable to mechanical harvesting.

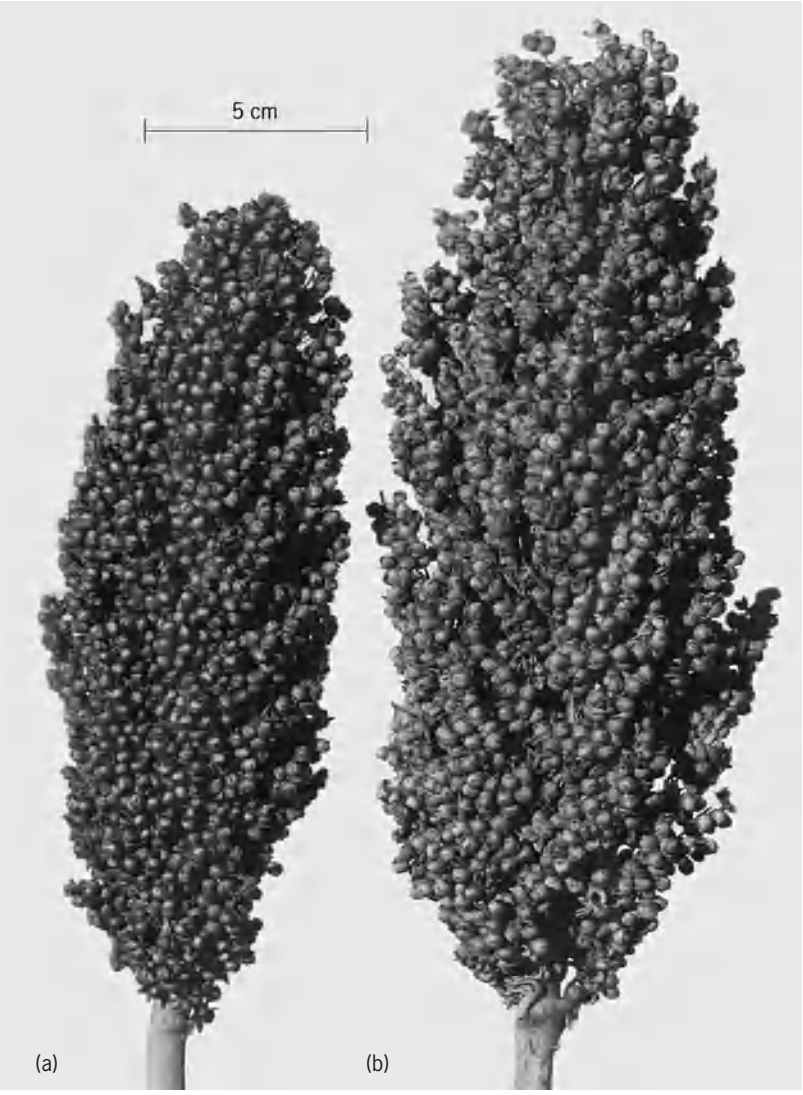

Fig. 2. Heads of grain sorghum. (*a*) Typical variety. (*b*) Typical hybrid.

main stem. The inflorescence (head) varies from a dense to a lax panicle, and the spikelets produce perfect flowers that are subject to both self- and cross-fertilization. The amount of natural cross-pollination ranges from 25 to 0% but averages about 5%. Mature grain in different varieties varies in size and color from white to cream, red, and brown. Color pigments are located in the pericarp (outer covering) of the grain or in a layer of cells beneath the pericarp called the testa. In some varieties the testa is absent; when the testa is present, however, the seed color is brown or some variation of brown. The endosperm (starch portion of the seed) is either white or yellow. There are endosperm modifications which cause the starch to be sugary or waxy, or to process a higher lysine content. The texture of the endosperm may vary from completely corneous to full floury. Most sorghums are intermediate in endosperm texture.

Grain sorghums are classified into types designated as milo, kafir, feterita, hegari, durra, shallu, kaoliang, and zerazera. This classification is based on morphological rather than cytological differences, since all types (including forage sorghums, broomcorn, and Sudangrass) have 10 pairs of chromosomes and freely cross.

Varieties and hybrids. A combine-height grain sorghum was developed before World War I but was not accepted by farmers or agriculturalists. Acceptance of combine grain sorghums was stimulated by the drought of the 1930s and by a farm labor shortage during World War II. A number of widely adapted productive varieties were developed during this period. Nonetheless, varieties disappeared rapidly when hybrids were introduced in the mid-1950s (**Fig. 2**).

The commercial production of seed of sorghum hybrids was made possible by the discovery of cytoplasmic male sterility in the early 1950s. This type of sterility, as used in corn and a few other crops, prevents the development of normal pollen grains and makes possible the formation of hybrid seed by cross-pollination. Because the flowers of sorghum are perfect, containing both staminate (male) and pistillate (female) parts, the production of commercial quantities of hybrid seed was not possible without a workable male-sterility mechanism. The first hybrid seed in quantity was sold to farmers in 1957, and within a period of less than 5 years hybrids had replaced most of the varieties previously grown. Sorghum hybrids yield at least 20% more than varieties, and concurrent improvements in cultural practices have combined to boost per-acre yield over 50% since their development. The emphasis on research in fertilization, irrigation, insect and disease control, and other areas has provided information to help account for the remarkable yield and production increases over the years.

Through a process of conversion many of the best varieties from around the world are being changed from tall, late, unadapted types to short, early, very useful cultivars. From this program plant breeders are expanding germ-plasm utilization and are developing parents of hybrids for the entire sorghum-producing world. *See* BREEDING (PLANT).

Planting. Grain sorghum seeds are small and should not be planted too deep since sorghum lacks the soil-penetrating ability of corn. A seeding depth of 1 in. (2.5 cm) is acceptable in moist and friable soil, but 2 in. (5 cm) may be necessary under dry soil conditions. The seeds are planted either in rows wide enough for tractor cultivation or in narrower rows if cultivation is not intended. Row planters for corn, cotton, field beans, and sugarbeets may be used when equipped with the proper seed plates. When wheat farming is practiced, much grain sorghum is planted with a grain drill with alternate or various feeder holes plugged to provide the desired row spacing. Soil temperature largely determines when the seed should be planted, assuming soil moisture conditions are adequate. Being tropical in origin, sorghum should not be planted in the spring until soil temperature is 65–70°F (18–22°C) at the planting depth and until there is little chance of subsequent lower temperatures. Dry-land grain sorghum is planted at a rate of 3–5 lb/acre (3.3–5.5 kg/hectare), and the rate is increased up to 10 lb/acre (11 kg/ha)

when planted under more favorable moisture conditions and irrigation.

Cultivation. Good seedbed preparation is essential for full stands and for weed control. Tilling fields improves soil structure in most cases and often aids in warming the soil. A rotary hoe is effective in controlling weeds when the plants are small. Subsequent cultivations are made as needed with the same equipment used for cultivating corn. Minimum tillage is practiced in many areas where control of weeds with chemicals is a part of the technology of sorghum production. When used correctly, herbicides are a boon to sorghum culture, but when used carelessly, disappointment may result. There are several chemical herbicides registered and approved for weed control in sorghum.

Harvesting. Nearly all grain sorghum is harvested standing in the field with a combine (Fig. 1). Harvest begins in southern Texas in early June and slowly proceeds northward. In the central and northern Great Plains, the crop is usually harvested after frost. The grain threshes freely from the head when the seed moisture content is 20–25% or lower. The grain should not contain more than 12% moisture to ensure safe storage after harvest. Grain dryers are often used when the grain at harvest is not dry enough for optimum storage. Proper storage must be maintained until the grain can be marketed. The industry's economic health lies in the ability to provide grain of the right quality in the right quantity at the proper time and place.

Frederick R. Miller

Sweet Sorghum

Commonly known as sorgo, sweet sorghum was introduced into North America from China in 1850, although its ancestry traces back to Egypt. It is an annual, rather drought-resistant crop. The culms are from 2 to 15 ft (0.6 to 4.6 m) tall, and the hard cortical layer, or shell, encloses a sweet, juicy pith that is interspersed with vascular bundles. At each node both a leaf and a lateral bud alternate on opposite sides; the internodes are alternately grooved on one side. Leaves are smooth with glossy or waxy surfaces and have margins with small, sharp, curved teeth. The leaves fold and roll up during drought. The inflorescence is a panicle of varying size having many primary branches with paired ellipsoidal spikelets containing two florets in each fertile sessile spikelet. The plant is self-pollinated. *See* CORTEX (PLANT); PITH.

Seed is planted in cultivated rows and fertilized similarly to corn. Maturity varies between 90 and 125 days. The juice contains about 12% sugar. The main sorghum-syrup-producing area is in the south-central and southeastern United States (**Fig. 3**).

Leonard D. Baver

Fig. 3. Sweet sorghum in Oklahoma. The scale indicates feet. 1 ft = 0.3 m. (*USDA*)

Diseases

Sorghums are plagued by a variety of diseases that vary in importance from year to year and among locations because of the environment, plant genotypes, cultural practices, variations in pathogens, or the interaction of any of these factors. These diseases may be classified into five general categories: those that rot the seed or injure seedling roots; those that attack the leaves, making the plants less productive; those that attack or destroy the grain in the heads; those that cause root and stalk rots; and those caused by viruses or viruslike organisms.

Fungi causing seed rotting and seedling diseases may be seed-borne or soil-inhabiting and are most destructive after planting, when the soil is cold and wet. Species of *Fusarium, Pythium, Helminthosporium*, and *Penicillium* are the most important fungi involved. Damage may be considerably reduced by planting sound hybrid seed treated with an approved fungicide in soil warm enough to ensure prompt germination.

Leaf diseases are caused by three species of bacteria and at least eight species of fungi. Many of these pathogens are favored by high temperatures and humid conditions, but a few are favored by cool, humid conditions. Disease lesions occurring as discolored spots or streaks may coalesce to involve the entire leaf. Rotation, seed treatment, and the use of resistant varieties are recommended control measures.

Three of the four known smuts of sorghum are found in the United States; covered kernel, loose kernel, and head smut. Kernel smuts, while historically important, are now controlled by routine seed treatments. Head smut (**Fig. 4**) destroys the entire head and continues to cause major crop losses. Resistant hybrids are providing control, although new strains of the fungus pathogen require the periodic development of new resistant hybrids.

Sorghum downy mildew (**Fig. 5**) has spread throughout the southern and central sorghum-growing regions. Losses result when the disease systemically invades the plant. Diseased plants are

barren. Fortunately, excellent resistance has been developed and used.

Maize dwarf mosaic, caused by an aphid-transmitted virus, spread throughout the sorghum-growing regions during the 1970s. Host tolerance reduces losses caused by this prevalent disease. Yellow sorghum stunt, caused by a mycoplasmalike organism and transmitted by leafhoppers, rarely reaches economically significant proportions.

Several diseases of the roots and stalks are of primary importance. Periconia root and crown rot, which caused extensive damage to milo and darso sorghums, is controlled by resistant varieties. *Pythium graminicola* causes a major root rot during periods of frequent rainfall in dryland sorghums. Charcoal rot, most evident as the plant approaches maturity under extreme conditions of heat or drought, causes shredding of the stalks and extensive lodging. Anthracnose, or red rot, develops in susceptible hybrids during wet years; plants lodge at the base of the peduncle. Development of resistant or tolerant hybrids appears to be the only effective method of control of the stalk rots if irrigation is not possible.

Grain mold is a disease of mature grain caused by species of *Fusarium* and *Curvularia*. These fungi infect at the flowering stage and rot the seed as it matures, particularly during wet weather at harvest time. *See* PLANT PATHOLOGY. Richard A. Frederiksen

Fig. 5. Sorghum downy mildew, a serious fungus disease that systematically invades sorghum plants, causing stripped leaves and barren stalks.

Fig. 4. Sorghum head smut, a disease that completely destroys the normal head and replaces it with masses of smut spores.

Bibliography. W. F. Bennett et al., *Modern Grain Sorghum Production*, 1990; H. Doggett, *Sorghum*, 1970; J. R. Quinby, *Sorghum Improvement and the Genetics of Growth*, 1974; J. S. Wall and W. M. Ross, *Sorghum Production and Utilization*, 1970; R. J. Williams, R. A. Frederiksen, and G. D. Bengston, *Proceedings of the International Sorghum Disease Workshop*, International Crops Research Institute for the Semi-Arid Tropics, Hyderabad, 1980; R. J. Williams, R. A. Frederiksen, and J. C. Girard, *Sorghum and Pearl Millet Disease Identification Handbook*, International Crops Research Institute for the Semi-Arid Tropics, Inform. Bull. 2, Hyderabad, 1978.

Sound

The mechanical excitation of an elastic medium. Originally, sound was considered to be only that which is heard. This admitted questions such as whether or not sound was generated by trees falling where no one could hear. A more mechanistic approach avoids these questions and also allows acoustic disturbances too high in frequency (ultrasonic) to be heard or too low (infrasonic) to be classed as extensions of those events that can be heard.

A source of sound undergoes rapid changes of shape, size, or position that disturb adjacent elements of the surrounding medium, causing them to move about their equilibrium positions. These disturbances in turn are transmitted elastically to

neighboring elements. This chain of events propagates to larger and larger distances, constituting a wave traveling through the medium. If the wave contains the appropriate range of frequencies and impinges on the ear, it generates the nerve impulses that are perceived as hearing. *See* HEARING (HUMAN).

Acoustic Pressure

A sound wave compresses and dilates the material elements it passes through, generating associated pressure fluctuations. An appropriate sensor (a microphone, for example) placed in the sound field will record a time-varying deviation from the equilibrium pressure found at that point within the fluid. The changing total pressure P measured will vary about the equilibrium pressure P_0 by a small amount called the acoustic pressure, $p = P - P_0$. The SI unit of pressure is the pascal (Pa), equal to 1 newton per square meter (N/m^2). Standard atmospheric pressure (14.7 lb/in.^{2}) is approximately 1 bar $= 10^6$ dyne/cm^2 $= 10^5$ Pa. For a typical sound in air, the amplitude of the acoustic pressure may be about 0.1 Pa (one-millionth of an atmosphere); most sounds cause relatively slight perturbations of the total pressure. *See* MICROPHONE; PRESSURE; PRESSURE MEASUREMENT; PRESSURE TRANSDUCER; SOUND PRESSURE.

Plane Waves

One of the more basic sound waves is the traveling plane wave. This is a pressure wave progressing through the medium in one direction, say the $+x$ direction, with infinite extent in the y and z directions. A two-dimensional analog is ocean surf advancing toward a very long, straight, and even beach. The surf looks like a long, corrugated surface advancing uniformly toward the shore but extending transversely in a series of parallel peaks and troughs. A plane wave has acoustic pressure controlled by the one-dimensional wave equation in cartesian coordinates, Eq. (1). An appropriate solution to this equation is Eq. (2), with f any function differentiable twice with

$$\frac{\partial^2 p}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} \tag{1}$$

$$p = f\left(t - \frac{x}{c}\right) \tag{2}$$

respect to x or t. This solution has the property that the acoustic pressure p has a single value for all pairs of x and t such that the phase $(t - x/c)$ is constant. At some point x_0 and time t_0, the acoustic pressure p has the value $p_0 = f(t_0 - x_0/c)$. As t increases from t_0 to $t_0 + \Delta t$, the value p_0 will move from x_0 to a new position $x_0 + \Delta x$, where Δx and Δt are related by Eq. (3). Solving for Δx in terms of Δt gives $\Delta x/\Delta t = c$,

$$(t_0 + \Delta t) - \frac{(x_0 + \Delta x)}{c} = t_0 - \frac{x_0}{c} \tag{3}$$

so that the specific value p_0 is translated through space with a speed of propagation c. Thus, c is the speed of sound of the wave. *See* WAVE (PHYSICS); WAVE EQUATION; WAVE MOTION.

Harmonic waves. A most important plane wave, both conceptually and mathematically, is the smoothly oscillating monofrequency plane wave described by Eq. (4). The amplitude of this wave is P.

$$p = P\cos\left[2\pi f\left(t - \frac{x}{c}\right)\right] \tag{4}$$

The phase (argument of the cosine) increases with time, and at a point in space the cosine will pass through one full cycle for each increase in phase of 2π. The period T required for each cycle must therefore be such that $2\pi fT = 2\pi$, or $T = 1/f$, so that $f = 1/T$ can be identified as the frequency of oscillation of the pressure wave. During this period T, each portion of the waveform has advanced through a distance $\lambda = cT$, and this distance λ must be the wavelength. This gives the fundamental relation (5) between

$$\lambda f = c \tag{5}$$

the frequency, wavelength, and speed of sound in any medium. For example, in air at room temperature the speed of sound is 343 m/s (1125 ft/s). A sound of frequency 1 kHz (1000 cycles per second) will have a wavelength of $\lambda = c/f = 343/1000$ m $= 0.34$ m (1.1 ft). Lower frequencies will have longer wavelengths: a sound of 100 Hz in air has a wavelength of 3.4 m (11 ft). For comparison, in fresh water at room temperature the speed of sound is 1480 m/s (4856 ft/s), and the wavelength of 1-kHz sound is nearly 1.5 m (5 ft), almost five times greater than the wavelength for the same frequency in air.

Because of many close analogies between sound and electricity, it is often convenient in practice to define the effective pressure amplitude $P_e = P/\sqrt{2}$ in a wave such as that of Eq. (4). Similarly, the frequency is often represented by the angular frequency $\omega = 2\pi f$ and the wavelength expressed reciprocally by the wave number $k = 2\pi/\lambda$. With these definitions, Eq. (4) could be written as Eq. (6), and Eq. (5) as Eq. (7).

$$p = \sqrt{2}P_e\cos(\omega t - kx) \tag{6}$$

$$\frac{\omega}{k} = c \tag{7}$$

See ALTERNATING-CURRENT CIRCUIT THEORY.

Transient and continuous waves. Monofrequency waves are the building blocks for more complicated waves that are either continuous or transient in time. For example, a sawtooth continuous wave has acoustic pressure which, during each cycle, begins at a positive value P_{max}, decreases uniformly (linearly with time) to a negative value $-P_{max}$ at the end of the period T, and then jumps instantaneously back to the positive value P_{max}, repeating this cycle indefinitely. It is a direct consequence of the wave equation that this waveform can be described equivalently as a Fourier superposition, or summation, of waves $p_1 + p_2 + p_3 + \cdots$, each of the form of Eq. (8), where

$$p_n = \frac{2}{\pi}\frac{P_{max}}{n}\sin\left[2\pi nf\left(t - \frac{x}{c}\right)\right] \tag{8}$$

the fundamental frequency $f = 1/T$ gives the repetition rate of the waveform and n has integer values 1, 2, 3, ... The waves p_n constitute the overtones of the waveform. In this case the frequencies nf of the overtones are integer multiples of the fundamental frequency f, and the overtones are termed harmonics. Any signal that is nonrepeating or is nonzero only with some limited duration of time can be written as a summation of waves that are not harmonically related or, in the extreme, an integration of waveforms of all frequencies. As an illustration, a very sharp positive pulse of pressure traveling in the x direction and lasting for an infinitesimally short duration of time is represented by the Dirac delta function $\delta(t - x/c)$. This function, which has unbounded value where $t = x/c$ and is zero elsewhere, can be expressed as an integral, as in Eq. (9). These considerations show that

$$\delta\left(t - \frac{x}{c}\right) = 2\int_0^\infty \cos\left[2\pi f\left(t - \frac{x}{c}\right)\right] df \qquad (9)$$

a study of monofrequency sound waves is sufficient to deal with all sound waves, and that the fundamental concepts of frequency and wavelength permeate all aspects of sound. *See* FOURIER SERIES AND TRANSFORMS; HARMONIC (PERIODIC PHENOMENA); NONSINUSOIDAL WAVEFORM; WAVEFORM.

Standing waves. In many situations, sound is generated in an enclosed space that traps the sound within or between boundaries. For example, if there is a boundary that causes the pressure wave given by Eq. (4) to be completely reflected back on itself, then there is an additional wave given by Eq. (10),

$$p' = P\cos\left[2\pi f\left(t + \frac{x}{c}\right)\right] \qquad (10)$$

which represents a monofrequency plane wave traveling in the $-x$ direction. This wave combines with the incident wave, resulting in a standing wave with a pressure dependence p_T given by Eq. (11). This kind

$$p_T = p + p' = 2P\cos\left(\frac{2\pi x}{\lambda}\right)\cos(2\pi f t) \qquad (11)$$

of wave would be found within a sounding organ pipe and is analogous to a vibrating guitar string. The pressure waveform p_T is zero at positions $x = \lambda/4$, $3\lambda/4$, $5\lambda/4$, ... These nodes (pressure nulls) occur every half-wavelength in space. Midway between them are antinodes at which the pressure waveform oscillates in time between its maximum and minimum values of $\pm 2P$. More complicated standing waves can be formed from waves traveling in a number of directions. For example, vibrating panels or drum heads support standing waves in two dimensions, and steady tones in rooms can excite three-dimensional standing waves. *See* VIBRATION.

Particle Speed and Displacement

As sound passes through a fluid, the small fluid elements are displaced from their equilibrium rest positions by the fluctuating pressure gradients. The motion of a fluid element is described by the particle velocity $\vec{u}$ with which it moves about its equilibrium position. This is related to the pressure by Newton's second law. With the neglect of some small nonlinear and viscous terms, this law can be written as Eq. (12), where ρ_0 is the equilibrium density of

$$\rho_o \frac{\partial \vec{u}}{\partial t} = -\nabla p \qquad (12)$$

the fluid and ∇ is the gradient operator. *See* CALCULUS OF VECTORS; FLUID-FLOW PRINCIPLES; NEWTON'S LAWS OF MOTION.

For a one-dimensional plane wave moving in the $+x$ direction, the acoustic pressure p and particle speed u are proportional and related by $p/u = \rho_0 c$. The product $\rho_0 c$ is a basic measure of the elastic properties of the fluid and is called the characteristic impedance. This is an index of the acoustic "hardness" or "softness" of a fluid or solid. (The term characteristic impedance is restricted to the plane-wave value of $\rho_0 c$; more generally, the term specific acoustic impedance is used.) Some representative values of the speed of sound, the density, and the specific acoustic impedance are given in **Table 1**. Significant differences among these quantities for gases, liquids, and solids are evident. *See* ACOUSTIC IMPEDANCE.

Because fluids cannot support shear (except for small effects related to viscosity), the particle velocity of the fluid elements is parallel to the direction of propagation of the sound wave and the motion is longitudinal. In contrast, solids can transmit shearing or bending motion—reeds, strings, drum heads, tuning forks, and chimes can vibrate transversely.

Description of Sound

The characterization of a sound is based primarily on human psychological responses to it. Because of the nature of human perceptions, the correlations between basically subjective evaluations such as loudness, pitch, and timbre and more physical qualities such as energy, frequency, and frequency spectrum are subtle and not necessarily universal.

Intensity, loudness, and the decibel. The strength of a sound wave is described by its intensity I. From basic physical principles, the instantaneous rate at which energy is transmitted by a sound wave through unit area is given by the product of acoustic pressure and the component of particle velocity perpendicular to the area. The time average of this quantity is the acoustic intensity, as in Eq. (13). For a

$$I = \frac{1}{t}\int_0^t pu\, dt \qquad (13)$$

plane monofrequency traveling wave, this is given by Eq. (14) in the direction of propagation. If all quanti-

$$I = \frac{1}{2}\frac{P^2}{\rho_0 c} = \frac{P_e{}^2}{\rho_0 c} \qquad (14)$$

ties are expressed in SI units (pressure amplitude or effective pressure amplitude in Pa, speed of sound in m/s, and density in kg/m^3), then the intensity will be in watts per square meter (W/m^2).

Because of the way the strength of a sound is perceived, it has become conventional to specify the

TABLE 1. Density, speed of sound, and specific acoustic impedance in selected materials*

Gases	Density (ρ_0), kg/m^3	Speed of sound (c), m/s
Air	1.21	343
Oxygen (0°C; 32°F)	1.43	317
Hydrogen (0°C; 32°F)	0.09	1,270

Liquids	Density (ρ_0), kg/m^3	Speed of sound (c), m/s
Water	998	1,481
Seawater (13°C; 55°F)	1,026	1,500
Ethyl alcohol	790	1,150
Mercury	13,600	1,450
Glycerin	1,260	1,980

Solids	Density (ρ_0), kg/m^3	Speed of sound (c), m/s		Specific acoustic impedance ($\rho_0 c$), N · s/m^3	
		bar	bulk	bar	bulk
Aluminum	2,700	5,150	6,300	13.9	17.0
Brass	8,500	3,500	4,700	29.8	40.0
Lead	11,300	1,200	2,050	13.6	23.2
Steel	7,700	5,050	6,100	39.0	47.0
Glass	2,300	5,200	5,600	12.0	12.9
Lucite	1,200	1,800	2,650	2.2	3.2
Concrete	2,600	—	3,100	—	8.0

*Temperature = 20°C = 68°F unless otherwise indicated. Pressure = 1 atm = 101.3 kPa = 14.7 lbf/in.^{2} 1 kg/m^3 = 0.0624 lbm/ft^3; 1 m/s = 3.281 ft/s; 1 N · s/m^3 = 6.366 × 10^{-3} lbf · s/ft^3.
SOURCE: After L. E. Kinsler et al., *Fundamentals of Acoustics,* 3d ed., Wiley, 1982.

intensity of sound in terms of a logarithmic scale with the (dimensionless) unit of the decibel (dB). An individual with unimpaired hearing has a threshold of perception near 10^{-12} W/m^2 between about 2 and 4 kHz, the frequency range of greatest sensitivity. As the intensity of a sound of fixed frequency is increased, the subjective evaluation of loudness also increases, but not proportionally. Rather, the listener tends to judge that every successive doubling of the acoustic intensity corresponds to the same increase in loudness. This is conveniently expressed by Eq. (15), where the logarithm is to base 10, I

$$L_I = 10 \log \left(\frac{I}{I_{\text{ref}}} \right) \tag{15}$$

is the intensity of the sound field in W/m^2, I_{ref} is 10^{-12} W/m^2, and L_I is the intensity level in dB re 10^{-12} W/m^2. On this scale, the weakest sounds that can be perceived have an intensity level of 0 dB, normal conversational levels are around 60 dB, and hearing can be damaged if exposed even for short times to levels above about 120 dB. Every doubling of the intensity increases the intensity level by 3 dB. For sounds between about 500 Hz and 4 kHz, the loudness of the sound is doubled if the intensity level increases about 9 dB. This doubling of loudness corresponds to about an eightfold increase in intensity. For sounds lying higher than 4 kHz or lower than 500 Hz, the sensitivity of the ear is appreciably lessened. Sounds at these frequency extremes must have higher threshold intensity levels before they can be perceived, and doubling of the loudness requires smaller changes in the intensity with the result that at higher levels sounds of equal intensities tend to have more similar loudnesses. It is because of this characteristic that reducing the volume of recorded music causes it to sound thin or tinny, lacking both highs and lows of frequency. *See* DECIBEL; LOUDNESS.

Since most sound-measuring equipment detects acoustic pressure rather than intensity, it is convenient to define an equivalent scale in terms of the sound pressure level. Under the assumption that Eq. (14) is valid for most commonly encountered sound fields, a reference effective pressure amplitude $P_{\text{ref}} = 20$ micropascals (μPa) generates the reference intensity of 10^{-12} W/m^2 in air (at standard temperature and pressure) and a sound pressure level (SPL) can be defined by Eq. (16), where P_e is the

$$\text{SPL} = 20 \log \left(\frac{P_e}{P_{\text{ref}}} \right) \tag{16}$$

effective pressure amplitude in μPa. The intensity level and sound-pressure level are usually taken as identical, but this is not always true (these levels may not be equivalent for standing waves, for example). For underwater sounds, the sound pressure level is also expressed by Eq. (16), but the reference effective pressure is defined as 1 μPa.

Frequency and pitch. How "high" sound of a particular frequency appears to be is described by the sense of pitch. A few minutes with a frequency generator and a loudspeaker show that pitch is closely related to the frequency. Higher pitch corresponds to higher frequency, with small influences depending on loudness, duration, and the complexity of the waveform. For the pure tones (monofrequency sounds) encountered mainly in the laboratory, pitch and frequency are not found to be proportional. Doubling the frequency less than doubles the pitch. For the more complex waveforms usually encountered,

however, the presence of harmonics favors a proportional relationship between pitch and frequency. *See* PITCH.

Consonance and dissonance. Two tones generated together cannot be distinguished from each other if their frequencies are the same. If their frequencies f_1 and (slightly higher) f_2 are nearly but not exactly identical, the ear will perceive a slow beating, hearing a single tone of slowly and regularly varying amplitude. The combination yields an equivalent signal given by Eq. (17), which is heard as a

$$\cos\left(2\pi f_1 t\right) + \cos\left(2\pi f_2 t\right) = 2\cos\left(2\pi\frac{f_2 - f_1}{2}\,t\right)\cos\left(2\pi\frac{f_2 + f_2}{2}\,t\right) \quad (17)$$

single tone having a frequency that is the average of the frequencies of the two individual tones and an amplitude that varies slowly according to the difference of the two frequencies. As f_2 increases more, the beating will quicken until it first becomes coarse and unpleasant (dissonant) and then resolves into two separate tones of different pitches. With still further increase, a sense of beating and dissonance will reappear, leading into a blending or consonance that then breaks again into beats and dissonance, and the whole cycle of events repeats. These islands of consonance surrounded by beating and dissonance are attained whenever the ratio of the two frequencies becomes that of small integers, $f_2/f_1 = 1/1, 2/1, 3/2, 4/3, \ldots$. The larger the integers in the ratio, the more subtle the effects become. *See* BEAT.

Frequency spectrum and timbre. Sounds can be characterized by many subjective terms such as clean, nasal, edgy, brassy, or hollow. Each term attempts to describe the nature of a complex waveform that may be of very short or long duration and that consists of a superposition or combination of a number of pure tones. The sound of a person's whistling or of a flute played softly often has a pure, clean, but somewhat dull quality. These sounds consist mainly of a pure tone with few or no harmonics. As described above, complex repetitive waveforms are made up of a fundamental tone and a number of harmonics whose frequencies are integer multiples of the fundamental frequency. Blown or bowed instruments such as flute, bowed violin, oboe, and trumpet provide good examples. Other sounds are transient, or nonrepetitive, and usually consist of a fundamental plus a number of overtones whose frequencies are not integer multiples of the lowest. Piano, tympani, cymbals, and plucked violin generate these kinds of sounds.

An important factor is the way in which a sound commences. When a chime is struck, there is a clear sharp onset much like a hammer hitting an anvil; the higher overtones are quite short in duration, however, and quickly die out, leaving only a few nearly harmonic lower overtones that give a clear sensation of pitch. Plucking a guitar string near the bridge with a fingernail yields a similar effect. A gong gives a very complex impression because there are many nonconsonant overtones present, dying away at different rates, so that the sound seems to shift in pitch and timbre, and may appear nonmusical or merely noise to some. It is the abundance of harmonics and overtones, the distribution of intensity among them, and how they preferentially die away in time that provide the subjective evaluation of the nature or timbre of the sound. *See* MUSICAL ACOUSTICS.

Propagation of Sound

Plane waves are a considerable simplification of an actual sound field. The sound radiated from a source (such as a loudspeaker, a hand clap, or a voice) must spread outward much like the widening circles from a pebble thrown into a lake.

Spherical waves. A simple model of this more realistic case is a spherical source vibrating uniformly in all directions with a single frequency of motion. The sound field must be spherically symmetric with an amplitude that decreases with increasing distance r from the source, and the fluid elements must have particle velocities that are directed radially. A solution of the wave equation with spherical symmetry describing this kind of motion for an outgoing monofrequency traveling wave is given by Eq. (18),

$$p = \frac{A}{r}\cos\left(\omega t - kr\right) \quad (18)$$

where $\omega = 2\pi f$ is the angular frequency and $k = 2\pi/\lambda$ the wave number. The pressure amplitude is A/r and diminishes inversely with distance from the source. If the spatially dependent pressure amplitude is defined by Eq. (19), then the intensity is still

$$P(r) = \frac{A}{r} \quad (19)$$

given by Eq. (14), but with P and P_e interpreted as A/r and $(A/r)/\sqrt{2}$, respectively. Thus, the intensity falls off with distance as $1/r^2$. This is consistent with conservation of energy. The acoustic power sent through a sphere of radius r surrounding the source is the intensity of the wave multiplied by the surface area through which it passes. This yields $4\pi r^2 I = 4\pi A^2/(2\rho_0 c)$, which is independent of r.

While the particle velocity for this wave is a relatively complicated function of r, at distances $r \ll \lambda$ the ratio p/u approaches $\rho_0 c$, the same as for a traveling plane wave. Further, in this limit the surfaces of constant phase, for which $(t - r/c)$ has constant value, become more and more planar. Thus, at sufficient distances from the source a spherical wave becomes indistinguishable from a plane wave when viewed over regions of the space whose dimensions are small with respect to r. This asymptotic behavior allows use of the simple plane-wave relationships for many situations.

Directional waves. Not all sources radiate their sound uniformly in all directions. When someone is speaking in an unconfined space, for example an open field, a listener circling the speaker hears the voice most well defined when the speaker is facing the listener. The voice loses definition when the

speaker is facing away from the listener. Higher frequencies tend to be more pronounced in front of the speaker, whereas lower frequencies are perceived more or less uniformly around the speaker.

Dipole source. The fields radiated by sources of more complicated shape and size can be calculated by considering the complicated source as being made up of a collection of small spherical sources, each radiating a pressure wave like that given by Eq. (18), and then adding the pressure fields together. A simple example is the dipole source, consisting of two small sources spaced very closely together and radiating 180° out of phase, so that one is shrinking as the other is expanding, and vice versa. The two sources will nearly cancel because as one is generating a positive pressure, the other is generating a negative pressure. Because the two sources are slightly separated, however, the fields will not exactly cancel. If θ is defined as the angle included between the desired direction in space and the line joining the centers of the two (out-of-phase) sources, and d is the distance between the two sources, then at large distances r away from the pair (**Fig. 1***a*) the total acoustic pressure field

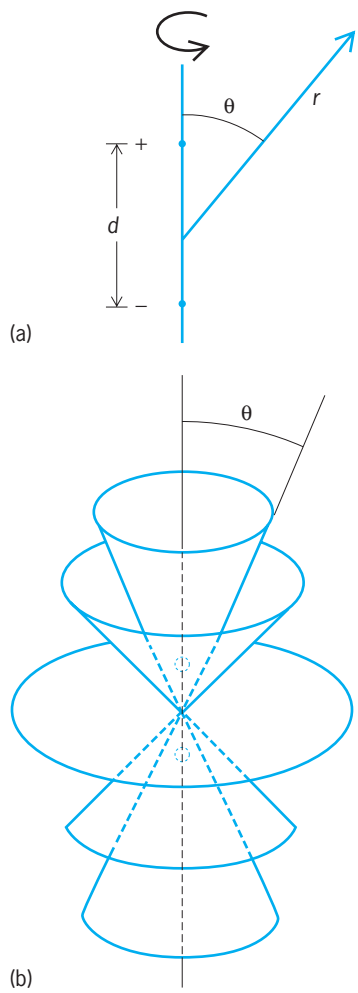

Fig. 1. Pressure field of a dipole source. (*a*) Coordinates θ and r used to describe the field. The distance between the component sources is d. (*b*) Nodal surfaces of the field when the distance d is sufficiently large with respect to the wavelength λ.

is given approximately by Eq. (20). This equation is

$$p = \frac{Akd}{r}(\cos\theta)\ \sin(\omega t - kr) \qquad (20)$$

valid when the wavelength of sound is much larger than the distance separating the sources ($kd < 1$) and r is much larger than d. The amplitude of the pressure $P(r)$is given by $(Akd/r)|\cos\theta|$. In any fixed direction, the amplitude of the pressure falls off as $1/r$, but at a fixed distance the pressure amplitude is modulated in direction by $|\cos\theta|$. In the direction $\theta = 0$ (or $\theta = \pi$ radians), the two sources lie one behind the other and the cancellation is least complete. If $\theta = \pi/2$, then the two sources lie side by side and the cancellation is total. There is a nodal surface, defined by the plane $\theta = \pi/2$, exactly midway between the two sources and perpendicular to the line joining them.

If the distance d is not significantly less than the wavelength λ, then Eq. (20) is not accurate, and a more complicated expression, Eq. (21), must be

$$p = 2\,\frac{A}{r}\ \sin\left(\frac{1}{2}\,kd\cos\theta\right)\sin(\omega t - kr) \qquad (21)$$

used. There is still a nodal surface at $\theta = \pi/2$, but if kd is large enough there may be additional nodal surfaces, each a cone concentric with the line joining the sources, in directions given by angles θ_n satisfying $kd\cos\theta_n = 2n\pi$ with $n = 1, 2, \ldots$ (Fig. 1*b*).

Generalized source. Generalizing from the above shows that the pressure field radiated by a source of arbitrary shape will have an amplitude that at large distance can be written as the product of a function of r and a function of direction, as in Eq. (22).

$$P(r, \theta, \phi) = P_{\mathrm{ax}}(r)H(\theta, \phi) \qquad (22)$$

Here, $P_{\mathrm{ax}}(r)$ falls off as $1/r$, and $H(\theta,\phi)$ is a function only of direction and has a maximum value of 1. At fixed distance r, the pressure amplitude has maximum value given by $P_{\mathrm{ax}}(r)$, termed the axial pressure amplitude. The direction for which $H(\theta,\phi)$ has its maximum magnitude of 1 defines the acoustic axis of the source. The acoustic axis may be a plane or a series of surfaces of constant θ or ϕ, but often the acoustic axis is a single line. The ratio of the intensity found at distance r in some arbitrary direction given by θ and ϕ to the value found at the same distance on the acoustic axis is simply $H^2(\theta,\phi)$, and the expression of this as an intensity level, called the beam pattern $b(\theta,\phi)$, is given by Eq. (23). The decibel value

$$b(\theta, \phi) = 20\log H(\theta, \phi) \qquad (23)$$

of the beam pattern is the reduction in intensity level in that direction compared to the value on the acoustic axis (at the same distance from the source).

Baffled piston source. A loudspeaker or a hole in a partition can be represented as a baffled piston. This is a flat, circular source of radius a mounted on a large, flat, rigid surface, commonly called a baffle. All portions of the source move together with uniform particle velocity normal to the baffle. It is convenient to

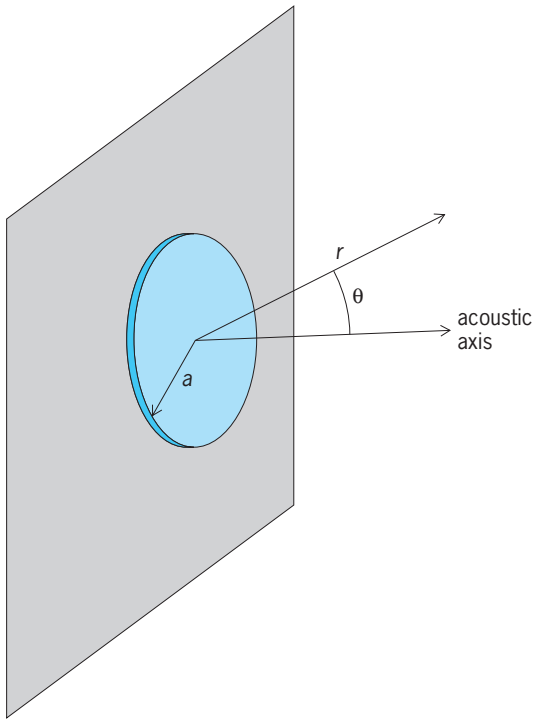

Fig. 2. Coordinates *r* and θ used to describe the pressure field of a baffled piston. The radius of the piston is *a*.

measure the distance r from the center of the piston and the direction θ with respect to the perpendicular to the piston and baffle (**Fig. 2**). The acoustic axis is the line $\theta = 0$, and the principal sound radiated by the piston propagates along this axis, falling off as $1/r$ at larger distances and spreading within a conical nodal surface surrounding the axis. This spreading beam of sound is the major lobe of the radiated pressure field. The angle θ_1 defining this nodal surface is given by $ka \sin\theta_1 = 3.83$. This surface will exist only if the frequency is high enough and the wavelength consequently small enough that ka is greater than 3.83. For sufficiently high frequencies, a number of nodal surfaces may exist, defined by angles θ_n satisfying Eq. (24). [Here, the quantities j_{1n} are the values of

$$ka \sin\theta_n = j_{1n} = 3.83,\ 7.02,\ 10.17,\ \ldots \qquad (24)$$

the argument of the first-order Bessel function J_1 that reduce it to zero: $J_1(j_{1n}) = 0$.] Between these nodal surfaces will be minor lobes of lower-intensity levels. Each of these minor lobes will be funnel- or bowl-shaped, surrounding the major lobe as a figure of revolution about the $\theta = 0$ acoustic axis. Some examples for a piston with different values of ka are shown in **Fig. 3**. The first side lobe, when it exists, has an intensity level about 18 dB less than the major lobe. Subjectively, sound heard by someone standing in the direction of the first side lobe is about one-fourth as loud as that heard in the direction of the major lobe at the same distance. For a given piston, it is apparent that higher frequencies (shorter wavelengths) will yield pressure fields that have narrower major lobes and more nodal surfaces and minor lobes. This is one reason why higher-frequency sounds are often radiated by smaller loudspeakers (tweeters) in high-fidelity systems. The smaller loudspeakers radiate the shorter wavelength sounds more uniformly over larger angles, helping to distribute the energy throughout the space. These small speakers are not useful for lower frequencies, however, because they cannot move enough air to generate sufficiently intense low-frequency sound. *See* BESSEL FUNCTIONS; LOUDSPEAKER.

In addition to the geometrically simple far field found at large distances from a source, there will also often be a rather complicated near field in the immediate vicinity of the source. For the piston source, the transition between near and far fields occurs around

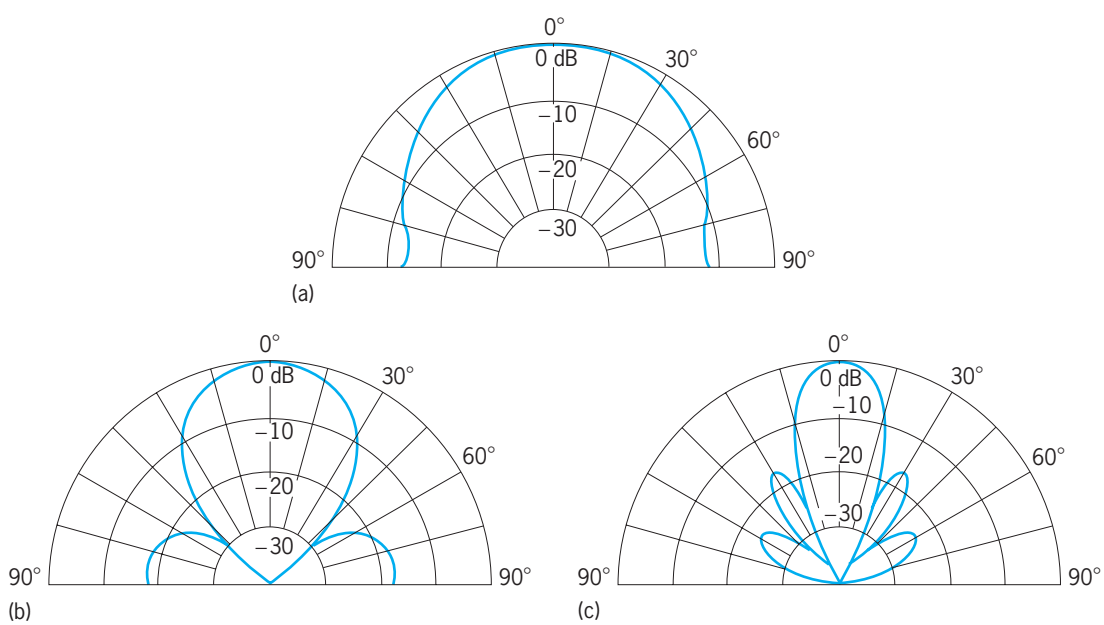

Fig. 3. Directivity patterns for a rigid circular piston mounted on an infinite baffle for different values of *ka*, where *a* is the radius of the piston and *k* is the wave number. (*a*) $ka = 3$. (*b*) $ka = 5$. (*c*) $ka = 10$.

a distance r_1 estimated by $r_1/a \sim a/\lambda$. *See* DIRECTIVITY.

Diffraction. It is possible to hear but not see around the corner of a tall building. However, higher-frequency sound (with shorter wavelength) tends to bend or "spill" less around edges and corners than does sound of lower frequency. The ability of a wave to spread out after traveling through an opening and to bend around obstacles is termed diffraction. This is why it is often difficult to shield a listener from an undesired source of noise, like blocking aircraft or traffic noise from nearby residences. Simply erecting a brick or concrete wall between source and receiver is often an insufficient remedy, because the sounds may diffract around the top of the wall and reach the listeners with sufficient intensity to be distracting or bothersome. The effectiveness of a barrier in blocking sound can be roughly estimated by calculation of an insertion loss. For a long solid wall (brick or concrete) of uniform height the insertion loss can be estimated from Eq. (25), where r is the direct-

$$\Delta L_I = 19 + 10 \log \left(\frac{D - r}{\lambda} \right) \tag{25}$$

path distance (the distance in a straight line from source to receiver), D is the new path length from the source just over the top of the wall to the receiver, and ΔL_I is the difference between the intensity level that would be experienced if there were no barrier and the intensity level with the barrier. The equation is accurate when $(D - r)/\lambda \ll 1$. *See* ACOUSTIC NOISE; DIFFRACTION.

Speed of sound. In fluids, the acoustic pressure and the density variation are related by Eq. (26),

$$p = \mathscr{R} s \tag{26}$$

where the condensation s is $(\rho - \rho_0)/\rho_0$ and $\mathscr{R}$ is a constant of proportionality. This is a three-dimensional Hooke's law stating that stress and strain are proportional. The propagation of sound requires such rapid density changes that there is not enough time for thermal energy to flow from compressed (higher-temperature) elements to decompressed (lower-temperature) elements. Because there is little heat flow, the process is nearly adiabatic and $\mathscr{R}$ is the adiabatic bulk modulus (rather than the more familiar isothermal modulus relevant to constant-temperature processes). A derivation of the wave equation for a fluid reveals that the speed of sound is given by Eq. (27). Equations (26) and (27) can be combined to give Eq. (28).

$$c^2 = \frac{\mathscr{R}}{\rho_0} \tag{27}$$

$$p = \rho_0 c^2 s \tag{28}$$

See ADIABATIC PROCESS; HOOKE'S LAW.

In most fluids, the equation of state is not known, and the speed of sound must be measured directly or inferred from experimental determinations of the bulk modulus. In distilled water the measured speed of sound in meters per second can be represented by a simplified equation (29), where T is the temper-

$$c = 1403 + 4.9T - 0.05T^2 + 0.16\,\mathscr{P}_0 \tag{29}$$

ature in degrees Celsius and $\mathscr{P}_0$ is the equilibrium pressure in bars. For most gases at reasonable pressures and temperatures, however, the equation of state is known to be very close to the perfect gas law, Eq. (30), where R is the specific gas constant

$$\mathscr{P}_0 = \rho_0 R T_K \tag{30}$$

and T_K the absolute temperature. (Absolute temperature T_K in kelvin and standard temperature T in degrees Celsius are related by $T_K = T + 273$.) Further, for sound in a gas, the instantaneous pressure and density are related by Eq. (31), where γ is the ratio

$$\frac{\mathscr{P}}{\mathscr{P}_0} = \left(\frac{\rho}{\rho_0} \right)^{\gamma} \tag{31}$$

of specific heats. For air, $\gamma = 1.4$. For small pressure fluctuations, Eq. (31) can be approximated by Eq. (32). Comparing this with Eq. (28), and using the

$$\frac{\mathscr{P}}{\mathscr{P}_0} \doteq 1 + \gamma s \tag{32}$$

perfect gas law, Eq. (30), gives a theoretical expression for the speed of sound in an ideal gas, Eq. (33).

$$c = \sqrt{\gamma R T_K} \tag{33}$$

Thus, the speed of sound in an ideal gas depends on temperature, increasing as the square root of the absolute temperature. Because many applications of Eq. (33) are for gases around room temperature, it is convenient to reexpress the formula in terms of degrees Celsius and the speed of sound c_0 at $T = 0°C$ (32°F), as in Eq. (34). This equation can be used to

$$c = c_0 \sqrt{1 + \frac{T}{273}} \tag{34}$$

calculate the speed of sound for a perfect gas at any temperature in terms of the speed of sound in that gas at 0°C (32°F). In air, for example, $c_0 = 331$ m/s (1086 ft/s), and at 20°C (68°F) the speed of sound is calculated to be 343 m/s (1125 ft/s). *See* GAS; HEAT CAPACITY; UNDERWATER SOUND.

In solids, the sound speed depends on the transverse extent of the solid with respect to the wave. If the solid is rodlike, with transverse dimensions less than the acoustic wavelength, the solid can swell and contract transversely with relative ease, which tends to slow the wave down. For this case, the speed of sound is given by Eq. (35) where $\mathscr{Y}$ is Young's mod-

$$c = \sqrt{\frac{\mathscr{Y}}{\rho_0}} \tag{35}$$

ulus. If, in the other extreme, the solid has transverse dimensions much larger than the acoustic wavelength, then a longitudinal wave will travel with a

greater speed of sound given by Eq. (36), where $\mathscr{G}$ is

$$c = \sqrt{\frac{\mathscr{B} + 4\mathscr{G}/3}{\rho_0}} \tag{36}$$

the shear modulus. These two different propagation speeds are often called bar and bulk sound speeds, respectively. *See* ELASTICITY; YOUNG'S MODULUS.

Ray and refraction. Since the speed of sound varies with the local temperature (and pressure, in other than perfect gases), the speed of a sound wave can be a function of position. Different portions of a sound wave may travel with different speeds of sound.

If a highly directional piston projects a sound wave horizontally into the air, and if the value of ka is large enough, the major lobe of the sound radiated by the piston is confined to directions within a very small angle θ_1 about the (horizontal) acoustic axis. The beam of sound has surfaces of constant phase, each resembling a disk, perpendicular to and propagating out along the acoustic axis with the speed of sound of the air. If, however, the air temperature decreases gradually with height above the ground, then the upper portions of the beam travel in air with a lesser sound speed than the lower portions. The top of a surface of constant phase gradually lags behind the bottom, and the disk tilts a little so that its normal points slightly upward. Since the direction of propagation is normal to the surface of constant phase, this surface gradually propagates upward, rising above the horizontal as it travels outward away from the piston. Thus, as a consequence of the speed of sound decreasing with altitude, the beam of sound is refracted upward in an arc with one end horizontal at the piston and the other tilting progressively upward as distance from the piston increases.

Each small element of a surface of constant phase traces a line in space, defining a ray along which acoustic energy travels. The sound beam can then be viewed as a ray bundle, like a sheaf of wheat, with the rays distributed over the cross-sectional area of the surface of constant phase (**Fig. 4**). As the major lobe spreads with distance, this area increases and the rays are less densely concentrated. The number of rays per unit area transverse to the propagation path measures the energy density of the sound at that point.

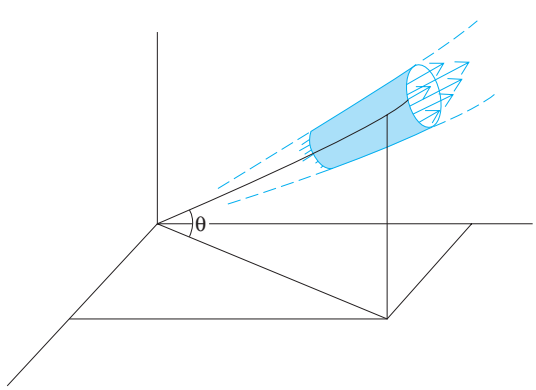

Fig. 4. Representation of an upwardly refracted sound beam as a ray bundle.

It is possible to use the concept of rays to study the propagation of a sound field. The ray paths define the trajectories over which acoustic energy is transported by the traveling wave, and the flux density of the rays measures the intensity to be found at each point in space. This approach, an alternative way to study the propagation of sound, is approximate in nature but has the advantage of being very easy to visualize.

An equation describing refraction can be obtained from the three-dimensional wave equation (37), with c now a function of space, by trying a solution of the form of Eq. (38). The pressure amplitude $P(x,y,z)$

$$\nabla^2 p = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} \tag{37}$$

$$p = P(x,y,z)\cos\left[\omega\left(t - \frac{\psi(x,y,z)}{c_0}\right)\right] \tag{38}$$

depends on location, and c_0 is a convenient constant reference value of sound speed. The quantity ψ has the dimensions of a length, and constant ψ determines a surface of constant phase. The local direction of propagation is perpendicular to this surface and is given by $\nabla\psi$, the gradient of ψ. For the plane traveling wave of Eq. (4), the surfaces, given by $x =$ constant, are planes perpendicular to the x axis. The gradient of x points in the $+x$ direction, the direction of propagation of the phase surface. For the spherical wave of Eq. (6), the surfaces are given by $r =$ constant, and each is a sphere concentric with the source at $r = 0$. The gradient of r is direct radially outward at each point. Substitution of Eq. (38) into the wave equation and some approximation results in the Eikonal equation (39) for the direction of propagation. The Eikonal equation is valid if the fractional changes in P and c in any direction over distances of a wavelength are very small. In general, solving this equation requires integral and differential calculus and results in rather difficult formulas.

$$|\nabla\psi|^2 \doteq \left(\frac{c_0}{c}\right)^2 \tag{39}$$

For the case of a one-dimensional speed of sound profile $c = c(z)$, where z represents altitude, the solution can be simply stated: If direction is defined by an angle θ of elevation ($\theta = 0$ points horizontally and $\theta = \pi/2$ points vertically), then the local direction of propagation of a ray must obey Snell's law, Eq. (40). If the elevation angle of the ray has

$$\frac{\cos\theta}{c} = \frac{\cos\theta_0}{c_0} \tag{40}$$

value θ_0 at an altitude where the speed of sound is c_0, then as the ray propagates to a new altitude where the speed of sound is $c(z)$, its directional angle will change to a new value $\theta(z)$ according to Eq. (40). This allows the ray to be traced point by point as it propagates through the fluid, each new directional angle at each new point being obtained from Snell's law. If the speed of sound varies linearly with altitude, that is, $c(z) = c_0 - gz$, where the gradient $g = \partial c/\partial z$ of the profile is a constant, then the ray path will be

the arc of a circle whose center lies at the altitude where the speed of sound would vanish. [This would never happen since $c(z)$ will stop decreasing linearly long before this altitude is reached. It is, however, a useful artifice for drawing the trajectory of the ray.] **Figure 5** provides an example of such a ray path. *See* ATMOSPHERIC ACOUSTICS; REFRACTION OF WAVES.

Reflection and transmission. If a sound wave traveling in one fluid strikes a boundary between the first fluid and a second, then there may be reflection and transmission of sound. For most cases, it is sufficient to consider the waves to be planar. The first fluid contains the incident wave of intensity I_i and reflected wave of intensity I_r, the second fluid, from which the sound is reflected, contains the transmitted wave of intensity I_t. The directions of the incident, reflected, and transmitted plane sound waves may be specified by the grazing angles θ_i, θ_r, and θ_t (measured between the respective directions of propagation and the plane of the reflecting surface).

The intensity reflection coefficient I_r/I_i and the intensity transmission coefficient I_t/I_i are found by imposing boundary conditions that require continuity of pressure and continuity of the normal component of the particle velocity at the interface between the two fluids. These boundary conditions then require that the grazing angle of the incident plane wave equals that of the reflected plane wave, Eq. (41). For

$$\theta_i = \theta_r \tag{41}$$

the transmitted wave, the grazing angle θ_t must be given by Eq. (42), where c_1 and c_2 are the values of the

$$\cos\theta_t = \left(\frac{c_2}{c_1}\right)\cos\theta_i \tag{42}$$

speed of sound in the first and second fluids, respectively. Both of these results are consistent with Snell's law. There are two possible situations depending on whether the speeds of sound c_1 and c_2 and the angle of incidence are such that $\cos\theta_t < 1$, which presents no conceptual problem, or $\cos\theta_t \ll 1$, which does.

If $\cos\theta_t < 1$, then the transmitted wave is a plane wave sent into the second fluid. The beam of sound is refracted according to Snell's law, being inclined toward the normal to the boundary if c_2 is less than c_1 and being deflected away from the normal if c_2 exceeds c_1. The reflected and transmitted intensities are given by Eqs. (43) and (44), where ρ_1 and ρ_2 are

$$\frac{I_r}{I_i} = \left(\frac{\dfrac{\rho_1 c_2}{\rho_2 c_2} - \dfrac{\sin\theta_t}{\sin\theta_i}}{\dfrac{\rho_1 c_1}{\rho_2 c_2} + \dfrac{\sin\theta_t}{\sin\theta_i}}\right)^2 \tag{43}$$

$$\frac{I_t}{I_i} = \frac{\left(\dfrac{4\rho_2 c_2}{\rho_1 c_1}\right)}{\left(\dfrac{\rho_1 c_1}{\rho_2 c_2} + \dfrac{\sin\theta_t}{\sin\theta_i}\right)^2} \tag{44}$$

the densities of the first and second fluids, respectively. Clearly, the intensities depend strongly on the angle of incidence. If there is a large mismatch in the characteristic acoustic resistances of the two fluids, regardless of which is the larger, the reflection coefficient approaches unity. If $\rho_2 c_2 < \rho_1 c_1$, the second medium appears rigid, and the pressure amplitude at the boundary (the amplitude of the sum of the incident and reflected waves evaluated at the interface) is as great as possible. In the other extreme, $\rho_2 c_2 \ll \rho_1 c_1$, the second medium offers virtually no opposition to the motion of the first, so there is no reactive force and therefore no pressure at the interface. In this case the boundary is called pressure release.

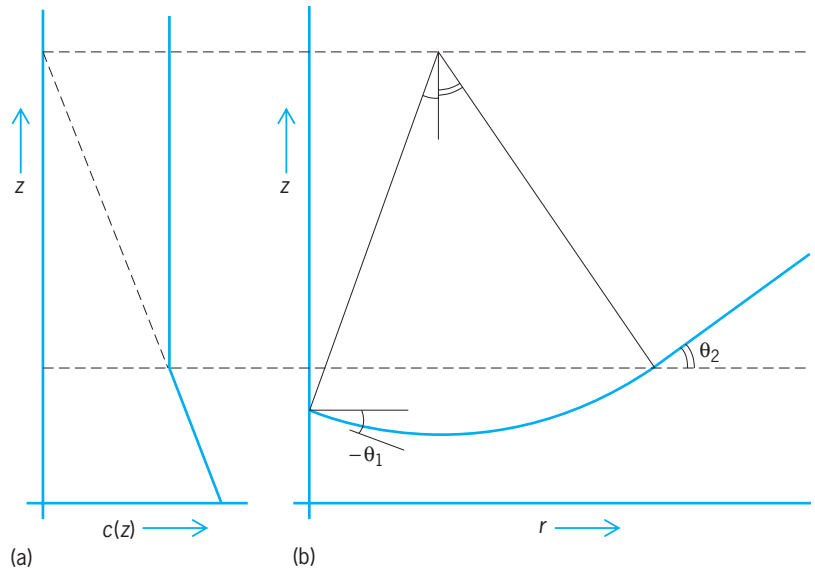

Fig. 5. Ray path for the refraction of sound in a speed-of-sound profile. (*a*) Speed of sound *c* as a function of altitude *z*. (*b*) Path of the ray. *r* = horizontal distance from source; θ = angle of elevation of the ray.

If $\cos\theta_t \ll 1$, then the transmission angle has no physical meaning. In this case the intensity reflection coefficient has unit value, and the transmitted wave travels parallel to the boundary as if it were attached to it with an amplitude that decays exponentially with increasing distance from the boundary. This case is the acoustic analog of the situation in optics when there is complete internal reflection. *See* REFLECTION AND TRANSMISSION COEFFICIENTS; REFLECTION OF ELECTROMAGNETIC RADIATION; REFLECTION OF SOUND.

Absorption. When sound propagates through a medium, there are a number of mechanisms by which the acoustic energy is converted to heat and the sound wave weakened until it is entirely dissipated. This absorption of acoustic energy is characterized by a spatial absorption coefficient α for traveling waves. For a monofrequency traveling plane wave, Eq. (4) must be modified to Eq. (45). [Similarly, a monofrequency traveling spherical wave like that of Eq. (18) would be multiplied by a factor $\exp(-\alpha r)$.]

$$p = Pe^{-\alpha x}\cos\left[2\pi f\left(t - \frac{x}{c}\right)\right] \tag{45}$$

The pressure amplitude diminishes exponentially with increasing distance and the sound pressure level decreases linearly with x, according

TABLE 2. Values of α/f^2 for selected gases and liquids*

	Classical thermoviscous value, $\alpha/f^2 \times 10^{11}$, $Np \cdot s^2/m$	Observed value, $\alpha/f^2 \times 10^{11}$, $Np \cdot s^2/m$
Gases		
Argon	1.85	1.87
Helium	0.53	0.54
Nitrogen	1.35	1.64
Oxygen	1.61	1.92
Liquids		
Acetone	7.0	30.0
Benzene		750
Ethyl alcohol		50
Glycerine	3000	3000
Mercury	6.0	8.0
Water (fresh)	8.1	25.0

*Temperature = 20°C = 68°F. Pressure = 1 atm = 101.3 kPa = 14.7 lbf/in.^{2} 1 $Np \cdot s^2/m$ = 0.3048 $Np \cdot s^2/ft$.

SOURCE: After L. E. Kinsler et al., *Fundamentals of Acoustics,* 3d ed., Wiley, 1982; and J. J. Markham, R. T. Beyer, and R. B. Lindsay, Absorption of sound in fluids, *Rev. Mod. Phys.,* 23:353–411, 1951.

to Eq. (46). The dimension of α is the inverse meter

$$\mathrm{SPL}(x) = \mathrm{SPL}(0) - 8.7\,\alpha x \tag{46}$$

(m^{-1}). For historical reasons, and to avoid ambiguity with other quantities like the wave number k, absorption is given in units of nepers per meter (Np/m), where the neper is a dimensionless label. Because of the use of sound pressure level in acoustics, spatial absorption can also be specified in terms of decibels per meter (dB/m). From Eq. (46), the absorption coefficient in dB/m is 8.7α. For standing waves, the sound field dies away with time according to a factor $\exp(-\beta t)$. The temporal absorption coefficient β, given in nepers per second (Np/s), is related to α by $\beta = \alpha c$. Some values of spatial absorption coefficients are given in **Table 2**. *See* NEPER.

Classical mechanisms. One mechanism of absorption is shear viscosity, the analog of friction for fluids. Molecules traveling between fluid elements of differing speeds will transfer momentum between the elements through intermolecular collisions, diffusing the differences in collective motion into random (thermal) motion of the individual molecules. This loss of acoustic energy into heat is described by the absorption coefficient for viscosity $\alpha\upsilon$. In fluids, this coefficient increases with the square of the frequency (up to about 10^{10} Hz). *See* VISCOSITY.

Another mechanism arises from the thermal conductivity of the fluid. In a perfectly insulating fluid (zero thermal conductivity), cold regions could not conduct thermal energy away from hot regions. In real materials, small amounts of thermal energy can flow between these hotter and colder regions, tending to equalize the temperatures a little and converting acoustic energy into heat. This energy conversion is described by a thermal absorption coefficient α_t, which also increases with the square of the frequency. The losses from shear viscosity and thermal conductivity add together to give the classical thermoviscous absorption coefficient $\alpha_c = \alpha\upsilon + \alpha_t$. *See* CONDUCTION (HEAT).

Structural relaxation. In certain liquids, including water, acetone, and alcohols, the observed absorption is significantly larger than that predicted by the classical absorption coefficient but has the same frequency dependence. A postulated mechanism is a structural relaxation, involving transitions between two preferred clusterings of nearest neighbors, one favored at lower density and the other at higher. When the sound wave travels through the fluid, the pressure and temperature fluctuations cause transitions between these two structural states. The net effect is to introduce a bulk viscosity coefficient additive to the shear viscosity. In water, the effect of bulk viscosity results in a measured absorption coefficient about three times the classical value.

Thermal molecular relaxation. In some gases with polyatomic molecules, like air, thermal molecular relaxation is an important mechanism of absorption. When two gas molecules collide, in addition to simply rebounding from each other like two billiard balls, one or the other of the molecules may be excited into internal vibrations, thus removing energy from the collective acoustic motion. Once excited, the molecule may release this energy back into kinetic energy of motion during another collision, but this release is not coordinated with the acoustic wave and so appears at random, contributing to the thermal motion of the molecules. It takes a characteristic amount of time, called the relaxation time τ, to accomplish the conversion of energy into and from these internal vibrations. If τ is sufficiently small compared to the period T of the acoustic wave, then the mechanism is fully activated and the resulting absorption increases with the square of the frequency. At higher frequencies such that $T < \tau$, the absorption from this effect becomes a constant independent of the frequency. Thus, while at lower frequencies molecular thermal relaxation can be an important source of absorption in certain gases, at higher frequencies the effect is completely dominated by the classical thermoviscous absorption. In air, collisions involving carbon dioxide with either nitrogen or oxygen, and oxygen with water vapor, are important in exciting internal energy states. The total absorption coefficient is a strong function of humidity and frequency throughout the audible range of frequencies. The effects are strong enough that the reverberation characteristics of a poorly ventilated concert hall can change during the course of a concert as the moisture exuded by the audience alters the humidity. **Figure 6** shows curves for the absorption coefficient in decibels per meter as a function of frequency for three different humidities and the classical thermoviscous value.

Chemical relaxations. In certain liquids, there can be chemical relaxations involving the concentrations of various chemical species and how they change with temperature. A sound wave passing through the liquid causes local temperature changes, and these in turn can cause changes in the ionization of dissolved salts. The basic mechanisms cause losses quite similar in effect to the molecular thermal relaxation

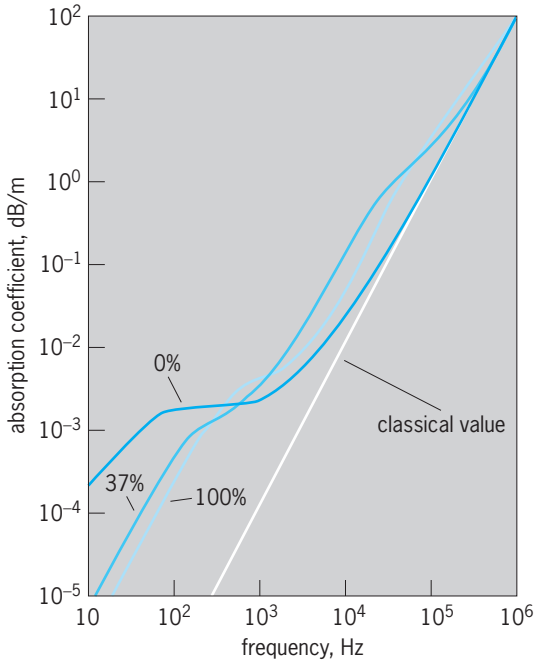

Fig. 6. Absorption of sound in air at 20°C (68°F) as a function of frequency for three humidities and the classical thermoviscous value. 1 dB/m = 0.3 dB/ft. (*After H. E. Bass, H.-J. Bauer, and L. B. Evans, Atmospheric absorption of sound: Analytical expressions, J. Acous. Soc. Amer., 52:821–825, 1972*)

described above. In seawater, chemical relaxations contribute appreciable additional absorption at frequencies below about 100 kHz for magnesium sulfate ($MgSO_4$) and below about 1 kHz for boric acid.

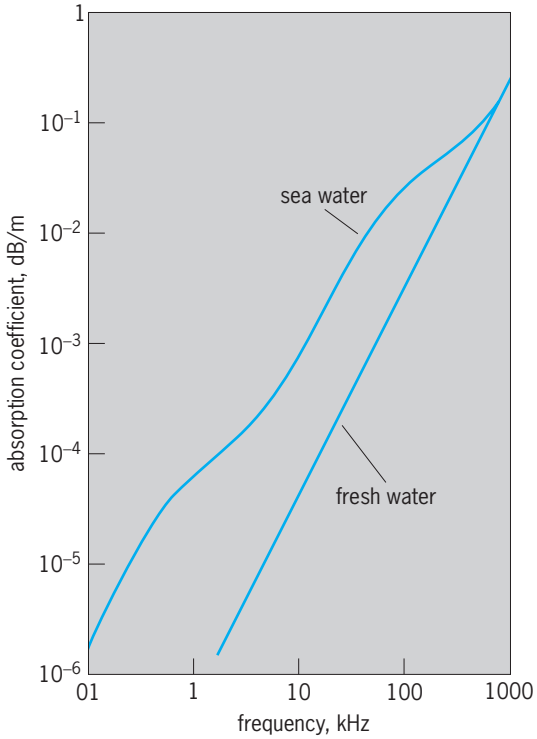

Fig. 7. Absorption of sound in fresh water and in seawater of 35 parts per thousand salinity at atmospheric pressure and 5°C (41°F). 1 dB/m = 0.3 dB/ft. (*After F. H. Fisher and V. P. Simmons, Sound absorption in sea water, J. Acous. Soc. Amer., 62:558–664, 1977*)

Figure 7 gives a representative curve of the frequency dependence of the absorption coefficient in decibels per meter for seawater. *See* SEAWATER.

Alan B. Coppens

Bibliography. L. L. Beranek, *Acoustics*, 1986; R. E. Berg and D. G. Stork, *The Physics of Sound*, 2d ed., 1995; D. E. Hall, *Basic Acoustics*, 1987, reprint 1993; L. E. Kinsler et al., *Fundamentals of Acoustics*, 4th ed., 2000; P. M. Morse and K. U. Ingard, *Theoretical Acoustics*, 1986; M. Rossi, *Acoustics and Electroacoustics*, 1988; T. D. Rossing and N. H. Fletcher, *Principles of Vibration and Sound*, 1994.

Sound absorption

The process by which the intensity of sound is diminished by the conversion of the energy of the sound wave into heat. The absorption of sound is an important case of sound attenuation. This article will address absorption mechanisms in varying media but will concentrate on absorption in ideal gases (such as air), since these effects are the most commonly observed.

Attenuation. Regardless of the material through which sound passes, its intensity, measured by the average flow of energy in the wave per unit time per unit area perpendicular to the direction of propagation, decreases with distance from the source. This decrease is called attenuation. In the simple case of a point source of sound radiating into an ideal medium (having no boundaries, turbulent fluctuations, and the like), the intensity decreases inversely as the square of the distance from the source. This relationship exists because the spherical area through which the energy propagates per unit time increases as the square of the propagation distance. This attenuation or loss may be called geometrical attenuation.

In addition to this attenuation due to spreading, there is effective attenuation caused by scattering within the medium. Sound can be reflected and refracted when incident on media of different physical properties, and can be diffracted and scattered as it bends around obstacles. These processes lead to effective attenuation, for example, in a turbulent atmosphere; this is easily observed in practice and can be measured, but is difficult to calculate theoretically with precision. *See* DIFFRACTION; REFLECTION OF SOUND; REFRACTION OF WAVES.

In actual material media, geometrical attenuation and effective attenuation are supplemented by absorption due to the interaction between the sound wave and the physical properties of the propagation medium itself. This interaction dissipates the sound energy by transforming it into heat and hence decreases the intensity of the wave. In all practical cases, the attenuation due to such absorption is exponential in character. Thus, if I_0 denotes the intensity of the sound at unit distance from the source, then I, the intensity at distance r in the same units, has the form of Eq. (1), where e represents the base

$$I = I_0 e^{-\alpha r} \qquad (1)$$

of napierian logarithms, 2.71828 . . . and α is called the intensity absorption coefficient. The α can be expressed by Eq. (2). Hence α has the dimensions

$$\alpha = \frac{-1}{r} \ln\left(\frac{I}{I_0}\right) \tag{2}$$

of reciprocal length or distance. However, it is customary to express the logarithm, which is a pure dimensionless number, in nepers, and hence α is in nepers per unit length. Since 10 $\log_{10}$ (I/I_0) is the number of decibels between the intensity I and I_0, the absorption coefficient α can also be expressed in decibels per unit length or distance. *See* DECIBEL; NEPER; SOUND INTENSITY.

Classical sources. The four classical mechanisms of sound absorption in material media are shear viscosity, heat conduction, heat radiation, and diffusion. These attenuation mechanisms are generally grouped together and referred to as classical attenuation or thermoviscous attenuation. They will be discussed first as they apply to fluids, namely, liquids and gases.

Viscosity. In the flow of a fluid such as air or water, viscosity is the tendency for the fluid flowing in a layer parallel to the direction of motion to retard the motion in adjoining layers. The viscosity of a liquid (indicated qualitatively by the stickiness or sluggishness of its flow) decreases as the temperature rises, whereas that of gases increases with the temperature. *See* VISCOSITY.

Appropriate modifications of the standard wave equation for plane sound waves to take account of the effect of viscosity of a fluid lead to Eq. (3) for

$$\alpha_v = \frac{8\pi^2 f^2 \mu}{3\rho_0 c^3} \tag{3}$$

the absorption coefficient as defined in Eq. (1), but specifically due to viscosity, where μ is the viscosity of the fluid, f is the frequency of the sound wave in hertz, c is the wave's velocity, and ρ_0 is the equilibrium density of the fluid. The expression for α_v holds only for frequencies that are not too high. For example, in the case of air, it fails at frequencies on the order of 10^9 Hz, which involve ultrasonics. *See* ULTRASONICS.

Heat conduction. The role of heat conduction can be qualitatively understood as follows. The compression associated with the high-pressure portion of a propagating sound wave produces a local increase in temperature, which sets up a temperature gradient through which heat will flow by conduction before the succeeding rarefaction cools the medium. This is an irreversible phenomenon, leading to energy dissipation into heat. The value of the sound absorption coefficient due to heat conduction is given by Eq. (4), where $\gamma = C_\rho/C_v$, the ratio of the molar specific heat at constant pressure C_p to that at constant volume C_v, ρ_0 is the equilibrium density of the fluid, and κ is the thermal conductivity. Equation (4) is again valid only for frequencies that are not too high. *See* CONDUCTION (HEAT); HEAT CAPACITY.

$$\alpha_b = \frac{2\pi^2 f^2 \kappa}{\rho_0 c^3 C_v}\left(\frac{\gamma - 1}{\gamma}\right) \tag{4}$$

Substitution of the appropriate numerical values for the physical parameters shows that for gases the value of α_b is somewhat less than α_v for the same frequency. For liquids, except for liquid metals, $\alpha_v \gg \alpha_b$ for the same frequency. It is usually considered appropriate to add α_v and α_b to get the total absorption due to these two processes.

Heat radiation. Thermal radiation results whenever one portion of a fluid acquires a temperature different from that of an adjacent portion. This effect is an electromagnetic-wave phenomenon different in character from heat transfer by conduction. Calculation shows that thermal radiation plays a very small role in sound absorption except possibly at very high frequencies in very low pressure gases. *See* HEAT RADIATION.

Diffusion. Diffusion operates mainly in the case of fluid mixtures, for example, air that is a mixture mainly of nitrogen and oxygen, in which the two gases have a slight tendency to diffuse into each other. For ordinary gases in which the components differ little in molecular weight, the effect of diffusion on absorption is negligible compared with that of viscosity and heat conduction. However, in gaseous mixtures in which the components differ markedly in molecular weight—for example, a mixture of the rare gases helium and krypton—the effect of diffusion can be much greater than that due to viscosity and heat conduction combined. *See* DIFFUSION.

Measurement of absorption in fluids. Sound absorption in fluids can be measured in a variety of ways, referred to as mechanical, optical, electrical, and thermal methods. All these methods reduce essentially to a measurement of sound intensity as a function of distance from the source.

Mechanical. The mechanical method, confined to the estimate of absorption in liquids, uses radiation pressure, proportional to the sound intensity, to "weigh" the sound in a vertical beam. *See* ACOUSTIC RADIATION PRESSURE.

Optical. The optical method, which is also mainly used for liquids, is based on the diffraction of a beam of light by the sound beam, the latter being treated essentially as a diffraction grating. The greater the intensity of the sound, the more diffraction bands are produced. *See* DIFFRACTION.

Electrical. The electrical methods, the most generally used for both gases and liquids, employ an acoustic transducer like a piezoelectric crystal to receive the sound beam as it traverses the medium whose absorption coefficient is to be evaluated, and to measure the variation in the sound intensity as the distance from the source is varied. An alternative is the use of an acoustic interferometer, which functions essentially in the same way as an optical interferometer. Still another scheme is based on the use of the reverberation time technique. In this method the medium is confined in a closed spherical vessel with highly reflective walls. The reverberation time, that is, the time taken by a sound pulse initially produced

at the center of the vessel by a transducer (which can also act as a sound receiver) to decay to one-millionth of its original intensity provides a measure of the intensity loss and leads to an evaluation of the absorption coefficient. This method has been successfully applied to liquids and gases. *See* INTERFEROMETRY; MICROPHONE; REVERBERATION.

Thermal. The thermal method undertakes to estimate the loss in intensity of a sound wave by direct measurement of the heat produced in its passage. This demands very careful experimental technique, and the difficulties associated with it have provoked controversy about the usefulness of the method.

Experimental values for gases. Measurements carried out by these methods show that the experimental values for the sound absorption coefficient (as a function of frequency) of a monatomic gas like helium agree rather well with the values calculated on the basis of viscosity and heat conduction, except at very low pressures and very high frequencies. On the other hand, for more complicated gases, like air (though in general α varies roughly directly as the square of the frequency, as predicted by classical theory), the actual numerical values are greater (and in many cases very much greater) than the theoretically predicted ones. For example, for air at 20°C (68°F) and normal atmospheric pressure, the theoretical value of $(\alpha_v + \alpha_b)/f^2$ is 1.27×10^{-12} neper s^2/cm, whereas the experimental value of α/f^2 for air of low humidity under the same conditions is approximately 8×10^{-12} in the same units.

Experimental values for liquids. The situation with respect to liquids is about the same. Mercury and the liquefied gases like argon, oxygen, and nitrogen have absorption coefficients in moderately good agreement (that is, within 10–15%) with the classically calculated values. But ordinary liquids like water and aqueous salt solutions show values much in excess of those classically calculated. The observed excess in the case of organic liquids like carbon disulfide, for example, is very large indeed, of the order of several hundred times. A satisfactory explanation for this excess absorption has been found in molecular and related relaxation processes.

Molecular relaxation in gases. Although sound propagation in a gaseous medium is phenomenologically a wave process, it actually takes place through the collisions of the molecules composing the gas. According to the kinetic theory, the molecules of a gas are in ceaseless random motion. When the gas is locally compacted by the compression portion of a propagating sound wave, the average translational kinetic energy of the molecules is locally increased. Energy has been introduced into the group of molecules. This energy increase is conveyed through collisions to the adjoining group of molecules, which in turn passes it on through the gas, and so on, and this continuous transfer represents the sound-wave energy.

A sound wave thus involves the momentary local compression of the medium, which by increasing the local density, according to kinetic theory, also leads to an increase in the average translational kinetic energy and hence to an increase in local pressure. In the case of a gas made up of polyatomic molecules such as oxygen and nitrogen, the increase in pressure cannot take place instantaneously, since some of the extra translational kinetic energy is transferred to the internal rotational and vibrational molecular-energy states. If one imagines the diatomic oxygen (O_2) molecule as a dumbbell-shaped weight, with a springy connecter between the weights, some of the collision energy acts to spin the molecule and some acts to vibrate the weights toward and away from each other. The dissipation of these energy states is referred to as rotational relaxation and vibrational relaxation, respectively.

After the local density has begun to decrease due to the rarefaction portion of the propagating wave, the pressure continues to build up as energy is transferred back to the translation form from the internal rotational and vibration states. The excess density and excess pressure thus get out of phase, as in the standard hysteresis phenomenon, which leads to an effective transformation of acoustic-wave energy into heat, corresponding to sound absorption. This process is termed a relaxation process, and the time taken for the density at any point to return to within the fraction $1 - (1/e)$ [approximately 63%] of its equilibrium value is termed the relaxation time, denoted by τ. *See* HYSTERESIS; MOLECULAR STRUCTURE AND SPECTRA.

Calculation of the effect. The results of the theoretical calculation of the effect of molecular relaxation will be discussed for the special case in which there are only two energy states of each molecule: (1) that in which it possesses translational kinetic energy only, and (2) that in which it possesses translational kinetic energy plus a single internal energy level. Here the relaxation time is given by Eq. (5), where κ_{12}^0 is the av-

$$\tau = \frac{1}{\kappa_{12}^0 + \kappa_{21}^0} \tag{5}$$

erage number of transitions per second from energy state 1 to energy state 2, and κ_{21}^0 is the average number of transitions per second from state 2 to state 1. The superscripts mean that the transition rates refer to the equilibrium condition of the gas as unaffected by the sound wave. Analysis then gives the sound absorption coefficient α due to this type of relaxation process in the form of Eq. (6), where f is the

$$\alpha = K\frac{4\pi^2 f^2 \tau}{1 + 4\pi^2 f^2 \tau^2} \tag{6}$$

frequency and K is a constant containing the molar gas constant R, the sound velocity, and the specific heats of the gas. If $2\pi f\tau \ll 1$, α varies as the square of the frequency, as in the case of viscous and heat conduction absorption. However, if $2\pi f\tau \gg 1$, the attenuation becomes a constant with respect to frequency and the effect is said to be frozen out. In the more general case, if $\alpha/2\pi f$ is plotted as a function of frequency, it has a maximum for $f = 1/2\pi\tau$, the plot taking the form of an approximate bell-shaped curve. Such a plot, carried out experimentally,

provides a method for evaluating the relaxation time τ.

The theoretical evaluation of τ from Eq. (5) involves essentially a calculation in quantum statistics. The agreement of theory with experiment for many gases showing excess absorption over that attributable to viscosity and heat conduction is satisfactory. The relaxation process described here is usually termed thermal relaxation, since the specific heats of the gas enter prominently into the process.

Examples. Examples of gases for which experimental values are in reasonable agreement with theory include hydrogen, for which the excess absorption over classical values arises primarily from rotational relaxation, with a relaxation time on the order of 10^{-8} s, and oxygen, for which the excess low-frequency absorption depends primarily on vibrational relaxation, with a relaxation time on the order of 3×10^{-3} s at room temperature. In more complicated polyatomic gases, it is rare for a single relaxation process to be solely effective, and rather elaborate analysis is required to take account of multiple relaxations.

Accompanying relaxation absorption in gases is sound dispersion, which is the variation of sound velocity with frequency. The dispersion leads to the distortion of waveforms and features prominently in areas such as sonic boom propagation. *See* NONLINEAR ACOUSTICS; SONIC BOOM.

Relaxation processes in liquids. The excess absorption of sound in liquids above the classically predicted values has also been attributed to relaxation processes. The assumption of thermal relaxation has been successful for certain nonpolar liquids such as benzene. It has not worked, however, for polar liquids such as water. For the latter, so-called structural relaxation has been successfully invoked. This assumes that liquid water can exist in two different states of mutual molecular orientation and that the passage of sound affects the average rate of energy transfer back and forth between the two states. The state of lower energy is assumed to be similar to the structure of ice, in which each water molecule is surrounded on the average (in spite of the random motion of the molecules) by its four nearest neighbors arranged at the corners of a regular tetrahedron with the given molecule in the center. The higher-energy state corresponds to a more closely packed configuration of the molecules, with an arrangement like that of a face-centered cubic crystal. In water in equilibrium, most of the molecules are in the lower-energy state, but the passage of sound causes a transfer of some molecules into the higher state and upsets the equilibrium. There is the usual lag in the return of the molecules to the lower state, leading to a relaxation time, and this produces absorption similar to that in gases. This theory has been successful in accounting for the excess sound absorption in water as a function of temperature in the range from 10 to 80°C (50 to 176°F). In this range, the appropriate relaxation time has been found to be of the order of 10^{-12} s.

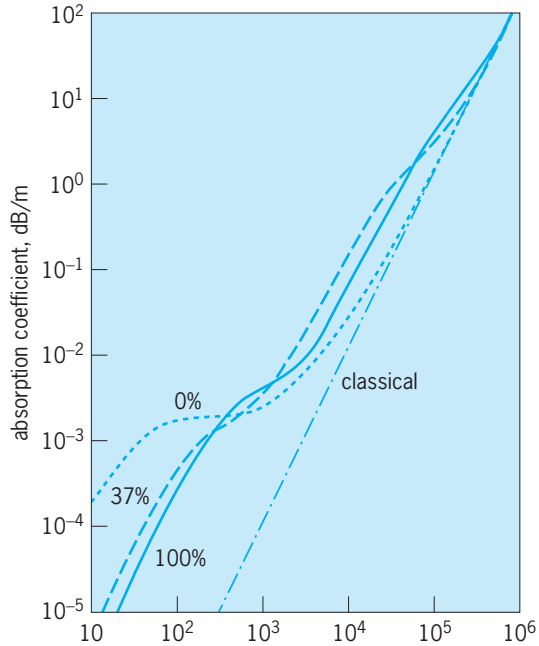

Fig. 1. Absorption of sound in air at 20°C (68°F) and various values of relative humidity. (*From L. E. Kinsler et al., Fundamentals of Acoustics, 3d ed., John Wiley, 1982; used by permission of John Wiley & Sons, Inc.*)

Excess absorption in aqueous solutions has been treated successfully by chemical relaxation. A chemical reaction in the liquid such as a dissociation-association process can provide two energy states, and the transfer of energy between these states will be affected by the compressive stress involved in the passage of a sound wave. The dissociation reaction of magnesium sulfate has been used to explain excess sound absorption in seawater.

In the atmosphere. The amount of sound that air absorbs increases with audio frequency and decreases with air density, but also depends on temperature and humidity. Sound absorption in air depends heavily on relative humidity, as seen in **Fig. 1**, which displays the predicted classical thermoviscous absorption as well as the actual absorption in air for varying relative humidity.

The reason for the strong dependence on relative humidity is molecular relaxation. One can note the presence of two transition regimes in most of the actual absorption curves, representing the relaxation effects of N_2 and O_2, the dominant constituents of the atmosphere. *See* ATMOSPHERIC ACOUSTICS.

In seawater. Sound absorption in water is generally much less than in air. It also rises with frequency, and it strongly depends on the amount of dissolved materials (in particular, salts in seawater), due to chemical relaxation (**Table 1**). *See* UNDERWATER SOUND.

In solids. The theory of sound attenuation in solids is complicated because of the presence of many mechanisms responsible for it. These include heat conductivity, scattering due to anisotropic material properties, scattering due to grain boundaries, magnetic domain losses in ferromagnetic materials,

TABLE 1. Sound absorption versus frequency for fresh water and seawater*

	Absorption coefficient, dB/m†		
	10 kHz	100 kHz	1 MHz
Fresh water	0.0001	0.006	0.2
Seawater	0.0008	0.04	0.3

*Reprinted with permission from L. Beranek, *Acoustical Measurements,* rev. ed., American Physical Society, 1988; © Acoustical Society of America.
†1 dB/m = 0.3 dB/ft.

interstitial diffusion of atoms, and dislocation relaxation processes in metals. In addition, in metals at very low temperature the interaction between the lattice vibrations (phonons) due to sound propagation and the valence of electrons plays an important role, particularly in the superconducting domain. *See* CRYSTAL DEFECTS; DIFFUSION; FERROMAGNETISM; SUPERCONDUCTIVITY.

At boundaries. When sound waves passing through air or water encounter another material or strike a surface, part of the incident wave energy is reflected, part is absorbed within the material, and still another part is transmitted. The general property of interest in acoustical design is the absorption coefficient, a, of the surface: the ratio of the energy that is absorbed and transmitted, to the energy incident on that boundary. For a surface in a room, this coefficient measures the fraction of the energy that is not returned to the room, independent of the mechanism of extraction. For example, an open window has $a = 1$, since none of the sound comes back into the room, even though none of it may be immediately absorbed by the air. On the other hand, the intensity absorption coefficient α, discussed previously, refers specifically to energy that has been transformed from acoustic energy to thermal energy within the medium during propagation.

There is no simple way of relating a, the sound absorption by a boundary, to α, the absorption of sound on passing through a continuous solid, liquid, or gas. The values of a used in acoustical design are typically empirically determined, being derived from laboratory or field tests of specimens of the particular material of interest.

Practical absorbing materials for airborne and water-borne sound are based on the mechanisms of viscosity and heat conduction. Most useful sound absorbers bring these mechanisms into play by being porous, perforated, or slotted yet having a significant resistance to the flow of air or water through their openings. The sound-wave movement through such openings dissipates energy by viscosity and thermal conduction with the walls. In the case of airborne sound, examples are mineral wool, fiberglass batts, acoustical ceiling tiles, and domestic drapes, all of which are very effective above 1 kHz. *See* ARCHITECTURAL ACOUSTICS; MUFFLER.

Material ratings. Full-sized material samples are typically tested in a laboratory reverberation room. The random-incidence sound absorption (absorption of sound waves coming from all directions) of acoustical materials is measured at frequencies between 125 and 4000 Hz, which corresponds to a significant fraction of the normal audio range encountered in speech and music. The average of the absorption at 250, 500, 1000, and 2000 Hz is called the noise reduction coefficient (NRC). Small material samples are tested in a device known as an impedance tube, which relates the incident sound wave to the reflected sound wave from the test specimen. In this case, the normal-incidence sound absorption (absorption of sound impinging along a line perpendicular to the absorber) is measured.

Applications to room design. Room walls, floors, and ceiling surfaces reflect and absorb sound differently in accordance with the acoustic properties of their construction material, that is, depending on whether they are concrete, glass, brick, or carpet, for example. This is depicted in **Fig. 2**, which shows how sound can arrive at a receiver from many paths with varying degrees of absorption along the route.

Sound generation in the average nonindustrial room occurs via speech, music, equipment or appliance noise, and occupant movements. The subsequent sound-level buildup or reverberation may be undesirably high or low depending on the intended use of the room. If the reverberation is too great, the room is typically referred to as a live room; if it is too low, the room is referred to as a flat room. Reverberation level can be modified by treatment of the room's interior surfaces by adding materials with greater or lesser sound absorptive qualities. The acoustic design of a room depends on its intended use, whether as a conference room, music room, lecture room, or open-plan office. For example, if the intended use is as a church, a greater reverberation may be desired for the typical music played. If the intended use is as an office space, less reverberation may be desired for speech communication. Conventional rugs, hanging drapes, stuffed furniture, and carpeting are typical sound absorption materials used in many applications. However, additional treatment is often necessary. *See* ACOUSTIC NOISE; SOUND FIELD ENHANCEMENT.

Once the appropriate materials are selected,

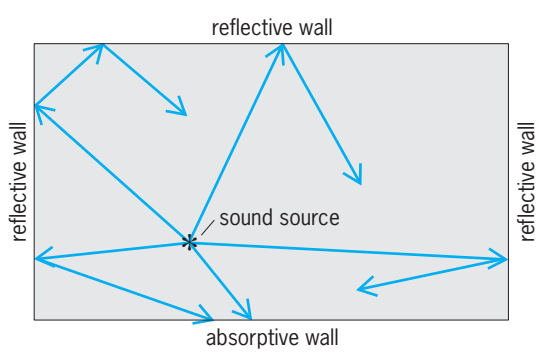

Fig. 2. Depiction of direct and reflected sound in a room with varying wall properties.

TABLE 2. Absorption coefficients of typical construction materials*

Material	Frequency, Hz					
	125	250	500	1000	2000	4000
Acoustic paneling	0.15	0.3	0.75	0.85	0.75	0.4
Brick	0.05	0.04	0.02	0.04	0.05	0.05
Brick, painted	0.01	0.01	0.02	0.08	0.02	0.03
Concrete block	0.36	0.44	0.31	0.29	0.39	0.25
Curtains	0.05	0.12	0.15	0.27	0.37	0.50
Floor, concrete	0.02	0.02	0.02	0.04	0.05	0.05
Floor, wood	0.15	0.2	0.1	0.1	0.1	0.1
Floor, carpeted	0.1	0.15	0.25	0.3	0.3	0.3
Glass fiber insulation, 2 in.	0.17	0.55	0.8	0.9	0.85	0.8
Glass, window	0.35	0.25	0.18	0.12	0.07	0.04
Gypsum board on studs	0.29	0.10	0.05	0.04	0.07	0.09
Plaster	0.03	0.03	0.02	0.03	0.04	0.05
Plywood paneling	0.28	0.22	0.17	0.09	0.10	0.11
Water surface	0.008	0.008	0.013	0.015	0.020	0.025
One person	0.18	0.4	0.46	0.46	0.51	0.46

*From C.E. Wilson, *Noise Control*, Krieger, 1994.

usually with the help of a professional acoustical consultant, the area of treatment, A, is calculated by means of the Sabine formula [Eqs. (7)] to determine

$$T = 0.049\frac{V}{aA} \qquad \text{(English)} \qquad (7a)$$

$$T = 0.161\frac{V}{aA} \qquad \text{(metric)} \qquad (7b)$$

the reverberation time T required for the intended use. In the Sabine formula, V is the room volume in cubic feet or cubic meters, a is the absorption coefficient of the acoustical material, and A is the area in square feet or square meters of the room. If the existing reverberation time for a room is unsatisfactory, acoustical materials with varying values of a can be used to replace portions of the room surface A to modify the reverberation time. The value of a for common acoustical materials depends on frequency; typical values are indicated in **Table 2**.

Industrial applications. In industrial settings it is the overall sound level rather than the quality and intelligibility of the sound which is typically the driving concern. Industrial noise is often at levels sufficient to cause damage to hearing if it is not reduced or if hearing protection (such as headphones, which absorb sound before it enters the ear) is not used. Noise reduction is often achieved by adding sound-absorbing materials and enclosures around the noise source. Qualitatively, the treatment of a noisy industrial setting is quite similar to the treatment of a noisy office space, but special care must be paid to material selection due to the harsher operating environment of a typical factory.

Henry E. Bass; Angelo J. Campanella; James P. Chambers; R. Bruce Lindsay

Bibliography. R. T. Beyer and S. V. Letcher, *Physical Ultrasonics*, 1969; M. J. Crocker (ed.), *Handbook of Acoustics*, 1998; L. E. Kinsler et al., *Fundamentals of Acoustics*, 3d ed., 1982; A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications*, 2d ed., 1989; C. E. Wilson, *Noise Control*, 1989.

Sound field enhancement

A system for enhancing the acoustical properties of both indoor and outdoor spaces, particularly for unamplified speech, song, and music. Systems are subdivided into those that primarily change the natural reverberation in the room, increasing its level and decay time (reverberation enhancement systems); and those that essentially replace the natural reverberation (sound field synthesis systems). Both systems may use amplifiers, electroacoustic elements, and signal processing to add sound field components to change the natural acoustics. Sound field enhancement is used to produce variable acoustics, to produce a particular acoustics which is not attainable by passive means, or because a venue has one or all of the following deficiencies: (1) unsuitable ratio between direct, early reflected, and reverberant sound; (2) unsatisfactory early reflection pattern; and (3) short reverberation times. *See* AMPLIFIER.

The purpose of sound field enhancement systems is not to provide higher speech intelligibility or clarity—in contrast to sound reinforcement systems—but to adjust venue characteristics to best suit the program material and venue, and in such a way optimize the subjectively perceived sound quality and enjoyment. *See* SOUND-REINFORCEMENT SYSTEM.

The use of sound field enhancement ranges from systems for homes (surround sound systems and home theater sound systems) to systems for rooms seating thousands of listeners (concert halls, performing arts centers, opera houses, and churches). Outdoor spaces may have elaborate designs to approach the sound quality of indoor spaces (advanced concert pavilions, outdoor concert venues). Systems may be used to improve conditions for both performers and listeners. *See* ARCHITECTURAL ACOUSTICS; REVERBERATION.

Basic elements. A sound field enhancement system may consist of one or more microphones, systems for amplification and electronic signal processing, and one or more loudspeakers. The signal processing equipment may include level controls, equalizers, delays and reverberators, systems for prevention of howl and ringing, as well as frequency dividing networks and crossover networks. *See* EQUALIZER; LOUDSPEAKER; MICROPHONE.

Prerequisites. Many indoor venues have acoustical characteristics unsuitable for communication such as speech, song, or music. Outdoor venues seldom have satisfactory acoustical conditions for unamplified acoustic communication. Satisfactory acoustic conditions include for speech, good speech intelligibility and sound quality; and for music, appropriate loudness, clarity, reverberation envelopment, spaciousness, and other sound quality characteristics for the music to be performed and enjoyed. Rooms with insufficient reverberation are often perceived

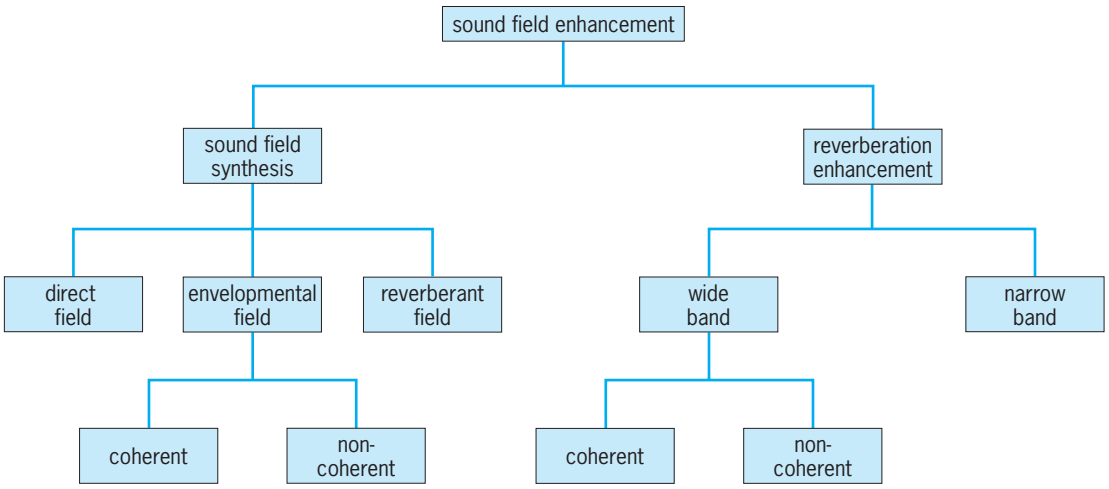

Fig. 1. Classification scheme for sound field enhancement systems.

to be acoustically "dry" or "dead." Insufficient reverberation usually is due to lack of room volume for the existing sound absorption areas in the room. The reverberation time is essentially proportional to room volume and inversely proportional to sound absorption, as indicated by Sabine's formula. In these cases, sound field enhancement systems may be used to improve acoustical conditions, by increasing the sound level and reverberation time. Absence of noise is also necessary for good acoustic conditions. *See* ACOUSTIC NOISE; SOUND ABSORPTION.

System design. A simple reverberation enhancement system uses a microphone to pick up the sound signal to be reverberated, amplified, and reradiated in the auditorium. In most cases, reverberation is added to the original signal by electronic digital signal processing. An advanced sound field enhancement system uses several, possibly directional microphones to pick up sound selectively from the stage area. Signal processing and many loudspeakers are used to obtain the desired sound field for the audience and performers. Most performing art centers need variable acoustic conditions that can be changed instantaneously using electronic controls.

Categories. Sound field enhancement systems can be classified according to different criteria. One possible scheme is shown in **Fig. 1**. Sound field synthesis systems are designed to provide the necessary sound field components, such as early reflections from side walls and ceiling (the latter are usually called envelopmental sound), and reverberation by using digital filters, which can be designed to imitate a specific desired room response. Reverberation enhancement systems are designed to primarily increase the reverberation time and sound level of the reverberant field, while having negligible influence on the direct sound and the early reflected sound. Reverberation enhancement systems can be further subdivided into regenerative systems, based on the acceptance of unavoidable positive feedback between loudspeakers and microphones to increase the reverberation time; and nonregenerative systems, based on the use of electronic reverberators. Regenerative systems can be designed so as to work over a wide or narrow frequency range; nonregenerative systems are usually wide-frequency-range systems. *See* SOUND; REFLECTION OF SOUND.

Routing. A sound field synthesis system can generally be modeled as shown in **Fig. 2**. The system uses interconnected microphones and loudspeakers to form a more or less complete matrix of channels, where signal processing varies between channels. The signals from the microphones to the loudspeakers can be sent either directly between a specific microphone and loudspeaker pair, or in some form of matrix system feeding the signal from a microphone to several loudspeaker channels. A matrixed system can be designed as a phased array loudspeaker system. Another possibility is for a matrixed system to act as multiple individual systems, with uncorrelated or correlated outputs. Digital signal processing may be used in the circuits for various reasons.

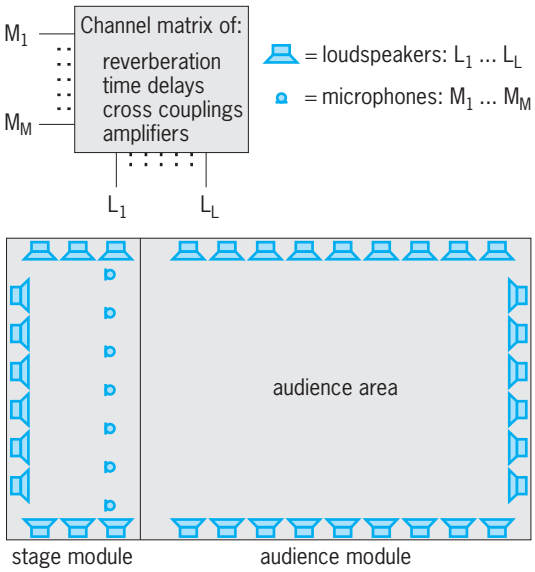

Fig. 2. A sound field enhancement system can generally be modeled as a matrix. The system uses microphones (M) and loudspeakers (L).

Direct sound and room reverberation. The concept of room impulse response is useful in understanding sound field enhancement. The room impulse response describes the sound pressure in a room as a result of a very short sound pulse. Hearing uses binaural analysis to separate various types of information such as direction of incidence, direct sound from reverberation, and so on. The binaural room impulse responses for various source and receiver positions characterize the acoustics of the venue. *See* ACOUSTICS; BINAURAL SOUND SYSTEM; SOUND PRESSURE.

It is common to subdivide the (binaural) room impulse response in the range above 250 Hz into three temporal sections. For music in large venues, these sections are often direct (0–20 milliseconds), envelopmental (20–150 ms), and reverberant sound (beyond 150 ms), all times referred to after the arrival of the direct sound from the source to a listener's location. The time ranges are approximate, somewhat individual, and depend on the frequency range considered. The desirable decibel levels of the three sound field components (direct, envelopmental, and reverberant sound) depend on the use of the venue—for example, whether it is to be used for an organ recital, a symphony, an opera, or a speech. Conditions for optimizing these levels can be attained instantaneously by electronic means using a sound field enhancement system, which can be used to change the reflection patterns as well as the relative decibel level between the different sound field components. The spatial characteristics of the acoustics, such as spaciousness and envelopment, are mainly determined by the directional properties of the envelopmental sound and the reverberant sound.

Direct sound. Direct sound is the combination of both the true direct sound (which has traveled directly from source to listener) and the various early reflections which follow within the first 20 ms. Direct sound usually determines the perceived direction of the sound source. In active systems, unwanted shift of the apparent source direction may occur due to excessively strong additional sound components from unwanted directions. Halls for classical music may need the direct sound to be weakened relative to the envelopmental and reverberant sound, if the clarity is excessive. Amplification of direct sound is needed when (1) the sound source is weak, (2) sound is attenuated while passing over an audience area, (3) the venue is very large, and (4) the performance space is outdoors.

Envelopmental sound. The envelopmental sound has been reflected against walls and ceilings relatively few times, particularly in typical halls seating 1000–3000 people. The envelopmental sound usually is a major part of the energy in the sound reaching the listener, and is therefore critical to the balance of the three sound field components. In many wide halls, lacking sufficiently reflecting side walls and ceiling, or in halls with absorptive balconies on the side walls, the envelopmental sound may be too weak, and it will then be necessary to add envelopmental sound. One of the major problems in enhancing envelopmental sound is obtaining equal coverage over the audience area.

Reverberant sound. The impression of reverberation on classical music is determined by the decay characteristics of the envelopmental sound and reverberant sound, as well as by the degree of correlation between these sounds at the listener's ears. The degree of interaural cross-correlation influences the perceived envelopment and spaciousness. Both the decay rate and relative level of the reverberant sound are critical to the overall impression of reverberation. However, on "running" music it is usually only the envelopmental sound and the early part of the reverberation of a room which is heard, because new music signals continue to arrive to be analyzed by hearing. Therefore, the perceived reverberation characteristics for running music and those for a final chord may be different, the latter reverberation being audible longer.

Feedback. In any real environment there will be unavoidable positive acoustic feedback from the loudspeaker(s) back to the microphone(s) (**Fig. 3**.) The feedback gives rise to regeneration, which may be controlled in various ways. Uncontrolled regeneration, that is, system oscillation (usually called howl), is obtained at and above a critical channel gain setting. Other effects of regeneration are ringing (that is, temporal artifacts such as flutter echo, primarily in wide-bandwidth systems) and coloration (that is, spectral artifacts such as frequency response peaks). From a control viewpoint, the feedback is characterized by its loop gain, which depends on the microphones, loudspeakers, amplification, signal processing, and the room itself. *See* GAIN; OSCILLATION.

Stability and coloration. Because of the acuity of hearing, coloration is noticed at much lower loop gains than oscillation because sound at some frequencies is exaggerated. An example of frequency response curves at some different gain settings of a feedback system involving a room is shown in **Fig. 4**.

The regeneration also prolongs the decay time of the signal because the sound energy from the

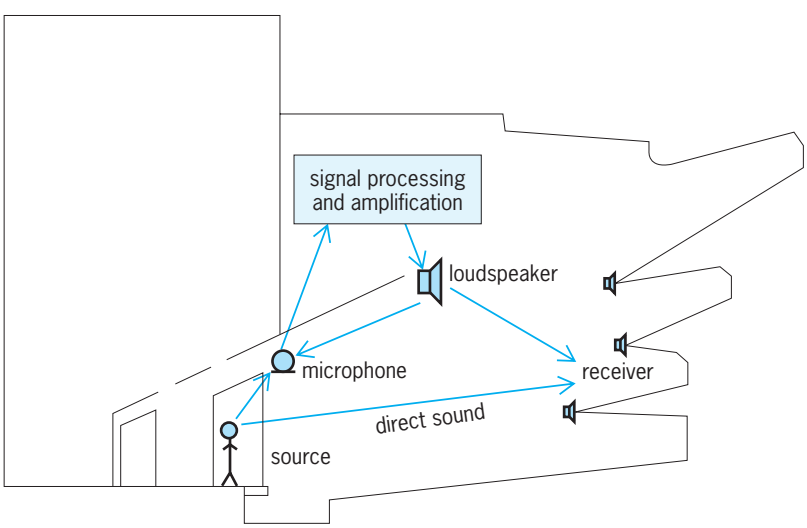

Fig. 3. Basic signal paths of a sound field enhancement system showing an acoustic feedback channel from a loudspeaker back to a microphone.

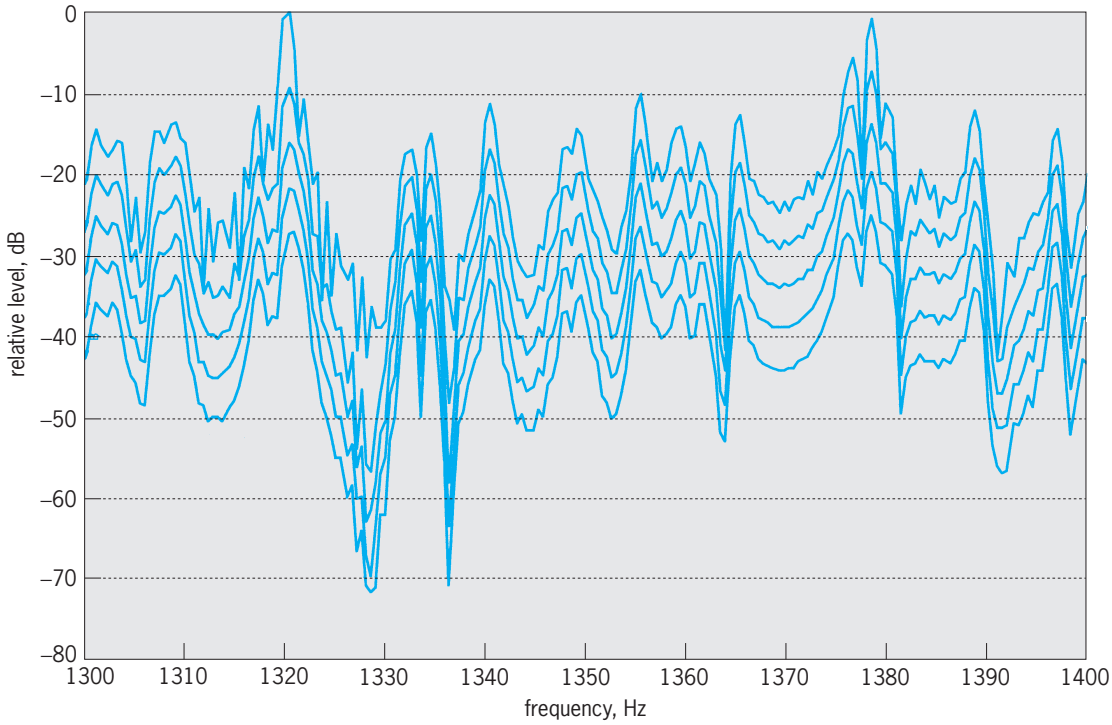

Fig. 4. Example of frequency response curves for a room with a feedback system. The upper curve shows response for loop gain setting close to oscillation; the lower curves shows response for gain settings of −5, −10, −15, −20 dB relative to this level.

loudspeakers compensates the sound absorbed by audience, chairs, and so on, leading to an increased reverberation time. In simple systems, regeneration effects are subjectively annoying since only sound at a few frequencies is subject to increase of reverberation time and level. In a system consisting of many channels, the gain can be adjusted so that each channel has low gain and coloration, and therefore the combined effect of such a system will yield both prolonged reverberation, over the entire operating frequency range of the channels, and negligible coloration.

Regenerative systems use this feedback as a means of achieving the desired sound field, whereas nonregenerative systems try to achieve the desired conditions with as little feedback as possible in order to have maximum control. Sound field synthesis systems are usually nonregenerative. To change acoustic conditions outdoors, only sound field synthesis systems are used.

Nonregenerative systems. Sound field synthesis can often be used to good advantage, provided that the environment is acoustically dry, and regeneration has been minimized by use of directional microphones and loudspeakers. It can also be used when the signal is obtained from a source other than microphones, for example in a surround sound system or a home theater system, in which case there is no possibility of regeneration. In these systems, both envelopmental and reverberant sound is simulated using signal processing and is radiated via side and rear loudspeakers.

Direct sound enhancement. Systems for amplification of direct sound may be configured as single-channel or multichannel systems. The sound level may be increased relative to that of the unamplified direct sound by using the masking effect of the direct sound. Any substantial positive acoustic feedback will give rise to coloration. Therefore, these systems require directive or close-mounted microphones, as well as highly directive loudspeakers radiating primarily at sound absorptive surfaces in the room, such as audience areas, to minimize feedback. Loudspeakers should be near the stage so that the audible directional impression will agree with the visual impression.

Envelopmental sound enhancement. Improved envelopmental sound can also be implemented along the same lines as direct sound amplification. The delayed microphone (or other) signals should arrive in a suitable time pattern at all places in the audience area. Digital filters are used to provide single or multiple delays. Because of the limited number of channels and the need for wide sound coverage and substantial envelopmental sound levels, systems have to be protected from positive feedback, which could otherwise give rise to coloration. *See* DIGITAL FILTER.

Envelopmental sound enhancement can be achieved by both coherent and noncoherent designs. In coherent designs, arrays of close-mounted loudspeakers are used.

An example of a coherent system is the wave field synthesis approach, the most advanced of the nonregenerative systems suggested so far. One can consider it as a holographic system because it uses an array of microphones covering the source, and an extensive, dense array of loudspeakers surrounding the listening area. It is possible to approximately

simulate the reflected waves of a space much larger than the one in which the system is installed. For the system to work as intended, the transducers must be very close and individually steered from the matrix. *See* TRANSDUCER.

It is also necessary to have some form of adaptive control over the microphones so that the microphone array can effectively track the sources (to avoid noise problems). In early implementation attempts, it was necessary to severely compromise these principles. A noncoherent approach is used by a beacon system, where integrated, small (wall-mounted) microphone-loudspeaker channels pick up sound and retransmit it with suitable delay and reverberation characteristics. Adaptive systems can be used to control and eliminate the internal feedback in beacon systems. If a number of beacons are used in a moderately sound absorbing or sound diffusing surface, one can create an illusion of a more remotely positioned sound reflecting surface or object. *See* ACTIVE SOUND CONTROL.

Reverberant sound enhancement. Reverberation is obtained along the same lines as the envelopmental sound enhancement systems except that many more and quite evenly spaced loudspeakers are necessary to achieve good coverage over the audience area. The signal is sent through long finite or infinite impulse response digital filters to be reverberated. Many techniques are available to design infinite impulse response filters with little coloration. In practice, however, finite impulse response filters offer superior sound quality, and can be based on calculated or measured room impulse responses. Usually several filters, having insignificant correlation, will be used.

In acoustically dry rooms, the reverberation responses of the filters dominate the reverberation characteristics heard by the audience. In practice, the unavoidable scattered sound by various sound reflecting surfaces (for example, chair back rails) will fill in and smooth the reverberant sound field from the loudspeaker system.

Regenerative systems. The primary reason for using regenerative systems is that some regeneration (positive feedback) will always be present in any sound field enhancement system. Thus, it might be better to try to turn the coloration due to regeneration into something useful instead of trying to eliminate the coloration by other, perhaps more audible methods. Regenerative systems are used only for enhancement of reverberant sound. Microphones are thus placed far from both sound sources and system loudspeakers.

Noncoherent systems. In noncoherent regenerative system, microphones and loudspeakers are not positioned in regular arrays, as is often the case in sound field synthesis systems, but are spread out over the room (hung from ceilings or wall-mounted).

In the simplest systems, each microphone–signal processor–loudspeaker channel in the system is allowed to act independently. The microphones and loudspeakers are interconnected to form a matrix of channels where only the diagonal contains nonzero coefficients. Each channel includes a level control, equalizer, and power amplifier (and possibly a system for minimizing regeneration). This is the approach of the classical reverberation enhancement systems, both narrow and wide frequency band. More complex systems may use microphones and loudspeakers to form a more or less complete matrix of channels of pair combinations, having nonzero transfer functions in all channels. *See* POWER AMPLIFIER.

Coherent systems. Coherent regenerative systems use the same basic approach as more complex noncoherent systems, but channels are interconnected using various signal processing to form many coherent channels, reminiscent of some sound field synthesis systems.

Control of regeneration. There are multiple ways of controlling the loop gain in order to reduce the problems of instability and coloration. The loop gains of the electroacoustic channels vary considerably with frequency and location. The expected value of the peak to mean ratio of the sound pressure level response of a room, as a function of frequency, is not very dependent on the room reverberation time for wide-frequency-range sounds. One can show that the instability condition in a system channel will be fulfilled approximately 10 dB before wideband instability. On the average, the frequency intervals between peaks are inversely proportional to reverberation time.

Directional transducers. The loop gain can be reduced, for example, by using highly directive microphone and loudspeaker systems. Because of sound reflections from walls and so forth, it is usually not possible to achieve a large reduction in the loop gain in this way. Even if loudspeaker systems for envelopmental sound are designed to radiate primarily at the (sound-absorbing) audience areas, their side lobes and the reflection of sound from room surfaces set a limit to the reduction of the loop gain. Microphone systems suffer the same limitations whether highly directive microphones or adaptive steered arrays are used. In any case, the directive properties of the sound sources cause problems. Most music sound sources are quite directional, and even speech is so directive that it is probably necessary to use personal microphones in order to avoid detrimental effects due to directivity when performers are moving.

Adaptive filters. It is impossible to predict the exact location in frequency of the loop gain peaks because of changes in room temperature, movement by persons, and so on. Commercial adaptive electronic notch (band rejection) filters can be used which automatically detect and remove unwanted loop gain peaks. This is necessary because after a peak has been controlled, once the channel gain is increased a new peak will quickly appear.

Modulation by linear periodic time-variant filters. Under typical conditions, the mean distance between adjacent sound pressure level maxima (in a narrow frequency band) is approximately a wavelength. Consequently, higher loop gain or less coloration can be obtained

by moving the sound source or the loudspeakers, or by adding a small shift to the signal frequency. Moving loudspeakers and sources is impractical. Instead, time variance is achieved by use of digital signal processing to modulate the loop gain. Several modulation methods are possible such as amplitude, phase, delay, and frequency modulation. All time variance introduces sidebands, that is, extra frequency content. *See* AMPLITUDE MODULATION; FREQUENCY MODULATION; MODULATION; PHASE MODULATION.

The modulation methods vary as to audibility. Generally, slow modulation (low modulation frequency) is heard as beats, whereas fast modulation (high modulation frequency) is heard as roughness. The sensitivity of hearing to modulation is highest for modulation frequencies around 2–4 Hz, which is unfortunate since this frequency range is often desirable from a technical viewpoint to average the room frequency response. To minimize unwanted audible effects, the modulation frequency should be kept as low as possible, preferably below 1 Hz. Low-pass, filtered random signals can be used for modulation. Most commercial reverberation enhancement systems use some form of modulation. A challenge for future room acoustics research is to design new modulation methods, inaudible yet effective at preventing regeneration. *See* SOUND-REPRODUCING SYSTEMS.

Mendel Kleiner

Bibliography. W. Ahnert and F. Steffen, *Beschallungstechnik, Grundlagen und Praxis*, S. Hirzel, Leipzig, 1993; M. Barron, *Auditorium Acoustics*, E & FN Spon, London, 1993; L. Beranek, *Concert and Opera Halls*, Acoustical Society of America, Woodbury, NY, 1996; H. Kuttruff, *Room Acoustics*, 4th ed., Spon Press, London, 2000; M. Kleiner and P. Svensson, A review of active systems in room acoustics and electroacoustics, *ACTIVE95 Proceedings of the International Symposium on Active Control of Sound and Vibration*, Newport Beach, CA, July 1995, included in *Proceedings of the Sabine Centennial Symposium on CD ROM* by Acoustical Society of America Publications, P.O. Box 1020, Sewickley, PA 15143-9998.

Sound intensity

A fundamental acoustic quantity which describes the rate of flow of acoustic energy through a unit of area perpendicular to the flow direction. The unit of sound intensity is watt per square meter. The intensity is calculated at a field point (x) as a product of acoustic pressure p and particle velocity u. Generally, both p and u are functions of time, and therefore an instantaneous intensity vector is defined by Eq. (1). The time-variable instantaneous intensity,

$$\vec{I}_t(x, t) = p(x, t) \cdot \vec{u}(x, t) \qquad (1)$$

$\vec{I}_t(x, t)$, which has the same direction as $\vec{u}(x, t)$, is a nonmoving static vector representing the instantaneous power flow through a point (x). *See* POWER; SOUND PRESSURE.

Frequency spectrum. Many acoustic sources are stable at least over some time interval so that both the sound pressure and the particle velocity in the field of such a source can be represented in terms of their frequency spectra. The static sound intensity vector $\vec{I}(x, \omega)$ represents the time-independent sound power, which propagates through the point (x) in the direction of the particle velocity. The vector $\vec{I}(x, \omega)$ can be calculated at a frequency ω (or frequency band $\Delta\omega$) from Eq. (2), where φ is the phase

$$\vec{I}(x, \omega) = p(x, \omega) \cdot \vec{u}(x, \omega) \cdot \cos(\varphi) \qquad (2)$$

shift between the real quantities p and u. It is practical to define both p and u as complex quantities so that the intensity vector $\vec{I}$ can be calculated from Eq. (3), where * denotes a complex conjugate.

$$\vec{I}(x, \omega) = {}^1\!/_2 Re[p(x, \omega) \cdot \vec{u}^*(x, \omega)] \qquad (3)$$

Because intensity represents power, $\vec{I}$ is always a real quantity. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES; PHASE (PERIODIC PHENOMENA).

Measurement. The applications of sound intensity were fully developed after a reliable technique for intensity measurement was perfected. Sound intensity measurement requires measuring both the sound pressure and the particle velocity. Very precise microphones for sound-pressure measurements are available. Because there are no microphones for a direct particle velocity measurement, an indirect technique based on Euler's equation (4), which links

$$\vec{u}(x, \omega) = -\frac{1}{jk\rho c}\frac{\partial p(x, \omega)}{\partial x} \qquad (4)$$

u with a pressure gradient, is used. Here, ρ is the density of the fluid medium, c is the sound speed, k is the wave number ($k = \omega/c$), and $j = \sqrt{-1}$. *See* MICROPHONE.

Measuring the sound pressure at close points with two precision microphones approximates the pressure gradient. To keep the measurement errors at an acceptable level, both microphones must have identical sensitivity and be phase-matched to 0.05°. The calculation of sound intensity is usually obtained by feeding the microphone output into a dual-channel fast-Fourier-transform (FFT) analyzer. The intensity is proportional to the imaginary part of the cross-spectrum of the input voltages. This measurement arrangement measures the intensity component in the direction of an axis connecting the two microphones. To determine the total intensity vector, three perpendicular components must be obtained by measuring along three mutually perpendicular axes. Microphone probes, consisting of six microphones, to implement such a measurement have been constructed. Both international and national standards define the requirements on the instrumentation to secure the needed measurement precision. *See* FOURIER SERIES AND TRANSFORMS.

The usefulness of the measurement technique based on two microphones goes beyond the measurement of sound intensity. The sound pressure can be obtained by interpolating the pressure measured by the two microphones. Similarly, both the potential and kinetic energy densities as well as the particle velocity vector can be obtained.

Applications. One essential application of the intensity technique is the measurement of sound power radiated from sources. Such a measurement can be done with a precision unobtainable from pressure measurements. The radiated sound power is obtained from intensity measurements over an area circumscribing the source. The knowledge of the radiated power makes it possible to classify, label, and compare the noise emissions from various pieces of equipment and products and to provide a reliable input into environmental design. The intensity technique is perfect for transmission and insertion-loss measurements. Because the measurements can be done in the near field of acoustical radiators, a reliable overall functional analysis and identification of sound radiating sources can be performed. General sound-field studies and mapping are essential for many applications. *See* SOUND. Jiri Tichy

Bibliography. F. J. Fahy, *Sound Intensity*, 2d ed., E & FN Spon (an imprint of Chapman & Hall), London, 1995.

Sound pressure

The incremental variation in the static pressure of a medium when a sound wave is propagated through it. Sound refers to small-amplitude, propagating pressure perturbations in a compressible medium. These pressure disturbances are related to the corresponding density perturbation via the material equation of state, and the manner in which these disturbances propagate is governed by a wave equation. Since a pressure variation with time is easily observed, the science of sound is concerned with small fluctuating pressures and their spectral characteristics. The unit of pressure commonly used in acoustics is the micropascal ($1\ \mu\text{Pa} = 1\ \mu\text{N/m}^2 = 10^{-5}\ \text{dyne/cm}^2 = 10^{-5}\ \mu\text{bar}$). One micropascal is approximately 10^{-11} times the normal atmospheric pressure. *See* PRESSURE; PRESSURE MEASUREMENT; WAVE MOTION.

The instantaneous sound pressure at a point can be harmonic, transient, or a random collection of waves. This pressure is usually measured with an instrument that is sensitive to a particular band of frequencies. If the pressure is measured in a specified frequency band over a short period of time, the mean value of the fluctuating pressure is found to be zero, but the mean value of the square of the pressure, $\langle p^2 \rangle$, is a positive quantity. The square root of $\langle p^2 \rangle$ is the root-mean-square pressure, p_{rms}, and is proportional to the time average of the sound intensity, $\langle I \rangle$. The measured pressures will depend on the bandwidth and filter response of the measurement instrument. The peak sound pressure for any specified time interval is the maximum absolute value of the instantaneous sound pressure in that interval and is equal to $\sqrt{2}\ p_{\text{rms}}$ for sinusoidal signals. Acoustic pressures are commonly studied as a function of frequency, where the frequency can range over the sonic band, which lies between 20 and 20,000 Hz and is bounded by the infrasonic and ultrasonic frequency bands. The Fourier transform and its digital implementation, the fast Fourier transform (FFT), are widely used to characterize the spectrum of the $\langle p^2 \rangle$ versus frequency. Thus $\langle p^2 \rangle$ in a given band, Δf, is often normalized by this bandwidth to yield the spectral density, $\langle p^2 \rangle / \Delta f$, with units of $[\mu\text{Pa}]^2/\text{Hz}$. *See* FOURIER SERIES AND TRANSFORMS; SOUND INTENSITY.

A concept widely used in acoustics is "level," which refers to the logarithm of the ratio of any two field quantities. When the ratio is proportional to a power ratio, the unit for measuring the logarithm of the ratio is called a bel, and the unit for measuring this logarithm multiplied by 10 is called a decibel (dB). The sound intensity, I, which describes the rate of flow of acoustic energy (acoustic power flow) per unit area, is given by the mean square pressure, $\langle p^2 \rangle$, divided by the acoustic impedance, defined as the product of the medium density, ρ, and compressional wave speed, C; that is, the intensity $I = \langle p^2 \rangle / \rho C$. A ratio proportional to power can be formed if a reference quantity is specified. The reference pressure for water is usually chosen to be 1 μPa with $\rho C \approx 1.545 \times 10^6\ \text{Pa} \cdot \text{s/m}$, and the reference pressure for air is 20 μPa with $\rho C \approx 413\ \text{Pa} \cdot \text{s/m}$. The effective sound pressure level (SPL, symbolized L_p) is a useful quantity since the power transmitted by a sound wave can be related to $\langle p^2 \rangle$ for a specified medium; however, when comparing levels in seawater to those in air, the acoustic impedance difference between the two media must be considered. With these considerations in mind, the SPL [dB re p_{ref}] is given by Eq. (1).

$$\text{SPL [dB re p}_{\text{ref}}] = L_p = 10 \log_{10}(\langle I \rangle / \langle I_{\text{ref}} \rangle)$$

$$= 10 \log_{10} \left(\langle p^2 \rangle \Big/ \left\langle p_{\text{ref}}^2 \right\rangle \right) \qquad (1)$$

See DECIBEL.

When the sound pressure level is restricted to a frequency band, the term "band sound pressure level" (BSPL, symbolized L_{pb}) is used. Air acoustics standards use a weighted filter response such as an A or C filter with either fast (F, 125-ms) or slow (S, 1000-ms) exponential time averaging. In these instances the levels are designated by L_A [dB re 20 μPa] or L_C [dB re 20 μPa]. When the mean square pressure is normalized by the bandwidth, the quantity becomes a density and is referred to as a spectral density or a spectrum level with the sound pressure spectrum level (SPSL, symbolized L_{ps}) defined by Eq. (2).

$$\text{SPSL [dB re} \langle p^2 \rangle_{\text{ref}}/\text{Hz}] = L_{ps}$$

$$= 10 \log_{10}[(\langle p^2 \rangle / \Delta f)/(\langle p^2 \rangle / \Delta f)_{\text{ref}})] \qquad (2)$$

See NOISE MEASUREMENT; SOUND. William M. Carey

Sound recording

The technique of entering sound, especially music, on a storage medium for playback at a subsequent time. The storage medium most widely used is magnetic tape. *See* MAGNETIC RECORDING.

Monophonic sound recording. In the simplest form of sound recording, a single microphone picks up the composite sound of a musical ensemble, and the microphone's output is recorded on a reel of $^1/_4$-in. (6.35-mm) magnetic tape. This single-track, or monophonic, recording suffers from a lack of dimension, since in playback the entire ensemble will be heard from a single point source: the loudspeaker.

Stereophonic sound recording. An improved recording, with left-to-right perspective and an illusion of depth, may be realized by using two microphones which are spaced or angled appropriately and whose outputs are routed to separate tracks on the tape recorder. These two tracks are played back over two loudspeakers, one each on the listener's left and right. Under ideal conditions this stereophonic system will produce an impressive simulation of the actual ensemble, and the listener may identify the apparent location of each integral part of the ensemble. *See* STEREOPHONIC SOUND.

As early as 1931, Alan Blumlein applied for a patent on several stereophonic microphone placement techniques. Despite the advances in recording technology that have been made since then, the methods described in his patent may still be found in use today. One example uses two microphones constructed within a single housing. The microphones' diaphragms are at right angles to each other, and when the device is placed at a proper location, the diaphragms point toward the left and right sides of the ensemble.

The microphone picks up essentially what listeners would hear if they were in the same room. Instruments at the rear are heard imperceptibly later than those placed closer to the microphone. A complex mixture of direct and reflected sound is heard, and although a listener would scarcely be conscious of these subtle relationships, the brain uses this information to analyze the size of the ensemble and the acoustic conditions of the playing room. *See* HEARING (HUMAN).

The successful use of this technique demands an acoustically excellent concert hall or recording studio, and an ensemble that is musically well balanced. The recording engineer has little or no technical control over the recorded sound, beyond moving the microphone device to a more favorable location. *See* SOUND-RECORDING STUDIO.

Binaural sound recording. In this system, two microphones are placed on either side of a small acoustic baffle in an effort to duplicate the human listening condition. The recording is played back over headphones, so that the microphones are, in effect, an extension of the listener's hearing mechanism. Although the reproduced program is most faithful to the original, headphone listening restricts the listener's freedom to move about and excludes others from hearing the program. Loudspeakers are, of course, more widely used, yet the distance between them, as well as their distance from the listener, somewhat defeats the binaural effect, and other stereophonic microphone placements are apt to prove more satisfactory.

Multitrack sound recording. To give the recording engineer more technical control over the recording medium, many recordings are now made using a multiple-microphone technique. In place of the stereo microphone, one or more microphones are located close to each instrument or group of instruments. *See* MICROPHONE.

In the control room the engineer mixes the outputs of all microphones to achieve the desired musical balance. As a logical extension of this technique, the microphone outputs may not be mixed at the time of the recording, but may be routed to 24 or more tracks on a tape recorder, for mixing at a later date.

When many microphones are so used, each instrument in the ensemble is recorded—for all practical purposes—simultaneously. The complex time delay and acoustic relationships within the room are lost in the recording process, and the listener hears the entire ensemble from one perspective only, as though he or she were as close to each instrument as is each microphone. Electronic signal processing devices may not be entirely successful in restoring the missing information, and the listener hears a recording that may be technically well executed, yet lacking the apparent depth and musical cohesiveness of the original.

However, the multitrack technique becomes advantageous when it is impractical or impossible to record the entire ensemble at once stereophonically. Especially in the case of new music, it may be impractical to spend the amount of recording time required to achieve an ideal balance. By recording the various sections on separate tracks, the engineer may experiment with balances later on, after the musicians have left.

In contemporary popular music, each member of a small group may wish to play or sing more than one part. For example, a basic rhythm section may be recorded, with each instrument routed to a separate track on the multitrack tape recorder. To keep the instruments acoustically isolated from each other, the engineer may place the microphones as close up as possible. For further isolation, acoustic baffles, or goboes, may be placed between instruments, and particularly subdued instruments may be placed within isolation booths.

As a consequence of this technique, each microphone picks up only the direct sound of the instrument in front of it. The echo and reverberation within the room, and the sound of the other instruments, are not heard. Therefore, by processing the output of any one microphone or tape recorder track, the engineer may make major changes in the sound of a musical instrument without affecting any of the others.

After the basic rhythm section has been recorded,

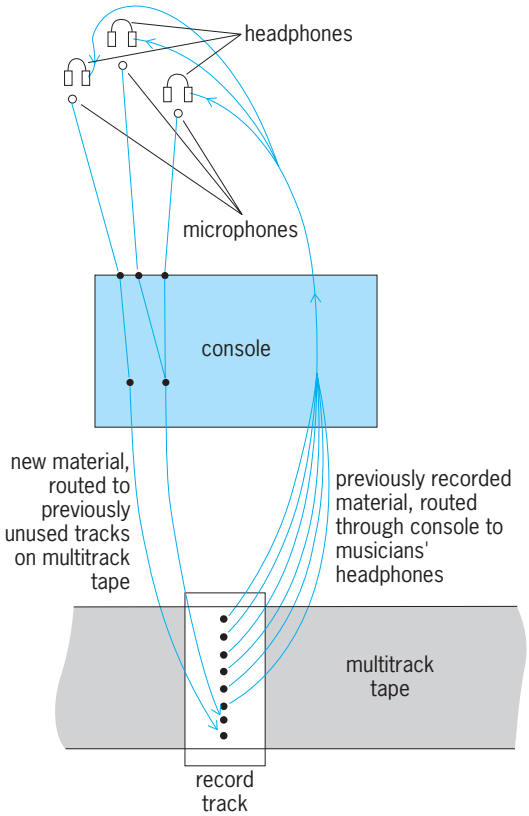

Fig. 1. Adding tracks to an unused recorder tape.

the group may wish to add some additional parts. Or at a later date an arranger may write a string or brass accompaniment, after listening to the rhythm tracks. In either case, the musicians listen over headphones to the material already recorded and play along in accompaniment. The accompaniment is recorded on one or more of the previously unused tracks (**Fig. 1**).

Since the tape recorder's playback head is some distance from the record head, there is a slight time delay after the tape passes the record head until it reaches the playback head. If the musicians listened to the previously recorded program via the playback head, this time delay would place the new material on the tape out of synchronization with the original program. Consequently, the record head is used for monitoring whenever new material is being added. Thus it performs a dual function simultaneously: playback of the previously recorded program, and recording of the new material on previously unused tracks.

Multitrack recording console. A modern recording console gives the engineer the flexibility required to meet the demands of multitrack sound recording. At the console, microphones may be quickly routed to any track on the tape recorder, and just as quickly reassigned to other tracks as the recording progresses. Since the musicians in the studio may not be able to hear each other due to the acoustic isolation between instruments, the console provides a studio headphone monitoring system, which the engineer may regulate to give the musicians the balance required. Previously recorded material is also routed through the console to the musicians' headphones when the performers are to play along in accompaniment. Although any of this material can be monitored over a loudspeaker in the studio, headphones are preferred so that the microphones do not again pick up the previously recorded program.

In the control room the engineer may wish to hear a somewhat different balance, and the console permits this flexibility. For testing purposes, the engineer may briefly listen to the output of just one microphone, without affecting the recorded program. During playback the balance of the entire program may be adjusted without affecting the recording setup.

Multitrack mixdown session. When all the recording work is completed, the engineer plays back the multitrack tape, mixing and balancing the many separate tracks to produce the desired stereophonic program (**Fig. 2**). During this mixdown session the outputs of the multitrack tape recorder are routed through the console, just as the microphones were during the original recording session. In effect, the multitrack tape now takes the place of the studio musicians, and the various tracks—each with its distinctive musical information—are mixed to create the final master tape. This master tape is used as the first step in the production of phonograph records.

During the mixdown session a wide variety of processing devices may be used to modify the signals that were recorded earlier. An equalizer changes the apparent frequency response of a musical instrument, compressors and limiters restrict the dynamic range, and expanders attenuate or remove low-level signals.

The natural ambience of a large room may be simulated with delay lines and reverberation chambers. One or two electronically delayed echo signals may be mixed with the direct program, while the

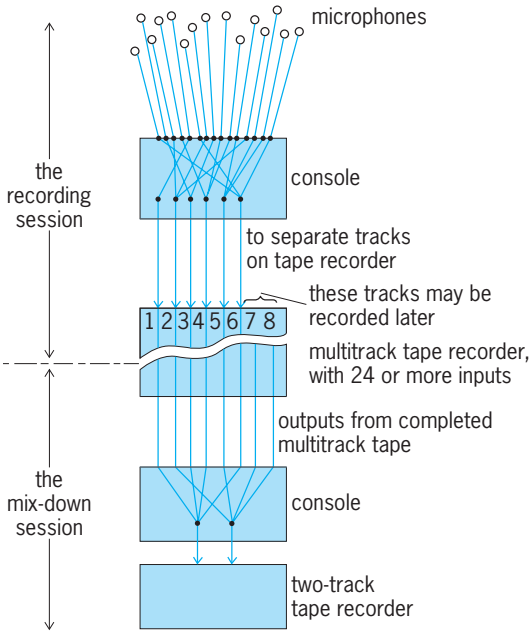

Fig. 2. Multitrack sound recording and mixdown session.

reverberation chamber simulates the multiplicity of reflections typical of a large concert hall. Signals that are delayed just slightly, perhaps 25 milliseconds or less, may be mixed with the direct sound to create the illusion of a larger group. The delay is not heard as such, but simulates the slight imprecision of attack that is characteristic of, say, a large group of violins. Excessive use of any of these signal-processing devices may create special effects that the musicians might find impossible to produce on their instruments.

In the case of a monophonic or stereophonic sound recording, the engineer is unable to isolate and modify, or remove entirely, one or more parts of the ensemble. However, the recording may convey an excellent sense of depth, with the complex relationships between instruments well preserved.

At the cost of losing some of the ensemble's spatial relationships within the room, the multitrack sound-recording technique offers the recording engineer an unprecedented degree of technical control over the entire recording and record production process. *See* SOUND-REPRODUCING SYSTEMS. John M. Woram

Bibliography. B. Bartlett, *Practical Recording Techniques*, 2d ed., 1998; J. Borwick, *Sound Recording Practice*, 4th ed., 1996; J. Eargle, *Handbook of Recording Engineering*, 2d ed., 1992; D. M. Huber, *Modern Recording Techniques*, 5th ed., 2001; J. Woram, *Sound Recording Handbook*, 1989; A. J. Zaza, *Mechanics of Sound Recording*, 1991.

Sound-recording studio

An enclosure treated acoustically to provide controlled recording of music or speech. Professional sound-recording studios range from small home studios (such as a converted garage or basement) to large facilities using state-of-the-art equipment. They consist of two rooms, the main studio or stage (performance area) and the audio control room. The stage includes one or more microphones, is acoustically isolated to ensure that external sounds do not interfere with the recording, and can be broken up into small booths to separate individual performers. The audio control room contains the mixing console and recording devices.

Motion picture sound stages. These large enclosures, with a volume of 1,000,000 ft^3 (28,000 m^3), are used for photography (to allow long shots), video recording, and sound recording. As well as keeping out extraneous noises, the interior walls and roof are treated with sound absorption material covering up to 85% of the total interior surface to achieve a short reverberation time (RT). This assures intelligible speech recording, as well as the realistic simulation of outdoor scenes. The sound mixing person (mixer) is on the floor of the stage. *See* CINEMATOGRAPHY; OPTICAL RECORDING; REVERBERATION.

Television stages. These are smaller than motion picture sound stages and require separate audio and video control rooms. Emphasis is on dialog; some television stages allow audience participation, such as panel shows and short comedies where audience clapping is an essential part of the program. The microphone operators are on the stage with the camera operators. Audio mixing of a video program takes place in the audio control room. Still another enclosure is required for the lighting console. Stages are often built in pairs joined by a passageway, allowing construction on one stage while recordings take place in the other. Sound insulation is provided between these stages so that hammering in one room will not disturb recordings on the other. Sound and video recording and lighting features are carefully rehearsed. Light fixture dimming is controlled electrically, often with high electrical current pulses that necessitate that microphone cables be specially shielded.

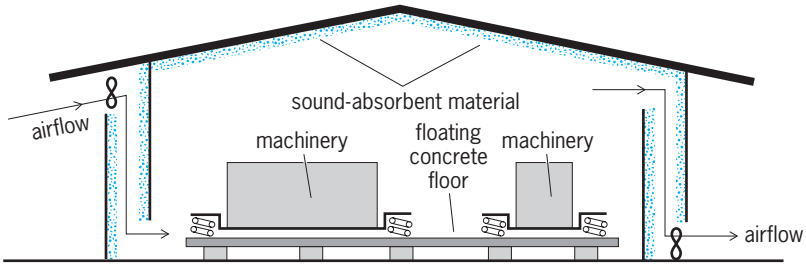

Fig. 1. Concrete housing for air-conditioning equipment to attenuate both airborne and solid-borne sound transmission to a nearby studio.

HVAC systems. The presence of several hundred people in a television stage requires a high-volume heating, ventilation, and air-conditioning (HVAC) system of low noise level, which is achieved by large ducts purveying air with a velocity less than 500 ft (150 m) per minute. HVAC condensers, chillers, and cooling towers are remote from the television stage in an isolated enclosure at ground level or underground (**Fig. 1**). Mechanical equipment is supported on vibration isolators. Enclosure walls are lined with sound absorption material. Heat removal ventilation may include noise reduction openings. *See* AIR COOLING.

Studio floors. Motion picture and television stages have concrete floors for durability, load-bearing capacity, and absence of squeaks. They are isolated from the walls to minimize the transmission of exterior mechanical vibrations. Floors may include drain pipes for rain scenes on the stage, and pits may be built into the floor to allow video recording of simulated below-ground scenes without building up the set.

Figure 2 shows a typical foundation detail of two adjoining television stages separated by a distance of at least 2 ft (0.6 m). Sand and decomposed granite form a cushion for the concrete floors, which are structurally isolated at the perimeters by a 1-in.-thick (2.5-cm) elastic seal. The walls are lined with 2-in.-thick (5-cm) fiberglass sound-absorbent treatment, covered with fireproof facing. The interspace between the stages may be covered with sheet lead. *See* TELEVISION STUDIO.

Music-recording studios. Music-recording studios vary in size up to 100,000 ft^3 (2800 m^3), the largest

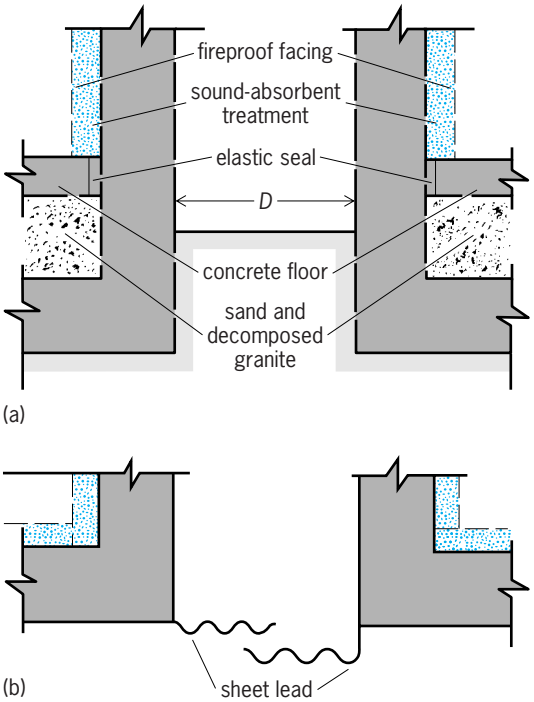

Fig. 2. Adjoining television stages. (*a*) Vertical cross section. (*b*) Top view.

being for the recording of symphonies and opera music by large recording companies. However, the majority of music-recording studios are owned and operated by smaller independent producers, or are rented studios or empty theaters. The reverberation time will vary. The recording microphones may be placed near the artists for individual clarity, or hung at strategic locations just above and in front of a stage.

Multitrack recording. Multitrack magnetic recording appeared first as only 3-track audio recordings, then 4-track, then 8-track. Now some of the larger recording companies employ 32 and 64 tracks. Each track contains the sound from only one or two musicians. The conductor is replaced by the recording engineer, who mixes the many music parts in the audio control room. Concomitant with this type of recording is the installation of drum cages, piano chambers, vocal booths, trumpet stalls, and various other cubicles to achieve clean individual section tracks. A number of adjacent enclosures have windows in the partitions for performers to see each other to keep in time. *See* MAGNETIC RECORDING.

Figure 3 illustrates an audio/video recording studio that allows the mixer a direct view of each instrument section. The video mixers have only a partial direct view. They watch their program material over a closed-circuit television circuit anyway for best videography.

Special acoustical devices. Special acoustical features, such as sound traps and bass traps, may be installed in small recording studios to achieve distinction in recording music for records, audio cassettes, video cassettes, and compact disks (CDs).

Built in a pit between adjoining instrument sections, a sound trap (**Fig. 4**) replaces a barrier that would not allow neighboring performers to see each other. The trap is generally filled with 2-in.-thick (5-cm) fiberglass panels spaced 1 in. (2.5 cm) apart. When a performer sits over a sound trap 3 to 4 ft (1 to 1.3 m) square, much of the sound from the instrument is absorbed into the trap due to the low sound pressure over the trap compared to

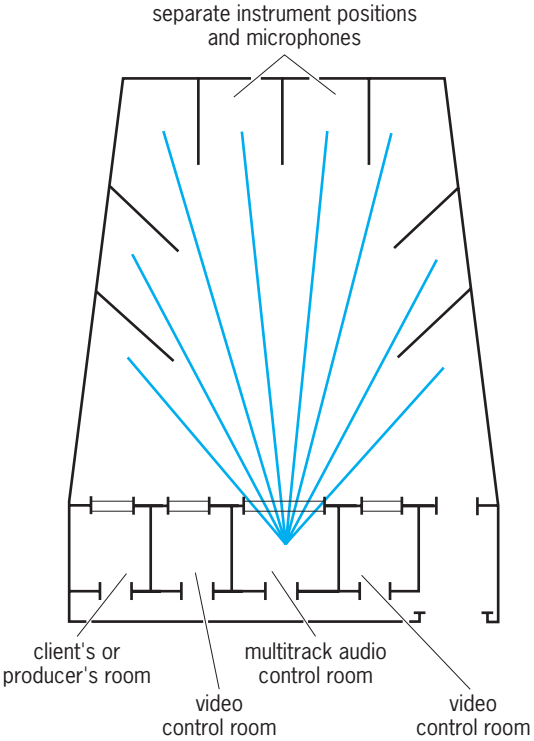

Fig. 3. Plan of audio/video recording studio for multitrack recording.

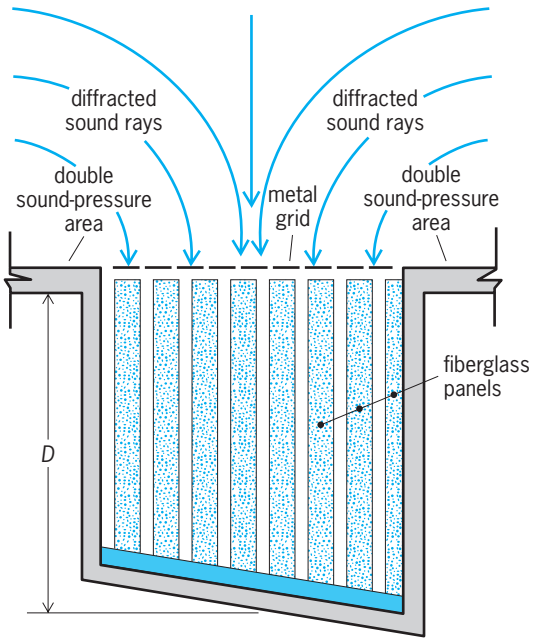

Fig. 4. Sound trap used to minimize reflections from overhead instrument to adjacent instrument microphones. Diffracted sound rays are absorbed into the pit where sound pressure is low.

the adjoining floor. Trap depth *D* should be a quarter wavelength of the lowest frequency to be absorbed. The trap is covered by a metal grid to support the performer. *See* SOUND.

An entire studio can be built from such sound traps and still have sound reflective walls, floor, and ceiling. Both the traps and the reflective surfaces are practically frequency-independent in their absorption and reflection characteristics so that such enclosures do not exhibit the coloration (spectral amplitude variations) from sound reflected from room boundaries.

A bass trap absorbs sound below about 100 Hz. This may be a sound trap deep enough to absorb low-frequency sound, a resonant absorber (Helmholtz chamber), or a thin panel that resonates on the layer of air trapped between it and a rigid wall near which it is placed. A common drywall on studs will act as a bass trap in the range 80–160 Hz.

Sidewall sections. Another device used in music recording is acoustically variable sidewall sections. These allow parts of a wall in a rectangular room to be turned to be either sound absorbent or reflective without producing discrete or structured echoes (**Fig. 5**).

Control room mixing and audition. A control room is used to monitor a musical program performed in a studio. A modern control room also serves as a venue for the mixer to later adjust a multitrack recording for volume levels, frequency response (equalization), and reverberatory notes, and to expand or compress certain musical passages. The mixer does this by listening to the program reproduced by one or more loudspeakers in the control room and therefore assumes the role of the conductor. The three optimal acoustic treatment theories for such a control room are live-end-dead-end (LEDE), EDEL (LEDE spelled backwards), and MOD (moderate).

For LEDE, a dead or absorbent treatment is placed on surfaces near the loudspeaker. The opposite end of the room is made reflective or diffusive (to scatter sound) [**Fig. 6***a*]. This arrangement is intended to prevent coloration of the reproduced program by first-order reflections from the sidewalls and ceiling about the loudspeaker. It is, of course, impossible to prevent first-order low-frequency reflections from the window below the loudspeaker, but the originators of the system claim that these reflections are not important for the purpose of mixing sound tracks.

For EDEL, the acoustic treatment has the same allotment of hard and soft surfaces, but with the absorbent end behind the mixer (Fig. 6*b*). Reflective or diffusive surfaces are placed near the loudspeaker and above the mixing console in the control room, while the rear wall is made absorptive.

For MOD, the acoustic treatment comprises a uniform disposition of the sound-absorbent material all about the control room.

There is little consensus whether LEDE, EDEL, or MOD treatment is best for the mixer. This is understandable, since the reverberation time in such a room is generally very short, approximately 0.1 (log *V*), where *V* is the volume of the control room in cubic feet (1 m^3 = 35 ft^3). This reverberation time may be only 0.4 second for a large control room of 10,000 ft^3 (280 m^3). None of these treatments will improve the intelligibility of speech, singing, or music. Rather, each treatment supports subjective preferences that may achieve the individual esthetic goal of the mixer.

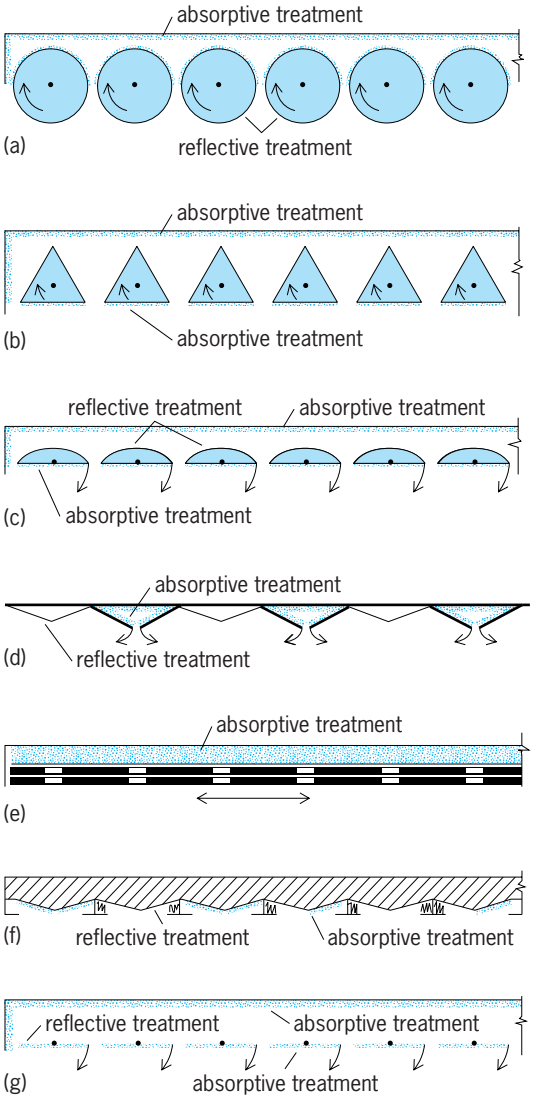

Fig. 5. Variable sidewall treatments. (*a*) Rotatable cylinders. (*b*) Rotatable triangular prisms. (*c*) Rotatable cylindrical sections. (*d*) Folding splays. (*e*) One stationary and one sliding perforated panel. (*f*) Retractable curtains. (*g*) Rotatable flat panels.

ADR room. The automatic dialog replacement (ADR) or postsynchronizing room (**Fig. 7**), traditionally of the motion picture industry, also has wide use in television production. The performer watches himself or herself on the screen and then repeats the original speech with lip synchronism. ADR employs forward-and-backward transportation of a film or magnetic tape sound track so that short scene bites can be repeatedly reproduced visually and aurally for an actor who did not intelligibly record his or her speech during original filming or video recording. This may have been because the performer had

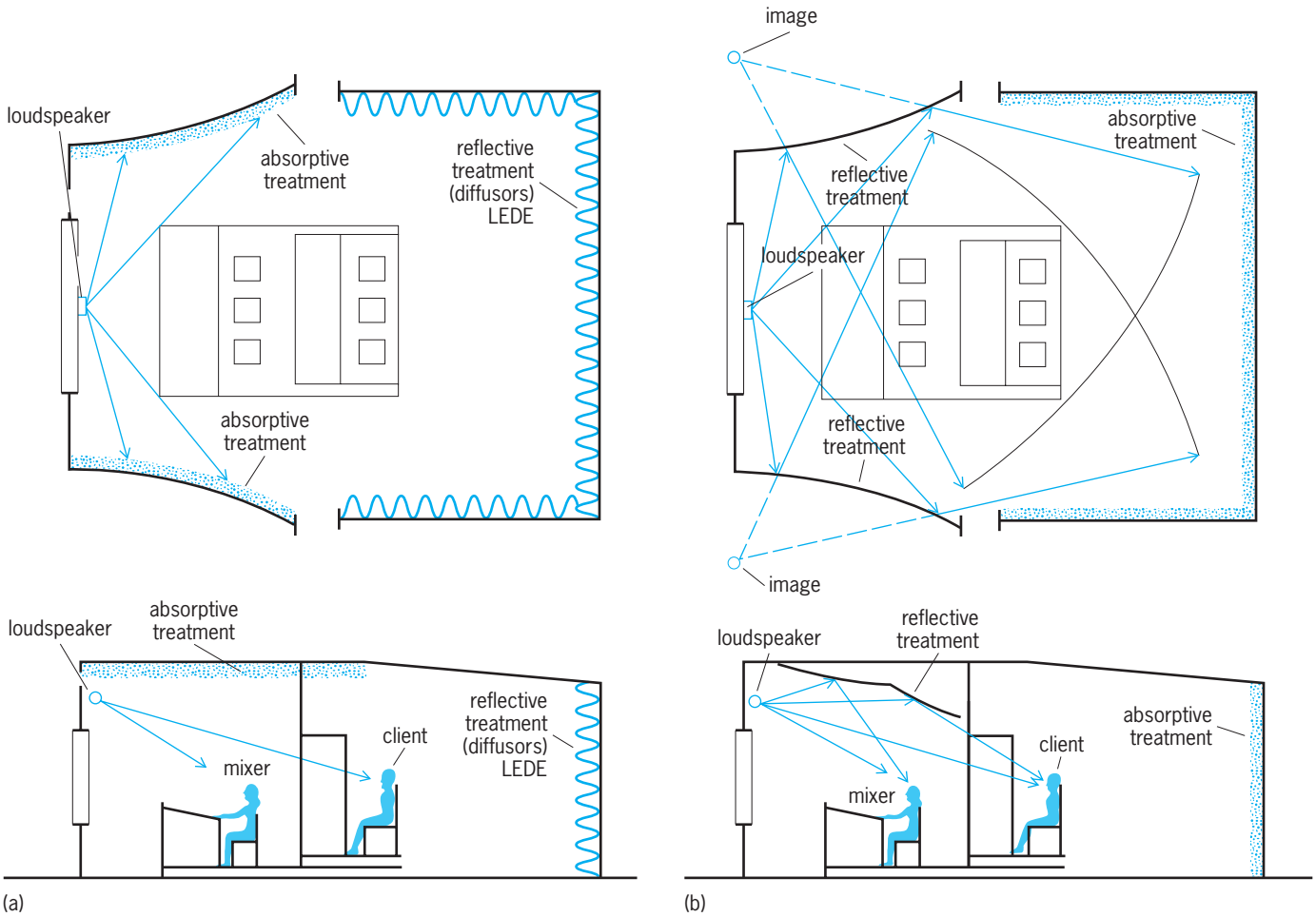

Fig. 6. Control rooms. Top and side views are shown for each type of room. (*a*) LEDE room with live rear wall and sound-absorbent environment at the frontal sidewalls to prevent lateral reflections from mixing with the direct signals from the loudspeaker. (*b*) EDEL room with absorbent rear wall and live front end as in conventional listening locales.

laryngitis or because unavoidable noise was present, such as seaside surf or nearby vehicular or aircraft traffic.

To achieve realism in the recording, it is necessary to credibly recreate the ambiance (acoustical liveness) and the soundscape (original background sound) where the original dialog took place. This can be done by placing the actor in an acoustically

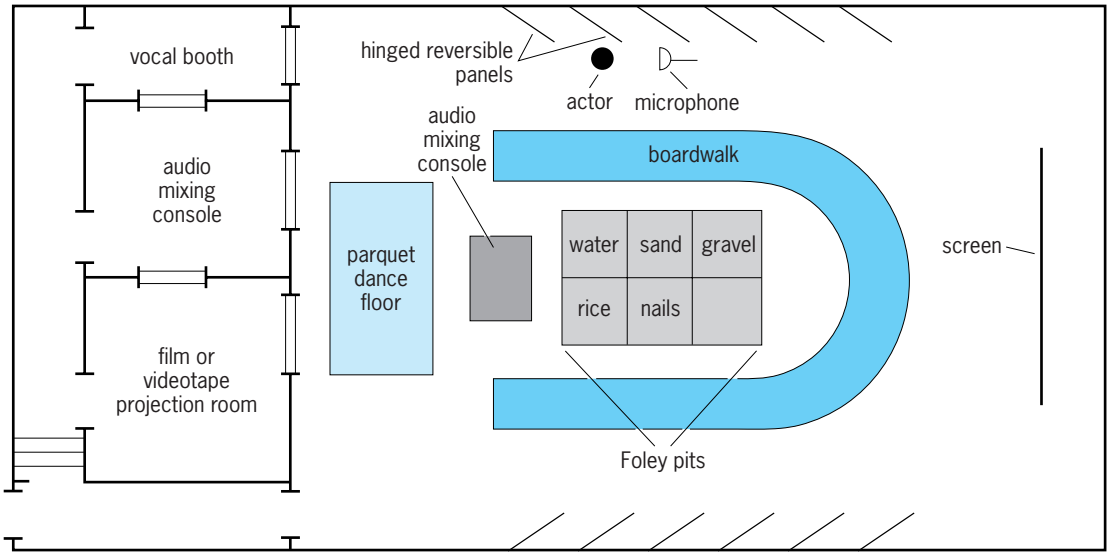

Fig. 7. ADR (automatic dialog replacement) or postsynchronizing studio with Foley pits containing various materials for generating sound effects.

adjustable environment within the ADR room. Such "variable acoustics" is achieved on one or more sidewalls of the ADR facility with hinged reversible panels or rotatable cylindrical columns that are absorbent on one side and reflective on the other (Figs. 5 and 7).

It is frequently necessary to add synthetic sound effects to a scene. When played back, the original recording of a slamming door or a pistol shot may not sound natural. ADR rooms are often equipped with Foley pits (Fig. 7) containing materials to imitate sound effects that could not be recorded well during filming or video recording, or devices that provide realistic sound signals for the events pictured on the screen. Such devices may include a boardwalk to allow the recording of footstep sounds on wood when the original scene was performed on a stage with a concrete floor. A parquet dance floor is used to duplicate footsteps shown on the screen, especially in tap dancing, where the dancer listens to the music over headphones, so that his or her recorded footstep signals are not marred by extraneous music. Soundscapes such as birds or surf may be added from a library of environmental sounds. Either mixing console shown in Fig. 7 may be used, depending on the mixer's preference or requirements for recording the sound of a scene.

Optimal studio room shape. Sound recording studios are often highly sound absorbent ("dead"), in which case room shape is unimportant. However, for studios that cannot be made dead, or where some reverberance is desirable, room shape is critical. Small rooms of regular shape (cubes and squares) have redundant resonances (degenerate modes) that will color sound reverberating in that room. **Table 1** shows some acoustically recommended dimension ratios to minimize the number of degenerate modes. Such rectangular rooms require treatment for flutter echoes. A recording studio having the harmonic ratio of 1:2:3 will provide the greatest floor area compared to other ratios of rooms of equal volume, accommodating the most performers in a given space. A room with a ratio of 2:3:5 has the highest ceiling compared to the other cited ratios, providing a maximum number of sidewall reflections (diffusion) and hence good stereophonic sound reception, as may be desired for a concert hall. Adherence to any of these ratios is more important for reverberant rooms than it is for studios having the prescribed short reverberation time of 0.1 (log V).

Reverberation time. This is predicted by the Sabine formula $RT = 0.049\ V/A$, where V is the room volume in cubic feet and A is the sum total of all the acoustic absorption in the room, measured in sabins. (In metric units, $RT = 0.161\ V^1/A^1$, where V^1 is the room volume in cubic meters and A^1 is the total absorption in metric sabins.) In large studios and stages, the reverberation time may be greater than predicted. The reported absorption coefficient, determined from tests on small samples, can be greater than unity, but when the same material is placed over the majority of walls and ceiling areas, the resulting reverberation time will be longer, indicating that the material absorption coefficient is then as low as 0.80.

Sound and television stages must have a very short reverberation time to simulate an outdoor acoustical atmosphere. Music recording studios, ADR chambers, and control rooms may have a more live interior. The recording engineer may wish to change the reverberation time. Performers often prefer a live environment for orchestral or ecclesiastical music. Jazz and rock-and-roll performers prefer a shorter reverberation time. Figure 5 shows ways to change the sound reflectivity of the walls in a music recording studio or an ADR room, thus changing the sound liveness near such walls. These devices minimally affect the reverberation time found in the middle of the room unless they are installed on all surfaces of that room. To change the reverberation time in a large studio, it is necessary to move either one of its walls or its ceiling. This is rarely done because of the expense. *See* ARCHITECTURAL ACOUSTICS.

Windows. Recording studio windows for control rooms and public viewing must present high sound insulation. They may consist of two thick glass panes widely spaced and sealed to the adjoining wall surface. To avoid first-order sound reflections, they may be triple-pane units, with the center pane made of a 3/4-in.-thick (19-mm) to 7/8-in.-thick (22-mm) noise barrier glass. The outside panes are 1/4-in.-thick (6-mm) Lucite or Plexiglas sheets cold-bent through an arc of 60° with a radius equal to the window width.

Noise control. Recording studios' background noise level must be very low. This is particularly important for digital recording, because of its very wide dynamic range of 80–90 dB. Heating, ventilating, and air-conditioning equipment must not be installed on the roof of such studios. Airflow control or augmenting devices (VAV boxes) must not be located at, over,

TABLE 1. Rooms with recommended ratios of dimensions

		Rooms normalized for equal volume		Rooms normalized for equal height	
Ratio of room dimensions	Name of ratio	Normalized ratio	Relative floor area	Normalized ratio	Relative floor area
1:2:3	Harmonic	1:2:3	6	1:2:3	6
1:6:3:4	Knudsen	1.08:2.03:2.71	5.5	1:1.88:2.5	4.7
3:5:8	European	1.1:1.84:2.95	5.43	1:1.67:2.67	4.45
1:1.6:2.5	Volkmann	1.14:1.82:2.85	5.19	1:1.6:2.5	4
$(5^{1/2} - 1):2:(5^{1/2} + 1)$	Golden section	1.12:1.82:2.94	5.35	1:1.62:2.62	4.24
2:3:5	Sabine	1.17:1.75:2.92	5.11	1:1.5:2.5	3.75

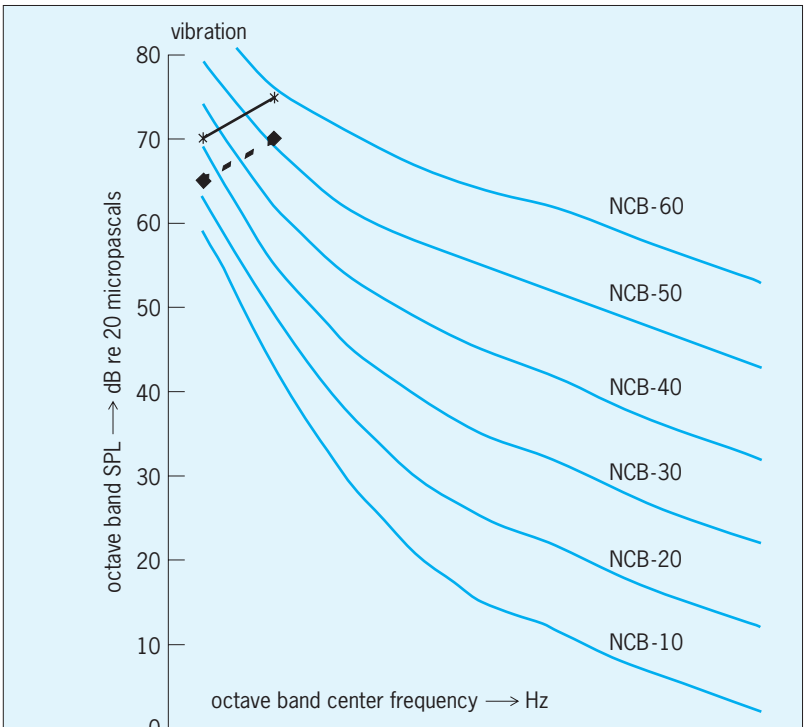

	32	63	125	250	500	1000	2000	4000	8000
NCB-10	59	43	30	21	15	12	8	5	2
NCB-20	63	49	38	30	25	22	18	15	12
NCB-30	69	55	46	40	35	32	28	25	22
NCB-40	74	62	54	49	45	42	38	35	32
NCB-50	79	69	62	58	55	52	49	46	43
NCB-60	85	76	71	67	64	62	59	56	53
Vibration	65	70							
Strong vibration	70	75							

Fig. 8. Room background noise balanced criterion curves.

or near the studio space. Large HVAC mechanical equipment (boilers, chiller pumps, condensers, and cooling towers) should be located on the ground at a distance from the studio (Fig. 1). It may be necessary to employ a sound attenuator to the outdoor ventilation intake and exhaust openings in that enclosure.

Noise criterion (NC) curves. L. Beranek's balanced NC curves (**Fig. 8**) provide a convenient measure of the annoyance of background noise in buildings. Matching the noise level found in a room to any of these curves is done as a tangent criterion, where the noise background level in every frequency band must not exceed the indicated curve. A dBA level computed from the sum of all bands is often numerically 5 decibels or more greater than the corresponding NCB value. **Table 2** shows the recommended noise criterion balanced (NCB) value for occupied spaces.

Use of the limiting-curve criterion alone can lead to dubious results for low-frequency mechanical equipment noise ("rumble"). Low-frequency sound (and vibration) above 60 dB can cause audible rattles in ceiling and studio fixtures. The noise from an air-conditioning installation could exceed an NCB-30 specification, or the alternative RC-30 specification, by 2 or 3 dB in the octave band whose midfrequency is 1000 Hz and thus be deemed unacceptable, while the spectrum of another installation may hug the NCB-30 at all points on the curve and be declared acceptable. Vibration and discomfort can result from low-frequency sound. Thus, in Fig. 8, additional limits (- - -, moderate; — —, severe) are placed on the lowest-frequency noises. To minimize these problems, the alternative RC curve set with suppressed lower and higher frequency levels is used as an HVAC design tool by mechanical engineers.

Background limit. This may be NCB-15 (about 20 dBA) for high-quality sound recording studios or NCB-25 (30 dBA) for television studios (Fig. 8). For an ordinary studio it should not exceed NCB-20 (25 dBA). When many performers are in the room or photography must take place there, the limit should be NCB-25 (30 dBA). *See* NOISE MEASUREMENT.

Exterior noise. The outdoor noise at a prospective site is surveyed over at least one 24-hour period and over several days if possible. **Figure 9** shows the 24-hour noise history of a prospective studio site. Hours of maximum noise exposure, especially at low frequencies (for example, from diesel locomotive and aircraft passages), are evident. The spectrum of these maximum noise events must be known for airborne sound insulation design of studio exterior walls, roof, doors, and windows.

Wall noise isolation design. Concrete walls are good noise barriers. They may be solid formboard-cast, ground-cast and later raised with a crane (tilt-up walls), or concrete block (CMU). One side of a CMU is plastered to seal joint leakage paths for sound. Solid

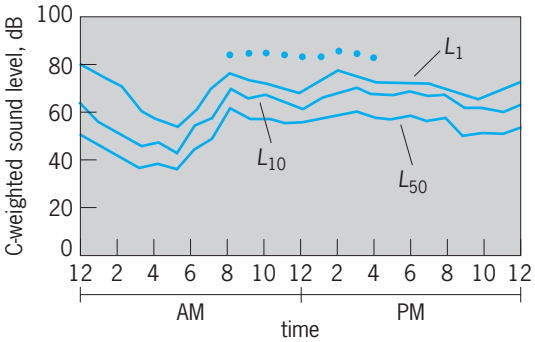

Fig. 9. Results of a 24-hour noise exposure level survey at a planned studio site. L_1, L_{10}, and L_{50} represent the C-weighted noise level that was exceeded 1, 10, and 50% of the observation period. Dots are aircraft flyover noise maximum values.

TABLE 2. Recommended background noise criteria for various rooms

Enclosure	Noise criterion balanced (NCB) value	A-weighted sound level, dB
Concert halls	15–20	25–30
Recording studios	15–20	25–30
Theaters	20–25	30–35
Homes	25–30	35–40
Offices	30–35	40–45
Stores	35–40	45–50

concrete walls approximately obey the acoustical mass law for the transmission line (TL) in decibels for a sound of frequency, f,

$$\begin{aligned} \mathrm{TL} &= 14.5 \log M + 20 \log f - 31 + 20 \log \cos\theta \\ &= 14.5 \log M^1 + 20 \log f - 41 + 20 \log \cos\theta \end{aligned}$$

where M is the surface density in lb/ft^2 (M^1 is the surface density in kg/m^2), θ is the angle of sound incidence. For instance, at 500 Hz ($\log f = 2.7$) and 0° angle of sound incidence ($\log \cos\theta = 0$), a 6-in.-thick (15-cm) solid concrete wall weighing 70 lb/ft^2 ($\log M = 1.85$) or 340 kg/m^2 ($\log M^1 = 2.53$) exhibits a transmission loss of 50 dB.

The enclosing floor, walls, and ceiling of studios located in buildings housing other activities require similar analysis and treatment. Ceiling noise isolation is improved by suspending a grid on vibration isolators that will carry several layers of drywall. The air gap so formed is filled with one or more layers of sound-absorbing insulation. *See* ACOUSTIC NOISE.

Ground vibrations. Studios should be isolated from ground vibrations from trains, subways, and industrial plants. Studios located on upper floor levels are subject to vibrations from mechanical equipment anywhere in the building such as HVAC, transformers, and pumps. Ground vibration and its spectral content is measured with an accelerometer that is either buried in the ground or attached to the affected floor, preferably over a 24-hour period. A studio concrete floor is isolated from these vibrations by "floating" it on pads 2–3 in. thick, on coil springs of appropriate deflection or on air springs over a similar air gap. The resonant frequency of this support and gap must be a quarter or less that of the lowest audio frequency expected to be reproduced, for example, a 10-Hz resonance for a 40-Hz lowest reproduced audio frequency. *See* ACCELEROMETER; MECHANICAL VIBRATION; SOUND RECORDING; VIBRATION ISOLATION. Michael Rettinger; Angelo Campanella

Bibliography. G. G. M. Alkin, *Sound Recording and Reproduction*, 1997; L. Beranek, Balanced noise criterion (NCB) curves, *J. Acous. Soc. Amer.*, 86(2):650–664, August 1989; L. L. Beranek and I. L. Ver (eds.), *Noise and Vibration Control Engineering: Principles and Applications*, 1992; D. Gibson and G. Petersen (eds.), *The Art of Mixing: A Visual Guide to Recording, Engineering and Production* (*Mix Magazine Pro Audio Series*), 1997; P. Newell, *Studio Monitoring Design: A Personal View*, 1995; D. N. Nordantonio, *Sound Studio Production Techniques*, 1990; A. Nisbett, *The Sound Studio*, 6th ed., 1995; P. White, *Creative Recording II: Microphones, Acoustics, Soundproofing, and Monitoring*, vol. 2, 1999.

Sound-reinforcement system

A system that amplifies or enhances the sound of a speaking person, singer, or musical instrument. Sound-reinforcement systems are used to increase the clarity of the original sound as well as to add loudness and reverberation. Those that improve loudness and clarity are often called public address systems, and those that improve reverberation characteristics are called reverberation enhancement systems. Systems for increased loudness are used when listeners are located far from the sound source, as in a large audiorium or outdoors. Systems for increased clarity are used to overcome the effects of unwanted noise or excessive reverberation in a room (such as concert halls and churches with reverberation characteristics designed for music). Reverberation enhancement systems are used to add sound field components, such as reverberation to rooms having unsuitable levels of direct, early reflected sound and reverberant sound or having short reverberation times. *See* ARCHITECTURAL ACOUSTICS; REVERBERATION; SOUND; SOUND FIELD ENHANCEMENT.

Basic elements. The basic elements of a sound-reinforcement system are (1) one or more microphones, (2) amplification and electronic signal processing systems, and (3) one or more loudspeakers. The signal processing equipment may include level controls, equalizers or tone controls, systems for prevention of howl and ringing, delays and reverberators, as well as frequency dividing and crossover networks.

Figure 1 illustrates the signal flow diagram of a simple system. In the simplest systems, all elements are joined as one physical unit. An example of a simple public address system is the hand-held battery-driven electronic megaphone, consisting of a microphone, a frequency-response-shaping filter, an amplifier, and a horn loudspeaker (**Fig. 2**). Typical use of the megaphone by police and others is to address a crowd of people. In order to obtain high sound levels in combination with high intelligibility, the frequncy response of a megaphone is limited to save battery power. This may result in poor sound quality.

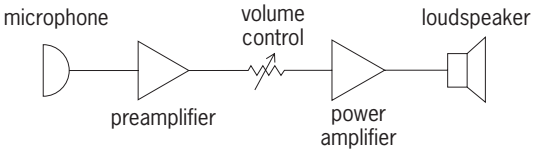

Fig. 1. Basic sound-reinforcement system.

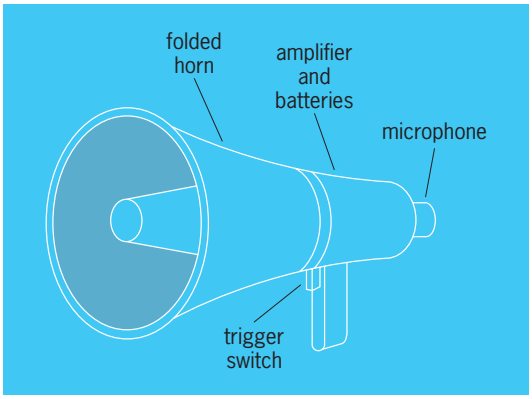

Fig. 2. Electronic megaphone, with all the basic elements of a sound system in one package.

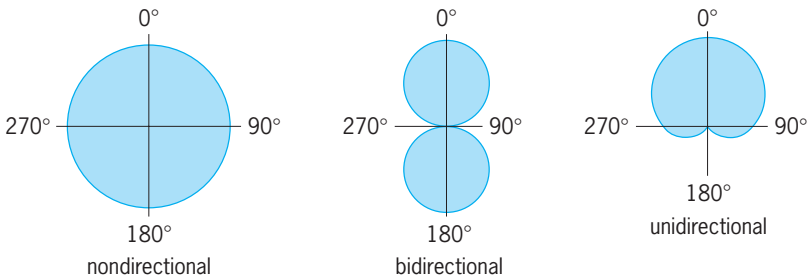

Fig. 3. Polar directivity patterns of microphones.

The "amplified speaker pulpit" is a stationary and only slightly more complex system. It is a sound-reinforcement system built into a pulpit. Because the system is stationary (and most system parameters are known to the designer), it can be tailored to give a very good and natural sound quality. Since the loudspeaker is close to the talker, there is no shift in apparent source location.

Most sound-reinforcement systems, however, are assembled from individual elements to optimize and adapt the system performance for the needs of a particular installation.

Microphones. Microphones often use dynamic, condenser, or electret types of transducers, and may have various polar directivity patterns such as omnidirectional, unidirectional, or bidirectional (**Fig. 3**). In some microphones, it is possible to switch the pattern by remote control. Microphones for use on stage floors, boardroom tables, and so on are boundary microphones with special frequency response characteristics, and their use can reduce or eliminate phase cancellation effects from reflections. *See* ELECTRET TRANSDUCER; TRANSDUCER.

Condenser microphones require a preamplifier close to the microphone element. They may be permanently polarized (electret) or require a polarizing voltage. In either case, dc power must be sent to the preamplifier, usually positive on the signal line pair and negative on the shield. *See* ELECTRET; PREAMPLIFIER.

Digital microphones use an analog transducer, but integrate analog-digital converters and other components to offer improved possibilities for signal identification, cabling, and routing, making them less susceptible to noise and hum. *See* ANALOG-TO-DIGITAL CONVERTER.

Radio links may substitute for wire to connect system components. The most familiar radio link is wireless microphone. Wearing such a microphone on the body permits the user to move around freely. The wireless microphone usually consists of a miniature omnidirectional microphone (either electret condenser or dynamic type) worn close to the neck on the outside of clothing; a compact FM transmitter, worn inside clothing; an antenna, usually a wire worn along a leg; and a receiver located in a fixed position convenient both for connecting to the rest of the system and for obtaining unobstructed radio pickup from the transmitter (**Fig. 4**). Miniature wireless microphones may be worn on the head and hidden by hair. *See* FREQUENCY-MODULATION RADIO; MICROPHONE.

Signal control equipment. Control and signal processing equipment may range from a single volume control in a preamplifier to complex control boards similar to those used in recording studios, with possibly hundreds of input channels, complex equalizers for every input channel, feedback control equipment, reverberation generation equipment and delay units, matrixed outputs, and so forth. Modern large and complex systems feature preset automatic digital control, with the audio signals determining real-time settings. *See* SOUND-RECORDING STUDIO.

Delay units use analog-digital conversion, digital delay, and digital-analog conversion to delay an input signal. Reverberation may be generated by using finite or infinite impulse response filters, corresponding to the addition of many delayed sound components. Finite impulse response filters use an assembly of delays, and can be based on natural (measured) or calculated room impulse responses. In infinite impulse response filters, some of the delay output signals are fed back to the delay inputs and added in special patterns to reduce coloration in the generated reverberation. Power amplifiers may range in output from 10 watts to thousands of watts, and a system may have only one power amplifier or hundreds of them. *See* ACOUSTICS; AMPLIFIER; POWER AMPLIFIER.

Because sound-reinforcement systems add energy to the sound field in a room, there is a possibility that some frequencies will be particularly emphasized due to in-phase feedback of sound to the microphone(s). At a critical gain setting, the system gain will be sufficient to let the system start to oscillate (feedback). The gain-before-feedback is a measure of the loudness the system can produce, with reference to the loudness of the live sound source, before the system starts to oscillate. At lower gain settings, close to oscillation, the system will sound "colored" because sound at some frequencies is exaggerated. The feedback also creates reverberation, that is, prolongs the decay time of the signal. These effects are subjectively annoying. Feedback control equipment may be based on techniques such as (1) adaptive

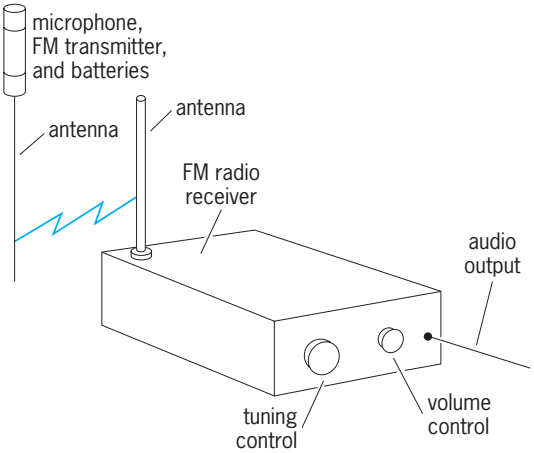

Fig. 4. Elements of a wireless microphone.

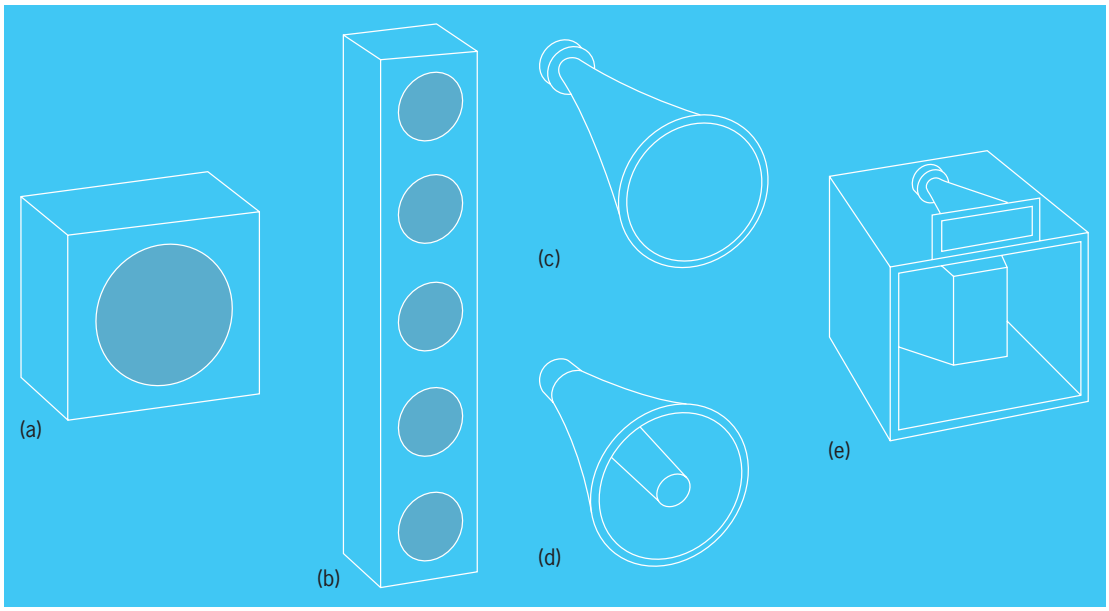

Fig. 5. Loudspeakers used in sound-reinforcement systems. (*a*) Direct-radiator loudspeaker in a cabinet. (*b*) Multiple direct radiators in a vertical line array or column. (*c*) Straight-axis horn loudspeaker. (*d*) Folded or reentrant horn loudspeaker. (*e*) Two-way horn loudspeaker system with a single high-frequency straight directional horn and a separate low-frequency horn enclosure with one or two bass loudspeakers.

notch filters, that is, very sharp (narrow-band) rejection filters, (2) frequency shifting, or (3) time-varying delay or amplification. All these techniques may introduce audible artifacts into the sound if not properly used, but may add up to approximately 6 dB of gain in a system. *See* GAIN; LOUDNESS.

Loudspeakers. Loudspeakers may be direct radiators or horn-loaded drivers. The direct radiators are usually cone-type loudspeakers, but are sometimes bending-wave-transducer, electrostatic, or other unconventional types. Horn loudspeakers use the reentrant horn for simple low-quality paging; diffraction horns for somewhat better-quality paging; front-loaded directional horns of various types, including the constant directivity horns (which attempt to provide the same coverage characteristics over a wide frequency range and which have largely displaced the older multicellular and radial horns in high-quality applications); and horn bass enclosures, which often employ folds or bends for compactness, or else supplement the horn loading with ports to bring up bass response (**Fig. 5**).

A particular loudspeaker system may consist of a single-radiator loudspeaker, a series of direct-radiator loudspeakers in either horizontal or radial lines (vertical line-source loudspeaker systems are often called column loudspeakers), combinations of direct radiators for bass radiation and horn loudspeakers for treble radiation, or large enclosures with large low-frequency loudspeakers, used with high-frequency horns (Fig. 5). Some systems used for high-level music reinforcement divide the frequency range into four parts, with subwoofers for the lowest bass frequencies, a second set of bass loudspeakers for midbass frequencies, midrange constant-directivity horns, and supertweeters for the extreme high-frequency end of the audio spectrum. Each of these loudspeakers may be driven by separate amplifiers, which are fed by a four-way crossover network (frequency dividing network). The use of separate amplifiers for different portions of the sound spectrum permits the lowest distortion and the greatest flexibility for adjustment. *See* LOUDSPEAKER; SOUND-REPRODUCING SYSTEMS.

Loudspeaker placement and selection. Fixed sound systems are of two basic types: central and distributed. The central system employs loudspeakers reasonably close to the source of live sound (but not so close that feedback is an unsolvable problem), while the distributed system has many loudspeakers fairly close to the audience. For central systems to work properly, the performing area must be defined, as in a typical auditorium or concert hall. Spaces where the sound source may be located anywhere, such as in hotel ballrooms where the head table might be along any wall, are usually best served by some type of distributed system.

Central systems. The most common type of central system employs a loudspeaker or loudspeaker system (sometimes called a cluster) directly over the front of the stage, with the directional characteristics of the loudspeaker or system chosen to cover the audience with good uniformity without directing too much energy toward the stage. Such a system is usually the first choice for most sound system designers because it can be installed most simply and usually provides the greatest intelligibility and naturalness (**Fig. 6**).

Variations on the central system include the use of vertical column loudspeakers on each side of the stage, the use of a horizontal line source along the top of the stage, and the use of directional horn loudspeakers on each side of the stage. All these solutions have their own peculiar problems. Any of the

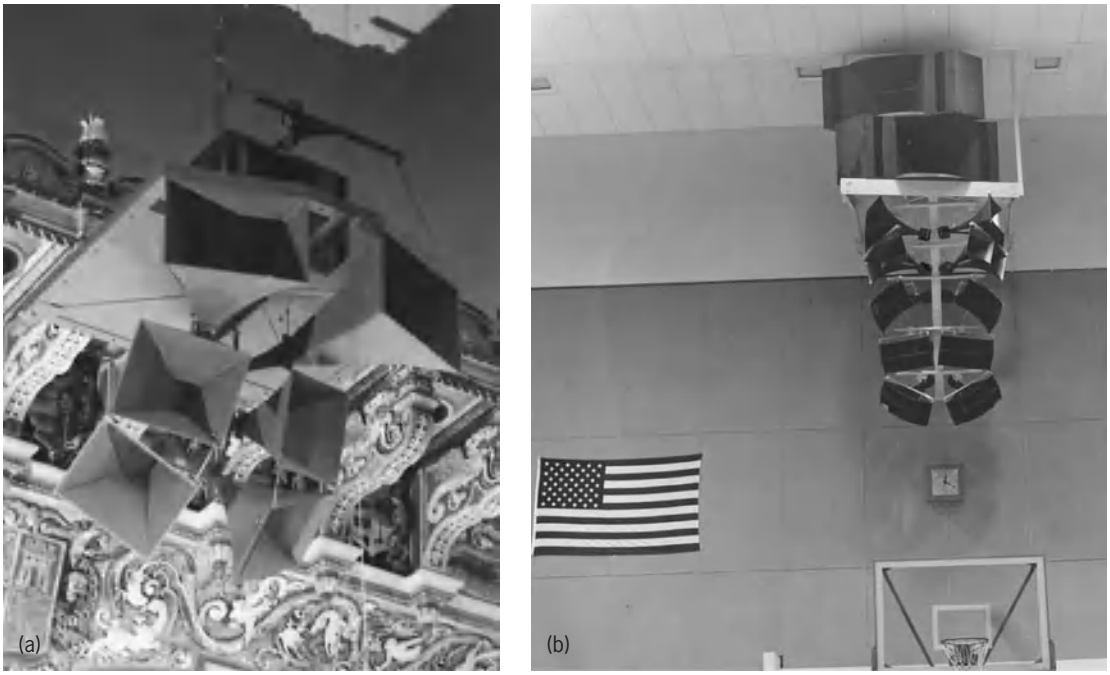

Fig. 6. Central cluster loudspeaker systems. (*a*) System in a theater. The constant-coverage high-frequency horns are below the supporting frame, and a two-loudspeaker straight-horn low-frequency enclosure is above (*courtesy of Amir Pishdad*). (*b*) System for a gymnasium. Ten multicellular high-frequency horns are below two double-loudspeaker straight-horn bass enclosures.

split-source methods (with loudspeakers on each side of the stage) can result in alternate zones of cancellation and reinforcement near the centerline of an auditorium. Modern column loudspeakers may offer steerable array technology, using digital signal processing so that the radiation pattern at various frequencies can be tailored to give optimal coverage over a wide area.

Distributed systems. The most common type of distributed system employs loudspeakers mounted in the ceiling, usually directed straight down. Less frequently used variations include loudspeakers in seat backs (the "pew-back" approach, most often used in churches) or in building columns. Horizontal rows of closely spaced loudspeakers have been used from one side of a theater to the other above the proscenium, as well as down the length of the apex of the ceilings of Mormon churches for reasonably uniform coverage in spaces without excessive reverberation. Loudspeakers in side walls are best used for adding reverberance in a room, and for surround sound and specific effects in cinema and theater. Low ceilings and chandelier systems require wide-dispersion loudspeakers, usually about 100-mm-diameter (4-in.) or larger coaxial (small high-frequency cone and larger bass cone) loudspeakers; however, above 10 m (33 ft) height directional horns are often required.

Reduction of reverberation. With proper placement and selection of loudspeakers, sound systems in live rooms can actually reduce the apparent reverberation by concentrating the amplified sound energy into the sound-absorbing audience or seats, with minimum energy directed to wall and ceiling surfaces. The constant-directivity horns have particular applications in central systems used for this purpose. Such systems can provide high speech intelligibility, with naturalness, in highly reverberant spaces that have poor speech acoustics.

Enhancement of reverberation. In a system consisting of many sound reinforcement channels, one can adjust the gain so that each channel has low gain and coloration. The combined effect of the system will yield a prolonged reverberation time over the entire operating frequency range of the channels. Systems with up to a hundred channels are sometimes used for reverberation enhancement purposes in rooms perceived to be acoustically "dead". The subjective perception of reverberation may also be enhanced by adding reverberation into the signals using a reverberator in the signal-processing system. Unless this reverberation is radiated incoherently by the use of many loudspeakers in different directions from the listener, the reverberation will be insufficiently diffuse and inadequate. Sometimes reverberation is added selectively to a singer's voice to enhance the quality.

Signal processing. In addition to proper loudspeaker placement and selection, the design of a sound-reinforcement system with the desired characteristics requires the use of signal-processing techniques such as signal delay and equalization.

Signal delay. Digital signal delay equipment is now often used with distributed systems. This equipment places the electronic audio signal in a memory device, producing it in time for the appropriate loudspeakers to be synchronized with the live sound from the performance area.

Haas or precedence effect. Use of this effect can contribute greatly to naturalness. If the live signal from

a talker or singer is heard first and the amplified signal arrives up to 25–30 milliseconds later and is not more than 10 dB louder, the listener will determine the direction of the composite signal as coming from the talker or singer. That is, the listener will not be aware that the amplified signal is coming from a different direction. This effect can be used with central systems by proper location of the loudspeakers. In some central systems and most distributed systems, use of the effect requires both proper loudspeaker placement and proper selection of delays.

Equalization. The equalization process can also contribute to the naturalness of a sound system and to its gain-before-feedback in many situations. Before equalization, almost every sound system exhibits some frequency response irregularities that are a departure from an ideal frequency response curve. In the past, such irregularities (due to deficiencies in either the room or the loudspeakers) were accepted. Now, broadband or narrowband equalization usually is applied to a sound system to ensure both the smoothest possible response and adequate gain-before-feedback. Various types of equalizers and frequency response analysis equipment are used in the equalization process. The most widely used appear to be octave-band or third-octave-band equalizers that reduce (or sometimes boost) the audio electrical signal in adjacent bands, and audio analyzers that display the relative acoustical signal level in the same bands via a visual readout. Such equipment is often used with a "pink" noise generator, giving a random electrical signal that has the same signal level in each band, during the analysis and adjustment process. In some cases, the equalization is supplemented by the use of (adaptive) notch filters to prevent oscillation. Equipment in digital mode can continually analyze signals to provide real-time adaptive equalization. *See* EQUALIZER.

Time-energy-frequency analyzers. Both the setting of correct signal delays and proper equalization can benefit from the use of analyzers, which display the energy of a signal as a function of time and frequency. Simple audio analyzers present the total sound energy arriving at the microphone with the sound system excited by a continuous random-noise generator. Modern analyzers, however, excite the system with a particular pulse, a pseudo-random noise, or a pure-tone signal swept in frequency. The analysis equipment uses information about the signal emitted to calculate the response as a function of time and frequency. This technique allows the time alignment of two separate signals to set delays and locate loudspeakers, and allows independent evaluation of the frequency response of the direct signal (the signal from the loudspeaker without reflections) and the various components of room sound (the reflected energy). There are situations where equalization of the direct sound is preferred to equalization of the total sound field. Even with modern equipment, the ear is the final judge of any equalization scheme.

Power requirements. The size of the power amplifiers needed for an indoor sound system depends on the acoustics of the room served, the sound pressure levels needed (much less for a speech-reinforcement system than for a music system), and the efficiency of the loudspeakers used. Direct-radiator loudspeakers may be only 1% efficient in converting electrical power to acoustical power, while certain types of horn loudspeakers may be as high as 25% efficient. Generally, low-frequency loudspeakers are less efficient than high-frequency components in a multiway system; thus, if electronic crossovers are used, larger amplifiers are needed for low frequencies than for high frequencies. A small church might use only 30–50 W of power amplification for speech reinforcement, while a small night club with less efficient loudspeakers and higher sound levels might require several thousand watts. Temporary outdoor systems set up for rock concerts may run into the hundreds of thousands of watts, as extremely high levels are produced over long distances for hundreds of thousands of listeners. Electronic reverberation and surround systems, particularly for organ and for orchestra, require tremendous amounts of bass power and large areas of bass-loudspeaker radiating surface to sound convincing. This also holds true for all outdoor music reinforcement. *See* SOUND PRESSURE.

Outdoor systems. Outdoor systems (**Fig. 7**) are similar to indoor systems, without the reflections that

Fig. 7. Cluster of high-frequency horns at the base of the loudspeaker tower at a stadium, ensuring even coverage and the projection of sufficient sound power to produce 100-dB levels 700 ft (210 m) away at home plate. The mouths of the horns are protected by mesh.

Fig. 8. Large control console used for both reinforcement and recording at a theater. The window enables the operator to hear the sound in the theater. (*Courtesy of Amir Pishdad*)

result from interior room surfaces. (The reflections in indoor systems produce reverberation, which can be a negative or positive effect, depending on the nature of the program. Reflections also can result in greater uniformity of sound pressure level throughout the audience area to be covered.) Outdoor systems may have to cope with such ambient noise problems as aircraft flyovers, nearby rail lines or highways, and factory noise. This limits the dynamic range, or range of loudness, from outdoor systems such that the softest passages of a musical work must be brought up in loudness to override intruding noise, while the loudest passages cannot be so loud as to bother some listeners.

Care must be taken in outdoor facilities to prevent the sound system from becoming a nuisance to the community. For example, outdoor rock concerts should not take place in the midst of populated areas. The use of directional loudspeakers (such as constant-directivity high-frequency horns), the proper placement and orientation of these loudspeakers, and the use of architectural barriers (such as walls and berms) can minimize the annoyance created by the use of an outdoor sound system.

System development. Development of a successful sound-reinforcement system begins with analysis of the uses of the system: just who or what requires reinforcement, and what is the spatial relationship of the live sound source with respect to the audience. Next comes the acoustical, architectural, and electronic design, including choice of a central or distributed system, locations for loudspeakers, their orientation, choice of microphones, preparation of a signal-flow chart which is developed into a line diagram for the electronic portion of the system, and proper selection of a sound control position if one is necessary.

The design of a system may be greatly simplified through the use of acoustic simulation by computer software—based on the image source or ray tracing methods—to optimize the design with respect to loudspeaker directional characteristics, room geometry, and acoustic conditions. Using auralization (audible sound field simulation based on such simulation), it is possible to listen binaurally to a simulation of the sound-reinforcement system's performance in advance of its being installed. *See* BINAURAL SOUND SYSTEM; VIRTUAL ACOUSTICS.

Often the sound-reinforcement installation must be put to competitive bidding, and this process requires detailed specifications indicating the overall performance expected from the system and from critical individual components, the system's functional requirements, wiring practices, details of installation, documentation, and so forth. This leads to bid review by an impartial expert, who may have designed the system. Careful installation is necessary, followed by adjustment and equalization, in turn followed by acceptance tests by the owner's representatives. Instruction of operators takes place at this time and is critical to the performance of the system. Good service can keep a high-quality system in operation for periods of 20 years or more. Usually service is best provided by the installing contractor.

System control. The sound system operator should be located where it is possible to hear the same mix of live and amplified sound as heard by the audience. Backstage is a poor location for the controls of a theater sound-reinforcement system. Giving the operator a monitor loudspeaker in the control booth, or even a "dummy head" (mannequin with two "ear" microphones) feeding headphones, can never be as good as a large open window in a correctly located control booth (**Fig. 8**). The use of binaural technology may help in controlling the sound in a large venue by the use of many dummy heads. Many performing arts centers have cockpits in the middle of the audience area, placing the sound system operator directly in the audience. Compromise positions at the rear or at the side are more frequent and can also work. A frequent control position in churches is at a balcony rail.

Several types of automatic mixers have been developed. These devices recognize which microphone is being addressed, increase the amplification from that microphone while suppressing the signal from others, and reduce the gain appropriately when two microphones are being addressed to keep the system from feeding back. The best of these devices also provide overall gain control, reducing the amplification from the loudest voices while bringing up amplification from the softest. For a simple church speech-reinforcement system, one of these devices can replace a human operator with good results. As sophistication of the equipment and techniques of use have improved, automatic mixers in theaters assist, but do not replace, the human operator.

David Lloyd Klepper; Mendel Kleiner

Bibliography. D. Davis and C. Davis, *Sound System Engineering*, 2d ed., 1997; G. Davis and R. Jones, *The Sound Reinforcement Handbook*, 2d ed., 1990; J. Eargle, *JBL Sound System Design Reference Manual*, 1999; D. L. Klepper, *Sound Reinforcement:*

An Anthology, vol. 1, 1978, vol. 2, 1996; H. F. Olson, *Acoustical Engineering*, 1957; S. H. Stark, *Live Sound Reinforcement: A Comprehensive Guide to P.A. and Music Reinforcement Systems and Technology*, 2d ed., 1996.

Sound-reproducing systems

Systems that attempt to reconstruct some or all of the audible dimensions of an acoustic event that occurred elsewhere. A sound-reproducing system includes the functions of capturing sounds with microphones, manipulating those sounds using elaborate electronic mixing consoles and signal processors, and then storing the sounds for reproduction at later times and different places.

The design of audio products has been progressively more scientific, thanks to a better understanding of psychoacoustics, the relationship between what we measure and what we hear. Such investigations can be productive only if it is possible to elicit consistent opinions from individual listeners and agreement among numbers of different listeners. Although some with deteriorated hearing certainly do "hear differently," the majority of listeners, given some practice and the opportunity to form opinions about sound in circumstances free from biases (such as seeing the products, loudness differences, and variable program sources), are remarkably consistent judges of sound quality. *See* PSYCHOACOUSTICS.

When opinions obtained in a controlled manner are studied, it is found that they correlate strongly with certain measurable quantities. Of all audio components, electronic devices are the closest to technical perfection, exhibiting the smallest audible differences. Transducers present many more problems. Microphones, which convert sound into an electrical signal, exhibit fewer problems than loudspeakers, which do the reverse. This is because microphones are relatively simple devices, and they are small compared to most sound wavelengths. Loudspeakers must be physically larger in order to generate the required sound levels. Because the sound interacts with the acoustical space (a room or automobile interior, for example), as it propagates to the ears, reflected sounds modify and sometimes dominate the audible impressions. The acoustical space is the final component in a sound-reproduction system. *See* ARCHITECTURAL ACOUSTICS; LOUDSPEAKER; MICROPHONE; TRANSDUCER.

The original event for most recordings occurs in the control room of a recording studio or the dubbing stage of a film studio, where the artistic decisions are made about what is to be stored on disk or tape. These decisions are made by listening through loudspeakers in rooms. The loudspeakers and room used in capturing the sound influence what is recorded. If these are significantly different from those used during the reproduction, the sound will not be heard as it was created. As there are no universal standards governing the design and use of loudspeakers and rooms, there are substantial variations in the quality of recordings, as well as how they sound when reproduced through different loudspeakers in different rooms. This is a problem for a sound-reproducing system (**Fig. 1**). *See* SOUND-RECORDING STUDIO.

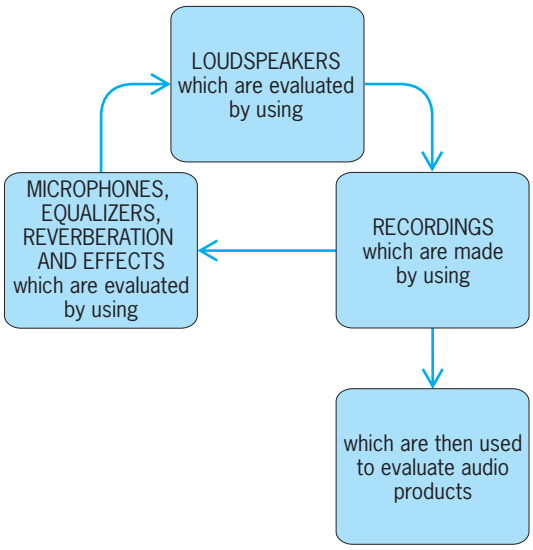

Fig. 1. Loudspeakers influence the sound of recordings while they are being made, and then those recordings are used to evaluate other audio products, including loudspeakers. Subjective evaluation, used alone, is a flawed process.

Technical variables. Sound-reproduction systems deal not only with the sounds of live performers but also with synthesized sounds. The same system can reasonably be expected to do justice to the nuances of chamber music, the dynamics and drama of a symphony, and World War II gunfire reenacted on film. The following technical variables are relevant to the reliable reconstruction of audio events.

Frequency range. Traditionally the audible frequency range has been considered to be 20 Hz to 20 kHz. However, lower frequencies can generate interesting tactile impressions, and it is now argued by some that higher frequencies (even 50 kHz and beyond) add other perceptual nuances. Whether these frequencies are significant, or even exist, is a matter of current debate. Most scientific data indicate that only exceptional persons, in the prime of youth, respond to sounds beyond about 20 kHz. *See* HEARING (HUMAN).

Dynamic range. In humans, this is the range of sound level from the smallest audible sound to the largest sound that can be tolerated. In devices, it is the range from background noise to unacceptable distortion. The contours of equal loudness show that from the threshold of audibility to the onset of discomfort is about 120 decibels at middle and high frequencies. In reality, much less than this is required to give a satisfactory impression of accurate reproduction. When recording a live performance, microphones are not a limiting factor since the best have dynamic ranges in excess of 130 dB. In the recording studio, dynamics are manipulated by electronic gain compensation, so the ultimate limitation is the storage medium. For example, a 16-bit compact disk

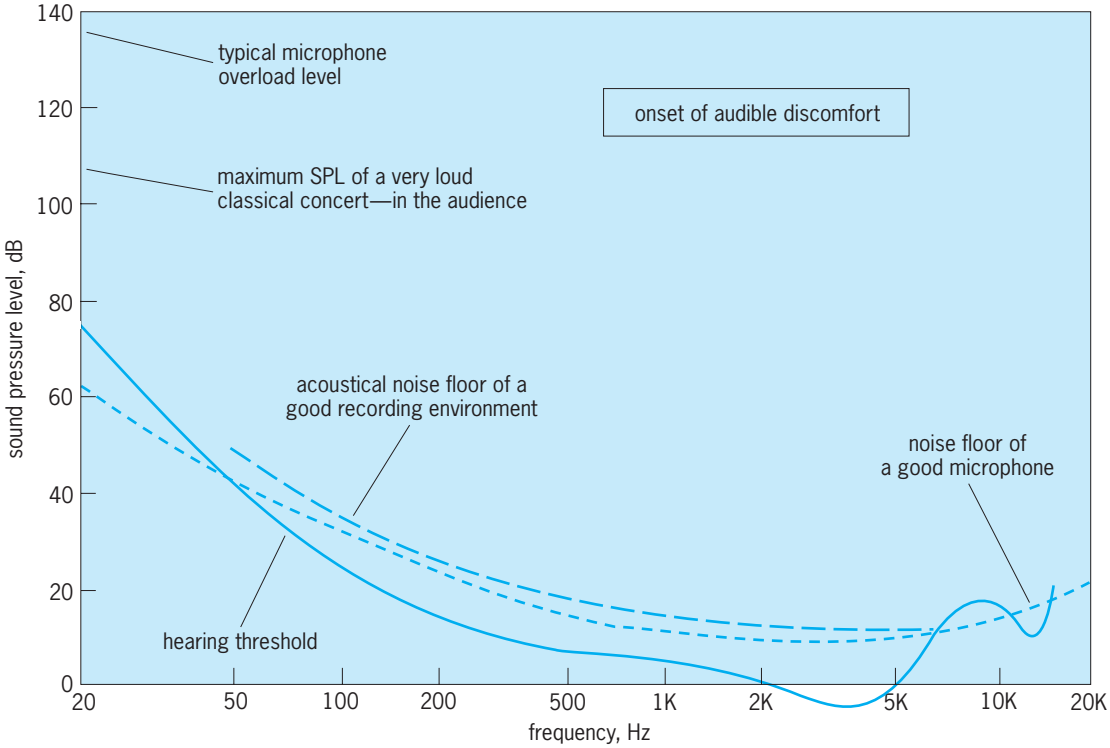

Fig. 2. Some important reference sound levels relating to what is recorded and what is heard.

(CD) can exceed 100 dB signal-to-noise ratio, while 24-bit digital media can exceed any reasonable need for dynamic range. Background acoustical noise in studios and concert halls sets the lower limit for recordings, just as background noise in homes and cars sets a lower limit for playback. Quiet concert halls and homes are around 25 dB sound pressure level (SPL) at middle and high frequencies. Crescendoes in music and movies approximate 105 dB, giving a working dynamic range of about 80 dB for fussy listeners and less for casual listeners. Today, this is no challenge for electronics, and it is not even difficult for loudspeakers in homes and cars, except perhaps at very low frequencies where a lot of air must be moved to fill a large room (**Fig. 2**). *See* ACOUSTIC NOISE; COMPACT DISK; DECIBEL; DISTORTION (ELECTRONIC CIRCUITS); GAIN; LOUDNESS; SIGNAL-TO-NOISE RATIO; SOUND; SOUND PRESSURE.

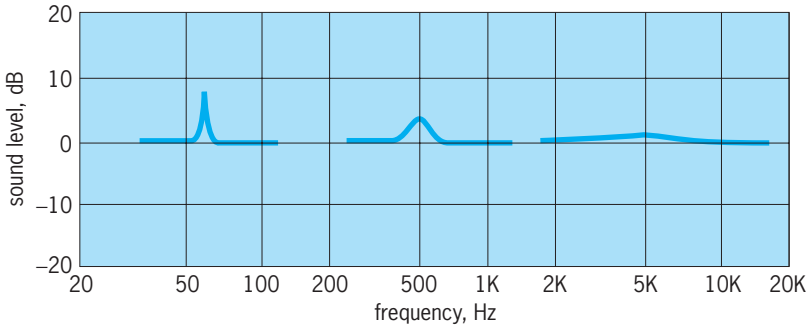

Fig. 3. Just-audible deviations in a measured frequency response when the program material is a complex orchestration in a reverberant space. From left to right are shown deviations caused by a high-*Q*, a medium-*Q*, and a low-*Q* resonance. The magnitudes of the deviations change little as a function of where they are located on the frequency scale.

Linear distortions—amplitude/frequency response. Variations in amplitude and phase as functions of frequency are known as linear distortions, since (ideally) they do not vary with signal level. In practice they often do, and in the process generate nonlinear distortions. In terms of their importance in preserving the timbre of voices and musical instruments, the amplitude versus frequency characteristic, known commonly as the frequency response, is the dominant factor. Humans are very sensitive to variations in frequency response, the amount depending on the bandwidth of the variation.

It is convenient to discuss this in terms of the quality factor (Q) of resonances, since most such variations are the result of resonances in loudspeaker transducers or enclosures. High-Q resonances (such as $Q = 50$) have a narrow bandwidth; they are very frequency specific. Such resonances are just audible when their deviation from an otherwise flat frequency response is about 10 dB for a sound that is not very revealing (nonreverberant, multimiked pop or jazz), to something less for complex orchestrations in reverberant spaces, to about 3 dB for pink noise. As the Q drops (wider bandwidth), so do the amplitudes at which the resonances are audible. A broadband Q of 1 is just audible at an amplitude deviation of about 0.3 dB with pink noise (**Fig. 3**). So, when we speak of frequency responses having deviations of +/− so many dB, over a frequency range, such a statement is meaningless unless there is a curve to show the bandwidth of the variations; a narrow spike can be much larger than a broad bump, for the same audible effect. Critical listeners report that they like the sound of devices with flat, smooth, broadband

frequency responses. *See* DISTORTION (ELECTRONIC CIRCUITS); Q (ELECTRICITY); RESONANCE (ACOUSTICS AND MECHANICS); RESPONSE; REVERBERATION.

Linear distortions—phase/frequency response. A pattern of phase shift accompanies the amplitude variation caused by a resonance. We can assume that this visual clue indicates something potentially audible. However, there is no clear evidence that phase shift, by itself, is an audible problem. Since both amplitude and phase responses must be linear to allow accurate waveform reproduction (Fourier transform), the conclusion is that the reproduction of accurate waveforms is not a requirement for good sound. The reality of everyday listening is that we hear a combination of direct sound and multiple reflections from room boundaries and furniture. Moving either the head or the sound source can cause huge changes in the waveform at the entrance to the ear, yet the perception of timbre is hardly, if at all, changed. *See* FOURIER SERIES AND TRANSFORMS; PHASE (PERIODIC PHENOMENA).

Nonlinear distortions. These distortions occur when a device behaves differently at different signal levels. The waveform coming out of a distorting device will be different from the one entering it. The difference in shape, translated into the frequency domain, is revealed as new spectral components. If the input waveform is a single frequency (pure tone), the additional spectral components will be seen to have a harmonic relationship to the input signal. Hence they are called harmonic distortion. If two or more tones are applied to the device, nonlinearities will create harmonics of all of the tones (harmonic distortion) and, in addition, more new spectral components that are sum-and-difference multiples of combinations of the tones. These additional components are called intermodulation distortion.

It should be remembered that it is the same nonlinearity that causes all of the distortion, and that the measurement methods described here are simply different ways to quantify the nonlinearity. In reality, we listen only rarely to anything resembling a pure tone, and only occasionally to multiples of pure tones. We listen to much more complicated sounds and, as a result, the distortions that result from their interaction with the nonlinearity are of many forms, continuously changing with time. It is therefore impossible to predict, from our simple distortion measures, just how offensive, or even audible, the unwanted spectral contamination will be.

Multiway loudspeaker systems complicate intermodulation distortion measurements because that form of distortion occurs only when the multiple test signals are reproduced by the same transducer. Thus it is possible for a nonlinear system to yield similar harmonic distortion data but very different intermodulation distortion data, depending on how many crossovers there are in the system and the frequencies at which they occur. For example, a woofer that becomes mildly nonlinear at low frequencies may be acceptable if it is restricted to frequencies below 80 Hz, but annoying if it operates to 1000 Hz where important voice and music frequencies can be modulated by simultaneous strong bass signals.

Perceptually, listeners are aided by a phenomenon called masking, in which loud sounds prevent some less loud sounds from being heard. This means that the music causing the distortion inhibits the ability to hear it. Rough guidelines suggest that, in music, much of the time we can be unaware of distortions measuring in whole percentages, but that occasionally small fractions of a percent can be heard. The only safe guideline seems to be to keep distortion, however it is measured, as low as possible. *See* MASKING OF SOUND.

Dispersion, directivity, sound power—of loudspeakers. The preceding technical issues apply to both electronic devices and loudspeakers. However, loudspeakers radiate sound in all directions, so that measurements made at a single point represent only a tiny fraction of the total sound output. In rooms, most of the sound we hear from loudspeakers reaches our ears after reflections from room boundaries and furnishings, meaning that our perceptions may be more influenced by measures of reflected and total sound than by a single measurement, say, on the principal axis.

At low frequencies, common box loudspeakers radiate omnidirectionally. At middle and high frequencies, loudspeakers can be designed to radiate in a predominantly forward direction, bidirectionally (forward and back) with the transducers in phase (the so-called bipole) or out-of-phase (the dipole), omnidirectionally, or any of several other multidirectional possibilities. Many of these designs evolved during the era of two-channel stereo, as means of enhancing the reflected sound field in rooms, compensating for the lack of true surround information. However, the most popular configuration has been loudspeakers with transducers arranged on the front surface, facing the listeners.

To be useful, technical measurements must allow us to anticipate how these loudspeakers might sound in rooms. Consequently, it is necessary to measure sounds radiated in many different directions at points distributed on the surface of an imaginary sphere surrounding the loudspeaker. From these data, it is possible to calculate the direct sound from the loudspeaker to the listener, estimates of strong early reflected sounds from room boundaries, and estimates of later reflected or reverberant sounds. It is also possible to calculate measures of total sound output, regardless to the direction of radiation (sound power) and of the uniformity of directivity as a function of frequency. All of these measured quantities are needed in order to fully evaluate the potential for good sound in rooms. *See* DIRECTIVITY.

Power compression of loudspeakers. As more power is delivered to a transducer, the voice coil heats up, the resistance increases, and the sound output is lower than it should be. This is called power compression. In systems with woofers, midranges, and tweeters, the different units may heat by different amounts, leading to audibly different spectral balance at different playback levels. Another power-related performance change has to do with the vent or port in

bass-reflex woofer systems. At high sound levels, the high air velocity in the port can generate turbulence, causing noise and changing the carefully designed tuning of the system. Thus loudspeaker systems can have quite distinctive sounds at different playback levels. *See* AERODYNAMIC SOUND.

System configurations. In the earliest systems, there was monophonic sound (monaural means one ear, not one channel). In the 1930s, Bell Telephone Laboratories published research results indicating that, to successfully preserve the "auditory perspective" of a frontal sound stage, there were two practical techniques: binaural (meaning two ears) and multichannel.

Binaural. This refers to a sound-reproduction technique in which a dummy head, equipped with microphones in the ear locations, captures the original performance. Listeners then audition the reproduced performance through headphones, with the left and right ears hearing, ideally, what the dummy head "heard." The system is good, but flawed in that sounds that should be out in front (usually the most important sounds) tend to be perceived to be inside, or close to, the head. Some laboratory exercises claim to have solved the problem, but common experience is different. Addressing this limitation, systems have been developed that use two loudspeakers, combined with signal processing to cancel the acoustical crosstalk from the right loudspeaker to the left ear, and vice versa. The geometry of the loudspeakers and listener is fixed. In this mode of listening, sounds that should be behind are sometimes displaced forward, but the front hemisphere can be very satisfactorily reproduced. *See* BINAURAL SOUND SYSTEM; EARPHONES; VIRTUAL ACOUSTICS.

Multichannel audio. This began with stereophonic systems using two channels, not because two was an optimum number, but because in the 1950s the technology was limited to one modulation for each groove wall of a record. Bell Labs, many years before, had recommended many channels and, as a practical minimum, three (left, center, and right) or, for a single listener, two. Good stereo imaging is possible only for listeners on the axis of symmetry, equidistant from both loudspeakers. At its best, it has been very enjoyable. Stereo auditioned through headphones is much less realistic. *See* MODULATION; STEREOPHONIC SOUND.

Quadraphonic systems. These appeared in the 1970s, and added two more loudspeakers behind the listeners, mirroring the ones in front. The most common systems exhibited generous interchannel crosstalk or leakage. To hear the proper spatial perspective, the preferred listening location was restricted to front/back as well as left/right symmetry. The lack of a center channel was a serious omission. The failure to agree on a single standard resulted in the system's demise.

Dolby Surround. For film sound applications, Dolby modified the matrix technology underlying one of the quadraphonic systems, rearranging the four matrixed channels into a left, center, right, and surround configuration. In cinemas, the single, limited-bandwidth, surround channel information is sent to several loudspeakers arranged along the side walls and across the back of the audience area. For home reproduction, the surround channel is split between two loudspeakers placed above and to the sides of the listeners. To enhance the spatial effect of the single surround channel, some systems introduced electronic decorrelation between the two loudspeakers, some used multidirectional loudspeakers in the surround channels, and some did both.

Digital discrete multichannel systems. Even with the best active (electronically steered) matrix systems, the channels are not truly discrete. Sounds leak into unintended channels where they dilute and distort directional and spatial impressions. Several competing digital recording systems now provide five discrete full-bandwidth channels, plus a low-frequency special effects channel for movies with truly big bass sounds. For homes, 5.1 channels refers to five satellite loudspeakers, operating with a common low-frequency subwoofer. The subwoofer channel is driven through a bass-management system that combines the low frequencies of all channels, usually crossing over at 80–100 Hz so that, with reasonable care in placement, it cannot be localized.

Movies provided the motivation for multichannel audio in homes, but music may be where it offers its greatest opportunities. Movies tend to limit the use of off-screen sounds mainly to vague spatial effects, ambient music, reverberation, and carefully rationed directional effects. Music, on the other hand, is not constrained. Traditional acoustical music takes advantage of the greater spatial realism that is possible with more channels. Modern music uses the additional channels to surround listeners with moving and static sounds, instruments, and voices. Listeners may find these new directional and spatial effects variously stimulating and annoying. At least now, with more channels and more loudspeakers, what we hear at home and in our cars is likely to more closely resemble what was created in the studio. *See* MUSICAL ACOUSTICS.

The 5.1 channels do not satisfy all listeners. The configuration of three across the front and two at the side/rear locations originated with films. It assumed that all listeners would face a screen. Music listeners are less disciplined in where they sit and in which direction they face. More channels improve the situation and open more creative opportunities (**Fig. 4**).

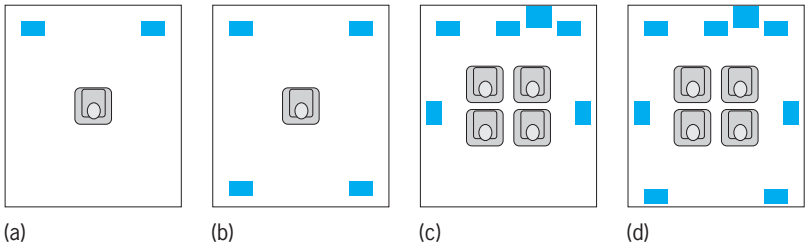

Fig. 4. Evolution of multichannel audio. (*a*) Two-channel stereophonic reproduction. (*b*) Quadraphonics. (*c*) 5.1-channel "home theater." The larger loudspeaker symbol is a subwoofer. (*d*) 7.1-channel multichannel movie and music entertainment system.

System components—electronics. The number of signal sources is large and growing. The control device of any multichannel system must accept analog and digital inputs from several stereo and multichannel storage devices and from real-time delivery systems, including terrestrial broadcasting, cable, satellite, and the Internet. Mass storage options now include "juke boxes" holding hundreds of CD's and DVD's, as well as computerlike hard drive storage for compressed audio and video. To this complexity must be added the variety of encoding/decoding options, including different lossy and lossless digital compression algorithms for multichannel audio and digital video. Confusion is common. These are all digital devices and, with enough intelligent programming, a comprehensible user interface should make it all workable. Today's audio/video receivers are much more complicated than the stereo devices of just a few years ago, yet they are easy to use. These, and especially the stand-alone surround processors, are really very powerful computers. Sound quality in electronic devices is currently excellent and is getting even better. *See* VIDEO DISK.

System components—loudspeakers. The one audio component that remains stubbornly evolutionary is the loudspeaker. In basic form, cone and dome transducers of today look and work remarkably the same as those of decades ago. Other variations in transducers—electrostatic and electromagnetic planar panels, horns, compression drivers, arrays, an so on—have been developed, but none has prevailed. The excellent fidelity, versatility, and cost effectiveness of moving-coil cone and dome transducers justify their widespread acceptance. Better mechanical and acoustical measurements, new engineering design aids, and more accurate performance targets have led to considerable improvements in sound quality. New diaphragm materials allow resonances to be better controlled or even eliminated by forcing them above the highest frequency reproduced by the transducer. Acoustical waveguides tailor the directivity of both direct radiators and compression drivers so that loudspeaker systems radiate sounds more uniformly at all frequencies. The benefit is a greatly improved interaction with rooms, large and small, matching the perceived timbral signatures of the direct, early reflected, and reverberant sounds. *See* SOUND FIELD ENHANCEMENT; SOUND-REINFORCEMENT SYSTEM.

The marriage of electronics and transducers is a welcome trend, as the performance of the entire system can then be optimized. Designers have more scope to perfect the crossovers between transducers in multiway systems. Digital techniques can refine the performance of the transducers by reducing residual linear and nonlinear distortions. Built-in equalization can compensate for acoustical problems specific to the listening room. Supplying the signals through a digital connection makes it possible to link the loudspeakers into a whole-room or whole-house plug-and-play audio/video network. All of these functions have been demonstrated. It remains to be seen which will become commonplace. Floyd Toole

Bibliography. K. B. Benson (ed.), *Audio Engineering Handbook*, 1988; J. Borwick (ed.), *Loudspeaker and Headphone Handbook*, 3d ed., 2001; S. E. Olive, B. Castro, and F. E. Toole, A New Laboratory for Evaluating Multichannel Systems and Audio Components, *105th AES Convention Audio Eng. Soc.*, Preprint 4842, September 1998; J. C. Steinberg and W. B. Snow, Auditory perspective: Physical factors, *Elec. Eng.*, 53:12–17, January 1934; F. E. Toole, The future of stereo, *Audio*, 81(5):126–142, May 1997, and 8(6):34–39, June 1997; F. E. Toole, Listening tests: Turning opinion into fact, *J. Audio Eng. Soc.*, 30:431–445, June 1982; F. E. Toole, Loudspeaker measurements and their relationship to listener preferences, *J. Audio Eng. Soc.*, 34:227–235, April 1986, and 34:323–348, May 1986; F. E. Toole, Subjective measurements of loudspeaker sound quality and listener performance, *J. Audio Eng. Soc.*, 33:2–32, January/February 1985; F. E. Toole and S. E. Olive, The modification of timbre by resonances: Perception and measurement, *J. Audio Eng. Soc.*, 36:122–142, March 1988.

Sourwood

A deciduous tree, *Oxydendrum arboreum*, of the heath family, indigenous to the southeastern section of the United States, and found from Pennsylvania to Florida and west to Indiana and Louisiana. Sourwood is hardy and often cultivated in the north. It is usually a small or medium-sized tree, but it sometimes attains

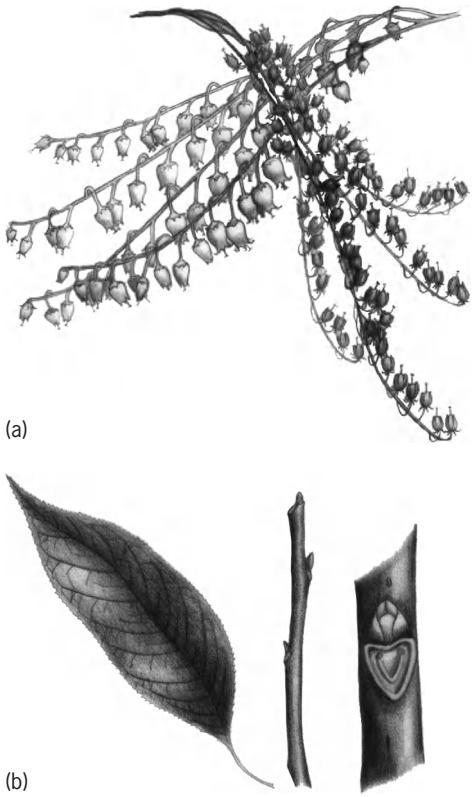

Sourwood (*Oxydendrum arboreum*). (*a*) Branch has urn-shaped flowers. (*b*) Leaf, twig, and enlargement of twig showing axial bud and leaf scar are below.

a height of 75 ft (23 m) and a diameter of nearly 2 ft (0.6 m). The leaves are simple, finely toothed, long-pointed, and 4–8 in. (10–20 cm) long, and have an acid taste which explains the popular name. They turn scarlet in the fall. The flowers are white and urn-shaped and grow in long clusters (see **illus.**). The dry fruit remains on the tree through the fall and winter. The wood is not used commercially. Sourwood is also known as sorrel tree, and it is widely planted as an ornamental. *See* ERICALES; FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

South America

The southernmost of the New World or Western Hemisphere continents, with three-fourths of it lying within the tropics. South America is approximately 4500 mi (7200 km) long and at its greatest width 3000 mi (4800 km). Its area is estimated to be about 7,000,000 mi^2 (18,000,000 km^2). South America has many unique physical features, such as the Earth's longest north-south mountain range (the Andes), highest waterfall (Angel Falls), highest navigable fresh-water lake (Lake Titicaca), and largest expanse of tropical rainforest (Amazonia). The western side of the continent has a deep subduction trench offshore, whereas the eastern continental shelf is more gently sloping and relatively shallow. *See* CONTINENT.

Regional characteristics. South America has three distinct regions: the relatively young Andes Mountains located parallel to the western coastline, the older Guiana and Brazilian Highlands located near the eastern margins of the continent, and an extensive lowland plains, which occupies the central portion of the continent. The regions have distinct physiographic and biotic features.

The Andes represent the highest and longest north-south mountain system on Earth. Altitudes often exceed 20,000 ft (6000 m) and perpetual snow tops many of the peaks, even along the Equator (**Fig. 1**). So high are the Andes in the northern half of the continent that few passes lie below 12,000 ft (3600 m). Over most of their length the Andes are not just a single range but two or three parallel ranges. Within the parallel peaks lie a vast series of intermontane basins and plateaus. Ecuador contains a string of 10 such basins. Bolivia has the Altiplano, an extensive basin about 400 mi (640 km) long and up to 200 mi (320 km) wide that is almost entirely surrounded by rugged and lofty peaks. The Altiplano is cold and high, averaging 12,000 ft (3600 m), and is mostly level. The Altiplano is distinctive in that it possesses the highest navigable fresh-water body, Lake Titicaca, measuring approximately 3488 mi^2 (8965 km^2) at 12,497 ft (3750 m) elevation and approximately 918 ft (275 m) in depth.

On the northeastern and eastern continental periphery lie the Guiana and the Brazilian Highlands. These vast areas of hilly uplands and low mountains are separated from each other by the Amazon River drainage (**Fig. 2**). Together these two plateaus form

Fig. 1. Andes Mountains, Chile. The Andes, which parallel the Pacific coast, constitute one of the Earth's most majestic features. Only the Himalayas exceed them in elevation and mass. (*United Press International*)

Fig. 2. Representative view of the hilly uplands and low mountain country of the Guiana Highlands. One of five waterfalls on the Canaima River, Gran Sabana in southern Venezuela. (*Venezuelan Information Service*)

the geologically oldest part of South America. Rocks of ancient igneous and metamorphic origin are partially covered by younger sedimentary beds and sandstones. The vegetation and animal life associated with the flat-topped mesas of the highlands, called *tepuis*, show a high proportion of endemicity (in-situ speciation) resulting from long periods of isolation and ecosystem stability.

More than 65% of South America's total area is characterized by lowland plains under 1000 ft (300 m) in elevation. Some 40% of the continent is less than 650 ft (200 m) above sea level. The plains lie from about 8°N to 40°S between the lofty Andean backbone on the west coast and the Guiana and the Brazilian Highlands on the east coast, and between the Río Orinoco in the north and the Río Colorado in the south. Also in this region are the Llanos of Venezuela and Colombia, the Amazon plain of Brazil, the Paraguayan Chaco, and the Pampas of Argentina. Some of these areas are quite flat while others are undulating. Some, such as the Llanos and Chaco, are alternately flooded and baked.

Physiographic features. The three Andes regions, the lowlands, and the eastern highlands are composed and bounded by many physiographic features, including mountains, intermontane basins, extensive lowland plains, coastlines, and rivers.

Mountains. Because of the vast extent of the Andes, a greater proportion of South America than of any other continent lies above 10,000 ft (3000 m). The young, rugged, folded Andean peaks stand in sharp contrast to the old, worn-down mountains of the eastern highlands. Although the Andes appear to be continuous, most geologists believe that they consist of several structural units, more or less joined. They are a single range in southern Chile, two ranges in Bolivia, and dominantly three ranges in Peru, Ecuador, and Colombia.

Except in Bolivia, where they attain their maximum width of 400 mi (640 km), the Andes are seldom more than 200 mi (320 km) wide. They do not equal the Himalayas in height, but have at least 30 peaks above 20,000 ft (6000 m). The average height of the Andes is estimated to be 13,000 ft (3900 m). However, it is only north of latitude 35°S that the mountains exceed elevations of 10,000 ft (3000 m) [see **table**].

Both active and quiescent volcanic formations are common in southern Colombia, Ecuador, central and southern Peru, and western Bolivia. These volcanic peaks are the surface expression of the subduction of the Nazca tectonic plate beneath the South American plate, forming the offshore Peru-Chile Trench. The Andes were raised up by tectonic forces from 6560 ft (1970 m) to an average 13,120 ft (3940 m) during the Pliocene uplift approximately 5 million years ago, and they continue to rise incrementally each year. As a result of these powerful tectonic forces, there is a spectacular series of volcanic peaks in western Bolivia and on each side of the structural depression in Ecuador. *See* PLATE TECTONICS.

Stratification of climate and vegetation with

altitude can be readily observed in the Andes. At their eastern base is a zone with hot, humid lowland and foothills up to 3300 ft (1000 m), known as *tierra caliente*. *Tierra templada*, the zone from 3300 to 6600 ft (1000 to 2000 m), has a relatively mild, frost-free climate and was preferred for European settlement and the production of plantation crops such as coffee and coca. From 6600 to 13,200 ft (2000 to 4000 m) is *tierra fria*, a montane zone with occasional frosts. *Tierra helada*, at 13,200 to 19,800 ft (4000 to 6000 m), occupies the zone between the daily frost line and the final zone of permanent snow called *tierra nevada*.

Because of the great north-south extent of the Andes, the processes of their erosion and denudation have varied. Southward from about 40°S, and especially in the far south, the Andes were heavily glaciated during the Ice Age, and an extensive area north of the Strait of Magellan still has a broad mantle of permanent ice. Glaciers descend to the heads of many fiords on the Pacific coast or into lakes on the eastern side of the mountains.

Coastal features. From the southern tip of Cape Horn north to 41°S latitude, the western coastal zone consists of a broad chain of islands where a mountainous strip subsided and the ocean invaded its valleys. This is one of the world's finest examples of a fiorded coast. Nowhere along the Pacific coast is there a true coastal plain. South of Arica, Chile, the bold, precipitous coast is broken by only a few deep streams, the majority of which carry no water for years at a time. Between Arica and Caldera, Chile, there are no natural harbors and almost no protected anchorages. In fact, South America's coastline is the least indented of all the continents except Africa's. The only exceptions are the fiorded coast and the offshore archipelagoes of southern Chile, the Gulf of Venezuela including Lake Maracaibo in the north, the Gulf of Guayaquil in the west, and the estuaries of the Amazon and the Río de la Plata in the east. *See* FIORD.

Fig. 3. Uplifted beaches on desert coast, east side of San Juan Bay, Departmento de Ica, Peru. The sea cliff is about 100 ft (30 m) high. Each horizontal line represents a former beach. (*Photograph by F. Atchley*)

Many raised beaches are seen along the coast (**Fig. 3**), supporting the premise that the Andes are of recent origin and that the forces that raised them are still at work. Nearly every port in coastal Peru is an open roadstead (less enclosed than a harbor). Coastal Ecuador, separated by mountain spurs and a narrow coastal plain east of the Gulf of Guayaquil, varies in width from 50 to 150 mi (80 to 240 km). Two important river systems cross it, the Esmeraldas in the north and the Guayas in the south. But the most important segment of the entire area is the Guayas Lowland, east and north of the Gulf of Guayaquil. Here are found one-quarter of Ecuador's total population, the most important commercial agricultural lands, and the principal port, Guayaquil. The climatic transition that occurs in coastal Ecuador is one of the most pronounced in the world. Within a distance of only 425 mi (680 km), one passes from areas having a rainy tropical climate, to wet and dry tropical, to semiarid tropical, and finally to arid tropical—in short, from the world's rainiest to its driest climates.

The Caribbean coast of Colombia is a lowland formed largely of alluvium, deposited by the Magdalena and Cauca rivers, and bounded by mountains on three sides. In Venezuela, the Central Highlands rise abruptly from the Caribbean, with lowlands around Lake Maracaibo, west of Puerto Cabello, and around the mouth of the Río Tuy of the Port of Guanta. The coastal region of Guyana, Suriname, and French Guiana is a low, swampy alluvial plain 10–30 mi (16–48 km) wide, and as much as 60 mi (96 km) wide along the larger rivers. This coastal plain is being built up by sediments carried by the Amazon to the Atlantic and then deflected westward by the equatorial current and cast upon the shore by the trade winds.

Principal Andean peaks*

Peak	Height, ft (m)
Aconcagua, Argentina	22,835 (6960)
Ampato, Peru	21,702 (6615)
Caca Aca, Bolivia	20,329 (6196)
Cachi, Argentina	21,326 (6500)
Chimborazo, Ecuador	20,577 (6272)
Cincel, Bolivia	20,102 (6127)
Condoriri, Bolivia	20,043 (6109)
Coropuna, Peru	22,802 (6950)
Cuzco (Ausangate), Peru	20,187 (6153)
Del Acay, Argentina	20,801 (6340)
Dos Conos, Argentina	22,507 (6860)
Falso Azufre, Argentina-Chile	22,277 (6790)
Huascaran, Peru	22,188 (6763)
Illampu, Bolivia	21,276 (6485)
Illimani, Bolivia	21,282 (6487)
Incahuasu, Argentina-Chile	21,720 (6620)
Llullaillaco, Argentina-Chile	22,015 (6710)
Mercedario, Argentina-Chile	21,884 (6670)
Ojos del Salado, Argentina-Chile	22,573 (6880)
Payachata, Bolivia	20,768 (6330)
Pissis, Argentina	22,245 (6780)
Porongos, Argentina-Chile	20,512 (6252)
Pular, Chile	20,342 (6200)
Sajama, Bolivia	21,390 (6520)
Sarmiento, Chile	20,670 (6300)
Socompa, Argentina-Chile	19,787 (6031)
Tocorpuri, Bolivia-Chile	22,163 (6755)
Tortolas, de las, Chile	20,018 (6101)
Tres Cruces, Chile	21,720 (6620)
Turpungato, Chile	21,490 (6550)
Valadero, Argentina	20,735 (6320)

*The elevations are approximate and some are controversial. Data based mostly on tables in *Goode's World Atlas*, 12th ed., Rand McNally, 1964.

There is no broad coastal plain south of the Amazon and east of the Brazilian Highlands to afford easy access to the interior as is characteristic of North America's coast. The rise from the coastal strip to the interior is quite gradual in northeastern Brazil; but southward, between Bahia and Río Grande do Sul, the steep Serra do Mar is a formidable obstacle to transportation. Nonetheless, the first railroad was constructed here connecting São Paulo in the coffee region with the port of Santos. Though only 40 mi (64 km) long, it was a remarkable feat since the abrupt escarpment of the Serra do Mar behind Río de Janeiro posed difficult engineering problems and required the use of cog tracks and cables.

Along coastal Uruguay there is a transition between the hilly uplands and plateaus of Brazil and the flat Pampas of Argentina, whereas coastal Argentina as far south as the Río Colorado, in Patagonia, is an almost featureless plain. In Patagonia, steep cliffs rise from the water's edge. Behind these cliffs lies a succession of dry, flat-topped plateaus, surmounted occasionally by hilly land composed of resistant crystalline rocks. Separating southern Patagonia from Tierra del Fuego is the Strait of Magellan, which is 350 mi (560 km) long and 2–20 mi (3–32 km) wide. Threading through numerous islands, the strait is lined on each side with fiords and mountains.

Major rivers and river systems. There are three great river systems in South America and a number of important rivers that are not a part of these systems. The largest river system is the Amazon which, with its many tributaries, drains a basin covering 2.7 million square miles (7 million square kilometers), or about 40% of the continent. The next largest is the system composed of the Paraguay, Paraná, and La Plata rivers, the last being a huge estuary. The third largest river system, located in southern Venezuela, is the Orinoco, which drains 365,000 mi^2 (945,000 km^2) of land, emptying into the Atlantic Ocean along the northeast edge of the continent.

Amazon River system. The Amazon River begins in the eastern Andean highlands and stretches across the north-central part of the continent, draining the interior lowlands. Not only is it the longest, but it carries the largest volume of water and silt of any river in the world. It is also navigable for oceangoing vessels for a greater distance than any other river. The Amazon discharges 3.4 billion gallons/min 13 billion liters/min and has an average flow of 7.5 million cubic feet/s (210,000 cubic meters/s). Its discharge is about 18% of that of all rivers of the world. In places, for example, between Obidos, Brazil (500 mi or 800 km upstream) and the mouth, its depth exceeds 300 ft (90 m). It drains an area of 2.3 million square meters (6 million square kilometers), and its gradient averages only 0.2 in./mi (3 mm/km) or about 35 ft (11 m) for the last 2000 mi (3200 km) to the sea. Considering the enormous amount of silt, sand, and alluvial debris the Amazon carries to the Atlantic Ocean, it might be expected that a huge delta would have formed. But the Amazon has no delta because of the steady submergence of the land where the river reaches the sea. The Amazon is joined by some 500 tributaries descending the Andes and the Brazilian and Guiana Highlands.

Paraguay-Paraná-La Plata system. From its headwaters in southwestern Brazil, the Paraguay River courses southward 1300 mi (1780 km), discharging into the Paraná at the southern edge of Paraguay. The western bank of the Paraguay is low, and during the rainy season the western side of the entire basin becomes inundated, flooding thousands of square miles of low-lying country.

The Paraná ranks among the world's major rivers. It is longer and carries more water than the Mississippi River in the United States. Its source in the Brazilian Highlands is 2450 mi (3920 km) from its outlet in the Río de la Plata. The Paraná has cut a deep canyon in the flat-topped plateau of its upper course. As the river drops over the edge of the lava formation, famous waterfalls are found—La Guayra and Iguazú, the latter on the Iguazú, a tributary of the Paraná. However, from Corrientes, where the water of the Paraguay is added, the gradient is gentle. Since the Paraná rises in areas with characteristically heavy summer rainfall, its water volume fluctuates widely, the variation from high to low water reaching 15 ft (4.5 m) in the lower course. At the mouth of the Paraná is an enormous delta, shared with the emptying Uruguay River. The delta consists of numerous low, flat islands that are submerged for weeks at a time.

The largest indentation on the east coast of South America occurs from the head of the estuary of La Plata to the open Atlantic and is almost 200 mi (320 km) long. At Buenos Aires the estuary is about 25 mi (40 km) wide, at Montevideo 60 mi (96 km), and at its mouth 138 mi (221 km). The Paraná and the Uruguay transport huge quantities of silt into La Plata, some of which settles in navigation channels, necessitating costly dredging.

Navigation for vessels drafting 6–7 ft (1.8–2.1 m) is possible as far as Corumba, Brazil, some 1800 mi (88 km) upstream from Buenos Aires via the Paraná and Paraguay rivers. The remainder of the Paraná is rarely used for navigation, and its hydroelectric potential is as yet undeveloped because markets are too distant. The Paraguay, however, serves the interior, and is considered the lifeline to the outside world for Paraguay.

Orinoco River system. The headwaters of the Orinoco River lie in southern Venezuela near the Brazilian border. Near Esmeraldas, in southern Venezuela, is a stretch of about 220 mi (350 km) known as the Casiquiare canal, joining the Orinoco with the Negro of the Amazon system.

The Orinoco, approximately 1600 mi (2560 km) long, is noted for its variability in volume from wet to dry season. It is a broad stream, and is a mile wide at Ciudad Bolivar, Venezuela, the "narrows." Rainfall from June to October is so heavy that the Orinoco and its many tributaries are unable to handle all the water, and enormous areas become inundated. The river rises 39 ft (12 m) at Puerto Ordaz, Venezuela. From January to March, the dry season,

the waters recede and only the larger rivers flow freely. The smaller rivers are gradually converted into chains of pools and swamps lying along the valley bottoms. Where the Orinoco flows into the Atlantic, it forms a low, wide delta consisting of small islands and swamps and an intricate maze of distributaries.

Biotic features. The plants and animals of the South American tropics are classified as Neotropical, defined by the separation of the South American and African continents during the Middle Cretaceous (95 million years ago). The Paraná basalt flow, which caps the Brazilian shield in southern Brazil and adjacent parts of Uruguay and Argentina, as well as western Africa, indicates the previous linkage between the South American and African continents. During the Middle Cretaceous Period, angiosperms (flowering plants) expanded in distribution across the Earth's landmasses, largely replacing the gymnosperms, which previously dominated the Gondwana landmass. While there are regions of important endemicity for plants and animals within South America, many evolved early enough to share significant similarities with North American, Eurasian, and African fauna and flora. South America has many biotic environments, including the constantly moist tropical rainforest, seasonally dry deciduous forests and savannas, and high-altitude tundra and glaciated environments.

Tropical rainforest. Amazonia contains the largest extent of tropical rainforest on Earth. It is estimated to encompass up to 20% of the Earth's higher plant species and is a critically important source of fresh water and oxygen. Structurally complex, the rainforest is composed of up to four distinct vertical layers of plants and their associated fauna. The layers often cluster at 10, 20, 98, and 164 ft (3, 6, 30, and 50 m) in height. The highest levels of canopy experience intense competition for light and water. Epiphytes and lianas are often found in the taller trees of this layer. The lower canopy and forest floor are usually open spaces because of the low intensity of light (around 1%) that reaches the forest floor. Over 75% of Amazonian soils are classified as infertile, acidic, or poorly drained, making them undesirable for agriculture because of nutrient deficiencies. Most of the nutrients in the tropical rainforest are quickly absorbed and stored in plant biomass because the high annual rainfall and associated leaching make it impossible to maintain nutrients in the soils. In addition to the high structural complexity of the tropical rainforest, there is considerable horizontal diversity or patchiness. As many as 300 separate species of trees can be found in a square mile (2.6 km^2) sample tract of rainforest in Brazil. The high complexity and species diversity of the rainforest are the result of long periods of relative stability in these regions. Many areas of Amazonia show unique concentrations of species that suggest these territories may have acted as Pleistocene refugia where plants and animals were able to survive the glacial advances of higher latitudes. *See* PLEISTOCENE.

Deciduous and conifer forest. Deciduous forests are found in areas where there is seasonal drought and the trees lose their leaves in order to slow transpiration. The lower slopes of the Andes, central Venezuela, and central Brazil are areas where these formations are found. Conifer forests occur in the higher elevations of the Andes and the higher latitudes of Chile and Argentina. Many of the species of these formations derive from the Late Cretaceous Southern Gondwana flora, including the *Araucaria* (star pine), *Podocarpus*, and *Dacrydium*. *See* DECIDUOUS PLANTS.

Savanna. Tropical savannas occupy an extensive range in northern South America through southeastern Venezuela and eastern Colombia. Temperate savannas are found in Paraguay, Uruguay, the Pampas of Argentina, and to the south, Patagonia. Savannas are composed of a combination of grass and tree species. The climate in these areas is often quite hot with high rates of evapotranspiration and a pronounced dry season. Most of the plants and animals of these zones are drought-adapted and fire-adapted. Tall grasses up to 12 ft (3.5 m) are common as are thorny trees of the Acacia (Fabaceae) family. Many birds and mammals are found in these zones, including anteater, armadillo, capybara (the largest rodent on Earth), deer, jaguar, and numerous species of venomous snake, including rattlesnake and bushmaster (*mapanare*).

Desert. South America is unique in having a west-coast desert that extends almost to the Equator, probably receiving less rain than any on Earth (the Atacama), and an east coast desert located poleward from latitude 40°S (the Patagonian).

The Atacama-Peruvian Desert dominates the Pacific coast for nearly 2000 mi (3200 km), lying between the ocean and the higher slopes of the Andean ranges. In many places, particularly in Chile, the land rises abruptly from the sea to an elevation of 2000–3000 ft (600–900 m). The driest stretch, one of the driest places in the world, is between Arica and Caldera, Chile (**Fig. 4**). Above the seaward scarp is a low range of hills and a desert trough, 30–50 mi (48–80 km) wide, from which nitrate deposits are extracted. The persistence of this desert through latitudes at which deserts are not found elsewhere is the result of the north-moving Humboldt Current (Peru Current), accompanied along the immediate shore by the upwelling of cold water from considerable depths, creating stable, high-pressure atmospheric

Fig. 4. Port of Matarani on the desert coast of southern Peru. This oblique aerial view looks coastwise southward over one of the world's driest and most barren coastal zones. (*Servicio Aerofotográfico Nacional*)

conditions. Prevailing winds, blowing from the west, carry cold air onto the warmer land with a drying effect.

South America's east-coast desert, the Patagonian, is found between the Río Colorado and the Strait of Magellan. This desert is partly located in the rain shadow of the Andes, and the meager annual rainfall is ordinarily 5–6 in. (13–15 cm). Also contributing to the region's aridity and to the presence of the desert near to the shores of the Atlantic are the cold waters of the Falkland Current, moving northward off the eastern coast, and the ceaseless winds, which hasten evaporation. Wind velocities can exceed 70 mi/h (31 m/s). What precipitation there is occurs mostly in the winter. The region's very cold winter temperatures are the result of high latitude and high elevation.

Tundra. In Bolivia and Peru the zone from 10,000 to 13,000 ft (3000 to 3900 m), though occasionally to 15,000–16,000 ft (4500 to 4800 m), is known as the *puna*. Here the hot days contrast sharply with the cold nights. Above the *puna*, from timberline to snowline, is the *paramo*, a region of broadleaf herbs and grasses found in the highest elevations of Venezuela, Colombia, and Ecuador. Many of the plant species in these environments are similar to those found at lower elevations; however, they grow closer to the ground in order to conserve heat and moisture.

Deforestation and environmental contamination. The 1992 Quincentennial and the 1992 Earth Summit (United Nations Conference on Environment and Development) have drawn widespead attention to environmental policy and economic development in South American countries that have encouraged the destruction of natural vegetation. Intense political discourse has focused on the removal of tropical rainforest and the contamination of both lowland and highland environments across the continent. An estimated 20% of the Amazonian forest has already been destroyed, and more than 7700 mi^2 (200,000 km^2) are estimated to be replaced by agriculture, cattle ranching, and urbanization each year. The loss of tropical rainforest in particular has been widely publicized, although many other biotic environments, including deciduous forest, coastal mangrove, and tundra environments, are also being degraded by human use.

Deborah A. Salazar; C. Langdon White

Bibliography. B. Box, *South American Handbook*, 82d ed., 2006; C. Caviedes and G. Knapp, *South America*, 1995; B. Cox and B. Moore, *Biogeography*, 7th ed., 2005; A. Morris, *South America*, 3d ed., 1991; S. E. Place, *Tropical Rainforests*, rev. ed., 2001; A. H. Strahler and A. Strahler, *Introducing Physical Geography*, 4th ed., 2005.

South Pole

That end of the Earth's rotational axis opposite the North Pole. It is the southernmost point on the Earth and one of the two points through which all meridians pass (the North Pole being the other point). This is the geographic pole and is not the same as the south magnetic pole, which is near the coast of Antarctica, south of Tasmania. The South Pole lies inland from the Ross Sea, within the landmass of Antarctica, at an elevation of about 9200 ft (2800 m).

Since the Earth's axis is inclined $23^1/_2{}^\circ$ from the vertical to the plane of the ecliptic, the South Pole experiences six months when the Sun is continuously above the horizon, reaching a maximum altitude of $23^1/_2{}^\circ$. It also has 6 months when the Sun remains continuously below the horizon, from March 22 to September 22. During the latter period, however, there is an interval of 7 weeks at the beginning and at the end when twilight conditions preclude total darkness. Twilight prevails when the Sun is located between 0° and 18° below the horizon.

There is no natural way to determine local Sun time because there is no noon position of the Sun, and shadows point north at all times, there being no other direction from the South Pole. *See* MATHEMATICAL GEOGRAPHY; NORTH POLE.

Van H. English

Bibliography. A. H. Strahler and A. Strahler, *Introducing Physical Geography*, 4th ed., 2005.

Southeast Asian waters

All the seas between Asia and Australia and the Pacific and the Indian oceans. They form a geographical and oceanographical unit because of their special structure and position, and make up an area of 3,450,000 mi^2 (8,940,000 km^2), or about 2.5% of the surface of all oceans.

Ocean floor. The ocean floor consists of two large shelves and a number of deep-sea depressions. The Sunda Shelf is covered with littoral sediments, and the submerged valleys of two large river systems are found on it. The Arafura Shelf connects New Guinea with Australia. The China Basin, whose maximal depth is 16,450 ft (5016 m), is connected with the Pacific Ocean over a sill about 6000 ft (2000 m) in depth. The Sulu Basin (maximal depth 18,300 ft or 5580 m) has the highest sill (1400 ft or 420 m). The Celebes Basin (maximal depth 20,400 ft or 6220 m) is connected with the Pacific Ocean over three sills of about 4200 ft (1400 m) each. The Banda Basin is subdivided into several smaller basins, including the Weber Deep of 24,400 ft (7440 m); its sill depth is 6170 ft (1880 m). In the Celebes, Banda, and Flores basins, volcanic ashes are found. *See* EAST INDIES.

Surface circulation. The surface circulation is completely reversed twice a year (**Figs. 1** and **2**) by the changing monsoon winds.

Monsoon current. During the north monsoon, which is fully developed in February, the monsoon current is formed in the northern China Sea, flows along the coast of Vietnam into the Java Sea, and into the Flores Sea. Parts of its water masses return into the Pacific Ocean; parts sink in the Banda Sea or flow into the Indian Ocean.

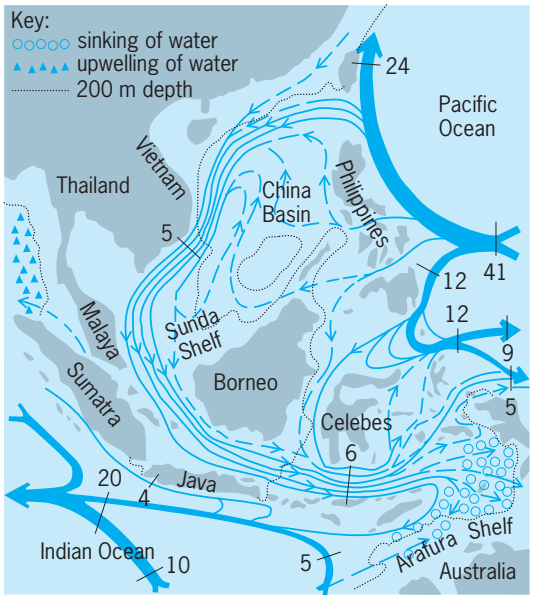

Fig. 1. Water circulation in South China and Indonesian seas, February. Transports in 10^6 m^3/s.

During the south monsoon in August, the monsoon current is formed by water flowing from the Java Sea into the South China Sea and farther into the Pacific Ocean.

The transports of the monsoon current are 110,000,000 ft^3/s (3,000,000 m^3/s) in August and 180,000,000 ft^3/s (5,000,000 m^3/s) in February, but are small compared with those in the adjoining parts of the Pacific and Indian oceans (Figs. 1 and 2).

Indonesian throughflow. Indonesian throughflow is a complex system of currents within the depth range of the thermocline to 1000 m (3300 ft) that allows water to flow from the Pacific to the Indian Ocean. It is the only place in the global oceans where warm equatorial waters pass from one ocean basin into another. Throughflow water must be replaced in the Pacific by inflowing, colder water at a high latitude in the South Pacific. Consequently, Indonesian throughflow exports a large amount of heat from the Pacific. The shift in heat between ocean basins, and the subsequent release of the heat to the atmosphere over the Indian Ocean, is recognized as a critical component of the global climate system. Consequently, Indonesian throughflow has been the subject of intensive oceanographic research. Attempts to understand why it exists led to fundamental theoretical advances in geophysical fluid dynamics. The studies showed that the mean throughflow is a response to the wind field over the whole Pacific Ocean.

The main path of Indonesian throughflow is from the Celebes Sea near Mindanao, through Makassar Strait, eastward along the Indonesian archipelago through the Flores and Banda seas, and finally into the Indian Ocean passing south of Timor. Some flow passes through the lesser Sunda Islands. The strength of throughflow is probably not larger that 360,000, 000 ft^3/s (10,000,000 m^3/s), while the temperature difference from the South Pacific is about 10°C (18°F). The large temperature difference is the factor that makes Indonesian throughflow critical in the global climate system. The throughflow is strongest during the south monsoon (Fig. 2), when upwelling water from the Banda Sea contributes to the flow. Its strength varies considerably from one year to the next in response to changing winds over the equatorial Pacific and Indian oceans. The variability is related to the El Niño Southern Oscillation phenomenon. *See* EL NIÑO.

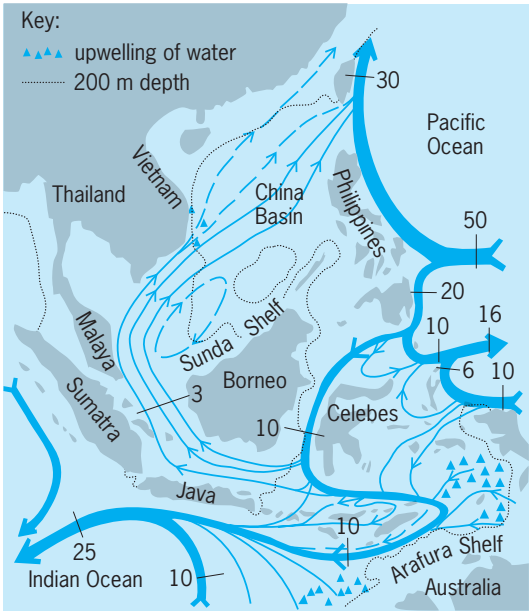

Fig. 2. Water circulation in South China and Indonesian seas, August. Transports in 10^6 m^3/s.

Salinity and temperature. The monsoons cause a pronounced rainy and dry season over most parts of the region and consequently strong annual variations of the surface salinity. During the north monsoon, rainfall is strongest to the south of the Equator and surface salinity decreases, while in the China Sea the excess of evaporation over rainfall increases salinity. During the south monsoon, conditions are reversed; strong rainfall over the China Sea decreases its salinity, while evaporation and upwelling increase salinity in the Java, Flores, and Banda seas. Regions of permanent low salinities are in the Malacca Straits, the Gulf of Thailand, and the waters between Borneo and Sumatra, because of the discharge of the large rivers. The surface temperature is always high, between 79 and 84°F (26 and 29°C), and drops only in the northern parts of the South China Sea and in the Arafura Sea to about 75°F (24°C) during the dry season. *See* MONSOON METEOROLOGY.

Deep circulation. The deep basins are supplied with Pacific Ocean deep water that enters over the various sills (**Fig. 3**). Only the Timor and Aroe basins are filled with Indian Ocean deep water. The renewal of water in the deep-sea basins is sufficiently rapid, on the order of 100 years, to maintain an adequate level of oxygen for the deep-sea fauna.

Tides. The tides are mostly of the mixed type. Diurnal tides are found in the Java Sea, Gulf of Tonkin, and Gulf of Thailand. Semidiurnal tides with high amplitudes occur in the Malacca Straits. Strong tidal

Fig. 3. Flow of bottom waters in the eastern part of the Indonesian Archipelago.

currents occur in many of the small shallow passages. *See* INDIAN OCEAN; PACIFIC OCEAN; TIDE.

Klaus Wyrtki; Gary Meyers

Geology and geophysics. The Southeast Asian seas are characterized by the presence of numerous major plate boundaries. Plate boundaries represent sites where the Earth's crust is created or destroyed, or where rigid blocks of crust slide past one another. In Southeast Asia the plate boundaries are identified, respectively, by young, small ocean basins with their spreading systems and associated high heat flow; deep-sea trenches and their associated earthquake zones and volcanic chains; and major strike-slip faults such as the Philippine Fault (similar to the San Andreas Fault in California). The Southeast Asian seas are thus composed of a mosaic of about 10 small ocean basins whose boundaries are defined mainly by trenches and volcanic arcs. The dimensions of these basins are much smaller than the basins of the major oceans. The major topographic features of the region are believed to represent the surface expression of plate interactions, the scars left behind on the sea floor. *See* BASIN; FAULT AND FAULT STRUCTURES.

The proximity of important metalliferous ore deposits in selected islands of Southeast Asia, coupled with the fact that other nearby areas show significant potential for oil and gas, has led many scientists to investigate models that could relate plate tectonic processes to the formation of ore deposits and oil.

Ocean basins. The small ocean basins which make up the Southeast Asian waters are relatively young when compared with the ages of the major ocean basins. The basins of Southeast Asia that have been dated with some confidence are all younger than approximately 50,000,000 years. As part of the Deep Sea Drilling Project, the vessel *Glomar Challenger* drilled several holes along an east-west line (near 18°N) from the central part of the Philippine Sea into the Pacific Basin. One purpose of the drilling program was to determine the age and chemical properties of the underlying sea floor for comparison with the sea floor of the major ocean basins. The ultimate goal of the drilling is to determine the evolutionary processes which have given rise to the extremely complex tectonic features which are observed in the Southeast Asian seas.

Sediments. The sediments that blanket the sea floor in the Southeast Asian seas primarily reflect the proximity of land (that is, sediment sources) and the age and depth of the underlying sea floor. The thickly sedimented areas are found close to the continental margins and island arcs. The sediments here are composed mostly of materials transported from the adjacent continents and islands. It is possible that an additional component of the sediments found near the margins was first deposited on the deep-sea floor but was subsequently uplifted onto the margins when certain plates collided. *See* CONTINENTAL MARGIN; MARINE SEDIMENTS.

Tectonic processes. The study and understanding of the tectonic processes active in this region are of fundamental importance to many unresolved problems of the Earth's history. For example, it is a widely considered hypothesis that the continental regions may have evolved through the subduction (underthrusting) and the obduction (overthrusting) of ancient island arc systems with one another and with preexisting continental fragments. Another important hypothesis regarding the origin of metalliferous ore deposits is known as the geostill hypothesis. This postulates that rich metalliferous deposits were formed on or in the ocean crust near sites of plate accretion, that they were transported laterally to the arc-trench areas by sea-floor spreading processes and there were thrust back into the Earth's interior; eventually they were remelted and migrated back to the Earth's surface. When these metals (for example, copper) are found in concentrations such as in the Philippines, they are of great economic importance.

Finally, the release of energy by earthquakes and volcanic eruptions at the collisional plate boundaries underlines the need for more studies so that the destructive effects of these events can be minimized through accurate monitoring and prediction programs. *See* MARINE GEOLOGY; PLATE TECTONICS.

Dennis E. Hayes

Bibliography. G. Dietrich, *General Oceanography: An Introduction*, 2d ed., 1980; A. B. Duxbury and A. C. Duxbury, *Fundamentals of Oceanography,* 2d ed., 1995; J. J. Gotthold and D. W. Gotthold, *Indian Ocean*, 1988; D. E. Hayes (ed.), *The Tectonic and Geologic Evolution of Southeast Asian Seas and Islands*, pt. 1, 1980, pt. 2, 1983; D. E. Hayes et al., *A Geophysical Atlas of East and Southeast Seas*, Geological Society of America, 1978; R. Lukas, T. Yamagata, and J. P. McCreary (eds.), *Pacific Low-Latitude Western Boundary Currents and the Indonesian Throughflow*, American Geophysical Union reprint from *J. Geophys. Res.*, 101(C5):12209–12488, May 15, 1996; H. U. Sverdrup, M. W. Johnson, and R. H. Fleming, *The Oceans: Their Physics, Chemistry, and General Biology*, 1946; K. Wyrtki, *Oceanographic Atlas of the International Indian Ocean Expedition*, 1987.

Soybean

Glycine max, a legume native to China that is a major source of vegetable protein and oil for human and animal consumption and for industrial usage. Introduced into North America in 1765, soybeans were first grown as a curiosity and then as a forage and soil-building crop. The valued portion of the plant is the seed, which contains about 40% protein and 20% oil on a dry-weight basis.

Economic importance. The United States is the world's largest producer of soybeans. Illinois, Iowa, Minnesota, Indiana, Ohio, Missouri, Nebraska, Arkansas, South Dakota, Kansas, Michigan, and Mississippi are the major soybean producers there. Ranking after the United States are Brazil, China, and Argentina. About 64% of the soybeans produced in the United States are crushed or processed domestically, 31% are exported as whole beans, 3% are used for seed, and 2% have miscellaneous uses. Countries that import large quantities of whole beans for processing include Japan, the Netherlands, Taiwan, Spain, Mexico, South Korea, Germany, and Belgium. Russia, Canada, Algeria, and Venezuela import large quantities of soybean meal; and Pakistan, Morocco, Tunisia, Ecuador, and the Dominican Republic are the major importers of soybean oil.

Livestock consume about 97% of the soybean protein that is utilized in the United States. There is a limited use of the protein for human consumption as soy flours, grits, and texturized proteins in combination with other foods. Soybean oil is widely used throughout the world for human consumption as margarine, salad and cooking oils, and shortening. A small amount of protein and oil is used for industrial purposes.

In China, Taiwan, Japan, Korea, and Indonesia, soybeans are a regular part of the human diet as a source of protein and oil. Soybean milk, made by soaking beans in water, grinding the soaked beans, and filtering the insoluble pieces out, is directly consumed primarily by the Chinese; it is the starting material for preparation of soybean curd, probably the most important and popular soybean food in Asia. The soft cheeselike curd, known as tou fu, tofu, tubu, and other local names, has a bland taste and can be flavored with seasonings or blended with other foods. Soybean curd is made into a variety of products by frying, dehydration, and freezing, and is consumed daily in the same manner as high-protein foods in Western cultures. Soybeans also are prepared by fermentation with microorganisms to form highly flavored foods and seasonings such as soy sauce and fermented soybean paste. Soybean sprouts, immature green seed, and deep-fried mature seed provide additional sources of food for human consumption.

Plant characteristics. The soybean is an annual plant that reaches a mature height of 25–50 in. (64–127 cm). The cotyledons are the parts of the seed that emerge from the soil as the seedling develops. Two unifoliolate leaves (single leaflets) develop opposite each other above the cotyledonary nodes. Other leaves produced on the plant generally are trifoliolate (three leaflets). Branches develop from the main stem in response to the amount of space between adjacent plants. A plant may produce few branches and only a few seeds in close spacings or many branches and over a thousand seeds in wide spacings.

Flowers of soybeans are self-fertilized and white or varying shades of purple. Pubescence (hair) on the stems, leaves, and pods is brown or gray. Pod color of commercial varieties in the United States is brown or tan, although genetic types are available with black pods. The pods generally contain one to four seeds ranging in size from about 800 to 10,000 seeds per pound (1800 to 22,000 seeds per kilogram); however, most commercial varieties have two or three seeds per pod with a size of 1600 to 3600 seeds per pound (3600 to 8000 seeds per kilogram). The seedcoat of commercial varieties is yellow, and the common colors of the hilum (seed scar showing point of attachment to pod) are yellow, gray, buff, brown, imperfect black, or black. Varieties of soybean native to Asia have yellow, green, black, brown, and varicolored seedcoats.

As a legume, soybeans can obtain nitrogen needed for growth from nitrogen-fixing bacteria that live in their roots. The bacteria can be applied at the time of planting by mixing them with the seed or by direct application to the soil. They can survive in most soils for several years and do not have to be reapplied when soybeans are grown in a field every 2 or 3 years. *See* LEGUME; NITROGEN FIXATION.

The growing season for soybean varieties is controlled by their response to photoperiod and temperature; consequently, the most productive varieties for the northern United States are different from those grown in the South. Number of days from planting to maturity can range from about 90 to 180, depending on the variety and environmental conditions.

More than 15,000 varieties and selections of soybeans have been introduced into the United States from other areas of the world, and in the past some of them were grown commercially. Current production, however, is entirely devoted to varieties produced by the breeding programs conducted by the U.S. Department of Agriculture, state agriculture experiment stations, and private companies.

Most soybeans in the United States are planted in a well-prepared seedbed, although in some locations they are planted in unplowed fields to reduce soil erosion. In areas where the growing season is long enough, soybeans are planted as a second crop immediately after wheat or other crops are harvested. Planting is done in rows ranging in width from 7 to 40 in. (18 to 102 cm) with equipment also used to plant corn, cotton, wheat, and similar crops. Weed control includes the use of chemical herbicides, mechanical cultivation, and hand labor. Harvest of soybeans with combines begins when the leaves have dropped and moisture in the seed is at or below the 13% level required for safe storage.

Walter R. Fehr

Diseases

Most diseases of soybean that occur throughout the world can be found in the United States. Diseases tend to be chronic in nature, occurring with some regularity from year to year and taking a small (and often unnoticed) toll. Annual losses from all diseases have been estimated at 12%, and this figure can fluctuate considerably, depending upon environmental conditions in any one particular season. Many soybean diseases are usually found in any one field, and a disease-free field is rare indeed.

Although the term disease is often difficult to define, it is usually considered that soybean diseases can be caused by both living and nonliving entities. In the first category are the parasitic diseases caused by fungi, nematodes, bacteria, and viruses. The nonliving causes of disease include conditions of weather and soil and the deficiency or excess of various minerals. Iron, potassium, and manganese deficiencies are common, and varying degrees of chlorosis accompany such deficits. Herbicide and lightning damage are also often reported by growers.

Fungus diseases. Fungi cause more diseases than other parasitic organisms. Leaves can be attacked by *Peronospora manshurica* and become severely spotted with yellowish-green areas on upper leaf surfaces. If the case is severe enough, the leaves can fall prematurely; the leaves more often remain intact, and the spots take on a dark-brown appearance with fungus growth apparent on the lower side of leaves. This symptom gives rise to the common term mildew. The fungus spreads rapidly from plant to plant and may overwinter in fallen leaves or invade pods and survive on seeds as a white crust. Many pathogenic variants of this fungus are found, and resistant varieties of soybean are being developed. Up to 8% reduction in yield has been reported from this disease alone in Iowa.

In the Midwest, brown spot, caused by *Septoria glycines*, is another disease of importance on leaves. It appears early in the season as angular lesions on primary leaves and may continue to move up the plant and cause defoliation, attack pods, and ultimately invade seed. Target spot, a disease caused by *Corynespora cassiicola*, is found farther south in the soybean-growing region and has caused losses in the range of 18–32% in Mississippi. Its name is derived from concentric dark rings formed as large brown and often irregular spots develop on leaves. Fungus leaf spots of minor importance include those caused by *Cercospora sojina*, *Phyllosticta sojicola*, and *Alternaria atrans*.

A series of fungus root and stem rots may plague soybeans. Root rot, caused by *Phytophthora megasperma*, is probably the most spectacular and serious disease in areas where it occurs. Plants die in all stages of growth, usually in low places in fields and along end turn rows. Several states and Canada have reported damage from this disease, and control is effected by the use of resistant varieties. Inheritance of resistance is controlled by a single factor, thus making it relatively simple to incorporate this resistance into promising (or preexisting) varieties of soybeans. The pathogen is not seed-borne.

By 1980, 19 pathogenic races had been described, and some researchers believe that single-gene resistance or tolerance no longer is a viable approach to control. Either a generalized resistance or tolerance has been considered. Most commercial cultivars possess considerable field resistance to *Phytophthora*.

Brown stem rot, caused by *Cephalosporium gregatum*, is a disease of major concern in the Midwest, and the fungus is widespread in Minnesota, Illinois, Iowa, and Indiana. Crop rotation (on a 3–4-year basis) is a well-known control. Resistance as a viable method of control was investigated in the 1970s, and by the end of that decade at least two major producing states were considering this method of control. Seed rot and root rot by *Pythium* and *Rhizoctonia* are common and chronic on seedlings over vast areas of the soybean region. Plants able to emerge usually outgrow this problem; however, stand reduction may be severe enough to warrant replanting. Stem canker and pod and stem blight are two stages of an important disease caused by *Diaporthe phaseolorum* which are favored by high humidity and can be controlled to some extent by chemical treatment. A widespread and serious seed quality problem in the warmer and more humid regions of the South and Midwest results from pod infection by *Diaporthe*. Systemic fungicides have provided some improvement in seed quality and yields.

Soybean rust, a fungus-induced disease, is assuming more importance as acreages expand in traditional soybean growing areas and as production moves into new areas. The disease is serious in Asia, especially in tropical and subtropical areas. In Taiwan, Thailand, the Philippines, and Java it is the most damaging fungal disease affecting soybean crops. The pathogen, *Phakopsora pachyrhizi*, can infect and sporulate on species in 17 known legume genera but is economically important only on soybean. The pathogen has been reported in Africa on cowpea but not on soybean. In 1976 soybean rust was found on soybean cultivars Biloxi, Hardee, Santa Rosa, and Williams in an experimental planting in the highlands of western Puerto Rico. This was the first record of rust on soybean in the Western Hemisphere, and is cause for concern in the Americas. It is likely that the pathogen will be found on soybean plantings elsewhere in the Caribbean area.

Bacterial diseases. Bacterial diseases usually are a chronic problem of this crop. Bacterial blight (**Fig. 1**) caused by *Pseudomonas glycinea* is found primarily in the upper Midwest and bacterial pustule caused by *Xanthomonas phaseoli* predominates in the southern United States. These pathogens enter plants through pores or injuries in leaves and are spread rapidly by windblown rain. Seed-borne infection is common and constitutes the major means of overwintering. Infected seedlings can thus emerge, and the disease cycle begins again. Antibiotic sprays are impractical, and resistant strains of the blight pathogen develop rapidly during their use. Some varieties of soybean are more resistant to blight than

Fig. 1. Bacterial blight symptoms on soybean leaves under field conditions. The condition results principally in chlorate spots and malformations.

others (**Fig. 2**), and inheritance of resistance in the soybean host depends upon the particular strain of the pathogen being tested.

A chlorosis (yellowing of leaves) may develop from strains of *Rhizobia*, which are abnormal in that they cannot fix nitrogen but are aggressive in entering soybean roots. These bacteria thus occupy space that efficient nitrogen fixers would otherwise occupy. A bacterial wilt caused by *Corynebacterium flaccumfaciens* has been reported in certain areas of the Midwest but is not of great economic importance.

Virus diseases. The viruses may enter soybean plants by any of several methods, although each virus usually is very specific in its mode of entrance. They may be injected into the plant by insects or nematodes or through wounds or may go from one generation to the next via the seed. Only a few of the viruses of this crop have been studied extensively.

Soybean mosaic is common and is easily transmitted mechanically or by insects or seeds. It usually appears on leaves as crinkling and distortion of young tissues, and damage is similar to that produced by sprays used in weed control. It affects the seed by producing a mottled appearance; the seed becomes discolored when pigments from the seed scar appear to streak across seed coat surfaces during their development. Control is difficult, but yield losses have not been excessive.

Bud blight, on the other hand, is a very destructive virus disease, and no resistance is known. It, too, is seed-borne; it is apparently transmitted by insects. Usually, isolated plants in a field are found scattered at random (originating from infected seed), or plants along the edge of fields bordering alfalfa or fence rows may be heavily infected. This phase is thought to be due to an insect vector moving the virus from other plants to soybeans. Loss in yield can be 100% of the crop. This virus causes the plant to produce a rank green stem which stays green after frost, and seed pods are either absent or small; if infected while young, the growing points of plants die and curve downward, giving rise to the common name "bud blight." The same virus causes ring spot in tobacco. *See* PLANT VIRUSES AND VIROIDS.

Nematodes. These pathogens are microscopic, eel-like worms in the soil. Members of the genus *Meloidogyne* cause root knots or galls which form as large swellings and cause severe yellowing and stunting of aboveground portions of the plant. Crop rotation and the use of resistant varieties constitute the usual control methods. The soybean cyst nematode *Heterodera glycines* represents a real threat to soybean production and is now so widespread that quarantine no longer is effective and has been abandoned. It was first found in 1952 in North Carolina and since then has been found in isolated places in Virginia, Mississippi, Arkansas, Illinois, Kentucky, Missouri, and Tennessee. It is now feared from southern to northern climates. It moved northward in Illinois for many years, and was found for the first time as far north as Minnesota in 1978 and in Iowa in 1979; it appears to be destructive in northern as well as southern climates. Symptoms vary but are not very dissimilar to root knot symptoms on aboveground parts of the plants. Below ground, small white cysts are noted; once present, these cysts can exist many years in the soil, and complete control becomes almost impossible. Yield loss may be severe. Resistant varieties and rotation with nonsusceptible crops are recommended. *See* NEMATA (NEMATODA); PLANT PATHOLOGY.

Bill W. Kennedy

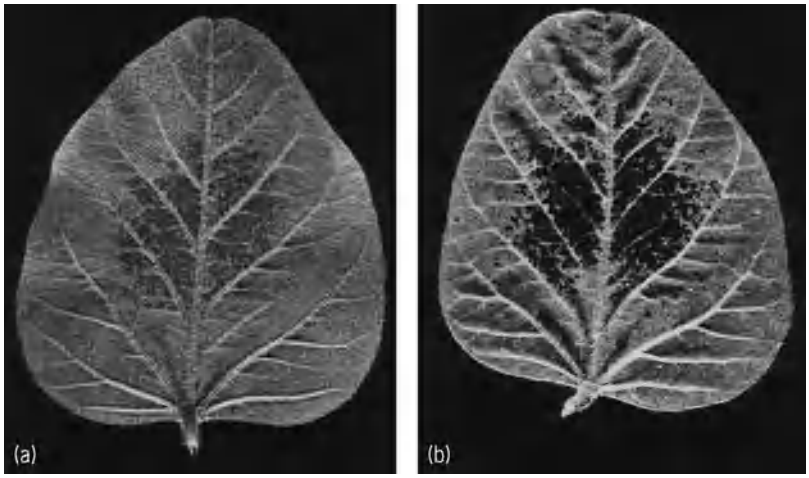

Fig. 2. Response of soybean varieties to bacterial blight. (*a*) Complete resistance. (*b*) Complete susceptibility.

Processing

Processing of soybeans into oil and defatted cake or meal began in the Orient. The oil was used primarily in foods, whereas the meal was used as a fertilizer and in animal feed. In the United States, soybeans have

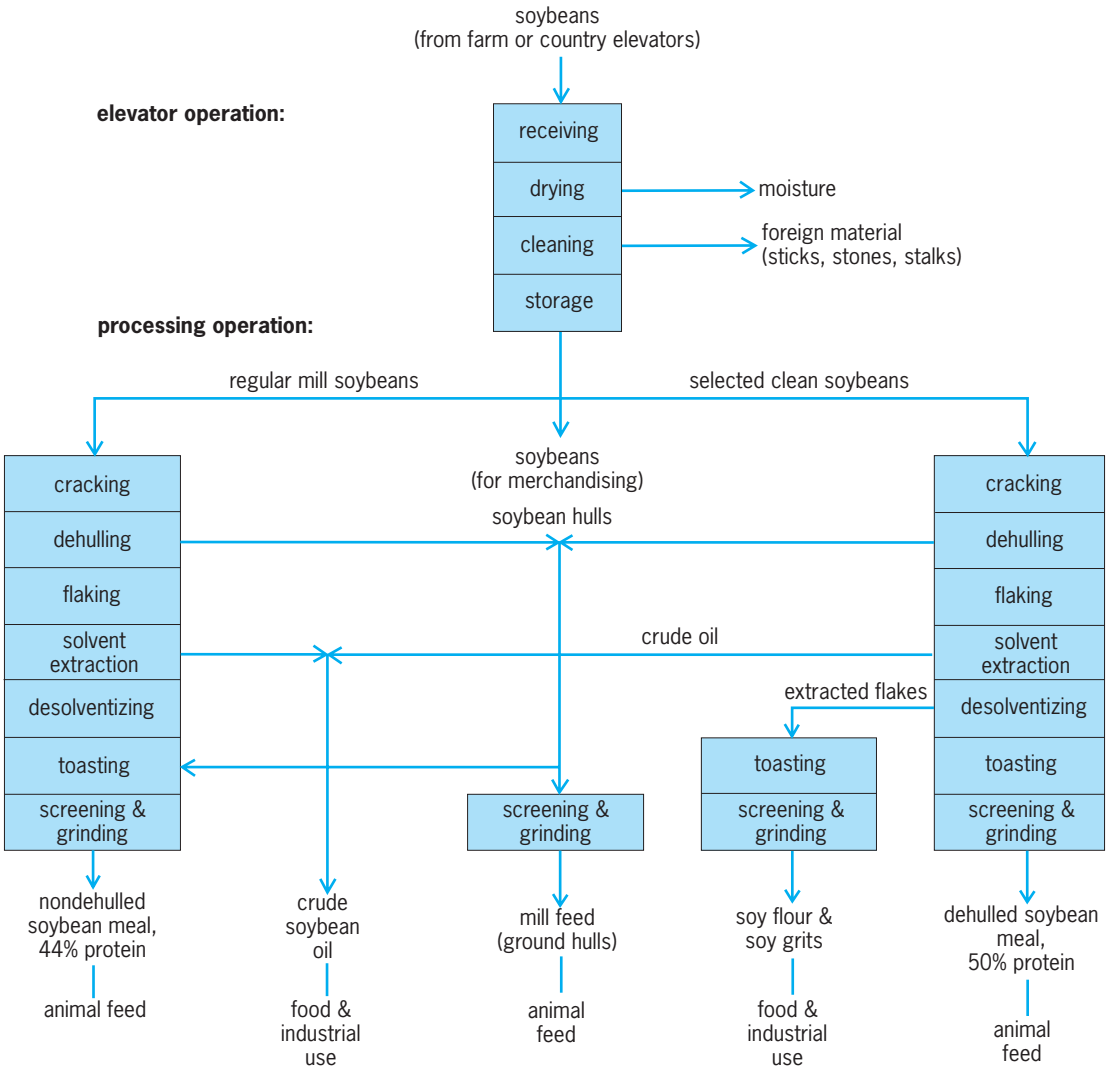

Fig. 3. Steps involved in the solvent extraction method for soybeans.

been processed primarily to obtain oil. The oil was processed for use in shortenings, margarines, and salad dressings. The defatted cake or soybean meal, once a by-product of soybean processing, has become the principal product from soybeans because of its protein value for livestock feeding. Soybean protein products are prepared primarily for direct human consumption. Direct food usage has been limited, but has been expected to increase dramatically because of population pressure on the food supply and limited land and energy resources available for expansion of the livestock industry.

The soybean processing industry converts the oil and meal from the soybean seed into numerous food and industrial products. The products can be generally classified into three categories: soybean oil, soybean meal, and soybean protein products. The processing steps are oil extraction, oil refining, and protein processing.

Oil extraction. The basic processing of soybeans consists of separation of the oil and meal. The equipment used varies considerably from plant to plant although the general process, solvent extraction of the oil, remains the same. Older, obsolete methods used either continuous screw presses or hydraulic presses to squeeze out the oil. Solvent extraction has almost universally been adopted because of its efficiency and low operating cost.

The steps involved in the solvent extraction of soybeans are outlined in **Fig. 3**. Storage of soybeans in elevators or silos is included as part of the processing since it ensures a continuous, year-round supply of beans for processing without the day-to-day price fluctuations of the commodity market. The soybeans, prior to storage, are cleaned and dried to 12–14% moisture for improved stability until processed.

Preparation. The processing operation consists of preparation, extraction, and meal-finishing phases. Preparation for extraction consists of steam-conditioning the beans, cracking, dehulling, and flaking. Cracking the beans into small particles, usually six to eight pieces, loosens the hulls from the soybean meats and assists in the flaking operation. Corrugated rollers are used for the cracking operation. Dehulling or separation of the hull material from the cracked meats is necessary to increase extraction

efficiencies. Dehulling is performed by a combination of air aspiration, screening, centrifugal separation, and density separation methods. Flaking the cracked, dehulled soybean meats facilitates the extraction process by disruption of the internal cellular structure of the bean and by allowing better contact of the extracting solvent with the internal parts of the seed. Flaking is done by heating and steaming the cracked meats and passing them through differential roller mills. The resulting flakes are paper-thin (0.009–0.02 in. or 0.25–0.40 mm).

Extraction. This operation consists of countercurrent washing of the flakes with a petroleum solvent, usually hexane. The solvent is kept near 120°F (50°C) during extraction to speed the solubilization of oil from the flakes. The oil-laden solvent, called micella, is then separated from the flakes and distilled for recovery of the crude soybean oil. The solvent is recovered for further use. The crude soybean oil is then pumped to storage tanks for further processing or to be sold as crude soybean oil. *See* SOLVENT EXTRACTION.

Finishing. The solvent-wet, defatted flakes are desolventized by steam and heat. The severity of the desolventizing depends on the uses for the meal. If the extracted flakes are to be used in animal feeds, moisture is adjusted and the flakes are heat-treated (toasted) to increase their feed value as a protein supplement. Flakes to be used for edible soy flour or protein products are subjected to less severe heating to prevent heat damage to the proteins present. After desolventizing, the flakes are screened and ground. The protein content may also be adjusted. Anticaking agents may be added to ensure ease in handling.

Yields. The processing yields based on a 60-lb (27.2-kg) bushel are 10.2–11.6 lb (4.6–5.3 kg) of crude oil and 48.4–49.8 lb (22–22.6 kg) of meal and feed products. The maximum yield of crude oil is always attempted. The meal from the extraction contains 49–50% protein. The hulls, accounting for about 5% of the beans, are ground and may be added back to the extracted meal for production of a 44% protein meal. The quantities of each meal prepared depend upon the market value at the time of production. The ground hulls may be stored for later use, sold as mill feed to be used as a source of fiber in mixed livestock feed products, or used in a patented process as a carrier for molasses in feed concentrates.

Soybean oil processing. Soybean oil is used in salad dressings, shortenings, and margarines and in various nonfood applications such as paints and varnishes. Most (93%) is used in foods. To be suitable for edible products, crude oil from the extraction requires further processing.

The processing of soybean oil, generally referred to as refining, is outlined in **Fig. 4**. Individual processes are degumming, alkali refining, bleaching (decolorizing), hydrogenation, deodorization, and

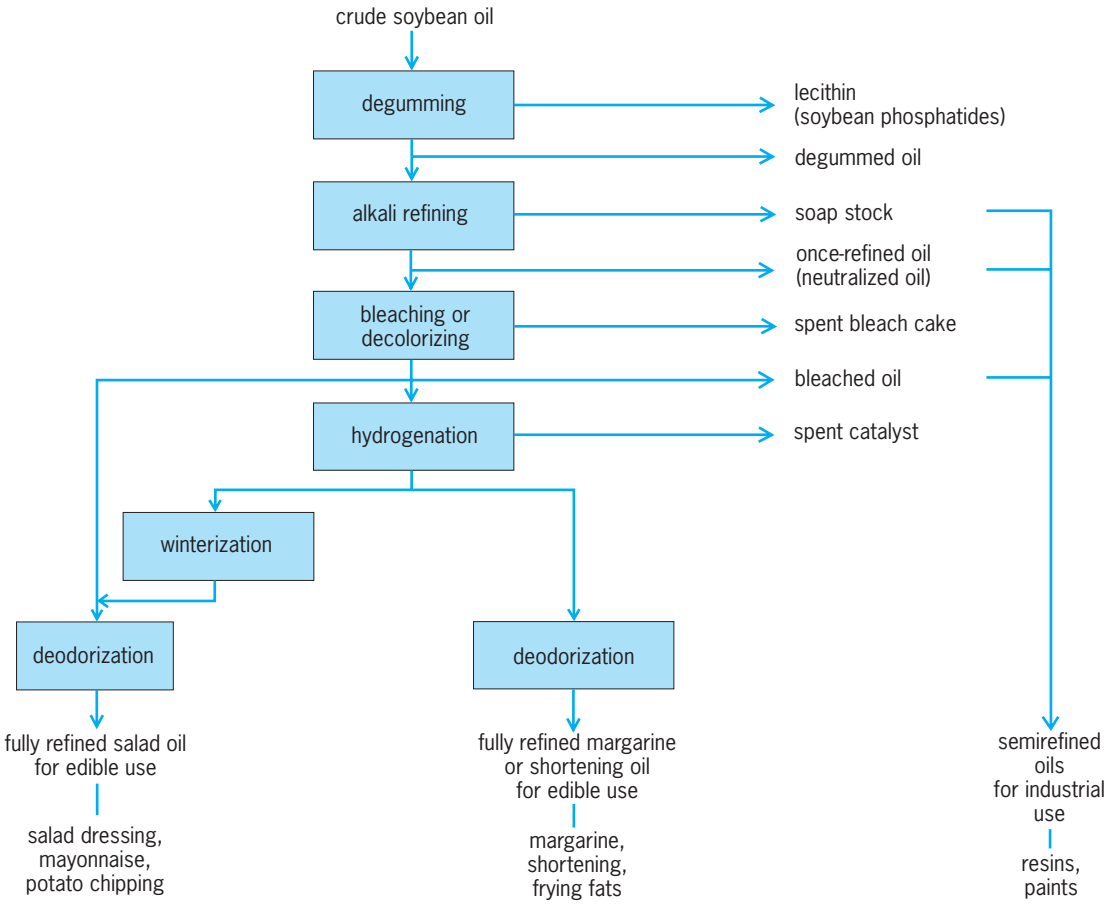

Fig. 4. Steps involved in the soybean oil refining process.

winterization. Semirefined oils are also marketed, depending upon the customer's demands. Hydrogenation is varied, depending upon the final use of the oil. All edible oils are deodorized.

Degumming. Phospholipids, or lecithin, present in the crude soybean oil will precipitate during storage and contribute to deterioration during use of the oil. During degumming, the soybean phospholipids are removed from the crude oil. Degumming itself refers to the gummy consistency of the wet phospholipids. The process involves the addition of a small amount of water or steam to the crude oil, resulting in the hydration of the phospholipids and their subsequent precipitation. Acidic catalysts are sometimes added with the water to improve the hydration. The precipitated phospholipids are then separated by centrifugation and are vacuum-dried. The dried phospholipids are sold as soybean lecithin.

Lecithin. Commercial soybean lecithin is widely used in foods, particularly chocolate, margarine, shortening, and greaseless frying compounds. Soybean lecithin from the degumming step contains 60–65% phospholipids and 30–35% entrained soybean oil. The principal phospholipids present are phosphatidylcholine, phosphatidylethanolamine, and phosphatidylinositol. The crude lecithin is dark-colored and may be bleached with hydrogen peroxide to produce a light, straw-colored product. Fatty acids may be added to make a fluid product, or the lecithin may be treated with ammonia to form a firm, plastic consistency. More refined grades (pharmaceutical lecithin) are produced by acetone extraction of the entrained soybean oil. The phospholipids are insoluble and precipitate. Further processing consists of desolventizing and granulating.

Alkali refining. Free fatty acids decrease the heat stability of the oil. The degummed soybean oil is alkali-refined to remove the free fatty acids, residual phospholipids, and color. Alkali refining consists of addition of caustic soda, soda ash, or a combination of the two to neutralize the free fatty acids. Careful addition of the alkali is required to prevent further degradation of the oil. The neutralized free fatty acids or soaps are removed from the oil by centrifugation or settling. The soaps, referred to as soapstock, are sold to soap manufacturers, fatty acid industries, or livestock feed formulators in the alkaline form, or in an acidulated form after treatment with sulfuric acid. The alkali-refined oil is further washed with water to remove all traces of alkali and soaps present and then is dried. The neutralized oil is often sold to industrial users for nonfood applications.

Bleaching. Bleaching (decolorizing) removes the color bodies or pigments present in the oil by adsorption on activated earth or carbon. The bleaching agent is simply dispersed in the oil and removed by filtration. The bleached oil is sent to the hydrogenation unit or to the deodorizers, depending upon the final use. Foods requiring improved oxidative stability or plastic fats such as margarines or bakery shortenings require hydrogenation.

Hydrogenation. Development of the hydrogenation process for vegetable oils has permitted the substitution of vegetable oils for animal-derived fats and the substitution of the various types of vegetable oils with soybean oil. Variations in the hydrogenation conditions can be used to produce high-stability soybean oil, hardened soybean oils that melt rapidly at body temperature, and shortening-like fats.

Hydrogenation is performed by purging the oil with hydrogen gas at an elevated temperature and pressure in the presence of a catalyst. Nickel catalysts are generally used. The addition of hydrogen to the fat saturates the molecular structure. The degree of saturation depends upon the time, temperature, and pressure of hydrogenation. *See* HYDROGENATION.

Deodorizing. Deodorization of the oil removes the undesirable flavor and odor constituents. Oils must be odorless and bland in flavor to be suitable for food use. Deodorization is carried out under high temperature and vacuum with steam injected to assist volatization of the undesirable components.

Winterization. Winterization is performed on lightly hydrogenated soybean oil that is to be used as a salad oil. The oil, in addition to conforming to oxidative stability requirements, must remain clear at refrigerator temperatures. Winterization consists of cooling of the oil and filtering off the cloudy haze that forms.

Reversion. Refined soybean oil is a highly unsaturated oil and will develop, after storage, characteristic flavors and odors described as painty or fishy. The development of these flavors is called reversion, and is possibly a result of the oxidation of the linolenic acid component of the oil. Care in the exposure of the oil to air during processing and controlled hydrogenation to partially saturate the linolenic acid minimizes the tendency of soybean oil to reversion. Metal sequestrants, such as citric acid, are sometimes added during deodorization to inactivate the prooxidant metals such as iron and copper. Antioxidants may also be added. *See* ANTIOXIDANT; FAT AND OIL (FOOD).

Soybean protein products. The soybean protein products are prepared from the extracted meal. Three distinct product types are prepared: soy flour and grits, soy protein concentrates, and soy protein isolates. These are used both in foods and in nonfood applications.

Flour and grits. Soy flour and soy grits contain approximately 50% protein based on the nonlipid fraction. Flours and grits are used in processed meats, bakery mixes, cereals, breads, and baby foods. A large amount has been utilized in nutrients for export to undeveloped, underfed countries.

The term soy flour refers to an extremely fine powdered product, whereas the term grits refers to a granular product. Various products are prepared varying in oil content, oil type, particle size, and water solubility of the protein. In general, the oil content varies from less than 1% (defatted) to 18–20% (full fat); the oil source is soybean oil or lecithin. Particle size varies from powders that pass a 200-mesh screen (flour) to those that pass a 14-mesh (grits) one. Water solubility of the protein ranges from 10% (toasted) to 80% (untoasted).

Soy flour or grits is produced similarly to soybean

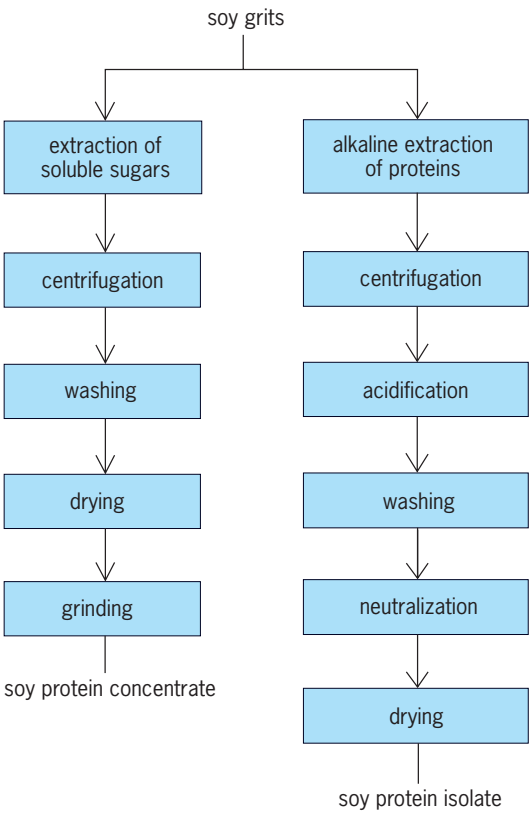

Fig. 5. Production of soy protein concentrates and isolates.

meal, except with thorough removal of the hulls before extraction. Desolventizing and toasting are varied to alter the water solubility of the protein, with minimum heating being required for high protein solubility. The defatted, desolventized flakes are ground and screened to the desired particle size.

The oil-containing products are generally prepared by blending soybean oil or lecithin back to the extracted meals. Full-fat soy flour may also be prepared by steaming the unextracted, dehulled flakes to remove the bitter flavors, and then drying, grinding, and screening them to the desired particle size.

Concentrates. Soybean protein concentrates and soybean protein isolates are generally prepared from soy flour or grits. The processing methods vary considerably between processors. A generalized process is shown in **Fig. 5**.

Soybean protein concentrates contain 70% protein. Concentrates are prepared by washing the water-soluble carbohydrates from the meal under conditions designed to ensure minimum protein solubility. Dilute acids or aqueous alcohol solutions have been used. The washing is performed in agitated tanks. The remaining insoluble proteins and carbohydrates are removed from the extraction by centrifugation or filtration, washed, dried, and ground. Products are marketed with ranges of particle sizes and protein solubilities similar to soy flour and grits. Soy protein concentrates are usually bland in flavor, possess a lighter color, and are more functional than soy flour (having emulsification, water-binding, gelling, and other properties).

Isolates. Soy protein isolates are the purest form of soy proteins marketed (90% protein minimum). The isolates are prepared by solubilizing the protein present in the grits by using mildly alkaline solutions, separating the protein solution from the fibrous residue by centrifugation or filtration, and acidifying the protein solution to precipitate the protein as a curd. The curd is recovered, washed, resolubilized, and spray-dried into a fine, flourlike powder.

Soy protein isolates are manufactured to possess specific functional properties such as gelling, whipping, foaming, emulsifying, thickening, and binding. Isolates are used in meats, simulated dairy products, and bakery products.

Textured proteins. Each of the soybean protein products can be used to form textured soy proteins. Textured soy proteins have a meatlike texture and appearance when hydrated and can be used for extension of ground and comminuted meats (such as ground meat and bologna) and for the preparation of simulated meats. When they are used as extenders, the textured proteins increase cook yield, aid in binding of meat constituents, and improve emulsion stability.

Two general processes (with variations) are used for texturizing soy proteins: extrusion cooking and spinning. Extrusion cooking consists of moistening the protein, working the doughlike mass under high temperature and pressure, and forcing the cooked mass through an orifice, which permits the heated compressed mass to expand. The product is then dried and sized. Colored and flavored extruded soy protein products are prepared by addition of these additives either before or after extrusion. Imitation bacon bits that are based on extruded soy flour are widely marketed.

Soy protein isolates are required for spinning. Spun proteins are somewhat more adaptable to simulated meat production than extruded proteins. Spun proteins are produced by solubilizing the isolates in an alkaline solution and forcing this solution through a multiholed die immersed in an acid bath. The solubilized protein immediately coagulates in the acid bath, forming a continuous filament. The filaments are stretched to orient the protein molecules within the filament. Toughness of the filaments is dependent upon the degree of protein orientation. Filament diameters are approximately 0.003 in. (0.076 mm). Groups of filaments when combined with egg white, flavors, and colors, followed by compression and heat setting, form imitation meats such as ham, chicken, or beef. *See* FOOD MANUFACTURING.

F. T. Orthoefer

Bibliography. D. Erickson et al., *Handbook of Soy Oil Processing and Utilization*, 1980; W. O. Scott and S. R. Aldrich, *Modern Soybean Production*, 2d ed., 1983; J. B. Sinclair, *Compendium of Soybean Diseases*, 4th ed., 1999; S. R. Singh et al. (eds.), *Soybeans for the Tropics: Research, Production, and Utilization*, 1988; H. E. Snyder and T. W. Kwon, *Soybean Utilization*, 1987; D. Swern (ed.), *Bailey's Industrial Oil and Fat Products*, 2 vols., 4th ed., 1986.

Space

Physically, space is that property of the universe associated with extension in three mutually perpendicular directions. Space, from a newtonian point of view, may contain matter, but space exists apart from matter. Through usage, the term space has come to mean generally outer space or the region beyond the Earth. Geophysically, space is that portion of the universe beyond the immediate influence of Earth and its atmosphere. From the point of view of flight, space is that region in which a vehicle cannot obtain oxygen for its engines or rely upon an atmospheric gas for support (either by buoyancy or by aerodynamic effects). Astronomically, space is a part of the space-time continuum by which all events are uniquely located. Relativistically, the space and time variables of uniformly moving (inertial) reference systems are connected by the Lorentz transformations. Gravitationally, one characteristic of space is that all bodies undergo the same acceleration in a gravitational field and therefore that inertial forces are equivalent to gravitational forces. Perceptually, space is sensed indirectly by the objects and events within it. Thus, a survey of space is more a survey of its contents. *See* EUCLIDEAN GEOMETRY; LORENTZ TRANSFORMATIONS; RELATIVITY; SPACE-TIME.

The space around the Earth and extending out perhaps 10 earth radii (40,000 mi or 64,000 km) has properties which differ from those of interplanetary space. The chief reasons for this difference are the existence of the gravitational and magnetic fields of Earth. These fields affect the densities and motions of neutral and charged particles in the vicinity of Earth. The corona of the Sun, which was once believed to be a static and limited atmosphere, fills the entire solar system, although in a highly tenuous form. This corona, in the form of a plasma of hydrogen particles, is pervaded by a magnetic field, and it flows past the Earth at speeds as high as 250 mi/s (400 km/s). *See* VAN ALLEN RADIATION.

The geomagnetic field of the Earth, unlike an undistorted dipole field, is confined to a limited region of space, as data supplied by satellite measurements suggest. The actual field is distorted by radiation belts and primarily by its interaction with the solar wind. This volume of space is called the magnetosphere, and is bounded by a thin layer called the magnetopause that fluctuates in response to the solar wind. The magnetosphere appears to have the shape of an elongated comet whose tail points away from the Sun.

Between 10 to 20 earth radii, a shock wave is produced by the flow of the solar wind around the magnetosphere.

The gravitational field allows Earth to retain its atmosphere but is not strong enough to prevent escape of the atmosphere completely. Above a level of about 310 mi (500 km), where the atmosphere is so rarefied that collisions between atoms can be neglected, the atoms describe ballistic orbits and form the exosphere. Light atoms, especially hydrogen, but also helium, occasionally obtain enough energy so that they escape completely from Earth's gravitational field. *See* ATMOSPHERE; MESOSPHERE.

Earth's magnetic field deflects moving charged particles. Particles of the highest energy, namely cosmic rays, are deflected by the magnetic field so that only those having an energy of more than 15 GeV can enter near the Equator. At the poles, however, cosmic rays of all energies may enter because there the lines of force of the magnetic field are vertical. As a result, cosmic ray intensity at the pole is 10–50 times higher than at the Equator. *See* GEOMAGNETISM.

Solar corpuscular radiation, containing particles of much lower energies, is affected more strongly by Earth's magnetic field and is deviated into zones at high latitudes. Many of these solar particles are trapped in Earth's magnetic field and remain in it for long periods of time, moving back and forth along a line of force, but being reflected from each end. In this process particles may be accelerated to higher energies and then be energetic enough to produce luminous displays in the upper atmosphere. In their trapped condition these particles contribute to Earth's radiation belt, first observed in early Earth satellite experiments. Another source for the radiation belt comes from cosmic rays which plunge into Earth's atmosphere, disintegrate atmospheric nuclei, and throw back into space debris which can be trapped. *See* ASTRONOMY; AURORA; COSMIC RAYS.

S. F. Singer

Space biology

Collectively, the various biological sciences concerned with the study of organisms and their components when exposed to the environment of space (or of a spacecraft beyond the immediate influence of Earth and its atmosphere). Space biology studies include all the biological science disciplines, such as molecular, cellular, developmental, environmental, behavioral, and radiation biology; physiology; medicine; biophysics; and biochemistry. Space biology encompasses both basic and applied sciences, and focuses on using space research to understand and resolve biological problems on Earth, as well as alleviating or preventing the deleterious physiological changes associated with space flight or habitation in non-Earth environments. *See* SPACE.

As crewed space flight developed, medical research focused on the health and well-being of astronauts was paramount. This medical research has historically been included in the area of aerospace medicine. From these research efforts, directed at survival of living systems in the space environment, has evolved the broader discipline of space biology, which uses the space environment to gain an understanding of basic biological processes. *See* AEROSPACE MEDICINE.

Space biology considers all of the inherent physical factors of the space environment and their effects on living organisms: (1) factors due to the environment of space, such as radiation and reduced pressure; (2) factors due to the flight dynamics of

spacecraft, including acceleration, weightlessness, and vibration; (3) factors due to the environment of the spacecraft, such as air composition, atmospheric pressure, toxic materials, noise, and confinement; and (4) factors due to the environment on nonterrestrial bodies, such as the level of gravity, composition of soil, and temperature extremes. Of these factors, only weightlessness and space radiation are unique to the space environment, but the effects of the other factors must be considered and mimicked in ground-based or flight controls when conducting space biological experiments. In addition, space experiments must take into consideration the physical effects of weightlessness on the environment surrounding and supporting the organisms. For example, neither gases nor liquids behave in a space environment as they do under the conditions found on Earth, as sedimentation, buoyancy, and convection do not occur in a weightless environment. *See* ACCELERATION; WEIGHTLESSNESS.

Space Gravitational Biology

Among the effects of space flight, weightlessness (free fall) associated with orbital flight presents the greatest scientific opportunity and poses the greatest challenge. All life has evolved under the most constant and ubiquitous of the Earth's environmental factors, gravity. The near-weightless (microgravity) environment of space flight provides the means to manipulate gravity below 1 *g* in order to understand its effects on biological systems. Studies in microgravity provide information as to the influence of gravity on living systems, and insights into possible life in environments beyond Earth.

Periods of near-weightlessness, ranging from a few seconds to about 15 minutes, can be achieved by using drop towers, parabolic airplane flights, and sounding rockets. To extend the exposure time period to days, weeks, and months requires orbital space flight. Biological specimens have been sent into space in a variety of satellites and crewed space ships (**Figs. 1** and **2**). The collection of biological data in space has been complicated by a number of limiting and confounding factors: few flight opportunities, short-duration space flights, adverse spacecraft environmental conditions, complicating spacecraft flight dynamics, accommodations for only a few species, limited means to manipulate specimens in space, and the general lack of a 1-*g* reference control centrifuge on board the spacecraft. Yet many significant observations have been made that establish a foundation for future space research.

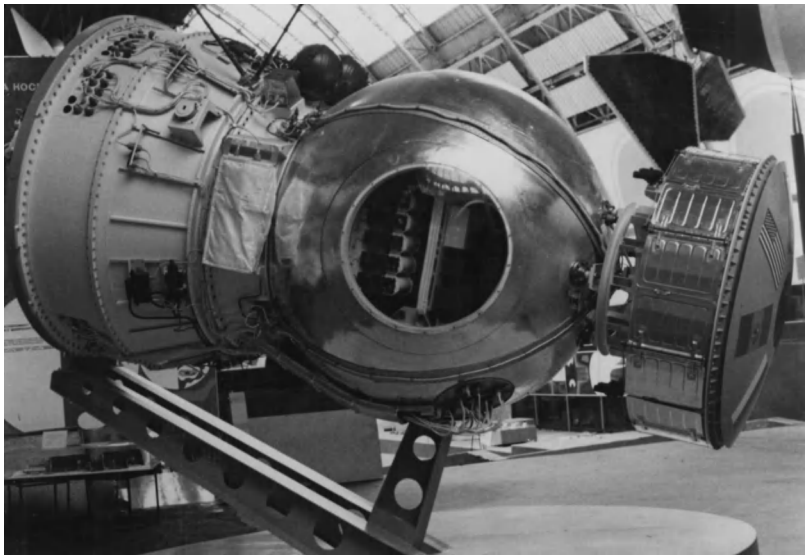

Fig. 1. ***Cosmos 782*****, one of a series of Soviet and Russian spacecraft that carried biological experiments. Launched on November 25, 1975, on a 20-day mission, it carried more than 20 different species. Scientists from France, Czechoslovakia, Hungary, Poland, Romania, the United States, and the Soviet Union participated in these investigations.**

Ground-based research. A variety of ground-based experimental methods have been used successfully to complement space experiments. Each method depends upon changing the force imposed on part of an organism or the entire organism—for example, by using centrifugation to increase the level of gravity or changing the gravitational vector by clinostating (randomly rotating). These methods are restricted by the untoward additional biological effects they create. Organs and tissues such as bone and muscle that have evolved to support animals against the Earth's gravitational force have been the most responsive to ground-based modeling. Rats exposed to space flight and those subjected to 1-*g* and 2-*g* centrifugation showed that total body calcium was decreased by reduced gravitational loads and increased by elevated gravitational loads. Computerized analysis of the morphology of the rat inner-ear vestibular system, crucial for orientation and balance, has shown that the number of nerve cells innervating these organs is also gravity-dependent. There are more nerve cells in space-flight-exposed tissue, for which the gravitational stimulus has been reduced, while the number of cells is decreased when the gravitational load is increased to 2 *g*. Further, bone cells unloaded by rotation demonstrate differential responsiveness to growth mediators compared to static cultures. Imposition of a load to the cells results in cell-type differentiation. During centrifugation, the increase in the level of gravity produces distinct differences in the behavior of rodents. Activity is reduced, and there are shifts in circadian rhythms opposite to those observed during space flight. Ground-based research using these types of experimental models provides direction for and insights into space flight experiments.

Effects on unicellular organisms. Space experiments have been conducted with a variety of unicellular organisms, including bacteria, fungi, protozoa, and algae. Developmental, physiological, and behavioral effects have been noted. In prokaryotic cells (cells without nuclei, such as bacteria), changes in metabolism have been repeatedly documented. Bacteria grow two to three times more rapidly in microgravity, are more resistant to antibiotics, and have more complete exchange of nuclear material during conjugation than on the ground. Decreased sporulation is observed, suggesting a change in bacterial differentiation in microgravity. While no differences are

found in the genetic recombination mechanisms of transformation or transduction of *Escherichia coli*, conjugation is increased three- to fourfold in space flight. *See* BACTERIAL GENETICS; BACTERIAL PHYSIOLOGY AND METABOLISM; TRANSDUCTION (BACTERIA); TRANSFORMATION (BACTERIA).

Under microgravity conditions the eukaryotic single-celled organism *Paramecium tetraurelia* displays an increased growth rate accompanied by increased cell volume and decreased dry weight, protein content, and calcium. There are possible changes in cell membrane assembly and extrusion of trichocysts as well. The alga *Chlorella vulgaris* shows changes in the ultrastructure of the cells, and rearrangement of organelles. The degree and spectrum of change are correlated with the duration of the space flight. Mitochondrial density and volume and the number of cristae in plastids are the first to undergo change. *Chlamydomonas* cultures show increased cell size and a prolongation of the active growth phase. The slime mold *Physarum polycephalum* demonstrates that both light and gravity regulate the contraction frequency of plasmodial strands. The results implicate the mitochondria, since they play a major role in changing the contraction rhythm, and suggest that the mitochondria not only supply most of the energy necessary for the contraction but also may be involved in its regulation and possibly even in the perception of the gravity stimulus. *See* CELL ORGANIZATION; CELL PLASTIDS; MITOCHONDRIA.

Effects on animals. Physiological and developmental effects of microgravity have been observed in many types of animals.

Physiological effects. Bone in growing rats is subtly altered in space flights lasting up to 21 days. The mineral concentration, size, and bone resorption are minimally changed, while bone formation, mineralization, and maturation are suppressed. While the bone increases in size in growing rats during flight, it has reduced strength or resistance to breaking. Bone vascularity, nerves, ligaments, tendons, and vertebral discs of young rats all respond in some way to space flight.

Skeletal muscles that normally function to overcome gravity atrophy during space flight. Although muscle mass changes occur quickly, the underlying physiological changes may continue even after the mass achieves maximum loss. There is a decrease in the cross-sectional area of muscle fibers (primarily of the slow-twitch oxidative fibers), associated with a substantial loss of muscle protein, especially myofibril proteins; a decrease in alpha-actin mRNA; increased glucose uptake in the presence of insulin; and increased activities of glycolytic enzymes; while oxidation of fatty acids appears to be reduced. In cardiac muscle, mitochondrial numbers decrease, and increased lipid and glycogen content indicate a reduced work capacity. Cardiac muscle cells appear deficient in their ability to utilize energy, based on adenosine triphosphate (ATP) content and fatty acid metabolism. Regulatory proteins in the ventricle show reduced responsiveness to hormone stimulation, and heart cells show mild degradation. *See* CARDIOVASCULAR SYSTEM; SKELETAL SYSTEM.

Fig. 2. Payload specialist Mamoru Mohri using a microscope to produce photomicrographs of mammalian cells during the Spacelab SL-J mission, a joint effort of Japan and the United States aboard flight *STS 47* of the space shuttle *Endeavour* in September 1992. The biopack contains a health-monitoring experiment. (*NASA*)

In rats flown in space, there is a reduction in the volume of red blood cells. Hemolysis of red blood cells is about three times greater in rats exposed to space flight than in control rats, and the mean life-span of these cells is decreased by about 4 days during flight. This reduction is in part due to a decrease in plasma levels of erythopoietin, reducing production of red blood cells. There is a reduced immune response to mitogens. Livers from rats exposed to weightlessness contain up to 100% more glycogen than those from control rats, and bone and other organs also show increases in lipid content. Analysis of the body constituents of space-flight-exposed young rats shows an increase in potassium, creatine, nitrogen, and magnesium and a decrease in calcium, phosphorus, and sodium when compared to 1-*g* controls on Earth. Conversely, rats subjected to an increase in gravity to 2 *g* by centrifugation exhibit a decrease in potassium, creatine, and nitrogen and an increase in calcium, phosphorus, magnesium, and sodium. Pituitary oxytocin and vasopressin are decreased, suggesting in-flight adjustments in the regulation of body water content. Rats, like humans, appear dehydrated after flight, with changes in water-salt metabolism. The extent of this water loss appears to be related to the time after landing at which the measurements are made. The longer the time after landing, the greater the loss of body water. This is corrected within 3–5 days after return to Earth. Upon return to Earth, animals have a pronounced increase in urine output, further supporting a reduction in regulation of water and electrolyte homeostasis.

Decreased levels of growth hormone are related to increased concentration of growth hormone in pituitary cells; these cells release less growth hormone when cultured postflight, suggesting a secretory dysfunction. These changes in growth hormone

also appear to be associated with a reduction in translation, as adequate amounts of mRNA are present. Most changes induced by space flight are corrected within 1 month of return to Earth, and an artificial gravity level of 1 *g* during flight appears to normalize many flight-induced effects. *See* ENDOCRINE MECHANISMS.

In rhesus monkeys, cardiovascular measurements taken during space flights indicate a decrease in heart rate, increase in blood pressure, and reciprocal changes in blood flow velocity associated with adaptation to space flight. Opposite responses are noted with exposure to increased gravity produced by centrifugation. After space flight, studies of muscle function and morphology show a shift in muscle typing from slow twitch to fast twitch. This is predominantly noted in those muscles used to support and move the body against gravity. In addition, the function of the muscle during movement is altered. Circadian rhythms of activity and body temperature persist during space flight, but both regulation and circadian timing of these systems are significantly influenced by microgravity. At the same time, heart rate maintains a normal 24-hour rhythm, indicating it is unaffected by microgravity. Studies of vestibular function suggest that acclimation to microgravity causes changes in the way that body orientation information is processed in the nervous system. These changes in vestibular function, body orientation, and muscle morphology and function contribute to altered posture and movement upon return to Earth. There are significant alterations in walking and running gait. Fine motor control also appears to be affected. These changes are similar to those reported in astronauts.

Reproduction and development. The experimental difficulty in initiating fertilization in space and the short duration of most space flights have been limiting factors in research on animal reproduction and development in space. The scientific questions fall in two main categories: (1) the effects on fertilization and early development, and (2) the effects on the different organ systems, since each has a critical development period that is especially sensitive to environmental perturbation. The majority of the research has focused on fertilization and early development, and in general has found that microgravity has minimal effect on the animals tested.

Normal fertilization and development of fruit flies can occur in microgravity. However, larval hatching and the number of adults emerging after recovery are significantly reduced, oogenesis is affected, and there is a slight increase in abnormal head morphology. Stick insects also show a decrease in hatching with apparent normal development. The developmental process in polyps and ephyrae of jellyfish (*Aurelia aurita*) appears to be normal. However, their pulsing movement is abnormal and may reflect abnormal development of the graviceptors, the nerve net, or the neuromuscular system. Successful fertilization of the sea urchins *Paracentrotus lividus* and *Lytechinus pictus* has been achieved; however, significant abnormalities were noted in the centrosome-centriole complex of dividing cells. These abnormalities suggest alterations in developmental processes in microgravity. The nematode *Caenorhabditis elegans* exhibits normal development and chromosome mechanics.

The fish *Fundlulsu* reproduces and develops normally, and gravid guppies give birth to normal young when returned to Earth. Female frogs of the species *Xenopus laevis* have been stimulated to ovulate in space, and their eggs have been fertilized in microgravity. Exposure to microgravity from oocyte maturation through hatching does not deleteriously affect amphibian morphogenesis or permanently alter behavior. The duration of the flights with developing frogs has been short, and some developmental differences have been noted that may affect morphogenesis. For example, the lungs of tadpoles reared in space are not inflated, either because they are not needed for buoyancy or because the young cannot break through the air-water interface to inflate them. Upon return to Earth, tadpoles fail to swim appropriately, possibly due to the difference in buoyancy.

Apparently normal chicks have hatched from fertilized Japanese quail eggs (*Coturnix coturnix japonica*) incubated in space. They react to visual and auditory stimuli and display normal motor activity but have orientation problems. Early chicken embryos have died during space flight following early preincubation on the ground. The deaths have been suggested to result from the absence of buoyancy during space flight. The yolk normally floats at the top of the egg, allowing the embryo close contact with the shell for easy diffusion of oxygen. During space flight the yolk does not float to the top of the egg, and thus the embryo may be starved for oxygen and die. Later in development, surviving chicks from embryos exposed to space flight show decreased vestibular sensitivity to gravitational stimuli.

In rats, early development during gestation does not appear to be affected by microgravity. Pregnant rats launched into space early in gestation deliver physiologically and behaviorally normal rat pups upon return to Earth. However, maternal differences occur that could affect the young. The metabolism of the mammary glands of rats flown in space is reduced and could result in a decrease in milk production. Young rats flown on the space shuttle prior to weaning do show deleterious effect. If they are too young, mortality is very high, as the microgravity environment affects behavioral interactions between the mother and neonate. Animals that survive appear to have neurological deficiencies that persist into later life when returned to Earth.

Effects on cultured animal cells. Most mammalian cells grown in microgravity (Fig. 2) show a reduced growth rate compared to ground controls. Reduced proliferation is observed with cultured human lung cells, monkey kidney cells, mouse hybridoma cells, and human lymphocytes. The rate of glucose utilization may be reduced, and there may be changes in membrane structure or receptors. Culture experiments with cells that are dependent upon attachment to a substrate are compromised in microgravity,

where all objects tend to float. Culturing of lymphocytes in microgravity is not similarly impacted because the cells are capable of autonomous motility. Immune-cell cultures in space produce interferon more rapidly and with four to eight times higher titers. Human mononuclear cells grown in space lose 50–90% of their ability to respond to stimulation with the mitogen concanavalin A, compared to control cells grown on Earth or in 1-*g* centrifuge in space. Immune cells also display major depression of blast cell transformation in response to mitogens, loss of cytokine production or function, and changes in natural killer cell activity and response to colony-stimulating factor. Changes in cellular function are associated with alteration of gene levels. Human kidney cells grown in space show alterations in signaling pathways, and structures within the cell possibly contribute to functional changes. *See* CELLULAR IMMUNOLOGY.

Effects on plants. Plants were the first organisms to demonstrate a biological role for gravity. Experimental interest centered on gravitropism, the differential growth induced by gravity that causes shoots to grow up and roots to grow down. This process allows plants to correct for changes in direction imposed by the environment, such as roots growing around rocks. *See* PLANT MOVEMENTS.

Although seeds and plants have been flown in space for over 40 years, experiments have been infrequent and experimental conditions limiting. Major environmental influences due to the physical properties of the space flight environment have been identified. Successes of early plant experiments were influenced by the physical properties of water in the microgravity environment. For example, free-floating water forms into a ball, while contact with a surface causes the water to form a film. These properties led to inappropriate nutrient delivery. Airflow was also limited in early flights due to the lack of convection, which allowed the creation of boundaries around leaves limiting nutrient delivery and increased concentrations of waste products. These problems were overcome by engineering interventions in the design of subsequent habitats to grow plants. Results of experiments on plants in space suggest an extensive direct influence of gravity on all aspects of plant growth, development, and metabolism.

Physiology, orientation, and tropisms. Plants are exquisitely sensitive to gravity. Graviperception (the magnitude of acceleration required to induce a directed growth response) has been calculated at approximately 10^{-3} *g* in shoots and 10^{-4} *g* in roots in flight experiments.

Orientation of roots and shoots (**Fig. 3**) is directed by external stimuli. In microgravity, light directs the orientation of wheat and pea shoots while roots remain disoriented. These studies suggest that, at least in these species, gravity directs the orientation of root growth, while gravity and light guide shoot orientation. Both circumnutation (a gravity-induced touch response) and epinastic movements (bending of leaves downward due to growth of their upper surface) occur in microgravity. *See* PLANT GROWTH.

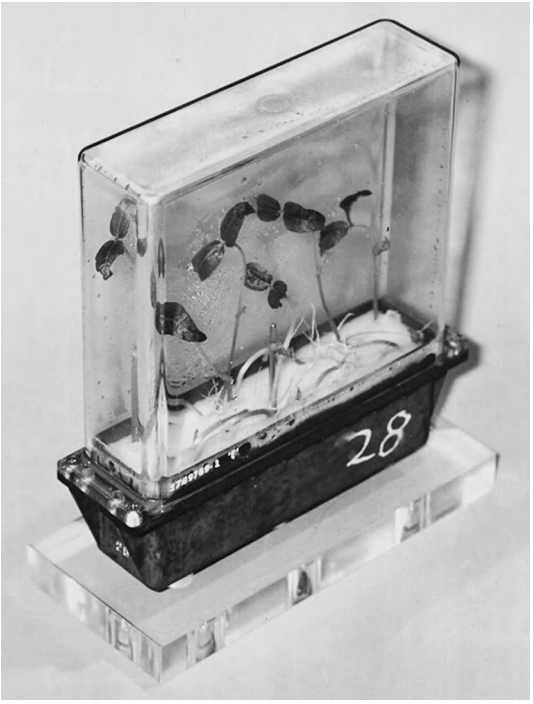

Fig. 3. Shoots in space.

Flight experiments indicate that microgravity affects a plant's mineral balance. It has been postulated that cytoplasmic calcium concentration increases, calcium ion adenosine triphosphatase activity is altered, and ion intake by roots is disrupted in microgravity. In fact, calcium has been linked to both the mechanisms of gravitropism and the mitotic defects seen in seedlings grown in microgravity.

Reproduction and development. Space flight does not affect germination, and its effects on growth are unclear. Generally, by gross observation plants appear to grow normally for short periods of time in microgravity. Over extended periods of microgravity, plants slow or stop growing and then die, most often at the flowering stage. Success was finally achieved with small, fast-growing *Arabidopsis* seeds which were sown in space, completed a full life cycle, and produced fertile seeds in 1982. Growth was not vigorous, few seeds were produced, and the seeds had a significant level of abnormalities. Many of these observations were demonstrated to be the product of the physical properties of the seeds' space environment. Stagnant air layers that form in the absence of convection during microgravity stress the plants by producing localized high humidity, diminished availability of metabolic gases, buildup of harmful gases, and high temperatures at the top of the plant near light sources. Implementation of protocols to compensate for these properties has greatly improved the reproductive development and germination of *Arabidopsis* such that, in every aspect examined, values were similar to those of ground control plants.

Cell division is generally reduced or inhibited in microgravity. In fact, in one space shuttle experiment, oat seedlings had only one-tenth as many

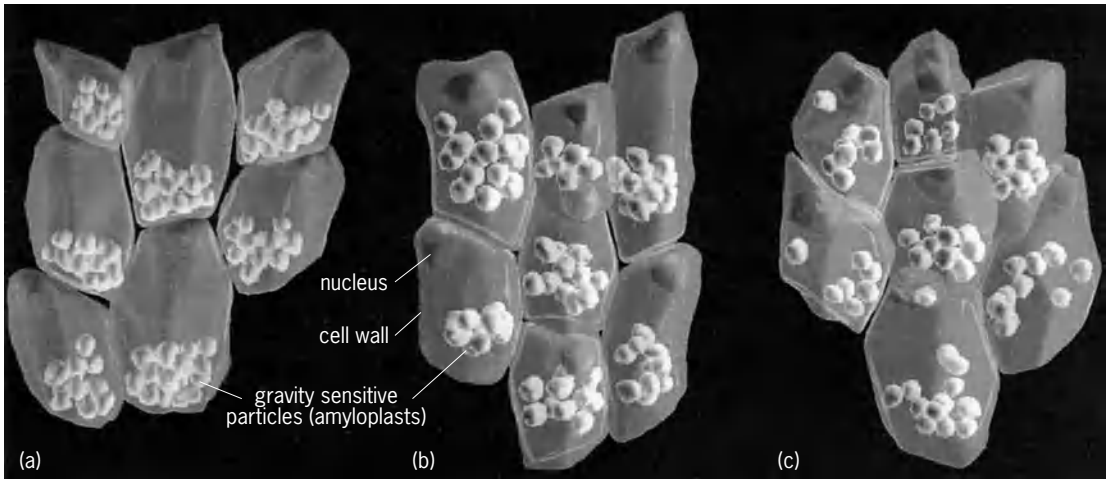

Fig. 4. Groups of reconstructed gravity-sensing cells from the root tips of 3-day clover seedlings. (*a*) Cells grown on Earth. Gravity-sensitive particles (amyloplasts) are sedimented to the bottom of the cell. (*b*) Cells grown in space (microgravity; no gravity direction), with larger particles that display grouping. (*c*) Cells grown in a clinostat where they are subjected to slow tumbling (gravity in all directions), with particles that are fewer in number and dispersed, showing that slow tumbling does not mimic the effect of space flight on plants. Particle size and number respond to gravity sensitivity, while particle distribution responds to biochemical signals.

dividing cells as ground controls. Carrot and day lily plants grown in microgravity from somatic cells develop a substantial number of binucleate cells, while the ground control cells are uniformly uninucleate. This may be due to a defect in the ability to put down wall material. Chromosomal perturbations, including breakage, translocations, and deletions, are seen in a variety of microgravity-grown plants derived from somatic embryos, tissue-culture plantlets, and seeds. Dosimetry data collected in conjunction with some of these experiments have ruled out radiation as a cause of the damage, and in-flight fixation of the plants eliminates reentry and 1-*g* ground recovery perturbations as possible causes of the observed chromosomal damage. Microgravity caused disturbances in the mitotic process. Abnormalities are believed to be associated with spindle disturbances, although this has yet to be demonstrated. *See* CELL DIVISION; CHROMOSOME ABERRATION; MITOSIS.

The pattern and rate of cellular differentation may be altered by microgravity. Undifferentiated clusters of anise cells differentiate more rapidly in space. Carrot and lettuce embryo roots and grape hyacinth gametophytes show accelerated development as well; and roots of lettuce, *Crepis*, and oat seedlings differentiate prematurely. This could indicate that microgravity accelerates organ formation and development. *See* CELL DIFFERENTIATION.

Cellular effects. The size of plant cells may be increased, decreased, or normal in microgravity. Rounded cells are seen in space-grown shoots of pine and maize and the roots of maize and *Crepis*, possibly due to the increased effect of surface tension and reduced cell-to-cell mechanical pressure in microgravity. Root caps do not regenerate on some plants in microgravity; however, this failure could be associated with the environment of their growth chamber.

Growth on Earth is responsible for the arrangement of at least some cell organelles, for in microgravity the distribution of not only amyloplasts but also endoplasmic reticulum is altered. Amyloplasts in root caps of plants grown in microgravity (**Fig. 4**) are always distributed randomly. Moreover, their structure and starch content are changed. Chloroplasts of space-grown pea plants show morphological changes. *See* ENDOPLASMIC RETICULUM.

Nuclei of space-grown pea and *Arabidopsis* root cells show abnormal distribution and increased amounts of condensed chromatin, which correlates well with the general observation of decreased mitosis and increased chromosomal aberrations in cells grown in microgravity. This suggests reduced functional activity that could indicate an effect on protein synthesis. *See* CELL NUCLEUS.

Mitochondria of microgravity-grown plants are characteristically swollen. This would suggest an effect of microgravity on cellular metabolism. Microgravity-grown seedling root cells contain 50–90% less dictyosomal volume than cells of ground control plants. This correlates with the thinning of cell walls in space.

It has been proposed that cellulose and lignin should be reduced in space because they function in a structural capacity that is unneeded there. Flight experiments have supported this assumption. *See* CELL (BIOLOGY); SPACE FLIGHT.

Ionizing Radiation

The physical environment of space includes a broad spectrum of ionizing radiation, ranging from the ultraviolet to galactic cosmic rays, that is able to penetrate the life-supporting enclosures in which organisms live (habitats, spacecraft, and space suits). In addition, biological specimens are exposed to products derived from interactions of radiation with other biological systems and materials in the spacecraft. These by-products, such as free radicals and superoxides, can impact biological processes.

The components of space radiation consisting of energetic charged particles are the ones of greatest

concern in terms of possible biological effects. They can be categorized as galactic cosmic radiation, geomagnetically trapped radiation (the Van Allen belts), and other particles of solar origin. The number of particles arriving from the Sun varies as a function of solar activity, which has a cycle of approximately 11 years. During solar particle events, large numbers of energetic charged particles, mainly protons, are ejected from the Sun and increase the interplanetary radiation levels by several orders of magnitude. *See* COSMIC RAYS; MAGNETOSPHERE; SUN; VAN ALLEN RADIATION.

On Earth the atmosphere shields the planet from short-wavelength ultraviolet radiation and low-energy charged particles. The magnetic field of the Earth shields the planet from low- and medium-energy charged-particle radiation by deflecting the particles along magnetic field lines, where they remain captured for long periods of time. The effectiveness of this magnetic shielding varies with latitude, being highest at the Equator and disappearing at the magnetic poles. The displacement of the magnetic poles with respect to the geographic poles results in regions of higher radiation (for example, the South Atlantic Anomaly) within some low-altitude spacecraft orbits.

Measurements of galactic cosmic radiation from space probes show that these particles consist of 87% protons (hydrogen nuclei), 12% alpha particles (helium nuclei), and 1% energetic heavy nuclei, also known as HZE (high charge and energy) particles, ranging from lithium to tin. The energy of individual particles is very high, reaching 10^{20} eV in some instances. Consequently, it is virtually impossible to passively shield against these particles.

Effects of HZE particles. The energy transfer (the number of ionizations) that occurs along the track of a charged particle is identified in terms of linear energy transfer. High linear energy transfer characterizes highly ionizing radiation, while low linear energy transfer characterizes sparsely ionizing radiation. When HZE particles enter cells, they produce intense ionization along their path and generate a cascade of secondary ionization events. The number of ions produced per unit length of track varies with the charge and velocity of the particles. This variation in ion distribution within a cell is an important factor in the induction of biological damage. *See* LINEAR ENERGY TRANSFER (BIOLOGY).

Radiation hazard and exposure risks are expressed in several interrelated units used to define radiation levels and biological effects due to exposure. For example, a rad of radiation corresponds to the absorption of 100 ergs of energy per gram of any material (10^{-2} joule per kilogram, or 10^{-2} gray). However, a unified theoretical understanding of the radiation quality of cosmic radiation has not been achieved. *See* RADIATION BIOLOGY; UNITS OF MEASUREMENT.

The major radiation hazards for living systems occurring during exploration-class missions outside the Earth's magnetosphere are due to solar-particle-event protons and to the HZE particles in galactic cosmic rays. These hazards are complicated by the fact that there may be a much lower probability of observing the biological effects of HZE particles (and, to some extent, of protons and helium nuclei) than of lightly ionizing radiation (x-rays, gamma rays).

The biological effects of radiation depend on complex energy transduction factors (including energy deposition or dose, dose rate, ionization density, radiochemical intermediaries, molecular biology lesions, and cellular response mechanisms and their regulation and expression in the intact organism). The difference in biological effects due to x-rays and gamma rays on the one hand and heavy charged particles on the other hand are largely derived from the much denser local pattern of ionization around charged-particle trajectories. For lightly ionizing radiation (x-rays and gamma rays), effects are mainly dependent on dose and dose rate. For HZE particles, the concept of dose is inadequate; in particular, the density of energy deposited around a single track is high enough that there is no such thing as a low dose, even for a single particle. *See* RADIATION INJURY (BIOLOGY).

Measurement. Measurements of radiation on board spacecraft are performed with a large variety of detectors. Dosimeters measuring the average energy deposited by radiation are used both inside and outside spacecraft in order to understand the way in which spacecraft materials modify the radiation field. Materials suffering permanent (or long-term) changes in their physical properties due to radiation are used as so-called passive dosimeters. Examples are film emulsions, certain plastics, and thermoluminescent materials. Active dosimeters are radiation detectors whose response to radiation can be monitored electronically as it occurs. These detectors are used to provide ongoing warnings of radiation hazards during space flight activities. *See* DOSIMETER.

Since the biological effects of HZE particles are strongly dependent on the particle type and energy (which result in variations in the linear energy transfer), instruments to detect these particles need to identify their charge and energy. Such instruments are called particle identification spectrometers. They generally consist of stacked semiconductor detectors in conjunction with some velocity-sensitive detector. The characteristic signature identifying the charge of the particle is given by its pattern of energy loss as it traverses the detector stack; for particles stopping in the stack, the total energy deposition is also a measure of the particle energy. *See* JUNCTION DETECTOR.

A simplified approximation of particles and their spectra is obtained by use of a low-pressure, gas-filled ionization chamber of precisely defined geometry designed to have the same electron density as the tissue of humans. The ionization produced in the chamber is proportional to the energy loss of the incident radiation. Theoretical models are then used to interpret these measurements in terms of the biological impact of the radiation. *See* IONIZATION CHAMBER; PARTICLE DETECTOR.

Ground-based experiments. Ground-based experiments provide the means to study the effects of

individual components of space radiation in a 1-*g* environment uncomplicated by all of the other physical factors of space flight, including microgravity. The availability of a ground-based accelerator facility to simulate galactic cosmic-ray irradiation by means of accelerated beams of HZE nuclei is essential for the theoretical understanding of the phenomena involved. Ground-based and space-flight studies of HZE effects have revealed that radiobiological mechanisms of densely ionizing radiation are qualitatively different from those of sparsely ionizing radiation. Ground-based experiments have shown that x-ray irradiation followed by heavy ion exposure is far more damaging to an organism than irradiation in the reverse order. *See* PARTICLE ACCELERATOR.

Experiments in space. The HZE-particle component of galactic cosmic radiation was discovered in 1948, and in 1952 it was predicted that humans exposed to these particles could experience a visual light flash sensation. The crews of *Apollo 1* and subsequent missions reported experiencing a visual light flash. Microscopic examination of the eyes of rats flown on space flights indicates that light flashes and eye damage result from cosmic rays passing through the eye retina. The particles stimulate the photoreceptors of the eye as they pass through the tissue and are thus perceived as light. Even more important to the understanding of HZE particle effects, however, is the finding that HZE particles produce channellike or tunnel lesions. Ground-based accelerator experiments show that each tunnel lesion is produced by a single HZE particle.

Packaged series of monolayers of bacterial spores, plant seeds, animal eggs, and lower animals interlayered with physical track detectors permit evaluation of individual tracks, and allow identification of each penetrating particle and determination of its possible effects on biological matter in its path. It is found that the very high concentration of absorbed energy produced by a single HZE particle of cosmic radiation can have serious biological effects. Mutations are produced in corn seeds, but in general bacterial spores and seeds are highly resistant. The eggs of the brine shrimp *Artemia salina* are extremely sensitive. An egg that survives a hit by an HZE particle demonstrates abnormal growth and behavior when it continues development. Two species of bacteria show reduced viability and antibiotic sensitivity that is believed to be due to defective DNA repair mechanisms. Space radiation also results in an increase in sex-chromosome-linked mutations in the fruit fly, *Drosophila melanogaster*. Radiobiological models derived from ground-based models fall significantly short of reproducing the observed space effects. *See* MUTATION.

Both synergistic and antagonistic relationships are observed between space radiation and the microgravity environment. However, embryonic development is more sensitive to HZE particle damage under microgravity conditions than inside a centrifuge simulating the acceleration of gravity at the Earth's surface (1 *g*). When the eggs of the stick insect, *Carausius morosus*, are exposed in space both at 1 *g* and in a weightless environment, statistically significant differences in development are observed between them. These results indicate that microgravity affects biological processes directly at the cell level, and that these changes in turn modify or interfere with the expression of the effects of the various components of space radiation. It has been proposed that microgravity may affect cell repair and recovery processes, as well as cytoskeleton and membrane functions, thereby making cells more sensitive to radiation damage.

Charles Edwin Wade

Bibliography. H. J. Appell, H. Hoppeler, and H. Oser, Muscle research in space, *Int. J. Sports Med.*, Suppl. 4, 18:S255–S334, 1997; A. Berthoz and A. Güell (eds.), Space neuroscience research, *Brain Res. Rev.*, 28:1–233, 1998; Cell and molecular biology research in space, *FASEB J.*, 13:S1–S178, 1999; International Workshop on Bone Research in Space, *Bone*, 22:71S–161S, 1998; R. Letaw, R. Silberberg, and C. H. Tsao, Radiation hazards on space missions, *Nature*, 303:709–710, 1988; S. A. Moody and C. Golden, Developmental biology in space: Issues and directions in the era of the International Space Station, *Dev. Biol.*, 228:1–5, 2000; A. Mortimer, International Workshop on Radiation Biology, *Mutation Res.*, 430:165–7, 1999; A. Sievers, B, Buchen, and T. K. Scott (eds.), Plant Biology in Space, *Planta*, Supplement to vol. 203, 1997; K. Souza, G. Etheridge, and P. X. Callahan, *Life into Space: Ames Research Center and Kennedy Space Center 1991–1998*, NASA/SP-2000–534, 2000; K. Souza, R. Hogan, and R. Ballard, *Life into Space: Ames Research Center 1965–1990*, NASA/RP-1372, 1992.

Space charge

The net electric charge within a given volume. If both positive and negative charges are present, the space charge represents the excess of the total positive charge diffused through the volume in question over the total negative charge. Since electric field lines end on electric charge, the space-charge density ρ may also be defined in terms of the divergence of the electric field **E** or the laplacian of the electric potential V by Eq. (1) (Poisson's equation). Here ϵ is

$$-\frac{\rho}{\epsilon} = -\text{div}\,\mathbf{E} = \nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} \qquad (1)$$

the dielectric constant of the medium and x, y, and z are rectangular coordinates defining the position of a point in space. If, under the influence of an applied field, the charge carriers acquire a drift velocity v, the space charge becomes j/v, where j is the current density. For current carried by both positive and negative carriers, such as positive ions and electrons, the space charge density is given by Eq. (2). Here the

$$\rho = \frac{j_+}{v_+} - \frac{j_-}{v_-} \qquad (2)$$

subscripts + and − indicate the current density and

drift velocity for the positive and negative carriers, respectively. Thus a relatively small current of slow-moving positive ions can neutralize the space charge of a much larger current of high-velocity electrons. *See* CHILD-LANGMUIR LAW. Edward G. Ramberg

Space communications

Communications between a vehicle in outer space and Earth, using high-frequency electromagnetic radiation (radio waves). Provision for such communication is an essential requirement of any space mission. The total communication system ordinarily includes (1) command, the transmission of instructions to the spacecraft; (2) telemetry, the transmission of scientific and applications data from the spacecraft to Earth; and (3) tracking, the determination of the distance (range) from Earth to the spacecraft and its radial velocity (range-rate) toward or away from Earth by the measurement of the round-trip radio transmission time and Doppler frequency shift (magnitude and direction). A specialized but commercially important application, which is excluded from consideration here, is the communications satellite system in which the spacecraft serves solely as a relay station between remote points on Earth. *See* COMMUNICATIONS SATELLITE; DOPPLER EFFECT; MILITARY SATELLITES; SATELLITE (SPACECRAFT); SCIENTIFIC AND APPLICATIONS SATELLITES; SPACE FLIGHT; SPACE NAVIGATION AND GUIDANCE; SPACE PROBE; SPACECRAFT GROUND INSTRUMENTATION; TELEMETERING.

General features. Certain characteristic constraints distinguish space communication systems from their terrestrial counterparts. Although only line-of-sight propagation is required, both the transmitter and the receiver are usually in motion. The movement of satellites relative to the rotating Earth, for example, requires geographically dispersed Earth stations to achieve adequate communication with the spacecraft on each orbit. The signal power received is expressed by

$$P_r = KP_tG_tG_r\left(\frac{\lambda}{4\pi R}\right)^2$$

where P_r = received signal power; P_t = transmitter power output; G_t = transmitter antenna gain; G_r = receiver antenna gain; K = a factor to allow for system losses, such as cable and connector losses, or losses due to antenna pointing error; R = range (the distance between transmitter and receiver); and λ = wavelength of the transmitted signal in the same units as the range.

Because enormous distances are involved (over a billion miles to the planets beyond Jupiter), the signal received on Earth from deep-space probes is so small that local interference, both artificial and natural, has to be drastically reduced. For this purpose, the transmitted frequency has to be sufficiently high, in the gigahertz range, to reduce noise originating in the Milky Way Galaxy (galactic noise background). The receiver site must be remote from technologically advanced population centers to reduce artificial noise,

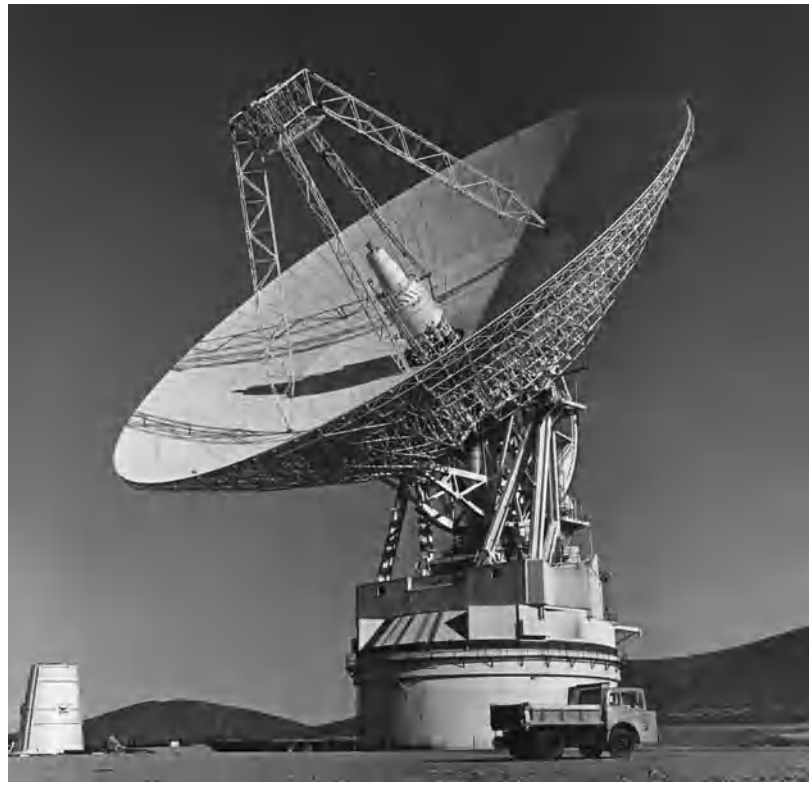

Fig. 1. Goldstone 210-ft (64-m) antenna. (*Jet Propulsion Laboratory*)

and at a dry location such as Goldstone, California (**Fig. 1**), or in the Australian desert, to avoid precipitation attenuation of the radio signal as well as the higher antenna thermal noise associated with higher atmospheric absolute humidity and relatively warm cloud droplets. The receiving antennas must be steerable and large, typically 85 ft (26 m) or at times 210 ft (64 m) in diameter, to enhance the received signal strength relative to the galactic noise background. Special low-noise preamplifiers such as cooled masers are mounted on the Earth receiver antenna feed to reduce the receiver input thermal noise background. Sophisticated digital data processing is required, and the ground-receiver complex includes large high-speed computers and associated processing equipment. *See* MASER; PREAMPLIFIER; RADIO RECEIVER; RADIO TELESCOPE.

The spacecraft communications equipment is constrained by severe power, weight, and space limitations. Typical communications equipment mass ranges from 25 to 220 lb (12 to 100 kg). Another major challenge is reliability, since the equipment must operate for years, sometimes for decades, unattended, in the difficult radiation, vacuum, and thermal environment of space. Highly reliable components and equipment have been developed, and redundancy is employed to eliminate almost all single-point failures. For example, it is not unusual to have as many as three redundant command receivers operating continuously, because without at least one such receiver in operation no command can get through, including a command to switch from a failed command receiver to a backup radio. Power

can be saved by putting some or all of the redundant radios on timers, and to switch to a backup receiver if no commands have been received through the primary receiver within a predetermined interval; but the saved power may come at the cost of a possible delay in emergency response initiation.

Spacecraft power is always at a premium, and other techniques must also be used to minimize its consumption by the communication system. The transmitter is a major power consumer, so its efficiency must be maximized. All aspects of data transmission must contribute to error-free (very low bit error rate) reproduction of the telemetry data using no more power or bandwidth than is absolutely essential. Pulse-code modulation is a common technique which helps meet this goal. In general terms, space communication systems are far less forgiving than terrestrial systems and must be designed, constructed, and tested to much higher standards. *See* PULSE MODULATION; SPACE POWER SYSTEMS; SPACE TECHNOLOGY.

Before describing several typical space communication systems, a brief examination of some important subsystems is in order.

Data sources. The kinds of data that must be transmitted by the space communication system vary, but certain basic types (internal spacecraft environment, scientific, and applications) may be distinguished.

Internal spacecraft environment (housekeeping) data. Onboard sensors provide information from selected locations about a number of functions of internal spacecraft behavior, such as switch position, temperature, pressure, acceleration, voltage, and current. Earth-based monitoring of such data is essential to detect and correct or bypass malfunctions as well as to evaluate trends such as battery, solar cell, and thermal coating degradation.

Scientific data. These data contain the information needed to accomplish the objectives of the mission being flown. For the most part, they originate in sensors of electromagnetic radiation at wavelengths as long as radio waves, ranging to photons as energetic as primary cosmic rays. The sources of this radiation vary from violent astrophysical events which occurred in deep space as long as 10 billion years ago, to solar flare outbursts whose timing and duration over the 23-year solar magnetic-field (sunspot) cycle are still very difficult to understand and to predict. Many of these data are collected and transmitted to Earth in a raster-scan image similar to that used on computer and television displays. Like terrestrial television, the desired data bandwidths are often in the multimegahertz range. Deep-space probes, for which such unchallenged bandwidth needs would produce prohibitively massive and costly spacecraft, trade more transmission time for less bandwidth. Should the mission be a planetary flyby, the data can be acquired in a few hours, stored in solid-state broadband recorders in the probe, and transmitted relatively slowly over the following days or weeks. The ratio of the data collection time interval to the transmission time interval equals the ratio of the required transmission bandwidth to the original data collection bandwidth. Additional bit-error-rate reduction may be accomplished by the statistical treatment of successive transmissions of the same image from the space probe.

Another class of scientific investigations has to do with the immediate external environment of the spacecraft, particularly involving the measurement of the complex terrestrial and solar vector (magnitude and direction) magnetic fields and associated energetic plasma (ionized particles). In lower-altitude denser terrestrial and other planetary atmospheres, measurements of the neutral and ionized atmospheric composition, pressure, temperature, and density are also valuable. Such investigations require telemetry bandwidths in the range of tens of kilohertz, substantially less than that which is required for images.

Applications data. A third category includes data which are supplied regularly as a service to users of images of Earth's land and water surfaces (Earth resources), and of the atmosphere (weather and climatology). The former use is exemplified by *Landsat* 7 images in the 0.45–12.50-micrometer bands, and the latter by the National Oceanic and Atmospheric Administration (NOAA) polar and Geostationary Operational Environmental Satellite (GOES) images in both visible and infrared spectral bands. Such images again use telemetry bandwidths in the multimegahertz range. Another applications data type is information transmitted from data collection platforms on Earth and relayed to ground processing stations via telemetry. Here the data bandwidth is in the kilohertz range. In general, applications data are part of a guaranteed service; that is, if an operational spacecraft fails, another already has been or shortly will be placed in orbit to continue the service.

Antennas. Antennas for space communication may be classified as ground or spacecraft antennas according to their function.

Ground antennas. The use of paraboloidal (dish) reflectors for large-aperture antennas is practically universal. As the size increases, so do cost and structural problems, in step with the volume which is proportional to the aperture diameter cubed. The signal collection area increases only as the diameter squared. The magnitude of reflector surface irregularities sets an upper limit to the usable frequency. Standard reflector diameters in the National Aeronatics and Space Agency (NASA) deep-space network are 85 ft (26 m) and 210 ft (64 m). The feed system has to be carefully designed for optimum performance. For large antennas a common and efficient feed system uses the Cassegrain configuration in which a hyperboloidal reflector is interposed between the focal point of the main antenna and its reflecting surface (Fig. 1). Its main advantage is that the feed horn and associated microwave receiving equipment are near the surface of the paraboloid rather than at its focus, which results in critical operational advantages and higher efficiencies.

Microwave antennas for near-Earth missions usually have smaller diameters, typically 40 or 30 ft (12 or 9 m). Smaller antennas are used for this purpose

because they cost considerably less, and larger systems are not needed because the range is so much smaller. For the reception of lower-frequency VHF (very high frequency) signals from near-Earth spacecraft, an array of crossed-dipole Yagi antennas is often used. *See* ANTENNA (ELECTROMAGNETISM); YAGI-UDA ANTENNA.

Spacecraft antennas. An important requirement for a spacecraft antenna is high reliability in the space environment, since it is usually a component in line to a single-point failure of the communication system. For deep-space missions, spacecraft microwave antennas have diameters up to 16 ft (5 m), depending upon transmitter power, data rate, and range involved. Near-Earth spacecraft use both high-gain and omnidirectional antennas to transmit telemetry data, the particular design depending mainly upon data rates, type of orbit, and the receiving system being employed. Nearly all spacecraft use omnidirectional antennas for receiving commands and transmitting low-data-rate housekeeping information. With this mode of operation, if a partial system failure occurs, the problem can be analyzed and corrective actions taken via the command system to keep the spacecraft in operation.

Modulation techniques. Pulse-code modulation (PCM), in which the data are time-multiplexed, sampled, digitized, and transmitted as a binary stream of pulses grouped into code words, is employed almost universally for transmitting scientific data and command signals. This information is usually phase-modulated on a very stable transmitter carrier to optimize system performance. Techniques such as block coding or convolutional coding are often used to improve the system sensitivity by up to 6 dB (a factor of 4 in power). These techniques require more bandwidth for a given data rate and additional complexity of signal-processing equipment.

Spacecraft transponders. The major function of a spacecraft transponder is to provide a means of measuring range and range rate of a spacecraft by a ground station. This information is then used to compute spacecraft orbits and trajectories. The basic principle of operation is to use a phase-locked loop (PLL) to lock an oscillator coherently to the carrier of the received signal by an exact ratio (typically 240:221 for NASA S-band systems). Comparison of the transmitted and received signals at the ground station provides a coherent two-way Doppler-shift measurement which translates directly into range rate. A ranging signal is modulated on the ground transmitter carrier and returned to the ground station via the transponder carrier; the time delay between transmission and reception on the ground provides a means of computing the range of the spacecraft. Ranging modulation is usually in the form of pseudonoise (PN) PCM codes or coherently related subcarriers. By suitable multiplexing, the transponder also performs the functions of a command receiver and a telemetry transmitter. *See* DOPPLER EFFECT; PHASE-LOCKED LOOPS.

Tracking and Data Relay Satellite System. The Tracking and Data Relay Satellite System (TDRSS) consists of a series of geostationary spacecraft and an Earth terminal located at White Sands, New Mexico. The purpose of TDRSS is to provide telecommunication services between low-Earth-orbiting (LEO) user spacecraft and user control centers (**Fig. 2**). A principal advantage of the system is the elimination of the need for many of the worldwide ground stations for tracking such spacecraft. The *Tracking and Data Relay Satellite* (*TDRS*) provides no processing of data; rather, it translates received signals in frequency and retransmits them. User orbits are calculated from range and range-rate data obtained through the *TDRS* by using transponders on the user spacecraft.

The first *TDRS* was launched in 1983, and the first-generation spacecraft will near the end of their operational life early in the first decade of the new millennium. The upgraded second-generation replacement series was initiated in 2000 with the launch of *TDRS H* (the eighth of the series).

Forward and return link data are transmitted over three basic services in the second-generation TDRSS.

Multiple access. The multiple-access (MA) service provides simultaneous return-link service from as many as 20 LEO user spacecraft, with data rates up to 3 megabits per second (Mbps) for each user. Forward-link service to the user spacecraft is time-shared with a maximum data rate of 300 kilobits per second (kbps), one user at a time. The multiple-access service is provided at S-band frequencies of 2106.4 MHz for the forward link and 2287.5 MHz for the return link. Multiple-access users operating on the same frequency are discriminated by unique pseudonoise codes and antenna beam pointing. Multiple-access antennas on each *TDRS* are adaptive arrays in which signals received by each antenna element are transmitted to White Sands on separate carriers. Ground-based processing provides the ability to form multiple antenna beams simultaneously.

S-band single access. The S-band single-access (SSA) service provides return-link data rates up to 6 Mbps for each user spacecraft. The forward S-band access link provides data at 300 kbps. Carrier frequencies are assigned in the bands from 2025 to 2120 MHz for the forward link and 2200 to 2300 MHz for the return link.

Ka- and Ku-band single access. The K-band single-access (KSA) service provides wide-band forward and return access links. Both forward KSA links provide data at 25 Mbps. The return-link data rates are up to 300 and 800 Mbps for the Ku and Ka bands, respectively. The Ku forward link carrier operates at 13.775 GHz, the return link at 15.0034 GHz. Corresponding Ka assignments are 22–23 and 25–27.5 GHz.

Complete systems. Two examples—the automated *Landsat* 7 and the crewed space shuttle—illustrate TDRSS service to modern space systems.

Landsat 7. This applications spacecraft is designed to monitor the resources and environment of Earth and supply these data to users worldwide. The U.S. Geological Service (USGS) and NASA share

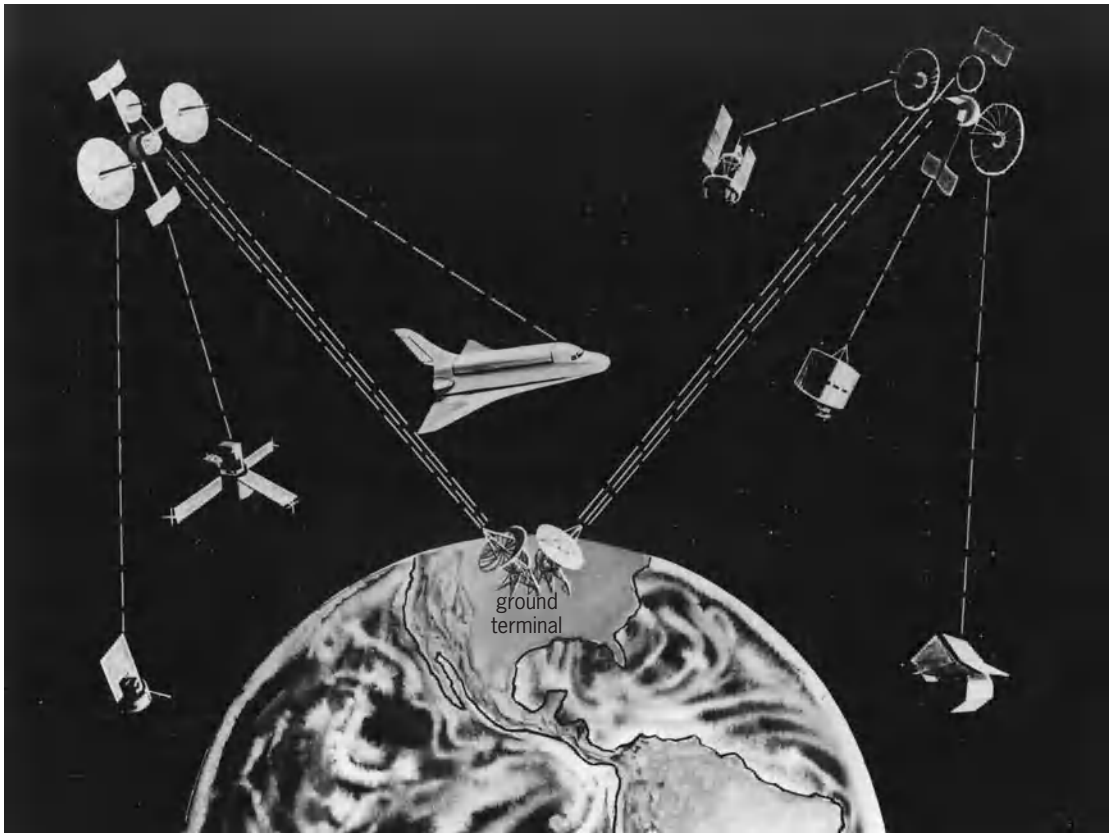

Fig. 2. Tracking and Data Relay Satellite System. (*Goddard Space Flight Center*)

the responsibility for *Landsat* 7 operations, data processing, archiving, and distribution. The Earth-observing sensor system on board *Landsat* 7 is the Enhanced Thematic Mapper Plus (ETM+), optimized for global change studies, land-cover monitoring and assessment, and large-area monitoring. Seven Earth-observing spectral bands are provided at wave-lengths between 0.45 and 12.50 μm, as well as a panchromatic band. Applications of *Landsat* 7 data include crop and forestry monitoring and resource mapping, land-use mapping, geological feature mapping, surface water coverage, coastal resource mapping, and environmental change monitoring. *See* AGRICULTURE; REMOTE SENSING.

Landsat 7 is controlled by the Mission Operations Center (MOC) at the NASA Goddard Space Flight Center. Communication with this satellite may be divided into two parts: the wideband user data generated by the ETM+ sensor, and the narrow-band housekeeping telemetry and command traffic.

Landsat 7 has the capability to transmit wideband ETM+ data to as many as three ground stations simultaneously. A ground network receives these data via an X-band direct downlink at a data rate of 150 Mbps. The primary receiving station is at the USGS Data Center in Sioux Falls, South Dakota. Processing of the image data is done after receipt at the data center by the *Landsat* 7 processing system (LPS). The image data may be transmitted as it is collected, in real time, or it may be stored on an onboard solid-state recorder (SSR) for delayed playback to the ground.

The X-band wideband data system transfer is done by the use of gimbaled X-band antennas (GXA) [**Fig. 3**]. Three frequencies are employed: 8082.5, 8212.5, and 8342.5 MHz. Quadrature phase-shift keying (QPSK) is used, permitting division of each 150-Mbps link into two 75-Mbps data streams, one of which is modulated on the in-phase (I) channel and the other on the quadrature (Q) channel. Each QPSK modulator drives a solid-state power amplifier (SSPA) which transmits the amplified modulated carrier at the appropriate frequency to the ground station through the designated GXA using right-hand circular polarization (RHCP).

Telemetry and commanding of *Landsat* 7 is done through two S-band transponders which operate continuously in their receive mode. Communications with *Landsat* 7 can be done either directly from ground stations of the NASA Spaceflight Tracking and Data Network (STDN) or through the TDRSS. Each transponder is equipped with a transmit/receive S-band omnidirectional antenna which is left-hand circular-polarized (LHCP). Each receiver in these S-band transponders simultaneously searches for TDRSS pseudonoise code correlation and STDN continuous-wave (CW) signals greater than the detection threshold. Upon acquisition of either command signal type, the other mode is inhibited until loss of the initial signal and a subsequent return to the search mode occurs.

The S-band telemetry downlink has three data rates: 1.216, 4.864, and 256.0 kbps. This downlink

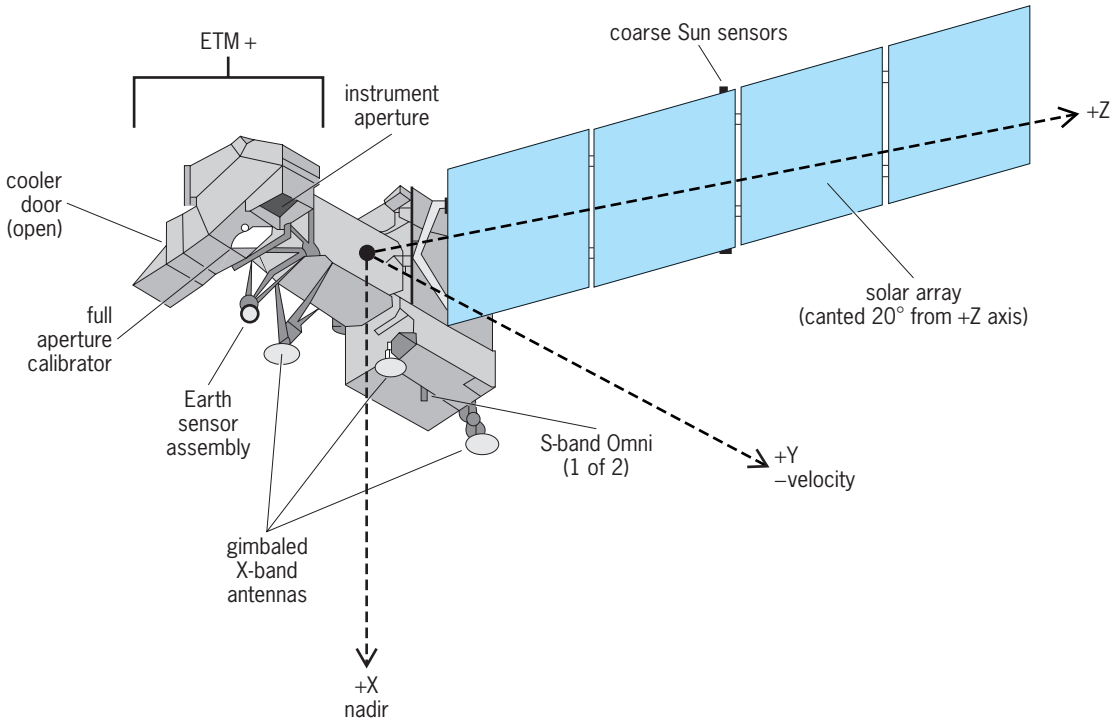

Fig. 3. ***Landsat 7*** **satellite. (*Goddard Space Flight Center*)**

is also used for two-way Doppler tracking. The nominal output frequency is 2287.5 MHz. The telemetry data are phase shift-key (PSK) modulated onto a 1.024-MHz subcarrier. When the recorded data playback is commanded, the 256-kbps SSR data are phase-modulated onto the baseband channel. The PSK subcarrier is summed with the baseband channel data, and this composite signal phase-modulates the transponder transmitter. Command uplinking to the user spacecraft is done through the S-band system at a data rate of 2.0 kbps.

Space shuttle. The space shuttle is a crewed spacecraft system which is used to provide a heavy-lift reusable space transportation capability to low Earth orbit. Since 2000 its primary use has been to support the construction, continued occupation, and long-term microgravity research aboard the *International Space Station* (*ISS*). The launch configuration consists of the orbiter, external tank, and solid rocket boosters (SRBs). Shortly after launch, the SRBs are separated from the space shuttle and are recovered by oceangoing tugs for reuse on later flights. The external tank is jettisoned when the main orbiter engines are shut down, and the orbital maneuvering system (OMS) engines provide the principal orbiter propulsion for the rest of the mission. The space shuttle carries up to seven crew and mission specialists, and each orbiter is designed to be reused for up to 100 missions. It carries a mass of up to 65,000 lb (29,000 kg) of payload in the cargo bay (**Fig. 4**) on ascent and up to 30,000 lb (14,000 kg) on landing. Major cargo includes logistics support items to and from the *ISS*; massive modules to the ISS during its construction, such as its robotic arm; and automated spacecraft for launch or recovery from orbit for refurbishment on Earth and possible reuse. *See* SPACE STATION.

The Orbiter Communications and Tracking System operates with the NASA ground-based STDN, the TDRSS, the U.S. Air Force Satellite Control Facility (SCF), the *ISS*, and crew members performing extravehicular activities. The shuttle uses UHF-, S-, and Ku-band frequencies.

The Ku-band system is designed to operate as a radar during a space rendezvous and, when not so employed, as a two-way link with the TDRSS. The Ku-band antenna is a 3-ft (0.9-m) monopulse paraboloid tracking antenna with a gain of 38 dB and a beam width of 0.36°. This antenna is deployed after the orbiter is in space and the cargo bay doors are opened. The transmitter uses a 50-W traveling-wave tube amplifier. The Ku-band data from the orbiter occupy a 192-kbps channel for voice and operations data, and a channel of up to 50 Mbps for science data. On command, the 50-Mbps channel can be replaced by a 4-MHz analog channel for television transmission. Television is used for monitoring onboard activities, for monitoring the cargo bay area, and by astronauts in extravehicular activities. *See* TRAVELING-WAVE TUBE.

The S-band communication system is composed of two independent subsystems: the network subsystem and the payload subsystem. The network subsystem provides tracking and two-way communication direct to ground stations or through TDRSS. The payload subsystem provides two-way communication with automated orbiting spacecraft in the vicinity of the orbiter. All S-band antennas are low-gain and flush-mounted. Locations are chosen to favor the desired direction of coverage. The flush antennas are

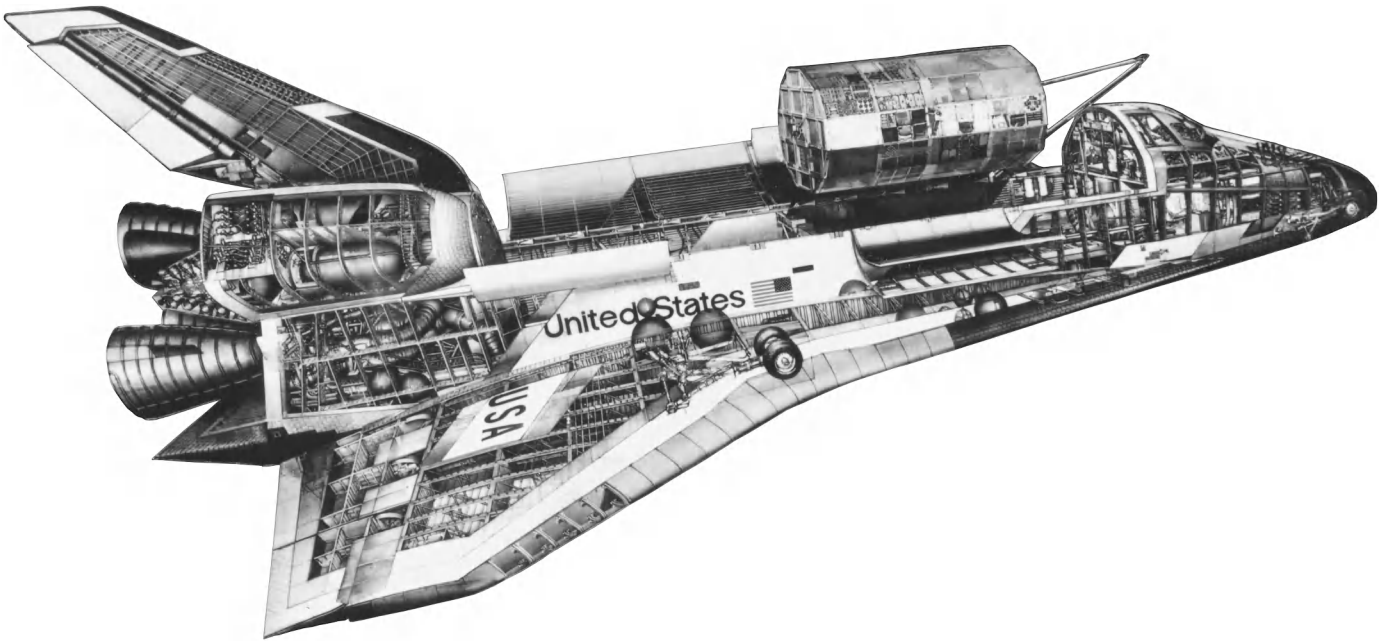

Fig. 4. Cutaway view of the space shuttle orbiter. (*NASA*)

overlaid with thermal protective material to enable them to survive the heat of entry.

The network subsystem is designed to be compatible with both the NASA and the Department of Defense (DOD) networks. The NASA-compatible transponders use a transmit-to-receive frequency ratio of 240:221 and are phase-modulated. When transmitting to NASA ground stations, a power output of 2 W is used. Transmissions to the TDRS use 100 W, and the data are convolutionally encoded to increase the link margins. The DOD-compatible transponders use a frequency ratio of 256:205 and transmit to the SCF with a power of 2 W. *See* PHASE MODULATION.

These transponders provide the capability of two digitally encoded voice channels plus a command channel to the orbiter, and a low-data-rate telemetry and two digital voice return channels. A special-purpose frequency-modulated (FM) transmitter operates on a frequency of 2250 MHz and has a minimum power output of 10 W. This transmitter is used primarily for television video data transmission to ground stations. *See* FREQUENCY MODULATION.

The S-band payload subsystem operates with a large variety of spacecraft communication systems. It is used during the release and recovery of spacecraft by the orbiter and is effective at ranges from a few feet to over a mile. The maximum power output is 2 W, with the ability to transmit at reduced power levels when the spacecraft is nearby.

The UHF communication system is used during extravehicular activities by astronauts and operates at frequencies of 260 and 297 MHz. These transmitters are amplitude-modulated (AM), with direct modulation of voice plus a 5.4-kHz subcarrier of biomedical data. Two orbiter ultrahigh-frequency (UHF) antennas are used, on the inside of the air lock and on the lower portion of the orbiter. *See* AMPLITUDE MODULATION; SPACE SHUTTLE.

John F. Clark; Donald Platt

Bibliography. B. H. Batson (ed.), *IEEE Trans. Commun.*, COM-26:11, November 1978; W. Froehlich, *New Space Network: The Tracking and Data Relay Satellite System*, NASA EP-251, 1986; *Interface Control Document Between Landsat 7 and the Landsat 7 Ground Network (LGN)*, NASA 430-14-01-001-0, 1997; H. B. Poza, TDRSS telecommunications payload: An overview, *IEEE Trans. Aerosp. Electr. Sys.*, AES-15:414–429, May 1979; *Tracking and Data Relay Satellite System (TDRSS) Users' Guide*, Rev. 4, STDN 101.2, Goddard Space Flight Center, January 1980; J. H. Yuen (ed.), *Deep Space Telecommunications Systems Engineering*, 1983.

Space flight

The penetration by humans into the reaches of the universe above the terrestrial atmosphere, and investigation of these regions by automated, remote-controlled and crewed vehicles.

The purpose of space flight is to provide significant contributions to the physical and mental needs of humanity on a national and global basis. Such contributions fall specifically in the areas of (1) Earth resources of food, forestry, atmospheric environment, energy, minerals, water, and marine life; (2) Earth and space sciences for research; (3) commercial materials processing, manufacturing in space, and public services. More general goals of space flight include expansion of knowledge; exploration of the unknown, providing a driving force for technology advancement and, hence, improved Earth-based

productivity; development and occupation of new frontiers with access to extraterrestrial resources and unlimited energy; strengthening of national prestige, self-esteem, and security; and providing opportunity for international cooperation and understanding.

This article focuses on crewed space flight. For discussion of other missions *see* APPLICATION SATELLITES; COMMUNICATIONS SATELLITE; METEOROLOGICAL SATELLITES; MILITARY SATELLITES; SATELLITE (SPACECRAFT); SATELLITE NAVIGATION SYSTEMS; SCIENTIFIC SATELLITES; SPACE PROBE.

To conduct crewed space flight, the two leading space-faring nations, the United States and Russia, formerly the Soviet Union, have developed spacecraft systems and the necessary ground facilities, research and development base, operational know-how, planning experience, and management skills. In the United States, crewed space programs are conducted by the National Aeronautics and Space Administration (NASA), a federal agency established in 1958 for the peaceful exploration of space. In Russia, crewed space flights have been under the auspices of the U.S.S.R. Academy of Sciences; they are now the responsibility of the Russian Space Agency (Rosaviakosmos, or RKA). The first spacecraft with a human on board, the *Vostok 1*, piloted by Yuri A. Gagarin, was launched on April 12, 1961, from the Baikonur Cosmodrome in Kazakhstan and returned after completing one revolution of the Earth. The first American space flight of a human took place 3 weeks later, when NASA launched Alan B. Shepard on May 5 on the *Mercury-Redstone 3* for a 15-min suborbital test flight.

The early spacecraft of both nations were built for only one space flight. The first multiply reusable space vehicle, the space shuttle, was launched by the United States on April 12, 1981. By December 31, 2006, 906 men and 109 women had flown a total of 254 missions. Of these, the United States alone, in a total of 148 missions, had put 668 men and 101 women into space (some of them more than once). Up until permanent occupancy of space began on November 2, 2000, with the boarding of the *International Space Station* (*ISS*) by its first resident crew, humans had spent a total of 11,008 days 3 hours in space (or about 519,525 crew-hours, accounting for crew sizes). By the beginning of permanent residency, the United States had accumulated a total of 1195 days in space (or about 149,174 crew-hours; **Table 1**).

In November 1987, the 13 member nations of the European Space Agency (ESA) agreed to become the third major power engaging in human space flight. European efforts today focus on a two-pronged program: participation as a partner in the *International Space Station*, primarily with the Columbus Orbital Facility (COF), and extension of the Ariane family of expendable launch vehicles with the heavy-lift carrier Ariane 5. By joining the original partnership behind the space station and assigning astronauts to United States space shuttle missions (as did the European Space Agency and Russia), Canada and Japan have also entered the ranks of space-faring nations, while other countries, such as Brazil, are preparing to join.

Crewed Spacecraft

A crewed spacecraft is a vehicle capable of sustaining humans above the terrestrial atmosphere. In a more limited sense, the term "crewed spacecraft" is usually understood to apply to vehicles for transporting and sustaining human crews in space for time periods limited by prestored on-board supplies, as distinct from orbital space stations which support theoretically unlimited habitation of humans in space by autonomous systems, crewed maintenance, and periodic resupply.

Reentry. The basic requirements of crewed spacecraft are quite different from those of crewless space probes and satellites. The presence of humans on board necessitates a life-sustaining environment and the means to return safely to Earth. The major common feature of all crewed spacecraft, therefore, is the atmospheric return element or reentry module. It consists basically of a pressure-tight cabin for the protection, comfort, and assistance of the crew, similar to the cockpits of fighter aircraft, but shaped externally in accordance with the desired airflow and its forces, and surrounded by heat-resistant material, the thermal protection system (TPS), to cope with the high frictional and radiative heating accompanying the energy dissipation phase of the atmospheric return flight to Earth.

To better deal with this heating, the design of the forward-facing portion of the reentry module is usually relatively blunt and symmetrical. Temperatures on and around the stagnation point of the blunt face can reach 3000°F (1700°C) and more. When they are too high for metallic materials, a class of materials called ablators is used. Ablators can be composed of phenolic, silicone rubber, or epoxy compounds. At high temperatures, a great deal of heat is absorbed and carried away from the body by outgassing, charring, and ablating. The remaining heat influx can be dealt with by the insulating properties of the thermal protection system to keep the backface temperatures within the limits of conventional structural materials. *See* NOSE CONE.

Ballistic reentry. A ballistic reentry, flown with a blunt, symmetrical heat shield, is the simplest, most direct mode of return from space, but it also incurs the highest heating on the thermal protection system and the heaviest deceleration loads on the crew since the orbital energy has to be dissipated in the short time of the plunge. This mode also allows no further control of the landing point once the atmosphere has been entered. The first generation of spacecraft of both the United States and the Soviet Union was designed for ballistic reentry and landing by parachute. While the United States since 1981 has gone on to lifting reentry for human space flight, Russia still depends on ballistic reentry with the *Soyuz* spacecraft for crewed missions and the automated *Progress* drones for logistics flights.

Lifting reentry. To provide more flexibility to the spacecraft for selecting its touchdown point and to

TABLE 1. Crewed space flights 1961–2006

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Apr. 12, 1961	*Vostok 1* Yuri A. Gagarin	10,419 lb 1 rev 203 mi	1 h 48 min	First person in space, conducted radio and TV communication with Earth. Spacecraft contained life-support systems, telemetry, ejection seat, and recovery system. Cosmonaut and spacecraft landed at preselected area. Automatic control with cosmonaut available as backup. Two-gas air supply at sea-level pressure.
May 5, 1961	*Freedom 7* Alan B. Shepard, Jr.	2855 lb — 116 mi	15 min	*Mercury* spacecraft launched in ballistic suborbital trajectory by Redstone booster (*Mercury-Redstone 3*). Downrange distance 302 mi. First United States person in space; weightless for 5 min with no ill effects. Astronaut exercised manual control of spacecraft.
July 2–21, 1961	*Liberty Bell 7* Virgil I. Grissom	2836 lb — 118 mi	15 min	*Mercury* spacecraft launched into ballistic trajectory by Redstone booster (*Mercury-Redstone 4*). Downrange distance 303 mi. Weightless for 5 min with no ill effects. Hatch opened prematurely during recovery, spacecraft filled with water and sank in 2500 fathoms in Atlantic Ocean. Astronaut was recovered.
Aug. 6–7, 1961	*Vostok 2* Gherman S. Titov	10,432 lb 16 rev 152 mi	25 h 11 min	First test of prolonged weightlessness; cosmonaut ate, worked, and slept in space. Monitored by TV and radio. Vestibular disturbances produced motion sickness but apparently no significant aftereffects.
Feb. 20, 1962	*Friendship 7* John H. Glenn, Jr.	2987 lb 3 rev 162 mi	4 h 55 min	*Mercury* spacecraft launched into orbit by Atlas booster (*Mercury-Atlas 6*). First United States crewed orbital flight, achieving original Project Mercury objectives of placing a human in orbit, observing reactions to space environment, and making a safe recovery. No adverse physiological effects from the 4.5 h of weightlessness.
May 24, 1962	*Aurora 7* M. Scott Carpenter	2975 lb 3 rev 167 mi	4 h 56 min	*Mercury* spacecraft launched into orbit by Atlas booster (*Mercury-Atlas 7*). Flight plan similar to *Friendship 7*. Manual attitude control exercised during retrofire. Landed 250 mi beyond predicted impact point.
Aug. 11–15, 1962	*Vostok 3* Andrian G. Nikolayev	10,412 lb 60 rev 146 mi	94 h 22 min	Launched as first half of tandem mission with *Vostok 4*. Radiation measurements taken and effects of biological specimens studied. Extensive physiological measurements recorded during prolonged exposure to space environment with no adverse effects noted. TV monitoring, radio and TV contact with *Vostok 4*. Cosmonaut floated free of restraint for a total of 3.5 h, was ejected to land separately from spacecraft.
Aug. 12–15, 1962	*Vostok 4* Pavel R. Popovich	10,425 lb 45 rev 147 mi	70 h 57 min	Launched from same facility as *Vostok 3* into tandem flight. Passed to within 4 mi of *Vostok 3* at closest point. Extensive radiation and physiological experiments conducted. TV monitoring, radio and TV contact with *Vostok 3*. Cosmonaut floated free of restraint for a total of about 3 h, was ejected to land separately from spacecraft.
Oct. 3, 1962	*Sigma 7* Walter M. Schirra, Jr.	3029 lb 5 rev 176 mi	9 h 13 min	*Mercury* spacecraft launched into orbit by Atlas booster (*Mercury-Atlas 8*). Automatic reentry mode utilized, landed 9000 yd from prime recovery ship for first recovery in Pacific Ocean. Experiments included camera, radiation sensors, and ablative samples on the neck of spacecraft.
May 15–16, 1963	*Faith 7* L. Gordon Cooper, Jr.	3033 lb 21 rev 166 mi	34 h 20 min	*Mercury* spacecraft launched into orbit by Atlas booster (*Mercury-Atlas 9*). Long-endurance mission with no adverse physiological effects as astronaut ate, slept, and worked in the space environment. Reentry sequence was with astronaut orienting spacecraft and firing retrorockets manually. A very accurate landing was made in the Pacific Ocean 7000 yd from prime recovery ship. Experiments included infrared and standard photography, radiation measurement, ejection of flashing light (sighted in fifth and sixth orbits), and observation of lights on the ground.
June 14–19, 1963	*Vostok 5* Valery F. Bykovsky	10,408 lb 76 rev 146 mi	119 h 6 min	New time and distance space-flight records set, last 45 revolutions in tandem with *Vostok 6*. Radio contact with *Vostok 6*. Direct TV broadcast to Earth in real time. Extensive medical and biological experiments conducted.
June 16–19, 1963	*Vostok 6* Valentina V. Tereshkova	10,392 lb 45 rev 143 mi	70 h 50 min	Orbited first woman and first person not trained as a pilot. Passed within proximity of *Vostok 5* on first orbit, in different orbital plane. Radio contact with *Vostok 5*. Direct TV broadcast to Earth in real time. Parachuted separately from spacecraft on landing.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Oct. 12–13, 1964	*Voskhod 1* Vladimir M. Komarov Konstantin P. Feoktistov Boris B. Yegorov	11,731 lb 15 rev 254 mi	24 h 17 min	First multicrewed space flight; first "shirt-sleeve" environment in space flight, no pressure suits; record altitude for crewed space flight. Feoktistov, scientist, and Yegorov, physician, as experimenters and observers in addition to Komarov as pilot. Both Feoktistov and Yegorov experienced motion sickness and disorientation illusions. Crew worked, ate, drank, and slept unrestrained. Crew remained in spacecraft through landing.
Mar. 18–19, 1965	*Voskhod 2* Pavel I. Belyayev Aleksei A. Leonov	12,529 lb 16 rev 308 mi	26 h 2 min	New record altitude for crewed space flight. During second orbit Leonov took first space walk with autonomous life-support system, moved 17 ft from spacecraft on tether, 23 min in space environment, including 11 min in air lock through which he exited from spacecraft. While outside, Leonov was observed through the spacecraft TV camera. First Soviet landing sequence using manual system in lieu of automatic. Crew remained in spacecraft through landing. Planned landing point was overshot: landed in deep snow.
Mar. 23, 1965	*Gemini 3* Virgil I. Grissom John W. Young	7111 lb 3 rev 140 mi	4 h 53 min	*Gemini* spacecraft launched into orbit by modified Titan 2 booster (*Gemini-Titan 3*). First United States two-person space flight, Grissom first person in space for second time. One orbital maneuver was conducted in each of the three orbits. Manual control was exercised throughout the reentry phase, using the limited lifting characteristics of the spacecraft to steer toward touchdown.
June 3–7, 1965	*Gemini 4* James A. McDivitt Edward H. White II	7879 lb 62 rev 184 mi	97 h 56 min	*Gemini* spacecraft launched into orbit by modified Titan 2 booster (*Gemini-Titan 4*). Endurance mission of 4 days to test crew and spacecraft in buildup to longer missions. Eleven scientific and engineering experiments, in addition to basic medical checks on 4 days of exposure to the space environment. White spent 22 min outside the spacecraft in EVA, propelling himself with an oxygen-jet gun. During space walk he was attached to the spacecraft with a 25-ft tether-and-oxygen line; he carried emergency oxygen and camera. Spacecraft computer malfunction necessitated a ballistic reentry sequence.
Aug. 21–29, 1965	*Gemini 5* L. Gordon Cooper, Jr. Charles Conrad, Jr.	7947 lb 120 rev 217 mi	190 h 55 min	*Gemini* spacecraft launched into orbit by modified Titan 2 booster (*Gemini-Titan 5*). Endurance mission of 8 days confirmed the physiological feasibility of Apollo lunar landing mission. First flight of fuel-cell electrical power system. On Aug. 23 a simulated rendezvous was conducted with a series of four maneuvers through two orbits which raised the perigee to 124 mi and apogee to 192.6 mi. Five medical experiments measured physiological effects; extensive weather and terrain photography was conducted. Voice communication was conducted with *Sea Lab 2* under the Pacific Ocean.
Dec. 4–18, 1965	*Gemini 7* Frank Borman James A. Lovell, Jr.	8076 lb 206 rev 204 mi	330 h 35 min	*Gemini* spacecraft boosted into orbit by modified Titan 2 booster (*Gemini-Titan 7*). Endurance mission of 14 days, longest scheduled flight in Gemini program. Station-keeping conducted with booster second stage after separation; spacecraft maneuvered to serve as rendezvous target for *Gemini 6*. Accurate controlled reentry was made, landing 7.6 mi from target. One astronaut at a time doffed his pressure suit to enhance comfort. No adverse physiological effects from 2 weeks in the space environment. Twenty experiments were incorporated into the mission, ranging from detailed medical measurements and sleep analysis to Earth photography and radiometry.
Dec. 15–16, 1965	*Gemini 6* Walter M. Schirra, Jr. Thomas P. Stafford	7817 lb 16 rev 193 mi	25 h 51 min	*Gemini* spacecraft launched into orbit by modified Titan 2 booster (*Gemini-Titan 6*). Rendezvous with *Gemini 7* as target. A series of posigrade maneuvers began the rendezvous chase and shortened the initial 1200 mi separation distance to 730 mi. During second revolution the distance was reduced to 300 mi with another thruster firing, followed by a plane change to match *Gemini 7's* orbit. Orbit was adjusted to near circular at beginning of third revolution; as fourth revolution began, the terminal phase of rendezvous was initiated, resulting in approach to within 1 ft of *Gemini 7*. Station-keeping and fly-around maneuvers were conducted for 5 h 19 min. Controlled reentry to landing within 13 mi of prime recovery ship.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Mar. 16, 1966	*Gemini 8* Neil A. Armstrong David R. Scott	8351 lb $6^1/_2$ rev 186 mi	10 h 41 min	*Gemini* spacecraft launched into orbit by modified Titan 2 booster, 1 h 41 min after Gemini-Agena target vehicle (GATV)- Atlas booster combination. Rendezvous with the GATV was accomplished in fourth orbit, 6 h after launch, and first docking of two vehicles in space was achieved 33 min later. After 28 min of flight in the docked configuration, a thruster in the orbital-attitude maneuvering system stuck open and the spacecraft began to spin. *Gemini* spacecraft was undocked from GATV and, after 25 min, was stabilized by use of reentry control system. Mission terminated early with landing in secondary recovery area in western Pacific.
June 3–6, 1966	*Gemini 9a* Thomas P. Stafford Eugene A. Cernan	8268 lb 45 rev 194 mi	75 h 21 min	Mission utilizing augmented target docking adapter (ATDA) as the target was placed into orbit on June 1 with an Atlas booster. The *Gemini* was launched into orbit on June 3 by its modified Titan 2 booster. Rendezvous with ATDA achieved on third orbit. Crew confirmed that shroud covering the docking collar had failed to separate, and docking was canceled. Two additional rendezvous were performed as planned, one using visual techniques, the last from above the ATDA. After 1 h of EVA, Cernan's visor accumulated fog, and communications between Stafford and Cernan were poor. Planned use of the self-contained life-support and propulsion systems to maneuver was canceled. Total EVA was 2 h 7 min. Controlled reentry to 0.42 mi of target.
July 18–21, 1966	*Gemini 10* John W. Young Michael Collins	8248 lb 43 rev 476 mi	70 h 47 min	GATV-Atlas booster combination launched 100 min before *Gemini* spacecraft launched into orbit by modified Titan 2 booster. Rendezvous and docking with GATV 10 accomplished in fourth orbit. GATV engine used to propel the docked combination to record altitude, then proper orbit to rendezvous with GATV 8. After separation from GATV 10, *Gemini 10* effected rendezvous with the passive target. First period of standup EVA was terminated at 50 min when both crew members suffered eye irritation. Second EVA period, after rendezvous with passive GATV, consisted of Collins on umbilical moving to GATV and recovering micrometeorite experiment package.
Sept. 12–15, 1966	*Gemini 11* Charles Conrad, Jr. Richard F. Gordon, Jr.	8509 lb 44 rev 853 mi	71 h 17 min	GATV-Atlas booster combination launched 1 h 37 min before *Gemini* spacecraft was launched into orbit by modified Titan 2 booster within 2-s launch window. Rendezvous accomplished on first orbit using on-board information exclusively. Docking at 1 h 34 min after launch. Docking accomplished twice by each crew member. During umbilical EVA, Gordon removed the nuclear-emulsion experiment package and attached tether between *Gemini* spacecraft and GATV. Excessive fatigue from these activities caused early termination of EVA at 33 min. Using GATV primary propulsion system, docked configuration was propelled to an altitude of 853 mi. Hatch opened third time for 2 h 10 min standup EVA, during which scheduled photography was accomplished. The two spacecraft undocked, and station-keeping by use of a 100-ft nylon tether was exercised with a slow spin rate imparted to the tethered combination. After separation of the tether, a second rendezvous with GATV was conducted from 25 mi separation. Reentry sequence was fully automatic, with impact 2.9 mi from the aiming point.
Nov. 11–15, 1966	*Gemini 12* James A. Lovell, Jr. Edwin E. Aldrin, Jr.	8297 lb 59 rev 187 mi	94 h 35 min	GATV-Atlas booster combination launched 1 h 38 min before *Gemini* spacecraft was launched into orbit by modified Titan 2 booster. Rendezvous was accomplished in third orbit with docking 4 h 14 min after *Gemini* launch. Two standup EVA periods, on first and third days, totaled 3 h 34 min. On second day Aldrin conducted 2 h 6 min umbilical EVA, testing restraint devices to overcome body-positioning problems experienced on prior flights. Resting frequently, Aldrin used portable hand rails, foot restraints, and various tethers. He completed 19 tasks, demonstrating that useful EVA work was feasible with proper planning and restraining devices. *Gemini* and GATV spacecraft undocked, remaining joined by 100-ft tether for station-keeping and gravity-gradient stabilization experiment. Fourteen experiments were conducted. Automatic reentry sequence was used for second time.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Apr. 23–24, 1967	*Soyuz 1* Vladimir M. Komarov	14,000 lb 18 rev 139 mi	24 h 45 min	After a 2-year hiatus, Soviet Union crewed space flight resumed with *Soyuz 1*. During reentry on its eighteenth orbit, after 25.2 h of flight, *Soyuz 1*'s parachutes fouled and the 14,000-lb spacecraft crashed, killing Komarov.
Oct. 11–22, 1968	*Apollo 7* Walter M. Schirra, Jr. Donn Eisele Walter Cunningham	87,382 lb 163 rev 244 mi	260 h 9 min 3 s	First crewed flight of *Apollo* command and service modules. Launched by a Saturn 1B, this was an engineering test flight and was a complete success. Rendezvous and simulated docking exercises with the Saturn upper stage were practiced. A TV camera was carried, and a series of in-flight programs were staged by the astronauts, showing their activities within the command module. Transmissions were relayed live to home receivers all over the United States.
Oct. 26–30, 1968	*Soyuz 3* Georgi T. Beregovoi	14,000 lb 64 rev 151 mi	94 h 51 min	First successful crewed flight of the *Soyuz* spacecraft. Accomplished orbital rendezvous with the uncrewed *Soyuz 2*, which was orbited several days previously. Docking was not attempted. Automatic control was used to bring the two craft within 650 ft of each other. At the point the pilot assumed manual control, and he made at least two simulated dockings. Television pictures of the inside of the spacecraft were relayed to home receivers in the Soviet Union and Europe.
Dec. 21–27, 1968	*Apollo 8* Frank Borman James A. Lovell, Jr. William Anders	87,000 lb 10 rev 231,000 mi (lunar module)	146 h 59 min 49 s	First crewed circumlunar orbital mission. Also first use of Saturn 5 launch vehicle for a crewed flight. Remained in an orbit about 69 mi above the lunar surface for approximately 20 h (10 revolutions). The ability to use lunar landmarks for navigation to the lunar landing sites was successfully verified. Communications with Earth were excellent, and TV pictures of good quality, of both the lunar surface and the Earth from lunar and translunar distances, were transmitted to the Earth from circumlunar reentry speed. Landing was within 3 mi of the recovery ship.
Jan. 14–17, 1969	*Soyuz 4* Vladimir A. Shatalov	14,000 lb 48 rev 158 mi	71 h 14 min	Rendezvoused and docked with *Soyuz 5* to form the "first experimental space station." Cosmonaut Shatalov in *Soyuz 5* performed the actual docking. The two spacecraft remained docked for 4 h 35 min, during which time cosmonauts Khrunov and Yeliseyev transferred from *Soyuz 4* to *Soyuz 5*, spending about 1 h outside the spacecraft in the process.
Jan. 15–18, 1969	*Soyuz 5* Boris V. Volynov Yevgeny V. Khrunov Aleksei S. Yeliseyev	14,000 lb 50 rev 158 mi	72 h 46 min	After docking with *Soyuz 4*, cosmonauts Khrunov and Yeliseyev transferred to *Soyuz 5* and cosmonaut Volynov remained in the spacecraft. Details of the transfer operation were transmitted over TV.
Feb. 28–Mar. 13, 1969	*Apollo 9* James A. McDivitt David R. Scott Russell L. Schweikart	95,231 lb 151 rev 240 mi	241 h 1 min	Earth orbital flight designed to give the lunar module its first crewed flight test. This vehicle was crewed for 28 h by astronauts Schweikart and McDivitt. During this period the hatches of both command module and the lunar module were opened. Schweikart got out onto the "front porch" of the lunar module and thus was able to exercise the suit and backpack designed for operation on the lunar surface. Some curtailment of this activity was deemed advisable because of Schweikart's previous nausea. Subsequently, the lunar module was separated from the command module, and after 6 h 24 min was rendezvoused and docked with the command module. The astronauts in the lunar module reentered the command module and then jettisoned the lunar module. Numerous navigation and control exercises and Earth photography experiments were accomplished. Because of weather conditions in the prime recovery areas, the landing area was moved 500 mi south. Splashdown occurred within 3 mi of the recovery ship.
May 18–26, 1969	*Apollo 10* Thomas P. Stafford John W. Young Eugene A. Cernan	96,000 lb 31 rev 251,000 mi (lunar module)	192 h 3 min 23 s	Lunar module tested in lunar orbit. Astronauts Stafford and Cernan entered the lunar module, which separated from the command module, descended to within 50,000 ft of the lunar surface, and then rendezvoused and docked with the command module, which had remained in an orbit about 69 mi above the lunar surface. Transmitted to the Earth were color TV pictures of the lunar surface, of the Earth from lunar and translunar distances, and of the activities of the astronauts inside the command module. This flight successfully verified that the *Apollo* spacecraft was ready to attempt a lunar landing on the next flight.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
July 16–24, 1969	*Apollo 11* Neil A. Armstrong Michael Collins Edwin E. Aldrin, Jr.	96,000 lb 31 rev 250,000 mi (lunar CM)	195 h 17 min 12 s	The objective of the Apollo program (a crewed lunar landing) was achieved in this flight. Astronaut Armstrong emerged from the lunar module on July 20, 1969, followed shortly afterward by Aldrin. This event was recorded by a TV camera and relayed to home receivers in many parts of the world by satellite. The astronauts gathered samples of the lunar surface material for return to Earth. After deploying a United States flag, they installed a passive seismic instrument and a laser reflector on the lunar surface so that data of scientific interest could be obtained after their return to Earth. The astronauts found that mobility on the lunar surface in the one-sixth gravity environment presented no unanticipated problems, and walking was relatively easy. During the whole flight all the systems functioned correctly and the other parts of the flight followed *Apollo 10* pattern closely.
Oct. 11–16, 1969	*Soyuz 6* Georgi S. Shonin Valeri N. Kubasov	14,000 lb 79 rev 144 mi	118 h 42 min	Tested effectiveness of various welding techniques in space. Along with *Soyuz 7* and *8*, tested multiple launch capacity of Soviet Union and performed various scientific investigations.
Oct. 12–17, 1969	*Soyuz 7* Anatoly V. Filipchenko Vladislav N. Volkov Viktor V. Gorbatko	14,000 lb — 141 mi	118 h 41 min	Along with *Soyuz 6* and *8*, tested capacity of Soviet Union for multiple launch and rendezvous; also performed navigation experiments and various scientific investigations.
Oct. 13–18, 1969	*Soyuz 8* Vladimir A. Shatalov Aleksei S. Yeliseyev	14,000 lb — 140 mi	119 h	Along with *Soyuz 6* and *7*, tested capacity of Soviet Union for multiple launch and rendezvous; also tracked other ships' maneuvers and performed various scientific investigations.
Nov. 14–22, 1969	*Apollo 12* Charles Conrad, Jr. Richard F. Gordon, Jr. Alan L. Bean	96,000 lb 45 rev 250,000 mi (lunar CM)	142 h 52 min	More lengthy exploration of the Moon's surface.
Apr. 11–17, 1970	*Apollo 13* James A. Lovell, Jr. Fred W. Haise, Jr. John L. Swigert, Jr.	97,000 lb — 250,000 mi	142 h 52 min	Rupture of a fuel-cell oxygen tank in service module forced cancellation of the third crewed trip to the Moon. The rupture was caused by high pressure which developed when the failure of two thermal switches responsible for regulating the tank heaters led to abnormally high temperatures within the tank, produced by the combustion of Teflon in the supercritical (cryogenic liquid) oxygen. Lunar module life-support system enabled crew to return to Earth.
June 1–19, 1970	*Soyuz 9* Andrian G. Nikolayev Vitali I. Sevastyanov	14,000 lb 300 rev 167 mi	424 h 59 min	Predominantly biomedical flight, testing human ability to live and work for prolonged periods of time in spacecraft environment. Earth resources applications experiments were also performed.
Jan. 31–Feb. 9, 1971	*Apollo 14* Alan B. Shepard Jr. Stuart A. Roosa Edgar D. Mitchell	98,000 lb 34 rev 250,000 mi (lunar CM)	216 h 2 min	Third lunar landing. Precision touchdown accomplished by Shepard and Mitchell in "Antares." On two EVAs, crew deployed Apollo lunar surface scientific experiment package (ALSEP), completed other scientific tasks. First extraterrestrial use of a two-wheeled cart, modular equipment transporter (MET). New record surface time, 9 h 24 min. Surface samples returned to Earth, 94 lb.
Apr. 19, 1971	*Salyut 1* Crewless	40,000 lb — 176 mi	175 days	First Soviet space station. Prototype. Launched crewless. Reentered Oct. 11, 1971.
Apr. 23–24, 1971	*Soyuz 10* Vladimir A. Shatalov Aleksey S. Yeliseyev Nikolai Rukavishnikov	14,500 lb 32 rev 165 mi	47 h 46 min	Three-person spacecraft, docked with *Salyut 1* space station. Remained docked for 5.5 h, but no crew transfer took place, probably because of mechanical difficulties.
June 6–30, 1971	*Soyuz 11* Georgi T. Dobrovolskiy Vladislav N. Volkov Viktor I. Patsayev	14,300 lb 385 rev 140 mi	570 h 22 min	Second docking with *Salyut 1*. Crew boarded *Salyut* and remained for 23 days 17 h 40 min, breaking previous *Soyuz 9* record. After extensive on-board activities, crew transferred to *Soyuz* for Earth return and was killed during reentry by pulmonary embolisms, caused by rapid decompression of cabin through leaking hatch valve.
July 26–Aug. 7, 1971	*Apollo 15* David R. Scott Alfred M. Worden James B. Irwin	107,000 lb 74 rev 250,000 mi (lunar CM)	295 h	Perfect landing by Scott and Irwin in "Falcon" at Hadley Rill. First use of lunar roving vehicle (LRV), driven a total of 17.5 mi during three surface EVAs. Brought back more than 170 lb of samples. First subsatellite, carrying particles and fields (P&F) experiment, launched in lunar orbit by crewed spacecraft. First inflight EVA performed by Worden in deep space on return leg.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Apr. 16–27, 1972	*Apollo 16* John W. Young Thomas K. Mattingly II Charles M. Duke	103,000 lb 64 rev 250,000 mi (lunar CM)	265 h 51 min 5 s	Landing by Young and Duke in "Orion" at Descartes. Second use of LRV, for a total of 16.6 driven miles. Three EVAs totaling 20 h 14 min yielded 209 lb of samples. Fourth ALSEP station was deployed on surface and second P&F subsatellite launched in lunar orbit. On return leg, Mattingly undertook second deep-space EVA to retrieve film from external bay.
Dec. 7–19, 1972	*Apollo 17* Eugene A. Cernan Ronald E. Evans Harrison H. Schmitt	103,000 lb 75 rev 250,000 mi (lunar CM)	301 h 51 min	Last *Apollo* lunar landing, by Cernan and Schmitt, a geologist, in "Challenger" in mountainous region at Taurus-Littrow. Third LRV was driven 21 mi; fifth ALSEP deployed. Longest lunar stay time (74 h 59 min 38 s), with three EVAs setting new record of 22 h 5 min 6 s. Most samples returned to Earth, 250 lb. Third deep-space film retrieval EVA, by Evans.
Apr. 3, 1973	*Salyut 2* Crewless	40,000 lb —	—	Second Soviet space station. Failed in orbit.
May 14, 1973–July 11, 1979	*Skylab 1* Crewless	170,000 lb (without payload shroud) — 276 mi	—	First United States (experimental) space station, launched crewless on two-stage Saturn 5. At 60 s after launch, the workshop meteoroid/thermal shield tore loose. Primary solar array wing was lost, the other jammed closed by debris. As a result, inside temperatures became too high and power output too low for long-term human habitation. Station was saved through determined effort by NASA engineers on ground and astronauts in space, to become successful beyond expectations. Station reentered July 11, 1979.
May 25–June 22, 1973	*Skylab 2* Charles Conrad, Jr. Joseph P. Kerwin Paul J. Weitz	31,000 lb 404 rev 276 mi	28 days 49 min	Crew launch on Saturn 1B postponed from May 15 to analyze problems. After entering *Skylab*, crew deployed parasol-like thermal blanket on workshop outside through small air lock. On June 7, Conrad and Kerwin performed hazardous EVA to repair jammed solar wing. Three EVAs totaled 5 h 51 min.
July 28–Sept. 25, 1973	*Skylab 3* Alan L. Bean Jack R. Lousma Owen Garriott	31,000 lb 858 rev 270 mi	59 days 1 h 9 min	Launch vehicle, Saturn 1B. Crew erected new twin-pole sun shield on EVA, installed new rate gyros. Extended solar observations with Apollo telescope mount (ATM), and 39 Earth observation passes. Three EVAs totaled 13 h 44 min. Space manufacturing and maneuvering unit tests added to mission schedule.
Sept. 27–29, 1973	*Soyuz 12* Vasily G. Lazarev Oleg G. Makarov	14,460 lb 32 rev 216 mi	Approx. 48 h	First flight after *Soyuz 11* tragedy. Short-duration test of a modified *Soyuz* with new hatch/valve design, improved maneuvering and manual control capabilities. Tested new space suit design and cabin suit loop.
Nov. 16, 1973–Feb. 8, 1974	*Skylab 4* Gerald P. Carr William R. Pogue Edward Gibson	31,000 lb 1214 rev 270 mi	84 days 1 h 16 min	Launch vehicle, Saturn 1B. Last *Skylab* mission with 4 EVAs for film exchange and mechanical repairs, totaling 22 h 21 min. Mission included first exoatmospheric observation of a comet, Kohoutek.
Dec. 18–26, 1973	*Soyuz 13* Pyotr I. Klimuk Valentin V. Lebedev	14,400 lb — 169 mi	7 days 20 h 55 min	Similar to *Soyuz 12*, with addition of *Orion 2* telescope system for solar, stellar, and cometary studies. Also included were biomedical studies, such as Levka brain-bloodflow device.
June 25, 1974	*Salyut 3* Crewless	40,000 lb — 172 mi	214 days	Third Soviet space station. Military type. New solar panel and other systems.
July 3–19, 1974	*Soyuz 14* Pavel R. Popovich Yuri P. Artukhin	14,400 lb — 169 mi	15 days 17 h 39 min	Crew docked and transferred to *Salyut 3* space station. Purpose was to test the new space station. Stay on board was $14^1/_2$ days.
Aug. 26–28, 1974	*Soyuz 15* Gennady V. Sarafanov Lev S. Demin	14,400 lb 32 rev 183 mi	48 h 12 min	Intended as second crew transfer to *Salyut 3*, but unable to dock because remotely controlled docking system failed.
Dec. 2–8, 1974	*Soyuz 16* Anatoly V. Filipchenko Nikolai Rukavishmikov	14,400 lb — 140 mi	5 days 22 h 24 min	Soviet test and simulation of ASTP mission, including new compatible docking system and reduction of spacecraft pressure from 14.7 to 10 lb/in.^{2} in flight. To approximate the United States docking system for docking tests, a dummy structural ring was taken along and separated in orbit.
Dec. 26, 1974	*Salyut 4* Crewless	40,000 lb — 226 mi	770 days	Fourth Soviet space station. Civilian type.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Jan. 11–Feb. 9, 1975	*Soyuz 17* Aleksei Gubarev Georgiy Grechko	14,400 lb — 225 mi	29 days 13 h 20 min	Crew docked and transferred to *Salyut 4* space station.
Apr. 5, 1975	*Soyuz 18A* Vasiley Lazarev Oleg Makarov	14,400 lb — —	Approx. 20 min	Suborbital flight. Mission aborted after launch due to vehicle failure after approximately 20 min of flight.
May 24–July 26, 1975	*Soyuz 18* Pyotr I. Klimuk Vitali I. Sevastyanov	14,400 lb — 220 mi	62 days 23 h 20 min	Second crew transfer to *Salyut 4*, which had completed 3352 revolutions at the time. Flew part of the time simultaneously with ASTP.
July 15–24, 1975	Apollo-Soyuz Test Project (ASTP) *United States*: Thomas P. Stafford Vance D. Brand Donald K. Slayton *Soviet Union*: Aleksely A. Leonov Valerity N. Kubasov	 32,500 lb 138 rev 143 mi 15,000 lb 96 rev —	 9 days 1 h 28 min 5 days 22 h 31 min	First international link-up in space (docking at 31 mi max. distance from Earth for 1 day 23 h 17 min) and last flight of an *Apollo* spacecraft. *Apollo* launched on Saturn 1B $7^1/_2$ h after *Soyuz*. First handshake between Stafford and Leonov occurred 3:19 p.m. EDT on July 17. Crews performed interspacecraft transfer four times and conducted joint activities. A total of 31 experiments were performed by American crew, some of them jointly with Soviets, and docking tests with *Soyuz* were practiced. After final undocking, each spacecraft continued on own mission. *Soyuz* returned July 21.
June 22, 1976	*Salyut 5* Crewless	40,000 lb — 169 mi	412 days	Fifth Soviet space station. Civilian type.
July 6–Aug. 24, 1976	*Soyuz 21* Vitaly Zholobov Boris V. Volynov	14,4000 lb — 169 mi	49 days 6 h 24 min	Crew docked on July 7 and transferred to *Salyut 5* space station, launched June 22. On-board activities included earth resources photography, weather observations; formation of crystals and metals-melting in zero *g*; plant and insect growth for genetic studies. Crew reportedly suffered sensory deprivation effects. Flight had been preceded by uncrewed *Soyuz 20* (launched Oct. 17, 1975, docked to *Salyut 4* on Oct. 19).
Sept. 15–23, 1976	*Soyuz 22* Valery F. Bykovsky Vladimir Aksenov	14,400 lb 127 rev 175 mi	7 days 21 h 54 min	Earth surveillance mission. Test of a special East German camera system for use on space stations.
Oct. 14–16, 1976	*Soyuz 23* Vyacheslav Zudov Valery Rozhdestvenskiy	14,400 lb 32 rev 172 mi	2 days 6 min	Attempt to dock with *Salyut 5* space station failed. Crew returned to Earth.
Feb. 7–25, 1977	*Soyuz 24* Viktor V. Gorbatko Yuri Glazkov	14,400 lb 285 rev 176 mi	17 days 17 h 23 min	Crew transferred to *Salyut 5* on Feb. 8. Space station had mostly military mission with special photo reconnaissance runs.
Sept. 29, 1977	*Salyut 6* Uncrewed	42,000 lb — 221 mi	—	Uncrewed launch of new Soviet space station. Reentered July 29, 1982.
Oct. 9–11, 1977	*Soyuz 25* Vladimir Kovalyonok Valery Ryumin	14,400 lb 32 rev 221 mi	2 days 46 min	Attempt to dock with *Salyut 6* failed. Crew returned to Earth.
Dec. 10–Jan. 16, 1977	*Soyuz 26* Yuri Romanenko Georgiy Grechko	14,400 lb 590 rev 221 mi	37 days 10 h 6 min Crew duration: 96 days 10 h	Crew transferred to *Salyut 6* for a 96-day mission, breaking United States record of 84 days on Mar. 4, 1978. Crew returned on Mar. 16, 1978, with *Soyuz 27*.
Jan. 10–16, 1978	*Soyuz 27* Vladimir Dzhanibekov Oleg G. Makarov	14,400 lb 1024 rev 219 mi	64 days 22 h 53 min Crew duration: 5 days 22 h 59 min	Crew transferred to *Salyut 6* to join *Soyuz 26* crew. Resupplied by first crewless logistics flight, *Progress 1* (launched Jan. 20). On return to Earth, crew used the "older" *Soyuz 26* spacecraft, leaving *Soyuz 27* for the long-duration crew (returned Mar. 16).
Mar. 2–10, 1978	*Soyuz 28* Aleksei Gubarev Vladimir Remek	14,400 lb 125 rev 221 mi	7 days 22 h 17 min	Crew transferred to *Salyut 6* on Mar. 4 to join *Soyuz 26* crew. First non-Soviet cosmonaut: Remek (Czechoslovakia).
June 16–Sept. 3, 1978	*Soyuz 29* Vladimir Kovalyonok Alexander Ivanchenko	14,400 lb 125 rev 230 mi	79 days 15 h 23 min Crew duration: 140 days	Crew transferred to *Salyut 6* on June 17 to establish new duration record (140 days). Returned on Nov. 2 with *Soyuz 31* spacecraft. Resupplied by *Progress 2* (launched July 7), *Progress 3* (launched Aug. 7), *Progress 4* (launched Oct. 3).

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
June 27–July 5, 1978	*Soyuz 30* Pyotr I. Klimuk Miroslav Hermaszewski	14,400 lb 125 rev 225 mi	7 days 22 h 4 min	Crew transferred to *Salyut 6* to join *Soyuz 29* crew. Second non-Soviet cosmonaut: Hermaszewski (Poland).
Aug. 26–Nov. 2, 1978	*Soyuz 31* Valery F. Bykovsky Sigmund Jaehn	14,400 lb 1070 rev 221 mi	67 days 20 h 14 min Crew duration: 7 days 20 h 49 min	Crew transferred to *Salyut 6* to join *Soyuz 29* crew. Third non-Soviet cosmonaut: Jaehn (East Germany). Crew returned with *Soyuz 29* spacecraft.
Feb. 25–June 13, 1979	*Soyuz 32* Vladimir Lyakhov Valery Ryumin	14,400 lb 1711 rev 232 mi	108 days 4 h 24 min Crew duration: 175 days	Crew transferred to *Salyut 6* to establish new endurance record (175 days). Returned on Aug. 19 with *Soyuz 34*. *Soyuz 32* returned uncrewed. Resupplied by *Progress 5* (launched Mar. 12) and *Progress 6* (launched May 13).
Apr. 10–12, 1979	*Soyuz 33* Nikolai Rukavishnikov Georgi Ivanov	14,400 lb 31 rev 221 mi	1 day 23 h 1 min	Docking with *Salyut 6* failed. Crew returned to Earth. Fourth non-Soviet cosmonaut: Ivanov (Bulgaria).
June 6–Aug. 19, 1979	*Soyuz 34*	14,400 lb 1152 rev 257 mi	73 days 18 h 17 min	Launched crewless and docked to *Salyut 6*. Returned *Soyuz 32* crew. Resupplied by *Progress 7* (launched June 28). New type of crewless *Soyuz* (*Soyuz T1*) launched Dec. 16 on flight test to dock with *Salyut 6*.
Apr. 9–June 3, 1980	*Soyuz 35* Leonid Popov Valery Ryumin	14,400 lb 869 rev 220 mi	55 days 22 min Crew duration: 184 days 20 h 12 min	Crew transferred to *Salyut 6* for record-duration mission (184 days). Returned Oct. 11 with *Soyuz 37* spacecraft. Resupplied by *Progress 8* (launched Mar. 27), *Progress 9* (launched Apr. 27), *Progress 10* (launched June 29), *Progress 11* (launched Sept. 28).
May 26–July 31, 1980	*Soyuz 36* Valery Kubasov Bertalan Farkas	14,400 lb 1040 rev 220 mi	65 days 20 h 54 min	Crew transferred to *Salyut 6* to join with *Soyuz 35* crew. Returned June 3 with *Soyuz 35* spacecraft. Fifth non-Soviet cosmonaut: Farkas (Hungary).
June 5–9, 1980	*Soyuz T2* Yuriy Malyshev Vladimir Aksenov	14,400 lb 62 rev 220 mi	3 days 22 h 41 min	First crewed test of new type spacecraft. Crew transferred to *Salyut 6* on June 6. Tenth crewed visit to *Salyut 6*.
July 23–Oct. 11, 1980	*Soyuz 37* Viktor V. Gorbatko Pham Tuan	14,400 lb 1258 rev 220 mi	79 days 15 h 17 min Crew duration: 7 days 20 h 42 min	Crew transferred to *Salyut 6* on July 25 to join with *Soyuz 35* crew. Returned July 31 with *Soyuz 36* spacecraft. Sixth non-Soviet cosmonaut: Pham Tuan (Vietnam).
Sept. 18–26, 1980	*Soyuz 38* Yuri Romanenko Arnaldo T. Mendez	14,400 lb 124 rev 220 mi	7 days 20 h 43 min	Crew transferred to *Salyut 6* to join with *Soyuz 35* crew. Returned Sept. 26 with *Soyuz 38*. Seventh non-Soviet cosmonaut: Mendez (Cuba).
Nov. 27–Dec. 10, 1980	*Soyuz T3* Leonid Kizim Oleg G. Makarov Gennady Strekalov	14,400 lb 202 rev 220 mi	12 days 19 h 8 min	Crew transferred to *Salyut 6*. Thirteenth crewed visit to *Salyut 6*. First three-person mission since ill-fated *Soyuz 11* in 1971. Returned Dec. 10. *Salyut 6* resupplied by *Progress 12* (launched Jan. 24).
Mar. 12–May 26, 1981	*Soyuz T4* Vladimir Kovalyonok Viktor Savinykh	14,400 lb 1183 rev 220 mi (estimated)	75 days	Crew transferred to *Salyut 6* on Mar. 14. Savinykh became 100th human to go into space. To date: 43 U.S. astronauts, 57 East-bloc cosmonauts.
Mar. 22–30, 1981	*Soyuz 39* Vladimir Dzhanibekov Jugderdemidiyn Gurragcha	14,400 lb 125 rev 220 mi	7 days 20 h 43 min	Crew transferred to *Salyut 6* on Mar. 23 to join with *Soyuz T4* crew. Returned on Mar. 30 with *Soyuz 39*. Eighth non-Soviet cosmonaut: Gurragcha (Mongolia).
Apr. 12–14, 1981	*STS 1* John Young Robert Crippen	205,000 lb 36 rev 172 mi	2 days 6 h 20 min 59 s	First crewed orbital test flight of the U.S. space shuttle *Columbia*, for a full, successful orbital check of the revolutionary reusable spacecraft and its many subsystems. Return to Earth, featuring atmospheric entry at Mach 25, hypersonic/supersonic maneuvering flight and subsonic unpowered landing at Edwards AFB, California, was celebrated by millions around the world. Time of landing on Apr. 14: 1:20 P.M. EST.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
May 14–22, 1981	*Soyuz 40* Leonid Popov Dumitru Prunariu	14,400 lb 124 rev 220 mi	7 days 20 h 41 min	Crew transferred to *Salyut 6* on May 15 to join with *Soyuz T4* crew. Returned May 22 with *Soyuz 40*. Ninth non-Soviet cosmonaut: Prunariu (Romania).
Nov. 12–14, 1981	*STS 2* Joe Engle Richard Truly	210,000 lb 36 rev 166 mi	2 days 6 h 13 min	Second crewed orbital test flight of the U.S. space shuttle *Columbia*. Launch at Cape Canaveral after two postponements and a liftoff delay due to hardware problems. Mission duration reduced from the planned 124 h to 54 h due to malfunction of on-board fuel cell. Despite shortened mission, crew achieved all objectives. Landing on Rogers Dry Lake, Edwards AFB.
Mar. 22–30, 1982	*STS 3* Jack R. Lousma C. Gordon Fullerton	210,000 lb 130 rev 150 mi	8 days 4 min	Third crewed orbital test of space shuttle *Columbia*. Launch at Cape Canaveral after 1 h delay. Fully successful. Flight extended 1 day to enable landing at Northrup Strip, New Mexico, after rains rendered Rogers Dry Lake, California, too soft.
Apr. 19, 1982	*Salyut 7* Crewless	42,000 lb — 163 mi	—	Crewless launch of new Soviet space station, replacing *Salyut 6* (deorbited July 29, 1982). Eleventh anniversary of launch of *Salyut 1*. *Salyut 7* and an attached module *Cosmos 1686* remained in orbit for almost 9 years, reentering on Feb. 6, 1991, after about 50,200 orbits of the Earth.
May 13–Dec. 10, 1982	*Soyuz T5* Anatoliy Berezovoy Valentin Lebedev	14,400 lb Approx. 3380 rev (crew) 216 mi	106 days Crew duration: 211 days	Crew transferred to *Salyut 7* on May 14 and went on to establish new duration record of 211 days. Was host to two visiting crews, *T6* and *T7*, and returned with *Soyuz T7* spacecraft.
June 25–July 2, 1982	*Soyuz T6* Vladimir Dzhanibekov Aleksandr Ivanchenkov Jean-Loup Chrétien	14,400 lb 125 rev 194 mi	7 days 21 h 51 min	Crew transferred to *Salyut 7* on June 25 to join with *T5* crew. Returned July 2 with *T6* spacecraft. Tenth non-Soviet cosmonaut: Jean-Loup Chrétien (France).
June 27–July 4, 1982	*STS 4* Thomas K. Mattingly Henry W. Hartsfield	210,000 lb 112 rev 150 mi	7 days 1 h 10 min	Fourth and last orbital test flight of *Columbia*, after 74 workdays of refurbishment. Mission accomplished and flight test program concluded with great success. Landing at Edwards AFB.
Aug. 19–27, 1982	*Soyuz T7* Leonid Popov Aleksandr Serebrov Svetlana Savitskaya	14,400 lb 125 rev (crew) 194 mi	113 days Crew duration: 7 days 21 h 52 min	Crew transferred to *Salyut 7* on Aug. 20 to join *T5* crew. Returned Aug. 27 with *T5* spacecraft. Second woman in space: Svetlana Savitskaya.
Nov. 11–16, 1982	*STS 5* Vance D. Brand Robert Overmyer Joseph Allen William Lenoir	210,000 lb 81 rev 150 mi	5 days 2 h 14 min	First operational mission of shuttle *Columbia* successfully delivers reimbursable cargo of two commercial communications satellites (comsats) to orbit. First four-person crew. After landing at Edwards, *Columbia* returns to manufacturer for changes necessitated by *Spacelab* payload (*STS 9*).
Mar. 10–Aug. 14, 1983	*Kosmos 1443* Crewless	42,000 lb — 214 mi	157 days	*Salyut 7* temporarily enlarged by docking of crewless "Kosmodul" *Kosmos 1443* resupply/reentry vehicle.
Apr. 4–9, 1983	*STS 6* Paul Weitz Karol Bobko Donald Peterson Story Musgrave	250,000 lb 80 rev 183 mi	5 days 24 min	First flight of space shuttle *Challenger* featuring first EVA by shuttle astronauts (Musgrave and Peterson) after 9 years. Cargo of 2-ton *Tracking and Data Relay Satellite* (*TDRS*) deployed successfully but attained geostationary orbit (GEO) later as planned after major rescue program by NASA (58 days) involving 39 maneuvers by mini-thrust-control jets.
Apr. 20–22, 1983	*Soyuz T8* Vladimir Titov Gennady Strekalov Alexandr Serebrov	14,400 lb 32 rev 194 mi	48 h 18 min	Mission aborted. Crew returned after failure of link-up with *Salyut 7* space station when rendezvous radar deployed incorrectly.
June 18–24, 1983	*STS 7* Robert Crippen Frederick Hauck John Fabian Sally Ride Norman Thagard	250,000 lb 97 rev 187 mi	6 days 2 h 24 min	Shuttle *Challenger* with first five-person crew including first American woman in space (Ride). Successful deployment of two commercial comsats (*Anik C-2, Palapa B-1*). First deployment/retrieval of free-flying experiment-carrying platform, *Shuttle Palette Satellite* (*SPAS 01*) of West Germany, with Canadian-built robot-arm remote manipulator system.
June 27–Nov. 24, 1983	*Soyuz T9* Vladimir Liakov Aleksandr Aleksandrov	14,400 lb 2390 rev 214 mi	149 days 9 h 46 min	Crew transferred to *Salyut 7/Kosmos 1443* on June 28, conducted close to 300 experiments during their stay—which was extended unexpectedly when a planned follow-on mission did not come because of explosion of carrier rocket during the launch attempt

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
June 27–Nov 24, 1983 (*cont.*)				on Sept. 26. Its crew, Titov and Strekalov, already ill-fated on *T8*, escaped serious injury. *T9* crew added solar arrays, delivered by *Kosmos 1443*, to *Salyut 7*, and returned on Nov. 24 with *T9* spacecraft.
Aug. 30–Sept. 5, 1983	*STS 8* Richard Truly Daniel Brandenstein Dale Gardner William Thornton Guion Bluford	250,000 lb 97 rev 193 mi	6 days 1 h 8 min	Third flight of *Challenger*, featuring first night launch (Canaveral) and first night landing (Edwards). First black NASA astronaut in space (Bluford). Successful delivery of Indian comsat *Insat 1-B*, and test of robot-arm remote manipulation system with 4-ton proof mass. First inflight investigations into space motion sickness phenomenon (space adaptation syndrome; SAS) by physician Thornton.
Nov. 28–Dec. 8, 1983	*STS 9/SL 1* John Young Brewster Shaw Byron Lichtenberg Ulf Merbold Owen Garriott Robert Parker	250,000 lb (*Spacelab*: 26,000) 166 rev 156 mi	10 days 7 h 47 min	First flight of European-built (ESA) *Spacelab* aboard shuttle *Columbia*. First six-person crew. First foreign shuttle astronaut: Merbold (West Germany). Work aboard *Spacelab* on more than 70 experiments in many scientific-technical disciplines went so well that mission was extended by 1 day. Landing at Edwards, after mission had conclusively proved the practical worth of humans in space.
Feb. 3–11, 1984	*STS 41-B** Vance D. Brand R. Gibson Bruce McCandless Robert Stewart Ron McNair	250,000 lb 127 rev 185 mi	7 days 23 h 16 min	Tenth flight of shuttle and fourth flight of *Challenger*, carrying two reimbursable commercial comsats (*Westar 6, Palapa B-2*). Satellites deployed successfully, but payload assist module (PAM) upper stage failed in both cases. Comsats retrieved later by mission 51-A. First free (untethered) EVA performed by McCandless, wearing manned maneuvering unit (MMU) back-pack system, to 300 ft from orbiter. Second black astronaut (McNair). First landing at Kennedy Space Center, Florida.
Feb. 8–Oct. 2, 1984	*Soyuz T10* Leonid Kizim Vladimir Solovyov Oleg Atkov	14,400 lb 3790 rev (crew) 200 mi	63 days Crew duration: 236 days 21 h 50 min	Crew transferred to *Salyut 7* on Feb. 9 and went on to set new endurance record of 237 days. During the 34-week stay, Kizim and Solovyov made six EVAs and spent almost 23 h outside the station, mainly for major repair work on *Salyut 7*. Crew took more than 25,000 photographs of Earth and was joined twice by visiting crews (*T11* and *T12*). Resupplied by *Progress 19* on Feb. 23, *Progress 20* on Apr. 17, *Progress 21* on May 8, *Progress 22* on May 30, and *Progress 23* on Aug. 15. Returned on Oct. 2 with *Soyuz T11* spacecraft.
Apr. 2–11, 1984	*Soyuz T11* Juri Malishev Grennady Strekalov Rakesh Sharma	14,400 lb 125 rev (crew) 200 mi	183 days Crew duration: 7 days 21 h 42 min	Crew transferred to *Salyut 7* to join *T10* crew. Returned with more than 1000 photographs of India and other Earth areas with *Soyuz T10* spacecraft. Eleventh non-Soviet cosmonaut: Sharma (India).
Apr. 6–13, 1984	*STS 41-C* Robert Crippen Richard Scobee George Nelson Terry Hart James van Hoften	250,000 lb (LDEF: 21,000) 107 rev 313 mi	6 days 23 h 41 min	Fifth flight of *Challenger*. Flew new "direct" ascent to higher-than-usual altitude for rendezvous with solar telescope satellite *Solar Maximum Mission* (*SMM*), inoperative since end of 1980. Crew deployed large Long Duration Exposure Facility (LDEF), then captured and repaired *SMM* with full success, setting new EVA record of 7 h 18 min.
July 17–19, 1984	*Soyuz T12* Vladimir Dzhanibekov Sventlana Savitskaya Igor Volk	14,400 lb 196 rev 221 mi	12 days 7 h 14 min	Crew transferred to *Salyut 7* on July 18 to join *T10* crew. Second space mission of Savitskaya, who also became world's first female spacewalker (3 h 35 min), performing experiments with electron-beam tool for space assembly and repair. Crew returned with *T12* spacecraft.
Aug. 30–Sept. 5, 1984	*STS 41-D* Henry W. Hartsfield Michael Coats Richard Mullane Stephen Hawley Judith Resnik Charles Walker	250,000 lb 96 rev 185 mi	6 days 56 min	First flight of space shuttle *Discovery*. Cargo: three reimbursable comsats (*SBS 4, Telstar 3-C, Leasat 1*) successfully deployed. Also: OAST-1 experiment pallet. Second U.S. woman in space (Resnik). First "passenger" from industry (Walker, of McDonnell Douglas).
Oct. 5–13, 1984	*STS 41-G* Robert Crippen Dave Walker Kathryn Sullivan Sally Ride David Leestma Marc Garneau Paul Scully-Power	250,000 lb 132 rev 220 mi	8 days 5 h 24 min	Sixth flight of *Challenger*. First seven-person crew, including two women, one Canadian (Garneau), one Australian (Scully-Power). Payload: *Earth Radiation Budget Satellite* (*ERBS*), successfully deployed; OAST-3 experiments. First U.S. woman to walk in space (Sullivan). Second landing at Kennedy Space Center.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

*Beginning with 41-B, shuttle flights were no longer numbered consecutively. In the new system, the first digit indicates the last digit of the current fiscal year (begins Oct. 1 of previous calendar year), that is, 4 = 1984. The second digit indicates the launch site: 1 = Kennedy Space Center; 2 = Vandenberg Air Force Base. The letter indicates which serial flight in the fiscal year; for example, B = second flight.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Nov. 8–16, 1984	*STS 51-A* Frederick Hauck Dave Walker Joseph Allen Anna Fisher Dale Gardner	250,000 lb 126 rev 220 mi	7 days 23 h 45 min	Second flight of *Discovery*. Two reimbursable comsats *(Telesat-H, Leasat/Syncom IV-1)* successfully deployed. Two others, *Palapa B-1* and *Westar 6* from mission 41-B, successfully retrieved and returned to Earth for repair. Fifth U.S. woman in space (Fisher). Third landing at Kennedy Space Center.
Jan. 24–27, 1985	*STS 51-C* Thomas Mattingly Loren Shriver Gary Payton James Buchli Ellison Onizuka	— 47 rev —	3 days 1 h 34 min	Third flight of *Discovery*. First dedicated Department of Defense mission, classified secret. Fourth landing at Kennedy Space Center.
Apr. 12–19, 1985	*STS 51-D* Karol Bobko Donald Williams Margaret Rhea Seddon Jeffrey Hoffman David Griggs Charles Walker E.J. "Jake" Garn	250,000 lb 109 rev 283 mi	6 days 23 h 55 min	Fourth flight of *Discovery*. Payload of two reimbursable comsats (*Telesat-I/Anik, Syncom IV-3*) successfully deployed, but one (*Syncom*) failed to ignite its upper stage despite heroic attempts by crew with remote manipulator system. First flight of a U.S. politician (Senator Garn, Utah). Second flight of industry representative Walker. Sixth U.S. woman in space (Seddon). Landing at Kennedy Space Center resulted in damaged tires.
Apr. 29–May 6, 1985	*STS 51-B/Spacelab 3* Robert Overmyer Frederick Gregory Don Lind Norman Thagard William Thornton Lodewijk van den Berg Taylor Wang	250,000 lb 110 rev 220 mi	7 days 8 min	Second flight of *Spacelab* pressurized module. Crew conducted over a dozen major experiments and tested research-animal holding facility with rats and two squirrel monkeys. Mission accomplished with great success. Landing of *Challenger* at Edwards AFB.
June 6–Sept. 26, 1985	*Soyuz T13* Vladimir Dzhanibekov Viktor Savinykh	14,400 lb — 200 mi	112 days/ Dzhanibekov; 168 days 3 h 51 min/Savinykh	Crew transferred to *Salyut 7* space station on June 8, after testing new flight control systems in *Soyuz* spacecraft. Performed extensive repair/maintenance of power system on station. Resupplied by *Progress 24* (launched June 21), *Cosmos 1669*, a new type of free-flying, crewless platform (docked July 21, separated Aug. 29). On Sept. 26, Dzhanibekov returned with "visiting" Georgi Grechko in *Soyuz T13*, leaving Savinykh with Vasyutin and Volkov in *Salyut* until Nov. 21.
June 17–24, 1985	*STS 51-G* Daniel Brandenstein John Creighton Shannon Lucid Steven Nagel John Fabian Patrick Baudrey Sultan Salman Al-Saud	250,000 lb 111 rev 240 mi	7 days 1 h 40 min	Fifth flight of *Discovery*, with cargo of three reimbursable comsats (*Morelos-A/Mexico; Arabsat 1B/Arabia; Telstar 3D/AT&T*). Crew also deployed and later retrieved free-flying instrument platform SPARTAN 1 (Shuttle Pointed Autonomous Research Tool for Astronomy). Payload included six scientiflc-technological experiments from U.S., France, and Germany. Two foreign astronauts: Prince Al-Saud (Saudi Arabia), Baudry (France). Seventh U.S. woman (Lucid).
July 29–Aug. 6, 1985	*STS 51-F/Spacelab 2* Charles G. Fullerton Roy D. Bridges, Jr. Karl G. Henize Anthony W. England F. Story Musgrave Loren W. Acton	252,800 lb 126 rev 197 mi	7 days 22 h 45 min	Eighth flight of *Challenger*, first flight of *Spacelab* pallets-only version. Lift-off after 1 h 37 min delay due to computer problem. One of three orbiter engines shut down prematurely (after 5 min 45 s) but not early enough to cause mission abort. Flight extended 1 day tohelp recover from on-board experiment problems. Mission accomplished. Landing at Edwards AFB.
Aug. 27–Sept. 3, 1985	*STS 51-I* Joe H. Engle Richard O. Covey James D. A. van Hoften John M. Lounge William F. Fisher	337,000 lb 111 rev 280 mi	7 days 2 h 17 min	Sixth flight of *Discovery*. Liftoff after two launch delays, one (Aug. 24) caused by weather, the other (Aug. 25) by backup computer problem. Payload of three reimbursable comsats [*ASC 1, AUSSAT 1* (Australia), *Leasat (Syncom) IV-4*] successfully deployed, the first two on the same day—a first. Crew performed rendezvous with *Leasat (Syncom) IV-3*, deployed Apr. 13 by *STS 51-D* and drifting lifeless, and successfully conducted salvage and repair operation in space on Aug. 31 and Sept. 1, redeploying the comsat for normal operation. Landing at Edwards AFB.
Sept. 17–Nov. 21, 1985	*Soyuz T14* Vladimir Vasyutin Georgy Grechko Alexander Valkov	14,400 lb — 200 mi	9 days/Grechko; 64 days 21 h 52 min/Vasyutin and Volkov; 168 days 3 h 51 min/Savinykh	Crew transferred to *Salyut 7/Soyuz T13* on Sept. 18. Veteran cosmonaut Grechko returned with Dzhanibekov on Sept. 26 in *Soyuz T13*, making this the first successful space station crew rotation. Vasyutin, Volkov, and Savinykh returned suddenly on Nov. 21 with *Soyuz T14* because of illness of Vasyutin, requiring hospitalization.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Oct. 3–7, 1985	*STS 51-J* Karol Bobko Ronald J. Grabe Robert Stewart William A. Pailes David C. Hilmers	— 63 rev 320 mi	4 days 1 h 45 min	First flight of fourth shuttle orbiter *Atlantis*, setting new altitude record (after *41-C*). Payload: military (classifled). Landing at Edwards AFB.
Oct. 30–Nov. 6, 1985	*STS 61-A/Spacelab D1* Henry W. Hartsfield Steven R. Nagel Guion Bluford Bonnie J. Dunbar James F. Buchli Reinhard Furrer Ernst Messerschmid Wubbo Ockels	250,000 lb (*Spacelab*: 26,000) 111 rev 202 mi	7 days 32 min	Ninth flight of *Challenger* with the West German *D-1* mission, the flrst foreign-chartered *Spacelab* flight. Crew included three foreign astronauts: Furrer, Messerschmid (both W. Germany), and Ockels (Netherlands). Eighth U.S. woman (Dunbar). First eight-member crew. Mission concluded on Nov. 6 with great success. Landing at Edwards AFB with test of new nose wheel steering.
Nov. 26–Dec. 3, 1985	*STS 61-B* Brewster H. Shaw Bryan D. O'Connor Sherwood C. Spring Mary L. Cleave Jerry L. Ross Rudolfo Neri Vela Charles D. Walker	261,455 lb 109 rev 235 mi	6 days 22 h 54 min	Second flight of *Atlantis*, carrying three commercial comsats: *Morelos-B* (Mexico), *Aussat 2* (Australia), and *Satcom KU-2*. Ross and Spring conducted major outboard activities demonstrating space-station-type assembly and handling of large structures in space (payload designations: EASE/ACCESS). One foreign payload specialist: Neri (Mexico). Third flight of industry representative Walker. Ninth U.S. woman (Cleave). First flight of *PAM-D2* upper stage. Heaviest PAM payload (Satcom). Landing at Edwards AFB after full mission accomplishment.
Jan. 12–18, 1986	*STS 61-C* Robert L. Gibson Charles F. Bolden, Jr. George D. Nelson Steven A. Hawley Franklin E. Chang-Diaz Robert J. Cenker C. William Nelson	250,000 lb 96 rev 201 mi	6 days 2 h 4 min 9 s	First flight of *Columbia* after major modiflcations since its last flight (*STS 9*) in Nov. 1983. Originally scheduled for Dec. 18, 1985, launch was delayed seven times due to systems problems and inclement weather. Crew included first Hispanic American in space (Chang-Diaz, Costa Rica), a payload specialist from industry (Cenker, RCA), and the second U.S. congressman in space (Florida U.S. Rep. Bill Nelson). Crew deployed advanced *Satcom K-1* communications satellite, took photographs of Comet Halley, and performed numerous on-board experiments. Landing at Edwards AFB after two postponements due to bad weather at Kennedy Space Center in Florida. Second night landing.
Jan. 28, 1986	*STS 51-L* Francis R. Scobee Michael J. Smith Judith A. Resnik Ellison S. Onizuka Ronald E. McNair Gregory B. Jarvis Christa McAuliffe	250,000 lb — —	1 min 13 s	*Challenger* was launched at 11:38 a.m. EST after a night of below-freezing temperatures. Its right solid-rocket booster developed a catastrophic leak due to malfunction of a joint seal between two motor elements. The shuttle was destroyed 73 s after liftoff, killing its crew of seven and bringing the United States space program to a halt for 32 months.
Feb. 19, 1986	*Mir* Crewless	235,000 lb 210 rev —	—	Crewless launch of new Soviet space station *Mir* (Peace), similar in size and design to *Salyut 7*, with estimated length of 57 ft. Represents a new-generation station, with six docking ports for future buildup with add-on modules and docking of *Soyuz* and *Progress* logistics fiights.
March 13–July 16, 1986	*Soyuz T15* Leonid Kizim Vladimir Solovyev	14,400 lb 2000 rev (est.) 210 mi	125 days	Crew docked at *Mir* on Mar. 15 about 49 h after liftoff and transferred to the space station. First resupply flight followed on Mar. 19 with launch of crewless *Progress 25* which docked to *Mir* on Mar. 21. *Progress 26* was launched on Apr. 23, and the tanker-transport docked with *Mir* on Apr. 26. On May 5, Kizim and Solovyev transferred from *Mir* to *Salyut 7* using *Soyuz T15*. Their stay in the *Salyut 7/Soyuz T15/Cosmos 1686* complex, trailing *Mir* in orbit by 15 min, included 3 h 50 min spacewalk on May 28 to demonstrate the ability to work with a 15-m-long triangular lattice. A crewless *Soyuz TM*, launched on May 21, docked with *Mir* on May 23 and returned to Earth on May 30—a test of a modifled version of the *Soyuz* with improved computers, rendezvous-docking systems, and landing parachute. A second EVA on May 31, 5 h long, also demonstrated structural deployment. *Progress 26* undocked on June 22 and deorbited on June 23. On June 26, the crew returned to *Mir* with *Soyuz T15*. On July 16, the crew returned in *Soyuz T15* to Earth. Kizim had set a new record of 373 days of cumulative time in space by a human. Landing was 34 mi northeast of Arkalyk in Kazakhstan.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Feb. 6–July 30, 1987	*Soyuz TM-2* Yuri Romanenko Alexander Laveikin	14,400 lb Approx. 145 rev 230 mi	174 days 1 h 22 min/ Laveikin; 8 days 20 h 1 min/Faris and Viktorenko	First fiight to *Mir* in 1987 was the crewless *Progress 27* tanker/transport, launched on Jan. 16, and docked to *Mir* on Jan. 18, to provision the station for the subsequent crew visit. After liftoff in a new, more advanced *Soyuz* version, crew docked at *Mir* on Feb. 8 for a long-duration attempt. After an initial zero-*g* adaptation problem experienced by Laveikin, crew settled down to conduct materials processing work. On Mar. 5, *Progress 28*, a new tanker/transport, docked with *Mir's* rear port and was unloaded. On Mar. 9, its engines boosted *Mir* to increase its orbit altitude.
March 31, 1987	*Kvant* Crewless	43,000 lb — 250 mi		An astrophysics module, *Kvant*, launched on Mar. 31, failed to dock to the *Mir* core on Apr. 5. After a second fruitless attempt on Apr. 9, the crew corrected the problem—a piece of cloth left by the ground crew—in a 3 h 40 min EVA on Apr. 12, and docking was accomplished. *Kvant's* propulsion module was jettisoned Apr. 13, making the module's rear docking port accessible. Resupply ship *Progress 29*, launched Apr. 21, docked to this port on Apr. 23. After unloading and taking on stored waste, it burned up in reentry on May 11. On May 13, *Progress 30* was launched with fresh supplies and mail. In a second and third EVA, on June 12 and 16, the crew installed a third solar power array on the station.
July 21–Dec. 29, 1987	*Soyuz TM-3* Alexander Viktorenko Alexander P. Alexandrov Mohammed Faris	14,400 lb Approx. 2576 rev 230 mi	161 days/ Alexandrov; 326 days/ Romanenko; 8 days/ Levchenko	After launch from Baikonur, the crew, which included the first Syrian cosmonaut (Faris), docked with *Mir* on July 24. On July 29, Viktorenko and Faris, accompanied by Laveikin who had shown an abnormal electrocardiogram reading, undocked *Soyuz TM-2* and landed in Kazakhstan early on July 30. Romanenko and Alexandrov continued the mission on *Mir.* Resupply: *Progress 31*, docked Aug. 7, separated Sept. 22; *Progress 32*, launched Sept. 22, docked Sept. 26. On Sept. 30, Romanenko broke the current record of 237 days. After separation of *Progress 32*, resupply ship *Progress 33* docked on Nov. 23. When Yuri Romanenko returned Dec. 29, 1987, in *TM-3* with Alexandrov and Anatoliy Levchenko (of *TM-4*), he had set a new record for the longest crewed space flight to date–326 days.
Dec. 21, 1987–June 17, 1988	*Soyuz TM-4* Vladimir Titov Musa Manarov Anatoliy Levchenko	14,400 lb — 230 mi	366 days 22 h 39 min/Titov and Manarov; 7 days 21 h 2 min/ Levchenko	Docked on Dec. 23, after 33 revolutions, to *Mir*. When, on Dec. 29, Levchenko (a shuttle test pilot), Alexandrov, and Romanenko returned to Earth in *TM-3*, landing at 12:16 P.M. about 50 mi from Arkalyk in Kazakhstan. Romanenko had set a new world record of 326 days in space. Titov and Manarov went on for a long-duration stay in *Mir*, slightly over 1 year. Resupply ship *Progress 34* was launched Jan. 21, 1988, and docked early on Jan. 23, *Progress 35* followed on Mar. 24 (undocked May 5), and *Progress 36* on May 15.
June 7–Sept. 7, 1988	*Soyuz TM-5* Anatoly Solovyov Viktor Savinykh Alexander Alexandrov	14,400 lb 162 rev 225 mi	10 days 8 h 5 min (crew)	Docked. to *Mir* on June 9. The 10-day mission was the 63d crewed flight since Gagarin. Alexandrov was the second Bulgarian (after Georgi Ivanov) and the 13th foreign national flown by the Soviets since 1978. Countdown and launch were widely televised. Visitor crew returned to Earth on June 17 in *TM-4*, leaving the "fresh" *TM-5* docked to *Mir.* On June 18, Titov and Manarov transferred the *TM-5* from the aft to the forward docking port. During a 5-h EVA on June 30, Titov and Manarov did not succeed in replacing a faulty detector unit on a British/Dutch x-ray telescope outside *Mir*, because of a broken wrench. *Progress 37*, the next resupply ship, was launched July 19.
Aug. 29–Dec. 21, 1988	*Soyuz TM-6* Vladimir Lyakhov Valery Polyakov Abdullah Ahad Mohmand	14,400 lb 138 rev 221 mi	8 days 20 h 27 min/Lyakov and Mohmand; 241 days/ Polyakov	*TM-6* was launched for a trip to *Mir*, with a crew including a physician, Polyakov, who was to stay aboard *Mir* to observe Titov and Manarov, and an Afghan cosmonaut, Mohmand. Docking occurred on Aug. 31. Since Lyakhov had received special training as a rescuer-cosmonaut, he was able to handle *TM-6* alone. After undocking from *Mir* in *Soyuz TM-5* of the previous mission, Lyakhov and Mohmand prepared for reentry at 3 a.m. but became stranded in orbit due to malfunction of a guidance system sensor, followed by computer input errors. With oxygen supply threatening to run out after 48 h, the twice-failed retroburn was achieved 24 h later after preventive measures. *TM-5* landed on Sept. 7, 100 mi from Dzhezkazgan in Kazakhstan. On Sept. 12, the crewless resupply ship *Progress 38* docked with *Mir* after *Soyuz TM-6* had been moved, on Sept. 8, from the *Kvant* astrophysics module to the airlock of the station. *Progress 39* was launched Sept. 9, carrying French equipment for an upcoming visit by a French cosmonaut.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Aug. 29–Dec. 21, 1988 (*cont.*)				In a 4-h EVA on Oct. 20, Titov and Manarov successfully repaired a telescope in the *Kvant* module, finishing an earlier attempt of June 30.
Sept. 29–Oct. 3, 1988	*STS 26*[†] Frederick H. Hauck Richard O. Covey John M. (Mike) Lounge David C. Hilmers George D. Nelson	253,700 lb 645 rev 188 mi	4 days 1 h 57 min	The United States' return to space. After a 32-month hiatus, NASA astronauts were back in space on the orbiter *Discovery*, after a flawless launch at 11:37 a.m. EDT. At 6 h 13 min into the mission Nelson deployed the prime payload, NASA's second *Tracking and Data Relay Satellite* (*TDRS*) for transfer to geostationary orbit. Secondary payloads were 11 experiments, including 2 by students. It was the seventh flight of *Discovery*, with over 200 engineering changes since the *Challenger* loss. Touchdown occurred at 12:37 P.M. EDT at Edwards, with a rollout distance of 7451 ft.
Nov. 15, 1988	*Buran* Crewless	224,000 lb 2 rev 156 mi	3 h 25 min	First flight of Soviet reusable space shuttle *Buran* (Blizzard). Launched by second flight of Energia launch vehicle at 6 a.m. local time, landed two orbits later on runway at Baikonur, 7.5 mi from launch pad.
Nov. 26, 1988–Apr. 27, 1989	*Soyuz TM-7* Alexandrovitch Volkov Sergei Krikalev Jean-Loup Chrétien	14,400 lb — 230 mi	25 days/ Chrétien; 152 days/ Volkov and Krikalev	Second flight of French cosmonaut Chrétien to a Soviet space station. Mission, named "Aragatz," docked on *Mir* on Nov. 28. Chrétien became first French citizen to perform an EVA when he installed the deployable antenna structure ERA in space on Dec. 9. On Dec. 21, Chrétien returned in *Soyuz TM-6* with the long-duration crew, Titov and Manarov, who on this day completed 366 days on *Mir*. Volkov and Krikalev stayed in space with Polyakov until Apr. 27, 1989, when they returned in *TM-7*, leaving *Mir* crewless for the first time in 2 years. During its occupied period, the space station had been resupplied by a total of 17 *Progress* flights with about 40 tons of goods.
Dec. 2–6, 1988	*STS 27* Robert L. Gibson Guy S. Gardner Mike Mullane Jerry L. Ross William M. Shepherd	— 70(?) rev —	4 days 9 h 6 min 19 s	Launched as second flight after the *Challenger* accident, *Atlantis* carried a radar imaging reconnaissance satellite on a classified Department of Defense mission into a high-inclination orbit. Returned to Edwards after highly successful mission.
March 13–18, 1989	*STS 29* Michael L. Coats John E. Blaha James F. Buchli Robert C. Springer James P. Bagian	263,290 lb 79 rev 206 mi	4 days 23 h 39 min	*Discovery*, on this third flight after the *Challenger* disaster, deployed its main payload, the third *Tracking and Data Relay Satellite*, *TDRS-D*, completing the constellation of NASA's new communications system. After conducting a number of physical, biological, medical, and materials experiment, *Discovery* returned to Edwards AFB.
May 4–8, 1989	*STS 30* David M. Walker Ronald J. Grabe Norman E. Thagard Mary L. Cleave Mark C. Lee	217,513 lb 64 rev 185 mi	4 days 56 min 38 s	Originally scheduled for Apr. 28, the launch of *Atlantis* was aborted at T-31 due to failure of the hydrogen recirculation pump of a main engine and rescheduled for May 4. On that day, *Atlantis* lifted off and, $6\frac{1}{2}$ h later, placed the Venus radar mapper probe *Magellan* on its trajectory to the planet Venus, the first interplanetary probe launched by the United States in nearly 11 years. *Atlantis* landed at Edwards after a highly successful mission.
Aug. 8–13, 1989	*STS 28* Brewster H. Shaw, Jr. Richard N. Richards James C. Adamson Mark N. Brown David C. Leestma	— 80 rev 192 mi	5 days 1 h 56 s	Fifth flight after the *Challenger* accident. The extensively refurbished *Columbia*, on its first trip after over 3 years of inactivity, carried a classified Department of Defense payload into a 57° inclination orbit. Landing at Edwards took place after a successful mission.
Sept. 6, 1989–Feb. 19, 1990	*Soyuz TM-8* Alexander Viktorenko Alexander Serebrov	14,400 lb Approx. 2662 rev 230 mi	166 days 7 h 58 min	After a pause of almost 5 months in its crewed program, the Soviet Union launched *TM-8* to *Mir*, the 67th crewed flight of the Soviet Union, carrying *Mir*'s 20th and 21st occupants. Docking on *Mir* occurred on Sept. 8, executed manually after a last-minute technical problem with the automatic rendezvous/docking system of the *TM-8*. The arrival of the crew had been preceded by the docking of *Progress M1*, the first of a new series of crewless resupply craft, able to carry 5952 lb of cargo (882 lb more than the earlier version), and equipped with a recoverable entry capsule (not used on this flight).

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

[†] After the *Challenger* accident, shuttle mission designations returned to consecutive numbering. Unavoidable changes in flight assignments, however, resulted in numbered flights not being flown in sequence.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Oct. 18–23, 1989	*STS 34* Donald E. Williams Michael J. McCulley Shannon W. Lucid Franklin R. Chang-Diaz Ellen S. Baker	264,775 lb 79 rev 206 mi	4 days 23 h 40 min	Launch of *Atlantis* with the *Galileo* Jupiter probe as primary payload occurred after two launch postponements, one from Oct. 12 due to a main engine controller problem, the other from Oct. 17 due to inclement weather. At 7:15 P.M., the 38,500-lb payload was successfully deployed; at 8:15 P.M. the Inertial Upper Stage (IUS) first stage was fired, followed by second-stage ignition at 8:20. The 2-ton *Galileo* separated from the IUS at 9:05 with deployed booms, was spun up to 2.8 revolutions per min at 9:30 P.M., and began its 6-year voyage to Jupiter, which includes one Venus and two Earth flybys for gravity assist. After conducting in-space experiments on material processing, plant life, ozone monitoring, ice crystals, and human physiology, *Atlantis* returned to Earth two orbits early due to adverse wind at the landing site.
Nov. 22–27, 1989	*STS 33* Frederick D. Gregory John E. Blaha Manley L. (Sonny) Carter Kathryn C. Thornton F. Story Musgrave	— 78(?) rev —	5 days 7 min	On the seventh shuttle flight after the *Challenger* accident, the *Discovery* lifted off in darkness, carrying a classified military payload into a low-inclination orbit. After a successful mission, the vehicle landed at Edwards. It was the third night launch of the shuttle program.
Nov. 26, 1989	*Kvant 2* Crewless	43,000 lb — 250 mi	—	*Kvant 2*, the second of four large building blocks of the Soviet *Mir* complex, was launched on a Proton. It docked to *Mir* on Dec. 6, after a lengthy delay caused by a multistep trouble-shooting program to deploy a stuck solar array and subsequent rendezvous problems. After the successful docking. *Kvant 2*, which carried systems and equipment to expand *Mir*'s operational capability and a crewed maneuvering unit for extravehicular activity, generated 3.5 kW of power from each of its two solar arrays.
Jan. 9–20, 1990	*STS 32* Daniel C. Brandenstein James D. Wetherbee Bonnie J. Dunbar Marsha S. Ivins G. David Low	229,848 lb 172 rev 224 mi	10 days 21 h 52 s	Longest shuttle mission to date. Originally scheduled for Dec. 18, 1989, but delayed due to unfinished launch pad work and one weather-caused slip (on Jan. 8). Launched on its mission to deploy the Navy communications satellite *Syncom IV-F5* (on Jan. 10) and retrieve the Long Duration Exposure Facility (LDEF), a 12-sided, open-grid aluminum satellite, 30 ft long, 14 ft in diameter, and 10.5 tons in mass, carrying 57 experiments. LDEF was originally orbited by *Challenger* (*STS 41C*) on Apr. 6, 1984, but its planned retrieval 1 year later was delayed by schedule slips and the *Challenger* accident. After *Columbia*'s rendezvous with LDEF on Jan. 12, Mission Specialist Dunbar used the Remote Manipulator System arm to grapple the research facility, 2093 days and 32,419 orbits after LDEF's launch. *Columbia* landed at Edwards, delayed 2 days by ground fog. Heaviest landing weight (by 5 tons) to date.
Feb. 11–Aug. 9, 1990	*Soyuz TM-9* Anatoliy Solovyev Alexander Balandin	14,400 lb 230 rev —	179 days	Docked with *Mir* on Feb. 13 to relieve Viktorenko and Serebrov. Earlier in February, they had conducted several spacewalks (EVA) to test a crewed maneuvering unit named *Ikar* (Icarus): Serebrov on four flights on Feb. 1 during a 5-h EVA, Viktorenko several times on a 3-h 45-min EVA on Feb. 5. The cosmonauts also tested a modified space suit and portable life-support system. The Soviet crewed maneuvering unit weighed 440 lb and was propelled by 36 compressed-air thrusters. Solovyev and Balantin became the 22d and 23d occupants of *Mir*. Viktorenko and Serebrov returned to Earth on Feb. 19, landing in *Soyuz TM-8* near Arkalyk, 205 mi northeast of Baikonur, in good health after 166 days in space, including five spacewalks and dozens of experiment programs.
Feb. 28–Mar. 4, 1990	*STS 36* John O. Creighton John H. Casper Michael Mullane David C. Hilmers Pierre J. Thuot	— 71 rev —	4 days 10 h 19 min	Launched as ninth flight since *Challenger* loss, after five delays due to head cold of the Commander, bad weather, and a ground computer glitch. *Atlantis* deployed a secret intelligence-gathering satellite on a Department of Defense mission in a highly inclined orbit (62°). Returned to Edwards AFB after completing a fully successful mission.
Apr. 24–29, 1990	*STS 31* Loren J. Shriver Charles F. Bolden Steven A. Hawley	259,229 lb 76 rev 381 mi	5 days 1 h 15 min	The long-awaited launch of the *Hubble Space Telescope*, 20 years in planning and development, took place after an earlier launch attempt, on Apr. 10, was scrubbed due to malfunction of a valve in one of *Discovery*'s auxiliary power units. The

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Apr. 24–29, 1990 (*cont.*)	Bruce McCandless Kathryn D. Sullivan			12.5-ton spacecraft, 42.5 × 14 ft, was placed into space by the manipulator arm on Apr. 25 at 8:45 a.m. EDT and released, after deployment of its solar arrays and a checkout, at 3:38 p.m. For the remainder of the mission, the crew carried out medical and engineering tests in the shuttle middeck. *Discovery* landed at Edwards AFB.
May 31, 1990	*Kristall* Crewless	43,000 lb 250 rev —	—	As third major addition to *Mir*, the technology module *Kristall* docked to the Soviet space station on June 10. Nearly identical to the previous module, *Kvant 2*, with length of 45 ft, width of 14.3 ft, and volume of 2000 ft^3, *Kristall* forms the second of four spokes of the *Mir* complex. Equipped with furnaces for materials experiments in space, it also has a docking port for the *Buran* space shuttle. Docking took two attempts, using the backup system to a failed maneuvering jet. On June 11, *Kristall* was moved by the Ljappa pivot arm from forward docking port to a lateral port opposite the *Kvant 2* module. With core, *Kvant 1*, *Kvant 2*, *Kristall*, and *Soyuz-TM*, *Mir* now weighed 183,000 lb. *Kristall* also carried equipment to repair loose insulation panels on *Soyuz TM-9*. Repair took place on July 17, but Solovyev and Balandin encountered difficulties in closing the *Kvant 2* outer airlock door, damaged on exit, and exceeded the life-support capability of their spacesuits. Crew used emergency entry method (depressurizing *Kvant 2*), conducted 3-h EVA to repair hatch on July 26, and returned to Earth on Aug. 9, one week after arrival of *TM-10*, landing at Arkalyk in Kazakhstan with 286 lb of material.
Aug. 1–Dec. 10, 1990	*Soyuz TM-10* Gennaidy Manakov Gennaidy Strekalov	14,400 lb Approx. 2096 rev 235 mi	131 days	Launched to replace *TM-9*, it docked with *Mir* on Aug. 3 to continue work of previous crew (which had completed 506 of 520 planned science experiments), concentrating mostly on materials manufacturing with the *Kristall* module and switching solar panels from *Kristall* to *Kvant 1*. Crewless *Progress M-4* was launched Aug. 15 and docked Aug. 17, to bring supplies and boost *Mir* again up to 248 mi altitude. *Progress M-4* was used to check out a new feature, a returnable capsule for returning material to Earth from *Mir*, prior to its use on *Progress M-5*. Manakov and Strekalov returned on Dec. 10.
Oct. 6–10, 1990	*STS 41* Richard N. Richards Robert D. Cabana Bruce E. Meinick William M. Shepherd Thomas D. Akers	256,330 lb 65 rev 206 mi	4 days 2 h 10 min	Replacing the postponed *STS 35* (*Columbia*), *Discovery* was launched with the European-built out-of-the-ecliptic solar probe *Ulysses* plus additional scientific experiments on board. *Ulysses* was deployed during the fifth orbit and launched by its own IUS/PAM rocket system. The fastest artificial object, it started its 5-year mission via a gravity-assist slingshot maneuver around Jupiter in Feb. 1992. It passed out of the Sun in summer 1994 and spring 1995, respectively.
Nov. 15–20, 1990	*STS 38* Richard O. Covey Frank L. Culbertson Carl J. Meade Robert C. Springer Charles D. ("Sam") Gemar	— 79 rev 130 mi	4 days 21 h 54 min 32 s	On this fifth night launch, *Atlantis*, with a classified DOD payload, lifted off at 6:48 p.m. EST on it seventh flight. The planned landing at Edwards on Nov. 19 was canceled due to high winds and look place 1 day later at Kennedy at about 4:43 p.m. EST. It was the sixth Kennedy landing, $5^1/_2$ years after the last one (*Discovery*, on Apr. 5, 1985).
Dec. 2–11, 1990	*STS 35* Vance D. Brand Guy S. Gardner John M. Lounge Robert A. R. Parker Ronald A. Parise Jeffrey A. Hoffman Samuel T. Durrance	267,513 lb 143 rev 220 mi	8 days 23 h 6 min 5 s	The first shuttle mission in 7 years dedicated to a single science discipline: astrophysics. After 6 months delay due to a faulty cooling system valve, hydrogen leaks, and payload electronics repair, *Columbia*'s tenth flight took off with ASTRO 1 (two *Spacelab* pallets with three ultraviolet telescopes on a pointing system, and one x-ray telescope) to explore the hottest, albeit invisible parts of the universe from the Orbiter cargo bay. The flight featured more than 80 h of observations (about 394) of about 135 science targets in the universe, starting 23 h after orbit insertion and ending 4 h before deorbit burn. Also, an experiment of communicating with amateur radio stations within line-of-sight of *Columbia* in voice and data mode (SAREX) was conducted by Parise. During orbit 39 on Dec. 4, *Columbia* and *Mir* were only 31 mi apart, and *Mir* was sighted twice from the shuttle. *STS 35* landed at Edwards one day earlier than planned because of bad weather that was forecast for later, after a mission troubled by computer, pointing, and plumbing problems.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Dec. 2–May 26, 1990	*Soyuz TM-11* Viktor Afanasyev Musa Manarov Toyohiro Akiyama	14,400 lb Approx. 2740 rev 230 mi	8 days/Akiyama; 131 days/ Manakov and Strekalov	Lifting off with the first Japanese in space on board, a journalist, *Soyuz TM-11* docked at *Mir* 2 days later, making up a record number of 12 humans in space at the same time (5 in *Mir*, 7 in *STS 35*). Cosmonauts Manakov, Strekalov, and Akiyama returned on Dec. 10, in *TM-10* with Afanasyev and Manarov taking over on *Mir*, Resupply ship *Progress M7*, launched Mar. 19, failed twice to dock to *Mir*'s *Kvant* module. On the second attempt on Mar. 23, ground controllers barely avoided a collision between *Mir* and *M7*. On March 26, the crew relocated *Soyuz TM-11* from the main port to the *Kvant 1* module and docked *M7* successfully to that port on Mar. 28.
Apr. 5–11, 1991	*STS 37* Steven Nagel Kenneth D. Cameron Jerry L. Ross Jerome Apt Linda M. Godwin	240,809 lb 92 rev 281 mi	5 days 23 h 32 min	*Atlantis* lifted off with the 35,000-lb *Gamma Ray Observatory* (*GRO*), the heaviest NASA science satellite ever carried by a shuttle. Equipped with four instruments to observe celestial high-energy sources with 10 times the sensitivity of prior experiments, *GRO* was set free on April 7 after an unplanned emergency EVA by Ross and Apt to deploy *GRO*'s stuck high-gain antenna—the first United States spacewalk since Dec. 1985. On Apr. 8, Ross and Apt performed a scheduled 6-h EVA to test various crew equipment translation aids such as carts on a track, for use on the exterior of the space station.
Apr. 28–May 6, 1991	*STS 39* Michael L. Coats L. Blaine Hammond, Jr. Gregory J. Harbaugh Donald R. McMonagle Guion S. Bluford C. Lacy Veach Richard J. Hieb	236,623 lb 133 rev 162 mi	8 days 7 h 22 min	*Discovery* made the first unclassified DOD-dedicated shuttle mission. The mission was highlighted by around-the-clock observations of the atmosphere, gas release from subsatellites, shuttle engine firings, and the shuttle's orbital environment in wavelengths ranging from infared to the far-ultraviolet, including the aurora australis.
May 18–Oct. 10, 1991	*Soyuz TM-12* Anatoly Artsebarsky Sergei Krikalev Helen Sharman	14,400 lb Approx. 2265 rev 230 mi	7 days 19 h 5 min/ Sharman; 175 days 2 h 48 min/ Afanasyev and Manarov	With the first British space traveler, food technologist Helen Sharman, as part of Britain's commercial project Juno, *Soyuz* docked at *Mir* on May 20, the ninth crew to visit the station. Earlier, on Apr. 25, the *Mir* crew, during a 5-h EVA, inspected the automatic docking system antenna that had nearly caused the collision with the *Progress M7* automated transport on Mar. 23. *M7* was discarded to burn up in the atmosphere. Sharman, Afanasyov, and Manarov returned on May 26 in *TM-11*, landing in Kazakhstan. In the first 3 months of their stay, Artsebarsky and Krikalev made six spacewalks totalling a record 32 h, for such tasks as repairing a docking antenna and a television camera, deploying instruments, and assembling a 46-ft-tall towerlike structure called Sofora, to carry reaction jets, on *Mir*. On Aug. 16, *Progress M8* was separated from *Mir*, and on Aug. 20 new supplies were brought up on *M9* in a night launch. *Soyuz TM-12* returned on Oct. 10 with Artsebarsky, Aubakirov, and Viehboeck.
June 5–14, 1991	*STS 40* Bryan D. O'Connor Sidney M. Gutierrez Tamara E. Jernigan M. Rhea Seddon James P. Bagian F. Drew Gaffney Millie Huges-Fulford	241,175 lb 145 rev 185 mi	9 days 2 h 15 min	*Columbia* carried Spacelab Life Sciences 1 (SLS 1), with three medical doctors, one biochemistry researcher, and one physicist on board, to perform 18 major experiments on the cardiovascular, renal/endocrine, blood, immune, musculoskeletal, and neurovestibular systems. Test objects included white rats and jellyfish.
Aug. 2–11, 1991	*STS 43* John E. Blaha Michael A. Baker Shannon W. Lucid G. David Low James C. Adamson	247,965 lb 140 rev 207 mi	8 days 21 h 21 min	The *Atlantis* flight put NASA's fourth *Tracking and Data Relay Satellite* (*TDRS-E*) into space, boosted successfully to a geosynchronous orbit by an Inertial Upper Stage (IUS). Other payloads included an atmospheric ultraviolet calibration instrument to measure the Earth's ozone layer, a heat pipe experiment, and an optical communications experiment. *Atlantis* landed at Kennedy.
Sept. 12–18, 1991	*STS 48* John O. Creighton Kenneth S. Reightler Mark N. Brown James F. Buchli Charles D. Gemar	224,141 lb 81 rev 356 mi	5 days 8 h 27 min	*Discovery* carried the giant *Upper Atmosphere Research Satellite* (*UARS*) into a 57°-inclination orbit, the first major flight element of NASA's Mission to Planet Earth program. The 14,400-lb satellite carried nine complementary scientific instruments to provide a more complete understanding of the chemistry, dynamics, and energy input of the upper atmosphere, including the ozone layer, plus a radiometer measuring the Sun's energy output. Other on-board work included microgravity research on crystal growth, fluids and structures behavior, life sciences, and radiation effects.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Oct. 2–Mar. 25, 1991	*Soyuz TM-13* Alexander Volkov Toktar Aubakirov Franz Viehboeck	14,400 lb Approx. 2740 rev 230 mi	8 days/ Viehboeck and Aubakirov; 144 days 19 h/ Artsebarsky	*Soyuz* docked with *Mir* on Oct. 4. Cosmonauts Artsebarsky, Viehboeck (first space traveler from Austria), and Aubakirov (first Kazakh crew member) returned on Oct. 10 in *TM-12,* leaving Volkov and Krikalev in control of the space station.
Nov. 24–Dec. 1, 1991	*STS 44* Frederick D. Gregory Terence T. Henricks Story Musgrave Mario Runco, Jr. James S. Voss Thomas J. Hennen	247,087 lb 108 rev 195 mi	6 days 22 h 51 min	*Atlantis*, on a DOD-dedicated mission, carried a Defense Support Program (DSP) satellite for detection, from geosynchronous orbit, of nuclear detonations, and missile and space launches. After its deployment on the first day of the mission, the crew worked with a variety of secondary payloads, including naked-eye Earth observations, radiation measurements, and medical investigations. Due to malfunction of one of three Inertial Measurement Units (IMUs), the flight had to be shortened by 3 days.
Jan. 22–30, 1992	*STS 42* Ronald J. Grabe Stephen S. Oswald Norman E. Thagard William F. Readdy David C. Hilmers Roberta L. Bondar Ulf D. Merbold	231,497 lb 128 rev 188 mi	8 days 1 h 16 min	*Discovery* carried the International Microgravity Laboratory 1 (IML 1), dedicated to worldwide research in the behavior of materials and life in weightlessness. 225 scientists from 15 countries, experimenting in 42 research disciplines, cooperated in this fight. Along with IML 1, the payload consisted of 12 Get Away Special (GAS) containers, the large-format IMAX camera, and two student experiments. The crew included foreign astronauts Bondar (Canada) and Merbold (Germany).
Mar. 17–Aug. 10, 1992	*Soyuz TM-14* Alexander Viktorenko Alexander Kaleri Klaus-Dietrich Flade	14,400 lb Approx. 2280 rev 245 mi	312 days 6 h 1 min/Krikalev; 175 days 3 h 52 min/Volkov; 7 days 21 h 57 min/Flade	The crew, Including the German guest cosmonaut Flade, docked with *Mir* on Mar. 19. Flade conducted a crowded science program and returned with former *Mir* occupants Volkov and Krikalev on Mar. 25, to land in *Soyuz TM-13* in Kazakhstan, leaving Viktorenko and Kaleri in *Mir*. In space since May 18, 1991, Krikalev was originally slated to return in late 1991, but the schedule was revised, leaving him in orbit for five extra months. On Feb. 20–21 he conducted his seventh spacewalk, fetching experiments left outside during previous EVAs, and setting a new record.
Mar. 24–Apr. 2, 1992	*STS 45* Charles Bolden Brian Duffy C. Michael Foale David Leestma Kathryn Sullivan Dirk Frimout Byron Lichtenberg	222,086 lb 142 rev 182 mi	8 days 22 h 9 min 20 s	The multinational flight of *Atlantis*, including a Belgian payload specialist (Frimout), carried the first ATLAS (Atmospheric Laboratory for Applications and Science) associated with NASA's Mission to Planet Earth initiative. The ATLAS payload of 14 instruments remained attached to the Orbiter, most of them mounted on two *Spacelab* pallets. A Far Ultraviolet Space Telescope (FAUST) was also employed.
May 7–16, 1992	*STS 49* Daniel C. Brandenstein Kevin P. Chilton Richard J. Hieb Bruce E. Melnick Pierre J. Thuout Kathryn C. Thornton Thomas D. Akers	246,008 lb 140 rev 181 mi	8 days 21 h 18 min 36 s	This was the first mission featuring four EVAs, and the maiden voyage of the orbiter *Endeavour*. The main mission was reboosting the 11.7-ft-diameter, 17.5-ft-high communications satellite *INTELSAT VI* (F-3) launched 2 years earlier by a Titan rocket, but stranded in a useless orbit when the Titan's second stage failed to separate. After rendezvous with the satellite, Thuot and Hieb, on two spacewalks on flight days 4 and 5, failed to capture the 8960-lb satellite. Joined by Akers on May 13, however, they succeeded in grasping it with their hands, berthing it in the shuttle cargo bay with the remote manipulator arm, attaching it to a new perigee kick motor, and redeploying it. This first three-person EVA in history lasted 8 h 36 min, the longest spacewalk ever conducted. The solid-fuel motor was fired remotely from the ground on May 14, taking *INTELSAT VI* to its final destination in a geosynchronous orbit. In a fourth EVA on May 14, Thornton and Akers evaluated space-station assembly equipment and techniques as well as several prototype crew self-rescue devices.
June 25–July 9, 1992	*STS 50* Richard N. Richards Kenneth D. Bowersox Bonnie J. Dunbar Carl J. Meade Ellen S. Baker Lawence J. DeLucas Eugene H. Trinh	245,902 lb 220 rev 185 mi	13 days 19 h 30 min 30 s	The first extended-duration (EDO) shuttle flight. *Columbia*, specially modified for 2-week missions, carried the first U.S. Microgravity Laboratory 1 (USML 1), a *Spacelab* module equipped with 31 experiments for advanced research in zero *g*. The longest shuttle flight to date, it supported investigations of the effects of weightlessness on plants, humans, and materials. The mission also carried investigations into polymer membrane manufacturing, the Space Shuttle Amateur Radio Experiment II (SAREX-II), and numerous other cargo bay, middeck, and secondary payloads.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
July 27–Feb. 1, 1992	*Soyuz TM-15* Anatoliy Solovyev Sergei Avdeyev Michel Tognini	14,400 lb 2946 rev 245 mi	146 days/Kaleri and Viktorenko; 14/ Tognini	The third Russian-French mission since 1982, *Soyuz* carried French researcher Tognini, to conduct 10 biomedical, fluid and materials science and technological experiments on board *Mir*. Tognini returned on Aug. 10 in *Soyuz TM-14* along with Viktorenko and Kaleri. At touchdown in Kazakhstan, the spherical *TM 14* rolled over, briefly trapping but not injuring the crew. Solovyev and Avdeyev remained in *Mir* with numerous tasks such as installing new gyrodynes and rocket thrusters for attitude control. *Progress M-13* was launched on Aug. 14, carrying supplies and medical gear to *Mir*.
July 31–Aug. 8, 1992	*STS 46* Loren J. Shriver Andrew M. Allen Jeffrey A. Hoffman Franklin R. Chang-Diaz Marsha S. Ivins Claude Nicollier Franco Malerba	241,797 lb 126 rev 266 mi	7 days 23 h 14 min 12 s	*Atlantis* successfully deployed (on Aug. 2) the first European Retrievable Carrier (EURECA) platform, carrying seeds, spores, crystals, and other experimental materials for research in zero *g*, to be retrieved about 9 months later. Due to data errors, EURECA's orbital self-transfer to its final altitude of 313 mi occurred later than planned but was accomplished on Aug. 7, Deployment of the main payload, the 1142-lb Italian Tethered Satellite System (TSS 1) on a 12.5-mi-long cable from the *Atlantis*, however, was terminated when a breakdown in the reel mechanism halted the tether after only 850 ft. The crew included two foreign astronauts, Malerba (Italy) and Nicollier (Switzerland) for ESA.
Sept. 12–20, 1992	*STS 47* Robert L. Gibson Curtis L. Brown, Jr. Mark C. Lee Jay Apt N. Jan Davis May C. Jemison Mamoru Mohri	232,661 lb 126 rev 192 mi	7 days 22 h 30 min	*Endeavour* carried the Japan-chartered Spacelab SL-J with 44 materials and life sciences experiments and nine Get Away Special (GAS) payload canisters. The cargo also included the Israeli Space Agency Investigation About Hornets (ISAIAH) experiment and another flight of the Shuttle Amateur Radio Experiment (SAREX). The crew included the first Japanese shuttle astronaut (Mohri).
Oct. 22–Nov. 1, 1992	*STS 52* James D. Wetherbee Michael A. Baker C. Lacy Veach William M. Shepherd Tamara E. Jernigan Steven G. MacLean	239,178 lb 158 rev 189 mi	9 days 20 h 56 min	*Columbia* carried the U.S. Microgravity Payload (USMP 1) as main cargo. The crew, with the third Canadian astronaut (MacLean), performed materials and science experiments in micro-*g* with the U.S. Microgravity Payload 2 (USMP 2), deployed the second Italian/U.S. *Laser Geodynamic Satellite* (*LAGEOS 2*), and conducted other physical experiments including the Canadian Experiment #2 (CANEX 2), and space-station-related investigations.
Dec. 2–9, 1992	*STS 53* David M. Walker Robert D. Cabana Guion S. Bluford James S. Voss Michael Richard Clifford	230,731 lb 116 rev 237 mi	7 days 7 h 20 min	As the last dedicated DOD mission, *Discovery* carried the satellite *DOD 1* as primary and two major experiments, the Glow/ Cryogenic Heat Pipe Experiment (GCP) and Orbital Debris Radar Calibration Spheres (ODERACS), as secondary payloads. With GCP, remotely commanded special sensors observed orbiter surfaces, contamination events, and airglow, and measured the zero-*g* performance of liquid oxygen heat pipes. ODERACS had to be canceled when its container in the cargo bay failed to open.
Jan. 13–19, 1993	*STS 54* John H. Casper Donald R. McMonagle Mario Runco, Jr. Susan J. Helms Gregory J. Harbaugh	248,338 lb 188 rev 95 mi	5 days 23 h 39 min	*Endeavour* carried the sixth *Tracking and Data Relay Satellite* (*TDRS*), its Inertial Upper Stage (IUS), and the Diffuse X-Ray Spectrometer (DXS), consisting of two instruments attached to the sides of the cargo bay. After successful deployment of *TDRS-F*, crew activities included several microgravity experiments on materials and rodents, and the first in a series of test spacewalks leading up to space-station construction. Harbaugh and Runco, assisted by Helms from the cabin, tested EVA training procedures for the space station and the *Hubble* repair mission (*STS 61*). Using children's toys, the crew responded to students from various elementary schools asking questions about physics in zero *g*.
Jan. 24–July 22, 1993	*Soyuz TM-16* Gennadiy Manakov Alexander Polishchuk	14,400 lb Approx. 2790 rev 242 mi	189 days/ Solovyev and Avdeyev	Equipped with the new androgynous docking system APAS-89, to be tested for use in the newly planned *Mir*/space shuttle docking mission, *Soyuz* docked with *Mir*, for the first time, at the *Kristall* module's docking port. The previous crew, Solovyev and Avdeyev, returned on Feb. 1 in *Soyuz TM-15* after 189 days in space. On Feb. 4, *Progress M-15* was undocked from *Mir*'s aft port and used to deploy, by centrifugal force, a 66-ft-diameter solar sail made from superthin aluminum foil. The purpose of the test, which reflected a visible light beam to Earth, was to assess the force of photons from the Sun for propulsive application.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Apr. 8–17, 1993	*STS 56* Kenneth D. Cameron Stephen S. Oswald Michael Foale Kenneth D. Cockrell Ellen Ochoa	225,597 lb 147 rev 185 mi	9 days 6 h 8 min	*Discovery* carried ATLAS 2, the second in NASA's series of Atmospheric Laboratory for Applications and Science missions. The remote sensing instruments of ATLAS studied the Sun's energy output and Earth's middle-atmosphere chemical makeup, particularly with regard to ozone depletion. The crew on Apr. 11 deployed Spartan-201, a free-flying platform carrying instruments to study the Sun's corona and the solar wind, and recovered the platform on Apr. 13.
Apr. 26–May 6, 1993	*STS 55* Steven R. Nagel Terence T. Henricks Jerry L. Ross Charles J. Precourt Bernard A. Harris Ulrich Walter Hans Wihelm Schlegel	244,156 lb 158 rev 187 mi	9 days 23 h 40 min	*Columbia* carried the *Spacelab* module on its eleventh flight, chartered for the second time by the German space agency DARA, with two German payload specialists (Walter, Schlegel) among its crew. The Spacelab D2 mission was devoted to zero-*g*-relevant investigations in a wide range of science disciplines, including materials research; life sciences; technology experiments in optics, robotics, automation, and medical applications; Earth observations; atmospheric physics; and astronomy, for a total of about 90 different projects. Many investigations were aimed at developing, testing, and verifying future operations/utilization in permanent space-station laboratories.
June 21–July 1, 1993	*STS 57* Ronald Grabe Brian Duffy David Low Nancy Sherlock Peter (Jeff) J. K. Wisoff Janice Voss	239,319 lb 158 rev 289 mi	9 days 23 h 46 min	Highlights of the *Endeavour's* flight were the first flight of the privately developed middeck module *Spacehab* and the retrieval of the European space platform EURECA, in orbit since its deployment from *STS 46* on Aug. 1, 1992. *Spacehab* experiments included 13 commercial payloads, ranging from drug improvement to feeding plants, cell splitting, the first soldering experiment in space by United States astronauts, and high-temperature melting of metals; and six NASA experiments involving biotechnology, human factors, and water recycling for the future space station. Other payloads included SHOOT (Superfluid Helium On-Orbit Transfer) to demonstrate the resupply of liquid helium containers in space. A 5½-h EVA by Low and Wisoff on June 25 served to latch two EURECA antennas and to refine procedures for the planned *Hubble* repair mission.
July 1–Jan. 14, 1993	*Soyuz TM-17* Vasili Tsibliyev Alexander Serebrov Jean-Pierre Haignere	14,400 lb Approx. 2790 rev 242 mi	179 days/ Polishchuk and Manakov; 21 days/ Haignere	Carrying the fourth French cosmonaut since 1982, *Soyuz* docked with *Mir* on July 3. Resupply craft *Progress M-18*, docked on May 24, was undocked on July 4 and reentered destructively, after separating a small ballistic Raduga ("Rainbow") reentry capsule carrying experiments and recordings for a soft landing in Russia. During his stay at the station, Haignere conducted 11 major Altair experiments, then returned to Earth on July 22 with the previous crew, Polishchuk and Manakov, in *TM-16*.
Sept. 12–22, 1993	*STS 51* Frank Culbertson William Readdy Jim Newman Dan Bursch Carl Walz	250,023 lb 157 rev 296 mi	9 days 20 h 11 min	*Discovery* carried two payloads. An *Advanced Communication Technology Satellite* (*ACTS*), a testbed for technology leading to a new generation of comsats, was deployed and injected by a Transfer Orbit Stage (TOS) toward its geosynchronous location. The *German Orbiting and Retrievable Far and Extreme Ultraviolet Spectrometer–Shuttle Pallet Satellite* (*ORFEUS–SPAS*) was set free by the robot arm and was retrieved 6 days later, the first of a series of such missions. Other mission tasks included a 7-h EVA by Walz and Newman on Sept. 16 as part of planning for the *Hubble* repair mission.
Oct. 18–Nov. 1, 1993	*STS 58* John E. Blaha Richard A. Searfoss M. Rhea Seddon William S. McArthur, Jr. David A. Wolf Shannon W. Lucid Martin J. Fettman	244,860 lb 224 rev 179 mi	14 days 13 min 40 s	*Columbia* carried SLS-2, the second *Spacelab* dedicated to life sciences research. NASA's longest shuttle mission to date conducted, in a 39° inclination orbit, a series of experiments to gain more knowledge on how the human body adapts to the weightless environment of space. The 14 experiments were performed on the crew, which included two medical doctors (Seddon, Wolf), a physicist/biochemist (Lucid), and a veterinarian (Fettman), and on laboratory animals (48 rats).
Dec. 2–13, 1993	*STS 61* Richard O. Covey Kenneth D. Bowersox Tom Akers Jeffrey A. Hoffman F. Story Musgrave Claude Nicollier Kathryn Thornton	236,670 lb 162 rev 370 mi	10 days 19 h 58 min 35 s	*Endeavour* made the first in a series of planned visits to the orbiting *Hubble Space Telescope* (*HST*). The crew included four astronauts with considerable EVA experience (Hoffman, Musgrave, Thornton, Akers) and Swiss astronaut Nicollier (ESA). On Dec. 4, Nicollier grappled the *HST* with the shuttle's robot arm, berthing it in the cargo bay. In five space walks during the following 5 days, the astronauts installed the Wide Field/Planetary Camera II and the Corrective Optics Space Telescope Axial Replacement (COSTAR)

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Dec. 2–13, 1993 (*cont.*)				mechanism, both designed to compensate for the erroneous spherical aberration of the primary mirror. They also replaced the spacecraft's solar arrays and installed new gyroscopes, additional computer memory, and other systems to increase *HST* reliability. Total EVA time accumulated by all crews: 35 h 28 min. *HST* was set free on Dec. 10.
Jan. 8–July 9, 1994	*Soyuz TM-18* Viktor Afanasyev Yuri Usachov Valery Polyakov	14,400 lb 3072 rev 242 mi	197 days 3 h 2 min/ Tsibliyev and Serebrov; 437 days 17 h 59 min/ Polyakov	This was the fifteenth mission to the *Mir* station. Docking occurred on Jan. 10. Polyakov, a physician cosmonaut who had spent 241 days in space in 1988 (*TM-6*), intended to conduct the longest human space flight in history to date to test the effects of long periods of weightlessness on the human body. He returned on *TM-20* on Mar. 22, 1995, after 14.6 months in space. During the mission, three *Progress* supply ships were to be sent to *Mir*, which was repaired and maintained by the crews. The previous crew, Vasily Tsibliyev and Alexander Serebrov, returned to Earth on Jan. 14 after 6 months, despite a last-minute slight collision of their *Soyuz TM-17* with *Mir*.
Feb. 3–11, 1994	*STS 60* Charles F. Bolden Kenneth S. Reightler, Jr. N. Jan Davis Ronald M. Sega Franklin R. Chang-Diaz Sergei K. Krikalev	233,290 lb 130 rev 220 mi	8 days 7 h 9 min	This *Discovery* mission included the first Russian cosmonaut (Krikalev). The mission also carried the commercial *Spacehab* facility (on its second flight) with 12 biotechnology and materials processing payloads, and the free-flying Wake Shield Facility, a 12-ft-diameter disk designed to generate an ultravacuum environment in space in its wake, within which to grow thin semiconductor crystal films of gallium arsenide. WSF 1 deployment, however, did not take place due to problems with a horizon sensor on the satellite; crystals were grown with WSF attached to the robot arm. The mission also included deployment of the German satellite *BREMSAT*.
Mar. 4–18, 1994	*STS 62* John H. Casper Andrew M. Allen Charles D. Gemar Marsha S. Ivins Pierre J. Thuot	245,457 lb 223 rev 188 mi	13 days 23 h 17.5 min	This *Columbia* mission was dedicated to microgravity research with the U.S. Microgravity Payload 2 (USMP 2) and the Office of Aeronautics and Space Technology 2 (OAST 2) payloads. The four experiments on USMP 2 were designed to study solidification techniques, crystal growth, and gas behavior in zero *g*. OAST 2 carried five experiments for various other effects of materials and environmental phenomena in space. The Orbiter also carried special acceleration measurement systems, and the Shuttle Solar Backscatter Ultraviolet (SSBUV) instrument.
Apr. 9–20, 1994	*STS 59* Sidney M. Gutierrez Kevin P. Chilton Jerome Apt Michael R. Clifford Linda M. Godwin Thomas D. Jones	237,048 lb 181 rev 139 mi	11 days 5 h 50 min	*Endeavour*, in a 57° orbit, carried the first Space Radar Laboratory (SRL 1) for internationally conducted studies of the Earth's global environment. SRL consisted of the Spaceborne Imaging Radar-C (SIR-C) and the X-Band Synthetic Aperture Radar (X-SAR). Other payloads included microgravity experiments (CONCAP-IV) and medical investigations.
July 1–Nov. 4, 1994	*Soyuz TM-19* Yuri Malenchenko Talgat Musabayev	14,400 lb 1960 rev 242 mi	125 days 20 h 53 min; 182 days/ Afanasyev and Usachov	This mission docked with *Mir* on July 3, carrying a Kazakhstan cosmonaut (Musabayev) besides a Russian. The two crew members replaced Afanasyev and Usachev, who returned to Earth in *Soyuz TM-18* on July 9, after 6 months in space. Valery Polyakov continued his orbital stay. Malenchenko and Musabayev returned in *TM-19* on Nov. 4, 1994 (with Ulf Merbold of *TM-20*).
July 8–23, 1994	*STS 65* Robert D. Cabana James Donald Halsell Richard J. Hieb Carl E. Walz Leroy Chiao Donald A. Thomas Chiaki Naito-Mukai	228,640 lb (landing) 235 rev 188 mi	14 days 17 h 55 min	*Columbia* carried the second mission of the International Microgravity Laboratory (IML 2) with furnaces and other facilities to study the behavior of materials and life in the weightless environment, and to produce a variety of material structures, from crystals to metal alloys. More than 200 scientists contributed 82 investigations, of which 81 were successfully completed in 19 IML 2 facilities, plus other middeck experiments. The crew included the first Japanese woman in space (Naito-Mukai).
Sept. 9–20, 1994	*STS 64* Richard N. Richards L. Blaine Hammond Jerry Linenger Susan J. Helms Carl J. Meade Mark C. Lee	210,916 lb (landing) 176 rev 162 mi	10 days 22 h 50 min	*Discovery* was launched for a mission of atmospheric research, robotic atmospheric research, robotic operations, and an EVA. LITE (Lidar In-Space Technology Experiment) was flown and used in space. Operating from the shuttle payload bay, the neodymium YAG laser, an optical radar, probed the Earth's atmosphere to collect measurements on clouds and airborne dust. Other payloads were the Robot-Operated Processing System

1 lb = 0.45 kg. 1 ml = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.8 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Sept. 9–20, 1994 (*cont.*)				(ROMPS), a Simplified Aid for EVA Rescue (SAFER) backpack used experimentally, and the astronomy satellite *Spartan 201*, deployed and retrieved after 40 h of free flight.
Sept. 30–Oct. 11, 1994	*STS 68* Michael A. Baker Terrence W. Wilcutt Thomas D. Jones Steven L. Smith Daniel W. Bursch Peter (Jeff) J. K. Wisoff	223,040 lb (landing) 182 rev 139 mi	11 days 5 h 46 min	*Endeavour* carried the Space Radar Laboratory into space for the second time. An international team used its instruments to study how the Earth's global environment is changing.
Oct. 4–Mar. 22, 1994	*Soyuz TM-20* Alexander Viktorenko Elena Kondakova Ulf Merbold	14,400 lb 2637 rev 247 mi	31 days 11 h 36 min/Merbold; 169 days 5 h 22 min/ Viktorenko and Kondakova	This mission docked at *Mir* on Oct. 6, its crew including the third Russian woman (Kondakova) since inception of the Soviet/Russian space program, and ESA-astronaut Merbold on his third space trip. *TM-20* was the first of two ESA Euro-Mir missions, and its launch had depended on successful docking of *Progress M 24*, accomplished by *Mir* commander Malenchenko on Sept. 2 manually. Malenchenko and Musabayev then performed two EVAs, on Sept. 9 and 13, to inspect the (undamaged) orbital station where it had been struck by *M 24* on Aug. 30 and by *Soyuz TM-17* after its undocking in January, as well as *Mir's* Safora control mast. On Oct. 11, the crew had to cope with a serious power and attitude control failure of the station's core module. After performing experiments in life sciences, materials processing, and technology development, though limited by a failure of the CSK 1 materials furnace, Merbold, with Malenchenko and Musabayev, returned to *TM-19* on Nov. 14, leaving Kondakova and Viktorenko in space with Polyakov (to return on Mar. 22, 1995).
Nov. 3–14, 1994	*STS 66* Donald R. McMonagle Curtis L. Brown Ellen Ochoa Joseph R. Tanner Jean-François Clervoy Scott E. Parazynski	209,857 lb (landing) 174 rev 193 mi	10 days 22 h 35 min	*Atlantis* carried the Atmospheric Laboratory for Applications and Science (ATLAS) on its third flight for Mission to Planet Earth. The crew, with a French astronaut (Clervoy) for ESA, used the remote-sensing laboratory for studying the Sun's energy output, the middle atmosphere's chemical makeup, and the effects of these factors on global ozone levels. On Nov. 4, they deployed the German-built reusable satellite *CRISTA-SPAS*, carrying cryogenic infrared spectrometers, telescopes, and a high-resolution spectrograph for atmosphere research. They recovered it on Nov. 12, using a new rendezvous method in preparation for later rendezvous/docking flights to the Russian *Mir* station.
Feb. 3–11, 1995	*STS 63* James D. Wetherbee Eileen M. Collins Bernard A. Harris C. Michael Foale Janice Voss Vladimir G. Titov	211,318 lb (landing) 129 rev 250 mi	8 days 6 h 28 min 15 s	*Discovery* lifted off with the second Russian cosmonaut (Titov) and the first woman shuttle pilot (Collins) within a narrow launch window of 5 min, as required by its mission: rendezvous and fly-around of the Russian space station *Mir* at 51.6° inclination. The rendezvous, a dress rehearsal of docking missions to follow, took place on Feb. 6, validating new flight techniques required for those missions, though no docking was performed as yet. Also on board: *Spacehab* on its third flight and *Spartan 204*, a satellite carrying various science experiments. It was deployed by Titov on Feb. 7 and retrieved with the robot arm on Feb. 9. Foale and Harris performed a 4.5-h EVA on Feb. 9.
Mar. 2–18, 1995	*STS 67* Stephen S Oswald William G. Gregory Tamara E. Jernigan John M. Grunsfeld Wendy B. Lawrence Samuel T. Durrance Ronald A. Parise	217,683 lb (landing) 262 rev 220 mi	16 days 15 h 9 min	Setting a shuttle duration record of about $16\frac{1}{2}$ days in orbit, *Endeavour* carried the second flight of the ASTRO observation platform with three unique telescopes for ultraviolet astronomy. Its tasks included mapping areas of space still largely uncharted in ultraviolet and measuring intergalactic gas abundance. Ultraviolet spectra were taken of about 300 celestial objects, including Jupiter and the Moon. Other experiments concerned large-structure control dynamics and materials processing.
Mar. 14–Sept. 11, 1995	*Soyuz TM-21* Vladimir Dezhurov Gennady Strekalov Norman Thagard	14,400 lb 2820 rev 246 mi	115 days 9 h 44 min	This eighteenth crew visit to *Mir* carried the first American (Thagard) on a Russian spacecraft. Rendezvous and docking took place on Mar. 16. With the current *Mir* crew of 3 cosmonauts and the 7 astronauts on *STS 67*/*Endeavour*, the number of humans simultaneously in space was 13, a record (until the *STS 67* landing on Mar. 18). *Mir 17* crew Viktorenko, Kondakova, and Polyakov returned on Mar. 22 in *TM-20*, with a new long-duration record of 437 days 17 h 59 min set by Polyakov, and longest

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.8 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Mar. 14–Sept. 11, 1995 (*cont.*)				space flight by a woman (Kondakova) of 169 days 5 h 22 min. After a 3-month stay, Thagard, Dezhurov, and Strekalov were to be returned to Earth by the shuttle *Atlantis* (*STS 71*). Known as *Mir 18*, Thagard's mission was part of Phase One of joint activities leading to the *International Space Station.* At the end of his mission, Thagard had established a new United States record for time spent in space of 115 days 9 h 44 min.
May 20, 1995	*Spektr* Crewless	43,000 lb — 257 mi	—	Russia's new four-solar-array research module, launched on a heavy-lift Proton, carried extensive atmospheric science instrumentation and about 1760 lb of United States equipment for cardiovascular and other biomedical research by astronaut Thagard, later by his replacement, Bonnie Dunbar. *Spektr* docked automatically to *Mir* on June 1, requiring five spacewalks by Dezhurov and Strekalov between May 12 and June 2 and the repositioning of the older module, *Kristall*. This prepared the station for the arrival and docking of the shuttle *Atlantis* (*STS 71*).
June 27–July 7, 1995	*STS 71* Robert L. (Hoot) Gibson Charles J. Percourt Ellen S. Baker Gregory J. Harbaugh Bonnie Dunbar Anatoly Y. Solovyev (up) Nikolai M. Budarin (up) Vladimir Dezhurov (down) Gennady Strekalov (down) Norman Thagard (down)	214,700 lb (landing) 153 rev 246 mi	9 days 19 h 22 min 17 s; 75 days 11 h 19 min/ Solovyev and Budarin	*Atlantis* opened Phase One of the *International Space Station* program as the first of a series of shuttle–*Mir* linkups, 20 years after the joint United States–Russian ASTP mission. Docking took place on June 29, forming the largest human-made structure in space. Joint scientific (mostly medical) investigations were carried out inside the *Spacelab* module in the *Atlantis* cargo bay. Solovyev and Budarin then took over as new occupants of *Mir* (*Mir 19*), while the *Mir* 18 crew of Dezhurov, Strekalov, and Thagard returned to Earth with *Atlantis*. Undocking was on July 4, 4 days 21 h 35 min after linkup.
July 13–22, 1995	*STS 70* Terrence (Tom) Henricks Kevin Kregel Donald Thomas Nancy Currie Mary Ellen Weber	195,195 lb (landing) 142 rev 206 mi	8 days 22 h 21 min 2 s	*Discovery* carried the seventh *Tracking and Data Relay Satellite* (*TDRS 7*). It was released out of the payload bay and boosted into geosynchronous orbit with its Inertial Upper Stage (IUS). Other payloads included a Commercial Protein Crystal Growth (CPCG) facility, a Bioreactor Demonstrator, and various biological experiments.
Sept. 3–Feb. 29, 1995	*Soyuz TM-22* Yuri Gidzenko Sergei Avdeyev Thomas Reiter	14,400 lb 2789 rev 246 mi	179 days	EUROMIR'95 docked to *Mir* on Sept. 5, after *Progress M-28* was undocked and jettisoned. ESA astronaul Reiter (Germany) served as flight engineer, marking the first time for Europe in Russia's space program. On Oct. 20, he (with Avdeyev) installed a European experiment on the outside of *Spektr*. Two cassettes exposed to natural and human-made dust and debris were returned by Reiter and Gidzenko during another EVA (the third) on Feb. 8, 1996. The *Mir 19* crew (Solovyev and Budarin) returned in *TM-21* on Sept. 11. On-board activities included experiments in life sciences and micro-*g* materials processing. Most of ESA's equipment had been brought to *Mir* by *Progress M-28* in July and by *Spektr*. Another load arrived on *Progress M-29* (launch Oct. 8, docking Oct. 10, at aft end of *Kvant 1*), including the Munich Space Chair, set up Oct. 11 in *Spektr*. In November, the crew hosted the visiting shuttle *Atlantis/STS 74*. The crew returned 6 days after arrival of *TM-23* with the *Mir 21* crew.
Sept. 7–18, 1995	*STS 69* David M. Walker Kenneth D. Cockrell James S. Voss Jim Newman Michael L. Gernhardt	21,971 lb (landing) 170 rev 232 mi	10 days 20 h 29 min	*Endeavour* carried the *Spartan 210* free-flyer on its third flight and the Wake Shield Facility (WSF 2) on its second flight as main payloads, *Spartan* released on Sept. 8 and retrieved 2 days later, observed the solar wind and the Sun's outer atmosphere in conjunction with the *Ulysses* probe while it passed over the Sun's north polar region. The WSF, deployed by the robot arm on Sept. 11 and retrieved on Sept. 14, was partially successful in growing thin semiconductor films in the ultravacuum created in its wake. Other payloads included the International Extreme Ultraviolet Hitchhiker (IEH 1) and the Capillary Pumped Loop 2/Gas Bridge Assembly (CAPL 2/GBA) for microgravity studies of fluid systems. Voss and Gernhardt on an EVA tested construction tools, an arm-sleeve strap-on computer, and suit modifications for space-station assembly.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Oct. 20–Nov. 5, 1995	*STS 73* Kenneth D. Bowersox Kent Rominger Kathryn Thornton Catherine Coleman Michael Lopez-Alegria Fred Leslie Albert Sacco	230,158 lb (landing) 255 rev 174 mi	15 days 21 h 53 min	*Columbia* carried the second U.S. Microgravity Laboratory (USML 2) Spacelab with science and technology experiments in areas such as fluid physics, materials science, biotechnology, combustion science, and commercial space processing technologies. Students interacted with the crew, discussing and comparing on-board microgravity experiments with similar ground-based experiments.
Nov. 12–20, 1995	*STS 74* Kenneth Cameron James Halsell Chris Hadfield Jerry Ross William McArthur	205,000 (landing) 128 rev 185 mi	8 days 4 h 31 min	*Atlantis* carried a Russian-built Docking Module (DM), two new solar arrays, and supplies for *Mir* on the second shuttle-*Mir* docking mission. After attaching the DM to *Mir*'s *Kristal* module with the robot arm and linking the shuttle's airlock to it on Nov. 15, the *Atlantis* crew, including a Canadian (Hadfield), was welcomed by Gidzenko, Avdeyev, and Reiter. During 3 days of joint operation, the crews transferred the United States biomedical and microgravity science samples and data collected by the *Mir 18*, *19*, and *20* resident crews from the station to the shuttle, and fresh supplies to *Mir*. On Nov. 18, *Atlantis* undocked, leaving the DM attached to the station for future linkups.
Jan. 11–20, 1996	*STS 72* Brian Duffy Brent W. Jett, Jr. Leroy Chiao Winston E. Scott Koichi Wakata Daniel T. Barry	217,000 lb (landing) 141 rev 290 mi	8 days 22 h 1 min	*Endeavour* retrieved, on Jan. 13, the Japanese satellite *SFU* (*Space Flyer Unit*) launched by an H-2 rocket in March 95. On Jan. 14, the Japanese crew member (Wakata) released the *OAST-Flyer*, a free-flying platform with four autonomous experiments. On Jan. 16, the satellite was retrieved. On Jan. 14, Chiao and Barry performed a spacewalk to evaluate a portable work platform. A second EVA, by Chiao and Scott, took place on Jan. 17 for further testing of spacesuit modifications and space station assembly tools and techniques.
Feb. 21–Sept. 2, 1996	*Soyuz TM-23* Yuri Onufrienko Yuri Usachev	14,400 lb 3027 rev 247 mi	193 days 19 h 11 min	The *Mir 21* crew docked to the *Mir* station on Feb. 23, joining the three crew members of *Mir 20*, who returned in *TM-22* on Feb. 29. In the first month, on-board activities focused on national research, followed by the United States science program NASA 2 when astronaut Shannon Lucid joined the crew on Mar. 23, On May 21 and 24, Onufrienko and Usachev performed their second and third spacewalk to install and unfurl the Mir Cooperative Solar Array brought up by *Atlantis* (*STS 74*) and a fourth on May 30 to install a European earth resources camera on the *Priroda* module. On June 13, they installed a truss structure called Rapana to *Kvant 1*, taking the place of the similar Strela. Later, they worked on the French research program Cassiopeia when *Mir 22* arrived with French astronaut Andre-Deshays. Onufrienko and Usachev returned with her on Sept. 2.
Feb. 22–Mar. 9, 1996	*STS 75* Andrew Allen Scott Horowitz Franklin Chang-Diaz Jeffrey Hoffman Claude Nicollier Maurizio Cheli Umberto Guidoni	229,031 lb (landing) 250 rev 185 mi	15 days 17 h 39 min	*Columbia* carried three foreign astronauts among its crew (Cheli/ESA-Italy, Nicollier/ESA-Switzerland, Guidoni/ASI-Italy). On Feb. 25, the Tethered Satellite System, on a reflight from *STS 46* in July 1992, was deployed on the end of its thin 12.5-mi-long tether line. Experiments could not be completed, however, since, with 12 mi of tether deployed, the cable broke close to the shuttle, and the satellite was lost in orbit. On-board activities then focused on the operation of the third U.S. Microgravity Payload (USMP 3), continuing research aimed at improving basic knowledge of materials under zero-*g* conditions.
Mar. 22–31, 1996	*STS 76* Kevin Chilton Richard Searfoss Ronald Sega M. Richard Clifford Linda Godwin Shannon Lucid	246,335 lb (landing) 144 rev 246 mi	9 days 5 h 15 min 53 s	*Atlantis* lifted off for the third linkup with Russia's *Mir* station as part of Phase 1 of the *International Space Station* program. After docking on Mar. 23 and hatch opening, the two crews of eight conducted joint operations and transferred 5000 lb of supplies and water to *Mir*. The first spacewalk by United States astronauts (Godwin, Clifford) with a shuttle attached to the Russian station took place on Mar. 27, to mount four experiments to its outside for assessing its environment. After concluding science and technology experiments in the *Spacehab* module in the cargo bay, *Atlantis* undocked on Mar. 28, leaving Lucid, the first United States astronaut to be ferried to *Mir* by shuttle, aboard the station to join Onufrienko and Usachev for the next 188 days.
Apr. 23, 1996	*Priroda* Crewless	43,000 lb — 246 mi	—	The *Priroda* module of *Mir* was launched on a Proton rocket. Docking occurred on the first attempt on Apr. 26, despite the failure of two of the module's batteries on Apr. 25. On Apr. 27, *Priroda*

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Apr. 23, 1996 (*cont.*)				was rotated from an axial to a radial docking port. The module carries a set of remote-sensing instruments and some microgravity experiments.
May 19–29, 1996	*STS 77* John H. Casper Curtis L. Brown Andrew S. W. Thomas Daniel W. Bursch	254,879 lb (landing) 161 rev 177 mi	10 d 39 min 18 s	*Endeavour* carried into orbit the commercial *Spacehab* module (fourth flight), with nearly 3000 lb of experiments and support equipment, an Inflatable Antenna Experiment (IAE) on the free-flying *Spartan 207* platform, and four technology experiments called TEAMS. The crew included a Canadian (Garneau). The *Spartan* was released on May 20, and the 50-ft-diameter IAE inflated on its three 92-ft-long struts, the most complex and precise inflatable space structure ever. It was separated from *Spartan* later and burned up in the atmosphere. The satellite was retrieved on May 21, followed by further experiments which included rendezvous with the passive aerodynamically stabilized, magnetically damped satellite PAMS/STU (Satellite Test Unit).
June 20–July 7, 1996	*STS 78* Terence Henricks Kevin Kregel Richard M. Linnehan Susan J. Helms Charles E. Brady, Jr. Jean-Jacques Favier Robert Brent Thirsk	256,170 lb (landing) 271 rev 177 mi	16 d 21 h 48 min	*Columbia* carried the Life and Microgravity Sciences (LMS) Spacelab, equipped to focus on both living organisms in the low-gravity environment and the processing of various substances when gravity effects are greatly reduced. Crew included a French and a Canadian payload specialist (Favier, Thirsk). Mission set a shuttle duration record.
Aug. 17, 1996–Mar. 2, 1997	*Soyuz TM-24* Valeriy Korzun Aleksandr Kaleri Claudie André-Deshays	14,400 lb 243 rev —	32 d 18 h 27 min/André-Deshays; 196 d 17 h 26 min/Korzun and Kaleri	The *Mir 22* crew included a French female researcher-cosmonaut (André-Deshays). The three cosmonauts joined the *Mir 21* crew on board *Mir* on Aug. 19. Prior to their arrival, on Aug.1, the crewless cargo ship *Progress M-31* was undocked and deorbited over the Pacific; on Aug. 2, *Progress 32* arrived. It was undocked on Aug. 18 to free up a docking port, and linked up again with *Mir* on Sept. 3.
Sept. 16–26, 1996	*STS 79* William F. Readdy Terrence W. Wilcutt Jay Apt Tom Akers Carl E. Watz John E. Blaha (up) Shannon Lucid (down)	249,352 lb (landing) 159 rev 239 mi	10 d 3 h 18 min; 188 d 5 h/Lucid	*Atlantis* took off for its fourth docking with Russia's *Mir*, the second flight to carry a commercial *Spacehab* module in its cargo bay, and the first to carry a double version of it. Main mission purposes: to replace Lucid on *Mir* with Blaha for a 4-month stay; to carry 4600 lb of logistics, including food, clothing, experiment supplies, and spare equipment, to the space station; and to return 2200 lb of science samples and hardware to Earth. Five different experiments were turned off on the shuttle and rapidly transferred to *Mir*, the first such transfer of powered scientific apparatus. The two vehicles docked on Sept. 18, and separated on Sept. 23. The mission began a new era of permanent space occupancy by U.S. astronauts. Lucid set records for longest space flight, both for women and for U.S. astronauts (188 days 5 hours).
Nov. 19–Dec. 7, 1996	*STS 80* Kenneth D. Cockrell Kent V. Rominger Tamara E. Jernigan Thomas D. Jones Story Musgrave	226,954 lb (landing) 278 rev 220 mi	17 d 15 h 53 min	*Columbia* carried the 7876-lb (3600-kg) reusable German *ORFEUS-SPAS II* ultraviolet observatory and the 4650-lb (2100-kg) Wake Shield Facility (WSF), on its third flight, with its 4750-lb (2155-kg) carrier system. The *ORFEUS-SPAS* satellite was released on Nov. 19 and recovered with the RMS-robot arm on Dec. 3. The free-flying WFS was deployed on Nov. 22 and retrieved on Nov. 25. Both satellites were highly successful in their missions. The crew conducted biomedical and genetic experiments in weightlessness. Astronaut Musgrave tied NASA astronaut John Young's record for most space flights by any human being (6), and at age 61, he became the oldest person to date to fly in space. The flight set a new duration record for the space shuttle.
Jan. 12–22, 1997	*STS 81* Michael A. Baker Brent W. Jett, Jr. Peter (Jeff) Wisoff John M. Grunsfeld Marsha S. Ivins Jerry M. Linenger (up) John Blaha (down)	249,936 lb (landing) 160 rev 245 mi	10 d 4 h 56 min	*Atlantis* on the fifth mission (of 9) to *Mir* replaced astronaut John Blaha with Jerry Linenger as U.S. *Mir* resident and delivered various supplies. After docking on Jan. 14, the largest transfer of cargo items to date took place: 1400 lbs (635 kg) of water, 1137 lb (516 kg) of U.S. science equipment, 2206 lb (1000 kg) of Russian logistics, and 268 lb (122 kg) of miscellaneous material.
Feb. 10–Aug. 14, 1997	*Soyuz TM-25* Vasiliy Tsibliev Aleksandr Lazutkin Reinhold Ewald (*Mir '97*)	14,400 lb 2885 rev 246 mi	19 d 16 h 35 min/Ewald; 184 d 22 h 6 min/Tsibliev and Lazutkin	The *Mir 23* crew joined the three station occupants, which included U.S. astronaut Jerry Linenger. On Feb. 23, a small flash fire broke out in the *Kvant* module when a lithium perchlorate "candle," a secondary oxygen-generating system, malfunctioned. The crew extinguished the flames within 90 seconds, but had to wear

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Feb. 10–Aug. 14, 1997 (*cont.*)				breathing apparatus and subsequently filtering masks because of smoke. A few hours later, operations returned to normal. Ewald returned to Earth with Korzun and Kaler; in *Soyuz TM-24* on Mar. 2. The two Elektron primary oxygen systems (based on electrolysis of waste water) having failed previously, only one of the secondary systems burning chemicals remained until arrival of the crewless resupply ship *Progress M-34* on Apr. 8. On June 25, the automated *Progress M-34* collided with *Mir* when Tsibliev attempted to link it to the *Spektr* module during a redocking exercise trying out the TORU manual guidance system. The collision damaged one of *Spektr*'s four solar arrays, a radiator and its hull, resulting in its depressurization—a major emergency. The crew was able to disconnect power and data cables and close the airtight hatch in time. During the subsequent weeks, on-board operations were severely curtailed because of the power shortage and were further beset by heart problems of the *Mir* commander (Tsibliev) and various subsystems malfunctions. Tsibliev and Lazutkin returned to Earth on *Soyuz TM-25*, one week after the arrival of the replacement crew (*Soyuz TM-26*, *Mir-24*).
Feb. 11–21, 1997	*STS 82* Kenneth D. Bowersox Scott J. Horowitz Joseph R. Tanner Steven A. Hawley Gregory J. Harbaugh Mark C. Lee Steven L. Smith	213,312 lb (landing) 150 rev 370 mi	9 d 23 h 36 min	*Discovery* caught up with and retrieved the *Hubble Space Telescope* (*HST*) on Feb. 13 for its second scheduled servicing mission since its launch in April 1990. After installing the observatory with the manipulator arm on a turntable in the cargo bay, Lee, Smith, Harbaugh, and Tanner performed critical maintenance of the *HST* and significantly upgraded its scientific capabilities by exchanging eight degraded or failed components and replacing two major instruments, the Goddard High Resolution Spectrometer (GHRS) and the Faint Object Spectrograph (FOS) with the more advanced Space Telescope Imaging Spectrograph (STIS) and the Near Infrared Camera and Multi-Object Spectrometer (NICMOS), respectively. The *HST* was deployed and released in perfect condition on Feb. 19.
April 4–8, 1997	*STS 83* James D. Halsell Susan L. Still Janice Voss Michael L. Gernhardt Donald A. Thomas Roger K. Crouch Gregory T. Linteris	259,705 lb (landing) 63 rev 185 mi	3 d 23 h 33 min 30 s	*Columbia* carried the *Spacelab* module configured as the Microgravity Science Laboratory-1 (MSL-1) with 33 investigations from four space agencies in the areas of crystallography, combustion science, materials science, and life sciences. The mission was to test some of the hardware, facilities, and procedures to be used on the *International Space Station* (*ISS*). In its international complexity, the flight was intended to mirror the future work aboard the *ISS*, bringing together academic, industrial, and governmental partners from around the world. One of the important new facilities designed to enable research in zero-*g* was the EXPRESS Rack (EXpedite Processing of Experiments to the Space Station). However, on Apr. 6, a malfunction of one of the three power-generating fuel cells of the shuttle necessitated shutdown of the unit and early mission termination. *Columbia* had to return on Apr. 8 without accomplishing its mission.
May 15–24, 1997	*STS 84* Charles J. Precourt Eileen M. Collins C. Michael Foale (up) Jerry Linenger (down) Carlos I. Noriega Edward T. Lu Jean-François Clervoy Elena V. Kondakova	221,087 lb (landing) 145 rev 247 mi	9 d 5 h 21 min	*Atlantis* successfully conducted the sixth *Mir* docking mission of the *ISS*/Phase I program, and remained linked to *Mir* for 5 days, unloading 1025 lb (465 kg) water, 845 lb (383 kg) U.S. science equipment, 2576 lb (1168 kg) Russian logistics including a much-needed Elektron oxygen generator unit and a vacuum cleaner, and 393 lb (178 kg) miscellaneous material. The mission also served to exchange U.S. astronaut Jerry Linenger, who had spent over 132 days in *Mir*, against Michael Foale for another 4 months of continuing U.S. presence in space.
July 1–17, 1997	*STS 94* James D. Halsell Susan L. Still Janice Voss Michael L. Gernhardt Donald A. Thomas Roger K. Crouch Gregory T. Linteris	259,705 lb (landing) 251 rev 185 mi	15 d 16 h 45 min	Repeat of mission *STS 83* of Apr. 4–8, 1997, with same crew and payload. *Columbia* carried the Microgravity Science Laboratory (MSL-1), a continuation of NASA's micro-*g* research program in the fields of materials science, protein crystal growth, and physics, with 33 investigations developed by scientists from four space agencies. Main focus: combustion and other processes important to advanced industrial operations and engine development.
Aug. 5, 1997–Feb. 19, 1998	*Soyuz TM-26* Anatoly Solovyev Pavel Vinogradov	14,400 lb ~3093 rev 244 mi	197 d 15 h 35 min/ Solovyev and Vinogradov	*Soyuz TM-26* carried a specially trained crew for repairing the damaged space station. The *Mir 24* crew worked the next $6^1/_2$ months to make repairs and reestablish proper function to all onboard systems, except for the *Spektr* module. When Solovyev and Vinogradov returned to Earth on Feb. 19 *Mir* was again in satisfactory condition for safe human occupancy.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Aug. 7–19, 1997	*STS 85* Curtis L. Brown Kent V. Rominger N. Jan Davis Robert L. Curbeam, Jr. Stephen K. Robinson Bjami V. Tryggvason	217,910 lb (landing) 189 rev 185 mi	11 d 20 h 27 min	*Discovery* lifted off with the Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere-Shuttle Pallet Satellite-2 (*CRISTA-SPAS-2*) on its second flight (and the fourth mission in a cooperative venture between the German Space Agency DARA and NASA). The 7724-lb (3500-kg) satellite was released in orbit to measure infrared radiation emitted by the atmosphere; on Aug. 16 it was retrieved with the shuttle's mechanical arm. Other payloads included the Japanese Manipulator Flight Demonstration (MFD), that is, three separate tests of a robot arm being developed by Japan for its module JEM of the *International Space Station* (*ISS*); the Technology Applications and Science-1 (TAS-1) package with seven separate experiments for studies of atmospheric and topographic physics, solar energy, and testing of new thermal control devices; and the International Extreme Ultraviolet Hitchhiker-02 (IEH-02) with four experiments to research UV radiation from the stars, the Sun, and other sources in the solar system.
Sept. 25–Oct. 6, 1997	*STS 86* James D. Wetherbee Michael J. Bloomfield Vladimir G. Titov Scott E. Parazynski Jean-Loup Chrétien Wendy B. Lawrence David A. Wolf (up) C. Michael Foale (down)	251,730 lb (landing) 169 rev 243 mi	10 d 19 h 21 min	*Atlantis* performed the seventh docking mission with *Mir*. On Oct. 2, Parazynski and Titov performed a 5 h 1 m EVA to retrieve four suitcase-sized Mir Environmental Effects Payload (MEEP) experiments from the exterior of the docking module where they had been attached by *STS 76* crew members in March 1976. They also left a large (121 lb or 55 kg) solar array cover cap at the outside for planned use in a future EVA as part of Russia's attempts to repair the damaged *Spektr* module, and evaluated the small jet-backpack called SAFER (Simplified Aid for EVA Rescue).
Nov. 19–Dec. 5, 1997	*STS 87* Kevin R. Kregel Steven W. Lindsey Kalpana Chawla Winston E. Scott Takao Doi Leonid K. Kadenyuk	226,449 lb (landing) 251 rev 174 mi	15 d 16 h 34 min 4 s	*Columbia* lifted off with the U.S. Microgravity Payload-4 (USMP-4), the *Spartan 201-04* free-flyer, a Collaborative Ukrainian Experiment (CUE) for plant space biology research, and several other hitchhiker payloads. The launch, for the first time, incorporated a new flight maneuver by which the shuttle, at about 6 min into the flight, rolled to a "heads up" position, allowing early acquisition of the Tracking and Data Relay Satellite System (TDRSS) in geostationary orbit in lieu of the Bermuda tracking station and resulting in cost savings. On Nov. 21, the *Spartan* satellite was deployed with the RMS, Intended for investigations of physical conditions and processes of the solar corona and the hot outer layers of the Sun's atmosphere, the spacecraft did not self-activate and began to tumble on a failed attempt by Chawla to regrapple it with the RMS. It was manually recaptured on Nov. 24 by Scott and Doi during a 7 h 43 m spacewalk. On Dec. 3, Scott and Doi conducted a second EVA of 4 h 59 m duration to further evaluate equipment and procedures for the *ISS*, including a telescoping (to 17.5 ft) crane for moving large Orbital Replacement Units (ORUs), and a prototype free-flying TV camera (AERCam/Sprint) for exterior inspections.
Jan. 22–31, 1998	*STS 89* Terrence W. Wilcutt Joe F. Edwards, Jr. James F. Reilly II Michael P. Anderson Bonnie J. Dunbar Salizhan S. Sharipov Andrew S. Thomas (up) David A. Wolf (down)	251,612 lb (landing) 138 rev 247 mi	8 d 19 h 47 min	*Endeavour* completed the eighth docking mission to *Mir*. The crew transferred 774 items of cargo including a new computer and air-conditioning equipment. Wolf replaced Thomas (who became the seventh U.S. astronaut, NASA 7, to stay on Mir for a long-duration mission). Wolf had spent 126 days aboard *Mir*, bringing the total U.S. crew time on *Mir* to 843 days. Other payloads on *STS 89* included the *Spacehab* double module with science experiments and new research equipment for NASA 7.
Jan. 29–Aug. 25, 1998	*Soyuz TM-27* Talgat Musabayev Nikolai Budarin Leopold Eyharts	14,400 lb ~3290 rev 239 mi	207 d 51 min/ Musabayev and Budarin	*Soyuz TM-27* launched with the new *Mir 25* crew and a researcher (Eyharts) from France's space agency CNES. After docking to *Mir* on Jan. 31, the space station had five occupants until Feb. 19, when the *Mir 24* crew (Solovyev and Vinogradov), along with Eyharts, departed in *Soyuz TM-26*.
Apr. 17–May 3, 1998	*STS 90* Richard A. Searfoss Scott D. Altman Richard M. Linnehan Kathryn P. Hire Dafydd Rhys Williams Jay Clark Buckey, Jr. James A. Pawelczyk	232,500 lb (landing) 255 rev 178 mi	15 d 21 h 50 min	*Columbia*'s prime mission for the *Spacelab* module (on its sixteenth flight) was to conduct microgravity research in the human nervous system, including brain, spinal cord, nerves, and sensory organs. Acting both as subjects and operators, the crew performed 26 experiments organized in eight groups: Adult Neuronal Plasticity, Mammalian Development, Aquatic, Neurobiology, Autonomic Nervous System, Sensory Motor and Performance, Vestibular, and Sleep. Included were studies on a variety of species, such as rats, mice, toad fish, zebra fish, crickets, and snails. Secondary objectives included measurement of shuttle vibration forces, demonstration of the Bioreactor system used for cell growth, and three Getaway Special (GAS) payloads.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
June 2–12, 1998	*STS 91* Charles J. Precourt Dominic L. Pudwill Gorie Franklin R. Chang-Diaz Wendy B. Lawrence Janet Lynn Kavandi Valery V. Ryumin Andrew Thomas (down)	259,834 lb (landing) 154 rev 237 mi	9 d 19 h 48 min	The last of nine docking missions with *Mir* ended Phase 1 of the *International Space Station* (*ISS*) program as *Discovery* linked up with the space station on June 4. Payloads flying on *Discovery* included the Alpha Magnetic Spectrometer (AMS) of the Department of Energy to study high-energy particles from deep space, several experiments of advanced technology, human life sciences, *ISS* risk mitigation, and microgravity physics, four Get-Away Specials (GAS) experiments, and two Space Experiment Module (SEM) canisters with experiments from U.S. middle school, high school, and college students. U.S. astronauts performed the majority of science research on *Mir* while their Russian hosts primarily concerned themselves with station upkeep and maintenance. This international cooperation provided a wealth of new knowledge about space flight, particularly with regard to space station design, operation, and maintenance, and was the first of its kind on this magnitude.
Aug. 13, 1998–Feb. 28, 1999	*Soyuz TM-28* Gennady Padaika Sergei Avdeev Yuri Baturin	14,400 lb ~3109 rev 225 mi	11 d 19 h 41 min/ Baturin; 198 d 15 h 30 min/ Padalka	The *Mir 26* crew docked to *Mir*, using the manual backup system TORU due to prior failure of one of two automatic KURS systems on *Mir*. Baturin returned with the *Mir 25* crew on Aug. 25.
Oct. 29–Nov. 7, 1998	*STS 95* Curt Brown, Jr. Steven W. Lindsey Scott Parazynski Stephen Robinson Pedro Duque Chiaki Mukai John Glenn	227,783 lb (landing) 134 rev 348 mi	8 d 21 h 45 min	*Discovery* lifted off with a Japanese woman astronaut on her second flight (Mukai), and payload specialist Senator John Glenn, at 77 years the oldest human ever to fly into space and also the first American to fly into Earth orbit. To take scientific advantage of Glenn's advanced age and his unbroken medical history records with NASA over the last 40 years, about 10 of the flight's 83 experiments and payloads investigated questions pertaining to geriatrics in space, that is, the phenomena of natural aging vs. weightlessness-induced effects closely resembling them. The other experiments made up one of the most diverse sets of commercial and scientific research investigations addressing biotechnology, life sciences, solar physics, and astronomy. The payload included the *Spartan 201-04* free-flyer on its reflight after its failed deployment on mission *STS 87*. It was deployed on Nov. 1 and successfully retrieved 2 days later. The *Discovery* mission was the last U.S. flight prior to the new era of the *International Space Station*.
Nov. 20, 1998	*FGB/Zarya* Crewless	44,088 lb — 240 mi	—	Funded by the U.S. and built by Russia (Khrunichev State Research and Production Center, KhSC, in Moscow), the power and control module *FGB* (*Funktionalnyi-grusovoi blok*) named *Zarya* (Dawn) was launched on a Proton rocket from Russia's cosmodrome in Kazakhstan as the first component of the *International Space Station*.
December 4–15, 1998	*STS 88* Robert D. Cabana Frederick W. Sturckow Jerry L. Ross Nancy J. Currie James H. Newman Sergei K. Krikalev	263,927 lb (orbit) 185 rev 243 mi	11 d 19 h 18 min	*Endeavour* lifted off with the connecting module Node 1/Unity, the first U.S.-built station element for the *ISS*. Rendezvous and connection with *ISS/Zarya* occurred on Dec. 6. On Dec. 14, the crew deployed two satellites: an Argentinian satellite (*SAC-A*) on Dec. 13 and a U.S. Air Force test satellite (*MightySat*). Additional payloads were an IMAX camera for filming from the cargo bay, an educational Space Experiment Module (SEM-07), and other experiments.
Feb. 20–Aug. 27, 1999	*Soyuz TM-29* Viktor Afanasyev Jean-Pierre Haignere Ivan Bella	14,400 lb ~2950 rev 227 mi	7 d 21 h 56 min/ Bella; 187 d 22 h 18 min/ Afanasyev and Haignere; 379 d 14 h 52 min/ Avdeev	The new *Mir 27* crew docked on Feb. 22, and began some 100 experiments, with an incubator with 60 quail eggs hatching a few days after docking, to study effects of space mission on embryonic and postembryonic developments. Fledglings returned to Earth (only two survived) with Padalka and Bella in *Soyuz TM-28* on Feb. 28. On July 18, the Progress M-42 cargo drone docked to deliver a backup computer intended to provide motion control redundancy during a forthcoming period of approx. 6 months of crewless (mothballed) operation of *Mir*, for the first time in 10 years.
May 27–June 6, 1999	*STS 96* Kent V. Rominger Rick D. Husband Tamara E. Jernigan Ellen Ochoa Daniel T. Barry Julie Payette (Canada) Valery Tokarev (Russia)	262,035 lb (orbit) 153 rev 234 mi	9 d 19 h 13 min	The first visit in history to the *International Space Station* occurred when *Discovery* docked flawlessly to the Unity/*Zarya* complex on May 29. The crew began a spacewalk to transfer two cranes (one Russian, one U.S.-made) from the shuttle's cargo bay to the station's outside, installed two new portable foot restraints, and attached three bags with tools and handrails for future assembly operations. The shuttle's hatch to the *ISS* mating adapter PMA 2 was opened on May 29 for leak and pressurization tests. On May 30 the crew members began transferring supplies, equipment,

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
May 27–June 6, 1999 (*cont.*)				and water. They also replaced 18 recharge controller units for the six *Zarya* batteries, installed mufflers over noisy *FGB* fans, and replaced a power distribution unit and transceiver of the Node communication system. At work conclusion, the crew had transferred a total of 3567 lb (1618 kg) material to the *ISS*, including 686 lb (311 kg) water. After releasing the student satellite *Starshine* on June 5, *Discovery* returned.
July 23–27, 1999	*STS 93* Eileen M Collins Jeffrey S. Ashby Catherine G. Coleman Steven A. Hawley Michel Tognini	219,980 lb (landing) 177 rev 79 mi	4 d 22 h 49 min 37 s	*Columbia* launched, with the *Chandra X-Ray Observatory* (AXAF), the third in NASA's series of Great Observatories. The observatory was successfully deployed and subsequently boosted itself to a final orbit of 86,992 by 6034 mi (139,188 by 9655 km), with a period of 63 h 28 m. Secondary objectives included firing of *Columbia*'s jet thrusters at various times during the flight to help an Air Force satellite gather data on the characteristics of jet plumes in space. Also, the Southwest Ultraviolet Imaging System (SWUIS) collected data on ultraviolet light originating from a variety of planetary bodies.
Dec. 19–27, 1999	*STS 103* Curtis L. Brown Scott J. Kelly Steven L. Smith Jean-François Clervoy John M. Grunsfeld Michael Foale Claude Nicollier	211,129 lb (landing) 119 rev 367 mi	7 d 23 h 11 min	*Discovery* conducted the third maintenance/repair mission to the *Hubble Space Telescope* (*HST*). The mission, originally planned for 2000, became necessary when the *HST* had to be put in "safe" (dormant) mode with only two of its six gyroscopes functioning. The crew installed six new gyros and six voltage/temperature improvement kits; replaced the main computer, the Fine Guidance Sensor, and a radio transmitter; and installed a new digital recorder as well as new insulation on two equipment bay doors.
Feb. 11–22, 2000	*STS 99* Kevin R. Kregel Dominic L. P. Gorie Gerhard P. J. Thiele Janet L. Kavandi Janice Voss Mamoru Mohri	225,669 lb (landing) 181 rev 146 mi	11 d 5 h 38 min	Shuttle Radar Topography Mission (SRTM), with primary objective to acquire a high-resolution topographic map of Earth's land mass between 60°N and 56°S, and to test new technologies for deployment of large rigid structures and measurement of their distortions to extremely high precision. The 87 cube-shaped "bays" of the carbon fiber-reinforced plastic, stainless steel, alpha titanium, and Invar radar mast carried active radar antennas for C-band and X-band wavelengths. SRTM will produce topographic maps of Earth 30 times as precise as the best global maps to date.
Apr. 4–June 16, 2000	*Soyuz TM-30* Sergei Zalyotin Alexander Kaleri	14,400 lb 1139 rev 210 mi	72 d 19 h 43 min	First crew return to *Mir* after the 14-year-old station had floated for 223 days without occupants in "mothballed" condition. Prior to their flight, the crewless *Progress M1-1* (first of an upgraded version with double tankage volume) docked automatically with the station on Feb. 3. A second Progress flight to *Mir* occurred on Apr. 27. The crew had the task of returning the station to safe operating condition. After 2 weeks of tests, they succeeded in sealing an air leak in a hatch between the core module's transfer compartment and the airless *Spektr* module. On May 12, an EVA was conducted to test new techniques of leak repair. The *Mir 28* crew returned to Earth in *TM-30*, leaving the renovated station sealed up and again on automatic control.
May 19–29, 2000	*STS 101* James D. Halsell Scott J. Horowitz Mary Ellen Weber Jeffrey N. Williams James S. Voss Susan J. Helms Yuri V. Usachev	225,504 lb (landing) 239 rev 155 mi	9 d 20 h 9 min 9 s	The primary objectives of the second crewed visit to the *International Space Station* (*ISS*) were to remove and replace *FGB*/*Zarya* hardware that had failed, such as battery systems, and to reboost the *ISS* to an altitude consistent with the projected Service Module/*Zvezda* launch and docking. *Atlantis* docked with *ISS* on May 21, and Williams and Voss reinstalled a loose U.S. crane, assembled and installed a Russian "Strela" cargo crane, replaced a communications antenna outside the Node Unity, and installed eight handrails on the outside wall of Unity. Repair and replacement activities extended over the following 4 days. In three separate maneuvers, *ISS* altitude was increased 28.7 mi (45.9 km) to 237.7 × 229.4 mi (380.3 × 367.0 km). The crew transferred over 3300 lb (1500 kg) of gear from *Atlantis* to *ISS* and returned other cargo to the shuttle.
July 12, 2000	*Service Module* (*SM*) *Zvezda* Crewless	42,000 lb — —	—	The Service Module (SM), named *Zvezda* (Star), the first Russian-funded and -owned element of the *ISS*, was launched on a Proton rocket from Baikonur Cosmodrome in Kazakhstan. The docking of *Zvezda* with the *ISS* on July 25, executed very gently at 8 in./s (0.2 m/s) by the *FGB*/*Zarya* as active partner, marked the arrival of the first human living quarters for the new outpost. The new *ISS* stack spanned 119 ft (36 m) in length, and its mass had grown to almost 60 tons.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Sept. 8–20, 2000	*STS 106* Terrence W. Wilcutt Scott D. Altman Daniel C. Burbank Edward T. Lu Richard A. Mastracchio Yuri I. Malenchenko Boris V. Morukov	221,271 lb (landing) 185 rev 200 mi	11 d 19 h 11 min 2 s	*Atlantis* flew the third logistics/outfitting flight to the space station. Lu and Malenchenko conducted a space walk to prepare the *ISS* exterior for future operations. During a week of docked operations the crew worked as movers, cleaners, plumbers, electricians, and cable installers, stowing more than 6600 lb (3 metric tons) of supplies and filling the disposable *Progress M1-3* ship (which had docked at *Zvezda*'s aft port on August 8) with trash and discarded materials.
Oct. 11–22, 2000	*STS 92* Brian Duffy Pamela A. Melroy Leroy Chiao William S. McArthur Peter J. K. Wisoff Michael E. Lopez-Alegria Koichi Wakata	204,455 lb (landing) 202 rev 200 mi	12 d 21 h 43 min	*Discovery* brought the first truss structure segment (called Z1, for its zenith port location on Node Unity), and the third docking adapter (PMA-3, pressurized mating adapter) to the space station. The Z1 truss was installed on the Node, with necessary electrical connections being hooked up and other deployment activities being accomplished during a space walk on Oct. 15. PMA-3 was installed on the Node's nadir port during a second space walk. Two additional space walks completed the assembly tasks for this mission.
Oct. 31, 2000–Mar. 20, 2001	*Soyuz TM-31*, 2R/1S William Shepherd Yuri Gidzenko Sergei Krikalev	14,400 lb — 236 mi	140 d 23 h 38 min	The first resident crew (Expedition One) was launched to the *International Space Station.* After liftoff in Baikonur, the Soyuz-U carrier injected the *Soyuz TM-31* (*ISS* mission 2R) into orbit. Docking occurred on Nov. 2. The second crewless tanker ship *Progress M1-4* (2P) arrived in November—followed by *STS 97* in December with the P6 photovoltaic power module; *STS 98* (5A) with the U.S. Laboratory Destiny in February 2001; a third crewless logistics flight, *Progress M-244* (3P), in the same month; and the Lab outfitting mission 5A.1/*STS 102* in March.
Nov. 30–Dec. 11, 2000	*STS 97* Brent W. Jett Michael J. Bloomfield Joseph R. Tanner Marc Garneau Carlos I. Noriega	197,879 lb (landing) 170 rev 234 mi	10 d 19 h 58 min	*Endeavour* flew Mission 4A to the *ISS*, the first shuttle visit to the station with an *ISS* crew awaiting it. Garneau used the remote manipulator system (RMS) to "park" the 17-ton ITS (integrated truss segment) P6 above the cargo bay to allow for its temperature adjustment. Then Garneau attached P6, attached to the Z1 ITS, on top of the Node, assisted by Noriega and Tanner on an EVA. Jett deployed the two solar array wings, jointly spanning 240 ft (73 m). Deployment of the second wing was delayed one day when the tensioning of the two array blankets of the first wing malfunctioned. Two additional EVAs were conducted, the latter to correct the starboard wing's tension and to install an electrical potential sensor. The P6 provided the *ISS* with up to 62 kW of power.
Feb. 7–20, 2001	*STS 98* Kenneth D. Cockrell Mark L. Polansky Robert L. Curbeam Marsha S. Ivins Thomas D. Jones	254,694 lb (orbiter at liftoff) 201 rev 239 mi	12 d 21 h 20 min	*Atlantis*/5A carried U.S. Laboratory module Destiny to the *International Space Station.* The 28 × 14 ft (8.5 × 4.3 m) Lab was installed at the Node by Ivins using the (RMS) robot arm, assisted by space walkers Curbeam and Jones. Two more EVAs followed as the *ISS* crew completed setup and activation of critical and noncritical Lab systems.
Mar. 8–21, 2001	*STS 102* James D. Wetherbee James M. Kelly Andrew S. W. Thomas Paul W. Richards James S. Voss (up) Susan J. Helms (up) Yury V. Usachev (up) William Shepherd (down) Sergei Krikalev (down) Yuri P. Gidzenko (down)	198,507 lb (orbiter at liftoff) 201 rev 242 mi	12 d 19 h 49 min 166 d 6 h 41 min/Usachev, Voss, Helms	*Discovery*/5A.1, the eighth shuttle mission to the *ISS,* carried resupply and systems/science racks for outfitting the U.S. Lab Destiny, all packed in Leonardo, the first of three Italian-built multipurpose logistics modules (MPLMs) used as "moving vans". During the docked period, the MPLM was transferred to the *ISS* Node's nadir port, unloaded, filled with trash and discarded equipment, and returned to the shuttle. In two space walks, crew members worked outside to support *ISS* assembly. In the first crew rotation, Usachev, Voss, and Helms became the new station crew, replacing Shepherd, Krikalev, and Gidzenko who, after 141 days in space, returned to Earth.
Apr. 19–May 1, 2001	*STS 100* Kent V. Rominger Jeffrey S. Ashby Chris A. Hadfield John L. Phillips Scott E. Parazynski Umberto Guidoni Yuri V. Lonchakov	228,188 lb (orbiter at liftoff) 186 rev 253 mi	11 d 21 h 30 min	*Endeavour* on Mission ISS-6A carried the 3960-lb (1800-kg) Canada-built Space Station Remote Manipulator System (SSRMS), the second Italian-built MPLM Raffaello (total weight loaded: 9000 lb or 4100 kg), and three non-U.S. crew members. The mission featured two EVAs by Parazynski and Hadfield, one to support transfer and deployment of the SSRMS Canadarm2 from the shuttle payload bay to the *ISS* Lab module and install a UHF antenna, the second for external connections and removal of an ECS (Early Communications System) antenna from the Node. After the MPLM was berthed at the Node, the crew transferred two EXPRESS payload racks and individual components from it to the *ISS* and loaded it with return cargo. *Endeavour* undocked on April 29.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Apr. 28–May 6, 2001	*Soyuz TM-32*, 2S Talgat A. Musabaev Yuri M. Baturin Dennis Tito	14,400 lb 123 rev 252 mi	7 d 22 h 4 min	The first "taxi" flight to the *ISS* brought a fresh *Soyuz* returnvehicle for the *ISS* crew. The flight also carried the first commercial space tourist, U.S. businessman Dennis Tito. The *Soyuz* capsule docked at the *FGB* nadir port on Apr. 30. On May 5, the three visitors climbed into the old *Soyuz TM-31*, docked at the Service Module aft end, and returned to Earth, landing in Kazakhstan.
July 12–24, 2001	*STS 104* Steven W. Lindsey Charles O. Hobaugh Michael L. Gernhardt Janet L. Kavandi James F. Reilly	258,222 lb (orbiter at liftoff) 201 rev 250 mi	12 d 18 h 35 min	*Atlantis*, on mission ISS-7A, carried the 6.5-ton Joint Airlock *Quest*, which was installed at the starboard berthing port of the Node, completing Phase 2 of the *ISS* program. Also installed were four high-pressure tanks with oxygen and nitrogen. These activities were supported by three space walks by Gernhardt and Reilly, the last one from the new airlock.
Aug. 10–22, 2001	*STS 105* Scott J. Horowitz Frederick W. Sturckow Patrick G. Forrester Daniel T. Barry Frank L. Culbertson (up) Vladimir N. Dezhurov (up) Mikhail Tyurin (up) Yury V. Usachev (down) James S. Voss (down) Susan J. Helms (down)	257,748 lb (orbiter at liftoff) 185 rev 250 mi	11 d 21 h 13 min; 128 d 20 h 45 min/Culbertson, Dezhurov, Tyurin	*Discovery* delivered the Expedition 3 crew (Culbertson, Dezhurov, Tyurin) to the *International Space Station* and picked up the second station crew (Usachev, Voss, Helms) on the second crew rotation mission. As ISS assembly mission 7A.1, *Discovery* also brought equipment and supplies, carried in the MPLM (multipurpose logistics module) Leonardo on its second flight, which was docked temporarily at the Node for unloading and reloading with return cargo. Barry and Forrester conducted two space walks to install an ammonia servicer, materials experiment carriers, handrails, and cables.
Sept. 14, 2001	*4R* (*DC-1*) Crewless	14,500 lb — 250 mi	—	The Russian Docking Compartment (DC-1) *Pirs* was launched on a Soyuz-U rocket from Baikonur/Kazakhstan at 7:35 p.m. EDT. The Stikovochnly Otsek No. 1 (SO1) linked up with the *ISS* on Sept. 16. The new 16-ft-long (4.88-m) module, with a mass of 8000 lb (3.6 metric tons), is an additional docking port for future Soyuz and Progress vehicles arriving at the station, plus an added stowage area and an airlock for EVAs in Russian Ortan-M spacesuits. DC-1 was launched attached to the cargo ship *Progress M-SO1*, a Soyuz-derived instrumentation and propulsion section of about 3000 lb (1.36 metric tons), which was jettisoned on Sept. 26. *ISS* mass now totaled 303,500 lb (137.67 metric tons).
Oct. 21–30, 2001	*Soyuz TM-33,* 3S Viktor Afanasyev Konstantin Kozeyev Claudie Haigneré (France)	14,400 lb ~153 rev 257 mi	9 d 19 h 0 min	Second *ISS* "taxi" flight, bringing a fresh Soyuz crew return vehicle (CRV) to the station, good for another 200 days. Haigneré was flying commercially. Crew assignment included the French Andromedé (Andromeda) science mission. Afanasyev, Kozeyev, and Haigneré departed in the old *Soyuz TM-32* and landed in Kazakhstan under automatic control.
Dec. 5–17, 2001	*STS 108* Dominic L. Gorie Mark E. Kelly Linda M. Godwin Daniel M. Tani Yuri Onufrienko (up) Carl E. Walz (up) Daniel W. Bursch (up) Frank Culbertson (down) Vladimir Dezhurov (down) Mikhail Tyurin (down)	~257,000 lb (orbiter at liftoff) 186 rev 334 mi	11 d 19 h 36 min; 196 d 20 h 39 min/ Onufrienko, Walz, Bursch	On flrst utilization fiight (UF-1) of the *ISS* program, *Endeavour* carried MPLM *Raffaello* with approximately 3 tons (2.7 metric tons) of cargo plus the fourth expeditionary crew (Onufrienko, Bursch, and Walz) to rotate with the Expedition 3 crew (Culbertson, Dezhurov, and Tyurin). The MPLM was transferred to the Node nadir port with the shuttle RMS. Tani and Godwin conducted an EVA on Dec. 10. After cargo unloading and its replacement with discarded equipment and waste, *Raffaello* was returned to the shuttle cargo bay. Before its undocking, *Endeavour* conducted three reboosts of the station, with a total mean altitude increase of 7.6 n.mi. (14 km).
Mar. 1–12, 2002	*STS 109* Scott D. Altman Duane G. Carey John M. Grunsfeld Nancy J. Currie Richard M. Linnehan James H. Newman Michael J. Massimino	257,917 lb (orbiter at liftoff) 165 rev 359 mi	10 d 22 h 10 min	Fourth "service call" to the *Hubble Space Telescope* (*HST*) was performed by *Columbia* after undergoing more than 100 improvements since its last fiight (*STS 93*). During flve space walks, *HST* received new, more durable solar arrays (SA) with their diode boxes (DBAs), a large gyroscopic Reaction Wheel Assembly (RWA) for pointing, a new power control unit (PCU), a new advanced camera for surveys (ACS), and an experimental cryocooler to restore the dormant Near Infrared Camera and Multi-Object Spectrometer (NICMOS) to operation. After reboost on Mar. 8 by about 4 mi (6.4 km), *HST* was deployed from *Columbia* on Mar. 9.
Apr. 8–19, 2002	*STS 110* Michael J. Bloomfleld Stephen N. Frick Rex J. Walheim Ellen Ochoa Lee M.E. Morin Jerry L. Ross Steven L. Smith	257,079 lb (orbiter at liftoff) 171 rev 250 mi	10 d 19 h 43 min	*Atlantis* flew mission 8A to the *ISS*, carrying the S0 truss and the Mobile Transporter (MT). Four EVAs were conducted to install the S0 and the MT. On Apr. 14 the MT rail cart traversed a distance of 72 ft (22 m) during this flrst operation of a "railroad in space." This was also the flrst flight during which the SSRMS robot arm of the *ISS* was used to maneuver spacewalkers around the station, and the flrst where all EVAs of a shuttle crew were performed from the station's own airlock, *Quest*.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Apr. 25–May 4, 2002	*Soyuz TM-34* Yuri Gidzenko Roberto Vittori (Italy) Mark Shuttleworth (South Africa)	14,400 lb ~154 rev 256 mi	9 d 21 h 25 min	Third "taxi" fiight to deliver a fresh CRV to *ISS*. Shuttleworth was the second "space tourist." A science program, sponsored by Italy and South Africa, included 11 experiments. Gidzenko, Vittori, and Shuttleworth returned in *Soyuz TM-33*, which had been docked to the *ISS* since October 2001.
June 5–19, 2002	*STS 111* Kenneth D. Cockrell Paul S. Lockhart Franklin Chang-Diaz Philippe Perrin (France) Valery Korzun (up) Peggy Whitson (up) Sergei Treschev (up) Yuri Onufrienko (down) Carl E. Walz (down) Daniel W. Bursch (down)	256,889 lb (orbiter at liftoff) 217 rev 240 mi	13 d 21 h 35 min; 184 d 22 h 14 min 23 s/Korzun, Whitson, Treschev	*Endeavour*, on *ISS* Mssion UF-2, carried the flfth *ISS* resident crew (Korzun, Whitsun, and Treschev), returning with the fourth station crew (Onufrienko, Bursch, and Walz). Its cargo comprised MPLM *Leonardo* on its third fiight, loaded with 5600 lb (2500 kg) of equipment and consumables; the Mobile Base System (MBS); a replacement wrist roll (WR) joint for the SSRMS; meteoroid/orbital debris protection shields (ODPs) for the *Zvezda* Service Module; and new science payloads. During three spacewalks, astronauts installed the MBS on the MT (mobile transporter), temporarily stowed the ODPs, and replaced the SSRMS WR joint.
Oct. 7–18, 2002	*STS 112* Jeffrey S. Ashby Pamela A. Melroy David A. Wolf Sandra H. Magnus Piers J. Sellers Fyodor N. Yurchikhin (Russia)	256,917 lb (orbiter at liftoff) 172 rev 252 mi	10 d 19 h 58 min	*Atlantis* flew Mission 9A to the *ISS*. On Oct. 10, its main payload, the 14.5-ton (13.1-metric-ton) S1 truss element, was transferred with the Canadarm2 from the shuttle cargo bay to the *ISS* and flrmly attached to the S0 center truss segment atop the Lab module. The installation for the flrst time featured the SSRMS operating from its new MBS. The crew also transferred about 1000 lb (450 kg) of supplies, equipment and other cargo to the station. Three spacewalks were performed to connect the S1 and S0 trusses with electrical, flber-optical, data, and coolant umbilicals, and to install external systems.
Oct. 29–Nov. 9, 2002	*Soyuz TMA-1*, 5S Sergei Zalyotin Yuri Lonchakov Frank De Winne (ESA, Belgium)	15,920 lb ~170 rev 260 mi	10 d 20 h 53 min	Fourth taxi flight to the *ISS*, to replace the previous CRV, *Soyuz TM-34*. *TMA-1* was the first of a new version of the Soyuz, with modifications financed by NASA for improved safety and widened crewmember size range. A science program sponsored by Belgium and the Russian space agency Rosaviakosmos (RSA) included 23 experiments. The crew returned on the "old" *Soyuz TM-34.*
Nov. 23–Dec. 7, 2002	*STS 113* Jim Wetherbee Paul Lockhart Michael Lopez-Alegria John Herrington Kenneth Bowersox (up) Nikolai Budarin (up) Donald Pettit (up) Valery Korzun (down) Peggy Whitson (down) Sergei Treschev (down)	256,747 lb (orbiter at liftoff) 215 rev 247 mi	14 d 6 h 47 min; 161 d 1 h 14 min 38 s/ Bowersox, Budarin, Petit	Mission *ISS*-11A on *Endeavour* carried the Expedition 6 crew (Bowersox, Budarin, and Pettit), the fourth truss element, P1 (of 11 total), and other equipment and resupply. The 14.5-ton (13.1-metric-ton) P1 was transferred and installed on the S0 truss portside, and three spacewalks were conducted to connect the P1 and install other gear. After *Endeavour* undocked on Dec. 2, it launched two tethered mini-satellites called MEPSI (microelectromechanical systems-based Picosat Inspector), and returned with the Expedition 5 crew (Korzun, Whitson, and Treschev).
Jan. 16–Feb. 1, 2003	*STS 107* Rick D. Husband William C. McCool David M. Brown Kalpana Chawla Michael P. Anderson Laurel B. Clark Ilan Ramon (Israel)	263,706 lb (orbiter at liftoff) 255 rev 177 mi	15 d 22 h 21 min	*Columbia* flew on a research mission with 32 payloads and 59 separate investigations. Both in the shuttle middeck and in the *Spacehab* Research Double Module (RDM), on its first flight in the cargo bay, the crew worked on a mixed complement of competitively selected and commercially sponsored research in the space, life, and physical sciences. *Columbia* was lost with its crew during reentry on Feb. 1, 16 min before planned landing at Kennedy Space Center, when it violently disintegrated over Texas. Debris from the orbiter fell in Texas and other states. The Columbia Accident Investigation Board (CAIB) later concluded that one of the left wing's leading edge RCC (reinforced carbon-carbon) elements or an associated filler strip of the heat shield had been compromised during ascent to orbit by a piece of debris, possibly a suitcase-sized chunk of foam insulation blown off the external tank by the supersonic air stream, hitting the leading edge and rendering the wing unable to withstand reentry heating longer than about 8 min after entry interface.
Apr. 25–Oct. 27, 2003	*Soyuz TMA-2*, 6S Yuri Malenchenko Ed Lu	15,920 lb ~2889 rev 256 mi	184 d 22 h 47 min	The first crew rotation flight by a Soyuz carried Expedition 7, a two-man "caretaker" crew, to stretch out *ISS* consumables during the shuttle stand-down. *Soyuz TMA-2* docked to the *ISS* on April 28, replacing the previous CRV, *Soyuz TMA-1*. The Expedition 6 crewmembers (Bowersox, Budarin, and Pettit) returned to Earth in *Soyuz TMA-1*, experiencing an unexpected switch of their onboard computer to the backup reentry mode of pure ballistic descent, which almost doubled their peak deceleration (~9 g), and missed the primary landing site by an "undershoot" of ~300 mi (480 km), landing on May 3. Recovery forces reached the landing site after a delay of 4.5 h, finding the crew in good health outside the capsule.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Oct. 15–16, 2003	*Shenzhou 5* ("Divine Vessel 5") Yang Liwei	17,000 lb 14 rev 213 mi	21 h 23 min	Orbital launch of the first Chinese "taikonaut" on the new human-rated Long March (Chang Zheng, CZ) 2F rocket. Landing was about 600 mi (1000 km) east of the launch area, the Jinquan Satellite Launch Center in the Gobi Desert.
Oct. 18, 2003-Apr. 29, 2004	*Soyuz TMA-3*, 7S Michael Foale Alexander Kaleri Pedro Duque (ESA, Spain)	15,920 lb ~3053 rev 249 mi	9 d, 21 h 3 min/ Duque; 194d 18 h 35 min/ Foale and Kaleri	Second *ISS* crew rotation flight by a Soyuz carried Expedition 8 (Foale and Kaleri), the second "caretaker" crew, and visiting crewmember Duque. On Oct. 27, the previous CRV, *Soyuz TMA-2*, undocked from the of the *FGB*/*Zarya* module nadir port, where it had stayed for 182 days, and landed in Kazakhstan with Malenchenko, Lu, and Duque.
Apr. 18–Oct. 24, 2004	*Soyuz TMA-4*, 8S Gennady Padalka Michael Fincke André Kuipers (ESA, Netherlands)	15,920 lb ~2951 rev 241 mi	10 d 20 h 52 min/ Kuipers; 187 d 21 h 17 min/ Padalka and Fincke	Third crew rotation flight by a Soyuz carried Expedition 9 crew (Padalka and Fincke) plus visiting cosmonaut/researcher Kuipers. *TMA-4* docked to the *ISS* on Apr. 21, replacing the previous CRV, *Soyuz TMA-3*, which, when it undocked on Apr. 29, had reached an in-space time of 194 days. In it, Expedition 8 crewmembers Foale and Kaleri plus Kuipers landed in Kazakhstan.
Oct. 13, 2004–Apr. 24, 2005	*Soyuz TMA-5*, 9S Leroy Chiao Salizhan Sharipov Yuri Shargin (Russia)	15,920 lb ~3031 rev 241 mi	9 d 21 h 30 min/ Shargin; 192 d 19 h 35 min/ Chiao and Sharpov	Fourth *ISS* crew rotation flight by a Soyuz carried Expedition 10 (Chiao and Sharipov), the fourth "caretaker" crew, plus visiting cosmonaut/researcher Shargin. On Oct. 16, *TMA-3* docked to the *ISS* at the DC-1 docking compartment's nadir (downward)-pointing port. On Oct. 23, the previous CRV, *Soyuz TMA-4*, undocked from the *FGB* nadir port, where it had stayed for 186 days (187 days in space), and landed in Kazakhstan with Padalka, Fincke, and Shargin.
Apr. 14–Oct. 10, 2005	*Soyuz TMA-6*, 10S Sergei Krikalev John Phillips Roberto Vittori (ESA, Italy)	15,920 lb ~2818 rev 235 mi	9 d 21 h 22 min/ Vittori; 179 d 23 min/ Krikalev and Phillips	Fifth crew rotation flight by a Soyuz carried Expedition 11 crew (Krikalev and Phillips) plus visiting crewmember Vittori. TMA-6 docked to the *ISS* on Apr. 16, replacing the previous CRV, Soyuz *TMA-5*, which, when it undocked on Apr. 24, had reached an in-space time of 192 days. In it, Chiao, Sharipov, and Vittori landed in Kazakhstan.
July 26–Aug. 9, 2005	*STS 114* Eileen M. Collins James M. Kelly Andrew S. W. Thomas Charles J. Camarda Wendy B. Lawrence Stephen K. Robinson Soichi Noguchi (Japan)	267,825 lb (orbiter at liftoff) 219 rev 218 mi	13 d 21 h 32 min 22 s	Return-to-flight mission of *Discovery*. Prior to final approach to the *ISS*, *Discovery* performed an R-bar or rendezvous pitch maneuver (RPM) at ~600 ft (180 m) distance under the *ISS*, a 360° backflip to allow digital imagery of its thermal protection system from the *ISS*. About 15,000 lb (6800 kg) of cargo was transferred from the shuttle's flight container, the MPLM *Raffaelo*, to the *ISS* and approximately 8600 lb (3900 kg) from the station to the shuttle for return to Earth. Crewmembers performed three EVAs, tested new safety equipment and repair procedures, and carried out a first-of-its kind spacewalking heat shield repair.
Sept. 30, 2005–Apr. 8, 2006	*Soyuz TMA-7*, 11S William McArthur Valery Tokarev Gregory Olsen (United States)	15,920 lb ~2992 rev 229 mi	9 d 21 h 14 min/ Olsen; 189 d 19 h 53 min/ McArthur, Tokarev	Sixth *ISS* crew rotation flight by a Soyuz carried Expedition 12 crew (McArthur, Tokarev) and the third "space tourist" (Olsen). *TMA-7* docked to the *ISS* on Oct. 3. On Oct. 10, *Soyuz TMA-6* undocked from the *FGB* nadir port, where it had stayed for 177 days, and landed in Kazakhstan with Krikalev, Phillips, and Olsen.
Oct. 12–17, 2005	*Shenzhou 6* Fei Junlong Nie Haisheng	17,000 lb 75 rev 209 mi	115 h 32 min	Second orbital flight of a crewed spacecraft on China's Long March 2F rocket, this time carrying two "taikonauts." Both crewmembers left their seats, floated in space, and at one time took off their space suits.
Mar. 29–Sept. 28, 2006	*Soyuz TMA-8*, 12S Pavel V. Vinogradov Jeffrey Williams Marcos Cezar Pontes (Brazil)	15,920 mi ~2885 rev 229 mi	9 d 22 h 18 min./ Pontes; 182 d 22 h 43 min/ Vinogradov and Williams	Seventh *ISS* crew rotation flight by a Soyuz carried Expedition 13 crewmembers Vinogradov and Williams, and visiting crewmember Pontes. On Apr. 8 the previous CRV, *Soyuz TMA-7*, undocked from the *ISS* and returned to Kazhakstan, carrying McArthur, Tokarev, and Pontes.
July 4–17, 2006	*STS 121* Steven Lindsey Mark Kelly Stephanie Wilson Michael Fossum Piers Sellers Lisa Nowak Thomas Reiter (ESA, Germany) [up]	266,962 lb (orbiter at liftoff) 202 rev 220 mi	12 d 18 h 37 min; 171 d 3 h 54 min/ Reiter	*Discovery*, on Mission ULF1.1 to the *ISS*, flew the first of two test flights after return to flight. The joint crews transferred the fully loaded MPLM *Leonardo*, conducted three EVAs, reberthed the loaded MPLM in the shuttle cargo bay, performed external inspections of the orbiter heat shielding, and completed cargo transfers. Reiter remained on the *ISS* when *Discovery* returned to Earth, joining Expedition 13.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

TABLE 1. Crewed space flights 1961–2006 (*cont.*)

Dates launched and recovered	Designation and crew	A. Weight B. Revolutions C. Max. distance from Earth	Duration	Remarks
Sept. 9–21, 2006	*STS 115* Brent Jett Chris Ferguson Dan Burbank Heidemarie Stefanyshyn-Piper Joe Tanner Steve MacLean (Canada)	269, 840 lb (orbiter at liftoff) 187 rev 217 mi	11 d 19 h 6 min	*Atlantis*, on Mission 12A to the *ISS*, flew Shuttle Flight Test 2. Prior to final approach, *Atlantis* performed an RPM, similar to that of *STS 114*. Preliminary views of the thermal protection system did not indicate any signs of damage. Crews installed the P3/P4 truss structure with the second set of solar arrays (of four), deployed the two solar array wings and a huge radiator panel from the P4 truss, conducted three EVAs, and resupplied the station.
Sept. 18, 2006 –	*Soyuz TMA-9*, 13S Michael Lopez-Allegria Mikhail Tyurin Anousheh Ansari	15,920 lb — 216 mi	10 d 21 h 4 min 17 s/Ansari	This flight carried new Expedition 14 crewmembers Lopez-Alegria and Tyurin, replacing Expedition 13 crewmembers Vinogradov and Williams, plus Iranian-American "taxi cosmonaut" Ansari, the first female spaceflight participant to visit the space station, flying under contract with the Russian Federal Space Agency. Expedition 13 crewmember Reiter remained on the station until Dec. 20, joining Expedition 14. On Sept. 28, *Soyuz TMA-8* undocked from the *ISS* FGB nadir port and landed in Kazakhstan with Vinogradov, Williams, and Ansari. Return and landing of *Soyuz TMA-9* was scheduled for Mar. 19, 2007, which would give Lopez-Alegria and Tyurin 182 d in space.
Dec. 9–22, 2006	*STS 116* Mark Polansky William (Bill) Oefelein Nicholas Patrick Robert (Bob) Curbeam Joan Higginbotham Christer Fuglesang (ESA, Sweden) Sunita Williams (up) Thomas Reiter (down)	265,466 lb (orbiter at liftoff) 202 rev 222 mi	12 d 20 h 45 min	*Discovery*, on Mission 12A.1 to the *ISS*, delivered and installed the P5 truss segment (serving as spacer at the end of the P4 truss), reconfigured the electrical power and thermal control systems to their final configurations and activated them, and transferred extra oxygen for EVA prebreathing and nitrogen to storage tanks on the outside of the airlock. Astronauts conducted four EVAs, including one that was not originally planned to support continuing attempts by ground controllers to retract the 4B solar array of the P6 truss. Williams replaced Reiter as a member of Expedition 14 on the *ISS*.

1 lb = 0.45 kg. 1 mi = 1.6 km. 1 ft = 0.3 m. 1 yd = 0.9 m. 1 fathom = 1.8 m. 1 lb/in.^{2} = 6.9 kPa.

reduce g loads and heat pulse, especially for higher-energy missions such as return from high altitudes and lunar or planetary trips, advanced reentry vehicles fly a lifting reentry. Some lift can even be obtained from a basically ballistic capsule. A basic feature of the Soviet *Soyuz* and the American *Gemini* and *Apollo* spacecraft, as well as the United States space shuttle and future crewed vehicles, is that the return trajectory is extended by applying to the vehicle an aerodynamic crosswind or lifting force L that is created by offsetting the center of gravity from the center of aerodynamic pressure of the spacecraft (located at the centerline for symmetrical bodies) in order to let the airflow hit the vehicle asymmetrically at an angle of attack. The direction and magnitude of this force is controlled by signals from an on-board guidance computer. Although often only a small fraction of the drag force D (L/D for *Apollo*-type bodies is typically $^1/_4$ to $^1/_5$), this lift is sufficient to modulate the trajectory for landing point control and reduction of heating (not total, or integrated, heat load) and deceleration load peaks (**Fig. 1**). *See* AEROTHERMODYNAMICS; ATMOSPHERIC ENTRY; SPACE SHUTTLE.

Subsystems. In addition to the thermal protection system and reentry and landing systems, basic features of crewed spacecraft include the following subsystems: (1) crew equipment (couches and restraints, operational aids, clothing, food and water, hygiene equipment, medical supplies, and survival equipment); (2) displays and controls; (3) electrical power system; (4) environmental control and life support system; (5) reaction control system (for attitude maneuvers); (6) on-board propulsion system (for velocity-change maneuvers); (7) telecommunications system (on-board voice and data/telemetry radio-frequency equipment and antennas); (8) guidance and navigation system, which is interrelated with (9) the stabilization and control system (inertial sensors, optical instruments, and computer); (10) mechanical systems (escape, docking, landing); and (11) extravehicular activity (EVA) systems, including space suits, life support umbilicals or backpacks, mobility aids, restraints, and handholds. *See* SPACE NAVIGATION AND GUIDANCE; SPACE TECHNOLOGY.

	Apollo	*Gemini*
Thickness at stagnation point	1.7 in. (4.3 cm)	1 in. (2.5 cm)
Max. heat rate		
Btu/(ft^2)(s)	620	120
J/(m^2)(s)	7.04×10^6	1.36×10^6
Heat load		
Btu/ft^2	36,000	12,100
J/m^2	4.09×10^8	1.37×10^8

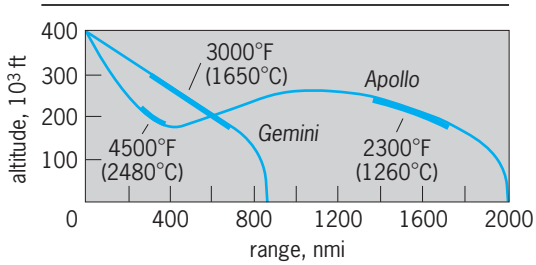

Fig. 1. ***Apollo*** **and** ***Gemini*** **reentry conditions; 1 ft = 0.3 m; 1 nmi = 1.852 km. (*NASA*)**

Environmental controls. The main system that distinguishes crewed spacecraft from other spacecraft is the environmental control and life support system. It provides the crew with a suitably controlled atmosphere. Because all atmospheric supplies must be carried into space, it is essential to recirculate and purify the spacecraft atmosphere to keep the total weight of the vehicle within reasonable limits. In future years, partially or completely closed life support systems will even recover oxygen and water from the atmosphere for reuse, reducing resupply requirements.

Soviet and, today, Russian spacecraft have used a nitrogen-oxygen environment at normal sea-level pressure. United States spacecraft have used pure oxygen at a pressure of 5 lb/in.^{2} (34.5 kilopascals) during space flight. It provides the lightest, simplest system possible, and was adopted for this reason. Although some medical disadvantages were feared, they did not materialize. However, the technique requires complete control of oxidizable (inflammable) materials on board. For the extended mission duration of post-*Apollo*-type vehicles, such as *Skylab*, the space shuttle and the *International Space Station*, NASA has used a more Earth-like atmosphere. *See* SPACE BIOLOGY.

Supplies. All supplies for the entire flight must be carried in the spacecraft, unless special supply missions are to be flown, for example, to a station in Earth orbit. An Earth-like environment is required because of human dependence on certain gaseous constituents for breathing. There must be provision for controlling the metabolically produced carbon dioxide (CO_2) in the spacecraft atmosphere; controlling water vapor and maintaining a comfortable humidity; reducing odor and removing toxic products; maintaining a comfortable thermal environment, atmospheric pressure, sound-pressure level, and radiation level; meeting drinking water and nutritional requirements; disposing of waste; supporting the crew during emergencies; and remedying weightlessness.

Space suits. Space suits or pressure suits are mobile spacecraft or chambers that house the astronauts and protect them from the hostile environment of space. They provide atmosphere for breathing, pressurization, and thermal control; protect astronauts from heat, cold, glare, radiation, and micrometeorites; contain a communication link and hygiene equipment; and must have adequate mobility. The suit is worn during launch, docking, and other critical flight phases. During noncritical flight phases, the space suit is removed for greater comfort. This practice has also been adopted by Soviet cosmonauts.

The use of two gases at atmospheric pressure on board, while excellent for intravehicular activities (IVA), becomes somewhat disadvantageous when extravehicular activity is required, for two principal reasons. First, adequate mobility in a space suit

pressurized to over 3–4 lb/in.^{2} (21–28 kPa) has proved extremely difficult. Second, at lower suit pressures, suitable for mobility, only pure oxygen can be used for breathing. In this case, because of the lower gas density, cooling may have to be provided to the astronaut with water, circulating in tubes sewn into an undergarment.

Power. The various life-support, communications, and other systems in a spacecraft depend upon adequate power. For very short missions conventional battery supplies may be used, but for longer-duration missions battery supplies must be constantly recharged from some other source such as a solar or nuclear supply. Fuel cells, which generate electric power from the chemical combination of suitable reactants such as hydrogen and oxygen, have proved effective in *Apollo*, *Skylab*, and the shuttle. Solar power is particularly useful within the orbit of Mars. Nuclear power may be used anywhere in space, thus removing the requirement to stay close to the Sun, but requires special provision for shielding spacecraft occupants and equipment from the radiations. *See* FUEL CELL; NUCLEAR BATTERY; SOLAR POWER; SPACE POWER SYSTEMS.

Communications. Radio and television communications channels permit a two-way exchange of information and instructions that are important for both normal flight and emergencies. Through telemetering links, the operations center monitors the health and operation of the spacecraft and its crew, and receives various kinds of engineering, operational, and research data.

A space flight communications system requires suitable transmitters, receivers, display units, recorders, power supplies, and antennas. The antennas impose challenging design requirements. The ground antennas must be able to track the spacecraft, and should be large enough to minimize the size and complexity of the spacecraft antenna. For near-Earth orbits the spacecraft or its antenna must be sufficiently maneuverable to keep the antenna pointed at the receiving station on the ground; for deep-space missions rapid maneuverability is not required, but accurate pointing is necessary. *See* SPACE COMMUNICATIONS; SPACECRAFT GROUND INSTRUMENTATION.

Reliability. Because of the potential hazards of space flight to personnel, crewed spacecraft must meet stringent requirements of safety and hardware reliability. Reliability is associated with the probability that the spacecraft systems will operate properly for the required length of time and under the specified conditions. To assure survival and return in all foreseeable emergencies, the design of a crew-rated spacecraft includes standby systems (double or triple redundancy) and allows for launch escape, alternative and degraded modes of operation, contingency plans, emergency procedures, and abort trajectories to provide maximum probability of mission success. The priority order of this reliability requirement is (1) crew safety, (2) minimum achievable mission fulfillment, (3) mission data return, and (4) minimal degradation. *See* RELIABILITY (ENGINEERING).

Crew recovery. For launch escape, the *Mercury* used a solid rocket mounted on a tower (**Fig. 2**) to provide the necessary propulsive force to escape from a malfunctioning launch vehicle. This device was jettisoned after the critical flight phase. The *Gemini* and early Soviet spacecraft had ejection seats with parachutes. These were used on some of the early *Vostok* flights to recover the crew on land, but in *Gemini* they were intended only for launch escape. Because a vehicle launched from the Russian sites passes over land during the extremely critical flight phase immediately after launch, the Russians provide for landing on land to effect a safe recovery from an abort in this region. It is also used as the prime recovery method after reentry. However, *Mercury* and *Gemini* were designed for launch over the Atlantic Ocean, eastward from Cape Canaveral, Florida, so that a water-landing capability was required. Also, because much of the orbital track was over water, a water-landing technique would be necessary to cope with the majority of possible random abort situations. The same logic applied to the lunar missions of *Apollo*. Although the United States technique of water recovery has been highly developed and has proven eminently satisfactory for non-reusable spacecraft, more advanced spacecraft, such as the space shuttle, are designed for landing on runways much like an airplane.

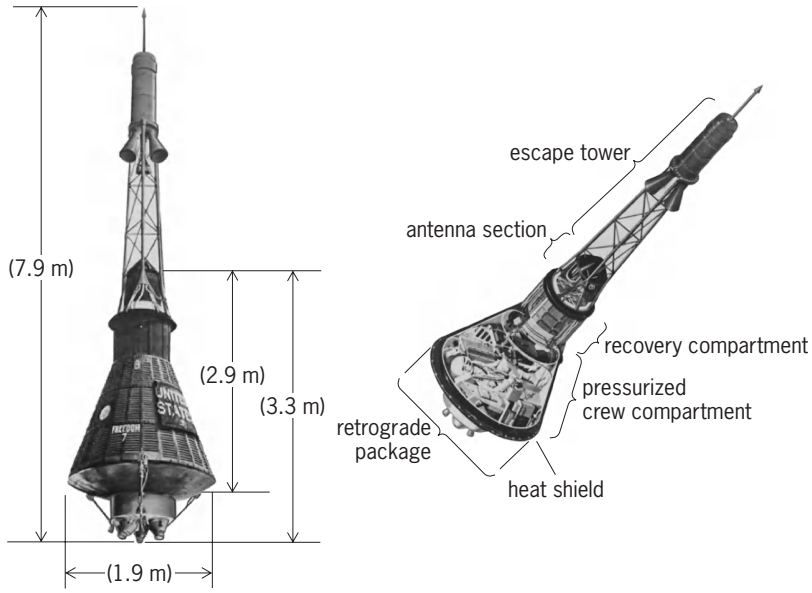

Fig. 2. ***Mercury*** **spacecraft, showing approximate dimensions. 1 m = 3.3 ft. (*NASA*)**

In-flight maintenance. The mission reliability of spacecraft can be increased by providing for inflight maintenance and repairs by the crew. The decision to use in-flight repair on crewed spacecraft, however, must take into consideration the effect on the overall probability of mission success and flight crew safety, and the associated weight penalties which result from this reliability improvement technique.

Emergencies. Both equipment and operational procedures are carefully planned and designed to minimize harmful results from emergencies. On the operational side, procedures are worked out in advance for a wide range of imaginable occurrences. On the

equipment side, ample use is made of redundancies, so that if one piece of equipment fails an alternate means of performing that equipment's functions is available. The value of redundancy was evident in the *Apollo 13* emergency. When an explosion occurred in one of the *Apollo* fuel cell oxygen tanks, substantially disabling the spacecraft, the ability of the attached lunar module to provide oxygen, power, orientation, and propulsion made it possible to return the astronauts safely to the ground after circumnavigating the Moon.

Preventive measures. As an example of the extreme complexity of crewed space flight, the Saturn 5 launch vehicle of the United States, with the *Apollo* spacecraft combined, had between 5 and 6 million functional parts, miles of wiring, and thousands of connections. To assure the functional reliability required for the lunar landing of humans, preventive measures had to include conservative design (with a wide margin of safety), stringent technical and administrative controls, and extensive premission simulation and testing, far beyond the required safety level, under various mission conditions. Without high-speed digital computers, a successful lunar landing program would not have been possible and the space shuttle could not fly. *See* HUMAN-MACHINE SYSTEMS; SERVOMECHANISM; SIMULATION.

Physiological stresses. Experience gained from both the United States and Russia's crewed space flight programs has provided valuable physiological data. *Apollo*, *Skylab*, *Salyut*, and *Mir* were especially fruitful. Voice, heart rate, respiratory rate, blood pressure, and body temperature are monitored during periods of dynamic stress (for example, launch, reentry, and extravehicular activity). Extensive ground studies and laboratory tests have been performed.

The stress imposed by long exposure to weightlessness is still a major question. Astronauts adapt fairly quickly to the sensation of weightlessness, but the complete role played by the inner ear system in this adaptation has not been determined.

Cardiovascular deconditioning was detected postflight during the Gemini program. Conditioning equipment was developed subsequently to be used in space to measure the amount of deconditioning and to act as a countermeasure to weightlessness. The most significant type of this equipment is a lower-body negative pressure device which exposes the lower half of the body to reduced pressure; engorgement of the veins in the legs and abdominal viscera results, thereby counteracting blood pooling in the chest cavity and increased pressure in the left atrium at 0 *g*. This effect is believed to prevent the loss in blood volume otherwise produced as engorgement of the chest veins stimulates diuresis.

Mineral balance is also markedly affected by weightlessness. Loss of the bone calcium secreted in urine and feces is observed, calling for countermeasures. Similar loss of nitrogen indicates loss of muscle tissue, which is shown clearly by circumferential measurement of the astronauts' legs. These effects are considered serious, and although *Skylab*, *Salyut*, and *Mir* have shown that humans can stay in space for appreciable periods of time, it is the considered opinion of those experts who have reviewed the data that for crewed missions lasting much longer than a year the causes of mineral loss have to be better understood and methods developed to prevent and control these losses. Since similar mineral losses are observed during prolonged bed rest, bed-rest subjects are being used to study this problem further.

Biomedical monitoring throughout a space flight is considered important to the safety and effectiveness of the crew. Although life support and environmental parameters (for example, total pressure, oxygen partial pressure, temperature, and radiation level) can be expected to provide the earliest indication of non-nominal conditions, their significance and urgency for corrective action usually require correlation with biomedical data. The medical evaluation of potential microbial and chemical contamination aboard the spacecraft requires microbiological, pulmonary, and metabolic studies. Basic hematological and biochemical measurements are being obtained at frequent intervals during the initial portion of a space flight to obtain a database. *See* AEROSPACE MEDICINE.

Soviet/Russian Programs

In the former Soviet Union, ideas and concepts of crewed space flight date back at least to the last century, to K. E. Tsiolkovskiy, now considered one of the pioneers of space flight. After the first *Sputnik* launch on October 4, 1957, developments of crewed space flight capability followed in quick succession, leading from the first-generation *Vostok* (East) to the second-generation *Voskhod* (Ascent) and to the third-generation *Soyuz* (Union) spacecraft. Originally engaged in an aggressive program to land a cosmonaut on the Moon before the United States lunar orbit mission of *Apollo 8* in December 1968, the Soviet Union redirected its aims, after four test failures of its N1 heavy-lift launcher, toward the development of permanent human presence in Earth orbit.

Vostok. The *Vostok* (**Fig. 3**), a single-seater for short-duration missions and ballistic reentry from Earth orbits, consisted of a near-spherical cabin covered entirely with thermal protection system material, having three small viewports and external radio antennas. It contained a life-support system, radios, simple instrumentation, and an ejection seat. The 7-ft (2-m) sphere of the cabin was attached to a service module resembling two truncated cones base to base, with a ring of gas pressure bottles on the upper cone close to the cabin. The module contained chemical batteries, orientation rockets, the main retropropulsion system, and additional support equipment.

In all, six *Vostok* cosmonauts were launched from 1961 to 1963, including the first woman in space, V. V. Tereshkova (*Vostok 6*). The launch vehicle was a two-stage rocket with four additional strap-on booster units surrounding the first stage, forming a pyramidlike assemblage. All five propulsion units, feeding into 20 rocket nozzles, were started on the launch pad; later, the four strap-on units fell away,

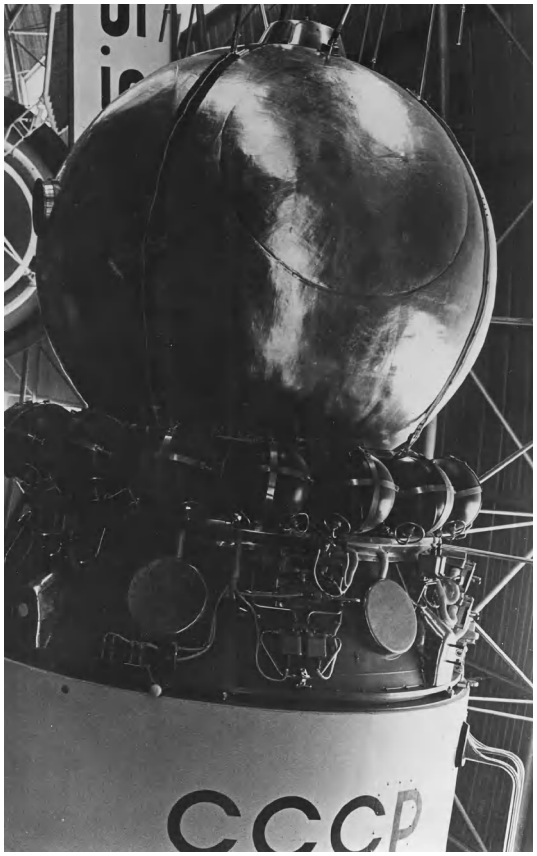

Fig. 3. ***Vostok*, a first-generation Soviet spacecraft. (*NASA*)**

leaving the core unit to continue burning for a time. After its burnout, the single-engine second stage took over, inserting the payload into the target orbit. Payload capability to orbit was about 10,400 lb (4750 kg).

Voskhod. The second-generation *Voskhod*, essentially a greatly modified *Vostok* (for example, the heavy ejection-seat system was removed), was a short-duration multiperson craft and was designed to permit extravehicular activity or spacewalking by one of the crew. The first three-member crew flew on *Voskhod 1*, while on the double-seater *Voskhod 2* the place of the third seat was used for an expandable airlock to permit A. A. Leonov on March 18, 1965, to egress into space without evacuating the main cabin of air, making him the first human to perform extravehicular activity. The two *Voskhod* spacecraft were launched in 1964–1965 by an improved version of the *Vostok* launch vehicle, fitted with a new upper stage, that had the capability to lift the heavier *Voskhods* and—upon further uprating—the subsequent *Soyuz* vehicles.

Soyuz. The *Soyuz* design is much heavier, larger, and more advanced in its orbital systems, permitting extended orbital stay times. It consists of three main sections or modules: a descent vehicle, an orbital module, and an instrument-assembly module (**Fig. 4**), with a total habitable volume of about 350 ft^3 (10 m^3). The three elements are joined together, with the descent vehicle in the middle. Shortly before atmospheric reentry, the two outer modules are jettisoned. The spacecraft has a total length of 24.5 ft (7.5 m), and the span of its solar panel wings is 27.5 ft (8.4 m).

Descent vehicle. The descent vehicle is intended for crew location during launch, orbital flight, descent, and landing. Roughly bell-shaped, it is covered with an ablative thermal protection system and designed for lifting reentry, permitting a controlled descent with the blunt end, at an angle of attack, facing forward. There are two portholes on its side, and one opening with a protruding periscope sighting device for the navigation system. All the control organs and most of the displays are located in this module.

Orbital module. The forward orbital module, about 7 ft (2 m) in diameter, is used by the crew as scientific laboratory and living quarters for orbital work, rest, and recreation. Consisting of two hemispheres joined by a cylindrical insert, it contains a "sofa" and folding table, a "sideboard," a control panel with instrumentation and equipment of the main life support system, and scientific and photographic equipment. A docking system is located at the front end of the orbital module. Originally of the probe-and-socket design, like the United States system on Apollo, for the American-Soviet Apollo-Soyuz Test Project (ASTP), this system was replaced by the peripheral "international" (androgynous) docking system designed by NASA and Soviet engineers.

Launch vehicle. The *Soyuz* launch vehicle (**Fig. 5**), capable of placing 15,000 lb (6800 kg) in orbit, has a total height of 161.7 ft (49.3 m), with 33.8 ft (10.3 m) maximum diameter. Its first stage, assisted by the four strap-on units, has a total vacuum thrust of 504 tons (4.47×10^6 newtons), generated by five engines with a total of 20 nozzles. In addition, there are 12 steering nozzles. The side units have a length of 62 ft (19 m), the central unit 92 ft (28 m). The second stage is 26 ft (8 m) in length and 8.5 ft (2.6 m) in diameter; its four-nozzle engine creates 30 tons (2.7×10^5 N) of thrust. Total lift-off weight of the *Soyuz* is about 300 metric tons. *See* ROCKET PROPULSION; ROCKET STAGING.

Flight history. Since its first, ill-fated flight in April 1967, resulting in the death of cosmonaut V. M.

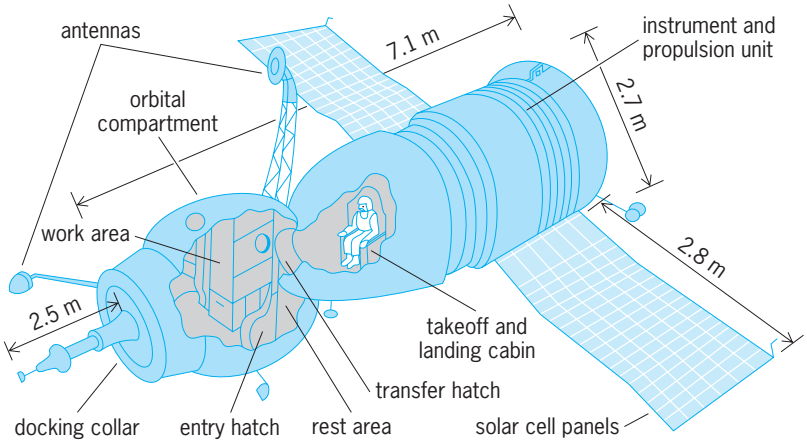

Fig. 4. ***Soyuz*, a third-generation Soviet spacecraft. 1 m = 3.3 ft.**

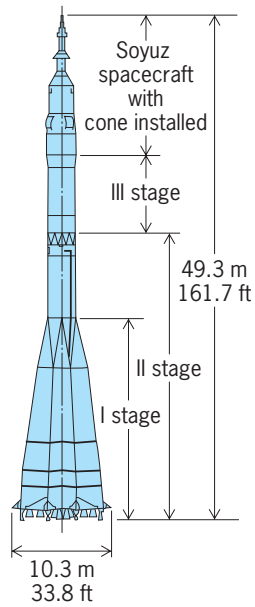

Fig. 5. ***Soyuz*** **launch vehicle SL-4. (*NASA*)**

Komarov because of failure of the parachute landing system, the *Soyuz* spacecraft has served as the mainstay of the Soviet/Russian crewed space program in many functions, for example, as a development test bed of autonomous navigation, guidance, maneuvering, and rendezvous procedures; research laboratory for biomedical and other scientific and technical studies; and logistics vehicle for the *Salyut* space stations.

Soyuz spacecraft carried originally up to three cosmonauts. However, when an accidental explosive decompression of the descent cabin during reentry caused the death (by pulmonary embolisms) of the *Soyuz 11* crew—G. T. Dobrovolskiy, V. N. Volkov, and V. I. Patsayev—on June 30, 1971, the wearing of pressure suits during launch, ascent to orbit, descent to Earth, and other critical maneuvers was reintroduced for safety. The third seat was removed to make room for the necessary additional "suit loop" of the life-support system and the emergency air supply system, and the *Soyuz* thereafter carried only two cosmonauts. An improved version, the *Soyuz T* (for "Transport") was introduced in 1980. A new version, the *Soyuz-TM*, with extended mission duration and new subsystems, replaced this vehicle in February 1987, after a crewless test flight in May 1986. (The third seat was reintroduced with *Soyuz TM-24* in August 1996.) In October 2002, the *Soyzuz TM* was, in turn, replaced by the *Soyuz TMA*, with modifications financed by NASA for improved safety and widened crewmember size range.

By December 31, 2006, there had been 94 crewed *Soyuz* missions (up to and including *TMA-9*) with 229 cosmonauts (some of them on more than one flight, and including foreign guest cosmonauts) and two aborted launch attempts (*Soyuz 18A* and an attempt on September 26, 1983). Thirty-two of these flights docked with, and transferred crews to, *Salyut* space stations, 49 with the *Mir* space station, and 13 with the *International Space Station* (*ISS*). The first mission to the *ISS* carried the first resident crew and the second carried the first "space tourist," Dennis Tito. The *Soyuz TMA* craft have served as lifeboats, providing assured crew return capability to the *ISS*, and from April 2003 through September 2006 they shouldered the burden of providing crew rotation for the station, after the loss of the shuttle *Columbia* temporarily brought shuttle operations to a standstill.

Salyut. These orbital stations were 20 m (65 ft) long and 4 m (13 ft) wide, and weighed 19 tons. Between 1971 and 1981, five of them operated in low Earth orbit successively for some time. By the end of that period, the most successful of them, *Salyut 6*, had seen seven host (long-stay) crews and nine visiting crews, including eight non-Soviet cosmonauts (from Czechoslovakia, Poland, East Germany, Hungary, Vietnam, Bulgaria, Cuba, Mongolia, and Romania). Also during that period, *Salyut 6* was resupplied by 12 *Progress* crewless automatic cargo carriers.

Salyut 6 was the first of a "second generation" of *Salyut* space stations. Unlike its predecessors, it had two docking ports, instead of only one. This enabled it to receive visiting crews and resupply ships. Together with other new on-board systems, this feature was the key to missions of considerably extended duration.

Its successor, *Salyut 7*, was launched on April 19, 1982. By the end of 1985, *Salyut 7* had seen five long-stay host crews and four visiting crews, which included the first cosmonauts from India and France as well as two flights of Svetlana Savitskaya, the second woman in space and the first female spacewalker. After serious systems problems arose in 1985, the crew of *Soyuz T13* performed extensive repair and replacement work on the station and saved it from becoming inoperative. With the flights of *Soyuz T13* and *T14*, *Salyut 7* saw the first successful space station crew rotation, although the mission of the *T14* crew had to be abruptly terminated because of illness of one of the cosmonauts. The crew of *Soyuz T15* conducted the first transfer between two space stations when they visited *Salyut 7* in May and June 1986 from the new permanent orbital facility, *Mir*.

Mir. On February 19, 1986, the Soviet Union launched the core vehicle in its new *Mir* (Peace) space station complex series. This complex represented a new-generation space station which evolved from the *Salyut* and *Kosmos* series vehicles. The core, an advanced version of *Salyut*, had six docking ports, a large number of portholes, and was 43 ft (13.13 m) long and 13.6 ft (4.15 m) wide, with a mass of 21 metric tons (23 tons). It consisted of four sections: transfer section [8.2 ft (2.2 m) diameter], work section [10.3 ft (3.15 m)], intermediate section, and equipment section [both 13.6 ft (4.15 m)], the latter containing two rocket engines of 66 lbf (294 N) thrust each. Its two winglike solar panels were larger than those of *Salyut 7*: 818 ft^2 (76 m^2) versus 441 ft^2 (41 m^2), with a wingspan of 97.5 ft (29.73 m).

The core could be expanded by individual laboratory modules launched separately and connected to

its various docking ports. These modules were usually dedicated to different scientific and technical disciplines or functions. These included technological production with a shop, astrophysics, biological research, and medical. Launch of the first block, the service module *Kvant 2*, took place on November 26, 1989. The second block, *Kristall*, launched on May 31, 1990, became the third major addition to *Mir*. Nearly identical to *Kvant 2*, this building block was equipped with furnaces for materials experiments in space and a docking port for the originally planned *Buran* shuttle, later abandoned. The third module, *Spektr*, was launched on May 20, 1995; and the fourth module, *Priroda*, was launched on April 23, 1996.

After launch, *Mir*'s laboratory modules were docked initially to the axial position on the multiport docking fixture on the forward end of the station. They then were transferred by means of their own articulated robot arm to one of the four outward-facing ports on the docking fixture of the core module, forming a spokelike array around the core.

The station complex had about 23 kW power, generated by solar-cell arrays, and about 10,600 ft^3 (300 m^3) of pressurized (habitable) volume. Maximum crew size, for relatively short durations, was six persons, limited by the number of triple-crewed *Soyuz-TM* crew transport-escape vehicles that could be docked to the station. With the launch of *Mir*, the Soviet Union had had eight different stations in Earth orbit, including one failed attempt (*Salyut 2* in April 1973).

The first *Mir* crew, Leonid Kizim and Vladimir Solovyev, was launched on board *Soyuz T15* on March 13, 1986. Docking and transfer to the station took place about 49 h later on March 15. More crews followed in succession, keeping the station continuously occupied until April 1989, when the crew of *Soyuz TM-7* left *Mir* for a crewless period until the arrival of *TM-8* in September 1989. By the time *Mir* was decommissioned and deorbited on March 23, 2001, having been in operation for 15 years, the station had been visited by 39 crews, totaling 88 cosmonauts, including seven times by the space shuttle *Atlantis* and once each by *Endeavour* and *Discovery*. During this time, *Mir* occupants set new marks three times for the longest human stay in space: 326 days by Yuri Romanenko in 1987, followed by 366 days by Vladimir Titov and Musa Manarov in 1988, and 437 days by Valery Polyakov in 1944–1995. In its later years, *Mir* required increasing maintenance and repairs by its crews, particularly after two serious emergencies in 1997: an on-board fire, caused by a backup oxygen generator device, and a collision of the automated Progress M-34 freighter with *Mir*'s *Spektr* module during a docking exercise. When U.S. support of *Mir* ended at the conclusion of the joint Shuttle/Mir program (ISS Phase 1), further crewed operations of the space station were increasingly difficult and impractical for the Russian Space Agency. The last-best-one crew—cosmonauts Sergei Avdeev and Viktor Afanasyev, and French guest cosmonaut Jean-Pierre Haigneré—returned to Earth on August 27, 1999, in *Soyuz TM-29*.

Most of *Mir*'s on-board systems were turned off on September 8. The station entered, for the first time in 10 years, a period of crewless dormancy, and preparations began for the possibility of a final decision by its Russian owners to terminate it with a controlled reentry into the Earth's atmosphere in early spring 2000. *Mir* gained another few months of life with the arrival of the final crew, Sergei Zalyotin and Alexander Kaleri in *Soyuz TM-30*, who prepared it for a flight extension. However, a deal sought by a commercial firm, MirCorp, fell through, and *Mir*'s end came on March 23, 2001, when it was deorbited and splashed into the South Pacific Ocean. *See* SPACE STATION.

Energia. With the successful first test flight, from Baikonur, of the powerful expendable heavy-lift launcher *Energia* on May 15, 1987, even if its payload did not reach orbit, the Soviet Union gained a tremendous new capability in launching large and heavy satellites, space station components, interplanetary spacecraft, and even the new Soviet space shuttle. The launcher has two stages, in parallel arrangement, and the payload is mounted not on top but on the side of the second core stage. Four liquid-fueled boosters (kerosine and liquid oxygen) make up the first stage. They are "strapped" to the cryogenic core stage, which is powered by four engines of 148 metric tons (1.45 meganewtons) thrust each (liquid hydrogen and liquid oxygen). With a total thrust at liftoff of about 3000 metric tons (30 MN) and a liftoff mass of 2650 tons (2400 metric tons), a height of only 197 ft (60 m), and a largest diameter of 65.6 ft (20 m), the *Energia* can carry 110 tons (100 metric tons) to Earth orbit or inject 20 tons (18 metric tons) to geostationary orbit, 35 tons (32 metric tons) to the Moon, and up to 31 tons (28 metric tons) to Venus and Mars.

With its second, and last, flight on November 15, 1988, the *Energia* launched the Soviet space shuttle *Buran*.

Buran. Described as the first of a fleet of Soviet shuttles, the space glider *Buran* (Blizzard) accomplished its first orbital test flight on November 15, 1988. The test mission was flown automated, that is, without crew, and lasted 3 h 25 min. The flight, while concluding with a return to the intended landing site, was only partially successful. It did not lead to further development efforts of this vehicle. The delta-winged *Buran* appeared to copy the major features of the United States shuttle, without, however, its own main propulsion system (which was provided by the expendable *Energia* core stage). The main mission of the Buran program reportedly was to support the Soviet space station *Mir* and its successors. After the collapse of the Soviet Union, the Energia/Buran programs were terminated.

European Program

European efforts in space have long had the basic goal of achieving European space autonomy: the capability of Europe to develop, launch, and operate its own satellites and space vehicles independent

of the space systems of other nations, including the United States. This goal has been accomplished for only crewless launches of commercial and scientific satellites. In order to conduct human space activities, European astronauts have made use of the United States space shuttle and the Soviet/Russian *Soyuz* launcher.

European Space Agency (ESA). The multinational European space program is the responsibility of the European Space Agency, representing 13 member nations (Austria, Belgium, Denmark, France, Germany, Ireland, Italy, Netherlands, Norway, Spain, Sweden, Switzerland, and the United Kingdom, with Finland as associate member and Canada a cooperative nation). The ESA conducts operations at field installations in the Netherlands, Germany, and Italy; tracking stations in a number of countries; and its launch center at Kourou, French Guyana.

Ariane 5. The growing weight of geostationary satellites and the economic advantage of launching two satellites on one carrier led the European Space Agency to develop a more powerful version of the successful commercial expendable launcher Ariane. Ariane 5 is a $2^1/_2$-stage launch vehicle, with a core stage (propelled by a cryogenic engine), two reusable solid boosters, similar to those of the United States space shuttle, and an upper stage. Developed by the French space agency [Centre Nationale d'Etudes Spatiales (CNES)] for the European Space Agency, Ariane 5 is able to launch payloads of up to 23 tons (21 metric tons) to low Earth orbit, and is also used for launching geostationary satellites.

The ECA (enhanced capability) version of the Ariane 5, designed to lift 10 tons (9 metric tons) to geostationary transfer orbit, enough for two big communications satellites at once, uses a new cryogenic upper stage, an improved Vulcain 2 main stage engine, and solid boosters loaded with 10% more propellant. By December 31, 2006, the Ariane 5-G (generic) had flown 24 flights and the Ariane 5-ECA had flown 7 (after initial failures of both vehicles, in June 1996 and December 2002, respectively).

United States Programs

During the first decade after its inception in 1961, the United States crewed space program was conducted in three major phases—Mercury, Gemini, and Apollo. Each crewed flight led to increased knowledge of the systems and techniques needed to operate successfully in space, and each phase represented a significant advancement.

Development. Many differences in the three crewed United States spacecraft are readily apparent, such as size and weight (**Fig. 6**). The major differences, however, are in the complexity and refinement of subsystems. Apollo's requirement for hardware "maturity" was significantly higher than for the earlier programs. Each subsystem became progressively more complex, with many more demands made upon it and a correspondingly greater capability. For example, only *Apollo* had its own guidance and navigation system. Another good example of increased system complexity was electrical power. For the *Mercury* craft, it was supplied by six batteries; *Gemini* required seven batteries and two chemical fuel-cell power plants. For *Apollo*, electrical power was supplied by five batteries and three fuel cells.

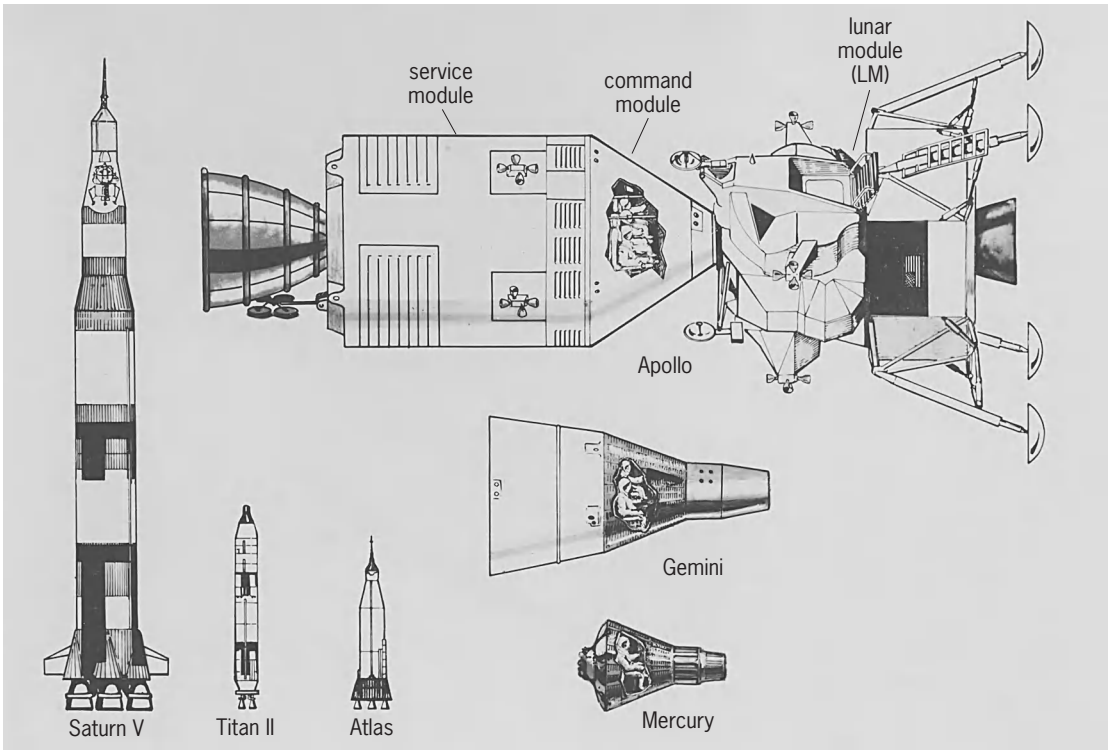

Fig. 6. Comparison of United States spacecraft and launch vehicle configurations through 1972. (*NASA*)

Physically, the three systems do not appear greatly different, but in terms of operational demand they differed considerably. The *Mercury* electrical system had to supply power to sustain the 3000-lb (1360-kg) spacecraft and its single astronaut for $1^1/_2$ days. In *Gemini*, sufficient power had to be provided to operate a 7000-lb (3175-kg) craft containing two astronauts for as long as 2 weeks, while the *Apollo* system was designed to support a 100,000-lb (45-metric-ton) spacecraft carrying three persons for up to 2 weeks.

Perhaps the greatest difference between *Apollo* and the earlier spacecraft was in reliability. Everything had to work and keep working no matter what the circumstances. Unlike the previous crewed space programs, in which the crew could return to Earth almost within minutes if an emergency arose, the *Apollo* crew needed as much as 6 days to get back from a position outbound toward the Moon from the Earth's vicinity, requiring lunar circumnavigation to return safely to Earth. Such a situation occurred in the flight of *Apollo 13* in 1970 (Table 1).

Mercury. After Alan Shepard's first ride on *Mercury 3* atop a Redstone rocket, John Glenn in 1962 became the first American to orbit the Earth. His vehicle, a *Mercury*, was in its basic characteristics similar to the Soviet *Vostok*, but it weighed only about a third as much, as necessitated by the smaller missiles of the United States at that time (the Redstone and Atlas).

The one-person *Mercury* capsules (Fig. 3) used ballistic reentry and were designed to answer the basic questions about humans in space: how they were affected by weightlessness, how they withstood the gravitational forces of boost and entry, how well they could perform in space. The *Mercury* flights proved that humans not only could survive in space, but could also greatly increase humanity's knowledge of the universe. Between 1961 and 1963, a total of six astronauts flew on *Mercury* spacecraft, the last flight (Gordon Cooper's *Faith 7*) lasting a record 34 h. *See* WEIGHTLESSNESS.

Gemini. The second United States step into space was the Gemini Program. With the two-person *Gemini* capsule (**Fig. 7**), for the first time a crewed spacecraft had been given operational maneuverability in space. Its reentry module flew a lifting reentry trajectory for precise landing point control (Fig. 1). In addition, its design permitted extravehicular activity by one of the crew. The goal of the *Gemini* missions was to find out how humans could maneuver themselves and their craft, and to increase knowledge about such things as celestial mechanics, space navigation, and rendezvous and docking, as required later for the Apollo Program. All these extra capabilities required that the weight of the spacecraft be increased by 130% over the *Mercury*. This in turn called for a more powerful launch vehicle, the Titan 2, which had its origin as an Air Force ICBM.

Gemini has a record of 10 successful crewed flights between 1965 and 1966, with 20 astronauts and a total mission duration of 969 h 52 min in space.

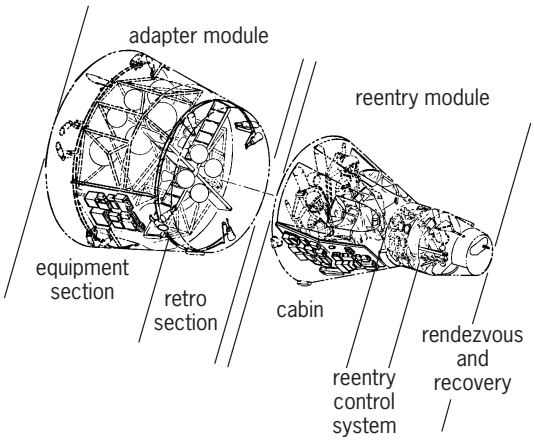

Fig. 7. ***Gemini*** **spacecraft. (*NASA*)**

It set many records, including the longest duration (almost 14 days, *Gemini 7*), the first rendezvous by two maneuverable spacecraft (*Gemini 6* and *7*), and the first docking (*Gemini 8*). Moreover, on June 3, 1965, *Gemini 4* crew member Edward White performed the first spacewalk by an American astronaut, drifting and maneuvering in space for 21 min at the end of a 25-ft (7.5-m) umbilical/tether. On later *Gemini* missions, EVAs were performed by E. A. Cernan, M. Collins, R. F. Gordon, and E. E. Aldrin, totaling 13 h 23 min, including open-hatch or stand-up EVAs (SEVAs).

Apollo. The third-generation spacecraft, *Apollo*, was 82 ft (25 m) tall and had five distinct parts: the command module (CM), the service module (SM), the lunar module (LM), the launch escape system (LES), and the spacecraft/lunar module adapter (SLA). The three modules made up the basic spacecraft; the LES and SLA were jettisoned early in the mission after they had outlived their function.

Command module. The command module served as the control center for the spacecraft and provided 210 ft^3 (6 m^3) of living and working quarters for the three-member crew for the entire flight, except for the period when two persons entered the lunar module for the descent to the Moon and return. The command module was the only part of the spacecraft that returned to Earth, flying a lifting trajectory with computer-steered maneuvers (Fig. 1).

Resembling a cone of almost 10-ft (3-m) height and 12.7-ft (3.9-m) width, the command module consisted of two shells separated by a layer of insulation: an inner crew compartment (pressure vessel) and an outer heat shield.

Service module. Attached to the command module at its heat shield end was the service module, a 22.8-ft-tall (7-m), 12.7-ft-wide (3.9-cm) cylinder housing the electrical power subsystem, most of the reaction control engines, part of the environmental control subsystem, and the service propulsion subsystem, including the main rocket engine for insertion into orbit around the Moon, for return from the Moon, and for course corrections.

Lunar module. The lunar module (**Fig. 8**) carried two astronauts from the orbiting command/service mod-

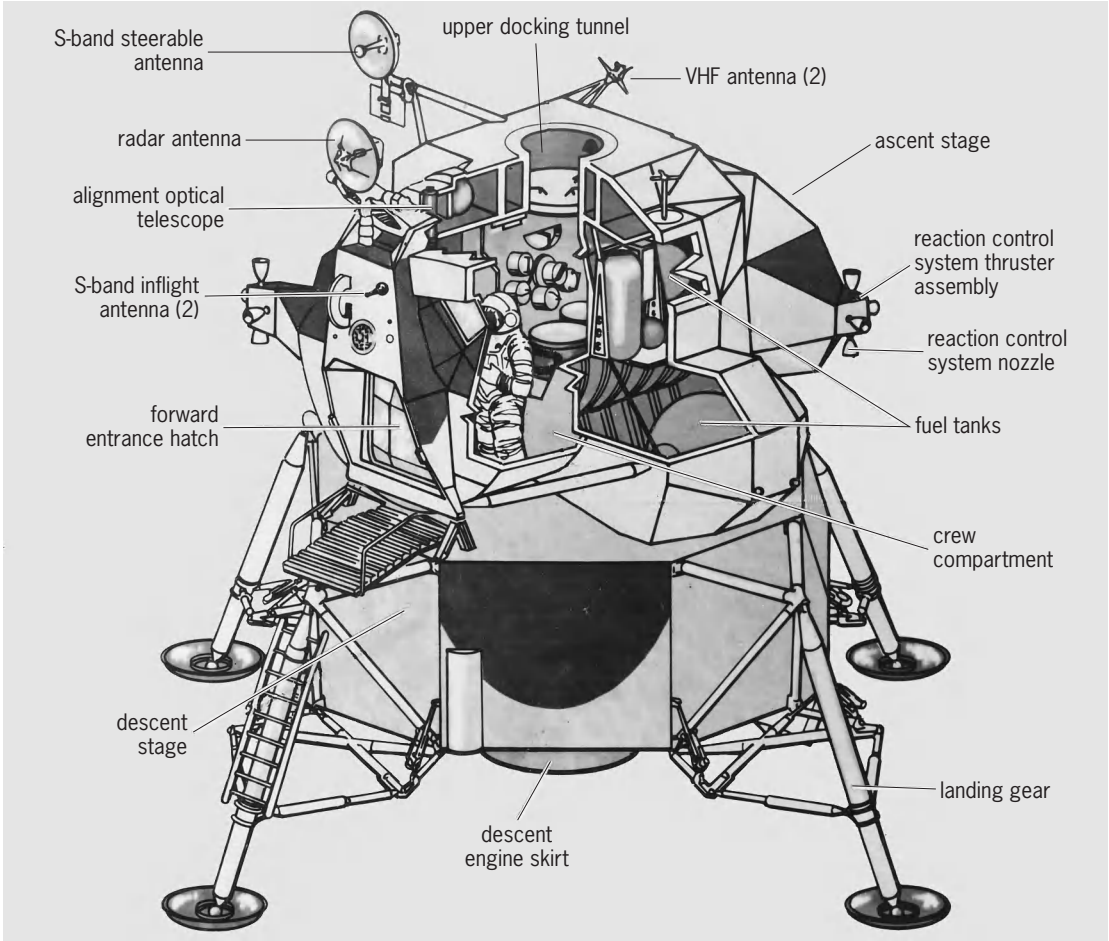

Fig. 8. ***Apollo*** **lunar module. (*NASA*)**

ule (CSM) down to the surface of the Moon, provided a base of operations there, and returned the two astronauts to a rendezvous with the command service module in orbit. It consisted of two components: the ascent stage (on top) and the descent stage. The latter had a descent engine and propellant tanks, landing gear assembly, a section for scientific equipment for use on the Moon, another section to transport a collapsed wheeled vehicle on the last three *Apollo* flights (the lunar rover), and extra oxygen, water, and helium tanks. The ascent stage contained the crew compartment with about 160 ft^3 (4.6 m^3) of habitable volume, the ascent engine with propellant tanks, and all lunar module controls. Its subsystems were essentially the same kinds as those of the command module and service module. The lunar module measured about 23 ft (7 m) in height (legs extended), weighed about 32,500 lb (14.6 metric tons), and was constructed of aluminum with titanium fittings.

Flight history. Including Earth-orbital flights by *Apollo 7* and *9*, a total of 11 crewed missions were launched in the Apollo Program between 1968 and 1972 (Table 1). Of these, nine crews (27 astronauts) went to the Moon, with six landings (**Fig. 9**). Twelve astronauts walked on the Moon, accumulating almost 300 h of lunar stay time and 166 person-hours of surface exploration. They traversed almost 60 mi (97 km) and brought about 850 lb (385 kg) of rock and soil samples back to Earth. Approximately 33,000 lunar photographs and 20,000 reels of magnetic tapes of geophysical data were collected, causing a major revolution in planetary science by providing the first deep understanding of a celestial body other than Earth. *See* MOON.

Lunar roving vehicle. On the last three lunar landings, *Apollo 15–17*, lunar exploration was supported by the lunar roving vehicle (LRV; **Fig. 10**). This vehicle weighed 460 lb (209 kg) on Earth (77 lb or 35 kg Moon weight) and 1600 lb (725 kg) fully loaded. Measuring 122 in. (3 m) in length and 6 ft (2 m) in width, it could carry two astronauts in their space suits and backpack life support units, scientific equipment, and lunar samples. Each wheel was driven by its own separate traction drive assembly consisting of a $^1/_4$-horsepower (0.2-kW) dc brush-type electric motor, harmonic drive reduction gear, and brake assembly. Power was provided by two nonrechargeable 36-V silver-zinc batteries, each containing 23 cells, having a capacity of 115 ampere-hours and weighing 59 lb (27 kg). Both rear and front wheels were steered by a hand-operated control stick through two separate electric motors on the axles. Turning radius was 122 in. (3.1 m). The lunar roving

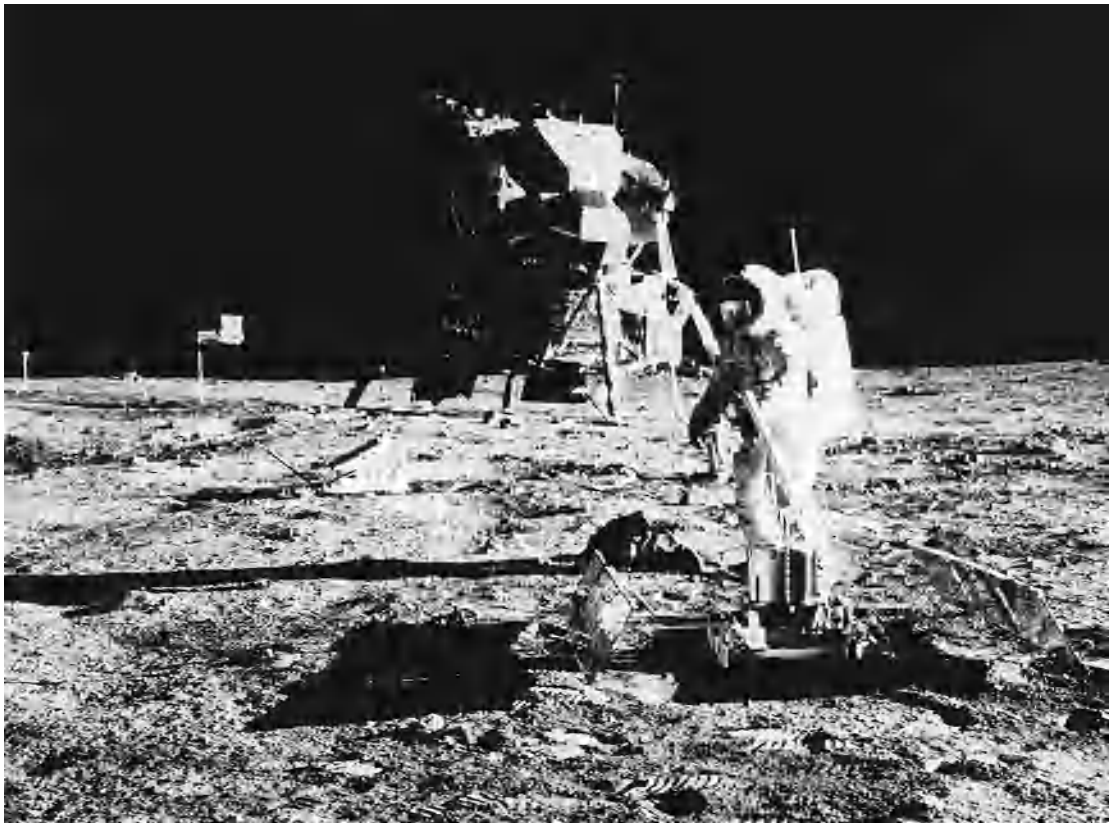

Fig. 9. Astronaut Edwin Aldrin, Jr., stands next to the passive seismometer experiment on the surface of the Moon. On July 20, 1969, Aldrin and Neil Armstrong became the first humans to land on the Moon. A television camera, American flag, laser reflector, and lunar module Eagle are behind Aldrin. (*NASA*)

vehicle had a range of 36 mi (58 km), traveled at a maximum speed of 8.5 mi/h (13 km/h), and negotiated slopes as steep as 25. During its transport to the Moon, it was stowed in folded and collapsed condition in the lunar module descent stage in a volume measuring 4.4 × 4.4 × 5.4 ft (1.3 × 1.3 × 1.6 m). After touchdown, semiautomatic deployment was accomplished by the astronauts' pulling nylon tapes and releasing cables. By using the lunar roving vehicle, the astronauts were able to spend more time on the Moon's surface during exploration.

Extravehicular activity. Extravehicular activities became of paramount importance in Apollo. In addition to the 12 astronauts who walked on the Moon, four crew members performed in-flight EVAs, with a total of 3 h 46 min. Counting the nine space exits by *Gemini* astronauts, a total of 41 individual EVAs had been accumulated by the end of the Apollo Program. To these, *Skylab* astronauts later added 19 more (10 occasions, **Table 2**).

Launch vehicle. Prior to Apollo, the United States' launch vehicle inventory included the Atlas-Agena B with 5000 lb (2250 kg) Earth-orbital payload capability, and the Titan 2 with 7800 lb (3500 kg). Clearly, a more powerful booster was required to lift the *Apollo* spacecraft to Earth orbit and thence to the Moon.

At the Army Ballistic Missile Agency (ABMA) in 1958, a team of engineers under Wernher von Braun set out to prove that vastly more powerful space

Fig. 10. Lunar roving vehicle used in the *Apollo 15* mission. (*NASA*)

TABLE 2. *Skylab* mission summary

Crewed periods	First	Second	Third	Total
Launch	5/25/73 9:00 a.m. EDT	7/28/73 7:10 a.m. EDT	11/16/73 9:01 a.m. EDT	
Splashdown	6/22/73 9:49 a.m. EDT	9/25/73 6:19 p.m. EDT	2/8/74 11:17 a.m. EDT	
Duration (day:h:min)	28:0:49	59:01:9	84:01:16	171:13:14
Revolutions	404	858	1214	2476
Distance, 10^6 mi (10^6 km)	11.5 (18.5)	24.5 (39.4)	34.5 (55.5)	70.5 (113.5)
SEVA*	0:37 (5/25/73)			
EVA† 1 duration (h:min)	3:30 (6/7/73)	6:29 (8/6/73)	6:33 (11/22/73)	
EVA 2 duration	1:44 (6/19/73)	4:30 (8/24/73)	7:01 (12/25/73)	
EVA 3 duration		2:45 (9/22/73)	3:28 (12/29/73)	
EVA 4 duration			5:19 (2/3/74)	
Total EVA	5:51	13:44	22:21	41:56
Solar observatory photos	30,242	77,600	75,000	182,842
Earth resources photos	8,886	14,400	17,000	40,286

* Standup (in spacecraft hatch) extravehicular activity.
† Extravehicular activity (completely outside of spacecraft).

rockets could be built from existing hardware by clustering engines and tanks. The project evolved into the Saturn Program of NASA. There were three types of launch vehicles in the Saturn family: the Saturn 1 which had a perfect record of ten successful flights, the Saturn 1B with eight successful flights, and the Saturn 5, which became the Apollo lunar launch vehicle and had the capability to lift 250,000 lb (113 metric tons) into low Earth orbit and to send 100,000 lb (45 metric tons) to the Moon. The development of the Saturn 5 proceeded through three phases, with the two preceding launch vehicles as stepping stones. Saturn 1 used primarily modified existing equipment. Saturn 1B used a modified first stage of the Saturn 1 with uprated engines and a new second stage and instrument unit. Saturn 5 introduced new first and second stages, departing from the clustering approach, and used as its third stage and instrument unit the Saturn 1B's second stage and instrument unit (**Table 3**).

TABLE 3. Saturn launch vehicles*

Launch vehicle and stage	Length, ft	Weight, 10^3 lb	Payload to low orbit, lb	Diameter, ft	Engines, number and type	Propellants, weight and, type†	Stage thrust,‡ 10^3 lb
Saturn 1, block I (with payload)	162.5	1120 (gross)	No orbital capability				
1st stage, S-1	81.6	103 (dry)	—	21.4	8 H-1's	875,000 lb, RP-1/LOX	1504 s.l.
2d stage, dummy	44	142 (ballast)		18.3			
Saturn 1, block II (with payload)	187.5	1128 (gross)	24,000				
1st stage, S-1	80.3	105.7 (dry)		21.4	8 H-1's	875,000 lb, RP-1/LOX	1504 s.l.
2d stage, S-4	41.4	13.9 (dry)		18.3	6 RL10A-3's	100,000 lb, LH_2/LOX	90 v.
Instrument unit	2.8	3.1		21.4			
Saturn 1B (with spacecraft)	224	1313 (gross)	40,000				
1st stage, S-1B	80.3	84.4 (dry)		21.4	8 H-1's	885,000 lb, RP-1/LOX	1650 s.l.
2d stage, S-4B	58.4	25 (dry)		21.7	1 J-2	230,000 lb, LH_2/LOX	200 v.
Instrument unit	3	4.3		21.7			
Saturn 5 (with spacecraft)	363	6200 (gross)	270,000				
1st stage, S-1C	138	303 (dry)		33	5 F-1's	4,578,000 lb, RP-1/LOX	7600 s.l.
2d stage, S-2	81.5	95 (dry)		33	5 J-2's	942,000 lb, LH_2/LOX	1160 v.
3d stage, S-4B	58.13	33.6 (dry)		21.7	1 J-2	232,000 lb, LH_2/LOX	230 v.
Instrument unit	3	4.3		21.7			

†1 ft = 0.3 m. 1 lb = 0.45 kg. 1 lbf = 4.4 newtons.
†RP-1 = type of kerosine, LOX = liquid oxygen (−297°F, −183°C), LH_2 = liquid hydrogen (−422.9°F, −252.8°C).
‡s.l. = thrust at sea level, v. = thrust in vacuum.

Fig. 11. Space station *Skylab* in orbit. The sunshade rigged by crew repair is visible. (*NASA*)

In addition to its lunar mission in the Apollo program, the Saturn 5, in a two-stage version, served also—in its last and thirteenth flight—to launch the first United States space station, *Skylab*.

Skylab

The experimental space station *Skylab* (**Fig. 11**) was the largest object ever placed in space, and the first crewed project in the U.S. Space Program with the specific purpose of developing the utility of space flight in order to expand and enhance humanity's well-being on Earth. To that end, *Skylab's* equipment included Earth resources remote sensing instrument and the first crewed solar telescopes in space. A total of three crews of three astronauts each carried out experiments and observations on *Skylab* in a number of different areas, among them observations of the Earth, solar astronomy, stellar astronomy, space physics, geophysics, biomedical studies, zero-gravity, biological studies, zero-gravity technology studies, and spacecraft environment. Originally begun in the early 1960s, under the names Apollo Extension Support Study and Apollo Applications Program, *Skylab* emerged as a cluster of five major components: orbital workshop (OWS), airlock module (AM), multiple docking adapter (MDA), Apollo telescope mount (ATM), and command/service module (CSM). The CSM, adapted from the Apollo Program, served as crew carrier, launched by Saturn 1B's.

Skylab was launched crewless by Saturn 5 on May 14, 1973 (*Skylab 1*), and suffered a structural failure of its combined meteoroid/thermal shield during the first minute of flight which eliminated most of its electrical power supply. Subsequently visited by three astronaut crews for periods of 28, 59, and 84 days (Tables 1 and 2), the space station underwent repair in orbit, a first for the space program, which proved that humans are eminently capable of performing heavy emergency repair, service, and maintenance operations in space, thereby setting the stage for the future of orbital operations by space shuttle crews.

After the third crew had returned to Earth on February 8, 1974, the United States first space station remained dormant for 5 years while NASA worked on plans to save *Skylab* from orbital decay and reentry for future missions. The intention was given up in December 1978 when it became clear that the estimated decay rate of the orbit was approaching faster than anticipated and the likely date of space shuttle rescue mission was moving ahead. Under active control by NASA, *Skylab* finally returned to Earth on July 11, 1979, impacting in the Indian Ocean, with a few pieces landing in Australia.

Apollo-Soyuz Test Project

The last flight of an *Apollo* spacecraft occurred as the United States part in the world's first international crewed space flight, the Apollo-Soyuz Test Project on July 15–24, 1975. During that time, three American astronauts and two Soviet cosmonauts linked up in space for a total joint period of 47 h 17 min (**Fig. 12**).

The ASTP mission was planned to accomplish spacecraft rendezvous, docking, undocking, and crew transfer between spacecraft of disparate technical and management philosophies, the Soviet *Soyuz* and the United States *Apollo*, as well as interaction of spacecrafts crews. The successfully conducted mission tested compatible rendezvous and docking systems being developed for future United States and Soviet crewed spacecraft and space stations; in addition, the crews practiced docking and undocking, conducted intravehicular crew transfer, and performed 31 scientific experiments and engineering investigations.

Docking. The new compatible rendezvous and docking system provided the basis for a standardized international system for docking of crewed spacecraft. Thus, it enhanced the safety of crewed flights in space and provided the opportunity for conducting subsequent joint missions and experiments.

To enable *Apollo* and *Soyuz* to dock, NASA developed and constructed a third element, the docking module (DM), that also served both crews as an airlock and transfer tunnel between the spacecraft (**Fig. 13**). It was 10 ft (3.2 m) long and 4.6 ft (1.4 m) wide, and weighed 4660 lb (2097 kg). It held two suited crew members and was attached to the forward end of *Apollo*. On the *Soyuz* side, it used the compatible docking system designed by NASA and Soviet engineers (**Fig. 14**). This docking system was of the androgynous type, which uses the principle of

Fig. 12. An artist's concept of the first international "handshake" in space, by the American and Soviet crews of the Apollo-Soyuz Test Project mission, July 17, 1975. (*NASA*)

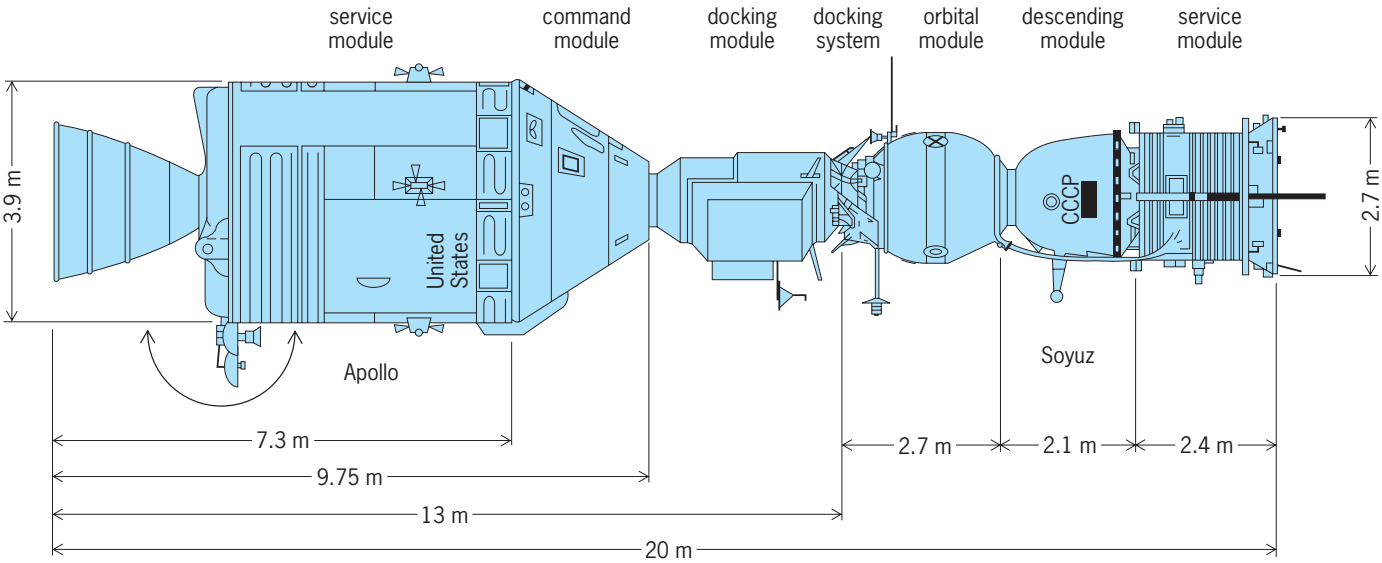

Fig. 13. *Apollo* and *Soyuz* spacecraft in docked configuration during the Apollo-Soyuz Test Project mission with docking module serving as connecting link and airlock. 1 m = 3.3 ft. (*NASA*)

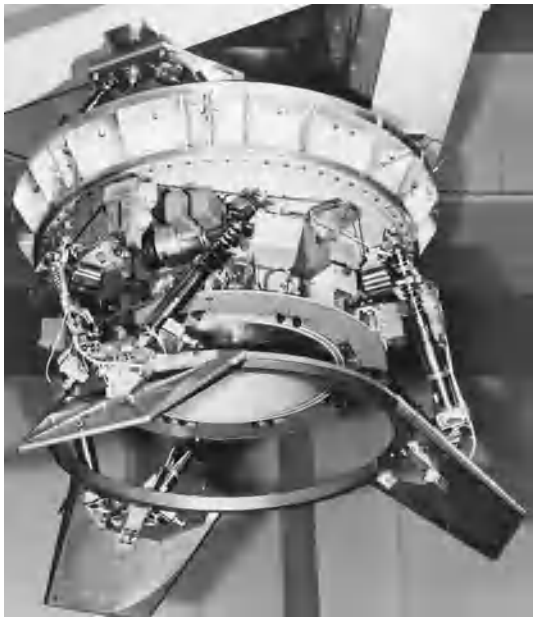

Fig. 14. Universal docking system, as used on Apollo-Soyuz Test Project (American version), suspended for testing. (*NASA*)

reverse symmetry to allow the mechanical linkage of otherwise dissimilar spaceships. Docking is usually performed by one (active) spacecraft on the other (passive) partner.

Since the *Soyuz* internal atmosphere consisted of nitrogen-oxygen at sea-level pressure as opposed to *Apollo's* pure oxygen environment at 5 lb/in.^{2} (34.5 kPa), the docking module was required to serve as an air lock. To reduce lengthy waiting periods for denitrogenation, the *Soyuz* pressure was reduced to about 10 lb/in.^{2} (69 kPa) before docking. During the mission itself, four transfers between spacecraft were accomplished by crew members of both nations.

Spacecraft modifications. To fit the mission needs, several modifications were made to both spacecraft. For *Apollo*, they included an increased number of propellant tanks for the reaction control system, added equipment to operate the docking module and the American-Soviet rendezvous and docking systems, and provisions for scientific and technical experiments. Among the major modifications of *Soyuz* were the new docking mechanism, additional communications equipment to accommodate the United States ultrahigh frequency of 296 MHz, a transponder (a combined receiver-transmitter that beams a signal when triggered by another radio signal) to allow *Apollo* to calculate distance during rendezvous, and alignment devices to aid *Apollo* in performing its active docking role.

Accomplishments. In addition to the rendezvous and docking test, the ASTP demonstrated that precision launch of complex vehicles from widely separated sites to meet the trajectory alignment requirements for a joint space mission was feasible, and that hardware, operating, and language differences can be surmounted for joint international space efforts, with enough planning and training in advance.

Space Transportation System

After the end of the joint American-Soviet space mission, the United States crewed space flight program had reached the end of a pioneering era of capability development. Attention began to focus on the routine application of newly acquired know-how systems and experience, specifically in the form of the emerging Space Transportation System (STS), with the ability to transport inexpensively a variety of useful payloads to orbit, as the mainstay and "work horse" of the United States space program in the 1980s and 1990s. Its design is aimed at reducing the cost and increasing the effectiveness of using space for commercial, scientific, and defense needs.

The two major components of the Space Transportation System are the space shuttle and the *Spacelab*. This system is complemented by a number of upper stages depending on the mission beyond Earth orbit, for example, the Payload Assist Module (PAM), the Transfer Orbit Stage (TOS), and the Inertial Upper Stage.

Space shuttle. The space shuttle (**Fig. 15**), as a reusable system, has the capability to conduct space missions in response to national and worldwide needs and the flexibility to respond to policy,

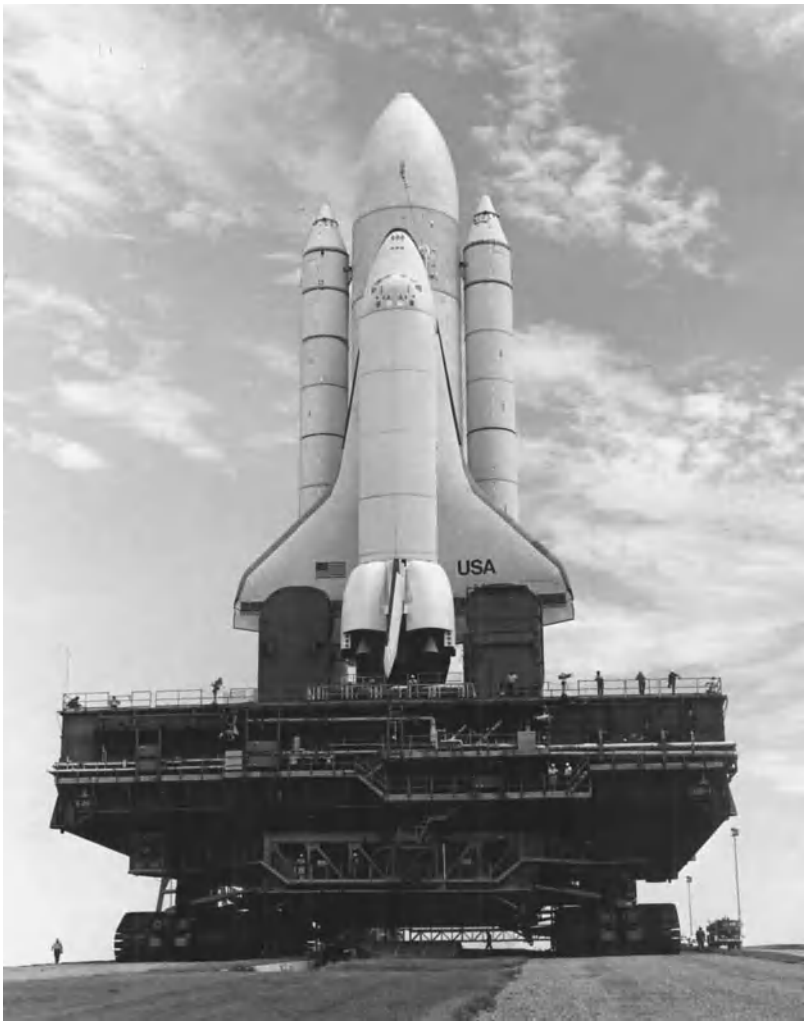

Fig. 15. Space shuttle on its crawler/transporter/launch platform. (*NASA*)

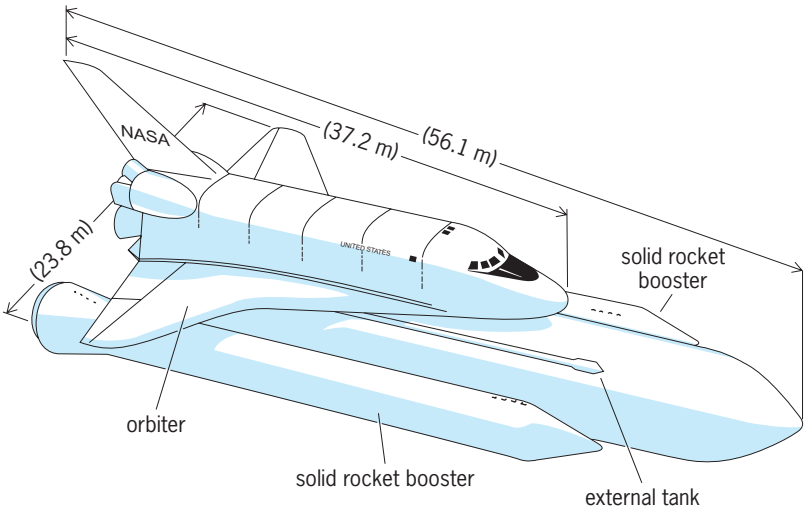

Fig. 16. Space shuttle vehicle. 1 m = 3.3 ft. (*NASA*)

discovery, innovation, and newly emerging needs. The shuttle provides extended capability in the areas of (1) easy access to and from low Earth orbit for nonastronaut passengers, (2) multipurpose reusable transportation, (3) standard equipment for a variety of spacecraft, (4) standard equipment for a variety of missions, (5) modularized subsystems for easy refurbishment, (6) refurbishable spacecraft/payloads in orbit, and (7) a technology base for more flexible future missions. The space shuttle flight system (**Fig. 16**) has three major elements: the orbiter, an external tank (ET) containing the liquid propellants to be used by the orbiter main engines for ascent, and two solid-propellant rocket boosters (SRBs) "strapped on" to the external tank. The orbiter and solid rocket booster casings are reusable; the external tank is expended on each launch. The total weight of the system at launch is 4,406,000 lb (1,998,500 kg).

Orbiter and external tank. The orbiter contains the crew and payloads. In size and weight (150,000 lb or 68,000 kg when empty of consumables) comparable to a DC-9 jet transport, it delivers payloads of 65,000 lb (29,500 kg) with lengths to 80 ft (18 m) and widths to 15 ft (4.5 m). The crew compartment accommodates seven crew members and passengers for some missions, but as many as ten persons in emergency operations.

The three main propulsion rocket engines in the aft fuselage of the orbiter, with a vacuum thrust of 470,000 lbf (2,090,000 N) each, use liquid hydrogen and liquid oxygen from the external tank, which is jettisoned shortly before orbit insertion after approximately 8 min of burn time. The orbiter also carries 46 auxiliary engines with a total vacuum thrust of 45,200 lbf (201,000 N). The external tank has a length of 153.7 ft (46.85 m) and a diameter of 331.0 in. (8.41 m). Its weight is 1,619,000 lb (734,400 kg) at launch and 69,000 lb (31,300 kg) when empty of fuel.

Solid rocket boosters. The two solid rocket boosters burn in parallel with the main engines of the orbiter to provide initial ascent thrust. Primary elements of each solid rocket booster are the motor, including casing, propellant, igniter and nozzle, forward and aft structures, separation and recovery systems, and thrust vector control subsystems. Each solid rocket booster has a length of 149.1 ft (45.45 m), a diameter of 146 in. (3.71 m), a weight of 1,265,000 lb (573,800 kg) at launch and 182,000 lb (82,600 kg) after the fuel has been expended, and a thrust of 2,700,000 lbf (12,000,000 N) at launch. The solid propellant grain is shaped to reduce thrust some time after liftoff to prevent overstressing the vehicle during the period of maximum dynamic pressure. The solid rocket boosters separate after about 2 min of flight from the external tank at an altitude of 30 mi (50 km) by pyrotechnic devices and eight small rockets; they descend on a system of parachutes and splash in the ocean about 170 mi (274 km) from the launch site—Kennedy Space Center in Florida or Vandenberg Air Force Base in California—to be retrieved for reuse.

Testing and flight. The first shuttle orbiter, named *Enterprise*, was rolled out at Palmdale, California, on September 17, 1976. Used for ground tests and approach and landing flights tests during 1977 and 1978, the *Enterprise* was superseded by the first space mission orbiter, *Columbia*, for the first orbital flight on April 12–14, 1981, piloted by astronauts John Young and Robert Crippen (**Fig. 17**). Three additional shuttles—*Challenger, Discovery*, and *Atlantis*—were added subsequently, so that by the end of 1985 the total shuttle fleet consisted of four vehicles. Planning for the construction of a possible fifth orbiter was held in abeyance pending a better understanding of future launch-frequency requirements.

By mid-January 1986, 24 shuttle missions to low Earth orbits has been conducted (up to and including *61-C*), all of them successfully, although a number of problems had to be overcome in the course of the program (Table 1). Highlights of these flights included the retrieval of two errant communications satellites (on *STS 51-A*), the in-space repair of two satellites (on *STS 41-C* and *STS 51-I*), the first untethered extravehicular activity of a human with an autonomous maneuvering system (on *STS 41-B*; **Fig. 18**), assembly of two large space structures (*STS 61-B*), and three flights of the space laboratory *Spacelab* (*STS 9, STS 51-B*, and *STS 61-A*).

With the launch of the twenty-fifth shuttle mission, *51-L*, on January 28, 1986, however, tragedy struck the American space program. At approximately 11:40 a.m. EST, 73 s after liftoff, the flight of *Challenger*, on its tenth mission, abruptly ended in an explosion triggered by a leak in the right solid rocket booster, killing the seven-member crew.

The loss of *Challenger* reduced NASA's shuttle fleet by one-fourth. An independent commission to investigate the causes of the *51-L* accident was established. Continuation of the United States crewed space program was suspended pending a thorough reassessment of flight safety issues and implementation of necessary improvements. The recovery period lasted 32 months (compared to 18

Fig. 17. Launch of the space shuttle *Columbia* on its first orbital flight, April 12, 1981. (*NASA*)

months after the *Apollo 1* fire in 1967 which killed three astronauts). During this time, all shuttle systems, operations, and its entire management were subjected to meticulous reviews and modifications, overviewed by outside expert committees. The solid booster joints, whose inherent design weaknesses had caused the explosion of the *Challenger*, were redesigned with 73 changes and tested thoroughly. The orbiter itself, in the process, underwent 228 modifications, not counting 100 software changes. The main propulsion system (63 changes) and even the external tank were improved, along with the assembly and launch facilities, as well as all preflight procedures. In July 1987, NASA ordered a new orbiter as *Challenger*'s replacement, called *Endeavour*.

Shuttle flight operations resumed on September 29, 1988, with the long-awaited launch of *Discovery* on the mission *STS 26* to place a *Tracking and Data Relay Satellite* (*TDRS*) in orbit. Subsequent flights followed with increasing confidence in the new systems and operations. At the end of 1989, after seven highly successful missions since the *Challenger* accident, NASA was ready to resume full-scale space shuttle operations, signaling a total recovery from the accident.

Highlights of post-*Challenger* missions included the launch of the interplanetary research probes *Magellan* to Venus (on *STS 30*) and *Galileo* to Jupiter [on *STS 34*), the retrieval of the 10.5-metric-ton (11.5-ton) research satellite Long Duration Exposure Facility (LDEF) [on *STS 32*], the placement in orbit of the *Hubble Space Telescope* (on *STS 31*), the launch of the European solar probe *Ulysses* (on *STS 41*), and the intricate repair and maintenance missions to the *Hubble* telescope (on *STS 61, 82, 103, and 109*). On *STS 60*, in February 1994 the first Russian cosmonaut, Sergei Krikalev, participated as a crew member, and in March 1995 *Endeavour*, on mission *STS 67*, established a long-duration record for the space shuttle program of 16 days 15 h 9 min. On *STS 71*, in June 1995, 20 years after the Apollo-Soyuz Test Project, *Atlantis* carried out the first of a series of dockings with the Russian *Mir* space station (**Fig. 19**), forming the largest human-made structure in space so far. Joint activities were carried out and crew members exchanged between the two spacecraft before separation. Eight more dockings between space shuttles and *Mir* were conducted in 1995–1998 as Phase 1 of the *ISS* program, with seven U.S. astronauts stationed for extended durations on the Russian facility. Since 2000, most of the shuttle missions have been in support of the assembly and operation of the *International Space Station*.

On February 1, 2003, after completing a successful research mission (*STS 107*), *Columbia* was lost with its crew during reentry. The Columbia Accident Investigation Board, chaired by Admiral (ret.) Harold W. Gehman, Jr., concluded that, unbeknown to crew

Fig. 18. Astronaut Bruce McCandless performing first free (untethered) extravehicular activity with an autonomous maneuvering system during shuttle flight *STS 41-B* on February 7, 1984.

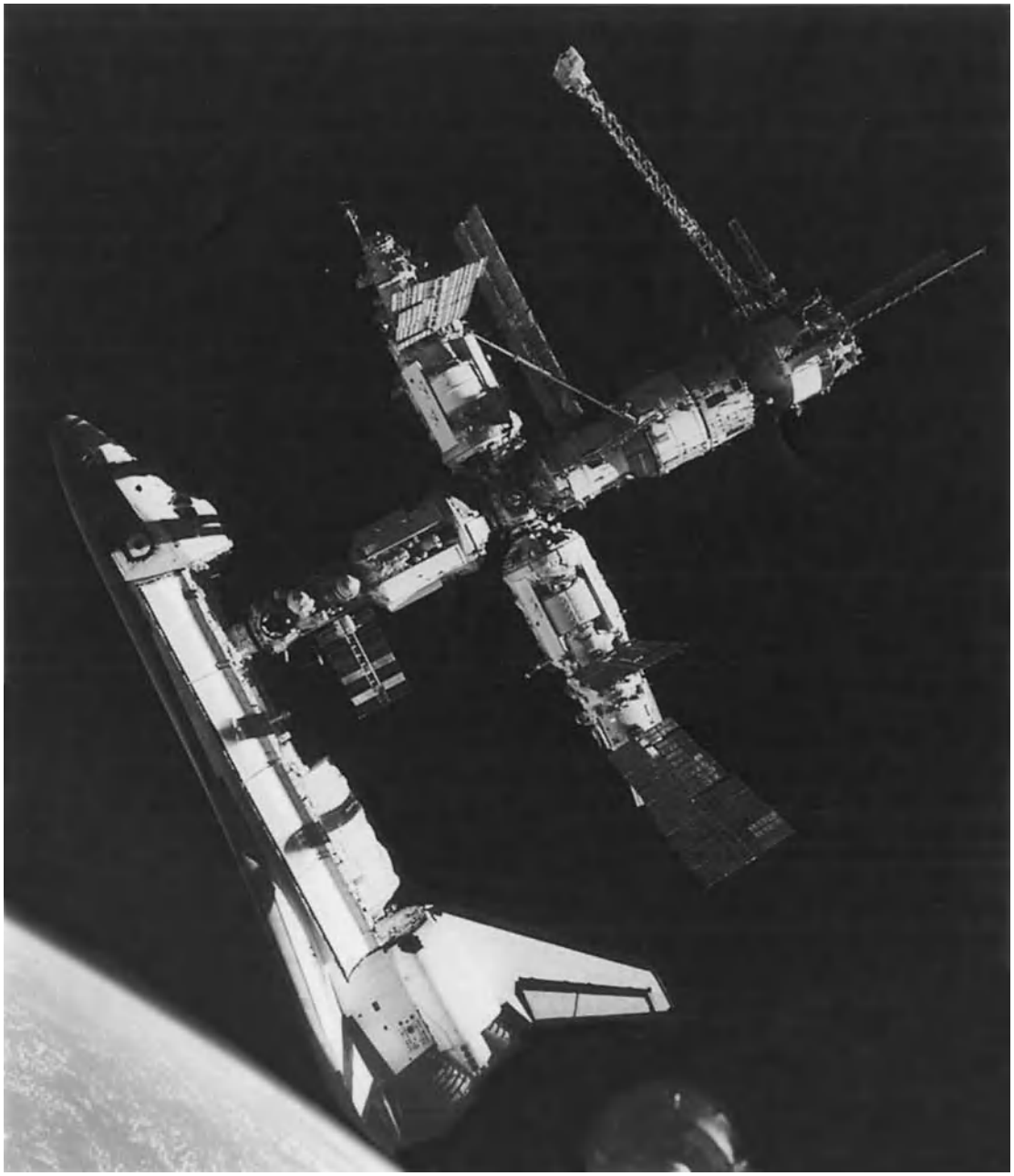

Fig. 19. Space shuttle *Atlantis* docked to the *Kristall* module on the Russian space station *Mir*, photographed by cosmonant Nikelai Budarin on July 4, 1995. (*NASA*)

and ground, one of the left wing's leading-edge reinforced carbon-carbon (RCC) elements had been punctured during ascent to orbit by a "suitcase-sized" chunk of foam insulation, blown off the external tank by the supersonic airstream, hitting the leading edge and rendering the wing unable to withstand reentry longer than about 8 min after entry interface. Shuttle operations were halted as NASA and its contractors labored on intensive return-to-flight (RTF) efforts.

The space shuttle returned to the skies on July 26, 2005, with the liftoff of *Discovery* on mission *STS 114* to the *ISS*. The flight was successful but engineers were concerned about the shedding of some insulating foam material off the external tank. Prior to its final approach to the station, *Discovery* performed a maneuver to allow digital imagery of its thermal protection system from the *ISS*, and during docking crewmembers performed a first-of-its-kind spacewalking heat shield repair. With more work required on redesign of the foam-based external tank insulation, new sensors for detailed damage inspection, and a boom to allow astronauts to inspect the vehicle externally during flight, the second RTF test flight was delayed for nearly a year. However, with test flights in July and September 2006 and an additional flight in December, the shuttle fleet resumed its operations.

As of December 31, 2006, the four U. S. shuttles, in over 24 years of operation, have flown a total of 117 missions, carrying 597 men and 101 women into space (including the lost crews of seven on *STS 51-L/Challenger* and *STS 107/Columbia*). Participating

in these flights were 81 crew members from 16 foreign nations. *See* SPACE SHUTTLE.

Spacelab. The *Spacelab* was a major adjunct of the shuttle. Developed and funded by member nations of the European Space Agency, the large pressurized *Spacelab* module with an external equipment pallet was designed to be the most important payload carrier during the space shuttle era. With its large transport capacity (in weight, volume, and power supply), the *Spacelab* was intended to support a wide spectrum of missions in science, applications, and technology by providing versatile and economical laboratory and observation facilities in space for many users. Another objective was to reduce significantly the time from experiment concept to experiment result, compared with previous space practice, and also to reduce the cost of space experimentation. It allowed the direct participation of qualified men and women scientists and engineers to operate their own equipment in orbit. Primary uses of the *Spacelab* were identified in the areas of process engineering and materials research, biomedical research, Earth observations and meteorology, communication and navigation, space flight technology, astrophysics, astronomy, solar physics, energy research, atmospheric and space physics, and basic chemical and physical sciences.

The *Spacelab* design consisted of three elements: (1) a pressurized module which could be assembled from modular sections in two sizes—a short version for missions requiring only minimal supplemental subsystems in a "shirt-sleeve" environment, and a longer version for more extensive experiments; (2) a pallet or instrument mounting platform to support sensors and antennas requiring vacuum and uninterrupted space exposure; and (3) support equipment such as air locks, viewports, storage vaults, and pointing systems. Depending on assembled size, the *Spacelab* was designed to have a total pressurized volume of 880–3500 ft^3 (25–100 m^3). In some missions, a pallet-only mode could be flown. *Spacelab*'s average power use was designed to be in the 3–6-kW range. Orbiting the Earth at 115–445 mi (185–555 km) altitude and 28–104° inclination on missions lasting 7–30 days, the *Spacelab* was designed to remain firmly attached inside the open shuttle cargo bay. A connecting tunnel was designed to provide access to and from the shuttle orbiter crew compartment. The laboratory was designed to provide working space for up to four payload specialists, not counting the commander, pilot, and mission specialist (flight engineer) or the space shuttle itself.

The first flight of the European-built *Spacelab, SL-1*, took place November 28–December 8, 1983, on board the *Columbia* as the ninth shuttle mission. Its six-member crew included the first foreign astronaut in the United States space program, the German mission specialist Ulf Merbold. *SL-1* and subsequent *Spacelab* flights, notably the first foreign chartered mission, the German *D-1*, but also the Japanese *SL-1* on *STS 47* and the second German *Spacelab* mission, *D-2* on *STS 55*, proved the concept of the crewed space laboratory and confirmed the unique worth of humans in space to operate, repair, maintain, and troubleshoot scientific equipment and conduct complex research programs. *Spacelab* flew 16 missions before it was decommissioned in 1998. *See* SPACE BIOLOGY; SPACE PROCESSING.

International Space Station

Interest in the development of a permanent crewed platform in Earth orbit dates back to the very beginnings of human space flight. While the practical realization of the concept was accomplished and proven by the Soviet Salyut/Mir and the American Skylab programs, the real breakthrough happened on January 29, 1998, when representatives of 16 nations (United States, Russia, Canada, Japan, the 11 nations of the European Space Agency (ESA), and Brazil) signed a partnership agreement for the joint development and operation of an *International Space Station* (*ISS*). With so many nations sharing in this unique megaprogram, the *ISS* has become a focal point of international understanding, cooperation, and collaboration, and thereby a strong catalyst for peace enhancement and conflict reduction.

The goal of the *ISS* program is to establish a permanent platform for humans to live and work in space in support of science and technology research, business, education, and exploration. Its objectives are to provide a world-class laboratory complex uniquely located in a microgravity and vacuum environment where long-term scientific research can be carried out to help fight diseases on Earth, unmask fundamental processes leading to new manufacturing processes and products to benefit life on Earth, and observe and understand the Earth's environment and the universe. The station is expected to create new opportunities for the private sector for business development and entrepreneurial ventures. As a "virtual classroom" in space, the *ISS* will inspire and support educators and students worldwide. As a testbed for studies of long-term effects of space on humans and for the development of new space-flight technologies, it will also become an essential gateway to future human space exploration.

When completed, the *International Space Station* will have a mass of about 1,040,000 lb (470 metric tons). It will be 356 ft (109 m) across and 290 ft (88 m) long, with almost an acre (0.4 hectare) of solar panels to provide up to 110 kilowatts of power to six state-of-the-art laboratories.

1998. The on-orbit assembly of the *ISS* began with the launch of the U.S.-owned, Russian-built Functional Cargo Block (FGB) *Zarya* ("Dawn") on November 20, 1998, from the Baikonur Cosmodrome in Kazakhstan on a Proton rocket. The second element, the U.S.-built Node (connecting module) Unity followed on the space shuttle *Endeavour*, launched on December 4, 1998, from the Kennedy Space Center, Florida. The shuttle crew attached Unity and *Zarya* during a 12-day mission, beginning the station's orbital construction which will extend over 12 years until its completion in 2010.

1999. Accomplishments during 1999 included the first crewed logistics-supply flight of a space shuttle to the early assembly in orbit in May and June; testing

of hardware and software for station elements to be launched in 2000 and 2001; successful implementation of a newly developed Multi-Element Integrated Test (MEIT), which functionally connects the first several U.S. elements of the station on the ground in a series of integrated tests; the completion of the Mission Control Center and the Payload Operations Integration Center; and delivery of the U.S. laboratory module Destiny and the airlock module to the launch site in Florida in preparation for MEIT, along with other components. Interface testing between the U.S. Mission Control Center in Houston and the Russian Mission Control Center in Moscow was successfully completed.

2000. Among the key events in 2000 was the July 12 launch of the service module *Zvezda* (Star), the first Russian-funded and owned element of the station, which linked up with the station on July 25. The docking, executed very gently at 0.6 ft/s (0.2 m/s) by *Zarya* as the active partner, marked the arrival of the first human living quarters and an initial command post for the station. *Zvezda* is also equipped with the propulsion system that will maintain the attitude and orbit of the space station. The station now spanned 119 ft (36 m) in length, and its mass had grown to almost 60 tons (54 metric tons). On August 8 the automated *Progress MI-3* docked with the station, carrying fuel for *Zvezda*'s propellant tanks, as well as equipment and supplies.

On October 31, a Soyuz-U carrier lifted off from Baikonur and placed in orbit *Soyuz TM-31*, carrying the first resident crew for the space station. The Expedition One crew consisted of American Commander William Shepherd and veteran Russian cosmonauts Yuri Gidzenko and Sergei Krikalev. On November 2, the *Soyuz* craft docked with the station and the crew entered, to begin what is intended to be a new era of permanent human habitation in space. The Expedition One crew remained on the station for 136 days, until March 18, 2001.

An *Endeavour* assembly mission on November 30-December 11 installed two solar array wings, jointly spanning 240 ft (73 m), at the space station (**Fig. 20**). The array is now providing the station with up to 62 kW of power.

2001. *ISS* assembly continued in 2001, notably with the installation of the United States Laboratory module *Destiny* on February 10. The 28 × 14 ft (8.5 × 4.3 m) laboratory was transported on *Atlantis*. The Canada-built Space Station Remote Manipulator System (SSRMS) and the U.S. airlock module *Quest* also arrived on space shuttles, and the Russian Docking Compartment (DC-1) *Pirs* was carried on a Russian Soyuz-U rocket, *Quest* provided station residents with their own "door" to space for extravehicular activities (EVAs), and DC-1 gave the station a second airlock for spacewalks in Russian Orlan-M spacesuits.

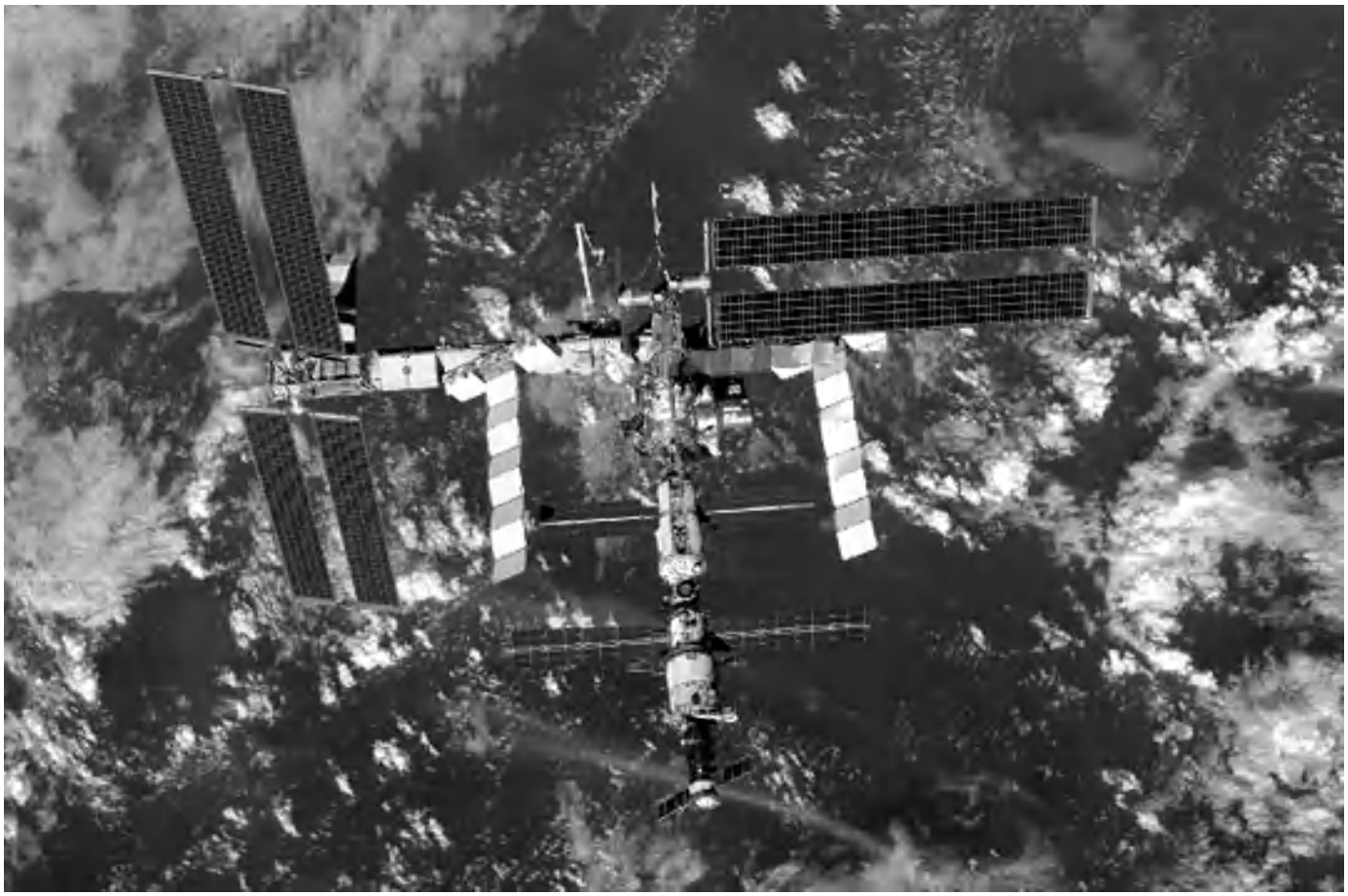

Fig. 20. ***International Space Station*, photographed from the space shuttle *Discovery* on December 19, 2006. (*NASA*)**

In the first crew rotation, the Expedition Two crew, consisting of Russian commander Yuri Ussachev and American flight engineers James Voss and Susank Helms, replaced Expedition One in March. In August, Expedition Two was replaced by the Expedition Three crew, consisting of American commander Frank Culbertson and Russian cosmonauts Vladimir Dezhurov and Mikhail Tyurin. The crews in both rotations were transported on the space shuttle *Discovery*. Successive expedition crews have served on the *ISS* during stays of four to six months.

2002. In April 2002, the first of eleven truss elements, S0 (Starboard Zero), was attached on top of *Destiny*, becoming the centerpiece of the 109-m-long (356-ft) truss for carrying the solar cell arrays of the station. In June, a new station crew (Expedition 5) delivered the Mobile Base System (MBS), to provide mobility for the SSRMS. The second truss segment, S1, arrived in October and was attached to S0 on the starboard side. Its counterpart on port, P1, followed in November and was also successfully mounted.

2003–2005. Early in 2003, further progress in *ISS* assembly was brought to a halt by the standdown of the space shuttles after the *Columbia* loss. Crew rotation flights could be accomplished only by Russian Soyuz vehicles. As an immediate consequence of the unavoidable reduction in resupply missions to the station, which now could be supported only by Russian crewless automated Progress cargo ships, station crew size was reduced from three to a two-person "caretaker" crew per expedition, except for brief stays by visiting researchers or commercial "tourists" arriving and departing on the third seat of Soyuz spacecraft. Operations procedures had to be revised accordingly, and such vital areas as onboard systems maintenance and spares provision had to be replanned carefully to continue crewed occupancy and a viable science research program on board despite the sudden constriction in logistics.

2006. With the return to flight of the shuttle, *ISS* assembly resumed. In July the shuttle carried a third astronaut (Thomas Reiter) to the station to join Expedition 13. In September, the P3/P4 truss structure was installed with a second set of solar arrays (of four), and the two solar array wings and a huge radiator panel were deployed from the P4 truss. In December, the P5 truss segment, serving as a spacer at the end of the P4 truss, was installed, and the electrical power and thermal control systems were reconfigured to their final configurations and activated.

As currently planned, the station will require a total of 80 assembly flights. Interspersed among these missions will be logistics missions by the shuttle, Russian *Soyuz* flights to launch crews, and multiple *Progress* tanker flights for refueling and supplying the growing structure in orbit.

Vision for Space Exploration

On January 14, 2004, U.S. President Bush announced the Vision for Space Exploration, a long-term plan for returning astronauts to the Moon to prepare for voyages to Mars and other planets in the solar system. The plan called for completion of the *International Space Station* and retirement of the space shuttle by 2010, and the development of new spacecraft and launch systems.

Project Constellation. To support this initiative, NASA is developing a family of spacecraft, launchers, and associated hardware, collectively known as Project Constellation.

The Orion spacecraft will be shaped like an Apollo capsule but will be significantly larger. It will have a crew module shaped similarly to the Apollo command module, capable of holding four to six crew members, and a cylindrical service module, containing the spacecraft's propulsion systems and supplies. It will be reusable for up to 10 flights, allowing the construction of a fleet of such spacecraft. Plans call for it to be launched into low-earth orbit, either for solo flights or for missions to the *International Space Station*, using the proposed Ares I rocket, which would consist of a single solid rocket booster derived from those used in the space shuttle system, and a liquid-fueled second stage.

Lunar flights would involve the use of a larger launch vehicle, the proposed Ares V, which would use five RS-68 engines (more powerful and less expensive than the space shuttle main engines) and a pair of five-segment solid rocket boosters. (Alternative plans have also been considered which would replace the Ares I and V rockets with a single launcher, the Ares IV.) The Ares V would carry a combination of the Earth Departure Stage (EDS), an enlarged version of the S-IVB upper stage of the Saturn V rocket, and the Lunar Surface Access Module (LSAM). The Ares V would fly for about 8 minutes before separating from the EDS/LSAM combination, which the EDS would then place into low-Earth orbit. The crewed Orion spacecraft would be launched separately on the Ares I at least two to four weeks later, and dock with the EDS/LSAM combination. After the systems were configured for lunar flight, the EDS would inject the combination into a translunar orbit and then be jettisoned.

Like the Apollo lunar module, the LSAM would consist of an ascent stage, which would house the four-person crew, and a descent stage, which would have the landing legs, most of the crew's oxygen and water, and scientific equipment. Unlike the Apollo lunar module, the LSAM's engines would place the Orion/LSAM combination in lunar orbit. The entire crew would transfer to the LSAM, which would undock from Orion and descend to the lunar surface. After completing operations there, the crew would enter the LSAM's ascent stage and lift off, powered by a single engine and using the descent stage as a launch pad. After docking with and transferring to the Orion spacecraft in lunar orbit, the crew would jettison the LSAM and fire the Orion's single engine for the return journey to Earth.

Lunar base. In December 2006, NASA announced plans to eventually construct a solar-powered lunar base located near one of the poles of the Moon,

rather than landing at several sites, as was done in the Apollo missions. An incremental buildup is envisioned, with astronauts making several seven-day visits to the Moon until their power supplies, rovers, and living quarters are operational. These visits would be followed by 180-day missions to prepare for journeys to Mars. Robotic precursor missions would provide landing site reconnaissance, natural resource assays, and risk reduction for the human landers. A decision on a site will probably not be made until after collecting data from several such missions. The *Lunar Reconnaissance Orbiter* (*LRO*) is scheduled to launch in 2008 to scout the lunar surface and search for resources. The *Lunar Crater Observation and Sensing Satellite* is to launch with the LRO and then crash into the lunar surface to search for water ice.

Jesco von Puttkamer; Jonathan Weil

Bibliography. W. Ley, *Rockets, Missiles, Space Travel*, 1969; NASA, *Report on the 90-Day Study on Human Exploration of the Moon and Mars*, 1989; Office of Exploration, NASA Headquarters, *Beyond Earth's Boundaries: Human Exploration of the Solar System in the 21st Century*, 1989; S. Ride, *Leadership and America's Future in Space*, Report to the Administrator, NASA Headquarters, 1987; *Round Trip to Orbit: Human Spaceflight Alternatives*, Special Report, OTA-ISC-419, U.S. Congress, 1989; W. Sweetman (ed.), *Jane's Space Directory 2006–2007*, 22d ed., 2006; Synthesis Group, *America at the Threshold: America's Space Exploration Initiative*, 1991; T. D. Thompson (ed.), *TRW Space Log*, TRW Space and Technology Group, annually; W. von Braun et al., *Space Travel: A History*, rev. ed., 1985.

Space navigation and guidance

The determination of the position and velocity of a space probe with respect to a target body or a known reference body such as the Earth (navigation) and, based upon this determination, the application of propulsive maneuvers to alter the subsequent path of the probe (guidance).

Navigation. If the position and velocity of a space probe, plus some secondary quantities such as the positions and masses of any perturbing celestial bodies, can be determined at a given instant, then the future path of that vehicle can be predicted for all time by the application of the laws of celestial mechanics. A simple way of understanding this is that, an instant later, the new position of the space probe will be the original position incremented by the velocity times the time duration. Similarly, the new velocity will be the original velocity incremented by the acceleration acting on the probe times the time duration. Additionally, the acceleration at the new time will be slightly different since gravitational accelerations acting on the probe depend on the position. Other accelerations may also affect the probe, such as thrust due to engine firings, solar radiation pressure, drag, and so forth. Thus, at the future time, the probe will arrive having a new position and a new velocity; these coordinates can be used as the basis for predicting an additional short interval into the future. By taking sufficiently short steps, a very accurate path of the future trajectory can be calculated. *See* CELESTIAL MECHANICS.

Space navigation can be viewed as determining the current position and velocity of the probe and then using that determination as the basis of predicting future motion. This determination is made by taking a series of measurements relating to the probe's location and motion and combining these measurements in such a manner as to make the most accurate estimate of the probe's current position and velocity, taking into account possible small errors or inaccuracies in the measurements themselves. One of the most powerful (and accurate) measurements which can be made is the relative velocity between an Earth tracking station and the space probe itself. This is accomplished by broadcasting to the probe an electromagnetic signal which consists of a single tone having a stable frequency. The probe will receive this signal shifted in frequency in exact proportion to the relative station-probe velocity and then immediately will rebroadcast it back to the Earth. This frequency shift is known as the Doppler effect. It is also possible to measure the station-probe distance (or range) using electromagnetic signals. Modern systems can make these measurements to an accuracy of about 0.02 mm/s (0.001 in./s) for the velocity and 1 m (3 ft) for the range at distances extending to the outer limits of the solar system. An additional, very powerful measurement is known as delta-differential-one-way-range (denoted Δ-DOR or delta-DOR). It is performed by simultaneously observing both a space probe and a nearby extragalactic radio source (quasar) from two tracking stations. These four simultaneous measurements are differenced to form an extremely accurate estimate of the angular distance between the probe and quasar as seen in the sky. The determination of the position and velocity of the probe relative to the Earth and hence the Sun can be made by observing these quantities over a period of time. *See* DOPPLER EFFECT.

The probe's orbit relative to the Earth, thus determined, is most useful when the target body's position relative to the Earth is also accurately known. This is true for lunar missions and missions to the planets. For missions to the satellites of the outer planets (Jupiter and beyond) or Pluto, for a mission to a body such as a comet or asteroid, the position of the target may be sufficiently uncertain as to make a strictly Earth-relative navigation scheme inadequate. Here it is necessary to make measurements that directly involve the target. Typical of these is to obtain optical measurements of the locations of the target relative to a star as seen from the probe itself using an onboard camera. Again, a series of measurements of the angular distance between the target and a known star as the probe approaches can be used to determine the probe's position and velocity, this time relative to the target.

Guidance. In its simplest form, guidance consists of comparing the predicted future motion of the probe against that which is desired, and if these are sufficiently different, executing a propulsive maneuver to modify that future position. Such maneuvers are called trajectory correction maneuvers (TCMs). Typically, the probe will contain a number of small rocket motors which can be fired in any desired direction by first rotating the spacecraft away from its cruise orientation and then holding the new attitude fixed while the rocket motor is firing. The cruise attitude is achieved by orienting the spacecraft relative to celestial references. Typically, the roll axis of the probe is oriented toward the Sun. A roll of the probe until a predetermined star enters the field of view of a fixed star sensor is then sufficient to fix the probe's complete attitude in space. The reorienting of the probe for the desired rocket firing is accomplished by gyroscopes within the probe which measure and command the desired turns from the celestial reference orientation. An alternative, for both probes cruising in a fixed orientation as well as probes that are designed to spin about a primary axis, is to fire multiple thrusters oriented in different directions either in sequence or simultaneously such that the combination of the several maneuvers adds up to the desired modification to the trajectory. *See* STAR TRACKER.

Modern missions generally employ sequences of these basic navigation and guidance functions. For example, for a typical trajectory to Mars, the navigation plan will call for a maneuver (TCM-1) shortly after launch (usually about 10 days) to correct for most of the errors due to the launch vehicle. A month or two later, TCM-2 is performed to correct for most of the errors from TCM-1. As the probe approaches Mars, one to three or even four more TCM's will be performed depending on a variety of requirements related to the specific mission. For each TCM, Earth-based tracking (Doppler, ranging, and usually Δ-DOR) will be used to estimate the trajectory prior to performing the maneuver. Missions to bodies requiring optical navigation will also include this data for the final approach maneuvers.

An entirely different method of making very large changes to the orbit of a space probe utilizes the gravity of intermediate bodies rather than the engines on the probe. These multiplanet missions require not only complex trajectory design techniques but also very accurate navigation to the intermediate targets. Such missions may gather scientific information at both the intermediate bodies and the ultimate target or may simply use the intermediate bodies as an extremely efficient means of getting to the final target. At its simplest, the method of "gravity-assist" uses a close flyby of a massive body to exchange energy and momentum between the space probe and the body. Although large changes to the orbit of the probe are made, no discernible effect is observed on the massive body due to the enormous difference in mass between the probe and flyby body. **Illustration** *a* shows that in the frame of reference of the flyby body (Jupiter in this case) the speed of the probe both approaching and departing Jupiter is the same although the direction changes. In illus. *b* it is seen that since Jupiter is moving in the Sun's frame of reference both the speed and direction of the probe are changed. In the 1970s the first missions to use gravity-assist were *Mariner Venus/Mercury* and both the *Pioneer* and *Voyager* missions to the outer solar system. Since that time *Galileo* (launched in 1989) and *Cassini-Huygens* (launched in 1997) have used gravity-assist both for the interplanetary transfer and in orbit at Jupiter and Saturn, respectively. Missions have been launched to the extremes of the Solar System (*MESSENGER* in 2004 to Mercury orbit using Earth, Venus, and Mercury flybys, and *New Horizons* in 2006 to Pluto using a Jupiter flyby), and the *Stardust* mission utilized an Earth flyby on its way to comet Wild 2 and returned a sample of comet dust to the Earth. Other than further missions to Mars, almost all of the future exploration of the solar system may make use of gravity-assist. All of these orbit designs

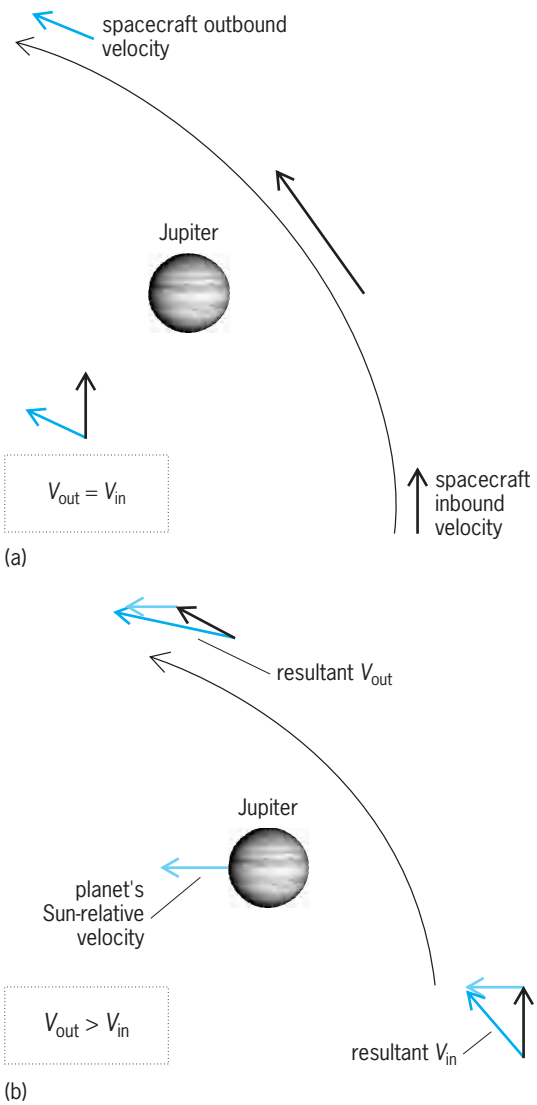

Jupiter flyby of a space probe in (*a*) Jupiter's frame of reference and (*b*) the Sun's frame of reference.

include a series of trajectory correction maneuvers on approach to each flyby body as well as one or more "clean-up" trajectory correction maneuvers shortly after the flyby.

The guidance maneuvers discussed so far are designed to make small corrections to the planned orbit. For missions designed to orbit another body, a large terminal guidance maneuver, also called an orbit insertion maneuver (for example, MOI for Mars orbit insertion, or JOI for Jupiter orbit insertion), is applied at the encounter with the target body. These maneuvers are quite large and may use the large majority of the propellant carried by the probe. For example, the *Galileo* JOI was about 0.65 km/s (0.4 mi/s), while MOIs for *Mars Odyssey* (2001), *Mars Express* (2003), and *Mars Reconnaissance Orbiter* (2006) were about 1.4 km/s (0.9 mi/s), 0.8 km/s (0.5 mi/s), and 1.0 km/s (0.6 mi/s), respectively. For orbiter missions, navigation and guidance takes place much as for interplanetary cruise with both orbit determination and correction maneuvers being performed on a regular basis. *See* GUIDANCE SYSTEMS; SPACE PROBE.

Nomenclature. Often "navigation" is used to mean all of orbit design, orbit determination, and maneuver design and execution. The terminology "guidance and control (G&C)" or "guidance, navigation, and control (GN&C)" is then used to denote the operation of the space probe itself—knowledge and control of its pointing and orientation in space plus the timing and firing of the onboard engines.

Dennis V. Byrnes; David W. Curkendall

Bibliography. R. H. Battin, *An Introduction to the Mathematics and Methods of Astrodynamics*, 1987, paper 1999; M. Kayton (ed.), *Navigation: Land, Sea, Air and Space*, 1990; C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*, rev. ed., 1995; W. G. Melbourne, Navigation between planets, *Sci. Amer.*, 234(6):58–74, June 1976; W. E. Wiesel, *Spaceflight Dynamics*, 2d ed., 1996.

Space power systems

On-board assemblages of equipment to generate, store, and distribute electrical energy on satellites and spacecraft. A reliable source of electrical power is required for supplying energy to the spacecraft and its payloads during launch and through several years of operational lifetime in a space environment. Present-generation spacecraft fly power systems from tens of watts to several kilowatts. Each of the three fuel cells on the space shuttle delivers 12 kW continuous power and 16 kW peak power. The *International Space Station*'s solar arrays will generate 110 kW total power, with approximately 46 kW available for research activities. *See* SATELLITE (SPACECRAFT); SPACE FLIGHT; SPACE PROBE; SPACE SHUTTLE; SPACE STATION.

With few exceptions, the power systems for United States satellites have used photovoltaic generation for power, batteries for energy storage, and a host of electrical equipment for appropriate regulation, conversion, and distribution. Fuel cells have been limited primarily to use in the crewed space program, supplying power for *Gemini, Apollo*, and the space shuttle. Radioisotope thermoelectric generators (RTGs) have powered, or augmented solar power on, many planetary missions and probes, powered lunar instruments left on the Moon by the *Apollo* missions, and augmented the solar-array battery power system on at least one Earth-orbiting spacecraft.

Many factors influence the final configuration of a power system. Basic to the initial consideration are the nature of the mission (Earth-orbiting or planetary) and mission lifetime. Other relevant factors include (1) spacecraft and payload requirements with consideration to average and peak power loads; (2) effect of environment such as orbit period, length of time the spacecraft is in sunlight and shadow, radiation particle damage, and space charging; and (3) constraints imposed by the spacecraft such as weight, volume, spacecraft shape and appendages, satellite attitude control, electromagnetic radiation limitations, characteristics of payloads, and thermal dissipation. For spacecraft that are shuttle-launched, additional considerations include compatibility with shuttle payload safety requirements and, in some cases, the retrievability by the shuttle for return to Earth. The weight of the power system ranges from 15 to 25% of the spacecraft weight.

Solar cells. The photovoltaic effect was discovered in 1877, but was not exploited until the silicon cell was invented in 1954. The first cells had conversion efficiencies of approximately 6%, and were little more than laboratory curiosities until the space program began in 1958.

The first solar cells flown, *pn* cells (a thin *p* layer diffused into an *n*-type substrate), were selected because of the lower sensitivity of the material. However, these cells degraded in a radiation environment much more rapidly than the *np* cells (a thin *n*-layer diffused into a *p*-type substrate). Since the early 1960s, the *np* cell has been used almost exclusively for United States spacecraft.

Development efforts for solar cells have emphasized improved conversion (sunlight to electrical) efficiency, increased resistance to charged-particle radiation, reduced manufacturing costs, and reduced weight. Conversion efficiency obtained in the conventional solar cells was approximately 11% until 1972, when the violet cell was perfected with an efficiency of over 13%.

The following additions have pushed silicon solar-cell efficiency up to about 14%: a p^+ layer (called a black-surface field) that reduces hole-electron recombinations at the cell's back surface; a textured front surface and multilayer antireflection coatings that decrease reflected light; and a back-surface mirror that reduces operating temperature. For silicon cells, maximum achievable efficiency appears to be 16–18%. Consequently most development efforts are focused on improving the design of the array or perfecting the gallium aluminum arsenide/gallium arsenide (GaAlAs/GaAs) cell technology.

The potential of GaAlAs/GaAs solar cells to obtain efficiencies of 20–25% has prompted major developments in this technology. A liquid-phase epitaxially grown GaAlAs/GaAs cell has achieved efficiencies as high as 18%. A metal/organic chemical vapor deposition process for cell fabrication and a multibandgap gallium aluminum arsenide/gallium indium arsenide (GaAlAs/GaInAs) cell that uses aluminum and indium material to enhance efficiency are other GaAlAs/GaAs developments. Developments anticipated for GaAlAs/GaAs cells indicate higher-performance cells with significant cost savings for space application. *See* SOLAR CELL.

Solar arrays. The process of assembling individual cells into series and parallel circuits to obtain the required working voltage and power is intricate and complex. There are two basic cell-mounting configurations. Body-mounted arrays, because of limited available space, occasionally use a flexible shingle arrangement. This arrangement usually provides more cells per unit area, thereby providing a higher output for a given area. However, flat mounting, with space provided between cells for electrical interconnects, is most frequently used (**Fig. 1**).

Many solar arrays now in use are in the 1–10-kW range, such as the array for the *Solar Maximum Mission* (**Fig. 2**). The four solar arrays for the *International Space Station* produce 31 kW each.

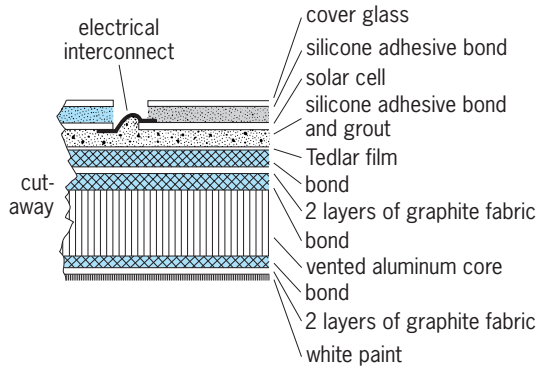

Fig. 1. Cross section of a solar array, illustrating the materials and components used in assembly with the flat-mounting technique.

The flexible roll-up solar-array (FRUSA) technology was first demonstrated in 1971 on an Air Force flight-test validation program. Solar arrays such as the FRUSA have a specific power of 20 W/lb (44 W/kg). Studies for a solar-array-powered propulsion stage have demonstrated the feasibility of a 12.5-kW array with a specific power of 30 W/lb (66 W/kg), and work has been undertaken to eventually achieve a specific power of 90 W/lb (200 W/kg).

The *Hubble Space Telescope* used two rolled-up arrays. They developed cooling-heating problems

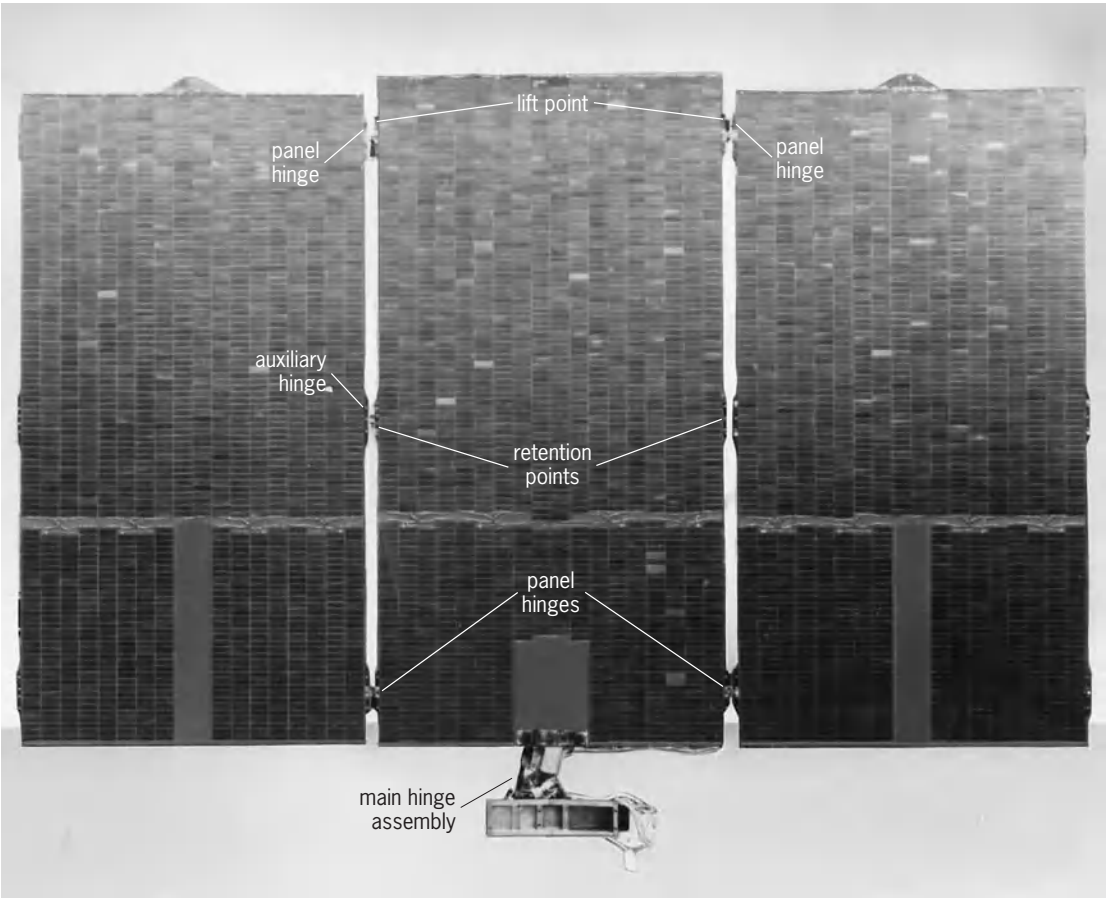

Fig. 2. One of the two solar-array wings flown on the *Solar Maximum Mission* spacecraft. Overall dimensions are 7.2 by 11.5 ft (2.2 by 3.5 m).

which caused vibration during the transitions of the telescope from shadow to sunlight. The arrays were redesigned and replaced during a telescope reservicing mission. *See* HUBBLE SPACE TELESCOPE.

With the advent of GaAlAs/GaAs cells, solar arrays using solar concentrators have been studied. The useful range of conventional solar arrays for deep-space missions is limited to about 1.5 astronomical units (AU) [Mars orbit]. Studies have shown that, with solar concentration, the photovoltaic power system's useful range can be extended to at least 9 AU [Saturn orbit]. A concentrator system can be used with either silicon or GaAlAs/GaAs cells. However, the GaAlAs/GaAs solar arrays potentially have several advantages over silicon solar arrays, including higher beginning-of-life (BOL) efficiencies, less degradation in a radiation environment, higher operating temperatures, more rapidly increasing performance with concentrator ratios, and greater ability to recover from radiation damage at lower temperatures. With concentrator ratios of 8:25, an overall power density at beginning of life of 9 W/lb (20 W/kg) is projected. The potentially low cost of concentration may make this approach economically more attractive than RTGs.

One concentrator system is aboard an HS 702 geosynchronous spacecraft, designed by Hughes Space and Communications Company. The satellite, *Galaxy XI*, was launched on December 21, 1999. The solar arrays were deployed on January 21, 2000, and for the first time television cameras were used to watch the deployment. Most concentrator designs use a lens to focus more sunlight onto the cells. The Hughes design used reflector panels on each side of the solar panel. The cells were Ga-As and produced 12.4 kW at approximately 25% efficiency.

Another mission that used concentrators was *Deep Space 1*. It had 720 lenses to focus sunlight onto the solar arrays, producing 2500 W at Earth's orbit.

Other proposals have been made to improve the efficiencies and thus power outputs of solar arrays. One proposal is to stack cells on top of each other, each efficient at absorbing certain bandgap frequencies (colors) of light but allowing other frequencies to be transmitted to the next cell. A GaAs three-junction multibandgap cell has a calculated efficiency of over 34%. Another proposal would use the same principle, but instead of stacking, the cells would be placed side by side and a prism would spread the white light into the rainbow of colors that the cells would absorb. Five different cells would have approximately 32% efficiency in normal sunlight, and if concentrators were used before the prisms to concentrate the light by eight times, the efficiency would be 43%. *See* SOLAR ENERGY.

Batteries. Although both primary and secondary (rechargeable) batteries have been used in the space program, satellites now in use almost always employ rechargeable batteries. Five different rechargeable electrochemical systems are used in space power systems: nickel-cadmium (NiCd), silver-cadmium (AgCd), silver-zinc (AgZn), and nickel-hydrogen (NiH), and lithium-ion (Li-ion).

The predominant factor in the use of nickel-cadmium batteries is the exceptional long-cycle life capability (**Fig. 3**). An excess of 8 years (over 40,000 cycles) has been obtained on nickel-cadmium batteries operating in a near-Earth environment. At synchronous altitudes, several satellites have operated in excess of 10 years with nickel-cadmium batteries.

Sealed silver-cadmium batteries have seen limited applications in space power systems, even though the energy density is higher than that of a nickel-cadmium battery. This limited use is due primarily to its very low cycle life (less than one-tenth that of the nickel-cadmium battery). Missions that use the silver-cadmium battery usually require a spacecraft with ultraclean magnetic properties. With an appropriately designed and wired silver-cadmium battery, the magnetic field is typically less than 1 gamma (1 nanotesla) at 1 ft (30 cm). Lifetimes in excess of 4 years have been obtained on missions with low battery duty cycles and a "cold storage" of the battery during nonduty cycles.

Because silver-zinc batteries offer excellent energy density but have a very limited life, particularly when used in a cycling mode, they have been used only in space power systems that require only a few days to a few months of operation.

Nickel-hydrogen batteries have replaced nickel-cadmium batteries on some space missions. They have a higher energy density and energy per volume than nickel-cadmium. They have been used on the *Satcom V* communication satellite, the *Earth Observation Satellite,* and the *Hubble Space Telescope*. In fact, the original batteries for the *Hubble Space Telescope* were nickel-cadmium, but they were changed to nickel-hydrogen because the telescope is colder than was expected, leading to more power being used by the heaters.

Lithium-ion batteries were formerly considered too dangerous for space missions since lithium can catch fire or explode. However, this changed with the launch of *Syracuse IIIA* from French Guiana in October 2005, the first spacecraft to carry a lithium-ion battery. The advantage of lithium-ion batteries is that they can store the same amount of energy as a nickel-hydrogen battery in a smaller volume with 50% less mass. Also, after a month of storage, they still have 98% of their charge. This means that no recharge is needed on the launch pad.

There are several batteries under development to replace nickel-cadmium (since cadmium is toxic) and nickel-hydrogen (since such batteries are expensive to produce). Nickel-metal hydride (NiMH), silver-metal hydride (AgMH), and sodium-sulfur (NaS) batteries are the prime contenders. Nickel-metal hydride batteries are similar to nickel-hydrogen except that the hydrogen is stored in the metal hydride solid rather than as a gas. The pressure of nickel-metal hydride is about 50 lb/in.^{2} (350 kilopascals), compared to 1000 lb/in.^{2} (7 megapascals) for nickel-hydrogen. The nickel-hydrogen batteries

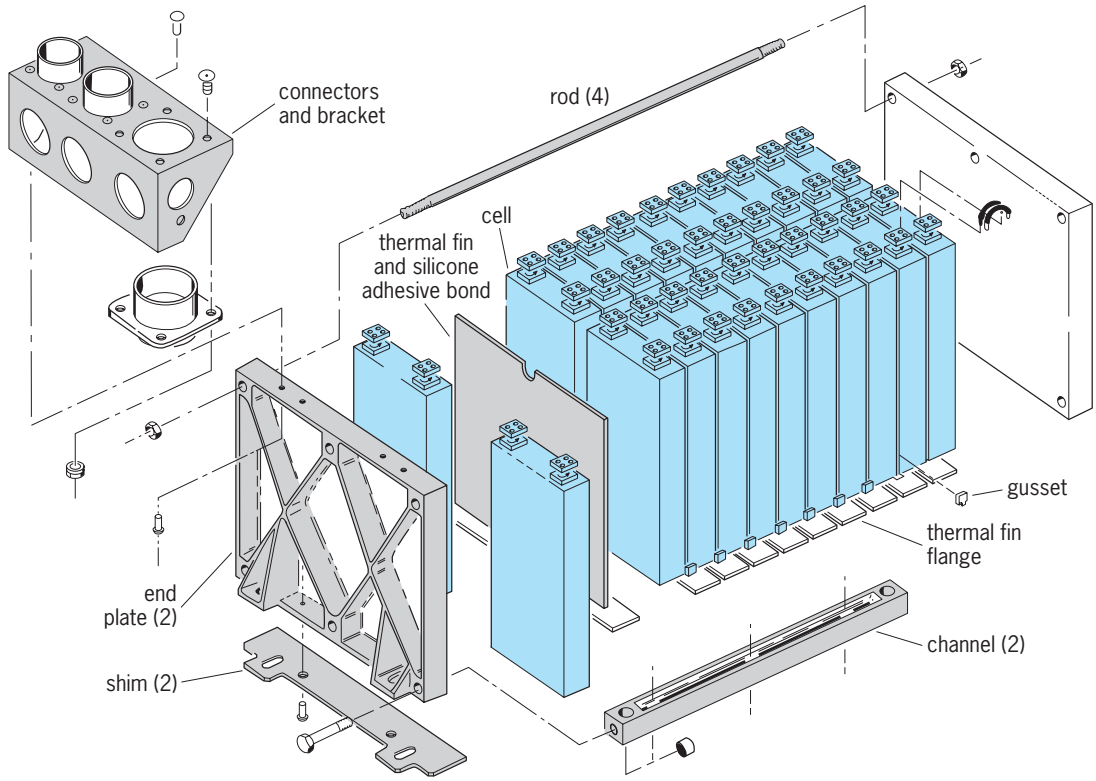

Fig. 3. A 20-Ah battery designed for space flight. Hermetically sealed nickel-cadmium cells are series-connected to provide spacecraft power at approximately 27 V.

are therefore in the form of a round pressure vessel, whereas the nickel-metal hydride batteries can be made rectangular. Nickel-metal hydride batteries have twice the energy density of nickel-cadmium and twice the energy per unit volume. They would also have twice as much energy per unit volume as nickel-hydrogen. Silver-metal hydride batteries might have three times the energy density of nickel-cadmium. *See* ENERGY STORAGE.

Flywheels. Flywheels store mechanical energy using a spinning wheel. When connected to a motor/generator, the solar arrays spin up the motor/flywheel. When the arrays are in shadow, the flywheel spins down and the generator produces electricity. Thus, the flywheel can be used to "store" electricity like a battery. Prior to 2001, NASA worked on development of a flywheel energy storage (FES) system for the *International Space Station*. The solar arrays provide power for the station, and part of that electricity is used to charge the batteries. When the station enters the shadow of the Earth, the batteries supply the power to the station through the Battery Charge/Discharge Unit (BCDU). The flywheel system was intended to replace the batteries and the BCDU. The flywheels were expected to store more energy than the batteries, allowing extra electrical capacity for the experiments. Batteries have a limited amount of charge/discharge cycles, whereas flywheels can be charged and discharged virtually forever. Therefore they last longer and were expected to save hundreds of millions of dollars over the life of the station. Testing at Johnson Space Center was scheduled to start in 2001 with a planned launch to the *International Space Station* in 2004. However, this program was canceled by the Bush Administration in 2001. *See* SPACE STATION.

Fuel cells. The technology resulting from the *Gemini* and *Apollo* fuel-cell programs served as the basis for the space shuttle fuel-cell development. The shuttle fuel cell has three sets of 32 series-connected cells that provide an operating voltage range of 27.5–32.5 V and power of 16 kW (with maximum 21 kW). Each cell is designed for up to 2400 h of operation and weighs 281 lb (127 kg). The shuttle power plant consists of three fuel cells that feature a startup time of 15 min (compared with 24 h for *Apollo*) and instantaneous shutdown (17 h for *Apollo*).

Fuel cells have four distinct advantages for crewed missions: high power, high specific power (56.9 W/lb or 125.8 W/kg), high efficiency (70%), and a benign by-product—water. In fact, the astronauts drink the water produced by the shuttle fuel cells. *See* FUEL CELL.

Tethers. Tethers are a space power system currently in development. A conducting tether many kilometers long is deployed from a spacecraft in orbit about a planet with a magnetic field, such as Earth. The solar wind issues electrons, which travel in spiral paths around the magnetic field lines. As the conducting tether crosses the field lines, it induces an electric field, the same principle that terrestrial generators use. Using an electron gun on the spacecraft to shoot electrons back into space completes the circuit and current flows. The electromotive

force (emf) for low-Earth-orbit (LEO) satellites is a few tenths of a volt per meter. A 6-mi (10-km) tether would produce several kilovolts, leading to several kilowatts of power. The TSS-1 (Tethered Satellite System 1) shuttle payload tried to test this idea but had problems reeling out the tether. Tether voltages as high as 3500 V and current levels of 480 mA were obtained in a subsequent flight, generating a modest amount of power, before the tether severed at 12.2 mi (19.6 km). This was better than was predicted by the analytical models. *See* ELECTROMAGNETIC INDUCTION.

Solar dynamic systems. Solar dynamic systems have been tested to achieve higher efficiencies and power while using less mass and volume than a solar array would use. The principle is that a parabolic or nearly parabolic reflector would concentrate solar energy on a receiver. The receiver has a thermal storage area (for when the station enters Earth's shadow) and a fluid. The solar energy heats the fluid, which turns a turbine-generator to supply electricity. Efficiencies of roughly 30–40% are possible.

A joint NASA and Russia program to put a solar dynamic system on the space station *Mir* was canceled in February 1996 after several successful tests. The flight hardware built for the *Mir* program is currently in storage at the NASA Glenn Research Center. Plans by NASA to use most of the existing hardware and redesign the system to be installed on the *International Space Station* were canceled.

Nuclear power systems. Nuclear power system concepts encompass radioisotopes and nuclear reactor heat sources coupled with thermal-to-electric converters. Nuclear power is now limited to a few experimental applications. Significant advances in materials must be made before a prototype converter can be demonstrated. The relatively heavy weight (2.5–3 W/lb or 5.5–6.6 W/kg), high cost, and lack of demonstrated reliability make nuclear power systems unattractive for Earth-orbiting space power systems below the kilowatt range.

Estimates indicate that the nuclear system could become competitive with solar power at 25 kW or greater. However, with the reentry of a nuclear-reactor-powered Soviet satellite (*Cosmos 954*) in Canada during the summer of 1979, the use of nuclear reactors in space moved from the technical domain to the political arena. The need and desirability of nuclear power in space is a controversial issue with worldwide ramifications.

RTGs have been used for space-borne power primarily for special or unique applications. Present RTGs are descended from power sources developed by the Space Nuclear Auxiliary Power (SNAP) program starting in the 1950s. Although RTGs have been used occasionally in Earth-orbiting satellites, their principal space use has been to power space probes and lunar equipment. RTGs remain on the lunar surface from *Apollo 12, 14, 16*, and *17* missions. These units used lead telluride thermoelectric material and produced approximately 75 W of power. The *Galileo* probe to Jupiter, the *Ulysses* probe of the Sun, and the *Cassini* probe to Saturn use the most advanced RTG so far, the general-purpose heat source (GPHS) RTG. These use silicon-germanium thermoelectric material and produce 300 W of power. *See* NUCLEAR BATTERY; THERMOELECTRICITY.

For power levels above 1 kW, a dynamic isotope power system (DIPS) is desirable. The radioisotope is used to power a Rankine, Brayton, or Stirling device. This system has roughly three times the efficiency of a thermoelectric generator. NASA is currently (2006) testing Stirling generators. *See* BRAYTON CYCLE; RANKINE CYCLE; STIRLING ENGINE.

Jeffrey C. Mitchell

Bibliography. J. A. Angelo, Jr., and D. Buden, *Space Nuclear Power*, Orbit Book Company, Malabar, FL, 1985; S. W. Angrist, *Direct Energy Conversion*, 3d ed., Allyn and Bacon, Boston, 1976; M. D. Griffin and J. R. French, *Space Vehicle Design*, 2d ed., American Institute of Aeronautics and Astronautics, Washington, DC, 2004.

Space probe

An automated crewless vehicle, the payload of a rocket-launching system, designed for flight missions to other planets, to the Moon, and into interplanetary space, as distinguished from Earth-orbiting satellites (see **table**).

The space probe is used primarily for scientific purposes, which are stated as the mission objectives. Missions have been designed to explore many diverse targets, such as the rocky bodies of the inner solar system, including Mercury, Venus, Mars, and the asteroids; the giant gaseous planets in the outer solar system, Jupiter, Saturn, Uranus, and Neptune; and icy bodies originating in the outer solar system, such as comets, Kuiper Belt objects, and the dwarf planet Pluto. Other missions are designed to study solar physics, the properties of interplanetary space, or targets far beyond the solar system. Most spacecraft launched to a planet or other body also study the environment of charged particles and electromagnetic fields in interplanetary space during their cruise phase en route to the destination.

Missions may also be categorized by complexity. The simplest are flyby spacecraft, which study their target body during a relatively brief encounter period from a distance of hundreds to thousands of miles as they continue past. Next are orbiters, which circle a planet or other body for extended study; some may carry atmospheric descent probes. Even more complex are lander missions, which touch down on a planet or other body for the collection of on-site data; some may bear exploration rovers designed to range beyond the immediate landing site. Finally, the most complex space probes envisaged are sample-return missions, which would collect specimen material from a target body and return it to Earth for detailed study.

Spacecraft Subsystems

In the broadest terms, a space probe may be considered a vehicle that transports a payload of sensing

Important space probes

Name	Launch date	Comments
Luna 1	Jan. 2, 1959	Lunar flyby; passed within 3278 mi (5257 km) of the Moon
Pioneer 4	Mar. 3, 1959	Cosmic rays; passed 37,300 mi (60,000 km) from Moon
Luna 2	Sept. 2, 1959	Impacted Moon
Luna 3	Oct. 4, 1959	Photographed far side of Moon
Pioneer 5	Mar. 11, 1960	Study of space between Earth and Venus; magnetic fields and cosmic rays
Mariner 2	Aug. 26, 1962	First planetary flyby; Venus probe
Ranger 7	July 28, 1964	Lunar impact and approach photography
Mariner 4	Nov. 28, 1964	Mars flyby; photography, magnetic fields, cosmic rays
Ranger 8	Feb. 17, 1965	Lunar impact and approach photographs
Ranger 9	Mar. 21, 1965	Lunar impact in Alphonsus; approach photography
Zond 3	July 18, 1965	Flyby of Moon; photographs, other data
Pioneer 6	Dec.16, 1965	Solar orbiter
Luna 9	Jan. 31, 1966	First soft landing on Moon
Luna 10	Mar. 31, 1966	First probe to orbit Moon; collected data, but no camera instrument
Surveyor 1	May 30, 1966	Soft landing on Moon; environmental data and photography
Lunar Orbiter 1	Aug. 10, 1966	Lunar photographs
Pioneer 7	Aug. 17, 1966	Solar orbiter
Luna 11	Aug. 24, 1966	Lunar orbiter; TV camera failed
Luna 12	Oct. 22, 1966	Lunar orbiter; successful return of photographs and other data
Lunar Orbiter 2	Nov. 6, 1966	Lunar orbital photography
Luna 13	Dec, 21, 1966	Lunar lander; surface photography and soil information
Lunar Orbiter 3	Feb. 5, 1967	Lunar orbital photography
Venera 4	June 12, 1967	Analysis of Venus atmosphere; lander failed during descent
Mariner 5	June 14, 1967	Venus flyby; atmospheric and magnetospheric data
Surveyor 3	Apr. 17, 1967	Lunar surface photography and surface properties
Lunar Orbiter 4	May 4, 1967	Lunar orbital photography
Lunar Orbiter 5	Aug. 1, 1967	Lunar orbital photography
Surveyor 5	Sept. 8, 1967	Lunar surface photography and surface properties, including elemental analysis of surface
Surveyor 6	Nov. 7, 1967	Same as *Surveyor 5*; landing in Sinus Medii
Pioneer 8	Dec. 13, 1967	Solar orbiter
Surveyor 7	Jan. 7, 1968	Same as *Surveyor 5*
Luna 14	Apr. 7, 1968	Lunar orbiter
Pioneer 9	Nov. 8, 1968	Solar orbiter
Zond 6	Nov. 10, 1968	Lunar flyby; crashed on return to Earth
Venera 5	Jan. 5, 1969	Same as *Venera 4*
Venera 6	Jan. 10, 1969	Same as *Venera 4*
Mariner 6	Feb. 25, 1969	Photography and analysis of surface and atmosphere of Mars
Mariner 7	Mar. 27, 1969	Same as *Mariner 6*
Luna 15	July 14, 1969	Lunar orbiter (crashed during attempted lunar landing)
Zond 7	Aug. 8, 1969	Reentered Aug. 14, 1969; third crewless circumlunar flight; recovered in the Soviet Union
Venera 7	Aug. 17, 1970	Lander capsule transmitted 23 min from surface of Venus, Dec. 15, 1970
Luna 16	Sept. 12, 1970	Reentered Sept. 24, 1970; crewless Moon lander touched down on Sea of Fertility Sept. 20, 1970; returned lunar soil samples
Zond 8	Oct. 20, 1970	Circled Moon; recovered Oct. 27, 1970
Luna 17	Nov. 10, 1970	Landed on Moon Nov. 17, 1970; crewless Moon rover
Mars 2	May 19, 1971	First Mars landing
Mars 3	May 28, 1971	Mars probe
Mariner 9	May 30, 1971	Mars probe
Luna 18	Sept. 2, 1971	Lunar lander; failed during descent
Luna 19	Sept. 28, 1971	Lunar photography mission
Luna 20	Feb. 14, 1972	Lunar sample return
Pioneer 10	Mar. 2, 1972	Jupiter encounter; transjovian interplanetary probe
Venera 8	Mar. 27, 1972	Venus landing July 22, 1972
Luna 21	Jan. 8, 1972	Lunar lander with rover
Pioneer 11	Apr. 5, 1973	Jupiter encounter and trasjovian interplanetary probe; also Saturn encounter
Mars 4	July 21, 1973	Mars orbiter
Mars 5	July 25, 1973	Mars orbiter
Mars 6	Aug. 5, 1973	Mars lander
Mars 7	Aug. 9, 1973	Mars lander
Mariner 10	Nov. 3, 1973	Venus and Mercury encounter
Luna 22	May 29, 1974	Lunar orbiter
Helios 1	Dec. 10, 1974	Inner solar system, solar wind exploration
Venera 9	June 8, 1975	Venus probe
Venera 10	June 14, 1975	Venus probe
Viking 1	Aug. 20, 1975	Mars lander and orbiter
Viking 2	Sept. 9, 1975	Mars lander and orbiter
Helios 2	Jan. 15, 1976	Interplanetary; similar objectives to those of *Helios 1*
Luna 24	Aug. 9, 1976	Lunar sample return
Voyager 2	Aug. 20, 1977	Jupiter, Saturn, Uranus, and Neptune encounters; also satellites and ring systems
Voyager 1	Sept. 5, 1977	Same objectives as *Voyager 2* with some orbital differences giving differing encounter trajectories
Pioneer Venus Orbiter	May 20, 1978	Returning atmospheric, surface, and particle and field information
Pioneer Venus Multi-Probe Bus	Aug. 8, 1978	Penetration of Venus atmosphere by four probes; returned atmospheric data
Venera 11	Sept. 8, 1978	Venus lander; returned information on surface properties; detection of lightning and thunderlike sounds
Venera 12	Sept. 14, 1978	Similar mission to *Venera 11*

(*conti.*)

Important space probes (*cont.*)

Name	Launch date	Comments
Venera 13	Oct. 30, 1981	Venus lander
Venera 14	Nov. 4, 1981	Venus lander
Venera 15	June 2, 1983	Venus lander; surface topography
Venera 16	June 7, 1983	Similar mission to *Venera 15*
International Cometary Explorer (*ICE*)	—	Originally *International Sun-Earth Explorer 3* (*ISEE 3*) Earth satellite, redirected using a lunar swingby on Dec. 22, 1983, to encounter with Comet Giacobini-Zinner; plasma and magnetic field measurements
Vega 1	Dec. 15, 1984	Venus probe–Halley intercept
Vega 2	Dec. 21, 1984	Venus probe–Halley intercept
Sakigake	Jan. 8, 1985	Halley intercept; precursor to *Suisei*, upgraded to full mission
Giotto	July 2, 1985	Halley intercept
Suisei	Aug. 19, 1985	Halley intercept; plasma and magnetic field measurements
Phobos 1	July 7, 1988	Mars/Phobos probe, lost by command error
Phobos 2	July 12, 1988	Mars/Phobos probe; some data, but communications lost
Magellan	May 4, 1989	Venus radar mapper
Galileo	Oct. 18, 1989	Jupiter orbiter and atmospheric probe
Hiten/Hagoromo	Jan. 24, 1990	Moon orbiter and relay probe; orbiter transmitter malfunctioned
Ulysses	Oct. 6, 1990	Solar polar orbiter
Mars Observer	Sept. 25, 1992	Contact lost 3 days before Mars arrival
Clementine	Jan. 25, 1994	Orbited Moon; thruster malfunction prevented asteroid flyby
Solar and Heliospheric Observatory (*SOHO*)	Dec. 2, 1995	Orbits L1 libration point to study the Sun
Near Earth Asteroid Rendezvous	Feb. 17, 1996	Asteroid orbiter
Mars Global Surveyor	Nov. 7, 1996	Mars orbiter
Mars 96	Nov. 16, 1996	Mars orbiter and landers; launch vehicle failed
Mars Pathfinder	Dec. 4, 1996	Mars lander and rover
Advanced Composition Explorer	Aug. 25, 1997	Orbits L1 libration point to study charged particles
Cassini	Oct. 15, 1997	Saturn orbiter/Titan descent probe
Lunar Prospector	Jan. 6, 1998	Lunar orbiter
Nozomi	July 4, 1998	Mars orbiter; failed to reach Mars orbit
Deep Space 1	Oct. 24, 1998	Test of ion engine and 11 other advanced technologies; asteroid and comet flybys
Mars Climate Orbiter	Dec. 11, 1998	Lost during Mars arrival
Mars Polar Lander	Jan. 3, 1999	Lost during Mars arrival
Stardust	Feb. 7, 1999	Comet flyby, dust sample return
Mars Odyssey	April 7, 2001	Mars orbiter
Wilkinson Micowave Anisotropy Probe	June 30, 2001	Solar orbiter studying cosmic background radiation
Genesis	Aug. 8, 2001	Solar wind sample return; hard impact upon landing in Utah
Comet Nucleus Tour	July 3, 2002	Intended for two comet flybys; failed shortly after launch
Hayabusa	May 9, 2003	Comet sample return mission; sample collection uncertain
Mars Express	June 2, 2003	Mars orbiter
Beagle 2	June 2, 2003	Lander carried on Mars Express; failed during descent
Spirit (*Mars Exploration Rover A*)	June 10, 2003	Mars rover
Opportunity (*Mars Exploration Rover B*)	July 7, 2003	Mars rover
Spitzer Space Telescope	Aug. 25, 2003	Infrared telescope in Earth-trailing heliocentric orbit
Smart 1	Sept. 27, 2003	Lunar orbiter
Rosetta	March 2, 2004	Comet orbiter-lander
Messenger	Aug. 3, 2004	Mercury orbiter
Deep Impact	Jan. 12, 2005	Released impactor to excavate crater in comet nucleus
Mars Reconnaissance Orbiter	Aug. 12, 2005	Mars orbiter
Venus Express	Nov. 9, 2005	Venus orbiter
New Horizons	Jan. 19, 2006	Pluto flyby
Solar Terrestrial Relations Observatory	Oct. 26, 2006	Twin spacecraft in solar orbit

instruments to the vicinity of a target body. Thus, the spacecraft must include a number of subsystems to provide power, to communicate with Earth, to maintain and modify attitude and perform maneuvers, to maintain acceptable on-board temperature, and to manage the spacecraft overall. *See* SPACE TECHNOLOGY.

Scientific instruments. The scientific payload may be divided into remote-sensing instruments such as cameras, and direct-sensing instruments such as magnetometers or dust detectors. They may be classified as passive instruments, which detect radiance given off by a target body, or active, which emit energy such as radar pulses to characterize a target body.

Examples of passive remote-sensing instruments are cameras and other imagers, polarimeters, photometers, spectrometers, and radiometers. Passive direct-sensing instruments include those that study magnetic fields and charged particles, such as high- and low-energy particle detectors, plasma instruments, plasma wave detectors, dust detectors, and magnetometers, as well as instruments on landers

such as those of the *Viking* Mars mission. Active remote-sensing instruments include synthetic-aperture radar and altimeters.

Early missions up to and including *Voyager* used vidicon tubes similar to those in terrestrial television cameras of that period as the detector of the on-board camera. Some, including Lunar Orbiters and Luna missions, actually used film which was developed and scanned on board. The image information was then digitized before being transmitted to Earth. Subsequent missions have used solid-state charge-coupled devices (CCDs), similar to those in contemporary television cameras. A spacecraft's camera system frequently is outfitted with two sets of optics to provide wide- and narrow-field views; the narrow-field camera is the equivalent of a small on-board telescope. In addition, the camera system includes a filter wheel with various colored filters which may be placed in the optical path. Color images are made by rapidly taking a series of three exposures through different filters; these are transmitted to Earth and combined into a single frame during image processing. Most spacecraft cameras, such as those on *Voyager*, *Galileo*, or *Cassini*, capture light in discrete snapshot-type exposures. By contrast, some orbiters, such as *Mars Global Surveyor* and *Mars Reconnaissance Orbiter*, carry pushbroom-type cameras; these collect data continuously, building up a long ribbonlike image as the spacecraft passes over the planet's surface. *See* ASTRONOMICAL IMAGING; CHARGE-COUPLED DEVICES; COLOR FILTER; REMOTE SENSING; TELEVISION CAMERA TUBE.

Power subsystem. Electrical power is required for all spacecraft functions. The total required power ranges from about 300 to 2500 W for current missions, depending on the complexity of the spacecraft. The power subsystem must generate, store, and distribute electrical power.

All space probes launched so far have generated power either via solar panels or via radioisotope power systems. Solar panels are generally employed on missions to the solar system's inner planets, where sunlight remains relatively strong. The strength of sunlight diminishes as the square of the distance from the Sun. Missions beyond the orbit of Mars have therefore relied on radioisotope systems. In addition, radioisotope systems have been used on the *Viking* Mars landers because of the unreliability of sunlight during dust storms and other periods.

Solar panels typically employ photovoltaic devices constructed of materials such as crystalline silicon or gallium arsenide. A silicon cell 2.4 in. (6 cm) in diameter, for example, produces a current of about 1 A at 0.25 V under direct sunlight at a distance of 1 astronomical unit (1 AU, the average distance from Earth to the Sun, is about 9.3×10^7 mi or 1.5×10^8 km). Solar panels are mounted on arrays which may be articulated to optimize pointing at the Sun; alternatively, they may be off-pointed during periods to generate less electrical energy in the interest of on-board power management. The performance of solar panels degrades gradually at a rate of perhaps 1–2% per year. On the solar-powered *Magellan* mission, power loss was unusually great. After a 15-month interplanetary cruise and 4 years in orbit at Venus, careful power management was required in the mission's final days before *Magellan* was intentionally directed into the planet's atmosphere to burn up in October 1994. *See* ASTRONOMICAL UNIT; SOLAR CELL.

The other principal power source, the radioisotope power system, produces electrical power through the radioactive decay of plutonium-238. The radioisotope device is not a nuclear reactor, however; as the plutonium naturally decays, it gives off heat which is converted to electricity. In early devices this conversion was accomplished by thermocouples made of silicon-germanium junctions, whereas one newer design uses thermocouples composed primarily of lead telluride, and another envisages the use of a dynamic Stirling converter. Waste heat not converted to electricity is radiated into space via metal fins on the generator casing. The output of radioisotope power systems gradually decays at about 1–2% per year, slightly more rapidly than actively used solar panels. *See* NUCLEAR BATTERY.

Because they typically go through alternating sunlight and shadow, spacecraft powered by solar panels must carry batteries to store electricity for periods when the panels are in shade. Nickel hydrogen, nickel metal hydride, and lithium batteries are often used. Their performance degrades after many cycles of charging and partial discharging, but the batteries may be reconditioned by discharging them completely, much like rechargeable batteries in terrestrial applications. Because the electrical power output of radioisotope power systems does not vary significantly from hour to hour, spacecraft equipped with them do not require batteries, but may require short-term energy-storage devices, such as capacitors, to provide for high-pulse current needs.

The other function of the power subsystem is to distribute electrical power. A typical spacecraft supplies its various on-board subsystems with a specific constant bus voltage, although some instruments may require special voltages. During the course of the mission, various instruments are switched on and off as they are used to make observations. This switching is performed with solid-state or mechanical relays. A shunt regulator is generally used to maintain a constant voltage on the main bus; this converts electrical energy into heat, which is radiated into space. In some cases it may be necessary to switch instruments on and off in a particular order to manage the voltage on the main bus. *See* RELAY; SPACE POWER SYSTEMS.

Telecommunications subsystem. In order to accomplish its mission, the spacecraft must maintain communications with Earth. This includes receiving commands sent from ground controllers, as well as transmitting scientific data and routine engineering "housekeeping" data. All of these transmissions are made in various segments of the microwave spectrum. Early missions typically used lower frequencies in the L or S band (1–4 GHz), while subsequent missions have used the X band (7–9 GHz). The

design of the telecommunications subsystem takes into account the volume of data to be transmitted and the distance from Earth at which the spacecraft will operate, dictating such considerations as the size of antennas and the power of the on-board transmitter. *See* MICROWAVE; TELEMETERING.

Advanced planetary probes have carried a dish-shaped high-gain antenna which is the chief antenna used to both transmit and receive. These antennas typically consist of a large parabolic reflector, with a subreflector mounted at the main reflector's focus in a Cassegrain-type configuration. Most large high-gain antennas are fixed to the spacecraft's body; the entire spacecraft must maintain Earth pointing to avoid disrupting communication. Smaller parabolic antennas used in relatively close missions such as those to Mars, however, may be on a steerable mount. Most large parabolic antennas are rigid; the high-gain antenna for *Cassini*, for example, is fabricated from a lightweight composite material. *Galileo* used a 16-ft (4.8-m) high-gain antenna designed to unfold like an umbrella after launch; although similar antennas functioned well on Earth-orbiting satellites, the antenna on *Galileo* did not open fully when ground controllers attempted to deploy it in April 1991, rendering it unusable. The beam width of parabolic high-gain antennas is typically very narrow, requiring an Earth pointing to within a fraction of a degree. The advantage of the high-gain antenna is that it allows a much greater transmission rate than do other antenna types. *See* ANTENNA (ELECTROMAGNETISM).

In the interest of redundancy and in the event that Earth pointing is lost, spacecraft virtually always carry other on-board antennas. These may be low-gain antennas, which typically offer nearly omnidirectional coverage except for blind spots shadowed by the spacecraft body. *Galileo* was outfitted with two low-gain antennas which combined to afford completely omnidirectional capability. Other spacecraft have used medium-gain antennas, which provide a beam width of perhaps 20–30°, with a gain midway between that of high- and low-gain antennas. *Magellan*, for example, was equipped with a conical feed horn antenna which it used when its high-gain antenna was off Earth point.

The spacecraft's on-board transmitter typically provides a signal in the tens of watts. The final stage of the transmitter is a power amplifier which may be a traveling-wave tube or a solid-state amplifier. The carrier signal is modulated with a subcarrier which contains the measured scientific or engineering data. Ground controllers frequently make use of the technique of Doppler tracking, wherein slight changes in the frequency of the probe's signal are measured to reveal subtle acceleration changes by the spacecraft. In some cases, spacecraft may be commanded to transmit a pure carrier signal for purposes such as Doppler tracking; this was used with great success in the *Magellan* mission to map Venus's gravity field by measuring slight acceleration changes in the spacecraft as it passed over various portions of the planet. *See* DOPPLER EFFECT; MICROWAVE SOLID-STATE DEVICES; POWER AMPLIFIER; RADIO TRANSMITTER; TRAVELING-WAVE TUBE.

The spacecraft's receiver is designed to be sensitive to a narrow range of frequencies within a very small percentage of a specified frequency. Many spacecraft receivers incorporate phase-lock-loop circuitry, which locks on to a signal detected within the specified bandwidth and follows any slight changes in frequency, particularly the Doppler shift produced by the motion of the probe relative to Earth. Although this design has worked well for most missions, *Voyager 2* presented special challenges because one of its redundant receivers failed midway through the mission, and the second receiver's frequency varied depending on the temperature of the on-board electronics. Ground controllers thus had to calculate the expected frequency to which the on-board receiver would be sensitive and had to tailor transmissions of spacecraft commands accordingly. *See* PHASE-LOCKED LOOPS; RADIO RECEIVER.

In most modern cases, the transmitter and receiver may be combined into a single unit called a transponder. Some components of the telecommunications subsystem may be used for purposes other than communications; there may in fact be a radio-science team of investigators who use the on-board radio as a scientific instrument. As noted above, *Magellan* measured Venus's gravity by means of Doppler tracking of the radio carrier signal. On many missions such as *Voyager*, the radio signal is tracked as the spacecraft passes behind a planet, revealing information on the chemical makeup of the planet's atmosphere. *Magellan*'s main scientific instrument, its imaging radar, used the spacecraft's high-gain antenna to send and receive radar pulses to and from the planet's surface. *See* RADAR; RADAR ASTRONOMY; SPACE COMMUNICATIONS.

Attitude-control subsystem. It would be impossible to navigate the spacecraft successfully or point its scientific instruments or antennas without closely controlling its orientation in space, or attitude. Many specialized techniques have been developed to accomplish this.

Some spacecraft, particularly earlier ones, have been spin-stabilized; during or shortly after launch, the spacecraft is set spinning at a rate on the order of a few revolutions per minute. Much like a rotating toy top, the spacecraft's orientation is stabilized by the gyroscopic action of its spinning mass. Thruster jets may be fired to change the spacecraft's attitude. Spin stabilization is preferred for scientific instruments that measure electromagnetic fields and charged particles, because they are continuously swept in a circle in such a way as to cancel out interference from the spacecraft itself. The stabilization method is not as good for scientific instruments that require pointing, such as optical cameras. The *Pioneer 10* and *11* spacecraft were spin-stabilized, capturing photolike images with sweep polarimeters. *Ulysses* exclusively carries fields and particles instruments, and thus is also spin-stabilized.

Most planetary spacecraft, however, are three-axis-stabilized, meaning that their attitude is fixed in

relation to space. This method makes the operation of pointed instruments much simpler, although it is not as preferable for fields and particles instruments. The spacecraft's attitude is maintained and changed via on-board thruster jets or reaction wheels, or a combination of both. Reaction wheels are electrically powered devices similar to gyroscopes which rotate the spacecraft about a given axis by spinning up or slowing down a massive wheel. Although most planetary spacecraft are three-axis-stabilized, *Galileo* was the only spacecraft to use a complex dual-spin design. Most of the *Galileo* spacecraft spins at three revolutions per minute, while a despun segment carrying the camera and other pointed instruments is held in a three-axis-stabilized attitude. The two segments of the spacecraft are joined by a complex spin bearing that must transfer hundreds of electrical signals. *See* GYROSCOPE.

Other parts of the attitude-control subsystem include Sun sensors and star trackers which are used to establish points of reference in space. The subsystem's functions are usually overseen by a dedicated computer system distinct from the spacecraft's main command and data computer. Because spacecraft must operate at great distances from Earth where radioed commands would take many minutes or even hours to arrive, the attitude-control subsystem is usually one of the most autonomous parts of the spacecraft, maintaining the spacecraft's orientation with only infrequent intervention from the ground. *See* STAR TRACKER.

Propulsion subsystem. Most spacecraft are outfitted with a series of thruster jets, each of which produces approximately 0.2–2 pounds-force (1–10 newtons) of thrust. The thrusters are placed so that at least two lie in opposite directions on each of the spacecraft's three orthogonal axes. Additional thrusters are often added to lend redundancy. Thrusters are usually fueled with a monopropellant, hydrazine, which decomposes explosively when it contacts an electrically heated metallic catalyst within the thruster. The system also includes helium used to pressurize the propellant.

In addition to maintaining the spacecraft's attitude, on-board thrusters are used for trajectory-correction maneuvers. These maneuvers are usually carried out at several points during interplanetary cruise to progressively fine-tune the spacecraft's aim point at its destination. Maneuvers are conducted by firing one or more thrusters for periods from several seconds to several minutes, accelerating or decelerating the spacecraft typically on the order of several meters per second.

Spacecraft designed to orbit a planet or similar target body must carry a larger propulsion element capable of decelerating the spacecraft into orbit upon arrival. *Magellan* used a solid-fuel motor to insert it into orbit at Venus, whereas spacecraft such as *Galileo* and *Cassini* use bipropellant hypergolic substances, monomethylhydrazine and nitrogen tetroxide, which expand rapidly in a flameless combustion when mixed. The main engine on *Galileo* and *Cassini* used to place the spacecraft in planetary orbit provided 90 and 110 lbf (400 and 490 N) of thrust, respectively. Mars orbiters have taken advantage of different propulsion approaches. *Mars Global Surveyor* and *Mars Odyssey* each used a large main engine [135 and 154 lbf (600 and 685 N), respectively] to enter orbit at Mars. *Mars Reconnaissance Orbiter*, by contrast, braked into orbit by firing a group of six thrusters supplying 38 lbf (170 N) of thurst each. *See* SPACECRAFT PROPULSION.

Thermal control subsystem. The environment of space is very harsh, ranging from temperatures near absolute zero in the outer solar system to the extremely hot locales of inner planets such as Venus and Mercury. In order to minimize the impact of temperature variations on the electronics onboard, spacecraft nearly always incorporate some form of thermal control. Mechanical louvers, controlled by bimetallic strips similar to those in terrestrial thermostats, are often used to selectively radiate heat from the interior of the spacecraft into space. Other thermal strategies include painting exterior surfaces; *Magellan* used white thermal blankets to reflect heat and keep interior units cooler. *Magellan* also carried optical solar reflectors composed of quartz mirror tiles on a number of surfaces, including the back sides of its solar arrays. Internal components are frequently painted black to encourage heat exchange between these components.

In some cases, spacecraft may also carry one or more active forms of heating to maintain temperature at required minimums. Often this is an electric unit that radiates heat from a resistive element similar to electric heaters on Earth; it may be controlled autonomously by on-board logic or commanded from the ground. Some spacecraft carry several radioisotope heating units (RHUs), each containing a small amount of plutonium which heats nearby spacecraft components. Because of their power needs, active cooling systems are not generally feasible on planetary spacecraft.

Command and data subsystem. This designation is given to the main computer that oversees management of spacecraft functions and handling of collected data. Blocks of commands transmitted from Earth are stored in memory in the command and data subsystem and are executed at prescribed times. This subsystem also contains the spacecraft clock in order to accurately pace its activities, as well as all the activities of the spacecraft.

In its data-handling function, the subsystem receives data destined for the ground on the spacecraft's data bus from other on-board subsystems, processes them, formats them for telemetry, and delivers them to the telecommunications subsystem for transmission to Earth. Included are data from scientific instruments as well as housekeeping engineering data such as voltage and temperature readings. Data are usually stored on tape recorders on board while awaiting transmission to Earth. Since the 1990s, solid-state data recorders have taken the place of traditional tape-driven devices for this purpose. Most spacecraft encode data to be sent of Earth in order to correct transmission errors;

Viterbi, Golay, and Reed-Solomon encoding schemes have been used, and Turbo coding is now the most common. Some spacecraft also perform data compression to reduce the number of bits to be transmitted. This technique was used extensively on *Galileo* to compensate partially for the fact that the spacecraft's high-gain antenna did not deploy completely. *See* DATA COMPRESSION; INFORMATION THEORY.

Another critical function of the command and data subsystem is fault protection for the spacecraft in the event of a mishap. One of the chief strategies designed to respond to a contingency is safing, which shuts down or reconfigures subsystems to prevent damage either from within or from the external environment. After carrying out these activities, the fault-protection software usually commands the spacecraft to begin an automated search to regain Sun reference and reestablish communications. Spacecraft safing events that have taken place in actual flight missions have usually been caused by single-event upsets, probably resulting from a cosmic ray interfering with on-board electronics or from transient electrical phenomena from on-board sources. Spacecraft are usually recovered after safing events, although one of the *Phobos* probes was believed to be lost because of a faulty software command. Another example of on-board fault protection is the command-loss timer, which counts how much time has passed since a command has been received from Earth. If this count passes a designated value, it is assumed that a failure has occurred in the telecommunications or command subsystems. The spacecraft then switches to backup systems and takes other actions in an attempt to reestablish reliable contact.

Structure subsystem. The spacecraft's physical structure is considered a subsystem itself for the purposes of planning and design. Usually the heart of this structure is a spacecraft bus, often consisting of a number of bays, which houses the spacecraft's main subsystems. Besides providing physical support, the bus generally offers an electrical grounding reference, radio-frequency interference shielding, and protection from radiation and meteoroids. The bus also provides attachment points for appendages such as booms and scanning platforms bearing scientific instruments. Magnetometers, for example, are usually mounted on a boom which may extend 30 ft (10 m) or more away from the spacecraft bus. For launch, this boom collapses and is stowed in a compact canister. The spacecraft bus is almost always covered with one or more layers of thermal blankets to help stabilize on-board temperature and to shield the spacecraft from micrometeoroid impacts. Special shielding may be added to protect the spacecraft from radiation in environments such as that at Jupiter, where ionized particles may degrade on-board subsystems, or from temperature extremes at planets such as Venus or Mercury. Spacecraft designed as landers usually also carry aeroshells and parachutes to aid during descent. *See* ATMOSPHERIC ENTRY; SPACECRAFT STRUCTURE.

Miscellaneous subsystems. Among other miscellaneous spacecraft subsystems, the pyrotechnics subsystem includes electrically initiated devices, called squibs, that are used to deploy components such as booms, explode bolts that separate or jettison hardware such as descent probes or launch adapters, ignite solid rocket motors, and operate some types of valves.

Redundancy. As noted above, redundancy traditionally has been a prominent feature of planetary spacecraft, which typically carry two transmitters, two receivers, two command computers, and so on. In the event of a mishap, the spacecraft can switch to a backup system and carry on; this has happened many times. In an effort to reduce spacecraft mass and complexity, some designers have considered dropping redundancy and allowing spacecraft to operate on a single string of components. Although this may save mass, it also usually means a spacecraft cannot function if any of its principal subsystems fails.

Mission Stages

The process of conceiving, designing, building, launching, and operating a spacecraft mission takes many years. *Voyager*, for example, was conceived during the late 1960s and given formal project approval—called a "new start"—in 1972 by the National Aeronautics and Space Administration (NASA). Five years later, the twin *Voyager 1* and *2* were launched, and executed their planetary flybys between 1979 and 1989. In the 1990s and the first decade of the twenty-first century, they continued to return data as they left the solar system. Efforts have been made to simplify missions and to shorten the lead time required to design and launch spacecraft, but prelaunch activities still involve a minimum of 2–3 years even for modest missions.

In the United States, the programmatic steps leading to preparation of a mission are commonly designated by a series of terms. A conceptual study is conducted, which outlines the general scheme of a mission concept. If approved, this proceeds to a phase A study, also called a preliminary analysis, which creates a tentative design and project plan outlining the mission. This specifics what to build, when to launch, what course the spacecraft is to take, what is to be done during cruise, when the spacecraft will arrive at its destination, and what operations will be carried out. The project then moves into phase B, also called the definition phase, in which more formal specifications and requirements are determined. If the project wins formal approval, it progresses to phase C/D, or design and development, where the spacecraft is designed, assembled, tested, and launched. The process concludes with the mission's operations phase, phase E, including its interplanetary cruise and eventual flyby, orbit insertion, or landing. If the spacecraft is healthy at the conclusion of its specified mission period, it may be approved to continue into an extended mission.

Throughout the design process, mission planners must be sensitive to various constraints. These include the project's overall acceptable budget,

which will influence many decisions on spacecraft and mission complexity. Another constraint is the availability of launch vehicles, which in turn will dictate the maximum spacecraft mass and feasible trajectories. The launch period during which planetary geometry is appropriate to carry out the mission must also be taken into account. The launch period for a transfer orbit to Mars, for example, occurs once every 26 months. Contention for available resources must also be considered, such as tracking time on the giant dish antennas of the U.S. Deep Space Network used to communicate with deep-space probes. *See* SPACECRAFT GROUND INSTRUMENTATION.

Launch. Virtually all planetary launches in the United States take place at the Cape Canaveral Air Force Station in Florida, which is situated relatively close to the Equator with an eastward launch range extending over the Atlantic Ocean. Launches conducted by Russia take place at the Baikonur Cosmodrome in Kazahkstan. In the 1980s, the United States and the Soviet Union were joined by Japan and the European Space Agency in launching solar system probes. Japan's launch facility is at the Uchinoura Space Center, near the southwest tip of the island of Kyushu. European launches take place at the Center Spatial Guyanais at Kourou, French Guiana, on the northeast coast of South America. *See* LAUNCH COMPLEX.

United States space probes traditionally have been launched on such vehicles as the Atlas, Delta, and Titan rockets. In 1989 and 1990, the *Magellan*, *Galileo*, and *Ulysses* spacecraft were launched from the payload bay of the United States' space shuttle, using specially adapted solid-fuel upper-stage engines to inject them on their interplanetary trajectories. The Russian launch fleet includes the liquid-fuel Proton. Japan has launched planetary missions on its solid-fueled Mu-5, while Europe has developed the Ariane 5. *See* ROCKET PROPULSION; SPACE SHUTTLE.

After a spacecraft arrives as the launch site, it is generally tested to ensure that all systems are working normally after being transported from the facility where it was assembled. The probe is then mated to its upper-stage engine, and the combined stack is placed atop the launch vehicle. The probe's propulsion tanks are filled with fuel and pressurant. After launch, the main rocket usually places the probe and upper stage in a relatively low orbit around Earth. The upper stage then fires to send the probe on its interplanetary trajectory. At that point, pyrotechnic squibs may be fired to deploy on-board appendages, and the entire spacecraft goes through a checkout phase.

Cruise. Many activities take place while the spacecraft is en route to its destination planet or other target. Some scientific instruments, such as those studying fields and particles, may be fully powered to collect data during the interplanetary transit. Other instruments used only at the target body may be switched off. It is usually necessary to conduct one or more trajectory correction maneuvers in which on-board thrusters are fired to fine-tune the spacecraft's flight path.

Although it is common to think of a spacecraft trajectory as a simple path from Earth to the target body, the fact that Earth and all other bodies are moving around the Sun makes trajectory design more complex. To travel from Earth outward to Mars, for example, using the least energy possible, mission planners usually take advantage of a Hohmann transfer orbit. This involves adjusting the existing orbit around the Sun so that it is more elliptical. Under the new orbit, the spacecraft's perihelion, or closest approach to the Sun, will coincide with the orbit of Earth, while its aphelion, or farthest point from the Sun, will coincide with the orbit of Mars. In order to arrive at a time when Mars is at the same point as the spacecraft, it must be launched during an appropriate time window.

Beginning in the 1970s, mission planners took advantage of a technique called gravity assist trajectories to send spacecraft on to more than a single planetary target. In this technique, the spacecraft passes within a few hundred miles of a given planet; this accelerates (or decelerates) the spacecraft and changes its trajectory. (Because no energy is ever created or lost, the encounter also results in the planet being slowed, or speeded, by an immeasurably small amount in its orbit around the Sun.) Gravity assist was first used by *Mariner 10* in 1973, when the spacecraft used the gravity of Venus to send it on to three flybys of Mercury in 1974 and 1975. The technique was then used very extensively by *Voyager* to send spacecraft past Jupiter, Saturn, Uranus, and Neptune. In the 1990s, gravity assist became useful to send heavier spacecraft to destinations that would not be possible via direct trajectories by using available launch vehicles. *Galileo*, for example, flew by Venus once and Earth twice on its way to Jupiter. *See* SPACE NAVIGATION AND GUIDANCE.

Arrival. Activities at the target body vary, depending on whether the spacecraft executes a flyby, goes into orbit, or lands on the body's surface. Flyby activities are generally divided into a far-encounter period, which may extend for several weeks, and a close encounter, which includes the most hectic days directly surrounding the spacecraft's closest approach to its target body.

Spacecraft designed to go into orbit around the target body must fire a large engine to brake their velocity sufficiently so that the target body's gravity can capture them. In a typical mission scenario, the spacecraft may enter an initial capture orbit which may be highly elliptical; over a period of several weeks, thruster firings will be used to circularize the orbit so that the spacecraft can carry out mapping operations. In addition, it has become routine for Mars orbiters to circularize their orbits after arriving at the planet by using the drag of Mars's atmosphere to reduce the orbit's apoapsis, or high point, over time. This technique is called aerobraking. Most spacecraft are placed into a polar or fairly highly inclined orbit at a planet to maximize the amount of surface that can be studied. *Magellan* was placed in

a near-polar orbit at Venus that caused the planet to rotate once beneath it every 243 days. In the *Mars Reconnaissance Orbiter* mission, the spacecraft was placed in a Sun-synchronous orbit that took it over regions uniformly at about 3 p.m. Mars local time. In this orbit type, the spacecraft's orbit plane precesses with nearly the same period as the planet's solar orbit period. This is useful for many scientific instruments because it makes the incidence angle of sunlight on surface features uniform from orbit to orbit.

Landing a spacecraft on a planet requires even greater deceleration than in the case of orbit insertion. Spacecraft approaching planets with an atmosphere, such as Mars, usually take advantage of the technique of aerobraking, or using friction with the atmosphere, to slow the spacecraft. An aeroshell is usually required to protect the spacecraft from the effects of heating. After jettisoning the aeroshell, the spacecraft may use a propulsion system to achieve a soft landing, perhaps augmented by a parachute system. Although the Soviet Union's *Luna* spacecraft returned soil samples from the Moon during the 1970s, sample return missions have not yet been mounted to other planets, although samples of cometary dust and solar and interstellar particles have been brought to Earth. Japan's *Hayabusa* spacecraft apparently failed to collect a sample when it landed on the asteroid Itokawa in 2005, but probably obtained some dust from the asteroid that will be carried to Earth in 2010.

Fig. 2. Full-scale model of the *Pioneer 3* and *4* space probes, showing the antenna, the paint stripes for thermal control, and the despin device about the major diameter. (*JPL/NASA*)

Significant Missions

Soon after the Soviet Union launched the Earth-orbiting *Sputnik* in 1957 and the United States responded with *Explorer 1* in early 1958, both countries turned to developing probes that would escape Earth orbit and travel to the Moon and beyond.

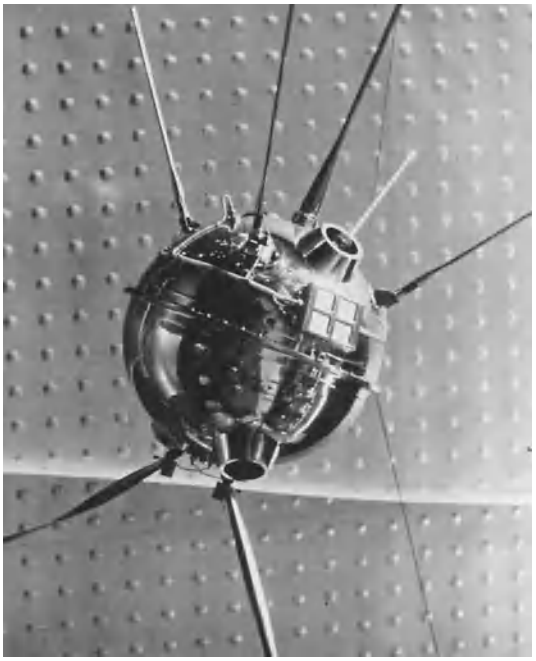

Fig. 1. Model of *Luna 1* (Mechta). Long metal strips are antennas, the tapered post is the magnetometer boom, and raised windowlike apertures are ion detectors.

Luna. The first successful mission to escape Earth orbit was the Soviet Union's *Luna 1* (**Fig. 1**), which executed a 3728-mi (5275-km) lunar flyby after launch on January 2, 1959. Later that year, the Soviet Union launched *Luna 2*, which made the first lunar impact, and *Luna 3*, which studied the far side of the Moon. *Luna 4*, launched in 1963, missed the Moon by 5282 mi (8500 km). The Soviets had a string of failures in 1965 in attempting a lunar soft landing; *Luna 5* failed when the spacecraft made a hard impact on the Moon, *Luna 6* missed the Moon by 100,000 mi (160,000 km), and *Luna 7* and *8* failed because retrorockets fired too early or late. *Luna 9* made the first lunar soft landing on February 2, 1966, and returned photographs from the Moon's surface. *Luna 10*, *11*, *12*, *14*, *15*, and *19* were lunar orbiters, while *Luna 13* was a soft lander. In 1970 the Soviet Union launched *Luna 16*, which carried out that country's first successful lunar sample return. Later the same year, *Luna 17* made a soft landing on the Moon with a robotic rover vehicle. *Luna 18* was intended as a sample-return mission but failed during lunar descent. *Luna 20* in 1972 executed a successful sample return. The series continued with *Luna 21*, a lander and rover, in 1973, and *Luna 22*, an orbiter, in 1974; and concluded with *Luna 23* and *24*, sample return missions conducted in 1974 and 1976 (*Luna 24* was successful, while *Luna 23* was not). Much information about the missions was

withheld by the Soviet government for many years, but some details have become available since the dissolution of the Soviet Union in the early 1990s. *See* MOON.

Pioneer. This name was given to a family of largely dissimilar United States spacecraft used mainly for interplanetary physics, although some later missions in the series achieved success as planetary probes. *Pioneer 1*, *2*, and *3* were failed lunar probes in 1958, although *Pioneer 1* and *3* returned data on the environment of space before reentering Earth's atmosphere. *Pioneer 4* (**Fig. 2**) passed within 37,300 mi (60,000 km) of the Moon in 1959, followed by *Pioneer 5* in 1960, which relayed solar system data to a distance of 22.5×10^6 mi (36.2×10^6 km). Following a hiatus of 5 years, *Pioneer 6* studied the space environment in solar orbit, as did *Pioneer 7* in 1966, *Pioneer 8* in 1967, and *Pioneer 9* in 1968.

The series took a very different direction with *Pioneer 10* and *11*, the first space probes sent into the outer solar system in 1972 and 1973, respectively. *Pioneer 10* executed a flyby of Jupiter in December 1973, while *Pioneer 11* flew by Jupiter in December 1974 and Saturn in 1979. In 1978 the series took yet another departure with *Pioneer Venus* (**Fig. 3**), which used two launches to send an orbiter and

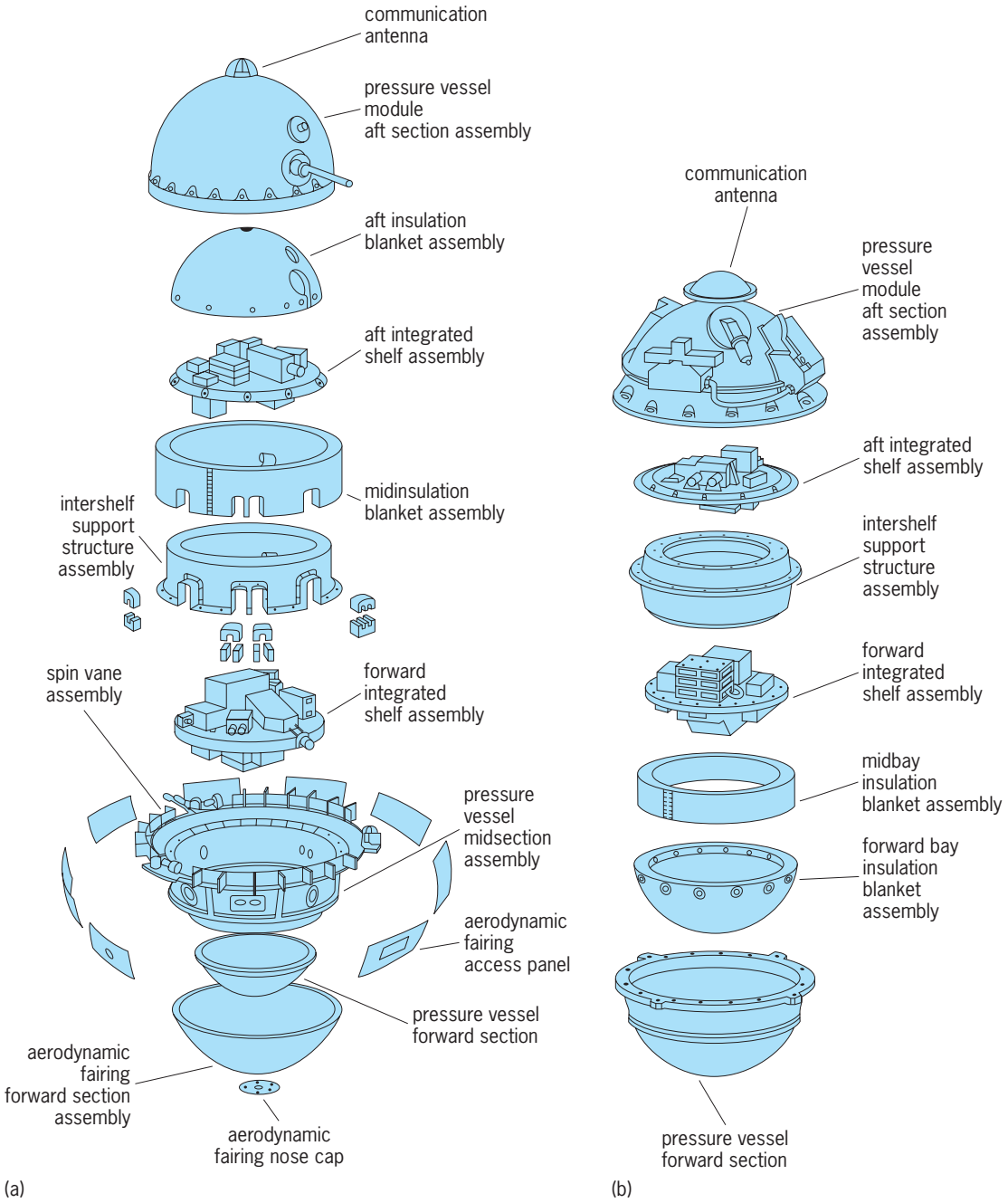

Fig. 3. Expanded views of the structure of (*a*) the large probe and (*b*) one of the smaller probes launched into the Venus atmosphere from the *Pioneer Venus* bus. (*After L. J. Nolte and S. C. Sommer, Probing a Planetary Atmosphere: Pioneer-Venus Spacecraft Description, 1975*)

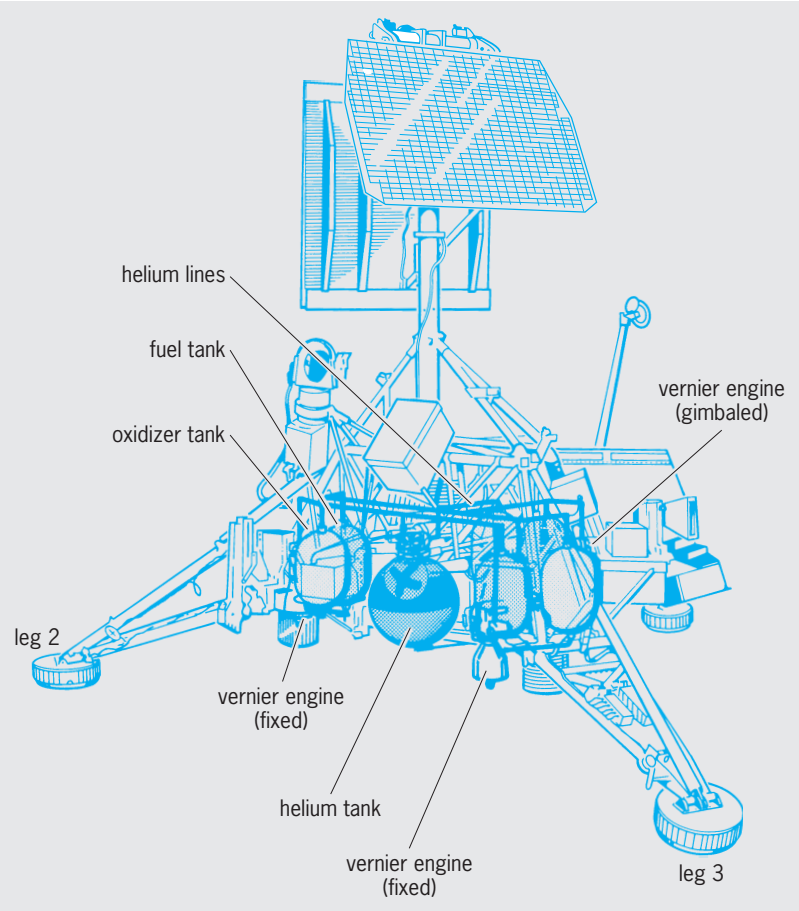

Fig. 4. Partial schematic of *Surveyor* to show part of the propulsion system associated with the vernier engines which provided midcourse trim and attitude control during firing of the large retrorocket, and final velocity removal near the lunar surface. (*JPL/NASA*)

multiple atmospheric probes to Venus. *See* JUPITER; SATURN.

Ranger. In the early 1960s, the United States carried out this series of missions to relay photographs of the Moon as the spacecraft descended to a crash landing. *Ranger 1* and *2* ended in launch vehicle failures; *Ranger 3* missed the Moon by 22,862 mi (36,793 km); *Ranger 4* impacted the Moon but experiments proved inoperative when a timer failed; *Ranger 5* missed the Moon by 450 mi (725 km); and *Ranger 6* impacted the Moon but its television camera failed. *Ranger 7* was the first success, returning 4308 photos during its lunar descent in July 1964. In 1965 *Ranger 8* and *9* sent back 7137 and 5814 photos, respectively.

Surveyor. These United States missions (**Fig. 4**) soft-landed robot spacecraft on the Moon as precursors to the crewed Apollo missions. *Surveyor 1* landed on the Moon on June 2, 1966, and returned 11,150 photos as it operated for 6 weeks. *Surveyor 2* suffered from a vernier engine failure and crashed into the lunar surface. *Surveyor 3* carried a soil sampler for its soft landing in 1967. *Surveyor 4* touched down on the Moon, but its signal was lost $2^1/_2$ min later. *Surveyor 5* soft-landed in 1967, and relayed 19,000 photos and conducted soil analyses. *Surveyor 6* executed the first takeoff from the Moon during its 1967 mission. *Surveyor 7*, which soft-landed in 1968, was the last in the series.

Lunar Orbiter. These spacecraft launched during the same era as the *Surveyors* were the first attempt to make detailed maps of the Moon. *Lunar Orbiter 1* was launched August 10, 1966, and relayed photos for the next 20 days. *Lunar Orbiter 2* sent 205 photographic frames later that year; *Lunar Orbiter 3* sent 182 photos in 1967, followed by *Lunar Orbiter 4* and *5*.

Mariner. This designation was given to the United States' early planetary probes, which chiefly studied Venus and Mars, although the final in the series also visited Mercury. *Mariner 1* was lost in 1962 after a typographical error in a software program caused the launch vehicle to veer off course. *Mariner 2* became the first spacecraft to encounter another planet when it flew by Venus at a distance of 21,594 mi (34,752 km) later that year. It continued to relay data to a distance of 53.9×10^6 mi (86.7×10^6 km). *Mariner 3* and *4* (**Fig. 5*a***) targeted Mars; the first experienced a shroud failure, while the second became the first spacecraft to encounter Mars in 1964. *Mariner 5* (Fig. 5*b*) executed a 2480-mi (3991-km) Venus flyby in 1967. *Mariner 6* and *7* (Fig. 5*c*) again targeted Mars, sending back 75 and 126 television pictures, respectively, during flybys at a distance of about 2150 mi (3450 km) in 1969. *Mariner 9* was a Mars orbiter, sending back detailed pictures after its launch in 1971. *Mariner 10* was the first spacecraft to use a gravity-assist trajectory, encountering Venus before going on to three flybys of Mercury in 1974 and 1975. The United States then planned a mission with the working title *Mariner Jupiter-Saturn*, but the name was changed to *Voyager* before launch. *See* MARS; MERCURY (PLANET); VENUS.

Venera, Mars, and Zond. These missions were the Soviet Union's equivalent of the Mariners. The Soviets had the greatest success with *Venera*, which lofted a number of spacecraft to Venus over more than two decades. Although the first three *Venera* failed in 1961 and 1965, *Venera 4* transmitted during its 94-min descent, though it failed at a considerable altitude above to Venus's extremely hot surface. The series continued with *Venera 5* and *6* in 1969. *Venera 7* made the first soft-landing of Venus in 1970 and transmitted for 23 min, which was remarkable considering the planet's 900°F (500°C) surface temperature. *Venera 8* made another successful landing in 1972. In 1975 the Soviet Union launched *Venera 9* and *10*, which combined Venus orbiters and landers; *Venera 9* returned the first pictures from the surface of another planet. In 1978 and 1981 the landers *Venera 11*, *12*, *13*, and *14* followed. *Venera 15* and *16*, launched in 1983, were orbiters that used synthetic-aperture radar to pierce Venus's opaque atmosphere and map the planet's surface.

The Soviets' Mars program achieved more mixed results. After the failure of *Mars 1* in 1962, the Soviets attempted to make soft landings in 1971 with *Mars 2* and *3*. The first failed when it evidently impacted the planet's surface with excessive force; the

second touched down but failed 110 s later, having transmitted only a partial photograph that revealed no detail. In 1973 the Soviets made another attempt with a four-spacecraft armada, the orbiters *Mars 4* and *5* and the landers *Mars 6* and *7*. *Mars 4* and *7* missed the planet entirely, while *Mars 6* stopped transmitting during its descent. *Mars 5* entered orbit, but communications proved to be erratic.

Zond was the designation given to several Soviet missions with various targets from 1964 through 1970. *Zond 1* and *2* were failed Venus and Mars probes, respectively, while *Zond 3* returned photos from a lunar flyby in 1965. *Zond 5* flew around the Moon and returned to Earth, splashing down in the Indian Ocean in 1968. In addition to other objectives, this and later *Zond* missions carried tortoises and other organisms for biological research. *Zond 6* flew around the Moon before returning to crash in the Soviet Union. *Zond 7* also was quite successful. *Zond 8* performed another lunar return in 1970.

Viking. One of the most ambitious planetary explorations ever mounted was the United States' *Viking* program. Two launches of combined orbiter-landers, *Viking 1* and *2*, took place in August and September 1975, with arrival at Mars in June and August 1976. *Viking*'s chief purpose was the search for biological activity on the Martian surface. Both the orbiters and landers also carried out extensive imaging; the orbiters reconnoitered landing sites in preparation for descent of the landers and also carried out a global mapping program. The landers (**Fig. 6**) sent back the first extensive detailed pictures from the surface of another planet.

The spacecraft failed to detect any sign of life on the planet. Many scientists, however, later concluded that *Viking*'s search methodology was very narrow, and began to devise other ways to search for past or present life on Mars. The *Viking* orbiters revealed a landscape that appeared to have been heavily refigured by liquid water, but it appeared for many years that none exists currently on the frigid planet. In 2006, images taken by *Mars Global Surveyor* provided evidence suggesting that on current-day Mars water has broken out of cliffs or crater walls and run downhill before evaporating. Mar's polar caps have been found to be a mixture of carbon dioxide ice and water ice.

Helios. This is a German-built and United States–launched solar system probe which because of its close approach to the Sun (0.29 AU) is called a solar probe. *Helios 1* was launched in 1974, followed by *Helios 2* in 1976. The experiments, directed toward phenomena taking place in the solar wind, were chiefly magnetometers and particle detectors. *See* SOLAR WIND.

Voyager. In the late 1960s, mission designers began to contemplate a coming opportunity that would not be repeated for another 177 years. In the 1970s it would be possible to launch a spacecraft that could sequentially visit all four giant gaseous outer planets—Jupiter, Saturn, Uranus, and Neptune—by using the technique of gravity assist to send the probe on to its next destination.

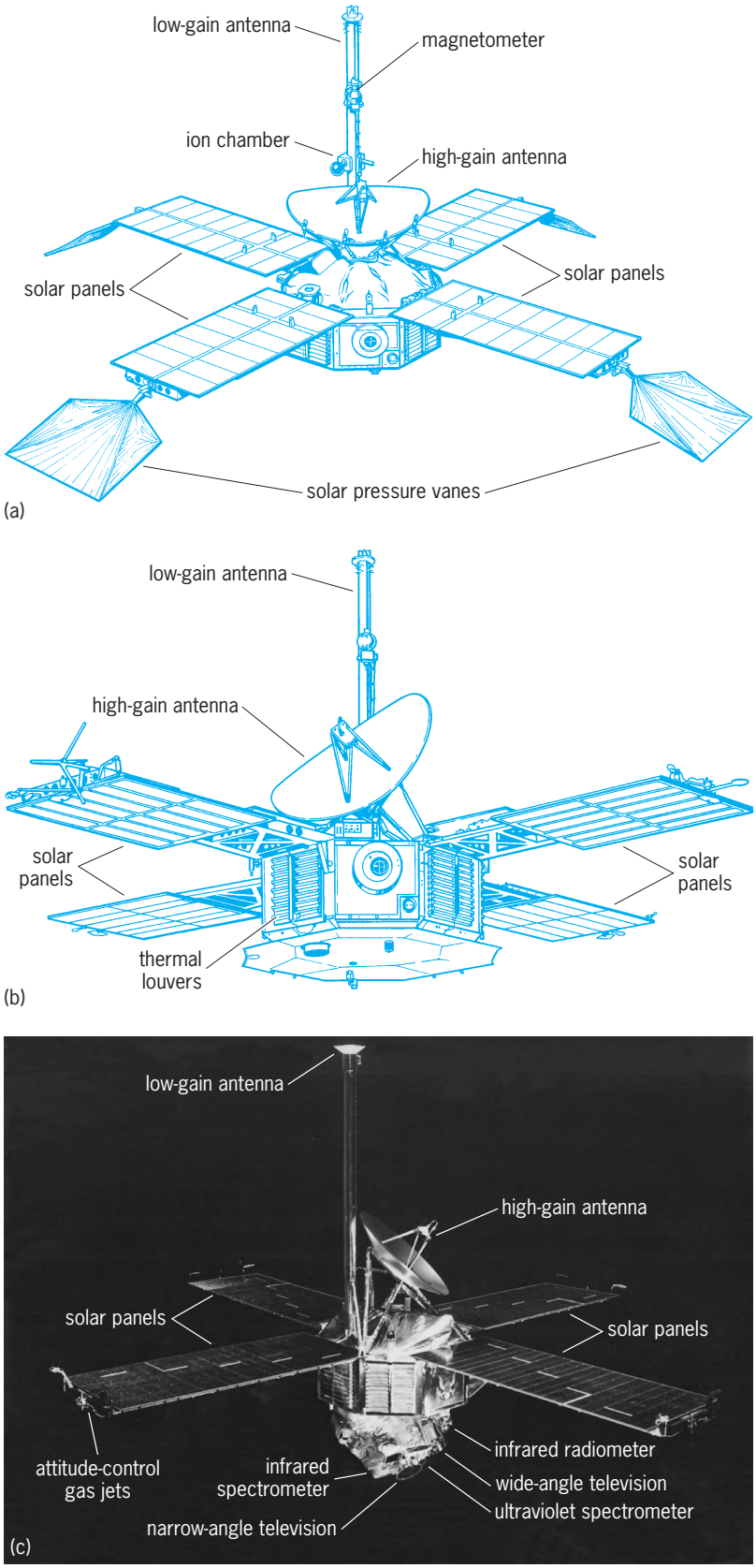

Fig. 5. Evolution of the *Mariner* spacecraft. (*a*) *Mariner 4*. (*b*) *Mariner 5*. Active thermal louvers, shown on the sides of the main body, are driven by temperature-sensitive motors which control the degree of opening, and thus the heat-emission rate from the spacecraft. Many changes from *Mariner 4* are not seen in this gross view from the outside. (*c*) *Mariner 6* and *7*. Although only about 300 lb (135 kg) heavier than *Mariner 4* and *5*, this type of spacecraft is much more complex because of advances in electronics. (*NASA*)

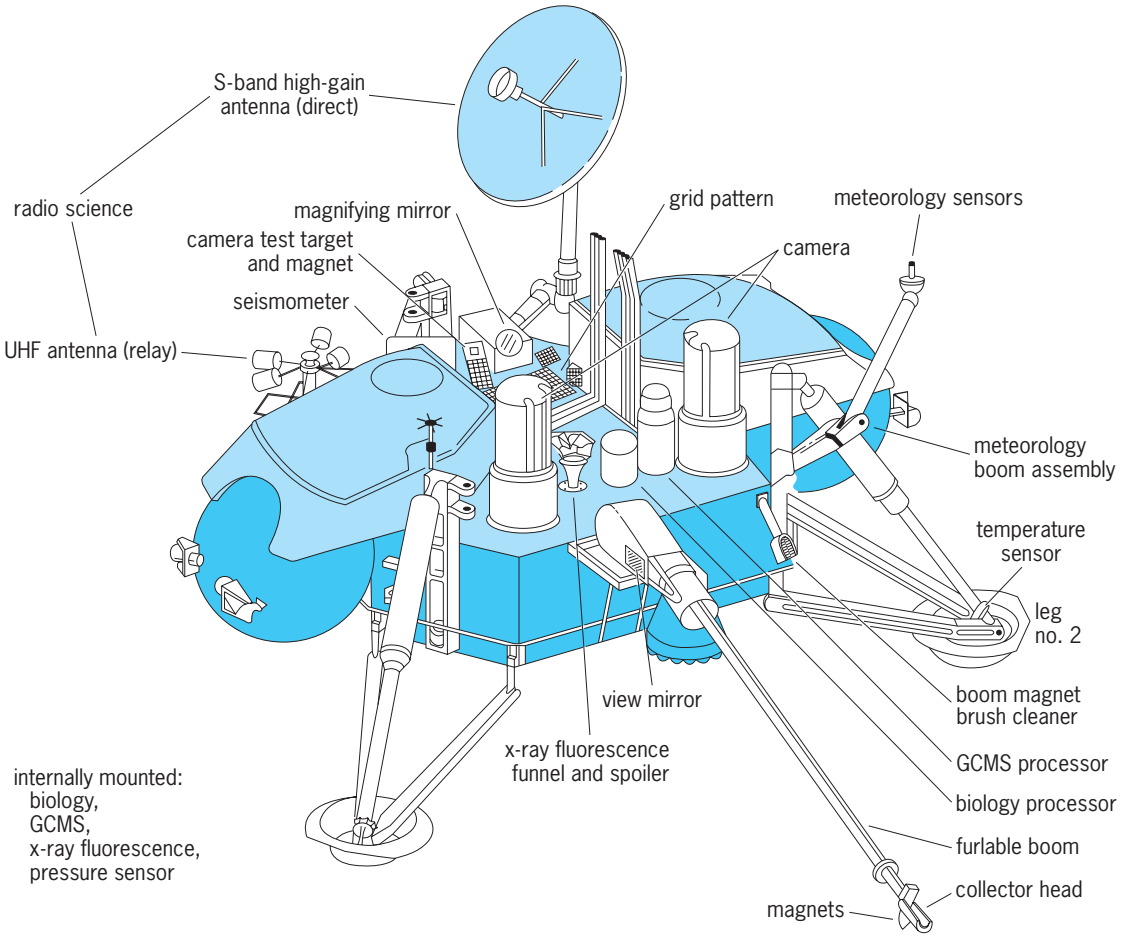

Fig. 6. ***Viking*** **lander, with scientific experiments. (*Viking Press Kit, NASA News Release 75–183, 1975*)**

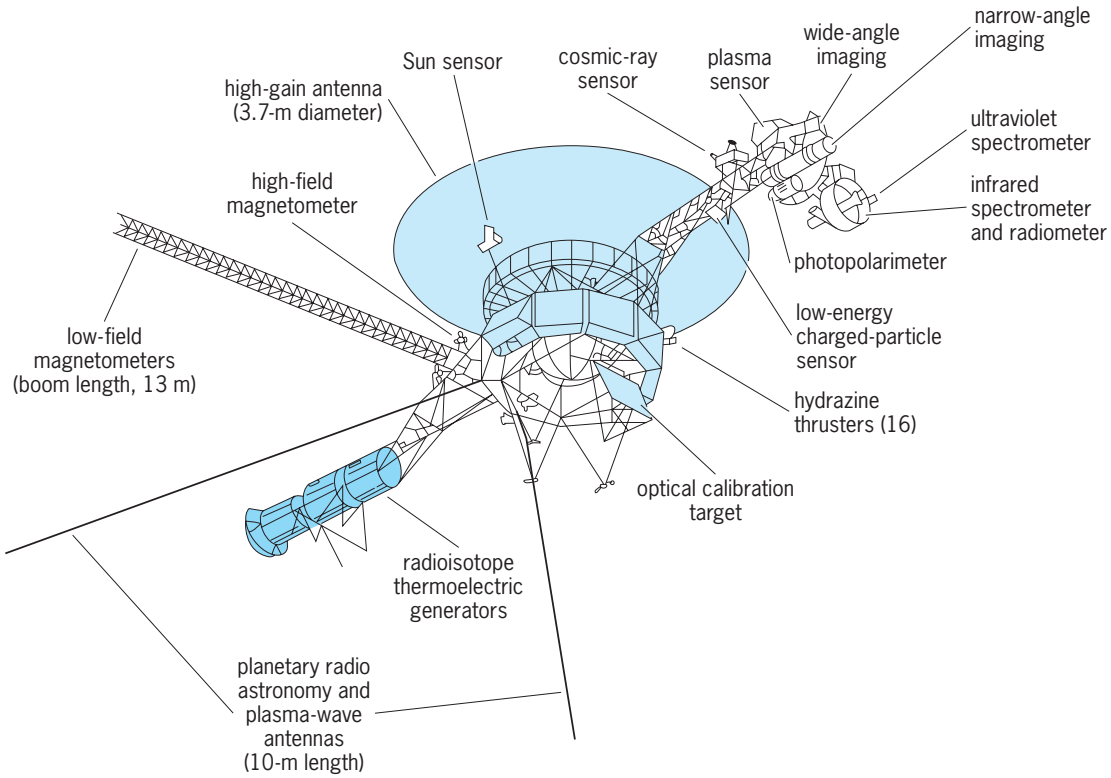

Fig. 7. ***Voyager*** **spacecraft showing primary subsystems. 1 m = 3.3 ft. (*NASA*)**

Efforts culminated with the launch in 1977 of *Voyager 1* and *2* (**Fig. 7**), which began a decade-long "grand tour" of the outer solar system. Although *Voyager 2* was launched first, *Voyager 1* took a shorter trajectory to Jupiter and arrived in 1979; it went on to encounter Saturn in 1980. *Voyager 1*'s trajectory at Saturn was optimized to study the planet's major satellite, Titan, which was found to contain complex organic chemistry in its opaque atmosphere. This flight path precluded the spacecraft from later planetary encounters. *Voyager 2* visited Jupiter in 1979 and Saturn in 1981, then continued on to flybys of Uranus in 1986 and Neptune in 1989. During the 1990s and the first decade of the twenty-first century, both spacecraft continued to relay back fields and particles data as they exited the solar system. *See* NEPTUNE; URANUS.

Comet missions. The arrival of Comet Halley to the inner solar system in the mid-1980s prompted a great deal of international interest in sending spacecraft to meet it. The Soviet Union and Japan each dispatched two spacecraft, and the European Space Agency a single craft. The efforts were notable for a high degree of international cooperation, including the use of data obtained by the NASA Deep Space Network from the early-arriving Soviet craft to target more precisely the flyby of the European craft. *See* COMET; HALLEY'S COMET.

Plans originally called for a United States spacecraft to join the Halley armada, but that mission was the victim of a budgetary cancellation. United States mission planners developed an ingenious scheme to redirect the *International Sun-Earth Explorer 3* (*ISEE 3*) satellite to conduct a flyby of the tail of Comet Giacobini-Zinner on September 11, 1985, about 6 months before the Comet Halley flybys.

International Comet Explorer (ICE). This was the new name given the *ISEE 3* (**Fig. 8**) when it was redirected to make its comet visit. The craft was originally placed in a halo orbit, a conditionally stable circular orbit with its center at the Lagrange point upstream (toward the Sun) from the Earth-Moon system. In 1982 and 1983 a series of maneuvers in the Earth-Moon system culminated in a final close flyby of the Moon that sent the craft on a new trajectory to fly through the tail of Comet Giacobini-Zinner, some 6000 mi (10,000 km) from the comet's nucleus. The craft's payload was a half-dozen instruments probing plasma electrons, plasma ions, energetic protons, magnetic fields, plasma waves, and radio waves. *See* CELESTIAL MECHANICS.

Suisei and Sakigake. These two spacecraft represented the first solar system exploration mission launched by the Japanese Institute of Space and Astronautical Science (ISAS). *Sakigake* was launched in January 1985 and flew within a few million miles of Comet Halley on March 11, 1986. The spin-stabilized craft carried a solar-wind ion detector, a magnetometer, and a plasma-wave probe.

The *Suisei* craft (**Fig. 9**), also called *Planet-A*, was launched in August 1985 and approached within about 90,000 mi (150,000 km) of Comet Halley on March 9, 1986. Like its sibling craft, *Suisei* was spin-

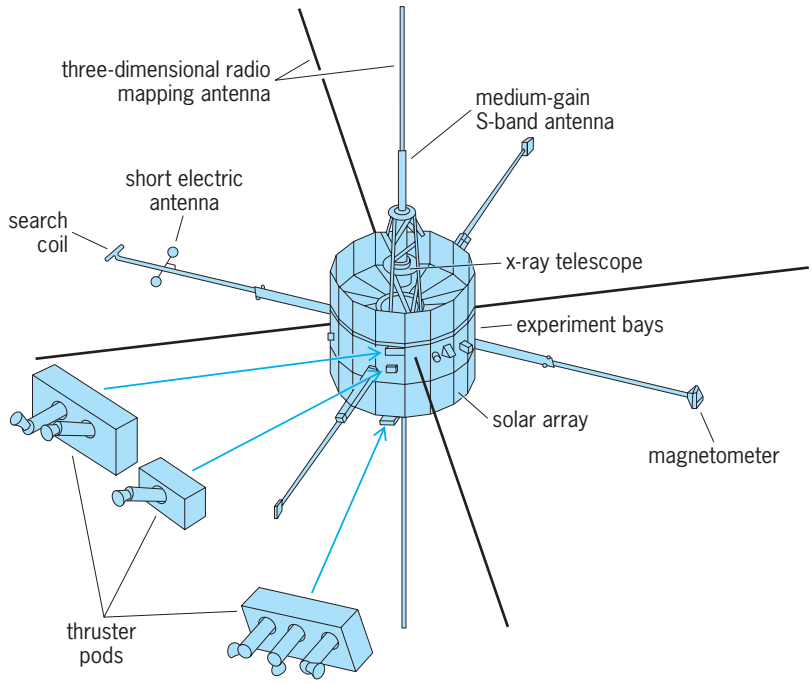

Fig. 8. ***International Sun-Earth Explorer 3*** **(*ISEE 3*), the last satellite of a coordinated series of three designed to study the Earth's magnetosphere. This satellite was redirected toward Comet Giacobini-Zinner by a lunar encounter, and was renamed *International Cometary Explorer* (*ICE*). (*ISEE Science Working Team*)**

stabilized, but it carried a different science payload: an ultraviolet camera and a solar-wind experiment.

VEGA. This enterprising mission by the Soviet Union sent a pair of spacecraft on flybys of Venus, where each released a lander and an atmospheric balloon, before continuing on to encounter Comet Halley (**Fig. 10**).

VEGA 1 and *2* were launched in December 1984 and flew by Venus in June 1985. *VEGA 1*'s Halley encounter took place March 6, 1986, while *VEGA 2*'s followed 3 days later. Each craft carried optical cameras, spectrometers, dust detectors, and fields and particles instruments. Unlike the Japanese and European craft, the *VEGA*s were three-axis-stabilized.

The *VEGA*s achieved a closer flyby distance from Halley than did the Japanese probes, encountering the comet at distances of 5523 and 4990 mi (8889 and 8030 km). They were able to provide highly useful targeting data for the European probe, which was to make by far the closest pass by the comet.

Giotto. This spacecraft (**Fig. 11**) was the European Space Agency's representative in the Comet Halley fleet. At its closest approach, *Giotto* came within about 360 mi (600 km) of the comet.

The spin-stabilized craft's payload included a camera, spectrometers, a photopolarimeter, dust experiments, and fields and particles instruments. Launch took place in July 1985, with the Halley flyby on March 14, 1986.

Because the environment near the comet's nucleus was known to be very dusty, *Giotto* was designed to endure a high number of particle impacts. Although only 2–3% of the spacecraft's solar-panel capacity was lost during the encounter, the mirror on the camera began to suffer degradation when

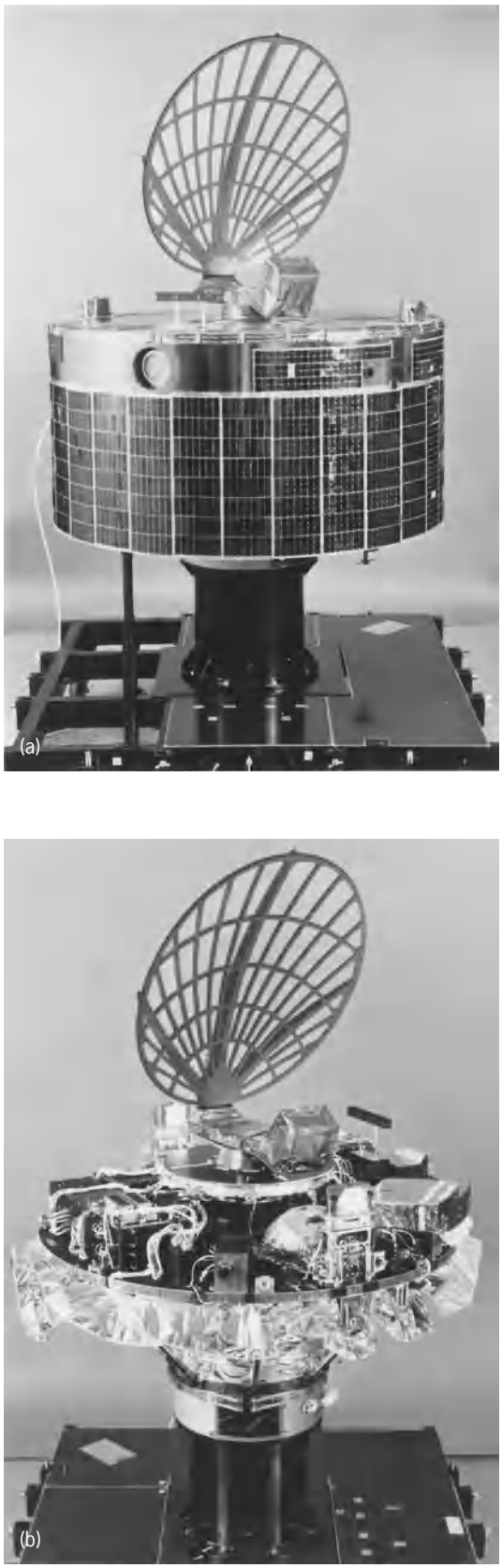

Fig. 9. Model of the Japanese comet probe *Suisei* (*Planet-A*). (*a*) Assembled spacecraft with the solar cells assembly, high-gain antenna, and ports for instruments. (*b*) Inside view of the spacecraft with instruments arranged cylindrically about the spin axis. The extensive foil is for thermal protection and heat control. (*Institute of Space and Astronautical Science*)

Giotto was inbound at a distance of about 3000 mi (5000 km), and the last usable image was taken at a distance of about 1220 mi (1963 km). In addition, particle impacts caused the craft to wobble with a nutation of 0.8°, resulting in a loss of radio signal 14 s before closest approach. The wobble was overcome, however, about 32 min later, and *Giotto* remained a highly functional probe years after the encounter.

Phobos. An ambitious project for the exploration of Mars and its natural satellite, Phobos was undertaken by the Soviet Union in the late 1980s. Unfortunately, the two spacecraft were not to fulfill completely their planned activities.

Phobos 1 and *2* were launched 5 days apart in July 1988. The two nearly identical craft were to arrive at Mars in January 1989 and begin orbiting the planet. A series of maneuvers would have taken the craft near Phobos, where they would approach for a surface-contour-following flyby about 150 ft (50 m) from the rocky object's surface. The craft would also have deployed a 110-lb (50-kg) long-duration lander, expected to operate for about a year, and a 112-lb (51-kg) hopper, which would have moved about the surface of Phobos in 60-ft (20-m) leaps. Altogether, the craft included a total of 37 scientific experiments.

A series of commands sent to *Phobos 1* in August 1988 apparently contained a coding error that instructed the spacecraft to turn off its attitude-control system. The craft then began to tumble slowly; as its solar panels were no longer pointed toward the Sun, it lost power and all contact was permanently lost.

Problems arose with *Phobos 2* by the end of 1988 when its television system malfunctioned en route to Mars. Those difficulties were resolved by ground command. More serious was the loss of the on-board high-power radio transmitter, which would have permitted a high rate of data transmission. It was still possible, however, to use the slower low-power transmitter to send data to Earth.

Phobos 2 survived the transit of Mars, entered orbit in late January 1989, and began its scientific surveys of the red planet. By late March, the craft was positioned near the satellite Phobos, capturing high-detail images from a distance of several hundred miles. On March 27, 1989, however, a failure in the attitude control system caused the craft to begin tumbling erratically in an orbit just outside that of Phobos. Ground controllers spent a week trying to resume communications before giving up the craft as lost.

Magellan. In the late 1970s and the 1980s, improvements were made in synthetic-aperture radar, a technology that uses radar pulses to capture images where optical systems are not practical. This technology is ideal for the study of the planet Venus, which is perpetually shrouded by a thick layer of opaque clouds. *See* SYNTHETIC APERTURE RADAR (SAR).

The first craft to perform radar imaging at Venus, *Pioneer Venus* of the United States was followed by the Soviet *Venera 15* and *16* probes, which were able to improve resolution by about a factor of 10. They were followed by the United States' *Magellan*

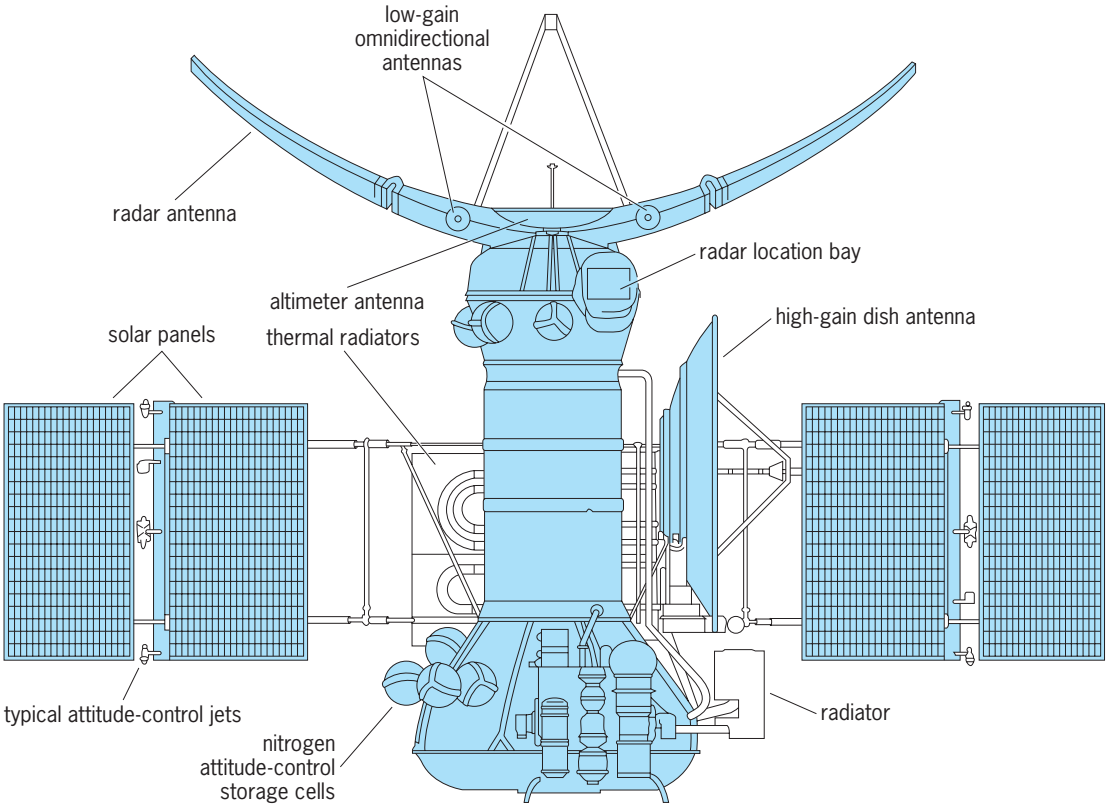

Fig. 10. ***VEGA*** **spacecraft in fully deployed view.**

mission, which in turn achieved a factor-of-10 improvement over the resolution delivered by the Soviet missions.

Launched in May 1989, *Magellan* (**Fig. 12**) was also the first of several solar system probes to be lofted by the United States space shuttle. The probe was carried into Earth orbit in the cargo bay of the shuttle and released. A solid-fuel upper-stage engine attached to the probe then injected it into its interplanetary trajectory.

Magellan entered Venus orbit in August 1990, initiating a highly successful 4-year study of the planet. Each 8 months, Venus rotated once underneath the spacecraft as it passed overhead in a polar orbit. *Magellan* spent several of these 8-month orbital cycles producing radar maps of 98% of the Venusian surface with a resolution of 1000 ft (300 m) before compiling a high-resolution, comprehensive gravity field map for 95% of the planet. In May 1994, ground controllers experimented with aerobraking, making use of the planet's atmosphere to lower the spacecraft's orbit. With solar panels and other components gradually failing, mission controllers decided to plunge the spacecraft into Venus's atmosphere to study the craft's behavior and atmospheric characteristics. Contact was lost with *Magellan* on October 12, 1994, as the craft presumably descended and broke apart.

Galileo. After the initial reconnaissance of the solar system performed by the *Mariners* and *Voyagers* from the 1960s through the 1980s, mission planners

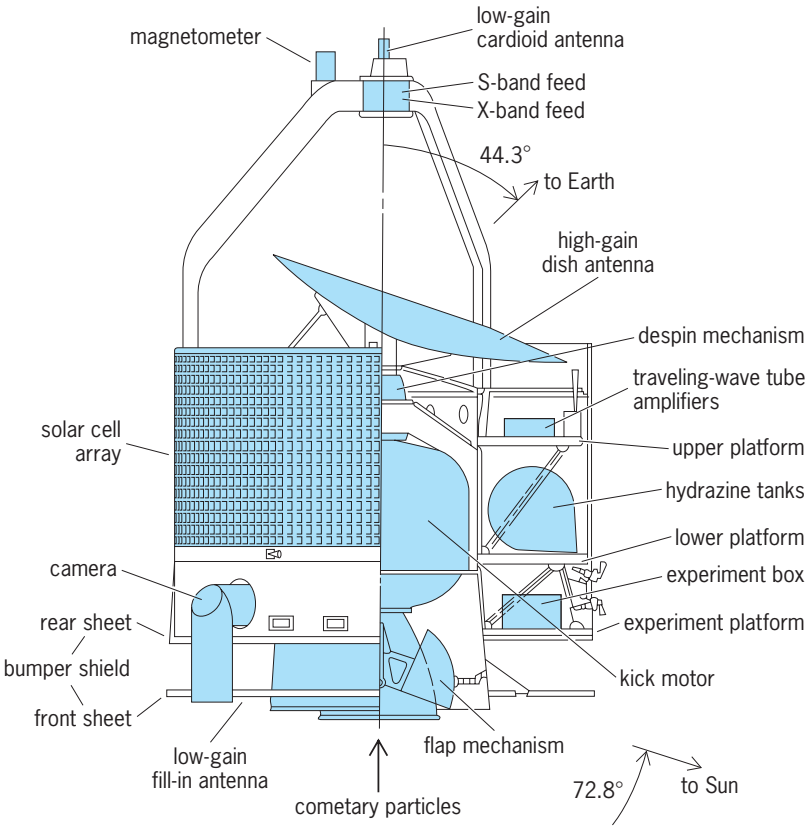

Fig. 11. Cutaway view of the ***Giotto*** **comet-encounter spacecraft showing the primary subsystems. (*European Space Agency*)**

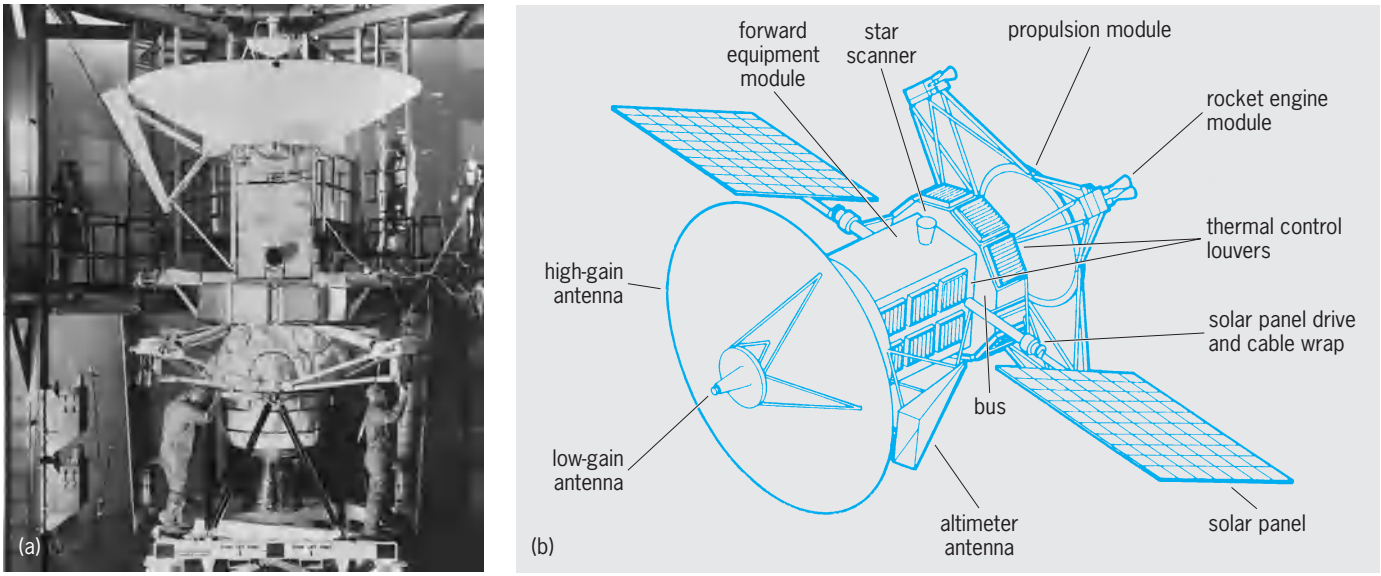

Fig. 12. ***Magellan*** **spacecraft. (*a*) Undergoing testing. (*b*) Deployed, showing primary subsystems. (*JPL/NASA*)**

began to consider the next step in planetary exploration. Their plans called for a series of orbiters that would return to each of the planets visited by flyby spacecraft to conduct detailed studies. The first such approved mission was *Galileo*, which would orbit the solar system's largest planet, Jupiter.

Galileo (**Fig. 13**) had a complex dual-spin design in which most of the spacecraft was spin-stabilized to optimize fields and particles experiments, while the platform that carried pointed instruments such as cameras was capable of being three-axis-stabilized. The orbiter also carried a descent probe designed to relay data as it dived into the giant planet's atmosphere.

En route to Jupiter, *Galileo* performed a flyby of Venus and two flybys of Earth. It also performed the first-ever flyby of an asteroid when it encountered the rocky body Gaspra in October 1991. In August 1993, *Galileo* flew by a second asteroid, Ida. When images were sent to Earth from the craft's on-board tape recorder 6 months later, the first-ever natural satellite of an asteroid was discovered when the tiny moon Dactyl was found to be orbiting Ida. *See* ASTEROID.

The failure of *Galileo*'s umbrella-like high-gain antenna to open fully prompted an extensive redesign of planned observations at Jupiter. After its arrival in December 1995, the mission's highest priority was

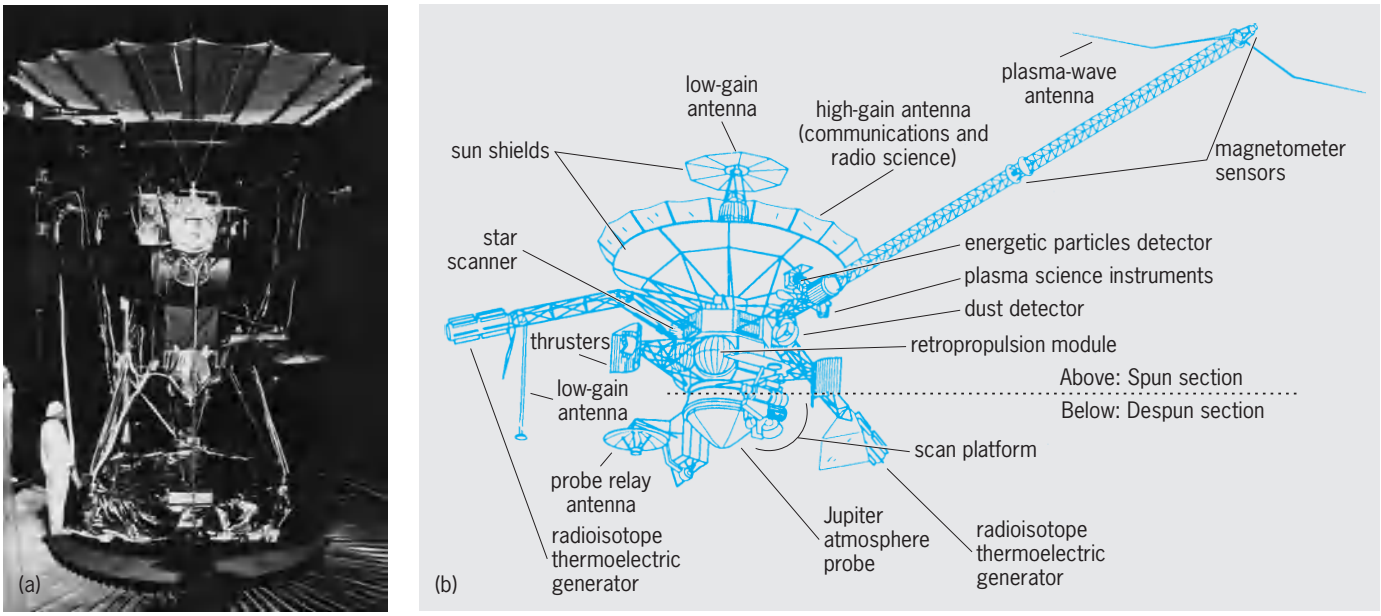

Fig. 13. ***Galileo*** **spacecraft. (*a*) Being tested. (*b*) Deployed, showing primary subsystems. The scan platform contains the photopolarimeter radiometer, near-infrared mapping spectrometer, solid-state imaging camera, and ultraviolet spectrometer. (*JPL/NASA*)**

to relay data from the atmospheric probe during its descent. *Galileo* then began sending selected images and other scientific data that had been compressed in on-board computers to make the best use of the craft's available low-gain antennas.

Galileo made about 30 orbits of Jupiter in the first 5 years after its arrival. After completing an initial $2^1/_2$-year prime mission, *Galileo* was approved for an extended mission focusing on the Jovian satellites Europa and Io. The mission was extended a second time to allow the spacecraft to focus on observations of Jupiter's largest satellite, Ganymede, and on collaborative studies with *Cassini* as that spacecraft passed Jupiter in late 2000 en route to Saturn. *Galileo* discovered strong evidence that Europa has a melted saltwater ocean under an ice layer on its surface. The spacecraft also found indications that Ganymede and another Jovian satellite, Callisto, have layers of liquid saltwater as well. Other major science results from the mission include details of varied and extensive volcanic processes on the moon Io, measurements of conditions within Jupiter's atmosphere, and the discovery of a magnetic field generated by Ganymede.

On September 21, 2003, *Galileo* was intentionally plunged into Jupiter's atmosphere to avoid the risk of impacting and possibly contaminating Europa.

Hiten and Hagoromo (Muses-A). After their success with its two Comet Halley craft, mission planners in Japan turned to other projects they could execute as the nation's space program developed. Their next effort was *Muses-A*, a mission to study the Moon.

An expendable rocket launch in January 1990 carried aloft two spacecraft, *Hiten* and *Hagoromo*. The latter, a basketball-sized package carrying an instrument payload, was to go into orbit around the Moon. The larger *Hiten* was to occupy a highly elliptical orbit around Earth, ranging in distance between 19,000 and 470,000 mi (30,000 and 760,000 km). This orbit would take *Hiten* out toward the Moon, where it would receive and relay to Earth signals from *Hagoromo*. As *Hagoromo* approached the Moon, however, its radio transmitter failed, ending all contact with the craft.

Ulysses. By the 1980s, the Sun had been studied for many years by ground-based equipment. Missions such as the *Orbiting Solar Observatory* satellites and the crewed *Skylab* in the 1970s extended these observations to Earth orbit. All observing equipment, however, was limited to the two-dimensional surface of the ecliptic, the plane in which most planets orbit the Sun.

Ulysses was designed to extend that point of view by sending a spacecraft out of the ecliptic and into the unexplored regions above and below the Sun's poles. After collaborative work between NASA and the European Space Agency and the budgetary cancellation of a United States craft that was to have taken part, a project emerged wherein the European agency built and provided the single *Ulysses* spacecraft (**Fig. 14**) and contributed a computer system

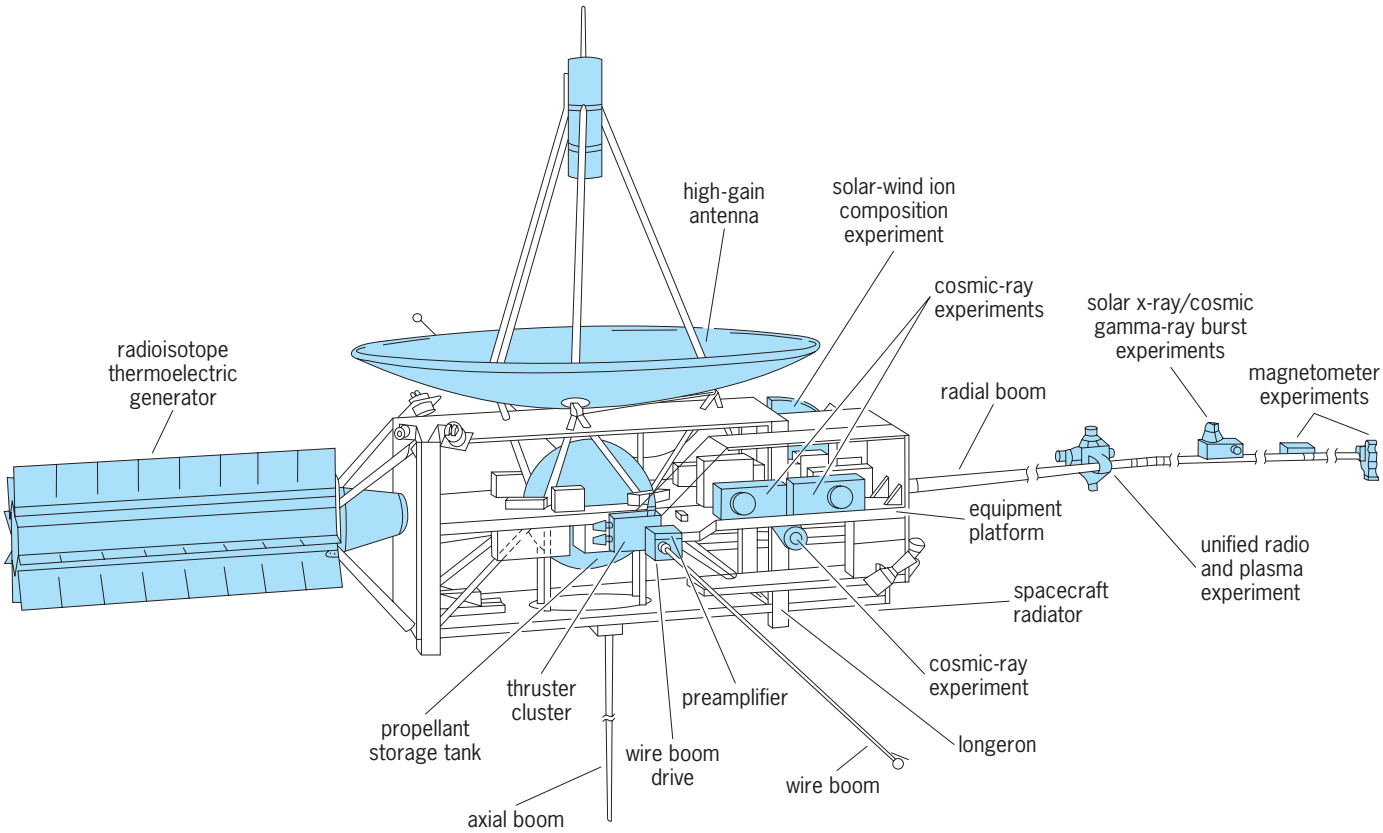

Fig. 14. ***Ulysses*** **spacecraft, showing primary subsystems. The propellant storage tank is part of the reaction control equipment (for attitude and so forth), and the thruster cluster is essentially a set of small engine nozzles for this equipment. The preamplifier is part of the unified radio and plasma wave experiment. (*European Space Agency*)**

to control the craft from the ground. NASA provided some instruments, the radioisotope power source, launch by a United States space shuttle, tracking via its global Deep Space Network, and ground facilities at the Jet Propulsion Laboratory in Pasadena, California.

Launched in October 1990, *Ulysses* flew first to Jupiter, where the giant planet's gravity sent the craft out of the ecliptic in February 1992. The encounter of the Sun's south pole took place in 1994, with an encounter of the north pole in 1995. As the mission progressed, one of the key scientific findings revealed a Sun with a homogeneous magnetic field lacking latitudinal variations, very different from the magnetic climate observed by ground-based instruments. *Ulysses*'s observations continued after its flight over the Sun's poles, and it made another passage over the Sun's polar regions under conditions of high solar activity in 2000 and 2001. In 2006 it began its third passage over the Sun's poles.

Mars Observer. This mission was to be the United States' first return to the red planet since the *Viking* effort of the 1970s. A polar orbiter, *Mars Observer* carried a high-resolution camera and other instruments to study Mars's geology and atmosphere. The spacecraft was adapted from an assembly-line commercial Earth-orbiting satellite in an effort to streamline design and save costs.

Launched on a Titan 3 rocket in September 1992, the mission proceeded according to plan until 3 days before Mars arrival in August 1993, when all contact was lost with the craft. The loss took place during a period in which *Mars Observer*'s propulsion system was being pressured in preparation for the firing of its main engine to place it on orbit; the craft's radio transmitter was intentionally turned off at the time to protect it from any forces during the pressurization. A review panel concluded that a failure probably occurred in the propulsion system that either ruptured a fuel line or caused an explosion.

Clementine. Unlike most United States missions, carried out by NASA, this novel effort was the work of the Department of Defense. *Clementine*'s flight plan called for it to orbit Earth for several days before departing for the Moon, where it would spend more than 2 months in orbit capturing high-resolution images of the lunar surface. The spacecraft was then to leave once more on a trajectory that would take it within 60 mi (100 km) of the asteroid 1620 Geographos.

After launch in January 1994, *Clementine* (**Fig. 15**) arrived in lunar orbit in February and carried out systematic mapping; on May 5, it left lunar orbit. Two days later, a malfunction of the on-board computer caused several of the craft's thrusters to be activated during a 20-min period when *Clementine* was out of touch with ground controllers. This depleted all fuel from the craft's tanks, preventing the asteroid flyby.

Although it returned a wealth of lunar imagery, the chief objective of *Clementine* was not scientific, but testing 23 advanced technologies of interest to the military. Moreover, an aim was to demonstrate a streamlined project management approach that would make it possible to build and launch a small yet highly capable spacecraft cheaply in less than 2 years.

Cassini. Considered the last of the its era of large United States planetary exploration missions, *Cassini* (**Fig. 16**) was launched in 1997 to study the ringed planet Saturn and its major natural satellite, Titan. The *Cassini* orbiter carried an instrumented probe provided by the European Space Agency, *Huygens*, designed to descend to the surface of Titan. Shrouded by an opaque atmosphere containing a rich diversity of organic compounds, Titan possibly contains liquid oceans of methane. Like *Galileo*, the craft has made use of multiplanet gravity assists, flying by Venus, Earth, and Jupiter on the way to its destination.

The *Cassini* and *Galileo* spacecraft made coordinated observations of Jupiter on *Cassini*'s flyby of the giant planet on December 30, 2000. *Cassini* data were used to track daily changes of some of the planet's most visible storms and to study patterns in natural radio emissions near the edge of Jupiter's magnetic environment.

After 7 years in transit, *Cassini* entered orbit at Saturn on July 1, 2004. Six months later, on January 14, 2005, the *Huygens* probe took photos as it descended to the surface of Titan, where it touched down and continued to send data for about 90 min. Preliminary findings confirmed that the landing site is near the shoreline of a liquid ocean.

The *Cassini* orbiter, meanwhile, continued on numerous close approaches to Saturn's moons and studies of its atmosphere, rings, and magnetic field. One notable early science finding was the discovery that Saturn's moon Enceladus may have liquid water reservoirs that erupt in geysers.

Fig. 15. ***Clementine*** **spacecraft being prepared for testing at Naval Research Laboratory, Washington, DC.**

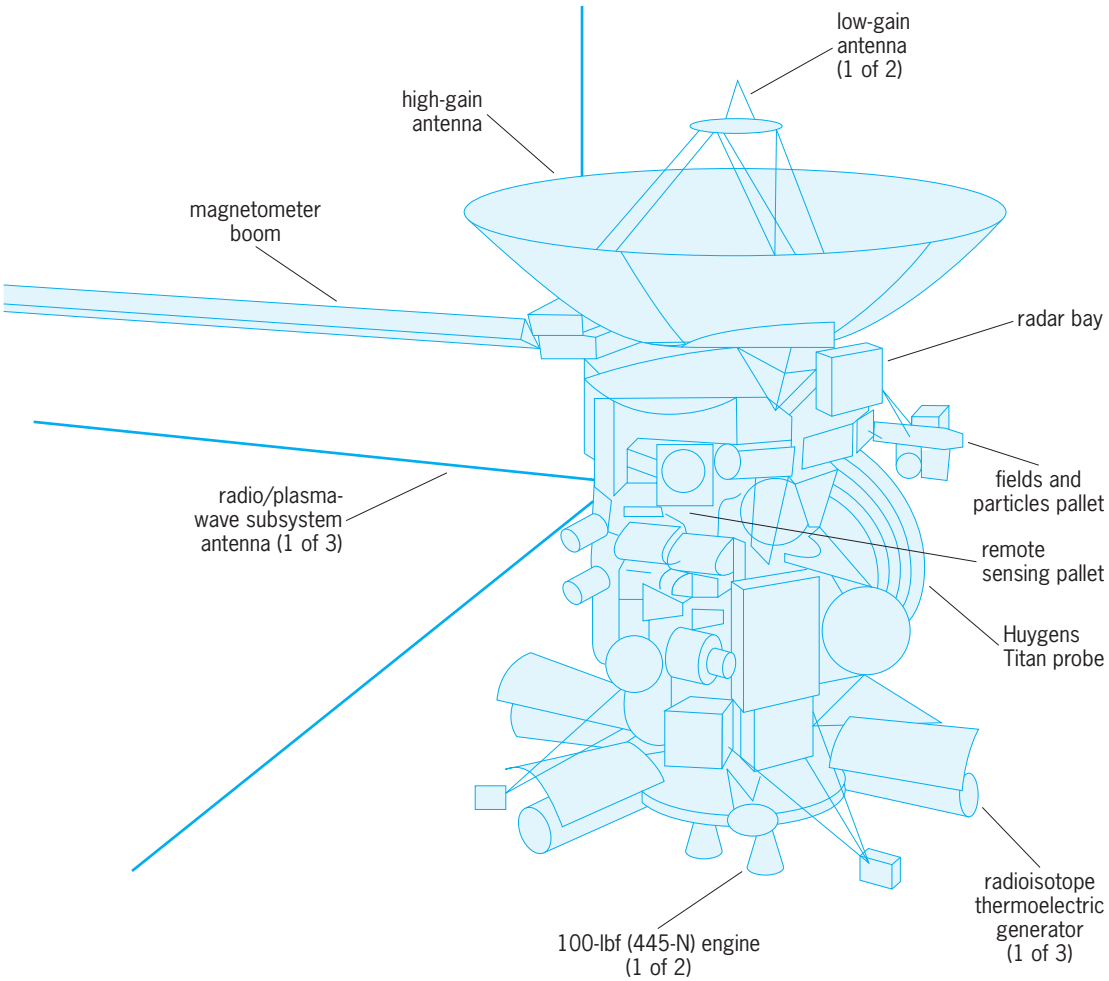

Fig. 16. ***Cassini*** **spacecraft, showing primary subsystems. (*JPL/NASA*)**

Discovery Programs. Faced with constrained budgets but increasingly complex and costly planetary spacecraft, NASA in the mid-1990s undertook a conscious program of downsizing missions. At the center of this effort was the Discovery Program, designed to sponsor low-cost solar system missions with highly focused science objectives. The program periodically makes competitive selections from proposals submitted by teams led by scientific principal investigators. Early selected missions included *Near Earth Asteroid Rendezvous*, *Mars Pathfinder*, *Lunar Prospector*, *Stardust*, *Comet Nucleus Tour*, *Messenger*, and *Deep Impact*.

Near Earth Asteroid Rendezvous (NEAR). The *NEAR* probe was launched in 1996 to rendezvous with and orbit the asteroid Eros, one of the small rocky bodies that pass relatively near Earth. After passing asteroid 253 Mathilde in 1997, the probe was scheduled to enter orbit around Eros in late 1998. Due to a problem with flight software, however, the orbital insertion engine firing did not take place correctly, and *NEAR* sailed past Eros. However, the craft was able to try again about a year later after circling once around the Sun. *NEAR* correctly entered orbit around Eros on February 14, 2000, and began a series of maneuvers to take it close to the surface of the small rocky body. The spacecraft, renamed *NEAR-Shoemaker* in honor of the late astronomer Eugene Shoemaker, gently landed on asteroid Eros on February 12, 2001, after orbiting the rocky body for a year. It captured 69 images during the final 3 mi (5 km) of its descent, revealing features as small as 0.4 in. (1 cm) across. The final signals from the spacecraft before it was shut down were received February 28, 2001.

Lunar Prospector. One of the smallest of the early Discovery missions, *Lunar Prospector* was launched in 1998 and reached the Moon after 5 days in flight. Orbiting 60 mi (100 km) above the lunar surface, the probe carried five instruments designed to help understand the origin and evolution of the Moon and to determine whether water ice may be present in shaded areas near the Moon's poles. After a year and a half in orbit, *Lunar Prospector* was intentionally crash-landed in a perpetually shadowed crater near the Moon's south pole; scientists hoped that the impact would throw off water vapor detectable from ground- and space-based observatories. No visible debris plume was detected, however.

Stardust. Launched in 1999, this probe (**Fig. 17**) was designed to fly through the cloud of dust that surrounds the nucleus of a comet and, for the first time, collect cometary material for return to Earth.

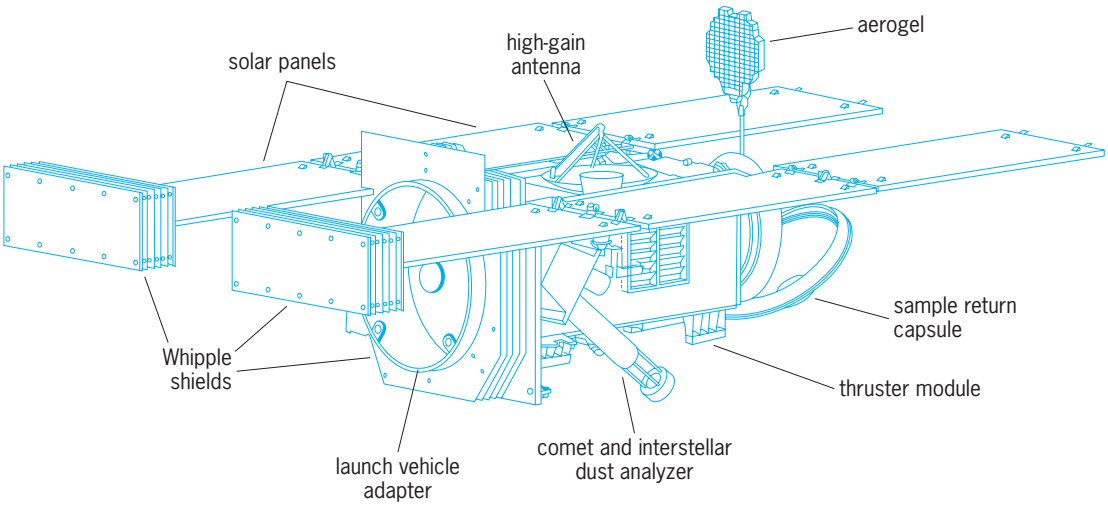

Fig. 17. ***Stardust*** **spacecraft, showing primary subsystems. (*JPL/NASA*)**

The spacecraft had a pair of tennis-racket-shaped collectors using a substance called aerogel to snag minuscule dust particles. During the course of the mission, the spacecraft collected dust streaming through the solar system from interstellar sources as well as dust thrown off by Comet Wild-2 during *Stardust*'s flyby on January 2, 2004. On January 15, 2005, the spacecraft returned to Earth, where a sample capsule descended by parachute to the salt flats of the Utah desert.

Genesis. Somewhat similar in concept to *Stardust*, *Genesis* was a mission designed to collect particles and return them to Earth. Launched August 8, 2001, *Genesis* traveled sunward to the Lagrange 1 point, where it spent more than 2 years collecting samples of solar wind particles. On September 8, 2004, the spacecraft brought its sample return capsule to Earth and put it on a trajectory to descend over the salt flats of Utah. The capsule's drogue parachute and parafoil did not deploy, however, and the capsule impacted the ground at a speed of 193 mi/h (311 km/h). Despite this setback, technicians found that much of the payload was intact, allowing the science team to fulfill its highest objectives, including the measurement of the isotopic ratios of oxygen and nitrogen in the solar wind.

Comet Nucleus Tour. The *Comet Nucleus Tour* (*Contour*) mission was designed to visit two diverse comets. The spacecraft was launched July 3, 2002, but 6 weeks later contact was lost after a maneuver intended to send the probe out of Earth orbit.

Messenger. Messenger—short for Mercury Surface, Space Environment, Geochemistry, and Ranging—was conceived as the first probe to orbit Mercury, the closest planet to the Sun. Mercury was previously studied at close range in the 1970s during three flybys by the *Mariner 10* spacecraft. *Messenger* was launched August 3, 2004, and was to fly by Earth once, Venus twice, and Mercury three times before entering orbit at Mercury in 2011.

Deep Impact. This novel mission concept called for a spacecraft to fly to a comet and dispatch a copper projectile that would excavate a substantial crater in the comet's nucleus, giving scientists a unique view of the nucleus' makeup. Launched January 12, 2005, *Deep Impact* encountered Comet Tempel 1 on July 4, 2005, delivering the impactor was planned. The collision—documented both by the penetrator before its demise and by the flyby spacecraft—threw off a large bright cloud of debris. Investigators were surprised to find that the comet nucleus was buried in a deep layer of powdery dust.

New Millennium Program. At about the same time that NASA established the Discovery Program to sponsor smaller planetary missions, it also created a program called New Millennium designed to launch missions flight-testing advanced technologies. The first New Millennium mission was *Deep Space 1*, a probe launched in 1998 to test an ion engine and 11 other technologies. After an initial problem in starting the propulsion system, the ion engine worked well, and the technology validation effort was viewed as highly successful, although a pointing problem prevented the probe's camera from obtaining useful photos during an asteroid flyby in 1999. *See* ION PROPULSION.

After the failure of the star tracker on *Deep Space 1*, the spacecraft's science camera was reprogrammed to maintain the craft's three-dimensional orientation in flight. After completing its primary mission of flight-testing the new space technologies, *Deep Space 1* executed a flyby of Comet Borrelly in September 2001.

Mars initiatives. Following the failure of *Mars Observer*, NASA quickly approved an orbiter mission, *Mars Global Surveyor*, to duplicate most of the scientific objectives of the lost spacecraft. In addition, the space agency established a program to send pairs of orbiters and landers to Mars during each biannual launch period, culminating in sample return missions in the early twenty-first century.

Mars Global Surveyor. Launched in 1996, *Mars Global Surveyor* spent 2 years circularizing its orbit using the technique of aerobraking, wherein the spacecraft

is slowed by frictional drag as it skims through the planet's thin upper atmosphere at each orbital periapsis. First tested at the end of *Magellan*'s mission at Venus in 1994, aerobraking allows planetary orbiters to use less propellant to carry out orbital insertion at the time of arrival. *Mars Global Surveyor* embarked on its main mapping mission in 1999, collecting high-resolution images and elevation data. Among early findings, scientists determined that the Martian magnetic field is not globally generated in the planet's core but is localized in particular areas of the crust. But the most startling science results came shortly before contact was lost with *Mars Global Surveyor* in 2006 following nearly a decade in orbit around the Red Planet. By comparing photos taken over a series of several years, scientists found evidence of bright new deposits in two gullies on Mars suggesting that liquid water flowed there briefly sometime during those years.

Mars Pathfinder. This mission was initially conceived as an engineering demonstration of a way to deliver a payload to Mars using the innovative technique of inflatable airbags for landing, and was subsequently equipped with an instrumented robotic rover to carry out a science mission. Launched in 1996, *Mars Pathfinder* (**Fig. 18**) landed on July 4, 1997, in Ares Vallis, an ancient floodplain that scientists selected because it was believed to contain a wide variety of rocks. The last successful data transmission from *Pathfinder* was completed on September 27, 1997. Both the lander and rover outlived their design lifetimes, the lander by nearly three times and the rover by twelve times. *Pathfinder* returned 2.3×10^9 bits of information, including more than 16,500 images from the lander and 550 images from the rover, as well as more than 15 chemical analyses of rocks and soil and extensive data on winds and other weather factors. Findings from the investigations carried out by scientific instruments on both the lander and the rover suggest that Mars was at one time in its past warm and wet, with water existing in its liquid state and with a thicker atmosphere.

Mars 96. During the same biannual launch period when *Mars Pathfinder* and *Mars Global Surveyor* were launched, Russia launched a complex spacecraft called *Mars 96*, which combined an orbiter and several landers. Originally named *Mars 94*, the mission was renamed due to a launch delay. The *Mars 96* probe failed to leave Earth orbit when booster rockets misfired, and reentered Earth's atmosphere.

Mars 98. Mars 98 consisted of the *Mars Climate Orbiter* and the *Mars Polar Lander*. The orbiter carried the balance of science instruments from the payload of the lost *Mars Observer* that were not reflown on *Mars Global Surveyor*. The lander, meanwhile,

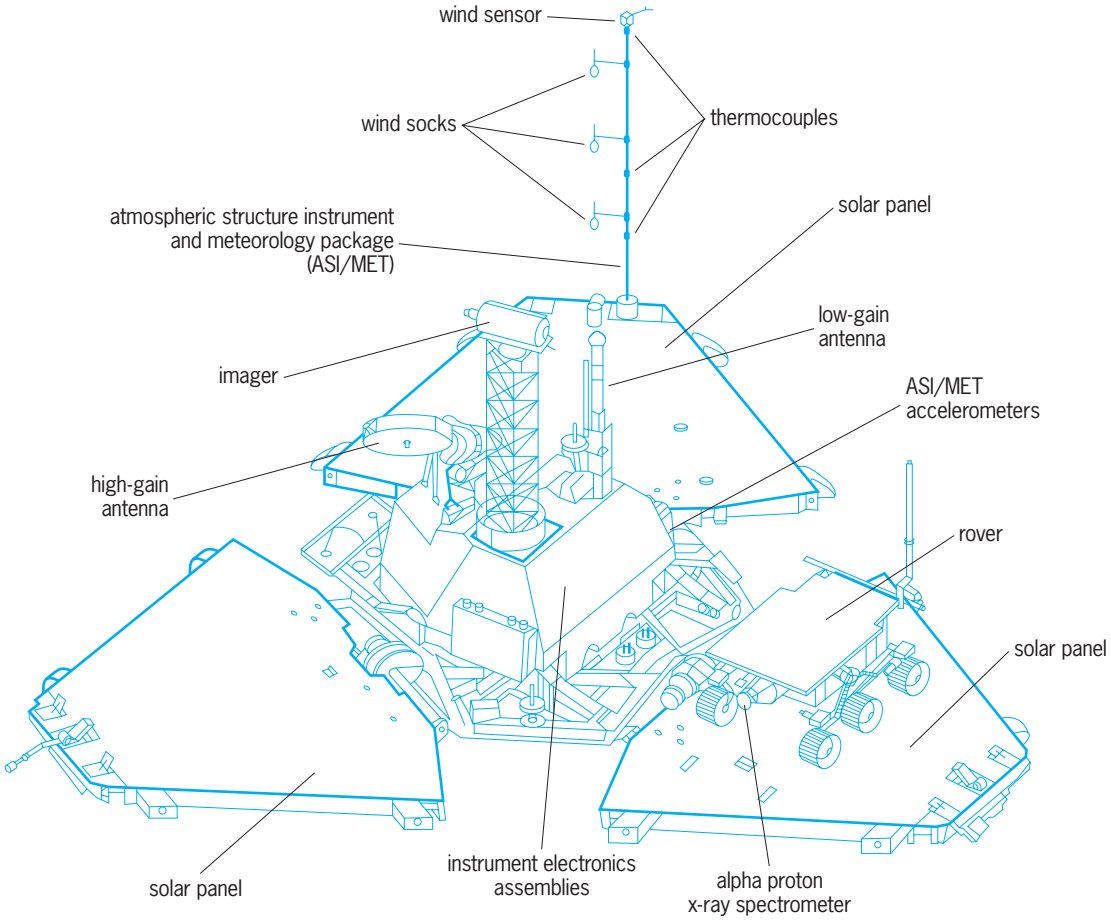

Fig. 18. ***Mars Pathfinder*, as deployed on Mars, showing primary subsystems. (*JPL/NASA*)**

was designed to be the first spacecraft to settle down near the planet's south pole to study the history of water on Mars. In addition, the lander carried two *Deep Space 2* microprobes, developed under NASA's New Millennium Program, designed to crash into the planet's surface independently and test for the presence of water ice as a technology demonstration.

Launched in December 1998, *Mars Climate Orbiter* was lost when it entered the Martian atmosphere too low and presumably burned up; a subsequent investigation established that the failure was caused because navigation-related data were not properly converted from U.S. Customary to metric units. *Mars Polar Lander* was launched in January 1999 and was also lost upon arrival. Although the cause of this failure was less certain, investigators concluded that flight software did not properly take into account the behavior of sensors on the spacecraft's landing legs, and as a result thrusters were probably cut off too early during the craft's final descent. No signal was received from the *Deep Space 2* microprobes, which were to be jettisoned just before atmospheric entry.

Mars Odyssey. A follow-up to *Mars Global Surveyor*, the *Mars* Odyssey spacecraft was an orbiter launched April 7, 2001, that reached the planet 6 months later. Aerobraking ended in January 2002, and the science mission commenced a month later. *Mars Odyssey* carries three science instruments, including one package designed to study the radiation environment in space and at Mars for the purpose of protecting future astronauts; it also serves as a communication relay for the Mars Exploration Rovers.

Mars Express/Beagle 2. For the 2003 Mars launch opportunity, the European Space Agency conceived an orbiter, *Mars Express*, equipped with a science payload including an imaging radar that could penetrate the planet's surface. In addition, *Mars Express* carried *Beagle 2*, a United Kingdom–built lander designed to search for signs of Martian life, past or present. Launched June 2, 2003, on a Soyuz-Fregat rocket from Kazakhstan, the orbiter reached the Red Planet on December 25, 2003. *Beagle 2* entered Mars's atmosphere the same day, but contact was never established with it following its descent. *Mars Express*, meanwhile, embarked on its orbital study, producing early results such as evidence of buried impact craters from the imaging radar.

Mars Exploration Rovers. During the same period in which Europe launched *Mars Express* and *Beagle 2*, the United States lofted two identical robotic rovers designed to explore Mars, *Spirit* and *Opportunity*. The travel capabilities of the rovers were much greater than those of *Sojourner*, the United States' previous Mars rover that landed in 1997. *Spirit* was launched June 10, 2003, and landed January 4, 2004, while *Opportunity* was launched July 7, 2003, and landed January 25, 2004. They touched down on opposite sides of the planet—*Spirit* in Gusev Crater and *Opportunity* in Merdiani Planum, both sites being viewed as promising for the study of rocks that could reveal past water activity on Mars. Despite a software problem that prevented *Spirit* from performing science for 10 days shortly after its landing, both rovers went on to travel around studying rocks for several years, far beyond their originally planned 90-day missions.

Mars Reconnaissance Orbiter. On August 12, 2005, the United States launched *Mars Reconnaissance Orbiter*, which reached the planet on March 10, 2006, and, following aerobraking, began its science mission in November. One of its instruments is a high-resolution camera that is the most powerful ever carried on a planetary mission. In addition, the orbiter is equipped with a subsurface radar designed to operate in conjunction with the radar on *Mars Express*. Soon after arrival, the science team released photos taken by the orbiter of several landers on the planet's surface, demonstrating the power of the high-resolution camera.

Other missions. Although the United States' Discovery and Mars programs account for many of the missions sent beyond Earth orbit in the early twenty-first century, several other projects have been launched under other sponsorships.

Wilkinson Microwave Anisotropy Probe. Intended to survey the sky at infrared wavelengths to measure heat left over from the big bang, the *Wilkinson Microwave Anisotropy Probe* was launched June 30, 2001. It orbits the Lagrange 2 point some 9×10^5 mi (1.5×10^6 km) from Earth. The mission was able to measure many cosmological parameters with higher accuracy than was previously possible; it puts the age of the universe, for example, at 13.7×10^9 years, plus or minus 200×10^6 years. *See* WILKINSON MICROWAVE ANISOTROPY PROBE.

Hayabusa. Known before launch as *Muses-C*, *Hayabusa* is a Japanese mission to collect samples from an asteroid and return them to Earth. The spacecraft was launched May 9, 2003, and approached the small near-Earth asteroid 25143 Itokawa in September 2005. It did not enter orbit around the asteroid, but took up a heliocentric orbit nearby. After making observations of the object, the spacecraft prepared for a series of touchdowns to collect samples. It was uncertain whether the sample collection was carried out as planned, but the spacecraft was believed to have probably obtained at least some dust in its sample collector. The spacecraft is scheduled to deliver a reentry capsule in June 2010 to Earth, where it will land via parachute in Australia.

Spitzer Space Telescope. Similar in concept to a previous Earth-orbiting infrared observatory, the *Spitzer Space Telescope* was launched August 25, 2003, and occupies an Earth-trailing heliocentric orbit, drifting away from Earth at about 0.1 AU per year. The telescope features a 33.5-in. (85-cm) mirror and three instruments that allow it to conduct a variety of photometric, spectroscopic, and spectrophotometric studies. *See* SPITZER SPACE TELESCOPE.

Smart 1. Smart 1 (the name derives from Small Missions for Advanced Research in Technology) was a European spacecraft that orbited the Moon.

The mission's primary objective was testing of a solar-powered ion thruster. Launched September 27, 2003, from French Guiana, *Smart 1* spent 13 months in an orbit gradually spiraling outward from Earth, shaped by the ion propulsion. In November 2004 the ion thruster was activated to decelerate it into orbit around the Moon—a maneuver that was expected to take $4^1/_2$ days but which required approximately 3 months. It was intentionally crash-landed onto the Moon on September 3, 2006.

Rosetta. The European *Rosetta* probe was launched March 2, 2004, on a mission to rendezvous with Comet 67P/Churyumov-Gerasimenko more than 10 years later, in May 2014. En route, the spacecraft is scheduled to fly by Earth three times and Mars once, in addition to a flyby of the asteroid 2867 Steins in 2008. *Rosetta* carries a lander named *Philae*, which will use harpoons and drills to secure itself to the surface of the comet nucleus.

Venus Express. Another European mission, *Venus Express* was launched November 9, 2005, and entered orbit around Venus on April 11, 2006. The purpose of the mission is to make detailed studies of the planet's atmosphere, clouds, plasma environment, and surface characteristics.

New Horizons. During the 1990s, NASA began conducting studies of how to carry out a mission to Pluto, which was then regarded as the solar system's only planet that had not been visited by a robotic spacecraft. The mission concept that was eventually approved and moved forward was *New Horizons*. Launched January 19, 2006, *New Horizons* was to fly by Jupiter in 2007 and fly by Pluto in July 2015, continuing on and out of the solar system. One or two encounters with Kuiper Belt objects may be possible in the event of an extended mission. A few months after the launch, the International Astronomical Union reclassified Pluto as a dwarf planet. *See* KUIPER BELT; PLUTO.

Solar Terrestrial Observatory. The *Solar Terrestrial Observatory* (*Stereo*), is a mission designed to study the Sun. Two identical spacecraft were launched October 26, 2006, into highly elliptical geocentric orbits with an apogee at about the distance of the Moon. After a few orbits, they were to pass close to the Moon, where a gravity-assist effect would send them into heliocentric orbits inside and outside the Earth's orbit. The dual spacecraft allow stereoscopic imaging of the Sun and solar phenomena such as coronal mass ejections.

Trends. Missions of the early twenty-first century demonstrate ongoing refinement in a number of areas. One area that has seen considerable development is that of increasing on-board autonomy. For example, the Mars Exploration Rovers are able to execute simple decision-making about their environments, and how to carry out science observations that makes operations proceed more quickly than if every move was commanded from Earth. In addition, newer spacecraft have great capabilities in their anomaly response. The fault protection system of an older spacecraft might direct it simply to enter a protective or "safing" mode in response to unexpected conditions. Newer spacecraft are programmed to carry out additional diagnostics and take corrective actions.

Another new initiative is in the area of in situ science instruments. The development of greatly miniaturized microprocessors, memory, and sensors has enabled the creation of more sophisticated science instruments. Many of these are designed to search for evidence of exobiological processes on Mars and other bodies. Franklin O'Donnell

Bibliography. J. K. Beatty, C. C. Petersen, and A. Chaikin, *The New Solar System*, 4th ed., Cambridge University Press, 1999; W. E. Burrows, *Exploring Space: Voyages in the Solar System and Beyond*, 1990; D. Doody and G. Stephan, *Basics of Space Flight*, JLP D-9774, 1993; W. T. Huntress, Jr., Grand challenges for space exploration, *Space Times*, American Astronautical Society, May-June 1999; Jet Propulsion Laboratory, *Cassini Mission Plan*, JPL D-5564, 1994; D. A. Mendis and R. T. Tsurutani, The spacecraft encounters of Comet Halley, *Eos*, 67(20):478–481, 1986; B. C. Murray, *Journey into Space: The First Three Decades of Space Exploration*, 1989; F. O'Donnell, *Mariner 2* and beyond: Planetary exploration's first 25 years, *IEEE Aerosp. Electr. Sys.*, 2(8):2–6, August 1987; M. D. Rayman, The successful conclusion of the *Deep Space 1* mission, *Space Technol.*, 23(2–3):185, 2003; A. Rosenthal, *Satellite Handbook: A Record of NASA Space Missions 1958–1980*, Goddard Space Flight Center, 1981; Space Studies Board, *New Frontiers in the Solar System: An Integrated Exploration Strategy*, National Research Council, 2003; E. C. Stone, *Mars and the Search for Life Elsewhere: Innovations in the Third Era of Space Exploration*, American Institute of Aeronautics and Astronautics, 1999.

Space processing

Experiments conducted in space in order to take advantage of the reduced gravity in studies of the growth, behavior, and properties of materials. Spacelab, developed by the European Space Agency (ESA), was a laboratory module that flew in the space shuttle payload bay. First launched in 1983 on the *Columbia*, the module became the workhorse for United States and international science missions emphasizing low gravity. Twelve such missions were flown, and an additional seven missions utilized Spacelab components (such as a pallet to support experiments not requiring direct crew interaction). The experiments in fluids, combustion, materials science, and biotechnology conducted on these missions, together with their related ground-based research, resulted in more than 2200 scientific publications, about 1200 of which appeared in peer-reviewed journals.

The Spacehab laboratory module, which also flies in the shuttle payload bay, was privately developed by Spacehab, Inc., beginning in 1993. A number

of commercial experiments have been conducted on the 21 Spacehab missions (as of January 2007), several of which carried experiments to the Russian *Mir* space station.

Biotechnology. Since biotechnology experiments generally have modest space and power requirements, they were able to be accommodated by middeck lockers in other shuttle missions as well as on the *Spacelab* microgravity emphasis missions, which provided them with many more flight opportunities. Even though the microgravity environment on some of these missions was not ideal, biomolecular crystal growth experiments for structural analysis as well as cell and tissue culturing experiments produced many interesting results, some of which could have commercial applications.

Protein crystal growth. It has now been established beyond question that at least some proteins and other macromolecular systems form higher-quality crystals in a microgravity environment. A survey of 221 macromolecular growth systems flown on 19 shuttle missions found that 57.1% produced crystals large enough for diffraction analysis, and 27.2% (or almost half of those that grew large enough for diffraction analysis) showed improvement in the diffraction data compared to the best crystals ever grown on the ground by any means. One criterion used for selecting growth systems for flight was the ability to grow diffraction-sized crystals in the flight apparatus in a time commensurate to the time available on orbit. Since crystals growing under diffusion-limited transport in space would be expected to grow more slowly than those under convective enhanced transport on the ground, there is little question that many more of the space-grown crystals would have grown to a size large enough for x-ray analysis if more time had been available. This supposition has been confirmed by crystal growth experiments that were carried to the *Mir* station. Even though conditions on *Mir* were not ideal for crystal growth, several of these growth systems have produced centimeter-sized crystals that are large enough for neutron diffraction analysis. (Neutron diffraction can directly reveal the positions of hydrogen atoms, whereas the positions of these atoms must be inferred from the positions of heavier atoms seen by x-ray diffraction.) Crystal size limitations have restricted the number of protein structures that have been determined by neutron diffraction analysis to fewer than a dozen. *See* CRYSTAL GROWTH; CRYSTAL STRUCTURE; NEUTRON DIFFRACTION; X-RAY CRYSTALLOGRAPHY; X-RAY DIFFRACTION.

In some cases the improvement in diffraction resolution seen in the space-grown crystals has been key to solving the structure of a particular protein for the first time, and in other cases the space-grown crystals contributed significantly to the refinement of the structure of certain proteins. Several drugs based on structures in which space-grown crystals played a part are in clinical trials.

Cell culturing experiments. A number of experiments have demonstrated that cell metabolism and biosynthesis of cell products are altered in microgravity. For example, the analysis of the essential oils produced by a rose flower on orbit found products that are not normally found in normal gravity. Several experiments have also shown that gene regulation is significantly altered on exposure to microgravity. Included are genes that regulate cell-cycle, apoptosis, heat shock response, metabolism, structural proteins, cancer cell growth, intracellular signaling, mitochondrial proteins, and a number of transcription factors and DNA binding proteins that are either up-regulated or down-regulated. In particular, it was found that a specific cell-cell signaling matrix material, important to the layering of cells during heart development, was up-regulated in microgravity. This information could be important to the development of tissue-engineered heart patches to replace damaged cardiac tissue. *See* BIOMEDICAL ENGINEERING; CELL (BIOLOGY); TISSUE CULTURE.

Normal cells do not tend to form three-dimensional structures in cultures in vitro. However, a new rotating bioreactor developed by NASA mimics one aspect of microgravity by keeping cells in suspension as they continuously fall through their culture medium. This device has been used on the ground by a variety of investigators at the NASA/NIH Center for Three-Dimensional Tissue Culture and at several universities to study a broad spectrum of biomedical research issues. However, it has also been found that the shear forces on the cells in the rotating reactor often inhibit their attachment to the three-dimensional assembly. This was justification for flying the rotating bioreactor on several shuttle flights to the *Mir* space station. In microgravity, the rotation rates required to maintain the cells in suspension are orders of magnitude smaller than on Earth. Techniques developed to provide nutrients to the culture and remove waste products were successfully demonstrated, and several cultures were carried out which had proven difficult in normal gravity.

Combustion experiments. There are two compelling reasons for the study of combustion in microgravity. One is the issue of fire safety in the design and operation procedures of orbiting laboratories; the other is to take advantage of the weightless state to study certain combustion phenomena in more detail and to test various models in which convection has been ignored in order to be mathematically tractable. Examples of the latter are a number of droplet combustion experiments in which the burning of free-floating or tethered droplets was studied. The absence of gravity allowed the droplet size to be increased to as much as 5 mm (0.2 in.) so that more detailed observation could be made. The objective is to test theories of droplet combustion and soot formation that are of importance to improving the efficiency of internal combustion engines, gas turbine engines, and home and industrial oil-burning heating systems. *See* COMBUSTION; GAS TURBINE; INTERNAL COMBUSTION ENGINE.

In an interesting variation of this experiment, a container was filled with various combustible mixtures near their lean limit of combustion (the lowest concentration of fuel-to-air mixture that will ignite

and/or support combustion). A flame ball was created by an electrical spark. A stationary spherical flame front develops as fuel and oxygen diffuse into, and heat and combustion products diffuse out of, the flame ball. This is the simplest possible geometry in which to study the chemical reactions and the heat and mass transport of lean combustion processes. It is expected that these experiments will provide new insight into combustion processes in the lean burning limit, which are important in improving the efficiency of engines and heating systems.

Materials science. Processing of metallic alloys and composites has been carried out in space in order to study their microstructure, thermal properties, and crystal growth.

Evolution of microstructure. The strength and other properties of an alloy depend on its microstructure, which is characterized by the size, orientation, and composition of the grains that make up the solid. One of the main tasks of a materials scientist is to design solidification processes to produce the microstructure that will give the material the desired properties. With current computational capabilities, it is possible to design a complex mold so that the heat and mass flow will produce the desired microstructure throughout the final casting. However, in order to do this the basic laws governing the development of the microstructure must be known, along with the thermophysical properties of the components. Establishing the physical basis for the various laws that describe the solidification process has, over the last half-century, transformed metallurgy from an industrial art based on empiricism to a more exact science. Because of the complicating effects of convection, many of the laws in use today are based on theories that (for simplicity's sake) assume no convective flows. The effects of convection are then added to the calculation to make the result somewhat more realistic. However, most of these theories have never been tested in the absence of convection, so various subtleties may have been overlooked. The ability to experiment in microgravity provides an opportunity to test some of these theories to make sure they are basically correct. Later Spacelab experiments investigated theories that predict the lamella spacing in eutectic structures, the growth rate of dendrites in castings, the coarsening and particle agglomeration in composites formed by liquid-phase sintering, the particle-solidification interactions and ripening in alloys with a dispersed second phase, and the phase-separation mechanisms that act in monotectic alloys (systems which exhibit liquid phase immiscibility). *See* ALLOY; COMPOSITE MATERIAL; CONVECTION (HEAT); MATERIALS SCIENCE AND ENGINEERING; METAL CASTING.

Measurement of thermal properties. A number of self-diffusion and binary diffusion experiments in liquid metal systems have been carried out on the Spacelab missions. It was found that diffusion coefficients measured in space were generally 30–50% lower than the accepted values based on ground-based experiments carried out in capillaries to minimize convection. This would suggest that most, if not all, diffusion measurements in normal gravity have been contaminated to some degree by convection. Furthermore, the increased precision of the space data allowed an accurate determination of the temperature dependence of diffusion in liquid metal systems. Surprisingly, instead of following the Arrhenius behavior typical of diffusion in solids, it was found that the diffusion data appear to fit a power law in which temperature is raised to exponents ranging from 1 to 2, depending on the system. This unexpected power law behavior of diffusion may shed some light on the structure of the liquid state of metals, which is still poorly understood. *See* DIFFUSION.

Conducting materials can be levitated by electromagnetic induction in normal gravity, but the current required for levitation also heats the sample. In microgravity, the levitation and heating may be decoupled, which allows samples to be deeply undercooled in the liquid state with no physical contact that would induce the nucleation of the solid phase. Thus it becomes possible to measure thermophysical properties of liquids well below their normal freezing temperatures, which are important in developing new metallic glasses and other novel nonequilibrium solids. Several very clever techniques have been developed for measuring the temperature, surface tension, viscosity, and electrical resistivity of levitated samples with no physical contact using the TEMPUS facility developed by the Institute for Space Simulation in Cologne, Germany. These properties of several metallic glass–forming systems were measured in the undercooled state with this apparatus. One particularly interesting result was evidence for magnetic ordering of ferromagnetic $Co_{80}Pd_{20}$ in the undercooled liquid state near the Curie temperature of the solid, which demonstrates that ferromagnetism can indeed exist in the liquid state. *See* ELECTROMAGNETIC INDUCTION; FERROMAGNETISM.

Crystal growth experiments. On four Spacelab missions was flown a sophisticated solidification experiment apparatus developed by the French Space Agency (CNES) and the French Atomic Energy Commission (CEA) called Materials for the Study of Interesting Phenomena of Solidification on Earth and in Orbit (MEPHISTO). It was used by French and American investigators to study the stability and breakdown of the solidification interface of metal alloy systems. The Seebeck voltage generated by the melt-solid interfacial was measured, which served as an indicator of incipient interfacial breakdown. The results were used to test and refine theories of interfacial stability of both faceting and nonfaceting alloy systems. Such theories are important in developing crystal growth processes for alloy-type semiconductor and photonic systems in which interfacial breakdown must be avoided. *See* ALLOY STRUCTURES; SEEBECK EFFECT; SEMICONDUCTOR; THERMOELECTRICITY.

A number of attempts were made to grow a variety of compound and alloy-type semiconducting systems from the melt with the aim of improving the compositional uniformity through diffusion-controlled transport. This goal has proven to be elusive, even in microgravity, which underscores the

extreme sensitivity of such systems to buoyancy-driven flows even at acceleration levels on the order of a micro-*g*. It was found, however, that dislocations and twinning were significantly reduced in many of the space experiments, especially in regions where the forming solid was not in contact with the wall of the growth ampoule. This observation would implicate wall effects as a major source of imperfections in terrestrially grown crystals, and has encouraged some crystal growers to try to develop growth methods that avoid or least minimize wall contact. *See* CRYSTAL DEFECTS.

Several semiconductor systems were grown in space by the floating zone process to eliminate wall contact. The absence of hydrostatic pressure in space allows much larger floating zones to be deployed than are possible under normal gravity. Well-ordered crystals of gallium arsenide (GaAs) with diameters up to 20 mm were grown using a mirror furnance to focus light from an arc lamp onto the molten zone. The dislocation densities were very low; however, dopant striations attributed to nonsteady Marangoni convection (convective flows driven by surface tension gradients) were observed. Such striations were avoided in other growth systems that had a native oxide skin which served to suppress the Marangoni flows. *See* ZONE REFINING.

Robert J. Naumann

Bibliography. P. A. Curreri and D. E. McCauley (eds.), Second United States Microgravity Payload: One Year Report, *NASA Tech. Mem.*, no. 4737, 1996; P. A. Curreri, D. E. McCauley, and C. Walker (compilers), Third United States Microgravity Payload: One Year Report, *NASA Tech. Mem.*, no. 4737, 1996; P. Downey (ed.), Life and Microgravity Spacelab (LMS) Final Report, *NASA Conf. Publ.*, NASA CP-1998–206960, 1998; NASA Microgravity Research Program 1998 Annual Report, *NASA Tech. Mem.*, TM-1999–209757, 1999; N. Ramachandran et al. (eds.), Proceedings of the Joint Launch + One Year Science Review of USML-1 and USMP-1 with the Microgravity Measurement Group, *NASA Conf. Publ.*, no. 3272, 1994; M. B. Robinson (compiler), Microgravity Science Laboratory (MSL-1) Final Report, *NASA Conf. Publ.*, NASA/CP-1998–208868, 1998; P. R. Sahm et al. (eds.), *Proceedings of the Symposium on Scientific Results of the German Spacelab Mission D-2*, Norderney, March 14–16, 1994 (available from the European Space Agency); R. Snyder (compiler), Second International Microgravity Laboratory (IML-2) Final Report, *NASA Ref. Publ.*, no. 1405, July 1997; *Spacelab Science Results Study, Final Report*, vol. 2: *Microgravity Science*, Contract no. NAS8-97095, Task no. H-30194D, August 18, 1999 (available from George C. Marshall Space Flight Center, Huntsville, AL 35812).

Space shuttle

A reusable crewed orbital transportation system. The space shuttle, along with crewless (robotic) expendable launch vehicles such as Atlas and Delta, make up the United States Space Transportation System (STS). The shuttle provides the unique capability for in-flight rendezvous and retrieval of faulty or obsolescent satellites, followed by satellite repair, update, and return to orbit or return to Earth for repair and relaunch. The space shuttle also plays an essential continuing role in the construction and provisioning of the *International Space Station* (*ISS*) by transporting major components, such as the giant solar-cell arrays and the Canadian computer-driven double-ended robotic arm, from Earth to the *ISS* and installing them using extended extravehicular activity (EVA, or space walks) by trained *ISS* resident and shuttle-visiting astronauts. *See* SATELLITE (SPACECRAFT); SPACE STATION.

Early in its history the space shuttle was expected to be a low-cost replacement for expendable launch vehicles. Following the *Challenger* shuttle accident in 1986, it became clear that crewless vehicles would have a continuing place in the United States launch vehicle fleet, that they have advantages over the shuttle such as lower cost and shorter lead time for tasks within their capability, and that the shuttle fleet would be fully occupied for much of the first decade of the twenty-first century doing complex jobs requiring human presence, in particular associated with the *ISS*, which no other launch system can do. For such reasons the STS was expanded to include the expendable launch vehicle families in use by the Department of Defense (DOD) and the National Aeronautics and Space Administration (NASA).

Prior to permanent occupation of the *ISS*, the shuttle carried attached payloads, such as the European Space Agency (ESA) Spacelab module, which made use of the low-gravity environment for mission durations of up to 10 days or so. Such experiments are now almost always flown on the *ISS* to take advantage of the longer microgravity duration implicit in the *ISS* "continuing human presence in space" philosophy.

The space shuttle orbiter accommodates a crew of four to seven for orbital mission durations up to about 10 days. The space shuttle flight system (**Fig. 1**) consists of the orbiter, which includes the three liquid-fueled shuttle main engines; an external fuel tank; and two solid-fuel strap-on booster rockets. The external tank is discarded during each launch. The solid-fuel rocket casings are recovered and reused. The orbiter lands horizontally as an unpowered ("dead stick") aircraft on a long runway at the NASA Kennedy Space Center in Florida or, when conditions for landing these are unacceptable, at the Edwards Air Force Base in California. *See* ROCKET PROPULSION.

Flight sequence. A nominal flight sequence is shown in **Fig. 2**. The shuttle is launched with all three main engines and both strap-on solid-fuel booster rockets burning. The booster rockets are separated 2 min after liftoff at an altitude of 30 mi (48 km), 28 mi (44 km) downrange. Main engine cutoff occurs about 8 min after liftoff, nearly 70 mi (110 km) above the Atlantic Ocean, 890 mi

(1430 km) downrange from Kennedy. The external tank (ET) is separated from the orbiter shortly after the main engine cutoff, before orbital velocity (speed) is reached, so that the relatively light empty tank can burn up harmlessly during its reentry into the atmosphere above the ocean. Main propulsion for the rest of the mission is provided by the orbital maneuvering system (OMS) engines, whose first task is to complete insertion of the shuttle into its final, nearly circular path during its first orbit (flight once around Earth) after liftoff.

After the shuttle orbiter completes the orbital phase of its mission and is ready to return to Kennedy, its pilot rotates the spacecraft 180° (tail-first relative to the orbiter's direction of motion) and fires the orbital maneuvering system engines to decrease the space vehicle's speed. This maneuver reduces the orbiter's speed just enough to allow the orbiter to fall into the tenuous outer atmosphere, where atmospheric drag causes the orbit to decay (altitude and velocity decrease) along a spiral path. This reentry slowdown occurs in a precise, computer-controlled manner so that the landing may be made manually, halfway around the Earth, on the desired runway at Kennedy or Edwards. At orbital altitudes, small reaction-control engines (jets) maintain the desired orientation of the orbiter. As the spacecraft-becoming-airplane descends into the denser lower air, its aerodynamic control surfaces gradually take over their normal maintenance of heading, pitch, and roll, and the small jets are turned off. Landing is made at a speed of 220 mi/h (100 m/s). The brakes are helped to bring the vehicle to a stop by a parachute (drag chute) which is deployed from the tail after touchdown of the nose wheel and released well before the vehicle comes to a complete stop.

Vehicle data. The orbiter carries into space the crew and the payload for the space shuttle vehicle system. It is comparable in size to a medium commercial jet transport, with a length of 121 ft (37 m) and a wing span of 78 ft (24 m). Its dry (unfueled) mass is 161,400 lb (73,200 kg). The orbiter can deliver cylindrical payloads of up to 65,000 lb (29,500 kg) of mass, with lengths as great as 60 ft (18 m) and diameters of up to 15 ft (4.6 m).

The external tank is 157 ft (48 m) long by 28 ft (7.6 m) in diameter and contains a combined mass of 1,550,000 lb (703,000 kg) of available liquid oxygen (LOX) and liquid hydrogen propellants. The solid-fuel booster rockets are each 149 ft (45.4 m) long by 12 ft (3.7 m) in diameter and have a loaded mass of 1,292,600 lb (586,200 kg). Together, the three space shuttle main engines and the two solid-fuel rocket boosters generate a thrust of 6,400,000 lbf (28,500,000 newtons) at liftoff. The total mass of the shuttle system at liftoff is 4,491,000 lb (2,037,000 kg).

Orbiter vehicle systems. Major orbiter systems are the environmental control and life support, electric power, hydraulic, avionics and flight control, and the space shuttle main engines. Each system is designed with enough redundancy to permit either continuation of mission operations or a safe return to Earth after any single-element failure. For example, three fuel-cell power plants generate the in-flight electric power. Each fuel cell feeds one of three independent electrical distribution buses. Similarly, three independent hydraulic systems, each powered by an independent auxiliary power unit, operate the aerosurfaces. *See* FUEL CELL; SPACE POWER SYSTEMS.

Fig. 1. Space shuttle lifting off from the Kennedy Space Center.

The avionics system uses five general-purpose computers, four of which operate in a redundant set while the fifth operates independently. These computers perform guidance, navigation, and control calculations, and operate the attitude (orientation) controls and other vehicle systems. They also monitor system performance and can automatically reconfigure redundant systems in the event of a failure during flight-critical mission phases. Each of the three space shuttle main engines has an independent computer for engine control during ascent. *See* DIGITAL COMPUTER; MULTIPROCESSING.

The thermal protection system, which is not redundant, presented a major technical challenge

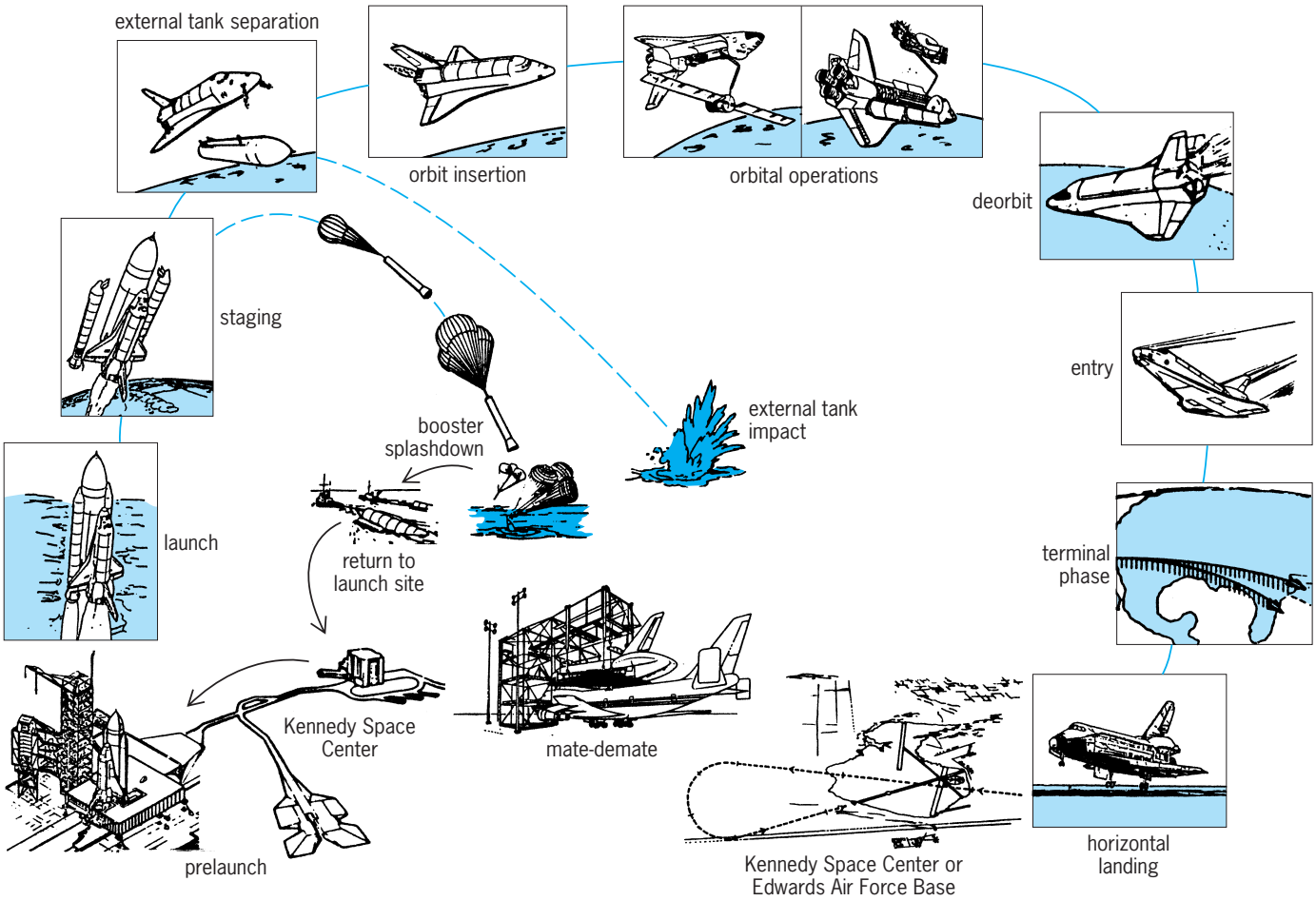

Fig. 2. Typical space shuttle mission profile.

to the orbiter development schedule. This system, unlike those of previous single-use spacecraft, has a design requirement of reuse for 100 missions. Performance requirements also dictate that the system withstand temperatures as high as 3000°F (1650°C) while maintaining the vehicle's structure at no more than 350°F (177°C). *See* ATMOSPHERIC ENTRY.

The key to meeting this challenge was to develop a material possessing an extremely low specific heat capacity and thermal conductivity, together with adequate mechanical strength to withstand the launch and reentry vibration and acceleration. These thermal characteristics must remain stable up to temperatures of at least 3000°F (1650°C). The solution was found in silica ceramic tiles, some 24,000 of which cover most of the orbiter's surface. During the highest thermal load portion of the reentry trajectory (path), the orbiter's wings are level and the nose is elevated well above the flight path. This causes the temperature of the undersurface to be substantially higher than that of the upper surface. For this reason, the undersurface is covered with black borosilicate glass-coated high-temperature tiles. Most of the upper surface of the shuttle is covered with a lower-temperature silica blanket made of two outer layers of woven fabric and an insulating buntinglike center layer stitched together in a quiltlike pattern. Finally, the nose cap and wing leading edges are subject to the highest (stagnation) temperatures, which requires the use of molded reinforced carbon-carbon material.

Flight testing. Like any commercial or military transport aircraft, the space shuttle system had to pass a series of flight tests before it could be committed to operational service. These tests were conducted in two phases: approach and landing, and orbital flight.

For the approach and landing phase, which tested the capabilities of the shuttle as an unpowered aircraft, the first test vehicle, *Enterprise*, was air-launched from a modified Boeing 747 airplane which was used later to airlift ("piggyback") operational shuttle orbiters between California and Florida. Five successful flights in 1977 verified the subsonic (low-speed) aerodynamics and handling qualities of the orbiter.

In 1981 the first protoflight powered space shuttle *Columbia* was launched from Kennedy on the first of four orbital test missions, STS 1 through STS 4. All major test objectives were successfully accomplished. The most significant problem that surfaced involved damage to *Columbia*'s thermal tiles, but the orbiter's underlying structure suffered no

temperature problems and the repaired areas survived subsequent flights without further thermal difficulties.

Flight operations. STS 5 initiated the operational shuttle flight phase in 1982. Many tests of the orbiter remained to be accomplished and were scheduled in conjunction with operational flights. The ability to do deployment, rendezvous, proximity operation, and retrieval using the remote manipulator system (robotic arm), was demonstrated in 1983 using the West German *Shuttle Pallet Satellite Spacecraft* (*SPAS*). One year later the first shuttle landing took place on Kennedy's 3-mi (5-km) runway. Longer-term problems were found and fixed. For instance, an unexpected chemical reaction with the tile waterproofing agent caused degradation of the tile bonding adhesive, requiring replacement of the degraded material and development of a new waterproofing agent. In 1985 the planned four-shuttle fleet was completed with the delivery of the orbiter *Atlantis*.

Payload accommodations. The orbiter payload bay is divided into four cylindrical sections, each of which can support one major payload. Standard services to each section include electric power, data recording, and two-way telemetry (data exchange). Optional services, such as active payload cooling, may also be purchased by the customer.

Small experiments for researchers are accommodated by a variety of carriers. These are normally manifested on a space-available basis and flown at a relatively modest cost. Self-contained experiments requiring only turn-on and turn-off by the crew are carried in canisters mounted along the forward portion of the payload bay sidewall. They are known informally as "get-away special" (GAS) canister (CAN) payloads. Small experiments can also be acommodated in the middeck section of the pressurized cabin. Larger experiments are mounted on a carrier in the payload bay, making orbiter power and telemetry services available.

Experiments requiring precision pointing can be mounted on a free-flying carrier that is deployed and retrieved with the robotic arm. The carrier platform provides highly accurate pointing and a microgravity environment.

Figure 3 shows a typical collection of payloads and experiments to illustrate the diversity of purposes which have been served on each shuttle flight. Specifically, it shows the STS 61-B cargo configuration, which was made up of three communications satellites: the Australian *AUSSAT 2*; the Mexican *MORELOS-B*; and the United States commercial *SATCOM KU-2*. The attached payloads are Experiment Assembly of Structure in Extravehicular Activity (EASE), Assembly Concept for Construction of Erectable Space Structure (ACCESS), IMAX Cargo Bay Camera (ICBC), and G-479 Telesat-Canada Getaway Special (GAS). In the crew compartment are the Continuous Flow Electrophoresis System (CFES), Diffusive Mixing of Organic Solutions (DMOS), Morelos Payload Specialist Experiments (MPSE), and Orbiter Experiments (OEX).

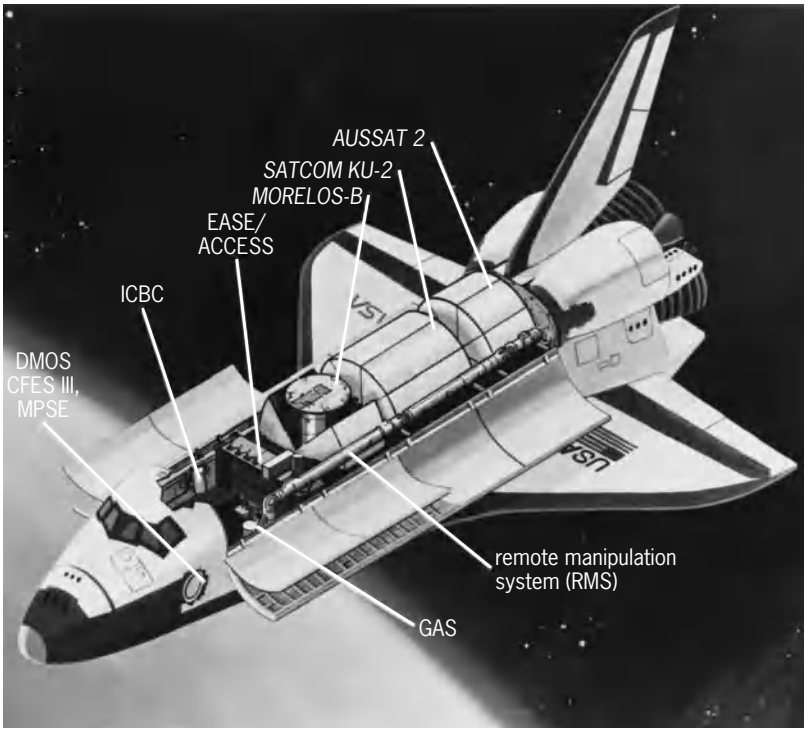

Fig. 3. STS 61-B cargo configuration. Acronyms are explained in the text.

It should be noted that, were this mission (which was launched in November 1985) planned after 1986, probably each of the communications satellites would be handled by an expendable launcher, and after 2000 many of the experiments would be done on the *ISS*.

Challenger disaster. Shortly before noon on January 28, 1986, *Challenger* lifted off from Kennedy Space Center on the twenty-fifth shuttle mission. At 72 s into the flight, with no apparent warning from real-time data, the external tank exploded and *Challenger* was destroyed.

All seven crew members perished, the first to die in NASA's human space-flight program spanning 25 years and 56 crewed launches. Further flight operations were suspended while a Presidential (Blue-ribbon) Commission reviewed the circumstances surrounding the accident, determined its probable cause, and developed recommendations for corrective actions.

Recovery. Nearly 5 months later the Commission submitted its report which contained recommendations to be implemented before returning the space shuttle to flight status.

The cause of the accident was determined to be inadequate design of the solid rocket motor field joint. "Field joint" refers to the assembly of the solid rocket motor at Kennedy rather than at the manufacturer's factory, because the massive motor must be shipped in cylindrical sections. Deflection of the joint with deformation of the seals at cold temperature allowed hot combustion products to bypass both O-rings, resulting in erosion and subsequent failure of both primary and secondary seals.

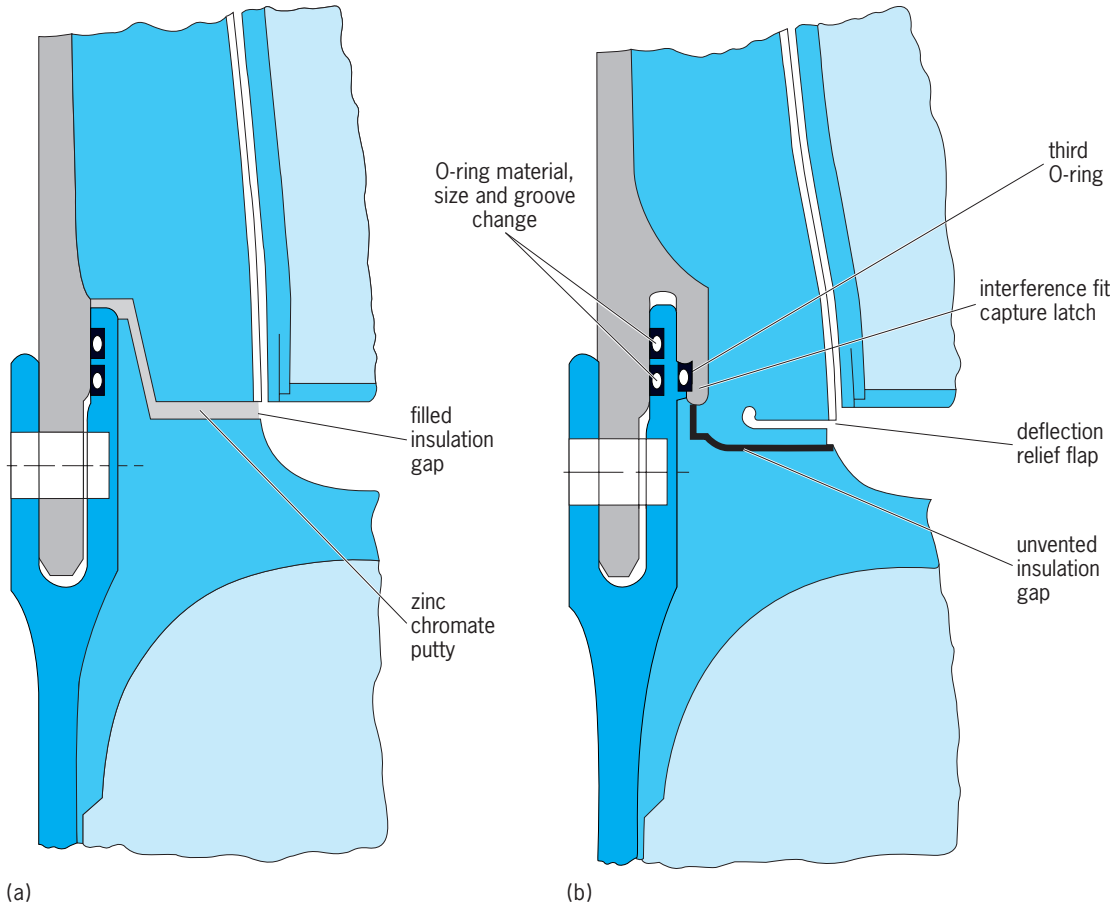

Fig. 4. Solid-fueled rocket motor field joint. (*a*) Original design. (*b*) New design.

The redesigned field joint (**Fig. 4**) has an overlapping tang that helps prevent separation during the startup phase of motor operation when the case is responding dynamically to the internal pressure buildup. A third O-ring seal, in the overlapping tang, was also included in the new design, along with the ability to verify the seal's integrity after installation. Other improvements were also made in the design of the solid rocket motor.

NASA completely reviewed all activities associated with launch aborts and established an additional transatlantic abort landing site at Ben Guerir, Morocco. An escape system was devised that allows the crew to parachute out of the orbiter once the vehicle has achieved controlled gliding flight. The system uses a pole extended out of the orbiter side hatch (**Fig. 5**).

NASA also modified its shuttle program management structure and established an Office of Safety, Reliability, Maintainability, and Quality Assurance, reporting directly to the NASA Administrator.

Return-to-flight actions. Before resuming space shuttle flights, in addition to actions mandated by the Presidential Commission, NASA initiated a Systems Design Review (SDR), a detailed review of all shuttle program hardware and software. Its objective was to establish a series of changes that would result in improved safety, with reliability and efficiency being of second and third priority, respectively. As a result, 254 modifications were made to the orbiter, 24 to each main engine, 30 to each solid rocket booster, and 13 to the external tank.

Because the solid rocket cases are recovered and refurbished, it was known that the primary (but not the secondary) O-ring had failed on several earlier flights. For human space flight systems, a basic requirement of redundancy for safety, which includes the field joint O-rings, is that should the designed redundancy be known or suspected of being compromised, the mission must be delayed until the redundancy is known to be restored. This requirement was waived by NASA for several shuttle missions, including *Challenger*'s final flight.

On September 29, 1988, 32 months after *Challenger*'s final flight, space shuttle *Discovery* successfully resumed shuttle operations. During the next 14 years, more than 80 missions were flown with no anomalies more serious than several foreshortened missions because of the failure of a redundant component or subsystem.

Columbia disaster. On February 3, 2003, while the shuttle *Columbia* was reentering the atmosphere, it disintegrated. All seven crew members perished, and pieces of the orbiter fell in a long swath over East Texas. As much of the debris as possible was collected and taken to Kennedy Space Center, where pieces were placed in their appropriate position on a hanger floor. The Columbia Accident Investigation

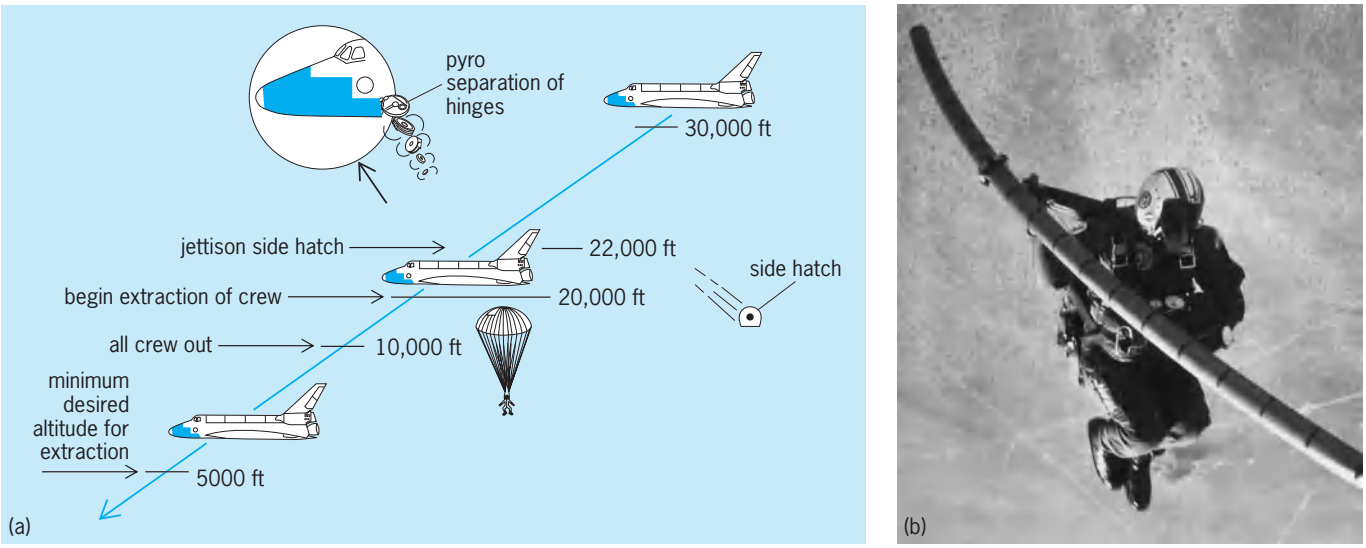

Fig. 5. Orbiter crew escape system. (*a*) Typical profile. 1 ft = 0.3 m. (*b*) U.S. Navy jumper demonstrating use of the crew escape system pole in a test aircraft.

Board (CAIB) was formed to determine the cause of the accident and to present recommendations. As during the *Challenger* investigation, all flights were suspended.

Recovery. Seven months after the disaster, the CAIB issued its report. The cause of this accident was a breach in the thermal protection sytem on the leading edge of the left wing, caused by a piece of insulating foam that separated during launch from the left bipod ramp section on the external tank. This breach in the reinforced carbon-carbon panel allowed superheated air to penetrate through the leading edge of the left wing, which melted the aluminum structure of the wing. Aerodynamic forces caused loss of control of the shuttle and its subsequent breakup.

The CAIB made 15 recommendations. The independent NASA Return to Flight Task Group agreed that NASA had met the intent of 12 of the recommendations. NASA could not meet the major recommendation, to eliminate foam from leaving the external tank, due to engineering problems. NASA did remove the foam on the bipod ramp and made improvements in foam application to lessen the foam debris. NASA also met other recommendations for better ground-based imagery of the vehicle, and better high-resolution imagery of the orbiter and the external tank.

STS 114 return to flight. When *STS 114* launched in July 2005, the new cameras on the external tank, installed for high resolution of the orbiter, recorded a large piece of foam coming off the tank. On orbit, the crew used the new Orbiter Boom Sensor System (OBSS) to inspect the reinforced carbon-carbon and the tiles for damage. Only two gap fillers were found, extending between the tiles. NASA decided that the fillers had to be removed or cut off. An astronaut on the station arm pulled them out easily. On another spacewalk, NASA also tested several methods to repair the tiles and the reinforced carbon-carbon, should further damage be detected in the future.

Six in-flight anomalies were taken due to the unexpected foam loss. The analysis and wind tunnel testing required to understand and eliminate the loss led to further delays in the launch of the next mission. Worse, when hurricane Katrina struck the Gulf Coast, it damaged the Michoud Assembly Facility, which builds the external tanks, causing further delays. The next mission, *STS 121*, did not launch until July 4, 2006, almost exactly one year after *STS 114*. The most significant change in the tank was the elimination of the liquid hydrogen and liquid oxygen protuberance air load (PAL) ramps. The tank was shipped from Kennedy Space Center back to Michoud for the modifications.

Technical achievements. The first operational shuttle accomplished the first commercial satellite deployments in 1982. In 1984 the first in-orbit satellite repair was accomplished on the *Solar Maximum Mission*'s control system and a primary experiment sensor. Later that same year *Palapa* and *Westar*, two communications satellites in useless orbits, were recovered and returned to Earth on the same space shuttle mission. These repair and recovery missions demonstrated the usefulness of humans in these new space tasks.

In September 1988, a *Tracking and Data Relay Satellite* (*TDRS*) was launched by the first space shuttle to fly after the *Challenger* accident. The initial *TDRS* network was completed 6 months later by the second shuttle launch of a *TDRS*, which enabled the closing of many Spaceflight Tracking and Data Network (STDN) ground stations and a major improvement in the duration and bandwidth of communication with robotic and crewed spacecraft. *See* SPACE COMMUNICATIONS; SPACECRAFT GROUND INSTRUMENTATION.

The year 1989 was a banner one for solar system exploration, thanks to space shuttle launches of

Magellan to Venus and *Galileo* to Jupiter. In 1990 the first mission, *Ulysses*, was undertaken to investigate the magnetic field configuration and variations around the Sun's poles, after the craft's shuttle launch. STS 32 launched the long-awaited *Hubble Space Telescope* (*HST*) in 1990; however, this telescope's most important work was not accomplished until later that decade after two shuttle EVA visits—the first in 1993 to correct a bad case of astigmatism produced by the *HST* mirror manufacturer, and the second in 1997 to update the sensors and replace major spacecraft subsystem components that were failing. Additional servicing missions were carried out in 1999 and 2002. In 2006, after much debate, NASA decided to perform one more *Hubble* servicing mission, enhancing its capabilities and extending its life by replacing failed components and updating some sensors and data processors.

Science accomplishments. Successful launches of *Magellan*, *Galileo*, *Ulysses*, and *HST* (together with the two in-orbit servicing calls) constitute a major contribution to the space sciences. The deployment of the Long Duration Exposure Facility (LDEF) as a free-flying satellite by a shuttle in 1984 and its shuttle recovery in 1990 alerted materials researchers to the rapid degradation of many materials used in space, even though some data were lost due to the delay of the LDEF recovery because of the hiatus in shuttle operations in 1986–1988. Finally, the experience of the bioscientists and materials researchers developing research methods and equipment for experiments aboard *ISS* is helping to get this work off to a good start. *See* REMOTE MANIPULATORS; SPACE BIOLOGY; SPACE PROCESSING.

Safety upgrades. Until 2001, official NASA statements projected the availability of a single-stage-to-orbit space shuttle replacement by 2012. Cancellation or suspension in 2001 of the X-33 and other steps toward a shuttle replacement led to speculation that the space shuttle might be the major crewed link with *ISS* until 2020 or later. Consequently, the Space Flight Advisory Committee advised NASA to give higher priority and funding to space shuttle high-priority safety upgrades, which NASA had already identified and was investigating.

In 2001, NASA planned to develop four high-priority safety upgrades for the shuttle fleet. An electric auxiliary power unit would replace the hazardous hydrazine-fueled high-speed turbine-driven hydraulic systems. New engine sensors in the Main Engine Advanced Health Monitoring system could cut the risk of a catastrophic engine failure by about 25%. The Solid Rocket Booster Advanced Thrust Vector Control project would shift away from hydrazine in the turbines that drive the nozzle hydraulics. Cockpit upgrades would install advanced avionics in the shuttle fleet, including the flat-panel cockpit displays already available on *Atlantis* to ease the crew workload during ascent.

In 2004, President Bush announced the Vision for Space Exploration. Part of the Vision required the shuttle to be retired no later than 2010, only 5 years after the first return-to-flight mission. Consequently, some of the safety upgrades were canceled, and other upgrades were added. The electric auxiliary power units and Solid Rocket Booster Advanced Thrust Vector Control programs were canceled. The Advanced Health Monitoring System was flown in monitor-only mode on *STS 116* in December 2006. The results will be examined before it is placed in operational mode. All three orbiters received the Cockpit Avionics Upgrade.

A new upgrade was added, the Station/Shuttle Power Transfer System (SSPTS). Since the *ISS* is powered from an inexhaustible supply of sunlight and the shuttle is limited in power by the amount of hydrogen and oxygen carried onboard, a system was designed to power the shuttle from the *ISS*. The SSPTS converts the 125-volt DC *ISS* power to 28-volt DC shuttle power. This will allow the shuttle to stay at the *ISS* for longer periods of time. The shuttle *Edeavour* has been modified with the SSPTS and is scheduled to fly in June 2007 to test it. Installation of the SSPTS is also planned on the shuttle *Discovery*, but not on *Atlantis*, due to the retirement of the shuttles.

John F. Clark; Jeffrey C. Mitchell

Bibliography. *Aviation Week and Space Technology*, March 19, 2001; H. Benedict, *NASA: The Journey Continues*, 2d ed., 1992; *Columbia Accident Investigation Board Report*, 6 vols., Government Printing Office, 2003; D. R. Jenkins, *Return-to-Flight Space Shuttle Discovery*, 2006; D. R. Jenkins, *Space Shuttle*, 3d ed., 2001; NASA, *Shuttle Flight Data and Inflight Anomaly List*, Rep. JSC-19413, Revision M, January 1990; J. R. Wertz and W. J. Larson (eds.), *Space Mission Analysis and Design*, 3d ed., 1999.

Space station

A complex physical structure specifically designed to serve as a multipurpose platform in low Earth orbit. Functioning independently and often without a crew actively involved in onboard operations, a space station contains the structures and mechanisms to operate and maintain such support systems as command and data processing, communications and tracking, motion control, electrical power, and thermal and environmental control. Evolving together with technology and increasing in scope and complexity, the space station has a history in which each program was based upon the developments and achievements of its predecessor.

Salyut. The Soviet Union began construction of the world's first space station, called *Salyut*, in 1970. It was the primary Soviet space endeavor for the next 15 years. The design was retained not only through a series of Salyuts that were launched within that decade and later, but also through the development of *Mir*, the most famous, long-lived Russian achievement in space. The Salyuts were cylindrical in shape and contained compartments with specialized functions for operating a space station. The docking compartment was designed to accept the Soyuz spacecraft that transported the cosmonauts. The transfer

compartment gave the cosmonauts access to the various work compartments. The work compartments contained the mechanisms that operated and controlled the station, as well as the laboratories in which the cosmonauts performed experiments while they were onboard. *Salyut 1* initially carried a crew of three, but after three cosmonauts died when a valve in the crew compartment of their descent module burst and the air leaked out, the crews were reduced to two cosmonauts, and both crew members were outfitted with pressurized space suits.

During the Cold War, Soviet Salyuts were launched as missions to serve both civilian and military purposes. These missions never were linked in a flight, and many of the details about the flights were shrouded in secrecy. Since the end of the Cold War those details have been declassified, revealing the extent and scope of Soviet activity in space, which included astronomical and medical experiments; testing of a broad range of communications and telemetry, guidance, and navigation systems for orbiting spacecraft; Earth observation using cameras and satellites; manufacture of space-processed products; and, perhaps most importantly, evaluations such as the effects of zero gravity over an extended period of time on a human body.

The Soviet Salyuts set the record of 8 months for extended duration of human space flight. At the end of their usefulness, all Salyut spacecraft were deorbited by Mission Control Moscow and permitted to burn up during reentry into the Earth's atmosphere.

Skylab. Starting later than the Soviets, the United States also had plans for a space station. The program, known as Skylab, was developed by the National Aeronautics and Space Administration (NASA) building on the success of its heavy-lift rocket, the Saturn, which had boosted the Apollo rockets and helped place the first human being on the Moon. *Skylab* weighed just less than 100 tons (90 metric tons). It was launched on May 14, 1973, from the Kennedy Space Center aboard a Saturn 5 rocket, which enjoyed a reputation for never having failed. Although the launch was flawless, a shield designed to shade *Skylab's* workshop deployed about a minute after liftoff and was torn away by atmospheric drag. That began a series of problems, most involving overheating, that had to be overcome before *Skylab* was safe for human habitation. Eventually three crews served aboard *Skylab* throughout 1973 for periods of 28, 59, and 84 days, respectively. The single greatest contribution made by each crew was the extravehicular activities that restored *Skylab's* ability to serve as a space station.

Apollo-Soyuz Test Program. In 1975, during the period of dètente, plans were made for a joint United States–Soviet space venture, known as the Apollo-Soyuz Test Program (ASTP). For the first time, United States astronauts and Soviet cosmonauts became members of each other's flight teams, training together, touring their launch sites, and working in their mission control rooms. For 15 months, they trained and prepared for their historical joint mission in space. In July 1975, the astronauts of *Apollo 18* linked with the cosmonauts of *Soyuz 22*; their docked configuration functioned as a miniature international space station.

Mir. Throughout the 1970s, the goal of the Soviet space program appeared to be the development of a permanently deployed space station. It would have both civilian and military applications, and it would be based upon the technology gleaned from the experience with the Salyut and the Soyuz vehicles. A major task that the Soviets sought to accomplish in their quest for a permanent space station was the optimization of conditions that would enable cosmonaut crew members to live in space for extended periods of time. The crew then would be able to operate at maximum efficiency and effectiveness when performing space-based experiments. Soviet decision makers sought to take advantage of the unique opportunities presented by a permanent space station to exploit the space environment and benefit their economy. For example, Soviet scientists studied methods to direct solar power by means of equipment onboard a space station to Soviet facilities for consumption at home.

Mir (which in Russian means either "peace" or "world") was the name of the vehicle for the world's first multipurpose, permanently operating crewed space station. Based upon the Salyut design and configuration, *Mir* essentially was an extension and expansion of the shell of the basic Soviet space vehicle (**Fig. 1**). It would incorporate the standard Salyut/Soyuz/Progress profile. External ports were added for docking of Soyuz vehicles which would carry crew members, and Progress (resupply) vehicles which would bring foodstuffs, drinking water, extra equipment and apparatus, sanitary requisites, medical apparatus, and propellant. These same Progress vehicles would return space "junk" to Earth. Construction was based upon a modular design, permitting the Soviets to replace modules whenever significant improvements in technology made the earlier modules obsolete. Two cylindrical modules formed the basic shape of *Mir* and served as the living and central control compartments for the crews. Additional modules were used for scientific experiments.

The *Mir* core station was a 49-ft (15-m) module. Its launch aboard a Soviet Proton rocket on February 20, 1986, was televised internationally for the first time in Soviet space history. *Mir* was assembled in space and was composed of six modules. In addition to the core, the Kvant-1 module (launched in April 1987) was a 19-ft (6-m) pressurized lab used for astrophysics research. The Kvant-2 module was launched in December 1989, and was used for biological research and for Earth observation. The Kristall module was launched in August 1990, and provided the docking port for the United States space shuttles that visited *Mir*. Between 1994 and 2001, there were nine dockings between space shuttle vehicles and *Mir*. The final two modules, Spektr and Priroda (launched in June 1995 and April 1996, respectively) were remote sensing modules used to study the Earth. *Mir*

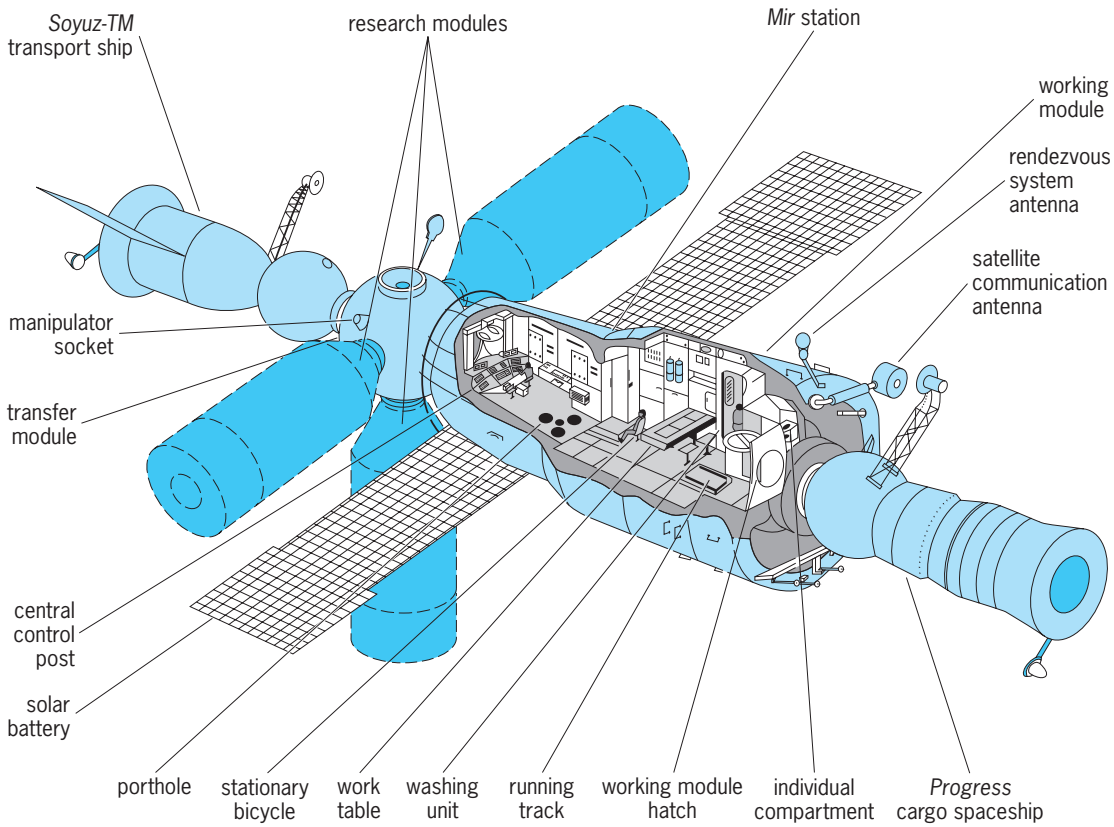

Fig. 1. ***Mir*** **complex, in orbit 1986–2001. (*After Commercial Uses of Space Exploration* [*brochure*], *Glavcosmos, 1989*)**

provided a platform for long-term microgravity research and increased knowledge about living and working in space. By the end of its function in March 2001, the 143-ton (130-metric-ton) *Mir* had spent 15 years in orbit and had orbited the Earth more than 87,600 times with an average orbiting speed of 17,885 mi/h (28,800 km/h). Its superstructure, 109 ft (33 m) long and 90 ft (27 m) wide, burned up upon reentry in to the Earth's atmosphere, and scraps were scattered into the Pacific Ocean just east of New Zealand. *See* AEROSPACE MEDICINE; SPACE BIOLOGY; SPACE PROCESSING; SPACE SHUTTLE.

International Space Station. Through the late 1980s, the United States continued to compete with the Soviet Union for dominance in space. The Soviet successes with *Mir* prompted the United States to respond with what became known as *Space Station Freedom*. Moved from the NASA drawing board in 1993, components of the space station were tested by shuttle crews during a variety of missions. Supporters of this space station recognized that the benefits of space exploration are far-reaching and long-term, and they advocated such unique opportunities as manufacturing drugs, scientific materials research in microgravity, and studying the health and status of the Earth's environment from outer space. At the direction of President Clinton, the United States transformed the single-nation concept for *Space Station Freedom* into a multinational partnership with the European Space Agency and the Russian Space Agency to create what is now known as the *International Space Station* (*ISS*). *See* SPACE PROCESSING.

Five space agencies [NASA in the United States, the Canadian Space Agency (CSA), the European Space Agency (ESA), the Japan Aerospace Exploration Agency (JAXA), and the Russian Federal Space Agency (RSA)] and sixteen nations have united to build the *International Space Station*. More complex than any ground-based civil engineering project built to date, the *ISS* will take approximately 12 years to construct (**Fig. 2**).

The design evolved over more than a decade. When completed, the *ISS* will be the largest human-made object ever to orbit the Earth. It will include six laboratories built by a consortium of nations. The internal research space will be equivalent to the cabin space inside a 747 aircraft. Approximately 80 United States space shuttle, Russian Soyuz, and partner-nation heavy transport vehicle flights will transport the structures and mechanism necessary to construct the station over those 12 years. The United States space shuttle must deliver most *ISS* modules and major components. The *ISS* will weigh approximately 1 million pounds (almost 450,000 kilograms); it will orbit the Earth at an inclination of 51.6° to either side of the Equator; it will fly at an altitude between 230 and 285 mi (370–460 km); and it will move at a speed of 17,500 mi/h (28,000 km/h), orbiting the Earth every 90 min. It will host a maximum crew of six.

As of January 2007, there had been 18 United States space shuttle, 14 Russian Soyuz, and

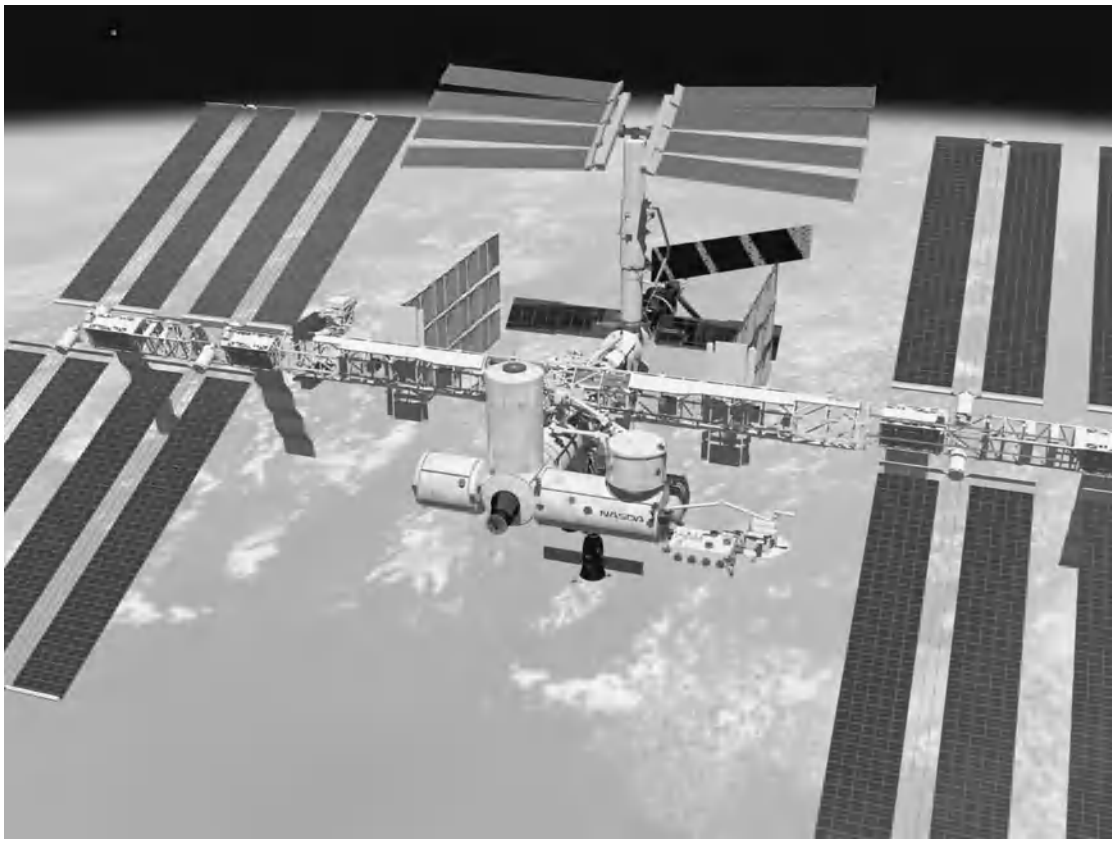

Fig. 2. Artist's rendering of how the *International Space Station* may look when its assembly sequence is completed in 2010. (Changes to the final station configuration will likely occur before then.) The 1-million-pound (450,000-kg) station will have a pressurized volume equal to two jumbo jets, and an acre of solar panels. (*NASA*)

25 Russian Progress flights to build and support the *International Space Station*. The first flight, from the Baikonur Cosmodrome (a launch facility in Kazakhstan), was a Russian Proton rocket in November 1998, which carried the Zarya module or Functional Cargo Block (FGB). This was a self-contained, independently operating power and communications module. It formed the base to which further components were added. A month later, the space shuttle *Endeavour* carried up the Unity module, which was attached to Zarya. Additional flights added components and conducted repairs. The Russian service module (SM), Zvezda, was launched and docked with the FBG in July 2000. It provided the station's early living quarters and systems to support power, data processing, flight control, propulsion, and life support. Consistent with the Russian historical design and architecture of the *Salyut* and *Mir* platforms, Zarya and Zvezda featured automated operation.

Expedition flights support the ISS. Astronauts and cosmonauts assigned to flights to the *ISS* compose an Expedition crew. As of January 2007, there had been 14 Expedition crews posted to the *ISS*. Expedition crews are composed of either two or three members: a Commander and either one or two Flight Engineers. The primary missions of each Expedition crew are to assemble the station and to learn how to live and work in space. These are essential tasks if crewed space programs are to venture deeper into the solar system for the purposes of human exploration and scientific discovery. To this end, Expedition crew members serve as research scientists. They use the *ISS* as an orbiting laboratory and as a test platform for future technologies.

Expedition One begins building the ISS. The first permanent crew of the *International Space Station* also launched from the Baikonur Cosmodrome, on October 31, 2000, aboard a three-stage, 310-ton (280-metric-ton) Soyuz rocket. Known as Expedition One, the crew was composed of one American and two Russians. U.S. Navy Captain William (Bill) Shepherd served as the commander in what was his fourth trip into space. The Russians were Air Force Lieutenant Colonel Yuri Gidzenko, who served as Soyuz commander, and Sergei Krikalev, who served as flight engineer. This was the second flight for Gidzenko, whose first flight was as the commander for *Mir* on which he served for 179 days. Krikalev had been in space four times previously, logging 484 days in space. One of his flights was to the *International Space Station* aboard space shuttle *Endeavour* for the mission which mated the Zarya and Unity modules. Shortly after their arrival, the Expedition One crew named the *International Space Station* the "Space Station Alpha."

During flight the Expedition One crew used Greenwich Mean Time (GMT) as its reference. This

was designed to provide the best combination of communications coverage for the American and Russian flight control teams.

While aboard the *International Space Station*, the Expedition One crew hosted three visiting space shuttle crews. Those crews arrived with the mission of continuing the construction of the station. The vast array of photovoltaic cells that provide the bulk of the power for the station arrived and was installed by one crew. An additional crew brought the American laboratory, Destiny, which will be the center for scientific research. Within the laboratory are housed a variety of scientific racks, containing hardware and equipment for research and study. The primary mission of the Expedition One crew was to prepare the station for long-term habitation, and there were several significant tasks that had to be accomplished. For crew command and control, all onboard computer control systems and all means to identify emergencies, warnings, and cautions were inspected and tested for function and reliability. For crew comfort and safety, all air purification and revitalization devices, heaters, and water supply systems were inspected and made serviceable. For communications within the modules and down to Mission Controls in Houston and Moscow, all internal and external communications, command and data handling, and telemetry systems were tested and made operational. In addition, the Expedition One crew was to refurbish the *International Space Station*. That was accomplished by unloading three crewless Russian Progress resupply vehicles, all of which arrived and docked automatically to the ports available on the Russian Zarya and Zvezda modules.

Because of the amount of time required by the Expedition One crew to prepare the station for nominal operations, there was not much time planned for scientific activities and experiments. However, some of the experiments planned for the first expedition included research into protein crystal growth, analysis of seed growth, study of the Earth by means of high-resolution images taken from the space station, spatial differences in carbon dioxide concentration, physiological tests involving a resistive exercise device, and treadmill and medical evaluations of locomotion and heart rate. Scientific activities prepared independently by the Russian Space Agency involved research in the fields of biotechnology, biomedical research, geophysical research, and engineering.

Expedition Two continues construction. Expedition One was succeeded by Expedition Two, which launched aboard space shuttle *Discovery* on March 8, 2001, and docked with the *International Space Station* the following day. The Expedition One crew returned to Earth aboard *Discovery*, which undocked from the *International Space Station* on March 18, 2001, and landed at Edwards Air Force Base 2 days later. Their highly successful mission had lasted 138 days 17 hours 39 minutes.

Expedition Two also had a three-member crew. The commander was Russian cosmonaut Yuri Usachev, and the two flight engineers were U.S. Air Force Lieutenant Colonel Susan Helms and U.S. Army Colonel James Voss. On their arrival flight, the space shuttle *Discovery* carried the first multipurpose logistics module, which was built by the Italian Space Agency and named Leonardo. The logistics modules are reusable cargo carriers designed specifically for space-flight operations. The space shuttle *Endeavour* with a crew of seven launched from Kennedy Space Center on April 19, 2001, to deliver additional station-building materials and instrumentation. Unique among the cargo in this flight was the space station's remote manipulator system or robotic arm, and another multipurpose logistics module, called Rafaello.

Follow-on Expeditions. Through Expedition Fourteen and continuing with Expeditions until the *ISS* is no longer viable as a platform for exploration and research, the Expedition crews will serve onboard the *ISS* for stays between 4 and 6 months. In addition to the tasks of helping build, maintain, and sustain the *ISS*, the Expedition crews conduct science activities daily in a variety of fields such as human research, physical and life sciences, technology demonstrations, Earth observations, and education activities. These efforts were interrupted briefly after the loss of the space shuttle *Columbia* in February 2003, and United States space shuttle flights to the *International Space Station* were suspended until July 2005.

Given the *ISS* status as an advanced orbiting laboratory operating in zero gravity, the opportunities to study human endurance in space and to test new technologies and techniques for space exploration are endless. Human medical research, which develops the knowledge essential to send space explorers beyond Earth's orbit and deep into the solar system, focuses both on the effects of living in space and zero gravity and the countermeasures necessary to reduce the health risks associated with space travel. The goal is to defeat the harmful effects of space on human health. Research in the physical and life sciences is intended to use the environment of microgravity to study the growth of proteins, cells, crystals, and other biological structures. Technology demonstrations help develop new technologies used in spacecraft materials and systems that function in the microgravity environment. Finally, Earth observations and education activities link the global population with the work and mission on board the station, inspiring humankind to continue to explore and use space as a resource for improving life on Earth.

Future development. In addition to serving as a test bed and springboard for future space exploration, the *ISS* will serve as a research laboratory for innovation in science and technology. Early research on the *ISS* has focused on such topics as biomedical research and countermeasures to understand and control the effects of space and zero gravity on crew members; biological study of gravity's influence on the evolution, development, and internal processes of plants and animals; biotechnology to develop superior protein crystals for drug development; and

fluid physics. Advanced research will be oriented toward topics such as human support technology, materials science, combustion science, and fundamental physics.

Construction of the *International Space Station* is projected to be completed in 2010. The ISS Program will continue to coordinate the training and operations of international flight crews, use vehicles built in multiple countries and launched at multiple sites, integrate a worldwide communications network that extends into deep space, and unite the international scientific and engineering communities.

When the *International Space Station* is completed, it will host an international crew of six members. Training for these crews is projected for a period between 3 and 5 years. Expedition crews will live and work aboard the station for a period of 3 to 6 months. The flight records of these expeditions and the results of the research conducted by the crews will be available to the international community. *See* SPACE FLIGHT; SPACE TECHNOLOGY; SPACECRAFT STRUCTURE. James J. Pelosi

Bibliography. V. Gorshenin, *History of Soviet Space Exploration in Manned Orbiting Facilities*, Soviet Academy of Sciences Press, Moscow, 1994; V. Grenowitz, The triumph of *Mir*: A Russian national treasure, *Space Tech. J.* (Moscow), no. 33, April 2001; M. Lechliter, *Whence the Journey into Whirlwind: Cosmonaut Challenges in Space*, Moscow State University Press, 2000; Library of Congress, Federal Research Division, *History of Skylab*, Item no. 1167, November 1975; N. R. Pellis and R. M. North, Recent NASA research accomplishments, aboard ISS, *54th International Astronautical Congress of the International Astronautical Federation, the IAA, and the International Institute if Space Law*, Bremen, Germany. IAC-03-T.4.07, Sept. 29–Oct. 3, 2004; J. L. Rhatigan, J. A. Robinson, and C. F. Sawin, *Exploration-Related Research on ISS: Connecting Science Results to Future Missions*, NASA/TP-2005-213166, 1–34, 2005; J. A. Robinson and D. A. Thomas, NASA utilization of the International Space Station and the Vision for Space Exploration, ISTS 2006-0-1-06, *International Symposium on Space Technology and Science*, Kanazawa, Japan, June 4–10, pp. 1–6, 2006.

Space technology

The systematic application of engineering and scientific disciplines to the exploration and utilization of outer space. Since 1957, space satellites and probes have been launched beyond the Earth's atmosphere to orbit the Earth (velocity of 7 km/s or 4.4 mi/s) or to escape from it (velocity 11 km/s or 6.9 mi/s). Due to the severe environmental conditions in space, such as high vacuum, intense solar radiation, and extremes of temperature, a variety of disciplines are required to design, build, test, and launch a successful spacecraft. *See* ESCAPE VELOCITY.

Space technology developed so that spacecraft and humans could function in this environment that is so different from the Earth's surface. Conditions that humans take for granted do not exist in outer space. Objects do not fall. There is no atmosphere to breathe, to keep people warm in the shade, to transport heat by convection, or to enable the burning of fuels. Stars do not twinkle. Liquids evaporate very quickly and are deposited on nearby surfaces. The solar wind sends electrons to charge the spacecraft, with lightninglike discharges that may damage the craft. Cosmic rays and solar protons damage electronic circuits and human flesh. The vast distances require reliable structures, electronics, mechanisms, and software to enable the craft to perform when it gets to its goal—and all of this with the design requirement that the spacecraft be the smallest and lightest it can be while still operating as reliably as possible.

Basic Technology

All spacecraft designs have some common features: structure and materials, electrical power and storage, tracking and guidance, thermal control, and propulsion.

Structure and materials. The spacecraft structure is designed to survive the forces of launch (3 *g* for a space shuttle launch) and to be lifted by cranes, attached to the vehicle, and otherwise handled during ground processing. The structure is made of metals (aluminum, beryllium, magnesium, titanium) or a composite (boron/epoxy, graphite/epoxy). It must also fit the envelope of the launcher. For the space shuttle this is 5 m (16.5 ft) in diameter and the length of a bus. For expendable launchers, the spacecraft is enclosed in a fairing, a cylindrical volume that varies with the launcher. The volume of the fairing and shuttle is not large enough to permit fully deployed solar arrays and other devices during launch. The spacecraft must also have structures that extend out so that sensors will be far enough away from sources of interference, such as radioisotope thermal electric generators (RTGs), to be useful, or outside the magnetic fields of the spacecraft. *See* SPACE SHUTTLE.

The use of fluids and gases presents unique problems. In the case of the Hubble Space Telescope, the optics are so sensitive that traditional thrusters cannot be used—they would deposit contaminants on the mirrors and ruin the clarity of the optics. Batteries, heat pipes, ammonia coolant loops, and thrusters must be analyzed to determine how their leakage would affect the mission. For long space missions, care must be taken to have exterior surfaces that are resistant to atomic oxygen, high solar intensity, or extremes of heat and cold. Sliding surfaces or mechanical bearings requiring lubricants are usually placed inside a sealed container. When they are not, liquid lubricant replenishment is provided, or special solid lubricants may be used. Graphite, which is an excellent lubricant on Earth, becomes a harsh lubricant in space when absorbed gases leave its surface. Special dry lubricants or magnetic bearings need to

be used. *See* SPACE TELESCOPE, HUBBLE; SPACECRAFT STRUCTURE.

Thermal control. The average temperature of a spacecraft is determined by equating the heat input with the heat radiated away. The major power input for spacecraft located between Earth orbit and the orbit of Mercury is solar radiation. This is equal to the product of the solar constant (approximately 1370 W/m^2 near the Earth and 5900 W/m^2 near Mercury), the area facing the Sun, a, and the average absorptivity of the surface, α (0 to 1). The heat radiated away is the product of the emissivity of the surface, ϵ (0 to 1), the surface area, A, the Stefan-Boltzmann constant σ [5.67×10^{-8} $W(m^2/K^4)$], and the fourth power of the absolute temperature. For a sunlit spacecraft that does not generate heat, the temperature can be obtained from Eq. (1). Note that the

$$\sigma T^4 = \frac{\alpha}{\epsilon}\frac{a}{A}S \tag{1}$$

temperature varies with the properties of the surface (α/ϵ), the orientation of the spacecraft (a/A), and the intensity of the solar radiation (S). As the spacecraft's time in space increases, the surface characteristics change due to ultraviolet exposure, atomic oxygen attack, and micrometeoroid impact. *See* SOLAR CONSTANT.

Spacecraft generally radiate heat internally; thus the above equation is a simplification. Radio transmitters generate many watts of heat, as do other electrical devices on board. Externally, the reflected light from the Earth or other celestial bodies adds to the external heat absorbed. The spacecraft's temperature will drop radically when these power inputs drop, as when the spacecraft passes into the shadow of the Earth and the solar radiation is eclipsed. For example, the temperature of an eclipsed deployed solar array of a spacecraft in geostationary orbit may drop to −100 K (−280°F).

Temperatures inside the spacecraft depend on thermal radiation between surfaces, thermal conduction through solid components, and heat dissipation of electrical components at various locations. These in turn depend on the surface coating (paint, metallic surface, or other coating). To maintain temperatures at acceptable limits, various active and passive devices are used: coatings or surfaces with special absorptivities and emissivities, numerous types of thermal insulation, such as multilayer insulation (MLI) and aerogel, mechanical louvers to vary the heat radiated to space, heat pipes, electrical resistive heaters, or radioisotope heating units (RHUs).

Multilayer insulation has been used extensively on spacecraft. It consists of layers of insulation between layers that reflect radiation. The atmosphere between the layers is removed in the vacuum of space, which prevents the conduction of heat. It is tens to hundreds of times more effective than foam or fiberglass batting.

Heat pipes are sealed tubes or rectangular pipes in which a liquid is evaporated at the hot end and the vapor flows to the cold end where it condenses and travels back to the hot end via a wick. They have the advantage of no moving parts and transferring 200 times the heat of a solid pipe of the same cross-sectional area. *See* HEAT PIPE; TEMPERATURE.

Tracking and positioning. The location of a spacecraft can be measured by determining its distance from the transit time of radio signals or by measuring the direction of received radio signals, or by both. Data are usually taken over a period of hours or days, and the best orbit fit is determined numerically by computer. Distances are often measured by transmitting tones of different frequencies and finding the phase relations of the returned signal. The Intelsat ranging system uses four frequencies from 35 to 27,777 Hz, while the Goddard Space Flight Center range and range rate (GRARR) system transmits tones from 8 to 500,000 Hz. The direction of a spacecraft can be determined by turning the Earth station antenna to obtain the maximum signal, or by other equivalent and more accurate methods. *See* ASTRONAUTICS; CELESTIAL MECHANICS; GUIDANCE SYSTEMS; ORBITAL MOTION; SPACE NAVIGATION AND GUIDANCE.

Propulsion. The velocity of a spacecraft is changed by firing thrusters. Solid propellant thrusters are rarely used. Liquid propellant thrusters are either monopropellant or bipropellant. Monopropellant thrusters use a fuel such as hydrazine or hydrogen peroxide that decomposes when it comes in contact with a catalyst. Bipropellant thrusters use a fuel and oxidizer that ignite when the two combine in the thruster. Electric thrusters, such as mercury or cesium ion thrusters, have also been used. Electric thrusters have the highest efficiency (specific impulse) but the lowest thrust. The *Deep Space 1* spacecraft, which visited an asteroid and was expected to visit Comet Borrelly in late September 2001, uses electric thrusters to accelerate xenon ions to 109,435 km/h (68,000 mi/h). *See* HYDRAZINE; INTERPLANETARY PROPULSION; ION PROPULSION; PROPELLANT; SPACECRAFT PROPULSION; SPECIFIC IMPULSE.

Attitude control. Most spacecraft are spin-stabilized or are three-axis body-stabilized. The former uses the principles of a gyroscope; the latter uses sensors and thrusters to maintain orientation. Some body-stabilized spacecraft (such as astronomical observatories) are fixed in inertial space, while others (such as Earth observatories) have an axis pointed at the Earth and rotate once per orbit. A body-stabilized spacecraft is simpler than a spinner but requires more hardware. The orientation of a spacecraft is measured with Sun sensors (the simplest method), star trackers (the most accurate), and horizon (Earth or other body) or radio-frequency (rf) sensors (usually to determine the direction toward the Earth). Attitude corrections are made by small thrusters or by reaction or momentum wheels; as the motor applies a torque to accelerate or decelerate the rotation, an equal and opposite torque is imparted to the spacecraft. Torques can also be exerted by turning the axis of a momentum wheel. When a wheel reaches its mechanical limit of rotation speed,

momentum is "dumped" by firing thrusters; that is, as the wheel decelerates, the thrusters are fired to exert a torque equal and opposite to the torque of the decelerating wheels. Momentum wheels are used on the Hubble Space Telescope because thrusters would deposite contaminants on its mirror. To reduce size and weight, momentum wheels are being considered for both attitude control and power storage. *See* STAR TRACKER.

Electrical power. Primary electrical power is most often provided by solar cells made from a thin section of crystalline silicon protected by a thin glass cover. Gallium arsenide solar cells are beginning to replace silicon, despite their increased weight, because they are more efficient. The crystal contains dopants to produce extra electrons on one side and holes on the other side, forming a *pn* junction. The *n*-type material is on the sunlit surface, since it is more resistant to the damage caused by the radiation in the Van Allen belts. Sunlight creates hole-electron pairs in the junction, and the electric field caused by the space charge layer in the junction sends the electrons toward a metal grid, and the holes toward the base electrode. A 2×4 cm (0.8×1.6 in.) cell provides about 150 mW of power at 10–18% efficiency for silicon cells and up to 23% for gallium arsenide cells. *See* SOLAR CELL.

Excess power from the solar cells is stored in rechargeable batteries so that when power is interrupted during an eclipse, it can be drawn from the batteries. Nickel-cadmium batteries were used widely but are being replaced by nickel-hydrogen batteries, which provide higher energy per weight and higher energy per volume, properties that are very desirable for spacecraft. Other batteries with even better properties are being tested, such as lithium ion batteries, which up to now have been considered too dangerous, since they can explode or catch fire. Electrical power can be stored with momentum wheels. On the *International Space Station*, a flywheel energy storage system was being designed to replace the nickel-hydrogen batteries. The lifetime of batteries is limited by the temperature control maintained and the depth of discharge during operation. The latter is the fraction of power taken from the battery during each charge-discharge cycle. Flywheels are not limited by temperature or depth of discharge. They can be charged and discharged indefinitely. *See* BATTERY.

Other sources of power generation include fuel cells, radioisotope thermoelectric generators (RTGs), tethers, and solar dynamic power. Fuel cells have been used on the Apollo and space shuttle programs and produce a considerable amount of power, with drinkable water as a by-product. Fuel cells use the same principles as a battery, but employ a fuel and oxidizer to produce electricity and water at very high efficiencies (70%). RTGs use the heat of decay of radioisotopes to produce electricity. They are very useful for spacecraft at distances of Mars and beyond where the solar flux would require impractically huge arrays of solar cells to produce the same amount of power. RTGs have low efficiencies (9%), and the radioisotope of choice, plutonium, is very expensive. Tethers use a kilometers-long conducting wire moving through the magnetic field of the Earth to generate electricity, the same principle as a generator. They have been tested on the *TSS-1R* mission and produced voltages as high as 3500 V and current levels of 480 mA. Solar dynamic power is being considered for the space station. It uses the Sun's energy to heat a fluid to turn a turbine, much like terrestrial power plants. These devices have the advantage of high power and high efficiency (30%) while being small and not producing as much atmospheric friction as the large solar arrays. *See* ELECTROMAGNETIC INDUCTION; FUEL CELL; NUCLEAR BATTERY; SOLAR ENERGY; SPACE POWER SYSTEMS.

Telemetry and command. The status and condition of a spacecraft are determined by telemetry. Temperatures, voltages, switch status, pressures, sensor data, and many other measurements are transformed into voltages, encoded into pulses, and transmitted to Earth. This information is received and decoded at the spacecraft control center. Desired commands are encoded and transmitted from the control center, received by the satellite, and distributed to the appropriate subsystem. Commands are often used to turn equipment on or off, switch to redundant equipment, make necessary adjustments, and fire thrusters and pyrotechnic devices. *See* SPACE COMMUNICATIONS; SPACECRAFT GROUND INSTRUMENTATION; TELEMETERING.

Special-Purpose Technology

Many spacecraft missions have special requirements and hence necessitate special equipment.

Space probes. Satellites that leave the Earth's gravitational field to travel around the Sun and visit other planets have special requirements. Distances are greater (on the order of 1.8×10^8 mi or 3×10^8 km), mission times are longer (years), and the intensity of the Sun's radiation may change markedly (inversely proportional to the square of the distance from the Sun). With greater distances, communications are more difficult, and information must be transmitted at a slower rate. For probes that go closer to the Sun, the higher intensity of sunlight poses a severe thermal problem. For probes that go to the outer planets (Jupiter, Saturn, and beyond), the decrease in solar intensity makes solar cells less effective, requiring other power sources such as radioisotope thermoelectric generators. *See* SPACE PROBE.

Reentry vehicles. Spacecraft that return to Earth require special protection. Retro-rockets are used to reduce the orbital velocity so that the satellite falls and enters the Earth's atmosphere. Most of the decrease in velocity is produced by atmospheric friction; in the last stage this may be augmented by deploying parachutes. On entering the atmosphere, the velocity is extreme (4.4–6.9 mi/s or 7–11 km/s), and the friction generates high temperatures (thousands of degrees). One type of protection is an ablative

shield. This shield, which was used on *Apollo* capsules, was designed to melt and vaporize, dissipating heat in the process. On the shuttle, a special thermal shield made of glass fibers can withstand the high temperatures, with enough insulation to protect the interior.

The shape of a reentry spacecraft is designed to maintain a stable and desired attitude during reentry. Communications during part of the reentry are cut off by the high-density plasma sheath formed around the spacecraft, through which radio-frequency signals will not penetrate. Final recovery may be accomplished in the air (film reconnaissance satellite, as in the Discoverer program), in the ocean (Apollo astronauts), or on land (shuttle and Russian cosmonauts).

For missions to other planets, aerobraking has been used. Aerobraking uses the atmosphere of the planet instead of using retrorockets to slow the spacecraft. The *Magellan* spacecraft used aerobraking in 1994 to achieve orbit of Venus, and the *Mars Global Surveyor* used aerobraking in the much thinner atmosphere of Mars to achieve orbit. *See* ATMOSPHERIC ENTRY; NOSE CONE.

Crewed spacecraft. Technology must be developed to meet the many needs associated with human survival in space. Oxygen equivalent to one-fifth of an atmosphere is needed for breathing; the remaining nitrogen is desirable but not essential. The early Soviet cosmonauts had an atmospheric pressure essentially equivalent to that on Earth; United States astronauts had much less nitrogen, with a total pressure of only one-third of an atmosphere. This lower pressure is satisfactory for breathing, but increases the fire hazard and does not carry sound as far (only 16 ft or 5 m for normal conversation). Provision must be made for adding oxygen and removing carbon dioxide from the air, for supplying food and water, and for disposing of human wastes. For extravehicular activity (EVA), a special suit is required to provide oxygen for breathing and to maintain adequate pressure.

The lack of gravity produces a number of technological requirements for astronauts living in space. Physically, it is just as easy (or just as hard) to walk on the ceiling as to walk on the floor, but an astronaut feels more comfortable in a room with a local vertical, that is, with one surface that looks like a floor and another that looks like a ceiling. Objects tend to float in space so that restraints are necessary: mechanical, elastic, or magnetic. Showers require special vacuum hoses to remove the water, since it does not run down the drain. Astronauts can be weighed by rocking in a special chair fitted with springs; the astronaut's mass is determined from the frequency of rocking. *See* SPACE BIOLOGY; SPACE FLIGHT; SPACE STATION; WEIGHTLESSNESS.

Rendezvous. In some missions one spacecraft must find, approach, and make contact with another spacecraft. One active spacecraft performs the maneuvering by firing thrusters, while the other spacecraft (target) remains passive. During the far approach, the orbits of the two spacecraft are determined by conventional means, and maneuvers are calculated to bring the spacecraft closer together. During the near approach, the position of the target spacecraft is determined directly from the active spacecraft, sometimes by radio and sometimes by visual contact. For final approach, a closing velocity is maintained with the proper attitude, so that contact is made with the desired surfaces. Usually a probe mates with an appropriate drogue, hooks hold the two spacecraft together, and mechanical devices then firmly attach one spacecraft to the other.

For the space shuttle, the Ku-band antenna also serves as a radar antenna for final approach and rendezvous. When the shuttle docks with the *International Space Station*, it uses docking hardware manufactured in Russia.

Reliability

Space is distant not only in kilometers but also in difficulty of approach. Large velocity changes are needed to place objects in space, which are then difficult to repair and expensive to replace. Therefore spacecraft must function when they are launched, and continue to function for days, months, or years. The task is similar to that of building a car that will go 125,000 mi (200,000 km) without requiring mechanical repair or refueling. Not only must space technology build a variety of parts for many missions, but it must achieve a reliability far greater than the average. This is accomplished by building inherent reliability into components and adding redundant subsystems, supported by a rigorous test schedule before launch. Efforts are made to reduce the number of single points of failure, that is, components that are essential to mission success and cannot be bypassed or made redundant.

Electronics. The failure rate for an electronic subsystem is often estimated by adding up the failure rates of its components. The failure rate of each component is assumed to be independent of that of other components, and the failure of a subsystem. While electronic parts are highly reliable, a subsystem has so many parts that there may be an appreciable chance of failure. Component failure rates are minimized by special manufacture, quality control, and derating, that is, operation at voltages or temperatures which are lower than the maximum permitted.

If an electronic subsystem does not have the desired reliability, a redundant unit can be included. If either of the two subsystems continues to operate, the mission can still be accomplished. Mathematically, the probability that at least one of the subsystems is operating, R_{12}, is given by Eq. (2), where R is

$$R_{12} = 1 - (1 - R)^2 \tag{2}$$

the reliability of the single subsystem. For example, if each subsystem has a reliability, R, of 0.98, then the probability that at least one subsystem is working is 0.9996, which is a significant improvement. The equation above is based on the assumption that failures of the two subsystems are not related; this may not be true when two units are built by the

same manufacturer at the same time. *See* RELIABILITY, AVAILABILITY, AND MAINTAINABILITY.

Mechanical systems. While there are fewer mechanical parts than electrical parts on most spacecraft, mechanical parts tend to be less reliable than electrical ones. Deployment mechanisms usually work only once, but must work correctly the first time. Bearings must be lubricated for their expected lifetimes, and special lubricants must be used. Torque margins are increased, so that even if unusually high torques are encountered, the motors will still keep the device turning. Electromechanical relays should have currents reduced before they are opened, to prevent arcing and welding of contacts in vacuum.

Testing. The key to a reliable spacecraft is extensive testing before launch. This ensures not only that every spacecraft part is adequately functioning, but also that the spacecraft is subjected to the differing environments expected in space, and in some cases is overstressed to find the weak components. Tests are performed first on components, then subsystems, and finally, when feasible, on fully assembled spacecraft. The testing program requires many weeks and consists of three main parts: electrical performance, vibration, and thermal-vacuum tests. Electrical testing includes evaluation of the spacecraft's ability to perform a variety of operations which will be required in space. Antennas may be separately tested on an antenna range, where the effect of reflections from nearby objects is minimized. Steady-state acceleration may be evaluated with a static test, in which mechanical loads are placed on various parts of the structure, or with a large centrifuge. Vibration tests, simulating the launch environment, subject the spacecraft to low-frequency vibration (5–100 Hz), acoustic vibration (20–10,000 Hz), and pyrotechnic shock conditions. Many of these tests can be accomplished with an electromechanical vibrator. Thermal-vacuum tests consist of placing the spacecraft in a high-vacuum chamber (less than 10^{-5} torr or 10^{-3} pascal) and simulating the temperatures that will occur in space. *See* ENVIRONMENTAL TEST; INSPECTION AND TESTING; MECHANICAL VIBRATION; QUALITY CONTROL.

Sometimes the first spacecraft tested (prototype) remains on the ground, and only subsequent (flight) spacecraft are launched. The prototype spacecraft is often less reliable because of modifications made, tests which must be repeated, and planned overstressing in tests. Most spacecraft perform their missions successfully, but often unexpected anomalies occur. Redundant components may fail, and subsystems, although performing adequately, may not quite meet design specifications. Telemetered performance is analyzed, problems are diagnosed, and changes are proposed in the design and testing of future spacecraft. In a few cases, space technology improves the spacecraft performance in space, and appropriate rf commands are sent to modify its operation, to reset parameters, to modify software, or to fire thrusters which will reorient the spacecraft or shake it to expel a bubble in a fuel line.

Testing must be timed correctly during the assembly process. The Tethered Satellite System 1 was fully tested to ensure that the satellite with its tether would deploy properly. A modification was then approved that ultimately caused the tether not to deploy fully, resulting in a partially failed mission. The *International Space Station* has the unique problem that some of the components are not connected to each other until they are in orbit. NASA has a system using state-of-the-art computer-aided-design-and-manufacturing (CAD/CAM) software and photography to determine that the components will mate to each other mechanically. NASA also decided to conduct tests to ensure that the electrical connections from one component to another were compatible and that the software would command and control each of the elements. NASA and Boeing had to build emulators of the components already in orbit, and will continue to do so as the station is assembled. These multielement integrated tests add further expense, but are necessary to prove that the space station will work when fully assembled. *See* ASTRONAUTICAL ENGINEERING; COMPUTER-AIDED DESIGN AND MANUFACTURING; SATELLITE (SPACECRAFT); SPACE.

Jeffrey C. Mitchell; Gary D. Gordon

Bibliography. J. A. Angelo, Jr., and D. Buden, *Space Nuclear Power*, Orbit Book Company, 1985; G. D. Gordon and W. L. Morgan, *Principles of Communications Satellites*, 1993; M. D. Griffin and J. R. French, *Space Vehicle Design*, 2d ed., 2004; C. M. Harris and A. G. Piersol, *Harris' Shock and Vibration Handbook*, 5th ed., 2002; M. H. Kaplan, *Modern Spacecraft Dynamics and Control*, 1976; G. Maral and M. Bousquet, *Satellite Communication Systems*, 4th ed., 2002; W. L. Morgan and G. D. Gordon, *Communications Satellite Handbook*, 1989; G. P. Sutton and O. Biblarz, *Rocket Propulsion Elements*, 7th ed., 2001.

Space-time

A term used to denote the geometry of the physical universe as suggested by the theory of relativity. It is also called space-time continuum. Whereas in newtonian physics space and time had been considered quite separate entities, A. Einstein and H. Minkowski showed that they are actually intimately intertwined. I. Newton's ideas on space and time are summarized in the following list:

1. Given two events, each of which is clearly localized in space and lasts only for an instant in time, such as two strokes of lightning striking small targets, all observers of the events will be in agreement as to which of the two events took place earlier in time, or whether they were actually simultaneous.

2. If the events were not simultaneous, the interval of time between them is an absolute entity, agreed on by all competent observers.

3. The spatial distance between two simultaneous events is an absolute entity, agreed on by all competent observers.

Of these three, the first assumption, concerning the concept of simultaneity of distant events, is the crucial one, the other two depending on it. Simultaneity, however, can be given an unambiguous meaning only if there is available some instantaneous method of signaling over finite distances. Actually, according to the theory of relativity, the greatest speed of transmission of intelligence of any kind is the speed of light c, equaling about 3×10^{10} cm/s. Moreover, any signal traveling precisely at the speed c appears to travel at that same speed to all conceivable observers, regardless of their own states of motion. This is the only reasonable interpretation of the results of the Michelson-Morley experiment and the effect of aberration. Accordingly, the question of whether two given events are simultaneous or not can be decided only with the help of signals that at best have traveled from the sites of these events to the station of the observer at the speed of light. *See* ABERRATION OF LIGHT; LIGHT; RELATIVITY.

Under these circumstances, Einstein showed that in general two observers, each using the same techniques of observation but being in motion relative to each other, will disagree concerning the simultaneity of distant events. But if they do disagree, they are also unable to compare unequivocally the rates of clocks moving in different ways, or the lengths of scales and measuring rods. Instead, clock rates and scale lengths of different observers and different frames of reference must be established so as to assure the principal observed fact. Each observer, using his or her own clocks and scales, must measure the same speed of propagation of light. This requirement leads to a set of relationships known as the Lorentz transformations. *See* LORENTZ TRANSFORMATIONS.

In accordance with the Lorentz transformations, both the time interval and the spatial distance between two events are relative quantities, depending on the state of motion of the observer who carries out the measurements. There is, however, a new absolute quantity that takes the place of the two former quantities. It is known as the invariant, or proper, space-time interval τ and is defined by Eq. (1), where

$$\tau^2 = T^2 - \frac{1}{c^2}R^2 \tag{1}$$

T is the ordinary time interval, R the distance between the two events, and c the speed of light in empty space. Whereas T and R are different for different observers, τ has the same value. In the event that Eq. (1) would render τ imaginary, its place may be taken by σ, defined by Eq. (2). If both τ and σ are

$$\sigma^2 = R^2 - c^2T^2 \tag{2}$$

zero, then a light signal leaving the location of one event while it is taking place will reach the location of the other event precisely at the instant the signal from the latter is coming forth.

The existence of a single invariant interval led the mathematician Minkowski to conceive of the totality of space and time as a single four-dimensional continuum, which is often referred to as the Minkowski universe. In this universe, the history of a single space point in the course of time must be considered as a curve (or line), whereas an event, limited in both space and time, represents a point. So that these geometric concepts in the Minkowski universe may be distinguished from their analogs in ordinary three-dimensional space, they are referred to as world curves (world lines) and world points, respectively.

Minkowski geometry. The geometry of the Minkowski universe in some respects resembles the geometry of ordinary (euclidean) space but differs from it in others. The Minkowski universe has four dimensions instead of the three dimensions of ordinary space; that is to say, for a complete identification a world point requires four pieces of data, for instance, three space coordinates and a time reading. But there are in the Minkowski universe world points, world lines, two-dimensional surfaces (including planes), three-dimensional surfaces (often called hypersurfaces), and four-dimensional domains. A hypersurface may, for instance, be a spatial domain (volume) at one instant in time, or it may be a two-dimensional (ordinary) surface for an extended period of time. Thus one may form all the geometric figures that are also possible in a four-dimensional euclidean space.

The indefinite metric. In a euclidean space there are the cartesian coordinate systems, those rectilinear systems of coordinates that are mutually perpendicular and whose coordinate values correspond to real lengths. The distance S between two points whose coordinate differences are X, Y, and Z, respectively, is given by Eq. (3). The coordinate transformations

$$S^2 = X^2 + Y^2 + Z^2 \tag{3}$$

that lead from one cartesian coordinate system to another are called orthogonal coordinate transformations. Formally, they are the coordinate transformations that preserve the precise form of Eq. (3). Likewise, the Lorentz transformations preserve the precise form of Eqs. (1) and (2), respectively. A coordinate system in which these two equations hold is called a Lorentzian frame of reference.

Whereas the form on the right of Eq. (3) is positive definite, that is, always greater than or equal to zero, the right-hand sides of Eqs. (1) and (2) are indefinite; that is, they may be positive or negative. This fact represents the single but all-important difference between a four-dimensional euclidean space and a Minkowski universe. The forms (1), (2), and (3) are called metrics. Hence the Minkowski universe is said to possess an indefinite metric.

As a result of the indefinite character of the metric, a triangle in the Minkowski universe may possess one side that is longer than the sum of the two others; conversely, one of the three sides may have the length zero. Depending on whether τ or σ is real, or both vanish, an interval is classified as timelike, spacelike, or a null-interval (see **illus.**)

Improper Lorentz transformations. In the Minkowski universe there are three different types of Lorentz transformations involving some kind of reflection, in addition to the more usual Lorentz transformations

(called proper Lorentz transformations). The first type of improper Lorentz transformation changes the sign of all three spatial coordinates but leaves the sense of the time axis unchanged. This transformation changes a right-handed screw into a left-handed screw; it is also called a parity transformation. The second type of improper Lorentz transformation interchanges the future with the past but leaves the space coordinates unchanged. This transformation is called time reversal. The third improper Lorentz transformation reflects both the space and the time coordinates; it bears no special name of its own. The original arguments which led to the formulation of the special theory of relativity all support the proposition that the laws of nature are invariant under proper Lorentz transformations. They are inconclusive as to whether the laws of nature should also be invariant under the improper Lorentz transformations. The laws of mechanics and of electrodynamics have this property; the second law of thermodynamics distinguishes between past and future in a nonsymmetric manner, but it is usually assumed that this feature of thermodynamics is to be explained by its statistical nature and that it has no bearing on the properties of the underlying basic dynamics. *See* SYMMETRY LAWS (PHYSICS).

Until recent times, it had therefore been taken for granted that the basic laws of nature should make no distinction between right-handed and left-handed screws, and that they should not discriminate between future and past. Certain difficulties in the interpretation of the decay of *K* mesons led T. D. Lee and C. N. Yang to suspect that these assumptions might not be tenable, and they suggested some experiments on meson decay and on radioactive beta decay, which showed that in these particle transformations nature certainly discriminates between left-handed and right-handed screws. *See* PARITY (QUANTUM MECHANICS).

Neutrinos are electrically neutral particles which travel at the speed of light and are endowed with an intrinsic spin equal to that of electrons. It appears that the neutrino always spins counterclockwise when viewed in the direction of its travel, while its antiparticle, the antineutrino, always spins clockwise. (As far as is known, this holds true of each of the presently known varieties of neutrinos.) The exact role of the improper Lorentz transformations in nature is under intensive theoretical and experimental investigation. *See* NEUTRINO.

Curved space-time. Whereas the Minkowski universe is the appropriate geometric model for the special theory of relativity, the general theory of relativity makes use of a further generalization. In the Minkowski universe a particle that is not subject to external forces and which therefore travels along a straight line (in ordinary space) and at a uniform speed is represented by a straight world line. In general relativity, an external gravitational force is indistinguishable from an inertial force, which in special relativity would arise if a non-lorentzian frame of reference were to be employed. Accordingly one requires a four-dimensional space in which it is impossible to distinguish between a lorentzian and a non-lorentzian frame of reference whenever a gravitational field is present. Such a space cannot be flat, as the Minkowski universe is, but must be curved. Such an aggregate is called a riemannian space. In a general riemannian space there are no straight lines but only curves. *See* FRAME OF REFERENCE; GRAVITATION; RIEMANNIAN GEOMETRY. Peter G. Bergmann

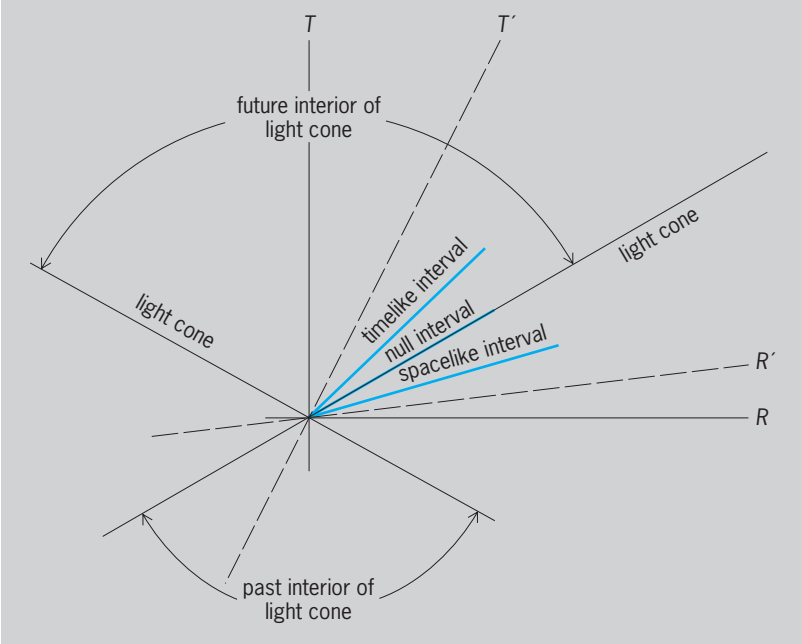

Two-dimensional sketch of three possible classes of intervals. The null directions actually form a three-dimensional surface having the shape of a double cone, called the light cone. An alternative Lorentz frame is indicated by broken lines.

Spacecraft ground instrumentation

Instrumentation located on the Earth for monitoring, tracking, and communicating with satellites, space probes, and crewed spacecraft. Radars, communication antennas, and optical instruments are classified as ground instrumentation. They are deployed in networks and, to a lesser extent, in ranges. Ranges are relatively narrow chains of ground instruments used to follow the flights of missiles, sounding rockets, and spacecraft ascending to orbit. Some ranges are a few miles long; others, such as the U.S. Air Force's Eastern Test Range, stretch for thousands of miles. Networks, in contrast, are dispersed over wide geographical areas so that their instruments can follow satellites in orbit as the Earth rotates under them at 15° per hour, or space probes on their flights through deep space.

Networks are of two basic kinds: networks supporting satellites in Earth orbit, and networks supporting spacecraft in deep space far from Earth. A third concept was added in the 1980s with the Tracking and Data Relay Satellite System (TDRSS), also called the Space Network (SN). TDRSS replaced most of the ground stations used for Earth orbital support. The *Tracking and Data Relay Satellites* (*TDRS*) are

placed in geosynchronous orbits to relay signals to and from other orbiting spacecraft during more than 85% of each orbit, to and from a single ground station.

Functions. Ranges and networks have various technical functions:

1. Tracking: determination of the positions and velocities of space probes and satellites through radio and optical means.
2. Telemetry: reception of telemetered signals from scientific instruments and spacecraft housekeeping functions.
3. Voice reception and transmission: provision for communication with the crew of a spacecraft, such as the space shuttle.
4. Command: transmission of coded commands to spacecraft equipment, including scientific instruments.
5. Television reception and transmission: provision for observation of the crew, spacecraft environment, and so on.
6. Ground communications: telemetry, voice, television, command, tracking data, and spacecraft acquisition data transmission between network sites and the central mission control center, and payload information to user facilities.
7. Computing: calculation of orbital elements and radar acquisition data prior to transmission to users; also, computation of the signals that drive visual displays at a mission control center. *See* TELEMETERING.

Mission control. From the mission control center, highly trained mission personnel, observing computer-generated displays, dispatch commands and analyze acquired data through ground communication with the appropriate network stations. From these stations, command data are dispatched to orbiting satellites.

Tracking can be accomplished by a great variety of instruments, broadly classified as optical or radio-frequency (rf). In the optical class are cinetheodolites, ballistic cameras, Baker-Nunn cameras, telescopes, and laser ranging systems. Radio-frequency tracking systems are even more varied and include radar, active and passive radio interferometers, radio Doppler systems, and range-and-range-rate systems, which resemble Doppler radars except that they do not measure spacecraft direction. *See* DOPPLER RADAR; OPTICAL TRACKING SYSTEMS; RADAR.

Ranges. The best-known range in the United States is the Eastern Test Range (ETR; formerly called the Atlantic Missile Range), which begins at Cape Canaveral, Florida. In addition, the Department of Defense operates the Western Test Range (WTR) and several others. The National Aeronautics and Space Administration (NASA) launches most of its spacecraft from the Eastern Test Range, but also launches some sounding rockets and small satellites from its Wallops Island facility on the Virginia east coast. During the shuttle era, the Eastern Space and Missile Center (ESMC) within the Eastern Test Range performs the functions of both the Eastern Test Range and the Western Test Range for the shuttle. Each of the United States ranges possesses tracking equipment, telemetry receivers, and electronic equipment to perform its functions. Many other countries have missile and sounding rocket ranges.

The Eastern Test Range is one of the longest ranges in the world. The range head begins with the NASA and military launch pads and blockhouses, on Cape Canaveral proper, and those of NASA on nearby Merritt Island, Florida. A missile or spacecraft launched from one of these areas climbs vertically, then pitches toward downrange tracking and communication antennas. As the rocket moves over the Atlantic, it passes within range of stations instrumented to track, command, and receive telemetry from the rising launch vehicle and its payload. *See* LAUNCH COMPLEX.

Cape Canaveral itself is well supplied with instruments such as the C-band AN/FPS-16 radar, many tracking telescopes, and several large Doppler and interferometer tracking systems. Similar instruments along the range watch each flight. The range stations are connected to each other by submarine cable and radio and satellite communication links. Geodetic surveys, timing signals from radio station WWV of the National Institute of Standards and Technology, and standardized operating procedures also unify the range stations.

Spacecraft networks. NASA operates two ground-based networks (**Fig. 1**) and the TDRSS. The ground-based networks are the Spaceflight Tracking and Data Network (STDN), which tracks, commands, and receives telemetry from United States and foreign satellites in Earth orbit; and the Deep Space Network (DSN), which performs the same functions for all types of spacecraft sent to explore deep space, the Moon, and solar system planets. The TDRSS provides the same support to Earth orbital spacecraft as the STDN. The U.S. Department of Defense operates two classified networks: the Satellite Control Facility (SCF); and the National Range Division Stations, which include those of all United States military ranges. Russia and European Space Agency (ESA) also maintain similar networks.

Spaceflight Tracking and Data Network (STDN). The STDN is a complex of ground-based stations designed to support NASA's Earth-orbiting scientific and applications spacecraft and crewed spaceflight programs such as the space shuttle. The operations control center and the central computing facility for operation and analysis are located at the Goddard Space Flight Center in Greenbelt, Maryland (Fig. 1).

The STDN, operational since May 1971, combined the former Space Tracking and Data Acquisition Network (STADAN) and the Manned Space Flight Network (MSFN). The crewless STADAN tracked spacecraft in Earth orbits from ground facilities distributed over much of the globe. The minimum tracking (minitrack) facilities, implemented in 1958, used antennas composed of steel rails parallel to the ground. The minitrack's narrow fan pattern received radio transmissions from the spacecraft as it passed overhead. The STADAN stations used parabolic (dish) antennas whose conical beams swept the sky to follow the orbiting spacecraft.

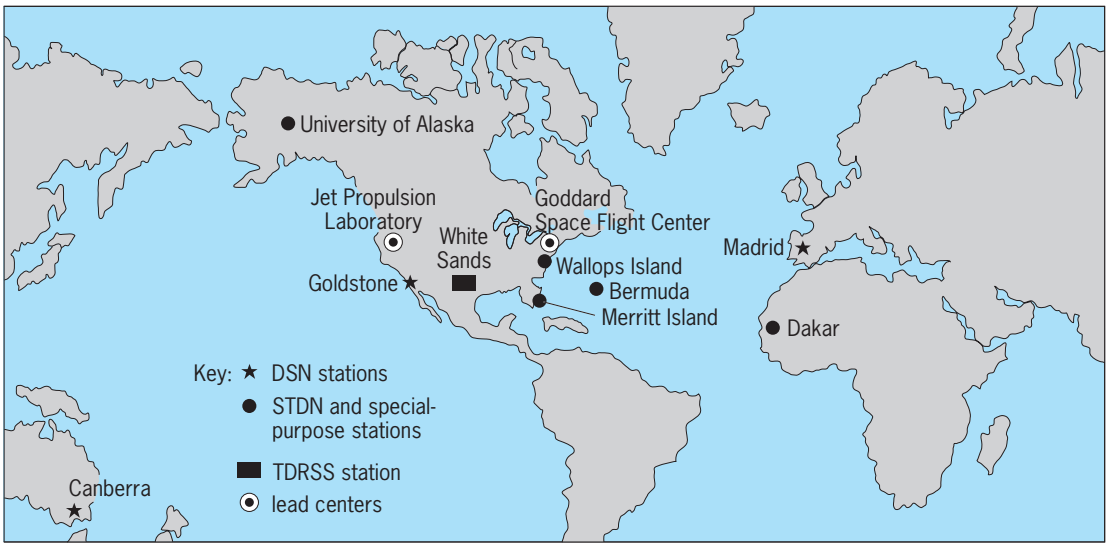

Fig. 1. World map showing the tracking networks for space operations of the National Aeronautics and Space Administration (*NASA*).

The MSFN provided two-way contact between the ground and space for the astronauts in the Mercury, Gemini, and Apollo programs. It consisted of both ground stations and tracking and support ships that could be positioned where ground stations were unavailable or impractical. The stations were configured to support the individual requirements of each mission whether the mission involved Earth orbit or Moon landing.

In general, the systems needed at each STDN station are determined by the support that the station provides and whether the support is directed primarily toward crewed or crewless space efforts. Within the STDN, the unified S-band (USB) system is the prime equipment dedicated to tracking, telemetry, and two-way communications. It is so named because voice and data can be handled on one S-band carrier. Stations that use the USB system are further identified by the size of their USB antenna systems: 85-ft (26-m) or 30-ft (9-m) dish diameter.

The significant disadvantage of the STDN is that the antennas are bound to the surface of the Earth. This creates a limited field of view because the antenna can usefully transmit and receive only when a spacecraft is overhead. Thus, the average coverage that a low-altitude spacecraft may expect to receive from the tracking stations is limited to approximately 15% of its orbit. This amount of coverage is insufficient to support newer, more sophisticated spacecraft with instruments that generate much more information. It became impractical to store this information on board the spacecraft during the long periods when the spacecraft is out of contact with ground stations. The most practical solution was to construct the TDRSS.

The TDRSS network has replaced most of the STDN stations, substituting the space segment and two ground stations for most of the ground stations within the STDN. The remaining STDN stations provide support to older spacecraft, which are not compatible with the TDRSS, and the special support, such as launch and landing support to the shuttle.

Deep Space Network (DSN). NASA, through the Jet Propulsion Laboratory of the California Institute of Technology in Pasadena, has established a Deep Space Network (DSN), consisting of three Deep Space Communication Complexes (DSCC) located at Goldstone, California; Madrid, Spain; and Canberra, Australia. The complexes are approximately 120° apart in longitude, which makes it possible to provide continuous communications for planetary spacecraft as the Earth rotates. From its beginning in 1959, the DSN has supported all the deep space probes such as *Mariner*, *Pioneer*, *Viking*, and *Voyager*, as well as assisting during the *Apollo* crewed lunar missions. While TDRSS began replacing STDN in 1985, DSN services continue to be used for Earth orbiters above 7500 mi (12,000 km) because TDRSS antenna pointing is designed only for spacecraft up to this altitude.

The Network Control Center at the Jet Propulsion Laboratory is the single point of control for the DSN. The Network Control Center handles messages transmitted to and from the spacecraft, and validates DSN-generated radio navigation data. The data are then sent to Mission Operations for flight path control. Additionally, the Network Control Center performs research, development, and engineering for the DSN.

Each DSCC has three or four antennas. A main data-capture instrument is the 230-ft (70-m) transmit-receive antenna with a parabolic surface. The precisely contoured parabolic surface permits the antenna to operate efficiently at short-wavelength radio frequencies (for example, at the X-band frequency, principally used to receive pictures from the outer planets, the wavelength is 3.5 cm). The combination of large area and high-frequency operations gives the 230-ft (70-m) antenna an extremely narrow field of view (one ten-millionth of the sky). The receiver system that receives such weak signals from deep space has a minuscule noise temperature (20 K), hundreds of

times less than that in the best high-fidelity home FM radio receivers. The total power received by each 230-ft (70-m) antenna in a typical outer planet encounter, such as *Voyager* at Uranus, can be as weak as 10^{-15} W. Yet, the noise-combating codes allow high-resolution pictures to be reconstructed with few errors. For example, in January 1986, the NASA *Voyager 2* spacecraft flew by Uranus. A good data rate was maintained by arraying or simultaneously combining signals from the DSN 230-ft (70-m) antenna and the two 112-ft (34-m) DSN antennas at Canberra and the Australian Parkes Observatory 210-ft (64-m) radio astronomy antenna 200 mi (320 km) away. This provided 29,000 bits of information per second, equivalent to a color picture every 5 min. The result was hundreds of color pictures during the brief encounter. Because Uranus was in the southern part of the sky in 1986, the Australian stations viewed the spacecraft longer than the Goldstone or Madrid stations.

In addition to receiving data, stations must send commands to the distant spacecraft to instruct the on-board computers what maneuvers to perform and what experiments to conduct. The antennas at each DSCC can simultaneously send commands to and receive data from the spacecraft. Different frequencies are used for transmitting and receiving; therefore, even though hundreds of kilowatts can be transmitted, the effect on the extremely weak signals being received from the spacecraft is negligible.

While the antenna is receiving data and transmitting commands, radio measurements of spacecraft paths are conducted for navigation. Because the very distant spacecraft often must be targeted to pass within small (5–10-mi or 8–16-km) regions near the target planets or their satellites or rings, extreme directional accuracy is required, in some cases near a ten-millionth of a degree. By receiving signals from the spacecraft at two DSCCs simultaneously, the time delay between reception at the two stations can be measured. Since the speed of light and the distance and direction between the two stations are known, the position of the spacecraft can be determined from trigonometry.

The DSN is also used as a scientific facility in radio science experiments. In these experiments, the changes in the received signal are carefully measured as they pass through a planet's atmosphere or ionosphere, interplanetary space, the solar corona, or the rings around planets. Using these measurements, radio scientists can determine the atmospheric pressure as a function of altitude above the surface. Coupled with spacecraft instruments and Earth-based measurements, such radio science has been a very powerful tool for the exploration of the solar system.

The DSN planetary radar is also used to characterize the bodies in the solar system. Its noteworthy achievements include calibrating the length scale of the solar system (the astronomical unit) to greatly improve radio navigation of planetary spacecraft, discovering the slow backward rotation of the planet Venus, finding craters on Venus, measuring altitudes on Mars, selecting Martian landing sites for the *Viking* Lander, determining that Saturn's rings were composed of large icelike particles, and studying asteroids and comets to determine surface properties and rotation.

By using techniques similar to two-station precision radio navigation, changes in relative positions of particular locations on Earth are measured. This provides a deeper understanding of crustal plate motion and, ultimately, of earthquake mechanisms. The DSN is also used as a radio astronomy observatory, either alone or in cooperation with other observatories. *See* GEODESY; RADIO ASTRONOMY.

Much advanced technology has arisen from the DSN, including the first field-operational amplifying masers for low-noise, weak-signal reception; the first field-operational hydrogen maser clocks for frequency and timing in navigation; power-efficient coded communications, including the first use of block and convolutional coding; on-line signal processing by general-purpose digital computers; and high-speed determination of signal strength as a function of frequency across a very wide frequency band. *See* ATOMIC CLOCK; MASER.

In addition to providing communications for United States deep-space craft, joint deep-space missions with other countries are supported on a cooperative basis. As the only worldwide network that can work with extremely weak signals, the DSN is a valuable resource for other countries.

Laser Tracking Network. The Laser Tracking Network consists of a series of laser systems used for ranging to retroreflector-equipped satellites in highly stable orbits. Laser stations obtain ranging data for these satellites by bouncing a highly concentrated pulse of laser light off the retroreflector corner cube installed on the spacecraft exterior. The exact position of the spacecraft in orbit can then be mathematically determined for a given point in time. By comparing several ranging operations, orbital predictions can be interpolated and extrapolated.

The resultant data have a variety of applications, such as precise prediction of satellite orbit and measurement of the Earth's gravitational field, polar motion, earth tides, Earth rotation, tectonic plate motion, and crustal motion to accuracies within the centimeter range. The Laser Tracking Network is a multinational cooperative effort with over 30 laser sites located in North and South America, Europe, China, Japan, and Australia.

The Laser Tracking Network consists of both fixed and mobile laser ranging systems. The mobile stations can be deployed worldwide to prepared sites. A typical laser mobile system consists of a laser transmitter, a receiver subsystem, the optical equipment, a computer, a timing subsystem, and a surveillance radar. The total system is housed in mobile vans, with a power-generation van required at some sites.

TDRSS. In the plan developed for 1990 and beyond, the DSN provides support to high-altitude Earth-orbiting spacecraft. Other than for launch support from the stations at Merritt Island, Florida; Bermuda; and Dakar, Senegal, all support to low-altitude Earth-orbiting satellites is from the TDRSS. The TDRSS

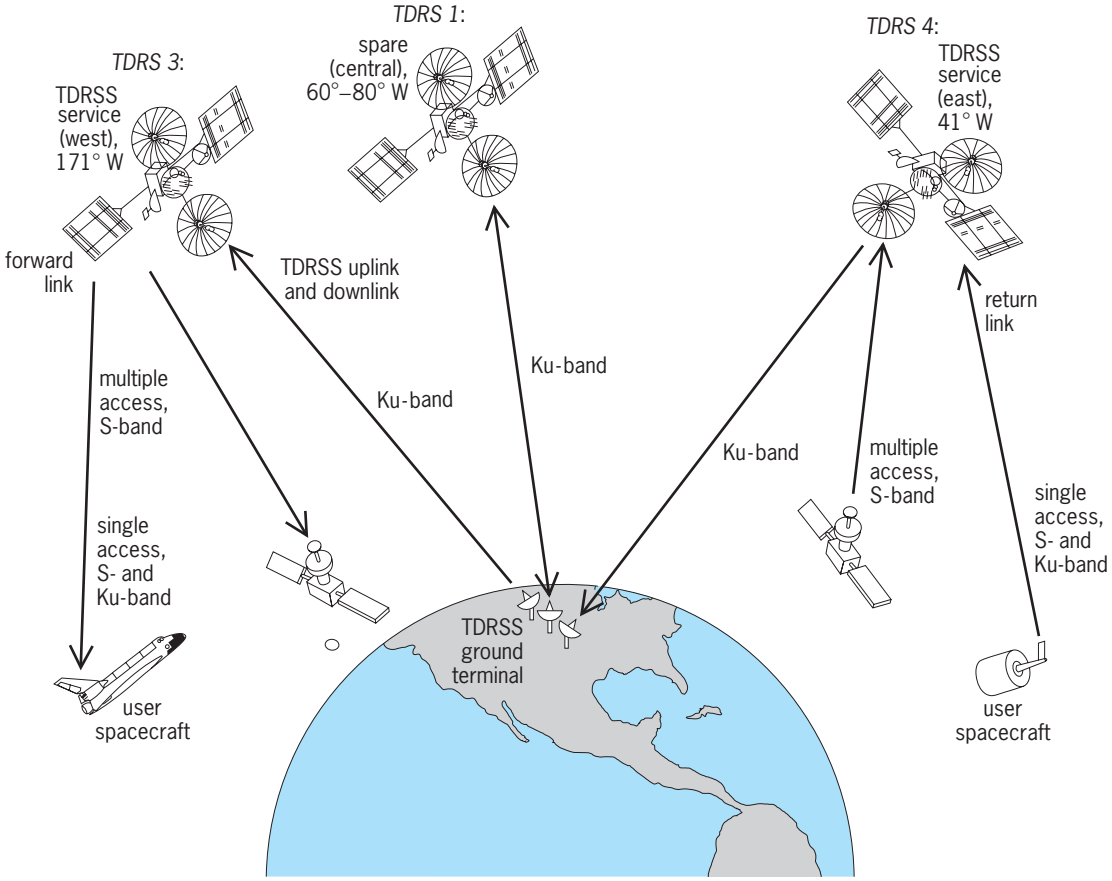

Fig. 2. Tracking and Data Relay Satellite System (TDRSS) typical user configuration.

(**Fig. 2**) consists of two operational geostationary satellites, an in-orbit spare, and two colocated ground terminals with associated operations control facilities at White Sands, New Mexico. The two active satellites, called *TDRS-East* and *TDRS-West*, are in geosynchronous orbit roughly above the Equator and at 41 and 171°W longitude, respectively. The TDRSS spare is located at 83°W longitude. The TDRSs relay signals between the ground station and satellites orbiting generally below 7500 mi (12,000 km) above Earth during nearly all of each orbit, except for a zone of exclusion which varies with spacecraft altitude.

Communication between the ground terminals and TDRSS is at Ku-band frequencies via ground terminal antennas. TDRS on-orbit tracking, telemetry, and command are provided by the same facilities. A separate antenna is provided for S-band operations during transfer orbits and as a backup to the Ku-band capability. Each TDRS can support two high-data-rate users simultaneously at either S- or Ku-band with the steerable single-access antennas. Multiple access permits up to 20 users at a time, but with much lower data rates.

The advantage of the TDRSS is that support of low-orbiting spacecraft operations is no longer restricted to short and infrequent periods of mutual STDN-spacecraft visibility. Orbital coverage for low-altitude Earth-orbiting spacecraft is increased from a maximum of 15% of each orbit, using conventional tracking stations, to 85% with TDRSS. This increase, in turn, made possible the elimination of the majority of the STDN stations and consolidation of their remaining functions with those of the DSN, which can take over responsibilities of both.

The space shuttle *Challenger* carried the first of the three satellites into space in April 1983. Since that time the two other satellites required to complete the constellation were launched, and the system is operational.

Worldwide network. The common element connecting all networks is the worldwide NASA communication system (NASCOM), which links all NASA stations, on all continents except Antarctica, through the use of submarine cables, land lines, microwave data links, and satellite and fiber-optic transmissions. The trend is progressively toward satellite communications as the data rates to be handled continue to increase. The use of satellites permits an Earth terminal to be placed at the user's facility and thus permits the transfers of data in fewer hops. Ground instrumentation provides the vital link between the spacecraft and mission controllers and scientists on the ground. *See* SATELLITE (SPACECRAFT); SPACE COMMUNICATIONS; SPACE FLIGHT; SPACE NAVIGATION AND GUIDANCE; SPACE PROBE; SPACE STATION.

H. William Wood

Bibliography. M. Gerard and P. W. Edwards (eds.), *Space Information Systems in the Space Station Era*, 1988.

Spacecraft propulsion

A system that provides control of location and attitude of spacecraft by using rocket engines to generate motion. Spacecraft propulsion systems come in various forms depending on the specific mission requirements. Each exhibits considerable variation in such parameters as thrust, specific impulse, propellant mass and type, pressurization schemes, cost, and materials. All of these variables must be considered in deciding which propulsion system is best suited to a given mission. Typical spacecraft applications include communications satellites, science and technology spacecraft, and Earth-monitoring missions such as weather satellites. Orbital environments range from low-Earth to geosynchronous to interplanetary. *See* ASTRONAUTICAL ENGINEERING; ROCKET PROPULSION; SATELLITE (SPACECRAFT); SPACE FLIGHT; SPACE PROBE; SPECIFIC IMPULSE; THRUST.

Design parameters. The two fundamental variables that define the design of spacecraft propulsion systems are the total velocity change to be imparted to the spacecraft for translational purposes, and the impulse necessary to counteract the various external torques imposed on the spacecraft body. From these, the required quantity of a given propellant combination can be specified by Eqs. (1) and (2).

$$M_{p1} = \left[1 - \exp\left(\frac{-\Delta V}{I_{sp}\, g}\right)\right] M_{s/c} \qquad (1)$$

$$M_{p2} = \frac{I}{I_{sp}} \qquad (2)$$

In Eq. (1), M_{p1} is the mass of propellant that is required to be expended to perform the translational maneuver of velocity increment ΔV; I_{sp} is the specific impulse, which is dependent on the propellant combination and engine design; g is the acceleration of gravity; and $M_{s/c}$ is the mass of the spacecraft prior to the rocket engine burn. In Eq. (2), M_{p^2} is the mass of propellant required to counteract the total disturbance impulse I. The total propellant requirement for a given mission is the sum of these results. A typical propellant budget for a communications satellite is given in the **table**. Propellant accounts for almost 60% of the lift-off mass of a communications satellite. In this case, the translational ΔV propellant constitutes over 95% of the spacecraft's needs, and station-keeping (translational) burns use over 95% of the fuel required during the operational life of the spacecraft (over 10 years).

Most spacecraft require translation in one or more axes as well as rotation about the spacecraft center of mass. The ΔV requirements are dictated by the mission design, launch-site selection, and in-orbit perturbations. For instance, some spacecraft require repositioning throughout the mission. A launch from, for example, French Guiana in South America results in a similar mission to a launch from Cape Canaveral, Florida, but the propellant requirements are substantially affected because the launch vehicles leaving from each site put the spacecraft in different orbits. The spacecraft must then perform translations of different magnitude to place themselves in the final location.

Importance of specific impulse. From Eqs. (1) and (2) the significant effect of the specific impulse I_{sp} on the total propellant load that a spacecraft must carry to perform its assigned mission can be seen. Since a massive satellite must be boosted into space by the use of expensive launch vehicles, such as the space shuttle and Ariane, significant cost savings may be gained if smaller, less expensive launch vehicles may be used. The size of the required launch vehicle is directly proportional to the mass of the payload. Since most of the other components that make up spacecraft are relatively fixed in weight, it is critical to utilize propellant combinations that maximize specific impulse.

Propellant combinations. For modern spacecraft the choices are either bipropellants, which utilize a liquid oxidizer and a separate liquid fuel; solid propellants, which consist of oxidizer and fuel mixed together; or monopropellants, which are liquid fuels that are easily dissociated by a catalyst into hot, gaseous reaction products. High specific impulse is offered by bipropellants, followed by solid propellants and monopropellants. For spacecraft the specific impulse values range from 315 s down to 220 s. The liquid bipropellants are nitrogen tetroxide (N_2O_4) as the oxidizer and either hydrazine (N_2H_4) or monomethylhydrazine ($CH_3N_2H_3$) as the fuel. Solid rocket motors use hydroxy-terminated polybutadiene (HTPB) as a binding agent for the ammonium percholorate (NH_4ClO_4) oxidizer and aluminum fuel. Monopropellants are presently limited to hydrazine, which is catalyzed on a bed of iridium-coated pellets. *See* PROPELLANT.

Propulsion system designs. Spacecraft attitude control schemes play an important role in defining the

Typical propellant budget for a geosynchronous communications satellite

Function	ΔV, ft/s (m/s)	Fuel, lbm (kg)	Oxidizer, lbm (kg)
Transfer orbit burns	—	4 (2)	—
Apogee injection	4927 (1502)	1200 (544)	1320 (599)
Preoperational	—	20 (9)	—
Station acquisition	43 (13)	25 (11)	—
Station keeping	2240 (683)	900 (408)	—
Attitude control	—	20 (9)	—
Total	7210 (2198)	2169 (984)	1320 (599)

detailed characteristics of spacecraft propulsion systems. Essentially, there are three methods for stabilizing a spacecraft: three-axis control, spin control, and gravity gradient. In three-axis systems the body axes are inertially stabilized with reference to the Sun and stars, and utilize rocket engines for control in all six degrees of freedom. Spin-stabilized spacecraft use the inertial properties of a gyroscope to permanently align one of the axes by rotating a major portion of the spacecraft body about this axis. This approach significantly reduces the number of thrusters needed for control. Gravity gradient control is a nonactive technique that relies on the Earth's tidal forces to permanently point a preferred body axis toward the Earth's center. *See* GYROSCOPE; INERTIAL GUIDANCE SYSTEM; SPACECRAFT STRUCTURE.

Translation of a spacecraft, independent of its control technique, requires thrusters aligned parallel to the desired translational axis. Usually, all three axes require translational capability. Combining the two requirements for attitude control and translation results in the minimum number of rocket engines required to perform the mission. These are supplemented with additional thrusters to allow for failures without degrading the performance of the propulsion system. Thrust levels vary and are dependent on spacecraft maximum and minimum acceleration levels, which in turn are dependent on allowable burn efficiencies, power availability, and the structural capability of deployed appendages, such as solar arrays or scientific instruments. In spacecraft applications, bipropellant and monopropellant engines are significantly lower in thrust by several orders of magnitude compared to solid rocket motors. This implies that burn durations, which are inversely proportional to thrust, are much longer for liquid propulsion systems than solid systems for a given velocity increment ΔV. Solid rocket motors are almost completely limited to performing large velocity changes. Bipropellant rocket engine thrust levels are between 100 lbf (445 newtons) for large-ΔV burns and 2 lbf (10 N) for attitude and translational maneuvers. Monopropellant thrusters range from 100 lbf (445 N) down to 0.05 lbf (0.2 N). **Figure 1** shows the schematic thruster layout for a typical three-axis controlled bipropellant propulsion system for a communications system. The propellant budget shown in the table is based on the same spacecraft mission.

Minimization of propulsion system mass. As discussed above, a critical parameter in the ultimate design of a spacecraft propulsion system is the mass of the propellant. This is dictated by the engine specific impulse and the spacecraft launch weight. Simplistically then, it would be reasonable to assume that the propellant–engine combination with the highest specific impulse would be the preferable choice. However, the ultimate requirement is the lowest possible mass for the entire propulsion system. The complexity of the system is greatly influenced by, and is roughly proportional to, the specific impulse, since bipropellants require more tanks, valves, and so forth than either solid systems or monopropellant systems. This is primarily due to the differences in density

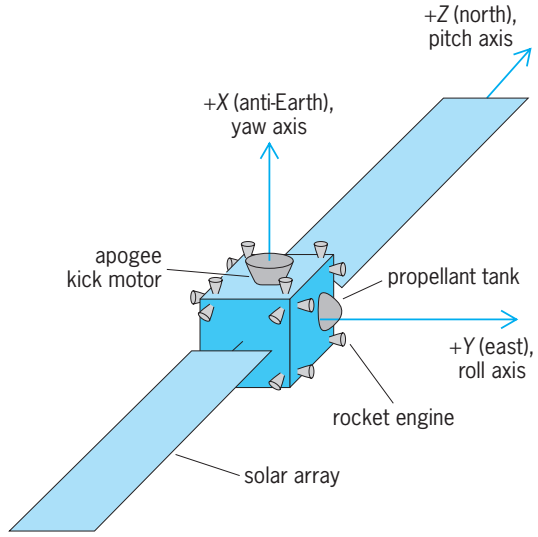

Fig. 1. Typical communications satellite thruster layout. Thrust is provided to control rotation about all three axes and translation, in six degress of freedom.

between liquids and solids, and the fact that bipropellants require high-pressure gas sources to expel the fluid from the tanks and into the rocket engine chamber. Typical three-axis stabilized bipropellant and monopropellant systems are compared in **Fig. 2**. The definition of the mission requirements, in the form of total spacecraft launch mass and the mission total velocity increment ΔV, are translated into the hardware elements needed to control the system, and the mass of these elements is estimated. Subtracting the propellant and system hardware mass from the total launch mass provides an estimate of the spacecraft mass that is available for distribution among all the other systems that constitute a spacecraft (for example, power, attitude control, and payload). For communications satellites in the lift-off weight range of 3000 lbm (1360 kg), the trade-off between specific impulse and system mass dictates the use of a solid rocket motor for the main-orbit circularizing burn and a monopropellant propulsion system for on-orbit attitude control and translation. Spacecraft launch masses above about 5000 lbm (2268 kg) require the use of all-bipropellant systems.

Blowdown versus pressure-regulated systems. From Fig. 2 it can be seen that there are two approaches to expelling propellant from spacecraft in a zero-*g* environment. The simple monopropellant systems use a pressurant gas stored within the same tank as the hydrazine fuel. Operation of the rocket engine valves allows the fuel to flow by virtue of the pressure difference between the tanks and the spacecraft exterior (vacuum). As the fluid is expended, the tank pressure decreases exponentially until all the fluid is removed. This technique is called a blowdown system. It is lightweight but requires a significant gas volume to ensure that the pressure remains at some reasonable value, since the specific impulse decreases with feed pressure. For most monopropellant systems the ratio of initial to final tank pressure is 4:1. By using the ideal gas law, it can be shown that the

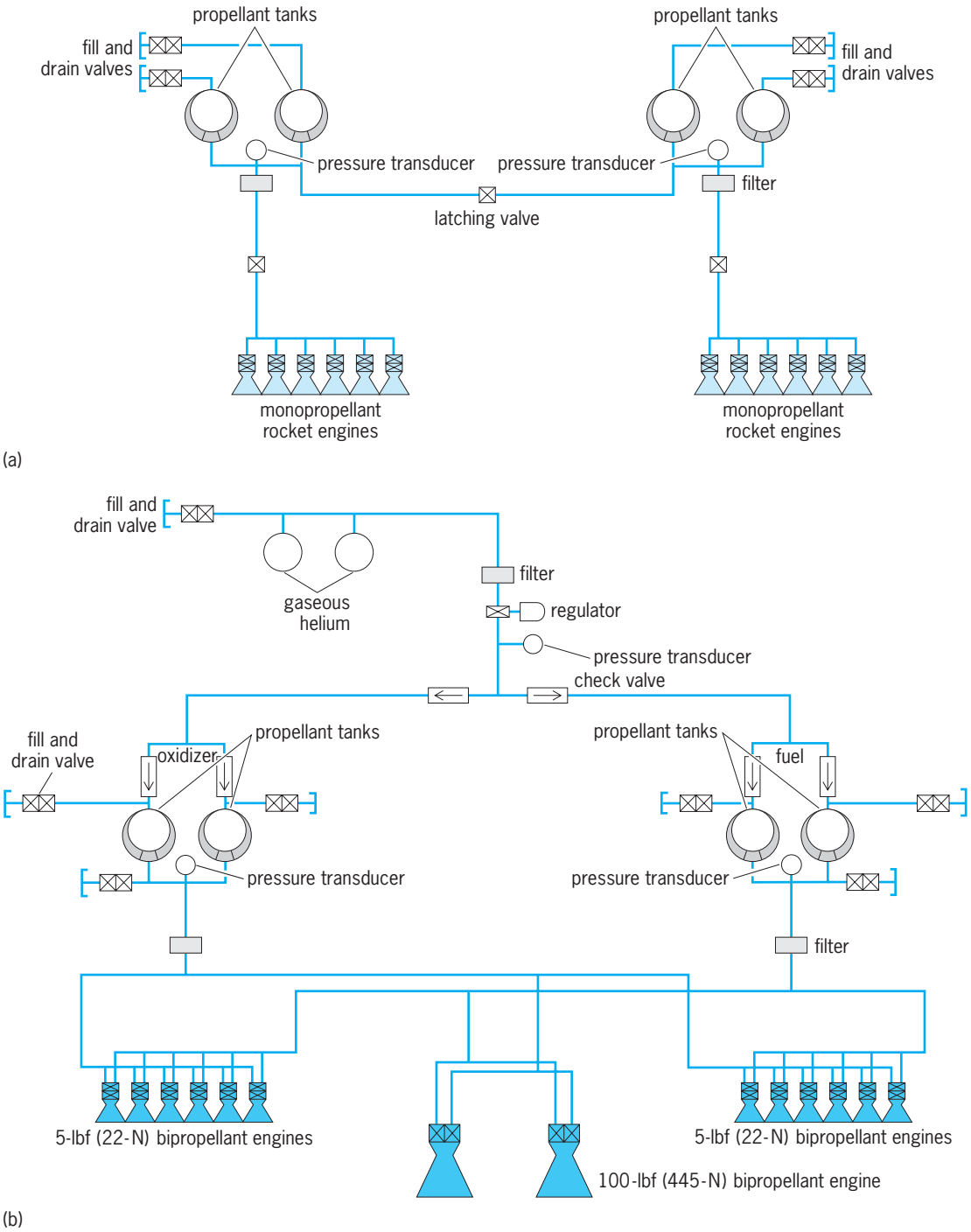

Fig. 2. Typical propulsion system designs for communications, Earth observation, and interplanetary spacecraft missions. (*a*) Monopropellant system. (*b*) Bipropellant system.

initial gas volume must be 25% of the tank volume. *See* GAS.

In systems with larger propellant mass, which must be bipropellant systems, the tank volumes required for a simple blowdown system become impractical. Therefore, bipropellant systems usually employ separate gas reservoirs, which store the gas (helium) required to expel the propellant at very high pressures; 4500 lb/in.^{2} absolute (31 megapascals) is typical. The storage-tank pressure is reduced to the engine operating pressure of 220 lb/in.^{2} absolute (1.5 MPa) by a pressure regulator, hence the name pressure-regulated system. Although the storage-tank mass is high, this is offset by the gain in propellant-tank mass compared to that in the blowdown approach and the ability to operate the system at the optimum engine pressure to maximize specific-impulse performance.

At the propellant loads typical of current spacecraft propulsion systems (2000 lbm or 907 kg), the use of turbopumps, which are often used in launch vehicles, has not proved to be weight-competitive

with the blowdown or the pressure-regulated systems. In the future, as propellant mass continues to increase as spacecraft mass increases, pumps are likely to be utilized.

Implications of zero-g. A primary feature that distinguishes spacecraft propulsion technology from that utilized on launch vehicles is the need to cope with the extremely low acceleration values that exist in orbit. (This environment is termed zero-g, although this is not really accurate, since a constant acceleration level of the order of 10^{-5} g is permanently experienced by Earth-orbiting spacecraft.) Such a low-acceleration environment primarily affects the behavior of any fluid on board (generally propellant), due to the domination of surface-tension forces. In 1-g, Earth-bound experience, fluid geometry is controlled totally by gravitational forces. Hence, water in a cup exhibits essentially a flat surface. Any difference in fluid height is removed by fluid flowing under the influence of gravity to a common surface. The relationship defining whether surface tension or gravitational forces are dominant is provided by the dimensionless Bond number B_0, defined by Eq. (3).

$$B_0 = \frac{\rho a L^2}{\sigma} \tag{3}$$

Here, ρ is the density of the fluid, a is the local acceleration, L is a characteristic length of the system (for example, the tank diameter), and σ is the fluid surface tension. For environments where B_0 is much greater than 1, the system is said to be gravity-dominated. When B_0 is much less than 1, the system is surface-tension-dominated. In those cases where the Bond number is close to 1, the system is intermediate and neither force clearly dominates. *See* SURFACE TENSION.

Clearly, in a spacecraft environment gravity is no longer significant, and the fluid internal forces define the steady-state liquid surface. In this situation, fluids attempt to minimize their total energy by reducing their surface area, and therefore the surface tension, resulting in spherical propellant shapes. The fluid spheres can attach anywhere on the internal tank surface. In particular, the propellant may not be located over the outlet of the tank, and hence gas may be expelled instead of propellant, which is clearly undesirable.

To keep the propellant "puddle" from locating itself in arbitrary locations, the principle of surface tension is often used to preferentially locate the propellant over the outlet. One way is by using channels constructed of metal with a single open face covered with a fine-weave metal screen (**Fig. 3***a*). The screen, which is woven like cloth, has pores of only about 15 micrometers that are wetted by the propellant. From Eq. (3) it follows that the Bond number is very low, and the pores are surface-tension-dominated. Figure 3*b* shows schematically the liquid configuration within a pore. When the channel is surrounded by fluid, the liquid flows through the pores with little frictional loss. In this case, gas-free propellant is provided to the propulsion system since only

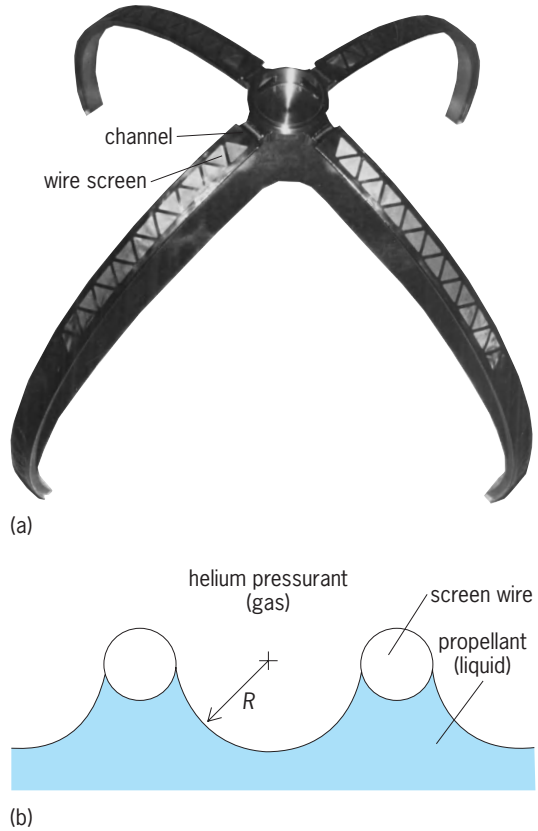

Fig. 3. Surface-tension propellant management device. (*a*) Typical device. Liquid propellant flows through the wire screens and through the channels to the circular device at the center, where it is expelled from the spacecraft. (*b*) Configuration of liqud and gas within one of the pores (radius = *R*) of the screen.

liquid is flowing out of the tank. In the more general case of the channel and the screen exposed to gas, the surface-tension forces restrict the ability of the gas to penetrate the liquid surface within the pore. The gas pressure required to ingest bubbles across the fluid barrier is termed the bubble point and is defined by Eq. (4), where ΔP is the bubble point

$$\Delta P = \frac{\sigma}{2R} \tag{4}$$

for the given configuration, σ is the surface tension of the propellant, and R is the characteristic radius (in this application it is the radius of the pore). For nitrogen tetroxide, the bubble point is about 2 lb/in.^{2} (14 kilopascals). Hence, the flow losses in the fluid moving through the screen must equal this value before the gas can break through the fluid surface and enter the propulsion system tubing. By proper design, it is possible to make channel devices that can expel 99.9% of the fluid content of a tank before the gas breaks through. When the gas does enter the tubing, and finally the engine, the spacecraft can no longer be predictably controlled, and this is the definition of the end of the mission. Typical mission durations are 10–15 years of operation for communications satellites and 1–5 years for low-Earth missions. Interplanetary spacecraft life requirements often exceed 10 years.

Keith Davies

Bibliography. B. N. Agrawal, *Design of Geosynchronous Spacecraft*, 1986; P. G. Hill and C. R. Peterson, *Mechanics and Thermodynamics of Propulsion*, 2d ed., 1992; G. P. Sutton, *Rocket Propulsion Elements*, 7th ed., Wiley, 2001.

Spacecraft structure

The supporting structure for systems capable of leaving the Earth and its atmosphere, performing a useful mission in space, sometimes returning to the Earth and sometimes landing on other bodies. Among the principal technologies that enter into the design of spacecraft structures are aerodynamics, aerothermodynamics, heat transfer, structural mechanics, structural dynamics, materials technology, and systems analysis. In applying these technologies to the structural design of a spacecraft, trade studies are made to arrive at a design which fulfills system requirements at a minimum weight with acceptable reliability and which is capable of being realized in a reasonable period of time.

The structural aspects of space flight can be divided into six broad regions or phases: (1) transportation, handling, and storage; (2) testing; (3) boosting; (4) Earth-orbiting flight; (5) reentry, landing, and recovery on the Earth; (6) interplanetary flight with orbiting of or landing on other planets. Each phase has its own structural design criteria requiring detailed consideration of heat, static loads, dynamic loads, rigidity, vacuum effects, radiation, meteoroids, acoustical loads, atmospheric pressure loads, foreign atmospheric composition, solar pressure, fabrication techniques, magnetic forces, sterilization requirements, accessibility for repair, and interrelation of one effect with the others. Heavy reliance is placed on computer-generated mathematical models and ground testing.

The basic spacecraft structural design considerations apply equally well to both crewed and crewless spacecraft. The degree of reliability of the design required is, however, much greater for crewed missions. Also, the spacecraft structures in the case of crewed missions must include life-support systems, and reentry and recovery provisions. In the case of lunar or planetary missions where landing on and leaving the foreign body are required, additional provisions for propulsion, guidance, control, spacecraft sterilization, and life-support systems must be realized, and the structure must be designed to accommodate them.

Various factors must be considered when selecting spacecraft structural materials for the three principal phases of space flight—launch, space journey, and reentry (**Table 1**).

Transportation, handling, storage. For ground transportation and storage, special load isolation systems and hermetically sealed carriers are often built to control shock and prevent contamination. The spacecraft is monitored continuously to prove that it has remained within its design envelope as it travels to the launch site by truck, airplane, or barge.

TABLE 1. Materials selections factors*

Factors	Phase I: launch	Phase II: space journey	Phase III: reentry
Technological factors	Rocket power	Satellite all-up weight	Resistance to high kinetic heating
	Thrust alignment	Impact by meteoroids,particles, atoms, ions, and electrons	Dissipation and redistribution ofheat
	Structural stiffness		Compatibility of structural materials under thermal shock
	Operation of the launcher stages	Electromagnetic radiation	
		Radiation heating and cooling	Operation of control devices
Materials factors	Elastic modulus-density ratio	Behavior of solids at low pressures	Thermal stresses
	Rocket nozzle erosion, particularly nonuniform erosion	Radiation damage, effect on the physical properties of materials	Material survival under thermal shock at temperatures of ~8000 K (14,000°F)
	Low-temperature conditions in fuel tanks	Transparencies	Erosion
	Construction of solid-propellant rocket-motor casings	Emissivity	Ablation
		Radiation shielding	Thermal insulation
	Acoustic fatigue	Resistance to impact: surface damage, spalling	Heat sinks
			Protective coatings
		Behavior of bearing surfaces, friction	
		Seals	
		Behavior of electrical materials, semiconductors	
		Lightweight vehicle construction	
		Composite assemblies	
Materials of special interest	Graphite	Light alloys: aluminum, magnesium, titanium, beryllium	Graphite
	Refractory oxides	Polymers and elastomers	Plastics
	Carbides, nitrides, and borides		Ceramics, particularly reinforced ceramics
	Structural plastics		Molybdenum
	Steels		Niobium
	Nickel alloys		
	Cobalt alloys		
	Some structural low-density alloys		

*After A. J. Kennedy, *Materials Background to Space Technology*, George Newnes Limited Tower House, London, 1964.

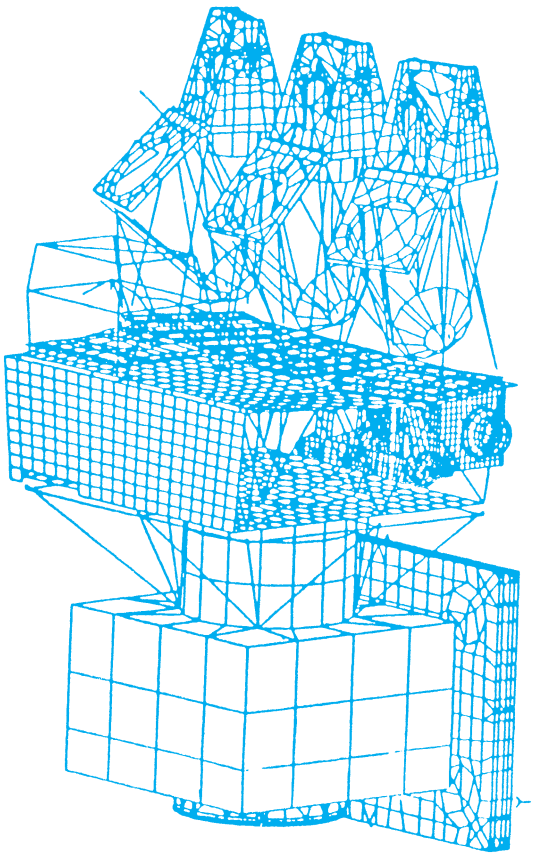

Fig. 1. Finite element model of the *SPOT 5* spacecraft in its stowed launch configuration. (*European Space Agency/ Matra Morconi*)

Testing. To ensure that spacecraft structures will meet mission requirements criteria in general requires testing to levels above the expected environmental conditions by a specific value. The test level must be set to provide for variations in materials, manufacture, and anticipated loads. In cases where structures are required to perform dynamic functions repeatedly, life testing is required to ensure proper operation over a given number of cycles of operation. *See* INSPECTION AND TESTING.

A major function of the test program is to validate the structural mathematical models. A typical model (**Fig. 1**) is constructed with a powerful finite element computer program such as NASTRAN or PATRAN. The model is used to predict dynamic loads, stress levels, natural frequencies in both the stowed and deployed configurations, on-orbit jitter and stability, temperature extremes and deformations, and operation loads for mechanisms. *See* COMPUTER-AIDED ENGINEERING; ENGINEERING DESIGN; FINITE ELEMENT METHOD.

Boost. The purpose of the boost phase is to lift the vehicle above the sensible atmosphere, to accelerate the vehicle to the velocity required, and to place the spacecraft at a point in space, heading in the direction required for the accomplishment of its mission. For space missions, the required velocities range from 26,000 ft/s (8 km/s) for nearly circular Earth orbits to 36,000 ft/s (11 km/s) for interplanetary missions. Achievement of these velocities requires boosters many times the size of the spacecraft itself. Generally, this boosting is accomplished by a chemically powered rocket propulsion system using liquid or solid propellants. Multiple stages are required to reach the velocities for space missions. Vertical takeoff requires a thrust or propulsive force that exceeds the weight of the complete flight system by approximately 30%. An example of a multiple-stage booster is the Delta launch vehicle used for the crewless missions (**Fig. 2**). The Delta II 7925 vehicle has the capability to place 4000 lb (1800 kg) into a geosynchronous transfer orbit. *See* INTERPLANETARY PROPULSION; PROPELLANT; ROCKET PROPULSION.

The boost trajectory (**Fig. 3**) can be controlled to some extent by proper selection of the thrust-to-weight ratio and launch angle. The higher thrust-to-weight ratios are more efficient, but they result in a higher dynamic pressure which in turn results in increased spacecraft loading and heating.

Boost environments. The basis for design of the primary spacecraft structure is the ability to withstand the boost environment. Typical launch vehicles produce longitudinal acceleration loadings in excess of five times the Earth's gravitational field, or 5 g. Lateral load factors are typically less than half of the longitudinal load factor, but cause the greatest portion of stress induced by the spacecraft base bending moment. A coupled-loads analysis, which is a dynamic simulation of the critical transient loading events, is performed in order to predict actual component

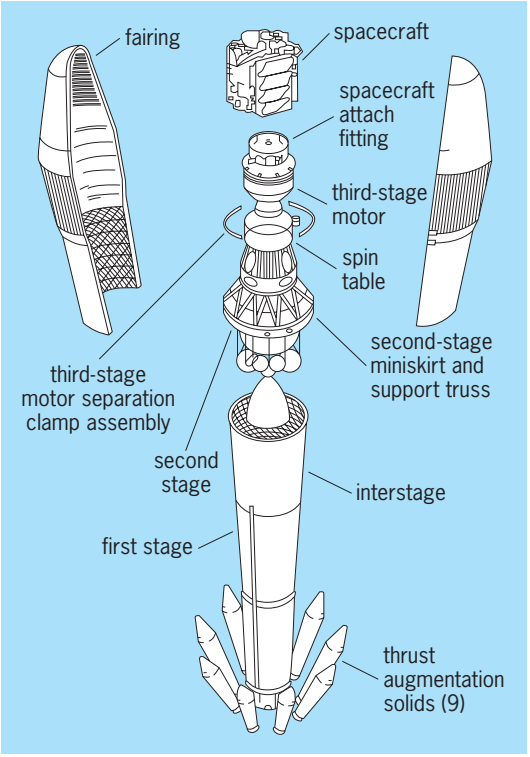

Fig. 2. Exploded view of typical Delta II 7925 three-stage structure. (*After Commercial Delta II Payload Planners Guide, McDonnell Douglas Commercial Delta Inc., 1990*)

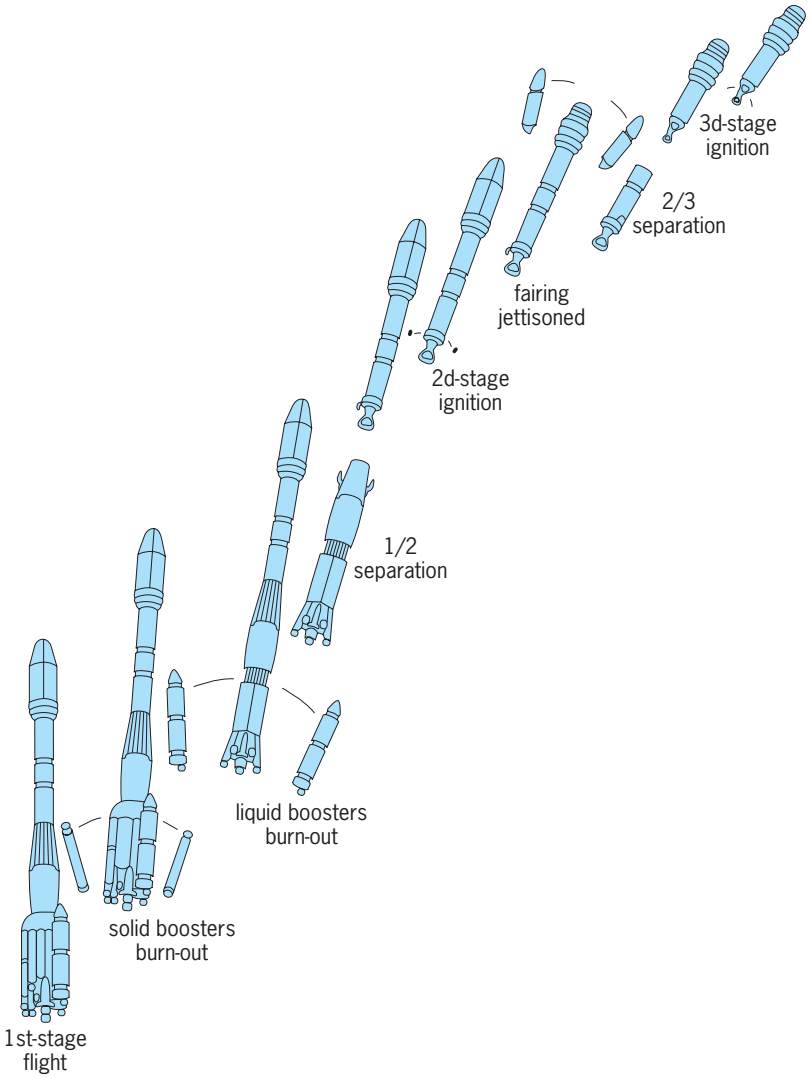

Fig. 3. Typical flight profile trajectory of the Ariane 4. (*Ariane 4 User's Manual*)

acceleration, force, stress, or displacement response for each unique launch configuration. Sophisticated finite element modeling techniques are used to represent the launch vehicle and spacecraft in such analyses.

Vibroacoustic loads. Spacecraft structures are exposed to an acoustic environment throughout the boost phase of flight until the vehicle is out of the sensible atmosphere. The highest acoustical sound-pressure levels occur immediately after liftoff because of impingement on the payload fairing of sound originating from engine exhaust energy which is reflected off the launch pad. Another significant acoustic event occurs during transonic flight and is caused by aerodynamic shock waves and a highly turbulent boundary layer. The sound-pressure level within the fairing creates random vibration of the spacecraft components. This acoustically generated random vibration environment is the most significant loading for small components which are typically mounted on panel-type secondary structures sensitive to high-frequency excitation. *See* BOUNDARY LAYER FLOW; SHOCK WAVE; SOUND; TRANSONIC FLIGHT; VIBRATION.

Space phase and design considerations. The space phase begins after the boost phase and continues until reentry. In this phase, the structures that were stowed for launch are deployed. Important design considerations include control system interaction, thermally induced stress, and minimization of jitter and creaks.

Deployment of loads. In order to fit within the tight confines of the boost vehicle, the spacecraft must be tightly folded and stowed. When in Earth orbit, it must extend and deploy into its operational shape (**Fig. 4**). This operation places high reliability requirements on mechanisms, motors, and joints. In space, the temperature changes can be extreme. The deployment may be complicated and involve pyrotechnic devices, coiled springs, redundantly driven motors, gearboxes, latches, and cables.

Extensive testing of the deployment is required because there are probably more failures of satellites from this phase than from any other. Of course, some space structures will not support themselves in Earth's gravitational field, so accurate deployment simulation can be challenging.

If an astronaut is involved, then the deployment operation must meet restricted force and access requirements. To assist the task, the astronaut is provided with a special tool that produces torque at several preset levels.

Control system. The spacecraft control system imparts inertial loads throughout the structure. In the zero gravity environment, every change in loading or orientation must be reacted through the structure.

The structural stiffness of the spacecraft must be considered, especially as it affects the control system. The control system sensor location and the higher-order modes must be considered so that control system coupling, which may result in overall spacecraft instability, is avoided. The liquid propellant rocket systems, using fairly low pressurization, result in thin-shell spacecraft structures which are not as rigid as solid propellant cases, which must withstand high internal pressures.

For orbiting telescopes, line-of-sight jitter and control is just as important as stability. As larger space structures are developed, they are increasingly difficult to control. This is due to two main factors: (1) their large size creates very low natural frequencies, which are slow to respond and are usually lightly damped, and (2) the structures often change configuration when telescopes point to new targets or solar arrays track the Sun. *See* CONTROL SYSTEMS.

Pressure vessels. Spacecraft structural design usually requires that part of the principal structure be a pressure vessel. Efficient pressure vessel design is therefore imperative. An important material property, especially in pressure vessel design, is notch sensitivity. Notch sensitivity refers to the material's brittleness under biaxial strain. This apparent brittleness contributed to premature failure of some early boosters. *See* PRESSURE VESSEL.

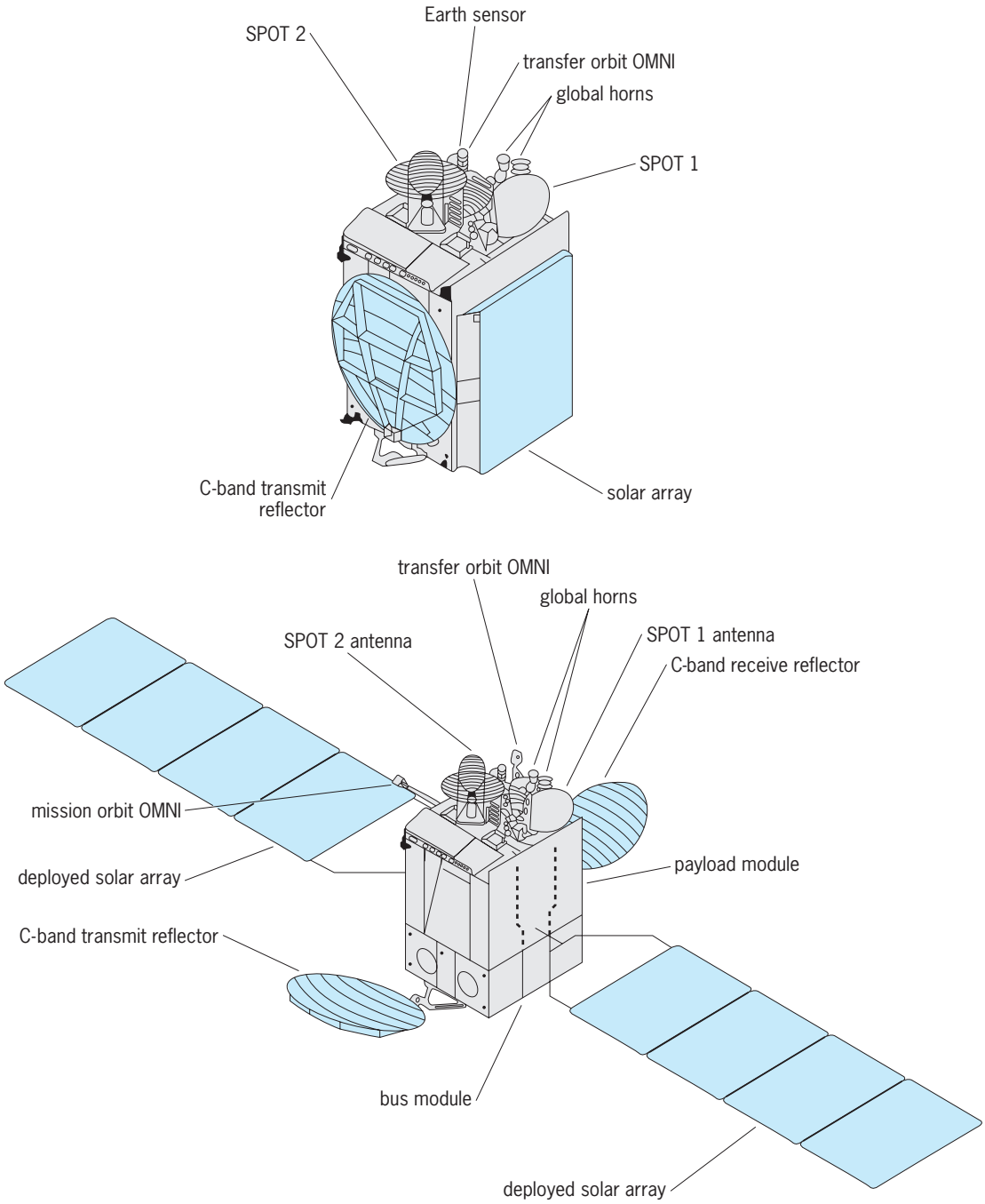

Fig. 4. ***INTELSAT VII*** **in its stowed and deployed configurations. (*Loral, LM Astro*)**

In pressure vessel design the variation of ultimate tensile strength with temperature is important. The exposure to elevated temperature is short.

Tank structures for spacecraft differ with the type of propellant used. With solid propellants the tank is primarily a pressure vessel; it must be able to contain the high pressures (500–1000 lb/in.^{2} or 3.4–6.9 megapascals) generated by combustion. Other loads are usually of secondary importance, and the bursting strength of the tank is usually the most significant factor. Care must be taken in design to avoid stress concentrations, particularly near joints, which might cause premature failure.

Tanks for liquid propellants, on the other hand, need to contain relatively low pressures (10–100 lb/in.^{2} or 69–690 kilopascals). However, flight loads in the form of thrust and bending moment become important in the design of the tank. One approach is to use the tank pressure to assist in carrying these loads, and to stabilize the light-gage unstiffened tank shell. The skin gage is chosen to be just adequate for containing the tank pressure stresses; the pressure is chosen on the basis of providing adequate stability under flight loads. Such pressures are generally in excess of those required to stabilize the tank to the point where the buckling strength is

developed. Alternative approaches are to provide stiffening in the form of stringers or frames to stabilize the tank, and lowering tank pressures only to that level required by the propulsion system.

Interstages between tanks are generally designed for flight loads; their construction is similar to that of airplane fuselages. They may be monocoque unstiffened shells or may have longitudinal (stringer) or circumferential (frame) stiffening.

Gravity gradient loads. Gravity gradient loads are particularly significant in long flexible structures located within a significant gravity field such as the Earth's. Properly considered, these loads can work for the spacecraft designer by providing a method of passive spacecraft control. An excellent example of the use of this technique was the radio astronomy explorer spacecraft (*RAE-Explorer 38*) which gravity-stabilized by use of its 750-ft (229-m) antennas (**Fig. 5**).

Solar radiation pressure. Solar radiation pressure can cause long flexible structural members, such as the *RAE* antennas and booms, to deflect quite significantly. Solar pressure can also cause significant disturbing torques on spacecraft that have large surface areas. This effect cannot be ignored in spacecraft structural design or in control systems design.

Space material and debris. Meteoroids are natural bodies remaining from the formation of the solar system or from collisions of other natural bodies. These particles may have extremely high velocities relative to the spacecraft (up to 225,000 ft/s or 68 km/s).

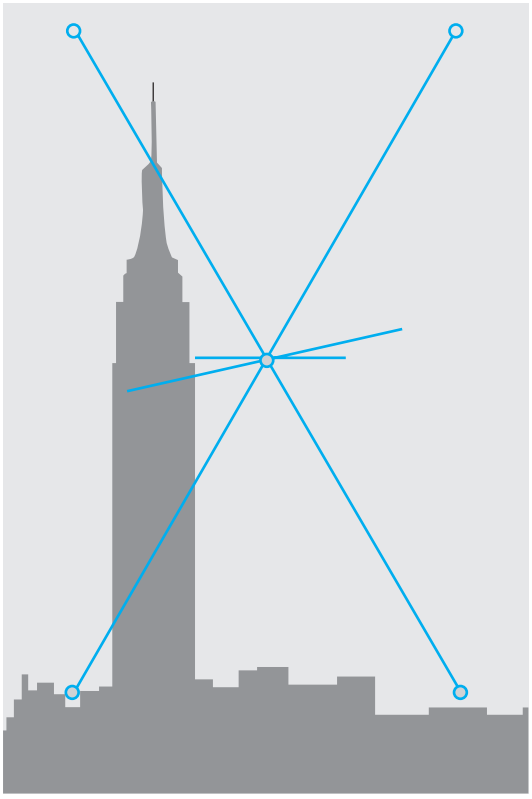

Fig. 5. Radio Astronomy Explorer spacecraft with antennas extended in comparison to the Empire State Building in New York City.

Orbital debris comprises residual particles resulting from human space-flight activities. Collisions involving these bodies and a space station and other long-duration orbiting spacecraft are inevitable. The worst-case effects of such collisions include the degradation of performance and the penetration of pressure vessels, including high-pressure storage tanks and habitable crew modules. An essential parameter in the design of these structures is the mitigation of these effects. *See* METEOR.

The probability of an object 0.4 in. (1 cm) or larger impacting any part of a structure with a 53,820-ft^2 (5000-m^2) exposed surface area (roughly one football field) and a 250-mi (400-km) operating altitude is predicted to be one impact in 71 years. However, an impact would not necessarily cause a failure in a space station since the core components of the largest space station designed cover only about 21,500 ft^2 (2000 m^2) and the design life is only 15 years. Further, space-station design features shielding to protect against objects as large as 0.55 in. (1.4 cm). This represents protection from 99.8% of the objects that would otherwise penetrate an unshielded component (roughly 0.04 in. or 0.1 cm). These factors make a catastrophic penetration extremely unlikely.

The great majority of the debris is far smaller than the millimeter scale. Impacts from these objects, most of which are the size of a grain of sand, too small to cause significant structural damage, are the most frequent. These impacts cause the degradation of sensitive surfaces such as optical components and solar panels. Mitigation of these effects can be dealt with by routine maintenance operations.

Another important consideration is protection for crew members during extravehicular activities (EVA). Presently the risk appears very small because of the relatively small exposed area of the space suit and short duration of the exposure. However, increases in the debris environment could pose operational and design constraints on extravehicular activity.

Energetic particles. Radiation shielding is required for some vehicles, particularly those operating for extended times within the Earth's magnetically trapped radiation belts or during times of high sunspot activity. The shielding may be an integral part of the structure. Computer memories are particularly susceptible to radiation and cosmic-ray activity and must be shielded to survive. This is a minor concern for scientific spacecraft but a major concern for defense satellites. Effects of radiation on most metallic structures over periods of 10–20 years is not severe. The durability of composite structures in space is a major concern for long life. Based on available data, the synergistic effects of vacuum, heat, ultraviolet, and proton and electron radiation degrade the mechanical, physical, and optical properties of polymers. *See* MAGNETOSPHERE; RADIATION HARDENING; VAN ALLEN RADIATION.

Thermal extremes. Temperatures in space can vary from 212°F (100°C) on sunlit surfaces to −292°F (−180°C) on dark surfaces. As a result, temperature

is a major design consideration for all moving parts, such as motors, hinges, and valves. When a satellite enters or exits the Earth's shadow, the rapid temperature change can induce a phenomenon called thermal jitter, where parts of the structure deform so fast that they oscillate about their neutral positions and shake the entire spacecraft. This happened in the early days of the *Hubble Space Telescope*. It was corrected with updated control software and new solar arrays. *See* HUBBLE SPACE TELESCOPE.

Temperature extremes in the structure and the enclosed environment are controlled by several techniques. Passive control is accomplished by surface coatings and multilayer thermal blankets which control the radiation transfer from the spacecraft to space and vice versa. Because incident solar radiation varies inversely with the square of distance from the Sun, means of adjusting surface conditions are required for interplanetary missions. Heat generated by internal equipment or other sources must be considered in the heat balance design. Other techniques used to actively control spacecraft temperatures are thermal louvers and heat pipes. *See* HEAT PIPE.

Thermal gradients must be considered in spacecraft design, especially when the spacecraft has one surface facing the Sun continuously. In some cases it is desirable to slowly rotate the spacecraft to eliminate such gradients.

Inertial loads. An important consideration in spacecraft structural design is the distribution of weight which determines the balance and inertia ratios of the spacecraft. In general the structural weight of spacecraft ranges 15–30% of the total spacecraft weight. By proper design, additional weight needed to balance the spacecraft and obtain the proper inertia ratios can be minimized. For spin-stabilized spacecraft the spin moment of inertia should be greater than the principal lateral moment of inertia by approximately 5% in order to ensure stability about the spin axis. Consideration must be given to expendables such as jettisoned hardware and gas depletion. As spacecraft become larger and the equipment contained therein becomes more massive, inertia loading is increasingly important.

Sloshing. Sloshing in the case of liquid propulsion systems must be carefully evaluated because it affects the structural integrity of the thin shell and the overall stability of the spacecraft. The fuel mass movements result in shifts of the center of gravity, causing pitch or yaw of the system. Control system coupling with either pitch or lateral sloshing can occur unless proper baffling or damping is provided.

Material considerations. The materials of general interest used in spacecraft structures are alloys of aluminum, magnesium, and titanium; steel; stainless steel; superalloys; plastics; and graphite-epoxy, Kevlar, and fiberglass composites. The fabrication and design techniques that are employed to obtain the most efficient structure for a given job often include honeycomb sandwich construction (**Fig. 6**), helical wound fiberglass and metal, expanded plastic sandwich construction, and woven fabrics (**Fig. 7**).

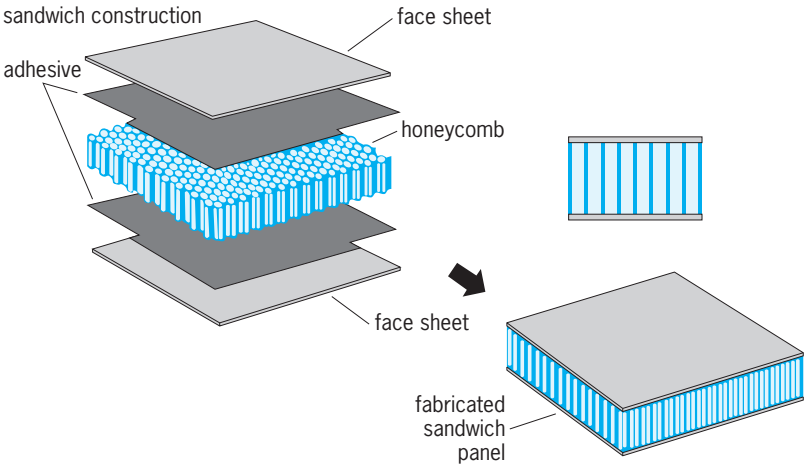

Relative stiffness (D)	100	700	3700
Relative strength	100	350	925
Relative weight	100	103	106

Fig. 6. Honeycomb sandwich construction provides strong and stiff but lightweight structures. (*Hexcel Corp.*)

Spacecraft structures usually are required to be lightweight and rigid, which results in the selection of high-modulus materials.

Titanium and beryllium have low densities and relatively high modulus-density ratios. Alloys of these metals are relatively difficult to fabricate, and

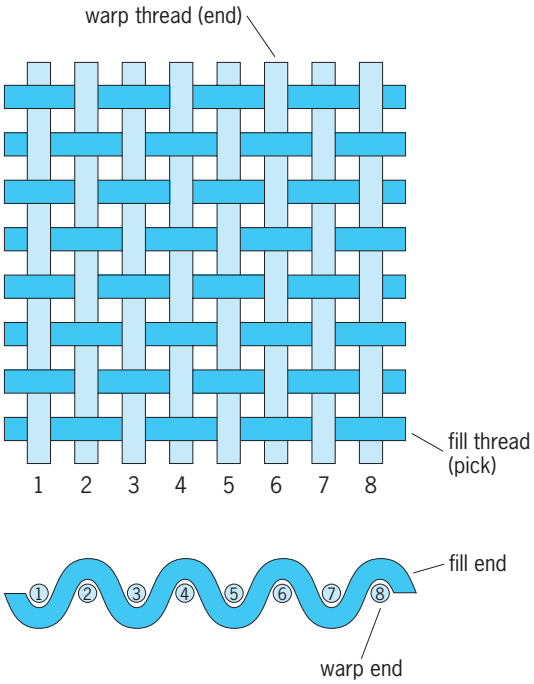

Fig. 7. Woven Kevlar fabric is light and strong.

therefore their application is quite limited. The more common aluminum, magnesium, and stainless steel alloys are basic spacecraft structural materials. They are easy to fabricate, relatively inexpensive, and in general quite suitable for use in the space environment.

Plastics are used in spacecraft structures when radio-frequency or magnetic isolation is required. They are also used in situations where some structural damping is desired. The thermal blankets, mentioned earlier as passive temperature control methods, are primarily Mylar plastic.

Composite materials. The modern requirements for low weight, high strength, high stiffness, and low thermal expansion (for precision optical pointing) have prompted the use of composite materials for spacecraft structure. These materials consist of high-strength reinforcement fibers which are supported by a binder material referred to as the matrix (Fig. 7). The fibers are typically made of glass, graphite, or carbon, and the matrix is an epoxy resin. The material is usually produced with the fibers configured in a unidirectional ply which has been pre-impregnated with the matrix. Such plies exhibit tremendous strength in the fiber direction but are relatively weak in the transverse direction. As a result, the plies are bonded together to form a laminate in which the fiber angle in each unique ply can be specified to provide the desired combined properties in each direction. The combined laminate is cured in an autoclave under stringent temperature and pressure conditions to ensure constant properties. With knowledge of the loading due to a unique application, the material properties can be tailored by optimizing the ply angles within the laminate to produce a material which may be substantially less dense than an equivalent-strength metallic structure.

An example of a spacecraft structure relying on composite materials is the Earth Observation System (EOS-AM1). The graphite-epoxy tubes composing the primary structure have been optimized for the substantially compressive loading induced by the Atlas launch vehicle. The tubes are made from a 70-ply rolled laminate consisting of 80% plies in the axial direction for high tensile and compressive strength and 20% plies in the 45° orientation for high torsional strength. Further, the EOS-AM1 instruments require precision pointing accuracies in the range of 0.1°. Under the extreme temperature gradients imposed on the spacecraft because of its polar orbit, the thermally induced expansion would preclude the use of typical metals for the primary structure. The thermal expansion coefficient for the graphite-epoxy tubes is 1.8 microstrain/°F (3.2 microstrain/°C) in comparison to 12.0 microstrain/°F (21.6 microstrain/°C) for typical aluminum alloys. Graphite-epoxy laminate is also used in the *Hubble Space Telescope*, where the laminates are arranged to minimize the coefficient of thermal expansion, preventing deformation when temperature changes occur (**Fig. 8**). In the *Hubble*, only 1 milliarcsecond (3×10^{-7} degree) of line-of-sight angular deformation is allowed for thermal expansion. For spacecraft requiring high precision in pointing angle, graphite-epoxy composites are the materials of choice. *See* COMPOSITE MATERIAL.

Reentry phase. Although the atmospheric layer of the Earth is relatively thin, it is responsible for the reduction of vehicle velocity and the resulting deceleration loads, as well as for the severe heating experienced by reentering vehicles. A body entering the Earth's atmosphere possesses a large amount of energy. This energy must be dissipated in a manner which allows the reentering vehicle to survive. Most of the vehicle's original energy can be transformed into thermal energy in the air surrounding the vehicle, and only part of the original energy is retained in the vehicle as heat. The fraction that appears as heat in the vehicle depends upon the characteristics of the flow around the vehicle. In turn, the flow around the vehicle is a function of its geometry, attitude, velocity, and altitude. Several different approaches are used for reentry vehicles with different trajectories (**Fig. 9**). *See* ATMOSPHERIC ENTRY.

Atmospheric drag. Spacecraft are seldom designed to reenter the Earth's atmosphere (the space shuttle being an exception), but may be designed to enter extraterrestrial atmospheres (**Fig. 10**). In either case, the structural design is similar.

High-speed reentry causes extreme friction and heat buildup on spacecraft that must be dissipated by using high-temperature ceramic or ablative materials. The space shuttle has reentered the Earth's atmosphere over a hundred times without accident (**Fig. 11**). It is covered with special thermal insulating tiles that allow the structural elements to remain cool when the surface reaches 1200°F (650°C), and its leading edges are protected by a carbon-carbon reinforced material that can withstand temperatures as high as 2300°F (1260°C).

Satellites whose orbits decay into the Earth's upper atmosphere become flaming objects as they rapidly descend. Generally, most or all of the satellite is consumed before it reaches the surface, but there are exceptions such as the March 22, 2001, reentry of the Russian space station, *Mir*.

Ablation. The normal application of ablation is to limit the temperature of the structural shell or internal components of the payload, in order to withstand the extreme thermal environment during ballistic zero-lift reentry. This temperature limitation is the principal criterion. The most desirable material, therefore, is one which has a high heat of ablation and low thermal conductivity. Material properties, however, do not necessarily possess these characteristics simultaneously, to the same degree (**Table 2**). Therefore, specific investigations or trade studies must be made to determine the minimum weight of ablation material required to maintain the structural temperature within design limits. Care must be exercised to match the ablation material to the environment. For example, if the maximum temperature that would be reached without ablation is below 3000°F (1650°C), quartz would not be considered because it does not ablate appreciably below 3000°F. Many materials will rapidly deteriorate without ablating if

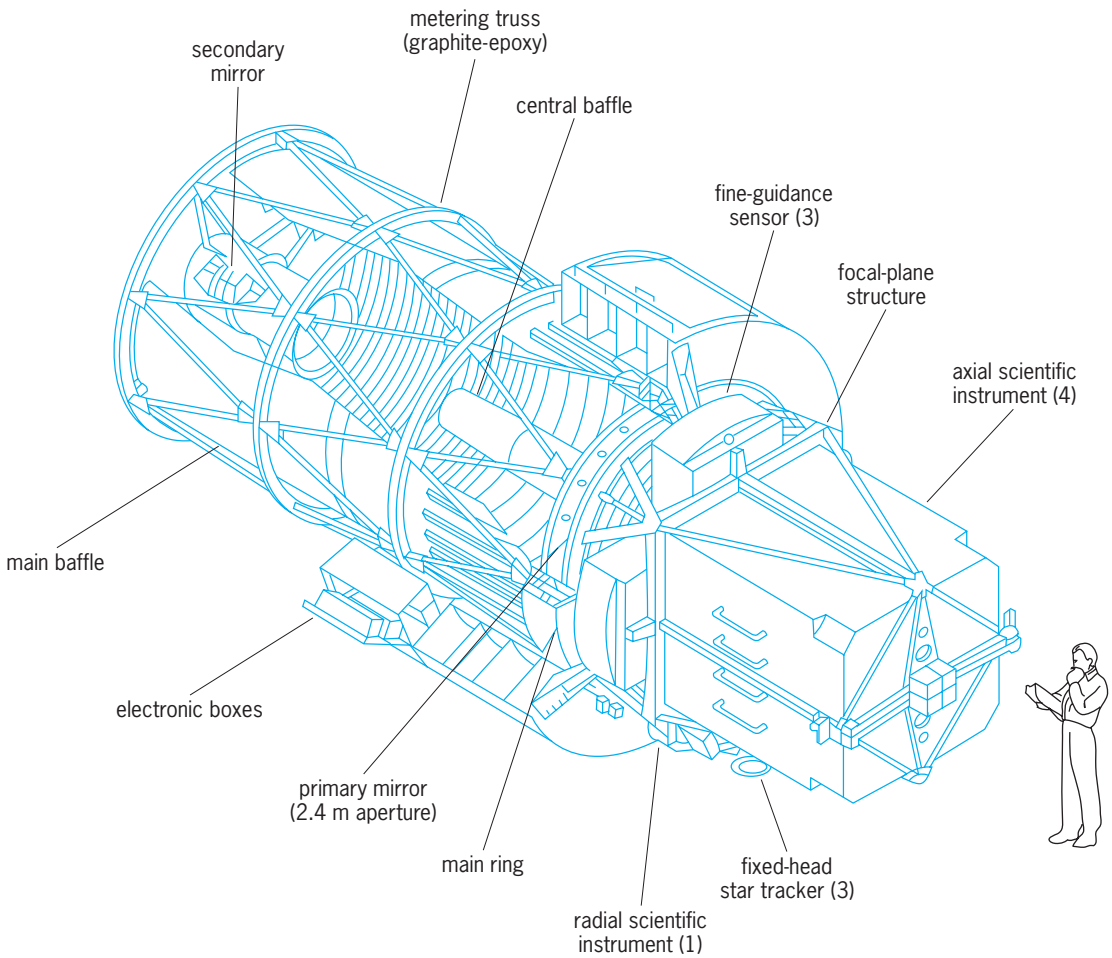

Fig. 8. Internal telescope of the *Hubble Space Telescope*, whose structural members are made of graphite-epoxy. (*NASA/Lockheed*)

exposed to temperatures below their ablation temperatures, and thus will not provide the needed protection for the structure.

Cooling and transpiration. Cooling the vehicle by circulation of water, lithium, or other media, or by transpiration of gas or vapor is limited by system complexity, reliability, and adverse weight penalties.

Thermal protective systems are time-limited because the heat-absorption capacity of the protection system is continuously expended during the process.

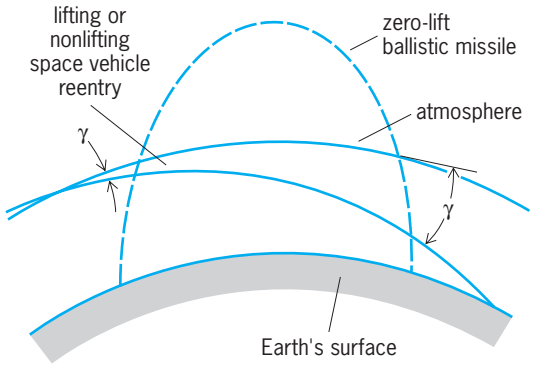

Fig. 9. Comparison of various space reentry and ballistic missile trajectories.

These systems, therefore, apply primarily to short-duration (5–10 min) flights.

Lifting reentry. In crewed applications, vehicles employing aerodynamic lift during reentry have several advantages over zero-lift ballistic bodies: (1) The use of lift allows a more gradual descent, thus reducing the deceleration forces on both vehicle and occupants. (2) The vehicle's ability to glide and maneuver within the atmosphere gives it greater accuracy in either hitting a target or landing at a predetermined spot. (3) It can accommodate greater errors of guidance systems because for a given deceleration it can tolerate a greater range of entry angles. (4) Greater temperature control is afforded because aerodynamic lift may be varied to control altitude with velocity.

The total time required for a lifting vehicle to descend through the atmosphere depends on the drag characteristics of the vehicle. With high drag coefficients the time may be as low as 30 min, and with low drag it may be approximately 1–2 h. In the case of grazing reentries, several orbital loops may be required to bleed off the energy, and reentry duration can be measured in days. The relatively long times required for descent make time-dependent ablation or cooling impractical.

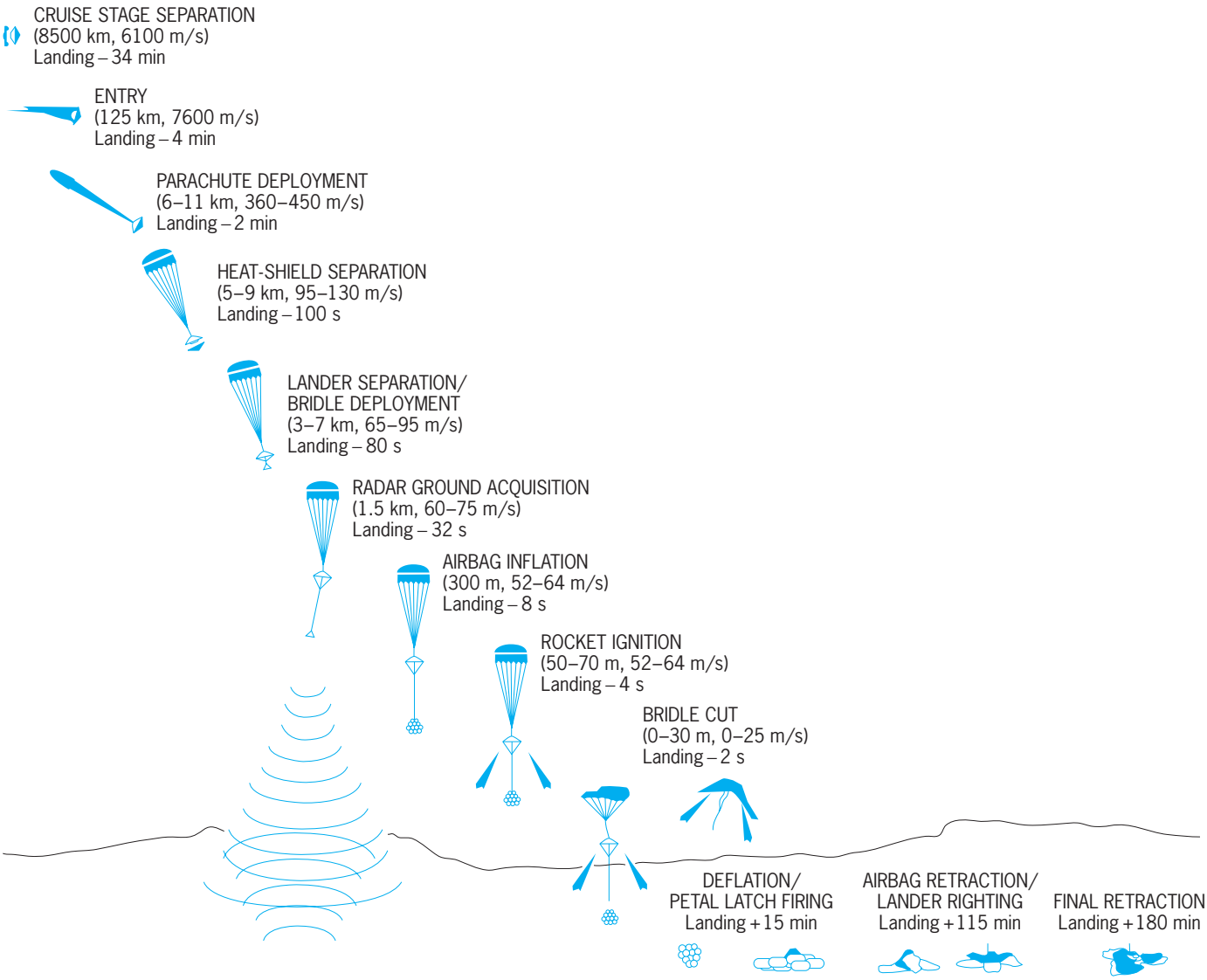

Fig. 10. ***Mars Pathfinder*** **entry, descent, and landing. (*NASA/Jet Propulsion Laboratory*)**

Thermal stress during reentry. One of the most severe structural problems is thermal stress, which arises as a result of temperature gradients. These temperature gradients result from causes such as heat sinks, transient heating, and crewed-compartment cooling. Thermal stress is a function of the coefficient of expansion, proportional to the modulus of elasticity and the temperature gradient.

Thermal stresses will be produced in sandwich or conventional sheet-stringer construction if the thermal deformations are externally constrained. Thermal stresses will also be produced in the absence of external constraints, solely because of the incompatible deformations of the different parts of the body, for example, a rectangular anisotropic sandwich panel with a temperature gradient through its thickness.

The elastic thermal stresses are computed for the typical temperature gradients associated with the maximum temperatures. If the vehicle is intended for repeated use, the compressive yield point cannot be exceeded. As an example, sandwich construction of high-nickel alloy M-252 cannot be used above 1660°F (900°C) because the thermal stress equals the compressive yield stress at that temperature. The air-load stresses should be added to the thermal stresses, thus reducing even further the allowable temperature for material and type of construction under these conditions. The same curves are drawn for a titanium-molybdenum alloy. Its more favorable lower coefficient of thermal expansion and its higher allowable stress in the higher-temperature region result in a better sandwich. However, there are many practical problems to be solved in the brazing of molybdenum alloys.

If the vehicle were designed for a single shot, the thermal stresses could exceed the yield point considerably without seriously jeopardizing the mission. However, tremendous benefits accrue if thermal stresses are eliminated.

High-temperature materials. The extreme thermal environment requires materials with higher temperature

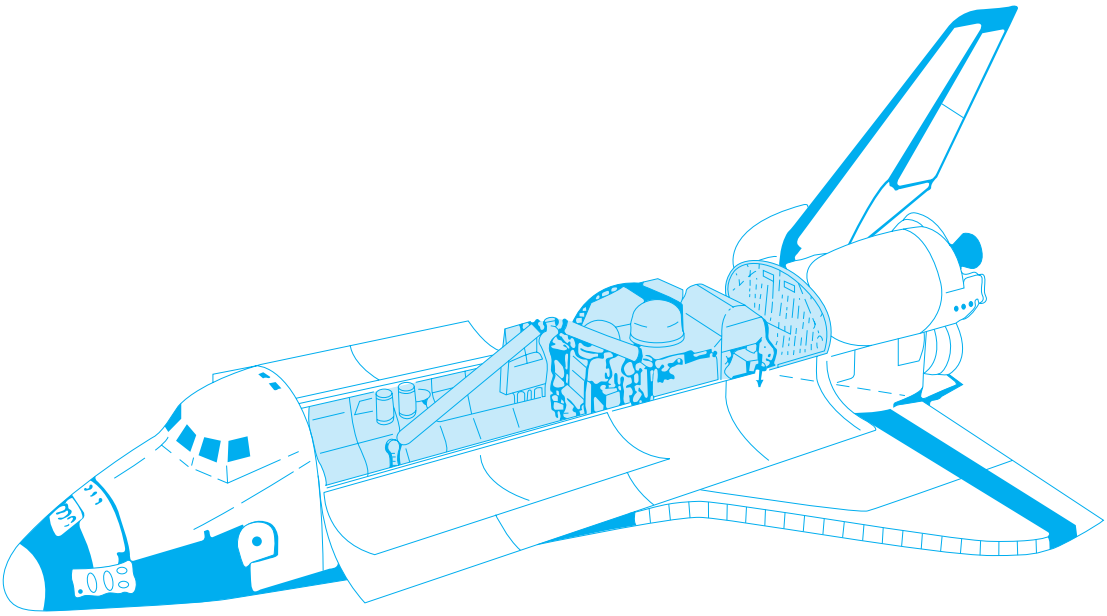

Fig. 11. The space shuttle (shown with cargo bay doors open) can return to Earth with help from insulation tiles that keep it from burning up. (*NASA/Johnson Space Center*)

capability than standard aircraft steels and aluminum alloys. These materials are the superalloys, refractory alloys, and ceramics.

The superalloys which contain high percentages of either nickel or cobalt have a maximum useful temperature range, under low stresses, of 1600–2000°F (870–1090°C). These alloys are readily available in thin gages and can be formed, riveted, brazed, and fusion- or spot-welded with essentially standard procedures. The surface of these alloys is rapidly oxidized above 1300°F (700°C). However, the oxide forms an adherent coating that protects the metal from further oxidation and provides a high surface emissivity (0.8–0.9).

The refractory alloys (niobium, molybdenum, tantalum, and tungsten) have higher densities than the superalloys but retain their useful strength from 3300°F or 1820°C (niobium) to over 4000°F or 2200°C (tungsten). All refractory metals are oxidized catastrophically at high temperatures, and surface protection is required. Coatings have been developed which protect molybdenum for approximately 300 h at 3000°F (1650°C) and niobium for several hours at 2600°F (1430°C). Niobium alloys are available in thin gages and can be formed by standard procedures. The quality of thin-gage molybdenum alloys is not consistent and requires hot forming because of their low ductility at room temperature. Tungsten alloy, available in sheet, bar, and wire form, has the highest melting point of all the metallic elements. This material offers extremely high strength but at a high mass density of 19 g/cm^3. Tantalum is similar to niobium except that it has a much higher density. The refractory alloys can be joined by riveting, but fusion welding requires such special procedures as inert atmosphere. Spot welding is extremely difficult.

Conventional ceramics retain fairly high strengths up to 4000°F (2200°C), and the more exotic ceramics are potentially usable at over 8000°F (4430°C). Their high melting points and relative inertness to oxidizing atmospheres make them attractive materials for leading edges and nose cones of glide vehicles. However, the lack of ductility presents serious design problems.

Cost is not necessarily the criterion for material selection. Availability, strength, stiffness, weight, producibility, and cost of fabrication must also be considered. The final material selected will be the one which results in the least expensive end item

TABLE 2. Typical ablation materials and properties

Material	Ablation temperature, °F (°C)	Thermal conductivity, Btu/(in.)(s)(°F) [J/(m)(s)(°C)]	Specific heat, Btu/(lb)(°F) [kJ/(kg)(°C)]	Density, lb/in.^{3} (g/cm^3)
Quartz	4400 (2430)	0.1292×10^{-4} (0.9660)	0.25 (1.05)	0.08391 (2.323)
Acrylic	250* (120)	0.4033×10^{-3} (30.15)	0.38 (1.59)	0.04282 (1.185)
Poly(tetrafluoroethylene)	1000* (540)	2.896×10^{-6} (0.2165)	0.25 (1.05)	0.07523 (2.082)
Refrasil–11% phenolic resin	2000° (1090)	3.475×10^{-6} (0.2598)	0.23 (0.96)	0.06366 (1.762)

*Estimated.

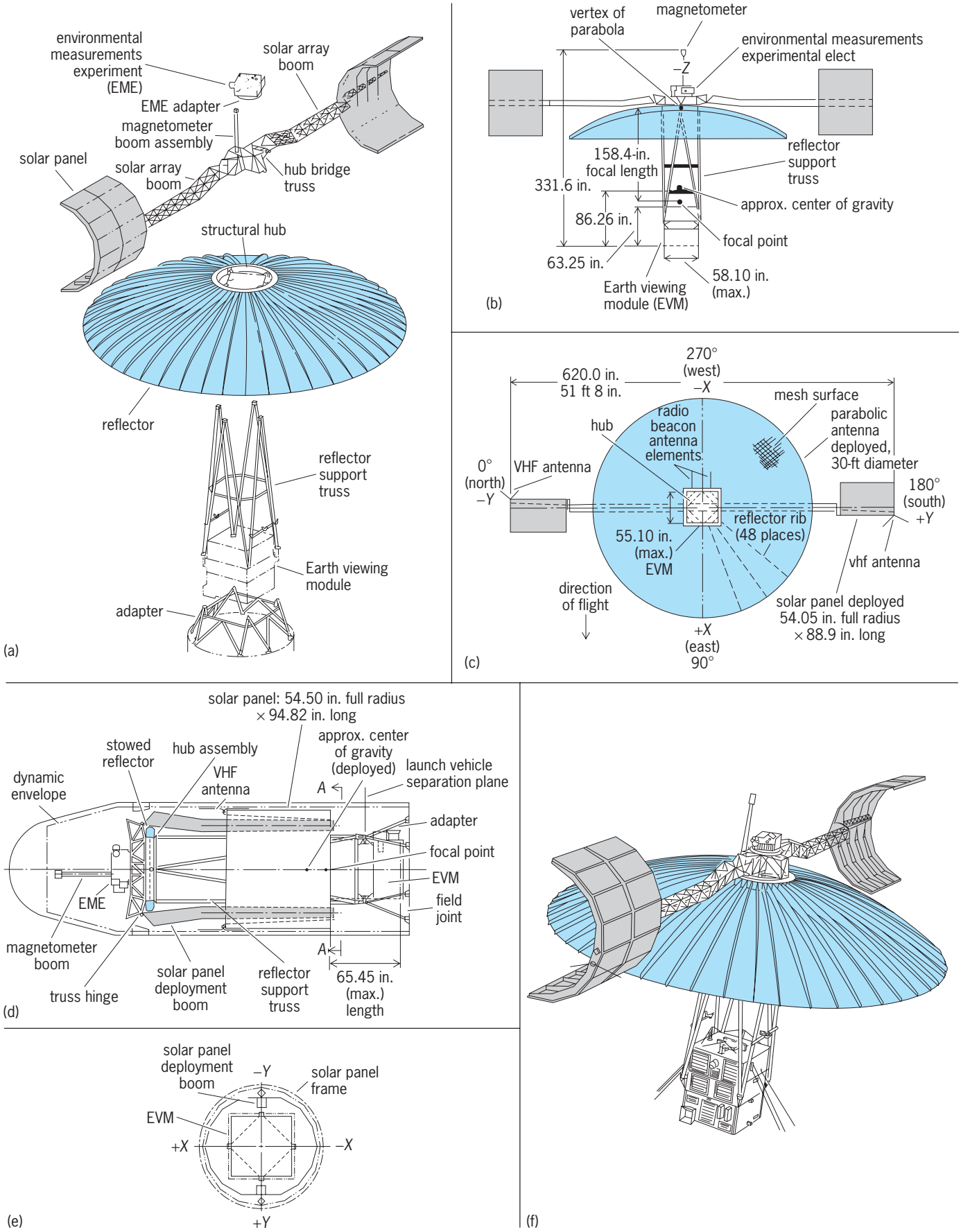

Fig. 12. ***ATS 7.*** **(*a*) Structural assemblies. (*b*) Side view and (*c*) bottom view of deployed configuration. (*d*) Side view and (*e*) section A-A of launch configuration. (*f*) View of deployed spacecraft. 1 in. = 25 mm; 1 ft = 0.3 m.**

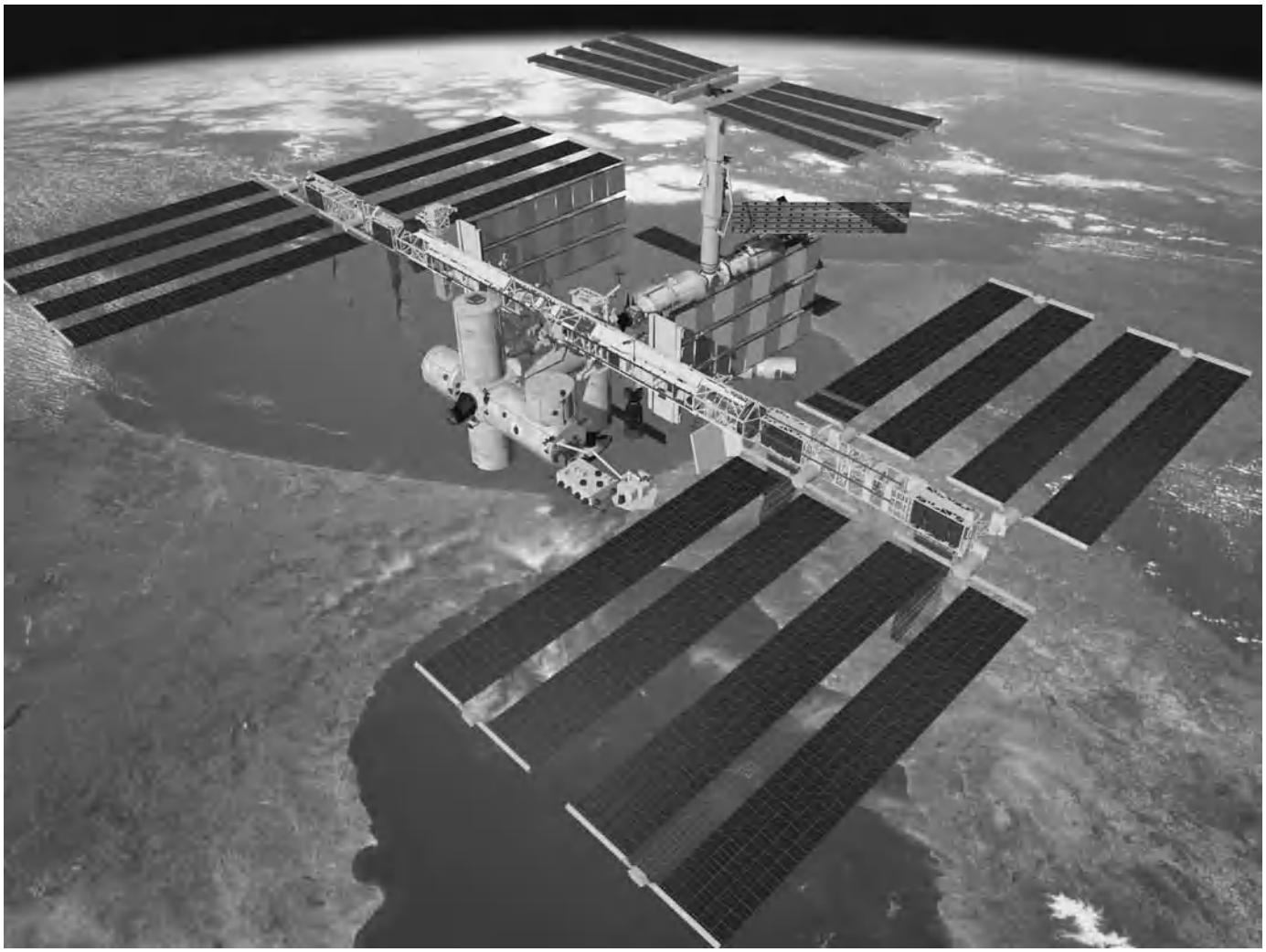

Fig. 13. The *International Space Station* is an example of a large structure of changing configuration that cannot be completely assembled on Earth. (*NASA*)

and best meets the design requirements. *See* HIGH-TEMPERATURE MATERIALS.

Erectable structures. These take many and varied shapes. They are sometimes relatively simple hinged booms, while on other occasions they become quite large and massive. One type of unique erectable structure is the antennas on the radio astronomy explorer (*RAE*) spacecraft (Fig. 5). The four 750-ft (229-m) V antennas were erected in flight, as were two 300-ft (91-m) damper tubes and two 60-ft (18-m) dipole antennas. Each of these antenna booms was stored as a 0.002-in.-thick (50-micrometer) beryllium copper alloy tape approximately 2 in. (5 cm) wide in a mechanism within the spacecraft. Upon commands from the Earth, the booms were deployed from the mechanism to form $^1/_2$-in.-diameter (13-mm) tubes. The weight of the $^1/_2$-in.-diameter (13-mm) tube is approximately $^1/_4$-oz/ft (23 g/m); therefore the 3750 ft (1143 m) of elements aboard the *RAE* weighed only 60 lb (27 kg). The 750-ft (225-m) antennas were designed so that upon deployment the flat tape interlocked into a torsionally rigid structural member. In addition to serving the function of radio astronomy antennas, the long booms provided gravity stabilization so that one axis of the spacecraft always faced the Earth.

Rigid spacecraft structures. Many more spacecraft structures are rigid than erectable or inflatable. Some examples of rigid spacecraft structures are the *Orbiting Astronomical Observatory* (*OAO*), the *Nimbus* weather spacecraft, the *Telstar* communications spacecraft, the *Syncom* communications spacecraft, the *Orbiting Solar Observatory* (*OSO*), and the lunar and command service modules.

Typical of a rigid spacecraft structure is the *Applications Technology Satellite* (**Fig. 12**), which includes a flexible, erectable, parabolic reflector 30 ft (9 m) in diameter. It is used to broadcast television signals into remote areas.

Another spacecraft structure is the space shuttle (Fig. 11) or Space Transportation System (STS), which can carry 65,000 lb (30,000 kg) of cargo to and from low Earth orbit.

The *International Space Station* (*ISS*) is a cooperative, 16-nation effort (**Fig. 13**). It will include six laboratories and weigh a million pounds when

assembled as currently projected in 2010. *See* SPACE STATION.

The *Viking* lander is a prime example of an interplanetary spacecraft designed for orbiting and soft landing on Mars. *See* COMMUNICATIONS SATELLITE; METEOROLOGICAL SATELLITES; MILITARY SATELLITES; SATELLITE NAVIGATION SYSTEMS; SCIENTIFIC AND APPLICATIONS SATELLITES.

William Haile; Peter S. Weinberger

Bibliography. K. G. Budinski and M. K. Budinski, *Engineering Materials: Properties and Selection*, 1998; V. A. Chobotov, *Spacecraft Attitude Dynamics and Control (Orbit, a Foundation Series)*, 1991; P. Conley (ed.), *Space Vehicle Mechanisms: Elements of Successful Design*, 1997; D. Jenkins, *Space Shuttle: The History of the National Space Transportation System: The First 100 Missions*, 2001; R. Meyer, *Elements of Space Technology for Aerospace Engineers*, 1999; R. Neri, *Manned Space Stations: Their Construction, Operation, and Potential Applications*, 1990; T. Sarafin (ed.), *Spacecraft Structures and Mechanisms, from Concept to Launch*, 1995.

Spallation reaction

A nuclear reaction that can take place when two nuclei collide at very high energy (typically 500 MeV per nucleon and up), in which the involved nuclei are either disintegrated into their constituents (protons and neutrons), light nuclei, and elementary particles, or a large number of nucleons are expelled from the colliding system resulting in a nucleus with a smaller atomic number. This mechanism is clearly different from fusion reactions induced by heavy or light ions with modest kinetic energy (typically 5 MeV per nucleon) where, after formation of a compound nucleus, only a few nucleons are evaporated. A spallation reaction can be compared to a glass that shatters in many pieces when it falls on the ground. The way that the kinetic energy is distributed over the different particles involved in a spallation reaction and the process whereby this results in residues and fluxes of outgoing particles are not well understood. *See* NUCLEAR FUSION.

Reactions in space. Spallation reactions take place in interstellar space when energetic cosmic rays (such as high-energy protons) collide with interstellar gas, which contains atoms such as carbon, nitrogen, and oxygen. This leads to the synthesis of light isotopes, such as ^{6}Li, ^{9}Be, ^{10}Be, and ^{11}B, that cannot be produced abundantly in nucleosynthesis scenarios in the big bang or stellar interiors. *See* NUCLEOSYNTHESIS.

Laboratory reactions. In terrestrial laboratories spallation reactions are initiated by bombarding targets with accelerated light- or heavy-ion beams, and they are used extensively in basic and applied research, such as the study of the equation of state of nuclear matter, production of energetic neutron beams, and radioactive isotope research.

Heavy-ion reactions. When a spallation reaction occurs as the result of a heavy nucleus colliding at high impact energy with another heavy nucleus (for example, two lead nuclei with 82 protons and 126 neutrons), the kinetic energy of the bombarding nucleus is transformed into compression, thus increasing the density of the nuclear matter, and into thermal motion of the nucleons. New particles are thereby created, some of which escape the nuclear medium. By studying these escaping particles, information can be extracted concerning the equation of state of nuclear matter. As in ordinary gases, liquids, and solids, this equation of state gives information about properties of nuclear matter, such as the compressibility, as a function of density and temperature. The equation of state of nuclear matter is relevant to better understanding the creation of neutron stars and black holes. At extremely high bombarding energy, so high a temperature might be reached that the nucleons (protons and neutrons) would disintegrate into their constituents, known as quarks, and, together with the gluons that carry the strong interaction between the quarks—like small strings—form the so-called quark-gluon plasma. At present it is assumed that such a state was present at the very beginning of the universe, immediately after the big bang. *See* BIG BANG THEORY; QUARK-GLUON PLASMA; RELATIVISTIC HEAVY-ION COLLISIONS.

Light-ion-induced reactions. When a light nucleus (such as a proton) hits a heavy nucleus (such as lead) with an energy of a few hundred megaelectronvolts or more, spallation can also take place. The resulting neutrons that escape the reaction are intensively used in condensed-matter, chemistry, geology and biology research at different spallation neutron source facilities. *See* NEUTRON; NEUTRON DIFFRACTION; RADIATION DAMAGE TO MATERIALS; SLOW NEUTRON SPECTROSCOPY.

The light-ion-induced spallation reaction is less violent, and there is a nonnegligible probability that only a number of nucleons (10–40) will be removed from the target nucleus, resulting in a radioactive isotope. Such spallation reactions are used intensively to produce new isotopes. On the chart of nuclides, a representation of the different isotopes as a function of number of neutrons and protons, these radioactive isotopes can lie far away from the line of stable isotopes and have very short half-lives, on the order of a few milliseconds. Their study yields information about the manifestation of the electromagnetic, strong, and weak interactions in the nuclear medium, and their properties determine to a certain extent the nucleosynthesis processes taking place in explosive stellar environments. However, spallation of heavy atoms can lead to the production of more then 100 different nuclei. Therefore the desired isotopes must be separated from the debris. Electromagnetic mass spectrometers combined, for example, with laser ion sources are used, and numerous new isotopes have been discovered and studied. Several projects around the world have been set up to accelerate radioactive isotopes that are produced in proton-induced spallation reactions (among

others), in order to carry out studies of nuclear reactions that, in turn, involve these reaction products. *See* MASS SPECTROSCOPE; NUCLEAR REACTION; RADIOACTIVE BEAMS; RESONANCE IONIZATION SPECTROSCOPY. Piet Van Duppen

Bibliography. H. L. Ravn and B. W. Allardyce, On-line mass separators, in D. A. Bromley (ed.), *Treatise on Heavy Ion Science*, vol. 8, p. 363, Plenum Press, New York, 1989; C. E. Rolfs and W. S. Rodney, *Cauldrons in the Cosmos*, pp. 508–514, University of Chicago Press, 1988; R. Silberberg and C. H. Tsao, Spallation process and nuclear interaction products of cosmic rays, *Phys. Rep.*, 191:351–408, 1990.

Sparganosis

An infection by the plerocercoid, or sparganum, of certain species of the genus *Spirometra*. Members of this genus have a life cycle much like that of *Diphyllobothrium latum*. The adult normally occurs in the intestine of dogs and cats, the procercoid in copepods, and the plerocercoid in the musculature of frogs, snakes, or aquatic mammals, but not in fish. The whitish, unsegmented plerocercoid usually is an inch or so in length, but may reach a length of a foot or more. Humans may become infected accidentally by drinking water with infected copepods or by eating hosts containing plerocercoids. Since humans are not suitable hosts for the adult stage, the plerocercoids leave the gut and enter the musculature, body cavities, or other sites and remain as plerocercoids. In Eastern countries, freshly opened frogs are sometimes used as poultices on sores, particularly around the eyes, and if plerocercoids are present, they may invade the human tissues. Human sparganosis is rare in North America. *See* PSEUDOPHYLLIDEA. Reino S. Freeman

Spark gap

The region between two electrodes in which a disruptive electrical spark may take place. The gap should be taken to mean the electrodes as well as the intervening space. Such devices may have many uses. The ignition system in a gasoline engine furnishes a very important example. Another important case is the use of a spark gap as a protective device in electrical equipment. Here, surges in potential may be made to break down such a gap so that expensive equipment will not be damaged. The dependence of spark-gap operation on such factors as pressure, length, electrode characteristics, and time dependence of potential is quite complicated. *See* BREAKDOWN POTENTIAL; ELECTRIC SPARK. Glenn H. Miller

Spark knock

The sound and related effects produced in spark-ignited internal combustion engines by instantaneous ignition and combustion (autoignition) of the gaseous fuel–air mixture ahead of the advancing flame front. Spark knock is a complex physicochemical phenomenon.

After spark ignition, a flame travels outward from the spark plug and, under normal combustion, will progressively burn the entire fuel–air charge. The burned gas liberates heat and expands, leading to increased pressure and temperature in the unburned gas ahead of the flame front. In addition, cylinder pressure and temperature are increasing due to the upward, compressive motion of the piston. The result is that the unburned gas may be raised above its autoignition temperature. However, before autoignition occurs, time must be allowed for the chemical reactions that precede knock. If the flame front passes through the unburned gas before these reactions are completed, normal combustion results. If these reactions occur too quickly or if the flame front velocity is too small, the unburned gas spontaneously ignites and burns instantaneously (see **illus.**). While this process is often called detonation, it is not identical to the detonation of explosives. The instantaneous combustion results in a very intense pressure wave that is internally reflected within the cylinder at a characteristic frequency (5–9 kHz) related to the velocity of sound within the cylinder and the dimensions of the combustion chamber. The flexing of the cylinder wall and cylinder head produces the audible, high-frequency (pinging) sound known as spark knock. *See* COMBUSTION CHAMBER; EXPLOSIVE.

Besides sound, spark knock can result in pitting, or erosion, of the combustion chamber, damage to spark plug electrodes, and possible structural damage to the engine. Spark knock also leads to loss of engine efficiency by inducing spark plug preignition, resulting in overly advanced spark timing. Spark knock also causes intense turbulence within the cylinder, aggravating heat loss from the burned gas to the colder cylinder and head surfaces and reducing efficiency.

There are three general factors affecting the existence of spark knock: the effect of engine design, the effect of fuel, and the effect of engine operating variables.

Effect of engine design. Engine design variables influencing spark knock are compression ratio, combustion chamber design, spark plug location, and in-cylinder gas velocity. Designs that increase the temperature, pressure, and chemical residence time of the unburned gas (end gas) increase spark knock. Increased compression ratio, off-center plug location, and slow-burn combustion chambers all lead to increased spark knock. A faster-burning chamber, due to higher in-cylinder gas velocity, and central plug location increase knock resistance. Characteristics of faster-burning chambers include rotational motion (swirl) of the charge (due to off-cylinder axis charge admission); in-cylinder jetting action at the top of the compression stroke caused by nonuniform squeezing of the charge (due to proximity of portions of the piston and cylinder head); and use of two or more spark plugs. A spark plug should be

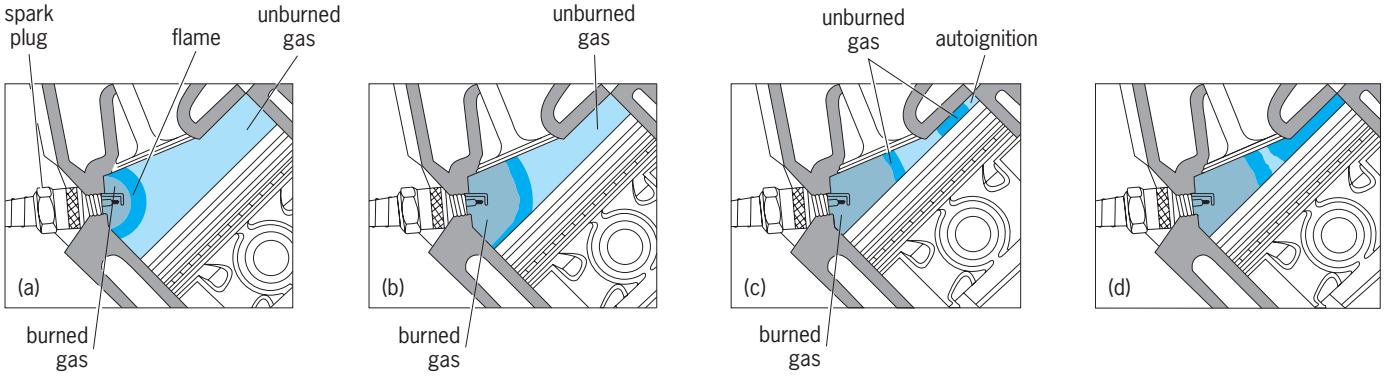

Sequence of spark knock. (*a*) Initial flame growth around the end of the spark plug. (*b*) Beginning of combustion and progression of flame through unburned gas. (*c*) Autoignition in the unburned end gas ahead of the flame. (*d*) Spark knock, or detonation. (*Champion Spark Plug Co.*)

located closer to the exhaust valve than the intake valve, since the former is hotter. Turbocharged and supercharged engines have increased spark-knock tendency because of higher intake-manifold air pressure. Stratified charge spark–ignited engines, having inhomogeneous air–fuel ratio distribution in each cylinder, have decreased knock tendency. *See* IGNITION SYSTEM; SPARK PLUG; SUPERCHARGER; TURBOCHARGER.

Effect of fuel. The fuel octane number represents its resistance to spark knock. Higher octane numbers are due to higher autoignition temperatures or longer end-gas chemical reaction times. Spark knock can be eliminated by substituting a high-octane fuel for one of lower octane. *See* OCTANE NUMBER.

Either fuel structure or fuel additives determine octane number. More compact hydrocarbon molecules have higher octane numbers than do long-chain molecules. For many years, one of the most popular and effective antiknock additives was tetraethyllead. However, tetraethyllead use for new automobile engines sold in the United States has been substantially eliminated because of lead's ability to poison platinum catalysts. Platinum-based catalysts are required for aftertreatment of HC, CO, and NO_x exhaust emissions to render them harmless, as required by federal regulations. With the introduction of catalysts, engine compression ratios were reduced to permit extensive use of lower-octane nonleaded gasoline. Increasingly, nonleaded octane improvers, such as ethanol, methanol, and methyl butyl tertiary ether, are added to refinery gasoline. However, considerably larger amounts of these than of tetraethyllead are required to provide equivalent antiknock. *See* ALCOHOL FUEL; GASOLINE.

Effect of operating variables. Engine operating variables, such as spark advance, revolutions per minute, throttle angle, coolant temperature, intake air temperature and humidity, and air/fuel ratio, can have a major influence on spark knock. To reduce or eliminate spark knock, spark advance can be reduced (retarded), revolutions per minute reduced, throttle angle decreased, coolant temperature decreased, intake air temperature reduced, intake air humidity increased, and air/fuel ratio reduced.

Many automotive engines, including most supercharged and turbocharged spark-ignition engines, have a piezoelectric knock sensor. When detonation occurs, the sensor sends a small signal voltage to the detonation-control system or the electronic engine control system. This retards the ignition timing until the detonation stops. On some supercharged or turbocharged engines, the detonation-control system reduces the boost pressure in the intake manifold. On other supercharged or turbocharged engines, detonation is controlled by first retarding ignition timing and then reducing boost pressure. *See* AUTOMOTIVE ENGINE; COMBUSTION; INTERNAL COMBUSTION ENGINE. Joseph R. Asik; Donald Anglin

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; A. By, B. Kempinski, and J. M. Rife, *Knock in Spark Ignition Engines*, Soc. Automot. Eng. Pap. 810147, 1981; E. F. Obert, *Internal Combustion Engines and Air Pollution*, 1973; C. F. Taylor, *The Internal Combustion Engine in Theory and Practice*, 2d rev. ed, vol. 2, 1984.

Spark plug

A device that screws into the combustion chamber of an internal combustion engine to provide a pair of electrodes between which an electrical discharge is passed to ignite the combustible mixture. The spark plug consists of an outer steel shell that is electrically grounded to the engine and a ceramic insulator, sealed into the shell, through which a center electrode passes (see **illus.**). The high-voltage current jumps the gap between the center electrode and the ground electrode fixed to the outer shell.

The electrodes are made of nickel and chrome alloys that resist electrical and chemical corrosion. Some center electrodes have a copper core, while others have a platinum tip. Many spark plugs have a resistor in the center electrode to help prevent radio-frequency interference. The parts exposed to the combustion gases are designed to operate at temperatures hot enough to prevent electrically conducting deposits but cool enough to avoid ignition of the

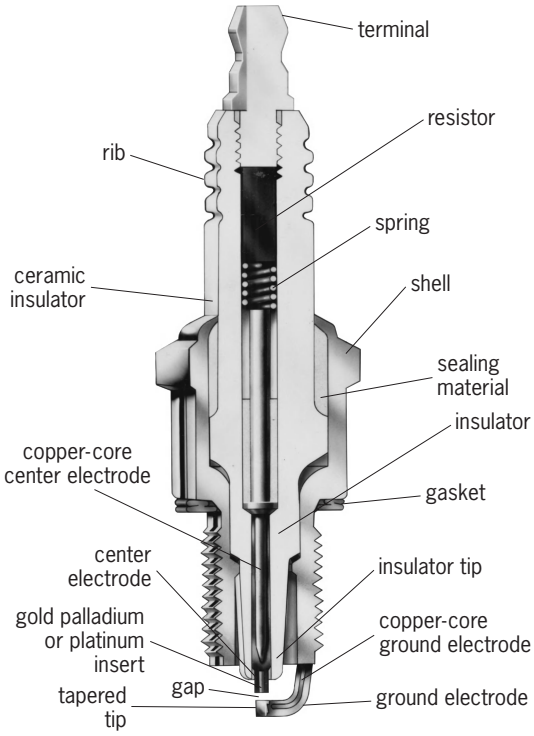

Cross section of a typical spark plug. (*Champion Spark Plug Co.*)

mixture before the spark occurs. *See* IGNITION SYSTEM. Donald L. Anglin

Spatangoida

An order of exocyclic Euechinoidea in which the posterior ambulacral plates form a shield-shaped area behind the mouth. The plates are arranged in two similar parallel longitudinal series. The structure is termed an amphisternous plastron. Teeth are lacking; phyllodes develop, but without bourrelets. Thus there is no floscelle. The apical system is compact. The families are defined mainly by reference to the fascioles, which are ribbonlike bands of minute, close-set uniform ciliated spinules on various parts of the test (see **illus.**). According to their position, they are defined as anal, subanal, inner (surrounding the apical system), peripetalous (surrounding the petals), or marginal. The earliest forms were the Toxasteridae of the Lower Cretaceous, which lacked fascioles and petals. The surviving Palaeopneustidae are deep-sea forms with an oval test, long spines, and weakly developed fascioles and petals. The Spatangoida are more specialized, heart-shaped forms which live in mud or sand. The order reached its maximum in the mid-Tertiary, but is still richly represented today in most seas. *See* ECHINODERMATA; ECHINOIDEA; EUECHINOIDEA. Howard B. Fell

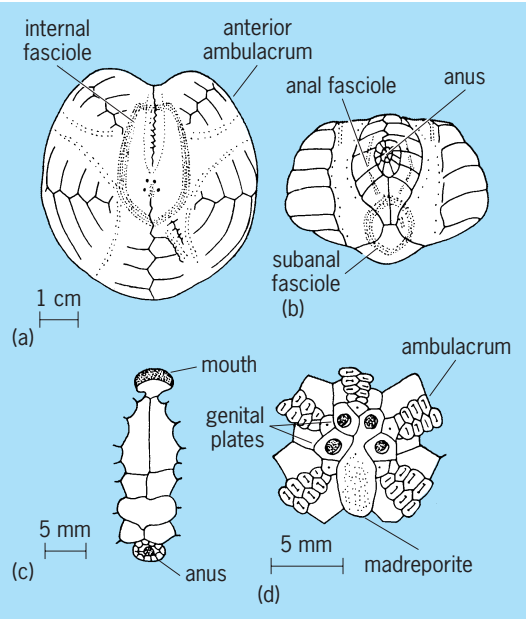

***Echinocardium cordatum*, showing diagnostic characters of spatangoids. (*a*) Aboral view. (*b*) Posterior aspect. (*c*) Amphisternous plastron. (*d*) Apical system.**

Spearmint

Either of two vegetatively propagated, clonal cultivar species of mints of the family Lamiaceae (Labiatae). They are grown primarily in Idaho, Indiana, Michigan, Washington, and Wisconsin as a source of essential oil of spearmint.

Cultivation. In agricultural production the dormant stolons, or leafless underground stems (often misnamed roots), of these perennial species are planted in a 6-in.-deep (15-cm) trench in rows 38 in. (97 cm) apart from February to April and chemically weeded with a terbacil herbicide; the leaf crop is mowed and windrowed in July or August at the time of flowering and plant maturity, partly dried, and then steam-distilled. Fields are shallow-plowed in early November, and the second-year, third-year, and subsequent crops are a solid stand, or meadow mint, in all areas that do not use rill irrigation.

Cultivars. The original cultivar "Native spearmint" with flowers in terminal inflorescences (see **illus.**), introduced from Europe into New England in pioneer days, is a sterile allotriploid hybrid having 36 somatic chromosomes, an example of a botanical mule. Native cultivar is usually called *Mentha spicata* ($2n = 48$) but is actually a hybrid between this species and *M. longifolia* ($2n = 24$). Native spearmint comprises about one-third of the production in the United States. Scotch spearmint cultivar (illus. *b*) with flowers in the axils of the leaves, *M. cardiaca* ($2n = 72$, some strains have 60 chromosomes), was introduced in 1910, apparently from Scotland. *Mentha cardiaca* has hybrid vigor and usually yields 50–65 lb (23–30 kg) of oil per acre in the Midwest as opposed to 30–40 lb (14–18 kg) per acre for Native, with yields up to 80–125 lb (36–57 kg) in the Far West. Both varieties have similar but not identical oil composition, with 60–69% carvone, 7–20% limonene, and 1–2% cineole as main compounds.

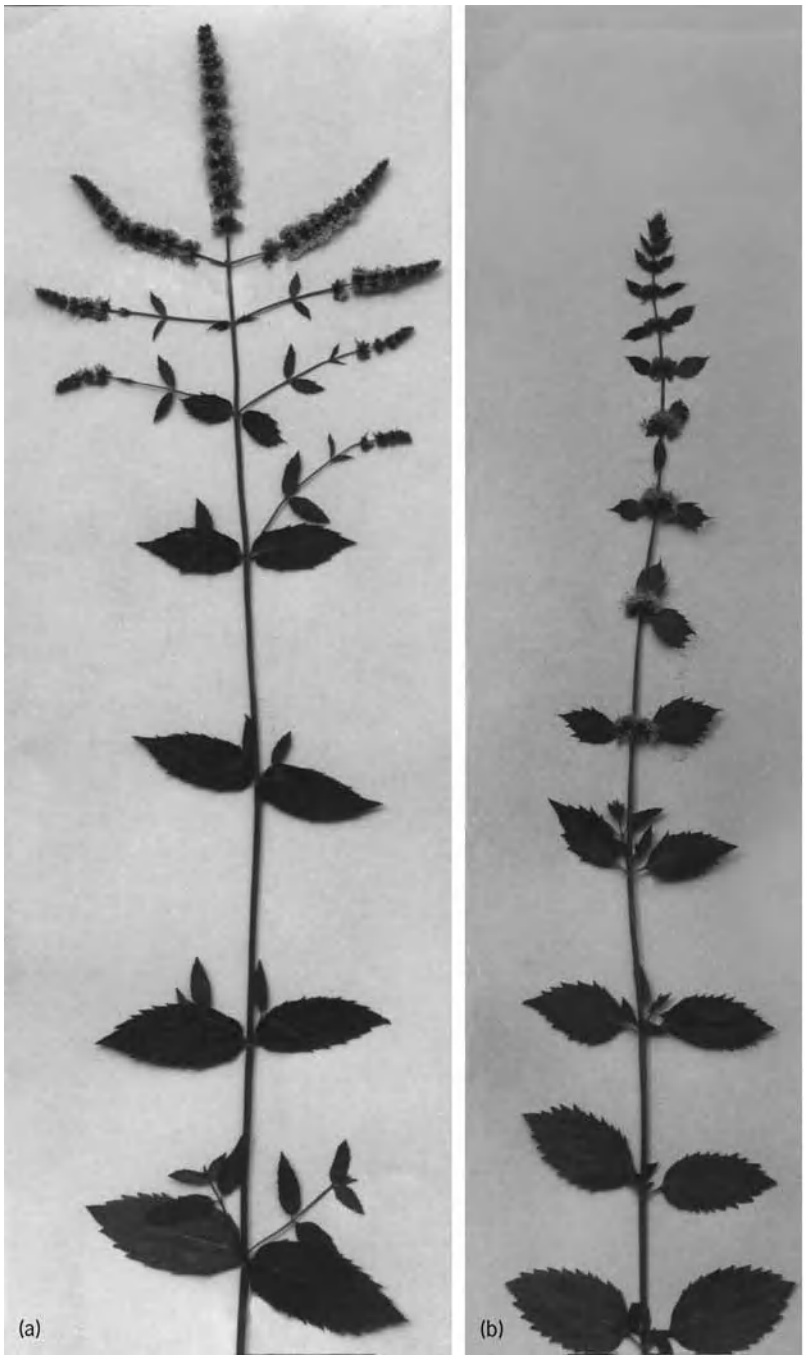

Spearmint. (*a*) Native spearmint (*Mentha spicata*). (*b*) Scotch spearmint (*M. cardiaca*). (*Courtesy of A. M. Todd Co.*)

Research has shown that Scotch is a very specific, very select, sterile F_1 hybrid between *M. arvensis* ($2n = 96$ or 72) and *M. spicata*. Both cultivars are subject to Puccinia rust and powdery mildew, and Scotch is also susceptible to Verticillium wilt. Mutation breeding is in progress to obtain resistant strains for these major diseases.

Uses. Principal uses of the oil are in flavoring gum, toothpaste, and candy. Chopped fresh leaves of *M. spicata* preserved in vinegar are used as a condiment served with lamb, especially in England, and dried or freeze-dried leaves of several strains are used in flavoring soups, stews, tea, or sauces. Sprigs of the decorative curly mint *M. crispa* (or *M. spicata* var. *crispata*) are often used in mixed drinks such as mint juleps.

Other mint species, as *M. niliaca*, *M. rotundifolia*, and *M. cordifolia*, may have carvone and a spearmint odor, but the presence of too much dihydrocarvone or of the alcohols and esters of carvone and dihydrocarvone (as carveol and carvyl acetate) or other hydrocarbons such as myrcene makes the commercial utilization of their oils unlikely. Although the cultivar species have been introduced throughout the world, in the United States they grow best between 40 to 48° north latitude, as in northern Europe, or farther north to 55° in the milder climate of Great Britain. *See* LAMIALES; SPICE AND FLAVORING.

Merritt J. Murray

Special functions

Functions which occur often enough to acquire a name. Some of these, such as the exponential, logarithmic, and the various trigonometric functions, are extensively taught in school and occur so frequently that routines for calculating them are built into many pocket calculators. *See* DIFFERENTIATION; LOGARITHM; TRIGONOMETRY.

The more complicated special functions, or higher transcendental functions as they are often called, have been extensively studied by many mathematicians because they arose in the problems which were being studied. Among the more useful functions are the following: the gamma function defined by Eq. (1), which generalizes the factorial; the related beta function defined by Eq. (2), which generalizes

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt \tag{1}$$

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt \tag{2}$$

the binomial coefficient; and elliptic integrals, which arose when mathematicians tried to determine the arc length of an ellipse, and their inverses, the elliptic functions. The hypergeometric function and its generalizations include many of the special functions which occur in mathematical physics, such as Bessel functions, Legendre functions, error functions, and the classical orthogonal polynomials of Jacobi, Laguerre, and Hermite. The zeta function defined by Eq. (3) has many applications in number theory,

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} \tag{3}$$

and it also arises in M. Planck's work on radiation. *See* BESSEL FUNCTIONS; ELLIPTIC FUNCTION AND INTEGRAL; GAMMA FUNCTION; HEAT RADIATION; HYPERGEOMETRIC FUNCTIONS; LEGENDRE FUNCTIONS; NUMBER THEORY; ORTHOGONAL POLYNOMIALS.

Bernoulli polynomials and numbers. Bernoulli polynomials, defined by the generating function in

Eq. (4), and Bernoulli numbers, $B_n = B_n(0)$, arise

$$\frac{te^{xt}}{e^t - 1} = \sum_{n=0}^{\infty} B_n(x)\frac{t^n}{n!} \qquad |t| < 2\pi \qquad (4)$$

in many applications. (Different notations are used, especially for Bernoulli numbers.) They occur in the Euler-Maclaurin summation formula (5), where a, m,

$$\sum_{j=a}^{n} f(j) = \int_a^n f(x)\,dx + {}^1/_2 f(a) + {}^1/_2 f(n)$$
$$+ \sum_{k=1}^{m} \frac{B_{2k}}{(2k)!}\left[f^{(2k-1)}(n) - f^{(2k-1)}(a)\right] + R_m(n) \qquad (5)$$

and n are arbitrary integers, $a < n$, $m > 0$, and notation (6) applies. The symbol $[x]$ stands for the largest

$$R_m(n) = -\int_a^n \frac{B_{2m}(x - [x])}{(2n)!} f^{(2m)}(x)\,dx \qquad (6)$$

integer less than or equal to x. This formula is very useful in finding numerical values of some slowly convergent infinite series or partial sums of divergent series. An important sum which can be evaluated using Bernoulli numbers is given by Eq. (7).

$$\sum_{k=1}^{\infty} \frac{1}{k^{2n}} = \frac{(2\pi)^n B_{2n}}{2(2n)!} \qquad (7)$$

Bernoulli polynomials and numbers also play an important role in number theory, combinatorial analysis, and the study of spline functions. Splines are piecewise polynomials which are spliced together in a smooth manner, up to a certain degree of differentiability. They are very useful in fitting curves to numerical data and are the solutions to a number of important extremal problems. *See* COMBINATORIAL THEORY.

More complicated functions. All of the special functions mentioned above can be given by explicit representations, either as a series or an integral. There are other functions, among them Lamé functions, spheroidal and ellipsoidal wave functions, and Mathieu functions, which arise as solutions to important differential equations, and for which explicit formulas are lacking. At least the formulas are not nearly as explicit as for the above functions. The ordinary differential equations satisfied by these functions arise from Laplace's equation or the wave equation when it is solved by separation of variables in certain systems of curvilinear coordinates. Instead of having explicit integral representations, these functions satisfy special integral equations. *See* DIFFERENTIAL EQUATION; INTEGRAL EQUATION; LAPLACE'S DIFFERENTIAL EQUATION; WAVE EQUATION.

For example, the prolate spheroidal wave functions of order zero, $\psi_n(x)$, $n = 0, 1, \ldots,$ are the bounded continuous solutions of differential equation (8) and integral equation (9), where χ_n and

$$(1 - x^2)\frac{d^2\psi_n}{dx^2} - 2x\frac{d\psi_n}{dx} + (\chi_n - c^2x^2)\psi_n = 0 \qquad (8)$$

$$\lambda_n\psi_n(x) = \int_{-1}^{1} \frac{\sin c\,(x-y)}{\pi\,(x-y)}\psi_n(y)\,dy \qquad (9)$$

λ_n are eigenvalues. Differential equation (8) is the usual starting place for the study of these functions. It arises when solving the wave equation in prolate spheroidal coordinates, and the integral equation has classically been derived as a property of these functions. Integral equation (9) is of interest because of its applications to stochastic processes, the study of lasers, the uncertainty principle, antenna theory, and the statistical theory of energy levels of complex systems. *See* STOCHASTIC PROCESS; UNCERTAINTY PRINCIPLE.

Historical development. Typically, special functions are discovered in the course of working on one problem. Many properties are then discovered, not only to aid in the solution of the problem which gave rise to them, but as a mathematical exercise, and then some of the new properties are used to solve completely different problems.

Each generation of mathematicians has a way of looking at mathematics which is different from previous generations. In the nineteenth century, complex analysis was being developed, and many important properties of special functions were discovered in the course of this development. In the second half of the twentieth century, Lie group and Lie algebra methods have been applied to obtain other important properties of special functions. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES; LIE GROUP.

Richard Askey

Bibliography. M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, 10th ed., 1972; L. Andrews, *Special Functions of Mathematics for Engineers*, 2d ed., 1998; R. A. Askey, T. M. Koornwinder, and W. Schempp (eds.), *Special Functions: Group Theoretical Aspects and Applications*, 1984; A. Erdélyi et al., *Higher Transcendental Functions*, 3 vols., 1953–1955, reprint 1981; G. Szegö, *Orthogonal Polynomials*, 4th ed., 1975; Z. X. Wang and D. R. Guo, *Special Functions*, 1989.

Speciation

The process by which new species of organisms evolve from preexisting species. It is part of the whole process of organic evolution. The modern period of its study began with the publication of Charles Darwin's and Alfred Russell Wallace's theory of evolution by natural selection in 1858, and Darwin's *On the Origin of Species* in 1859.

There was no problem of speciation during the period when it was believed that species were divinely created and immutable. Belief in the fixity of species was almost universal before the middle of the nineteenth century. Then it was gradually realized that all species continuously change, or evolve; however, the causative mechanism remained to be discovered.

Darwin proposed a mechanism. He argued that (1) within any species population there is always some heritable variation; the individuals differ among

themselves in structure, physiology, and behavior; and (2) natural selection acts upon this variation by eliminating the less fit. Thus if two members of an animal population differ from each other in their ability to find a mate, obtain food, escape from predators, resist the ravages of parasites and pathogens, or survive the rigors of the climate, the more successful will be more likely than the less successful to leave descendants. The more successful is said to have greater fitness, to be better adapted, or to be selectively favored. Likewise among plants: one plant individual is fitter than another if its heritable characteristics make it more successful than the other in obtaining light, water, and nutrients, in protecting itself from herbivores and disease organisms, or in surviving adverse climatic conditions. Over the course of time, as the fitter members of a population leave more descendants than the less fit, their characteristics become more common.

This is the process of natural selection, which tends to preserve the well adapted at the expense of the ill adapted in a variable population. The genetic variability that must exist if natural selection is to act is generated by genetic mutations in the broad sense, including chromosomal rearrangements together with point mutations. *See* GENETICS; MUTATION.

If two separate populations of a species live in separate regions, exposed to different environments, natural selection will cause each population to accumulate characters adapting it to its own environment. The two populations will thus diverge from each other and, given time, will become so different that they are no longer interfertile. At this point, speciation has occurred: two species have come into existence in the place of one. This mode of speciation, speciation by splitting, is probably the most common mode. Two other modes are hybrid speciation and phyletic speciation; many biologists do not regard the latter as true speciation.

Splitting. There are four ways in which a species population can become separated into two (or more) parts that can undergo genetic divergence and evolve into separate species (**Fig. 1**). G. L. Bush has named them types Ia, Ib, II, and III.

In type Ia an extensive population becomes split into comparatively large parts (Fig. 1*a*). In terrestrial species this may happen because movement of the Earth's crustal plates rafts the parts away from each other; or else a barrier such as a mountain range or a tract of desert may form and split a once-continuous population. A marine population is split if a land bridge forms which cuts through it; for example, the fall in sea level that accompanied the Pleistocene glaciations created a land bridge between Alaska and Siberia.

A type Ib split occurs if a small, outlying segment of a population becomes detached from the rest (Fig. 1*b*). Such a peripheral isolate is usually a colony founded by a few emigrants from the main body of the species population. The colony can evolve into a new species even if it is not totally isolated from the parent population, provided its necessarily small gene pool chances to be well adapted to the habitat in which it finds itself. The "preadapted" possessors of these genes will be selectively favored within their small local habitat, and will quickly proliferate.

Population splits of both types Ia and Ib cause the diverging populations to be geographically separated, or allopatric. The resultant genetic separation is known as allopatric speciation.

In type II a population becomes split into geographically contiguous parts between which there is a limited amount of gene exchange. The environments of the daughter populations are assumed to differ sufficiently for natural selection to override the effects of the restricted gene flow, and speciation can therefore occur. This is parapatric speciation (Fig. 1*c*).

A type III split occurs if, within a single geographic region, parts of a population in some way become so isolated from each other that gene exchange between them falls below the level needed to maintain the integrity of the original population. The result is known as sympatric speciation (Fig. 1*d*). It is probably common among parasites and parasitoids, but is thought to be rare or nonexistent in free-living organisms. Even though closely related species adapted to different habitats, or with different modes of life, or with different resource requirements, often co-occur in the same geographic area, their co-occurrence should not be ascribed to sympatric speciation unless a mechanism can be found that could have initiated their divergence in the first place. A more likely explanation for such co-occurrences is that the species evolved allopatrically or parapatrically, that ecological divergence accompanied speciation, and that the new species then became secondarily sympatric by migrating into each other's areas after the geographic barrier that formerly separated them disappeared. *See* ISLAND BIOGEOGRAPHY.

Descriptions of allopatric, parapatric, and sympatric speciation often tacitly assume that in the

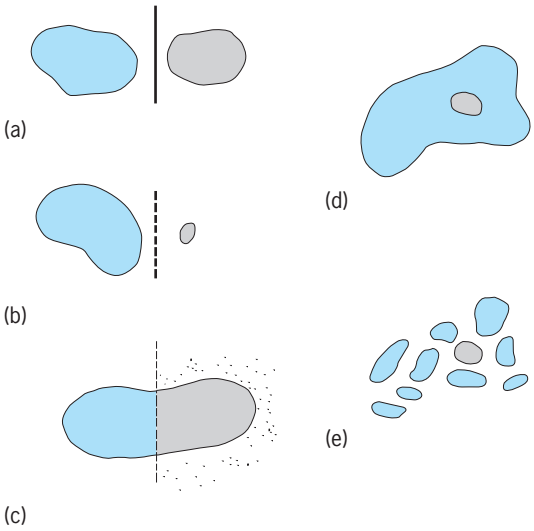

Fig. 1. Different modes of speciation; *a–e* are described in the text. The two tints represent the geographic ranges of diverging and speciation species populations.

absence of clear, unmistakable barriers, gene exchange occurs freely throughout very extensive populations; often an entire species is treated as a single breeding population. However, P. R. Ehrlich and P. H. Raven have argued that usually the total population of a whole species is broken into a large number of small breeding populations (demes) that are fairly isolated from each other and among which gene exchange is comparatively rare. Thus a deme rather than a species population is the unit of evolution. As a result, sympatric speciation will sometimes seem to have occurred if a deme near the center of a species' range chances to diverge from the surrounding demes; the process may be called quasisympatric speciation (Fig. 1*e*). The fact that all demes do not diverge from one another, giving as many species as there are demes, presumably results from stabilizing selection. Stabilizing selection causes groups of demes to persist in their resemblance to each other, and such groups are the "populations" constituting the visible units of evolution. *See* POPULATION GENETICS.

Isolating mechanisms. Two genetically diverging populations are not separate species so long as it remains possible for them to lose their distinctness, by interbreeding, if they chance to become sympatric. For speciation to be complete, the populations must evolve isolating mechanisms that prevent gene exchange between them; this allows them to maintain their genetic integrity.

There are many kinds of isolating mechanisms. The diverging populations may become genetically incompatible, so that hybrids either are never produced or are sterile. Even when the diverging populations can be shown to be infertile, they may be effectively isolated by a variety of mechanisms; they may breed at different times or in different habitats; they may have evolved different patterns of mating behavior, different mating calls, or different sex attractants (pheromones), so that individuals of the two populations never attempt to copulate; copulation may be mechanically impossible because of the structure of the genitalia (this occurs in some insects); and in insect-pollinated plants, diverging populations sometimes become adapted to different species of insect pollinators.

Hybrid speciation. In plants, new species are sometimes formed abruptly by the following mechanism. The accidental crossing of two diploid plants of different, though related, species yields a healthy but sterile diploid hybrid. The hybrid is sterile because the chromosomes it has received from its parents cannot form normal bivalent pairs; meiosis is upset, and functional gametes cannot be produced. But now chromosome doubling occurs in the sterile hybrid, making it tetraploid; normal meiosis becomes possible; fertile gametes are produced, and a new species has appeared.

In symbols, let the genomic constitution of the parents be *AA* and *BB* respectively. That of their sterile hybrid is *AB*, in which none of the chromosomes has a homologous partner to pair with at meiosis. After chromosome doubling, however, the hybrid, now tetraploid, has the constitution *AABB*; meiosis can proceed normally, and the gametes are fertile. The new species is called an amphidiploid (or allotetraploid).

A whole range of similar, but more elaborate, variants of this process occur as well, yielding hybrids with different numbers of chromosome sets. The general name for the process, with numbers unspecified, is amphiploidy (or allopolyploidy or alloploidy).

The process happens because chromosome doubling is a common occurrence in plants; doubling results either from somatic doubling at mitosis (the paired chromosomes fail to separate) or from nonreduction at meiosis. Amphiploids may be produced repeatedly, with new lineages being added in a succession of outcrossings, backcrossings, and chromosome doublings. The result is a polyploid complex, an array of interrelated diploid and polyploid species. *See* MEIOSIS.

Speciation in uniparental organisms. Many plants reproduce uniparentally for long periods of time. Uniparental reproduction occurs in two ways: by autogamy (self-fertilization) and by apomixis (asexual reproduction); the latter term includes both vegetative multiplication and agamospermy (the asexual formation of viable seeds). A uniparental plant is not a member of a "biological species" as the term is usually defined, namely, a reproductively isolated collection of interbreeding individuals. This has led to the definition, by G. G. Simpson, of the "evolutionary species," a more inclusive category, defined so as to apply to both uniparental and biparental organisms. An evolutionary species is a collection of genotypically similar individuals of common ancestry whose similarity is maintained by stabilizing selection in a given environment.

In describing speciation in uniparental organisms, one must consider how a new uniparental lineage arises in the first place, and how its members can subsequently speciate. A uniparental lineage arises when two sexual species hybridize and the hybrids, though sexually sterile, reproduce uniparentally. If many crossings occur between the two sexual species, the result is a collection of hybrid individuals that are similar, but not identical, genetically. The collection constitutes an evolutionary species. Suppose now that each individual becomes the ancestor of a succession of uniparentally produced descendants. If the progenitors of these lineages begin growth in a number of slightly different habitats, natural selection will lead to differentiation among them: the survivors in each habitat will be adapted to that habitat. Thus the plants in any one habitat are likely to differ slightly from those in other habitats. The differences are genetic, and they persist. Each distinguishable group of plants is called a microspecies. An array of microspecies, descended from the same ancestral pair of sexual species, is an agamic complex.

Microspecies that reproduce asexually (apomictically) are usually highly heterozygous; but autogamous microspecies lose the heterozygosity of their hybrid origin after several generations of

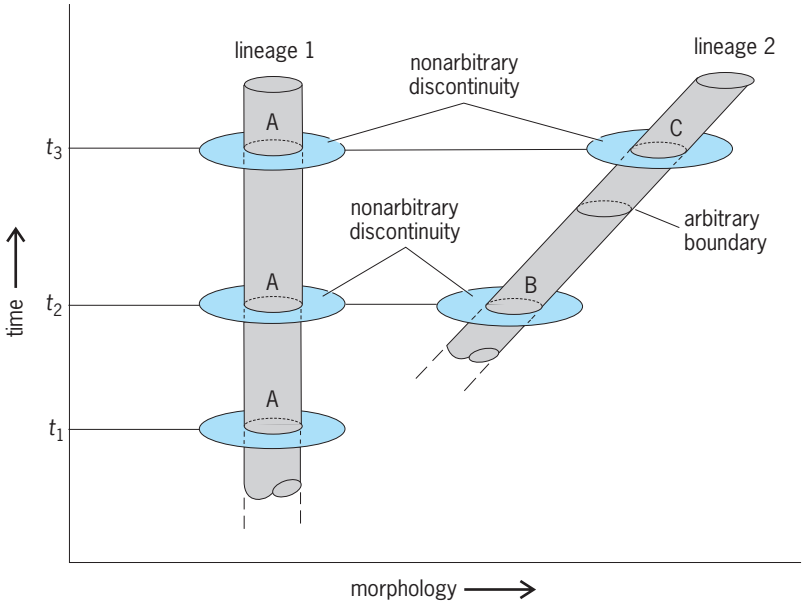

Fig. 2. Evolution of two lineages. Lineage 1 (species A) persists unchanged. Lineage 2 changes with time and has been divided, at an arbitrary boundary, into two chronospecies, B and C. (*After P. D. Gingerich, Univ. Mich. Pap. Paleontol., 15:1–140, 1976*)

self-fertilization and become complete, or nearly complete, homozygotes.

Phyletic evolution and chronospecies. Consider a series of fossils spanning a long interval of time (long in the geological sense) and believed to represent the successive forms of a single lineage. If there has been appreciable evolution within the lineage, the change is described as phyletic evolution or anagenesis.

Fossils from a single lineage but of different ages may differ from each other so conspicuously that even though the younger is directly descended from the older, paleontologists find it convenient to treat them as taxonomically different species. Such species are called chronospecies (or successional species, paleospecies, or evolutionary species). They are not species in the usual sense; questions as to whether or not they can interbreed do not arise, since they are never alive at the same time. In a lineage evolving gradually, decisions on where to put the dividing line between each chronospecies and its successor are arbitrary. A "vanished" chronospecies, one that has disappeared and been replaced by a successor chronospecies of the same lineage, is said to have undergone pseudoextinction.

Figure 2 contrasts the behavior of a lineage evolving phyletically (lineage 2) with one persisting unchanged (lineage 1). The "species" labeled B and C in lineage 2 are chronospecies. At an arbitrary point on the time scale, B becomes pseudoextinct and is succeeded by C.

The transition from a chronospecies to its successor is not usually regarded as speciation in the strict sense. Speciation as usually defined entails the splitting of lineages.

Quantum speciation versus phyletic gradualism. When a lineage splits, the two descendant lineages may diverge slowly or rapidly. There has been much debate on whether one or other of these modes has dominated the course of evolution and, if so, which. The two modes to be contrasted are shown diagrammatically in **Fig. 3**.

The tree on the right in the figure represents evolution and speciation according to the model known as phyletic gradualism. It supposes that all lineages undergo continuous gradual change (that is, phyletic evolution as described above) at a fairly constant rate. When splits occur, the descendant lineages diverge gradually from each other. There is no appreciable change in evolutionary rate at the time a split takes place. While the separated lineages are still capable of interbreeding to a limited extent, though with reduced fertility, they are known as semispecies. Their morphological divergence may sometimes be so slow that they remain practically indistinguishable even after they have lost the capacity to interbreed and have therefore reached specific rank; if so, they are known as cryptic species.

In Fig. 3*a* the tree represents the punctuated equilibrium model of evolution and speciation, so called by N. Eldredge and S. J. Gould. It supposes that the splitting of a parent species population is always (or nearly always) accompanied by quantum speciation, the extremely rapid (in the geological sense) and pronounced divergence of one of the descendant lineages from the ancestral form. It is assumed that the divergent species originates most often by the chance splitting off of a small peripheral isolate from an extensive parental population; this is allopatric speciation of type Ib (Fig. 1). If the isolate is small, its small gene pool may not be at all representative of the parent gene pool from which it is drawn. It may, simply by chance, deviate markedly. Further, because of its peripheral location, the environment of the isolate is likely to differ from that of the main body of the species population, and natural selection will then tend to magnify the deviation.

Evolutionary changes associated with the appearance of new species are thus, according to this model, concentrated in speciation events of very

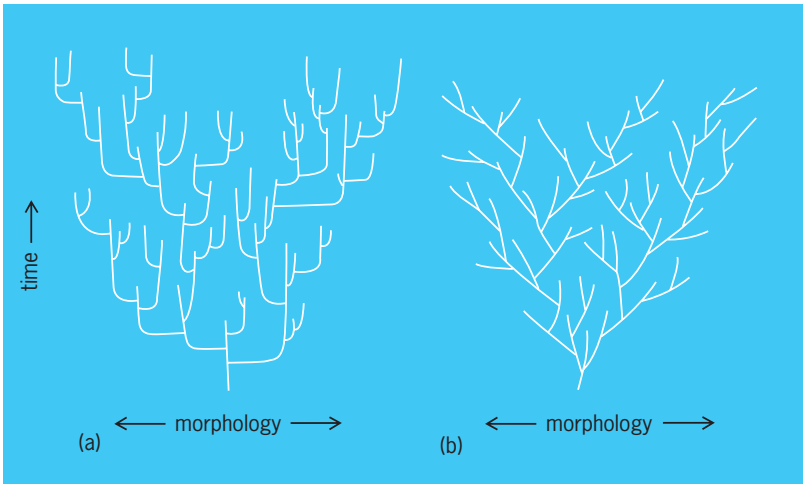

Fig. 3. Two contrasted modes of species evolution: (*a*) punctuated equilibrium model and (*b*) phyletic gradualism model. (*After S. M. Stanley, Macroevolution, Pattern and Process, W. H. Freeman, 1979*)

short duration; they are represented by the horizontal segments of the tree's branches (Fig. 3*a*). During the long time intervals between successive speciation events, a species remains almost unchanged, as shown by the vertical segments of the branches. While a species is in this static condition, phyletic evolution may bring about minor adjustments; this within-species evolution is known as microevolution. But the magnitude of these phyletic changes is negligible compared with that accompanying macroevolution, the formation of new species. Ernst Mayr introduced the terms maintenance evolution and switch evolution to describe these two evolutionary modes.

Many students of evolution are of the opinion that most groups of organisms evolve in accordance with the punctuated equilibrium model rather than by phyletic gradualism. There are two chief arguments for this view. First, it is clear from the fossil record that many species persist without perceptible change over long stretches of time and then suddenly make large quantum jumps to radically new forms. Second, phyletic gradualism seems to be too slow a process to account for the tremendous proliferation of species needed to supply the vast array of living forms that have come into existence since life first appeared on Earth. *See* ANIMAL EVOLUTION; SPECIES CONCEPT.

E. C. Pielou

Bibliography. G. L. Bush, Modes of animal speciation, *Annu. Rev. Ecol. Systemat.*, 6:339–364, 1975; T. Dobzhansky, *Genetics of the Evolutionary Process*, 1972; D. H. Erwin and R. L. Anstey, *New Approaches to Speciation in the Fossil Record*, 1995; V. Grant, *Plant Speciation*, 2d ed., 1981; M. L. McKinney, *Evolution of Life: Processes, Patterns, and Prospects*, 1993.

Species concept

The idea that the diversity of nature is divisible into a finite number of definable species. In general, species concepts grow out of attempts to understand the very nature of biological organization above the level of the individual organism. There are two basic questions: (1) What does it mean to be a species in general? Do all species have certain characteristics, such as forming genealogical lineages, just as all atoms have certain characteristics, such as the ability to undergo chemical reactions? (2) What does it mean to be a particular species? The first question addresses species concepts. The second question addresses how to apply a species concept to living organisms of the world. Does the name *Homo sapiens* apply to a group of organisms existing in nature? If so, does it belong, as a member, to a natural kind that can be characterized by some set of properties? The difference between the questions of what it is to be a species in general versus what does it mean to be a particular species represents the dividing line between what it is to be a natural kind and what it is to be a natural individual. To understand this distinction, we must first take up the more general question of the nature of kinds. We can then return to the question of species concepts. *See* SPECIATION.

Kinds and particulars. Kinds are attempts to divide the world of particulars according to a definition that specifies properties that separate individuals into two classes: those particulars that have the properties are members of the kind, and those particulars that do not have the properties belong to another kind. What do we mean when we say "particulars"? Particulars are individual things in nature. We can usually point to particulars, such as pointing to a cat or pointing to a truck. A particular belongs to a kind if the particular has the properties that define the kind. For example, we might define the kind "black cat" as those cats that have the property of being covered with black fur. My black cat belongs to this kind, but my gray cat does not since it lacks the property of the kind. In fact, black cats in general (every black cat that has ever lived or will live in the future) belong to the kind "black cat."

As it turns out, there are at least two types of kinds: nominal kinds and natural kinds. The difference between nominal and natural kinds lies in how the kind functions. Natural kinds function in scientific theories of how nature operates, because the theory asserts that certain particulars (individuals) have certain properties. For example, the kind "atom" functions in theories of physics and chemistry as the smallest kind of matter whose members (particular individual atoms) can take part in chemical reactions (the property). In turn, hydrogen is a kind of atom whose members have one proton (property), while helium is a kind of atom whose members have two protons (another property). Kinds such as hydrogen and helium are natural kinds because members of each kind (individual atoms of hydrogen and helium) are predicted to exist under the right circumstances and have their properties as an assertion of our theories of chemistry and physics. In contrast to natural kinds, most of our language is filled with nominal kinds. Nominal kinds are not predicted by scientific theories, but function to allow us to communicate ideas. For example, no scientific theory predicts the kinds "truck" and "bicycle." Yet, nominal kinds have properties (trucks have four wheels, bicycles have two) that allow us to distinguish trucks from bicycles. Thus, nominal kinds may have useful functions in communication, but they are not directly derived from scientific theories. *See* SCIENTIFIC METHODS.

Concepts and theories. Since Aristotle, biologists have attempted to develop a concept of "species" that will function in biology as a natural kind in a manner similar to natural kinds in other disciplines such as "quark" in physics and "atom" in chemistry. This search was inhibited because philosophers have traditionally thought of particular species (*Homo sapiens*, *Pinus ponderosa*, etc.) as natural kinds, or sets; or they thought that particular species were nominal—arbitrary but useful inventions of taxonomists to organize the real entities, that is, individual organisms. However, work by Michael

Ghiselin, David Hull, and others asserted that particular species were, in fact, particulars that function in evolutionary theory in the same manner as individual atoms function in chemical theory—as interactors or participants in process. If this is so, then each of these individual species would be members of a natural kind, "species," that had properties predicted from evolutionary theory. Attempts to capture the properties of the kind "species" result in a species concept.

The ideas that organisms could be grouped into more or less discrete units, and the idea that some of these units were more similar to each other than to other units, long predated ideas about evolution. That there was a pattern to life's diversity gave rise to the first species concept: members of one species are different in their physical characteristics from members of other species, and their offspring look like them (like begets like). This concept was the basis for both folk taxonomies (names of different groups of individual organisms in common language) and scientific taxonomies (formal names meant to have scientific import). As the practice of scientific taxonomy developed, species were gathered into larger groups based on similarities and differences. These groupings were hierarchical and seemed to reflect patterns of similarities and differences seen in nature. Current evolutionary theories seek to explain why life's diversity seems to be arranged in the hierarchical manner that we see in nature. *See* TAXONOMIC CATEGORIES; TAXONOMY; ZOOLOGICAL NOMENCLATURE.

Like all scientific theories, theories of evolution predict natural kinds with attendant properties that are found in individuals. Theories of evolution can be generally divided into two basic themes. Microevolutionary theory is concerned with evolutionary change within and between populations of individual organisms. Processes include natural selection, gene flow, and stochastic processes such as genetic drift. These theories do not address the hierarchical patterns that predate Darwin and have fascinated biologists since Linnaeus. Rather, they predict that populations of individual organisms will change through mutation, selection, and drift in predictable ways. Macroevolutionary theory is concerned with the origin and fate of species. It directly addresses the hierarchical nature of diversity. Process theories that purport to explain the hierarchy include speciation and extinction. These two themes, micro- and macroevolution, intersect at the level of species. From the macroevolutionary perspective, species are those particulars that originate from various processes of speciation and disappear through extinction. Macroevolutionary processes follow the "rules" of microevolutionary processes, but add an extra level of processes such as response to a physical barrier that divides a once continuous population into two populations. Thus, the natural kind "species" has members (particular species) that undergo such processes. *See* MACROEVOLUTION; ORGANIC EVOLUTION.

The problem with species concepts derives from the problem with changing ideas of the world in general. Our concepts of natural kinds, and indeed whether we think that a certain kind is real (has members in the world) or unreal (properties can be listed, but nothing exists that has the properties, like unicorns or fairies), depend on our current theories. The recent controversy concerning the definition of the kind "planet" is an example. Pluto has been rejected (at least for the moment) as being a planet based on current theory, and a new kind, dwarf planet, has been proposed that includes not only Pluto but other smaller bodies in the solar system, such as the asteroid Ceres. The hope is that the closer the theory is to nature, the more predictive will be the kinds and their properties. So it is with species concepts.

Types of species concepts. Species concepts, in one form or another, have been around for several hundred years. Each concept (and there are over 20 current ones) is an attempt to capture the properties of the largest (or smallest) kind and thus permit a search for particulars (*Homo sapiens*, etc.) that are members of the kind. Of the plethora of species concepts, evolutionary biologists (including systematists and taxonomists) seem to be converging on the concept of species-as-lineages, the Evolutionary Species Concept—species are those things that exist between speciation events and that originate through speciation events (Henning, 1996; Wiley and Mayden, 2000). Other concepts capture part of this more general concept. For example, the venerable Morphological Species Concept claims that different species have different or slightly different physical characteristics. This is true for many species (humans versus chimpanzees), but not true for others where speciation has resulted in behavioral or genetic changes with little or no detectable change in physical characteristics (as in some closely related fruit flies). The popular Biological Species Concept asserts that species have the property of reproductive isolation. This seems to be true for a great number of species, but not for all species and especially not for many species of recent origin. This has led workers such as Richard Mayden to assert that all currently valid species concepts (kinds for which examples can be found in nature) are different reflections of the more general natural kind, that is, the Evolutionary Species Concept. Of course, even if all evolutionary biologists agreed with Mayden, this does not end controversies as to whether a particular species is real. For example, is *Pinus ponderosa* (a pine tree) real? We might vigorously disagree as to whether or not a speciation event has occurred, based on the evidence presented. We might argue that *Pinus ponderosa* is not a species but is actually two or more species, or simply a population of another species. However, if we can agree, our agreement will be founded both on the empirical data and on a species concept that applies in an unambiguous manner, leading to a consilience (a synthesis of knowledge) between practice and theory.

E. O. Wiley

Bibliography. M. T. Ghiselin, A radical solution to the species problem, *Systematic Zool.*, 23:536–544,

1974; W. Hennig, *Phylogenetic Systematics*, University of Illinois Press, Urbana, 1966; D. L. Hull, Are species really individuals?, *Systematic Zool.*, 25:174–191, 1976; R. L. Mayden, A hierarchy of species concepts: The denouement in the saga of the species problem, pp. 381–424 in *Species, the Units of Biodiversity*, ed. by M. F. Claridge, H. A. Dawah, and M. R. Wilson, Chapman and Hall, London, 1997; G. G. Simpson, *Tempo and Mode in Evolution*, Columbia University Press, New York, 1944; E. O. Wiley and R. L. Mayden, The evolutionary species concept, pp. 70–89 in *Species Concepts and Phylogenetic Systematics*, ed. by Q. D. Wheeler and R. Meier, Columbia University Press, New York, 2000.

Specific charge

The ratio of charge to mass, expressed as e/m, of a particle. The acceleration of a particle in electromagnetic fields is proportional to its specific charge. Specific charge can be determined by measuring the velocity υ which the particle acquires in falling through an electric potential V ($\upsilon = \sqrt{2eV/m}$); by measuring the frequency of revolution in a magnetic field H (the so-called cyclotron frequency $\omega = eH/mc$, where c is the velocity of light); or by observing the orbit of the particles in combined electric and magnetic fields. In the mass spectrograph, the fields are arranged so that particles of differing velocities but of the same e/m are focused at a point. *See* ELECTRON MOTION IN VACUUM; ELEMENTARY PARTICLE; MASS SPECTROSCOPE.

Charles J. Goebel

Specific fuel consumption

The ratio of the fuel mass flow of an aircraft engine to its output power, in specified units. Specific fuel consumption (abbreviated sfc or SFC) is a widely used measure of atmospheric engine performance. For reciprocating engines it is usually given in U.S. Customary units of pound-mass per hour per horsepower [(lbm/h)/hp or lbm/(hp·h)], and International System (SI) units of kilograms per hour per kilowatt [(kg/h)/kW]. *See* RECIPROCATING AIRCRAFT ENGINE.

For the gas turbine family of atmospheric aircraft engines, and for ramjets, performance is usually given in terms of thrust specific fuel consumption (abbreviated tsfc or TSFC) expressed as fuel mass flow per unit thrust output with Customary units of pound-mass per hour per pound-force [(lbm/h)/lbf] or SI units of kilograms per hour per newton [(kg/h)/N; 1 N equals approximately 0.225 lbf]. For high-supersonic and hypersonic ramjets, specific fuel consumption is sometimes given in pound-mass per second per pound-force [(lbm/s)/lbf] or kilograms per second per newton [(kg/s)/N]. Specific thrust (or specific impulse), which is the inverse of specific fuel consumption, is often used especially for supersonic combustion ramjet (scramjet) performance at Mach numbers above 6 (the Mach number N_{Ma} equals 1.0 at the local speed of sound). *See* AIRCRAFT PROPULSION; TURBINE ENGINE SUBSYSTEMS; MACH NUMBER; PROPULSION; SPECIFIC IMPULSE; TURBINE PROPULSION; TURBOFAN; TURBOJET.

Since the combustion process in atmospheric aircraft engines is supported by oxygen in the ambient air, only the hydrocarbon fuel (gasoline for reciprocating engines and kerosine for turbine engines and ramjets), which must be carried on board the aircraft, needs to be accounted for in determining engine performance. Methane and hydrogen may be used in the future as fuels for atmospheric engines. To obtain a low specific fuel consumption, the fuel should have both a high heat of combustion (that is, energy content per unit mass of fuel) and high cycle efficiency. The cycle efficiency is related to the operating compression ratio of the engine. *See* AIRCRAFT FUEL; GASOLINE; KEROSINE.

A typical specific fuel consumption value for reciprocating engines and turboprops (in terms of equivalent shaft horsepower) is approximately 0.5 (lbm/h)/hp [0.3 (kg/h)/kW]. Gas turbines range between 0.6 (lbm/h)/lbf [0.06 (kg/h)/N] for turbofans and 0.8 (lbm/h)/lbf [0.08 (kg/h)/N] for turbojets. Ramjet engines with hydrocarbon fuel range upward from a minimum of about 2.0 (lbm/h)/lbf [0.20 (kg/h)/N] at flight speeds between Mach numbers 2 and 4. Values of specific fuel consumption for ramjets and scramjets with hydrogen fuel beginning at Mach 4 trend from approximately 0.9 to 2.4 (lbm/h)/lbf [0.09 to 0.24 (kg/h)/N] at Mach 16 and beyond.

J. Preston Layton

Bibliography. D. G. Shepherd, *Aerospace Propulsion*, 1972.

Specific heat

A measure of the heat required to raise the temperature of a substance. When the heat ΔQ is added to a body of mass m, raising its temperature by ΔT, the ratio C given in Eq. (1) is defined as the heat capacity

$$C = \frac{\Delta Q}{\Delta T} \tag{1}$$

of the body. The quantity c defined in Eq. (2) is called

$$c = \frac{C}{m} = \frac{1}{m}\frac{\Delta Q}{\Delta T} \tag{2}$$

the specific heat capacity, following the recommendation by the International Union of Pure and Applied Physics (IUPAP). However, this quantity is also often referred to with the shorter name specific heat. A commonly used unit for heat capacity is joule · kelvin^{-1} ($J \cdot K^{-1}$); for specific heat capacity, the unit joule · gram^{-1} · K^{-1} ($J \cdot g^{-1} \cdot K^{-1}$) is often used. Joule should be preferred over the unit calorie = 4.18 J. As a unit of specific heat capacity, Btu · lb^{-1} · °F^{-1} = 4.21 J · g^{-1} · K^{-1} is also still in use in English-language engineering literature. If the heat capacity is referred to the amount of substance in the body, the molar heat capacity c_m results, with the unit

$J \cdot mol^{-1} \cdot K^{-1}$. The IUPAP recommends distinguishing this quantity from the specific heat capacity c through the subscript m, although this recommendation is usually disregarded in an attempt to conserve subscripts.

If the volume of the body is kept constant as the energy ΔQ is added, the entire energy will go into raising its temperature. If, however, the body is kept at a constant pressure, it will change its volume, usually expanding as it is heated, thus converting some of the heat ΔQ into mechanical energy. Consequently, its temperature increase will be less than if the volume is kept constant. It is therefore necessary to distinguish between these two processes, which are identified with the subscripts V (constant volume) and p (constant pressure): C_V, c_V, and C_p, c_p. For gases at low pressures, which obey the ideal gas law, the molar heat capacities differ by R, the molar gas constant, as given in Eq. (3) where

$$c_p - c_V = R \tag{3}$$

$R = 8.31\ J \cdot mol^{-1} \cdot K^{-1}$; that is, the expanding gas heats up less. For monatomic gases, for example, $c_p = {}^5/_2R$, considerably smaller than $c_V = {}^3/_2R$.

For solids, the difference between c_p and c_V is of the order of 1% of the specific heat capacities at room temperature, and varies roughly linearly with temperature. It results from the thermal expansion. This small difference can often be ignored. *See* CALORIMETRY; CHEMICAL THERMODYNAMICS; HEAT CAPACITY; SPECIFIC HEAT OF SOLIDS; THERMODYNAMIC PROCESSES. Robert O. Pohl

Bibliography. International Union of Pure and Applied Physics, Document U.I.P. 20, 1978; G. N. Lewis and M. Randall, *Thermodynamics*, 2d ed., 1961; F. W. Sears and G. L. Salinger, *Thermodynamics, Kinetic Theory, and Statistical Mechanics*, 3d ed., 1975; K. Wark and D. E. Richards, *Thermodynamics*, 6th ed., 1999; M. W. Zemansky and R. H. Dittman, *Heat and Thermodynamics*, 7th ed., 1996.

Specific heat of solids

The specific heat (short for specific heat capacity) of a solid is the amount of heat required to increase the temperature of a unit mass of the solid by a unit amount. It is determined by the vibrations of its atoms or excitations of its electrons, and also by a variety of phase transitions. *See* SPECIFIC HEAT.

Contribution of atomic vibrations. If a solid object, say a piece of rock salt, is said to have a certain temperature, it means that its constituent atoms perform motions around their equilibrium positions. The larger the amplitude of these motions, the larger the temperature.

Classical case. In the classical picture of this thermal motion, each atom is considered to be a harmonic, three-dimensional oscillator with six degrees of freedom, three of kinetic and three of potential energy. Each degree of freedom contains on average the energy $(1/2)k_BT$, where T is the absolute temperature [measured in kelvins (K)] and k_B is Boltzmann's constant ($k_B = 1.38 \times 10^{-23}\ J \cdot K^{-1}$). Thus, an object containing 1 mole or $N_A = 6.02 \times 10^{23}$ atoms, contains the thermal energy Q in the form of atomic vibrations given by Eq. (1), and has the molar heat capacity c_v given by Eq. (2), where R is the molar gas constant,

$$Q = 6(1/2)RT \tag{1}$$

$$c_V = 3R \tag{2}$$

and $R = N_Ak_B = 8.31\ J \cdot mol^{-1} \cdot K^{-1}$. The molar heat capacity can be converted to specific heat capacity by dividing the molar heat capacity by the mass of 1 mole of the substance. Equation (2) is called the rule of Dulong and Petit, proposed in 1819 on the basis of measurements near room temperature. *See* DEGREE OF FREEDOM (MECHANICS); KINETIC THEORY OF MATTER; TEMPERATURE.

Quantized atomic vibrations. When H. F. Weber (1875) extended the measurements over a wide temperature range (220 to 1300 K or −84 to 1880°F), he found that Eq. (2) holds only in the limit of high temperatures. **Figure 1** shows measurements on copper, extended to even lower temperatures: at ~20 K (−424°F) the specific heat capacity has dropped to ~3% of its room-temperature value which agrees with the value based on the rule of Dulong-Petit; the arrow on the right ordinate of Fig. 1 indicates the molar heat capacity of $3R$. The explanation for the decrease at low temperatures was offered by A. Einstein (1907). He suggested that the atomic vibrations are quantized; the atoms cannot vibrate with continuously variable amplitude (or energy). Rather, the atoms can only have energies varying in integer steps of $h\nu$, where h is Planck's constant, $h = 6.63 \times 10^{-34}\ J \cdot s$, and ν is the classical frequency of

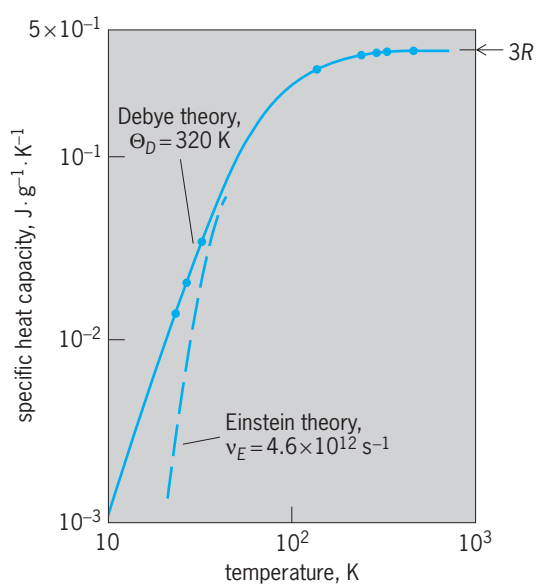

Fig. 1. Specific heat capacity for copper (data points), and predictions according to theories of Einstein and Debye (curves), on logarithmic scales. Above ~300 K (80°F), the specific heat capacity approaches the value based on the rule of Dulong-Petit, which for copper is $J \cdot g^{-1} \cdot K^{-1}$. The molar heat capacity of $3R$ for copper (molecular weight 63.546) is indicated.

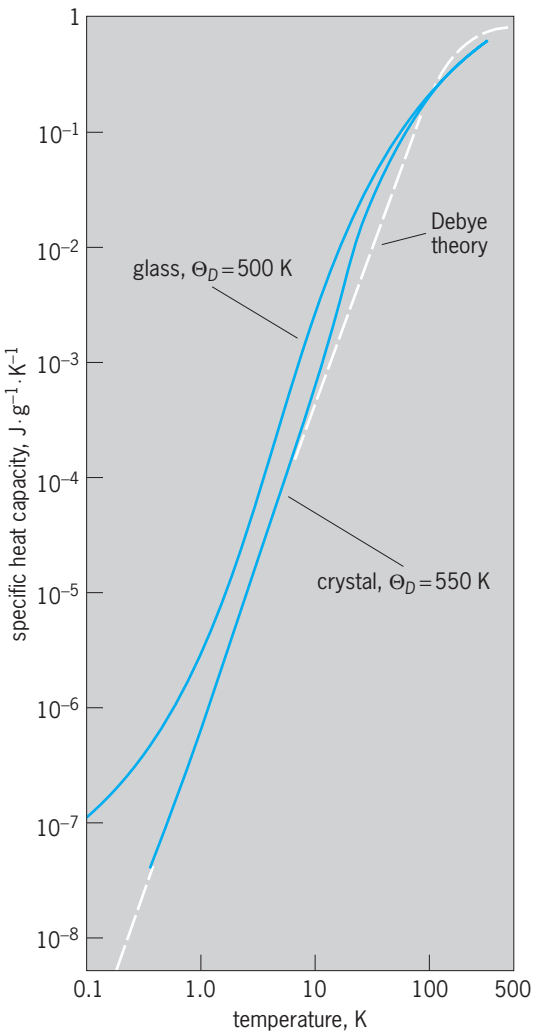

Fig. 2. Specific heat capacity of silicon dioxide (SiO_2) in its crystalline and amorphous (glass) phases. The Debye characteristic temperatures Θ_D, determined from elastic constants, are so similar that the Debye specific heat prediction is shown as a single curve for both solids.

vibration. As the temperature decreases, an increasing fraction of the atoms will not be excited at all; they will be frozen out. This leads to an exponential decrease of the specific heat capacity. Figure 1 shows Einstein's fit to the data, which he obtained by using the frequency $\upsilon_E = 4.6 \times 10^{12}\ s^{-1}$ (Einstein frequency for copper) as an adjustable parameter. Through this work, the validity of the quantum concept was demonstrated for the motion of material particles. *See* PLANCK'S CONSTANT; QUANTUM MECHANICS.

Quantized elastic waves. In 1911, Einstein realized a shortcoming in his theory: not only did it lead to a drop-off of the specific heat capacity at low temperatures faster than observed experimentally (Fig. 1), but it also led to a thermal conductivity of the wrong magnitude and temperature dependence. He also realized the cause for this problem: he had assumed that each atom vibrates independently of its neighbors; however, every atom, as it vibrates, will push and pull on its neighbors, and this will lead to a coupling between neighboring atoms. M. Born and T. von Kármám (1912) showed that this coupling will lead to elastic waves, which will propagate with the velocity of sound through the solid. Using the analogy with electromagnetic waves, P. Debye (1912) showed that these elastic waves lead to a slower temperature dependence of the specific heat capacity at low temperatures. According to Debye, the molar heat capacity is given at low temperatures ($T < 0.1\Theta_D$) by Eq. (3), where Θ_D is called

$$c = 234R\left(\frac{T}{\Theta_D}\right)^3 \tag{3}$$

the Debye characteristic temperature and is determined by the elastic constants of the solid, which are measured.

Figure 1 shows the specific heat capacity of copper predicted by Debye's theory. Another comparison is shown in **Fig. 2**, for crystalline quartz. The agreement below 10 K (−442°F) for quartz is excellent and demonstrates the validity of the theoretical model. The discrepancy between theory and experiment above this temperature is also well understood: Not all the elastic waves in quartz have the same speeds of sound; the high-frequency ones, which will be excited predominantly at high temperatures, are particularly variable. This variation of the speeds of sound leads to variations in the number density of the elastic waves that can be thermally excited, and hence will lead to variations in the specific heat.

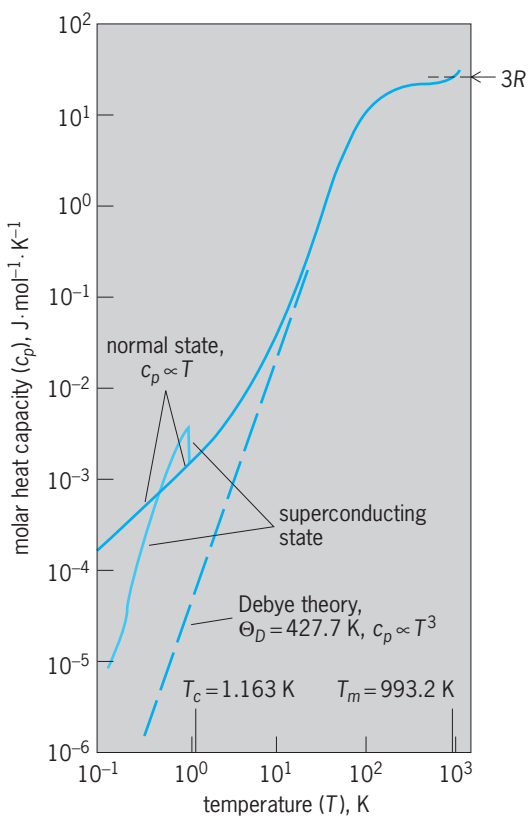

Fig. 3. Molar heat capacity at constant pressure (c_p) of aluminum. The specific heat capacity is obtained by dividing the molar heat capacity by the molecular weight of aluminum (26.982).

See LATTICE VIBRATIONS; THERMAL CONDUCTION IN SOLIDS.

Difference of specific heats. While in gases the difference between the specific heat capacity at constant pressure and that at constant volume is significant, in solids the difference is relatively small. Therefore, the specific heat is usually measured under constant pressure p (usually $p = 0$, vacuum). In the example shown in **Fig. 3**, the molar heat capacity under constant pressure (c_p) of aluminum above room temperature does not saturate at the value given by the Dulong-Petit rule, which is strictly valid only for c_v; the gradual increase of c_p above room temperature is the result of the thermal expansion of the solid. Above 700 K (800°F), there is also a sudden increase of the specific heat as the melting point T_m (933.2 K or 1220.1°F) is approached. *See* THERMAL EXPANSION.

Electronic specific heat. A typical metal like copper or aluminum contains roughly as many free electrons as atoms. According to classical theory, each electron should contain a thermal energy of $(^1/_2)k_BT$ for each degree of freedom, three in the case of translational motion. It is, therefore, remarkable that the Debye theory yields such excellent agreement with the experiment for copper, a metal (Fig. 1), considering that this theory completely ignores an electronic contribution. In fact, free electrons are noticed only at lower temperatures. In aluminum, the lattice specific heat according to Debye's theory, shown in Fig. 3, agrees well with the measurements above 20 K (−424°F). Below 1 K (−458°F), the experimental molar heat capacity approaches a linear temperature dependence, also shown in Fig. 3, that is given by Eq. (4). At 1.163 K, aluminum becomes superconducting: its electrical resistivity suddenly drops to zero. By applying a magnetic field (∼100 gauss or 0.01 tesla), this transition can be suppressed, and the linear temperature dependence of the electronic specific heat can be observed to the lowest temperatures. The electronic specific heat in the normal (that is, nonsuperconducting) state can be understood on the basis of the Pauli exclusion principle: each quantum state can be occupied by at most one electron, and thus only electrons occupying the uppermost energy states have unoccupied energy states nearby, to which they can move by picking up thermal energy. According to a theory due to A. Sommerfeld, this leads to a molar specific heat given by Eq. (5), where T_F is a characteristic temperature called the Fermi temperature, which is of the order of 10^5 K. In deriving this theory, it has been assumed that each atom in the metal supplies one conduction electron. *See* EXCLUSION PRINCIPLE; FERMI-DIRAC STATISTICS; FREE-ELECTRON THEORY OF METALS; SUPERCONDUCTIVITY.

$$c_p = 1.35 \times 10^{-3} \frac{\mathrm{J}}{\mathrm{mol} \cdot \mathrm{K}^2} T \tag{4}$$

$$c_V = \frac{\pi^2}{2} \frac{T}{T_F} R \tag{5}$$

When the metal sample is cooled in the absence of a magnetic field below its superconducting transition temperature T_c (1.163 K in aluminum), the specific heat rises abruptly as the electronic energy states are rearranged. Upon further cooling, the specific heat decreases exponentially. However, even at 0.2 K the specific heat of the electrons exceeds that of the atomic vibrations in aluminum (according to the Debye theory) by at least tenfold.

Specific heat of phase transitions. The transition from the normal to the superconducting state shown in Fig. 3 is an example of a phase transition occurring among the metal electrons. Other phase transitions in the solid state are, for example, those from magnetic disorder (paramagnetic) to magnetic order (ferro-, ferri-, or antiferromagnetic), or from electric dipolar disorder (paraelectric) to an ordered phase (for example, ferroelectric). The transition from an ordered to a disordered phase ordinarily occurs at a critical temperature T_c, near which fluctuations between regions of order and disorder cause an increase of the specific heat capacity. Thus a graph of specific heat capacity versus temperature T has a sharp peak at $T = T_c$. *See* CRITICAL PHENOMENA; PHASE TRANSITIONS.

While much progress has been achieved in the understanding of the approach to phase transitions in the solid state, including the specific heat near the critical temperature T_c, the solid-liquid phase transition (for example, the rapid rise of the specific heat of aluminum near the melting point, shown in Fig. 3) is still only poorly understood.

Amorphous solids. In contrast to crystals, these solids lack long-range order. Their structure resembles that of frozen-in liquids. The specific heat of these solids exceeds that of the Debye prediction, even if they are electrical insulators, notably at the lowest temperatures; below 1 K, it varies approximately linearly with the temperature and has almost the same magnitude (to within a factor of 10) in all amorphous solids. (In amorphous metals, the low-temperature specific heat will be increased further by the contribution by the electrons.) Figure 2 shows a comparison of the specific heat capacity of the electric insulator silicon dioxide (SiO_2) in the amorphous (silica) and the crystalline (α-quartz) phases. The Debye theory predicts nearly the same heat capacity for both phases, which is therefore shown as a single curve. Obviously, the Debye model is only a poor approximation in the amorphous phase. The physical nature of the excitations giving rise to the linear specific heat anomaly is not yet understood. A particular challenge is the universality of these excitations. *See* AMORPHOUS SOLID.

Robert O. Pohl

Bibliography. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1976; A. Cezairliyan (ed.), *Specific Heat of Solids*, 1988; E. S. R. Gopal, *Specific Heats at Low Temperatures*, 1966; C. Kittel, *Introduction to Solid State Physics*, 8th ed., 2005; Y. S. Touloukian et al. (eds.), *Thermophysical Properties of Matter*, vols. 4–6, 1970; G. K. White and P. J. Meeson, *Experimental Techniques in Low-Temperature Physics*, 4th ed., 2002.

Specific impulse

The impulse produced by a rocket divided by the mass m_p of propellant consumed. Specific impulse I_{sp} is a widely used measure of performance for chemical, nuclear, and electric rockets. It is usually given in seconds for both U.S. Customary and International System (SI) units, but this is not strictly correct.

The impulse produced by a rocket is the thrust force F times its duration t in seconds. The specific impulse is given by Eq. (1). Its equivalent,

$$I_{sp} = \frac{Ft}{m_p} \tag{1}$$

specific thrust F_{sp}, that is sometimes used alternatively, is the rocket thrust divided by the propellant mass flow rate $F/\dot{m}_p$. *See* IMPULSE (MECHANICS); THRUST.

The rationale for using seconds as the units of specific impulse is the lack of distinction in U.S. Customary usage between pound-force (lbf) and pound-mass (lbm) that are numerically equivalent, so that cancellation leaves units of seconds; however, this procedure cannot be justified in SI units, where the newton is the base unit of force and the kilogram the unit of mass, so that the unit of specific impulse is N·s/kg. *See* UNITS OF MEASUREMENT.

Usage of the quantity effective jet velocity V_{je} is preferred to specific impulse, since it is a more fundamental and better conceived measure of performance. It is related to specific impulse by Eq. (2),

$$V_{je} = g_c I_{sp} \tag{2}$$

where g_c, the standard gravitational acceleration, is 32.18 ft/s^2 or 9.807 m/s^2, to give V_{je} in units of ft/s and m/s, respectively.

Effective jet velocity for thermal rockets can be calculated from Eq. (3), where R is the universal gas constant, k is the propellant specific heat ratio, T_m is the maximum propellant temperature, M is the molecular mass of the propellants, and p_e/p_c is the ratio of nozzle exit to chamber pressure. The importance of high propellant temperature and low molecular mass of the propellant is shown by this equation.

$$V_{je} = \sqrt{\frac{2g_c Rk}{k-1}\frac{T_m}{M}\left[1-\left(\frac{p_e}{p_c}\right)^{\frac{(k-1)}{k}}\right]} \tag{3}$$

Nominal rocket performance characteristics of typical space propulsion systems

Type	Propellants	Specific impulse lbf-s/lbm	Specific impulse N-s/kg	Effective jet velocity ft/s	Effective jet velocity m/s	Status*	Examples
Thermal rockets							
Cold gas rockets	Nitrogen	50	490	1,609	490	O	Manned Maneuvering Unit (MMU)
Chemical rockets							
Solid propellant	Ammonium perchlorate/organic polymers and powdered aluminum	290	2,844	9,332	2,844	O	Minuteman Space shuttle boosters Titan boosters
Liquid propellant							
Monopropellant	Hydrazine	230	2,256	7,401	2,256	O	Attitude control of communication satellites
Bipropellant	Nitrogen tetroxide/monomethyl hydrazine (hydrazine)	310 (315)	3,040 (3,089)	9,976 (10,137)	3,040 (3,089)	O	Titan launch vehicle
	Oxygen/kerosine (RP-1)	350	3,432	11,260	3,432	O	V-2†
	Oxygen/hydrogen	450	4,413	14,480	4,413	O	Space shuttle main engine
Nuclear rockets	Hydrogen	900	8,826	28,960	8,826	R&T	
Electric rockets							
Electrothermal rockets							
Resistojet	Hydrazine	300	2,942	9,654	2,942	O	Station keeping for *Satcom* communications satellite
	Ammonia	500	4,904	16,090	4,904	R&T	
	Hydrogen	700	6,865	22,530	6,865	R&T	
Arcjet	Hydrazine	500	4,904	16,090	4,904	R&T	
	Ammonia	830	8,140	26,710	8,140	R&T	
	Hydrogen	1,250	12,260	40,220	12,250	R&T	
Electromagnetic rockets							
Pulsed plasma	Poly(tetrafluoroethylene)	1,000	9,807	32,180	9,807	O	Orbit control for *Nova* navigation satellite
	Argon	1,500	14,710	48,270	14,710	R&T	
Magnetoplasmadynamic							
Electrostatic rockets							
Ion	Mercury, argon, xenon	2,000 to 6,000	19,610 to 58,840	64,360 to 193,100	19,610 to 58,840	ST&E	

*O = operational; R&T = research and technology; ST&E = space test and evaluation.
† No longer used.

Calculation of specific impulse for the various forms of electric rockets involves electrothermal, resistance or arc heating of the propellant or its ionization and acceleration to high jet velocity by electrostatic or electromagnetic body forces. Ions in the exhaust jets of these devices must be neutralized so the spacecraft will not suffer from space charging or other effects from the plumes of the devices' operation. Nominal values of performance of typical space propulsion systems with various types of rockets and their propellants are given in the **table**. *See* ELECTROTHERMAL PROPULSION; ION PROPULSION; PLASMA PROPULSION; ROCKET PROPULSION; SPACECRAFT PROPULSION. J. Preston Layton

Bibliography. G. P. Sutton, *Rocket Propulsion Elements*, 7th ed., Wiley, 2001.

Speckle

The generation of a random intensity distribution, called a speckle pattern, when light from a highly coherent source, such as a laser, is scattered by a rough surface or inhomogeneous medium. Although the speckle phenomenon has been known since the time of Isaac Newton, the development of the laser is responsible for the present-day interest in speckle. Speckle has proved to be a universal nuisance as far as most laser applications are concerned, and only in the mid-1970s did investigators turn from the unwanted aspects of speckle toward the uses of speckle patterns, in a wide variety of applications. *See* LASER.

Basic phenomenon. Objects viewed in coherent light acquire a granular appearance. The detailed irradiance distribution of this granularity appears to have no obvious relationship to the microscopic properties of the illuminated object, but rather it is an irregular pattern that is best described by the methods of probability theory and statistics. Although the mathematical description of the observed granularity is rather complex, the physical origin of the observed speckle pattern is easily described. The surfaces of most materials are extremely rough on the scale of an optical wavelength (approximately 5×10^{-7} m). When nearly monochromatic light is reflected from such a surface, the optical wave resulting at any moderately distant point consists of many coherent wavelets, each arising from a different microscopic element of the surface. Since the distances traveled by these various wavelets may differ by several wavelengths if the surface is truly rough, the interference of the wavelets of various phases results in the granular pattern of intensity called speckle. If a surface is imaged with a perfectly corrected optical system (**Fig. 1**), diffraction causes a spread of the light at an image point, so that the intensity at a given image point results from the coherent addition of contributions from many independent surface areas. As long as the diffraction-limited point-spread function of the imaging system is broad by comparison with the microscopic surface variations, many dephased coherent contributions add at each image point to give a speckle pattern.

The basic random interference phenomenon underlying laser speckle exists for sources other than lasers. For example, it explains radar "clutter," results for scattering of x-rays by liquids, and electron scattering by amorphous carbon films. Speckle theory also explains why twinkling may be observed for stars, but not for planets. *See* COHERENCE; DIFFRACTION; INTERFERENCE OF WAVES; TWINKLING STARS.

Applications. The principal applications for speckle patterns fall into two areas: metrology and astronomical speckle interferometry.

Metrology. In the metrology area, the most obvious application of speckle is to the measurement of surface roughness. If a speckle pattern is produced by coherent light incident on a rough surface, then surely the speckle pattern, or at least the statistics of the speckle pattern, must depend upon the detailed surface properties. If the surface root-mean-square roughness is small compared to one wavelength of the fully coherent radiation used to illuminate the surface, the roughness can be determined by measuring the speckle contrast. If the root-mean-square roughness is large compared to one wavelength, the radiation should be spatially coherent polychromatic, instead of monochromatic; the roughness is again determined by measuring the speckle contrast.

An application of growing importance in engineering is the use of speckle patterns in the study of object displacements, vibration, and distortion that arise in nondestructive testing of mechanical components. The key advantage of speckle methods in this case is that the speckle size can be adjusted to suit the resolution of the most convenient detector, whether it be film or a television camera, while still retaining information about displacements with an accuracy of a small fraction of a micrometer. Although several techniques for using laser speckle in metrology are available, the basic idea can be explained by the following example for measuring vibration. A vibrating surface illuminated with a laser beam is imaged (Fig. 1). A portion of the laser beam is also superimposed on the image of the vibrating surface. If the object surface is moving in and out with a travel of one-quarter of a wavelength or more, the speckle pattern will become blurred out. However, for areas of the surface that are not vibrating (that is, the nodal regions) the eye will be able to distinguish the fully developed high-contrast speckle pattern.

Electronic speckle interferometry (ESPI) is finding much use in industrial applications involving

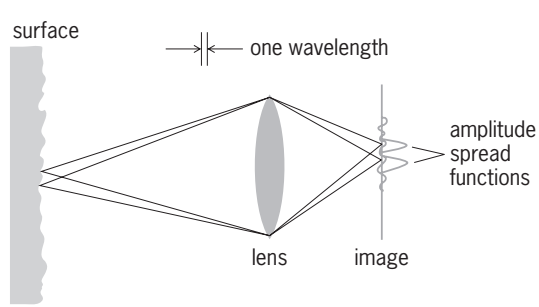

Fig. 1. Physical origin of speckle for an imaging system.

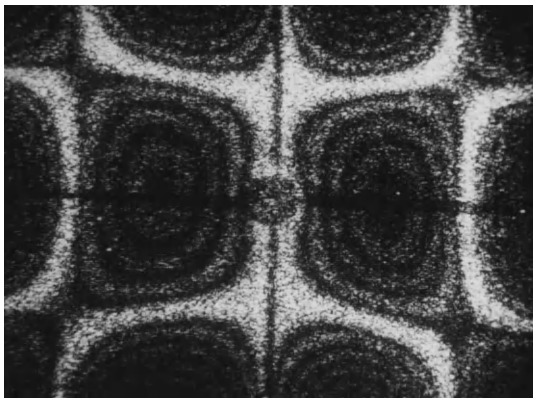

Fig. 2. Electronic speckle interferogram of vibrating object.

measurement of the vibrational properties of structures. In this technique, a television camera is used to view a speckle pattern. The output of the television camera is processed electronically to emphasize the high spatial frequency structure. The result is viewed on a television monitor (**Fig. 2**). The bright fringes correspond to the region of the object which are at the nodes of the vibration pattern. By knowing the angles at which the object is illuminated and viewed, it is possible to calculate the amplitude of the vibration from the number of dark fringes present. Thus, simply by looking at the television monitor it is possible to determine the vibrational pattern over the entire object. The amplitude and frequency of the drive source can be changed and instantly the new vibrational pattern can be observed. *See* NONDESTRUCTIVE EVALUATION.

Speckle patterns can also be used to examine the state of refraction of the eye. If a diffusing surface illuminated by laser light moves perpendicular to the line of sight of an observer, the speckles may appear to move with respect to the surface. For a normal eye, movement in a direction opposite to that of the surface indicates underaccommodation, and movement with the surface indicates overaccommodation. If the speckles do not move, but just seem to "boil," the observed surface is imaged on the retina.

James C. Wyant

Astronomical speckle interferometry. Astronomical speckle interferometry is a technique for obtaining spatial information on astronomical objects at the diffraction-limited resolution of a telescope, despite the presence of atmospheric turbulence.

In the absence of turbulence, a telescope of aperture D operating at wavelength λ would produce a diffraction-limited image of a point source of an angular size $\alpha = 1.22\ \lambda/D$ (the Rayleigh criterion). When observed through an Earth-based telescope, the image of a source in space is corrupted by distortions introduced during propagation of its light through the Earth's atmosphere. These distorting effects result from variations in the density, temperature, and index of refraction of the atmosphere, and give a blurred image, or seeing disk, which is independent of telescope aperture and seldom less than 1–2 arc-seconds in diameter. A sub-arc-second seeing disk is observed only under exceptionally good seeing conditions, at the most favorable high-altitude telescope locations, when the telescope environment is characterized by uniform temperature and nonturbulent conditions. Even in these optimal conditions, the seeing disk is dozens of times larger than the diffraction-limited angular size of the point source produced by a telescope of several meters' aperture. For example the diffraction-limited angular size of a point source produced at $\lambda = 500$ nanometers with a 157-in. (4-m) telescope is approximately 0.03 arc-second, and the seeing disk obtained at the very best observing conditions is at least 10 times larger.

At very short exposures (10–20 milliseconds), astronomical images consist of a fine-scale structure that fills the seeing disk, very similar to the speckle structures observed with laser-illuminated diffusers. These individual speckles contain spatial frequency information up to the diffraction limit of the telescope. Speckle interferometry techniques extract this information, and therefore reach spatial detail down to the diffraction-limited resolution of the telescope. In this process, long sequences of short-exposure images (each exposure a few milliseconds in duration) are recorded through magnifying optics and narrow-band filters by using detectors such as charge-coupled devices (CCDs) and photon-counting devices. These images are then combined by using various computational algorithms to reconstruct a single image with diffraction-limited resolution. The number of images required for convergence of speckle image reconstruction algorithms increases as the fourth power of the atmospheric seeing. *See* CHARGE-COUPLED DEVICES; IMAGE PROCESSING.

Speckle interferometry techniques have proven to be an invaluable tool for astronomical research, allowing studies of a wide range of scientifically interesting problems. They have been widely used to determine the separation and position angle of binary stars, and for accurate diameter measurements of a large number of stars, planets, and asteroids. Speckle imaging techniques have successfully uncovered details in the morphology of a range of astronomical objects, including the Sun, planets, asteroids, cool giants and supergiants, young stellar objects, the supernova SN1987A in the Large Magellanic Cloud, Seyfert galaxies, and quasars. *See* BINARY STAR; INTERFEROMETRY.

Margarita Karovska

Bibliography. M. Francon, *Laser Speckle and Applications in Optics*, 1979; R. Jones and C. M. Wykes, *Holographic and Speckle Interferometry*, 2d ed., 1989; S. R. Robinson (ed.), *The Infrared and Electro-Optical Systems Handbook*, 1993; R. S. Sirohi, *Speckle Metrology*, 1993.

Spectral type

A label used to indicate the physical and chemical characteristics of a star, as indicated by study of the star's spectrum. Stars possess a remarkable variety of spectra, some simple, others extraordinarily

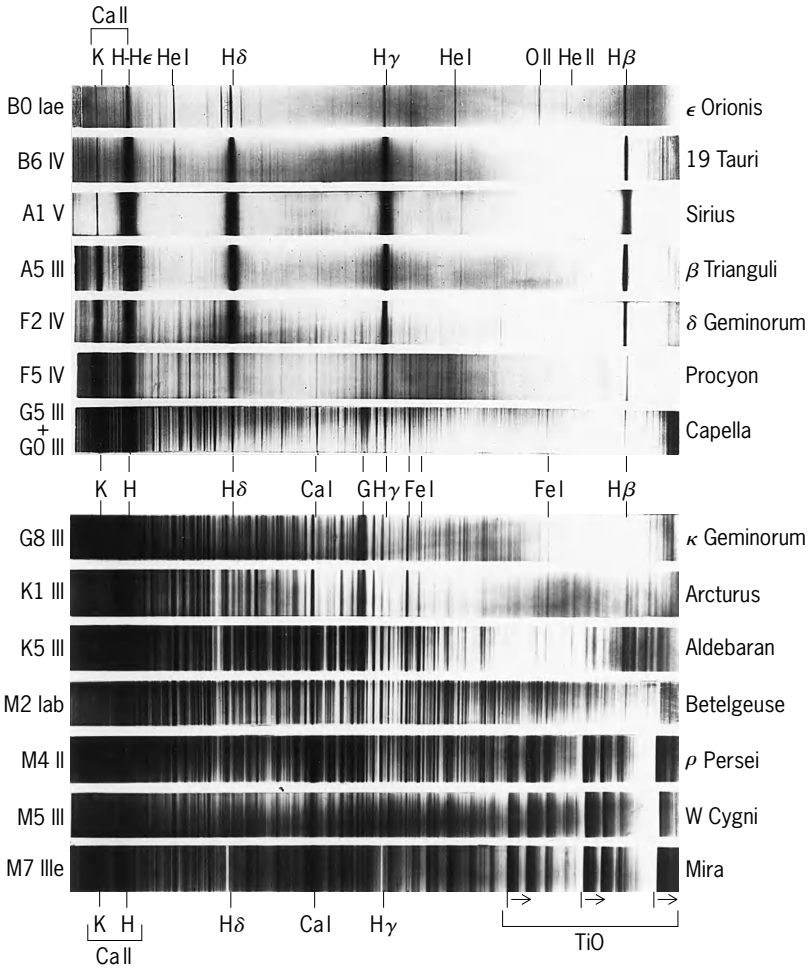

Fig. 1. Spectral sequence from B to M. (*University of Michigan*)

complex. To understand the natures of the stars, it was first necessary to bring order to the subject and to classify the spectra.

Standard sequence. Based in part on earlier work by A. Secchi, who had developed a system that correlated with stellar color, W. C. Pickering initiated the modern system starting about 1890. He ordered the spectra by letter, A through O, largely on the basis of the strengths of the hydrogen lines. He and Mrs. W. P. Fleming then found that several of the letters were unneeded or redundant. On the basis of continuity of lines other than hydrogen, A. Maury and A. J. Cannon found that B preceded A and O preceded B. The result is the classical spectral sequence, OBAFGKM (**Fig. 1**). Cannon also decimalized the classes, setting up the sequence O5, . . . , O9, B0, . . . , B9, A0, and so forth. (Not all the numbers are used, however.) The modern Harvard sequence, after the observatory where it was formulated, runs from O2 to M9. Over 350,000 stars, classified by Cannon, were included in the final result, the *Henry Draper Catalog*. Classes L and T were added to the end of the sequence in 1999. *See* ASTRONOMICAL CATALOGS.

Class A has the strongest hydrogen lines, B is characterized principally by neutral helium (with weaker hydrogen), and O by ionized helium. Hydrogen weakens notably through F and G, but the metal lines, particularly those of ionized calcium, strengthen. In K, hydrogen becomes quite weak, while the neutral metals grow stronger. The M stars effectively exhibit no hydrogen lines at all, but are dominated by molecules, particularly titanium oxide (TiO). L stars are dominated by metallic hydrides and neutral alkali metals, while T is defined by methane, ammonia, and water. At G, the sequence branches into R and N, whose stars are rich in carbon molecules; class S, in which the titanium oxide molecular bands of class M are replaced by zirconium oxide (ZrO), was introduced in 1923 (**Fig. 2**). Classes R and N are commonly combined into class C.

Spectral variation with temperature. At first appearance, the different spectral types seem to reflect differences in stellar composition. However, within the sequence OBAFGKMLT the elemental abundances are roughly similar. The dramatic variations seen in Fig. 1 are the result of changes in temperature (**Fig. 3**). The strength of an ion's absorption lines depends on the number of electrons capable of absorbing at those wavelengths, which in turn depends on the product of the atomic absorption probability, the number of electrons excited to the appropriate level of the ion, and the number of ions. In classes M and L (which divide near 2000 K), the temperature is too low to allow a significant number of electrons in the second level of hydrogen, from which the optical Balmer series arises. As a result, there are no hydrogen lines. As temperature climbs and collisional activity increases, more electrons are knocked into level 2 and the Balmer lines strengthen from K through A. Above class A, however, the temperature becomes so high that the hydrogen becomes ionized. That ionization and changes in the opacity of the stellar atmosphere cause the hydrogen lines to weaken.

The other ions are subject to the same kind of variation. The second level of neutral helium lies at a very high energy. As a result, the helium lines do not become visible until class B. It also takes a great deal of energy to ionize helium, so the ionized helium lines are not observed until nearly class O. Ionized metal lines are produced at intermediate temperatures and neutral metal lines at low temperatures. Molecules, which are relatively fragile, cannot exist unless the temperature is relatively low.

Spectral variation with composition. The different spectra of the R, N, and S stars, however, are caused by true and dramatic variations in the chemical composition. The M stars have more oxygen than carbon. In the N stars, which have similar temperatures, the ratio is reversed and carbon dominates oxygen. The

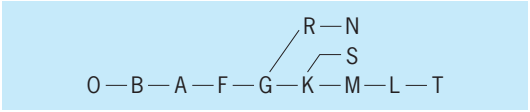

Fig. 2. Spectral classes. The stars in the sequence O through T have common chemical compositions. Those in the side branches are rich in carbon.

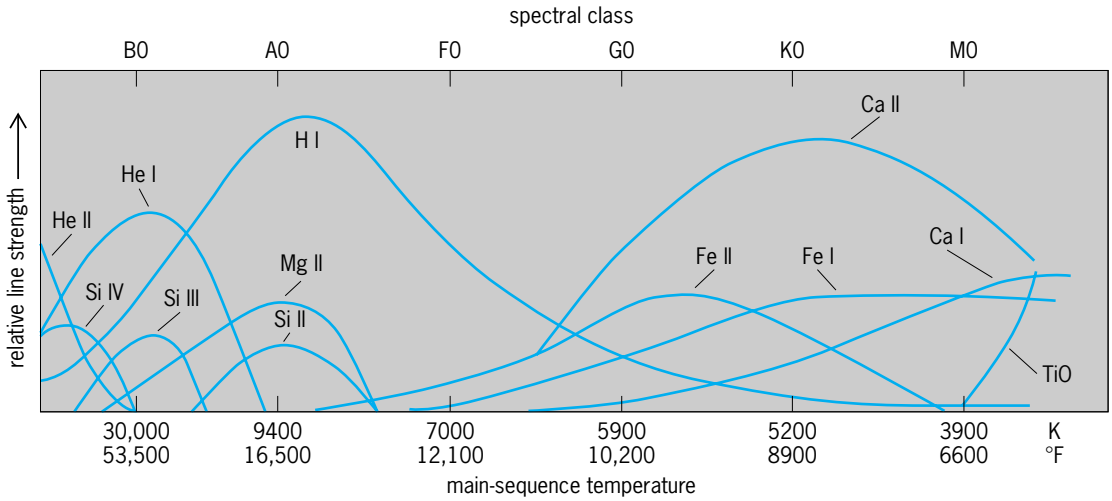

Fig. 3. Variation in line strength for different ions as a function of temperature and spectral class. In class L (off the right end of the graph), TiO disappears and is replaced by metallic hydrides. (*After H. Voigt, ed., Landolt-Bornstein Numerical Data and Functional Relationships in Science and Technology, Group VI, vol. 1: Astronomy and Astrophysics, Springer-Verlag, 1965*)

R stars are carbon-rich versions of classes G and K. In the cool S stars, carbon and oxygen have about the same abundance. That condition and the increased abundance of zirconium (which has a greater affinity for oxygen than titanium) yield strong zirconium oxide lines. These composition variations are the result of internal thermonuclear processing and convection. *See* STELLAR EVOLUTION.

Morgan-Keenan-Kellman system. In the 1940s, W. W. Morgan, P. C. Keenan, and E. Kellman expanded the Harvard sequence to include luminosity. A system of roman numerals is appended to the Harvard class to indicate position on the Hertzsprung-Russell diagram: I for supergiant, II for bright giant, III for giant, IV for subgiant, and V for dwarf or main sequence. The Morgan-Keenan-Kellman class, which includes the refined Harvard type, is defined by a set of standard stars in the MKK atlas, to which a program star must be compared. On this system, the Sun is a G2 V star. Though some dwarfs are carbon-rich, the R, N, and S stars are all giants. L stars are all dwarfs and brown dwarfs, and T stars are all brown dwarfs. Additional refinements have been added to indicate a variety of spectral peculiarities. *See* ASTRONOMICAL SPECTROSCOPY; HERTZSPRUNG-RUSSELL DIAGRAM; STAR.

James B. Kaler

Bibliography. J. B. Hearnshaw, *The Analysis of Starlight: One Hundred Fifty Years of Astronomical Spectroscopy*, Cambridge University Press, 1986; C. Jaschek and M. Jaschek, *The Classification of Stars*, 1990; J. B. Kaler, *Stars and Their Spectra*, Cambridge University Press, 1989, paperback 1997; J. B. Kaler, *The Cambridge Encyclopedia of Stars*, Cambridge University Press, 2006; W. W. Morgan, P. C. Keenan, and E. Kellman, *An Atlas of Stellar Spectra*, University of Chicago Press, 1943.

Spectrograph

An optical instrument that consists of an entrance slit, collimator, disperser, camera, and detector and that produces and records a spectrum. A spectrograph is used to extract a variety of information about the conditions that exist where light originates and along its paths. It does this by revealing the details that are stored in the light's spectral distribution, whether this light is from a source in the laboratory or a quasistellar object a billion light-years away.

Proper spectrograph design takes into account the type of light sources to be measured, and the circumstances under which these measurements will be made. Since observational astronomy presents unusual problems in these areas, the design of astronomical spectrographs may also be unique.

Astronomical versus laboratory design. Astronomical spectrographs have the same general features as laboratory spectrographs (**Fig. 1**). The width of the entrance slit influences both spectral resolution and the amount of light entering the spectrograph, two of the most important variables in spectroscopy. The collimator makes this light parallel so that the disperser (a grating or prism) may properly disperse it. The camera then focuses the dispersed spectrum onto a detector, which records it for further study.

Laboratory spectrographs usually function properly only in a fixed orientation under controlled environmental conditions. By contrast, most astronomical spectrographs are used on a moving telescope operating at local temperature. Thus, their structures must be mechanically and optically insensitive to orientation and temperature.

The brightness, spectral characteristics, and geometry of laboratory sources may be tailored to

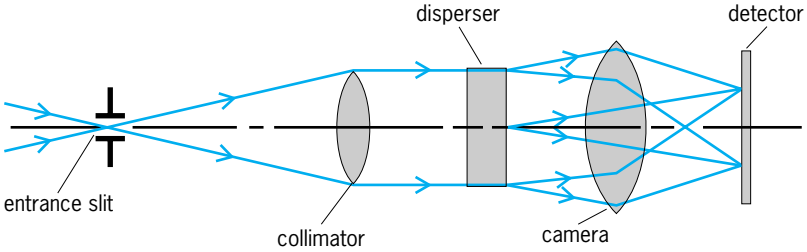

Fig. 1. Basic optical components of a spectrograph.

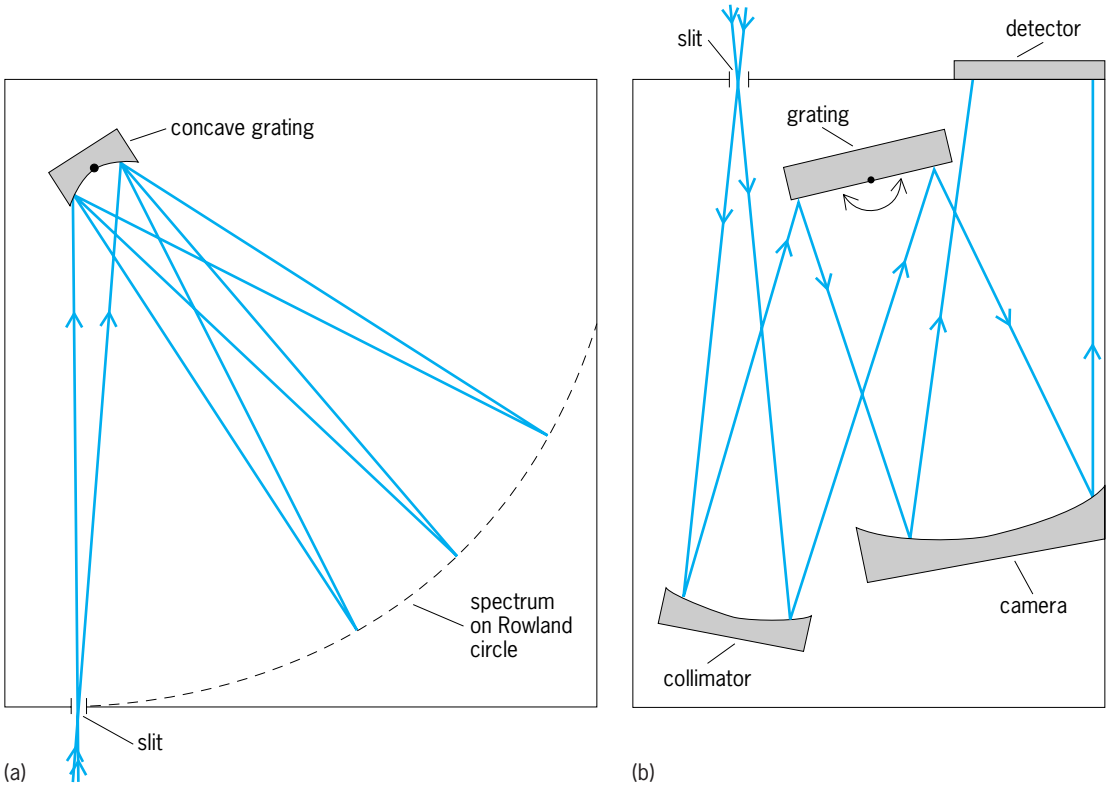

Fig. 2. Typical laboratory spectrographs. (a) Classical concave grating spectrograph. (b) Czerny-Turner plane grating spectrograph.

experimental requirements and to the capabilities of a spectrograph. Astronomical sources, in the form of images at the focus of a telescope, cannot be manipulated, and their faintness and spectral diversity make unusual and difficult demands on spectrograph performance.

Typical laboratory spectrographs use either concave gratings, which effectively combine the functions of collimator, grating, and camera in one optical element, or plan reflection gratings with spherical reflectors for collimators and cameras. An example of each is shown in **Fig. 2**. These designs may severely constrain the optical characteristics of the collimator and camera, and sometimes the grating, to obtain acceptable imaging at the detector. As a result, they are generally too limited in the choice of important spectrograph parameters for use in astronomy. In particular, they tend to preclude the use of very short cameras, an essential feature of spectrographs intended for study of faint astronomical objects. The Czerny-Turner design (Fig. 2*b*) is used in some solar spectrographs, where it offers some advantages and ample light is available. *See* SPECTROHELIOGRAPH; SUN.

To provide maximum flexibility in its application, the collimators, gratings, and cameras of an astronomical spectrograph must be relatively free of aberrations when used in collimated light. Then the optical subassemblies can be interchanged without disturbing the spectrograph's optical performance. Grating interchange changes spectral resolution and wavelength coverage with minimal effect on efficiency. Camera and detector interchange achieves operating characteristics that may be crucial for certain applications. Examples of typical astronomical spectrographs are shown in **Fig. 3**.

Details of astronomical spectrographs. Astronomical spectrographs typically use a pair of adjustable polished slit jaws, which are tilted to the telescope beam so that unused light is reflected to guide optics for accurate image positioning. The collimator may be either an off-axis paraboloid or a two-mirror Cassegrain-type optical system. It has a relatively long focal length and an $f/$ ratio matching that of the telescope's for optimum slit transmission.

The disperser is usually a plane reflection grating. In laboratory applications, larger gratings are generally used for their increased spectral resolution. However, in astronomy the resolution of the grating (number of grooves times interference order) is seldom realized, and larger gratings are used to obtain longer collimator focal lengths in order to preserve or increase slit demagnification at the camera's focus.

The spectral resolution of many laboratory spectrographs can be increased (grating permitting) only by decreasing the slit width. However, this decrease reduces available light and taxes the imaging capabilities of the optical system and the resolution capabilities of the detector. A better approach, usually used in astronomical spectrographs, is to keep the slit width within a relatively small range of values, for a given camera focal length, and vary spectral resolution by using gratings having different angular dispersions.

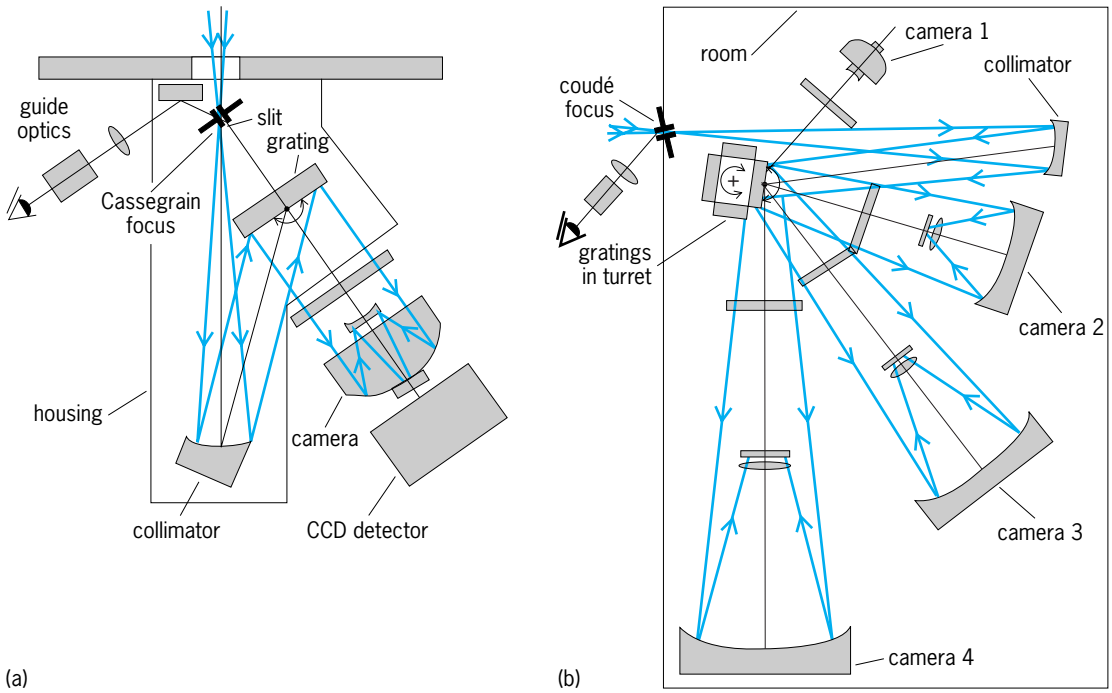

Fig. 3. Typical astronomical spectrographs. (a) Spectrograph for use at the Cassegrain focus (often called a Cassegrain spectrograph) using a parabolic collimator and semisolid Schmidt camera. (b) Plan view of a coude spectrograph with four Schmidt cameras and four interchangeable gratings. Camera 1 is the shortest camera and camera 4 is the longest. The spectrograph is shown using camera 2.

The cameras used in astronomy are often quite short (fast) to allow the entrance slit to be as wide as possible for maximum light transmission and to maximize spectral coverage for a given detector size. The camera is often a variation of the classical Schmidt telescope, using a combination of mirrors and ultraviolet transmitting optics. Some spectrographs may have a selection of cameras to permit higher resolutions (with narrower slit openings) when there is sufficient light, or to improve imaging in some wavelength ranges. In particular, very large spectrographs, used at the coude or Nasmyth foci, may have an array of camera focal lengths (Fig. 3*b*). When this arrangement is combined with interchangeable gratings, it permits a large selection of spectral resolutions and wavelength coverages. *See* SCHMIDT CAMERA.

In the past, the detector is astronomical spectroscopy was almost always the photographic emulsion. However, image intensifiers and, more recently, the very efficient and accurate cooled charge-coupled devices (CCDs) have become the preferred detectors. *See* CHARGE-COUPLED DEVICES; IMAGE TUBE (ASTRONOMY).

Observing sites and telescopes. Spectrographs are particularly appropriate for use at poor observing sites because, although the atmosphere might reduce available light, it has little effect on its spectrum. Light pollution is also less of a problem, since dispersing the light both dilutes the intensity of the unwanted background and identifies the spectrum of this background so that it can be subtracted later.

Astronomical spectrographs can be used effectively on smaller telescopes. The reason is that star images are typically larger than the entrance-slit width, and the amount of light passing through the slit is thus proportional to the diameter, not area, of a telescope's aperture. Thus, the magnitude limit of a 50-in. (1.25-m) telescope will be only about 1 stellar magnitude fainter than that of a far less costly 20-in. (0.5-m) telescope using the same spectrograph. *See* ASTRONOMICAL SPECTROSCOPY.

Ron Hilliard

Bibliography. W. A. Hiltner (ed.), *Astronomical Techniques*, 2d ed., 1962; J. James, *Spectrograph Design Fundamentals*, Cambridge University Press, 2007; J. F. James and R. S. Sternberg, *The Design of Optical Spectrometers*, 1969; C. R. Kitchin, *Astrophysical Techniques*, 4th ed., 2003; J. A. Samson and D. L. Ederer (eds.), *Vacuum Ultraviolet Spectroscopy*, Academic Press, 2000; D. J. Schroeder, *Astronomical Optics*, 2d ed., Academic Press, 2000.

Spectroheliograph

A spectrographic instrument that produces monochromatic images of the Sun. The principle of operation was first suggested by J. Janssen in 1869, and a photographic version was constructed by G. E. Hale at Mount Wilson Observatory in 1892. Around the same time, H. A. Deslandres devised a closely related instrument for operation at l'Obervatoire de Paris, Meudon. In a simple form of the instrument, an image of the Sun from a solar telescope is focused on a plane containing the entrance slit of the spectroheliograph (see **illustration**). The light passing through the slit is collimated by a concave mirror

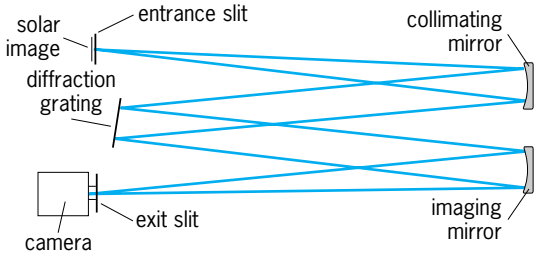

Optical configuration of a simple spectroheliograph.

that is tilted such that the light is incident on a plane diffraction grating. Part of the dispersed light from the grating is focused by a second concave mirror, identical to the first mirror, at an exit slit identical to the entrance slit. By symmetry of the optical system, the portion of the solar disk imaged on the entrance slit is reimaged in the plane of the exit slit with the same image scale but in dispersed wavelength. The light imaged along the exit slit then corresponds to the portion of the solar image falling on the entrance slit, but in the light of only a narrow region of the spectrum, as determined by the spectrographic dispersion. The particular wavelength sampled is set by the grating angle. By uniform transverse motion of the instrument such that the entrance slit is scanned across the solar image, the light passing through the exit slit maps out a corresponding monochromatic image of the Sun, which can be recorded photographically with a stationary camera. Alternatively, the solar image can be scanned across the entrance slit of a stationary spectroheliograph, the camera then synchronously moved in step with the image. *See* DIFFRACTION GRATING.

Narrow-bandpass tunable filters are also used to obtain monochromatic images of the Sun, with the advantage that the full image is recorded in one relatively short exposure. However, the transmittance of such filters is not as spectrally pure as that of a spectroheliograph of the same nominal bandwidth, and extremely narrow passband filters have restricted angular fields. Spectroheliographs also have the advantage that the slit widths can be adjusted to match lines of different spectral widths or other specific observing requirements, or can be set for different image-quality observing conditions. A disadvantage is that the time required to build up an image of the full Sun with a spetroheliograph that uses a photographic film-type camera can be many minutes, during which time critical solar features observed might change. *See* INTERFERENCE FILTER.

Digital recording has greatly expanded the performance capabilities, versatility, and range of applications of spectoheliographs, reducing the observing time and allowing flexibility in observing programs. Sequential readout of a linear array that is aligned with the exit slit, and that moves with the slit as the solar image is scanned at the entrance slit, provides a fast and convenient observing system. The maximum scan rate is set by the array readout rate. A charge-coupled-device (CCD) array, used directly at the focal plane without a slit, allows simultaneous recording of an image at the wavelength of a spectral line as well as the adjacent spectral region. Computer processing of the digital data can then provide velocity and magnetic field information, and extensive solar fields can be recorded in a few seconds. Digital video camera systems have been used for such applications. Without an exit slit, the optical configuration in this case is simply that of a spectrograph, except for an image scanning capability. *See* ASTRONOMICAL IMAGING; CHARGE-COUPLED DEVICES; SPECTROGRAPH.

Spectroheliographs have been designed that have more than one camera to record several spectral regions simultaneously, either in the same spectral order or in more than one order of the diffraction grating. Spectroheliographs have also been used at ultraviolet, extreme ultraviolet (EUV), and x-ray wavelengths, and a miniature version has been used successfully in spacecraft observations of the solar corona.

Raymond N. Smartt

Bibliography. R. Beck et al. (eds.), *Solar Astronomy Handbook*, Willmann-Bell, Richmond, Virginia, 1995; P. Foukal, *Solar Astrophysics*, 2d ed., Wiley, New York, 2004; K. J. H. Phillips, *Guide to the Sun*, Cambridge University Press, 1992; M. Stix, *The Sun*, Springer-Verlag, Berlin, 1989.

Spectroscopy

An analytic technique concerned with the measurement of the interaction (usually the absorption or the emission) of radiant energy with matter, with the instruments necessary to make such measurements, and with the interpretation of the interaction both at the fundamental level and for practical analysis. Mass spectroscopy is not concerned with the interaction of light with matter, but was so named because the appearance of the data resembles that of the spectroscopic data as just defined.

A display of such data is known as a spectrum, that is, a plot of the intensity of emitted or transmitted radiant energy (or some function of the intensity) versus the energy of that light. Spectra due to the emission of radiant energy are produced as energy is emitted from matter, after some form of excitation, then collimated by passage through a slit, then separated into components of different energy by transmission through a prism (refraction) or by reflection from a ruled grating or a crystalline solid (diffraction), and finally detected. Spectra due to the absorption of radiant energy are produced when radiant energy from a stable source, collimated and separated into its components in a monochromator, passes through the sample whose absorption spectrum is to be measured, and is detected. Instruments which produce spectra are known as spectroscopes, spectrometers, spectrographs, and spectrophotometers. *See* SPECTRUM.

Interpretation of spectra provides fundamental information on atomic and molecular energy levels, the distribution of species within those levels, the nature of processes involving change from one level to another, molecular geometries, chemical bonding,

and interaction of molecules in solution. At the practical level, comparisons of spectra provide a basis for the determination of qualitative chemical composition and chemical structure, and for quantitative chemical analysis.

Early history. In the historical development of spectroscopy, following the fundamental studies of crude spectra of sunlight by Isaac Newton in 1672, certain contributions and achievements are especially noteworthy. The significance of using a narrow slit instead of a pinhole or round aperture so as to produce spectra lines, each one an image of the slit and representing a different color or wave length, was demonstrated independently by W. H. Wollaston in 1802 and by Joseph Fraunhofer in 1814. Fraunhofer made many subsequent contributions to optics and spectroscopy, including first observation of stellar spectra, discovery and construction of transmission diffraction gratings, first accurate measurements of wavelengths of the dark lines in the solar spectrum, and invention of the achromatic telescope. The origin of the dark Fraunhofer lines in the solar spectrum was accounted for by G. R. Kirchhoff in 1859 on the basis of absorption by the elements in the cooler Sun's atmosphere of the continuous spectrum emitted by the hotter interior of the Sun. Further studies by Kirchhoff with R. Bunsen demonstrated the great utility of spectroscopy in chemical analysis. By systematically comparing the Sun's spectrum with flame or spark spectra of salts and metals, they made the first chemical analysis of the Sun's atmosphere. In 1861, while investigating alkali metal spectra, they discovered two new alkali metals, cesium and rubidium. These achievements by Kirchhoff and Bunsen provided tremendous stimulus to spectroscopic researches. The adoption in 1910 of the first international standards of wavelength gave further impetus. These and later standards made possible the measurement of wavelengths of any electromagnetic radiation with unprecedented accuracy. Since World War II, remarkable developments in spectroscopy have occurred in instrumentation, achieved largely through advances in electronics and in manufacturing technology. Direct reading, automatic recording, improved sensitivity with good stability, simplicity of operation, and extended capabilities are features provided by many commercial instruments, many of which now are microprocessor-controlled. Many newer instruments have dedicated data systems. Predictably, these developments, by facilitating widespread use of spectroscopic techniques, have had an enormous influence in promoting developments in both applied and theoretical spectroscopy.

The ultimate standard of wavelength is that of the meter, defined since 1983 as the length of the path traveled by light in vacuum during a time interval of 1/299,792,458 of a second. *See* WAVELENGTH STANDARDS.

Spectroscopic units. The change in energy of an ion, atom, or molecule associated with absorption and emission of radiant energy may be measured by the frequency of the radiant energy according to Max Planck, who described an equality $E = h\nu$, where E is energy, ν is the frequency of the radiant energy, and h is Planck's constant. The frequency is related to the wavelength λ by the relation $\nu\lambda = c/n$, where c is the velocity of radiant energy in a vacuum, and n is the refractive index of the medium through which it passes; n is a measure of the retardation of radiant energy passing through matter. The units most commonly employed to describe these characteristics of light are the following:

Wavelength:	1 micrometer (μm)	$= 10^{-6}$ m
	1 nanometer (nm)	$= 10^{-9}$ m
		(= 10 angstroms)
Frequency:	1 hertz (Hz)	$= 1\ s^{-1}$

For convenience the wave number $\bar{\nu}$ (read nu bar), the reciprocal of the wavelength, may be used; for this, the common units are cm^{-1}, read as reciprocal centimeters or occasionally kaysers. This number equals the number of oscillations per centimeter.

Spectral regions. Visible light constitutes only a small part of the spectrum of radiant energy, or electromagnetic spectrum; the human eye responds to electromagnetic waves with wavelengths from about 380 to 780 nm, though there is individual variation in these limits. The eye cannot measure color and intensity quantitatively, even for the visible portion of the electromagnetic spectrum; therefore, instruments are used for measurements and recording.

There are broad regions of the electromagnetic spectrum with associated wavelengths greater or less than the visible region (see **table**).

Origin of spectra. Atoms, ions, and molecules emit or absorb characteristically; only certain energies of these species are possible; the energy of the photon (quantum of radiant energy) emitted or absorbed corresponds to the difference between two permitted values of the energy of the species, or energy levels. (If the flux of photons incident upon the species is great enough, simultaneous absorption of two or more photons may occur.) Thus the energy levels may be studied by observing the differences between them. The absorption of radiant energy is accompanied by the promotion of the species from a lower to a higher energy level; the emission of radiant energy is accompanied by falling from a higher to a lower state; and if both processes occur together, the condition is called resonance.

Transitions. Transitions between energy levels associated with electrons or electronic levels range from the near infrared, the visible, and ultraviolet for outermost, or highest-energy, electrons, that is, those which can be involved in bonding, to the x-ray region for the electrons nearest the nucleus. At low pressures such transitions in gaseous atoms produce sharply defined lines because the energy levels are sharply defined. Transitions between energy levels of the nucleus are observed in the gamma-ray region. In the absence of an applied electric or magnetic field, these electronic and nuclear transitions are the only ones which atoms can undergo. *See* ATOMIC SPECTROMETRY; ATOMIC STRUCTURE AND SPECTRA.

Electronic transitions in molecules are also observed. In addition, transitions can occur between

Principal spectral regions and fields of spectroscopy

Spectral region	Approximate wavelength range	Typical source	Typical detector	Energy transitions studied in matter
Gamma	1–100 pm	Radioactive nuclei	Geiger counter; scintillation counter	Nuclear transitions and disintegrations
X-rays	6 pm–100 nm	X-ray tube (electron bombardment of metals)	Geiger counter	Ionization by inner electron removal
Vacuum ultraviolet	10–200 nm	High-voltage discharge; high-vacuum spark	Photomultiplier	Ionization by outer electron removal
Ultraviolet	200–400 nm	Hydrogen-discharge lamp	Photomultiplier	Excitation of valence electrons
Visible	400–800 nm	Tungsten lamp	Phototubes	Excitation of valence electrons
Near-infrared	0.8–2.5 μm	Tungsten lamp	Photocells	Excitation of valence electrons; molecular vibrational overtones
Infrared	2.5–50 μm	Nernst glower; Globar lamp	Thermocouple; bolometer	Molecular vibrations: stretching, bending, and rocking
Far-infrared	50–1000 μm	Mercury lamp (high-pressure)	Thermocouple; bolometer	Molecular rotations
Microwave	0.1–30 cm	Klystrons; magnetrons	Silicon-tungsten crystal; bolometer	Molecular rotations; electron spin resonance
Radio-frequency	10^{-1}–10^{3} m	Radio transmitter	Radio receiver	Molecular rotations; nuclear magnetic resonance

levels associated with the vibrations and rotations of molecules. Spacings between electronic levels are greater than between vibrational levels, and those between vibrational levels are greater than between rotational levels; each vibrational level has a set of rotational levels associated with it, and each electronic level a set of vibrational levels (**Fig. 1**). Transitions between vibrational levels of the same electronic level correspond to photons in the infrared region; transitions between rotational levels, to photons in the far-infrared and microwave region. Rotational spectra of gaseous molecules consist of sharp lines; vibrational spectra consist of bands, each of which arises from a given transition between vibrational levels altered in energy by increments due to changes in rotation occurring when the vibrational level changes. Likewise, molecular electronic spectra consist of bands due to transitions between electronic energy levels, altered by increments resulting from changes in vibration and rotation of the molecule on changing electronic state. *See* BAND SPECTRUM; INFRARED SPECTROSCOPY; LINE SPECTRUM; MICROWAVE SPECTROSCOPY; MOLECULAR STRUCTURE AND SPECTRA; SPECTRUM ANALYZER.

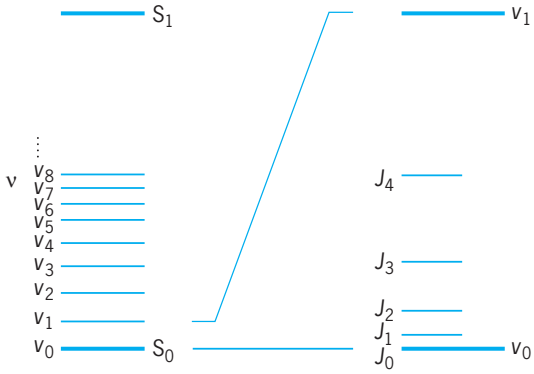

Fig. 1. Spacing of singlet electronic states *S*, vibrational states *v*, and rotational states *J*.

External fields. The application of a magnetic or electric field to the sample often separates normally indistinguishable, or degenerate, states in energy from each other. Thus, the orientation of the spin of an unpaired electron in an atom, ion, or molecule with respect to an applied magnetic field may have different values, and in the magnetic field the atom or molecule may have different energies. For typical field strengths, the difference in energy between these newly separated energy levels occurs in the microwave region. Similarly, for an atomic nucleus in an ion or molecule, differences in the orientation of the spin of the nucleus with respect to a magnetic field give rise to different energy levels of the nucleus in that field, so that energy differences will be found in the microwave region. The former phenomenon produces electron spin resonance or electron paramagnetic resonance spectra, and the latter produces nuclear magnetic resonance spectra (**Fig. 2**).

Nuclear magnetic resonance has been advanced by pulsed irradiation techniques which permit the identification of solid samples and complex biological molecules, and which are useful for the production of three-dimensional plots to permit analysis across the interior of objects without destruction. In medicine the imaging of internal organs is known as magnetic resonance imaging. Electron paramagnetic spectroscopy is used to establish structures of species containing unpaired electrons in solution or in a crystalline solid; some qualitative and quantitative analysis is also performed. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; MEDICAL IMAGING; NUCLEAR MAGNETIC RESONANCE (NMR).

Electronic spectra are also altered by external fields; removal of degeneracy by an externally applied electric field is termed the Stark effect, and removal of degeneracy by an externally applied magnetic field is termed the Zeeman effect. *See* STARK EFFECT; ZEEMAN EFFECT.

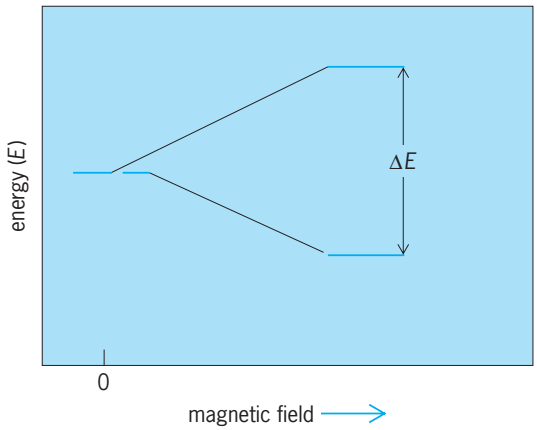

Fig. 2. Splitting degenerate energy levels of an appropriate atom in a magnetic field. Transitions between the separated levels require interaction with a photon of energy ΔE.

Spontaneous emission. The spontaneous emission of light from a sample by the decay of an electron from a higher level to the lowest available level is known either as fluorescence or as phosphorescence. The fundamental difference between these two terms is associated with the orientation of the spin of this electron. If the spin of the electron is such that it falls to the lowest state without a change of spin (the atom or molecule having the same spin before and after the transition), fluorescence occurs, and the process is characterized as allowed and is relatively fast. If the spin is such that the electron can fall to the lower state only with a change in spin dictated by the Pauli exclusion principle, the process is phosphorescence, characterized as a forbidden process, which is relatively slow (**Fig. 3**). In practice, other factors also govern the time the electron spends in the higher level, and the time intervals associated with each process overlap; that for fluorescence is 10^{-4} to 10^{-8} s, and that for phosphorescence is 10^{-4} to 10 s. The spontaneous emission process thus cannot be completely distinguished merely on the basis of the time interval associated with the emission. Time-resolved emission studies may allow separation of the processes. Fluorescence and phosphorescence are measured at a right angle to the incident beam exciting the sample, in order to avoid background problems. The emission of light beyond blackbody radiation when a sample is heated (thermoluminescence) or when it is strained or fractured (triboluminescence) has limited applicability. Similar emission during chemical reactions (chemiluminescence; in biochemical systems, bioluminescence) has broader analytical use if the reactions in which products emit light are, for example, catalyzed by a small amount of the substance to be analyzed. *See* BIOLUMINESCENCE; CHEMILUMINESCENCE; EXCLUSION PRINCIPLE; FLUORESCENCE; PHOSPHORESCENCE.

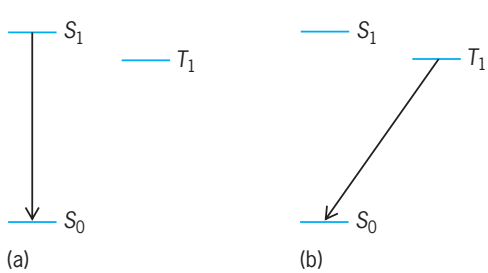

Fig. 3. Spontaneous emission of light by the decay of an electron from a higher to the lowest available level. (*a*) Fluorescence. (*b*) Phosphorescence.

Light rotation. Compounds whose molecules have a structure which cannot be superimposed on its reflection in a mirror (asymmetric molecules) rotate the plane of polarized light. If a device can be included in an instrument to plane-polarize the light from the source, that is, to produce light oscillating in only one plane, the amount of rotation of that plane by the sample can be studied as a function of the wavelength of light used in the experiment. A plot of this rotation versus wavelength is called an optical rotatory dispersion curve; such curves are useful in studying fundamentals of light rotation and in establishing the absolute structure of asymmetric molecules empirically. Asymmetric samples whose absorption of the components of plane-polarized light differs are said to exhibit circular dichroism; analysis of a sample in a magnetic field (magnetic circular dichroism) has a broader applicability in structural analysis. *See* COTTON EFFECT; OPTICAL ACTIVITY.

Instrumentation. Spectrometers require a source of radiation, a dispersion device, and a detector. For emission spectra, the source may be taken as the sample to be measured itself, although another source of radiation may be needed to excite the sample to an excited state. Absorption spectra require a source to generate appropriate radiation across the range of energies needed for the experiment. These radiant sources produce continuous spectra and can be distinguished from those producing discontinuous spectra.

Continuous spectra. Of the sources listed in the table, most of those from the vacuum ultraviolet region to the far-infrared region produce continuous spectra. Within this type, sources for the near ultraviolet, visible, and infrared regions consist of hot, incandescent solids; the maximum intensity of their emission occurs at a wavelength dependent on their temperature and emissivity. Electrical discharges in gases at higher pressures provide sources of continuous radiation for the vacuum ultraviolet and ultraviolet regions; hydrogen or deuterium gas discharges are especially used for the latter. X-ray continuous emission can be produced by collision of electrons or other particles accelerated through tens of kilovolts with a target. *See* BEAM-FOIL SPECTROSCOPY.

Discontinuous spectra. The sources of high-energy radiation which yield discontinuous spectra emit because of discrete processes in individual atoms, ions, or molecules. Such sources may include flames and furnaces to vaporize and atomize samples, and electrical discharges in gases at lower pressures. Other atomizing sources may be produced by high-frequency radio discharges and microwave discharges through

a flowing gas. For sources of radiation connected with nuclear processes, solid samples may be used containing the element to be studied, either a radioactive isotope or one which can be made radioactive by particle bombardment. *See* ARC DISCHARGE.

Among these so-called line sources are hollow-cathode lamps, in which a few layers of a cylindrical cathode containing the same element as is to be analyzed are vaporized by collisions with ions generated from rare gases and accelerated electrons. The gaseous atoms, also excited to higher energy states by this process, emit characteristic line spectra useful of analysis of that element. Lasers have become an especially useful source, and have opened different areas of spectroscopy, not only because of the intensity of the radiation supplied to the sample, which permits absorption of several photons in a single process, but also because of the coherent properties of the radiation, which permit the study of events occurring on a picosecond time scale. *See* LASER; ULTRAFAST MOLECULAR PROCESSES.

The sources for radio-frequency studies are tunable radio-frequency oscillators, designed with special regard for stability. The microwave sources may be spark discharges, magnetrons, or most commonly klystrons. An electron beam traveling through a closed cavity of defined geometry sets up electromagnetic oscillations, of which a certain value will be reinforced as a function of the dimension of the containing unit; the value can be varied somewhat by adjusting the size of the cavity. *See* KLYSTRON; MAGNETRON.

Dispersive elements. There are two kinds of dispersive elements, prisms and diffraction elements. The earlier and more commonly available prism has been supplanted in many cases by diffraction elements.

After collimation of the source radiation through a slit, diffraction is achieved in x-ray spectroscopy by the use of a pure crystal. The distances between atomic nuclei in the crystal are on the same order as the wavelength of the radiant energy; under these conditions an efficient dispersion of the radiation over an arc of many degrees is possible. For a distance d between repeating planes of atoms, the wavelength λ will be diffracted through an angle θ related by the Bragg law $n\lambda = 2d \sin \theta$, where n is an integer called the diffraction order. Materials used include gypsum ($CaSO_4 \cdot 2H_2O$), ammonium dihydrogen phosphate ($NH_4H_2PO_4$), and alkali halides (lighter elements). Each of these gives a wide angular dispersion in a portion of the 0.03–1.4-nm range of the spectrum commonly used for x-ray studies; the appropriate one must be chosen according to the experiment to be performed. *See* X-RAY OPTICS.

For spectroscopic techniques using visible light and the regions adjacent to it, ultraviolet and infrared, ruled diffraction gratings are often employed. Because their production has become economical, they have largely replaced prisms, which may require the use of several materials to get adequate angular dispersion and wavelength resolution over the entire range of interest, for example, 2–15 μm for an infrared spectrum. The dispersion of a prism increases near the limit of its transparency, and there is always a compromise in efficiency between transmission and dispersion. The resolution R, defined as the quotient of the wavelength of a line and the wavelength difference between it and another line just separated from it, is given by $T/(dn/d\lambda)$, where T is the thickness of the prism base and $dn/d\lambda$ is the variation of the refractive index with respect to wavelength. Efficient prisms are made of quartz for the ultraviolet, glass for the visible, and various salts for the infrared. *See* DIFFRACTION GRATING; OPTICAL MATERIALS; OPTICAL PRISM.

Very short- or very long-wavelength spectra are usually produced without a dispersive element; gamma-ray detection uses a wavelength-sensitive detector (a pulse height discriminator, for example), and microwave and radio-frequency detection may use a variable tuning radio receiver; in the latter case the source is tuned to emit a highly resolved frequency as well.

Detectors. Detectors commonly used for the various spectral regions are listed in the table. The devices for the infrared and microwave regions are basically heat sensors. Except for the eye and photographic methods used sometimes in the regions for gamma ray to visible, each detector converts the signal into an electrical response which is amplified before transmittal to a recording device. Direct imaging of the response by vidicon detection has also been used. Otherwise an oscilloscope or oscillograph is used to record very rapidly obtained signals. *See* GAMMA-RAY DETECTORS; OSCILLOSCOPE; TELEVISION CAMERA TUBE.

Instruments. Spectroscopic methods involve a number of instruments designed for specialized applications.

Spectroscope. An optical instrument consisting of a slit, collimator lens, prism or grating, and a telescope or objective lens which produces a spectrum for visual observation is called a spectroscope. The first complete spectroscope was assembled by Bunsen in 1859 (**Fig. 4**).

Spectrograph. If a spectroscope is provided with a photographic camera or other device for recording the spectrum, the instrument is known as a spectrograph. For recording ultraviolet and visible spectra, two types of quartz spectrographs are in common use. One type utilizes a Cornu quartz prism constructed of two 30° prisms, left- and right-handed, as

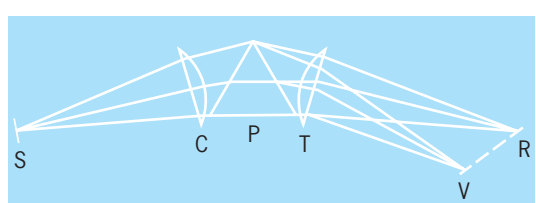

Fig. 4. Prism spectroscope or spectrograph. S, slit; C, collimator lens, P, 60° prism; T, telescope; RV, red to violet spectrum. For producing ultraviolet spectra, the glass items in C, P, and T must be replaced by equal amounts of left- and right-handed crystal quartz.

well as left- and right-handed lenses, so that the rotation occurring in one-half the optical path is exactly compensated by the reverse rotation in the other. The other type of quartz spectrograph employs a Littrow 30° quartz prism with a rear reflecting surface that reverses the path of the light through prism and lens, thus compensating for rotation of polarization in one direction by equal rotation in the opposite direction. Thus, in either type the effect of optical activity and birefringence in crystal quartz which produces double images of the slit is eliminated.

Grating spectrographs cover a much broader range of wavelengths (vacuum ultraviolet to far infrared) than do prism instruments. Various-type mountings are employed. The most common mounting for a plane-reflection grating is in a Littrow mount, entirely analogous to that of the Littrow quartz spectrograph. Mountings for concave reflection gratings require that the grating, slit, and camera plate all lie on the Rowland circle (imaginary circle of radius equal to one-half the radius of curvature of the grating) in order to achieve proper focus of spectral lines (slit images) on the plate. Paschen, Rowland, Eagle, and Wadsworth mountings are common in grating spectrographs.

Spectrometers. A spectroscope that is provided with a calibrated scale either for measurement of wavelength or for measurement of refractive indices of transparent prism materials is called a spectrometer. Also, the term frequently is used to refer to spectrographs which incorporate photoelectric photometers instead of photographic means to measure radiant intensities.

Spectrophotometer. A spectrophotometer consists basically of a radiant-energy source, monochromator, sample holder, and detector (**Fig. 5**). It is used for measurement of radiant flux as a function of wavelength and for measurement of absorption spectra. The wavelength is scanned by the monochromator at a constant rate (geared to chart speed of recorder) while the detector responds to, and discriminates between, the two alternating beams, one from the sample and the other from the attenuator. The servomechanism drives the attenuator, as well as the recorder pen, so that the intensities of the alternating signals are kept identical. The displacement of the pen on the chart can be calibrated in terms of relative intensity, percent transmittance, or absorbance. Generally a linear attenuator is employed, and the pen displacement is directly proportional to the attenuator displacement and thus the transmittance.

Interferometer. This optical device divides a beam of radiant energy into two or more parts which travel different paths and then recombine to form interference fringes. Since an optical path is the product of the geometric path and the refractive index, an interferometer measures differences of geometric path when two beams travel in the same medium, or the difference of refractive index when the geometric paths are equal. Interferometers are employed for high-resolution measurements and for precise determination of relative wavelengths: they are capable of

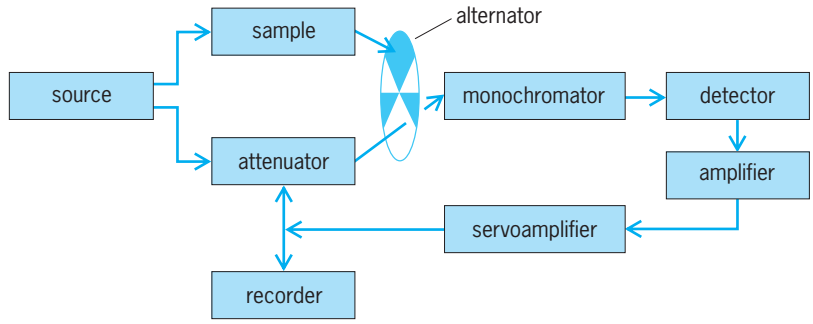

Fig. 5. Schematic diagram of a double beam (in time) spectrophotometer.

distinguishing between two spectral lines that differ by less than 10^{-6} of a wave. *See* INTERFERENCE OF WAVES; INTERFEROMETRY.

Quantitative relationships. For practical analysis, the absorption of radiant energy is related to the concentration of the sample by the relationship $-\log P/P_0 = abc$, where P and P_0 are respectively the power of the light attenuated after passing through the sample and its unattenuated power before reaching the sample, c is the concentration of the sample, b is the path length of the light through the sample, and a is a proportionality constant called the absorptivity of the sample. If c is in moles per liter, the absorptivity is called the molar absorptivity and symbolized ϵ. This relationship is known as the Beer-Lambert-Bouguer law, or simply Beer's law. For x-ray spectroscopy, the right-hand variables are grouped differently.

The practical relation between emission of light and concentration is given by the equation $F = k\phi I_0 abc$, where F is the fluorescence intensity, k is an instrumental parameter, ϕ is the fluorescence efficiency (quanta emitted per quantum absorbed), I_0 is the intensity of the incident light upon the sample, and the other units are as defined previously. The choice between absorption and fluorescence techniques for determining concentration is sometimes made on the basis of the accuracy of the linear method (fluorescence) versus the versatility of the logarithmic method (absorption). *See* LUMINESCENCE ANALYSIS.

Other methods and applications. Since the early methods of spectroscopy there has been a proliferation of techniques, often incorporating sophisticated technology.

Acoustic spectroscopy. When modulated radiant energy is absorbed by a sample, its internal energy increases. The loss of that excess produces a temperature increase that can be monitored as a periodic pressure change in the gas around the sample by using a microphone transducer. This is the optoacoustic effect. Its application provides a rapid method for study of some difficult samples. *See* PHOTOACOUSTIC SPECTROSCOPY.

Astronomical spectroscopy. The radiant energy emitted by celestial objects can be studied by combined spectroscopic and telescopic techniques to obtain information about their chemical composition,

temperature, pressure, density, magnetic fields, electric forces, and radial velocity. Radiation of wavelengths much shorter than 300 nm is absorbed by the Earth's atmosphere, and can only be studied by spectrographs transported into space by rockets. *See* ASTRONOMICAL SPECTROSCOPY; SPECTROHELIOGRAPH.

Atomic absorption and fluorescence spectroscopy. This branch of electronic spectroscopy uses line spectra from atomized samples to give quantitative analysis for selected elements at levels down to parts per million, on the average. However, detection limits vary greatly from element to element. Plasmas, especially inductively coupled plasmas, provide a stable, efficient, novel method for atomization and ionization.

Attenuated total reflectance spectroscopy. Spectra of substances in thin films or on surfaces can be obtained by the technique of attenuated total reflectance or by a closely related technique called frustrated multiple internal reflection. In either method the sample is brought into contact with a total-reflecting trapezoid prism and the radiant-energy beam is directed in such a manner that it penetrates only a few micrometers of the sample one or more times.

The technique is employed primarily in infrared spectroscopy for qualitative analysis of coatings and of opaque liquids.

Electron spectroscopy. This area includes a number of subdivisions, all of which are associated with electronic energy levels. The outermost or valence levels are studied in photoelectron spectroscopy, which uses photons of the far-ultraviolet region to remove electrons from molecules and to infer the energy levels of the remaining ion from the kinetic energy of the expelled electron. This technique is used mostly for fundamental studies of bonding. Electron impact spectroscopy uses low-energy electrons (0–100 eV) to yield information similar to that observed by visible and ultraviolet spectroscopy, but governed by different selection rules.

X-ray photoelectron spectroscopy (XPS), also called electron spectroscopy for chemical analysis (ESCA), uses x-ray photons to remove inner-shell electrons in a similar way, and the kinetic energy of the removed electron is measured. In Auger spectroscopy, an x-ray photon removes an inner electron. Another electron falls from a higher electron to take its place, and a third is ejected to conserve energy, since the second electron gains energy when it falls to the lower level. The kinetic energy of the third electron is measured. The chemical elements contained within a sample can then be identified, except for hydrogen. The energies of inner-shell electrons are dependent (in the third or fourth significant figure) upon the oxidation state of the atom, and so the kinetic energy of an XPS or Auger electron often provides information on the oxidation states of the elements in a compound too. Both techniques are used to study surfaces, since x-rays penetrate many objects for only a few atomic or molecular layers. The x-rays may be directed at the sample at various angles to the normal, and because they penetrate the sample for a more or less fixed distance, x-rays that are directed at the sample at a large angle from the normal (for example, 80°) encounter far fewer atomic or molecular layers than those that enter at smaller angles from the normal. This angle-resolved x-ray photoelectron spectroscopy permits more refinement in the analysis of the first few surface layers of a material. Ion neutralization spectroscopy uses protons or other charged particles instead of photons. *See* AUGER EFFECT; ELECTRON SPECTROSCOPY; SURFACE AND INTERFACIAL CHEMISTRY; SURFACE PHYSICS.

Fourier transform spectroscopy. This family of techniques consists of a fundamentally different method of irradiation of the sample, in which all pertinent wavelengths simultaneously irradiate it for a short period of time and the absorption spectrum is obtained by mathematical manipulation of the cyclical power pattern so obtained. It has been applied particularly to infrared spectrometry and nuclear magnetic resonance spectrometry, and allows the acquisition of spectra from smaller samples in less time, with high resolution and wavelength accuracy. The infrared technique is carried out with an interferometer of conventional instrumentation. *See* FOURIER SERIES AND TRANSFORMS.

Gamma-ray spectroscopy. Of special note in this category are the techniques of activation analysis, which are performed on a sample without destroying it by subjecting it to a beam of particles, often neutrons, which react with stable nuclei present to form new unstable nuclei. The emission of gamma rays for each element at different wavelengths with different half-lives as these radioactive nuclei decay is characteristic. *See* ACTIVATION ANALYSIS; NEUTRON SPECTROMETRY.

Mössbauer spectroscopy results from resonant absorption of photons of extremely high resolution emitted by radioactive nuclei in a crystalline lattice. The linewidth is extremely narrow because there is no energy loss for the nucleus recoiling in the lattice, and differences in the chemical environment of the emitting and absorbing nuclei of the same isotope may be detected by the slight differences in wavelength of the energy levels in the two samples. The difference is made up by use of the Doppler effect as the sample to be studied is moved at different velocities relative to the source, and the spectrum consists of a plot of gamma-ray intensity as a function of the relative velocity of the carriage of the sample. It yields information on the chemical environment of the sample nuclei and has found application in inorganic and physical chemistry, solid-state physics, and metallurgy. *See* DOPPLER EFFECT; MÖSSBAUER EFFECT.

Laser spectroscopy. Laser radiation is nearly monochromatic, of high intensity, and coherent. Some or all of these properties may be used to acquire new information about molecular structure. In multiphoton absorption the absorption of more than one photon whose energy is an exact fraction of the distance

between two energy levels allows study of transitions ordinarily forbidden when only a single photon is involved. Multiphoton ionization refers to the removal of an election after the absorption of several photons. Raman spectroscopy has also become more common because of the use of lasers as sources, especially with the development of resonance Raman spectroscopy.

Information on processes which occur on a picosecond time scale can be obtained by making use of the coherent properties of laser radiation, as in coherent anti-Stokes-Raman spectroscopy. Lasers may also be used to evaporate even refractory samples, permitting analysis of the atoms so obtained, and since the radiation may be focused, such a technique may be applied to one small area or a sample surface at a time, permitting profiling of the analysis across the sample surface. Laser fluorescence spectroscopy provides the lowest detection limits for many materials of interest in biochemistry and biotechnology, and intense work on development of rapid, sensitive methods proceeds. This will have application in the research that involves mapping of the human genome. *See* GENETIC MAPPING.

Ultrafast laser spectroscopy may be used to study some aspects of chemical reactions, such as transition states of elementary reactions and orientations in bimolecular reactions. *See* LASER SPECTROSCOPY.

Mass spectrometry. The source of a mass spectrometer produces ions, often from a gas, but also in some instruments from a liquid, a solid, or a material adsorbed on a surface. Ionization of a gas is most commonly performed by a current of electrons accelerated through about 70–80 V, but many other methods are used. The dispersive unit may consist of an electric field, or sequential electric and magnetic fields, or a number of parallel rods upon which a combination of dc and ac voltages are applied; these provide either temporal or spatial dispersion of ions according to their mass-to-charge ratio. *See* MASS SPECTROMETRY; SECONDARY ION MASS SPECTROMETRY (SIMS); TIME-OF-FLIGHT SPECTROMETERS.

Multiplex or frequency-modulated spectroscopy. The basis of this family of techniques is to encode or modulate each optical wavelength exiting the spectrometer output with an audio frequency that contains the optical wavelength information. Use of a wavelength analyzer then allows recovery of the original optical spectrum. The primary advantage of such a system is that it enables the detector to sense all optical wavelengths simultaneously, resulting in greatly decreased scan times and the possibility for signal averaging. Frequency modulation can be achieved, for example, by means of a special encoding mask as in Hadamard transform spectroscopy or by use of a Michelson interferometer as in Fourier transform spectroscopy.

Raman spectroscopy. Consider the passage of a beam of light, of a wavelength not corresponding to an absorption, through a sample. A small fraction of the light is scattered by the molecules, and so exits

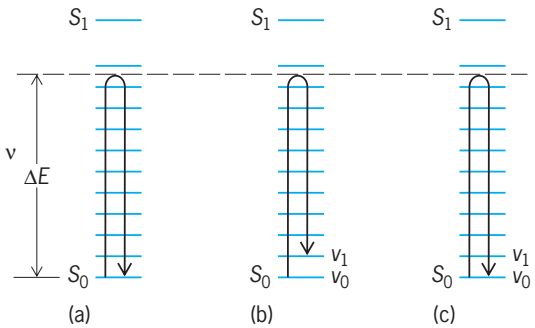

Fig. 6. Interactions of a molecule with a photon of energy ΔE which is not absorbed. (*a*) Rayleigh scattering. (*b*) Raman scattering, producing a Stokes line. (*c*) Raman scattering, producing an anti-Stokes line.

the sample at a different angle. This is known as Rayleigh scattering if the wavelength of the scattered light is the same as the original wavelength; but if the wavelength is different, it is called Raman scattering (**Fig. 6**). The fraction of photons thus scattered by the Raman effect is even smaller than by Rayleigh scattering. Differences in wavelength correspond to the wavelengths of certain vibrational and rotational processes. If the scattered light has a longer wavelength, the new line in the spectrum is known as a Stokes line; if shorter, an anti-Stokes line. The information obtained is of use in structural chemistry, because the rules allowing and forbidding transitions are different from those in absorption spectroscopy. Use of lasers for sources has revived this technique. A related process, resonance Raman spectroscopy, makes use of the fact that Raman probabilities are greatly increased when the exciting radiation has an energy which approaches the energy of an allowed electronic absorption. Raman scattering is also enhanced by surfaces, providing a valuable tool for surface analysis. *See* RAMAN EFFECT.

X-ray spectroscopy. The excitation of inner electrons in atoms is manifested as x-ray absorption; emission of a photon as an electron falls from a higher level into the vacancy thus created is x-ray fluorescence. The techniques are used for chemical analysis. Extended x-ray absorption fine structure, obtained under conditions of high resolution, shows changes in the energy of inner electrons due to the presence of neighboring atoms, and can be used in structural chemistry for the determination of the nature and location of neighbor atoms nearest to that being studied. Generation of x-rays by bombarding a sample with a proton beam gives the microanalytical technique of proton-induced x-ray emission. *See* X-RAY FLUORESCENCE ANALYSIS; X-RAY SPECTROMETRY.

Maurice M. Bursey

Bibliography. J. D. Ingle and S. R. Crouch, *Spectrochemical Analysis*, 1988; S. P. Parker (ed.), *Spectroscopy Source Book*, 1988; F. A. Settle et al., *Instrumental Methods of Analysis*, 8th ed., 2007; *Spectroscopy* (trade journal), monthly; D. H. Williams and I. Flemming, *Spectroscopic Methods in Organic Chemistry*, 5th ed., 1995.

Spectrum

The term spectrum is applied to any class of similar entities or properties strictly arrayed in order of increasing or decreasing magnitude. In general, a spectrum is a display or plot of intensity of radiation (particles, photons, or acoustic radiation) as a function of mass, momentum, wavelength, frequency, or some other related quantity. For example, a beta-particle spectrum represents the distribution in energy or momentum of negative electrons emitted spontaneously by certain radioactive nuclides, and when radionuclides emit alpha particles, they produce an alpha-particle spectrum of one or more characteristic energies. A mass spectrum is produced when charged particles (ionized atoms or molecules) are passed through a mass spectrograph in which electric and magnetic fields deflect the particles according to their charge-to-mass ratios. The distribution of sound-wave energy over a given range of frequencies is also called a spectrum. *See* MASS SPECTROSCOPE; SOUND.

In the domain of electromagnetic radiation, a spectrum is a series of radiant energies arranged in order of wavelength or of frequency. The entire range of frequencies is subdivided into wide intervals in which the waves have some common characteristic of generation or detection, such as the radio-frequency spectrum, infrared spectrum, visible spectrum, ultraviolet spectrum, and x-ray spectrum. Spectra are also classified according to their origin or mechanism of excitation, as emission, absorption, continuous, line, and band spectra. *See* ELECTROMAGNETIC RADIATION.

An emission spectrum is produced whenever the radiations from an excited light source are dispersed. Excitation of emission spectra may be by thermal energy, by impacting electrons and ions. or by absorption of photons. Depending upon the nature of the light source, an emission spectrum may be a continuous or a discontinuous spectrum, and in the latter case, it may show a line spectrum, a band spectrum, or both.

An absorption spectrum is produced against a background of continuous radiation by interposing matter that reduces the intensity of radiation at certain wavelengths or spectral regions. The energies removed from the continuous spectrum by the interposed absorbing medium are precisely those that would be emitted by the medium if properly excited. This reciprocity of absorption and emission is known as Kirchhoff's principle; it explains, for example, the absorption spectrum of the Sun, in which thousands of lines of gaseous elements appear dark against the continuous-spectrum background.

A continuous spectrum contains an unbroken sequence of waves or frequencies over a long range (see **illus.**).

In illus. *a* the spectra for 1000 and 2000°C (1832 and 3632°F) were obtained from tungsten filament, that for 4000°C (7232°F) from a positive pole of carbon arc. The upper spectrum in illus. *b* is that of the source alone, extending roughly from 400 to 650 nanometers. The others show the effect on this spectrum of interposing three kinds of colored glass.

All incandescent solids, liquids, and compressed gases emit continuous spectra, for example, an incandescent lamp filament or a hot furnace. In general, continuous spectra are produced by high temperatures, and under specified conditions the

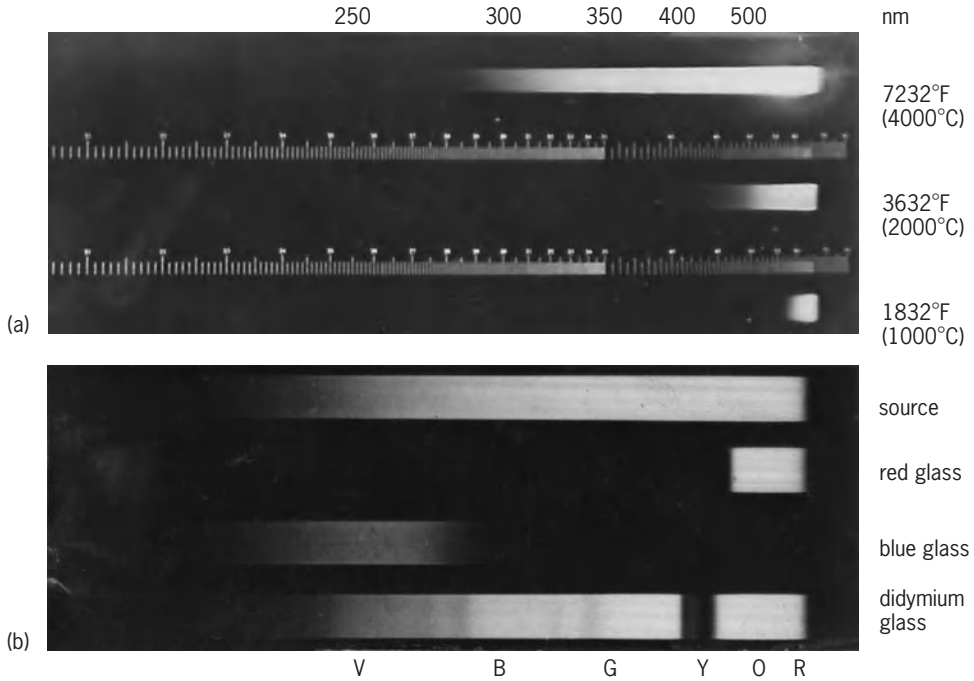

Spectra of a blackbody. (*a*) Continuous emission spectra of a solid, taken with a quartz spectrograph. (*b*) Continuous absorption spectra of the source alone and three kinds of colored glass interposed on it. (*From F. A. Jenkins and H. E. White, Fundamentals of Optics, 4th ed., McGraw-Hill, 1976*)

distribution of energy as a function of temperature and wavelength is expressed by Planck's law. *See* HEAT RADIATION; PLANCK'S RADIATION LAW.

Line spectra are discontinuous spectra characteristic of excited atoms and ions, whereas band spectra are characteristic of molecular gases or chemical compounds. *See* ATOMIC STRUCTURE AND SPECTRA; BAND SPECTRUM; LINE SPECTRUM; MOLECULAR STRUCTURE AND SPECTRA; SPECTROSCOPY.

W. F. Meggers; W. W. Watson

Spectrum analyzer

An instrument for the analysis and measurement of signals throughout the electromagnetic spectrum. Spectrum analyzers are available for subaudio, audio, and radio-frequency measurements, as well as for microwave and optical signal measurements.

Generally, a spectrum analyzer separates the signal into two components: amplitude (displayed vertically) and frequency (displayed horizontally). On some low-frequency analyzers, phase information can also be displayed. Low-frequency analyzers are sometimes grouped under the heading "harmonic analyzers," although this term is becoming less common.

Operation. On a conventional spectrum analyzer, a screen with a calibrated graticule displays the components of the input signal. The vertical scale displays the amplitude of each component, and the chosen frequency band is displayed horizontally. Components of the signal being analyzed are displayed as vertical lines whose height is proportional to amplitude and whose horizontal displacement equates to frequency. Originally, cathode-ray tubes were used for the display; solid-state displays such as liquid-crystal displays now are used. *See* CATHODE-RAY TUBE; ELECTRONIC DISPLAY.

Swept heterodyne technique. The traditional audio-frequency, radio-frequency, and microwave spectrum analyzer is essentially a swept heterodyne analyzer (**Fig. 1**). A ramp generator simultaneously increases the frequency of a local oscillator and moves the spot across the display. When the ramp voltage is zero, the spot will be on the left-hand side of the display and the local oscillator will be at a low frequency. Any signal at the input, at a frequency such that the difference between its frequency and the local oscillator is within the bandwidth of an intermediate-frequency filter, will be detected and will vertically deflect the spot on the display by an amount proportional to the amplitude of the input signal being analyzed. As the ramp voltage increases, the spot moves across the display and the analyzer tunes to increasingly higher-frequency signals. Each component of a complex waveform is thus isolated and displayed. *See* HETERODYNE PRINCIPLE.

Digital techniques. Traditionally, audio-frequency spectrum analyzers used a lower-frequency version of the swept heterodyne technique. Now, lower-frequency instruments generally use a digital

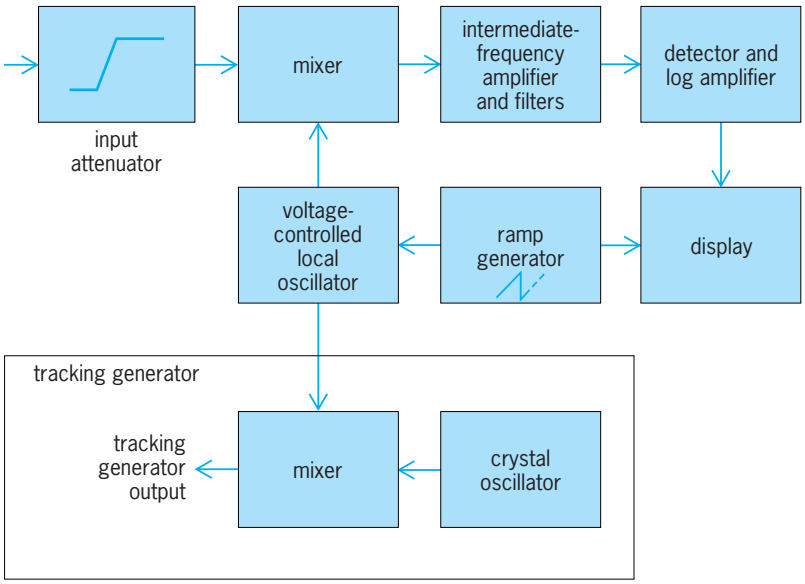

Fig. 1. Simplified block diagram of a spectrum analyzer with a tracking generator.

techniques known as the fast Fourier transform (FFT). The signal to be analyzed is converted to a digital signal by using an analog-to-digital converter, and the digital signal is processed by using the FFT algorithm. The input to this algorithm consists of blocks of digitized data. The algorithm analyzes the time-domain waveform, computes the frequency components present, and displays the results. All frequency components are analyzed simultaneously, in contrast to a swept instrument, which analyzes only a narrow portion of the spectrum at any one time. *See* ANALOG-TO-DIGITAL CONVERTER.

The FFT technique is much faster and the measurement is virtually real time, but it is restricted to lower frequencies. Real-time display is vital to analyze complex nonrepetitive signals such as speech or random vibrations. Digital processing techniques are used in radio-frequency and microwave instruments, but an analog input is required to down-convert the high-frequency signals to a signal low enough in frequency to be processed by currently available analog-to-digital converters. *See* FOURIER SERIES AND TRANSFORMS.

Controls. Modern spectrum analyzers incorporate sophisticated microprocessor control and can be remotely programmed (operated), so that a computer can control all functions and read the measured values. There are normally four basic groups of controls: frequency, horizontal scale, vertical scale, and markers. Both digital keypads and rotary controls are generally provided. The frequency controls determine the center or reference frequency of the display. The horizontal-scale controls determine the portion of the spectrum that can be viewed; this can be a scan from, say, 100 Hz up to 20 GHz or beyond. Separate controls determine the vertical scale and the sensitivity. Markers can be placed on spectral lines to measure accurately absolute and relative frequencies and levels. *See* MICROPROCESSOR.

Further controls allow the operator to optimize operation for different conditions; for example, the intermediate-frequency bandwidth can be changed to alter the resolution in order to resolve closely spaced signals. The input attenuation and gain of the intermediate-frequency section can be separately changed to optimize for either low-noise or low-intermodulation operation.

Digital storage. In order to obtain high resolution over a relatively wide band of frequencies, it is necessary to sweep slowly, especially in radio-frequency and microwave instruments. A single sweep may take as long as 100 seconds. Digital storage gives the effect of a constant display, even though a very slow sweep may have been used to acquire the displayed data.

Tracking generator. The tracking generator enhances the applications of spectrum analyzers (Fig. 1). Its output delivers a swept signal whose instantaneous frequency is always equal to the input tuned frequency of the analyzer. The swept signal is used to measure the frequency response of circuits such as filters and amplifiers over a wide dynamic range, which would not be possible with a broadband nonselective measurement method.

Harmonic analyzers. Spectrum analyzers covering up to typically 100 kHz can also be called harmonic analyzers. This term refers to instruments such as wave analyzers and distortion factor meters. Such instruments are still available, but the FFT spectrum analyzer has largely supplanted them, assisted with digital techniques using microprocessors. Faster and more accurate measurements are generally obtained using modern digital instrumentation.

Other harmonic analyzers that may be encountered include the tunable filter analyzer, stepped filter analyzer, parallel filter analyzer, and time compression analyzer.

Applications. Early radio-frequency and microwave analyzers were developed to measure the performance of microwave radar transmitters and to analyze signals from single-sideband transmitters. *See* RADAR; SINGLE SIDEBAND.

Fig. 2. Typical radio-frequency spectrum analyzer. (*IFR Ltd.*)

A typical use for radio-frequency and microwave spectrum analyzers is the measurement of spurious radiation (noise) from electrical machinery and circuits, known as radio-frequency interference (RFI). Other uses include monitoring and surveillance to detect unauthorized or unintended transmissions, such as the civil monitoring of broadcast and communication channels and the detection of electronic warfare signals. Another application is the analysis of radio communication transmitters and receivers, including those used in radio and television broadcasting, satellite systems, and mobile radio and cellular telephone communications (**Fig. 2**). *See* ELECTRICAL INTERFERENCE; ELECTRONIC WARFARE.

Spectrum analyzers also fulfill a variety of general-purpose uses in research, development, manufacture, and field maintenance of electrical and electronic equipment. For certain applications they are used instead of oscilloscopes, since the spectrum analysis display of a complex waveform can provide more information than an oscilloscope display. Oscilloscopes are typically limited to frequencies up to 1 GHz, while spectrum analyzers can measure microwave signals over 300 GHz. *See* OSCILLOSCOPE.

Low-frequency spectrum analyzers are used in a variety of applications. The most obvious use is the measurement of distortion and unwanted signals in all types of audio equipment, from recording and broadcast studios to amplifiers used in the home. *See* SOUND RECORDING; SOUND-REPRODUCING SYSTEMS.

Further uses include the analysis of speech waveforms, measurement of vibration and resonances in mechanical equipment and structures, determination of echo delays in seismic signals, investigation of noise such as from aircraft engines or from machinery in factories, analysis of sonar signals used to detect objects underwater, and study of ultrasonic waves to determine the internal structure of objects such as human tissue and metal castings. *See* BIOMEDICAL ULTRASONICS; MECHANICAL VIBRATION; NOISE MEASUREMENT; NONDESTRUCTIVE EVALUATION; SEISMOLOGY; SONAR; ULTRASONICS.

Optical spectrum analyzers use techniques such as a collimating mirror and a diffraction grating or a Michelson interferometer to separate out the light-wave components. They are used for a variety of applications, including measurements on lasers and light-emitting diodes, and for the analysis of optical-fiber equipment used to carry multichannel, digital telephony. *See* DIFFRACTION GRATING; INTERFEROMETRY; LASER; LIGHT-EMITTING DIODE; OPTICAL COMMUNICATIONS; OPTICAL FIBERS. S. J. Gledhill

Bibliography. J. S. Bendat and A. G. Piersol, *Engineering Applications of Correlation and Spectrum Analysis*, 2d ed., 1993; M. Engelson, *Modern Spectrum Analyzer Theory and Applications*, 1984; S. Goldman, *Vibration Spectrum Analysis: A Practical Approach*, 2d ed., 1999; A. D. Helfrick, *Electrical Spectrum and Network Analyzers*, 1991; B. Kibble et al., *A Guide to Measuring Direct and Alternating Current and Voltage Below 1 MHz*, Institute of Measurement and Control, 2003; T. S. Laverghetta,

Modern Microwave Measurements and Techniques, 1988; R. A. Witte, *Electronic Test Instruments: Analog and Digital Measurements*, 2d ed., 2002; R. A. Witte, *Spectrum and Network Measurements*, 2001.

Speech

A set of audible sounds produced by disturbing the air through the integrated movements of certain groups of anatomical structures. Humans attach symbolic values to these sounds for communication. There are many approaches to the study of speech.

Speech Production

The physiology of speech production may be described in terms of respiration, phonation, and articulation. These interacting processes are activated, coordinated, and monitored by acoustical and kinesthetic feedback through the nervous system.

Respiration. Most of the speech sounds of the major languages of the world are formed during exhalation. Consequently, during speech the period of exhalation is generally much longer than that of inhalation. In providing a sublaryngeal air supply, the respiratory muscles coordinate with the laryngeal and supralaryngeal muscles to produce suitable driving pressures for the formation of many speech sounds (**Fig. 1**). The aerodynamics of the breath stream influence the rate and mode of the vibration of the vocal folds. This involves interactions between the pressures initiated by thoracic movements and the position and tension of the vocal folds. *See* RESPIRATION.

The pressure pattern of the sublaryngeal air is closely related to the loudness of the voice and appears to be correlated with the perception of stress. For example, the word "permit" may be either a noun or a verb, depending on the placement of stress. Various attempts have been made to correlate units such as the syllable and the phrase to the contractions of specific respiratory muscles. However, the relation between the thoracic movements and the grouping of speech sounds is poorly understood.

Experimental studies on respiration for speech production employ techniques such as pressure and electromyographic recording at various points along the respiratory tract as well as x-ray photography to investigate anatomical movements.

Phonation. The phonatory and articulatory mechanisms of speech may be regarded as an acoustical system whose properties are comparable to those of a tube of varying cross-sectional dimensions. At the lower end of the tube, or the vocal tract, is the larynx. It is situated directly above the trachea and is composed of a group of cartilages, tissues, and muscles. The upper end of the vocal tract may terminate at the lips, at the nose, or both. The length of the vocal tract averages 6.5 in. (16 cm) in men and may be increased by either pursing the lips or lowering the larynx.

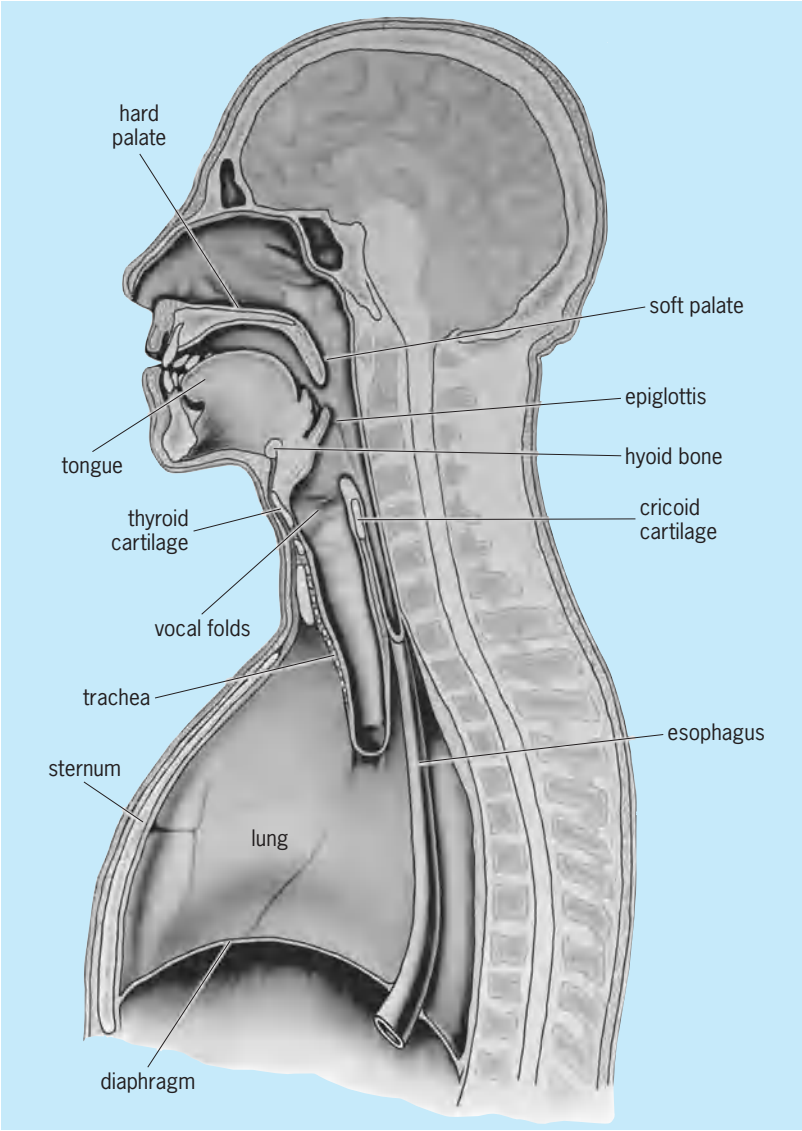

Fig. 1. Human head and thorax. (*Courtesy of G. E. Peterson*)

Mechanism of phonation. The larynx is the primary mechanism for phonation, that is, the generation of the glottal tone. The vocal folds consist of connective tissue and muscular fibers which attach anteriorly to the thyroid cartilage and posteriorly to the vocal processes of the arytenoid cartilages. The vibrating edge of the vocal folds measures about 0.92–1.08 in. (23–27 mm) in men and considerably less in women. The aperture between the vocal folds is known as the glottis (**Fig. 2**). The tension and position of the vocal folds are adjusted by the intrinsic laryngeal muscles, primarily through movement of the two arytenoid cartilages. By contraction of groups of muscles, the arytenoid cartilages may be pulled either anterior-posteriorly or laterally in opposite directions. They may also be rotated about the vertical axis by the cricoarytenoid muscles. During whispering and unvoiced sounds, the glottis assumes the shape of a triangle, with the apex directly behind the thyroid cartilage. *See* LARYNX.

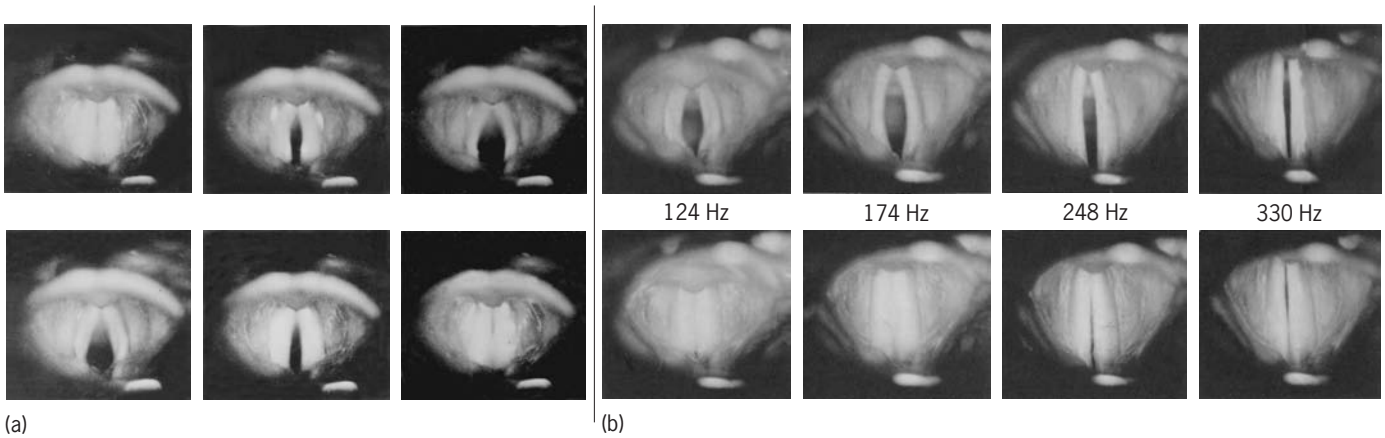

Fig. 2. Action of the glottis during speech. (*a*) Six successive positions of the vocal folds in one vibration cycle at 120 Hz. (*b*) Relation between pitch and vocal fold tension. (*Courtesy of D. W. Farnsworth*)

Air pressure. When the vocal folds are brought together and there is a balanced air pressure to drive them, they vibrate laterally in opposite directions. During phonation, the vocal folds do not transmit the major portion of the energy to the air. They control the energy by regulating the frequency and amount of air passing through the glottis. Their rate and mode of opening and closing are dependent upon the position and tension of the folds and the pressure and velocity of airflow. The tones are produced by the recurrent puffs of air passing through the glottis and striking into the supralaryngeal cavities.

As the air passes through the glottis with increasing velocity, the pressure perpendicular to the direction of airflow is reduced. This allows the tension of the folds to draw them together, either partly or completely, until sufficient pressure is built up to drive them apart again. It is generally assumed that the volume flow is linearly related to the area of the glottis and that the glottal source has a waveform that is approximately triangular.

Voiced sound. Speech sounds produced during phonation are called voiced. Almost all of the vowel sounds of the major languages and some of the consonants are voiced. In English, voiced consonants may be illustrated by the initial and final sounds in the following words: "bathe," "dog," "man," "jail." The speech sounds produced when the vocal folds are apart and are not vibrating are called unvoiced; examples are the consonants in the words "hat," "cap," "sash," "faith." During whispering all the sounds are unvoiced.

Frequency and pitch. The rate of vibration of the vocal folds is the fundamental frequency of the voice (F0). It correlates well with the perception of pitch. The frequency increases when the vocal folds are made taut (Fig. 2*b*). Relative differences in the fundamental frequency of the voice are utilized in all languages to signal some aspects of linguistic information. In **Fig. 3**, the fundamental frequency of the voice for the phrase "speech production" is shown as a dotted line. Note that frequency is highest on the vowels "ee" and "u," indicating that these are stressed. Also, the fundamental frequency drops to the lowest point at the end of the phrase, indicating the phrase is a statement rather than a question.

Many languages of the world are known as tone languages, because they use the fundamental frequency of the voice to distinguish between words. Chinese is a classic example of a tone language. There are four distinct tones in Chinese speech. Said with a falling fundamental frequency of the voice, *ma* means "to scold." Said with a rising fundamental frequency, it means "hemp." With a level fundamental frequency it means "mother," and with a dipping fundamental frequency it means "horse." In Chinese, changing a tone has the same kind of effect on the meaning of a word as changing a vowel or consonant in a language such as English.

The mode of vibration of the vocal folds can be varied to change the acoustical properties of the voice. For example, if the folds remain sufficiently apart during the vibration, the voice is perceived to be breathy in quality.

Study methods. The properties of the movements of the vocal folds have been studied by placing microphones at different points in relation to the larynx. Experiments have also been conducted with electrodes inserted directly into the laryngeal structures of hemilaryngectomized subjects and simultaneously recording their speech. Much information about the glottal vibration has been gathered through direct observation of the vocal folds by means of x-ray photography, the stroboscope, and high-speed motion-picture photography. The relation between the volume flow, the mode of glottal vibration, and the spectrum of the glottal tone has been studied by these techniques. Several electronic and mechanical models have been constructed to simulate phonation, based on the collated evidence from the acoustics and physiology of the larynx.

Articulation. The activity of the structures above and including the larynx in forming speech sound is known as articulation. It involves some muscles of the pharynx, palate, tongue, and face and of mastication (Fig. 1).

The primary types of speech sounds of the major languages may be classified as vowels, nasals,

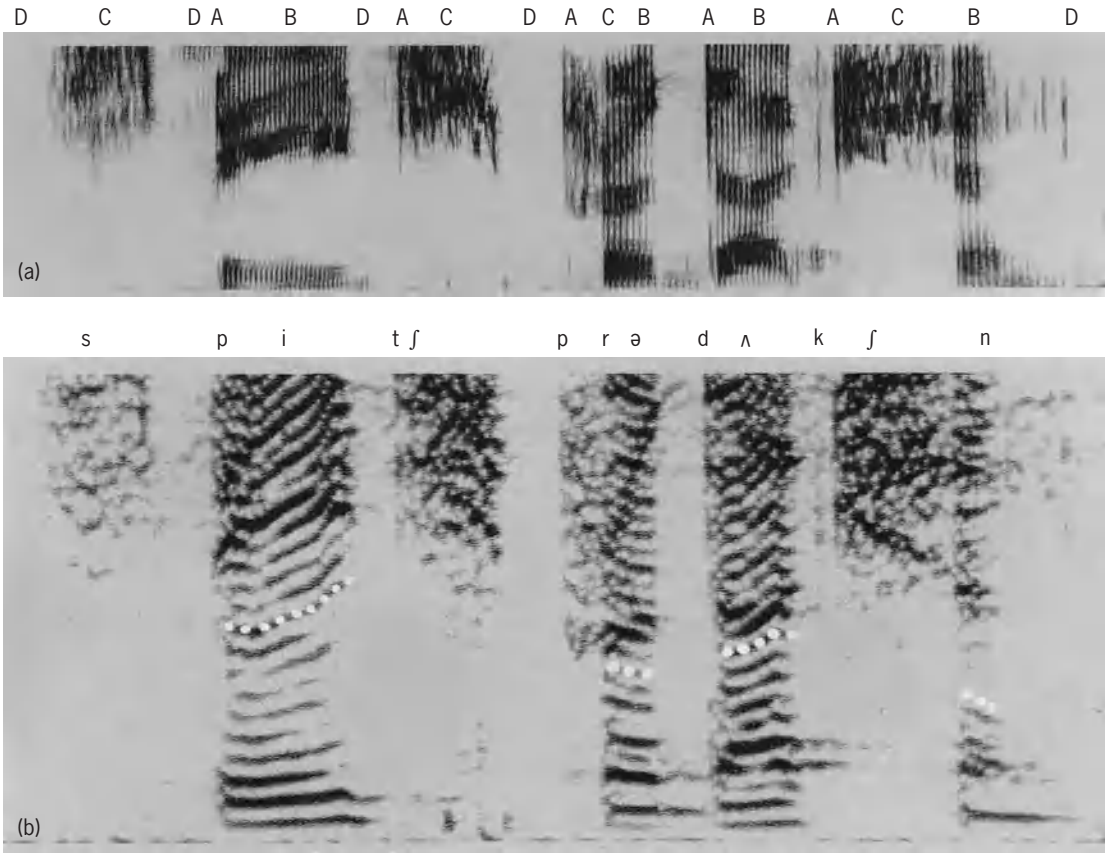

Fig. 3. Sound spectrograms of the utterance "speech production." (*a*) Energy patterns: A, pulse; B, periodic; C, aperiodic; and D, silence. (*b*) The dotted lines represent the pitch pattern of the utterance.

plosives, and fricatives. They may be described in terms of degree and place of constriction along the vocal tract. *See* PHONETICS.

Vowels. The only source of excitation for vowels is at the glottis. During vowel production the vocal tract is relatively open and the air flows over the center of the tongue, causing a minimum of turbulence. The phonetic value of the vowel is determined by the resonances of the vocal tract, which are in turn determined by the shape and position of the tongue and lips.

The point of constriction of a vowel is where the cross-sectional area of the vocal tract is minimized by the humping of the tongue. At this point the vocal tract is divided approximately into an oral cavity in front and a pharyngeal cavity behind. Vowels are known grossly as back, central, and front as the point of constriction moves from the pharyngeal wall anteriorly along the palate, as in the words "boot" (back), "but" (central), "beet" (front). Vowels are referred to as high, mid, and low as the height of the tongue hump is increasingly lowered, as in the words "beet" (high), "bet" (mid), "bat" (low). As the two parameters, point of constriction and tongue height, are varied, the relations between the oral and pharyngeal cavities change to produce characteristic resonances for different vowels.

Vowels are grossly described as rounded when produced with the contraction of the muscles of the lips, as in "he." Since rounding lengthens the vocal tract, the resonances of the rounded vowel are generally lower than those of the corresponding unrounded vowel.

Nasals. The nasal cavities can be coupled onto the resonance system of the vocal tract by lowering the velum and permitting airflow through the nose. Vowels produced with the addition of nasal resonances are known as nasalized vowels. Nasalization may be used to distinguish meanings of words made up of otherwise identical sounds, such as *bas* and *banc* in French. If the oral passage is completely constricted and air flows only through the nose, the resulting sounds are nasal consonants. The three nasal consonants in "meaning" are formed with the constriction successively at the lips, the hard palate, and the soft palate.

Plosives. These are characterized by the complete interception of airflow at one or more places along the vocal tract. The pressure which is built up behind the intercepting mechanism may not be immediately released or may be released through the oral or nasal orifice. The places of constriction and the manner of the release are the primary determinants of the phonetic properties of the plosives. The words "par," "bar," "tar," and "car" begin with plosives.

When the interception is brief and the constriction is not necessarily complete, the sound is classified as a flap. By tensing the articulatory mechanism in proper relation to the airflow, it is possible to set the mechanism into vibrations which quasi-periodically

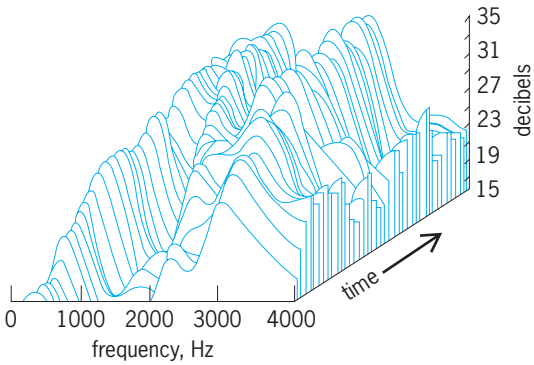

Fig. 4. Computer-generated three-dimensional spectral display of the spoken syllable "yah."

intercept the airflow. These sounds are called trills and can be executed with the velum, the tongue, and the lips. They usually are produced around 25 Hz, whereas the fundamental frequency of the male speaking voice ranges from 80 to 160 Hz.

Fricatives. These are produced by a partial constriction along the vocal tract which results in turbulence. Their properties are determined by the place or places of constriction and the shape of the modifying cavities. The fricatives in English may be illustrated by the initial and final consonants in the words "vase," "this," "faith," "hash."

Acoustical analysis. The muscular activities of speech production influence each other both simultaneously and with respect to time. It is frequently difficult to segment a sequence of sounds because the physiological activities which produce them form a continuum. Consequently the physical aspects of a sound type are determined to some extent by neighboring sounds. Each physiological parameter—for example, lip rounding and fundamental frequency of the glottal tone—may be varied continuously. Therefore, within its physiological limits the speech mechanism is able to produce a continuum of different sounds. The physiological and acoustical criteria in a classification of these sounds are therefore dependent on external conditions such as the threshold of the analyzing instruments.

There is a complex, many-to-one relation between the physiology of production and the resultant acoustical waves. Acoustically, speech sounds may be regarded as the simultaneous and sequential combinations of pulse, periodic, and aperiodic forms of energy interrupted by silence of varying duration. These energy patterns are labeled as A, B, C, and D, respectively, in the sound spectrogram of Fig. 3*a*. The horizontal bands in the periodic portions indicate the frequency positions of the vocal tract resonances. On the average, there is one resonance per kilohertz for vowels produced by a male vocal tract, though these resonances are differently spaced for various vowels. The pitch of the utterance is illustrated in the sound spectrograms of Fig. 3*b*. Each horizontal line indicates a harmonic of the glottal tone. The spectrum of the glottal tone is usually taken to have a slope of −6 to −12 decibels per octave when the effects of the vocal tract resonances are discounted. The sound spectrograph has been a valuable instrument for research on the acoustical aspect of speech. Essentially, it makes a Fourier type of analysis on the acoustical wave. The results are then translated into graphic form, usually with frequency along the ordinate and time along the abscissa.

Advances in computer technology have also been applied to the analysis of speech. **Figure 4** shows a three-dimensional display of the syllable "yah." In this display, the amplitude of each vocal tract resonance is indicated by the height of the peak. Thus it can be clearly seen that as the syllable progresses in time, the resonances change not only in frequency but in amplitude as well.

Novel methods have been introduced from signal-processing technology, such as the so-called speech flakes shown in **Fig. 5**. The flakes are of the vowels as in "far," "food," and "fee." These three vowels are by far those most often used in the languages of the world. The redundant symmetry in such flakes may turn out to be quite helpful for visualizing speech sounds.

Neurology. The ability to produce meaningful speech is dependent in part upon the association areas of the brain. It is through them that the stimuli which enter the brain are interrelated. These areas are connected to motor areas of the brain which send fibers to the motor nuclei of the cranial nerves and hence to the muscles. Three neural pathways are directly concerned with speech production, the pyramidal tract, the extrapyramidal, and the cerebellar motor paths. It is the combined control of these pathways upon nerves arising in the medulla and ending

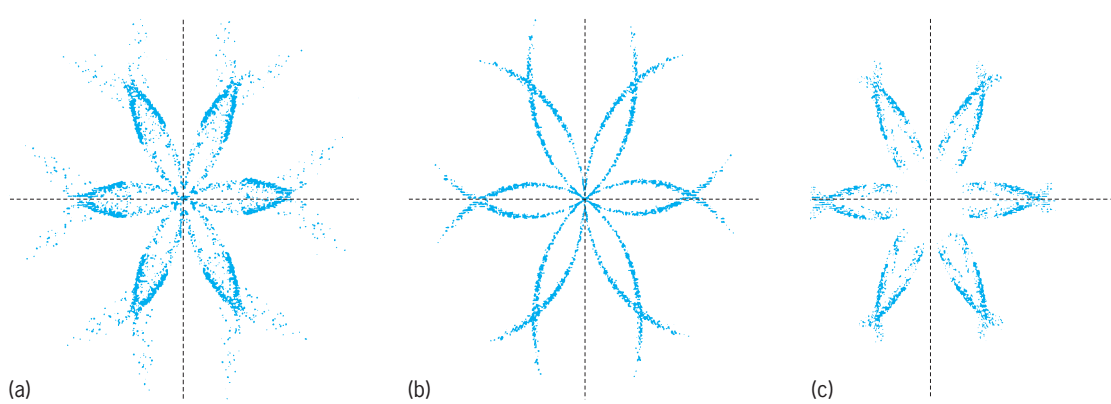

Fig. 5. Speech flakes of vowels in the words (*a*) "far," (*b*) "food," and (*c*) "fee." (*Courtesy of Charles Wooters*)

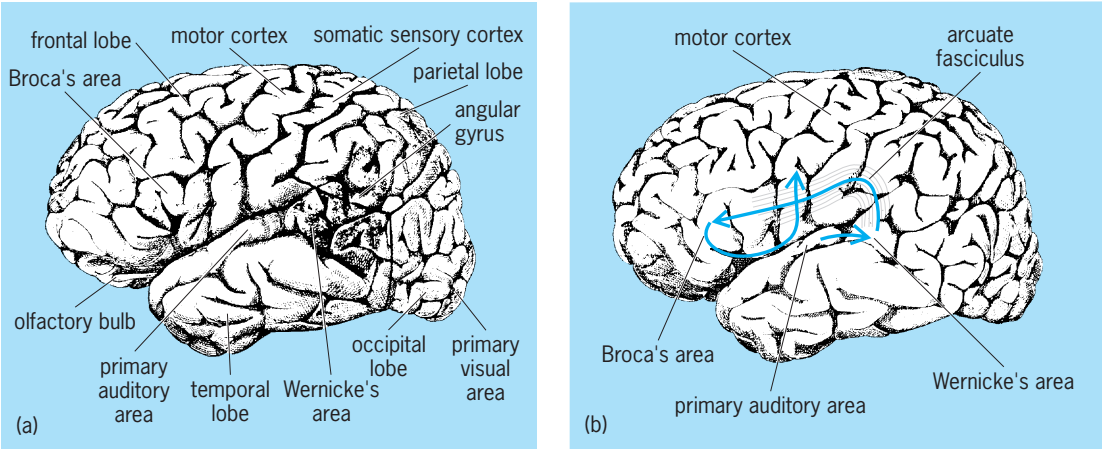

Fig. 6. Regions of the brain involved in speech. (*a*) Left hemisphere, usually more involved in speech functions than the right hemisphere. (*b*) A model for the neural processes involved in speaking a heard word. Arrows indicate the path of the relayed signal from the primary auditory area to the motor cortex.

in the muscles of the tongue, lips, and larynx which permits the production of speech. *See* NERVOUS SYSTEM (VERTEBRATE).

The part of the pyramidal tract which is most important for speech has its origin in the lower portion of the precentral gyrus of the cerebral cortex. In proximity to the precentral gyrus on the left side of the brain is Broca's area. This area is one of several which are believed to activate the fibers of the precentral gyrus concerned with movements necessary for speech production.

In most people, the left hemisphere of the brain is more involved in speech functions than the right hemisphere. **Figure 6***a* shows the regions on the surface of the left hemisphere that are especially pertinent for speech. Figure 6*b* shows a possible model for speaking a heard word. First the signal is received at the primary auditory area. From there the signal is relayed to Wernicke's area, situated posteriorly in the temporal lobe, where it is recognized as a word. The heard word is then relayed to Broca's area, situated inferiorly in the frontal lobe, via bundles of nerve fibers collectively known as arcuate fasciculus. Finally it is relayed to the motor cortex, which ultimately structures the muscular activities that produce the sounds, that is, speak the word.

Areas of the extrapyramidal tract which contribute to speech are the caudate nucleus, globus pallidus, and the thalamus. Some of the fibers from these centers also go to the same motor nerves as those of the pyramidal tract. The chief function of these paths in regard to speech is their regulatory and refining actions. It is through the influence of the extrapyramidal system that the regulation and inhibition of opposing sets of muscles are controlled. The cerebellar motor paths also contribute to the coordination and tonus of muscle structures requiring movements of paired muscles. In addition, the cerebellar motor paths are important in coordinating breathing for speech production.

Six of the 12 cranial nerves send motor fibers to the muscles that are involved in the production of speech.

These nerves are the trigeminal, facial, glossopharyngeal, vagus, spinal accessory, and the hypoglossal. They represent the link between the neural activity which begins in the cerebral cortex and the coordinated muscular movements which produce speech. The relations between the neurology, physiology, and acoustics of speech are extremely complex and poorly understood. Within normal anatomical variation, the skill in producing particular speech sounds is almost entirely determined by the early linguistic environment of the speaker. Through the interactions between the linguistic environment and the speech of the individual, the sounds and their symbolic values change from community to community and through time. *See* PSYCHOACOUSTICS; PSYCHOLINGUISTICS.

Development

In the early stages of speech development the child's vocalizations are quite random. The control and voluntary production of speech are dependent upon physical maturation and learning.

It is possible to describe the development of speech in five stages. In the first stage the child makes cries in response to stimuli. These responses are not voluntary but are part of the total bodily expression. The second stage begins between the sixth and seventh week. The child is now aware of the sounds he or she is making and appears to enjoy this activity. During the third stage the child begins to repeat sounds heard coming from himself or herself. This is the first time that the child begins to link speech production to hearing. During the ninth or tenth month the child enters the fourth stage and begins to imitate without comprehension the sounds that others make. The last stage begins between the twelfth and eighteenth month, with the child intentionally employing conventional sound patterns in a meaningful way. The exact time at which each stage may occur varies greatly from child to child.

Although most children begin to use speech meaningfully by the eighteenth month, they are not able to articulate all the necessary sounds. The first sounds the child uses are mostly front vowels. The back vowels become more frequent as the child develops. The

ability to produce vowels correctly seems to be almost completely learned by 30 months. The earliest consonants the child uses are those formed with both lips. By 54 months, the child can produce all the necessary consonants. The approximate age level at which each of 23 consonant sounds of English is mastered by American children follows.

Age, months	Sounds
$3\frac{1}{2}$	b(*b*aby), p(*p*a*p*a), m(*m*a*m*a), w(*w*et), h(*h*e)
$4\frac{1}{2}$	d(*d*a*d*a), t(*t*wo), n(*n*ose), g(*g*o), k(*c*at), ng(si*ng*), y(*y*et)
$5\frac{1}{2}$	f(*f*un)
$6\frac{1}{2}$	v(*v*ery), th(*th*at), z(a*z*ure), sh(*sh*oe), l(*l*ie)
$7\frac{1}{2}$	s(*s*ee), z(*z*oo), r(*r*ain), th(*th*in), wh(*wh*o)

The close coupling of age and sound acquisition, as exhibited above, together with other observations, has led some researchers to post innate, genetic mechanisms which serve exclusively speech. This innate hypothesis is a focal point for discussion on how speech emerged, in the child as well as in the human species.

Speech Technology

Since about 1990, technologies have arisen that are taking over an ever-expanding share of the marketplace. These technologies are based on the knowledge that has accumulated from many decades of research on how speech is produced and perceived. Speech technology has been developing within three areas.

One area has to do with identifying a speaker by analyzing a speech sample. Since the idea is analogous to that of identifying an individual by fingerprint analysis, the technique has been called voice print. However, fingerprints have two important advantages over voice prints: (1) they are based on extensive data that have accumulated over several decades of use internationally, whereas no comparable reference exists for voice prints; and (2) it is much easier to alter the characteristics of speech than of finger prints. Consequently, this area has remained largely dormant. Most courts in the United States, for instance, do not admit voice prints as legal evidence. Some progress may be made in voice print identification after a better understanding of the process of phonation is achieved, since this is more specific to individuals.

In contrast, the two other areas of speech technology, synthesis and recognition, have seen explosive growth. Humans prefer to interact with speech, which they have mastered since early childhood, rather than via keyboard and screen. This preference becomes a necessity when visibility is poor, or when the hands are otherwise occupied. With the rapid pace in computer advancement, speech technology is expected to extend into more and more sectors of human life in the years to come.

In many applications where a limited repertoire of speech is required, computer-synthesized speech is used instead of human speakers. These applications began with automobile devices which remind the driver to fasten his seat belt, and telephone directory services which tell customers the numbers they seek. The area has advanced greatly. Coupled with a machine that can read the printed page, the conversion of text to speech is almost a reality. A common technology currently used in speech synthesis involves an inventory of pitch-synchronized, prestored human speech. These prestored patterns are selected according to the particular requirements of the application and recombined with some overlap into the desired sentence by computer, almost in real time.

The quality of synthesized speech for English is remarkably good, though it is limited at present to neutral, emotionless speech. Many other languages are being synthesized with varying degrees of success. Much research will be required to endow synthesized speech with the rich panorama of emotion that human speech can convey.

The recognition of speech by computer is much more difficult than synthesis. Instead of just reproducing the acoustic wave, the computer must understand something of the semantic message that the speech wave contains, in order to recognize pieces of the wave as words in the language. Humans do this easily because they have a great deal of background knowledge about the world, because they are helped by contextual clues not in the speech wave, and because they are extensively trained in the use of speech. Although there are various systems in the marketplace at present which purport to perform dictation, their levels of performance are typically quite low. This is especially true when the speech takes place in noisy conditions.

Nonetheless, given various constraints, some of the existing systems do remarkably well—even to the extent of helping surgeons in the operating room, where reliability is of paramount concern. These constraints include (1) stable acoustic conditions in which speech is produced, (2) a speaker trained by the system, (3) limited inventory of utterances, and (4) short utterances. The research here is strongly driven by the marketplace, since all sorts of applications can be imagined where spoken commands are required or highly useful. It is expected that speech technology will continue to grow at a fast pace in the coming decades. *See* SPEECH DISORDERS.

William S.-Y. Wang

Bibliography. G. Fant, *Speech Sounds and Features*, 1973; O. Fujimura (ed.), *Vocal Fold Physiology*, vol. 2 of *Vocal Physiology: Voice Production, Mechanisms and Functions*, 1988; J.-C. Jungva and J.-P. Haton, *Robustness in Automatic Speech Recognition: Fundamentals and Applications*, 1996; P. Lieberman, *Speech Physiology and Acoustic Phonetics*, 1977; W. S.-Y. Wang, *The Basis of Speech*, 1989; W. S.-Y. Wang (ed.), *The Emergence of Language: Development and Evolution*, 1991.

Speech disorders

Typically, problems of speech production, articulation, or voice, or disruptions of the ability to use language. These disorders can be categorized in a variety of ways, but most belong to one of the following types: articulation, voice, fluency, and language.

Another way to categorize these disorders is on the basis of the cause, organic or nonorganic. Organic speech and language disorders typically result from disease or injury to the structures, muscles, and nerves necessary to move and control the mechanisms for speech and voice production. Damage to the brain, particularly the left hemisphere, as the result of stroke, dementia, or traumatic head injury can result in disruptions of normal language function. In children, failure of the brain or systems subserving language and speech to develop adequately can result in significant developmental delays in the acquisition of appropriate language and speech functions. Disorders of language and speech for which there is no demonstrable organic etiology are often the result of poor learning or inadequate speech and language models. In some instances, there may be a psychogenic basis for the disorder.

Articulation. Articulation for speech production involves the appropriate sequential movement of structures known as the articulators: mandible, tongue, lips, velum, and so forth. There is a wide range of acceptable productions of a given sound (phoneme) in the language. The variations, sometimes known as dialect, reflect regional, ethnic, and socioeconomic characteristics; they should not be considered as deviant productions. In most of these instances, production is precise and falls within the accepted bounds of the individual's speech and language community. However, when the production of one or more sounds in the language is imprecise and cannot be understood, then an articulation disorder is present. Disorders can range from mild problems such as the distortions associated with a lisp, or can be so severe as to render the speaker unintelligible. Most articulation problems can be categorized in terms of how the sounds are produced. Speech sounds can be misarticulated by distortion, omission, substitution, or addition. These problems can arise from organic causes, such as impairment to motor control of the speech musculature following a stroke or cerebral palsy (dysarthria), cleft palate or lip, or hearing loss. Many articulation problems have no discernible organic basis; these are labeled as functional, since the etiology of the problem is unclear. Many of these problems begin in childhood and can persist into adult life if not treated appropriately. *See* CLEFT LIP AND CLEFT PALATE; HEARING IMPAIRMENT; PHONETICS.

There are a number of approaches for treating these disorders, both organic and functional. For example, discrimination training is based on the concept that the basis of the problem is poor sound discrimination and the inability to match the auditory feedback from production with the patterns that others are able to produce. Treatment focuses on training auditory discrimination to discover the differences and then modifying production to match the pattern. Other approaches include behavior modification; production training, which requires relearning specific motor patterns; linguistic approaches based on theories of phonological structure; and psychological approaches, including the psychoanalytic model. The psychological approach has not been widely accepted; however, psychological counseling can sometimes be of great help to children who have not been accepted because of their poor articulatory behaviors. Finally, in cases of very severely impaired persons, such as those with cerebral palsy and accompanying dysarthria, communication boards and alternative communicative methods, including computers using synthetic speech, have become increasingly common. In persons with orofacial abnormalities, such as cleft lip or palate, that interfere with articulation, surgical and dental repair or protheses may be necessary before speech therapy can be initiated. *See* PSYCHOLINGUISTICS.

Production of speech sounds is a complex process that requires precise programming and direction by the brain. When there is damage to those brain regions controlling articulatory programming because of stroke or trauma, a disorder known as apraxia may ensue. This is in contrast to dysarthria, where articulation is distorted because of paralysis or paresis of the musculature necessary for speech production. *See* APRAXIA; BRAIN.

Voice. Voice disorders occur when the phonatory mechanism is not functioning properly. Some impairments of function can be the result of vocal abuse or misuse. These voice disorders are said to be functional, although prolonged abuse or misuse can result in temporary or permanent damage to the phonatory mechanism. There are, however, voice disorders that have a true organic etiology. These disorders result from disease or some other condition that affects the laryngeal, oral, or respiratory structures, which are involved in the production of voice. Phonation problems associated with cleft palate, which is a structural defect in the palate resulting in hypernasality, are not included in functional or organic disorders. Aphonia is a complete absence of voice. More commonly, individuals with voice disorder are dysphonic, retaining some voice quality. Whether functional or organic in origin, voice disorders are described in terms of pitch, quality (resonance), or loudness. Instruments have become available that permit sensitive and objective measurement of vocal characteristics. These instruments are of value in both diagnosis and treatment of voice disorders.

Pitch. Disorders of voice in which pitch is the predominant factor occur when the speaker's voice pitch is inappropriate for age and gender. In these instances, the pitch is described as too high, too low, monotonous, or demonstrating repeated inappropriate pitch patterns. A voice that is pitched too high is often the result of psychological tension. The

muscles of the larynx are contracted in such a way as to raise pitch significantly beyond its normal range. It may also result when the larynx is small for age and gender and the vocal folds are thin and short. In some instances, the speaker's inability to perceive pitch and pitch patterns may result in a high-pitched voice, monotonous voice, or inappropriate use of repeated pitch patterns. *See* PITCH.

Quality. Disorders of voice quality (resonance) can be divided into two major groups. In one group the problem is at the point of origin of the voice, the vocal folds and larynx; in the other, at the supralaryngeal resonating cavities. An example of the first type is functional dysphonia, related to emotional stress, in which breathy voice indicates vocal cord vibration without a closed phase. Vocal abuse may also cause vocal fold nodules, which do not permit adequate closure of the folds, yielding a breathy and harsh vocal quality. In contrast, when the vocal cords are squeezed too tightly, they cannot vibrate properly, and the voice sounds hoarse and low-pitched. Hoarseness associated with excessive vocal cord closure is known as hyperfunctional dysphonia. Hoarseness or breathiness can also result from organic disease affecting the vocal cords, such as paralysis of one vocal cord, trauma, myasthenia, or cancer of the larynx.

Disorders of quality related to the resonating cavities are typically of two types, hypo- or hypernasal. Hyponasal voice results when there is inadequate nasal resonance. This may be due to a physiologic or anatomic blockage of the vibrating air that ordinarily passes through the nasal cavities. In this case, the vibrating airstream does not reach the nasal passages. Hypernasality is the opposite condition, where too much air comes through the nasal cavities during phonation. This is a common problem associated with cleft palate. Speakers who are imprecise in their articulation and speak rapidly also demonstrate a mild form of this problem. Another type of hypernasality occurs when there is obstruction in the anterior nasal cavities (pinching the nostrils), which produces a quality similar to that of a nasal twang. In this case, the vibrating airstream reaches the nasal cavities but is unable to pass through them unimpeded.

Loudness. Disorders of loudness relate to the intensity of voice production. Here the voice is either not loud enough to be audible or too loud. The latter problem may occur in persons who work in very noisy environments and must shout in order to be heard. These persons carry this loud production over to the nonwork environment. Persons whose voice intensity is significantly below that which can be heard are typically those individuals whose feel inadequate and insecure. Loudness changes can often by symptomatic of hearing loss. When there is a conductive hearing impairment, the voice is often not loud enough, while in the case of sensoryneural impairment the voice is too loud. *See* LOUDNESS.

Alaryngeal speech. One other disorder of vocal production that should be considered is alaryngeal or esophageal speech. This results following surgical removal of the larynx, usually because of cancer. The individual who has undergone such surgical removal can, by means of esophageal speech or an artificial larnyx, produce adequate but limited voice. *See* LARYNX.

Treatment. Treatment of the voice disorders begins with diagnosis. It is essential that any organic basis or the disorder be determined and treated before beginning voice therapy with a certified speech-language pathologist. Beyond medical treatment, therapy may include modification of the environment, psychological counseling, and direct therapy. The last will usually include the development of listening skills, relaxation exercises, and specific vocal exercises to eliminate vocal abuse and determine the optimal pitch and resonance. Listening exercises are particularly useful in treating disorders where the voice is too loud.

Fluency. Among the most difficult disorders of speech to treat are those of fluency, which are characterized by a disruption of the normal rate of speech production. There are three major types of fluency disorders: stuttering, cluttering, and those related to motor speech disorders of neurological origin (dysarthria and apraxia). Of these, stuttering is the most commonly observed.

Stuttering. No one definition of stuttering encompasses the views of scholars devoted to its study. However, the following features appear in most definitions: a disruption or blockage of the normal flow of speech because of repetitions, prolongations of sounds or syllables, and periods of silence. All of these features are involuntary and appear to be associated with effortful struggles to avoid the behavior.

There is no universal agreement as to the cause of the disorder. Three prominent theories attempt to explain the disorder. One assumes that it is a learned behavior, another that it represents a psychological disorder, and the third that it has an organic basis in disordered neurologic function.

There is no cure for stuttering. However, many persons with this disorder can be treated so that the degree of dysfluency and the associated distracting secondary behaviors (for example, facial twitches or body movement) can be significantly reduced. No single accepted mode for treatment of stuttering is generally accepted. Almost everything from surgery of the tongue to psychoanalysis and hypnotherapy has been used to try to cure the disorder. Interestingly, of the children who reportedly stutter prior to puberty, only 20% continue to do so after the onset of puberty. This is the case even in the absence of therapy.

The approach to treatment is different for children and adults; however, most therapeutic approaches fall into one of three types: psychological approaches, modification of the stuttering, and modification of speech. In the first, the emphasis is on counseling and psychotherapy to alter the stutterer's self-image and attitude toward the dysfluency and to reduce avoidance behaviors. In treatment

directed at modifying the stuttering itself, the focus is on enabling the individual to stutter more efficiently and fluently. When treatment is directed at modifying speech, the goal is to achieve "stutter-free" production by altering the rhythm, rate, and voice. This approach draws heavily on the principles of behavior modification. Most speech pathologists use a combination of these general approaches when working with the dysfluent individual.

Cluttering. This is distinguished from stuttering by the fact that it is produced without awareness or concern on the part of the speaker. Cluttering is characterized by very rapid production accompanied by alterations in articulatory patterns, particularly involving substitutions and omissions. This combination, especially in the case of children, frequently renders speech unintelligible. It is often the classroom teacher who is the first to detect the problem. In adults who have been unrecognized clutterers, it may be an employer, friend, or colleague who recognizes the problem. Treatment focuses on recognition of the cluttering by the speaker, followed by specific techniques to alter rate and articulatory patterns.

Other disorders. Speech rate and rhythm can be significantly impeded by motor speech disorders. Diseases of the nervous system such as multiple sclerosis or parkinsonism and conditions such as cerebral palsy interfere with the normal sensory-motor control over the articulatory and rhythmic production of speech. Because the etiology of this group of fluency disorders varies so much, there are no specific patterns that can be described. Treatment therefore focuses on having the individual attempt to gain as much control over production as possible, given the limits imposed by the cause. In some cases, such as Parkinson's disease, the use of drugs such as L-dopa can have a marked effect on speech production. Muscle relaxants are sometimes used along with speech therapy for patients with multiple sclerosis. However, in cases of progressive neurologic disease, treatment can provide only temporary respite from the disorder. *See* MULTIPLE SCLEROSIS; PARKINSON'S DISEASE.

Language. Language disorders involve impairment of the ability to voluntarily use or comprehend the language of the individual's linguistic community. Such disorders need not involve a concomitant impairment of speech or voice production. These disorders can be divided into two groups. In delayed language acquisition, normal language development fails to occur because of neurologic or intellectual problems. Aphasia typically occurs in the adult following a stroke or traumatic injury to the brain after the individual had acquired and was using language appropriate to age, gender, and linguistic community. In addition, there has been increasing recognition of language disorders associated with the dementias. The language disorders associated with delayed acquisition in children and those seen in individuals who had previously acquired adequate language share few similarities, and the approaches to treatment are dissimilar.

Aphasia. The aphasias can be categorized in a variety of ways. Nonfluent aphasics show good comprehension but have difficulty with confrontation naming and produce cryptic "telegraphic" speech. Fluent aphasics show great language output, often meaningless, with very poor comprehension. However, there are some aphasics whose only apparent difficulty is in the use of names, and while they have reactively good comprehension, their expressive language depends on circumlocutions and nonspecific words. Finally, there are global aphasics who are nonfluent, with poor comprehension and repetition, and often able only to utter a meaningless syllable.

Because the event that produces the aphasia usually involves more than one region of the brain, the aphasic often exhibits other problems such as hemiplegia (paralysis of one side of the body); motor speech problems (dysarthria); visual problems such as blindness of one visual field (hemianopsia); or apraxia (difficulty in programming, planning, initiating, and sequencing motor behaviors, especially those required for articulation). In addition, because the onset of the neurologic event producing the aphasia is frequently sudden, the aphasic is often frightened and devastated. The impact of the disorder is felt not only by the patient but also by family members who will have to play a significant role in rehabilitation.

Some spontaneous reduction of the aphasia often occurs in the first few days and weeks following the onset of the disorder. The amount of improvement is governed by the extent of the brain damage, its locus, and the patient's overall condition and motivation. Most of these changes take place during the first 3 months following onset, with some continuation at a slower pace through the sixth month. However, language therapy during this period will enhance the spontaneous recovery process. Continued improvement, with the goal of having the aphasic reach the best level of communicative function possible, requires language therapy by a qualified speech–language pathologist. The first step in this process is a detailed evaluation of the aphasic's current language status. From this evaluation a treatment plan can be developed. Both individual and group treatment approaches are helpful. There are no uniform treatment prescriptions for aphasia. Techniques such as programmed learning, behavior modification, and cognitive and language stimulation have all been used with varying degrees of success. Some have recommended the use of Amerind (American Indian sign language) for severely impaired aphasics. Another approach attempts through a series of specific steps to achieve improved communication by a program known as Promoting Aphasic Communicative Effectiveness (PACE). There have also been efforts to incorporate personal computers into aphasia rehabilitation. Most of these techniques, as well as others, do function to reduce the communicative impairment associated with aphasia. However, it is unusual for aphasics to make a complete recovery, and most must learn to function with limited language abilities. *See* APHASIA.

Childhood disorders. Language disorders of childhood are often associated with other disorders, such as mental retardation, hearing loss, and autism. However, there are some children who show specific language impairments. They are considered to be normal in other aspects of development, in particular, hearing and nonverbal intelligence. Terms uesd to categorize the childhood disorders include language delay, developmental aphasia, or language impairment. In the mentally retarded child, most of the linguistic features of the language are slow to emerge. However, when they do appear, they are similar to those of the normal child. The more abstract the features, the greater the delay in acquisition. Some features may never emerge. This is dependent in part on the severity of the retardation. Autistic children, on the other hand, suffer from marked failure to acquire and use language in the ordinary sense. Some use echolalia, unsolicited repetitive or imitative verbal patterns echoing a previously heard utterance. Since language is dependent upon hearing, children with hearing loss or deafness will show mild to profound delays in language acquisition, with the degree of impairment relative to the degree of hearing loss. *See* AUTISM; MENTAL RETARDATION.

Therapy for children with language delay often focuses on production of specific linguistic structures. This may involve modeling the structure for the child. Some speech–language pathologists advocate a behavioral–learning approach using aspects of behavior modification. Others focus on an interpersonal–interactive methodology. Because of the variety of problems presented by these children, treatment must be tailored to their specific needs and environments.

Dementia. Persons with dementia, such as Alzheimer's type or multi-infarct dementia, often show a disturbance of language. These disturbances can take a variety of forms, such as the inability to recall and use names or significantly impaired comprehension and language output. Linguistic deficits associated with dementia, except in rare instances, are related to the deterioration of intellect and memory rather than a primary deficit in language function. Treatment for these deficits is often indirect, working to create a language environment that is carefully structured to the present level of linguistic and cognitive function. Often therapy is directed at helping the family and care givers appreciate the problem and cope with it. *See* ALZHEIMER'S DISEASE; SPEECH.

R. S. Tikofsky

Bibliography. D. R. Boone, *Human Communication and Its Disorders*, 1987; G. H. Shames, *Human Communication Disorder*, 5th ed., 1998.

Speech perception

A term broadly used to refer to how an individual understands what others are saying. More narrowly, speech perception is viewed as the way a listener can interpret the sound that a speaker produces as a sequence of discrete linguistic categories such as phonemes, syllables, or words. *See* PHONETICS; PSYCHOLINGUISTICS.

Classical work in the 1950s and 1960s concentrated on uncovering the basic acoustic cues that listeners use to hear the different consonants and vowels of a language. It revealed a surprisingly complex relationship between sound and percept. The same physical sound (such as a noise burst at a particular frequency) can be heard as different speech categories depending on its context (as "k" before "ah," but as "p" before "ee" or "oo"), and the same category can be cued by different sounds in different contexts. Spoken language is thus quite unlike typed or written language, where there is a relatively invariant relationship between the physical stimulus and the perceived category.

The reasons for the complex relationship lie in the way that speech is produced: the sound produced by the mouth is influenced by a number of continuously moving and largely independent articulators. This complex relationship has caused great difficulties in programming computers to recognize speech, and it raises a paradox. Computers readily recognize the printed word but have great difficulty recognizing speech. Human listeners, on the other hand, find speech naturally easy to understand but have to be taught to read (often with difficulty). It is possible that humans are genetically predisposed to acquire the ability to understand speech, using special perceptual mechanisms usually located in the left cerebral hemisphere. *See* HEMISPHERIC LATERALITY.

Building on the classical research, the more recent work has drawn attention to the important contribution that vision makes to normal speech perception; has explored the changing ability of infants to perceive speech and contrasted it with that of animals; and has studied the way that speech sounds are coded by the auditory system and how speech perception breaks down in those with hearing impairment. There has also been substantial research on the perception of words in continuous speech.

Hearing and vision. Traditionally, speech perception has concentrated on the problem of turning sound into linguistic categories; but an important influence on the field has been the need to integrate visual information about articulation with sound.

Although lip-reading is a commonly recognized skill of the hearing-impaired, the contribution that is made by visible face movements to normal speech perception was not fully realized until H. McGurk serendipitously discovered the illusion that has become known as the McGurk effect. The sound of someone saying "bah" was dubbed onto a video recording of the same person saying either "dah" or "gah." Listeners who simply heard the sound track reported correctly "bah," but those who also saw the face reported that the person had said "dah."

Vision seems to give the listener clear information that can overrule or combine with the information that is being received by the ear. In this example, vision clearly indicates that the speaker's lips did not close. Since the lips should have closed if a "b" had been produced, perception produces a conciliatory percept that is both compatible with the seen information and auditorially similar to the actual sound.

Vision gives very clear information about certain speech sounds (primarily those involving the lips and jaw); however, it provides very little about other sounds and none about the presence or the value of voice pitch. The presence of voicing conveys important information about consonants, distinguishing "b" from "p" for instance; and the pitch contour indicates the position of stressed syllables, and varies the semantic and pragmatic value of an utterance. Prostheses for the profoundly deaf that simply convey information about voice pitch can substantially improve their ability to understand speech.

The McGurk effect is much weaker in young children than in adults, but very young infants are surprisingly sensitive to the relationship between speech sounds and the corresponding face movements. Even at a few months old, they will prefer to look at faces that match the sounds they are hearing than at those that do not. Attention to the visible as well as the audible consequences of articulation is undoubtedly helpful to the child in learning language. A visually simple gesture, such as closing the lips, can influence many different auditory cues; seeing the single gesture may help the child to grasp the underlying integrity of the sound. Visually handicapped children are slower than normal children in learning to produce sounds that have a visible articulation.

Perception and development of speech categories. Adult listeners are exquisitely sensitive to the differences between sounds that are distinctive in their language. The voicing distinction in English (between "b" and "p") is cued by the relative timing of two different events (stop release and voice onset). At a difference of around 30 milliseconds, listeners hear an abrupt change from one category to another, so that a shift of only 5 ms can change the percept. On the other hand, a similar change around a different absolute value, where both sounds are heard as the same category, would be imperceptible. The term categorical perception refers to this inability to discriminate two sounds that are heard as the same speech category.

Categorical perception can arise for two reasons: it can have a cause that is independent of the listener's language—for instance, the auditory system may be more sensitive to some changes than to others; or it can be acquired as part of the process of learning a particular language. The example described above appears to be language-independent, since similar results have been found in animals such as chinchillas whose auditory systems resemble those of humans. But other examples have a language-specific component. The ability to hear a difference between "r" and "l" is trivially easy for English listeners, but Japanese perform almost at chance unless they are given extensive training. How such language-specific skills are developed has become clearer following intensive research on speech perception in infants.

Newborn infants are able to distinguish many of the sounds that are contrasted by the world's languages. Their pattern of sucking on a blind nipple signals a perceived change in a repeated sound. They are also able to hear the similarities between sounds such as those that are the same vowel but have different pitches. The ability to respond to such a wide range of distinctions changes dramatically in the first year of life. By 12 months, infants no longer respond to some of the distinctions that are outside their native language, while infants from language communities that do make those same distinctions retain the ability. Future experience could reinstate the ability, so it is unlikely that low-level auditory changes have taken place; the distinctions, although still coded by the sensory system, do not readily control the infant's behavior.

Listening in the speech mode. Individual speech categories (such as "p" or "b") are cued by a wide variety of different aspects of sound. For example, "b" has a longer preceding vowel and a shorter silence than does "p." In addition, when the lips open, the initial burst of noise energy has a lower amplitude; and when voicing starts the peak of the low-frequency energy is at a lower frequency for "b" than for "p." Experiments with synthetic speech show that these different cues can be traded off against each other: a longer silence can be compensated for by a lower first format onset frequency. Such so-called trading relations between these different cues may sometimes be explained by basic sensory mechanisms such as adaptation in the auditory nerve, but others require more specialized mechanisms specific to speech. A particularly useful way of demonstrating this is to synthesize a caricature of speech known as sine-wave speech; it consists of sine waves at the formant frequencies. Short utterances of sine-wave speech can be heard either as non-speech whistles or as (peculiar) speech. Some of the trading relations between cues in normal speech are also found in sine-wave speech, but can be shown to arise only when the listener hears the sound as speech. They are therefore a result of special speech-processing mechanisms, rather than purely auditory ones.

Words and continuous speech. Continuous speech is not simply a concatenation of isolated syllables; syllables change their pitch, their duration, and their phonetic character in continuous speech. In extreme circumstances, they can vanish or be completely transformed—for example, "Do you want to?" becoming "Djawona?" Although such changes make it more difficult to understand how the lexicon can be accessed from the speech waveform, the changes in duration and pitch convey essential information about words, grammar, and intention.

Under good listening conditions and with clearly articulated speech, it is often possible to identify words before they have been completed verbally. A word such as "penguin" becomes identifiable in principle before the final vowel, since there is no other word in English that begins the same way (except for derived forms such as "penguins"). In practice, listeners can identify isolated words at their uniqueness points, but may not do so in normal speech where the necessary phonetic information can be absent or distorted. Subsequent context can then be used to limit the possibilities before the word is "heard." Speakers adjust the clarity of their speech depending on the context, relying on listeners' ability to use that context to make up for the decreased information in the signal. Stressed syllables are most resistant to such adulteration; they may serve as starting points for the perceptual process and, in English at least, as candidates for the beginnings of words.

The hearing-impaired. Although conductive hearing losses can generally be treated adequately by appropriate amplification of sound, sensorineural hearing loss involves a failure of the frequency-analyzing mechanism in the inner ear that humans cannot yet compensate for. Not only do sounds need to be louder before they can be heard, but they are not so well separated by the ear into different frequencies. Also, the sensorineurally deaf patient tolerates only a limited range of intensities of sound; amplified sounds soon become unbearable (loudness recruitment).

These three consequences of sensorineural hearing loss lead to severe problems in perceiving a complex signal such as speech. Speech consists of many rapidly changing frequency components that normally can be perceptually resolved. The **illustration** shows a spectrogram of speech that has been put through a filter bank similar to that found in the normal human ear. In the low-frequency region the individual harmonics of the voice are separated; at higher frequencies the ear cannot resolve individual harmonics, but it still separates the different formants (or resonances of the vocal tract).

The lack of frequency resolution in the sensorineural patient makes it harder for the listener to identify the peaks in the spectrum that distinguish the simplest speech sounds from each other; and the use of frequency-selective automatic gain controls to alleviate the recruitment problem reduces the distinctiveness of different sounds further. These patients may also be less sensitive than people with normal hearing to sounds that change over time, a disability that further impairs speech perception.

Although normal listeners can still recognize speech against a background of other noises (including speech, as in the cocktail-party effect), sensorineural patients are disproportionately affected by competing sounds. Their impaired frequency and temporal selectivities reduce their ability to separate

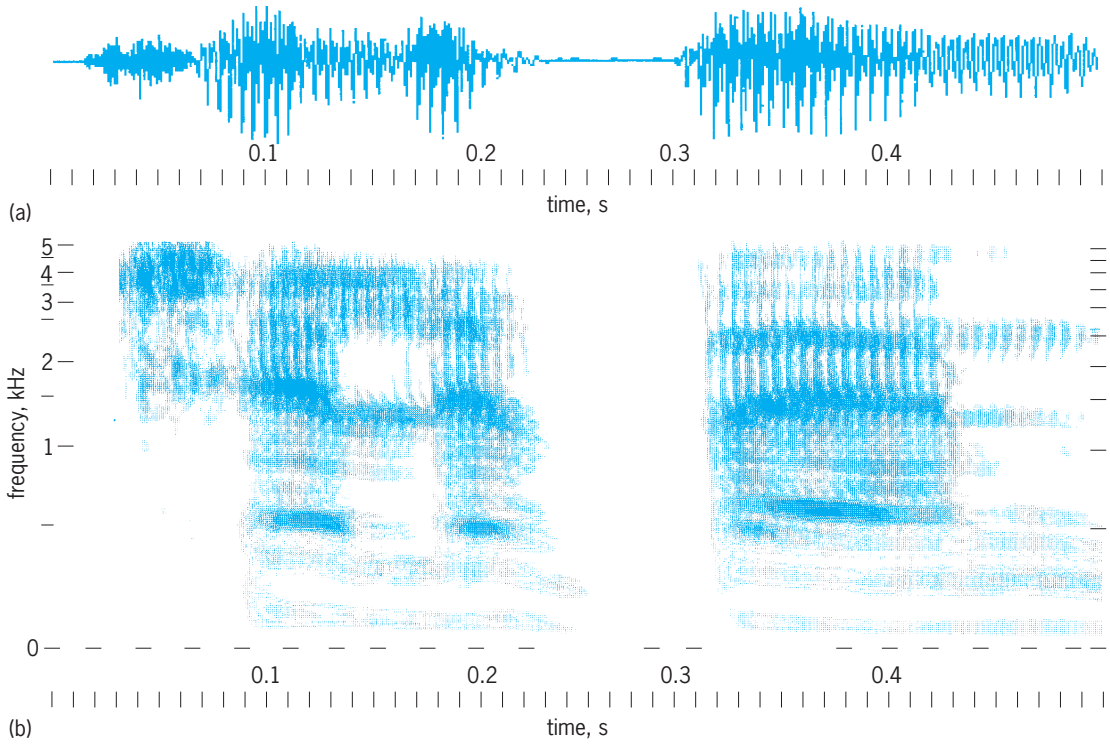

An auditory spectrogram of the world "telephone." (*a*) The vertical striations in the bands represent amplitude modulation at the frequency of the voice pitch. Such striations are absent from the initial noise portion corresponding to the voiceless initial "t." The medial "f" is too weak to appear on the spectrogram. (*b*) The first eight or so harmonics of the voice pitch are resolved and appear as dark, roughly horizontal bands. Higher harmonics are not resolved (here above about 1.5 kHz). The dark horizontal bands now represent the formant frequencies.

the required signal from its background. Hearing aids that can listen in to sound from a particular direction can help to filter out unwanted sounds, but they do not perform as well when echoes and reverberation are present. A practical solution to this problem has not been developed; however, it is possible in artificial situations to separate the speech of one speaker from another on the basis of voice pitch, so that a deaf listener can hear the selected voice more clearly. The same technique is also used by normal listeners as one way of solving the problem of competing background sounds. Extension of this method to normal situations is complicated by the fact that not all of speech has a pitch. Other, more complex information must be used to separate voiceless sounds and to link the voiced segments appropriately.

Some profoundly deaf patients can identify some isolated words by using multichannel cochlear implants. Sound is filtered into different frequency channels, or different parameters of the speech are automatically extracted, and electrical pulses are then conveyed to different locations in the cochlea by implanted electrodes. The electrical pulses stimulate the auditory nerve directly, bypassing the inactive hair cells of the damaged ear. Such devices cannot reconstruct the rich information that the normal cochlear feeds to the auditory nerve, but they promise to put profoundly deaf patients back into contact with the world of sound and to help them to communicate more effectively by speech. *See* HEARING (HUMAN); HEARING AID; HEARING IMPAIRMENT; PERCEPTION; PSYCHOACOUSTICS; SPEECH.

C. J. Darwin

Bibliography. W. D. Brooks and R. W. Heath, *Speech Communication*, 7th ed., 1993; B. Dodd and R. Campbell, (eds.), *Hearing by Eye*: *The Psychology of Lip-Reading*, 1987; U. H. Frauenfelder and L. K. Tyler (eds.), *Spoken Word Recognition*, 1978; I. G. Mattingly and M. G. Studdert-Kennedy (eds.), *Modularity and the Motor Theory of Speech Perception*, 1990; M. E. H. Schouten (ed.), *The Psychophysics of Speech Perception*, 1987.

Speech recognition

Speech recognition, in a strict sense, is the process of electronically converting a speech waveform (as the acoustic realization of a linguistic expression) into words (as a best-decoded sequence of linguistic units). At times it can be generalized to the process of extracting a linguistic notion from a sequence of sounds, that is, an acoustic event, which may encompass linguistically relevant components, such as words or phrases, as well as irrelevant components, such as ambient noise, extraneous or partial words in an utterance, and so on. Applications of speech recognition include an automatic typewriter that responds to voice, voice-controlled access to information services (such as news and messages), and automated commercial transactions (for example, price inquiry or merchandise order by telephone), to name a few. Sometimes, the concept of speech recognition may include "speech understanding," because the use of a speech recognizer often involves understanding the intended message expressed in the spoken words. Currently, such an understanding process can be performed only in an extremely limited sense, often for the purpose of initiating a particular service action among a few choices. For example, a caller's input utterance "I'd like to borrow money to buy a car" to an automatic call-routing system of a bank would connect the caller to the bank's loan department.

Fundamental concept. Converting a speech waveform into a sequence of words involves several essential steps (**Fig. 1**). First, a microphone picks up the acoustic signal of the speech to be recognized and converts it into an electrical signal. A modern speech recognition system also requires that the electrical signal be represented digitally by means of an analog-to-digital (A/D) conversion process, so that it can be processed with a digital computer or a microprocessor. This speech signal is then analyzed (in the analysis block) to produce a representation consisting of salient features of the speech. The most prevalent feature of speech is derived from its short-time spectrum, measured successively over short-time windows of length 20–30 milliseconds overlapping at intervals of 10–20 ms. Each short-time spectrum is transformed into a feature vector, and the temporal sequence of such feature vectors thus forms a speech pattern. **Figure 2** illustrates the speech analysis process.

The speech pattern is then compared to a store of phoneme patterns or models through a dynamic programming process in order to generate a hypothesis (or a number of hypotheses) of the phonemic unit sequence. (A phoneme is a basic unit of speech and a phoneme model is a succinet representation of the signal that corresponds to a phoneme, usually embedded in an utterance.) A speech signal inherently has substantial variations along many dimensions. First is the speaking rate variation—a speaker cannot produce a word of identical duration at will. Second, articulation variation is also abundant, in terms both of talker-specific characteristics and of the manner in which a phoneme is produced. Third, pronunciation variation occurs among different speakers and in various speaking contexts (for example, some phonemes may be dropped in casual conversation). Dynamic programming is performed to generate the best match while taking these variations into consideration by compressing or stretching the temporal pattern (a process often called dynamic time warping) and by probabilistically conjecturing how a phoneme may have been produced. The latter includes the probability that a phoneme may have been omitted or inserted in the utterance. The knowledge of probability (often called a probabilistic model of speech) is obtained via "training," which computes the statistics of the speech features from a large

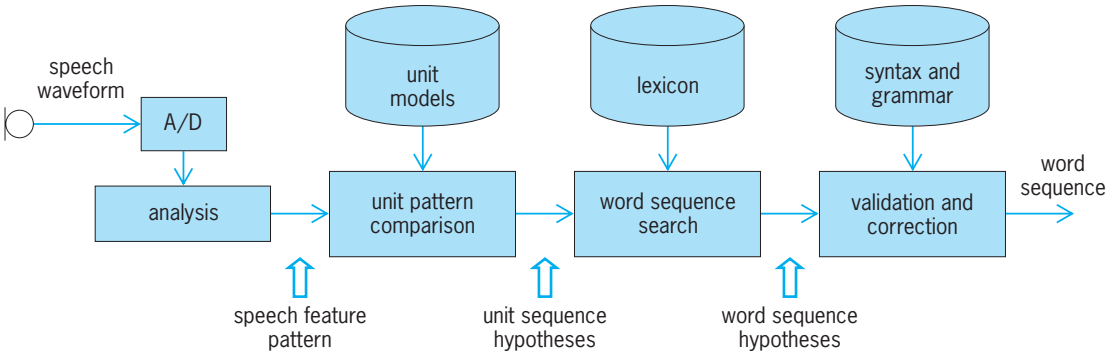

Fig. 1. Basic block diagram of a speech recognition system.

collection of spoken utterances (of known identity) according to a mathematical formalism. At the present time, the most widely adopted formalism is the hidden Markov model.

The hypothesized phoneme sequence is then matched to a stored lexicon to reach a tentative decision on the word identity. The decoded word sequence is further subject to verification according to syntactic constraints and grammatical rules, which in turn define the range of word hypotheses for lexical matching. This process of forming hypotheses about words and matching them to the observed speech pattern, and vice versa, in order to reach a decision according to a certain criterion is generally referred to as the "search." In limited domain applications in which the number of legitimate expressions is manageably finite, these constraints can be embedded in an integrated dynamic programming process to reduce search errors.

Categorization of tasks. Most current speech recognition systems are essentially designed to convert a speech waveform into one or more words in order to perform a specific task. Linguistic expressions via speech can take the form of word utterances spoken individually in isolation, or phrases and sentences spoken naturally without intentionally controlled word boundaries. A task involving the former case is called isolated (or discrete) word recognition, while the latter case is generally referred to as continuous speech recognition. However, the speaker's articulation effect can be substantially different when reading a prepared text and when conversing spontaneously with another person. Also at times, a task may involve connected words without the usual syntactic constraints, such as a digit sequence (for example, one's credit card number). Thus, a speech recognition task can be categorized, in terms of the way that speech is rendered, into four essential types: isolated word, connected word, continuously read speech, and spontaneous speech.

The degree of sophistication of a speech recognition task is largely a function of the size of the vocabulary it has to deal with. A task involving, say, less

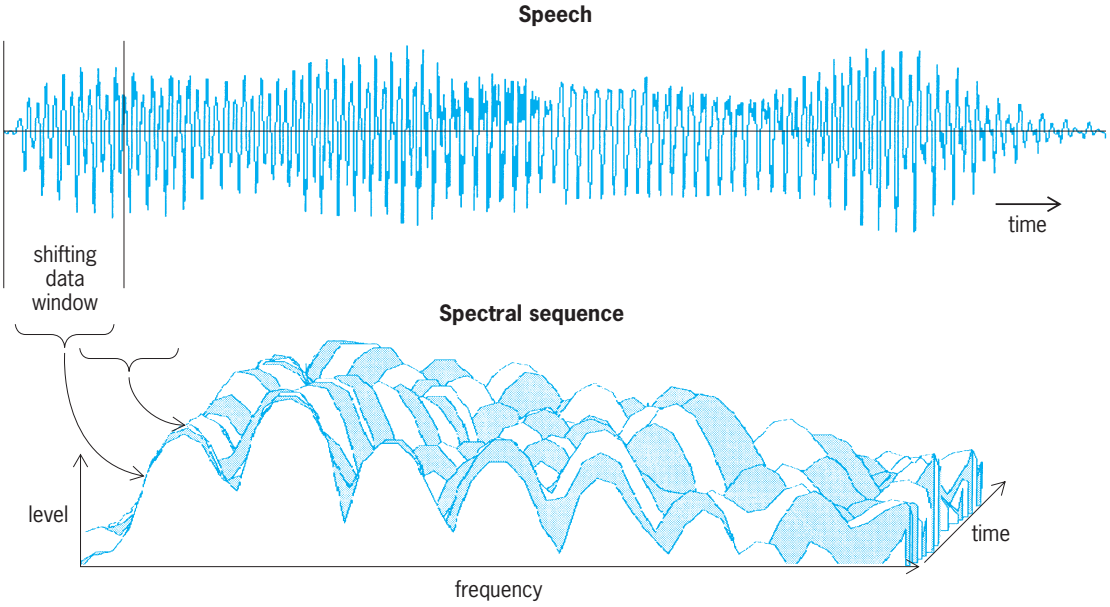

Fig. 2. Short-time spectral analysis of speech.

than 100 words is called small vocabulary recognition and is mostly for command-and-control applications with isolated word utterances as input. There are usually very few grammatical constraints associated with these types of limited tasks. When the vocabulary size grows to 1000 words, it is possible to construct meaningful sentences, although the associated grammatical rules are usually fairly rigid. For dictation, report writing, or other tasks such as newspaper transcription, a speech recognition system with a large vocabulary (on the order of tens of thousands of word entries) is needed.

A speaker-independent recognition system is designed to serve nonspecific users and is suitable for public service applications where the user population is large. A voice dialing system embedded in a telephone central office switch, a telephone call-routing system, and a customer service unit that can respond to a caller's voice without having to know the user's identity are all in this category. It has to be able to handle a broad range of accents and voice characteristics from various talkers. Another type of speech recognizer, called a speaker-dependent recognition system, may be designed to serve only a specific user; for example, a personal word-processor that responds to its authorized user's voice. A speaker-dependent system requires speaker-specific training and is thus also referred to as a speaker-trained system. Such a system can also be designed by adapting a speaker-independent system to the voice patterns of the designated user via an enrollment procedure; it is often called a speaker-adaptive system. For a limited number of frequent users of a particular speech recognizer, such an enrollment procedure generally leads to good recognizer performance.

Understanding of speech. Another type of speech recognition system has to do with the fundamental idea of task-oriented speech understanding. Using the popular call-routing system that sorts operator-assisted calls into five types (credit-card, collect, third-party billing, person-to-person, and operator) as an example, a recognition system design engineer can use a large-vocabulary continuous speech recognition system to convert the request utterance (for example, "I'd like to make a collect call") into a word sequence as best it can and then employ a natural language processing system to parse and understand the intent of the caller. Alternatively, the system designer can make the assumption that in such a limited task the intent of the caller will most likely be expressed in a "keyword" (for example, the word "collect" in the above sentence) embedded in the otherwise naturally spoken sentence. A system that can detect or spot the keyword would thus suffice for the task. This alternative approach is referred to as keyword spotting.

The keyword-spotting system has proved to be very useful, because it makes the interaction between the user and the machine natural and robust. Experience shows that many users of a speech recognition system in telecommunication applications often do not speak words or phrases according to the prescribed vocabulary, thus creating out-of-vocabulary (OOV) or out-of-grammar (OOG) errors. Rather than attempting to recognize every word in the utterance, a keyword-spotting system hypothesizes the presence of a keyword in appropriate portions of the speech utterance and verifies the hypothesis by a likelihood ratio test. By avoiding forced recognition decisions on the unrecognizable and inconsequential regions of the speech signal, the system can cope with natural sentences, obtaining satisfactory results in constrained tasks. A more sophisticated keyword spotting system may involve "gisting," that is, trying to compile the detected keywords into a possible message. Similarly, it is possible to perform a primitive analysis on the decoded words based on their correlation with a particular semantic notion in order to derive an understanding for limited domain applications.

Implementation. Realization of a speech recognition system involves two essential steps: training and decoding. The objectives of training include reference models of the phonetic units, which characterize (probabilistically) how linguistic units are acoustically realized (acoustic modeling), and probabilistic relationships among words that form grammati-cally legitimate sentences of a language (language modeling). These acoustic unit models and language models are the basis of reference for inferring a decision when given an unknown input utterance. Training is usually performed off-line in non–real time.

The hidden Markov model is the prevalent choice for the reference structure in acoustic modeling. A hidden Markov model is a doubly stochastic process involving (at least) two levels of uncertainty. It is a finite state (say N-state) machine in which each state is associated with a random process. The random process in each state accounts for the variation in the realization of a speech code (such as a phoneme), including both the spectral (pertaining to the vocal tract) and the temporal (pertaining to the rate of articulation) behavior. The finite state machine is a Markov chain, which characterizes the probabilistic relationship among the states in terms of how likely one state is to follow another state. **Figure 3** illustrates a three-state hidden Markov model. *See* STOCHASTIC PROCESS.

Training of a hidden Markov model requires a large collection of sample utterances, each with a known label. In the case of isolated-word recognition, each utterance would consist of a word. The identity or label of the utterance is usually manually obtained by a human listener. For continuous speech recognition, each utterance in the training set would be a sentence, represented by a manually annotated word sequence. Model training involves a procedure in which locating the proper unit boundaries in the utterance (segmentation) and adjusting the model parameter values for each encountered models (model estimation) interleave and iterate until a certain criterion is optimized.

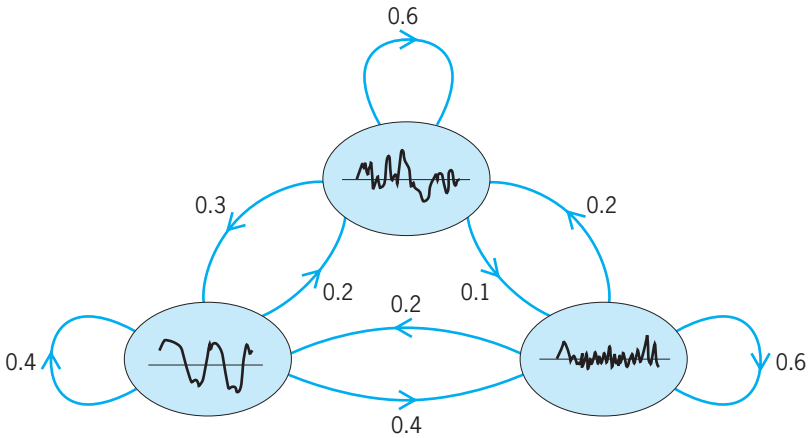

Fig. 3. Three-state hidden Markov model.

A similar principle applies in the case of language modeling, which is imperative in large-vocabulary continuous speech recognition. Language modeling is usually performed on a text database consisting of tens or hundreds of millions of words, sufficient to allow derivation of the probabilistic characteristics of a language (for example, the likelihood of a particular word appearing after another).

Decoding is the process of evaluating the probabilities that the unknown utterance manifests the same characteristics as each of the stored references and inferring the identity of the utterance based on the evaluated probabilities. For large-vocabulary, continuous speech recognition, the reference models (such as sentence hypotheses) are constructed from unit models and can be prohibitively large in number. **Figure 4** shows an example of representing a hypothesized sentence ("Show all alerts") by a sequence of context-dependent phoneme units. Efficient search algorithms need to be invoked such that the decoding result can be obtained within a reasonable computational bound. For a recognition system to be useful, it needs to be able to produce the recognition result without obvious latency, preferably in real time.

Except for some extremely simple tasks, a speech recognition system is implemented as a stored program for execution with a computer or a digital signal processor. The reference models are stored in read-only devices, with easy data access for decoding purposes. For telecommunication applications in which the system often handles multichannel traffic or sessions, implementation efficiency is achieved by sharing the reference models and proper division of the computational task among multiple channels.

Difficulties and remedies. Speech recognition research has made great progress since 1970. However, the current technology is still far from achieving its ultimate goal: accurate recognition of unconstrained speech from any speaker in any environment or operating condition. There are several key reasons why this remains a challenging problem.

One primary difficulty in speech recognition is that the variation in articulation is difficult to characterize. While phonemes are considered the fundamental speech units demonstrating regular acoustic-phonetic behavior (as determined by human ears), their acoustic realization usually contains ambiguity and uncertainty. A phoneme recognition system that attempts to convert a speech waveform into a phoneme sequence without incorporating additional knowledge, such as vocabulary entries and syntax, cannot achieve high accuracy. Grouping these error-prone phoneme or subword sequences into a sequence of words from a standard lexicon is a nontrivial task. Units as large as words and phrases have been suggested for limited task environments with good success. Large speech units, however, are unsuitable for large-vocabulary, continuous speech recognition. The longer the basic unit, the more the number of unit-classes there will be for the system to deal with. (Consider the extreme case of using the sentence as the unit in continuous speech recognition. The number of possible sentences becomes astronomical even with a moderate vocabulary size.) Having more unit classes also means higher complexity in training and decoding. For practical

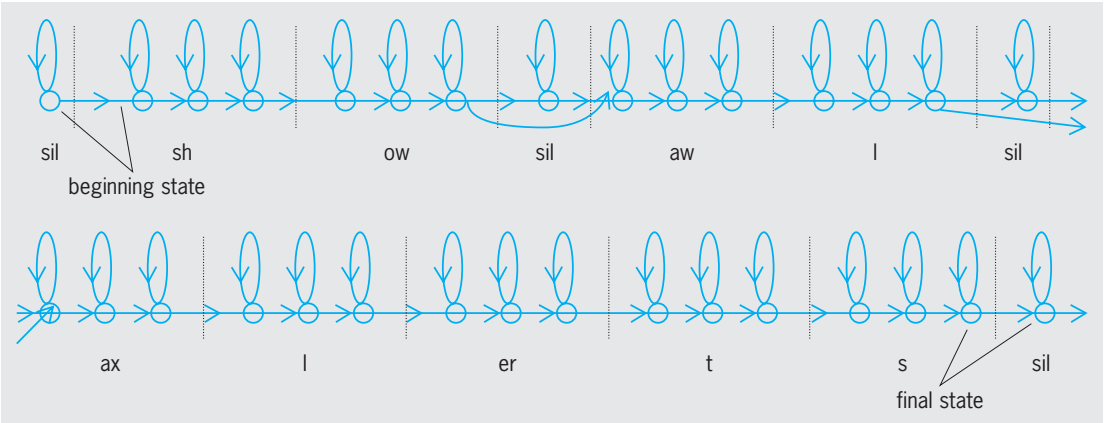

Fig. 4. Sentence representation of "show all alerts" with a finite state network using context-dependent units; "sil" denotes silence.

reasons, subword units such as phonemes, diphones (pairs of phonemes), or syllables, which are limited in number almost independent of the vocabulary size, are preferred. To circumvent the articulation problem, modern large-vocabulary, continuous speech recognition systems almost universally use context-dependent unit models, an example of which is shown in Fig. 4. A context-dependent phoneme model is a statistical characterization of a phoneme when uttered in a particular context; for example, a phoneme unit /a/ that follows a /n/ and precedes a /p/, as in "snap," is a member of the unit classes with model denoted by 'n-**a**-p.' This class is treated as a class different from, say, 'm-**a**-d' in "mad." Context dependency helps reduce the degree of acoustic-phonetic variation within the class and often leads to higher recognition accuracy.

Another major difficulty is the variation in the characteristics of the transmission medium over which recognition is performed (for example, over a telephone line, in an airplane cockpit, or using a hands-free speakerphone). This kind of variation includes the acoustic background noise, the various distortions inherently caused by the microphones that are used to pick up the acoustic signal, reverberation that varies substantially from one room to another, and the transmission distortion existing in the telephone network. To cope with this type of variation, assumptions, such as that the noise is additive and that the distortion is linear, are usually made. Another implicit assumption about the characteristics of the interference is that they do not change as rapidly as speech. For example, in telephone applications, most of the transmission properties are considered virtually constant over a reasonable period of time once the phone circuit connection is made. Signal processing methods to remedy these adverse effects include spectral subtraction to reduce the impact of additive noise, and cepstral bias or mean removal to suppress the influence of spectral distortions. The principle behind spectral subtraction and cepstral bias removal is quite similar. Spectral subtraction operates in the power spectrum domain and assumes that the power spectrum of the received signal is the sum of the speech spectrum and the noise spectrum. Cepstral bias removal operates in the log power spectrum domain and hence assumes that the log power spectrum of the received signal is the sum of the log power spectrum of the speech and the log frequency response of the transmission channel. In both cases, the system estimates the additive noise/bias component from an average of the signal (in the corresponding domain taken over a relatively long period, perhaps the whole utterance, based on the aforementioned assumption on the slow-changing nature of these adverse effects) and subtracts the bias estimate from each of the short-time speech representations that form the speech pattern of the utterance.

Talker variation is yet another source of difficulty. Variations due to the talker include regional accents, articulation and pronunciation, and mode and mood of speaking. Talker adaptation provides means for the system to adjust its operating parameters, such as the stored reference models, to suit the particular user. However, very limited knowledge exists at present on the effect of psycholinguistic factors, such as the contextual or acoustic environment (for example, a noisy café) in which the speech conversation takes place, upon the manifestation of acoustic-phonetic features in the speech signal. Some studies exist on the Lombard effect, which is used to describe the observation of change in the talker's articulation effort when he or she speaks in a noisy environment (for example, trying to raise the voice or to make his or her voice better understood by the listener).

Similarly, studies to establish systematic understanding of the difference between read and spontaneous speech are just beginning. The difference is intuitively obvious, but the science behind it is not. Unconstrained automatic speech recognition by a machine that converses with people like a human is still a challenging undertaking.

State of the art. Speech recognition by the end of the twentieth century achieved modest success in limited applications, particularly with systems designed to recognize small- to moderate-size vocabularies for well-defined tasks (such as word processing and document editing on a personal computer using voice, dictating a readout in a radiologist's office, and telephone dialing by speaking the name or the number).

A summary of typical speech recognition performance in terms of word error rate is provided in **Tables 1** and **2**. Table 1 shows the average word error rate for recognition of isolated words without contextual constraints for both speaker-dependent (SD) and speaker-independent (SI) systems. In general, the performance is more sensitive to vocabulary complexity than to vocabulary size. The 39-word alpha-digits vocabulary (letters A–Z, digits 0–9, three command words) contains highly confusable word subsets, such as the E-set (B, C, D, E, G, P, T, V, Z) and the digit 3, and has an average accuracy of 93–96%. In comparison, a task involving an 1109-word basic English vocabulary has an average accuracy of 96%.

For connected-word recognizers applied to tasks with various types of syntactic or grammatical constraints as in natural language, an additional parameter called perplexity is needed to indicate roughly the degree of difficulty of the task. Perplexity is defined as the average number of words that can follow an arbitrary word. It is usually substantially less than the size of the vocabulary because of the constraints of the task-specific language. Table 2 shows the typical performance, again in terms of word accuracy, of connected and continuous speech recognition systems for various tasks.

The resource management task involves a highly stylized language used in military duties with a perplexity of no more than 60. Sentences in the air travel information system are usually simple query utterances aiming at flight information such as

TABLE 1. Performance of isolated word recognition systems

Recognition task	Vocabulary size	Mode	Word accuracy
Digits (0–9)	10	SI	~100%
Voice dialing	37	SD	100%
Alpha-digits and command words	39	SD	96%
		SI	93%
Computer commands	54	SI	96%
Air travel words	129	SD	99%
		SI	97%
City names (Japanese)	200	SD	97%
Basic english	1109	SD	96%

airfare, flight time, and the origin and destination of the trip. The vocabulary includes a number of city or airport names in the task database. While the application language appears to be natural, the grammatical structure in the query sentence is actually rather limited with a perplexity of less than 25. In the business newspaper transcription task, the input is "read" speech; speakers read the text of news articles in the paper. Read speech has less variability in its acoustic realization, and is thus easier to recognize than spontaneous, conversational speech. The broadcast news transcription task involves "found" speech, such as the news announcements from radio or television. The signal may be degraded by noise (such as background music or conversation) and distortion (for example, due to radio transmission). Obviously, this is a more difficult task than transcribing a news article in a quiet office.

Despite progress, the performance of current systems is barely acceptable for large vocabulary systems. Tasks listed in the tables are all considered substantially easier than transcribing and understanding truly spontaneous conversational speech.

Speaker recognition. The technology of speech recognition often finds applications in speaker recognition tasks as well. Speaker recognition applications can be classified into two essential modes, speaker identification and speaker verification. The goal of speaker identification is to use a machine to find the identity of a talker, in a known population of talkers, using the speech input. Speaker verification aims to authenticate a claimed identity from the voice signal. These two modes call for different methodologies. Applications of speaker recognition include controlled entry to restricted premises, and authorization of access to privileged information or other service transactions. For these applications, speaker verification is used in most practical systems.

There are also two types of talker authentication with regard to the underlying technology. Whereas speaker verification (SV) verifies a speaker's identity based on speaker-specific characteristics reflected in spoken words or sentences (sometimes referred to as a talker's voice print), another method, known as verbal information verification (VIV), checks the content of the spoken password or pass-phrase [such as a personal identification number (PIN), a social security number, or a mother's maiden name]. In the deployment of a VIV system, the user assumes the responsibility in protecting the confidentiality of the password as in many electronic transactions applications.

Research aimed at finding acoustic and linguistic features in the speech signal that carry the characteristics of the speaker has a long history. The advent of a statistical pattern matching approach (such as the hidden Markov model) and its dominance in the 1990s has caused a convergence in the technological approach to this age-old problem. Statistical modeling methods are data-driven and are able to automatically "learn" speaker-specific (for SV) or password-specific (for VIV) characteristics from the provided spoken utterances with a known identity. During verification, the talker first makes an identity claim, such as a name or an account number, and then, in response to the system prompt or question, speaks a test phrase to the system. The system evaluates the input speech against the stored model for the person with the claimed identity (in SV) or the stored model for the password (in VIV), and decides to accept or reject the talker.

Key factors affecting the performance of a speaker verification system are the nature of input (whether the utterance is based on a specific text and how long the test utterance is) and the operating environment in which the verification system is used. The performance of a verification system is usually

TABLE 2. Performance benchmark of state-of-the-art automatic speech recognition systems for various connected and continuous speech recognition tasks

Recognition task	Vocabulary size	Perplexity	Word accuracy
Connected digit strings	10	10	~99%
Resource management	991	<60	97%
Air travel information system	1,800	<25	97%
Business newspaper transcription	64,000	<140	94%
Broadcast news transcription	64,000	<140	86%

measured by the equal error rate (EER), which is defined as the error rate of the system when the operating threshold for the accept/reject decision is adjusted such that the probability of false acceptance and that of false rejection become equal. A state-of-the-art system using a text-dependent sentence-long test utterance is capable of achieving an EER of 1–2%. With short, unconstrained spoken utterances in a noisy environment, current systems would have an EER of 4–8%.

The performance of a VIV system depends on how many passwords or pass phrases the user is asked to give as input to the system in a verification session. Experiments showed that a VIV system could achieve perfect (0 error) verification after three rounds of question and answer.

The deployment of speaker verification often involves nontechnical issues, such as the trade-off between false rejection and false acceptance. In business applications, such as an automated banking machine where a false rejection may mean unjustified inconvenience causing the loss of a legitimate long-term customer, the operating threshold may be set to allow a high false-acceptance rate. Determination of the operating point is thus a matter of business decision. *See* LINGUISTICS; PSYCHOACOUSTICS; SPEECH.

Biing Hwang Juang

Bibliography. J. L. Flanagan, *Speech Analysis, Synthesis, and Perception*, Springer-Verlag, New York, 1972; S. Furui, *Digital Speech Processing, Synthesis, and Recognition*, Marcel Dekker, New York, 2001; B. H. Juang and S. Furui (eds.), Spoken Language Processing, special issue of *IEEE Proc.*, 2000; L. R. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition*, Prentice Hall, Englewood Cliffs, NJ, 1993; L. R. Rabiner and B. H. Juang, An introduction to hidden Markov models, *IEEE Signal Proc. Mag.*, 3(1):4–16, 1986.

Speed

The time rate of change of position of a body without regard to direction. It is the numerical magnitude of a velocity and hence is a scalar quantity. Linear speed is commonly measured in such units as meters per second, miles per hour, or feet per second. It is the most frequently mentioned attribute of motion.

Average linear speed is the ratio of the length of the path Δs traversed by a body to the elapsed time Δt during which the body moved through that path, as in Eq. (1), where s_0 and t_0 are the initial position

$$\text{Speed (average)} = \frac{s_f - s_0}{t_f - t_0} = \frac{\Delta s}{\Delta t} \qquad (1)$$

and time, respectively, s_f and t_f are the final position and time, and Δ stands for "the change in."

Instantaneous speed, defined by Eq. (2), is the lim-

$$\text{Speed (instantaneous)} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = \frac{ds}{dt} \qquad (2)$$

iting value of the foregoing ratio as the elapsed time approaches zero. *See* VELOCITY.

Rogers D. Rusk

Speed regulation

The change in steady-state speed of a machine, expressed in percent of rated speed, when the load on the machine is reduced from rated value to zero. The definition of regulation usually means the net change in a steady-state characteristic, and does not include any transient deviation or oscillation that may occur prior to reaching the new operation point. This same definition is used for stating the speed regulation of electric motors and for certain prime movers, such as steam turbines.

In specifying the speed regulation of dc motors, for example, one would compute the steady-state regulation R by Eq. (1), where N_R is the speed (in any

$$R = \frac{N_0 - N_R}{N_R}(100) \qquad \text{percent} \qquad (1)$$

convenient units) at rated load and N_0 is the speed (in the same units used to specify N_R) after the load is removed and a new steady-state value is reached. Regulation is usually stated as a positive quantity, even though the regulation characteristic of speed versus load, when plotted as a straight line, would have a negative slope.

For small load changes about an operating point, an incremental steady-state regulation is also defined by Eq. (2), where both the speed N and power P

$$R = \frac{\Delta N}{\Delta P}(100) \qquad (2)$$

are normalized by using rated values as the base quantities. Equation (2) recognizes the fact that the speed-versus-power regulation characteristic is not necessarily a straight line. In computing the regulation it should also be specified that all adjustments to the speed-changing mechanism should remain unchanged throughout the test. *See* MOTOR.

The speed regulation characteristic of steam turbines driving synchronous generators is an important parameter in the control of frequency for large power systems. Here the speed regulation, often called the droop characteristic, determines not only the overall systems frequency response characteristic to load changes but also how any load change is shared by the connected turbine-generator units. The droop characteristic is important in estimating steady-state response, but this parameter enters prominently into the determination of transient behavior as well. *See* GENERATOR.

The definition of steady-state regulation (1) is also used in the specification of the performance of automatic control systems, where it is used as the measure of goodness of a controlled variable, such as speed. *See* CONTROL SYSTEMS.

Paul M. Anderson

Bibliography. N. Cohn, *Control of Generation and Power Flow on Interconnected Power Systems*, 1966; IEEE Power Engineering Society, Power Generation Committee, *Recommended Practice for Functional and Performance Characteristics of Control Systems for Steam Turbine-Generator Units*, IEEE Std. 122–1991; *IEEE Standard Dictionary of Electrical and Electronics Terms*, 7th ed., IEEE Std.

100–2001; L. K. Kirchmayer, *Economic Control of Interconnected Systems*, 1959.

Speedometer

A device for indicating the speed of a vehicle. The mechanical analog speedometer (**Fig. 1**) was introduced around 1906, when automobiles were beginning to increase in numbers. While this type of speedometer continues in wide use, changing technology has resulted in increasing numbers of quartz electric analog speedometers and digital microprocessor speedometers.

Mechanical speedometer. The mechanical speedometer is driven by a cable housed in a casing and connected to a gear at the transmission. This gear is designed for the particular vehicle model, considering the vehicle's tire size and rear axle ratio.

In most cases, the speedometer is designed to convert 1001 revolutions of the drive cable into registering 1 mi on the odometer, which records distance traveled by the vehicle. Thus, 1001 revolutions of the cable per minute will indicate a speed of 60 mi/h (96 km/h). Odometers with six wheels record 99,999 mi (or kilometers), while those with seven wheels record 999,999 mi (or kilometers).

The speed-indicating portion of the speedometer operates on the magnetic principle. In the speedometer head, the drive cable attaches to a revolving permanent magnet that rotates at the same speed as the cable. Floating on bearings between the upper frame and the revolving permanent magnet is a nonmagnetic movable speed cup. (There is almost no mechanical connection between the revolving magnet and the speed cup.) The magnet revolves within the speed cup, producing a rotating magnetic field. The magnetic field is constant, and the amount of speed cup movement is at all times in proportion to the speed of the magnet rotation. The magnetic field exerts drag on the speed cup, causing it to revolve in the same direction and in direct ratio to the speed of the revolving magnet and the vehicle itself. A pointer, attached to the speed cup spindle, indicates the speed on the speedometer dial. The speed cup movement is opposed by a hair spring, so that the speed cup is held steady and the pointer indicates the true speed. The hair spring also pulls the pointer back to a point less than 5 mi/h (8 km/h) when the magnet stops rotating.

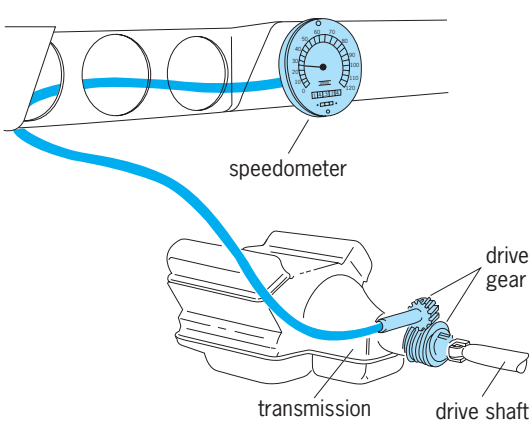

Fig. 1. Components of a mechanical speedometer system.

Quartz electric speedometer. The quartz electric speedometer and odometer assembly (**Fig. 2**) displays the speed of the vehicle, total vehicle mileage, and trip mileage. A conventional (analog) dial and pointer assembly, along with odometer wheels, are used to display this information. However, due to improved drive methods, the need for the conventional speedometer (flexible shaft) cable is eliminated.

To provide accurate vehicle information, the quartz speedometer utilizes an accurate clock signal supplied by a quartz crystal, along with integrated electronic circuitry to process an electrical speed signal. This signal, used by the circuitry to drive the air-core gage and odometer stepper motor, is generated by a permanent-magnet generator mounted in the transmission. This permanent-magnet generator, designed to be used with both quartz and digital speedometers, provides a sinusoidal speed signal that is proportional to vehicle speed at the rate of 4004 pulses per mile (2503 per kilometer). This speed signal is transmitted to a buffer amplifier which filters and squares the speed signal. After conditioning, the signal is transmitted to the speedometer circuitry at a rate of 1.111 Hz/(mi)(h) [0.694 Hz/(km)(h)].

The odometer assembly for the quartz electric speedometer consists of conventional odometer wheels driven by an electric stepper motor. The stepper motor receives a frequency-modulated square-wave signal of 0.556 Hz/(mi)(h) [0.348 Hz/(km)(h)] from the speedometer circuitry. This signal controls the rate at which the motor turns the odometer gears.

The air-core gage mechanism in the electric speedometer consists of a coil, bobbin, and case assembly that utilizes electrical signals supplied by the circuit assembly to display vehicle speed. Current flow through the coils of the air-core gage causes a flux field to be induced. As the flux field varies, the armature magnet inside of the coils rotates to align itself with the induced field. A spindle and pointer is attached to the magnet and indicates information on the dial. By controlling the flow of current through the coils (via the electronic circuit), the air-core gage can be used to display accurately the desired analog information. Once the current potential across the coils is removed, the pointer returns to zero.

Digital microprocessor speedometer. In this type of instrument (**Fig. 3**), the vehicle speed is monitored by the permanent speed sensor mounted in the transmission. Like the quartz speedometer system, the speed signal is filtered and squared by the buffer amplifier. After conditioning, the signal is transmitted to the microprocessor where the counter converts the speed signal to a digital signal and stores it in memory. The timing circuit has the capacity to handle

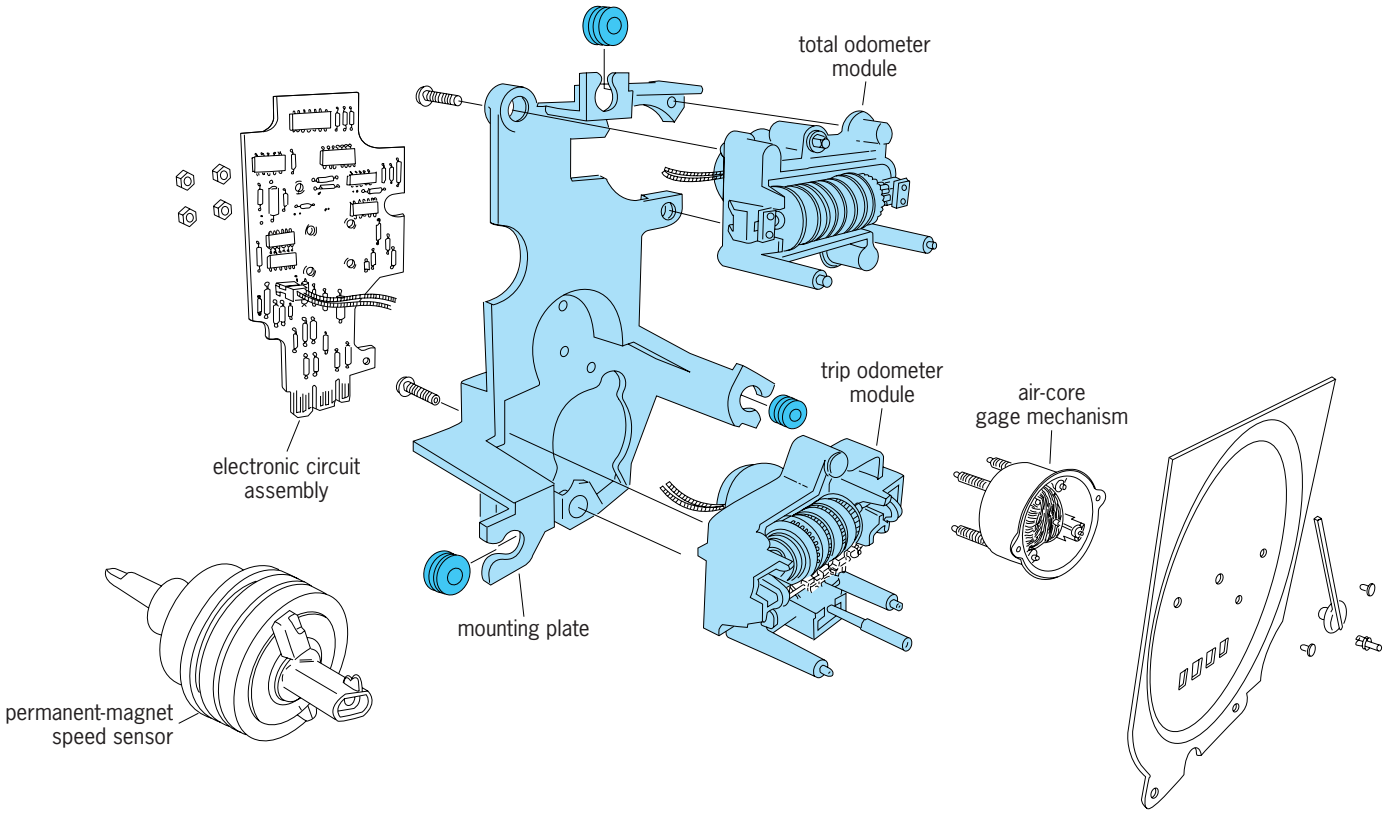

Fig. 2. Components of a quartz electric speedometer system.

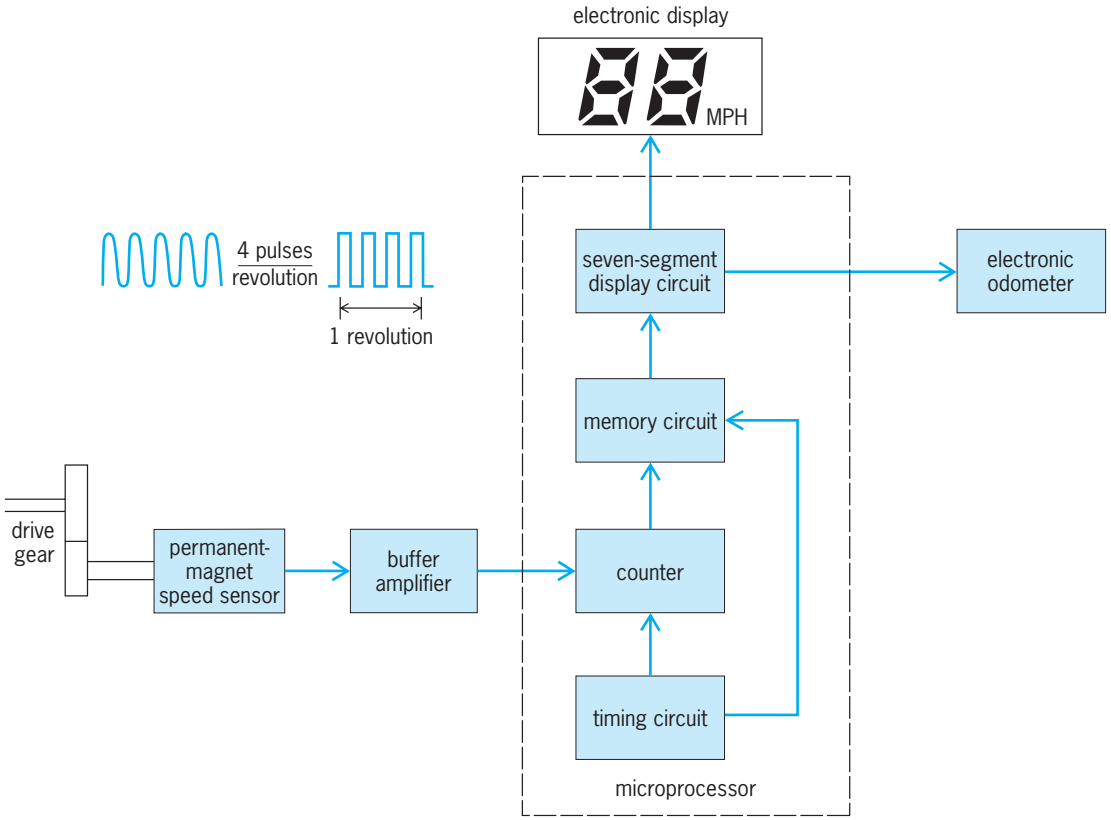

Fig. 3. Schematic diagram of the operation of a digital electronic speedometer. The permanent-magnet speed sensor generates four pulses for every revolution; at 60 mi/h (97 km/h), 4004 pulses are generated.

the counter and memory storage in less than 0.25 s. Memory circuit signals are sent to the display circuit, which selects the display numerals representing the vehicle's speed, according to the number of pulses received from the speed sensor. The electronic displays are light-emitting diodes, fluorescent tubes, liquid crystals, plasma gas discharge tubes, or electroluminescent panels. Odometers used in conjunction with digital speedometers are stepper motor–driven drum registers or digital displays employing nonvolatile memory. *See* AUTOMOTIVE TRANSMISSION; ELECTRONIC DISPLAY; TACHOMETER.

Robert A. Grimm

Bibliography. D. Maschino, D. Taylor, and G. Thoman, *Practical Electric Speedometer System for Passenger Car Application*, SAE Tech. Pap. 830326, February 1983; Society of Automotive Engineers, *SAE Handbook Nineteen Eighty-Eight: Parts and Components*, vol. 2., 1988; W. K. Toboldt, *Automotive Encyclopedia*, rev. ed., 2000.

Spelaeogriphacea

A crustacean order within the class Malacostraca, superorder Peracarida. Few spelaeogriphacean species are known, four extant species and two supposed fossil species. *Spelaeogriphus lepidops* is from a stream in a cave in Table Mountain, South Africa. *Potiicoara brasiliensis* inhabits a pool in a cave in the Mato Grosso, Brazil. *Mangkurtu mityula* (see **illustration**) and *M. kutjarra* have been reported from aquifers in the Pilbara region of Western Australia. *Acadiocaris novascotica* is a fossil from Carboniferous marine sediments in Canada, and *Liaoningogriphus quadripartitus* is a fossil from lacustrine (lake-associated) deposits of Jurassic age in China. It is probable that *Acadiocaris* is truly a spelaeogriphacean, but it is less certain that *Liaoningogriphus* is; on the basis of characters visible in fossils, its classification in this order or any other is impossible to justify.

Spelaeogriphus lepidops is the largest of the living species, reaching 9 mm (0.36 in.) in length; the others attain half that length at most. All have a slender flexible body of similar-sized segments. The head, incorporating the first thoracic segment, supports a short carapace that extends backward over the top and sides of the second thoracic segment. The carapace encloses a branchial space in which sit a pair of maxillipedal gills, less developed in *Potiicoara* than in the other genera. A pair of articulating disc-like eyestalks sit on top near the front of the head but without pigmented eyes. The limbs on the head are a pair of antennules with a prominent scaphocerite (scale), a pair of long biramous (having two branches) antennae, complex mandibles with grinding molar surfaces, incisor and palp (sensory appendage near the mouth), setose maxillae and maxillules and maxillipeds each with a

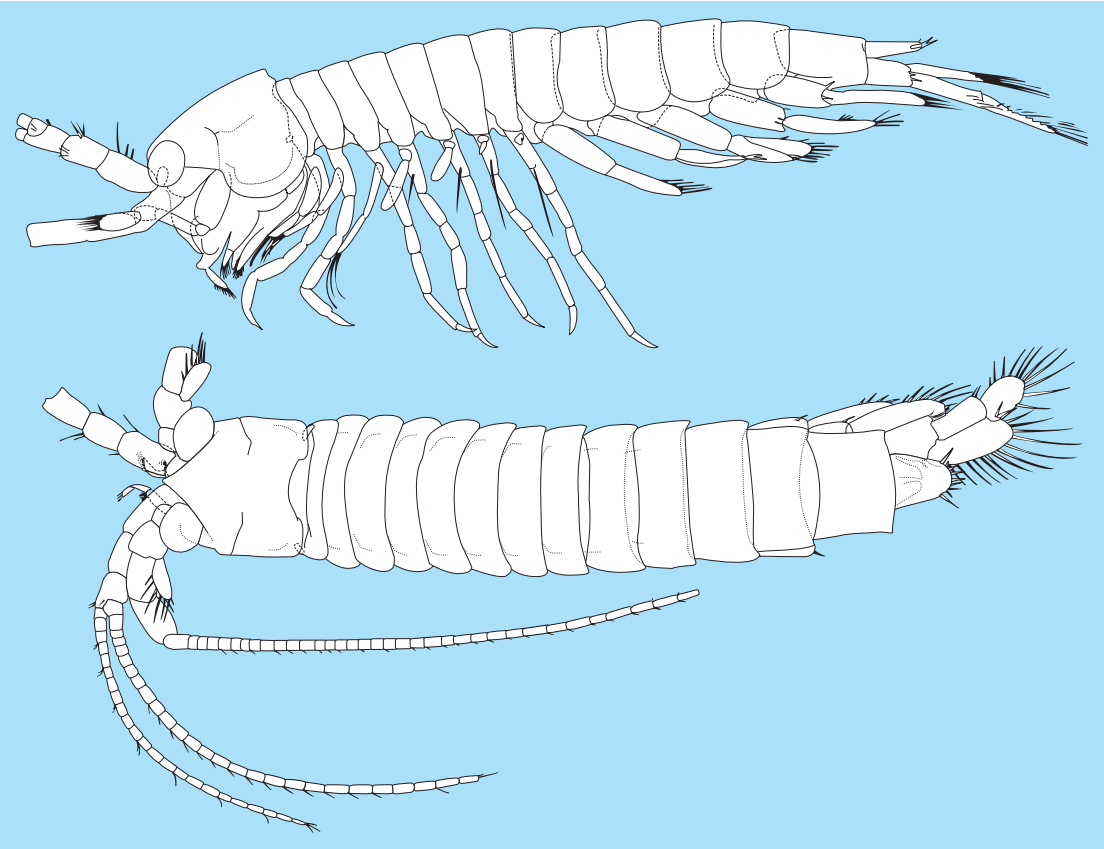

Lateral (top) and dorsal views of the Western Australian spelaeogriphacean species *Mangkurtu mityula*.

prominent basal endite. The patterns of setae (bristles) on the mouthparts distinguish speleaogriphaceans from the most closely related peracaridan order, Mictacea. The seven pairs of walking legs on the remaining thoracic segments are thin and biramous, with the long exopods on the more anterior ones generating a respiratory current under the carapace and past the reduced posterior exopods. Each of the first five abdominal segments bears a pair of biramous flattened pleopods (abdominal appendages used for swimming) except in *Spelaeogriphus* where the last pair is rudimentary. The last abdominal segment carries a tail-fan of paired uropods and terminal telson. Females incubate 10–16 eggs in four or five pairs of oostegites, ventral plates between the thoracic legs (as in most Peracarida). Males and females differ only in features of the antennae and pleopods, which are not the same in all genera. The food consists of detritus manipulated by the mouthparts. The animals creep over the bottom with their walking legs and swim by undulations of the whole body. All extant species are inhabitants of fresh water in subterranean environments, caves, or aquifers. It is hypothesized that the group arose in the sea but dispersed to Gondwanan continents (the continents that formed from the fragmentation of the ancient southern supercontinent Gondwana, including South America, Africa, Antarctica, and Australia), where they were stranded in localized fresh-water refugia (areas that escaped the great changes that occurred in the regions as a whole, providing conditions for relic colonies to survive). Because all species are found in such small areas and are dependent on the persistence of ground water, it could be advocated that they are critically endangered. *See* PERACARIDA. Gary C. B. Poore

Bibliography. I. Gordon, On *Spelaeogriphus*, a new cavernicolous crustacean from South Africa, *Bull. Brit. Mus. Nat. Hist. (Zool.)*, 5:31–47, 1957; J. R. Grindley and R. R. Hessler, The respiratory mechanism of *Spelaeogriphus* and its phylogenetic significance (Spelaeogriphacea), *Crustaceana*, 20:141–144, 1971; A. M. S. Pires, *Potiicoara brasiliensis*: A new genus and species of Spelaeogriphacea (Crustacea: Peracarida) from Brazil with a phylogenetic analysis of the Peracarida, *J. Nat. Hist.*, 21:225–238, 1987; G. C. B. Poore and W. F. Humphreys, First record of Spelaeogriphacea from Australasia: A new genus and species from an aquifer in the arid Pilbara of Western Australia, *Crustaceana*, 71:721–742, 1998; Y.-b. Shen et al., A new spelaeogriphacean (Crustacea: Peracarida) from the Upper Jurassic of China, *Contrib. Zool.*, 68:19–35, 1998.

Sperm cell

The male gamete. The typical sperm of most animals has a head containing the nucleus and acrosome, a middle piece with the mitochondria, and a tail with the 9 + 2 microtubule pattern (**Fig. 1**). Sperm, as well as the acrosome shape, varies with the species

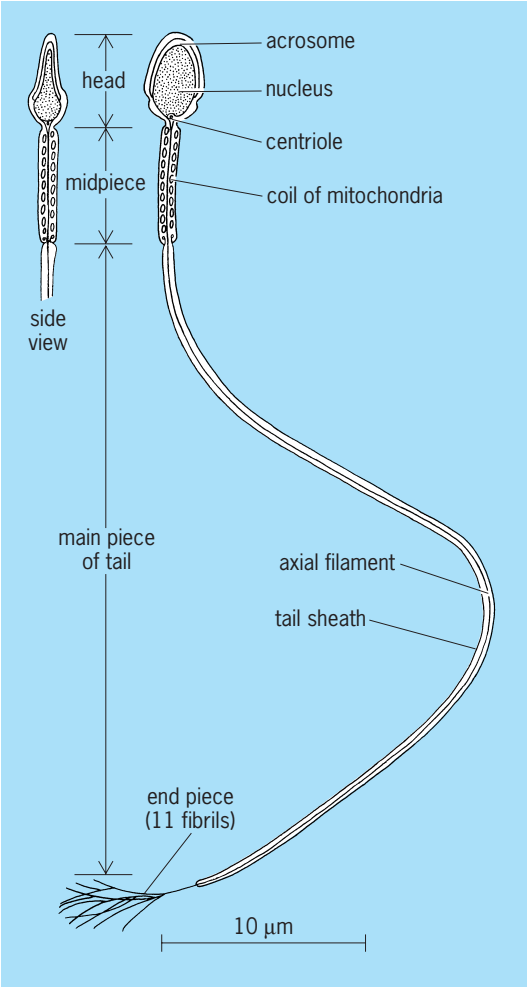

Fig. 1. Diagram of a human spermatozoon, based on electron micrographs.

(**Fig. 2**). The nucleus consists of condensed chromatin (deoxyribonucleic acid, DNA) and histone proteins. The acrosome, which is derived from the Golgi complex, contains hydrolytic enzymes, that is, hyaluronidase capable of lysing the egg coats at fertilization. Actin molecules which aid in the interaction between sperm and egg are found in the area between the acrosome and nucleus. The mitochondria in the middle piece apparently provide the energy necessary for the motility created by the tail. The tail has a central core, or axial filament, made up of nine double tubules and two central tubules. *See* CILIA AND FLAGELLA.

Many groups, including nematodes, myriapods, and crustaceans, have atypical sperm which lack a flagellum and are presumably nonmotile. The sperm of *Ascaris* is round and moves by ameboid means. The crustacean sperm have a large acrosome of several components. In the anomurans a middle with mitochondria and arms filled with microtubules precedes the nucleus. In the true crabs the nucleus forms arms which possess microtubules and also surrounds the many component acrosomes. The nucleus of crustacean sperm does not condense with

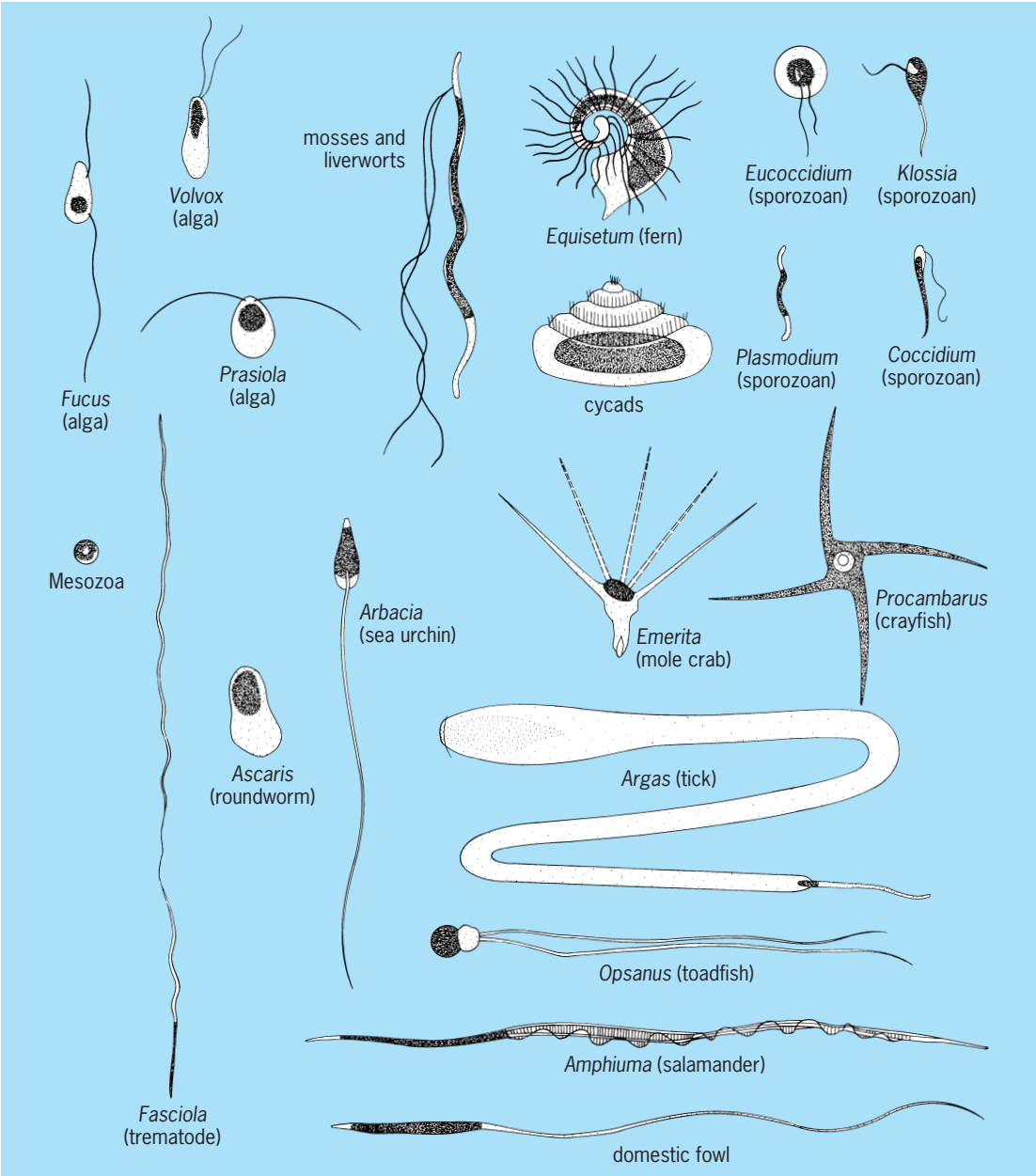

Fig. 2. Forms of spermatozoa. (*After C. R. Austin, Fertilization, Prentice-Hall, 1965*)

maturation as it does in typical sperm mentioned above. *See* SPERMATOGENESIS. Gertrude W. Hinsch

Spermatogenesis

The differentiation of spermatogonial cells into spermatozoa. This process occurs within the testes of the male (see **illus.**). The primordial germ cells are designated spermatogonia when the gonad is identifiable as a testis.

Spermatogonial divisions occur continuously throughout the life of mammals; these divisions both maintain the stem cell population (spermatogonial cells) and supply cells which develop into sperm. Clusters of spermatogonia maintain communication through cytoplasmic bridges, and these groups become primary spermatocytes when they synchronously enter the first meiotic prophase. The first meiotic prophase is characterized by a series of remarkable changes in chromosome morphology, which are identical to those seen in the corresponding stage of oogenesis. The secondary spermatocyte produced by this division then undergoes a division in which the chromosomes are not replicated; the resulting spermatids contain half the somatic number of chromosomes. *See* MEIOSIS.

The spermatids become embedded in the cytoplasm of Sertoli cells, and there undergo the distinctive changes which result in formation of spermatozoa. These morphological transformations include the conversion of the Golgi apparatus into the acrosome and progressive condensation of the chromatin in the nucleus. A centriole migrates to a position

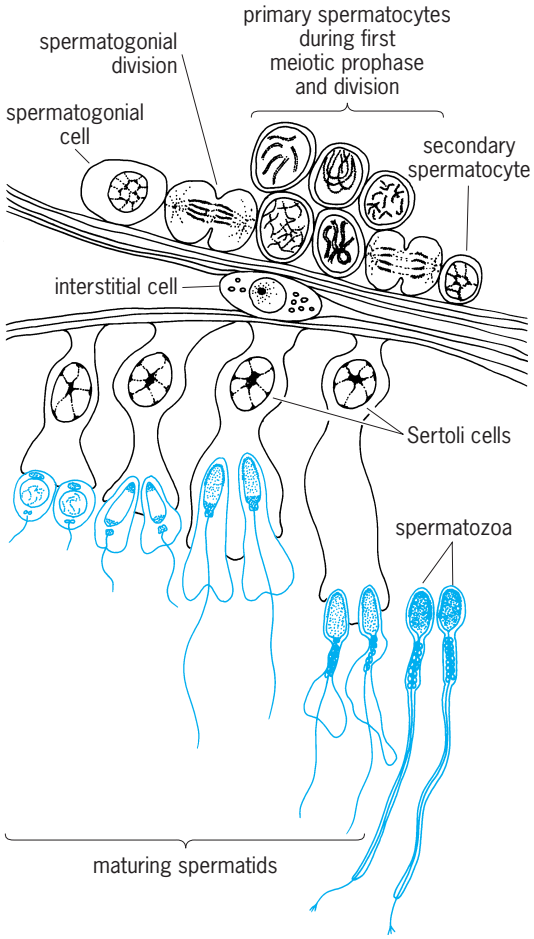

Cellular events in human spermatogenesis. Cell types are labeled.

distal to the nucleus and begins organizing the axial filament which will form the motile tail of the sperm. Mitochondria may fuse to form a nebenkern as is the case for many vertebrates, or there may be less extensive fusion as in mammals. In all cases the resulting structures become located around the axial filament in the midpiece. The cytoplasm of the spermatid is reflected distally away from the nucleus during spermatid maturation; eventually, most of the cytoplasm is sloughed off and discarded (see illus.).

The Sertoli cells are thought to provide nutrition for the developing sperm, because their cytoplasm contains large stores of glycogen which diminish as spermatids mature. There is no direct evidence for this nutritive function, but some forms of male sterility are associated with the failure to produce normal Sertoli cells. Electron microscopy has revealed distinct plasma membranes surrounding the two cell types at the points of contact, and thus the Sertoli cell–spermatid relationship is not syncytial as once thought.

Spermatogenesis is cyclical to a varying extent depending on the species, and under endocrine control. The periodic function of the human testis is not nearly as clearly definable as the menstrual cycle in the female, and can only be detected by examining deoxyribonucleic acid (DNA) synthesis and the changes in proportions of the various stages present. In seasonal-breeding animals, the testis regresses, and spermatogenesis ceases during the quiescent season. Spermatogenesis is maintained and regulated by male steroid hormones such as testosterone, which is produced by the interstitial or Leydig cells found in the connective tissue of the testis. Interstitial cells, in turn, are stimulated by luteinizing hormone (LH) which is produced by the pituitary gland. The male testis-regulating hormone was formerly known as interstital cell–stimulating hormone (ICSH), but it is now known to be identical to LH. *See* ENDOCRINE MECHANISMS; GAMETOGENESIS; MENSTRUATION; TESTIS.

Spencer J. Berry

Sphaeractinoidea

An extinct group of fossil organisms generally referred to as the Mesozoic stromatoporoids. As originally proposed by Othmar Kühn in 1927, the order included the Mesozoic stromatoporoids (Sphaeractiniidae), the heterastridiids (a group of obscure spherical Mesozoic fossils commonly regarded as hydrozoans), and the spongiomorphs. Kühn placed all of these in the class Hydrozoa. The discovery of spicules (spikelike supporting structures) in some genera of the sphaeractinids has allowed some genera formerly considered Mesozoic stromatoporoids to be classified in the Porifera as representatives of the class Demospongiae and the subclasses Ceractinomorpha and Tetractinomorpha. This discovery has opened the probability that all the sphaeractinids will eventually be referable to living sponge families and orders, if traces of spicules can be found in them. Both Mesozoic and Paleozoic stromatoporoids are now considered to be poriferans; the Paleozoic representatives are best considered a class of the phylum separated from the Mesozoic forms by time, lack of spicules, and microstructure. Those genera of the Mesozoic group that, at present, cannot be assigned to a living poriferan taxon owing to the lack of (or our failure as yet to discover) spicules are a disparate orphan group that can best be informally referred to as sphaeractinids or Mesozoic stromatoporoids.

These Mesozoic stromatoporoids appear in the late Permian with the genus *Circopora*, but did not reach abundance until the Jurassic and early Cretaceous. They declined rapidly in late Cretaceous time, and their presence in younger beds is in doubt. They grew large, calcareous, internally laminate skeletons in encrusting, dome-shaped, bulbous, fingerlike, and branched forms. Internally, the skeleton is composed of an irregularly repetitive three-dimensional network of longitudinal pillars or walls, tangential laminae with beams and cyst plates (see **illus.**) much like that of their Paleozoic look-alikes. In some genera, longitudinal canals crossed by tabulae are prominent. The microstructure of the structural elements is fibrous in well-preserved fossils and has suggested to some that they secreted aragonite rather than calcite. However, some living sponges secrete calcite in a similar microstructure, and it may be that the Mesozoic stromatoporoids did the same.

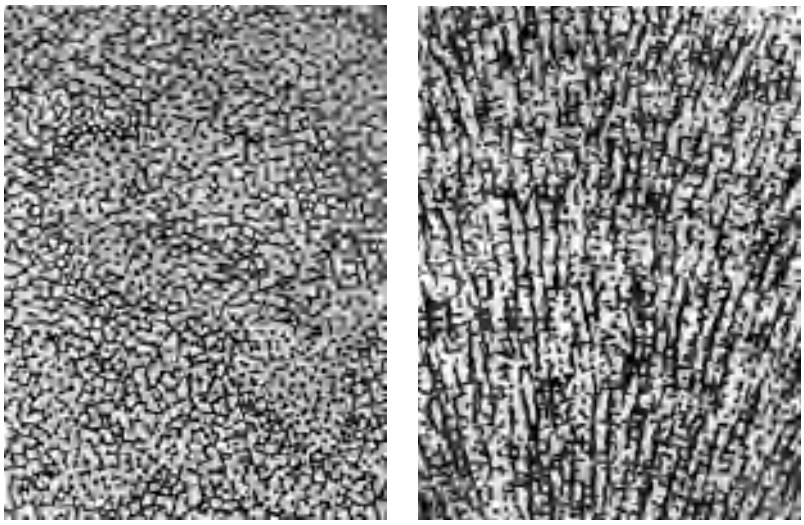

(a) (b)

Sphaeractinoid *Actinostromaria rhodaclada*, (*a*) tangential and (*b*) longitudinal section, ×6. (*From Steinor, 1932, Cretaceous, Jura Mountains*)

The Mesozoic stromatoporoids are important reef-building organisms in the Middle East, particularly in areas that surrounded the Mesozoic Tethys seaway. They are known from the Permian of Pakistan (Salt Range); the Triassic of the Caucasus, Bulgaria, and Indonesia; the Jurassic of Israel, Arabia, the Middle East–Mediterranean region, and North America (the eastern continental shelf); and the Lower Cretaceous of the Mediterranean region. *See* DEMOSPONGIAE; PORIFERA; REEF; SPONGIOMORPHIDA; STROMATOPOROIDEA.

C. W. Stearn

Bibliography. C. W. Stock et al., Annotated checklist of post-Devonian "Stromatoporoid genera" and their type species, *Fossil Cnidaria Porifera Newsl.*, 21(1.2):1–22, 1992; D. Turnšek and A. Barbulescu, Upper Jurassic hydrozoa in central Dubrogea (Romania), *Geologija*, 12:73–84, 1969; R. A. Wood, Biology and revised systematics of some late Mesozoic stromatoporoids, *Spec. Pap. Palaeontol.*, no. 37, pp. 1–89, 1987; R. A. Wood and J. Reitner, Poriferan affinities of Mesozoic stromatoporoids, *Palaeontology*, 29:469–478, 1986.

Sphaerocarpales

An order of liverworts in the subclass Marchantiidae, consisting of two families, the terrestrial Sphaerocarpaceae (*Sphaerocarpos* and *Geothallus*) and the aquatic Riellaceae (*Riella*). The plants are characterized by envelopes surrounding each antheridium and archegonium, absence of elaters, poor development of seta, and absence of thickenings in the unilayered wall of an indehiscent capsule.

The gametophytes consist of a simple, lobed thallus of rosette form, or a stemlike axis with lateral lobes or appendages. The rhizoids are smooth, and air chambers and pores are lacking. Oil bodies may be present. The antheridia are separately contained within receptacles on the dorsal surface or at the margins, while the archegonia are separately produced in receptacles dorsally or behind the growing point. The sporophyte consists of a bulbous foot, short or obsolescent seta, and globose indehiscent capsule with a unilayered wall. There are no elaters, but minute green cells are mingled with the spores. The haploid chromosome number is 8 or 9. *See* BRYOPHYTA; HEPATICOPSIDA; MARCHANTIIDAE.

Howard Crum

Sphagnopsida

A class of the plant division Bryophyta containing plants commonly called peatmosses. The spongelike plants (**Fig. 1**) grow as perennials in soft cushions or lawns in wet habitats (rarely they grow submerged). The class consists of a single genus, *Sphagnum*, of some 200 species. The thallose protonema, fascicled branches, dimorphous leaf cells, and spores developed from amphithecial tissue are characters unique to the class. The plants are ecologically important, owing to the spongelike construction of leaves and outer cells of stems and branches, and the ability, whether living or dead, to create acid conditions by exchanging hydrogen ions for cations in solution. The dried plant parts are used as a mulch in horticultural practice, and as fuel where the plant is abundant.

Peatmoss stems are simple or forked, with branches usually of two kinds, some spreading and others pendant, in fascicles spirally arranged around the stem and crowded in a headlike apex. Rhizoids (present only in early stages) are multicellular with slanted crosswalls. The stems and branches are enveloped in a cortex of large, empty, thin-walled hyaline cells, which often have pores to the outer

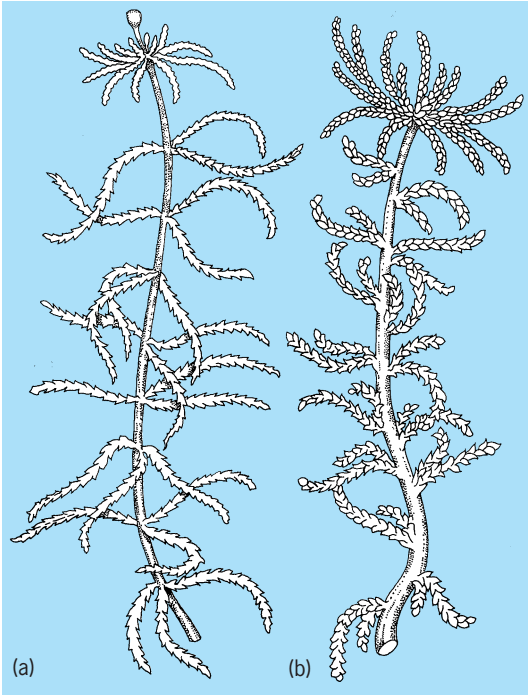

Fig. 1. Sphagnobrya. (*a*) *Sphagnum capillaceum*. (*b*) *Sphagnum palustre*. (*After E. L. Palmer, Fieldbook of Natural History, McGraw-Hill, 1949*)

surface. Internal to the cortex is a wood cylinder of long, thick-walled cells surrounding an inner tissue of more or less thin-walled and undifferentiated cells.

The stem and branch leaves are generally differentiated. The branch leaves consist of a network of linear green cells (chlorophyllose cells) enclosing large, empty hyaline cells that are nearly always reinforced internally by annular fibrils and perforated by pores on the free surface (**Fig. 2**). The globose, stalked antheridia occur among the leaves of somewhat differentiated spreading branches, and are often catkinlike and highly colored; one antheridium is produced at the side of each leaf. The archegonia are few and flask-shaped on very short branches and enclosed by large perichaetial leaves.

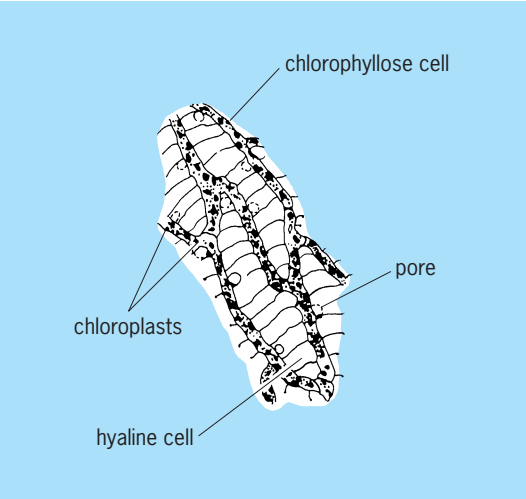

Fig. 2. Cellular structure of a leaf. (*After F. W. Emerson, Basic Botany, 2d ed., McGraw-Hill, 1954*)

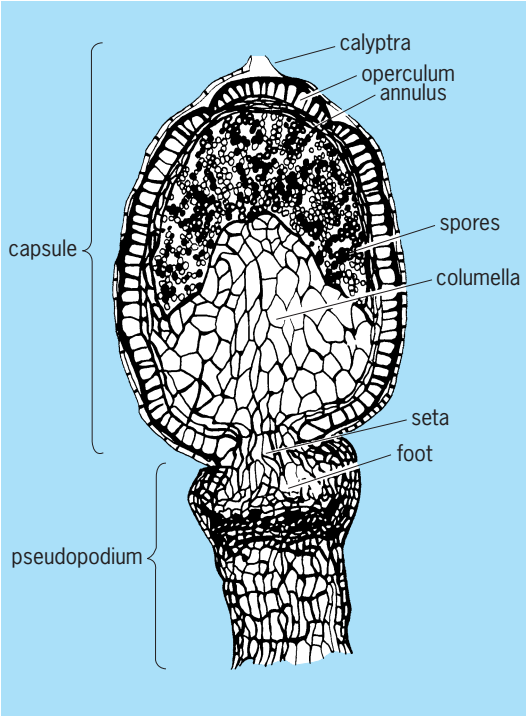

Fig. 3. Sporophyte of *Sphagnum*. (*After F. W. Emerson, Basic Botany, 2d ed., McGraw-Hill, 1954*)

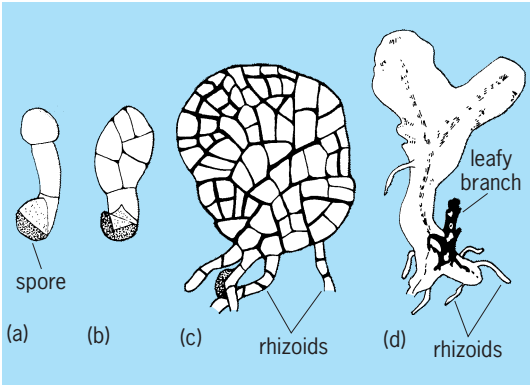

Fig. 4. Stages in the development of a protonema. (*a*) Germination of spore; (*b*) division of terminal cell to form an apical cell; (*c*) formation of multicellular rhizoids (*after A. Engler and K. Prantl, Die natürlichen Pflanzenfamilien, Bd. 10, 2 Aufl., 1924*). (*d*) Thalloid protonema with young, upright leafy branch forming (*after F. W. Emerson, Basic Botany, 2d ed., McGraw-Hill, 1954*).

The sporophytes are anchored by a foot, and are raised beyond the perichaetial leaves at maturity by an extension of the gametophyte called a pseudopodium (**Fig. 3**). The globose capsules dehisce by a lidlike operculum. The capsule wall consists of several layers of cells forming a solid tissue, with large numbers of nonfunctional pseudostomata (having two guard cells but no pore) scattered over the surface. The spores are derived from the amphithecium, or outer portion of the capsule. The spores are tetrahedral, with triradiate scars marking a tetrad origin. A membranous calyptra invests the capsule until its maturity. The protonema is normally a small, lobed thallus producing one or rarely two leafy plants but sometimes proliferating marginal filaments that produce at their tips secondary protonemata (**Fig. 4**). The haploid chromosome number is 19 and 38. *See* BRYOPHYTA.

Howard Crum

Bibliography. H. Crum, Sphagnopsida, Sphagnaceae, *No. Amer. Flora 2*, 11:1–180, 1984; B. Lange, Key to the northern boreal and arctic species of *Sphagnum*, based on characteristics of the stem leaves, *Lindbergia*, 8:1–29, 1982; D. H. Vitt and R. E. Andrus, The genus *Sphagnum* in Alberta, *Can. J. Bot.*, 55:331–357, 1977.

Sphalerite

A mineral, β-ZnS, also called blende. It is the low-temperature form and more common polymorph of ZnS. Pure β-ZnS on heating inverts to wurtzite, α-ZnS, at 1020°C (1868°F), but this temperature can be lowered substantially by impurity-atom solid solution (especially Cd^{2+} and Fe^{2+}) and sulfur fugacity. Sphalerite crystallizes in the hextetrahedral class of the isometric system with a structure similar to that of diamond. The space group is $F\bar{4}3m$, and the cubic unit cell has an edge $a = 0.543$ nanometer, which contains four ZnS molecules. Zinc atoms occupy the positions of half the carbon atoms of diamond, and sulfur atoms occupy the other half. Each zinc atom is bonded to four sulfur atoms, and each sulfur atom is bonded to four zinc atoms. The

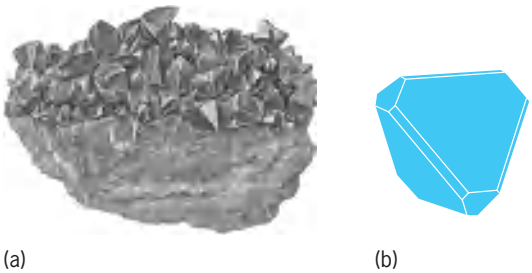

Sphalerite. (*a*) Crystals in limestone from Joplin, Missouri (*specimen from Department of Geology, Bryn Mawr College*). (*b*) Crystal habit (*after C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*).

common crystal forms of sphalerite are the tetrahedron, dodecahedron, and cube, but crystals are frequently complex and twinned (see **illus.**). The mineral is most commonly in coarse to fine, granular, cleavable masses. The luster is resinous to submetallic; the color is white when pure, but is commonly yellow, brown, or black, darkening with increased percentage of iron. It has been shown that excess sulfur can also contribute to the darkening of the color. There is perfect dodecahedral cleavage; the hardness is $3^1/_2$ on Mohs scale; specific gravity is 4.1 for pure sphalerite but decreases with increasing iron content.

Pure sphalerite contains 67% zinc and 33% sulfur, but iron is usually present, substituting in the structure for zinc and may amount to 36%. The amount of iron in sphalerite varies directly with temperature at the time of crystallization. Cadmium and manganese may be present in small amounts.

Sphalerite is a common and widely distributed mineral associated with galena, pyrite, marcasite, chalcopyrite, calcite, and dolomite. It occurs both in veins and in replacement deposits in limestones. As the chief ore mineral of zinc, sphalerite is mined on every continent. The United States is the largest producer, followed by Canada, Mexico, Russia, Australia, Peru, the Congo, and Poland. *See* ZINC.

Cornelius S. Hurlbut, Jr.; Paul B. Moore

Sphenisciformes

The penguins, a small monotypic order of flightless marine swimming birds found in the colder southern oceans (see **illustration**). They have most likely evolved from the order Procellariiformes, perhaps from a diving petrel-like ancestor. Classification schemes that hypothesize a link between penguins and loons have no factual support. At one time, the Sphenisciformes had been placed in a separate superorder, Impennes, but such a designation suggested too great a divergence of the penguins from their ancestral group of tube-nosed swimmers. Some taxonomists have enlarged the Ciconiiformes to include the Sphenisciformes penguins, as in the Sibley-Ahlquist taxonomy (a radical new approach to taxonomy, based on DNA hybridization studies). *See* CICONIIFORMES; GAVIIFORMES; PROCELLARIIFORMES.

Fossil record. The fossil record of the penguins begins in the late Eocene and continues throughout the Tertiary. Fossils have been found in Antarctica, Australia, and South Africa, which approximates the range of present-day penguins. The earliest fossil penguins were already well specialized as flightless, wing-propelled swimming birds, and their fossil history provides little new information about the evolutionary history and distribution of the group. Some fossil penguins were much larger than the emperor penguin, the largest living species. Some Eocene species were gigantic and had a long spearlike bill.

Characteristics. Penguins are medium-sized to large birds. They are completely flightless, their wings having been modified into stiff flattened flippers. They stand upright on legs that are placed far posterior and that terminate in four toes, the anterior three of which are webbed. Penguins swim and dive well, using only their wings for propulsion: their feet are used only for steering during swimming. Terrestrial locomotion is by walking, hopping, or sliding on the belly while pushing with the wings. The plumage consists of dense scalelike feathers that are black dorsally and white ventrally. A distinctive pattern or crest, often yellow, occurs on the head.

Penguins are gregarious, breeding in large colonies along the coast. The males and females, which are identical, form strong pair bonds and share in the incubation and care of the downy nestlings. The older young of some species are kept in large groups, or creches. The emperor penguin (*Aptenodytes forsteri*) breeds on the ice pack along the Antarctic coasts during the fall. Incubation of the one egg is the responsibility of the male, which remains on the nest for over 2 months in the winter without eating. The females leave the breeding colony during this period to feed, and return fat and laden with

***Aptenodytes patagonicus*, the king penguin. (*Photo by Gerald and Buff Corsi, © California Academy of Sciences*).**

food for the young. Males then leave for the ocean and return with food to relieve the females. This alternation continues until the young are old enough to enter the water. By this time, the sea ice has melted back close to site of the nesting colony.

Habitat. Penguins are found only in the cold southern oceans, on the Antarctic continent and its surrounding islands and northward to Australia, New Zealand, South America, and Africa. One species, the Galápagos penguins (*Spheniscus mendiculus*), is found in the Galápagos Islands, which are on the Equator but are surrounded by the cold Humboldt Current. *See* AVES. Walter J. Bock

Bibliography. L. S. David and M. Renner, *Penguins*, Yale University Press, 2004; L. Davis and J. Darby (eds.), *Penguin Biology*, Academic Press, 1990; T. D. Williams, *The Penguins*, Oxford University Press, 1995.

Sphenophyllales

An extinct order of articulate land plants, common during Late Pennsylvanian and Early Permian times. They are typified by *Sphenophyllum* (see **illus.**), a small, branching plant, probably of trailing habit. The long, jointed stems rarely exceeded 0.4 in. (1 cm) in diameter and had superposed, longitudinal, surficial ribs between nodes. The vascular system contained a solid xylem core with triangular primary wood. The leaves were wedge-shaped, usually shorter than 0.8 in. (2 cm), and had toothed, notched, or rounded distal margins. They were attached at the nodes by their narrow ends, in whorls of usually 6 or 9 leaves each, rarely 18. Long, terminal cones, usually called *Bowmanites* when found detached, contained sporangia and spores. The sporangia terminated slender stalks, forming concentric whorls that alternated with whorls of sterile bracts. Most species were homosporous (produced spores of a single type). *See* PALEOBOTANY; PLANT KINGDOM; SPHENOPHYTA.

Sergius H. Mamay

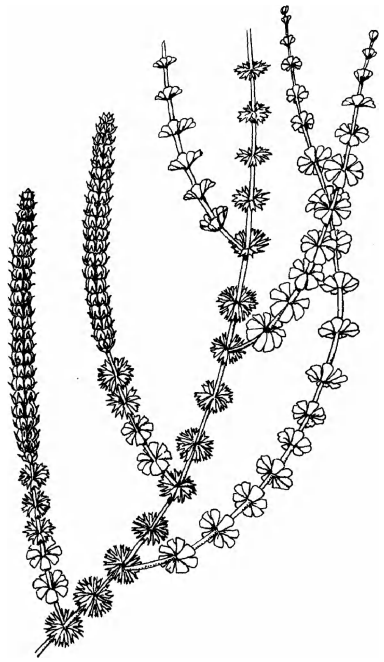

Reconstruction of *Sphenophyllum cuneifolium*, showing whorled leaves and two cones. (*After G. M. Smith, Cryptogamic Botany, vol. 2, 2d ed., McGraw-Hill, 1955*)

Sphenophyta

One of the major divisions (formerly known as Equisetophyta) of vascular plants that includes both living and fossil representatives. The three principal orders are Pseudoborniales (Devonian), Sphenophyllales (Devonian-Triassic), and Equisetales (Devonian-Recent); the Hyeniales (Devonian) may also be sphenophytes. *See* EQUISETALES; HYENIALES; PSEUDOBORNIALES; SPHENOPHYLLALES.

All sphenophytes are characterized by axes with distinct nodes that produce whorls of small leaves or branches; branches often contain longitudinal ribs and furrows. Internally the stems of sphenophytes are characterized by longitudinally oriented canals, some of which functioned in gaseous exchange. Secondary tissues were produced in a few fossil forms. The reproductive organs of this group are loosely arranged strobili or cones consisting of a central axis bearing whorls of modified branches, each terminating in a recurved, thick-walled sporangium. Most sphenophytes produced one type of spore (homospory), although a few fossil forms were heterosporous.

Pseudoborniales are known from a single species collected from a few Devonian localities south of Spitsbergen and in Alaska. Impression fossils suggest that the plant was a tree about 66 ft (20 m) tall, constructed of whorls of branches and leaves. The reproductive parts of *Pseudobornia* are not well known, but appear to contain recurved sporangia with one type of spore.

Sphenophyllales, less than 3 ft (1 m) tall, formed a portion of the Carboniferous forest understory. The leaves were often wedge shaped, and many were produced in a whorl. Some forms produced secondary tissues in the stems. All were homosporous, with the sporangia produced in cones.

During the Carboniferous, the sphenophytes were represented by numerous species in the order Equisetales, ranging from herbaceous forms to trees. Both homosporous and heterosporous forms were present, but the spores of all species contained extra exinous structures termed elaters. During the Mesozoic Period, the group declined, with fewer, smaller plants.

Today all that remains of the sphenophytes is the genus *Equisetum*, commonly called the horsetail or scouring rush. Except in Australia and New Zealand, the members are worldwide in distribution and typically grow in damp habitats along the edges of streams, although some species are adapted to mesic conditions. A few species attain considerable size, but none of these produce secondary tissues. All species are characterized by one type of spore (homospory), each bearing two pairs of elaters on the surface. *See* EMBRYOBIONTA. Thomas N. Taylor

Bibliography. T. N. Taylor and E. L. Taylor, *The Biology and Evolution of Fossil Plants*, 1993.

Sphere

Both in euclidean solid geometry and in common usage the word sphere denotes a solid of revolution obtained by revolving a semicircle of radius r about its diameter. Its total volume is $V = 4/3\pi r^3$.

However, in analytic geometry, and more generally in modern mathematics, the word sphere denotes a spherical surface that bounds a solid sphere. In this sense a sphere is the locus of all points P in three-dimensional space whose distance from a fixed point O (called the center) is equal to a given number. The word radius may refer either to one of the segments OP, or to their common length r. A plane that intersects a sphere in just one point is called a tangent plane and is perpendicular to the radius drawn from the center of the sphere to that point. A plane that intersects a sphere in more than one point intersects it in a circle. The circle is called a great circle or a small circle of the sphere according to whether the plane does or does not pass through the center of the sphere. If two parallel planes intersect a sphere (see **illus.**), the spherical surface between them is called a zone, and its area is $2\pi rh$, where the height h of the zone is the distance between the parallel planes. The volume bounded by the zone and the two base planes is called a spherical segment, and is measured by the formula $V = (\pi h/6)(3r_1^2 + 3r_2^2 + h^2)$, where r_1 and r_2 are the radii of the two base circles. It is also measured by the prismoidal formula $V = (h/6)(B_1 + 4M + B_2)$, where B_1, B_2 denote the areas of the two bases, and M the area of the midsection halfway between the bases.

Any great circle of a sphere divides it into two hemispheres. A second great circle cuts a hemisphere into two lunes, each having vertices at the two points where the great circles intersect and each having an area proportional to the angles at the vertices between the tangents to the great circle boundaries. A third great circle cuts each lune into two spherical triangles. In general a spherical triangle is the figure formed by connecting any three points A, B, C on the surface of a sphere (but not on the same great circle) by great circle arcs called sides. The sides BC, CA, AB are measured by the angles a, b, c that they subtend at the center of the sphere. If the angles between the pairs of sides at the vertices A, B, C are denoted by α, β, and γ, respectively, and if $2\sigma = \alpha + \beta + \gamma - 180^\circ$ denotes the excess of the sum of these angles over two right angles, then the area of a spherical triangle of excess 2σ on a sphere of radius r is $(\sigma/90^\circ)\pi r^2$ (or $2\sigma r^2$ if σ is measured in radians). Important relations between the sides and angles of a spherical triangle are as follows:

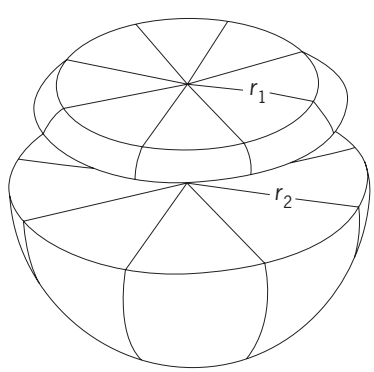

Segment of a sphere formed by intersection of two parallel planes with sphere.

Law of sines: $$\frac{\sin\alpha}{\sin a} = \frac{\sin\beta}{\sin b} = \frac{\sin\gamma}{\sin c}$$

Law of cosines:
$$\cos a = \cos b \cos c + \sin b \sin c \cos\alpha$$

J. Sutherland Frame

Spherical harmonics

A spherical harmonic or solid spherical harmonic of degree n is a homogeneous function, $R_n(x,y,z)$, of degree n which satisfies Laplace's equation (1).

$$\Delta R \equiv \frac{\partial^2 R}{\partial x^2} + \frac{\partial^2 R}{\partial y^2} + \frac{\partial^2 R}{\partial z^2} = 0 \qquad (1)$$

Here n is any number and

$$(x^2 + y^2 + z^2)^{(-n-1)/2}\, R_n(x, y, z)$$

is a spherical harmonic of degree $-n - 1$. There are analogous definitions for spaces of any number of dimensions. In the present article, n will be a nonnegative integer and R_n a polynomial in x,y,z (polynomial spherical harmonic). In terms of spherical coordinates r, θ, ϕ, $R_n(x,y,z) = r^n S_n(\theta,\phi)$ where S_n, a polynomial in $\cos\theta$, $\sin\theta$, $\cos\phi$, $\sin\phi$ is a spherical surface harmonic of degree n. There are $2n + 1$ linearly independent spherical surface harmonics of degree n; any spherical surface harmonic of degree n is a linear combination of these, and conversely any linear combination of spherical surface harmonics of degree n is again a spherical surface harmonic of degree n.

Applications. Spherical harmonics occur in potential theory. They occur in connection with Laplace's equation not only in spherical coordinates but also in spheroidal coordinates (spheroidal harmonics) and confocal coordinates (ellipsoidal surface harmonics). In the latter case their occurrence is due to the circumstance that the natural affine mapping of an ellipsoid onto a sphere carries the partial differential equation of ellipsoidal surface harmonics into the partial differential equation satisfied by spherical surface harmonics. In spherical coordinates, spherical surface harmonics occur in connection with Laplace's and Poisson's equations, the wave equation, the Schrödinger equation, and generally in connection with partial differential equations of the form $\Delta U + f(r)U = 0$. In the latter case one has special solutions of the form $F(r)S_n(\theta,\phi)$, where F satisfies the ordinary differential equation (2). In geometry,

$$\frac{d^2F}{dr^2} + \frac{2}{r}\frac{dF}{dr} + \left[f(r) - \frac{n(n+1)}{r^2}\right]F = 0 \qquad (2)$$

spherical surface harmonics are used in the theory of

surfaces. In mathematical physics, spherical harmonics appear in the theories of gravitation, electricity and magnetism, hydrodynamics, and in other fields.

Spherical harmonics of degree n. Notation (3), with

$$\int_{-\pi}^{\pi} (x\cos u + y\sin u + iz)^n f(u)\, du \tag{3}$$

$f(u)$ an integrable function, is a polynomial spherical harmonic of degree n, and every such spherical harmonic can be so represented. The representation is not unique. If c_n is a constant, $h_1, h_2, \ldots, h_n$ are n directions (not necessarily distinct), and $\partial/\partial h$ denotes directional differentiation in the direction h, then notation (4) is a polynomial spherical harmonic of de-

$$c_n r^{2n+1} \frac{\partial^n}{\partial h_1 \cdots \partial h_n} \frac{1}{r} \tag{4}$$

gree n, and every such spherical harmonic can be so represented. The representation is unique. In a zonal spherical harmonic the n directions coincide, and in a sectorial spherical harmonic they are in a plane at angles of π/n. If $n - m$ directions coincide in the axis and the remaining directions are in the plane perpendicular to the axis at angles of π/m, one has a tesseral spherical harmonic of degree n and order m.

Explicit forms. For spherical harmonics whose axis is the z axis ($m = 0, 1, 2, \ldots, n$), Eq. (5) represents

$$\begin{aligned} S_n^{\pm m}(\theta, \phi) &\equiv \frac{(-1)^{n-m} r^{n+1}}{(n-m)!} \frac{\partial^{n-m}}{\partial z^{n-m}} \left(\frac{\partial}{\partial x} \pm i \frac{\partial}{\partial y} \right)^m \frac{1}{r} \\ &= P_n^m(\cos\theta) e^{\pm im\phi} \end{aligned} \tag{5}$$

a linearly independent system of spherical surface harmonics of degree n. $S^m{}_n \pm S^{-m}{}_n$ is a zonal, sectorial, tesseral spherical surface harmonic of degree n and order m according as $m = 0$, $m = n$, $1 \leqslant m \leqslant n - 1$. The $P_n^m(w)$ are associated Legendre functions which satisfy the associated Legendre equation (6).

$$(1 - w^2)\frac{d^2P}{dw^2} - 2w\frac{dP}{dw} + \left[n(n+1) - \frac{m^2}{1 - w^2} \right] P = 0 \tag{6}$$

$P_n^0 = P_n$ is the Legendre polynomial of degree n.

Properties. A function is said to be harmonic in a region if it is a twice continuously differentiable solution of Laplace's equation there, and if, in addition, it vanishes at infinity in case the point at infinity is an interior point of the region. Every function harmonic inside a sphere about the origin can be expanded in a series

$$\sum_{n=0}^{\infty} r^n S_n(\theta, \phi)$$

convergent inside that sphere. Every function harmonic outside a sphere about the origin can be expanded in a series

$$\sum_{n=0}^{\infty} r^n S_n(\theta, \phi)$$

convergent outside that sphere. The reciprocal distance of the points (x,y,z) and $(0, 0,a)$ is harmonic in the regions $r < a$ and $r > a$ and possesses in these regions the expansions of Eqs. (7).

$$\frac{1}{\sqrt{a^2 - 2ar\cos\theta + r^2}} = \begin{cases} \displaystyle\sum_{n=0}^{\infty} \frac{r^n}{a^{n+1}} P_n(\cos\theta) & r < a \\ \displaystyle\sum_{n=0}^{\infty} \frac{a^n}{r^{n+1}} P_n(\cos\theta) & r > a \end{cases} \tag{7}$$

Here θ, ϕ determine a point on the unit sphere. The scalar product of two functions, f and g, on the unit sphere is suitably defined as Eq. (8) where $\overline{g}$

$$(f, g) = \int_0^{\pi} \int_{-\pi}^{\pi} f(\theta, \phi)\overline{g(\theta, \phi)} \sin\theta\, d\theta\, d\phi \tag{8}$$

is the complex conjugate of g. If $(f,g) = 0$, f and g are orthogonal. Spherical surface harmonics are functions on the unit sphere. Any two spherical surface harmonics of different degrees are orthogonal. The spherical surface harmonics of Eq. (9) form

$$S_n^m(\theta, \phi) m = -n, -n + 1, \ldots, n \tag{9}$$

$$n = 0, 1, \ldots$$

an orthogonal system; that is, $(S_n^m, S_{n'}^{m'}) = 0$ unless $m = m'$ and $n = n'$. This orthogonal system is complete; that is, a continuous function which is orthogonal to all the S_n^m vanishes identically. With an integrable function f is associated the Laplace expansion (10)

$$\sum_{n=0}^{\infty} \sum_{m=-n}^{\infty} c_{mn} S_n^m \tag{10}$$

where $$C_{mn} = \frac{\left(f, S_n^m\right)}{\left(S_n^m, S_n^m\right)}$$

Under suitable conditions, the Laplace expansion will converge to f. For instance, if f is continuous and continuously differentiable on the unit sphere, then the Laplace expansion converges to f uniformly.

Let (θ_0,ϕ_0) be a fixed point, and let $\cos\gamma = \cos\theta\cos\theta_0 + \sin\theta\sin\theta_0\cos(\phi - \phi_0)$ be the spherical distance of (θ,ϕ) and (θ_0,ϕ_0). The Laplace expansion of Eq. (11) is the addition theorem of Legendre

$$P_n(\cos\gamma) = P_n(\cos\theta)P_n(\cos\theta_0) + 2\sum_{m=-n}^{n} \frac{(n-m)!}{(n+m)!} \cdot P_n^m(\cos\theta)P_n^m(\cos\theta_0)\cos m(\phi - \phi_0) \tag{11}$$

polynomials and expresses the change to a new axis through the point (θ_0,ϕ_0). Other spherical surface harmonics have corresponding addition theorems.

Let $\cos\gamma$ be the spherical distance of (θ,ϕ) and (θ_0,ϕ_0), and let $k(w)$ be a continuous function for $-1 \leqslant w \leqslant 1$. Then for any spherical surface harmonic

S_n of degree n, Eq. (12) holds.

$$\int_0^{\pi}\int_{-\pi}^{\pi} K(\cos\gamma)\, S_n(\theta,\phi)\sin\theta\, d\theta\, d\phi = \lambda_n S_n(\theta_0,\phi_0) \quad (12)$$

where $$\lambda_n = 2\pi\int_{-1}^{1} K(w)P_n(w)\, dw$$

See DIFFERENTIAL EQUATION. A. Erdélyi

Bibliography. A. Erdélyi et al., *Higher Transcendental Functions*, 3 vols., 1953–1955, reprint 1981; E. W. Hobson, *Spherical and Ellipsoidal Harmonics*, 1955; M. N. Jones, *Spherical Harmonics and Tensors for Classical Field Theory*, 1985; E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 4th ed., 1927.

Sphingolipid

Any lipid containing the long-chain amino alcohol sphingosine (structure **1**) or a variation of it, such as dihydrosphingosine, phytosphingosine (structure **2**), or dehydrophytosphingosine. Sphingosine itself is synthesized by condensing a long-chain fatty acid with the amino acid serine.

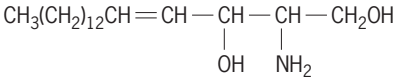

(1)

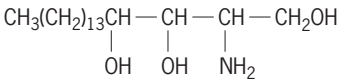

(2)

Sphingosine is converted into a variety of derivatives to form the family of sphingolipids. The simplest form is a ceramide which contains a sphingosine and a fatty acid residue joined by an amide linkage. The structure of the ceramide *N*-acylsphingosine is shown in structure **3**.

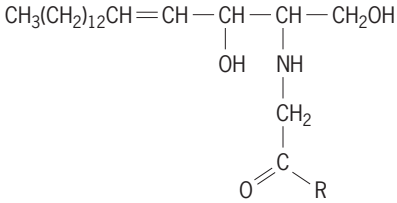

(3)

Ceramide is the basic building block of practically all of the naturally occurring sphingolipids. Ceramide can be further modified by the addition of a phosphorylcholine at the primary alcohol group to form sphingomyelin (structure **4**), a ubiquitous phospholipid in the plasma membranes of virtually all cells. Modification of a ceramide by addition of one or more sugars at the primary alcohol group converts it to a glycosphingolipid, which occurs widely in both the plant and animal kingdoms. The simplest form of a glycosphingolipid is a cerebroside which usually contains a glucose or a galactose residue joined to the ceramide moiety through a glycosidic linkage. The carbohydrate unit can be further extended to form more complex structures. A notable example of modification is the addition of one or more sialic acid residues to an existing carbohydrate chain to form the family of gangliosides which are particularly abundant in the nervous system. The structure of a typical ganglioside, GM1, is shown (structure **5**). *See* GLYCOSIDE; LIPID.

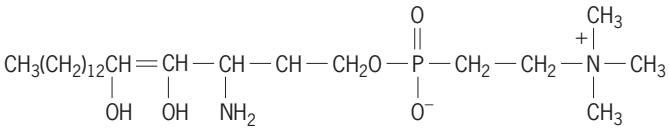

(4)

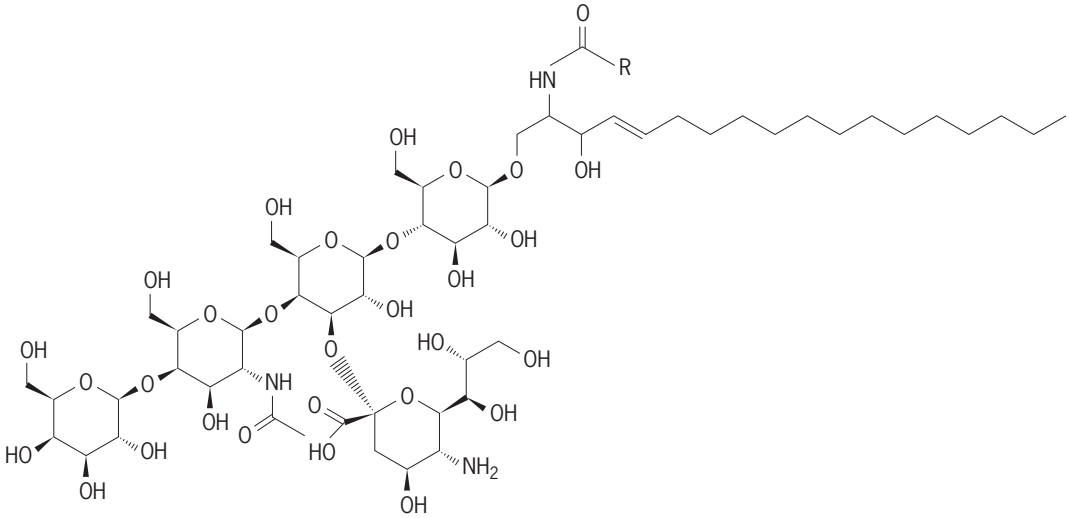

(5)

In addition to being important structural components of cellular membranes, sphingolipids participate in diverse cellular functions. A number of inheritable diseases that can cause severe mental retardation and early death occur as the result of a deficiency in one or more of the degradative enzymes, resulting in the accumulation of a particular sphingolipid in tissues. These diseases are collectively called sphingolipidoses and include Niemann-Pick disease, Gaucher disease, Krabbe disease, metachromatic leukodystrophy, and several forms of gangliosidoses, such as Tay-Sachs disease. Functionally, glycosphingolipids are known to serve as important cell-surface molecules for mediating cell-to-cell recognition, interaction, and adhesion. They also serve as receptors for a variety of bacterial and viral toxins. Many glycosphingolipids can modulate immune responses as well as the function of hormones and growth factors by transmitting signals from the exterior to the interior of the cell. A number of glycolipids are also found to participate in a variety of immunological disorders by serving as autoantigens. Other sphingolipids and their metabolites, notably ceramide, sphingosine-1-phosphate, sphingosylphosphorylcholine, and *N,N*- and *N,N,N*-methylsphingosine, may serve as second messengers in several signaling pathways that are important to cell survival or programmed cell death (apoptosis). *See* AUTOIMMUNITY; METABOLIC DISORDERS.

Robert K. Yu

Spica

A 1st-magnitude star in the constellation Virgo. Spica (α Virginis) is a hot main-sequence star of spectral type B1, effective temperature about 25,000 K (45,000°F), at a distance of 80 parsecs from the Sun (2.59×10^{15} km or 1.61×10^{15} mi). *See* SPECTRAL TYPE.

It is a double-line spectroscopic binary with an orbital period of almost exactly 4 days. The pair has been resolved with an intensity interferometer, which allows an accurate determination of the fundamental parameters of the system. The primary has a mass that is 10.8 times that of the Sun, and a radius of 8.1 solar radii; the mass and radius of the secondary are 6.8 and 4.4, respectively, in solar units. The system constitutes an example of an ellipsoidal variable, where the separation between the two stars is so small that powerful tidal forces distort the stars. The changing aspect of the primary as seen from Earth causes light variations with a period equal to the orbital period. The separation between the two stars is only about 3.5 times the radius of the primary. Both components are rotating around their axes very rapidly at projected equatorial velocities of 160 km/s (100 mi/s) and 100 km/s (60 mi/s), respectively. *See* BINARY STAR; ECLIPSING VARIABLE STARS; STELLAR ROTATION.

The primary itself is an intrinsic variable of the β Cephei type and is pulsating with a period of about 4 h. As a result, the observed light curve is a superposition of two different effects. There are also changes in the profiles of the spectral lines, which are related to the same mechanisms. *See* OCCULTATION; STAR; VARIABLE STAR.

David W. Latham

Spice and flavoring

Ingredients added to food to provide all or a part of the flavor. Spices, a unique category of flavorings which are given preferred legal status because of the long history of their use in foods, are pungent or aromatic substances of vegetable origin used in foods at levels that yield no significant nutritive value. Flavor is the perception of those characteristics of a substance taken orally that affect the senses of taste and olfaction. The term flavoring refers to a substance which may be a single chemical species or a blend of natural or synthetic chemicals whose primary purpose is to provide all or part of the particular flavor effect to any food or other product taken orally. Flavorings are categorized by source: animal, vegetable, mineral, and synthetic.

Chemical nature of flavor. The sensation of flavor has a chemical basis that involves both taste and smell.

Taste. The elements of taste consist of sourness, caused by hydrogen ions found in organic and inorganic acids; sweetness, caused by organic chemicals containing hydroxyl, carbonyl and amino groups, for example glycerol, sugars, amino acids, saccharin, hexamic acid salts; saltiness, a characteristic of sodium chloride, but also exhibited by other inorganic salts; and bitterness, which is caused by inorganic salts and some relatively high-molecular-weight organic compounds, such as alkaloids, for example, quinine, tannins, and sucrose octacetate. The taste-producing substances are usually nonvolatile.

Sensations of pain, heat, and coolness are also caused by chemicals. Several fatty acid amides of vanillylamine are responsible for the heat or bite in red peppers, while another amide, piperine, and its relatives provide the pungency of black pepper. Menthol produces a coolness at low concentrations in the mouth. The astringency that is produced by alum is also produced by tannins from tea and other botanicals.

Smell. The smell (odor) of a food is judged by the nose as volatile components of the food enter either through the nostrils or through the mouth. The chemical constituents which cause the odor component of flavor therefore are volatile. No simple characterization of basic odors has yet been found such as is available for tastes. Attempts to reduce the thousands of complex identifiable odors to a simple classification have been made, but have not been widely accepted.

Types of flavors. The application of the physical processes of distillation and extraction developed during the Middle Ages led to a wide range of essential oils, extracts, and oleoresins. Research involving the chemistry of natural products in the second half of the nineteenth century provided a nucleus

TABLE 1. Volatiles identified in flavorful foods*

Type	Hydrocarbons	Alcohols	Aldehydes and ketones	Acids	Esters	Bases	Sulfur compounds	Miscellaneous	Total
Apple	8	40	39	39	80	4	—	11	221
Cheddar cheese	32	11	25	28	25	14	6	12	153
Cinnamon	24	9	10	15	8	—	—	15	81
Cocoa	45	27	52	37	38	103	18	43	383
Onions and garlic	1	5	15	2	5	4	59	8	99
Potato	24	24	31	13	1	9	23	9	134
Tomato	17	31	90	12	28	4	8	9	217
Cooked beef	74	32	76	22	25	25	45	30	329
Vanilla	37	24	21	15	39	2	1	23	162
Roasted peanuts	37	25	72	29	11	63	16	25	278

*From S. van Straten (ed.), *Volatile Compounds in Food,* 4th ed., 1977, and supplements.

of chemicals for use as flavoring ingredients. The reactions studied at that time (Knoevenagel condensation, Knorr pyrrole synthesis, Perkins reactions and rearrangements, Kolbe-Schmitt reaction) produce the characterizing ingredients of many spices and fruits. Spices contain the largest concentration of flavor volatiles and were the first to be studied. This early work confirmed the presence of, and the structure of, vanillin in vanilla beans at 2% concentration, eugenol in clove buds at 12%, and cinnamic aldehyde in cinnamon at 1–2% concentration.

The first half of the twentieth century brought a continuous effort to identify the flavorful volatiles in food by classical organic chemical methods. The many ingredients in citrus peel oils and noncitrus fruits began to yield their flavor secrets. Creative flavorists utilized this knowledge and filled in the gaps with information supplied by their noses and their palates. Reasonably representative flavors of the fruits became possible to produce. These fruits contained volatiles which effected their flavor at concentrations of 0.01–0.1% (100–1000 ppm).

The second half of the twentieth century, especially since the application of gas-liquid chromatography, has seen an explosion of information concerning the identity of the volatile components of all flavorful foods. The sensitivity of equipment for detecting ingredients and the speed with which equipment can identify these ingredients has confirmed the complex nature of flavor. Many of the volatiles in the less highly flavored foods have been examined, isolated, and identified. As many as several hundred volatile organic chemicals have been identified in common flavorful foods (**Table 1**). These components range from simple hydrocarbons to complex chemicals with sulfur or nitrogen (**Table 2**). While many of these chemicals are not present in sufficient quantities to affect the flavor and some key ingredients have not yet been identified, flavorists have used this information in attempts to prepare artificial flavors.

A so-called nature-identical chemical is self-defining. If the flavor of foods were due to stable chemicals, flavors could be made from natural-identical chemicals alone. But the natural flavors of strawberries, grilled hamburger, or roast beef are not stable. Research involving development of synthetic shelf-stable flavors continues. Many of the types of flavors available are listed in **Table 3**.

Classification. There are several ways of dividing flavorings into subgroups other than source. They can be divided into two groups; one group affects primarily the sense of taste, and the other affects primarily the sense of olfaction. The members of the first group are called seasonings, the members of the second group are called flavors.

The same terms can be used to divide flavorings in another way. The term flavors can be applied to those products which provide a characterizing flavor to a food or beverage. The term seasoning can then be applied to those products which modify or enhance the flavor of a food or beverage—spices added to meats, blends added to potato chips, lemon added to apple pie.

Flavorings can also be classified by their physical form, as solids, liquids, or pastes. The form will be determined by the nature of the flavoring, and the form of food to be flavored.

TABLE 2. Flavor components of various types of flavorful foods

Type	Component	Species	Total flavor components, ppm
Cooked and processed meats	Amino acids, sugars	Sulfides, pyrazines, thiazoles, furanones	50
Fish	Amino acids	Amines, substituted phenols	50
Nuts	Texture	Pyrazines, lactones, substituted phenols	50
Vegetables	Starch, sugars	Unsaturated aldehydes, sulfur compounds	10–100
Dairy	Fats, acids	Acids, ketones, esters	100–1000
Fermented	Alcohols, acids	Alcohols, aliphatic esters	100–1000
Fruits	Sucrose, citric acid	Esters of aliphatic acids, aliphatic aldehydes, aromatic aldehydes, terpenes	100–1000
Spices	Amides	Terpenes, phenols, aldehydes, ketones, ethers	2000–200,000

TABLE 3. Examples of artificial flavors and their uses

Type	Example	Use
Cooked and processed meat	Beef	Gravy mix
	Bacon	Bacon bit or analog, snack
Seafood	Shrimp	Snack cracker
Nut	Walnut	Toppings, cake mixes
	Peanut	Cookies, candies
Vegetable	Tomato	Sauces
	Cucumber	Dressings
Dairy	Butter	Margarine, baked goods
	Cheese	Snacks, analogs
Fermented	Vinegar	Salad dressings
	Bread	Chemically raised bread
	Beer and wine	Sauces
	Rum	Baked goods
Fruits	Lemon	Candy, baked goods
	Apple	Baked goods, chewing gum
	Strawberry	Candy, brandy
	Papaya	Noncarbonated beverages
Spices	Anise	Baked goods, candy
	Dill	Pickles, sauces
	Nutmeg	Sauces, condiments, processed meats

In the United States most flavorings are processed, since undried herbs and fresh fruits for desired flavor effects are available only on a small scale. Therefore another classification is by method of manufacture. There is general agreement on the several commercial designations of flavorings, indicating the method of manufacture of each and the type of flavors which are available by using ingredients from each category.

Chemicals. These are produced by chemical synthesis or physical isolation and include such substances as vanillin, salt, monosodium glutamate, citric acid, and menthol. Artificial or natural, they are only occasionally used alone, and usually are compounded or blended with other ingredients or diluted to achieve a specific flavor character at a workable concentration. *See* SALT (FOOD).

Concentrates. Powders are manufactured by dehydrating vegetables, such as onions and garlic, or concentrated fruit juices. These expensive natural products are available only in limited quantities.

Condiments. These are single ingredients or blends of flavorful (sometimes exotic) foods, spices, and seasonings, some of which may have been derived by fermentation, enzyme action, roasting, or heating. They are usually designed to be added to prepared food at the table (for example, chutney, vinegar, soy sauce, prepared mustard). Some are completely natural, while others are a blend of natural and synthetic ingredients.

Spices. These are natural substances that have been dehydrated and consist of whole or ground aromatic and pungent parts of plants, for example, anise, cinnamon, dill, nutmeg, and pepper. Such spices are also classified as herbs if grown in temperate climates (**Table 4**).

Extracts. These are natural substances produced by extraction from solutions of the sapid constituents of spices and other botanicals in food-grade solvents. They are also available as synthetics. Extracts are used in home food preparation, and industrially by beverage manufacturers. Examples include vanilla, lemon, kola nut, and coffee.

Oleoresins. These are the natural volatile and nonvolatile flavor constituents of spices and other botanicals produced by volatile solvent extraction and subsequent removal of the solvent by distillation. They are used industrially in processed meat, canning, and baking industries.

Essential oils. These are the volatile oils obtained from spices and other plant materials which have the characteristic odors of their sources. They are usually obtained by distillation, although the citrus peel oils are obtained by expression. Essential oils are used in confectionery and beverage industries, and in compounded flavors for many applications. *See* ESSENTIAL OILS.

Hydrolyzed vegetable proteins. These flavors are produced by optimizing the development of a basic meaty flavor through hydrolysis of various vegetable proteins. They are used in soup formulations and in compounding flavors and seasonings.

Process flavors. These flavors are natural or artificial products of roasting, heating, fermentation, or enzymolysis of foods or normal food ingredients.

Compounded or blended flavorings. These are a group of substances produced by mixing.

Seasonings. Available in dry or liquid form, these are natural or artificial flavorings produced by the blending of various categories of flavoring agents with flavorful foods and flavoring adjuncts, for example, free-flow agents and antioxidants. They are suitable for altering the flavor of foods to which they are added, with the emphasis on the taste of the food; for example, barbecue seasoning for potato chips or chili seasoning for beans.

Flavors. These are natural or artificial flavorings produced by the compounding of various categories of flavoring ingredients and flavoring adjuncts (for example, solvents, antioxidants, and preservatives) suitable for imparting, modifying, or augmenting the flavor of foods to which they are added, with the emphasis on smell (odor, aroma), for example, artificial strawberry flavor for a gelatin dessert, natural orange flavor for a breakfast drink mix, artificial bacon flavor for a flavored dairy dip, or artificial meat flavor for a hamburger extender. Further processing before or after compounding can provide these flavors in dry or liquid form. They may be water- or oil-soluble, and are usually designed to be suitable for specific food products and processes. *See* FOOD ENGINEERING; FOOD MANUFACTURING.

Regulations. Perhaps because spices were at one time used to mask deterioration of foods, and perhaps because other flavorings were at one time used to hide economic adulterations, a complex, not

TABLE 4. Economically important spices

Spice	Plant and source	Plant part used	Principal use
Allspice* (*Pimenta*)	Evergreen of myrtle family from Jamaica and Guatemala	Dried fruit	Pickles, roast meat, ketchup
Anise*	Parsley family from Mediterranean	Seed	Baked goods, anisette (a liqueur)
Basil	Mint family grown in U.S. and other parts of world	Leaves and tender stems	Sauces and soups, especially tomato-based
Bay	Evergreen of laurel family from Mediterranean, Turkey, Greece, and Portugal	Leaf	Pickles, stews, soups, sauces
Caraway*	Biennial of parsley family from north-central Europe and southern England	Seed	Bread and baked goods, cheese, sauces, kummel (a liqueur)
Cardamon*	Ginger family from Guatemala and India	Seed	Baked goods, coffee blends, curry powders
Cayenne	Capsicum family grown over most of the world (name used for those high in pungency)	Pod or fruit	Many foods, pickles, sauces, meats, curry powder
Celery*	Parsley family principally from France and India	Seed	Sauces, salads, pickles, soups
Cinnamon*	Evergreen member of laurel family from Ceylon	Bark	Baked goods, pickles, candy
Clove*	Evergreen tree of the myrtle family from Madagascar and Zanzibar	Unopened bud flower	Pork products, pickles, stews, meats, and gravies
Coriander*	Plant of parsley family from former Yugoslavia and Morocco	Dried fruit	Frankfurters, bologna, baked goods
Cumin*	Annual plant of parsley family from Iran and Morocco	Dried fruit	Chili powder
Dill*	Member of parsley family from India and domestic sources	Seed and leaves	Pickles, soups, sauces
Fennel*	Perennial of parsley family from India and Rumania	Seed	Baked goods, salad dressings, meat products
Garlic*	Member of lily family from U.S. and Mediterranean	Bulb	Baked meats, sauces, dressings, soups
Ginger*	Tuberous perennial from Jamaica, India, and Africa	Rhizome	Soft drinks, baked goods, pickles, puddings
Mace	Tropical tree similar to rhododendron from Indonesia and West Indies	Aril covering nutmeg	Baked goods, processed meats
Marjoram	Perennial of mint family from France, Peru, and Chile	Leaves	Soups, stews, sauces
Mint	Perennial herb of many varieties grown in U.S.	Leaves	Confections, sauces, jellies, gums
Mustard*	Annual of mustard family from U.S. in Montana and California	Seeds	Sauces, baked meats, processed meats and gravies
Nutmeg*	Tropical tree similar to rhododendron from Indonesia or West Indies	Seed	Baked goods, processed meats, meat, fruit sauces
Oregano	Perennial of mint family from U.S., Mexico, Italy, and France	Leaves	Sauces and stews, especially tomato-based
Paprika*	Member of capsicum family from Spain, U.S., and Hungary	Pod or fruit	Red-colored dressings, sauces, meats, condiments as ketchup
Pepper* black white	Perennial climbing vine from India and Indonesia	Berry (black); berry with cortex removed (white)	Virtually all food products, table use
Poppy seed	Annual of poppy family from Poland, Argentina, Iran, and Turkey	Seed	Toppings for baked goods and confections
Rosemary	Evergreen of mint family from France, Spain, and Portugal	Leaf	Meats, sauces, gravies
Saffron*	Plant of crocus family from Spain	Stigma of flower	Coloring rice and other specialties
Sage*	Member of mint family from Greece and former Yugoslavia	Leaf	Sausage, poultry, poultry stuffing
Savory	Member of mint family from France and Spain	Leaf	Meats, stuffings, salads, sauces
Sesame	From sesame plant in U.S., Nicaragua, Salvador, Egypt, and Brazil	Seed	Baked goods and confections
Thyme	Perennial of mint family from France and U.S.	Leaf	Sauces for shellfish and with fresh tomato
Turmeric*	Member of ginger family from India, Haiti, Jamaica, and Peru	Rhizome	Yellow color in pickles, sauces, and fish

*See individual articles for further information.

always logical, series of laws and regulations has been enacted in the United States and most other developed countries to prevent the use of flavorings in a deceptive manner. The basis of the laws and regulations varies from country to country, causing problems in international trade. The *Codex Alimentarius* represents an effort to standardize the composition and identification of many major foods and food ingredients. It is jointly sponsored by the United Nations Food and Agriculture Organization (FAO) and the World Health Organization (WHO) with support from other groups such as the International Organization of Flavor Industries (IOFI).

Laws and definitions differ widely from country to country. In the United States the addition of spices and natural and artificial flavors to food (including beverages and chewing gum) is permitted by law and governed by regulations. These regulations are

included in the Code of Federal Regulations. Natural flavors are those which contain only natural flavoring ingredients; artificial flavors are those which contain at least some artificial flavoring ingredients. Natural flavoring ingredients include essential oils, oleoresins, essences or extractives, protein hydrolysates, distillates or any products of roasting, heating or enzymolysis which contain the flavoring constituents derived from spices, fruits, or fruit juices, vegetables or vegetable juices, edible yeast, herbs, barks, buds, roots, leaves or similar plant materials, meats, seafood, poultry, eggs, dairy products, or fermentation products whose significant function is flavoring rather than nutritive. An artificial flavor is any substance whose function is to impart flavor which is not derived as a defined natural flavor ingredient. In addition, any flavor ingredient—natural or artificial—must either be generally recognized as safe (GRAS) or covered by a regulation of the United States Food and Drug Administration. Earl J. Merwin

Bibliography. K. Bauer and D. Garbo (eds.), *Common Fragrance and Flavor Materials*, 3d ed., 1997; D. B. Birch and M. G. Lindley (eds.), *Developments in Food Flavours*, 1987; H. B. Heath and G. A. Reineccius, *Flavor Chemistry and Technology,* 1986; J. K. Jank, *Spices*: *Their Botanical Origin*, *Their Chemical Composition*, *Their Commercial Use*, *Including Seeds*, *Herbs*, *and Leaves*, 1980; D. B. Min (ed.), *Flavor Chemistry of Lipid Foods,* 1989; I. D. Morton and A. J. MacLeod (eds.), *Food Flavours*, pt. A: *Introduction*, 1982, pt. B: *The Flavour of Beverages*, 1986; R. Taranishi, R. G. Buttery, and F. Shadidi, *Flavor Chemistry: Trends and Developments*, 1989.

Spider silk

Spider silks are fine fibers produced from fibrous-protein solutions for use by spiders in the various functions required for their survival. Spiders are the only animals that use silk for a variety of functions in their daily lives, including reproduction, food capture, and construction. Silk fibers produced by silkworms to build their cocoons have been used in the textile industry for more than 5000 years. But in recent years, spider silk, particularly silk from the major ampullate glands used to make elaborate "orb" webs, has attracted the attention of scientists and engineers in many fields from evolutionary biology to mechanical engineering and material science. The exceptional toughness, stiffness, and extensibility of spider-silk fibers outperform even the best synthetic materials. With the recent and expected advances in molecular genetics and proteomics, as well as in the understanding of material properties in the nanoscale size range, orb-web-weaving spiders are ideal organisms for studying the relationship among the physical properties of protein-based materials in terms of their function and evolution. *See* ARANEAE; FIBROUS PROTEIN; NATURAL FIBER; SILK.

Mechanical properties. Over the past 400 million years, many different kinds of silk have evolved in the nearly 37,000 known spider species. For every kind of spider silk that has been characterized in any detail, there are more than 1000 that have not been characterized. Silk from only a few species, such as *Nephila clavipes* (the golden orb-weaving spider) and *Araneus diadematus* (common garden spider), have been examined in detail. Orb-weaving spiders produce silk for different functions from as many as seven different glands (see **illustration** and **Table 1**). Dragline and viscid silk fibers have been the focus of recent research because they are among the toughest materials known. Toughness is a measure of a material's energy absorption capacity. A single silk fiber is only 1–10 micrometers in diameter, compared to a human hair which typically is close to 50 μm. Dragline silk will support the spider's weight and is used to make the radial threads in webs. Viscid silk is the sticky, spiral threads in webs used for capturing prey. **Table 2** lists the characteristic mechanical properties of dragline and viscid silk, compared to other materials. Most of the other materials are either very strong but not tough and extensible (such as steel) or are highly extensible but not strong (such as rubber). The extraordinary characteristic of dragline silk is that it has an exceptional combination of high stiffness, strength, extensibility, and toughness. The breaking stress or strength of dragline silk is comparable to steel, Kevlar® (used in bullet-proof clothing), and carbon fibers, but it is more extensible and tougher. Viscid silk is even more extensible, with stiffness comparable to rubber, but it is much stronger than rubber. The high stiffness of dragline and the rubberlike extensibility of viscid silk with the exceptional ability to absorb energy appear to be optimized for their required functions. Scientists seek to understand how spiders are able to spin silk fibers from an aqueous protein solution with the desired properties. Using spider silk as a guide, it should be possible to make the next-generation high-performance (very strong and yet very extensible) materials. *See* STRENGTH OF MATERIALS; STRESS AND STRAIN; YOUNG'S MODULUS.

Structure. All types of silk have the same basic building blocks—protein polymers made of amino acids. Although the different silk glands in an evolutionarily advanced orb-weaving spider, such as *N. clavipes*, possibly evolved from a single type of gland, silk glands now differ significantly from one another in morphology and composition, and produce different kinds of silk. The best-characterized silk protein and associated fiber is the dragline silk of *N. clavipes*. Dragline-silk fibers are composed primarily of two protein polymers: spidroin I and spidroin II. Both spidroin I and II consist of repeated alternating protein sequences rich in alanine (Ala) and glycine (Gly) amino acids. These proteins are stored in the glands as concentrated solutions (30–50 wt %). A combination of physicochemical and rheological processes (including ion exchange and elongational flow) cause the protein molecules to self-assemble and transform from an unaggregated and soluble state in the gland, through an intermediate liquid-crystalline state, into an insoluble fiber, while passing through the spinning canal.

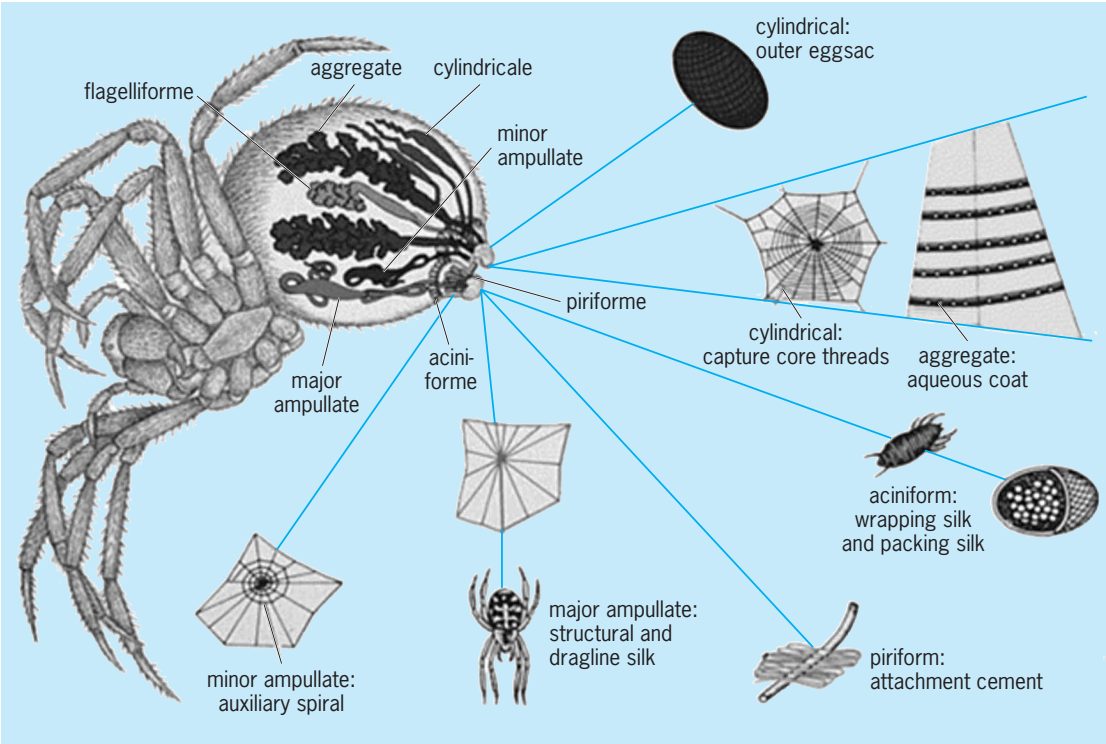

Silk fibers and silk glands in *Araneus diadematus*. (*Reprinted from F. Vollrath, Strength and structure, Rev. Mol. Biotechnol., 74(2):67–83, 2000*)

After spinning, the dry silk fibers consist of stiff nanometer-scale crystals embedded in a softer protein matrix. The nanocrystals are formed by the aggregation of hydrophobic polyalanine regions into a highly ordered crystalline β-sheet via self-assembly, while the glycine-rich protein regions form the softer rubbery matrix. This combination of hard crystallites embedded in a soft matrix is considered the primary reason for the exceptional mechanical strength and extensibility of dragline fibers. It is still not clear how spiders and silkworms prevent and regulate the formation of β-sheet aggregates during the spinning process, but understanding of the mechanism is of great relevance for protein folding. An understanding of this mechanism could help in developing therapies for protein-folding diseases, such as Alzheimer's, which result from the uncontrolled aggregation of proteins into β-sheet aggregates known as amyloid plaques. *See* AMINO ACIDS; POLYMER; PROTEIN.

Mimicking spider silk. Although spidroin I and spidroin II have been sequenced, a complete understanding of the various physicochemical processes during the spinning of silk fibers is still lacking. Before we can completely mimic spider silk, we must fully understand the biochemistry, processing strategy, and molecular basis of the mechanical properties of silk. Advances are being made on different fronts to emulate different aspects of spider silk. Three main approaches are being pursued to make materials with mechanical properties comparable to dragline silk. The first approach is to use recombinant DNA methods to produce proteins with structure similar to spidroin I and II in other organisms, including potato, tobacco, *Escherichia coli*, and goat's milk. The second approach is to synthesize high-molecular-weight block copolymers that mimic the molecular structure of silk proteins. The third approach is to develop strong, yet extensible synthetic composites by reinforcing elastomeric polymers with inorganic nanoparticles. To date, these approaches have had limited success. One resulting application is the spinning of fibers from recombinant silk proteins.

TABLE 1. Types of common spider silk and their function

Silk	Gland	Function
Dragline	Major ampullate	Radial threads of web frame safety line
Viscid	Flagelliform	Prey capture
Gluelike	Aggregate	Prey capture, attachment
Minor ampullate	Minor ampullate	Orb-web frame construction
Cocoon	Cylindriform	Reproduction
Wrapping	Aciniform	Wrapping captured prey
Attachment	Pyriform	Attachment to environmental substrates

TABLE 2. Tensile mechanical properties of spider silk and other materials

Material*	Stiffness or Young's modulus, GPa (kilo psi)	Strength or breaking stress, GPa (kilo psi)	Extensibility or maximum strain, %	Toughness or maximum energy absorption, MJ/m^3 (Btu/in.^{3})
AD dragline silk	10 (1450)	1.1 (159)	27	160 (2.48)
AD viscid silk	0.003 (0.43)	0.5 (72)	270	150 (2.32)
NC dragline silk	22 (3190)	1.1 (160)	9	150 (2.32)
NE dragline	8 (1160)	1.1 (160)	28	165 (2.56)
Silkworm silk	7 (1015)	0.6 (87)	18	70 (1.08)
Tendon, collagen	1.5 (217)	0.15 (22)	12	7.5 (0.12)
Bone	20 (2900)	0.16 (23)	3	4 (0.06)
Elastin	0.001 (0.145)	0.002 (0.29)	150	2 (0.03)
Synthetic rubber	0.001 (0.145)	0.05 (7.25)	850	100 (1.55)
Nylon fiber	5 (725)	0.95 (137)	18	80 (1.24)
Kevlar 49	130 (18850)	3.6 (522)	2.7	50 (0.77)
Carbon fiber	300 (43500)	4 (580)	1.3	25 (0.39)
High tensile steel	200 (29000)	1.5 (217)	0.8	6 (0.09)

*AD: *Araneus diadematus*; NC: *Nephila clavipes*; NE: *Nephila edulis*.

In addition to the material aspects, the silk-spinning process is far superior and more sophisticated than industrial spinning processes for making polymeric fibers. Conventional industrial processes, such as wet spinning or dry spinning, operate only under very controlled conditions during which the polymer filaments are rapidly solidified or "quenched" to provide sufficient mechanical strength to be drawn into fibers. Spiders are able to spin silk fibers under widely different ambient conditions, and the silk filaments rarely break during spinning. Even in the wet state during spinning, the silk filaments are strong enough to support the weight of the spider. Microscopic analysis shows that the fibers have radially graded mechanical properties, with aligned, stiffer molecules surrounding a softer elastic core. Spiders also produce different silk fibers using the same apparatus through highly regulated protein self-assembly. A complete understanding of spiders' spinning processes would help in developing new technologies that produce nanometer- and micrometer-scale fibers from self-assembling polymers and nanocomposite materials. *See* ELASTICITY; MANUFACTURED FIBER; NANOSTRUCTURE.

Nitin Kumar; Gareth H. McKinley

Bibliography. C. L. Craig, *Spiderwebs and Silk: Tracing Evolution from Molecules to Gene to Phenotypes*, Oxford University Press, New York, 2003; J. Gosline et al., The mechanical design of spider silks: From fibroin sequence to mechanical function, *J. Exp. Biol.*, 202:3295–3303, 1999; D. Kaplan et al., *Silk Polymers: Material Sciences and Biotechnology*, American Chemical Society, Washington DC, 1994; F. Vollrath, Strength and structure of spiders' silks, *Rev. Mol. Biotechnol.*, 74:67–83, 2000; F. Vollrath and D. P. Knight, Liquid crystalline spinning of spider silk, *Nature*, 410:541–548, 2001.

Spilite

An aphanitic (microscopically crystalline) to very fine-grained igneous rock, with more or less altered appearance, resembling basalt but composed of albite or oligoclase, chlorite, epidote, calcite, and actinolite.

In spite of the highly sodic plagioclase, spilites are generally classed with basalts because of the low silica content (about 50%). Spilites also retain many textural and structural features that are characteristic of basalt.

Spilites are found most frequently as lava flows and more rarely as small dikes and sills. Spilitic lavas typically show pillow structure, in which the rock appears composed of closely packed, elongated, pillow-shaped masses up to a few feet across. Pillows are typical of subaqueous lava flows. Vesicles, commonly filled with various minerals, may give the rock an amygdaloidal structure. *See* SYNCLINE.

Once thought by some petrologists to be of uniquely igneous origin, spilites are now recognized as metamorphosed basalts. The metamorphism which takes place consists of the reaction of primary Ca-rich plagioclase to produce the assemblage of albite and one or more calcic minerals such as epidote, calcite, prehnite, and pumpellyite. Primary pyroxene and olivine in the basalt convert typically to chlorite, actinolite, and serpentine. In many cases, except for the filling of vesicles, the texture of the original basalt is essentially preserved. Some basalt flows have endured spilitization of their permeable, vesicular tops, yet remain unaltered in their denser interior. The conversion of basalt to spilite is a hydrothermal metamorphism that results from the action of hot briny water which permeates the rock, taking advantage of the pillow or vesicular structures to gain access to the primary minerals. In places where former oceanic crust is now exposed on land, spilite and keratophyre (spilitized andesites) pillow lavas are seen repeatedly. This indicates that low-grade hydrothermal metamorphism is common in the oceanic crust. *See* BASALT; IGNEOUS ROCKS; METAMORPHISM; METASOMATISM; PETROLOGY.

William Ingersoll Rose, Jr.

Bibliography. G. C. Amstutz et al. (eds.), *Spilites and Spilitic Rocks*, 1974; M. G. Best, *Igneous and Metamorphic Petrology*, 2d ed., 2002.

Spin (quantum mechanics)

The intrinsic angular momentum of a particle. It is that part of the angular momentum of a particle which exists even when the particle is at rest, as distinguished from the orbital angular momentum. The total angular momentum of a particle is the sum of its spin and its orbital angular momentum resulting from its translational motion. The general properties of angular momentum in quantum mechanics imply that spin is quantized in half integral multiples of $\hbar$ ($h/2\pi$, where h is Planck's constant); orbital angular momentum is restricted to half *even* integral multiples of $\hbar$. A particle is said to have spin $^3/_2$, meaning that its spin angular momentum is $^3/_2 h$. *See* ANGULAR MOMENTUM.

A nucleus, atom, or molecule in a particular energy level, or a particular elementary particle, has a definite spin; for instance, a deuteron has spin 1, a $_3Li^{+5}$ nucleus in its ground state has spin $^3/_2$, and an electron has spin $^1/_2$. The spin is an intrinsic or internal characteristic of a particle, along with its mass, charge, and isotopic spin. *See* I-SPIN; SYMMETRY LAWS (PHYSICS).

A particle of spin s has $2s + 1$ spin states, since according to quantum mechanics the projection of an angular momentum of magnitude j along an axis can have the $2j + 1$ integrally spaced values $j, j - 1 \cdots -j + 1, -j$. These spin states represent an internal degree of freedom of a particle, in addition to its external freedom of motion in three-dimensional space.

In field theory, in which particles are regarded as quanta of a field, the spin of the particle is determined by the tensor character of the field. For instance, the quanta of a scalar field have spin 0 and the quanta of a vector field have spin 1. A celebrated theorem of quantum field theory, proved first by W. Pauli, states a connection between spin and statistics: A particle with half even integral spin obeys Bose-Einstein statistics and is called a boson; a particle with half odd integral spin obeys Fermi-Dirac statistics and is called a fermion. *See* ELECTRON SPIN; QUANTUM FIELD THEORY; QUANTUM MECHANICS; QUANTUM STATISTICS.

Charles J. Goebel

Spin-density wave

The ground state of a metal in which the conduction-electron-spin density has a sinusoidal variation in space, with a wavelength usually incommensurate with the crystal structure. This antiferromagnetic state normally occurs in metals, alloys, and compounds with a transition-metal component. It occurs also, however, in quasi-one-dimensional organic conductors. *See* ANTIFERROMAGNETISM; CRYSTAL STRUCTURE; ELECTRON SPIN.

In a metal, the conduction-electron charge densities for up and down spin are ordinarily equal. Their sum, $\rho_0(\mathbf{r})$, exhibits a dependence on $\mathbf{r}$ having the same spatial periodicity as that of the lattice. A metal with a spin-density wave has, instead, charge densities given by Eq. (1). The spin density $\sigma(\mathbf{r})$ is the

$$\rho_\pm(\mathbf{r}) = {}^1/_2\rho_0(\mathbf{r})[1 \pm p\cos\mathbf{Q}\cdot\mathbf{r}] \qquad (1)$$

difference between $\rho_+(\mathbf{r})$ and $\rho_-(\mathbf{r})$, and is therefore given by Eq. (2),

$$\sigma(\mathbf{r}) = \hat{\varepsilon}p\rho_0(\mathbf{r})\cos\mathbf{Q}\cdot\mathbf{r} \qquad (2)$$

where $\hat{\varepsilon}$ is a unit vector defining the axis of spin quantization, and p is the amplitude of the spin-density wave. The wave vector $\mathbf{Q}$ is determined by the conduction-electron Fermi surface. The wavelength, $\lambda = 2\pi/Q$, of the spin-density wave is thus in general not commensurate with the lattice period a. The ratio λ/a varies with temperature, pressure, alloy composition, and other parameters, because of changes in the electronic structure or in the population of electron states as a function of these parameters. *See* FERMI SURFACE.

Origin. The Pauli exclusion principle keeps electrons with parallel spin apart. The resulting reduction in Coulomb interaction is called exchange energy. This reduction is increased when the conduction-electron charge densities are modulated according to Eq. (1). However, modulation requires an increase in conduction-electron kinetic energy. According to the spin-density-wave instability theorem, energy can always be further reduced if the quantum theory is treated in an approximation that neglects electron–electron scattering. Were it not for these latter correlations, every metal would have one or more spin-density waves. *See* EXCHANGE INTERACTION; EXCLUSION PRINCIPLE.

The correlation energy, which contributes to a metal's stability, arises primarily from scattering of up-spin electrons by down-spin electrons. Since a spin-density wave tends to stratify opposite spin densities in alternating layers, as indicated in Eq. (1), the scattering is reduced, thereby tending to suppress the spin-density-wave instability. This contrasts with a charge-density wave, where shifts of exchange and correlation energy act in unison. Consequently, spin-density-wave states are less likely to occur than charge-density-wave states. *See* CHARGE-DENSITY WAVE.

Detection. Since the total electronic charge density, the sum of ρ_+ and ρ_- in Eq. (1), is equal to $\rho_0(\mathbf{r})$, a spin-density wave cannot be detected in an x-ray diffraction experiment. It leads, however, to a sinusoidal magnetic field, from which neutrons, having a magnetic moment, can experience Bragg diffraction. At incommensurate points in wave-vector space, two magnetic satellites are observed on opposite sides of each crystallographic Bragg reflection.

The magnetic origin of spin-density-wave satellites can be demonstrated in two ways. The first is their absence in an x-ray experiment. The second is by means of neutron diffraction with a spin-polarized beam, since only a magnetic reflection can cause reversal of neutron polarization during diffraction. *See* NEUTRON DIFFRACTION; X-RAY DIFFRACTION.

Occurrence. There are well over 100 materials which, over a temperature range, support a

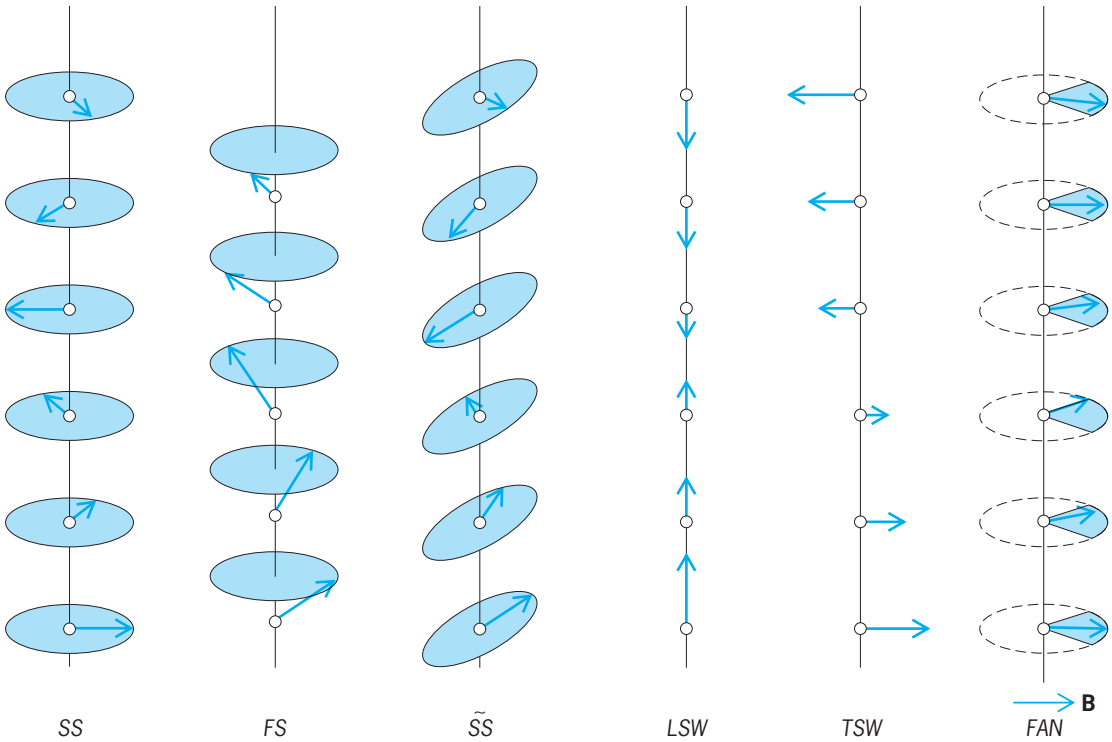

Types of spin-density wave observed in metals: *SS* = spiral (helical); *FS* = ferromagnetic (cone); *S̃S* = skewed; *LSW* = longitudinal; *TSW* = transverse; *FAN* = stable only in a magnetic field B in the direction shown.

spin-density wave. These include some of the rare-earth elements of the lanthanide series and the 3*d* transition metals, manganese and chromium, the latter being the prototypical itinerant electron antiferromagnet. Many types of spin-density wave occur, since in general not only may the vector $\hat{\varepsilon}$ in Eq. (2) be parallel or perpendicular to the wave vector **Q**, corresponding to a longitudinal or transverse spin-density wave, respectively, but in the latter case $\hat{\varepsilon}$ may rotate about **Q**, corresponding to a helical spin-density wave. A uniform ferromagnetic magnetization may be superposed on the spin-density wave (see **illus.**).

The occurrence of inelastic neutron-scattering peaks at incommensurate points indicates the existence of spin-density-wave fluctuations in some metals thought to be nonmagnetic (for example, copper and yttrium) when doped with magnetic impurities (manganese and gadolinium, respectively). This behavior suggests that the spin-density-wave instability may be common, even in nontransition metals.

Spin-density-wave fluctuations have been proposed as an essential feature for some forms of high-temperature superconductivity. In one class of quasi-one-dimensional organic conductors, the Bechgaard's salts, the spin-density-wave states give way to a superconducting state under pressure. In another case, $(TMTSF)_2ClO_4$ (TMTSF = tetramethyltetraselenafulvalene) is superconducting at low temperatures and at ambient pressure but develops a spin-density wave in a magnetic field. *See* ORGANIC CONDUCTOR; RARE-EARTH ELEMENTS; SUPERCONDUCTIVITY; TRANSITION ELEMENTS.

Eric Fawcett

Bibliography. E. Fawcett, Spin-density-wave antiferromagnetism in chromium, *Rev. Mod. Phys.*, 60:209–283, 1988; E. Fawcett, Spin-density waves and SDW fluctuations, *Sov. J. Low Temp. Phys.*, 15:441–444, 1989; Yu. A. Izyumov, Modulated, or long-periodic, magnetic structures of crystals, *Sov. Phys. Usp.*, 27:845–867, 1984; D. Jérome and H. J. Schulz, Organic conductors and superconductors, *Adv. Phys.*, 31:299–490, 1982; N. I. Kulikov and V. V. Tugushev, Spin-density waves and itinerant antiferromagnetism in metals, *Sov. Phys. Usp.*, 27:954–976, 1984; A. W. Overhauser, Spin-density waves in an electron gas, *Phys. Rev.*, 128:1437–1452, 1962.

Spin glass

One of a wide variety of materials that contain interacting atomic magnetic moments and also possess some form of disorder, in which the temperature variation of the magnetic susceptibility undergoes an abrupt change in slope, that is, a cusp, at a temperature generally referred to as the freezing temperature, T_f. At temperatures below T_f the spins have no long-range magnetic order, but instead are found to have static or quasistatic orientations which vary randomly over macroscopic distances. The latter state is referred to as spin-glass magnetic order. Below T_f these systems also exhibit magnetic hysteresis and slow relaxation phenomena. Analogous effects are observed in measurements of torque in an applied magnetic field and in electron paramagnetic resonance.

Spin-glass ordering is usually detected by means of magnetic susceptibility measurements, although additional data are required to demonstrate the absence of long-range order. Closely related susceptibility cusps can also be observed by using neutron diffraction. It is not generally agreed whether spin glasses undergo a phase transition or not. *See* MAGNETIC SUSCEPTIBILITY; NEUTRON DIFFRACTION; PHASE TRANSITIONS.

Spin-glass transition. An example of a spin-glass ordering transition is shown in **Fig. 1**, where magnetic susceptibility χ is plotted as a function of absolute temperature T for the case of dilute alloys of iron in gold. The iron solute atoms carry a magnetic moment and occupy sites at random in the gold lattice. Neighboring moments interact with one another via the Ruderman-Kittel-Kasuya-Yosida (RKKY) exchange interaction U, given byEq. (1), which is

$$U = -\sum_{ij} J_{ij}\vec{S}_i \cdot \vec{S}_j \qquad (1)$$

mediated by the conduction electrons of the gold. The exchange constant J_{ij} oscillates rapidly in sign with increasing distance between spins $\vec{S}_i$ and $\vec{S}_j$. Anisotropic interactions analogous to the RKKY exchange as well as dipolar and crystal field interactions are also believed to have important consequences for spin-glass properties. The sharp susceptibility peak in Fig. 1 occurs at a temperature T_f such that $k_B T_f$ has the value of a typical exchange constant J_{ij}, where k_B is Boltzmann's constant. In this case the necessary disorder arises from the random placement of the magnetic moments. The resulting configuration of exchange couplings is such that at low temperatures the minimum energy state is one where the moment orientations are only correlated locally and are effectively random over macroscopic distances.

As seen in Fig. 1, the sharpness of the transition is destroyed by the application of a relatively small magnetic field, $H_0 \sim 100$ oersteds (1 Oe = 79.6 ampere-turns/m). This is evidence that the transition is a cooperative effect involving a large number of moments, since the motion of an isolated moment would be largely unaffected by a field of 100 Oe at this temperature.

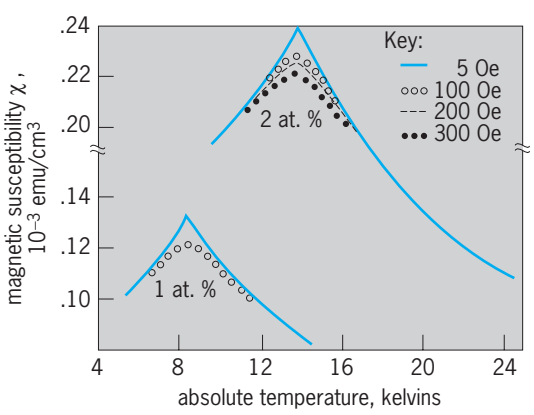

Fig. 1. Magnetic susceptibility versus absolute temperature for two different concentrations of iron alloyed with gold. The data curves indicate measurements in several applied field strengths as shown. 1 Oe = 79.6 units. Magnetic susceptibility in SI units = magnetic susceptibility in emu multiplied by 4π = 12.57. (*After V. Cannella and J. A. Mydosh, Magnetic ordering in gold-iron alloys, Phys. Rev., B6:4220–4237, 1972*)

Some of the basic elements of the spin-glass transition are reproduced in a mean-field model calculation. In this model one calculates the thermodynamic behavior of a system of classical spins which interact via exchange couplings having the form given by Eq. (1) with, however, a distribution of J_{ij} values which is taken to be a gaussian random function. Below the temperature T_f, the system is found to be ordered in the sense that each spin $\vec{S}_i$ will retain some memory of its orientation $\vec{S}_i(t_1)$ at time t_1 at a much later time t_2, even though $\langle \vec{S}_i(t_1) \rangle = 0$, where $\langle\ \rangle$ denotes an average over all spins. The ordering effect is expressed by the Edwards-Anderson order parameter $q = \langle \vec{S}_i(t_1) \cdot \vec{S}_i(t_2) \rangle$. Their theory gives Eq. (2) for the magnetic susceptibility χ. Here $\chi_c = a/T$ is the Curie

$$\chi = \chi_c(1 - q) \qquad (2)$$

susceptibility, where a is the Curie constant. For $T > T_f$, $q = 0$; and for $T < T_f$, $q = {}^1/_2\,[(T_f/T)^2 - 1]$. Thus $\chi(T)$ is just the Curie susceptibility for $T > T_f$, and for $T < T_f$ is given by Eq. (3) for $T_f - T < T_f$, where b is a

$$\chi(T) \cong \frac{a}{T_f} - b(T_f - T)^2 \qquad (3)$$

constant. Equation (2) therefore exhibits a cusp similar to that shown in Fig. 1. The Edwards-Anderson model also predicts a singularity in the specific heat at T_f. However, this is not observed experimentally.

Spin-glass ordering. Spin-glass ordering has been found to occur in a wide variety of disordered magnetic materials, exhibiting several distinct types of disorder. The following list is intended to be illustrative but not exhaustive. The dilute alloy Au:Fe cited above is disordered as a consequence of random placement of iron atoms in the gold lattice. For similar reasons other dilute alloys with noble-metal hosts (for example, Cu:Mn, Ag:Mn, and Au:Mn) are also spin glasses. Some of these systems show spin-glass behavior over a wide range of composition, limited only by the Kondo effect and possible ferromagnetic behavior at low and high concentrations of the magnetic ion, respectively. At low concentrations the exchange interactions of these alloys are well characterized as to their strength and distribution. Another type of metallic spin glass is one which contains a relatively high concentration of magnetic ions and is structurally amorphous. Examples of this are $GdAl_2$ and $(Fe_xMn_{1-x})_{0.75}(P_{16}B_6Al_3)_{0.25}$. In the former case the gadolinium ions are the magnetic constituent. Special techniques are required to prepare these materials in a structurally amorphous state. The second example cited is from a family of materials known as metglass and is a reentrant spin glass, as discussed below. *See* AMORPHOUS SOLID; KONDO EFFECT.

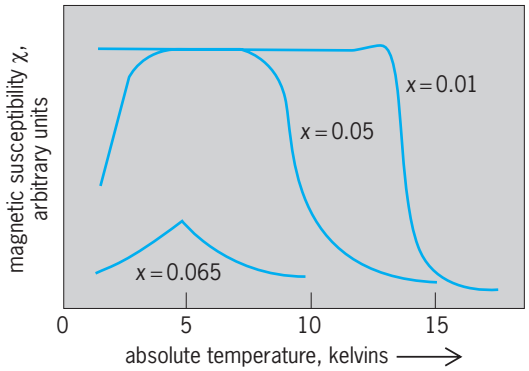

Fig. 2. Magnetic susceptibility versus absolute temperature for the compound $(Pd_{0.9965}Fe_{0.0035})_{1-x}Mn_x$ for three values of the manganese atomic fraction *x*. The scale for the curve marked *x* = 0.065 has been magnified four times. (*After G. J. Nieuwenhys, B. H. Verbeek, and J. A. Mydosh, Towards a uniform magnetic phase diagram for magnetic alloys with mixed types of order, J. Appl. Phys., 50:1685–1690, 1979*)

There are also a number of interesting cases of insulating spin glasses. The best-characterized and most thoroughly studied insulating spin glass is the system $Eu_xSr_{1-x}S$. This compound is a ferromagnet for $x = 1$ as a consequence of the ferromagnetic nearest neighbor exchange interaction among the europium, while the original crystalline structure is retained. At sufficient dilution ($x < 0.5$) the system becomes a spin glass through the agency of antiferromagnetic second-neighbor exchange couplings. At higher europium concentrations this system also becomes reentrant.

Ferromagnetic to spin-glass transition. A number of compounds are found to form ferromagnetic order at some Curie temperature T_c, then exhibit a collapse of that order as the temperature is lowered below T_c. Such a system is termed a reentrant spin glass, because the spin-glass phase appears below a temperature where the magnetically ordered higher temperature phase collapses. An example of this behavior is shown in **Fig. 2**, where magnetic susceptibility versus absolute temperature is plotted for three concentrations of manganese (Mn) in the compound $(Pd_{0.9965}Fe_{0.0035})_{1-x}Mn_x$. For $x = 0.01$ the compound is a ferromagnet, with χ saturated for all temperatures below 13 K. For $x = 0.05$ the compound becomes ferromagnetic at $T \sim 9$ K, but the ferromagnetism collapses into spin-glass behavior for $T < 3$ K. This is the reentrant-spin-glass case. For higher manganese concentrations (for example, $x = 0.065$) the ferromagnetism disappears, and the material is a simple spin glass as in Fig. 1. Whether the spin-glass transition of the ferromagnetic to spin-glass transition is a true phase transition is a matter of some controversy. Reentrant-spin-glass behavior has been observed in a variety of other systems as well, including $Eu_xSr_{1-x}S$, $Au_{1-x}Fe_x$, Fe_xCr_{1-x}, and amorphous $(Fe_{1-x}Mn_x)_{0.75}(P_{16}B_6Al_3)_{0.25}$. The first of these is a nonconductor, and the others are magnetic alloys. *See* FERROMAGNETISM. R. E. Walstedt

Spin label

A molecule which contains an unpaired electron spin which can be detected with electron spin resonance (ESR) spectroscopy. Molecules are labeled when an atom or group of atoms which exhibit some unique physical property is chemically bonded to a molecule of interest. Groups containing unpaired electrons include organic free radicals and a variety of types of transition-metal complexes (such as vanadium, copper, iron, and manganese). Molecules with unpaired electron spins are readily detected with electron spin resonance spectroscopy. Through analysis of ESR spectra, rates of molecular motion whose motion is restrained by surrounding molecules can be determined. Measurement of rates of molecular motion and molecular orientation has proved to be very important in the study of a variety of types of biologically important problems.

Electron spin resonance. This is a spectroscopic technique which detects transitions between electron spin energy levels. When a molecule possessing an unpaired electron spin is placed in an external magnetic field, the external field interacts with the magnetic moment of the unpaired electron spin, producing two distinct energy states. In one state the unpaired electron is oriented with its magnetic moment along the direction of the external field, while the other state has the moment of the unpaired electron oriented against the external field direction. The separation of these two energy states is about 3 cm^{-1} in an external field of 3000 gauss (0.3 tesla), and an absorption of radiation can be observed when a sample is irradiated with microwave radiation of about 3 cm wavelength. This absorption is observed as a single line in an electron spin resonance spectrum. The wavelength at which absorption of energy is observed depends on the magnitude of the external field and on a molecular property called the g value.

Crystals. If a crystalline sample is used, different g values may be observed along the various axes of the crystal. When directional dependence is observed the g value is anisotropic, with principal values along the x, y, and z axes of the crystal denoted by g_{xx}, g_{yy} and g_{zz}. If $g_{xx} = g_{yy}$, the sample has an axial symmetric g tensor, and the g values are denoted as $g_{\parallel}$ and $g_{\perp}$. The unpaired electron spin is often delocalized over two or more atoms in the spin label. If the nuclei have nuclear spin other than zero, the magnetic moment of the nuclei can interact with the magnetic moment of the electron spin and split the electron spin energy levels. When this is the case, the ESR spectrum shows a series of lines that depend on the number and type of nuclear spins in the molecule. The interaction between the electron and nuclear spins is called a hyperfine interaction, and the separation of the energy levels and ESR lines is called the hyperfine splitting. The magnitude of the interaction between nuclear and electron spins may depend on the orientation of the molecule with respect to the external field. The components of the

electron-nuclei anisotropic splitting are denoted by A_{xx}, A_{yy}, and A_{zz}, or $A_\perp$ and $A_\parallel$ when the molecule has axial symmetry.

Solutions. When a spin-labeled molecule is free to rotate in solution, the anisotropic g value and anisotropic hyperfine interaction are averaged by rapid reorientation of the label, and one observes the isotropic (angle-independent) g value and hyperfine splitting in the ESR spectrum. As rotation of the molecule is slowed, the anisotropic interactions start to contribute, and the ESR spectrum is changed. A continuous change in the ESR spectrum is observed as the sample goes from the case of free rotation to the case of a rigid sample where motion is very slow. Analysis of ESR spectra permits determination of motional correlation times from about 10^{-12} to 10^{-3} s. In many cases, determination of the direction of motion of the labeled group (that is, rotation in all directions or preferred rotation around a single axis) is also possible.

Biological studies. Analysis of the rate and type of motion of a spin label is important for a wide variety of biological problems. The type of label used in these studies is generally a nitroxide free radical. The unpaired electron spin in a nitroxide radical is delocalized over an N-O group, and an electron-nuclei hyperfine splitting is observed from the spin = 1 nitrogen nucleus. The oxygen atom has a spin of zero. In solution an isotropic nitrogen hyperfine splitting of about 15 gauss (1.5 millitesla) is observed, while in a crystalline solid anisotropic splitting is observed with $A_{zz} = A_\parallel = 32$ gauss (3.2 millitesla) and $A_{xx} = A_\perp = 6$ gauss (0.6 mT). The z direction is defined by the $2p$ orbital of the nitrogen atom.

Enzymes. Nitroxide spin labels have been bonded to a number of enzyme substrates with connecting chains of varying length between the label and the substrate. The labeled substrate is allowed to form a complex with the enzyme, and the ESR spectrum is monitored. The receptor site on the enzyme is often in a type of well within the protein, within which the substrate must fit. If the nitroxide label is connected to the substrate by a short chain, it lies within the well and rotational motion is inhibited. When a sufficiently long connecting chain is used, the nitroxide lies outside the protein and rotational motion is relatively free. Changes in motion of the label can be monitored with ESR, and the transition from hindered to free rotation as a function of the length of the connecting chain can be determined. The distance from the substrate to the label can be calculated from molecular models for labeled molecules with different connecting chain lengths. If the distance at which free rotation starts is determined, the depth of the receptor site well of the enzyme can be calculated. Spin-labeling studies of this type provide a powerful technique for the study of the geometry and dimensions of receptors in enzymes. *See* ENZYME.

Membranes. Spin labels have been used extensively to study the structure of membranes. Biological membranes consist mainly of a phospholipid bilayer with embedded proteins. Other molecules (for example, cholesterol) may also be found in the lipid bilayer, and a variety of types of molecules may be bound to the surface. The phospholipids have polar head groups (hydrophilic) which are on the outside of the membrane, and nonpolar aliphatic carbon chains which are on the inside of the bilayer. The lipids form ordered structures which are readily oriented in an external magnetic field. Spin labels have been bonded to the carbon atoms along the aliphatic chains of the lipids to study internal motion in the membrane bilayer. The ESR spectra of spin labels bound to lipids which have been oriented in an external field show that the lipid can rotate around its long axis, but that head-to-tail rotation is very slow. The relative motional correlation times of labels bound at different positions along the chain can be determined from ESR spectra. These studies show that motion is freer on the inside of the lipid away from the polar head group, indicating greater fluidity on the inside. Membrane structures can be changed by a number of kinds of perturbations, and spin labels have been used to monitor the effect of the perturbation. The structure of the phospholipid bilayer changes with temperature, and spin-labeling studies have been conducted to determine phase transition temperatures. Molecules such as cholesterol or ethanol change the fluidity of the membrane, and spin-labeling studies have been conducted to monitor the effect of chemical perturbations on the fluidity of membranes. *See* CELL MEMBRANES.

When nitroxide labels are relatively close to one another, the odd electron spins in the nitroxide groups may be interchanged between molecules. This spin exchange reaction dramatically changes the appearance of the ESR spectrum of the labels, since the energy states are averaged by this interaction. This interaction may be used to study the rate of diffusion of labeled molecules from an area of high concentration (strong exchange) to an area of low concentration (weak exchange). This type of investigation has been conducted to determine the rate of lateral diffusion of lipids in membranes. A high concentration of labeled lipid is initially introduced into a membrane, and an ESR spectrum from strongly exchange-coupled spin labels is observed. As the labeled molecules diffuse apart, the ESR spectrum changes to that characteristic of labels with no exchange coupling. By monitoring the change in the ESR spectrum of the label as a function of time, it is possible to determine how fast lipids diffuse through membranes. Spin-labeling studies of these types have provided important information about the structure, organization, and rates of motion in membranes.

Other applications. Spin labels have also been used to study the structure and organization of synthetic polymers and to study phase transitions. Investigations using deuterium substitution for hydrogen with nuclear magnetic resonance detection have been conducted to study problems similar to those investigated with nitroxide labels. Spin-labeling studies allow investigation of a variety of types of problems which cannot readily be studied with other

techniques. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; FREE RADICAL; TRIPLET STATE.

Robert Kreilik

Bibliography. L. J. Berliner, *Biological Magnetic Resonance*, vol. 14: *Spin Labeling*, 1998; L. J. Berliner (ed.), *Spin Labeling: Theory and Applications*, vols. 1 and 2, 1976, 1979; C. P. Poole, Jr., *Electron Spin Resonance*, 2d ed., 1997.

Spinach

A cool-season annual of Asiatic origin, *Spinacia oleracea*, belonging to the plant order Caryophyllales. It is grown for its foliage (**Fig. 1**) and served as a cooked vegetable or as a salad. New Zealand spinach (*Tetrogonia expansa*) and Mountain spinach (*Atriplex hortense*) are also called spinach but are less commonly grown. Spinach plants are usually dioecious. *See* CARYOPHYLLALES.

Cultivation and harvesting. Propagation is by seed, commonly planted in rows 10–20 in. (25–50 cm) apart. Cool weather favors maximum production. High temperatures and long daylight periods encourage seedstalk formation and reduce vegetative growth. Fall-seeded spinach is overwintered and harvested in the spring in areas where the weather is mild or the crop is protected by snow cover. *See* PHOTOPERIODISM; SEED.

Varieties (cultivars) are classified according to seed (smooth or prickly), leaves (smooth or savoyed), and seeding tendencies (early or late, or long standing). Popular varieties are Long Standing Bloomsdale, Early Hybrid 7, Hybrid 424, Savoy Hybrid 621, and Viroflay.

Harvesting by hand or machine begins when the plants have reached full size and before seedstalks form, usually 40–50 days after planting.

H. John Carew

Disease. Diseases may reduce yield and quality, and be limiting factors in the production of spinach for processing and fresh market. Some of the most important diseases of spinach are Fusarium wilt, caused by *Fusarium oxysporum* f. sp. *spinaciae*; white rust, caused by *Albugo occidentalis*; downy mildew, caused by *Peronospora spinaciae*; blights, mosaics, and stunts caused by viruses and yellows and decline (**Fig. 2**). Several factors may cause a yellows and decline syndrome; these include soil-borne pathogenic fungi and nematodes, nutrient imbalances, excess soil moisture, abrupt temperature changes, and prolonged periods of cloudy, wet weather.

Fig. 1. Broad, mature leaves of spinach plants in the field. (*Editorial Photocolor Archives*)

Fig. 2. Four seedlings with root and hypocotyl decay (left) and three healthy seedlings (right).

Crop rotations, with spinach being planted only once every 3 years, and the use of disease-free seed are the most effective control measures for white rust, downy mildew, and other leaf spot diseases that survive in seed, plant debris, and soil. The wilt pathogen will survive in the soil indefinitely, so crop rotation is not an effective method of control.

Some varieties are resistant to downy mildew and blight, but market type and quality are not always acceptable. Fungicides can be used to control downy mildew and white rust, but fungicides are not available to control other foliage diseases of less economic importance, such as Cercospora leaf spot and anthracnose. To help reduce the spread of epidemics of foliage diseases, spring spinach should not be planted in the vicinity of winter spinach.

Viral diseases may be produced by destroying volunteer spinach plants and weeds that serve as alternative hosts. Controlling insect vectors may slow the spread of viruses within a field and between fields. Early spring crops can be grown before blight becomes prevalent in the northern United States, but the practice is ineffective in southern Texas, where vectors may be blown in by northerly winds. *See* PLANT PATHOLOGY.

Donald R. Sumner

Bibliography. A. F. Sherf and A. A. MacNab, *Vegetable Diseases and Their Control*, 2d ed., 1986.

Spinal cord

The portion of the central nervous system within the spinal canal of the vertebral column, that is, the entire central nervous system except the brain. The spinal cord extends from the foramen magnum at the base of the skull to a variable level of the spinal canal; it terminates at the lumbar level in humans and extends well into the caudal region in fishes. Paired spinal nerves enter the spinal canal between each pair of vertebrae and connect with the spinal cord. The number of spinal nerves varies widely in vertebrates; in humans there are 31 pairs (8 cervical,

12 thoracic, 5 lumbar, 5 sacral, and 1 coccygeal). In tetrapods there are spinal cord enlargements where the large spinal nerves extend to the limbs, a cervical enlargement cranially, and a lumbosacral enlargement caudally.

The internal organization of the spinal cord can be seen in a transverse section (**Fig. 1**). The outer portion of the spinal cord is made up of nerve fibers, most of which are oriented longitudinally and carry information between parts of the spinal cord, between spinal cord and brain, and between brain and spinal cord. The outer white matter is divided into dorsal, lateral, and ventral columns. The interior of the spinal cord consists of gray matter and is divided into a dorsal sensory horn (or column) and a ventral motor horn (or column). In the thoracic and lumbar regions of the cord there is also a small lateral horn (or column) which contains preganglionic sympathetic neurons. In the very center of the bilaterally symmetrical spinal cord is a small central canal, lined with ependymal cells and containing cerebrospinal fluid. *See* SYMPATHETIC NERVOUS SYSTEM.

Each spinal nerve divides into a dorsal sensory root and a ventral motor root before entering the spinal cord. The dorsal root contains the spinal ganglion (**Fig. 2**), and then its axons enter the dorsolateral aspect of the cord. Upon entering the spinal cord, each dorsal root axon divides and ramifies to send terminal branches to one or all of the following neuron groups: sensory neurons of the dorsal horn; motor neurons of the ventral horn; and long terminal branches to neurons in the brain. The motor neurons of the ventral horn, in addition to receiving synapses from dorsal root axons, also receive synaptic endings from neurons in other parts of the spinal cord and from long axons coming from the brain. The axons of the ventral horn neurons leave the cord through the ventral root of the spinal nerve and run with peripheral nerves to innervate the muscles of the body. With this complex synaptic and fiber organization, the spinal cord can act as the integrating center for spinal reflexes (such as the knee jerk reflex), send sensory information from the brain, and receive information from the brain to initiate or inhibit muscular activity. *See* MOTOR SYSTEMS; SENSATION.

Other than size and shape differences, the spinal cords of vertebrates are basically similar. However, in fishes some or all of the visceral motor fibers pass through the dorsal roots rather than the ventral roots of spinal nerves. *See* AUTONOMIC NERVOUS SYSTEM; NERVOUS SYSTEM (VERTEBRATE). Douglas B. Webster

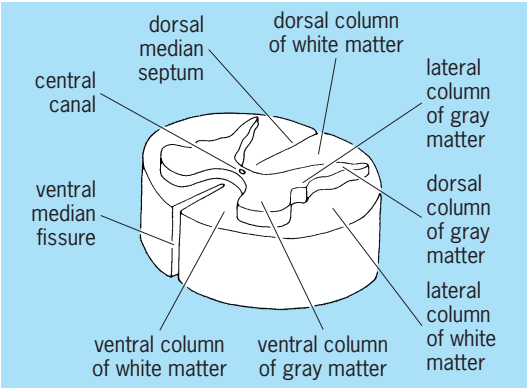

Fig. 1. Human spinal cord sectioned to show the white and gray matter. (*After G. M. Wyburn, The Nervous System, Academic, 1960*)

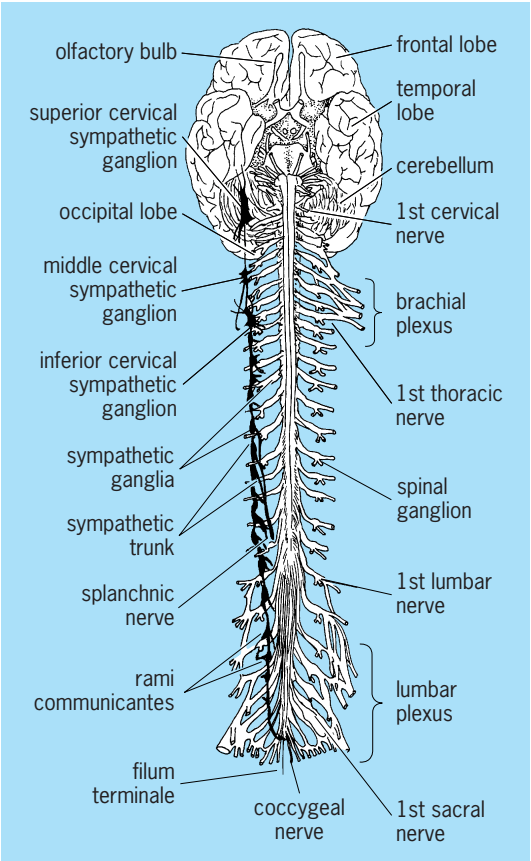

Fig. 2. Brain and spinal cord, illustrated in ventral aspect. Nerve roots and chief autonomic ganglia are shown. (*After H. Morris et al., Comparative Anatomy, Blakiston–McGraw-Hill, 1936*)

Spinal cord disorders

In addition to those disorders common to the brain, the spinal cord is subject to certain lesions because of its position or structure. A few of the more important are mentioned.

Injury. Spinal cord injury results from dislocations, fracture, or compression in many cases, but a special form, called spinal shock, may result from a severe blow without actual distortion of adjacent tissue. In this case, there is a temporary paralysis which gradually clears in a length of time related to the severity of the paralysis. In direct damage, the cord may be slightly, partially, or completely damaged at one or more levels. Typical motor and sensory losses follow, with a poor prognosis for recovery if the nerve tissue is severely injured. Occasionally, similar damage may result from a hemorrhage produced by injury.

A fairly common type of potential cord injury is seen in a number of cases of slipped disks, in which the inner, soft part of the vertebral column extrudes into the spinal canal. If this compresses the cord, functional loss of temporary or permanent degree

can follow; more often, pressure is exerted on spinal roots so that pain, numbness, and some type of muscle weakness intervene.

Tumors. Spinal cord tumors are not infrequent and most of these are of two types, the metastatic, from a primary source elsewhere in the body, and the tumors of the meninges or connective tissue related to the cord. The latter include neurofibromas, meningiomas, and gliomas, which occur most often. The signs and symptoms and the extent of damage relate largely to the physical compression of the cord at a particular level. *See* TUMOR.

Congenital defects. A few of the more common congenital defects involving the cord include an unclosed neural canal, or spina bifida, and reduplicated or otherwise malformed cords, such as those caught in an external sac of other tissues, the meningomyelocoele.

Inflammation. Inflammations may result from known or unknown agents and in meningitis may involve primarily the coverings; in myelitis, the cord itself. The meningococcus, pneumococcus, streptococcus, tubercle bacillus, and other microorganisms frequently cause meningitis. The most widely known cause of myelitis, of course, is the poliomyelitis virus group, although other agents will also produce inflammations, such as typhus, spotted fever, and *Treponema pallidum*, the cause of syphilis. Other diseases such as malaria, amebiasis, trichinosis, elephantiasis, toxoplasmosis, and blastomycosis may display their effects on the spinal cord.

Vascular diseases. Not uncommonly motor or sensory deficits in the spinal cord may result from specific vascular diseases and from such nonspecific conditions as injury, aneurysms, or congenital defects, and rarely, from blood vessel tumors.

Nerve tract degeneration. An ill-defined group of disorders characterized by degeneration of nerve tracts or myelin sheaths of the cord is found more often than one would suspect. In these, some unknown or poorly understood mechanism causes the deterioration of cells and fibers so that function is altered, then lost, and the nervous tissue is either replaced by a scar or a softening of cystlike area remains. Multiple sclerosis, combined degeneration associated with pernicious anemia, Parkinson's disease, postinfectious encephalomyelitis, and syringomyelia are examples. *See* MULTIPLE SCLEROSIS; PARKINSON'S DISEASE.

Drugs. Finally, the effect of drugs or chemicals and metabolic disorders is often seen as cord depression or stimulation with variable prognoses and permanence of effects. These include alcoholism, carbon monoxide poisoning, lead intoxication, vitamin deficiency, and certain familial diseases. *See* ALCOHOLISM; POISON; SPINAL CORD; VITAMIN.

Edward G. Stuart; N. Karle Mottet

Spinicaudata

An order of fresh-water branchiopod crustaceans formerly included in the order Conchostraca. The entire body and its appendages are covered by a bivalve carapace up to about 17 mm (0.7 in.) long, hence the name clam shrimp. The carapace is often bilaterally compressed and usually displays several lines of growth. The head cannot be protruded. It usually bears paired sessile eyes, but these are fused in *Cyclestheria*. Locomotion is by means of large antennae. The mandibles are of the rolling, grinding type.

There are more than 16 pairs of similar foliaceous trunk limbs (at least 32 in some species); these draw in particles of food by beating in metachronal rhythm. In the male, the first two pairs (the first pair only in *Cyclestheria*) are modified as claspers. The trunk terminates in a robust telson that bears a pair of stout claws.

Reproduction is usually bisexual. Usually, the female carries a clump of numerous, very small, drought-resistant eggs, sometimes of complex shape, that are shed at the next molt. In *Cyclestheria* a few larger parthenogenetic eggs are usually brooded in a dorsal brood pouch beneath the carapace. Eggs, except the parthenogenetic eggs of *Cyclestheria*, hatch as nauplii. All species frequent temporary pools in many parts of the world, but *Cyclestheria* also occurs in permanent waters. *See* BRANCHIOPODA.

Geoffrey Fryer

Spinning (metals)

A production technique for shaping and finishing metal. In the spinning of metal, a sheet is rotated and worked by a round-ended tool, controlled manually or mechanically. The sheet is formed over a mandrel.

Spinning operations are usually carried out on a special rigid lathe fitted only with a driving headstock, tail spindle, and tool rest. Surface speeds of 500–5000 ft/min (2.5–25 m/s) are used, depending on material and diameter. The work is rubbed with soap, lard, or a similar lubricant during working. The operation can be set up quickly, and thus is desirable for short runs or for experimental units subject to change. Spinning may serve to smooth wrinkles in drawn parts, provide a fine finish, or complete a forming operation as in curling an edge of a deep-drawn part. Other operations include smoothing, necking, bulging, burnishing, beading, and trimming. Spun products range from precision reflectors and nose cones to kitchen utensils. Such materials as steel, aluminum, copper, and their softer alloys are spun in thicknesses up to 18 in. (3 mm). It may be necessary to anneal the metal during the spinning. *See* SHEET-METAL FORMING.

Ralph L. Freeman

Spinning (textiles)

The fabrication of yarn (thread) from either discontinuous natural fibers or bulk synthetic polymeric material. In a textile context the term spinning is applied to two different processes leading to the yarns used to make threads, cords, ropes, or woven or knitted textile products.

Natural fibers, such as wool, cotton, or linen, are generally found as short, entangled filaments. Their conversion into yarn is referred to as spinning. After a carding operation on the raw material to disentangle the short filaments, the filaments are drawn (drafted) to promote alignment in an overlapping pattern and then twisted to form, by mechanical interlocking of the discontinuous filaments, a resistant continuous yarn. *See* COTTON; LINEN; WOOL.

The term spinning is also used for the production of monofilaments from synthetic polymers—for example, polyamides or nylons, polyesters, and acrylics—or modified natural polymers, such as cellulose-rayon. The process involves forcing a pressurized fluid—either a melt, solution, or chemical intermediate—through many small orifices of a spinnerette and subsequently solidifying the emerging monofilaments by cooling, drying, or coagulation. Generally the monofilaments are then stretched (drawn) to increase their strength by promoting molecular orientation and are wound as yarn which can be used directly for threads, cords, or ropes. Such yarn, however, is often cut into relatively short lengths (staple) and reformed by a process similar to that used for natural fibers into a yarn more suitable, in terms of appearance and feel, for making certain textile products. *See* MANUFACTURED FIBER; NATURAL FIBER; TEXTILE. J. M. Charrier

Bibliography. K. L. Hatch, *Textile Science*, 1993; M. L. Joseph et al., *Joseph's Introductory Textile Science*, 6th ed., 1992; M. Raheel (ed.), *Modern Textile Characterization Methods*, 1996; B. Wulfhorst, T. Gries, and D. Veit, *Textile Technology*, 2006.

Spiral

A term used generically to describe any geometrical entity that winds about a central point or axis while also receding from it. Spiral staircases, helices, nonplanar loxodromes (curves that intersect those of a given class at a constant angle, for example, rhumb lines, in case the curves are on a sphere whose meridians form the given class) are examples of spirals whose windings do not lie in a plane. The planar spiral curve $\rho = a\theta$, in polar coordinates ρ and θ (see **illus.**), was introduced by Archimedes in his book *On Spirals*. Other planar spiral curves are logarithmic, $\rho = e^{\theta}$, the tangents to which make a constant angle with the radii vectors drawn to the points of contact; hyperbolic, $\rho\theta = a$; and the lituus, $\rho^2\theta = a^2$. *See* HELIX. Leonard M. Blumenthal

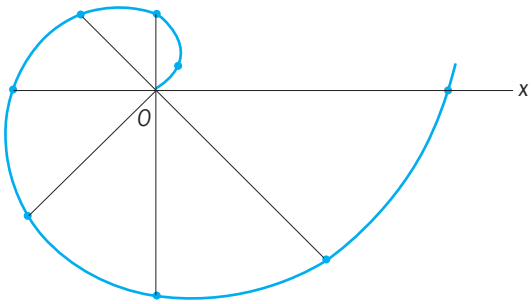

Planar spiral curve, or Archimedes' spiral.

Spiriferida

An extinct order of brachiopods, in the subphylum Rhynchonelliformea, that inhabited shallow seas of the Paleozoic and early Mesozoic. It was the most diverse group of spire-bearing brachiopods (those that contain a spirally coiled calcareous structure called the brachidium used to support the lophophore).

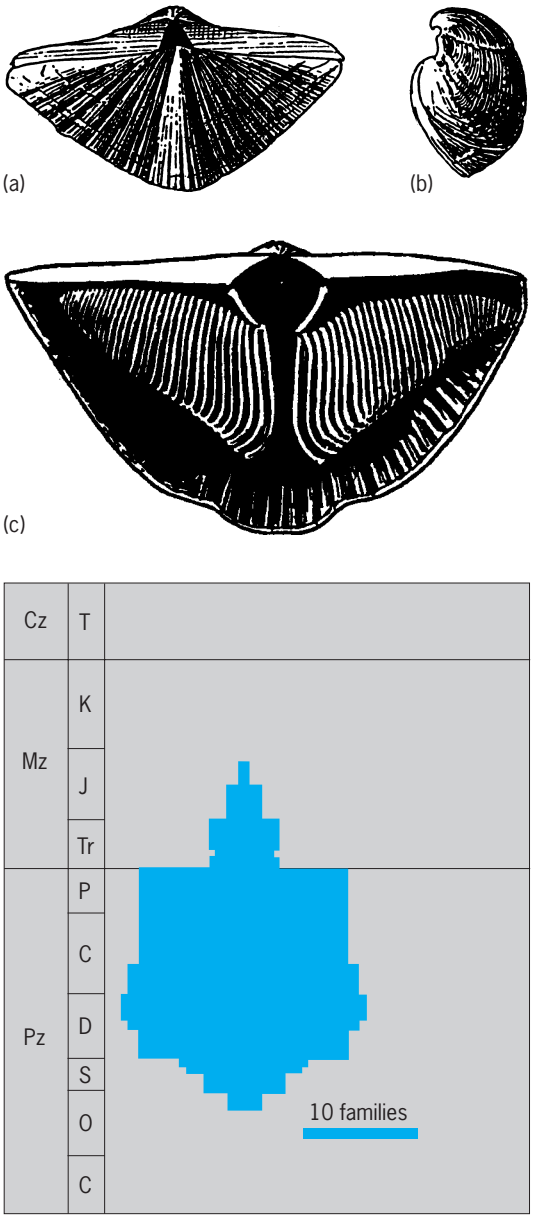

Spiriferida. (*a-b*) Dorsal and lateral views of *Neospirifer*, Permian, and (*c*) with dorsal valve removed showing spiral lophophore supports. (*d*) Diversity history of spiriferid families. (*a–c after R. C. Moore, C. G. Lalicker, and A. G. Fischer, Invertebrate Fossils, McGraw-Hill, New York, 1952; d after J. J. Sepkoski, Jr., A Compendium of Fossil Marine Animal Families, 2d ed., Milwaukee Public Mus. Contrib. Biol. Geol., no. 83, 1992*)

Spiriferids also possess unequally biconvex valves that are externally smooth or radially ribbed. The valves are generally strophic (straight hinge) with a well-developed interarea commonly limited to the ventral valve (**illus.** *a–c*). Shells may be punctate (possessing small perforations filled by outer epithelial tissue) or impunctate. Punctate shells arose in stratigraphically younger spiriferids; the presence or absence of punctae serves to distinguish two major taxonomic groups.

The evolutionary origin of spiriferids is unclear. One hypothesis has the spiriferids derived from the rhynchonellids via another spire-bearing group called the athyrids. A competing hypothesis has the spiriferids derived directly from the orthids and the other spire-bearing groups derived from the rhynchonellids. Clarification of this problem will require additional analyses of new fossil material.

Spiriferids were sessile, attached, epifaunal suspension feeders. Most had a functional pedicle, used for attachment to the substrate. The earliest spiriferids appeared in deep-water environments in the Late Ordovician and thus did not take part in the Ordovician radiation of marine life. Spiriferids first diversified in the Silurian following the Late Ordovician mass extinction (illus. *d*). They, along with strophomenids and pentamerids, became dominant members of marine benthic communities, replacing the orthids. They reached their acme in diversity and dominance of communities in the Devonian. They suffered vast extinction in the Late Devonian, but rebounded in diversity and were still dominant members of communities in the late Paleozoic. Spiriferid diversity dropped dramatically following the late Permian mass extinction and never returned to Paleozoic levels. Spiriferids suffered another big decline in the Late Triassic mass extinction, and finally died out in the Middle Jurassic. *See* BRACHIOPODA; ARTICULATA (ECHINODERMATA). Mark E. Patzkowsky

Bibliography. P. Copper and R. Gourvennec, Evolution of the spire-bearing brachiopods (Ordovician-Jurassic), in P. Copper and J. Jin (eds.), *Brachiopods*, A. A. Balkema, Rotterdam, 1996.

Spirometry

The measurement, by a form of gas meter, of volumes of gas that can be moved in or out of the lungs. The classical spirometer is a hollow cylinder (bell) closed at its top. With its open end immersed in a larger cylinder filled with water, it is suspended by a chain running over a pulley and attached to a counterweight (see **illus.**). The magnitude of a gas volume entering or leaving is proportional to the vertical excursion of the bell. In another type of spirometer, the bell is wedge shaped and is constrained to rotate on an axis located along its apex. Volume changes are proportional to the arc traversed.

Volume changes can also be determined from measurements of flow, or rate of volume change. Flow may be measured continuously by a pneumotachograph. A resistive element, such as a bundle of fine-bore tubes, is inserted in the tube through which breathing occurs. The pressure drop across the resistance is proportional to flow and can be sensed and recorded continuously by a transducer that generates an electrical signal. The flow signal can be continuously integrated to yield a volume trace.

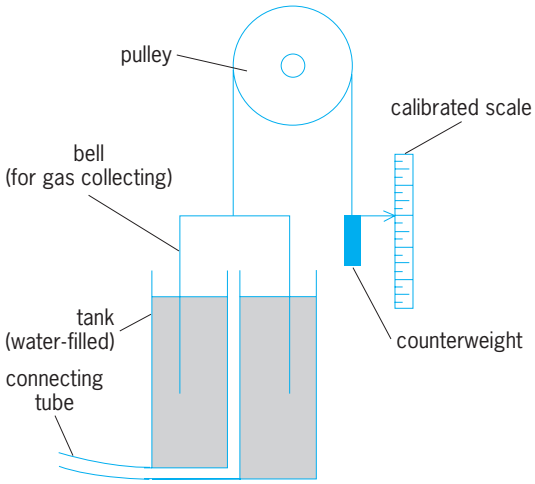

Diagram of a spirometer.

The volume of gas moved in or out with each breath is the tidal volume; the maximal possible value is the vital capacity. Even after the most complete expiration, a volume of gas that cannot be measured by the above methods, that is, the residual volume, remains in the lungs. It is usually measured by a gas dilution method or by an instrument that measures blood flow in the lungs. Lung volumes can also be estimated by radiological or optical methods.

At the end of an expiration during normal resting breathing, the muscles of breathing are minimally active. Passive (elastic and gravitational) forces of the lungs balance those of the chest wall. In this state the volume of gas in the lungs is the functional residual capacity or relaxation volume. Displacement from this volume requires energy from natural (breathing muscles) or artificial (mechanical) sources. *See* RESPIRATION. Arthur B. Otis

Spirophorida

An order of sponges of the class Demospongiae, subclass Tetractinomorpha, with a globular shape and a skeleton of oxeas and triaenes as megascleres, and microspined, contorted, sigmalike microscleres.

Internal buds or gemmules may be formed. The eggs become affixed to substratum by means of long pseudopodia and undergo cleavage and morphogenesis without a free larval stage.

Species of *Tetilla* may have basal spicule tufts functioning to anchor the sponge in sandy bottoms. Species of *Cinachyra*, with ostia and oscules restricted to surface depressions, are common inshore inhabitants of coral reef lagoons. Members of this order range down to depths of at least 5900 ft (1800 m). *See* DEMOSPONGIAE. Willard D. Hartman

Spirurida

An order of nematodes in which the labial region is most frequently provided with two lateral labia or pseudolabia; in some taxa there are four or more lips; rarely lips are absent. Because of the variability in lip number, there is variation in the shape of the oral opening, which may or not be surrounded by teeth. The amphids are most often laterally located; however, in some taxa they may be located immediately posterior to the labia or pseudolabia. The stoma may be cylindrical and elongate or rudimentary. The esophagus is generally divisible into an anterior muscular portion and an elongate swollen posterior glandular region, where the multinucleate glands are located. Eclosion larvae are usually provided with a cephalic spine or hook and a porelike phasmid on the tail.

All known spirurid nematodes utilize an invertebrate in their life cycle; the definitive hosts are mammals, birds, reptiles, and, rarely, amphibians. The order is divided among four superfamilies: Spiruroidea, Physalopteroidea, Filarioidea, and Drilonematoidea.

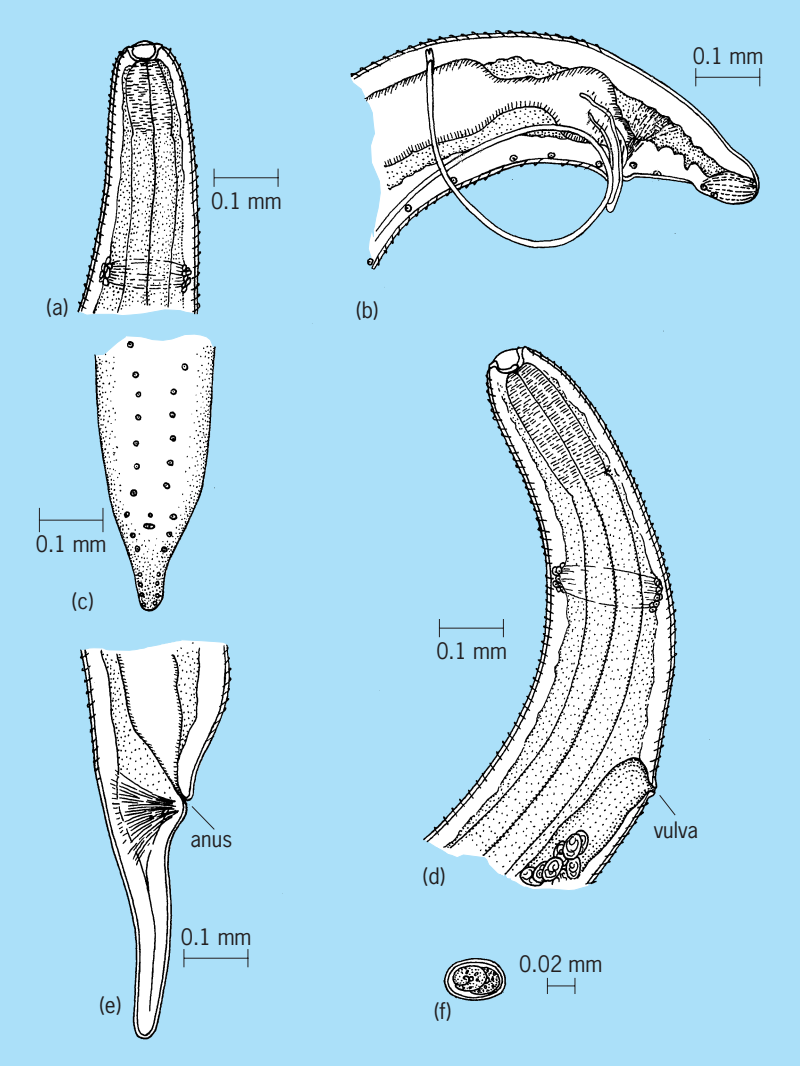

Fig. 1. ***Thelazia platyptera*, a parasitic spiruroid nematode. (*a*) Anterior end of male, lateral view; (*b*) caudal end of male, lateral view; (*c*) caudal end of male, ventral view; (*d*) anterior end of female, lateral view; (*e*) caudal end of female, lateral view; (*f*) egg. (*Parasite Laboratory, Agricultural Research Service, Beltsville, Maryland*)**

Spiruroidea. The superfamily Spiruroidea comprises parasitic nematodes whose life cycle always requires an intermediate host for larvae to the third stage. The definitive hosts are mammals, birds, fishes, reptiles, and rarely amphibians; and spiruroids may be located in the host's digestive tract, eye, or nasal cavity or in the female reproductive system (**Table 1**). Morphologically (**Figs. 1** and **2**), the lip region is very variable in Spiruroidea, ranging from four lips to none. When lips are present, the lateral lips are well developed and are referred to as pseudolabia. The cephalic and cervical region may be ornamented with cordons, collarettes, or cuticular rings. The stoma is always well developed and is often provided with teeth just inside the oral opening. In birds the nematodes are often associated with the gizzard, and the damage caused results in death, generally by starvation. When the muscles of the gizzard are destroyed, seeds pass intact and cannot be digested.

This superfamily contains the largest of all known nematodes, *Placentonema gigantissima*, parasitic in the placenta of sperm whales. Mature females attain a length of 26 ft (8 m) and a diameter of 1 in. (2.5 cm). They are unusual not only because of their great length but also because the adult female has 32 ovaries, which produce great numbers of eggs. The eggs are not unusually large. Armand Maggenti

Filarioidea. The superfamily Filarioidea contains highly specialized parasites of most groups of vertebrates. They are particularly common in amphibians, birds, and mammals. While they cannot be classified as completely harmless, most of the many hundreds of known species are not associated with any recognized disease. A limited number of species produce serious diseases in humans, and a few others produce serious diseases in domestic or wild animals (**Table 2**). The filarial parasites of humans are found almost exclusively in the tropics, with some extension into the subtropics. This is by no means the universal pattern. Heartworm disease in dogs has become established in the north temperate zone in the United States and Canada and in the south temperate zone in Australia. Several other enzootic species are known in northern amphibians, birds, and mammals. *See* HEARTWORMS.

Morphology of adults. Among the Filarioidea there are no conspicuous divisions into distinct body regions. True lips are absent, and the mouth cavity is greatly reduced. The cuticule is devoid of prominent ornaments found on many nematodes. The female genital opening is usually located in the region of the esophagus-intestine junction in the anterior portion of the worm. Sexual dimorphism is the rule; in common with other nematodes, the female filarioid is at least twice as long as the male, and often the difference is much greater. Among the species found in humans or dogs, the females vary from 1 to 28 in. (2.5 to 70 cm) long. Sensory organs, the cephalic papillae, are reduced in size and number. The eggs in a

TABLE 1. Common spiruroid parasites of domestic animals

Parasite	Definitive host	Intermediate host	Distribution	Pathology
Ascarops strongylina (Rudolphi, 1819)	Swine	Beetles	Cosmopolitan	Inflammation of stomach mucosa
Cheilospirura hamulosa (Diesing, 1851)*	Fowl, turkey	Grasshoppers, beetles	N. America, S. America, Asia, Europe	Soft nodules in musculature of gizzard
Dispharynx nasuta (Rudolphi, 1819)	Fowl, turkey, pigeon, guinea fowl, pheasant	Isopods	N. America, S. America, Europe, Asia	Ulcers and gland destruction in proventriculus
Echinuria uncinata (Rudolphi, 1819)	Duck, goose, swan	Water fleas	N. America, Europe, Asia, N. Africa	Inflammation and nodules in proventriculus and gizzard
Draschia megastoma (Rudolphi, 1819)	Equines	*Musca* sp.	Cosmopolitan	Tumors in stomach wall
Draschia microstoma (Schneider, 1866)	Equines	Stable fly, housefly	Cosmopolitan	Catarrhal gastritis, stomach ulcers
Habronema muscae (Carter, 1861)	Equines	*Musca* sp.	Cosmopolitan	Catarrhal gastritis
Oxyspirura mansoni (Cobbold, 1879)	Fowl, turkey, peafowl	Cockroaches	Southern U.S., S. Asia, Australia, S. America	Lesions in eyes, blindness, loss of eyeball
Tetrameres americana (Cram, 1927)	Fowl, turkey	Grasshoppers, cockroaches	N. America, S. America, Africa, Asia, Europe	Irritation and inflammation of proventriculus
Thelazia californiensis (Price, 1930)	Sheep, cat, dog	*Fannia thelazia*	California	Scar tissue formation in eyes, blindness
Thelazia callipaeda (Railliet and Henry)	Dog, rabbit	Unknown	Far East	Scar tissue formation in eyes, blindness
Thelazia rhodesii (Desmarest, 1828)	Cattle, sheep, goat, buffalo	*Musca* sp.	Europe, Asia, Africa	Scar tissue formation in eyes, blindness

*Occasional parasite of humans.

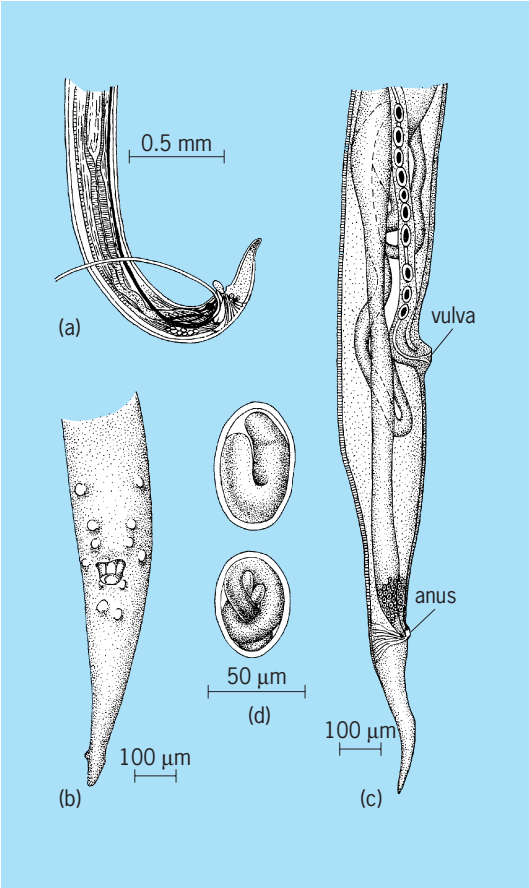

Fig. 2. ***Oxyspirura mansoni*, a common parasite of poultry. (*a*) Tail of male, lateral view; (*b*) tail of male, ventral view; (*c*) tail of female; (*d*) eggs. (*Parasite Laboratory, Agricultural Research Service, Beltsville, Maryland*)**

few species have thick shells but more commonly are covered only by a thin vitelline membrane. The eggs characteristically undergo partial embryonation in the uterus, and precociously active embryos (microfilariae) are discharged. The adult worms are found in a wide variety of places in the body of the vertebrate host, but each particular species has its preferred host and preferred location within that host (Table 2).

Microfilariae. These precociously active embryos are devoid of any alimentary canal (**Fig. 3**). The different species range from less than 100 to more than 700 micrometers long and 2 to 8 μm wide; they may or may not have a cephalic spear. The body is filled with a long column of somatic cells. Among them may be distinguished, with special staining, four cells variously termed R cells or G cells from which the alimentary canal and the reproductive system will develop. In some species the microfilariae are discharged still enclosed in the vitelline membrane and are termed "sheathed"; without the membrane they are "unsheathed." Most of the known forms have microfilariae in the blood regardless of the tissue in which the adults occur. These microfilariae may continue to circulate for months to years depending on the species and the level of host immune response. They do not necessarily circulate in constant numbers. Some species show a conspicuous daily periodicity, either nocturnal or diurnal, in the peripheral circulation. It has been shown with microfilariae of *Wuchereria bancrofti* of humans and with the unidentified microfilariae in the crow that the periodicity can be reversed in less than a week by reversing the active period (work, feeding, light, and so on) and the resting period (sleeping, darkness).

TABLE 2. Some common and better-known Filarioidea

Species	Definitive host	Location in the host	Disease
Wuchereria bancrofti			
Periodic	Human	Deep lymphatics	Elephantiasis
Aperiodic	Human	Deep lymphatics	Elephantiasis
Brugia malayi	Human	Deep lymphatics	Elephantiasis
B. timori	Human	Deep lymphatics	Elephantiasis
B. pahangi	Wild felines, human	Deep lymphatics	Elephantiasis
Onchocerca volvulus	Human	Subcutaneous lymphatics	Fibrous-nodules, blindness
O. cervicalis	Horse	Ligamentun nuchae	Slight irritation
Loa loa	Human	Subcutaneous tissues, eyes	Calabar swellings (nonpitting edema)
Stephanofilaria stilesi	Cattle	Skin	Dermatitis
Elaeophora schneideri	Mule deer, sheep, moose, elk	Distal arteries	Dermatitis, "clear-eyed blindness"
Dipetalonema perstans	Human	Peritoneal and pleural cavities	Minor allergies
D. streptocerca	Chimpanzee	Subcutaneous	Slight edema
D. reconditum	Dog	Subcutaneous tissues	Nonreported
Mansonella ozzardi	Human	Body cavities	Very little
Dirofilaria immitis	Dog	Right heart	Chronic debilitation (heartworm disease)
Dirofilaria repens	Dog	Subcutaneous tissues	Eczema
Litomosoides carini	Cotton rat	Pleural cavity	Pleural adhesions

Furthermore, an annual periodicity has been shown to occur with one species, the microfilariae of *Dirofilaria immitis*, in the dog, in the temperate zones; the microfilariae are most plentiful in the peripheral blood during the summer with its higher temperatures and longer periods of daylight. When the microfilariae are not in the peripheral circulation, they are sequestered in the capillary bed of internal organs, notably the lungs. They are incapable of further development as long as they remain in the definitive host.

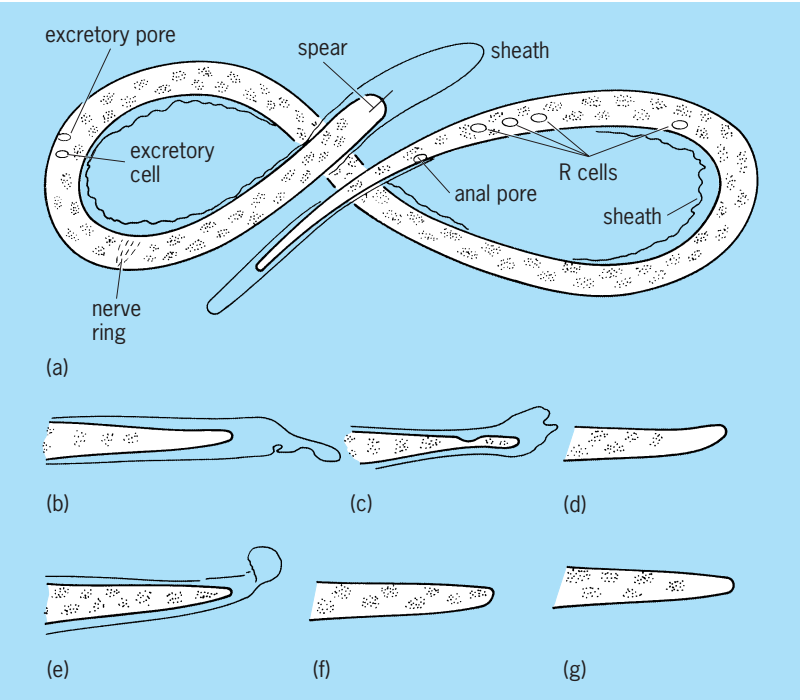

Fig. 3. Microfilariae. (*a*) *Wuchereria bancrofti*. (*b-g*) Posterior ends of the six species normally found in humans: (*b*) *W. bancrofti*; (*c*) *Brugia malayi*; (*d*) *Mansonella ozzardi*; (*e*) *Loa loa*; (*f*) *Dipetalonema* (*Acanthocheilonema*) *perstans*; (*g*) *Onchocerca volvulus*. Of these, *b-f* are found in blood, *g* is found in skin.

Life cycle. All the known filariae require a blood-sucking arthropod intermediate host, usually an insect and commonly a dipteran, in which to complete embryonation. The microfilariae are ingested as the arthropod feeds. After embryonation is completed, the resulting infective larvae gain entrance into the definitive vertebrate in association with the next feeding of the arthropod (**Fig. 4**).

Filariasis. Filariasis is a disease caused by filarioidea in humans or lower animals. More loosely, the term is used to indicate mere infection by such organisms. In human medicine, filariasis commonly refers to the disease caused by, or to infection with, one of the mosquito-borne, elephantoid-producing Filarioidea—most frequently *W. bancrofti*, less frequently *Brugia malayi*, and more recently including *B. timori*.

1. *Diagnosis.* Serological and intradermal tests do not give any clue to specific identification. The only specific laboratory aid to diagnosis is the detection and identification of the microfilariae. *Brugia timori* is a sheathed microfilaria, but the sheath does not stain with the common blood-parasite stain, Giemsa. This microfilaria has the two detached terminal nuclei seen in *B. malayi* but is nearly half again as long as the latter, with the cephalic space nearly twice

Intermediate host	Microfilariae			Geographical distribution
	Location	Sheath*	Periodicity	
"Domestic" night-biting mosquitoes	Blood	+	Nocturnal	Nearly worldwide in tropics
Sylvan day-biting mosquitoes	Blood	+	Slightly diurnal	South Pacific islands
Sylvan night-biting mosquitoes	Blood	+	Nocturnal	Southern Asia
Domestic night-biting mosquitoes	Blood	+	Nocturnal	Indonesia
Sylvan night-biting mosquitoes	Blood	+	Nocturnal	Southern Asia
Simulium	Skin	–	–	Tropical Africa, Mexico, Guatemala, Venezuela, Columbia
Culicoides	Skin	–	–	Cosmopolitan
Chrysops	Blood	+	Diurnal	Equatorial Africa
Unknown	Cutaneous lymph vessels	+	–	Southern and western U.S.
Unknown	Skin	–	–	Mountains of southwestern U.S.
Culicoides	Blood	–	–	Tropical Africa, South and Central America
Culicoides	Skin	–	–	Africa
Fleas, ticks (?)	Blood	–	–	Southern Europe, North Africa, U.S.
Culicoides	Blood			Central and South America, West Indies
Mosquitoes	Blood	–	Nocturnal and seasonal	Tropics and subtropics, temperate zone in U.S.
Mosquitoes	Blood	–	–	Southern Europe, North Africa, Vietnam
Mites	Blood	+	–	Southern U.S.

*+indicates presence of sheath.

as long. The microfilariae of *Dipetalonema streptocerca*, a subcutaneous parasite of chimpanzees, are unsheathed and nearly half again as long as those of *D. perstans*. However, the posterior terminal column of nuclei is more than twice as long (with 9–12 nuclei) as in the latter, and the posterior end is curved like a shepherd's crook. The parasite has been reported frequently in humans in the rainforests of western Africa. Over 40% infection with microfilariae was shown by one set of skin snips from a small group of natives in Ghana.

2. *Distribution and history.* The species *Wuchereria bancrofti* has worldwide distribution in the tropics and extends into both north and south temperate zones. However, the very high infection rates and the high prevalence of the resulting grotesque elephantiasis (characterized by cutaneous and subcutaneous tissue enlargement due to lymphatic obstruction) are localized at sea-level areas with year-round uncontrolled high mosquito densities in the tropics. Introduction into the temperate zones commonly results in a few secondary cases before the infection dies out. This parasite is important also in medical and biological history. A century ago Patrick Manson noted that the common occurrence of tropical elephantiasis was restricted to areas where *W. bancrofti* was common (with 30–50% microfilaremia) but that microfilaremia commonly did not occur in the older subjects with advanced disease symptoms. This was one of the earliest indications of host (immune) response to repeated parasitic infections. Manson's demonstration that the microfilariae after ingestion with a blood meal continued embryonic development within the mosquito vector was also the first incrimination of mosquito transmission of any disease-causing organism.

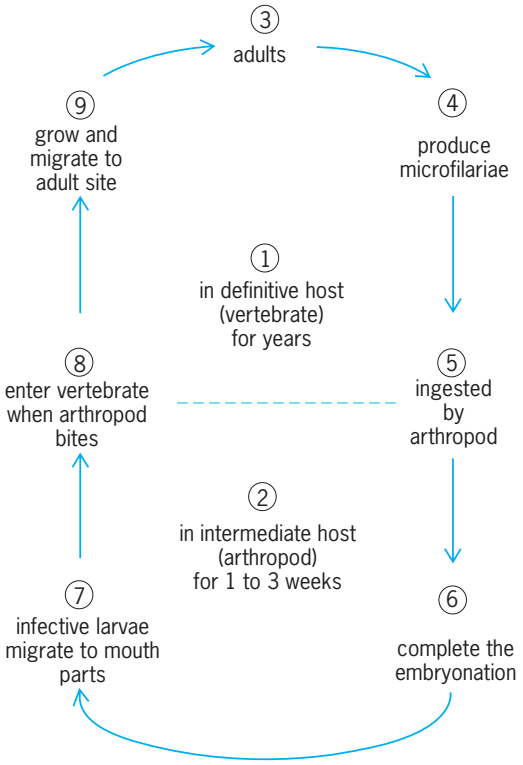

Fig. 4. Life cycles of the Filarioidea.

3. *Periodicity.* In most areas of the world, the microfilariae of *W. bancrofti* have a conspicuous nocturnal periodicity. There are 50 to 150 times as many microfilariae in the midnight blood as in noon blood. Often none may be found in midday blood. However, this well-defined periodicity has been completely reversed within a week by reversing the human quiescent period (sleep and darkness) and the active period (exercising, eating, and daylight). This periodic variety of *W. bancrofti* is transmitted by domestic, anthropophilous, night-biting mosquitoes, which readily enter houses to feed. *Culex fatigans* (*C. quinquefasciatus*) and some of the important malaria vectors such as *Anopheles darlingi, A. funestus*, and *A. farauti* are important vectors. Nevertheless, even in areas dominated by the periodic parasite, individual cases of the aperiodic form occur. (In 1906 such a case was found in the Philippines, and since it was considered that the case might involve a distinct species, the specific name *philippinensis* was coined.)

On some of the geologically newer islands of the South Pacific east of the 176th meridian (Samoa, Marquesas, Fiji, and the Gilbert and Ellice among them) the dominant mosquitoes are sylvan daytime feeders. Nocturnal mosquitoes are rare or absent. In this environment an evolutionary change has taken place. The microfilariae have either no periodicity or only a slight diurnal periodicity. *Aedes polynesiensis* and close relatives are the only vectors, and the parasite has almost completely lost its capacity to develop in *C. fatigans*. This parasite is evidently evolving into a new species. Perhaps it has already evolved into distinct species, despite the fact that the loss of nocturnal periodicity is the only unequivocal distinction. There is some suggestion that acute febrile attacks after repeated bites by the mosquito intermediate host are more common than with the periodic forms. At any event, to the earlier specific name, *philippinensis, pacifica* was added in 1932.

4. *Control.* The fundamental difference in the epidemiology of the two forms is clearly depicted by the control measures. The periodic form has been successfully reduced or eliminated by use of bednets, household screening, residential insecticides, and control of the readily detected domestic breeding sites. The use of community-wide household spraying to eliminate malaria from Guyana also eliminated transmission of *W. bancrofti*. Acquired resistance to insecticides has complicated this approach.

Neither attempts at larval destruction nor attacks on adult mosquitoes have had any success in controlling the aperiodic infection. Only communitywide medication has had any success. Diethylcarbamazine (Hetrazan) in various regima has reduced the pool of microfilariae for mosquito infection and is active against the infective larvae or early developmental stages.

Brugia malayi occurs in both periodic and aperiodic strains, but it does not appear to be as completely isolated as is the case with *W. bancrofti*. Another parameter is added by the fact that species of *Mansonia* are the principal vectors. The larval stages of these mosquitoes are attached to submerged vegetation; herbicides have some value in control. Diethylcarbamizine has also been used. It appears to be very successful in areas where *B. timori* has been introduced recently.

Onchocerciasis. This disease is caused by *Onchocerca volvulus* in the subcutaneous lymphatics. It is characterized by subcutaneous nodules which are most conspicuous where the skin lies closely over bony structures, via cranium, pelvic girdle, joints, and shoulder blades. When they are on the head, the microfilariae reach the eyes. Ocular disturbances vary from mild transient bleary vision to total and permanent blindness. This is a major problem in the provinces of Oaxaca and Chiapas in Mexico and the nearby Pacific slopes of Guatemala, where infection is an occupational hazard associated with coffee growing.

Termed "river blindness," onchocerciasis also affects field workers in the Ivory Coast, Ghana, Mali, Togo, and Upper Volta. The World Health Organization has estimated that there are over a million cases in Upper Volta alone. Attempts to control the blackfly vector have had only some local successes. Surgical removal of the nodules is curative and on a communitywide basis eliminates or reduces transmission.

Loa loa. The African eye worm, *Loa loa*, is the filarioidan worm most commonly acquired by Caucasian immigrants, including missionaries, in Africa. Transmission is by daytime-feeding sylvan deer-flies, *Chrysops*. The only preventive measures are protective clothing, including head nets. Repellents have some value. Fortunately, serious damage is rare even when the worm gets into the eye. The areas of pitting edema known as calabar swellings are painful and diagnostic. They commonly occur on the wrists, hands, arms, or orbital tissues. Adult worms move freely. When the worm is accessible, surgical removal is indicated. This is one of the few Filarioidea in which adult worms are killed by diethylcarbamazine. *See* NEMATA.

Gilbert F. Otto

Bibliography. H. W. Brown and F. A. Neva, *Basic Clinical Parasitology*, 6th ed., 1996; L. S. Garcia and D. A. Bruckner, *Diagnostic Medical Parasitology*, 4th ed., 2001; P. E. Sartwell (ed.), *Maxcy-Rosenau's Preventive Medicine and Public Health*, 11th ed., 1980.

Spitzer Space Telescope

A high-performance infrared telescope that is one of the four Great Observatories launched by the National Aeronautics and Space Administration (NASA) between 1990 and 2003. It takes advantage of dramatic advances in infrared detectors that have occurred over the last 20 years; it utilizes modern detector arrays in space, where they are limited only by the faint glow of the zodiacal dust cloud. Ground-based infrared telescopes can operate only at the wavelengths where the atmosphere is transparent, lying

between 1 and 25 micrometers. Even within these windows, the thermal emission of the atmosphere is more than a million times greater than the dilute emission of the zodiacal cloud; there is additional foreground thermal emission from the telescope itself. High-sensitivity detectors are blinded by these bright foreground signals. Operating in space eliminates the atmospheric absorption and emission; also, a telescope in the vacuum of space can be cooled sufficiently to virtually eliminate its emission. *See* ZODIACAL LIGHT.

Spitzer was known for most of its life as *SIRTF* (*Space Infrared Telescope Facility*). *SIRTF* was launched on August 25, 2003, and was renamed after Lyman Spitzer, Jr., the first prominent astronomer to advocate putting telescopes into space.

Development. The *Infrared Astronomy Satellite* (*IRAS*) and the *Infrared Space Observatory* (*ISO*) were important predecessors of the *Spitzer Telescope*. *IRAS*, launched in 1983, had four sets of infrared detectors, operating at wavelengths of 12, 25, 60, and 100 μm. The detectors and a 60-cm (24-in.) telescope were mounted inside a liquid-helium Dewar vessel that maintained them at a temperature of ~2 K (−456°F) for the 10-month mission, allowing a very sensitive all-sky survey. The European Space Agency built *ISO* to follow up on the *IRAS* survey. *ISO* too had a 60-cm telescope mounted in a liquid helium Dewar. It carried four instruments designed for detailed study of individual objects through imaging, photometry, and spectroscopy. *ISO* was launched in 1995 and had a 28-month mission that returned a wealth of data over the 3–200-μm spectral range. *See* DEWAR FLASK.

The *Spitzer Telescope* was also conceived to follow up on *IRAS*. Although its science team was selected in 1984, the mission was repeatedly delayed. Throughout this period, the mission was revived repeatedly with new technical concepts. Finally, in 1996 the telescope got started officially toward construction under the direction of the Jet Propulsion Laboratory. By that time, NASA had downsized all its science missions under the guidelines of its "faster-better-cheaper" philosophy. Spitzer had been descoped by about a factor of 6 in cost from the original plans.

Design. The approved concept introduces a number of innovations developed in response to the pressure to reduce cost. It uses an Earth-trailing orbit to get far from the thermal radiation of the Earth (**Fig. 1**). The instruments are cooled inside a Dewar containing liquid helium, as with *IRAS* and *ISO*. However, the telescope is mounted on the outside of this Dewar and was launched warm. The very cold environment away from the Earth allows the outer shell of the satellite to cool passively, greatly reducing the heat load on the telescope. The thermal loads on the helium Dewar are also very small. This concept of a telescope launched warm that cools by radiating into space provides the technical foundation for future, larger infrared telescopes such as the *James Webb Space Telescope (JWST)*.

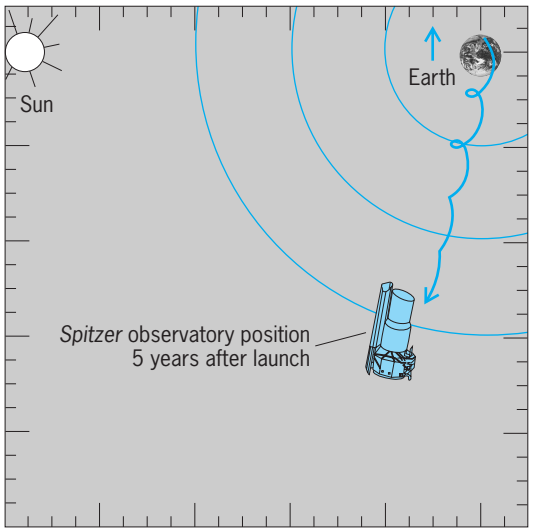

Fig. 1. The orbit of *Spitzer*. The Earth orbits the Sun in a counterclockwise direction in this figure, and the frame of reference rotates with the Earth-Sun line to show how *Spitzer* trails behind. Viewed in a solar frame of reference, the *Spitzer* orbit is a normal Keplerian ellipse. In the Earth-centered frame here, the observatory makes a *loop* or *kink* every year when it partially overtakes the Earth. The observatory will fall farther and farther behind during its 5-year mission. The circles centered on the Earth are 0.2, 0.4, and 0.6 astronomical unit in radius, respectively.

The 85-cm (33.5-in.) telescope and 360-liter (95-gallon) Dewar are held off the spacecraft by a thermally insulating truss, and a thermal shield blocks radiation from the spacecraft (**Fig. 2**). The solar panel is cantilevered off the spacecraft, and the telescope is protected from its heat by another thermal shield. The Dewar protected the liquid helium from thermal loads while the observatory was on the ground. The instruments are mounted on top of the helium vessel, and a vacuum-tight aperture door sealed out the air when on the ground. Upon reaching orbit, the aperture door was opened to allow light from the telescope to reach the instruments. The outer shell cooled to 34 K (−398°F) during the first month of the mission. Venting the cold helium vapor from the Dewar through heat exchangers cools the telescope further, to about 6° above absolute zero, or 6 K (−449°F). The extremely efficient thermal design is expected to yield a 5-year lifetime at these temperatures.

Spitzer has three instruments. The Infrared Array Camera (IRAC) images in bands at wavelengths of 3.6, 4.5, 5.8, and 8 μm. The Infrared Spectrograph (IRS) provides spectra from 5 to 40 μm at low resolution (1–2%) and from 10 to 38 μm at moderate resolution (about 0.15%). The Multiband Imaging Photometer for *Spitzer* (MIPS) provides imaging at 24, 70, and 160 μm. The instruments use advanced infrared detector arrays that operate at or near the fundamental natural background limit in space. The IRAC arrays are 256 × 256 pixels in format, using indium antimonide (InSb) photodiodes for the two shorter bands and silicon:arsenic (Si:As) impurity band conduction (IBC) devices for the

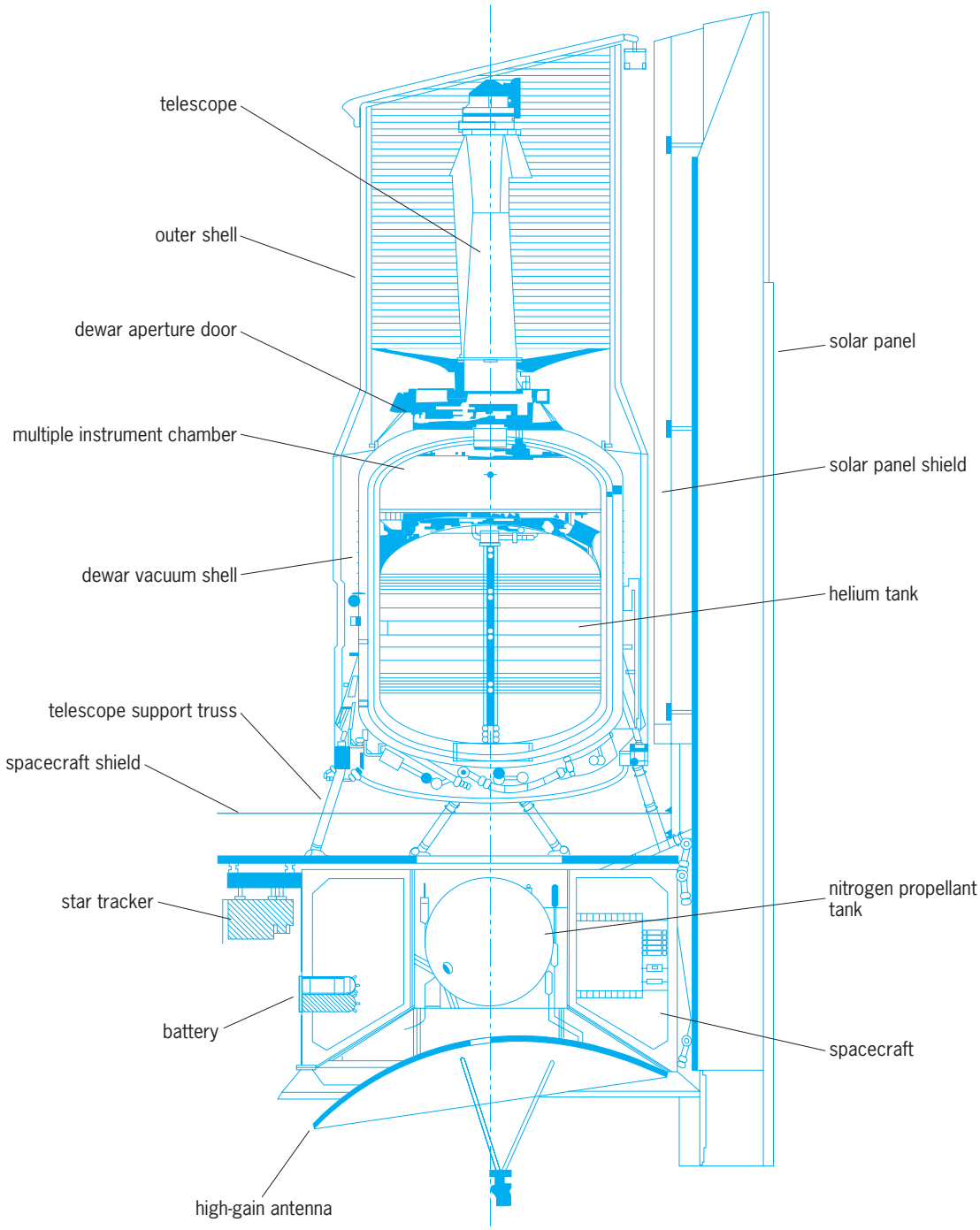

Fig. 2. Cutaway drawing of *Spitzer*, showing the major subsystems.

longer ones. IRS uses 128 × 128 arrays of Si:As and silicon:antimony (Si:Sb) IBC devices. MIPS uses a similar Si:As array at 24 μm and germanium:gallium (Ge:Ga) arrays for the two longer bands. The high performance of these detector arrays allows *Spitzer* to deliver its anticipated breakthrough in science capability, despite its history of descopes. *ISO* had about 50 times more detectors than *IRAS*, and *Spitzer* has about 100 times more than *ISO*. The sensitivity per detector has also improved by an order of magnitude with each new mission.

Objectives. *Spitzer* is operated as a general-access observatory from the Jet Propulsion Laboratory and a science center at the California Institute of Technology. Eighty percent of the observing time is assigned by competitive application from the astronomical community. The science program includes investigations of star formation, planetary systems, and the assembly and evolution of galaxies.

Star formation. Young stars form in dense interstellar molecular clouds, carrying a heavy burden of interstellar dust. Initially, forming stars are cold, only tens

of degrees above absolute zero. Even after gravitational contraction has warmed them to more typical stellar temperatures, they are hidden by the surrounding cocoons of interstellar gas and dust and can be seen only in the infrared. *Spitzer* is mapping many nearby star-forming regions to identify very new stars that are too cold and faint to be known from other types of observation, documenting the very first stages of star formation. It is finding luminous massive stars and the outflows of gas that occur as accretion stops on these objects. It is also locating very-low-mass substellar brown dwarfs, revealing the processes by which stellar-mass clumps in molecular clouds can fragment into much smaller pieces before collapsing. *See* BROWN DWARF.

Figure 3*a* shows a star-forming region about 14,000 light-years away that has a classic structure. Within the core of a molecular cloud, clumps of gas have collapsed into massive stars. Winds from these stars have cleared away the remaining gas, punching a hole into the cloud. From the *Spitzer* images, more than 300 newly formed stars have been identified in the central hole and surrounding molecular cloud. The remains of the cloud are heated by the young stars and glow at wavelengths of 4.5, 5.8, and 8 μm in this image. *See* MOLECULAR CLOUD; STELLAR EVOLUTION.

Planetary systems. Embryo planets are believed to form during the first few million years of the life of their star, from a protoplanetary disk made of dusty and gaseous material with too much angular momentum to fall into the star itself. This stage in the planet formation process stops as the disk is dissipated by evaporation of grains of material within it, by accretion of material into the star or the embryo planets, and by ejection of material from the system. Thereafter, planets grow by accretion of material from the embryos that is released when they collide with each other. A by-product of these collisions is circumstellar disks of dust and debris. The dust grains in the circumstellar disk are warmed by the star and glow in the infrared, where *Spitzer* can detect them. Initial *Spitzer* measurements of debris disks show a broad variety of behavior, indicating a few cases that appear to be dominated by recent and dramatic collisions. The disk signals damp down after about 150 million years. These observations are reminiscent of theories for the formation of the planets in the solar system, and in particular of the extreme bombardment of the Earth over the first few hundred million years of its existence. *See* EXTRASOLAR PLANETS; INTERPLANETARY MATTER; PLANET; PROTOSTAR; SOLAR SYSTEM.

Figure 3*b* and *c* shows *Spitzer* images of Fomalhaut, a bright star about 25 light-years away and about 200 million years old. The image of the star itself has been removed from the figure panels. Figure 3*b*, at a wavelength of 70 μm, shows the dusty debris from a shattering collision of asteroid-sized embryo planets. A *Hubble Space Telescope (HST)* image of scattered light from this ring shows it to be off-center relative to the star, possibly maintained in this configuration by the gravitational action of a massive planet. Because the ring is closer to the star in the lower left of the 70-μm image (Fig. 3*b*), the dust is warmer there and glows more brightly than elsewhere. At a wavelength of 24 μm (Fig. 3*c*), the ring is filled in with warmer dust that is probably falling into the star. This dust is very dilute and is not seen in the *Hubble* image.

Assembly and evolution of galaxies. In nearby galaxies, the shortest-wavelength *Spitzer* bands reveal the distribution of the stars that dominate the visible mass,

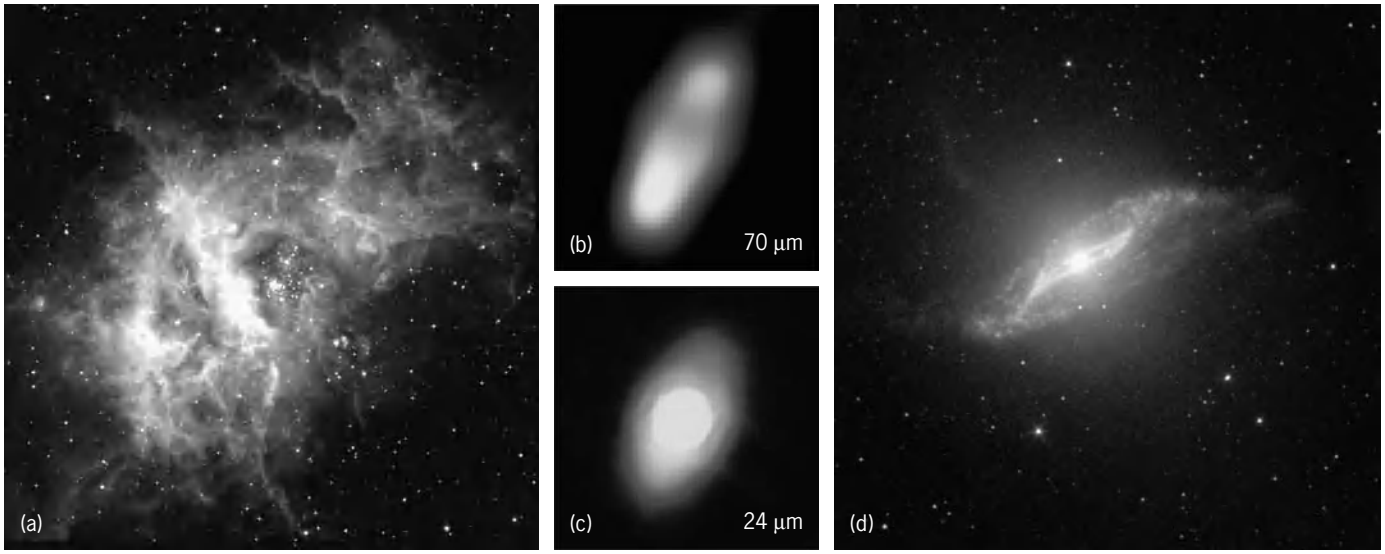

Fig. 3. Illustrative science results from the *Spitzer* observatory. (*a*) Star formation in the nebula RCW 49 (*courtesy of NASA/JPL-Caltech; E. Churchwell, University of Wisconsin*). (*b*) Fomalhaut circumstellar disk at wavelengths of 70 μm and (*c*) 24 μm (*courtesy of NASA/JPL-Caltech; K. Stapelfeldt, JPL*). (*d*) Dusty, elliptical galaxy Centaurus A (*courtesy of NASA/JPL-Caltech; J. Keene, Spitzer Science Center & Caltech*).

improving understanding of the structure of the galaxies. The longer-wavelength bands glow in the infrared emission of various forms of interstellar dust, from tiny aromatic hydrocarbon powder heated to hundreds of degrees from the absorption of a single ultraviolet photon, to large grains at about 20 K (−424°F).

Figure 3*d* shows how *Spitzer* can see into the nearest large elliptical radio galaxy, Centaurus A, about 10 million light-years away. A striking trapezoidal structure glows in the 8-μm band, which traces the distribution of polyaromatic hydrocarbon powder. A computer model indicates that this beautiful structure remains from a spiral galaxy that has been engulfed by the larger elliptical galaxy, a process that may have been responsible for triggering the nuclear activity and radio emission.

In the distant universe, *Spitzer* is locating a population of very luminous, dusty galaxies at redshifts of $z = 1$–4. These galaxies are identified because of their very red colors in the IRAC bands (resulting from the combination of high redshift and absorption by interstellar dust) or their detection in the MIPS bands, along with information on the redshift from spectra, or inferred from the IRAC colors. They represent a stage when smaller galaxies were merging at a high rate to form what have become large elliptical galaxies and spiral galaxy bulges. *Spitzer* is also detecting galaxies and quasars at even higher redshifts and will contribute to knowledge of the most distant objects known in the universe. *See* GALAXY, EXTERNAL; GALAXY FORMATION AND EVOLUTION; QUASAR. George Rieke

Bibliography. J. C. Mather (ed.), *Optical, Infrared, and Millimeter Space Telescopes*, Proc. SPIE, vol. 5487, 2004; G. H. Rieke, *The Last of the Great Observatories: Spitzer and the Era of Faster, Better, Cheaper at NASA*, University of Arizona Press, Tucson, 2006; *Spitzer Space Telescope* Mission (special issue), *Astrophys. J. Suppl. Ser.*,154(1):1–474, September 2004.

Splachnales

An order of the true mosses (subclass Bryidae), whose members are remarkable perennial plants that grow mainly on nitrogenous substrates, such as dung, and show considerable differentiation of neck tissue below the spore-bearing part of the capsule. Many of the species show adaptations that encourage insect dispersal of spores. The order consists of two families with about eight genera.

The plants are gregarious or dense-tufted, erect, and often forked. The leaves are soft, lanceolate or obovate, sometimes bordered at the margins, and commonly toothed. The single costa may be excurrent or terminate at or below the apex. The perichaetia are terminal, with bracts much like the upper stem leaves. The setae are elongate and the capsules erect with a noticeably differentiated neck. Stomata are very numerous in the neck, and have two guard cells. A single peristome is usually present, entire or forked, with 16 teeth sometimes paired or joined in twos and fours. In *Splachnum* the teeth are chambered since segments of the inner peristome lie opposite the teeth and are fused with them. The calyptrae are generally mitrate and sometimes hairy. *See* BRYIDAE; BRYOPHYTA; BRYOPSIDA. Howard Crum

Spleen

An organ of the circulatory system present in most vertebrates, lying in the abdominal cavity usually in close proximity to the left border of the stomach. It exhibits wide variation in size, shape, color, and location, depending upon species and age.

Anatomy. In humans the spleen normally measures about 1 × 3 × 5 in. (2 × 7 × 12 cm) and weighs less than $^1/_2$ lb (230 g). It is a firm organ with an oval shape and is indented on its inner surface to form the hilum, or stalk of attachment to the peritoneum. This mesentery fold also carries the splenic artery and vein to the organ. The spleen is composed of a complex pulp lying between fibrous partitions, or trabeculae. These partitions project inwardly from the dense fibrous capsule that surrounds the organ.

Blood circulation is unusual in that the splenic arterioles open into thin-walled dilations, called sinusoids, which in turn drain into small veins. This arrangement of distensible blood-filled vessels, clusters of cells, and developing follicles accounts for the pulpy character of the spleen.

The pulp consists of a white pulp and a red pulp. The white pulp (lymphatic tissue) forms irregular nodules or follicles, often called Malpighian corpuscles and consisting of aggregations of developing lymphocytes. These nodules form initially around the small branches of the splenic arteries. In addition to the lymphocytes, other white cells such as monocytes, histiocytes, and giant cells are found in quantity. The red pulp consists of the sinusoids filled with blood and of splenic or Billroth cords filling the spaces between the sinusoids. Lymphocytes are found in the splenic cords that merge gradually into the white pulp. The distribution and amount of white and red pulp vary greatly, depending upon certain infections, intoxications, and disturbances in blood cell formation (anemia and leukemia) in the body. *See* ANEMIA; LEUKEMIA.

Function. The spleen is an important part of the blood-forming, or hematopoietic, system; it is also one of the largest lymphoid organs in the body and as such is involved in the defenses against disease attributed to the reticuloendothelial system. This system consists of many different types of cells in the body having the power to neutralize or engulf foreign particles or bacteria and sometimes to act as filters through which blood or lymph must pass. *See* HEMATOPOIESIS.

Although the chief functions of the spleen appear to be the production of lymphocytes, the probable

formation of antibodies, and the destruction of worn-out red blood cells, other less well-understood activities are known. For example, in some animals it may act as a reservoir for red blood cells, contracting from time to time to return these cells to the bloodstream as they are needed. In the fetus and sometimes in later life, the spleen may be a primary center for the formation of red blood cells. *See* ANTIBODY.

Another function of the spleen is its role in biligenesis. Because the spleen destroys erythrocytes, it is one of the sites where extrahepatic bilirubin is formed. Bilirubin is the principal pigment of bile formed by the reduction of biliverdin. Bile is normally present in the feces. *See* BILIRUBIN.

Lower vertebrates. In the lower vertebrates splenic tissue appears as scattered lymphoid masses in the wall of the digestive tract and apparently acts as the primary blood-forming organ. Bone marrow appears in certain amphibians and in most higher forms, taking over the formation of red cells in adults under normal conditions. The spleen then produces cells which are principally of the lymphoid type and engages in a number of other activities, some of which have been mentioned above.

This organ is not necessary to life for any vertebrate because its activities may be shared or assumed by other organs under proper conditions. Many of the splenic functions, such as lymphocyte formation, are common to all lymphoid tissue. Splenectomy may result in a temporary depression of antibody formation or by a recrudescence of latent or low-grade infection. *See* SPLEEN DISORDERS.

Walter Bock

Spleen disorders

The spleen is rarely the site of primary disorders except those of vascular origin, but it is frequently involved in systemic inflammations, metabolic diseases, and generalized blood disorders.

Among vascular disturbances, acute and chronic congestion are prominent, particularly chronic congestion caused by cardiac failure, cirrhosis of the liver, and obstruction of the blood flow from the spleen by thrombi, scarring, or tumor tissue. Obstruction of the splenic artery or its branches by thrombi may result in an infarct caused by either cardiac or blood disease.

Inflammations include acute and chronic forms. The characteristic engorgement of blood often causes a marked enlargement of the organ. Bacteremias frequently produce this enlargement, or splenomegaly, and inflammation, but any severe infectious disease such as diphtheria or pneumonia may do so. Tuberculosis, syphilis, typhoid fever, and malaria, as well as many other specific infections, may cause splenomegaly, often with characteristic gross or microscopic changes for each disease. *See* MALARIA; SYPHILIS; TUBERCULOSIS.

The spleen is involved in certain types of pigment metabolism and is frequently the site of specific kinds of degeneration when amyloidosis, hemosiderosis, or argyrosis is present.

Similarly, in disorders of lipid metabolism, abnormal kinds or amounts of fat may appear in connection with Gaucher's disease, the hereditary Neimann-Pick disease, and others. *See* LIPID METABOLISM.

Because the spleen normally aids in the destruction of worn-out red blood cells and, under certain conditions, in the formation of red cells, it is not uncommon to find splenic involvement in blood disorders such as sickle-cell anemia, congenital hemolytic jaundice, polycythemia, and Mediterranean anemia. *See* HEMATOLOGIC DISORDERS.

The leukemias, especially when of the lymphocytic or neutrophilic varieties, cause some of the most prominent cases of splenomegaly as well as other changes. In myeloid leukemia, for instance, spleen weight of 13–18 lb (6000–8000 g) is not rare and the spleen may fill the entire abdomen. *See* LEUKEMIA.

Hypersplenism is an obscure disorder in which one or more of the blood cell types are destroyed to excess. Primary hypersplenism results from unknown causes; secondary hypersplenism may follow inflammatory diseases, chronic congestion, or tumor invasion.

Tumors originating in the spleen are rare and usually limited to such benign growths as hemangiomas, lymphangiomas, and fibromas, but malignant lymphomas and lymphosarcomas also occur.

Secondary tumors, which originate elsewhere and metastasize to the spleen, are not uncommon, particularly the lymphoma group. Other varieties are seen less often and include carcinomas, particularly from a bloodstream invasion or growth in the neighboring stomach or intestine. *See* TUMOR.

Trauma to the spleen is more common than suspected, and the consequences of rupture of the organ are often tragic because this may follow an apparently moderate abdominal injury.

Many other conditions produce splenic change, notably splenomegaly of some degree. The presence of an enlarged spleen indicates that a thorough diagnostic evaluation should be made. *See* SPLEEN.

Edward G. Stuart; N. Karle Mottet

Splines

A series of projections and slots used instead of a key to prevent relative rotation of cylindrically fitted machine parts. Splines are several projections machined on the shaft; the shaft fits into a mating bore called a spline fitting. Splines are made in two forms, square and involute (see **illus.**). Since there are several projections (integral keys) to share the force in transmitting power, the splines can be shallow, thereby not weakening the shaft as much as a standard key.

Square splines have 4, 6, 10, or 16 splines. The external part (shaft) may have the splines formed by milling and the internal part (bore) by broaching. Three classes of fits are used: sliding (as for gear

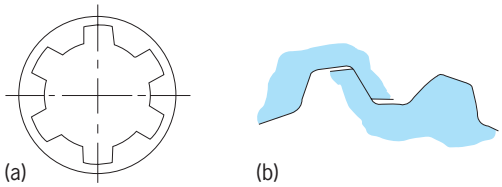

Diagrams of (*a*) square spline and (*b*) involute spline profile.

shifting) under load, sliding when not loaded, and permanent fit. Square splines have been used extensively for machine parts. In the automotive industry, square splines have been replaced generally by involute splines, which cost less to make accurately for good fit and interchangeability.

Involute splines are used to prevent relative rotation of cylindrically fitted machine parts and have the same functional characteristics as square splines. The involute spline, however, is like an involute gear, and the spline fitting (internal part) is like a mating internal gear. Profiles are the same as for gear teeth of the stub (fractional pitch) form with 30° pressure angle.

Involute splines on the shaft are generated by a hob or a gear shaper, and internal splines are formed by a broach or a gear shaper. Three classes of fit are standard: sliding, close, and press.

Involute serrations are similar to involute splines except that the pressure angle is 45° and, while there are three standard fits (loose, close, and press), serrations are usually press-fitted and used for permanent assembly. They are used for both parallel and tapered shafts. *See* MACHINE KEY. Paul H. Black

Bibliography. American National Standards Institute (ANSI), *Involute Splines and Inspections*, *Inch Version*, B92.1-1970; ANSI, *Involute Splines and Inspections*, *Metric Version*, B92.1M-1970; ANSI, *Involute Splines*, *Metric Modules*, B92.2M-1981; C. W. Ham et al., *Mechanics of Machinery*, 4th ed., 1958; C. R. Mischke and J. E. Shigley, *Mechanical Engineering Design*, 6th ed., 2000; G. H. Ryder and M. D. Bennnett, *Mechanics of Machines*, 2d ed., 1990.

Spodumene

The name given to the monoclinic lithium pyroxene $LiAl(SiO_3)_2$. Spodumene commonly occurs as white to yellowish prismatic crystals, often with a "woody" appearance, exhibiting the 87° pyroxene (110) cleavages. It is easily identified during heating in a flame by the red color given off, accompanied by a marked swelling of the fragment. Spodumene usually contains an appreciable quantity of hydrogen substituting for lithium. At 720°C (1330°F), spodumene inverts to a tetragonal form, β-spodumene, which is accompanied by a 30% increase in volume. Spodumene is capable of forming immense crystals in nature. A single crystal 47 ft (14.3 m) in length and 5 ft (1.5 m) in diameter, as well as others almost as large, has been found at the Etta mine in South Dakota. This implies the remarkable ability of a single crystal to replace a large variety of preexisiting minerals and yet maintain the integrity of a single crystal, a crystal growth unequaled elsewhere in nature.

Spodumene is usually found as a constituent in certain granitic pegmatites in association with quartz, alkali feldspars, mica, beryl, phosphates, and a large variety of rare minerals. It is also known to occur as disseminated grains in some granite gneisses. Spodumene often alters to a fibrous mass composed of eucryptite $LiAlSiO_4$ and albite, or eucryptite and muscovite. The emerald-green variety, hiddenite, and a lilac variety, kunzite, are used as precious stones. Spodumene from pegmatites is used as an ore for lithium. *See* LITHIUM; PYROXENE. George W. DeVore

Spongiomorphida

An extinct group of genera of Mesozoic fossils. Informally known as the spongiomorphs, they were established as a family of scleractinian corals by Fritz Frech in 1890 on the basis of specimens from the Zlambach beds of late Triassic age in Austria. He grouped the fossils into four genera (*Spongiomorpha, Heptastylopsis, Heptastylis*, and *Stromatomorpha*). The fossils consist of closely spaced longitudinal rods of calcite or aragonite that are perpendicular to the growth surface and that give off lateral expansions which join the rods and in some forms unite to form laminae parallel to the growth surface (see **illustration**). Frech noted the similarity of his group of fossils to the stromatoporoids, particularly to the genus *Actinostroma*, and named one of the species *Heptastylis stromatoporoides*. In tangential section, the rods of the spongiomorphs have a central dark dot, if round, or a line, if oval, and concentric microstructure. Their fibrous microstructure suggests that they

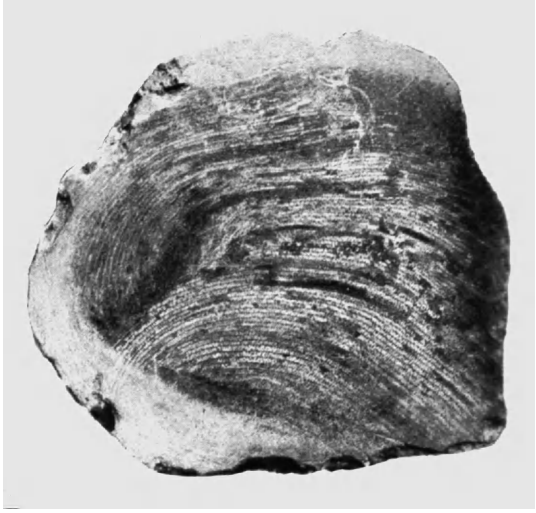

Spongiomorpha acyclica **(Triassic, Austria), length 8 cm. (*From J. P. Smith, Upper Triassic Marine Invertebrate Faunas of North America: Spongiomorphs, U.S. Geol. Surv. Prof. Pap. 1927*)**

were originally aragonite, and their water-jet appearance in longitudinal section is similar to that of the septa of scleractinian corals. Some spongiomorphs from the middle Triassic Cassian beds still preserve the aragonite mineralogy of the living corals.

In some spongiomorphs, the rods in tangential sections seem to be grouped in sixfold symmetry about a central rod with their axial dark lines radiating away from the central rod. This radial pattern was thought by some investigators to be homologous to the astrorhizal systems of the Paleozoic stromatoporoids. To those who believed the stromatoporoids were hydrozoans, this radial arrangement was taken as evidence that the spongiomorphs were hydrozoans. However, the radial pattern is not produced by convergent canals, as in the stromatoporoids, but by the alignment of rods. Further, the hexameral radial symmetry is typical of the Mesozoic corals, not of hydrozoans. Frech recognized the similarity of his specimens to the scleractinian coral genus *Astraeomorpha* of the family Poritidae. In this family of living corals, the septa tend to break down into a series of rods, and the boundaries of the corallites disappear to produce a structure very similar to the spongiomorphs. The affinity of Frech's genera to the poritid corals has been confirmed by the work of P. Gautret and J.-P. Cuif on the microstructure of these fossils.

The term spongiomorph should thus be used only in an informal sense for the group of genera in which the septa are reduced to closely spaced rods and the radial structure and individuality of the corallites is obscure.

Spongiomorph corals are known from the Middle Triassic of the Alps; the Upper Triassic of Austria, Bulgaria, Romania, the United States (California, Alaska, Oregon), Turkey, Morocco, Iran, Japan, south China, central Asia, and Tajikistan; the Upper Jurassic of China (Tibet), Austria, Spain, and Japan; and the Upper Cretaceous of Sweden and the West Indies. *See* SCLERACTINIA; SCLEROSPONGE; SPHAERACTINOIDEA; STROMATOPOROIDEA. C. W. Stearn

Bibliography. E. Flügel and E. Sy, Die Hydrozoen der Trias, *Neues Jahrbuch für Geologie und Paläontologie, Abhandlungen*, 109:1–108, 1959; P. Gautret, F. Ezzoubair, and J.-P. Cuif, Recherche sur les affinités des spongiomorphidae Frech, 1890—I: Charactéristiques microstructurales et minérologiques de *Spongiomorpha acyclica* Frech, 1890, *Geobios*, 25(3):345–355, 1992; J. P. Smith, Upper Triassic marine invertebrate faunas of North America: Spongiomorphs, *U.S. Geol. Surv. Prof. Pap.*, 141:132–135, 1927; G. D. Stanley and M. T. Whalen, Triassic corals and spongiomorphs from Hell's Canyon, Wallowa Terrane, Oregon, *J. Paleontol.*, 63:800–819, 1989.

Sporozoa

A subphylum of Protozoa, typically with spores. The spores are simple and have no polar filaments. There is a single type of nucleus. There are no cilia or flagella except for flagellated microgametes in some groups. In most sporozoa there is an alternation of sexual and asexual stages in the life cycle. In the sexual stage, fertilization is by syngamy, that is, the union of male and female gametes. All sporozoa are parasitic. *See* PROTOZOA.

Taxonomy. The subphylum is divided into three classes—Telosporea, Toxoplasmea, and Haplosporea.

Telosporea. The class Telosporea is by far the largest and most important. It contains two subclasses: members of the subclass Gregarinia, which is composed of three orders (Archigregarinida, Eugregarinida, and Neogregarinida), are parasites of insects and other invertebrates; members of the subclass Coccidia, which is composed of two orders (Protococcida and Eucoccida), are parasites mostly of vertebrates, although a few are found in invertebrates. *See* INSECT DISEASES; TELOSPOREA.

The order Eucoccida is the largest in the subclass Coccidia. Two of its suborders (Eimeriina and Haemosporina) are of great importance. The suborder Eimeriina contains several hundred species of coccidia, which cause disease in various domestic and wild animals. Most of them multiply in the cells of the intestine and cause diarrhea, dysentery, and even death. They are monoxenous, multiplying asexually by schizogony in their host cells and eventually producing gametes which unite to form oocysts that pass out in the feces, mature on the ground, and infect other animals when eaten. A typical species is *Eimeria tenella*, which occurs in the ceca of chickens and may cause their death.

The suborder Haemosporina includes the malaria parasites and related forms. Their life cycle is basically similar to that of the Eimeriina, but they generally multiply asexually in vertebrate blood cells and produce sexual stages in invertebrate hosts. A typical species is *Plasmodium vivax*, which causes benign tertian malaria in humans (called benign because it seldom kills its victims, in contrast to *P. falciparum*, which causes malignant tertian malaria and has a relatively high death rate). *See* MALARIA.

Spores are present in the Telosporea, and reproduction is typically both sexual and asexual, with alternation of generations. Some species occur only in a single type of host (monoxenous), while others require two types of host (heteroxenous), with sexual development in one and asexual in the other. Some groups have flagellated microgametes; if these are missing, locomotion is by body flexion or gliding; pseudopods are ordinarily absent, but if present are used for feeding, not locomotion. The name Telosporea comes from Greek words which indicate that spore formation occurs at the end of the life cycle.

Toxoplasmea. The class Toxoplasmea contains "sporozoa" which lack spores and sex. There are no pseudopods or flagella at any stage, and locomotion is by body flexion or gliding. Cysts or pseudocysts (with a wall formed by the host cell) are formed and contain many trophozoites. Species in this class have a single type of host.

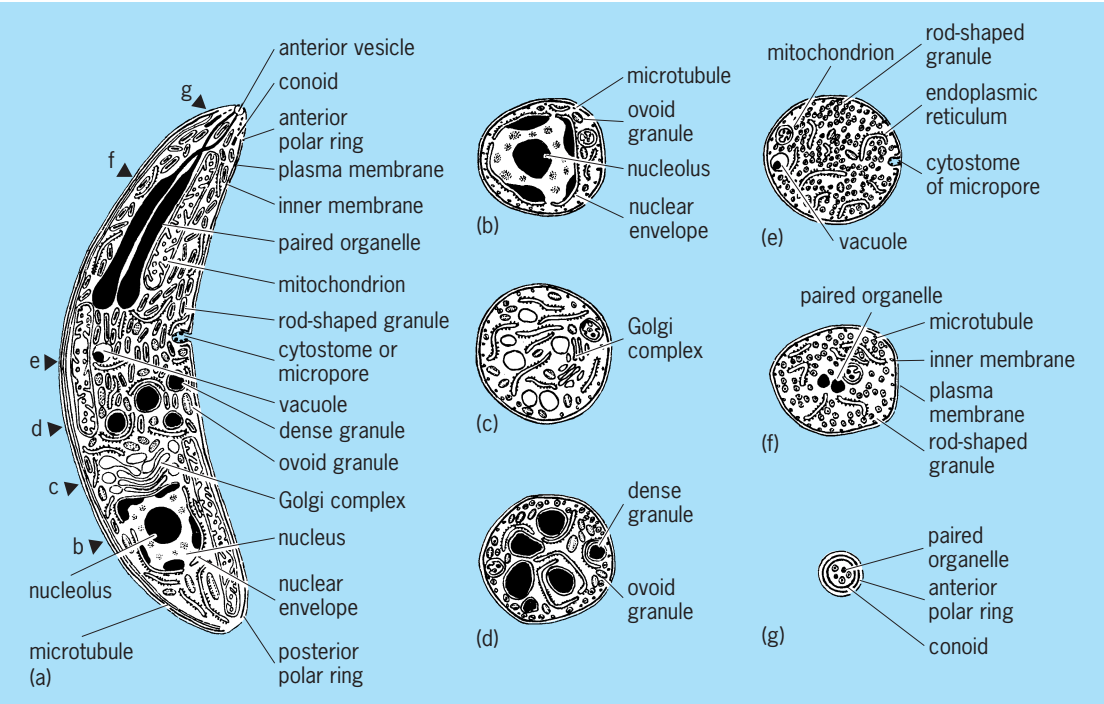

Fig. 1. Fine structure of a coccidian merozoite (*Eimeria miyairii* from the rat). (*a*) Longitudinal section. Cross sections at levels of (*b*) nucleus, (*c*) Golgi complex, (*d*) large dense granules, (*e*) cytostome, (*f*) paired organelle, and (*g*) conoid. (*After J. Andreassen and O. Behnke, J. Parasitol., 54:150–163, 1968*)

There seem to be only three genera in this class, all of which parasitize vertebrates. *Toxoplasma* occurs mostly in mammals. By far the most important species is *T. gondii*, which is common throughout the world and causes toxoplasmosis in humans and many lower mammals. Most infected individuals show no symptoms, but the disease has various manifestations ranging up to acute, fatal encephalitis in newborn infants. Its normal mode of transmission is still unknown. *Sarcocystis* is found in many mammals and birds, especially sheep, cattle, swine, and ducks. The cysts, which may be visible to the naked eye, are found in the muscles. Hunters sometimes see them in wild ducks, where they make the breast muscles appear striated. They are harmless when cooked and eaten, but they are esthetically undesirable, and heavily infected ducks are rarely eaten. The third genus, *Besnoitia*, is relatively uncommon, although one species may cause a skin disease of cattle and another may pit the bones of reindeer. *See* TOXOPLASMEA.

Haplosporea. It is not certain whether the last class, Haplosporea, really belongs in this subphylum. Its members have spores but apparently do not reproduce sexually. They have no flagella, but pseudopods may be present. Some people think that this class belongs with the Cnidospora, even though its members lack a polar filament, while others think it may be related to the fungi. At any rate, there are only a few species in the group, and they are mostly parasites of mollusks, annelids, and other invertebrates. *See* CNIDOSPORA; HAPLOSPOREA.

Structure. The structure of the Sporozoa varies with the species and stage of development, and the variations are too diverse to be given here. At one stage or another all (except perhaps the Haplosporida) have a characteristic conoid and polar ring at the anterior end. Subpellicular fibrils or microtubules extend almost the whole length of the body just beneath the pellicle and are presumably responsible for locomotion. There is a Golgi apparatus which presumably has to do with secretion or excretion, one or more mitochondria which have to do with metabolism, a number of micronemes, rod-shaped granules, convoluted tubules, sarconemes or lankesterellonemes extending backward in the anterior part of the organism, a paired organelle or series of toxonemes originating inside the conoid and

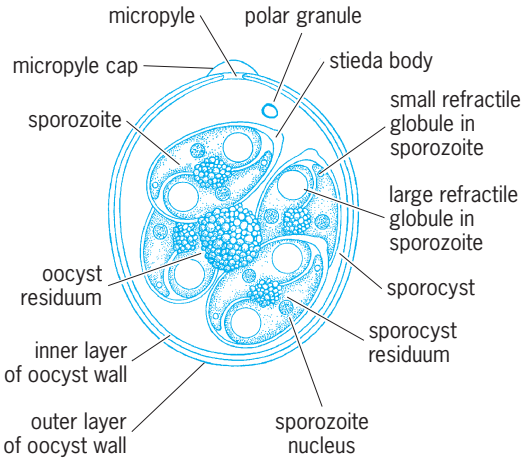

Fig. 2. Mature (sporulated) oocyst of a coccidium (*Eimeria*). (*After N.D. Levine, Protozoan Parasites of Domestic Animals and Man, Burgess, 1961*)

running backward in the cytoplasm, one or more micropores (cytostomes, micropyles) in the side of the body, several types of granule in the cytoplasm, and endoplasmic reticulum (**Fig. 1**). The electron microscope is needed to see most of these structures.

The Telosporea have an oocyst which is produced by fusion of male and female zygotes. Its structures can be seen with the light microscope. The characters of the oocyst vary with the species. In the Eimeriina there are generally one or more sporocysts within the oocyst, and one or more sporozoites in each sporocyst. The oocyst may have a micropyle, which may be capped. There may be a polar granule within the oocyst and its wall may be composed of one or more layers. There may be residual material in the oocyst and sporocyst, left over after formation of the sporocysts or sporozoites (**Fig. 2**).

Life cycle. The life cycle of a typical telosporean sporozoon involves alternation of sexual and asexual generations. The sexual forms may occur in one type of host and the asexual forms in another, or there may be only a single type of host. The latter situation is characteristic of the eimeriin Eucoccida. A typical example is *Eimeria tenella*, which lives in the epithelial cells of the chicken cecum (**Fig. 3**). When a chicken eats a mature (sporulated) oocyst, the oocyst breaks in the gizzard, releasing the sporozoites. A sporozoite (stage 1 in Fig. 3) enters an intestinal epithelial cell (2) and rounds up and turns into a first-generation schizont (3), which forms a number of merozoites (4) by multiple fission. These break out of the host cell (5), enter new host intestinal epithelial cells (6), and turn into second-generation schizonts (7, 8), which produce second-generation merozoites (9, 10), which break out of the host cell (11), and may form either third-generation schizonts (12, 13), which produce third-generation merozoites (14, 15), or gamonts. All third-generation schizonts produce gamonts. The gamonts are either male or female. The male ones (16, 17) produce a large number of flagellated microgametes (18), while the female ones simply grow, becoming macrogametes (19, 20). The microgametes fertilize the macrogametes, and the resultant zygote (21) rounds up, secretes a wall around itself to become an immature oocyst, and leaves the host cell (22). It passes out of the chicken and becomes mature (sporulates) on the ground. The sporont undergoes meiosis and becomes haploid, throws off a polar body, and forms four sporoblasts (23), each of which forms a sporocyst containing two sporozoites. When this sporulated oocyst is eaten by the chicken, the sporozoites are released in the gut and the cycle begins again.

Eucoccida. In the suborder Eimeriina, since each species has a definite number of asexual generations, the infection is self-limiting. This is not true in the suborder Haemosporina, which contains the malaria parasites. In this group the number of asexual generations is indefinite, a moderate number of microgametes is formed, asexual stages occur in a vertebrate host, and sexual stages and sporozoite formation occur in an invertebrate, usually an insect. In human malaria, for instance, the vertebrate host is a human and the invertebrate host an *Anopheles* mosquito. The vertebrate host is infected when the invertebrate host bites it. In the third suborder of this group, the Adeleina, few microgametes are formed and they are produced when the male and female gamonts are attached to each other, that is, when they are in syzygy.

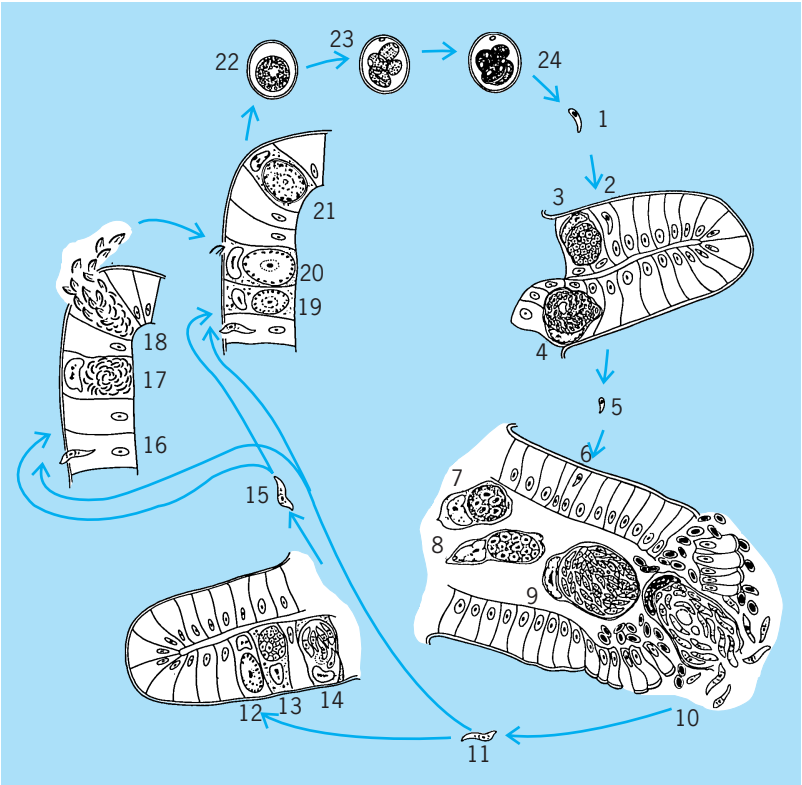

Fig. 3. Life cycle of the chicken coccidium *Eimeria tenella*. The numbered steps are explained in the text. (*After N. D. Levine, Protozoan Parasites of Domestic Animals and Man, Burgess, 1961*)

Gregarinia. In the second large telosporean subclass, the Gregarinia, the mature trophozoites (vegetative stages) are large and extracellular. They occur in the digestive tract or body cavity of invertebrates. Typically (in the order Eugregarinida), no schizonts are formed, the trophozoites producing gamonts and then gametes, which fuse to form a zygote which produces an oocyst and then a number of sporozoites.

Toxoplasmea. In the class Toxoplasmea there is no sexual reproduction. The trophoxoites divide by external or internal binary or multiple fission, and cysts or pseudocysts with many naked trophozoites are formed.

Norman D. Levine

Bibliography. J. O. Corliss, *The Ciliated Protozoa: Characterization, Classification, and Guide to the Literature*, 2d ed., 1979; P. C. C. Garnham, *Malaria Parasites and Other Haemosporidia*, 1966; R. R. Kudo, *Protozoology*, 5th ed., 1977; N. D. Levine, *Protozoan Parasites of Domestic Animals and of Man*, 2d ed., 1973; N. D. Levine, *Textbook of Veterinary Parasitology*, 1978; N. D. Levine et al., A newly revised classification of the protozoa, *J. Protozool.*, 27:35–58, 1980; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Sports medicine

A branch of medicine concerned with the effects of exercise and sports on the human body, including treatment of injuries. Sports medicine can be divided into three general areas: clinical sports medicine, sports surgery, and the physiology of exercise. Clinical sports medicine includes the prevention and treatment of athletic injuries and the design of exercise and nutrition programs for maintaining peak physical performance. Sports surgery is also concerned with the treatment of injuries from contact (human or object) sports. Exercise physiology, a growing field of sports medicine, involves the study of the body's response to physical stress. It comprises the science of fitness, the preservation of fitness, and the role of fitness in the prevention and treatment of disease.

The various aspects of sports medicine touch on many different branches of medicine and surgery, including cardiology, neurology, orthopedics, pulmonary medicine, and endocrinology. Nevertheless, sports medicine is becoming a distinct discipline.

Advances in instrumentation also led to increased understanding of cardiac function and its response to exercise. The pressure drum and sphygmograph mechanically transmit and record arterial pulses. The modern electrocardiogram and other techniques now give information on the mechanical and electrical properties of heart action at rest and during exercise.

Investigations of the effects of exercise on systems other than the heart and lungs were initiated in the 1930s. The notion of exercise as a route to good health became widespread during the latter part of the nineteenth century.

Diagnosis and therapy. The health problems arising from sports include injuries resulting from overuse, such as ill-defined aches and pains, tendinitis, and chondromalacia (softening of the cartilage), one cause of "runner's knee." Such injuries are becoming increasingly common as more people engage in amateur athletics; they also afflict some musicians who develop overuse of muscle-tendon units resulting in pain of the hand and fingers, commonly seen in string players. Sports medicine also treats acute ailments such as dehydration and heat stroke, injury to the musculoskeletal system, the eyes, the brain and spinal cord, and (rarely) cardiac arrest. The usual therapy for overuse injuries is rest. Injuries caused by contact can require urgent surgical care. *See* PERFORMING ARTS MEDICINE.

Strategies to prevent injuries and maintain peak performance have become an important part of sports medicine. Careful attention to diet and fluid intake is routine for professional and amateur athletes during both training and sports events. (The use of performance-enhancing drugs, such as steroids and stimulants, by some athletes is an unfortunate result of the competition for improved athletic performance.) Studies of biomechanics have led to preventive measures such as improved helmets, shoes, and supportive clothing. *See* BIOMECHANICS; STEROID.

Exercise physiology. This subdivision of sports medicine is involved with the responses of healthy individuals to physical stress and aging in all parts of the body. Hormonal, cardiovascular, renal, neurologic (including mental function), musculoskeletal, and other bodily systems are studied. Modern technologies are applied, because many are noninvasive and allow repeated measurements.

Exercise and wellness. The role of sports in the prevention of disease and the improvement of health is a major concern of sports medicine. It is widely believed that physical fitness can help ward off many of the illnesses common in industrial society, mainly obesity and cardiovascular disease but also osteoporosis (brittleness of bones) and such chronic complaints as back pain and headaches. In addition, it is believed that a regimen in which regular exercise is accompanied by beneficial habits, such as a diet that is high in fiber and low in fats, moderate consumption of alcohol, and an avoidance of smoking, can lower the incidence of some illnesses, including cancers.

Conclusive evidence of exercise's capacity to prevent disease is lacking, but the circumstantial evidence is strong. Weight loss, lowered cholesterol, and improved neurologic, gastrointestinal, and general body function all seem to contribute to this effect.

Otto Appenzeller

Bibliography. O. Appenzeller (ed.), *Sports Medicine*: *Training*, *Fitness*, *Injuries*, 3d ed., 1988; E. Jokl, Running, psychology and culture, *Ann. N. Y. Acad. Sci.*, 301:970–1001, 1977.

Spot welding

A resistance-welding process in which coalescence is produced by the flow of electric current through the resistance of metals held together under pressure. A low-voltage, high-current energy source is required (**Fig. 1**). Usually the upper electrode moves and applies the clamping force. Pressure must be maintained at all times during the heating cycle to prevent flashing at the electrode faces. Electrodes are

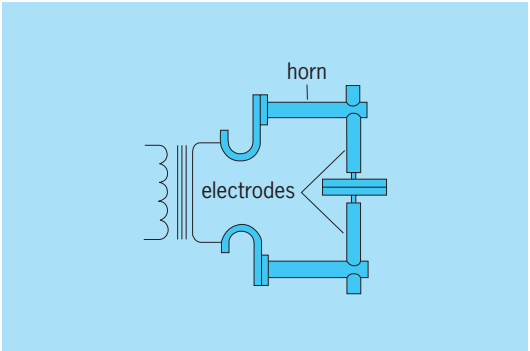

Fig. 1. Spot-welding circuit. When electrodes are closed on the workpiece, the circuit is completed for welding.

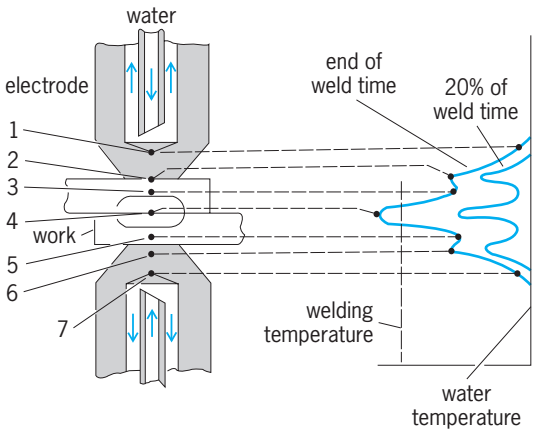

Fig. 2. Distribution of temperature in local (numbered) elements of a spot-welding operation.

water-cooled and are made of copper alloys because pure copper is soft and deforms under pressure. *See* RESISTANCE WELDING.

The electric current flows through at least seven resistances connected in series for any one weld. They are (1) upper electrode, (2) contact between upper electrode and upper sheet, (3) body of upper sheet, (4) contact between interfaces of sheets,(5) body of lower sheet, (6) contact between lower sheet and electrode, and (7) lower electrode (**Fig. 2**). Heat generated in each of the seven sections will be in proportion to the resistance of each. The greatest resistance is at the interfaces in 4, and heat is most rapidly developed there. The liquid-cooled electrodes 1 and 7 rapidly dissipate the heat generated at the contact between electrodes and sheets 2 and 6, and thus contain the metal that is heated to fusion temperature at the interfaces.

After the metals have been fused together, the electrodes usually remain in place sufficiently long to cool the weld. An exception is in welding quench-sensitive metals, where it is desirable to remove the electrodes as soon as possible to allow the heat to be conducted in the surrounding metal, preventing steep quench gradients. *See* WELDING AND CUTTING OF MATERIALS. Eugene J. Limpel

Bibliography. L. Jeffus, *Welding: Principles and Applications*, 5th ed., 2002; H. Zhang and J. Senkara, *Resistance Welding*, 2005.

Spray flow

A special case of a two-phase (gas and liquid) flow in which the liquid phase is the dispersed phase and exists in the form of many droplets. The gas phase is the continuous phase, so abstract continuous lines (or surfaces) can be constructed through the gas at any instant without intersection of the droplets. The droplets and the gas have velocities that can be different, so both phases can move through some fixed volume or chamber and the droplets can move relative to the surrounding gas. *See* GAS; LIQUID; PARTICULATES; TWO-PHASE FLOW.

Applications. Spray flows have many applications. Sprays are used to introduce liquid fuel into the combustion chambers of diesel engines, turbojet engines, liquid-propellant rocket engines, and oil-burning furnaces. They are used in agricultural and household applications of insecticides and pesticides, for materials and chemicals processing, for fire extinguishing, for cooling in heat exchangers, for application of medicines, and for application of coatings (including paint and various other types of layered coatings). Common liquids (such as water, fuels, and paints) are used in sprays. It is sometimes useful to spray uncommon liquids such as molten metals. In the various applications, the approximately spherical droplets typically have submillimeter diameters that can be as small as a few micrometers. *See* METAL COATINGS.

Formation. Sprays are formed for industrial, commercial, agricultural, and power generation purposes by injection of a liquid stream into a gaseous environment. In addition, sprays can form naturally in a falling or splashing liquid. Injected streams of liquid tend to become unstable when the dynamic pressure (one-half of the gas density times the square of the liquid velocity) is much larger than the coefficient of surface tension divided by the transverse dimension. Typically, the liquid stream disintegrates into ligaments (coarse droplets) and then into many smaller spherical droplets. The breakup (or atomization) process is faster at higher stream velocity, and the final droplet sizes are smaller for higher stream velocities. Spray droplet sizes vary and typically are represented statistically by a distribution function. The number of droplets in a spray can be as high as a few million in a volume smaller than a liter. *See* ATOMIZATION; JET FLOW; SURFACE TENSION.

Drag force on droplets. The droplets experience a drag force as they move relative to the gas. The magnitude and direction of the drag depend on the magnitude and direction of the relative velocity between the gas and the droplet. The drag force is proportional to the square of the droplet diameter, while the droplet mass is proportional to the cube of the diameter; therefore, the acceleration, which equals the drag per unit mass, is inversely proportional to the diameter. Drag causes the relative velocity to decrease faster for smaller droplets. Droplets of only a few micrometers in size move at the gas velocity. *See* FLUID-FLOW PRINCIPLES.

Flow in and around droplets. A droplet moving relative to the surrounding gas behaves similarly to an aerodynamic body. An observer fixed to the droplet would see that the gas flow first decelerates as it moves toward the droplet and then accelerates as it passes the droplet. The **illustration** shows a cross section of the flow around and within a droplet. The frictional force at the gas-liquid interface retards the flow in the gas layer at the droplet surface. The approximate thickness of this boundary layer is proportional to the droplet radius divided by the square root of the Reynolds number. The nondimensional Reynolds number is the triple product of the gas

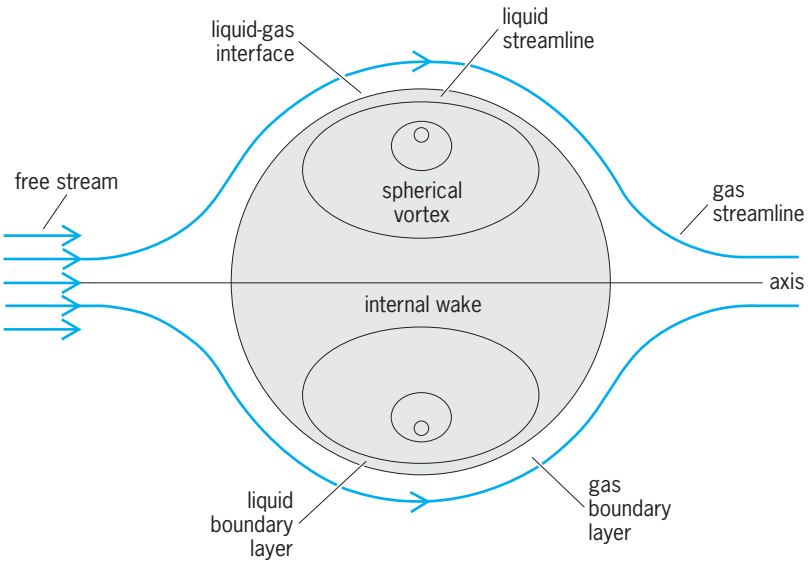

Cross section of a vaporizing droplet, showing the relative gas-droplet motion and internal circulation.

density, relative velocity, and droplet radius divided by the coefficient of gas viscosity. For droplets, the range of the Reynolds number is 0–200. The flow stream in the boundary layer can separate behind the droplet surface, causing a wake with recirculating flow in the gas. The wake formation causes a pressure drop across the droplet. This pressure difference and the friction force result in the drag force. The friction force exerted on the liquid surface of the droplet causes the internal liquid to circulate. The streamlines in the illustration show the spherical vortex pattern that emerges, with recirculation through a region called the internal wake. The convective current affects the transport of heat to the droplet interior from its surface region. *See* BOUNDARY LAYER FLOW; REYNOLDS NUMBER; WAKE FLOW.

Droplet trajectories. The trajectories of the droplets in a spray depend on the aerodynamic forces (lift and drag) and the force of gravity that they experience. The aerodynamic forces on a droplet are modified by the proximity of neighboring droplets. The drag force is reduced for a droplet moving in tandem behind another droplet. The wake behind the first droplet is moving in the same direction as the second droplet, thereby reducing the relative velocity and the drag force. (This same phenomenon occurs for trucks moving in a convoy.) Droplets moving side by side can experience lift forces, causing them to either attract or repel each other. The lift force is caused by the modification of the pressure due to the Bernoulli effect as the gas flows between the two droplets. As the two droplets come closer, the gas velocity first increases, reducing the inside pressure below the outside pressure, thereby pushing the droplets toward each other. As the droplets come still closer, blockage of the inside flow occurs, increasing the inside pressure above the outside pressure and pushing the droplets apart. The aerodynamic forces can cause droplets to collide but, generally, the droplets in a spray are diverging. The conduction and convection of heat and the diffusion of vapors are also affected by neighboring droplets. *See* AERODYNAMIC FORCE.

Heat transfer and vaporization. In many applications (such as combustors and heat exchangers), a high-temperature environment exists in the gas surrounding each droplet of a spray. Substantial heat transfer to the liquid surface results; some of the heat is required to vaporize the liquid at the surface, and the remainder of the heat is transferred to the droplet interior, thereby raising its temperature with time. The droplet lifetime is defined as the time required to vaporize the total droplet after its introduction to the hot environment. The droplet temperature is lower than the ambient gas temperature, with continuous increase in the gas temperature through a thermal gas layer surrounding the droplet. An increase in the relative velocity, or equivalently in the Reynolds number, causes the thickness of the thermal layer to decrease, thereby increasing the heating rate and the vaporization rate and decreasing the droplet lifetime. *See* HEAT TRANSFER.

The temperature of a rapidly vaporizing droplet (for example, a fuel droplet in a combustor) will vary throughout its lifetime in a transient manner. The internal liquid circulation provides convective augmentation to the conduction of heat from the droplet surface toward the interior. A temperature variation through the liquid interior can remain during the lifetime of the droplet. For a droplet vaporizing slowly in a low-temperature gas environment, the droplet temperature can be approximately uniform through the interior at the theoretical maximum temperature, which is called the wet-bulb temperature. A submillimeter-size water or fuel droplet in a gas environment with a temperature above 1000 K (1340°F) can vaporize in milliseconds. Vaporization time varies with droplet diameter to a power between 1.5 and 2.0.

The vapor that comes from the droplet is typically different chemically from the surrounding gas. It must convect and diffuse away from the droplet and mix with the surrounding gas. Sometimes the vapor also reacts with the surrounding (for example, fuel vapor in a heated air environment). *See* EVAPORATION.

Supercritical behavior. At very high pressures and temperatures, the distinction between a vapor and its liquid disappears. That is, above the critical temperature and critical pressure values, the addition of heat results in a continuous change in density rather than the discontinuous phase change (vaporization) that is experienced at lower pressures and temperatures. In many applications, a liquid stream is injected into a chamber at supercritical pressure. If the initial liquid temperature is subcritical but the pressure through the gas film and the ambient gas temperature are supercritical, heating of the liquid can result, with a continuous change of density as the temperature increases from subcritical to supercritical values. The gas mixture in the film surrounding the droplet can have substantially higher critical pressures than

either the pure vapor or the ambient gas. Sometimes the gas film is subcritical even if the ambient gas is supercritical. Vaporization, with a discontinuous density change, can occur in this situation. As the liquid droplet approaches the critical temperature, surface tension forces are substantially reduced. Therefore, distortion of the droplet shape and secondary atomization become more likely. *See* CRITICAL PHENOMENA; SUPERCRITICAL FLUID.

William A. Sirignano

Bibliography. A. Frohn and N. Roth, *Dynamics of Droplets*, Springer-Verlag, New York, 2000; A. H. Lefebvre, *Atomization and Sprays*, 1989; W. A. Sirignano, *Fluid Dynamics and Transport of Droplets and Sprays*, 1999; Z. Zapryanov and S. Tabakova, *Dynamics of Bubbles, Drops and Rigid Particles*, vol. 50 of *Fluid Mechanics and Its Applications*, 1999.

Spread spectrum communication

A means of communicating by purposely spreading the spectrum (frequency extent or bandwidth) of the communication signal well beyond the required bandwidth of the data modulation signal. Spread spectrum signals are typically transmitted by electromagnetic waves in free space, with usage in both nonmilitary and military systems.

Motivation for using spread spectrum signals is based on the following:

1 Spread spectrum systems have the ability to reject hostile as well as unintentional jamming by interfering signals so that information can be communicated.

2 Spread spectrum signals have a low probability of being intercepted or detected since the power in the transmitted wave is "spread" over a large bandwidth or frequency extent.

3 Since these signals cannot be readily demodulated without knowing the code and its precise timing, a level of message privacy is obtained.

4 The wide bandwidth of the spread spectrum signals provides tolerance to multipath (reflected waves that typically take longer to arrive at the receiver than the direct desired signal so that the two can be distinguished).

5 A high degree of precision in ranging (distance measuring) can be obtained by using one type of spread spectrum signaling called direct sequence, with applications to navigation.

6 Multiple access, or the ability to send many independent signals over the same frequency band, is possible in spread spectrum signaling. *See* BANDWIDTH REQUIREMENTS (COMMUNICATIONS); COMMUNICATIONS SCRAMBLING; ELECTRICAL INTERFERENCE; ELECTRONIC WARFARE; RADIO-WAVE PROPAGATION.

There are four generic types of spread spectrum signals: direct sequence (DS) or pseudonoise (PN), frequency hopping (FH), linear frequency modulation (chirp), and time hopping (TH). The first two methods are much more commonly used today than the other two.

Chirp modulation. Chirp modulation is a spread spectrum method which was developed for radar use in the mid-1940s. The basic idea is to transmit a long rectangular pulse whose carrier frequency is linearly increased. The frequency-modulated signal returned from the target passes through a filter in the receiver at a velocity of propagation proportional to frequency. The result is a pulse that is much shorter in time duration than the transmitted pulse with a larger peak power. Unchirped pulses such as interference or jamming pulses do not "compress" at the receiver, so that this method yields a processing gain or advantage for the chirped signal. *See* RADAR.

Direct sequence systems. Direct sequence modulation is characterized by phase-modulating a sine wave by an unending string of pseudonoise code chips (symbols of much smaller time duration than a bit). This unending string is typically based on a pseudonoise code that generates an apparently random sequence of code chips that repeats only after the pseudonoise code period. If T_b is the duration of a bit expressed in seconds, then typically the duration of a pseudonoise code chip, T_c, would be much smaller than T_b, so that many chips would occur during one bit (typically from hundreds to millions). An important parameter in direct sequence systems is the processing gain (PG), which is given by the ratio T_b/T_c. Larger values of processing gain give greater immunity to interference.

Transmitter. In many systems both the data and the pseudonoise code phase-modulate the carrier by 0° or 180°, which is called binary phase-shift keying (BPSK). In a basic binary phase-shift keyed, pseudonoise-encoded spread spectrum system (**Fig. 1***a*), digital data representing the information to be transmitted are binary phase-shift keyed onto the carrier. Then the pseudonoise code

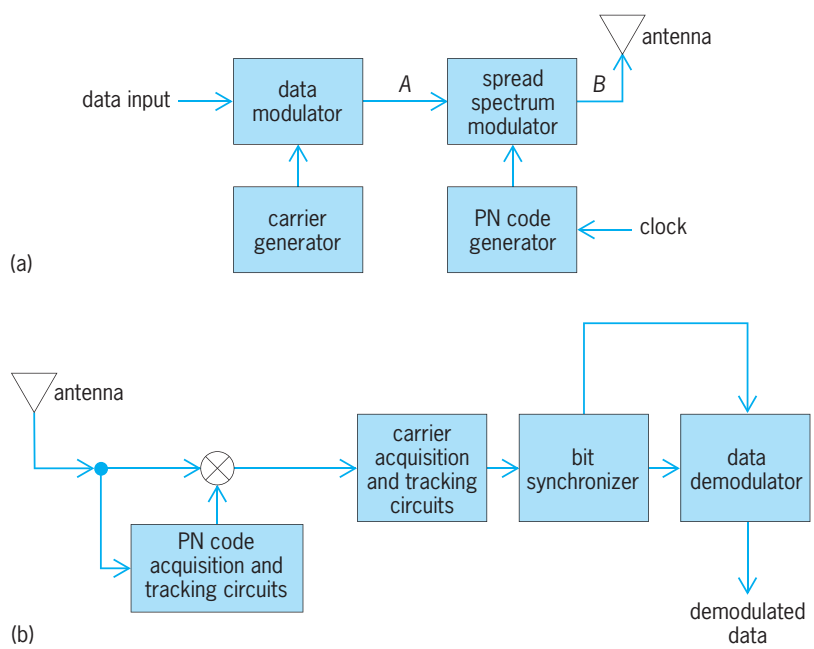

Fig. 1. Equipment for a basic direct-sequence (pseudonoise-encoded), binary phase-shift keyed, spread spectrum system. (*a*) Transmitter. (*b*) Receiver.

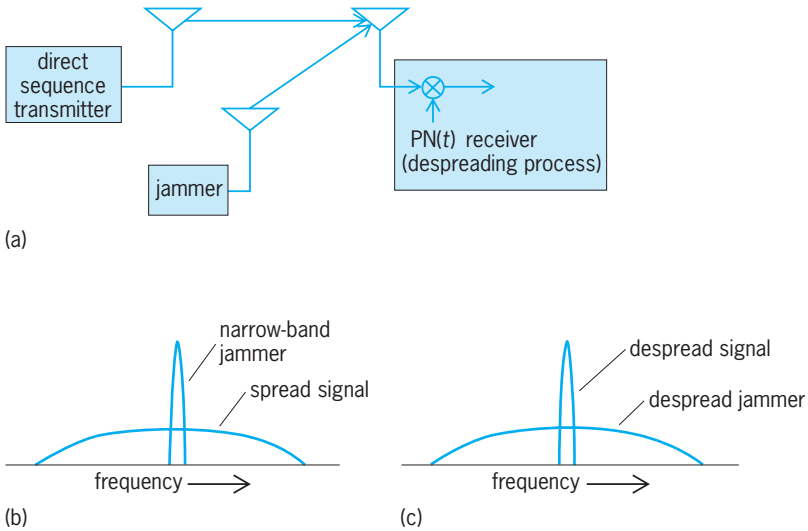

Fig. 2. Example of how a narrow-band jammer is despread in the front end of a receiver. (*a*) System model. (*b*) Transmitted signal and jammer. (*c*) Despread signal and jammer.

generator also binary phase-shift keys the carrier, and the composite signal is transmitted. *See* ELECTRICAL COMMUNICATIONS; MODULATION; PHASE MODULATION.

Receiver. For the receiver to function, it must first align the receiver pseudonoise code generator timing with the received pseudonoise code signal timing (Fig. 1*b*). This is accomplished in two distinct steps, code acquisition and code tracking. Code acquisition is process of searching in time by correlating the locally generated code against the received signal plus noise. The search continues until the correlation process produces an output above a predetermined threshold. This occurs either when the local and received codes are in time alignment or there is a false alarm due to noise. In either case the code loop is engaged and given a fixed length of time to track the received code. When true tracking occurs, the received signal is said to be despread, leaving only the data to be demodulated.

The code tracking circuit, which is called a delay lock loop, tracks the code (keeps the local receiver code aligned with the received code) once acquisition has been achieved. When the received code is being tracked, it can be multiplied by the locally generated pseudonoise code which then removes the code on the received signal.

At this point a BPSK-data-modulated carrier signal remains to be demodulated. In order to do this, a carrier tracking loop must be utilized, such as a Costas loop or a squaring loop. Their function is to track the carrier and thereby allow phase-coherent demodulation of the data, that is, multiplication by a signal exactly in phase with the received carrier, which leaves only the baseband data to be detected.

A bit synchronizer then tracks (provides bit timing for) the data bit sequence, and a data demodulator utilizes a matched filter to detect the data. The matched filter for binary phase-shift keyed data is a T_b-second integration over each bit time. The output of the integrator produces either a positive or a negative output, which is then interpreted as a binary 1 or a 0. This bit sequence of 0's and 1's determines the message being sent. The data bits transmitted are either 1 or −1, which maps to either a binary 0 or a 1. If the signal-to-noise ratio is sufficiently large, very few errors will occur in the bit detection process.

The above discussion has obviously been oversimplified. For example, the reconstructed pseudonoise code is never perfectly aligned with the received code due to noise and filtering effects, and the locally generated carrier likewise is not exactly in phase with the received carrier for similar reasons. Degradations therefore occur to the system when compared to an idealized system. Other modulations besides BPSK are used.

Protection against jamming. It is instructive to see how the effects of a narrow-band jammer (narrow-band compared to the spreading-code bandwidth) are diminished at the receiver due to the despreading action in a direct sequence system. A jamming signal, which interferes with the reception of the desired signal, may be intentional or unintentional. **Figure 2** illustrates a simplified direct-sequence transmitter and receiver including a signal with a narrow-band jammer. For this example it is assumed that the data and direct spread spectrum signal are BPSK-modulated onto the carrier and transmitted via the transmitter antenna. A jammer transmits its narrow-band jammer in the direction of the receiver antenna. At the receiver the sum of the signal and the jammer are captured at the receiver antenna and despread by a synchronous despreading signal, PN(t). The process of synchronously despreading (mathematically multiplying) the received signal returns the signal to the prespread bandwidth. However, the process of despreading the jammer, since it is unrelated to the signal, spreads the jammer to the bandwidth of the original spread signal, and only a small amount of the jammer power enters the receiver in the data bandwidth. Thus the processing gain of the direct-sequence system (PG $= T_b/T_c$) greatly reduces the effect of a narrow-band jammer on its performance. Furthermore, it follows that the spectral height of the spread signal is much lower than if it were not spread; hence, it is much harder to detect by an unauthorized user in the constant presence of the receiver noise. *See* JAMMING.

Attributes. The six attributes of spread spectrum systems that were discussed above apply in this direct-sequence, pseudonoise spread spectrum system. Suppose an unmodulated tone at the carrier frequency is present at the receiver along with the spread signal. Then in the pseudonoise code despreading process, the undesired tone is spread or multiplied by the local pseudonoise code, thereby spreading its spectrum to a bandwidth of $2/T_c$. However, the received signal is collapsed to a bandwidth of $2/T_b$ ($T_b \gg T_c$) so that the data detector "sees" only a small fraction (T_c/T_b) of the original tone energy. Low probability of intercept is achieved by virtue of the fact that the spread signal is very wide band and has low spectral height, and therefore

cannot be easily detected in a noise background. Message privacy is enhanced if the actual pseudonoise code sequence used by the transmitter and receiver is cryptographically derived and is unknown to unwanted listeners. Multipath tolerance is achieved by searching first with minimum delay and then expanding the search to larger delays. The desired signal will have less delay than the multipath signal and will be acquired first. Additional processing can be used to greatly minimize the multipath problems. If a set of essentially orthogonal codes, such as Gold codes (a special type of pseudonoise code), are used, many users can occupy the same bandwidth and only transmitter-receiver pairs with the same code can communicate. Since the pseudonoise chips are very short in time, ranging can be made accurate to a small fraction (5%) of a chip time. The higher the chip rate, the greater the range accuracy. *See* CRYPTOGRAPHY.

Time hopping. In time hopping, a time interval T is selected to be large compared to the data bit time, T_b, and is divided into, say, K slots. Then, during one slot time (T/K seconds), T/T_b bits are all burst-transmitted. The slot time that is transmitted is selected pseudorandomly so that only the transmitter and the receiver know the slot time code. In time hopping, the power increases by K times compared to what a continuous signal would require, but the time-hopping signal is on for only $1/K$ of the time, resulting in the same average power. The processing gain is K for BPSK modulation in this case, since bits are transmitted at K times the rate before time hopping. This scheme has high bandwidth efficiency, but high peak power and critical timing are required.

Frequency-hopped systems. In a direct-sequence system, the phase of the carrier changes pseudorandomly with the pseudonoise code. In a frequency-hopping system, the frequency of the carrier changes according to a pseudonoise code with a consecutive group of pseudonoise code chips defining a particular frequency. Typically either multiple frequency-shift keying (MFSK) or differential phase-shift keying (DPSK) is used as the data modulation. Multiple frequency-shift keying is a modulation scheme in which one of a number of tones (2, 4, 8, and so forth) is transmitted at a given time according to a group of consecutive data bits (n bits produce 2^{+n} tones). At each hop frequency one of the 2^{+n} tones is selected according to the n bits, and one of 2^{+n} corresponding frequencies, centered about the hop frequency, is transmitted. In conjunction with frequency hopping, multiple frequency-shift keying would imply, at each instant of time, a given carrier frequency that depends on the hop-pseudonoise code sequence and the consecutive group of the most recent n data bits. Differential phase-shift keying is similar to phase-shift keying except that only the differences of the phases (not the actual phases) are encoded and noncoherent techniques (not requiring a carrier loop) can be employed at the receiver.

Frequency-hopped systems are classified as fast or slow. Slow hopping occurs if one or more data symbols are transmitted in the time interval of one hop (T_b). Fast hopping occurs if two or more hops occur per data symbol. Although fast hopping is more complex to implement, it offers greater jammer protection. The processing gain for a frequency-hopped system is defined by $PG = W/R_b$, where W is the hopped bandwidth and R_b is the data rate.

Figure 3 illustrates a simple frequency-hopped system with binary frequency-shift keying (BFSK) data modulation. In this example a data bit 1 is denoted by a positive offset of Δf hertz, and a 0 by a negative offset of Δf hertz from the selected hop frequency. In Fig. 3 a 0 is sent in the first hop period since it occurs in the lower portion of the frequency bin, and in the second hop period a 1 is sent (at a different hop frequency). This process continues indefinitely. A well-designed hopping pattern will cover all parts of the hopping bandwidth with the same frequency of occurrence.

A device called a frequency synthesizer achieves the actual frequency selection. For example, a 12-bit segment of the pseudonoise code may correspond to one of 2^{12} different frequencies, so that one of approximately 4000 (2^{12}) frequencies is selected each hop time. Frequency synthesizers are used in both the transmitter and the receiver (**Fig. 4**). The transmitter modulates the data by typically using either multiple frequency-shift keying or differential phase-shift keying modulation, which in turn is frequency-hopped by the frequency synthesizer. At the receiver, an acquisition process is utilized to synchronize the receiver frequency synthesizer with the received hopping signal, and then a tracking system maintains synchronism. Finally a bit synchronizer provides timing for the data demodulator which demodulates the original, transmitted data bits.

Hybrid systems. Besides the four basic systems discussed above, hybrid schemes combining two or more of these systems have been used. The most

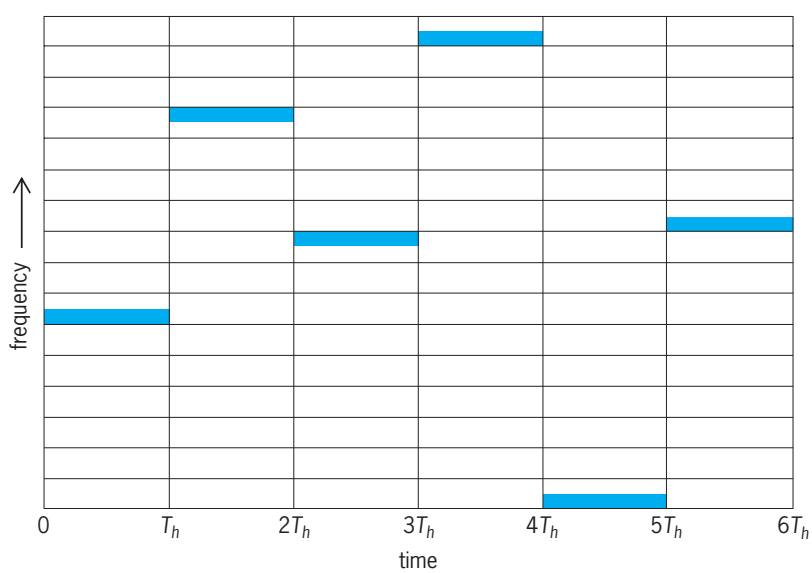

Fig. 3. Example of a frequency-hopped signal pattern.

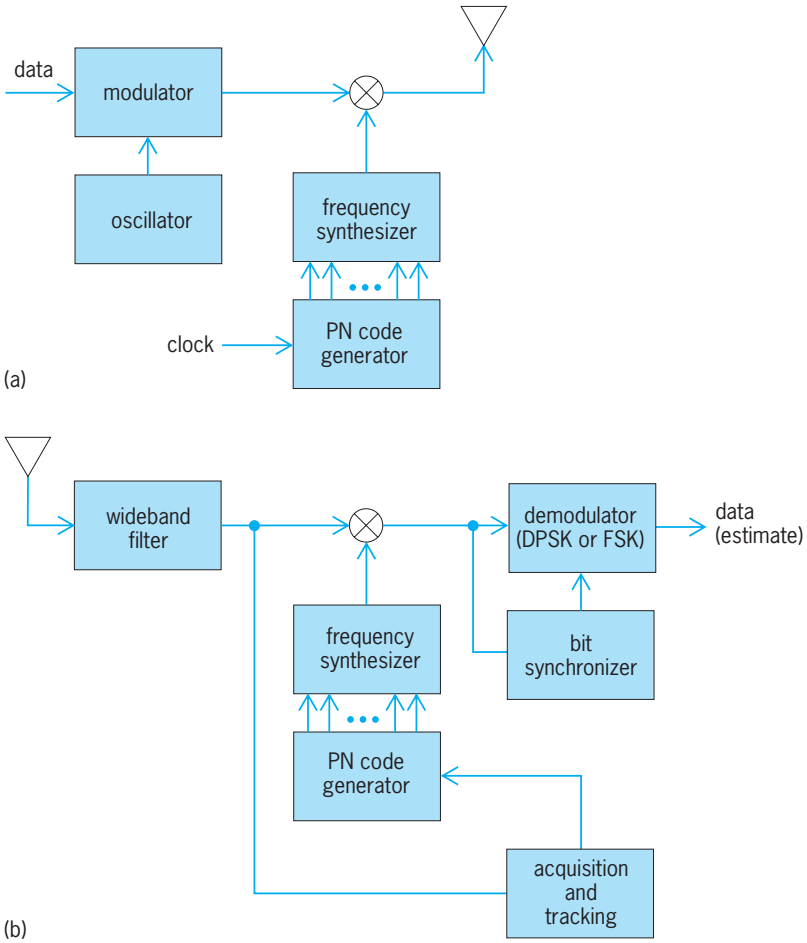

Fig. 4. Frequency-hopping equipment for (*a*) transmitter and (*b*) receiver.

common hybrid system is frequency hopping in conjunction with direct sequence; that is, each frequency is "spread" by a direct-sequence code.

Code division multiple access. An important aspect of spread spectrum communications is multiple access. Code-division multiple access (CDMA) is a method by which spread spectrum signals are utilized to allow the use of multiple signals over the same frequency band. The two common types are direct-sequence CDMA (DS/CDMA) systems and frequency-hopped CDMA (FH/CDMA) systems. *See* MULTIPLEXING AND MULTIPLE ACCESS.

DS/CDMA. This method is based on modulating the carrier phase via the multiple access code and the data modulation. Multiple codes can be used over the same channel because the different codes are nearly uncorrelated. That is to say, when any distinct pair of codes are multiplied together and the product is compared chip by chip over one data symbol, the code chips (+1's or −1's) have about the same number of "plus ones" as "minus ones." Thus two distinct codes are uncorrelated. However, when the same code is correlated against itself, its product is all ones, and thus is completely correlated. This property of different codes being uncorrelated is required for each code pair used in the multiple access system in order to minimize crosstalk or mutual interference effects. When multiple CDMA signals exist on the same channel, CDMA noise (background noise) occurs. *See* CROSSTALK.

FH/CDMA. This method uses spread spectrum codes to control the hop frequencies at various values across the spread bandwidth in such a way that, most of the time, they do not collide. Typically the data modulation is multiple frequency-shift keying, as discussed above. In many applications it is not possible to construct codes that will not collide (or will have zero correlation in the DS/CDMA case). These occasional collisions cause a degradation in the bit error rate performance. However, coding and interleaving/deinterleaving mitigate the effects of the collisions. The more channels that are used on the same band, the greater the interference level, which is sometimes viewed as background noise. *See* ELECTRICAL NOISE.

Personal communication systems. In 1983 an analog cellular radio system, called the Advanced Mobile Phone System (AMPS), was introduced in the United States. Other analog systems were later introduced in Europe and Asia. In 1993 a CDMA spread spectrum personal cellular service was introduced in North and South America, Asia, and elsewhere. This standard, called TIA IS-95, is a second-generation (2G) standard. Each mobile user is assigned one member of a set of 64 orthogonal functions know as Walsh functions. The Walsh functions are generated as rows of special square matrices called Hadamard matrices. On the forward channel (from the base station to the mobile user) the Walsh functions are used for spreading, and on the reverse channel (user to base station) they are used as orthogonal functions for the modulation. Third-generation (3G) CDMA systems are in the beginning stages, and offer more features and greater data rates than second-generation systems. *See* MOBILE COMMUNICATIONS.

Other applications. Although the early evolution of spread spectrum systems was motivated primarily by military interests, nonmilitary applications have enjoyed considerable development. One important example that has both military and nonmilitary users is the Global Positioning Systems (GPS), which is a direct-sequence, CDMA, spread spectrum system for transmitting the satellite ranging codes. *See* SATELLITE NAVIGATION SYSTEMS.

The space shuttle utilizes a direct-sequence spread spectrum communication system on its forward link. It relays data through the geostationary *Tracking and Data Relay Satellites* (*TDRS*) by using either S-band or Ku-band spread spectrum links. Another example in the military arena is the Milstar system, which utilizes frequency-hopping spread spectrum communication over a very large bandwidth to achieve considerable immunity from unfriendly jamming signals. *See* MILITARY SATELLITES; SPACE COMMUNICATIONS; SPACE SHUTTLE.

Globalstar is a commercial satellite system that utilizes a CDMA signal structure along with a bent-pipe transponder to provide communications much like a cell phone, except that Globalstar is satellite-based. Its signal structure is similar to IS-95. Currently the

handsets are considerably more expensive than cell phones. However, these phones have the advantage that they can be reached beyond the range of cell phones. *See* COMMUNICATIONS SATELLITE.

The main reception source of news stories for radio and newspapers is small spread spectrum ground stations. Another application of code-division multiple access is transmission directly by satellite from a bank's automatic teller machine to that bank's computer facility. A home security system using spread spectrum techniques imposed on the ac power line has been used in Japan. Many cordless telephones utilize direct-sequence spread spectrum techniques. Jack K. Holmes

Bibliography. S. Glisic and B. Vucetic, *Spread Spectrum CDMA Systems for Wireless Communications*, Artech House Publishers, 1997; J. K. Holmes, *Coherent Spread Spectrum Systems*, Krieger Press, 1990; R. L. Peterson et al., *Introduction to Spread Spectrum Communications*, Prentice Hall, 1995; M. K. Simon et al., *Spread Spectrum Communications Handbook*, rev. ed., 1994.

Spring (hydrology)

A place where groundwater discharges upon the land surface because the natural flow of ground water to the place exceeds the flow from it. Springs are ephemeral, discharging intermittently, or permanent, discharging constantly.

Springs are usually at mean annual air temperatures. The less the discharge, the more the temperature reflects seasonal temperatures. Spring water usually originates as rain or snow. Meteoric water compositions for the deuterium (D) and oxygen-18 (^{18}O) are given by the empirical equation

$$\delta D = 8\delta^{18}O + 10$$

where δ is the isotope composition in parts per thousand referred to standard mean ocean water.

Standard mean ocean water is defined as zero in both O^{18} and D. Meteoric water from evaporation of seawater is depleted in D and ^{18}O for kinetic reasons and hence the values of δD and $\delta^{18}O$ are negative and increasingly so the farther from the Equator.

Hot springs. Hot-spring water may differ in composition from meteoric water through exchange between the water and rocks. Common minerals consist of component oxides. Oxygen of minerals has more ^{18}O than meteoric water. Upon exchange, the water is enriched in ^{18}O. Most minerals contain little deuterium, so that slight deuterium changes occur. Water in hot rock is buoyant relative to surrounding water and isues as hot springs. Temperatures of the deep water may be estimated. Estimates stem from mineral solubilities with temperature. If saturation with a mineral is assumed and the chemical composition of the water known, a temperature may be calculated. The solubility of quartz, SiO_2, is commonly used for estimating the deep temperatures of hot springs. Common minerals also contain silica, so that temperature estimates are ambiguous. Some hot-spring waters are acid from the oxidation of hydrogen sulfide to sulfate.

Mineral springs. Mineral spring waters have high concentrations of solutes. Mineral springs have long been used for therapy; the study of such use of springs is balneology. Pliny the Younger reported mineral springs in Belgium and Anatolia, which are still in use. Although mineral waters are still valued for therapy in much of the world, in North America such use has declined greatly since the early twentieth century.

Mineral springs have wide ranges in chemistry and temperatures, and hot mineral springs may be classified as hot springs as well as mineral springs. Most mineral springs are high either in sodium chloride or sodium bicarbonate (soda springs) or both; other compositions are found, such as a high percentage of calcium sulfate from the solution of gypsum. The water and the solutes may be of different origins. The stable isotope compositions of the water are the best guides to the sources of the water. For the mineral springs of Europe, the United States, and Asia Minor, deuterium and ^{18}O compositions show that the waters are meteoric.

Carbon dioxide. Carbon dioxide is commonly found dissolved in or issuing through mineral water, and is identifiable by its ^{13}C composition. Carbon of carbonate minerals deposited in oceans is near zero in ^{13}C compared to an international standards, the Peedee belemnite. The carbon of the carbon dioxide in most of the mineral springs of Europe, the United States, and Asia Minor is close to zero in ^{13}C isotope composition. It is inferred that the carbon dioxide results from the breakdown (metamorphism) of the carbonate minerals that are present in rocks of marine sedimentary origin.

Carbon dioxide from mineral springs of volcanic islands of the Atlantic Ocean seems to come from the source of the volcanic rocks, the Earth's mantle, because of the similarity of the isotope compositions of the carbon dioxide in the springs and in carbon dioxide bubbles in volcanic rocks erupted on the deep ocean floor in the -4 to -8 per thousand range. Carbon dioxide from great depth also discharges through springs in the Sierra Nevada, California, and the Alaska peninsula. In mineral springs from the Rocky Mountains, carbon dioxide comes from both very deep sources and the shallower chemical reactions of marine rocks.

A third source of carbon dioxide is organic material. Springs in the Coast Ranges of California, along the Copper River of Alaska, and in New Zealand yield mixtures of carbon dioxide and methane. The carbon dioxide is depleted in carbon-13 and probably comes from the breakdown of organic (woody) materials in the sedimentary rocks.

Salts. Other solutes in the mineral springs are inferred to come from the rocks, but isotope data do not show origins. Salt deposits are sources for solutes in some instances. There are many mineral springs for which there is no possible salt deposit source for the solutes. Chemical compositions may be used to infer the sources of the solutes. Chloride-rich waters

in marine sedimentary rocks also contain bromide, iodide, and boron. During metamorphism of the rock the waters may be retained as fluid inclusions or as films along grain boundaries. Flushing of the retained brines may result in the saline springs occasionally found in metamorphic terraines.

Chemical equilibrium. The chemical compositions of spring waters are seldom in chemical equilibrium with the air. Ground waters whose recharge is through grasslands may contain a thousand times as much CO_2 as would be in equilibrium with air, and those whose recharge is through forests may contain a hundred times as much as would be in equilibrium with air. The high CO_2 content, along with other solutes, makes some spring waters quite nutrient to aquatic plants. If the CO_2-rich ground water has dissolved calcite, loss of CO_2 in a spring (chiefly by photosynthesis) may cause the calcite to precipitate to make travertime. Ground water flowing through rocks or sediments containing organic material or reduced minerals such as sulfides loses its dissolved oxygen by oxidizing the organic matter and the sulfides. When the dissolved oxygen is lost, iron enters the groundwater as ferrous ion, Fe^{2+}. The Fe^{2+} is not chemically compatible with air, and springs containing Fe^{2+} lose it as $Fe(OH)_3$ upon oxidation. Thin films of $Fe(OH)_3$ on the water surface are iridescent and resemble oil films. Unlike oil films, the $Fe(OH)_3$ films are brittle and if ruptured do not reform. Accumulations of the iron precipitates may lead to bog iron ores. Sulfate in ground water may be reduced in the presence of organic matter to H_2S, giving some springs the odor of rotten eggs. *See* GEYSER; GROUNDWATER HYDROLOGY. Ivan Barnes

Bibliography. D. K. Todd and L. W. Mays, *Groundwater Hydrology*, 3d ed., 2004; P. Younger, *Groundwater in the Environment: An Introduction*, 2006.

Spring (machines)

A machine element for storing energy as a function of displacement. The flywheel, in contrast, is a means for storing energy as a function of angular velocity. Force applied to a spring member causes it to deflect through a certain displacement thus absorbing energy. *See* FLYWHEEL.

A spring may have any shape and may be made from any elastic material. Even fluids can behave as compression springs and do so in fluid pressure systems. Most springs take on specific and familiar shapes such as helix, flat, or leaf springs. All mechanical elements behave to some extent as springs because of the elastic properties of engineering materials.

Uses. Energy may be stored in a spring for many uses: to be released later, to be absorbed at the instant the energy first appears, and so on.

Motive power. One of the early and still the most frequent uses of springs is to supply motive power in a mechanism. Common examples are clock and watch springs, toy motors, and valve springs in auto engines. In these, energy is supplied to, and stored in, the spring by applying a force through a suitable mechanism to deflect or deform the spring. The energy is released from the spring by allowing it to push (as in the valve) or twist (as in a clock) a mechanism through a required displacement.

Return motion. A special case of the spring as a source of motive power is its use for returning displaced mechanisms to their original positions, as in the door-closing device, the spring on the cam follower for an open cam, and the spring as a counterbalance. To a certain extent the springs in vehicles of transportation are in this category. They are designed to keep the car at a certain level with respect to the road or rails, returning the vehicle to this position if displaced by applied forces.

Shock absorbers. Frequently a spring in the form of a block of very elastic material such as rubber absorbs shock in a mechanism. For example, the four legs of a punch press rest on four blocks of rubber. The rubber pads prevent the die-closing inertia forces of the press from transferring down through the legs to the floor with impact or hammer blow proportions. With the rubber pads under the press legs, the force on the floor builds up relatively slowly and no shock is evident. As the acceleration of the die block goes to zero, the inertia force goes to zero and the rubber pad springs, which were deflected by the press blow, are relaxed and ready for the next stroke of the press die. The whole press moves up and down relative to the floor, but by proper selection of the rubber pad the elastic constant is such that this motion is small. *See* SHOCK ABSORBER.

Vibration control. Springs serve an important function in vibration control by supplying the necessary flexibility in the support of the vibrating mechanism and the required opposing forces as a result of their deflection. In controlling vibration, the body or mass of the mechanism must be freely supported so as to generate forces opposing the vibrating forces. These opposing forces tend to bring the sum of the forces on the vibrating body to zero. Vibration-absorbing mounts are available in a wide range of sizes and spring constants to meet most requirements. These do not prevent vibration, which is a function of speed and the balance of the mechanism, but rather they minimize its effect on the machine frame or the mounting. *See* SHOCK ISOLATION.

Force measurement. Springs have long been used in simple weighing devices. Accurate weighing is usually associated with dead-weight devices or balances, but modern spring scales have received wide acceptance and certification for commercial use. Extremely accurate springs for heavy loads are used to calibrate testing machines, and they are employed in scales over crane hooks for weighing material as it is hoisted. Carefully calibrated springs are used in instruments such as electric meters and pressure gages. *See* BALANCE.

Retaining rings. A relatively modern machine part in which the spring function is used as a holding means is the retaining or snap ring. This device is a split ring of square, rectangular, or special cross section. It fits in a groove on a cylindrical surface or in a bore

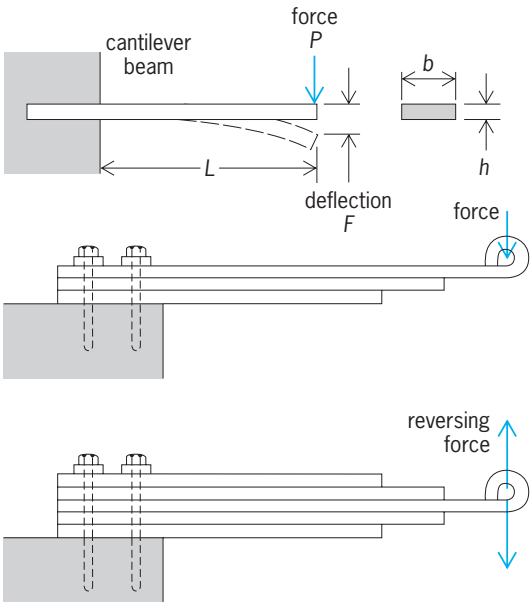

Fig. 1. Flat or leaf spring responds to perpendicular forces in either direction.

and stays in place by spring force. A retaining ring prevents or restrains relative axial motion between a shaft or bore and the components on the shaft or in the bore. *See* SNAP RING.

Types. Springs may be classified into six major types according to their shape. These are flat or leaf, helical, spiral, torsion bar, disk, and constant force springs.

Flat or leaf spring. A leaf spring is a beam of cantilever design with a deliberately large deflection under a load (**Fig. 1**). One end of a leaf spring is usually firmly anchored to the frame of the machine and the other end is linked to the moving machine elements by a two-force (pin-ended) link. Force may be tension or compression with no modification of the design. This push-pull feature is the great advantage of a leaf spring, plus the fact that a relatively large amount of energy can be stored in a small space.

Helical spring. The helical spring consists essentially of a bar or wire of uniform cross section wound into a helix. The last turn or two at each end of the spring is modified to a plane surface perpendicular to the helix axis, and force can then be applied to put the helix in compression. The ends of the spring helix may be modified into hooks or eyes so that a force may be applied in tension. In general it is necessary to design helical springs so that force may be transmitted either in tension or in compression but not both ways for the same spring (**Fig. 2**). Where reversing forces occur in a spring mechanism, it is better to use a bar or leaf spring and in some cases a disk spring.

Spiral spring. In a spiral spring, the spring bar or wire is wound in an Archimedes spiral in a plane. Each end of the spiral is fastened to the force-applying link in the mechanism.

A spiral spring is unique in that it may be deflected in one of two ways or a combination of both of them (**Fig. 3**). If the ends of the spiral are deflected by forces perpendicular to the spiral plane, the spiral is distorted into a conical helix. For better stability and ease of applying the forces, the spring is often made as a conical helix to start with and is deflected into a plane spiral.

A spiral spring may also have the forces act tangent

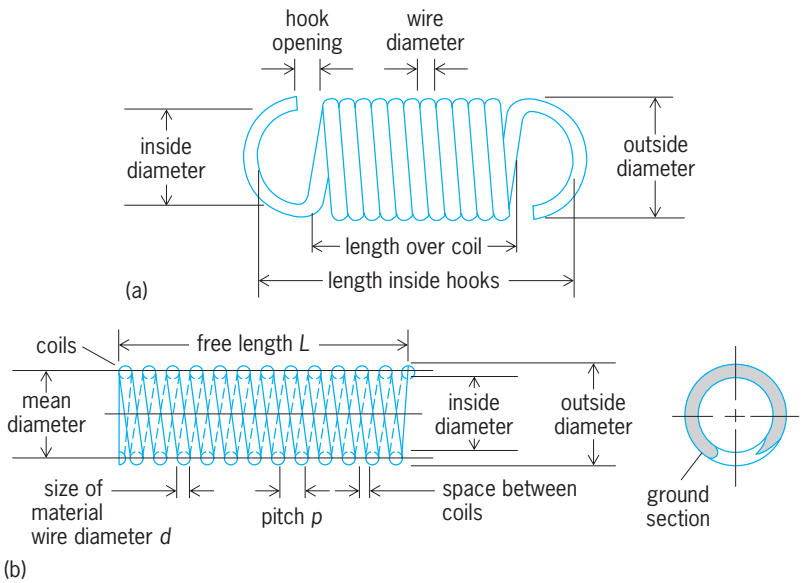

Fig. 2. Helical spring. (*a*) Wound tight to extend under axial tension. (*b*) Wound loose to contract under axial compression.

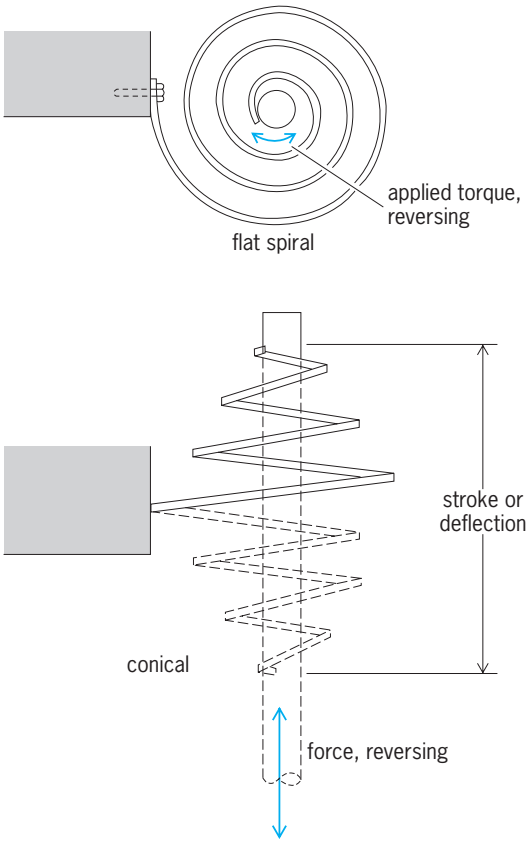

Fig. 3. Spiral spring is unique in responding to torsional or translational forces.

to the spiral as in a clock spring. The spiral is wound quite open and the tangential force in tension on the spiral tends to close the gaps between successive turns. The spring behaves like a beam, bending to a shorter radius of curvature and thus storing energy.

Torsion bar. A torsion bar spring consists essentially of a shaft or bar of uniform section. It stores energy when one end is rotated relative to the other. It is used in the spring system of the chassis of modern motor cars. *See* TORSION BAR.

Disk spring. Where large forces are present and space is at a premium, the disk spring may be used, although it is usually expensive to design and build. It consists essentially of a disk or washer supported at the outer periphery by one force (distributed by a suitable chuck or holder) and an opposing force on the center or hub of the disk (**Fig. 4**). For greater deflections several disks may be stacked with the forces transmitted from one to the next by inner and outer collars.

Constant-force spring. Many mechanisms require that a constant force be applied regardless of displacement. The counterbalancing of vertically moving masses against the force of gravity is a typical example. The Neg'ator spring of Hunter Spring Co. provides such a constant force; it uses a tight coil of flat steel spring stock. When the outer free end is extended and the coil allowed to rotate on its shaft or pintle, the spring presents a constant restoring force.

Design. Each type spring has its special features and design refinements. Common to all forms of springs are the basic properties of elastic materials. Within the elastic range of a material, the ratio of applied force to resulting deflection is constant. Spring systems can be designed to have a variable ratio. The ratio is the spring rate or scale and has the dimension of force per unit length. *See* HOOKE'S LAW.

In a helical spring the elastic action stresses the wire in torsion. The following variables are subject

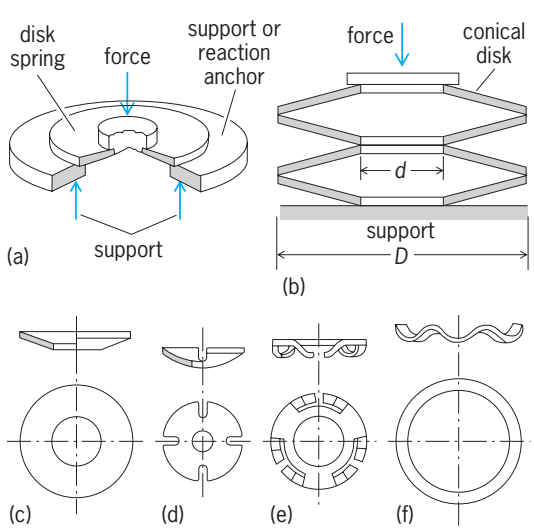

Fig. 4. Disk spring. (*a*) Single tapered disk; (*b*) multiple disk; (*c*) Belleville spring; (*d*) slotted curved washer; (*e*) finger washer; (*f*) wave spring washer.

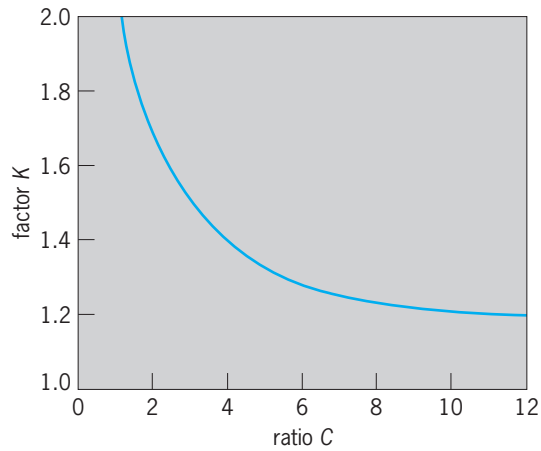

Fig. 5. Correction factor for curvature and shear. (*After A. M. Wahl, Mechanical Springs, Penton, 1944*)

to the designer's action and decision:

free length	L	spring rate	k
mean diameter	D	modules of elasticity in torsion	G
wire diameter	d	number of active turns	N
allowable stress	S	deflection at load P	F
applied load	p		

These variables are related by Eq. (1), which

$$S = \frac{8PD}{\pi d^3} \tag{1}$$

neglects the effect of coil curvature on the stress, and by Eq. (2). For springs subject to frequent cycling, as

$$F = \frac{8PD^3N}{Gd^4} \tag{2}$$

a valve spring, the correction factor for wire curvature must be included. The generally used correction is the factor proposed by A. M. Wahl. Introducing this factor changes Eq. (1) to the form of Eq. (3), in

$$S_{\max} = \frac{8PD}{\pi d^3}\left(\frac{4C-1}{4C-4} + \frac{0.615}{C}\right) \tag{3}$$

which $C = D/d$. The portion in brackets is Wahl's factor K whence maximum stress is K times the torsional stress (**Fig. 5**).

Various combinations of spring dimensions for the variables listed above will produce an acceptable spring. Usually the design is chiefly limited by allowable stress of materials and space.

Where compression and tension springs are cycled at very high frequencies, surging may cause high local stress and result in early fatigue failure. Surging is the inability of all parts of the spring to deflect at the same rate due to the inherent inertia in the coils. This phenomenon is closely associated with the natural frequency of the spring; springs should be designed so that their natural frequency and cyclic rate are as far apart as practical.

L. Sigfred Linderoth, Jr.

Bibliography. Society of Automotive Engineers, *Spring Design Manual*, 2d ed., 1995; Spring Manufacturers Institute, *Handbook of Spring Design*,

rev. ed., 2002; A. M. Wahl, *Mechanical Springs*, 2d ed., reprinted 1991.

Spruce

Evergreen tree belonging to the genus *Picea* of the pine family. The needles are single, usually four-sided, and borne on little peglike projections; the cones are pendulous. Resin ducts in the wood may be seen with a magnifying lens, but they are fewer than in *Pinus*. *See* EVERGREEN PLANTS; PINALES; PINE.

Eastern species. The white spruce (*P. glauca*), ranging from northern New England to the Lake States and Montana and northward into Alaska, is distinguished by the somewhat bluish cast of its needles, small cylindrical cones (**Fig. 1***c*), and gray or pale-brown twigs without pubescence (hairs). Of erect pyramidal habit, it usually attains a height of 60–70 ft (18–21 m) with trunk diameters of 1–2 ft (0.3–0.6 m), but occasionally in British Columbia and Alberta the tree reaches heights of 80–140 ft (24–43 m) with diameters up to 4 ft (1.2 m). When bruised, the leaves of some individuals emit a disagreeable odor, hence the names cat or skunk spruce. The pale, straw-colored wood is soft and straight-grained without different coloration of the heartwood. The wood is used for paper pulp, construction, and interior finish. The evenly grained wood is used in the manufacture of musical instruments.

Red spruce (*P. rubens*) is a similar tree but with greener foliage; smaller, more oval cones (Fig. 1*b*); and more or less pubescent twigs. Occurring naturally with white spruce in the northeastern United States and adjacent Canada, red spruce extends southward along the Appalachians into North Carolina. As in white spruce, the greatest use of the wood is for paper pulp.

Black spruce (*P. mariana*) ranges from northern New England and Newfoundland to Alaska. However, it occurs sparingly in the Appalachians to West Virginia. The cones are smaller than in the white and red species and are egg-shaped or nearly spherical (1–1$^1/_2$ in. or 2.5–3.8 cm long) and persistent (Fig. 1*a*). The twigs are pubescent. Known also as bog or swamp spruce in the southern part of its range, black spruce appears at high elevations in the north. The long-fibered wood is ideal for paper pulp and is also used for paddles, oars, construction, and shipbuilding. It is short-lived as an ornamental tree.

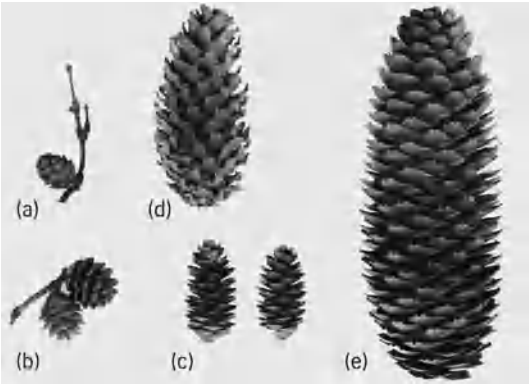

Fig. 1. Cones of eastern and western spruce. (*a*) Black spruce (*Picea mariana*). (*b*) Red spruce (*P. rubens*). (*c*) White spruce (*P. glauca*). (*d*) Colorado blue spruce (*P. pungens*). (*e*) Norway spruce (*P. abies*). (*Brooklyn Botanic Garden*)

Fig. 2. Blue spruce (*Picea pungens*) in Bryce Canyon National Park, Utah. (*National Park Service*)

Western species. Blue spruce (*P. pungens*) (**Fig. 2**), also known as Colorado blue spruce, is probably the best known of the western species (Fig. 1*d*) because of its wide use as an ornamental tree. The twigs are glabrous (without pubescence). The tree is popular as an ornamental in northern Europe, as well as throughout the United States. Individuals vary in the depth of blue coloration, the most pronounced being that of the Koster variety. Most of the ornamental specimens come from grafts made on Norway spruce.

Engelmann spruce (*P. engelmanni*) has needles usually of a deep blue-green color, sometimes much like those of the blue spruce, but the young twigs are slightly hairy. The cones, although cylindrical, are smaller than in blue spruce, being about 1$^1/_2$–3 in. (3.8–7.6 cm) long. This species is also a Rocky Mountain tree like the blue spruce, but it is more widely distributed from British Columbia to Arizona

and also in the mountains of Oregon and Washington. It is an important timber species of the Rocky Mountain region.

Sitka spruce (*P. sitchensis*) is the largest spruce in the Northern Hemisphere. The leaves have a pungent odor, are considerably flattened, and stand out from the twig in all directions. Ranging from Alaska to northern California, it occupies a coastal strip about 40–50 mi (64–80 km) wide. Mature trees are 180–200 ft (55–61 m) tall and $3^1/_2$–6 ft (1.1–1.8 m) in diameter, but in some instances larger dimensions have been recorded. Timber volume in the United States is nearly all in the states of Washington and Oregon. The wood is used for furniture, doors, blinds, and paper pulp, and in piano manufacture of sounding boards.

The Norway spruce (*P. abies*), the common spruce of Europe, is much planted in the United States for timber, as well as for ornamental purposes. It can be recognized by the dark-green color of the leaves; glabrous, pendent, short branchlets; and cones 4–6 in. (10–15 cm) in length (Fig. 1*e*), usually near the top of the tree. An important source of paper pulp, it has often been planted for that purpose in the northern and eastern United States. It is also used for shelter belts and Christmas trees, although it is not as desirable for the latter purpose as the firs and Douglas-fir, since it sheds its needles quickly in a warm room. There are many cultivated varieties. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Bibliography. A. Farjon, *Pinaceae: Drawings and Descriptions of the Genera Abies, Cedrus, Pseudolarix, Keteleeria, Nothotsuga, Tsuga, Cathaya, Pseudotsuga, Larix, and Picea*, Koeltz Scientific Books, Königstein, 1990; R. J. Taylor, Sections on Picea and Tsuga, in *Flora of North America: North of Mexico*, vol. 2, ed. by Flora of North America Editorial Committee, Oxford University Press, 1993.

Sputtering

The ejection of material from a solid or liquid surface following the impact of energetic ions, atoms, or molecules. Sputtering was first noted in 1853 by W. R. Grove and in 1854 by M. Faraday as the occurrence of a metallic coating on the glass walls of a discharge tube. E. Goldstein in 1902 showed that the sputtering effect was caused by positive ions of the discharge striking the cathode and ejecting cathode material. Sputtering now is the basis of a large variety of methods for the synthesis and analysis of materials.

Classification. Sputtering can be classified according to the mode of energy loss of the incident (primary) particle. Nuclear stopping involves billiard ball–like atomic collisions in which a significant momentum transfer occurs; it dominates for incident ion energies below about 1–2 keV per nucleon. Electronic stopping involves collisions in which little momentum is transferred, but significant electronic excitation is caused in the target; it dominates for energies above about 10 keV per nucleon.

Sputtering has also been classified into physical and chemical sputtering. Physical sputtering involves a transfer of kinetic energy from the incident particle to the surface atoms leading to ejection, while chemical sputtering occurs when the incident species react chemically with the target surface leading to the formation of a volatile reaction product which evaporates thermally from the surface.

The main models for physical sputtering are the thermal sputtering and the collision cascade models. The thermal sputtering model suggests that the ion impact instantaneously heats a small region of the surface to a temperature well above the boiling point of the material, so that a small amount of material can evaporate before the heat is dissipated to the surrounding solid. The collision cascade model suggests that the impinging ion collides with surface atoms, transferring kinetic energy to these atoms which then collide with their neighbors, and so on. Atoms at or near the surface which are set in motion out of the surface with a kinetic energy great enough to overcome the chemical binding to their neighbors are sputtered. There is little evidence that thermal sputtering is important for most materials; in contrast, the collision cascade model accounts for the vast majority of observed sputtering phenomena. In particular, sputtering of single crystals leads to ejection in certain preferred directions correlated with the crystal lattice, which clearly would not occur in a random disordered hot zone. *See* CRYSTAL STRUCTURE.

Yields. In the nuclear stopping regime, sputtering yields, Y, can be expressed as in the equation

$$Y \propto \frac{S_n}{U} \qquad \text{atom/ion}$$

where S_n is the nuclear stopping coefficient (energy deposited in the solid per unit distance traversed) and U is the surface binding energy of the target material (assumed equal to the heat of sublimation per atom). Sputtering threshold energies—the incident ion energy at which sputtering is just detectable—are on the order of 10 times the sublimation energy, or some tens of electronvolts, for most elements. Sputtering yields are a maximum for incident ion energies around 200–300 eV per nucleon, where the nuclear stopping coefficient is a maximum. For argon ions (atomic mass 40) this corresponds to energies around 10 keV, at which energy sputtering yields for most materials under argon ion bombardment are in the range 1–10 atoms sputtered per incident ion. For incident ion energies below about 10 keV, a major fraction (80–90%) of the sputtered atoms come from the outermost surface atom layer, most of the remainder from the second layer, while essentially all the sputtered atoms come from no deeper than about four atomic layers; this is the basis of the use of sputtering in surface analysis.

Sputtering yields in the electronic stopping regime cannot be expressed simply, as mechanisms are still obscure. Hard, "billiard-ball" collisions are rare at high energies, because the incident ion travels through the solid as a bare nucleus, or certainly with many of its electrons stripped away, and is thus

extremely small. However, highly energetic particles can deposit up to several thousand electronvolts per nanometer into electronic excitation and ionization in the substrate. Mechanisms suggested for coupling of this electronic excitation to the nuclear motion necessary for sputtering include Coulomb repulsion of ionized atoms from one another along the particle track, or excitation to antibonding (repulsive) electronic states.

Complex materials. Sputtering of complex materials—metal alloys, inorganic and organic compounds and polymers, and minerals—can produce complex results. The relative efficiencies with which different elemental species are ejected following ion impact can differ, giving rise to preferential sputtering. When preferential sputtering occurs, the species sputtered with the lower efficiency accumulates to a higher concentration at the surface. Subsurface collisions of the incident ion, which penetrates several nanometers into the solid at kiloelectronvolt energies, cause atomic motion leading to ion beam–induced atomic mixing of surface and subsurface layers over the ion penetration depth. Chemical bonds can be broken, and sometimes new bonds can be formed. Sputtering of solids which have multiple phases, or which are polycrystalline, leads to the development of surface roughness due to the differences in sputtering yields between different regions. *See* ION BEAM MIXING.

Semiconductor devices. Sputtering is widely used in the manufacture of semiconductor devices; sputter deposition is used to deposit thin films with a high degree of control by sputtering material from a target onto a substrate; sputter etching is used to remove unwanted films in a reversal of this process. Most removal and deposition systems generate the energetic sputtering ions in radio-frequency (rf) or direct current (dc) glow discharges in argon gas. Reactive ion etching is a chemical sputtering process in which chemically active sputtering species form volatile compounds with the target material leading to significantly higher etch rates and great selectivity. For example, fluorine-containing compounds etch silicon rapidly by forming volatile silicon tetrafluoride (SiF_4) but do not etch aluminum or other metals used to make electrical interconnections between devices on a semiconductor chip because the metal fluorides are involatile. Sputter etching and reactive ion etching have the useful advantage of being anisotropic—that is, they etch only in one direction (usually normal to the surface) so that very fine surface features can be delineated. *See* CRYSTAL GROWTH; GLOW DISCHARGE; INTEGRATED CIRCUITS.

Materials characterization. In materials characterization, sputtering is used to remove surface material controllably, allowing in-depth concentration profiles of chemical composition to be determined with a surface-sensitive sampling technique such as Auger electron spectroscopy, x-ray photoelectron spectroscopy, secondary ion mass spectrometry, or ion scattering spectrometry. Sputtering ejects not only atomic but also molecular species from surfaces, and a fraction of these ejected species are ionized as a result of the energetic collisions leading to ejection. Analysis of these secondary ions forms the basis of secondary ion mass spectrometry. Inorganic materials can be analyzed by identification of atomic secondary ions; organic materials can give characteristic molecular ions. Degradation of organic materials by the ion beam irradiation is reduced by dispersal in a liquid matrix or by operation at very low irradiation rates. Projectile energies both in the nuclear stopping and in electronic stopping regimes result in ejection of very large molecular ions up to 5000 atomic mass units for nuclear stopping and over 15,000 atomic mass units for electronic stopping, allowing analysis of very large biomolecules by mass spectrometry. *See* AUGER EFFECT; SURFACE PHYSICS.

Peter Williams

Bibliography. R. Behrisch and W. Eckstein (eds.), *Sputtering by Particle Bombardment*, 2007; B. Chapman, *Glow Discharge Processes: Sputtering and Plasma Etching,* 1980; M. Ohring, *The Materials Science of Thin Films*, 2d ed., 2001; K. Wasa, M. Kitabatake, and H. Adachi, *Thin Films Materials Technology: Sputtering of Compound Materials*, 2004.

Squalene

A C_{30} triterpenoid hydrocarbon. Squalene is made up of six (*trans*-1,4)-isoprene units linked as two farnesy1 (head-to-tail) groups that are joined tail to tail in the center (**Fig. 1**).

Squalene can be isolated in large quantities from the liver oils of the shark (*Squalus* spp.) and other elasmobranch fishes, and is a relatively inexpensive compound. It undergoes oxidation reactions typical of a polyene. Complete hydrogenation of the liver oil gives the saturated hydrocarbon squalane, which is used in lotions and skin lubricants.

The major significance of squalene is its role as a central intermediate in the biogenesis of all steroids and triterpenoids. Thus, although squalene accumulates in a narrow phylogenetic group of primitive fish, it is a ubiquitous metabolite in all organisms that contain steroids.

The biosynthetic pathway begins with acetic acid units and leads via mevalonate and isopentenyl pyrophosphate to the sesquiterpene farnesol. Two farnesyl pyrophosphate units are then coupled tail to tail. This coupling involves the formation of presqualene followed by rearrangement and reduction of the cyclopropane ring (**Fig. 2***a*).

The transformation of squalene to polycyclic triterpenoids and steroids is initiated by epoxidation of one double bond followed by cyclization. The product depends on the conformation, or folding, of the

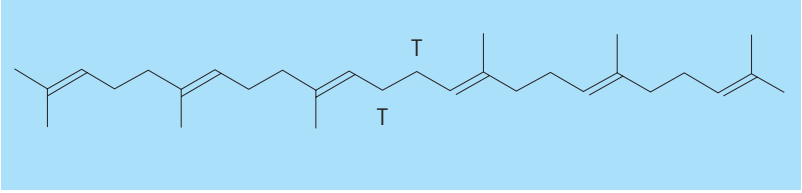

Fig. 1. Structure of squalene; the tail-to-tail joining is indicated by T.

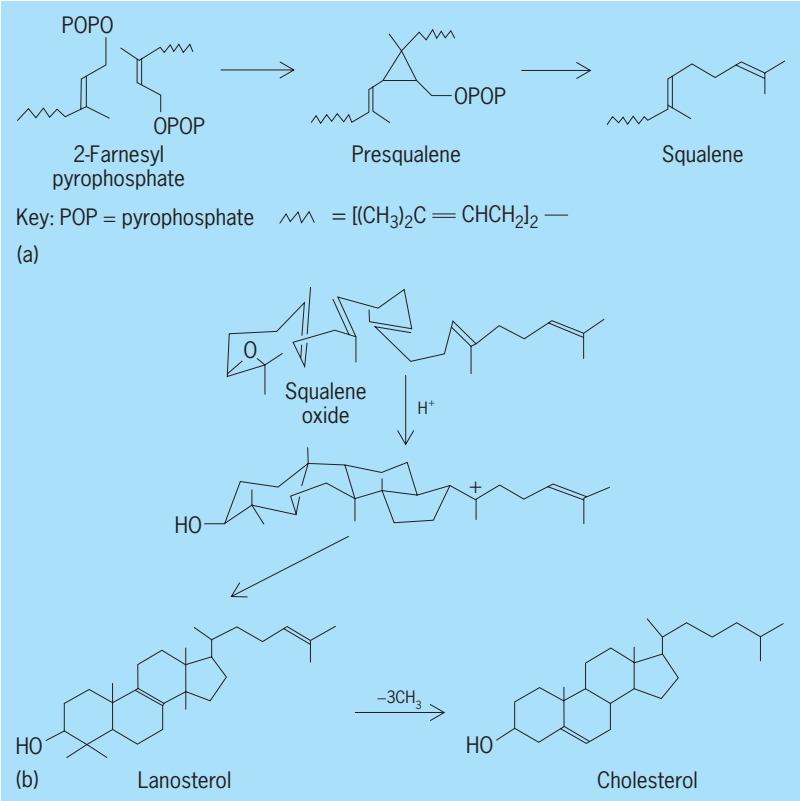

Fig. 2. Biosynthetic pathways. (*a*) Formation of squalene. (*b*) Transformation of squalene to cholesterol.

chain. The pathway from squalene oxide to cholesterol is indicated in Fig. 2*b*. A concerted sequence of bond formation and methyl group rearrangements leads to the triterpenoid lanosterol, which is a major constituent of sheep wool fat. The final steps to cholesterol include oxidative loss of three methyl groups. Further enzymatic oxidations in specific tissues give rise to the several classes of steroid hormones. *See* STEROID; TERPENE; TRITERPENE.

James A. Moore

Bibliography. S. V. Bhat, B. A. Nagasampagi, and S. Minakshi, *Chemistry of Natural Products*, 2005; T. W. Graham Solomons and C. B. Fryhle, *Organic Chemistry*, 8th ed., 2003; R. K. Murray et al., *Harper's Biochemistry*, 25th ed., 1999.

Squaliformes

An order of sharks comprising six families, 24 genera, and 97 species, commonly known as dogfish sharks. The order differs from all other sharks by the following combination of characters: two dorsal fins, with or without spines; no anal fin; spiracle behind eye; and lower nictitating eye membrane absent. For the most part, squaliform sharks inhabit continental and insular shelves and slopes, and sea mounts. Some species live at great depths. So far as known, squaliforms are ovoviviparous and produce from one to 22 pups per litter, depending on the species. The dwarf lantern shark (*Etmopterus perryi*), at 17 cm (6.7 in.) in length, probably is the world's smallest shark; the largest squaliform is the Pacific sleeper shark (*Somniosus pacificus*), which is up to 7 m (23 ft) in length. The diet consists of bony fishes, cephalopods, crustaceans and other invertebrates, mammals, and other sharks. *See* ELASMOBRANCHII; SELACHII.

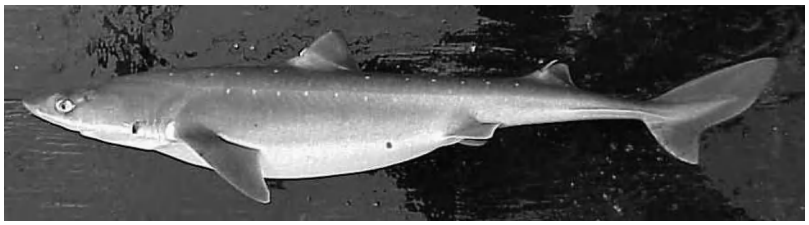

Spiny dogfish (*Squalus acanthias*). (*Photo courtesy of Alaska Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, U.S. Department of Commerce*)

Squalidae (dogfish sharks). Squalidae has a spine (without grooves) preceeding both dorsal fins, teeth in the lower jaw only slightly larger than those in the upper jaw, usually a precaudal pit, and a lateral keel on each side of the caudal peduncle. The family comprises two genera. *Cirrhigaleus* (two species) is very distinctive in having a robust body and barbels in the form of large nasal flaps that reach the mouth. By comparison, *Squalus*, comprising eight species, has a more slender body and much smaller nasal flaps. The species *S. acanthias* (spiny dogfish) (see **illustration**) is probably is the world's most abundant shark and also the largest of the family, attaining a length of 160 cm (63 in.), although most are much smaller. Squalids are antitropical (found in both hemispheres but not in equatorial regions) in the eastern and western Atlantic and Pacific oceans, but apparently are absent in the Indian Ocean.

Centrophoridae (gulper sharks). The family Centrophoridae is distinguished from other squaliforms by having grooved spines in both dorsal fins, teeth in the lower jaw larger than those in the upper jaw, and precaudal pit and lateral keels absent from the caudal peduncle. Gulper sharks comprise two genera and 14 species, which are represented in the eastern Atlantic from Iceland to South Africa, and in the Indian and western Pacific oceans. They range in size from about 79 to 160 cm (31 to 63 in.).

Etmopteridae (lantern sharks). Etmopteridae is a family of small deep-water sharks comprising five genera and 41 species. The family is characterized by the following combination of characters: both dorsal fin spines grooved; spines usually quite large, with second spine larger and both spines free of fin; upper lobe of caudal fin with subterminal notch; and luminous organs usually present on body. Species range in adult total length from 17 to 84 cm (6.7 to 33 in.). Lantern sharks occur on continental and insular slopes from tropical to temperate zones of the Atlantic (northward to Iceland), Indian, and Pacific oceans. *Etmopterus*, the principal genus, comprises 31 species.

Somniosidae (sleeper sharks). The family Somniosidae comprises seven genera and 17 species of relatively small sharks, ranging in length from about 49 to 114 cm (19 to 45 in.). Sleeper sharks differ from other squaliform sharks by usually lacking dorsal fin spines; however, if present, spines are small and in both fins.

They also have a lateral ridge on each side between the pectoral and pelvic fins, as well as luminous organs in most species. Sleeper sharks occur mostly on continental and insular shelves and slopes (some are oceanic) of the Atlantic, Indian, and Pacific oceans, and on Arctic and sub-Antarctic shelves. One of the sleeper sharks, the Portuguese shark (*Centroscymnus coelolepis*), which occurs in the western North Atlantic and eastern Atlantic from Iceland to South Africa, ranges in depth from 270 to 3675 m (885 to 12,055 ft).

Oxynotidae (rough sharks). Oxynotidae, comprising one genus (*Oxynotus*) and five species, is the most bizarre family of squaliform sharks. The family is characterized as follows: trunk deep and compressed, triangular-shaped in cross section; prominent ventrolateral ridge between pectoral and pelvic fins; dorsal fins saill-ike, with the first extending forward over the pectoral fins and gill slits; both dorsal fins with spines embedded in the fin, with only tips exposed; skin very rough; and most species with luminous organs. The family occurs on continental and insular slopes of the western and eastern Atlantic (including the Mediterranean Sea) and the western Pacific at depths of 500 to 720 m (1640 to 2360 ft). The species range in length from 49 to 150 cm (19 to 59 in.).

Dalatiidae (kitefin sharks). Dalatiidae, comprising seven genera and 10 species (perhaps more), is characterized as follows: small size, ranging from 25 to 50 cm (10 to 20 in.); dorsal fins without spines except in genus *Squaliotus*, which may have small spines in the first dorsal; and luminous organs, mostly on belly. Kitefin sharks occur on continental and insular shelves and slopes of tropical and temperate zones of the Atlantic, Indian, and Pacific oceans. The pygmy shark (*Euprotomicrus bispinatus*), only 27 cm (10.6 in.) in maximum length, occurs at or near the surface at night and may make lengthy daily vertical migrations of 1500 m (4900 ft) or more, a round-trip of almost 2 mi. Species of the genus *Isistius* are called cookiecutter sharks (after the cookie-shaped bite wounds that they inflict on the bodies of prey). They have a short bulbous snout, and their large, fleshy suctorial mouth allows the shark to attach itself to prey, bony fishes and cetaceans, much in the manner employed by lampreys. When attached, the cookiecutter shark bites circular chunks of flesh from its prey. Its lower teeth, quite different from the uppers, are very large (in fact, in proportion to its size, the largest teeth of all sharks), erect, bladelike, and interlocking, making a formidable weapon for this relatively small shark.

Herbert Boschung

Bibliography. L. J. V. Compagno, *FAO Species Catalogue*, vol. 4: *Sharks of the World: An Annotated and Illustrated Catalogue of Shark Species Known to Date, Part 1, Hexanchiformes to Lamniformes*, FAO Fish. Synop. no. 125, pp. 212–249, 1984; L. J. V. Compagno, Checklist of Chondrichthyes, pp. 503–547 in W. C. Hamlett (ed.), *Reproductive Biology and Phylogeny of Chondrichthyes: Sharks, Batoids, and Chimaeras*, Science Publishers, Enfield, NH, 2005; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Squall

A strong wind with sudden onset and more gradual decline, lasting for several minutes. Wind speeds in squalls commonly reach 30–60 mi/h (13–27 m/s), with a succession of brief gusts of 80–100 mi/h (36–45 m/s) in the more violent squalls. Squalls may be local in nature, as with isolated thunderstorms, or may occur over a wide area in the vicinity of a well-developed cyclone, where the squalls locally reinforce already strong winds. Because of their sudden violent onset, and the heavy rain, snow, or hail showers which often accompany them, squalls cause heavy damage to structures and crops and present severe hazards to transportation.

The most common type of squall is the thundersquall or rain squall associated with heavy convective clouds, frequently of the cumulonimbus type. Such a squall usually sets in shortly before onset of the thunderstorm rain, blowing outward from the storm and generally lasting for only a short time. It is formed when cold air, descending in the core of the thunderstorm rain area, reaches the Earth's surface and spreads out. Particularly in desert areas, the thunderstorm rain may largely or wholly evaporate before reaching the ground, and the squall may be dry, often associated with dust storms. *See* DUST STORM; SQUALL LINE; THUNDERSTORM.

Squalls of a different type result from cold air drainage down steep slopes. The force of the squall is derived from gravity and depends on the descending air which is colder and more dense than the air it replaces. So-called fall winds of this kind are common on mountainous coasts of high latitudes, where cold air forms on elevated plateaus and drains down fiords or deep valleys. Channeling of the air through narrow valleys increases their force. Squall winds (bora) of 100 mi/h (45 m/s) or more are observed in winter, when arctic air masses from Russia intermittently spill over the mountains of former Yugoslavia into the warm coastal regions of the Adriatic.

Violent squalls also characterize the warm foehn winds of the Alps and the similar chinook winds on the eastern slopes of the Rocky Mountains. Sometimes destructive, especially in winter, they are characteristically very gusty and variable over short distances. Strong surface wind speeds are a response to the distribution of pressure established in a standing breaking wave, somewhat analogous to standing waves in water flowing over and around submerged obstacles. Momentum transferred to the surface layers is dissipated in small-scale turbulence, an effective dissipative process affecting jet streams worldwide. *See* CHINOOK; JET STREAM; STANDING WAVE; WIND.

Chester W. Newton; Edwin Kessler

Bibliography. C. D. Ahrens, *Meteorology Today: An Introduction to Weather, Climate, and the Environment*, 6th ed., 1999; C. D. Whiteman, *Mountain Meteorology: Fundamentals and Applications*, 2000.

Squall line

A line of thunderstorms, near whose advancing edge squalls occur along an extensive front. The thundery region, 12–30 mi (20–50 km) wide and a few hundred to 1250 mi (2000 km) long, moves at a typical speed of 15 m/s (30 knots) for 6–12 h or more and sweeps a broad area. In the United States, severe squall lines are most common in spring and early summer when northward incursions of maritime tropical air east of the Rockies interact with polar front cyclones. Ranking next to hurricanes in casualties and damage caused, squall lines also supply most of the beneficial rainfall in some regions. *See* SQUALL.

A squall line may appear as a continuous wall of cloud, with forerunning sheets of dense cirrus, but severe weather is concentrated in swaths traversed by the numerous active thunderstorms. Their passage is marked by strong gusty winds, usually veering at onset, rapid temperature drop, heavy rain, thunder and lightning, and often hail and tornadoes. Turbulent convective clouds, 6–9 mi (10–15 km) high, present a severe hazard to aircraft, but may be circumnavigated with use of radar. *See* AIRBORNE RADAR.

Formation requires an unstable air mass rich in water vapor in the lowest 0.6–2 mi (1–3 km), such that air rising from this layer, with release of heat of condensation, will become appreciably warmer than the surroundings at upper levels. Broad-scale flow patterns vary; **Fig. 1** typifies the most intense outbreaks. In low levels, warm moist air is carried northward from a source such as the Gulf of Mexico. This process, often combined with high-level cooling on approach of a cold upper trough, can rapidly generate an unstable air mass. *See* THUNDERSTORM.

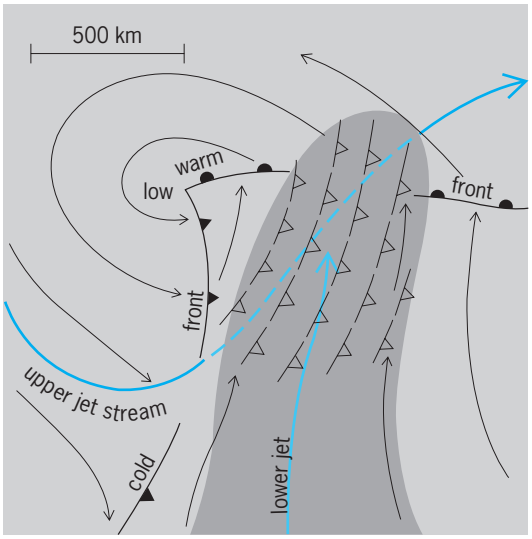

Fig. 1. Successive locations of squall line moving eastward through unstable northern portion of tongue of moist air (dark area) in warm sector of a cyclone. Black arrows show general flow in low levels; colored arrows, axes of strongest wind at about 0.6 mi (1 km) above ground and at 6–7 mi (10–12 km).

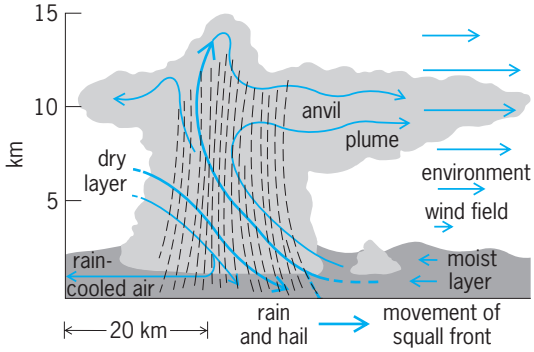

Fig. 2. Section through squall-line-type thunderstorm. Precipitation is represented by black broken lines. 1 km = 0.6 mi.

The instability of this air mass can be released by a variety of mechanisms. In the region downstream from an upper-level trough, especially near the jet stream, there is broad-scale gentle ascent which, acting over a period of hours, may suffice; in other cases frontal lifting may set off the convection. Surface heating by insolation is a contributory mechanism; there is a preference for formation in midafternoon although some form at night. By combined thermodynamical and mechanical processes, they often persist while sweeping through a tongue of unstable air, as shown in Fig. 1. Squall lines forming in midafternoon over the Plains States often arrive over the midwestern United States at night. *See* STORM.

Figure 2 shows, in a vertical section, the simplified circulation normal to the squall line. Slanting of the drafts is a result of vertical wind shear in the storm environment. Partially conserving its horizontal momentum, rising air lags the foot of the updraft on the advancing side. In the downdraft, air entering from middle levels has high forward momentum and undercuts the low-level moist layer, continuously regenerating the updraft. Buoyancy due to release of condensation heat drives the updraft, in whose core vertical speeds of 66–132 mi/h (30–60 m/s) are common near the midlevel of the storm. Rain falling from the updraft partially evaporates into the downdraft branch, which enters the storm from middle levels where the air is dry, and the evaporatively chilled air sinks, to nourish an expanding layer of dense air in the lower 0.6–1.2 mi (1–2 km) that accounts for the region of higher pressure found beneath and behind squall lines. In a single squall-line thunderstorm about 12 mi (20 km) in diameter, 1–2 $\times$ 10^7 lb/s (5–10 kilotons/s) of water vapor may be condensed, half being reevaporated within the storm and the remainder reaching the ground as rain or hail. *See* MESOMETEOROLOGY.

Chester W. Newton

Bibliography. C. D. Ahrens, *Meteorology Today: An Introduction to Weather, Climate, and the Environment*, 6th ed., 1999; E. Kessler (ed.), *Thunderstorm Morphology and Dynamics*, 2d ed., 1992; F. K. Lutgens, E. J. Tarbuck, and D. Tasa, *The Atmosphere: An Introduction to Meteorology*, 10th ed., 2006.

Squamata

The dominant order of living reptiles, composed of the lizards and snakes. The group first appeared in Jurassic times and today is found in all but the coldest regions. Various forms are adapted for arboreal, burrowing, or aquatic lives, but most squamates are fundamentally terrestrial. There are over 7700 extant species: 4800+ lizards and 2900+ snakes. *See* REPTILIA.

The order is readily distinguished from all known reptiles by its highly modified skull. The immediate ancestors of primitive lizards are members of the diapsid order Eosuchia, in which a pair of temporal foramina are present on both sides of the head. A similar pattern is found among living reptiles in the tuatara, order Sphenodonta (Rhynchocephalia), and the crocodylians, order Crocodylia. In the Squamata, however, in the case of the quadratojugal and jugal bones, which border the lower margin of the inferior temporal opening, the former has been lost completely and the latter greatly reduced. In consequence, there is only a single temporal opening present, and the quadrate has become enlarged and movable. Even this temporal opening is lost or reduced in many forms. No other reptiles show these modifications, which allow for great kinesis in the lower jaw since it articulates with the quadrate. In addition, the order is distinct from other living reptile groups because its members have no shells or secondary palates, and the males possess paired penes (hemipenes). Another distinguishing characteristic of the squamates is the synchronous shedding of the outer layer of the integument, called ecdysis. The rate of this shedding is not cyclic, but is based fundamentally on the rate of growth of the individual organism, which in turn is based on resource acquisition. *See* CROCODYLIA; EOSUCHIA; LEPIDOSAURIA; SPHENODONATA.

Traditionally, the order Squamata has been divided into two major subgroups: the lizards, suborder Sauria (= Lacertilia), and the snakes, suborder Serpentes. The latter group is basically a series of limbless lizards, and it is certain that snakes are derived from some saurian ancestor. There are many different legless lizards, and it has been suggested that more than one line has evolved to produce those species currently grouped together as snakes. Besides this possibility, there is the additional problem of separating the two suborders. No single attribute will suffice for this purpose, but the features in the **table**, used in combination, will definitely allocate any questionable form to one or the other group.

Sauria

The majority of lizards are insectivorous, but a few species feed on plants. Others, notably the Varanidae and allies, feed on larger prey, including birds and mammals. The largest living lizard is the Komodo dragon (*Varanus komodoensis*) [**Fig. 1**], which attains a length of 10 ft (3 m) and a weight of 250 lb (113 kg). Several small geckos (**Fig. 2**) and African chameleons are about 1.5 in. (4 cm) long when fully grown. *See* CHAMELEON; GECKO.

Fig. 1. Komodo dragon (*Varanus komodoensis*), Komodo National Park, Indonesia.

Physiology. The senses of lizards include the usual ones found in all terrestrial vertebrates. Thermal, tactile, and pain receptors are distributed over the body surface.

Olfaction. The principal chemoreceptors are the nasal organ proper, and the organ of Jacobson or the vomeronasal organ, a specialized derivative of the olfactory apparatus. The nasal organs open into the mouth cavity through the internal nares and are connected to the outside by the external nares. Inspired air passes through the organ on its way to the throat and lungs. Small particles of substances in the air are trapped by the nasal mucosa and analyzed by sense cells in this lining. The paired organs of Jacobson lie anterior to the olfactory sacs and are separated from them. They open only into the mouth, and particles in watery solution from this region are analyzed by them. In many lizards the tongue has a divided tip, and the two divisions are used to pick up particles outside the mouth and bring them back into it. The tips are then rubbed against the openings of the vomeronasal organs. Jacobson's organs are indicated as separate areas of the nasal organ in Amphibia and turtles, are partially separated from the main sac in Rhynchocephalia, and are vestigial in crocodiles, birds, and most mammals. *See* CHEMORECEPTION; OLFACTION; SENSE ORGAN.

Hearing. The auditory apparatus of lizards shows remarkable variation correlated with the mode of life. In the majority of forms, a well-developed eardrum is present in a slight recess at the posterior margin of the head near the quadrate bone. A ligament attaches the small middle-ear ossicle, the stapes, to the inner face of the tympanum, and the other end

Fig. 2. Western banded gecko (*Coleonyx variegatus*). (*Photo by NPS/Tonto National Monument, Arizona*)

Comparison of Sauria and Serpentes

Determining structure	Sauria	Serpentes
Limbs	Usually two or four; sometimes none	None
Eyelids	Usually movable; sometimes immovable	Immovable
External ear opening	Usually present; sometimes lacking	None
Pectoral girdle	Usually present; rarely none	None
Braincase	Usually open anteriorly, bounded by connective tissue sheath instead of bone	Completely bony anteriorly
Mandible	Two halves, usually articulating	Two halves connected by elastic ligament, not articulating

fits into an opening into the inner-ear region. Sound is transmitted from the eardrum along the stapes to the fluid of the inner ear. In many burrowing lizards, the eardrum is partially or completely covered by skin and scales or is lost entirely. Most of these forms are insensitive to airborne sounds and receive their auditory stimuli from the substratum through the lower jaw to the quadrate and then to the stapes. *See* EAR (VERTEBRATE); HEARING (VERTEBRATE); PHONORECEPTION.

Vision. Lizards have moderately good vision except for burrowing forms in which the eyes are reduced or covered over by skin, scales, or even bone. Most lizards with unreduced eyes have color vision, but some nocturnal families such as the Gekkonidae (geckos) and Xantusiidae (night lizards) are colorblind. In certain forms, the movable eyelids have a transparent window through which an image may be seen even when the eyes are closed. In several groups, the two eyelids have fused and are completely transparent. The fused lids form a protective spectacle in geckos, night lizards, certain burrowing skinks, and other forms. *See* COLOR VISION; EYE (VERTEBRATE); VISION.

Locomotion. The majority of lizards are quadrupedal in locomotion and are usually ambulatory scamperers or scansorial (adapted for climbing) [**Fig. 3**]. Some forms are bipedal, at least when in haste. Perhaps the most famous of the bipedal forms is the Jesus Cristo lizard, *Basiliscus*, of Middle America which runs for some distance across the surface of small streams when frightened. In many burrowing lizards, the limbs are reduced so that sometimes only the anterior pair is present or all are gone. In these species, locomotion is of a serpentine, curvilinear type. One genus of Asian lizards, *Draco*, has large membranes between its limbs that can be used as gliding surfaces to increase the distances of leaps and to reduce the rate of descent.

Fig. 3. Tree lizard (*Urosaurus ornatus*). (*Photo by NPS/Tonto National Monument, Arizona*)

Thermoregulation. Lizards, in common with all other reptiles, are dependent upon external sources for maintenance of their body temperatures. They are ectothermic. Thermoregulation is attained by means of gross behavioral adjustments. A lizard moves into shade if its temperature rises, and moves into areas of higher temperatures as its temperature drops. Individuals are most active between temperatures of 60 and 108°F (16 and 42°C) with optimum activity between 80 and 90°F (27 and 32°C). Temperatures above 110°F (43°C) are lethal, and no lizard can survive in direct desert sunlight for more than 15–30 min. *See* TEMPERATURE ADAPTATION; THERMOREGULATION.

Respiration and excretion. Respiration is achieved by means of paired lungs. The excretory product of protein metabolism is an insoluble substance, uric acid. The feces are dark, oblong, and capped by a white mass of uric acid when deposited. *See* EXCRETION; RESPIRATION; RESPIRATORY SYSTEM.

Coloration. The coloration of each species of lizard is characteristic. Most forms exhibit marked differences in coloration between the sexes, at least during the breeding season, and frequently the young are markedly different from the parents. Color changes occur in rapid fashion among some species, and all are capable of metachrosis or changing color to a certain extent. Contrary to popular legend, the changes are not made to match the color of the background; they seem, rather, to be under neurohormonal control and to fluctuate with temperature and light and with the activities of the lizard. *See* CHROMATOPHORE; SEXUAL DIMORPHISM.

Sound production. Many lizards make sounds ranging from hisses that are produced by simple expulsion of air through the glottis, to squeaks of small forms and loud cluckings of the geckos. Sounds are used for defensive bluffing and probably for recognition as well.

Defense mechanisms. The basic defense against discovery utilized by most lizards is "freezing" motionless in position. The majority are protectively colored and when still are extremely difficult to see. However, if they are located, rapid flight becomes the second line of defense. Even if a predator catches a lizard, the game is not over, because many forms bite and claw. Among the more interesting specialized

defensive devices of saurians are the bluffing techniques of certain forms. The Australian frilled lizard, *Chlamydosaurus*, of the family Agamidae, can expand a huge nuchal collar which when presented with open mouth to a predator is an effective deterrent to aggression. The agamid, *Uromastix*, of Africa and the iguanid, *Ctenosaura*, of Middle America have spiny tails that may lash out to injure a potential enemy. The horned lizards, *Phrynosoma*, of the United States and Mexico are armed with numerous spines, particularly on the head, and make a rather unpleasant mouthful. In addition, they have the capacity of squirting blood from ruptured vessels in the lower eyelids into the mouth or face of a predator. The blood contains chemical substances, secreted into it in the eye region, that are extremely repugnant to coyotes and dogs. A great many lizards are able to throw off their tails voluntarily (tail autotomy), and reflexes in the tail make it thrash about to distract a predator from its real objective. In many cases, the tail is brightly colored to add to the illusion. The tail will regenerate on the lizard, but the new tail will lack a bony support. *See* PROTECTIVE COLORATION; REGENERATION (BIOLOGY).

Venomous species. There are only two species of venomous lizards, both members of the genus *Heloderma*, in the family Helodermatidae. The Gila monster (*H. suspectum*) [**Fig. 4**] is found in western New Mexico, Arizona, extreme southern Nevada, southwestern Utah, eastern California, and northwestern Mexico. The beaded lizard (*H. horridum*) is a Mexican species. Venom is produced in glands along the lower jaw and penetrates wounds inflicted by the anterior recurved teeth (that is, teeth that curve back toward the animal). No mechanism for injection of the venom into the wound is present, contrary to the situation in most snakes. Although occasionally fatal, the bites of these large lizards, of which a length of 3 ft (0.9 m) is not uncommon, are not as dangerous as those of snakes, nor are human beings bitten frequently.

Rhythmic activity. Saurians show rather definite rhythms in their activities. In temperate regions, seasonal cycles are correlated with temperature and sexual activity, and hibernation occurs in the winter. General activity continues year-round in tropic climates. Daily activity cycles are typical of most species and include a definite daily pattern of feeding, drinking, resting, basking, and excreting. Most lizards are diurnal, and their daily cycles are usually correlated with temperature. Certain families, notably the Gekkonidae and Xantusiidae, tend strongly toward nocturnal habits.

Fig. 4. Gila monster (*Heloderma suspectum*). (*Photo by NPS/Tonto National Monument, Arizona*)

Territoriality. Most lizards spend their entire life within a rather restricted area and return at the end of each day to the same site to sleep. The males of the majority of forms exhibit territorial behavior and defend a central portion of their home range against trespass by other males of the same species. Defense of the territory usually involves much bluffing and noise, and the "defending" lizard is usually the victor. Various color patches, scale crests, and other distinctive marks play a part in intimidation and display at these times. Fights frequently occur and generally consist of attempts at biting. Male monitor lizards, Varanidae, rear up on their hindlegs and bite and slash with the front claws. No truly gregarious lizards are known, and none are migratory. *See* TERRITORIALITY.

Courtship activity. Males usually breed with any female that they meet in moving about their home ranges during the breeding season. A complex courtship pattern involving a chase and considerable prenuptial nuzzling and biting is typical. During copulation the male frequently grasps the head or neck of the female in his mouth, and the tails of the breeding pair are entwined so that their cloacas are brought into contact. The male has paired hollow hemipenes which normally lie retracted into the base of the tail. When engorged with blood, these organs are everted to be used singly.

Reproduction. Most species produce eggs that have leathery or calcareous shells. The eggs are buried in the soil or hidden in decaying logs or under bark. In many cases, the female guards the eggs against predators. Some species are ovoviviparous and retain the eggs within the body until they hatch. Others are truly viviparous with a seroallantoic placental connection between mother and embryo. Gestation or incubation is generally 5–8 weeks. In oviparous and ovoviviparous forms, at least, the young have a special egg tooth which grows up from the tip of the upper jaw. This tooth is used to cut through the egg membranes and the shell at the time of hatching and is lost shortly thereafter.

Classification of lizards. The following list indicates the major evolutionary lines, families, and distribution of lizards. Families indicated by an asterisk contain limbless, snakelike species (based on M. J. Benton, 2005).

Iguania line
- Family Iguanidae: the Americas, Madagascar, Fiji
  - Agamidae: Africa, southern and central Asia, Australia
  - Chamaeleonidae: Africa, Madagascar, southern India, Near East, southern Spain

Gekkota line
- Family Gekkonidae: circumtropical, all continents and most continental and oceanic islands
  - Pygopodidae: Australia, New Guinea

Scincomorpha line
Family Xantusiidae: North and Middle America, Cuba
Teiidae: the Americas
Lacertidae: Africa, Eurasia
*Gerrhosauridae: Africa, Madagascar
*Scincidae: cosmopolitan, except frigid areas
Gymnophthalmidae: Central and South America
Cordylidae: southern Africa
Annulata line
*Family Amphisbaenidae: tropics of Africa and America, Iberian Peninsula
*Bipedidae: Mexico
Trogonophidae: northern Africa and Middle East
Rhineuridae: Florida (U.S.)
Anguimorpha line
*Family Anguidae: the Americas, Eurasia
*Anniellidae: the Californias
Xenosauridae: southeastern China, Central America
Helodermatidae: North America
Varanidae: Africa, southern Asia, Australia
Lanthonotidae: North Borneo

Serpentes

Snakes (**Figs. 5** and **6**) are basically specialized, limbless lizards which probably evolved from burrowing forms but have now returned from subterranean habitats to occupy terrestrial, arboreal, and aquatic situations. In addition to those features already mentioned as distinguishing the snakes from lizards, the following characteristics are typical of all serpents. There is no temporal arch so that the lower jaw and quadrate are very loosely attached to the skull. This gives the jaw even greater motility than is the case in lizards. The body is elongate with 100–200 or more vertebrae, and the internal organs are elongate and reduced. A spectacle covers the eye, and there is no tail autotomy. In most features of their biology, serpents resemble lizards; the two groups are closely related.

The largest living snake is the Indian python (*Python reticulatus*), which reaches 30 ft (9 m) in length and a weight of 250 lb (113 kg). The largest venomous snake is the king cobra (*Ophiophagus hannah*) of southern Asia, which is known to attain a length of 18 ft (5.5 m). A number of small tropical burrowing snakes are vastly smaller by comparison, being about 5 in. (12.5 cm) in length when fully grown. The longevity of snakes is not well known, but one captive cobra lived 28 years and many of the larger snakes probably live at least 25 years.

Fig. 5. Western coral snake (*Micruroides euryxanthus*). (*Photo by NPS/Tonto National Monument, Arizona*)

Fig. 6. Eastern ribbon snake (*Thamnophis sauritus*). (*Photo by NPS/Little River Canyon National Preserve, Alabama*)

Physiology. The senses of snakes are fundamentally similar to those of lizards. Great dependence is placed upon olfaction and the Jacobson's organs. The tongue of all snakes is elongate and deeply bifurcated. When not in use it can be retracted into a sheath located just anterior to the glottis, but it is protrusible and is constantly being projected to pick up samples for the organs of Jacobson from the surrounding environment. Snakes are deaf to airborne sounds and receive auditory stimuli only through the substratum via the bones of the head. The eyes are greatly modified from those in lizards, and there is no color vision. Some groups are totally blind and have vestigial eyes covered by scales or skin.

Thermoreception. Specialized thermoreceptors, not found in any other vertebrates, are present in many boas and pythons (family Boidae, and in all pit vipers) (subfamily Crotalinae). In the boids, these receptors form a series of depressions in the lip scales and are called labial pits. In the crotalids, there is a single, large loreal pit located about midway between the eye and nostril and somewhat below them on the side of the head. These organs are sensitive to temperature differences of about 2°F (1.1°C) and are apparently used to locate prey whose temperature may be higher or lower than that of the surrounding environment.

Locomotion. Four basic patterns of locomotion are found in snakes, and several may be used by a particular individual at different times. The most familiar type is serpentine curvilinear (lateral undulatory). In this pattern, the snake moves forward by throwing out lateral undulations of the body and pushing them against any irregularity in the surface. Snakes using rectilinear locomotion move forward in a straight line, without any lateral undulations, by producing wavelike movements in the belly plates. Laterolinear locomotion, or sidewinding, is used primarily on smooth or yielding surfaces and is very complex. In essence, the snake anchors a portion of the body in the substratum and lifts the rest out laterally to a new

position. When the lifted part is anchored again, the portions of the body left behind are lifted to the new position. By a constant lifting and anchoring of alternate parts, the snake moves in a lateral direction. Concertina locomotion movement resembles the expansion and contraction of that musical instrument.

Other activities. The basic bodily functions of snakes—thermoregulation, respiration, excretion, and elimination—are the same as those in lizards.

Defense mechanisms. Snakes and lizards are alike in utilizing procrypsis and flight as their principal means of defense. Protective coloration is of prime importance in the first of these responses. Although all snakes are capable of some color change, none show the marked changes characteristic of many lizards. Snakes are excellent bluffers, and hissing and striking, sometimes with the mouth closed, may be used to intimidate or discourage possible predators. Among the most effective deterrents to attack are the specialized warning devices of the rattlesnakes (**Fig. 7**). Rattlesnakes, however, are not bluffing, because they can back up their threats with an injection of venom. These rattles are actually an extension of the integument; moreover, each time the individual organism sheds (ecdysis), a new rattle is added. Contrary to popular lore, the age of the snake cannot be determined by the number of rattles. Some snakes may shed once over a year period, while other snakes may shed three times.

Ecology. The seasonal and daily activity patterns of snakes are similar to those described above for lizards. More snakes are nocturnal than diurnal, and they generally have a larger home range than do lizards. Home sites are found in most species, but no territorial behavior is known. A "combat dance" in which the males rear upright and push and weave around one another occurs frequently. The dance has often been described as being a courtship display, but it always involves two males. Its exact function remains a mystery. Snakes are generally nongregarious, although the sea snakes, Elapidae, travel in large schools. Migrations occur in temperate regions to and from den sites, and the sea snakes migrate annually to breeding grounds.

Fig. 7. Arizona black rattlesnake (*Crotalus viridis cerberus*). (*Photo by NPS/Tonto National Monument, Arizona*)

Courtship. Courtship is distinctive for each form but usually involves tracking of the female by the male, considerable nuzzling and rubbing of the female by the male, and a wrapping of the bodies of the breeding pair tightly around one another for mating. Male snakes utilize their hemipenes one at a time. Eggs are buried or hidden and are leathery or calcareous. Development is oviparous, ovoviviparous, or viviparous. Incubation or gestation takes about 60–100 days.

Nutrition. Food is almost always of animal material such as amphibians, reptiles, birds and their eggs, and mammals. However, some forms eat lizard and snake eggs to a large extent, as well as insects, mollusks, or fish; most snakes probably eat carrion. The majority of species are rather specific in food preference. Prey may simply be grabbed and swallowed, or specialized techniques may be employed, depending upon the species. Constriction of the victim is practiced by many forms. In this case, the food organism is large and is grasped in the snake's mouth while coils of the muscular body are wrapped around the prey. The coils are gradually tightened until the victim's heart stops or suffocation occurs. Venom may be injected into the body of the prey by certain forms. Food articles are always swallowed whole. The process is complex but consists primarily of a series of coordinated movements of the head and jaws. The lower jaws are connected to one another by an elastic ligament which allows each half to be moved independently of the other. The teeth on the jaws are recurved so that by alternately moving the right half of the jaw forward and then the left and hooking the teeth into the prey the snake pulls itself forward over the prey's body.

Venomous snakes. The vast majority of living snakes are harmless to humans, although a number are capable of inflicting serious injury with their venomous bites. The venom apparatus has evolved principally as a method of obtaining food, but it is also advantageous as a defense against attackers. Venomous snakes may be placed in three categories on the basis of their venom apparatus. Opisthoglyphodont snakes have fangs in the rear of the mouth. Proglyphodont snakes have fangs in the front of the mouth. This group can be subdivided into the proteroglyphodont in which the anterior fangs are fixed and immovable—for example, cobras, mambas, kraits, coral snakes, and sea snakes—and the solenoglyphodont in which the fangs are inserted on the movable maxilla so that they may be rotated forward when the mouth is open and folded back against the roof of the mouth when not in use, as in vipers.

Fangs are teeth modified for the injection of venom into the victim, and the venom glands are modified salivary glands connected to the grooved fangs by a duct. Special muscles are present in all proglyphous snakes to force the venom into the wound. The

venom itself is a complex substance containing a number of enzymes. Certain of these enzymes attack the blood, others the nervous system, and some are spreaders. The venom of each species is distinct and contains different kinds, combinations, and proportions of the deleterious enzymes.

The rear-fanged and proteroglyphodont forms have relatively short fangs and require some chewing of the victim in order to inject the venom. Small rear-fanged forms are consequently harmless to human beings because they cannot get their mouths around any part of a person's body to bring the fangs into contact. Vipers, however, have very long fangs, and they usually strike or bite with blinding speed. Their injection mechanism is efficient, and a bite usually means that a considerable amount of venom has been injected.

Classification of snakes. The following list indicates the major groups of living snakes and their distribution.

Family Typhlopidae: circumtropical
- Leptotyphlopidae: circumtropical
- Anomalepididae: southern Central America and northern South America
- Aniliidae: southwestern Asia and Central America
- Anomochilidae: Malaysia, Sumatra, Borneo
- Boidae: circumtropical, western United States
- Bolyeriidae: Round Island (Indian Ocean)
- Tropidophiidae: West Indies, Central America, South America
- Acrochordidae: India, Solomon Islands, Australia
- Colubridae: cosmopolitan, except frigid areas; some species are opisthoglyphodont
- Elapidae: Africa, Asia, Australia, the Americas; all species are proglyphodont
- Viperidae: Eurasia, Africa; all species are proglyphodont

W. Ben Cash; Jay M. Savage

Bibliography. M. J. Benton, *Vertebrate Paleontology*, 3d ed., Blackwell Publishers, Oxford, 2005; F. H. Pough et al., *Herpetology*, 3d ed., Prentice Hall, Upper Saddle River, NJ, 2004; G. R. Zug, L. J. Vitt, and J. P. Caldwell, *Herpetology*, Academic Press, San Diego, 2001.

Squash

The common name for edible fruits of several species of the genus *Cucurbita*: *C. pepo*, *C. moschata*, *C. maxima*, and *C. mixta*. Those species originated in the Americas but are now grown in most countries around the world. Within squash there is a tremendous variation in size, shape, color, and usage (see **Fig. 1**).

Summer squash. The most clearly defined group is summer squash, fruit of any species of *Cucurbita* eaten as a vegetable when immature. It is most commonly *C. pepo*, and its plants are usually bushy rather than having long runners. In some countries fruits may be picked as early as the day the flowers open, but they are eaten at all stages up to the time when the rind starts to harden. Fruit color may be white, yellow, or light or dark green, and the green may be solid or striped. Shapes may be flattened disks as in Pattypan, cylindrical as in Zucchini and Cocozelle, or with necks as in the straightneck and crookneck types. Summer squash has mild flavor, high water content (about 95%), and relatively low nutritional value. It is not especially high in any one nutrient but contributes modest amounts of several vitamins and minerals.

Winter squash. Winter squash is fruit of *Cucurbita* eaten when mature and derives its name from its ability to be stored for several weeks or months without consumption. Varieties of winter squash are found in all four species. The Table Queen group,

Fig. 1. Three varieties of squash. (*a*) Zucchini. (*b*) St. Pat Scallop squash (*W. Atlee Burpee Co.*). (*c*) Crookneck squash (*Asgrow Seed Co.*, *The Upjohn Co.*).

synonymous with Acorn, is *C. pepo*, Butternut belongs to *C. moschata*, Green-striped Cushaw is *C. mixta*, while *C. maxima* has the widest range of types, including Buttercup, Hubbards, and Delicious of various colors, Banana, and Boston Marrow. Flesh color varies from light yellow to dark orange, and the edible portion ranges from thin to very thick, depending on variety.

Characteristics. The percentage of soluble solids, which are mostly sugars, in cooked winter squash is commonly 8–10% and may be as high as 16%, putting squash in the same range as good melons. Squash is a good source of calories and of vitamin A, particularly those with dark orange flesh, which are high in carotene. Varieties tend to have their own characteristic flavors: the greatest difference is between the Table Queen group and the other winter squashes. Plants of winter squash are characteristically viny in habit with runner lengths of 12 to 20 ft (3.6 to 6 m) under good growing conditions, but bush or semiviny forms have been developed in some of the types. The bush forms of winter squash tend to be lower in solids and quality, especially if they set a heavy load of fruit.

There is some confusion in terminology between winter squash and pumpkin, and the use of these names differs from place to place. Both names refer to mature fruits, and varieties called squash and others called pumpkins may be found in all four *Cucurbita* species. In the United States, varieties eaten as a vegetable are generally called squash, while those used for pies, stock feed, or autumn decorations are usually called pumpkins. Pumpkins tend to be round with orange exterior color and flesh that is paler, more fibrous, and less sweet than squash. *See* PUMPKIN.

Propagation. Squash is propagated by seed and its flowering habit is monoecious, that is, with separate male and female flowers on the same plant. Male flowers usually appear several days before the first female. Pollination is done naturally by insects, and within a species types which look very different cross readily with each other. However, the fruit is not affected in the year crossing occurs. Natural crossing occurs only rarely between the different species. Separation of the sexes makes controlled crossing relatively easy, and first-generation hybrids are increasingly being used as varieties, especially in summer squash where hybrid vigor is manifested in earlier flowering. One result of hybridization is that Butternut, Table Queen, Zucchini, and others which were originally single varieties have become groups of varieties with similar fruit shape but differing fruit sizes, colors, times of maturity, and plant sizes. *See* BREEDING (PLANT).

Squash seed is usually planted directly in the field after the soil is warm, in rows 4–6 ft (1.2–1.8 m) apart for bush type and 6–9 ft (1.8–2.7 m) apart for vining types. Plants may be spaced from 2–3 ft (0.6–0.9 m) apart in both cases. In cool regions, plants may be started in a greenhouse, provided they are grown in containers which permit transplanting without much disturbance of roots and before they become too large. The best size for transplanting may be reached in 10–14 days under good growing conditions.

Harvesting and storage. Summer squash begins to bear in 50–60 days from planting. The fruit grows very rapidly and must be harvested at frequent intervals (3–4 times per week) if one wishes the plants to continue bearing for an extended period.

Winter squash should be harvested when fully mature for best quality. Unless planted very early, it is often left in the field until the first light frost occurs, then harvested before the exposed fruit can be damaged by subsequent frost. Storage temperatures of 50–60°F (10–16°C) are best because chilling injury occurs at lower temperatures and shortens storage life. For all varieties except Table Queen, a curing period of 1–2 weeks at 70–80°F (21–27°C) may improve quality by hastening conversion of starch to sugar. H. M. Munger

Diseases. There are more than 40 diseases of squash and about 16% of the annual crop is lost due to diseases. Diseases caused by the mosaic group of viruses are of primary importance because of their widespread occurrence in all areas where the crop is grown, both in the United States and elsewhere. These viruses produce such striking symptoms in infected plants, such as malformation and mottling of the leaves (**Fig. 2**) and distortion and abnormal discoloration of the fruits (**Fig. 3**). These viruses can be transmitted by seed, insects, and mechanical contact. Effective control measures are the use of virus-free seed, eradication of weed hosts of the viruses, and elimination of insect vectors.

Seed decay is common when squash is planted in cool wet soil, so treatment of the seed with registered fungicides is essential. The fungus *Fusarium solani* causes stunting, yellowing, and wilting of squash plants. The most effective control measures against this disease are the use of noncontaminated seed and extended crop rotations. In the arid southwestern United States and in squash production areas throughout the world with similar climates, powdery mildew, *Sphaerotheca fuliginea*, is a severe production-limiting disease, This disease can be easily recognized by the abundant production of white masses of spores on the leaves of infected plants so that they appear to have been sprinkled with a

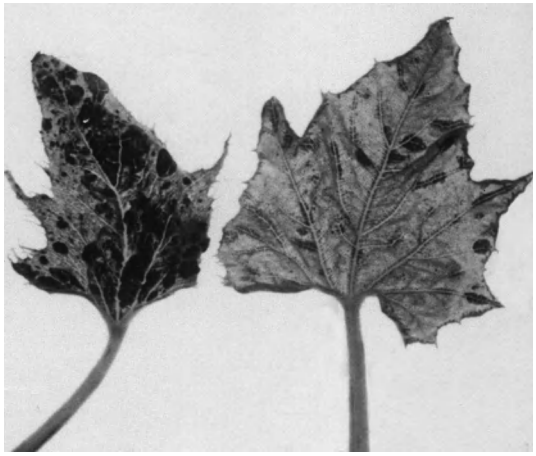

Fig. 2. Squash leaves, mottled by mosaic virus.

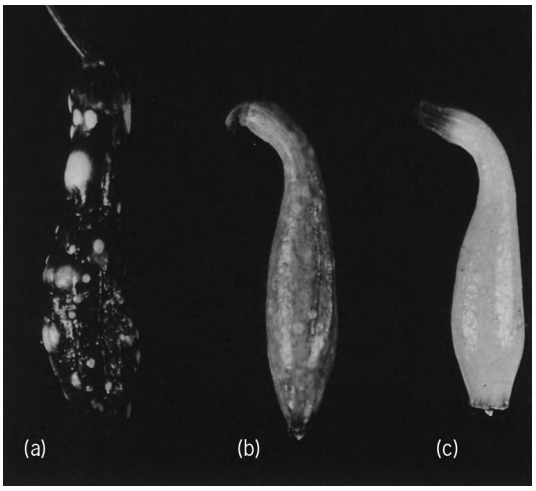

Fig. 3. Crookneck squash. (*a*, *b*) Fruit mottled by mosaic virus. (*c*) Healthy fruit.

heavy white powder. In the late 1960s, effective systemic fungicides were developed that gave excellent control of this disease. By the mid-1970s, however, the fungus had begun to develop strains resistant to these fungicides, and growers returned to the use of older, more conventional, protective fungicides. There have been efforts by plant breeders to develop squash varieties and hybrids with genetic resistance to powdery mildew.

When heavy, prolonged dew periods are prevalent during the growing season, the downy mildew fungus, *Pseudoperonospora cubensis*, severely defoliates the plants. Since no effective level of resistance to this disease is available in commercial varieties, the only practical control procedure is the application of protective fungicides at regular intervals throughout the growing season. *See* PLANT PATHOLOGY. Claude E. Thomas

Bibliography. D. M. Bates, R. W. Robinson, and C. Jeffrey (eds.), *Biology and Utilization of the Cucurbitaceae*, 1989.

Squatiniformes

An order of bizarre sharks, commonly known as angel sharks, comprising one family, one genus (*Squatina*), and 15 species that superficially resemble rays (see **illustration**). The following combination of characters identify the angel sharks: body flattened dorsoventrally; pectoral fins winglike and spread forward alongside head but not attached to it; mouth almost terminal; nostrils terminal, with barbels on anterior margin; eyes and large spiracles on top of head; five gill slits partly on sides of head and partly on ventral surface; two small spineless dorsal fins on tail; anal fin absent; and lower lobe of caudal fin longer than upper lobe. Maximum lengths of three species vary from 63 to 89 cm (25 to 35 in.), and 12 species vary from 120 to 200 cm (47 to 79 in.).

Angel sharks are found primarily in cool temperate to tropical zones of the eastern and western Pacific, eastern and western Atlantic, and extreme southwestern Indian Ocean, but they do not occur in most of the Indian Ocean or in the central Pacific reaches of Oceania (that is, the archipelagoes and islands of the central and south Pacific). They have a wide depth range, from nearshore to the upper continental slopes at 1300 m (4265 ft). They feed on a variety of small bony fishes, crustaceans, cephalopods, gastropods, and bivalves by using their highly protrusible jaws to snap up prey at high speed. *See* ELASMOBRANCHII; SELACHII. Herbert Boschung

Bibliography. L. J. V. Compagno, Checklist of Chondrichthyes, pp. 503–547 in W. C. Hamlett (ed.), *Reproductive Biology and Phylogeny of Chondrichthyes: Sharks, Batoids, and Chimaeras*, Science Publishers, Enfield, NH, 2005; L. J. V. Compagno, Checklist of living elasmobranches, pp. 471–498 in W. C. Hamlett (ed.), *Sharks, Skates, and Rays: The Biology of Elasmobranch Fishes*, Johns Hopkins University Press, Baltimore, 1999; L. J. V. Compagno, *FAO Species Catalogue*, vol. 4: *Sharks of the World: An Annotated and Illustrated Catalogue of Shark Species Known to Date, Part 2, Carcharhiniformes*, FAO Fish. Synop. no. 125, pp. 251–655, 1984; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

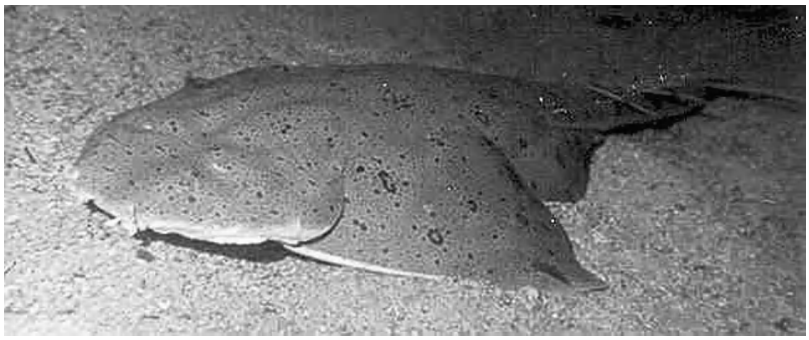

Pacific angel shark (*Squatina californica*). (*Photo by Tony Chess/NOAA Fisheries, Southwest Fisheries Science Center*)

Squeezed quantum states

Quantum states for which certain variables can be measured more accurately than is normally possible.

All matter and radiation fluctuate. Much random fluctuation derives from environmental influence, but even if all these influences are removed, there remains the intrinsic uncertainty prescribed by the laws of quantum physics. The position and momentum of a particle, or the electric and magnetic components of an electromagnetic field, are conjugate variables that cannot simultaneously possess definite values (Heisenberg uncertainty principle). It is possible, however, to have the position of a particle more and more accurately specified at the expense of increasing momentum uncertainty; the same applies to electromagnetic field amplitudes. This freedom underlies the phenomenon of squeezing or the possibility of having squeezed quantum states. With squeezed states, the inherent quantum fluctuation

may be partly circumvented by focusing on the less noisy variable, thus permitting more precise measurement or information transfer than is otherwise possible. *See* UNCERTAINTY PRINCIPLE.

Vacuum and laser light. According to quantum electrodynamics, the vacuum is filled with a free electromagnetic field in its ground state that consists of fluctuating field components with significant noise energy. If ϕ is a phase angle and $a(\phi)$ and $a[\phi + (\pi/2)]$ are two quadrature components of the field (for example, the electric and magnetic field amplitudes), the vacuum mean-square field fluctuation is given by Eq. (1), independently of the phase angle.

$$\langle \Delta a^2(\phi) \rangle = \frac{1}{4} \tag{1}$$

Equation (1) is normalized to a photon; the corresponding equivalent noise temperature at optical frequencies is thousands of kelvins. Equation (1) also gives the general fluctuation of an arbitrary coherent state, which is the quantum state of ordinary lasers. Further environment-induced randomness is introduced in addition to Eq. (1) for other conventional light sources, including light-emitting diodes. *See* COHERENCE; ELECTRICAL NOISE; LASER; QUANTUM ELECTRODYNAMICS.

Squeezed-state light. In a squeezed state, the quadrature fluctuation is reduced below Eq. (1) for some ϕ, as given in Eq. (2). At that point, squeezing, that is, reduction of field fluctuation below the coherent state level, occurs. The fluctuation of the conjugate quadrature is correspondingly increased to preserve the uncertainty relation, Eq. (3). In a two-

$$\langle \Delta a^2(\phi) \rangle < \frac{1}{4} \tag{2}$$

$$\langle \Delta a^2(\phi) \rangle \left\langle \Delta a^2 \left(\phi + \frac{\pi}{2} \right) \right\rangle \geq \frac{1}{16} \tag{3}$$

photon coherent state, or squeezed state in the narrow sense, Eq. (3) is satisfied with equality. As a consequence, for a fixed energy E, the field quadrature signal-to-noise ratio is maximized by a two-photon coherent state. In particular, it becomes $4E(E + 1)$ compared to $4E$ for a coherent state. The two-photon coherent states are nothing but a slight generalization of the well-known minimum uncertainty states already studied by E. Schrödinger. The designation squeezed state is partly derived from the fact that the noise circle of Eq. (1) is squeezed to an ellipse when Eq. (3) is satisfied with equality (see **illus.**). Such states are called two-photon coherent states partly because only even counts of photons are possible for them when the mean field amplitude is removed, which is very characteristically a quantum feature.

Generation and detection. Squeezed light can be generated by a variety of processes, especially nonlinear optical processes. To detect squeezing, an improved form of optical homodyne detection is employed called balanced homodyning. The first successful experimental demonstration of squeezing, in 1985, involved a four-wave mixing process in an atomic beam of sodium atoms. Since then, squeezing has been observed in a variety of other systems; the most pronounced was seen in parametric oscillation and amplification with about a factor of 4 below Eq. (1). *See* NONLINEAR OPTICAL DEVICES; NONLINEAR OPTICS.

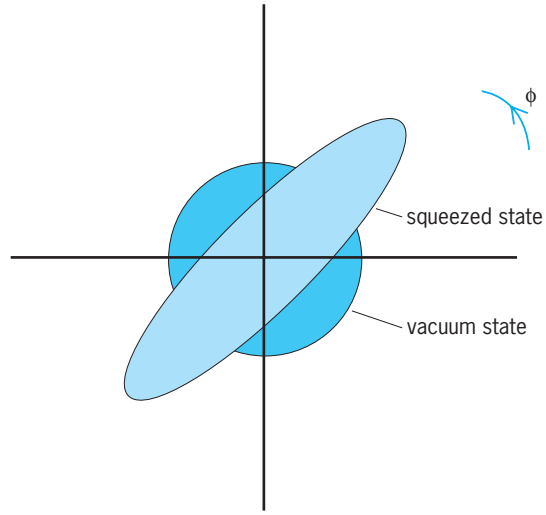

Field-amplitude fluctuation $\langle \Delta a^{+3}(\phi) \rangle$ as a function of phase angle ϕ. The noise circle of the vacuum state, Eq. (1) in the text, is squeezed to an ellipse according to Eq. (3).

Potential applications. Squeezing was first studied in connection with optical communication, although it was evident that reduced quantum fluctuation might find applications in precision measurements. Utilization of squeezed light has been proposed for interferometry, gravitational-wave detection, and spectroscopy, among other areas. The full realization of these applications is still remote. *See* OPTICAL COMMUNICATIONS.

Intensity squeezing. A phenomenon closely related to squeezing, which is also exhibited by some two-photon coherent states, is antibunching, where the photon number fluctuation in a sub-poissonian state is reduced below the coherent-state Poisson level. This sub-poissonian behavior is sometimes called amplitude squeezing or intensity squeezing. The generation, propagation, and detection of sub-poissonian light and squeezed light have similar problems; applications are similar and complementary. Strongly sub-poissonian light has been observed from constant-current-driven semiconductor lasers. *See* DISTRIBUTION (PROBABILITY).

Coherence and loss. Squeezed light and sub-poissonian light constitute the most important classes of nonclassical light, specifically, those lights whose quantum state cannot be represented as a random superposition of coherent states. Nonclassical light with space-time coherence can be generated by nonlinear optical processes. On the other hand, the quantum fluctuation characteristic is sensitive to the losses introduced during propagation and detection. Loss not only attenuates the signal energy on nonclassical light, but it also mixes in vacuum (coherent-state) fluctuation quantum-mechanically,

thus degrading the desirable quantum noise characteristic of nonclassical light. This could be a serious problem for obtaining good sources and for many applications. The loss problem could be alleviated in some situations by using appropriate quantum amplifiers, specifically, parametric amplifiers for squeezed light and photon-number amplifiers for sub-poissonian light. These amplifiers, in particular the photon-number amplifier, are not yet sufficiently developed for such applications. *See* PARAMETRIC AMPLIFIER.

Contractive states. In the context of a particle, the position and momentum take the roles of $a(\phi)$ and $a[\phi + (\pi/2)]$. A two-photon coherent state is then just a slightly generalized minimum-uncertainty wave packet. A subclass of these states called contractive states have been found to be of potential importance in gravitational-wave detection via the motion of a free mass. It has been shown that quantum measurements involving such states could, in principle, lead to arbitrarily precise monitoring of the position of a free mass, which was formerly thought to be impossible. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS. Horace P. Yuen

Bibliography. R. Loudon and P. L. Knight, Squeezed light, *J. Mod. Opt.*, 34:709–759, 1987; Special issue: Quantum noise reduction in optical systems: Experiments, *Appl. Phys. B*, vol. 55, no. 3, 1992; Special issue: Squeezed states of the electromagnetic field, *J. Opt. Soc. Amer. B*, vol. 4, no. 10, 1987; H. P. Yuen, Two-photon coherent states of the radiation field, *Phys. Rev. A*, 13:2226–2243, 1976.

Squid

The common name applied to cephalopods of the order Teuthoidea. They are marine mollusks that inhabit the oceans of the world. Squids (see **illustration**) are characterized by having eight arms and two longer tentacles around the mouth; an elongated, tapered, usually streamlined body; an internal rod- or bladelike shield (gladius); and fins on the body (mantle). The arms have two (infrequently four or six) rows of suckers and occasionally clawlike hooks, and the tentacles have terminal clubs with suckers, hooks, or both. The muscular, elastic tentacles are contractile, not retractile into pockets like those of cuttlefishes (Sepioidea). *See* SEPIOIDEA.

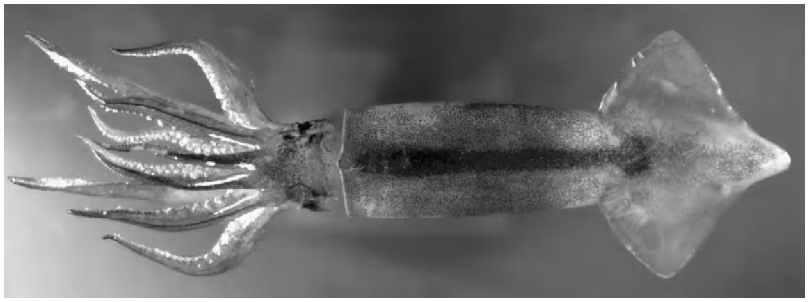

Short-finned squid (*Illex borealis*). (*Donald Flescher; image courtesy of Marine Biological Laboratory Archives*)

Squids have an exceptionally well-developed brain and organs of the central nervous system that approach in complexity and function those of fishes and even some birds and mammals. The eyes are large and provide acute vision. Thousands of pigment cells (chromatophores) in the skin are individually innervated to accommodate very rapid color and color pattern changes important in camouflage and behavior. Highly innervated photophores, capable of producing bioluminescent light, may be present on the skin, arms, or tentacles, or on the internal organs and eyes, and function to warn predators, eliminate the silhouette of deep-sea forms, or enable recognition of species or sexes. *See* BIOLUMINESCENCE; CHROMATOPHORE; NERVOUS SYSTEM (INVERTEBRATE); PHOTORECEPTION.

Squids are active, powerful swimmers, driven by jet propulsion as water taken into the mantle cavity is forcefully expelled through the funnel. Prey, normally shrimps, fishes, or other squids, are captured with the two tentacles and held with the arms while the beaks, like parrot beaks, cut off bites that the radula and tongue shove down the throat. Squids are preyed upon by toothed whales (sperm whales, dolphins), pinnipeds (seals, sea lions), fishes (tunas, mackerels, swordfish, groupers, sharks), and pelagic birds (petrels, albatrosses).

Two groups (suborders) of squids are recognized: Myopsida, like *Loligo pealei* and *Lolliguncula brevis* from the east coast of North America and the Caribbean and *Loligo opalescens* from the west coast of North America; and Oegopsida, the open-ocean and deep-sea forms like *Illex illecebrosus* of the western North Atlantic ocean and *Dosidicus gigas* of the eastern Pacific Ocean from off Chile to California.

Squids are important in biomedical research, especially that concerning neurophysiology and biophysics, in which the giant nerve axon plays a major role. Squids are eaten around the world; annual catches total about 1 million metric tons. An excellent source of protein, squid is very low in fats and high in essential minerals. *See* CEPHALOPODA; COLEOIDEA; TEUTHOIDEA. Clyde F. E. Roper

Bibliography. L. Bonnaud, R. Boucher-Rodoni, and M. Monnerot, Phylogeny of cephalopods inferred from mitochondrial DNA sequences, *Mol. Phylog. Evol.*, 7:44–54, 1997; D. L. Gilbert, W. J. Adelman, and J. M. Arnold (eds.), *Squid as Experimental Animals*, Plenum Press, New York, 1990; R. T. Hanlon and J. B. Messenger, *Cephalopod Behaviour*, Cambridge University Press, 1996; J. B. Messenger, Cephalopod chromatophores: Neurobiology and natural history, *Biol. Rev.*, 76:473–528, 2001.

SQUID

An acronym for superconducting quantum interference device, which actually refers to two different types of device, the dc SQUID and the rf SQUID.

The dc SQUID consists of two Josephson tunnel junctions connected in parallel on a

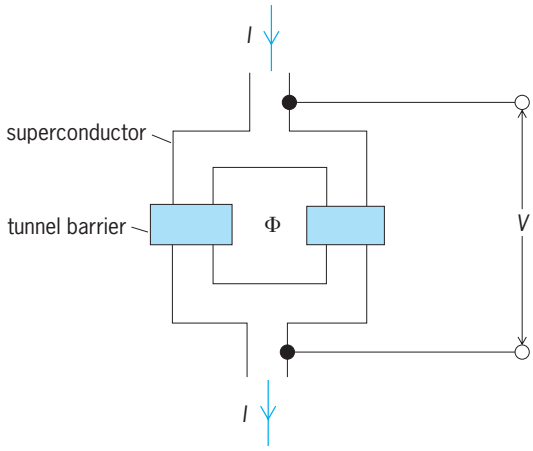

Fig. 1. Direct-current (dc) SQUID with enclosed magnetic flux Φ. *I* = applied current; *V* = generated voltage.

superconducting loop (**Fig. 1**). A small applied current flows through the junctions as a supercurrent, without developing a voltage, by means of Cooper pairs of electrons tunneling through the barriers. However, when the applied current exceeds a certain critical value, a voltage is generated. When a magnetic field is applied so that a magnetic flux threads the loop, the critical value oscillates as the magnetic flux is changed (**Fig. 2**), with a period of one flux quantum, $\Phi_0 = h/(2e) \approx 2.07 \times 10^{-15}$ weber, where h is Planck's constant and e is the electron charge. The oscillations arise from the interference of the two waves describing the Cooper pairs at the two junctions, in a way that is closely analogous to the interference between two coherent electromagnetic waves. *See* INTERFERENCE OF WAVES; JOSEPHSON EFFECT; SUPERCONDUCTIVITY.

The rf SQUID consists of a single junction interrupting a superconducting loop. In operation, it is coupled to the inductor of an *LC*-tank circuit excited at its resonant frequency by a radio-frequency (rf) current. The rf voltage across the tank circuit oscillates as a function of the magnetic flux in the loop, again with a period of one flux quantum. Although SQUIDs were for many years operated while immersed in liquid helium, ceramic superconductors with high transition temperatures make possible devices operating in liquid nitrogen at 77 K.

SQUIDs have important device applications. Usually with the addition of a superconducting input circuit known as a flux transformer, both dc and rf SQUIDs are used as magnetometers to detect tiny changes in magnetic field. The output of the SQUID is amplified by electronic circuitry at room temperature and fed back to the SQUID so as to cancel any applied flux. This makes it possible to detect changes in flux as small as 10^{-6} of one flux quantum with SQUIDs based on low-transition-temperature superconductors, corresponding to magnetic field changes of the order of 1 femtotesla in a 1-hertz bandwidth. Suitable modifications to the input circuit enable the SQUID to measure other physical quantities, including voltages, displacement, or magnetic susceptibility. SQUIDs are also used for logic and switching elements in experimental digital circuits and high-speed analog-to-digital converters. *See* ANALOG-TO-DIGITAL CONVERTER; INTEGRATED CIRCUITS; MAGNETOMETER; SUPERCONDUCTING DEVICES.

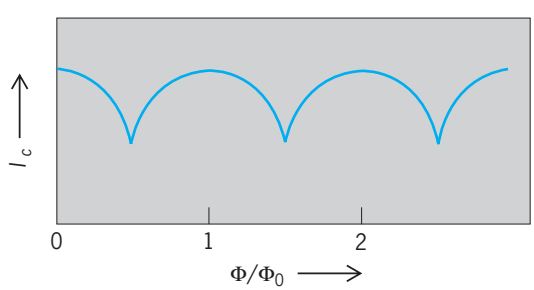

Fig. 2. Maximum supercurrent I_c versus Φ/Φ_0 for dc SQUID, where Φ = enclosed magnetic flux, Φ_0 = flux quantum.

John Clarke

Bibliography. A. Barone (ed.), *Principles and Applications of Superconducting Quantum Interference Devices*, 1992; A. Barone and G. Paterno, *Physics and Applications of the Josephson Effect*, 1982; J. Clarke, SQUIDs, *Sci. Amer.*, 271(2):46–53, August 1994; T. Van Duzer and C. W. Turner, *Principles of Superconductive Devices and Circuits*, 2d ed., 1999.

Squirrel

A rodent in the family Sciuridae. This family includes all tree squirrels and flying squirrels, as well as squirrels that live on the ground such as the woodchuck, marmot, chipmunk, and prairie dog. There are 51 living genera and 272 species living in a wide variety of habitats ranging through tropical rainforests, tundra, alpine meadows, and semiarid deserts in Europe, Asia, Africa, North America, and most of South Africa. They are absent from Australia, Madagascar, Polynesia, and the Sahara Desert. Locomotion is plantigrade (walking with the whole sole of the foot touching the ground). Sensitive vibrissae are present on the head, feet, and outsides of the legs. The dental formula is I 1/1, C 0/0, Pm 1-2/1, M 3/3 for a total of 20 or 22 teeth. The incisors grow continually throughout the animal's life. *See* CHIPMUNK; DENTITION; MARMOT; PRAIRIE DOG; RODENTIA.

Tree and flying squirrels. Tree and flying squirrels are mostly small and slender-bodied with bushy tails often as long as their bodies (see **illustration**). The tail is primarily used to maintain and correct the balance of the animal in leaping from branch to branch, but it may also be used as a rudder when leaping, as a flag to communicate social signals, and as a wraparound blanket when the animal sleeps. Soft pads are present on the soles of the feet, and sharp claws are present on all toes. These squirrels range in size from the African pygmy squirrel (*Myosciurus pumilio*) of Cameroon, Gabon, Congo, and Equatorial Guinea, which has a body length of about 60–70 mm (2–3 in.), a narrow tail of about 50–60 mm (2 in.), and weighs about 16 g (0.6 oz), to the large Oriental giant

Delmarva fox squirrel (*Sciurus niger cinereus*). (*Photo by W. H. Julian/U.S. Fish and Wildlife Service*)

squirrel (*Ratufa*) of India, Nepal, Borneo, and the Malay Peninsula, which may reach a length of over 0.9 m (3 ft) and weigh up to 3 kg (6.5 lb). Habitats vary considerably. The European pine squirrel (*Sciurus vulgaris*) and the North American red squirrel (*Tamiasciurus hudsonicus*) dwell primarily in pine and evergreen forests, whereas the gray squirrel (*S. carolinesis*) of eastern North America is a resident of groves of oak, maple, and hickory trees. The largest tree squirrel in eastern North America, the fox squirrel (*S. niger*), favors open upland groves of oak, hickory, beech, and longleaf pine rather than the deep unbroken forest.

Tree squirrels are agile and are active during the day, unlike flying squirrels, which are nocturnal and have large eyes for seeing in dim light. They hoard nuts, seeds, and fruits in holes in trees or in underground storage sites. Food caches are located by memory and also by the keen sense of smell. Nests are constructed primarily of leaves and twigs and are located either in tree cavities or in the fork of a tree. Tree squirrels are active all year, although they may remain in their nests for extended periods during severe winter weather. Breeding usually occurs in February with litters being born in March or April. Newborn squirrels are hairless, blind, and helpless. Their eyes open in about 5 weeks.

Other tree squirrels include the flame and pygmy squirrels of South America; the Indian palm (*Funambulus*) and common oriental (*Callosciurus*) squirrels of Oriental Asia and the Pacific Islands; and the oil palm (*Protoxerus*), pygmy (*Myosciurus*), sun (*Heliosciurus*), and bush (*Paraxerus*) squirrels of Africa.

Flying squirrels are not really fliers but gliders. Their "flying" equipment consists of a thin, loose furry membrane (patagium) extending along each side of the body between the fore and hind limbs. As the animal launches into the air from some lofty bough, it spreads its legs, drawing the membranes taut. This wide expanse gives the squirrel enough support to permit it to glide in a downward course for a considerable distance. Before landing, the squirrel checks its speed by manipulating its tail and comes to rest head up on the trunk of a tree. The flying squirrel cannot sustain its flight for any long period of time; it lacks propelling organs like the wings of a bat or bird. It can, however, glide through the air up to a distance of 80 yards (73 m) or more. Flying squirrels are found in forested regions of Europe, Asia, and North America south to Honduras. They are absent from South America, Africa, and Australia. Nests are usually constructed in tree cavities but may be in birdhouses. Flying squirrels are active all year and feed on nuts, seeds, berries, and insects. Lichens and fungi may be eaten during the winter. On occasion, they kill and eat sleeping birds and bird eggs.

Ground squirrels. Ground-dwelling squirrels are usually larger and more heavy-bodied than tree squirrels. They have powerful forelimbs and large claws for digging. They live in subterranean burrows or have their underground dens among rocks. Burrows usually consist of a nest chamber and several side rooms. Ground-dwelling squirrels are omnivorous. Grasses, clover, forbs, flowers, and bulbs make up the bulk of the diet, but insects, bird eggs, and even carrion may also be eaten. They are diurnal, and some, such as woodchucks and marmots, hibernate. The 14 species of marmots (*Marmota*) are found throughout most of the United States and Canada as well as in the Alps of Europe and eastward in the mountains to China and southeastern Siberia.

The five species of prairie dogs (*Cynomys*) are stout bob-tailed rodents that were famous for their "towns" in the midwestern United States. These were actual underground cities with miles of well-worn tunnels and dens extending in every direction. Some large towns may have had millions of inhabitants. A few towns still remain, but prairie dog populations were decimated by cattle ranchers who killed them with strychnine because they consumed the same forage as the herds of cattle.

Chipmunks (*Tamias*) inhabit sunlit woodlands and open pastures. They favor stone walls, rocks, and fallen timber. The 25 species are found in Asia and North America. Chief foods are seeds, grains, nuts, and berries. Food is transported in their cheek pouches to underground storage sites.

Ground squirrels (*Spermophilus*, *Ammospermophilus*, *Atlantoxerus*, *Xerus*, and *Spermophilopsis*) are robust mammals that have rounded heads, short ears, short legs, and somewhat bushy tails that are short to medium in length. The 51 species are social animals that live in densely populated towns beneath the ground. They feed largely on green vegetation, seeds, grasshoppers, cicadas, and many other insects.

Most ground squirrels hibernate, with some entering hibernation as early as July and August and not coming out until February. One litter of 7 to 10 young is produced annually. *See* HIBERNATION AND ESTIVATION.

Donald W. Linzey

Bibliography. B. Grzimek (ed.), *Grzimek's Encyclopedia of Mammals*, McGraw-Hill, 1990; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., The Johns Hopkins University Press, 1999.

SS 433

A binary star system in the Milky Way Galaxy consisting of a neutron star or a black hole accreting matter in the form of a disk from an early-type companion star. The most remarkable feature of the system is the presence of bipolar narrow outflows of matter (jets), with a velocity of 26% the speed of light, whose ejection axes change with time. *See* BLACK HOLE; NEUTRON STAR.

Discovery. SS 433 owes its name to a catalog published in 1977 by C. B. Stephenson and N. Sanduleak. They made an optical survey of the sky (the SS survey), searching for sources with a peculiarity in their emission spectrum, namely a strong hydrogen emission line. SS 433 is the 433rd object in their list of sources. Between 1967 and 1981, SS 433 was detected during other observing campaigns at different wavelengths, and appeared in many other catalogs as a radio source, an x-ray source, and as an optical variable star. Yet, the discovery that really drew the attention of scientists to this source was related to the characteristics of its hydrogen emission lines. *See* ASTRONOMICAL ATLASES; ASTRONOMICAL SPECTROSCOPY; RADIO ASTRONOMY; VARIABLE STAR; X-RAY ASTRONOMY.

Further optical observations of SS 433 in 1978 showed, besides the very strong hydrogen emission line in the spectrum (the so-called Hα line, that is, the first atomic transition in the hydrogen Balmer series, at a wavelength of 656.3 nm), a few other lines whose wavelengths did not correspond to any known element. The extraordinary discovery, at first incomprehensible, was that the wavelengths of these "strange" lines also moved with time. A well-known effect that can cause shifts in the wavelength of a line is the Doppler effect. If the gas that emits a line is moving toward the observer, the line in the spectrum appears at a shorter wavelength than the wavelength at which it has been emitted at rest with the gas (that is, blueshift of the line), while if the gas is moving away from the observer, the line appears at a larger wavelength (that is, redshift of the line). What was difficult to understand was the fact that these moving lines appeared always in pairs, and moved smoothly in a periodic pattern. The solution of the mystery came one year later. *See* ATOMIC STRUCTURE AND SPECTRA; DOPPLER EFFECT.

Geometry of the jets. All the above mentioned unusual spectral features of SS 433 can be well fit with a relatively simple geometrical model, the kinematic model (**Fig. 1**). The system is ejecting matter in form of two antiparallel narrow (that is, with an opening angle less than 5°) jets. These jets precess in a 21° half-opening angle cone with a period of about 162 days. The components of the jets (probably discrete blobs of matter) move in a straight line away from the system. Since the ejection axes rotate on a cone (like a spinning top) and the angle between the precession axes and the line of sight is about 78°, the projection of the jets onto the plane of the sky results in a twisted trace (**Fig. 2**). The blobs of matter have a constant velocity of about 0.26c, where c is the speed of light. For the first time, relativistic jets (that is, moving at a significant fraction of the speed of light, in this case 26%), which had already been discovered in distant galaxies, were observed relatively close-by, from a stellar object in our own galaxy.

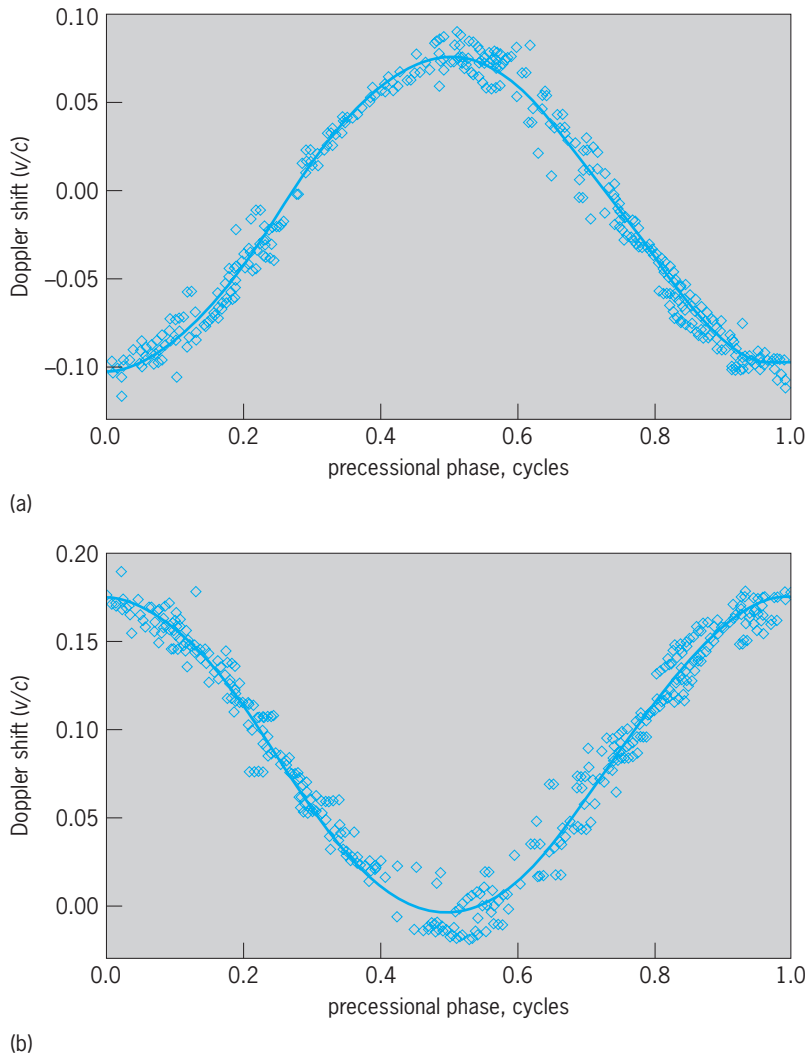

Fig. 1. Doppler shifts of the "moving" Hα lines of (*a*) the approaching jet and (*b*) the receding jet of SS 433 versus the precession phase. The Doppler shift gives the ratio *v/c*, where *v* is the approach or recession velocity and *c* is the speed of light. The data are shown as open diamonds, and the solid lines are fits of the data using the kinematic model. (*From S. S. Eikenberry et al., Twenty years of timing SS 433, Astrophys. J., 561:1027–1033, 2001. © The American Astronomical Society. Reproduced by permission of the American Astronomical Society.*)

Electromagnetic radiation from jets. The physical processes involved in the first steps of collection, launch, acceleration, and collimation of matter in the form of a jet are still poorly understood. Most jets are thought to originate in the innermost regions of an

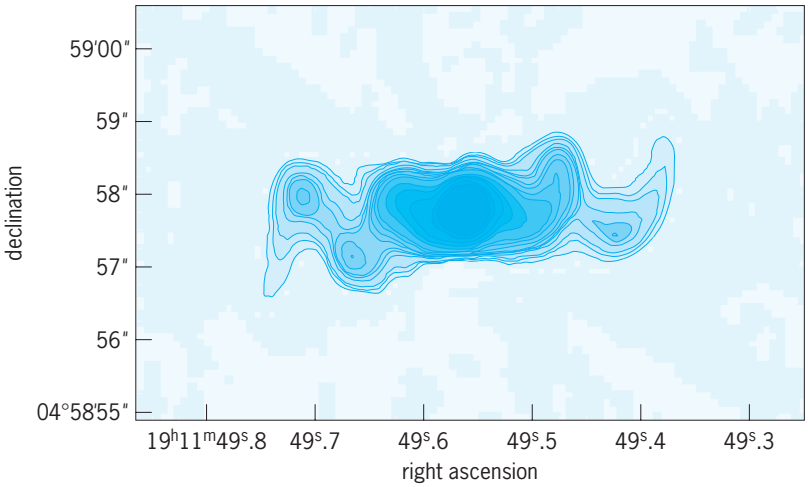

Fig. 2. Radio image of SS 433 obtained in 2003 with the Very Large Array. A contour plot of the intensity of radio emission is superimposed on a pixel image. The binary core is the brightest spot in the center, from which matter is outflowing in two anitparallel jets. On each side of the core is a twisted trace that is the projection onto the plane of the sky of the precessing jets.

accretion disk, and the outflowing axis to lie in a direction approximately perpendicular to the disk plane. Jet models for SS 433 that can be tested by observations concern the description of the electromagnetic radiation (light) emitted by the jets when they have already formed, say farther then 10^6 km (1 km = 0.6 mi) from the core of the system.

Emission lines in the spectrum of SS 433 are observed not only in the optical band but also in x-ray spectra, revealing the presence in the jets of heavy elements (such as iron) and highly ionized atoms. Emission lines indicate that: (1) the jets of SS 433 comprise, besides light particles like electrons, atomic nuclei; and (2) the gas that emits lines is thermal (that is, the state of the matter can be described by a macroscopic quantity like the temperature). The process that emits the light observed in the radio band from the jets of SS 433 is, instead, nonthermal and is due to synchrotron emission: Charged particles wrapping around a magnetic field, accelerate and thus emit light. *See* SYNCHROTRON RADIATION.

The model developed and generally accepted to describe the emission of radiation due to the thermal component of the jets is the adiabatic cooling model. The hypothesis of adiabatic cooling means that the gas ejected in the jets does not exchange energy with the environment. The gas has a certain density and temperature at the time of launch at the base of the jet, and both temperature and density decrease with the distance from the binary core only because of the expansion of the gas in a conical shape. The more dense and hot is the gas, the more it emits at shorter wavelengths. Close to the binary core, at distances less than 10^6–10^7 km, the jets' gas is at a temperature of about 10^8 K, and emits light in the x-ray band. Then the gas in the jets cools, and at distances of about 10^{10} km from the core the temperature is of the order of 10^4 K and the gas emits optical radiation. (This is the only really model-independent constraint on temperature and distance, coming from optical spectral analysis.) Based on the adiabatic cooling model, farther out the gas should be too cool to thermally emit optical or x-ray radiation. The nonthermal radio emission superimposed on this thermal radiation is probably caused by the highly relativistic free electrons in the tail of the thermal distribution of the gas.

The adiabatic cooling model breaks down beyond a certain distance from the binary core. In fact, spatially resolved x-ray observations of the jets of SS 433 showed highly ionized Doppler-shifted iron emission lines, emitted in the jets at distances of about 10^{12} km from the binary core, from gas with a temperature of the order of 10^8 K. This evidence indicates that reheating of atoms takes place in the jet flow, which is still moving with relativistic velocity more than 100 days after launch from the binary core (**Fig. 3**).

A possible explanation has been suggested for this reheating. Random variations of about 15% in the velocity of the blobs launched in the jets are inferred from long-term optical monitoring of the source. Therefore, a faster blob could catch up with a slower blob launched 162 days before with the same precession phase; this would produce a shock and a reheating of matter in the jet a few hundred days after its launch from the binary core. *See* SHOCK WAVE.

Further out, at distances of 10^{14}–10^{15} km (10–100 light-years) from the binary core, x-ray observations show bright termination shocks, caused by the matter in the cold jets that hits the dense interstellar medium. This process stops the jets' run, releasing frictional energy in form of light. At those distances the two jets deform, like a seashell, the surrounding W50 nebula, which is thought to be a supernova remnant associated with the formation of the neutron star or black hole in SS 433. *See* INTERSTELLAR MATTER; NEBULA.

Binary system. As already noted, SS 433 is an x-ray binary system, that is, a binary star system in which a neutron star or a black hole is accreting matter in the form of a disk from a companion star. It is also an eclipsing binary system with an orbital period of about 13 days. Astronomers agree that the companion star is an early-type high-mass star (O-type or A-type stars are among the current hypotheses), with a mass that should exceed 5 times the mass of the Sun. However, the proper identification and classification of this star, and consequently (using the so-called mass function) the classification of the compact object as a neutron star or a black hole, is still controversial. In fact, the binary system is imbued with a dense envelope of gas probably coming from a strong stellar wind from the high-mass companion star. This envelope makes difficult any deep investigation of the two stars of the system. *See* BINARY STAR; ECLIPSING VARIABLE STARS.

Equatorial outflow. Radio observations have also shown outflows of matter in a direction perpendicular to the jets' axes, in the disk plane. This equatorial outflow can be associated to a windlike emission from the system (the high-mass companion) or to an extension of the accretion disk. It is not clear

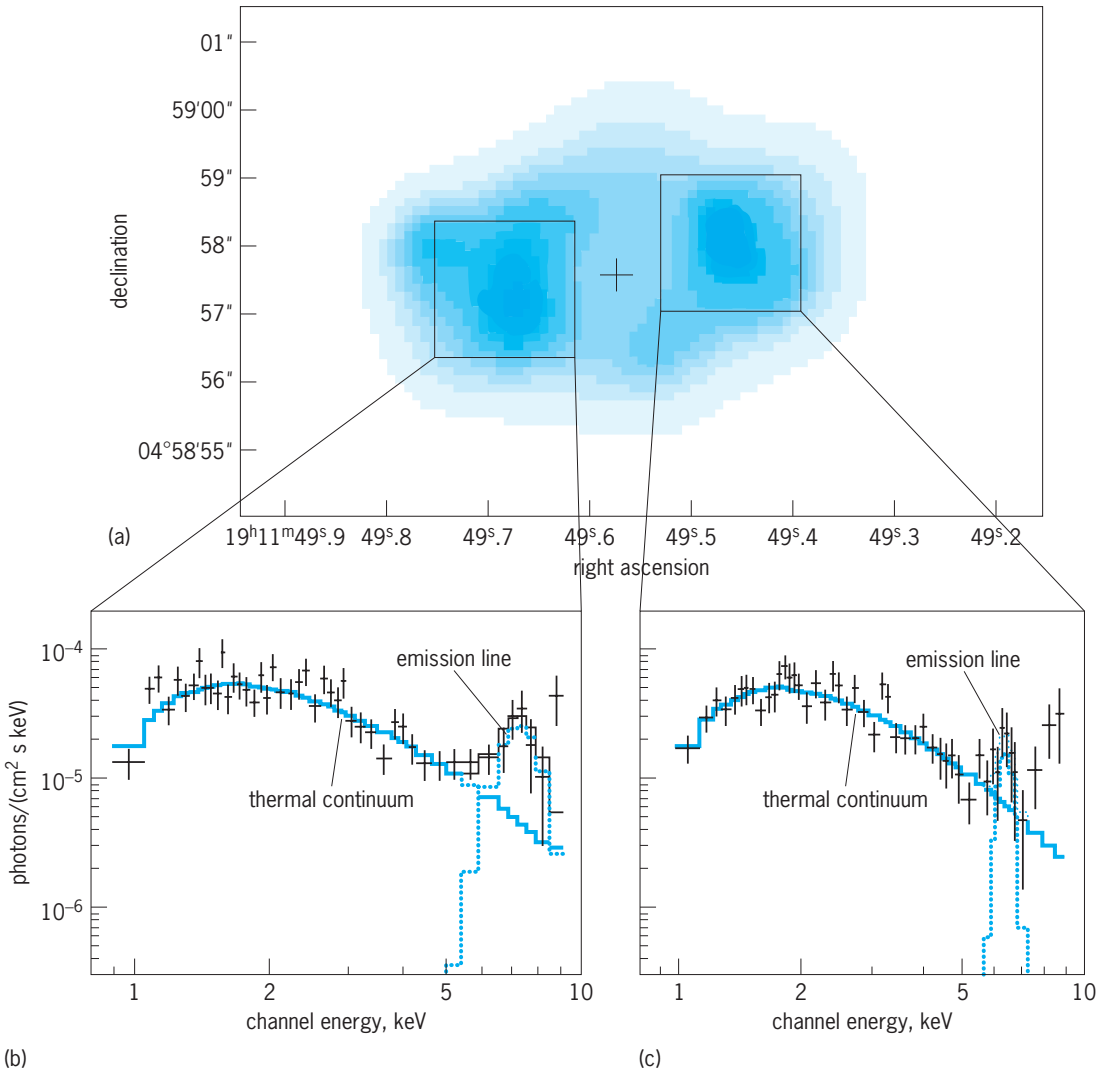

Fig. 3. X-ray observations of SS 433 performed in 2000 with the *Chandra X-ray Observatory*. (*a*) X-ray image. The binary core of SS 433 is indicated with a cross and is not bright due to an instrumental effect. The lobes of the x-ray jets are visible on either side of the core. The squares indicate the regions of the jets from which the energy spectra are extracted. (*b*) Energy spectrum of east lobe. (*c*) Energy spectrum of west lobe. The spectra (indicated by crosses) are consistent with a thermal continuum (indicated by solid color lines) and Doppler-shifted highly ionized emission lines (indicated by color dotted lines).

yet if these outflows consist also of thermal matter, although x-ray observations seem to support this hypothesis.

Based on the principle that the physical processes behind the formation and evolution of relativistic jets are the same regardless the scale of the astrophysical objects involved (either a supermassive black hole in the center of a distant galaxy or a stellar collapsed object in a galactic binary system), the discovery of SS 433 gave astronomers the first near-by laboratory to study in more detail one of the most powerful phenomena of the universe, relativistic jets of matter.

Simone Migliari

Bibliography. S. S. Eikenberry et al., Twenty years of timing SS 433, *Astrophys. J.*, 561:1027–1033, 2001; B. Margon, Observations of SS 433, *Annu. Rev. Astron. Astrophys.*, 22:507–536, 1984; S. Migliari, R. Fender, and M. Méndez, Iron emission lines from extended x-ray jets in SS 433: Reheating of atomic nuclei, *Science*, 297:1673–1676, 2002.

Stability augmentation

The alteration of the inherent behavior of a system. As an example, ships tend to exhibit significant rolling motions at sea. To dampen these rolling motions, a roll stabilization (feedback) system can be used. Such a system consists of a set of vanes (that is, small wings) extending outward from the hull, below the waterline. By varying the vane incidence angle relative to the hull, a hydrodynamic lift is generated on the vane. The vanes are driven by a feedback system so that the rolling motions are opposed by creating positive lift on one side of the hull and negative lift on the other side.

As a second example, nearly all satellites require some form of stability augmentation to help in keeping the antennas or sensors aligned with receiving equipment on Earth. The stability augmentation is effected by thrusters which receive their commands

from a feedback system. *See* SPACECRAFT PROPULSION.

As a third example, stability augmentation systems are used on aircraft while airborne. This is usually achieved by a system which controls one or more flight-control surfaces (or engines) automatically without inputs from the pilot. The inherent stability and response behavior of many modern airplanes tends toward low damping or even instability. The physical reasons have to do with the configuration of the airplane and the combination of flight speed and altitude at which the airplane is operated. Several modern fighters and even some transports are intentionally designed with no or little inherent stability. There are a number of reasons for such a design condition. In the case of fighters, excellent maneuverability in combat is essential. By making a fighter intentionally inherently unstable, it is easy to design the control system so that load factors in pull-ups or in turns can be built up rapidly. In the case of transports, the motivation to design for little or no inherent stability is to lower the size of the tail and thereby achieve a reduction in drag and weight. *See* AIRPLANE; MILITARY AIRCRAFT.

Yaw-damper operation. If, for whatever reason, an airplane is deficient in the damping of its response or is divergent instead of stable, the flying qualities are compromised. This can be corrected by installing a stability augmentation system. Yaw dampers were among the first such systems designed into airplanes. The dutch-roll motion of an airplane consists of simultaneous oscillations of the bank (or roll) angle, the sideslip angle, and the heading angle; when this motion becomes poorly damped, it is difficult for a pilot to maintain manual control over the airplane when flying through turbulence. A yaw damper can usually solve this problem. It normally drives the rudder with the help of a servo (electrohydraulic or electromechanical). The servo receives its signal to move from a computer-amplifier combination. The computer has as its inputs the desired yaw rate and the actual yaw rate. The desired yaw rate of the airplane depends on the flight condition. If the desired condition is straight-line flight, the desired yaw rate will be zero. If the desired flight condition is a turn, the desired yaw rate depends on the steepness of the turn but is not zero. The actual yaw rate of the airplane is sensed by a yaw-rate gyro. The desired yaw rate is normally derived from the automatic pilot computer. The error signal, which is the difference between the two inputs (**Fig. 1**), is gradually driven to zero so that the output rate of the yaw damper equals the input rate. Generically, such systems are called feedback systems. *See* CONTROL SYSTEMS; FLIGHT CHARACTERISTICS; GYROSCOPE; SERVOMECHANISM.

Contrast with flight control. The control exercised by the stability augmentation system contrasts with that exercised by the pilot. The pilot may be connected with the flight-control surface via a direct mechanical link. Alternatively, in many modern airplanes the pilot cockpit control movement is sensed by a position transducer. The output of the position transducer in turn is sent, via a computer-amplifier combination, to a hydraulic actuator, referred to as a servo, which drives the flight-control surface (**Fig. 2**). Command signals which come from the pilot or from the stability augmentation system are sent by wire (fly-by-wire) or by optical conduit (fly-by-light) to the electromagnetic valve. A valve distributes high-pressure hydraulic fluid either to the left or to the right of the piston so that the piston is forced to move. The piston in turn moves the flight-control surface. *See* FLIGHT CONTROLS.

Types of aircraft systems. When an airplane flies through turbulence, an important task is keeping the wings level. This becomes a tedious job for a pilot during long flights. For that reason, even many small airplanes are equipped with an automatic control device (called a wing leveler) which keeps the wings level.

Yaw dampers typically address lack of damping around the yaw axis (or z axis) of an airplane. Many airplanes also require improvement in damping around the pitch axis (or y axis). Some even require extra damping around the roll axis (or x axis). These dampers are referred to as pitch dampers and roll dampers, respectively. Several fighter and transport airplanes now have negative or low static longitudinal stability. Such problems can be addressed by an angle-of-attack or load-factor feedback system. The reliability of these stability augmentation systems and the associated computer hardware and software has improved to the point where only three such systems are required in commercial airplanes.

The early digital feedback systems still used fly-by-wire systems. Such systems are susceptible to electromagnetic interference and therefore must be protected (that is, hardened) against such interference. Because hardened systems are both costly and heavy, more advanced digital feedback systems are fly-by-light systems.

With the introduction of fast in-flight digital computers, it has become possible to equip airplanes

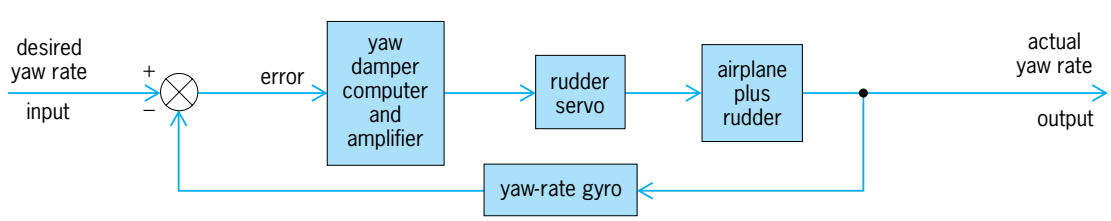

Fig. 1. Block diagram for a yaw damper. The rudder deflects to oppose the error, and the airplane reacts with a different yaw rate.

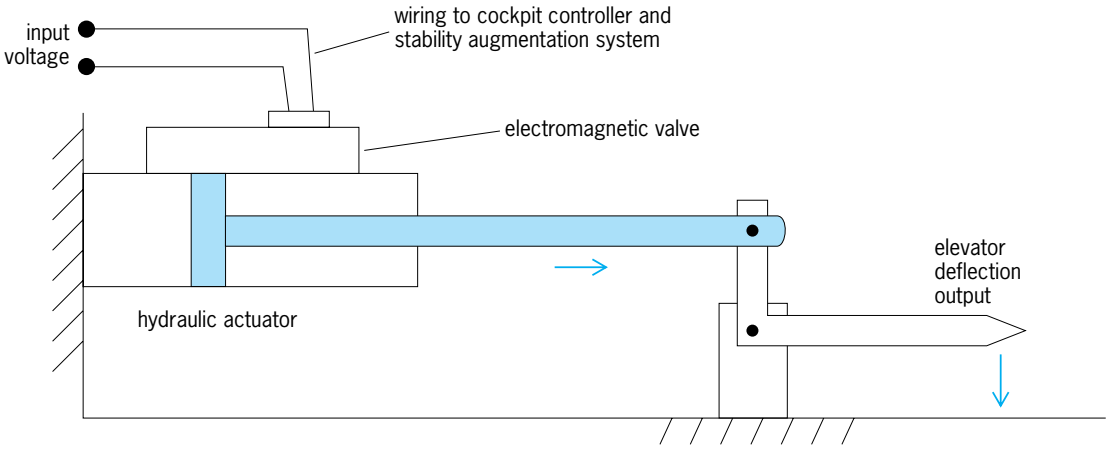

Fig. 2. Electrohydraulically powered elevator system. Hydraulic lines to the actuator are not shown. The actuator is attached to a horizontal stabilizer structure.

with so-called full flight envelop protection systems. Such systems are designed to refuse any pilot input which might get the airplane into a flight condition from which recovery is no longer possible. Such systems can easily be arranged to prevent a pilot from rolling a commercial airplane too much or to prevent the pilot from stalling the airplane. Such systems can also be arranged so that loads acting on the wing or tail do not approach dangerously high levels. In that case the system is referred to as a load-alleviation system.

Flight simulations. A major design problem in all of these types of systems is assurance against unforeseen failures or unforeseen actions by the pilot. Very detailed failure cause-and-effect analyses are carried out during the detailed design and development stages of modern flight-control systems. However, it is not always possible to foresee every possible type of failure or unintentional wrong action by the cockpit crew. For that reason extensive so-called iron-bird simulations, with and without pilots in the loop, are carried out long before the first flight. During these simulations, all possible worst-case scenarios are, in effect, test-flown on the simulator, and the reaction of the system to these events in monitored. If necessary, design changes are introduced and tested. *See* AIRCRAFT TESTING; SIMULATION. Jan Roskam

Bibliography. J. H. Blakelock, *Automatic Control of Aircraft and Missiles*, 2d ed., 1991; D. McRuer, I. Ashkenas, and D. Graham, *Aircraft Dynamics and Automatic Control*, 1974, paper 1990; R. C. Nelson, *Flight Stability and Automatic Control*, 2d ed., 1997; J. Roskam, *Airplane Flight Dynamics and Automatic Control*, parts I and II, 1995.

Stabilizer (aircraft)

The horizontal or vertical aerodynamic wing surfaces that provide aircraft stability and longitudinal balance in flight. Horizontal and vertical stabilizers (fins) are similar to the aircraft wing in structural design and function of providing lift at angle of attack to the wind. However, stabilizers are not required to supply lift to overcome aircraft weight during flight, and when the wing-fuselage center of pressure is behind the aircraft center of gravity, the aerodynamic load on the horizontal stabilizer may be downward.

Stabilizer arrangements. Stabilizers may be swept back like wings for improved high Mach-number characteristics. The horizontal stabilizer may be found in a conventional or tail-last arrangement or in a canard or tail-first arrangement, or it may be dispensed with to give a tailless arrangement. Vertical stabilizers, however, are invariably present at the rear of the aircraft in single or multiple units. Aircraft designed for very high speeds may have auxiliary vertical stabilizers or ventral fins below the wing or fuselage to avoid the large losses in directional stability that occur with conventional arrangements at positive angles of attack at Mach numbers above about 2. Various stabilizer arrangements are shown in the **illustration**.

Stabilizing function. The stabilizing function of the horizontal and vertical stabilizers may be

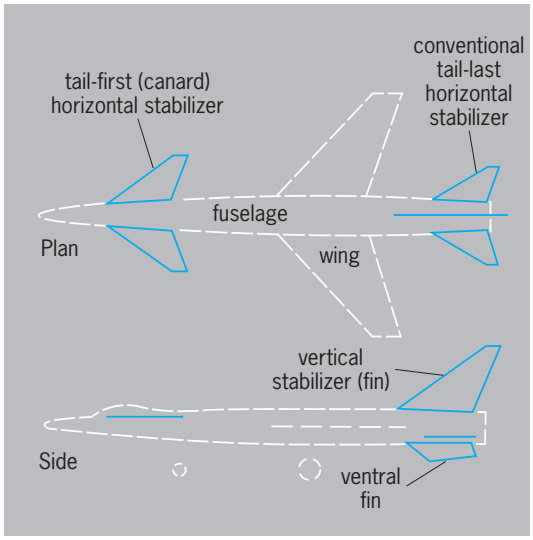

Various stabilizer arrangements.

represented by the equation below,where $\partial C_m/\partial\alpha$

$$\frac{\partial C_m}{\partial \alpha} = -\left(\frac{\partial C_L}{\partial \alpha}\right)\frac{S'}{S}\frac{l'}{l}(1-\epsilon_\alpha)\frac{q'}{q}$$

is the dimensionless rate of change of stabilizing or weathervane moment with angle of attack of sideslip, $\partial C_L/\partial\alpha$ is the dimensionless lift-curve slope of the isolated stabilizer based on stabilizer dimensions, and S'/S and l'/l are the ratios of stabilizer area to wing area, and of distance from the stabilizer center of pressure to the aircraft center of gravity to a characteristic length, respectively. The quantities ϵ_α and q'/q are flow conditions at the stabilizer and are the downwash or sidewash and dynamic pressure, respectively, referred to the free stream.

Control function. On aircraft provided with elevator and rudder control surfaces, the stabilizers may be either fixed to the fuselage or adjustable in incidence at a slow rate (about 1°/s) for trim. In either case, the stabilizers provide the support for the elevator and rudder, carrying their loads to the fuselage. Supersonic aircraft and guided missiles generally dispense with hinged trailing-edge elevator and rudder control surfaces, using the stabilizers for control as well as stabilization. In this case the stabilizers are actuated by the primary flight control system at rates up to about 50°/s. *See* FLIGHT CHARACTERISTICS; FLIGHT CONTROLS.

Malcolm J. Abzug

Bibliography. A. W. Babister, *Aircraft Dynamic Stability and Response*, 1980; B. Etkin and L. D. Reid, *Dynamics of Flight-Stability and Control*, 3d ed., 1995; R. C. Nelson, *Flight Stability and Automatic Control*, 2d ed., 1997.

Stabilizer (chemistry)

Any substance that tends to maintain the physical and chemical properties of a material. Degradation, that is, irreversible changes in chemical composition or structure, is responsible for the premature failure of materials. Stabilizers are used to extend the useful life of materials as well as to maintain their critical properties above the design specifications. Oxygen and water are the principal degradants, but ultraviolet radiation also can have a significant effect (photodegradation).

There are three general approaches to stabilization: surface modification, modification of chemical composition, and use of additives. The relative importance of these approaches varies with the material required to be protected and with the mode of degradation.

Surface modification. Wood and metals are protected against destructive weathering and corrosion by covering with paints or other surface finishes. Primers and rust-inhibiting paints are specialized formulations designed to provide corrosion resistance for metals. Aluminum is stabilized against oxidative degradation by a natural process in which an oxide film forms over the surface which then restricts oxygen penetration into subsurface layers. *See* CORROSION; PAINT.

Laminated structures are also used to minimize degradation. For example, moisture-resistant polymers have been laminated over metals to provide corrosion resistance. In other instances, naturally resistant polymer films have been laminated over the surface of more sensitive polymers to stabilize the latter against photodegradation. Polymers have also been stabilized by crosslinking at the surface to restrict penetration of reactants into the bulk of the material.

The protection of food products by plastic wrappings is an extension of the approach to stabilization by surface modification. Plastic wrappings limit the amount of oxygen that reaches the food, thus inhibiting those oxidative reactions that lead to spoilage. *See* FOOD MANUFACTURING.

Modification of composition. Stainless steels are stabilized steels. Modifying the structure of steel is done by alloying with chromium or nickel corrosive chemical reactions. *See* STAINLESS STEEL.

Both natural and synthetic polymers degrade during processing and long-term exposure. Different polymers exhibit considerable variation in the level of stability to hostile environments, their stability being directly related to the dissociation energies of chemical bonds within the structure. When strong bonds are present, such as the C—F bonds in fluorinated polymers, a greater level of stability is observed in contrast to that of polymers in which the weaker C—H bonds predominate. This accounts for the superior thermal stability of tetrafluoroethylene over that of hydrocarbon polymers, which exhibit significant differences in the strength of their C—H bonds. The weakest C—H bonds in polypropylene occur at the tertiary carbon atoms where methyl groups are attached. Since one such bond is present in each repeating unit, polypropylene is much less stable to oxidative degradation than is linear polyethylene. The stability of hydrocarbon polymers could be modified by controlling the extent of branching. However, an improvement in stability by this approach will be accompanied by changes in other potentially important properties, for example, modulus.

There is evidence that some synthetic polymers have structural irregularities that could serve as sites for degradation to begin. Reduction of such irregularities during synthesis or replacement with less labile groups could also be used as an approach to stabilization. *See* POLYMER.

Additives. A wide variety of additives has been developed to stabilize polymers against degradation. Stabilizers are available that inhibit thermal oxidation, burning, photodegradation, and ozone deterioration of elastomers. Research in the chemistry of the low-temperature oxidation of natural rubber has revealed that hydrocarbon polymers oxidize by a free-radical chain mechanism. In pure, low-molecular-weight hydrocarbons, an added initiator is required to produce the first radicals. In contrast, initiation of polymer oxidation occurs in the complex molecules of elastomers through impurities already present, for example, hydroperoxides. *See* CHAIN REACTION (CHEMISTRY); FREE RADICAL.

Stabilization of hydrocarbon polymers can be accomplished with preventative or chain-breaking antioxidants. Preventative antioxidants stabilize by reducing the number of radicals formed in the initiation stage. Where hydroperoxides are responsible for initiation, the induced decomposition of these reactive intermediates into nonradical products provides effective stabilization. Suppressing the catalytic effects of metallic impurities that increase the rate of radical formation can also provide stabilization. Chain-breaking antioxidants interrupt the oxidative chain by providing labile hydrogens to compete with the polymer in reaction with the propagating radicals. The by-product of this reaction is a radical which is not capable of continuing the oxidative chain.

Stabilization of polymers against photooxidation, the principal component of outdoor weathering, is accomplished by addition of light screens, ultraviolet absorbers, or radical scavengers. Carbon black is the most effective of all light screens, functioning by preventing the penetration of ultraviolet radiation into the polymer. Correctly applied, this pigment can extend the outdoor life of hydrocarbon polymers from 1–2 years to over 50 years. Ultraviolet absorbers, which absorb and harmlessly dissipate damaging radiation, are not as effective as carbon black. However, another class of additives, known as hindered-amine light stabilizers, has become available. These stabilizers are believed to function by scavenging destructive radicals, and they may approximate the protection that is provided by carbon black. *See* ANTIOXIDANT; CARBON BLACK; INHIBITOR (CHEMISTRY); PHOTODEGRADATION. W. Lincoln Hawkins

Bibliography. N. S. Allen and M. Edge, *Fundamentals of Polymer Degradation and Stabilisation*, 1993; V. A. Goldade et al., *Plastics for Corrosion Inhibition*, 2005; A. Hamidi, *Handbook of Polymer Degradation*, 2d ed., 2000; P. Marcus (ed.), *Corrosion Mechanisms in Theory and Practice*, 2d ed., 2002; Y. A. Shlyapnikov, S. G. Kiryushkin, and A. P. Mar'in, *Antioxidative Stabilization Of Polymers*, 1996; M. P. Stevens, *Polymer Chemistry: An Introduction*, 3d ed., 1998.

Stain (microbiology)

Any colored, organic compound, usually called dye, used to stain tissues, cells, cell components, or cell contents. The dye may be natural or synthetic. The object stained is called the substrate. The small size and transparency of microorganisms make them difficult to see even with the aid of a high-power microscope. Staining facilitates the observation of a substrate by introducing differences in optical density or in light absorption between the substrate and its surround or between different parts of the same substrate. In electron microscopy, and sometimes in light microscopy (as in the silver impregnation technique of staining flagella or capsules), staining is accomplished by depositing on the substrate ultraphotoscopic particles of a metal such as chromium or gold (the so-called shadowing process); or staining is done by treating the substrate with solutions of metallic compounds such as uranyl acetate or phosphotungstic acid. This type of staining is not dealt with in this article. *See* ELECTRON MICROSCOPE.

Stain classification. Stains may be classified according to their molecular structure. For example, there are the triphenylmethane dyes, such as the fuchsins and the methyl violets; the oxazine dyes, such as Nile blue; and the thiazine dyes, such as thionine and methylene blue.

The stains may also be classified according to their chemical behavior into acid, basic, neutral, and indifferent. This classification is of more practical value to the biologist. An acid dye, such as Congo red and eosine, is usually the sodium or potassium salt of a dye acid; a basic dye, such as methylene blue and Nile blue, is usually the chloride or sulfate of a dye base; a neutral dye, such as the eosinate of methylene blue, is a complex salt of a dye acid with a dye base; and an indifferent dye, such as Sudan III, is one with little chemical activity. At ordinary pH values, cells stain with basic dyes; at low pH values, they stain with acid dyes.

Staining procedures. The staining of microorganisms usually begins by making a smear. When it is desired to preserve intercellular relations, the smear may be replaced by a microculture. The smear or its substitute is usually fixed by heat and treated with a solution of some suitable biological stain as in simple stains or, in more elaborate procedures, it may be chemically fixed, treated with mordants, stained, partially decolorized, and counterstained. In certain procedures parts of a substrate are chemically removed prior to staining.

Smear preparation. The material containing the microorganisms is smeared or spread on the surface of a glass slide. Liquid material may be spread by means of a freshly flamed and cooled loop of platinum or some suitable alloy. Solid material may be smeared directly, or it may be used to make a suspension in water or some liquid medium, a loopful of which is spread on the slide. The film produced is allowed to dry in the air and is fixed by passing the slide, film side up, three or four times over the flame of a Bunsen burner. Some cytological procedures require chemical fixation. After cooling, the film is stained and examined.

Simple stains. These are procedures in which the smear, or the substrate, is stained with a dye solution for a given length of time, washed with water, air-dried, and examined. The most widely used solutions for simple staining are (1) Löffler's methylene blue, a solution of 0.3 g of methylene blue in 30 ml of 95% ethanol, mixed with 100 ml of 0.01% potassium hydroxide; (2) Ziehl-Neelsen's carbol fuchsin, a solution of 0.3 g basic fuchsin in 10 ml of 95% ethanol, mixed with a solution of 5 g phenol in 95 ml of distilled water; and (3) Hucker's crystal violet, prepared by mixing a solution of 0.2 g of crystal violet in 20 ml of 95% ethanol with a solution of 0.8 g ammonium oxalate in 80 ml of distilled water.

Complex staining procedures. Among complex staining procedures are the flagella and capsule stains. Both include the use of a special mordant to increase the density of their substrates and their affinity for the dye. This also increases the thickness of the flagella to a range that is resolved with the light microscope. There are many procedures for demonstrating both structures. In addition, most flagella stains recommend precautions to prevent destruction of these delicate structures while making the smear.

Nuclear stains. Most nuclear stains are based on the staining properties of nucleic acids. In many organisms, particularly among the bacteria, the cytoplasm is rich in ribonucleic acid. As a result, the entire cell, in young and mature cultures, tends to stain like a nucleus. Consequently, most nuclear stains include a step in which the ribonucleic acid is removed, or the bacteria are grown in media deficient in nitrogen supply to hinder the synthesis of ribonucleic acid.

Ribonucleic acid may be removed by treatment with the enzyme, ribonuclease, or a strong mineral acid. This is the basis of the widely used Feulgen reaction and of the HCl-Giemsa stain. In both methods the fixed smear is placed in normal HCl, usually at 60°C for about 10 min. In the Feulgen procedure this is followed by placing the smear for several hours in a solution of basic fuchsin that has been decolorized with sodium or potassium metabisulfite. The nuclei appear purple to violet in a colorless cell.

In the HCl-Giemsa procedure, the smear is stained with dilute Giemsa's stain. Stock solutions of this stain contain chiefly the eosinates of methylene azure and of methylene blue dissolved in a mixture of methyl alcohol and glycerol. It is a valuable cytological stain because of its metachromatism; it stains chromatin purple to red and the cytoplasm blue.

Differential stains. These are staining procedures that bring out color differences between a substrate and its background or between different parts of the same substrate. This may be accomplished by simple staining with a metachromatic dye; Nile blue, for example, stains the neutral fat droplets red and the rest of the cell blue. Or it may require controlled decolorization or differentiation with acids, alkalies, neutral solvents, or some other agents. Examples of the latter procedures are Gram's stain and acid-fast stain, and the spore stain.

Spore stains. These are based on the fact that the spore does not take up dyes readily, but once stained it resists decolorization. The differentiating agent used may be a dilute solution of an organic acid, an acid dye, or another basic dye.

Macchiavello stain. Another differential stain is that of Macchiavello for demonstrating rickettsiae. The fixed tissue smear is stained with basic fuchsin, differentiated with a 0.5% solution of citric acid, and counterstained with a 1% solution of methylene blue. Rickettsiae stain red, and tissue cells stain blue.

Negative stain. Another differential stain is the negative stain. In this procedure the cells appear colorless against a colored background. The cells are suspended in a droplet of the negative stain, such as India ink, water-soluble nigrosine, or Congo red in 1% aqueous solutions, on a glass slide. The droplet is spread out into a thin film and allowed to dry in the air. When pathogenic organisms are to be examined, the preparation is dipped in a mixture of 1 ml concentrated hydrochloric acid (HCl) and 100 ml of 95% ethanol. The acid alcohol changes Congo red preparations to blue, giving better contrast. Georges Knaysi

Fluorochrome stain. The fluorochrome staining procedure is used for the early and rapid detection of disease-producing *Mycobacterium tuberculosis* and *M. leprae*, together with other species of *Mycobacteria* and *Nocardia asteroides*. The method has been used worldwide since the mid-1970s. Widely recognized as the truant auramine-rhodamine (A—R) acid-fast staining procedure, it not only permits the detection of tubercle bacilli more quickly in clinical specimens, such as sputa, urines, and tissues, but also gives a higher yield of positive findings than the conventional Ziehl-Neelson acid-fast procedure, which employs carbol fuchsin as the primary dye. The truant auramine-rhodamine procedure has been especially popular and relatively inexpensive because of the improvement in the development of fluorescent microscopes. The auramine and rhodamine dyes appear to have structures better suited for penetrating and stacking along the receptor macromolecules of the mycobacterial peptidolipid or peptidoglycolipid than does carbol fuchsin. The intracellular concentration possible with auramine and rhodamine may be greater than with carbol fuchsin, which may result in the superior staining qualities. The interpretation is that the cell wall lipids of mycobacterial cells in their native state containing mycolic acid react with the two fluorescing dye molecules of auramine and rhodamine and therefore absorb the stain. The dyes may well confer the property of acid fastness because they also function as a permeability barrier and impede the penetration of the bleaching agent or decolorizer. J. P. Truant

Bibliography. M. E. Boon and J. S. Drijver, *Routine Cytological Staining Techniques: Background and Practice*, 1986; G. Clark, *Staining Procedures*, 4th ed., 1980; P. R. Lewis and D. P. Knight, *Staining Methods for Sectioned Material*, 1977; P. Murray (ed.), *Manual of Clinical Microbiology*, 7th ed., 1999.

Stainless steel

The generic name commonly used for that entire group of iron-base alloys which exhibit phenomenal resistance to rusting and corrosion because of chromium content. The metallic element chromium (Cr) has been used in small amounts to strengthen steel since the famed Eads Bridge spanned the Mississippi at St. Louis, Missouri, in 1872; but only in the present century was it discovered that contents with chromium exceeding 10%, with carbon (C) held suitably low, make iron effectively rustproof. Onset

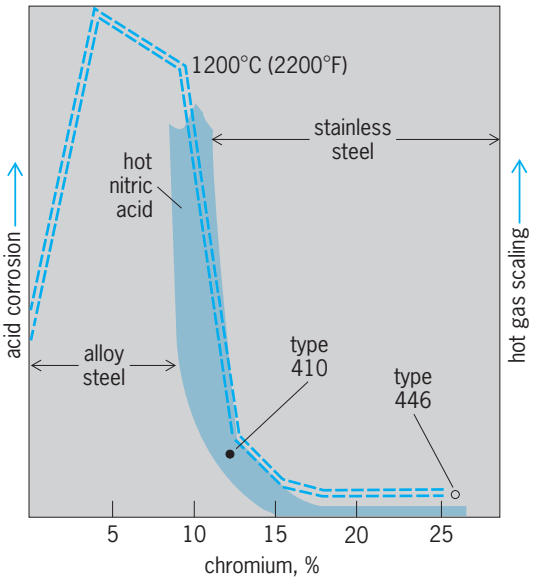

Fig. 1. Passivity phenomenon in stainless steel. 1200°C = 2200°F.

of the property is striking (**Fig. 1**), and the chromium can then be increased to about 27% with even further chemical advantages before the metal becomes structurally useless. *See* ALLOY; STEEL.

Other alloy elements, notably nickel (Ni) and molybdenum (Mo), can also be added to the basic stainless composition to produce both variety and improvement of properties. Over 100 different stainless steels are produced commercially, about half as standardized grades. Some are more properly classed as stainless irons since they do not harden as steel; others are true steels to which corrosion resistance becomes an added feature. Still others that are neither properly steels nor irons introduce totally new classes of materials, from both mechanical and chemical standpoints.

In fact, the great modern developments of heat-resisting superalloys and refractory metals stand as extreme extensions of the two heat-resisting families of stainless steels. That same resistance to rusting at ordinary temperatures shows as resistance to oxidation at elevated temperatures (Fig. 1), producing a family of oxidation-resistant high-temperature alloys; and with this basic feature contributed by the chromium content, further additions of nickel and other elements contribute remarkable properties of workability and strength, which extend not only to high temperatures but to cryogenic temperatures as well. The nickel-base heat-resisting alloys, the most prominent in that field, project from stainless steel in the sense that nickel increasingly replaces the iron (Fe). Similarly, the important titanium (Ti) and cobalt (Co) alloy systems, such refractory metals as niobium (Nb) and tungsten (W), and also chromium and molybdenum, represent the far extremes for use of these alloy elements in stainless steel.

The shaded band in Fig. 1 covers typical values for corrosion penetration of iron alloys with increasing chromium content, the actual data referring to hot nitric acid. The width of the band is due to differences in temperature and concentration. Attack almost ceases near 12% chromium, the alloy type 410; and it becomes scarcely measurable as chromium increases to type 446. The double-dash boundary in Fig. 1 is a plot of actual data for hot-gas scaling at 2190°F (1200°C). The sudden drop occurs in nearly the same chromium range as for aqueous corrosion, type 446 again being the most resistant alloy. The reversal at the far left is also similar for aqueous corrosion.

In service, the stainless steels cover three major fields: corrosion resistance at ordinary temperatures, scaling resistance at high temperatures, and mechanical strength over the entire temperature range of metallurgical engineering.

Types. Metallurgically, the alloys fall within four groups: ferritic, martinsitic, austenitic, and precipitation-hardening.

Ferritic. When chromium is added to iron to develop corrosion resistance without steel-type hardening, ferritic alloys are obtained. *See* IRON ALLOYS.

The typical alloys in the table are standards listed by the American Iron and Steel Institute (AISI). The first is the least expensive, and a little aluminum (Al) is usually added to help the chromium suppress steel-type hardening and ensure the ferritic condition, since this avoids welding problems from stray martensite. The second is the most popular commercial ferritic grade, widely used for automotive trim; and some fractional martensitic hardening is deliberately courted for additional strength. Type 446 is primarily used for oxidation resistance, for above this chromium content the engineering value of the chromium-iron alloys rapidly vanishes.

Martensitic. When chromium is added to steels with carefully chosen carbon contents, the martensitic alloys result (see **table**). Again, these cover a range from the least expensive and most widely used type

Typical compositions of standard stainless steels in percent

AISI no.	Cr	C
Ferritic		
405	13	0.08
430	17	0.12
446	26	0.20
Martensitic		
410	12	0.12
420	13	0.25
440-C	17	1.00

AISI no.	Cr	Ni	Other elements
Austenitic			
302	18	8	C = 0.08
310	25	20	C = 0.20
316	16	13	Mo = 2.75
347	18	11	Nb = 0.50
Precipitation-hardening			
630 (17-4 PH)	17	4	Cu = 3.2, Nb = 0.25
633 (AM-350)	16.5	4.3	Mo = 2.75, N = 0.10
(174-4 Cu Mo)	17	14	Cu = 3.0, Mo = 2.5 Nb = 0.50, Ti = 0.25

410, through the more selected type 420, ideal for cutlery, and on to the highly specialized type 440-C, whose great hardness makes it suitable for ball bearings and surgical instruments.

Austenitic. Nickel becomes the next most important alloying element. Its commonest use is in the renowned "18–8" containing 18% Cr and 8% Ni. Unlike the atomic structure of ferrite, which has the form of a neatly body-centered cube, or the martensite with its atomic disarrangement responsible for strength and hardness, the austenite of this third class of stainless steel has a face-centered cubic form made stable by the nickel. Completely new mechanical properties appear, outstanding for workability, notch toughness, high-temperature strength, and a frequently superior corrosion resistance. These are generally the most prized and popular of all stainless steels and also the most expensive. Only a few of the score of standards appear in the table. The parenthetical C for type 302 refers to a strange sensitivity to carbon content, such that 18–8 today is more commonly type 304 with carbon held below 0.08, and even type 304 with carbon held below 0.03. Type 310 lies well toward the nickel-base alloys and is accordingly notable for its heat resistance. The molybdenum in type 316 confers a special resistance to pitting corrosion, while the niobium in type 347 is a highly specialized addition correcting the carbon problem in welding 18–8.

While these alloys are called steels, no martensite develops on cooling; neither can the alloys be hardened by the usual quenching and tempering. Contrarily, they are put into their softest state by producing austenite at high temperature, with subsequent quenching to retain it. Strengthening is developed instead by cold working, and by adjusting the 18–8 composition slightly downward toward 17–7. The austenite then tends to decompose inside the cold steel; the martensite crowds the slip planes loosened by cold working, locking them against further movement and thereby increasing the strength (**Fig. 2**). Hence, ferritic and martensitic steels are little strengthened by cold working, while austenitic steels are greatly strengthened by it, notably type 301.

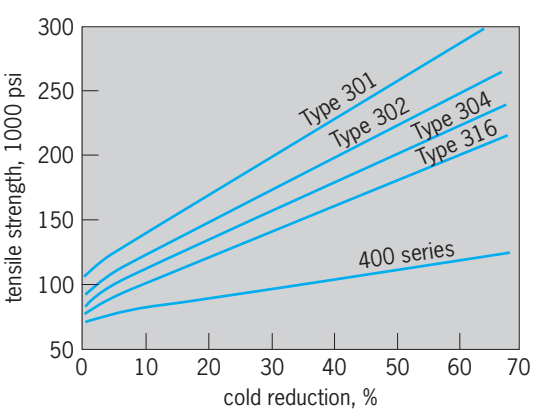

Fig. 2. Graph of the effect of cold-work on tensile strength of stainless steels. 1000 psi = 6.9 MPa.

Precipitation hardening. Exploiting this approach to strengthening, and following the example of duralumin in a neighboring metallurgical field, metallurgists have developed the fourth group of stainless steels known as precipitation hardening. A score of alloys became available within two decades of their commercial acceptance, and trade names rather than generic designations are still commonly used. The first listed in the table is subclassed martensitic because the metal behaves primarily as martensite which further experiences internal precipitation; the second is typical of what are called semiaustenitic alloys for the reason that they rather closely follow the instability pattern of 17–7; and the third belongs to the austenitic subgroup, with hardening essentially due to precipitation within austenite. The special properties are developed through carefully controlled heat treatments that always begin with a "solution anneal" at some high temperature and are followed by internal precipitation at some lower temperature. The latter may have multiple stages, ranging down to subzero treatments, and effects of cold work are often also exploited, particularly for the semiaustenitic alloys. These precipitation-hardening stainless steels are the hallmark since the third quarter of the twentieth century, bringing mechanical properties to levels unmatched by any other metal or alloy. Tensile strengths are actually exceeding the phenomenal figure of 300,000 psi (2.1 gigapascals).

Fabrication. As with steels in general, the so-called wrought stainless steels come from the melting furnaces in the form of either ingot or continuously cast slabs. Ingots require a roughing or primary hot working, which the other form commonly bypasses. All then go through fabricating and finishing operations such as welding, hot and cold forming, rolling, machining, spinning, and polishing. No stainless steel is excluded from any of the common industrial processes because of its special properties; yet all stainless steels require attention to certain modifications of technique.

Hot working. Hot working is influenced by the fact that many of the stainless steels are heat-resisting alloys. They are stronger at elevated temperatures than ordinary steel. Therefore they require greater roll and forge pressure, and perhaps less reductions per pass or per blow. The austenitic steels are particularly heat-resistant.

Welding. Welding is influenced by another aspect of high-temperature resistance of these metals—the resistance to scaling. Oxidation during service at high temperatures does not become catastrophic with stainless steel because the steel immediately forms a hard and protective scale. But this in turn means that welding must be conducted under conditions which protect the metal from such reactions with the environment. This can be done with specially prepared coatings on electrodes, under cover of fluxes, or in vacuum; the first two techniques are particularly prominent. Inert-gas shielding also characterizes widely used processes, among which at least a score are now numbered. As for weld cracking, care

must be taken to prevent hydrogen absorption in the martensitic grades and martensite in the ferritic grades, whereas a small proportion of ferrite is almost a necessity in the austenitic grades. Metallurgical "phase balance" is an important aspect of welding the stainless steels because of these complications from a two-phase structure. Thus a minor austenite fraction in ferritic stainless can cause martensitic cracking, while a minor ferrite fraction in austenite can prevent hot cracking. However, the most dangerous aspect of welding austenitic stainless steel is the potential "sensitization" affecting subsequent corrosion, to be discussed later. *See* WELDING AND CUTTING OF MATERIALS.

Machining and forming. These processes adapt to all grades, with these major precautions: First, the stainless steels are generally stronger and tougher than carbon steel, such that more power and rigidity are needed in tooling. Second, the powerful work-hardening effect (Fig. 2) gives the austenitic grades the property of being instantaneously strengthened upon the first touch of the tool or pass of the roll. Machine tools must therefore bite surely and securely, with care taken not to "ride" the piece. Difficult forming operations warrant careful attention to variations in grade that are available, also in heat treatment, for accomplishing end purposes without unnecessary work problems. Spinning, for example, has a type 305 modification of 18–8 which greatly favors the operation; while at the opposite extreme is type 301 or 17–7, for those who wish to take advantage of the strengthening due to cold working. *See* MACHINABILITY OF METALS.

Finishing. These operations produce their best effects with stainless steels. No metal takes a more beautiful polish, and none holds it so long or so well. Stainlessness is not just skin-deep, but body-through. And, of course, coatings are rendered entirely unnecessary. At this point the higher initial cost of these alloys rapidly falls behind the long-range cost of other metals, often making them cheaper to buy on the market.

Service. Service of the stainless steels covers almost every facet of modern metallurgical engineering. They began in the chemical industry for reasons of their corrosion resistance, then graduated into high-temperature service for reasons of resistance to both surfacial sealing and internal creep. They have finally proliferated in every field of mechanical operation, ranging from high strength and hardness in combination with good ductility, at elevated as well as cryogenic temperatures, through cutlery and spring and nonmagnetic materials, to the purely ornamental applications of modern architecture.

However, when stainless steel does break down, the action is usually dramatic, and nothing exceeds the phenomenology of corrosion in the chemical industry. The situation can be understood by recognizing that the passivity of the stainless steels—that which prevents or restrains their reaction with environmental chemicals—is not the nonactivity or nobleness which characterizes the noble metals of the platinum group. It comes as a surprise to the nonscientists that the chromium content makes stainless steel even more disposed to rust than plain iron. But it is the very tendency which immediately fixes the oxygen or rust so tightly upon the surface that a protective coating results. The same quality applies to titanium and aluminum and their alloys. If something disturbs the quite special conditions needed for this passivity, corrosion takes place. The damage then usually takes one of the three major forms illustrated in **Fig. 3**.

Sensitization. Some phenomenon, not yet fully understood, causes chromium carbide to precipitate in the grain boundaries of the austenitic grades when heated even for short periods of time in the range 800–1500°F (427–816°C); these grain boundaries thereupon become so susceptible to chemical attack that a piece of formerly solid steel can fall apart in the hand like granulated sugar. The most workable explanation is that the carbon has depleted the chromium content in forming the carbide, thus leaving the metal locally nonstainless. Because welding necessarily produces a heat zone alongside the bead which somewhere must cover this critical range, welded steels often are susceptible to sensitization when later subjected to chemical service. The principal cures thus far are as follows: (1) The carbon is lowered to values where carbide cannot form; this has been accomplished with the introduction of type 304L. (2) Other and even more powerful carbide formers—notably titanium (Ti) and niobium (Nb)—have been added to the alloys to tie up carbon at high temperatures before chromium carbide forms. Nevertheless, the problem persists as the principal concern in the chemical service of austenitic stainless steel.

Stress-corrosion cracking. Known for a hundred or more years in the brass industry as season cracking, the same phenomenon seems to have been discovered in stainless steel. In marked contrast to the intergranular nature of sensitization, stress-corrosion cracking cuts right through the body of the grain; and although the austenitic steels are by far the most susceptible, with the martensitic and ferritic following in that order, it appears that any metal or alloy can probably be made to fail in this manner if subjected to the right chemical under appropriate conditions of stress. Present corrections or preventions lie principally in directions of reducing stress levels, both internal and external, and avoiding exposure to chemicals known to be particularly dangerous. A close approach to immunity is indicated through the use of highly purifying melting processes in steelmaking.

Pit and crevice corrosion. These are forms specifically following from oxygen "starvation" in local areas. When the steel can no longer rust, the passive condition converts to an active one; and to worsen the situation, all of the remaining surface then takes on the role of cathode in a galvanic cell which concentrates all the anodic or dissolving activity right at the tiny point of breakdown. If this is just a local spot, the attack takes the form of pitting, but if two flat

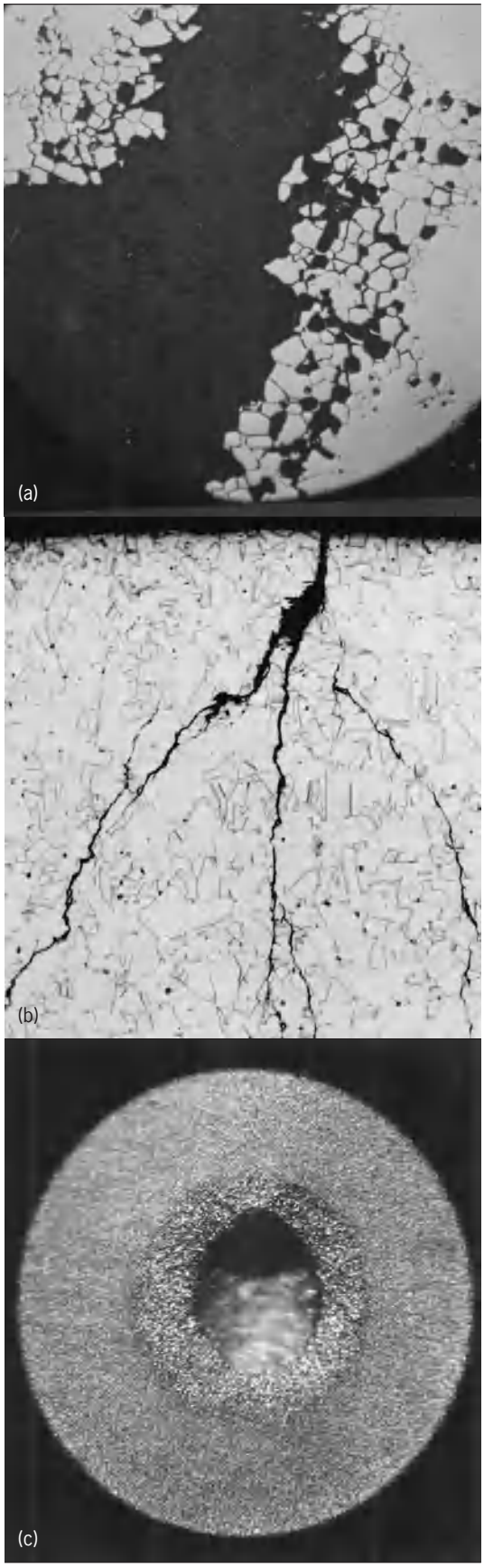

Fig. 3. Three great corrosion effects on stainless steel. (*a*) Sensitization. (*b*) Stress-corrosion cracking. (*c*) Pit corrosion.

surfaces come close together as in bolt-and-washer assemblies, the thin film of entrapped liquid can run out of oxygen and cause a general breakdown called crevice corrosion. *See* CORROSION.

Production. Stainless steel has undergone revolutionary changes in its melting technology. There are two major problems in the melting of stainless steel which distinguish it markedly from melting ordinary steel: chromium must be brought to high levels, and carbon levels must be held low. With minor exceptions in the high-carbon martensitic grades, stainless cannot be produced in ordinary fuel-fired furnaces such as the open hearth. The technology arose solely because of the closer controls possible with electrically heated furnaces, specifically the arc furnace where tonnage production is concerned. Thus stainless melting arose historically out of ordinary steelmaking by graduating into specialty furnaces. Following World War II the technological demands upon stainless steel pressed the technology into still greater specialization, such that steelmaking in this field displayed the most brilliant technical achievements in the history of metallurgy.

Ordinary steelmaking is concerned only with removal of carbon from the liquid metal, and this is easily accomplished by the addition of iron ore—the carbon seizes the oxygen and escapes as a gas. But chromium also reacts strongly with oxygen, thereby removing itself from the metal to join the slag. Therefore, this branch of steelmaking had to develop a two-stage process: melt down ordinary steel and drive off the carbon, and add the chromium after the oxidizing of carbon was completed. This in turn brought two severe handicaps: scrap stainless steel—the cheapest chromium carrier—could not be used in the first-stage charge because the chromium would be lost; and an expensive chromium-rich alloy with low carbon content would have to be added in the second stage—so-called low-carbon ferrochrome. *See* STEEL MANUFACTURE.

In the 1930s Alexander Feild developed the rustless process in Baltimore, based upon the discovery that the competition between chromium and carbon for oxygen shifts strongly in favor of carbon at very high temperatures. By using chromium oxide itself for the hearth lining and then carefully pressing the arc heating some hundreds of degrees above normal operating temperatures, Feild found that the first-stage charge could be loaded not only with scrap but also with high-carbon ferrochrome, which is much cheaper than the low-carbon grade. Minimal chromium losses to the slag were then "kicked back" with a shower of ferrosilicon, which easily took the oxygen away from the chromium and left silicates in the slag, while returning metallic chromium to the bath.

In the 1940s an abundance of liquid oxygen became available from developments in war industries, and E. J. Chelius developed the idea of the "oxygen lance." A long, insulated steel pipe was plunged beneath the surface of the molten steel, churning the metal with pure oxygen. The difference between iron oxide and gaseous oxygen, with respect to

oxidizing a steel bath, is that the latter is a gas, and so is carbon oxide. The bubbling of the one aids the escape of the other, while to chromium oxide it makes no difference. Furthermore, the exceedingly high temperatures developed locally around the lance brought the advantage of the Feild effect in removing carbon while conserving chromium. The 1950s then witnessed a virtual conversion of the entire industry to the oxygen lance, spreading into the open hearth as well as the furnaces used for making stainless steel. At last all stainless could be made with high scrap additions in the charge, right up to 100%. As for carbon contamination and the use of high-carbon ferrochrome, the higher the first-stage carbon the better. It served as a fuel, and a convenient source of heat during lancing.

Scarcely had these achievements come into widespread use when successful development of a porous refractory plug permitted blowing a cleansing gas through the bottom of a ladle. At first this technique involved only the use of an inert gas, such as argon, for removing damaging quantities of hydrogen often dissolved in steels of all kinds. But shortly the method evolved to the use of oxygen as a substitute for lancing. However, the violence of the reaction, from entry so deep in the liquid, made it necessary to dilute the oxygen with an inert gas. Again argon was chosen, and the argon-oxygen process was born. New achievements were electrifying: (1) Almost the entire electric furnace processing was dispensed with, the furnace being needed only to melt down the charge, with all further work done in the ladle. The expensive furnace time was cut by factors of three and four, or more. (2) The vastly greater facilities for quality control in a ladle, as compared to a furnace, shortly permitted the further bridging of this technology with the long-developed and highly sophisticated vacuum techniques of physics and laboratory furnaces. The argon-oxygen process then developed variations using vacuum chambers and reduced pressures instead of positive pressures, and in some cases substituting nitrogen or other gases for argon.

Concurrently two major developments entered the field, this time from directions of those highly specialized heat-resistant metals and alloys known as the superalloys and the refractory metals. The superalloys generally comprise systems based upon nickel, cobalt, or titanium, as well as highly complex stainless steels, while the term "refractory metals" typically refers to wolfram (tungsten), tantalum, molybdenum, and niobium (columbium). Frequently the superalloys have quality demands so high as to require remelting a cast form in order to purify it a second time. One process arose naturally out of the vacuum technology just mentioned, and the more critical grades of stainless steel are now often subjected to vacuum-arc remelting. The original melt is cast in the form of a huge electrode which, as with a consumable welding electrode, melts off at the tip and falls into a water-cooled mold in a vacuum chamber. The metal is thus purified, and its solidification structure is very favorable to good mechanical properties. The second process also arose out of a welding concept, namely the use of a slag cover to protect the metal from environmental reactions while molten. First conceived by R. E. Hopkins in America several decades ago, the process was developed in the Soviet Union, then captured worldwide interest in the 1960s. Again the metal in electrode form is remelted electrically, but under a carefully prepared slag blanket instead of a vacuum. Solidification is again well controlled, and the metal is highly purified. *See* HIGH-TEMPERATURE MATERIALS.

Technological advances continued at a rapid rate, with such sophisticated developments as melting with electron beams, plasma arcs, and lasers—even levitation melting, where crucibles or containers are no longer necessary. The metal just suspends itself in mid-air! While levitation melting remains in curiosity stages and lasers still serve principally for punching holes and welding, the plasma-arc technique did enter commercial usage in welding, with beginnings in steelmaking. The electron-beam furnace has led to remarkable accomplishments in producing commercial quantities of certain ferritic stainless steels whose properties were previously believed impossible.

Most stainless is cast as ingots, although continuous casting is rapidly becoming popular. Virtually all wrought grades have their casting counterparts in the foundry industry, separately standardized by the Alloy Casting Institute (ACI). The metallurgy of the cast alloys is essentially the same as that for the wrought, with the principal exception that certain liberties can be taken in alloy composition because the castings are not hot-worked. In designing casting alloys, one can forget about mechanical problems of forming the part and concentrate on the chemical challenges of corrosion resistance and heat resistance. *See* METAL CASTING.

Carl A. Zapffe

Bibliography. American Society for Metals, *Metals Handbook*, *Properties and Selection: Irons and High Performance Alloys*, 10th ed., vol. 1, 1990; J. R. Davis (ed.), *Stainless Steels*, 1994; R. A. Lula, *Stainless Steel*, 1985; F. B. Pickering (ed.), *The Metallurgical Evolution of Stainless Steels*, 1979; L. L. Shreir (ed.), *Corrosion*, vols. 1 and 2, 3d ed., 1994.

Stalactites and stalagmites

Stalactites, stalagmites, dripstone, and flowstone are travertine deposits in limestone caverns (see **illus.**), formed by the evaporation of waters bearing calcium carbonate. Stalactites grow down from the roofs of caves and tend to be long and thin, with hollow cores. The water moves down the core and precipitates at the bottom, slowly extending the length while keeping the core open for more water to move down. Stalactites are banded concentrically to the center.

Stalagmites grow from the floor up and are commonly found beneath stalactites; they are formed from the evaporation of the same drip of water that forms the stalactite. Stalagmites are thicker and

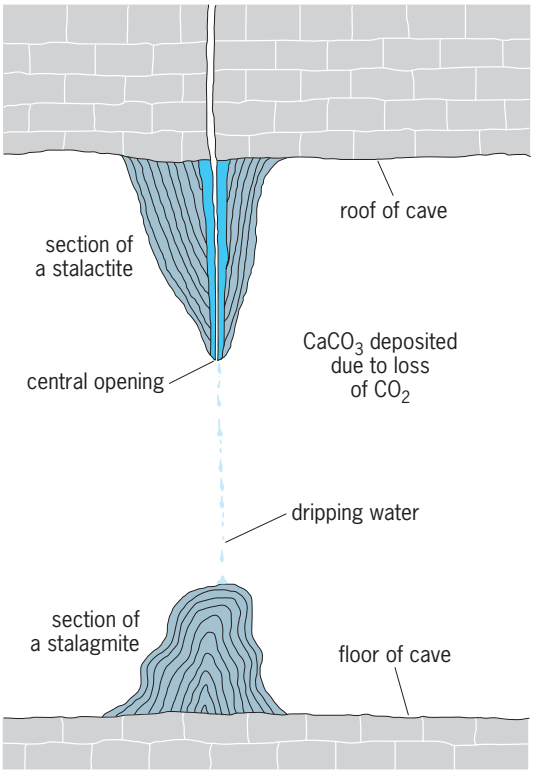

Stalactite and stalagmite deposits. Along the central opening are deposits made by water as it dripped from the crack in the limestone. Outside layers were deposited by water flowing over the surface. Water may drip to the floor and create a stalagmite. (*After E. B. Branson et al., Introduction to Geology, 3d ed., McGraw-Hill, 1952*)

shorter than stalactites and have no central hollow core. The banding of stalagmites is parallel to the surface at the time of deposition of each band.

Dripstone and flowstone are travertine accumulations of much the same origin as stalactites and stalagmites (evaporation from cave waters) but are irregular. Their irregularity is due to deposition on irregular sloping walls or surfaces in the cave rather than drips from the roof. *See* CALCITE; CAVE; LIMESTONE; TRAVERTINE. Raymond Siever

Bibliography. D. Gillieson, *Caves: Processes, Development, Management*, 1997; C. A. Hill and P. Forti, *Cave Minerals of the World*, 2d ed., 1997.

Stall-warning indicator

A device that determines the critical angle of attack for a given aircraft, at which point the lift stops increasing and the aircraft will no longer sustain itself in steady-state condition (level flight or climb/descent). The indicator usually operates from vane sensors, airflow pressure sensors, tabs on the leading edge of the wings, and computing devices which include accelerometers, airspeed detectors, and vertical gyros. The indicator may be a pointer over a dial, lights, or an audible signal. In some cases, the computer may actuate a control-column shaker to alert the pilot. *See* AIRCRAFT INSTRUMENTATION; AIRPLANE; AUTOPILOT. James W. Angus

Standard (physical measurement)

An accepted reference sample which is used for establishing a unit for the measurement of physical quantities. A physical quantity is specified by a numerical factor and a unit; for example, a mass might be expressed as 8 g, a length as 6 cm, and a time interval as 2 min. Here the gram is a mass unit defined in terms of the international kilogram, which serves as the primary standard of mass. The centimeter is defined in terms of the international meter, which is the primary standard of length and is defined as the length of path traveled by light in a vacuum during a time interval of 1/299 792 458 of a second. In similar fashion, the minute is a time interval defined as 60 s, where the second is the international standard of time and is defined as the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine energy levels of the ground state of the cesium-133 atom.

The National Institute of Standards and Technology in the United States and comparable laboratories in other countries are responsible for maintaining accurate secondary standards for various physical quantities. *See* ELECTRICAL UNITS AND STANDARDS; LIGHT; METRIC SYSTEM; PHYSICAL MEASUREMENT; TIME. Dudley Williams

Standard model

The theory that explains the three major interactions of elementary particle physics—the strong interaction responsible for nuclear forces, the weak interaction responsible for radioactive decay, and the electromagnetic interaction—in terms of a common physical picture. The model for this picture is quantum electrodynamics, the fundamental theory underlying electromagnetism. In that theory, electrons, viewed as structureless elementary constituents of matter, interact with photons, structureless elementary particles of light. The standard model extends quantum electrodynamics to explain all three interactions of subnuclear physics in terms of similar basic constituents.

Quantum electrodynamics (QED). In the 1940s, Richard P. Feynman, Julian S. Schwinger, and Shin-Ichiro Tomonaga constructed quantum electrodynamics, the quantum theory of the interactions of electrons with electromagnetic fields. They represented the electron as a pointlike particle, possessing no size but only mass and intrinsic angular momentum, or spin. Similarly, they represented electromagnetic waves as composed of pointlike quanta, photons, possessing spin but no mass. They assigned these particles the simplest equations of motion consistent with Einstein's special relativity, and postulated also the simplest relativistic coupling which allowed electrons to emit and absorb photons one at a time. This showed that these ingredients reproduce the observed properties of electromagnetism and its interaction with electrons. The experimental validation of this theory has only improved with time.

Computations in quantum electrodynamics of the intrinsic magnetic field of an electron agree with experiment to the level of parts per trillion. *See* ELECTRON; LIGHT; PHOTON; QUANTUM ELECTRODYNAMICS.

Spin. It may seem strange that a particle of zero size can have intrinsic spin. This is one of the unexpected and profound consequences of combining relativity theory and quantum mechanics. In quantum mechanics, angular momentum associated with motion takes only the discrete values which are multiples of $\hbar$, Planck's constant divided by 2π. Relativistic particles can also have intrinsic angular momentum, and this may come in either units or half-units of $\hbar$. In quantum theory, a relativistic particle is viewed as a quantum of the quantized oscillations of fields extending throughout space. For each possible value of the spin, there is a unique simplest form for the equations of propagation of the field. The very simplest, the Klein-Gordon equation, is associated with a spinless (spin-0) particle. For a particle with intrinsic angular momentum of one unit of $\hbar$ (spin 1), the required propagation equations are precisely Maxwell's equations for electromagnetic waves. For a particle with a spin of half a unit of $\hbar$ (spin $^1/_2$), the characteristic equation is called the Dirac equation. Particles whose spin is an integer multiple of $\hbar$ are called bosons; particles with an extra half-unit of spin are called fermions. It is a consequence of relativistic quantum mechanics that multiple bosons can occupy the same location. When present in great numbers, these particles can form macroscopic fields, such as the electric field. On the other hand, fermions in the same spin state are forbidden from occupying the same location in space. This property accounts for the Pauli exclusion principle, which leads to the rigidity of atoms and of atomic nuclei. It is thus natural to build up matter from elementary fermions, and to represent fields of force by elementary bosons. In quantum electrodynamics, the electrons have spin $^1/_2$ and the photons have spin 1. *See* BOSE-EINSTEIN STATISTICS; EXCLUSION PRINCIPLE; FERMI-DIRAC STATISTICS; MAXWELL'S EQUATIONS.

The spin also affects the ability of a particle to obtain mass. For spin-0 particles, there is a zero-mass Klein-Gordon equation, but one may add mass simply by adding one extra term to this equation. For spin $^1/_2$, the simplest Dirac equation involves two massless states; a particle which spins in the left-handed sense and an antiparticle which spins in the right-handed sense. The electron of quantum electrodynamics is described by putting together two of these pairs of states, to make up the left- and right-handed quantum states of the electron and the left- and right-handed states of the antielectron or positron. Only when the two halves are present is it possible to put a mass into the equation of motion. As an outcome, the electron and positron are predicted to have equal mass and opposite electric charge. (These relations are confirmed experimentally to an accuracy of 40 parts per billion.) The equation of motion for spin-1 particles does not allow a mass to be added at all, except under a special circumstance, explained later, in which the spin-1 particle couples to a spin-0 particle or field. This barrier can explain why the photon has zero mass. (Experimentally, the photon is observed to have mass less than 10^{-22} times the electron mass.) *See* RELATIVISTIC QUANTUM THEORY; SPIN (QUANTUM MECHANICS); WAVE EQUATION.

Quarks and leptons. The subnuclear world contains many more constituents than electrons and photons. First of all, two heavier spin-$^1/_2$ particles have been observed which have the same electromagnetic charge as the electron (e) but larger mass. These particles, the μ and τ, are unstable; they decay to the electron with the simultaneous emission of very light, uncharged spin-$^1/_2$ particles called neutrinos. Experiments have demonstrated that there is one neutrino species for each of the three charged particles e, μ, and τ. The three charged and three neutral fermions and their antiparticles are collectively called leptons. *See* LEPTON; NEUTRINO.

Next, experiments have observed a very large number of particles which experience the strong interactions. These particles, collectively called hadrons, include the proton and neutron, the π and K mesons, and many short-lived heavier species. It has been established that the hadrons are composite systems built from more elementary fermions, called quarks. The known list of hadrons can be built up from six species of quarks: u (up), d (down), s (strange), c (charm), b (bottom), and t (top). (The t quark was discovered in 1995.) For example, the proton is a bound state of three quarks (uud); the π^+ meson is a bound state of a quark and an antiquark (u–d). To give the correct electric charge assignments for the hadrons, the quarks must be assigned fractional electric charges, $^2/_3$ or $-^1/_3$. In addition, each species of quark must have three possible states of an invisible quantum number, called color. Hadrons built of three quarks, called baryons, must contain one quark of each color. With this rule, the bound states of three quarks consistent with the Pauli exclusion principle give precisely the observed set of baryons. *See* BARYON; COLOR (QUANTUM MECHANICS); HADRON; MESON; QUARKS.

Experiments at very high energy have confirmed the quark picture of hadrons. The pattern of scattering of high-energy electrons and neutrinos from protons confirms that quarks have fractional electric charge and spin $^1/_2$. When electrons and positrons collide at high energy and annihilate, the resulting burst of energy produces hadrons at a rate three times larger than the simple prediction from electrodynamics, indicating the presence of the three color states. The hadrons form two back-to-back jets of particles, following the directions that would be preferred by spin-$^1/_2$ quarks and antiquarks.

In the standard model, all matter is built from the quarks and the leptons. These 12 fermions organize themselves into three families, each containing the same set of electric charges. **Table 1** shows this organization and gives the mass of each particle. The masses of the u, d, and s quarks are quite uncertain; because most of the rest energy of a proton, for example, comes from the strong-interaction binding

TABLE 1. Elementary matter particles of the standard model*

Quarks				Leptons		
Charge = $\frac{2}{3}$		Charge = $-\frac{1}{3}$		Charge = 0 (neutrinos)	Charge = −1	
u	3	d	5	ν_e	e	0.5
c	1200	s	100	ν_μ	μ	105.7
t	174,000	b	4200	ν_τ	τ	1777.0

*The number next to each charged particle shows its observed mass, in MeV. (The scale may be referred to the proton mass: m_p = 938 MeV.) The neutrinos have very small masses.

rather than from the intrinsic masses of the quarks, these values must be inferred indirectly. Nevertheless, there is a clear hierarchical pattern to the spectrum of masses. The origin of this pattern is still a mystery.

Symmetry groups and quantum states. To discuss further the relations between the various quarks and leptons, it is necessary to review how symmetry operations are treated in quantum mechanics. Rotations in space provide a convenient example. A classical arrow can point in any direction; a classical rotation moves the arrow from one definite direction to another. In quantum mechanics, the Heisenberg uncertainty principle forbids one from specifying the orientation of an arrow. Instead, one starts with a fixed number of quantum states corresponding roughly to different orientations. The number of these states depends on the angular momentum. For example, a system with angular momentum 1, in units of $\hbar$, has three possible quantum states; a system with spin $^1/_2$ has two. If the system starts in one of these states and is rotated, the result is a weighted sum of the possible states. By rotating far enough, it is possible to convert the original state completely into one of the others. Conversely, a rotation in space can be thought of as built up out of three basic operations: small rotations about the x, y, and z axes (**Fig. 1**). A continuous family of symmetry operations built up in this way is called a Lie group. The basic operations are called the generators of the group. The mathematical notion of a Lie group was actually invented in the nineteenth

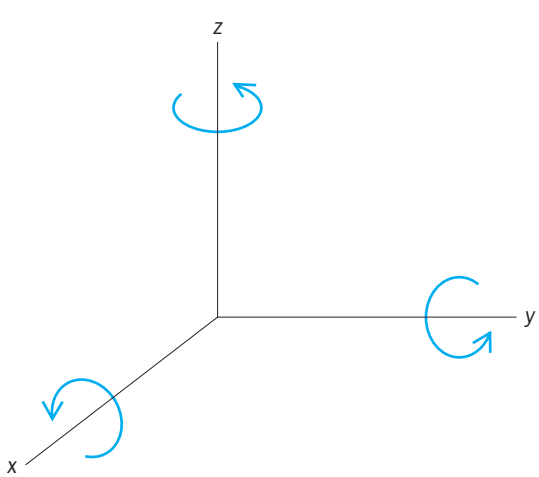

Fig. 1. Basic rotations in three-dimensional space.

century. After the discovery of quantum mechanics, Hermann Weyl and Eugene P. Wigner found that this description of symmetries is the one that fits naturally into quantum theory. *See* GROUP THEORY; LIE GROUP; SYMMETRY LAWS (PHYSICS).

To understand the symmetry structure of the standard model, imagine setting all of the quark and lepton masses equal to zero. Then, as far as the information in Table 1 is concerned, the three families of quarks and leptons would be exactly equivalent. In this limit, there are four types of symmetry operations that could be applied to the states in the table. The most obvious, called a flavor rotation, transforms or rotates the states of one family into those of another. Another possible operation transforms the leptons into the quarks of one particular color. A third, called a color rotation, transforms the quarks of one color state to those of another. A fourth, called an isospin rotation, transforms the two quarks of each family into one another, and, similarly, the two leptons. These last two operations are especially important for the structure of the standard model. These operations have a mathematical structure equivalent to that of the group of rotations in a space with, respectively, three and two complex-valued coordinates. Mathematicians refer to these rotation groups as $SU(3)$ and $SU(2)$. The two groups are generated, respectively, from eight and three basic operations, in the sense discussed at the end of the previous paragraph.

Yang-Mills theory. In quantum electrodynamics, photons do not interact with one another, except indirectly through their coupling to matter. It can be shown that relativity and quantum mechanics put severe constraints on the possible interactions of spin-1 bosons. The unique consistent prescription was discovered by Chen Ning Yang and Robert L. Mills in 1954. To construct a Yang-Mills theory, one begins with a Lie group and assigns a spin-1 particle to each generator of the group. Matter particles are assigned to quantum states which are rotated into one another by the group operations. Then the spin-1 bosons are coupled to matter in such a way that the emission or absorption of such a boson converts one matter state into another according to the group transformation law. This prescription also causes the bosons to be coupled to one another: For rotations in space, as can be seen from Fig. 1, a rotation about x causes a rotation about y to be transformed into a rotation about z. Then the spin-1 boson associated

with x rotations can absorb a spin-1 boson associated with y and transform it into the spin-1 boson associated with z. The probability for boson absorption or emission is governed by a number g called the coupling constant, which generalizes the electric charge of electrodynamics. In fact, electrodynamics can be considered as a particularly simple case of Yang-Mills theory in which the Lie group is the group of rotations of a circle, which has only one generator; mathematicians call this group U(1). In the 1970s, Gerardus 't Hooft and Martinus J. G. Veltman, Andrei A. Slavnov, John C. Taylor, Benjamin Lee, and others showed that the Yang-Mills prescription gives a well-defined quantum theory of spin-1 bosons, and that prescriptions which lack some elements of this structure lead to inconsistencies. The spin-1 bosons which play the central role in Yang-Mills theory are called gauge bosons. *See* GAUGE THEORY.

The standard model can now be defined as the theory of elementary particles in which matter is constructed from the spin-$^1/_2$ fermions in Table 1 and the interactions of these particles are generated from a Yang-Mills theory based on the product of three independent groups of symmetry operations, SU(3) × SU(2) × U(1). The SU(3) group is taken to be the group of color rotations, as defined at the end of the previous section. The SU(2) group is taken to be the group of isospin rotations, but acting only on the left-handed quark and lepton states. The U(1) factor refers to one extra gauge boson similar to the photon of quantum electrodynamics. Using the mathematical structure described by Yang and Mills, this prescription leads to a compact set of equations describing massless fermions and massless spin-1 bosons. **Table 2** lists, on the left, the three component symmetry groups of the standard model and a measure of their intrinsic coupling strengths; on the right, the three corresponding interactions of subnuclear physics, with their relative strengths and the spin-1 gauge bosons that mediate each.

Quantum chromodynamics (QCD). At this stage, it may seem that the standard model is an elegant physical theory, but it is not apparent that the model has anything to do with nature. The strong and weak interactions are both short-range; in quantum theory, this implies that they are mediated by massive, not massless, bosons. And a symmetry which acts only on left-handed and not right-handed fermions precludes combining these states to form the observed massive quarks and leptons. To show that this mathematical theory does describe nature requires new physical mechanisms.

The strong interactions will be considered first. In the standard model, these interactions are described by an $SU(3)$ Yang-Mills theory acting on color, called quantum chromodynamics. The eight gauge bosons are called gluons. The most important qualitative feature of the strong interactions is that the observed physical particles, the hadrons, are bound states of quarks, while the individual quarks are not observed. Remarkably, this feature can be derived from quantum chromodynamics when its quantum-mechanical aspects are fully taken into account.

In quantum theory, the size of charges or coupling constants is not fixed but can vary as a function of the energy scale. In quantum electrodynamics, the electric charge e is fixed at macroscopic distances but becomes larger at very small distances. This is the result of an effect called vacuum polarization. Since particle-antiparticle pairs can be created by quantum processes, such pairs are constantly being created and destroyed, everywhere in space, by quantum-mechanical fluctuations. If there is an electric field present, the electrons and positrons created by these fluctuations will move to partially screen out the field. Then the field observed at larger distances will be smaller. An electron-positron pair in such a fluctuation can have any separation up to the electron Compton wavelength, 4×10^{-12} m. If two charged objects are separated by a distance smaller than this, they will be inside the fluctuating cloud of particles and antiparticles and will feel a stronger electric interaction. The strength of the electromagnetic force can be expressed by the dimensionless ratio called the fine structure constant, $\alpha = e^2/\hbar c$, where c is the speed of light. At macroscopic distances, $\alpha = 1/137$, but in experiments in which high energies are used to bring electrons and positrons to within distances less than 10^{-18} m, the value $\alpha = 1/129$ is observed.

In Yang-Mills theories in which the gauge bosons interact directly, there is a compensatory quantum effect in the opposite direction. Through this effect, the interaction of a gauge electric field with a quantum fluctuation can reinforce its strength. This effect, called asymptotic freedom because it implies that the coupling constant is smaller at short distances, was discovered in 1973 by David Politzer, David J. Gross, and Frank Wilczek. In quantum chromodynamics, this effect overwhelms the vacuum polarization effect and makes the coupling grow dramatically as

TABLE 2. Elementary interactions of standard model

Symmetry group	Coupling strength (α)*	Bosons	Interaction	Strength
SU(3)	1/8.4	Gluons	Strong	1
SU(2)	1/29.6	Photon	Electromagnetic†	10^{-2}
U(1)	1/98.4	W^+, W^-, Z^0	Weak†,‡	10^{-5}

*Coupling strengths are given at a high energy, equal to the rest energy of the Z^0 boson, 91,000 MeV, and are written as pure numbers by defining $\alpha = g^2/c$.
†Weak and electromagnetic interactions are combinations of the basic SU(2) and U(1) interactions, as explained in text.
‡ The small strength of the weak interactions is due to the large masses of the W and Z bosons.

objects with color are separated. States with widely separated quarks or gluons acquire very high energy, so that it is always energetically favorable to create additional quark-antiquark pairs to neutralize the color charges. Then these configurations decay rapidly to quark and antiquark bound states. The separation at which the quantum chromodynamics coupling becomes strong, which sets the size of typical hadrons, is about 10^{-15} m. This corresponds to a binding energy of roughly $m_p c^2$, where m_p is the proton mass. The massive bosons which mediate the nuclear force are then quark-antiquark bound states, mainly π and ρ mesons, which acquire their mass from their internal quantum chromodynamics energy.

Asymptotic freedom implies that experiments at energies well above the proton rest energy should be able to see quarks and gluons. At the end of the discussion of quarks, it was noted that the hadrons resulting from the annihilation of electrons and positrons at high energy form pairs of collimated jets that follow the directions expected for a quark-antiquark pair. The detailed properties of a jet indicate that it originates as a single quark or antiquark and acquires its complexity from the production of additional pairs along the gauge electric field direction. In 1979, experiments at the German accelerator PETRA observed events with three jets of hadrons, with the distribution of the third jet in energy and angle very similar to that expected for radiation of photons from the produced quark and antiquark. These additional jets originated as radiated gluons. *See* GLUONS; QUANTUM CHROMODYNAMICS.

Higgs mechanism. The massive particles which mediate the weak interactions appear experimentally to be spin-1 bosons with the simple couplings of Yang-Mills gauge bosons. There are three of these massive particles, called W^+, W^-, and Z^0. Given the fundamental difficulties noted above in writing a mass term for spin-1 particle, it is not clear how these states could arise from a Yang-Mills theory. There is one way known in which Yang-Mills bosons can acquire mass, discovered in the 1960s by Peter W. Higgs and a number of others. Though in a Yang-Mills theory the basic equations must be symmetric, it is possible that nature realizes these equations in an asymmetrical way. For example, with respect to rotations in a three-dimensional space (Fig. 1), if we lived in a medium which randomly chose a preferred direction in this space, say, the z direction, only the rotation about z would be recognizable as a symmetry. The other two rotation symmetries are said to be spontaneously broken. Such spontaneous symmetry breaking occurs in magnetic materials when they choose an axis of magnetization, and in crystals when they choose axes of alignment.

If the symmetries that are spontaneously broken are those of a Yang-Mills theory, the corresponding gauge bosons become massive. For example, if some field took up an orientation that preferred the z axis in Fig. 1, the gauge bosons corresponding to x and y rotations would become massive, while the one associated with z rotations would remain massless. It is possible to turn this picture into a candidate for a unified model of weak and electromagnetic interactions, in which the two massive bosons are the W^+ and W^-, and the massless spin-1 boson is interpreted as the photon. Experiment chooses a different theory built up around the same idea, constructed in the 1960s by Sheldon L. Glashow, Steven Weinberg, and Abdus Salam (GWS). This theory makes use of the SU(2) and U(1) pieces of the standard model Yang-Mills symmetry. This symmetry group has four generators in all. Three of these are spontaneously broken and one remains unbroken. This gives the massive W^+, W^-, and Z^0, plus the massless photon. The field that is responsible for the symmetry breaking is still a complete mystery. In the simplest model, the symmetry breaking can be due to one spin-0 field, called the Higgs field; the quanta of this field are called Higgs bosons. In this model, the Higgs field also is responsible for coupling the left- and right-handed components of quarks and leptons and thus for generating the fermion masses. *See* HIGGS BOSON; SYMMETRY BREAKING.

Electroweak interactions. The Glashow-Weinberg-Salam theory describes electroweak interactions. This theory contains and relates three different types of interactions: electromagnetism, mediated by the photon; charge-changing processes, mediated by the W bosons; and the neutral-current interaction, mediated by the Z boson. The latter two processes involve couplings that are stronger than those of electromagnetism, but they appear weak in nuclear physics because the W and Z have large masses, 80,4000 and 91,188 MeV, respectively, in units in which the proton mass is 938 MeV.

The charge-changing processes include the radioactive beta-particle decay of nuclei, which corresponds in this picture to a d quark inside a neutron changing to a u quark and a quantum of the W^- field, which materializes as an electron and an electron antineutrino. The fundamental coupling involves only left-handed species, and this fact leads to predictions for the energy and angular distributions of the decay products which are in good accord with experiment. Similar predictions can be made for the products of more exotic reactions such as μ and τ decay and the scattering of high-energy neutrinos, and again experiments confirm the simple picture in which only the left-handed fermion states couple to the W bosons. *See* RADIOACTIVITY.

The neutral-current interaction, a short-ranged scattering interaction between fermions which supplements the electromagnetic interaction, is a characteristic feature of the Glashow-Weinberg-Salam theory. The charges by which the various fermions couple to this interaction, or to the Z boson, are given by the electric charges combined with an extra factor for left-handed species. The neutral-current interaction was first observed in the 1970s in experiments on the scattering of high-energy neutrinos. Accelerators at the Stanford Linear Accelerator Center (SLAC) in California and at CERN, in Geneva, Switzerland, have produced individual Z bosons by the annihilation of electrons and positrons and have observed

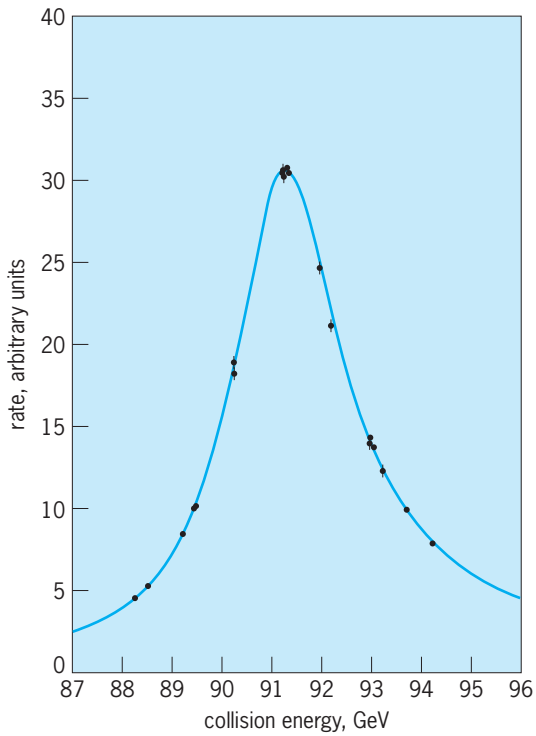

Fig. 2. Measurements of the rate for electron-positron annihilation into a Z^0 boson as a function of the collision energy, from the data of the OPAL experiment at CERN.

the decay of the Z to all of the fermion species of the standard model (except for the very heavy t quark). The results confirm the Glashow-Weinberg-Salam charge assignments in detail. One piece of this confirmation begins from the fact that in quantum mechanics an unstable state with lifetime τ has an uncertainty in energy given by $\hbar/\tau$. Then a reaction which produces the unstable state proceeds over a range of energies, with the rate varying in a characteristic shape, called a resonance. **Figure 2** shows the rate for electron-positron annihilation to form a Z boson, as measured by the OPAL experiment at CERN. The width of the resonance measures the lifetime of the Z. The agreement shown between theory and experiment actually tests all three of the interactions of the standard model—the weak interactions, because the decay rate of the Z depends on the coupling to each fermion species; the strong interactions, because the decay rate to quarks is enhanced by the color factor 3 and another 4% correction due to gluon effects; and the electromagnetic interactions, because the radiation of photons from the incident electrons and positrons produces the asymmetrical shape of the resonance curve. In this and other comparisons, the Z decay rate agrees with the prediction of the standard model to an accuracy of 0.1%. *See* ELECTROWEAK INTERACTION; INTERMEDIATE VECTOR BOSON; NEUTRAL CURRENTS; WEAK NUCLEAR INTERACTIONS.

Open questions. The standard model brings together diverse aspects of the behavior of subnuclear particles in a common framework. We can truly claim to understand the nature of the strong, weak, and electromagnetic interactions. The explanation, though, sets up a number of new questions. The first among these is the nature of the agent that is responsible for spontaneously breaking SU(2) × U(1) symmetry. Is it the Higgs field described above or a more complex set of interactions? Whatever this agent is, it is responsible for the masses of the heavy spin-1 bosons W and Z, and for all of the quark and lepton masses. A related question is the origin of the pattern of fermion masses shown in Table 1. The origin of the Yang-Mills symmetry group SU(3) × SU(2) × U(1) is also a mystery. In 1973, Howard Georgi and Glashow and Jogesh C. Pati and Salam proposed that this group orginates as the set of unbroken symmetries that remain when a larger symmetry group, called the grand unification symmetry, is spontaneously broken at very short distances. The choice of SU(5) for the grand symmetry group explains the pattern of electric charge assignments of the quarks and leptons. New measurements of the standard model coupling constants have turned out to be inconsistent with the simplest form of the grand unification hypothesis, but they fit the expectation if the assumption is added of an enhanced space-time symmetry called supersymmetry, which interconnects bosons and fermions. *See* GRAND UNIFICATION THEORIES; SUPERSYMMETRY.

At least a part of the answers to these questions should be found at the next generation of particle accelerators. If there is a Higgs boson, its mass should not be far above the masses of the W and Z. If grand unification is correct, new particles predicted by supersymmetry, or another principle which extends the standard model, should also appear. With luck, these experiments should lead to an even simpler and more all-encompassing mathematical description of nature. *See* PARTICLE ACCELERATOR.

Michael E. Peskin

Bibliography. G. D. Coughlan, J. E. Dodd, and B. M. Gripaios, *The Ideas of Particle Physics*, 3d ed., 2006; R. P. Crease and C. C. Mann, *The Second Creation*, rev. ed., 1996; J. F. Donoghue, E. Golowich, and B. R. Holstein, *Dynamics of the Standard Model*, 1995; D. Griffths, *Introduction to Elementary Particles*, 1987; F. Halzen and A. D. Martin, *Quarks and Leptons*, 1984; G. Kane, *Modern Elementary Particle Physics*, 2d. ed., 1993; G. Kane, *The Particle Garden*, 1995; M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, 1995.

Staphylococcus

A genus of bacteria containing at least 28 species that are collectively referred to as staphylococci. Their usual habitat is animal skin and mucosal surfaces. Although the genus is known for the ability of some species to cause infectious diseases, many species rarely cause infections. Pathogenic staphylococci are usually opportunists and cause illness only in compromised hosts. *Staphylococcus epidermidis* is a major component of the human skin flora. *Staphylococcus aureus*, the most pathogenic species, is

usually identified by its ability to produce coagulases (proteins that affect fibrinogen of the blood-clotting cascade). Since most other species of staphylococci do not produce coagulase, it is useful to divide staphylococci into coagulase-positive and coagulase-negative species. Coagulase-negative staphylococci are not highly virulent but are an important cause of infections in certain high-risk groups. Although *Staphylococcus* infections were once readily treatable with antibiotics, some strains have acquired genes making them resistant to multiple antimicrobial agents. *See* BACTERIA.

Cell structure, physiology, and genetics. *Staphylococcus* cells are spherical with a diameter of 0.5–1.5 micrometers. Clumps of staphylococci resemble bunches of grapes when viewed with a microscope, owing to cell division in multiple planes. The staphylococci have a gram-positive (retain violet color of Gram stain after destaining with alcohol) cell composition, with a unique peptidoglycan structure that is highly cross-linked with bridges of amino acids. *See* STAIN (MICROBIOLOGY).

Most species are facultative anaerobes (grow in the presence or absence of oxygen). Within a single species, there is a high degree of strain variation in nutritional requirements. Staphylococci are quite resistant to desiccation and high-osmotic conditions and can grow in the presence of 10% sodium chloride. These properties facilitate their survival in the environment, growth in food, and communicability.

In addition to genetic information on the chromosome, pathogenic staphylococci often contain accessory elements such as plasmids, bacteriophages, pathogenicity islands (DNA clusters containing genes associated with pathogenesis), and transposons. These elements harbor genes that encode toxins or resistance to antimicrobial agents and may be transferred to other strains. Genes involved in virulence, especially those coding for exotoxins and surface-binding proteins, are coordinately or simultaneously regulated by loci on the chromosome such as the accessory gene regulator, *agr*, and the *S. aureus* regulator, *sar*. *See* BACTERIAL GENETICS; BACTERIOPHAGE; PLASMID; TRANSPOSONS.

Resistance to antimicrobial agents. Staphylococci have a remarkable ability to acquire resistance to antimicrobial agents. Presently, most *S. aureus* isolates are resistant to penicillin and many other antibiotics. The most common resistance mechanism is mediated by production of beta-lactamase enzymes, which inactivate antibiotics by preventing their inhibition of bacterial cell wall synthesis. Methicillin is the drug of choice for strains that are resistant to multiple antimicrobial agents. Unfortunately, methicillin-resistant *S. aureus* is being isolated with increasing frequency, especially from patients with nosocomial (hospital-acquired) infections. For most of these isolates, vancomycin is the drug used as a last resort. In the past few years, a few isolates with partial resistance to vancomycin have been found, causing concern that this organism may become untreatable unless new drugs become available. *See* DRUG RESISTANCE.

Strain typing. Infections caused by staphylococci, especially *S. aureus*, may be communicable, and it is important to trace the source of organism in outbreaks. Several typing methods have been used for epidemiologic studies. The classical phenotypic method, phage typing, was employed successfully for decades and is based on the fact that *S. aureus* may carry bacteriophages that can infect some less-related isolates. Based on susceptibility profiles, strains are grouped into five types. Because of technical difficulty and the emergence of nontypeable isolates, phage typing is being replaced by molecular techniques. One technique, pulsed-field gel electrophoresis, analyzes chromosomal DNA fragments generated by the restriction enzyme *Sma*I. Since only a few sites are recognized by *Sma*I in the *S. aureus* chromosome, a limited but variable number of fragments are generated. These fragments can be separated by pulsed-gel electrophoresis, and the profile generated represents a "fingerprint" for the strain in question. *See* GENETIC MAPPING.

Coagulase-positive (S. aureus) infections. *Staphylococcus aureus* causes approximately 8.9 million cases of human disease annually in the United States. Many people carry *S. aureus* and can develop endogenous infections if they become immunocompromised or if the organism enters the body through a break in the skin or mucous membrane. Infections acquired from exogenous sources are also common. Exogenous sources can be other people, fomites (inanimate objects, such as dishes, clothing, towels), the environment, or occasionally animals. *Staphylococcus aureus* strains are variable in expression of multiple-potential virulence factors; therefore, an encounter with this organism can lead to three types of illness: pyogenic infections, toxigenic infections, and food-borne intoxications.

Pyogenic infections. *Staphylococcus aureus* infections are common, and most develop into a pyogenic (pus-forming) lesion caused by acute inflammation. Inflammation helps eliminate the bacteria but also damages tissue at the site of infection. Typical pyogenic lesions are abscesses with purulent centers containing leukocytes, fluid, and bacteria. The abscess is enclosed by fibrin and cells. Many *S. aureus* products trigger inflammation, but the most important are cell wall components and several staphylococcal exotoxins (proteins that are secreted and not associated with the cell). Exotoxins with hemolytic or cytotoxic activity destroy tissue by damaging host cell membranes. Some also act on leukocyte membranes and prolong survival of *S. aureus* in the abscess.

Pyogenic infections can occur anywhere in the body. Skin infections are common and may be communicable. Styes and boils are uncomplicated infections of hair follicles in the eyelid or sweat or sebaceous glands in the skin. More extensive skin infections are termed carbuncles or furuncles. Infections may also develop internally if staphylococci enter through a wound or escape from an abscess. Invasion may occur through a direct extension of an abscess or via the bloodstream. Blood infections

(septicemia) can disseminate the organism throughout the body and abscesses can form internally. Infections in certain parts of the body, such as in the bone (osteomyelitis), respond poorly to antimicrobial therapy.

Toxigenic infections. Certain strains of *S. aureus* produce exotoxins that mediate two illnesses, toxic shock syndrome and staphylococcal scalded skin syndrome. In both diseases, exotoxins are produced during an infection, diffuse from the site of infection, and are carried by the blood (toxemia) to other sites of the body, causing symptoms to develop at sites distant from the infection.

Toxic shock syndrome is an acute life-threatening illness mediated by staphylococcal superantigen exotoxins. Toxic shock syndrome patients develop fever, hypotension, and multiorgan involvement and subsequent desquamation (peeling of the skin). The occurrence of staphylococcal toxic shock syndrome in the United States is approximately 0.3 to 0.8 case per 100,000 people. The *S. aureus* superantigen family contains nine antigenic variants of staphylococcal enterotoxins and toxic shock syndrome toxin-1 (TSST-1). Superantigens are unique molecules that bind simultaneously to major histocompatibility complex (MHC) Class II molecules on antigen-presenting cells and to receptors on T cells. Unlike conventional antigens, superantigens are neither processed by antigen-presenting cells nor expressed within the MHC Class II peptide groove. Instead, superantigens bind to peripheral surfaces of MHC Class II molecules and to the T-cell receptor at semiconserved elements of the β-chain variable region (Vβ). Binding in this manner stimulates a high percentage of T cells. The activated immune cells express excessive quantities of cytokines, which generate fever and adverse effects on the cardiovascular system, resulting in shock. Menstrual toxic shock syndrome is caused by TSST-1 produced by *S. aureus* colonizing the vagina or cervix. This illness is linked to tampon use during menstruation, with absorbency being a risk factor. Nonmenstrual toxic shock syndrome occurs when *S. aureus* causes an infection elsewhere in the body and produces TSST-1 or enterotoxins. *See* CELLULAR IMMUNOLOGY; TOXIC SHOCK SYNDROME.

Staphylococcal scalded skin syndrome, also known as Ritter's disease, refers to several staphylococcal toxigenic infections. It is characterized by dermatologic abnormalities caused by two related exotoxins, the type A and B exfoliative (epidermolytic) toxins. Exfoliative toxins are superantigens, but they are not typically associated with toxic shock syndrome. Instead, they are proteases that cause the separation of granular cells within the epidermis and skin peeling in a single plane at the stratum corneum. Bullous impetigo is a localized form of staphylococcal scalded skin syndrome in which the toxin acts only at the site of infection to form a liquid-filled lesion termed a bulla. Systemic staphylococcal scalded skin syndrome occurs when the exfoliative toxin disseminates via toxemia. One form of systemic illness, generalized exfoliative disease, is characterized by skin peeling in large sheets over most parts of the body. In an alternate form, staphylococcal scarlantiniform eruption, patients develop a generalized roughened sandpaper appearance of the skin without bulla formation or generalized exfoliation. Since nearly all people develop antibodies early in life, systemic staphylococcal scalded skin syndrome is rare after the age of 5.

Standard therapy for toxigenic infections is to identify and control the infection to minimize the amount of toxin produced. Toxic shock syndrome and severe staphylococcal scalded skin syndrome are typically treated with intravenous administration of beta-lactamase-resistant, cell-wall-active agents such as methicillin. In toxic shock syndrome, it is also important to provide supportive treatment for shock and counteract serum chemistry abnormalities. Intravenous immunoglobulin has proven successful in reducing severity of toxic shock syndrome; however, there is currently no vaccine for toxic shock syndrome or staphylococcal scalded skin syndrome.

Food-borne intoxications. At least 1 million cases of staphylococcal food poisoning occur annually in the United States. Staphylococcal food poisoning is not an infection, but an intoxication that results from ingestion of staphylococcal enterotoxins in food. The enterotoxins are produced when food contaminated with *S. aureus* is improperly stored under conditions that allow the bacteria to grow. Although contamination can originate from animals or the environment, food preparers with poor hygiene are the usual source.

The staphylococcal enterotoxins are stable and survive moderate heating. If ingested, they act in the gastrointestinal tract to stimulate the vagus nerve and signal the emetic center in the medulla. Vomiting, diarrhea, and abdominal cramps are the usual symptoms of staphylococcal food poisoning. Symptoms begin 1–6 hours following ingestion and usually resolve within 24 hours. Most cases do not require treatment. Antimicrobial agents have no effect. Severe cases accompanied by dehydration should be treated with intravenous fluids and antiemetic agents to control vomiting. Fatalities are rare and usually occur only in the very young or elderly. Because of toxin stability, the illness is not routinely preventable by heating foods. Pasteurization is effective in eliminating organisms in dairy products, although post-pasteurization contamination by human sources is possible. *See* FOOD POISONING.

Effective methods for preventing staphylococcal food poisoning are aimed at eliminating contamination through common hygiene practices, such as wearing gloves, and proper food storage to minimize toxin production.

Coagulase-negative diseases. Because they frequently colonize human skin and mucosal surfaces, *Staphylococcus* species other than *S. aureus* are important opportunistic pathogens for humans. Most of these have low virulence and do not produce coagulase. Although several coagulase-negative staphylococci occasionally cause infections, *S. epidermidis* and *S. saprophyticus* are frequently associated with

certain human diseases. *Staphylococcus epidermidis*, the most virulent, coagulase-negative streptococcus, is a major cause of infected prosthetic devices, shunts, and catheters. The more frequent use of these devices has led to an increased occurrence of infections. Strains that cause these infections typically produce a variety of extracellular substances that facilitate their adherence to, and subsequent accumulation on, artificial materials. *Staphylococcus saprophyticus* causes at least 10% of urinary tract infections in nonhospitalized women aged 16–25 years. The source of this pathogen is unclear, but recent sexual intercourse is the leading risk factor. *Staphylococcus saprophyticus* is not highly virulent otherwise. Its surface components that bind to the urinary epithelium are important for disease. *See* MEDICAL BACTERIOLOGY.

Animal diseases. Coagulase-positive staphylococci are the most important *Staphylococcus* pathogens for animals. Certain diseases of pets and farm animals are very prominent. *Staphylococcus aureus* is the leading cause of infectious mastitis in dairy animals. This udder infection causes nearly $2 billion in lost revenue to the United States dairy industry per year. *Staphylococcus intermedius* causes skin pyoderma in dogs and cats. *Staphylococcus aureus* and other coagulase-positive and coagulase-negative *Streptococcus* species are also important pathogens for goats, sheep, horses, pigs, poultry, and dolphins.

Greg Bohach

Bibliography. J. E. Alouf and J. H. Freer (eds.), *The Comprehensive Sourcebook of Bacterial Protein Toxins*, 1999; J. C. Buzby et al., *Bacterial Food-Borne Disease, Medical Costs and Productivity Losses*, 1996; K. B. Crosby and G. L. Archer (eds.), *The Staphylococci in Human Disease*, 1997; M. P. Doyle, L. R. Beuchat, and T. J. Montville (eds.), *Food Microbiology: Fundamentals and Frontiers*, 1997; V. R. Fischetti et al., *Gram Positive Pathogens*, ASM Press, 2000; J. G. Holt et al. (eds.), *Bergey's Manual of Systematic Bacteriology*, vol. 9, 1994; S. Ladhani et al., Clinical, microbiological, and biochemical aspects of the exfoliative toxins causing staphylococcal scalded skin syndrome, *Clin. Microbiol. Rev.*, 12:224–242, 1999; R. P. Novick (ed.), *Molecular Biology of the Staphylococci*, 1990; J. F. Timoney et al., *Microbiology and Infectious Diseases of Domestic Animals*, 8th ed., 1988.

Star

A self-luminous body that during its life generates (or will generate) energy and support by thermonuclear fusion.

Sun as reference. The fundamental reference for properties of stars is the Sun. The Sun's distance of 1.496×10^8 km (9.30×10^7 mi) from the Earth defines the astronomical unit (AU) and gives a solar radius (R) of 6.96×10^5 km (4.32×10^5 mi). From the energy received on Earth (1367 W m^{-2}), the solar luminosity (L) is 3.85×10^{26} W. Now, $L = 4\pi R^2$ (surface area) $\times\ \sigma T^4$ (luminosity per unit area according to the Stefan-Boltzmann law, where σ is the Stefan-Boltzmann constant); therefore, the effective blackbody "surface" temperature (T) is 5780 K (9950°F). (The "surface" is where the solar gases become opaque thanks to absorption by the negative hydrogen ion.) The Earth's orbital parameters and Kepler's third law give a solar mass of 1.99×10^{30} kg, the solar volume leading to an average density of 1.4 g cm^{-3} (1.4 times that of water). *See* ASTRONOMICAL UNIT; HEAT RADIATION; SUN.

The Sun's temperature and density rise inwardly as a result of gravitational compression, and at the still-gaseous solar center reach values of 15.7×10^6 K (28.3×10^6°F) and 150 g cm^{-3}. Application of atomic theory to the solar spectrum and analysis of the solar wind and oscillations show the Sun to be 92% hydrogen (by number of atoms), 8% helium, and about 0.015% everything else (oxygen dominating, then carbon, neon, and nitrogen). The solar chemical composition provides a standard for other stars as well as an observational benchmark against which to compare theories of the chemical evolution of the Milky Way Galaxy.

Above about 8×10^6 K (14×10^6°F), within the inner quarter of the radius and about half the mass, hydrogen (the proton, ^{1}H) fuses to helium primarily via the proton-proton chain:

$$^1\mathrm{H} + {}^1\mathrm{H} \rightarrow {}^2\mathrm{H} + \text{positron (positive electron)} + \text{neutrino}$$

$$\text{positron} + \text{electron} \rightarrow \text{two gamma rays}$$

$$^2\mathrm{H} + {}^1\mathrm{H} \rightarrow {}^3\mathrm{He} + \text{gamma ray}$$

$$^3\mathrm{He} + {}^3\mathrm{He} \rightarrow {}^4\mathrm{He} + 2({}^1\mathrm{H})$$

The resulting release of energy supports the Sun against gravitational contraction. The radiation works its way out, degrading in energy with decreasing temperature, until it is released largely at optical wavelengths. The neutrinos, nearly massless particles, exit the Sun at close to the speed of light. Though nearly everything is transparent to them, enough are captured by terrestrial neutrino telescopes to confirm the theory of the Sun's energy production. There was initially enough hydrogen to run the chain for 10 billion years. With the Sun 4.6 billion years old (from radioactive dating of meteorites), it has over 5 billion years of core hydrogen fusion left. *See* NEUTRINO; POSITRON; PROTON-PROTON CHAIN; SOLAR NEUTRINOS.

Names. Over 6000 stars can be seen with the unaided eye. The brightest carry proper names from ancient times, most of Arabic origin that reflect the position of the star within its constellation. A more general system (from Johannes Bayer in the early seventeenth century) names stars within constellations by Greek letters roughly in accord with apparent brightness, followed by the Latin genitive of the constellation name (for example, Vega, in Lyra, is also Alpha Lyrae). More generally yet, brighter stars carry numbers, derived from the work of the astronomer John Flamsteed (1646–1719), in easterly order within a constellation (Vega also 3 Lyrae). All naked-eye stars also have HR (Harvard Revised)

TABLE 1. First-magnitude stars*

Proper name	Greek-letter name	Apparent visual magnitude (V)	Distance, light-years	Absolute visual magnitude (M_v)	Spectral class
Sirius	α Canis Majoris	−1.46	8.6	1.43	A1 V
Canopus	α Carinae	−0.72	313	−5.63	F0 II
Rigil Kentaurus	α Centauri	−0.29	4.4	4.06	G2 V + K1 V
Arcturus	α Boötis	−0.04	37	−0.30	K1.5 III
Vega	α Lyrae	0.03	25	0.58	A0 V
Capella	α Aurigae	0.08	42	−0.48	G8 III + G0 III
Rigel	β Orionis	0.12	775	−6.8	B8 Iab
Procyon	α Canis Minoris	0.34	11	2.62	F5 IV−V
Achernar	α Eridani	0.46	145	−2.76	B3 V
Hadar	β Centauri	0.61	335	−4.44	B1 III
Betelgeuse	α Orionis	0.70	425	−4.9	M2 Iab
Acrux	α Crucis	0.76	320	−4.2	B0.5 IV + B1 V
Altair	α Aquilae	0.77	17	2.21	A7 V
Aldebaran	α Tauri	0.85	65	−0.65	K5 III
Antares	α Scorpii	0.96	600	−5.9	M1.5 Ib
Spica	α Virginis	1.04	260	−3.49	B1 III−IV + B2 V
Pollux	β Geminorum	1.14	34	1.07	K0 III
Fomalhaut	α Piscis Austrini	1.16	25	1.73	A3 V
Deneb	α Cygni	1.25	2600	−8	A2 Ia
Mimosa	β Crucis	1.25	350	−3.9	B0.5 III
Regulus	α Leonis	1.35	78	−0.53	B7 V
Adhara	ϵ Canis Majoris	1.50	430	−4.2	B2 II

*As seen by eye, doubles combined into one. Doubles are indicated by two entries in the spectral class. Low-luminosity companions are not included.

numbers assigned in order east of the vernal equinox. A variety of catalogs list millions of telescopic stars. *See* ASTRONOMICAL CATALOGS; CONSTELLATION.

Magnitudes and colors. About 130 B.C., Hipparchus assigned naked-eye stars to six brightness groups or "apparent magnitudes" (m), with first magnitude the brightest (**Table 1**). This scheme is now quantified as a logarithmic system such that five magnitudes correspond to a factor of 100 in brightness, rendering first magnitude 2.512... times brighter than second, and so on; calibration extends the brightest stars into negative numbers. The faintest stars now observed are around magnitude 30, 10^{12} times fainter than magnitude zero.

Stars assume subtle colors from red to blue-white, reflecting different spectral energy distributions that result from temperatures ranging from under 2000 K (3100°F) to over 100,000 K (180,000°F). The magnitude of a star therefore depends on the detector's color sensitivity. A blue-sensitive detector "sees" blue stars brighter than does the eye, which is most sensitive to yellow light, and will assign them lower magnitudes. Numerous magnitude systems range from the ultraviolet into the infrared, though the apparent visual magnitude ($m_v = V$) is still standard. The differences among the systems allow measures of stellar color and temperature. The total brightness at the Earth across the spectrum is given by the apparent bolometric magnitude (m_{bol}), which is dependent on the visual magnitude and temperature. *See* COLOR INDEX; MAGNITUDE (ASTRONOMY).

Distances. The fundamental means of finding stellar distances is parallax. As the Earth moves in orbit around the Sun, a nearby star will appear to shift its location against the background. The parallax (p in arcseconds) is defined as one half the total shift, and is the angle subtended by the Earth's orbital radius as seen from the star. Distance in parsecs (pc) is 1/pc, where 1 pc = 206,265 AU = 3.26 light-years. (A light-year is the distance that a ray of light will travel in a year.) The nearest star, a telescopic companion to Alpha Centauri, is 1.34 pc = 4.39 light-years away (**Table 2**). Parallaxes of hundreds of thousands of

TABLE 2. Stars within 10 light-years

Star	Distance, light-years	Spectral class	Apparent visual magnitude (V)	Absolute visual magnitude (M_v)
Proxima Centauri	4.21	M5.5 V	11.05	15.49
Alpha Centauri A	4.40	G2 V	−0.01	4.34
Alpha Centauri B	4.40	K1 V	1.33	5.68
Barnard's Star	5.94	M4 V	9.54	13.24
Wolf 359	7.66	M6 V	13.54	16.68
BD+36°2147	8.32	M2 V	7.49	10.47
Sirius A	8.61	A1 V	−1.47	1.43
Sirius B	8.61	DA†	8.44	11.33
L726-8A	8.64	M5.5 V	12.57	15.45
L726-8B	8.64	M5.5 V	12.52	15.40
Ross 154	9.69	M3.5 V	10.95	13.58

†Hydrogen-rich white dwarf.

Fig. 1. The Milky Way in Sagittarius, composed of millions of stars in the disk of the Galaxy. (© *Anglo-Australian Observatory; photograph by David Malin*)

stars found with the *Hipparcos* satellite allow fair accuracy to distances of over 1000 light-years. Parallax measures then allow the calibration of other distance methods (such as spectroscopic distance, main-sequence cluster fitting, and use of Cepheid variables) that extend much farther. *See* LIGHT-YEAR; PARALLAX (ASTRONOMY); PARSEC.

Distribution and motions. All the unaided-eye stars and 200 billion more are collected into the Milky Way Galaxy (or simply, the Galaxy), 98% concentrated into a thin disk over 100,000 light-years across. From the Sun, inside the disk and 27,000 light-years from the Galaxy's center in Sagittarius, the disk appears as the Milky Way. The appearance of the Milky Way varies considerably. It is brightest and widest toward the galactic center, and faintest in the opposite direction toward the anitcenter in Taurus and Auriga (**Fig. 1**). Surrounding the disk is a vast but sparsely populated halo.

Since stars orbit the Galaxy's center, they are seen to move relative to the Sun. Angular proper motions across the line of sight range from 10 arcseconds per year to below a thousandth of an arcsecond. Proper motions (μ, in arcseconds per year) depend on velocities across the line of sight (v_t, in kilometers per second) and distances (D, in parsecs) according to Eq. (1), which allows v_t to be found. Velocities along

$$\mu = \frac{v_t}{4.74D} \tag{1}$$

the line of sight (v_r) are measured from Doppler shifts of stellar spectrum lines. The combination yields space velocities (v_s) relative to the Sun according to Eq. (2). Statistical analysis of these motions shows

$$v_s^2 = v_t^2 + v_r^2 \tag{2}$$

the Sun to be moving through the local stars at a speed of 15–20 km/s (9–12 mi/s), roughly toward Vega. From radial velocities of sources outside the Galaxy, it is found that the Sun moves in a roughly circular orbit at 220 km/s (137 mi/s), which when combined with the space motions of other stars allows their galactic orbits to be determined. Stars in the disk have closely circular orbits; those in the halo have elliptical ones. *See* DOPPLER EFFECT; MILKY WAY GALAXY.

Absolute magnitudes. The apparent visual magnitude of a star depends on its intrinsic visual luminosity and on the inverse square of the distance. Knowledge of the distance allows the determination of the true visual luminosity, expressed as the absolute visual magnitude, M_v. This quantity is defined as the apparent visual magnitude that the star would have at a distance of 10 pc, and is found from the magnitude equation (3). Absolute visual magnitudes

$$M = m + 5 - 5 \log d \text{ (pc)} \tag{3}$$

from class O to mid L range from $M_v = -10$ to $+23$ (a factor of 10^{13}). The Sun's absolute visual magnitude is in the middle of this range, $+4.83$. When the invisible infrared radiation from faint cool stars and the equally invisible ultraviolet from the ultraluminous hot stars are taken into account, the range in absolute bolometric luminosities (expressed in magnitudes as M_{bol}) is 30 times lower. The quantity M_{bol} can be calibrated in watts; for the Sun, $M_{\text{bol}} = +4.75$.

Spectral classes. Stars exhibit a variety of absorption-line spectra. The absorptions, narrow cuts in the spectra, are produced by atoms, ions, and molecules in the stars' thin, semitransparent outer layers, or atmospheres. Some stars show powerful hydrogen absorptions, others weak hydrogen and metals, yet others molecules but no hydrogen. Over a century ago, Edward C. Pickering lettered them according to the strengths of their hydrogen lines. After he and his assistants dropped some letters and rearranged others for greater continuity, they arrived at the standard spectral sequence, OBAFGKM, which was later decimalized (the Sun is in class G2). Classes L and T were added in 1999 (**Fig. 2; Table 3**). Since the sequence correlates with color, it must also correlate with temperature, which ranges from near 50,000 K (90,000°F) for hot class O, through about 6000 K (10,000°F) for solar-type class G, to under 2000 K (3100°F) for class L, and below 1000 K (1300°F) for class T (all of which are brown dwarf substars).

The majority of stars in the galactic disk have chemical compositions like that of the Sun. Differences in spectra result primarily from changes in molecular and ionic composition and in the efficiencies of absorption, all of which correlate with temperature. Molecules exist only at low temperatures. At higher temperatures they break up, and neutral atoms become ionized, so the absorptions of the neutral state eventually weaken with increasing temperature. When the temperature becomes

high enough, single ionization is replaced by double, and so on. Hydrogen reaches maximum strength in class A. Hotter than class A, hydrogen becomes ionized and the neutral hydrogen lines weaken in classes B and O. Even the absorptions of individual ions are affected, as absorbing electrons are pumped toward higher-energy states at higher temperatures.

Most stars have more oxygen than carbon. In some cool stars, however, the ratio is reversed, carbon molecules replacing those of metallic oxides. Stars whose temperatures track class M are assigned class N, warmer ones class R; both are now called carbon stars and lumped into class C. In between, where oxygen and carbon are in equal abundance, leftover oxygen attaches preferentially to zirconium (whose abundance is also raised above normal); the zirconium oxide stars are called class S. *See* ASTRONOMICAL SPECTROSCOPY; CARBON STAR; SPECTRAL TYPE.

Hertzsprung-Russell diagram. Shortly after the invention of the spectral sequence, H. N. Russell and E. Hertzsprung showed that luminosity correlates with spectral class. A graph of absolute visual (or bolometric) magnitude (luminosity increasing upward) plotted against spectral class (or temperature, decreasing toward the right) is called the Hertzsprung-Russell (HR) diagram (**Fig. 3**). The majority of stars lie in a band in which luminosity climbs up and to the left with temperature (with the Sun in the middle). But another band begins near the location of the Sun and proceeds up and to the right (to lower temperature), luminosities increasing to thousands solar as temperature drops to class M. To be bright and cool, such stars must have large radiating areas and radii. To distinguish between the bands, the larger stars are called giants, while those of the main band are termed dwarfs (or main-sequence stars). Most carbon stars are giants, and stars of classes R, N, and S are mixed in with those of class M. Yet brighter stars to the cool side of the main sequence, with luminosities approaching 10^6 solar, are called supergiants. In between the giants and the dwarfs lie a few subgiants. At the top, superior to the supergiants, are the very rare hypergiants. *See* DWARF STAR; GIANT STAR; SUBGIANT STAR; SUPERGIANT STAR.

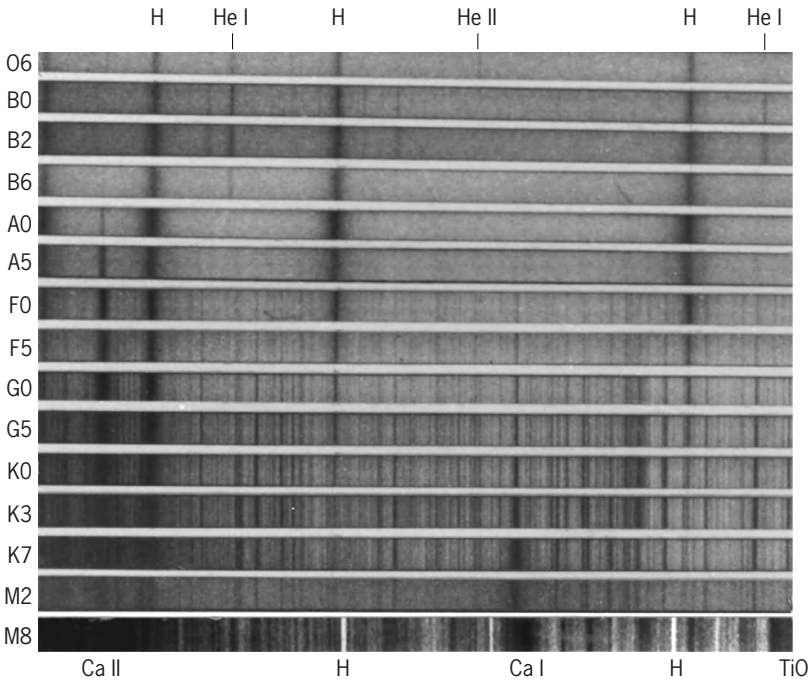

Fig. 2. Spectral sequence ranging from hot class O to cool class M. The hydrogen lines increase in strength from O to A and then weaken, disappearing in class M, where molecules take over. (*From N. Ginestet et al., Atlas de Spectres Stellaires, Observatoire Midi-Pyrénées and Observatoire de Strasbourg, 1992; Y. Yamashita et al., An Atlas of Representative Stellar Spectra, University of Tokyo Press, 1978; J. B. Kaler, Extreme Stars, Cambridge University, Press, 2001*)

White dwarf stars. In the lower left corner of the diagram, beneath the main sequence, are stars so dim that they must be very small. Since the first ones found were hot and white, they became known as white dwarfs in spite of their actual temperatures or colors. White dwarfs cannot be classified on the standard system. Most have almost pure hydrogen atmospheres caused by the downward settling of heavier helium atoms. Others have no hydrogen at all and display only helium lines. White dwarfs must therefore be positioned on the HR diagram according to

TABLE 3. Spectral classes

Class	Characteristic absorption lines	Color[a]	Temperature, K[b]
O	H, He^+, He	Blue	31,000–49,000
B	H, He	Blue-white	9800–31,000
A	H	White	7500–9800
F	Metals, H	Yellow-white	6000–7500
G	Ca^+, metals	Yellow	5300–6000
K	Ca^+, Ca	Orange	3900–5300
M	TiO, other molecules, Ca	Orange-red	2200–3900
L[c]	Metal hydrides, alkalai metals	Red	<2200
T[d]	Methane, water, ammonia bands	Infrared	1000
R[e]	Carbon	Orange	4000–5800
N[e]	Carbon molecules	Red	2000–4000
S[e]	ZrO and other molecules	Orange-red	2000–4000

[a]The color refers to energy distribution; visual colors are subtle.
[b]Main sequence.
[c]A mixture of red dwarf stars and brown dwarfs.
[d]Brown dwarfs.
[e]R and N are combined into carbon stars, class C. Classes R, N, and S all consist of giants.

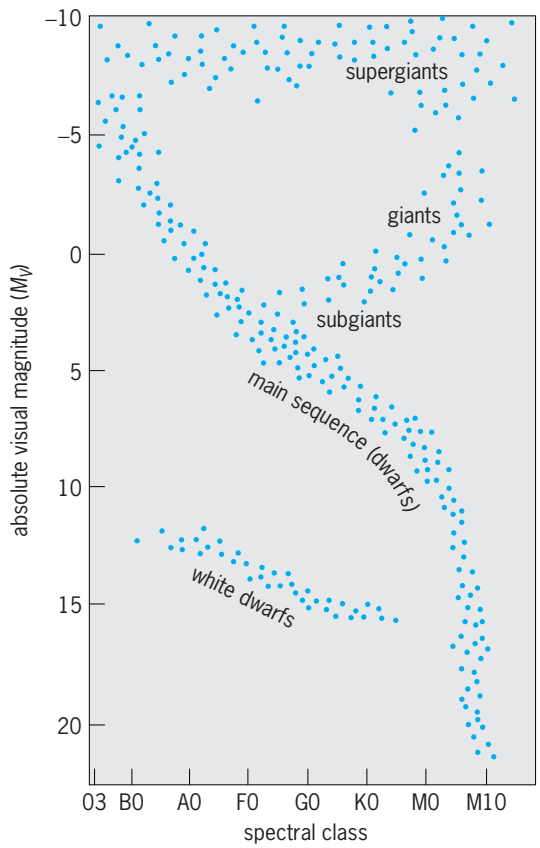

Fig. 3. Hertzsprung-Russell (HR) diagram, showing the positions of the major kinds of stars.

their temperatures (rather than their spectral classes). *See* WHITE DWARF STAR.

Subdwarfs. Stars in the galactic halo are deficient in heavy elements. Typical iron contents are a hundredth that of the Sun, and in the extreme a hundred-thousandth. Low metal content makes halo dwarfs bluer than those of the standard main sequence, shifting them to the left and seemingly downward on the HR diagram, where they are known as subdwarfs. *See* STELLAR POPULATION.

Stellar sizes. Since, for any spherical star, luminosity $L = 4\pi R^2 \sigma T^4$, the luminosity in solar units is given by Eq. (4). Temperature (T) from spectral class

$$(L/L_{\text{sun}}) = (R/R_{\text{sun}})^2 (T/T_{\text{sun}})^4 \qquad (4)$$

or spectral analysis plus the luminosity from absolute bolometric magnitude M_{bol} gives the stellar radius. Dwarf radii increase from about 10% that of the Sun in class M to nearly 15 solar in class O. Giant radii range past 10 times solar to over 1 AU, supergiants to nearly 10 AU. White dwarfs are about the size of Earth. The radii of stars are confirmed by direct measurement of angular size through interferometry and analysis of eclipsing double (binary) stars. *See* ECLIPSING VARIABLE STARS.

Spectral lines are sensitive to stellar density, so stars can be classified by size and luminosity on the basis of their spectra alone. In the standard Morgan-Keenan-Kellman (MKK) system, roman numerals I through V stand for supergiant, bright giant, giant, subgiant, and dwarf (VI is sometimes used for subdwarf). Once luminosity and absolute visual magnitude are known from a star's spectrum, its spectroscopic distance can be calculated from the magnitude equation (which must contain a correction for dimming by interstellar dust).

Stellar frequency distribution. The unaided-eye sky is filled with class A and B dwarfs, and with giants and supergiants that range from class O to M. Per unit volume of space, however, dim M and L dwarfs (none visible to the unaided eye and not counting dimmer brown dwarfs) dominate, constituting over 80% of all stars (Table 2). Of the standard sequence (discounting L and T), some 10% of stars are G dwarfs like the Sun and only 1% are class A stars. Rarest, making up only 0.0001%, are the O stars, which are so bright that they can be seen for great distances. There are about as many giants as A dwarfs, and supergiants are as rare as O stars. *See* HERTZSPRUNG-RUSSELL DIAGRAM.

Double and multiple stars. Most stars are members of some sort of community, from doubles through multiples (double-doubles, and so forth) to clusters, which themselves contain doubles. Separations between components of double stars range from thousands of astronomical units (with orbital periods of a million years) to stars that touch each other (and orbit in hours). Orbits of the more widely spaced binaries (periods of months to centuries) can be determined directly (or through interferometry). Limited orbital parameters of the more closely spaced ones are found by means of Doppler shifts caused by high orbital velocities. Some of the latter eclipse each other, the eclipses establishing orbital tilts, true orbital speeds, and stellar radii. The sum of the masses of a binary is found from the period and orbital dimension through Kepler's third law; the ratio of the masses is found by locating the center of mass of the system; and the combination of these two leads to determination of the individual masses. *See* BINARY STAR; KEPLER'S LAWS; ORBITAL MOTION.

Masses and main-sequence properties. Observations of hundreds of binary stars show that mass (M) increases upward along the main sequence from about 7.5% solar at cool class M to over 20 solar in the cooler end of class O. Extrapolation by theory, as well as through observations of clusters, suggests masses of 120 solar at the extreme hot end (class O3) of the main sequence. Fainter than the Sun, luminosity is proportional to M^3; from the Sun and brighter, luminosity goes as M^4 and then back to M^3. This mass-luminosity relation is the result of higher internal temperatures and pressures in more massive stars caused by gravitational compression. Given the Sun's energy source and that luminosity varies smoothly with mass, the main sequence must be a hydrogen-fusing mass sequence. Giants, supergiants, and white dwarfs must therefore have some other energy source. *See* MASS-LUMINOSITY RELATION.

Carbon cycle. As internal temperature climbs above the 8×10^6 K limit, hydrogen fuses to helium via

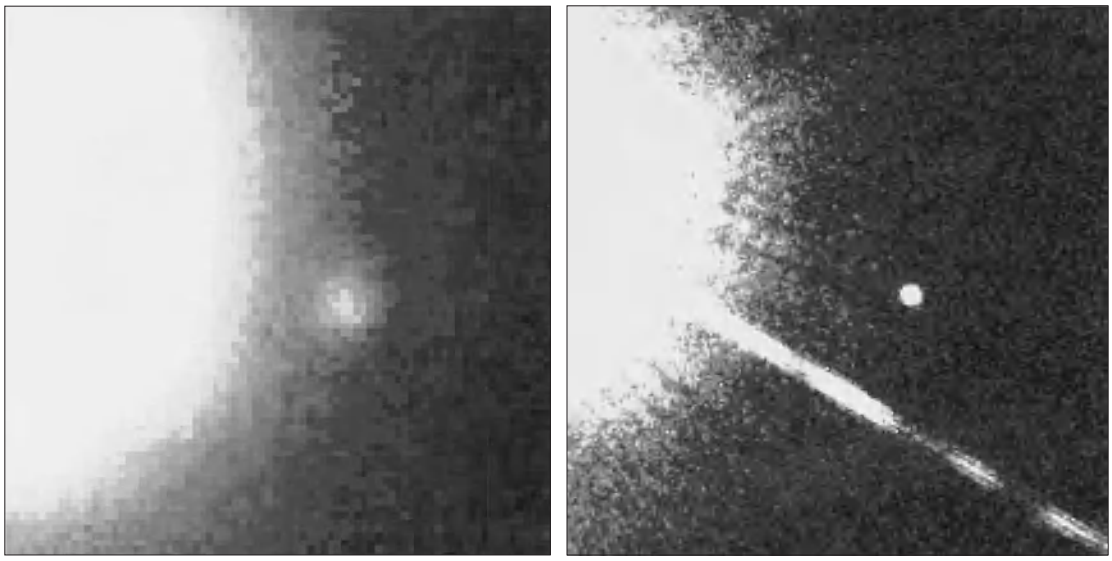

(a) (b)

Fig. 4. Images of the brown dwarf Gliese 229B, a companion to a much brighter M dwarf. (*a*) *Discovery* image, taken at Palomar Observatory in California (*T. Nakajima and S. Kulkarni*). (*b*) Hubble Space Telescope image (*S. Durrance and T. Golimowski; NASA*).

the proton-proton cycle at an ever-increasing rate. Above about 15×10^6 K (27×10^6°F), so does fusion by the carbon cycle, in which carbon acts as a nuclear catalyst:

$$^{12}C + {}^{1}H \rightarrow {}^{13}N + \text{gamma ray}$$

But ^{13}N is radioactive and decays, so the cycle proceeds:

$$^{13}N \rightarrow {}^{13}C + \text{positron} + \text{neutrino}$$
$$\text{positron} + \text{electron} \rightarrow \text{two gamma rays}$$
$$^{13}C + {}^{1}H \rightarrow {}^{14}N + \text{gamma ray}$$
$$^{14}N + {}^{1}H \rightarrow {}^{15}O + \text{gamma ray}$$

Again, ^{15}O is unstable, so the cycle concludes:

$$^{15}O \rightarrow {}^{15}N + \text{positron} + \text{neutrino}$$
$$\text{positron} + \text{electron} \rightarrow \text{two gamma rays}$$
$$^{15}N + {}^{1}H \rightarrow {}^{12}C + {}^{4}He$$

Four protons are therefore converted to an atom of helium. Together with side chains that create heavier oxygen and fluorine, the carbon cycle overtakes the proton-proton chain around class F5 (close to 1.3 solar masses), greatly shortening the star's life. *See* CARBON-NITROGEN-OXYGEN CYCLES.

Stellar convection. The onset of carbon-cycle dominance coincides with a change in stellar structure. The Sun has a radiatively stable core surrounded by an envelope whose outer parts are in a state of convection that helps produce a magnetic field and magnetic sunspots. Hotter dwarfs have shallower convection layers, and above class F, envelope convection disappears, the cores becoming convective. The convective layers of cooler dwarfs deepen, until below about 0.3 solar mass (class M3), convection takes over completely and the stars are thoroughly mixed. *See* CONVECTION (HEAT).

Brown dwarfs. Below a lower mass limit of 7.5% solar, internal temperatures and densities are not great enough to allow the proton-proton chain to operate. Such stars, called brown dwarfs, glow dimly and redly from gravitational contraction and from the fusion of natural deuterium (^{2}H) into helium (**Fig. 4**). Long sought, brown dwarfs are now being found in abundance with infrared technologies. These and the stars at the bottom of the main sequence together inhabit class L. All class T stars are brown dwarfs. The lower mass limit to brown dwarfs is unknown; they may even overlap the masses of planets. *See* BROWN DWARF; EXTRASOLAR PLANETS; INFRARED ASTRONOMY.

Stellar rotation. The Sun rotates with an equatorial velocity of 2.0 km/s (1.25 mi/s) and a period of 25 days. Rotation broadens stellar spectrum lines, allowing the observation of projected velocities which, since axial inclinations are not generally known, are lower limits to the true velocities. The assumption of random inclinations allows rotation speeds of stars in different spectral classes to be found. Main-sequence stars also exhibit magnetically induced sunspots, whose rotation in and out of sight allows accurate knowledge of rotation periods. Cooler dwarfs rotate slowly because the magnetic fields they generate couple with stellar winds and slow them down. Above the "rotation break" in class F, rotation speeds increase, reaching maximum values of 300 km/s (190 mi/s) or more in class B. Class B stars are commonly surrounded by rotation-related disks, the "Be" stars exhibiting emission lines in their spectra. Because of their large sizes, giants and supergiants rotate slowly. *See* STELLAR ROTATION.

Clusters. Doubles and multiples are highly structured. Clusters are not, the member stars orbiting a common center of mass. Open clusters are

Fig. 5. The Jewel Box open cluster (Kappa Crucis), which lies next to the Southern Cross. (© *Anglo-Australian Observatory; photograph by David Malin*)

fairly small collections in which a few hundred or a thousand stars are scattered across a few tens of light-years. Examples are the Pleiades and Hyades in Taurus, which are among thousands of such clusters occupying the Galaxy's disk (and thus the Milky Way) [**Fig. 5**]. About 150 globular clusters occupy the Galaxy's halo, the poorest about as good as a rich open cluster, the best containing over a million stars within a volume 100 light-years across (**Fig. 6**). The main-sequence stars of globular clusters are subdwarfs, and like other halo stars are deficient in metals.

The HR diagrams of clusters are radically different from the HR diagrams of the general stellar field. Those of open clusters differ among themselves in having various portions of the upper main sequence removed. Some main sequences go all the way from class M to O, while others extend only from M through G. The effect is related to the cluster's age, since high-mass stars die first. Globular clusters, which lack an upper main sequence and are therefore all old (up to about 12×10^9 years), contain a distinctive "horizontal branch" composed of modest giants.

All clusters disintegrate with time as their stars escape through a form of evaporation aided by tides raised by the Galaxy. Old open clusters (more than 10^9 years old) are rare, while the compact globulars can survive for the age of the Galaxy.

If the HR diagram of a distant cluster is constructed using apparent magnitude rather than absolute, then comparison of the cluster's main sequence with that found from local stars using parallax allows the determination of the difference between absolute and apparent magnitude for the cluster, and therefore of the cluster's distance (after correction for interstellar absorption). Knowing the cluster's distance, in turn, makes it possible to learn the absolute magnitudes of the kinds of stars that are not within parallax distance. By locating the center of the distribution of globular clusters, most of which lie outside the obscuring dust of the galactic disk, it is possible to determine the distance to the center of the Milky Way Galaxy. *See* STAR CLUSTERS.

Variable stars. Dwarfs are generally stable. Giants and supergiants, however, can have structures that allow them to pulsate. Cepheids (named after the star Delta Cephei) are F and G supergiants and bright giants that occupy a somewhat vertical instability strip in the middle of the HR diagram. They vary regularly by about a magnitude over periods ranging from 1 to 50 days. The pulsation is driven by a deep layer of gas in which helium is becoming ionized. This layer alternately traps and releases heat from the outgoing radiation. Larger and more luminous Cepheids take longer to pulsate. Once this period-luminosity relationship is calibrated through parallax and main-sequence cluster fitting, the period of a Cepheid gives its absolute magnitude, which in turn gives its distance. Cepheids are vital in finding distances to other galaxies (**Fig. 7**). *See* CEPHEIDS; HUBBLE CONSTANT.

As the instability strip crosses the class F and A giants found in globular clusters, it produces the low-metal RR Lyrae stars, which have periods of a day or less. Since they all have absolute magnitudes around 0.7, they too are valuable distance indicators. Farther down the HR diagram, the instability strip crosses the class A and F main sequence. Here and in the giants just above it are the subtly varying Delta Scuti

Fig. 6. Globular cluster 47 Tucanae, one of the grandest in the Milky Way Galaxy, containing over 1 million stars. (© *Anglo-Australian Observatory; photograph by David Malin*)

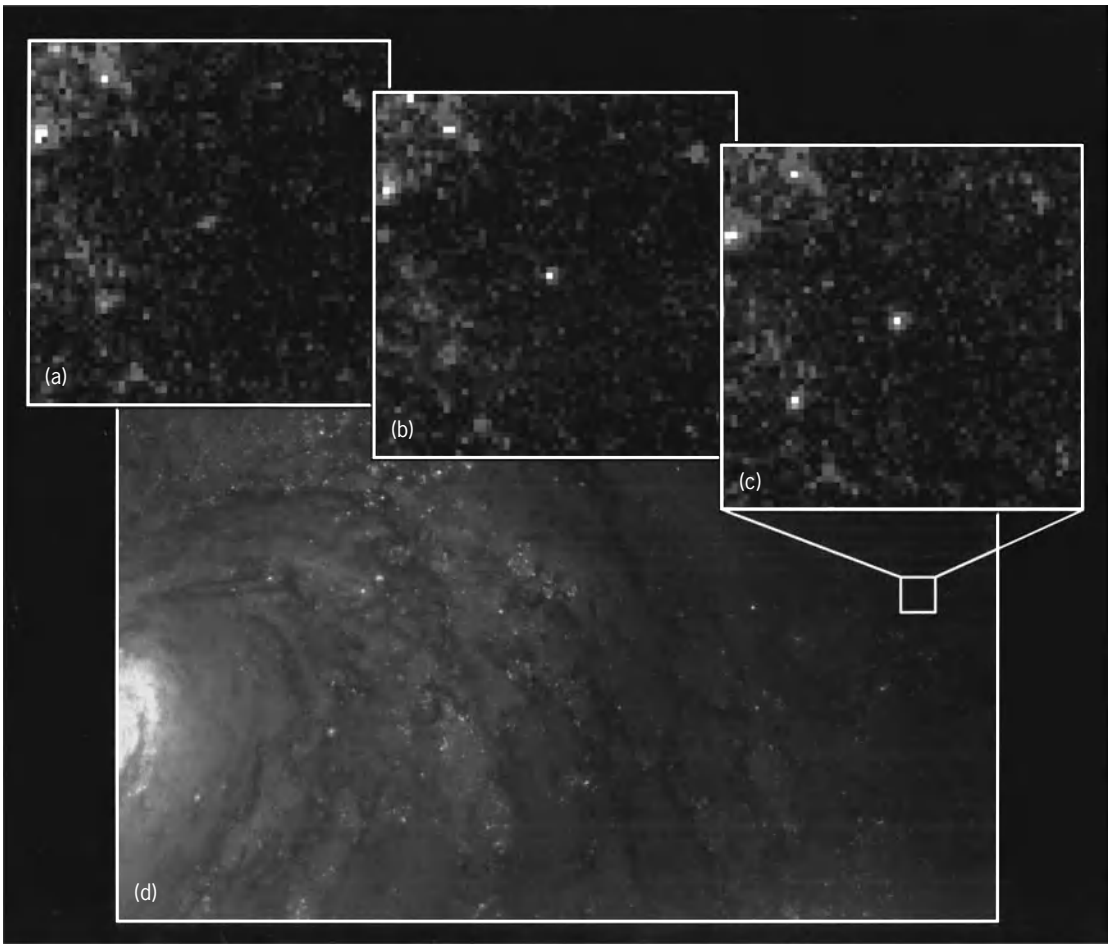

Fig. 7. A Cepheid 51 million light-years away in the spiral galaxy M100. (*a–c*) Hubble Space Telescope images of the star (in center of each image) on three different nights, showing changes in brightness. (*d*) Location of the star in one of the galaxy's spiral arms. (*Wendy L. Freedman, Observatories of the Carnegie Institution of Washington; NASA*)

stars. Where the instability strip hits the hydrogen-rich white dwarfs, at temperatures between 10,500 and 13,000 K (18,400 and 22,900°F), it produces the ZZ Ceti stars, which pulsate nonradially with periods of minutes. (Hydrogen-poor white dwarfs have their own instability zone around 22,000 K or 39,100°F.)

Miras (after the star Mira, Omicron Ceti), or long-period variables, are luminous M, C, or S giants that can vary visually by more than 10 magnitudes over periods that range from 50 to 1000 days, the pulsations again driven by deep ionization layers. Much of the visual variation is caused by small changes in temperature that shift the starlight from the visual to the infrared and that also alter the formation of obscuring titanium oxide molecules. The infrared (and bolometric) variation amounts to only two magnitudes or so. Other giants, and especially supergiants, are semiregular or even irregular variables with no periodicities at all. Still other stars, such as the class B subgiant Beta Cephei stars, are multiply periodic. *See* MIRA.

Duplicity produces its own set of intrinsic variations. If the members of a binary system are close enough together, one of them can transfer mass to the other, and instabilities in the transfer process cause the binary to flicker. If one companion is a white dwarf, infalling compressed hydrogen can erupt in a thermonuclear runaway, producing a sudden nova that can reach absolute magnitude −10. If the white dwarf gains enough matter such that it exceeds its allowed limit (the Chandrasekhar limit) of 1.4 solar masses, it can even explode as a supernova (discussed below). *See* CATACLYSMIC VARIABLE; NOVA; VARIABLE STAR.

Off the HR diagram. Various kinds of stars are not placeable on the classical HR diagram. The most common examples are the central stars of planetary nebulae, which are complex shells and rings of ionized gas that surround hot blue stars that range in temperature from 25,000 to over 200,000 K (45,000 to over 360,000°F) [**Fig. 8**]. While they can have luminosities as high as 10,000 solar, these central stars appear relatively dim to the eye, as they produce most of their radiation in the high-energy ultraviolet portion of the electromagnetic spectrum. This radiation ionizes the surrounding nebulae and causes them to glow. On the theoretician's extended HR diagram, in which luminosity is plotted directly against temperature, the planetary nebula central stars clearly connect the red giants with the white dwarfs. *See* PLANETARY NEBULA.

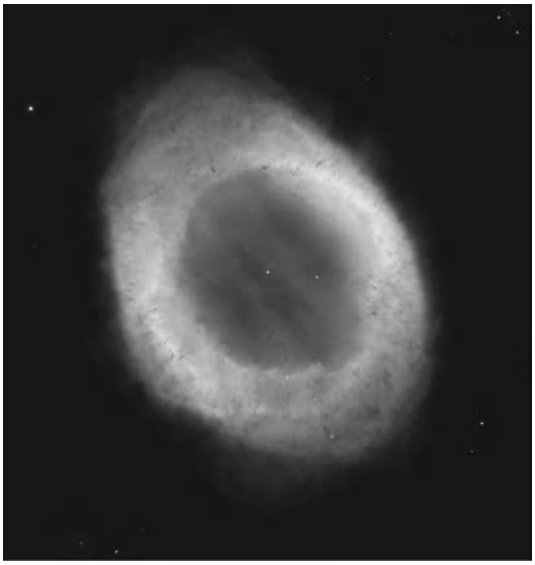

Fig. 8. Hubble Space Telescope image of the Ring Nebula in Lyra, a planetary rebula made of glowing ionized gas that surrounds a hot blue star. (*Hubble Heritage Team; Associated Universities for Research in Astronomy; Space Telescope Science Institute; NASA*)

Even hotter neutron stars (10^6 K, 1.8×10^6°F) are found associated with the exploded remains of supernovae (supernova remnants). Only 25 or so kilometers (15 mi) across, the visible ones spin rapidly, and are highly magnetic, with fields 10^{12} times the strength of Earth's field. Radiation beamed from tilted, wobbling magnetic axes can strike the Earth to produce seeming pulses of radiation (**Fig. 9**). Over time, the spin rates slow and these pulsars disappear. The remains of exploded stars, quiet pulsars (and active ones whose radiation does not hit Earth) litter the Galaxy. *See* NEUTRON STAR; PULSAR.

Evolution. The different kinds of stars can be linked through theories of stellar evolution. Stars are born by the gravitational collapse of dense knots within cold dusty molecular clouds found only in the Galaxy's disk. When the new stars are hot enough inside to initiate fusion, their contraction halts and they settle onto the main sequence. The higher the stellar mass, the greater the internal compression and temperature, and the more luminous the star. But the higher the internal temperature, the greater the rate at which hydrogen is fused, and the shorter the star's life. The fuel supply is proportional to mass, M, and the rate of fusion is proportional to the luminosity, L. Lifetime t is proportional to (fuel supply)/(rate of use) and thus to M/L. But L is on average proportional to $M^{3.5}$, so t is proportional to $1/M^{2.5}$. For example, a 100-solar-mass star will exhaust its core hydrogen in only 2.5 million years. *See* MOLECULAR CLOUD; PROTOSTAR.

Main-sequence lifetime is so long for stars under 0.8 solar mass that none has ever evolved in the history of the Galaxy. When a star between 0.8 and about 10 solar masses uses all its core hydrogen, its outer layers expand and cool to class K (**Fig. 10**). The star then brightens as it becomes a giant, lower masses brightening by larger factors and first becoming subgiants. The ascent to gianthood is terminated when the central temperature is high enough (10^8 K or 1.8×10^8 °F) to initiate the fusion of core helium via the triple-alpha process, wherein three helium nuclei (alpha particles) combine to make one carbon atom. This process temporarily stabilizes the star with a helium-burning core and a hydrogen-burning shell. (Fusion with another ^{4}He nucleus makes oxygen.)

When the helium is fused to carbon and oxygen, the core contracts, helium fusion spreads outward

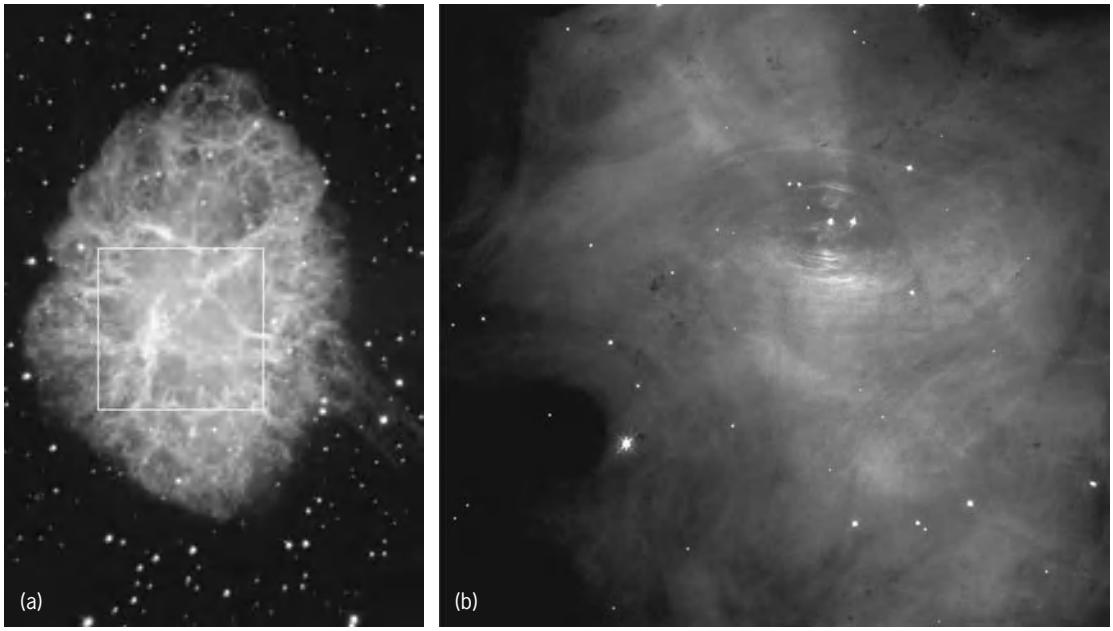

Fig. 9. Hubble Space Telescope image of the inner parts of the Crab Nebula supernova remnant. A spinning neutron star (pulsar) is the left of the two stars near the center of the image. Part *b* is blowup of boxed area in *a*. (*J. Hester and P. Scowen; NASA*)

into a shell, and the star again climbs the giant branch, the hydrogen- and helium-fusing shells sequentially turning on and off. Such asymptotic giant branch (AGB) stars become larger and brighter than before, passing into class M, where they eventually become unstable enough to pulsate as Miras. Within a certain mass range, convection in giants can bring carbon from helium fusion (as well as other new elements) to the surface, and the star can become a class S and then a carbon star. Powerful winds strip the star nearly to its fusion zone, which is protected from the outside by a low-mass hydrogen envelope. As the inner region becomes exposed, the star heats and eventually illuminates the surrounding outflowing matter to produce a planetary nebula, which dissipates into the interstellar medium, leaving a carbon-oxygen white dwarf behind.

When high-mass stars, those above about 10 solar masses, use up their core hydrogen, they too migrate to the right on the HR diagram, becoming not so much brighter but larger, cooling at their surfaces and turning into supergiants. Below about 40 solar masses they become class M red supergiants, losing huge amounts of matter through immense winds. Some stabilize there under the action of helium fusion; others loop back to become blue supergiants. Above 60 solar masses, so much mass is lost that the supergiants do not make it much past class B, stalling there as helium fusion begins. Supergiants are so massive that, whatever their outward disposition, when the internal helium is fused to carbon and oxygen, the core shrinks and heats to the point that carbon fusion can begin. The mixture converts to oxygen, neon, and magnesium, which then burns to a mixture of silicon and sulfur, and finally to iron, each fusion stage moving outward to surround the core in successive shells.

Since iron cannot fuse and produce energy, the core suddenly collapses and somehow "bounces" to explode away all the outer layers. From the outside, the observer sees the explosion as a type II supernova that can reach absolute magnitude −18. Temperatures within the exploding layers are high enough to create all the chemical elements, including iron, nickel, and uranium. The debris—the supernova remnant—expands for centuries into space (Fig. 9), its shock wave helping to compress the interstellar matter and create new stars. The remaining core is so dense that it converts to a neutron star or, if the mass is great enough, even to a black hole. *See* BLACK HOLE; CRAB NEBULA; NEBULA.

Other routes to stellar explosion derive from double stars. Mass transfer from a low-mass main-sequence star to a white dwarf companion can cause the white dwarf to exceed the Chandrasekhar limit of 1.4 solar masses (beyond which it cannot support itself by way of electron degeneracy). The resultant collapse of the white dwarf produces a hydrogenless type Ia supernova that can reach $M_v = -19$ with such reliability that such supernovae are used to measure distances to distant galaxies. Stellar evolution feeds huge quantities of enriched matter back into the star-generating clouds of interstellar space, causing the metal content of the Galaxy to increase and providing the raw materials from which the Earth is made. *See* NUCLEOSYNTHESIS; STELLAR EVOLUTION; SUPERNOVA.

James B. Kaler

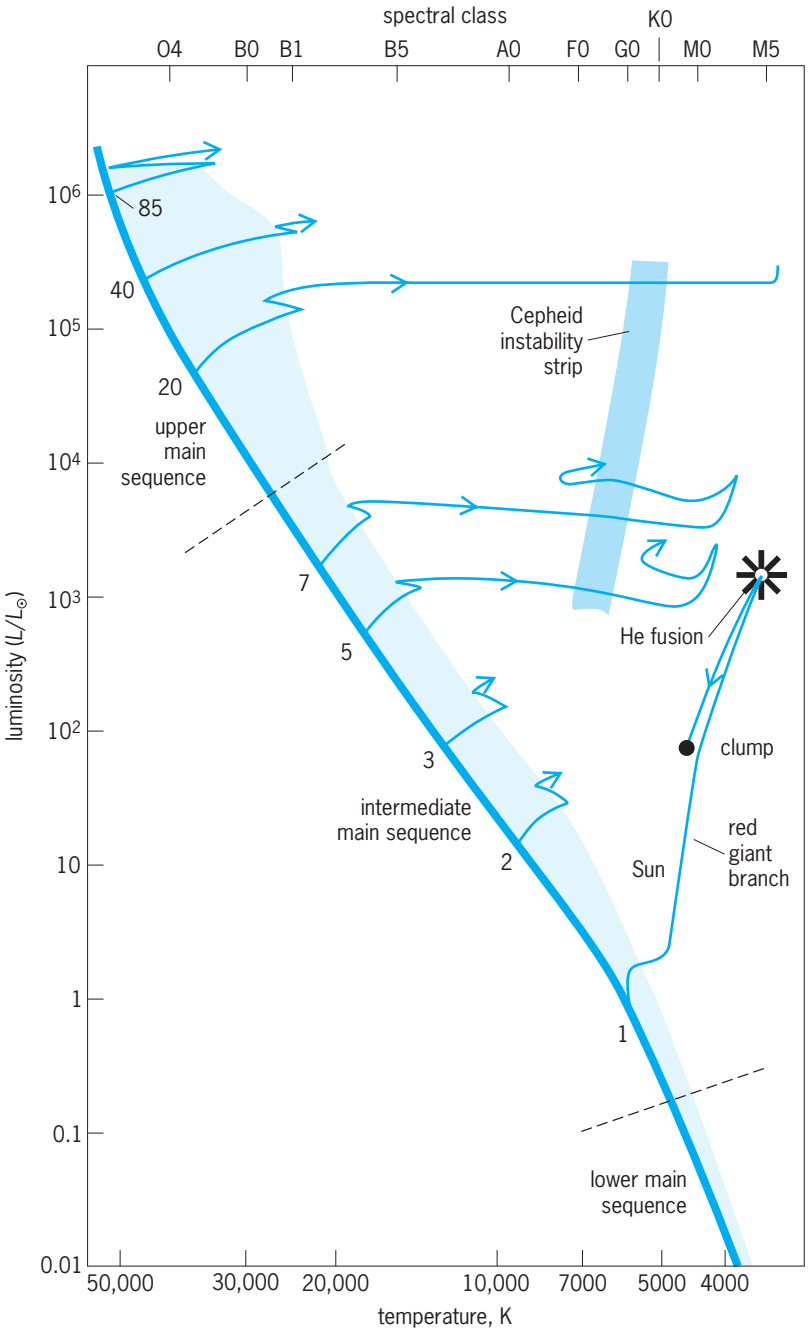

Fig. 10. Evolution of stars on the HR diagram. Solar masses of stars are the indicated numbers along the main sequence. Lower main-sequence stars have never had time to evolve. Intermediate-mass stars evolve from the main sequence to become giants that stabilize during helium fusion in a "clump." (Subsequent evolution, when helium fusion is done, is not shown here.) High-mass stars evolve to the right as supergiants and then explode. (*After J. B. Kaler, Astronomy!, Addison-Welsey-Longman, 1994, from work of I. Iben, Jr.*)

Bibliography. L. H. Aller, *Atoms, Stars, and Nebulae*, 3d ed., Cambridge University Press, 1991; E. Böhm-Vitense, *Introduction to Stellar Astrophysics*, 3 vols. Cambridge University Press, 1989, 1992; J. B. Kaler, *Extreme Stars*, Cambridge University Press, 2001; J. B. Kaler, *Stars*, Freeman, New York, 1992, paperback, revised 1998; J. B.

Kaler, *Stars and Their Spectra*, Cambridge University Press, 1989; J. B. Kaler, *The Cambridge Encyclopedia of Stars*, Cambridge University Press, 2006; J. B. Kaler, *The Hundred Greatest Stars*, Copernicus, New York, 2002; R. J. Tayler, *The Stars: Their Structure and Evolution*, 2d ed., Cambridge University Press, 1994.

Star clouds

Large groupings of stars with dimensions of 1000 to several thousand light-years. Star clouds are not stable clusterings of stars or even groups of a common origin, such as stellar associations. Instead, they are large areas where the stellar density is higher than average in a galaxy. They are nevertheless physical entities in that they result from large-scale star formation events or series of events that are to some extent limited in size by the characteristics of the galactic environment.

Large star clouds are found commonly in the spiral arms of galaxies such as the Milky Way Galaxy and the Andromeda Galaxy (M31), but are not found in very old galaxies such as elliptical galaxies M32 or NGC 3379. Conspicuous star clouds in the Milky Way Galaxy are found in the constellations Carina, Cygnus, Sagittarius, and Scutum. They are particularly spectacular when photographed with a wide-field telescope, when tens of thousands of stars crowd the photo. The densest star clouds, such as those in the constellation Sagittarius, are also spectacular when photographed with the *Hubble Space Telescope* (see **illustration**). In some galaxies, such as the giant spiral galaxy M101, they are immense, ranging up to 6000 light-years (5.7×10^{16} km or 3.6×10^{16} mi) in diameter. *See* ANDROMEDA GALAXY; MILKY WAY GALAXY.

Some astronomers believe that the evidence suggests that star groups are hierarchical, with open clusters the smallest units; they are grouped together in OB associations, and the associations are grouped in star clouds, which in turn are grouped in galaxies. This is sometimes interpreted to indicate that star formation proceeds from the large to the small scale, with the interstellar gas first breaking up into large clouds that gravitationally separate into protostar clouds, and then episodes of star formation occur within these clouds to form OB associations, the nuclei of which then can become stable star clusters. *See* GALAXY, EXTERNAL; INTERSTELLAR MATTER; MOLECULAR CLOUD; STAR; STAR CLUSTERS; STELLAR EVOLUTION.

Paul Hodge

Hubble Space Telescope **image of the center of the Sagittarius star cloud. (*Space Telescope Science Institute; NASA*)**

Bibliography. A. E. Dolphin and D. A. Hunter, Recent star formation in Shapley Constellation III in the Large Magellanic Cloud, *Astron. J.*, 116:1275–1285, 1998; Y. Efremov, Star complexes and associations, *Astron. J.*, 110:2757–2770, 1995; D. A. Gouliermis et al., The star-forming region NGC 346 in the Small Magellanic Cloud with *Hubble Space Telescope* ACS observations, I. Photometry, *Astrophys. J. Suppl.*, 166:549–556, 2006; R. C. Kennicutt, F. Bresolin, and A. E. Dolphin, Systematic properties of young clusters and associations in galaxies, in D. Geisler, E. K. Grebel, and D. Minniti (eds.), *Extragalactic Star Clusters*, IAU Symp. 207, pp. 410–420, Astronomical Society of the Pacific, San Francisco, 2002.

Star clusters

Groups of stars held together by mutual gravitational attraction. There are two basic morphological types: open clusters and globular clusters. Typical densities of field stars near the Sun are 0.1 star per cubic parsec. (One parsec equals 1.9×10^{13} mi or 3.1×10^{13} km.) Open clusters, with dozens to thousands of stars, have central densities of 0.3 to 10 stars per cubic parsec, and are often elongated or amorphous in shape (**Fig. 1**). Globular clusters, with a thousand to several million stars, and central densities of a few hundred to over 100,000 stars per cubic parsec, are generally spherical (**Fig. 2**). Associations are even looser assemblages, with a few hundred stars and central densities lower than the field. They are often recognized because of unusually large numbers of special types of stars. OB associations are dominated by hot, luminous, young O and B stars. T associations are dominated by young T Tauri variable stars. *See* SPECTRAL TYPE; T TAURI STAR.

Star clusters are important because, within each cluster, member stars probably are at the same distance from the Sun, and have the same age and the same initial chemical composition. Stars with larger masses begin life hotter and brighter than lower-mass stars and end their lives earlier. Most stars within clusters lie along a band in the luminosity-temperature plane (where temperature is measured by spectral type or color). In older clusters, this main-sequence band terminates at fainter and cooler levels. Ages derived from such observations provide a chronology of which clusters formed first. Differences observed among stars away from the main sequence,

such as in luminosity and temperature or in chemical composition, reflect changes to the stars during the ends of their lives and provide laboratories for the study of stellar evolution. Apparent luminosities of main-sequence stars may be converted into intrinsic luminosities by distance estimates obtained via geometric methods, including stellar parallaxes or stellar motions, or by photometric methods. Such distance determinations map out the locations of clusters in the Milky Way Galaxy. *See* HERTZSPRUNG-RUSSELL DIAGRAM; MILKY WAY GALAXY; PARALLAX (ASTRONOMY); STELLAR EVOLUTION.

Open clusters. Open clusters, once called galactic clusters, are found mostly along the band of Milky Way, and are rarely found more than 1000 parsecs above or below the plane. A dozen are visible to the naked eye, such as Ursa Major, the Hyades (the horns of Taurus), the Pleiades (also in Taurus), and Praesepe ("the Beehive," in Cancer). Over a 1100 are cataloged, and the Milky Way probably contains tens of thousands, most hidden by interstellar obscuration. *See* HYADES; PLEIADES.

Distances and dimensions. The nearest, the Ursa Major cluster, lies about 25 parsecs away. The Hyades and Pleiades cluster lie about 45 and 125 parsecs away, respectively. Open clusters have been found at distances as large as 15,000 parsecs (Berkeley 29). Open clusters have diameters of a few to about 20 parsecs.

Spectral characteristics. The brightest main-sequence stars in the youngest clusters are luminous, violet O stars. In slightly older clusters, such as the Pleiades, the brightest main-sequence stars are blue B stars. In the oldest clusters, the brightest main-sequence stars are yellow G stars. Post-main-sequence stars include luminous red giants. About a dozen clusters contain Cepheid variables. Measuring distances to such clusters provide the intrinsic luminosities of the Cepheids, and their detection and measurement in distant galaxies provides excellent distance estimates to those galaxies. This is a key component in measuring the rate of the universal expansion. The Pleiades and other clusters of the same age or older often contain white dwarf stellar remnants. *See* CEPHEIDS; COSMOLOGY; HUBBLE CONSTANT; WHITE DWARF STAR.

Embedded clusters. The youngest clusters may be hidden within dark dust clouds opaque to optical light. Infrared methods can reveal such young clusters, including T Tauri variables, Herbig-Haro objects, and bipolar outflows associated with them. *See* HERBIG-HARO OBJECTS; INFRARED ASTRONOMY; PROTOSTAR.

Motions. Proper motions, distances, and radial velocities provide a snapshot of the motion of a cluster. Open-cluster motions are not very dissimilar from that of the Sun and appear to confine the clusters to the plane of the Milky Way.

Ages and dissolution. Comparison of main-sequence data with stellar evolution models yields age estimates. Embedded clusters have ages of less than 10^6 years. The Pleiades is about 3×10^7 years old, and the oldest open clusters, NGC 6791 and Berkeley 17, may reach 10^{10} years. Few old clusters are known, partly because of the dissolution of clusters with time. The lower-mass stars can be said to float to the outer boundaries of a cluster and may be stripped away by passing stars, clusters, or giant molecular clouds. The oldest clusters are farther from the galactic plane, probably since they are more likely to survive in the lower densities there. *See* MOLECULAR CLOUD.

Chemistry. Most open-cluster stars have chemical compositions similar to that of the Sun, beginning life with 88% hydrogen, 10% helium, and 2% in so-called heavy elements such as oxygen, carbon, and iron. Open-cluster heavy-element abundances range from a factor of about 3 below that of the Sun to

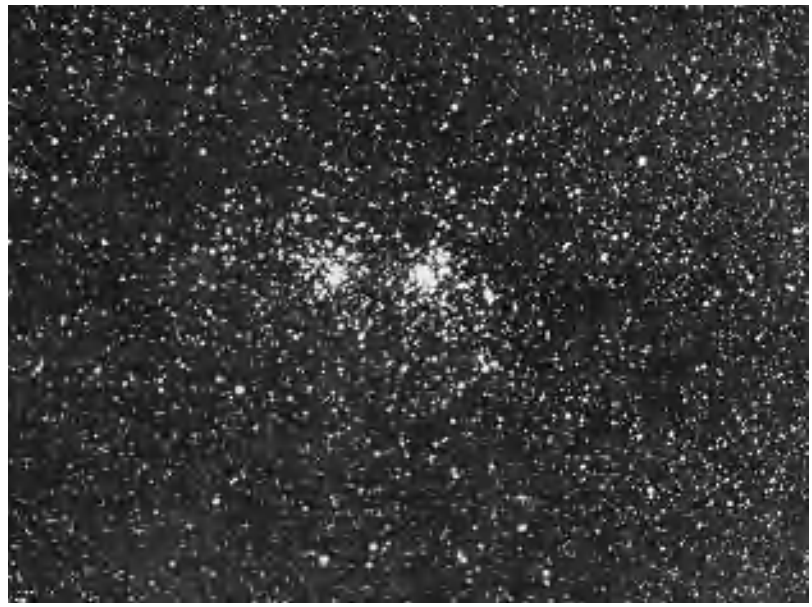

Fig. 1. Double cluster in Perseus. (*Yerkes Observatory*)

Fig. 2. Great globular cluster in Hercules, Messier 13, photographed with 200-in. (500-cm) telescope. (*California Institute of Technology, Palomar Observatory*)

about 2 times above. Open clusters in the outer part of the plane of the Milky Way Galaxy tend to have lower heavy-element abundances.

Stellar associations. These are very low density stellar ensembles of young stars, first named by V. A. Ambartsumian in 1947. About 80 associations are cataloged, 8 of which contain Cepheid variables. Associations are no longer gravitationally bound, probably because of expulsion by strong stellar winds and supernovae of large amounts (over 50%) of gas and dust out of which the stars formed. Open clusters often appear as subconcentrations within some associations. The double cluster h and χ Persei (Fig. 1) lies within the Perseus OB1 association. *See* SUPERNOVA.

Globular clusters. Generally more distant, more massive, older, and more deficient in heavy elements than open clusters, globular clusters are dispersed all over the sky, but with a strong concentration in the direction of the galactic center (Sagittarius). In 1917, H. Shapley estimated distances to globulars to infer the distance to the galactic center. A total of 153 globular clusters are known, some of which are visible to the unaided eye, such as M13 in Hercules and especially ω Centauri. One to three dozen more clusters may be awaiting discovery behind interstellar obscuration. A significant fraction of globular clusters may have been formed in dwarf galaxies that have been absorbed by the Milky Way Galaxy. The Sagittarius dwarf galaxy is in the process of being absorbed, and several globular clusters are or have been part of it. The globular cluster M54 may be the Sagittarius dwarf galaxy's nucleus. *See* GALAXY, EXTERNAL.

Distances and dimensions. The nearest globular (M4 in Scorpius) lies about 2000 parsecs away, while the most distant ones in the Milky Way Galaxy, AM-1 and Palomar 3, lie over 100,000 parsecs from the Sun. The distances are estimated by photometric means since they are too distant for geometric methods. RR Lyrae variable stars, often found in globulars, are a frequently used "standard candle" to estimate their distances. Globular-cluster diameters range from about 10 to about 100 parsecs. The typical globular cluster contains about 100,000 stars, but the range is from 1000 (AM-4) to over 3 million (ω Centauri).

Spectral characteristics. The brightest main-sequence stars in globulars are yellow G stars, signifying that all such clusters in the Galaxy are old. Red giants are the brightest stars. Some clusters contain RR Lyrae variables, and some contain a few blue stragglers, stars hotter and more luminous than the brightest and hottest normal main-sequence stars. Some globular clusters contain a type of cepheid variable, but they differ from those found in open clusters and are not frequently used as distance indicators. *See* BLUE STRAGGLER STAR; VARIABLE STAR.

Structure. Some globular clusters have low central concentrations; other have very high values, approaching 100,000 stars per cubic parsec. Some show signs of having experienced what is known as core collapse. As the stars at the outer edges evaporate from the cluster, the central regions contract, densities become extremely high, and binary star systems form. About 20% of the globular clusters have experienced core collapse. The high frequencies of blue stragglers in some cluster cores, and the presence of millisecond pulsars (17 known in nine clusters) may be related to this dynamical evolution of the cluster. Heavy stellar remnants such as neutron stars are probably also present nearer the center. Bright x-ray sources may result from binary star transfer of mass onto a neutron star. Globular clusters are unusually rich in such sources, and all are near the cluster centers. *See* ASTROPHYSICS, HIGH-ENERGY; BINARY STAR; NEUTRON STAR; PULSAR; X-RAY ASTRONOMY.

Chemistry. Heavy-element abundances in globular clusters cover a wide range, from 50% above solar to 300 times below solar. There seem to be two chemically distinct classes of globular clusters. The metal-deficient class is typically 30 times poorer in heavy elements than the Sun, while the metal-rich class is typically about 3 times more deficient. Those associated with the Sagittarius dwarf galaxy have different element-to-iron abundance patterns than most other globular clusters.

Motions. The metal-deficient globulars have very high velocities relative to the Sun, and so do not belong to the disk of the Milky Way; they are found in a spherical distribution around the galactic center. The more metal-rich globulars have motions characteristic of disk stars and clusters, and are found within a few kiloparsecs of the plane. The metal-rich clusters are all within 10,000 parsecs of the galactic center.

Ages. All globular clusters in the Milky Way are very old, typically about 1.2×10^{10} years. There may be a modest age spread of 3×10^{9} years. The globular clusters are a primary method of estimating the lower limit to the age of the galaxy and of the universe. This may be compared to the age derived by the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) mission and the expansion rate derived from studies of extragalactic Cepheids. *See* COSMIC BACKGROUND RADIATION; WILKINSON MICROWAVE ANISOTROPY PROBE.

Clusters in other galaxies. Giant associations are known in spiral galaxies (such as M31 and M33) and irregular galaxies (such as Magellanic Clouds). Large open clusters are also known in spiral and irregular galaxies throughout the Local Group. Globular clusters are known in most galaxies, including ellipticals, throughout the Local Group and in galaxies out to disances of 10^8 parsecs. The galaxy M87 has the largest number (6000 counted, 16,000 estimated). Generally all extragalactic globular clusters are old, like those in the Milky Way, although there is a class of blue populous clusters that are young, roundish, and fairly massive (10,000 stars) in some star-forming galaxies. *See* ANDROMEDA GALAXY; LOCAL GROUP; MAGELLANIC CLOUDS; STAR.

Bruce W. Carney

Bibliography. *Astronomical Almanac,* annually; W. E. Harris, Globular cluster systems in galaxies beyond the Local Group, *Annu. Rev. Astron. Astrophys.*, 29:543–579, 1991; S. R. Majewski, Galactic structure surveys and the evolution of the Milky Way, *Annu. Rev. Astron. Astrophys.*, 31:575–639, 1993; C. Payne-Gaposchkin, *Stars and Clusters*, 1979.

Star tracker

A device that automatically measures the angular separation of stellar observations with respect to a reference platform. It is also referred to as an as-trotracker.

By utilizing the star tracker in conjunction with a precise time reference (chronometer) and a dead-reckoning device consisting of gyroscopes and accelerometers (inertial navigator), a digital computer can correct many of the inertial navigator errors so that precise, autonomous (free from any radio position aids), terrestrial navigation can be achieved. The major errors corrected by the star tracker are introduced by the inertial navigator's gyroscopes that result in attitude deviations. In this configuration, called a stellar inertial system, very high precision aircraft autonomous navigation is achieved. Since the availability of radio position aids poses no problem for commercial aircraft, stellar inertial navigation technology is applied only on military vehicles. These navigation devices are used when radio position aids, such as Loran and the Global Positioning System (GPS), may be unavailable. *See* CHRONOMETER; ELECTRONIC WARFARE; INERTIAL GUIDANCE SYSTEM.

Star trackers are also utilized extensively for both military and nonmilitary applications on space probes, space-based interceptors, and satellites. In these applications the precise attitude capabilities of these devices provide the fiducial precision reference for pointing of the vehicle and Earth or planet sensors. On space missions, star trackers are the only sensor presently available that can provide arc-second attitude accuracy. *See* SPACE NAVIGATION AND GUIDANCE.

Stellar visibility. Terrestrial star trackers must be capable of making star measurements at sea level and in daylight. Cloud cover represents an impediment to terrestrial stellar observations. The probability of seeing a star at any given time on the Earth at sea level is 50%. At altitude greater than 45,000 ft (13.7 km), which is above the clouds, stellar observations are always available. This guaranteed availability of stellar observations at high altitudes has led to these devices being used on high-altitude aircraft (such as reconnaissance vehicles), ballistic missiles, satellites, and space probes.

An example of a terrestrial star tracker application is the stellar inertial navigation system of the U.S. Air Force's B-2 stealth bomber. When the B-2 is flying at altitudes above 45,000 ft, the inertial navigation system is continually updated by stellar observations. When it is flying a terrain-following, low-altitude mission, the star tracker is capable of making stellar observations 50% of the time and using these observations to correct the inertial navigation system's gyroscopic errors. Thus, the system is not dependent on any radio navigation aids which may be rendered unavailable in a wartime scenario; in particular, it is not susceptible to radio jamming of the Global Positioning System (GPS) or natural electromagnetic disturbances causing GPS outages. *See* MILITARY AIRCRAFT; SATELLITE NAVIGATION SYSTEMS.

Mounting. A gimbaled stellar inertial system is mounted on the stable element of an inertial navigator. The tracking system measures the telescope azimuth rotation angle and elevation angle to the acquired star with respect to a stable inertial reference. The tracker is programmed to observe different, widely angular separated stars in order to achieve accurate operation. The deviation of the measured stellar observations from their ideal stellar positions is utilized to enhance the performance of the inertial navigator.

The preponderance of modern systems do not contain gimbals. These systems are completely strapped down (that is, with no moving parts) and observe stars at random as the stars come into the rigid, fixed star tracker's field of view. Modern star trackers can operate in two modes. The first (lost in space) uses sophisticated star pattern matching algorithms to determine initial position of the vehicle. The second uses its knowledge of the telescope's location and direction to know which stars should be in the field of view and where their images are located on the star tracker's optical detector (**Fig. 1**). In order to make a three-axis attitude correction, there should be a least two stellar observations ideally separated by 90°. Smaller angular separations lead to a dilution of attitude correction precision. The multiple measurements need not be performed simultaneously since the star tracker corrects the attitude of each axis based on the stellar observations available. Thus, these strapdown startracker systems consist of either multiple telescopes with moderate fields of view or a single telescope with a very wide field of view for proper operation. *See* CELESTIAL NAVIGATION.

Processor functions. Digital computers, sometimes referred to as processors, perform many functions in a stellar inertial system. The computer has stored in its memory the ideal locations of the stars in a catalog called the ephemeris. Depending on the ephemeris used for the star tables, it usually stores the color-corrected star magnitude, declination, and right ascension. Since the declination and right ascension change slightly during a year, these parameters are periodically altered. In some systems the proper motions of the stars are stored in the star catalog in order to minimize the catalog updates. The catalog deliberately culls out stars that are located close to each other, such as binary stars, in order to eliminate stellar observation ambiguity. Using accurate time supplied to the computer from a chronometer, the processor indexes the ephemeris with the apparent position of the platform so that it can determine which stellar observations will be processed. *See* ALMANAC; DIGITAL COMPUTER; EPHEMERIS.

For a gimbaled star tracker, pointing and tracking commands are issued to precisely direct the telescope at a star. Since the strapdown star tracker has no gimbals to point the telescopes, the computer indicates the ideal locations of the stars that should be in the field of view of the fixed telescopes, along with their color-corrected brightness. The color correction in the ephemeris modifies the brightness

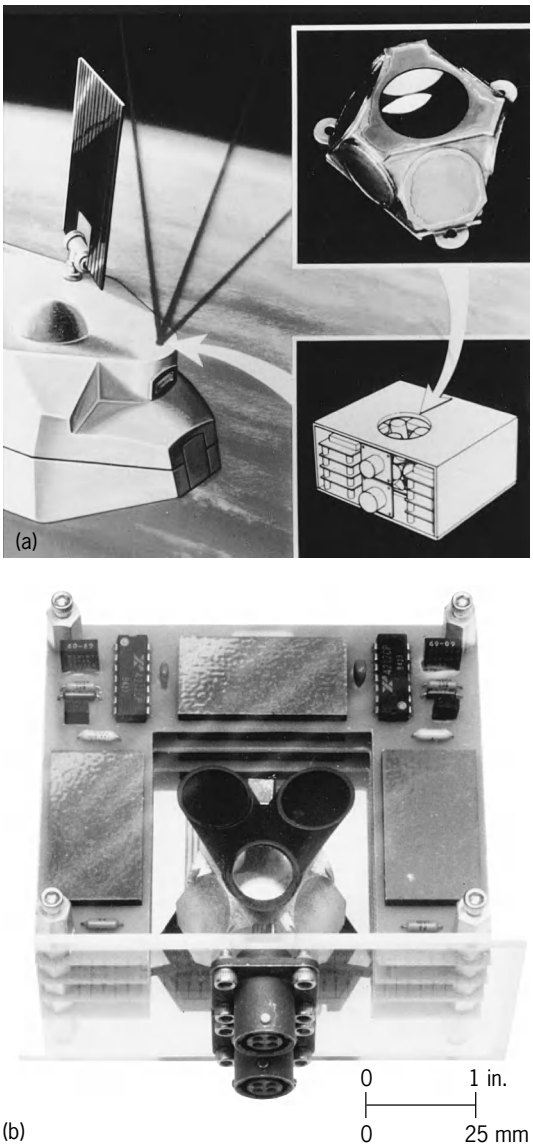

Fig. 1. Strapdown satellite star tracker. (*a*) Mounting of system on the satellite. (*b*) Star tracker in housing. (*Northrop Grumman Corp.*)

according to the responsivity of the telescope and its optical detector. Because of the lack of a pointing capability, the strapdown star-tracker ephemeris must contain as many as 10,000 or more stars, whereas the star tracker in the gimbaled, or pointing, system may contain only a table of less than 70 of the brightest stars. Thus, the strapdown star-tracker ephemeris is far more extensive than the gimbaled one, but the complexity and cost are more than offset by the reduction in cost, parts, size, and weight as well as enhanced reliability gained by the elimination of the gimbaled system's moving parts and gimbal shaft angular encoder readouts.

The computer processes the stellar data and chronometer data to arrive at the measured attitude. It then processes any known misalignments and compares the corrected measurements with the inertial estimate of attitude to provide enhanced attitude performance. Care is taken so that no star tracking is performed on stars in the vicinity of the Sun, Moon, or zodiacal light. This is done to prevent their high irradiance from corrupting the detection of the much weaker stellar irradiance signal. *See* ZODIACAL LIGHT.

Detectors. The preponderance of star-tracker photodetectors used to measure the stellar irradiance are solid-state, silicon semiconductor photosensors. These devices have their peak responsivity in the near-infrared region (0.6–1.0 micrometer). Imaging focal-plane arrays can have several million photosensors on a single semiconductor silicon chip arranged in a rectangular grid matrix (**Fig. 2**). Each of the individual photosensors on the focal-plane array is called a pixel. The devices are similar to the photosensors that have found wide utilization in home video-camera recorders. Two types of focal-plane arrays used for star trackers are charge-coupled devices (CCD) and charge-injection devices (CID). These rectangular devices are less than 0.125 in. (3 mm) thick and are approximately 0.25 in. (6 mm) per side for a 65,000-pixel sensor to 1 in. (25 mm) per side for a megapixel sensor. The advantages of the imaging photosensor can be illustrated by applying it to a standard 4° field-of-view star-tracker telescope. For a star tracker with a star magnitude 6 capability (some strapdown star trackers have the capability of measuring stars as dim as magnitude 10), on the average, there will be over three stars on the photosensor image plane. Thus, these imaging focal-plane arrays can process multiple stars with each stellar observation. The silicon focal-plane arrays are also more reliable than earlier photodetectors based on photomultipliers and image dissector electron tubes, and occupy only a small fraction of their volume. *See* CHARGE-COUPLED DEVICES; IMAGE TUBE (ASTRONOMY); MAGNITUDE (ASTRONOMY); PHOTOMULTIPLIER; TELEVISION CAMERA.

Accuracy. Star-tracker accuracy ranges from subarcsecond to 30 arcseconds. For terrestrial astro-inertial navigation, which must track stars in daylight at sea level, the predominant noise source is the sky background. The sky background noise is usually mitigated by large telescope apertures and by minimizing the field of view. In the aircraft applications of astro-inertial navigation systems, there are number of other error sources besides the star tracker accuracy, such as accelerometer noise, that degrade the system's performance. For these applications a star-tracker accuracy of 1 arcsecond is compatible with providing a latitude and longitude navigation accuracy of less than 600 ft (180 m) for mission flight times of up to 12 hours, although 1 arcsecond of the Earth's circumfereance is only about 100 ft (30 m). On the opposite end of the spectrum are the space trackers used for satellites and deep-space probes. These space-probe trackers do not contend with sky background noise when the stellar observation is sufficiently displaced from looking through a planet's atmosphere. Also, no star tracking is performed with the telescope looking in the vicinity of the Sun or Moon. This is done by selecting the location of the tracker with respect to its orbit or by providing multiple trackers to assure that the darkness of space provides the background illumination

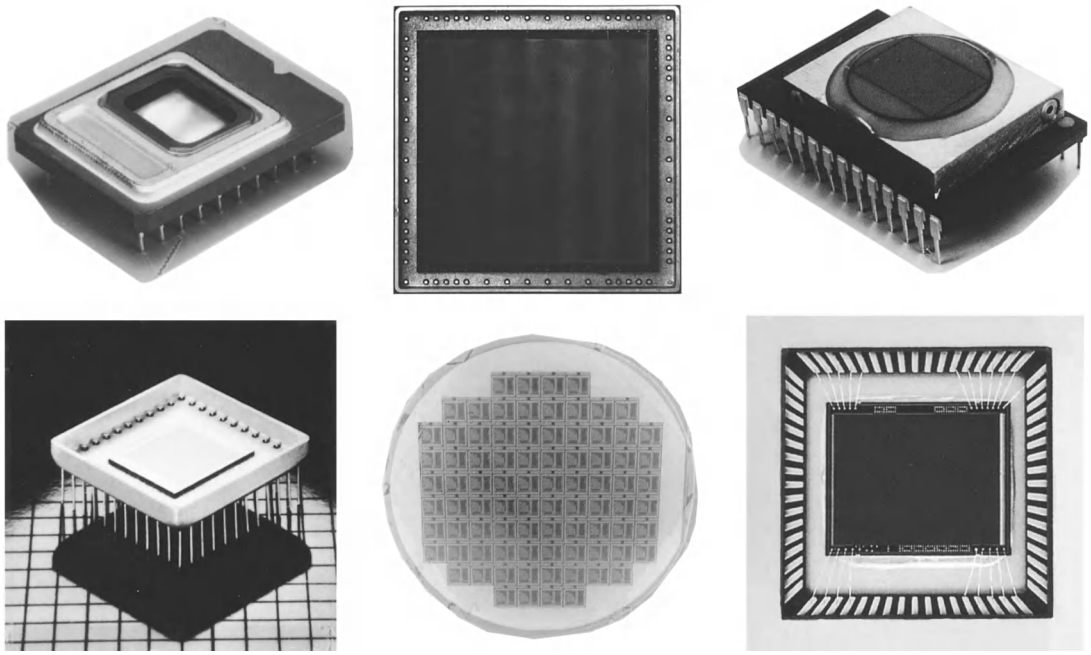

Fig. 2. Silicon focal-plane arrays. (*Northrop Grumman Corp.*)

of the telescope that is tracking stars. The accuracy of these devices ranges from 30 arcseconds for communications satellites to sub-arcsecond trackers that are used on reconnaissance satellites and space probes which track scientific targets. The high precision is required on reconnaissance satellites operating in high Earth orbits (altitudes of thousands of miles) to steer Earth-observation sensors and to report the location of terrestrial regions of interest. In the reconnaissance and mapping satellite applications, the angular accuracy of the satellite's star tracker represents a pointing location error for objects located on the surface of the Earth. For a 1-arcsecond star tracker the terrestrial location reporting errors of objects range from 30 ft (9 m) for a 1000-nautical-mile-altitude (1850-km) low-Earth-orbit (LEO) satellite to 600 ft (180 m) for a 19,325-nautical-mile-altitude (35,800-km) geosynchronous- or geostationary-orbit (GEO) satellite. *See* MILITARY SATELLITES.

Seymour Levine

Bibliography. K. Bachman, R. Dennis, and S. Levine, Strapdown astro-inertial navigation utilizing the optical wide-angle lens startracker, *Navig. J. Inst. Navig.*, 37(4):347–362, Winter 1990–1991; International Society of Optical Engineering (SPIE), *Space Guidance, Control and Tracking Conference* (vol. 1949), Orlando, Florida, April 12–14, 1993; M. Kuritsky and M. Goldstein, Inertial navigation, *Proc. IEEE*, 71:1156–1176, 1993; G. Quasius and F. McCanless, *Star Trackers and System Design*, 1966.

Starburst galaxy

A galaxy that is observed to be undergoing an unusually high rate of formation of stars. It is often defined as a galaxy that, if it continues to form stars at the currently observed rate, will exhaust its entire supply of star-forming material, the interstellar gas and dust, in a time period that is very short compared to the age of the universe. For a typical starburst galaxy, this gas and dust exhaustion time scale is less than 10^8 years, that is, less than 1% of the age of the universe. Since such a galaxy must shortly run out of star-forming material, the high star formation rate currently observed not only must end soon but also must have started relatively recently or the gas supply would have run out long ago. It follows that such galaxies must be undergoing a passing burst of star formation. The rates of star formation that are estimated to occur in starbursts can exceed 100 times the mass of the Sun per year for the most energetic bursts observed in the nearby universe. This may be compared to about 2–3 solar masses per year in the Milky Way Galaxy. However, there are large uncertainties in such estimates because there is only weak evidence on the range of masses of stars that form in starburst events. There are theoretical limits to the total rate of star formation that can occur in a starburst. One limit is derived from the mass of all the available interstellar material turning into stars during the time taken for that material to fall inward under gravity to form a typically sized galaxy. Another limit is related to the expectation that when large numbers of newly formed massive stars reach supernova stage, they will blow away much of the remaining star-forming material, thus inhibiting the burst.

The term "starburst," coined in the 1981 by Daniel W. Weedman and his collaborators, usually implies that the burst of star formation is occurring in the nuclear regions of the galaxy, the term first being used to describe a sample of luminous spiral galaxies with bright pointlike nuclei. The prototype

starburst nucleus is NGC 7714. Indeed, it has been confirmed by extensive studies that starbursts are very frequently confined to the inner approximately 1000-light-year-diameter regions of massive galaxies. However, there exist galaxies that meet the definition of short gas-exhaustion time scale but exhibit more galaxy-wide star formation. These include extragalactic H II regions, clumpy irregular galaxies, blue compact dwarf galaxies, and the nearest, best-studied starburst galaxy, M82 (**Colorplate 1**). The related term "starburst region" is also now often used to refer to large regions within galaxies where exceptionally energetic star formation is occurring.

Radiative emission. Starburst galaxies often appear blue because of the dominance of their energy output by hot massive stars. Many blue starburst galaxies were discovered in the extensive ultraviolet-sensitive objective prism surveys of B. E. Markarian. Most starburst galaxies are also strong emitters of mid- and far-infrared radiation, because the newly formed stars illuminate their natal clouds of dust and gas with ultraviolet and optical radiation, and as the dust warms, it reradiates this energy at far-infrared wavelengths. *The Infrared Astronomical Satellite* (*IRAS*), which surveyed the entire sky at far-infrared wavelengths in 1983, discovered over 100,000 galaxies that emit strongly at 60 micrometers, the majority of which are likely to be starburst galaxies. The most luminous of these galaxies are commonly termed ultraluminous infrared galaxies (ULIRGs). More detailed and more sensitive mapping in the mid-infrared with the *Infrared Space Observatory* (*ISO*), which operated from 1995 to 1998, helped to confirm that a large fraction of the newest star formation activity in starburst galaxies is hidden from view in the optical waveband by the dense, dusty molecular clouds which spawn the new stars. The *Spitzer Space Telescope* has been able to build on the foundations laid by the *IRAS* and *ISO* observatories, imaging starburst galaxies in the mid- and far-infrared with unprecedented sensitivity and detail (**Colorplate 2**). Only the slightly older young star complexes (a few million years) are clearly visible in the optical as bright blue complexes of stars and hot gas; these stars have had time to move away from their nurseries, or expel their natal clouds by stellar wind or supernova wind pressure. One of the most spectacular results from *Spitzer* is the multicolor imaging of starbursts, which show regions with intense emission from large molecules, probably polycyclic aromatic hydrocarbons (PAHs). This emission, usually colored red in the *Spitzer* false color images, dramatically highlights regions where the molecules are excited to shine by strong shocks occurring in star-forming regions. *See* INFRARED ASTRONOMY; INTERSTELLAR MATTER; MOLECULAR CLOUD; PROTOSTAR; SPITZER SPACE TELESCOPE.

Strong millimeter-wavelength emission lines of carbon monoxide, often highly concentrated to the nuclear regions, have also shown that starburst galaxies are rich in molecular gas, the dense, relatively cool material in which star formation takes place. In several galaxies there is evidence that these nuclear molecular clouds are disklike in shape, rotating around the galactic nucleus. At radio wavelengths, starburst galaxies show emission that is strongly correlated in strength with the far-infrared emission. Supernovae are common in starbursts because the most massive young stars have very short lives that end spectacularly by supernova. It is thought that the radio emission from starbursts arises from the effects of the supernova explosions on the surrounding interstellar medium. Extremely high resolution observations at radio wavelengths using the techniques of very long baseline interferometry (VLBI) have demonstrated the detection of individual young and highly luminous radio supernovae in a small number of galaxies, including one of the most energetic local ultraluminous infrared galaxies, Arp 220. In the submillimeter wavelength range, polarized emission—emission which has a preferred orientation for the light waves to oscillate in—has been found in Messier 82, thought to be caused by strong magnetic fields within the galaxy. The revealed orientation of the magnetic field lines indicates that the fields may help drag interstellar material into the center of the system. In the x-rays, starburst galaxies are weak compared to active galactic nuclei, and the x-ray emission that they do radiate is thought to come from superwinds and from binary-star radiation processes such as seen in the Milky Way Galaxy. *See* BINARY STAR; POLARIZATION OF WAVES; RADIO ASTRONOMY; STELLAR EVOLUTION; SUPERNOVA; X-RAY ASTRONOMY.

Spectrum. The existence of a starburst in a galaxy is often deduced from the presence of strong, narrow emission lines in the optical spectrum. These lines originate in hot gaseous nebulae, called H II regions, which form around the hot young stars. Stronger proof of the existence of a large population of hot young stars in a galaxy comes from the detection in the optical spectrum of broad emission lines of helium, characteristic of very massive and short-lived Wolf-Rayet stars. Over 130 Wolf-Rayet galaxies are known. In a small number of galaxies, the ultraviolet absorption features caused by hot young stars have been detected by the *International Ultraviolet Explorer* (*IUE*) satellite. Starbursts can also be recognized spectroscopically in the mid-infrared by the detection of broad emission features from PAH molecules, using the Infrared Spectrograph on board *Spitzer*. *See* ASTRONOMICAL SPECTROSCOPY; NEBULA; ULTRAVIOLET ASTRONOMY; WOLF-RAYET STAR.

Superwinds and stellar smoke. Energetic outflows of hot material have been discovered along the minor axes of many nearby starburst galaxies. Revealed in x-rays, emission lines of hydrogen, and interstellar absorption lines, these so-called superwinds are believed to be ejected by the combined energy of many supernova explosions and stellar winds in the starburst. Outflow wind velocities seem to be in the range 400–800 km/s (250–500 mi/s), and the rate of gas expulsion can be tens of solar masses per year, perhaps even exceeding the rates at which

gas is being converted into stars inside the host galaxies. These results have implications for the development over time of the heavy-element enrichment of the gas, both within galaxies and in the intergalactic medium. For example, they may help to explain the observation that galaxy spheroids of larger mass tend to have stars that are more enriched in heavy elements than the stars in lower-mass spheroids. Accompanying the hot x-ray wind in the nearest starburst galaxy, M82, *Spitzer* has discovered great plumes of cool "stellar smoke," revealed in mid-infrared images sensitive to dust emission and to the PAH molecule (Colorplate 1).

Super star clusters. In some starburst galaxies, images from the *Hubble Space Telescope* indicate the existence of super star clusters. These are knots of emission thought to be extraordinarily luminous and compact clusters of young stars, possibly young analogs of globular clusters. Some of the most beautiful examples lie along the extended tails of interacting galaxy systems, like "strings of pearls." *See* HUBBLE SPACE TELESCOPE; STAR CLUSTERS..

Causes of starburst episodes. There are several theories as to why starbursts occur. One likely cause is the interaction between two galaxies as they pass close to or collide with one another. The tidal forces generated result in shock-wave compression of the interstellar material, loss of angular momentum and infall of material into the central regions of the galaxy, and star formation in the compressed clouds. Many starburst galaxies show evidence of interactions, including distorted appearance and long wispy tails of material. Not all starbursts can be due to interactions, however, since some display no evidence for any recent disturbance. Other mechanisms that are thought responsible for high star-formation rates in galaxies are very strong spiral density waves and central bar instabilities. *See* GALAXY FORMATION AND EVOLUTION.

Starbursts in the early universe. In the local universe, starbursts are rare events that contribute less than 1% of the total amount of energy radiated by all galaxies, so they may seem exotic and interesting but not very important. However, dramatic observational breakthroughs made possible by the *Hubble Space Telescope*, the Keck Telescopes on Mauna Kea in Hawaii, the *Cosmic Background Explorer* (*COBE*) satellite, the Submillimeter Common User Array (SCUBA) on the James Clerk Maxwell Telescope in Hawaii, and the *Spitzer Space Telescope* have demonstrated that star formation rates in galaxies were probably much higher at very early times in the universe than they are now. In particular, evidence has been found that ultraluminous infrared galaxies were much more common and much more important at earlier times in the history of the universe than now. It is not known whether starbursts are responsible for the tremendous energy of these distant, young, highly infrared-luminous galaxies; if they are, then, starburst events may even have dominated the processes of galaxy building long ago, and much of the total energy radiated by early galaxies may be detectable only at infrared or submillimeter wavelengths because ultraviolet and optical light would have been very strongly obscured by the dense clouds of gas and dust involved in the starburst. *See* SUBMILLIMETER ASTRONOMY; TELESCOPE; UNIVERSE.

Relationship to other objects. Nuclei of starburst galaxies resemble active galactic nuclei in being bright and pointlike. However, active galactic nuclei are believed to differ from starbursts in harboring a central massive black hole, since they display, among other phenomena, an emission-line spectrum that is excited by a more powerful source of ultraviolet energy than can be provided by young stars. Many galaxies display the characteristics of both starbursts and active galactic nuclei, with a ring of star formation surrounding an active galactic nucleus. It is very likely that there exists a physical relationship between starbursts and the events leading to active galactic nuclei because both are frequently seen in the same galaxy, but the details of the connection remain to be fully explained. Possibilities include the following: (1) a galaxy-galaxy interaction triggers both phenomena by channeling gaseous material into the central regions of the galaxy; (2) the starburst event promotes the formation of a nuclear massive black hole; (3) the starburst event promotes the fueling of an existing, but previously quiescent, nuclear massive black hole by providing it with fuel; and (4) an active galactic nucleus turns on first, and then triggers a starburst in the surrounding interstellar medium. *See* ASTROPHYSICS, HIGH-ENERGY; BLACK HOLE; GALAXY, EXTERNAL.

Carol J. Lonsdale

Bibliography. B. Brandl et al., The mid-infrared properties of starburst galaxies from *Spitzer*-IRS spectroscopy *Astrophys. J.*, 653:1129–1144, 2006; R. Genzel, D. Lutz, and L. Tacconi, Star formation triggered by galaxy collisions, *Nature*, 395:859–862, 1998; J. S. Greaves, Magnetic field surrounding the starburst nucleus of the galaxy M82 from polarized dust emission, *Nature*, 404:732–733, 2000; T. Jarrett et al., Remarkable disk and off-nuclear starburst activity in the Tadpole Galaxy as revealed by the *Spitzer Space Telescope, Astron. J.*, 131:261–281, 2006; D. W. Weedman et al, NGC 7714—The prototype star-burst galactic nucleus, *Astrophys. J.*, 248:105–112, 1981.

Starch

A carbohydrate that occurs as discrete, partially crystalline granules in the seeds, roots (tubers), stems (pith), leaves, fruits, and pollen grains of higher plants. Starch functions as the main storage or reserve form of carbohydrate; it is second in abundance only to cellulose, a major structural component of plants. Cereal grains, tuber and root crops, and legumes (seeds) have long been used as major sources of carbohydrate in human diets. *See* CELLULOSE.

Starch is isolated commercially from the following sources: cereal grain seeds [maize (corn), wheat, rice, sorghum], roots and tubers [potato, sweet

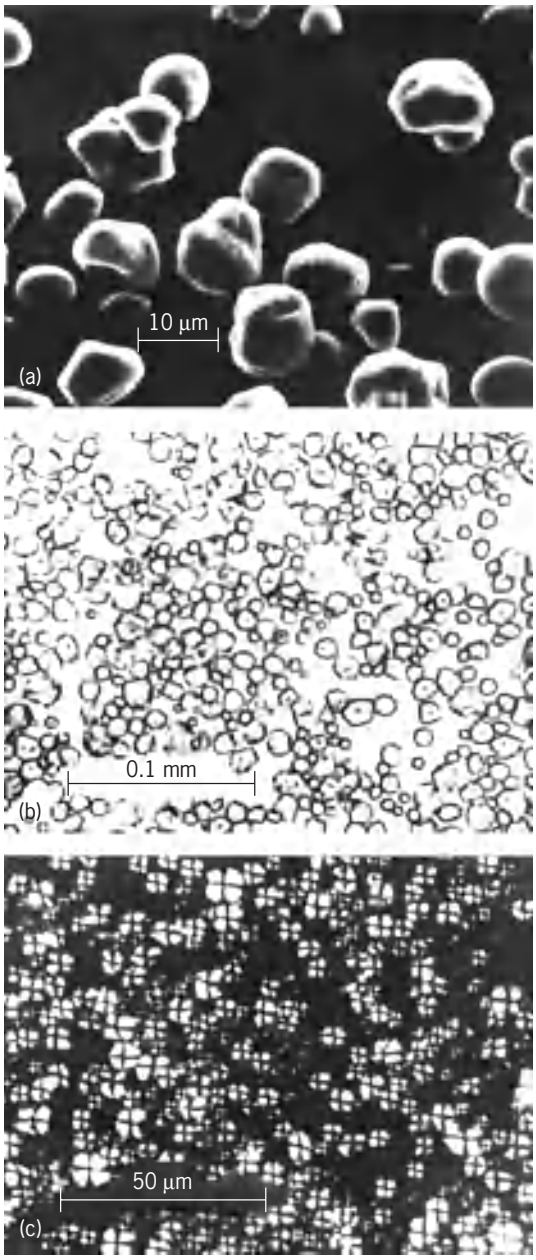

Fig. 1. Cornstarch as seen by (*a*) scanning electron microscope, (*b*) light microscope, and (*c*) polarized light. (*From R. L.* Whistler, *E. F. Paschall and J. N. BeMiller, eds., Starch: Chemistry and Technology, 2d ed., Academic, 1984*)

potato, tapioca (cassava), arrowroot], and stems and pith (sago). Cereal grains are often steeped first to loosen the starch granules in the endosperm matrix, followed by wet grinding or milling. Roots and tubers are ground to give a suspension containing starch granules. Then follows screening or sieving, washing, centrifuging, dewatering, and drying.

Characteristics. Starch, normally a white powder of 98–99.5% purity, is insoluble in cold water, ethanol, and most common solvents. Starch is stained a blue to reddish-purple color, depending on the amylose content, by dilute aqueous iodine.

The size and shape of starch granules depend on the botanical source of the starch. Rice starch, a small-granule starch, has polygonal granules ranging in size from 3 to 5 micrometers, while potato starch, a large-granule starch, has ellipsoidal or oval granules from 15 to 120 μm. Normal cornstarch has spherical or polygonal granules from 5 to 30 μm (**Fig. 1**). When viewed through a microscope under polarized light, starch granules are birefringent and show a characteristic "maltese cross" (Fig. 1*c*). Starch granules have characteristic x-ray diffraction patterns and have been estimated to be 20–50% crystalline.

Structure and architecture. Starch, a polymer of glucose, is an alpha-glucan, predominantly containing alpha-1,4-glucosidic linkages with a relatively small amount of alpha-1,6-glucosidic linkages forming branch points (**Fig. 2**). Two major polymeric components are present: amylose and amylopectin. *See* GLUCOSE.

Amylose, long thought to be a linear polymer, contains a limited amount of long-chain branching. Normal corn amylose contains about 900–1000 glucose units and appears to be a mixture of nearly equal numbers of large branched (5.3 chains on average) molecules and small unbranched molecules. In many respects amylose behaves like a linear polymer. In aqueous solution, it complexes with iodine to yield a deep blue color and associates with itself (retrogrades) to form precipitates or gels.

Amylopectin is a high-molecular-weight (10^7–10^8) branched polymer with an average chain length of 20–26 glucose units. The structure is composed of clusters of branch points with individual chains in each cluster containing 12–16 glucose units. Some longer chains extend through more than one cluster. Association of chains in amylopectin clusters forms the crystalline regions of the granule. The branch points are located in the less organized, amorphous regions of the granule.

The amounts of amylose and amylopectin vary with the source of the starch. Common starches

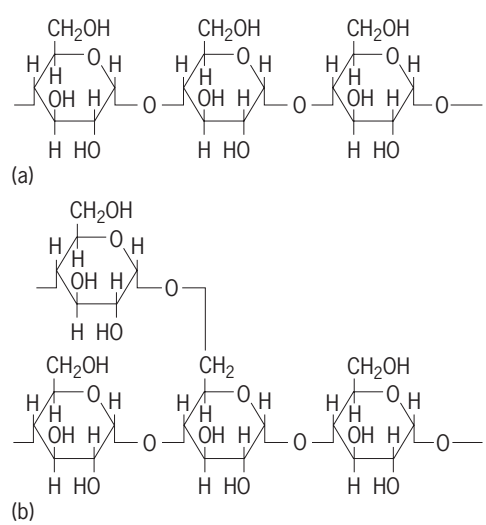

Fig. 2. Structures involved in starch. (*a*) Linear chain of alpha-1,4-glucan. (*b*) Alpha-1,6 branch point. (*After Y. Pomerzanz, ed., Wheat: Chemistry and Technology, 2d ed., American Association of Cereal Chemists, 1971*)

contain 15–30% amylose, the remainder being amylopectin. Special hybrids, with altered ratios of amylose to amylopectin, have been produced. Waxy starches from maize (commercially available), barley, rice, and sorghum contain only amylopectin. High-amylose maize starches containing 50–70% amylose are available.

Gelatinization. As a dilute aqueous suspension of starch is heated above 50°C (122°F), swelling of the granule begins in the less organized, amorphous regions with loss of birefringence and crystallinity. This is known as gelatinization and occurs over an interval of about 10°C (18°F) within the range of 55–75°C (131–167°F) for most common starches. As the temperature increases, irreversible swelling occurs, the crystalline areas undergo disorganization, and amylose leaches from the granule. When a sufficiently high temperature is reached (cooking, pasting), granule structure is destroyed, with apparent dissolution of the amylose and amylopectin components. The increase in viscosity occurring during gelatinization is an important property of starch, resulting in its common use as a thickener in foods.

Use in food systems. Starches are involved in important roles in foods, either naturally occurring in an ingredient or added to achieve a desired functional characteristic. Often the desired functional characteristic (thickening; gelling; adhesive; binding; improving acid, heat, and shear stability) cannot be achieved by using a native starch. Starches may be altered physically, chemically, or enzymatically to produce modified starches with improved functional properties.

Modified starches, commonly known as starch derivatives, are controlled in the United States by Food and Drug Administration (FDA) regulations. Modification of starch may involve a change in physical form (pregelatinization), controlled degradation (dextrinization, oxidation, acid or enzyme modification) or chemical reaction (oxidation, substitution, cross-linking). Most modifications are accomplished in aqueous suspensions under conditions in which the granule is not destroyed. Often, only minor modification is required to generate significant changes in starch properties. Modified starches can be produced from any starch, but the common commercial sources in the United States are normal maize, waxy maize, tapioca, and potato starches. *See* FOOD.

Starches are modified for the following reasons: to improve the resistance of swollen granules to acid, heat, and mechanical processing conditions; to eliminate unwanted gelling characteristics; to reduce or increase starch gelatinization temperature; to generate freeze–thaw stability; and to maximize, reduce, or inhibit viscosity.

Modified starches that produce solutions of lower viscosity are obtained by (1) treating aqueous suspensions of starch with dilute acid below the gelatinization temperature (acid-thinned starches); (2) treating a gelatinized starch slurry with an alpha-amylase (enzymatically modified starch), commercially accomplished in a continuous process; (3) heating (roasting) dry starch with small amounts of acid (dextrinization) to produce dextrins (pyrodextrins, British gums, white dextrins, yellow dextrins) that often have a pale yellow to tan color; or (4) oxidizing starch in an aqueous suspension, which results in the introduction of carboxyl and carbonyl groups and partial depolymerization of the starch. Each of these processes produces starch derivatives with the desired altered properties, for example, viscosity, paste clarity, or gel-forming ability, for different uses.

Cross-linked starch derivatives are produced by treating an aqueous suspension of starch granules with cross-linking reagents. The resulting modified starches have improved cooking characteristics (higher gelatinization temperatures); increased stability to acid, heat, or shear, with resulting control of viscosity during processing; and improved textural characteristics of starch pastes. Starch ethers or esters (substituted starches) are formed by treating an aqueous suspension of starch with appropriate reagents in the presence of an alkaline catalyst. These stabilized starch derivatives have freeze–thaw stability and a reduced tendency to undergo syneresis (formation of water on the surface of a gel or paste due to retrogradation).

Glucose (dextrose), glucose syrups, maltodextrins, maltose syrups, and high-fructose corn syrup are obtained from starch by using acid or enzymes. Enzymes are used industrially with increasing frequency because of the mild conditions required for processing and the range of products that can be obtained. High-fructose corn syrup finds extensive use as sweetener, particularly in soft drinks. *See* ENZYME; FOOD MANUFACTURING; FRUCTOSE; MALTOSE.

Nonfood uses. The paper, textile, adhesive, chemical, pharmaceutical, and polymer industries use starch and starch derivatives. Organic acids and organic solvents for use as chemical intermediates, enzymes, hormones, antibiotics, and vaccines are industrially produced from starch. *See* CARBOHYDRATE.

David R. Lineback

Bibliography. T. L. Barsby, A. M. Donald, and P. J. Frazier, *Starch: Advances in Structure and Function*, 2002; A.-C. Eliasson (ed.), *Starch in Food*, 2004; M. N. Sivak, J. Preiss, and S. Taylor (eds.), *Advances in Food and Nutrition Research*, vol. 14: *Starch—Basic Science to Biotechnology*, 1998; R. L. Whistler, E. F. Paschall, and J. N. BeMiller (eds.), *Starch: Chemistry and Technology*, 3d ed., 2001; O. B. Wurzburg, *Modified Starches: Properties and Uses*, 1986.

Stark effect

The effect of an electric field on spectrum lines. The electric field may be externally applied, but in many cases it is an internal field caused by the presence of neighboring ions or atoms in a gas, liquid, or solid. Discovered in 1913 by J. Stark, the effect is most easily studied in the spectra of hydrogen and helium, by observing the light from the cathode dark space of an electric discharge. Because of the large potential

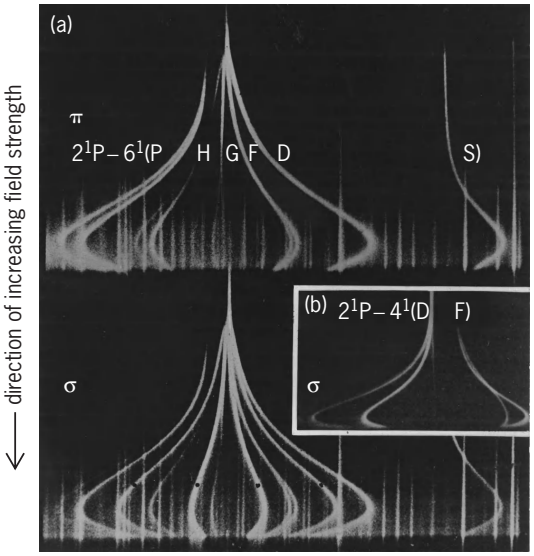

Fig. 1. Stark effect for helium lines. (*a*) The 414.4- and 416.9-nm lines at field strengths 0–85,000 V/cm. Note the disappearance of ^{1}P–^{1}H line at a certain field strength, and symmetry of the s components (those with polarization perpendicular to the field), indicated by dots. (*b*) The 492.2-nm line at 0–40,000 V/cm. Note the crossing-over of two ^{1}P–^{1}F components. (*J. S. Foster and C. Foster, McGill University*)

drop across the region, the lines are split into several components. For observation perpendicular to the field, the light of these components is linearly polarized. The splitting can be easily resolved with a spectrograph of moderate dispersion, amounting in the case of the H_γ line at 434 nanometers to a total spread of 3.2 nm for a field of 10,400 kV/m.

Linear Stark effect. This effect exhibits large, nearly symmetrical patterns. Examples are shown in **Fig. 1**, where π and σ refer to the two states of polarization. The interpretation of the linear Stark effect was one of the first successes of the quantum theory. According to this theory, the effect of the electric field on the electron orbit is to split each energy level of the principal quantum number n into $2n - 1$ equidistant levels of separation proportional to the field strength. Thus the higher members of a series show a larger number of components and greater overall splittings. The criterion for the occurrence of the linear Stark effect is that the splitting of the levels shall be large compared to the natural separation of levels of the same n but different L, where L is the quantum number of orbital angular momentum. *See* QUANTUM NUMBERS; ZEEMAN EFFECT.

Quadratic Stark effect. This occurs in lines resulting from the lower energy states of many-electron atoms. Here the large separation of states of different L results from the penetration of the valence electrons into the core of other electrons, with the result that the permanent dipole moment associated with hydrogenlike orbits no longer exists. There is, however, a small induced dipole moment due to polarization of the atom. This moment is proportional to the electric field strength, and since the energy change is proportional to the product of the dipole moment and the field strength, the energy levels shift by an amount depending on the square of the field. Thus all levels have shifts of the same sign and therefore are displaced to lower energies. Each field-free level is also split, as a result of the space quantization of the angular momentum vector J, so that there are $J + 1$ components if J is integral, or $J + {}^1/_2$ components if it is half-integral. Because the lower levels are usually less displaced than the higher ones, the quadratic Stark effect ordinarily shows itself as a shift of the lines toward the red end of the spectrum, with an accompanying separation into several components. The quadratic Stark effect is basic to the explanation of the formation of molecules from atoms, of dielectric constants, and of the broadening of spectral lines.

Intermolecular Stark effect. Produced by the action of the electric field from surrounding atoms or ions on the emitting atom, the intermolecular effect causes a shifting and broadening of spectrum lines. The molecules being in motion, these fields are inhomogeneous in space and also in time. Hence the line is not split into resolved components but is merely widened. Particularly in the electric discharge through gases with high currents, the large ion density may cause very wide lines. The amount of the broadening is found to run parallel to the sensitivity of the line to the Stark effect and thus is greatest for those lines susceptible to the linear effect.

Inverse Stark effect. This is the effect as observed with absorption lines. It has been detected, for example, by applying an electric field to potassium vapor and measuring a small displacement of the absorption lines toward the red. The displacements are found to be proportional to the square of the field strength, as in the quadratic Stark effect. In addition, certain lines of large n, which are normally forbidden by the selection rule for L, are observed to appear in the presence of the field. This type of transition is said to be produced by forced dipole radiation. Such forbidden lines also may be obtained in emission. In Fig. 1, the lines 2^1P–6^1S and 2^1P–4^1F are examples of forced dipole transitions. F. A Jenkins; W. W. Watson

Quantum-confined Stark effect. This is the Stark effect observed in structures in which the hydrogenic system is confined in a layer of thickness much less than its normal diameter. This is not practical with atoms, but the effect is observed with excitons in semiconductor quantum-well heterostructures. It is important that quantum-confined Stark shifts can be much larger than the binding energy of the hydrogenic system. The resulting shifts of the exciton optical absorption lines can be used to make optical beam modulators and self-electrooptic-effect optical switching devices.

Wannier excitons behave like very large hydrogen atoms; for example, in gallium arsenide (GaAs) their diameter is approximately 30 nm, and their binding energy is approximately 4 meV. In a quantum well, consisting, for example, of a 10-nm gallium arsenide well layer between aluminum gallium arsenide (AlGaAs) barrier layers, the exciton is confined primarily in the gallium arsenide layer, becoming overall

smaller. This smaller size, incidentally, gives the exciton a larger binding energy (for example, 10 meV), and in contrast to most semiconductors, exciton absorption resonances can be seen clearly even at room temperature in many quantum-well structures. When an electric field is applied perpendicular to the layer planes, the electron and hole in the exciton are polarized, being displaced toward opposite sides of the well. This polarization by the field gives essentially a quadratic Stark shift of the ground-state exciton in a symmetric quantum well. Such Stark shifts in quantum wells can be much larger than the binding energy (typically 40 meV).

In normal semiconductors with unconfined excitons, some quadratic Stark shifts can also be seen, but these are limited to a small fraction (10%) of the binding energy because of line broadening resulting from field ionization. Fields so large that they correspond to a potential drop of more than one binding energy over the exciton diameter (for example, about 100 kV/m) field-ionize the unconfined exciton very rapidly, pulling the electron and hole away from one another in much less than a classical orbit time. Because the particle lifetime is so short, the optical absorption spectrum shows primarily strong broadening of the exciton lines, according to the Heisenberg uncertainty principle. With the confined exciton, however, the barriers prevent the exciton from being rapidly field-ionized, even at fields of about 10 MV/m. Hence the exciton absorption lines in the quantum-confined Stark effect are not strongly broadened, even though the Stark shifts in such large fields can be many times the binding energy of the exciton (**Fig. 2**). The exciton absorption lines in Fig. 2 correspond to the creation of excitons in their ground state, and the shifts of the absorption line correspond to the (quadratic) quantum-confined Stark shift of this exciton ground state. The absorption peak height is reduced with field because there is less spatial overlap of the electron and hole as they are pulled apart, lowering the absorption strength of the transition.

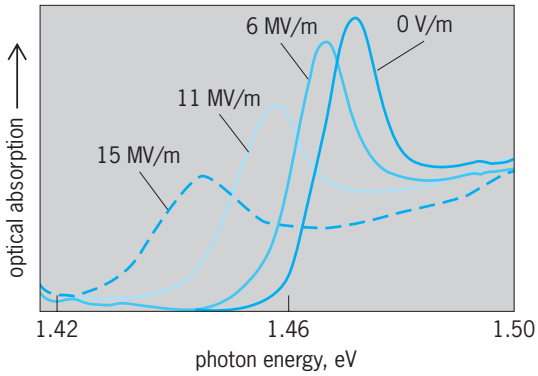

Fig. 2. Optical absorption spectra of 9.4-nm gallium arsenide (GaAs) quantum wells for different fields perpendicular to the layers, showing the quantum-confined Stark effect shift of the exciton absorption peak. (*After D. A. B. Miller, J. S. Weiner, and D. S. Chemla, Electric-field dependence of linear optical properties in quantum will structures: Waveguide electroabsorption and sum rules, IEEE J. Quant. Electron., QE-22:1816–1830, 1986*)

A typical structure for observing the quantum-confined Stark effect would contain about 50 gallium arsenide quantum-well layers of total thickness of 1 micrometer within a semiconductor diode structure. Reverse-biasing the diode with about 10 V applies a 10-MV/m electric field; this can change the transmitted power of a light beam of appropriate wavelength (for example, 850 nm) by a factor of 2 to make an optical modulator. This offers a method for modulating light beams with very low operating energy in small devices at room temperature. *See* ARTIFICIALLY LAYERED STRUCTURES; ELECTROOPTICS; EXCITON; OPTICAL MODULATORS; SEMICONDUCTOR HETEROSTRUCTURES. David A. B. Miller

Bibliography. B. H. Bransden and C. J. Joachain, *Physics of Atoms and Molecules*, 2d ed., 2003; E. U. Condon and H. Odabasi, *Atomic Structure*, 1980; R. D. Cowan, *The Theory of Atomic Structure and Spectra*, 1981; H. Haug (ed.), *Nonlinear Optical Properties of Semiconductors*, 1988; P. H. Heckman and E. Trabert, *Introduction to the Spectroscopy of Atoms*, 1989.

Static

A hissing, crackling, or other sudden sharp noise that tends to interfere with the reception, utilization, or enjoyment of desired signals or sounds. Perhaps the commonest form of static is that heard in ordinary broadcast receivers during electrical storms. Interference in radio receivers caused by improperly operating electric devices in the vicinity is sometimes also called static. *See* SFERICS.

The crackling sounds heard when long-playing plastic phonograph records are played are also called static. These sounds are caused by sudden deflection of the phonograph needle by dust particles, which are attracted to the grooves of the record by surface electric charges that are caused by friction on dry days. Static appears as momentary white specks in a television picture. *See* ELECTRICAL INTERFERENCE; ELECTRICAL NOISE. John Markus

Static electricity

Electric charge at rest, generally produced by friction or electrostatic induction. In the sixth century B.C., Thales of Miletus, who is usually credited with the discovery of static electricity, described experiments in which rubbed amber induced movement in nearby particles. It was not until the eighteenth century, when Benjamin Franklin carried out his kite-flying experiments in which he "captured" some thunderstorm electricity in a Leyden jar, that it was verified that thunderstorm electricity and static electricity were one and the same. *See* ELECTRIC CHARGE.

Triboelectrification is the process whereby charge transfer between dissimilar materials, at least one of which must have a high electrical resistivity, occurs due to rubbing or mere contact. Contact and

separation of materials occur during industrial operations such as powder processing and the manufacture of plastic and other materials in sheet form. *See* ELECTRICAL RESISTIVITY.

Most organic and polymeric materials have volume resistivities greater than 10^8 ohms per meter and retain charge for periods of many hours. Triboelectric charging is not fully understood, and the magnitude of charging may be strongly influenced by surface contamination of the contacting materials, relative humidity of the atmosphere, and the energy of rubbing. For example, during the grinding of material during powder manufacture, charge levels of about a microcoulomb per kilogram are typical. This is a relatively low level of charge. The more energetic process of pneumatically conveying a powder along a pipe may give rise to charge levels a thousand times greater and an electrostatic hazard may arise. Hazards are due to sparks to ground after a relatively large amount of static charge has accumulated on a body that is improperly grounded. *See* ELECTRIC INSULATOR; ELECTRIC SPARK.

In modern industry, highly insulating synthetic materials, such as plastic powders and insulating liquids, are used in large quantities in an ever increasing number of applications. Such materials charge up readily, and large quantities of electrical energy may develop with an attendant risk of incendiary discharges. When, for example, powder is pneumatically transported along pipes, charge levels of up to about 100 microcoulombs per kilogram can develop and potentials of thousands of volts are generated within powder layers and the powder cloud. Energetic sparking from charged powder may initiate an explosion of the powder cloud. Similar problems occur when insulating liquids, such as certain fuels, are pumped along pipes, and it is essential that strict grounding procedures are followed during the refueling of aircraft, ships, and other large vehicles.

The capacity of a person for retaining charge depends upon stature, but is typically about 150 picofarads. Even the simple operations of removing items of clothing or sliding off a chair can lead to body discharges to ground of about 0.1 μC, which are energetic enough to ignite a mixture of natural gas and air. Human body capacitance is sufficiently high that, if poorly conducting shoes are worn, body potential may rise to 15,000 V or so above ground during industrial operations such as emptying bags of powder. Sparking may then occur with energy exceeding the minimum ignition energy of powder or fumes, so initiating a fire or explosion. Conducting footware should be used to prevent charge accumulation on personnel in industrial situations where triboelectrification may occur. *See* CAPACITANCE.

In the microelectronics industry, extremely low-energy discharges, arising from body potentials of only a few tens of volts, can damage microelectronics systems or corrupt computer data. During the handling of some sensitive semiconductor devices, it is imperative that operators work on metallic grounded surfaces and are themselves permanently attached to ground by conducting wrist straps. Whenever films of insulating material are wound over rollers, such as in the photographic industry, surface charging and subsequent discharging occur. The discharges observed from charged insulator surfaces are brushlike and are known as brush discharges. *See* ELECTROSTATICS.

A. G. Bailey

Bibliography. M. Glor, *Electrostatic Hazards in Powder Handling*, 1988; D. M. Taylor and P. E. Secker, *Industrial Electrostatics: Fundamentals and Measurements*, 1994.

Static var compensator

A thyristor-controlled (hence static) generator of reactive power, either lagging or leading, or both. The word var stands for volt ampere reactive, or reactive power. The device is also called a static reactive compensator.

Need for reactive compensation. Reactive power is the product of voltage times current where the voltage and current are 90° out of phase with one another. Thus, reactive power flows one way for one-quarter of a cycle, the other way for the next quarter of a cycle, and so on (in contrast to the real power, or active power, which flows in one direction only). This back-and-forth flow results in no net power being delivered by the generator to the load. However, current associated with reactive power does flow through the conductor and creates extra losses. *See* ALTERNATING-CURRENT CIRCUIT THEORY; ELECTRIC POWER MEASUREMENT; VOLT-AMPERE.

Most loads, particularly motor loads, draw lagging reactive power along with active power. Increased lagging reactive power causes electric power system voltage to sag, and under heavy load conditions the voltage can decrease below desirable levels. (Utilities usually hold system voltage to within ±5% of the nominal value.) On the other hand, under light loads; the capacitance of high-voltage lines can create excessive leading reactive power, causing the voltage at some locations to rise above the nominal value.

Finally, it is prudent to keep reactive power flows to a minimum in order to allow the lines to carry more active power. For example, cables have very high capacitance to ground (compared to overhead lines), and the resulting leading reactive power loads a 10–20-mi (15–30-km) transmission cable to its thermal limit when voltage is applied, leaving it with no ability to carry real power.

Mechanical versus static compensation. Utilities frequently install capacitors connected from line to ground to compensate for lagging reactive power and reactors connected from line to ground to compensate for leading reactive power. These reactors and capacitors are switched in and out with mechanical switches based on the level of line loading as it varies throughout the day. However, frequent operation of these mechanical switches may reduce their reliability.

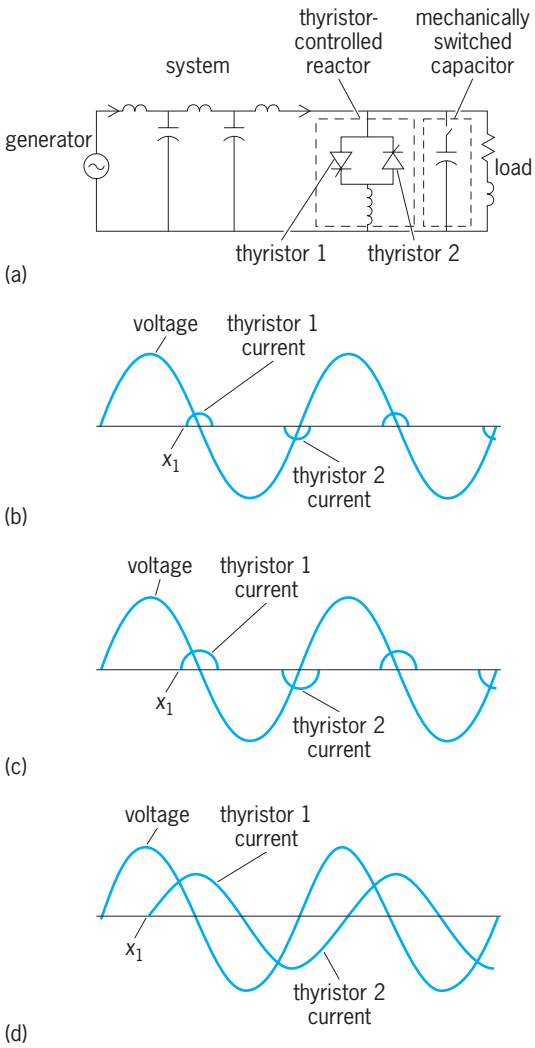

Fig. 1. One phase of a static var compensator (thyristor-controlled reactor). (*a*) Circuit diagram of thyristor-controlled reactor in the electric power system. (*b*) Voltage and current waveforms with partial conduction. (*c*) Waveforms with greater, but still partial, conduction. (*d*) Waveforms with full conduction.

It is desirable to have a controllable source of reactive power (leading or lagging); and the static var compensator, controlled with static switches, called thyristors, for higher reliability, fulfills this function. It is more expensive than mechanically switched capacitors and reactors (due to the cost of thyristor valves and associated equipment), and hence its use is based on an economic trade-off of benefits versus cost. *See* SEMICONDUCTOR RECTIFIER.

Operation. **Figure 1***a* shows one phase of a typical static var compensator (thyristor-controlled reactor). Lagging reactive power is produced by controlling the current flow through a reactor with thyristors.

By controlling the start of conduction of thyristor 1, current is allowed to flow at the desired moment, say x_1 (Fig. 1*b*). The current rises, reaches its peak at the voltage zero, and then falls back to zero in about the same time period that it took to reach its peak. Thyristor 2 similarly controls the current flow in the opposite direction. More current and hence an increasing amount of lagging reactive power can be allowed to flow by firing the thyristors at an earlier time (Fig. 1*c*). Maximum current will flow when the thyristors are fired 90° before the respective voltage zero (Fig. 1*d*). Thus, lagging reactive power can be rapidly controlled by the appropriate firing of thyristors. Control circuits, including firing circuits and sophisticated controls, vary the firing time. Since power systems are three-phase, three single-phase compensators must be constructed (**Fig. 2**).

Voltage. Since thyristors for the high voltages of transmission lines can be expensive, it is customary to provide a step-down transformer to lower the voltage to about 10–20 kV for connection of a static var compensator. Often a power transformer connected to a transmission circuit is provided with a tertiary winding of low voltage, to which the static var compensator can be connected. The relatively low voltage capability of thyristors requires connecting several of them in series. *See* TRANSFORMER.

Use of filters. Except when the thyristors are fired at 90° for maximum current, the current is not a pure 60-Hz sine wave, but rather contains undesirable harmonics of the 60-Hz component. Filters must be provided to eliminate most of these harmonics (Fig. 2). *See* ELECTRIC FILTER.

Since these filters have large capacitors, they also become a source of leading reactive power. Thus, the required rating of the controlled reactor is correspondingly increased to compensate for them. If smooth reactive power control over a wide range of leading and lagging reactive power is desirable, then additional mechanically switched capacitors (Fig. 1) and a larger controlled reactor are provided.

If mechanically controlled capacitors are not desirable because of the need for frequent operation, the capacitors can be controlled with thyristors (Fig. 2). However, because the current in the

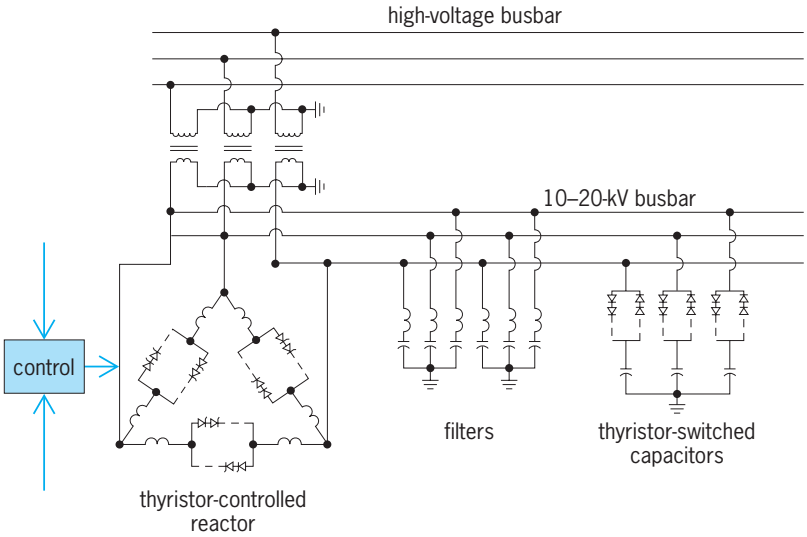

Fig. 2. Complete three-phase static var compensator, constructed from three single-phase compensators.

capacitors cannot be controlled as conveniently as in the reactors, several small capacitor banks are provided. Narain G. Hingorani

Bibliography. R. Hauth and R. Moran, *Static VAR Systems for Voltage and VAR Control*, IEEE Tutorial 78EH0135-4-PWR, IEEE Summer PAS Meeting, Los Angeles, July 1978; Institute of Electrical and Electronics Engineers, Power Systems Engineering Committee, *Application of Static VAR Systems for System Dynamic Performance* (symp.), IEEE Spec. Doc. 87TH0187-5-PWR, 1987; *Proceedings of the Transmission Static VAR Systems Seminar*, Elec. Power Res. Inst. Spec. Rep. EL-1047-SR, April 1979; *Transmission System Static Var Control*, Elect. Power Res. Inst. EL-2754 (final report), Dec. 1982.

Statics

The branch of mechanics that describes bodies which are acted upon by balanced forces and torques so that they remain at rest or in uniform motion. This includes point particles, rigid bodies, fluids, and deformable solids in general. Static point particles, however, are not very interesting, and special branches of mechanics are devoted to fluids and deformable solids. For example, hydrostatics is the study of static fluids, and elasticity and plasticity are two branches devoted to deformable bodies. Therefore this article will be limited to the discussion of the statics of rigid bodies in two and three spatial dimensions. *See* BUOYANCY; ELASTICITY; HYDROSTATICS; MECHANICS.

Mechanics is the study of motions in terms of mass, length, time, and forces. In statics the bodies being studied are in equilibrium. Positions of points in space, velocities, forces, and torques are all vector quantities, since each has direction, magnitude, and units. Beams, bridges, machine parts, and so on are not really rigid, but whenever the largest change in length of a portion of a body Δl is much smaller than the length l, that is, $l > \Delta l$, then it is a satisfactory approximation to treat that body as a rigid body. Many of the objects used in architecture, engineering, and physics can be satisfactorily idealized as rigid bodies, and much of building, machine, bridge, and dam design is based upon the study of statics.

The equilibrium conditions are very similar in the planar, or two-dimensional, and the three-dimensional rigid-body statics. These conditions are that the vector sum of all forces acting upon the body must be zero, and the resultant of all torques about any point must be zero. Thus it is necessary to understand the vector sums of forces and torques.

Force as a vector. The physical effect of a force is a push or pull on an object at its point of application. The effect of a force is to change the velocity of the body. The SI units of force are newtons (1 N = 1 kilogram-meter per second squared). The vector nature of forces is expressed by their direction in space. This direction is the direction of the push or pull exerted by the force. The notation for a vector quantity is an overarrow, for example, $\vec{F}$ or $\vec{P}$. In **Fig. 1** a two-dimensional vector $\vec{P}$ is shown as a directed line segment $\overrightarrow{AB}$ in the plane of the paper. A three-dimensional vector $\vec{P}$ is shown in **Fig. 2** as the directed line segment $\overrightarrow{OD}$. A complete description of a vector is obtained by specifying a coordinate system and giving its components in that coordinate system. The sum of any two vectors $\vec{A} + \vec{B}$ is the sum of their components, so that $\vec{A}$ and $\vec{B}$ must have the same number of components.

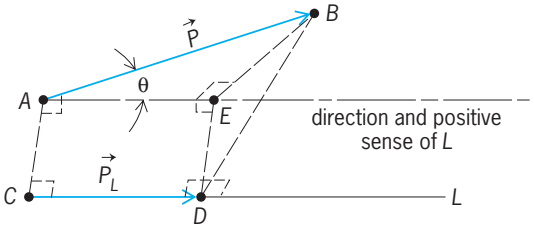

Fig. 1. Directed line $\overrightarrow{AB}$ or $\vec{P}$ is a vector quantity in the plane.

Superposition and transmissibility. In studying statics problems (and general mechanics problems), two principles, superposition and transmissibility, are used repeatedly on force vectors. They are applicable to all vectors, but specifically to forces and torques (first moments of forces).

1 The principle of superposition of d-dimensional vectors is that the sum of any two d−vectors is another d-vector. Of course, some or all components can be zero. This principle is illustrated in **Fig. 3**. The two vectors labeled Q, which can be shown applied to a rigid structural member D, add up to zero, and so do the two labeled R. Superposition applies to all sums of vectors, not just sums which vanish.

2 The principle of transmissibility of a force applied to a rigid body is that the same mechanical effect is produced by any shift of the application of the force. The principle of transmissibility is illustrated in Fig. 3. The solid rectangle C represents a solid structural member, and the line AB is the line

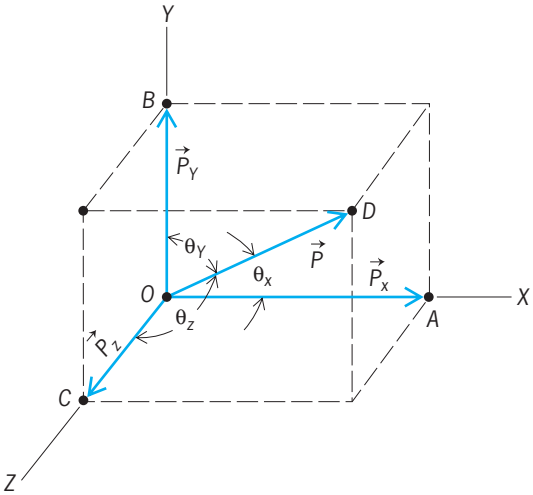

Fig. 2. Rectangular components of a force vector $\vec{P}$ in three-dimensional space.

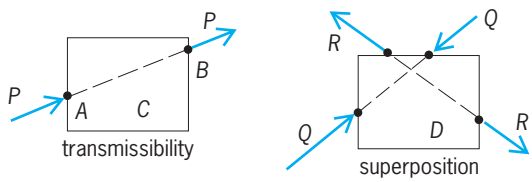

Fig. 3. Illustration of two theorems of statics.

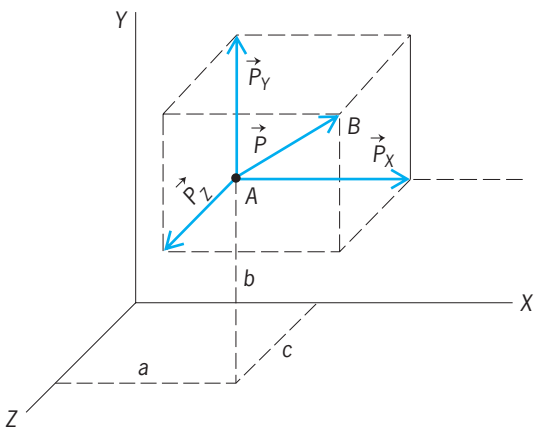

Fig. 4. Moments about the axis in rectangular coordinates.

of action for the equivalent forces P. Since these two forces have the same magnitudes and direction and are applied along the same line, by transmissibility, all of their mechanical effects are equivalent. To use the superposition principle to add two vectors, the principle of transmissibility is used to move some vectors along their line of action in order to add to their components.

Components of a force. Figure 1, the two-dimensional example, shows the construction by which orthogonal force components are defined. In this figure $\vec{P}$ or $\overrightarrow{AB}$ is a force vector, and L a directed line whose positive sense is toward its labeled end. Construction lines AC and BD are in planes (not shown) normal to L, and θ is the direction angle of $\vec{P}$ relative to L; it is a plane angle between the directions in the positive senses of $\vec{P}$ and L; further, $0 \leq \theta \leq 180^\circ$.

The orthogonal vector component of force $\vec{P}$ on directed line L is a force of direction and magnitude given by P_L or CD, where P_L is in the direction of L. Its magnitude is given by Eq. (1).

$$P_L = \overline{CD} = \overline{AE} = \overline{AB}|\cos\theta| = P|\cos\theta| \qquad (1)$$

The component of $\vec{P}$ on L is $P_L = P\cos\theta$, where P_L is positive if $0^\circ \geq 90^\circ < 90^\circ$ and negative if $90^\circ < \theta \leq 180^\circ$. The absolute magnitude of P_L is designated $|P_L|$.

The rectangular components of a force are its components on mutually perpendicular lines. In Fig. 2, $\vec{P}$ or $\overrightarrow{OA}$, $\vec{P}$ or $\overrightarrow{OB}$, and $\vec{P}$ or $\overrightarrow{OC}$ are the rectangular vector components of $\vec{P}$ or $\overrightarrow{OD}$ in the directions of lines (axes) X, Y, and Z, respectively.

The corresponding components have the magnitudes $P_X = P\cos\theta_X$, $P_Y = P\cos\theta_Y$, and $P_Z = P\cos\theta_Z$, and are in the X, Y, Z directions, respectively. The magnitude of the three-dimensional vector $\vec{P}$ is given by Eq. (2).

$$P = \sqrt{P_X^2 + P_Y^2 + P_Z^2} \qquad (2)$$

Moment of a force. The moment of a force about a directed line is a signed number whose value can be obtained by applying these two rules:

1 The moment of a force about a line parallel to the force is zero.

2 The moment of a force about a line normal to a plane containing the force is the product of the magnitude of the force and the least distance from the line to the line of the force. Conventionally, the moment is positive if the force points counterclockwise about the line as viewed from the positive end of the line.

In **Fig. 4** the moments of $\vec{P}$ about the X, Y, and Z coordinate axes are $M_X = bP_Z - cP_Y$ about the X axis; $M_Y = cP_X - aP_Z$ about the Y axis; and $M_Z = aP_Y - bP_X$ about the Z axis. The resultant of the two torques or moments of force, $\vec{T}_1$ and $\vec{T}_2$ about a point P, is shown as $\vec{T}_r$ in **Fig. 5**. The components of these torques are due to forces $\vec{F}_1$ and $\vec{F}_2$ oriented with respect to $\vec{T}$ as shown.

Figure 6 illustrates cases where the forces in a structure may or may not depend on the relative strength of its parts. Case 6*a* is called statically determinate, and case 6*b* statically indeterminate. The terms are actually misnomers, for the forces can be found in both cases. Better terminology would be rigidly determinable and indeterminable. In case 6*b*, if each member were rigid (infinitely strong), the forces in each could not be determined.

In summary, the statics of rigid bodies in statically determinate structures is carried out by summing vector forces and vector torques and setting the

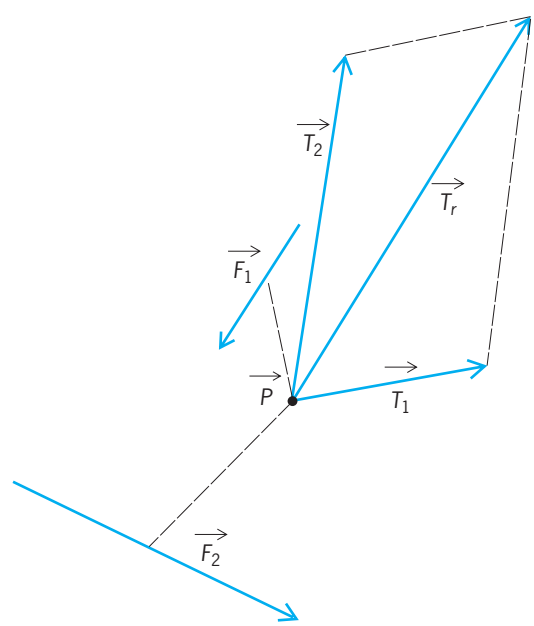

Fig. 5. Resultant of two torques.

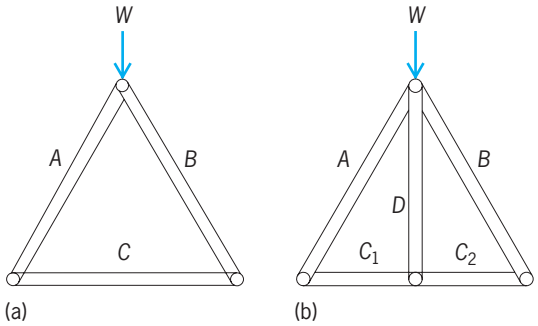

Fig. 6. Statically determinate and indeterminate structures. (*a*) The forces in members *A* and *B* are independent of their strengths (for small deformations). (*b*) The relative strength of the added member *D* will determine how much of the load *W* it supports.

resultants equal to zero. *See* COUPLE; EQUILIBRIUM OF FORCES; FORCE; TORQUE. Brian De Facio

Bibliography. F. P. Beer and E. Russell Johnston, *Vector Mechanics for Engineers*: *Statics*, 6th ed., 1996; W. L. Fowler and A. Bedford, *Engineering Materials: Statics*, 2d ed., 1998; R. C. Hibbeler, *Engineering Mechanics*: *Statics*, 8th ed., 2000; D. J. McGill and W. W. King, *Engineering Mechanics: An Introduction to Statics*, 3d ed., 1995; J. L. Meriam and L. G. Kraige, *Engineering Mechanics: SI-English Version*, vol. 1: *Statics*, 3d ed., 1992; A. Pytel and K. Kiusalaas, *Engineering Mechanics: Statics*, 1994.

Statistical mechanics

That branch of physics which endeavors to explain the macroscopic properties of a system on the basis of the properties of the microscopic constituents of the system. Usually the number of constituents is very large. All the characteristics of the constituents and their interactions are presumed known; it is the task of statistical mechanics (often called statistical physics) to deduce from this information the behavior of the system as a whole.

Scope of Statistical Mechanics

Elements of statistical mechanical methods are present in many widely separated areas in physics. For instance, the classical Boltzmann problem is an attempt to explain the thermodynamic behavior of gases on the basis of classical mechanics applied to the system of molecules. Historically, this was the first systematic investigation of a statistical problem, and many of the procedures and methods originate from this investigation. It is important to realize that statistical mechanics gives more than just an explanation of already known phenomena. By using statistical methods, it often becomes possible to obtain expressions for empirically observed parameters, such as viscosity coefficients, heat conduction coefficients, and virial coefficients, in terms of the forces between molecules. This kind of result, a direct relation between an observed macroscopic entity and an intermolecular potential, constitutes one of the main achievements of statistical physics. *See* BOLTZMANN STATISTICS; INTERMOLECULAR FORCES; KINETIC THEORY OF MATTER.

Statistical considerations also play a significant role in the description of the electric and magnetic properties of materials. In this case the objective is to deduce such characteristics as the permittivity, electrical conductivity, and magnetic permeability from the known properties of the atom. To obtain the Maxwell equations, which are macroscopic equations for matter in the bulk, from the Lorentz (electron) equations is typically a statistical question; new information which can be obtained from this approach includes, for example, the temperature dependence of the permittivity. *See* ELECTRICAL CONDUCTIVITY OF METALS; MAGNETIC SUSCEPTIBILITY; MAXWELL'S EQUATIONS; PERMITTIVITY; RELATIVISTIC ELECTRODYNAMICS.

If the problem of molecular structure is attacked by statistical methods, the contributions of internal rotation and vibration to thermodynamic properties, such as heat capacity and entropy, can be calculated for models of various proposed structures. Comparison with the known properties often permits the selection of the correct molecular structure. The statistical description of the activated complex in reaction mechanisms has enabled chemists to formulate a theory of absolute reaction rates. *See* MOLECULAR STRUCTURE AND SPECTRA.

Perhaps the most dramatic examples of phenomena requiring statistical treatment are the cooperative phenomena or phase transitions. In these processes, such as the condensation of a gas, the transition from a paramagnetic to a ferromagnetic state, or the change from one crystallographic form to another, a sudden and marked change of the whole system takes place. The appropriate description of such processes is one of the most difficult but most interesting problems of statistical physics. *See* PHASE TRANSITIONS.

Statistical considerations of quite a different kind occur in the discussion of problems such as the diffusion of neutrons through matter. In this case, the probability of the various events which affect the neutron are known, such as the capture probability and scattering cross section. The problem here is to describe the physical situation after a large number of these individual events. The procedures used in the solution of these problems are very similar to, and in some instances taken over from, kinetic considerations. Similar problems occur in the theory of cosmic-ray showers. Here the probabilities of the basic processes—pair production, annihilation, ionization, meson production—are all presumed known. The major task is the computation of the combined effect of a large number of these individual events. Special techniques of statistical physics, the use of distribution functions and transport-type equations, find extensive applications in these areas. *See* BOLTZMANN TRANSPORT EQUATION; COSMIC RAYS; REACTOR PHYSICS.

It happens in both low-energy and high-energy nuclear physics that a considerable amount of energy is suddenly liberated. An incident particle may

be captured by a nucleus, or a high-energy proton may collide with another proton. In either case, there is a large number of ways (a large number of degrees of freedom) in which this energy may be utilized. In the nuclear case, there are usually many decay and excitation modes; in the case of the proton-proton collision, there is enough energy available for the creation of a number of mesons. To survey the resulting processes, one can again invoke statistical considerations. The statistical problem here is to calculate the probability that, given a total amount of energy, a particular process will occur. Of course, the statistical method cannot help in the investigation of the actual mechanisms of these processes. However, the statistical factors must be considered in the interpretation of experiments. *See* NUCLEAR REACTION; SCATTERING EXPERIMENTS (NUCLEI).

Of considerable importance in statistical physics are the random processes, also called stochastic processes or sometimes fluctuation phenomena. The brownian motion, the motion of a particle moving in an irregular manner under the influence of molecular bombardment, affords a typical example. The process may be described in terms of a fluctuating force acting on a particle, perhaps in addition to a systematic force. The interest in this problem centers around a calculation of the position and velocity of a particle after it has experienced many collisions or after the fluctuating force has acted for a long time. The stochastic processes are in a sense intermediate between purely statistical processes, where the existence of fluctuations may safely be neglected, and the purely atomistic phenomena, where each particle requires its individual description. All statistical considerations involve, directly or indirectly, ideas from the theory of probability of widely different levels of sophistication. The use of probability notions is, in fact, the distinguishing feature of all statistical considerations. *See* BROWNIAN MOVEMENT; PROBABILITY; PROBABILITY (PHYSICS); STATISTICS; STOCHASTIC PROCESS.

Methods of Statistical Mechanics

In the following sections procedures used in statistical mechanics are discussed.

Phase space. For a system of N particles, each of mass m, contained in a volume V, the positions of the particles may be labeled $x_1, y_1, z_1, \ldots, x_N, y_N, z_N$, their cartesian velocities $\upsilon_{x1}, \ldots, \upsilon_{zN}$, and their momenta $p_{x1}, \ldots, p_{zN}$. The simplest statistical description concentrates on a discussion of the distribution function $f(x,y,z;\upsilon_x,\upsilon_y,\upsilon_z;t)$. The quantity $f(x,y,z;\upsilon_x,\upsilon_y,\upsilon_z;t)$ $(dxdydzd\upsilon_x d\upsilon_y d\upsilon_z)$ gives the (probable) number of particles of the system in those positional and velocity ranges where x lies between x and $x + dx$; υ_x between υ_x and $\upsilon_x + d\upsilon_x$, and so on. These ranges are finite. The six-dimensional space defined by $x, y, z, \upsilon_x, \upsilon_y, \upsilon_z$ is called the μ space, and the behavior of the gas is represented geometrically by the motion of N points in the μ space. It was shown by L. Boltzmann that in the course of time f approaches an equilibrium distribution f^0 as in Eq. (1). Here $d^3x = dxdydz$, A and β are parameters,

$$f^0(\mathbf{x}, \mathbf{v})\, d^3x d^3v = A \exp\left[-{}^1/_2\beta m\mathbf{v}^2 - \beta U(x, y, z)\right] d^3x d^3v \quad (1)$$

and $U(x,y,z)$ is the potential energy at the point x, y, z. Boltzmann also could interpret Eq. (1) as the most probable distribution.

Actually, important as the use of the distribution function is in practice, there are in principle serious limitations and difficulties associated with it. The limitations refer to the fact that the description by means of a single distribution function is correct only for noninteracting particles; also, the neglect of triple and higher collisions restricts the applicability to very dilute systems. The indiscriminate use of the ideas of Boltzmann leads to paradoxical results.

In the further study of these questions, the phase space of a dynamical system plays an important role. This is the $6N$-dimensional euclidean space, whose $6N$ axes are given in notation (2). The state of the

$$x_1, y_1, z_1, \ldots, x_N, y_N, z_N; p_{x_1}, \ldots, p_{z_N} \quad (2)$$

system at a given time is completely specified by these $6N$ numbers. These numbers define a point (the representative point, or phase point) in the $6N$-dimensional phase space. In the course of time the variables $x_1, \ldots, p_{zN}$ change, hence the phase point moves. If the system is conservative, Eq. (3) defines, for the energy E, a $(6N - 1)$-dimensional surface in

$$\sum_i \frac{\mathbf{p}_i^2}{2m_i} + U(x_1, \ldots, z_N) = E \quad (3)$$

the phase space called the energy surface. For a conservative system, the phase point in the course of time wanders over the energy surface. Intuitive as this geometrical representation of an involved mechanical system might seem, the actual trajectory of the phase point is extremely complicated. Also, the use of the phase space does not by itself imply any statistical considerations. A knowledge of the trajectory is precisely equivalent to a solution of the complete mechanical problem, for it determines each p and x as a function of time. However, many different microscopic states all correspond to the same macroscopic situation.

Observations made on a system always require a finite time; during this time the microscopic details of the system will generally change considerably as the phase point moves. The result of a measurement of a quantity Q will therefore yield the time average, as in Eq. (4). The integral is along the tra-

$$\overline{Q}_t = \frac{1}{t}\int_0^t Q\, dt \quad (4)$$

jectory in phase space; Q depends on the variables $x_1, \ldots, p_{zN}$ and t. To evaluate the integral, the trajectory must be known, and as mentioned above, this requires the solution of the complete mechanical problem.

For many years one of the major problems in statistical mechanics was precisely to recast the expression for a time average in a more tractable form. Generally attempts were made to replace time averages by phase space averages, by trying to show that the time spent in a region was proportional to the size of the region. The legitimacy of this procedure was studied extensively by mathematicians, but although these studies opened up significant areas in mathematics, they did not really help elucidate problems of physical interest.

The subtlety of the problems involved is illustrated by the so-called ergodic theorem of G. D. Birkhoff, which asserts that $\lim_{t\to\infty}\overline{Q}_t$ exists for almost all trajectories, although this limit varies discontinuously from trajectory to trajectory. The Poincaré recurrence theorem, which states that a system within a finite time will return as closely as desired to its initial state, is of special interest in statistical mechanics. Apart from providing the mathematical framework for these theorems, the phase space provides also the appropriate framework for the formulation of ensemble theory, which is the basis of modern statistical mechanics.

Ensembles; Liouville's theorem. J. Willard Gibbs first suggested that instead of calculating a time average for a single dynamical system, a collection of systems, all similar to the original one, should instead be considered. Such an ensemble of systems is to be constructed in harmony with the available knowledge of the single system, and may be represented by an assembly of points in the phase space, each point representing a single system. If, for example, the energy of a system is precisely known, but nothing else, the appropriate representative example would be a uniform distribution of ensemble points over the energy surface, and no ensemble points elsewhere. An ensemble is characterized by a density function $\rho(x_1, \ldots, z_N; p_{x1}, \ldots, p_{zN}; t) \equiv \rho(x,p,t)$. The significance of this function is that the number of ensemble systems dN_e contained in the volume element $dx_1 \ldots dz_N; dp_x \ldots dp_{zN}$ of the phase space (this volume element will be called $d\Gamma$), at time t is as given in Eq. (5).

$$\rho(x, p, t)d\Gamma = dN_e \tag{5}$$

The ensemble average of any quantity Q is given by Eq. (6). The basic idea now is to replace the time

$$\overline{Q}_{\text{ens}} = \frac{\int Q\rho\, d\Gamma}{\int \rho\, d\Gamma} \tag{6}$$

average of an individual system by the ensemble average, at a fixed time, of the representative ensemble. Stated formally, the quantity $\overline{Q}_t$ defined by Eq. (4), in which no statistics is involved, is identified with $\overline{Q}_{\text{ens}}$ defined by Eq. (6), in which probability assumptions are explicitly made. Another form of this same connection between the behavior of the individual system and the ensemble is that the probability that the individual system at time t will be in a region R of the phase space is given by the fraction of ensemble systems contained in R at time t, Eq. (7).

$$P(R, t) = \frac{\int_R \rho\, d\Gamma}{\int \rho\, d\Gamma} \tag{7}$$

It is clear, on the basis of these relations, that the complete statistical behavior of a system is known once its representative ensemble has been obtained. In the construction of such ensembles, it is always assumed that accessible parts of the phase space should be weighted equally. There is one other general requirement that the density function must satisfy. Inasmuch as the number of ensemble members remains constant (no ensemble members are created or destroyed in the course of time), there is a continuity equation, Eq. (8), for the density

$$\frac{\partial\rho}{\partial t} + \sum_{i=1}^{3N}\left[\frac{\partial}{\partial x_i}\left(\rho\frac{dx_i}{dt}\right) + \frac{\partial}{\partial p_i}\left(\rho\frac{dp_i}{dt}\right)\right] = 0 \tag{8}$$

function ρ which simply states that the change in the number of systems in a volume of phase space per unit time equals the difference of the number of systems flowing in and flowing out of that volume. If the time derivatives dx_i/dt and dp_i/dt are now expressed through the hamiltonian equations, Eq. (9) is obtained. Here $d\rho/dt$ is the substantial,

$$\frac{\partial\rho}{\partial t} + \sum_{i=1}^{3N}\left(\frac{\partial\rho}{\partial x_i}\frac{dx_i}{dt} + \frac{\partial\rho}{\partial p_i}\frac{dp_i}{dt}\right) \equiv \frac{d\rho}{dt} = 0 \tag{9}$$

or total derivative. *See* EQUATION OF CONTINUITY; HAMILTON'S EQUATIONS OF MOTION.

Equation (9), called the Liouville theorem, asserts that the time rate of change of ρ is zero along a streamline in phase space. It can also be deduced from Eq. (9) that the ensemble systems move in the course of time in a volume-preserving fashion through the phase space. If at some initial time a set of ensemble points occupy a volume V_0, then in the course of time each representative point will move in a complicated fashion. However, at time t the volume V_t made up by the ensemble points which initially were in V_0 is still equal to V_0. The shape of the volume changes, but the actual volume is constant. In the attempts which have been made to justify the postulate that $\overline{Q}_t = \overline{Q}_{\text{ens}}$, this volume-preserving property plays a dominant role. The proof of the Poincaré recurrence theorem also depends crucially on this property. Yet in spite of the general significance of the Liouville theorem, Eq. (9) generally does not help in the task of setting up an appropriate representative ensemble. (It is relevant, however, to the subsequent discussion on nonequilibrium theory of liquids.) The choice of the ensemble is usually determined by physical considerations and by the already mentioned postulate of equal a priori probabilities in phase space. Perhaps most important is the a posteriori justification of the statistical procedures, obtained through a comparison with the experimental facts. It is important that inasmuch as the methods are statistical, not only average values but also fluctuations around these average values can be computed. It turns out that usually these fluctuations are

extremely small; however, in those instances in which they are appreciable, they can be compared with experiments, providing additional support for the statistical procedures. Hence, despite the fact that no completely rigorous mathematical justification of the ensemble methods exists, these methods are physically so plausible and lead to such agreement with experiments for a variety of systems that they reasonably must be considered as a well-established part of statistical physics.

Relation to thermodynamics. It is certainly reasonable to assume that the appropriate ensemble for a thermodynamic equilibrium state must be described by a density function which is independent of the time, since all the macroscopic averages which are to be computed as ensemble averages are time-independent. Thus an equilibrium ensemble is realized when $\partial\rho/\partial t = 0$. In that case it follows from Liouville's theorem that Eq. (10) holds. It is easy to

$$\sum_{i=1}^{3N}\left(\frac{\partial\rho}{\partial x_i}\frac{dx_i}{dt}+\frac{\partial\rho}{\partial p_i}\frac{dp_i}{dt}\right)=0 \tag{10}$$

see that if ρ is a function of a quantity α (which is a function of x and p and which is a constant of the motion, such as the energy, so that $\partial\alpha/\partial t = 0$), then any $\rho(\alpha)$ satisfies Eq. (10). Hence any density function ρ which is a function of x and p through the dependence of the energy E on p and x satisfies the Liouville equation (9). The functional form of $\rho(E)$ is left completely unspecified.

Microcanonical ensemble. In the case of an isolated system, where the energy is specified within a well-defined range, the most obvious example to use would be the so-called microcanonical ensemble defined by Eq. (11*a*), where c is a constant, for the energy E between E_0 and $E_0 + \Delta E$; for other energies Eq. (11*b*) holds. By using Eq. (6), any microcanonical average may be calculated. The calculations, which involve integrations over volumes bounded by two energy surfaces, are not trivial. Still, many of the results of classical Boltzmann statistics may be obtained in this way.

$$\rho(p,x)=c \tag{11a}$$

$$\rho(p,x)=0 \tag{11b}$$

Canonical ensemble. For applications and for the interpretation of thermodynamics, the canonical ensemble is much more preferable. This ensemble describes a system which is not isolated but which is in thermal contact with a heat reservoir. By describing the complete system (system plus reservoir) by a microcanonical ensemble, it may be shown that the system itself may be represented by Eq. (12*a*). Here

$$\rho(x,p)=N_e\exp\left[\frac{\psi-E(x,p)}{\theta}\right] \tag{12a}$$

$$\int\rho(x,p)\,d\Gamma=N_e \tag{12b}$$

N_e is the total number of ensemble systems, $E(x,p)$ is the mechanical energy of the system, and ψ and θ are parameters independent of x and p characterizing the canonical ensemble. From the fact that Eq. (12*b*) is valid, it follows that the two parameters ψ and θ are not independent, but that Eq. (13) applies.

$$e^{-(\psi/\theta)}=\int d\Gamma\exp\left[-\frac{E(x,p)}{\theta}\right]\equiv Z \tag{13}$$

The quantity Z is usually called the partition function. It is customary to define Z for a system of N identical particles with a different multiplicative constant by Eq. (14). Here h is the Planck constant. It is now

$$Z'=\frac{1}{N!h^{3N}}\int\exp\left[-\frac{E(p,x)}{\theta}\right]d\Gamma \tag{14}$$

very important that the parameters θ and ψ be identified with definite thermodynamic functions. This may be done by considering some of the thermodynamic relations. If η is the entropy, ϵ the thermodynamic energy (internal energy), and ψ' the free energy (Helmholtz free energy), then some examples of thermodynamic relations are, for gases, described by pressure, volume, and temperature (P, V, T) as in Eqs. (15).

$$\psi'=\epsilon-T\eta \tag{15a}$$

$$d\psi'=-\eta\,dT-P\,dV \tag{15b}$$

$$\frac{\partial\psi'}{\partial T}=-\eta \qquad \frac{\partial\psi'}{\partial V}=-P \tag{15c}$$

$$\epsilon=\psi'-T\frac{\partial\psi'}{\partial T} \tag{15d}$$

See ENTROPY; FREE ENERGY; INTERNAL ENERGY; THERMODYNAMIC PRINCIPLES.

From Eqs. (12*a*), (6), and (13) the average energy $\overline{E}$ may be calculated, Eq. (16). For an ideal gas, in which the energy of a molecule is $\mathbf{p}^2{}_i/2m$, Eq. (17)

$$\overline{E}=\theta^2\frac{\partial}{\partial\theta}\log Z \tag{16}$$

$$\begin{aligned}Z&=\int\cdots\int dx_1\cdots dp_{z_N}\exp\left(-\tfrac{1}{2m\theta}\sum_{i-1}^{N}\mathbf{p}_i^2\right)\\&=V^N(2\pi m\theta)^{3N/2}\end{aligned} \tag{17}$$

holds. (The evaluation of the integrals is elementary; the space integrations merely contribute V^N, and the momentum integrals are gaussian.) Equation (18*a*) is

$$\overline{E}=\frac{3N}{2}\theta \tag{18a}$$

$$\epsilon={}^3/_2NkT\equiv{}^3/_2RT \tag{18b}$$

then obtained as the energy of an ideal gas, having N molecules in a vessel of any volume V. Now it is known by experiment that the thermodynamic energy of 1 mole of an ideal gas is given by Eq. (18*b*). Here R is the ideal gas constant, T is the absolute temperature, N is Avogadro's number, and k is the Boltzmann constant. Comparison of Eqs. (18*a*) and (18*b*) leads to the identification of $\theta = kT$. It is also

easy to deduce the Maxwell-Boltzmann distribution. Applying Eq. (7) yields Eq. (19) for the probability

$$P\left(\mathbf{p}_1\mathbf{x}_1\right) d^3p_1 d^3x_1 = \frac{\int\cdots\int d_3x_1\, d_3p_1\, d^3x_2\cdots d^3p_N e^{-E/\theta}}{\int\cdots\int d^3x_1\cdots d^3p_N e^{-E/\theta}} = \frac{1}{V} e^{-\mathbf{p}^2/(2m\theta)}(2m\theta)^{-3/2}\, d^3p_1 d^3x_1 \tag{19}$$

that molecule 1 will have a given range of positions and momenta. The canceled symbols in the numerator mean that the coordinates and momenta of the molecule 1 are not integrated over. The result given by Eq. (19) is again true for an ideal gas. The identification of $\theta = kT$, however, is made in general. By differentiating Eq. (13) with respect to θ, it is possible to establish from Eq. (18*a*) the general connection (using $\theta = kT$) written as Eq. (20). When this general

$$\overline{E} = \psi - T\frac{\partial\psi}{T} = \epsilon \tag{20}$$

relation, deduced from the canonical ensemble, is compared with the general thermodynamic relation (15*d*), the parameter θ and the free energy θ' are indeed seen to play identical roles. Again, for an ideal gas, ψ can be computed, and it is the same as the ideal gas free energy. Hence ψ and ψ' are identified in general. Equation (13) relates the thermodynamic free energy to an integral containing the mechanics of the microscopic problem. It is also interesting to observe that a quantity η' defined by Eq. (21) plays in

$$\eta' = \frac{\int \rho \log \rho\, d\Gamma}{\int \rho\, d\Gamma} \tag{21}$$

all instances the role of the entropy. In fact, it can be shown that the equation $\psi = \epsilon - T\eta'$ is in harmony with Eq. (15*a*). These thermodynamic analogies can be followed through in all detail to see that the canonical ensemble combined with the relevant definitions does indeed reproduce the formal aspects of thermodynamics. The average energy $\overline{E}$ for a canonical ensemble has already been calculated. It is important to know just how often appreciable deviations from this average energy can be expected. For this, the fractional fluctuations in energy given by notation (22) must be calculated. The definition of $\overline{E}^2$, in Eq. (23), follows again from Eq. (6). From Eq. (23)

$$\frac{\overline{E}^2 - (\overline{E})^2}{(\overline{E})^2} \tag{22}$$

$$\overline{E}^2 = \frac{\int E^2 e^{-E/\theta}\, d\Gamma}{\int e^{-E/\theta}\, d\Gamma} \tag{23}$$

the quantity $\overline{E}^2$ may be expressed directly in terms of Z, the partition function. For an ideal gas it is then possible to obtain, using Eq. (17), an explicit expression, Eq. (24),for the fluctuations in energy. Since the

$$\left(\frac{\overline{E}^2 - (\overline{E})^2}{(\overline{E})^2}\right)^{1/2} = \sqrt{\frac{2}{3N}} \tag{24}$$

number of particles N is about 10^{22} in systems dealt with in practice, the chance of observing a sizable deviation from the average energy is extremely small. Using these same methods, defining in addition the specific heat $C_v = \partial\overline{E}/\partial T$, it may be shown generally that the fractional fluctuation in energy is given by Eq. (25). Even though the energy fluctuations are

$$\left(\frac{\overline{E}^2 - (\overline{E})^2}{(\overline{E})^2}\right)^{1/2} = \left(\frac{kT^2C_v}{(\overline{E})^2}\right)^{1/2} \tag{25}$$

negligible for an ideal gas, this is not always so for other systems. For solids at low temperatures, the fluctuations may be appreciable. When phase transitions (of the first order) take place, C_v becomes infinite, indicating via Eq. (25) that the fluctuations become very large. It is clear that the main problem of the applications is reduced to the calculation of the partition function Z; all thermodynamic entities follow from Eqs. (16) and (13).

Grand canonical ensemble. There is yet another ensemble which is extremely useful and which is particularly suitable for quantum-mechanical applications. Much work in statistical mechanics is now based on the use of this so-called grand canonical ensemble. The grand ensemble describes a collection of systems; the number of particles in each system is no longer the same, but varies from system to system. The density function $\rho(N,p,x)\, d\Gamma_N$ gives the probability that there will be in the ensemble a system having N particles, and that this system, in its $6N$-dimensional phase space Γ_N, will be in the region of phase space $d\Gamma_N$. The function ρ is given by Eq. (26). Here θ, Ω, and μ are the parameters char-

$$\rho(N,p,x) = \exp\left(\frac{\Omega + \mu N - E(p,x)}{\theta}\right) \tag{26}$$

acterizing the ensemble, just as ψ and θ characterize the canonical ensemble. It is again true that these parameters are directly related to thermodynamic state functions. The detailed argument follows the identical pattern indicated in the discussion of the canonical ensemble. The normalization condition is now given by Eq. (27). The grand canonical average of

$$\sum_N \int d\Gamma_N \rho(N,p,x) = \sum_N \int d\Gamma_N \exp\left(\frac{\Omega + \mu N - E(p,x)}{\theta}\right) = 1 \tag{27}$$

any quantity Q is given by Eq. (28) and the grand partition function by Eq. (29*a*). Again Eq. (29*b*) is

$$\overline{Q}_{\text{gr}} = \sum_N \int d\Gamma_N \rho(N,p,x)\, Q_N(p,x) \tag{28}$$

$$Z_{\text{gr}} = \sum_N e^{\mu N/\theta} \int d\Gamma_N e^{-(E/\theta)} \tag{29a}$$

$$Z_{\text{gr}} = \sum_N \frac{1}{N!}\frac{1}{h^{3N}} e^{\mu N/\theta} \int d\Gamma_N e^{-(E/\theta)} \tag{29b}$$

customarily used for a system consisting of N identical particles. From Eq. (27) it follows immediately

that Eq. (30) holds. From Ω, sometimes called the

$$Z_{\text{gr}} = e^{-(\Omega/\theta)} \tag{30}$$

grand potential, all other thermodynamic functions may be computed. The parameter μ is the chemical potential. It is defined by $\mu = (\partial\xi/\partial N)_{P,T}$ where ξ is the Gibbs free energy or thermodynamic potential; $\xi = \psi + PV$. Formally, these results are written as Eqs. (31). From a knowledge of the grand partition

$$P = -\left(\frac{\partial\Omega}{\partial V}\right)_{\mu,T} \tag{31a}$$

$$\eta = -\left(\frac{\partial\Omega}{\partial T}\right)_{V,\mu} \tag{31b}$$

$$\overline{N}_{\text{gr}} = -\left(\frac{\partial\Omega}{\partial \mu}\right)_{V,T} \tag{31c}$$

function as a function of V, T, and μ, all thermodynamic functions follow. For an ideal gas, for example, Eq. (29*b*) in conjunction with Eq. (17) yields Eq. (32*a*). Here λ is the thermal de Broglie wave-

$$Z_{\text{gr}} = \sum_{n=1}^{\infty} \frac{1}{N!}\left(\frac{2\pi m\theta}{h^2}\right)^{(3/2)N} V^N e^{\mu N/\theta}$$

$$= \exp\left(e^{\mu/\theta}\frac{V}{\lambda^3}\right) \tag{32a}$$

$$\lambda = \frac{\hbar}{\sqrt{2\pi mkT}} \tag{32b}$$

length, defined by Eq. (32*b*). From Eqs. (32*a*) and (30) it follows that Eq. (33) holds. Application of

$$\Omega = -\theta e^{\mu/\theta}\frac{V}{\lambda^3} \tag{33}$$

Eqs. (31) yields the usual results for ideal gases, such as $PV = \overline{N}kT$; however, the grand average of $\overline{N}$ enters the relations now, rather than just $\overline{N}$. From the definitions (28) and (29*b*) it can be shown that the fractional fluctuation in the number N is given by Eq. (34). Thus for gases, the fluctuations in num-

$$\left(\frac{\overline{N^2_{\text{gr}}} - (\overline{N}_{\text{gr}})^2}{\overline{N}^2_{\text{gr}}}\right)^{1/2} = \frac{1}{\sqrt{N_{\text{gr}}}} \tag{34}$$

ber are negligibly small, showing that the number of particles in the grand canonical ensemble is sharply peaked around the average value. The physical results deduced from the canonical and grand canonical ensemble can therefore be expected to be the same; the calculations, however, are frequently (especially in quantum problems) simpler in the grand canonical than in the canonical scheme. *See* CHEMICAL EQUILIBRIUM; CHEMICAL THERMODYNAMICS.

Applications. Two of the newest and most important applications of statistical physics involve the theory of nonideal gases and the theory of nonequilibrium states of dense gases and liquids.

Theory of nonideal gases. The importance of relations such as Eq. (13), which connects the thermodynamic free energy to the partition function, lies in the general validity of these connections. Specifically, when dealing with a gas of interacting molecules, where the energy is given by Eq. (35) with the potential U

$$E(x_1, \ldots, z_N; p_{x_1}, \ldots, p_{z_N}) = \sum_{i=1}^{N} \frac{p^2_{x_i} + p^2_{y_i} + p^2_{z_i}}{2m} + U(x_1, \ldots, z_N) \tag{35}$$

a known function, the relation is still valid. In that case, the partition function may be obtained directly in terms of U by performing the integrations over the momenta as in Eq. (36). Since from Z_N all thermody-

$$Z_N = \frac{1}{N!}\frac{1}{\lambda^{3N}}\int \cdots \int d^3x_1, \ldots, d^3x_N \times \exp\left[-\frac{U(x_1, \ldots, z_N)}{kT}\right] \tag{36}$$

namic relations follow, Eq. (36) is the appropriate starting point for all studies of nonideal gases. The canonical ensemble has been used in obtaining Z_N, which depends on V through the integration limits, on T through λ [see Eq. (32*b*)], and on N through the number of integrations. Since Z_N gives the free energy immediately, the equation of state (the relation between P, V, and T) can be derived by computing $P = -\partial\psi/\partial V$. The experimental results on the equations of state may be expressed in terms of the so-called virial expansion, Eq. (37). Here $B(T)$ and $C(T)$

$$P = \frac{RT}{V}\left(1 + \frac{B(T)}{V} + \frac{C(T)}{V^2} + \cdots\right) \tag{37}$$

are the second and third virial coefficients; they are determined experimentally. It is clear that Eq. (37) is a development starting from the ideal gas equation. A first task for statistical mechanics is to deduce the form (37), as well as an explicit expression for $B(T)$, and so on. A first step in carrying this out is to set up relation (38*a*), and to assume that the potential

$$\exp\left[-\frac{U(x_1, \ldots, z_N)}{kT}\right] \equiv W(1, 2, \ldots, N) \tag{38a}$$

$$U = U(\alpha) + U(\beta) \tag{38b}$$

$$W = W(\alpha)W(\beta) \tag{38c}$$

$$\begin{aligned}
W(1) &= S(1) = 1\\
W(1,2) &= S(1,2) + S(1)S(2)\\
W(1,2,3) &= S(1,2,3) + S(1)S(2,3) + S(2)S(1,3)\\
&\quad + S(3)S(1,2) + S(1)S(2)S(3)\\
W(1,\ldots,N) &= S(1,\ldots,N)\\
&\quad + S(1)S(2,\ldots,N)\\
&\quad + \cdots + S(1,2)S(3,\ldots,N)\\
&\quad + \cdots + S(1)\cdots S(N)
\end{aligned} \tag{38d}$$

$U(x_1, \ldots, z_N)$ is such that it is possible to make an unambiguous distinction between interacting and noninteracting particles. Any potential which vanishes

for a separation larger than a critical amount has this property, and any short-range potential approximates it. A separated configuration may be defined as one in which none of the molecules of one group interacts with any of another. If two such groups are called α and β, Eq. (38*b*) is satisfied for such a configuration. Hence, by relation (38*a*), Eq. (38*c*) holds. A set of functions S may now be defined by Eqs. (38*d*). From Eqs. (38*c*) and (38*d*) the following important property can be proven: the S functions vanish for a separated configuration. For instance $S(1,2,\ldots,l)$ is zero, unless the molecules $1, 2, \ldots, l$ form a nonseparated configuration. This is an important property, for it allows the integration of the S functions in a simple fashion. This can be illustrated by the example of integral (39*a*). To evaluate this integral it is useful

$$\int \cdots \int d^3x_1 \cdots d^3x_l S(1,\ldots,l) \tag{39a}$$

$$\int \cdots \int dx_1 \cdots dz_1 S(1,\ldots,l) = Vl!b_l(T) \tag{39b}$$

to first imagine that molecule 1 is fixed somewhere in the middle of the vessel. Since the S functions vanish for a separated configuration, the l molecules must all be quite near to molecule 1. Roughly speaking, if a is the range of the molecular forces, integral (39*a*) gets contributions only from a range of about la around molecule 1. If la is smaller than $V^{1/3}$, the first $l-1$ integrations are independent of V. The last integrations over the coordinates of 1 just contribute V, for molecule 1 can be placed anywhere in V. (Wall effects are ignored.) Thus Eq. (39*b*) holds. The $l!$ is a normalization factor and b_l is independent of V. The b_l are called cluster integrals. Using Eqs. (36), (38*a*), (38*d*), and (39*b*), the pressure may be obtained in terms of the cluster integrals as in Eq. (40). Thus, the experimental form of the virial

$$P = \frac{RT}{B} \times \left(1 - \frac{Nb_2}{V} = 4\left(N^2b_2^2 - 2N^3b_3\right)\frac{1}{V^2} + \cdots\right) \tag{40}$$

development has been deduced, while the cluster integrals are related to the experimental virial coefficients by Eqs. (41*a*) and (41*b*). Using Eq. (39*b*) for

$$B(T) = -Nb_2 \tag{41a}$$

$$C(T) = N^2(4b_2^2 - 2b_3) \tag{41b}$$

$$B(T) = 2\pi N \int_0^\infty r^2\, dr(1 - e^{-[U(r)/kT]}) \tag{41c}$$

$l = 2$ gives a direct relation between B and the intermolecular force potential $U(r)$ as in Eq. (41*c*). This is a "perfect" statistical formula. In the case of helium, where both $U(r)$ and B are known, Eq. (41*c*) may in fact be checked. [At low temperatures, quantum effects, not included in Eq. (41*c*), begin to play an important role.]

The complete equation of state can be obtained using a similar procedure. The result comes out in an implicit form, as in Eqs. (42*a*) and (42*b*), derived

$$P = kT\Sigma b_l z^l \tag{42a}$$

$$\frac{N}{V} = \sum_{l=1}^{N} lb_l z^l \tag{42b}$$

by J. E. Mayer. In principle, the auxiliary variable z can be eliminated between Eqs. (42*a*) and (42*b*) to obtain a relation between P, V, and T. Of course, to have a useful relation it is necessary to know the general character of the cluster integrals, and, in fact, this is not known, except for special molecular models. A question which has concerned physicists for some time is whether or not the system of equations (42) actually predicts a condensation phenomenon. It is well known that every real gas at a low enough temperature will condense when the volume is decreased. During the condensation process the pressure remains constant, the onset of condensation being marked by a discontinuity in the slope of the isotherm. The problem is now whether this information can be obtained from Eqs. (42). An interesting clue was obtained by Mayer. He showed that for low enough temperatures, $b_l(T) \cong \text{constant} \times b^l{}_0$. Substituting this relation in Eq. (42*b*) yields Eq. (43).

$$\frac{N}{V} = \sum_{l=1}^{N} l(\text{constant})(b_0 z)^N \tag{43}$$

Since the last power in the series is $N \cong 10^{24}$, the series given by Eq. (43) is radically different for $b_0 z < 1$ and $b_0 z > 1$. If $b_0 z > 1$, $(b_0 z)^N$ is a very large number and a change in N/V will cause a very slight change in z, hence a very slight change in P by Eq. (42*a*). Therefore, $b_0 z = 1$ separates two ranges: If $b_0 z < 1$, changes in N/V cause reasonable changes in P; if $b_0 z > 1$, P becomes quite insensitive to changes in N/V. This qualitative idea has been refined in many ways, by rigorously studying limit (44). These stud-

$$\lim_{\substack{N\to\infty \\ V\to\infty}} (Z_N)^{1/N} \tag{44}$$

ies have clarified some aspects of the situation, but even so the problem is still not completely settled. *See* GAS.

Nonequilibrium theory of liquids. It was observed in the earlier discussion of Liouville's theorem that the Liouville equation usually does not help in setting up an appropriate ensemble. Yet the Liouville equation has become the starting point for an important development aimed at an understanding of the nonequilibrium states of dense gases and liquids. This development starts from the Liouville equation (45). The interpretation that $\rho(x_1, \ldots,$

$$\frac{\partial\rho}{\partial t} + \sum_{i=1}^{3N}\left(\frac{\partial\rho}{\partial x_i}\frac{dx_i}{dt} + \frac{\partial\rho}{\partial p_i}\frac{dp_i}{dt}\right) = 0 \tag{45}$$

$p_{z_N};t)d\Gamma$ is the probability of finding the molecules of the system in their prescribed momentum and position ranges may now be used. A set of probability

functions (or distribution functions) can be defined by Eqs. (46). The function f_1 is (apart from constants)

$$f_1(\mathbf{x}_1, \mathbf{p}_1, t) = \int \cdots \int d(1)d(2)\cdots dN\ \rho(x, p, t) \qquad (46a)$$

$$f_1(\mathbf{x}_1, \ldots, \mathbf{p}_l; t) = \int \cdots \int d(1)\cdots d(l)d(l+1)\cdots d(N)\ \rho(x, p, t) \qquad (46b)$$

the Boltzmann distribution function. The canceled symbol $d(1)$ means that the variables x_1, y_1, z_1, p_{x1}, p_{y1}, p_{z1} are not integrated over. The higher distribution functions become important for dense systems. If it is assumed that the potential energy of the system consists of an external potential V_0, and a potential U which is additive, then integration of Eq. (45) over all coordinates except those of molecule 1 yields an equation for f_1, Eq. (47).

$$\frac{\partial f_1}{\partial t} + \frac{p_\alpha}{m_1}\frac{\partial f_1}{\partial x_\alpha} - \sum_{\alpha=1}^{3}\frac{\partial V_0}{\partial x_\alpha}\frac{\partial f_1}{\partial p_\alpha}$$

The equation for f_l involves f_{l+1}, and so on.

$$= \iint d^3p_2 d^3x_2 \sum_\alpha \frac{\partial U(x_1x_2)}{\partial x_\alpha}\frac{\partial f_2(p_1p_2, x_1x_2)}{\partial p_\alpha} \qquad (47)$$

The discussion of this hierarchy forms the basis of the study of the nonequilibrium phenomena in dense gases. The basic difficulty is that the equations for f_l always involve a function f_{l+1}. To obtain an equation for a single f_l function, it is necessary to make a guess or assumption about the way in which a higher f function can be expressed in terms of lower ones, if indeed this can be done at all. A frequently discussed possibility is the so-called superposition approximation relation (48). An appeal to probabilities of independent events makes relation (48) appear reasonably plausible. However, it has not been proven that the approximation of relation (48) is indeed consistent with the system of Eq. (47). The use of relation (48) in Eq. (47) leads to (nonlinear) integral equations. In spite of the considerable amount of work done with these equations, the results still have not led to great advances in the theory of dense gases. However, for additive central potentials, defined by $U(1, \ldots, N) = \Sigma U(r_{ij})$, it is possible to express the equation of state in terms of the pair distribution function. This function is defined by Eq. (49a). Here

$$f_2(p_1x_1, p_2x_2) \cong f_1(p_1x_1)f_1(p_2x_2) \qquad (48)$$

$$n_2(\mathbf{x}_1, \mathbf{x}_2) = N(N-1) \times \frac{\int \cdots \int d(1)\, d(2)\, d(3)\cdots d(N)W(1, \ldots, N)}{\int \cdots \int d(1)\cdots d(N)W(1, \ldots, N)} \qquad (49a)$$

$$n_2(\mathbf{x}_1, \mathbf{x}_2) = \iint d^3p_1\, d^3p_2\, f_2(\mathbf{x}_1\mathbf{p}_1, \mathbf{x}_2\mathbf{p}_2) \qquad (49b)$$

$W(1, \ldots, N)$ is defined by Eq. (38). The function n_2 gives the positional distribution of molecular pairs. As such it is less detailed than f_2, which gives the distribution of pairs both in positions and momenta. Apart from combinatorial factors, Eq. (49b) is valid. For central additive potentials, it may now be shown that n_2 is in fact a function of the distance r_{12} between the molecules: $n_2 \propto g(r)$, where $g(r)$ is defined as giving the number of molecules between r and $r + dr$. *See* INTEGRAL EQUATION.

An analysis similar to the one given in the preceding discussion on nonideal gases now leads to an equation of state, Eq. (50), in terms of $g(r)$ alone.

$$PV = NkT - \frac{2\pi N(N-1)}{3V}\int g(r)\frac{dU}{dr}r^3\, dr \qquad (50)$$

The integral equation referred to for the f_2 functions (once the superposition approximation is made) may now be expressed as an approximate integral equation for $g(r)$. This equation, obtained by M. Born, H. S. Green, and J. Yvon, leads to approximations for the virial coefficient, which are fairly good but not excellent. The pair distribution function $g(r)$ may be obtained directly from experimental x-ray scattering data. In principle experimental x-ray data could be used to obtain via Eq. (50) data about the equation of state. For this to be a feasible procedure would demand accuracies much beyond the present limits, as well as extensive measurements over a range of temperatures and densities. Unfortunately, therefore, this is not a possible procedure at present, although the formal relation (50) is generally valid. A current theory of liquid helium makes use of the relation between the pair distribution function and neutron-scattering data. *See* LIQUID HELIUM; NEUTRON DIFFRACTION; X-RAY DIFFRACTION.

Quantum statistical mechanics. The ensemble techniques can also be applied to systems which are described in terms of quantum mechanics. Consider an ensemble, where each system is described by a hamiltonian operator H. Let $\psi^\alpha(x,t)$ be a wave function (x stands for $\mathbf{x}_1, \ldots \mathbf{x}_N$, and α characterizes an ensemble member). The Schrödinger equation is written as Eq. (51).

$$H\psi^\alpha(x, t) \doteq -\frac{\hbar}{i}\frac{\partial \psi^\alpha(x, t)}{\partial t} \qquad (51)$$

See NONRELATIVISTIC QUANTUM THEORY.

If $\varphi_n(x)$ is a complete orthogonal set, Eq. (52), the

$$\int \varphi_n^* \varphi_m = \delta_{nm} \qquad (52)$$

function ψ^α may be developed in terms of the set $\{\varphi\}$ as in Eq. (53). Here $|a_n^\alpha(t)|^2$ is the probability

$$\psi^\alpha(x, t) = \sum_n a_n^\alpha(t)\varphi_n(x) \qquad (53)$$

that ensemble member α, at time t, is in state n. Of course, Eq. (54) is valid.

$$\sum_n \left|a_n^\alpha(t)\right|^2 = 1 \qquad (54)$$

If G is a quantity which is defined for each ensemble system by G^α, then an ensemble average of

G is defined by Eq. (55a), where N_e is the number

$$\overline{\overline{G}} = \frac{1}{N_e}\Sigma(G^\alpha) \tag{55a}$$

$$\overline{G}^\alpha = \int (\psi^\alpha)^* G_{\text{op}} \psi^\alpha \tag{55b}$$

of ensemble members. It is important to distinguish the ensemble average from the ordinary quantum-mechanical average, which is defined for a single system as in Eq. (55b). The notation, double bar for ensemble average, single bar for quantum-mechanical average, stresses the difference. Of special importance in the calculation of averages is the density matrix. The matrix elements of ρ, the density matrix, relative to the set of functions φ are defined by Eq. (56a). The sum of the diagonal elements of ρ, called the trace of ρ, is given by Eq. (56b). The density matrix ρ is the counterpart of the classical density function. From the Schrödinger equation (53) the analog of the Liouville theorem, Eq. (57), can

$$\rho_{mn} = \frac{1}{N_e}\sum_\alpha (a_n^\alpha)^* a_n^\alpha = \overline{\overline{a_n^* a_m}} \tag{56a}$$

$$\text{Tr}(\rho) = \sum_n \rho_{mn} = \frac{1}{N_e}\sum_\alpha \sum_n |a_n^\alpha|^2 = 1 \tag{56b}$$

$$i\hbar \frac{d\rho_{mn}}{dt} = [H, \rho]_{mn} \tag{57}$$

be deduced immediately. Here $[a,b] \equiv ab - ba$ is the commutator. Finally the basic relation giving the ensemble average of any operator, Eq. (58), can be

$$\overline{\overline{G}} = \text{Tr}(\rho G) \tag{58}$$

obtained. The fact that observable entities, such as ensemble averages, come out in the form of a trace, implies that the calculated results are independent of the set of functions φ. This is true because a change in this basic set will induce a similarity transformation on both ρ and G, and the trace operation is invariant under such a transformation. *See* DENSITY MATRIX.

An ensemble in which all ensemble members are in the identical state (ψ^α is independent of α) constitutes a pure case. An ensemble in which the ψ^α do depend on α is a statistical mixture. A necessary and sufficient condition for a pure case is that $\rho^2 = \rho$.

At equilibrium, the system is the counterpart of the canonical ensemble. Then the density operator [see Eq. (12a)] is written as Eq. (59a). The free energy is

$$\rho = e^{(\psi - H)/kT} \tag{59a}$$

$$Z = e^{-(\psi/kT)} = \text{Tr}\left(e^{-(H/kT)}\right) \tag{59b}$$

again directly related to the partition function as in Eq. (59b). In Eq. (60), Q is an operator e^Q defined by

$$e^Q = \sum_n \frac{Q^n}{n!} \tag{60}$$

the series expansion. The thermodynamic relations follow as before.

The construction of an ensemble in quantum mechanics must be performed in harmony with the knowledge of the system. If the system is known to be in either one of two quantum states, the ensemble members must be evenly distributed over these states with, in addition, random phases. It is important to include the requirement of random phases; if this were not done, a uniform stationary ensemble would not even remain stationary. *See* QUANTUM STATISTICS.

Max Dresden

Bibliography. D. S. Betts, *Introduction to Statistical Mechanics*, 1992; C. Garrod, *Statistical Mechanics and Thermodynamics*, 1995; W. Greiner, L. Neise, and H. Stocker, *Thermodynamics and Statistical Mechanics*, 1995; K. Huang, *Statistical Mechanics,* 2d ed., 1987; R. Kubo et al., *Statistical Physics*, 2 vols., 2d ed., 1991; L. D. Landau and E. Lifshitz, *Statistical Physics*, 3d ed., 1980; R. C. Tolman, *The Principles of Statistical Mechanics*, 1938, reprint 1980; D. H. Trevens, *Statistical Mechanics: An Introduction*, 1993.

Statistical process control

A collection of strategies and statistical methods for ensuring high-quality results in manufacturing and service processes. Traditionally, the statistical methods associated with statistical process control (SPC) have focused on process monitoring. As a result, some experts argue that statistical process control is a misnomer and that the term statistical process monitoring is more appropriate. Others argue that statistical quality control methods have become so entwined with strategic decisions and daily operations that the term statistical process control is entirely appropriate. Regardless, statistical process control methods and philosophies are essential tools for survival in today's highly competitive environment. *See* PROCESS CONTROL; QUALITY CONTROL; STATISTICS.

The origins of statistical process control coincide with a shift in focus by manufacturing companies from quantity to quality. The pioneers in this transformation include W. Shewhart, W. E. Deming, J. Juran, K. Ishakawa, G. Taguchi, and others. At the heart of this transformation were two simple ideas: understanding the customer's needs and meeting the customer's needs consistently. Accomplishing both of these tasks while minimizing wasted time, effort, and resources insured a competitive advantage. The success of organizations employing quality-oriented management styles led to a revolution in management tools and philosophies.

Total quality management. Total quality management (TQM) gained strength in the 1980s. The expression total quality management refers to both the philosophy and practices associated with continuously improving processes within a company. An all-encompassing view of processes included not

only manufacturing operations but also service areas such as sales and accounting functions. Companies adopted various strategies for continuous improvement with varied levels of commitment under an assortment of names. Total quality control, continuous quality improvement, and total quality excellence are synonyms for total quality management. These programs had several common themes. An important commonality was reliance upon actual process data to monitor systems, inform decision makers, and guide the continuous improvement activities. The tools associated with statistical process control are fundamental building blocks in these quality-driven programs. International standards were eventually established to ensure that total quality management programs, while varied, followed certain basic principles and practices. The standards known as ISO 9000:2000 are a product of the International Standards Organization (ISO), the American Society for Quality (ASQ), and the American National Standards Institute (ANSI).

Six Sigma. The ideas surrounding total quality management grew into the Six Sigma quality movement, which continues to evolve. Among the unique features of Six Sigma programs are the certification levels for managers trained in statistical problem solving and the use of the DMAIC process. The acronym DMAIC refers to five steps: define/deploy, measure, analyze, improve, and control.

The certification process begins with training in basic statistical methods and continuous quality improvement concepts. Managers mastering these initial topics are called "green belts." Through a series of training classes and real-time projects, managers ascend to the ranks of "black belt" or "master black belt." Managers receiving the higher-level designations must demonstrate an understanding of the philosophies and tools of continuous improvement, as well as the ability to put these ideas into practice with substantial success, as measured in long-term savings in real dollars.

The DMAIC problem-solving approach is central to Six Sigma's successful use of statistical process control methods throughout an organization. The define/deployment phase involves selecting a project leader (black belt), defining the project, and creating a project support structure. The measurement phase captures current project-related wisdom. Helpful tools used in this phase include many traditional quality control tools such as flow charts, check sheets, cause-and-effect diagrams, histograms, scatter plots, and control charts. These tools help identify key input and output variables used to control and measure process performance. The analysis phase provides a comprehensive understanding of process data and procedures. Process changes (improvements) are then recommended and tested using designed experiments or optimization techniques. Changes that result in improvements are permanently incorporated in the control phase, as are activities such as determining control plans, updating standard operation procedures, and mistake-proofing processes. *See* CONTROL CHART.

The preceding paragraphs highlight the seamless incorporation of statistical tools into the day-to-day working of management. Traditional quality control techniques such as control charts are now complemented by appropriate statistical methods such as designed experiments and modern technologies. Hence, statistical process control tools are used beyond quality engineering. The result is a blurring of the boundaries that traditionally defined statistical process control.

Quality control charts. All processes display variability. Processes include all meaningful activities that may be measured in some way. The process of getting to work in the morning or serving a customer may be measured in time. The process of surgery may be measured by survival rates, recovery times, or infection rates. The process of manufacturing an object may be measured by yield rates, dimensional properties of the object, or conformance to customer specifications. Determining whether a process is performing in a stable and predictable manner or displaying unusual behavior is the traditional domain of statistical process control.

All measurable processes exhibit variation that can be characterized as belonging to one of two classes: common-cause variation and special-cause variation. Common-cause variation is often unexplainable and unavoidable without fundamental changes to the underlying process. Special-cause variation is attributable to a specific cause that may be identified and corrected. The quality control chart is used to detect and address special-cause variation. Additionally, the control chart is used to quantify and avoid overreaction to common-cause variation. Overadjustment of equipment by operators can result in a deterioration of overall quality.

Underlying logic. To illustrate the usefulness of control charting, consider a process for manufacturing an automobile part. The logic of control charting is simple, as outlined in the following steps:

1. Select an appropriate measure for a given quality characteristic. One may be interested in the length of the manufactured part.

2. Collect samples at regular intervals. One might randomly select four parts every 2 hours for inspection.

3. Calculate an appropriate statistic from the sample data. The sample mean (average of the lengths for the four parts) would help determine if parts are of appropriate lengths. The sample range (difference in the lengths of the longest and shortest parts in the sample) would help determine if the parts are consistently the correct length.

4. Plot the calculated statistic(s) over time and compare the plotted values to predetermined values, called control limits. Unusual patterns or extreme observations beyond the control limit could indicate special-cause variation to be identified and corrected by operators. Such patterns or extreme values are called signals. Values falling between the control limits indicate common-cause variation, and

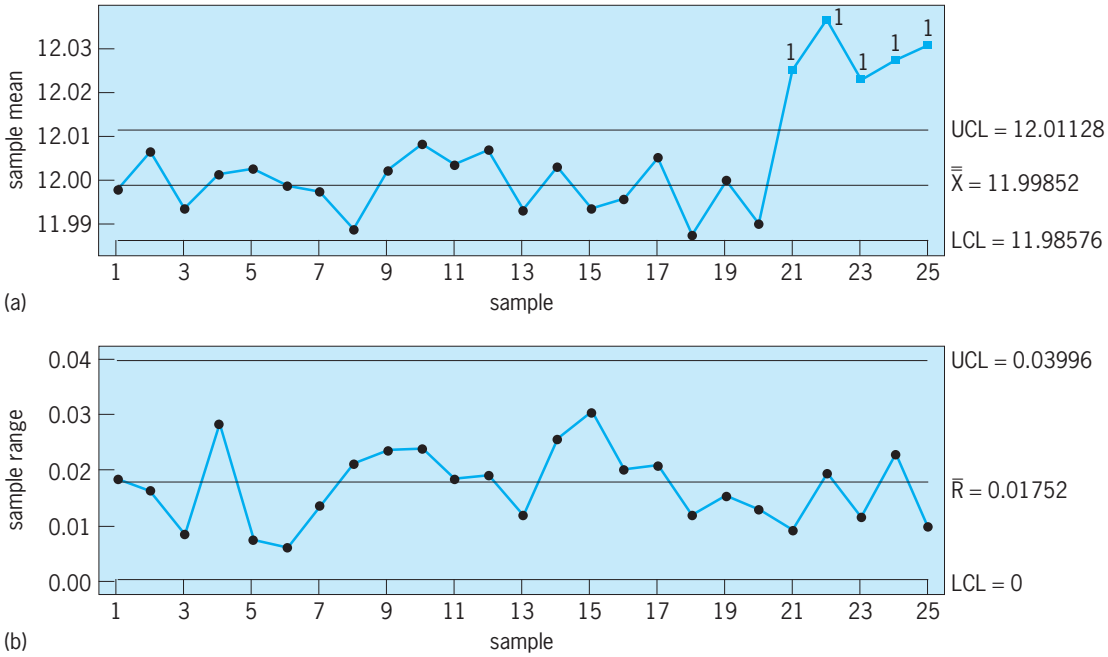

Control charts for length: (*a*) mean and (*b*) range. UCL = upper central limit. LCL = lower central limit. $\bar{\bar{X}}$ = average of all $\bar{X}$ points. $\bar{R}$ = average of all R points.

operators would allow the process to run without interruption.

The control chart is displayed graphically to allow easy use by operators. For example, the **illustration** shows the control charts for the mean length of parts and the range in the length of parts. For each sample, the average length for four randomly selected parts is calculated and plotted on the mean chart, and the sample range for the lengths is plotted on the range chart. Observations falling beyond the control limits of a control chart are called signals. The mean chart signal at time = 21 suggests that something unusual has occurred that requires the operator's attention. No values are beyond the control limits of the range chart, which indicates that the problem affects only the part length, not the process consistency (that is, the process is consistently making bad parts).

Guidelines for important questions. While the steps outlined in the previous section appear straightforward, careful consideration is required at each step.

Step 1: Get the right data. Selecting the best measurement method for a quality characteristic with regard to process information and practical considerations, such as safety, time, and ease of obtaining measurements, is obviously process-specific. There are, however, guidelines governing variable selection.

Control charts typically fall into two groups: attributes and variables. Attribute data involve counts or proportions. For example, a sample of 100 items might be inspected and the number of items not meeting customer requirements reported. Each item would be simply classified as good or bad depending on some quality attribute (length is acceptable or not acceptable). On the other hand, continuous variable measurements involve lengths, weights, times, or measurements that could obtain any value within a given range. For example, actual lengths of parts might be recorded. Continuous variable measurements are inherently more informative than attribute measurements. For example, suppose customer specifications require that parts have lengths of 12 ± 0.03 in. Simply knowing that a part is not the correct length is less informative than knowing that the part length is 12.04 in., which indicates both the direction and magnitude of the problem. The additional information contained in variable measurement allows more powerful monitoring schemes and helps operators in diagnosing problems. Continuous measurements should be used where possible. *See* INSPECTION AND TESTING.

Step 2: Get the right amount of data from the right place. The idea of rational subgroups is helpful in determining sampling locations. One should collect data at locations that allow operators to isolate potential problems. For example, if three production lines produce identical parts, sampling at the end of each line is preferred to sampling after parts from the three lines have been merged. After the parts have been mixed together, it may be difficult to determine which production line is experiencing problems.

Selecting how many items should be in each sample and how often samples should be collected requires technical knowledge of control chart performance. Generally, small samples taken frequently are preferred to large samples taken infrequently.

Step 3: Summarize the data in an appropriate manner. This involves the statistics that should be calculated. Data summaries (sample statistics) should answer two questions: Is the process on target (for example, parts are on average the correct length)? Is the process behaving in a consistent manner?

Continuous variables data provide better answers to these questions than do attribute data. Data summarized by averaging sample values (or weighted averages of past values) are used to determine if a process is on target. Practitioners refer to these statistics as measures of location. The sample range or sample standard deviation is useful in quantifying consistency of process behavior. Such statistics are referred to as measures of variation.

Step 4: Plot the summary statistics and control limits. The value(s) of the control limit(s) depends on the statistic being plotted, the sample size, and the tolerable frequency of false alarms. A signal on a control chart is called a false alarm if no actual problem exists. The average number of samples required for a control chart to signal a process problem is called the average run length (ARL). A large ARL is desired for problem-free (in-control) processes, while small ARLs are desired for processes with problems (out-of-control processes). General guidelines for determining control limits are provided in most introductory texts on statistical process control methods.

Additional comments. The appropriate selection and design of control charts has received much attention over the years. Performances of the most common control charts are well understood. Information concerning the performance and selection of control charts may be found in many textbooks. In addition to common control charts, there are many specialized monitoring techniques available to practitioners. Multivariate control charts are used when several important process variables are related to one another. Forecast-based or regression-based charts are useful when data values change systematically over time. Process change due to tool wear is an example. Practitioners are advised to select an appropriate technique for each process.

A second area of caution associated with control charting is the relationship between control limits and specification limits. The relationship is extremely simple because there is no relationship. Specification limits reflect the customer's requirements. Control limits reflect the common-cause variation in a process and the statistic plotted on the control chart. The customer's needs are not determined by the process, nor are process variations determined by the customer.

The primary goals of control charting are to identify and eliminate special causes of variation, ensure that the process remains in control, and quantify common-cause variation. Once a process is behaving in a stable and predictable manner, one may evaluate the capability of the process and use designed experiments to gain further improvements.

Capability indices. One is always interested in knowing if inherent process variation allows one to meet customer needs. A capability index is a numerical measure relating the demands of the customer to the capability of the producer. There are many versions of capability indices, just as there are many types of control charts. The appropriate index depends on the properties displayed by process data. Most capability indices are a ratio of customer specifications and common-cause process variation. For example, one capability index is defined by

$$C_p = \frac{(\mathrm{USL} - \mathrm{LSL})}{6\hat{\sigma}}$$

where $\hat{\sigma}$ represents the estimated standard deviation associated with common-cause variation, and USL and LSL represent the customer's upper specification limit and lower specification limit, respectively. A large C_p value indicates that the inherent process variation ($6\hat{\sigma}$) is small relative to customer specification limits (VSL−LSL). Thus, the process is very capable of meeting customer requirements.

A small capability ratio indicates that a process is not capable of meeting customer requirements without fundamental changes in the process and/or operation procedures. Process capability must be improved by identifying the most efficient and appropriate process changes that ensure that the output is on target, with minimal variation.

Designed experiments. Designed experiments are powerful techniques for improving process capability. Statistically designed experiments allow systematic investigation of factors that affect process performance. These efforts can improve processes in several ways. Greater understanding of process variables aids diagnostic efforts, leads to process optimization, and increases process robustness to nuisance variables that are difficult or impossible to control. As processes are improved, control charts must be updated to reflect the new and improved process behavior. Process capability ratios must be recalculated to reflect the increased process capability.

The cycle of monitoring and stabilizing processes, determining their capability, and improving process capability through statistically designed experiments is a powerful approach to quality improvement. Understanding customer needs in a quantifiable manner, communicating these requirements to operations, and focusing attention on consistently meeting customer needs seem only logical. Statistical process control techniques not only maximize the chance of gaining a competitive advantage but also allow gaining an advantage at a minimum of expense. These techniques have been useful in pure research and development for many years. As understanding of their usefulness grows statistical process control methods will remain a key tool in the management of manufacturing and service processes. *See* EXPERIMENT; INDUSTRIAL ENGINEERING; MANUFACTURING ENGINEERING.

Benjamin M. Adams

Bibliography. E. Grant and R. Leavenworth, *Statistical Quality Control*, 7th ed, 1996; G. Gryna, *Quality Planning & Analysis: From Product Development Through Use*, 4th ed., 2001; J. M. Juran and A. B. Godfrey, *Juran's Quality Handbook*, 5th ed., 1999; D. Montgomery, *Introduction to Statistical Quality Control*, 5th ed., 2005; T. Pyzdek, *The Six Sigma Handbook, Revised and Expanded: The Complete*

Guide for Greenbelts, Blackbelts, and Managers at All Levels, 2d ed., 2003; K. Yang and B. S. El-Haik, *Design for Six Sigma: A Roadmap for Product Development*, 2003.

Statistics

The field of knowledge concerned with collecting, analyzing, and presenting data. Not only workers in the physical, biological, and social sciences, but also engineers, business managers, government officials, market analysts, and many others regularly use statistical methods in their work. The methods range from simple counting to complex mathematical systems designed to extract the maximum amount of information from very extensive data.

In an important sense statistics may be regarded as a field of application of probability theory. The common problem faced by a physicist reading a meter, an engineer testing a material, an agronomist measuring the yield of a hybrid corn, a chemist determining the concentration of ascorbic acid, and an interviewer studying public opinion is the problem of random variation which prevents repetition of exactly the same result when a measurement is repeated. Statistical methods are employed to assess the magnitude of random variation, to minimize it, to balance it out, to remove it by calculation procedures, and to analyze it by suitably arranged patterns of observation. The theory of probability is concerned with the properties of random variables and hence furnishes the basis for developing techniques for controlling them. *See* PROBABILITY.

Viewing statistics from another direction, it is the science of deriving information about populations by observing only samples of those populations. A population is any well-specified collection of elements. Thus, one may refer to the population of adults in the continental United States viewing television screens at 8:14 P.M. on August 6, 1970; the population of automobiles less than 2 years old registered in Los Angeles County on a certain date; the population of vineyards in France; the hypothetical population of outcomes of tossing a given coin endlessly. Populations may be finite or infinite. An element of a univariate population is characterized by the value of a random variable which measures some single attribute of interest in the population. Thus, one may be interested in whether or not individuals of the television audience were or were not viewing program A; with each individual one may associate a random variable, let it be X, which takes on the value of 1 if the individual is watching A and 0 if he is not. If one were interested in a second characteristic of the elements of the television audience (such as age), one would be said to be dealing with a bivariate population; a third characteristic (such as economic status) would make it a trivariate or, less specifically, a multivariate population.

Random variables are either continuous, which means they can take on any numerical value (the length of a room), or discrete, which means they can take on only a restricted set of values (number of windows in a room).

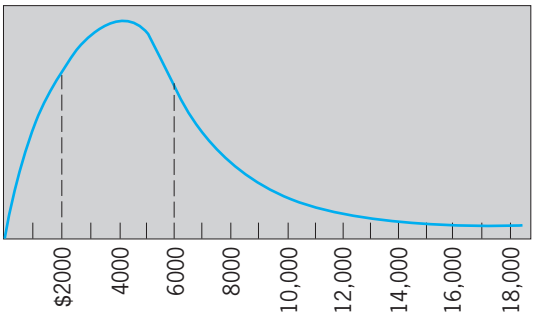

Fig. 1. Distribution of incomes.

Distributions. In a univariate population, the population distribution is a curve (function of the random variable which characterizes the elements of the population) from which one can determine the proportion of the population which has elements in a certain range of the random variable. For example, the curve of **Fig. 1** provides the distribution of annual incomes of family units. The total area under the curve is 1. The area under the curve between any two vertical lines gives the proportion of the families having annual incomes between the two values marked on the horizontal scale by the two vertical lines. Thus, the fact that the area under the curve between $2000 and $6000 is 0.541 means that 54.1% of the family units have incomes in that range.

The distribution is also referred to as the distribution function, the density function, the frequency function, or the probability density.

The total area under the distribution curve to the left of each point can also be plotted to give a curve which starts at zero and reaches unity as the variable becomes large; the resulting curve is sometimes called the cumulative distribution function, the probability distribution, or simply the distribution. The cumulative form of the curve of Fig. 1 is shown in **Fig. 2**; the height of the curve at any point on the horizontal scale equals the area to the left of that point under the curve of Fig. 1 and is the proportion of the population having incomes less than the value at that point. The distribution (in either frequency or cumulative form) gives complete information about the way the characterizing variable is spread through the population.

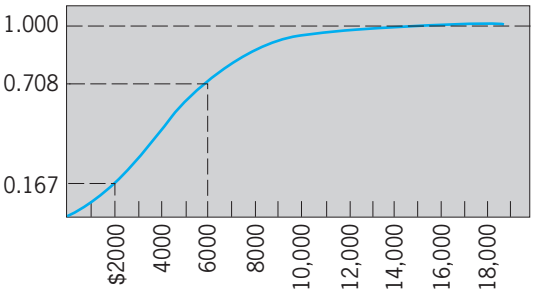

Fig. 2. Cumulative distribution of incomes.

Population parameters. Populations (or population distributions) are often specified incompletely by certain population parameters. Some of these parameters are location parameters or measures of central tendency; a second class of important parameters consists of measures of dispersion or scale parameters.

The most widely used location parameters are the mean, the median, and the mode. The mean is the average over all the population of the values of the random variable. It is often represented by the Greek letter μ. In mathematical terms, if x is the random variable, $f(x)$ the frequency function for a given population, and $F(x)$ its cumulative form, then the mean is as shown in Eq. (1).

$$\mu = \int_{-\infty}^{\infty} x f(x)\, dx = \int_{-\infty}^{\infty} x\, dF(x) \qquad (1)$$

The median, often designated by Med, M, or $X_{.50}$, is a number such that, at most, one-half the values of the variable associated with the elements of the population fall above or below it, as in Eq. (2). The mode

$$\int_{-\infty}^{M} dF(x) \geq {}^1\!/_2 \leq \int_{M}^{\infty} dF(x) \qquad (2)$$

is the most frequent value of the random variable; if the frequency function has a unique maximum value, the mode is the value of the random variable at which the frequency function reaches its maximum. Location parameters are numbers near the center of the range over which the random variable of the population varies; different ones arise from different definitions of center; generally the mean is used unless special circumstances make some other location parameter more appropriate.

The extent to which a population is scattered on either side of its center is roughly indicated by measures of dispersion such as the standard deviation, the mean deviation, the interquartile range, the range, and sometimes others. The standard deviation is the square root of the mean square of the deviations from the mean; it is usually denoted by the Greek letter σ; σ^2 is called the variance and is expressed by Eq. (3). The mean deviation, shown in Eq. (4), is the average over the population of the

$$\begin{aligned} \sigma^2 &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx \\ &= \int_{-\infty}^{\infty} (x-\mu)^2\, dF(x) \qquad (3) \end{aligned}$$

$$\begin{aligned} \text{Mean deviation} &= \int_{-\infty}^{\infty} |x-\mu|\, f(x)\, dx \\ &= \int_{-\infty}^{\infty} d\,|x-\mu|\; dF(x) \qquad (4) \end{aligned}$$

deviations from the mean, all taken to be positive. The interquartile range (often denoted by Q) is the difference $X_{.75} - X_{.25}$, where $X_{.75}$ is the value of the random variable such that one-quarter of the population has values larger than $X_{.75}$, and $X_{.25}$ is the number such that one-quarter of the population has values smaller than $X_{.25}$. The three numbers, $X_{.25}$, $X_{.50}$, $X_{.75}$, are called quartiles; these divide the population into quarters. The range is the difference between the largest and the smallest of the population elements.

Samples. If one examines every element of a population and records the value of the random variable for each, complete information is obtained about the distribution of the random variable in the population, and there is no statistical problem. It is usually impossible or uneconomical to make a complete enumeration (or census) of a population, and one must therefore be content to examine only a part or sample of the population. On the basis of the sample, one draws conclusions about the entire population; the conclusions thus drawn are not certain in the sense that they would likely have been somewhat different if a different sample of the population had been examined. The problem of drawing valid conclusions from samples and of specifying their range of uncertainty is known as the problem of statistical inference.

Statisticians distinguish two kinds of samples. A survey sample is one chosen from a population, all the elements of which actually exist. An experiment is a sample chosen from a hypothetical population. Thus if one was interested in the total number of pigs being fattened in a given season by the farmers of the state of Iowa, the researcher would survey-sample the existing population of pigs in Iowa at that time. If the effect of a certain hormone on the growth rate of pigs was the subject, an experiment would be performed by giving a group of pigs the hormone for a period of time; using the sample of observations obtained from the experiment, the researcher would draw inferences about the hypothetical population of observations that would have resulted had all pigs been given the hormone.

Random sampling. In planning a sample survey, the manner in which the data to be gathered will fulfill the purpose should be clearly stated. The population to be sampled must be explicitly defined. The method of sampling should be efficient and lead to a straightforward analysis. The question of what elements should be included in the population depends on the purpose of the survey. Thus the population of vineyards in France might include as a vineyard a dozen vines in the backyard of a person living in the heart of Paris if the purpose were to estimate total grape production of France; but the population might be defined to exclude such a small vineyard if the purpose were to estimate the size of the harvest to be available to commercial wineries.

The type of sampling of interest in statistics is probability sampling, because it eliminates subjective aspects from the selection of the sample. In probability sampling, all possible distinct samples are known, the selection of the sample is done randomly according to a preassigned probability, and the method of analysis is predetermined and unambiguous. Only from such samples can inferences about populations be made with measurable precision.

Simple random sampling is a method of selecting a sample of n elements out of a population of N elements so that all such samples have an equal probability of being drawn. This may be done by selecting a first element at random from the population, then a second element at random from the remaining population, and so on until the n elements are selected. Because an element cannot appear more than once in the sample, this is a form of sampling without replacement. The sampling ratio or sampling fraction is n/N.

Often the purpose of drawing a sample is to estimate the mean or average of a characteristic of the population. If y_i is the value of the characteristic of the ith unit, then the population mean is as shown in Eq. (5). The sample mean is the average of the n units

$$\mu = \frac{1}{N}\sum_{i=1}^{N} y_i$$

$$= \frac{1}{N}(y_1 + y_2 + y_3 + \cdots + y_N) \tag{5}$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}(x_1 + x_2 + \cdots + x_n) \tag{6}$$

in the sample and is expressed in Eq. (6). Here the n x's are the n y's selected from the population as the sample. The population total is $N\mu$ and is estimated by $N\bar{x}$. The population variance is defined as shown in Eq (7), and the sample variance is usually defined as in Eq. (8), although some authors use n instead

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - \mu)^2 \tag{7}$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{8}$$

of $n - 1$. The quantities $\bar{x}$ and s^2, the sample mean, and the sample variance are called sample statistics; they are the first and second sample moments and are also estimators of the corresponding population parameters, μ and σ^2.

The estimators are themselves random variables. If one repeatedly drew samples of size n from the population and computed $\bar{x}$ from each sample, he would obtain a population of $\bar{x}$'s with its own distribution which would differ from the distribution of x. It can be shown that the $\bar{x}$ population has exactly the same mean μ as the x population. Further, the variance of the $\bar{x}$ population is as shown in notation (9), where

$$\frac{(N-n)}{(N-1)}\frac{\sigma^2}{n} \tag{9}$$

σ^2 is the variance of the x population. The first fraction is ordinarily nearly unity so that the variance of $\bar{x}$ is approximately the fraction $1/n$ of the original population variance. As n, the sample size, becomes large, the $\bar{x}$ population becomes more concentrated about μ and the reliability of a particular value of $\bar{x}$ as an estimate of the population mean increases.

The observations of a sample, besides providing estimates of population parameters, can also be used to obtain an estimate of the population's frequency function. This estimate is determined by dividing the range of the sample observations into several intervals of equal length L and counting the number of observations occurring in each interval; these numbers are then divided by nL to determine fractions giving the relative density of the sample occurring in each interval; then on a sheet of graph paper one lays out the intervals on a horizontal axis and plots horizontal lines above each interval at a height equal to the fraction corresponding to the interval; finally the successive plotted horizontal lines are connected by vertical lines to form a broken line curve known as a histogram (**Fig. 3**). The area under the curve is unity, and the area between any two points gives the fraction of the sample observations lying between those two points. If one takes larger and larger samples, the chosen intervals can be made smaller and the broken line curve will come closer and closer to the underlying population frequency function. Often one does not trouble to divide the interval frequencies by nL to normalize the area of the histogram, but merely plots the frequencies themselves; the resulting broken line curve is still referred to as a histogram.

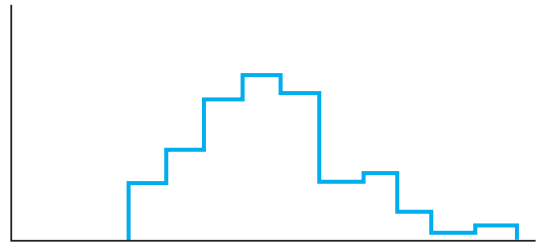

Fig. 3. Histogram.

Sampling techniques. When a population can be regarded as being made up of several nonoverlapping subpopulations, one may draw a sample from it by drawing a simple random sample from each subpopulation (or stratum); this procedure is called stratified random sampling. The method is employed when it is desired to have specific information about each stratum individually, when it is administratively convenient to subdivide the population, when there are natural strata, or when a gain in precision would be realized because each stratum is more homogeneous than the whole population.

Let n be the number of units sampled in the entire population of N units, and let n_b be the number of units sampled in stratum b containing N_h units. Therefore, the sum of all the n_b is n, and of all the N_b is N. Let σ_b be the true standard deviation within stratum b. When $n_b/n = N_b/N_n$, the n_h are said to be proportionately allocated. When $n_b/n = N_b\sigma_b/\Sigma N_b\sigma_b$, the n_b are said to be optimally allocated because the variance of the estimated mean is then smallest for fixed total size of sample. Ordinarily, if the σ_b are well estimated, optimum allocation gives greater precision than proportional allocation, and proportional allocation gives greater precision than simple random sampling.

Systematic sampling may be regarded as a sampling of units at regular intervals in the population,

for example, every tenth unit. Usually this is done with a random start; that is, the first unit in the sample is selected at random from a small group at the beginning. A systematic sample is usually relatively easy and fast to execute, and if the analogy to stratified sampling is valid, a systematic sample will be more accurate. However, because of the essentially nonrandom nature of a systematic sample, it is difficult to estimate the sample variance unless certain assumptions are made about randomness in the order of the population. If the population has any periodicity, then the estimates made from a systematic sample can be quite poor.

When the elements of a population are in mutually exclusive groups called the primary units of the population, then a sample of these primary units might be made, and then, from those selected, a sample of the individual elements would be made. This procedure is called two-stage sampling or subsampling. Multistage sampling can involve more than two stages of sampling. Although more complicated and difficult to apply and to analyze, multistage sampling offers considerable flexibility for balancing between statistical precision and the cost of sampling.

When too little is known about a population to plan a sample, then two samples are made. The manner in which the second sample is taken is determined from the results of the first sample. In Stein's method of two-stage sampling, the size of the additional sample is decided by using the estimate of the variance from the first sample. In double sampling or two-phase sampling, the first sample is used to gather information on a variate associated with the variate of interest. This information is used to establish strata for drawing a stratified random sample involving the variate of interest.

Ratio and regression estimates are often used when observations are made on one or more variates related to the variate of primary interest.

Sequential sampling and sequential analysis refer to a method of sampling in which the size of the sample is not specified in advance. Observations are drawn one by one until a specified degree of confidence in the information to be obtained has been achieved.

Sampling distributions. Many important sampling distributions are derived for random samples drawn from a normal or gaussian distribution, which is a bell-shaped symmetrical distribution centered at its mean μ. The distribution is illustrated in **Fig. 4** for three different values of σ, the standard deviation. Equation (10) is the mathematical equation for these

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-u)^2/\sigma^2} \tag{10}$$

curves. The area under the curves for the indicated limits is as follows:

Area	*Limits*
0.50	$\mu - 0.675\sigma$ to $\mu + 0.675\sigma$
0.683	$\mu - \sigma$ to $\mu + \sigma$
0.90	$\mu - 1.645\sigma$ to $\mu + 1.645\sigma$
0.95	$\mu - 1.960\sigma$ to $\mu + 1.960\sigma$
0.99	$\mu - 2.326\sigma$ to $\mu + 2.326\sigma$

Referring to the above limits, 0.675σ is called the probable error inasmuch as the odds are even that a randomly drawn observation will lie within 0.675σ of the mean.

If samples of size n are drawn from a normal population and the sample mean $\bar{x}$ is computed for each, the $\bar{x}$'s will have a normal distribution with the same mean and with variance σ^2/n. Thus, if samples of size 4 were drawn from a population distributed by the curve marked $\sigma = 1$ in Fig. 4, then the means of those samples would be distributed by the curve marked $\sigma = 0.5$ (that is, $1/\sqrt{4}$). Further, the probability would be 0.95 that a particular $\bar{x}$ so drawn would lie with $(1.960)(0.5) = 0.98$ of the population mean μ.

The central limit theorem of the theory of probability states that under very general conditions the sample mean $\bar{x}$ is approximately normally distributed whatever may be the distribution function for the underlying population. This powerful theorem enables one to make probability statements such as the one at the end of the preceding paragraph in ignorance of the actual population distribution.

Besides the distribution of the mean, two other sampling distributions derived from the normal distribution have wide application. One is the chi-square distribution, χ^2, which provides the distribution of the sample variance, as shown in Eq. (11).

$$s^2 = \sum_{i=1}^{n} \frac{(x_1 - \bar{x})^2}{n-1} \tag{11}$$

Here $x_1, x_2, \ldots, x_n$ are the observations of a random sample of size n drawn from a normal population.

The quantity given in Eq. (12) has a distribution

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \tag{12}$$

illustrated in **Fig. 5** for three values of n. The mathematical form for the distribution curve is shown in notation (13) and is often referred to as the chi-square

$$\frac{1}{[(n-3)/2]!}\left(\frac{n-1}{2\sigma^2}\right)^{(n-1)/2} (\chi^2)^{(n-3)/2} e^{-(n-1)\chi^2/2\sigma^2} \tag{13}$$

distribution with $n - 1$ degrees of freedom.

The most useful random variable for interval

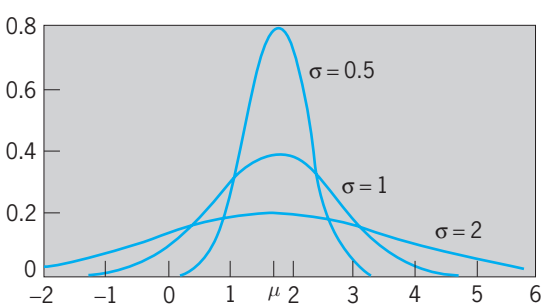

Fig. 4. Normal distribution.

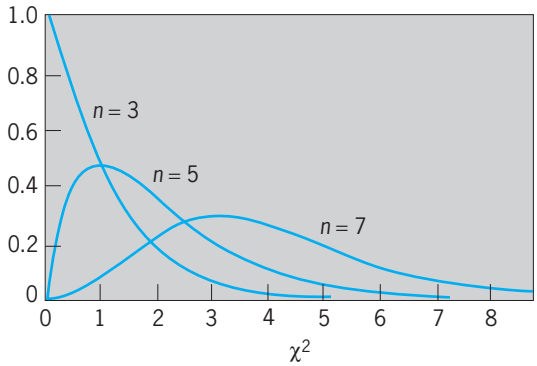

Fig. 5. Chi-square distributions.

estimation of a population mean is shown in Eq. (14),

$$t = \frac{\sqrt{n}(\bar{x} - \mu)}{s} \tag{14}$$

which has a symmetrical distribution very similar in appearance to the curves plotted in Fig. 4. The mathematical form for the distribution is notation (15). This is referred to as the t distribution or the Student

$$\frac{[(n-2)/2]!}{\sqrt{(n-1)\pi}\,[(n-3)/2]!} \Big/ \left(1 + \frac{t^2}{n-1}\right)^{n/2} \tag{15}$$

distribution with $n - 1$ degrees of freedom. The following gives limits which include 95% of the area under the curve for a few values of n:

n	*Limits*
2	−12.71 to + 12.71
3	−4.30 to + 4.30
4	−3.18 to + 3.18
10	−2.20 to + 2.20
25	−2.06 to + 2.06
120	−1.98 to + 1.98

For very large sample sizes, the limits become −1.96 to +1.96, as for the normal distribution. An illustration of the use of these limits is given below.

Estimation. In making an estimate of the value of a parameter of a population from a sample, a function (called the estimator) of the observations is used. For example, for estimating the mean of a normal population the mean of the sample observations is usually taken as the estimator. Another estimator is the average of the two most extreme observations. In fact, there is an infinitude of estimators. The problem of estimation is to find a "good" estimator.

A good estimator may be regarded as one which results in a distribution of estimates concentrated near the true value of the parameter and which can be applied without excessive effort. There is no single way of deciding how good an estimator is, but there are several criteria by which an estimator may be judged.

An unbiased estimator is one which results in a distribution of estimates which has a mean exactly equal to the value of the parameter being estimated. Otherwise the estimator is called biased. The bias is the mean of the distribution of the estimator minus the value of the parameter it estimates.

An estimator is said to be consistent if the probability that an estimate will differ from the value of the parameter by more than any fixed amount can be made arbitrarily small by increasing the number of observations.

The variance of an estimator is the mean squared deviation of the estimates from the value of the parameter. The estimator with the smallest variance is called most efficient. The relative efficiency of two estimators is the ratio of the variances. When the numerator of this ratio is the variance of a most efficient estimator, this ratio is simply called the efficiency of the other estimator.

An estimator is said to be sufficient if it contains all the information in the sample regarding the parameter. This is so when the conditional distribution of the sample for a given estimate is independent of the parameter.

There are several methods of constructing estimators of parameters. The method of moments is applied by assuming that the first few sample moments are equivalent to the moments of some distribution, and then solving for the parameters of that distribution. The methods known as least squares, minimum variance, and minimum chi-square all have as their basis the estimation of the values of the parameters which minimize some linear function of the squares of deviations of the observations from the values of the parameters. In applying Bayes' method, the distribution of possible values of the parameter before the sample is taken, called the a priori distribution, is used in conjunction with the observations in the sample to yield an estimator. The method of maximum likelihood uses as the estimate that value of the parameter for which the probability of the sample is highest.

A confidence interval is an interval constructed in such a way that the true parameter value is within this interval with a predetermined probability in repeated sampling; this probability is called the confidence level of the interval. For example, a sample mean μ known to be normally distributed with variance σ^2 will differ less than 1.96σ from the true but unknown mean μ with probability 0.95. Thus expression (16) can be formulated. On solving the two inequalities for μ, expression (16) may be written as expression (17).

Probability that

$$(-1.9\sigma < \bar{x} - \mu < 1.96\sigma) = 0.95 \tag{16}$$

Probability that

$$(\bar{x} - 1.96\sigma < \mu < \bar{x} + 1.96\sigma) = 0.95 \tag{17}$$

When a particular value of x^- is computed from the observations, then $x^- \pm 1.96\sigma$ is a pair of numbers which define a particular interval called the confidence interval. Because μ is a fixed number, it is within this particular interval with probability zero or one. However, μ is unknown, so that the confidence to be associated with the interval is stated in terms of the proportion of all intervals constructed in this manner which would include μ, rather than

this particular interval. A probability of this sort, valid for a population of outcomes, when used in reference to a particular outcome is called a fiducial probability.

Ordinarily σ is unknown and the estimate s derived from the sample must be used to form a confidence interval; in this case, the t distribution rather than the normal distribution must be used. As an example, suppose a chemist has made four determinations of the atomic weight μ of hydrogen as follows: 1.0066, 1.0090, 1.0084, and 1.0086. The average of these values is given in Eq. (18), and the sample estimate of

$$\bar{x} = {}^1/_4(1.0066 + 1.0090 + 1.0084 + 1.0086) = 1.0082 \tag{18}$$

the variance of his technique is as shown in Eq. (19), the estimate of σ being therefore Eq. (20). It follows

$$s^2 = {}^1/_3[(1.0066 - 1.0082)^2 + (1.0090 - 1.0082)^2 + (1.0084 - 1.0082)^2 + (1.0086 - 1.0082)^2] = .00000123 \tag{19}$$

$$s = \sqrt{0.00000123} = 0.0011 \tag{20}$$

from the definition of t in the preceding section that in this case Eq. (21) holds. Using the short t table

$$t = \frac{2(1.0082 - \mu)}{0.0011} \tag{21}$$

given in the preceding section, one finds that expression (22) can be formulated. Solving these inequalities for μ gives expression (23), and the chemist can

$$\text{Probability that } \left(-3.18 < \frac{2(1.0082 - \mu)}{0.0011} < 3.18\right) = 0.95 \tag{22}$$

$$\text{Probability that } (1.00645 < \mu < 1.00995) = 0.95 \tag{23}$$

assert with 95% confidence that the atomic weight of hydrogen lies between 1.00645 and 1.00995. For greater precision, more observations are required.

In a similar fashion one may obtain a confidence interval for σ by using s^2 in connection with the chi-square distribution.

A confidence region is the generalization of a confidence interval and refers to the simultaneous estimation of several population parameters. The confidence level is the proportion of the time that the region actually includes the true values of the parameters.

In general, the most desirable confidence interval or region is the smallest one which can be constructed for the selected confidence level.

Tests of hypotheses. Besides estimation of parameters, another major area of statistical inference is the testing of hypotheses. A hypothesis is merely an assertion that a population has a specific property. The test consists of drawing a sample from the population and determining whether or not it is consistent with the assertion. Very often the hypothesis is a statement about the mean of a population; that it has a given value, that it is the same as that of another population, that it exceeds that of another population by at least 10 units, and the like. Thus, one may be comparing a new blend of gasoline with a current blend, a new drug with a standard one, a new manufacturing process with an existing one.

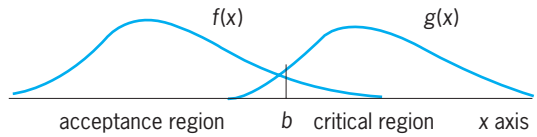

Fig. 6. Simple test of a hypothesis.

For a very simple illustration of the basic ideas involved, suppose a population is known to have either of two distributions which differ mainly in location. Let the random variable be x and let the two functions of x which determine the two distributions be represented by the symbols $f(x)$ and $g(x)$ as indicated in **Fig. 6**; that is, $f(x)$ represents the height of the left curve at x, and $g(x)$ represents the height of the right curve at x. Assume that the population has the distribution $f(x)$; there is in this case only a single alternative, $g(x)$, so that rejection of the hypothesis implies acceptance of the alternative. Suppose that for testing the hypothesis one can afford only a single observation, that is, a sample of size one.

It is obvious in this instance that one should choose some number b in advance, then accept the hypothesis if the observation falls to the left of b and reject it if the observation falls to the right. Values of x to the left of b are said to constitute the acceptance region of the test; values to the right of b constitute the critical region. The area under $f(x)$ to the left of b, which is denoted mathematically by notation (24), is

$$\int_{-\infty}^{b} f(x)\,dx \tag{24}$$

the probability that the hypothesis will be accepted when it is true. The area to the right of b under $f(x)$ is the probability that the hypothesis will be rejected when it is true; this probability is called the type I error of the test. The area to the left of b under $g(x)$ is the probability that the hypothesis will be accepted when it is false and is called the type II error of the test. The area to the right of b under $g(x)$ is called the power of test and is, of course, one minus the type II error.

By moving the point b to the right one can make the type I error as small as one likes, but in doing so the type II error is increased. This dilemma is characteristic of the construction of tests of hypotheses. Ordinarily one arbitrarily chooses the type I error to be some small number such as 0.05 or 0.01 and then chooses the critical region so as to minimize the type II error.

The fraction $f(x)/g(x)$ is called the likelihood ratio or simply the likelihood function of x. From Fig. 6 it is evident that the likelihood is relatively large when the hypothesis is true and small when it is false. For

a sample of size n with observations $x_1, x_2, \ldots, x_n$ the sample likelihood is defined to be the product of the individual likelihoods given in expression (25). Let all possible samples of size n be divided

$$\frac{f(x_1)}{g(x_1)}\frac{f(x_2)}{g(x_2)}\cdots\frac{f(x_n)}{g(x_n)} \tag{25}$$

into two sets, with one set (the acceptance region) containing those for which the likelihood is larger than some number b and with the other set (the critical region) containing those samples for which the likelihood is less than b. It can be proved mathematically that this is the best critical region for testing the hypothesis in the sense that it minimizes the type II error for given type I error. The likelihood criterion therefore furnishes a procedure for constructing specific tests of hypotheses. Other approaches to the testing problem, such as Bayes' method or the minimax principle of game theory, lead to the same criterion.

If, for example, one applies the likelihood criterion to the problem of testing whether a normal population has a given mean, perhaps 10, and specifies that the type I error shall be 0.01, then the following procedure results: Draw a random sample of size n; construct an 0.99 confidence interval for the population mean μ; if the number 10 lies within the interval, accept the hypothesis; otherwise, reject it.

An important class of hypotheses has to do with tests of independence in multivariate populations. As an example, one may consider the population of registered voters in the United States as a bivariate population with one variable being their opinions (yes, no, undecided) on some political proposition of the moment, and the other variable being geographical location. Elements of the population may be classified into a so-called contingency table as shown in **Fig. 7**. The hypothesis of independence asserts that the division of political opinion is unaffected by location. To test that hypothesis, a random sample of the population may be interviewed and classified into the 12 categories or cells of the contingency table. Let n_{ij} be the number of individuals falling in the cell in the ith row and the jth column of the table. The sum of all the n_{ij} is, of course, n, the sample size. The row sums may be denoted by r_1, r_2, r_3, r_4, and the column sums by c_1, c_2, c_3; the sum of each of the sets is n. On applying the likelihood criterion to the test of the hypothesis, one finds that it rests on

location	yes	no	undecided
east	n_{11}	n_{12}	n_{13}
south	n_{21}	n_{22}	n_{23}
midwest	n_{31}	n_{32}	n_{33}
west	n_{41}	n_{42}	n_{43}

Fig. 7. Contingency table.

expression (26), which is approximately distributed

$$2\left(n\ \log n + \sum_{ij} n_{ij}\log n_{ij} - \sum_i r_i \log r_i - \sum_j c_j \log c_j\right) \tag{26}$$

according to the chi-square distribution with six degrees of freedom. To test at the 0.05 level for the type I error, one would compute the above expression and compare it with 12.6, which is the value that marks 95% of the area under the chi-square distribution curve. If the expression turned out to be less than 12.6, one would accept the hypothesis of independence; otherwise, one would reject it.

Had there been R rows and C columns instead of four and three, one would have used the chi-square distribution with $(R-1)(C-1)$ degrees of freedom. For a trivariate population there would have been a three-way contingency table and four hypotheses of independence that could have been tested; three of them assert that a given criterion of classification is independent of the other two, and the fourth asserts that all three are mutually independent.

Design of experiments. An experiment is performed to obtain information about the relations between several variables. For example, one may study the effect of storage temperature and duration of storage on the flavor of a frozen food. Three variables (flavor, temperature, and duration) are involved; one (flavor) is called the subject of the experiment; the other two are called factors which influence the subject. Sometimes the factors have intrinsic value in themselves; sometimes they are merely nuisance variables which must be taken into account because it is impossible to perform the experiment without them.

There exist in the statistical literature great numbers of specific experimental designs. These are patterns for making experimental observations; the actual construction of the designs requires quite advanced mathematics based on group theory, finite geometries, and combinatorial analysis. The mathematical problem is to find a pattern from which it is possible to extract the desired information and yet minimize the number of observations.

Experimental designs are most important to the experimenter when observations are expensive to make and when more than one factor is involved in the experiment. In the past it was believed that the best experimental procedure was to vary factors one at a time. Thus, in the frozen food experiment one might have held storage temperature constant and studied the effect of duration of storage only on flavor. Having determined that relationship, one would then hold duration constant and study the temperature effect. This procedure is not only wasteful of time and resources but may very well lead to erroneous conclusions because in all likelihood there is interaction between the two factors; that is, duration effect probably changes in a not obvious way when temperature is changed. Even if one believes that one can extrapolate accurately

the duration effect to other temperatures, there is much to be said for checking the extrapolation procedure in the experiment because it can probably be done at no additional cost if the experiment is well designed.

As an illustration of the use of a design, consider a very simple and useful one called the randomized block design. Suppose a manufacturer contemplates purchasing a machine which can be obtained from one of three sources, X, Y, or Z. He obtains one of each on trial in order to compare their performance using several operators from his own plant. Perhaps five operators, A, B, C, D, and E, are to be used in the proposed experiment, which is a two-factor experiment with performance being the subject and the two factors being humans and machines. The operators are not really a factor of interest in evaluating the machines but are, of course, necessary to the experiment.

It would be a mistake to use perhaps 15 operators, putting 5 on each machine, because operators differ in ability and a machine that turned out well might have been fortunate in having a good set of operators assigned to it. It is necessary that the machines be compared in as nearly equivalent circumstances as possible; in this case, that is done by having each operator operate every machine. The machines may then be compared in blocks (operators) which are individually homogeneous although they may be quite different among themselves.

The data are obtained by measuring production of each operator on each machine for a specified period of time and filling in the table of **Fig. 8**. It is essential that the order in which an operator works on the three machines be randomized (by tossing dice, for example) so that minor factors not taken into account in the design will not bias the results. If, for example, learning is an important factor in operating the machines, then the experiment must be three-factor and a more elaborate design is required.

In the resulting experimental observations the effects of humans and of machines are entangled, but there exists a computational technique for well-designed experiments known as the analysis of variance which will disentangle them. This enables the total variation to be broken down into parts, a variance attributable to machines, one to humans, and one to interactions; if the experiment were duplicated, one could also break out a variance corresponding to experimental error. One can estimate the individual main effects of machines, the main effects of humans, and individual interaction effects. In experiments involving more factors, one may be able (depending on the experimental design selected) to estimate higher-order interactions such as the second-order interactions between the main effects of one factor and the first-order interactions of two other factors.

machine / operator	X	Y	Z
A			
B			
C			
D			
E			

Fig. 8. Data form for randomized block experiment.

A common hypothesis in an analysis of variance is the null hypothesis that the variance associated with some factor or interaction is not larger than the experimental error variance; the test criterion is essentially the ratio of the two variances and has the so-called F distribution when the null hypothesis is true. Rejection of the null hypothesis implies that the factor or interaction in question had a significant effect on the subject of the experiment.

Regression and correlation. The regression problem is that of estimating certain unknown constants or parameters occurring in a function which relates several variables; the variables may be random or not. By far the most easily handled cases are those in which the function is linear in the unknown parameters, and it is worth considerable effort to transform the function to that form if at all possible.

The eventual adult height H of a 5-year-old boy may be quite well predicted by a linear function of three variables: his own present height B, his father's height F, and his mother's height M. The linear function is given in Eq. (27), where a, b, c, and d are

$$H = a + bB + cF + dM \tag{27}$$

the unknown parameters. The three variables B, F, and M are called independent variables, the variable H is called the dependent variable, and the parameters b, c, and d are often referred to as regression coefficients.

To estimate the parameters, one might draw a random sample of 5-year-old boys, measure their heights and those of their parents, then some years later measure their adult heights. One would then have sufficient data from which the parameters could be estimated by procedures entirely analogous to the methods for estimating location and scale parameters described above. In practice, one would not take so long to get the data but would use men whose childhood records were available. This method involves subtle sampling problems, however; for example, persons whose records are available may have been better cared for on the average and hence taller on the average.

The data which supply the estimates of regression coefficients can also be used to estimate the standard deviation, or standard error of regression, σ. Using that estimate and a table of the t distribution, one can compute a prediction interval (analogous to a confidence interval) which will have the desired probability of including the correct adult height of a given boy.

Tall fathers sometimes have short sons and short fathers sometimes have tall sons, but generally a father's height is a good indicator of his son's height;

that is, fathers' heights and their sons' heights are positively correlated. For a different example: The price of a commodity in a free economy is negatively correlated with the supply of that commodity; there is not a fixed relation between price and supply, but as a general proposition one goes up when the other goes down.

Statisticians have developed measures of the degree of such imprecise relationships called coefficients of correlation. The most widely used one is the Pearson, or product moment, correlation, which is generally denoted by ρ. It measures the degree of linear association or correlation between two random variables of a bivariate or multivariate population and is defined as the mean over all the population of the product of the deviations of the two variables from their means divided by the product of their standard deviations. Mathematically, it is as shown in Eq. (28),

$$\rho = \int (x - \mu)(y - \nu) f(x, y)\, dx \frac{dy}{\sigma_x \sigma_y} \tag{28}$$

where x and y are the random variables, μ and ν their means, σ_x and σ_y their standard deviations, and $f(x,y)$ their frequency distribution.

If the two variables are completely unrelated, then $\rho = 0$; if they have a fixed linear relation so that one can be calculated directly from the other, then $\rho = +1$ or $\rho = -1$ depending upon whether their relation is direct or reverse. Otherwise ρ will be some fraction between -1 and $+1$ with the fraction being near zero if there is poor correlation between them.

When the independent variables of a regression function are random variables, there is an equivalence between correlation coefficients and regression coefficients; after one set has been defined, the definition of the other follows automatically; mathematical formulas connecting them may be found in any statistics textbook.

Nonparametric inference. Most techniques of statistical inference rely on the central limit theorem or on sampling distributions derived from normal populations. They are practically always valid for large samples and are often valid for samples of intermediate size. However, there are occasions when one cannot rely on these techniques, particularly when samples are small or when some evident peculiarity of the population (such as marked asymmetry) makes ordinarily used sampling distributions suspect. In such instances one uses nonparametric methods, which are valid whatever form the population distribution might take.

Nonparametric techniques use the so-called order statistics, which are merely the sample observations arranged in ascending order of magnitude. One may let x_1 be the smallest sample observation, x_2 the next smallest, and so on with x_n being the largest. A basic theorem states that on the average the sample divides the population into $n + 1$ equal parts, that is, that $1/(n + 1)$ of the population lies between any two successive order statistics. Many nonparametric methods rest on this fact and its consequences.

As a simple illustration of a nonparametric estimate and confidence interval, one may consider the estimation of the population median, M, given an ordered sample of size 5 consisting of x_1, x_2, x_3, x_4, and x_5. The estimate of M among the observations is simply the central sample observation x_3. The extreme observations x_1 and x_5 provide limits for a confidence interval for M; the probability level for the interval is calculated as follows. One-half of the population lies to the right of M, hence the probability is $^1/_2$ that a randomly drawn observation will lie to its right. The probability is $(^1/_2)^5$ that all five observations will lie to its right. Similarly, the probability is $(^1/_2)^5$ that all five will lie to its left. In all other cases the sample will have at least one observation on each side of M; therefore, expression (29) can be formulated.

Probability that

$$\begin{aligned}(x_1 < M < x_5) &= 1 - (1/2)^5 - (1/2)^5 \\ &= 30/32 \cong 0.94\end{aligned} \tag{29}$$

See ANALYSIS OF VARIANCE; BIOMETRICS; DISTRIBUTION (PROBABILITY); EXPERIMENT; QUALITY CONTROL.

Alexander M. Mood

Bibliography. K. A. Brownlee, *Statistical Theory and Methodology in Science and Engineering*, 2d ed., 1965, reprint 1984; W. J. Dixon and F. J. Massey, *Introduction to Statistical Analysis*, 4th ed., 1983; M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics*, vol. 2, 5th ed., 1991, vol. 3, 4th ed., 1983, reprint 1987, vol. 4, 1983; A. M. Mood, *Introduction to the Theory of Statistics*, 3d ed., 1974; D. H. Sanders, *Statistics: A First Course*, 5th ed., 1995; M. F. Triola, *Elementary Statistics*, 8th ed., 2000.

Statoblasts

Chitin-encapsulated bodies, resistant to freezing and a limited amount of desiccation. They serve as a special means of asexual reproduction in the Phylactolaemata, a class of fresh-water Bryozoa. They are 0.01–0.06 in. (0.26–1.5 mm) long, and their shape and structure are important to the taxonomy of the group.

Statoblasts are classified as follows: sessoblasts, those which attach to zooecial tubes (structures of the outer layer or zooecial of an individual in the colony) or the substratum; floatoblasts and spinoblasts, both of which have a float of "air" cells and therefore are free-floating; and piptoblasts, which are free but have no float (see **illus.**).

These bodies are produced in enormous quantities from spring to autumn. They develop by organization of masses of peritoneal cells and epidermal cells that bulge into the coelom. Each mass then secretes protective upper and lower chitinous valves, the rims of which often project peripherally. Statoblasts are therefore somewhat disk-shaped.

These bodies remain dormant for variable periods of time and serve to tide species over adverse ecological conditions, such as freezing or drying, which

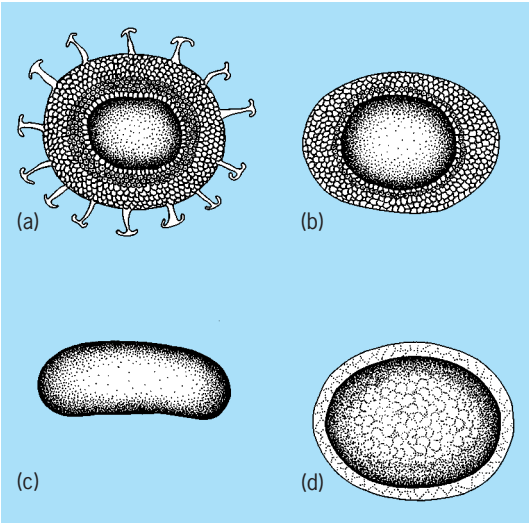

Some types of statoblasts. (*a*) Spinoblast of *Pectinatella magnifica*. (*b*) Floatoblast of *Plumatella repens*. (*c*) Piptoblast or sessoblast of *Fredericella sultana*. (*d*) Sessoblast of *Stolella indica*.

kill the colony. During this time they may be dispersed over considerable distances, being carried by animals, floating vegetation, or the action of water currents. When environmental conditions become favorable, which is usually in the spring, statoblasts germinate and a zooid develops from the mass of cells lying between the two valves. Statoblasts of *Lophopodella* have germinated after 50 months of drying. *See* BRYOZOA. Sybil P. Parker

Staurolite

A nesosilicate mineral occurring in metamorphic rocks. The chemical formula of staurolite may be written as $A_4B_4C_{18}D_4T_8O_{40}X_8$, where A = Fe^{2+}, Mg, □ (□ = vacancy, □ > 2); B = Fe^{2+}, Zn, Co, Mg, Li, Al, Fe^{3+}, Mn^{2+}, □; C = Al, Fe^{3+}, Cr, V, Ti; D = Al, Mg, □ (□ > 2); T = Si, Al; X = OH, F, O. Staurolite is an order-disorder series between an ordered monoclinic structure (with $\beta = 90.64°$) and a disordered orthorhombic structure (with $\beta = 90°$). The structure of staurolite consists of slabs of the kyanite structure [Al_2SiO_5] and layers of Fe-Al oxyhydroxide [$\sim^{[6]}Al_{0.67}\ ^{[4]}Fe_2\ (OH)_2$] alternating along [010], the monoclinic symmetry axis. The kyanite slab consists of edge-sharing chains of AlO_6 octahedra (**Fig. 1***a*, regular dots) extending along the *c* direction, flanked by AlO_6 octahedra (random dots) and cross-linked by SiO_4 tetrahedra (crosses); only small amounts of Mg and Fe substitute for Al in this slab. The oxide-hydroxide layer consists of edge-sharing chains of (Al,□)O_6 octahedra (Fig. 1*b*, regular dots) extending along the *c* axis, alternating with chains of (Al,□)O_6 octahedra (random dots) and (FeO_4) tetrahedra (crosses). In the latter chains, the (Al,□)O_6 octahedra share edges; the (FeO_4) tetrahedra share faces with adjacent (Al,□)O_6 octahedra, and polyhedra adjacent across a shared face cannot both be locally occupied. The extensive coupling of substitutions involving the cations Al, Fe^{2+}, H and □ (vacancy) gives rise to extensive short-range order, and local bond-valence arrangements are optimal for the formula $[Fe^{2+}{}_4Al_{18}\text{-}Si_8O_{44}(OH)_4]^{2+}$. However, at the long-range level, a crystal cannot have a net charge, and hence the principal heterovalent substitutions in staurolite [Al ↔ Si, Mg ↔ Al, □ ↔ H, Li ↔ Fe^{2+}, □ ↔ Fe^{2+}] all act such as to reduce this charge to zero. Thus the chemical complexity of staurolite results from the interaction of long-range and short-range charge-balance requirements.

Staurolite occurs as well-formed, often-twinned, prismatic crystals (**Fig. 2**). It is brown-black, reddish brown, or light brown in color and has a vitreous to dull luster. Light color and dull luster can result from abundant quartz inclusions. There is no cleavage, specific gravity is 3.65–3.75, and hardness is 7–7.5 (Mohs scale).

Staurolite occurs in pelitic (former shale) schists of medium grade (degree) of metamorphism. In typical rocks, it forms as temperature increases from the reaction of muscovite, garnet, and chlorite to form staurolite, biotite, and water, or from the breakdown

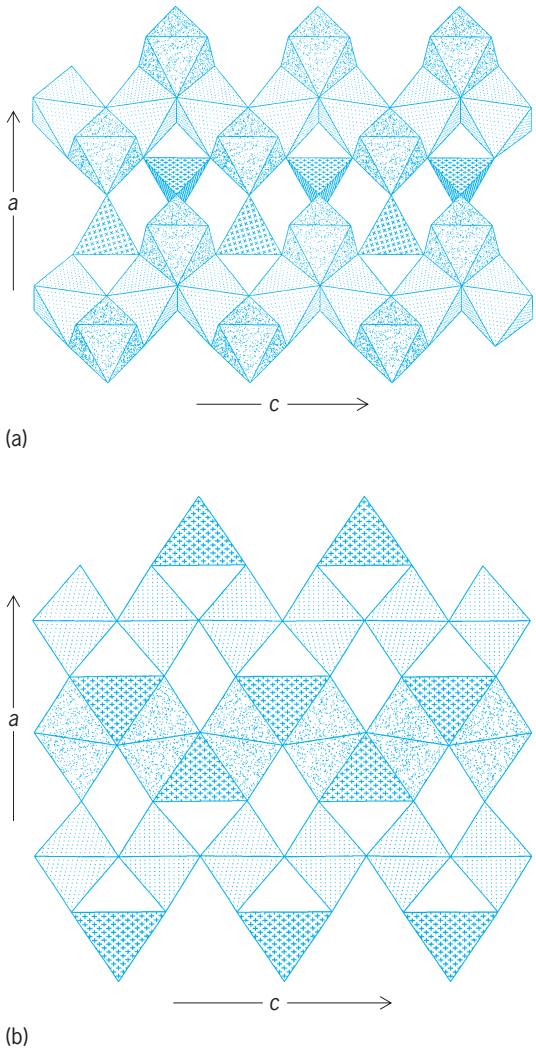

Fig. 1. Staurolite crystal structure: (*a*) kyanite slab and (*b*) oxide-hydroxide layer, where *a* is the shortest axis and *c* is the longest axis.

Fig. 2. Staurolite crystal habits. (*After C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley, 1993*)

of chloritoid to form staurolite, garnet, and water at temperatures of about 930°F (500°C) as shown by laboratory experiments. With increasing temperature in such rocks, staurolite reacts with muscovite to form garnet, biotite, and aluminosilicate minerals (kyanite, sillimanite, or andalusite, depending on pressure) at temperatures near 1110°F (600°C). Thus, most staurolite forms between 930 and 1110°F (500 and 600°C). Staurolite may form at pressure as low as 2.5 kilobars (250 megapascals), corresponding to depths in the Earth's crust of 5.5. mi (9 km); at lower pressure, cordierite and andalusite take its place in pelitic schists or contact metamorphic rocks. Staurolite may occur at pressures corresponding to the deepest levels in the Earth's crust (>10 kbar). However, it is less common and tends to be Mg-rich in deep crustal rocks. Typical minerals occurring with staurolite are quartz, micas (muscovite and biotite), garnet (almandine), tourmaline, and kyanite, sillimanite, or andalusite. Staurolite is common where pelitic schists reach medium-grade metamorphism. Examples are the Swiss and Italian Alps (notable at Saint Gotthard, Switzerland), and all the New England states, Virginia, the Carolinas, Georgia, New Mexico, Nevada, and Idaho. *See* METAMORPHISM; SILICATE MINERALS. Frank C. Hawthorne

Bibliography. W. A. Deer, R. A. Howie, and J. Zussman, *Rock Forming Minerals*: *Orthosilicates*, vol. 1A, 1997; F. C. Hawthorne et al., The crystal chemistry of staurolite, *Can. Mineral.*, 31:551–616, 1993; M. J. Holdaway, B. L. Dutrow, and P. Shore, A model for the crystal chemistry of staurolite, *Amer. Mineral.*, 71:1142–1159, 1986; C. Klein and C. S. Hurlbut, Jr., *Manual of Mineralogy*, 21st ed., 1993.

Stauromedusae

An order of the cnidarian class Scyphozoa, usually found in circumpolar regions. *Haliclystus auricula* is typical. The egg develops into a planula which can only creep since it lacks cilia. The planula changes into a polyp that metamorphoses directly into a combined polyp and medusa form. The medusa is composed of a cuplike bell called a calyx (medusan part) and a stem (polyp part) which terminates in a pedal disk (see **illus.**). The calyx is eight-sided and has eight groups of short, capped tentacles and eight sensory bodies, called anchors (rhopalioids), on its margin. The mouth, situated at the center of the calyx, has four thin lips and leads to the stomach in which gastral filaments are arranged in a row on either side of each interradius. Though sessile, the medusa can

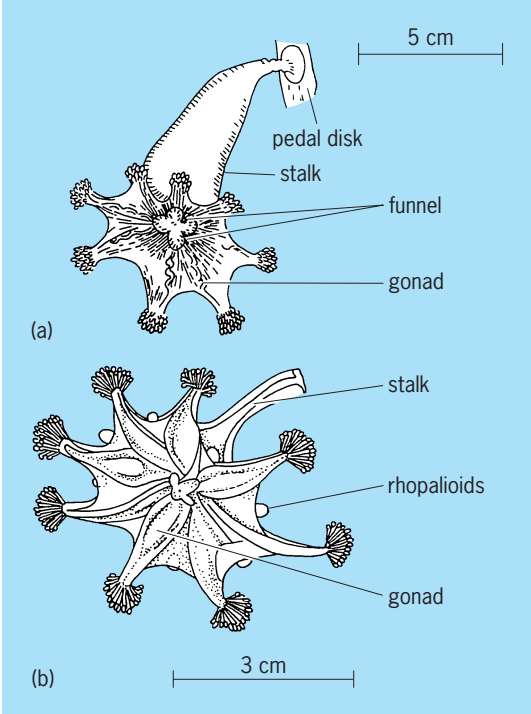

Stauromedusae. (*a*) *Lucernaria*. (*b*) *Haliclystus*. (*After L. H. Hyman, The Invertebrates, vol. 1, McGraw-Hill, 1940*)

move in a leechlike fashion by alternate attachment and release of the pedal disk, using the substratum as an anchor. *See* SCYPHOZOA. Tohru Uchida

Steam

Water vapor, or water in its gaseous state. Steam is the most widely used working fluid in external combustion engine cycles, where it will utilize practically any source of heat, that is, coal, oil, gas, nuclear fuel (uranium and thorium), waste fuel, and waste heat. It is also extensively used as a thermal transport fluid in the process industries and in the comfort heating and cooling of space. The universality of its

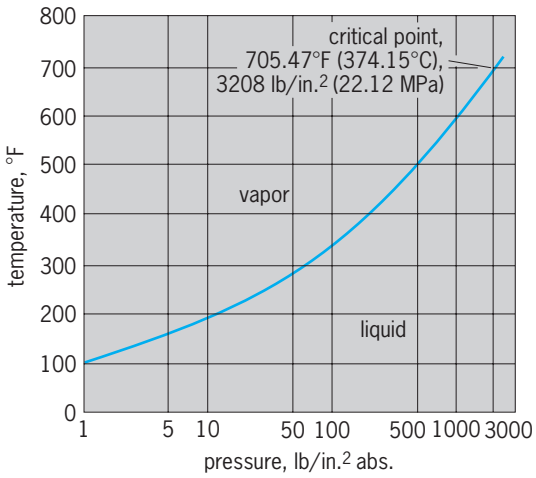

Fig. 1. Vapor-tension curve for water. 1 lb/in.^{2} abs. = 6.895 kPa, absolute pressure. °C = (°F − 32)/1.8.

Thermal properties of steam (approximate)*

Pressure, lb/in.^{2} abs. (kPa)	Temperature, °F (°C)	Enthalpy, Btu/lb‡ Saturated liquid†	Latent heat	Saturated vapor†
1 (6.895)	102 (39)	69.7	1036	1106
10 (68.95)	193 (89)	161	982	1143
14.7 (101.3)	212 (100)	180	970	1150
100 (689.5)	328 (164)	298	889	1187
1000 (6895)	545 (285)	542	649	1192
3208 (22,120)	705 (374)	906	0	906

*For more exact values see J. H. Keenan and F. G. Keyes, *Thermodynamic Properties of Steam,* Wiley, 1936, or American Society of Mechanical Engineers, *Thermodynamic and Transport Properties of Steam* (tables and charts), 1967.
†Enthalpy is measured from saturated liquid at 32°F (0°C).
‡1 Btu/lb = 2326 J/kg.

availability and its highly acceptable, well-defined physical and chemical properties also contribute to the usefulness of steam. The engineering profession has developed so much evaporative and condensing equipment for changing the state of water that its properties can be utilized from pressures ranging from 0.18 in. Hg or 0.61 kilopascal (ice point) to supercritical (>3208.2 lb/in.^{2} absolute or 22.120 megapascals absolute pressure and 705.47°F or 374.15°C). These temperature ranges are readily adaptable for use with ferrous and nonferrous metal structures at costs which dominate the design, selection, and use of many applications.

The temperature at which steam forms depends on the pressure in the boiler. The steam formed in the boiler (and conversely steam condensed in a condenser) is in temperature equilibrium with the water. The relation is rigorously specified by the vapor tension curve (**Fig. 1**). (The **table** gives a few selected values in the first two columns.) Under these conditions, with steam and water in contact and at the same temperature, the steam is termed saturated. Steam can be entirely vapor when it is 100% dry, or it can carry entrained moisture and be wet. It may also be contaminated, as in all practical and engineering applications, with other gases, especially air, in which case it is a mixture and Dalton's law prevails. After the steam is removed from contact with the liquid phase, the steam can be further heated without changing its pressure. If initially wet, the additional heat will first dry it and then raise it above its saturation temperature. This is a sensible heat addition, and the steam is said to be superheated.

The heat energy in steam can be divided into three parts: (1) enthalpy of the liquid required to raise the water from its initial temperature (usually considered to be 32°F or 0°C) to the boiling temperature, (2) enthalpy of vaporization required to convert the water to steam at the boiling temperature, and (3) enthalpy of superheat that raises the steam to its final temperature. As the steam performs its thermodynamic function of giving up its heat energy, it loses its superheat, becomes wet, and finally condenses to hot water. While wet, the steam has a quality that decreases as the percent of dry, saturated steam present in the wet steam decreases. Dry steam has a quality of 100% (see the table and **Fig. 2**). *See* ENTHALPY.

Properties as a gas. Superheated steam at temperatures well above the boiling temperature for the existing steam pressure follows closely the laws of a perfect gas. Thus, $pv \approx 85.8T$ and $k \approx 1.3$, where p is pressure in lb/ft^2, v is volume in ft^3/lb, T is absolute temperature in °R, and k is the ratio of specific heat at constant pressure to specific heat at constant volume. (In SI units, $p'v' \approx 461.5\ T$; where p' is pressure in Pa, v' is volume in m^3/kg, and T' is absolute temperature in K.) However, the behavior of dry, saturated steam departs from that of a perfect gas, and wet steam is a mixture. Therefore, the properties of steam near its vaporization temperature are determined experimentally.

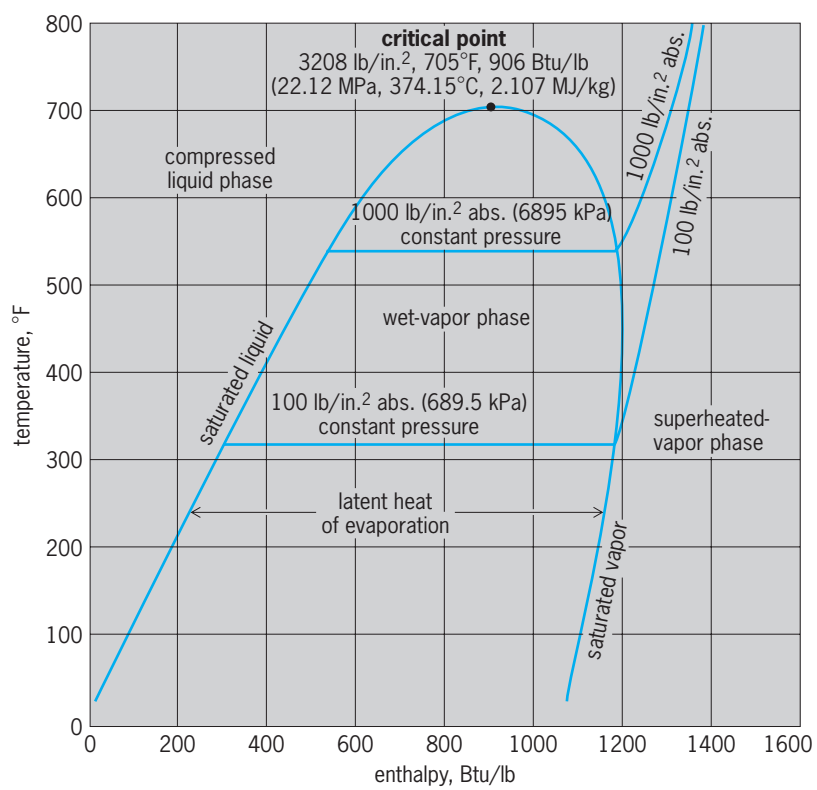

Fig. 2. Temperature-enthalpy chart for steam. 1 Btu/lb = 2326 J/kg. °C = (°F − 32)/1.8.

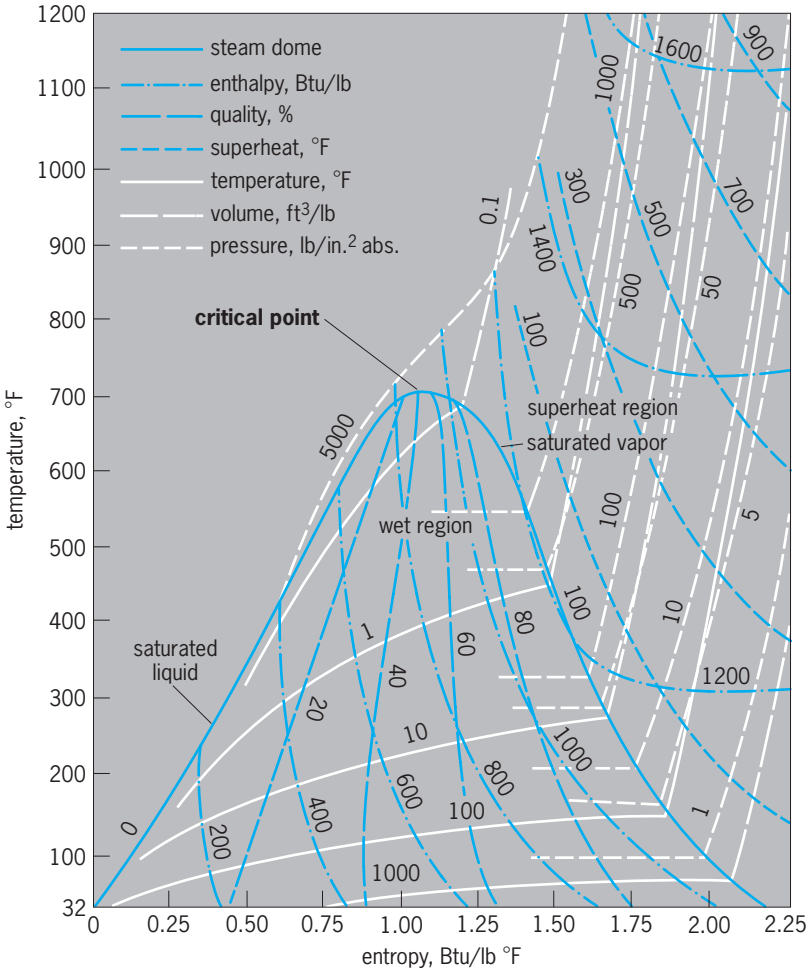

Fig. 3. Temperature-entropy chart for steam. °C = (°F − 32)/1.8; 1 Btu/lb °F = 4186.8 J/kg °C; 1 Btu/lb = 2326 J/kg; 1 ft³/lb = 62.4 × 10⁻³ m³/kg; 1 lb/in.² = 6.895 kPa. (*After J. H. Keenan and F. G. Keyes, Thermodynamic Properties of Steam, John Wiley and Sons, 1936*)

Data can be presented in tabular form, giving pressure, volume, entropy, enthalpy, and temperature for saturated liquid (water at the boiling point), saturated steam, and steam vapor at various temperatures of superheat. Alternatively, thermodynamic properties of steam can be presented diagrammatically. For analysis of thermodynamic cycles, the temperature-entropy chart is widely used (**Fig. 3**). Area on this chart is proportional to heat energy. At the critical point, steam condenses directly into water without releasing energy. In the uncharted area the pressure is higher than is usually encountered in commercial practice.

For engineering application the Mollier diagram presents steam data in a convenient form (**Fig. 4**). In the wet region, lines of constant temperature and of constant pressure are straight and coincide with each other. Total enthalpies above 32°F (0°C) are plotted as ordinates and total entropies as abscissas.

The specific volume of steam varies widely with temperature and pressure. At 1 lb/in.^{2} absolute (6.895 kPa absolute pressure) the specific volume of saturated steam is 334 ft^3 (20.85 m^3/kg), at atmospheric pressure it is 26.8 (1.29), at 1000 lb/in.^{2} absolute (6.895 MPa) it is 0.45 (28×10^{-3}), at 2000 lb/in.^{2} absolute (13.79 MPa) it is 0.19 (12×10^{-3}), and at the critical pressure (3208 lb/in.^{2} absolute or 22.12 MPa) it is 0.05 (3×10^{-3}). Superheating substantially increases values of specific volume of a steam at saturation conditions.

Viscosity of steam increases with temperature and pressure. Saturated steam at atmospheric pressure has a viscosity of about 2.6×10^7 (lb)(s)/ft^2 (1.2×10^9 Pa·s). At 1000°F (538°C) and 2000 lb/in.^{2} (13.79 MPa) pressure it is 13.4×10^7 (lb)(s)/ft^2 (6.4×10^9 Pa·s). Similarly, the heat conductivity of steam increases with temperature and pressure. The thermal conductivities for the conditions for which viscosities are given are about 14, 37, and 109 Btu/(h)(ft)(°F) or 24, 64, and 189 W/(m)(°C).

Application. Chiefly because of its availability, but also because of its nontoxicity, steam is widely used as the working medium in thermodynamic processes. It has a uniquely high latent heat of vaporization: 1049 Btu/lb (2.440 MJ/kg) at 1 in. Hg abs. (3.39 kPa) and 79°F (26°C), 970 (2.256 MJ/kg) at 14.7 lb/in.^{2} abs. (101 kPa) and 212°F (100°C), 889 (2.068 MJ/kg) at 1000 lb/in.^{2} abs. (6.895 MPa) and 328°F (164°C), 649 (1.510 MJ/kg) at 1000 lb/in.^{2} abs. (6.895 MPa) and 545°F (285°C), and 463 (1.077 MJ/kg) at 2000 lb/in.^{2} abs. (13.79 MPa) and 636°F (336°C). Steam has a specific heat in the vicinity of half that of water. For comparison, its specific heat is about twice that of air and comparable to that of ammonia. Except for a few gases such as hydrogen, with a specific heat seven times that of steam, the specific heat of steam is relatively high so that it can

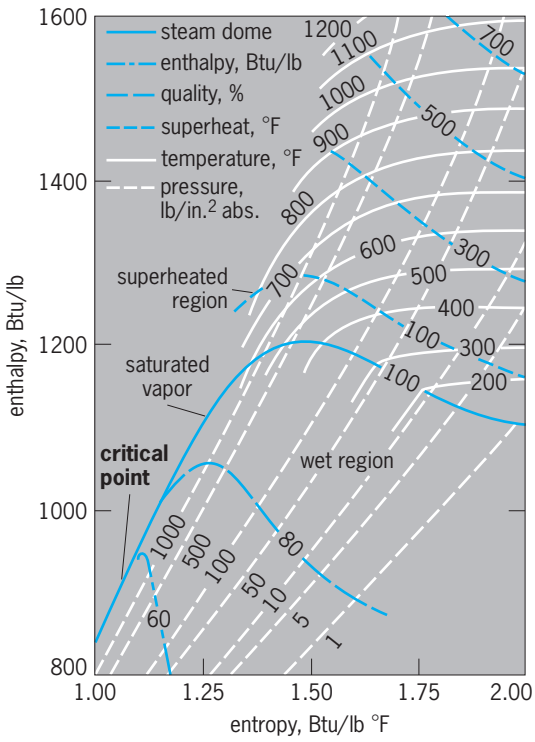

Fig. 4. Mollier diagram for steam. 1 Btu/lb = 2326 J/kg; 1 Btu/lb °F = 4186.8 J/kg °C; °C = (°F − 32)/1.8; 1 lb/in.² = 6.895 kPa. (*After Mollier's Steam Tables and Diagrams, Pitman, 1927*)

carry more thermal energy at practical temperatures than can other usable gases. *See* BOILER; BOYLE'S LAW; CHARLES' LAW; DALTON'S LAW; ENTROPY; STEAM ENGINE; STEAM-GENERATING UNIT; STEAM HEATING; STEAM TURBINE; THERMODYNAMIC CYCLE; THERMODYNAMIC PRINCIPLES; WATER. Theodore Baumeister

Bibliography. American Society of Heating, Refrigeration, and Air-Conditioning Engineers, *Handbook*: *Applications*, 1993, *Equipment*, 1992, *Fundamentals*, 1994, *Systems*, 1992; E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; J. H. Keenan et al., *Steam Tables: Thermodynamic Properties of Water Including Vapor, Liquid, and Solid Phases*, 1978; K. S. Pitzer, *Thermodynamics*, 3d ed., 1994; E. Schmidt, *Properties of Water and Steam*, 1981.

Steam condenser

A heat-transfer device used for condensing steam to water by removal of the latent heat of steam and its subsequent absorption in a heat-receiving fluid, usually water, but on occasion air or a process fluid. Steam condensers may be classified as contact or surface condensers.

In the contact condenser, the condensing takes place in a chamber in which the steam and cooling water mix. The direct contact surface is provided by sprays, baffles, or void-effecting fill such as Rashig rings. In the surface condenser, the condensing takes place separated from the cooling water or other heat-receiving fluid (or heat sink). A metal wall, or walls, provides the means for separation and forms the condensing surface.

Both contact and surface condensers are used for process systems and for power generation serving engines and turbines. Modern practice has confined the use of contact condensers almost entirely to such process systems as those involving vacuum pans, evaporators, or dryers, and to condensing and dehumidification processes inherent in vacuum-producing equipment such as steam jet ejectors and vacuum pumps. The steam surface condenser is used chiefly in power generation but is also used in process systems, especially in those in which condensate recovery is important. Air-cooled surface condensers are used increasingly in process systems and in power generation when the availability of cooling water is limited.

Water-cooled surface condensers used for power generation are of the high-vacuum type. Their main purpose is to effect a back pressure at the turbine exhaust. To achieve an economical station heat rate and fixed cost, these condensers must be designed for high heat-transfer rates, minimum steam-side-pressure loss, and effective air removal. The usual power plant steam condenser incorporates a single steam-condensing chamber into which the steam turbine exhausts. Multipressure condensers, having two or more steam-condensing chambers operating at different pressures, have been found to be more economical than single-pressure condensers for large installations, which generally use cooling towers as the final heat sink. Most multipressure condensers are compartmented internally, on the steamcondensing side, with the cooling water arranged to flow through the tubes in one direction from inlet to outlet. The mean cooling water temperature progressively increases as the water flows from compartment to compartment, with back pressure on the turbine increasing proportionately. Economical performance is effected with multiple-exhaust pressure turbines when each exhaust pressure section of the turbine is designed for the back pressure produced in its respective pressure section in the condenser. A rise in water temperature in the order of 20–30°F (11–17°C) is necessary to effect required pressure differences in the turbine.

Air-cooled surface condensers used for power generation are designed to operate at higher back pressures than water-cooled condensers. This is consistent with their lower overall heat-transfer rate, which is of the order of 10–12 Btu/(°F)(ft^2)(h) [57–68 W/(°C)(m^2)] in contrast with the normal 600–700 Btu/(°F)(ft^2)(h) [3.4–4.0 kW/(°C)(m^2)] characteristic of water-cooled steam condensers. To compensate for these large differences in heat transfer, greater temperature differences are required between the condensing steam and the cooling air than those needed with conventional water cooling. Steam is condensed on the inside of the heat-transfer tubes, and airflow is usually produced by fans across their outside surface. The tubes are of the extended-surface type with fins on the outside. The normal ratio of outside to inside surface is on the order of 10 or 12 to 1. The use of air as the cooling medium for steam condensers used in smaller industrial steam power plants is increasing rapidly. Air-cooled steam condensers for the production of electric power have been confined to smaller plants; however, with the decreasing availability of water for cooling, the application of air-cooled steam condensers to the larger steam electric plants may be anticipated.

Condenser sizes have increased with the increase in size of turbine generators. The largest, in a single shell, is approximately 485,000 ft^2 (45,000 m^2); the steam-condensing space occupies a volume of 80,000 ft^3 (2265 m^3). Large units have been built with the surface divided equally among two or three shells. Power plant condensers require 35–100 lb of cooling water to condense 1 lb (35–100 kg for 2 kg) of steam, or, if air-cooled, 2500–5000 ft^3 of air for 1 lb (150–300 m^3 for 1 kg) of steam. Normally, about 0.5 ft^2 (0.05 m^2) of surface is required for water-cooled condensers, and about 10 ft^2 (0.9 m^2) of inside tube surface for air-cooled condensers is required for each kilowatt of generating capacity. *See* STEAM; STEAM TURBINE; VAPOR CONDENSER. Joseph F. Sebald

Bibliography. H. P. Bloch, *A Practical Guide to Steam Turbine Technology*, 1995; J. P. Holman, *Heat Transfer*, 9th ed., 2001.

Steam electric generator

An alternating-current (ac) synchronous generator driven by a steam turbine for 50- or 60-Hz electrical generating systems operation. *See* STEAM TURBINE.

Basic operation and construction. The synchronous generator is made of two parts: a stator (stationary) and a rotor (rotating) [see **illus.**]. The stator is a cylindrical steel frame, inside of which a cylindrical iron core made of thin insulated laminations is mounted on a support system. The iron core has equally spaced axial slots on its inside circumference, and wound within the core slots is a stator winding. The stator winding copper is electrically insulated from the core. The rotor consists of a forged solid steel shaft. Wound into axial slots on the outside diameter of the shaft is a copper rotor winding that is held in the slots with wedges. Retaining rings support the winding at the rotor body ends. The rotor winding, referred to as the field, is electrically insulated from the shaft and is arranged in pole pairs (always an even number) to form the magnetic field which produces the flux. The rotor shaft (supported by bearings) is coupled to the steam turbine, and rotates inside the stator core. *See* ELECTRIC ROTATING MACHINERY; MAGNETIC FLUX; WINDINGS IN ELECTRIC MACHINERY.

The stator winding (armature) is connected to the ac electrical transmission system through the bushings and output terminals. The rotor winding (field) is connected to the generator's excitation system. The excitation system provides the direct-current (dc) field power to the rotor winding via carbon brushes riding on a rotating collector ring mounted on the generator rotor. The synchronous generator's output voltage amplitude and frequency must remain constant for proper operation of electrical load devices. During operation, the excitation system's voltage regulator monitors the generator's output voltage and current. The voltage regulator controls the rotor winding dc voltage to maintain a constant generator stator output ac voltage, while allowing the stator current to vary with changes in load. Field windings typically operate at voltages between 125 and 575 V dc. The synchronous generator's output frequency is directly proportional to the speed of the rotor,

$$\text{Frequency} = (\text{rpm} \times \text{number of poles}) \div (120)$$

and the speed of the generator rotor is held constant by a speed governor system associated with the steam turbine. *See* ALTERNATING-CURRENT GENERATOR; GENERATOR.

The first commercial use of the ac synchronous generator was in the mid-1880s. Synchronous generators range in size from a few kilovolt-amperes to 1,650,000 kVA. 60-Hz steam-driven synchronous generators operate at speeds of either 3600 or 1800 rpm; for 50-Hz synchronous generators these speeds would be 3000 or 1500 rpm. These two- and four-pole generators are called cylindrical rotor units. For comparison, water (hydro)-driven and air-driven synchronous generators operate at lower speeds,

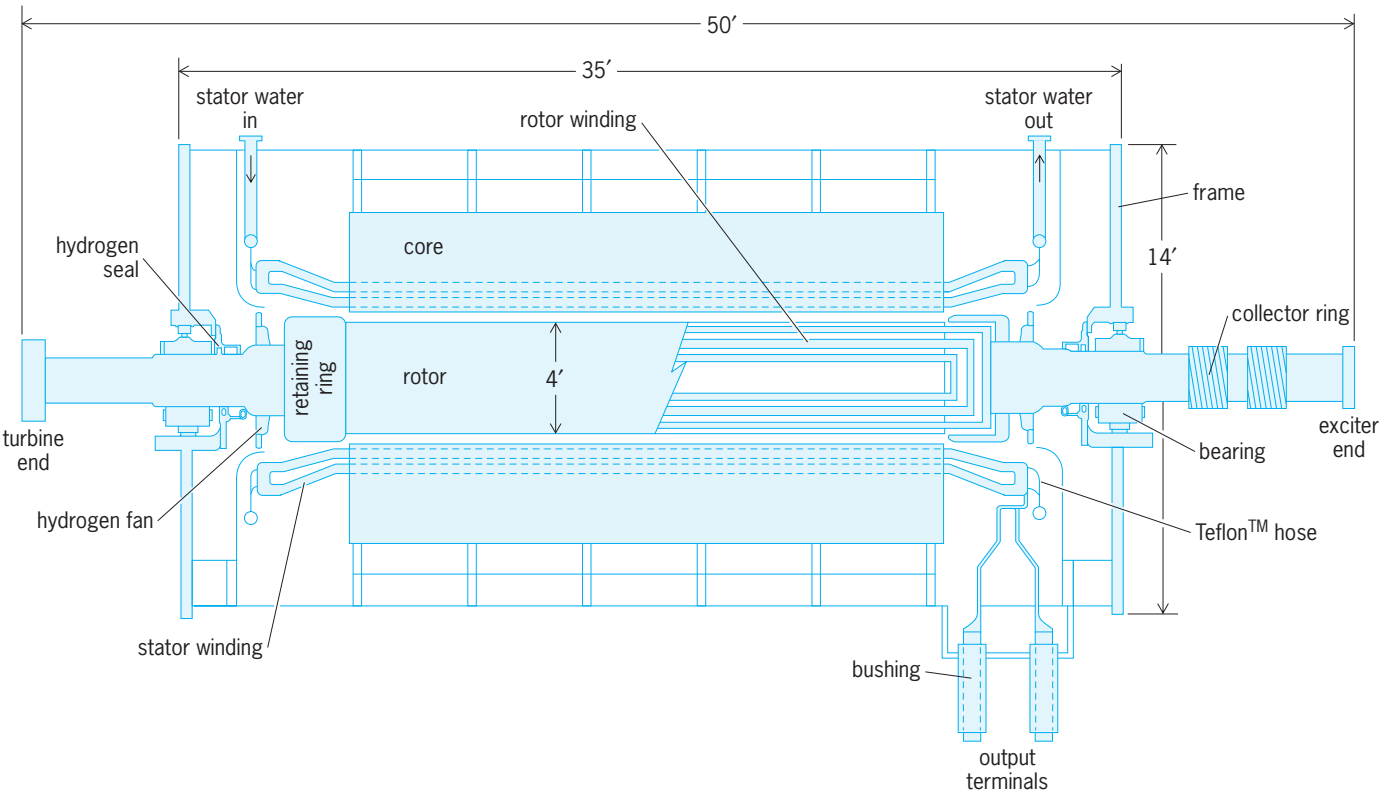

Cross section of a typical hydrogen- and water-cooled, two-pole, 3600-rpm, steam electric generator. The dimensions shown are for a 900,000-kVA synchronous generator. The hydrogen coolers are not shown.

some as low as 62 rpm (116 poles). The stator output voltage of large (generally greater than 100,000 kVA) units ranges 13,800–27,000 V. The operating characteristics of the steam-turbine-driven synchronous generators are defined in IEEE Std. C50.13-2005, *IEEE Standard for Cylindrical-Rotor 50 Hz and 60 Hz Synchronous Generators Rated 10 MVA and Above*. *See* ELECTRIC POWER GENERATION; HYDROELECTRIC GENERATOR.

The physical size of a steam generator is determined by the transportation limitations of size and weight. For the generator shown in the illustration, the stator would weigh about 800,000 lb (363,000 kg), and the rotor about 140,000 lb (64,000 kg).

Cooling. There are five sources of heat loss in a synchronous generator: stator winding resistance, rotor winding resistance, core, windage and friction, and stray losses. Removing the heat associated with these losses is the major challenge to the machine designer.

As the demand for power increased, generators grew in output rating and physical size. The cooling requirements for the stator windings, rotor windings, and core increase proportionally to the cube of the machine size. The early synchronous generators were air-cooled. Later, air-to-water coolers were required to remove the heat.

In 1937, hydrogen gas was introduced as a cooling medium in place of air. The generator frame had to be sealed against hydrogen leakage. The first units were filled with pure hydrogen at pressures of 0.5–5 psi (3.5–35 kPa), with a hydrogen purity of greater than 97%. In today's large generators, hydrogen pressures of 75 psi (520 kPa) are common. An elaborate seal-oil system (gland seal) is used to form a hydrogen seal near the bearings where the rotating shaft exits the frame. The seal-oil system operates at a pressure that is slightly higher than the hydrogen pressure. Hydrogen is the most effective gas available for generator cooling, and is used because it removes approximately ten times more heat than air; its density is one-tenth that of air; windage losses are lower; hydrogen is a better electrical insulator; it is a nonoxidizing agent; and pure hydrogen will not support combustion in the case of a winding failure. The hydrogen is circulated throughout the stator frame by blowers or fans mounted on the rotor, and is cooled by hydrogen-to-water heat exchangers mounted within the stator frame. In the early machines, the hydrogen gas cooled the surface of the stator and rotor windings (as did air), relying on conduction for heat transfer through the winding insulation. This cooling method is called indirect cooling. For larger machines, the stator windings were designed with hydrogen cooling passages manufactured within the copper winding, which allowed the hydrogen to flow through the winding in direct contact with the copper. These designs were called directly cooled windings. The rotor windings were also directly cooled with hydrogen gas. Helium is not used because it has higher density, has higher specific heat, and costs five times more than hydrogen. *See* HYDROGEN.

Synchronous generators continued to increase in output, and in 1957 oil was introduced as a cooling medium for stator windings. Oil was quickly replaced by water in 1961. The stator winding losses are responsible for about 25–40% of the total losses. The copper windings were manufactured with hollow copper conductors (strands). Water flows through the conductors in direct contact with the copper windings, thus removing the heat. Water enters and leaves the stator winding through Teflon hoses, which electrically insulate the steel water piping from the stator winding copper. A stator cooling water system contains a storage tank and pumps to move the water through the stator winding, filters, strainers, heat exchangers, and demineralizers to maintain low conductivity. The demineralizer contains a resin bed which removes minerals and other conductive impurities from the water.

James R. Michalec

Bibliography. T. C. Elliott, K. Chen, and R. Swanekamp (eds.), *Standard Handbook of Powerplant Engineering*, McGraw-Hill, New York, 2d ed., 1998; A. E. Fitzgerald, C. Kingsley, Jr., and S. D. Umans, *Electric Machinery*, McGraw-Hill, New York, 6th ed., 2003; A. J. Gonzalez et al., *Monitoring and Diagnosis of Turbine-Driven Generators*, Prentice Hall, Englewood Cliffs, NJ, 1995; Institute of Electrical and Electronics Engineers, *IEEE Guide for Operation and Maintenance of Turbine Generators*, IEEE Standard 67–1990, 1990; Institute of Electrical and Electronics Engineers, *IEEE Standard Dictionary of Electrical and Electronics Terms*, 7th ed., IEEE Std. 100–2001, 2001; Institute of Electrical and Electronics Engineers, *IEEE Standard for Cylindrical-Rotor 50 Hz and 60 Hz Synchronous Generators Rated 10 MVA and Above*, IEEE Std. C50.13–2005, 2005.

Steam engine

A machine for converting the heat energy in steam to mechanical energy of a moving mechanism, for example, a shaft. The steam engine dominated the industrial revolution and made available a practical source of power for application to stationary or transportation services. The steam power plant could be placed almost anywhere, whereas other means of power generation were more restricted, experiencing such site limitations as an elevated water supply, wind, animal labor, and so on. The steam engine can utilize any source of heat in the form of steam from a boiler. It was developed in sizes which ranged from that of children's toys to 25,000 hp (18.6 MW), and it was adaptable to pressures up to 200 lb/in.^{2} (1.4 megapascals). It reached its zenith in the nineteenth century in stationary services such as drives for pumping plants; drives for air compressor and refrigeration units; power supply for factory operations with shafting for machine shops, rolling mills, and sawmills; and drives for electric generators as electrical supply systems were perfected. Its adaptability to portable and transportation services rested largely on its development of full-rated torque at any

speed from rest to full throttle; its speed variability at the will of the operator; and its reversibility, flexibility, and dependability under the realities of stringent service requirements. These same features favored its use for many stationary services such as rolling mills and mine hoists, but the steam engine's great contribution was in the propulsion of small and large ships, both naval and merchant. Also, in the form of the steam locomotive, the engine made the railroad the practical way of land transport. Most machine elements known today had their origin in the steam engine: cylinders, pistons, piston rings, valves and valve gear crossheads, wrist pins, connecting rods, crankshafts, governors, and reversing gears. *See* LOCOMOTIVE.

The twentieth century saw the practical end of the steam engine. The steam turbine with (1) its high speed (for example, 3600 rpm); (2) its utilization of maximum steam pressures (2000–5000 lb/in.^{2} or 14–34 MPa), maximum steam temperatures (~1100°F or 600°C), and highest vacuum (~29 in. Hg or 98 kilopascals); and (3) its large size (~1,000,000 kW) led to such favorable weight, bulk, efficiency, and cost features that it replaced the steam engine as the major prime mover for electric generating stations. The internal combusion engine, especially the high-speed automotive types which burn volatile (gasoline) or nonvolatile (diesel) liquid fuel, offers a self-contained, flexible, low-weight, low-bulk power plant with high thermal efficiency that has completely displaced the steam locomotive with the diesel locomotive and marine steam engines with the motorship and motorboat. It is the heart of the automotive industry, which produces in a year 10,000,000 vehicles that are powered by engines smaller than 1000 hp (750 kW). Because of the steam engine's weight and speed limitations, it was excluded from the aviation field, which has become the exclusive preserve of the internal combustion piston engine or the gas turbine. *See* DIESEL ENGINE; GAS TURBINE; INTERNAL COMBUSTION ENGINE; STEAM TURBINE; TURBINE.

Cylinder action. A typical steam reciprocating engine consists of a cylinder fitted with a piston (**Fig. 1**). A connecting rod and crankshaft convert the piston's to-and-fro motion into rotary motion. A flywheel tends to maintain a constant-output angular velocity in the presence of the cyclically changing steam pressure on the piston face. A D slide valve

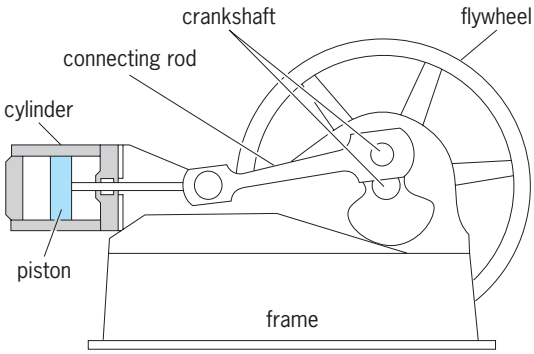

Fig. 1. Principal parts of horizontal steam engine.

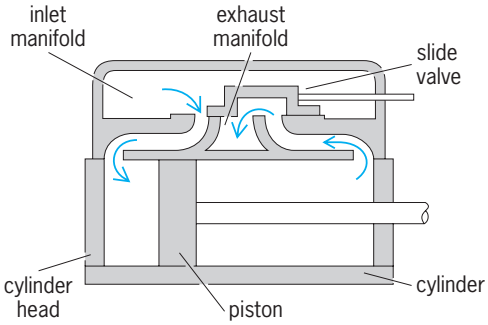

Fig. 2. Single-ported slide valve on counterflow double-acting cylinder.

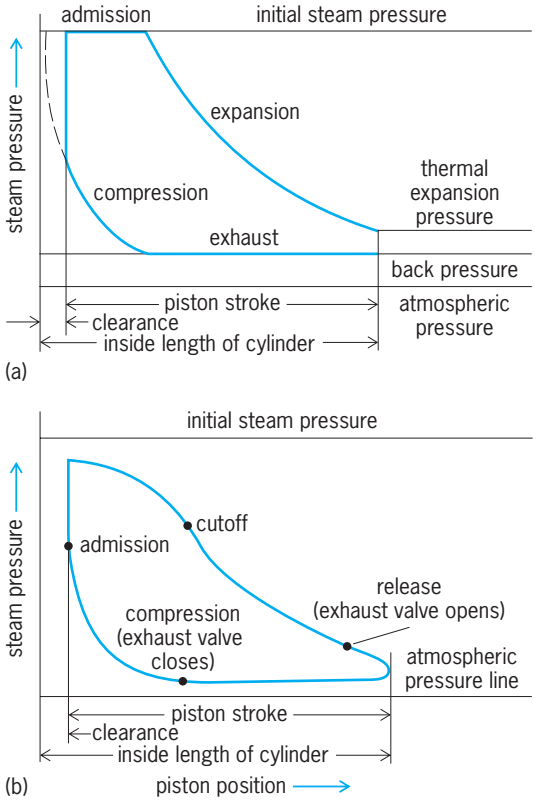

Fig. 3. Events during one cycle of a piston operation. (*a*) In ideal engine. (*b*) As depicted on the indicator card of a noncondensing steam engine.

admits high-pressure steam to the cylinder and allows the spent steam to escape (**Fig. 2**). The power developed by the engine depends upon the pressure and quantity of steam admitted per unit time to the cylinder.

Indicator card. The combined action of valves and piston is most conveniently studied by means of a pressure-volume diagram or indicator card (**Fig. 3**). The pressure-volume diagram is a thermodynamic analytical method which traces the sequence of phases in the cycle. It may be an idealized operation (Fig. 3*a*), or it may be an actual picture of the phenomena within the cylinder (Fig. 3*b*) as obtained with an instrument commonly known as a steam engine indicator. This instrument, in effect, gives a graphic picture of the pressure and volume for all phases of steam admission, cutoff, expansion, release, exhaust, and compression. It is obtained as

the engine is running and shows the conditions which prevail at any instant within the cyclinder. The indicator is a useful instrument not only for studying thermodynamic performance, but for the equally important operating knowledge of inlet and exhaust valve leakage and losses, piston ring tightness, and timing correctness. *See* THERMODYNAMIC PROCESSES.

The net area of the indicator card shows, thermodynamically, the work done in the engine cylinder. By introducing the proper dimensional quantities, power output can be measured. Thus if the net area within the card is divided by the length, the consequent equivalent mean height is the average pressure difference on the piston during the cycle and is generally called the mean effective pressure p, usually expressed in lb/in.^{2} With a cylinder dimension of piston area a (in in.^{2}), length of piston stroke l (in ft), and n equal to the number of cycles completed per minute, the equation given below holds. It has often

$$\text{Indicated horsepower} = \frac{plan}{33,000}$$

been referred to as the most important equation in mechanical engineering.

Engine types. Engines are classified as single- or double-acting, and as horizontal (Fig. 1) or vertical depending on the direction of piston motion. If the steam does not fully expand in one cylinder, it can be exhausted into a second, larger cylinder to expand further and give up a greater part of its initial energy. Thus, an engine can be compounded for double or triple expansion. In counterflow engines, steam enters and leaves at the same end of the cylinder; in uniflow engines, steam enters at the end of the cylinder and exhausts at the middle.

Steam engines can also be classed by functions, and are built to optimize the characteristics most desired in each application. Stationary engines drive electric generators, in which constant speed is important, or pumps and compressors, in which constant torque is important. Governors acting through the valves hold the desired characteristic constant. Marine engines require a high order of safety and dependability.

Valves. The extent to which an actual steam piston engine approaches the performance of an ideal engine depends largely on the effectiveness of its valves. The valves alternately admit steam to the cylinder, seal the cylinder while the steam expands against the piston, and exhaust steam from the cylinder. The many forms of valves can be grouped as sliding valves and lifting valves (**Fig. 4**). *See* CARNOT CYCLE; THERMODYNAMIC CYCLE; VALVE.

D valves (Fig. 2) are typical sliding valves where admission and exhaust are combined. A common sliding valve is the rocking Corliss valve; it is driven from an eccentric on the main shaft like other valves but has separate rods for each valve on the engine. After a Corliss valve is opened, a latch automatically disengages the rod and a separate dashpot abruptly closes the valve. Exhaust valves are closed

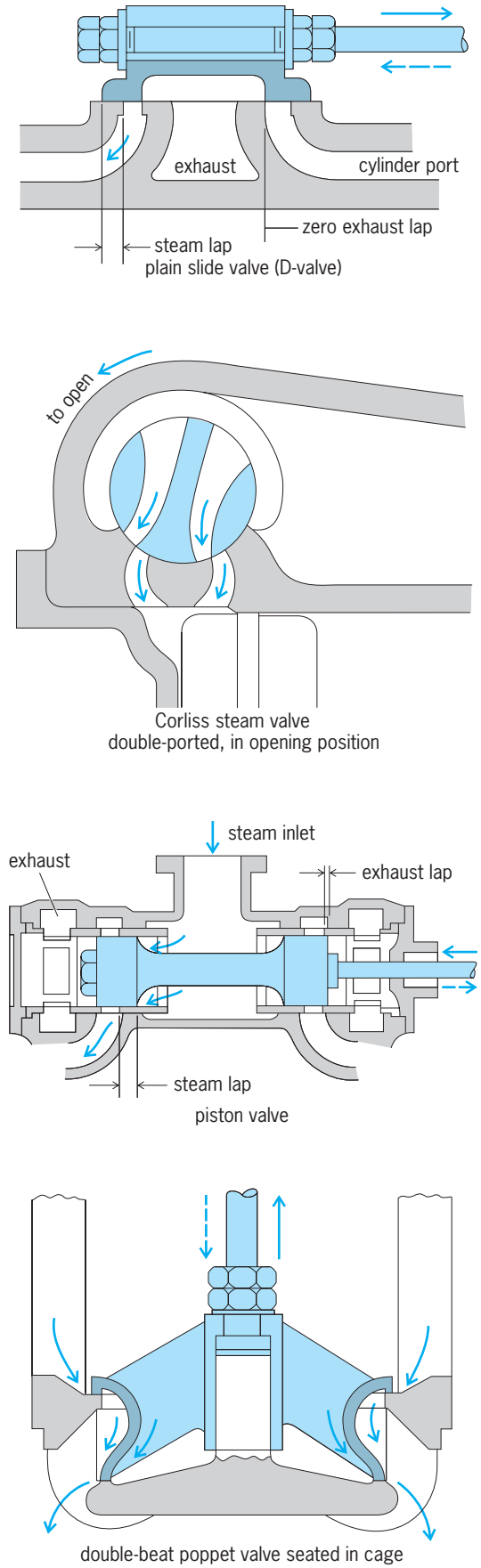

Fig. 4. Steam engine valves in closed positions. The arrows indicate the path steam will travel when the valves (colored parts) are open.

by rods, as are other sliding valves.

Lifting valves are more suitable for use with high-temperature steam. They, too, are of numerous forms. The poppet valve is representative.

Valves are driven through a crank or eccentric on the main crankshaft. The crank angle is set to open the steam port near dead center, when the piston is at its extreme position in the cylinder. The angle between valve crank and connecting-rod crank is slightly greater than 90°, the excess being the angle of advance. So that the valves will open and close quickly, they are driven at high velocity with consequently greater travel than is necessary to open and close the ports. The additional travel of a sliding valve is the steam lap and the exhaust lap. The greater the lap, the greater the angle of advance to obtain the proper timing of the valve action.

Engine power is usually controlled by varying the period during which steam is admitted. A shifting eccentric accomplishes this function or, in releasing Corliss and in poppet valves, the eccentric is fixed and cutoff is controlled through a governor to the kickoff cams or to the latch that allows the valves to be closed by their dashpots.

For high engine efficiency, the ratio of cylinder volume after expansion to volume before expansion should be high. The volume before expansion into which the steam is admitted is the volumetric clearance. It may be determined by valve design and other structural features. For this reason, valves and ports are located so as not to necessitate excessive volumetric clearance. *See* BOILER; STEAM.

Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996.

Steam-generating furnace

An enclosed space provided for the combustion of fuel to generate steam. The closure confines the products of combustion and is capable of withstanding the high temperatures developed and the pressures used. Its dimensions and geometry are adapted to the rate of heat release, to the type of fuel, and to the method of firing so as to promote complete burning of the combustible and suitable disposal of ash. In water-cooled furnaces the heat absorbed materially affects the temperature of gases at the furnace outlet and contributes directly to the generation of steam. *See* STEAM.

Furnace walls. Prior to 1925, most furnaces were constructed of firebrick. As the capacity and the physical size of boiler units increased and the suspension burning of pulverized coal was developed, limitations were imposed by the height of the refractory walls that could be made self-supporting at high temperatures and by the inability of refractories to resist the fluxing action of molten fuel ash. Thus refractory is used primarily for special-purpose furnaces, for the burner walls of small boilers, or for the arches of stoker-fired boilers to promote the ignition of the fuel. *See* REFRACTORY.

The limitations of refractory constructions can be extended somewhat by cooling the brickwork with air flowing through channels in the structure, or by sectionalizing the wall into panels and transferring the structural load to external air-cooled steel or cast-iron supporting members. The heat absorbed can be recovered by using the cooling air for combustion, and the low rate of heat transfer through the refractory helps to maintain high furnace temperatures, thus accelerating ignition and the burning of the fuel.

The low tensile strength of refractories restricts the geometrical shapes that can be built, making it difficult to provide overhanging contours or roof closures. Thus sprung arches or shaped tiles suspended from steel must be used. Many refractory mixtures having air-setting or hydraulic-setting properties are available, and they can be used to form monolithic structures by ramming, guniting, or pouring in forms.

Water-cooling of all or most of the furnace walls is almost universally used in boiler units of all capacities and for all types of fuel and methods of firing. The water-cooling of the furnace walls reduces the transfer of heat to the structural members and, consequently, their temperatures are within the limits that provide satisfactory strength and resistance to oxidation. Water-cooled tube constructions facilitate large furnace dimensions and optimum arrangements of the roof, hopper, and arch, as well as the mountings for the burners and the provision for screens, platens, or division walls to increase the amount of heat-absorbing surface exposed in the combustion zone. External heat losses are small and are further reduced by the use of insulation.

Furnace wall tubes are spaced on close centers to obtain maximum heat absorption and to facilitate ease of ash removal (see **illus.**). The so-called tangent tube construction, wherein adjacent tubes are almost touching with only a small clearance provided for erection purposes, has been frequently used. However, most new boiler units use membrane-tube walls, in which a steel bar or membrane is welded

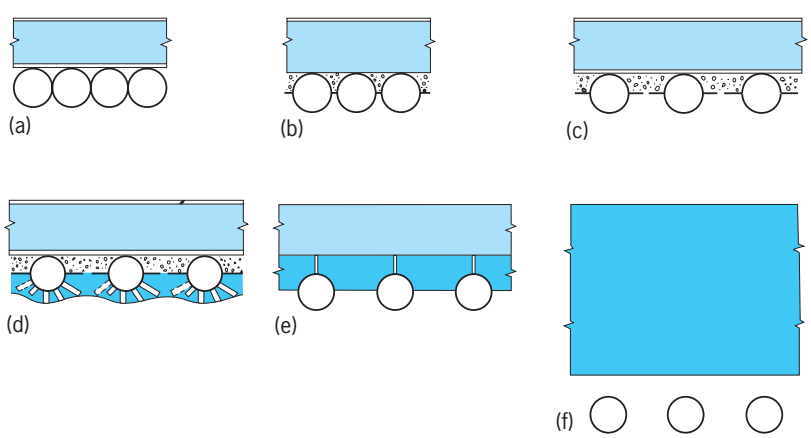

Water-cooled furnace-wall constructions. (*a*) Tangent tube. (*b*) Welded membrane and tube. (*c*) Flat stud and tube. (*d*) Full stud and refractory-covered tube. (*e*) Tube and tile. (*f*) Tubes backed by refractory. (*After T. Baumeister, ed., Marks' Standard Handbook for Mechanical Engineers, 8th ed., McGraw-Hill, 1978*)

between adjacent tubes. This construction facilitates the fabrication of water-cooled walls in large shop-assembled tube panels. Less effective cooling is obtained when the tubes are wider spaced and extended metal surfaces, in the form of flat studs, are welded to the tubes; if even less cooling is desired, the tube spacing can be increased and refractory installed between or behind the tubes to form the wall enclosure.

Furnace walls must be adequately supported, with provision for thermal expansion and with reinforcing buckstays to withstand the lateral forces resulting from the difference between the furnace pressure and the surrounding atmosphere. The enclosure walls must prevent air infiltration when the furnace is operated under suction, and must prevent gas leakage when the furnace is run at pressures above atmospheric.

Additional furnace cooling, in the form of tubular platens, division walls, or wide-spaced screens, often is used; in high heat input zones the tubes may be protected by refractory coverings anchored to the tubes by studs. *See* FURNACE CONSTRUCTION.

Heat transfer. Heat-absorbing surfaces in the furnace receive heat from the products of combustion and, consequently, lower the furnace gas temperature. The principal mechanisms of heat transfer take place simultaneously and include intersolid radiation from the fuel bed or fuel particles, nonluminous radiation from the products of combustion, convection from the furnace gases, and conduction through any deposits and the tube metal. The absorption effectiveness of the surface is greatly influenced by the deposits of ash or slag.

Analytical solutions to the transfer of heat in the furnaces of steam-generating units are extremely complex, and it is very difficult, if not impossible, to calculate furnace-outlet gas temperatures by theoretical methods. Nevertheless, the furnace-outlet gas temperature must be predicted accurately because this temperature determines the design of the remainder of the unit, particularly that of the superheater. Calculations therefore are based upon test results supplemented by data accumulated from operating experience and judgment predicated upon knowledge of the principles of heat transfer and the characteristics of fuels and slags. *See* HEAT TRANSFER; STEAM-GENERATING UNIT. George W. Kessler

Bibliography. E. A. Avallone and T. Baumeister III, *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; R. D. Reed, *Furnace Operations*, 3d ed., 1981; M. W. Thring, *The Science of Flames and Furnaces*, 1962.

Steam-generating unit

The wide diversity of parts, appurtenances, and functions needed to release and utilize a source of heat for the practical production of steam at pressures to 5000 lb/in.^{2} (34 megapascals) and temperatures to 1100°F (600°C), often referred to as a steam boiler for brevity. *See* STEAM.

The essential steps of the steam-generating process include (1) a furnace for the combustion of fuel, or a nuclear reactor for the release of heat by fission, or a waste heat system; (2) a pressure vessel in which feedwater is raised to the boiling temperature, evaporated into steam, and generally superheated beyond the saturation temperature; and (3) in many modern

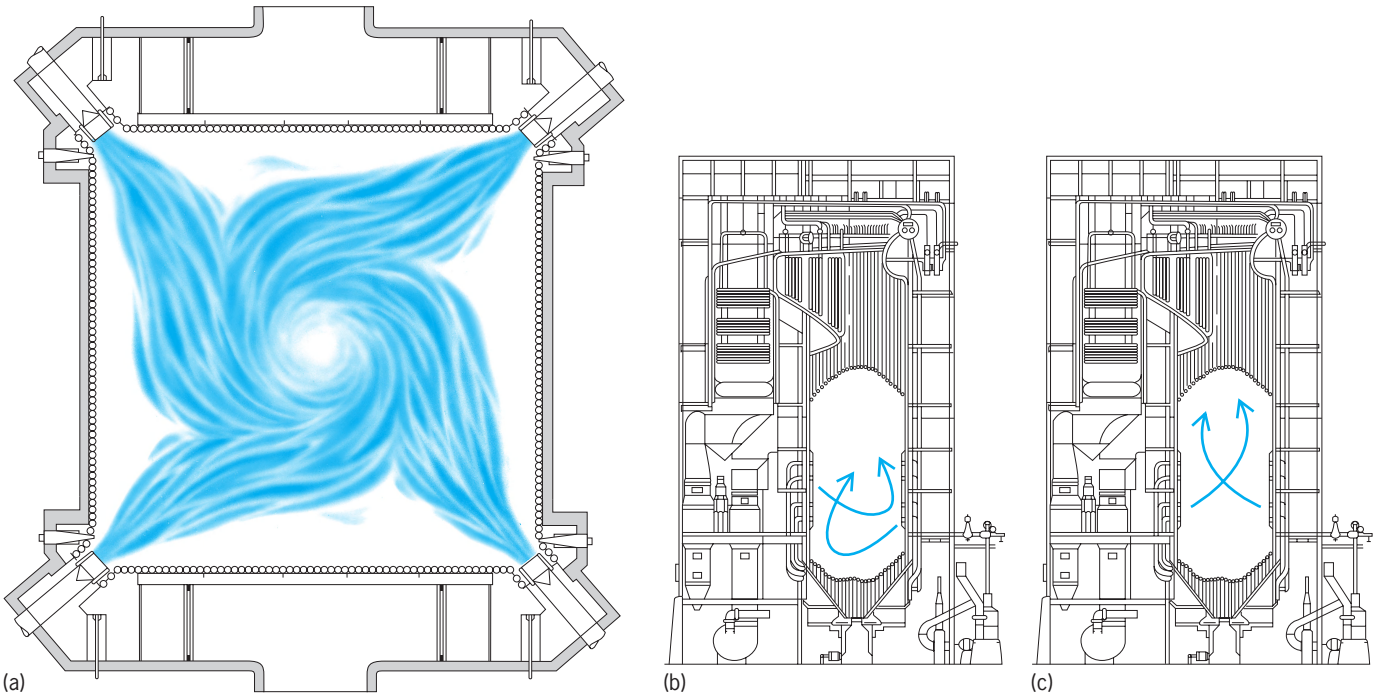

Fig. 1. Large, central-station steam-generating unit, with pump-assisted circulation, corner-fired with tangential, tilting burners for pulverized coal, oil, or gas. (*a*) Plan; (*b*) burners tilted down; (*c*) burners tilted up. Output is 2,000,000 lb steam/h (250 kg/s) at 2400 lb/in.^{2} (16.5 MPa), 1000°F (540°C), and 1000°F (540°C) reheat. (*Combustion Engineering, Inc.*)

central station units, a reheat section or sections for resuperheating steam after it has been partially expanded in a turbine. This aggregation of functions requires a wide assortment of components, which may be variously employed in the interests, primarily, of capacity and efficiency in the steam-production process. The selection, design, operation, and maintenance of these components constitute a complex process which, in the limit, calls for the highest technical skill.

Service requirements. The wide diversity of commercial steam-generating equipment stems from attempts to accommodate design to the dictates of the imposed service conditions.

Safety. The pressure parts, a potential bursting hazard, must be designed, maintained, and operated under the most stringent codes, such as the ASME Boiler and Pressure Vessel Code. The codes and regulations are backed by the police power of the state.

Shape. Physical shape must fit the limits imposed. The differences between marine and stationary applications are an example.

Bulk. Space occupied has various degrees of value, and the equipment must best comply with these dictates.

Weight. Frequently, weight is the lowest common denominator in determining portability and economic suitability.

Setting. Confinement of heat source operations and the ensuing transport function will vary the details of the setting, for example, the use of metal or of refractory enclosure.

Character of labor. Highly skilled labor, as in atomic power plants, can utilize design details that would be prohibitive in areas where the labor is primitive.

Cleanability. Impurities on the heat-transfer and heat-release surfaces must not impair safe or efficient operation; surfaces must be cleanable to give acceptable reliability and performance.

Life. Short or long life and short or long operating periods between overhauls are vital to the selection of a preferred design.

Efficiency. Inherent high efficiency can be built into a design, but economics will dictate the efficiency chosen; the full realization of built-in high efficiency requires the services of skilled operators and firefighters.

Cost. The overall cost of steam in cents per 1000 lb or per 1000 kg is the ultimate criterion, reflecting investment, operating, and maintenance expenses in commercial operations.

Adaptation to use. These service requirements differ in each installation so that a wide variety of designs is available in the commercial markets (**Figs. 1** and **2**). There is no one best design for universal application, but propriety rests in the proper selection (see **table**). The designs are characterized by the various degrees of applicability of component parts such as steam drums and tubes; fuel-burning equipment and furnaces; draft systems for air supply and flue gas removal; blowers, fans, and stacks; heat traps such as economizers and air preheaters; structural steel for the support of parts with ample provision for expansion; a casing or setting with possible utilization of water walls, refractory, and insulation; pumps for boiler feed and for water circulation; fuel- and refuse-handling systems; feedwater purification systems; blowdown systems and soot-blowing equipment; and a wide assortment of accessories and instruments such as pressure gages, safety valves, water-level indicators, and sophisticated automatic control with its elaborate interlocks for foolproof, efficient operation and complete computerization.

The maintenance of a safe working temperature for metal parts requires ample circulation of water over the steam-generating parts and ample circulation of steam over the superheating and reheating parts. Water circulation in the generating sections may be by natural convection processes where ample physical height gives the circulatory pressure difference between a column of water in the

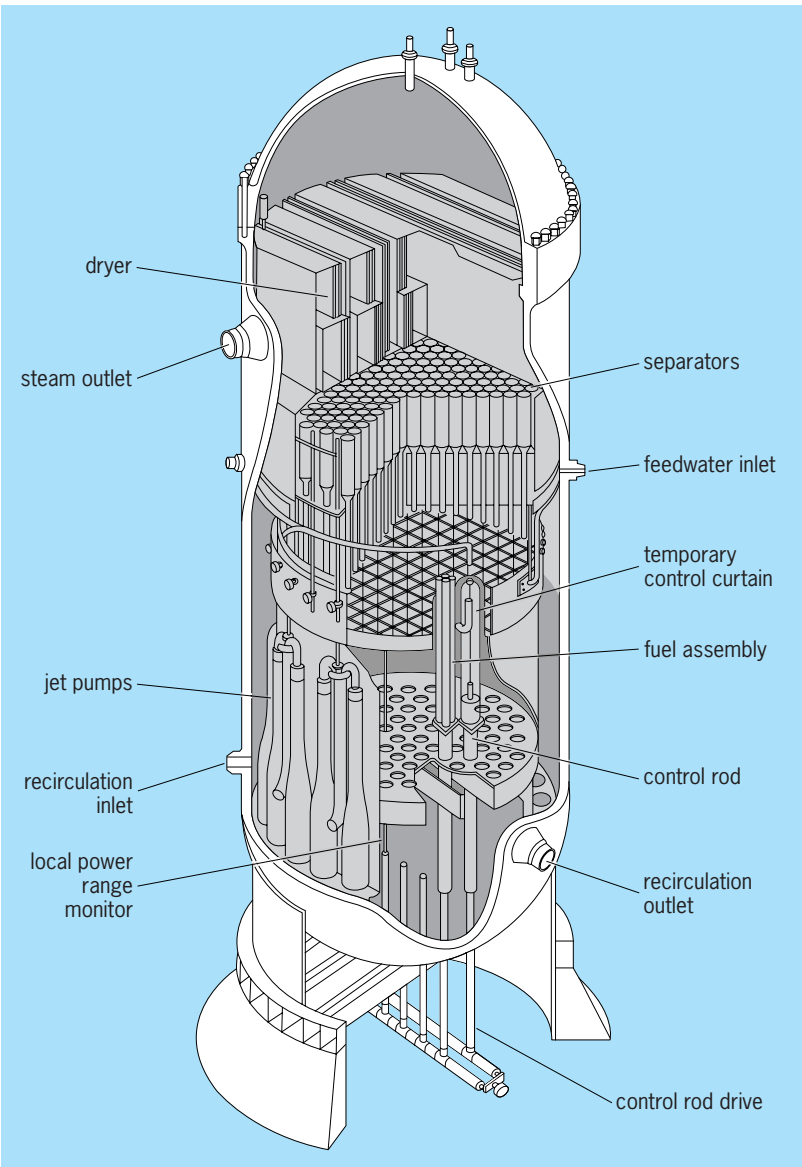

Fig. 2. Boiling-water reactor (BWR) for a large central-station electric generating plant showing essential components. Typical performance is 6,000,000 lb/h (750 kg/s) at 1000 lb/in.^{2} (7 MPa), saturated steam. (*General Electric Co.*)

Performance characteristics of some steam-generating units

Boiler type	Fuel (potential)	Steam output, lb/h (kg/s)	Steam pressure, lb/in.^{2} (MPa)	Steam temperature, °F (°C)	Boiler efficiency, %
Horizontal return tubular (HRT), fire-tube	Coal, oil, gas	10,000 (1.3)	150 (1.0)	Dry and saturated	70
Large industrial, natural-circulation, water-tube	Coal, oil, gas	300,000 (38)	1400 (9.7)	900 (480)	88
Large central station, natural-circulation	Coal, oil, gas	2,000,000 (250)	2700 (18.6)	1050/1050 (570/570)	90
Large central station, supercritical, once-through	Coal, oil, gas	5,000,000 (630)	3600 (24.8)	1000/1000/1000 (540/540/540)	90
Boiling-water reactor (BWR)	Uranium	6,000,000 (750)	1000 (6.9)	Dry and saturated (550 ± °F or 290 ± °C)	98±
Pressurized-water reactor (PWR), hot water, 4 loops	Uranium	6,000,000 (750)	700 (4.8)	Dry and saturated (500 ± °F or 260 ± °C)	98
Electric boiler, small	200± volts	2,000 (0.25)	150 (1.0)	Dry and saturated	98±

downcomer and a column of steam mixed with water in the riser parts. With high operating steam pressures, such as 1400 lb/in.^{2} (9.7 MPa), pump-assisted circulation is often selected. With supercritical operation (above 3200 lb/in.^{2} or 22.1 MPa) there is no difference in density between liquid and vapor so that forced-flow, once-through principles supersede water recirculation practice. Such supercritical steam generators have no steam or separating drums and, like nuclear steam generators, require the highest purity of feed and boiler water (typically 10^{-9} part of impurities) to avoid deposits in heated circuits or the transport of solids to the turbine. *See* BOILER; BOILER ECONOMIZER; FIRE-TUBE BOILER; MARINE BOILER; RAW WATER; REHEATING; STEAM SEPARATOR; STEAM TEMPERATURE CONTROL; SUPERHEATER; WATER-TUBE BOILER. Theodore Baumeister

Bibliography. American Society of Mechanical Engineers, *Boiler And Pressure Vessel Code 2004*, 2003; E. A. Avallone, T. Baumeister III, and A. Sadegh (eds.), *Mark's Standard Handbook for Mechanical Engineers*, 11th ed., 2006; J. J. Jackson, *Steam Boiler Operation; Principles and Practices*, 2d ed., 1987; E. B. Woodruff, H. B. Lammers, and T. F. Lammers, *Steam Plant Operation*, 8th ed., 2004.

Steam heating

A heating system that uses steam generated from a boiler. The steam heating system conveys steam through pipes to heat exchangers, such as radiators, convectors, baseboard units, radiant panels, or fan-driven heaters, and returns the resulting condensed water to the boiler. Such systems normally operate at pressure not exceeding 15 lb/in.^{2} gage or 103 kilopascals gage, and in many designs the condensed steam returns to the boiler by gravity because of the static head of water in the return piping. With utilization of available operating and safety control devices, these systems can be designed to operate automatically with minimal maintenance and attention.

One-pipe system. In a one-pipe steam heating system, a single main serves the dual purpose of supplying steam to the heat exchanger and conveying condensate from it. Ordinarily, there is but one connection to the radiator or heat exchanger, and this connection serves as both the supply and return; separate supply and return connections are sometimes used. Because steam cannot flow through the piping or into the heat exchanger until all the air is expelled, automatic air-venting valves must be provided on all exchangers and at the ends of all mains. These valves may be of a type which closes whenever steam or water comes in contact with the operating element but which also permits air to flow back into the system as the pressure drops. A vacuum valve closes against subatmospheric pressure to prevent return of air.

Two-pipe system. A two-pipe system is provided with two connections from each heat exchanger, and in this system steam and condensate flow in separate mains and branches (**Fig. 1**). A vapor two-pipe system operates at a few ounces above atmospheric pressure, and in this system a thermostatic trap is located at the discharge connection from the heat exchanger which prevents steam passage, but permits air and condensation to flow into the return piping.

When the steam condensate cannot be returned by gravity to the boiler in a two-pipe system, an alternating return lifting trap, condensate return pump, or vacuum return pump must be used to force the condensate back into the boiler. In a condensate return-pump arrangement, the return piping is arranged for the water to flow by gravity into a collecting receiver or tank, which may be located below the steam-boiler waterline. A motor-driven pump controlled from the boiler water level then forces the condensate back to the boiler.

In large buildings extending over a considerable area, it is difficult to locate all heat exchangers above the boiler water level or return piping. For these systems a vacuum pump is used that maintains a suction below atmosphere up to ±25 in. or ±64 cm (max) of mercury in the return piping, thus creating a positive return flow of air and condensate back to the pumping unit. Subatmospheric systems are similar to vacuum systems, but in contrast provide a means of partial vacuum control on both the supply and

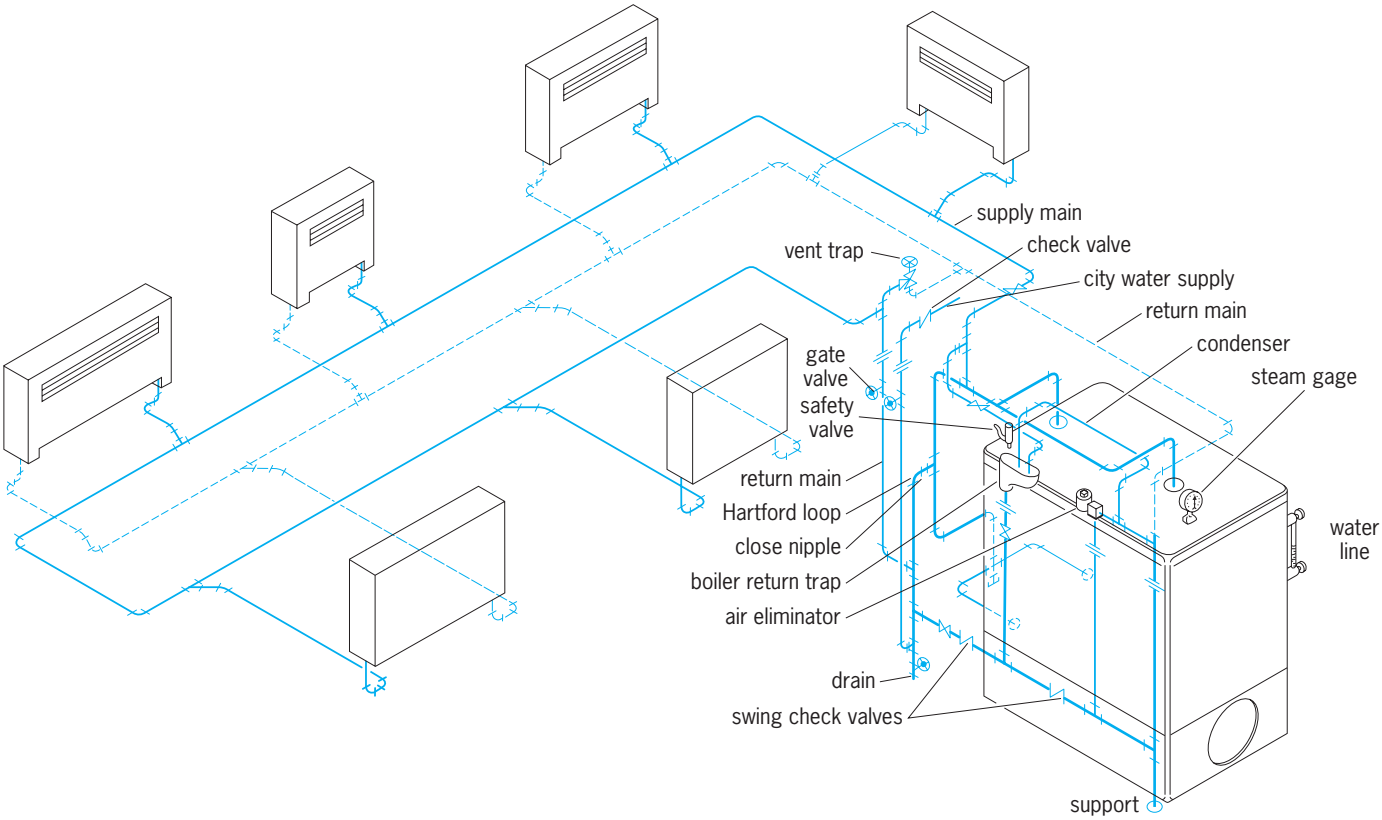

Fig. 1. Two-pipe up-feed system with automatic return trap.

return piping so that the steam temperature can be regulated to vary the heat emission from the heat exchanger in direct proportion to the heat loss from the structure.

Figure 2 depicts a two-pipe vacuum heating system which uses a condensation pump as a mechanical lift for systems where a part of the heating system is below the boiler room. Note that the low section of the system is maintained under the same vacuum conditions as the rest of the system. This is

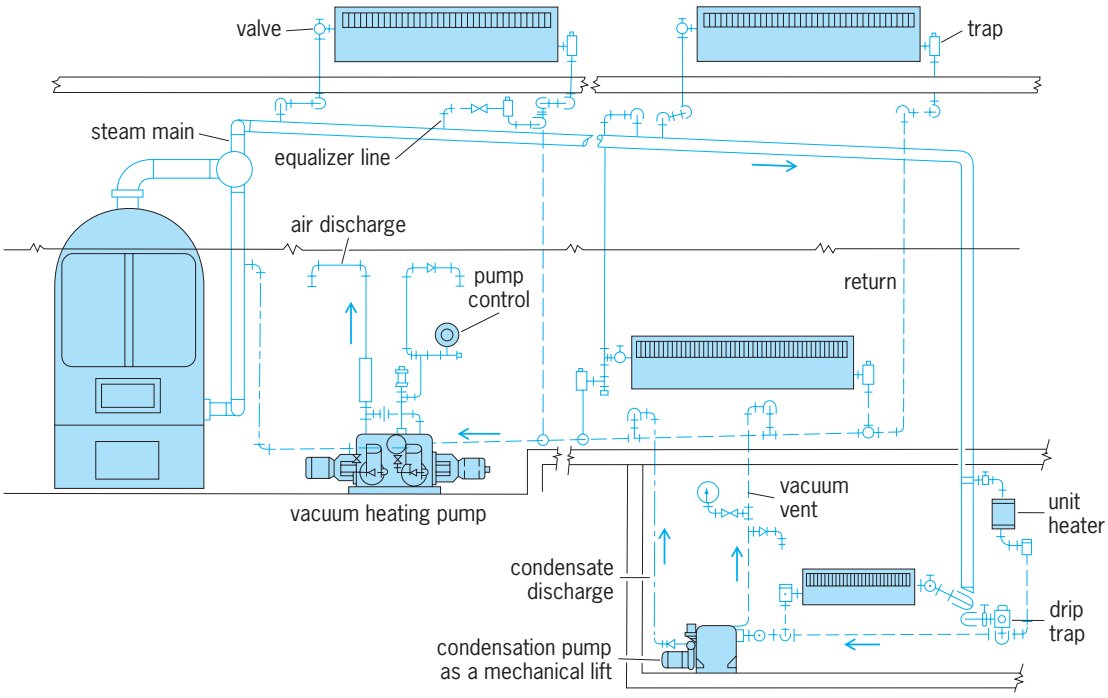

Fig. 2. Layout of a vacuum heating system with condensation and vacuum pumps.

accomplished by connecting the vent from the receiver and pump discharge to the return pipe located above the vacuum heating pump.

With the wide acceptance of all-year air conditioning, low-pressure steam boilers have been used to produce cooling from absorption refrigeration equipment. With this system the boiler may be used for primary or supplementary steam heating.

Exhaust from gas- or oil-driven turbines or engines may be used in waste heat boilers or separators, along with a standby boiler to produce steam for a heating system.

Another source for steam for heating is from a high-temperature water source (350–450°F or 180–230°C) using a high-pressure water to low-pressure steam heat exchanger. *See* BOILER; COMFORT HEATING; OIL BURNER; STEAM-GENERATING FURNACE.

John W. James

Bibliography. M. Greenwald, *Residential Hot Water and Steam Heating with Gas, Oil and Solid Fuels,* 1987; E. B. Woodruff, H. B. Lammers, and T. F. Lammers, *Steam-Plant Operation*, 8th ed., 2004.

Steam jet ejector

A steam-actuated device for pumping compressible fluids, usually from subatmospheric suction pressure to atmospheric discharge pressure. A steam jet ejector is most frequently used for maintaining vacuum in process equipment in which evaporation or condensation takes place. Because of its simplicity, compactness, reliability, and generally low first cost, it is often preferred to a mechanical vacuum pump for removing air from condensers serving steam turbines, especially for marine service. *See* VACUUM PUMP.

Principle. Compression of the pumped fluid is accomplished in one or more stages, depending upon the total compression required. Each stage consists of a converging-diverging steam nozzle, a suction chamber, and a venturi-shaped diffuser. Steam is supplied to the nozzle at pressures in the range 100–250 lb/in.^{2} gage (0.7–1.7 megapascals). A portion of the enthalpy of the steam is converted to kinetic energy by expanding it through the nozzle at substantially constant entropy to ejector suction pressure, where it reaches velocities of 3000–4500 ft/s (900–1400 m/s). The air or gas, with its vapor of saturation, which is to be pumped and compressed is entrained, primarily by friction in the high-velocity steam jet. The impulse of the steam produces a change in the momentum of the air or gas vapor mixture as it mixes with the motive steam and travels into the converging section of the diffuser. In the throat or most restricted area of the diffuser, the energy transfer is completed and the final mixture of gases and vapor enters the diverging section of the diffuser, where its velocity is progressively reduced. Here a portion of the kinetic energy of the mixture is reconverted to pressure with a corresponding increase in enthalpy. Thus the air or gas is compressed to a higher pressure than its entrance pressure to the ejector. The compression ratios selected for each stage of a steam jet ejector usually vary from about 4 to 7. *See* DIFFUSER.

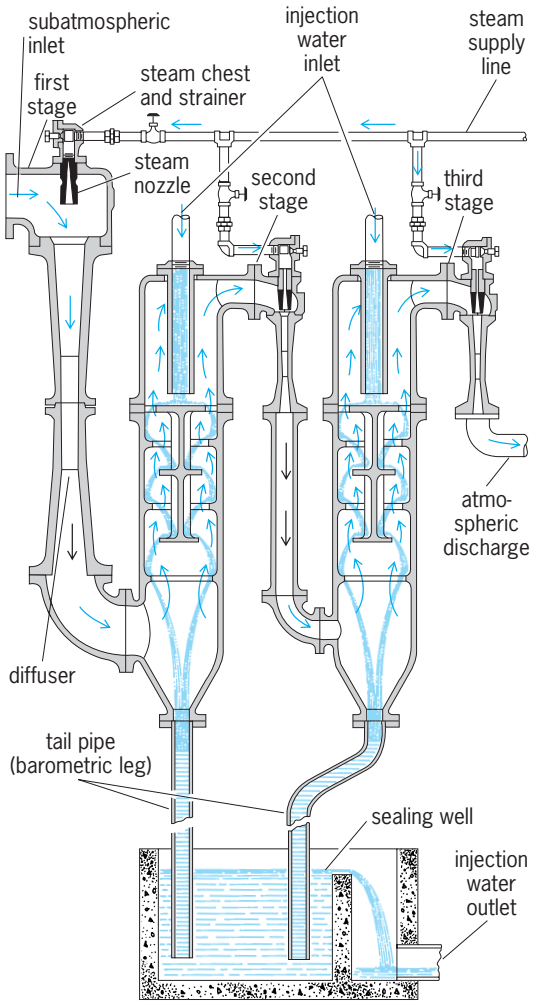

Fig. 1. Typical multistage steam jet ejector with contact barometric condensers; first and second stages condensing and third stage noncondensing.

Application. Two or more stages may be arranged in series, depending upon the total compression ratio required (**Fig. 1**). Two or more sets of series stages may be arranged in parallel to accommodate variations in capacity.

Vapor condensers are usually interposed between the compression stages of multistage steam jet ejectors to condense and remove a significant portion of the motive steam and other condensable vapors (**Fig. 2**). This action reduces the amount of fluid to be compressed by the next higher stage and results in a reduction in the motive steam required. Both surface and contact types of vapor condensers are used for this purpose. *See* VAPOR CONDENSER.

Ejectors used as air pumps for steam condensers that serve turbines are usually two-stage and are equipped with inter- and aftercondensers of the surface type. The steam condensed is drained through traps to the main condenser and returned to the boiler feed system. Ejectors used as vacuum pumps in process systems may be equipped with either surface condensers or contact condensers of the barometric type between or after stages or both. They

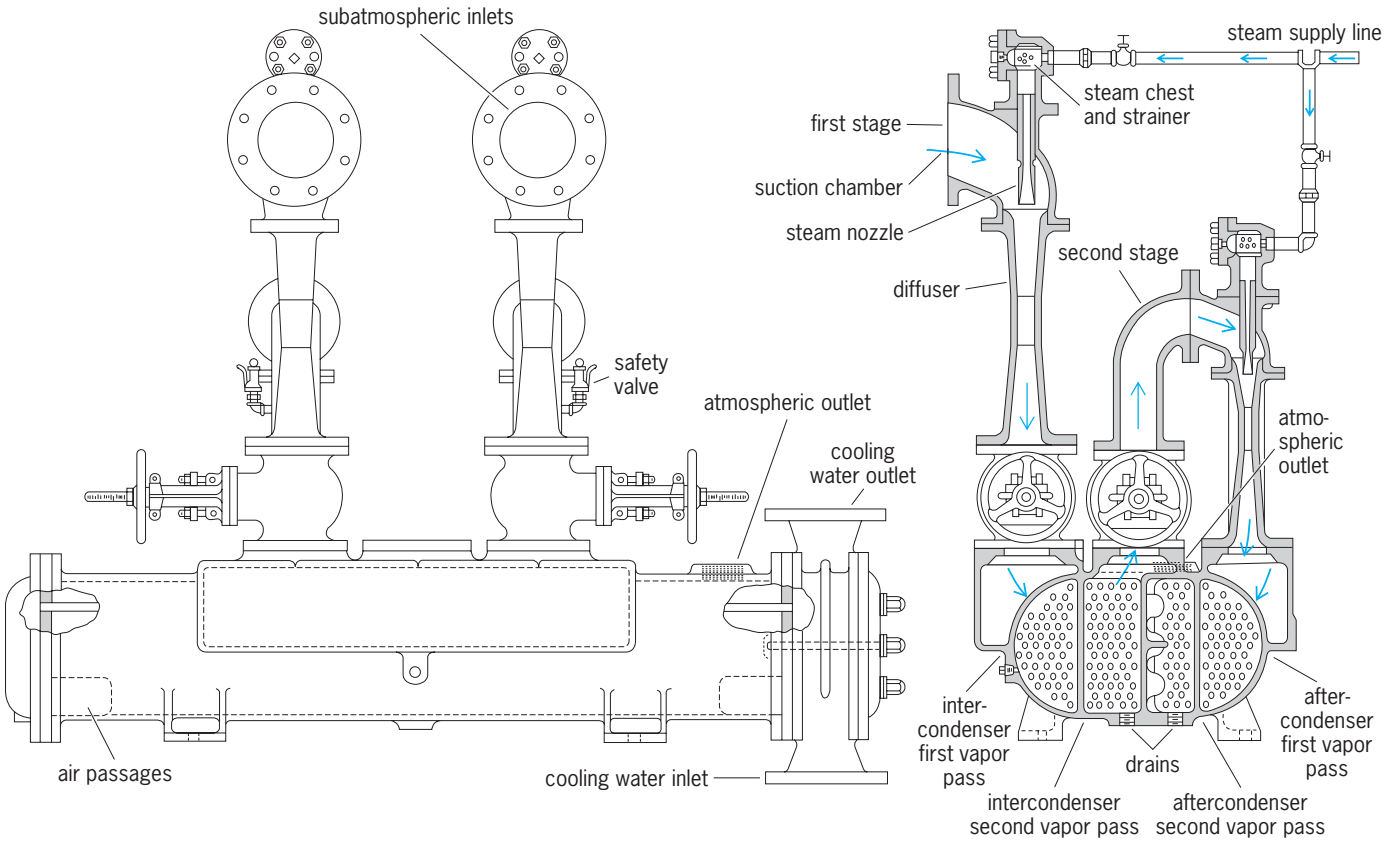

Fig. 2. Multistage steam jet ejector with surface inter- and aftercondenser; first and second stages condensing.

may be single or multistage machines. High-vacuum process ejectors with as many as seven stages in series have been built. Industrial or process ejectors are frequently used instead of mechanical vacuum pumps to pump corrosive vapors because they can be manufactured economically from almost any corrosion-resistant material.

The number of stages of compression usually used for various suction pressures with atmospheric discharge pressure is as follows:

No. of ejector stages	*Range of suction pressure*
1	3–30 in. Hg (10–102 kPa) abs.
2	0.4–4 in. Hg (1.4–14 kPa) abs.
3	1–25 mmHg (0.13–3.3 kPa) abs.
4	0.15–3 mmHg (20–400 Pa) abs.
5	20–300 μm (2.7–40 Pa)
6	5–20 μm (0.7–2.7 Pa)
7	1–5 μm (0.13–0.7 Pa)

Ejectors are also made which use air or other gases instead of steam as the energy source.

Joseph F. Sebald

Steam separator

A device for separating a mixture of the liquid and vapor phases of water. Steam separators are used in most boilers and may also be used in saturated steam lines to separate and remove the moisture formed because of heat loss.

In boilers, steam separators must perform efficiently because both steam-free water and water-free steam are required. The force promoting water circulation in a natural-circulation boiler is derived from the density difference between the water in the downcomer tubes and the steam-water mixture in the riser tubes. Solid (steam-free) water in the downcomers maximizes this force. Steam-free water is equally important in a forced-circulation boiler to prevent cavitation in the pump.

Water-free steam is quite necessary because dissolved and suspended solids entering the boiler with the feedwater concentrate in the boiler water, and such solids contaminate the steam if all the water is not separated. Impurity concentrations as low as a few parts per billion can cause solid deposits to form in turbines or can cause product contamination if the steam is used in direct contact with a product. Solid deposits in turbines decrease turbine capacity and efficiency and may even cause unbalance of the rotor in extreme cases. Consequently, the sampling of steam and the measurement of impurities in steam have received much attention.

Many low-pressure boilers operate satisfactorily without steam separators because their capacity ratings are quite conservative. Steam separation in such boilers may be compared with a pan of water boiling on a stove at a moderate rate. However, it is known that if the fire under the pan is turned too high or

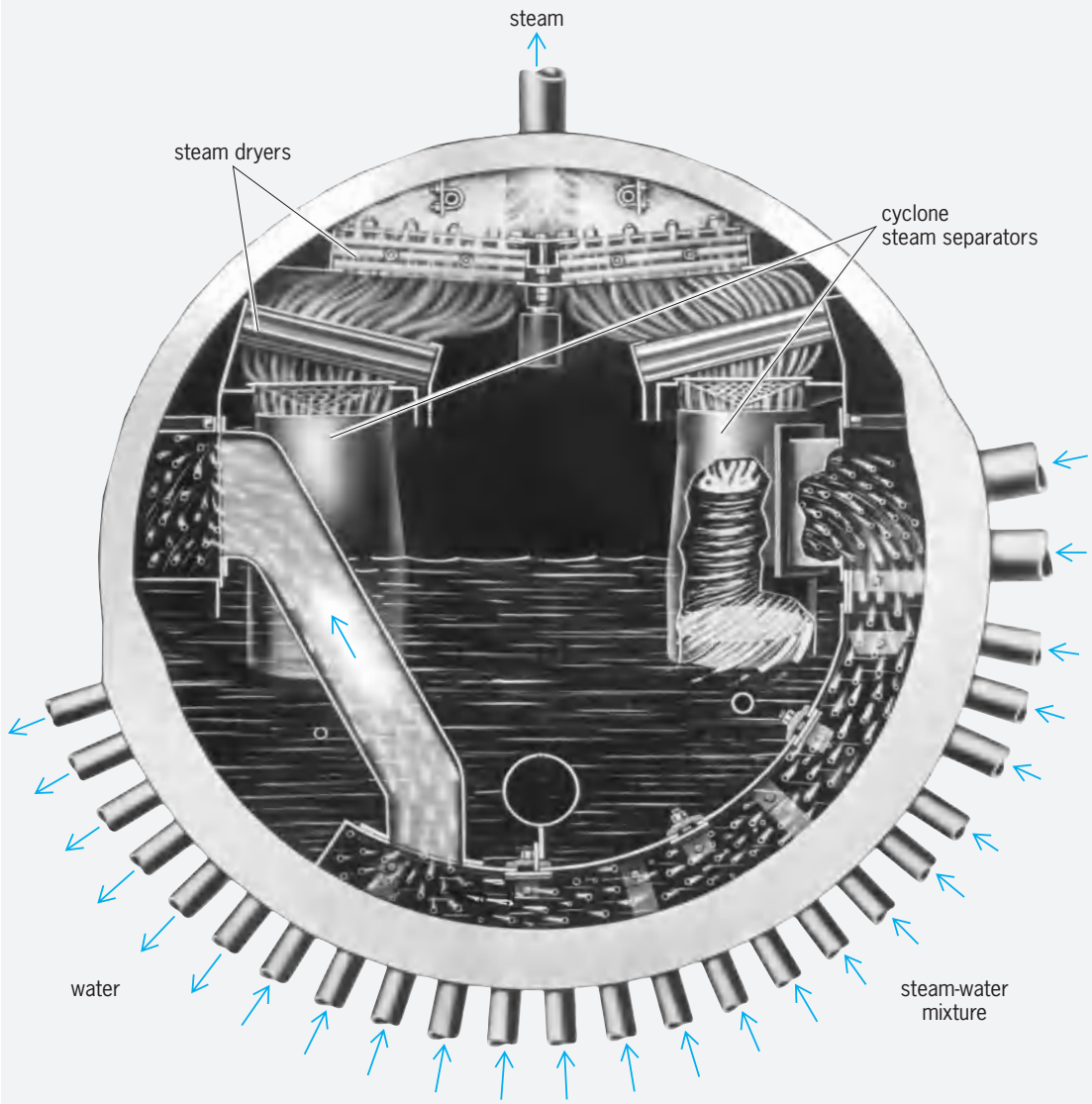

Steam dryers and cyclone steam separators separate steam and water in high-capacity boiler unit. Arrows show direction of flow.

if the water contains certain solids (like milk, for example), separation becomes inadequate and the pan boils over. Similar phenomena take place in natural separation boilers if the load is too high or the concentration of solids in the water is too great.

Operating pressure also influences natural separation. The force effecting separation results from the difference in density between water and steam. This difference is quite great at low pressures, but as pressure is increased, the density of water decreases while that of steam increases, resulting in a steadily declining density difference.

Because of these effects, steam separators are used on the majority of boilers. Steam separators have many forms and may be as fundamental as a simple baffle that utilizes inertia caused by a change of direction. Most modern, high-capacity boilers use a combination of cyclone separators and steam dryers, as shown in the **illustration**.

Cyclone separators have two major advantages. First, they slice the steam-water mixture into thin streams so that the steam bubbles need travel only short distances through the mixture to become disengaged. Second, they whirl the mixture in a circular path, creating a centrifugal force many times greater than the force of gravity.

Cyclone separators remove practically all of the steam from the water, producing the steam-free water needed for the boiler downcomer tubes. They also remove the major portion of the water from the steam. After leaving the cyclone separators, the steam is passed through the dryers, where the last traces of moisture are removed. Steam dryers remove small droplets of water from the steam by providing a series of changes in direction and a large surface area to intercept the droplets.

Because of the efficient operation of steam separators, steam is one of the purest mass-produced commodities made. *See* BOILER; BOILER FEEDWATER; STEAM; STEAM TURBINE. Earl E. Coulter

Bibliography. American Society for Testing and Materials, *Methods of Sampling Steam*, ASTM

Designation D1066; ASTM, *Methods of Test for Deposit-Forming Impurities in Steam*, ASTM Designation D2186; ASTM, *ASTM Standards*, pt. 23: *Water:Atmospheric Analysis*, annually; P. M. Brister, F. G. Raynor, and E. A. Pirsh, *The Influence of Boiler Design and Operating Conditions on Steam Contamination*, ASME Pap. 51–A–95; M. J. Moore and C. H. Sieverding, *Two-Phase Steam Flow in Turbines and Separators*: *Theory*, *Instrumentation*, *Engineering*, 1976.

Steam temperature control

Means for regulating the operation of a steam-generating unit to produce steam at the required temperature. The temperature of steam is affected by the change in the relative heat absorption as load varies, by changes in ash or slag deposits on the heat-absorbing surface, by changes in fuel, by changes in the proportioning of fuel and combustion air, or by changes in feedwater temperature. Low steam temperature lowers the efficiency of the thermal cycle. However, high steam temperature, which increases thermal efficiency, is restricted by the strength and durability of materials used in superheaters. Control of steam temperature is, therefore, a matter of primary concern in the design of modern steam-generating units.

Steam temperature can be controlled by one or more of several methods (see **illus.**). These include

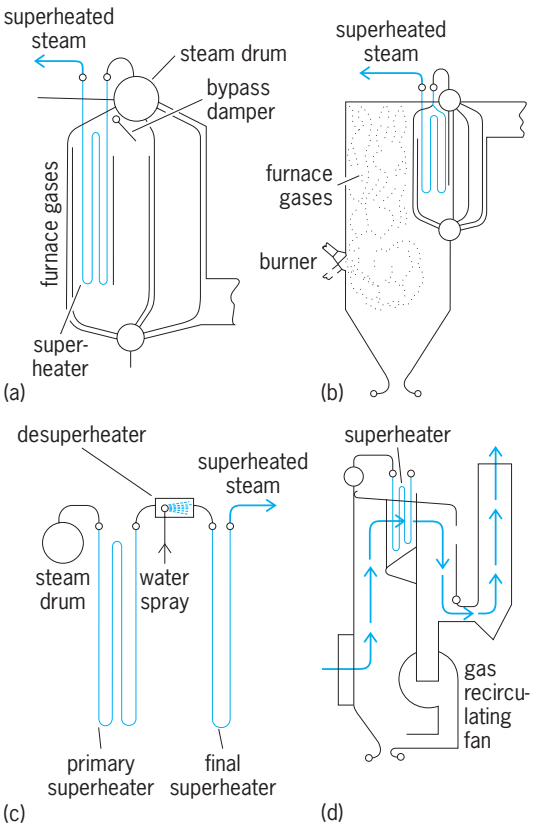

Methods of controlling steam temperature. (*a*) Bypass dampers. (*b*) Tilting burners. (*c*) Spray attemperation. (*d*) Gas recirculation.

(1) the damper control of gases to the superheater, to the reheater, or to both, thus changing the heat input; (2) the recirculation of low-temperature flue gas to the furnace, thus changing the relative amounts of heat absorbed in the furnace and in the superheater, reheater, or both; (3) the selective use of burners at different elevations in the furnace or the use of tilting burners, thus changing the location of the combustion zone with respect to the furnace heat-absorbing surface; (4) the attemperation, or controlled cooling, of the steam by the injection of spray water or by the passage of a portion of the steam through a heat exchanger submerged in the boiler water; (5) the control of the firing rate in divided furnaces; and (6) the control of the firing rate relative to the pumping rate of the feedwater to forced-flow once-through boilers. Generally, these various controls are adjusted automatically. *See* STEAM-GENERATING UNIT.

George W. Kessler

Steam turbine

A machine for generating mechanical power in rotary motion from the energy of steam at temperature and pressure above that of an available sink. By far the most widely used and most powerful turbines are those driven by steam. In the United States well over 85% of the electrical energy consumed is produced by steam-turbine-driven generators. Individual turbine ratings historically have tended to follow the increasing capacity trend but have reached limits imposed by material and machine design considerations. The largest unit shipped during the 1950s was rated 500 MW. Units rated about 1100 MW were in service by the close of the 1960s, and ratings up to 1300 MW saw frequent application in the 1970s and early 1980s. Units of all sizes, from a few horsepower to the largest, have their applications.

Until the 1960s essentially all steam used in turbine cycles was produced in boilers burning fossil fuels (coal, oil, and gas) or, in minor quantities, certain waste products. The 1960s marked the beginning of the introduction of commercial nuclear power. *See* ELECTRIC POWER GENERATION; NUCLEAR POWER.

Turbine parts. Figure 1 shows a small, simple mechanical-drive turbine of a few horsepower. It illustrates the essential parts for all steam turbines regardless of rating or complexity: (1) a casing, or shell, usually divided at the horizontal center line, with the halves bolted together for ease of assembly and disassembly; it contains the stationary blade system; (2) a rotor carrying the moving buckets (blades or vanes) either on wheels or drums, with bearing journals on the ends of the rotor; (3) a set of bearings attached to the casing to support the shaft; (4) a governor and valve system for regulating the speed and power of the turbine by controlling the steam flow, and an oil system for lubrication of the bearings and, on all but the smallest machines, for operating the control valves by a relay system connected with the governor; (5) a coupling to connect with the driven machine; and (6) pipe connections to the steam supply

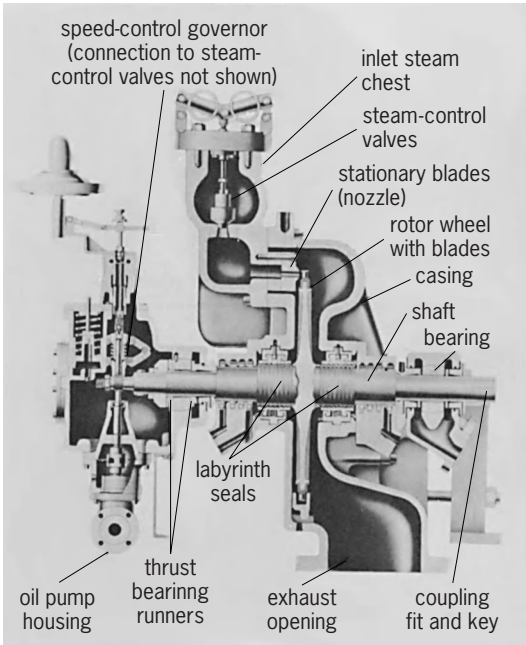

Fig. 1. Cutaway of small single-stage steam turbine. (*General Electric Co.*)

at the inlet and to an exhaust system at the outlet of the casing or shell.

Applications. Steam turbines are ideal prime movers for applications requiring rotational mechanical input power. They can deliver constant or variable speed and are capable of close speed control. Drive applications include centrifugal pumps, compressors, ship propellers, and, most importantly, electric generators.

The turbine shown in Fig. 1 is a small mechanical-drive unit. Units of this general type provide 10–1000 hp (7.5–750 kW) with steam at 100–600 lb/in.^{2} gage (0.7–4.1 megapascals) inlet pressure and temperatures to 800°F (427°C). These and larger multistage machines drive small electric generators, pumps, blowers, air and gas compressors, and paper machines. A useful feature is that the turbine can be equipped with an adjustable-speed governor and thus be made capable of producing power over a wide range of rotational speeds. In such applications efficiency varies with speed (**Fig. 2**), being 0 when the rotor stalls at maximum torque and also 0 at the runaway speed at which the output torque is 0. Maximum efficiency and power occur where the product

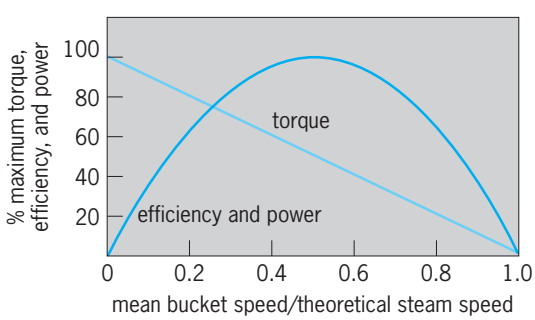

Fig. 2. Illustrative stage performance versus speed.

of speed and torque is the maximum.

Many industries need steam at one or more pressures (and consequently temperatures) for heating and process purposes. Frequently it is more economical to produce steam at high pressure, expand it partially through a turbine, and then extract it for further processing, rather than using a separate boiler at the process steam pressure. **Figure 3** is a cross section through an industrial automatic extraction turbine. The left set of valves admits steam from the boiler at the flow rate to provide the desired electrical load. The steam flows through five stages to the controlled extraction point. The second set of valves acts to maintain the desired extraction pressure by varying the flow through the remaining 12 stages. Opening these internal valves increases the flow to the condenser and lowers the controlled extraction pressure.

Industrial turbines are custom-built in a wide variety of ratings for steam pressures to 2000 lb/in.^{2} gage (14 MPa), for temperatures to 1000°F (538°C), and in various combinations of nonextracting, single and double automatic extraction, noncondensing and condensing. Turbines exhausting at or above atmospheric pressure are classed as noncondensing regardless of what is done with the steam after it leaves the turbine. If the pressure at the exhaust flange is less than atmospheric, the turbine is classed as condensing.

Turbines in sizes to about 75,000 hp (56 MW) are used for ship propulsion. The drive is always through reduction gearing (either mechanical or electrical) because the turbine speed is in the range of 4000–10,000 rpm, while 50–200 rpm is desirable for ship propellers. Modern propulsion plants are designed for steam conditions to 1450 lb/in.^{2} gage (10.0 MPa) and 950°F (510°C) with resuperheating to 950°F (510°C). Fuel consumption rates as low as 0.4 lb of oil per shaft-horsepower-hour (0.07 kg/megajoule) are achieved. *See* SHIP POWERING, MANEUVERING, AND SEAKEEPING.

Central station generation of electric power provides the largest and most important single application of steam turbines. Ratings smaller than 50 MW are seldom employed today; newer units are rated as large as 1300 MW. Large turbines for electric power production are designed for the efficient use of steam in a heat cycle that involves extraction of steam for feedwater heating, resuperheating of the main steam flow (in fossil-fuel cycles), and exhausting at the lowest possible pressure economically consistent with the temperature of the available condenser cooling water.

In fossil-fuel-fired cycles, steam pressures are usually in the range of 1800–3500 lb/in.^{2} gage (12.4–24.1 MPa) and tend to increase with rating. Temperatures of 950–1050°F (510–570°C) are used, with 1000°F (540°C) the most common. Single resuperheat of the steam to 950–1050°F (510–570°C) is almost universal. A second resuperheating is occasionally employed. **Figure 4** shows a typical unit designed for fossil-fuel steam conditions. Tandem-compound double-flow machines of this general

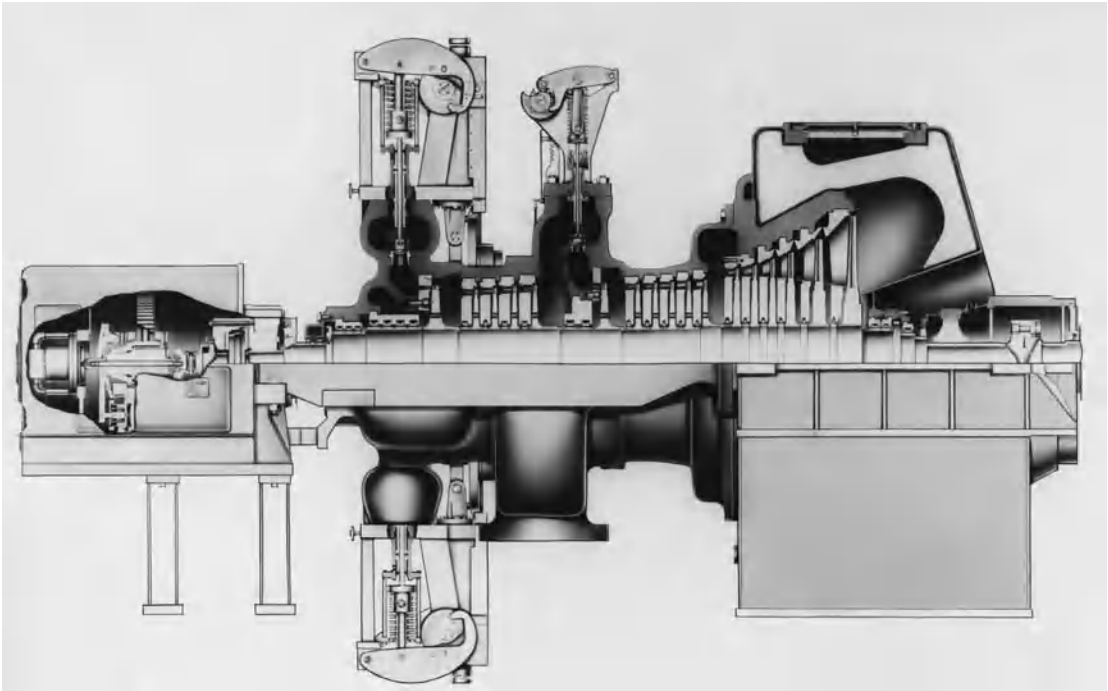

Fig. 3. Cross-section view of single-automatic-extraction condensing steam turbine. (*General Electric Co.*)

arrangement are applied over the rating range of 100–400 MW. Initial steam flows through the steam admission valves and passes to the left through the high-pressure portion of the opposed flow rotor. After resuperheating in the boiler it is readmitted through the intercept valves and flows to the right through the reheat stages, then crosses over to the double-flow low-pressure rotor, and exhausts downward to the condenser. *See* SUPERHEATER.

The water-cooled nuclear reactor systems common in the United States provide steam at pressures of about 1000 lb/in.^{2} gage. (6.9 MPa), with little or no initial superheat. Temperatures higher than about 600°F (315°C) are not available. Further, reactor containment and heat-exchanger considerations preclude the practical use of resuperheating at the reactor. The boiling-water reactor, for example, provides steam to the turbine cycle at 950-lb/in.^{2} gage (6.55-MPa) pressure and the saturation temperature of 540°F (282°C). Such low steam conditions mean that each unit of steam flow through the turbine produces less power than in a fossil-fuel cycle.

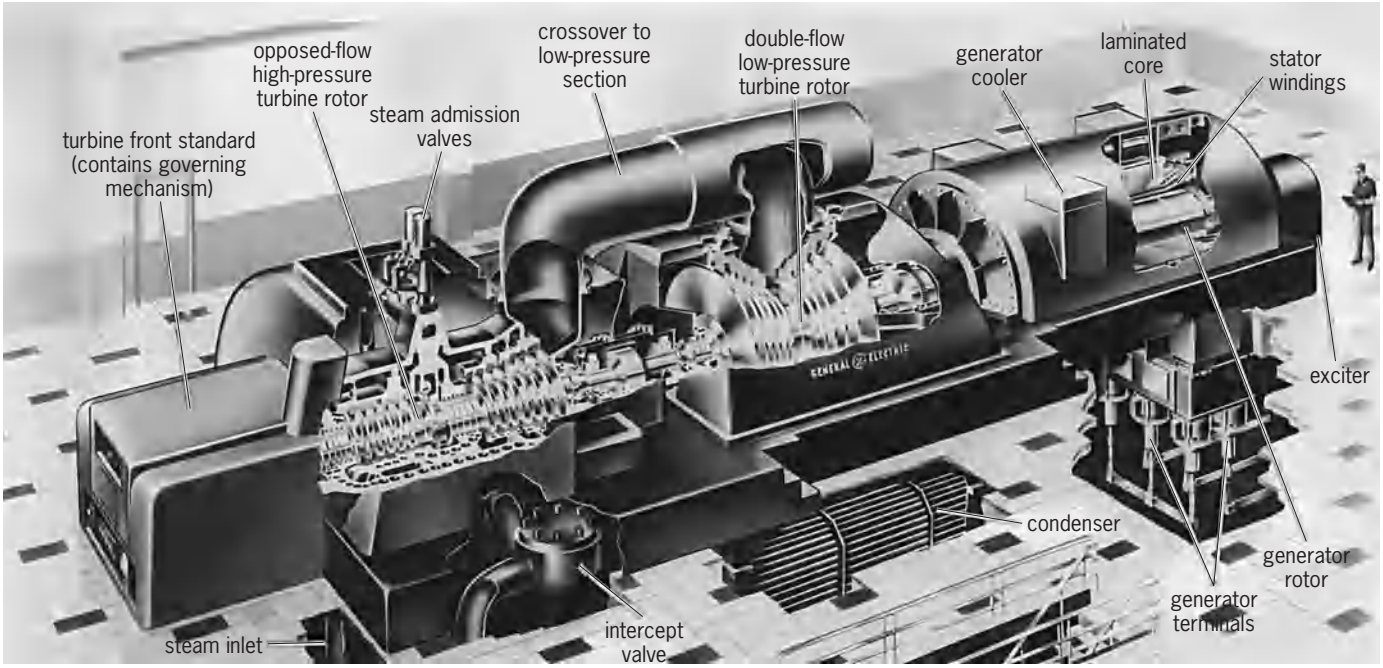

Fig. 4. Partial cutaway view of 3600-rpm fossil-fuel turbine generator. (*General Electric Co.*)

Fewer stages in series are needed but more total flow must be accommodated for a given output. Nuclear steam conditions often produce a turbine expansion with water of condensation present throughout the entire stream path. Provisions must be made to control the adverse effects of water: erosion, corrosion, and efficiency loss. In consequence of the differences in stream conditions, the design for a nuclear turbine differs considerably from that of a turbine for fossil fuel application. The former tend to be larger, heavier, and more costly. For example, at 800 MW, a typical fossil-fuel turbine can be built to run at 3600 rpm, is about 90 ft (27 m) long, and weighs about 1000 tons (900 metric tons). A comparable unit for a water-cooled nuclear reactor requires 1800 rpm, is about 125 ft (38 m) long, and weighs about 2500 tons (2300 metric tons). *See* NUCLEAR REACTOR.

Figure 5 represents a large nuclear turbine generator suitable for ratings of 1000 to 1300 MW. Steam from the reactor is admitted to a double-flow high-pressure section, at the left, through four parallel pairs of stop and control valves, not shown. The stop valves are normally fully open and are tripped shut to prevent dangerous overspeed of the unit in the event of loss of electrical load combined with control malfunction. The control valves regulate output by varying the steam flow rate. The steam exhausting from the high-pressure section is at 150 to 200 lb/in.^{2} abs. (1.0 to 1.4 MPa absolute pressure) and contains about 13% liquid water by weight. The horizontal cylinder alongside the foundation is one of a pair of symmetrically disposed vessels performing three functions. A moisture separator removes most of the water in the entering steam. Two steam-to-steam reheaters follow. Each is a U-tube bundle which condenses heating steam within the tubes and superheats the main steam flow on the shell side. The first stage uses heating steam extracted from the high-pressure turbine. The final stage employs reactor steam which permits reheating to near initial temperature. Alternate cycles employ reheat with initial steam only or moisture separation alone with no reheat. Reheat enhances cycle efficiency at the expense of increased investment and complexity. The final choice is economic, with two-stage steam reheat, as shown, selected most frequently.

Reheated steam is admitted to the three double-flow low-pressure turbine sections through six combined stop-and-intercept intermediate valves. The intermediate valves are normally wide open but provide two lines of overspeed defense in the event of load loss. Exhaust steam from the low-pressure sections passes downward to the condenser, not shown in Fig. 5.

Turbine cycles and performance. **Figures 6** and **7** are representative fossil-fuel and nuclear turbine thermodynamic cycle diagrams, frequently called heat balances. Heat balance calculations establish turbine performance guarantees, provide data for sizing the steam supply and other cycle components, and are the basis for designing the turbine generator.

The fossil-fuel cycle (Fig. 6) assumes a unit rated 500 MW, employing the standard steam conditions of 2400 lb/in.^{2} gage (2415 lb/in.^{2} abs. or 16.65 MPa absolute) and 1000°F (538°C), with resuperheat to 1000°F (538°C). As can be seen in the upper left corner, the inlet conditions correspond to a total heat content, or enthalpy, of 1461 Btu/lb (3.398 MJ/kg) of steam flow. A flow rate of 3,390,000 lb/h (427.1 kg/s) is needed for the desired output of 500 MW (500,000 kW). For efficiency considerations the regenerative feedwater heating cycle is used. Eight heaters in series are employed so that water is returned to the boiler at 475°F (246°C) and 459 Btu/lb (1.068 MJ/kg) enthalpy, rather than at the condenser temperature of 121°F (49°C). Because of the higher feedwater temperature, the boiler adds heat to the cycle at a higher average temperature, more closely approaching the ideal Carnot cycle, in which all heat is added at the highest cycle temperature. The high-pressure turbine section exhausts to the resuperheater at 530 psia (3.65 MPa absolute) pressure and 1306 Btu/lb (3.038 MJ/kg) enthalpy. The reheat flow of 3,031,000 lb/h (381.9 kg/s) returns to the reheat or intermediate turbine section at 490 lb/in.^{2} abs. (3.38 MPa absolute) pressure and 1520 Btu/lb (3.536 MJ/kg) enthalpy. These data are sufficient to calculate the turbine heat rate, or unit heat charged against the turbine cycle. The units are Btu of heat added in the boiler per hour per kilowatt of generator output. Considering both the initial and reheat steam, the heat rate is given by Eq. (1).

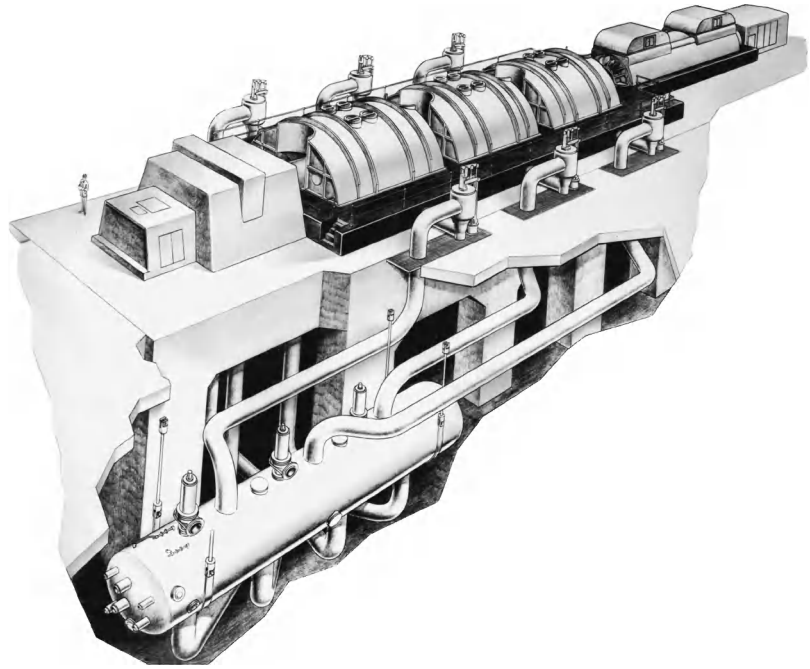

Fig. 5. An 1800-rpm nuclear turbine generator with combined moisture separator and two-stage steam reheater. (*General Electric Co.*)

$$\text{Turbine heat rate} = \frac{3{,}390{,}000\,(1461-459)+3{,}031{,}000\,(1520-1306)}{500{,}000}$$

$$= 8090 \text{ Btu/kWh}$$

$$= 2.37 \text{ kW of heat/kW of output} \qquad (1)$$

The typical power plant net heat rate is poorer

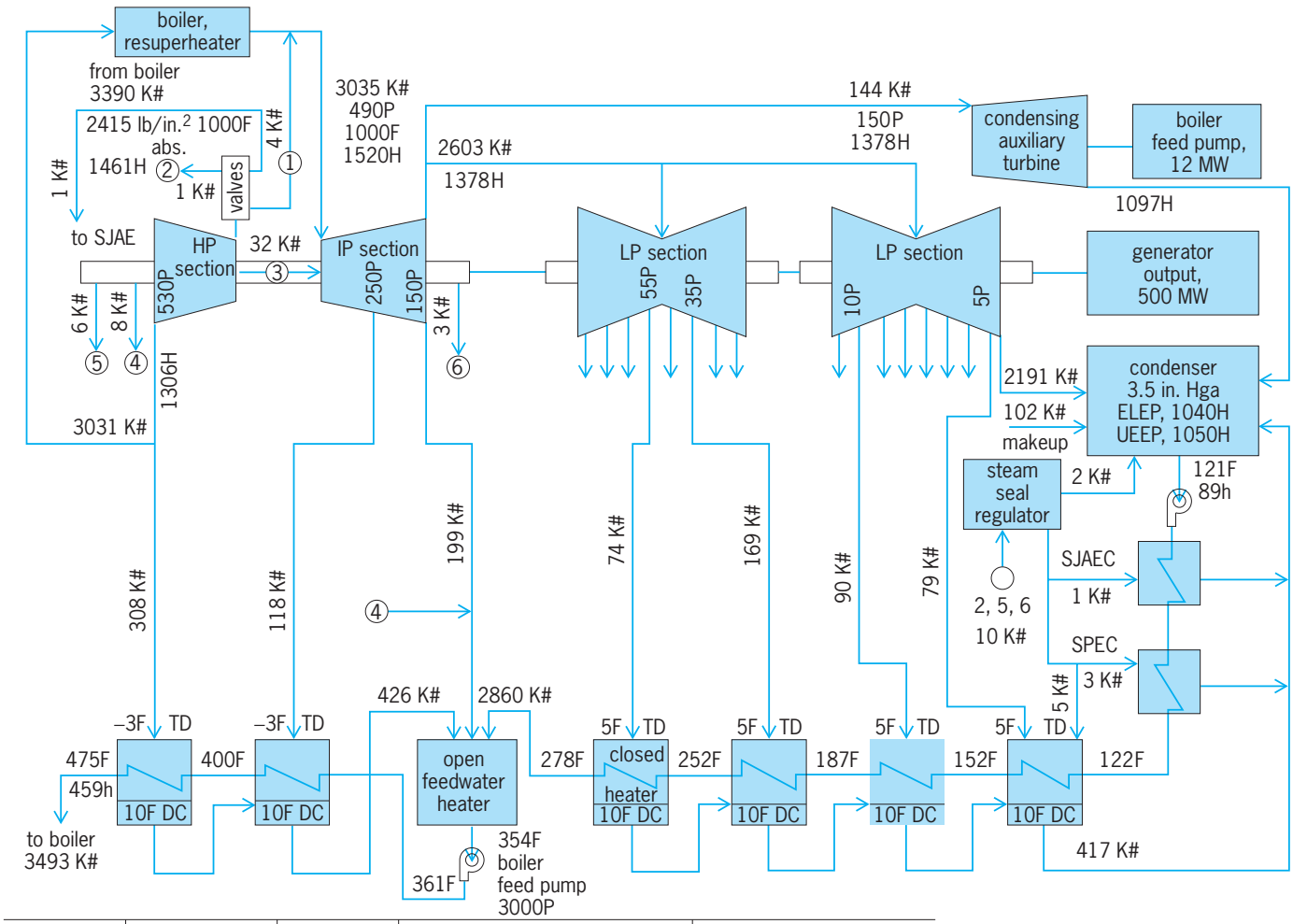

Symbol	Quantity	U.S. Customary Unit	SI equivalent	Abbreviations
#	Mass flow	lb/h	1.2600×10^{-4} kg/s	DC = feedwater heater, drain cooler
K#	Mass flow	10^3 lb/h	0.12600 kg/s	temperature approach
F	Temperature	°F	$^{t}C = (^{t}F - 32)/1.8$	TD = feedwater heater, terminal temperature difference
H	Enthalpy,steam	Btu/lb	1.05506 kJ	
h	Enthalpy, water	Btu/lb	1.05506 kJ	
P	Pressure	lb/in.^{2} abs.	6.8948 kPa	
	Pressure	in. Hga	3.3864 kPa	
	Heat rate	Btu/kWh	2.9307×10^{-4} kJ/(kW · s)	

Fig. 6. Typical fossil-fuel steam turbine cycle.

than the turbine heat rate because of auxiliary power required throughout the plant and because of boiler losses. Assuming 3% auxiliary power (beyond the boiler-feed pump power given in the turbine cycle in Fig. 6) and 90% boiler efficiency, the net plant heat rate is given by Eq. (2).

$$\text{Net plant heat rate} = \frac{3{,}390{,}000\,(1461 - 459) + 3{,}031{,}000\,(1520 - 1306)}{500{,}000\,[(100 - 3)/100][90/100]} = 9270 \text{ Btu/kWh} = 2.717 \text{ kW of heat/kW of output} \quad (2)$$

The heat rates of modern fossil-fuel plants fall in the range of 8600–10,000 Btu/kWh (2.52–2.93 kW of heat/kW of input). Considering that the heat-energy equivalent of 1 kWh is 3412 Btu, the thermal efficiency of the example is given by Eq. (3).

$$\eta_t = (3412/9270)\,100 = 37\% \quad (3)$$

Figure 6 shows 2,191,000 lb/h (276.1 kg/s) of steam exhausting from the main unit to the condenser at an exhaust pressure of 3.5 in. of mercury (11.9 kilopascals) absolute. The theoretical exhaust enthalpy (ELEP), without considerations of velocity energy loss and friction loss between the last turbine stage and the condenser, is 1040 Btu/lb (2.419 MJ/kg). The actual used energy end point (UEEP) is 1050 Btu/lb (2.442 MJ/kg). The exhaust heat at the condenser pressure is thermodynamically unavailable and is rejected as waste heat to the plant's surroundings. The exhaust steam is condensed at a constant 121°F (49°C) and leaves the condenser as water at 89 Btu/lb (0.207 MJ/kg) enthalpy.

On a heat-rate basis, this cycle rejects heat to the

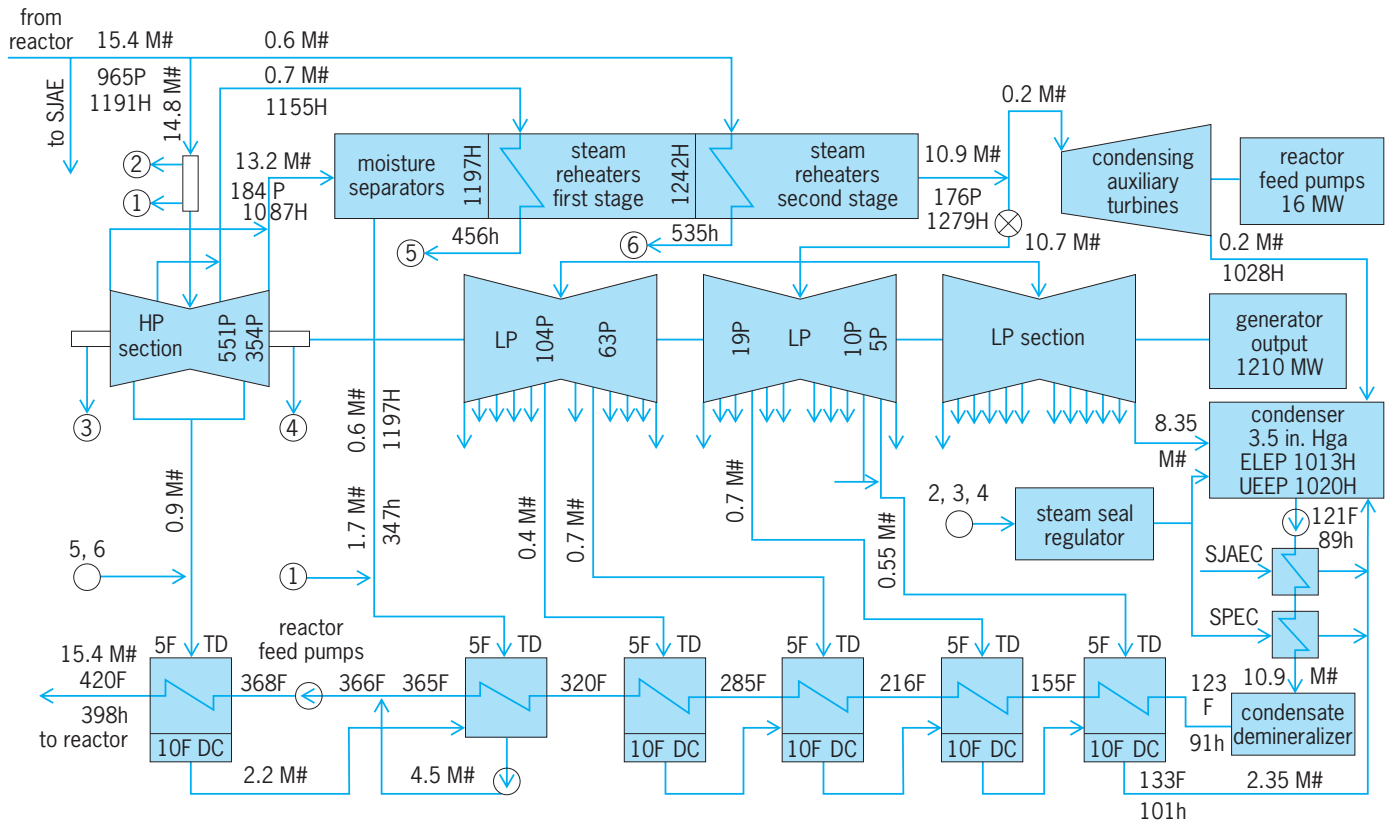

Symbol	Quantity	U.S. Customary Unit	SI equivalent	Abbreviations
#	Mass flow	lb/h	1.2600×10^{-4} kg/s	DC = feedwater heater, drain cooler
M#	Mass flow	10^6 lb/h	0.12600 kg/s	temperature approach
F	Temperature	°F	$^{t}C = (^{t}F - 32)/1.8$	TD = feedwater heater, terminal temperature difference
H	Enthalpy, steam	Btu/lb	1.05506 kJ	
h	Enthalpy, water	Btu/lb	1.05506 kJ	
P	Pressure	lb/in.^{2} abs.	6.8948 kPa	
	Pressure	in. Hga	3.3864 kPa	
	Heat rate	Btu/kWh	2.9307×10^{-4} kJ/(kW · s)	

Fig. 7. Typical nuclear steam turbine cycle.

condenser at the approximate rate which is given by Eq. (4).

$$\text{Net station condenser heat rejection rate}$$
$$= \frac{2{,}191{,}000\,(1050 - 89) + 144{,}000\,(1097 - 89) + 417{,}000\,(100 - 89)}{500{,}000\,(0.97)}$$
$$= 4650 \text{ Btu/kWh}$$
$$= 1.363 \text{ kW of heat/kW of output} \qquad (4)$$

If evaporating cooling towers are used, each pound of water provides about 1040 Btu (each kilogram provides about 2.419 MJ) cooling capacity, which is equivalent to a required minimum cooling-water flow rate of 4.5 lb/kWh (0.57 kg/MJ). The cooling-water needs of a large thermal plant are a most important consideration in plant site selection.

The nuclear cycle (Fig. 7) assumes a unit rated 1210 MW and the steam conditions of the boiling-water reactor. Many similarities can be seen to Fig. 6. The major differences include moisture separation and steam reheating and the lack of need for an intermediate pressure element. The low steam conditions are apparent. The consequent turbine heat rate is given in Eq. (5).

$$\text{Turbine heat rate}$$
$$= \frac{15{,}400{,}000\,(1191 - 398)}{1{,}210{,}000}$$
$$= 10{,}090 \text{ Btu/kWh}$$
$$= 2.957 \text{ kW of heat/kW of output} \qquad (5)$$

A typical nuclear plant also requires about 3% auxiliary power beyond the reactor feed-pump power already included in Fig. 7. The equivalent boiler efficiency approaches 100%, however, and leads to the equivalent net plant heat rate given by Eq. (6).

$$\text{Net plant heat rate}$$
$$= \frac{15{,}400{,}000\,(1191 - 398)}{1{,}210{,}000\,[(100 - 3)/100]}$$
$$= 10{,}400 \text{ Btu/kWh}$$
$$= 3.048 \text{ kW of heat/kW of output} \qquad (6)$$

The corresponding thermal efficiency is given by Eq. (7).

$$\eta_t = (3412/10{,}400)\ 100 = 33\% \qquad (7)$$

Heat is rejected at the condenser at a rate given approximately by Eq. (8).

$$\begin{aligned}&\text{Net station condenser heat rejection rate}\\ &= \frac{[8{,}350{,}000\ (1020-89) + 200{,}000\ (1028-89) + 2{,}350{,}000\ (101-89)]}{1{,}210{,}000\ (0.97)}\\ &= 6810\ \text{Btu/kWh}\\ &= 1.996\text{kW of heat/kW of output} \qquad (8)\end{aligned}$$

Comparison of the heat rates shows that the nuclear cycle requires about 12% more input heat than does the fossil-fuel cycle and rejects about 46% more heat to the condenser, thus requiring a correspondingly larger supply of cooling water. It consumes heat priced at about half that from coal or one quarter that from oil, and is essentially free of rejection to the atmosphere of heat and combustion products from the steam supply.

Turbine classification. Steam turbines are classified (1) by mechanical arrangement, as single-casing, cross-compound (more than one shaft side by side), or tandem-compound (more than one casing with a single shaft); (2) by steam flow direction (axial for most, but radial for a few); (3) by steam cycle, whether condensing, noncondensing, automatic extraction, reheat, fossil fuel, or nuclear; and (4) by number of exhaust flows of a condensing unit, as single, double, triple flow, and so on. Units with as many as eight exhaust flows are in use.

Often a machine will be described by a combination of several of these terms.

The least demanding applications are satisfied by the simple single-stage turbine of Fig. 1. For large power output and for the high inlet pressures and temperatures and low exhaust pressures which are required for high thermal efficiency, a single stage is not adequate. Steam under such conditions has high available energy, and for its efficient utilization the turbine must have many stages in series, where each takes its share of the total energy and contributes its share of the total output. Also, under these conditions the exhaust volume flow becomes large, and it is necessary to have more than one exhaust stage to avoid a high velocity upon leaving and consequent high kinetic energy loss. Figure 5 is an example of a large nuclear turbine generator which has six exhaust stages in parallel.

Machine considerations. Steam turbines are high-speed machines whose rotating parts must be designed for high centrifugal stress. Difficult stress problems are found in long last-stage blading, hot inlet blading, wheels, and rotor bodies.

Casing or shell stresses. The casings or shells at the high-pressure inlet end must be high-strength pressure vessels to contain the internal steam pressure. The design is made more difficult by the need for a casing split at the horizontal center line for assembly. The horizontal flange and bolt design must be leakproof. Shell design problems lead to the use of small-diameter, high-speed turbines at high pressure, and the use of double shell construction (Fig. 4).

Rotor buckets or blades. Turbine buckets must be strong enough to withstand high centrifugal, steam bending, and vibration forces. Buckets must be designed so that their resonant natural frequencies avoid the vibration stimulus frequencies of the steam forces, or are strong enough to withstand the vibrations.

Sealing against leakage. It is necessary to minimize to the greatest possible extent the wasteful leakage of steam along the shaft both at the ends and between stages. The high peripheral velocities between the shaft and stationary members preclude the use of direct-contact seals. Seals in the form of labyrinths with thin, sharp teeth on at least one of the members are utilized. In normal operation these seals do not touch one another, but run at close clearance. In the case of accidental contact, the sharp teeth can rub away without distorting the shaft.

Vibration and alignment. Shaft and bearings should be free of critical speeds in the turbine operating range. The shaft must be stable and remain in balance.

Governing. Turbines usually have two governors, one to control speed and a second, emergency governor to limit possible destructive overspeed. The speed signal is usually mechanical or electrical. A power relay control, usually hydraulic, converts speed signals to steam valve position. Great reliability is required. *See* GOVERNOR.

Lubrication. The turbine shaft runs at high surface speed; consequently its bearings must be continuously supplied with oil. At least two oil pumps, a main pump and a standby driven by a separate power source, are usually provided on all but the smallest machines. A common supply of oil is often shared between the governing hydraulic system and the lubrication system.

Aerodynamic design. The vane design for highest efficiency, especially for the larger sizes of turbines, draws upon modern aerodynamic theory. Classic forms of impulse and reaction buckets merge in the three-dimensional design that is required by modern concepts of loss-free fluid flow. To meet the theoretical steam flow requirements, vane sections change in shape along the bucket. To minimize centrifugal forces on the vanes and their attachments, long turbine buckets are tapered toward their tips. *See* CARNOT CYCLE; HEAT BALANCE; STEAM-GENERATING UNIT; TURBINE.

Frederick G. Baily

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; F. G. Baily, K. C. Cotton, and R. C. Spencer, Predicting the performance of large steam turbine-generators operating with saturated and low superheat steam conditions, *Combustion*, 3(3):8–13, 1967; H. P. Bloch, *A Practical Guide to Steam Turbine Technology*, 1996; R. H. Kehlhofer et al., *Combined-Cycle Gas and Steam Turbine Power Plants*, 2d ed., 1999; J. K. Salisbury, *Steam Turbines and Their Cycles*, 1950, reprint 1974; B. G. A. Strotzki, Steam turbines, *Power*, 106(6):S1–S40, 1962.

Steel

Any of a great number of alloys that contain the element iron as the major component and small amounts of carbon as the major alloying element. These alloys are more properly referred to as carbon steels. Small amounts, generally on the order of a few percent, of other elements such as manganese, silicon, chromium, molybdenum, and nickel may also be present in carbon steels. However, when large amounts of alloying elements are added to iron to achieve special properties, other designations are used to describe the alloys. For example, large amounts of chromium, over 12%, are added to produce the important groups of alloys known as stainless steels.

Phases. There are three unique thermodynamically stable arrangements of iron and carbon atoms in carbon steels. The arrangements are crystalline and referred to as phases. Ferrite is the phase with a body-centered cubic unit cell, austenite is the phase with a face-centered cubic unit cell, and cementite is the phase with an orthorhombic unit cell. The phase or combination of phases that actually exists in a given steel depends on temperature and composition. A phase diagram is a type of map that shows where the different phases exist under equilibrium conditions. Equilibrium assumes that a steel has been held long enough at a temperature to ensure that the most stable arrangement of atoms and phases has developed (**Fig. 1**).

The phase diagram for the Fe-C system shows that austenite is stable only at high temperatures and that it is replaced at low temperatures by other phases. This characteristic of iron-carbon alloys is the fundamental basis for the processing and great versatility of steels. When a steel is heated in the austenite phase field, it becomes quite ductile, and heavy sections can be hot-rolled or forged to smaller sizes and more complex shapes. On cooling after hot work, the austenite transforms to structures of higher strength. Another important use of the austenite phase field is the heat treatment of steels. By heating a finished steel part into the austenite phase, that is, by heating above the critical temperatures identified by A_3 and A_{cm} on the phase diagram (Fig. 1), the microstructure existing at room temperature is replaced by austenite. The resulting austenite may then be cooled at various rates to form new arrangements of the phases with a variety of mechanical properties. Cooling at very high rates or quenching produces the hardest structure in steels, and is discussed in the sections on medium- and high-carbon steels. *See* ALLOY STRUCTURES.

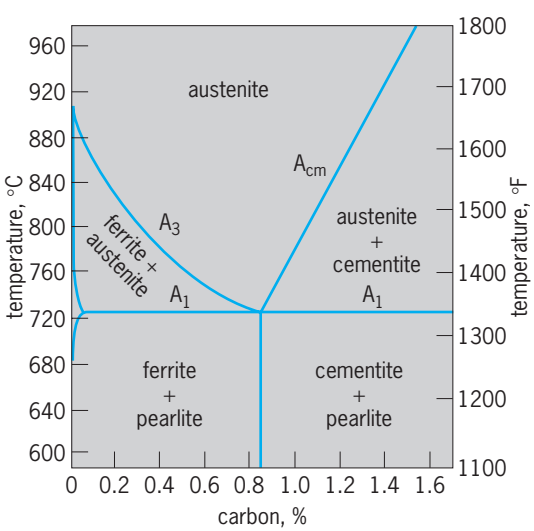

Fig. 1. Critical temperatures in plain carbon steels and constituents which are present in iron-carbon alloys upon slow cooling.

Microstructure and properties. When steels are cooled slowly from the austenite phase field, the austenite transforms to microstructures of ferrite and cementite. In steels with carbon content below 0.8%, the austenite first transforms to ferrite, as indicated by the ferrite-plus-austenite field of the phase diagram for the Fe-C system (Fig. 1). Steels with carbon content above 0.8% first form cementite. At or below the A_1 temperature, any austenite not transformed to ferrite or cementite transforms to a mixture of ferrite and cementite referred to as pearlite. The ferrite and cementite form as parallel plates, or lamellae, in pearlite. Thus, when a steel reaches room temperature, its microstructure consists of either ferrite and pearlite or cementite and pearlite, depending on its carbon content (Fig. 1).

Typical microstructures exist for steels cooled very slowly in a furnace, a process referred to as full annealing, or cooled in air, a process referred to as normalizing. The 0.06% carbon steel consists almost entirely of ferrite in the form of fine crystals or grains. The higher-carbon steels contain increasingly greater amounts of pearlite which, at the low magnifications (**Fig. 2**), appears dark because the lamellae are too close together to be resolvable.

The mechanical properties of annealed, normalized, or hot-rolled steels are directly dependent on the amounts of pearlite present in the microstructure. Cementite is a hard carbide phase and, when present in the pearlite, significantly strengthens a steel. The yield and tensile strengths increase with increasing carbon content or increasing pearlite content of steels containing up to 0.8% carbon. However, as strength increases, ductility falls and the steels become less formable, as shown by the decrease in reduction of area and elongation (**Fig. 3**).

The change in mechanical properties with carbon content separates steels into low-, medium-, and high-carbon classifications that are related to the need for formability in manufacturing and strength and toughness for good performance in service.

Low-carbon steels. These steels, sometimes referred to as mild steels, usually contain less than 0.25% carbon, and therefore have microstructures that consist mostly of ferrite and small amounts of pearlite. These steels are easily hot-worked in the austenite condition and are produced in large tonnages for beams and other structural applications. The relatively low strength and high ductility (Fig. 3)

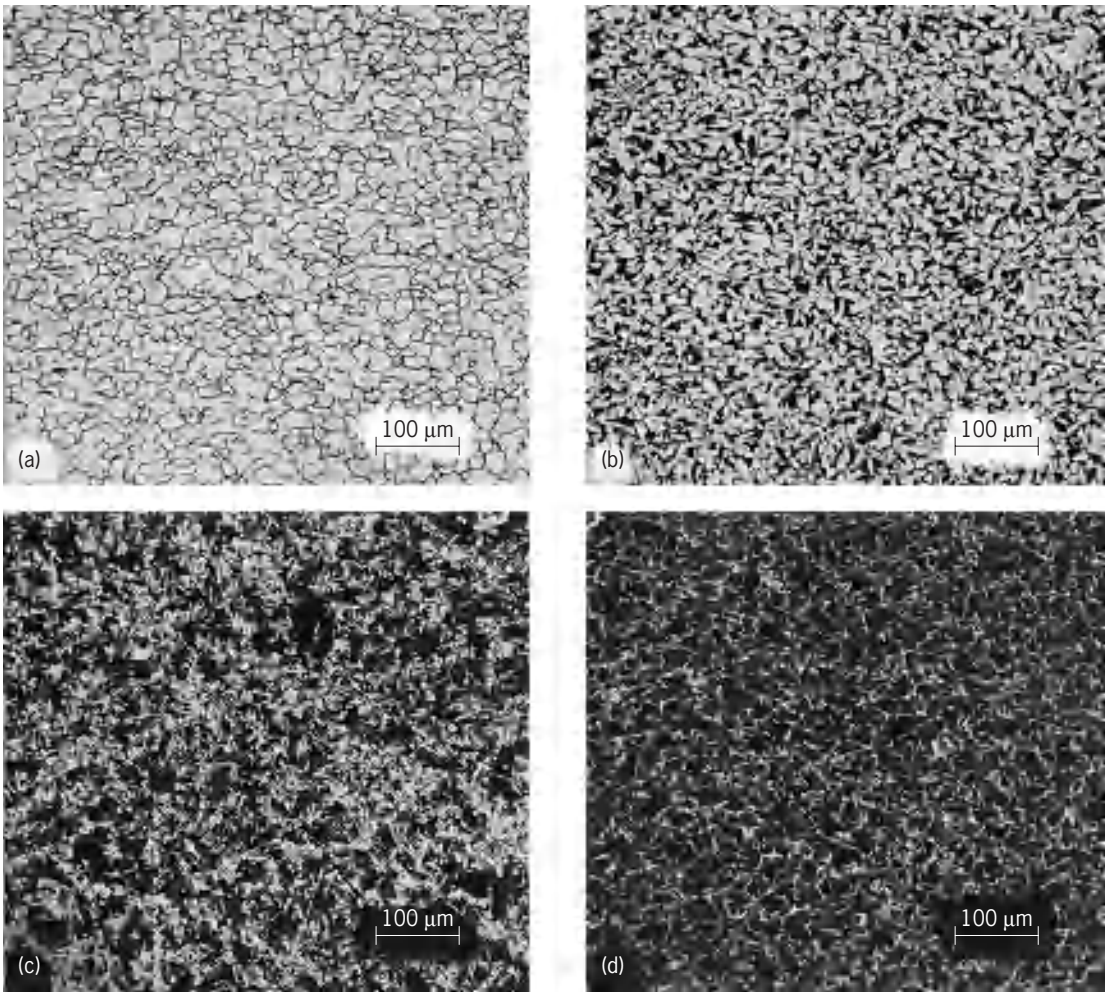

Fig. 2. Micrographs illustrating appearance of the constituents shown in Fig. 1. (*a*) 0.06% carbon steel: white is ferrite, grain boundaries appear as a dark network. (*b*) 0.18% carbon steel: white is ferrite, dark is pearlite. (*c*) 0.47% carbon steel: same as *b*, but larger amount of pearlite in this higher-carbon steel. (*d*) 1.2% carbon steel: dark is pearlite, white grain boundaries are cementite.

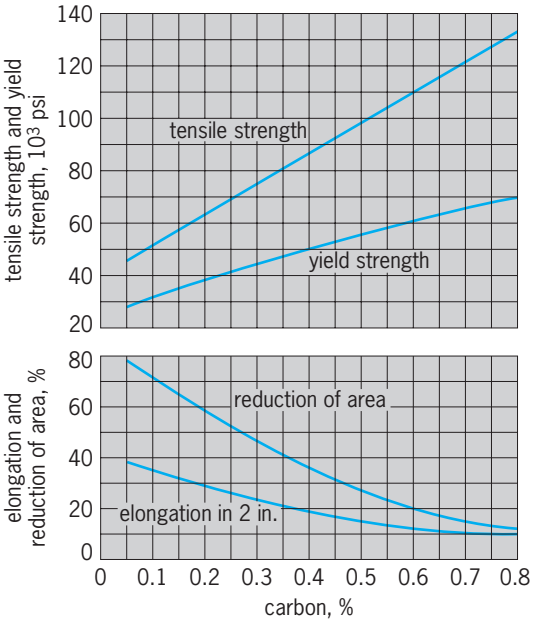

Fig. 3. Effect of carbon content on the tensile properties of hot-worked carbon steels. 10^3 psi = 6.9 MPa; 2 in. = 5 cm.

of the low-carbon steels make it possible also to cold-work these steels in the ferritic condition. Cold-rolled low-carbon steels are extensively used for sheet applications in the appliance and automotive industries. Cold-rolled steels have excellent surface finishes, and both hot- and cold-worked mild steels are readily welded.

The low-carbon ferritic steels have undergone interesting developments in order to meet the challenges of the energy crisis. Although mild steels have excellent formability, their low strength requires relatively thick sections for a given load-bearing application. New compositions of ferritic steels combined with new processing have significantly raised the strength of low-carbon steels, and therefore made possible significant reductions in vehicle weight by permitting the use of thinner sections.

One approach to improved low-carbon ferritic steels has led to the development of a class of steels referred to as high-strength low-alloy (HSLA). HSLA steels are alloyed with very small amounts of strong carbide-forming elements, such as vanadium and niobium. Very fine carbides are formed when the microalloyed steels are finish-hot-worked low in the

austenite phase field. The alloy carbides limit the growth of the austenite grains, and therefore the size of the ferrite grains that form from the austenite. The resulting very fine ferrite grain size, together with the fine-alloy carbides, significantly increases strength. Special attention is also paid to inclusion content and shape in HSLA steels, and as a result of inclusion control and the very fine ferritic grain sizes, HSLA steels have excellent toughness for critical applications.

Another approach to producing high strength and good formability in low-carbon steel is to heat a steel with a ferrite-pearlite microstructure into the ferrite-plus-austenite phase field (Fig. 1). Such intercritically annealed steels are referred to as dual-phase steels. The intercritical annealing treatment converts the pearlite areas into pools of austenite and, by controlling cooling rate or alloying the austenite, or both, transforms to structures other than pearlite on cooling. Good formability and high strength are achieved by the elimination of the pearlite and by the introduction of dislocations into the ferrite matrix surrounding the tranformed austenite. The dislocations are crystal defects that make possible the shaping of metals.

Medium-carbon steels. The medium-carbon steels contain between 0.25 and 0.70% carbon, and are most frequently used in the heat-treated condition for machine components that require high strength and good fatigue resistance (resistance to cyclic stressing). The heat treating consists of austenitizing, cooling at a high enough rate to produce a new phase, martensite, and then tempering at a temperature below the A_1. The martensite is essentially a supersaturated ferrite with carbon atoms trapped between the iron atoms. This structure is very hard and strong. The hardness and strength of martensite increase rapidly with the carbon content of the steel, and the greatest benefits of forming martensite occur above 0.3% carbon content (**Fig. 4**).

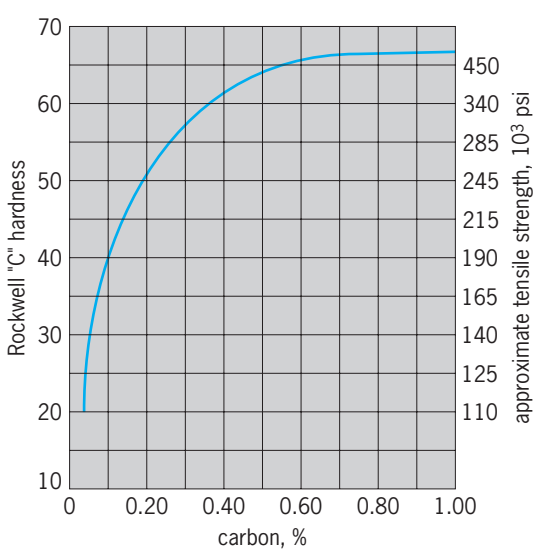

Fig. 4. Relation of hardness and tensile strength to the carbon content in steels which are rapidly quenched from above critical temperature. 10^3 psi = 6.9 MPa.

However, the same structural factors that make martensite very hard also make it very brittle. Therefore, martensitic steels are tempered to increase their toughness. The tempering essentially makes it possible for the carbon atoms to diffuse from the martensitic structure to form independent carbide particles. Thus the microstructure changes to a mixture of ferrite and carbides on tempering. The extent of this process is controlled primarily by temperature, and a large range of hardness and toughness combinations can be produced.

Martensite can be formed only when austenite transformation to ferrite or pearlite is suppressed. The ferrite and pearlite phases require the diffusion of carbon, a time-dependent process, for their formation. In plain-carbon steels the diffusion can be suppressed only by quenching or cooling very quickly in brine or water. Even when martensite is formed on the surface of a bar of plain-carbon steel, however, the center of the bar may cool too slowly to form martensite.

Small amounts of alloying elements such as chromium, nickel, and molybdenum reduce the rate at which ferrite or pearlite form, and therefore make it possible to form martensite in heavier sections or at slower cooling rates. Slower cooling reduces the tendency to distortion or cracking that sometimes accompanies the high residual stresses introduced by severe quenching. Hardenability is the term used to define the ease of martensite formation relative to cooling rates, section sizes, and steel composition.

High-carbon steels. Steels containing more than 0.7% carbon are in a special category because of their high hardness and low toughness. This combination of properties makes the high-carbon steels ideal for bearing applications where wear resistance is important and the compressive loading minimizes brittle fracture that might develop on tensile loading.

Two types of microstructure are produced in high-carbon steels. One consists entirely of a pearlite with a very fine lamellar spacing produced in steels with about 0.80% carbon. Figure 1 shows that no ferrite or cementite forms above A_1 at this carbon content. An important application of pearlitic 0.8% carbon steels is in railroad rail. Another application in which pearlitic microstructure is important is high-strength wire for cables and wire rope. In a process referred to as patenting, rods of 0.8% carbon steel are transformed to very fine pearlite. The rods are then drawn to wire to produce an aligned pearlite with strengths up to 350 ksi (2415 MPa).

The other type of microstructure produced in high-carbon steels is tempered martensite. The most common bearing steel contains 1% carbon and about 1.5% chromium. This steel is oil-quenched to martensite and tempered to retain as high a hardness as possible. Austenitizing is performed in the austenite-cementite phase field. As a result, the microstructure contains very fine cementite particles which not only contribute to wear resistance but also help to maintain a fine austenite grain size. *See* CAST IRON;

FERROALLOY; HEAT TREATMENT (METALLURGY); IRON; IRON ALLOYS; METAL, MECHANICAL PROPERTIES OF; STAINLESS STEEL; STEEL MANUFACTURE.

George Krauss

Mechanical properties. The mechanical properties of steel are determined from a uniaxial tension test in which a cylindrical bar of measured diameter is subjected to tension along its length while the magnitude of the force applied and the resulting elongation are measured. The force, divided by the original cross-sectional area of the bar, is referred to as the stress, and the elongation of a specific portion of the bar, divided by the original length of that portion, is the strain. The strain is clearly dimensionless, but the stress has the units of force per unit area. Since the forces involved are large (in the American system, forces are expressed in kips, where 1 kip is 1000 lb), stresses are expressed in kips/in.^{2} (ksi). The SI unit of force is the newton, and 1 N/m^2 is the pascal (Pa). For the large stresses in steel, the usual unit is the megapascal, (MPa); 1 ksi = 6.895 MPa. A diagram plotted with stress as the ordinate and strain as the abscissa is referred to as a stress-strain diagram. For all steels with carbon generally making up less than 2% (structural steels), the initial portion of such a diagram is linear. The slope of this straight-line portion, the ratio of stress to strain, is the modulus of elasticity; it is about 29,000 ksi (200,000 MPa). At this time, the test specimen is also elastic, recovering its original dimensions if the load is removed, with unloading following down the same straight line plotted during loading. *See* ELASTICITY; STRESS AND STRAIN; STRUCTURAL STEEL.

In the low-carbon and medium-carbon steels, linearity ceases abruptly in the yield zone, with a large increase in strain with little or no increase in load. The magnitude of the strain occurring during yield is 10–20 times as large as the strain at the end of the elastic region.

At the conclusion of the yield, strain-hardening occurs, and additional load is once again required to produce additional deformation, but at a much lower ratio of stress to strain. Eventually, the load reaches a maximum value, the ultimate tensile strength, where the specimen begins to neck down, developing a severely reduced cross section at some point, and quickly elongates a great deal and fractures. The total elongation is an order of magnitude larger than that during yield, two orders of magnitude above the elastic strain.

In the higher-carbon steels, there is generally no abrupt yield. The linear zone makes a transition to a curve going up to the ultimate strength and down to fracture. It is still important to mark the end of elastic behavior, and therefore a yield is defined for these steels, either in terms of a measured amount of strain (typically 0.005) or as an amount of inelastic, nonlinear strain, found as the point where the stress-strain curve intersects a straight line having the same slope as the initial portion but with an offset (typically 0.002).

The behavior of steel in compression is similar, provided that buckling of the specimen is prevented. Of course, there is no neck-down or fracture, the specimen simply being squeezed down and shortened, and flowing out.

Commercially available steels have yield points ranging 33–100 ksi (230–700 MPa) and ultimate tensile strengths up to 130 ksi (900 MPa), although building codes and specifications generally restrict operating stress levels to values below the yield.

G. Donald Brandt

Bibliography. American Society for Metals, *Metals Handbook, Properties and Selection: Irons, Steels, and High Performance Alloys*, 10th ed., vol. 1., 1990; Association of Iron and Steel Engineers, *The Making, Shaping, and Treating of Steel*, 11th ed., 1998; H. K. D. H. Bhadeshia and R. W. K. Honeycombe, *Steels: Microstructure and Properties*, 3d ed., 2006; D. Llewellyn and R. Hudd, *Steels: Metallurgy and Applications*, 3d ed., 1998.

Steel manufacture

A sequence of operations in which pig iron and scrap steel are processed to remove impurities and are separated into the refined metal and slag.

Ironmaking

Ironmaking, one of the earliest of all metallurgical operations, is the initial step in most steel-producing operations. The blast furnace process for ironmaking came into being with the stone-built furnaces of the Middle Ages. Since then, ironmaking furnaces have evolved into towering structures that dominate the integrated steel mill.

Blast furnace. The iron blast furnace process consists of charging iron ore, coke, and limestone flux into the top of a shaft furnace and blowing heated air or blast into the bottom. This process yields a high-carbon iron containing silicon and manganese which is converted to steel in steelmaking furnaces.

The reactions occurring in iron- and steelmaking primarily involve adjustment of the oxygen potential—first, to decrease the oxygen potential and reduce the iron ore to an impure pig iron, and then to increase the oxygen potential and oxidize impurities from that iron in the steelmaking operation. In the iron blast furnace, the oxygen potential is extremely low in the hearth zone and continually increases as the reducing gases move up the furnace, reacting with iron oxide and coke.

Direct reduced iron. Another source of iron units for production of steel, in addition to pig iron and scrap, is direct-reduced iron. Metallic iron is produced from iron ores by removing most of the oxygen at temperatures below the melting points of materials in the process. Direct-reduced iron can be refined to steel or used to increase productivity of the blast furnace or other melting processes. Several processes have been developed to contact reducing agents with iron oxides, including fluidized beds, moving-bed shaft furnaces, fixed-bed retorts, and rotary-kiln processes.

Direct-reduction processes represent a small

fraction of iron production, but are being adopted where suitable ores and reducing agents, particularly natural gas, are available. Direct-reduced iron is very suitable as a charge material in steelmaking processes, and has special advantages in electric furnace steelmaking. *See* FLUIDIZATION; IRON METALLURGY; PRESSURIZED BLAST FURNACE.

Steelmaking Processes

Reduction of iron ores by carbonaceous fuel directly to a steel composition was practiced in ancient times, but liquid processing was unknown until development of the crucible process, in which iron ore, coal, and flux materials were melted in a crucible to produce small quantities of liquid steel. Modern steelmaking processes began with the invention of the airblown converter by H. Bessemer in 1856. The Thomas process was developed in 1878; it modified the Bessemer process to permit treatment of high-phosphorus pig iron. The Siemens-Martin process, also known as the open-hearth process, was developed at about the same time. The open-hearth process utilizes regenerative heat transfer to preheat air used with a burner; it can generate sufficient heat to refine solid steel scrap and pig iron in a reverberatory furnace. After World War II, various oxygen steelmaking processes were developed.

Steelmaking can be divided into acid and basic processes depending upon whether the slag is high in silica (acid) or high in lime (basic). The furnace lining in contact with the slag should be a compatible material. A silica or siliceous material is used in acid processes, and a basic material such as burned dolomite or magnesite is used in basic processes. Carbon, manganese, and silicon, the principal impurities in pig iron, are easily oxidized and separated; the manganese and silicon oxides go into the slag, and the carbon is removed as carbon monoxide and carbon dioxide in the off-gases. Phosphorus is also oxidized but does not separate from the metal unless the slag is basic. Removal of sulfur occurs to some extent by absorption in a basic slag. Thus, the basic steelmaking processes are more versatile in terms of the raw materials they can handle, and have become the predominant steelmaking processes.

A typical pig iron charged to the steelmaking process might contain roughly 4% carbon, 1% manganese, and 1% silicon. The phosphorus and sulfur levels in the pig iron vary. The composition of the steel tapped from the steelmaking furnace generally ranges from 0.04 to 0.80% carbon, 0.06 to 0.30% manganese, 0.01 to 0.05% phosphorus, and 0.01 to 0.05% sulfur, with negligible amounts of silicon. The chemistry of steelmaking can be represented as reactions (1)–(5), which are treated in terms of oxidation by

$$C + (FeO) \rightarrow CO + Fe(l) \quad (1)$$

$$Si + 2(FeO) \rightarrow (SiO_2) + 2Fe(l) \quad (2)$$

$$Mn + (FeO) \rightarrow (MnO) + Fe(l) \quad (3)$$

$$2P + 4(CaO) + 5(FeO) \rightarrow 4CaO \cdot P_2O_5 + 5Fe(l) \quad (4)$$

$$S + CaO \rightarrow CaS + O \quad (5)$$

liquid iron oxide in the slag. Several alternative mechanisms are considered possible, such as direct oxidation by gaseous oxygen, reaction (6), or alternatively,

$$C + {}^{1}/_{2}O_2(g) = CO(g) \quad (6)$$

the oxygen may dissolve in the liquid iron and react according to reaction (7). The removal of phosphorus from the bath requires a basic oxidizing slag as illustrated by reaction (4). The actual form of the phosphorus in the slag is thought to be as tricalcium or tetracalcium phosphate, although evidence is not entirely clear in this regard. Sulfur is absorbed by a basic slag, although it may or may not form calcium sulfide as indicated by reaction (5).

$$C + O = CO \quad (7)$$

Bessemer converter processes. The original Bessemer converter employed an acid lining. This acid process is no longer used extensively. The process requires a very low phosphorus- and sulfur-bearing pig iron since these elements are not removed in the process. A cross section of a Bessemer converter is shown in **Fig. 1**. Oxidation of the charge is accomplished by blowing air through a wind box on the bottom of the vessel. Elements are eliminated essentially in the same order as the stability of the oxide, that is, silicon is eliminated first, along with manganese, and finally carbon. The Bessemer process is autogenous—that is, no additional heat is supplied to the process—and the oxidation of the metalloids; particularly silicon, is sufficient to keep the charge molten and raise the temperature as purification proceeds so that a low-carbon steel at a temperature near 2900°F (1600°C) is produced at the end of the process. A small amount of cold scrap can be charged.

The Thomas or basic Bessemer process, used more extensively in Europe, is carried out in a basic lined vessel, and utilizes lime or limestone as a charge material to form a basic slag. Pig iron charged to this process contains up to 1 wt % phosphorus, which is oxidized from the bath after most of the carbon has been removed in a so-called afterblow. The

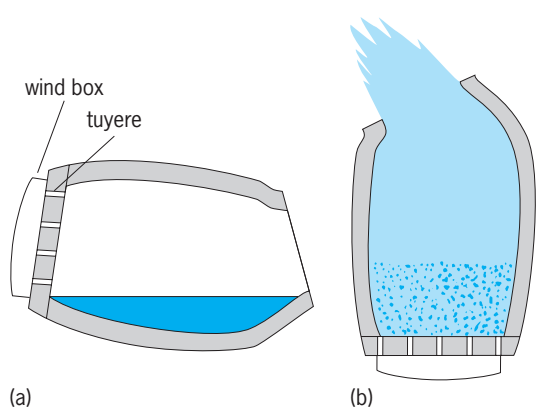

Fig. 1. Operating positions of a Bessemer converter. (*a*) Charging and pouring. (*b*) Blowing. (*After R. D. Pehlke, Unit Processes of Extractive Metallurgy, p. 89, American Elsevier, 1973*)

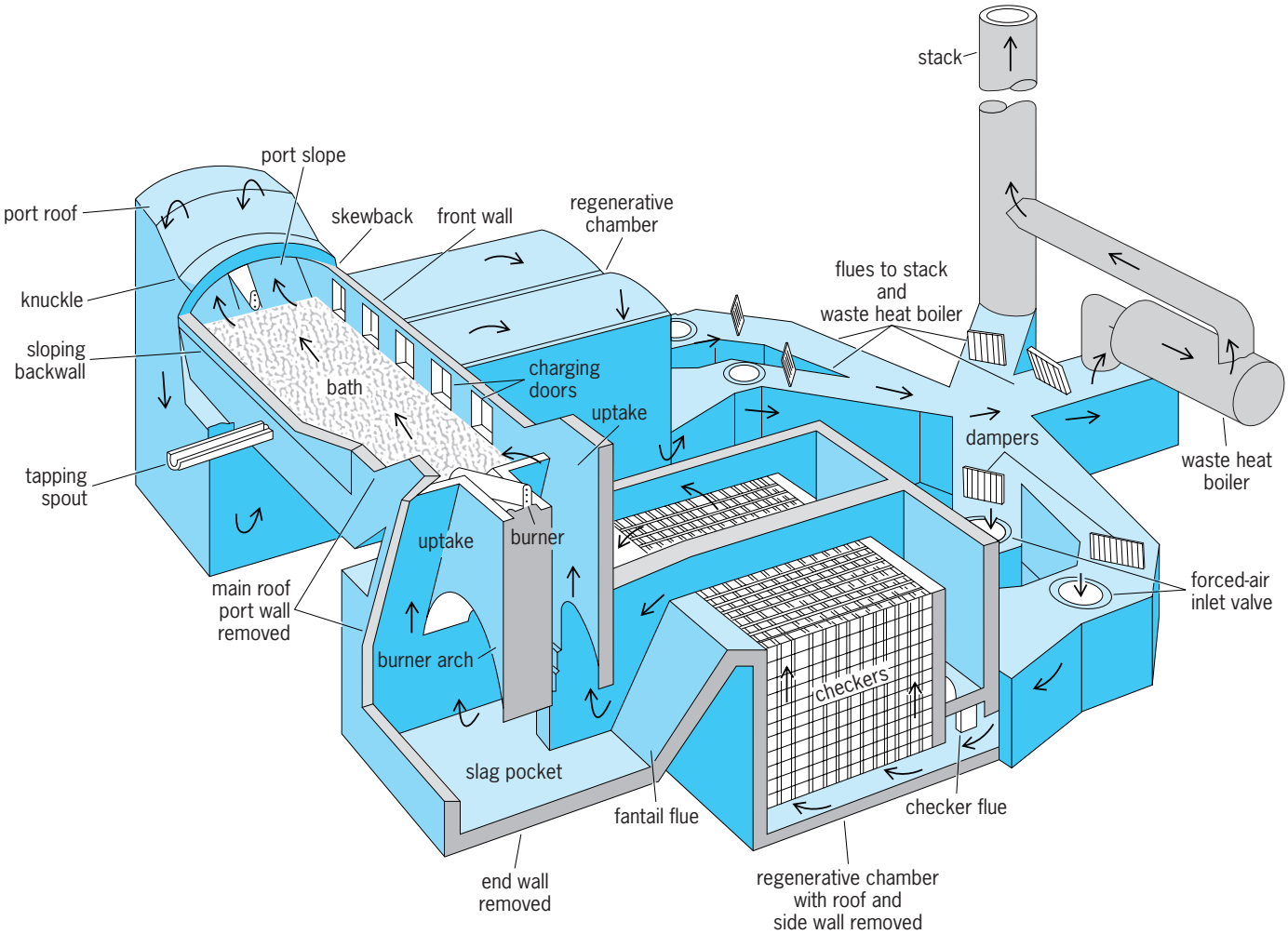

Fig. 2. Diagram of the principal parts of an open-hearth furnace. The arrows indicate the direction of the flow of preheated air, flame, and waste gases when a liquid fuel is fired through a burner in the trench at the right end of the furnace. The five charging doors shown are in the front wall of the furnace, and the checker chambers extend under the charging floor, which is not shown. When reversal of firing takes place, the function of the uptake is reversed, producing the downtake. (*After H. E. McGannon, ed., The Making, Shaping and Treating of Steel, United States Steel Corporation, 8th ed., p. 471, 1964*)

oxidation of the phosphorus is exothermic and adds considerable thermal energy to the process.

One of the disadvantages of Bessemer steelmaking is the resulting higher phosphorus and nitrogen contents in the steel. The nitrogen content of a Bessemer steel is approximately 0.015 wt %, although lower nitrogen contents can be achieved by careful control of temperature and by carrying out the final oxidation of the bath with iron ore additions.

An important aspect of converter steelmaking is control, because of the rapidity with which the end point of the process is approached, and the necessity for providing a careful thermal balance for the autogenous process. End-point control has been greatly enhanced by the use of photocells to observe the character of the flame and emission from the vessel toward the end of the blow.

Open-hearth process. From the turn of the century through the late 1950s, the basic open-hearth process was the workhorse of the American industry, accounting for approximately 85% of the steel produced in the United States in 1955. Because of its melting capabilities, the basic open-hearth process can handle a wide range of pig iron and scrap ratios, the optimum operation being a metallic charge of 40 to 60% molten pig iron with the balance as steel scrap.

A typical open-hearth furnace is shown in **Fig. 2**. In an effort to obtain a high flame temperature and good thermal efficiency, the open hearth employs a regenerative heat transfer system. By reversing the direction of firing in the furnace, hot combustion gases pass out of the furnace through a set of brick checkers which recover heat from the expanded combustion gases. The incoming air is passed through the hot checker work on the opposite end of the furnace before being reacted with fuel at the burner. This technique for preheating the combustion air provides fuel economy.

In a basic operation, a wide range of raw materials can be treated, and the phosphorus and sulfur concentrations are kept under suitable control. An advantage of the open hearth over the Bessemer process is that the steels produced contain about

0.005 wt % nitrogen, which is low enough to provide good properties in deep-drawing sheet material.

The solid constituents—steel scrap, limestone, and iron ore—are charged to the furnace, and the burner turned on. Burners are built into both ends of the furnace, and the direction of the flame is changed from time to time to take advantage of the regenerative heat transfer system for the air which is located on either side of the furnace at a lower level. As soon as the charge starts to melt down, pig iron is added and the iron oxide reacts with the molten pig iron to produce carbon monoxide in a so-called ore boil. As the temperature increases, the decomposition of limestone at the bottom of the furnace occurs and a lime boil starts, with carbon dioxide being evolved at a relatively rapid rate. As the calcined lime floats up through the charge, phosphorus is removed as well as some sulfur. The productivity of the open-hearth furnace had been limited by both charging time and the extensive melting and refining periods, requiring approximately 6 to 8 h for melting and refining alone. The availability of low-cost oxygen has permitted oxygen lances to be inserted through the roof of open-hearth furnaces, which decreases the time required to process a heat (single batch of steel) from about 12–14 h to approximately 6 h. Oxygen blowing, however, creates fumes and requires gas-cleaning equipment. This limitation on productivity and the added expense of gas-cleaning equipment in an older installation have led to rapid displacement of the open-hearth furnace by the oxygen top-blowing process.

Oxygen steelmaking. The high dissolved nitrogen content in Bessemer steel limits its application, and a nitrogen-free gas for blowing in the converter is needed. The use of pure oxygen in a bottom-blown vessel will result in extremely severe erosion of the bottom lining and tuyere area. An alternative is to blow a Bessemer converter with oxygen and steam to produce low-nitrogen steel. This process has received considerable attention in Europe.

In the early 1950s, commercial development of the oxygen steelmaking process was achieved at Linz, Austria. Since Austrian ores are low in phosphorus, they cannot support the basic Bessemer process thermally; furthermore, a shortage of scrap in Austria following World War II eliminated economic application of the open hearth. These factors along with availability of oxygen at much lower cost in tonnage quantities led to development of the top-blown oxygen process, which is based on experience with a small vessel at Gerlafingen, Switzerland. The first large-scale oxygen steelmaking converters were 30-ton vessels. Further development of the process, particularly in the United States, has increased the size of the vessel until most oxygen steelmaking shops employ 250-ton vessels, with some vessels as large as 400 tons in operation.

A schematic cross section of an oxygen steelmaking vessel is shown in **Fig. 3**. **Figure 4** shows the oxygen jetting of a 300-ton heat of steel. The process is autogenous and uses about 25 to 35% scrap, the remainder being basic pig iron which is relatively low in sulfur. Burned lime is used as a flux, with enough added to produce a lime-silica ratio of around 3 to 1 in the final slag. Oxygen is blown down vertically into the bath through a Laval nozzle, which produces supersonic flow. With larger vessels, the use of a three-hole lance has improved operation.

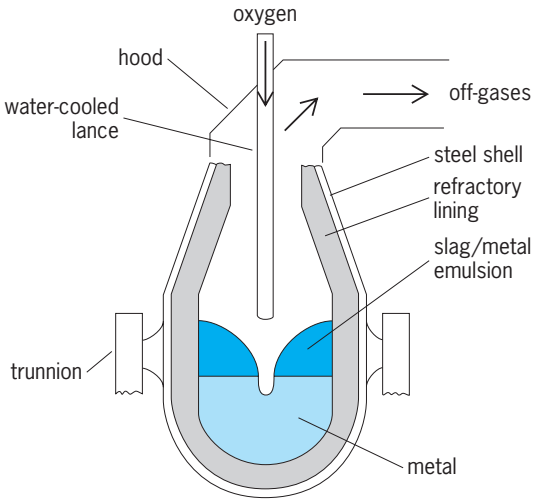

Fig. 3. Schematic section of oxygen steelmaking vessel.

Copious fumes in the form of iron oxide particles around a micrometer in size are evolved during the oxygen blow. This fume loss amounts to 2 to 3% of the metallic charge, and gas-cleaning equipment represents a large portion of the capital expenditure for a basic oxygen furnace plant. Significant operating variables include the lance height and oxygen flow

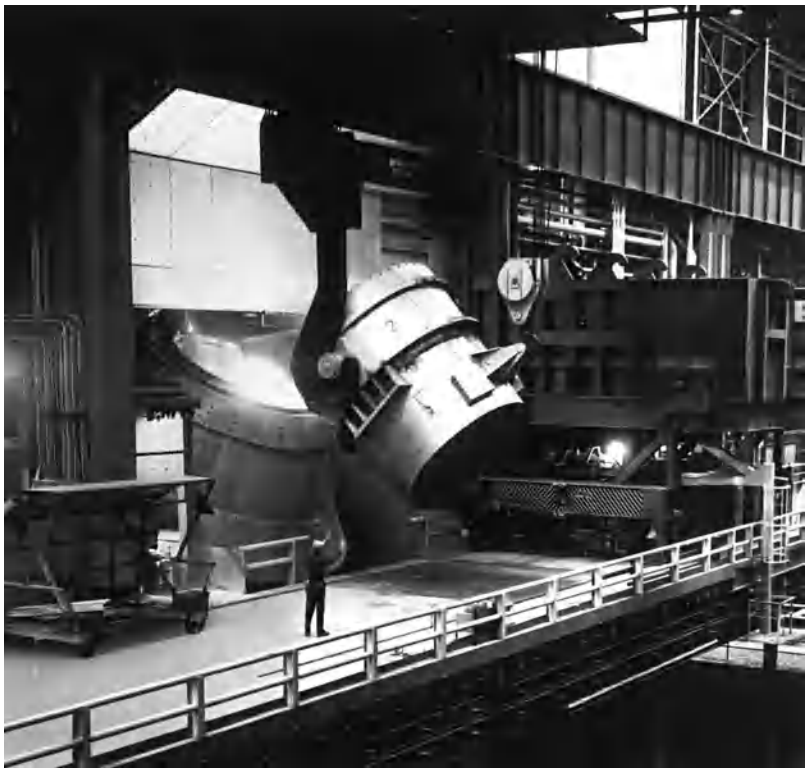

Fig. 4. Large oxygen steelmaking furnace in operation. (*Bethlehem Steel Corporation*)

rate. The required blowing time for a heat, including those approaching 250 tons, is approximately 20 min. By suitable adjustment of the lance height, some iron can be oxidized early in the blow to dissolve the lime and produce an early basic oxidizing slag, which assists in phosphorus removal. Later in the heat, lowering the lance or increasing the oxygen flow rate provides the deep penetration necessary to agitate metal and slag and provide for relatively low iron losses in the slag.

The processing of high phosphorus irons in the top-blown oxygen steelmaking process can be accomplished by injecting powdered lime into the oxygen stream. This technique, known as the LD-AC process, is used in Europe to process pig iron containing more than 0.5 wt % phosphorus.

Oxygen converter gas recovery. Many oxygen steelmaking vessels have open hoods into which air is ingested to burn the hot evolved converter gases, which contain about 90% carbon monoxide during much of the oxygen-blowing period. Considerable effort has been made, particularly in Japan, to recover the converter gas, which averages over 80% carbon monoxide, for use as a fuel in the steel plant. The use of a closed hood and special precautions to prevent explosion of the carbon monoxide has resulted in an economical process for recovering the thermal energy in the carbon monoxide. In most hoods, a waste heat boiler is employed to recover the sensible heat of the stack gas; but a marked improvement in thermal efficiency is achieved by collecting the carbon monoxide gas for use as a supplementary fuel in the steel plant. Furthermore, the gas-cleaning problems are decreased because the volume of gas is only approximately one-sixth as great as when air is ingested in an open-hood system.

Submerged injection processes. Submerged injection processes were developed first for several European plants, in which bottom-blown Thomas (basic Bessemer) vessels were converted to use oxygen and hydrogen-bearing gases such as hydrocarbons, steam, or natural gas. In France, G. Savard and R. Lee in 1966 patented a dual tuyere with which oxygen could be injected through an inner tuyere and a gaseous hydrocarbon injected through the surrounding annular space formed by the central pipe and the surrounding outer pipe. This arrangement, utilizing the effect of cracking the hydrocarbon to cool the surrounding refractory, permitted bottom blowing of oxygen in steelmaking converters.

Several conversions of Thomas converter plants to the bottom-blown oxygen process were made by using natural gas in the outer surrounding tuyere. These steelmaking vessels varied in size up to 50 tons capacity; in 1973 the scale-up to 200-ton vessels was accomplished by the United States Steel Corporation.

The bottom-blown oxygen steelmaking process utilizes a converter with a replaceable bottom containing dual tuyeres with oxygen introduced through the central outlet. The number of tuyeres increases with the size of the furnace. Tuyere placement is critical in order to take advantage of the stirring capability of bottom injection, particularly under conditions where oxygen flow rates are decreased. Replacement of the bottom refractory plug for the vessel, which contains the tuyeres, is required only as often as the replacement of the remaining vessel refractory lining, typically after more than 1000 heats.

The absence of a lance in the bottom-blowing process has offered opportunities to decrease total plant height and has allowed relatively easy conversion of older open-hearth buildings to oxygen steelmaking shops.

Simultaneous top and bottom blowing has been explored to combine the advantages of each type of oxygen steelmaking. The joint effects of simultaneous top and bottom blowing of oxygen and mixed gases have resulted in higher metallic yield, improved manganese recovery, better control of oxygen content, and lower oxygen concentrations, as well as improved dephosphorization.

Electric furnace process. Electric arc furnace technology began late in the nineteenth century with the original design of P. L. T. Heroult. The three-graphite electrode furnace with a swinging roof for top charging and a rocker base for tilting to tap the finished molten steel has been continuously improved and developed further. *See* ELECTRIC FURNACE.

In its early form, the electric arc furnace had a small capacity and was used only for high-alloy and stainless steels, and for 5- or 10-ton heats of various steels cast into sand molds to produce shaped castings in foundries. Modern direct-arc furnaces ranging in size up to 400 tons are used to produce all types of steel. The electric furnace may be either acid- or basic-lined (**Fig. 5**). An advantage over the open-hearth and oxygen converter processes is that very high temperatures can be produced. The high temperature and the flexibility of using a 100% solid charge, or up to 60% molten pig iron, provides considerable versatility for this melting unit. The very high temperatures are particularly useful in producing alloy steels. If large amounts of alloying elements are added to a ladle from an oxygen converter or an open-hearth furnace, the steel is often cooled to the extent that good-quality steel cannot be poured, or if the alloying elements are added in the furnace, they may be oxidized into the slag. This is particularly true for elements such as niobium (columbium), vanadium, and chromium. Furthermore, the requirement in certain types of stainless steel to provide for very low carbon contents requires that the carbon be removed without excessive chromium losses. This is possible at high temperatures such as can be achieved in the electric arc furnace.

The rapid development of steelmaking technology using the electric arc furnace, not only for alloy and stainless steels but especially for carbon steel production, has increased its share of production capacity to about 20% of the steel industry. This increased capacity reflects the ascension of the "mini-mill" as a major steel supplier. The size of the mini-mill also has dramatically increased to the point where this label is not descriptive of this type of steel plant,

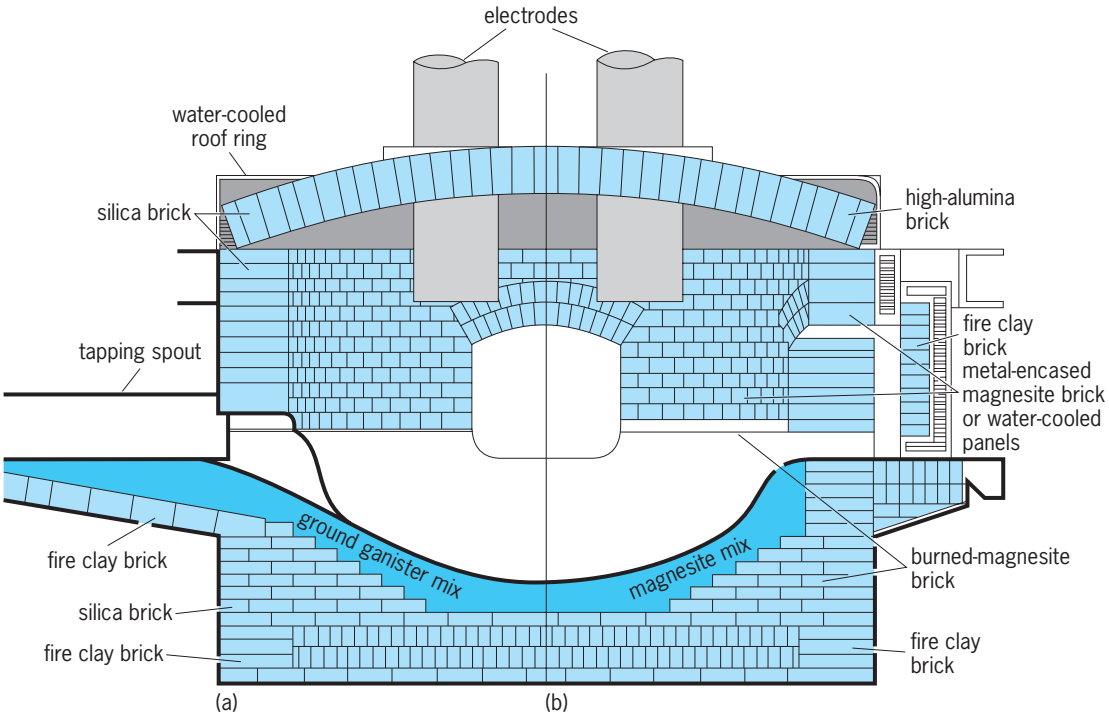

Fig. 5. Schematic cross section of a Heroult electric arc furnace with a flat-bottomed shell and stadium-type subhearth construction, indicating typical refractories employed in (*a*) an acid lining and (*b*) a basic lining. Only two of the three electrodes are shown in this section. (*After H. E. McGannon, ed., The Making, Shaping and Treating of Steel, 8th ed., United States Steel Corporation, 1964*)

several such mills being located within large integrated plants and having two or more 200-ton electric arc furnaces.

The modern mini-mill includes one or more electric furnaces with high-power energy sources; a 100-ton electric furnace may have a 75-MW transformer. These furnaces are water-cooled, at least in the upper wall where cooled panels substitute for refractory above the slag line. Water-cooled roofs are in use, and water-cooled electrode holders are being developed to provide reduced electrode consumption.

The product line of the mini-mill includes special-quality bars and rods, as well as reinforcing bars and small structural shapes. Some mills are producing plate from electric furnace remelting operations. This technology, which may depend on direct-reduced iron in the future, is important in steel production.

Special processes. Remelting and refining of special alloys are carried out in duplex or secondary processes; the principal ones are argon-oxygen decarburization, electroslag refining, vacuum arc remelting, and vacuum induction melting.

In the argon-oxygen decarburization process, a liquid metal is refined in a refractory-lined vessel by injection of gas mixtures, in particular of varying argon-oxygen mixtures, through tuyeres in the side or bottom of the unit. Usually the liquid metal is provided from an electric arc furnace, or in smaller installations from an induction furnace. The argon-oxygen decarburization process was developed originally for the production of stainless steels, but has been modified for the production of many other grades. The argon dilution of the gases in the process provides excellent degassing characteristics which allow achievement of low carbon contents in the liquid metal without chromium losses by oxidation to Cr_2O_3 into the slag; also hydrogen is removed from crack-sensitive grades of steel. Argon-oxygen decarburization provides good slag-metal mixing for removal of sulfur and separation of nonmetallic inclusions.

In the electroslag refining process, an impure electrode prepared from the material to be refined is remelted in a slag which is resistively heated. An electric current passed through the slag melts the electrode. Molten droplets of metal form on the tip of the electrode and are refined as they fall through the slag, forming an ingot in a water-cooled mold. *See* ELECTROMETALLURGY.

In the vacuum arc remelting process, a prepared electrode is arc-melted in a vacuum. The droplets of metal from the electrode are exposed to the vacuum and are solidified in a water-cooled mold. The process provides refining of volatile constituents and removal of dissolved gases, and promotes chemical homogeneity and improved structure.

Control of temperature, pressure, and stirring can be exercised over wide ranges and for extended periods of time in the vacuum induction melting process, allowing considerable flexibility for refining of an alloy. Vacuum induction melting is applied to small melts, as well as large ones ranging up to 60 tons (54 metric tons). The molds are usually enclosed in an adjoining vacuum chamber connected by large-scale access doors and valves; this minimizes the time

required to melt and prepare molds, since they can be done independently. Vacuum induction melting systems, essentially the same as the conventional induction melting systems, utilize high-frequency current in a water-cooled coil surrounding the crucible containing the melt; however, the processing is conducted under low vacuum pressures. *See* VACUUM METALLURGY.

Ladle metallurgy. Ladle metallurgy was used first to produce high-quality steels, but has been extended to producing many grades of steel because of the economic advantages of higher productivity. The purpose of these ladle treatments is to produce clean steel; introduce reactive additions, such as calcium or rare earths; add alloying additions, as for microalloyed steels, with high recovery; and increase furnace utilization, allowing higher-productivity smelting operations of the blast furnace, and melting and refining operations in steelmaking, followed by final ladle treatment of the steel to achieve controlled low levels of contained inclusions. Ladle treatments in steel production generally are classified as synthetic slag systems; gas stirring or purging; direct immersion of reactants, such as rare earths; lance injection of reactants; and wire feeding of reactants. These are often used in combination to produce synergistic effects, for example, synthetic slag and gas stirring for desulfurization followed by direct immersion, injection, or wire feeding for inclusion shape control.

A major aspect of steel development since the mid-1970s has been the effort directed at the improvement of physical and mechanical properties through steel cleanliness and reduced anisotropy, which are largely dependent upon the amount and form of oxygen and sulfur in the steel. Such improved properties are required by steels used for thin-sheet materials, heavy plate, high-strength low-alloy (HSLA), and for linepipe. Thin sheet steels require good formability; heavy plate requires guaranteed through-thickness properties; and linepipe steels require high-transverse-impact energy values. Removal of reaction products is most critical during the course of deoxidation of molten steel. Nonmetallic inclusions in the form of oxides of silicon, manganese, and aluminum may adversely influence final properties.

Sulfur in steel has always presented a problem in steelmaking. Improved surface quality can be achieved by lowering sulfur levels to 0.015% or below. For example, the achievement of low sulfur contents with two-slag practices in electric furnaces presents operating difficulties and production delays.

In the 1950s steel was produced primarily in the open-hearth furnace, where sulfur levels were normally in the range from 0.03 to 0.04%. Sulfur contents of 0.02 to 0.03% could be achieved with lower sulfur fuels and oxygen roof lances, permitting shorter heat times and improved mixing of metal and slag. The basic oxygen steelmaking process brought about an even further improvement within the range of 0.015–0.025% sulfur, and with the bottom-blown oxygen converter Q-BOP (Quelle-Basic oxygen process), sulfur contents in the 0.01–0.02% range can be achieved. However, even lower sulfur contents are required.

In killed steels with low oxygen content, such as when aluminum is used for deoxidation and grain size control, sulfur combines with manganese as highly deformable manganese sulfides. These manganese sulfides have a low melting point and, as the last liquid to solidify in the steel, collect as films at grain boundaries. During hot rolling, the manganese sulfides are plastically deformed into elongated stringers extending parallel to the rolling direction. This shape and distribution of sulfides can have a marked effect on the directional properties of steel products.

Two methods can be utilized to minimize the directionality problem. One is to reduce the sulfur content in the steel product to a minimum, lowering the total volume of manganese sulfide inclusions. This can be accomplished by decreasing the total sulfur input, treating the hot metal to remove sulfur, and controlling the scrap charge; by adjusting melting practices, including the use of additional quantities of flux; and by direct treatment of the liquid steel with desulfurizing agents, such as calcium or magnesium. A second approach involves modifying the shape or morphology of sulfide inclusions by producing a relatively nondeformable complex sulfide or oxysulfide by the addition of calcium or rare-earth metals.

Reoxidation and inclusion minimization. Since there is a strong affinity between oxygen and liquid steel, and the necessity to restrict pickup of oxygen during teeming operations, several steps to control reoxidation are advisable. The first is the protection of the steel pouring stream, whether from ladle to ingot mold or tundish, or from tundish to continuous casting mold. The separation of reoxidation products, or deoxidation products formed during cooling of the steel in the teeming operation, can be promoted. In particular, in continuous casting the appropriate design of tundish in terms of residence time of the steel in the tundish and the appropriate fluid flow patterns can enhance separation of oxides from the liquid steel. *See* METAL CASTING; PYROMETALLURGY; REFRACTORY; STAINLESS STEEL; STEEL.

Robert D. Pehlke

Bibliography. Association of Iron and Steel Engineers (AISE), *The Making, Shaping, and Treating of Steel*, 10th ed., 1985; A. Ghosh, *Secondary Steelmaking: Principles and Applications*, 2000; Iron and Steel Society, *Electric Furnace Steelmaking: Collective Work*, 1985; T. Kawaguchi and K. Sugiyama, *Steel Industry I: Manufacturing System* 1989.

Stellar evolution

The large-scale, systematic, and irreversible changes with time of the structure and composition of a star.

Types of stars. Dozens of different kinds of stars populate the Milky Way Galaxy. The most common

are main-sequence dwarfs like the Sun that fuse hydrogen into helium within their cores (the core of the Sun occupies about half the mass). Dwarfs run the full gamut of stellar masses, from perhaps as much as 120 solar masses (120 $M_\odot$) down to the minimum of 0.075 solar mass (beneath which the full proton-proton chain does not operate). They occupy the spectral sequence from class O (maximum effective temperature 50,000 K or 90,000°F, maximum luminosity 5×10^6 solar), through classes B, A, F, G, K, and M, to the new class L (2000 K or 3100°F and under, minimum luminosity 4×10^{-5} solar). Within the main sequence, they break into two broad groups, those under 1.3 solar masses (class F5) whose luminosities derive from the proton-proton chain, and higher-mass stars that are supported principally by the carbon cycle. Below the end of the main sequence (masses less than 0.075 $M_\odot$) lie, the brown dwarfs that occupy half of class L and all of class T (the latter around 1000 K or 1300°F). These shine both from gravitational energy and from fusion of their natural deuterium. Their low-mass limit is unknown. *See* CARBON-NITROGEN-OXYGEN CYCLES; PROTON-PROTON CHAIN; SPECTRAL TYPE; STAR.

Scattered among the dwarfs are different kinds of stars distinguishable (through comparison with the dwarfs) by their sizes and luminosities: bright giants (with maximum radii equaling that of the inner solar system), brilliant supergiants (maximum radii comparable to the orbits of Jupiter and Saturn), subgiants (which fall between giants and dwarfs), and dim white dwarfs (the size of Earth). All these classifications may be broken into subclasses. The list also includes the central stars of planetary nebulae, double stars, novae, supernovae, neutron stars, and black holes. The study of stellar evolution seeks to determine where they all came from. Do they stand alone, or are they somehow connected? If the latter, what kinds of stars precede and succeed what other kinds?

Stellar models. Stars age so slowly that it is impossible—except under rare circumstances—to see the transformations take place. Instead, stellar theorists, using the laws of physics, build numerical models of stars and then age them. The procedure is in principle simple, in practice difficult. At the start, a star is in hydrostatic equilibrium in which each layer supports the layers above it (the pressure difference across a layer equaling the weight of the layer), and is internally supported by thermonuclear fusion. The relation between pressure, density, and temperature is given by the perfect gas law, $P = NkT$, where P is pressure, N is the number density of atomic particles, k is Boltzmann's constant, and T is temperature. As fusion proceeds, and the rate of energy generation and the internal chemical composition change, the structure of the star changes as well. Stepwise calculations that predict luminosity, radius, and effective temperature then allow the star to be followed as it "moves" (changes effective temperature and luminosity) on the Hertzsprung-Russell (HR) diagram. When appropriate mass loss is taken into account, most of the different kinds of stars fall into place. *See* GAS; HERTZSPRUNG-RUSSELL DIAGRAM; HYDROSTATICS; THERMODYNAMIC PROCESSES.

Star formation. Stars are born from compact knots within dark molecular clouds that are refrigerated by dust that blocks heating starlight. If the random knots, compressed by supernovae or other means, are dense enough, they can contract under their own gravity. Conservation of angular momentum demands that as they collapse they must spin faster. Star formation requires that angular momentum be removed such that the new "young stellar objects" do not completely tear themselves apart before they can become fully developing protostars. *See* ANGULAR MOMENTUM; MOLECULAR CLOUD.

High-speed particles (cosmic rays) from exploding stars partially ionize the dusty knots. The ions grab onto the weak magnetic field of the Galaxy and, as a result of their physical interaction with neutral atoms and molecules, provide the initial means to slow the rotation. If the rotation is still too fast, the contracting body may split into a double (or more complex) star, though the origins of doubles are not clearly solved. A contracting protostar still indeed rotates progressively faster until the part of its mass not accreted by the star itself is spun out into a dusty disk, from which planets might later accumulate. From the disk shoot powerful molecular flows that slow the star still more (**Fig. 1**). *See* PROTOSTAR.

When the protostar's interior reaches about 10^6 K (1.8×10^6 °F), it can fuse its internal deuterium. That and convection, which brings in fresh deuterium from outside the nuclear-burning zone, bring some stability, and a star can now be said to be born. Stars like the Sun shrink at constant temperature until deuterium fusion dies down. Then they heat at roughly constant luminosity until the full proton-proton chain begins, which provides the stars' luminosity and stops the contraction. The stars settle onto the zero-age main sequence (from which they

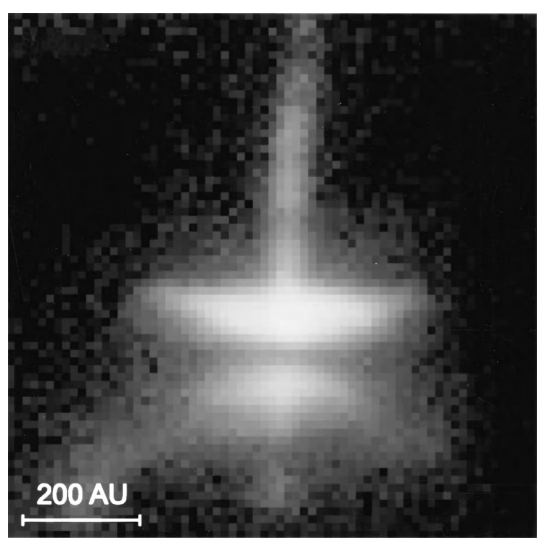

Fig. 1. Focused, even created, by a circumstellar disk, a jet pours from the youthful star Herbig-Haro 30 (HH 30). (*C. Burrows, Space Telescope Science Institute; WPFC2 Science Team; NASA*)

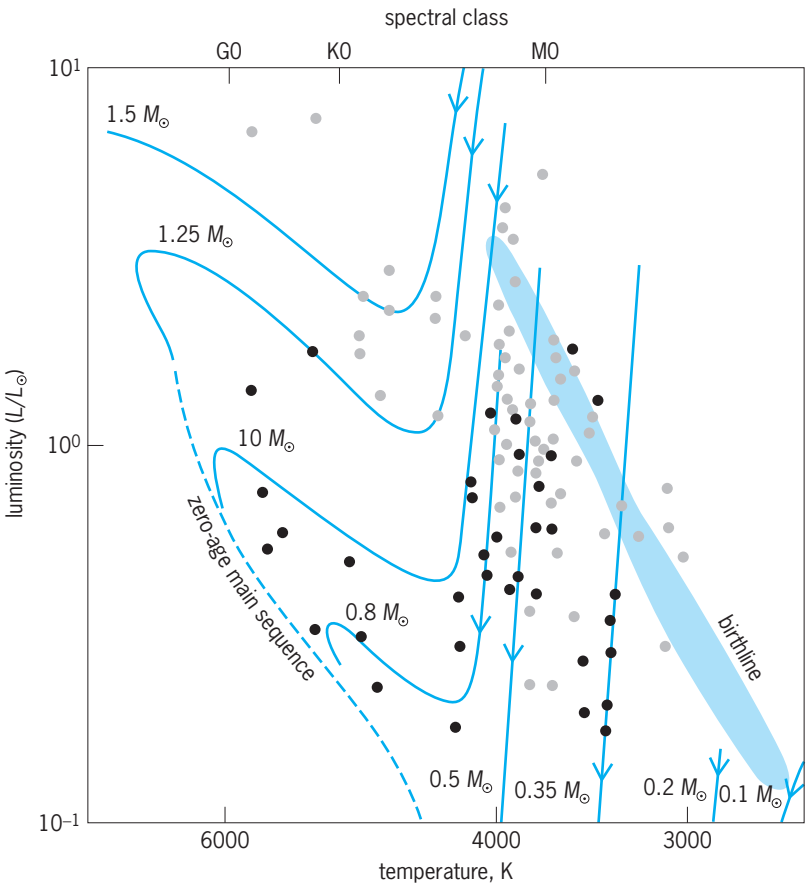

Fig. 2. Hertzsprung-Russell diagram of star formation. Stars begin to stabilize at the birthline and then descend on evolutionary tracks as T Tauri stars (dots) toward the zero-age main sequence. (*After J. B. Kaler, Cosmic Clouds, W. H. Freeman, 1997, from work of S. W. Stahler and I. Iben, Jr.*)

will later evolve). At the same time, the surrounding dusty cloud is clearing, allowing new, accreting, and highly active T Tauri stars to be seen flocking around their birthclouds (**Fig. 2**). The whole process takes only 10 or so million years, the mature stars then leaving their birthplaces, destined both to age and to orbit the Galaxy. High-mass stars proceed similarly, but at such a great pace that the death process begins even as the birth process is ending. *See* T TAURI STAR.

Early evolution. The main sequence is that zone on the HR diagram in which stars are stabilized against gravitational contraction by fusion. The higher the stellar mass, the greater the internal compression and temperature, and the more luminous the star. Hydrogen fusion is highly sensitive to temperature, a small increase in stellar mass meaning a much higher fusion rate. Although greater mass means a greater nuclear-burning core mass and therefore a larger fuel supply, the increased fusion rate is more than offsetting and thereby shortens stellar life. A star's fuel supply is proportional to mass (M), while the rate of fusion is expressible by the luminosity (L). The average mass-luminosity relation found from binary stars is $L \propto M^{3.5}$. If a star's lifetime t is proportional to (fuel supply)/(rate of use), then $t \propto M/L \propto M/M^{3.5} = 1/M^{2.5}$. While the new Sun was destined to survive on the main sequence for 10^{10} years, a 0.1-solar-mass star will live there for 10^{13} years, while a 100-solar-mass star will exhaust its core hydrogen in only 2.5×10^6 years.

Main-sequence life, while stable, is not altogether quiet. Even there stars change and evolve. As hydrogen "burns" to helium, four particles (protons) are converted to one (a helium nucleus). The pressure of a gas depends on the number of particles per unit volume, not on their kind. The result is a slow shrinkage of the core, which increases the temperature, raises the fusion rate, and causes the core to eat into the surrounding hydrogen envelope (incorporating fresh fuel). As a result, main-sequence dwarfs slowly brighten and eventually expand and cool some at their surfaces. (The Sun will more than double in brightness, dooming life on Earth long before core hydrogen is gone.) As a result, the main sequence spreads itself into a band toward the cool side (to the right) of the zero-age main sequence, the band widening toward the top.

Clusters of stars are born together in both space and time, most (presumably) with fully intact main sequences. As a cluster ages, stars peel off the main sequence (to become giants and supergiants) from the top down. A star cluster can be dated by where its main sequence ends. Open clusters, which occupy the Galaxy's disk, range from just-forming to about 10^{10} years old, which gives the disk's age. The globular clusters of the Galaxy's halo, however, have burned their main sequences down to around 0.8 solar. This process takes roughly 12×10^9 years, which must be close to the age of the Galaxy. *See* MILKY WAY GALAXY; STAR CLUSTERS.

The main sequence is divided into three parts. Below 0.8 solar mass (roughly class G8), no star has ever had time to evolve. Low-mass (K and M dwarfs) evolution is therefore of only academic interest. Stellar evolution deals with less than 15% of the stars. Their fates depend again on mass. Between 0.8 solar mass and around 10 solar masses (classes G8 to B1), the stars die as white dwarfs. Above it (classes O and B0), they explode. Binary stars contribute further to the richness of stellar phenomena.

Intermediate-mass evolution. Main-sequence life lasts until the core hydrogen is almost gone, at which time hydrogen fusion rather suddenly shuts down. With no support, the now-quiet helium core can contract more rapidly under gravity's force. It heats, causing hydrogen fusion to spread into a thick enclosing shell that runs on the carbon cycle. With a new (though temporary) energy source, the star first dims some while it expands and cools at the surface, changing its spectrum to class K. The transition takes only a few hundred million years or less, leaving few stars in the middle of the HR diagram, the lower masses appearing as F, G, and K subgiants.

The consequences again depend on mass. As core contraction proceeds beyond the rightward transition in the HR diagram, stars from about 1 to 5 solar masses (still fusing hydrogen in a shell) suddenly and dramatically increase their luminosities. The future Sun will eventually grow 1000 times brighter

than it is today, and a 5-solar-mass star (which begins at about 600 solar luminosities) will reach nearly 3000 solar. At the same time, the stars swell to become red giants. While the core (roughly half a solar mass) of the future Sun shrinks to the size of Earth, the radius will expand to that of Mercury's orbit, or even beyond. When the core temperature climbs to 10^8 K, helium nuclei (alpha particles) begin fusing to unstable beryllium (^{8}Be), which quickly decays back to alpha particles, setting up an equilibrium. The tiny amount of ^{8}Be present reacts with additional alpha particles to create helium via the triple alpha process:

$$^4\text{He} + {}^4\text{He} \leftrightarrow {}^8\text{Be}$$

$$^8\text{Be} + {}^4\text{He} \rightarrow {}^{12}\text{C} + \text{gamma ray}$$

(The "skip" over lithium, beryllium, and boron renders these elements rare.) The initiation of the triple-alpha process is explosive in stars like the Sun, since the core electrons become degenerate, like those that support white dwarfs. (The sudden burst of energy is absorbed and does not reach the stellar surface.) Fusion with additional helium nuclei creates oxygen and even neon. Above 2 solar masses, helium burning starts more quietly (**Fig. 3**). *See* NUCLEOSYNTHESIS.

The star, now stabilized by a helium-burning core that is surrounded by a hydrogen-fusing shell, retreats about halfway down the red giant branch. The numerous lower-mass stars reside in the class K "red giant clump." Low-mass metal-deficient globular cluster stars that have suffered different rates of mass loss spread out from the clump toward higher temperatures to create the distinctive horizontal branch. Energy-generating fusion reactions will try to proceed toward iron (the most stable of all nuclei). Some 80% of the energy is generated in hydrogen burning, so (discounting further burning modes) the helium-burning stage lasts only around 20% of the main-sequence lifetime. From 5 solar masses up, evolution proceeds similarly, but instead of settling into a distinct location on the HR diagram, the stars loop to the blue (higher temperatures) where they fuse helium as class G, F, and A giants. This change in evolutionary style is not abrupt, but gradual with increasing mass. *See* GIANT STAR.

Asymptotic giant branch (AGB). When the helium has fused to carbon and oxygen, the core again contracts. Helium fusion spreads outward into a shell, and for the second time the star climbs the HR diagram's giant branch. Since the second climb is roughly asymptotic to the first, the second climb creates the asymptotic giant branch. The shrinking carbon-oxygen core is now surrounded by a shell of fusing helium, while the old hydrogen-fusing shell expands, cools, and shuts down. Eventually, however, the helium shell runs out of fuel, and the hydrogen shell reignites. Hydrogen burning feeds fresh helium into the space between it and the carbon core, and when there is enough of it, helium burning reignites explosively in a helium flash (or thermal pulse) that can affect the star's surface. The

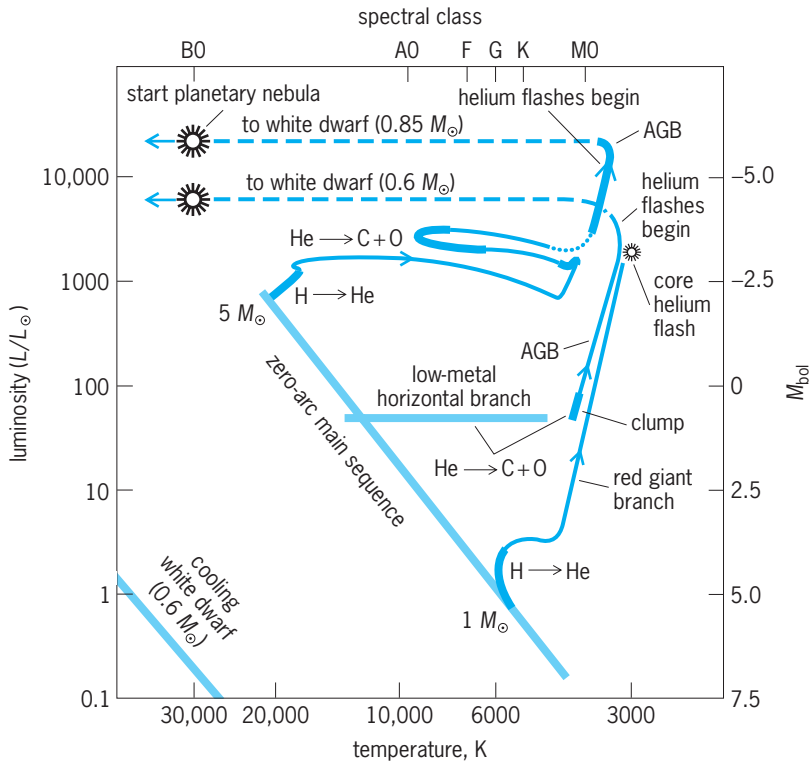

Fig. 3. Hertzsprung-Russell diagram of evolution of intermediate-mass stars. These stars evolve from the main sequence to become giants that stabilize during helium fusion in a "clump." When helium fusion is done, they brighten again as AGB stars, lose their envelopes, and evolve to the left to produce planetary nebulae, finally at lower left becoming white dwarfs. (*After J. B. Kaler, Stars, W. H. Freeman, 1992, 1993, from work of I. Iben, Jr.*)

flash squelches hydrogen fusion, and the whole process starts again, helium flashes coming at progressively shorter intervals. AGB stars become larger and brighter than before, passing into the cool end of class M, where they eventually become unstable enough to pulsate as long-period variables (Miras). The Sun will become 5000 times brighter than now and will reach out to the Earth's orbit, perhaps destroying the planet. More massive Miras can exceed the size of the Martian orbit. *See* MIRA; VARIABLE STAR.

During this activity, stars go through various stages of convective dredge-up, in which they can raise elements created in their nuclear-burning zones to the stellar surfaces. The first of these dredge-ups alters the nitrogen and carbon isotope ratios while the star is on the first ascent of the red giant branch. Above about 3 solar masses, a second dredge-up in the early AGB stage can increase both surface nitrogen and helium. In the helium-flashing state, even carbon from helium fusion can be brought upward. Moreover, a host of elements created by slow neutron capture are elevated as well, including zirconium, strontium, barium, and many others. When the carbon abundance equals that of oxygen, an S star is seen (loaded with fresh zirconium); when carbon exceeds oxygen, a genuine carbon star is seen. *See* CARBON STAR.

Mass loss and planetary nebulae. During the giant stages, stellar winds greatly increase. Mira pulsations cause shock waves that help drive mass from the

stellar surfaces, where the cooled gas becomes ever richer in molecules, some even condensing into dust grains. High luminosity pushes the dust outward, and the dust couples with the gas, resulting in slow (10 km/s or 6 mi/s) thick winds tens of millions of times stronger than the solar wind (up to 10^{-5} solar mass per year). Because stars are in this state for hundreds of thousands of years, they will lose much of themselves back into space—the Sun about half of its mass, a 10-solar-mass star over 80%. Advanced giants and Miras lose so much that they can visually disappear within dust clouds. Carbon star winds contain various kinds of carbon dust as well as organic molecules, while oxygen-rich stars produce silicate dust. The dust blown into interstellar space helps make new generations of stars. The elements created in the nuclear-burning shells are added to the interstellar inventory, where they too find their way into new generations of both stars and planets.

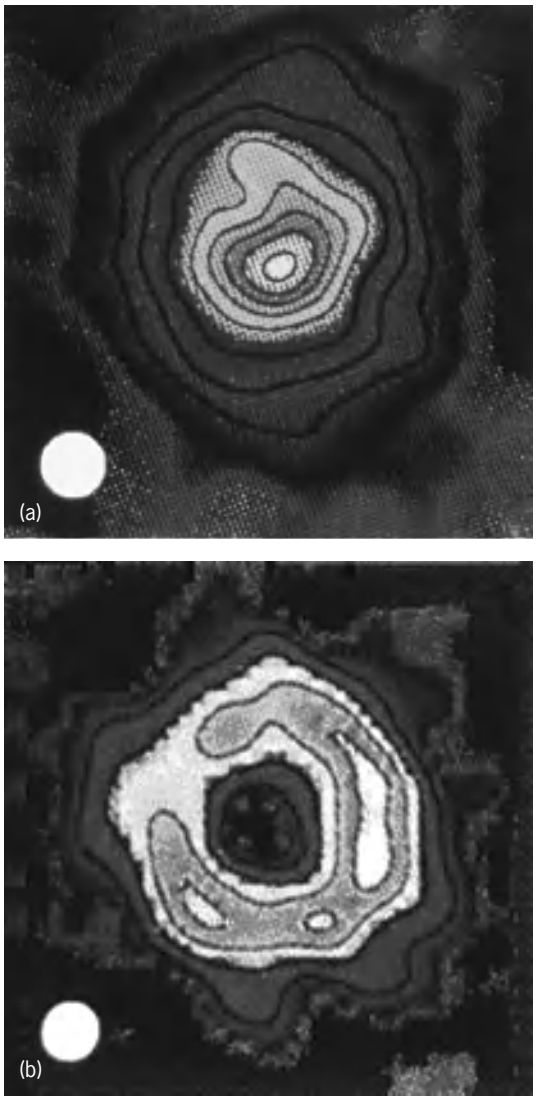

Fig. 4. Thick molecule-filled cloud surrounding the carbon star IRC+10 216. (*a*) Image made in radiation from hydrocyanic acid (HCN). (*b*) Image in radiation from cyanoacetylene (HC_3N). Each map is 6000 astronomical units across. (*J. Bieging; Berkeley-Illinois-Maryland Association*)

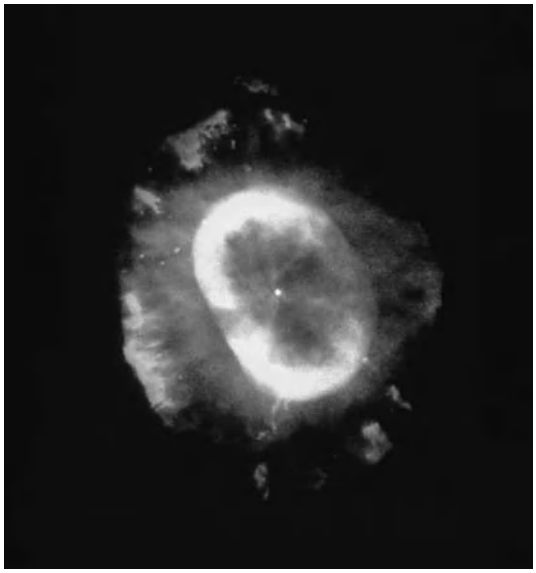

Fig. 5. Hubble Space Telescope image of planetary nebula NGC 7662. Planetary nebulae are formed when giant stars slough their outer envelopes, leaving a white dwarf behind. (*B. Balick et al.; Space Telescope Science Institute; NASA*)

A good portion of the carbon, most of the nitrogen, and significant fractions of the Galaxy's other elements come from such winds (**Fig. 4**).

So much mass is lost that an evolving star becomes stripped nearly to its fusion zone, which is protected from the outside by a low-mass hydrogen envelope. As the inner region becomes exposed, the wind diminishes in mass but increases in speed and temperature. Hammering at the surrounding dusty, molecule-filled shroud of lost mass, the high-speed wind compresses the inner edge into a thick ring. Eaten away from the top by the wind and from below by fusion, the stellar envelope shrinks, slowly exposing the hot shell-core structure beneath. When the stripped star's surface reaches 25,000 K (45,000°F), the dense ring that its high-speed wind had previously created is ionized. Subsequent recapture of electrons by ions, along with collisional excitation of heavy atoms, causes the shell to glow, and a planetary nebula is born. The first of these was announced by William Herschel (who named the class on the basis of their disklike shapes) in 1785. Over a thousand are now known (**Fig. 5**). *See* PLANETARY NEBULA.

As the planetary nebula, expanding at about 20 km/s (12 mi/s), grows in size, it eventually becomes so tenuous that the ionizing radiation bursts through, allowing much of the full structure to be seen. After 50,000 years, the nebula invisibly merges with the interstellar medium. During this time, the star inside first heats at constant luminosity to over 100,000 K (180,000°F), the luminosity and final temperature depending on the old core's mass (which ranges from around 0.5 $M_\odot$ to nearly 1.4 $M_\odot$, the Chandrasekhar limit, above which white dwarfs cannot exist). As residual nuclear fusion shuts down, the star cools and dims at constant radius to become a white dwarf. The luminosity-mass relation is now

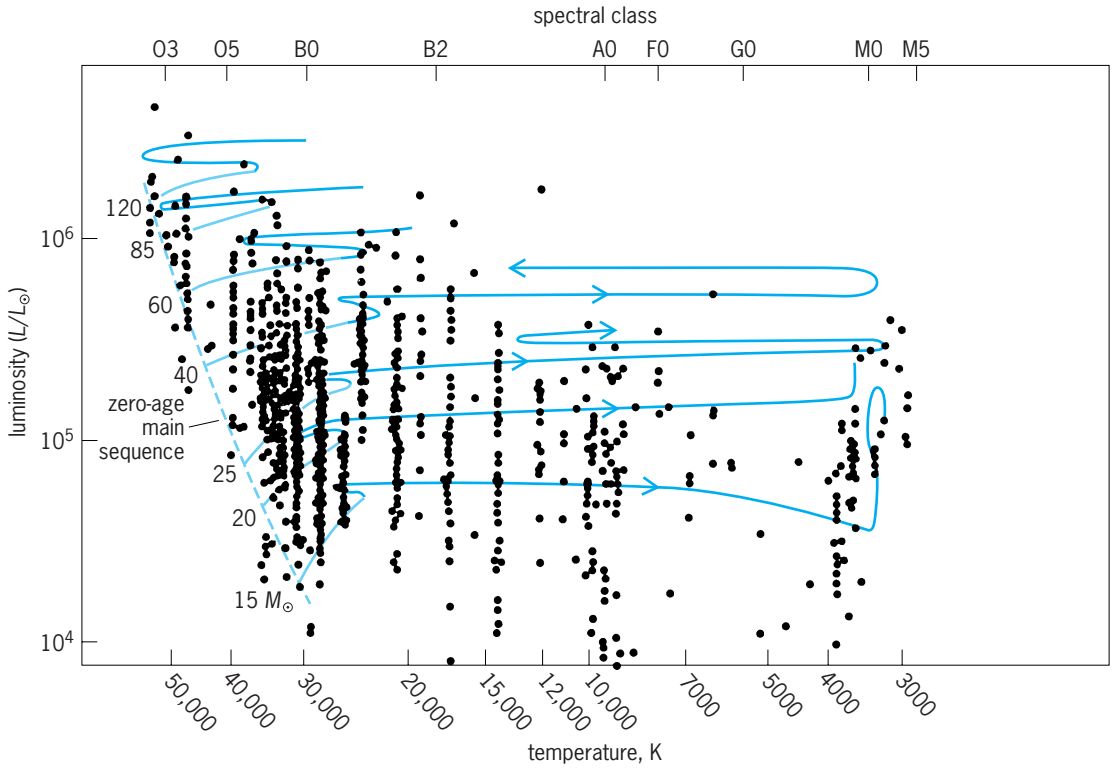

Fig. 6. Hertzsprung-Russell diagram for evolution of high-mass stars. These stars evolve to red supergiants, some evolving back to blue supergiants. Above about 60 solar masses, evolution is stalled by extreme mass loss. (*From work of A. Maeder and G. Meynet; and R. Humphreys and K. Davidson*)

reversed, as higher mass means higher gravitational compression and smaller radius.

Made of carbon and oxygen surrounded by a thin hydrogen or helium atmosphere, the white dwarf is supported forever by degenerate electron pressure. Since the evolutionary and cooling times of white dwarfs can be calculated theoretically, the oldest of them provides an age measure for the galactic disk, which accords well with that found from open clusters. *See* WHITE DWARF STAR.

High-mass evolution. As the mass of a star increases, so does the mass of the core. The Sun will turn into a white dwarf of around 0.6 solar mass. At 10 solar masses, the core reaches the Chandrasekhar limit, and the star cannot become a white dwarf. At first, high-mass evolution proceeds similarly to that of stars of lower mass. As high-mass stars use their core hydrogen, they too migrate to the right on the HR diagram, becoming not so much brighter but larger, cooling at their surfaces and turning into supergiants. Below about 40 solar masses, they become class M red supergiants, losing huge amounts of mass through immense winds. Some (depending on mass) stabilize there by the start of helium fusion; others loop back to become blue supergiants. Above 60 solar masses, so much mass is lost through winds that the stars do not make it much past class B, stalling there as helium fusion begins (**Fig. 6**). As the stars—as well as bright giants—loop through the middle of the diagram, some become Cepheid variables. *See* CEPHEIDS; SUPERGIANT STAR.

Though almost all these supergiants vary to some extent, the most massive become unstable and undergo huge eruptions. In 1846, Eta Carinae, near 100 solar masses (and a probable double star), brightened to match Sirius and Canopus, lost a solar mass, and then faded to near the edge of naked-eye vision (**Fig. 7**). Such luminous blue variables (LBVs), plus similar lesser lights, are losing their hydrogen envelopes and may be turning into helium-rich Wolf-Rayet stars, highly luminous stars that have been

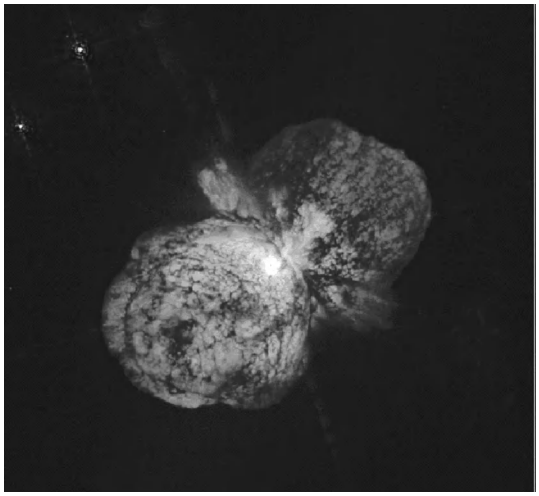

Fig. 7. Hubble Space Telescope image of a luminous blue variable, the supergiant Eta Carinae, buried in the middle of a giant expanding cloud of its own making. (*J. Morse, University of Colorado; Space Telescope Science Institute; NASA*)

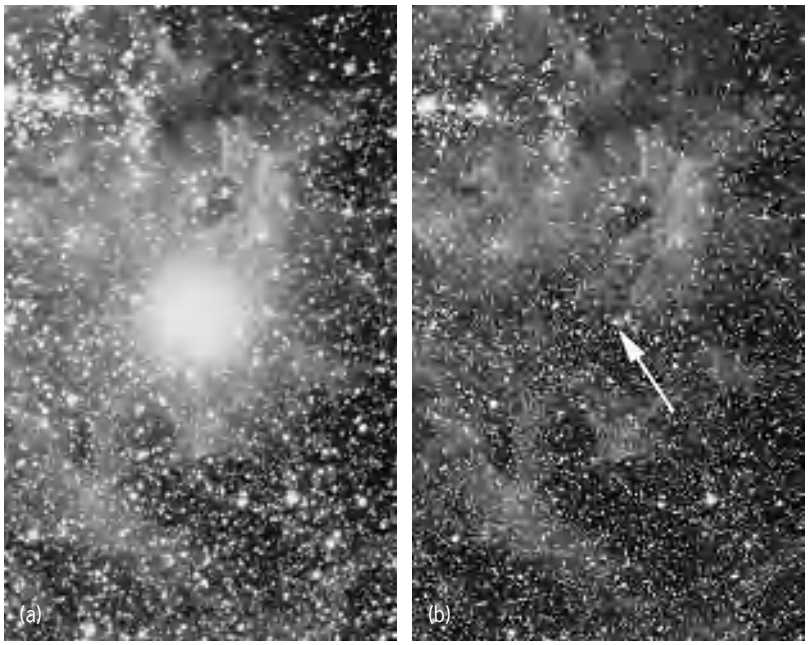

Fig. 8. Field around the type II core-collapse supernova SN 1987A. (*a*) Two weeks after the supernova exploded in the Large Magellanic Cloud in 1987. The supernova far outshines its neighboring stars, and was easily visible 150,000 light-years away. (*b*) Before the explosion it was a rather ordinary B1 blue supergiant (indicated by arrow). (© *Anglo-Australian Observatory, photograph by David Malin*)

almost completely stripped of their outer hydrogen. Though the sequence is not clear, the luminous blue variables probably first turn into the nitrogen-rich WN variety (whose nitrogen has been enriched from hydrogen fusion by the carbon cycle) and then into the carbon-rich WC species (the carbon from more advanced helium fusion).

While intermediate-star fusion stops at carbon and oxygen, supergiants continue onward. The carbon-oxygen core shrinks and heats to the point that carbon fusion can begin, and then carbon and oxygen convert to a more complex mix dominated by oxygen, neon, and magnesium. Helium fusion now continues in a surrounding shell that is nested in one that is fusing hydrogen. Once carbon burning has run its course, the unsupported oxygen-neon-magnesium core shrinks and heats, and now it is carbon burning's turn to move outward into a shell. When hot enough, the oxygen-neon-magnesium mix fires up to burn to one dominated by silicon and sulfur. Continuing the process, the developed silicon and sulfur core finally becomes hot enough to fuse to iron, the silicon-burning core wrapped in oxygen-neon-magnesium, helium, and hydrogen-burning shells.

Supernova. Each nuclear fusion stage generates less energy, and since each takes on the role of supporting the star, each lasts a shorter period of time. While hydrogen fusion takes millions of years, the iron core is created from silicon fusion in a matter of weeks. Iron cannot fuse and produce energy. The core, about 1.5 solar masses and the size of Earth, suddenly collapses at a speed a good fraction that of light. The iron atoms are broken back to neutrons and protons. Under crushing densities, free electrons merge with protons to make yet more neutrons. When the collapsing neutron core hits nuclear densities of 10^{14} g/cm^3 (10^{14} times the density of water), it violently bounces, and the rebound tears away all the outer layers.

From the outside, the observer sees the explosion as a type II supernova that can reach an absolute magnitude of −18 (**Fig. 8**). Temperatures within the exploding layers are high enough that a great variety of nuclear reactions, from equilibrium fusion of one element into another to rapid neutron capture, create all the chemical elements up to and beyond uranium. The optical glow eventually comes mainly from the manufacture of a tenth or so of a solar mass of radioactive nickel, ^{56}Ni, which quickly decays to radioactive cobalt and then to stable iron, ^{56}Fe, which gets blasted into the cosmos along with everything else. *See* RADIOACTIVITY.

The nature of the rebound is far from understood. By itself, it does not account for the explosion of a supernova. However, the merger of core protons and electrons generates vast numbers of near-massless neutrinos. Nearly everything is transparent to them. Those created in the solar core by the proton-proton chain go right through the Earth. The densities around the rebounding neutron core, however, are so high that the neutrinos are trapped, and may aid in shoving the surrounding shells and envelope outward. Other ideas invoke violent convection within the neutron core to provide the energy. However it is done, the outflowing neutrinos carry away 100 times more energy than does the blast itself. The power output of a single supernova for a brief moment equals that of the rest of the stars in the visible universe. At the highest progenitor masses, supernovae are believed to eject their energy in bipolar jets, leading to longer-duration gamma-ray bursts (GRBs). The progeny of O stars, supernovae are rare, occurring in the Milky Way Galaxy at a rate of only one or two per century, so the odds of one being close to Earth are fortunately small. *See* GAMMA-RAY BURSTS; NEUTRINO; SUPERNOVA.

Supernova remnants and compact remains. The debris of the explosion, the supernova remnant, highly enriched in iron and other heavy elements, expands for centuries into space. Its shock wave energizes interstellar space, heats vast bubbles of gas to hundreds of thousands of degrees, and helps compress interstellar matter and create new stars. Its heavy elements become incorporated into molecular clouds. Most of the Earth came from previous generations of supernovae, each one contributing perhaps a mountain mass. Accelerated by the shock, ambient interstellar electrons, protons, and heavier nuclei approach the speed of light and speed through the Galaxy as cosmic rays, which in their turn also help initiate star formation (**Fig. 9**). *See* CRAB NEBULA; INTERSTELLAR MATTER; NEBULA; SHOCK WAVE.

Exposed by the explosion is the compact hot neutron star. Only 25 km (15 mi) across, it is stabilized by the pressure of neutron degeneracy. Conservation of angular momentum makes the little star spin dozens of times per second, its magnetic field compacted to

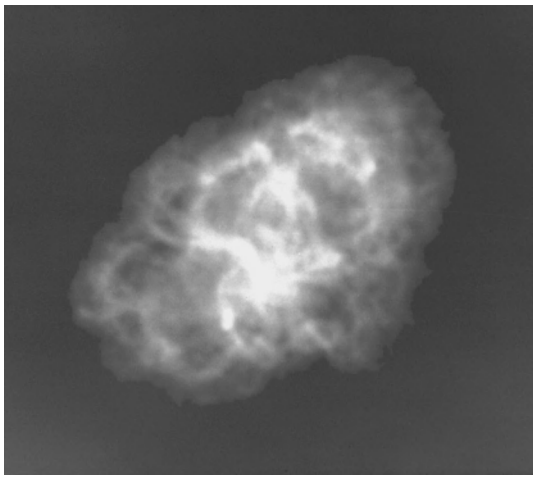

Fig. 9. Crab Nebula, the remnant of the supernova of 1054, as seen in radio radiation. (*National Radio Astronomy Observatories*)

10^{12} or more times stronger than Earth's. Radiation beams out along a tilted, wobbling magnetic axis, and if Earth is in its path, a "pulse" of radiation is observed, the neutron star now a "pulsar." As the neutron star radiates away its energy, it spins more slowly and finally disappears from view. The most highly magnetized neutron stars, the magnetars, can have magnetic field strengths greater than 10^{14} times that of Earth, and are related to anomalous x-ray pulsars (AXPs) and to soft gamma-ray repeaters (SGRs), the latter occasionally releasing bursts so powerful that they can affect the Earth. The Galaxy must contain over 10^8 quiet, near-invisible neutron stars. Off-center detonation in the supernovae that created them may give them kicks that can send them off at speeds far higher than those of most of the Galaxy's stars. *See* NEUTRON STAR; PULSAR.

Beyond the Chandrasekhar limit of $1.4\ M_\odot$, electron degeneracy pressure cannot stabilize a white dwarf, and it must collapse. Neutron stars are close to this limit, and are themselves similarly limited to around 3 solar masses. Stars at the upper end of the main sequence are expected to create iron cores that exceed even this limiting mass. The dense remains cannot stabilize, and they too must therefore collapse. When such a collapsing star passes a critical radius at which the escape velocity is about that of light, it become invisible, and a black hole is born. About half a dozen black hole candidates are recognized in binary systems in which the black hole affects the companion. *See* BLACK HOLE.

Double-star evolution. Double (binary) stars are a common fact of stellar life. Within wide margins, about half the Galaxy's stars have companions. If the components of a binary are well separated, by tens or hundreds of astronomical units, the stars evolve separately as if each were single. If the stars are close enough together, however (the critical limits dependent on mass), they can profoundly affect each other and their courses of evolution.

A binary is first made of a pair of dwarf stars. The more massive evolves first into a giant or even a supergiant. If the two are close enough to start with, the growing giant can be affected by the smaller star. As it swells, it approaches a teardrop-shaped tidal surface in which the gravitational pulls from the two stars (including the forces resulting from orbital motion) are effectively equal. If the giant or supergiant actually reaches the zero-gravity surface, mass can flow through the point of the teardrop toward the smaller star. If the two still have sufficient separation, the incoming matter will first flow into a disk around the dwarf, from which the dwarf will accrete mass. If the two are close enough, matter will impact the dwarf directly without forming a disk. In either case, the result is that mass will be transferred, resulting both in interesting spectral activity and in the more massive star eventually becoming the less massive. Among the most dramatic examples of this kind of behavior is the eclipsing binary Algol, in which a class B dwarf accretes mass from a K giant. Mass loss can also wrap the pair in a common envelope. That and angular momentum losses resulting from the close interaction as well as from interacting magnetic fields can draw the two closer together and perhaps even make them merge.

The giant eventually produces a planetary nebula and then becomes a white dwarf. Mass loss has left the white dwarf with almost no hydrogen, perhaps only a thin skin on its surface. If the two stars have been drawn close enough, the white dwarf can tidally stretch the remaining dwarf (usually a G, K, or M star) to the zero-gravity surface, which can then transfer mass, mostly hydrogen, back to the white dwarf (**Fig. 10**). Instabilities in the accretion disk, as well as ionization and illumination of the flowing stream by a hot spot created by interacting gases, make the pair a flickering cataclysmic variable (CV). When the fresh hydrogen layer is thick, compressed, and hot enough, it explosively burns via the carbon cycle, the thermonuclear runaway creating a nova that brightens to absolute visual magnitude $M_V = -10$ or so. Several are seen each year, one of first magnitude every few decades. Since a nova is only a surface phenomenon and does not much affect the stars of the binary, the system settles back into its prior CV behavior, and after 10^5 years may produce

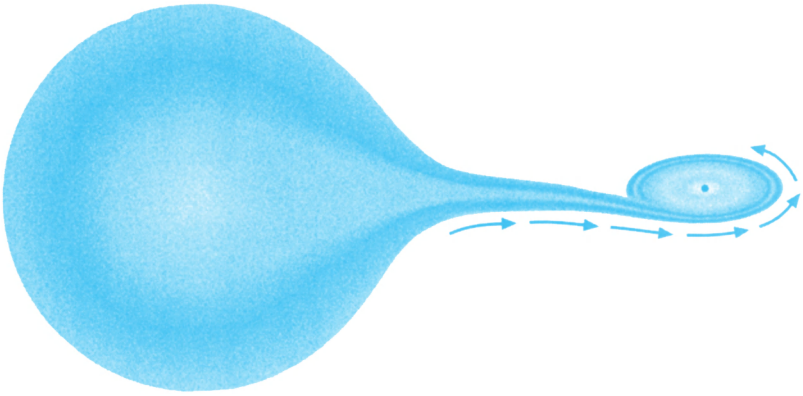

Fig. 10. A star fills its tidal surface and passes mass onto a white dwarf in preparation for a nova or even a type Ia supernova. (*From J. B. Kaler, The Little Book of Stars, Copernicus Books, New York, 2000*)

a nova explosion again. *See* CATACLYSMIC VARIABLE; NOVA.

In wider dwarf–white dwarf combinations, the ordinary dwarf will eventually begin its own evolution and become a giant. If it had received mass from its companion when it was a dwarf (perhaps through a wind), the giant may have been contaminated with by-products of nuclear fusion and appear as a barium star, anomalously rich in barium and other elements. As the giant grows, it passes its mass to the white dwarf through its vigorous wind or via an accretion disk. The hot white dwarf, or a hot boundary layer between an accretion disk and the white dwarf's surface, now gives the binary dual characteristics, those of both a hot and a cool star at the same time. These symbiotic stars can vary dramatically, and the transferred hydrogen can erupt into long-term fusion reactions. *See* SYMBIOTIC STAR.

White dwarf supernovae. There are two quite different kinds of supernovae. Type II supernovae have hydrogen in their spectra. They occur in galactic disks, where young high-mass stars reside, and are produced by iron-core collapse. Type I events, however, do not exhibit spectral hydrogen. They are divided into three categories: Type Ia supernovae display an absorption line of ionized silicon, whereas type Ib exhibits helium instead of silicon, and Ic has neither (or weak helium). While Ib and Ic supernovae are also confined to the disks of host galaxies, Ia supernovae occur in both disks and galaxy haloes where there are no high-mass stars. Tycho's Star of 1572 and possibly Kepler's Star of 1604 were type Ia (though some argue that the latter was type II).

The Ib variety is probably produced by core collapse in massive hydrogen-poor Wolf-Rayet stars. White dwarfs in double systems are obvious candidates for type Ia. The more massive the receiving white dwarf, the more quickly the infalling fresh hydrogen reaches the nova flash point. If massive enough, the interval between successive novae becomes less than the time over which they have been observed, and the novae become recurrent. For example, RS Ophiuchi and its cohort produce novae every 20 years or so. Each nova leaves a little residual matter, so the white dwarf increases in mass. If it is pushed over the Chandrasekhar limit, the star must catastrophically collapse and burn, producing a type Ia supernova. A second scenario that might produce type Ia supernovae is the merger of the stars in a double white dwarf system in which the sum of the masses exceeds the Chandrasekhar limit. Type Ia supernovae are even brighter than core-collapse type II events, and can reach $M_V = -19$. They also deposit more iron into the cosmos than do the type II, typically three-tenths of a solar mass. All the iron in the universe is made by supernovae. *See* WOLF-RAYET STAR.

Though the origins of type Ia explosions are not very well understood, the supernovae have such predictable absolute magnitudes that they can be used to measure distances to distant galaxies and determine not just the Hubble constant but also more subtle expansion characteristics of the universe.

Heavy-element enrichment. Many of the above scenarios are still shrouded in mystery. Not all the doubles have all the behavior patterns described, and there are great numbers of categories of each kind of phenomenon. The overall result of stellar evolution is clear, however. As stars become giants that turn into white dwarfs, or as they become supergiants that create neutron stars or black holes, they feed huge quantities of enriched matter back into the star-generating clouds of interstellar space. Later generations of stars therefore have more heavy elements than do earlier generations. Computer modeling of the chemical composition of the Galaxy, starting with the hydrogen and helium of the big bang, can closely replicate the chemical composition of the Sun, providing powerful evidence that the theories are correct. The Earth, made almost entirely of heavy elements, is a distillate of solar gases, and exists only as a result of the combined action of the stars.

James B. Kaler

Bibliography. L. H. Aller, *Atoms, Stars, and Nebulae*, 3d ed., Cambridge University Press, 1991; E. Böhm-Vitense, *Introduction to Stellar Astrophysics*, 3 vols., Cambridge University Press, 1989, 1992; J. B. Kaler, *Stars*, Freeman, New York, 1992, paperback, revised 1998; J. B. Kaler, *The Cambridge Encyclopedia of Stars*, Cambridge University Press, 2006; R. J. Tayler, *The Stars: Their Structure and Evolution*, 2d ed., Cambridge University Press, 1994.

Stellar magnetic field

A magnetic field, far stronger than the Earth's magnetic field, which is possessed by many stars. Magnetic fields are important throughout the life cycle of a star. Initially, magnetic fields regulate how quickly interstellar clouds collapse into protostars. Later in the star formation process, circumstellar disk material flows along magnetic field lines, either accreting onto the star or flowing rapidly out along the rotation axis. Outflowing material (stellar winds) carries away angular momentum, slowing rotation at a rate that depends on stellar magnetic field strength. On the Sun, dark sunspots, prominences, flares, and other forms of surface activity are seen in regions where there are strong magnetic fields. There is some evidence that long-term variations in solar activity may affect the Earth's climate. Turbulence in the solar atmosphere drives magnetic waves which heat a tenuous corona (seen during eclipses) to millions of degrees. About 10% of hotter stars (with temperatures of about 10,000 K) with stable atmospheres are Ap stars, which have stronger magnetic fields that control the surface distribution of exotic elements. Even after stars end their internal fusion cycle and become compact remnants, magnetic fields channel accreting material from binary companions, occasionally producing spectacular novae. Despite the enormous gravity around pulsars, magnetic forces far exceed gravitational forces, creating intense electromagnetic beams that spin down (slow the rotation of) the pulsar. *See* BINARY STAR; CATACLYSMIC

VARIABLE; NOVA; PROTOSTAR; PULSAR; STELLAR EVOLUTION.

Measurement techniques. Stellar magnetic fields are measured remotely using optical, infrared, and radio spectra. Magnetic fields alter atomic and molecular energy levels via the Zeeman effect, causing spectral lines to split and shift relative to their nonmagnetic locations. The size of the shift depends directly on wavelength and magnetic field strength. When the splitting is large compared to line width (as in sunspots, some Ap stars, and magnetic white dwarfs), the location and relative strengths of individual components yield both field strength and orientation. When the splitting is too small to resolve, measurements of spectral line polarization (polarimetry) can sometimes put limits on magnetic field strength, but only if the field geometry is simple (as in Ap stars). For distant stars that resemble the Sun, the polarization signal is small because the surface as a whole contains nearly equal amounts of oppositely directed magnetic field. In this case, the only recourse is to look for subtle enhancements of spectral line width or strength due to the Zeeman effect. *See* ASTRONOMICAL SPECTROSCOPY; POLARIMETRY; ZEEMAN EFFECT.

Observations. Broadening of spectral lines in sunspots was first observed by J. Norman Lockyer in 1866, but the magnetic origin of this broadening was not understood until after the theoretical work of Pieter Zeeman in 1897 and the first polarimetry by George Ellery Hale in 1908. Field strengths are typically 0.3 tesla (3 kilogauss) in sunspots. Zeeman broadening implies similar field strengths for several other cool stars, but current stellar measurement techniques are effective only if at least 5% of the surface contains strong magnetic fields (compared to 0.5% for the Sun). For cool stars, polarimetry is useful only if specific areas with uniform magnetic field orientation can be localized (through spatial maps of the Sun, Doppler images of rapidly rotating stars, and emission from flares or accretion sites). Ap stars have magnetic fields as strong as 3.3 T (33 kG), organized into global dipoles that are often offset from the center of the star and tilted with respect to the rotation axis. For reference, the Earth has a dipole magnetic field that is only 6×10^{-5} T (0.6 G) at the pole. Finally, dense stellar remnants can have tremendous fields, reaching up to 10^4 T (10^8 G) for white dwarfs and 10^8 T (10^{12} G) for neutron stars. *See* GEOMAGNETISM; NEUTRON STAR; SOLAR MAGNETIC FIELD; WHITE DWARF STAR.

Origin and dynamics. Large magnetic structures on the Sun vary in location, number, and magnetic polarity with an average cycle period of 22 years. Similar cycles have been detected on several other cool stars with turbulent (convective) atmospheres. In these stars, magnetic heating of coronae increases with rotation rate. Such behavior is expected of hydromagnetic dynamos, but details are sketchy and more than one type of dynamo may operate in the Sun. In warmer or more compact stars with convectively stable envelopes, magnetic fields are assumed to be concentrated fossils from earlier stages of evolution. Ultimately, the magnetic fields in these stable stars are attributed to the original interstellar gas out of which the stars formed. *See* MAGNETOHYDRODYNAMICS; STAR; SUN.

Jeff A. Valenti

Bibliography. G. Chanmugam, Magnetic fields of degenerate stars, *Annu. Rev. Astron. Astrophys.*, 30:143–184, 1992; I. Rüedi et al., Magnetic field measurements on moderately active cool dwarfs, *Astron. Astrophys.*, 318:429–442, 1997; J. C. del Toro Iniesta, On the discovery of the Zeeman effect on the Sun and in the laboratory, *Vistas Astron.*, 40:241–256, 1996.

Stellar population

One of the categories into which stars may be classified, based on their based on their spatial distribution and dynamics within a galaxy and their epoch of formation. The stellar component of the Milky Way Galaxy consists of three populations: the thin disk, the thick disk, and the halo. The thin disk, originally referred to as population I, is the youngest, and includes stars currently forming as well as older stars such as the Sun. It is located amid most of the molecular and cold atomic gas, and confined to a height of order 1 kiloparsec (3.3×10^3 light-years or 3.1×10^{16} km or 1.9×10^{16} mi) from the galactic plane. The bulge and halo, roughly corresponding to the original population II, is a much older, far more extended structure, having an approximately spherical distribution with a scale length of order 3.5 kpc.

Development of concept. The existence of different populations in the stellar component of external galaxies was announced by W. Baade in 1944. During the wartime blackouts of Los Angeles, he was able to take the first deep multiwavelength photographs of the Andromeda Galaxy, M31. He found that the blue stars are confined to the disk of that galaxy, while the redder stars whose colors are more like the Sun or the globular clusters are more concentrated toward the galactic center and nuclear bulge. Baade referred to these stellar components as populations and labeled the younger blue disk stars population I. Those in the halo and nuclear spheroid, identified with the RR Lyrae variables (also known as cluster variables because of their preponderance in globular clusters), were called population II. A hypothetical population III has also been proposed to represent the first stars formed during the collapse of a galaxy. Some evidence for these stars is found in the extremely high abundances of *r*-process elements found in stars in which the overall abundance of heavy elements (elements other than hydrogen and helium) is less than 10^{-4} that of the Sun. Further kinematic studies of space motions demonstrated the distinctness of the two principal populations and accounted for their different distributions but left open the interpretation of these effects. *See* ANDROMEDA GALAXY; NUCLEOSYNTHESIS; VARIABLE STAR.

The population concept is now central to the understanding of chemical evolution of the Milky Way Galaxy. The globular clusters, the prototypical old stellar objects, have main-sequence turn-off ages of

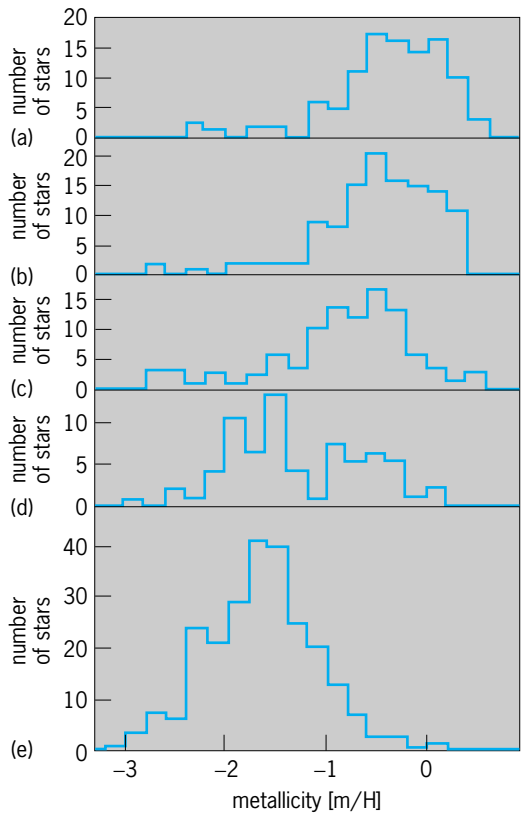

The metallicity distribution of stars as a function of velocity *V* perpendicular to the galactic plane. (Negative values of *V* indicate motion toward the plane.) (a) -30 km/s $\leq V \leq -10$ km/s. (b) -60 km/s $\leq V \leq -30$ km/s. (c) -100 km/s $\leq V \leq -60$ km/s. (d) -150 km/s $\leq V \leq -100$ km/s. (e) $V \leq -150$ km/s. (*After B.W. Carney, D. W. Latham, and J. R. Laird, A survey of proper-motion stars, VIII. On the galaxy's third population, Astron. J., 97:423-430, 1989*)

at least 10^{10} years and are among the most metal-deficient stars in the Milky Way Galaxy. These are widely distributed throughout the system and serve as markers in the halo to the mass distribution. On the other hand, the OB associations contain many massive stars with ages not greater than 10^7 years and have the highest metal abundances, often as much as a factor of 2 higher than the Sun. (At an age of about 4.6×10^9 years, the Sun was formed as a thin disk star as well.) The subdwarfs, the metal-deficient main-sequence stars, are a systematically older group of stars than those found confined to the thick galactic disk. The **illustration** shows how the metallicity distribution of stars decreases with increasing magnitude of velocity perpendicular to the galactic plane. Here, the metallicity [m/H] is the logarithmic difference in the overall abundance of heavy elements with respect to the Sun. For example, in a star with [m/H] $= 0$, this abundance is equal to that of the Sun, while in a star with [m/H] $= -3$ it is 10^{-3} that of the Sun. *See* STAR CLUSTERS.

Populations in galaxies. Fundamentally, the presence of these distinct stellar populations demonstrates that the Milky Way Galaxy has undergone an extended period of active star formation. The spheroid population is quite similar to elliptical galaxies, which are assumed to have ended active star formation many billions of years ago. Starburst galaxies give a counterexample of the fixity of the population distinction, however, since they may have completed their primary star-forming phase long ago, but the injection of a new mass of gas through collisions or accretion in a large-scale flow can produce a dramatic turn-on of the star formation. There is also increasing evidence that the Galactic populations may be contaminated by accretion of other, lower-mass satellite galaxies within the Local Group whose star-forming histories differed from that of the Milky Way. *See* GALAXY, EXTERNAL; GALAXY FORMATION AND EVOLUTION; STARBURST GALAXY.

Characteristics of populations. The thin disk is the youngest component of the Galactic stellar population. Still actively forming massive stars from molecular clouds, it is confined to within about 0.35 kpc of the plane. All of the stars have metallicities lying between about one-fifth and twice the solar value (illus. *a*), similar to that of the interstellar medium diffuse gas, and star formation appears to have remained relatively constant in this population for about the past 8×10^9 years, although there have been some excursions of a factor of order 50% during that time. One reason for the relatively small thickness of the disk is the low velocity dispersion of the component stars, about 22 km/s (14 mi/s); their systematic motion is completely dominated by the differential rotation of the disk. Since the most massive stars are also the youngest, having the shortest nuclear lifetimes, these hot stars are found associated with H II regions and OB associations as well as in open clusters. Massive stars are uniquely formed in the thin disk, supporting the attribution of youth to the population. Finally, type II supernovae, the death throes of massive stars, show a thin-disk distribution throughout the Milky Way Galaxy, and there is an indication that the remnants of massive stars, the pulsars (neutron stars with strong magnetic fields), are also confined to the thin disk. *See* INTERSTELLAR MATTER; MOLECULAR CLOUD; PULSAR; SUPERNOVA.

The thick disk is an older population, approximately 9–10 $\times 10^9$ years, roughly corresponding to the range between what was once called population II and population I. Its metallicity lies between about one-tenth and one-third of the solar value. The stars in this population are distributed over greater distances from the plane, up to 1.5 kpc, and have correspondingly larger velocity dispersion, about 45 km/s (28 mi/s). This population also includes globular clusters and subdwarfs that overlap at the lowest end of the abundances with the properties of the halo globulars, although the system of old disk globulars is distributed differently than those of the halo. Type I supernovae are associated with this population. In the Milky Way Galaxy the spatial structure of this population roughly resembles that of an E7 elliptical galaxy.

Finally, enclosing the disk and the nuclear spheroidal bulge, there is a halo that extends to considerable distances from the plane, with some stars as distant as 30 kpc. This population has an age of order 10–15 $\times 10^9$ years and a scale height of order 3.5 kpc or greater. The stars in this region have very large velocity dispersions, about 130 km/s (80 mi/s),

and do not appear to participate in the differential rotation as much as other stars. Their metallicities are all lower than about one-twentieth that of the Sun and may extend down to 10^{-3} of the solar value (illus. *e*). The most metal-poor globular clusters belong to this population. This stellar halo is not the same as the so-called dark matter halo, but is probably embedded within it. *See* DARK MATTER.

Delineation of populations. The population paradigm has become progressively more blurred with time, as the complex history of Galactic star formation has become clearer, and there has been considerable debate over whether any of these populations are truly distinct. Indeed, the separation based on elemental abundances renders this question somewhat circular since these abundances are already embedded in the definition of the membership in the different populations. Nonetheless, there are some distinct differences in these populations that permit their delineation. For instance, the globular clusters, although separated into halo and disk systems, are distinctly older than any of the clusters found in the disk. The oldest disk open clusters, NGC 188 and M69, show abundances nearly as low as the most metal-rich globulars like 47 Tucanae, but are younger by nearly 3×10^9 years than the youngest of the disk globulars. The most metal-deficient globulars are considerably older than any of the stars in the disk and much lower in abundance. *See* MILKY WAY GALAXY; STAR.

Steven N. Shore

Bibliography. R. Cayrel et al., First stars V: Abundance patterns from C to Zn and supernova yields in the early Galaxy, *Astron. Astrophys.*, 416:1117–1138, 2004; K. Freeman and J. Bland-Hawthorn, The new Galaxy: Signatures of its formation, *Annu. Rev. Astron. Astrophys.*, 40:487–537, 2002; A. Helmi et al., Pieces of the puzzle: Ancient substructure in the galactic disc, *Mon. Not. Roy. Astron. Soc.*, 365:1309–1323, 2006; B. Nordström et al., The Geneva-Copenhagen survey of the solar neighbourhood: Ages, metallicities, and kinematic prpperties of ~14,000 F and G dwarfs, *Astrom. Astrophys.*, 418:989–1019; A. Renzine, Stellar population diagonostics of elliptical galaxy formation, *Annu. Rev. Astron. Astrophys.*, 44:141–192, 2006.

Stellar rotation

The spinning of stars, due to their angular momentum. Stars do not necessarily rotate as solid bodies, and their angular momentum may be distributed nonuniformly, depending on radius or latitude. However, it is nearly impossible to resolve the surfaces of stars, or see their interiors, and the limited ability to observe them means that their rotation is generally expressed as a single number, $v_{eq} \sin i$. In this measured quantity, v_{eq} is the star's equatorial rotational velocity (in kilometers per second), and i is the angle between the star's rotation axis and the line of sight to the star. In other words, $v_{eq} \sin i$ is the component of a star's rotation that is projected onto the line of sight between the star and the observer. Measurements of $v_{eq} \sin i$ in stars range from as little as 1 km s^{-1} (0.6 mi s^{-1}) up to 400 km s^{-1} (250 mi s^{-1}) or more. A more physically useful measure of rotation is Ω, a star's angular velocity, or P_{rot}, the rotation period (the inverse of Ω). In some cases it is possible to measure P_{rot} directly. *See* ANGULAR MOMENTUM; STAR.

A $v_{eq} \sin i$ value is determined from the breadth of absorption lines in the star's spectrum. The Doppler effect causes the lines to be broadened because one limb of the star is receding and the other approaching. As a reference, the Sun is a very slowly rotating star ($v_{eq} \sin i = 1.8$ km s^{-1} or 1.1 mi s^{-1}), and solar absorption lines have a full width at half maximum (FWHM) of about 7 km s^{-1} (4 mi s^{-1}), which arises from thermal motions, the physics of line formation, and turbulent motions in the Sun's atmosphere. This means that if a star's rotation rate exceeds 15–20 km s^{-1} (9–12 mi s^{-1}), then rotation dominates the broadening of the lines and other effects need not be considered, in general. To determine smaller values of $v_{eq} \sin i$ reliably, it is necessary to observe absorption lines carefully and to then fit them with models that take into account all the broadening mechanisms. By doing this it is possible to measure $v_{eq} \sin i$ down to the Sun's level and below. *See* DOPPLER EFFECT; SUN.

Late-type stars (cooler than about 6500 K or 11,200°F) exhibit spots on their surfaces analogous to sunspots, and these can be large enough to produce observable changes in the light of the star. In such cases it is possible to measure P_{rot} directly and, in a few instances, changes in P_{rot} have led to estimates of differential rotation (the dependence of Ω on latitude). Differential rotation is critical for the mechanisms that generate magnetic fields in low-mass stars. Rotation periods are independent of aspect and so provide a true indication of the intrinsic spread of rotation within star clusters, for instance. The high quality of P_{rot} measurements means that more is now known about the rotation in low-mass stars than is known for massive stars, the opposite of the case a generation ago. *See* STAR CLUSTERS; STELLAR MAGNETIC FIELD.

Stars get their angular momentum from the material from which they form, and one of the key problems in star formation is understanding how a cloud of material in the interstellar medium can contract enough to become a star. This is because a tiny degree of spin in the natal cloud leads to huge rotation rates by the time the material is dense enough to be a star. Stars appear to regulate their rotation in these early phases by interacting with the disk of material that surrounds them (the debris left over from star formation) and through jets, although this is an active area in current research. Stars are observed to have their rotation axes distributed randomly, which means that they do not acquire their rotation from the rotation of the Milky Way Galaxy but from turbulent motions in the parent cloud. This is confirmed by noting that the angular momentum axes of binary stars are also oriented randomly and that there is no tendency for stars in clusters to have rotation axes that align with the Galaxy's. *See* BINARY STAR; PROTOSTAR; STELLAR EVOLUTION.

Each type of star has its own rotational characteristics, and there is a broad trend in which typical rotation rates decline with the mass of the star in going down the main sequence. There is an especially prominent break in rotation such that stars with masses below about 1.2 times the solar mass have very low rotation rates, due to their intrinsic structure, as explained below.

Early-type stars. The early-type stars (types O, B, A, and early-F) typically have $v_{eq} \sin i$ values of 50 km s^{-1} (30 mi s^{-1}) or more, and so their rotation is easily discerned on moderate-resolution spectrograms. For this reason, and because the early-type stars were easier to study with photographic plates, they were the first to have their rotational properties examined in detail, and most of the extant data dates from the photographic era. The most massive stars (of types O and B) sometimes exhibit nonradial pulsations and strong winds. They also can have surface abundance anomalies that result from internal mixing. All these effects are related to rotation, and they can complicate the interpretation of the observations in very massive stars. At the same time, rotation rates in O and B stars can approach the breakup velocity, reducing the effective gravity in the star and thereby affecting the star's evolution. *See* ASTRONOMICAL SPECTROSCOPY; SPECTRAL TYPE.

For the O stars, the distribution of $v_{eq} \sin i$ has two peaks, centered at about 100 and 300 km s^{-1} (60 and 190 mi s^{-1}). The highest velocities seen are near 400 km s^{-1} (250 mi s^{-1}), and the high $v_{eq} \sin i$ values (200 km s^{-1} or 125 mi s^{-1} and more) are seen only in the main-sequence O stars. These very high rates extend into the Be stars. O stars can also have significant turbulence (20 km s^{-1} or 12 mi s^{-1}), and so there are none with sharp-lined spectra. O stars have strong radiation-driven winds and significant mass loss so that these stars can lose angular momentum during their main-sequence lifetimes.

The highest rotation rates can reduce the central pressure within a star; this can lead to slightly lower luminosity (by as much as 3%). The shape of the star can also change so that the surface temperature changes with latitude. A more significant effect is that rapid rotation influences mixing within the star and transports angular momentum. In particular, higher rotation leads to a higher degree of turbulent diffusion that can double a star's luminosity in some cases.

The B stars have an average $v_{eq} \sin i$ of about 140 km s^{-1} (87 mi s^{-1}), but the Be stars (B stars with emission lines, comprising about 15–20% of the B stars) rotate at significantly higher rates (the average v_{eq} for Be stars, corrected for inclination, is 265 km s^{-1} or 165 mi s^{-1}). Even this high rate is still about a factor of 2 below the breakup velocity (500–550 km s^{-1} or 310–340 mi s^{-1}).

The A and early-F stars have moderate rotation rates (50–100 km s^{-1} or 30–60 mi s^{-1}) except for those exhibiting peculiar chemical abundances. The Am, or metallic-line A stars, and the Ap, or magnetic A stars, rotate at an average rate of about 40 km s^{-1} (25 mi s^{-1}), a factor of 3–4 lower than normal stars of similar colors. Only a very few chemically peculiar stars have $v_{eq} \sin i$ exceeding 100 km s^{-1} (60 mi s^{-1}). Also, in these peculiar stars rotation declines with age, but the relationship is neither well delineated nor understood. The opposite is also true: very few normal A stars have low $v_{eq} \sin i$ values.

Late-type stars. Late-type stars are cooler than about 6500 K (11,200°F) and have spectral types of late-F, G, K, or M. The key property of late-type stars that determines much of their phenomenology is the presence of a convective envelope. In such stars, rotation (especially differential rotation) interacts with the convection to produce complex motions in the electrically conductive plasma of the star. These circulatory patterns enable the star to regenerate a seed magnetic field, the so-called dynamo mechanism. The magnetic field can grip an ionized wind beyond the star's surface, leading to gradual angular momentum loss, a process that is seen occurring on the Sun. More rapidly rotating stars produce stronger magnetic fields, as evidenced by spectroscopic indicators of magnetic-related activity, such as emission in the cores of the Ca II H and K lines. There is, therefore, a strong correlation between rotation and activity, in the various forms that activity is observed (Ca II H and K, Hα, x-rays, and so forth). These activity indices are often observed to measure P_{rot} for stars because they vary more pronouncedly than the star's light overall. *See* MAGNETISM; STELLAR MAGNETIC FIELD.

In this scenario, late-type main-sequence stars spin down over their lifetimes, and that is observed to occur because steadily declining rotation rates can be seen among stars in clusters of increasing age. Moreover, stars that are born together but with different rotation rates will tend to end up with the same P_{rot} because the faster-rotating star will lose angular momentum more quickly. Young stars tend to rotate rapidly because they have not had time to lose the angular momentum with which they formed, while old low-mass stars rotate slowly.

Young late-type stars are mostly observed in clusters because there are few of them elsewhere, and because clusters provide rich samples of the same age. The Pleiades, for example, is 100 million years old, meaning that its low-mass stars have only recently reached the zero-age main sequence (ZAMS). Among the solar-type stars of the Pleiades, an intrinsic spread in rotation of more than a factor of 20 is seen at any one mass, with the most rapidly rotating stars spinning at 100 times the solar rate. These ultrafast rotators are a minority constituent of the cluster (about 20% of the stars), but their presence along with many slowly rotating stars is difficult to account for in models of the angular momentum evolution of low-mass stars. It has been speculated that some of these young stars may experience a decoupling of the radiative core from the convective envelope, with the core continuing to spin at a high rate for an extended time. The α Persei cluster is also at the zero-age main sequence, with an age of 50 million years, and its stars confirm that what is seen in the Pleiades is typical for young low-mass stars. By the time that stars become as old as the Hyades (age 600 million years), their rotation rates have largely converged and show little spread at any

one mass, although there is a strong mass dependence. *See* HYADES; PLEIADES.

The most rapidly rotating low-mass stars seen outside clusters spin at about 10 times the solar rate, but the vast majority, being old like the Sun, rotate very slowly (at the solar level or below). Evolved late-type stars (giants and supergiants) rotate at very low rates, in part because they have swelled so much since leaving the main sequence. A few rare exceptions spin rapidly, probably because they acquired the orbital angular momentum of their companion when a binary system coalesced. *See* GIANT STAR; SUPERGIANT STAR.

Very low mass stars and brown dwarfs. Significant numbers of very low mass stars and brown dwarfs have been found and studied at high spectroscopic resolution. These objects are fully convective from their surfaces all the way to their cores, in contrast to the Sun and similar stars, which have a convective envelope. This change in structure corresponds to a change in the observed rotation in that very low mass stars and brown dwarfs appear not to lose angular momentum and so continue to spin rapidly. These objects are very small in size, so the rotation rates seen—15 to 60 km s^{-1} (10 to 40 mi s^{-1})—imply high angular velocities of as much as 100 times that of the Sun. *See* BROWN DWARF.

Presence of planets. There is no relationship between the rotation of a star and the presence of planets around it. It is true that planets have been found primarily around old, slowly rotating stars like the Sun, but that is in part because such stars have the very sharp absorption lines needed to measure the reflex motion to the necessary high precision. Young solar-type stars have broader lines that prevent reaching the necessary precision, and they also exhibit effects related to their activity that add noise to the velocity measurements. *See* EXTRASOLAR PLANETS.

Solar rotational history. The rotational history of the Sun is not known in detail, but the broad features of that history can be delineated from studying stars of solar mass of many ages. The Sun formed from interstellar material, most probably in concert with many other stars at the same time. For about the first 10 million years of its life, it kept a disk of debris material around it, and this formed into planets before about 50 million years was reached. At an age of 10 million years the Sun rotated at two or three times its present angular velocity, and it was well above the main sequence. The Sun reached the zero-age main sequence when it was about 50 million years old, at which point it could have rotated at a rate of 3 to 100 times the current solar rate, although it probably was at the low end of this range. Over its 4.5-billion-year main-sequence lifetime, the Sun has gradually lost angular momentum to get to the rate that is seen today. Helioseismological studies of the Sun's interior show it to rotate as a solid body, so if it ever had a rapidly rotating core, it does no longer. In about 5 billion years the Sun will exhaust most of the hydrogen in its core and leave the main sequence to become a giant star. It will bloat to 20 or more times its present size, and conservation of angular momentum will mean that its rotation rate will become extremely small.

The Sun remains the standard of comparison for other stars. The rotation of the Sun was among Galileo's first discoveries as he traced the motions of spots across the surface, and only a little later the different rates of rotation at different solar latitudes led to the fundamental notion that the Sun is a fluid, not a glowing solid. Angular momentum is central to the formation and evolution of stars, and work now taking place will lead to a fuller picture of what happens in stars and how the Sun fits into that picture.

David Soderblom

Bibliography. D. F. Gray, *The Observation and Analysis of Stellar Photospheres*, 3d ed., 2005; A. Maeder and P. Eenens (eds.), *Stellar Rotation*, IAU Symp. 215, Astronomical Society of the Pacific, San Francisco, 2004 ; R. Pallavicini, G. Michela, and S. Sciortino (eds.), *Stellar Clusters and Associations: Convection, Rotation, and Dynamos*, Astronomical Society of the Pacific, San Francisco, 2000; J.-L. Tassoul, *Stellar Rotation*, Cambridge University Press, 2000.

Stem

The organ of vascular plants that usually develops branches and bears leaves and flowers. On woody stems a branch that is the current season's growth from a bud is called a twig. The stems of some species produce adventitious roots. *See* ROOT (BOTANY).

General Characteristics

While most stems are erect, such as aerial structures (**Fig. 1***a*), some remain underground (**Fig. 2***a–b*), others creep over or lie prostrate on the surface of the ground (Fig. 1*c–e*), and still others are so short and inconspicuous that the plants are said to be stemless, or acaulescent (Fig. 1*b*). When stems lie flattened immediately above but not on the ground, with tips curved upward, they are said to be decumbent, as in juniper (Fig. 1*c*). If stems lie flat on the ground but do not root at the nodes (joints), the stem

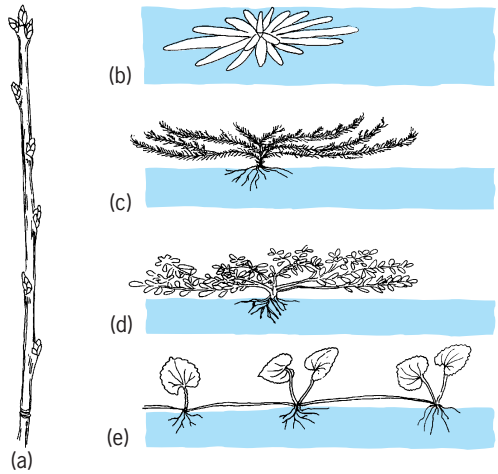

Fig. 1. Kinds of stems. (*a*) Woody (plum). (*b*) Acaulescent (evening primrose). (*c*) Decumbent (juniper). (*d*) Procumbent (purslane). (*e*) Repent (ground ivy).

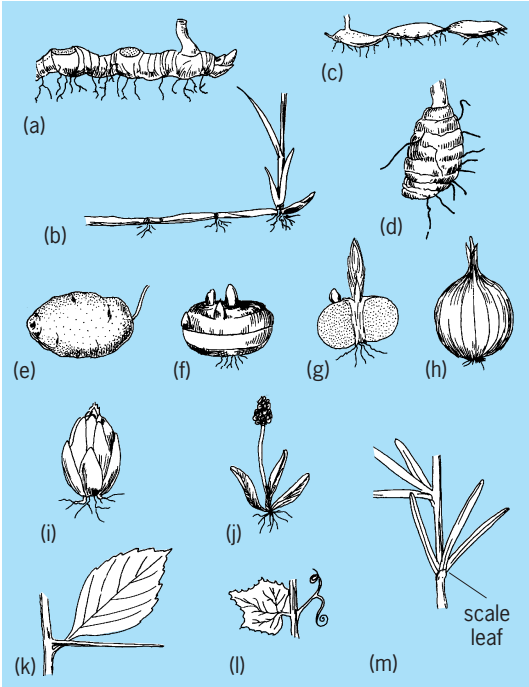

Fig. 2. Specialized stems. (*a*) Rhizome (Solomon's seal). (*b*) Rhizome (grass). (*c*) Moniliform rhizome (pepperroot). (*d*) Caudex (trillium). (*e*) Tuber (potato). (*f*) Corm (crocus). (*g*) Section of corm. (*h*) Tunicated bulb (onion). (*i*) Scaly bulb (lily). (*j*) Scape (English plantain). (*k*) Thorn (hawthorn). (*l*) Tendril (grape). (*m*) Cladode (asparagus).

is called procumbent or prostrate, as in purslane (Fig. 1*d*). If a stem creeps along the ground, rooting at the nodes, it is said to be repent or creeping, as in ground ivy (Fig. 1*e*).

Shape and texture. Most stems are cylindrical and tapering (terete), appearing circular in cross section; others may be quadrangular, as in mints, or triangular, as in some sedges.

Herbaceous stems (annuals and herbaceous perennials) die to the ground after blooming or at the end of the growing season. They usually contain little woody tissue. Woody stems (perennials) have considerable woody supporting tissue and live from year to year. A woody plant with no main stem or trunk, but usually with several stems developed from a common base at or near the ground, is known as a shrub. Suffrutescent stems are intermediate between herbaceous and shrubby and become partly woody and perennial at the base, as in teaberry.

Specialized stems. The stems of some plants are highly modified in various ways (Fig. 2). An underground horizontal stem is a rhizome or rootstock. It may be thickened, as in Solomon's seal; slender, as in Bermuda grass; or contracted at regular intervals, as in pepperroot, in which case it is a moniliform rhizome. A caudex is an upright perennial underground stem; it is much like a rhizome but grows vertically, as in trillium. The term caudex is applied also to the trunks of palm trees which grow in a similar manner but aboveground. Tubers are enlarged ends of rhizomes in which food accumulates, as in white potato. A corm is a short, erect, fleshy subterranean stem usually broader than high and often covered with dry membranous coats, as in crocus. A bulb may be regarded as a short, subterranean stem with many overlapping fleshy leaf bases or scales, as in onion. A scape is a leafless flowering stem arising from the ground, as in plantain or dandelion. A thorn is a rigid, sharp-pointed modified branch of a woody plant, as in hawthorn. A tendril is a slender coiling structure capable of twining about a support to which the plant is then attached, as in grape. A cladophyll, or phylloclade, is a usually flattened stem resembling a leaf that arises in the axil of a minute leaf (scale), as in asparagus. In some instances the resemblance of a cladophyll to a leaf is noted only or mainly in its green color, as in certain species of cacti. *See* PLANT ORGANS.

Nelle Ammons

External Features

A shoot or branch usually consists of a stem, or axis, and leafy appendages. Stems have several distinguishing features. They arise either from the epicotyl of the embryo in a seed or from buds. The stem bears both leaves and buds at nodes, which are separated by leafless regions or internodes, and sometimes roots and flowers. A woody twig (**Fig. 3**) shows leaf scars and lenticels on its surface. Eventually, secondary growth thickens the stem axis and obscures its division into nodes and internodes and the relation between buds and leaves. *See* LEAF.

Nodes. The nodes are the regions of the primary stem where leaves and buds arise. The number of leaves at a node is usually specific for each plant species. In deciduous plants which are leafless during winter, the place of former attachment of a leaf is marked by the leaf scar. The scar is formed in part by the abscission zone formed at the base of the leaf petiole. Small bundle scars left by the broken ends of the vascular (conducting) tissues between the stem and leaf are seen within the leaf scar. Flower and fruit scars, marking the points of flower and fruit attachment, may also be visible.

Internodes. The stem regions between nodes are called internodes. Internode length varies greatly among species, in different parts of the same stem, and under different growing conditions. One or many internodes may grow rapidly during stem elongation. Generally, an individual internode matures

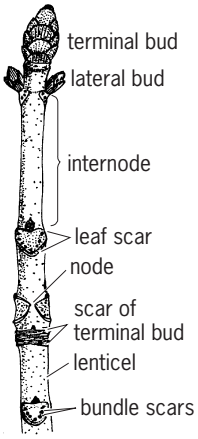

Fig. 3. Winter woody twig (horse chestnut) showing apical dominance. (*After E. W. Sinnott and K. S. Wilson, Botany: Principles and Problems, 5th ed., McGraw-Hill, 1955*)

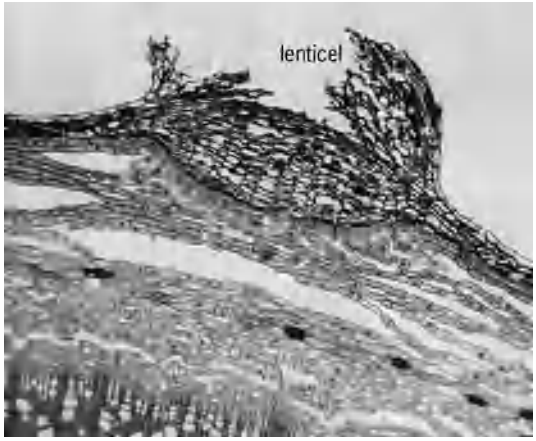

Fig. 4. Portion of a cross section of a young stem of elder (*Sambucus canadensis*) showing a lenticel. (*G. H. Conant, Triarch Products*)

from the base upward in dicotyledons and many monocotyledons. Conversely, internodes of some monocotyledons (grasses and sedges) and horsetails (*Equisetum*) have prolonged growth from intercalary meristems located at the base of each internode.

Lenticels. There are small, slightly raised or ridged regions of the stem surface that are composed of loosely arranged masses of cells in the bark (**Fig. 4**). Their intercellular spaces are continuous with those in the interior of the stem, therefore permitting gas exchange similar to the stomata that are present before bark initiation. The size and shape of lenticels vary among species and with age they are absent in a few species. *See* PERIDERM.

Bud scales. In most temperate, perennial plants, buds are protected by a series of modified leaves or bud scales. Wax or hairs on their surface further protect the apical meristem within the bud. The buds of many tropical trees are often protected by the bases of mature leaves or simply by the expanding, immature leaves. Such naked buds are also present in most herbaceous plants.

Most temperate trees form a terminal bud at the end of a season's growth. The cluster of scars that mark the position of the closely packed terminal bud scales also delimit the segment of stem produced during one growing season (Fig. 3). The age and yearly growth of a twig can be determined by observing the number of sets of these terminal bud scale scars and the distances between them.

Bud types. Buds are partly developed shoots or undifferentiated masses of meristematic cells. They may be classified as terminal, lateral, and adventitious. *See* BUD.

Terminal or apical buds. These are buds at the tips of branches, often the largest buds on the plant, which usually contribute to the growth in length of the stem bearing them.

Lateral buds. These are buds located on the sides of the stem and are initiated within the terminal bud from superficial tissues at the nodes. They are further classified according to their position. Most plants have a bud in the axil of each leaf directly above the leaf (or leaf scar). Such axillary buds usually remain dormant for some time. If the axillary bud develops into a branch, it becomes organized into a new terminal bud which itself produces new lateral buds. In some species, additional buds are produced above or on either side of the axillary bud. Such accessory or supernumerary buds may eventually develop as flower buds, thorns, or vegetative branches like the axillary bud.

Adventitious buds. These are buds which are not produced within the terminal bud and lack a definite relation to leaf position. They arise from differentiated (mature) tissues of the root, stem, or leaf, or from wound callus. The root suckers of many trees (oak, sweet gum, aspen) and troublesome weeds arise from adventitious buds. Such buds are also produced on leaf cuttings of African violet and begonia.

Bud development. Most buds of a plant remain undeveloped for indefinite periods and are known as dormant, latent, or potential buds in contrast to active or developing buds.

In many trees and shrubs, especially those with large, overwintering buds, the whole year's growth is laid down in rudimentary form in the bud during early summer. These buds expand by elongation of internodes the following spring. Some buds, especially flower buds, may form in the spring of the same season during which they expand. In temperate trees most lateral buds remain dormant for the first season, and only grow into branches the following season because of the apical dominance of the terminal bud. Such delayed bud outgrowth is characteristic of proleptic branches. In contrast, many tropical trees have sylleptic branches, in which lateral buds grow out immediately without a rest period. *See* APICAL DOMINANCE; PLANT GROWTH.

Branching. There are three major types of stem branching: dichotomous, monopodial, and sympodial. Dichotomy occurs by a division of the apical meristem to form two new axes. Such branching without axillary buds occurs in some lower vascular plants and in a few angiosperms (some palms and cacti, and *Strelitzia*). If the terminal bud of an axis continues to grow and lateral buds grow out as branches, the branching is called monopodial. If the apical bud terminates growth in a flower or dies back and one or more axillary buds grow out, the branching is called sympodial. Often only one bud develops so that what appears to be a single axis is in fact composed of a series of lateral branches that are arranged in linear sequence.

In many woody plants there is a distinction between long and short shoots, which results from the degree of internodal elongation. Short shoots are commonly the site of reproduction, and the spur shoots of apple, because of their congested nodes, bear a tight cluster of leaves.

Stem form. The large and conspicuous stems of trees and shrubs assume a wide variety of distinctive forms.

Columnar stems are basically unbranched and form a terminal leaf cluster, as in palms, or lack obvious leaves, as in cacti.

Branching stems have been classified either as excurrent, when there is a central trunk and a conical

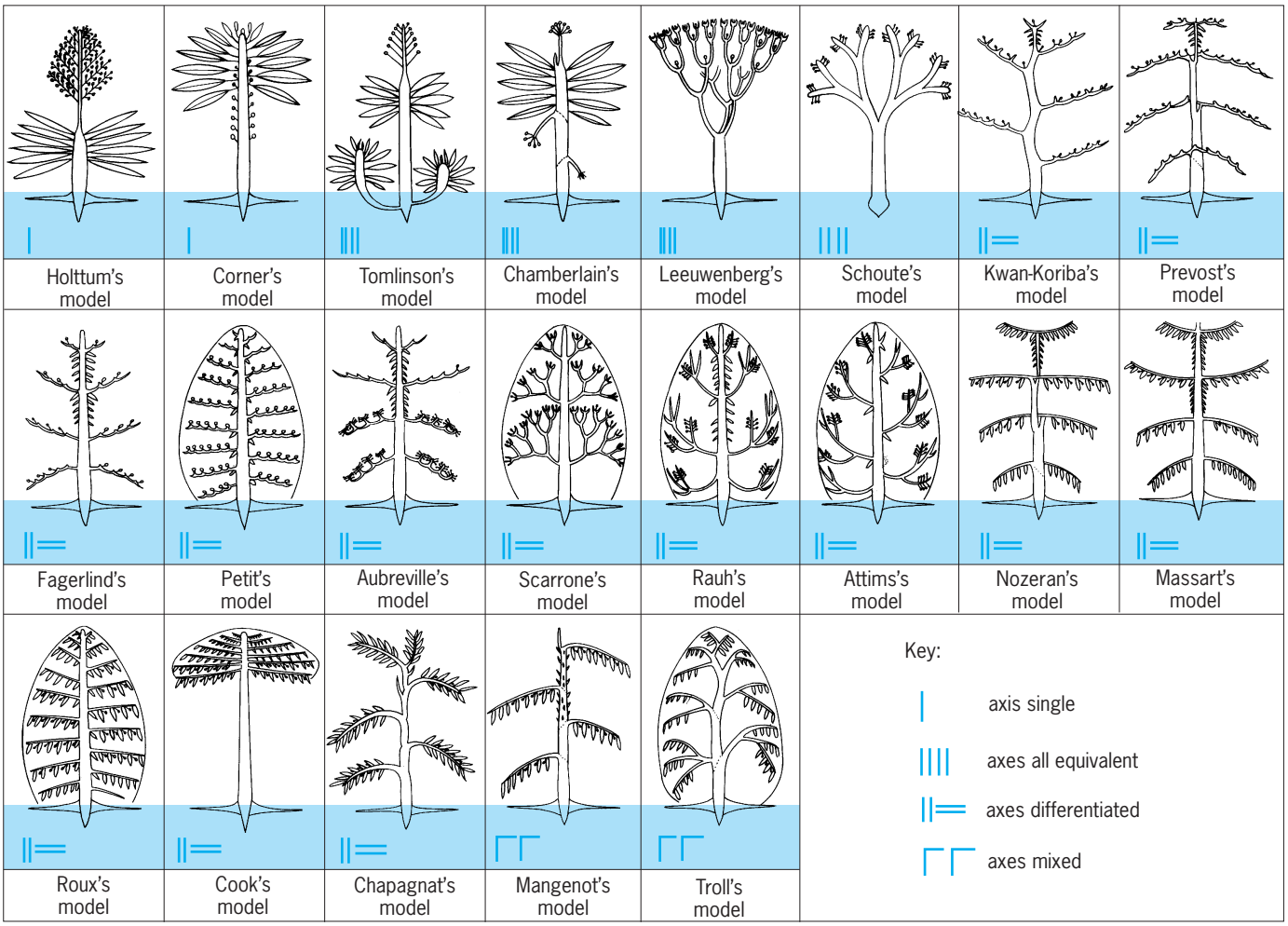

Fig. 5. Tree architecture. The idealized models of tree architecture can account for most types of branching and growth in woody plants. (*After F. Hallé and R. A. A. Oldeman, An essay on the architecture and dynamics of growth of tropical trees, Penerbit Universiti Malaya, 1975*)

leaf crown, as in firs and other conifers, or as decurrent (or deliquescent), when the trunk quickly divides up into many separate axes so that the crown lacks a central trunk, as in elm.

Another approach to the study of stem form is the concept of tree architecture which uses models that can describe most living and fossil trees and shrubs (**Fig. 5**). Each architectural model is defined by the orientation, growth sequence, and branching pattern of the trunk and lateral branches of shoot. In at least some trees (*Terminalia*), the observed branching pattern approaches the theoretical ideal for the optimum filling of space and exposure of leaf surfaces, important adaptive characteristics for stem form that have probably been evolutionarily selected for. *See* TREE.

Jack B. Fisher

Internal Features

The stem is composed of the three fundamental tissue systems that are found also in all other plant organs: the dermal (skin) system, consisting of epidermis in young stems and periderm in older stems of many species; the vascular (conducting) system, consisting of xylem (water conduction) and phloem (food conduction); and the fundamental or ground tissue system, consisting of parenchyma and sclerenchyma tissues in which the vascular tissues are embedded. The arrangement of the vascular tissues varies in stems of different groups of plants, but frequently these tissues form a hollow cylinder enclosing a region of ground tissue called pith and separated from the dermal tissue by another region of ground tissue called cortex. *See* CORTEX (PLANT); EPIDERMIS (PLANT); PARENCHYMA; PHLOEM; PITH; SCLERENCHYMA; XYLEM.

Primary and secondary tissues. Part of the growth of the stem results from the activity of the apical meristem located at the tip of the shoot. The derivatives of this meristem are the primary tissues: epidermis, primary vascular tissues, and the ground tissues of the cortex and pith. In many species, especially those having woody stems, secondary tissues are added to the primary. These tissues are derived from the lateral meristems, oriented parallel with the sides of the stem: cork cambium (phellogen), which gives rise to the secondary protective tissue periderm, which consists of phellem (cork), phellogen (cork cambium), and phelloderm (secondary

cortex) and which replaces the epidermis and vascular cambium, which is inserted between the primary xylem and pholem and forms secondary xylem (wood) and phloem (**Fig. 6**). *See* APICAL MERISTEM; LATERAL MERISTEM.

Stele. The vascular tissues and the closely associated ground tissues—pericycle (on the outer boundary of vascular region), interfascicular regions (medullary or pith rays), and frequently also the pith—may be treated as a unit called the stele. The variations in the arrangement of the vascular tissues serve as a basis for distinguishing the stelar types. The word stele means column and thus characterizes the system of vascular and associated ground tissues as a column. This column is enclosed within the cortex, which is not part of the stele. *See* PERICYCLE.

Primary State

The leafy stem of the plant is initiated as a shoot primordium in the embryo. This primordium is called the plumule, or epicotyl, and is located above the insertion of the cotyledons (seed leaves) on the hypocotyl. The epicotyl has an apical meristem that gives rise to all subsequent primary parts of the leafy stems. In the primary growth of the stem, nodes and internodes are differentiated, and the leaves appear in a characteristic arrangement (phyllotaxis). The cells derived from the apical meristem compose the meristematic precursors of the primary tissues of the stem and are called primary meristems (protoderm, ground meristem, procambium) [Fig. 6]. These meristems differentiate into the primary tissue systems.

Shoot apex or apical meristem. This so-called growing point consists of the meristematic initials and their derivatives. The organization of the apical meristem varies in different groups of plants. In the angiosperms the initials are usually arranged in two or more layers (tiers), frequently three in gymnosperms, a large group of apical initials without a tunica and with some cell stratification; and in the ferns and still lower vascular plants the shoot tip may bear a single apical initial. The organization in the angiosperms is referred to as that of tunica and corpus. The outermost tier or tiers of initials and their immediate derivatives constitute the tunica; the innermost tier and its immediate derivations compose the corpus. The tunica is a mantlelike region enclosing the corpus or core. The tunica increases only in surface by anticlinal (at right angles to the surface) divisions; the corpus shows volume growth, for its cells divide in various planes. All the initials and their first derivatives together are often designated as a promeristem. Below this promeristem are partially differentiated tissues, the primary meristems, named from the periphery inward: the protoderm; the outer ground meristem; the procambium, or provascular tissue; and the inner ground meristem. As stated previously, the primary peristems differentiate into primary tissue systems: the protoderm into the epidermis, the ground meristem into tissues of cortex and pith, and the procambium into primary vascular tissues (Fig. 6).

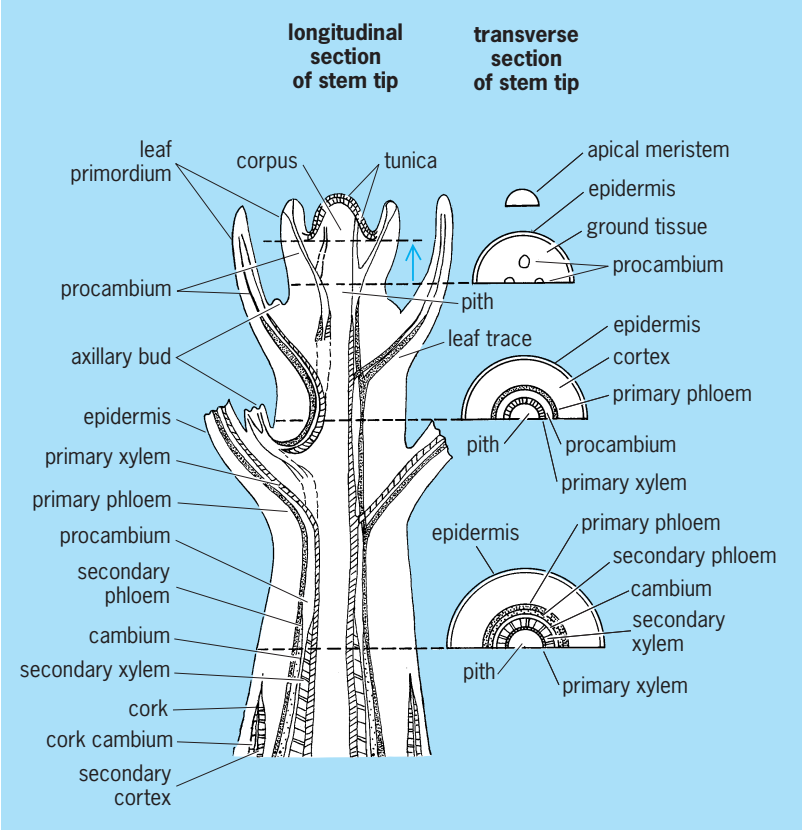

Fig. 6. Diagrammatic representation of a woody, dicotyledonous stem in transverse and longitudinal sections. It illustrates the gradual transition (arrow) from primary growth (growth in length) to secondary growth (growth in diameter). The primary plant body is derived from the apical meristem; the secondary plant body is produced by the vascular cambium and the cork cambium.

Differentiation. The complex of physiological and morphological changes that occur in a cell as it develops from a meristematic to a mature state is called differentiation. Common evidences of differentiation of cells are vacuolation (water uptake), accumulation of ergastic substances (metabolic products), formation of plastids, and increase in thickness of the cell wall and its impregnation with new wall substances. In highly specialized cells the nucleus may break down (sieve elements in the phloem), or even the entire protoplasm may be lost (water-conducting cells in the xylem). *See* PLANT CELL.

Primary tissue systems. As seen in the stem as a whole, the primary tissues form integrated systems and are referred to as the primary tissue systems: epidermis (or epidermal system), ground tissue system, and primary vascular system.

Epidermis. The epidermal system differentiates from the protoderm. The epidermis is usually composed of a single layer of cells. Occasionally, the protoderm divides periclinally (parallel to the surface) and produces a multiple epidermis. Epidermal cells may be variously modified into guard cells with stomata or the diverse epidermal appendages called hairs, or trichomes. The walls of the epidermis are cutinized, that is, impregnated with a waxy substance, and the outer walls are covered with a continuous layer of the same substance, the cuticle. *See* EPIDERMIS (PLANT).

Cortex. Between the epidermis and the vascular system is the cortex, formed by the outer ground tissue. However, there is no distinguishable cortex in plants in which the vascular bundles have a scattered arrangement, as in corn. The outer layers of the cortex frequently form collenchyma, a mechanical, or strengthening, tissue. The cortex may also include chlorenchyma (tissue containing chlorophyll). Still other cells may develop as sclereids, fibers, or laticifers. Secretory canals may be present. In water plants the cortical parenchyma may contain large air spaces and is then called aerenchyma. The outermost layers of the cortex are sometimes cutinized, like the epidermis, and are referred to as the hypodermis. Cortical parenchymal cells contain various inclusions such as starch, tannins, and crystals. *See* COLLENCHYMA; CORTEX (PLANT).

Endodermis. The innermost layer of the cortex (or possibly, in the lower vascular plants, the outermost layer of the stele) forms a histochemically distinct and sometimes a morphologically specialized layer, the endodermis. If an endodermis is morphologically distinct, it forms a compact layer and the anticlinal walls of its cells bear a suberized and lignified band, or Casparian strip. The endodermis is found in the stems and roots of lower vascular plants, the rhizomes and roots of seed plants, and a few herbaceous stems of angiosperms. It is not differentiated in the aerial stems of gymnosperms and most angiosperms. In young stems of angiosperms a homologous layer contains abundant starch and is called the starch sheath. *See* ENDODERMIS.

Pericycle. If a distinct layer or layers of cells are present between the ground-tissue endodermis and the vascular tissue, the tissue constitutes the pericycle. This tissue is interpreted as part of the ground tissue of the stele. The pericycle may consist of parenchyma or sclerenchyma. It may be seen in the stems and roots of lower vascular plants, but is usually absent in the stems of gymnosperms and angiosperms. The fibers located on the periphery of the vascular systems in many dicotyledons are often called pericycle fibers. In many species, however, these fibers originate in the phloem. *See* PERICYCLE.

Pith. The ground tissue enclosed by the vascular tissues is the medulla or pith. In stems with no cortex the pith is not distinguishable. The vascular bundles may be scattered throughout the central ground tissue, as in some monocotyledons, or there may be strands of internal phloem or complete vascular bundles located within the pith, as in some dicotyledons. The pith may consist of parenchyma or may become sclerotic (hard-walled), especially at the nodes. The outer zone of the pith often consists of small, persistent cells with more or less sclerified walls. Many herbaceous plants have hollow stems (internodes) because the growth of the pith fails to keep pace with stem expansion. Some woody plants, such as the walnut and *Liriodendron*, have a diaphragmed pith, that is, pith with thick-walled parenchyma cells and sclereids arranged in disks, or diaphragms, of firm tissue.

Primary vascular system. In the ferns and the seed plants the primary vascular system, derived from the procambium, is composed of anastomosing (interconnected) vascular bundles. At each node some of the vascular bundles diverge into the leaves and axillary branches. A bundle in the stem which leads directly into a leaf is called a leaf trace; one leading into a branch is a branch trace. One or more leaf traces may be associated with each leaf. Where the leaf trace diverges from the stem into a leaf at a node, the vascular cylinder commonly has a more or less circumscribed region of parenchyma. This region is called a leaf gap or lacuna. The lateral branches of most dicotyledons and gymnosperms have two branch traces. Some species have one trace per branch, others more than two. In monocotyledons, for example, the axillary shoot connection consists of many vascular strands. At the node the branch traces may have a single branch gap which, moreover, is commonly confluent with the leaf gap.

Separate parts of the primary vascular system of the stem constitute vascular strands or vascular bundles. The phloem and xylem within a vascular bundle may be arranged in various ways. The most common type in gymosperms and angiosperms is the collateral bundle in which the phloem occurs only on one side of the xylem. If phloem occurs on both sides of the bundle, as in the Cucurbitaceae, the bundle is bicollateral. If one kind of vascular tissue completely surrounds the other, the bundles are concentric. The concentric bundle is amphicribral if the phloem surrounds the xylem, as in ferns, or amphivasal if the xylem surrounds the phloem, as in some monocotyledons and dicotyledons.

The primary vascular system with associated fundamental tissue is termed the stele. The simplest type of stele, and perhaps also the most primitive, is the protostele, in which a solid central core of xylem is surrounded by phloem, as found generally in the roots, in the stems of lower vascular plants such as the fern *Gleichenia,* and the stems of some water plants (**Fig. 7***a*). In the plant group Lycopsida the lycopods or "ground pines" have a modified protostele with alternating bands of xylem and phloem (Fig. 7*b*), and the "club mosses" (*Selaginella*) have separate protosteles supported in air spaces by radially elongated endodermal cells designated trabeculae. In the stems of vascular plants located higher than Lycopsida on the evolutionary scale, the center of the stele is occupied by a pith, and the vascular system has a tubular structure. The stele, therefore, is called a siphonostele (tubular stele). If the phloem occurs only on the outer side of the xylem of a siphonostele (gymnosperms and angiosperms), the stele is called an ectophloic siphonostele (Fig. 7*c*). If the phloem differentiates also on the inner side of the xylem (in such ferns as *Adiantum*, *Dicksonia*, and *Marsilia*), the stele is designated an amphiphloic siphonostele or a solenostele (Fig. 7*d*). In the ferns, in which the stem is short and the leaves are set close together, the leaf gaps overlap, and the tubular stele is dissected into a network of distinct strands. Such a stele is called dictyostele, as in *Polypodium*

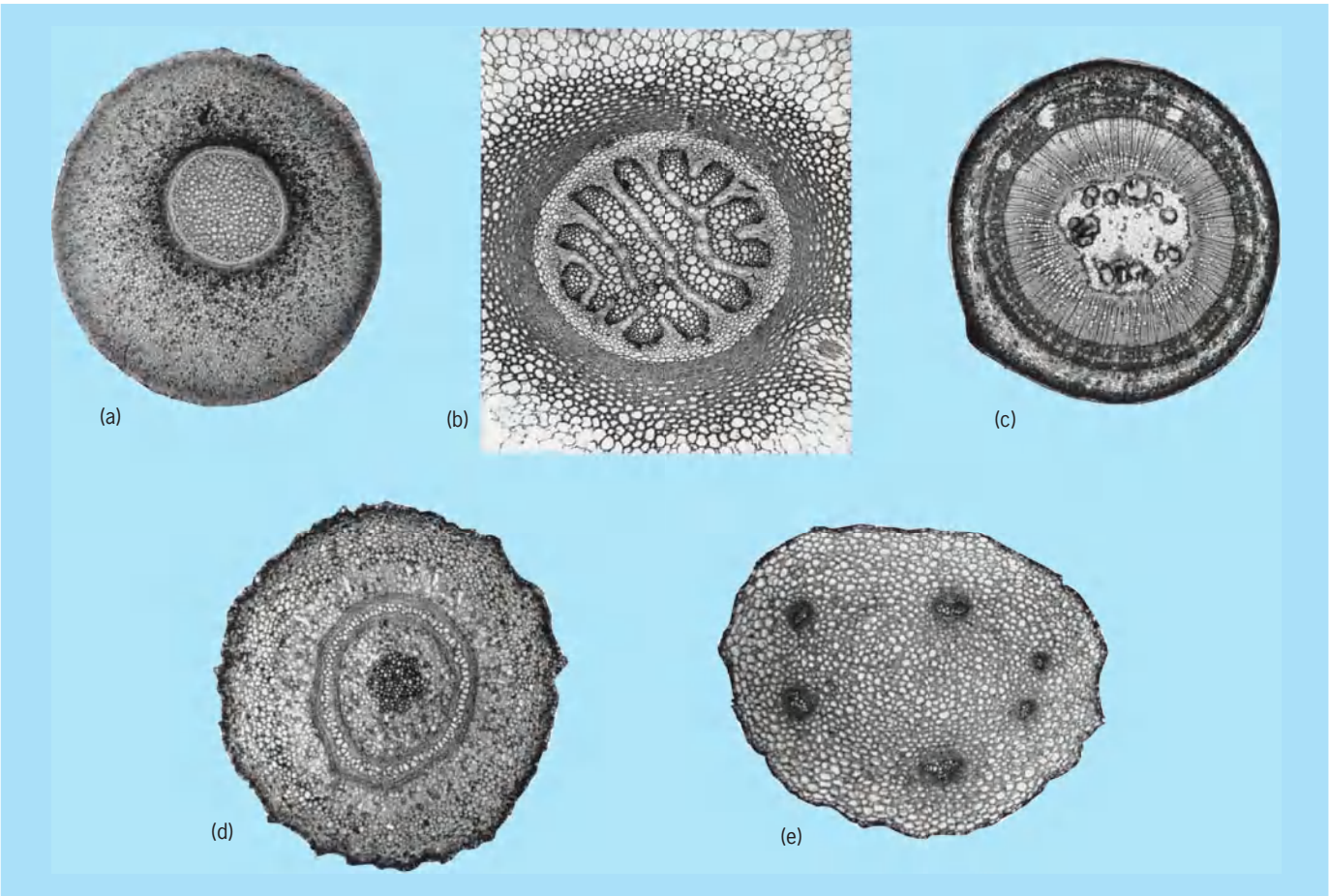

Fig. 7. Types of steles as seen in photomicrographs of cross sections of stems. (*a*) Protostele (*Gleichenia*). (*b*) Radial (modified) protostele (*Lycopodium*). (*c*) Ectophloic siphonostele (*Tilia*). (*d*) Amphiphloic siphonostele (*Dicksonia*). (*e*) Dictyostele (*Polypodium*).

and *Dryopteris* (Fig. 7*e*). Another modification of the siphonostele is the eustele, in which the vascular system consists of collateral or bicollateral strands, and both interfascicular regions and leaf gaps are present. Eustele is characteristic of gymnosperms and angiosperms. The most complex stele, the atactostele, has a dispersed arrangement of strands, as in many monocotyledons.

Primary development. The origin of the stem from the apical meristem is closely associated with the development of leaves. At initiation the leaf primordium is merely a small protuberance on the flank of the shoot tip and cannot be delimited from the stem part to which it is attached. Only later, as growth continues, does the primordium become an organ distinguishable externally and internally from the stem. The stem itself shows at first no division into nodes and internodes. It appears like a complex of superimposed disks (future nodes), each bearing one or more leaf primordia, depending on leaf arrangement. Later, during elongation, the disklike zones become separated from one another by cell division and enlargement between the levels at which the leaf primordia are inserted. Thus, the internodal elongation is initiated, and the stem becomes differentiated into nodes and internodes. The elongation of internodes is the phenomenon that brings about the characteristic rapid extension of new shoots during the spring growth of trees and shrubs.

Internally, the differentiation of stem tissues is more or less closely correlated with the differentiation of leaves. In lower vascular plants with protosteles in their stems, as the lycopods, the vascular system differentiates beneath the apical meristem as a column of provascular tissue (procambium), and the leaf traces, which connect the small leaves with this column, differentiate later through the cortex. In its differentiation the provascular tissue becomes distinct from the surrounding ground tissue of the cortex because the cells of the latter show an increase in vacuolation and an overall enlargement, whereas the cells of the provascular tissue retain their dense cytoplasm and they increase in length but remain narrow.

In the ferns and seed plants, in which the leaves are more prominent parts of the shoot than in the lycopods and the vascular system of the stem consists of a cylinder of strands (siphonostele structure) connected with the leaves at regular intervals (at the nodes), the vascular system is initiated in relation to the leaves. Vacuolation of the differentiating cortex and pith delimits the future vascular region as a still highly meristematic cylinder of cells. Then within this cylinder, procambial strands differentiate

beneath the emerging leaf primordia chiefly by longitudinal cell division and cell elongation; that is, the leaf traces are initiated. The cells between the bundles that do not become procambial cells differentiate as parenchyma cells of the interfascicular regions and the leaf gaps. If followed downward in the stem, the young leaf traces will be found joined with one another and with older traces. Thus, the procambial strands form an interconnected system similar to that of the mature primary vascular system, except that all its parts are still short. The procambial system elongates in coordination with the other tissues of the internodes by further division and elongation of cells. Eventually, primary phloem and primary xylem differentiate in the procambial strands, the pholem in continuity with that of the older parts of the stem (acropetal differentiation, that is, from the base of the stem toward its apex), the xylem first appearing at the bases of leaves then differentiating in two directions: downward (basipetally) toward a connection with the xylem of older stem regions and upward in the leaves. *See* LYCOPODIALES; MAGNOLIOPHYTA; POLYPODIALES.

Differentiation of lateral buds. The lateral or axillary buds arise more or less close to the apical meristem of the terminal bud, depending on the type of branching characteristic of the plant. The buds are initiated by cell division in the stem to produce a lens-shaped group of cells in the outer cortex near the axils of subtending leaves. The divisions occur in such a manner that eventually an apical meristem of the bud is organized. The bud meristem may develop immediately into a side shoot by producing leaves and increments of stem, or it may interrupt growth after variable degrees of initial development. The lateral bud may also develop into an inflorescence, or flower, by appropriate changes in growth pattern as compared with that of a vegetative shoot. The bud and parent stem establish a vascular connection through the bud traces, the differentiation of which is more or less similar to that of the leaf traces of the same plant.

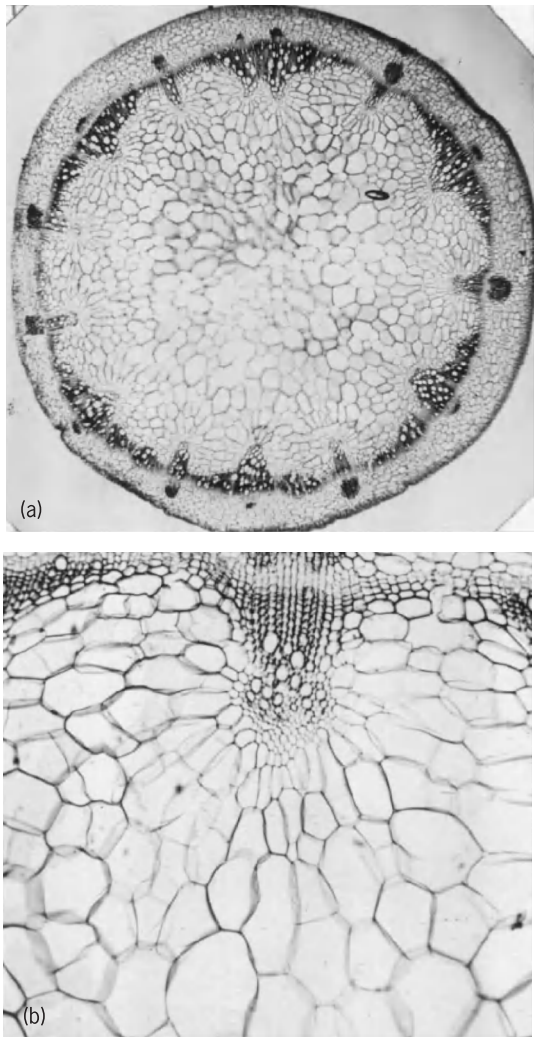

Fig. 8. Secondary tissue in the stem of sunflower (*Helianthus annuus*). (*a*) Transverse section of the entire stem showing the thicker-walled, dark, secondary xylem cells as they are added to the primary, more in the fascicular than in the interfascicular areas. (*b*) Enlarged view of a portion of *a*. (*From J. B. Hill, L. O. Overholts, and H. W. Popp, Botany: A Textbook for Colleges, 2d ed., McGraw-Hill, 1950*)

Secondary State

The stems of gymnosperms, most dicotyledons, and some monocotyledons show an increase in thickness by secondary growth. A lateral meristem, the vascular cambium, produces secondary vascular tissues which constitute the secondary body. Also, a cork cambium, or phellogen, produces a secondary protective tissue called periderm, a tissue which replaces the primary protective tissue, the epidermis.

The vascular cambium originates from the procambial and certain parenchyma cells of the primary body. Procambial cells that remain meristematic after the stem completes its elongation and the primary vascular tissues become mature proceed to divide periclinally (parallel with the surface). These cells are designated as cambial initials. In addition, periclinal divisions are initiated by parenchymal cells between the vascular bundles; these cells also are cambial initials. Collectively, all these initials constitute the vascular cambium. The part of cambium originating within the vascular bundles is the fascicular cambium, that arising in the parenchyma between vascular bundles, the interfascicular cambium (**Fig. 8**).

The arrangement of the secondary vascular tissues produced by the cambium varies in different plants. This tissue may appear as (1) a continuous cylinder; (2) individual strands, with secondary activity limited to the bundles (no interfascicular cambium is formed); (3) individual strands, with secondary vascular tissues produced in the fascicular regions and secondary ray parenchyma in the interfascicular regions; or (4) anomalous (atypical) arrangements characterized by uneven growth of xylem or phloem or by formation of successive cambia, one farther outward than the preceding.

The cork cambium or phellogen originates in the epidermis or in the cortex from parenchyma or collenchyma cells. It produces phellem or cork

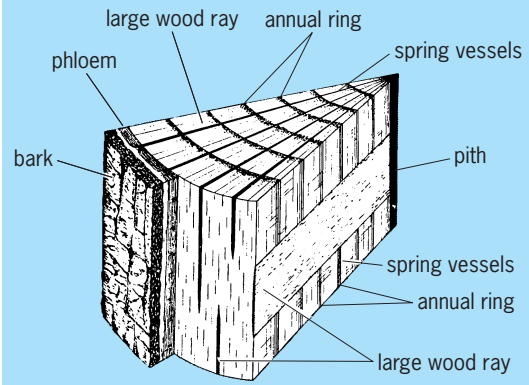

Fig. 9. Diagram of a segment of an oak log showing the grain in three dimensions, that is, the pattern formed by the annual rings and the wood rays. At the right, the block has been cut radially, and above, transversely. At the left (the surface of the log), a portion of the phloem and bark has been removed, showing a tangential view of the wood underneath. (*After E. W. Sinnott and K. S. Wilson, Botany: Principles and Problems, 5th ed., McGraw-Hill, 1955*)

outwardly and phelloderm or secondary cortex inwardly.

Woody dicotyledons. The cambium here is a continuous cylinder. The fascicular cambium produces secondary xylem and phloem; the interfascicular cambium forms either the same kind of tissue as the fascicular cambium or only vascular ray tissue.

As seen in cross sections, the secondary xylem, or wood, is formed in distinct concentric rings called annual, or growth, rings (**Fig. 9**). Each annual ring is composed of an inner portion called the spring, or early, wood, in which the vessels and tracheids are relatively large, and an outer portion called the summer, or late, wood, in which the cells have much smaller diameters and thicker walls than those of the early wood. Usually only the outermost growth rings function in conduction of water. The outer, light-colored, relatively soft, functioning part of the wood is called sapwood. The nonconducting older (early formed) wood, called heartwood, is often filled with gums, resins, tannins, and mineral salts, and is dark in color. Growth rings in oblique or horizontal branches become asymmetrical, the rings being wider in the upper half (in most angiosperms) or in the lower half (in conifers). The wood with the wider rings has a different histologic and chemical structure from that with the narrower and is called tension wood in the angiosperms and compression wood in the conifers.

The wood varies in structure, chiefly depending on (1) the relative amounts of vessels, tracheids, and wood fibers, (2) the distribution of wood parenchyma, and (3) the presence and character of vascular rays.

Two patterns of vessel arrangement are recognized. If the early wood contains wide vessels and the late wood narrow vessels, so that each growth ring begins with a zone of large vessels, the wood is called ring-porous (**Fig. 10***a*). If vessels of more or less uniform widths are formed throughout the year, the wood is called diffuse-porous (Fig. 10*b*).

Wood texture, coarse or fine, refers to the size and number of the xylem cells, especially vessels. The term grain refers to the arrangement of the cells, especially fibers, making up the wood. A wood is hard or soft, depending upon the number of lignified fibers. *See* WOOD ANATOMY.

Among the many wood products, plywood and veneer are particularly important. In making plywood, the wood is softened and the log rotated against a heavy blade to produce a continuous sheet of very thin wood which is cut to suitable lengths; pieces alternately stacked with grain at right angles are bonded together in varying odd numbers of layers, pressed, and dried. Plywood, or other wood bases, may be covered by thin sheets of fine-grained wood or veneer. Characteristic and interesting grain patterns for fine woods or veneers are obtained by cutting wood in transverse, radial, and tangential section (**Fig. 11**). Veneer wood with the most interesting pattern or grain comes from the base of a tree,

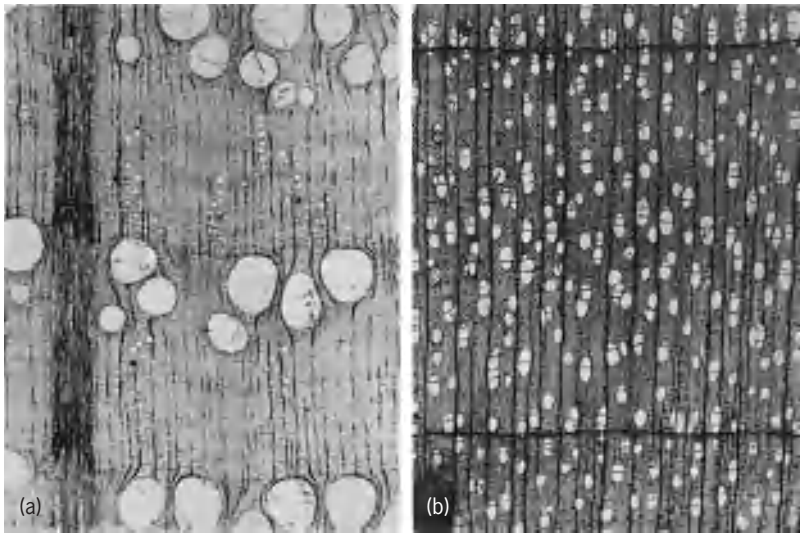

Fig. 10. Transverse sections, magnified, of two kinds of dicotyledonous wood. (*a*) White oak, ring-porous wood with large spring vessels and narrow (uniseriate) and wide (multiseriate) rays. (*b*) Yellow birch, a diffuse-porous wood with narrow (mostly uniseriate) rays. (*U. S. Forest Products Laboratory*)

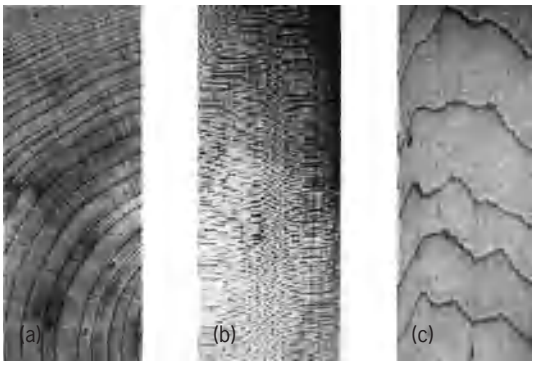

Fig. 11. Silver maple (*Acer saccharinum*) wood in different sections. Dark and light colors and grain irregularities result from cutting through fibers and vessels of annual rings in different planes. Also, the rays, particularly in radial section, give the wood a pleasing and interesting appearance. (*a*) Transverse section. (*b*) Radial section. (*c*) Tangential section. (*From J. B. Hill, L. O. Overholts, and H. W. Popp, Botany: A Textbook for Colleges, 2d ed., McGraw-Hill, 1950*)

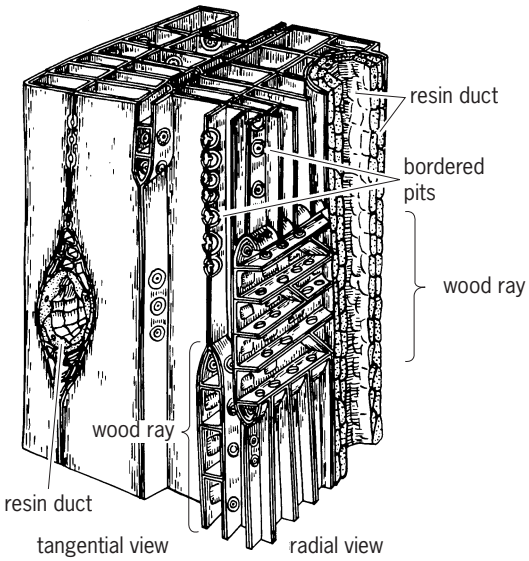

Fig. 12. Block of white-pine wood, in three dimensions, showing relationships of tracheids and ray elements. (*After J. B. Hill, L. O. Overholts, and H. W. Popp, Botany: A Textbook for Colleges, 2d ed., McGraw-Hill, 1950*)

from burls, or from crotch wood, where the vascular arrangements may be wavy or spiral. Some of the choice curly grains, such as curly maple, result from the development of adventitious buds. Injury to the cambium may produce a similar pattern. *See* WOOD PRODUCTS.

Gymnosperms. Most gymnosperms are similar to the woody dicotyledons in that a vascular cambium forms a continuous cylinder of secondary xylem and phloem. Commercial gymnosperm wood is usually derived from conifers. Conifer wood is composed of tracheids, fiber tracheids, and uniseriate wood rays, but no vessels (**Fig. 12**). Because of the absence of vessels, the conifer wood is more homogeneous in sections than the angiosperm wood (**Fig. 13**).

As stems grow and form new branches, those near the base are likely to be crowded, shaded, and die, particularly in pine. The dead branches fall off, and their bases are covered by successive layers of wood, forming knots. Knots weaken the wood and are not desirable unless for ornamental paneling.

The secondary growth of other gymnosperms is similar to that of the conifers. The cycads are interesting because growth rings are added at intervals of 2–20 years and are difficult to distinguish, except in *Dioon*. Many of the cycads form only one ring of wood; but others, such as *Cycas*, *Dioon*, *Macrozamia*, and *Encephalartos*, produce several rings.

Monocotyledons. Some monocotyledons, as the bamboos and palms, may form rather woody trunks by primary thickening. The thickening growth may occur rather generally throughout the stem, or there may be a primary thickening meristem, a cambium-like zone of cells originating from the ground tissue beneath young leaf bases. It increases the width of the stem directly beneath the apex. Additional division of cells throughout the stem and the enlargement of cells complete the growth in thickness. In certain monocotyledons, such as *Aloe*, *Yucca*, *Agave*, and *Dracaena*, secondary growth also occurs. However, the cambium does not originate in the bundles, which are scattered in the ground tissue. It appears in the ground parenchyma outside the vascular bundles and by tangential divisions forms parenchyma on the outside and both vascular strands and parenchyma on the inside. The basic structure of the primary and secondary bodies is the same—in both, discrete vascular strands are embedded in ground tissue, usually sclerified (hardened) parenchyma.

Effect of secondary growth. The secondary xylem covers the primary xylem and pith, usually without changing them; whereas the primary phloem and cortex are pushed outward and are more or less compressed by the centrifugal growth of the secondary vascular tissues. The epidermis and cortex may keep pace with secondary vascular growth, or they may be ruptured and replaced by periderm.

Cambium develops in the parenchyma of the leaf gap, and its derivatives close the leaf gaps and break and bury the leaf traces.

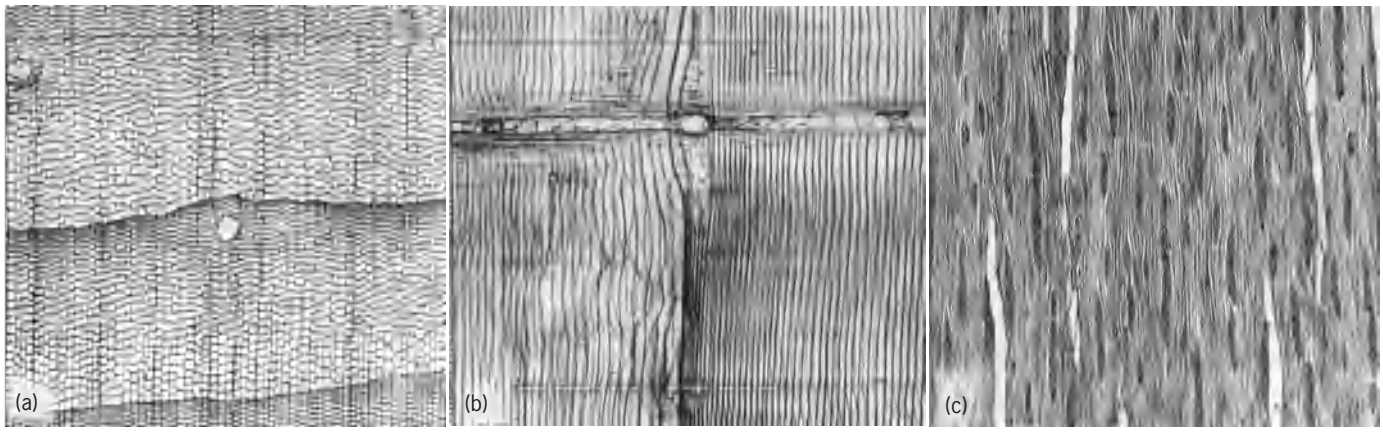

Fig. 13. Enlarged photomicrographs of pine wood. (*a*) Transverse section. The wood cells are cut across at right angles to their length. Note the thin-walled cells (tracheids) in the springwood and the thick-walled cells in the summerwood (fiber tracheids). The uniseriate wood rays run at right angles to the annual rings. The large openings are resin canals. (*b*) Radial section. (*c*) Tangential section. In the last two sections the tracheids are cut lengthwise. In the radial section the rays run in a horizontal direction across the tracheids. In the tangential section the rays appear as lens-shaped clusters of cells. (*U. S. Forest Products Laboratory*)

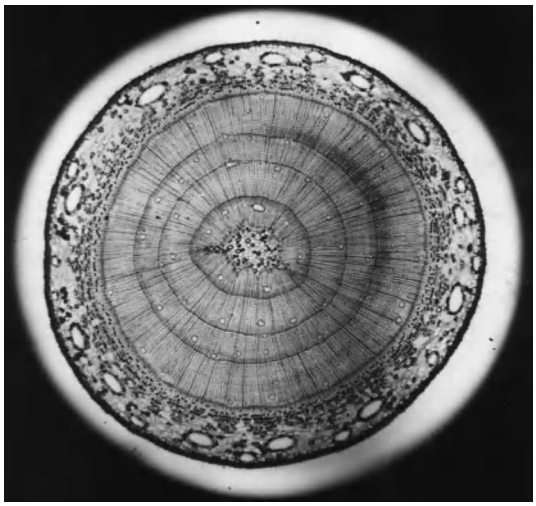

Fig. 14. Cross section of 4-year-old pine stem, showing annual xylem rings and other tissues. (*From J. B. Hill, L. O. Overholts, and H. W. Popp, Botany: A Textbook for Colleges, 2d ed., McGraw-Hill, 1950*)

Types

The stems of vascular plants, or Tracheophyta, show an endless variety of form and structure. In this article only a few are considered as type examples.

Gymnosperm stems. This group of Tracheophyta contains some of the largest plants in the world (especially in Pinales), as well as *Welwitschia*, in which there is no aerial stem at all. No annual or perennial herbs are recognized in this group. *See* PINOPSIDA.

Pinales. The gymnosperms of this order are much-branched, small-leaved, woody tree forms, except some shrubby junipers. The stem has typically an ectophloic siphonostele. Most conifers produce a conspicuous amount of secondary tissues which form a solid continuous cylinder (**Fig. 14**). Resin ducts occur in the cortex and the vascular tissues, except in the yews. The wood is dense and massive, composed of tracheids, fiber tracheids, and uniseriate wood rays. *See* SECRETORY STRUCTURES (PLANT).

Cycadales. The stem types range from tuberous and partly or wholly subterranean to relatively tall, aerial stems. The stem is usually unbranched and bears a crown of large leaves. The cortex and pith are large, loose, and parenchymatous. The vascular cylinder is narrow and has broad medullary rays and a loose-textured wood. The stele is a siphonostele. The primary xylem usually contains scalariform tracheids, the secondary xylem tracheids with bordered pits. Leaf traces girdle the stem that is, they have a nearly horizontal course in the cortex. The stem is rigid mainly because of persistent leaf bases.

Ginkgoales. The only living representative, *Ginkgo biloba*, is a large profusely branched tree with both long shoots (normal vegetative) and short shoots (restricted vegetative or reproductive). Secondary growth is vigorous, and the pith and cortex are relatively small.

Gnetales. The three genera included in this order are highly specialized. *Ephedra* is usually a much branched shrub, *Gnetum* species are mostly lianas, and in *Welwitschia* most of the stem is buried in the ground. A feature which distinguishes these plants from other gymnosperms is the presence of vessels among the xylem elements.

Angiosperms. The angiosperms (flowering plants) include two main groups: the dicotyledons (plants with two seed leaves) and the monocotyledons (plants with one seed leaf). The stems of these two groups of plants show some differences. It is customary also to divide the stems of angiosperms into woody and herbaceous types. Although the various stem types intergrade in their character, some approximate generalizations regarding their differences may be made, as shown in the **table**.

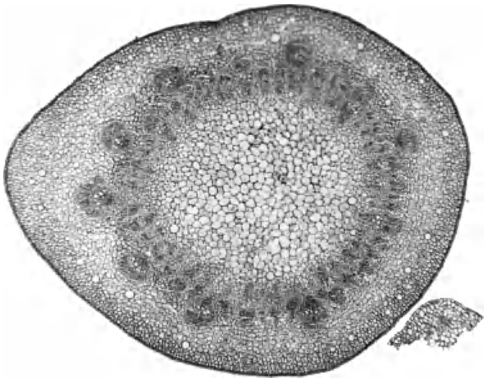

Fig. 15. Woody dicotyledon; 1-year-old woody stem of *Liriodendron*, showing characteristics of ectophloic siphonostele; epidermis not yet replaced by periderm.

Comparison of dicotyledon and monocotyledon stems

Dicotyledons	Monocotyledons
Stems woody, herbaceous, vines	Stems herbaceous, few woody
Secondary growth common; phloem, vascular cambium, and xylem in concentric layers; separate vascular bundles common in the primary body, uncommon in secondary	Some woody types with secondary growth; vascular cambium not between xylem and phloem; in the primary body vascular bundles often scattered through ground tissue, sometimes in two circles; in the secondary body also separate vascular bundles
Distinct pith and cortex	No definable cortex and pith
Endodermis and pericycle usually absent; sometimes pericycle fibers on outer limit of vascular cylinder; usually phloem fibers	Endodermis and pericycle usually absent; vascular bundles with sheaths of sclerenchyma
Number of leaf traces to each leaf is variable, rarely large	Number of leaf traces to each leaf is large; sheathing leaf bases enclose the stem
Periderm frequently present	Periderm not common

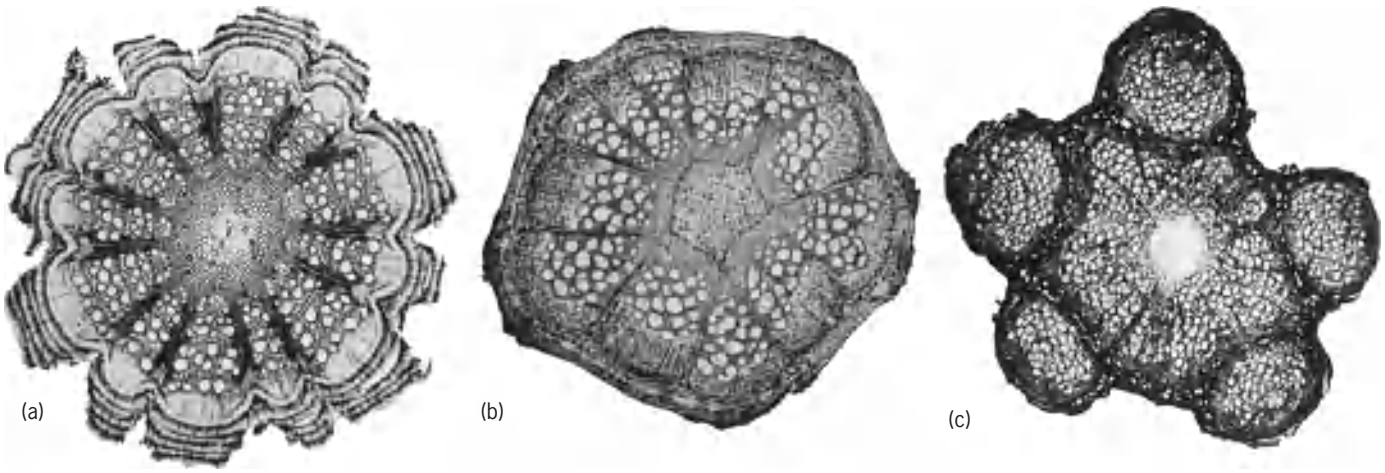

Fig. 16. Transverse sections of woody vines. (*a*) Clematis. (*b*) *Serjania mexicana*; mature stem shows irregular development (unequal amounts) of secondary xylem and phloem. (*c*) *Serjania subdentata*; mature stem shows five well-developed cortical steles.

Woody dicotyledons. The primary vascular cylinder in the woody dicotyledons is an ectophloic siphonostele, and the pith and cortex are well defined. *Liriodendron* illustrates this type of structure (**Fig. 15**). The stem from the inside out consists of (1) a central pith of large, loose, parenchyma cells; (2) a vascular system containing closely arranged vascular bundles, with the leaf traces often recognizable as rather large bundles protruding into the cortex; (3) a cortex; and (4) an epidermis. When secondary growth occurs, a continuous vascular cambium and continuous layers of secondary xylem and phloem are formed. The epidermis is eventually replaced by a periderm. There is no endodermis and no pericycle. Primary phloem fibers form the outer limit of the vascular region.

Woody vines or lianas. Vines are generally characterized by a primary and secondary vascular system dissected into strands by broad rays (**Fig. 16***a*). Also, they often have a ringbark which results from concentric successively formed periderms. Lenticels may be absent in ringbark.

Tropical lianas (climbing woody vines) often have a more or less unusual structure. There may be, for example, a highly irregular development of secondary xylem and phloem, so that the outline of the xylem is uneven (Fig. 16*b*). A still more striking structure results when the leaf traces develop complete cylinders of cambium, xylem, and phloem at levels where they have diverged from the central cylinder. The stem appears as though composed of several central cylinders (Fig. 16*c*).

Herbaceous dicotyledons. These forms differ from the arborescent, or woody, dicotyledons chiefly in having a smaller amount of cambial activity or none. They are often similar to 1-year-old woody stems. The primary vascular cylinder forms an ectophloic siphonostele with the vascular strands often more widely separated by interfascicular regions than those in the arborescent species (**Fig. 17**). Vascular cambium may be absent, as in *Ranunculus*; may be present but form only parenchyma and sclerenchyma in the interfascicular regions, as in *Medicago*, or may form a continuous vascular cylinder, as in *Vinca* or *Digitalis*.

Herbaceous monocotyledons. In these plants the vascular system is composed of widely spaced strands arranged in one of four ways. First, as in most grasses, the vascular bundles are arranged in two circles, with the outer smaller bundles embedded in a continuous sheath of sclerenchyma close to the epidermis (**Fig. 18***a*). The vascular bundles are collateral, each enclosed in a sheath of sclerenchyma. The pith may break down in the internodes but not in the nodes. Transverse bundles in the nodal region interconnect the leaf traces.

Second, in a few monocotyledons, for example *Clintonia* (Fig. 18*b*), a single, concentrically arranged series of bundles may occur. Each bundle has its own complete endodermis.

Third, in a few monocotyledons the bundles are grouped in the center of the stem, as in the rhizome

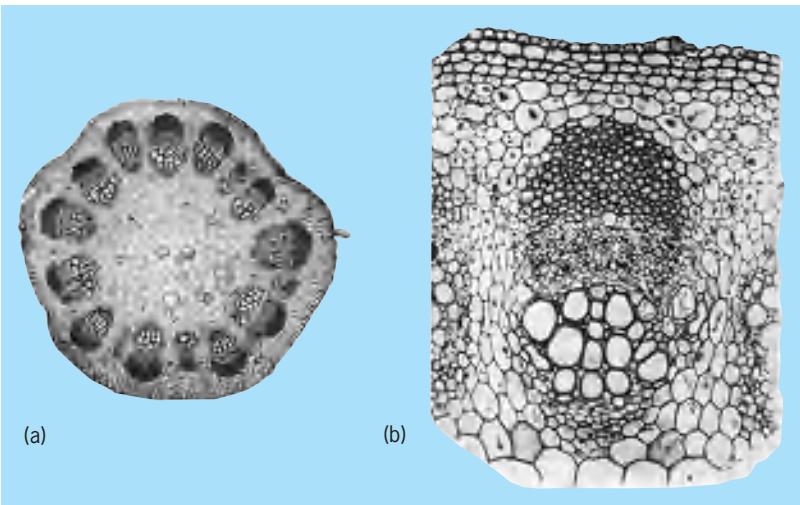

Fig. 17. The primary tissues of the stem of a typical herbaceous dicotyledonous plant, sunflower (*Helianthus annuus*), showing the characteristics of an ectophloic siphonostele. (*a*) Transverse section of entire stem showing the arrangement of vascular tissues in one circle. (*b*) Enlarged portion of *a* showing details of the structure, including one vascular bundle. (*From J. B. Hill, L. O. Overholts, and H. W. Popp, Botany: A Textbook for Colleges, 2d ed., McGraw-Hill, 1950*)

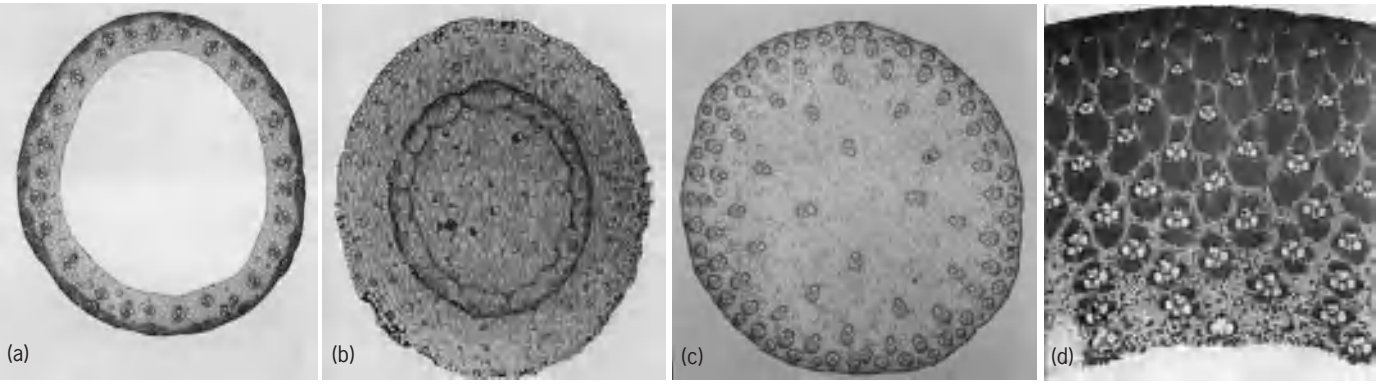

Fig. 18. Photomicrographs of transverse sections of monocotyledonous stems showing various arrangements of vascular bundles. (*a*) Wheat (grass). (*b*) *Clintonia*. (*c*) Corn. (*d*) Bamboo. (*From J. B. Hill, L. O. Overholts, and H. W. Popp, Botany: A Textbook for Colleges, 2d ed., McGraw-Hill, 1950*)

of *Acorus calamus*. The bundles are amphivasal, the phloem being surrounded by xylem.

Fourth and most commonly, vascular bundles are numerous and scattered in a ground tissue, as in corn or bamboo (Fig. 18*c* and *d*), with no pith or cortical regions being evident. The complex arrangement of the vascular bundles is related to the large number and variation in size of leaf traces. There may or may not be a sclerenchyma cylinder about each bundle, but subepidermal parenchyma is strongly sclerified. Endodermis and pericycle are absent.

Woody monocotyledons. Some monocotyledons, as the bamboos and palms, may form rather woody trunks by thickening of cell walls with age until nearly the whole ground tissue is heavily sclerified. Despite their large size these trunks are composed of primary tissues only. There are, however, woody monocotyledons with secondary growth, but the secondary tissues resemble the primary tissues in that they are composed of vascular bundles embedded in ground tissue, usually sclerified parenchyma.

Specialized erect stems. Two kinds of erect stems have special names. Culm is a name applied to hollow, but solid-jointed, stems of grasses, woody and herbaceous. Caudex means the axis of a plant, consisting of root and stem. The term is sometimes applied to short, enduring stems or stocks which throw up new stalks each year from persistent buds at ground level. Also, the name is used to describe the trunks of palms and tree ferns which consist of persistent leaf bases forming a column.

Perennial stems. Perennials which have woody stems are generally called trees or shrubs. Shrubs are woody perennials with several main stems branched from or near the ground and generally not more than 20–25 ft (6.0–7.5 m) high. Trees are woody perennials with a single main trunk or axis which rises some distance above the ground before branching. Trees generally exceed shrubs in height. The two words are convenient but not exact terms because trees and shrubs intergrade in height and form. Herbaceous plants also may be perennial. Their stems commonly die down to the ground each year.

Modified aerial stems. Stems may be variously modified for photosynthesis, storage, propagation, support, and protection.

Cladode. This term is applied to a branch with a single internode which is flattened and serves as a leaf, as is found in asparagus.

Phylloclade. A flattened leaflike shoot that serves to replace a leaf as a photosynthetic organ, as found in the *Opuntia* cactus, is called a phylloclade.

Thorn. This hard, pointed projection may be a modified branch, or leaf, or stipule. Thorns may be unbranched, as in the osage orange, or branched, as in the honey locust. A lateral branch may form leaves and flowers, then terminate growth by producing a thorn at the apex.

Creeping or prostrate stems. Stems that trail along the surface of the ground and that may take root at the nodes are called creeping or prostrate stems.

Runner. A runner, as in the strawberry, is a horizontally growing, sympodial stem system; that is, it appears single but actually is composed of a series of lateral branches arranged in lineal order. In this system the stem forms adventitious roots near the tip, then gives rise to a rosette of leaves from a bud. A new runner emerges from the axil of a reduced leaf at the base of this rosette.

Stolon. The term stolon is used for creeping stems, shorter than runners. They also root adventitiously (raspberry, currant). The term has no precise meaning.

Climbing stems or vines. These are long, slender stems which usually climb by special devices.

Rambler. The ramblers rest on the tops of other plants and some, as certain roses, have spines or prickles which help them to adhere to their support.

Root climbers. English ivy and posion ivy have stems which climb by means of adventitious roots.

Tendril climbers. These plants climb by means of modified leaf tendrils, as in the garden pea, or by stem tendrils, as in the grape.

Twiners. In the twiners the whole stem winds about its support, as pole beans or many tropical lianas.

Underground stems. Some stems grow underground and are often mistaken for roots. The main kinds of underground stems are rhizomes or rootstocks, tubers, corms, bulbs, and rhizomorphic droppers.

Rhizomes are usually quite varied, plagiotropic (growing horizontally), perennial, underground

stems found in a vast number of plants. Their cauline (stem) nature is evident by well-defined nodes and internodes. Roots, scale leaves, and axillary buds form at the nodes. Some buds form leafy, upright shoots while others form new, sympodially branched underground shoots. Among different species, rhizomes may be (1) thin, tough, and rapidly growing; (2) fleshy though rapidly growing; and (3) short and fleshy, serving more for storage than for spreading. The anatomy of rhizomes, except for some dorsiventrality, is much like that of the aerial stem of the same species. The other types of underground stems, such as tubers, corms, and bulbs, are essentially modified rhizomes. James E. Gunckel

Bibliography. J. D. Mauseth, *Plant Anatomy*, 1988.

Stem cells

Cells that have the ability to self-replicate and to give rise to mature cells. The concept of stem cells was originally based on renewing tissues. Many adult tissues, such as the skin, blood, and intestines, consist of mostly mature and short-lived cells that must be continuously replaced. Stem cells were postulated as the source of the self-renewal. In the early 1960s, Canadian scientists Ernest A. McCulloch and James E. Till provided the first experimental proof of the existence of stem cells in the blood system. They revealed that a type of cell in bone marrow possesses the capacity to replicate itself and to differentiate to various lineages of mature blood cells. Self-renewal, together with the capacity for differentiation, defined the properties of stem cells. This definition is generally used in stem cell biology today.

Stem cells can be found at different stages of fetal development and are present in a wide range of adult tissues. Many of the terms used to distinguish stem cells are based on their origins and the cell types of their progeny. There are three basic types of stem cells. Totipotent stem cells, meaning that their potential is total, have the capacity to give rise to every cell type of the body and to form an entire organism. Pluripotent stem cells, such as embryonic stem cells, are capable of generating virtually all cell types of the body but are unable to form a functioning organism. Multipotent stem cells can give rise only to a limited number of cell types. For example, adult stem cells, also called organ- or tissue-specific stem cells, are multipotent stem cells found in specialized organs and tissues after birth. Their primary function is to replenish cells lost from normal turnover or disease in the specific organs and tissues in which they are found. *See* CELL DIFFERENTIATION; HISTOGENESIS.

Totipotent and embryonic stem cells. Totipotent stem cells occur at the earliest stage of embryonic development. The union of sperm and egg creates a single totipotent cell. This cell divides into identical cells in the first hours after fertilization. All these cells have the potential to develop into a fetus when they are placed into the uterus. [To date, no such totipotent stem cell lines (primary cell cultures) have been developed.] The first differentiation of totipotent cells forms a sphere of cells called the blastocyst, which has an outer layer of cells and an inner cell mass. The outer layer of cells will form the placenta and other supporting tissues during fetal development, whereas cells of the inner cell mass go on to form all three primary germ layers: ectoderm, mesoderm, and endoderm. The three germ layers are the embryonic source of all types of cells and tissues of the body. Embryonic stem cells are derived from the inner cell mass of the blastocyst. They retain the capacity to give rise to cells of all three germ layers. However, embryonic stem cells cannot form a complete organism because they are unable to generate the entire spectrum of cells and structures required for fetal development. Thus, embryonic stem cells are pluripotent, not totipotent, stem cells. *See* EMBRYOLOGY; EMBRYONIC DIFFERENTIATION; GERM LAYERS.

Embryonic germ cells. Embryonic germ cells differ from embryonic stem cells in the tissue sources from which they are derived, but appear to be similar to embryonic stem cells in their pluripotency. Human embryonic germ cell lines are established from the cultures of the primordial germ cells obtained from the gonadal ridge of late-stage embryos, a specific part that normally develops into the testes or the ovaries. Embryonic germ cells in culture, like cultured embryonic stem cells, form embryoid bodies, which are dense, multilayered cell aggregates consisting of partially differentiated cells. The embryoid body–derived cells have high growth potential. The cell lines generated from cultures of the embryoid body cells can give rise to cells of all three embryonic germ layers, indicating that embryonic germ cells may represent another source of pluripotent stem cells.

Growing mouse embryonic stem cells. Much of the knowledge about embryonic development and stem cells has been accumulated from basic research on mouse embryonic stem cells. The techniques for separating and culturing mouse embryonic stem cells from the inner cell mass of the blastocyst were first developed in the early 1980s. To maintain their growth potential and pluripotency, mouse embryonic stem cells can be grown on a feeder layer, usually consisting of mouse embryonic fibroblast cells. The feeder cells support embryonic stem cells by secreting a cytokine growth factor, the leukemia inhibitory factor, into the growth medium. Alternatively, purified leukemia inhibitory factor can be added to the growth medium without the use of a mouse embryonic feeder layer. (The leukemia inhibitory factor serves as an essential growth factor to maintain embryonic stem cells in culture.) A line of embryonic stem cells can be generated from a single cell under culture conditions that keep embryonic stem cells in a proliferative and undifferentiated state. Embryonic stem cell lines can produce indefinite numbers of identical stem cells. When mouse embryonic stem cells are integrated into an embryo at the blastocyst stage, the introduced embryonic stem cells can contribute to cells in all tissues of the resulting mouse. In the absence of feeder cells and the

leukemia inhibitory factor in cultures, embryonic stem cells undergo differentiation spontaneously. Many studies are focused on directing differentiation of embryonic stem cells in culture. The goal is to generate specific cell types. Formation of cell aggregates with three-dimensional structure during embryonic stem cell differentiation in culture may allow some of the cell-cell interaction to mimic that of in vivo development. The culture conditions can be designed to support and select specific cell types. With these experimental strategies, preliminary success has been achieved to generate some cell types, such as primitive types of vascular structures, blood cells, nerve cells, and pancreatic insulin-producing cells. *See* GENETIC ENGINEERING; SOMATIC CELL GENETICS.

Growing human embryonic stem cells. Since 1998, research teams have succeeded in growing human embryonic stem cells in culture. Human embryonic stem cell lines have been established from the inner cell mass of human blastocysts that were produced through in vitro fertilization procedures. The techniques for growing human embryonic stem cells are similar to those used for growth of mouse embryonic stem cells. However, human embryonic stem cells must be grown on a mouse embryonic fibroblast feeder layer or in media conditioned by mouse embryonic fibroblasts (see **illustration**). There are a number of human embryonic stem cell lines being generated and maintained in laboratories in the United States and other nations, including Australia, Sweden, India, South Korea, and Israel. The National Institutes of Health has created a Human Embryonic Stem Cell Registry, which lists stem cell lines that have been developed and can be used for research. Human embryonic stem cell lines can be maintained in culture to generate indefinite numbers of identical stem cells for research. As with mouse embryonic stem cells, culture conditions have been designed to direct differentiation into specific cell types (for example, neural and hematopoietic cells).

Adult stem cells. Adult stem cells, also referred to as somatic stem cells, occur in a wide variety of mature tissues in adults as well as in children. Like all stem cells, adult stem cells can self-replicate. Their ability to self-renew can last throughout the lifetime of individual organisms. Unlike embryonic stem cells, though, it is usually difficult to expand adult stem cells in culture. Adult stem cells reside in specific organs and tissues but account for a very small number of the cells in tissues. They are responsible for maintaining a stable state of the specialized tissues. To replace lost cells, stem cells typically generate intermediate cells called precursor or progenitor cells, which are no longer capable of self-renewal. However, they continue undergoing cell division, coupled with maturation, to yield fully specialized cells. Such stem cells have been identified in many types of adult tissues, including bone marrow, blood, skin, gastrointestinal tract, dental pulp, retina of the eye, skeletal muscle, liver, pancreas, and brain. Adult stem cells are usually designated according to their source and their potential. Adult stem cells are multipotent because their potential is normally limited to one or more lineages of specialized cells. However, a special multipotent stem cell that can be found in bone marrow, called the mesenchymal stem cell, can produce all cell types of bone, cartilage, fat, blood, and connective tissues. *See* REGENERATION (BIOLOGY).

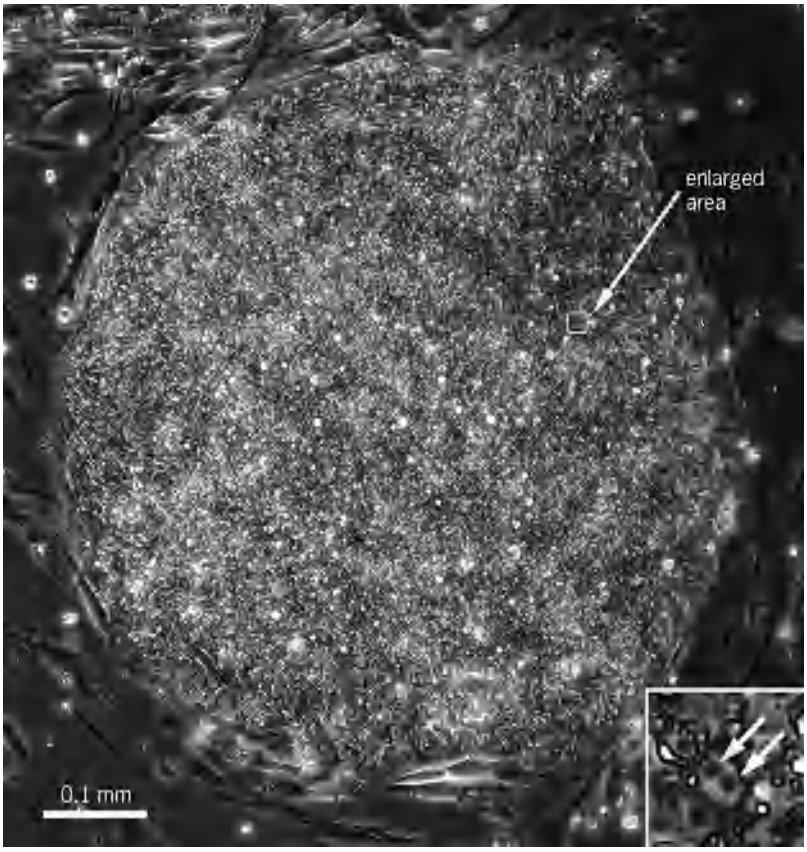

A colony of the human embryonic stem cell line H9 growing in culture. The elongated cells surrounding the colony are part of the mouse embryonic fibroblast feeder layer on which the stem cells are grown. The inset shows the enlarged area, with arrows pointing to the two nuclei of a single stem cell. The H9 cell line meets the criteria for the use of human embryonic stem cells by researchers funded by the U.S. government. (*Courtesy of Clay Glennon*)

Blood stem cells. Blood stem cells, or hematopoietic stem cells, are the most studied type of adult stem cells. The concept of hematopoietic stem cells is not new, as it has been long realized that mature blood cells are constantly lost and destroyed. Billions of new blood cells are produced each day to make up the loss. This process of blood cell generation, called hematopoiesis, occurs largely in the bone marrow. The presence of hematopoietic stem cells in the bone marrow was first demonstrated by E. A. McCulloch and J. E. Till in a mouse model in the early 1960s. The first experimental work on stem cells was an unexpected outcome from their study for measuring the effects of radiation. They found that the blood system of a mouse that has been subjected to heavy radiation can be restored by infusion of bone marrow. The stem cells responsible for reconstituting the blood system generate visible cell colonies on the spleen of the recipient mouse. Each of the spleen colonies consists of one or more types of blood cells, and all the cells in a colony are

derived from a single cell. Self-renewal capacity of the colony-forming cells is demonstrated by their ability to form secondary spleen colonies. Such blood stem cells, known as colony forming unit–spleen cells, qualify as pluripotent hematopoietic stem cells because they can replicate and give rise to multiple types of mature blood cells. A definitive proof of blood stem cells is their ability to reconstitute the blood system. Bone marrow transplantation demonstrates the restorative powers of blood stem cells in humans. *See* HEMATOPOIESIS; TRANSPLANTATION BIOLOGY.

Isolating blood stem cells. Like other adult stem cells, blood stem cells are rare and difficult to isolate. Only about 1 in 100,000 cells in the bone marrow is a stem cell. Scientists have used cell-sorting methods to enrich and purify blood stem cells. Stem cells differ from mature cells in their surface markers, which are specific protein molecules on the cell membrane that can be tagged with monoclonal antibodies. By using a set of surface markers, some expressed mainly on stem cells and others on mature blood cells, nearly pure populations of stem cells can be separated from bone marrow. The stem cells purified by this approach can engraft (begin to grow and function) and reconstitute the blood system in the recipient. In animal studies, as few as 30 purified stem cells can rescue a mouse that has been subjected to heavy radiation. Besides the bone marrow, a small number of blood stem cells can be found in circulating blood. In addition, stem cells in the bone marrow can be mobilized into the bloodstream by injecting the donor with certain growth factors or cytokines. This approach can result in a large number of stem cells circulating in peripheral blood, from which they can be collected and used for transplant therapy.

Umbilical cord blood and cord blood banks. An alternative source of blood stem cells is human umbilical cord blood, a small amount of blood remaining in the placenta and blood vessels of the umbilical cord. It is traditionally treated as a waste material after delivery of the newborn. However, since the recognition of the presence of blood stem cells in umbilical cord blood in the late 1980s, its collection and banking has grown quickly. Similar to bone marrow, umbilical cord blood can be used as a source material of stem cells for transplant therapy. In 1989, the first successful cord blood transplant was reported for treating a 6-year-old boy suffering from Fanconi's anemia (an inherited disease that primarily affects the bone marrow, resulting in decreased production of blood cells) in Paris. Since then, over 6000 cord blood stem cell transplants have been performed worldwide, mainly in patients with blood conditions and in some cancer therapies. However, because of the limited number of stem cells in umbilical cord blood, most of the procedures are performed on young children of relatively low body weight. A current focus of study is to promote the growth of umbilical cord blood stem cells in culture in order to generate sufficient numbers of stem cells for adult recipients.

Many blood banks have been established to collect and cryopreserve cord blood cells. Commercial banks offer services of storing cord blood of healthy newborns for potential future use by themselves or their siblings. Although it is considered a biological insurance, the chance of a child using his or her own cord blood is estimated at 1 per 20,000 collections. Of the estimated 6000 cord blood transplants, only 14 were performed using autologous sources.

Public banks encourage donation of cord blood for unrelated transplants. The Stem Cell Research and Therapeutic Act of 2005 (H.R. 2520) established a national umbilical cord blood program, providing federal funding to collect and store cord blood for blood cell transplants. The program functions to provide a national inventory of 150,000 cord blood units for public use and to establish a registry network integrated with the national marrow donor registry administered by the National Marrow Donor Program (NMDP).

Mesenchymal stem cells. Mesenchymal stem cells (MSCs) are a type of multipotent adult stem cells, and they are defined by the capacity to give rise to a variety of connective tissue lineages, including bone, cartilage, tendon, muscle, and fat cells. Classic studies found a type of cells in bone marrow stroma capable of generating fibroblast-like cell colonies. These clonogenic cells were termed colony forming unit–fibroblasts (CFU-F). CFU-F share some characteristics of MSCs. MSCs appear as fibroblast-like spindle-shaped cells. CFU-F assay is still used to evaluate MSCs in cell cultures. MSCs can be distinguished and isolated from other cells based on phenotypic characteristics. Typically, MSCs express specific surface antigens SH2, SH4, and STRO-1 and lack blood cell markers CD45 and CD34. MSCs can replicate as multipotent cells. The mesenchymal cell lineage potential can be demonstrated in vitro with appropriate culture conditions. Differentiation can be induced to osteocytes by dexamethasone and ascorbate, to chrondrocytes by transforming growth factor-β3, or to adipocytes by dexamethasone and insulin.

MSCs are primarily obtained from bone marrow stromal cells. They are also found in small numbers in umbilical cord blood. In addition, adipose-derived stem cells (ASCs) have been shown to be similar to MSCs. Fat tissue is of mesenchymal origin and contains stromal components. ASCs can be isolated from fat tissue by the method of liposuction. Human ASCs have been shown to exhibit the capacity to give rise to fat, bone, cartilage, muscle, and possibly neurons. Thus, ASCs may provide a potential source of multipotent adult stem cells.

Neural stem cells. Neural stem cells, the multipotent stem cells that generate nerve cells, are a new focus in stem cell research. Active cellular turnover does not occur in the adult nervous system as it does in renewing tissues such as blood or skin. Because of this observation, it had been dogma that the adult brain and spinal cord were unable to regenerate new nerve cells. However, since the early 1990s, neural stem cells have been isolated from the adult brain as well as from fetal brain tissues. Stem cells in the adult brain are found in the areas called the subventricular zone and the ventricle zone. Brain

ventricles are small cavities filled with cerebrospinal fluid. Another location of brain stem cells occurs in the hippocampus, a special structure of the cerebral cortex related to memory function. Stem cells isolated from these areas are able to divide and to give rise to nerve cells (neurons) and neuron-supporting cell types in culture.

Plasticity. Stem cell plasticity refers to the phenomenon of adult stem cells from one tissue generating the specialized cells of another tissue. The longstanding concept of adult organ-specific stem cells is that they are restricted to producing the cell types of their specific tissues. However, a series of recent studies have challenged the concept of tissue restriction of adult stem cells. Much of the experimental evidence is derived from transplant studies with blood stem cells. Bone marrow stem cells have been shown to contribute to liver, skeletal muscle, and cardiac cells in human recipients. In mouse models, purified blood stem cells have been demonstrated to generate cells in nonblood tissues, including the liver, gut, and skin. Although the stem cells appear able to cross their tissue-specific boundaries, crossing occurs generally at a low frequency and mostly only under conditions of host organ damage. The finding of stem cell plasticity is unorthodox and unexpected (since adult stem cells are considered to be organ/tissue-specific), but it carries significant implications for potential cell therapy. For example, if differentiation can be redirected, stem cells of abundant source and easy access, such as blood stem cells in bone marrow or umbilical cord blood, could be used to substitute stem cells in tissues that are difficult to isolate, such as heart and nervous system tissue. However, the concept of plasticity has been the subject of controversy. The observed frequency of lineage conversion is generally low. An alternative explanation to plasticity is the phenomenon of fusion of host and donor cells. Recent findings suggest that blood cells contribute to other tissues by fusing with preexisting cells rather than by converting to other cell lineages.

Potential clinical applications. Stem cells hold great potential for developing cell therapies to treat a wide range of human diseases. Already in clinical use is blood stem cell transplant therapy, well known as bone marrow transplant therapy for the treatment of patients with certain types of blood diseases and cancers. The discovery of stem cells in various adult tissues, stem cell plasticity, and human embryonic stem cells brings new excitement and opportunities. Stem cells offer the possibility of cell replacement therapy for many human diseases, such as Parkinson's and Alzheimer's diseases, spinal cord injury, diabetes, heart disease, and arthritis, that result from loss or damage of cells in a specialized tissue of the body. Stem cell therapy might revolutionize the treatment and outcome of these diseases. Stem cell science is still in the very early stage. Much more research is required to understand the biological mechanisms that govern cell differentiation and to identify factors that direct cell specialization. Future cell therapy will depend largely on advances in the understanding of stem cell biology and the ability to harness the process of stem cell growth and differentiation.

Somatic cell nuclear transfer (SCNT) stem cells. SCNT involves a micromanipulation procedure in which the nucleus of an egg is removed and replaced by a nucleus taken from somatic cells, typically skin cells. Successful nuclear transfer requires reprogramming of the donor nucleus. The cells so created may divide in cultures to generate embryonic stem cells that can initiate embryogenesis. This is the technique being used in cloning animals, such as the first cloned mammal, Dolly the sheep. However, cloning by nuclear transfer is observed with extremely low efficiency, probably due to faulty and incomplete reprogramming of the donor nucleus. The mechanisms governing the transition from a differentiated genome to a totipotent state remain largely unknown.

A major interest in SCNT is the prospect of creating patient-specific embryonic stem cells. These cells would be genetically identical to the nuclear donor except for maternal mitochondrial deoxyribonucleic acid (mtDNA) of the oocyte. Therefore, the problem of graft rejection would be avoided if the cells could be used for transplant therapy for the donor patients. The concept of using SCNT to generate customized stem cells for cell therapy is also referred to as therapeutic cloning. However, there are hurdles and limitations to using embryonic stem cells in clinical applications. A major challenge is to achieve the directed differentiation and controlled growth before stem cells can be used for transplant therapy. Another issue on SCNT in human stem cells is the sourcing of human eggs. The procedure requires a large number of eggs from women, and poses an ethical and technical challenge. *See* HISTOCOMPATIBILITY.

The success in producing embryonic stem cell lines by the SCNT technique has been demonstrated in mice. In an article published in *Science* in 2005, a team led by Hwang Woo Suk of South Korea claimed the establishment of patient-specific stem cell lines by using the SCNT technique. However, the paper was later retracted as the results were fabricated and the claim a fraud. The field is still left uncertain if somatic nuclear replacement is feasible in humans.

Ethical and regulatory issues. The use of human embryonic stem cells raises ethical, social, and legal issues. The major concern centers on the source of stem cells. Human embryonic stem cell lines are made from the inner cell mass of a blastocyst stage embryo. Most embryos used to produce stem cells are left over from in vitro fertilization (IVF) treatment. The embryos are destroyed by the procedure of extracting stem cells. The early embryo has the biological potential to develop into a person. However, society has not reached consensus on when human life begins. The attention on stem cell research and cloning calls for regulation and legislation from governments. In the United States, current policy allows federal funds to be used for research only on existing human embryonic stem lines. The human embryonic stem cell lines that meet the eligibility criteria are

listed in the Human Embryonic Stem Cell Registry by the National Institutes of Health (NIH).

One concern about SCNT is that it may lead to the reproductive cloning of humans. In theory, the embryo created via SCNT could be used to clone a human if it were implanted into a womans uterus. In the United States, the legislators in the House of Representatives and the Senate have introduced bills proposing a ban of all forms of cloning, including research cloning, or inhibiting reproductive cloning while preserving therapeutic cloning research. However, these bills have not been passed, and no federal law has been established on human embryonic stem cell research. The Canadian Parliament has passed Bill C-6 that prohibits creation of a human clone, sale of sperm or ova, and commercial surrogacy. The bill permits the use of stem cells obtained from discarded products of in vitro fertilization, that is, excess and unused embryos. In the United Kingdom, a law permits the use of embryos in research and therapeutic cloning research but bans reproductive cloning, and implanting a cloned embryo in a human uterus is liable to criminal prosecution. Cloning research must be licensed from the Human Fertilization and Embryology Authority that governs embryonic and stem cell research in the United Kingdom. Chen Wang

Bibliography. D. Kennedy, Editorial retraction [retraction of W. S. Hwang et al., *Science*, 308(5729):1777–1783, 2005], *Science*, 311(5759):335, 2006; J. Kurtzberg, A. D. Lyerly, and J. Sugarman, Untying the Gordian knot: Policies, practices, and ethical issues related to banking of umbilical cord blood, *J. Clin. Investig.*, 115(10):2592–2597, 2005; E. A. McCulloch and J. E. Till, Perspectives on the properties of stem cells, *Nat. Med.*, 11(10):v–vii, 2005; J. M. Polak and A. E. Bishop, Stem cells and tissue engineering: Past, present, and future, *Ann. N. Y. Acad. Sci.*, 1068:352–366, 2006; J. Rossant, Stem cells from the mammalian blastocyst, *Stem Cells*, 19:447–482, 2001; G. J. Spangrude, S. Heimfeld, and I. L. Weissman, Purification and characterization of mouse hematopoietic stem cells, *Science*, 241:58–62, 1988; J. A. Thomson et al., Embryonic stem cell lines derived from human blastocysts, *Science*, 282:1145–1147, 1998; I. L. Weissman, D. J. Anderson, and F. Gage, Stem and progenitor cells: Origins, phenotypes, lineage commitment, and transdifferentiations, *Annu. Rev. Cell Dev. Biol.*, 17:387–403, 2001.

Stenolaemata

A class of Bryozoa confined to fully marine water. The Stenolaemata include several thousand species distributed among four orders: Cystoporata, Trepostomata, Cryptostomata, and Cyclostomata (=Tubuliporata). The commonly recognized order Fenestrata likely is part of the Cryptostomata. First appearing late in the Early Ordovician, stenolaemates expanded quickly to dominate bryozoan assemblages through the Early Cretaceous. By mid-Cretaceous, the calcified gymnolaemate order Cheilostomata, which originated in the Late Jurassic, had increased in diversity and abundance to equal that of stenolaemates, eventually surpassing the stenolaemates, which are now a minor part of the bryozoan fauna in most marine environments. *See* BRYOZOA; CRYPTOSTOMATA; CYCLOSTOMATA (BRYOZOA); CYSTOPORATA; TREPOSTOMATA.

Colony morphology. Stenolaemate colonies vary greatly in size and shape. Many Paleozoic through mid-Mesozoic forms were large and massive, but more recent representatives are commonly small and delicate. Colonies are encrusting, erect, or free-living. Some erect forms have single or regularly spaced, flexible cuticular joints; segments between joints and all other entire colonies are rigidly calcified and enclosed within a thin cuticular membrane. The outer cuticle is secreted by a thin layer of epithelial cells that on colony basal (underside) surfaces switch secretion from cuticle to calcite crystals bound within ultrathin cuticular sheets. This may also happen locally on other surfaces of colonies, generating fixed-wall regions. In actively growing zones and over entire nonbasal surfaces of many colonies, all skeleton above the basal surface is secreted by single layers of epithelial cells that are ultimately folded in from the basal wall, with a fluid-filled gap between the skeleton and the outer cuticle, generating free-walled colonies or regions.

Zooidal morphology. Living stenolaemate colonies are made up of individual feeding units called zooids with lophophores (rings of tentacles) that are circular in basal outline (see **illustration**). Lophophores of feeding zooids bear 8 to fewer than 20 tentacles that flare into usually simple conical or bell shapes when in the feeding position. The lophophore barely clears the skeletal aperture (terminal opening) when fully protruded, with the mouth most commonly at or just below the skeletal aperture. Tentacles of living stenolaemates (Cyclostomata) have lateral cilia tracts, but unlike other bryozoans lack frontal cilia.

Zooidal body cavities are enclosed within tubular or prismatic zooecia (skeletons of individual zooids), gradually expanding from their proximal ends. The outer cuticle folds into each zooid as a vestibular wall and generally attaches by ligaments to the skeleton

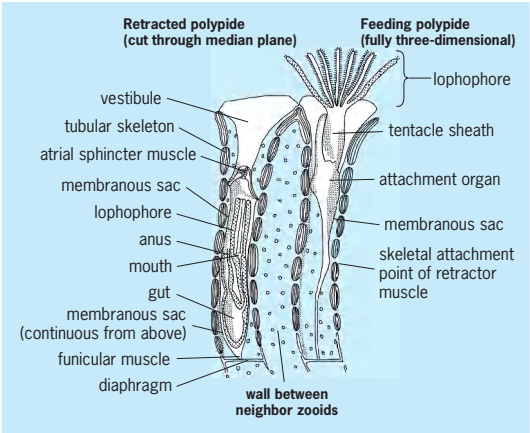

Some morphologic features of autozooids of Cyclostomata, the only living order of stenolaemates.

at some distance below the zooecial aperture. In all but one Mesozoic group, the zooidal orifice (wall opening through which lophophore and tentacles protrude) is closed by a ring of muscles, instead of a covering flap (epistome or operculum). Two membranes continue inward from the ligament, one extending as a sleeve-shaped sheet of ectoderm and mesoderm attached at its inner end around the base of tentacles. The other inward-extending membrane (the membranous sac) divides the body cavity and surrounds the gut, and consists only of mesodermal cells, interpolated annular muscle cells, and a basement membrane. Contraction of the muscle cells of the membranous sac appears to be critical in protrusion of the lophophore to its feeding position.

Zooecia typically are relatively thin-walled in colony interiors (endozone) and relatively thick-walled in outer regions (exozone). Inner surfaces of zooecia are lined by a single layer of epithelium, and it has been hypothesized that as a prerequisite for calcification of the tubular zooecia, the outer body wall of the stenolaemate ancestor split into the inner, mesodermal membranous sac and the outer, calcite-secreting ectoderm. Most Paleozoic stenolaemates had imperforate zooecia, but skeletal walls of most post-Paleozoic forms have pores (except basal walls). Ancient stenolaemates variously had planar transverse partitions (diaphragms), variously curved partitions or blisters, and spines subdividing or projecting into zooecial cavities. Undivided zooecial cavities function as living chambers, otherwise only the portion of the zooecial cavity on the outer side of the last-formed basal diaphragm functions as the living chamber. Living chambers vary across the class from about 2×10^{-3} mm^3 to 1 mm^3.

Polymorphism. Many stenolaemate colonies are constituted of skeletally monomorphic zooids, but polymorphs are present in some. The most common polymorphs are structural (various kenozooids), or females (gonozooids) with brooding chambers. More rarely a few other types of polymorphs occur, such as nanozooids (reduced zooids with single tentacles) and eleozooids (avicularium-like zooids).

Reproduction and growth. Individual colonies are hermaphroditic, some only with skeletally monomorphic, apparently hermaphroditic feeding zooids, and others with female zooids distinguished by chambers that brood polyembryonic embryos (multiple embryos developed by fission from a single fertilized egg). Larvae do not feed and are short-lived. They settle and metamorphose into a hemispherical protoecium from which a cylindrical tube extends to complete the ancestrular zooid (initial zooid of the colony). Colonies grow by asexual budding, and the polypide (organs and tissues) of each zooid is formed by proliferation of cells on the underside of the cuticle that overlies the skeletal orifice. Polypides usually live only a few weeks, after which they degenerate into a roughly spherical, brown-stained mass (brown body). A series of polypides may bud during the life of each zooid. Frank K. McKinney

Bibliography. R. S. Boardman, General features of the Class Stenolaemata, in R. A. Robison (ed.), *Treatise on Invertebrate Paleontology*, pt. G: *Bryozoa* (revised), vol. 1, Geological Society of America and University of Kansas, Boulder, 1989; R. S. Boardman, Reflections on the morphology, anatomy, evolution, and classification of the Class Stenolaemata (Bryozoa), *Smithson. Contrib. Paleobiol.*, 86:1–59, 1998; R. S. Boardman, A. H. Cheetham, and A. J. Rowell (eds.), *Fossil Invertebrates*, Blackwell Scientific Publications, Palo Alto, 1987; A. Ernst, Bryozoans, in B. D. Webby et al. (eds.), *The Great Ordovician Biodiversification Event*, pp. 147–156, Columbia University Press, New York, 2004; C. Nielsen and K. J. Pedersen, Cystid structure and protrusion of the polypide in *Crisia* (Bryozoa, Cyclostomata), *Acta Zool.* (Stockholm), 60:65–88, 1979.

Stephanoberyciformes

An order of fishes containing the pricklefishes and whalefishes. This order is recognized by G. D. Johnson and C. Patterson, but interrelationships of the taxa placed within it are debated among ichthyologists. Stephanoberyciformes is recognized as the sister group of all remaining acanthomorphs. It is defined by the following characters: pelvic girdle attached to cleithrum or coracoid of pectoral girdle; skull bone usually exceptionally thin (fragile bones often characteristic of pelagic fishes); subocular shelf absent; supramaxilla absent or reduced; and body shape variable, from elongate to short and rounded. The biology of stephanoberyciformes as a group is poorly known, and many species are known from only a few specimens. These are small deep-water fishes, ranging in total length from 2.1 to 43 cm (0.8 to 17 in.), but most are less than 10 cm (4 in.), and occupying depths to 5300 m (17,400 ft or 3.3 mi). Several species are known to be oviparous (the assumed mode of reproduction of the entire order), with planktonic eggs and larvae. The order comprises nine families, 28 genera, and about 75 described species, plus many others known but yet undescribed. Included in the order are families taken from the Beryciformes, Lampriformes, and the Trachichthyiformes taxa as described by J. A. Moore. *See* DEEP-SEA FAUNA; OSTEICHTHYES; TELEOSTEI.

Melamphaidae (bigscale fishes). Melamphaidae, comprising five genera and 36 species, can be distinguished by the following features: dorsal fin with 1–3 weak spines; pelvic fins thoracic or subthoracic, with 1 spine and 6 to 8 soft rays; caudal fin preceded dorsally and ventrally by 3 or 4 posteriorly directed spines; cycloid scales, often large and deciduous; lateral line absent, or limited to 1 or 2 pored scales; and lengths of 2.1 to 18 cm (0.8 to 7 in.), but most less than 10 cm (4 in.). Some species make daily vertical migrations from 700 m (2300 ft), and others are bathypelagic to depths exceeding 3500 m (11,500 ft). They occur in all oceans, but are unknown in Arctic seas and the Mediterranean Sea.

Stephanoberycidae (pricklefishes). Stephanoberycidae (**Fig. 1**), comprising three monotypic genera, can be identified by the following characteristics: 8 to 11 spines preceding the principal caudal fin rays

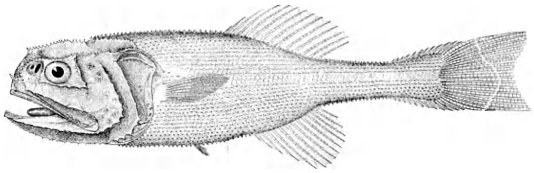

Fig. 1. Pricklefish (*Stephanoberyx monae*). (*From G. Brown Goode and T. H. Bean, Oceanic Ichthyology, plate 55, 1896*)

dorsally and ventrally; dorsal and anal fin spines, if present, weak; 10 to 14 soft rays in each dorsal and anal fin; pelvic fins abdominal or subabdominal, each with 5 soft rays; scales smooth or spiny; and lateral line faint. The three species represent the family in the eastern and western Atlantic (including the Gulf of Mexico), the Indian Ocean off South Africa, and the western and central Pacific. They are 8.1 to 13 cm (3.2 to 5.1 in.) in total length and have a depth range of 1655 to 5300 m (5430 to 17,400 ft).

Hispidoberycidae. Hispidoberycidae is a monotypic family known from the northeastern Indian Ocean and the South China Sea. The sole species, *Hispidoberyx ambagiosus*, has scales covered with small spines, an operculum with a stout spine, a dorsal fin with 4 or 5 spines and 10 soft rays, an anal fin with 3 spines and 9 soft rays, a total length of 18 cm (7 in.), and a depth range of 560 to 1000 m (1840 to 3280 ft).

Gibberichthyidae (gibberfishes). Gibberichthyidae comprises one genus and two species of rare deep-sea fishes. *Gibberichthys latifrons* occurs in the Indo-West Pacific, where adults occupy depths between 750 and 2000 m (2460 and 6560 ft), and *G. pumilus* occurs in the western Central Atlantic (including the Gulf of Mexico) and the western Pacific at depths between 320 and 1100 m (1050 and 3600 ft). Both species are about 13 cm (5.1 in.) in total length. Larvae are usually found at depths less than 50 m (164 ft). The family is distinguishable by the following features: subabdominal pelvic fins, each with 1 spine and 5 or 6 soft rays; a series of 5 to 8 and 4 or 5 partly isolated short spines preceding the soft dorsal and anal fins, respectively; 7 to 9 dorsal and anal soft rays; cycloid scales; flanks with vertical rows of papillae over the vertical lateral line tubes; and swim bladder partially filled with fat. Small crustaceans are known foods of both prejuveniles and adults. *See* SWIM BLADDER.

Whalefishes. The following three families are called whalefishes; however, they are far from whale size, with the species ranging from only 11 to 39 cm (4 to 15 in.). The whalefishes share the following characteristics: whale-shaped body; very large mouth and highly distensible stomach; eyes well developed or degenerate; lateral line of large hollow tubes; luminous tissue on body; dorsal and anal fins far posteriorly and opposite one another; swim bladder absent; and black body with red and orange-tipped fins. Whalefishes are bathypelagic.

Rondeletiidae (redmouth whalefishes). Rondeletiidae is distinguished by a lateral line of 14 to 26 pores in a vertical series, no fin spines, and the presence of pelvic fins. It consists of two species: *Rondeletia bicolor* (**Fig. 2**), from the western Atlantic off North Carolina and Suriname, which has a forward-directed spine above the eye and occurs at depths to 3000 m (9840 ft); and *R. loricata*, which occurs in temperate and tropical waters of the Atlantic, Indian, and Pacific oceans to depths of 1130 m (3700 ft), makes diurnal migrations from 750 m (2460 ft) up to 100 m (328 ft), and is known to feed on isopods and crustacean remains. Both species are only about 14 cm (5.5 in.) in total length.

Barboursiidae (red whalefish). Barboursiidae is a family of only one species, *Barbourisia rufa*, which occurs in tropical and temperate latitudes throughout the world's oceans, including the western Atlantic (Gulf of Mexico) and eastern Pacific (off north-central California). Identifying features are as follows: fin spines absent; large mouth, with the maxilla extending far beyond the eye; lower jaw projecting beyond the upper; jaw teeth in broad villiform bands; body and fins with minute spinules, resulting in a velvety feel; lateral line with distinct pores; maximum total length of about 43 cm (17 in.); and juveniles mesopelagic and adults benthopelagic to depths of 2000 m (6560 ft).

Cetomimidae (flabby whalefishes). Cetomimidae (**Fig. 3**), with nine genera and 35 species (15 of which are known but yet undescribed), plus 36 species in the family Melamphaidae account for 75% of the order Stephanoberyciformes. The following features identify the family: skin loose and without scales; eyes very small or vestigial; no pelvic fins; 3 or 4 pairs of gill arches; photophores absent, but luminous organs often present around anus and dorsal and anal fin bases; and in live specimens brown or orange color with bright orange or red jaws and fins. The very few males that are known are only 3 to 5 cm (1.2 to 2 in.) in length and were previously classified as juveniles. The maximum total length is about 43 cm (17 in.). These fishes are oceanic in the

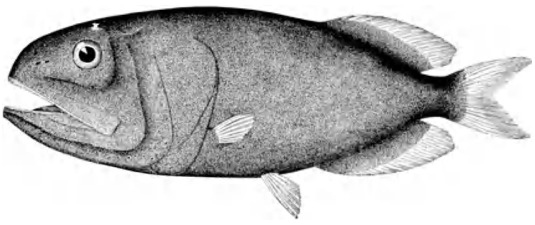

Fig. 2. Redmouth whalefish (*Rondeletia bicolor*). (*From G. Brown Goode and T. H. Bean, Oceanic Ichthyology, plate 77, 1896*)

Fig. 3. Flabby whalefish (Cetomimidae sp.). (*Photo from Alaska Fisheries Science Center, National Oceanic and Atmospheric Administration, National Marine Fisheries Service*)

Atlantic, Indian, and Pacific oceans, and are one of the deepest-dwelling families of the order, with one species, *Cetichthys parini*, at 5000 m (16,400 ft). *See* BIOLUMINESCENCE; PHOTOPHORE GLAND.

Mirapinnidae (tapertails). Mirapinnidae, comprising three genera and five species, occurs in tropical and subtropical zones of the Atlantic and Indian oceans and the western Pacific. The following characteristics identify the family: no scales; fins without spines; dorsal and anal fins opposite one another; dorsal fin with 16 to 33 soft rays and anal fin with 14 to 29 rays; pectoral fin with 13 to 24 rays; and pelvic fin jugular with 4 to 10 rays. Two subfamilies are recognized.

Mirapinninae, with one species, *Mirapinna esau*, is called hairyfish because of the short hairlike pile covering the body. It is further identified by pectoral fins inserted very high on the body, broad pelvic fins inserted in the jugular position well above the ventral profile, and overlapping halves of the caudal fin.

Eutaeniophorinae has two genera and four species. The body is eel-like, with the dorsal and anal fins placed far posteriorly and opposite one another.

Megalomycteridae (longnose fishes). Megalomycteridae comprises four genera and eight species, five of which are yet undescribed. The longnoses occupy depths between 1500 and 4000 m (4900 and 13,000 ft) in the Atlantic and Pacific oceans. They are identified by the following: exceptionally large olfactory organs; lack of fin spines; usually lacking pelvic fin, but if present thoracic with 1 to 3 rays; and less than 6 cm (2.4 in.) in total length. Herbert Boschung

Bibliography. A. W. Ebeling, Melamphaidae, pp. 427–432, in M. M. Smith and P. C. Heemstra (eds.), *Smiths' Sea Fishes*, Springer-Verlag, Berlin, 1986; A. W. Ebeling and W. H. Weed, Order Xenoberyces (Stephanoberyciformes), in *Fishes of the Western North Atlantic*, Mem. Sears Found. Mar. Res., Mem. 1 (pt. 6):397–478, 1973; R. Froese and D. Pauly (eds.), FishBase, World Wide Web electronic publication, version (05/2005); G. D. Johnson and C. Patterson, Percomorph phylogeny: A survey of acanthomorphs and a new proposal, *Bull. Mar. Sci.*, 52(1):554–626, 1993; J. A. Moore, Phylogeny of the Trachichthyiformes (Teleostei: Percomorpha), *Bull. Mar. Sci.*, 52(1):114–136, 1993; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; J. R. Paxton, Barbourisiidae (p. 1740), Cetomimidae (pp. 1171–1173), Mirapinnidae (pp. 1174–1175), in K. E. Carpenter (ed.), *The Living Marine Resources of the Western Central Atlantic*, FAO Species Identification Guide for Fishery Purposes, vol. 2, FAO, Rome, 2003; J. R. Paxton, G. D. Johnson, and T. Trnski, Larvae and juveniles of the deepsea "whalefishes" *Barbourisia* and *Rondeletia* (Stephanoberyciformes: Barbourisiidae, Rondeletiidae), with comments on family relationships, *Rec. Aust. Mus.*, 53:407–425, 2001; J. R. Paxton and T. Trnski, Rondeletiidae (pp. 1168–1169), Megalomycteridae (pp. 1176–1177), in K. E. Carpenter (ed.), *The Living Marine Resources of the Western Central Atlantic*, FAO Species Identification Guide for Fishery Purposes, vol. 2, FAO, Rome, 2003.

Stepping motor

An electromagnetic incremental-motion actuator which converts digital pulse inputs to analog output motion. The device is also termed a step motor. When energized in a programmed manner by a voltage and current input, usually dc, a step motor can index in angular or linear increments. With proper control, the output steps are always equal in number to the input command pulses. Each pulse advances the rotor shaft one step increment and latches it magnetically at the precise point to which it is stepped. Advances in digital computers and particularly microcomputers revolutionized the controls of step motors. These motors are found in many industrial control systems, and large numbers are used in computer peripheral equipment, such as printers, tape drives, capstan drives, and memory-access mechanisms. Step motors are also used in numerical control systems, machine-tool controls, process control, and many systems in the aerospace industry. *See* COMPUTER GRAPHICS; COMPUTER NUMERICAL CONTROL; COMPUTER PERIPHERAL DEVICES; CONTROL SYSTEMS; DIGITAL-TO-ANALOG CONVERTER; PROCESS CONTROL.

Major types. There are many types of step motors. Most of the widely used ones can be classified as variable-reluctance, permanent-magnet, or hybrid permanent-magnet types. These three categories can be compared with regard to various properties (see **table**).

Variable-reluctance motor. A variable-reluctance step motor is simple to construct. The stator and rotor have different numbers of teeth. For example, a three-phase, single-stack motor can have 12 stator teeth and 8 rotor teeth, and be wound in a four-pole configuration (**Fig. 1**). When phase A is energized

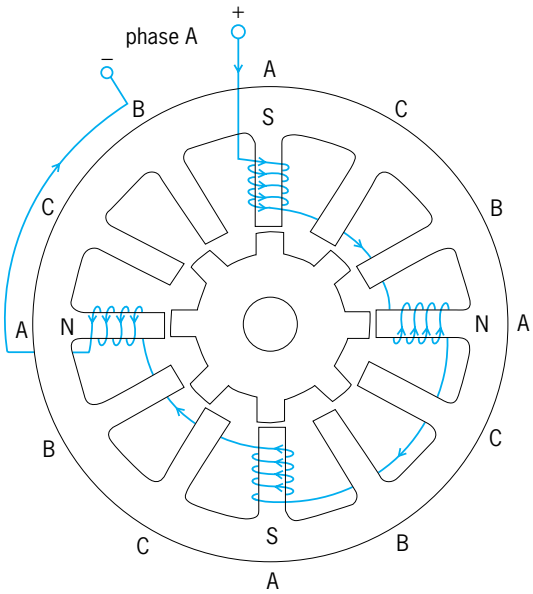

Fig. 1. Three-phase, single-stack, variable-reluctance step motor with 12 stator teeth and 8 rotor teeth, wound in a four-pole configuration. Only the windings on phase A are shown. A, B, C are phases; N and S are north and south magnetic poles.

Comparison of major types of step motors

Property	Variable-reluctance	Permanent-magnet	Hybrid permanent-magnet
Complexity	Simple construction		Relatively complex construction
Performance and efficiency	Low efficiency; requires relatively high current per unit torque	Low performance due to high inductance and stamped motor parts	Efficient; low current per unit torque requirement
Windings and drives	Unifilar windings and unipolar driver	Bifilar windings and unipolar driver; or unifilar windings and bipolar driver	Bifilar windings and unipolar driver; or unifilar windings and bipolar driver
Detent	No inherent detent	Inherent detent	Inherent detent
Rotor inertia	Low rotor inertia due to lack of a permanent magnet	High rotor inertia due to permanent magnet	High rotor inertia due to permanent magnet

by a dc voltage across the terminals, four of the rotor teeth will line up with the four stator teeth of the phase. When phase B is energized and phase A is deenergized, the rotor will rotate clockwise by 15° so that the other four rotor teeth will line up with the four stator teeth of phase B. Energizing phase C will cause the rotor to rotate counterclockwise by 15°. The principle of operation of the single-stack variable-reluctance step motor requires that some of the rotor teeth seek the minimum-reluctance path with the stator teeth of the energized phase. *See* RELUCTANCE.

Permanent-magnet motor. The permanent-magnet step motor (**Fig. 2**) has a rotor which is radially magnetized. Each stator section may have a single winding (unifilar winding) or a winding with a center tap, in which case the winding is called bifilar. A unifilar-wound motor must be driven by a bipolar drive; that is, the drive current must be reversed back and forth to advance the rotor. In contrast, a bifilar-wound permanent-magnet motor can be driven with a unipolar drive.

Hybrid permanent-magnet motor. This type of step motor also has a permanent-magnet rotor. Hybrid refers to the motor's operation under the combined principles of permanent-magnet and variable-reluctance step motors. The motor was originally designed as a two-phase ac synchronous motor for low-speed applications. It was found that the motor could also step incrementally if the phase windings are excited with dc pulses. The rotor of this step motor (**Fig. 3**) has a permanent magnet positioned in the axial direction to produce a north-south unidirectional magnetic field. The teeth on the two sections of the rotor are offset by one-half of a rotor tooth pitch. In general, the stator teeth are all uniformly distributed with a tooth pitch which is different than that of the rotor. The stator may have unifilar two-phase windings or bifilar windings. In the bifilar case, each phase has two windings, and the motor is often referred to as a four-phase motor.

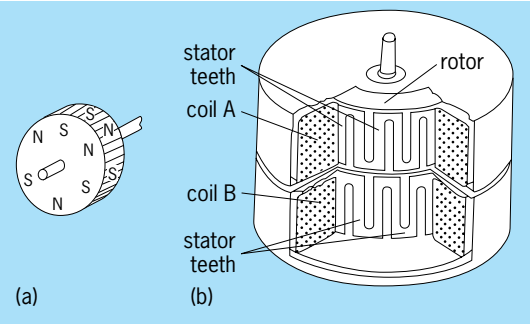

Fig. 2. Two-phase permanent-magnet step motor. (*a*) Rotor, magnetized with six radial poles. N and S are north and south magnetic poles. (*b*) Cutaway view.

Winding configurations. Most variable-reluctance step motors are wound with bifilar windings. Permanent-magnet and hybrid permanent-magnet step motors can be wound with either bifilar or unifilar windings.

Unifilar windings. A unifilar winding is a single coil wound on a pole of a motor. When a direct current is sent through the winding, a flux is produced in the pole if the magnetic path is closed. This produces a magnetic polarity at the pole. If the direction of the current is reversed, the direction of the flux will also be reversed. Thus, in order to reverse the polarity of the magnetic pole, with a unifilar winding, the direction of the current must be reversed. This requires a bipolar driver, that is, a driver that can reverse polarity between positive and negative.

Bifilar windings. A bifilar winding consists of two windings in opposite sense on the same pole. When a driver sends a current into one coil, a magnetic polarity is induced on the pole. By applying the same current to the other coil instead, the magnetic polarity of the pole is reversed. Thus, with a bifilar winding the magnetic pole polarity can be reversed without reversing the polarity of the drive voltage. The advantage is that only a single-ended, or unipolar, power supply is needed. A disadvantage of the bifilar winding is that only one-half of the total

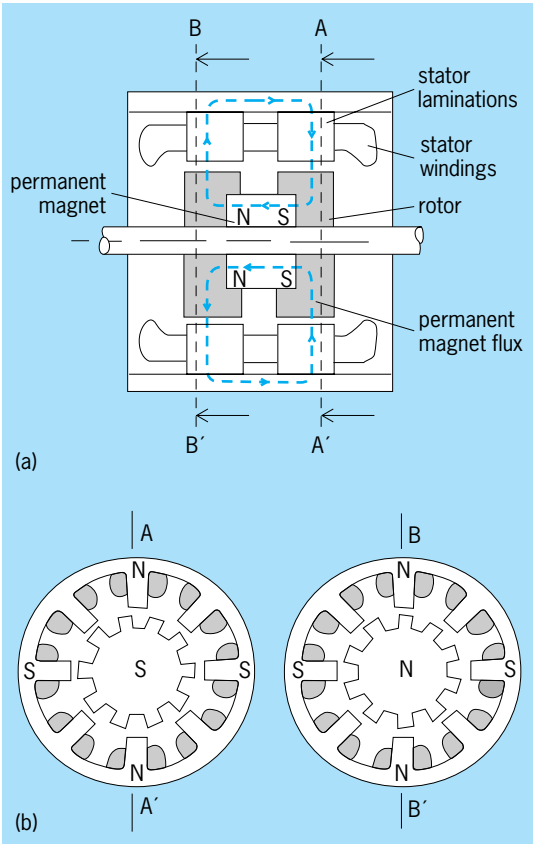

Fig. 3. Hybrid permanent-magnet step motor with eight-pole stator. N and S are north and south magnetic poles. (*a*) Axial cross section. (*b*) Radial cross sections at A-A′ and B-B′ on part *a*.

windings are used at any time so that for a given frame size the motor is derated. However, bifilar windings can be connected in series or in parallel so that the entire windings can be utilized. Since the operations of permanent-magnet and hybrid permanent-magnet step motors depend on the reversal of the magnetic polarities of the stator poles, these motors must have either bifilar windings with a unipolar driver or unifilar windings with a bipolar driver. *See* WINDINGS IN ELECTRIC MACHINERY.

Torque. Step motors are torque-generating devices. The torque can be classified as static and dynamic. When a phase of a variable-reluctance step motor is energized with a direct current, a static holding torque is developed. As the rotor is forced to rotate, ideally a sinusoidal static holding torque curve is generated for this phase. Similarly, torque is generated when the other two phases of a three-phase motor are energized in turn. The equilibrium positions are also the detent positions of the individual phases. In practice, the static holding torque may not be sinusoidal, depending on the design of the motor. A variable-reluctance motor must have at least three phases to rotate continuously.

The torque components of a hybrid permanent-magnet motor are complex. The permanent magnet in the rotor generates a fourth-harmonic detent torque when the stator windings are not energized relative to the torque generated when the stator windings are energized.

The pull-out torque of a step motor represents the dynamic performance of the motor. It is defined as the maximum frictional torque that can be driven by the motor at a given speed. The pull-out torque (torque-speed) characteristic of a step motor is similar in shape to that of a dc motor. *See* DIRECT-CURRENT MOTOR; TORQUE.

Drivers. A driver stage for a variable-reluctance step motor has a power transistor for each motor phase. To energize a certain motor phase, a logic sequencer simply sends out a signal to turn on the corresponding power transistor so that a current supplied by the power source will flow through the motor phase windings. To drive the motor in a particular sequence, the phases are energized accordingly one at a time. The phases can also be energized two at a time.

The unipolar driver for a variable-reluctance step motor can also be used for permanent-magnet and hybrid permanent-magnet step motors with bipolar windings. As noted above, a bipolar driver is required for unifilar motor windings.

The power transistors of the step motor driver can be controlled by a chopper driver to control the average current to the motor windings. Typically, the motor is switched to a high supply voltage, several times the rated value, causing the current to build up rapidly. A current-sensing network turns the high voltage to a low voltage when the current reaches some preset threshold. The frequency of chopping can range from 1 to 30 kHz. *See* CHOPPING; MOTOR; SERVOMECHANISM.

Benjamin C. Kuo

Bibliography. P. P. Acarnley, *Stepping Motors: A Guide to Modern Theory and Practice*, 2d ed., 1984; T. Kenjo and A. Sugawara, *Stepping Motors and Their Microprocessor Controls*, 2d ed., 1995; B. C. Kuo (ed.), *Proceedings of the Incremental Motion Control Systems and Devices Symposium*, 1972–1995; B. C. Kuo (ed.), *Step Motors and Control Systems*, 1979.

Stereochemistry

The study of the three-dimensional arrangement of atoms or groups within molecules and the properties which follow from such arrangement. Molecules that have identical molecular structures (that is, the same kind, number, and sequential arrangement of atoms) but differ in the relative spatial arrangement of component parts are stereoisomers. Inorganic and organic compounds exhibit stereoisomerism. Examples are structures (**1**) to (**8**). In the example, pairs

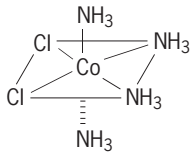

(**1**)

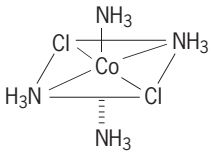

(**2**)

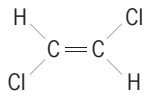
(3)

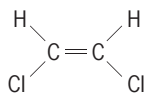
(4)

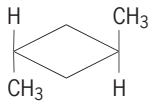
(5)

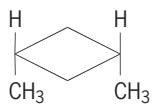
(6)

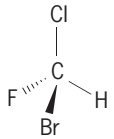
(7)

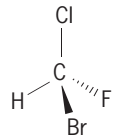
(8)

of stereoisomers are related by permutation (transposition) of bonded atoms or groups. In structures (**3**) and (**4**), a single permutation of Cl and H about either C atom produces two stereoisomers. Atom C_2 may be called a stereogenic atom (or a stereocenter).

Significance. Stereochemistry has played a significant role in the historical development of theories of molecular structure. The optical activity of substances such as sugar and turpentine was discovered by J. B. Biot in 1813, and L. Pasteur was the first to separate or resolve enantiomers (nonsuperimposable mirror images) from one another (1848). While Pasteur recognized that enantiomers must differ in symmetry, it was not until 1874 that J. H. van't Hoff and J. A. LeBel were able to relate symmetry, structure, and properties into the hypothesis of the tetrahedral carbon which formed the cornerstone of modern stereochemistry. The foundation of inorganic stereochemistry rests on the coordination theory of A. Werner (1893).

Since the latter part of the nineteenth century, stereochemical studies have been prominent in the evolution of both inorganic and organic chemistry. These studies have concerned themselves with the determination of configuration, the interconversion of diastereomers and the assessment of their energy differences, and conformational analysis.

Methods for the analysis and separation of stereoisomers have been devised, and stereochemical principles have been applied to the elucidation of reaction mechanisms, to the development of asymmetric syntheses, and to attempts to understand biological processes at the molecular level. *See* ASYMMETRIC SYNTHESIS; COORDINATION CHEMISTRY; RACEMIZATION.

Symmetry. The nature and the number of stereoisomers of a molecule are determined by the permutation of atoms or groups (called ligands) at stereocenters as well as by the symmetry of the molecule. The symmetry elements to be considered are: planes of symmetry, axes of symmetry, centers of symmetry, and rotation-reflection (or alternating) axes of symmetry. Two types of stereoisomers are known. Those such as (**7**) and (**8**), which are devoid of reflection symmetry—which cannot be superposed on their image in a mirror—are called enantiomers. All other stereoisomers, such as the pairs (**1**)–(**2**), (**3**)–(**4**), and (**5**)–(**6**), are called diastereomers. The configuration of a stereoisomer designates the relative position of the atoms associated with a specific structure. The structures of stereoisomers (**1**) and (**2**) differ only in configuration. The same is true for (**3**) and (**4**), (**5**) and (**7**), and (**7**) and (**8**).

Enantiomers are related to one another as a left hand is to a right hand. Such structures are said to be chiral (Greek *cheir* = hand) or dissymmetric. Unlike diastereomers, which differ in most physical and chemical properties, enantiomers have identical properties other than the sign (+ or −) of their optical activity. This identical behavior exists toward all agents and processes which are themselves achiral. Chiral agents act differently toward enantiomers, however. Biological specificity toward enantiomers is dramatic; for example, the enantiomers of the amino acid leucine have different tastes: one is bitter, while the other is sweet. *See* BIOLOGICAL SPECIFICITY; MOLECULAR ISOMERISM.

Configuration. In order to understand chemical processes involving stereoisomers, it is necessary to know the configurations of reactants and products. The configuration of stereoisomers (**1**) and (**2**) are designated cis (chlorines on the same side) and trans (chlorines on opposite sides), respectively; (**3**) and (**5**) are trans; (**4**) and (**6**) are cis. Configurations of diastereomers can often be determined from their physical properties. For example, the isomer of $C_2H_2Cl_2$ with zero dipole moment is (**3**), while (**4**) has a finite dipole moment. *See* DIPOLE; DIPOLE MOMENT.

The absolute configuration of chiral molecules such as (**7**) and (**8**) is the actual order of atoms about the stereogenic atom (tetrahedral carbon atom) which defines the specific enantiomer. While the three-dimensional picture of a molecule defines its stereochemistry just as the picture of a hand defines its handedness (right or left), it is not convenient or useful to rely on full three-dimensional representations in most cases, and a shorthand notation to designate configuration is widely accepted. Groups and atoms which surround the stereocenter are assigned priorities related by unambiguous rules to atomic number, for example, Br > Cl > F > H, and to substitution pattern. The molecule is viewed from the tetrahedral face opposite the carbon substituent with the lowest priority, that is, from the face opposite hydrogen in the case of (**7**). The counterclockwise descent (**7*a***) from high- to low-priority substituent (Br followed by Cl followed by F) is designated S (Latin *sinister* = left). Enantiomer (**8**) exhibits a similar but clockwise order (**8*a***) designated R (Latin

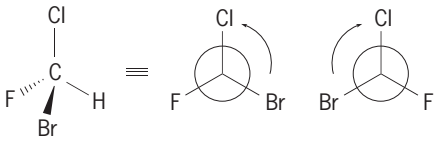
(7) (7*a*) (8*a*)

rectus = right). When the order of groups immediately bound to the stereocenter is equivocal, the priority is determined by atoms or groups attached to the atoms whose priorities are equivocal.

The same priorities may be used in the designation of the configurations of diastereomers such as (**3**) and (**4**). When the atoms or groups of highest priority lie on opposite sides of the double bond, as in (**3**), the stereoisomer is designated E (German *entgegen* = opposite). Isomer (**4**) in which the reference atoms are on the same side of the double bond are termed Z (German *zusammen* = together). This system is less ambiguous than the cis-trans system which is still applicable in simple cases.

Enantiomers are characterized by the sign of their optical activity at a given wavelength, temperature, and specified solvent. The assignment of configuration to specific enantiomers is carried out by chemical transformations of known stereochemistry, by spectroscopic means such as circular dichroism and optical rotatory dispersion, and by diffraction of single crystals employing anomalous x-ray scattering. From such studies it is known that the lactic acid isomer that is experimentally found to be levorotatory, that is, whose optical activity is (−)- or counterclockwise-rotating, has structure (**9**) corresponding

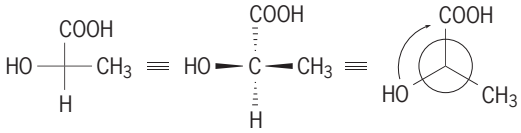

(**9*b***) (**9**) (**9*a***)
[priority OH > COOH > CH_3]

to the R configuration (**9*a***). Configuration has no simple connection to the sign of the optical rotation. *See* COTTON EFFECT; DICHROISM; OPTICAL ACTIVITY; OPTICAL ROTATORY DISPERSION; X-RAY DIFFRACTION.

Two-dimensional (planar) projections of three-dimensional structures are very commonly used to represent configurations. Projection formulas compress the tetrahedral geometry into the plane of the paper. In (**9**) the horizontal atoms or groups are forward of the plane (wedges) of the stereogenic atom, and the vertical groups are behind (solid or broken lines). Tipping formula (**9**) forward places the COOH group in the plane of the paper and the H atom behind, corresponding to structure (**9*a***). The planar projection of (**9**) may be written as (**9*b***). To avoid confusion in their use, the following conventions apply to planar projections: they must not be rotated in this plane through any but integral multiples of 180°; an odd number of exchanges of any pair of substituents is equivalent to transformation into its mirror image, and an even number of exchanges leaves the configuration unchanged.

The relative configurations of stereoisomers applicable to diastereomers are illustrated by the tartaric acids (**10**) and (**11**). Structure (**10**) represents (+)-tartaric acid (both stereocenters ∗ have R configurations), and (**11**) represents *meso*-tartaric acid, the diastereomer which is optically inactive

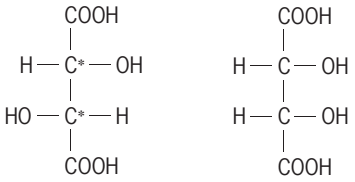

(**10**) (**11**)

(the two stereocenters, one R and one S, are mirror images of one another. Older configurational symbolism (D and L) involving intermolecular comparisons to reference substances such as glyceraldehyde and serine persists, but it is giving way to the R-S designations.

Stereoregular polymers with defined configurations at stereogenic atoms or at carbon-carbon double bond stereocenters differ substantially in properties from their diastereomers. Polymeric carbohydrates such as starch and cellulose and hydrocarbons such as rubber and gutta-percha exemplify natural diastereomer pairs. Synthetic polystyrene diastereomers (**12**) [isotactic] and (**13**) [syndiotactic] differ only in the configuration at stereogenic atoms. The high-melting, fiber-forming isomer is isotactic polystyrene (**12**).

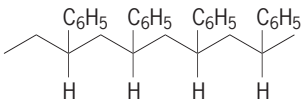

(**12**)

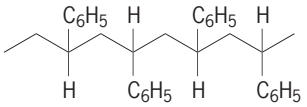

(**13**)

Configurational studies are powerful probes in the elucidation of reaction mechanisms. Replacement of substituent atoms or groups in a reaction is often attended by a change in configuration at the stereocenter (Walden inversion), or it may result in the loss of optical activity (racemization). In some cases retention of configuration prevails. Such results provide evidence for the geometry of transition states or for the intervention of intermediates (ions or radicals).

Resolution. Most chiral substances occur in nature as only one enantiomer. Few natural products are represented in nature by both enantiomers, and few also exist as mixtures with equal proportions of enantiomers called racemates. On the other hand, synthesis of the chiral substances in the laboratory or in industry in the absence of chiral agents or conditions results in racemates which exhibit no measurable optical activity, since the activity of one enantiomer cancels that of the other.

The separation of one or both enantiomers from a racemate, called resolution, requires the intervention of an optically active reagent, catalyst, or other chiral influence. The most common resolution

procedure involves the reversible conversion of an enantiomer mixture into a pair of diastereomers, with as widely different physical properties as practicable, by reaction with a single pure enantiomer. For example, a (±)-amine reacts with a (+)-acid to give two diastereomeric salts: (+)-$RNH_3^{\oplus}$(+)-$R'COO^{\ominus}$ and (−)-R-$NH_3^{\oplus}$(+)-$R'COO^{\ominus}$.

These diastereomeric salts may be separated by fractional crystallization and may then be converted to the individual enantiomeric amines by cleavage with a strong base. The (+)-acid resolving agent may be recovered and reused. Covalent diastereomers may also be separated by crystallization or increasingly by gas, liquid, or thin-layer chromatography.

Kinetic resolution takes advantage of differences in rates of reaction of enantiomers with chiral reagents. Enzymes are chiral catalysts which can preferentially catalyze reactions of one enantiomer. Enzymatic resolutions play a major role in the syntheses of substances of biochemical importance. Chromatographic resolution on chiral stationary phases, which bind or dissolve one enantiomer more strongly than the other, constitutes another useful type of resolution.

In resolution by preferential crystallization or entrainment, one enantiomer preferentially crystallizes upon seeding of supersaturated solutions with seed crystals of that enantiomer. Though relatively few racemates are amenable to this type of resolution, it is nevertheless of considerable importance on an industrial scale. *See* CHROMATOGRAPHY; CRYSTALLIZATION.

Conformational analysis. Molecules are not rigid collections of atoms. Torsional stereoisomerism arises as a consequence of the rotation (torsion) of atoms about bonds within molecules. This gives rise to stereoisomers called conformers which are interconvertible by rotation about carbon-to-carbon single bonds. The existence of two extreme forms of ethane (in terms of energy), (**14**) [staggered] and (**15**) [eclipsed], was predicted by K. S. Pitzer in

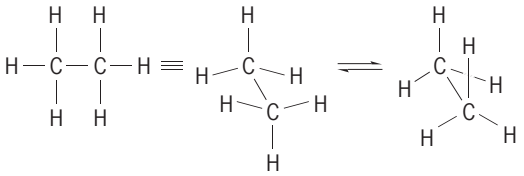

(**14**) (**15**)

1936 and later experimentally verified. An infinite number of forms (conformations) intermediate between the equilibrium (low-energy) conformation (**14**) and the higher-energy conformation (**15**) is possible. At room temperature most ethane molecules resemble (**14**). Though their presence is experimentally readily demonstrated, separation of conformers is normally not possible since the small energy difference between them [3 kcal/mole; 12.6 kilojoules/mole in the case of (**14**), and (**15**) makes their interconversion facile]. The ease of interconversion explains why some enantiomeric structures, such as the *gauche* conformers (large groups, CH_3, 60° apart) of *n*-butane (**16**), are incapable of resolution.

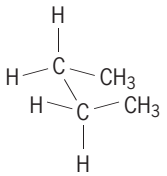

(**16**)

The relationship between the physical and chemical properties of substances and their preferred conformations (conformational analysis) has been the subject of many studies since 1950. Cyclohexane (**17**) exists overwhelmingly in the form of conformer (**18**; the chair form), which is the structural subunit

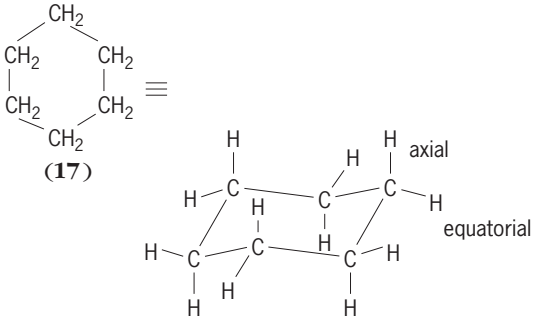

(**17**) (**18**)

from which the carbon atom lattice of diamond is constructed. Substituents on six-membered rings are designated as either axial or equatorial, as in (**18**). Conformational analysis of synthetic hydrocarbon polymers such as polypropylene and biopolymers such as deoxyribonucleic acid (DNA) has led to an understanding of their properties. Much of the biological activity of enzymes, for example, is made possible by the specific conformations adopted by these macromolecules. *See* CONFORMATIONAL ANALYSIS; ENZYME; POLYMER.

Stereocontrolled synthesis. An understanding of the stereochemical consequences of chemical reactions and the determination of configuration of complex natural products has permitted the total synthesis of many stereochemically complex substances since around 1940. For example, cholesterol and gibberellic acid each contain eight stereogenic carbon atoms, which makes possible, in principle, $2^8 = 256$ stereoisomers (128 racemates). Disregard for stereochemical consequences during synthesis would lead to extremely complex mixtures requiring repeated separation and to very low yields.

Living organisms unerringly synthesize (biosynthesis) just one of these isomers in fully stereoselective processes. Laboratory and commercial syntheses of substances which have desirable biological properties have increasingly been devised so as to mimic biosyntheses by taking advantage of stereocontrolled reactions. These are of two types: stereospecific reactions and stereoselective reactions. Stereospecific reactions are those in which different stereoisomers are transformed into stereochemically different products. For example, *Z*-2-butene

(**19**) [cis] is epoxidized to the *cis*-epoxide (**20**) only.

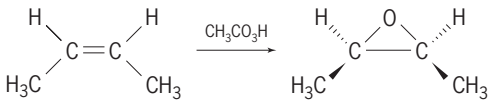

(**19**) (**20**)

The diastereomeric *E*-2-butene [trans] forms only the *trans*-epoxide.

Stereoselective reactions are those in which a single reactant gives rise to a mixture of diastereomers in which one predominates. Those stereoselective reactions in which new stereocenters are formed in unequal amounts are called asymmetric syntheses. Transformations of optically active compounds into more complex substances which incorporate additional stereocenters are also asymmetric syntheses.

Syntheses of natural products which originate in achiral starting materials require at least one resolution step if nonracemic products are desired. Alternatively, a total synthesis may well originate in a single enantiomer of a compound having but one or two stereogenic atoms. Additional stereocenters are introduced during the synthesis through stereospecific and especially through stereoselective steps. Separation of diastereomers at various stages may be required. *See* STEREOSPECIFIC CATALYST.

Samuel H. Wilen

Bibliography. E. L. Eliel, S. H. Wilen, and M. P. Doyle, *Basic Organic Stereochemistry*, 2001; E. Eliel, S. Wilen, and L. Mander, *Stereochemistry of Organic Compounds*, 1994; R. E. Gawley and J. Aube, *Principles of Asymmetric Synthesis*, 1996; M. M. Green et al. (eds.), *Topics in Stereochemistry: Materials-Chirality*, 2003; K. Mislow, *Introduction to Stereochemistry*, 2003; D. G. Morris, *Stereochemistry*, 2002; M. North, *Principles and Applications of Stereochemistry*, 1998.

Stereophonic radio transmission

The transmission of stereophonic audio signals over broadcast radio systems, including terrestrial analog amplitude-modulation (AM), frequency-modulation (FM) systems, and digital radio systems (terrestrial and satellite). *See* STEREOPHONIC SOUND.

Analog systems. Two-channel stereo broadcasting of audio was first implemented for FM radio signals in the 88–108-MHz frequency band (allocated for FM radio broadcasting) in the 1960s, and became the dominant mode of transmission, in particular for music programming, in the 1970s. A multiplexing method was used to accomplish this and allowed the signal transmitted by a stereo broadcast station to be decipherable by either a stereo receiver or a monophonic receiver. For the monophonic receiver to receive a stereo broadcast, some component of the broadcast must be a single signal that includes information from both the left and the right channels. This is called the $L + R$ signal. A monophonic receiver simply demodulates the $L + R$ signal and delivers it to the listener.

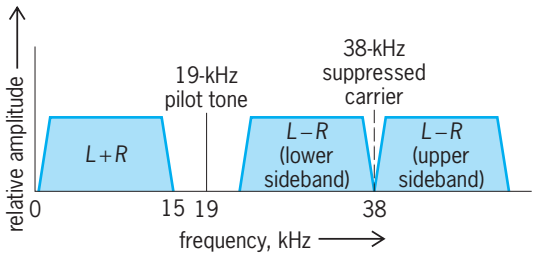

Baseband spectrum of an FM stereo signal. This audio and supersonic spectrum is frequency-modulated onto a radio-frequency carrier.

The stereo receiver must use the $L + R$ signal and other transmitted information to demodulate separate L and R signals. For this, stereo broadcasts include an $L - R$ signal. In the stereo receiver, after detection of the $L + R$ and $L - R$ signals, the L and R signals are separated as shown in Eqs. (1) and (2).

$$(L + R) + (L - R) = 2L \tag{1}$$

$$(L + R) - (L - R) = 2R \tag{2}$$

The subtraction used in Eq. (2) is realized through phase shifting and addition. In FM stereo transmissions, the $L + R$ signal is directly frequency-modulated onto the radio-frequency carrier, assuring compatibility with monophonic systems (see **illustration**). The $L - R$ signal is used to amplitude-modulate a 38-kHz hypersonic subcarrier. The 38-kHz carrier is suppressed, creating a double-sideband (DSB), suppressed-carrier AM signal at 38 kHz. This hypersonic signal is then frequency-modulated onto the radio-frequency carrier. A 19-kHz pilot signal is also transmitted to indicate that the broadcast is stereo and to provide an accurate frequency reference for demodulation of the 38-kHz $L - R$ signal. *See* AMPLITUDE MODULATION; FREQUENCY MODULATION.

In the FM stereo receiver, the entire baseband signal that has been frequency-modulated onto a radio-frequency carrier ($L + R$, pilot, and double-sideband $L - R$) is first detected with an FM discriminator. The $L + R$ signal is isolated by a low-pass filter. The pilot tone is doubled and mixed with the double-sideband suppressed-carrier $L - R$ signal, so that the $L - R$ signal can be demodulated by an AM envelope detector. The $L + R$ and $L - R$ audio signals are then passed through a network that implements the functions expressed in Eqs. (1) and (2). The resulting left- and right-channel audio signals are amplified and presented to the listener. *See* AMPLITUDE-MODULATION DETECTOR; ELECTRIC FILTER; FREQUENCY-MODULATION DETECTOR; FREQUENCY-MODULATION RADIO.

It is also possible to send stereo audio signals using more than two channels over FM radio transmissions, although this technology is not standardized for use in the United States by the Federal Communications Commission (FCC), as is the two-channel technology just described. SRS Circle Surround, Neural Audio, and Dolby Laboratories have all developed proprietary surround sound technologies which are

compatible with two-channel stereo FM broadcasts and which provide up to "6.1" channels of surround sound information (the ".1" represents a seventh, low-frequency channel with frequency response limited to approximately 100 Hz).

Stereo broadcasting for the AM band (550–1701 kHz) was first authorized by the FCC in 1982, when five different competing (and incompatible) systems were authorized because the FCC was unwilling to select just one and chose instead to let the marketplace decide. This proved to be a fateful decision since over the next 11 years the various AM stereo systems competed for market share. In the end, due in part to the confusion created by so many systems being available, AM stereo never became a popular service. In 1993, the FCC revisited their decision and finally selected a single system, called compatible quadrature amplitude modulation (C-QUAM), but receiver manufacturers and broadcasters never fully embraced this technology. In 2002 the FCC authorized a new digital radio service for the AM band, called in-band/on-channel (IBOC) digital radio, which supports stereo audio but is incompatible with C-QUAM, effectively ending the analog AM stereo saga in the United States.

Digital systems. A number of digital radio systems are being implemented around the world, and in most cases these systems support the transmission of not only two-channel stereo but also "surround sound" stereo signals with five or more channels. In 2005, satellite digital audio radio service (SDARS) provider XM Radio began broadcasting some of its music channels in 5.1-channel surround using technology developed by Neural Audio, and a number of terrestrial FM radio broadcasters were also experimenting with surround sound using the IBOC digital radio broadcast system authorized for the FM band by the FCC in 2002. David H. Layer; Jeffrey Ward

Bibliography. Federal Communications Commission, *Rules and Regulations*, Parts 73 and 74, and *Amendments*, 2004; D. G. Fink and W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; G. Jones, *A Broadcast Engineering Tutorial for Non-Engineers*, NAB and Focal Press, 2005; *NAB Engineering Handbook*, 9th ed., 1999; M. E. Van Valkenburg (ed.), *Reference Data for Engineers: Radio, Electronics, Computer, and Communications*, 9th ed., 2002.

Stereophonic sound

A system of sound recording or transmission in which signals are captured, mixed, or synthesized using two or more audio channels in such a way as to deliver a spatial or three-dimensional auditory impression to a listener when these audio channels are connected to loudspeakers in a listening room. Binaural or "head related" stereo is a term used to describe the capture or synthesis of a pair of signals that closely resemble those present at a listener's ears in a natural spatial sound field. Such signals are intended for reproduction using a pair of headphones worn by the listener, although they can be reproduced using loudspeakers if suitable signal processing is employed. *See* BINAURAL SOUND SYSTEM.

Stereo recording techniques were patented in the 1930s, but commercial development was not undertaken at that time. In the late 1930s and early 1940s, a few motion picture films were produced using two optical tracks for stereo playback. A variety of multichannel film systems were introduced in the 1950s. The commercial recording industry introduced two-track stereo tape recordings in 1954 and single-groove, two-channel stereo disks in 1957. The first stereo broadcasts on FM radio were made in 1961. In the 1970s, several quadraphonic (four-channel) systems were introduced, but none achieved lasting commercial success. From the late 1990s, "surround sound" stereo systems involving five or more channels began to achieve commercial success, primarily as an adjunct to home television systems. Wavefield synthesis systems, which reached the market in the early years of the twenty-first century, use very large numbers of loudspeakers in an attempt to recreate acoustical wavefronts in a physically accurate fashion. *See* STEREOPHONIC RADIO TRANSMISSION.

Auditory cues in stereophony. The aim of stereophonic recording and reproduction is to deliver consistent auditory spatial cues to the listener to give rise to an intended spatial impression. It is nonetheless common for stereophony to rely on a degree of auditory illusion for this purpose, as it is not usually possible to reconstruct acoustical sound fields in a physically accurate manner over a wide listening area when using a limited number of loudspeakers.

In order to determine the origin (location) of a sound, the brain considers the differences in sound pressure level and arrival time at each ear. The external ear (pinna) also gives rise to modifications to the frequency spectrum of sound signals, which are unique to each source location and can represent height information as well as horizontal cues. Fluctuations in these cues contribute to a sense of spaciousness or width. *See* HEARING (HUMAN).

Stereophonic signals are created either by capturing the relevant cues using two or more microphones, or by artificially introducing the necessary interchannel differences using audio signal processing. The perceived source images that result from these cues during reproduction are known as phantom images because they are not real sources. In binaural stereo or wavefield synthesis, they are sometimes known as virtual sources.

Stereophonic sound pickup. There are three basic microphone techniques for two-channel stereophonic pickup: coincident, spaced, and individual instrument (also called close miking). The techniques are sometimes combined.

The coincident technique uses two directional microphones located very close together, often enclosed in the same case. The coincident technique relies only on capturing differences in sound level between the channels to encode different source

positions. This can give rise to very precise stereo imaging. *See* MICROPHONE.

Spaced microphone techniques encode spatial cues using both time and level differences between the channels. These differences lead to directional perception based either on the "precedence effect," whereby the sound is localized toward the earlier loudspeaker signal, or on "summing localization," in which the perceived source position is the result of a trade-off between the two cues. Summing localization tends to work with interchannel time delays up to just over 1 millisecond, combined with interchannel level differences of up to 18 dB. Techniques can rely on large time differences and small level differences, or vice versa, but the quality of the stereo image resulting from different combinations is not identical. Techniques relying on summing localization tend to be known as "near-coincident" techniques because of the relatively small microphone spacings involved. Because of the time delay, spaced techniques can cause noticeable phase cancellations in monophonic reproduction.

Since the late 1960s, it has become common for much stereo program material to be recorded using a large number of microphones, each located close to one source or a small group of sources. By proportionally mixing the microphone outputs between the stereo channels, it is possible to provide the interchannel level differences necessary for perceiving phantom images. Artificial effects and reverberation may be added subsequently to create a spatial impression, or signals from microphones located in the reverberant field of a studio may be mixed with the dry signals.

Surround sound recording extends these concepts to more than two channels to create an immersive sound field. However, it is common to find that accurate stereophonic imaging is attempted only in front of the listener, using techniques similar to those described above, while the rear channels are used to enhance the experience of spaciousness and envelopment. This is particularly true of systems based on the International Telecommunications Union-Radiocommunication Sector document ITU-R BS.775 layout, known as 3-2 stereo or 5.1 surround, described below. *See* SOUND RECORDING; SOUND-RECORDING STUDIO.

Stereophonic reproduction. The ability to create convincing stereophonic reproduction is highly dependent upon the performance of the loudspeakers, the acoustics of the listening room, and the hearing process of the listener.

The loudspeakers should be identical and capable of broad frequency response with low distortion. Imaging is degraded by narrow or irregular dispersion of frequencies above about 500 Hz. Most high-quality loudspeaker systems have a horizontal coverage angle of 90 to 120°, above 500 Hz. The vertical angle is usually less, and often is 40 to 60°. The uniformity of the coverage pattern is an important measure of the loudspeakers' quality. All loudspeakers have wider coverage at low frequencies. *See* LOUDSPEAKER.

For two-channel stereo, the ideal arrangement is one where the loudspeakers subtend an angle of ±30° at the listening position. This arrangement is maintained in the ITU BS.775 configuration for 3-2 Stereo. (5.1 surround), in which a center loudspeaker is added between the left and right front loudspeakers, as well as two rear-side loudspeakers at ±110°. The rear-side locations were chosen as a compromise between the requirements for movie applications and those for music reproduction. In movie applications, the capacity to deliver effects from the rear is desirable, whereas for music reproduction the research evidence suggests a need to recreate lateral reflections to optimize envelopment.

The loudspeakers should be located in similar acoustical surroundings. The surfaces near each loudspeaker should be similar in acoustical quality in order that sound reflected from the area of each loudspeaker will be the same for the listener. The reverberation time of the listening room should ideally be less than about 0.5 s. Hard surfaces near the loudspeakers can reflect sound energy, which when delayed only a few milliseconds by its slightly longer path will combine with the direct sound in a frequency-dependent manner to produce irregularities in what should be a smooth frequency response at the listener's position. This "comb filter" effect can be reduced or avoided by making all surfaces around the loudspeakers acoustically very absorbent. In some professional listening rooms, such as recording or broadcast control rooms, this surface treatment extends from the loudspeakers to the listener's position; all wall surfaces behind the listener are made reflective and diffusive.

A modern high-quality loudspeaker system consists of two or more radiating units in the same enclosure, each reproducing only a part of the sound spectrum. These radiators cannot coexist in the same location; therefore, there can be path length differences (hence, time differences) between them and the listener. Because the hearing process is sensitive to short time delays, phantom images of sound sources being partly reproduced by two or more radiators in each loudspeaker will tend to smear or split in space. However, even among trained listeners, not all people can hear, or are distracted by, this effect. For those who are annoyed, several manufacturers have introduced time- (or phase-) aligned loudspeakers. Time alignment requires that path lengths be made equal by physical location of the radiators or by use of an electronic delay in order to synchronize arrival times at the listener's position. *See* SOUND-REPRODUCING SYSTEMS.

Rollins Brook; Francis Rumsey

Bibliography. B. and J. Barlett, *Practical Recording Techniques*, 4th ed, Focal Press, 2005; J. Eargle (ed.), *Stereophonic Techniques: An Anthology*, Audio Engineering Society, 1986; D. Howard and J. Angus, *Acoustics and Psychoacoustics*, 3d ed, Focal Press, 2006; D. M. Huber, *Modern Recording Techniques*, 6th ed., 2005.

Stereoscopy

The phenomenon of simultaneous vision with two eyes, producing a visual experience of the third dimension, that is, a vivid perception of the relative distances of objects in space. In this experience the observer seems to see the space between the objects located at different distances from the eyes. The stereoscopic effect is unique and cannot be easily described to one who does not possess it. Stereopsis, or stereoscopic vision, provides the individual with the most acute sense of relative depth and is of vital importance in visual tasks requiring the precise location of objects.

Stereopsis is believed to have an innate origin in the anatomic and physiologic structures of the retinas of the eyes and the visual cortex. It is present in normal binocular vision because the two eyes view objects in space from two points, so that the retinal image patterns of the same object points in space are slightly different in the two eyes. The stereoscope, with which different pictures can be presented to each eye, demonstrates the fundamental difference between stereoscopic perception of depth and the conception of depth and distance from the monocular view. In the **illustration**, each of the two eyes views a pair of vertical lines A and B drawn on cards. The separation of these lines for the right eye is greater than that for the left eye. If the difference in separation is not too great, the images of the lines fuse when the two targets are observed by the two eyes. There is almost an immediate stereoscopic spatial experience that the two lines are located in space, line B being definitely more distant than line A. No hint of this spatial experience occurs in the observation of each target alone. This difference between the images in the two eyes provides the stimulus for emergence of the stereoscopic experience.

Stereoscopic acuity and the presence of stereopsis may be tested by several methods or instruments such as the ordinary hand stereoscope and similar devices, vectograph targets, the Howard-Dolman peg test, the Verhoeff stereopter, the Hering falling bead test, pins with colored heads stuck in a board to test for near vision, and afocal meridional magnifying lenses placed before one eye while the subject views a special field (leaf-room or tilting table). *See* VISION.

Kenneth N. Ogle

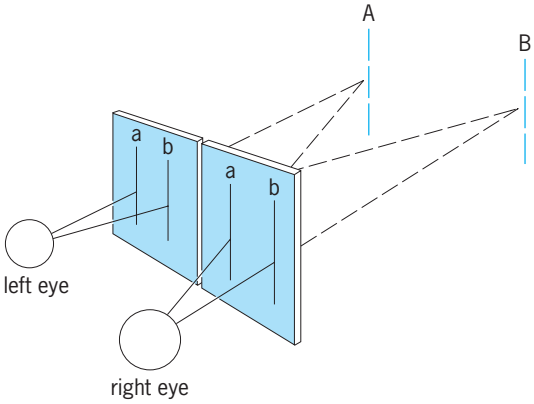

Diagram of stereoscopic vision.

Bibliography. C. H. Graham (ed.), *Vision and Visual Perception*, 1965; W. L. Gulick and R. B. Lawson, *Human Stereopsis: A Psychological Approach*, 1976; E. Hering, *The Theory of Binocular Vision*, 1977; K. N. Ogle, *Optics: An Introduction for Ophthalmologists*, 2d ed., 1976; K. N. Ogle, *Research in Binocular Vision*, 1950, reprint 1964; R. W. Reading, *Binocular Vision: Foundations and Applications*, 1983; H. Solomon, *Binocular Vision*, 1980.

Steric effect (chemistry)

The influence of the spatial configuration of reacting substances upon the rate, nature, and extent of reaction. The sizes and shapes of atoms and molecules, the electrical charge distribution, and the geometry of bond angles influence the courses of chemical reactions. The steric course of organochemical reactions is greatly dependent on the mode of bond cleavage and formation, the environment of the reaction site, and the nature of the reaction conditions (reagents, reaction time, and temperature). The effect of steric factors is best understood in ionic reactions in solution. The nucleophilic substitution reaction at a saturated carbon atom can serve as an illustration.

Saturated nucleophilic substitution. While the reacting carbon atom is in an electron-deficient state in the transition state (state of highest energy, somewhere between starting material and product) in a nucelophilic substitution process, the reaction can be varied from a two-step, unimolecular ionization to a one-step, bimolecular transformation. The former mode of reaction, a solvolysis or S_N1 process, converts a tetrahedral carbon into a solvated planar carbonium ion intermediate, and hence leads to racemization (randomization of configuration); for example, reaction (1) where the R terms represent organic groups.

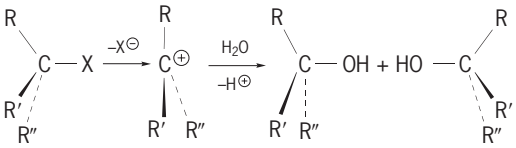

(1)

See REACTIVE INTERMEDIATES.

The second reaction path proceeds by a simultaneous rupture of the old bond and creation of a new one, either by inversion of configuration, a Walden inversion or S_N2 process, for example, reaction (2);

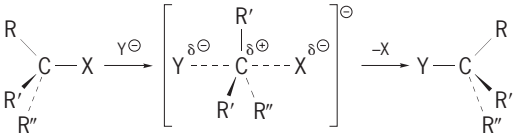

(2)

or in a few cases by a front-side displacement of part of the substituent already present, an $S_N i$ process, and hence, retention of configuration,

for example, reaction (3).

$$\mathrm{ROH} \xrightarrow{\mathrm{SOCl_2}} \mathrm{ROSOCl} \longrightarrow \underset{\text{(ion pair)}}{\mathrm{R^{\oplus}OSOCl^{\ominus}}} \xrightarrow{-\mathrm{SO_2}} \underset{\text{(ion pair)}}{\mathrm{R^{\oplus}Cl^{\ominus}}} \longrightarrow \mathrm{RCl} \qquad (3)$$

These substitution reactions are highly solvent-dependent; for example, tert-butyl chloride undergoes solvolysis close to 500,000 times faster in water, a solvent of high ionizing power, than in ethanol, a solvent of low dielectric properties. The nature of R, R′, and R″ is one of the factors determining which of the above pathways a compound prefers for its substitution. The larger the size of these three groups, the greater is the tendency to relieve steric strain by extrusion of X, and thus the more the need for an S_N1 process. The smaller the size of the environment, the greater is the accessibility of the reagent from the backside, and thus the more tendency toward an S_N2 process. As a consequence, tertiary compounds undergo racemization readily, whereas primary systems prefer inversion.

Both the rate and the steric course of an ionic displacement may depend often on the ability of groups adjacent to the reaction site to accommodate a positive charge. Substitution of α-halo ethers and allyl or benzyl halides occurs much faster than a similar reaction of unsubstituted halides because of the intermediacy of the stabilized cations or cationlike transition states in notation (4).

In view of the charge distribution over more than one atom in these cases, sometimes the incoming reagent forms a bond at a site different from that of the leaving group. As a consequence, an S_N2' process, reaction (5), or an S_Ni' path, reaction (6), may result. Both processes yield retention of configuration; that is, the orientation of the new substituent on its carbon atom is identical with that originally held by the former functional group on a site that is two carbon atoms removed.

In the presence of participating neighboring groups, even solvolyses can lead to retention of configuration. Certain rigidly held homoallyl systems undergo substitution at a site that is three carbons removed from the position of the leaving group, but with retention of configuration, as represented by reaction (7).

Double inversion is responsible for the retained configuration of the products of solvolysis of α-halo acid salts, reaction (8).

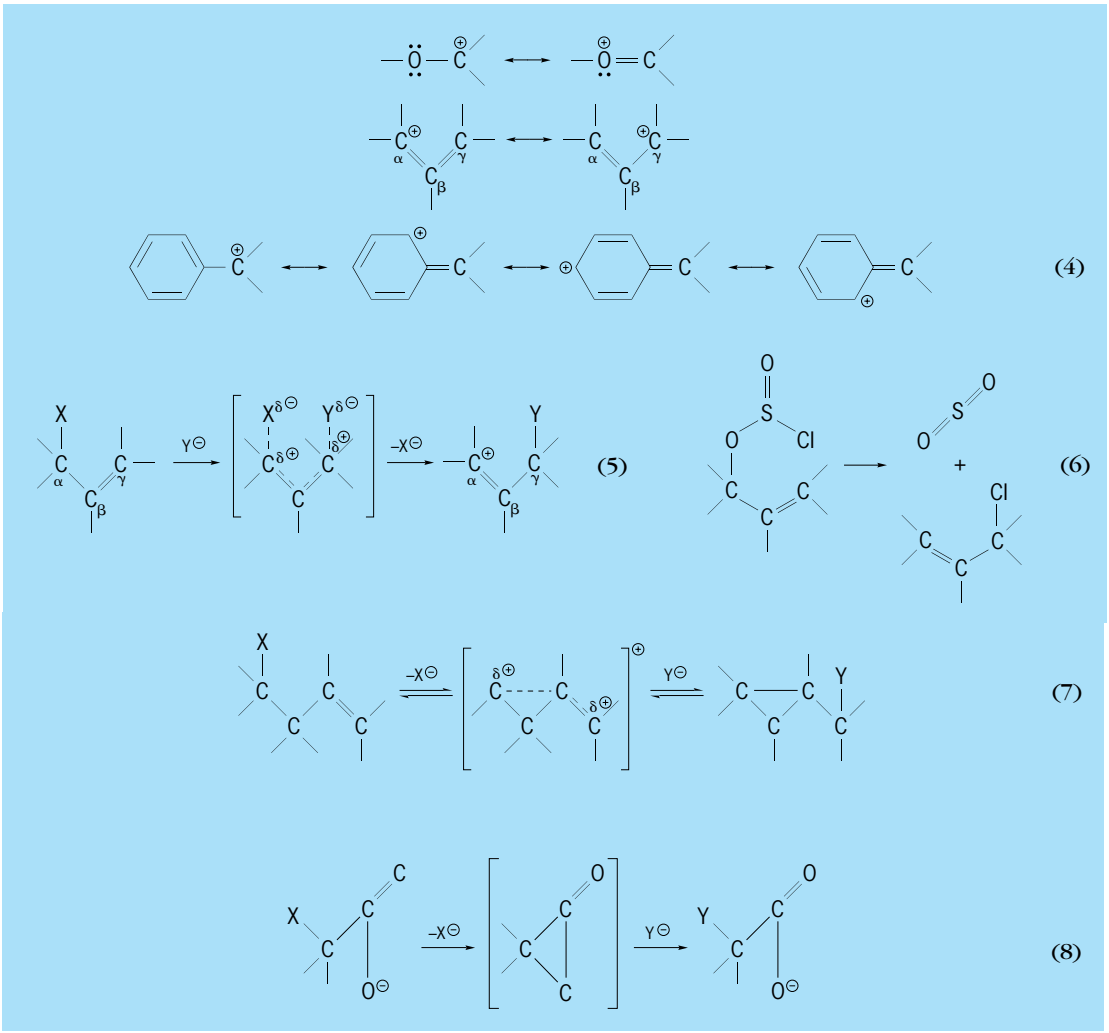

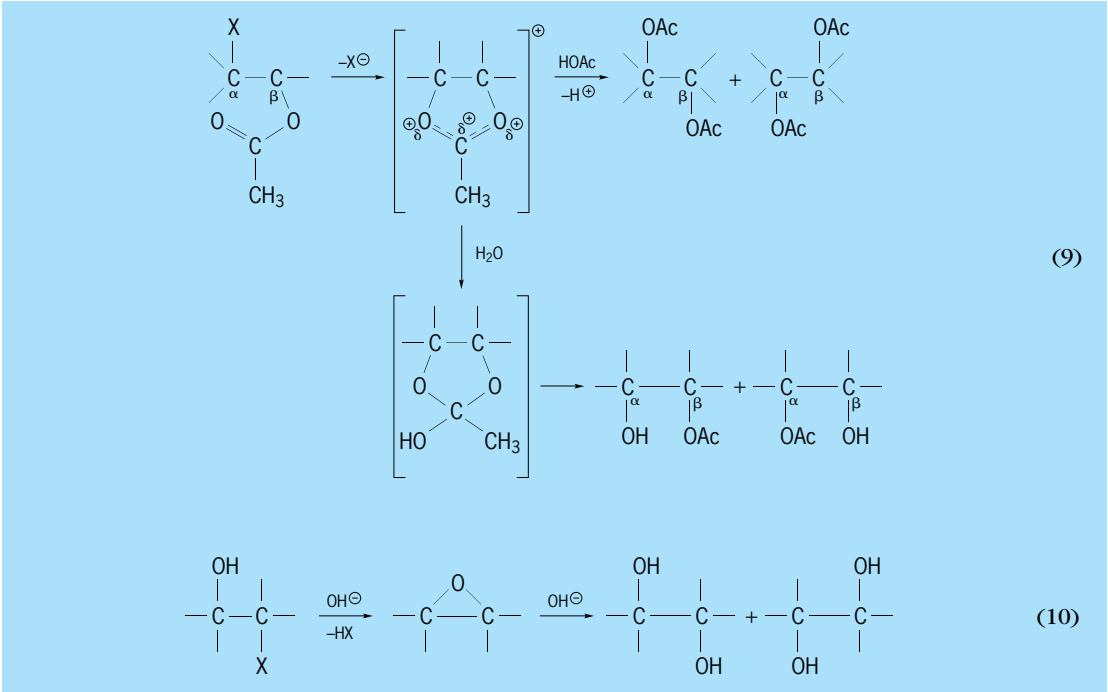

(9)

(10)

Solvolysis of *trans*-β-acetoxy systems in nonaqueous media leads to trans products; in the presence of water, cis compounds are obtained, reactions (9), where Ac = acetyl group (CH_3CO—).

Base treatment of *trans*-halohydrins leads to *trans*-vicinal glycols. The intermediate epoxide is isolable. Although ring opening of the latter may yield two different trans products, the diaxial one is formed preferentially in cycloalkane cases, reaction (10).

Rearrangement. Migration of neighboring groups toward the reaction site, resulting in skeletal rearrangements, is a common occurrence. Both the internal displacement of the leaving group by the migrating group and the subsequent external displacement of the migrating group by the solvent or added reagent proceed in a trans sense, that is, by back-side approach. Thus, the overall steric consequence of one migration sequence is retention of configuration.

There are several examples of 1,2-hydride migration, for example, reaction (11).

Transannular hydride shifts are quite similar in nature, as in reaction (12).

The 1,2 migration of phenyl groups proceeds by way of fairly stable phenonium ions, reaction (13).

The Wagner-Meerwein rearrangement of saturated neighboring groups shows directional effects similar to those of the above migrations; for example, the conversion of camphene hydrochloride to bornyl chloride is stereospecific, with retention of configuration, reaction (14).

Neopentyl halides solvolyze to tertiary amyl derivatives, reaction (15).

Cyclohexane systems with equatorial leaving groups may undergo contraction to five-membered rings, reaction (16).

Organic compounds possessing potential leaving groups at the bridgehead of small bicyclic ring systems undergo substitution processes only sluggishly. In the absence of ready access at the backside of the reaction center, the S_N2 pathway is excluded. The inability of the compounds to form planar carbonium ions precludes a S_N1 route. However, displacements

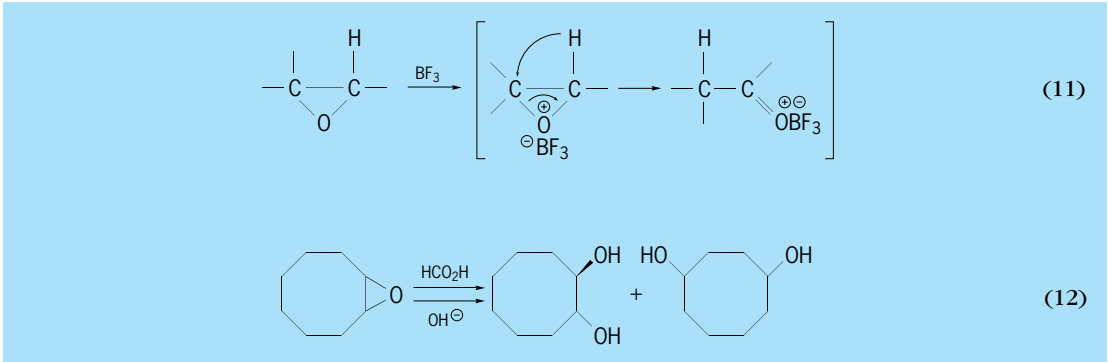

(11)

(12)

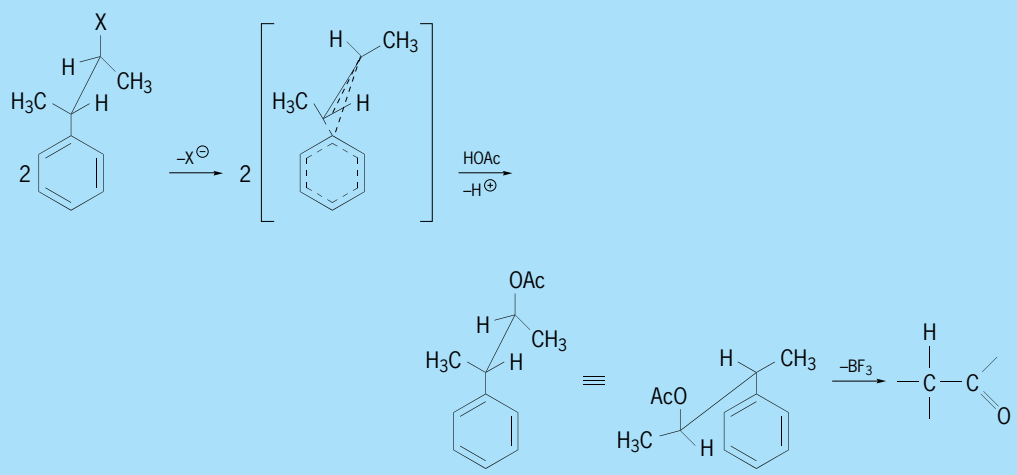

(13)

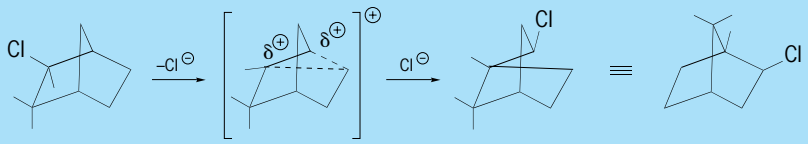

(14)

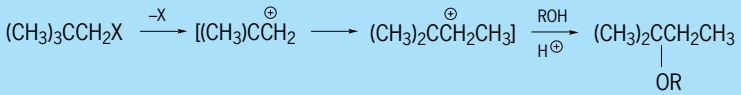

(15)

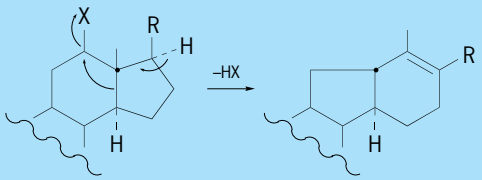

(16)

do occur slowly at elevated temperatures, presumably via nonplanar cations.

Unsaturated nucleophilic substitution. Nucleophilic substitution reactions at unsaturated carbon atoms can take place by two possible mechanistic routes, an elimination-addition process and an addition-elimination scheme. The former route is best illustrated by the transformation of aromatic halides into anilines, reaction (17).

The latter is encountered in the interconversion of carboxylic acids and their derivatives, for example, reaction (18). Because the central carbon atom has greater steric requirements in the reaction intermediate than in the starting material, the reaction velocity is strongly dependent on the size and number of neighboring groups; that is, an increase in the bulkiness of R is reflected in a decrease of the rate of the reaction.

The addition-elimination mechanism is portrayed also by the aromatic nucleophilic substitution reaction, for example, (19). In order to be able to stabilize the reaction intermediate, the all-important nitro group must be coplanar with the benzene ring. As a consequence, ortho substituents, which

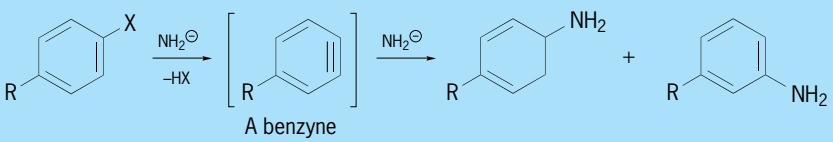

(17)

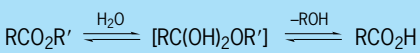

(18)

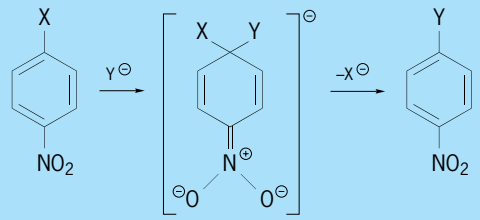

(19)

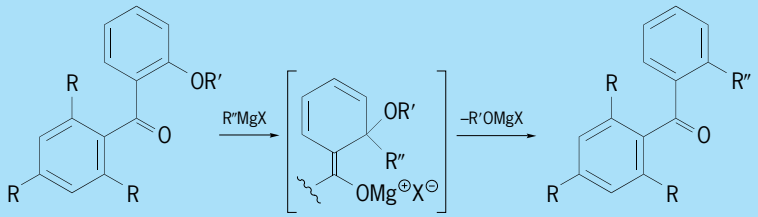

(20)

may block this steric requirement, retard the reaction rate. Unusual aromatic nucleophilic substitutions have been observed in cases where steric hindrance by ortho substituents has prevented addition to aromatic ketones to occur, for example, reaction (20).

Addition reactions. The steric course of addition reactions at unsaturated sites depends largely on the reagent. Catalytic hydrogenation, a nonhomogeneous process of undetermined mechanism, occurs in a cis manner. In the absence of any steric interference, it leads to thermodynamically stable products. In the presence of steric hindrance, the two new hydrogen atoms are usually introduced on the least hindered side of the unsaturated compounds. However, sometimes some bulky polar groups actually aid, rather than retard, adsorption of the catalyst on their side of the reducing compound, thereby leading to products of opposite configuration.

The oxidation of olefins to vicinal glycols by permanganate salts or osmium tetroxide also proceeds in a cis fashion and involves the least-hindered side of the reacting substrate. The Diels-Alder reaction behaves similarly, for example, reaction (21). All addition processes, during which two new bonds are formed more or less simultaneously, yield cis adducts. *See* DIELS-ALDER REACTION.

Ionic addition reactions of olefins and acetylenes occur in a trans manner. The mode of addition is such as to lead to product via the most stable cations (Markovnikoff addition), for example, reaction (22).

Halogen addition to cyclic olefins leads to trans diaxial dihalides which, on standing, isomerize to the more stable trans diequatorial dihalides, reaction (23).

Ionic addition reactions of carbonyl compounds follow a steric course very similar to those of olefins. However, the reagents are mostly nucleophilic, and some reactions are equilibrium processes, for example, reaction (24). The orientation of attack and the

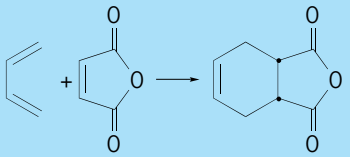

(21)

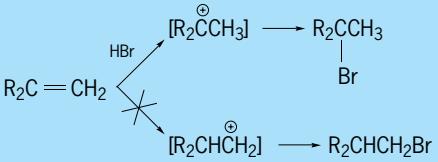

(22)

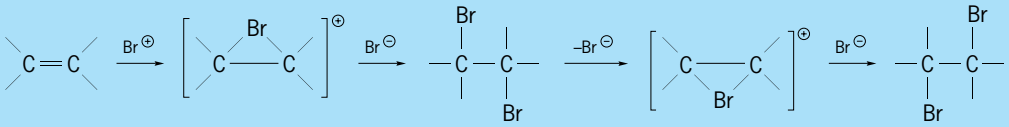

(23)

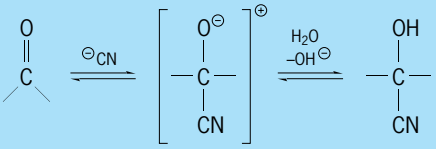

(24)

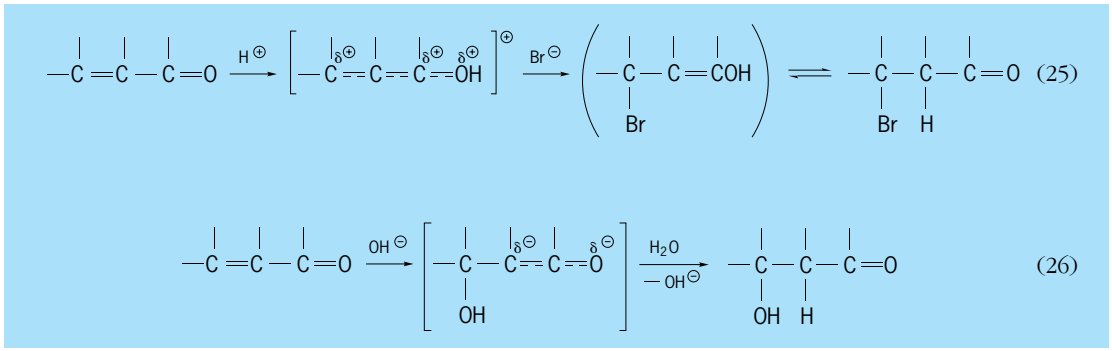

(25)

(26)

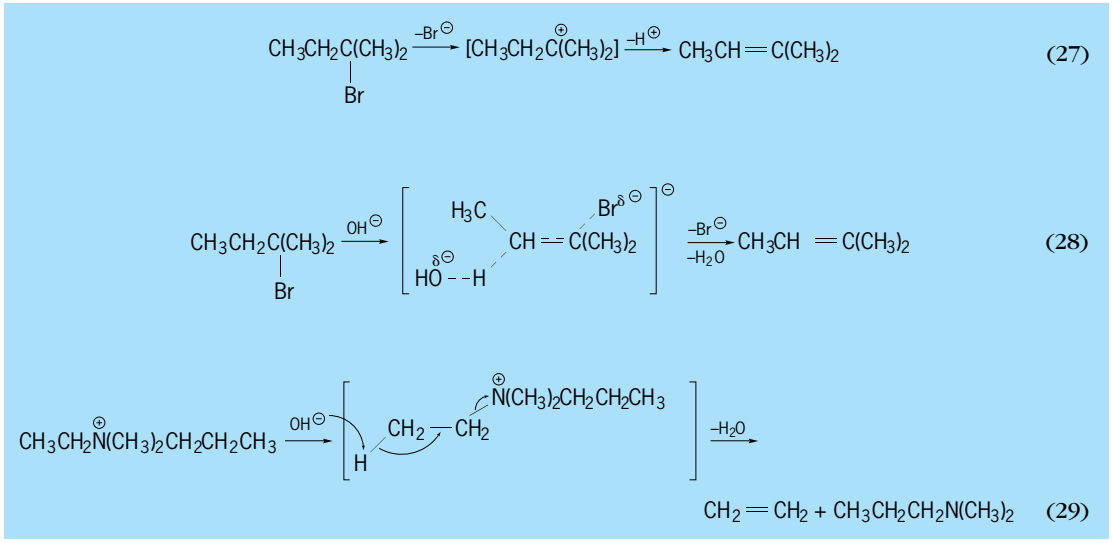

(27)

(28)

(29)

reaction rate are governed by the environment of the carbonyl group (C═O).

Addition reactions of conjugated carbonyl systems can occur through cation as well as anion intermediates, but they uniformly place the nucleophilic part of the reagents on the β-carbon atoms, for example, reactions (25) and (26).

The reaction of carbonyl compounds as enol or enolate anions with electrophilic reagents can take two different courses. If the process is kinetically controlled, the electrophile, a proton, halonium ion, or others, interacts with the substrate on its least-hindered side. However, if the reaction is thermodynamically controlled, then, independent of mechanism, the most stable product is obtained.

Elimination reactions. Elimination reactions can be carried out by pyrolysis of esters, halides, or amine oxides in the liquid or vapor phase. These eliminations always involve a rupture of vicinal cis bonds.

Alternatively, similar cleavage processes can be made to occur ionically in solution, in which case they proceed in a trans fashion. The direction of elimination depends greatly on the molecularity of the process, as well as on the sizes of the leaving group and the attacking base. The two-step, unimolecular cleavage, an *E*1 process, leads predominantly to the more substituted, hence more stable, olefins, for example, reaction (27).

The one-step bimolecular elimination (an *E*2 process) of neutral compounds yields similar products, for example, reaction (28). However, an *E*2 reaction on positive ions, ammonium or sulfonium salts, affords the less substituted olefin in preponderant yield, reaction (29).

Reactions leading to the more stable products are said to follow the Saytzeff rule, whereas those yielding less stable olefins obey the Hofmann rule. Because the transition state in the *E*2 reaction is of lowest energy when all atoms involved in the elimination are in a plane, the fastest rates among cyclic compounds are encountered in the cases which permit a diaxial alignment of vicinal substituents.

Many ionic elimination reactions are known which involve the rupture of more than two bonds, for example, reaction (30).

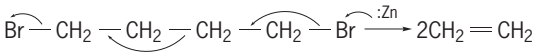

(30)

The *E*2′ processes, eliminations of two groups of carbon atoms separated by an olefinic linkage, appear to be cis in nature, reaction (31).

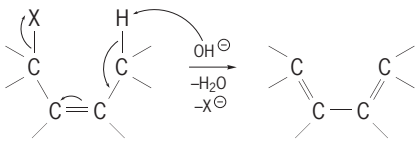

(31)

Electrophilic substitution. Steric factors have a fair control over the course of the aromatic electrophilic substitution reaction. In reactions of compounds containing ortho-para directing substituents, the *ortho/para* product ratio is usually greater than 1, and increases with the size of the substituent and that of the reacting species. The rate-accelerating participation of electrondonating groups, located ortho or para to the incoming substituent, in stabilizing the transition state is greatly diminished in the presence of bulky ortho neighbors which would prevent the groups from attaining coplanarity with the benzene, for example, reaction (32).

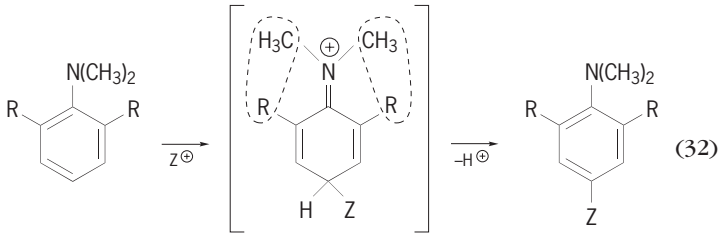

(32)

See CHELATION; CONFORMATIONAL ANALYSIS; ORGANIC SYNTHESIS; ORGANIC REACTION MECHANISM; PHYSICAL ORGANIC CHEMISTRY; STEREOCHEMISTRY.

Ernest Wenkert

Bibliography. A. T. Balaban et al. (eds.), *Steric Fit in Quantitative Structure-Activity Relations*, 1980; F. A. Carey and R. J. Sundberg, *Advanced Organic Chemistry*, pt. A: *Structures and Mechanisms*, pt. B: *Reactions and Synthesis*, 4th ed., 2000; E. L. Eliel and S. H. Wilen, *Stereochemistry of Organic Compounds*, 1994; J. E. McMurry and E. E. Simanek, *Fundamentals of Organic Chemistry*, 6th ed., 2006; G. Naray-Szabo and K. Simon (eds.), *Steric Aspects of Biomolecular Interactions*, 1987; S. D. Peyerimhoff (ed.), *Interactions in Molecules: Electronic and Steric Effects*, 2003.

Sterilization

An act of destroying all forms of life on and in an object. A substance is sterile, from a microbiological point of view, when it is free of all living microorganisms. Sterilization is used principally to prevent spoilage of food and other substances and to prevent the transmission of diseases by destroying microbes that may cause them in humans and animals.

Microorganisms can be killed either by physical agents, such as heat and irradiation, or by chemical substances. Regardless of the manner in which they are killed, they generally die at a constant rate under specified environmental conditions. If the logarithm of the number of survivors is plotted against time, the resulting curve is a straight line.

When testing a substance for sterility, care must be taken to employ appropriate techniques. A bacterial cell is considered to be killed when it is no longer capable of reproducing itself under suitable environmental conditions. If an inadequate medium is employed to subculture the treated bacteria, the substance being tested may be wrongly considered to be sterile.

By far the most resistant of all forms of life, to both physical and chemical killing agents, are some of the bacterial endospores. If they did not exist, sterilization of such materials as bacteriological media and equipment, hospital supplies, and canned foods would be much simpler.

Heat sterilization. This is the most common method of sterilizing bacteriological media, foods, hospital supplies, and many other substances. Either moist heat (hot water or steam) or dry heat can be employed, depending upon the nature of the substance to be sterilized. Moist heat is also used in pasteurization, which is not considered a true sterilization technique because all microorganisms are not killed; only certain pathogenic organisms and other undesirable bacteria are destroyed. *See* PASTEURIZATION.

Moist-heat sterilization. Some bacterial endospores are capable of surviving several hours at 212°F (100°C). Therefore, for moist-heat sterilization, an autoclave, pressure cooker, or retort, with steam under pressure, is required to achieve higher temperatures (**Fig. 1**).

Most bacteriological media are sterilized by autoclaving with steam at 250°F (121°C) with 15 lb pressure (10^2 kilopascals) for 20 min or more, depending upon the volume of material being heated. Some spores are capable of surviving moist-heat sterilization equivalent to at least 7 min at 250°F (121°C).

Steam under atmospheric pressure in an Arnold sterilizer is sometimes employed for specialized bacteriological media that are easily heat damaged. Because many bacterial spores survive this treatment, it is obviously inadequate to ensure sterility.

Tyndallization. The food or medium is steamed for a few minutes at atmospheric pressure on three or four successive occasions, separated by 12- to 18-h intervals of incubation at a favorable growing temperature. In theory the intervals of incubation allow any surviving bacterial spores to germinate into more heat-sensitive, vegetative cells, which then would be killed during the next heat treatment. However, spores like vegetative cells may require special conditions such as an appropriate medium, proper oxygen tension, or proper temperature to germinate and reproduce. These conditions may not be realized during the intervals between heat treatment, and no matter how often the steaming is repeated ungerminated spores may survive and eventually germinate when conditions have been changed. The survival of ungerminated spores reduces the effectiveness of this method and it has been supplanted by other methods.

Hot-air sterilization. Glassware and other heat-resistant materials which need to be dry after treatment are usually sterilized in a hot-air sterilizer. Dry sterilization requires heating at higher temperatures and for longer periods of time than does sterilization by steam under pressure. Temperatures of 320–330°F (160–165°C) for at least 2 h is generally employed in hot-air sterilization. Dry heat kills the germs through

denaturation of protein which may involve oxidative processes.

Radiation sterilization. Many kinds of radiations are lethal, not only to microorganisms but to other forms of life. These radiations include both high-energy particles as well as portions of the electromagnetic spectrum. The mechanism of the lethal action of these radiations is not entirely clear. It may involve a direct energy absorption at some vital part of the cell (direct target theory), or the production of highly reactive, ionized, free radicals near some vital part of the cell (indirect effect). Bacterial endospores are relatively resistant to all types of radiation. *See* RADIATION BIOLOGY.

Ultraviolet radiation. Radiant energy in the ultraviolet region of the spectrum is highly bactericidal, especially at wavelengths of approximately 265 nanometers. Lamps which generate ultraviolet radiation in this region are useful for the sterilization of air and smooth surfaces. Ultraviolet rays have very low penetrative capacity, since even a thin layer of glass absorbs a high percentage of the rays. Some irradiated cells that are presumably dead may be photoreactivated with visible light.

Gamma rays. These are high-energy, electromagnetic radiations similar to x-rays. They have great penetrative capacity, and their energy is dissipated in the production of ionized particles from the material being irradiated. Radioactive isotopes, such as cobalt-60, are a common source of gamma rays (**Fig. 2**). Gamma irradiation of foods has received much attention as a means of sterilizing foods without cooking them. Sterilization requires a radiation dose of approximately 5,000,000 rads (5×10^8 ergs of energy absorbed per gram).

Cathode rays. These high-speed electrons (beta rays) may be generated with various types of electron accelerators. This type of radiation has relatively low penetrative capacity, depending upon the energy level of the emitted electron beam. Cathode rays sterilize in a manner identical to gamma rays and without significantly raising the temperature of the material being irradiated. They have received some application in the sterilization of surgical supplies and drugs and in the experimental sterilization of foods (**Fig. 3**).

Filtration sterilization. This is the physical removal of microorganisms from liquids by filtering through materials having relatively small pores. Sterilization by filtration is employed with liquid that may be destroyed by heat, such as blood serum, enzyme solutions, antibiotics, and some bacteriological media and medium constituents. Examples of such filters are the Berkefeld filter (diatomaceous earth), Pasteur-Chamberland filter (porcelain), Seitz filter (asbestos pad), and the sintered glass filter. Most of these filters are available in different pore sizes.

The mean pore size of bacteriological filters is not the only determinant in their effectiveness. The electric charge of the pore surfaces tends to adsorb the bacteria and thus prevent their passage. Most bacteria have a net negative electrical charge on their surfaces. Usually, bacteriological filters will permit the passage of viruses, which are then called filterable.

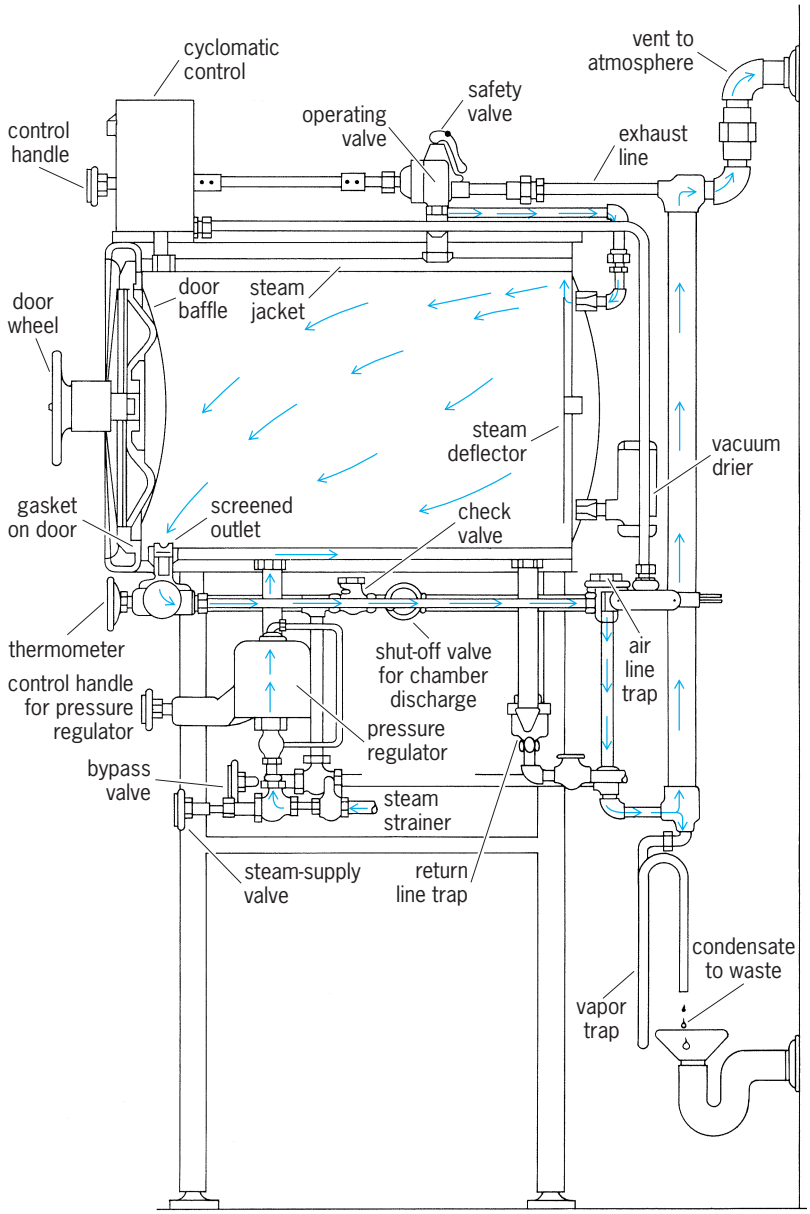

Fig. 1. Diagrammatic sketch of an autoclave. (*American Sterilizer Co.*)

A Millipore filter is a specially prepared membrane molecular filter designed to remove bacteria from water, air, and other materials, for the purpose of estimating quantitatively the bacterial population. A sterile filter disk is assembled in a filtration unit, and a specified volume of water or solution is drawn through the disk which then retains the bacteria. The filter disk is removed and placed in a sterile petri dish containing an absorbent pad, previously saturated with an appropriate bacteriological medium. Upon incubation, colonies will develop on the filter disk wherever bacteria were entrapped at the time of filtering. Special differential, or selective, media can be employed for the purpose of detecting quantitatively specific types of bacteria from the original material.

Chemical sterilization. Chemicals are used to sterilize solutions, air, or the surfaces of solids. Such chemicals are called bactericidal substances. In lower concentrations they become bacteriostatic rather than bactericidal; that is, they prevent the growth of bacteria but may not kill them. Other terms having similar meanings are employed. A disinfectant is a chemical that kills the vegetative cells of pathogenic microorganisms but not necessarily the endospores of spore-forming pathogens. An antiseptic is a chemical applied to living tissue that prevents or retards the growth of microorganisms, especially pathogenic bacteria, but which does not necessarily kill them.

The death of microorganisms subjected to bactericidal substances can be expressed exponentially, in that a straight-line graph is produced when the logarithm of survivors is plotted against time. The more concentrated the chemical employed, the greater is the rate of death.

There are hundreds of chemicals that may be considered to have sterilizing or bactericidal properties, depending upon the particular use for which they are intended. A chemical may be particularly useful for one purpose but not for another. Many are widely used for sterilization or disnfection of air, water, tabletops, surgical instruments, and so on.

The desirable features sought in a chemical sterilizer are toxicity to microorganisms but nontoxicity to humans and animals, stability, solubility, inability to react with extraneous organic materials, penetrative capacity, detergent capacity, noncorrosiveness, and minimal undesirable staining effects. Rarely does one chemical combine all these desirable features.

Fig. 2. Cobalt-60 furnace which is designed for experimental food sterilization.

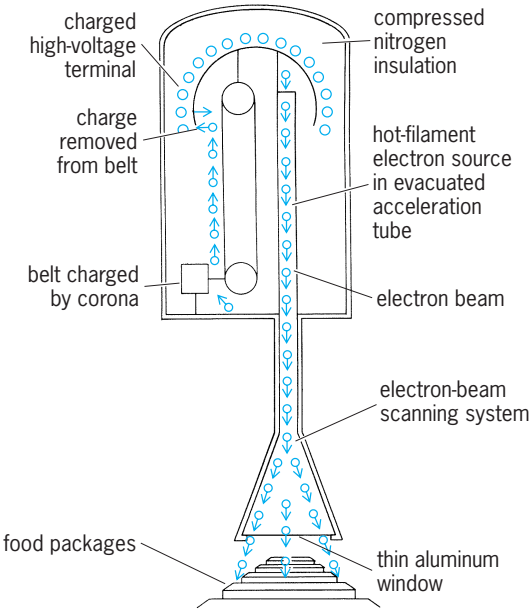

Fig. 3. Van de Graaff generator principle; packaged products sterilized by means of high-intensity electron beam which penetrates packages. (*After W. M. Urbain, Food Eng., vol. 25, February 1953*)

Among chemicals that have been found useful as sterilizing agents are the phenols, alcohols, chlorine compounds, iodine, heavy metals and metal complexes, dyes, and synthetic detergents, including the quaternary ammonium compounds.

Chlorine. Chlorine and chlorine-containing compounds represent the most widely used group of disinfectants. Chlorine gas is often used to purify municipal water supplies.

Various compounds of chlorine, such as the hypochlorites and chloramine, have many industrial and domestic uses as disinfectants or antiseptics.

Ozone. This is a highly oxidizing gas (O_3) used as a deodorizer, and also for disinfection of air and water. It has found some use in the food fields, but effective bactericidal concentrations may be irritating and toxic to humans.

Hydrogen peroxide. This chemical (H_2O_2) has high oxidizing and bleaching qualities and is usually employed in a 3% solution for topical application and disinfection of cuts, scratches, and minor wounds. It decomposes into water and oxygen and therefore is used where no taste, odor, or toxic residues are permitted.

Volatile organic compounds. Such compounds as formaldehyde and ethylene oxide have been used for the disinfection of sickrooms occupied by patients suffering from contagious disease (terminal disinfection), and of solids that do not permit heat treatment. These volatile substances have the advantages of effective penetrative capacity and ease of removal after treatment.

Volatile organic substances are sometimes employed as bacteriostatic agents to preserve

bacteriological medium constituents until they are heat-sterilized. Charles F. Niven, Jr.

Bibliography. S. S. Block (ed.), *Disinfection, Sterilization, and Preservation*, 5th ed., 2001; P. Murray et al., *Medical Microbiology*, 2000.

Steroid

Any of a subset of biological molecules known as lipids that share a specific, fused four-ring structure comprising three rings of six carbon atoms and one ring of five carbon atoms. Different types of steroids have different chemical modifications to this common ring structure (see **illustration**). Steroids are found in animals, plants, fungi, and protozoa. In mammals, steroids are components of cell membranes, aid in digestion, and act as hormones to control a wide range of physiological responses. Because of their role in governing physiological processes, naturally occurring and synthetic steroids are used extensively in medicine to treat a wide range of conditions. *See* LIPID; STEROL.

Cholesterol. Cholesterol is the most abundant steroid in animals. It is found primarily in the plasma membrane of cells, where it is a major component of the lipid bilayer. It is also found, usually to a lesser extent, in the membranes of subcellular organelles such as mitochondria. Because cholesterol possesses a fused ring structure, it has greater rigidity than the other lipid components of these membranes and therefore tends to reduce the fluidity of membranes by restricting the movements of the more flexible lipid components. Thus, cholesterol plays an important role in determining the physical characteristics of biological membranes. In addition, cholesterol is one of the major components of lipid rafts, localized regions within the plasma membrane that are more ordered and rigid than their surroundings. Lipid rafts associate with specific proteins and are believed to play an important role in signal transduction. While plants have very little cholesterol, a similar steroid, stigmasterol, plays an analogous role in the membranes of their cells. *See* CELL (BIOLOGY); CELL MEMBRANES.

Because of its nonpolar nature, cholesterol has low solubility is water. Consequently, cholesterol must be transported through the bloodstream attached to carrier proteins known as lipoproteins. Low-density lipoproteins (LDL) transport cholesterol from the liver to various tissues in the body. High-density lipoproteins (HDL) transport cholesterol from various body tissues to the liver. HDL can be thought of as a "cholesterol scavenger" because it can take excess cholesterol directly from cell membranes.

High levels of cholesterol-containing LDL in the bloodstream are strongly correlated with arteriosclerosis (hardening of the arteries) and elevated risk of arterial blockage and heart attack. This is due to the deposition of excess cholesterol in the smooth muscle cells of the inner arterial wall. Over time, these deposits become calcified and can ultimately lead to artery blockage. High HDL levels have been correlated with reduced risks of heart disease. *See* CHOLESTEROL; LIPOPROTEIN.

Steroid hormones. In addition to being an important component of cell membranes, cholesterol is a precursor of a variety of critical steroid hormones. Among these are glucocorticoids, androgens, estrogens, progestins, and mineralocorticoids. Each family of steroid hormones elicits a different physiological response.

Glucocorticoids. Cortisol (also known as hydrocortisone) is an example of a glucocorticoid. This hormone is produced by the adrenal gland and participates in the regulation of carbohydrate

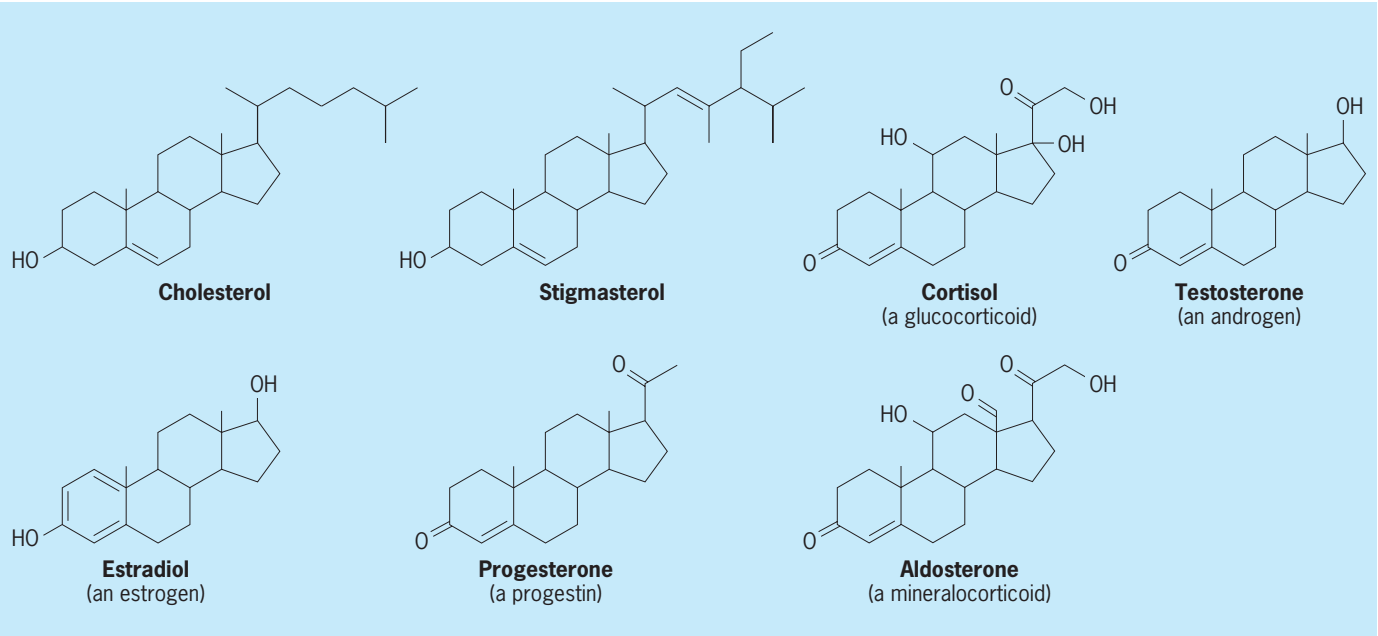

Structures of some important steroids.

metabolism. Cortisol causes an increase in the release of glucose from the liver and a decrease in the uptake of glucose by other tissues of the body. Consequently, cortisol effectively raises blood sugar (glucose) levels and acts in opposition to insulin. Cortisol also has anti-inflammatory and antiallergic effects, and is widely used medicinally to treat asthma, allergies, and various skin conditions.

Androgens. Testosterone is the primary androgen hormone. It is produced in the gonads and is a male sex hormone. Testosterone is required for the fetal development of male genitalia. During puberty, testosterone is required for the development of secondary sex characteristics in males, such as facial and pubic hair and a lowered voice. Testosterone also promotes protein synthesis that in turn leads to muscle production. Synthetic androgenic steroids that mimic testosterone are used medicinally to treat the severe weight loss that accompanies certain wasting diseases. To increase muscle mass and strength, many athletes have abused synthetic androgenic steroids. *See* ADRENAL GLAND.

Estrogens. Estradiol is the primary estrogen hormone. It is produced by the ovaries and placenta and is a female sex hormone. During puberty, estradiol is required for the development of female secondary sex characteristics, such as pubic hair and enlarged breasts. Estradiol causes the thickening of the uterine lining that occurs during the early phase of the menstral cycle. It also suppresses the secretion of follicle-stimulating hormone by the pituitary gland, thereby preventing ovulation. Because of its ability to suppress ovulation, estradiol is a major component of oral contraceptives. *See* ESTROGEN.

Progestins. Progesterone is the primary progestin. This hormone is produced by the adrenal gland and the corpus luteum (an ovarian follicle that has ruptured after releasing an egg). Progesterone is a female sex hormone. It causes changes in the uterine lining that enable implantation of a fertilized egg. If pregnancy results, the placenta starts to produce progesterone to prevent spontaneous abortion and prepare the mammary glands for milk production. Synthetic progestins are also major components of oral contraceptives. *See* PROGESTERONE.

Mineralocorticoids. Aldosterone is an example of a mineralocorticoid hormone. It is produced by the adrenal gland and regulates electrolyte balance. It acts on the kidneys to stimulate potassium ion excretion and to repress sodium ion excretion. *See* ALDOSTERONE.

Mechanism of action. Steroid hormones elicit the physiological responses described above by altering the expression of various genes. Because steroids are nonpolar, they readily pass through the cell membrane. Once inside the cell, the steroid molecules bind to specific protein receptors in the cytoplasm. The hormone-protein receptor complexes then enter the nucleus and bind to specific regions of the DNA known as response elements. In many cases, binding of the hormone-protein receptor complex to response elements in the DNA induces the expression of genes adjacent to the response element. In other cases, this binding represses the expression of genes adjacent to the response element.

Bile acids. In addition to being a precursor of steroid hormones, cholesterol is a precursor of bile acids (also known as bile salts), such as cholic acid and glycocholic acid. Cholesterol is converted to bile acids in the liver, and these compounds are secreted into the small intestine via the gall bladder. In the small intestine, bile acids act as emulsifying agents to aid in the digestion and absorption of fats. While most of the bile acid in the small intestine is reabsorbed and recycled, some is not. This nonabsorption represents the major means of excretion of excess cholesterol by the body. *See* LIVER.

M. Todd Washington

Bibliography. T. M. Devlin (ed.), *Textbook of Biochemistry with Clinical Correlation*, 5th ed., Wiley, New York, 2002; D. Voet and J. G. Voet, *Biochemistry*, 3d ed., Wiley, New York, 2004.

Sterol

Any of a group of naturally occurring or synthetic organic compounds with a steroid ring structure, having a hydroxyl (—OH) group, usually attached to carbon-3. This hydroxyl group is often esterified with a fatty acid (for example, cholesterol ester). The hydrocarbon chain of the fatty-acid substituent varies in length, usually from 16 to 20 carbon atoms, and can be saturated (containing no double bonds) or unsaturated (containing one or more double bonds). Sterols commonly contain one or more double bonds in the ring structure and also a variety of substituents attached to the rings. Sterols and their fatty-acid esters are essentially water insoluble. For transport in an aqueous milieu (for example, the bloodstream of mammals), sterols and other lipids are bound to specific proteins, forming lipoprotein particles. These particles are classified based on their composition and density. One lipoprotein class is abnormally high in the blood of humans prone to heart attacks.

Sterols are widely distributed in nature. The biosynthetic pathway leading to the steroid ring structure is utilized by higher and lower forms of life. Modifications of the steroid ring structure are made by specific enzyme systems, producing the sterol characteristic for each species such as ergosterol in yeast. The common biosynthetic pathway begins with the two-carbon-atom compound, acetate, which is converted by several enzymic steps to a six-carbon-atom compound, mevalonate. Next, mevalonate is converted to a five-carbon-atom intermediate, isopentenyl pyrophosphate. There follows a series of condensations of six isopentenyl pyrophosphate molecules to yield a 30-carbon-atom intermediate, squalene. Squalene is next cyclized to a 30-carbon-atom sterol, lanosterol. The structure of lanosterol is then modified to yield the principal

sterol of a species. The major regulatory step in the sterol biosynthetic pathway occurs early in the process. Drugs that lower blood cholesterol levels in humans are designed to inhibit this regulatory enzyme. In addition to their conversion to sterols, several intermediates in the pathway are precursors of other important biological compounds, including chlorophyll in plants, vitamins A, D, E, and K, and regulators of membrane functions and metabolic pathways.

A universal role of sterols is to function as part of membrane structures. In addition, some insects require sterols in their diets. Cholesterol also serves as a precursor of steroid hormones (estrogens, androgens, glucocorticoids, and mineralocorticoids) and bile acids. *See* CHOLESTEROL; STEROID.

Mary E. Dempsey

Bibliography. K. Bloch, The biological synthesis of cholesterol, *Science*, 150:19–28, 1965; G. W. Patterson and W. D. Nes (eds.), *Physiology and Biochemistry of Sterols*, 1992.

Stichosomida

An order of nematodes (roundworms) formerly constituting at least two important groups of animal parasites: the mermithids, which are parasites of invertebrates, and the trichinellids, which are parasites of vertebrates. These nematodes are characterized by a pharynx that is narrow and thin-walled anteriorly and is surrounded posteriorly by unicellular, glandular stichocytes, each with a duct into the pharyngeal lumen (**Figs. 1** and **2**). The pharynx extends one-fourth to nine-tenths of the body length in various taxa and is almost devoid of musculature. The region of the pharynx surrounded by stichocytes is known as the stichosome. Recent phylogenetic analysis based on a synthesis of molecular and morphological data suggests that the stichosome may be an example of parallel evolution and that the trichinellids and mermithids are more appropriately separated as two orders: Trichinellida, with at least six families, and Mermithida, with two families. However, debate on the higher classification of nematodes continues, and opinion will no doubt refocus as new data emerge and are applied. *See* MEDICAL PARASITOLOGY; NEMATA (NEMATODA).

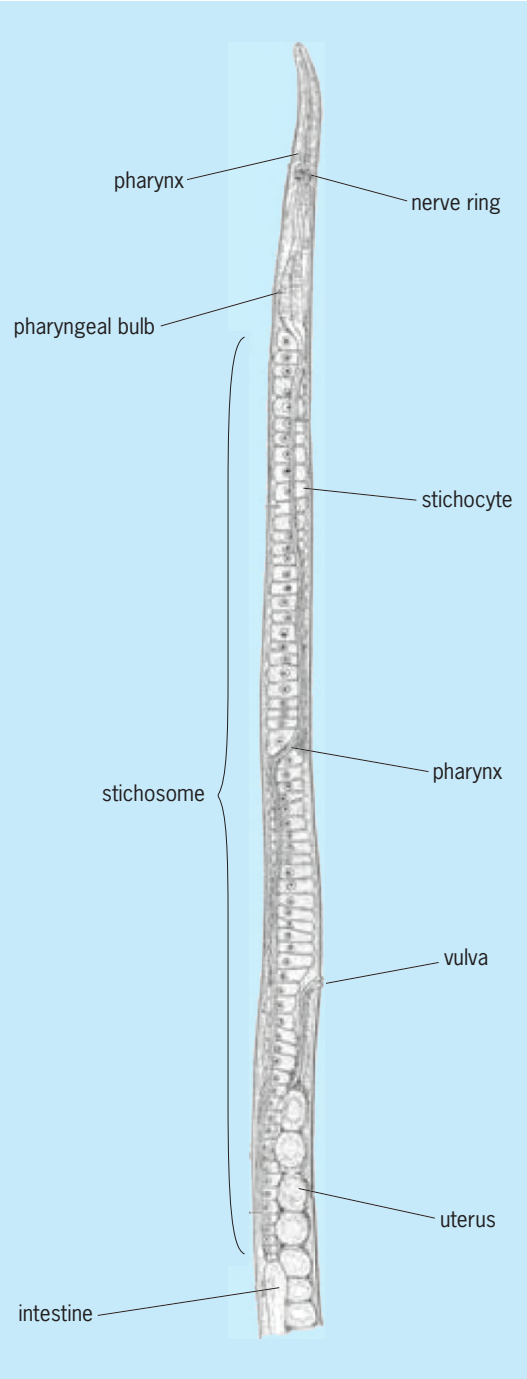

Fig. 1. Anterior end of *Trichinella spiralis* showing stichostome. (*Adapted from B. G. Chitwood, 1930*)

Two life history examples provide some insight into the biology of these organisms. In *Trichinella spiralis*, encysted larvae are ingested in infected muscle tissue (raw or undercooked pork is the classic example in the case of trichinosis in humans). The cyst surrounding the larva is digested in the new host, and the larvae molt to adults, mate, and embed in the intestinal epithelium; then females produce eggs which hatch (1000 larvae per female in 5 days). The hatched larvae are distributed via the circulatory system and migrate into surrounding cells; the larvae die unless they are in striated muscle fibers. Secretory products of the stichocytes induce DNA endoreduplication (replication of nuclear DNA without mitotic cell division) and transformation of the muscle fiber into a multinucleate nurse cell, which becomes encapsulated by collagen and supplied with capillaries. The life cycle continues when the muscle is eaten by another animal.

In *Romanomermis culicivorax*, the preparasitic larva is equipped with a stylet which allows penetration through the cuticle of the mosquito larva host. The stylet is lost in subsequent larval stages,

Fig. 2. Portion of the stichosome of the whipworm, *Trichuris trichiura*. (*Photomicrograph by Sung-Jong Hong, The Korean Society for Parasitology*)

and the stichosome separates from the intestine, which becomes a trophosome (a modified nutritive storage organ) with no anal opening to the exterior. Nutrient uptake from the insect hemolymph is through a very thin cuticle, perhaps enhanced by small-diameter pores. Interestingly, the epidermal cells adjacent to the cuticle have their absorptive surface area increased by outwardly directed microvilli. The continued role of the stichosome is unclear, although it appears to be involved in protein synthesis. The larva undergoes one molt in the insect host and then ruptures the body wall as it emerges. After a final molt to the adult stage, the female deposits eggs.

The nature and function of stichocyte secretions, particularly of vertebrate parasites, constitute an area of continued interest in research on host-parasite relationships. Howard Ferris

Bibliography. B. G. Chitwood, The structure of the esophagus in the Trichuroidea, *J. Parasitol.*, 17:35–42, 1930; P. De Ley and M. Blaxter, Systematic position and phylogeny, pp. 1–30 in D. L. Lee (ed.), *The Biology of Nematodes*, Taylor and Francis, London and New York, 2002; D. D. Despommier, *Trichinella* and *Toxocara*, pp. 597–607 in F. E. G. Cox, J. P. Kreier and D. Wakelin (eds.), vol. 5: *Parasitology*, in *Topley and Wilson's Microbiology and Microbial Infections*, Arnold, London, 1998; D. L. Lee, Life cycles, pp. 61–72 in D. L. Lee (ed.), *The Biology of Nematodes*, Taylor and Francis, London and New York, 2002; E. A. Munn and P. D. Munn, Feeding and digestion, pp. 211–232 in D. L. Lee (ed.), *The Biology of Nematodes*, Taylor and Francis, London and New York, 2002; G. O. Poinar, *The Natural History of Nematodes*, Prentice Hall, Englewood Cliffs, NJ, 1983; G. O. Poinar and R. Hess, *Romanomermis culicivorax*: Morphological evidence of transcuticular uptake, *Exp. Parasitol.*, 42:27–33, 1977.

Stickleback

Any member of the fish family Gasterosteidae, order Gasterosteiformes. The body may be elongate, naked, or with bony scutes along the sides. Three to 16 well-developed isolated dorsal spines precede a normal dorsal fin having 6 to 14 rays. Additional identifying features are as follows: 12 or 13 caudal fin rays; a single spine and 1 or 2 soft rays in the pelvic fin; 3 branchiostegal rays; circumorbital ring incomplete posteriorly; epineurals present; and 28 to 42 vertebrae. A maximum length of about 18 cm (7 in.) is reached in *Spinachia spinachia*. *See* GASTEROSTEIFORMES; OSTEICHTHYES; TELEOSTEI.

Sticklebacks are small freshwater and marine fishes of the cold and temperate waters of the Northern Hemisphere. All species are of some economic importance since their diet includes mosquito larvae, and the freshwater species are especially popular aquarium fishes. Also, sticklebacks are famous as subjects of numerous studies regarding evolution, genetics, behavior, and physiology. *See* INSTINCTIVE BEHAVIOR; SEXUAL DIMORPHISM.

Notable species. The fifteen-spined stickleback (*Spinachia spinachia*) and the four-spined stickleback (*Apeltes quadracus*) are marine shorefishes. *Spinachia spinachia* is found from Norway to the Bay of Biscay (France-Spain). The adult is about 18 cm (7 in.) in length and has a life span of 1 year. In low tides waters, the male builds a nest of algae stuck together with his mucus secretions. The female enters the nest and deposits 150 to 200 eggs, which are then fertilized by the male. The female dies shortly after egg deposition, and the male guards the nest and eggs. *Apeltes quadracus* occurs in freshwaters along the Atlantic coast from Nova Scotia to Virginia.

Pungitius pungitius (**Fig. 1**) measures about 7.5 cm (3 in.) in length and occurs in fresh and brackish waters of North America and Europe. This species is known as the nine-spine stickleback, but the number of spines varies. Nest building is relegated to the male, which also protects the eggs and fry until they hatch in about 1 week and protects the fry.

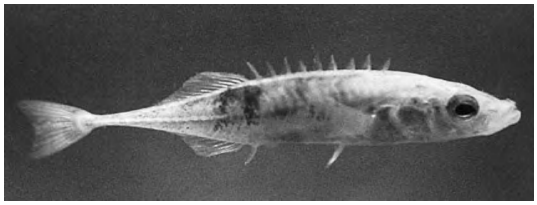

Fig. 1. Nine-spine stickleback (*Pungitius pungitius*). (*Photo by Konrad P. Schmidt*)

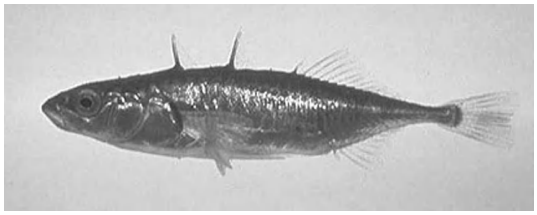

Fig. 2. Three-spine stickleback (*Gasterosteus aculeatus*). (*Bob Jenkins and Noel Burkhead; EFISH <http://www.cnr·vt.edu/efish/>*)

The best-known species is the three-spine stickleback (*Gasterosteus aculeatus*) [**Fig. 2**], the behavior of which has been extensively studied. The male makes a depression in the sandy bottom; here he builds a nest of vegetation cemented together by mucous strands produced by the kidneys, which are enlarged at breeding time. During the breeding season, the abdomen of the males becomes red and, when exhibited in the territory of the nesting area, evokes an innate behavioral response and attracts the females. The male also performs an elaborate courtship dance to entice the female to enter the nest. When the female does so, the male nuzzles her near the base of the tail and thus stimulates her to oviposit. The male then fertilizes the eggs and remains to guard and aerate them. *See* REPRODUCTIVE BEHAVIOR. Herbert Boschung; Charles B. Curtin

Bibliography. M. A. Bell and S. A. Foster (eds.), *The Evolutionary Biology of the Threespine Stickleback*, Oxford University Press, 1994; R. Froese and D. Pauly (eds.), FishBase, World Wide Web electronic publication, version (05/2005); J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; R. J. Wootton, *The Biology of the Sticklebacks*, Academic Press, 1976; R. J. Wootton, *A Functional Biology of Sticklebacks*, University of California Press, 1984.

Stilbite

A mineral belonging to the zeolite family of silicates. It crystallizes in the monoclinic system in crystals that are tabular parallel to the side pinacoid. Most characteristic are sheaflike aggregates of thin tabular crystals. There is perfect cleavage parallel to the side pinacoid, and here the mineral has a pearly luster; elsewhere the luster is vitreous. The color is usually white but may be brown, red, or yellow. Hardness is $3^1/_2$–4 on Mohs scale; specific gravity is 2.1–2.2. Stilbite is a calcium-sodium aluminum silicate, $Ca(Al_2Si_7O_{18}) \cdot 7H_2O$, but usually contains some sodium replacing calcium.

Stilbite is a secondary mineral usually found in cavities in basalts and related rocks. Much less commonly it is found in granites, gneisses, and metal-bearing veins. It is associated with other zeolites, datolite, prehnite, and calcite. Some notable localities are in Iceland, India, Scotland, Nova Scotia, and in the United States at Bergen Hill, New Jersey, and the Lake Superior copper district in Michigan. *See* SILICATE MINERALS; ZEOLITE. Cornelius S. Hurlbut, Jr.

Stirling engine

An engine in which work is performed by the expansion of a gas at high temperature to which heat is supplied through a wall. Like the internal combustion engine, a Stirling engine provides work by means of a cycle in which a piston compresses gas at a low temperature and allows it to expand at a high temperature. In the former case the heat is provided by the internal combustion of fuel in the cylinder, but in the Stirling engine the heat (obtained from externally burning fuel) is supplied to the gas through the wall of the cylinder (**Fig. 1**). *See* INTERNAL COMBUSTION ENGINE.

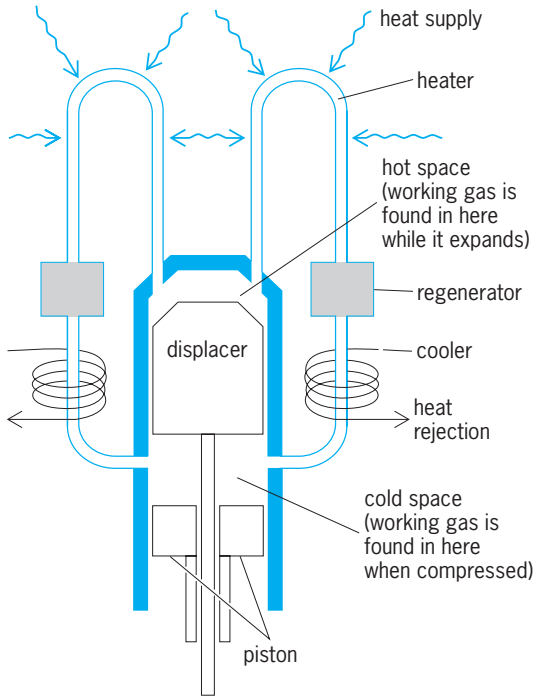

Fig. 1. Principle of Stirling engine, displacer type.

The rapid changes desired in the gas temperature are achieved by means of a second piston in the cylinder, called a displacer, which in moving up and down transfers the gas back and forth between two spaces, one at a fixed high temperature and the other at a fixed low temperature. When the displacer in **Fig. 2** is raised, the gas will flow from the hot space via the heater and cooler tubes into the cold space. When it is moved downward, the gas will return to the hot

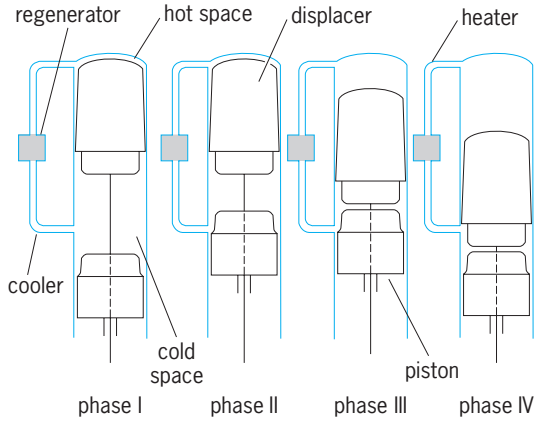

Fig. 2. Stylized Stirling process.

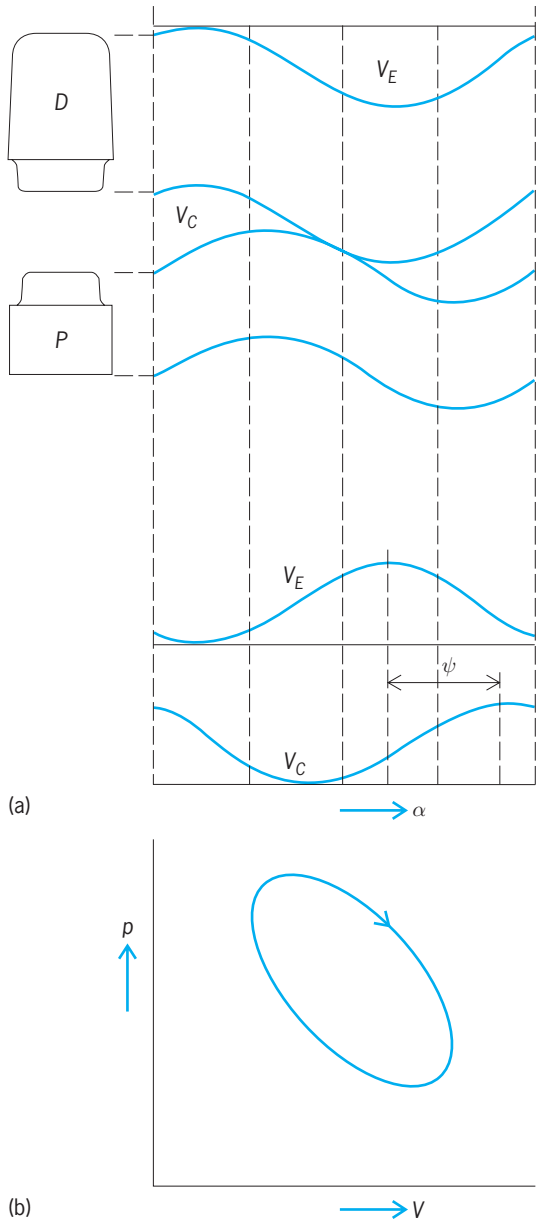

Fig. 3. Piston and displacer movement for a Stirling engine. (*a*) The continuous movement of the piston and displacer shown as a function of crank angle α. There is no clear-cut division between the various phases of the cycle. The variations in volume of the hot space V_E and the cold space V_C are plotted separately in the lower part of the diagram. (*b*) The *P–V* diagram of the cycle.

space along the same path. During the first transfer stroke the gas has to yield up a large amount of heat to the cooler; an equal quantity of heat has to be taken up from the heater during the second stroke.

The regenerator shown in Figs. 1 and 2 has been inserted between the heater tube and cooler tube in order to prevent unnecessary waste of this heat. It is a space filled with porous material to which the hot gas gives up its heat before entering the cooler; the cooled gas recovers the stored heat on its way back to the heater.

The displacer system serves to heat and cool the gas periodically; associated with it is a piston which compresses the gas while it is in the cold space of the cylinder and allows it to expand while in the hot space. Since expansion takes place at a higher temperature than compression, it produces a surplus of work over that required for the compression.

Stirling cycle. Any practical version of the engine will embody some kind of crank and connecting rod mechanism, in consequence of which there will be no sharp transitions between the successive phases indicated in Fig. 2; but this will not alter the principle of the cycle (nor detract from its efficiency). If, for the sake of simplicity, the piston and displacer are assumed to move discontinuously, the cycle can be divided into the following four phases. (1) The piston is in its lowest position, the displacer in its highest. All the gas is in the cold space. (2) The displacer is still in its highest position; the piston has compressed the gas at low temperature. (3) The piston is still in its highest position; the displacer has moved down and transferred the compressed gas from the cold to the hot space. (4) The hot gas has expanded, pushing the pistons, followed by the displacer, to their lowest positions. The displacer is about to rise and return the gas to the cold space, the piston remaining where it is, to give phase I again.

The actual piston and displacer movements might be as indicated in **Fig. 3**, which shows that the only essential condition for obtaining a surplus of work is that the maximum volume of the hot space occur before that of the cold space. This condition shows that more configurations with pistons and cylinders are possible than just the type with displacer in order to get a Stirling cycle. One of the most compact systems is shown in **Fig. 4**, which is the system known as the double-acting engine. In the double-acting engine there is a hot space (expansion space) at the top and a cold one (compression space) at the bottom of each of the four cylinders shown. The hot space of a cylinder is connected to the cold one through a heater, a regenerator, and a cooler. The pistons P_n of the cylinders move with a suitable phase shift between them. In the case of four cylinders, as shown in Fig. 4, this shift is 90°.

The theory of an actual engine is very complicated. In order to provide an understanding of the quantities that play a role, the formulas of the power and efficiency of an engine of highly idealized form will be given. On the assumption that the volumes of the hot (expansion) space V_E and that of the cold (compression) space V_C (Fig. 3) vary with the crank angle in a purely sinusoidal way, that both expansion and compression are isothermal at respectively T_E and T_C, and that all kinds of losses caused by the flow resistances in the tubes and regenerator, heat losses in the regenerator, and so on are ignored, the power P can be expressed as in Eq. (1). Here $\tau = T_C/T_E$,

$$P = (1-\tau)\frac{\pi n}{60} V_{E_{\max}} \cdot p_m \frac{\delta}{1+\sqrt{1-\delta^2}} \sin\Theta \qquad (1)$$

n is the speed in rpm, p_m is the mean pressure of the pressure variation, and δ and Θ are functions of the temperature ratio, swept volume ratio, dead spaces, and phase angle between the two swept volumes. Under these conditions the thermal efficiency η is, of course, that of the Carnot cycle, and is expressed in Eq. (2).

$$\eta = \frac{T_E - T_C}{T_E} = 1 - \tau \qquad (2)$$

See CARNOT CYCLE; THERMODYNAMIC CYCLE.

Engine design. Actual engines have been built in the Philips Laboratories at Eindhoven, Netherlands, as prototypes in the range of 10–500 hp (7.5–375 kW) per cylinder. After years of research on the Stirling engine, actual thermal efficiencies of 30–45% (depending on the specific output and temperature ratio) and a specific power of 1.88 hp/in.^{3} (85 kW/liter) swept volume of the piston were obtained with the displacer type of engine equipped with rhombic drive, as shown in **Fig. 5**. In the figure the piston and the displacer drive concentric rods, which are coupled to the rhombic drive turning twin timing gears. The cooler, regenerator, and heater are arranged as annular systems around the cylindrical working space. The heater tubes surround the combustion chamber. In the preheater the gas at 1470°F (800°C) from the heater is cooled to 300–390°F (150–200°C) while heating the combustion air to about 1200°F (650°C).

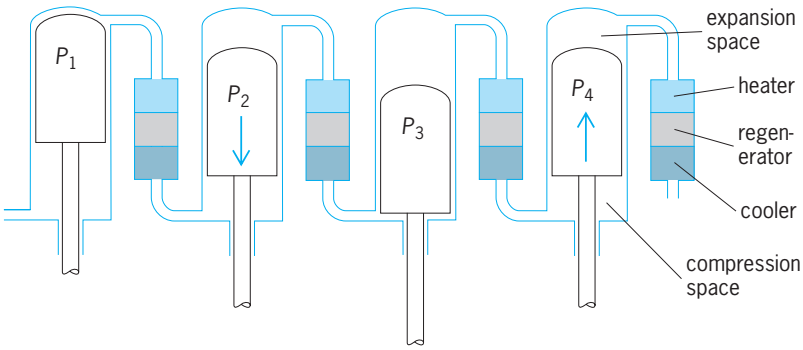

Fig. 4. Principle of the double-acting engine.

The rhombic drive mechanism allows complete balancing even of a single-cylinder engine and of a separate buffer space, thus avoiding heavy forces acting on the drive. The results of measurements on the first engine of this type, with hydrogen as the working fluid, are shown in **Fig. 6**. An approximate heat balance is output, 40%; exhaust and radiation, 10%; and heat rejection by cooling water, 50%. Control of engine output is by regulation of the pressure of the working fluid in the engine, while the temperature of the heater is kept constant by a thermostat;

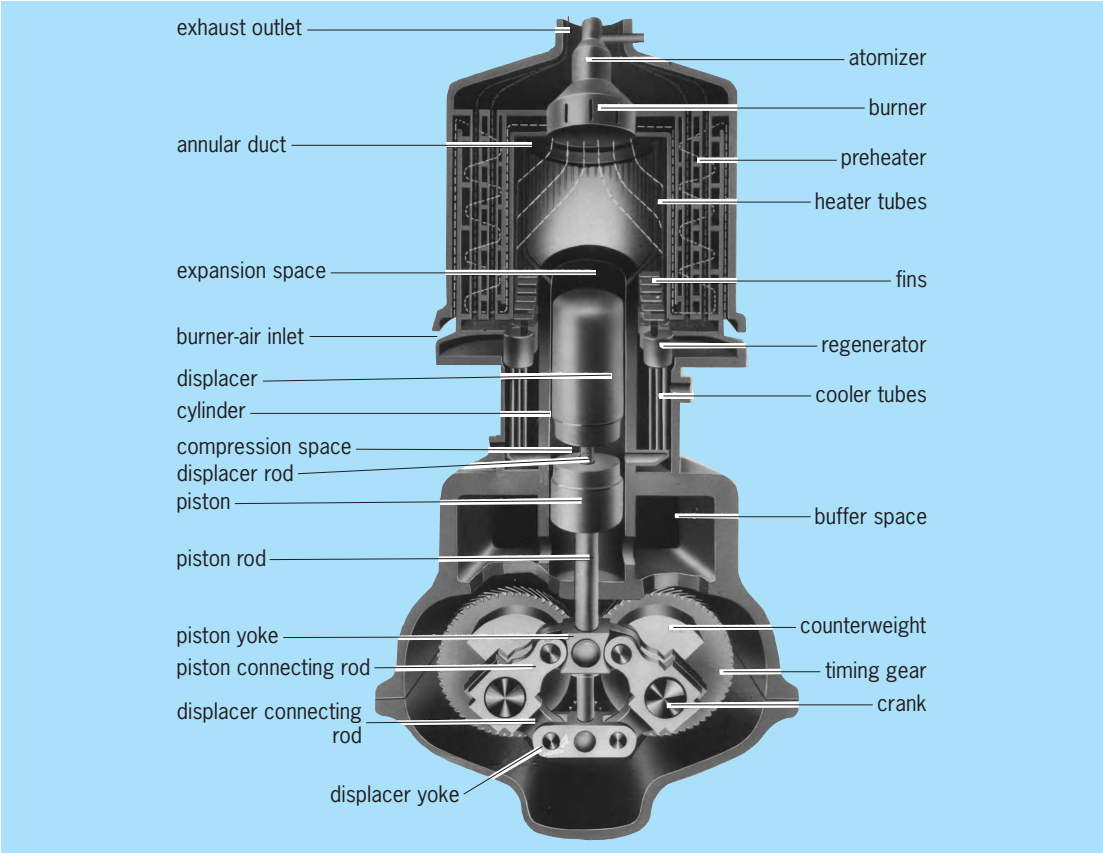

Fig. 5. Philips Stirling engine in cross section. (*Philips Gloeilampenfabrieken*)

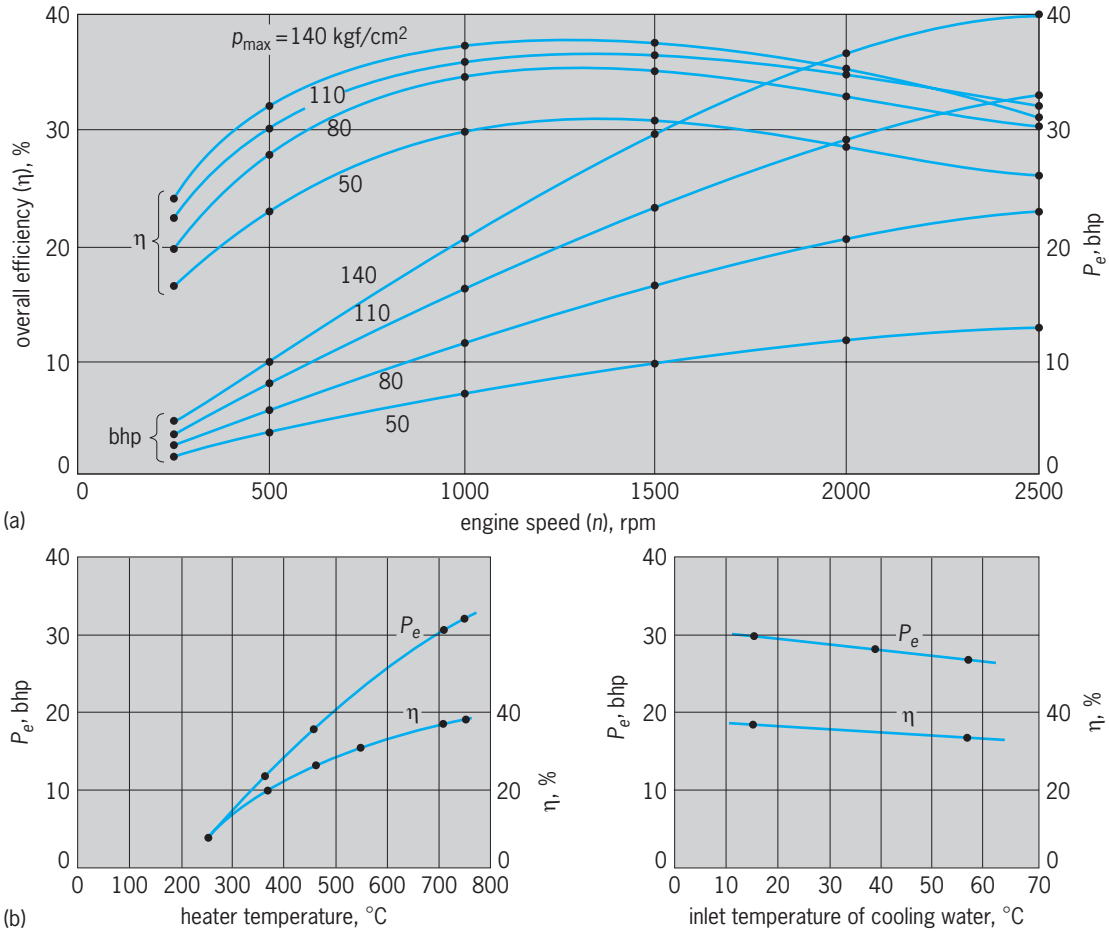

Fig. 6. Test measurements for the Stirling engine. (*a*) Measured shaft power P_e and efficiency of a 40-hp (30-kW) single-cylinder test engine with rhombic drive, plotted as a function of the engine speed *n*, at different values of the maximum pressure of the working fluid in the engine (p_{max}). 1 hp = 0.75 kW; 1 kgf/cm² = 98 kPa = 14.2 lbf/in.² (*b*) Power and efficiency of the 40-hp (30-kW) test engine, given as a function of the heater temperature and as a function of the inlet temperature of the cooling water. The curves apply to *n* = 1500 rpm and p_{max} = 140 kgf/cm² = 13.7 MPa = 1990 lbf/in.²; °F = (°C × 1.8) + 32.

hence the efficiency shows little dependence on the load.

The closed system of the Stirling engine endows this engine with many advantages and also some shortcomings. The continuous external heating of the closed system makes it possible to burn various kinds of liquid fuels and gases, without any modification whatsoever. This multifuel facility can be demonstrated with a 10-hp or 7.5-kW (at 3000 rpm) generator set (**Fig. 7**). The engine can operate on alcohol, various lead-containing gasolines, diesel fuel, lubricating oil, olive oil, salad oil, crude oil, propane, butane, and natural gas. Furthermore, it allows combustion to take place in such a way that air pollution is some orders of magnitude less than that due to internal combustion engines. Through the intermediary of a suitable heat transport system (for example, heat pipes) any heat source at a sufficiently high temperature can be used for this engine—radioisotopes, a nuclear reactor, heat storage, solar heat, or even the burning of coal or wood.

The almost sinusoidal cylinder pressure variation and continuous heating make the Stirling engine very quiet in operation. An engine having four or more cylinders gives a virtually constant torque per revolution, as well as a constant dynamometer torque over a wide speed range, which is particularly valuable for traction purposes. The present configuration makes complete balancing possible, thus eliminating vibrations. There is no oil consumption and virtually no contamination because a new type of seal for the reciprocating rods shuts off the cycle hermetically from the drive mechanism. **Figure 8** shows an engine of this configuration.

Where direct or indirect air cooling is required, the closed cycle has the drawback that more heat has to be removed from the cooler than in comparable engines with open systems, where a greater quantity of heat inevitably escapes through the exhaust.

If it is envisaged as someday taking the place of existing engines, the Stirling might be ideal as a propulsion engine in yachts and passenger ships, and in road vehicles, such as city buses, where a large radiator is acceptable. The system of continuous

Fig. 7. Phillips Stirling engine with generator to demonstrate its multifuel capacity. (*Philips Gloeilampenfabrieken*)

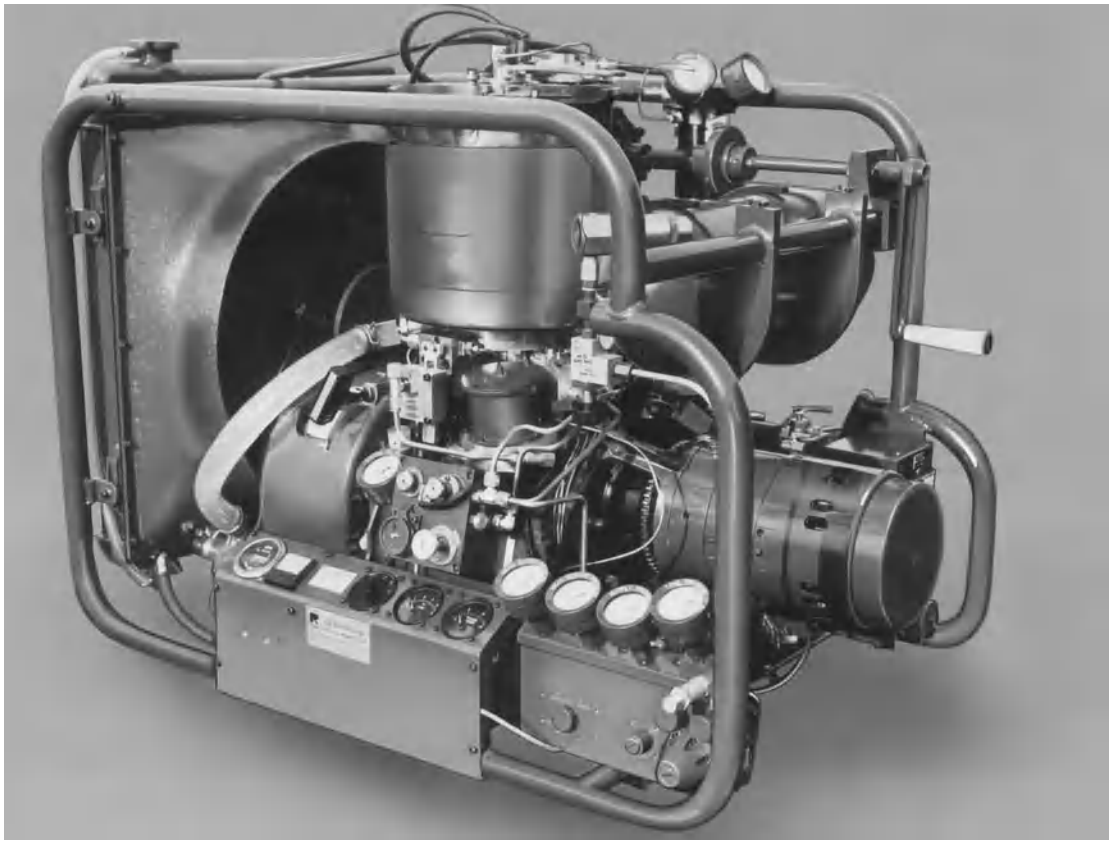

Fig. 8. A 3-kW Stirling engine generator set. (*General Motors Corp.*)

external heating is also able to open fields of application inaccessible to internal combustion engines. Roelof J. Meijer

Bibliography. R. Darlington, *Stirling and Hot Air Engines*, 2005; T. Finkelstein and A. J. Organ, *Air Engines: The History, Science, and Reality of the Perfect Engine*, 2001; W. R. Martini, *Stirling Engine Design Manual*, 2004.

Stishovite

Naturally occurring stishovite, SiO_2, is a mineral formed under very high pressure with the silicon atom in sixfold, or octahedral, coordination instead of the usual fourfold, or tetrahedral, coordination. It was discovered in the same type of Coconino sandstone from the Meteor Crater of Arizona. Stishovite has since been found in rocks from the Ries Crater of Germany. The presence of stishovite indicates formation pressures in excess of 10^6 lb/in.^{2} (7.5 gigapascals). The possibility of the existence of stishovite at great depths strongly influences the interpretations of geophysicists and solid-state physicists regarding the phase transitions of mineral matter, as well as the interpretation of seismic data in the study of such regions of the interior of the Earth. *See* SILICA MINERALS.

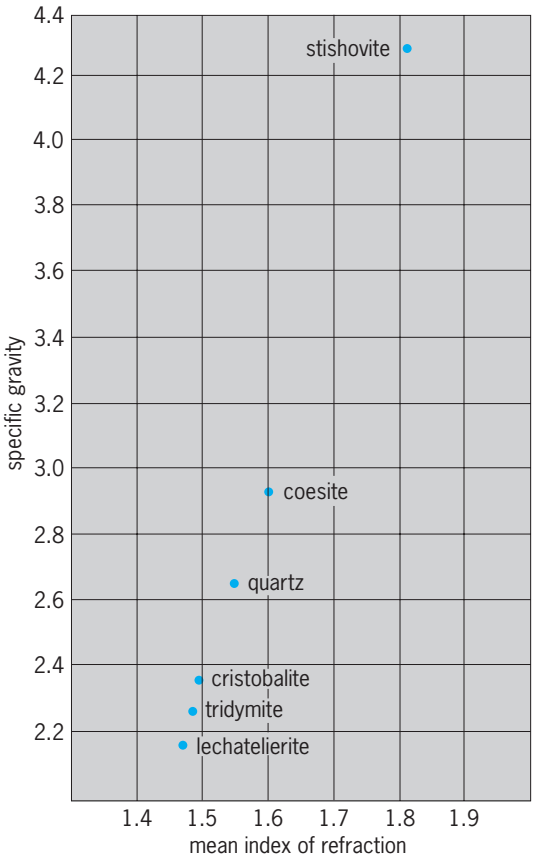

Comparison of the mean index of refraction and specific gravity of stishovite and other polymorphic modifications of silica.

Natural occurrence. Stishovite occurs in submicrometer size in very small amounts (less than 1% of the rock) in samples of Coconino sandstones from the Meteor Crater of Arizona, which contains up to 10% of coesite, the other high-pressure polymorph of silica. Besides stishovite and coesite, these rocks consist mainly of quartz, with small amounts of silica glass (lechatelierite) and traces of zircon and, in some cases, rutile. Because of its extremely fine grain size and because of the sparsity of this mineral in the rock, positive identification of the mineral is possible only by the x-ray diffraction method after chemical concentration. *See* COESITE; METEORITE.

Properties. Stishovite is much more resistant to attack by concentrated hydrofluoric acid than coesite. It can therefore be extracted or concentrated by prolonged treatment of the stishovite-bearing sandstone in hydrofluoric acid. Concentrates of stishovite, all of submicrometer size, are pale gray and have a roughly prismatic habit. Because of its fine-grained habit, only the mean index of refraction of the natural material can be measured. It is 1.80.

Crystallographically, stishovite is SiO_2 in the rutile structure, tetragonal with $a = 0.41790$ nanometer and $c = 0.26649$ nm (with a standard error of ± 0.00004 nm). The axial ratio is c/a is 0.6377, and the volume of the tetragonal unit cell is 0.046541 nm^3. The specific gravity of stishovite, calculated from the x-ray data, is 4.28, as reported by Chao and others, compared with the value of 4.35 for the synthetic material reported by Stishov and Popova. It is 46% denser than coesite and much denser than other modifications of silica (see **illus.**). Edward C. T. Chao

Bibliography. G. Beggerow, *High-Pressure Properties of Matter*, 1980; E. C. T. Chao et al., Stishovite, SiO_2, a very high pressure new mineral from Meteor Crater, Arizona, *J. Geophys. Res.*, 67(1):419–421, 1962; P. J. Heaney, C. T. Prewitt, and G. V. Gibbs (eds.), *Silica: Physical Behavior, Geochemistry, and Materials Applications (Reviews in Mineralogy, Vol. 29)*, 1994.

Stochastic control theory

A branch of control theory which aims at predicting and minimizing the magnitudes and limits of the random deviations of a control system output through optimizing the design of the controller. Such deviations occur when random noise and disturbance processes are present in a control system, so that the system does not follow its prescribed course but deviates from the latter by a randomly varying amount.

In contrast to deterministic signals, random signals cannot be described as given functions of time such as a step, a ramp, or a sine wave. The exact function is unknown to the system designer; only some of its statistical properties are known.

A random signal may be generated by one of nature's processes, for instance, radar noise and wind- or wave-induced forces and moments on a radar

antenna or a ship. Alternatively, a random signal may be generated by human intelligence, for instance, the bearing of a zigzagging aircraft, or the contour to be followed by a duplicating machine.

One outstanding experimental fact about nature's random processes is that these signals are approximately gaussian. The word gaussian refers to a mathematical concept which describes one or more signals $i_1, i_2, \ldots, i_n$ having the following properties:

1. The amplitude of each signal is normally distributed.

2. The joint distribution function of any number of signals at the same or different times taken from the set is a multivariate normal distribution. This experimental fact is not surprising in view of the fact that a random process of nature is usually the sum total of the effects of a large number of independent contributing factors. For instance, thermal noise is due to the thermal motions of billions of electrons and atoms. An ocean-wave height at any particular time and place is the sum of wind-generated waves at previous times over a large area. Gaussian signals have the property that when passed through linear systems they remain gaussian. *See* DISTRIBUTION (PROBABILITY); ELECTRICAL NOISE; STOCHASTIC PROCESS.

In the following, the mathematical characterization of the random processes, their transmission properties, measurements and simulation, effects on the system output, and control system optimization will be introduced.

Mathematical characterization. The random processes encountered in control work can be classified as stationary and nonstationary processes. The underlying mechanism that generates a random process can usually be described in physical or mathematical terms. For instance, the underlying mechanism that generates shot-effect noise is thermionic emission; the underlying mechanism that generates ocean waves is essentially wind force in conjunction with the Earth's gravity; and the outcome of dice throwing is determined by a probability of 1/6 for each face of each die independent of all others.

If the generating mechanism does not change with time, any measured average property of the random process such as mean or variance is independent of the time of measurement aside from statistical fluctuations, and the random process is called stationary. For instance, ocean-wave height in a given sea state, thermal noise, and a telegraphic signal of a certain language are examples of stationary random processes. If the generating mechanism does change, the random process is called nonstationary, in which case the generating mechanism may change in a predetermined fashion or at random. An example of the former is found in missile launching, where the unwanted fluctuations in vertical and lateral thrusts are generated by one or a group of rocket motors of which the operating condition is a predetermined function of time. Alternatively, the generating mechanism may change at random, as in the course of flight of an aircraft, where the wind disturbance depends on the local atmospheric condition which may be calm or turbulent.

Statistical ensemble. The random processes or events generated by the same underlying mechanism form a statistical ensemble. The ensemble average is the averaged value over all possible events weighted according to the probabilities of such events. For a stationary process there is little reason to distinguish between the ensemble average at any particular time and the time average of a single event. The life history of a single event can be chopped up into an infinite number of events. Since these events are generated by the same underlying mechanism which does not change with time, they can be thought of as arbitrarily selected from the entire ensemble, and therefore they represent the latter. The assumption of equivalence between time and ensemble averages is called the ergodic hypothesis. *See* STATISTICAL MECHANICS; TOPOLOGICAL DYNAMICS.

The ensemble average of a random variable or function is referred to as its mean value. For stationary systems, the mean value can be obtained by taking either time or ensemble average. Let $\bar{x}$ denote the mean value of a random variable x, and $\tilde{x}$ its random variation from the mean: $x = \bar{x} + \tilde{x}$. As $\bar{x}$ can be absorbed as part of the deterministic input, the random signals may be defined as variations about the mean values, and thus they have zero mean themselves.

Correlation functions. The statistical properties of a set of random signals (which may include the random inputs and system variables in response to these inputs) can be described by their correlation functions. The correlation function $\phi_{ab}(t_1,t_2)$ is defined as the ensemble average of signal $a(t)$ at time t_1 and $b(t)$ at time t_2, Eq. (1). When the two signals $a(t)$ and $b(t)$

$$\phi_{ab}(t_1, t_2) \equiv \overline{a(t_1)b(t_2)} \tag{1}$$

are different, $\phi_{ab}(t_1,t_2)$ is called the cross-correlation function. When $a(t)$ and $b(t)$ are one and the same, $\phi_{ab}(t_1,t_2)$ is called the autocorrelation function.

For stationary systems, the only significant independent variable is $t_2 - t_1$, denoted as τ, Eq. (2). The

$$\phi_{ab}(\tau) = \overline{a(t)b(t+\tau)} \tag{2}$$

time interval τ is called the correlation time. The useful properties of the correlation function given in Eqs. (3)–(7) follow from Eq. (2). Equation (6)

$$\phi_{aa}(0) = \overline{a(t)^2} \tag{3}$$

$$\phi_{ab}(-\tau) = \phi_{ba}(\tau) \tag{4}$$

$$\phi_{aa}(-\tau) = \phi_{aa}(\tau) \tag{5}$$

$$|\phi_{ab}(\tau)| \le \sqrt{\phi_{aa}(0)\phi_{bb}(0)} \tag{6}$$

$$|\phi_{aa}(\tau)| \le \phi_{aa}(0) \tag{7}$$

is obtained by taking the ensemble average of Eq. (8), yielding Eq. (9), and combining this with

Eq. (10). Equation (7) is a special case of Eq. (6) with $b(t) = a(t)$.

$$0 \le \left\{ \frac{|a(t)|}{\sqrt{\phi_{aa}(0)}} - \frac{|b(t+\tau)|}{\sqrt{\phi_{bb}(0)}} \right\}^2$$
$$= \frac{a(t)^2}{\phi_{aa}(0)} - \frac{2\,|a(t)b(t+\tau)|}{\sqrt{\phi_{aa}(0)\phi_{bb}(0)}} + \frac{b(t+\tau)^2}{\phi_{bb}(0)} \quad (8)$$

$$0 \le 2 - \frac{\overline{2\,|a(t)\,b(t+\tau)|}}{\sqrt{\phi_{aa}(0)\,\phi_{bb}(0)}} \quad (9)$$

$$|\phi_{ab}(\tau)| = \overline{|a(t)b(t+\tau)|} \quad (10)$$

Equations (3)–(7) show that the autocorrelation function is even in τ, and its value at $\tau = 0$ is the mean-square value of the random signal itself. The magnitude of the cross-correlation function cannot exceed the product of the root-mean-square values of the two signals.

Spectral-density functions. If the chopped signals $a_T(t)$ and $b_T(t)$ are defined as the same random signals $a(t)$ and $b(t)$, respectively, in the interval $0 \le t \le T$ but zero outside this interval, and $A_T(s)$ and $B_T(s)$ represent their Laplace transforms, then a function $\Phi_{ab}(s)$ can be defined by Eq. (11). Equation (11) implies that Eqs. (12) and (13) are valid. Equations (12)

$$\Phi_{ab}(s) \equiv \lim_{T\to\infty} \frac{1}{T}\overline{A_T(-s)B_T(s)} \quad (11)$$

$$\Phi_{ab}(j\omega) = \Phi_{ba}(-j\omega) \quad (12)$$

$$\Phi_{aa}(j\omega) = \Phi_{aa}(-j\omega)\rangle 0 \quad (13)$$

and (13) show that $\Phi_{ab}(j\omega)$ is the complex conjugate of $\Phi_{ba}(j\omega)$ and that $\Phi_{aa}(j\omega)$ is a real positive even function of ω. The function $\Phi_{ab}(s)$ is called the spectral-density function when a and b are the same, or the cross-spectral-density function when $a(t)$ and $b(t)$ represent different signals.

It is readily shown that $\phi_{ab}(\tau)$ and $\Phi_{ab}(s)$ are related by Laplace transforms, Eqs. (14) and (15). The

$$\phi_{ab}(\tau) = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \Phi_{ab}(s)e^{s\tau}\,ds \quad (14)$$

$$\Phi_{ab}(s) = \int_{-\infty}^{\infty} \phi_{ab}(\tau)e^{-s\tau}\,d\tau \quad (15)$$

transform relationship between correlation functions and spectral densities as represented by Eqs. (14) and (15) is called Wiener's theorem.

Transmission properties. When stochastic processes are applied to a physical system, the random output variations are related to the inputs and to each other. These relationships are expressed in terms of the transmission properties.

The **illustration** shows a multiple-input and multiple-output system where the vector $\mathbf{x}(t)$ denotes inputs $x_1(t), x_2(t), \ldots, x_n(t)$ and $\mathbf{y}(t)$ denotes outputs $y_1(t), y_2(t), \ldots, y_m(t)$, through filters with transfer functions $H_{ik}(s)$, Eq. (16). The filters are

$$Y_i(s) = \sum_k H_{ik}(s)X_k(s) \quad (16)$$

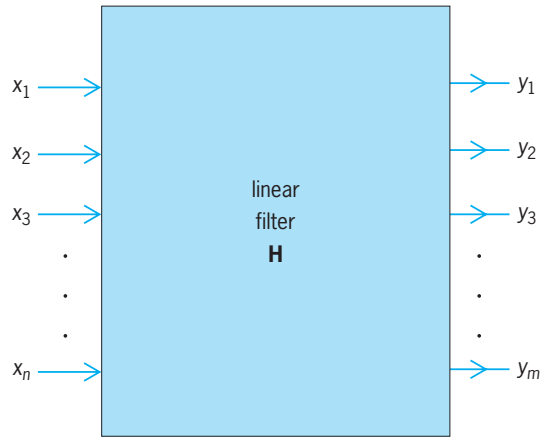

Transmission of stochastic processes through the linear filter H.

assumed to be stable (with all their poles in the left half of the s-plane). The spectral densities of components of $\mathbf{y}$ are expressed in terms of that of $\mathbf{x}$ by Eq. (17). Equation (17) is readily derived from

$$\Phi_{y_iy_j}(j\omega) = \sum_k \sum_l H_{ik}(-j\omega)\,\Phi_{x_kx_l}(j\omega)H_{jl}(j\omega) \quad (17)$$

Eq. (11) and can be written in matrix form, Eq. (18),

$$\Phi_y(j\omega) = \mathbf{H}(-j\omega)\Phi_x(j\omega)\mathbf{H}^T(j\omega) \quad (18)$$

where $\mathbf{\Phi}_y$, $\mathbf{\Phi}_x$, and $\mathbf{H}$ are $m \times m$, $n \times n$, and $m \times n$ matrices with $\Phi_{y_iy_j}$, $\Phi_{x_ix_j}$, H_{ij} as their i–jth components respectively. Note that $\mathbf{H}(-j\omega)$ and $\mathbf{H}^T(j\omega)$ are hermitian adjoints of each other. Equations (17) and (18) are perfectly general and can be used to obtain cross-spectral density between input and output of a filter as well. For instance, by regarding $\mathbf{x}$ as the $m+1$- to $m+n$-th components of an enlarged output vector, Eq. (18) readily leads to Eq. (19), where

$$\Phi_{xy}(j\omega) = \Phi_x(j\omega)\mathbf{H}^T(j\omega) \quad (19)$$

$\boldsymbol{Phi}_{xy}$ is an $n \times m$ matrix with $\boldsymbol{Phi}_{x_iy_j}$ as its i-jth component. *See* LINEAR SYSTEM ANALYSIS.

If the process is nonstationary, frequency domain analysis cannot be used. Equation (18) is replaced by Eq. (20), where $h_{kl}(t,\tau)$ denotes the response of filter

$$W_y(t,t') = \int_{-\infty}^{t}\int_{-\infty}^{t'} \mathbf{h}(t,\tau)\phi_x(\tau,\tau') \times \mathbf{h}^T(t',\tau')\,d\tau\,d\tau' \quad (20)$$

η_{kl} at time t to unit impulse input at time τ.

White stochastic process. A stochastic process x with n components $x_1, x_2, \ldots, x_n$ is said to be white if there is no correlation between any of its components at different times, Eq. (21). This includes auto-

$$\phi_{ik}(t,t') = 0 \text{ if } t \ne t' \quad (21)$$

correlations with $i = k$. Equation (21) has an interesting implication: at $t = t'$, the autocorrelation function for a white process must have infinite instantaneous amplitude if it is to produce any physically significant effect.

This can be proven as follows: A physical system can be represented by (or approximated as) a stable filter with finite impulse response $h(t,\tau)$. The mean-square value of its response $a(t)$ to a white noise input $x(t)$ can be obtained from Eq. (20), yielding Eq. (22). If it is assumed to the contrary that $x(t)$ is

$$\overline{[a(t)]^2} = \phi_{aa}(t, t) = \int_{-\infty}^{t}\int_{-\infty}^{t} h(t, \tau)h(t, \tau')\phi_{xx}(\tau, \tau')\, d\tau\, d\tau' \quad (22)$$

finite, then $\phi_{xx}(\tau,\tau') \leq M^2$. This condition, together with Eq. (21), gives the integral of Eq. (22) to be zero.

Equation (21) and the requirement of finite physical effect lead to the correlation function of Eq. (23)

$$\phi_{ik}(t, t') = W_{ik}(t)\delta(t - t') \quad (23)$$

for the white noise process, where the matrix $\mathbf{W}(t)$ with $W_{ik}(t)$ as its elements is positive semidefinite for all t (that is, $\mathbf{v}^T\mathbf{W}\mathbf{v} \geq 0$ for any vector $\mathbf{v}$ with real components); and $\delta(t)$ is the Dirac function defined by Eqs. (24) and (25). For a stationary white process,

$$\delta(t) = 0 \text{ if } t \neq 0 \quad (24)$$

$$\int_{-a}^{a} \delta(t)\, dt = 1 \text{ for all } a\rangle 0 \quad (25)$$

Wiener's theorem and Eq. (23) give the spectral density as Eq. (26).Equation (26) illustrates the origin of

$$\Phi_{ik}(j\omega) = W_{ik} \quad (26)$$

the description "white" in the sense that the spectral density is independent of frequency.

A white process is purely a mathematical concept as no physical process can have infinite amplitude or infinite bandwidth. It is nevertheless a very important concept in simulation of stochastic processes and in system optimization. A random process can be regarded as white in an approximate sense if Eq. (21) is satisfied for $|t-t'| > \Delta$, where Δ is a time interval much smaller than the response time of the control system under consideration. An equivalent condition is for Eq. (26) to hold for all $\omega < \omega_k$, where ω_k is at least an order of magnitude larger than the system bandwidth.

Simulation of stochastic processes. A white process can be simulated either digitally or by an analog method. Digitally it is simulated by a sequence of random numbers at a rate at least a decade higher than the system bandwidth. The numbers are then used directly in a simulation study or converted into an electrical signal in actual tests. In analog white noise generators, shot effect noise or thermal noise is amplified by a wide-band amplifier. Correlated white processes with matrix $\mathbf{W}$ are generated by using independent white noise sources and a resistive network. The required number of independent sources is equal to the rank of $\mathbf{W}$.

Nonwhite stochastic processes can be simulated as outputs of filters with white noise sources as inputs. Assuming $Phi_x(j\omega)$ to be the identity matrix, Eqs. (18) and (20) give respectively Eqs. (27) and (28). The problem is then to find $H(j\omega)$ so that $\phi_y(j\omega)$

$$\phi_y(j\omega) = \mathbf{H}(-j\omega)\mathbf{H}^T(j\omega) \quad (27)$$

$$\phi_y(t, t') = \int_{-\infty}^{t} \mathbf{h}(t, \tau)\mathbf{h}^T(t', \tau)\, d\tau \quad t' \geq t \quad (28)$$

is equal to the spectral density matrix of the stochastic processes to be simulated or to find $h(t,\tau)$ for the nonstationary case. *See* SIMULATION.

Regulator problem. The deterministic regulator problem is encountered in the optimal control of linear systems. Its state equation may be modified to include a white stochastic process $\mathbf{w}(t)$, as shown in Eq. (29). Here $\mathbf{x}(t)$ is the state vector, $\mathbf{u}(t)$ is the

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u} + \mathbf{w}(t) \quad (29)$$
$$\mathbf{x}(t_0) \text{ specified}$$

input vector, and $\mathbf{w}(t)$ is a white stochastic process with correlation function given by Eq. (23). The performance index J, as defined by Eq. (30), where $\mathbf{S}_1$

$$J = \mathbf{x}^T(t_1)\mathbf{S}_1\mathbf{x}\,(t_1) + \int_{t_0}^{t_1} \mathbf{x}^T\mathbf{Q}(t)\mathbf{x} + \mathbf{u}^T\mathbf{R}(t)\mathbf{u}\, dt \quad (30)$$

and $\mathbf{Q}(t)$ are symmetric, nonnegative definite matrices which penalize the system for large excursions, depends now on $\mathbf{w}(t)$. Therefore its expected value or ensemble average over all possible $\mathbf{w}(t)$ must be used, Eq. (31).

$$\text{Performance index} = E(J) \quad (31)$$

See OPTIMAL CONTROL (LINEAR SYSTEMS).

As $\mathbf{w}(t)$ is white, its future value has no correlation with its past, and consequently has no correlation with $\mathbf{x}(t)$. Furthermore, $\mathbf{w}(t)$, $t > t_0$, is completely unpredictable and nothing can be done about it. The best that can be done is to minimize the contribution to $E(J)$ by $\mathbf{x}(t)$. Thus, the optimal control law of Eq. (32) with $\mathbf{S}$ defined by the Riccati differential equation (33) and the boundary condition $\mathbf{S}(t_1) = \mathbf{S}_1$

$$\mathbf{u}(t) = -\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{S}(t)\mathbf{x}(t) \quad (32)$$

$$\dot{\mathbf{S}} + \mathbf{S}\mathbf{A}(t) + \mathbf{A}^T(t)\mathbf{S} - \mathbf{S}\mathbf{B}(t)\mathbf{R}^{-1}(t)\mathbf{B}^T(t)\mathbf{S} + \mathbf{Q}(t) = 0 \quad (33)$$

is still valid. This heuristic reasoning has been substantiated by rigorous mathematical proof.

However, the minimum value of $E(J)$ is different from that given by Eq. (34) because of the added

$$J_{\min}(\mathbf{x}(t_0), t_0) = \mathbf{x}^T(t_0)\mathbf{S}(t_0)\mathbf{x}(t_0) \quad (34)$$

disturbance $\mathbf{w}(t)$. Straightforward calculations give Eq. (35), where Tr stands for the trace of the matrix inside the braces.

$$\mathbf{E}(J)_{\min} = \mathbf{x}^T(t_0)\mathbf{S}(t_0)\mathbf{x}(t_0) + \int_{t_0}^{t_1} \mathrm{Tr}\,\{\mathbf{S}(t)\mathbf{W}(t)\}\, dt \quad (35)$$

In practical control systems, the state vector $\mathbf{x}(t)$ is not exactly known because of instrument noise and such, and the disturbance $\mathbf{w}(t)$ may not have white spectrum. However, with any given $Phi_{\mathrm{W}}(j\omega)$ or $\phi_{\mathrm{W}}(t,t')$, $\mathbf{w}(t)$ can be simulated as the output of linear filters with white noise processes $\mathbf{v}(t)$ as its inputs, as in Eqs. (36) and (37). A generalized vec-

$$\dot{\mathbf{w}}(t) = \mathbf{C}(t)\,\mathbf{w}(t) + \mathbf{v} \tag{36}$$

$$\phi_v(t, t') = \delta(t - t') \tag{37}$$

tor $\mathbf{z}(t)$ is defined by Eq. (38). If matrices $\mathbf{D}$, $\mathbf{E}$, and $\mathbf{F}$ are defined by Eqs. (39), where $\mathbf{1}$ and $\mathbf{0}$ denote

$$\mathbf{z}(t) = (\mathbf{x}(t)\mathbf{w}(t)) \tag{38}$$

$$\mathbf{D}(t) = \begin{pmatrix} \mathbf{A}(t) & \mathbf{1} \\ \mathbf{0} & \mathbf{C}(t) \end{pmatrix}$$
$$\mathbf{E}(t) = \begin{pmatrix} \mathbf{B}(t) \\ \mathbf{0} \end{pmatrix} \qquad \mathbf{F} = \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \end{pmatrix} \tag{39}$$

the identity matrix and null matrix of appropriate dimensions, then Eqs. (29) and (36) can be combined into one, giving Eq. (40). Since $\mathbf{Fv}$ is a white distur-

$$\dot{\mathbf{z}}(t) = \mathbf{D}(t)\mathbf{z}(t) + \mathbf{E}(t)\mathbf{u} + \mathbf{Fv} \tag{40}$$

bance, the problem is reduced to an optimal regulator problem with white disturbance. However, $\mathbf{z}(t)$ is not completely known.

One of the most important results in estimator theory is that the residue error $\mathbf{z}(t) - \hat{\mathbf{z}}(t)$, where $\hat{\mathbf{z}}(t)$ is the optimum estimate of $\mathbf{z}(t)$, cannot be inferred in any way from past measurements. Consequently the best that can be done is to apply regulatory action to the known part of $\mathbf{z}(t)$, using Eq. (41), where

$$\mathbf{u}(t) = -\mathbf{R}^{-1}(t)\mathbf{E}^T(t)\mathbf{S}(t)\hat{\mathbf{z}}(t) \tag{41}$$

$\mathbf{S}(t)$ is the solution to Riccati equation for the new problem defined by Eq. (40), and $\hat{\mathbf{z}}(t)$ is the best or optimum estimate of $\mathbf{z}(t)$. Equation (41) states that the optimal overall controller is obtained by applying the optimal control law to the separately optimally estimated state $\hat{\mathbf{z}}(t)$. It is known as the separation theorem.

For single-input, single-output, time-invariant systems with infinite time interval, both the estimation problem and the control problem may be reformulated, leading to an integral equation known as the Wiener-Hopf equation. The solution to this equation has been obtained and is a classical mathematical result in the theory of stochastic processes. Further, the method discussed above as well as the Wiener-Hopf approach may also be applied to discrete-time systems. *See* ESTIMATION THEORY.

Structural uncertainty. In Eq. (29), the random variable $\mathbf{w}(t)$ is an additive term. Quite often the system structure is not completely known, and $\mathbf{A}(t)$ or $\mathbf{B}(t)$ has stochastic, uncertain, or unknown, components. These are known as structural uncertainty. If the structural uncertainty is white gaussian, the optimal control law remains the same, with one or more terms added to the Riccati equation. If the structural uncertainty is nonwhite, or constant but unknown, it can be predicted to some extent from past measurements. Adaptive controllers are generally used. If the structural uncertainty varies in some unknown manner, minimax, least-sensitive, or guaranteed-cost controllers are used. Minimax means minimizing the largest possible value of the performance index. Least-sensitive means designing the controller in such a way that the performance index is not sensitive to the structural uncertainty. Guaranteed-cost aims at simple robust design which guarantees that the performance index does not exceed some given limit. *See* ADAPTIVE CONTROL; CONTROL SYSTEMS; OPTIMAL CONTROL THEORY.

S. S. L. Chang; Gerald Cook

Bibliography. A. Bagchi, *Optimal Control of Stochastic Systems*, 1993; A. Benoussian, *Stochastic Control of Partially Observable Systems*, 1992; J. A. Borrie, *Stochastic Systems for Engineers*, 1992; G. R. Chen, G. Chen, and S.-H. Hsu, *Linear Stochastic Control Systems*, 1995; M. J. Grimble and M. A. Johnson, *Optimal Control and Stochastic Estimation: Theory and Applications*, 2 vols., 1988; F. L. Lewis and V. L. Syrmos, *Optimal Estimation: With an Introduction to Stochastic Control Theory*, 2d ed., 1995; R. F. Stengel, *Stochastic Optimal Control: Theory and Application*, 1986.

Stochastic process

A physical stochastic process is any process governed by probabilistic laws. Examples are (1) development of a population as controlled by mendelian genetics; (2) brownian motion of microscopic particles subjected to molecular impacts or, on a different scale, the motion of stars in space; (3) succession of plays in a gambling house; and (4) passage of cars by a specified highway point.

In each case, a probabilistic system is evolving; that is, its state is changing with time. Thus the state at time t depends on chance: It is a random variable $x(t)$. The parameter set of values of t involved is usually (and will always be in this article) either an interval (continuous parameter stochastic process) or a set of integers (discrete parameter stochastic process). Some authors, however, apply the term stochastic process only to the continuous parameter case.

If the state of the system is described by a single number, $x(t)$ is numerical-valued. In other cases, $x(t)$ may be vector-valued or even more complicated. The discussion in this article will usually be restricted to the numerical case. As the state changes, its values determine a function of time, the sample function; and the probability laws governing the process determine the probabilities assigned to the various possible properties of sample functions.

A mathematical stochastic process is a mathematical structure inspired by the concept of a physical stochastic process, and studied because it is a mathematical model of a physical stochastic process or

because of its intrinsic mathematical interest and its applications both in and outside the field of probability. The mathematical stochastic process is defined simply as a family of random variables. That is, a parameter set is specified, and to each parameter point t a random variable $x(t)$ is specified. If one recalls that a random variable is itself a function, if one denotes a point of the domain of the random variable $x(t)$ by ω, and if one denotes the value of this random variable at ω by $x(t,\omega)$, it results that the stochastic process is completely specified by the function of the pair (t,ω) just defined, together with the assignment of probabilities. If t is fixed, this function of two variables defines a function of ω, namely, the random variable denoted by $x(t)$. If ω is fixed, this function of two variables defines a function of t, a sample function of the process.

Probabilities are ordinarily assigned to a stochastic process by assigning joint probability distributions to its random variables. These joint distributions, together with the probabilities derived from them, can be interpreted as probabilities of properties of sample functions. For example, if t_0 is a parameter value, the probability that a sample function is positive at time t_0 is the probability that the random variable $x(t_0)$ has a positive value. The fundamental theorem at this level is that, to any self-consistent assignment of joint probability distributions, there corresponds a stochastic process.

The concept of a stochastic process is so general that the study of stochastic processes includes all of probability theory. Although it would be impossible to make a useful but more restrictive definition, what the probabilist usually has in mind in using the term stochastic process (unless he is interested in the mathematical foundations of the general theory) is a stochastic process whose random variables have some sort of interesting mutual relations. For example, one such relation is that of independence, and one type of stochastic process which was studied long before the term was invented is a sequence of independent random variables. For historical reasons, such a sequence is not commonly thought of as a stochastic process (although its continuous parameter analog, a process with independent increments, to be discussed below, is). The rest of this article is devoted to a general discussion of specific types of stochastic processes which have received the most attention, because they are important in mathematical and nonmathematical applications.

Stationary processes. These are the stochastic processes for which the joint distribution of any finite number of the random variables is unaffected by translations of the parameter; that is, the distribution of $x(t_1 + h), \ldots, x(t_n + h)$ does not depend on h. For a more complete discussion *see* PROBABILITY

Markov processes. A Markov process is a process for which, if the present is given, the future and past are independent of each other. More precisely, if $t_1 < \cdots < t_n$ are parameter values, and if $1 < j < n$, then the sets of random variables $[x(t_1), \ldots, x(t_{j-1})]$ and $[x(t_{j+1}), \ldots, x(t_n)]$ are mutually independent for given $x(t_j)$. Equivalently, the conditional probability distribution of $x(t_n)$ for given $x(t_1), \ldots, x(t_{n-1})$ depends only on the specified value of $x(t_{n-1})$ and is in fact the conditional probability distribution of $x(t_n)$, given $x(t_{n-1})$. An important and simple example is the Markov chain, in which the number of states is finite or denumerably infinite. (The terminology varies somewhat here.) One simple type of discrete-parameter Markov chain is the following. Let (p_{ij}) be a set of numbers, where i and j range over a finite or infinite set of integers. (In physical language, the number p_{ij} will be the probability that some system has a transition from state i to state j in one step.) The numbers p_{ij} are to satisfy relation (1). The ran-

$$p_{ij} \geq 0 \qquad \sum_j p_{ij} = 1 \tag{1}$$

dom variables of the associated Markov process are integral-valued, denoted by $x(0), x(1), \ldots$. If i_0 is prescribed as the initial state, that is, if $x(0)$ is assigned the value i_0 identically, the probability that $x(k)$ has the value i_k, for $k = 1, \ldots, N$, is product (2). For

$$p_{i_0 i_1} \cdots p_{i_{N-1} i_N} \tag{2}$$

this process, p_{ij} is the probability that $x(n + 1)$ has the value j if $x(n)$ has the value i. The number p_{ij} is also (and this is the characteristic property of Markov processes) the probability that $x(n + 1)$ has the value j if $x(n)$ has the value i and if also $x(n - 1)$ has any prescribed value a_1, $x(n - 2)$ the value a_2, and so on. What makes this a special Markov chain, aside from the fact that the states are denoted by integers, is that the conditional probability just described does not depend on n. The chain is therefore described as having stationary transition probabilities. If the initial state is given a distribution, say by prescribing that $x(0)$ have the value i with probability p_i, evaluation (2) becomes sum (3).Here the p_i's are any nonnega-

$$\sum_i p_i p_{ii_1} \cdots p_{i_{N-1} i_N} \tag{3}$$

tive numbers with sum 1. If the initial distribution is chosen, as is not always possible, in such a way that the probability that $x(n)$ has the value j is p_j, not only for $n = 0$ but for all values of n, the resulting process is stationary.

In constructing the corresponding continuous-parameter Markov chain, it is supposed that, for each pair (i,j), there is a function $p_{ij}(.)$, defined for strictly positive t, satisfying relations (4). The equations of

$$\begin{gathered} p_{ij} \geq (t)0 \qquad \sum_j p_{ij}(t) = 1 \\ p_{ij}(s + t) = \sum_k p_{ik}(s)\, p_{kj}(t) \end{gathered} \tag{4}$$

the system in the last line are known as the Chapman-Kolmogorov equations. A Markov stochastic process with continuous parameter ranging from 0 to ∞ can be constructed for which, if again p_j is the probability that $x(0)$ has the value j, and if $0 < t_1 < \cdots < t_n$, the probability that $x(t_k)$ has the value i_k for

$k = 1, \cdots, N$ is given by sum (5). For this process,

$$\sum_i p_i p_{ii_1}(t_1) p_{i_1 i_2}(t_2 - t_1) \cdots p_{i_{N-1} i_N}(t_N - t_{N-1}) \quad (5)$$

if $s > 0$, the probability that $x(u + s)$ has the value j, if $x(u)$ has the value i, is $p_{ij}(s)$. The number $p_{ij}(s)$ is also the probability (and this is the characteristic property of Markov processes) that $x(u + s)$ has the value j if $x(u)$ has the value i; and if also $x(u_1)$ has any specified value a_1,$x(u_2)$, the value a_2, and so on, where $u_1, u_2, \ldots,$ are any positive numbers less than u. This example is not the general continuous-parameter Markov chain because the transition probability just described does not depend on u, that is, because the chain has stationary transition probabilities. The process is stationary if the probability that $x(u)$ has the value j does not depend on u, for all j. The second line of relation (4) has a simple interpretation: The probability, if $x(u)$ has the value i, that $x(u + s + t)$ has the value j, is the sum over k of the probability that $x(u + s)$ has the value k multiplied by the probability that, if $x(u)$ has the value i, and if $x(u + s)$ has the value k, then $x(u + s + t)$ has the value j. Without the Markov property, the second factor might depend on i. If the number of states is not finite or denumerably infinite, the preceding discussion is modified by replacing sums in notations (1), (3), (4), and (5) by integrals.

Typical applications. Typical questions that have been raised, and solved to a varying degree, about Markov processes with stationary transition probabilities are the following. They are phrased in the continuous-parameter case, for the Markov chain just described, and it is assumed that relation (6) holds. For convenience one defines $p_{ij}(0)$ as 1 if $i = j$ and as 0 otherwise.

$$\lim_{t \to 0} P_{ij}(t) = 1 \quad \text{for all } i \quad (6)$$

1. Does $p_{ij}(t)$ have a limit when $t \to \infty$? In other words, for each j is there a limiting probability that the system is in state j as time passes? The answer is yes, and the limiting probability depends only on the end state j, not on the initial state i, if transitions between all pairs of states are possible.

2. What are the asymptotic properties of $p_{ij}(t)$ as $t \to 0$? The answer is that $p'_{ij}(0)$ always exists and is finite except possibly when $i = j$. Under further hypotheses on the transition probability functions, always satisfied if there are only finitely many states, $p'_{ii}(0)$ is finite and relations (7) hold. These equations

$$\begin{aligned} p'_{ij}(t) &= \sum_k p_{ik}(t)\, p'_{kj}(0) \\ p'_{ij}(t) &= \sum_k p'_{ik}(0)\, p_{kj}(t) \end{aligned} \quad (7)$$

can be used to determine the transition probability functions in terms of assigned derivatives when $t = 0$. For example, if c is a strictly positive constant, and if $p'_{ij}(0)$ is specified as c if $j = i + 1$, as $-c$ if $j = i$, and as 0 otherwise, it is shown that $p_{ij}(t) = 0$ if $j < i$, and that otherwise relation (8) is true. The process

$$p_{ij}(t) = \frac{(ct)^{j-i} e^{-ct}}{(j - i)!} \quad (8)$$

with these transition probabilities is known as the Poisson process.

3. What are the properties of the sample functions? Under further restrictions on the process, the sample functions are constant on intervals, changing in jumps from one state to the next, and $-p'_{ij}(0)/p'_{ii}(0)$ is the probability that, if the system is in state i, its next jump will be into state j. If it is in the ith state, the time the system remains in this state thereafter is a random variable with density qe^{-tq}, where $q = -p'_{ii}(0)$. For example, in the case of the Poisson process described above, it is shown that, under proper normalization, and if $x(0) = 0$, the sample functions are integral-valued and monotone, increasing in unit jumps. This process is a mathematical model for the physical process of radioactive decay. That is, $x(t)$ can be interpreted as the number of radioactive disintegrations of a substance by time t. In other interpretations, $x(t)$ is taken as the number of telephone calls initiated by time t, or the number of cars that have passed a given highway point by time t. The constant c is the rate at which these various events occur. In fact, the expected value of $x(u + h) - x(u)$, that is, the expected number of events in a time interval of length h (h here is of course positive) is ch, and the probability that an event will occur in an interval of length h, regardless of the past history of the process, is ch up to higher powers of h.

There are many special types of Markov chains, for which more detailed questions become important. For example, consider the branching processes. In a system of particles, all of the same type, a particle will occasionally split, independently of its past history and of the other particles, into $j \geq 0$ particles with probability q_j. If a particle is observed at time t, the probability that it will split by time $t + h$ is ch, up to higher powers of h, where c is a strictly positive constant. Then the number of particles at time t is a random variable $x(t)$. The $x(t)$ process is a Markov process, and $p'_{ij}(0)$ is easily determined in terms of q_j and c. A more general branching process would permit particles of several types, and each particle would be allowed to split into particles of the various types. The rate c would depend on the particle type. In this case, $x(t)$ is defined as a vector whose ith component is the number of particles of type i at time t. The $x(t)$ stochastic process is a vector-valued Markov process. In studying branching processes, the most natural questions to ask are: What is the probability that the population will die out? If it does not die out, what is the asymptotic distribution of population as time passes? The answers are too technical to be given here.

Transition probabilities. If the states of a Markov process comprise all real numbers, the character of the process may be similar to that of a chain but may also be quite different. For example, the

sample functions of the process may be continuous. The most important examples of this type are the diffusion processes. Simplifying somewhat, but not assuming stationary transition probabilities, consider a Markov process for which the probability distribution of the state at time t, given state ξ at time $s < t$, has density $p(s, \xi, t, \eta)$. Then the basic conditions satisfied by the transition density, corresponding to relation (1), are relations (9). Now suppose that limits (10) and (11) exist and have the indicated

$$p(s, \xi, t, \eta) \geq 0 \qquad \int_{-\infty}^{\infty} p(s, \xi, t, \eta)\, d\eta = 1$$
$$p(s, \xi, t, \eta) = \int_{-\infty}^{\infty} p(s, \xi, u, \varsigma) p(u, \varsigma, t, \eta)\, d\xi \quad \text{if } s < u < t \tag{9}$$

$$\lim_{h \downarrow 0} \int_{-\infty}^{\infty} p(s, \xi, s+h, \eta)\,(\eta - \xi) \frac{d\eta}{h} = m(s, \xi) \tag{10}$$

$$\lim_{h \downarrow 0} \int_{-\infty}^{\infty} p(s, \xi, s+h, \eta)\,(\eta - \xi)^2 \frac{d\eta}{h} = \sigma^2(s, \xi) \tag{11}$$

values. These limit relations make m and σ^2 the instantaneous rates of change of the displacement and its variance, given a specified time and state. If m and σ are sufficiently regular, and if a further condition is imposed which, roughly, makes improbable significant sample function changes in short times, the corresponding Markov process, when properly normalized, will have continuous sample functions. Moreover, the transition density will then satisfy backward-diffusion equation (12) and forward diffusion equation (13), also known as the Fokker-Planck

$$\frac{\partial p(s, \xi, t, \eta)}{\partial s} + m(s, \xi) \frac{\partial p}{\partial \xi} + \frac{1}{2}\sigma^2(s, \xi) \frac{\partial^2 p}{\partial \xi^2} = 0 \tag{12}$$

$$\frac{\partial p(s, \xi, t, \eta)}{\partial s} + \frac{\partial}{\partial \eta}[m(t, \eta)p] - \frac{1}{2}\frac{\partial^2}{\partial \eta^2}[\sigma^2(t, \eta)p] = 0 \tag{13}$$

equation. Conversely, given a pair of coefficient functions m and 0 and σ these second-order parabolic equations can be used to derive the corresponding transition densities.

The simplest nontrivial example of a diffusion process corresponds to the specification $m = 0$ and σ a constant function. In this case, the diffusion process is the brownian motion process, or Wiener process: The increment $x(t) - x(s)$ has a gaussian distribution with mean value 0 and variance $\sigma^2|t - s|$. This is a mathematical model for the physical brownian motion. That is, if $x(t) - x(s)$ represents the displacement in a given direction of a brownian particle between times s and t, this process is a good model for the actual motion.

Martingales. A martingale is a stochastic process with the property that, if $t_1 < \cdots < t_n$ are parameter values, the expected value of $x(t_n)$, for given $x(t_1), \cdots, x(t_{n-1})$ is equal to $x(t_{n-1})$. That is, the expected future value, given present and past values, is equal to the present value. The interpretation that a martingale can be thought of as the fortune of a player after the successive plays of a fair gambling game is obvious.

Typical results on martingales are the following. If a sequence of random variables is a martingale, it converges under weak conditions which impose certain "bounds" on the random variables, for example, if the expectation of the absolute value of the nth random variable is bounded independently of n. The sample functions of a properly normalized continuous-parameter martingale do not have oscillatory discontinuities.

The applications of martingale theory are too technical to be given here, but one suggestive example will be given, which indicates at least how the theory can be usefully applied to information theory. Let $y, x_1, x_2, \ldots$ be random variables, and let y_n be the expected value of y knowing $x_1, \ldots, x_n$. Then y_n is a random variable which is a function of $x_1, \ldots, x_n$ and the sequence $y_1, y_2, \ldots$ is a martingale. That is, the expected values of a random variable, if one knows more and more, defines a sequence of random variables which is a martingale.

Processes with independent increments. Such a process is a continuous-parameter process with the property that, if $t_1 < \cdots < t_n$ are parameter values, the successive increments in notation (14) are mutually notation (14) independent.

$$x(t_2) - x(t_1), \ \ldots, x(t_n) - x(t_{n-1}) \tag{14}$$

If $y(t) = x(t) - x(t_0)$, where t_0 is fixed, the $y(t)$ process is then a Markov process. Both the Poisson and the brownian motion processes described above have independent increments.

Typical results on these processes include the following. If such a process is properly normalized, its sample functions do not have oscillatory discontinuities. Moreover, the distribution of any increment $x(t) - x(s)$ is infinitely divisible, and there is a standard form for the characteristic function of any such distribution. *See* GAME THEORY; INFORMATION THEORY; LINEAR PROGRAMMING; OPERATIONS RESEARCH.

Joseph L. Doob

Bibliography. R. N. Bhattacharya and E. C. Waymire, *Stochastic Processes with Applications*, 1990; J. L. Doob, *Stochastic Processes*, 1953, paper 1990; G. F. Lawler, *Introduction to Stochastic Processes*, 1995; J. Y. Medhi, *Stochastic Processes*, 2d ed., 1994; M. K. Ochi, *Applied Probability and Stochastic Processes in Engineering and the Physical Sciences*, 1990; A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3d ed., 1991; P. Todorovic and J. O. Berger, *An Introduction to Stochastic Processes and Their Applications*, 1992.

Stoichiometry

The field of chemistry that includes all chemical measurements, such as the measurements of atomic and molecular weights and sizes, gas volumes, vapor

densities, deviation from the gas laws, and the structure of molecules. In the long struggle to determine the relative weights of the atoms, scientists relied upon combining ratios, specific heats, and measurements of gas volumes. All such measurements, and the calculations that relate them to each other, constitute the field of stoichiometry. Since measurements are expressed in mathematical terms, stoichiometry can be considered to be the mathematics of general chemistry. Thus, stoichiometry is not part of inorganic, organic, physical, or analytical chemistry, but is an essential part of all of them. Chemistry is an exact science, and it depends upon exact measurements of weights, lengths, and volumes, and on the amounts of energy which are absorbed or evolved in chemical reactions.

In a more usual usage, the term stoichiometry refers to the relationships between the measured quantitites of substances or of energy involved in a chemical reaction; the calculations of these quantities include the assumption of the validity of the laws of definite proportions and of conservation of matter and energy. *See* CONSERVATION OF ENERGY; CONSERVATION OF MASS.

Laboratory measurements are made in terms of a unit of weight (grams, ounces, and such), volume (milliliters and such) or energy (ergs, calories, and such), but since atoms and molecules of different substances have different weights, it is necessary to convert numbers of atoms or molecules into the measured quantities. These quantities can be expressed in any convenient units. In laboratory work, weights are usually expressed in grams, but in chemical engineering, in pounds or tons. *See* ATOMIC MASS; GRAM-MOLECULAR WEIGHT; MOLECULAR WEIGHT.

A typical stoichiometric problem involves predicting the weight of reactant needed to produce a desired amount of a product in a chemical reaction. For example, phosphorus can be extracted from calcium phosphate, $Ca_3(PO_4)_2$, by a certain process with a 90% yield (some calcium phosphate fails to react or some phosphorus is lost). In a specific problem, it might be necessary to determine the mass of calcium phosphate required to prepare 16.12 lb of phosphorus by this process. The balanced equation for the preparation is shown in reaction (1). In this reac-

$$2Ca_3(PO_4)_2 + 10C + 6SiO_2 \rightarrow 6CaSiO_3 + 10CO + P_4 \quad (1)$$

tion, 2 moles of calcium phosphate are required to produce 1 mole of phosphorus. Two moles of calcium phosphate have a mass of 620 lb, and 1 mole of phosphorus as P_4 has a mass of 124 lb. Using these relationships, calculation (2) is made. Since the yield

$$\frac{2 \text{ moles } Ca_3(PO_4)}{1 \text{ mole } P_4} = \frac{620 \text{ lb } Ca_3(PO_4)_2}{124 \text{ lb } P_4} = \frac{X \text{ lb } Ca_3(PO_4)_2}{16.12 \text{ lb } P_4} \quad (2)$$

$$X = 80.6 \text{ lb}$$

of phosphorus is only 90%, extra $Ca_3(PO_4)_2$ must be used: 88.1 lb is the mass of calcium phosphate required to yield 16.12 lb of phosphorus by this process.

Calculations of this sort are important in chemical engineering processes, in which amounts and yields of products must be known. The same reasoning is used in calculations of energy generated or required. In this case, the energy involved in the reaction of a known weight of the material in question must be known or determined. There are many variations on this theme. For example, if it is necessary to determine the composition of a mixture of copper and zinc powders, a given weight of the material can be dissolved in hydrochloric acid, which reacts with the zinc, but not with the copper [reaction (3)]. From

$$Zn + 2HCl \rightarrow ZnCl_2 + H_2 \quad (3)$$

the volume of hydrogen gas liberated, the amount of zinc dissolved can be calculated, and hence the composition of the mixture. Again, if 5.0 g of sodium is placed in a vessel containing 5.0 g of chlorine, a violent reaction takes place [reaction (4)]. However,

$$2Na + Cl_2 \rightarrow 2NaCl \quad (4)$$

some of the sodium remains unattacked. It can be calculated that 5.0 g of chlorine can react with only 3.2 g of sodium; 1.8 g of sodium will remain. The chlorine is said to be the limiting reagent—when it is used up, the reaction must stop. *See* AVOGADRO'S LAW; COMBINING VOLUMES, LAW OF; DALTON'S LAW; DEFINITE COMPOSITION, LAW OF; MULTIPLE PROPORTIONS, LAW OF; SPECIFIC HEAT.

The calculations discussed in this article involve compounds in which the ratio of atoms is generally simple. For a discussion of compounds in which the relative number of atoms cannot be expressed as ratios of small whole numbers. *See* NONSTOICHIOMETRIC COMPOUNDS.. John C. Bailar, Jr.

Bibliography. J. A. Olmsted and G. M. Williams, *Chemistry*, 4th ed., 2004; K. C. Timberlake, *Chemistry*, 9th ed., 2005.

Stoker

A mechanical means for feeding coal into, and for burning coal in, a furnace. There are three basic types of stokers. Chain or traveling-grate stokers have a moving grate on which the coal burns; they carry the coal from a hopper into the furnace and move the ash out. Spreader stokers mechanically or pneumatically distribute the coal from a hopper at the furnace front wall and move it onto the grate which usually moves continuously to dispose of the ash after the coal is burned. Underfeed stokers are arranged to force fresh coal from the hopper to the bottom of the burning coal bed, usually by means of a screw conveyor. The ash is forced off the edges of the retort peripherally to the ashpit or is removed by hand. *See* STEAM-GENERATING FURNACE. George W. Kessler

Pages from

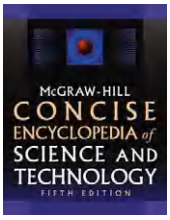

ISBN 0071429573
5th Edition (2005)

Stokes' theorem The assertion that under certain light restrictions the surface integral of $(\Delta \mathfrak{F} \times \mathfrak{F}) \cdot \mathbf{v}$ over a surface patch S is equal to the line integral of $\mathfrak{F} \cdot \tau$ taken around C, the boundary curve of S, provided the sense of transcription of C is right-handed relative to $\mathbf{v}$. This can be expresed as

$$\iint (\nabla \times \mathfrak{F}) \cdot \mathbf{v}\, dS = \int \mathfrak{F} \cdot \tau\, ds$$

Here $\mathfrak{F}$ is a vector function, $\mathbf{v}$ is one of the two unit normals to the two-sided surface S, s is arc length measured positively in the sense which is right-handed relative to $\mathbf{v}$, and τ is the unit tangent vector to C in the sense of increasing s. *See* CALCULUS OF VECTORS. [H.V.C.]

Stolonifera An order of the Alcyonaria which lacks coenenchyme. They form either simple or rather complex colonies. The polyp has a cylindrical body with a retractile oral portion which can withdraw into a solid anthostele or calyx protected by many calcareous spicules. The base of the mature polyp is attached to a creeping stolon which is a ribbonlike network or thin flat mat from which daughter polyps arise. Daughter polyps never bud from the wall of the primary polyp. Each polyp is connected by solenial tubes of the stolons, or by transverse platforms. *See* ALCYONARIA. [K.At.]

Stomach The tubular or saccular abdominal organ of the digestive system adapted for temporary food storage and preliminary stages of food breakdown.

In some primitive vertebrates the stomach may be little more than a simple tube quite similar to other portions of the gastrointestinal tract. In other forms the stomach is a distinct, and frequently large, saclike structure of variable shape. Carnivorous forms typically have a better-developed stomach than herbivores, probably reflecting the larger but less numerous feedings characteristic of the former, but exceptions are numerous.

In birds the stomach consists of a proventriculus and a gizzard. The former is well supplied with glands which secrete softening and digestive enzymes; the latter is a strong, muscular grinding organ whose action is often enhanced by the ingestion of small stones.

Mammals have stomachs which vary considerably in structure. Although a single chamber is most common, some mammals, such as cows and their relatives (ruminants), have as many as four. These chambers may have developed either from modifications of the posterior portions of the esophagus or from alterations of the stomach itself.

The human stomach is located beneath the diaphragm, through which the posterior, terminal end of the esophagus passes. The stomach appears as a dilated tube continuous with the distal end of the esophagus. The upper curvature of the stomach is usually above and to the left of the esophageal orifice. This expanded anterior portion is the fundus and is commonly filled with air or gas. The body (corpus) of the stomach is directed toward the attenuated right extremity or pyloric region and is subject to variations in size and shape, depending upon functional activities, habits, disease, and volume of the contents. The pyloric walls are marked by the heavy sphincter muscle which controls the passage of chyme (a semiliquid fluid produced by the mechanical and chemical changes of preliminary digestion) into the duodenum.

The stomach of vertebrates is lined by a mucous membrane that is usually thrown into longitudinal folds called rugae. Most of the surface is covered with mucus-secreting epithelial cells, but scattered throughout the lining are many small glandular pits which are lined with one or more types of secretory cells. *See* DIGESTIVE SYSTEM. [T.S.P.]

Stomatopoda The only order of the superorder Hoplocarida belonging to the subclass Malacostraca of the class Crustacea. This order of the mantis shrimps contains four families: Lysioquillidae, Bathysquillidae, Squillidae, and Gonodactylidae. Stomatopoda are among the larger Crustacea; the size of an adult ranges from about 0.5 in. (12 mm) to over 12 in. (300 mm).

The bodies of stomatopods are narrow and elongate, almost eruciform. Only part of the cephalic and thoracic somites are fused and covered by the dorsal shield, the carapace. Those cephalic somites which bear the antennulae and the eyes are free and visible anterior to the carapace, and the last four thoracic somites are similarly exposed. The tail fan consists of a well-developed, sometimes peculiarly sculptured or deformed, median plate, the telson, and the two uropods. Anteriorly, the carapace bears a flattened movable plate, the rostrum. The eyes are large, stalked, and movable. Of eight pairs of thoracic appendages the first is narrow, slender, and hairy, probably being used for cleaning purposes. The second thoracic leg is very strong and heavy. It has become a large raptorial claw that shows a great resemblance to that of the praying mantis, and for this reason the Stomatopoda are often given the name mantis shrimps.

The Stomatopoda are marine animals rarely found in brackish water. Most species are confined to tropical and subtropical areas, though some occur in the boreal and antiboreal regions. The majority of stomatopods live in the littoral and sublittoral zones, but a few species have been found in greater depths, down to 2500 ft (760 m). *See* MALACOSTRACA. [L.B.Ho.]

Stone and stone products The term stone is applied to rock that is cut, shaped, broken, crushed, and otherwise physically modified for commercial use. The two main divisions are dimension stone and crushed stone. Other descriptive terms may be used, for example, building stone, roofing stone, or precious stone. *See* GEM; ROCK.

The term dimension stone is applied to blocks that are cut and milled to specified sizes, shapes, and surface finishes. The principal uses are for building and ornamental applications. Granites, limestones, sandstones, and marbles are widely used; basalts, diabases, and other dark igneous rocks are used less extensively. Soapstone is used to some extent. Rock suitable for use as dimension stone must be obtainable in large, sound blocks, free from incipient cracks, seams, and blemishes, and must be without mineral grains that might cause stains as a result of weathering. It must have an attractive color, and generally a uniform texture.

Slate differs from other dimension stone because it can be split into thin sheets of any desired thickness. Commercial slate must be uniform in quality and texture and reasonably free from knots, streaks, or other imperfections, and have good splitting properties. Roofing slates are important products of most slate quarries. However, the roofing-slate industry has declined considerably because of competition from other types of roofing. Slate is also used for milled products such as blackboards, electrical panels, window and door sills and caps, baseboards, stair treads, and floor tile. *See* SLATE.

Nearly all the principal types of stone—granite, diabase, basalt, limestone, dolomite, sandstone, and marble—may be used as sources of commercial crushed stone; limestone is by far the most important. Crushed stone is made from sound, hard stone, free from surface alteration by weathering. Stone that breaks in chunky, more or less cubical fragments is preferred. Commercial stone should be free from certain deleterious impurities, such as opalescent quartz, and free from clay or silt. Crushed stone is used principally as concrete aggregate, as road stone, or as railway ballast. Other uses for limestone are as a fluxing material to remove impurities from ores smelted in metallurgical furnaces, in the manufacture of alkali chemicals, calcium carbide, glass, paper, paint, and sugar, and for filter beds and for making mineral wool. [R.L.B.]

Storage tank A container for storing liquids or gases. A tank may be constructed of ferrous or nonferrous metals or alloys, reinforced concrete, wood, or filament-wound plastics, depending upon its use. Tanks resting on the ground have flat bottoms: those supported on towers have either flat or curved bottoms. Standpipes, which are usually cylindrical shells of steel or reinforced concrete resting on the ground, are frequently of great height and comparatively small diameter. They are built to contain water for a distribution system, and height is required to maintain pressure in the system. Tanks for other liquids and for gases, where storage is more important than pressure, are generally lower and of greater diameter. [C.M.A.]

Storm An atmospheric disturbance involving perturbations of the prevailing pressure and wind fields on scales ranging from tornadoes (0.6 mi or 1 km across) to extratropical cyclones (1.2–1900 mi or 2–3000 km across); also, the associated weather (rain storm, blizzard, and the like). Storms influence human activity in such matters as agriculture, transportation, building construction, water impoundment and flood control, and the generation, transmission, and consumption of electric energy. *See* WIND.

The form assumed by a storm depends on the nature of its environment, especially the large-scale flow patterns and the horizontal and vertical variation of temperature; thus the storms most characteristic of a given region vary according to latitude, physiographic features, and season. Extratropical cyclones and anticyclones are the chief disturbances over roughly half the Earth's surface. Their circulations control the embedded smaller-scale storms. Large-scale disturbances of the tropics differ fundamentally from those of extratropical latitudes. *See* HURRICANE; SQUALL; THUNDERSTORM; TORNADO; TROPICAL METEOROLOGY.

Cyclones form mainly in close proximity to the jet stream, that is, in strongly baroclinic regions where there is a large increase of wind with height. Weather patterns in cyclones are highly variable, depending on moisture content and thermodynamic stability of air masses drawn into their circulations. Warm and occluded fronts, east of and extending into the cyclone center, are regions of gradual upgliding motions, with widespread cloud and precipitation but usually no pronounced concentration of stormy conditions. Extensive cloudiness also is often present in the warm sector. Passage of the cold front is marked by a sudden wind shift, often with the onset of gusty conditions, with a pronounced tendency for clearing because of general subsidence behind the front. Showers may be present in the cold air if it is moist and unstable because of heating from the surface. Thunderstorms, with accompanying squalls and heavy rain, are often set off by sudden lifting of warm, moist air at or near the cold front, and these frequently move eastward into the warm sector. *See* CYCLONE; JET STREAM; WEATHER.

Extratropical cyclones alternate with high-pressure systems or anticyclones, whose circulation is generally opposite to that of the cyclone. The circulations of highs are not so intense as in well-developed cyclones, and winds are weak near their centers. In low levels the air spirals outward from a high; descent in upper levels results in warming and drying aloft. Anticyclones fall into two main categories, the warm "subtropical" and the cold "polar" highs.

Between the scales of ordinary air turbulence and of cyclones, there exist a variety of circulations over a middle-scale or mesoscale range, loosely defined as from about one-half up to a few hundred miles. Alternatively, these are sometimes referred to as subsynoptic-scale disturbances because their dimensions are so small that they elude adequate description by the ordinary synoptic network of surface weather stations. Thus their detection often depends upon observation by indirect sensing systems. *See* METEOROLOGICAL SATELLITES; RADAR METEOROLOGY; STORM DETECTION. [C.W.N.]

Storm detection Identifying storm formation, monitoring subsequent storm evolution, and assessing the potential for destruction of life and property through application of various methods and techniques. Doppler radars, satellite-borne instruments, lightning detection networks, and surface observing networks are used to detect the genesis of storms, to diagnose their nature, and to issue warnings when a threat to life and property exists. *See* STORM.

Radar surveillance. Radars emit pulses of electromagnetic radiation that are broadcast in a beam, whose angular resolution is about 1° with a range resolution of about 0.5 km (0.3 mi). The radar beam may intercept precipitation particles in a storm that reflect a fraction of the transmitted energy to the transmitter site (generally called reflectivity or the scatter cross section per unit volume). As the transmitter sweeps out a volume by rotating and tilting the transmitting antenna, the reflectivity pattern of the precipitation particles embodied in the storm is defined. Doppler radars also can measure the velocities of precipitation particles along the beam (radial velocity). Reflectivity and velocity patterns of the storm hydrometeors then make it possible to diagnose horizontal and vertical circulations that may arise within the storm, and to estimate the type and severity of weather elements attending the storm, such as rainfall, hail, damaging winds, and tornadoes. *See* DOPPLER RADAR; PRECIPITATION (METEOROLOGY); RADAR; RADAR METEOROLOGY; WIND.

Satellite surveillance. Since the early 1960s, meteorological data from satellites have had an increasing impact on storm detection and monitoring. In December 1966 the first geostationary *Applications Technology Satellite (ATS 1)* allowed forecasters in the United States to observe storms in animation. A Geostationary Operational Environmental Satellite (GOES) program was initiated within the National Oceanic and Atmospheric Administration (NOAA) with the launch of *GOES 1* in October 1975. The visible and infrared spin scan radiometer (VISSR) provided imagery, which significantly advanced the ability of meteorologists to detect and observe storms by providing frequent-interval visible and infrared imagery of the Earth surface, cloud cover, and atmospheric moisture patterns.

The first of NOAA's next generation of geostationary satellites, *GOES 8* was launched in the spring of 1994. *GOES 8* introduced improved capabilities to detect and observe storms. The *GOES 8* system includes no conflict between imaging and sounding operation, multispectral imaging with improved resolution and better signal-to-noise in the infrared bands, and more accurate temperature and moisture soundings of the storm environment. The Earth's atmosphere is observed nearly continuously.

Derived-product images showing fog and stratus areas from *GOES 8* are created by combining direct satellite measurements, such as by subtracting brightness temperatures at two different wavelengths. *GOES 8* shows the fog and stratus much more clearly because of its improved resolution. This capability enables forecasters to detect boundaries between rain-cooled areas having fog or low clouds, and clear areas. Such boundaries are frequently associated with future thunderstorm development. The sounder on *GOES 8* is capable of fully supporting routine forecasting operations. This advanced sounding capability consists of better vertical resolution in both temperature and moisture, and improved coverage of soundings in and around cloudy weather systems. *See* CLOUD; FOG; METEOROLOGICAL SATELLITES; SATELLITE METEOROLOGY.

Surface observing systems. Larger convective storm systems such as squall lines and mesoscale convective systems can be detected (but not fully described) by the temperature, moisture, wind, and pressure patterns observed by appropriate surface instrumentation. Automatic observing systems provide frequent data on pressure, temperature, humidity, wind, cloud base, and most precipitation types, intensity, and accumulation. Analyses of these data, combined with improved conceptual models of convective storm systems, enable forecasters

to detect and monitor the intense mesoscale fluctuations in pressure and winds that often accompany the passage of convective weather systems such as bow echoes, derechos (strong, straight-line winds), and squall lines. A bow echo is a specific radar reflectivity pattern associated with a line of thunderstorms. The middle portion of the thunderstorm line is observed to move faster than the adjacent portions, causing the line of storms to assume a bowed-out configuration. Other analyses of these mesoscale data fields aid the forecaster in detecting favorable areas for thunderstorm cell regeneration, which may produce slowly moving mesoscale convective storms attended by heavy rains and flash floods. *See* SQUALL LINE; WEATHER OBSERVATIONS.

Cloud-to-ground lightning detectors. Lightning location stations provide forecasters with the location, polarity, peak current, and number of strokes in a flash to ground within seconds of the flash occurrence. Useful applications have emerged with regard to the detection and tracking of thunderstorms, squall lines, other mesoscale convective systems, and the weather activity that accompany these phenomena, such as tornadoes and hail. *See* LIGHTNING; MESOMETEOROLOGY; SFERICS; THUNDERSTORM; WEATHER FORECASTING AND PREDICTION. [C.F.Ch.]

Storm electricity Processes responsible for the separation of positive and negative electric charges in the atmosphere during storms, including the spectacular manifestation of this charge separation: lightning discharges. Cloud electrification is almost invariably associated with convective activity and with the formation of precipitation in the form of liquid water (rain) and ice particles (graupel and hail). The most vigorous convection and active lightning occurs in the summertime, when the energy source for convection, water vapor, is most prevalent. Winter snowstorms can also be strongly electrified, but they produce far less lightning than summer storms. Electrified storm clouds occasionally occur in complete isolation; more commonly they are found in convective clusters or in lines that may extend horizontally for hundreds of kilometers. *See* PRECIPITATION (METEOROLOGY).

Measurements of electric field at the ground and from instrumented balloons within thunderclouds have disclosed an electrostatic structure that appears to be fairly systematic throughout the world. The measurements show that the principal variations in charge occur in the vertical and are affected by the temperature of the cloud. The charge structure within a thundercloud is tripolar, with a region of dominant negative charge sandwiched between an upper region of positive charge and a subsidiary lower region of positive charge. In addition to the charge accumulations described within the cloud, electrical measurements disclose the existence of charge-screening layers at the upper cloud boundary and a layer of positive charge near the Earth's surface beneath the cloud. These secondary charge accumulations arising from charge migration outside the cloud are caused by electrostatic forces of attraction set up by the charges within the cloud.

Large differences of electric potential are associated with the distribution of charge maintained by active thunderclouds. These large differences in potential are maintained by charging currents that result from the motions of air and particles. The charging currents range from milliamperes in small clouds that are not producing lightning to several amperes for large storms with high rates of lightning. *See* CLOUD PHYSICS.

In response to charge separation within a thundercloud, the electric field increases to a value of approximately 10^6 V/m (300,000 V/ft) at which point dielectric breakdown occurs and lightning is initiated. Most lightning extends through the cloud at speeds of 10,000–100,000 m/s (22,000–220,000 mi/h). The peak temperature of lightning, which is a highly ionized plasma, may exceed 30,000 K (54,000°F). The acoustic disturbance caused by the sudden heating of the atmosphere by lightning is thunder. *See* LIGHTNING; THUNDER.

Meteorologists have shown a growing interest in the large-scale display of real-time lightning activity, since lightning is one of the most sensitive indicators of convective activity. Research has expanded into relationships between lightning characteristics and the meteorological evolution of different types of storms. The discovery of the sensitive dependence of local lightning activity on the temperature of surface air has led to research focused on the use of the global electrical circuit as a diagnostic for global temperature change. *See* ATMOSPHERIC ELECTRICITY; STORM DETECTION; THUNDERSTORM. [E.Wi.]

Storm surge An anomalous rise in water elevations caused by severe storms approaching the coast. A storm surge can be succinctly described as a large wave that moves with the storm that caused it. The surge is intensified in the nearshore, shallower regions where the surface stress caused by the strong onshore winds pile up water against the coast, generating an opposing pressure head in the offshore direction. However, there are so many other forces at play in the dynamics of the storm surge phenomenon, such as bottom friction, Earth's rotation, inertia, and interaction with the coastal geometry, that a simple static model cannot explain all the complexities involved. Scientists and engineers have dedicated many years in the development and application of sophisticated computer models to accurately predict the effects of storm surges.

The intensity and dimension of the storm causing a surge, and thus the severity of the ensuing surge elevations, depend on the origin and atmospheric characteristics of the storm itself. Hurricanes and severe extratropical storms are the cause of most significant surges. In general, hurricanes are more frequent in low to middle latitudes, and extratropical storms are more frequent in middle to high latitudes. *See* HURRICANE; STORM. [S.R.S.]

Straight-line mechanism A mechanism that produces a straight-line (or nearly so) output motion from an input element that rotates, oscillates, or moves in a straight line. Common machine elements, such as linkages, gears, and cams, are often used in ingenious ways to produce the required controlled motion. The more elegant designs use the properties of special points on one of the links of a four-bar linkage. *See* MECHANISM.

Four-bar linkages that generate approximate straight lines are not new. In 1784 James Watt applied the concept to the vertical-cylinder beam engine. By selecting the appropriate link lengths, the designer can easily develop a mechanism with a high-quality approximate straight line. Contemporary kinematicians have contributed to more comprehensive studies of the properties of the mechanisms that generate approximate straight lines. The work not only describes the various classical mechanisms, but also provides design information on the quality (the amount of deviation from a straight line) and the length of the straight-line output. *See* FOUR-BAR LINKAGE; LINKAGE (MECHANISM).

Gears can also be used to generate straight-line motions. The most common combination would be a rack-and-pinion gear. *See* GEAR.

Cam mechanisms are generally not classified as straight-line motion generators, but translating followers easily fall into the classical definition. *See* CAM MECHANISM. [J.A.Sm.]

Strain gage A device which measures mechanical deformation (strain). Normally it is attached to a structural element, and uses the change of electrical resistance of a wire or semiconductor under tension. Capacity, inductance, and reluctance are also used.

The strain gage converts a small mechanical motion to an electrical signal by virtue of the fact that when a metal (wire or foil) or semiconductor is stretched, its resistance is increased. The change in resistance is a measure of the mechanical motion. In addition to their use in strain measurement, these gages are used in sensors for measuring the load on a mechanical member,

forces due to acceleration on a mass, or stress on a diaphragm or bellows. *See* STRAIN. [J.H.Z.]

Strake The slender forward extension of the inboard region of the wing of a combat aircraft used to provide increased lift in the high angle-of-attack maneuvering condition. In contrast to the normal attached-flow design principles, these strakes are built to allow the flow to separate along the leading edge in the high angle-of-attack range and to roll up into strong leading-edge vortices. The illustration shows a typical strake installation and the vortices generated. These are highly stable, and

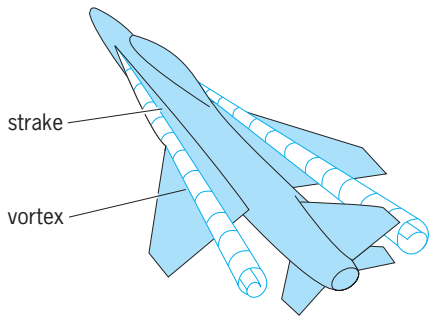

Strake and resulting vortex flow.

their strong swirling motion creates a lower pressure area on the strake upper surface, resulting in a large incremental increase in lift force known as vortex lift. This vortex lift increases rapidly in the high angle-of-attack range and is less susceptible to the normal stalling characteristics encountered with conventional lifting surfaces. *See* AIRFOIL; SUBSONIC FLIGHT; VORTEX; WING. [E.C.Po.]

Strand line The line that marks the separation of land and water along the margin of a pond, lake, sea, or ocean; also called the shoreline. The strand line is very dynamic. It changes with the tides, storms, and seasons, and as long-term sea-level changes take place. The sediments on the beach respond to these changes, as do the organisms that live in this dynamic environment. On a beach organisms move with the tides, and on a rocky coast they tend to be organized relative to the strand line because of special limitations or adaptations to exposure.

Geologists who study ancient coastal environments commonly try to establish where the strand line might be in the rock strata. This can sometimes be determined by a combination of the nature and geometry of individual laminations in the rock, by identifying sedimentary structures that occur at or near the strand line. The most indicative of these structures are swash marks, which are very thin accumulations of sand grains that mark the landward uprush of a wave on the beach. *See* FACIES (GEOLOGY); PALEOGEOGRAPHY; STRATIGRAPHY. [R.A.D.]

Strange particles Bound states of quarks, in which at least one of these constituents is of the strange (s) type. Strange quarks are heavier than the up (u) and down (d) quarks, which form the neutrons and protons in the atomic nucleus. Neutrons (udd) and protons (uud) are the lightest examples of a family of particles composed of three quarks, known as baryons. These and other composite particles which interact dominantly through the strong (nuclear) force are known as hadrons. The first strange hadron discovered (in cosmic rays in 1947) was named the lambda baryon, Λ; it is made of the three-quark combination uds. A baryon containing a strange quark is also called a hyperon. Although strange particles interact through the strong (nuclear) force, the strange quark itself can decay only by conversion to a quark of different type (such as u or d) through the weak interaction. For this reason, strange particles have very long lifetimes, of the order of 10^{-10} s, compared to the lifetimes of the order of 10^{-23} s for particles which decay directly through the strong interaction. This long lifetime was the origin of the term strange particles. *See* BARYON; HADRON; NEUTRON; PROTON; STRONG NUCLEAR INTERACTIONS.

In addition to strange baryons, strange mesons occur. The lightest of these are the kaons ($K^+ = u\bar{s}$ and $K^0 = d\bar{s}$) and the antikaons ($\bar{K}^0 = s\bar{d}$ and $\bar{K}^- = s\bar{u}$). Kaons and their antiparticles have been very important in the study of the weak interaction and in the detection of the very weak *CP* violation, which causes a slow transition between neutral kaons and neutral antikaons. *See* ELEMENTARY PARTICLE; MESON; QUARKS. [T.Bar.]

Strangles A highly contagious disease of the upper respiratory tract of horses and other members of the family Equidae, characterized by inflammation of the pharynx and abscess formation in lymph nodes. This disease occurs in horses of all ages throughout the world. The causative agent is *Streptococcus equi*, a clonal pathogen apparently derived from an ancestral strain of *S. zooepidemicus*. It is an obligate parasite of horses, donkeys, and mules. *See* STREPTOCOCCUS.

Strangles is most common and most severe in young horses, and is very prevalent on breeding farms. The causative agent has been reported to survive for 7 weeks in pus but dies in a week or two on pasture. Transmission is either direct by nose or mouth contact or aerosol, or indirect by flies, drinking buckets, pasture, and feed. The disease is highly contagious under conditions of crowding, exposure to severe climatic conditions such as rain and cold, and prolonged transportation. Carrier animals, although of rare occurrence, are critical in maintenance of the streptococcus and in initiation of outbreaks.

The mean incubation period is about 10 days, with a range of 3–14 days. The animal becomes quieter, has fever of 39–40.5°C (102–105°F), nasal discharge, loss of appetite, and swelling of one or more lymph nodes of the mouth. Pressure of a lymph node on the airway may cause respiratory difficulty. Abscesses in affected lymph nodes rupture in 7–14 days, and rapid clinical improvement and recovery then ensues. Recovery is associated with formation of protective antibodies in the nasopharynx and in the serum. *See* ANTIBODY.

Streptococcus equi is easily demonstrated in smears of pus from abscesses and in culture of pus or nasal swabs on colistin–nalidixic acid blood agar. Acutely affected animals also show elevated white blood cell counts and elevated fibrinogen.

Commercially available vaccines are injected in a schedule of two or three primary inoculations followed by annual boosters. However, the clinical attack rate may be reduced by only 50%, a level of protection much lower than that following the naturally occurring disease. *See* IMMUNITY.

Procaine penicillin G is the antibiotic of choice and quickly brings about reduction of fever and lymph node enlargement. *See* BIOLOGICALS. [J.F.T.]

Stratigraphy A discipline involving the description and interpretation of layered sediments and rocks, and especially their correlation and dating. Correlation is a procedure for determining the relative age of one deposit with respect to another. The term "dating" refers to any technique employed to obtain a numerical age, for example, by making use of the decay of radioactive isotopes found in some minerals in sedimentary rocks or, more commonly, in associated igneous rocks. To a large extent, layered rocks are ones that accumulated through sedimentary processes beneath the sea, within lakes, or by the action of rivers, the wind, or glaciers; but in places such deposits contain significant amounts of volcanic material emplaced as lava flows or as ash ejected from volcanoes during explosive eruptions. *See* DATING METHODS; IGNEOUS ROCKS; ROCK AGE DETERMINATION; SEDIMENTARY ROCKS.

Sedimentary successions are locally many thousands of meters thick owing to subsidence of the Earth's crust over millions of

years. Sedimentary basins therefore provide the best available record of Earth history over nearly 4 billion years. That record includes information about surficial processes and the varying environment at the Earth's surface, and about climate, changing sea level, the history of life, variations in ocean chemistry, and reversals of the Earth's magnetic field. Sediments also provide a record of crustal deformation (folding and faulting) and of large-scale horizontal motions of the Earth's lithospheric plates (continental drift). Stratigraphy applies not only to strata that have remained flat-lying and little altered since their time of deposition, but also to rocks that may have been strongly deformed or recrystallized (metamorphosed) at great depths within the Earth's crust, and subsequently exposed at the Earth's surface as a result of uplift and erosion. As long as original depositional layers can be identified, some form of stratigraphy can be undertaken. *See* BASIN; CONTINENTAL DRIFT; FAULT AND FAULT STRUCTURES; SEDIMENTOLOGY.

An important idea first articulated by the Danish naturalist Nicolaus Steno in 1669 is that in any succession of strata the oldest layer must have accumulated at the bottom, and successively younger layers above. It is not necessary to rely on the present orientation of layers to determine their relative ages because most sediments and sedimentary rocks contain numerous features, such as current-deposited ripples, minor erosion surfaces, or fossils of organisms in growth position, that have a well-defined polarity with respect to the up direction at the time of deposition (so-called geopetal indicators). This principle of superposition therefore applies equally well to tilted and even overturned strata. Only where a succession is cut by a fault is a simple interpretation of stratigraphic relations not necessarily possible, and in some cases older rocks may overlie younger rocks structurally. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS.

The very existence of layers with well-defined boundaries implies that the sedimentary record is fundamentally discontinuous. Discontinuities are present in the stratigraphic record at a broad range of scales, from that of a single layer or bed to physical surfaces that can be traced laterally for many hundreds of kilometers. Large-scale surfaces of erosion or nondeposition are known as unconformities, and they can be identified on the basis of both physical and paleontological criteria. *See* PALEONTOLOGY; UNCONFORMITY.

Most stratal discontinuities possess time-stratigraphic significance because strata below a discontinuity tend to be everywhere older than strata above. To the extent that unconformities can be recognized and traced widely within a sedimentary basin, it is possible to analyze sedimentary rocks in a genetic framework, that is, with reference to the way they accumulated. This is the basis for the modern discipline of sequence stratigraphy, so named because intervals bounded by unconformities have come to be called sequences.

Traditional stratigraphic analysis has focused on variations in the intrinsic character or properties of sediments and rocks—properties such as composition, texture, and included fossils (lithostratigraphy and biostratigraphy)—and on the lateral tracing of distinctive marker beds such as those composed of ash from a single volcanic eruption (tephrostratigraphy). The techniques of magnetostratigraphy and chemostratigraphy are also based on intrinsic characteristics, although these techniques require sophisticated laboratory analysis. Sequence stratigraphy attempts to integrate these approaches in the context of stratal geometry, thereby providing a unifying framework in which to investigate the time relations between sediment and rock bodies as well as to measure their numerical ages (chronostratigraphy and geochronology). Seismic stratigraphy is a variant of the technique of sequence stratigraphy in which unconformities are identified and traced in seismic reflection profiles on the basis of reflection geometry. *See* GEOCHRONOMETRY; SEISMIC STRATIGRAPHY; SEISMOLOGY.

Conventional stratigraphy currently recognizes two kinds of stratigraphic unit: material units, distinguished on the basis of some specified property or properties or physical limits; and temporal or time-related units. A common example of a material unit is the formation, a lithostratigraphic unit defined on the basis of lithic characteristics and position within a stratigraphic succession. Each formation is referred to a section or locality where it is well developed (a type section), and assigned an appropriate geographic name combined with the word formation or a descriptive lithic term such as limestone, sandstone, or shale (for example, Tapeats Sandstone). Some formations are divisible into two or more smaller-scale units called members and beds. In other cases, formations of similar lithic character or related genesis are combined into composite units called groups and supergroups.

Sequence stratigraphy differs from conventional stratigraphy in two important respects. The first is that basic units (sequences) are defined on the basis of bounding unconformities and correlative conformities rather than material characteristics or age. The second is that sequence stratigraphy is fundamentally not a system for stratigraphic classification, but a procedure for determining how sediments accumulate. *See* SEQUENCE STRATIGRAPHY.

[N.C.B.]

Stratosphere The atmospheric layer that is immediately above the troposphere and contains most of the Earth's ozone. Here temperature increases upward because of absorption of solar ultraviolet light by ozone. Since ozone is created in sunlight from oxygen, a by-product of photosynthesis, the stratosphere exists because of life on Earth. In turn, the ozone layer allows life to thrive by absorbing harmful solar ultraviolet radiation. The mixing ratio of ozone is largest (10 parts per million by volume) near an altitude of 30 km (18 mi) over the Equator. The distribution of ozone is controlled by solar radiation, temperature, wind, reactive trace chemicals, and volcanic aerosols. *See* ATMOSPHERE; TROPOSPHERE.

The heating that results from absorption of ultraviolet radiation by ozone causes temperatures generally to increase from the bottom of the stratosphere (tropopause) to the top (stratopause) near 50 km (30 mi), reaching 280 K (45°F) over the summer pole. This temperature inversion limits vertical mixing, so that air typically spends months to years in the stratosphere. *See* TEMPERATURE INVERSION; TROPOPAUSE.

The lower stratosphere contains a layer of small liquid droplets. Typically less than 1 micrometer in diameter, they are made primarily of sulfuric acid and water. Occasional large volcanic eruptions maintain this aerosol layer by injecting sulfur dioxide into the stratosphere, which is converted to sulfuric acid and incorporated into droplets. Enhanced aerosol amounts from an eruption can last several years. By reflecting sunlight, the aerosol layer can alter the climate at the Earth's surface. By absorbing upwelling infrared radiation from the Earth's surface, the aerosol layer can warm the stratosphere. The aerosols also provide surfaces for a special set of chemical reactions that affect the ozone layer. Liquid droplets and frozen particles generally convert chlorine-bearing compounds to forms that can destroy ozone. They also tend to take up nitric acid and water and to fall slowly, thereby removing nitrogen and water from the stratosphere. The eruption of Mount Pinatubo (Philippines) in June 1991 is believed to have disturbed the Earth system for several years, raising stratospheric temperatures by more than 1 K (1.8°F) and reducing global surface temperatures by about 0.5 K (0.9°F). *See* AEROSOL.

Ozone production is balanced by losses due to reactions with chemicals in the nitrogen, chlorine, hydrogen, and bromine families. Reaction rates are governed by temperature, which depends on amounts of radiatively important species such as carbon dioxide. Human activities are increasing the amounts of these molecules and are thereby affecting the ozone layer. Evidence for anthropogenic ozone loss has been found in the Antarctic lower stratosphere. Near polar stratospheric clouds, chlorine and bromine compounds are converted to species that, when the Sun comes up in the southern spring, are broken apart by ultraviolet

radiation and rapidly destroy ozone. This sudden loss of ozone is known as the anthropogenic Antarctic ozone hole. *See* STRATOSPHERIC OZONE. [M.H.H.]

Stratospheric ozone While ozone is found in trace quantities throughout the atmosphere, the largest concentrations are located in the lower stratosphere in a layer between 9 and 18 mi (15 and 30 km). Atmospheric ozone plays a critical role for the biosphere by absorbing the ultraviolet radiation with wavelength (λ) 240–320 nanometers. This radiation is lethal to simple unicellular organisms (algae, bacteria, protozoa) and to the surface cells of higher plants and animals. It also damages the genetic material of cells and is responsible for sunburn in human skin. The incidence of skin cancer has been statistically correlated with the observed surface intensity of ultraviolet wavelength 290–320 nm, which is not totally absorbed by the ozone layer. *See* OZONE; STRATOSPHERE; ULTRAVIOLET RADIATION (BIOLOGY).

Ozone also plays an important role in photochemical smog and in the purging of trace species from the lower atmosphere. Furthermore, it heats the upper atmosphere by absorbing solar ultraviolet and visible radiation ($\lambda < 710$ nm) and thermal infrared radiation ($\lambda \simeq 9.6$ micrometers). As a consequence, the temperature increases steadily from about -60°F (220 K) at the tropopause (5–10 mi or 8–16 km altitude) to about 45°F (280 K) at the stratopause (30 mi or 50 km altitude). This ozone heating provides the major energy source for driving the circulation of the upper stratosphere and mesosphere. *See* ATMOSPHERIC GENERAL CIRCULATION; TROPOPAUSE.

Above about 19 mi (30 km), molecular oxygen (O_2) is dissociated to free oxygen atoms (O) during the daytime by ultraviolet photons, ($h\nu$), as shown in reaction (1). The oxygen atoms produced then form ozone (O_3) by reaction (2), where M is an

$$O_2 + h\nu \rightarrow O + O \qquad \lambda < 242 \text{ nm} \tag{1}$$

$$O + O_2 + M \rightarrow O_3 + M \tag{2}$$

arbitrary molecule required to conserve energy and momentum in the reaction. Ozone has a short lifetime during the day because of photodissociation, as shown in reaction (3). However,

$$O_3 + h\nu \rightarrow O_2 + O \qquad \lambda < 710 \text{ nm} \tag{3}$$

except above 54 mi (90 km), where O_2 begins to become a minor component of the atmosphere, reaction (3) does not lead to a net destruction of ozone. Instead the O is almost exclusively converted back to O_3 by reaction (2). If the odd oxygen concentration is defined as the sum of the O_3 and O concentrations, then odd oxygen is produced by reaction (1). It can be seen that reactions (2) and (3) do not affect the odd oxygen concentrations but merely define the ratio of O to O_3. Because the rate of reaction (2) decreases with altitude while that for reaction (3) increases, most of the odd oxygen below 36 mi (60 km) is in the form of O_3 while above 36 mi (60 km) it is in the form of O. The reaction that is responsible for a small fraction of the odd oxygen removal rate is shown as reaction (4). A significant fraction

$$O + O_3 \rightarrow O_2 + O_2 \tag{4}$$

of the removal is caused by the presence of chemical radicals [such as nitric oxide (NO), chlorine (Cl), bromine (Br), hydrogen (H), or hydroxyl (OH)], which serve to catalyze reaction (4) (see illustration).

The discovery in the mid-1980s of an ozone hole over Antarctica, which could not be explained by the classic theory of ozone and had not been predicted by earlier chemical models, led to many speculations concerning the causes of this event, which can be observed each year in September and October. As suggested by experimental and observational evidence, heterogeneous reactions on the surface of liquid or solid particles that

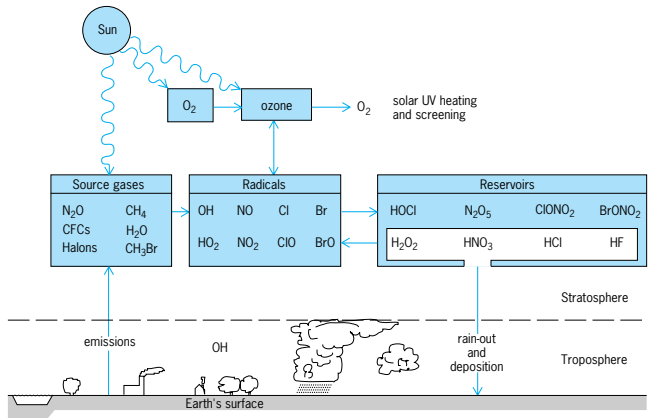

Principal chemical cycles in the stratosphere. The destruction of ozone is affected by the presence of radicals which are produced by photolysis or oxidation of source gases. Chemical reservoirs are relatively stable but are removed from the stratosphere by transport toward the troposphere and rain-out.

produce Cl_2, HOCl, and $ClNO_2$ gas, and the subsequent rapid photolysis of these molecules, produces chlorine radicals (Cl, ClO) which in turn lead to the destruction of ozone in the lower stratosphere by a catalytic cycle [reactions (5)–(7)]. Solar radiation is needed for these processes to occur.

$$Cl + O_3 \rightarrow ClO + O_2 \tag{5}$$

$$ClO + ClO \rightarrow Cl_2O_2 \tag{6}$$

$$Cl_2O_2 + h\nu \rightarrow 2Cl + O_2 \tag{7}$$

Sites on which the reactions producing Cl_2, HOCl, and $ClNO_2$ can occur are provided by the surface of ice crystals in polar stratospheric clouds (PSCs). These clouds are formed between 8 and 14 mi (12 and 22 km) when the temperature drops below approximately -123°F (187 K). Other types of particles are observed at temperatures above the frost point of -123°F (187 K). These particles provide additional surface area for these reactions to occur. Clouds are observed at high latitudes in winter. Because the winter temperatures are typically 20–30°F (10–15 K) colder in the Antarctic than in the Arctic, their frequency of occurrence is highest in the Southern Hemisphere. Thus, the formation of the springtime ozone hole over Antarctica is explained by the activation of chlorine and the catalytic destruction of O_3 which takes place during September, when the polar regions are sunlit but the air is still cold and isolated from midlatitude air by a strong polar vortex. Satellite observations made since the 1970s suggest that total ozone in the Arctic has been abnormally low during the 1990s, probably in relation to the exceptionally cold winter tempratures in the Arctic lower stratosphere recorded during that decade. [G.P.B.; R.G.Pr.]

Strawberry Low-growing perennials, spreading by stolons, with fruit consisting of a fleshy receptacle, and "seeds" in pits or nearly superficial on the receptacle. The strawberry in the United States is derived from two species: *Fragaria chihensis*, which grows along the Pacific Coast of North and South America, and *F. uirginiana*, the eastern meadow strawberry, both members of the order Resales. *See* ROSALES.

The strawberry is the most universally grown of the small fruits, both in the home garden and in commercial plantings. Home garden production is possible in nearly all of the states, provided water can be supplied where rainfall is insufficient. Commercial production is important in probably three-fourths of the states. The following states are large producers: Oregon, California, Tennessee, Michigan, Louisiana, Washington, Arkansas, Kentucky, and New York. *See* FRUIT. [J.H.C.]

Stream function In fluid mechanics, a mathematical idea which satisfies identically, and therefore eliminates completely, the equation of mass conservation. If the flow field consists of only two space coordinates, for example, x and y, a single and very useful stream function $\psi(x, y)$ will arise. If there are three space coordinates, such as (x, y, z), multiple stream functions are needed, and the idea becomes much less useful and is much less widely employed.

The stream function not only is mathematically useful but also has a vivid physical meaning. Lines of constant ψ are streamlines of the flow; that is, they are everywhere parallel to the local velocity vector. No flow can exist normal to a streamline; thus, selected ψ lines can be interpreted as solid boundaries of the flow.

Further, ψ is also quantitatively useful. In plane flow, for any two points in the flow field, the difference in their stream function values represents the volume flow between the points. *See* CREEPING FLOW; FLUID FLOW. [F.M.W.]

Stream gaging The measurement of water discharge in streams. Discharge is the rate of movement of the stream's water volume. It is the product of water velocity times cross-sectional area of the stream channel. Several techniques have been developed for measuring stream discharge; selection of the gaging method usually depends on the size of the stream. The most accurate methods for measuring stream discharge make use of in-stream structures through which the water can be routed, such as flumes and weirs.

A flume is a constructed channel that constricts the flow through a control section, the exact dimensions of which are known. Through careful hydraulic design and calibration by laboratory experiments, stream discharge through a flume can be determined by simply measuring the water depth (stage) in the inlet or constricted sections. Appropriate formulas relate stage to discharge for the type of flume used.

A weir is used in conjunction with a dam in the streambed. The weir itself is usually a steel plate attached to the dam that has a triangular, rectangular, or trapezoidal notch over which the water flows. Hydraulic design and experimentation has led to calibration curves and appropriate formulas for many different weir designs. To calculate stream discharge through a weir, only the water stage in the reservoir created by the dam needs to be measured. Stream discharge can be calculated by using the appropriate formula that relates stage to discharge for the type of weir used. *See* HYDROLOGY; SURFACE WATER. [T.C.Wi.]

Stream pollution Biological, or bacteriological, pollution in a stream indicated by the presence of the coliform group of organisms. While nonpathogenic itself, this group is a measure of the potential presence of contaminating organisms. Because of temperature, food supply, and predators, the environment provided by natural bodies of water is not favorable to the growth of pathogenic and coliform organisms. Physical factors, such as flocculation and sedimentation, also help remove bacteria. Any combination of these factors provides the basis for the biological self-purification capacity of natural water bodies.

Nonpolluted natural waters are usually saturated with dissolved oxygen. They may even be supersaturated because of the oxygen released by green water plants under the influence of sunlight. When an organic waste is discharged into a stream, the dissolved oxygen is utilized by the bacteria in their metabolic processes to oxidize the organic matter. The oxygen is replaced by reaeration through the water surface exposed to the atmosphere. This replenishment permits the bacteria to continue the oxidative process in an aerobic environment. In this state, reasonably clean appearance, freedom from odors, and normal animal and plant life are maintained.

An increase in the concentration of organic matter stimulates the growth of bacteria and increases the rates of oxidation and oxygen utilization. If the concentration of the organic pollutant is so great that the bacteria use oxygen more rapidly than it can be replaced, only anaerobic bacteria can survive and the stabilization of organic matter is accomplished in the absence of oxygen. Under these conditions, the water becomes unsightly and malodorous, and the normal flora and fauna are destroyed. Furthermore, anaerobic decomposition proceeds at a slower rate than aerobic. For maintenance of satisfactory conditions, minimal dissolved oxygen concentrations in receiving streams are of primary importance.

Municipal sewage and industrial wastes affect the oxygen content of a stream. Cooling water, used in some industrial processes, is characterized by high temperatures, which reduce the capacity of water to hold oxygen in solution. Municipal sewage requires oxygen for its stabilization by bacteria. Oxygen is utilized more rapidly than it is replaced by reaeration, resulting in the death of the normal aquatic life. Further downstream, as the oxygen demands are satisfied, reaeration replenishes the oxygen supply.

Polluted waters are deprived of oxygen by the exertion of the biochemical oxygen demand, which is defined as the quantity of oxygen required by the bacteria to oxidize the organic matter. Factors such as the turbulence of the stream flow, biological growths on the stream bed, insufficient nutrients, and inadequate bacteria in the river water influence the rate of oxidation in the stream as well as the removal of organic matter.

When a significant portion of the waste is in the suspended state, settling of the solids in a slow-moving stream is probable. The organic fraction of the sludge deposits decomposes anaerobically, except for the thin surface layer which is subjected to aerobic decomposition due to the dissolved oxygen in the overlying waters. In warm weather, when the anaerobic decomposition proceeds at a more rapid rate, gaseous end products, usually carbon dioxide and methane, rise through the supernatant waters. The evolution of the gas bubbles may raise sludge particles to the water surface. Although this phenomenon may occur while the water contains some dissolved oxygen, the more intense action during the summer usually results in depletion of dissolved oxygen.

Water may absorb oxygen from the atmosphere when the oxygen in solution falls below saturation. Dissolved oxygen for receiving waters is also derived from two other sources: that in the receiving water and the waste flow at the point of discharge, and that given off by green plants.

Unpolluted water maintains in solution the maximum quantity of dissolved oxygen. The saturation value is a function of temperature and the concentration of dissolved substances, such as chlorides. When oxygen is removed from solution, the deficiency is made up by the atmospheric oxygen, which is absorbed at the water surface and passes into solution. The oxygen balance in a stream is determined by the concentration of organic matter and its rate of oxidation, and by the dissolved oxygen concentration and the rate of reaeration. *See* ESTUARINE OCEANOGRAPHY; FRESHWATER ECOSYSTEM; SEWAGE DISPOSAL; WATER POLLUTION. [D.J.O.]

Stream transport and deposition The sediment debris load of streams is a natural corollary to the degradation of the landscape by weathering and erosion. Eroded material reaches stream channels through rills and minor tributaries, being carried by the transporting power of running water and by mass movement, that is, by slippage, slides, or creep. The size represented may vary from clay to boulders. At any place in the stream system the material furnished from places upstream either is carried away or, if there is insufficient transporting ability, is accumulated as a depositional feature. The accumulation of deposited debris tends toward increased ease of movement, and this tends eventually to bring into balance the transporting ability of the stream and the debris load to be transported. [L.B.L.]

Streaming potential The potential which is produced when a liquid is forced to flow through a capillary or a porous

solid. The streaming potential is one of four related electrokinetic phenomena which depend upon the presence of an electrical double layer at a solid-liquid interface. This electrical double layer is made up of ions of one charge type which are fixed to the surface of the solid and an equal number of mobile ions of the opposite charge which are distributed through the neighboring region of the liquid phase. In such a system the movement of liquid over the surface of the solid produces an electric current, because the flow of liquid causes a displacement of the mobile counterions with respect to the fixed charges on the solid surface. The applied potential necessary to reduce the net flow of electricity to zero is the streaming potential. [Q.V.W.]

Streamlining The contouring of a body to reduce its resistance (drag) to motion through a fluid.

For fluids with relatively low viscosity such as water and air, effects of viscous friction are confined to a thin layer of fluid on the surface termed the boundary layer. Under the influence of an increasing pressure, the flow within the boundary layer tends to reverse and flow in an upstream direction. Viscosity tends to cause the flow to separate from the body surface with consequent formation of a region of swirling or eddy flow (termed the body wake; illus. *a*). This eddy formation leads to a reduction in the downstream pressure on the body and hence gives rise to a force opposite to the body motion, known as pressure drag. *See* WAKE FLOW.

In general, streamlining in subsonic flow involves the contouring of the body in such a manner that the wake is reduced and hence the pressure drag is reduced. The contouring should provide for gradual deceleration to avoid flow separation, that is, reduced adverse pressure gradients. These considerations lead to the following general rules for subsonic streamlining: The forward portion of the body should be well rounded, and the body should curve back gradually from the forward section to a tapering aftersection with the avoidance of sharp corners along the body surface. These conditions are well illustrated by teardrop shapes (illus. *b*).

At supersonic speeds the airflow can accommodate sudden changes in direction by being compressed or expanded. Where this change in direction occurs at the nose of the body, a compression wave is created, the strength of which depends upon the magnitude of the change in flow direction. Lowering the body-induced flow angle weakens this compression shock wave. When the flow changes direction again at the midpoint of the body, the air will expand to follow the shape of the body. This change in direction creates expansion waves. At the tail of the body the direction changes again, creating another compression or shock wave. At each of these shock waves, changes in pressure, density, and velocity occur, and in this process energy is lost. This loss results in a retarding force known as wave drag. *See* SHOCK WAVE.

Bodies which are streamlined for supersonic speeds are characterized by a sharp nose and small flow deflection angles. Because the intensity of the shock wave and the drag level is dependent upon the magnitude of the change in flow direction, the width or thickness of the body should be minimal. *See* BOUNDARY-LAYER FLOW. [A.G.H.; D.M.Bu.]

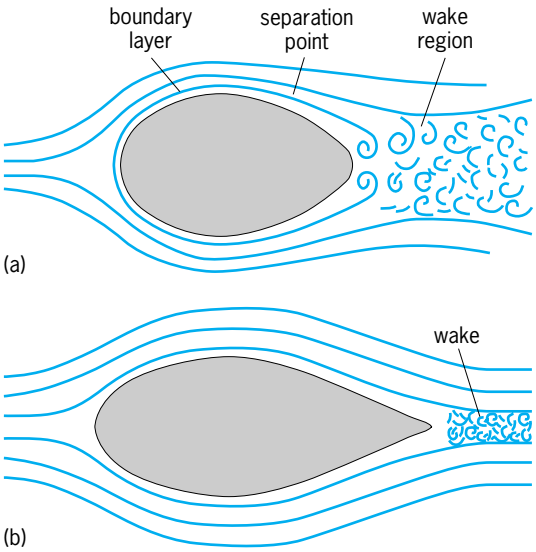

Flow about bodies in uniform subsonic flow. (*a*) Blunt body. (*b*) Streamlined body.

Strength of materials A branch of applied mechanics concerned with the behavior of materials under load, relationships between externally applied loads and internal resisting forces, and associated deformations. Knowledge of the properties of materials and analysis of the forces involved are fundamental to the investigation and design of structures and machine elements. *See* MACHINE DESIGN; STRUCTURAL MATERIALS.

Investigation of the resistance of a member, dealing with internal forces, is called free-body analysis. Determination of the distribution and intensity of the internal forces and the associated deformations is called stress analysis. *See* STRESS AND STRAIN.

A material offers resistance to external load only insofar as the component elements can furnish cohesive strength, resistance to compaction, and resistance to sliding. The relations developed in strength of materials analysis evaluate the tensile, compressive, and shear stresses that a material is called upon to resist. The most important factors in determining the suitability of a structural or machine element for a particular application are strength and stiffness. *See* SHEAR. [J.B.S.]

Strepsiptera An order of twisted-wing insects that spend most of their life cycle as internal parasitoids of other insects. The adult male is only a few millimeters in wing span, is free living, and has only one pair of full flight wings (the posterior, or metathoracic, pair). The front (mesothoracic) wings are reduced to narrow, clublike organs that may function as halteres, or flight balancers. The male eyes are coarsely faceted and berrylike, and the antennae have four to seven segments, with some segments having finger- or bladelike extensions. Most adult females are immobile, blind, and larviform, and live inside the insect host. Rarely, the female is free living, with legs and eyes but no wings. More than 600 species of Strepsiptera are known, many of them not yet formally described and named. The order's relationship to other insect orders remains uncertain.

Both male and female begin larval life as mobile first-stage organisms with eyes and three pairs of functional legs. These triunguloids eventually attack the immature or adult forms of the host and enter their bodies, where they molt to an apodous instar. Two subsequent molts result in the divergence of the sexes and the differentiation of a hardened forebody that eventually protrudes through the host integument; the molted larval integuments are not shed but are incorporated in the general puparial wall arid the neotenous adult female capsule. Males (and some adult females) emerge from the puparium and begin the search for a mate, probably through the mediation of airborne pheromones.

Strepsiptera are found worldwide in temperate and tropical habitats. They parasitize insects, mostly of orders Hemiptera and Hymenoptera, but also some cockroaches, mantids, orthopterans, flies, and silverfish. Their effect on the host is variable, ranging from reproductive failure to death. *See* ENDOPTERYGOTA; PHEROMONE. [W.L.Bro.]

Streptococcus A large genus of spherical or ovoid bacteria that are characteristically arranged in pairs or in chains resembling strings of beads. Many of the streptococci that constitute part of the normal flora of the mouth, throat, intestine,

and skin are harmless commensal forms; other streptococci are highly pathogenic. The cells are gram-positive and can grow either anaerobically or aerobically, although they cannot utilize oxygen for metabolic reactions. Glucose and other carbohydrates serve as sources of carbon and energy for growth. All members of the genus lack the enzyme catalase. Streptococci can be isolated from humans and other animals.

Streptococcus pyogenes is well known for its participation in many serious infections. It is a common cause of throat infection, which may be followed by more serious complications such as rheumatic fever, glomerulonephritis, and scarlet fever. Other beta-hemolytic streptococci participate in similar types of infection, but they are usually not associated with rheumatic fever and glomerulonephritis. Group B streptococci, which are usually beta-hemolytic, cause serious infections in newborns (such as meningitis) as well as in adults. Among the alpha-hemolytic and nonhemolytic streptococci, *S. pneumoniae* is an important cause of pneumonia and other respiratory infections. Vaccines that protect against infection by the most prevalent capsular serotypes are available. The viridans streptococci comprise a number of species commonly isolated from the mouth and throat. Although normally of low virulence, these streptococci are capable of causing serious infections (endocarditis, abcesses). *See* PNEUMONIA. [K.Ru.]

Streptothricosis An acute or chronic infection of the epidermis, caused by the bacterium *Dermatophilus congolensis*. which results in an oozing dermatitis with scab formation. Streptothricosis includes dermatophilosis, mycotic dermatitis, lumpy wool, strawberry foot-rot, and cutaneous streptothricosis—diseases having a worldwide distribution and affecting a wide variety of species, including humans.

The infectious form of the bacterium is a coccoid, motile zoospore that is released when the skin becomes wet. Thus, the disease is closely associated with rainy seasons and wet summers. Zoospores lodge on the skin of susceptible animals and germinate by producing filaments which penetrate to the living epidermis, where the organism proliferates by branching mycelial growth.

Early cutaneous lesion of dermatophilosis in cattle reveals small vesicles, papules, and pus formation under hair plaques. An oozing dermatitis then appears as the disease progresses and exudates coalesce to form scabs, which change to hard crusts firmly adherent to the skin. The crusts enlarge and harden, and are often devoid of hair. Lesions occur on most areas of sheep, but the characteristic lesions in the wooled areas occur as numerous hard masses of crust or scab scattered irregularly over the back, flanks, and upper surface of the neck. Lesion resolution has been found to correlate with the presence of immunoglobulin A–containing plasma cells in the dermis and with the antibody levels to *D. congolensis* at the skin surface of infected sheep and cattle.

No single treatment is considered specific for dermatophilosis. Some observers claim that topical agents are successful for sheep and cattle. A number of systemic antibiotics are effective in treating the disease. A combination of streptomycin and penicillin has given good therapeutic results in both bovine and equine infections. [S.S.S.]

Stress (psychology) Generally, environmental events of a challenging sort as well as the body's response to such events. Of particular interest has been the relationship between stress and the body's adaptation to it on the one hand and the body's susceptibility to disease on the other. Both outcomes involve behavioral and brain changes as well as psychosomatic events, that is, changes in body function arising from the ability of the brain to control such function through neural output as well as hormones. One problem is that both environmental events and bodily responses have been referred to interchangeably as stress. It is preferable to refer to the former as the stressor and the latter as the stress response. The stress response consists of a cascade of neural and hormonal events that have short- and long-lasting consequences for brain and body alike. A more serious issue is how an event is determined to be a stressor. One view is to define a stressor as an environmental event causing a negative outcome, such as a disease. Another approach is to view stressors as virtually any challenge to homeostasis and to regard disease processes as a failure of the normal operation of adapative mechanisms, which are part of the stress response. With either view, it is necessary to include psychological stressors, such as fear, that contain implied threats to homeostasis and that evoke psychosomatic reactions. These are reactions that involve changes in neural and hormonal output caused by psychological stress. Psychosomatic reactions may lead to adaptive responses, or they may exacerbate disease processes. Whether the emphasis is on adaptation or disease, it is essential to understand the processes in the brain that are activated by stressors and that influence functions in the body. *See* HOMEOSTASIS; PSYCHOSOMATIC DISORDERS.

Among the many neurotransmitter systems activated by stress is noradrenaline, produced by neurons with cell bodies in the brainstem that have vast projections up to the forebrain and down the spinal cord. Stressful experiences activate the noradrenergic system and promote release of noradrenaline; severe stress leads to depletion of noradrenaline in brain areas such as the hypothalamus. This release and depletion of noradrenaline stores results in changes at two levels of neuronal function: phosphorylation is triggered by the second-messenger cyclic AMP and occurs in the presynaptic and postsynaptic sites where noradrenaline is released and where it also acts; synthesis of new protein is induced via actions on the genome. Both processes enhance the ability of the brain to form noradrenaline when the organism is once again confronted with a stressful situation. Other neurotransmitter systems may also show similar adaptive changes in response to stressors. *See* NORADRENERGIC SYSTEM.

Stress also activates the neurally mediated discharge of adrenaline from the adrenal medulla and of hypothalamic hormones that initiate the neuroendocrine cascade, culminating in glucocorticoid release from the adrenal cortex. Thus, the activity of neurons triggered by stressful experiences, physical trauma, fear, or anger leads to hormone secretion that has effects throughout the body. Virtually every organ of the body is affected by stress hormones. The hypothalamic hormone (corticotrophin-releasing hormone) that triggers the neuroendocrine cascade directly stimulates the pituitary to secrete ACTH. In response to certain stressors, the hypothalamus also secretes vasopressin and oxytocin, which act synergistically with corticotrophin-releasing hormone on the pituitary to potentiate the secretion of ACTH. Various stressors differ in their ability to promote output of vasopressin and oxytocin, but all stressors stimulate release of corticotrophin-releasing hormone. Other hormones involved in the stress response include prolactin and thyroid hormone; the metabolic hormones insulin, epinephrine, and glucagon; and the endogenous opiates endorphin and enkephalin. *See* ENDORPHINS.

Of all the hormones in the endocrine cascade initiated by stress, the glucocorticoids are the most important because of their widespread effects throughout the body and in the brain. The brain contains target cells for adrenal glucocorticoids secreted in stress, and receptors in these cells are proteins that interact with the genome to affect expression of genetic information. Thus, the impact of stress-induced activation of the endocrine cascade that culminates in glucocorticoid release is the feedback of glucocorticoids on target brain cells. The effect is to alter the structure and function of the brain cells over a period of time ranging from hours to days.

In the case of noradrenaline, glucocorticoids have several types of feedback effects that modify how the noradrenergic system responds to stress. Glucocorticoids inhibit noradrenaline release, and they reduce the second-messenger response of brain

structures such as the cerebral cortex to noradrenaline. Glucocorticoid feedback also affects the serotonin system, facilitating serotonin formation during stress but at the same time altering the levels of several types of serotonin receptors in different brain regions, which has the net effect of shifting the balance within the serotonergic system. Taken together, evidence points to a role of glucocorticoid secretion in leading to restoration of homeostatic balance by counteracting the acute neural events such as increased activity of noradrenaline and serotonin, which are turned on by stressful experiences. Other neurotransmitter systems may also respond to glucocorticoid action. Moreover, the other hormones activated by stress have effects on the brain and body that must be considered. *See* SEROTONIN.

In general, stress hormones are protective and adaptive in the immediate aftermath of stress, and the organism is more vulnerable to many conditions without them. However, the same hormones can promote damage and accelerate pathophysiological changes, such as bone mineral loss, obesity, and cognitive impairment, when they are overproduced or not turned off. This wear-and-tear on the body has been called allostatic load. It is based upon the notion that allostasis is the active process of maintaining stability, or homeostasis, through change, and allostatic load is the almost inevitable cost to the body of doing so.

Stress hormone actions have important effects outside the brain on such systems as the immune response. Glucocorticoids and catecholamines from sympathetic nerves and the adrenal medulla participate in the mobilization and enhancement of immune function in the aftermath of acute stress. These effects improve the body's defense against pathogens but can exacerbate autoimmune reactions. When they are secreted chronically, the stress-related hormones are generally immunosuppressive; such effects can be beneficial in the case of an autoimmune disease but may compromise defense against a virus or bacterial infections. At the same time, glucocorticoids are important agents for containing the acute-phase response to an infection or autoimmune disturbance. In the absence of such containment, the organism may die because of the excessive inflammatory response. *See* IMMUNOLOGY.

Besides affecting the immune response, stressors are believed to exacerbate endogenous depressive illness in susceptible individuals. Major depressive illness frequently results in elevated levels of cortisol in the blood. It is not clear whether the elevated cortisol is a cause or strictly a result of the brain changes involved in depressive illness. *See* AFFECTIVE DISORDERS. [B.McE.]

Stress and strain Related terms defining the intensity of internal reactive forces in a deformed body and associated unit changes of dimension, shape, or volume caused by externally applied forces. Stress is a measure of the internal reaction between elementary particles of a material in resisting separation, compaction, or sliding that tend to be induced by external forces. Total internal resisting forces are resultants of continuously distributed normal and parallel forces that are of varying magnitude and direction and are acting on elementary areas throughout the material. These forces may be distributed uniformly or nonuniformly. Stresses are identified as tensile, compressive, or shearing, according to the straining action.

Strain is a measure of deformation such as (1) linear strain, the change of length per unit of linear dimensions; (2) shear strain, the angular rotation in radians of an element undergoing change of shape by shearing forces; or (3) volumetric strain, the change of volume per unit of volume. The strains associated with stress are characteristic of the material. Strains completely recoverable on removal of stress are called elastic strains. Above a critical stress, both elastic and plastic strains exist, and that part remaining after unloading represents plastic deformation called inelastic strain. Inelastic strain reflects internal changes in the crystalline structure of the metal. Increase of resistance to continued plastic deformation due to more favorable rearrangement of the atomic structure is strain hardening.

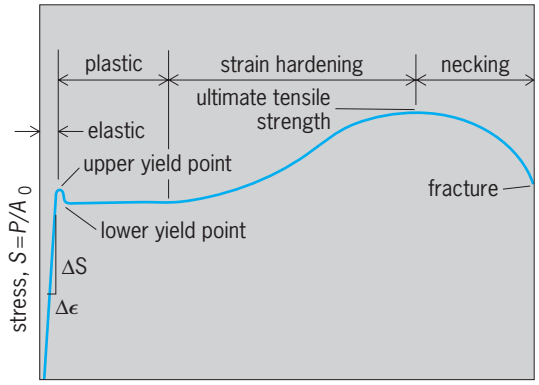

Stress-strain diagram for a low-carbon steel. ΔS = change in stress; $\Delta\epsilon$ = change in strain; P = force; A_0 = area of cross section.

A stress-strain diagram is a graphical representation of simultaneous values of stress and strain observed in tests and indicates material properties associated with both elastic and inelastic behavior (see illustration). It indicates significant values of stress-accompanying changes produced in the internal structure. *See* ELASTICITY; STRENGTH OF MATERIALS. [J.B.S.]

Strigiformes The owls, an order of nocturnal, worldwide predacious birds that are probably most closely related to the goatsuckers. The strigiformes are arranged in four families: Ogygoptyngidae (fossil), Protostrigidae (fossil), Tytonidae (barn owls; 11 species), and Strigidae (typical owls; 135 species). The bay owls (*Phodilus*), occurring in southern Asia and Africa, are somewhat intermediate between the Tytonidae and the Strigidae and are sometimes placed in a separate family; here they are treated as a subfamily, Phodilinae, of the Tytonidae.

Owls are small to medium in size, with soft plumage of somber colors. The large head has a facial disk that covers the feathered parabolic reflectors of the bird's acute directional hearing system. The eyes are large and capable of sight in very dim light, and the bill is strong and hooked. The wings are long and rounded, and the flight feathers are fringed for silent flight. The ulna has a unique bony arch on the shaft. Their strong legs are short to medium in length and terminate in strong feet. Owls are excellent fliers but walk poorly. Of the four toes on each foot, two point forward and two backward and bear strong claws. Prey is detected by acute night vision or by directional hearing; owls can locate and catch their prey in total darkness. Owls are generally nonmigratory and solitary, but some species live in small flocks. Courtship takes place at night with a male hooting to a female, which answers. A strong pair bond exists between the monogamous male and female, which usually build their nest in a tree cavity. Some species nest on the ground, on cliff ledges, or in abandoned crow or hawk nests. The clutch of up to seven eggs is incubated by both sexes, and after hatching, the young stay in the nest and are cared for by both parents. *See* AVES. [W.J.B.]

Stripping The removal of volatile component from a liquid by vaporization. The stripping operation is an important step in many industrial processes which employ absorption to purify gases and to recover valuable components from the vapor phase. In such processes, the rich solution from the absorption step must be stripped in order to permit recovery of the absorbed solute and recycle of the solvent. *See* GAS ABSORPTION OPERATIONS.

Stripping may be accomplished by pressure reduction, the application of heat, or the use of an inert gas (stripping vapor). Many processes employ a combination of all three; that is, after absorption at elevated pressure, the solvent is flashed to

atmospheric pressure, heated, and admitted into a stripping column which is provided with a bottom heater (reboiler). Solvent vapor generated in the rebolier or inert gas injected at the bottom of the column serves as stripping vapor which rises countercurrently to the downflowing solvent. When steam is used as stripping vapor for a system not miscible with water, the process is called steam stripping.

In addition to its use in conjunction with gas absorption, the term stripping is also used quite generally in technical fields to denote the removal of one or more components from a mixed system. Such usage covers (1) the distillation operation which takes place in a distilling column in the zone below the feed point, (2) the extraction of one or more components from a liquid by contact with a solvent liquid, (3) the removal of organic or metal coatings from solid surfaces, and (4) the removal of color from dyed fabrics. *See* DISTILLATION; ELECTROPLATING OF METALS; SOLVENT EXTRACTION. [A.L.K.]

Stroboscope An instrument for observing moving bodies by making them visible intermittently and thereby giving them the optical illusion of being stationary. A stroboscope may operate by illuminating the object with brilliant flashes of light or by imposing an intermittent shutter between the viewer and the object.

Stroboscopes are used to measure the speed of rotation or frequency of vibration of a mechanical part or system. They have the advantage over other instruments of not loading or disturbing the equipment under test. Mechanical equipment may be observed under actual operating conditions with the aid of stroboscopes. Parasitic oscillations, flaws, and unwanted distortion at high speeds are readily detected.

The flashing-light stroboscopes employ gas discharge tubes to provide a brilliant light source of very short duration. Tubes may vary from neon glow lamps, when very little light output is required, to special stroboscope tubes capable of producing flashes of several hundred thousand candlepower with a duration of only a few millionths of a second. *See* NEON GLOW LAMP; VAPOR LAMP. [A.R.E.]

Stroboscopic photography Stroboscopic or "strobe" photography generally refers to pictures of both single and multiple exposure taken by flashes of light from electrical discharges. Originally the term referred to multiple-exposed photographs made with a Stroboscopic disk as a shutter. One essential feature of modern Stroboscopic photography is a short exposure time, usually much shorter than can be obtained by a mechanical shutter.

High-speed photography with Stroboscopic light has proved to be one of the most powerful research tools for observing fast motions in engineering and in science. Likewise, the electrical system of producing flashes of light in xenon-filled flash lamps is of great utility for studio, candid, and press photography. *See* PHOTOGRAPHY; STROBOSCOPE. [H.E.E.]

Stromatolite A laminated, microbial structure in carbonate rocks (limestone and dolomite). Stromatolites are the oldest macroscopic evidence of life on Earth, at least 2.5 billion years old, and they are still forming in the seas. During the 1.5 billion years of Earth history before marine invertebrates appeared, stromatolites were the most obvious evidence of life, and they occur sporadically throughout the remainder of the geologic record. In Missouri and Africa, stromatolite reefs have major accumulations of lead, zinc, or copper; and in Montana, New Mexico, and Oman, stromatolites occur within oil and gas reservoirs. For geologists, the shapes of stromatolites are useful indications of their environmental conditions, and variations in form and microstructure of the laminations may be age-diagnostic in those most ancient sedimentary rocks that lack invertebrate fossils. *See* DOLOMITE; LIMESTONE; REEF.

Stromatolites are readily recognizable in outcrops by their characteristic convex-upward laminated structure. Individual, crescent-shaped laminations, which are generally about a millimeter thick, are grouped together to produce an enormous range of shapes and sizes.

The tiny, filamentous cyanobacteria (blue-green algae) that make present-day stromatolites, and similar filaments associated with the oldest stromatolites known, are considered one of the most successful organisms on Earth. Living stromatolites in the Bahamas and Western Australia possess laminations that record the episodic trapping and binding of sediment particles by the microbial mat. In the modern oceans, stromatolites develop almost exclusively in extreme marine conditions that exclude or deter browsing invertebrates and fish from destroying the microbial mats and inhibit colonization by competing algae. *See* ALGAE; CYANOBACTERIA. [R.Gi.]

Stromatoporoidea An extinct order thought to be sponges because of its close similarity to the class Sclerospongiae. Generally accepted stromatoporoids first appeared at the beginning of the Middle Ordovician and died out at the end of the Devonian. The Stromatoporoidea are worldwide in distribution, most commonly associated with bedded limestones and fossil reefs.

The stromatoporoid skeleton is called a coenosteum, which begins development as a thin, encrusting layer of irregular rods and plates called the peritheca. Most coenostea are hemispherical, or thick laminar or lens-shaped structures. However, some stromatoporoids are thin and irregularly undulatory and others are cylindrical, cone-shaped, branching, or anastomosing. *See* SCLEROSPONGIAE. [J.St.J.]

Strong nuclear interactions One of the fundamental physical interactions, which acts between a pair of hadrons. Hadrons include the nucleons, that is, neutrons and protons; the strange baryons, such as lambda (Λ) and sigma (Σ); the mesons, such as pion (π) and rho (ρ); and the strange meson, kaon (K). The nature of the interaction is determined principally through observations of the collision of a hadron pair. From this it is found that the interaction has a short range of about 10^{-15} m (10^{-13} in.) and is by far the dominant force within this range, being much larger than the electromagnetic interaction, which is next in magnitude. The strong interaction conserves parity and is time-reversal-invariant. *See* BARYON; HADRON; MESON; NUCLEON; PARITY (QUANTUM MECHANICS); STRANGE PARTICLES; SYMMETRY LAWS (PHYSICS); TIME REVERSAL INVARIANCE.

The interaction between baryons for distances greater than 10^{-15} m arises from the exchange of mesons. At relatively large distances, single-pion exchange dominates (illus. *a*). At shorter separation distances, the two-pion systems such as the ρ become important (illus. *b*). The interaction between the strange baryons, and between the strange baryons and the nucleons, is moderated by kaon exchange. To summarize this description, the interaction between baryons in the SU(3) multiplet is the consequence of the exchange of SU(3) spin 0 and spin 1 bosons. In a second

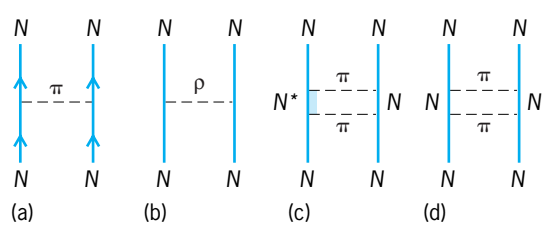

Interaction between nucleons (*a*) from exchange of single pion, (*b*) from exchange of ρ-meson, a two-pion system, (*c*) from exchange of two separate pions with formation of excited state of nucleon, and (*d*) without formation of excited state.

approximation, the exchange of two pions (illus. *c, d*), or more generally, the exchange of two members of the SU(3) spin 0 and spin 1 multiplet, is responsible for a component of the strong nuclear interaction. *See* UNITARY SYMMETRY.

The range of the interaction generated by these exchanges can be calculated by using the formula below, where m is the mass

$$\text{Range} = \frac{\hbar}{mc}$$

of the exchanged particles, $\hbar$ is Planck's constant divided by 2π, and c is the speed of light. According to the above equation, the range of the interaction developed when a single pion is exchanged (illus. 1*a*) is equal to 1.4×10^{-15} m (5.5×10^{-14} in.), while that due to two-pion exchange (illus. 1*d*) is 0.7×10^{-15} m (2.8×10^{-14} in.).

At short separation distances the quark-gluon structure of the baryons must be taken into account. The interaction must be considered as a property of the six-quark-plus-gluon system. The decisive elements are the Pauli principle obeyed by the quarks, and the mismatch between the six-quark wave function and the two-baryon wave function. Thus, at short distances the interaction is effectively repulsive or more generally independent of the kinetic energy of the baryons at infinite separations. *See* ELEMENTARY PARTICLE; EXCLUSION PRINCIPLE; FUNDAMENTAL INTERACTIONS; GLUONS; QUARKS. [H.F.]

Strongylida An order of nematodes in which the cephalic region may be adorned with three or six labia or the labia may be replaced by a corona radiata. All strongylid nematodes are parasitic. The order embraces eight superfamilies: Strongyloidea, Diaphanocephaloidea, Ancylostomatoidea, Trichostrongyloidea, Metastrongyloidea, Cosmocercoidea, Oxyuroidea, and Heterakoidea.

Strongyloidea. The Strongyloidea contain important parasites of reptiles, birds, and mammals. The early larval stages may be free-living microbivores, but the adults are always parasitic. Three species are important parasites of horses, *Strongylus vulgaris, S. equinus*, and *S. edentatus*. All three undergo direct life cycles; that is, infestations are acquired by ingestion of contaminated food.

Trichostrongyloidea. The Trichostrongyloidea comprise obligate parasites of all vertebrates but fishes. Normally they are intestinal parasites, but some are found in the lungs. The species are important parasites of sheep, cattle, and goats. The adult females lay eggs in the intestinal tract, which are passed out with the feces. In the presence of oxygen the eggs hatch in a few days. When the larvae are ingested by an appropriate host, their protective sheath is lost, and they proceed through the fourth larval stage to adulthood in the intestinal tract, where they may enter the mucosa. No migration takes place outside the gastrointestinal tract.

Metastrongyloidea. The Metastrongyloidea comprise obligate parasites of terrestrial and marine mammals, found commonly in the respiratory tract. In their life cycle they utilize both paratenic and intermediate hosts, among them a variety of invertebrates, including earthworms and mollusks. Two important species are *Metastrongylus apri* (swine lungworm) and *Angiostrongylus cantonensis* (rodent lungworm).

Heterakoidea. The Heterakoidea are capable of parasitizing almost any warm-blooded vertebrate as well as reptiles and amphibians. The species *Ascaridia galli* is the largest known nematode parasite of poultry; males are 2–3 in. (50–76 mm) long, and females 3–4.5 in. (75–116 mm). [A.R.M.]

Oxyuroidea. The Oxyuroidea constitute a large group of the phylum Nemata. Hosts include terrestrial mammals, birds, reptiles, amphibians, fishes, insects, and other arthropods.

The species are small to medium sized and thin bodied. With one exception, known life cycles are direct. Typically the eggs pass out of the host's alimentary tract onto the ground, where they become fully embryonated and infective. Normally the infective egg does not hatch until a susceptible animal ingests it. The cecum and colon of the host are the typical locations of these parasites. Larvae in all stages of development and adults occur in the gut.

The human pinworm, *Enterobius vermicularis*, is probably the most contagious of all helminthic diseases. It is estimated that 10% of the world's population suffer from this parasite, the majority being children. Indeed, incidence among schoolchildren in the cool regions of the world often approaches 100%. Infection occurs when eggs are inhaled or ingested. The most common method of transmission is from anus to mouth. Because of the aerial transmission, this disease is highly contagious. Though the infection is seldom serious, the behavioral symptoms are disturbing: nail biting, teeth grinding, anal scratching, insomnia, nightmares, and even convulsions. Several medical treatments are available, but there is often the danger of reinfestation from contaminated objects within the household or institution. *See* NEMATA. [J.T.L.]

Strontianite The mineral form of strontium carbonate, usually with some calcium replacing strontium. It characteristically occurs in veins with barite or celestite or as masses in certain sedimentary rocks. Strontianite is normally prismatic, but it may also be massive. It may be colorless or gray with yellow, green, or brownish tints. The hardness is $3^{1}/_{2}$ on Mohs scale, and the specific gravity of 3.76. It occurs at Strontian, Scotland, and in Germany, Austria, Mexico, and India and, in the United States, in the Strontium Hills of Calfornia. *See* CARBONATE MINERALS; STRONTIUM. [R.I.Ha.]

Strontium A chemical element, Sr, atomic number 38, and atomic weight 87.62. Strontium is the least abundant of the alkaline-earth metals. The crust of the Earth is 0.042% strontium, making this element as abundant as chlorine and sulfur. The main ores are celestite, $SrSO_4$, and strontianite, $SrCO_3$. *See* ALKALINE-EARTH METALS; PERIODIC TABLE; STRONTIANITE.

Strontium nitrate is used in pyrotechnics, railroad flares, and tracer bullet formulations. Strontium hydroxide forms soaps and greases with a number of organic acids which are structurally stable, resistant to oxidation and breakdown over a wide temperature range.

Properties of strontium

Property	Value
Atomic number	38
Atomic weight	87.62
Isotopes (stable)	84, 86, 87, 88, 90
Boiling point, °C	1638(?)
Melting point, °C	704(?)
Density, g/cm3 at 20°C	2.6

Strontium is divalent in all its compounds which are, aside from the hydroxide, fluoride, and sulfate, quite soluble. Strontium is a weaker complex former than calcium, giving a few weak oxy complexes with tartrates, citrates, and so on. Some physical properties of the element are given in the table. [R.F.R.]

Strophanthus A genus of woody climbers of the dogbane family (Apocynaceae), natives of tropical Asia and Africa. They are the source of arrow poisons. The dried, greenish, ripe seeds of *Strophanthus hispidus* and *S. kombe* contain the glucoside strophanthin, which is much used in treating heart ailments. Strophanthin acts directly on heart muscle, increasing muscular force. It causes the heart to beat more regularly and decreases the pulse rate. Strophanthin is a precursor of cortisone, which is used in the treatment of arthritis. *See* GENTIANALES. [P.D.St.; E.L.C.]

Strophomenida An extinct order of brachiopods, in the subphylum Rhynchonelliformea, that inhabited shallow shelf seas of the Paleozoic Era. It contains the greatest taxonomic diversity of all rhynchonelliform brachiopod orders, and its members exhibit a diverse range of morphologies and modes of life.

Strophomenida contains four suborders: Strophomenidina, Chonetidina, Productidina, and Strophalosiidina. A fifth suborder, Orthotetidina, shares characteristics with both strophomenids and primitive orthids. Strophomenids are distinguished from all other articulate brachiopods by their concavo-convex or plano-convex lateral profiles and the lack of a functional pedicle in adults. They also possess strophic (straight) hinge lines and variable ornamentation from fine-ribbed early Paleozoic forms to more elaborate spine-bearing late Paleozoic forms. Adult strophomenids were mostly free-living, sessile, epifaunal suspension feeders. The concavo-convex lateral profile and elaborate spines were adaptations for living on soft substrates. *See* BRACHIOPODA; RHYNCHONELLIFORMEA. [M.E.P.]

Structural analysis A detailed evaluation intended to assure that, for any structure, the deformations will be sufficiently below allowable values that structural failure will not occur. The deformations may be elastic (fully recoverable) or inelastic (permanent). They may be small, with an associated structural failure that is cosmetic; for example, the deflection of a beam supporting a ceiling may cause cracking of the plaster. They may be large, with an associated structural failure that is catastrophic; for example, the buckling of a column or the fracture of a tension member causes complete collapse of the structure.

Structural analysis may be performed by tests on the actual structure, on a physical model of the structure to some scale, or through the use of a mathematical model. Tests on an actual structure are performed in those cases where many similar structures will be produced, for example, automobile frames, or where the cost of a test is justified by the importance and difficulty of the project, for example, a lunar lander. Physical models are sometimes used where subassemblages of major structures are to be investigated. The vast majority of analyses, however, are on mathematical models, particularly in the field of structural engineering which is concerned with large structures such as bridges, buildings, and dams. *See* BRIDGE; BUILDINGS; DAM; STRUCTURE (ENGINEERING).

The advent of the digital computer made it possible to create mathematical models of great sophistication, and almost all complex structures are now so analyzed. Programs of such generality have been written as to permit the analysis of any structure. These programs permit the model of the structure to be two- or three-dimensional, elastic or inelastic, and determine the response to forces that are static or dynamic. Most of the programs utilize the stiffness method, in which the stiffnesses of the individual elements are assembled into a stiffness matrix for the entire structure, and analysis is performed in which all behavior is assumed to be linearly elastic. *See* DIGITAL COMPUTER; ELASTICITY.

The structural engineer's function continues to require training and experience in conceptualizing the structure, choosing the appropriate model, estimating the loads that will be of importance, coding the information for the program, and interpreting the results. The analyst usually enters the process after the conceptualization. Most structures consist of assemblies of members connected at joints. While all real members transmit axial, torsional, and bending actions, the majority of buildings and bridges are analyzed as trusses, beams, and frames with either axial or bending forces predominant. *See* BEAM; ENGINEERING DESIGN; STRESS AND STRAIN; STRUCTURAL DESIGN; TRUSS.

Whether the model selected is detailed or simplified, one extremely important part of the analysis consists of the estimate of the loads to be resisted. For bridges and buildings, the primary vertical loads are gravity loads. These include the weight of the structure itself, and such appurtenances as will be permanent in nature. These are referred to as dead loads. The loads to be carried, the live loads, may consist of concentrated loads (heavy objects occupying little space, for example, a printing press), or loads distributed over relatively large areas (such as floor and deck coverings). Horizontal loads on buildings are produced by wind and by the inertia forces created during earthquakes. In seismic analysis, computers are used to simulate the dynamic characteristics of the structure. The accelerations actually measured during earthquakes are then used to determine the response of the structure. *See* LOADS, DYNAMIC; LOADS, TRANSVERSE; SEISMIC RISK. [G.D.B.R.]

Structural chemistry Much of chemistry is explainable in terms of the structures of chemical compounds. The understanding of these structures hinges very strongly on understanding the electronic configurations of the elements. The union of atoms, and therefore the formation of compounds from the elements, is associated with interactions among the extranuclear electrons of the individual atoms. Electronic interactions among atoms may occur in two ways: Electrons may be transferred from one atom to another, or they may be shared by two (or more) atoms. The first type of interaction is called electrovalence and results in the formation of electrically charged monatomic ions. The second, covalence, leads to the formation of molecules and complex ions. *See* CHEMICAL BONDING. [D.H.B.]

In considering structures more complex than those derived from simple monoatomic ions, the logical step is to consider single polyhedral aggregates of atoms. In its most precise sense, structure is used to denote a knowledge of the bonding distances and angles between atoms in chemical compounds and, in turn, the geometrical arrangements which they form. These atomic arrangements and the associated distances and angles serve uniquely as "fingerprints" of these atom spatial configurations, and depend very much on the electronic configurations around atoms. The chemical combination of neutral atoms to produce uncharged species results in molecule formation, whereas the similar combination of atoms or ions possessing a net charge results in the formation of complex ions. A basic understanding of the species formed involves the concept of the coordination polyhedron, which allows a simple classification of the structures of many polyatomic molecules and ions. This type of classification is particularly useful because it conveniently explains the packing together of simple chemical molecules or ions in terms of highly symmetrical polyhedra. There is an obvious connection between polyhedra and the structures found in crystalline solids formed from them. Crystal formation often involves the linking of convex polyhedra by the sharing of corners, edges, or faces, ultimately forming space-filling assemblies in which all faces of each polyhedron are in contact with faces of other polyhedra. The most important simple polyhedrons are the tetrahedron, the trigonal bipyramid, the octahedron, the pentagonal bipyramid, and the cube. The most commonly observed of these polyhedral configurations are the tetrahedron (four faces) and the octahedron (six faces).

The simplest correlative device which accurately summarizes a very large number of structures and enables the chemist to predict, with a good chance of success, the geometric array of the atoms in a compound of known composition, is based on an extreme electrostatic model. This model, or theory, represents the bonds in a purely formal way. The central atom is considered to be a positive ion having a charge equal to its oxidation state. The groups attached to the central atom (the ligands) are then treated either as negative ions or as neutral dipolar molecules. The principal justification for this approach lies in its successful correlation of a vast amount of information.

A number of significant observations can be made with regard to these formulations. There are series of ions, or ions and molecules, having the same type of composition, differing only in the nature of the central ion and the net charge

on the aggregate. Examples are found in the series: NO_3^-, CO_3^{2-}, BO_3^{3-}, ClO_3^-, SO_3^{2-}, PO_3^{3-}; ClO_4^-, SO_4^{2-}, PO_4^{3-}, SiO_4^{4-}; AlF_6^{3-}, SiF_6^{2-}, PF_6^-. The numbers of atomic nuclei and of electrons are the same for all the members of each series; consequently, these are called isoelectronic series. Not only are the several chemical entities in such series isoelectronic, but they are usually identical in geometrical structure (isostructural).

It may also be observed that corresponding ions from a given vertical family of the periodic table commonly vary in coordination number. A useful example is found in N^{5+} and P^{5+} which form NO_3^- and PO_4^{3-}, respectively. In addition, some neutral molecules expand their coordination numbers to form stable anionic halo complexes, whereas others do not. Thus, SiF_4 reacts with fluoride ion to form SiF_6^{2-}, whereas CF_4 does not form a similar complex ion. The most satisfactory explanation of these and many related observations is conveniently formulated in terms of the electrostatic model chosen here.

The necessary condition for stability of the coordination polyhedron MA_n requires that the anions A are each in contact with the central atom M. As a consequence of this condition, the limit of stability of the structure arises in those cases where the anions are also mutually in contact. Larger ligands, or anions, would not be in contact with the central ion. This relationship is usually summarized in terms of the limiting ratio of the radius of the cation, r_M, to that of the anion, r_A, below which the anions would no longer be in contact with the cation.

According to the valence-bond theory, the principal requirements for the formation of a covalent bond are a pair of electrons and suitably oriented electron orbitals on each of the atoms being bonded. The geometry of the atoms in the resulting coordination polyhedron is correlated with the orientation of the orbitals on the central atom. The orbitals used depend on the energies of the electrons in them. In general, the order of increasing energy of the electron orbitals is $(n - 1)d < ns < np < nd$. It is concluded that a nontransition atom having one valence electron will form a covalent bond utilizing an *s* orbital. In those cases where an unshared pair of electrons may be assigned to the *ns* orbital, as many as three equivalent bonds may be formed by utilizing the three *np* orbitals of the central atom. Because of the orientation of these *p* orbitals with respect to each other, the three resulting bonds should be at 90° to each other. This expectation is nearly realized in PH_3. In order to account for four or six equivalent bonds, or for that matter in order to account for all the remaining polyhedral and polygonal structures, except the angular structure for a coordination number of 2 (with two unshared pairs of electrons on the central atom), an additional assumption is necessary. It is assumed that *s* and *p*, *s* and *d*, or *s*, *p*, and *d* orbitals, are replaced by new orbitals, called hybridized orbitals. These hybridized orbitals are derived from the original orbitals (mathematically) in such a way that the required number of equivalent bonds may be formed. In the simplest case, it is shown that *s* and *p* may be combined to form two equivalent *sp* hybridized orbitals directed at 180° to each other. Other sets of hybridized orbitals have been shown to be appropriate to describe the bonding in other structures. *See* LIGAND FIELD THEORY.

Among inert-gas ions of the first row of eight elements in the periodic table, there are four orbitals available for covalent bond formation, one 2*s* and three 2*p*. Consequently, a maximum of four bonds may be formed. This is in general agreement with the existence of the tetrahedron as the limiting coordination polyhedron among these elements, for example, BeF_4^{2-}, BF_4^-, CCl_4, NH_4^+. Although only Li^+ deviates from this pattern, having a coordination number of 6 in its crystalline halides, these compounds are best treated as simple electrovalent salts. In keeping with the limitation of only four orbitals, the formation of double or triple bonds between atoms of these elements reduces the coordination number of the central atom. Thus, the highest coordination number of a first-row element forming one double bond is 3. This is illustrated by the structures below.

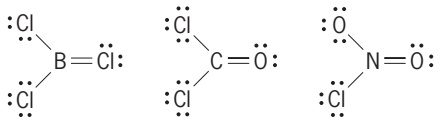

In these and similar examples, the geometric array is determined by the formation of three single bonds utilizing sp^2 hybridized orbitals on the central atom and *p* orbitals on the ligand. In general, the bonds determining the geometry of a molecule or ion (in this way) are called σ bonds. The double bond results from the superposition of a second bond, a π bond, between two atoms. In this example the formation of a π bond reduces the number of σ bonds from four to three, thus changing the geometries of the corresponding molecules or ions from tetrahedral to trigonal planar.

The formation of a second π bond (a triple bond or two double bonds) reduces the coordination number of the atom in question still further, resulting in the linear *sp* set of hybridized orbitals being utilized in σ-bond formation. *See* VALENCE.

With regard to the nature of doubly bonded compounds, another problem arises when such structures are viewed from the standpoint of valence-bond theory. In the species BCl_3, $COCl_2$, NO_2Cl, and many similar substances, nonequivalent bonds are predicted. The doubly bonded oxygen should be closer to the central carbon atom than the singly bonded ones. This is not found to be true experimentally so long as the similar atoms are otherwise equivalent. There is only one observable C—O distance in carbonate, one N—O distance in nitrate, and so on. To account for such facts as these, the concept of resonance must be introduced. If the π bond exists, it must exist equally between the central atom and all the equivalent oxygen atoms. The resonance method of describing this situation is to say that one of the pictorial structures is inadequate to describe the substance properly, but that enough pictorial structures (resonance structures) should be considered to permute the double bond about all the equivalent bonds. The true structure is assumed to be something intermediate to all the resonance structures and more stable than any of them because it exists in preference to any one of them. The resonance structures for CO_3^{2-} are the following:

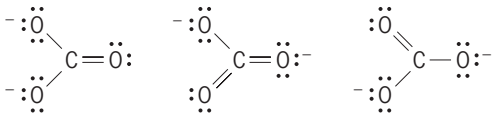

See CONJUGATION AND HYPERCONJUGATION; RESONANCE (MOLECULAR STRUCTURE).

The classic homologous series of compounds in organic chemistry provide useful examples involving the condensation of polyhedrons containing the same central element in the individual units. The general formula, $C_nH_{2n} + 2$, represents a large number of compounds extending from the lowest member, methane, CH_4, to polyethylene, a plastic of economic importance in which *n* is a very large number. Two ways exist for the linking together of these tetrahedrons. This gives rise to two molecular forms, both of which are stable, well-known compounds. It is

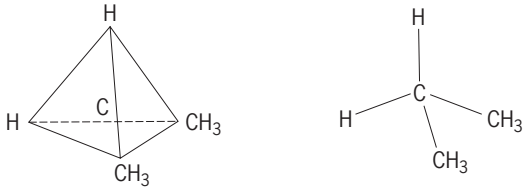

Fig. 1. Geometric structure of propane.

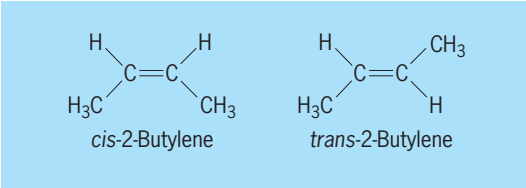

Fig. 2. Cis-trans isomerism among olefins.

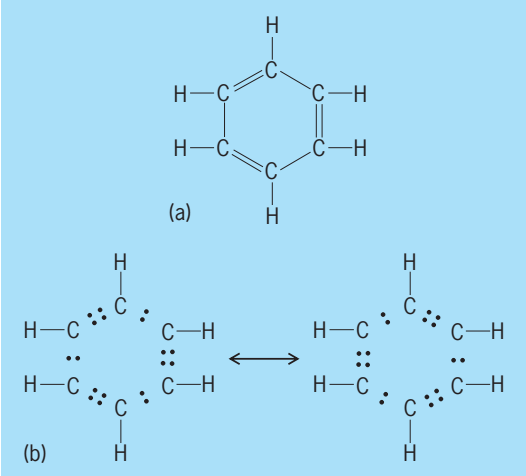

Fig. 3. Benzene molecule. (*a*) Structural formula. (*b*) Two forms in resonance.

an essential part of these structures that each —C link is linear (because it is merely a σ-bond); however, when two carbon atoms are linked to a third, the C—C—C angle is essentially determined by the bond angle of the central carbon atom (that is, the other carbons may be treated as ligands to the first (Fig. 1). The other familiar homologous series of organic chemistry differ from the saturated hydrocarbons in having at least one unique coordination polyhedron of a different type. The olefins contain two doubly bonded, or unsaturated, carbon atoms, whose polyhedral structures are trigonal planar, and the remainder of the carbons are tetrahedral. As in the case of the aliphatic hydrocarbons, the olefins exhibit an isomerism which is associated with the branching of the chain structure. In addition, the presence of two linked trigonal planar carbon atoms and the fact that the polyhedrons cannot rotate about the double bond give rise to a different kind of isomerism, called cis-trans isomerism (Fig. 2). *See* BOND ANGLE AND DISTANCE.

The existence of a predicted isomerism provides one of the most important confirmations of the theories of chemical structure. In general, the polyhedral view of molecular structure, as described here, has been thoroughly verified by the discovery of the many types of predicted isomerism. The first really convincing proof of the tetrahedral structures of saturated carbon atoms involved optical isomerism. *See* OPTICAL ACTIVITY.

The aromatic hydrocarbons are characterized by cyclic arrangements of trigonal planar carbon atoms (Fig. 3*a*). The highly symmetrical nature of the benzene molecule is not fully represented by such a structure. The figure indicates the presence of three double and three single bonds in the ring. It has been shown that the C—C bonds are all the same and, consequently, the true structure of the substance must be represented by two resonance structures which interchange the single and double bonds (Fig. 3*b*). *See* BENZENE. [J.M.Wi.]

Structural connections Methods of joining the individual members of a structure to form a complete assembly. The connections furnish supporting reactions and transfer loads from one member to another. Loads are transferred by fasteners (rivets, bolts) or welding supplemented by suitable arrangements of plates, angles, or other structural shapes. When the end of a member must be free to rotate, a pinned connection is used.

The suitability of a connection depends on its deformational characteristics as well as its strength. Rotational flexibility or complete rigidity must be provided according to the degree of end restraint assumed in the design. A rigid connection maintains the original angles between connected members virtually unchanged after loading. Flexible or nonrestraining connections permit rotation approximately equal to that at the ends of a simply supported beam. Intermediate degrees of restraint are called semirigid.

A commonly used form of connection for rolled-beam sections, called a web connection, consists of two angles which are attached to opposite sides of a member and which are in turn connected to the web of a supporting beam, girder, column, or framing at right angles. A shelf angle may be added to facilitate erection (Fig. 1).

A bracket or seat on which the end of the beam rests is a seat connection; it is intended to furnish the end reaction of the supported beam. Two general types are used: The unstiffened seat provides bearing for the beam by a projecting plate or angle leg which offers resistance only by its own flexural strength (Fig. 2); the stiffened seat is supported by a vertical plate or angle which transfers the reaction force to the supporting member without flexural distortion of the outstanding seat.

When the action line of a transferred force does not pass through the centroid of the connecting fastener group or welds, the connection is subjected to rotational moment which produces additional shearing stresses in the connectors. The load transmitted by diagonal bracing to a supporting column flange through a gusset plate is eccentric with reference to the connecting fastener group.

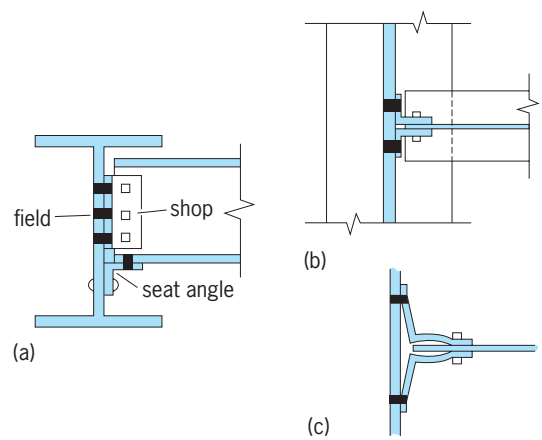

Fig. 1. Riveted or bolted web connections. (*a*) Beam-to-ginder, elevation view. (*b*) Plan view. (*c*) Bending of angle legs.

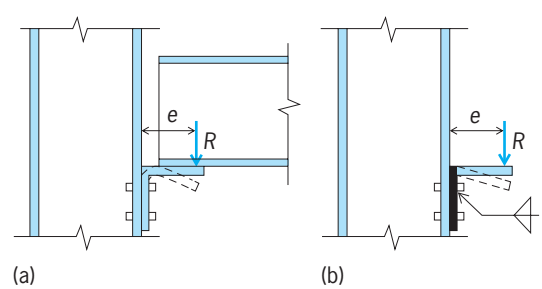

Fig. 2. Unstiffened seat connections. (*a*) Fastened angle seat. (*b*) Welded angle seat. *R* = reaction load; *e* = distance to reaction from column face.

In beam-to-column connections and stiffened seat connections or when members transfer loads to columns by a gusset plate or a bracket, the fasteners are subjected to tension forces caused by the eccentric connection. Although there are initial tensions in the fasteners, the final tension is not appreciably greater than the initial tension.

Rigidity and moment resistance are necessary at the ends of beams forming part of a continuous framework which must resist lateral and vertical loads. Wind pressures tend to distort a building frame, producing bending in the beams and columns which must be suitably connected to transfer moment and shear. The resisting moment can be furnished by various forms of angle T for fasteners or welded or bracket connections.

Where appreciably angular change between members is expected, and in special cases where a hinge support without moment resistance is desired, connections are pinned. Many bridge trusses and large girder spans have pin supports. *See* BOLT; JOINT (STRUCTURES); RIVET. [J.B.S.]

Structural deflections The deformations or movements of a structure and its components, such as beams and trusses, from their original positions. It is as important for the designer to determine deflections and strains as it is to know the stresses caused by loads. *See* STRESS AND STRAIN.

Deflections may be computed by any of several methods. Generally the computation is based on the assumption that stress is proportional to strain. As a result, deflection equations involve the modulus of elasticity E, which is a measure of the stiffness of a material.

The relation between deflections at different parts of a structure is indicated by Maxwell's law of reciprocal deflections. This states that if a load P is applied at any point A in any direction a and causes a shift of another point B in direction b, the same load applied at B in direction b will cause an equal shift of A in direction a (see illustration). The law is used in a number of ways such as in simplifying deflection calculations, checking the accuracy of computations, and producing influence lines. *See* STRUCTURAL ANALYSIS.

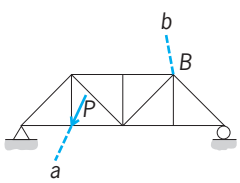

Example of Maxwell's law of reciprocal deflections.

Beam and truss deflections usually are computed by similar methods, except that integration is used for equations and summation for trusses. Beam deflection equations involve bending moments and moments of inertia. Truss deflection equations are based on the stresses and cross-sectional areas of chords and web members. Deflections may also be determined graphically. *See* BEAM; TRUSS. [J.B.S.]

Structural design The selection of materials and member type, size, and configuration to carry loads in a safe and serviceable fashion. In general, structural design implies the engineering of stationary objects such as buildings and bridges, or objects that may be mobile but have a rigid shape such as ship hulls and aircraft frames. Devices with parts planned to move with relation to each other (linkages) are generally assigned to the area of mechanical design.

Structural design involves at least five distinct phases of work: project requirements, materials, structural scheme, analysis, and design. For unusual structures or materials a sixth phase, testing, should be included. These phases do not proceed in a rigid progression, since different materials can be most effective in different schemes, testing can result in changes to a design, and a final design is often reached by starting with a rough estimated design, then looping through several cycles of analysis and redesign. Often, several alternative designs will prove quite close in cost, strength, and serviceability. The structural engineer, owner, or end user would then make a selection based on other considerations.

Before starting design, the structural engineer must determine the criteria for acceptable performance. The loads or forces to be resisted must be provided. For specialized structures this may be given directly, as when supporting a known piece of machinery, or a crane of known capacity. For conventional buildings, building codes adopted on a municipal, county, or state level provide minimum design requirements for live loads (occupants and furnishings, snow on roofs, and so on). The engineer will calculate dead loads (structure and known, permanent intallations) during the design process. For the structure to be serviceable or useful, deflections must also be kept within limits, since it is possible for safe structures to be uncomfortably "bouncy." Very tight deflection limits are set on supports for machinery, since beam sag can cause driveshafts to bend, bearings to burn out, parts to misalign, and overhead cranes to stall. Beam stiffness also affects floor "bounciness," which can be annoying if not controlled. In addition, lateral deflection, sway, or drift of tall buildings is often held within approximately height/500 (1/500 of the building height) to minimize the likelihood of motion discomfort in occupants of upper floors on windy days. *See* LOADS, DYNAMIC; LOADS, TRANSVERSE.

Technological advances have created many novel materials such as carbon fiber- and boron fiber-reinforced composites, which have excellent strength, stiffness, and strenth-to-weight properties. However, because of the high cost and difficult or unusual fabrication techniques required, glass-reinforced composites such as fiberglass are more common, but are limited to lightly loaded applications. The main materials used in structural design are more prosaic and include steel, aluminum, reinforced concrete, wood, and masonry. *See* COMPOSITE MATERIAL; MASONRY; PRECAST CONCRETE; PRESTRESSED CONCRETE; REINFORCED CONCRETE; STRUCTURAL MATERIALS.

In an actual structure, various forces are experienced by structural members, including tension, compression, flexure (bending), shear, and torsion (twist). However, the structural scheme selected will influence which of these forces occurs most frequently, and this will influence the process of material selection. *See* SHEAR; TORSION.

Analysis of structures is required to ensure stability (static equilibrium), find the member forces to be resisted, and determine deflections. It requires that member configuration, approximate member sizes, and material properties be known or assumed. Aspects of analysis include: equilibrium; stress, strain, and elastic modulus; linearity; plasticity; and curvature and plane sections. Various methods are used to complete the analysis.

Once a structure has been analyzed (by using geometry alone if the analysis is determinate, or geometry plus assumed member sizes and materials if indeterminate), final design can proceed. Deflections and allowable stresses or ultimate strength must be checked against criteria provided either by the owner or by the governing building codes. Safety at working loads must be calculated. Several methods are available, and the choice depends on the types of materials that will be used. Once a satisfactory scheme has been analyzed and designed to be within project criteria, the information must be presented for fabrication and construction. This is commonly done through drawings, which indicate all basic dimensions, materials, member sizes, the anticipated loads used in design, and anticipated forces to be carried through connections. [R.L.T.; L.M.J.]

Structural geology The branch of geology that deals with study and interpretation of deformation of the Earth's crust. Deformation brings about changes in size (dilation), shape

(distortion), position (translation), or orientation (rotation). Evidence for the changes caused by deformation are commonly implanted into geologic bodies in the form of recognizable structures, such as faults and joints, folds and cleavage, and foliation and lineation. The geologic record of structures and structural relations is best developed and most complicated in mountain belts, the most intensely deformed parts of the Earth's crust. *See* MOUNTAIN SYSTEMS.

The discipline of structural geology harnesses three interrelated strategies of analysis: descriptive analysis, kinematic analysis, and dynamic analysis. Descriptive analysis is concerned with recognizing and describing structures and measuring their orientations. Kinematic analysis focuses on interpreting the deformational movements responsible for the structures. Dynamic analysis interprets deformational movements in terms of forces, stresses, and mechanics. The ultimate goal of these interdependent approaches is to interpret the physical evolution of crustal structures, that is, tectonic analysis. A major emphasis in modern structural geology is strain analysis, the quantitative analysis of changes in size and shape of geologic bodies, regardless of scale.

There are many significant practical applications of structural geology. An understanding of the descriptive and geometric properties of folds and faults, as well as mechanisms of folding and faulting, is of vital interest to exploration geologists in the petroleum industry. Ore deposits commonly are structurally controlled, or structurally disturbed, and for these reasons detailed structural geologic mapping is an essential component of mining exploration. Other applications of structural geology include the evaluation of proposals for the disposal of radioactive waste in the subsurface, and the targeting of safe sites for dams, hospitals, and the like in regions marked by active faulting. *See* FAULT AND FAULT STRUCTURES; FOLD AND FOLD SYSTEMS; ORE AND MINERAL DEPOSITS; PETROLEUM GEOLOGY. [G.H.D.]

Structural materials Construction materials which, because of their ability to withstand external forces, are considered in the design of a structural framework.

Brick is the oldest of all artificial building materials. It is classified as face brick, common brick, and glazed brick. Face brick is used on the exterior of a wall and varies in color, texture, and mechanical perfection. Common brick consists of the kiln run of brick and is used behind whatever facing material is employed providing necessary wall thickness and additional structural strength. Glazed brick is employed largely for interiors where beauty, ease of cleaning, and sanitation are primary considerations. *See* BRICK.

Structural clay tiles are burned-clay masonry units having interior hollow spaces termed cells. Such tile is widely used because of its strength, light weight, and insulating and fire-protection qualities. *See* TILE.

Architectural terra-cotta is a burned-clay material used for decorative purposes. The shapes are molded either by hand in plaster-of-paris molds or by machine, using the stiff-mud process.

Building stones generally used are limestone, sandstone, granite, and marble. Until the advent of steel and concrete, stone was the most important building material. Its principal use now is as a decorative material because of its beauty, dignity, and durability. *See* GRANITE; LIMESTONE; MARBLE; SANDSTONE; STONE AND STONE PRODUCTS.

Concrete is a mixture of cement, mineral aggregate, and water, which, if combined in proper proportions, form a plastic mixture capable of being placed in forms and of hardening through the hydration of the cement. *See* CONCRETE; PRESTRESSED CONCRETE; REINFORCED CONCRETE.

The cellular structure of wood is largely responsible for its basic characteristics, unique among the common structural materials. When cut into lumber, a tree provides a wide range of material which is classified according to use as yard lumber, factory or shop lumber, and structural lumber. Laminated lumber is used for beams, columns, arch ribs, chord members, and other structural members. Plywood is generally used as a replacement for sheathing, or as form lumber for reinforced concrete structures. *See* LUMBER; PLYWOOD; WOOD ANATOMY; WOOD PRODUCTS.

Important structural metals are the structural steels, steel castings, aluminum alloys, magnesium alloys, and cast and wrought iron. Steel castings are used for rocker bearings under the ends of large bridges. Shoes and bearing plates are usually cast in carbon steel, but rollers are often cast in stainless steel. Aluminum alloys are strong, lightweight, and resistant to corrosion. The alloys most frequently used are comparable with the structural steels in strength. Magnesium alloys are produced as extruded shapes, rolled plate, and forgings. The principal structural applications are in aircraft, truck bodies, and portable scaffolding. Gray cast iron is used as a structural material for columns and column bases, bearing plates, stair treads, and railings. Malleable cast iron has few structural applications. Wrought iron is used extensively because of its ability to resist corrosion. It is used for blast plates to protect bridges, for solid decks to support ballasted roadways, and for trash racks for dams. *See* ALUMINUM; CAST IRON; MAGNESIUM ALLOYS; STEEL; STRUCTURAL STEEL; WROUGHT IRON. [C.M.A.]

Composite materials are engineered materials that contain a load-bearing material housed in a relatively weak protective matrix. A composite material results when two or more materials, each having its own, usually different characteristics, are combined, producing a material with properties superior to its components. The matrix material (metallic, ceramic, or polymeric) bonds together the reinforcing materials (whiskers, laminated fibers, or woven fabric) and distributes the loading between them. *See* COMPOSITE MATERIAL; CRYSTAL WHISKERS.

Fiber-reinforced polymers (FRP) are a broad group of composite materials made of fibers embedded in a polymeric matrix. Compared to metals, they generally have relatively high strength-to-weight ratios and excellent corrosion resistance. They can be formed into virtually any shape and size. Glass is by far the most used fiber in FRP (glass-FRP), although carbon fiber (carbon-FRP) is finding greater application. Although complete FRP shapes and structures are possible, the most promising application of FRP in civil engineering is for repairing structures or infrastructure. FRP can be used to repair beams, walls, slabs, and columns. *See* POLYMERIC COMPOSITE. [M.Sc.]

Structural petrology The study of the structural aspects of rocks, as distinct from the purely chemical and mineralogical studies that are generally emphasized in other branches of petrology. The term was originally used synonymously with petrofabric analysis, but is sometimes restricted to denote the analysis of only microscopic structural and textural features. *See* PETROFABRIC ANALYSIS; PETROGRAPHY. [J.M.Ch.]

Structural plate A simple rolled steel section used as an isolated structural element, as a support of other structural elements, or as part of other structural elements. When isolated plates are extremely thick, their design is controlled by shear; plates of moderate thickness are controlled by bending (with some torsion), and very thin plates carry their loads principally by tensile membrane action. Although stresses of all types exist in all plates, it is usually sufficient to deal with only the most significant. Bending is the most common design criterion.

Plates are commonly used as cover plates on wide-flange beams, as the flanges and webs of plate girders, and as the sides of tube-shaped beams and columns. In all these cases, serious consideration must be given to the fact that the plate may buckle when compressed. Fortunately, the plates have edge supports in the direction of the stress, so they function as panels rather than as beams. Their ratios of length to width are large enough that the resistance to local buckling of the plate element depends upon its width-thickness ratio, practically independent of its length. (The length of the overall section is still significant in

determining the member's capacity.) *See* BEAM; COLUMN; JOINT (STRUCTURES); LOADS, TRANSVERSE; STRUCTURAL STEEL. [G.D.Br.]

Structural steel Steel used in engineering structures, usually manufactured by either the open-hearth or the electric-furnace process. The exception is carbon-steel plates and shapes whose thickness is $^{7}/_{16}$ in. (11 mm) or less and which are used in structures subject to static loads only. These products may be made from acid-Bessemer steel. The physical properties and chemical composition are governed by standard specifications of the American Society for Testing and Materials (ASTM). Structural steel can be fabricated into numerous shapes for various construction purposes. *See* STEEL. [W.G.B.]

Structure (engineering) An arrangement of designed components that provides strength and stiffness to a built artifact such as a building, bridge, dam, automobile, airplane, or missile. The artifact itself is often referred to as a structure, even though its primary function is not to support but, for example, to house people, contain water, or transport goods. *See* AIRPLANE; AUTOMOBILE; BRIDGE; BUILDINGS; DAM.

The primary requirements for structures are safety, strength, economy, stiffness, durability, robustness, esthetics, and ductility. The safety of the structure is paramount, and it is achieved by adhering to rules of design contained in standards and codes, as well as in exercising strict quality control over all phases of planning, design, and construction. The structure is designed to be strong enough to support loads due to its own weight, to human activity, and to the environment (such as wind, snow, earthquakes, ice, or floods). The ability to support loads during its intended lifetime ensures that the rate of failure is insignificant for practical purposes. The design should provide an economical structure within the constraints of all other requirements. The structure is designed to be stiff so that under everyday conditions of loading and usage it will not deflect or vibrate to an extent that is annoying to the occupants or detrimental to its function. The materials and details of construction have durability, such that the structure will not corrode, deteriorate, or break under the effects of weathering and normal usage during its lifetime. A structure should be robust enough to withstand intentional or unintentional misuse (for example, fire, gas explosion, or collision with a vehicle) without totally collapsing. A structural design takes into consideration the community's esthetic sensibilities. Ductility is necessary to absorb the energy imparted to the structure from dynamic loads such as earthquakes and blasts. *See* CONSTRUCTION ENGINEERING; ENGINEERING DESIGN.

Common structural materials are wood, masonry, steel, reinforced concrete, aluminum, and fiber-reinforced composites. Structures are classified into the categories of frames, plates, and shells, frequently incorporating combinations of these. Frames consist of "stick" members arranged to form the skeleton on which the remainder of the structure is placed. Plated structures include roof and floor slabs, vertical shear walls in a multistory building, or girders in a bridge. Shells are often used as water or gas containers, in roofs of arenas, or in vehicles that transport gases and liquids. The connections between the various elements of a structure are made by bolting, welding or riveting. *See* COMPOSITE MATERIAL; CONCRETE; STEEL; STRUCTURAL MATERIALS. [T.V.G.]

Struthioniformes A small order of weak-flying, partridgelike birds and giant flightless ratite birds found in the southern continents. Their relationship to other birds is unknown. The struthioniforms are characterized by a palaeognathous palate which provides freedom of movement between segments of the palate and the jaw. For this reason they are frequently placed in a separate superorder, the Palaeognathae. However, that distinction places too much emphasis on their separation from other birds. Contrary to common opinion, no evidence supports the concept that the palaeognathous birds are primitive among living birds.

The order Struthioniformes can be divided into three suborders, Lithornithi, Tinami, and Struthioni, each of which includes both fossil and extant representatives.

The struthioniforms are medium-sized to giant birds. The ostriches are the largest extant birds, but some of the moas, elephant birds, and dromornithids were even larger. The head is small; a medium flattened bill is a common feature except in the kiwis, which have a long bill used for probing into the ground for worms. The wings are reduced in all forms except the lithionithids and tinamous, which can still fly, and the plumage is soft. Their legs are strong, and all forms run well. The struthioniforms eat a variety of foods, especially large fruits and other large food items. Breeding is polygamous, with two or more females laying eggs in a single nest. The males are responsible for incubation, and they assume the major or sole role in caring for the downy young, which leave the nest after hatching. *See* AVES; RATITES. [W.J.B.]

Strychnine alkaloids Alkaloid substances derived from the seeds and bark of plants of the genus *Strychnos* (family Loganiaceae). This genus serves as the source of poisonous, nitrogen-containing plant materials, such as strychnine (see structure; R = H). The seeds of the Asian species of *Strychnos* contain 2–3% alkaloids, of which about half is strychnine

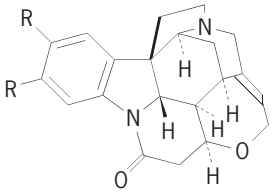

and the rest is closely related materials; for example, brucine (see structure; R = OCH_3) is a more highly oxygenated relative. Strychnine and brucine are isolated by extraction of basified plant residue with chloroform and then, from the chloroform solution, by dilute sulfuric acid. Precipitation from the dilute acid is accomplished with ammonium hydroxide. Strychnine is separated from brucine by fractional crystallization from ethanol. *See* CRYSTALLIZATION; STRYCHNOS.

At one time strychnine was used as a tonic and a central nervous system stimulant, but because of its high toxicity (5 mg/kg is a lethal dose in the rat) and the availability of more effective substances, it no longer has a place in human medicine. [D.D.]

Strychnos A genus of tropical trees and shrubs belonging to the Logania family (Loganiaceae). *Strychnos nux-vomica*, a native of India and Ceylon, is the source of strychnine. The alkaloid, strychnine, has been used medicinally in the treatment of certain nervous disorders and paralysis. Curare, used by the Indians to poison arrows, is obtained from *S. toxifera* and *S. castelnaei* in Guiana and Amazonas and from *S. tieute* in the Sunda Islands. Curare paralyzes the motor nerve endings in striated muscles and is used in medical practice in cases in which a state of extreme muscular relaxation or even immobility is desirable. It has become an important drug in the field of anesthesiology. *See* GENTIANALES; STRYCHNINE ALKALOIDS. [P.D.St.; E.L.C.]

Sturgeon Any of 10 species of large fish which comprise the family Acipenseridae in the order Acipenseriformes. These fish are found in North Temperate Zone waters, where they are almost exclusively bottom-living and feed on organisms such as mollusks, worms, and larvae. The body has five rows of bony plates, of which one is situated dorsally, two laterally, and two ventrally. The snout is elongate and there are four barbels on its lower surface; the mouth is ventrally located, and in the adult

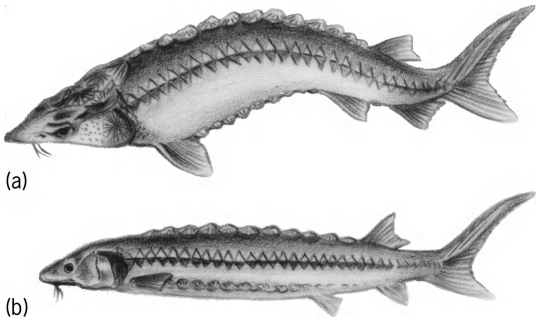

Two species of sturgeons found in United States coastal waters. (*a*) Atlantic sturgeon (*Acipenser oxyrhynchus*). (*b*) Short-nosed sturgeon (*A. brevirostrus*).

the jaws lack teeth (see illustration). The skeleton is mainly cartilaginous.

Sturgeons are known primarily for the roe, which is processed as caviar. A single female can produce millions of eggs, which may be removed from dead fish or stripped from living fish. The smoked flesh of the sturgeon is a delicacy. *See* ACIPENSERIFORMES. [C.B.C.]

Stylasterina An order of the class Hydrozoa of the phylum Coelenterata, including several brightly colored branching or encrusting "corals" of warm seas. (True corals belong to a different class, the Anthozoa). The calcareous skeleton is covered by living tissue and is penetrated by ramifying tubes. Nutritive polyps, the gastrozooids, lie in cups on one surface or along certain edges of the skeletal substance. A spine, or style, at the base of each cup gives the order its name.

Some authorities combine the Stylasterina with the Milleporina in a single order, the Hydrocorallina. *See* ANTHOZOA; HYDROZOA. [S.Cr.]

Stylolites Irregular surfaces occurring in certain rocks, mostly parallel to bedding planes, in which small toothlike projections on one side of the surface fit into cavities of like shape on the other side (see illustration). Stylolites are most common

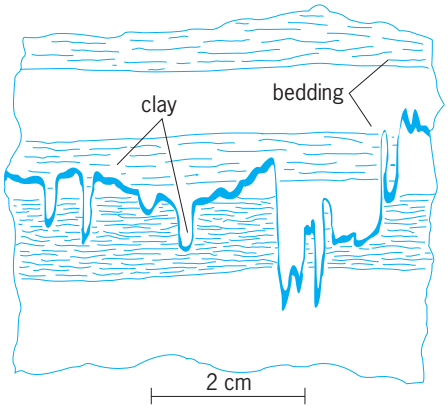

Stylolite in limestone.

in limestones and dolomites but are also present in many other kinds of rock, including sandstones, gypsum beds, and cherts. *See* DOLOMITE; LIMESTONE; SEDIMENTARY ROCKS. [R.Si.]

Stylommatophora A superorder of the molluscan subclass Pulmonata containing about 20,500 species that are grouped into 56 families. Nearly all land snails without an operculum are stylommatophorans. They have eyespots on the tips of a pair of retractile tentacles, hermaphroditic reproduction with partial fusion of the male and female system, and, normally, a second pair of retractile tentacles that function as chemoreceptors. All are air-breathing and most are truly terrestrial. Three orders, based on excretory structures, are recognized. The more primitive Orthurethra and Mesurethra have less efficient water conservation devices than do the more specialized members of the order Sigmurethra, in which a closed ureter functions in water conservation. The latter specialization apparently is a prerequisite for evolution toward a sluglike structure, since all 16 families with slugs or sluglike taxa belong to the Sigmurethra. *See* PULMONATA. [G.A.S.]

Styrene A colorless, liquid hydrocarbon with the formula $C_6H_5CH{=}CH_2$. It boils at 145.2°C (293.4°F) and freezes at −30.6°C (−23.1°F). The ethylenic linkage of styrene readily undergoes addition reactions and under the influence of light, heat, or catalysts undergoes self-addition or polymerization to yield polystyrene.

The majority of the styrene used is converted into polystyrene, but other thermoplastic or even thermosetting resins are prepared from styrene by copolymerization with suitable comonomers. A smaller quantity of styrene goes into the manufacture of elastomers or synthetic rubbers.

Styrene is a skin irritant. Prolonged breathing of air containing more than 400 ppm of styrene vapor may be injurious to health. *See* POLYMERIZATION; POLYSTYRENE RESIN. [C.K.B.]

Subduction zones Regions where portions of the Earth's tectonic plates are diving beneath other plates, into the Earth's interior. Subduction zones are defined by deep oceanic trenches, lines of volcanoes parallel to the trenches, and zones of large earthquakes that extend from the trenches landward.

Plate tectonic theory recognizes that the Earth's surface is composed of a mosaic of interacting lithospheric plates, with the lithosphere consisting of the crust (continental or oceanic) and associated underlying mantle, for a total thickness of about 100 km (60 mi). Oceanic lithosphere is created by sea-floor spreading at mid-ocean ridges (divergent, or accretionary, plate boundaries) and destroyed at subduction zones (at convergent, or destructive, plate boundaries). At subduction zones, the oceanic lithosphere dives beneath another plate, which may be either oceanic or continental. Part of the material on the subducted plate is recycled back to the surface (by being scraped off the subducting plate and accreted to the overriding plate, or by melting and rising as magma), and the remainder is mixed back into the Earth's deeper mantle. This process balances the creation of lithosphere that occurs at the mid-ocean ridge system. The convergence of two plates occurs at rates of 1–10 cm/yr (0.4–4 in./yr) or 10–100 km (6–60 mi) per million years (see illustration).

During subduction, stress and phase changes in the upper part of the cold descending plate produce large earthquakes in the upper portion of the plate, in a narrow band called the Wadati-Benioff zone that can extend as deep as 700 km (420 mi). The plate is heated as it descends, and the resulting release of water leads to melting of the overlying mantle. This melt rises to produce the linear volcanic chains that are one of the most striking features of subduction zones. *See* LITHOSPHERE.

Subduction zones can be divided in two ways, based either on the nature of the crust in the overriding plate or on the age of the subducting plate. The first classification yields two broad categories: those beneath an oceanic plate, as in the Mariana or Tonga trenches, and those beneath a continental plate, as along the west coast of South America (see illustration). The first type is known as an intraoceanic convergent margin, and the second is known as an Andean-type convergent margin. *See* MID-OCEANIC RIDGE; PLATE TECTONICS.

Active volcanoes are highly visible features of subduction zones. The volcanoes that have developed above subduction zones in East Asia, Australasia, and the western Americas surround the Pacific Ocean in a so-called Ring of Fire. At

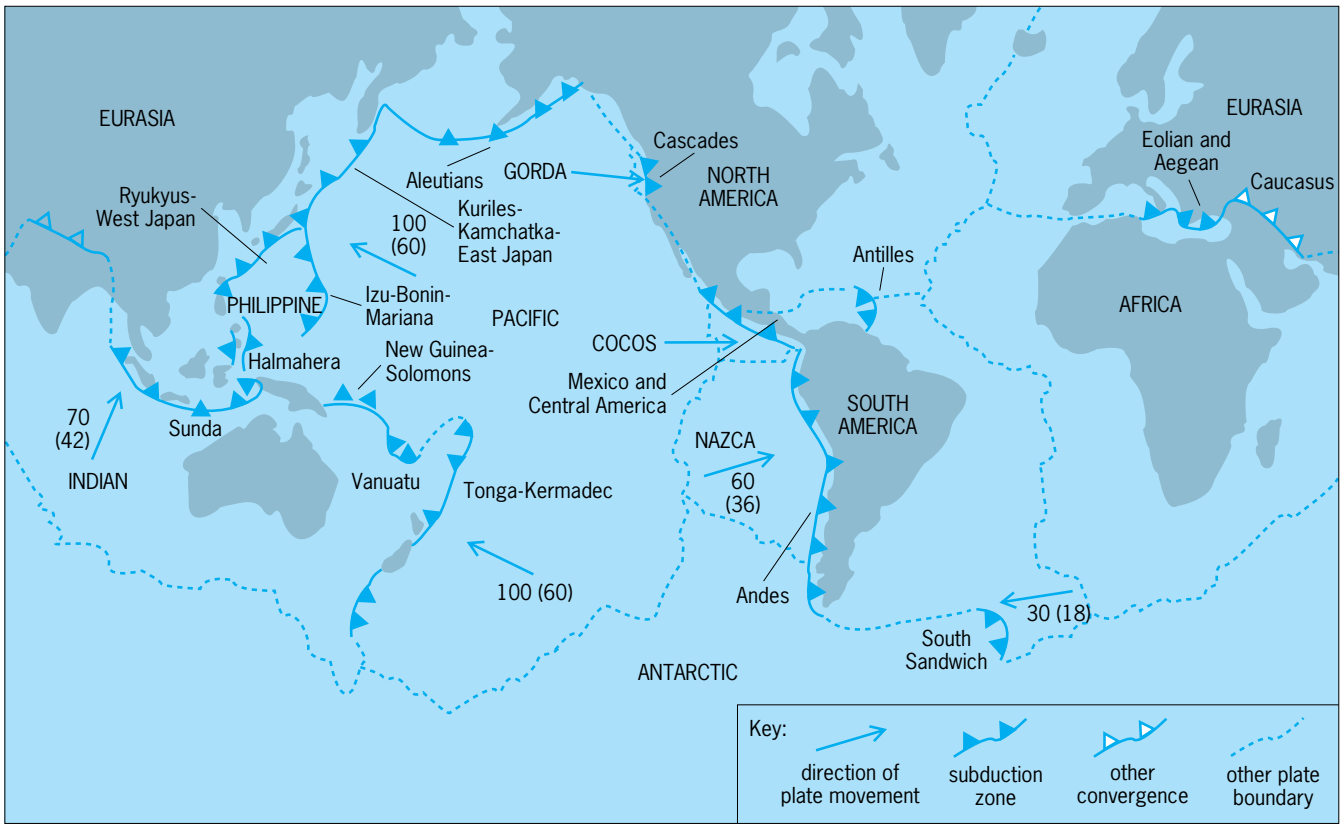

Principal subduction zones. Plate names are in all-capitals. Numbers at arrows indicate velocity of plate movement in kilometers (miles) per million years. (*After J. Gill, Orogenic Andesites and Plate Tectonics, Springer-Verlag, 1981*)

intraoceanic convergent margins, volcanoes may be the only component above sea level, leading to the name "island arc." The more general term "volcanic arc" refers to volcanoes built on either oceanic or continental crust. *See* VOLCANO.

An eventual consequence of subduction is orogeny, or mountain building. Subduction zones are constantly building new crust by the production of volcanic material or the accretion of oceanic sediments. However, the development of the greatest mountain ranges—the Alps or the Himalayas—occurs not during "normal" subduction but during the death of a subduction zone, when it becomes clogged with a large continent or volcanic arc. *See* OROGENY. [R.J.Ste.; S.H.B.]

Subgiant star An evolving star of luminosity class IV. Such a star is brighter than the main-sequence dwarfs and fainter than the true giants in its spectral class, lying between the two on the Hertzsprung-Russell diagram. The classic subgiants fall in a small region from class F to K (with effective temperatures ranging from 7000 to 4000 K or 12,000 to 7000°F). In class G they lie at absolute visual magnitude +3 with luminosities about five times the solar luminosity. Classic subgiants have masses around 1.3 times that of the Sun and violate the mass-luminosity relation as too bright for their masses. The concept is extended to the hot part of the Hertzsprung-Russell diagram from class F to O, in which the distinctions between subgiants and neighboring dwarfs and giants are much smaller, only a magnitude or less.

Main-sequence dwarfs run on the fusion of hydrogen to helium in their cores. The subgiant stage begins when the core hydrogen mass fraction drops to around 0.1, and then continues as the fraction goes to zero and the star evolves toward, but not onto, the red-giant branch with a contracting helium core. *See* DWARF STAR; GIANT STAR; HERTZSPRUNG-RUSSELL DIAGRAM; SPECTRAL TYPE; STAR; STELLAR EVOLUTION. [J.B.Ka.]

Sublimation The process by which solids are transformed directly to the vapor state without passing through the liquid phase. Sublimation is of considerable importance in the purification of certain substances such as iodine, naphthalene, and sulfur.

Sublimation is a universal phenomenon exhibited by all solids at temperatures below their triple points. For example, it is a common experience to observe the disappearance of snow from the ground even though the temperature is below the freezing point and liquid water is never present. The rate of disappearance is low, of course, because the vapor pressure of ice is low below its triple point. Sublimation is a scientifically and technically useful phenomenon, therefore, only when the vapor pressure of the solid phase is high enough for the rate of vaporization to be rapid. *See* PHASE EQUILIBRIUM; TRIPLE POINT; VAPOR PRESSURE. [N.H.N.]

Submarine A ship which can operate completely submerged in the water. The term formerly applied to any ship capable of operating completely underwater, but now usually describes a ship built for military purposes. The term "submersible" usually is applied to small, underwater vehicles that are built for research, rescue, commercial work, or pleasure.

By the end of World War II, antisubmarine warfare had progressed significantly by exploiting the limited underwater endurance and speed of the diesel-electric designs of that era. The application of nuclear power to submarines after World War II reestablished the near-invulnerability of the submarine to antisubmarine warfare from surface ships and aircraft. Nuclear power depends on nuclear fission rather than the oxidation of fossil fuels and thus requires no oxygen source as do diesel engines, allowing the submarine to operate submerged for very long periods. However, advances in submarine technology and

nonnuclear propulsion cause the nonnuclear submarine to remain highly attractive to the navies of many nations. *See* SHIP NUCLEAR PROPULSION.

Submarines can be classified by their primary military missions. Attack submarines are fast, long-range ships equipped with torpedo tubes or cruise missile launch tubes. They carry sensitive underwater sound receivers and transmitters (sonar) used to detect enemy submarines. They may be armed with torpedoes of various kinds, cruise missiles, mines, and equipment for deployment of small units of clandestine troops.

Ballistic-missile submarines carry long-range missiles fitted with nuclear warheads that can be launched while submerged. The submarine can remain submerged and undetected for many days and, on command, launch missiles on any target within range. The missiles are stowed in and launched from vertical tubes. *See* MISSILE.

Experimental submarines are occasionally built to test new designs of hull shape, deeper depth capability, power plants, or controls.

Submersibles are usually small, deep-diving vehicles. Their use is for exploration and study of the ocean depths, development of equipment, rescue, or commercial work. Some designs take advantage of the forces of gravity and buoyancy for vertical motion. Other designs use vertically oriented propellers to propel the craft up and down. Movement is restricted to short distances and slow speed because of small size and small battery capacity. *See* UNDERWATER VEHICLE.

Compared with surface ships, the submarine has features that enable it to submerge and resist great sea pressure. Submarines have a pressure hull and a nonpressure hull. The pressure hull is the watertight, pressure-proof envelope in which equipment operates and the officers and crew live. In certain areas of the submarine there is a nonpressure hull of lighter structure, forming the main ballast tanks. A nonwatertight superstructure provides a smooth, fair envelope to cover pipes, valves, and fittings on top of the hull. Above the superstructure the fairwater similarly encloses the bridge, the periscope, and multiple mast supports. *See* PERISCOPE; SHIP DESIGN.

The principal means of detecting the presence of a submerged submarine is to listen for sounds which may have been generated on board or by its movement through the water. Very small amounts of acoustic energy can be detected by sophisticated sonars. Therefore, modern submarines are designed with multiple features to greatly reduce the amount of noise they generate. [M.S.F.; J.H.W.]

Submarine cable A cable, primarily for communications, laid on the ocean floor to provide international links. This term is sometimes applied to power cables in water, but these are not usually of great length.

Coaxial cables with 1840 circuits (each having 3-kHz bandwidth) are used extensively for shorter cables. Cables using digital transmission over optical fibers (lightguide transmission) offer a significant cost-per-channel advantage over coaxial systems. Voice signals are transmitted by laser-generated digital light pulses over single-mode fiber pairs. *See* COAXIAL CABLE; COMMUNICATIONS CABLE; MULTIPLEXING AND MULTIPLE ACCESS; TRANSMISSION LINES.

SL280 digital lightguide transatlantic submarine cables utilize semiconductor laser diodes operating at a wavelength of 1.3-micrometers and carries 280 megabits per second on six hair-thin fibers, of which two pairs are active and the third is available as standby. Each fiber is a glass strand 125 μm in diameter, plastic-coated to 250 μm, and embedded in an elastomer that is part of the unit fiber structure. The fibers are helically wrapped around a central kingwire for strength during manufacture and an outer sheath of nylon to ensure dimensional stability. This structure is then helically wrapped with steel armor wires for strength, ensheathed in a copper conductor, and coated with a final protective layer of polyethylene insulation (see illustration).

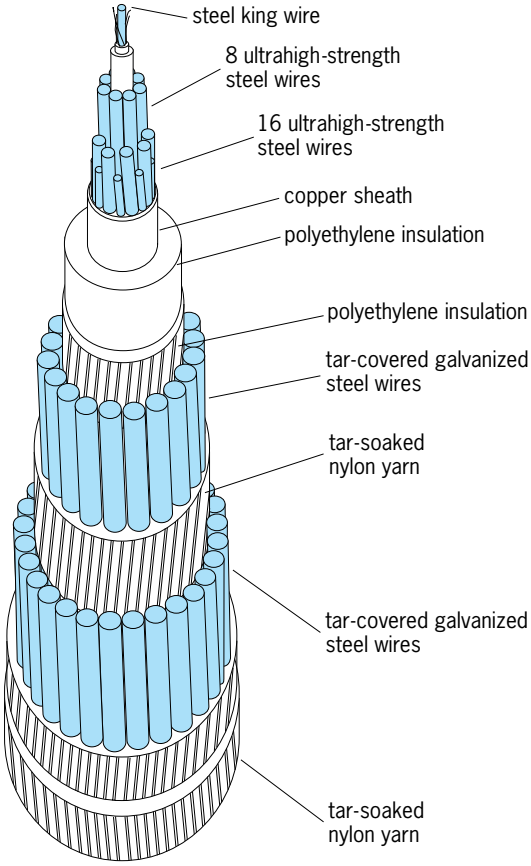

Cutaway view of double-armored SL cable.

The fiber-optic digital repeaters regenerate the signal in its original form, unlike the earlier analog repeaters which were signal amplifiers. The repeaters detect the optical signal entering the repeater, transform it into an electrical signal and amplify it, regenerate it by putting it through a high-speed clocked decision circuit, and then convert the regenerated electrical signal into an optical one by using a 1.3-μm laser transmitter. Transmission rates of 560 megabits per second are provided in the subsequent SL560 design, utilizing lasers operating at 1.5 μm.

Virtually all analog (coaxial) submarine cable systems were configured as a single line from one landing point to another. Traffic requirements, however, are not identical for all countries, and it was realized to be advantageous to implement a tree configuration, where the various branches go to different countries. Digital fiber-optic technology makes branching possible by dedicating certain fiber pairs to certain subsystems.

With the increased volume and capacity of fiber-optic systems, customers and system owners cannot afford or support the potential loss of business and revenue associated with a cable failure. To reduce this risk, many systems employ a ring-network design architecture. These systems use a self-healing loop network with built-in backup and redundancy, which provide fiber-on-fiber restoration and connectivity. In the unlikely event of a cable failure of any nature, communications traffic can be shifted from one fiber cable to another so that digital voice, data, or video communications can continue without disruption. *See* LASER; OPTICAL COMMUNICATIONS; OPTICAL FIBERS; PULSE MODULATION. [G.J.S.]

Submarine canyon A steep-sided valley developed on the sea floor of the continental slope. Submarine canyons serve as major conduits for sediment transport from land and the continental shelf to the deep sea. Modern canyons are relatively

narrow, deeply incised, steeply walled, often sinuous valleys with predominantly V-shaped cross sections. Most canyons originate near the continental-shelf break, and generally extend to the base of the continental slope. Canyons often occur off the mouths of large rivers, such as the Hudson River or the Mississippi River, although many other canyons, such as the Bering Canyon in the southern Bering Sea, are developed along structural trends. *See* CONTINENTAL MARGIN.

Modern submarine canyons vary considerably in their dimensions. The average lengths of canyons has been estimated to be about 34 mi (55 km); although the Bering Canyon is more than 680 mi (1100 km) long and is the world's longest submarine canyon. The shortest canyons are those off the Hawaiian Islands, and average about 6 mi (10 km) in length. Submarine canyons are characterized by relatively steep gradients. The average slope of canyon floors is 309 ft/mi (58 m/km). In general, shorter canyons tend to have higher gradients. For example, shorter canyons of the Hawaiian group have an average gradient of 766 ft/mi (144 m/km), whereas the Bering Canyon has a slope of only 42 ft/mi (7.9 m/km).

In comparison to modern canyons, dimensions of ancient canyons are considerably smaller. Deposits of ancient canyons are good hydrocarbon reservoirs. This is because submarine canyons and channels are often filled with sand that has the potential to hold oil and gas. Examples of hydrocarbon-bearing canyon-channel reservoirs are present in the subsurface in California, Louisiana, and Texas.

Physical and biological processes that are common in submarine canyons are mass wasting, turbidity currents, bottom currents, and bioerosion, mass wasting and turbidity currents being the more important. Mass wasting is a general term used for the failure, dislodgement, and downslope movement of sediment under the influence of gravity. Common examples of mass wasting are slides, slumps, and debris flows. Major slumping events can lead to formation of submarine canyons. The Mississippi Canyon in the Gulf of Mexico is believed to have been formed by retrogressive slumping during the late Pleistocene fall in sea level and has been partially infilled during the Holocene rise in sea level. Turbidity currents are one of the most important erosional and depositional processes in submarine canyons. There is considerable evidence to suggest that turbidity currents flow at velocities of 11–110 in./s (28–280 cm/s) in submarine canyons. Therefore, turbidity currents play a major role in the erosion of canyons. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; MARINE GEOLOGY; MARINE SEDIMENTS; MASS WASTING; TURBIDITY CURRENT.

[G.Sh.]

Submillimeter astronomy Investigation of the universe by probing the electromagnetic spectrum at wavelengths from approximately 0.3 to 1.0 millimeter: the submillimeter waveband. This waveband is bounded at longer wavelengths by the millimeter waveband (1–10 mm), and at shorter wavelengths by the far-infrared waveband (20–300 micrometers). Astronomical objects with temperatures between about 10 kelvins and several hundred kelvins, typically in the interstellar medium of galaxies, emit radiation strongly in the submillimeter waveband. *See* ELECTROMAGNETIC RADIATION.

The submillimeter waveband is one of the last parts of the electromagnetic spectrum to be investigated. This is due both to the technical challenge of making sensitive detectors, which must be held at temperatures close to absolute zero, and to the effects of the Earth's atmosphere. Atmospheric molecules, primarily water vapor, both absorb (dim) the signal from astronomical sources and emit their own radiation that acts to mask the astronomical signals. The effects are most severe at the shortest wavelengths, and only high mountain sites and air or space-borne platforms can be used for submillimeter astronomy. From high mountains, observations are possible in only three atmospheric windows, at wavelengths of about 0.35, 0.45, and 0.85 mm. Radiation of 0.85 mm penetrates down to about 2 km (1.2 mi) above sea level; while on the driest nights, only about one-half of photons with wavelengths of 0.35 or 0.45 μm that enter the upper atmosphere can be detected at the 4200-m (14,000-ft) summit of Mauna Kea, Hawaii.

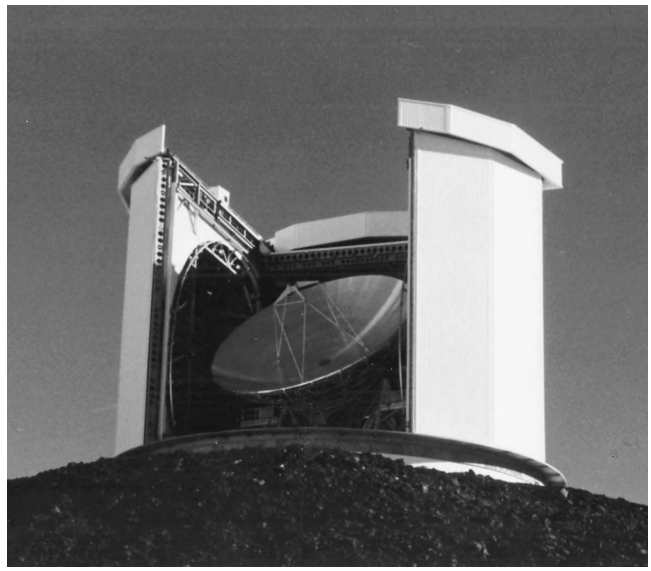

James Clerk Maxwell Telescope, 49-ft-diameter (15-m) submillimeter telescope situated at Mauna Kea, Hawaii.

Several submillimeter telescopes operate on high mountain tops, including the 15-m (49-ft) James Clerk Maxwell Telescope (JCMT; see illustration) and the 10.4-m (34-ft) Caltech Submillimeter Observatory (CSO) telescope, both on Mauna Kea, and the 15-m (49-ft) Swedish-ESO Submillimeter Telescope (SEST) in Chile. These telescopes look like satellite antennas, but their surfaces are smooth and precisely shaped to an accuracy of order 10 μm. These ground-based telescopes are subject to more atmospheric absorption than airborne and space telescopes; but because they have larger apertures, they can collect more radiation, boosting their sensitivity. *See* TELESCOPE.

In 2005 the Stratospheric Observatory for Infrared Astronomy (SOFIA), a 2.5-m-aperture (8.2-ft) telescope in a 747 aircraft, was expected to be operating. Like the former Kuiper Airborne Observatory (KAO), SOFIA will fly at up to 14 km (9 mi) altitude. Stratospheric balloons, flying at altitudes of about 40 km (25 mi), can also be used to carry submillimeter telescopes; an example is the BOOMERANG cosmic microwave background experiment reported in 2000. *See* INFRARED ASTRONOMY.

The signals from several submillimeter telescopes can be combined together in an interferometer to provide a telescope with resolution as fine as one with an aperture as large as the greatest separation between the individual telescopes. Examples of this type of telescope are the Smithsonian Millimeter Array (SMA) on Mauna Kea, and the United States-Europe-Japan Atacama Large Millimeter Array (ALMA) under development in Chile. *See* RADIO TELESCOPE.

Submillimeter observations are used by astronomers with a wide range of interests to study matter in the universe at temperatures of about 10–100 K and the cosmic microwave background radiation. Interstellar chemistry in clouds of gas and dust, the formation of protostars embedded in stellar nurseries within the Milky Way, the properties of the interstellar medium in even the most distant galaxies, and the evolution of galaxies are all studied in the submillimeter waveband. Submillimeter observations are especially important for observations of regions that are rich in dust and gas. These are opaque to visible and infrared light, but submillimeter light can still escape, allowing these regions to be studied.

[A.Bl.]

Submillimeter-wave technology The generation, detection, and application of radiation in the submillimeter-wave region of the electromagnetic spectrum. In practical terms, this region corresponds to wavelengths between about 30 micrometers and 3 millimeters (see illustration). At its extremes the region overlaps with the microwave and infrared, and a unique feature is that the distinct propagation methods of these adjacent regions, guided-wave and free-space propagation, are used throughout it. *See* ELECTROMAGNETIC RADIATION; INFRARED RADIATION; MICROWAVE.

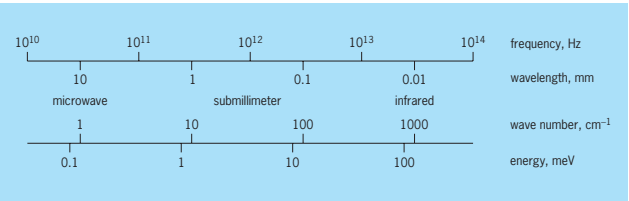

Electromagnetic spectrum in and near the submillimeter-wave region, showing its frequency, wavelength, wave number, and photon energy equivalents.

Widespread use of the submillimeter-wave region did not occur until two revolutionary developments in the late 1950s and early 1960s: (1) the practical realization of Fourier transform spectroscopy, which led to significant improvements in the quality of broad band submillimeter-wave spectroscopy; and (2) the discovery of two reasonably intense submillimeter-wave laser sources, the hydrogen cyanide and the water-vapor electrically excited lasers. *See* INFRARED SPECTROSCOPY; LASER.

Radiation sources. The main application of incoherent radiation sources is as broad band sources in Fourier transform spectrometers. The two most commonly used sources are the mercury-vapor lamp and the globar. *See* MERCURY-VAPOR LAMP; MICROWAVE TUBE.

There are several categories of coherent submillimeter-wave radiation sources: (1) nonrelativistic electron-beam tubes such as klystrons, extended interaction oscillators, orotrons, ledatrons, and carcinotrons or backward-wave oscillators, which oscillate with varying amounts of tunability and output power; (2) relativistic electron-beam devices such as gyrotrons, free-electron lasers, and synchrotrons; and (3) solid-state devices such as Gunn and IMPATT oscillators. With the exception of the free-electron laser, these are long-wave devices. The final category, gas lasers, covers the entire submillimeter-wave region. *See* GYROTRON; KLYSTRON; MICROWAVE SOLID-STATE DEVICES; SYNCHROTRON RADIATION.

Detectors. Most submillimeter-wave detectors are either thermal, rectifying, or photon detectors. The first respond to heat generated by absorbed radiation, the second by rectification of radiation-field-induced currents, the third by absorption of individual photons. They are commonly used in direct and heterodyne detection systems. Thermal detectors include room-temperature thermopile, Golay cell, and pyroelectric detectors, while more sensitive semiconductor bolometers operate at liquid-helium temperatures. Rectifying devices include metal-semiconductor and metal-metal point contacts and Schottky barrier diodes, all room-temperature devices, and Josephson-type devices operated at about 4.2 K (−452°F). Submillimeter-wave photon detectors are based on semiconductors and rely on the absorption of a photon changing the free-carrier density or mobility and hence the conductivity of the detector. *See* BOLOMETER; PYROELECTRICITY; RADIOMETRY; SUPERCONDUCTING DEVICES.

Atmospheric transparency. Between the visible and microwave regions the atmosphere is effectively opaque, with the exception of two windows in the near-infrared and three in the submillimeter-wave region, at 94, 140, and 220 GHz. The significant feature of the submillimeter-wave windows is that the levels of attenuation involved are not as seriously increased by adverse atmospheric conditions as those of the near-infrared. This makes them attractive for many applications. The attenuation associated with these windows is, nonetheless, quite high and means that ground-level low-frequency submillimeter-wave systems have a relatively limited range. This can be advantageous for some applications.

Applications. In the past, applications of radiation in the submillimeter-wave region were distributed evenly over the region and largely concerned research activities. Now, however, as a result of developments toward integrated, inexpensive transmitter and receiver systems, applications include many radar- and telecommunications-related activities centered on the 94-GHz atmospheric window. These applications are mainly military, but spin-offs into the civil area are expected.

Photon energies in the submillimeter-wave region correspond to activation energies of many physical, chemical, and bioloical phenomena, and the spectroscopic techniques of this region have been widely applied to their study. Other spectroscopic applications have concerned the development of radiometric methods to meet the requirements of fusion plasma diagnostics and astronomy. *See* SUBMILLIMETER ASTRONOMY.

The submillimeter-wave region has been extensively used for fundamental frequency metrology, relating the frequencies of submillimeter-wave gas lasers directly back to the primary frequency standard. Important consequences have been the definition of the speed of light as a fixed quantity and the redefinition of the meter as a unit derived from it. *See* LIGHT; PHYSICAL MEASUREMENT.

The submillimeter-wave region impinges on fusion plasma research in two ways. First, in machines such as tokamaks it provides a way of studying the spatial and temporal evolution of the electron temperature. The second application is electron cyclotron heating, in which electron heating by microwave radiation at the electron cyclotron frequency has been demonstrated. High-power gyrotron oscillators are generally used as the radiation sources. *See* NUCLEAR FUSION; PLASMA (PHYSICS); PLASMA DIAGNOSTICS.

In the study and applications of fast electromagnetic pulse propagation, a number of novel techniques for the submillimeter-wave region have been developed. The generation and detection of such pulses can be done with ultrashort optical pulses driving photoconductive switches, or by all-electronic systems based on nonlinear transmission lines. The submillimeter-wave radiation used in these systems is variously known as T waves or terahertz waves. *See* OPTICAL PULSES.

[J.R.Bi.]

Subsonic flight Movement of a vehicle through the atmosphere at a speed appreciably below that of sound waves. Subsonic flight extends from zero (hovering) to a speed approximately 85% of sonic speed corresponding to the ambient temperature. At higher vehicle velocities the local velocity of air passing over the vehicle surface may exceed sonic speed, and true subsonic flight no longer exists.

Vehicle type may range from a small helicopter, which operates at all times in the lower range of the velocity scale, to an intercontinental ballistic missile, which is operative throughout this and other velocity regimes, but is in subsonic flight for only a few seconds. The design of each is affected by the same principles of subsonic aerodynamics. Subsonic flow of a fluid such as air may be subdivided into a range of velocities in which the flow may be considered incompressible without appreciable error (below a velocity of approximately 300 mi/h or 135 m/s), and a higher range in which the compressible nature of the fluid becomes significant. In both cases the viscosity of the fluid is important. The theories which apply to compressible, inviscid fluids may be used almost without modification in some low-subsonic problems, and in other cases the results offered by these theories may be modified to account for the effects of viscosity and

compressibility. *See* COMPRESSIBLE FLOW; TRANSONIC FLIGHT; VISCOSITY.

A typical subsonic wing cross section (airfoil) has a rounded front portion (leading edge) and a sharp rear portion (trailing edge). Air approaching the leading edge comes to rest at some point on the leading edge, with flow above this point proceeding around the upper airfoil surface to the trailing edge, and flow below passing along the lower surface to the same point, where the flow again theoretically has zero velocity. The two points of zero local velocity are known as stagnation points. If the path from front to rear stagnation point is longer along the upper surface than along the lower surface, the mean velocity of flow along the upper surface must be greater than that along the lower surface. Thus, in accordance with the principle of conservation of energy, the mean static pressure must be less on the upper surface than on the lower surface. This pressure difference, applied to the surface area with proper regard to force direction, gives a net lifting force. Lift is defined as a force perpendicular to the direction of fluid flow relative to the body, or more clearly, perpendicular to the free-stream velocity vector. *See* AIRFOIL; BERNOULLI'S THEOREM.

The wing, as the lifting device, whether fixed, as in the airplane, or rotating, as in the helicopter, is probably the most important aerodynamic part of an aircraft. However, stability and control characteristics of the subsonic airplane depend on the complete structure. Control is the ability of the airplane to rotate about any of the three mutually perpendicular axes meeting at its center of gravity. Static stability is the tendency of the airplane to return to its original flight attitude when disturbed by a moment about any of the axes. [J.E.Ma.]

Substitution reaction One of a class of chemical reactions in which one atom or group (of atoms) replaces another atom or group in the structure of a molecule or ion. Usually, the new group takes the same structural position that was occupied by the group replaced.

Substitution reactions involve the attack of a reagent, which is the source of the new atom or group, on the substrate, the molecule or ion in which the replacement occurs. They involve the formation of a new bond and the breaking of an old bond. Substitution reactions are classified according to the nature of the reagent (electrophilic, nucleophilic, or radical) and according to the nature of the site of substitution (saturated carbon atom or aromatic carbon atom). *See* ELECTROPHILIC AND NUCLEOPHILIC REAGENTS.

Systematic names for substitution reactions are composed of the parts: name of group introduced + de + name of group replaced + ation, with suitable elision or change of vowels for euphony. Thus, the replacement of bromine by a methoxy group is called methoxydebromination. *See* ORGANIC REACTION MECHANISM. [J.F.B.]

Subsynchronous resonance The resonance between a series-capacitor-compensated electric system and the mechanical spring-mass system of a turbine-generator at subsynchronous frequencies. Beginning about 1950, series capacitors were installed in long alternating-current transmission lines [250 km (150 mi) or more] to cancel part of the inherent inductive reactance of the line. Until 1971, up to 70% of the 60-Hz inductive reactance was canceled by series capacitors in some long lines with little concern for side effects. (If 70% of a line's inductive reactance is canceled, the line is said to have 70% series compensation.) In 1970, and again in 1971, a turbine-generator at the Mohave Power Plant in southern Nevada experienced shaft damage that required several months of repairs on each occasion, following switching events that placed the turbine-generator radial on a series-compensated transmission line. The shaft damage was due to torsional oscillations between the two ends of the generator-exciter shaft. Shortly after the second event, it was determined that the torsional oscillations were caused by torsional interaction, which is a type of subsynchronous resonance. There have been no reported occurrences of two other types of subsynchronous resonance, the induction generator effect and torque amplification. *See* ALTERNATING-CURRENT CIRCUIT THEORY; ALTERNATING-CURRENT GENERATOR; TRANSMISSION LINES; TURBINE.

It has been clearly established that subsynchronous resonance can be controlled with the use of countermeasures, thus making it possible to benefit from the distinct advantages of series capacitors. About a dozen countermeasures have been successfully applied such that there has been no reported subsynchronous resonance event since 1971. Subsynchronous resonance countermeasures protect turbine-generator shafts from harmful torsional oscillations by one of two methods. First, the turbine-generator can be tripped when a subsynchronous resonance condition is detected. This limits the number of torsional oscillations experienced by the turbine-generator shafts. This type of countermeasure is relatively inexpensive but is not acceptable if the anticipated subsynchronous resonance conditions are expected to occur frequently. Generally, such a countermeasure will not be applied as the sole subsynchronous resonance protection if it is expected to cause a turbine-generator to be tripped more than once every 10 years. Three types of tripping countermeasures are applied: torsional motion relay, armature current relay, and the generator tripping logic scheme. The second method of protection does not involve turbine-generator tripping, but eliminates or limits harmful torsional oscillations. *See* ELECTRIC POWER SYSTEMS. [R.G.F.]

Subtraction One of the four fundamental operations of arithmetic and algebra. Subtraction is often regarded as an operation inverse to addition, that is, if a and b are numbers, the number $a - b$ is defined as that number which added to b gives a. The more modern viewpoint eliminates subtraction completely by considering the number $a - b$ as the sum of a and that number (denoted by $-b$) which added to b gives 0. The number symbolized by $-b$ is called the inverse of b (with respect to addition). Every real number has a unique inverse (the number 0 is its own inverse). In this sense "subtraction" may be performed on objects of many different kinds, and the original numerical operation greatly extended. *See* ADDITION; ALGEBRA; DIVISION; MULTIPLICATION; NUMBER THEORY. [L.M.Bl.]

Subway engineering The branch of transportation engineering that deals with feasibility study, planning, design, construction, and operation of subway (underground railway) systems. In addition to providing rapid and comfortable service, subways consume less energy per passenger carried in comparison with other modes of transportation such as automobiles and buses. They have been adopted in many cities as a primary mode of transportation to reduce traffic congestion and air pollution.

Subways are designed for short trips with frequent stops, compared to above-ground, intercity railways. Many factors considered in the planning process of subway systems are quite similar to those for railway systems. Subway system planning starts with a corridor study, which includes a forecast of ridership and revenues, an estimation of construction and operational costs, and a projection of the potential benefits from land development. *See* RAILROAD ENGINEERING; TRANSPORTATION ENGINEERING.

All subway systems have three major types of structures: stations, tunnels, and depots. The most important task in planning a new subway system or a new subway line is to locate stations and depots and to determine the track alignment. Subway lines are normally located within the right-of-way of public roads and as far away as possible from private properties and sites of importance. Because stations and entrances are usually located in densely populated areas, land acquisition is often a major problem. One solution is to integrate entrances into nearby developments such as parks, department stores, and public

buildings, which lessens the visual impact of the entrances and reduces their impediment to pedestrian flow.

Design of the permanent works includes structural and architectural elements and electrical and mechanical facilities. There are two types of structures: stations and tunnels. For stations, space optimization and passenger flow are important. The major elements in a typical station are rails, platform, staircases, and escalators. For handicapped passengers, provisions should be made for the movement of wheelchairs in elevators and at fare gates, and special tiles should be available to guide the blind to platforms.

In both stations and tunnels, ventilation is essential for the comfort of the passengers and for removing smoke during a fire. Sufficient staircases are required for passengers to escape from the station platform to a point of safety in case of a fire. The electrical and mechanical facilities include the rolling stock, signaling, communication, power supply, automated fare collection, and environmental control (air-conditioning) systems. Corrosion has caused problems to structures in some subways; therefore, corrosion-resistant coatings may be required. To minimize noise and vibration from running trains, floating slabs can be used under rails or building foundations in sections of routes crossing densely populated areas and in commercial districts where vibration and secondary airborne noise inside buildings are unacceptable. *See* VENTILATION.

Underground stations are normally constructed by using an open-cut method. For open cuts in soft ground, the sides of the pits are normally retained by wall members and braced using struts. The pits are fitted with decks for maintaining traffic at the surface. For new lines that pass under existing lines, it is not possible to have open cuts. In such cases, stations have to be constructed using mining methods (underground excavation). *See* TUNNEL.

Many modern subway systems are fully automated and require only a minimal staff. Train movements are monitored and regulated by computers in a control center. Therefore, engineering is limited to the function and maintenance of the electrical and mechanical facilities. The electrical and mechanical devices requiring constant care include the rolling stock, signaling, communication and broadcasting systems, power supply, elevators and escalators, automated fare collection, and environmental control systems. Also included are depot facilities, and station and tunnel service facilities. *See* ELECTRIC DISTRIBUTION SYSTEMS; RAILROAD CONTROL SYSTEMS. [Z.C.M.; R.N.Hw.]

Sucrose An oligosaccharide, α-D-glucopyranosyl-β-D-fructofuranoside, also known as saccharose, cane sugar, or beet sugar. The structure is shown below. Sucrose is very soluble in

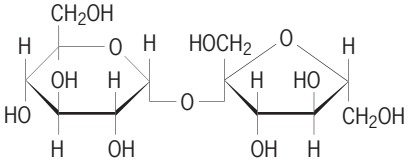

water and crystallizes from the medium in the anhydrous form. The sugar occurs universally throughout the plant kingdom in fruits, seeds, flowers, and roots of plants. Honey consists principally of sucrose and its hydrolysis products. Sugarcane and sugarbeets are the chief sources for the preparation of sucrose on a large scale. Another source of commercial interest is the sap of maple trees. *See* OLIGOSACCHARIDE; SUGARBEET; SUGARCANE. [W.Z.H.]

Suctoria A small specialized subclass of the protozoan class Ciliatea whose members were long considered entirely separate from the "true" ciliates. The sole order of this subclass is Suctorida. These forms show a number of highly specialized features. Most conspicuous are their tentacles, often numerous, which serve as mouths. These multiple organelles of ingestion

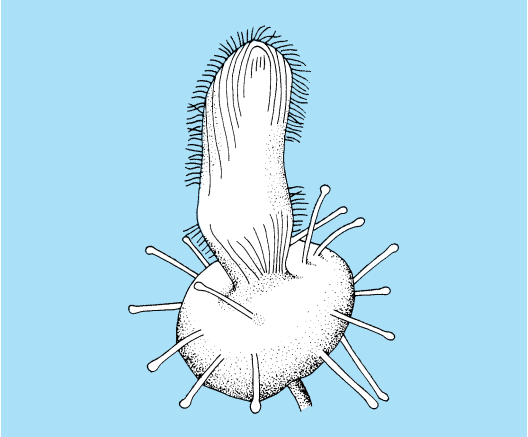

Endogenous budding in the suctorian *Podophrya*, a species which measure 10–28 micrometers.

fasten to the pellicle of prey organisms, generally passing ciliates. By forces not entirely understood, the tentacles are used to suck out the prey's protoplasm to provide sustenance for the suctorian. Nearly all species are stalked, and the sedentary, mature forms are devoid of any external ciliature. Young larval forms are produced by both endogenous and exogenous budding. These forms bear locomotor cilia and serve, as in the case of species of the Peritrichia, for dissemination (see illustration). *See* CILIATEA; PERITRICHIA. [J.O.C.]

Sudangrass An annual, warm-season grass of tropical origin said to have been grown in Egypt since early times, though its value was first recognized in Sudan only in 1909. In that same year it was introduced into the United States as a replacement for johnsongrass, which had become a noxious weed in many southern states. Sudangrass (*Sorghum bicolor* var. *sudanense*, also called *S. sudanense* and *S. vulgare* var. *sudanense*) and its hybrids are commonly used as pasture, greenchop, silage, or hay. They fill an important need in many regions of the United States, because they produce high-quality forage for cattle and sheep during the summer, when other pasture is in short supply or of low quality. Many of the varieties and hybrids produce forage until frost. *See* GRASS. [M.R.G.]

Sudden infant death syndrome (SIDS) The sudden and unexpected death of an apparently normal infant that remains unexplained after an adequate autopsy. Of a group of apparently healthy infants dying suddenly and unexpectedly, 15% will usually manifest pathologic evidence of a disease process which is sufficient to explain the death. The remaining 85% are unexplained and are classified as SIDS. In spite of the probable heterogeneity of diseases in SIDS cases, the consistent and distinctive characteristics of these infant deaths support the notion that many, if not the majority, represent a single disease process.

The incidence of SIDS in the United States is about 2.0 cases per 1000 live births, which makes SIDS the leading cause of death between the ages of 1 month and 1 year. Most SIDS deaths occur at 2–4 months of age, and about 90% occur by 6 months. SIDS is more common in males, prematurely born infants, multiple births, and the economically disadvantaged. SIDS is also increased in infants of teen-age or smoking mothers and in infants who have a history of a severe apparent life-threatening event, usually accompanied by marked cyanosis or pallor and limpness, and absence of breathing. SIDS also occurs more frequently during winter months. The rate among Native Americans is greater than among Blacks, which is greater than among Caucasians; Asians have the lowest rate. While there is slight familial clustering of SIDS, there is probably not a genetic predisposition to SIDS.

The cause of SIDS is unknown; leading hypotheses include respiratory, cardiac, and metabolic mechanisms. Much attention has been focused on the "apnea hypothesis," implicating a primary respiratory arrest due to chronic or transient insufficiency or irregularity of breathing. An imbalance between sympathetic and parasympathetic influences on cardiac activity, leading to potentially fatal cardiac arrhythmias, is a popular cardiac hypothesis.

While there is still no proof that SIDS can be prevented, electronic cardiorespiratory monitors have been prescribed for many infants in high-risk categories for SIDS. Home monitors are recommended only for infants at very high risk for SIDS. *See* CONGENITAL ANOMALIES; HUMAN GENETICS. [A.St.; K.Wi.; J.G.Br.]

Sugar Usually, sucrose, the common sugar of commerce. This sugar is a disaccharide, $C_{12}H_{22}O_{11}$, which is split, as shown in the reaction below, by hydrolysis into two monosaccharides,

$$\underset{\text{Sucrose}}{C_{12}H_{22}O_{11}} + \underset{\text{Water}}{H_2O} \rightarrow \underset{\text{Glucose}}{C_6H_{12}O_6} + \underset{\text{Fructose}}{C_6H_{12}O_6}$$

or simple sugars: glucose (dextrose) and fructose (levulose). Sucrose rotates the plane of polarized light to the right, as does glucose, but fructose is so strongly levorotatory that it overcomes the effect of glucose. Thus mixtures of equal amounts of glucose and fructose are levorotatory. The hydrolytic reaction is called inversion of sugar, and the product is invert sugar or, simply, invert. *See* CARBOHYDRATE; OPTICAL ACTIVITY.

Sucrose is widely distributed in nature, having been found in all green plants that have been carefully examined for its presence. The total quantity of all sugars formed each year has been estimated at a colossal 4×10^{11} tons (3.6×10^{11} metric tons). In spite of its availability in all green plants, sucrose is obtained commercially in substantial amounts from only two plants: sugarcane, which supplies about 56% of the world total, and the sugarbeet, which provides 44%.

Cane sugar manufacture. The manufacture of cane sugar is usually done in two series of operations. First, raw sugar of about 98% purity is produced at a location adjacent to the cane fields. The raw sugar is then shipped to refineries, where a purity that is close to 100% is achieved.

Raw cane sugar. The production of raw cane sugar begins with growing the cane in tropical or subtropical areas. The cane is harvested after a season varying from 7 months in subtropical areas to 12–22 months in the tropics. The cane stalks are harvested either mechanically or by heavy hand knives. The trend is toward mechanization. The stalks are transported to a mill by oxcart, rail, or truck. *See* SUGARCANE.

At the mill the stalks are crushed and macerated between heavy grooved iron rollers while being sprayed countercurrently with water to dilute the residual juice. The expressed juice contains 95% or more of the sucrose present. The fibrous residue, or bagasse, is usually burned under the boilers, although increasing amounts are being made into paper, insulating board, and hardboard as well as furfural, which is a chemical intermediate for the synthesis of furan and tetrahydrofuran.

Cane juice. The cane juice is treated with lime to bring its pH to about 8.2. This pH prevents the inversion reaction, which is favored by heat and acid and would lower the yield of crystallizable sugar. The juice is then heated to facilitate the precipitation of impurities, which are removed by continuous filtration.

The purified juice is concentrated by multiple-stage vacuum evaporation (usually four or five stages) and, when sufficiently concentrated, is boiled to grain or seeded with sucrose crystals in a single-stage vacuum pan. Usually three successive crops of crystals are grown, cooled, and centrifuged. The final mother liquor, which is resistant to further crystallization, is called blackstrap molasses. It is used principally as a feed for cattle. Relatively small amounts are still fermented to produce industrial alcohol and rum.

Raw cane sugar refining. The refining of raw sugar begins with dissolution of the molasses which remains in a thin film on the crystals in spite of the centrifugation. This step, called affination, brings the purity from about 98 to about 99%. The crystals are dissolved in hot water and percolated through bone char columns to remove color by adsorption. The sucrose is finally concentrated by vacuum evaporation, crystallized by seeding, centrifuged, and dried.

A major step forward has been the use of bone char in a continuous countercurrent manner (Grosvenor patent). The bone char is washed, dried, and burned to remove impurities; it is reused until it wears out mechanically and is discarded as fines. Even the fines have value as fertilizer because of their high calcium phosphate content.

Bone char has been replaced in some refineries by granular carbon derived from coal or wood waste. It is used in columns and regenerated by a process similar to the one used for bone char, which unlike granular carbon is primarily calcium phosphate. Granular carbon is frequently used in combination with decolorizing resins to reduce the color of process liquors coming from the carbon columns. Pulverized activated vegetable carbon is used in a few instances to remove color from raw sugar liquors, but its use is declining. Ion-exchange resins are used to remove ash from raw sugar liquors, especially when the liquor is to be converted directly into refined liquid sugar without going through the crystallization step first. *See* ION EXCHANGE.

Beet sugar. In the United States, sugarbeets are grown under contract by farmers from seed supplied by a beet sugar company. Because sugarbeets, like the other temperate-zone crops, thrive best under crop rotation, they are not well adapted to one-crop agriculture. *See* SUGARBEET.

Beet processing. When beets are delivered to a factory, they are washed and sliced, and the slices are extracted countercurrently with hot water to remove the sucrose. The resulting solution is purified by repeatedly precipitating calcium carbonate, calcium sulfite, or both, in it. Colloidal impurities are entangled in the growing crystals of precipitate and removed by continuous filtration. The resulting solution is nearly colorless, and the sucrose is concentrated by multiple-effect vacuum evaporation. The syrup is seeded, cooled, and centrifuged, and the beet sugar crystals are washed with water and dried.

Beet molasses differs from cane molasses in having a much lower content of invert sugar. It is, therefore, relatively stable to the action of alkali and, in the United States, is usually treated with calcium oxide to yield a precipitate of calcium sucrate. This is a mixture of loose chemical aggregates of sucrose and calcium oxide which are relatively insoluble in water. The precipitate is filtered, washed, and added to the incoming crude sugar syrup, where it furnishes calcium for the precipitations of calcium carbonate and sulfite referred to above, which remove impurities. Carbon dioxide in the form of flue gas is the other reagent, and sulfur dioxide from burning sulfur is used to produce sulfite. Ion exclusion processes have been developed for recovery of sucrose from beet molasses.

Beet residue use. The beet tops and extracted slices, as well as the molasses, are valuable as feeds. More feed for cattle and other ruminants can be produced per acre-year from beets than from any other crop widely grown in the United States. This is independent of the food energy in the crystallized sucrose, which exceeds that available from any other temperate-zone plant. It is for these reasons that the densely populated countries of Europe have expanded their beet sugar production, in spite of the ready availability of cane sugar from the tropics.

The increased use of nitrogenous fertilizer has resulted in augmenting the protein content of the molasses and other beet byproducts.

Nutritional value. Sucrose has, in the past, been attacked by some nutritionists on the ground that it provides only "empty calories," without protein, minerals, or vitamins. This argument lost much of its force when it was shown that all the vitamins

and minerals recommended by the National Research Council can be obtained by consuming any of a great variety of foods in amounts that yield a total of only one-half of the caloric requirements of an average person. The wide use by the public of vitamin supplements has caused some nutritionists to express an opposite worry over excessive vitamin consumption.

In addition to the charges that it provides empty calories, sugar has been blamed for such health problems as diabetes mellitus, coronary heart disease, dental caries, and obesity. These charges and the scientific evidence associated with them have been extensively reviewed by a special committee of the Federation of American Societies for Experimental Biology. A report of the committee findings, published by the Food and Drug Administration in 1976, includes the following: Sucrose is a contributor to the formation of dental caries when used at levels and in the manner prevailing in 1976; otherwise, there is no clear evidence in the available information that demonstrates a hazard to the public when sucrose is used at those levels.

Other sources of sugars. In Hawaii, the pineapple industry recovers both sucrose and citric acid from the rinds. The residue is fed to cattle. The total quantity of sucrose obtained from waste fruits is statistically negligible. On the other hand, a substantial fraction of the total sucrose consumed is that naturally present in a large number of fruits, vegetables, and nuts. *See* SUCROSE; SUGAR CROPS.

Lactose. Cow's milk, on a dry basis, is about 38% lactose or milk sugar. When milk is converted into cheese, the lactose remains in the whey, from which it may easily be isolated and purified. Lactose is a disaccharide which is split by hydrolysis into glucose and galactose. Lactose is about one-tenth as soluble in water as sucrose and it is one-sixth to one-half as sweet, depending on the concentration. Some of the major uses for lactose are as a direct-compression vehicle for the manufacture of pharmaceutical tablets and as a diluent for pharmaceuticals and synthetic sweeteners. *See* CHEESE; LACTOSE.

Starch. Starches can be hydrolyzed either by dilute acid or by enzymes. The product of acid hydrolysis varies with time and conditions but contains glucose, maltose, maltotriose, maltotetrose, and other sugars up to the dextrins. Only glucose, a monosaccharide, is readily isolated. Crystallized as the monohydrate, it is used in foods as a sweetener. *See* CORN; GLUCOSE.

Syrups high in maltose, a disaccharide, can be obtained by the action of amylases on starch. This hydrolysis has been of great importance for thousands of years in splitting starches for alcoholic fermentation. As yet, there is no large-scale production of pure maltose. *See* MALTOSE.

A major type of syrup was developed from starch hydrolyzates by the mid-1980s. An enzyme that isomerizes dextrose (glucose) was commercially developed to manufacture hydrolyzed starch syrups (corn syrups) containing varying amounts of fructose (levulose). The enzyme isomerizes dextrose into fructose theoretically to a 50-50 mixture of dextrose and fructose. In practice, the isomerization is allowed to progress until a syrup containing (on a solids basis) approximately 50% dextrose, 42% fructose, and 8% higher saccharide is achieved. Syrups containing more than 42% fructose (solid basis) can be made by subjecting the 42% fructose syrup to an ion-exclusion process that separates dextrose and fructose. The fructose-rich fraction from the separation is converted into a finished product, and the dextrose-rich fraction is returned to the isomerization process. Commercially produced syrups contain 42% or 55% fructose. The combined processes are also used to manufacture commercial fructose. The high-fructose corn syrups are competitive with sucrose and invert sugar in a variety of food products. The most widely used application for 55% high-fructose corn syrup is in soft drinks, displacing sucrose and invert sugar. *See* FRUCTOSE; STARCH.

Maple sugar. Before America was settled by Europeans, Indians were collecting and concentrating the juice of the hard maple (*Acer saccharum*), making maple syrup. The practice was copied by the new settlers, and production of maple syrup has been an industry ever since in the regions where hard maples are common, principally the northeastern United States.

The maple flavor does not exist in the sap but is developed by heating it. By additional heating at about 250°F (120°C), a flavor four or five times more intense can be developed. Maple syrup so produced is of special value for adding flavor to the less expensive products of the sucrose industry. Maple sugar is sucrose of about 95–98% purity; the characteristic flavor makes up only a small percent. Fairly satisfactory imitation maple flavors are available.

Honey. Honey is a form of relatively pure invert sugar dissolved in water to form a concentrated solution. However, honey also contains flavors derived from the nectar of the flowers from which it was obtained by the bee. Nutritionally, it is nearly equivalent in invert sugar but contains an excess of fructose over glucose. The sucrose found in the flowers is inverted by the enzyme honey invertase. Tupelo honey is remarkable in containing about twice as much fructose as glucose, and hence has little tendency to deposit glucose crystals.

The ready availability of the food energy in honey was known to athletes in ancient Greece. Only in recent times has it been discovered that, paradoxically, the energy of sucrose is still more quickly available. The flavors of various honeys run a wide gamut. That from the Mount Hymettus region, which is flavored by wild thyme, has been known and treasured since the poems of Homer.

Molasses. Virtually all molasses is distributed in the form of a concentrated viscous solution, but it can be reduced to a powder by means of spray-drying. It can then be handled without an investment by the consumer in tanks, pipes, and pumps. Unfortunately, later contact with moisture converts the dried molasses to a gummy mass. The availability of vaporproof bags (for example, those lined with polyethylene) has provided one solution to this problem. There are also various additives which, when mixed with molasses, reduce its tendency to pick up moisture. So far, dried molasses has made little headway against the practice of handling concentrated solutions.

Syrups. Syrups are relatively concentrated, somewhat viscous, solutions of various sugars, frequently in admixture to hinder crystallization. Large amounts of sucrose have been distributed in the form of high-purity syrups. These syrups consist of sucrose, invert sugar, or both. In some instances syrups include mixtures with dextrose and various corn syrups. The so-called liquid sugars are important for several reasons: the requirement for the use of syrups in a number of food processes (conversion of granular sucrose to a syrup was eliminated at the point of use); economy of handling, being moved with pumps, pipes, tank trucks, and tank cars; the high degree of sanitation inherent in the storage and distribution of sugar within the plant in a closed system, including the elimination of dust from the dumping of bags; and the availability of a combination of sugars with differing characteristics in a single product.

The quantity of sucrose distributed in syrup form has been reduced drastically because of the inroads made by high-fructose corn syrups, which have a composition somewhat comparable to that of invert sugar. These products are available only as syrups. They do not have taste characteristics identical to sucrose and invert sugar syrups, but lower prices catalyze their substitution for sucrose.

It is the aim of the sugar refiner and the beet processor to eliminate color and all flavors other than sweetness. The manufacture of table syrups, which are widely used on waffles and pancakes, aims at a broader spectrum of flavor. Corn syrup, which is somewhat lacking in sweetness, ordinarily has about 15% sucrose added. The high viscosity of the corn syrup, resulting from the content of dextrins, tends to hinder crystallization and is an advantage in the manufacture of certain candies.

Some sugar refiners manufacture so-called refiners' syrups for use in the manufacture of foods and table syrups. In addition to these syrups, a group of products generally referred to as cane

syrup and edible molasses are manufactured from cane juice. Cane syrup is concentrated whole cane juice, while edible molasses is concentrated cane juice from which some sugar has been removed by crystallization. These syrups and molasses contain both sucrose and invert sugar, and have a relatively dark color and a distinctive flavor. They also contain many of the nonsugars found in sugarcane juice. They are used principally by food processors, but substantial amounts are packaged for home use. Sorghum syrup is made by extracting juice from the stalk of sweet sorghum and evaporating the juice to syrup consistency. It is a dark, pungently flavored product that is not widely distributed. *See* FOOD ENGINEERING; FOOD MANUFACTURING; SUGAR CROPS. [H.B.H; Ch.B.B.]

Sugarbeet The plant *Beta vulgaris*, developed in modern times to fill the need for a sugar crop that could be grown in temperate climates. The sugarbeet was the source of only 5% of the world's commercial sugar in 1840, but by 1890 it supplied almost 50%, and since about 1920 this has dropped to about 40% with 60% derived from sugarcane. European countries produce most of their sugar from sugarbeet, with some countries having surplus sugar for export. *See* SUGARCANE.

The sugarbeet, fodder beet, garden beet, and leaf beet (Swiss chard) are cultivars of *B. vulgaris* and are genetically related. In addition to *B. vulgaris*, which includes all cultivated forms of beet, there are 12 species of wild beet in the Mediterranean and Middle Eastern regions. These wild beets are of interest in sugarbeet improvement as sources of disease resistance. Most varieties of sugarbeet in the United States are hybrid. [D.Ste.]

Sugarcane *Saccharum officinarum*, a member of the grass family. This crop originated in New Guinea about 15,000–8000 B.C., and was later moved by primitive peoples westward into Southeast Asia and India and eastward into Polynesia. Most current commercial varieties are interspecific hybrids involving primarily two or more of the following species: *S. officinarum*, *S. robustum*, and *S. spontaneum*.

Sugar is generally removed from sugarcane in large factories by either milling or diffusion. Milling crushes the stalks as they pass between a series of large metal rollers separating the fiber (bagasse) from the juice laden with sugar. The diffusion process separates sugar from finely cut stalks by dissolving it in hot water or hot juice. Processing of the juice is completed by clarification to remove nonsugar components, by evaporation to remove water, and by removal of molasses in high-speed centrifuges to produce centrifugal sugar. Bagasse and molasses are the principal byproducts from processing sugarcane. [R.E.Co.]

Sulfide phase equilibria The chemistry of the sulfides is rather simple inasmuch as most sulfide minerals involve only two major elements, although some involve three, and a few four. Understanding of the phase relations among these minerals is important both to the geologist, whose task it is to locate and exploit ore deposits, and to the metallurgist, whose task it is to extract the metals from the ores for industrial use.

Sulfide ore deposits are the most important sources of numerous metals such as lead, zinc, copper, nickel, cobalt, and molybdenum. In addition, sulfide ores provide substantial amounts of noble metals such as platinum, gold, and silver, and of other industrially important elements such as cadmium, rhenium, and selenium. Although iron sulfides usually are the most common minerals in such deposits, most commercial iron is mined from iron oxide ores and most sulfur from elemental sulfur deposits.

Some of the more common sulfide minerals are listed in the table. The first eight sulfide minerals listed may be plotted in the ternary system copper-iron-sulfur. Similarly, the first two and the last two minerals may be plotted in the ternary system iron-nickel-sulfur. Thus, by studying in detail the phase equilibria in two ternary systems a great deal of information can be obtained about many of the common sulfides occurring in ore deposits.

Common sulfide minerals

Mineral name	Chemical formula
Pyrite	FeS_2
Pyrrhotites (three or more varieties)	$Fe_{1-x}S$
Covellite	CuS
Digenite	Cu_9S_5
Chalcocite	Cu_2S
Bornite	Cu_5FeS_4
Chalcopyrite	$CuFeS_2$
Cubanite	$CuFe_2S_3$
Galena	PbS
Sphalerite and wurtzite	ZnS
Metacinnabar and cinnabar	HgS
Argentite and acanthite	Ag_2S
Molybdenite	MoS_2
Millerite	NiS
Pentlandite	$(Fe,Ni)_9S_8$

However, before the ternary systems can be explored in a systematic way, the binary systems bounding the ternaries have to be studied in detail. Similarly, it is necessary that the four bounding ternary systems be fully understood before a quaternary system can be systematically investigated. Thus, it is seen that before a systematic study of, for example, the immensely important quaternary system copper-iron-nickel-sulfur can proceed, much preliminary information is required. The prerequisite data include complete knowledge of the four ternary systems Cu-Fe-S, Cu-Fe-Ni, Fe-Ni-S, and Cu-Ni-S. In turn, the phase relations in these ternary systems cannot be systematically studied before the six binary systems Fe-Cu, Fe-Ni, Cu-Ni, Fe-S, Cu-S, and Ni-S have been thoroughly explored. *See* PHASE EQUILIBRIUM.

The enormous differences in the vapor pressures over the different phases occurring in sulfide systems add complications to the diagrammatic representation. For instance, in the Fe-S system the vapor pressure over pure iron is about 10^{-25} atm (10^{-20} pascal) at 450°C (840°F), whereas that over pure sulfur is a little more than 1 atm (10^5 pascals) at the same temperature. A complete diagrammatic representation of the relations in such a system, therefore, requires coordinates for composition and temperature, as well as for pressure. In a two-component system, such as the Fe-S system, such a representation is feasible because only three coordinates are necessary. However, in ternary (where such diagrams involve four-dimensional space) and in multicomponent systems, this type of diagrammatic representation is not possible. For this reason it is customary to use composition and temperature coordinates only for the diagrammatic representation of sulfide systems. The relations as shown in such diagrams in reality represent a projection from composition-temperature-pressure space onto a two-dimensional composition-temperature plane or onto a three-dimensional prism, depending upon whether the system contains two or three components.

Pyrite or pyrrhotite, or both, occur almost ubiquitously, not only in ore deposits but in nearly all kinds of rocks. Of the binary systems mentioned above, therefore, the iron-sulfur system is of the most importance to the economic geologist. *See* PYRITE. [G.Kul.]

Sulfonamide One of the group of organosulfur compounds, RSO_2NH_2. Many sulfonamides, which are the amides of sulfonic acids, have the marked ability to halt the growth of bacteria. The therapeutic drugs of this group are known as sulfa drugs. *See* ORGANOSULFUR COMPOUND.

The antibacterial spectrum of sulfonamides comprises a wide variety of gram-positive and gram-negative bacteria, including staphylococci, streptococci, meningococci, and gonococci, as well as the gangrene, tetanus, coli, dysentery, and cholera bacilli. They have only slight activity against *Mycobacterium tuberculosis*, while certain closely related sulfones are quite active against

M. leprae. The use of sulfonmethoxine has proved to be effective in the treatment of chloroquine-resistant malaria. The relative potency of sulfonamides against the different microorganisms varies, and their action is bacteriostatic rather than bactericidal.

The antibacterial effect, both in patients and in test tube cultures, is antagonized by *p*-aminobenzoic acid (PABA) and PABA-containing natural or synthetic products, such as folic acid and procaine. Accordingly, the mode of action of sulfonamides is considered to be an antimetabolite activity, dependent upon the inhibition of enzyme systems involving the essential PABA. *See* BACILLARY DYSENTERY; CHOLERA; LEPROSY; MALARIA; MENINGOCOCCUS; STAPHYLOCOCCUS; STREPTOCOCCUS; TETANUS; TUBERCULOSIS.

Bacterial resistance has developed to all known sulfonamides, and many sulfonamide-resistant strains are encountered among the gram-positive and gram-negative bacteria. The emergence of resistance to sulfonamides, however, seems less rapid and less widespread than resistance to most antibiotics.

Sulfonamides are used today mostly as auxiliary drugs or in combination with antibiotics. In certain infectious diseases, however (for instance, in meningococcal infections and most infections of the urinary tract), sulfonamides deserve preference over antibiotics. *See* ANTIMICROBIAL AGENTS; DRUG RESISTANCE. [R.L.M.; E.Gru.]

Sulfonation and sulfation Sulfonation is a chemical reaction in which a sulfonic acid group, —SO_3H, is introduced into the structure of a molecule or ion in place of a hydrogen atom. Sulfation involves the attachment of the —OSO_2OH group to carbon, yielding an acid sulfate, $ROSO_2OH$, or of the —SO_4— group between two carbons, forming the sulfate, $ROSO_2OR$.

Sulfonation of aromatic compounds is the most important type of sulfonation. This is accomplished by treating the aromatic compound with sulfuric acid, as in the reaction below. The product of sulfonation is a sulfonic acid.

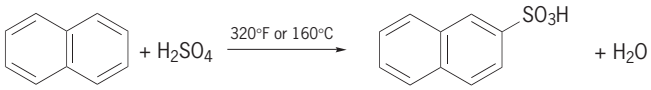

Napthalene β-Napthalenesulfonic acid

Sulfonation may also be defined as any chemical process by which the sulfonic acid group, —SO_2OH, or the corresponding salt or sulfonyl halide group, for example, —SO_2Cl, is introduced into an organic compound. These groups may be bonded to either a carbon or a nitrogen atom. The latter compounds are designated *N*-sulfonates or sulfamates.

Most sulfonates are employed as such in acid or salt form for applications where the strongly polar hydrophilic —SO_2OH group confers needed properties on a comparatively hydrophobic nonpolar organic molecule. Some sulfonates, such as methanesulfonic and toluenesulfonic acids, are used as catalysts. The major quantity of sulfonates and sulfates is both marketed and used in salt form. This category includes detergents; emulsifying, demulsifying, wetting, and solubilizing agents; lubricant additives; and rust inhibitors. *See* ORGANOSULFUR COMPOUND; SUBSTITUTION REACTION. [P.H.G./R.S.K.]

Sulfonic acid A derivative of sulfuric acid ($HOSO_2OH$) in which an OH has been replaced by a carbon group, as shown in the structure below. Sulfonic acids are strongly acidic,

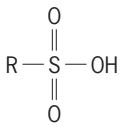

water-soluble, nonvolatile, and hygroscopic; they do not act as oxidizing agents and are typically highly stable compounds. Sulfonic acids rarely occur naturally. An exception, taurine, $NH_2CH_2CH_2SO_3H$, occurs in bile.

The aliphatic sulfonic acids are generally made by oxidation of thiols. Several have unique properties. For example, trifluoromethanesulfonic acid, CF_3SO_3H, is such a strong acid that it will protonate sulfuric acid. A compound derived from natural camphor, 10-camphorsulfonic acid, is used extensively in the optical resolution of amines.

Aromatic sulfonic acids are much more important than those of the aliphatic series. Aromatic sulfonic acids are produced by sulfonation of aromatic compounds with sulfuric acid or fuming sulfuric acid. Sulfonation of aromatic hydrocarbons is a reversible process; treatment of an aromatic sulfonic acid with superheated steam removes the —SO_3H group. This process can be used in purifying aromatic hydrocarbons. Aromatic sulfonic acids and their derivatives, especially metal salts, are important industrial chemicals. *See* SULFURIC ACID.

The most extensive use of the sulfonation reaction is in the production of detergents. The most widely used synthetic detergents are sodium salts of straight-chain alkylbenzenesulfonic acids. *See* DETERGENT.

Sulfonated polymers, particularly sulfonated polystyrenes, act as ion-exchange resins which have important applications in water softening, ion-exchange chromatography, and metal separation technology. Both sulfonated polymers and simple aromatic sulfonic acids, particularly *p*-toluenesulfonic acid, are frequently used as acid catalysts in organic reactions such as esterification and hydrolysis. *See* HOMOGENEOUS CATALYSIS; ION EXCHANGE.

The sulfonic group, in either acid or salt form, is capable of making many substances water soluble, increasing their usefulness. This application is particularly significant in the dying industry, in which a majority of the dyes are complex sodium sulfonates. Many acid-base indicators are soluble due to the presence of a sodium sulfonate moiety. Some pigments used in the paint and ink industry are insoluble metal salts or complexes of sulfonic acid derivatives. Most of the brighteners used in detergents compounded for laundering are sulfonic acid derivatives of heterocyclic compounds. *See* ACID-BASE INDICATOR; DETERGENT; DYE; INK; ORGANOSULFUR COMPOUND; PAINT. [C.R.J.]

Sulfur A chemical element, S, atomic number 16, and atomic weight 32.064. The atomic weight reflects the fact that sulfur is composed of the isotopes ^{32}S (95.1%), ^{33}S (0.74%), ^{34}S (4.2%), and ^{36}S (0.016%). The ratios of the various isotopes vary slightly but measurably according to the history of the sample. By virtue of its position in the periodic table, sulfur is classified as a main-group element. *See* PERIODIC TABLE.

The chemistry of sulfur is more complex than that of any other elemental substance, because sulfur itself exists in the largest variety of structural forms. At room temperature, all the stable forms of sulfur are molecular; that is, the individual atoms aggregate into discrete molecules, which in turn pack together to form the solid material. In contrast, other elements near sulfur in the periodic table normally exist as polymers (silicon, phosphorus, arsenic, selenium, tellurium) or as diatomic molecules (oxygen, nitrogen, chlorine). Selenium and phosphorus can exist as molecular solids, but the stable forms of these elements are polymeric.

At room temperature the most stable form of sulfur is the cyclic molecule S_8. The molecule adopts a crownlike structure, consisting of two interconnected layers of four sulfur atoms each. The S—S bond distances are 0.206 nanometer and the S—S—S bond angles are 108°. Three allotropes are known for cyclo-S_8. The most common form is orthorhombic α-sulfur, which has a density of 2.069 g/cm^3 (1.200 oz/in.^{3}) and a hardness of 2.5 on the Mohs scale. It is an excellent electrical insulator, with a room temperature conductivity of 10^{18} ohm^{-1} cm^{-1}. Sublimed sulfur and "flowers" of sulfur are generally composed of α-S_8. Sulfur is quite soluble in carbon disulfide (CS_2; 35.5/100 g or 1.23 oz/3.52 oz at 25°C or 77°F), poorly soluble in alcohols, and

practically insoluble in water. At 95.3°C (203°F), sulfur changes into the monoclinic β allotrope. This form of sulfur also consists of cyclic S_8 molecules, but it has a slightly lower density at 1.94–2.01 g/cm^3 (1.12–1.16 oz/in.^{3}). A third allotrope containing S_8 is triclinic γ-sulfur. The β and γ allotropes of sulfur slowly revert to the α form at room temperature. Crystals of sulfur are yellow and have an absorption maximum in the ultraviolet at 285 nm, which shifts to higher energy as the temperature decreases. At low temperatures, S_8 is colorless. Even at room temperature, however, finely powdered sulfur can appear to be nearly white.

The best-studied system is α-S_8, which converts to the β form at 90°C (194°F), which then melts at 120°C (248°F) to give a golden yellow liquid. If this melt is quickly recooled, it refreezes at 120°C (248°F), thus indicating that it consists primarily of S_8 molecules. If the melt is maintained longer at 120°C (248°F), then the freezing point is lowered about 5°C (9°F), indicating the formation of about 5% of other rings and some polymer. At 159.4°C (318.9°F), the melt suddenly assumes a red-brown color. Over the range 159.4–195°C (318.9–383°F), the viscosity of the melt increases 10,000-fold before gradually decreasing again. This behavior is very unusual, since the viscosity of most liquids decreases with increasing temperature. The strong temperature dependence of the viscosity is due to the polymerization and eventual depolymerization of sulfur. Polymeric sulfur retains its elastomeric character even after being cooled to room temperature. There are several polymeric forms of sulfur, but all of them revert to α-S_8 after a few hours.

Sublimination of S_8 occurs when it is maintained in a vacuum at a temperature below its melting point. It vaporizes at 444.61°C (832.30°F). Below 600°C (1110°F), the predominant species in the gas are S_8 followed by S_7 and S_6. Above 720°C (1328°F), violet S_2 is the major species.

Principal inorganic compounds. Hydrogen sulfide (H_2S) is the most important compound that contains only sulfur and hydrogen. It is a gas at room temperature with a boiling point of −61.8°C (−79.2°F) and a freezing point of −82.9°C (−117°F). The low boiling point of hydrogen sulfide is attributed to the weakness of intermolecular S···H hydrogen bonding; the O···H hydrogen bond is much stronger, as evidenced by the high boiling point of water. Gaseous hydrogen sulfide is 1.19 times more dense than air, and air-H_2S mixtures are explosive. Hydrogen sulfide has a strong odor similar to that of rotten eggs; its odor is detectable at concentrations below 1 microgram/m^3. At high concentrations, H_2S has a paralyzing effect on the olfactory system, which is very hazardous because H_2S is even more toxic than carbon monoxide (CO).

The most common compound that contains only carbon and sulfur is carbon disulfide (CS_2). Carbon disulfide molecules are linear, consisting of two sulfur atoms connected to a central carbon atom. Carbon disulfide is a toxic, highly flammable, and volatile liquid that melts at −111°C (−168°F) and boils at 46°C (115°F). Commercial carbon disulfide has a strong unpleasant odor due to impurities. It is manufactured from methane and elemental sulfur and is used for the production of carbon tetrachloride, rayon, and cellophane. Structurally related to carbon disulfide is carbonyl sulfide (SCO), which forms from carbon monoxide and elemental sulfur. The chlorination of CS_2 gives Cl_3CSCl, which can be reduced by H_2S to thiophosphene, $CSCl_2$. Thiophosgene ($CSCl_2$) [boiling point 73°C or 163°F] is a planar molecule with the carbon at the center of a triangle defined by the sulfur and two chlorine atoms. Thiocyanate, the linear anion NCS^-, is prepared by the reaction of cyanide (—CN) with elemental sulfur.

Several sulfur oxides exist, but the dioxide and trioxide are of preeminent importance. Sulfur dioxide (SO_2) is a colorless gas that boils at −10.02°C (113.97°F) and freezes at −75.46°C (−103.8°F). The density of liquid sulfur dioxide at −10°C (14°F) is 1.46 g/cm^3 (0.84 oz/in.^{3}). Liquid sulfur dioxide is an excellent solvent. The sulfur dioxide molecule is bent, with an O—S—O angle of 119°.

Sulfur trioxide (SO_3) is a planar molecule that is a liquid at room temperature that exists in equilibrium with a cyclic trimeric structure known as β-SO_3. When β-SO_3, actually S_3O_9, is treated with traces of water, it converts to either of two polymeric forms referred to as γ- and α-sulfur trioxide. These are fibrous materials, proposed to have the formula $(SO_3)_xH_2$, where x is in the thousands. Sulfur trioxide is prepared by the oxidation of sulfur dioxide, although at very high temperatures this reaction reverses. Exposure of sulfur trioxide to water yields sulfuric acid (H_2SO_4); exposure of SO_3 to sulfuric acid yields disulfuric acid ($H_2S_2O_7$). *See* SULFURIC ACID.

Chlorine and sulfur react to give a family of compounds with the general formula S_xCl_2, several members of which have been obtained in pure form. The structures of these compounds are based on an atom or chain of sulfur atoms terminated with Cl. Sulfur monochloride (S_2Cl_2), also known as sulfur monochloride, is the most widely available of the series. It is a yellow oil that boils at 138°C (280°F), and reacts with chlorine in the presence of iron(III) chloride ($FeCl_3$) catalyst to give sulfur dichloride (SC_2), which is a red volatile liquid with a boiling point of 59°C (138°F). Treatment of sulfur dichloride with sodium fluoride (NaF) gives SF_4.

Thionyl chloride ($OSCl_2$) is a colorless reactive compound with a boiling point of 76°C (169°F); it is used to convert hydroxy compounds to chlorides. Important applications include the preparation of anhydrous metal halides and alkyl halides. Sulfuryl chloride (O_2SCl_2; boiling point 69°C or 156°F) is used as a source of chlorine.

Organosulfur compounds. This family of compounds contains carbon, hydrogen, and sulfur, and it is a particularly vast area of sulfur chemistry. Thiols, also known as mercaptans, feature the linkage C—S—H. Mercaptans are foul-smelling compounds. They are the sulfur analogs of alcohols, but they are more volatile. They can be prepared by the action of hydrogen sulfide (H_2S) on olefins. Deprotonation of thiols gives thiolate anions, which form stable compounds with many heavy metals. Thiols and especially thiolates can be oxidized to form disulfides (persulfides), which have the connectivity of C—S—S—C. The organic persulfides are also related to organic polysulfides, which have chains of sulfur atoms terminated with carbon. The introduction of such mono-, di-, and polysulfide linkages is the basis of the vulcanization process, which imparts desirable mechanical properties to natural or synthetic polyolefin rubbers. This is accomplished by heating the polymer with sulfur in the presence of a zinc catalyst. *See* RUBBER.

Thioethers, also known as organic sulfides, feature the connectivity C—S—C and are often prepared from the reaction of thiolates and alkyl halides. Like mercaptans, thioethers often have strong unpleasant odors, but they are also responsible for the pleasant odors of many foods and perfumes. They are intentionally introduced at trace levels in order to impart an odor to gaseous hydrocarbon fuels. The reaction of alkyl dihalides and sodium polysulfides affords organic polysulfide polymers known as thiokols.

There are many organic sulfur oxides; prominent are sulfonic acids (RSO_3H), which are the organic derivatives of sulfuric acid. These compounds are prepared by the oxidation of thiols as well as by treatment of benzene derivatives with sulfuric acid, for example, benzene sulfonic acid. Most detergents are salts of sulfonic acids. *See* DETERGENT.

Biochemistry. Sulfur is required for life. Typical organisms contain 2% sulfur dry weight. Three amino acids contain sulfur, as do many prosthetic groups in enzymes. Some noteworthy sulfur compounds include the disulfide lipoic acid, the thioethers biotin and thiamine (vitamin B_1), and the thiol coenzyme A. Sulfide ions, S^{2-}, are found incorporated in metalloproteins and metalloenzymes such as the ferredoxins, nitrogenases, and hydrogenases. *See* AMINO ACIDS; ENZYME; PROTEIN.

Many bacterial species obtain energy by the oxidations of sulfides. Bacteria of the genus *Thiobacillus* couple the conversion of

carbon dioxide (CO_2) to carbohydrates to the aerobic oxidation of mineral sulfides to sulfuric acid. This activity can be turned to good use for leaching low-grade mineral ores. Often, however, the sulfuric acid runoff (such as in mines or sewers) has negative environmental consequences. The purple and green bacteria as well as the blue-green algae are remarkable because they are photosynthetic but anaerobic; they oxidize sulfide, not water (as do most photosynthetic organisms). Depending on the species, the sulfur produced in this energy-producing pathway can accumulate inside or outside the cell wall. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; PHOTOSYNTHESIS.

Minerals. Sulfide minerals are among the most important ores for several metals. These compounds are two- or three-dimensional polymers containing interconnected metal cations and sulfide S^{2-} or persulfido S_2^{2-} anions. In general, metal sulfides are darkly colored, often black, and they are not soluble in water. They can sometimes be decomposed by using strong acids, with liberation of hydrogen sulfide. Certain sulfides will also dissolve in the presence of excess sulfide or polysulfide ions.

Pyrites (FeS_2), also known as iron pyrites or fool's gold, are the most common sulfide minerals and can be obtained as very large crystals that have a golden luster. Sphalerite (zinc blende; ZnS) and galena (PbS) are major sources of zinc and lead. Orange cinnabar (HgS) and yellow greenockite (CdS) are the major ores for mercury and cadmium, respectively. Molybdenite (MoS_2) is the major ore of molybdenum.

The sulfur content of fossil fuels results from the sulfur in the ancient organisms as well as from subsequent incorporation of mineral sulfur into the hydrocarbon matrix. Gaseous fossil fuels are often contaminated with hydrogen sulfide, which is an increasingly important source of sulfur. Organic derivatives containing the C—S—C linkage are primarily responsible for the sulfur content of petroleum and coal. The so-called organic sulfur in petroleum can be removed by hydrodesulfurization catalysis, involving reaction with hydrogen over a molybdenum catalyst, to give hydrocarbons and hydrogen sulfide. [T.B.R.]

Sulfuric acid A strong mineral acid with the chemical formula H_2SO_4. It is a colorless, oily liquid, sometimes called oil of vitriol or vitriolic acid. The pure acid has a density of 1.834 at 25°C (77°F) and freezes at 10.5°C (50.90°F). It is an important industrial commodity, used extensively in petroleum refining and in the manufacture of fertilizers, paints, pigments, dyes, and explosives.

Sulfuric acid is produced on a large scale by two commercial processes, the contact process and the lead-chamber process. In the contact process, sulfur dioxide, SO_2, is converted to sulfur trioxide, SO_3, by reaction with oxygen in the presence of a catalyst. Sulfuric acid is produced by the reaction of the sulfur trioxide with water. The lead-chamber process depends upon the oxidation of sulfur dioxide by nitric acid in the presence of water, the reaction being carried out in large lead rooms.

Sulfuric acid reacts vigorously with water to form several hydrates. The concentrated acid, therefore, acts as an efficient drying agent, taking up moisture from the air and even abstracting the elements of water from such compounds as sugar and starch. The concentrated acid also acts as a strong oxidizing agent. It reacts with most metals upon heating to produce sulfur dioxide. *See* SULFUR. [F.J.J.]

Sum rules Formulas in quantum mechanics for transitions between energy levels, in which the sum of the transition strengths is expressed in a simple form. Sum rules are used to describe the properties of many physical systems, including solids, atoms, atomic nuclei, and nuclear constituents such as protons and neutrons. The sum rules are derived from quite general principles, and are useful in situations where the behavior of individual energy levels is too complex to describe by a precise quantum-mechanical theory. *See* ENERGY LEVEL (QUANTUM MECHANICS).

In general, sum rules are derived by using Heisenberg's quantum-mechanical algebra to construct operator equalities, which are then applied to particles or the energy levels of a system. *See* QUANTUM MECHANICS. [G.F.Be.]

Sun The star around which the Earth revolves, and the planet's source of light and heat, hence life. The Sun is a globe of gas, 1.4×10^6 km (8.65×10^5 mi) in diameter with a mass 333,000 times the Earth, held together by its own gravity. The surface temperature of the Sun is about 6000 K (10,000°F); since solids and liquids do not exist at these temperatures, the Sun is entirely gaseous. Almost all the gas is in atomic form, although a few molecules exist in the coolest surface regions, such as sunspots.

The Sun is a typical member of the spectral class dG2, stars of surface temperature 6000 K. The d stands for a dwarf, a normal star of that class. *See* SPECTRAL TYPE.

Solar structure. The interior of the Sun can be studied only by inference from the observed properties of the entire star. The great mass of the Sun presses down on the center, requiring a gas with a central density of near 90 g/cm^3 and 2×10^7 K (3.6×10^6°F) temperature to support it. At these huge temperatures and densities, nuclear reactions take place. The radiation produced flows outward till it is radiated into space by the surface (photosphere) at 6000 K (10,000°F).

Energy production. The energy of the Sun is produced by the conversion of hydrogen into helium. For each hydrogen atom converted, one neutrino is produced. These neutrinos are detected, but less than the expected number. *See* NEUTRINO; SOLAR NEUTRINOS.

The material at the center of the Sun is so dense that a few millimeters are opaque, so the photons created by nuclear reactions are continually absorbed and reemitted and thus make their way to the surface by a random walk. The atoms in the center of the Sun are entirely stripped of their electrons by the high temperatures, and most of the absorption is by continuum processes, such as the scattering of light by electrons. Because there are so many absorption and emission processes along the way, it can take as long as a million years to complete the random walk to the surface.

Convection. In the outer regions of the solar interior, the temperature is low enough for ions and even neutral atoms to form and, as a result, atomic absorption becomes very important. The high opacity makes it very difficult for the radiation to continue outward, so steep temperature gradients are established that result in convective currents. Most of the outer envelope of the Sun is in such convective equilibrium. These large-scale mass motions are responsible for the complex phenomena observed at the surface. *See* CONVECTION (HEAT).

Radiation. Electromagnetic energy is produced by the Sun in essentially all wavelengths. However, more than 95% of the energy is concentrated in the relatively narrow band between 290 and 2500 nm and is accessible to routine observation from ground stations on Earth. The maximum radiation is in the green region, and the eyes of human beings have naturally evolved to be sensitive to this range of the spectrum. The total radiation is called the solar constant. It is not exactly constant but varies slightly (±0.1%) with the solar cycle. The ultraviolet flux, however, varies by substantial factors depending on the exact wavelength, and this affects the Earth's upper atmosphere. *See* ELECTROMAGNETIC RADIATION; SOLAR CONSTANT.

Atmosphere. Although the Sun is gaseous, it can be seen only to the point at which the density is so high that the material is opaque. This layer, the visible surface of the Sun, is termed the photosphere. Light from father down reaches the Earth by repeated absorption and emission by the atoms, but the deepest layers cannot be seen directly. The surface is actually not sharp, but the Sun is so far away that the smallest distance that can be resolved with the best telescope is about 300 km (200 mi). Since

the density e-folding height (scale height) is less than 200 km (120 mi), the edge appears sharp. *See* PHOTOSPHERE.

Above the photosphere the atmosphere is transparent, and its density falls off much more slowly because magnetic fields support the ionized particles. The atmosphere can be seen by using a narrow-band filter or a spectrograph to pick out the isolated wavelengths absorbed by the atmospheric gases. In the upper photosphere it is cooler, and the lines are dark. If the light is imaged in the strongest lines, such as those of hydrogen, a region higher still is seen, called the chromosphere. The light from this region is dominated by the red hydrogen alpha (level $2 \rightarrow 3$ transition) line, which gives it a rosy color seen at a solar eclipse. The chromosphere is a rapidly fluctuating region of jets and waves coming up from the surface. When all the convected energy coming up from below reaches the surface, it is concentrated in the thin material and produces considerable activity. Where the magnetic field is stronger, these waves are absorbed, and raise the temperature to 7000–8000 K (12,000–14,000°F). The scale height of the chromosphere is 1000 km (600 mi) or more, so there no longer is a sharp edge. *See* CHROMOSPHERE; ECLIPSE.

When the Moon obscures the Sun at a total solar eclipse, the vast extended atmosphere of the Sun called the corona can be seen. The corona is a million times fainter than the photosphere, so it is visible only when seen against the dark sky of an eclipse or with very special instruments. Its density is low, but its temperature is high (more than 10^6 K or 1.8×10^6°F). The hot gas evaporating out from the corona flows steadily to the Earth and farther in the solar wind. *See* SOLAR CORONA; SOLAR WIND.

Coronal holes. Early coronal observations showed that the corona was occasionally absent over certain regions. In particular, at sunspot minimum it was quite weak over the poles. X-ray pictures revealed great bands of the solar surface essentially devoid of corona for many months. These proved to be regions where the local magnetic fields were connected to quite distant places, so the fields actually reached out to heights from which the solar wind could sweep the gas outward. Analysis of solar wind data showed that equatorial coronal holes were associated with high-velocity streams in the solar wind, and recurrent geomagnetic storms were associated with the return of these holes. Thus the relative intensity of the corona over sunspot regions is partly due to their strong, closed magnetic fields which trap the coronal gas.

Solar activity. There are a number of transient phenomena known collectively as solar activity. These are all connected with sunspots.

Sunspots. Sunspots were discovered around 1610. Heinrich Schwabe announced in 1843 that their number rose and fell with a 10-year period. Subsequent study of the old records revealed an 11-year period since the original discovery.

The number of sunspots peaks soon after the beginning of each cycle and decays to a minimum in 11 years. The first spots of a number cycle always occur at higher latitudes, between 20° and 35°, and the latitude of occurrence decreases as the cycle unfolds (Spörer's law). Almost no spots are observed outside the latitude range of 5–35°. The great majority are small and last a few days, but some last for two rotations. In 1908, George Ellery Hale discovered that sunspots had strong magnetic fields. Each spot group contains positive and negative magnetic polarity (monopoles are forbidden by Maxwell's laws). Hale found that the polarities were mirrored, with the same polarity generally leading in one hemisphere and following in the other. He found that with each new number cycle the lead polarity switches, so that the complete magnetic cycle lasts 22 years. But each new number cycle starts a few years before the end of the previous one, so the average duration of a half-cycle is nearly 14 years. *See* MAGNETISM.

The darkness of sunspots (Fig. 1) is probably due to the intense magnetic fields (3000 gauss or 0.3 tesla), which cool the surface by suppressing the normal convective energy flow from below. It takes several days for the darkening to occur.

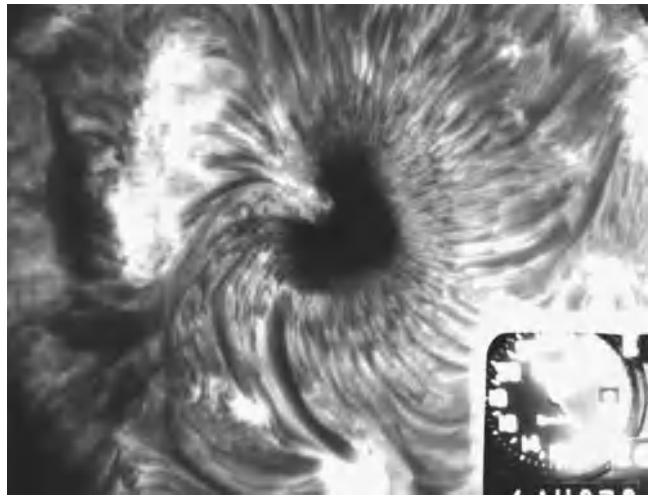

Fig. 1. Large symmetric sunspot photographed in Hα light. Clock indicates time of photograph. (*Big Bear Solar Observatory*)

Although the sunspot is cool, its neighborhood is the scene of the hottest and most intense activity, generally referred to as an active region. Magnetic energy is continually released there. The corona above an active region is hot and dense, roughly three times hotter and denser than in quiet regions.

Prominences. The term "prominence" is used for any cloud of cool gas in the corona, where it appears bright against the sky. Because these clouds absorb the chromospheric light and scatter it, they appear dark against the solar disk in Hα and other strong lines. In continuous light they are transparent. At the limb we see the chromospheric light they scatter against the dark sky. Since they are much denser than the corona, something must hold them up against gravity. Prominences are found only in regions of horizontal magnetic fields that support them. Thus filaments on the disk, which may last for weeks, are good markers of the magnetic boundaries. When the magnetic structure changes, prominences become unstable and erupt, always upward. They also may be ejected by solar flares or appear as graceful loops raining from the corona after flares. Erupting prominences are probably the source of coronal mass ejections, in which a bubble of coronal material erupts outward at several hundred kilometers per second and flows out into interplanetary space.

Plages. Just as prominences occur when the magnetic field changes from one sign to the other, plages occur whenever the magnetic field is vertical and relatively strong but not strong enough to form a sunspot. They are bright regions in any strong spectrum line, because the chromosphere is heated there. In a typical active region, the preceding magnetic field is clumped in a sunspot and the following field spread out in a plage. In Hα light, the plage is seen to be connected to the sunspot by dark fibrils outlining the lines of force.

Flares. The most spectacular activity associated with sunspots is the solar flare (Fig. 2). A flare is defined as an abrupt increase in the Hα emission from the sunspot region. The brightness of the flare may be up to eight times that of the chromosphere; the rise time is seldom longer than a few minutes. The Hα brightening results from heating of the chromosphere at the foot points of the magnetic field by a tremendous energy release in the atmosphere. While flares are usually visible only in chromospheric lines, the foot points of big flares can be seen in white light. From the foot points, a cloud of hot material, up to 3×10^7 K (5.4×10^7°F) arises and concentrates at the arch tops. This cloud condenses out in an array of loop prominences. An active sunspot

Fig. 2. The great "sea horse" flare of August 7, 1972, late in the flare, photographed in the blue wing of the Hα line. The neutral line between two bright strands is crossed by an arcade of bright loop prominences raining down from the corona. (*Big Bear Solar Observatory*)

group produces a hierarchy of flares, a few big and many small ones.

Flares are often associated with the eruption of filaments. A few minutes after the eruption begins, there is an abrupt acceleration and a storm of energetic particles is produced, heating the corona to flare brightness.

The flare produces a huge stream of solar energetic particles (SEP) as well as shock waves. A huge magnetohydrodynamic shock wave flies out at about 1000 km/s (600 mi/s) and continues into interplanetary space, often reaching the Earth. The wave produces a huge radio burst in the meter-wavelength range as it excites the coronal layers. The energetic nuclei produce gamma-ray lines from nuclear reactions as they penetrate to the photosphere. If they are sufficiently numerous, they heat the photosphere faster than it can reemit energy and a white light flare is observed, usually in the form of bright transient flashes at the foot points of the flare loops. The particles reach the Earth in a great particle storm. [H.Zi.]

Sun dog A bright spot of light that sometimes appears on either side of the Sun, the same distance above the horizon as the Sun, and separated from it by an angle of about 22° (see illustration). For higher Sun elevations, the angle increases slightly. These spots are known by many common names: sun dogs, mock suns, false suns, or the 22° parhelia. They usually show a red edge on the side closest to the Sun. On some occasions the entire spectrum of colors can be spread out in the sun-dog spot but, commonly, the red edge is followed by an orange or yellow band that merges into a diffuse white region. The effects result from the refraction by sunlight through small, flat, hexagonal-shaped ice crystals falling through the air such that their flat faces are oriented nearly horizontally. *See* HALO; METEOROLOGICAL OPTICS. [R.Gr.]

Sun dogs on either side of the Sun. Also visible are the 22° halo and the parhelic circle. (*Courtesy of Robert Greenler*)

Sundew Any plant of the genus *Drosera* (90 species) of the family Droseraceae. Sundews are small, herbaceous, insectivorous plants that grow on all the continents, especially Australia. Numerous glandular hairs (tentacles) on the leaf secrete a viscous fluid which traps a visiting insect. The tentacles then bend inward about the victim, bringing it into contact with the surface of the leaf where it is digested. The droplets secreted by the glands on the leaves glitter like dewdrops in the morning sunlight; hence, the name, sundew. *See* INSECTIVOROUS PLANTS; NEPENTHALES; SECRETORY STRUCTURES (PLANT). [P.D.St./E.I.C.]

Sundial An instrument for telling time by the Sun. It is composed of a style that casts a shadow and a dial plate, which is the surface upon which hour lines are marked and upon which the shadow falls. The style lies parallel to Earth's axis. The construction of the hour lines is based on the assumption that the apparent motion of the Sun is always on the celestial equator. The most widely used form is the horizontal dial that indicates local apparent time (Sun time). Other forms of the sundial indicate local mean time, and standard time. *See* TIME. [R.N.M.]

Sunfish A name given to certain members of the freshwater family Centrarchidae as well as to members of the marine family Molidae. There are two species of oceanic sunfishes in the genus *Mola*; both are large, stout, deep-bodied fishes with rough skin and are found in warm seas, where they are frequently seen near the surface. These fishes appear to be tailless.

The fresh-water sunfishes are grouped together with the crappies. The warmouth sunfish (*Chaenobryttus gulosus*) is the only species found in the United States, where it occurs in the Mississippi River drainage and in streams along the Atlantic coast. It differs from other sunfishes in having teeth on the tongue.

The crappies are game fish of considerable importance. Crappies are a favorite species for pond culture. They can be readily transplanted and, under favorable conditions, multiply prodigiously. *See* PERCIFORMES. [C.B.C.]

Sunflower *Helianthus annuus*, the most widely distributed of the 50 native North American species of this genus of the family Compositae. It is an extremely variable species, with two main divisions. The first involves wild weedy plants found along roadways and other recently disturbed areas; the second, domesticated plants grown in fields and gardens. *See* ASTERALES.

Within the domestic type there are two categories of plants: the ornamental, which has a few branches with larger heads than the wild, and the crop type, which has only a single stem and the largest head of all sunflowers (see illustration). Crop types are either oil or nonoil. Plant breeders have modified the plant for adaptation to modern, large-scale farming and have increased the oil content of the seeds. The present worldwide interest in growing sunflowers as a crop is due to the increased yield of the new commercially available oilseed hybrids.

Sunflowers are grown on all the continents and in many countries throughout the world. Russia is the major producer, followed by Argentina, the United States, and Canada. Sunflower oil (sunoil) is the second most important vegetable crop oil. It is

Maturing sunflower. (*U.S. Department of Agriculture*)

a high-quality oil and is high in linoleic fatty acid. The oil is used in cooking, salad dressing, mayonnaise, margarine, and soap. [B.H.B.]

Sunlamp A special form of mercury arc discharge lamp designed to produce ultraviolet radiation. These lamps also produce some radiant energy in the visible region of the spectrum, thus having a light output as well as an ultraviolet output. The lamps are principally used for producing a skin tan on the human body. The less common uses include therapeutically producing vitamin D in the body for the treatment of rickets and causing fluorescence or photochemical reactions.

Sunlamps have ultraviolet radiation at wavelengths above 280 nanometers. The lower limit is set by the fact that the quartz and high-silica glass used for lamp envelopes do not transmit below 280 nm. The two lamps designed for tanning are the FS-40 and the RS. The FS-40 tubular fluorescent lamp (illus. *a*) operates with a low-pressure mercury arc that causes a special chemical phosphor coating inside the tube to radiate ultraviolet energy. The RS sunlamp (illus. *b*) is a reflector unit containing a high-pressure mercury arc tube for generating ultraviolet energy plus a tungsten filament (similar to incandescent lamp filaments) in series with the arc tube to serve as a ballast. *See* MERCURY-VAPOR LAMP; ULTRAVIOLET LAMP. [G.R.P.]

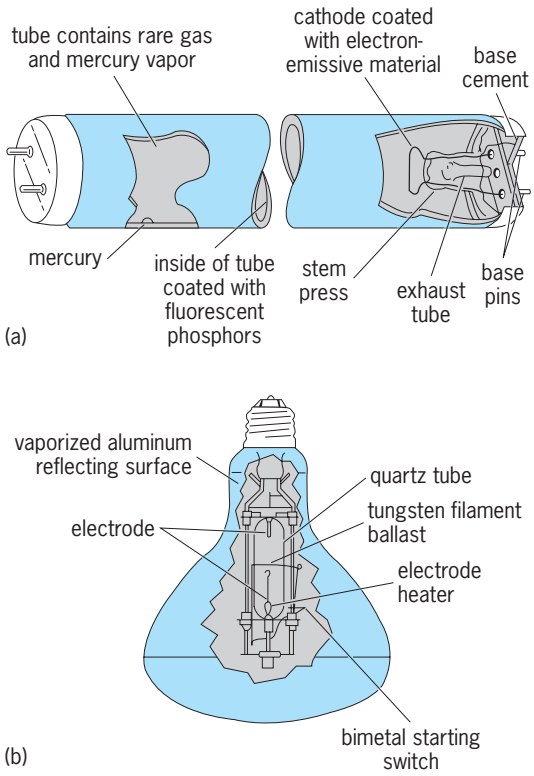

Structure of sunlamps. (*a*) FS-40. (*b*) RS.

Superacid An acid which has an extremely great proton-donating ability. It has proved convenient to define a superacid somewhat arbitrarily as an acid, or more generally, an acidic medium, which has a proton-donating ability equal to or greater than that of anhydrous (100%) sulfuric acid.

Superacids belong to the general class of proton or Brönsted acids. A proton acid is defined as any species which can act as a source of protons and which will therefore protonate a suitable base, as in reaction (1).

$$HA + B \rightleftharpoons BH^+ + A^- \tag{1}$$

The strengths of acids are often compared by measuring the extent of their ionization in water, that is, the extent to which they can protonate the base water, as in reaction (2).

$$HA + H_2O \rightleftharpoons H_3O^+ + A^- \tag{2}$$

However, all strong acids are fully ionized in dilute aqueous solution, and they therefore appear to have the same strength. Their strengths are said to be reduced or leveled to that of the hydronium ion (H_3O^+), which is the most highly acidic species that can exist in water. In any case, many of the superacids react with and are destroyed by water. For these reasons, the strengths of superacids cannot be measured by the conventional means of utilizing their aqueous solutions. The acidities of superacids can, however, be conveniently measured in terms of the Hammett acidity function. *See* ACID AND BASE.

The Hammett acidity function is a method of measuring acidity based on the determination of the ionization ratios of suitable weak bases (indicators), usually by means of the change in absorption spectrum that occurs on protonation of the base, although the nuclear magnetic resonance (NMR) spectrum has also been used. The Hammett acidity function (H_0) is defined by Eq. (3), where K_{BH+} is the dissociation constant of the acid

$$H_0 = pK_{BH+} - \log \frac{[BH^+]}{[B]} \tag{3}$$

form of the indicator and $[BH^+]/[B]$ is the ionization ratio of the indicator. Hammett acidity function (H_0) values for a number of superacids are given in the table. In each case the value refers to the 100% (anhydrous) acid. Each of the superacids in the table

Hammett acidity function values for several superacids

Superacid	Formula	$-H_0$
Sulfuric acid	H_2SO_4	11.9
Chlorosulfuric acid	HSO_3Cl	13.8
Trifluoromethane sulfonic acid	HSO_3CF_3	14.0
Disulfuric acid	$H_2S_2O_7$	14.4
Fluorosulfuric acid	HSO_3F	15.1
Hydrogen fluoride	HF	15.1

is a liquid at room temperature, and each forms the basis of a solvent system. *See* IONIC EQUILIBRIUM; SOLUTION. [R.J.Gi.]

Supercharger An air pump or blower in the intake system of an internal combustion engine. Its purpose is to increase the air-charge weight and therefore the power output from an engine of a given size. In an aircraft engine, the supercharger counteracts the power loss that results from the decrease of atmospheric pressure with increase of altitude. Various types of pumps and compressors may be used as superchargers, which are either mechanically driven by the engine crankshaft or powered by the engine exhaust gas. *See* COMPRESSOR; PUMP; TURBOCHARGER.

Some production automobile engines use a spiral-type supercharger, while others use a pressure-wave supercharger. However, automotive, marine, and stationary engines generally use a positive-displacement Roots blower driven from the engine crankshaft. *See* INTERNAL COMBUSTION ENGINE.

In a supercharged diesel engine, the increased air charge allows the engine to burn more fuel and produce greater power without creating excessive pressure inside the cylinder. Supercharging the diesel engine makes ignition of the fuel easier without requiring a fuel of better quality. *See* DIESEL ENGINE.

To enable a reciprocating aircraft engine to develop its rated sea-level power at altitude, a supercharger must be used to increase the pressure and weight of the intake air charge. Centrifugal compressors are generally used because of their relatively small size for a given capacity, and are driven either by a gear drive from the crankshaft or by the engine exhaust gas. With gear drive, a lower ratio is commonly used at low and medium altitudes, with a change to a higher ratio at high altitudes. *See* AIRCRAFT ENGINE. [J.A.B.; D.L.An.]

Supercomputer A computer which, among existing general-purpose computers at any given time, is superlative, often in several senses: highest computation rate, largest memory, or highest cost. Predominantly, the term refers to the fastest "number crunchers," that is, machines designed to perform numerical calculations at the highest speed that the latest electronic device technology and the state of the art of computer architecture allow.

The demand for the ability to execute arithmetic operations at the highest possible rate originated in computer applications areas collectively referred to as scientific computing. Large-scale numerical simulations of physical processes are often needed in fields such as physics, structural mechanics, meteorology, and aerodynamics. A common technique is to compute an approximate numerical solution to a set of partial differential equations which mathematically describe the physical process of interest but are too complex to be solved by formal mathematical methods. This solution is obtained by first superimposing a grid on a region of space, with a set of numerical values attached to each grid point. Large-scale scientific computations of this type often require hundreds of thousands of grid points with 10 or more values attached to each point, with 10 to 500 arithmetic operations necessary to compute each updated value, and hundreds of thousands of time steps over which the computation must be repeated before a steady-state solution is reached. *See* COMPUTATIONAL FLUID DYNAMICS; NUMERICAL ANALYSIS; SIMULATION.

Two lines of technological advancement have significantly contributed to what roughly amounts to a doubling of the fastest computers' speeds every year since the early 1950s—the steady improvement in electronic device technology and the accumulation of improvements in the architectural designs of digital computers.

Computers incorporate very large-scale integrated (VLSI) circuits with tens of millions of transistors per chip for both logic and memory components. A variety of types of integrated circuitry is used in contemporary supercomputers. Several use high-speed complementary metallic oxide semiconductor (CMOS) technology. Throughout most of the history of digital computing, supercomputers generally used the highest-performance switching circuitry available at the time—which was usually the most exotic and expensive. However, many supercomputers now use the conventional, inexpensive device technology of commodity microprocessors and rely on massive parallelism for their speed. *See* COMPUTER STORAGE TECHNOLOGY; CONCURRENT PROCESSING; INTEGRATED CIRCUITS; LOGIC CIRCUITS ; SEMICONDUCTOR MEMORIES.

Increases in computing speed which are purely due to the architectural structure of a computer can largely be attributed to the introduction of some form of parallelism into the machine's design: two or more operations which were performed one after the other in previous computers can now be performed simultaneously. *See* COMPUTER SYSTEMS ARCHITECTURE.

Pipelining is a technique which allows several operations to be in progress in the central processing unit at once. The first form of pipelining used was instruction pipelining. Since each instruction must have the same basic sequence of steps performed, namely instruction fetch, instruction decode, operand fetch, and execution, it is feasible to construct an instruction pipeline, where each of these steps happens at a separate stage of the pipeline. The efficiency of the instruction pipeline depends on the likelihood that the program being executed allows a steady stream of instructions to be fetched from contiguous locations in memory.

The central processing unit nearly always has a much faster cycle time than the memory. This implies that the central processing unit is capable of processing data items faster than a memory unit can provide them. Interleaved memory is an organization of memory units which at least partially relieves this problem.

Parallelism within arithmetic and logical circuitry has been introduced in several ways. Adders, multipliers, and dividers now operate in bit-parallel mode, while the earliest machines performed bit-serial arithmetic. Independently operating parallel functional units within the central processing unit can each perform an arithmetic operation such as add, multiply, or shift. Array processing is a form of parallelism in which the instruction execution portion of a central processing unit is replicated several times and connected to its own memory device as well as to a common instruction interpretation and control unit. In this way, a single instruction can be executed at the same time on each of several execution units, each on a different set of operands. This kind of architecture is often referred to as single-instruction stream, multiple-data stream (SIMD).

Vector processing is the term applied to a form of pipelined arithmetic units which are specialized for performing arithmetic operations on vectors, which are uniform, linear arrays of data values. It can be thought of as a type of SIMD processing, since a single instruction invokes the execution of the same operation on every element of the array. *See* COMPUTER PROGRAMMING; PROGRAMMING LANGUAGES.

A central processing unit can contain multiple sets of the instruction execution hardware for either scalar or vector instructions. The task of scheduling instructions which can correctly execute in parallel with one another is generally the responsibility of the compiler or special scheduling hardware in the central processing unit. Instruction-level parallelism is almost never visible to the application programmer.

Multiprocessing is a form of parallelism that has complete central processing units operating in parallel, each fetching and executing instructions independently from the others. This type of computer organization is called multiple-instruction stream, multiple-data stream (MIMD). *See* MULTIPROCESSING. [D.W.M.]

Superconducting devices Devices that perform functions in the superconducting state that would be difficult or impossible to perform at room temperature, or that contain components which perform such functions. The superconducting state involves a loss of electrical resistance and occurs in many metals and alloys at temperatures near absolute zero. An enormous impetus was provided by the discovery in 1986 of a new class of ceramic, high-transition-temperature (T_c)

superconductors, which has resulted in a new superconducting technology at liquid nitrogen temperature. Superconducting devices may be conveniently divided into two categories: small-scale thin-film devices, and large-scale devices which employ zero-resistance superconducting windings made of type II superconducting materials. *See* SUPERCONDUCTIVITY.

Small-scale devices. A variety of thin-film devices offer higher performance than their nonsuperconducting counterparts. The prediction and discovery in the early 1960s of the Josephson effects introduced novel opportunities for ultrasensitive detectors, high-speed switching elements, and new physical standards. Niobium-based devices, patterned on silicon wafers using photolithographic techniques taken over from the semiconductor industry, have reached a high level of development, and a variety of such devices are commercially available. These devices operate at or below 4.2 K (−452°F), the temperature of liquid helium boiling under atmospheric pressure. *See* LIQUID HELIUM.

The discovery of the high-transition-temperature superconductors has enabled the operation of devices in liquid nitrogen at 77 K (−321°F). Not only is liquid nitrogen much cheaper and more readily available than liquid helium, but it also boils away much more slowly, enabling the use of simpler and more compact dewars or simpler, relatively inexpensive refrigerators. Of the new ceramic superconductors, only $YBa_2Cu_3O_{7-x}$ (YBCO) has been developed in thin-film form to the point of practical applications, and several devices are available. Intensive materials research has resulted in techniques, notably laser-ablation and radio-frequency sputtering, for the epitaxial growth of high-quality films with their crystalline planes parallel to the surface of the substrate. Most of the successful Josephson-junction devices have been formed at the interface between two grains of YBCO. These so-called grain-boundary junctions are made by depositing the film either on a bicrystal in which the two halves of the substrate have a carefully engineered in-plane misalignment of the crystal axes, or across a step-edge patterned in the substrate. *See* CRYOGENICS; GRAIN BOUNDARIES.

Two types of superconducting quantum interference device (SQUID) detect changes in magnetic flux: the dc SQUID and the rf SQUID. The dc SQUID, which operates with a dc bias current, consists of two Josephson junctions incorporated into a superconducting loop. The maximum dc supercurrent, known as the critical current, and the current-voltage (*I-V*) characteristic of the SQUID oscillate when the magnetic field applied to the device is changed. The oscillations are periodic in the magnetic flux Φ threading the loop with a period of one flux quantum, $\Phi_0 = h/2e \approx 2.07 \times 10^{-15}$ weber, where h is Planck's constant and e is the magnitude of the charge of the electron. Thus, when the SQUID is biased with a constant current, the voltage is periodic in the flux. The SQUID is almost invariably operated in a flux-locked loop. A change in the applied flux gives rise to a corresponding current in the coil that produces an equal and opposite flux in the SQUID. The SQUID is thus the null detector in a feedback circuit, and the output voltage is linearly proportional to the applied flux. *See* JOSEPHSON EFFECT; SQUID.

The rf SQUID consists of a single Josephson junction incorporated into a superconducting loop and operates with an rf bias. The SQUID is coupled to the inductor of an *LC*-resonant circuit excited at its resonant frequency, typically 30 MHz. The characteristics of rf voltage across the tank circuit versus the rf current depends on applied flux. With proper adjustment of the rf current, the amplitude of the rf voltage across the tank circuit oscillates as a function of applied flux. The rf SQUID is also usually operated in a feedback mode.

SQUIDs are mostly used in conjunction with an input circuit. For example, magnetometers are made by connecting a superconducting pickup loop to the input coil to form a flux transformer. A magnetic field applied to the pickup loop induces a persistent current in the transformer and hence a magnetic flux in the SQUID. These magnetometers have found application in geophysics, for example, in magnetotellurics. *See* GEOPHYSICAL EXPLORATION; MAGNETOMETER.

Low-transition-temperature SQUIDs are widely used to measure the magnetic susceptibility of tiny samples over a wide temperature range. Another application is a highly sensitive voltmeter, used in measurements of the Hall effect and of thermoelectricity. Low-transition-temperature SQUIDs are used as ultrasensitive detectors of nuclear magnetic and nuclear quadrupole resonance, and as transducers for gravitational-wave antennas. So-called scanning SQUIDs are used to obtain magnetic images of objects ranging from single-flux quanta trapped in superconductors to subsurface damage in two metallic sheets riveted together. *See* HALL EFFECT; MAGNETIC SUSCEPTIBILITY; NUCLEAR MAGNETIC RESONANCE (NMR); NUCLEAR QUADRUPOLE RESONANCE; THERMOELECTRICITY; VOLTMETER.

Perhaps the single largest area of application is biomagnetism, notably to image magnetic sources in the human brain or heart. In these studies an array of magnetometers or gradiometers is placed close to the subject, both generally being in a magnetically shielded room. The fluctuating magnetic signals recorded by the various channels are analyzed to locate their source. These techniques have been used, for example, to pinpoint the origin of focal epilepsy and to determine the function of the brain surrounding a tumor prior to its surgical removal. *See* BIOMAGNETISM.

The most sensitive detector available for millimeter and submillimeter electromagnetic radiation is the superconductor-insulator-superconductor (SIS) quasiparticle mixer. In this tunnel junction, usually niobium–aluminum oxide–niobium, the Josephson supercurrent is quenched and only single electron tunneling occurs. The current-voltage characteristic exhibits a very sharp onset of current at a voltage $2\Delta/e$, where Δ is the superconducting energy gap. The mixer is biased near this onset where the characteristics are highly nonlinear and used to mix the signal frequency with the frequency of a local oscillator to produce an intermediate frequency that is coupled out into a low-noise amplifier. These mixers are useful at frequencies up to about 750 GHz (wavelengths down to 400 micrometers). Such receivers are of great importance in radio astronomy, notably for airborne, balloon-based, or high-altitude, ground-based telescopes operating above most of the atmospheric water vapor. *See* RADIO ASTRONOMY.

The advent of high-transition-temperature superconductors stimulated major efforts to develop passive radio-frequency and microwave components that take advantage of the low electrical losses offered by these materials compared with normal conductors in liquid nitrogen. The implementation of thin-film YBCO receiver coils has improved the signal-to-noise ratio of nuclear magnetic resonance (NMR) spectrometers by a factor of 3 compared to that achievable with conventional coils. This improvement enables the data acquisition time to be reduced by an order of magnitude. These coils also have potential applications in low-frequency magnetic resonance imaging (MRI). High-transition-temperature bandpass filters have application in cellular communications. *See* ELECTRIC FILTER; MOBILE RADIO. [J.Cl.]

Large-scale devices. Large-scale applications of superconductivity comprise medical, energy, transportation, high-energy physics, and other miscellaneous applications such as high-gradient magnetic separation. When strong magnetic fields are needed, superconducting magnets offer several advantages over conventional copper or aluminum electromagnets. Most important is lower electric power costs because once the system is energized only the refrigeration requires power input, generally only 5–10% that of an equivalent-field resistive magnet. Relatively high magnetic fields achievable in unusual configurations and in smaller total volumes reduce the costs of expensive force-containment structures. *See* MAGNET.

Niobium-titanium (NbTi) has been used most widely for large-scale applications, followed by the A15 compounds, which include niobium-tin (Nb_3Sn), niobium-aluminum (Nb-Al),

niobium-germanium (Nb-Ge), and vanadium-gallium (Va_3Ga). Niobium-germanium held the record for the highest critical field (23 K; −418.5°F) until the announcement of high-temperature ceramic superconductors. *See* A15 PHASES.

Significant advances have been made in high-temperature superconducting wire development. Small coils have been wound that operate at 20 K (−410°F). Current leads are in limited commercial use. Considerable development remains necessary to use these materials in very large applications.

MRI dominates superconducting magnet systems applications. Most of the MRI systems are in use in hospitals and clinics, and incorporate superconducting magnets. *See* MEDICAL IMAGING.

Some of the largest-scale superconducting magnet systems are those considered for energy-related applications. These include magnetic confinement fusion, superconducting magnetic energy storage, magnetohydrodynamic electrical power generation, and superconducting generators. *See* MAGNETOHYDRODYNAMIC POWER GENERATOR; NUCLEAR FUSION.

In superconducting magnetic energy storage superconducting magnets are charged during off-peak hours when electricity demand is low, and then discharged to add electricity to the grid at times of peak demand. The largest systems would require large land areas, for example, an 1100-m-diameter (3600-ft) site for a 5000-MWh system. However, intermediate-size systems are viable. A 6-T peak-field solenoidal magnet system designed for the Alaskan power network stores 1800 megajoules (0.5 MWh). High-purity-aluminum-stabilized niobium-titanium alloy conductor carrying 16 kiloamperes current is used for the magnet winding.

Superconducting magnets have potential applications for transportation, such as magnetically levitated vehicles. In addition, superconducting magnets are used in particle accelerators and particle detectors. *See* MAGNETIC LEVITATION; PARTICLE ACCELERATOR; PARTICLE DETECTOR. [A.M.Da.]

Superconductivity A phenomenon occurring in many electrical conductors, in which the electrons responsible for conduction undergo a collective transition into an ordered state with many unique and remarkable properties. These include the vanishing of resistance to the flow of electric current, the appearance of a large diamagnetism and other unusual magnetic effects, substantial alteration of many thermal properties, and the occurrence of quantum effects otherwise observable only at the atomic and subatomic level.

Superconductivity was discovered by H. Kamerlingh Onnes in Leiden in 1911 while studying the temperature dependence of the electrical resistance of mercury within a few degrees of absolute zero. He observed that the resistance dropped sharply to an unmeasurably small value at a temperature of 4.2 K (−452°F). The temperature at which the transition occurs is called the transition or critical temperature, T_c. The vanishingly small resistance (very high conductivity) below T_c suggested the name given the phenomenon.

In 1933 W. Meissner and R. Ochsenfeld discovered that a metal cooled into the superconducting state in a moderate magnetic field expels the field from its interior. This discovery demonstrated that superconductivity involves more than simply very high or infinite electrical conductivity, remarkable as that alone is. *See* MEISSNER EFFECT.

In 1957, J. Bardeen, L. N. Cooper, and J. R. Schrieffer reported the first successful microscopic theory of superconductivity. The Bardeen-Cooper-Schrieffer (BCS) theory describes how the electrons in a conductor form the ordered superconducting state. The BCS theory still stands as the basic explanation of superconductivity, even though extensive theoretical work has embellished it.

There are a number of practical applications of superconductivity. Powerful superconducting electromagnets guide elementary particles in particle accelerators, and they also provide the magnetic field needed for magnetic resonance imaging. Ultrasensitive superconducting circuits are used in medical studies of the human heart and brain and for a wide variety of physical science experiments. A completely superconducting prototype computer has even been built. *See* MEDICAL IMAGING; PARTICLE ACCELERATOR; SUPERCONDUCTING DEVICES.

Transition temperatures. It was realized from the start that practical applications of superconductivity could become much more widespread if a high-temperature superconductor, that is, one with a high T_c, could be found. For instance, the only practical way to cool superconductors with transition temperatures below 20 K (−424°F) is to use liquid helium, which boils at a temperature of 4.2 K (−452°F) and which is rather expensive. On the other hand, a superconductor with a transition temperature of 100 K (−280°F) could be cooled with liquid nitrogen, which boils at 77 K (−321°F) and which is roughly 500 times less expensive than liquid helium. Another advantage of a high-T_c material is that, since many of the other superconducting properties are proportional to T_c, such a material would have enhanced properties. In 1986 the discovery of transition temperatures possibly as high as 30 K (−406°F) was reported in a compound containing barium, lanthanum, copper, and oxygen. In 1987 a compound of yttrium, barium, copper, and oxygen was shown to be superconducting above 90 K (−298°F). In 1988 researchers showed that a bismuth, strontium, calcium, copper, and oxygen compound was superconducting below 110 K (−262°F), and transition temperatures as high as 135 K (−216°F) were found in a mercury, thallium, barium, calcium, copper, and oxygen compound.

Occurrence. Some 29 metallic elements are known to be superconductors in their normal form, and another 17 become superconducting under pressure or when prepared in the form of thin films. The number of known superconducting compounds and alloys runs into the thousands. Superconductivity is thus a rather common characteristic of metallic conductors. The phenomenon also spans an extremely large temperature range. Rhodium is the element with the lowest transition temperature (370 μK), while $Hg_{0.2}Tl_{0.8}Ca_2Ba_2Cu_3O$ is the compound with the highest (135 K or −216°F).

Despite the existence of a successful microscopic theory of superconductivity, there are no completely reliable rules for predicting whether a metal will be a superconductor. Certain trends and correlations are apparent among the known superconductors, however—some with obvious bases in the theory—and these provide empirical guidelines in the search for new superconductors. Superconductors with relatively high transition temperatures tend to be rather poor conductors in the normal state.

The ordered superconducting state appears to be incompatible with any long-range-ordered magnetic state: Usually the ferromagnetic or antiferromagnetic metals are not superconducting. The presence of nonmagnetic impurities in a superconductor usually has very little effect on the superconductivity, but the presence of impurity atoms which have localized magnetic moments can markedly depress the transition temperature even in concentrations as low as a few parts per million. *See* ANTIFERROMAGNETISM; FERROMAGNETISM.

Some semiconductors with very high densities of charge carriers are superconducting, and others such as silicon and germanium have high-pressure metallic phases which are superconducting. Many elements which are not themselves superconducting form compounds which are.

Certain organic conductors are superconducting. For instance, brominated polymeric chains of sulfur and nitrogen, known as $(SNBr_{0.4})_x$, are superconducting below 0.36 K. Other more complicated organic materials have T_c values near 10 K (−442°F). *See* ORGANIC CONDUCTOR.

Although nearly all the classes of crystal structure are represented among superconductors, certain structures appear to be especially conducive to high-temperature superconductivity. The so-called A15 structure, shared by a series of intermetallic compounds based on niobium, produced several superconductors

with T_c values above 15 K (−433°F) as well as the record holder, NbGe, at 23 K (−418°F). Indeed, the robust applications of superconductivity that depend on the ability to carry high current in the presence of high magnetic fields still exclusively use two members of this class: NbTi with $T_c = 8$ K (−445°F), and Nb_3Sn with $T_c = 18.1$ K (−427°F). *See* A15 PHASES.

After 1986 the focus of superconductivity research abruptly shifted to the copper-oxide-based planar structures, due to their significantly higher transition temperatures. Basically there are three classes of these superconductors, all of which share the common feature that they contain one or more conducting planes of copper and oxygen atoms. The first class is designated by the chemical formula $La_{2-x}A_xCuO_4$, where the A atom can be barium, strontium, or calcium. Superconductivity was originally discovered in the barium-doped system, and systematic study of the substitutions of strontium, calcium, and so forth have produced transition temperatures as high as 40 K (−388°F).

The second class of copper-oxide superconductor is designated by the chemical formula $Y_1Ba_2Cu_3O_{7-\delta}$, with $\delta < 1.0$. Here, single sheets of copper and oxygen atoms straddle the rare-earth yttrium ion and chains of copper and oxygen atoms thread among the barium ions. The transition temperature, 92 K (−294°F), is quite insensitive to replacement of yttrium by many other rare-earth ions. *See* RARE-EARTH ELEMENTS.

The third class is the most complicated. These compounds contain either single thallium-oxygen layers, represented by the chemical formula $Tl_1Ca_{n-1}Ba_2Cu_nO_{2n+3}$, where n refers to the number of copper-oxygen planes, or double thallium-oxygen layers, represented by the chemical formula $Tl_2Ca_{n-1}Ba_2Cu_nO_{2n+4}$. The number of copper-oxygen planes may be varied, and as many as three planes have been included in the structure. Thallium may be replaced by bismuth, thus generating a second family of superconductors. In all of these compounds, the transition temperature appears to increase with the number of planes, but T_c decreases for larger values of n.

The spherical molecule comprising 60 carbon atoms (C_{60}), known as a buckyball, can be alloyed with various alkaline atoms which contribute electrons for conduction. By varying the number of conductors in C_{60}, it is possible to boost T_c to a maximum value of 52 K (−366°F). *See* FULLERENE.

Superconductivity was discovered in magnesium diboride (MgB_2) in January 2001 in Japan. This material may be a good alternative for some of the applications envisioned for high-T_c superconductivity, since this compound has T_c of 39 K (−389°F), is relatively easy to make, and consists of only two elements.

Magnetic properties. The existence of the Meissner-Ochsenfeld effect, the exclusion of a magnetic field from the interior of a superconductor, is direct evidence that the superconducting state is not simply one of infinite electrical conductivity. Instead, it is a true thermodynamic equilibrium state, a new phase which has lower free energy than the normal state at temperatures below the transition temperature and which somehow requires the absence of magnetic flux.

The exclusion of magnetic flux by a superconductor costs some magnetic energy. So long as this cost is less than the condensation energy gained by going from the normal to the superconducting phase, the superconductor will remain completely superconducting in an applied magnetic field. If the applied field becomes too large, the cost in magnetic energy will outweigh the gain in condensation energy, and the superconductor will become partially or totally normal. The manner in which this occurs depends on the geometry and the material of the superconductor. The geometry which produces the simplest behavior is that of a very long cylinder with field applied parallel to its axis. Two distinct types of behavior may then occur, depending on the type of superconductor—type I or type II.

Below a critical field H_c which increases as the temperature decreases below T_c, the magnetic flux is excluded from a type I superconductor, which is said to be perfectly diamagnetic. For a type II superconductor, there are two critical fields, the lower critical field H_{c1} and the upper critical field H_{c2}. In applied fields less than H_{c1}, the superconductor completely excludes the field, just as a type I superconductor does below H_c. At fields just above H_{c1}, however, flux begins to penetrate the superconductor, not in a uniform way, but as individual, isolated microscopic filaments called fluxoids or vortices. Each fluxoid consists of a normal core in which the magnetic field is large, surrounded by a superconducting region in which flows a vortex of persistent supercurrent which maintains the field in the core. *See* DIAMAGNETISM.

Thermal properties. The appearance of the superconducting state is accompanied by rather drastic changes in both the thermodynamic equilibrium and thermal transport properties of a superconductor.

The heat capacity of a superconducting material is quite different in the normal and superconducting states. In the normal state (produced at temperatures below the transition temperature by applying a magnetic field greater than the critical field), the heat capacity is determined primarily by the normal electrons (with a small contribution from the thermal vibrations of the crystal lattice) and is nearly proportional to the temperature. In zero applied magnetic field, there appears a discontinuity in the heat capacity at the transition temperature. At temperatures just below the transition temperature, the heat capacity is larger than in the normal state. It decreases more rapidly with decreasing temperature, however, and at temperatures well below the transition temperature varies exponentially as $e^{-\Delta/kT}$, where Δ is a constant and k is Boltzmann's constant. Such an exponential temperature dependence is a hallmark of a system with a gap Δ in the spectrum of allowed energy states. Heat capacity measurements provided the first indications of such a gap in superconductors, and one of the key features of the macroscopic BCS theory is its prediction of just such a gap. *See* SPECIFIC HEAT OF SOLIDS.

Ordinarily a large electrical conductivity is accompanied by a large thermal conductivity, as in the case of copper, used in electrical wiring and cooking pans. However, the thermal conductivity of a pure superconductor is less in the superconducting state than in the normal state, and at very low temperatures approaches zero. Crudely speaking, the explanation for the association of infinite electrical conductivity with vanishing thermal conductivity is that the transport of heat requires the transport of disorder (entropy). The superconducting state is one of perfect order (zero entropy), and so there is no disorder to transport and therefore no thermal conductivity. *See* ENTROPY; THERMAL CONDUCTION IN SOLIDS.

Two-fluid model. C. J. Gorter and H. B. G. Casimir introduced in 1934 a phenomenological theory of superconductivity based on the assumption that in the superconducting state there are two components of the conduction electron "fluid" (hence the name given this theory, the two-fluid model). One, called the superfluid component, is an ordered condensed state with zero entropy; hence it is incapable of transporting heat. It does not interact with the background crystal lattice, its imperfections, or the other conduction electron component and exhibits no resistance to flow. The other component, the normal component, is composed of electrons which behave exactly as they do in the normal state. It is further assumed that the superconducting transition is a reversible thermodynamic phase transition between two thermodynamically stable phases, the normal state and the superconducting state, similar to the transition between the liquid and vapor phases of any substance. The validity of this assumption is strongly supported by the existence of the Meissner-Ochsenfeld effect and by other experimental evidence. This assumption permits the application of all the powerful and general machinery of the theory of equilibrium thermodynamics. The results tie together the observed thermodynamic properties of superconductors in a very satisfying way.

Microscopic (BCS) theory. The key to the basic interaction between electrons which gives rise to superconductivity was provided by the isotope effect. It is an interaction mediated by the

background crystal lattice and can crudely be pictured as follows: An electron tends to create a slight distortion of the elastic lattice as it moves, because of the Coulomb attraction between the negatively charged electron and the positively charged lattice. If the distortion persists for a brief time (the lattice may ring like a struck bell), a second passing electron will see the distortion and be affected by it. Under certain circumstances, this can give rise to a weak indirect attractive interaction between the two electrons which may more than compensate their Coulomb repulsion.

The first forward step was taken by Cooper in 1956, when he showed that two electrons with an attractive interaction can bind together to form a "bound pair" (often called a Cooper pair) if they are in the presence of a high-density fluid of other electrons, no matter how weak the interaction is. The two partners of a Cooper pair have opposite momenta and spin angular momenta. Then, in 1957, Bardeen, Cooper, and Schrieffer showed how to construct a wave function in which all of the electrons (at least, all of the important ones) are paired. Once this wave function is adjusted to minimize the free energy, it can be used as the basis for a complete microscopic theory of superconductivity.

The successes of the BCS theory and its subsequent elaborations are manifold. One of its key features is the prediction of an energy gap. Excitations called quasiparticles (which are something like normal electrons) can be created out of the superconducting ground state by breaking up pairs, but only at the expense of a minimum energy of Δ per excitation; Δ is called the gap parameter. The original BCS theory predicted that Δ is related to T_c by $\Delta = 1.76kT_c$ at $T = 0$ for all superconductors. This turns out to be nearly true, and where deviations occur they are understood in terms of modifications of the BCS theory. The manifestations of the energy gap in the low-temperature heat capacity and in electromagnetic absorption provide strong confirmation of the theory. [D.N.L.; R.J.So.; M.O.]

Supercontinent The six major continents today are Africa, Antarctica, Australia, Eurasia, North America, and South America. Prior to the formation of the Atlantic, Indian, and Southern ocean basins over the past 180 million years by the process known as sea-floor spreading, the continents were assembled in one supercontinent called Pangea (literally "all Earth"). Pangea came together by the collision, about 300 million years ago (Ma), of two smaller masses of continental rock, Laurasia and Gondwanaland. Laurasia comprised the combined continents of ancient North America (known as Laurentia), Europe, and Asia. Africa, Antarctica, Australia, India, and South America made up Gondwanaland (this name comes from a region in southern India). The term "supercontinent" is also applied to Laurasia and Gondwanaland; hence it is used in referring to a continental mass significantly bigger than any of today's continents. A supercontinent may therefore incorporate almost all of the Earth's continental rocks, as did Pangea, but that is not implied by the word. *See* CONTINENT; CONTINENTS, EVOLUTION OF.

Laurasia, Gondwanaland, and Pangea are the earliest supercontinental entities whose former existence can be proven. Evidence of older rifted continental margins, for example surrounding Laurentia and on the Pacific margins of South America, Antarctica, and Australia, point to the existence of older supercontinents. The hypothetical Rodinia (literally "the mother of all continents") may have existed 800–1000 Ma, and Pannotia (meaning "the all-southern supercontinent") fleetingly around 550 Ma. Both are believed to have included most of the Earth's continental material. There may have been still earlier supercontinents, because large-scale continents, at least the size of southern Africa or Western Australia, existed as early as 2500 Ma at the end of Archean times. *See* ARCHEAN; CONTINENTAL MARGIN.

The amalgamation and fragmentation of supercontinents are the largest-scale manifestation of tectonic forces within the Earth. The cause of such events is highly controversial. *See* PLATE TECTONICS. [I.W.O.D.]

Supercritical fields Static fields that are strong enough to cause the normal vacuum, which is devoid of real particles, to break down into a new vacuum in which real particles exist. This phenomenon has not yet been observed for electric fields, but it is predicted for these fields as well as others such as gravitational fields and the gluon field of quantum chromodynamics.

Vacuum decay in quantum electrodynamics. The original motivation for developing the new concept of a charged vacuum arose in the late 1960s in connection with attempts to understand the atomic structure of superheavy nuclei expected to be produced by heavy-ion linear accelerators. *See* PARTICLE ACCELERATOR.

The best starting point for discussing this concept is to consider the binding energy of atomic electrons as the charge Z of a heavy nucleus is increased. If the nucleus is assumed to be a point charge, the total energy E of the $1s_{1/2}$ level drops to 0 when $Z = 137$. This so-called $Z = 137$ catastrophe had been well known, but it was argued loosely that it disappears when the finite size of the nucleus is taken into account. However, in 1969 it was shown that the problem is not removed but merely postponed, and reappears around $Z = 173$. Any level $E(nj)$ can be traced down to a binding energy of twice the electronic rest mass if the nuclear charge is further increased. At the corresponding charge number, called Z_{cr}, the state dives into the negative-energy continuum of the Dirac equation (the so-called Dirac sea). The overcritical state acquires a width and is spread over the continuum. *See* ANTIMATTER; RELATIVISTIC QUANTUM THEORY.

When Z exceeds Z_{cr} a K-shell electron is bound by more than twice its rest mass, so that it becomes energetically favorable to create an electron-positron pair. The electron becomes bound in the $1s_{1/2}$ orbital and the positron escapes. The overcritical vacuum state is therefore said to be charged. *See* POSITRON.

Clearly, the charged vacuum is a new ground state of space and matter. The normal, undercritical, electrically neutral vacuum is no longer stable in overcritical fields: it decays spontaneously into the new stable but charged vacuum. Thus the standard definition of the vacuum, as a region of space without real particles, is no longer valid in very strong external fields. The vacuum is better defined as the energetically deepest and most stable state that a region of space can have while being penetrated by certain fields.

Superheavy quasimolecules. Inasmuch as the formation of a superheavy atom of $Z > 173$ is very unlikely, a new idea is necessary to test these predictions experimentally. That idea, based on the concept of nuclear molecules, was put forward in 1969: a superheavy quasimolecule forms temporarily during the slow collision of two heavy ions. It is sufficient to form the quasimolecule for a very short instant of time, comparable to the time scale for atomic processes to evolve in a heavy atom, which is typically of the order 10^{-18} to 10^{-20}. Suppose a uranium ion is shot at another uranium ion at an energy corresponding to their Coulomb barrier, and the two, moving slowly (compared to the K-shell electron velocity) on Rutherford hyperbolic trajectories, are close to each other (compared to the K-shell electron orbit radius). Then the atomic electrons move in the combined Coulomb potential of the two nuclei, thereby experiencing a field corresponding to their combined charge of 184. This happens because the ionic velocity (of the order of $c/10$) is much smaller than the orbital electron velocity (of the order of c), so that there is time for the electronic molecular orbits to be established, that is, to adjust to the varying distance between the charge centers, while the two ions are in the vicinity of each other. *See* QUASIATOM.

Giant nuclear systems. The energy spectrum for positrons created in, for example, a uranium-curium collision consists of three components: the induced, the direct, and the spontaneous, which add up to a smooth spectrum. The presence of the spontaneous component leads only to 5–10% deviations for normal nuclear collisions along Rutherford trajectories. This situation raises the question as to whether there is any way to get a clear

qualitative signature for spontaneous positron production. Suppose that the two colliding ions, when they come close to each other, stick together for a certain time Δt before separating again. The longer the sticking, the better is the static approximation. For Δt very long, a very sharp line should be observed in the positron spectrum with a width corresponding to the natural lifetime of the resonant positron-emitting state. The observation of such a sharp line will indicate not only the spontaneous decay of the vacuum but also the formation of giant nuclear systems ($Z > 180$). *See* LINEWIDTH; NUCLEAR MOLECULE.

Search for spontaneous positron emission. The search for spontaneous positron emission in heavy-ion collisions began in 1976. Of special interest are peak structures in the positron energy distribution. However, the issue of spontaneous positron production in strong fields remains open. If line structures have been observed at all, they are most likely due to nuclear conversion processes. The observation of vacuum decay very much depends on the existence of sufficiently long-lived (at least 10^{-20} s) giant nuclear molecular systems. Therefore, the investigation of nuclear properties of heavy nuclei encountering heavy nuclei at the Coulomb barrier is a primary task.

Other field theories. The idea of overcriticality also has applications in other field theories, such as those of pion fields, gluon fields (quantum chromodynamics), and gravitational fields (general relativity).

A heavy meson may be modeled as an ellipsoidal bag, with a heavy quark Q and antiquark $\bar{Q}$ located at the foci of the ellipsoid. The color-electric or glue-electric field lines do not penetrate the bag surface. The Dirac equation may be solved for light quarks q and antiquarks $\bar{q}$ in this field of force. In the spherical case the potential is zero, and the solutions with different charges degenerate. As the source charges Q and $\bar{Q}$ are pulled apart, the wave functions start to localize. At a critical deformation of the bag, positive and negative energy states cross; that is, overrcriticality is reached and the color field is strong enough that the so-called perturbative vacuum inside the bag rearranges so that the wave functions are pulled to opposite sides and the color charges of the heavy quarks are completely shielded. Hence two new mesons of types $\bar{Q}q$ and $Q\bar{q}$ appear; the original meson fissions. *See* GLUONS; QUANTUM CHROMODYNAMICS; QUARKS. [W.Gr.]

Supercritical fluid Any fluid at a temperature and a pressure above its critical point; also, a fluid above its critical temperature regardless of pressure. Below the critical point the fluid can coexist in both gas and liquid phases, but above the critical point there can be only one phase. Supercritical fluids are of interest because their properties are intermediate between those of gases and liquids, and are readily adjustable. *See* CRITICAL PHENOMENA; PHASE EQUILIBRIUM.

In a given supercritical fluid the thermodynamic and transport properties are a function of density, which depends strongly on the fluid's pressure and temperature. The density may be adjusted from a gaslike value of 0.1 g/ml to a liquidlike value as high as 1.2 g/ml. Furthermore, as conditions approach the critical point, the effect of temperature and pressure on density becomes much more significant. Increasing the density of supercritical carbon dioxide from 0.2 to 0.5 g/ml, for example, requires raising the pressure from 85 to 140 atm (8.6 to 14.2 megapascals) at 158°F (70°C), but at 95°F (35°C) the required change is only from 65 to 80 atm (6.6 to 8.1 MPa).

For a given fluid, the logarithm of the solubility of a solute is approximately proportional to the solvent density at constant temperature. Therefore, a small increase in pressure, which causes a large increase in the density, can raise the solubility a few orders of magnitude. While almost all of a supercritical fluid's properties vary with density, some of these properties are more like those of a liquid while others are more like those of a gas. *See* SUPERCRITICAL-FLUID CHROMATOGRAPHY.

In most supercritical-fluid applications the fluid's critical temperature is less than 392°F (200°C) and its critical pressure is less than 80 atm (8.1 MPa). High critical temperatures require operating temperatures that can damage the desired product, while high critical pressures result in excessive compression costs. In addition to these pure fluids, mixed solvents can be used to improve the solvent strength. [K.Jo.; R.Len.]

Spercritical-fluid chromatography Any separation technique in which a supercritical fluid is used as the mobile phase. For any fluid, a phase diagram can be constructed to show the regions of temperature and pressure at which gases and liquids, gases and solids, and liquids and solids can coexist. For the gas-liquid equilibrium, there is a certain temperature and pressure, known as the critical temperature and pressure, below which a gas and a liquid can coexist but above which only a single phase (known as a supercritical fluid) can form.

For example, the density of supercritical fluids is usually between 0.25 and 1.2 g/ml and is strongly pressure-dependent. Their solvent strength increases with density, so that molecules that are retained on the column can often be eluted simply by increasing the pressure under which the fluid is compressed.

Diffusion coefficients of solutes in supercritical fluids are tenfold greater than the corresponding values in liquid solvents (although about three orders of magnitude less than the corresponding values in gases). The high diffusivity of solutes in supercritical fluids decreases their resistance to mass transfer in a chromatographic column and hence allows separations to be made either very quickly or at high resolution. Molecules can be separated in a few minutes or less by using packed columns of the type used for high-performance liquid chromatography (HPLC). *See* DIFFUSION IN GASES AND LIQUIDS.

Despite the relatively large number of separations made using packed-column supercritical-fluid chromatography that have been described since 1962, supercritical-fluid chromatography became popular only in the late 1980s with the commercial introduction of capillary (open-tubular) supercritical-fluid chromatographs. While supercritical-fluid chromatography may never be as widely used as gas chromatography or high-performance liquid chromatography, it provides a useful complement to these chromatographies. *See* CHEMICAL SEPARATION TECHNIQUES; CHROMATOGRAPHY; FLUIDS; GAS CHROMATOGRAPHY; LIQUID CHROMATOGRAPHY. [P.R.G.]

Supercritical wing A wing with special streamwise sections, or airfoils, which provide substantial delays in the onset of the adverse aerodynamic effects which usually occur at high subsonic flight speeds.

When the speed of an aircraft approaches the speed of sound, the local airflow about the airplane, particularly above the upper surface of the wing, may exceed the speed of sound. Such a condition is called supercritical flow. On previous aircraft, this supercritical flow resulted in the onset of a strong local shock wave above the upper surface of the wing (illus. *a*). This local

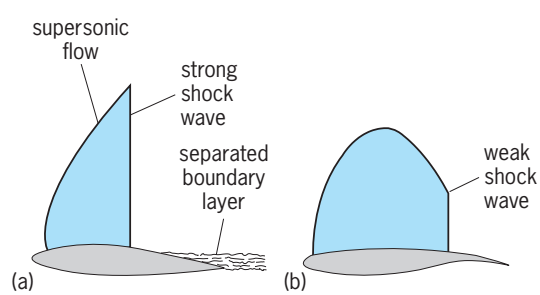

Comparison of airflow about airfoils. (*a*) Conventional airfoils, Mach number = 0.7. (*b*) Supercritical airfoils. Mach number = 0.8.

wave caused an abrupt increase in the pressure on the surface of the wing, which may cause the surface boundary-layer flow to separate from the surface, with a resulting severe increase in the turbulence of the flow. The increased turbulence leads to a severe increase in drag and loss in lift, with a resulting decrease in flight efficiency. The severe turbulence also caused buffet or shaking of the aircraft and substantially changed its stability or flying qualities. *See* AERODYNAMIC FORCE; AERODYNAMIC WAVE DRAG; LOW-PRESSURE GAS FLOW; SHOCK WAVE; TRANSONIC FLIGHT.

Supercritical airfoils are shaped to substantially reduce the strength of the shock wave and to delay the associated boundary-layer separation (illus. *b*). Since the airfoil shape allows efficient flight at supercritical flight speeds, a wing of such design is called a supercritical wing. *See* AIRPLANE; WING. [R.T.Wh.]

Superfluidity The frictionless flow of liquid helium at low temperature; also, the flow of electric current without resistance in certain solids at low temperature (superconductivity).

Both helium isotopes have a superfluid transition, but the detailed properties of their superfluid states differ considerably because they obey different statistics. ^{4}He, with an intrinsic spin of 0, is subject to Bose-Einstein statistics, and ^{3}He, with a spin of $^1/_2$, to Fermi-Dirac statistics. There are two distinct superfluid states in ^{3}He called A and B.

The term "superfluidity" usually implies He II or the A and B phases of ^{3}He, but the basic similarity between these and the "fluid" consisting of pairs of electrons in superconductors is sufficiently strong to designate the latter as a charged superfluid. Besides flow without resistance, superfluid helium and superconducting electrons display quantized circulating flow patterns in the form of microscopic vortices. *See* BOSE-EINSTEIN STATISTICS; LIQUID HELIUM; QUANTIZED VORTICES; SECOND SOUND; SUPERCONDUCTIVITY. [L.J.C.]

Supergiant star A member of a class of evolved stars that occupy the top of the Hertzsprung-Russell diagram to the right of the main sequence. The absolute visual magnitudes (M_V) of supergiants range approximately between -4 and -10, and they are the largest and brightest stars. They are recognized by their spectroscopic characteristics. For example, those in class A have narrow hydrogen lines. Supergiants are subdivided into classes Ib (M_V about -5) and Ia (M_V about -7). A "hypergiant class zero" was later added near $M_V = -10$; the use of transition class Ia-0 at absolute visual magnitudes near -9 is now common. Red supergiants are the largest of all stars, and at maximum can reach diameters approaching that of the orbit of Saturn.

In an evolutionary sense, supergiants are stars above about 10 solar masses and absolute visual magnitude -6 that began main-sequence life hotter than class B1 and cannot evolve to become white dwarfs. As progeny of O stars, supergiants are exceedingly rare. Supergiant masses allow further burning of carbon and oxygen to oxygen, neon, and magnesium; thence to silicon and sulfur; and thence to iron. The iron cores collapse to produce type II supernovae, the condensed remnants at the centers becoming either neutron stars or, from the most massive stars, black holes. *See* ASTRONOMICAL SPECTROSCOPY; BLACK HOLE; HERTZSPRUNG-RUSSELL DIAGRAM; NEUTRON STAR; SPECTRAL TYPE; STAR; STELLAR EVOLUTION; SUPERNOVA. [J.B.Ka.]

Supergranulation A system of convective cells, with typical diameters of 20,000 km (12,000 mi), that cover the Sun's surface. Solar convection cells are invisible in ordinary photographs.

High-resolution photographs of the Sun's visible surface reveal the granulation, a closely packed cellular grid having bright (hot) centers surrounded by dark (cool) lanes. Granules have lifetimes of 10–30 minutes and average diameters of 1000 km (620 mi). In the 1930s, L. Biermann suggested that, in the outer 30% of the Sun, heat from the Sun's interior is transported to the surface by convection, just like hot rising bubbles in a pot of boiling soup heated from below. The granules are the boiling bubbles (plumes) of this convective process. Theorists in the 1950s proposed that the bubble sizes are approximately equal to the local scale height H (the distance in which density or pressure changes by a factor $e \approx 2.7$). Since H varies from about 1000 km (620 mi) at the Sun's surface to 100,000 km (62,000 mi) at the base of the convection zone, a very large range of plume sizes was hypothesized. *See* CONVECTION (HEAT).

In 1959, R. Leighton modified the spectroheliograph to image the Sun using the Doppler and Zeeman effects. Images of line-of-sight (approaching and receding) gas motions (Dopplergrams) and of magnetic fields (magnetograms) could be obtained. The Dopplergrams showed a new cell structure with area 400 times that of the granulation, having a mainly horizontal flow pattern. Lifetimes of most supergranules are 1–2 days, but observations in 1998 with the Michelson Doppler Imager (MDI) on the *Solar and Heliospheric Observatory* (*SOHO*) spacecraft showed that some live longer than 4 days. The cells form an irregular polygonal structure, with most of the down-flow occurring at the polygon vertices. The supergranules fit neatly within an essentially identical network (grid) structure seen in magnetograms. The kinetic energy of the supergranular motions at the Sun's surface exceeds the magnetic field's ability to resist such motions, and thus the magnetic field is dragged forcibly to the supergranule boundaries until the two patterns coincide. This magnetic field, in turn, causes local heating of the upper solar atmosphere (chromosphere), producing a similar chromospheric network pattern seen in high-temperature spectral lines. *See* DOPPLER EFFECT; FRAUNHOFER LINES; SOLAR MAGNETIC FIELD; SPECTROGRAPH; SPECTROHELIOGRAPH; SUN; ZEEMAN EFFECT. [G.W.Si.]

Supergravity A theory that attempts to unify gravitation with the other fundamental interactions. The first, and only, completely successful unified theory was constructed by James Clerk Maxwell, in which the up-to-then unrelated electric and magnetic phenomena were unified in his electrodynamics. *See* FUNDAMENTAL INTERACTIONS; MAXWELL'S EQUATIONS.

Electroweak theory. The second stage of unification concerns the unification of electromagnetic and weak interactions, using Maxwell's theory as a guide. This was accomplished making use of the nonabelian gauge theories invented by C. N. Yang and R. L. Mills, and of spontaneous symmetry breaking. The symmetry of Maxwell's theory is very similar to spatial rotations about an axis, rotating the vector potentials while leaving the electric and magnetic fields unchanged. It is a local invariance because the rotations about a fixed axis can be made by different amounts at different points in space-time. Thus, Maxwell's theory is invariant under a one-parameter group of transformations U(1). In Yang-Mills theory this local invariance was generalized to theories with larger symmetry groups such as the three-dimensional rotation group $SO(3) \simeq SU(2)$ which has three parameters. The number of parameters of the local symmetry (gauge) group is also equal to the number of 4-vector potentials in the gauge theory based on that group. A detailed analysis of weak and electromagnetic forces shows that their description requires four 4-vector potentials (gauge fields), so that the gauge group must be a four-parameter group. In fact, it is the product $SU(2) \cdot U(1)$. *See* ELECTROWEAK INTERACTION; GAUGE THEORY; GROUP THEORY; SYMMETRY BREAKING.

Grand unified theories. In the third stage of unification, electroweak and strong forces are regarded as different components of a more general force which mediates the interactions of particles in a grand unified model. Strong forces are responsible for the interactions of hadrons and for keeping quarks confined inside hadrons. They are described by eight massless 4-vector potentials (gluons), the corresponding eight-parameter group being SU(3). This local symmetry is called color, and the corresponding theory quantum chromodynamics (QCD). The combination $SU(3) \cdot SU(2) \cdot SU(1)$ has strong experimental

support, and has come to be known as the standard model. Thus the gauge group of any grand unified model must include the standard model as a subsymmetry. The most dramatic prediction of these theories is the decay of protons. *See* GLUONS; GRAND UNIFICATION THEORIES; PROTON; QUANTUM CHROMODYNAMICS; QUARKS; STANDARD MODEL.

Supersymmetry and supergravity theories. A still higher and more ambitious stage of unification deals with the possibility of combining grand unified and gravity theories into a superunified theory, also known as supergravity. To achieve this, use is made of the dual role played by local internal symmetry groups. On the one hand, they describe the behavior of forces. On the other hand, they classify the elementary particles (fields) of the theory into multiplets: spin-zero fields in one multiplet, spin-1/2 fields in another multiplet, and so forth, but never fermions and bosons in a single irreducible multiplet of internal symmetry. This last restriction used to be a major obstacle on the way to superunification. This is because, of all the elementary particles, only the quanta of the gravitational field (gravitons) have spin 2, so that a multiplet of elementary particles including the graviton must of necessity involve particles of different spin. But then by an internal symmetry transformation, which is by definition distinct from space-time (Lorentz) transformations, it is possible to "rotate" particles of different spin into one another, thus altering their space-time transformation properties. This apparent paradox can be circumvented if both the internal symmetry and Lorentz transformations are part of a larger (supersymmetric) transformation group which also includes the spin-changing transformations. The irreducible multiplets of such supergroups naturally contain both fermions and bosons. This is how supersymmetry makes its appearance in supergravity theories. *See* GRAVITON; LORENTZ TRANSFORMATIONS; RELATIVITY; SUPERSYMMETRY; SYMMETRY LAWS (PHYSICS).

Effective theory. If supergravity models are regarded not as fundamental theories but as effective theories describing the low-energy behavior of superstring theories, it is possible to make a strong case for their usefulness. In that case, since supergravity is no longer a fundamental theory, it is no longer crucial that supergravity satisfy very stringent physical requirements such as renormalizability. In its role as an effective theory, supergravity has been used in a number of problems in particle physics. *See* ELEMENTARY PARTICLE; QUANTUM FIELD THEORY; SUPERSTRING THEORY. [F.M.]

Superheater A component of a steam-generating unit in which steam, after it has left the boiler drum, is heated above its saturation temperature. The amount of superheat added to the steam is influenced by the location, arrangement, and amount of superheater surface installed, as well as the rating of the boiler. The superheater may consist of one or more stages of tube banks arranged to effectively transfer heat from the products of combustion. *See* STEAM-GENERATING UNIT. [G.W.K.]

Superluminal motion Proper motion of an astronomical object apparently exceeding the velocity of light, c. This phenomenon is relatively common in the nuclei of quasars, many of which exhibit systematic changes in images of their radio-frequency emission over periods of months to years. In some cases, features in the image appear to separate at a speed inferred to be more than 10 times the speed of light, given the great distance of the quasars from Earth.

Superluminal motion was one of the most exciting discoveries to emerge from a technique in radio astronomy first developed in the late 1960s and called very long baseline interferometry (VLBI). This method involves the tape recording of radio signals from large antennas at up to 10–15 locations across the Earth, and the combination of these signals in a computer to form a radio image of the quasar at extremely high resolution (less than 0.001 arcsecond). *See* QUASAR; RADIO ASTRONOMY; RADIO TELESCOPE.

Superluminal motion is seen mostly in quasars but also in some other active galactic nuclei. This rapid motion is confined to within a few tens of parsecs of the nucleus, whose power source is believed to be a massive black hole. At least 30 examples of superluminal motion are now known. Most show apparent speeds less than $10c$, but examples of speeds above $20c$ have been found. A very few objects in the Milky Way Galaxy also show superluminal motion. An example is GRS 1915 + 105, a relativistic jet source, which emits strongly in the x-ray as well as in the radio spectrum. Because the object is within the Milky Way Galaxy, its apparent speed of $1.25c$ is detectable in less than a day. *See* GALAXY, EXTERNAL; X-RAY ASTRONOMY.

Announcement of the discovery of superluminal motion in 1972 caused widespread concern because of the apparent violation of Albert Einstein's special theory of relativity, even though the basic explanation still favored now was in fact predicted some years before the announcement. Many explanations were proposed (besides Einstein's theory being incorrect), but only the relativistic jet model has stood the test of time. Superluminal motion is explained in this model as primarily a geometric effect. A feature (perhaps a cloud of relativistic plasma) moves away from the nucleus of the quasar at high (relativistic) speed (but less than c) at a small angle to the line of sight to the Earth. Radio waves from the moving feature arrive only slightly later than waves from the nucleus, whereas the feature took a much longer time to reach its current position. The motion appears superluminal because the speed is calculated using this much shorter time interval. As the speed approaches c and the angle to the line of sight decreases, the apparent speed can be arbitrarily large. In this explanation, no material speeds faster than c are required, so there is no conflict with special relativity. *See* RELATIVITY. [S.C.U.]

Supermassive stars Hypothetical objects with masses exceeding 60 solar masses, the mass of the largest known ordinary stars (1 solar mass equals 4.4×10^{30} lbm or 2×10^{30} kg). The term is most often used in connection with objects larger than 10^4 solar masses that might be the energy source in quasars and active galaxies. These objects probably do not exist. However, their nonexistence is one of the major assumptions that makes the case for giant black holes, rather than supermassive stars, being the central engines of quasars. *See* BLACK HOLE; QUASAR; STAR. [H.L.Sh.]

Supermultiplet A generalization of the concept of a multiplet. A multiplet is a set of quantum-mechanical states, each of which has the same value of some fundamental quantum number and differs from the other members of the set by another quantum number which takes values from a range of numbers dictated by the fundamental quantum number. The number of states in the set is called the multiplicity or dimension of the multiplet. The concept was originally introduced to describe the set of states in a nonrelativistic quantum-mechanical system with the same value of the orbital angular momentum, L, and different values of the projection of the angular momentum on an axis, M. The values that M can take are the integers between $-L$ and L, $2L + 1$ in all. This is the dimension of the multiplet. If the hamiltonian operator describing the system is rotationally invariant, all states of the multiplet have the same energy. A supermultiplet is a generalization of the concept of multiplet to the case when there are several quantum numbers that describe the quantum-mechanical states. *See* ANGULAR MOMENTUM; SYMMETRY LAWS (PHYSICS).

Both concepts, multiplet and supermultiplet, acquire a precise mathematical meaning by the use of the theory of group transformations. A multiplet is an irreducible representation of a group, G. The quantum number called fundamental in the paragraph above labels the representation of the group. The other quantum number labels the representation of a subgroup G' of G. For angular momentum, the group G is the rotation group, called

special orthogonal group in three dimensions, SO(3), and its subgroup G' is the group of special orthogonal transformations in two dimensions, SO(2). A supermultiplet is a generalization to the case in which the group G is not a group of rank one but has larger rank. A group of rank one has only one quantum number to label its representations. The concept of a multiplet or supermultiplet is particularly useful in the classification of states of physical systems. *See* GROUP THEORY; NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; QUANTUM NUMBERS.

The term supermultiplet was first used by E. P. Wigner in 1932 in order to classify the quantum-mechanical states of light atomic nuclei. The constituents of these are protons, p, and neutrons, n. Each proton and neutron has an intrinsic spin, S, of $^1/_2$ in units of $\hbar$, which is Planck's constant divided by 2π. The projection of the intrinsic spin on an axis, S_z, is then $S_z = {}^1/_2$ or $-{}^1/_2$ (spin up or down). In addition to having the same spin, the proton and neutron have essentially the same mass but differ in that the proton is charged whereas the neutron is not. They can thus be regarded as different charge states of the same particle, a nucleon. The distinction can be made formal by introducing a quantum number called isotopic spin, T, which has the value $^1/_2$. The two charge orientations, T_z, are taken to be $^1/_2$ for the proton and $-{}^1/_2$ for the neutron. There are thus four constituents of nuclei, protons and neutrons with spin up and down, that is, $p \uparrow$, $p \downarrow$, $n \uparrow$, and $n \downarrow$. The set of transformations among these constituents forms a group called SU(4), the special unitary group in four dimensions. This is the group G for Wigner's theory. The representations of SU(4), that is, Wigner supermultiplets, are characterized by three quantum numbers $(\lambda_1, \lambda_2, \lambda_3)$ with $\lambda_1 \geq \lambda_2 \geq \lambda_3$. *See* I-SPIN; NUCLEAR STRUCTURE. [F.I.]

Supernova The catastrophic, explosive death of a star, accompanied by the sudden, transient brightening of the star to an optical luminosity comparable to that of an entire galaxy.

A supernova shines typically for several weeks to several months with a luminosity between 2×10^8 and 5×10^9 times that of the Sun, then gradually fades away. Each explosion ejects from one to several tens of solar masses at speeds ranging from thousands to tens of thousands of kilometers per second. The total kinetic energy, 10^{44} joules (2.5×10^{28} megatons of high explosive), is about 100 times the total light output, making supernovae some of the highest-energy explosions in the universe. Unlike its fainter relative, the nova, a supernova does not recur for the same object. *See* NOVA.

Supernovas may be grouped according to either their observational characteristics or their explosion mechanism. Basically, type I supernovae have no hydrogen in their spectrum; type II supernovae do. Two mechanisms are involved: thermonuclear explosion in white dwarfs and gravitational collapse in massive stars. Type I supernovae of different subclasses can occur by either mechanism, but it is thought that most type II supernovae are powered by gravitational collapse.

During the last thousand years, there have been approximately seven supernovae visible to the unaided eye, in 1006, 1054, 1181, 1408, 1572, 1604, and 1987. SN 1006 may have been as bright as the quarter moon. The first six of these occurred in the Earth's vicinity of the Milky Way Galaxy. But the last, and only, naked-eye supernova since the invention of modern instrumentation occurred in the Large Magellanic Cloud, a small satellite galaxy of the Milky Way about 160,000 light-years away. Supernovae are discovered in other galaxies at a rate of about 150 per year. Most supernovae in the Milky Way Galaxy are obscured by dust, but various arguments suggest that about two type II supernovae per century and one type Ia every other century occur in the Milky Way Galaxy. *See* MAGELLANIC CLOUDS; MILKY WAY GALAXY.

Type Ia supernovae. Type Ia supernovae may be regarded as nature's largest thermonuclear bombs. They occur when an accreting white dwarf, composed of carbon and oxygen, grows to a mass 1.38 times that of the Sun, almost the critical mass that can be supported by electron degeneracy pressure, and ignites carbon fusion near its center. Ignition occurs when carbon fusion at the center releases energy faster than neutrinos can carry it away. Because the pressure is insensitive to the temperature, a nuclear runaway occurs. Fusion releases energy, which raises the temperature, which makes fusion go faster, but the gas cannot expand and cool. The nuclear runaway spreads in about 1 second through the star. The energy released by this nuclear burning is more than enough to completely blow the white dwarf apart with high velocity. Nothing remains—no neutron star, no black hole, and no burst of neutrino emission. *See* BINARY STAR; THERMONUCLEAR REACTION.

Type II supernovae. A typical type II supernova results from a star somewhat over 8 solar masses, on the main sequence, that spends its last years as a red supergiant burning progressively heavier fuels in its center. The radius of the star, after hydrogen has burned and the star is part way through helium burning, is roughly 500 solar radii, and its luminosity is already about 100,000 times that of the Sun. Each burning stage is shorter than the previous one. The last stage turns silicon and sulfur into a ball of roughly 1.4 solar masses of iron. Once iron has been produced, no more nuclear energy is available. *See* SUPERGIANT STAR.

A combination of instabilities now leads to the implosion of the iron core to a neutron star. When the density at the center reaches several times that of the atomic nucleus, the collapse halts and briefly springs back owing to the short-range repulsive component of the nuclear force. But the energy of this bounce is soon dissipated, and a hot young neutron star remains which, over the next few seconds, radiates away its heat and binding energy as neutrinos. *See* NEUTRINO; STRONG NUCLEAR INTERACTIONS.

The energy output in these neutrinos is enormous, about 3×10^{46} joules or 15% of the rest mass of the Sun converted to energy; rivaling the luminosity of the rest of the observable universe in light. A small fraction of these neutrinos, about 0.3%, are absorbed in reactions with neutrons and protons in the regions just outside the neutron star and deposit their energy. Even this small amount of energy is much greater than the gravitational binding of the remaining part of the star external to the newly formed neutron star. A bubble of radiation is inflated by the neutrino energy deposition, the outer boundary of which expands supersonically, driving a shock wave through the rest of the star and ejecting it with high velocity. The main energy of the explosion, though, is carried away as neutrinos. This general picture was confirmed when a neutrino burst of the predicted energy and duration was detected February 23, 1987, from the Large Magellanic Cloud in conjunction with SN 1987A. *See* NEUTRINO ASTRONOMY; SHOCK WAVE.

Nucleosynthesis. Supernovae are major element factories, responsible for producing most of the elements in nature heavier than nitrogen. The largest yields are of the more abundant elements, including oxygen, silicon, magnesium, neon, iron, and a portion of carbon, but dozens of other elements are also made. *See* NUCLEOSYNTHESIS.

Type Ia cosmological applications. Because of their brightness and the regularity of their light curves, type Ia supernovae have long been used as standard candles to survey cosmological distances. More recently it has been realized that the relatively small variation that occurs in the peak brilliance of such supernovae may be correlated with their decline rates. Use of this so-called Phillips relation allows even greater precision in distance determination. Using type Ia supernovae in this fashion reveals a surprising result. Two independent analyses show that the expansion rate of the universe is not slowing as might be expected long after the big bang, but is actually accelerating. The pull of gravity can only cause deceleration, so the acceleration is attributed to an invisible form of dark energy that enters into the cosmological equations as a repulsive term. *See* ASTROPHYSICS, HIGH-ENERGY; COSMOLOGY; STAR; STELLAR EVOLUTION; UNIVERSE; VARIABLE STAR. [S.E.Wo.]

Superoxide chemistry A branch of chemistry that deals with the reactivity of the superoxide ion (O_2^-), a one-electron (e^-) adduct of molecular oxygen (dioxygen; O_2) formed by the combination of O_2 and e^-. Because 1–15% of the O_2 that is respired by mammals goes through the O_2^- oxidation state, the biochemistry and reaction chemistry of the species are important to those concerned with oxygen toxicity, carcinogenesis, and aging. Although the name superoxide has prompted many to assume an exceptional degree of reactivity for O_2^-, the use of the prefix in fact was chosen to indicate stoichiometry. Superoxide was the name given in 1934 to the newly synthesized potassium salt (KO_2) to differentiate its two-oxygens-per-metal stoichiometry from that of most other metal–oxygen compounds (NA_2O, Na_2O_2, NaOH, Fe_2O_3).

Ionic salts of superoxide (yellow-to-orange solids), which form from the reaction of dioxygen with metals such as potassium, rubidium, or cesium, are paramagnetic, with one unpaired electron per two oxygen atoms.

In 1969, by means of electron spin resonance (ESR) spectroscopy, superoxide ion was detected as a respiratory intermediate, and metalloproteins were discovered that catalyze the disproportionation of superoxide, that is, superoxide dismutases (SODs), as shown in the reaction below.

$$2O_2^- + 2H^+ \xrightarrow{\text{SOD}} O_2 + H_2O_2$$

The biological function of superoxide dismutases is believed to be the protection of living cells against the toxic effects of superoxide. The possibility that superoxide might be an important intermediate in aerobic life provided an impetus to the study of superoxide reactivity.

The most general and universal property of O_2^- is its tendency to act as a strong BrØnsted base. Its strong proton affinity manifests itself in any media. Another characteristic of O_2^- is its ability to act as a moderate one-electron reducing agent. *See* ACID AND BASE; OXIDATION-REDUCTION.

In general, superoxide ion chemistry does not appear to be sufficiently robust to make superoxide ion a toxin. However, it can interact with protons, halogenated carbons, and carbonyl compounds to yield peroxy radicals that are toxic. *See* OXYGEN TOXICITY; REACTIVE INTERMEDIATES.

Superoxide does not appear to have exceptional reactivity. Nevertheless, superoxide will continue to be an interesting species for study because of the multiplicity of its chemical reactions and because of its importance as an intermediate in reactions that involve dioxygen and hydrogen peroxide. *See* BIOINORGANIC CHEMISTRY; OXYGEN. [D.T.S.]

Superplastic forming A process for shaping superplastic materials, a unique class of crystalline materials that exhibit exceptionally high tensile ductility. Superplastic materials may be stretched in tension to elongations typically in excess of 200% and more commonly in the range of 400–2000%. There are rare reports of higher tensile elongations reaching as much as 8000%. The high ductility is obtained only for superplastic materials and requires both the temperature and rate of deformation (strain rate) to be within a limited range. The temperature and strain rate required depend on the specific material. A variety of forming processes can be used to shape these materials; most of the processes involve the use of gas pressure to induce the deformation under isothermal conditions at the suitable elevated temperature. The tools and dies used, as well as the superplastic material, are usually heated to the forming temperature. The forming capability and complexity of configurations producible by the processing methods of superplastic forming greatly exceed those possible with conventional sheet forming methods, in which the materials typically exhibit 10–50% tensile elongation. *See* SUPERPLASTICITY.

There are a number of commercial applications of superplastic forming and combined superplastic forming and diffusion bonding, including aerospace, architectural, and ground transportation uses. Examples are wing access panels in the Airbus A310 and A320, bathroom sinks in the Boeing 737, turbo-fan-engine cooling-duct components, external window frames in the space shuttle, front covers of slot machines, and architectural siding for buildings. *See* METAL FORMING. [C.H.Ha.]

Superplasticity The unusual ability of some metals and alloys to elongate uniformly thousands of percent at elevated temperatures, much like hot polymers or glasses. Under normal creep conditions, conventional alloys do not stretch uniformly, but form a necked-down region and then fracture after elongations of only 100% or less. The most important requirements for obtaining superplastic behavior include a very small metal grain size, a well-rounded (equiaxed) grain shape, a deformation temperature greater than one-half the melting point, and a slow deformation rate. *See* ALLOY; CREEP (MATERIALS); EUTECTICS.

Superplasticity is important to technology primarily because large amounts of deformation can be produced under low loads. Thus, conventional metal-shaping processes (for example, rolling, forging, and extrusion) can be conducted with smaller, and cheaper equipment. Nonconventional forming methods can also be used; for instance, vacuum-forming techniques, borrowed from the plastics industry, have been applied to sheet metal to form car panels, refrigerator door linings, and TV chassis parts. *See* METAL FORMING; PLASTICITY. [E.E.U.]

Superposition principle The principle, obeyed by many equations describing physical phenomena, that a linear combination of the solutions of the equation is also a solution.

An effect is proportional to a cause in a variety of phenomena encountered at the level of fundamental physical laws as well as in practical applications. When this is true, equations which describe such a phenomenon are known as linear, and their solutions obey the superposition principle. Thus, when $f, g, h, \cdots$, solve the linear equation, then s ($s = \alpha f + \beta g + \gamma h + \cdots$, where $\alpha, \beta, \gamma, \cdots$, are coefficients) also satisfies the same equation. *See* LINEAR ALGEBRA; LINEARITY.

For example, an electric field is proportional to the charge that generates it. Consequently, an electric force caused by a collection of charges is given by a superposition—a vector sum—of the forces caused by the individual charges. The same is true for the magnetic field and its cause—electric currents. Each of these facts is connected with the linearity of Maxwell's equations, which describe electricity and magnetism. *See* ELECTRIC FIELD; MAXWELL'S EQUATIONS.

The superposition principle is important both because it simplifies finding solutions to complicated linear problems (they can be decomposed into sums of solutions of simpler problems) and because many of the fundamental laws of physics are linear. Quantum mechanics is an especially important example of a fundamental theory in which the superposition principle is valid and of profound significance. This property has proved most useful in studying implications of quantum theory, but it is also a source of the key conundrum associated with its interpretation.

Its effects are best illustrated in the double-slit superposition experiment, in which the wave function representing a quantum object such as a photon or electron can propagate toward a detector plate through two openings (slits). As a consequence of the superposition principle, the wave will be a sum of two wave functions, each radiating from its respective slit. These two waves interfere with each other, creating a pattern of peaks and troughs, as would the waves propagating on the surface of water in an analogous experimental setting. However, while this pattern can be easily understood for the normal (for example, water or sound) waves resulting from the collective motion of vast numbers of atoms, it is harder to understand its origin in quantum mechanics, where the wave describes an individual quantum, which can be detected, as a single particle, in just one spot

along the detector (for example, photographic) plate. The interference pattern will eventually emerge as a result of many such individual quanta, each of which apparently exhibits both wave (interference-pattern) and particle (one-by-one detection) characteristics. This ambivalent nature of quantum phenomena is known as the wave-particle duality. *See* INTERFERENCE OF WAVES; QUANTUM MECHANICS. [W.H.Z.]

Superposition theorem (electric networks) Essentially, that it is permissible, if there are two or more sources of electromotive force in a linear electrical network, to compute at any element of the network the response of voltage or of current that results from one source alone, and then the response resulting from another source alone, and so on for all sources, and finally to compute the total response to all sources acting together by adding these individual responses.

Thus, if a load of constant resistance is supplied with electrical energy from a linear network containing two batteries, two generators, or one battery and one generator, it would be correct to find the current that would be supplied to the load by one source (the other being reduced to zero), then to find the current that would be supplied to the load by the second source (the first source now being reduced to zero), and finally to add the two currents so computed to find the total current that would be produced in the load by the two sources acting simultaneously.

By means of the principle of superposition, effects are added instead of causes. This principle seems so intuitively valid that there is far greater danger of applying superposition where it is incorrect than of failing to apply it where it is correct. It must be recognized that for superposition to be correct the relation between cause and effect must be linear. [H.H.Sk.]

Supersaturation A solution is at the saturation point when dissolved solute in its crystallizes from it at the same rate at which it dissolves. Under prescribed experimental conditions of temperature and pressure, a solution can contain at saturation only one fixed amount of dissolved solute. However, it is possible to prepare relatively stable solutions which contain a quantity of a dissolved solute greater than that of the saturation value provided solute phase is absent. Such solutions are said to be supersaturated. They can be prepared by changing the experimental conditions of a system so that greater solubility is obtained, perhaps by heating the solution, and then carefully returning the system to or near its original state. The addition of solute phase will immediately relieve supersaturation. Solutions in which there is no spontaneous formation of solute phase for extended periods of time are said to be metastable. There is no sharp line of demarcation between an unstable and metastable solution. The process whereby initial aggregates within a supersaturated solution develop spontaneously into particles of new stable phase is known as nucleation. The greater the degree of supersaturation, the greater will be the number of nuclei formed. *See* NUCLEATION; PHASE EQUILIBRIUM; PRECIPITATION (CHEMISTRY). [L.Go./R.W.Mu.]

Supersonic diffuser A passive compressor (or shaped duct) in which gas enters at a velocity greater than the speed of sound, is decelerated in a contracting section, and reaches sonic speed at a throat.

Supersonic compression systems can be categorized to three basic types (see illustration). External-compression inlets have the supersonic diffusion taking place at or ahead of the cowl lip (or throat station) and generally employ one or more oblique waves ahead of the normal shock. Internal-compression inlets accomplish supersonic diffusion internally downstream of the cowl lip. Deceleration of the flow is produced by a number of weak reflecting waves in a gradually convergent channel. The

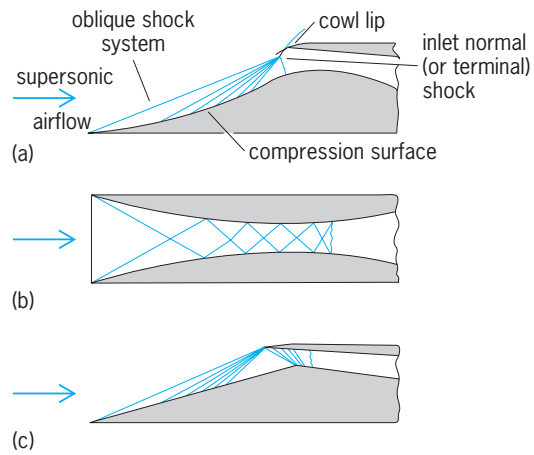

Basic supersonic compression systems. (*a*) External compression. (*b*) Internal compression. (*c*) Combination of external and internal compression.

third system is a combination of external and internal compression and appears to represent an effective compromise. [J.F.C.; L.J.O.]

Supersonic flight Relative motion of a solid body and a gas at a velocity greater than that of sound propagation under the same conditions. The general characteristics of supersonic flight can be understood by considering the laws of propagation of a disturbance or pressure impulse, in a compressible fluid.

If the fluid is at rest, the pressure impulse propagates uniformly with the velocity of sound in all directions, the effect always acting along an ever-increasing spherical surface. If, however, the source of the impulse is placed in a uniform stream, the impulse will be carried by the stream simultaneously with its propagation at sonic velocity relative to the stream. Hence the resulting propagation is faster in the direction of the stream and slower against the stream. If the velocity of the stream past the source of disturbance is supersonic, the effect of the impulse is restricted to a cone whose vertex is the source of the impulse and whose vertex angle decreases from 90° (corresponding to Mach number equal to 1) to smaller and smaller values as the Mach number of the stream increases (see illustration). If the source of the pressure impulse travels through the air at rest, the conditions are analogous. *See* MACH NUMBER.

Consider the supersonic motion of a wing moving into air at rest. Because signals cannot propagate ahead of the wing, the presence of the wing has no effect on the undisturbed air until the wing passes through it. Hence there must be an abrupt change in the properties of the undisturbed air as it begins to flow over the wing. This abrupt change takes place in a shock wave which

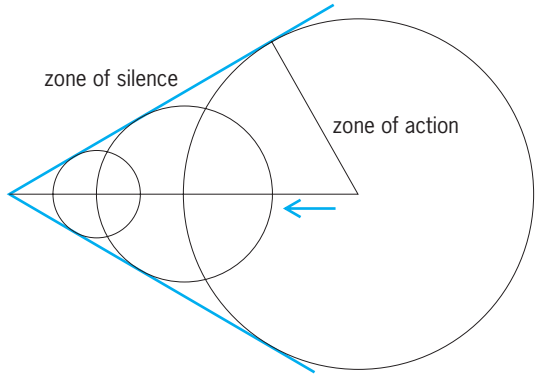

Generation of Mach wave by body at supersonic velocity; zones of action and silence are separated.

is attached to the leading edge of the wing, provided that the leading edge is sharp and the flight Mach number is sufficiently large. As the air passes through the shock wave, its pressure, temperature, and density are markedly increased.

Further aft of the leading edge, the pressure of the air is decreased as the air expands over the surface of the wing. Hence the pressure acting on the front part of the wing is higher than the ambient pressure, and the pressure acting on the rear part of the wing is lower than the ambient pressure. The pressure difference between front and rear parts produces a drag, even in the absence of skin friction and flow separation. The wing produces a system of compression and expansion waves which move with it. This phenomenon is similar to that of a speedboat moving with a velocity greater than the velocity of the surface waves. Because of this analogy, supersonic drag is called wave drag. It is peculiar to supersonic flight, and it may represent the major portion of the total drag of a body. *See* HYPERSONIC FLIGHT; SUBSONIC FLIGHT; TRANSONIC FLIGHT. [J.E.Sc.]

Supersonic flow Fluid motion in which the Mach number M, defined as the speed of the fluid relative to the sonic speed in the same medium, is more than unity. It is, however, common to call the flow transonic when $0.8 < M < 1.4$, and hypersonic when $M > 5$. *See* MACH NUMBER.

Mach waves. A particle moving in a compressible medium, such as air, emits acoustic disturbances in the form of spherical waves. These waves propagate at the speed of sound ($M = 1$). If the particle moves at a supersonic speed, the generated waves cannot propagate upstream of the particle. The spherical waves are enveloped in a circular cone called the Mach cone. The generators of the Mach cone are called Mach lines or Mach waves.

Shock waves. When a fluid at a supersonic speed approaches an airfoil (or a high-pressure region), no information is communicated ahead of the airfoil, and the flow adjusts to the downstream conditions through a shock wave. Shock waves propagate faster than Mach waves, and the flow speed changes abruptly from supersonic to less supersonic or subsonic across the wave. Similarly, other properties change discontinuously across the wave. A Mach wave is a shock wave of minimum strength. A normal shock is a plane shock normal to the direction of flow, and an oblique shock is inclined at an angle to the direction of flow. The velocity upstream of a shock wave is always supersonic. Downstream of an oblique shock, the velocity may be subsonic resulting in a strong shock, or supersonic resulting in a weak shock. The downstream velocity component normal to any shock wave is always subsonic. There is no change in the tangential velocity component across the shock.

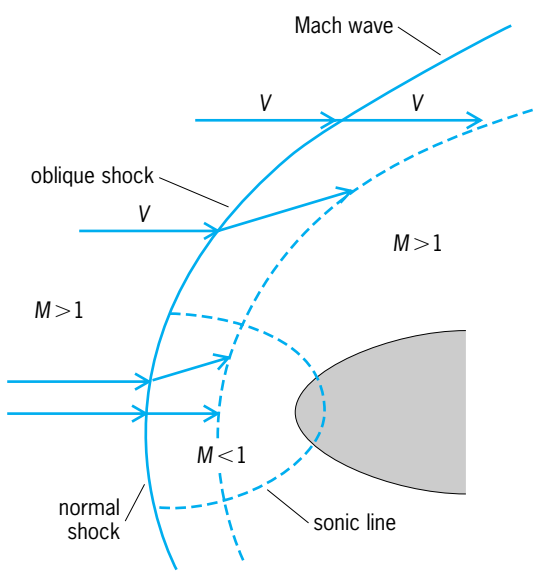

Typical normal shock, oblique shock, and Mach wave pattern in supersonic flow past a blunt body. *M* is the Mach number and *V* is the particle speed. The curved line parallel to normal and oblique shock waves indicates the end of the velocity vectors.

In a two-dimensional supersonic flow around a blunt body (*see* illustration), a normal shock is formed directly in front of the body, and extends around the body as a curved oblique shock. At a sufficient distance away, the flow field is unaffected by the presence of the body, and no discontinuity in velocity occurs. The shock then reduces to a Mach wave. *See* COMPRESSIBLE FLOW; FLUID FLOW; SUPERSONIC FLIGHT. [M.A.S.]

Superstring theory A proposal for a unified theory of all interactions, including gravity. At present, the strong, weak, and electromagnetic interactions are accounted for within the framework of the standard model. This model correctly describes experiments up to the highest energies performed so far, and gives a complete description of the elementary particles and their interactions down to distances of the order of 10^{-18} m. Nevertheless, it has serious limitations, and attempts to overcome them and to unify the forces of nature have been only partly successful. Moreover, these attempts have left standing fundamental difficulties in reconciling gravitation and the laws of quantum mechanics. Superstring theory represents an ambitious program to unify all of the interactions observed in nature, including gravitation, in a theory with no unexplained parameters. In other words, this theory, if successful, should be able to account for all of the particles observed in nature and their interactions. *See* ELEMENTARY PARTICLE; FUNDAMENTAL INTERACTIONS.

String concept. In string theory, the fundamental objects are not point particles, as in standard theories of elementary particles, but one-dimensional extended objects, the open and closed strings. In such a theory, what are usually called the elementary particles are simply particular quantum states of the string. In superstring theories, space-time is ten-dimensional (space is nine-dimensional). If such theories are to describe nature, six dimensions must be "curled up" or "compact." The main consequence of such extra dimensions is the existence of certain very massive particles. *See* SPACE-TIME.

The essential features of string theories can be understood by analogy with the strings of a musical instrument. Such strings vibrate at a characteristic frequency, as well as any integer multiple of that frequency. Each of these modes of vibration (so-called normal modes) can be excited by plucking or striking the string. In classical physics, the amplitudes of vibration of each mode can take on a continuum of values. If there were a string of atomic dimensions, subject to the laws of quantum mechanics, the energies of this quantum string could take on only discrete values, corresponding to particular quantum states. *See* QUANTUM MECHANICS; VIBRATION.

The strings of superstring theory are quite similar. The main difference is that they obey Einstein's principles of special relativity. As a result, since each quantum state has a particular energy, it has a definite mass. Thus, each state of the string behaves as a particle of definite mass. Because it is possible, in principle, to pump an arbitrarily large amount of energy into the string, the theory contains an infinity of different types of particles of arbitrarily large mass. The interactions of these particles are governed by the ways in which the strings themselves interact. To be consistent with the principles of relativity, a string can interact only by splitting into two strings or by joining together with another string to form a third string. As a result, the interactions of strings are nearly unique. This geometric picture of string interactions translates into a precise set of rules for calculating the interaction of individual string states, that is, particles. *See* RELATIVITY.

Classical solutions. Obtaining a description of superstring theory analogous to quantum field theory is an active topic of research. However, even though the equations that describe this

field theory are not completely known at present, it is known how to find classical solutions of these equations, and by various techniques, an enormous number of such solutions have been found. These include states in which space-time has any dimension between one and ten, and states with many bizarre symmetries and spectra. Each of these solutions then corresponds to a possible ground state of the system. The theories built around some of these states look very much like the real world. Not only are four dimensions flat while six are compact, but they possess gauge symmetries close to that of the standard model. Some have three or four generations of quarks and leptons, as well as light Higgs particles, which are of crucial importance in the standard model. Many of these solutions possess space-time supersymmetry. *See* GAUGE THEORY; HIGGS BOSON; LEPTON; QUARKS.

However, if the theory does describe nature, it must have some mechanism that chooses one of the possible ground states. Because the masses and couplings of the elementary particles depend only on the choice of ground state, determining this true ground state will yield a set of predictions for these quantities. If string theory is a correct theory, these predictions must agree with the experimental values. [M.Di.]

Supersymmetry A conjectured enhanced symmetry of the laws of nature that would relate two fundamental observed classes of particles, bosons and fermions.

All particles can be classified as fermions, such as the electron and quarks, or bosons, such as the photon and graviton. A fundamental characteristic distinguishing these two classes is that they carry different quantum-mechanical spin. If the amount of spin of an elementary particle is measured in terms of the fundamental quantum unit of angular momentum—$\hbar$, equal to Planck's constant divided by 2π—then bosons always have integer amounts of spin (that is, 0, 1, 2 . . .), while fermions have odd half-integer amounts of spin (that is, 1/2, 3/2, 5/2, . . .). *See* SPIN (QUANTUM MECHANICS).

There is seemingly a fundamental distinction between particles with differing amounts of spin. For example, bosons like to act collectively (Bose-Einstein statistics), producing such distinctive behavior as the laser, while, conversely, fermions obey the Pauli exclusion principle (and the Pauli-Dirac statistics), which disallows two identical fermions to be in the same state, and explains the stability of matter. Moreover, all the symmetries that are observed in the world relate different particles of the same spin. *See* BOSE-EINSTEIN STATISTICS; FERMI-DIRAC STATISTICS; QUANTUM STATISTICS; SYMMETRY LAWS (PHYSICS).

In contrast, supersymmetry would relate bosons and fermions. This would be a remarkable step forward in understanding the physical world. However, if supersymmetry were realized as an exact symmetry, the particles so related should have almost all their characteristics, such as mass and charge, preserved. Explicitly, any fermion of spin 1/2 should have a boson superpartner of spin 0, while any gauge boson of spin 1 should have a fermion superpartner of spin 1/2. This is apparently a disaster for the idea of supersymmetry since it predicts, for instance, that there should exist a spin-0 boson partner of the electron, the selectron, with electric charge and mass equal to that of the electron. Such a particle would be easy to detect and is certainly ruled out by very many experiments.

The crucial caveat to this negative result is the condition that supersymmetry be realized as an exact symmetry. A fundamental concept of modern physics is spontaneously broken symmetry. Physics displays many examples of symmetries that are exact symmetries of the fundamental equations describing a system, but not of their solutions. In particle physics the spontaneous breaking of a symmetry usually results in a difference in the masses of the particles related by the symmetry; the amount of breaking can be quantified by this mass difference. *See* SYMMETRY BREAKING.

If supersymmetry is broken by a large amount, then all the superpartners have masses much greater than the particles that are currently observed, and there is little hope of seeing evidence for supersymmetry. However, evidence that supersymmetry is broken by only a moderate amount comes from examination of the properties of the fundamental forces at high energy.

Of the four fundamental forces, the three excluding gravity are very similar in their basic formulation; they are all described by gauge theories, generalizations of the quantum theory of electromagnetism, and quantum electrodynamics (QED). The strength of electrical interaction between two electrons can be quantified in terms of a number, the coupling constant α_1. However, the quantity α_1 is actually not a constant, but depends on the energies at which the interaction strength is measured. In fact, the interaction strengths, α_1, α_2, and α_3, of the three forces (excluding gravity) all depend on energy, μ. The couplings $\alpha_{1,2,3}$ satisfy differential equations—renormalization group equations—that depend on the types of elementary particles that exist with mass at or below the energy scale μ and that are charged with respect to each of the three interactions. If the fundamental particles include not only the observed particles but also their superpartners, taken to have masses not greater than 1000 GeV heavier than their (observed) partners, then from the renormalization group equations, the couplings α_i are predicted to meet (unify) at a huge energy of 2×10^{16} GeV. In contrast, if either supersymmetry is not an underlying symmetry of the world, or it is very badly broken so that the superpartners are very massive, the couplings fail to unify at a single point. *See* FUNDAMENTAL INTERACTIONS; GAUGE THEORY; QUANTUM ELECTRODYNAMICS; RENORMALIZATION.

Although the unification of couplings is the most significant indication that supersymmetry is a new law of nature, there are a number of other hints in this same direction. By observing the large-scale motions of the galaxies, the average density of large volumes of the universe can be deduced, resulting in a value that is substantially greater than that directly observed in luminous matter such as stars and hot gas. Therefore, a substantial fraction of the mass of the universe must be composed of some form of nonluminous or dark matter. Remarkably, many attractive models of supersymmetry predict that the lightest of all the superpartners is a weakly interacting massive particle with just the right characteristics to be this dark matter. *See* COSMOLOGY; UNIVERSE; WEAKLY INTERACTING MASSIVE PARTICLE (WIMP). [J.M.R.]

Suppression (electricity) The process or technique of reducing electrical interference to acceptable levels or to situations having no adverse effect. Suppression techniques may be applied to the interference source, the intervening path, the victim or receptor, or any combination. Normal strategy for interference control is to first suppress the source, if possible, since it may disturb many victims.

For intentional transmitters, suppressing interference may include reducing or eliminating harmonic radiations, restricting the bandwidth, or restricting levels of unnecessary or excessive modulation sidebands. These are usually accomplished by radio frequency filters. *See* ELECTRIC FILTER.

For many devices involving incidental radiators, such as brush-type motors and fluorescent lights, interference suppression may require both filtering and shielding. Electrical filtering may take the form of transient surge suppressors, feed-through capacitors, electromagnetic interference (EMI) filters, ferrite absorbers, isolation transformers, or Faraday shielded isolation transformers. *See* ELECTRIC PROTECTIVE DEVICES.

Shielding to control radiation involves using metal boxes, cases, cabinet housings, or metalized plastic versions thereof. Since the interconnecting cables between equipment offer the greatest threat as an "antenna farm," the dominant suppression technique is to shield the cables, wires, or harnesses. For the best protection from electromagnetic interference, the cable shield should be designed as an extension of the box or equipment shield. Other forms of interference hardening of cables include twisting parallel wire pairs, multiple-layer shields, and

absorbing jackets. Where possible, best electromagnetic compatibility performance is obtained by replacing the signal cables with fiber-optic links. *See* ELECTRICAL SHIELDING.

From a system point of view, the radiated emission propensity or pickup susceptibility of interconnected equipment is proportional to the maximum length or dimension and frequency until this length corresponds to an electrical length of about one-half wavelength. Thus, to help suppress electromagnetic interference, the corresponding frequency should not fall within the passband of victims or receptors. *See* ELECTRICAL INTERFERENCE; ELECTROMAGNETIC COMPATIBILITY. [D.R.J.W.]

Supramolecular chemistry A highly interdisciplinary field covering the chemical, physical, and biological features of complex chemical species held together and organized by means of intermolecular (noncovalent) bonding interactions. *See* CHEMICAL BONDING; INTERMOLECULAR FORCES.

When a substrate binds to an enzyme or a drug to its target, and when signals propagate between cells, highly selective interactions occur between the partners that control the processes. Supramolecular chemistry is concerned with the study of the basic features of these interactions and with their implementation in biological systems as well as in specially designed nonnatural ones. In addition to biochemistry, its roots extend into organic chemistry and the synthetic procedures for receptor construction, into coordination chemistry and metal ion-ligand complexes, and into physical chemistry and the experimental and theoretical studies of interactions. *See* BIOINORGANIC CHEMISTRY; ENZYME; LIGAND; PHYSICAL ORGANIC CHEMISTRY; PROTEIN.

The field started with the selective binding of alkali metal cations by natural as well as synthetic macrocyclic and macropolycyclic ligands, the crown ethers and cryptands. This led to the emergence of molecular recognition as a new domain of chemical research that, by encompassing all types of molecular components and interactions as well as both oligo and polymolecular entities, became supramolecular chemistry. It underwent rapid growth with the development of synthetic receptor molecules of numerous types for the strong and selective binding of cationic, anionic, or neutral complementary substrates of organic, inorganic, or biological nature by means of various interactions (electrostatic, hydrogen bonding, van der Waals, and donor-acceptor). Molecular recognition implies the (molecular) storage and the (supramolecular) retrieval and processing of molecular structural (geometrical and interactional) information. *See* HYDROGEN BOND; MACROCYCLIC COMPOUND; MOLECULAR RECOGNITION.

Many types of receptor molecules have been explored (crown ethers, cryptands, cyclodextrins, calixarenes, cavitands, cyclophanes, cryptophanes, and so on), and many others may be imagined for the binding of complementary substrates of chemical or biological significance. They allow, for instance, the development of substrate-specific sensors or the recognition of structural features in biomolecules (for example, nucleic acid probes, affinity cleavage reagents, and enzyme inhibitors). *See* BIOPOLYMER; CYCLOPHANE; ENZYME INHIBITION.

A major step in the development of supramolecular chemistry over the last 20 years involved the design of systems capable of spontaneously generating well-defined, supramolecular entities by self-assembly under a given set of conditions.

The information necessary for supramolecular self-assembly to take place is stored in the components, and the program that it follows operates via specific interactional algorithms based on binding patterns and molecular recognition events. Thus, rather than being preorganized, constructed entities, these systems may be considered as self-organizing, programmed supramolecular systems.

Self-assembly and self-organization have recently been implemented in numerous types of organic and inorganic systems. By clever use of metal coordination, hydrogen bonding, and donor-acceptor interactions, researchers have achieved the spontaneous formation of a variety of novel and intriguing species such as inorganic double and triple helices termed helicates, catenates, threaded entities (rotaxanes), cage compounds, grids of metal ions, and so on.

Another major development concerns the design of molecular species displaying the ability to perform self-replication, based on components containing suitable recognition groups and reactive functions. Self-recognition processes involve the spontaneous selection of the correct partner(s) in a self-assembly event—for instance, the correct ligand strand in helicate formation.

A major area of interest is the design of supramolecular devices built on photoactive, electroactive, or ionoactive components, operating respectively with photons, electrons, or ions. Thus, a variety of photonic devices based on photoinduced energy and electron transfer may be imagined. Molecular wires, ion carriers, and channels facilitate the flow of electrons and ions through membranes. Such functional entities represent entries into molecular photonics, electronics, and ionics, which deal with the storage, processing, and transfer of materials, signals, and information at the molecular and supramolecular levels. Dynamic and mechanical devices exploit the control of motion within molecular and supramolecular entities. *See* INORGANIC PHOTOCHEMISTRY; ION TRANSPORT; PHOTOCHEMISTRY.

The design of systems that are controlled, programmed, and functionally self-organized by means of molecular information contained in their components represents new horizons in supramolecular chemistry and provides an original approach to nanoscience and nanotechnology. In particular, the spontaneous but controlled generation of well-defined, functional supramolecular architectures of nanometric size through self-organization—supramolecular nanochemistry—represents a means of performing programmed engineering and processing of nanomaterials. It offers a powerful alternative to the demanding procedures of nanofabrication and nanomanipulation, bypassing the need for external intervention. A rich variety of architectures, properties, and processes should result from this blending of supramolecular chemistry with materials science. *See* NANOCHEMISTRY; NANOTECHNOLOGY. [J.M.Le.]

Surface (geometry) A two-dimensional geometric figure (a collection of points) in three-dimensional space. The simplest example is a plane—a flat surface. Some other common surfaces are spheres, cylinders, and cones, the names of which are also used to describe the three-dimensional geometric figures that are enclosed (or partially enclosed) by those surfaces. In a similar way, cubes, parallelepipeds, and other polyhedra are surfaces. *See* CUBE; POLYHEDRON; SOLID (GEOMETRY).

Any bounded plane region has a measure called the area. If a surface is approximated by polygonal regions joined at their edges, an approximation to the area of the surface is obtained by summing the areas of these regions. The area of a surface is the limit of this sum if the number of polygons increases while their areas all approach zero. *See* AREA; CALCULUS; INTEGRATION; PLANE GEOMETRY; POLYGON.

Methods of description. The shape of a surface can be described by several methods. The simplest is to use the commonly accepted name of the surface, such as sphere or cube. In mathematical discussions, surfaces are normally defined by one or more equations, each of which gives information about a relationship that exists between coordinates of points of the surface, using some suitable coordinate system. *See* COORDINATE SYSTEMS.

Some surfaces are conveniently described by explaining how they might be formed. If a curve, called the generator in three-dimensional space, is allowed to move in some manner, then each position the generator occupies during this motion is a collection of points, and the set of all such points constitutes a surface that can be said to be swept out by the generator. In particular, if the generator is a straight line, a ruled surface is formed. If the generator is a straight line and the motion is such

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

18 SUR-TYP

www.MHEST.com

S

Surface (geometry) — Syzygy

Surface (geometry)

A two-dimensional geometric figure (a collection of points) in three-dimensional space. The simplest example is a plane—a flat surface. Some other common surfaces are spheres, cylinders, and cones, the names of which are also used to describe the three-dimensional geometric figures that are enclosed (or partially enclosed) by those surfaces. In a similar way, cubes, parallepipeds, and other polyhedra are surfaces. Often only the context indicates whether a surface or a solid is being referred to, but in modern mathematical usage such words refer only to surfaces. *See* CUBE; POLYHEDRON; SOLID (GEOMETRY).

Any bounded plane region has a measure called the area. If a surface is approximated by polygonal regions joined at their edges, an approximation to the area of the surface is obtained by summing the areas of these regions. The area of a surface is the limit of this sum if the number of polygons increases while their areas all approach zero. *See* AREA; CALCULUS; INTEGRATION; PLANE GEOMETRY; POLYGON.

Methods of description. The shape of a surface can be described by any of several methods. The simplest is to use the commonly accepted name of the surface, such as sphere or cube, if such a name exists. Information about a surface's shape might be only partially conveyed by the name, other information being necessary for a complete description. In mathematical discussions, surfaces are normally defined by one or more equations, each of which gives information about a relationship that exists between coordinates of points of the surface, using some suitable coordinate system. *See* COORDINATE SYSTEMS.

An equation (or equations) that defines a surface may have any of several forms. Relative to a three-dimensional rectangular coordinate system, a surface might be defined implicitly by an equation such as $F(x,y,z) = 0$ or explicitly by an equation such as $z = f(x,y)$ [here z was chosen as the dependent variable] or by parametric equations $x = x(u,v)$, $y = y(u,v)$, and $z = z(u,v)$, in which u and v are independent variables. Some surfaces are best defined by using a cylindrical or spherical coordinate system. *See* ANALYTIC GEOMETRY.

Some surfaces are conveniently described by explaining how they might be formed. If a curve, called the generator in R_3 (three-dimensional space), is allowed to move in some manner, then each position the generator occupies during this motion is a collection of points, and the set of all such points constitutes a surface that can be said to be swept out by the generator. In particular, if the generator is a straight line, a ruled surface is formed. If the generator is a straight line and the motion is such that all positions of the generator are parallel, a cylindrical surface (or just cylinder) is formed. A plane, therefore, is a ruled surface as well as a cylinder. If the generator is a straight line and all positions of the generator have a common point of intersection, a conical surface (or just cone) is formed. (To restrict the generator of a cylinder or cone from filling all

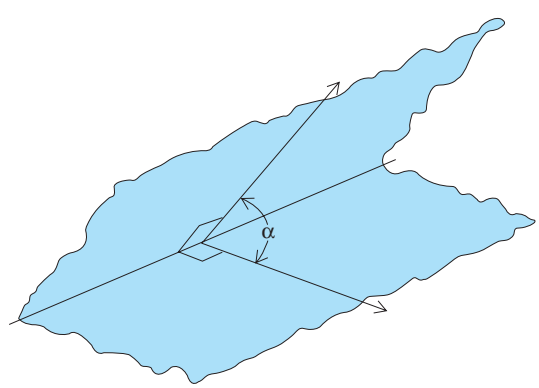

Fig. 1. Part of a dihedron (the rest has been torn away), showing dihedral angle.

of space, another condition is imposed: The generator must always contain a point of some particular curve in space.) A ruled surface that could be bent to lie in a plane (the bending to take place without stretching or tearing) is called a developable surface. Examples of developable surfaces include cylinders and cones, as well as other types. *See* CONE; CYLINDER.

Dihedron. A dihedron is the surface formed by bending a plane along a line in that plane. More formally, a dihedron is the union of two half-planes that share the same boundary line (**Fig. 1**).

A third plane, perpendicular to the boundary line, intersects the dihedral, forming an angle called the dihedral angle. Two lines that intersect generally form two supplementary pairs of equal-measure angles. The three-dimensional analog is the intersection of two planes, which define four dihedrals. Each of the four dihedral angles is either equal to, or supplementary to, the other three dihedral angles.

Quadric surfaces. A surface whose implicit equation $F(x,y,z) = 0$ is second degree is a quadric surface (**Fig. 2**), a three-dimensional analog of a conic section. A plane section of a quadric surface is either a conic section or one of its degenerate forms (a point, a line, parallel lines, or intersecting lines). With the proper choice of a rectangular coordinate system, an equation describing a quadric surface can have one of several basic forms (see **table**). *See* CONIC SECTION; QUADRIC SURFACE.

Surfaces of revolution. When a plane curve (the generator) is revolved about a line in that plane (the axis of revolution, or just axis), a surface of revolution can be said to be swept out (**Fig. 3**). The resulting surface will be symmetric about the axis of

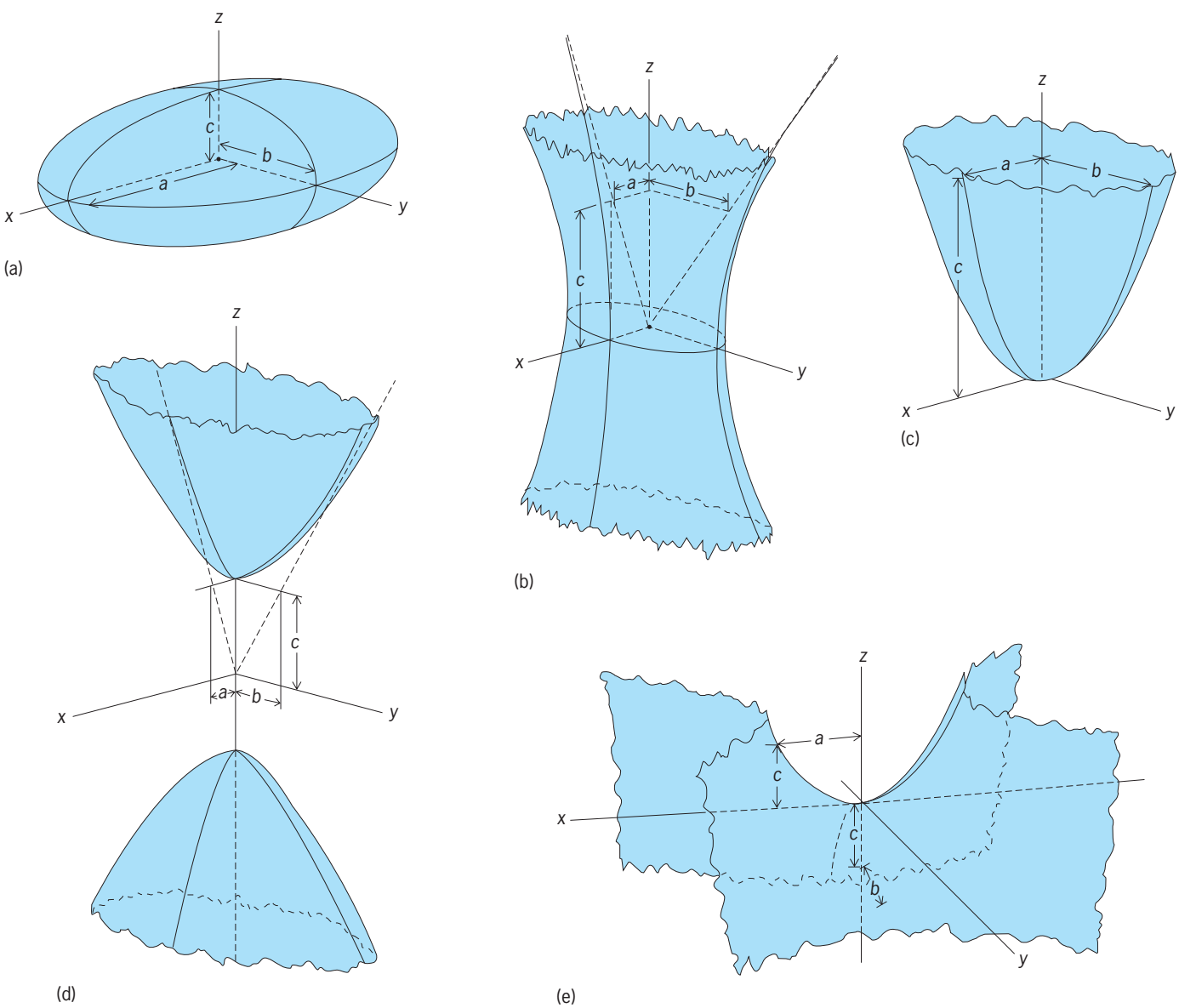

Fig. 2. Some quadric surfaces, with parameters *a*, *b*, and *c*. (*a*) Ellipsoid; $0 < c < b < a$. (*b*) Hyperboloid of one sheet; $0 < a < b < c$. (*c*) Elliptic paraboloid; $0 < a < b$ and $c > 0$. (*d*) Hyperboloid of two sheets; $0 < a < b < c$. (*e*) Hyperbolic paraboloid; $c > 0$.

Quadric surfaces

Surface type	Equation ($a, b, c \neq 0$)	Comments
Ellipsoid	$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$	If $a = b = c$, the surface is a sphere. If any two of a, b, c are equal, the surface is a spheroid.
Hyperboloid of one sheet	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$	A surface of revolution if $a = b$. A ruled surface.
Hyperboloid of two sheets	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$	A surface of revolution if $a = b$.
Elliptic cone	$z^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2}$	A surface of revolution if $a = b$. A ruled, and developable, surface.
Elliptic cylinder	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$	A surface of revolution if $a = b$. A ruled, and developable, surface.
Parabolic cylinder	$y = ax^2$	A ruled, and developable, surface.
Elliptic paraboloid	$\frac{z}{c} = \frac{x^2}{a^2} + \frac{y^2}{b^2}$	A surface of revolution if $a = b$.
Hyperbolic paraboloid	$\frac{z}{c} = \frac{x^2}{a^2} - \frac{y^2}{b^2}$	A ruled surface.

revolution, and this property leads to a more formal definition: A surface of revolution is any collection of points in three-dimensional space that, with a suitably chosen coordinate system, might be represented by an equation (or equations) in cylindrical coordinates (r,θ,z) in which the θ coordinate is absent. In such a representation, the z axis will be the axis of revolution. Planes perpendicular to the axis that intersect a surface of revolution will form parallel circles (or just parallels) whose centers lie on the axis; planes that contain the axis intersect the surface of revolution along congruent curves called meridian sections (or just meridians). If the generator lies in the xz plane (or yz plane) and has the equation $f(x,z) = 0$ [or $f(y,z) = 0$], an equation of the surface of revolution in cylindrical coordinates will be $f(r,z) = 0$.

There exist an unlimited number of surfaces of revolution. Mentioned below are only some of the common ones encountered in pure and applied mathematics.

Circular cylinder. A circular cylinder (a quadric surface) is formed when the generator and the axis of revolution are distinct parallel lines. A circular cylinder is an unbounded surface and so has infinite area.

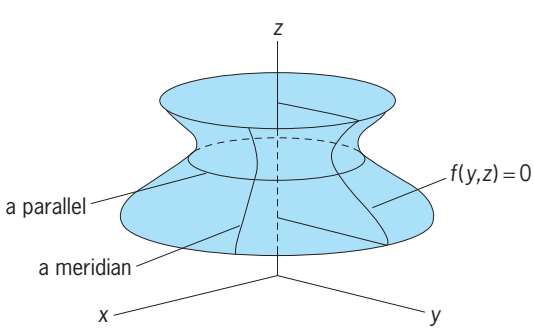

Fig. 3. Surface of revolution. The generating curve (in the *yz* plane) is $f(y,z) = 0$. The axis of revolution is the *z* axis.

However, if the generator is only a segment of a line (rather than the entire line), a bounded circular cylinder is generated. If R is the distance between the segment and the axis (this is called the radius of the cylinder) and h is the length of the segment (the height of the cylinder), then the area A of the surface is given by Eq. (1).

$$A = 2\pi Rh \tag{1}$$

Circular cone. A circular cone is a quadric surface formed when a straight-line generator intersects the axis of revolution at an acute angle. The cone consists of two parts, the nappes, joined at the point of intersection, which is the vertex of the cone. Although a cone is unbounded, a related bounded surface is formed when only a segment of a line is revolved about an intersecting line. If the segment has an end point on the axis, has length L, and makes an acute angle α with the axis, then the area A of this surface is given by Eq. (2).

$$A = \pi L^2 \sin \alpha \tag{2}$$

Sphere. A sphere (a quadric surface) is usually defined as a collection of points in three-dimensional space at a fixed distance (the radius, here denoted by R) from a given point (the center). However, a sphere can also be defined as the surface of revolution formed when a semicircle (or the entire circle) is revolved about its diameter.

The intersection of any plane with a sphere will be a circle (except for tangent planes). Such a circle is called, respectively, a great circle or a small circle, depending on whether or not the plane contains the center of the sphere. The plane area enclosed by a great circle is πR^2, and the area A of the sphere is four times that of a great circle, as in Eq. (3).

$$A = 4\pi R^2 \tag{3}$$

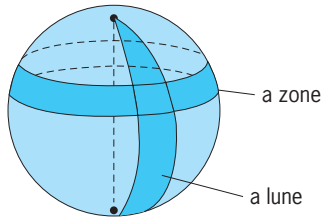

Fig. 4. Sphere, with a zone and a lune.

If only part of a semicircle is revolved about the diameter, a part of a sphere called a zone is formed (**Fig. 4**). This surface will be the same as that part of a sphere between two parallel planes that intersect the sphere. In the distance between these planes is h, the area A_z of the zone is given by Eq. (5).

$$A_z = 2\pi Rh \tag{5}$$

If a semicircle is revolved about its diameter through an angle less than one revolution, the surface swept out is a lune (Fig. 4). If θ is the radian measure of the angle through which the semicircle is rotated, the area A_l of the lune is given by Eq. (6).

$$A_l = 2\theta R^2 \tag{6}$$

See SPHERE.

Spheroid. A spheroid (also called an ellipsoid of revolution) is the quadric surface generated when an ellipse is revolved about either its major or minor axis. If the revolving is about the minor axis of the ellipse, the surface can be thought of as a flattened sphere, called an oblate spheroid. (The Earth, being slightly flattened at the poles, has roughly the shape of an oblate spheroid.) If the revolving is about the major axis, the surface can be thought of as a stretched sphere, called a prolate spheroid. (A watermelon has roughly the shape of a prolate spheroid.) *See* ELLIPSE.

Paraboloid. A circular paraboloid is the quadric surface formed when a parabola is revolved about its axis. Rays emanating from the focus of a parabola will be parallel to the axis after reflection off the parabola. This property is used to advantage by some reflecting telescopes and radio telescopes, which have surfaces with the shapes of circular paraboloids. Incoming rays parallel to the axis of revolution aimed at the concave side of the surface will be brought to a focus at the focus of the paraboloid. *See* PARABOLA; RADIO TELESCOPE.

Hyperboloid. A circular hyperboloid is the quadric surface formed when a hyperbola is revolved about either its transverse axis or its conjugate axis. The surface will be, respectively, a hyperboloid of one sheet (Fig. 2*b*) or two sheets (Fig. 2*d*), depending on whether the revolving is about the conjugate axis or the transverse axis of the hyperbola. *See* HYPERBOLA.

If one of two skew lines (lines in three-dimensional space that are not coplanar) is selected to be the axis of rotation and the other line is the generator, the surface swept out is a hyperboloid of one sheet. (The usual definition of a surface of revolution must be extended, by allowing the axis and the generating curve to lie in different planes.) Every point on a hyperboloid of one sheet is the point of intersection of two lines, each point of which lies on the hyperboloid. Either of these lines (which define equal angles with the axis of revolution) could be used as the generator.

Torus. A torus is generated when a circle is revolved about a line that does not intersect the circle. This doughnut-shaped surface has the property that not all points on the surface have the same sign of curvature. If R is the distance between the axis and the center of the generating circle, then the torus has respectively negative curvature, zero curvature, or positive curvature at points on the torus closer to the axis than R, at a distance R, or farther from the axis than R. *See* DIFFERENTIAL GEOMETRY; TORUS.

Catenoid. A suspended uniform slack rope has the shape of the curve called a catenary. If the axis of revolution is a horizontal line that underlies the suspended rope, then, using the catenary as the generator, the surface of revolution formed is a catenoid (**Fig. 5**). *See* CATENARY.

A soap film tries to attain a shape that minimizes the surface area subject to necessary constraints. An airborne soap bubble is spherical because that is the surface of minimum area that encloses a given volume of air. If two wire circles define the bases of a frustrum of a circular cone, and if the height of the frustrum is small, then when withdrawn from a soapy solution the soap film along the side of the wire-frame frustrum will form a surface of revolution having minimum area: a catenoid. *See* MINIMAL SURFACES.

Pseudosphere. A tractrix is the path followed by an object dragged slowly through a resisting medium by a constant-length cord, where the other end of the cord follows a specified path. If the cord is one unit long, and the end follows a straight-line path perpendicular to the initial position of the cord, and if this curve is revolved about that line, then a pseudosphere is generated. This surface has the property that its curvature is a negative constant at every nonsingular point of the surface. *See* TRACTRIX.

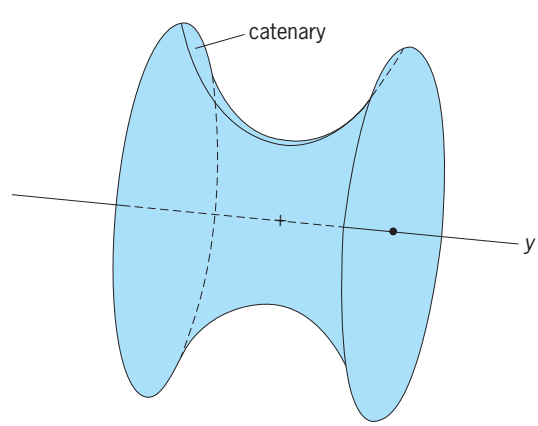

Fig. 5. Catenoid. The axis of revolution is the *y* axis.

Area. If ds is the element of arc length of a generator, and R represents the distance of that element of length from the axis of revolution, then the surface area A is given by Eq. (7), where a and b define

$$A = \int_a^b 2\pi R ds \tag{7}$$

the end points of the generator. When attempting to evaluate an area, the best variable of integration to use is normally determined by how the generator is described relative to the coordinate system chosen.

An alternative method for determining the areas of some surfaces of revolution is by using the surface theorem of Pappus: If a plane curve is revolved about a line that does not intersect the curve (except perhaps at an end point), then the area of the surface of revolution generated is equal to the length of the generating curve multiplied by the distance that the centroid of the curve moves during one revolution. To visualize the centroid of a plane curve, it may be imagined that the curve is made of a uniform thin wire, and a thin massless membrane encloses the wire. If the wire-membrane assembly is held with its containing plane horizontal, then the centroid is the so-called balance point, where the assembly could be balanced on a pencil point without tipping in any direction. It is necessary to know the distance of the centroid from the axis of revolution in order to use this theorem of Pappus, and usually it is just as much work to find this distance as it is to compute the surface area by integration. However, in some cases this distance can be easily found by using symmetry arguments. For example, the surface area of a torus can be computed with this theorem of Pappus by recognizing that the centroid of the generating circle is at the center. If d is the radius of the generating circle, the arc length is $2\pi d$; if R is the distance between the centroid and the axis of revolution, the distance traveled by the centroid during one revolution is $2\pi R$. The surface area A of a torus, therefore, is given by Eq. (8). Areas of other surfaces of revolution can be

$$A = 4\pi^2 dR \tag{8}$$

found by similar arguments. Harry L. Baldwin, Jr.

Bibliography. H. G. Ayre and R. Stephens, *A First Course in Analytic Geometry*, 1956; J. S. Frame, *Solid Geometry*, 1948; R. R. Middlemiss, J. L. Marks, and J. R. Stewart, *Analytic Geometry*, 3d ed., 1968; D. C. Murdoch, *Analytic Geometry, with an Introduction to Vectors and Matrices*, 1966; C. Smith, *An Elementary Treatise on Solid Geometry*, 1910; D. J. Struik, *Lectures on Classical Differential Geometry*, 1950, reprint 1988.

Surface-acoustic-wave devices

Devices that employ surface acoustic waves in the analog processing of electronic signals in the frequency range 10^7–10^9 Hz. Surface acoustic waves are mechanical vibrations that propagate along the surfaces of solids. In 1885, Lord Rayleigh discovered a type of surface acoustic wave that contains both compressional and transverse components 90° out of phase with one another. Since that time, other types of surface acoustic waves have been discovered and are an active area of intense research. A few notable examples include a wave propagating along a layer on a surface (Love wave), a wave propagating along an interface between two solids (Stoneley wave), and transverse guided waves on solids (Bleustein-Gulyaev-Shimizu waves). Love waves are shear-horizontal (SH) waves that have displacement only in a direction perpendicular to the plane of propagation.

Piezoelectric materials provide the required coupling between electrical signals and mechanical motion to generate surface acoustic waves. Crystalline piezoelectric materials, such as quartz, lithium niobate, and lithium tantalate, exhibit low attenuation and dispersion, and are therefore ideal for acoustic propagation. Surface acoustic waves in such a material are generated through a localized electric field at the surface that is created by applying voltage to an array of metal electrodes or fingers. This electrode array is known as an interdigital transducer (IDT). The IDT can also be used to detect surface waves, producing electrical output and hence an overall response. *See* PIEZOELECTRICITY.

Surface-acoustic-wave (SAW) devices have led to a versatile technology for analog signal processing in the frequency range 10^7–10^9 Hz. The much slower propagation velocity of acoustic waves as compared to electromagnetic waves permits time delays in SAW devices, as compared to electrical delay lines, that are crucial for signal processing applications. Notable devices include band-pass filters, resonators, oscillators, pulse compression filters, fast Fourier transform processors, and more recently chemical and biological sensors. Consumer application areas include mobile phones, television and satellite receivers, keyless entry systems (garage doors, cars, and so forth), and wireless applications. Commercial applications include fiber-optic communication, oscillators, local-area networks (LANs), test equipment, and chemical and biological detection systems, with military applications in radar, sonar, and advanced communications. *See* ELECTRIC FILTER; LOCAL-AREA NETWORKS; MOBILE COMMUNICATIONS; OPTICAL-FIBER COMMUNICATIONS; OSCILLATOR; RADAR; SIGNAL PROCESSING; SONAR; TELEVISION RECEIVER.

SAW transduction. A basic SAW device known as a delay line is shown in **Fig. 1***a*. A piezoelectric substrate has a polished upper surface on which two IDTs are deposited using photolithographic methods. The left-hand input transducer is connected, via fine bonded leads, to the electric source (V_s) through an electrical matching network (Z) and source resistance (R_s). The right-hand output transducer drives the load (R_L), usually 50 ohms, through another electrical matching network (Z). Advances in computer modeling have led to the development of 50-ohm IDT designs that do not require external electrical matching networks. The center

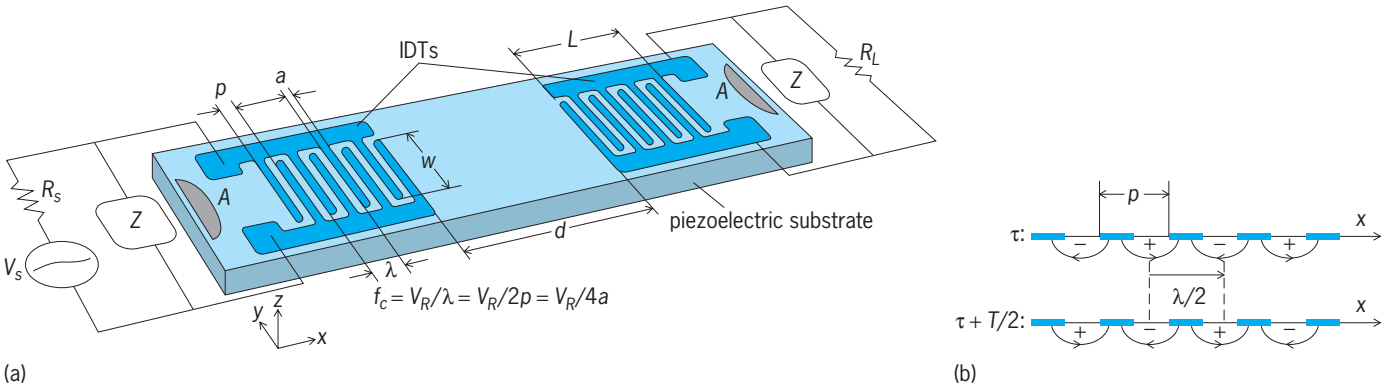

Fig. 1. Operating principle of a surface-acoustic-wave delay line. (*a*) Device layout. (*b*) Cross section of substrate and transducers, showing electric fields at times τ and $\tau + T/2$.

frequency (f_c) is governed by the Rayleigh wave velocity (V_R) on the piezoelectric substrate and the electrode width (a) of a single finger, according to Eq. (1). For SAW devices, the velocity of the wave de-

$$f_c = \frac{V_R}{4a} \tag{1}$$

pends on the properties of the piezoelectric crystal and its crystallographic orientation. Computer models have proven essential to iterate through numerous crystallographic orientations to search for the existence and type of a suitable acoustic wave. *See* DELAY LINE.

For filter applications, the width of the passband, which is governed by length of the IDT (L), is critical for frequency selection. Increasing L by the addition of more electrode pairs sharpens the filter response and reduces noise. IDT-to-IDT spacing (d) is used to select the delay time or phase slope of the filter. The IDT aperture (w) governs the diffraction behavior and determines transducer output power. Owing to symmetry, each transducer generates acoustic waves equally in two opposite directions, so that it is bidirectional. In this case, half of the power propagates in an unwanted direction, giving a loss of −3 dB [that is, $10 \cdot \log_{10}(0.50)$], and in a delay line with two IDTs this propagation contributes to −6 dB of insertion loss in the passband. Though waves are terminated by using an absorber (A), bidirectional emission is undesirable. (Absorber efficiency depends on the acoustic absorption properties of the material and the type of acoustic wave. For Rayleigh SAW devices, silicon rubber is a very effective absorber, reducing the amplitude of the backward-traveling waves by over 30 dB. However, SAW devices that use shear-horizontal waves require unidirectional transducers since such waves are not attenuated by absorbers.) In advanced SAW devices, through the use of unidirectional transducers, acoustic waves propagate preferentially in one direction, which dramatically reduces overall acoustic loss and eliminates the need for absorbers. Insertion losses for unidirectional transducers are around −4 dB or better, depending on the substrate type.

The voltage across the electrodes produces an electric field (Fig. 1*b*), which generates compressions and expansions near the surface. These give rise to various elastic waves. A sinusoidal supply voltage (V_s) produces vibrations that add constructively only if the distance (p) equals half the wavelength (λ). Stresses produced at time τ by a pair of fingers, for a given polarity of the voltage, travel distance $\lambda/2$ during half-period $T/2$ at the speed (V_R) of the Rayleigh wave (Fig. 1*b*). At time $\tau + T/2$, the stress arrives under the neighboring pair of fingers, where the voltage has just changed sign, producing a stress with the same phase. The stress due to the second pair of fingers adds constructively to the first.

The acoustic response at frequency f for the SAW delay line in Fig. 1*a* can be calculated approximately by regarding each IDT as having N electrodes or sources. Through summation of sources for their amplitude and phase, the frequency response can be determined as proportional to $|(\sin x)/x|$ [that is, $|\sin c(x)|$], where x is given by Eq. (2), providing a

$$x = \frac{N\pi(f - f_c)}{f_c} \tag{2}$$

band-pass filter characteristic. The electrical matching networks are normally arranged to minimize filter loss without compromising the acoustic response. The optimum number of periods, N, is inversely proportional to the piezoelectric coupling as is the filter band-pass width. *See* TRANSMISSION LINES.

Numerous IDT geometries have been designed to achieve optimal performance on a wide variety of piezoelectric materials. The simplest designs use single split electrode geometries (**Fig. 2***a*). A slightly more complex design uses a double split electrode configuration (Fig. 2*b*) to reduce reflection effects in the passband. In Fig. 2*a*, waves reflected from corresponding edges of two neighboring fingers of width $\lambda/4$, separated by an interval of $\lambda/4$, add constructively because the path difference causes a phase shift of 2π. The design in Fig. 2*b* suppresses reflections since each finger is composed of two strips of width $\lambda/8$. In this case the distance between two neighboring fingers produces a path difference of $\lambda/2$, resulting in destructive interference. However, the double split electrode configuration demands

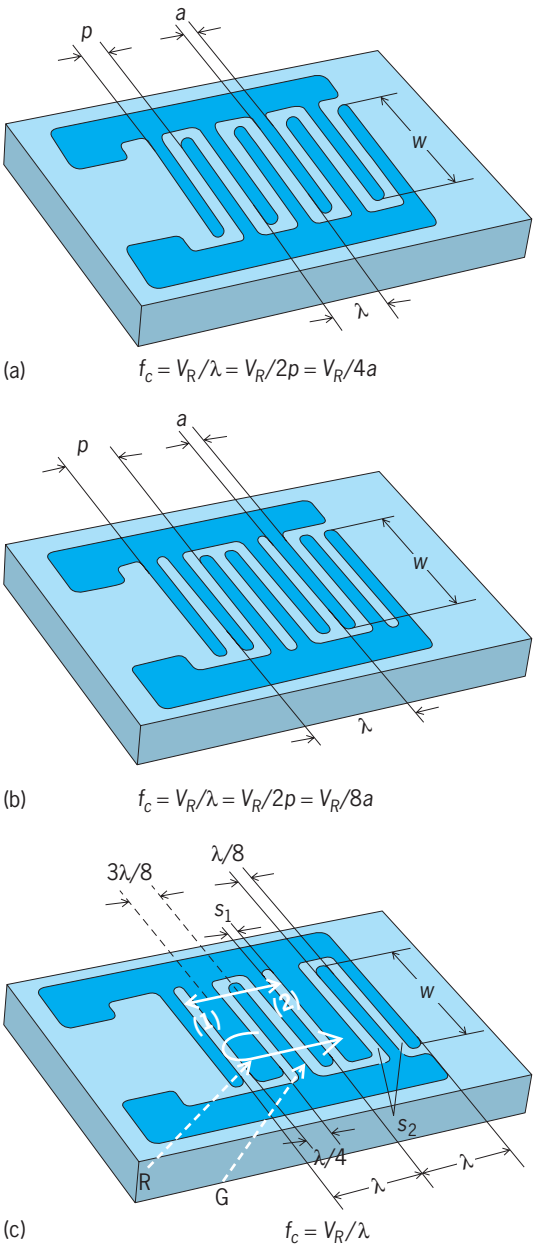

Fig. 2. Transducer arrangements based on (*a*) single split electrode, (*b*) double split electrode, and (*c*) Electrode-Width-Controlled/Single-Phase Unidirectional Transducer (EWC/SPUDT).

a higher photolithographic resolution to fabricate, since the IDT finger width (a) must reduced by a factor of 2 to achieve the same operating frequency as the layout of Fig. 2*a*. For example, for a 500-MHz filter on ST-quartz (where ST is the cut plane of the quartz) having Euler angles of (0°, 132.75°, 0°), with a velocity of $V_R = 3158$ m/s for x propagation in the x direction, the width (a) for the single split IDT is 1.6 μm, whereas for the double split IDT the width (a) reduces to 0.8 μm. [The angles given as (φ, θ, ψ) are referred to as Euler angles and used to designate the piezoelectric crystal cut. Euler angles describe the rotation of the crystal surface referenced to the crystal axes (X, Y, Z). Three main properties of the crystal must be transformed to the rotated axes using Euler angles to determine the SAW propagation behavior. These three material parameters are the elastic, piezoelectric, and permittivity constants of the crystal. The crystal density is not dependent on the rotation. By using the equation of motion and Laplace's equation, the SAW propagation velocity and wave type can be calculated. *See* CRYSTALLOGRAPHY; EULER ANGLES.

Figure 2*c* shows a common unidirectional transducer design that preferentially launches acoustic waves to the right. This IDT uses variable pitch (that is, finger-to-finger spacing) and finger width to achieve a sophisticated response. By using a reflection center (R) located at the middle of the $\lambda/4$ electrodes, a $-90°$ phase shift is introduced. Each generation center (G) launches waves designated (1) and (2) in both directions. By setting the distance between the reflector (R) and the launching electrode (G) to $3\lambda/8$, the waves moving to the left return in phase with the rightward-moving waves. In practice, the spaces s_1 and s_2 and the distance from G to R are optimized using computer simulation methods to achieve the best performance (for example, a highly linear phase response). *See* TRANSDUCER.

Piezoelectric substrates. SAW devices are fabricated using methods found in standard semiconductor processing facilities, such as photolithography, thin-film evaporation, and chemical etching techniques. Common piezoelectric substrates are ST-cut (0°, 132.75°, 0°) quartz, X-propagating for temperature stability; Y-cut Z-propagating (YZ) lithium niobate (0°, 90°, 90°) for high piezoelectric coupling; 128° Y-cut X-propagating lithium niobate (0°, 38°, 0°) for reduced bulk-wave excitation as compared with YZ lithium niobate; and gallium arsenide (GaAs) for compatibility of SAW devices with integrated circuits. *See* INTEGRATED CIRCUITS; MICROLITHOGRAPHY.

The **table** lists some common substrates for Rayleigh and leaky-SAW substrates. Rayleigh waves are called true surface waves since the propagation velocity is a real number. However, leaky waves refer to propagation velocities that are complex numbers, where the imaginary term contributes to additional loss within the crystal. In either case, the coupling efficiency is determined by the electromechanical coupling coefficient K^2, which is a measure of the efficiency for a specific piezoelectric substrate and orientation to convert an applied electrical signal into mechanical motion.

Current commercial photolithography is limited to approximately 0.3–0.5-μm line resolution, which limits operating frequency of Rayleigh-wave devices to about 2 GHz. This limitation has prompted the study of alternative wave-propagation mechanisms for use with existing photolithography. When acoustic waves with a higher velocity are used, the line resolution is not as critical. Leaky-SAW (LSAW) velocities can be much higher than SAW ones to the extent that LSAW devices can be designed for operation at frequencies up to 1.6 times higher than their SAW counterparts using the same photolithographic

Parameters of common piezoelectric substrates for SAW devices

Material	Crystal cut	SAW axis	Velocity, m/s	K^2, %	Temperature coefficient, ppm/°C	Applications
			Rayleigh-wave devices			
Quartz	ST	*X*	3158	0.11	~0	Precision oscillators, low-loss radio-frequency (RF) resonators
Lithium niobate	*Y*	*Z*	3488	4.5	94	Wideband intermediate-freqeuncy (IF) filters
Lithium niobate	128°	*X*	3992	5.3	75	Wideband IF filters
Gallium arsenide	(100)	<110>	2481	0.06	35	Semiconductor integrated circuit compatibility
Bismuth germanium oxide	(110)	<001>	1681	1.4	120	Long delay lines
			Leaky-wave devices			
LST-quartz	ST, $\theta = 15.7°$	*X*	3948	0.11	~0	Gigahertz-range precision oscillators, low-loss RF resonators
Lithium niobate	64°	*X*	4478	11.3	−81	Band-pass RF filters, mobile transceivers
Lithium niobate	41°	*X*	4840	21	−65	Band-pass RF filters, RF resonators
Lithium tantalate	36°	*X*	4112	4.7	−32	Ladder filters for antennas, mobile transceivers

geometry. Leaky surface acoustic waves propagate just beneath the piezoelectric surface, in the form of surface-skimming bulk waves (SSBW) for a free surface (that is, for an electrically open surface), and in the form of shear-horizontal waves for an electrically shorted surface. The type of propagation is a complex function of the crystal type, cut, and symmetry. LSAW devices have found extensive application in radio-frequency-filter devices for mobile and cellular phone systems into the 2-GHz regime. Some common LSAW piezoelectric substrates are LST-quartz (0°, 15.7°, 0°), 64° *YX*-lithium niobate (0°, 64°, 0°), 41° *YX* lithium niobate (0°, 41°, 0°), and 36° *YX* lithium tantalate (0°, 36°, 0°). In **Fig. 3**, an example of an optimized LSAW device is given using 36° *YX* lithium tantalate. This design implements an IDT similar to Fig. 2*c*, and the use of unidirectional transducers enables it to achieve an insertion loss of −4.8 dB at the center frequency, an improvement over the insertion loss of −6 dB when bidirectional IDTs are used.

Band-pass filters. Conventional (lumped-element) linear-phase passive *LC* filters all have some inherent degree of phase nonlinearity over their prescribed frequency range, and this nonlinearity increases with the order of the filter (that is, the number of reactive components). The resultant size of a passive *LC* filter, combined with cost and complexity, is often not suitable for integrated and mobile systems. However, the size of a SAW filter decreases with increasing frequency, with only the limitation of operating as a band-pass filter as opposed to a low-pass filter. In their band-pass operation, SAW filters have much more design versatility than *LC* band-pass filters. In contrast to conventional *LC* filters, SAW filters have the property that linear phase response can be achieved independently of the amplitude response and the amplitude response can be shaped to be asymmetric about its center frequency. Further, SAW filters are transversal filters in which the signal is repetitively delayed and added to itself, as in antenna arrays and digital filters. A design procedure used to model SAW filters is to compute the inverse Fourier transform of the prescribed filter response, giving the impulse response, which is the desired spatial image for the transducer. Due to finite piezoelectric substrate lengths, an infinitely long time-duration impulse response is not realizable. This limitation has necessitated the use of weighting functions to multiply and truncate the impulse response. *See* ANTENNA (ELECTROMAGNETISM); FOURIER SERIES AND TRANSFORMS.

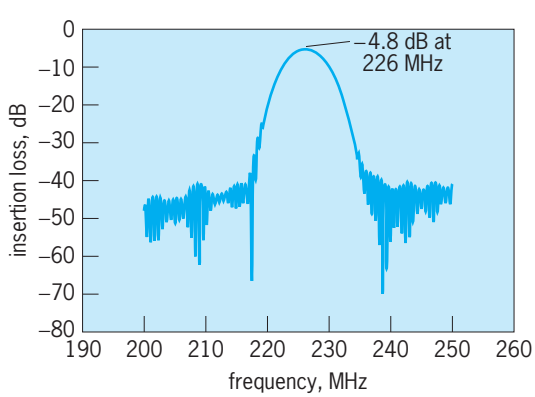

Fig. 3. Insertion loss for leaky-SAW device using 36° *YX* lithium tantalate.

A crucial factor in SAW filters is amplitude weighting, that is, alteration of the amplitude-frequency response to improve filter response. The SAW design in Fig. 1*a* has sidelobe suppression of about −24 dB below the main peak. Instead of using the uniform IDT with constant apodization (that is, finger-length overlap), the overlap of the fingers can be variable. If the apodization of the IDT is changed to a sinc function, the inverse Fourier transform will now approximate a rectangular band-pass response from the filter. This increases the sidelobe suppression to −30 dB,

but produces undesirable amplitude and phase ripple in the passband from the Gibbs phenomena as a result of having finite electrode length for the SAW aperture (w) in Figs. 1 and 2. The use of finite electrodes translates to termination of an infinite series expansion of a mathematical expression used to describe the IDT. Early termination of the infinite series results in the Gibbs phenomena, which occur in all orders of approximation. The result overshoots the original function.

To achieve a rectangular response without using excessively long apodized IDTs, windowing function techniques can be used to modify the IDT apodization pattern. Window functions are widely used in digital filter design to improve the shape of the passband response. Finite impulse response length corresponds to the truncation of an infinite Fourier series. This abrupt truncation of a Fourier series causes the Gibbs ripple phenomena. Window function techniques circumvent the problem of using excessive impulse response lengths by convolution of the IDT sinc function with a chosen window function (such as, Kaiser, Hamming, or cosine). Desirable window functions are those that have a narrow main frequency response and sidelobes that decrease rapidly with frequency. As a result, windowing function methods nearly eliminate passband ripple, while achieving high sidelobe suppression (greater than 45 dB). *See* DIGITAL FILTER.

Resonators. SAW resonators consist of either one or two IDTs bounded by two reflection gratings. One-port SAW resonators employ a single IDT for input and output, in conjunction with two SAW reflection gratings (**Fig 4**). Two-port SAW resonators can be formed using separate IDTs for input and output between two reflection gratings. In one-port SAW resonators, surface waves emitted from both sides of the excited IDT are constructively reflected at the center frequency by reflection gratings (R), which give rise to standing surface acoustic waves within the IDT.

Elements of the reflection grating are both periodically spaced and normal to the SAW propagation direction, to create narrow-band performance. Since the SAW grating reflects surface waves from both sides of an excited bidirectional IDT, the device insertion loss can be less than −6 dB. In practice, the reflection gratings can be either open or shorted metal strips, shallow grooves in the substrate, or shallow grooves filled with a metal such as gold to increase reflection efficiency. The key operation principle for SAW resonators is the phase of the grating reflection coefficient. Acoustic resonance occurs in a SAW resonator when the total phase shift φ of the surface wave is $2n\pi$ ($n = 1,2,\ldots$) within the cavity bounded by the reflection gratings. This is analogous to a microwave resonant cavity, where energy oscillation is between electric and magnetic fields, except in the acoustic case where the oscillation is between mechanical stress and strain fields. The resonator quality factor depends on the position of the IDTs, the reflection efficiency, and the cavity losses from material properties. Both resonant and antiresonant behavior

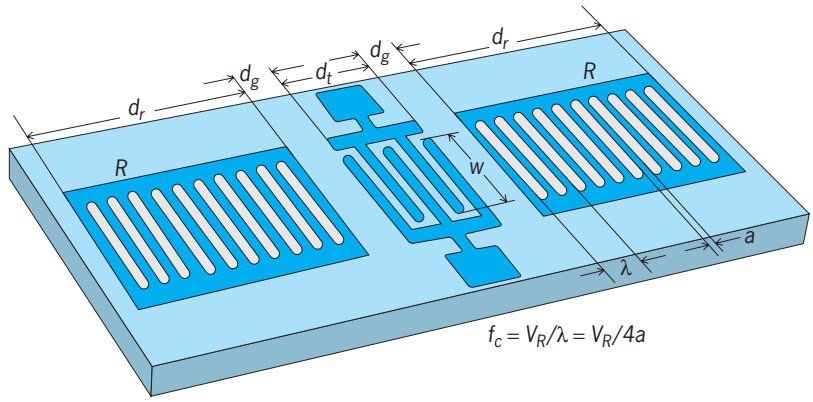

Fig. 4. One-port SAW resonator design. The key dimensional parameters are d_r, d_g, d_t, a, and w. The grating reflectors (R) are designed to create standing waves, forming a resonant cavity.

can be obtained, depending on whether the spacing between the arrays is an even or odd integral number of half-wavelengths of the surface acoustic wave. *See* CAVITY RESONATOR.

In SAW resonator applications, it is crucial to maintain short- and long-term frequency stability. For this reason, special care is required to isolate a SAW device from environmental vibration and temperature change. A critical aspect in packaging SAW resonators is the use of materials that isolate the device from vibration. Packaging methods often employ special support structures to achieve a high degree of vibration isolation. For ST-cut quartz SAW resonators, the vibrational sensitivity is about $\Delta f/f = 1 \times 10^{-9}/g$ [that is, fractional change in frequency per g (9.8 m/s^2 or 32.2 ft/s^2)]. For commercial resonators, the frequency tolerance at 25°C (77°F) ranges from about ±20 ppm to ±300 ppm depending on cost. For one-port SAW resonators in the range 100–1200 MHz, matched insertion loss is in the range 0.5–2.5 dB, with quality factor (Q_L) ~ 1600–7000. *See* VIBRATION ISOLATION.

In precision resonators, the effect of noise is a serious concern, and efforts are made to minimize noise through use of optimized designs and circuits. Flicker ($1/f$) noise is a type of noise that occurs in high-precision oscillators, causing the central frequency peak to broaden as a result of phase noise. One of the most common ways to measure phase noise is with a spectrum analyzer. Noise and spurious signals appear as jitter on either side of the output signal's center frequency on a spectral density plot. On this type of plot, pure frequencies would appear as a single spectral line, but this is not physically possible and instead phase noise causes broadening of the spectral line. Phase noise is specified in decibels relative to the carrier at a given frequency offset from the carrier (dBc/Hz). For example, if we shift 1 Hz from the carrier peak to the right, the dB value describes how far down the noise is compared to the signal. Attainable flicker noise ($1/f$) levels are −140 dBc/Hz at a frequency offset of 1 Hz. Over the same, frequency range, two-port resonator losses are 1–4 dB, and Q_L is ~3000–13,000. Two-port SAW

resonators have been reported at 2.6 and 3.3 GHz, with insertion loss less than 11 dB, $Q_L = 2000$, and insertion loss of 17 dB, $Q_L = 1600$, respectively. *See* SPECTRUM ANALYZER.

For commercial applications, one-port SAW resonators are useful for replacing bulk-wave resonators operating at high-overtone modes and for applications in narrow-band intermediate-frequency (IF) stages for analog cellular phone systems. Two-port resonators are used in oscillator applications for both fixed-frequency and tuning applications. The majority of SAW resonators are Rayleigh-wave designs fabricated on cuts of quartz for temperature stability.

SAW devices for chemical and biological detection. A more recent use of SAW devices is for chemical and biological detection. Since the acoustic energy is confined at the surface of the piezoelectric substrate, surface acoustic waves are highly sensitive to surface perturbations of the propagating medium. The boundary conditions at the solid or liquid interface govern the wave amplitude and velocity, allowing the surface wave device to respond to any shift in mechanical and electrical properties of the contacting medium. Chemical detection systems based on Rayleigh-wave SAW devices have largely used ST-quartz for its excellent temperature stability. To achieve specificity to a particular chemical agent, select polymer films are applied to the surface of the SAW device. Signal transduction occurs when chemicals partition into the polymer film, thereby altering the film properties, causing a change in the electrical response (that is, phase, frequency, time delay, or amplitude). For SAW resonators, acoustic loss (for example, due to fluids) at the surface contributes to a reduction of the Q, causing broadening of the response, and therefore degrades the detection limit.

SAW fluid detection systems require the use of piezoelectric substrates that minimize coupling of acoustic energy into fluids. To achieve this, a substrate must support shear-horizontal SAW (SH-SAW) excitation. A widely used substrate is the leaky-SAW substrate of 36° *YX* lithium tantalate, which can be converted to an SH-mode device by electrically shorting the surface where the acoustic wave propagates by application of a thin metal film. To increase sensitivity and electrically insulate the surface of SH-SAW devices for fluid-based biodetection systems, a thin layer or waveguide (often silicon dioxide or polymers) is applied. This creates waves that are dispersive or frequency-dependent, known as Love waves. Love-wave biosensors have been reported with detection limits of ~10–800 picograms/cm^2. *See* SEISMOLOGY.

Darren W. Branch

Bibliography. C. K. Campbell, *Surface Acoustic Wave Devices for Mobile and Wireless Communications*, Academic Press, San Diego, 1998; D. Royer and E. Dieulesaint, *Elastic Waves in Solids I: Free and Guided Propagation*, Springer-Verlag, Berlin, 2000; D. Royer and E. Dieulesaint, *Elastic Waves in Solids II: Generation, Acousto-optic Interaction, Applications*, Springer-Verlag, Berlin, 2000.

Surface and interfacial chemistry

Chemical processes that occur at the phase boundary between gas–liquid, liquid–liquid, liquid–solid, or gas–solid interfaces.

The chemistry and physics at surfaces and interfaces govern a wide variety of technologically significant processes. Chemical reactions for the production of low-molecular-weight hydrocarbons for gasoline by the cracking and reforming of the high-molecular-weight hydrocarbons in oil are catalyzed at acidic oxide materials. Surface and interfacial chemistry are also relevant to adhesion, corrosion control, tribology (friction and wear), microelectronics, and biocompatible materials. In the last case, schemes to reduce bacterial adhesion while enhancing tissue integration (**Fig. 1**) are critical to the implantation of complex prosthetic devices, such as joint replacements and artificial hearts. Selected technologies that involve surfaces and interfaces are listed in **Table 1**. *See* BIOMEDICAL CHEMICAL ENGINEERING; CRACKING; HETEROGENEOUS CATALYSIS; PROSTHESIS; SURFACE PHYSICS.

Interactions with the substrate may alter the electronic structure of an adsorbate. Those interactions that lower the activation energy of a chemical reaction result in a catalytic process. Adsorption of reactants on a surface also confines the reaction to two dimensions as opposed to the three dimensions available for a homogeneous process. The two-dimensional confinement of reactants in a bimolecular event seems to drive biochemical processes with higher reaction efficiencies at proteins and lipid membranes. *See* ADSORPTION.

A limitation in the study of surfaces and interfaces rests with the low concentrations of the participants in the chemical process. Concentrations of reactants at surfaces are on the order of 10^{-10} to 10^{-8} mole/cm^2. Such low concentrations pose a sensitivity problem from the perspective of surface analysis. Experimental techniques with high sensitivity are required to examine the low concentrations of a surface species at interfaces (**Fig. 2**).

Electron surface analysis techniques. Electron spectroscopy methods are widely used in the study of

TABLE 1. Some typical surface and interfacial processes

Interface	Processes	Significance
Liquid–liquid	Solute partitioning	Solvent extraction
Gas–solid	Adsorption/desorption Surface-mediated reactions	Catalysis
Liquid–solid	Electrochemistry Molecular recognition	Solar energy conversion Energy transfer and storage in cell membranes DNA sequencing

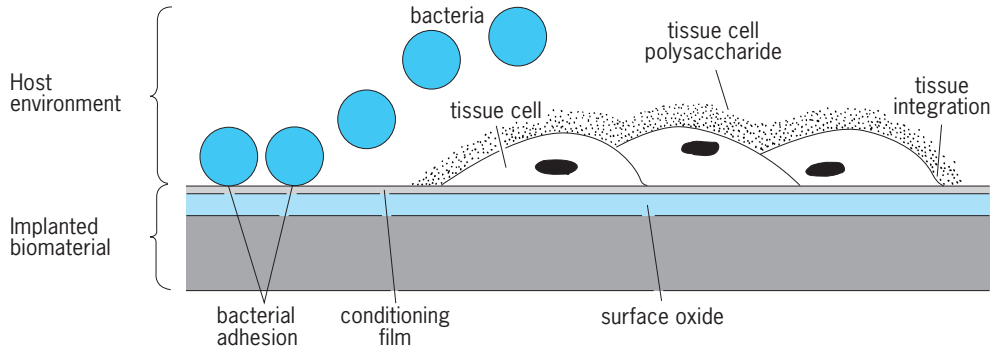

Fig. 1. Scheme for reducing bacterial adhesion at an implanted biomaterial.

surfaces because of the small penetration depth of electrons through solids. This attribute makes electron spectroscopy inherently surface-sensitive, since only a few of the outermost atomic layers are accessible. Electron spectroscopy requires propagation over distances of ~10 cm without collisions with gas-phase molecules. The mean free path of an electron, λ_e, through a gas is given by Eq. (1), where k

$$\lambda_e = \frac{4kT}{\pi \xi^2 P} \tag{1}$$

is Boltzmann's constant, T is absolute temperature, ξ is molecular diameter of the ambient gas, and P is ambient pressure. If an electron propagates 10 cm (a typical path length between the sample and detector) at room temperature, a vacuum of 7×10^{-2} pascal is required. Electron spectroscopy is typically carried out at much lower pressures of ultrahigh vacuum ($<1 \times 10^{-7}$ Pa) to ensure sample integrity.

The methods of electron spectroscopy used in surface studies have several common characteristics (**Table 2**). A source provides the incident radiation to the sample, which can be in the form of electrons, x-radiation, or ultraviolet radiation. Electron beams are generated from the thermionic emission of metal filaments or metal oxide pellets. The incident radiation induces an excitation at the surface of the sample, which alters the energy distribution of electrons that leave the surface. This distribution provides a diagnostic of the composition or structure of the interface. Typical electron energy analyzers include retarding-field, cylindrical-mirror, and hemispherical-mirror analyzers. *See* ELECTRON SPECTROSCOPY.

Auger electron spectroscopy (AES). This technique is used to determine the elemental composition of a surface species. As shown in **Fig. 3***a*, the incident electron beam (2–3 keV) ionizes a K-shell electron of an atom of energy, E_K, creating an electron vacancy. An incident x-ray can also initiate the same event. An electron in a higher-energy L-shell with energy E_L subsequently drops down to fill the vacancy, releasing the energy $E = E_K - E_L$. The excess energy can either be emitted as an x-ray or be released to an outer-shell electron. Known as an Auger electron, it is ejected with kinetic energy $E \sim (E_K - E_L - E_M)$. This is a KLM transition, based on the energy levels involved in this particular process. Other transitions involving different electrons also occur (for example, KLL and LMM). The kinetic energy of the Auger electrons provide a so-called fingerprint for elemental identification, as found in numerous reference compilations. *See* AUGER EFFECT; ELECTRON-HOLE RECOMBINATION.

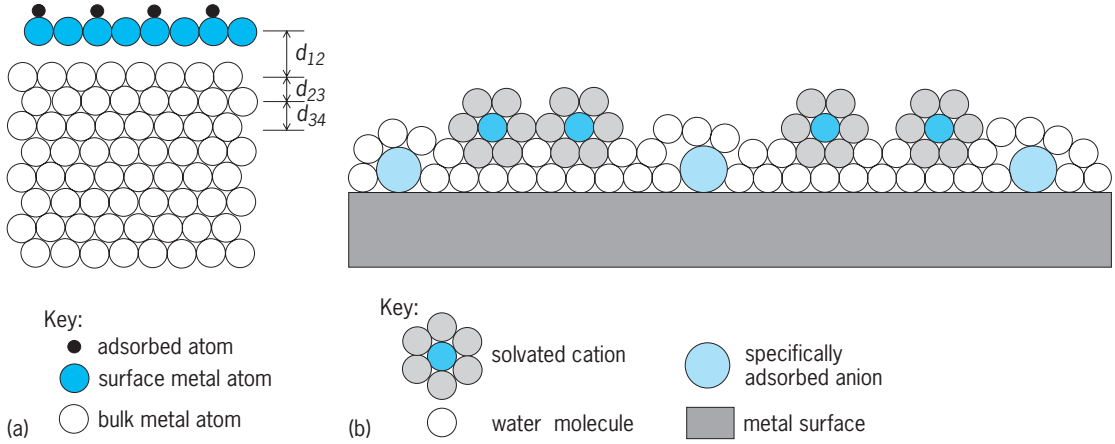

Fig. 2. Microscopic surface structures. (*a*) Surface atoms occupying positions other than the bulk-terminated locations. The first interlayer spacing d_{12} is greater than the bulk interlayer spacing d_{34}. The second interlayer spacing d_{23} is smaller than the bulk interlayer spacing. (*b*) Electrical double layer. The plane of closest approach is the inner Helmholtz plane. The adjacent layer of solvated cations is the outer Helmholtz plane.

TABLE 2. Surface-sensitive experimental techniques

Technique	Source*	Detectors	Level of Information
Auger electron spectroscopy (AES)	Electrons 2–3 keV	Cylindrical mirror or retarding field	Elemental composition
X-ray photoelectron spectroscopy (XPS)	X-rays 1254 eV (Mg) 1487 eV (Al)	Hemispherical or cylindrical mirror	Elemental composition and oxidation state
Ultraviolet photoelectron spectroscopy (UPS)	UV radiation 21 eV He(I) 41 eV He(II)	Hemispherical or cylindrical mirror	Electronic properties of adsorbate and/or bulk material
Energy loss spectroscopy (ELS)	Electrons 50–1000 eV	Electron energy analyzer	Electronic structure of surface
High-resolution electron energy loss spectroscopy (HREELS)	Electrons 1–10 eV	Electron energy analyzer	Vibrational losses
Low-energy electron diffraction (LEED)	Electrons 20–500 eV	Retarding fields and phosphorescent screen	Surface structure or periodicity
Infrared spectroscopy (IRS)	Photons	Mercury-cadmium-telluride or indium antimony	Molecular identity
Optical ellipsometry	Photons	Photomultiplier	Adsorbate layer thickness
Scanning tunneling microscopy (STM)	Tunneling current	Ammeter	Substrate roughness and texture

*Mg = magnesium; Al = aluminum; He = helium.

X-ray photoelectron spectroscopy (XPS). This technique is used to determine the oxidation state as well as the elemental composition of a surface species. In x-ray photoelectron spectroscopy, an inner-shell electron or photoelectron is ejected as a result of excitation by an incident x-ray photon (Fig. 3*b*). The x-ray source is usually a magnesium or aluminum anode. The ejected electrons have an energy $E = h\nu - E_i$, where $h\nu$ and E_i are the energies of the x-ray and the inner electron, respectively. The resulting vacancy can be filled by the Auger relaxation process. The energy of the photoelectron reflects both the oxidation state and elemental composition of the surface species. *See* X-RAY SPECTROMETRY.

Ultraviolet photoelectron spectroscopy (UPS). This is similar to x-ray photoelectron spectroscopy except that ultraviolet photons are used for excitation. The helium I and II lines are the most commonly used sources. Since ultraviolet radiation is less energetic than x-radiation, ultraviolet photoelectron spectroscopy probes only the valence electron levels of the surface species. The electron is ejected from the other outer shell with energy $E = h\nu - E_\upsilon$, where E_υ is the energy of the valence electron (Fig. 3*c*).

Electron energy loss spectroscopy (EELS). This technique probes the electronic and vibrational states of the surface species. It is divided into two categories, based on the energy E_p of the incident electron. Electronic levels are accessed with E_p = 50–1000 eV, whereas vibrational levels are probed with E_p = 1–10 eV (1 meV = 8.066 cm^{-1}). The technique that probes electronic levels is known as energy loss spectroscopy (ELS); the technique that probes vibrational levels is known as high-resolution electron energy loss spectroscopy (HREELS). Higher resolution is required to resolve the vibrational bands because the energy transitions are smaller, as shown in Fig. 3*d*. Most of the electrons are scattered elastically, and

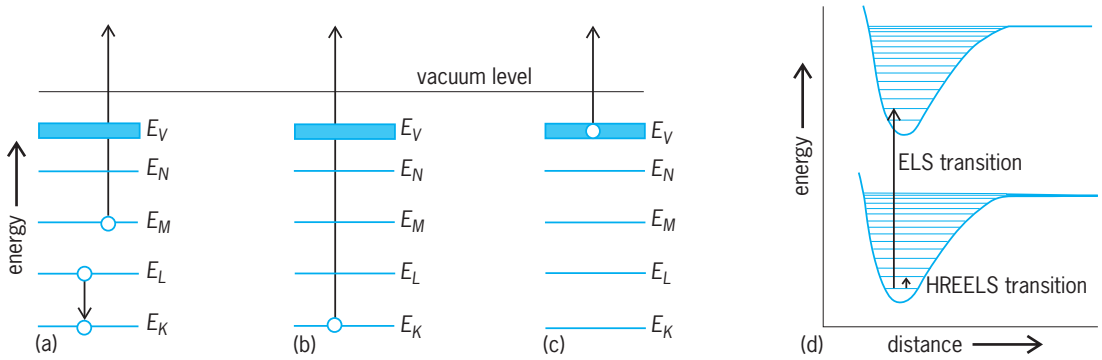

Fig. 3. Processes of electron spectroscopies. (*a*) Auger electron spectroscopy. A core-level electron is ejected by excitation from the incident high-energy electron. A second electron fills the vacancy left by the ejected electron. The third electron, the Auger electron, is ejected from the atom with the energy $E = E_K - E_L - E_M$. (*b*) X-ray photoelectron spectroscopy. A core-level electron is ejected from the atom as a result of excitation by the incident x-ray beam. Note that the Auger process, described in *a*, can also occur as a result of x-ray stimulation of the sample. (*c*) Ultraviolet photoelectron spectroscopy. Ultraviolet radiation excites the sample, causing ejection of weakly bound valence electrons. (*d*) Electron energy loss spectroscopy. An electron is excited into a higher electronic (ELS) or vibrational (HREELS) state by the incident electron beam. The scattered electrons lose the energy that was required for excitation. E_v = energy of the valence electrons.

they fail to induce an electronic or vibrational excitation of the surface species. However, in a few cases the incident electron undergoes an inelastic collision, losing energy via excitation of a vibration or electronic transition. The energy of the reflected beam is $E = E_p - E_{loss}$ (Fig. 3*d*). For high-resolution electron energy loss spectroscopy, only those vibrations that have a component of their transition dipole moment perpendicular to the surface can be excited. With this surface selection rule, the loss spectrum often provides information pertaining to the orientation of the surface species with respect to the surface normal.

Low-energy electron diffraction (LEED). This technique is the surface analog of x-ray diffraction; it probes the two-dimensional periodicity of the surface species. Applicable primarily to studies of single-crystal surfaces, low-energy electron diffraction arises from the elastic scattering of electrons incident on the surface. The electrons scatter in phase or out of phase, depending on the incident energy (wavelength) and the position of the atoms relative to one another. In-phase scattering, which interferes constructively to yield the low-energy electron diffraction pattern, occurs when the Bragg condition is satisfied. The Bragg condition is given by Eq. (2), where a is the distance

$$a(\sin\phi - \sin\phi_0) = n\lambda \qquad (2)$$

between surface atoms, ϕ and ϕ_0 are the backscattered and incident angles of the electron beam, λ is the wavelength of the electron beam, and n is an integer corresponding to the diffraction order. The diffracted electrons are accelerated and focused by a series of grids onto a phosphorescent screen for visualization of the low-energy electron diffraction pattern. *See* ELECTRON DIFFRACTION.

Optical surface analysis techniques. Optical spectroscopy techniques (visible and infrared) are also useful for probing the chemical composition and molecular arrangement of surface species. Typical application configurations are the transmission and reflection (both external and internal) modes as shown in **Fig. 4**. Transmission spectroscopy relies on the passage of the probe beam through the sample. External and internal reflection spectroscopies involve the reflection of the probe beam from a medium with a lower refractive index to a medium with a higher refractive index, and from a higher to lower refractive index, respectively. The sample support must be optically transparent to the probe beam for the internal reflection mode. In both cases, the substrates are polished to a smooth, mirrorlike finish. *See* SPECTROSCOPY.

Diffuse reflection spectroscopy, which is essentially reflection from a rough surface, has been utilized to examine supported-metal catalysts and high-surface-area oxides. In all cases, the power lost by the interaction of the incident radiation with the surface species gives rise to an absorption spectrum. Comparison with the spectrum of the adsorbate starting material provides insights into the nature of the surface species. Propagation distances are much longer than those for electron spectroscopy, allowing the application of these techniques to in-situ surface studies such as at the gas–solid interface of a high-pressure chemical reactor as well as the liquid–solid interface of an electrochemical cell. Raman spectroscopy also provides information about the molecular composition of the surface species, although at a lower sensitivity. *See* RAMAN EFFECT.

Optical probes are less surface-sensitive than electron spectroscopy because of the longer wavelength of the incident energy. This can result in the sampling of both bulk and surface species, complicating the interpretation of the observed spectrum. Penetration depths for optical spectroscopies are typically a fraction of the wavelength of the incident light, and they are a function of the optical properties of the sample and of the support and the angle of incidence. For visible radiation, penetration depths in a reflection experiment are tens to hundreds of nanometers, whereas those for infrared radiation are a few hundred to thousands of nanometers. For internal reflection spectroscopy, the penetration depth d_p is given by Eq. (3), where λ_1 is the wavelength of the incident radiation, θ is the angle of incidence, and n_{21} is the ratio of the refractive indices of media 2 and 1, respectively.

$$d_p = \frac{\lambda_1}{2\pi(\sin^2\theta - n_{21})^{1/2}} \qquad (3)$$

Infrared spectroscopy (IRS). This is a useful technique for the determination of the molecular composition and orientation of the surface species. In an external reflection mode, a beam of infrared light is reflected at a smooth metal, semiconductor, or insulator surface, as shown in Fig. 4*b*. For films with thicknesses

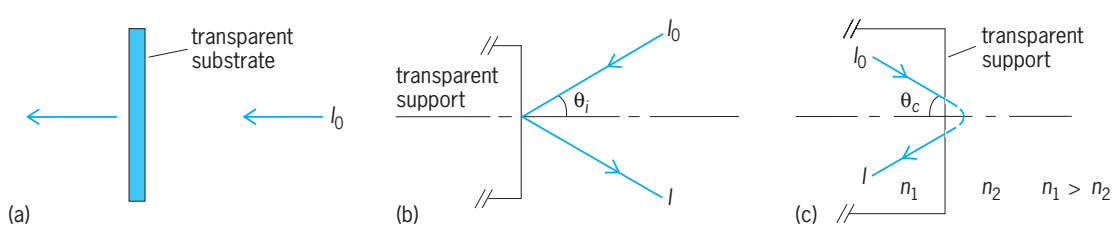

Fig. 4. Typical configuration for application of optical spectroscopies to surface analysis. (*a*) Transmission. (*b*) External or specular reflection. (*c*) Internal reflection. I = reflected light intensity. I_0 = initial light intensity. θ_i = angle of incidence at phase 1–phase 2 boundary. θ_c = angle of refraction at phase 1–phase 2 boundary. n_1 = refractive index of phase 1. n_2 = refractive index of phase 2.

greater than ~10 nm, conventional dispersive spectrometers can be used. However, studies of monolayers require a Fourier transform spectrometer and high-sensitivity, low-noise, photoconductive detectors, for example, mercury-cadmium-telluride.

The band shapes and intensities of reflection spectra can be calculated quantitatively with classical electromagnetic theory. Variables in the calculation, which is a reflection spectroscopy analog to the Beer-Lambert law in transmission spectroscopy, include the optical function of each phase and the polarization and angle of incidence of the incident beam. Comparison of the intensities of the observed and calculated spectra, assuming that the force constants of a given vibrational mode are unchanged by immobilization, provides an estimate of the average spatial orientation of the adsorbate with respect to the substrate. As with high-resolution electron energy loss spectroscopy, such interpretations at highly reflective surfaces such as metals are aided by the preferential excitation of those vibrational modes with a transition dipole moment perpendicular to the surface, according to the infrared surface selection rule. An orientational analysis at a surface with a low reflectivity can also be accomplished, but it is hindered by a complex variation of the relative degree of excitation of vibrations with modes both perpendicular and parallel to the surface. *See* INFRARED SPECTROSCOPY.

Optical ellipsometry. This is a variation of the external reflection measurement in that the state of polarization of the reflected light is measured as opposed to the reflected power. The change in polarization is caused by the difference in the intensity and phase change for the components of the incident electromagnetic wave that are parallel and perpendicular to the surface. Such changes, which are a function of the optical properties of the adsorbate, can be used to determine its thickness. More detailed treatments with effective medium theory can also provide insights into the density of the surface layer. Computerized data acquisition is advantageous, because numerical analysis requires the collection of large data sets and extensive complex algebraic manipulation. *See* ELLIPSOMETRY.

Scanning tunneling microscopy (STM). This is a surface analysis technique that provides information about the three-dimensional topography of a surface with resolution at the atomic level. Operation in ultrahigh vacuum, at liquid–solid interfaces, and in laboratory ambient (air) has been demonstrated.

The two general operation configurations are the constant-current and constant-height modes. In the constant-current mode (**Fig. 5***a*), a bias voltage (~10–10,000 mV) is applied between a sharp metallic tip and a conducting sample. A small tunneling current (~0.2–10 nanoamperes) flows when the tip is positioned within a few atomic diameters of the sample. The tip is then scanned laterally with a piezoelectric translator over a surface while a computer-controlled feedback mechanism adjusts the height of the tip to maintain a constant current. An image of the surface contour is constructed by plotting a

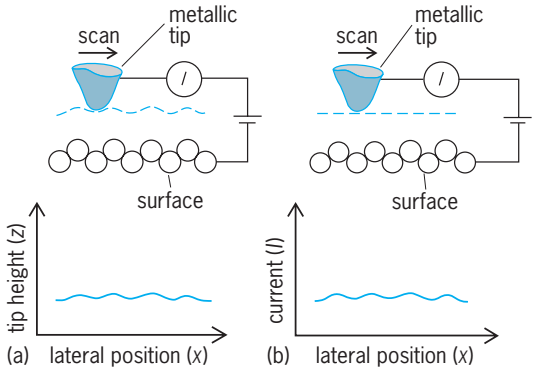

Fig. 5. Scanning tunneling microscopy and images for (*a*) constant-current mode and (*b*) constant-height mode; shown only in *x* dimension for simplicity.

weighted average of the tip height z versus the lateral positions x and y.

In the constant-height mode, a tip is scanned across a surface at constant height while monitoring the current I (Fig. 5*b*). A surface image is derived from a plot of multiple scans displaced from each other.

Each mode of operation has its own specific advantages. The constant-current mode provides topographic pictures of surfaces that are not atomically flat. This mode operates at scan rates of only a few hertz and is limited by the vertical movement of the tip by the z translator. In contrast, the constant-height mode can image atomically smooth surfaces at rates up to ~1 kHz. These high imaging rates are attainable because only the electronics for monitoring the tunneling current—and not the vertical movement of the z translator—must respond to the atoms passing under the tip. A high imaging rate allows the examination of real-time processes, reduces data acquisition time, and minimizes image distortion caused by thermal drift and piezoelectric creep. *See* INTERFACE OF PHASES; SCANNING TUNNELING MICROSCOPE.

Mary M. Walczak; Marc D. Porter

Bibliography. C. Bai, *Scanning Tunneling Microscopy and Its Application*, 2d ed., 2000; H. Bubert and H. Jenett (eds.), *Surface and Thin Film Analysis: A Compendium of Principles, Instrumentation, and Applications*, 2002; H. P. Hughes and H. Starnberg (eds.), *Electron Spectroscopies Applied to Low-Dimensional Materials*, 2000; A. J. Milling, *Surface Characterization Methods: Principles, Techniques, and Applications*, 1999; G. A. Somorjai, *Introduction to Surface Chemistry and Catalysis*, 1994.

Surface condenser

A heat-transfer device used to condense a vapor, usually steam, by absorbing its latent heat in a cooling fluid, ordinarily water. Most surface condensers consist of a chamber containing a large number of 0.5–1-in.-diameter (1.25–2.5-cm) corrosion-resisting alloy tubes through which cooling water flows. The vapor contacts the outside surface of the tubes and is

condensed on them. The tubes are arranged so that the cooling water passes through the vapor space one or more times. About 90% of the surface is used for condensing vapor and the remaining 10% for cooling noncondensable gases. Air coolers are normally an integral part of the condenser but may be separate and external to it. The condensate is removed by a condensate pump and the noncondensables by a vacuum pump. *See* HEAT TRANSFER; STEAM CONDENSER; VAPOR CONDENSER. Joseph F. Sebald

Surface hardening of steel

The selective hardening of the surface layer of a steel product by one of several processes which involve changes in microstructure with or without changes in composition. Surface hardening imparts a combination of properties to the finished product not produced by bulk heat treatment alone. Among these properties are high wear resistance and good toughness or impact properties, increased resistance to failure by fatigue resulting from cyclic loading, and resistance to surface indentation by localized loads. The use of surface hardening frequently is also favored by lower costs and greater flexibility in manufacturing.

Hardening processes. The principal surface hardening processes are (1) carburizing; (2) the modified carburizing processes of carbonitriding, cyaniding, and liquid carburizing; (3) nitriding; (4) flame hardening, induction hardening, and laser hardening; and (5) surface working.

Carburizing introduces carbon into the surface layer of low-carbon steel parts and converts that layer into high-carbon steel, which can be quench-hardened by appropriate heat treatment. Carbonitriding, cyaniding, and liquid carburizing, in addition to supplying carbon, introduce nitrogen into the surface layer; this element permits lower case-hardening temperatures and has a beneficial effect on the subsequent heat treatment. In nitriding, only nitrogen is supplied, and reacts with special alloy elements present in the steel.

Whereas the foregoing processes change the composition of the surface layer, flame hardening, induction hardening, and laser hardening depend on a heat treatment applied selectively to the surface layer of a medium-carbon steel. These processes are based on the same principle, but differ in the source of heat; also, in flame hardening and induction hardening, the heated layer is rapidly cooled (quenched) by an external cooling medium, whereas the layer heated by a laser beam is quenched (rapidly cooled) by conduction to the underlying cold metal. Surface working by shot peening, surface rolling, or prestressing improves fatigue resistance by producing a stronger case, compressive stresses, and a smoother surface.

Table 1 lists the surface-hardening processes for steel and their major characteristics. Processes related to the surface hardening of steel are hard-facing, metal spraying, electroplating, and various diffusion processes involving elements such as aluminum, silicon, and chromium. *See* CLADDING; ELECTROPLATING OF METALS; METAL COATINGS.

Carburizing. The oldest method of surface hardening steel, carburizing, introduces carbon into the surface layer of a low-carbon steel by heating above the transformation range in contact with a carbonaceous material. The carbon diffuses into the steel from the surface and thus converts the outer layer into high-carbon steel. The composite is then heat-treated by the procedures generally applicable to steels. In particular, it must be cooled from above the transformation temperature at a rate sufficiently fast to transform the high-carbon surface layer into a hard martensitic case while the low-carbon core remains tough and shock-resistant. The quench is usually followed by a low-temperature stress-relief anneal. *See* HEAT TREATMENT (METALLURGY).

Pack carburizing. In pack carburizing, carbon is supplied by charcoal or coke to which carbonates or organic materials are added; the mixture is known as the carburizing compound. Parts to be hardened are packed in a steel box with the carburizing compound and heated to the carburizing temperature, usually 1700–1750°F (925–955°C). Carbon is transferred to the steel by the formation of carbon monoxide at the compound surface and by its decomposition to carbon and carbon dioxide at the steel surface according to the reaction given below. The carbonates

$$C + CO_2 \rightleftharpoons 2CO$$

or organic materials in the carburizing compound decompose and increase the concentration of carbon oxide gases in the box required for the transfer of carbon to the steel. The depth of penetration of the carbon depends upon the time and temperature at which the treatment is carried out.

TABLE 1. Surface-hardening processes for steel

Process	Elements added	Hardening mechanism
Carburizing	Carbon	Formation of martensite
Carbonitriding	Carbon and nitrogen	Formation of martensite
Cyaniding	Carbon and nitrogen	Formation of martensite
Liquid carburizing	Carbon and nitrogen	Formation of martensite
Nitriding	Nitrogen	Precipitation of alloy nitrides
Flame hardening	None	Formation of martensite
Induction hardening	None	Formation of martensite
Laser hardening	None	Formation of martensite
Surface working	None	Work hardening

The principal advantage of pack carburizing is its simplicity; no expensive equipment is required. The results are almost certain to be satisfactory with proper temperature control and hardening practice. The hardening treatment usually requires that the parts, after removal from the box, be heated again above the transformation temperature, and that this process be followed by quenching.

Gas carburizing. In gas carburizing the parts are heated in contact with carbon-bearing gases, commonly carbon monoxide and hydrocarbons. The hydrocarbons may be methane, propane, butane, or vaporized hydrocarbon fluids. They are usually diluted with an inert carrier gas to control the amount of carbon supplied to the steel surface and to prevent the formation of soot.

The carbon monoxide and hydrocarbons are decomposed at the steel surfaces, the carbon thus liberated being absorbed by the steel. Close control of gas composition is required because the rate at which carbon is supplied to the steel surface controls the concentration of carbon in the carburized case. This control is an important advantage of gas carburizing. The process is cleaner and entails lower labor costs than pack carburizing. Direct quenching from the furnace is possible in gas carburizing; hence the process is particularly well suited to volume output for the mass production industries. In many installations large continuous furnaces with attached quenching and tempering equipment are used; the parts are charged into one end and leave at the other in the carburized and hardened state. Batch furnaces are also used for gas carburizing.

Steels for carburizing. The selection of steels for carburizing primarily concerns grain growth characteristics, carbon and alloy content, machinability, and cost. Steels that retain a fine-grained structure at the case-hardening temperature such as aluminum deoxidized steels are desirable because they permit simple heat-treatment procedures, in particular, hardening by a direct quench. A fine-grained structure in the finished product is essential for maximum shock resistance.

Plain carbon steels are satisfactory for many applications in which low distortion is not a critical requirement and for which optimum core properties are not required. The most common carbon contents are 0.20% in plain carbon steels and 0.08–0.20% in alloy steels. Steels with relatively high sulfur content are frequently used for improved machinability.

The most common alloy elements in carburizing steels are nickel (0.5–3.5%), chromium (0.5–1.6%), and molybdenum (0.1–0.25%). The nickel-molybdenum steels are particularly popular for strength and toughness of the core and toughness of the case. Alloy steels have less tendency to develop coarse-grained structures at the carburizing temperature. They also permit slower quenching rates in hardening, thus reducing distortion and the tendency to crack during quenching. Their higher cost is frequently more than offset by lower finishing costs due to reduced distortion. Improved fatigue resistance is another consideration in the selection of alloy steels.

Typical parts surface-hardened by carburization, as described above, or by modified carburizing processes, described below, include gears, ball and roller bearings, piston pins, sprockets and shafts, clutch plates, and cams.

Modified carburizing processes. Alloying the steel with nitrogen lowers the transformation temperature and reduces the transformation rate. Thus modification of the carburizing process by the diffusing of both carbon and nitrogen into the surface layer of the steel enables the process to be carried out at lower temperatures than with carbon alone. The sources of carbon and nitrogen distinguish the several processes from each other.

Carbonitriding. The carbonitriding process is the same as gas carburizing except that ammonia is added to the furnace atmosphere to provide the source of nitrogen. The amount of nitrogen absorbed by the steel can be controlled by the concentration of ammonia in the furnace atmosphere and the temperature. Thus the nitrogen content of carbonitrided cases may vary from small concentrations to the relatively high level characteristic of cyaniding.

Both because carbonitriding is conducted at lower temperatures than carburizing and because the slower transformation rate permits slower quenching rates, less distortion results from the hardening process. Consequently, plain carbon steel can frequently be substituted for alloy steels in carbonitrided parts.

Figure 1 shows a typical carbonitrided case. The microstructure is similar to a carburized case except that it contains a larger fraction of retained austenite near the surface.

Cyaniding and liquid carburizing. If the parts are immersed into molten baths consisting of solutions of cyanides and other salts, the cyanides supply both carbon and nitrogen to the steel. Thus the results from cyaniding and liquid carburizing are similar to

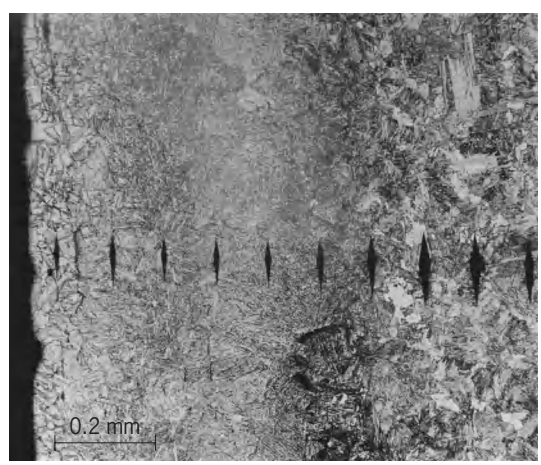

Fig. 1. Microstructure of typical carbonitrided steel. Dark border is specimen mount. Martensite needles (gray) and retained austenite (whitish) can be seen near surface. Microhardness indentations (black) are smaller in the case than in the core, indicating the greater hardness of the case. (*Courtesy of A. J. Gregor*)

those from carbonitriding. By controlling the composition and temperature of the bath, the amounts of carbon and nitrogen absorbed by the steel can be controlled within limits. The term cyaniding is usually applied to processing at temperatures of 1400–1550°F (760–845°C). Liquid carburizing is carried out at 1600–1750°F (870–955°C); at these temperatures the nitrogen absorption is lower so that the process, as far as the product is concerned, approaches pack and gas carburizing.

Nitrogen in cyanided cases results in the same advantages as those obtained from nitrogen in carbonitrided cases, but the control of nitrogen concentration is less precise. Cyaniding and liquid carburizing have the advantage of rapid heating of the charge because of the good heat transfer from the bath to the steel. These processes have considerable flexibility in that selective hardening may be accomplished by partial immersion and that different parts may be treated in the same bath for different times.

One disadvantage of these processes is that the case for a given treatment time is less deep when low operating temperatures are used. The size of salt baths is limited by the necessity to obtain uniform temperatures throughout. The baths are, therefore, less well adapted to quantity production than furnaces employing gaseous atmospheres. The nitrogen dissolved in the steel may cause an appreciable fraction of the steel to remain untransformed as retained austenite in the final product, particularly in the cyaniding process, but also in carbonitriding, especially at low temperatures.

Nitriding. This process is carried out by heating steels of suitable composition in contact with a source of active nitrogen at temperatures of 925–1100°F (495–595°C) for periods of 1–100 h, depending upon the steel being treated and the depth of case desired. Under these conditions, nitrides form if the steel contains alloying elements such as aluminum, chromium, molybdenum, vanadium, and tungsten. The formation of alloy nitrides at the nitriding temperature accounts for the hardened case. The microstructure of a typical nitrided steel is shown in **Fig. 2**.

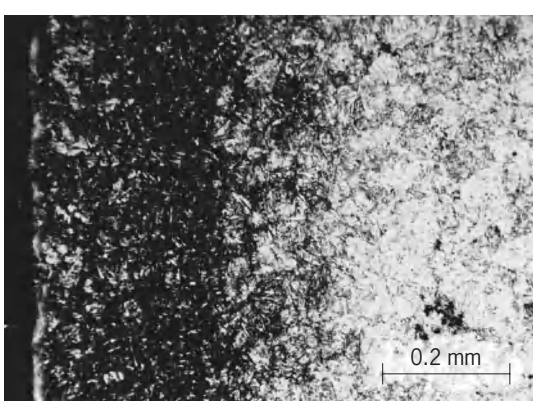

Fig. 2. Microstructure of a nitrided steel, which was heated to 1700°F (925°C) for 1 h, oil-quenched, tempered at 1250°F (675°C) for 2 h, air-cooled, and nitrided for 48 h at 975°F (525°C). Dark border is specimen mount.

The usual source of active nitrogen for nitriding is ammonia. However, mixtures of molten cyanide salts are also used, with similar advantages and disadvantages as in cyaniding and liquid carburizing. Salt mixtures with cyanates (which are less toxic than cyanides) are coming into use. A typical nitriding installation using ammonia consists of a reservoir of ammonia, a furnace, a retort containing the parts to be case hardened, and equipment to control the temperature and gas flow. At the nitriding temperature, part of the ammonia decomposes at the surface of the steel, liberating active nitrogen, some of which diffuses into the steel. The remainder passes into the molecular form, which is inert. For successful nitriding it is therefore necessary to control the gas flow in such a way as to continuously supply fresh ammonia to all steel surfaces. However, an oversupply of ammonia results in the formation of an excessively thick iron nitride layer on the surface. This so-called white layer can be controlled by regulation of the degree of ammonia dissociation, or it must be removed by grinding if it exceeds a depth of 0.0005 in. (12.7 micrometers).

In ion nitriding, the workpieces are cathodically charged to a potential of 400 to 1000 V in an evacuated furnace chamber. When a nitrogen-containing gas at low pressure is introduced into the chamber, it becomes ionized (a plasma). The positive nitrogen ions bombard the surface of the workpiece under the influence of the potential drop and cause the nitriding reaction which introduces nitrogen into the surface layer of the steel. The bombardment by the plasma discharge also generates the heat necessary to raise the temperature to the level required for the reaction to proceed. The basic features of the nitriding reaction are essentially the same as in conventional gas or liquid nitriding: nitrogen is absorbed at the surface, it diffuses into the workpiece and, given sufficient nitriding potential and time, a compound layer is formed at the surface. Ion nitriding is best suited for the production of shallow cases. It requires special equipment and appropriate expertise. Among the advantages claimed for ion nitriding, the use of which is still growing, are its flexibility, ability to be controlled, and low energy and gas consumption.

Steels that produce the hardest nitrided cases consistent with optimum case depth contain about 1% aluminum, 1.5% chromium, and 0.3% molybdenum. Alloy steels containing only chromium and molybdenum as nitride-forming elements are also popular. Stainless steels and high-speed cutting steels are nitrided for improved wear resistance. Among the parts hardened by nitriding are camshafts, fuel injection pump parts, gears, cylinder barrels, boring bars, spindles, splicers, sprockets, valve stems, and milling cutters.

Induction, flame, and laser hardening. The principle of both flame and induction hardening is to heat quickly the surface of the steel to above the transformation range followed by quenching rapidly. In induction hardening, heat is generated within the part by electromagnetic induction. The part (such

as a crankshaft, camshaft, axle, gear, or piston rod) is usually placed inside a copper coil or coils through which a rapidly alternating current is flowing. High-frequency currents are used because they confine the induced currents to the surface of the part to be heated; the higher the frequency, the shallower the case. Short heating cycles minimize conduction of heat to the interior and thus further restrict the heating to the surface layer. *See* INDUCTION HEATING.

In flame hardening, the steel to be surface-hardened is heated by direct impingement of a high-temperature gas flame. The surface layer is quickly heated to a temperature above the transformation range, followed by a quench. In both induction and flame hardening the quenching action is a combination of heat extraction by the cold metal beneath the case and by an external quenching medium.

Induction and flame hardening require close control of time and temperature of heating. In general, induction heating is used when large numbers of symmetrically shaped parts are to be processed. Flame hardening can be readily applied to large parts, such as large gears or lathe ways, and to parts of intricate design, such as camshafts. It is more economical than induction hardening when only a few parts are to be treated.

Steels for induction or flame hardening usually contain 0.4–0.75% carbon. Because no change in composition is involved, the steel is selected for both case and core properties. During hardening the core is not affected and, consequently, the core properties must be developed by proper heat treatment before surface hardening. Cast irons are also induction- or flame-hardened for certain applications. *See* CAST IRON.

In laser hardening, as in induction and flame hardening, a surface layer of the workpiece is heated above the transformation range without a change in composition, but laser hardening does not require an external quenching medium. The heat source is a laser which generates an intense energy flux and rapidly raises the temperature. The laser beam is manipulated by lenses and mirrors, and travels at a controlled speed across the surface area to be hardened. In order to reduce the energy loss by reflection, absorbing coating materials are applied or the surface is roughened. Since the heated surface layer is shallow, it loses heat rapidly to the interior (self-quenching) and hardens by the formation of martensite. The case depth depends mainly on the power density and traveling speed of the beam and usually is less than for induction and flame hardening. Advantages of laser hardening are its ability to harden selectively small areas as well as complex shapes and small bores. Drawbacks are the limited case depth and the capital costs of the equipment. *See* LASER.

Surface-working processes. The selective cold working of the surface layer of parts of steel and other metals increases surface hardness. The working requires a force exceeding the compressive yield strength of the material. Such a force can be applied by various methods, for example, hammer peening, mechanical peening, shot peening, and surface rolling. The surfaces must be accessible to the peening or rolling operation, but the processes can be applied to selected critical areas, such as the fillets of shafts. Prestressing involves stressing a part beyond its yield strength.

Surface-working processes result in substantial improvement of the properties, especially the fatigue resistance. The processes are applied to coil and leaf springs, shafts, gears, and steering knuckles. They have the advantage of being comparatively inexpensive. They must be carefully controlled to avoid overworking or underworking of the surface, both of which fail to give the desired improvement in the fatigue properties.

Properties. The characteristic properties of surface-hardened steels depend on the properties of both the case and the core. Case properties are determined mainly by composition, microstructure, and case depth. Core properties of carburized steels depend primarily on the transformation characteristics of the core during the hardening of the case. In nitriding and also in induction and flame hardening, the core properties are developed before the case hardening treatment.

Microstructure. The microstructures of the case and core are controlled by their composition and heat treatment. The structure of carburized, carbonitrided, and cyanided cases is typical of heat-treated high-carbon steel, and that of induction- and

TABLE 2. Typical case properties

Process	Typical case depth, in.*		Typical hardness
Carburizing	<0.020	Shallow	55–65 RC
	0.020–0.040	Medium	
	0.040–0.060	Heavy	
	>0.060	Extra deep	
Carbonitriding	0.003–0.020		55–62 RC
Cyaniding	0.001–0.010		
Nitriding	0.005–0.025		85–95 R-15N
	0.001–0.003	High-speed steel	
Induction hardening	0.010–0.25		50–60 RC
Flame hardening	0.030–0.25		50–60 RC
Laser hardening	0.008–0.050		
Surface working	0.020–0.040		

*1 in. = 25 mm.

flame-hardened cases is typical of medium-carbon steel. The desired microconstituent in each instance is martensite, which usually is tempered. The retained austenite, which is promoted by nitrogen in carbonitrided and cyanided cases, and by some other alloy elements, is generally undesirable. Austenite lowers the hardness and its subsequent transformation in service may cause brittleness and dimensional changes of the case-hardened part. *See* TEMPERING.

The microstructure of nitrided cases consists of finely dispersed alloy nitrides precipitated in a preexisting hardened and tempered structure. In induction- and flame-hardened cases, the fast rate of cooling from the austenitizing temperature results in the formation of martensite. The characteristic feature of surface-worked cases is their cold-worked structure and resulting increase in hardness. *See* PLASTIC DEFORMATION OF METAL.

Case depth. Typical case depths of surface-hardened steels are shown in **Table 2**. There has been a trend toward thinner cases, which are acceptable because of stronger cores and less distortion during processing. However, for some types of service, especially those involving contact of two loaded parts, the critical stress occurs below the surface; the case depth should preferably be sufficient to allow for this condition. Thick cases are essential in some wear applications.

Residual stresses. Compressive stresses at the surface are desirable because they provide protection against fatigue failure. The expansion accompanying the formation of martensite in carburizing and, similarly, that resulting from the precipitation reactions in nitriding cause these residual compressive stresses. Shot peening or other surface-working processes also cause residual compressive stresses at the surface.

Dimensional changes. In heat-treating carburized steels, distortion occurs primarily as a result of uneven quenching, which causes different parts of the surface to transform at different times. In nitriding, dimensional changes result primarily from the increase in volume of the case as the alloy nitrides are precipitated. Michael B. Bever; Carl F. Floe

Bibliography. American Society for Metals, *Metals Handbook*, desk edition, 1985; C. R. Brooks, *Heat Treatment of Ferrous Alloys*, 1979.

Surface mining

A mining method used to obtain valuable minerals from the earth by first removing the overlying soil and rock (overburden) and subsequently recovering the valuable mineral. Surface mining is done until the accumulated overburden removed becomes too thick (and expensive to remove) for the economic recovery of the mineral. For mineral deposits too deep to be surface-mined, underground mining methods are used. *See* MINING.

History. Large-scale surface mining in the United States started in the early 1900s, as the use of large mechanical excavators made the recovery of lower-grade ores economical. The Bingham Canyon copper mine in Utah, which opened in the late 1800s, was converted to an open-pit mine around 1905. Full-revolving steam shovels, purchased from the Panama Canal project, greatly improved the mining efficiency at this mine and allowed the open-pit concept to be implemented. Technology advancements in drilling and blasting methods also contributed to successful open-pit mining.

Surface coal mining became a significant contributor to coal production around 1910. A surface coal mine on Grapevine Creek near Danville, Illinois, may have been the first large-scale strip mine in the United States. Track-mounted steam shovels from the Panama Canal project were also used at this mine. It was not until the 1930s that steam shovels were replaced by electric shovels. Around the same time, trucking became the dominant transportation method, replacing steam-powered track haulage.

Mining methods. Surface mining methods include placer mining, open glory-hole mining or milling, open-pit mining, strip mining, and quarrying.

Placer mining. These methods are used to recover unconsolidated mineral-rich (placer) deposits, often in streams, rivers, or their associated flood plain. Hydraulic methods, such as high-pressure water, can be used to loosen dry, unconsolidated material. The water–ore mixture then flows by gravity to an associated facility for processing into a salable product. Due to environmental concerns, this mining method is seldom used today. Another method of mining placer deposits is to use a dredge for excavating unconsolidated material underwater. Sand and gravel mining often use dredges. *See* PLACER MINING.

Open glory holes or milling pits. These are combined surface and underground mining operations, where a large opening or shaft is excavated vertically from the surface through a mineral-rich deposit to an underground mine. Drilling and blasting of the ore is directed downward from the surface around the shaft. The broken ore falls through the shaft and is recovered in the underground mine. These methods are typically used for precious-metal ore deposits such as gold, silver, or copper.

Open-pit mining. This method is used to recover a large ore deposit by removing the overlying earth and rock in conjunction with removing the ore. An open-pit mine is designed in a circular or oval shape, creating a deep pit. Steep, unstable slopes are avoided since they could trigger a landslide and endanger the miners. The pit is kept dry by draining water to a sump and pumping it to a treatment facility outside the pit. The overburden and processed waste rock are placed in a landfill area. The mining operation can continue until the cost of removing the overburden exceeds the payment received for the recovered mineral. Open-pit mining typically is used at very large ore bodies, with the economic life sometimes exceeding 50 years. Open-pit mining is also called open-cut mining since the overburden is not returned to the pit and an open cut remains after mining is completed. *See* OPEN-PIT MINING.

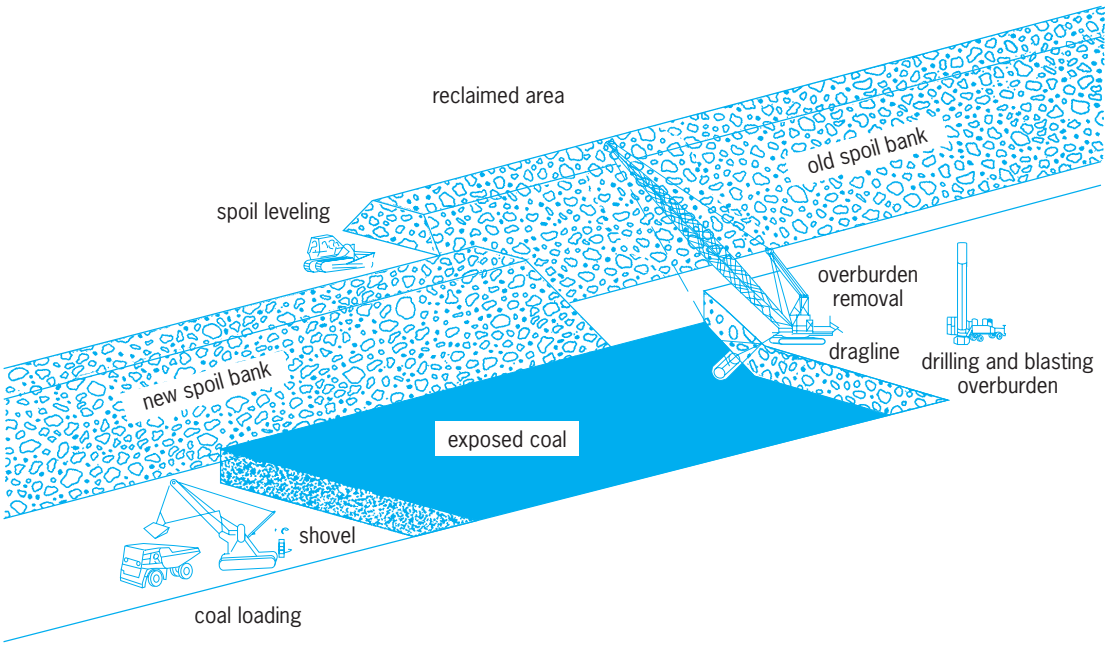

Fig. 1. Area mining.

Strip mining. Strip mining is the surface mining of coal. Before mining begins, ditches and ponds are constructed for containing sediment from erosion caused by rainstorms. The area is then cleared of trees, and the topsoil is removed and stockpiled. The three types of strip mining are area mining, contour mining, and mountaintop removal. *See* COAL; COAL MINING.

Area mining occurs on large sites that are flat or with rolling hills (**Fig. 1**). An initial opening is made by removing the overburden from a small area to uncover the coal. The opening is called a box cut since it is rectangular. It can also be called the first cut. Overburden from the first cut is stockpiled. After the coal is removed, overburden from a second cut is placed over the area of the first cut and graded to approximately the same contour as before mining. The area is covered with the stockpiled topsoil and reclaimed by planting grasses and trees. This operation continues until the limit of the mine is reached. The overburden from the first cut is then hauled to the last or final cut and used to restore that area. In this manner, the mined zone in a surface coal mine is returned to a condition similar to the one that existed before mining.

When a coal seam is located along a steep hillside and only one cut into the hill can be economically made, the contour mining method is used (**Fig. 2**). As the overburden cut advances, the excavated material is hauled past the exposed coal to the previously mined area and used for reclamation.

The mountaintop removal mining method is used to recover multiple coal seams that lie within the geologic strata on a mountain ridge (**Fig. 3**). The ridge's top is excavated, exposing the coal seams for recovery. Overburden is sequentially replaced to restore the ridge to approximately the same configuration as before mining. The overburden has a larger volume than the previously undisturbed strata because broken rock occupies more space than consolidated

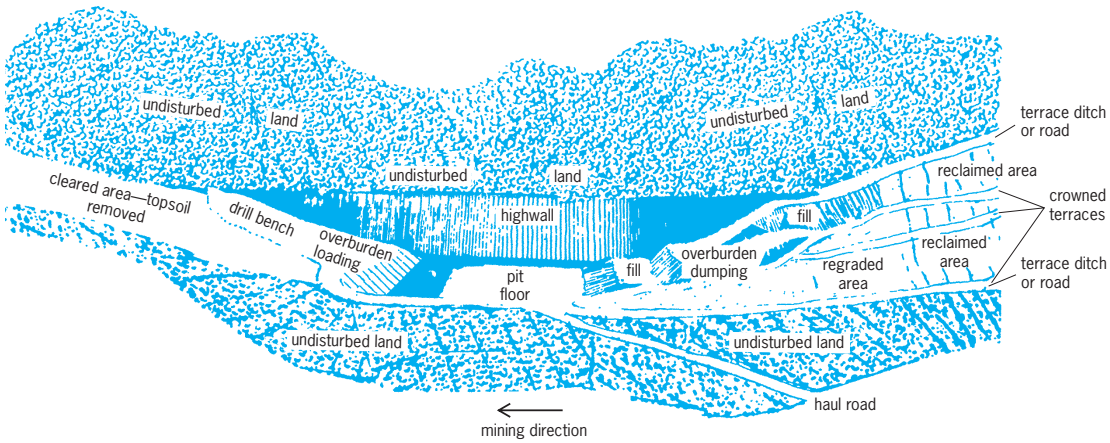

Fig. 2. Contour mining.

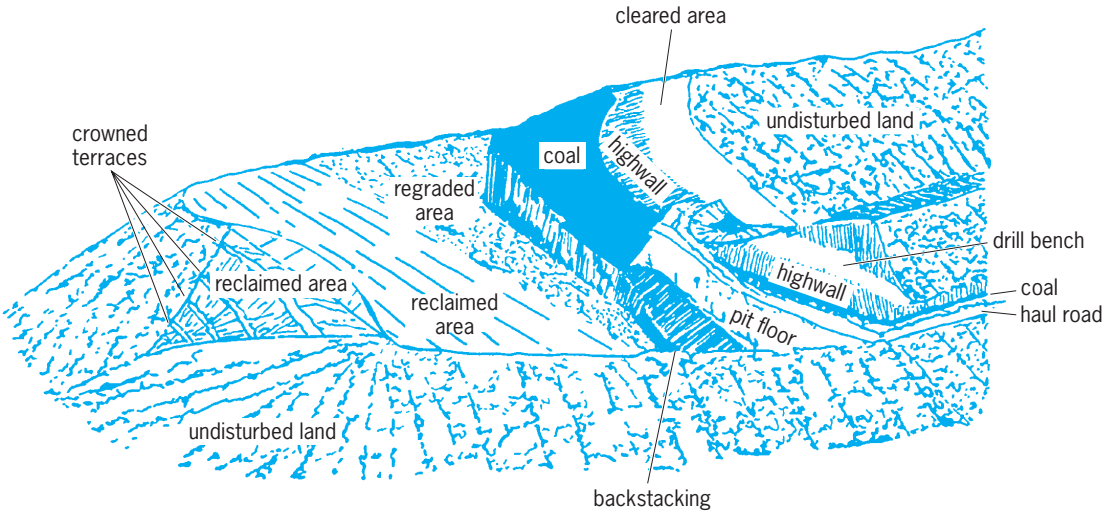

Fig. 3. Mountaintop mining.

rock, so excess overburden is placed in a landfill near the mine site.

Engineers plan the limits of strip mines by calculating economic strip ratios. The economic strip ratio is the thickness of the overburden divided by the thickness of the coal. Mining is economically feasible until the cost of excavation and reclamation becomes higher than the coal's worth. For example, if the price of coal restricts the mining of a coal seam that is 5 ft (1.5 m) thick to removing 100 ft (30 m) of overburden or less, the economic strip ratio is 20 to 1.

Quarrying. Surface mining of valuable stones, such as building stone and limestone, are called quarries (**Fig. 4**). A typical quarry operation removes only a small amount of overburden, while most of the rock mined is sold for crushed rock. The rock is typically broken by blasting, loaded into trucks, and hauled to a processing facility, often a crusher. Crushers are mechanical devices used to break rocks into smaller sizes. Large screens are used to separate the various rock sizes. Ornamental rock quarries, such as marble, may cut the rock to dimensions requested by a purchaser. Most of the rock mined is sold, with a large opening remaining after mining, as well as a near-vertical rock slope called a highwall. To restore a quarry, rock and soil are placed against the highwall to create a stable slope, which is subsequently covered by topsoil and revegetated. *See* QUARRYING.

Equipment. Productivity improvements in the surface mining industry have resulted principally from the increased size and efficiency of the equipment used. Steam-powered shovels were primarily responsible for creating the large-scale surface mining industry. They have been replaced by electric- or diesel-powered shovels. *See* POWER SHOVEL.

Surface mining sites use the same type of equipment as the construction industry. The major types of equipment include bulldozers, drilling and blasting equipment, wheel loaders, hydraulic excavators, draglines, off-road trucks, mechanical

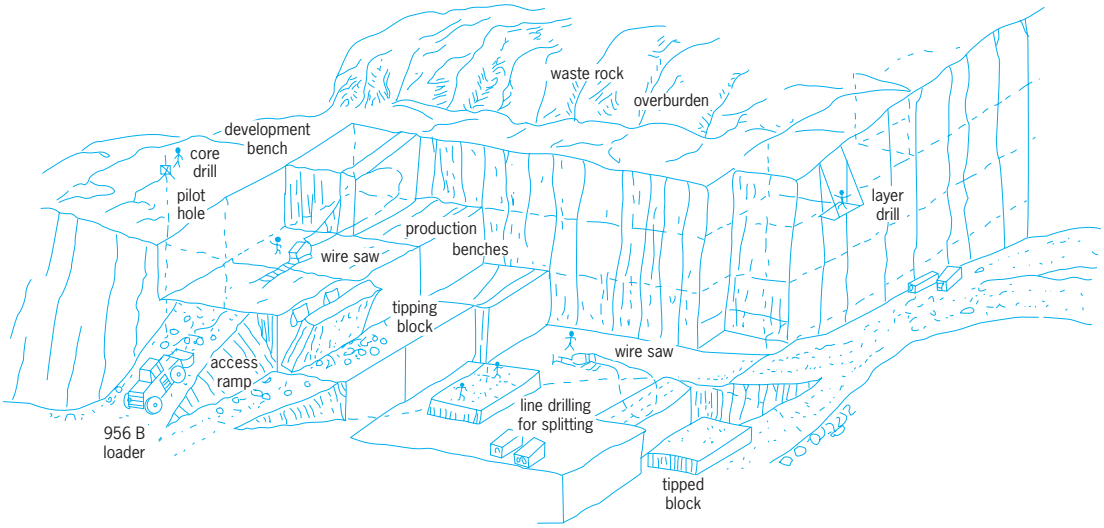

Fig. 4. Quarrying.

shovels, bucket wheel excavators, and scrapers. *See* CONSTRUCTION EQUIPMENT.

A bulldozer is a tractor with a large blade attached to its front, and is propelled using crawler tracks. Bulldozers can be used to relocate soil and rock over very short distances (less than 1000 ft or 300 m) and are typically used for leveling and grading soil during reclamation.

Large drilling machines are used to create vertical boreholes 6–12 in. (15–30 cm) in diameter in the overburden. Explosives inserted into the boreholes are detonated to break up the overburden so that it can be easily excavated.

Wheel loaders are commonly called front-end loaders. They are rubber-tired tractors with a large bucket in front. The bucket is filled by pushing it into overburden or ore, lifted and maneuvered by a hydraulic mechanism; and then its contents are dumped into a truck bed.

A hydraulic excavator is a backhoe with large bucket attached to a long boom. The bucket is filled with overburden or ore by being pulled back toward the operator by a hydraulic mechanism. These are quite useful when working on flat terrain and excavating material lower than the base of the excavator.

Draglines are machines that have a very large bucket suspended from a boom by cables (**Fig. 5**). The bucket is cast a considerable distance from the machine by swinging and dropping it in one motion. The bucket is filled with overburden by dragging it back toward the machine by retracting the cables. The filled bucket is then lifted and rotated to the disposal area and dumped by lifting the rear of the bucket, again by retracting the cables. Draglines have bucket capacity up to 220 yd^3 (168 m^3), and some weigh over 10,000 tons (9070 metric tons).

Haulage trucks used at mine sites have much higher capacity than trucks built for highway use, with some able to carry over 300 tons (272 metric tons).

Bucketwheel excavators can be used to excavate overburden that is unconsolidated and does not contain large rocks (**Fig. 6**). They continuously excavate using a rotating wheel of buckets (similar to a waterwheel) that scoops up and drops overburden onto a conveyor.

Scrapers are rubber-tired machines that can be used to remove and transport unconsolidated overburden. They are loaded by pushing forward a metal plate that scrapes and lifts dry, unconsolidated overburden into a hopper in the center of the machine. After hauling, the overburden is unloaded through an opening at the bottom of the hopper while moving forward. A scraper can be used, for example, to deposit overburden at an area being restored.

Industry statistics. According to the Mine Safety and Health Administration, there were about 11,000 surface mines in the United States in 2001. The sand and gravel industry had the most operations with 7126 mines, including 745 dredging operations. In the United States, all the sand and gravel is produced by surface mining. Stone quarries are the second most numerous type of surface mine. Coal-, nonmetal-, and metal-producing surface mines numbered less than 1000 mines each. In comparison, the total number of underground mining operations

Fig. 5. Large dragline operating from the top of the highwall.

in the United States during 2001 was 851, including 730 underground coal mines. *See* UNDERGROUND MINING.

In 2001, approximately 115,000 people were employed by surface mining operations in the United States (see **table**). The sand and gravel industry had the most employees. The largest surface mining operations, based on the average number of employees per mine, were metal-producing mines, principally the large western United States open-pit mines.

In the United States, the electric utility industry produces about 50% of its power from coal and consumes most of the coal produced there. Over 1 billion tons (900 million metric tons) of coal is produced annually in the United States, of which over 700 million tons (635 million metric tons) comes from surface mines. In 2001, the four states producing the most coal by surface mining were Wyoming, 350 million tons (318 million metric tons); West Virginia, 60 million tons (54 million metric tons); and Kentucky and Texas (primarily lignite) at 48 million tons (44 million metric tons) each.

Regulations. The surface mining industry is governed by safety and environmental laws, enacted by federal, state, and local governments. Agencies created by the governmental bodies typically issue regulations to enforce the laws. Federal agencies that regulate surface mining include the Mine Safety and Health Administration, Environmental Protection Agency, Office of Surface Mining, Bureau of Land Management, U.S. Forest Service, National Park Service, and Bureau of Indian Affairs.

Safety. Federal health and safety standards applicable to all surface mines are enforced by the Mine Safety and Health Administration. The Office of Surface Mining enforces the federal laws at coal mines concerning water and air pollution standards, while allowing individual states to have primacy over the industry if they enact laws that mostly mirror the federal laws. For these states, the Office of Surface Mining would merely oversee the state enforcement program. Many mining states have enacted such laws.

Environmental protection. Prior to 1977, individual states regulated the environmental practice of the surface coal mining industry. In 1977, a federal Surface Mining Law was enacted. This law established the Office of Surface Mining and provided program regulations, including mining and reclamation standards and the requirements to obtain written permits to mine.

Early surface mining in the United States did not employ the industry standards that are now used. Mining sites were generally not restored after the mining was completed. This resulted in unsightly sites, as well as hazardous conditions, poor vegetation, and water pollution. Modern surface mines have high reclamation standards. They restore the mined area to approximately the same contour as previously existed, revegetate with native grasses and trees, and comply with water pollution standards. Industry and environmental organizations such as the American Society of Mining and Reclamation were formed to promote reclamation practices for lands disturbed during mineral extraction.

Fig. 6. Bucketwheel excavators. (*a*) Large type, daily capacity 144,000 yd^3 (110,000 m^3). (*b*) Smaller model.

Surface mining statistics

	Number of surface mines	Number of employees	Average number of employees per mine
Sand and gravel	7,126*	37,508*	5
Coal	852	28,304	33
Stone	2,412	32,311	13
Metal	89	11,076	124
Nonmetal	418	5,561	13
Total	10,897	114,760	11

*Includes 745 dredging operations.
SOURCE: Mine Safety and Health Administration.

Previously unreclaimed surface mine sites are also being restored by the surface mining industry. Some are restored directly in conjunction with the mining of adjacent mineral resources. *See* LAND RECLAMATION.

Thomas Gray

Bibliography. S. M. Cassidy (ed.), *Elements of Practical Coal Mining*, SME/AIME, New York, 1973; J. T. Crawford III and W. A. Hustrulid (eds.), *Open Pit Mine Planning and Design*, SME/AIME, New York, 1979; *Glossary of Surface Mining and Reclamation Terminology*, 2d ed., Bituminous Coal Research, Monroeville, PA, 1983; T. A. Gray and R. E. Gray, Mine closure, sealing and abandonment (Chap. 8.7), in H. L. Hartman et al. (eds.), *SME Mining Engineering Handbook*, 2d ed., Society for Mining Metallurgy & Exploration, 1992; B. A. Kennedy (ed.), *Surface Mining*, 2d ed., Society for Mining Metallurgy & Exploration, 1990; C. T. Shaw and V. Pavlovic, *Surface Mining and Quarrying: Mechanization, Technology, and Capacity*, Ellis Horwood, 1993; P. W. Thrush (ed.), *A Dictionary of Mining, Mineral, and Related Terms*, Intertech Publishing, 1990.

Surface physics

The study of the structure and dynamics of atoms and their associated electron clouds in the vicinity of a surface, usually at the boundary between a solid and a low-density gas. Surface physics deals with those regions of large and rapid variations of atomic and electron density that occur in the vicinity of an interface between the two "bulk" components of a two-phase system. In conventional usage, surface physics is distinguished from interface physics by the restriction of the scope of the former to interfaces between a solid (or liquid) and a low-density gas, often at ultrahigh-vacuum pressures $p = 10^{-10}$ torr (1.33×10^{-8} newton/m^2 or 10^{-13} atm). *See* SOLID-STATE PHYSICS.

Surface physics is concerned with two separate but complementary areas of investigation into the properties of such solid-"vacuum" interfaces. Interest centers on the experimental determination and theoretical prediction of surface composition and structure (that is, the masses, charges, and positions of surface species), of the dynamics of surface atoms (such as surface diffusion and vibrational motion), and of the energetics and dynamics of electrons in the vicinity of a surface (such as electron density profiles and localized electronic surface states). As a practical matter, however, the nature and dynamics of surface species are determined experimentally by scattering and emission measurements involving particles or electromagnetic fields (or both) external to the surface itself. Thus, a second major interest in surface physics is the study of the interaction of external entities (that is, atoms, ions, electrons, electromagnetic fields, or mechanical probes) with solids at their vacuum interfaces. It is this aspect of surface physics that most clearly distinguishes it from conventional solid-state physics, because quite different scattering, emission, and local probe experiments are utilized to examine surface properties as opposed to bulk properties.

Physical principles of measurements. Since the mid-1960s, surface physics has enjoyed a renaissance by virtue of the development of a host of techniques for characterizing the solid-vacuum interface. These techniques are based on one of three simple physical mechanisms for achieving surface sensitivity. The first, which is the basis for field emission, field ionization, and scanning tunneling microscopy (STM), is the achievement of surface sensitivity by utilizing electron tunneling through the potential-energy barrier at a surface. This concept provides the basis for the development of STM to directly examine the atomic structure of surfaces by measuring with atomic resolution the tunneling current at various positions along a surface. It also has been utilized for direct determinations of the energies of individual electronic orbitals of adsorbed complexes via the measurement of the energy distributions either of emitted electrons or of Auger electrons emitted in the process of neutralizing a slow (energy $E \sim 10$ eV) external ion. *See* FIELD-EMISSION MICROSCOPY; SCANNING TUNNELING MICROSCOPE; TUNNELING IN SOLIDS.

The second mechanism for achieving surface sensitivity is the examination of the elastic scattering or emission of particles which interact strongly with the constituents of matter, for example, "low-energy" ($E \lesssim 10^3$ eV) electrons, thermal atoms and molecules, or "slow" (300 eV $\lesssim E \lesssim 10^3$ eV) ions. Since such entities lose appreciable ($\Delta E \sim 10$ eV) energy in distances of the order of tenths of a nanometer, typical electron analyzers with resolutions of tenths of an electronvolt are readily capable of identifying scattering and emission processes which occur in the upper few atomic layers of a solid. This second mechanism is responsible for the surface sensitivity of photoemission, Auger electron, electron characteristic loss, low-energy electron diffraction (LEED), and ion scattering spectroscopy techniques. The strong particle-solid interaction criterion that renders these measurements surface-sensitive is precisely the opposite of that used in selecting bulk solid-state spectroscopies. In this case, weak particle-solid interactions (that is, penetrating radiation) are desired in order to sample the bulk of the specimen via, for example, x-rays, thermal neutrons, or fast ($E \gtrsim 10^4$ eV) electrons. These probes, however, can sometimes be used to study surface properties by virtue of special geometry, for example, the use of glancing-angle x-ray diffraction to determine surface atomic structure. *See* AUGER EFFECT; ELECTRON DIFFRACTION; ELECTRON SPECTROSCOPY; PHOTOEMISSION; X-RAY CRYSTALLOGRAPHY.

The third mechanism for achieving surface sensitivity is the direct measurement of the force on a probe in mechanical contact or near contact with the surface. At near contact, the van der Waals force can be measured directly by probes of suitable sensitivity. After contact is made, a variety of other forces dominate, for example, the capillary force for solid

surfaces covered with thin layers of adsorbed liquid (that is, most solid surfaces in air at atmospheric pressure). When this mechanism is utilized via measuring the deflection of a sharp tip mounted on a cantilever near a surface, the experiment is referred to as atomic force microscopy (AFM) and results in maps of the force across the surface. Under suitable circumstances, atomic resolution can be achieved by this method as well as by STM. Atomic force microscopy opens the arena of microscopic surface characterization of insulating samples as well as electrochemical and biochemical interfaces at atmospheric pressure. Thus, its development is a major driving force for techniques based on surface physics. *See* INTERMOLECULAR FORCES.

Surface preparation. An atomically flat surface, labeled by $M(bkl)$, may be visualized as being obtained by cutting an otherwise ideal, single-crystal solid M along a lattice plane specified by the Miller indices (bkl), and removing all atoms whose centers lie on one side of this plane. On such a surface the formation of a "selvedge" layer can also be envisaged. Such a layer might be created, for example, by the adsorption of atoms from a contiguous gas phase. It is characterized by the fact that its atomic geometry differs from that of the periodic bulk "substrate." From the perspective of atomic structure, this selvedge layer constitutes the "surface" of a solid. A good example is the thin film of condensed water that exists on most surfaces at atmospheric pressure. In principle, the thickness of the selvedge layer is a thermodynamic variable determined from the equations of state of the solid and the contiguous gas phase. In practice, almost all solid surfaces are far from equilibrium, containing extensive regions (micrometers thick) of surface material damaged by sample processing and handling. *See* CRYSTALLOGRAPHY.

Another reason for the renaissance in surface physics is the capability to generate in a vacuum chamber special surfaces that approximate the ideal of being atomically flat. These surfaces may be prepared by cycles of fast-ion bombardment, thermal outgassing, and thermal annealing for bulk samples (for example, platelets with sizes of the order of 1 cm × 1 cm × 1 mm), molecular beam epitaxy of a thin surface layer on a suitably prepared substrate, or field evaporation of etched tips for field-ion microscopes. Alternatively, the sample may be cleaved in a vacuum chamber. In such a fashion, reasonable facsimiles of uncontaminated, atomically flat solid-vacuum interfaces of many metals and semiconductors have been prepared and subsequently characterized by various spectroscopic techniques. Such characterizations must be carried out in an ultrahigh vacuum ($p \sim 10^{-8}$ N/m^2) so that the surface composition and structure are not altered by gas adsorption during the course of the measurements. *See* EPITAXIAL STRUCTURES.

Experimental apparatus. Modern experimental surface physics is devoted to the determination of the chemical composition, atomic geometry, atomic dynamics, and electronic structure of surfaces. Since different measurements are required to assess each of these four aspects of a surface, the typical surface-characterization instrument consists of equipment for performing multiple measurements in a single ultrahigh-vacuum chamber or for moving a sample between multiple connected chambers without breaking vacuum. Two types of sample geometry are common. Platelet or wafer samples are studied using scattering and emission experiments. A typical modern apparatus (**Fig. 1**) contains an electron gun, an ion gun, an electron energy analyzer, a source of ultraviolet or x-ray electromagnetic radiation, a local probe microscope, and a sample holder permitting precise control of both its orientation and temperature. Occasionally, other features (such as a mass spectrometer) also are incorporated for special purposes. For specific applications in which less than a complete characterization of the surface is required, commercial instruments designed to embody only one or two measurements often are available. Such instruments commonly are utilized to determine the chemical composition of surfaces by, for example, ion scattering, secondary ion mass spectrometry, x-ray photoemission, or Auger electron emission. Similarly, local probe scanning tunneling microscopes and atomic force microscopes are available commercially. Obviously, the utility or such instruments is not limited to atomically flat or even crystalline surfaces, so that they find widespread applications in metallurgy and polymer science. Analogous measurements also are commonly utilized for the in situ characterization of materials made by vacuum deposition techniques, such as sputtering or molecular beam epitaxy. *See* MASS SPECTROMETRY; SPUTTERING.

The second common sample geometry is an etched tip, about 100 nm in radius. Such specimens are studied by field emission and ionization experiments. These experiments provide a direct magnified image of the surface structure in contrast to the statistical description of platelet or wafer surfaces afforded by instruments like that in Fig. 1 or the composite local images of such surfaces afforded by STM or AFM.

Data acquisition, analysis, and theory. Given the ability to perform surface-sensitive spectroscopic measurements, questions naturally arise concerning analysis of the raw spectra to extract parameters characterizing the structure of a given surface and the synthesis of such data to form a coherent picture of the behavior of electrons and atomic species at the surface. Thus, surface physics may be divided into three types of activity: the acquisition of surface-sensitive spectroscopic data, the analysis of these data using physical models of the appropriate scattering or emission spectroscopy, and the construction of theoretical models of surface structure and properties to be tested via critical comparison of their predictions with the results of such data analyses.

Ground- and excited-state properties. Theoretical models have been proposed for the description of two distinct types of surface properties. The stability of surface structures is examined by calculations of ground-state properties, such as surface energies

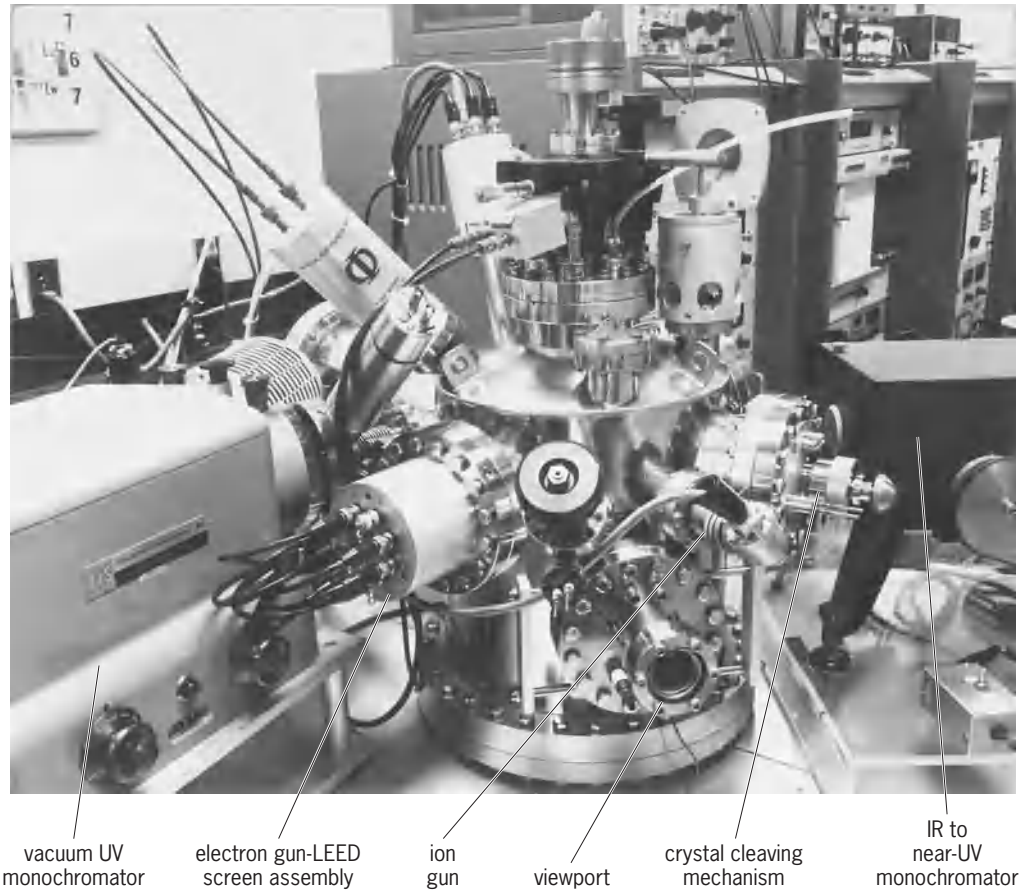

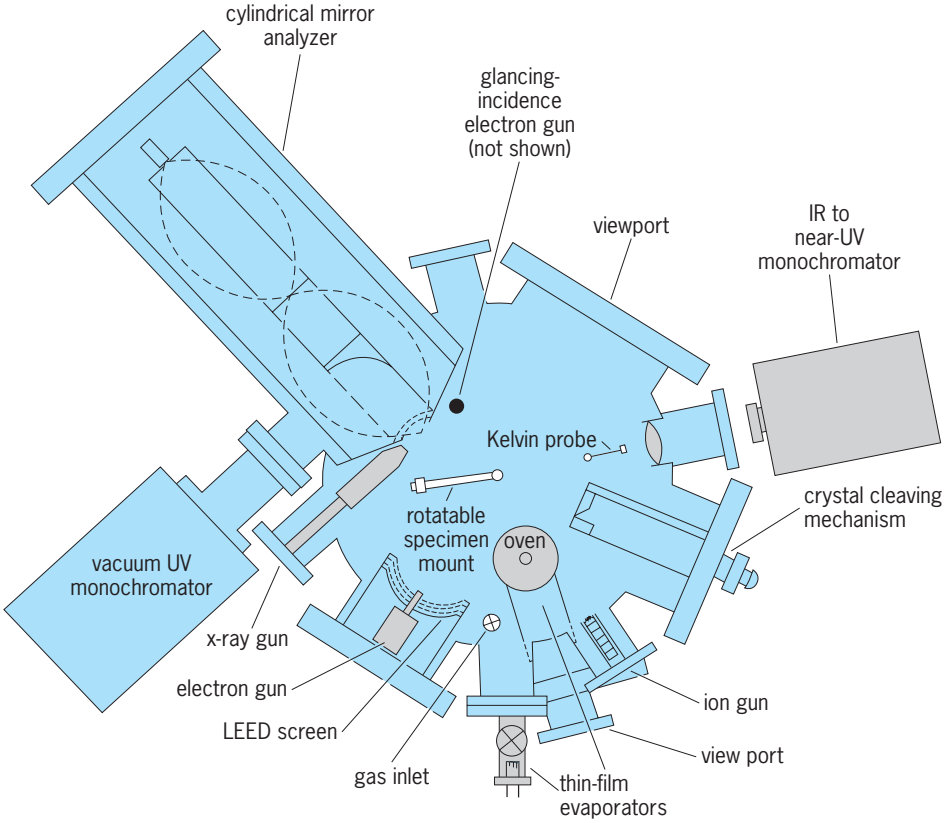

Fig. 1. Photograph and schematic diagram of a multiple-technique ultrahigh-vacuum surface characterization instrument for the study of semiconductor and insulator surfaces. (*Courtesy of L. J. Brillson, Xerox Corporation*)

or effective potential-energy diagrams for adsorbed species. These quantities are difficult to measure experimentally, although they are the most direct manifestations of the intrinsic behavior of an undisturbed surface. The interactions of external projectiles or fields with a solid create excited states of the electrons or atoms within the solid. Consequently, the associated scattering and emission spectra indicate the nature and energies of these excited states (called excitations) rather than of the ground state. Two kinds of excitations occur. Electronic excitations are generated when a disturbing force causes the electrons in the solid to alter their quantum states, whereas atomistic excitations are associated with the vibration or diffusion of atomic species (such as adsorbed atoms of molecules). It is important to distinguish between ground-state properties, electronic excitations, and atomic excitations because different models are used to describe each of these three types of phenomena. *See* EXCITED STATE; GROUND STATE.

Quantum theory of surfaces. The theory of the properties of solid surfaces does not differ in any fundamental way from the quantum theory of bulk solids. Specifically, the conventional quantum theory of interacting electron systems is thought to be applicable, although technical refinements are required because of the loss of translation symmetry and the presence of large electron density gradients normal to the surface. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM THEORY OF MATTER.

Macroscopic models. It is premature to speak of an embracing theory of surface phenomena. Rather, a diverse array of specific models has been proposed for the description of various properties. In the case of macroscopic models, the presence of a surface is treated as a boundary condition on an otherwise continuum theory of bulk behavior. Such models have found widespread use in semiconductor and insulator physics because the penetration depth of electrostatic fields associated with surface charges usually is large ($\lambda_e \sim 10^4$–10^5 nm) relative to the spatial extent of the charges themselves ($d \lesssim 1$ nm). Moreover, they continue to describe adequately the electrostatic fields involved in the operation of microelectronic semiconductor devices down to submicrometer dimensions. To describe the atomic and electronic properties associated with the upper few atomic layers at a surface, however, one must make use of a description of surfaces at the atomic or electronic level. This is becoming an issue for semiconductor device technology as the size of these devices approaches 0.1 μm and less. *See* SEMICONDUCTOR.

Microscopic models. Four major classes of microscopic models of surface properties have been explored. The simplest consists of models in which consideration of the electronic motion is suppressed entirely, and the solid is visualized as composed of atomic species interacting directly, for example, via two-body forces. While such models may suffice to describe the vibrational motion of atoms near a surface, they are inadequate to describe ground-state properties such as adsorbate potential-energy curves. Nevertheless, such models can be used to describe interactions between steps and kinks formed on a surface by growth or annealing. In this context they have found considerable favor to describe the evolution of the morphologies of surfaces during growth and processing.

The next more sophisticated models are empirical quantum-chemical models (such as tight-binding or empirical pseudopotential models, in solid-state terminology), in which electronic motions are considered explicitly but electron-electron interactions are incorporated either phenomenologically or not at all. Such models have proved useful in solid-state physics, although their value for surface physics is more limited because the large charge rearrangements (relative to the bulk) which occur at surfaces require an accurate, self-consistent treatment of both electron-electron and electron-ion interactions.

The simplest model in which electron-electron interactions are treated explicitly is the jellium model of metals, in which the positive charge associated with the ion cores immersed in the sea of conduction electrons is replaced by a uniform positive background charge terminating along a plane. This model, popular in the early 1970s, permits an informative but approximate treatment of electron-electron interactions at the expense of losing the effects of atomic lattice structure because of the uniform-positive-background hypothesis. Since 1990, computers have permitted the construction of realistic pseudopotential, local-density, and quantum-chemical models in which both the electron-electron and electron-ion interactions are treated in a self-consistent, if approximate, fashion. Such models are now routinely applied to predict ground-state electronic charge densities and surface excitation spectra (both electronic and atomic) at the low-index surfaces of essentially any metal, semiconductor, or insulator. The major tests of their adequacy arise from comparisons of their predictions with measured atomic geometries, work functions, photoemission spectra, and characteristic electron-loss spectra. Quite detailed calculations of the work functions, surface-state eigenvalue spectra, and surface geometries are now available for a variety of metal and semiconductor surfaces. The geometry of both metal and semiconductor surfaces can be predicted quantitatively by energy-minimization techniques applied to these models. Moreover, they are currently being extended to study atomic motions at surfaces associated with crystal growth and the dynamics of chemical reactions at the atomic level. *See* CRYSTAL GROWTH; QUANTUM CHEMISTRY.

Theoretical models for data analysis. Another group of theoretical models in surface physics consists of those utilized to analyze observed scattering and emission spectra in order to extract quantitative assessments of the atomic and electronic structure of surfaces. These models differ substantially from their bulk counterparts because of the necessity of strong particle-solid interactions to achieve surface-sensitive spectroscopies. Consequently, the fundamental assumption underlying the linear-response

theory of bulk solid-state spectroscopies—that is, the appropriate particle-solid interaction is weak and hence can be treated by low-order (usually first) perturbation theory—is invalid. This fact results in collision theories of surface-sensitive particle-solid scattering exhibiting a considerably more complicated analytical structure in order to accommodate the strong elastic as well as inelastic scattering of the particle by the various constituents of the solid.

Applications to LEED and photoemission. While the above considerations are quite general, the special case in which they have been developed in most detail is the coherent scattering (that is, diffraction) of low-energy electrons from the surfaces of crystalline solids. This is an important case because elastic low-energy electron diffraction (LEED) is the analog of x-ray diffraction for surfaces—that is, it is the major vehicle for the achievement of a quantitative surface crystallography. Since 1968 quite complete quantum field theory models of the LEED process have been developed, tested, and reduced to computational algorithms suitable for the routine analysis of LEED intensity data. From such analyses the surface atomic geometry of the low-index faces of a host of metals and semiconductors has been determined, as have the geometries of a wide variety of overlayer structures. Local surface structures of adsorbates have also been determined by applying these techniques to analysis of the diffraction of photoemitted electrons from the adsorbates, and to the diffuse scattering of incident electrons by disordered adsorbate overlayers. Atomic geometries of semiconductors generated in situ by molecular beam epitaxy are determined by the application of these techniques to analyze the diffracted intensities of glancing-incidence high-energy (keV) electrons via reflection high-energy electron diffraction (RHEED). Similar quantitative analyses of inelastic low-energy electron diffraction intensities have yielded the energy-momentum relations of collective surface electronic excitations (for example, surface plasmons). *See* PLASMON.

Data acquisition. It is the development of a host of novel surface-sensitive spectroscopic techniques, however, which has provided the foundation for the renaissance in surface physics. Having recognized that low-energy electrons, thermal atoms, and slow ions all constitute surface-sensitive incoming or exit entities in particle-solid collision experiments, one can envisage a wide variety of surface spectroscopies based on these plus quanta of electromagnetic radiation (photons) as possible incident or detected species. Most of these possibilities have been realized in some form. The selection of which technique to use in a particular application depends both upon what one wishes to learn about a surface and upon the relative convenience and destructiveness of the various measurements.

Typically, one wishes to determine the composition and atomic structure of a surface region, and often to ascertain its electronic structure as well. In the measurements of any of these quantities, important issues are the lateral and depth resolution of the possible techniques. In the case of scattering-based probes, the depth resolution is determined by the particle-solid force law of the incident and exit particles, high resolution being associated with stronger inelastic collision processes. The lateral spatial resolution depends on the ability to focus the incident beam. For typical focused-beam experiments, this is on the order of 1 cm^2 for photon beams, 10^{-8} to 10^{-12} cm^2 for electron beams, and 10^{-8} cm^2 for ion beams. Thus, scanning microscopies are both feasible and common with electron and ion beams. For photon beams, more sophistication is required. By using shaped fiber or semiconductor tips as light pipes and an appropriate local probe geometry, scanning near-field optical microscopy (SNOM) can yield useful images with lateral resolutions comparable to and below the wavelength of the incident light. Depth resolution is a single monolayer for thermal-atom and slow-ion scattering, and a few monolayers for slow-electron scattering. It can become 1000 atomic layers or more, however, for fast (MeV) ions and fast (10-keV) electrons. *See* OPTICAL MICROSCOPE.

The development of local probe STM and AFM methods has revolutionized surface physics by providing readily accessible maps of the geometry, from atomic to morphological distance scales (for example, steps or islands), of complex surfaces. The interpretation of the raw data is more complicated than commonly realized, because the depth resolution of these techniques is so high. For example, STM measures the electronic charge density in a region a fraction of a nanometer outside the geometrical surface defined by the position of the nuclei of the surface atoms. Similar caveats apply to SNOM, for which quantitative determination of the electromagnetic fields in the vicinity of the probe tip is required for precision data analysis. With the proper care in interpreting the data, however, these local probes offer the possibility of the application of surface physics techniques to characterize a wide range of surfaces, interfaces, and atomic processes (such as crystal growth or surface chemical reactions).

Surface composition. The elemental composition of surfaces is specified by measuring the masses or atomic numbers, or both, of resident species. Their masses may be ascertained either by the elastic backscattering of slow incident ions (ion scattering spectrometry, or ISS) or by using such ions to erode the surface, detecting the ejected surface species in a mass spectrometer (secondary ion mass spectrometry, or SIMS). *See* SECONDARY ION MASS SPECTROMETRY (SIMS).

The atomic numbers of surface species are determined by measuring the energy of tightly bound core electrons. A schematic diagram illustrating the nature and labeling of the various physical processes that can be utilized to accomplish this task is shown in **Fig. 2**. An electron, photon, or chemical species incident on a surface excites a low-energy core electron. The binding energy of this electron commonly is determined by measuring the energy loss of the incident electron (characteristic loss spectroscopy,

or CLS), the energy of the core electron ejected by an incident x-ray photon (x-ray photoelectron spectroscopy, or XPS, sometimes referred to as electron spectroscopy for chemical analysis, or ESCA), or the threshold energy of an incident particle necessary to generate a threshold in the secondary x-ray yield (soft x-ray appearance potential spectroscopy, or SXAPS). Alternatively, the binding energy of the core electron may be ascertained by secondary processes in which an initially empty core state (generated by a direct process) is filled by an electron in a higher-energy state. If the filling process is radiative recombination, then the energy of the emitted x-ray yields the binding energy (soft x-ray emission spectroscopy). If this process is radiationless, however, the energy of the electron excited by the Auger process indicates the binding energy of the initially empty core state (Auger electron spectroscopy, or AES).

These techniques operate on the dictionary premise; that is, calibration spectra are obtained on surfaces of independently known composition, with elemental analysis on unknown samples being performed by comparison of their spectra with the reference calibration spectra. Consequently, although the detailed interpretations of observed line shapes often has eluded surface physicists, the use of these spectroscopies for elemental analysis has proved both practical and eminently useful. Difficulties in interpretation usually preclude the use of these techniques for quantitative chemical analysis (for example, the determination of whether C and O are adsorbed on aluminum as CO, as CO_2 or C on Al_2O_3, and so on). Progress has been made, however, in developing this aspect of the core-electron spectroscopies. *See* ATOMIC SPECTROMETRY.

Surface atomic geometry. The atomic geometry of planar surfaces of crystalline solids usually is obtained by electron diffraction, although in certain cases slow-ion backscattering, megaelectronvolt ion channeling, photoelectron diffraction, surface-sensitive extended x-ray absorption fine structure, or valence-electron photoemission spectroscopy also may be employed. Two experimental configurations commonly are used, as indicated in **Fig. 3**. The RHEED configuration embodies glancing incidence electrons at kiloelectronvolt energies. It yields primarily the space-group symmetry of the surface and is quite sensitive to surface topography. The LEED experiment consists of measuring the backscattering intensities of electrons in the energy range 50 eV $\lesssim E \lesssim$ 500 eV. The configuration of diffracted beams reveals the space group symmetry of the surface structure, whereas analysis of their intensities permits determination of their atomic geometry.

Surface atomic geometry may also be determined by direct imaging. In the case of tip sample geometries, field ion microscopy permits the direct imaging of atoms on the tip surface. For platelet and wafer samples, the local probe STM and AFM methods afford scanning determinations of the local charge density and the force, respectively, just outside the surface. These are related to the atomic geometry of the surface species, but only indirectly. Nevertheless, maps of tip height (used in STM) or cantilever deflection (used in AFM) versus lateral position along the surface are commonly interpreted as reflecting the surface atomic geometry. While this is rarely adequate for precision determination of surface atomic positions, it is very useful for the observation of qualitative details of the surface atomic geometry, especially for surface defects.

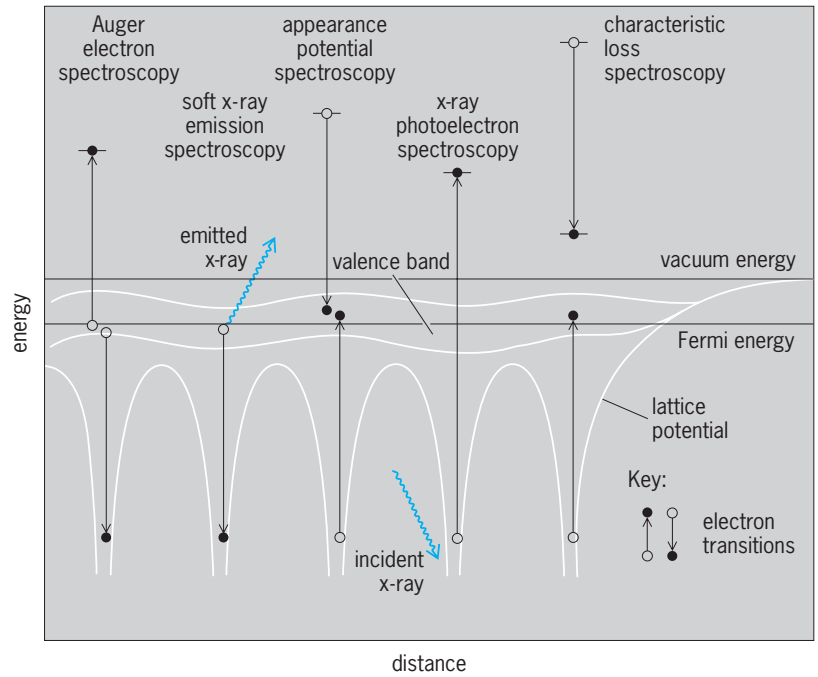

Fig. 2. Schematic diagram of the core-electron transitions utilized to ascertain the atomic number of surface species. (*After R. L. Park, Inner-shell spectroscopy, Phys. Today, 28(4):52–59, April 1975*)

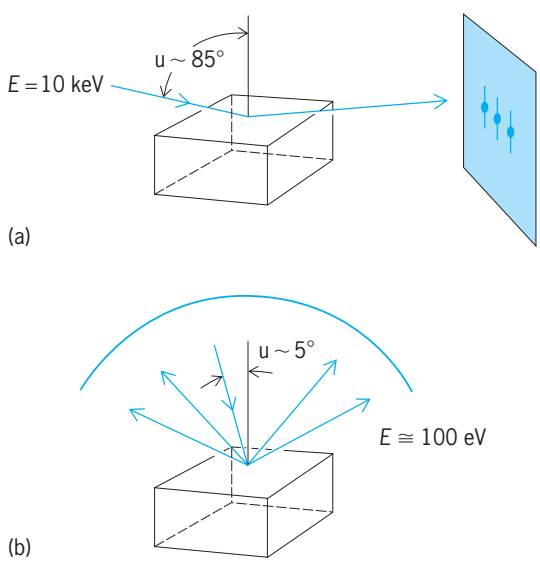

Fig. 3. Schematic diagram of the two electron diffraction techniques used to determine the precision atomic geometry of single-crystal surfaces. (*a*) Reflection high-energy electron diffraction (RHEED). (*b*) Elastic low-energy diffraction (LEED). (*After C. B. Duke, in I. Prigogine and S. A. Rice, eds., Aspects of the study of surfaces, Adv. Chem. Phys., 27(1):1–209, 1974*)

Surface atomic motion. The vibrational motion of surface species may be examined either by analysis of the temperature dependence of LEED intensities or by direct observation of small ($\Delta E \sim 0.01$ eV) electron energy losses caused by the excitation of a normal mode of vibration. The first approach provides the rms vibrational amplitudes of surface species, whereas the second yields the frequencies of localized surface normal modes of vibration. Both high-resolution electron energy-loss spectroscopy (HREELS: $\Delta E \sim 0.01$ eV) and infrared reflection and transmission spectroscopy have become popular techniques for the measurements of the vibrational spectra of adsorbed species. The complete energy momentum relations of surface vibrational excitations have been determined by inelastic helium (He) atom scattering. *See* LATTICE VIBRATIONS.

Surface electronic structure. The electronic structure of a solid-vacuum interface is studied by measuring the emission of valence electrons (induced by external fields, electrons, ions, or photons) or the inelastic scattering of an incident electron. A special situation arises when an emitted Auger electron is a valence electron. In this case, the initially empty core state is highly localized in space. Consequently, the emission line shape is a measure of the local electronic structure in the vicinity of this core state. The shifts in energy of core-level photoemission and Auger transitions (called chemical shifts) caused by the nearby electronic charge densities also yield an indication of the local electronic structure around a particular kind of surface atom.

In contrast to these emission processes involving localized core electrons, the photoemission, field emission, ion neutralization, and characteristic loss spectroscopies of valence electrons provide measures of their average behavior in the vicinity of a surface. Indeed, for precisely this reason such spectra from clean surfaces may be difficult to interpret because the distinction between bulk and surface features is often vague. Thus, their major use has occurred in the arena of chemisorption, in which case the changes in spectra upon adsorption can be monitored, and qualitative features of the electronic structure of the chemisorbed complexes can be inferred therefrom. A proper theoretical analysis of these valence electron emission processes is needed to convert the observed spectra into quantitative indicators of surface structure. Such analyses are not yet routine, although they have been given for a few exemplary materials such as clean nickel and gallium arsenide (GaAs) surfaces and a few chemisorption systems. The advent of synchrotron radiation has provided a major incentive for the development of computer programs to perform such analyses. *See* ADSORPTION; SYNCHROTRON RADIATION.

Several measurements probe the electric fields just outside the surface. Scanning tunneling microscopy measures the local electron density about 0.1 nm outside a surface. The effective potential seen by an atom scattering from a surface is also quite sensitive to the electronic structure in this region. Finally, work function measurements probe the dipole layer between the valence electrons and surface layer of ion cores in the solid. *See* WORK FUNCTION (ELECTRONICS).

Outlook. Since 1975, surface physics has advanced from a state in which the atomic composition, geometry, and atomic and electronic excitation spectra were essentially unknown, to one in which these quantities are routinely determined for the low-index faces of elements and binary alloys. Local probe microscopes have extended the reach of surface physics techniques from carefully prepared surfaces in ultrahigh vacuum to the domains of the biologist, chemist, and technologist. Surface physics is used to explore novel phenomena and characterize complex materials systems, especially various manifestations of atomic dynamics at surfaces and interfaces.

C. B. Duke

Bibliography. M. C. Desjonqueres and D. Spanjaard, *Concepts in Surface Physics*, 1993; C. B. Duke (ed.), *Surface Science: The First Thirty Years*, 1994; J. B. Hudson, *Surface Science: An Introduction*, 1991; J. Israelachvili, *Intermolecular and Surface Forces*, 1992; H. Luth, *Surfaces and Interfaces of Solids*, 1993; S. R. Morrison, *The Chemical Physics of Surfaces*, 2d ed., 1990; D. W. Pohl and D. Courjon (eds.), *Near Field Optics*, 1992; M. Prutton, *Introduction to Surface Physics*, 1994; W. N. Unertl (ed.), *Handbook of Surface Science I: Physical Structure*, 1996; M. A. van Hove et al. (eds.), *The Structure of Surfaces*, vols. 1–4, 1985–1994; J. M. Walls, *Surface Science Techniques*, 1994; D. P. Woodruff and T. A. Delchar, *Modern Techniques of Surface Science*, 2d ed., 1994.

Surface tension

The force acting in the surface of a liquid, tending to minimize the area of the surface. Surface forces, or more generally, interfacial forces, govern such phenomena as the wetting or nonwetting of solids by liquids, the capillary rise of liquids in fine tubes and wicks, and the curvature of free-liquid surfaces. The action of detergents and antifrothing agents, and the flotation separation of minerals depend upon the surface tensions of liquids.

Surface energy. In the body of a liquid, the time-average force exerted on any given molecule by its neighbors is zero. Even though such a molecule may undergo diffusive displacements because of random collisions with other molecules, there exist no directed forces upon it of long duration. It is equally likely to be momentarily displaced in one direction as in any other. In the surface of a liquid, the situation is quite different; beyond the free surface, there exist no molecules to conteract the forces of attraction exerted by molecules in the interior for molecules in the surface. In consequence, molecules in the surface of a liquid experience a net attraction toward the interior of a drop. These centrally directed forces cause the droplet to assume a spherical shape, thereby minimizing both the free energy and surface area.

From the macroscopic point of view, surface tension may be regarded either as a force exerted normally to a unit length in the surface, or as the work which must be expended upon the liquid to increase its area by unity. Accordingly, surface tension is expressed in SI units of newtons per meter (N/m) or joules per square meter (J/m^2). From the microscopic point of view, the surface tension (or its equivalent, surface energy) is the reversible isothermal work which must be done in bringing molecules from the interior of the liquid to the surface and creating 1 cm^2 of new surface thereby.

Most liquids have surface tensions of 20–40 millinewtons per meter at room temperature, but water has the exceptionally high value of 72.75 mN/m at 20°C (68°F). Condensed gases such as helium and nitrogen have quite low surface tensions [98 micronewtons per meter (−452°F) and 6.2 mN/m at 90.2 K (−297.2°F) respectively]. Liquid metals have large surface tensions by comparison: mercury, 0.47 N/m; and liquid copper at 1131°C (2068°F) has a surface tension of 1.103 N/m in hydrogen gas. Small but significant differences in the surface tensions of liquids depend upon the composition of the vapor phase.

In the wetting or nonwetting of solids by liquids, the criterion employed is the contact angle between the solid and the liquid (measured through the liquid) (**Fig. 1**). A liquid is said to wet a solid if the contact angle θ lies between 0 and 90°, and not to wet the solid if the contact angle lies between 90 and 180°. Three interfaces exist when a droplet of liquid contacts a solid, and three corresponding interfacial tensions exist: γ_{SL}, γ_{SV}, and γ_{LV}. The subscripts S, L, and V refer to solid, liquid, and vapor. At equilibrium, a balance of interfacial tensions exists at the line of common contact, which intersects the figures at point 0. For the case of a liquid which wets the solid ($\theta < 90°$), this equilibrium is expressed by Eq. (1).

$$\gamma_{SV} = \gamma_{SL} + \gamma_{LV} \cos\theta \tag{1}$$

Capillarity. Liquids which wet the walls of fine capillary tubes rise to a height which depends upon the tube radius, the surface tension, the liquid density, and the contact angle. In **Fig. 2**, a liquid of density ρ is shown as having risen to a height H in a capillary whose radius is R. A balance exists between the force exerted by gravity on the mass of liquid raised in the capillary and the opposing force caused by surface tension. The former is $\pi r^2 h\rho g$, whereas the latter is $2\pi r\gamma$, assuming the contact angle to be zero.

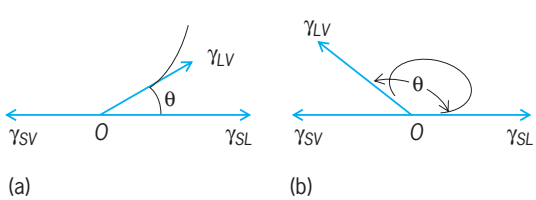

Fig. 1. Contact angle. (*a*) Liquid wets solid. (*b*) Liquid does not wet solid.

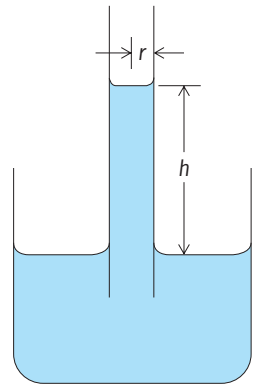

Fig. 2. Rise of liquid in capillary tube.

It is clear that $h = 2\gamma/r\rho g$, and that the capillary rise varies inversely with the tube radius and the liquid density. Liquids which do not wet the capillary walls are depressed in height according to the same equation.

The shape of the free surface of a liquid in a vessel is only an approximation to a plane. In narrow tubes the meniscus of a liquid is concave upward if the liquid wets the tube and, conversely, convex upward if it does not wet the tube. A pressure difference exists between the concave and convex sides of the surface, the excess pressure on the concave side over the convex side being given by Eq. (2), where r_1

$$p = \gamma \frac{1}{r_1} + \frac{1}{r_2} \tag{2}$$

and r_2 are the principal radii of curvature of the surface. The same equation applies for a bubble of gas within a liquid, with the consequence that the vapor pressure p is larger for small bubbles according to Eq. (3), where p_0 is the vapor pressure over a liquid

$$\ln\left(\frac{p}{p_0}\right) = \frac{2\gamma}{r\rho}\frac{M}{RT} \tag{3}$$

surface of infinite radius, R is the gas constant, and M is the molecular weight. *See* VAPOR PRESSURE.

Detergents, soaps, and flotation agents owe their usefulness to their ability to lower the surface tension of water, thereby stabilizing the formation of small bubbles of air. At the same time, the interfacial tension between solid particles and the liquid phase is lowered, so that the particles are more readily wetted and floated after attachment to air bubbles. *See* FLOTATION; INTERFACE OF PHASES; SURFACTANT.

Norman H. Nachtrieb

Bibliography. R. A. Alberty and R. J. Silbey, *Physical Chemistry*, 3d ed., 2000; K. L. Mittal, *Contact Angle, Wettability and Adhesion: In Honor of Professor Robert J. Good*, 1993; J. H. Noggle, *Physical Chemistry*, 3d ed., 1987.

Surface waves

Disturbances that are propagated at a gas-liquid interface and are dependent primarily upon the gravitational property of the liquid (surface tension and

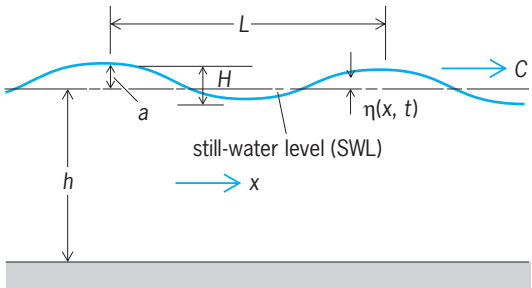

Fig. 1. Definition sketch for an oscillatory wave.

viscosity being of secondary importance). Wave motions that occur in confined fluids (either liquid or gaseous) are dependent primarily upon the elastic property of the medium. *See* WATER HAMMER; WAVE MOTION IN FLUIDS.

The fundamental concepts of gravity waves in liquids are presented in the one-dimensional form. For information on the generation of waves by wind in natural bodies of water and their transformation in coastal regions. *See* OCEAN WAVES.

Oscillatory waves. The term oscillatory implies a periodicity in the form of a disturbance moving past a fixed point. **Figure 1** is a definition sketch for an oscillatory wave propagating in a liquid of constant density ρ and depth h measured from the bottom to the still-water level (SWL). Wavelength L is the horizontal distance between successive crests of the wave. Wave height H is the vertical distance from crest to trough; amplitude a is the distance from the still-water level to the crest, and η is the elevation of the free surface with respect to the still-water level at any position x and instant of time t. In the linearized theory of small-amplitude waves ($H/L < 0.03$) the wave profile is sinusoidal and is givenby Eq. (1),

$$\eta = a \sin\left[2\pi\left(\frac{t}{T} - \frac{x}{L}\right)\right] \tag{1}$$

where T is the wave period. By definition, the celerity or speed of propagation $C = L/T$ and is given by Eq. (2) (in SI units or any other coherent system of

$$C = \sqrt{\left(\frac{\sigma}{\rho}\frac{2\pi}{L} + \frac{gL}{2\pi}\right)\tanh\left(2\pi\frac{h}{L}\right)} \tag{2}$$

units). The first term on the right expresses the influence of surface tension σ and need be considered only for waves of very small length (of the order of magnitude of 1 in. or 2.5 cm). In the remaining development, only gravity waves will be considered, as expressed by the second term of the celerity equation, where g is the acceleration of gravity (9.8 m/s^2 or 32.2 ft/s^2). *See* UNITS OF MEASUREMENT.

Oscillatory waves may be generated in a rectangular channel by a simple harmonic translation of a vertical wall forming one end of the flume. The wave amplitude will be determined by the displacement (stroke) of the wall, and the wavelength will be a function of the period of oscillation. The two major classes of oscillatory waves, deep-water and shallow-water waves, are determined by the magnitude of the ratio of liquid depth to wavelength h/L. An inspection of the celerity equation for gravity waves shows that as the depth becomes large in comparison with wavelength, Eqs. (3) hold. Hence, in a deep-water

$$\tanh\left(2\pi\frac{h}{L}\right) \to 1 \qquad C = \sqrt{\frac{gL}{2\pi}} \tag{3}$$

wave the celerity is a function only of the wavelength. This approximation is close if the depth is greater than one-half the wavelength. On the other hand, as the depth becomes small in comparison with wavelength, Eqs. (4) hold. Hence, the celerity

$$\tanh\left(2\pi\frac{h}{L}\right) \to 2\pi\frac{h}{L} \qquad C = \sqrt{gh} \tag{4}$$

depends only on the depth in a shallow-water wave. Somewhat arbitrarily, the limit $hL = 1/10$ is generally applied to this type of wave motion. In deep-water waves, individual fluid particles tend to move in circular orbits. The radius of the surface particle orbit is equal to the wave amplitude and the radius decreases exponentially with depth (**Fig. 2**). At a depth of one-half the wavelength, the orbital radius is about 1/20 of the amplitude. A zone of essentially zero fluid motion is rapidly approached and the character of the wave is therefore not affected by the total depth of the liquid.

In shallow water, no vertical particle motion can exist at the bottom; thus the wave characteristics are modified. The particle orbits are flat ellipses in which the minor axis is depressed to zero at the bottom (Fig. 2).

The energy of a wave consists of equal amounts of potential energy (due to particle position above or below the still water level) and kinetic energy (due to the motion of particles in their orbits). The rate of propagation of energy in the direction of wave travel is known as the group velocity to distinguish it from phase velocity C. In deep-water waves the group velocity is one-half the phase velocity; in shallow-water waves the two propagation velocities are equal.

Standing waves. A standing wave can be considered to be composed of two equal oscillatory wave trains traveling in opposite directions. The phase velocity of the resulting wave is zero; nevertheless the velocity of propagation of the component waves retains its usual meaning. In the notation of the previous section, the equation for the profile of a standing wave is obtained by adding the elevations of waves moving in the positive and negative x directions, given by Eqs. (5) and (6), respectively. Hence

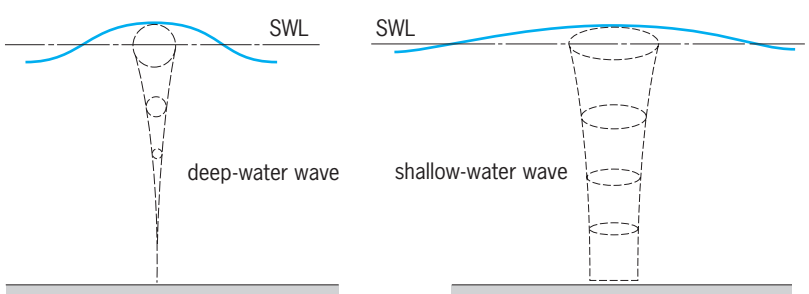

Fig. 2. Fluid particle orbital motions in deep- and shallow-water oscillatory waves. SWL = still-water level.

Eq. (7) holds. If the length of the basin l in which a dis-

$$\vec{\eta}_1 = a \sin\left[2\pi\left(\frac{t}{T} - \frac{x}{L}\right)\right] \tag{5}$$

$$\overleftarrow{\eta}_2 = a \sin\left[2\pi\left(\frac{t}{T} + \frac{x}{L}\right)\right] \tag{6}$$

$$\eta = \eta_1 + \eta_2 = H \sin\left(2\pi\frac{t}{T}\right)\cos\left(2\pi\frac{x}{L}\right) \tag{7}$$

turbance occurs is an integral number n of half wavelengths, a self-perpetuating (except for frictional dissipation) standing wave will result. Therefore, if $l = nL/2$, Eq. (8) is valid. For long waves, as with

$$\eta = H \sin\left(2\pi\frac{t}{T}\right)\cos\left(\frac{\pi n x}{l}\right) \tag{8}$$

shallow water, in a basin of uniform depth, the period of oscillation T is defined by Eq. (9). Standing waves

$$T = \frac{2l}{n\sqrt{gh}} \tag{9}$$

frequently occur in canal locks as a result of filling disturbances and in large lakes, bays, and estuaries as a result of wind or tidal action. *See* SEICHE.

Solitary waves. A solitary wave consists of a single crest above the original liquid surface which is neither preceded nor followed by another elevation or depression of the surface. Such a wave is generated by the translation of a vertical wall starting from an initial position at rest and coming to rest again some distance downstream. In practice, solitary waves are generated by a motion of barges in narrow waterways or by a sudden change in the rate of inflow into a river; they are therefore related to a form of flood wave. The amplitude of the wave is not necessarily small compared to the depth, and the wavelength is theoretically infinite because the elevation of the surface approaches the still water level asymptotically with distance as shown in **Fig. 3**. The profile of the solitary wave is given by Eq. (10) and the celerity by Eq. (11). When the solitary wave amplitude becomes

$$\eta = a\,\mathrm{sech}^2\left[\frac{x}{h}\sqrt{\frac{3}{4}\frac{a}{h}}\right] \tag{10}$$

$$C = \sqrt{g(h+a)} \tag{11}$$

approximately equal to the depth, the wave profile becomes unstable and a breaking wave results.

Surges. A surge is generated by the forward motion of a vertical wall, at a constant speed, as shown

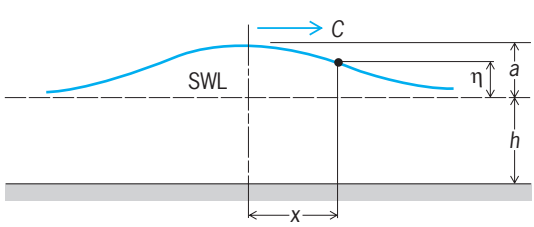

Fig. 3. Definition sketch for a solitary wave. SWL = still-water level.

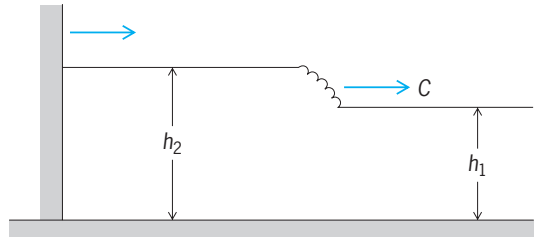

Fig. 4. Definition sketch for a surge wave.

schematically in **Fig. 4**. Surges in open channels are analogous to shock waves produced in a tube by the continuous motion of a piston. A zone of violent eddy motion occurs at the wavefront and the analysis of such motions must take into account the appreciable energy dissipation in this region. The velocity of propagation of a surge is given by Eq. (12). If a

$$C = \sqrt{gh_1}\left[\frac{1}{2}\frac{h_2}{h_1}\left(\frac{h_2}{h_1} + 1\right)\right]^{1/2} \tag{12}$$

velocity V_1 equal and opposite to C is imposed on the fluid upstream of the disturbance, the absolute velocity of the surge front will become zero. In this form the surge is known as a hydraulic jump and it is often used as a means of dissipating flow energy at the bottom of dam spillways. *See* HYDRAULIC JUMP.

Donald R. F. Harleman

Bibliography. R. G. Dean and R. A. Dalrymple, *Water Wave Mechanics for Engineers and Scientists*, 1991; Open University Course Team Staff (ed.), *Waves, Tides and Shallow-Water Processes*, 2d ed., 1999; J. J. Stoker, *Water Waves*, 1958, reprint 1992; A. Torum and O. T. Gudmestad (eds.), *Water Wave Kinematics*, 1990.

Surfactant

An amphiphilic (also called amphipathic) compound that adsorbs at interfaces to form oriented monolayers and shows surface activity. An amphiphilic compound is a molecule that has a hydrophilic (polar) head and a hydrophobic (nonpolar) tail. Common synonyms for the term "surfactant" include amphiphile, surface-active agent, and tenside. If a surfactant is placed into contact with both a polar medium, such as water, and a nonpolar medium, such as an oil, one part of its molecule has an affinity for the polar medium and one part that has an affinity for the nonpolar medium. An example is an organic molecule such as sodium dodecyl sulfate (SDS), which can be thought of as having a hydrocarbon (dodecyl) tail and a highly polar (sulfate) head group. If the molecule is placed into a system containing water and oil, the sulfate head group will have an affinity for the water, while the dodecyl tail will have an affinity for the oil. The energetically most favorable orientation for such molecules is at the interface between the polar and nonpolar media, so that each part of the molecule can reside in an environment for which it has the greatest affinity. In the case of SDS in a mixture of oil and water, the SDS molecules will preferentially adsorb at the water/oil interface, with the polar sulfate

group oriented into the water and the dodecyl tail group oriented into the oil. Three consequences of the amphiphilic nature of surfactants are their ability (1) to adsorb and form layers at interfaces, (2) to reduce the interfacial tension between fluids, and (3) to associate to form clusters, called micelles. *See* ADSORPTION; INTERFACE OF PHASES; MICELLE; SURFACE TENSION.

When surfactant molecules adsorb at an interface, they provide an expanding force that acts against the normal interfacial tension. At the interface between water and air, for example, there exists a natural interfacial tension that arises from chemical differences between the water and the air. The interfacial tension acts to cause the water/air interface to always strive toward a condition that minimizes the area of contact (interfacial area) between the two phases. This interfacial tension is what causes a waterdroplet to strive to achieve the shape of a sphere (the shape having the smallest interfacial area). When a small amount of a surfactant is added to the water/air system, surfactant molecules adsorb at the interface and reduce the interfacial tension. This makes it easier for the system to adopt shapes having a higher interfacial area. A waterdroplet in air that contains some surfactant can more easily distort and adopt shapes other than spherical. This makes it easier to form foams. The same thing happens when a small amount of surfactant is added to a mixture of water and oil. The reduced interfacial tension makes it easier for the system to adopt shapes having a higher interfacial area. This makes it easier to form emulsions either of oil-in-water or of water-in-oil. *See* EMULSION; FOAM.

Many hydrocarbon surfactants can lower the interfacial tension of air–water at 20°C (68°F) from 72.8 mN/m to about 28 mN/m. Polysiloxane surfactants can reduce it to about 20 mN/m, and perfluoroalkyl surfactants can reduce it still further to about 15 mN/m. Similarly, hydrocarbon surfactants can reduce the interfacial tension of water–mineral oil from about 40 mN/m to about 3 mN/m. In addition to lowering interfacial tension, adsorbed surfactant at interfaces can strongly influence other properties, such as interfacial elasticity and interfacial viscosity.

Lung surfactants, including phospholipids and proteins, are an example of natural surfactants present in the human body. The lung surfactants are necessary to maintain a low interfacial tension at the alveolar air–water interface and to help the alveolar spaces change size during the breathing cycle. These interfacial tensions also change rapidly in response to changes in alveolar radius, and can fall to values as low as 10 mN/m or less. A deficiency of lung surfactant causes the interfacial tension to be too high, which causes alveolar collapse at the end of expiration. The lung surfactants may also play a role in protecting the lung from injury and infection caused by inhalation of particles and microorganisms. *See* PHOSPHOLIPID; RESPIRATORY SYSTEM.

As surfactant is added to a multiphase system, the surfactant molecules will tend to adsorb at the interface(s), forming into an oriented monolayer at the interface and reducing interfacial tension. However, above a certain concentration, called the critical micelle concentration (cmc), the surfactant molecules will start to self-associate and form aggregates called micelles. The value of the cmc depends primarily on the nature of the surfactant. Other factors being equal, higher molar masses produce lower cmc values. Once micelles have been formed, almost all additional surfactant added to the system will become solubilized in the micelles. In aqueous solution the part of the surfactant molecules that have an affinity for the nonpolar medium (the tails) associates in the interior of the micelle, while that part that has an affinity for the water (the head groups) faces the aqueous medium. Micelles typically contain 20–100 surfactant molecules each. The insides of the micelles, being filled mostly with nonpolar tails, have properties similar to an oil phase.

Surfactant types. Depending on the nature of the polar (hydrophilic) part of the molecule, surfactants are classified as anionic (negatively charged), cationic (positively charged), nonionic (noncharged), or zwitterionic (able to contain either or both kinds of charge). A common anionic surfactant is sodium dodecyl sulfate, $CH_3(CH_2)_{11}SO_4^-Na^+$. A common cationic surfactant is cetyl trimethylammonium bromide, $CH_3(CH_2)_{15}N^+(CH_3)_3Br^-$, a common type of nonionic surfactant is polyoxyethylene alcohol, $C_nH_{2n+1}(OCH_2CH_2)_mOH$. An example of a zwitterionic surfactant is dodecyl betaine, $C_{12}H_{25}N^+(CH_3)_2CH_2COO^-$. A soap is a particular kind of anionic surfactant comprising any of the surface-active fatty-acid salts containing at least eight carbon atoms. Typically, the molar masses of surfactants range from a few hundreds to several thousands of grams per mole.

Although surfactants are not generally considered to be a serious threat to humans, they can be toxic to aquatic organisms. This can be a concern due to the large quantity of surfactants used. It has been estimated that the global use of synthetic surfactants is on the order of several million metric tons per year.

Applications. Of the numerous surfactant applications, the most common are emulsifying, foaming, suspending and floating, wetting, and detergency. Emulsion, foam, and suspension products span a diverse range of industries, including agriculture, food processing, cosmetics and personal care, pharmaceuticals, papermaking and deinking, water and wastewater treatment, environmental remediation, oil recovery, and minerals beneficiation.

Emulsifying and demulsifying. An emulsion is a system of small liquid droplets dispersed in a second, immiscible liquid. Usually one of the liquids is aqueous (that is, water containing dissolved substances), while the other is an oil, usually a hydrocarbon liquid. Surfactants are the most common agents used to stabilize emulsions. The surfactants adsorb at the interface between the two liquids and act to lower interfacial tension, which makes an emulsion easier to form. Surfactants may also act to increase surface elasticity, increase electric double-layer repulsion (in the case of ionic surfactants), and increase surface viscosity, among other properties. Surface elasticity can be thought of as elasticity operating only at a surface, in two dimensions. Increased elasticity

permits a surface to better withstand deformations without rupturing. The electric double layer includes the charged surface and the oppositely charged ions that are distributed in solution but close to the surface. Increased surface electric charge causes an increased repulsive force between emulsion droplets. In addition to influencing stability, some surfactants can have an influence on the sizes of the dispersed droplets and on the overall viscosity of the emulsion. *See* EMULSION.

Although surfactants have one part of their molecule that has an affinity for the water and one part that has an affinity for the oil, as mentioned above, most surfactants have a greater overall affinity for one or the other. Surfactants that have an overall greater affinity for water (hydrophilic surfactants) tend to stabilize emulsions of oil droplets dispersed in water (oil-in-water, or O/W). Surfactants that have an overall greater affinity for oil (oleophilic surfactants) tend to stabilize emulsions of water-droplets dispersed in oil (water-in-oil, or W/O). Skincare creams are usually formulated as O/W emulsions to promote easy spreading and adsorption while allowing a single product to contain both water- and oil-soluble components. Both water- and oil-soluble surfactants are used to stabilize the emulsions.

Surfactants can also be used to break (demulsify) already-formed emulsions, which are undesirable. For example, several kinds of emulsions are produced at wellheads during oil production operations. These emulsions have to be broken and the water and solids separated out before the crude oil can be sent to a refinery. Specialized surfactant-containing demulsifier formulations are almost always needed in the treatment of these emulsions. When crude oil is spilled at sea, a slick is formed that can turn into a particularly tenacious W/O emulsion, called a mousse emulsion. Surfactants used to treat oil spills tend to be aimed at breaking these mousse emulsions and/or at causing O/W emulsions to be formed. The O/W emulsions are desirable because they help disperse oil into the water column and away from sensitive shorelines and birds.

Foaming. A foam is a system of small gas bubbles dispersed in a liquid. Usually the liquid is aqueous while the gas is just air. Surfactants are the most common agents used to stabilize foams. The surfactants adsorb at the interface between the gas and the liquid and act to lower interfacial tension, which makes a foam easier to form. Effective foaming agents also act to increase surface elasticity, increase electric double-layer repulsion (in the case of ionic surfactants), and increase surface viscosity, among other properties. Just as for emulsions, in addition to influencing stability, the nature of the surfactant can have an influence on the sizes of the dispersed gas bubbles and on the overall viscosity of the foam. In general, ionic surfactants tend to be the best foaming agents, as are those with long, unbranched alkyl tails. *See* FOAM.

There are many examples of foams. Several kinds of foams have even been developed for use as temporary blankets. These foam blankets have been used in fire fighting (to shut off oxygen flow to the fire), over sanitary landfill sites (to suppress emissions and odors), at mine sites (to suppress dusts), at blasting sites (to suppress blast noise and pressure waves), and on petroleum products (to suppress flammable vapors), among others.

Suspending and floating. A suspension is a system of small, solid particles dispersed in a liquid. The liquid is usually aqueous. When added to a suspension, surfactants will adsorb at the interface between the solid and the liquids. For solids dispersed in water, ionic surfactants will usually stabilize the suspension by creating or increasing electric double-layer repulsion between particles. *See* SUSPENSION.

A complex example of a food suspension is ice cream, a partially frozen mixture containing many components, including proteins, fat, water, and air. The particles are ice crystals and corn-syrup solids. When the mix is whipped and partially frozen, a complex internal structure is achieved that has fat globules both adsorbed onto air bubbles and aggregated to each other, in addition to the ice crystals. The surfactants used to stabilize ice cream include casein and other proteins. In addition to being a suspension, ice cream is a foam and an emulsion.

Adsorbed surfactant may also change the wetting preference of the solid particles. If adsorbed surfactant causes finely divided particles to strongly prefer to be wetted by (in contact with) oil than water, then such particles will tend to float on water. This principle is used in the treatment of mineral ores by flotation, in which a small amount of surfactant, added during the grinding and slurrying process, alters the wettability of the ore particles so that they become hydrophobic and attach to air bubbles and float to the surface, where they are recovered by skimming. Flotation processes are used to concentrate a wide range of metallic and nonmetallic minerals, including copper, lead, zinc, nickel, calcium fluoride, barium sulfate, potash, sulfur, coal, phosphates, alumina, silicates, and clays. Flotation is also used in the removal of bacteria and particulates from water and wastewater.

Wetting. Another application in wetting is when adsorbed surfactant is used to make cloth fibers hydrophobic, which in turn makes a fabric water-repellent. Long-chain cationic surfactants are commonly used for this purpose, since the fiber surfaces are usually anionic. With the cationic heads adsorbed to the fiber surfaces, the hydrocarbon tails orient away from the surface, toward the water. As long as the weave is reasonably tight so the pore sizes between fibers are fairly small, water will enter only between the fibers under high pressure. The result is a fabric that is water-repellent but which allows air to pass through: the fabric "breathes."

In contrast, if the orientation is with the polar head away from a surface, a surfactant will make the surface hydrophilic—preferentially wetted by water. Some common wetting agents include barely water-soluble anionic hydrocarbon surfactants and some siloxane surfactants. Applications in which surfactants are used to promote water-wetting of otherwise waxlike surfaces include (1) wetting of fabrics to ensure even distribution of a textile dyes, (2) dips for treating animal hides, (3) adjuvants,

ensuring rapid wetting and spreading of insecticides and herbicides, and (4) ensuring uniform coverage when applying paints and coatings.

Detergency. Detergency involves the action of surfactants to adsorb at interfaces, alter wetting and/or interfacial tensions, and thereby reduce the energy needed to cause the removal of dirt from solid surfaces. Soaps were the first detergents, but are sensitive to acidic conditions and to the presence of hardness in the water (calcium and magnesium ions), causing soap scum. Modern detergents, such as alkyl sulfates and alkyl-aryl sulfonates, are much more effective and have almost entirely replaced soaps. These detergents comprise synthetic surfactants formulated with a number of other additives. An effective detergent formulation has to be able to induce water-wetting of the dirt (solids and/or oils) and encourage the dirt to be displaced from the surfaces to be cleaned into the wash water. Ideally, a detergent will also solubilize or otherwise prevent redeposition of the dirt. Surfactants cannot do all these things by themselves in practical situations, so commercial detergent formulations also contain additives such as builders (to enhance the detergent action), anti-redeposition agents, brighteners, and cosurfactants. *See* DETERGENT; SOAP.

Familiar applications of detergency include dish washing, clothes washing, shampooing, bathing, and hand washing. It is sometimes desirable for detergent formulations to contain surfactants that can act as germicides as well. Some cationic surfactants, like quaternary ammonium surfactants, are toxic to bacteria, fungi, and algae. An example is *N*-alkyldimethylbenzylammonium chloride. Detergency is also used in industry, for example, to clean hard surfaces like glass and metal, and soft surfaces like textiles.

Detergency can also form the basis for an enhanced oil recovery (EOR) process. In oil-containing reservoirs, the oil is distributed among pores in the rock. After conventional processes, such as waterflooding, there remain in the rock considerable amounts of oil trapped by interfacial forces. Surfactant-based EOR processes are intended to overcome these forces by altering the wetting and/or interfacial tension properties so that the oil is released from the pores in the form of emulsion droplets that can be driven toward a producing well. In the same way, surfactants can be used to flush nonaqueous phase liquid contaminants, such as gasoline or other fuels, out of underground aquifers as part of an environmental soil remediation process. *See* PETROLEUM ENHANCED RECOVERY. Laurier L. Schram

Bibliography. D. R. Karsa (ed.), *Industrial Applications of Surfactants*, Royal Society of Chemistry, London, 1987; D. Myers, *Surfactant Science and Technology*, VCH, 1988; M. J. Rosen, *Surfactants and Interfacial Phenomena*, 2d ed., Wiley, 1989; M. J. Rosen and M. Dahanayake, *Industrial Utilization of Surfactants*, AOCS Press, 2000; L. L. Schramm (ed.), *Surfactants: Fundamentals and Applications in the Petroleum Industry*, Cambridge University Press, 2000.

Surge arrester

A protective device which is connected between an electrical conductor and the ground with the intent of limiting the magnitude of transient overvoltages on equipment.

The valve arrester consists of disks of zinc oxide material that exhibit low resistance at high voltage and high resistance at low voltage. By selecting an appropriate configuration of disk material, the arrester will conduct a low current of a few milliamperes at normal system voltage. When insulation levels of equipment are coordinated with the surge arrester protective characteristics, and under lightning or switching surge overvoltage conditions, the surge current is limited by the electric circuit; and for the magnitudes of current that can be delivered to the arrester location, the resulting voltage will be limited to controlled values, and to safe levels as well.

A typical surge arrester consists of disks of zinc oxide material sized in cross-sectional area to provide desired energy discharge capability, and in axial length proportional to the voltage capability. The disks are then placed in porcelain enclosures to provide physical support and heat removal, and sealed for isolation from contamination in the electrical environment (see **illus.**).

System overvoltages may be of either external or internal origin, that is, lightning or switching. The arrester is not intended to determine the origin of the overvoltage and must attempt to limit the magnitude of all abnormal voltages above a specific level. Hence a lightning arrester is really a voltage-surge arrester. Important characteristics of arresters are steady-state voltage capability, discharge voltage

Surge arrester unit rated 40 kV, consisting of porcelain enclosure and multiple zinc oxide disks in series. (*General Electric Co.*)

across the arrester during passage of current, and maximum current discharge capability.

The discharge voltage characteristic determines the maximum transient voltage level permitted by the arrester. It is the measure by which the protective efficiency of arresters is determined. The discharge capability is the measure of the arrester endurance to severe lightning and switching surges.

Arresters are classified as station, intermediate, or distribution-type. Station-class arresters are used for the protection of apparatus in important substations where the highest level of protection is desired. Intermediate-class arresters are used for the protection of apparatus in small- and medium-size substations where economic considerations do not justify the use of station-type arresters. Distribution-class arresters are used to protect of pole-type distribution apparatus, principally distribution transformers. *See* LIGHTNING; LIGHTNING AND SURGE PROTECTION.

Glenn D. Breuer

Surge suppressor

A device that is designed to offer protection against voltage surges on the power line that supplies electrical energy to the sensitive components in electronic devices and systems. The device offers a limited type of protection to computers, television sets, high-fidelity equipment, and similar types of electronic systems.

A voltage surge is generally considered to be a transient wave of voltage on the power line. The amplitude of the surge may be several thousand volts, and the duration may be as short as 1 or 2 milliseconds or as long as about 100 ms. Typical effects can be damage to the electronics or loss of programs and data in computer memories. Many events can cause the surges, including lightning that strikes the power lines at a considerable distance from the home or office; necessary switching of transmission lines by the utilities; and rapid connections or disconnections of large loads, such as air conditioners and motors, from the power line, or even other appliances in the home. Lightning is perhaps the most common. *See* LIGHTNING.

Specifications. The suppressor acts to limit the peak voltage applied to the electronic device to a level that normally will not cause either damage to the device or software problems in the computers. The normal operating voltage and frequency of the power line in which the suppressor is to be used, the current rating, and the types of loads that are to be connected are usually specified, as are the types of protection that the suppressor will provide. Typical specifications include the maximum voltage that will be applied to the electronic system, the response time required (indicating how quickly the device responds), the maximum surge voltage against which the device offers protection, and the amount of energy that can be safely dissipated. For example, in countries where systems with an alternating-current voltage of 120 V and a frequency of 60 Hz are common, a typical set of specifications for a surge suppressor designed for the home or small office might include a peak clamping voltage of 250 V, a response time of 5 nanoseconds, a maximum transient protection level of 6000 V, and an energy dissipation of 40 joules. In countries having systems with an alternating-current voltage of 220 V, the clamping level would be higher, perhaps 500 V.

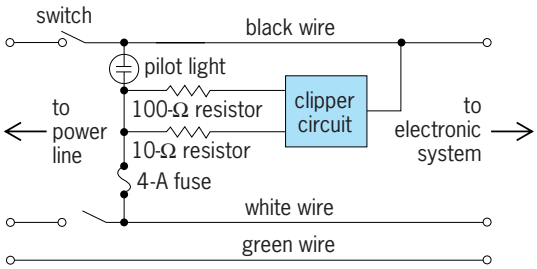

Typical circuit diagram for a surge suppressor; the colors are typically found on wires in North America.

Operation. The device may include a pilot light, a fuse, a clipping circuit, resistors, and a main switch (see **illus.**). The main switch serves to energize or deenergize the entire system. The pilot light is illuminated when the suppressor is operating correctly. The clipper circuit is the principal item, and the design of this portion is usually proprietary information. If a high-energy transient is absorbed by the clipper circuit, then it is likely that the fuse will be blown. Then the pilot light goes out, but the suppressor still supplies power to the electronic system. It does not, however, provide additional protection until the fuse is replaced. The unilluminated light is a signal to the operator that the suppressor has functioned as intended and that it must be checked. The minimum servicing is replacing the fuse. *See* FUSE (ELECTRICITY).

Limitations. A basic surge suppressor will not offer protection against massive, nearby lightning strokes. Lightning or surge arresters may provide such protection. It will not offer protection against sags or drops in line voltage, such as drops that cause lights to dim when, for example, an air conditioner switches on. Voltage regulators and constant voltage transformers serve this purpose. It will not reduce high-frequency electrical noise that may be on the power line. Filters are available to meet this need. *See* ELECTRIC FILTER; ELECTRONIC POWER SUPPLY; LIGHTNING AND SURGE PROTECTION; SURGE ARRESTER; TRANSFORMER; VOLTAGE REGULATOR.

Edwin C. Jones, Jr.

Bibliography. Institute of Electrical and Electronics Engineers, *Guidelines for Surge Voltages and Low-Voltage AC Power Circuits*, IEEE C62. 41, 1991; G. Mager, Saving your computer from surges, sags, or noise, *DEC Professional*, 3(7):26–38, December 1984; H. O. Nash, Jr., and F. M. Wells, Power systems disturbances and considerations for power conditioning, *IEEE Trans. Ind. Appl.*, IA-21(6):1472–1481, 1985; Underwriter's Laboratories, *Underwriters Standard 1449 for Surge Suppression*, rev. ed., 1995.

Surgery

General surgery includes broad operative expertise and experience with diseases of the gastrointestinal tract and abdomen; benign and malignant diseases of the breast; disorders of the endocrine system, including the thyroid, parathyroid, and adrenal glands; and vascular disease; as well as the comprehensive management of trauma. *See* BREAST DISORDERS.

Gastrointestinal surgery. This specialty encompasses the treatment of diseases of the esophagus, stomach, small intestine, large intestine, rectum, anus, pancreas, liver, and biliary tree. Traditionally, the organs of the abdomen (intraabdominal viscera) are exposed by making a major incision through the abdominal musculature to enter the peritoneal cavity, the direction and extent of the incision being individualized according to the condition being addressed. With the advent of video cameras, fiber-optic technology, and specialized instrumentation, some intraabdominal procedures are being performed through much smaller, less traumatic, and less uncomfortable incisions, resulting in lessened disability and minimal scarring. Videoscopic approaches have been successfully applied to cholecystectomy, surgery for peptic ulcer disease, surgery for gastroesophageal reflux, correction of inguinal hernia, colectomy for benign disease, and splenectomy for hematologic disease. *See* OPTICAL FIBERS.

In order to be able to visualize intraperitoneal structures with the video camera, carbon dioxide is pumped into the peritoneal cavity continuously to maintain constant distention. A series of small incisions are then made to introduce a laparoscope and video camera along with instrumentation such as retractors and dissectors. Coagulation of blood vessels by cauterization, that is, the use of a heated instrument, or laser is used regularly. Suturing also occurs through the laparoscope. The entire procedure is viewed by the surgeon on high-resolution television monitors which provide outstanding detail of intraabdominal structures. Videoscopic surgery is one of the most significant advances in general surgery since the mid-1980s (see **illus.**).

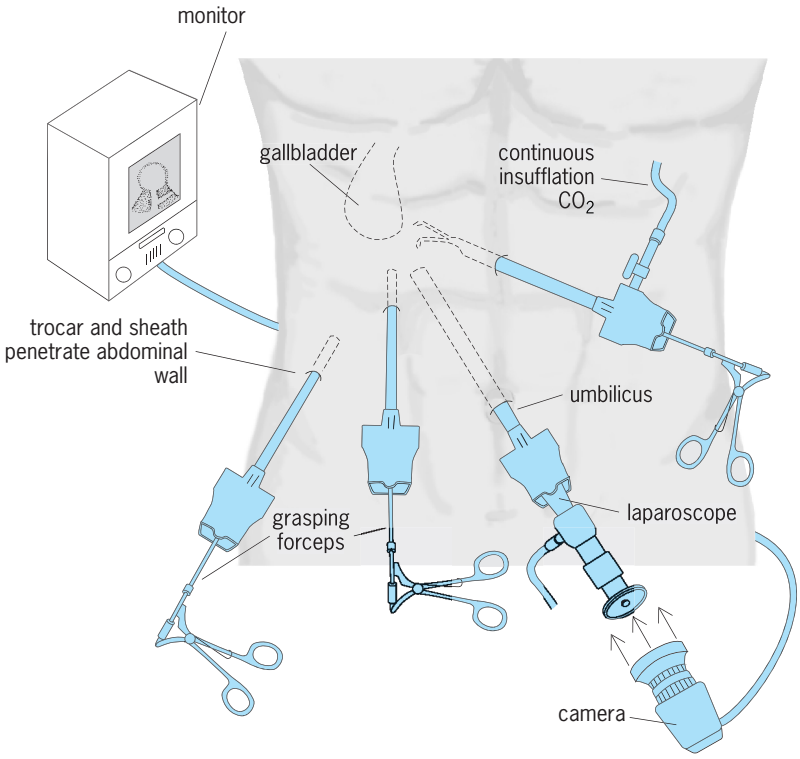

Position and function of laparoscopic instruments during laparoscopic cholecystectomy. (*After K. A. Zucker, ed., Surgical Laparoscopy, Quality Medical Publishing, 1991*)

Endoscopic procedures on the esophagus, stomach, duodenum, bile duct, and colon depend on fiber-optic technology to carry images from one end of a flexible telescope to the other, with a lens system at each end to enhance imaging. The ends of these endoscopes are flexible with control being maintained by wires in the wall of the instrument, permitting precise manipulation. Images are now usually viewed on digital television monitors; this is possible because a tiny television camera replaces the fiber-optic visualization system. A wide variety of instruments are available for passage through the endoscope to permit biopsy, excision, cauterization of bleeding points, and other major manipulations. *See* DIGESTIVE SYSTEM; MEDICAL IMAGING.

Vascular surgery. This specialty includes the treatment of diseases of the arteries, veins, and lymphatic vessels, exclusive of the heart, thoracic aorta (cardiothoracic surgery), and intracranial vessels (neurosurgery). A wide variety of blockages, aneurysms, and degenerative conditions of the vascular system are treated with a variety of techniques. Angiography is a frequently used procedure for planning an operative approach. Through direct puncture of an artery located in an extremity, catheters are passed centrally into the artery to be studied. This is followed by injection of contrast medium in order to achieve excellent visualization of the arterial system. While the procedure is performed by a radiologist, the vascular surgeon interprets and applies the information in order to develop a therapeutic plan. The vascular surgeon also uses a variety of noninvasive diagnostic procedures, including ultrasonic imaging of peripheral, carotid, and abdominal arteries and veins, as well as blood flow measurements with peripheral pressure determinations. With the improvement of noninvasive vascular techniques for the diagnosis and surveillance of vascular disease and their treatment, diagnoses are being made sooner, results of treatment are being followed systematically, failures or disease progression are being identified promptly, resulting in more timely intervention as indicated. *See* RADIOGRAPHY; ULTRASONICS.

Occlusion or blockage of the aorta or of the arteries to the legs is most often due to atherosclerosis, resulting in diminished flow with pain due to poor circulation (initially with exercise and then at rest) and possibly gangrene. Similar lesions occur in the carotid arteries, causing stroke; in the renal arteries, causing hypertension or renal failure; or in the intestinal arteries, causing chronic abdominal pain, weight loss, and gangrene. The vascular surgeon uses a variety of techniques to deal with these problems. Obstructions to the larger vessels in

the abdomen or upper leg are bypassed by using a tube graft made of synthetic material such as dacron or poly(tetrafluoroethylene). For bypass into the lower leg or to the foot, saphenous vein is preferred because of a higher long-term patency rate. For occlusive disease of the carotid arteries, endarterectomy is the preferred technique. This consists of actual removal of the atherosclerotic plaque from the opened artery, followed by restoration of blood flow to the brain. *See* ARTERIOSCLEROSIS.

Transplant surgery. Transplant surgery involves the surgical placement of an organ taken from a donor into a recipient with end stage failure of that organ. Donors are most often individuals who have been declared brain-dead; that is, there is irreversible brain damage incompatible with life. Following the consent of the donor or the family, normal organs can be removed for transplantation into a suitable recipient (cadaveric transplant). Less frequently, a healthy individual, often a close relative, serves as a donor (living related transplant). Diseases treated in this fashion include renal failure, liver failure, and insulin-dependent diabetes mellitus. Cardiac and pulmonary failure are also successfully treated by transplantation in selected cases.

Previously, early transplantation surgery was severely hindered by the recipient's aggressive immune response to the foreign organ which led to rejection of the organ and ultimate failure. The advent of a variety of effective and increasingly better-tolerated immunosuppressant medications has led to a blossoming of transplant surgery with generally excellent results in function and longevity of the transplanted organ. These agents include prednisone, azathioprine antithymocyte globulin, and especially cyclosporine. Unfortunately, these medications also impair resistance to infection, an ever-present potential problem for transplant recipients. The principal obstacle to widespread transplantation is a shortage of donors. *See* TRANSPLANTATION BIOLOGY.

Surgical oncology. This discipline deals with the treatment of the various forms of cancer amenable to surgical therapy. It encompasses all forms of gastrointestinal cancer, cancer of the breast, skin cancers such as melanoma, and many connective tissue malignancies (sarcomas) of the extremities. Traditionally, surgical oncologists removed large sections of tissues and regional lymph nodes along with the tumor. In addition, the surgical oncologist specializes in the application of multidisciplinary approaches to the management of malignancy, such as radiation therapy, chemotherapy, and nutritional support. In some instances, such as for many individuals with breast cancer, multimodality therapy has resulted in less aggressive, less destructive operations which, coupled with the use of chemotherapy and radiation therapy, offer the same or improved long-term survival benefit. More modern approaches include manipulation of the immune system and gene therapy. *See* CANCER (MEDICINE); CHEMOTHERAPY AND OTHER ANTINEOPLASTIC DRUGS; ONCOLOGY; TUMOR.

Trauma and critical care surgery. This discipline involves the postinjury and postoperative care of critically ill individuals. The goal is to prevent and treat multiorgan system failure and the other complications, as well as to address issues of nutrition, speech and communication, physical therapy, occupational therapy, wound care, and decubitus ulcer prevention, and to prescribe the appropriate rehabilitation plan for the individual.

Cardiothoracic surgery. Cardiothoracic surgery attempts to treat diseases of the heart and thorax, including the lungs, aorta, esophagus, thymus, pleura, diaphragm, and chest wall. The development of the heart-lung machine has allowed cardiac surgery to be a standard therapeutic option for many conditions, especially coronary atherosclerosis and valvular heart disease. This machine allows for cardiac and pulmonary function to be performed through the use of a pump and oxygenator, while the heart is stopped and operated upon. The most commonly performed procedure is coronary artery bypass grafting, in which saphenous veins from the leg or the internal mammary artery are used to bypass obstructed or narrowed coronary arteries, thereby increasing blood flow to the heart. This procedure is required for many despite the widespread use of balloon angioplasty to dilate narrowed segments of coronary arteries.

The application of videoscopic techniques to thoracic surgery using a miniature television camera, that is, the thoracoscope, and specialized instrumentation which can be introduced into the chest through small incisions has allowed surgeons to perform major thoracic procedures. Surgery of this type has been used for staging lung cancer, esophageal resection, limited pulmonary resection, resection of pulmonary cysts, and biopsy. The minimally invasive procedural approach markedly decreases morbidity, hospital stay, and postprocedure disability. *See* HEART DISORDERS.

Plastic surgery. The aim of plastic surgery is to repair, replace, and reconstruct defects of form and function of the skin and its underlying musculoskeletal system, with emphasis on the craniofacial structures, the oropharynx, the breast, the hand, and the external genitalia. It also includes esthetic surgery of structures with undesirable form. Plastic surgeons have extensive experience in the design and transfer of tissue flaps, in the transplantation of tissues, and in the replantation of various structures, along with skill in excisional surgery, the management of complex wounds, and the use of synthetic materials to augment the reconstructive process. The principles of plastic surgery can be applied to a wide variety of problems, such as congenital defects of the head and neck, including clefts of the lip and palate; neoplasms of the head and neck; facial trauma, including fractures; esthetic surgery of the head and neck, trunk, and extremities; and reconstructive surgery of the breast, hand, and extremities.

Pediatric surgery. This field involves surgical problems of the newborn, including premature infants, as well as children of all ages. Care of the pre- and postoperative child employs very different principles than those used for the adult. The pediatric surgeon treats congenital diseases in children, as well as a variety of illnesses and neoplasms that occur

in this population. There is special expertise in the management of congenital and acquired conditions of the gastrointestinal tract and abdominal cavity, the vascular system, the integument, the diaphragm and thorax exclulsive of the heart, the endocrine glands, the gonads and reproductive organs, and the head and neck, as well as in the comprehensive management of trauma.

Urology. Urology manages the treatment of the surgical diseases of the genitourinary tract. Infections, stone disease, organ dysfunction, and neoplasms of kidneys, ureters, bladder, prostate, urethra, and external genitalia all fall within the purview of the urologist. Techniques include the use of endoscopy, extracorporeal shock wave lithotripsy (ultrasonic fragmentation of kidney stones), videoscopic staging of neoplasms, and an emphasis on multidisciplinary management of neoplastic disease.

Orthopedic surgery. The surgical management of diseases of the musculoskeletal system, orthopedic surgery, includes congenital, neoplastic, traumatic, and degenerative diseases of long and short bones, the spine, the hand, the joints, and the associated musculature and tendons. The orthopedic surgeon is familiar with all aspects of adult orthopedics, pediatric orthopedics, fractures and trauma, surgery of the spine including disc surgery, hand and foot surgery, athletic injuries, and rehabilitation. There has been continued improvement in joint replacement surgery through the introduction of prostheses much more resistant to deterioration and significantly improved rehabilitation programs. Endoscopic instrumentation has been used for arthroscopic surgery, whereby miniaturized cameras and instruments are introduced into joints for diagnosis and therapy. Another advancement in orthopedics has been the development of distraction osteogenesis, that is, the formation of new bone to replace that which has been destroyed by infection to lengthen limbs in selected individuals. *See* PROSTHESIS.

Otolaryngology. Otolaryngology specializes in the treatment of diseases of the head and neck, the ears, nose, and vocal cords, the upper alimentary tract, and the upper respiratory tract. It also deals with many communicative disorders, especially those affecting hearing and voice production. The otolaryngologist is experienced in temporal bone surgery; paranasal sinus and nasal septum surgery; maxillofacial, plastic, and reconstructive surgery of the head and neck; surgery of the salivary glands; head and neck oncologic surgery; peroral endoscopy, both diagnostic and therapeutic; surgery of lymphatic tissues of the pharynx; and recognition and medical management of diseases and abnormal function of the ears, upper and lower respiratory tract, and upper alimentary tract. Treatments include videoscopic approaches to skull base tumors and microsurgery of the inner ear. Multimodality management of many head and neck cancers has also resulted in improved survival.

Neurological surgery. This discipline deals with the diagnosis, evaluation, and treatment of disorders of the central, peripheral, and autonomic nervous systems, including their supporting structures and vascular supply. Neurological surgery includes the operative and nonoperative management, diagnosis, evaluation, treatment, critical care, and rehabilitation of individuals with disorders of the nervous system and cerebrovascular disease, including aneurysms, trauma, and a wide range of tumors.

Minimally invasive approaches are being applied to a greater extent. Advances have been made in surgery for epilepsy, which has been unresponsive to medical treatment, through the precise excision of abnormal foci of electrical discharge in the brain. The gamma knife has been introduced to treat arterial-venous malformations and tumors located in parts of the brain inaccessible to direct surgery. The gamma knife consists of 202 individually focused beams of radiation which can be directed at such lesions with a very high degree of precision, sparing the surrounding normal brain almost completely. *See* MEDICINE.

Jeremy Steinbaum; Anthony L. Imbembo

Bibliography. R. W. Bailey et al., Laparoscopic cholecystectomy: Experience with 375 consecutive patients, *Ann. Surg.*, 214:531–541, 1991; M. J. Krasna and M. J. Mack, *Atlas of Thoracoscopic Surgery*, 1994; D. Sabiston (ed.), *Textbook of Surgery: The Biological Basis of Modern Surgical Practice*, 15th ed., 1997; S. I. Schwartz, G. T. Shires, and F. C. Spencer (eds.), *Principles of Surgery*, 6th ed., 1993; K. A. Zucker (ed.), *Surgical Laparoscopy Update*, 1993.

Surveillance radar

Any radar equipment used by civil or military authorities for locating aircraft, ships, or ground vehicles; most commonly, in civil air-traffic management, a radar used by controllers to indicate the position of aircraft aloft or on the airport surface. Two types of air-traffic control surveillance radar are employed internationally. Ground-based conventional, or primary, radar operates by transmitting microwave pulses and detecting the resulting energy reflected from the aircraft body (airframe). Secondary surveillance radar employs cooperative radio receiver-transmitter units (transponders) on the aircraft.

Primary radar. Primary radar was developed as an outgrowth of military surveillance efforts against hostile aircraft and ships in World War II. It can locate and track most aircraft without the need for special airborne equipment. Primary radar requires high-powered transmitters, capable of hundreds of kilowatts of peak power. It is sometimes limited in sensitivity by reflections (clutter) from terrain, ground objects, birds, and precipitation. However, clutter can be reduced by Doppler processing that discriminates between targets of different velocities and by adaptive clutter maps. *See* DOPPLER RADAR.

Secondary radar. Secondary radar, which evolved from military friend-or-foe identification techniques, provides several advantages. Sensitivity is improved

Characteristics of air-traffic surveillance radars in the United States

Characteristic	ASR-9	ARSR-4
Frequency, MHz	2800	1300
Primary reflector size, ft (m)	20 × 12 (6 × 3.7)	42 × 32 (13 × 10)
Secondary reflector size, ft (m)	28 × 4 (8.5 × 1.2)	42 × 32 (13 × 10)
Rotational rate, revolutions/min	12	5
Typical range, mi (km)	60 (100)	250 (400)

by the use of a transponder on the aircraft because the ground and airborne transmitters need only overcome the one-way path loss between the radar and the aircraft. (Radar transmitted power diminishes as the square of the range to the aircraft. In primary radar, the diminished power is reflected by the aircraft and is further diminished on the return trip to the radar for a total diminishment by the fourth power. In secondary radar, the power incident on the aircraft triggers a transponder, which then emits a fresh signal for the return trip, so that the power is only diminished as the square rather than the fourth power of the aircraft range.) Clutter is eliminated by the use of different frequencies for ground and airborne transmissions, 1030 MHz for interrogation and 1090 MHz for reply. The transponder antenna also makes sensitivity independent of the aircraft's size (radar cross section). In addition to sensitivity advantages, the transponder allows ground radar and aircraft to exchange coded messages.

The secondary surveillance radar system in current international use provides 4096 pilot-selected codes for limited aircraft identification and for altitude reporting in 100-ft (30-m) increments. An advanced international secondary radar standard operates on the same frequencies and provides 16 million codes to allow unique identifications for every aircraft in the world. This system, designated Mode S for selective address, also allows two-way exchange of data messages.

Mode S has been designed to be compatible with current ground and airborne equipment. Mode S ground sensors develop surveillance data for both current and Mode S airborne transponders using monopulse azimuth estimation techniques. The monopulse approach requires only four to six interrogations for each current transponder and one interrogation for each Mode S transponder; current ground interrogators use 20–30 interrogations. The time saved is used to selectively interrogate Mode S transponders and to conduct Mode S data-link transactions. *See* MONOPULSE RADAR.

Mode S transponders are designed to respond as a current transponder when interrogated by a conventional ground interrogator and to respond with Mode S replies when interrogated by either a ground Mode S sensor or an aircraft equipped with a Traffic Alert and Collision Avoidance System (TCAS). Thus there is complete compatibility between current ground and air equipment and Mode S. *See* AIRCRAFT COLLISION AVOIDANCE SYSTEM.

Monopulse azimuth estimation operates by simultaneously measuring the transponder reply amplitudes on sum and difference patterns of the ground antenna to determine the azimuth of the reply. The reply azimuth estimates are used to provide more accurate azimuth data (1 milliradian root-mean-square error) and assist in degarbling replies from current transponders to improve the reliability of the 4096-code and altitude data presented to the air-traffic controller.

Secondary radar surveillance is limited in that it operates only with transponder-equipped aircraft. All aircraft flying in United States controlled airspace above 10,000 ft (3048 m) and near major airports are equipped with transponders. However, 18% of general-aviation aircraft in the United States are not transponder-equipped. Another limitation of secondary radar results from shielding of the airborne antenna by the airframe in certain flight attitudes. Shielding sometimes causes temporary loss of coverage during takeoff or in steep banks. This problem can be overcome by using two aircraft antennas, with automatic antenna selection based on a comparison of the power received at the two antennas. All aircraft of large airlines in the United States carry Mode S transponders with diversity antennas.

Use. Although secondary radar generally provides more reliable and complete surveillance information than does primary radar, air-traffic authorities use both. In the United States, two major examples are the Federal Aviation Administration's Airport Surveillance Radar (ASR) and Air Route Surveillance Radar (ARSR), the latter shared at some locations with the U.S. Air Force (see **table**). Both normally operate unattended, their outputs fed by a landline or microwave link to their control positions. Primary radar provides backup coverage for nonequipped aircraft and for aircraft with airframe shielding. Primary radar can also detect potentially hazardous precipitation cells and potentially hostile aircraft or inadvertent intruders. Secondary radar is sometimes used without a primary radar backup for long-range, high-altitude surveillance where all aircraft are transponder-equipped, where airframe shielding is rare, and where the sensitivity disadvantage of primary radar requires costly high-power transmitters. The Federal Aviation Administration has authorized use of certain secondary radars having faster update rates and monopulse accuracy for monitoring aircraft conducting multiple parallel approaches in bad weather. *See* AIR-TRAFFIC CONTROL; RADAR.

Raymond R. LaFrey

Bibliography. J. L. Baker et al., Mode S system design and architecture, *Proc. IEEE*, 77(11):1684–1694, 1989; V. A. Orlando, The Mode S beacon radar system, *Lincoln Lab. J.*, 2(3):345–362, 1989; M. C. Stevens, *Secondary Surveillance Radar*, 1988.

Surveying

The measurement of dimensional relationships among points, lines, and physical features on or near the Earth's surface. Basically, surveying determines horizontal distances, elevation differences, directions, and angles. These basic determinations are applied further to the computation of areas and volumes and to the establishment of locations with respect to some coordinate system.

Surveying is typically used to locate and measure property lines; to lay out buildings, bridges, channels, highways, sewers, and pipelines for construction; to locate stations for launching and tracking satellites; and to obtain topographic information for mapping and charting.

Horizontal distances are usually assumed to be parallel to a common plane. Each measurement has both length and direction. Length is expressed in feet or in meters. Direction is expressed as a bearing of the azimuthal angle relationship to a reference meridian, which is the north-south direction. It can be the true meridian, a grid meridian, or some other assumed meridian. The degree-minute-second system of angular expression is standard in the United States.

Reference, or control, is a concept that applies to the positions of lines as well as to their directions. In its simplest form, the position control is an identifiable or understood point of origin for the lines of a survey. Conveniently, most coordinate systems have the origin placed west and south of the area to be surveyed so that all coordinates are positive and in the north-east quadrant.

Coordinate systems may be assumed, or established coordinate systems may be used. The most widely used systems in the United States are the various state plane coordinate systems established by the U.S. Coast and Geodetic Survey (now the National Geodetic Survey), which approximate in each case a small portion of the Earth's spherical form.

Vertical measurement adds the third dimension to an object's position. This dimension is expressed as the distance above some reference surface, usually mean sea level, called a datum. Mean sea level is determined by averaging high and low tides during a lunar month, but for certainty this must be carried out for as long as 19+ years.

Precision. Surveying devices and procedures possess individual limitations for accuracy of measurement. The term precision expresses the notion of degree of accuracy.

The choice of methods depends on the precision required for the result. Surveys of major installations or primary control surveys demand higher orders of precision than, for example, a preliminary highway profile or some lesser layout.

Instrumental surveys are designated first-order, second-order, and third-order. First-order is maximum precision, the others gradually less. Established criteria exist to differentiate the levels of precision.

Horizontal control. The main framework, or control, of a survey is laid out by traverse, triangulation, or trilateration. Some success has been achieved in locating control points from Doppler measurements of passing satellites, from aerial phototriangulation, from satellites photographed against a star background, and from inertial guidance systems.

In traverse, adopted for most ordinary surveying, a line or series of lines is established by directly measuring lengths and angles. In triangulation, used mainly for large areas, angles are again directly measured, but distances are computed trigonometrically. This necessitates triangular patterns of lines connecting intervisible points and starting from a baseline of known length. New baselines are measured at intervals. Trigonometric methods are also used in trilateration, but lengths, rather than angles, are measured. The development of electronic distance measurement (EDM) instruments has brought trilateration into significant use.

Traversing. Traverse lines are usually laid out as a closed polygon. This makes it possible to check the accuracy of the field measurements by calculating how nearly the figure closes mathematically. Commonly there are unavoidable random or accidental errors in angular and in distance measurements that tend to coalesce and militate against exact mathematical closure. Angular error is checked by geometric theory: the sum of all interior angles in any polygon equals the number of angles, less two, times 180°. The distance misclosure is determined by resolving the field-measured distance for each line into its north-south component (its latitude) and its east-west component (its departure). For the figure to close, components bearing north must equal in total length those bearing south, and components bearing east must equal in total length those bearing west. The net misclosure in the sum of latitudes and the net misclosure in the sum of departures are the two components of the linear error of closure of the traverse. If the total error of closure is within the prescribed error limit for the order of precision chosen for the survey, the total error is apportioned among the several angles and lengths of the traverse.

Triangulation. This begins with the selection of points whose connecting sight lines form one triangle or a series of triangles. In a series, each triangle has at least one side common to each adjacent triangle. An initial side length, the baseline, is measured, as are all the angles. By successive application of the law of sines, each successive side length is computed. The directions (azimuths) of the sides are also carried forward, so that the relative positions of all points can be determined.

Trilateration. In trilateration the figures are selected as for triangulation. An initial direction is determined and all side lengths are measured. Interior angles are computed by oblique-triangle formulas to obtain geometric checks on distances and to establish triangle-side directions. Ensuing latitude and departure computations yield the relative positions of the angle points of the figures.

Distance measurement. Traverse distances are usually measured with a tape or by EDM, but also may sometimes be measured by stadia, subtense, or trig-traverse.

Taping. The surveyor's tape, a steel ribbon graduated throughout its length, is usually 100 ft (30.48 m) long. However, 200-ft (60.90-m) and 300-ft (91.44-m) tapes are common.

Whether on sloping or level ground, it is horizontal distances that must be measured. Horizontal components of hillside distances are measured by raising the downhill end of the tape to the level of the uphill end. On steep ground this technique is used with shorter sections of the tape. The raised end is positioned over the ground point with the aid of a plumb bob.

Where slope distances are taped along the ground, the slope angle can be measured with the clinometer. The desired horizontal distance can then be computed. Precision in tape measuring is increased by refinements such as standardization, sag correction, tension control, and thermal-expansion correction.

Electronic distance measurement. In EDM the time a signal requires to travel from an emitter to a receiver or reflector and back to the sender is converted to a distance readout. It utilizes the precisely known speed of radio waves or of light, as well as a crystal-controlled chopped or sinusoidal variation in the emitted signal to measure distance in terms of number of wavelengths and a fractional part thereof. Microwave EDM requires a master instrument and a slave instrument, but has a capability of measuring up to 20 mi (32 km) or more. Shorter distances, from a few feet to about 2 mi (3.2 km), are best measured with ordinary light, laser light, or infrared EDM equipment. These require only a single instrument for emitting the signal and a simple retro-reflector or a reflector prism (corner cube) for returning the signal.

The great advantage of electronic distance measuring is its unprecedented precision, speed, and convenience. Further, if mounted directly onto a theodolite, and especially if incorporated into it and electronically coupled to it, the EDM with an internal computer can in seconds measure distance (even slope distance) and direction, then compute the coordinates of the sighted point with all the accuracy required for high-order surveying. The development of EDM is a great breakthrough, changing the nature and methods of a great part of modern surveying. Accuracies of 1 part in 10^6 are now assuredly within reach (**Fig. 1**).

Stadia. The stadia technique also requires no tape, although it is rather an imprecise distance-measuring method. A graduated stadia rod is held upright on a point and sighted through a transit telescope set up over another point. The distance between the two points is determined from the length of rod intercepted between two horizontal wires in the telescope.

Subtense. In the subtense technique the transit angle subtended by a horizontal bar of fixed length (usually 2.00000 m, approximately equivalent to 6.56168 ft) enables computation of the transit-to-bar distance (**Fig. 2**). With the 2-m (6.6-ft) bar and a theodolite to measure the angle (α) to 1 second, precise distance (D) can be calculated up to 200 m (660 ft) by the formula $D = 1.00000 \text{ cotan } {}^1/_2\alpha$ m. The subtense method can replace taping across a busy highway, canal, or ravine, although it is slower than EDM. For short to medium distances it is more precise than stadia.

Fig. 1. Instrument combines EDM and transit. (*Keuffel & Esser Co.*)

Trig-traverse. In trig-traverse the subtense bar is replaced by a measured baseline extending at a right angle from the survey line whose distance is desired. The baseline is at least long enough relative to the desired distance to assure the order of precision desired. The distance calculated in either subtense of trig-traverse is automatically the horizontal distance and so needs no correction because of difference of

Fig. 2. Subtense bar. (*Lockwood, Kessler, and Bartlett Inc.*)

elevation between theodolite and baseline or subtense bar.

Angular measurement. The most common instrument for measuring angles is the transit or theodolite. It is essentially a telescope that can be rotated a measurable amount about a vertical axis and a horizontal axis. Carefully graduated metal or glass circles concentric with each axis are used to measure the angles. Glass circles permit much greater magnification and more precise readings because concentrated light is passed through the portions of the circle being read. Further, superimposing the two opposite sides of the circle eliminates eccentricity error. The transit is centered over a point with the aid of either a plumb bob suspended by a string from the vertical axis or (on some theodolites) an optical plummet, which enables the operator to sight along the instrument's vertical axis to the ground through a right-angle prism.

Horizontal angles. To measure a horizontal angle between two intersecting lines, the transit is set up over the intersection. The telescope is sighted along one of the lines, the graduated circle is clamped against rotation, and the telescope is rotated to sight along the other line. The angle is indicated on the horizontal circle by another concentric circular plate, inscribed with an index and a vernier, that rotates with the telescope. Glass-circle transits and theodolites use optical micrometers instead of verniers, giving a much more precise reading. In either case, the angle can be read directly if the initial reading has been set at zero. Otherwise, the angle is found as the difference between the initial and final readings.

To lay off a predetermined angle from some reference line, the initial sight is taken along the line, the telescope is rotated through the angle desired, and a stake or other marker is set on the new line.

The special case of laying off a 180° angle is simply the extension of a straight line. It is done by backsighting along the reference line and rotating the telescope about its horizontal or elevation axis (transiting) for sighting ahead.

Vertical angles. These are measured by rotating the telescope about its elevation axis and reading the angle on the vertical circle. Vertical angles are measured from a hortizontal reference usually, though sometimes from the zenith or from the nadir.

Angular precision. Transit precision is denoted by the smallest angular distinction or resolution of which the instrument is capable. A 1-minute transit is so called because its circle is graduated to half-degrees, with the vernier measuring a thirtieth of the graduation, to enable an angular reading to 1 minute. Other transits read to 30 or 20 seconds. Theodolites with glass circles usually give readings to 0.1 minute (6 seconds); higher-order theodolites read to 1 second, even to a few hundredths of a second.

Elevation differences. Elevations may be measured trigonometrically in conjunction with reduction of slope measurements to horizontal distances, but the resulting elevation differences are of low precision.

Differential leveling. Most third-order and all second- and first-order measurements are made of differential leveling, wherein a horizontal line of sight of known elevation is sighted on a graduated rod held vertically on the point being checked. The transit telescope, leveled, may establish the sight line, but more often a specialized leveling instrument is used. For approximate results a hand level may be used.

In differential leveling, the rod is held on a bench mark, a point of known or assumed elevation. The level is set up and sighted for a reading on the rod. This reading, called the backsight or plus sight, is added to the bench-mark elevation to establish the height of the instrument. With the level remaining where it is, the rod is moved to a forward point (turning point). The reading to that point, called the foresight or minus sight, is subtracted from the height of the instrument to yield the elevation of the turning point. Distances to foresight and backsight, in pairs, are kept about equal so as to balance small instrumental errors and effects of Earth curvature and refraction of the line of sight. These are shown in exaggerated form in **Fig. 3**.

The foregoing process is duplicated successively as necessary to obtain the desired new elevation or elevations. Verification of the work is obtained by closing back on the original point by a rerun or by closing on another point of known elevation. Total error, if within allowable limits, is distributed along the level survey.

Reciprocal leveling, a variant of straight differential leveling, is used where the level cannot be set up for a long distance, as in crossing a ravine or a river. Here a short backsight and a long foresight (across the inaccessible distance) to a turning point on the opposite side are taken. Then the level is moved to the opposite side so that a long backsight is made to the original point of known elevation, and a short foresight is taken to the turning point. Since any errors of curvature and refraction in the long forward direction are canceled by those of the long backward sight, the average of the two elevation differences so obtained will be the actual difference.

Trigonometric leveling. Elevation difference can be determined also by trigonometric means, as in measuring the height of a mast, or even the elevation difference between two survey points. If the horizontal distance and the vertical angle are measured, the height of the sighted object above the

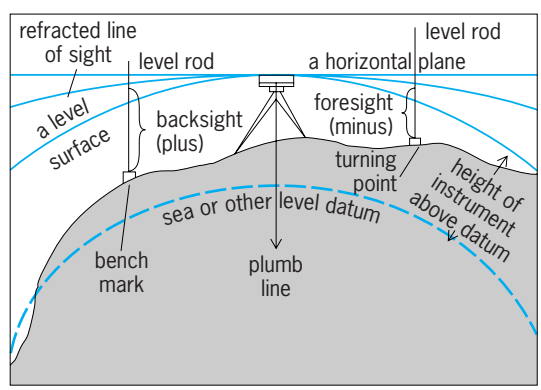

Fig. 3. Theory of differential leveling.

instrument can be calculated by right-triangle methods. In making long sights in this circumstance, a correction may be required for curvature and atmospheric refraction. An altitude observation (elevation angle) on a star or on the Sun will require refraction and parallax corrections.

Barometric leveling. Approximate elevation differences may be measured with the aid of a barometer. This method is particularly useful for reconnaissance of a substantial area.

Airborne profiling. A radar altimeter can give quite accurate elevation of an aircraft above the surface (of the ocean, for instance) by bouncing a signal from the surface back to the aircraft. Using it to give the altitude by bouncing a signal from any surface whose elevation is known can give the exact altitude as well. Thus by continuously recording a terrain profile on an overflight, the correct ground elevations can be obtained if a sufficient number of known ground elevations are scanned to provide control information. The technique is known as airborne profiling and renders fairly accurate results, adequate for a number of less demanding purposes.

Astronomical observations. To determine meridian direction and geographic latitude, observations are made by a theodolite or transit on Polaris, the Sun, or other stars. Direction of the meridian (geographic north-south line) is needed for direction control purposes; latitude is needed where maps and other sources are insufficient.

Meridian determination. The simplest meridian determination is made by sighting Polaris at its elongation, as the star is rounding the easterly or westerly extremity of its apparent orbit. At these times Polaris appears to be moving up or down the transit's cross hair, and there is ample time for assuring an accurate sighting. An angular correction is applied to the direction of sighting, which is referenced to a line on the ground. The correction value is found in an ephemeris. *See* EPHEMERIS.

The ephemeris also gives the noon declination of the Sun at Greenwich throughout the year. Declination is the angle of inclination of the Sun with respect to the equatorial plane. For direct solar observation the direction from the observer to the Sun is found from the equation below, where z is the

$$\cos z = \frac{\sin D}{\cos h \cos L} - \tan h \ \tan L$$

azimuth angle east or west from the meridian to the Sun; D is the declination at the instant of observation; h is the vertical angle to the Sun, corrected for refraction and parallax; and L is the observer's latitude.

A nonastronomic device for meridian determination is the portable gyroscope azimuth instrument, a newer development based on missile and aircraft inertial guidance systems.

Latitude determination. Latitude of a position may be determined directly at night by vertical-angle sighting on Polaris at upper or lower culmination (the northerly or southerly extremity of the star's apparent orbit) and application of the suitable vertical-angle correction from the ephemeris. In another method, the vertical angle to the Sun is observed at noon, and the Sun's declination is subtracted algebraically to yield a net angle, which is subtracted from 90° to give the latitude.

Longitude and time observations are made occassionally by noting the passage of the Sun across an established meridian. If Greenwich time is known, longitude can be computed.

Geoid. Astronomic observations from any station yield geographic latitude and longitude for that station; a great number of these stations connected by a network of triangles form a control network. An ellipsoid chosen most nearly to fit the geoidal shape of the Earth affords a base to calculate the interrelationship of the astronomic stations. Lack of conformity between measured distances and calculated distances between stations indicates disparity between astronomical and geoidal (geodetic) latitudes and longitudes, which in turn is traceable to deflection of the astronomic vertical from the ellipsoid vertical at the station or stations. These deflections of the vertical are due to undulations of the geoidal surface and its lack of conformity to the ellipsoid. Adjustments over a wide area (a continent) are usually made. The initial choice of a best ellipsoid renders the conformity closest, so that geodetic (ellipsoidal) computations will best fit the true geoid.

Several ellipsoids have been in use for different continents, but massive accumulations of measurements of gravity and deflections of the vertical through the space satellite program furnish the basis for an ellipsoid that more closely fits the geoid worldwide. It is the World Geodetic System Ellipsoid, the first of its kind. It permits the major geodetic networks of the world to be unified and the coordinates of points anywhere on Earth to be compatible. *See* GEODESY.

Types of surveys. These include geodetic control surveys, route surveys, construction surveys, and cadastral (property) surveys. *See* TOPOGRAPHIC SURVEYING AND MAPPING.

Geodetic control surveys. Control surveys provide the reference framework for lesser surveys. A traverse, with elevations of its points, may be the control for mapping a limited area. The broadest control surveys are the National Geodetic Survey (NGS) in the United States and by corresponding agencies in other countries, wherein the horizontal and vertical positions of points are established with first- or second-order accuracy.

In geodetic surveys, Earth curvature is taken into account. Coordinate positions are established in terms of latitude and longitude. Geodetic coordinates are convertible into state plane coordinates.

The traditional method of extending horizontal geodetic control is triangulation. Trilateration and traverse are supplementary techniques.

In geodetic surveys, refinements are added to the simple chain-of-triangles scheme. The most notable refinement is use of the quadrilateral as a geometric unit (**Fig. 4**). Diagonal directions as well as the sides are observed, so that the quadrilateral includes two

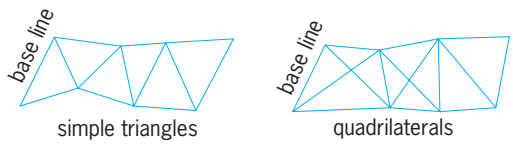

Fig. 4. Triangulation figures, used in geodetic surveys.

pairs of overlapping triangles, one being redundant although valuable as a check. In this way additional angular checks are available, and four separate calculations can be made for the entering side length of the next quadrilateral in the chain. The shapes of the triangles are important, small angles being avoided to preserve high strength of figure. A strength of figure analysis for the four different computational routes reveals the one that will best carry the length forward.

For precision and economy, geodetic triangulation stations are located as far apart as possible, consistent with the requirement of intervisibility. On the other hand, a precise control network for a builtup area may require numerous close-spaced auxiliary stations within quadrilateral figures.

Angles are measured by a precise theodolite, all angles around the horizon being observed several separate times for determination of the statistically best value of each angle. Long sightings are usually made on lighted station signals at night.

Because of the Earth's spheroidal shape, the interior angles of large triangles add up to more than 180°. A 75-mi^2 (200-km^2) triangle, for example, has 1 second more than the 180° total of interior angles. The extra second is called spherical excess.

Large triangulation chains of a primary network will have widely spaced ground stations of calculated position, which are of value as control to only a limited ground area. Thus, secondary triangulation networks are run to connect points on the primary. Then, to densify the control and make it available virtually everywhere, tertiary networks of triangulation or traverse are employed. Precise control for serving a city requires a very heavy saturation of auxiliary stations within the smallest triangulation figures.

Over-ocean geodetic control extension, formerly infeasible, now is accomplished by systems of simultaneous electronic distance measurements from the ground principles. Doppler observations of passing satellites have proved an astounding advance, enabling the location of ground stations to within less than a meter virtually anywhere on the Earth's surface. *See* SATELLITE NAVIGATION SYSTEMS.

In another method, simultaneously photographing a satellite against a star background from three camera stations (one, say, on a island) affords measurements to calculate the island station's position, again conquering the overwater problem. Offshore drilling platforms are located by one or several of these now feasible methods.

Each precisely determined distance between geodetic control stations is reduced to its sea-level projection, that is, the distance between the stations if projected vertically to sea level.

Geodetic data convert to state plane coordinate data with the aid of tables published by the NGS. Each plane coordinate system is expressed as a projection of the curved surface onto a plane intersecting either two standard parallels (latitudes) or two meridians. In both cases the meridian and parallel projections are at right angles to each other. At the standard meridians or parallels, scale is exact; elsewhere a known variable scale factor must be applied.

Advantages of state plane coordinate surveys include the ability to initiate a long route survey several places at once, and to interrelate survey points anywhere in the system. Proposed land data systems or a national cadastre would depend on the coordinate systems.

The vertical control system of the NGS consists of a first-order level network with supplementary second-order lines. In most parts of the United States, bench-mark elevations are available within a few miles of any point.

The geodetic leveling procedure is a refinement of that previously described, using more precise instruments and methods. The level instrument is built to rigid specifications. Rod graduations are on an Invar strip (Invar varies only negligibly under temperature change). Special care is taken to equalize foresight and backsight distances, to assure stable turning points, and to protect the instrument from minor stresses.

Route surveys. Surveys for the design and construction of linear works, such as roads, canals, pipelines, or railways, are called route surveys. They begin with reconnaissance and continue through preliminary, location, and construction surveys.

Reconnaissance for a new highway, for example, may be accomplished by study of existing maps together with a visual appraisal of field conditions, or even quick low-order horizontal and vertical field measurements. Controlling points, such as favorable ridge and river crossings, are found, and a preliminary line is selected. This line traditionally is laid out by transit and tape or EDM, being a linear traverse. It is profiled, that is, levels are run along the traverse line to find elevations. Transverse profiles (cross sections) are made at needed intervals. Structures and natural objects that would affect the final location are fixed by side shots. A side shot may be a direction and distance from a transit point, the intersection of two directions or two distances, or a perpendicular offset from a traverse line. Transverse profiles are made by hand level, tape, and level rod.

The result of the preliminary survey is a strip topographic map of sufficient precision to permit preliminary design of the final location, including approximate determination of earthwork quantities. More and more preliminary maps and designs are executed from topographic maps constructed by airborne photographic (photogrammetric) methods, with need for only limited ground surveying prior to construction layout.

The location survey line is conducted with at least third-order precision. Traverse procedures are followed, but curves are laid out and stationed. The

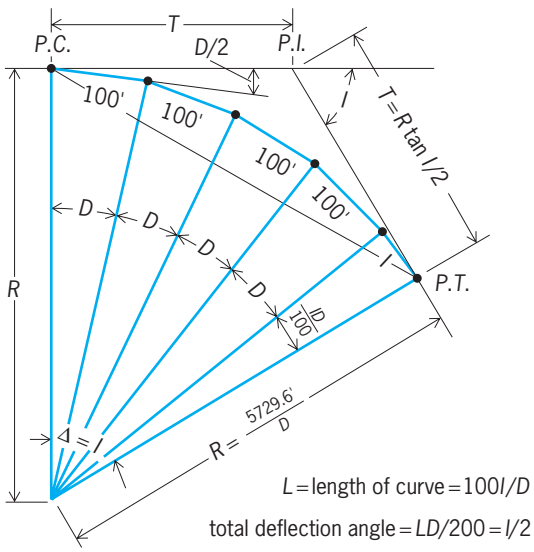

Fig. 5. Circular curve theory.

result is a staked centerline for the route to be constructed. Profile leveling and cross-sectioning for earthwork quantities also are part of the location survey.

Contract plans normally are based on the location survey. The construction stakeout is usually conducted later, but sometimes is carried out at the same time as the location survey. A prerequisite is final grade selection so that cuts and fills can be marked on centerline stakes and slope stakes can be set. The slope stake indicates the lateral limits of cut or fill at a given cross section. It is marked with the distance, plus or minus, from existing ground level to the proposed centerline grade. Being at the edge of earthwork, the slope stake supposedly is available for reference throughout construction.

A horizontal curve provides for change in direction for traffic. The directions, called tangents, connect at an angle point, called the point of intersection (P.I.). The intersection angle I is the deflection angle, right or left, from the forward projection of the tangent into the angle point (**Fig. 5**).

Any circular curve or arc, connecting two intersecting tangents, has a central angle, equal to the angle I formed by perpendiculars (radii) to the two tangents, one at the point of curvature (P.C.) and the other at the point of tangency (P.T.). The greater the radius, the more gradual the curve. The sharpness of the curve also is expressed as degree of curvature D, the central-angle increment that subtends 100 ft (30.48 m) of arc (highway definition) or 100 ft of chord (railroad definition). A 1° curve has a radius of about 5729.6 ft (1746.4 m) under either definition.

Circular curves are laid out by the following procedure, assuming D, and consequently radius R, have been chosen:

1. The tangent distance is measured back from the P.I. to establish the P.C.

2. The transit is set up at P.C., and a series of deflection angles, each equal to $D/2$, is turned in the direction of curvature to establish a point on the curve for each 100 ft (30.48 m) of chord or arc. For chords or arcs shorter than 100 ft the turned deflections angle is proportionally smaller.

With the advent of EDM equipment, alternatives have evolved for laying out points on a horizontal circular curve. The transit with an adjunct EDM device can be mounted at the P.C. or the P.T. (or indeed any control point of the survey), and distances and directions can be used directly to set curve stakes. The need for convenient and easy taping is no longer a constraint as long as sight distances are clear.

Vertical curves, connecting different grades, usually are parabolic. The parabola has an inherent transition, and it is easy to lay out with tape, level, and level rod after computation of ordinates for vertical-curve grade points along the centerline.

The traditional route survey procedures described above are economical for narrow routes in moderate terrain. Major dual highways with broad rights of way are located more efficiently on large-scale aerial topographic maps. On such maps, "paper" location surveys can be made, complete with earthwork quantity takeoff precise enough for contract bidding. Furthermore, digitized terrain data, directly available from some photogrammetric procedures, permit rapid electronic computation of quantities. With this approach it may be feasible to deter all field staking until construction time, the staking points being selected in advance by scaling or computation of map-coordinate relationships with points established during the small amount of ground survey required to control the photogrammetric survey. *See* PHOTOGRAMMETRY.

Construction surveys. Surveys for construction layout establish systems of reference points that are not likely to be disturbed by the work. Slope stakes are earthwork references. Buildings, bridge abutments, sewers, and many other structures traditionally are controlled by batter boards, horizontal boards fastened to two uprights. The top of a batter board is set at the elevation of the line to be established, and the horizontal position of the line is indicated by a mark or nail. Across the site another batter board is set up for the given line, and a string or wire stretched between is the line. The line may be a building face at first-floor level, or it may be a reference line. In trench work the line between batter boards may run at a fixed distance above the invert centerline.

In lieu of strings or wires, lines of sight may be used. A transit is set up on lines outside the work area, and points of the line are sighted as required. Further, a low-power laser, available as an attachment to the transit or as a separate special purpose instrument, can project a visible reference line or plane from an unattended position. A means of locating critical construction points, such as those for anchor bolts, is to compute their positions in a coordinate grid, then compute directions and distances from a reference point for their location by transit sighting and tape measurement.

Elevation controls are provided by bench marks near the construction area. The foregoing

construction techniques are adaptable into industrial plants for building large mock-ups and jigs as well as for the alignment of parts. Transits and levels on stable mounts may be used; however, related specialized equipment called optical tooling instruments are more readily applied. These include a transit or level telescope with an optical micrometer on the objective, to move the line of sight to a locus parallel with the initial sighting and flat self-reflecting mirrors for use as targets. Angular sight lines are established by a mirror mounted vertically on a transit base.

Underground surveys. Mine and tunnel surveys impose a few modifications on normal surveying techniques: Repeated independent measurements are made because normal checks (such as closed traverses) are not available; cross hairs illuminated because work is performed in relative darkness; vertical tape measurements and trigonometric levels, instead of differential levels, frequently must be relied upon; and in adits and tunnels survey points are placed overhead, rather than underneath, to save them from disturbance by traffic. On the mining transit, an auxiliary telescope outside the trunnion bracket facilitates steep sightings.

Traditionally the most exciting underground survey process has been the transferring of a direction from the surface. In shallow shafts, steep (but not vertical) sights may transfer the direction. Another technique is to hang two weighted wires down the shaft, observe the direction between them on the surface, and use this direction as a control below. The relative shortness of distance between wires makes the transfer geometrically weak, but procedural care enables satisfactory results. An alternative procedure applies inertial navigation principles to underground surveying. A north-seeking gyro mounted on a transit gives azimuth accuracy to about 6 seconds from a plumbed position at shaft bottom and carries tunnel lines forward without the angle-error accumulation in ordinary transit traverse lines (**Fig. 6**).

Hydrographic surveys. Data for navigation charts and underwater construction are provided by hydrographic surveys. The horizontal locations of depth measurements must be referenced to recognizable controls. Where the shoreline is visible, it is mapped and a system of triangulation stations is established on shore. Transits at two triangulation stations can be used to observe directions to the sounding vessel whenever it signals that a depth measurement is made. A check angle may be obtained by sextant observation of shore points from the deck of the vessel, or a third transit on shore may be used to provide a check intersection. Angular observations from a shore point, coupled with EDM lengths to the sounding vessel reflector or transponder, will also fix the sounding location. In another procedure, a sending microwave unit on the vessel can find distances from slave stations set on shore points, to be recorded (even plotted) instantly along with the depth sounded.

Cadastral surveys. To establish property boundary lines, cadastral surveys are made. Descriptions based on horizontal surveys are essential parts of any document denoting ownership or conveyance of land. The basis rule of property lines and corners is that they shall remain in their original positions as established on the ground. This basis rule is important because most land surveys are resurveys. They may follow the original description, but this description is merely an aid to the discovery of the originally established lines and corners. Substantial discrepancies are frequently found in original descriptions because low-order surveying devices such as the compass or the link chain were once used.

Fig. 6. North-seeking gyro eases underground survey tasks. (*Wild Heerbrugg Instruments, Inc.*)

Surveys in the original 13 states, plus Tennessee, Kentucky, and parts of others, are conducted on the metes and bounds principle. In the so-called public land states and in Texas, the basic subdivisions are rectangular.

If the boundaries to be described border an irregular line, such as a winding stream, a good mathematical description may be impossible. Such a line can, however, be located by a series of closely spaced perpendicular offsets from an auxiliary straight line.

In the rectangular system, the land parcels of a region are described by their relationship to an initial point. In public-land states the initial point is the intersection of a meridian (principal meridian) and a latitude (baseline), as in **Fig. 7**. Townships, normally 36 mi^2 (93.24 km^2), are designated by their position with respect to the initial point—the number of tiers north or south of a given baseline and the number of ranges east or west of the corresponding principal meridian.

Within the township each square mile, or section, has a number, from 1 to 36. Sections are subdivided into quarter sections (160 acres or 0.6475 km^2), and they may be subdivided further. North-south section lines are meridians originating at 1-mi (1.609-km) intervals along the baseline and along standard

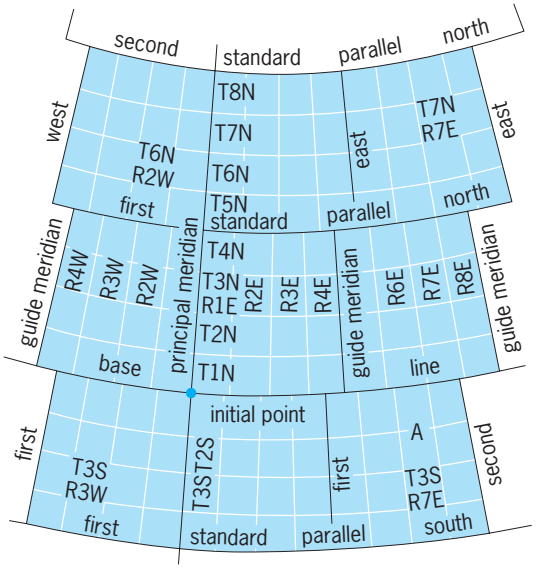

Fig. 7. Rectangular land layout. (*After P. Kissam, Surveying for Civil Engineers, McGraw-Hill, 1956*)

parallels of latitude (correction lines) generally spaced 24 mi (38.62 km) apart. The respacing of meridional lines every 24 mi reduces the effect of meridian convergence on the size of sections.

Resurveys become difficult where the corners are obliterated or lost. An obliterated corner is one for which visible evidence of the previous surveyor's work has disappeared, but whose original position can be established from other physical evidence and testimony. A corner is deemed lost when no sufficient evidence of its position can be found. Restoration requires a faithful rerun of the recorded original survey lines from adjacent points, distance discrepancies being adjusted proportionately. *See* CIVIL ENGINEERING.

B. Austin Barry

Bibliography. B. A. Barry, *Construction Measurements*, 2d ed., 1988; C. M. Brown, *Evidence and Procedures for Boundary Location*, 3d ed., 1993; H. Kahmen and W. Faig, *Surveying*, 1988; J. C. McCormac, *Surveying*, 3d ed., 1994; F. H. Moffitt and H. Bouchard, *Surveying*, 10th ed., 1997; P. R. Wolf and R. C. Brinker, *Elementary Surveying*, 9th ed., 1997.

Surveying instruments

Instruments used in surveying operations to measure vertical angles, horizontal angles, and distance. The devices used for these measurements were originally mechanical only, but advances in the technology led to the development of mechanical-optical devices, optical-electronic devices, and finally, electronic-only devices.

Level. Four types of levels are available: optical, automatic, electronic, and laser.

Optical level. An optical level is used to project a line of sight that is at a 90° angle to the direction of gravity. Both types, dumpy and tilting, use a precision leveling vial to orient to gravity (**Fig. 1**). The dumpy type was used primarily in the United States, while the tilting type was of European origin and used in the remainder of the world. The dumpy level has the leveling vial fixed to the telescope, which is fixed at 90° to a rotatable vertical spindle. Leveling screws, attached to the spindle, are used to center the leveling vial. The telescope can be pointed in any direction, and the line of sight through the telescope will then project at 90° to the direction of gravity. Tilting levels are similar to dumpy levels, with the only exception being that the leveling vial is connected to the telescope with a fine-movement (tilting) screw. Leveling screws are used to roughly center the vial, and the tilting screw is used for the final centering of the vial for each observation. *See* LEVEL MEASUREMENT.

Automatic level. Automatic levels (**Fig. 2**) use a pendulum device, in place of the precision vial, for relating to gravity. The pendulum mechanism is called a compensator. The pendulum has a prism or mirror, as part of the telescope, which is precisely positioned by gravity. The pendulum is attached to the telescope by using precision bearings or wires (metallic or nonmetallic). Leveling screws are used to roughly center a circular vial, and the optics on the pendulum then correct the line of sight through the telescope. Automatic levels are easy and fast to use, resulting in their domination of the optical-level market segment.

Fig. 1. Engineer's dumpy level. (*W. and L. E. Gurley Co.*)

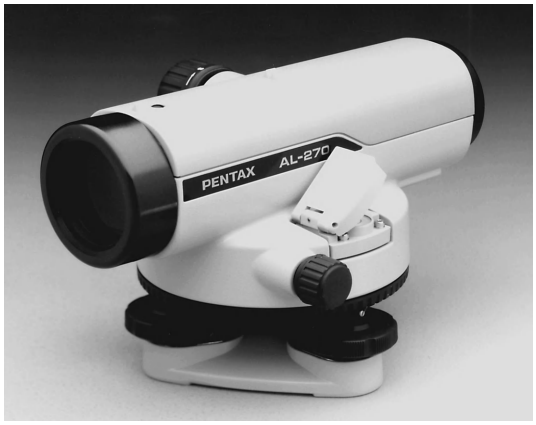

Fig. 2. Automatic level. (*Pentax Corp.*)

Electronic level. This type of instrument has a compensator similar to that on an automatic level, but the graduated leveling staff is not observed and read by the operator. The operator has only to point the instrument at a bar-code-type staff, which then can be read by the level itself. The electronic determination of the data is a further advantage because the data can be transferred to a data collector or stored on-board in a memory module or card. The electronic level eliminates human reading error and increases the speed at which leveling work can be performed. The only significant disadvantage is the high cost as compared to the optical automatic level.

Laser level. Although this type of instrument is categorized as laser, these levels actually employ three different types of light sources: tube laser, infrared diode, and laser diode. The instrument uses a rotating head to project the laser beam in a level 360° plane. The advantages are twofold: no operator is required once the instrument is set up; and different people in various locations can work by using a single light source. The disadvantages are that accuracy is less than that provided by other types of levels and that the cost is significantly higher. This type of instrument is used primarily for construction surveying.

Tube lasers have a visible beam with high power requirements. The power source is usually a 12-V car battery or an ac-to-dc converter. Infrared diode units have a nonvisible beam that can operate on flashlight batteries. The nonvisible beam requires that an electronic detector be used to "see" it. This instrument has a lower cost than tube lasers, and is physically smaller. The laser diode level has all the advantages of the other two types: low power requirement, small size, and visible beam.

There are three different mechanisms used to level-up laser levels. Leveling vials were used initially, and they are still used on the most inexpensive units. Moderate- to mid-priced units employ compensators. For the higher-priced units, electronic vials with servomotors mechanically level the mechanism inside the case. Servo types are also available in models that can tilt the rotating beam to a particular percent of grade. Dual-grade instruments, the most expensive type, can project two different grades at the same time, for example, (+)3% in the north direction and (−)5% in the east direction. *See* SERVOMECHANISM.

Transit. The primary purpose of a transit (**Fig. 3**) is to measure horizontal and vertical angles. Circles, one vertical and one horizontal, are used for these measurements. The circles are made of metal or glass and have precision graduations engraved or etched on the surface. A vernier is commonly used to improve the accuracy of the circle reading. *See* VERNIER.

The vertical circle, fixed to a horizontal axis that is part of and at 90° to a telescope, lets the telescope rotate in a vertical plane. The alidade, or the framework that supports the telescope axis, has a vertical axis (spindle) attached to its base, which allows the alidade to rotate with respect to the horizontal circle (Fig. 3). Level vials, attached to the alidade, are used to make the spindle vertical (in line with gravity), so that measurement of horizontal angles can be measured. A level vial is also attached to the telescope and provides the gravity index to measure vertical angles.

Theodolite. The theodolite serves the same purpose as the transit, and they have many similar features. The major differences are that the measuring circles are constructed only of glass and are observed through magnifying optics to increase the accuracy of angular readings. Theodolites are generally smaller than transits. Some models have a compensator, which allows quicker setups while maintaining vertical angle accuracy. Theodolites also have a separate, low-power telescope to view the setup point on the ground. This telescope replaces the use of a plumb bob to position the instrument over a point. Optical theodolites largely replaced transits but then were replaced by electronic theodolites.

The electronic theodolite uses electronic reading circles in place of the optically read ones. The circle readings are observed on a display located on the alidade. Some versions of this instrument can transmit the electronic readings via a serial port to an external data memory device or to an electronic distance meter attached to the theodolite.

Electronic distance meter. Electronic distance meters use either light waves or radio waves as their

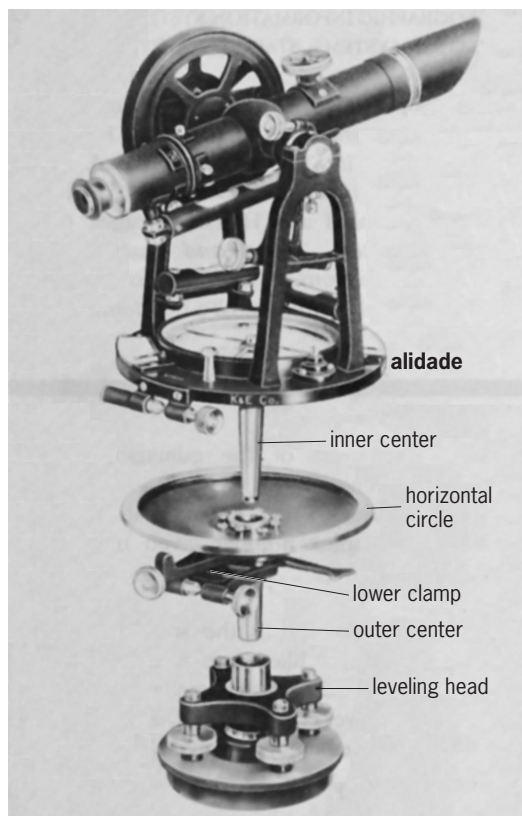

Fig. 3. Disassembled transit, showing its components. (*Keuffel and Esser Co.*)

measuring device. Radio-wave instruments dominated this category until the late 1960s, when lower-cost light-wave instruments became available.

Radio-wave instruments require similar units placed at each end of the line to be measured, are capable of longer measurements (up to 90 mi or 140 km), and do not require a direct line of sight between the units; however, they are more sensitive to atmospheric conditions.

Light-wave types use visible, laser, or infrared light. Only one unit is required, with the other end of the measured line occupied by a special type of prism reflector. Measurements are more precise, distance capability is shorter (up to 40 mi or 60 km for laser and 6 mi or 10 km for infrared), and a direct line of sight is required between the unit and the reflectors.

These instruments provide more stable results. Accuracy has improved to the degree that a basic accuracy of ±2 mm is commonly achieved. These instruments typically use several frequencies and phase-comparison techniques to determine distance, not time measurement.

Total stations. These units consist of combinations of devices, which can be optical, electronic, data-electronic, or motorized electronic.

Optical. The combination of a theodolite and an infrared electronic distance meter into one instrument is commonly referred to as an optical total station. The configuration can be one integrated unit, or it can be modular in design. Integrated types typically have longer-range distance capabilities matched with higher angle accuracy. Modular types can have any combination of short- to long-range distance measuring matched with low-to-high angle accuracy.

Electronic. An electronic total station (**Fig. 4**) comprises an electronic theodolite in combination with an infrared electronic distance meter. This type of instrument has become the standard surveying instrument, having replaced all other angle and distance instruments for most applications. Compensators are common on these instruments, and many models have dual-axis types that correct for misleveling in both directions. Some units have been developed that have many useful programs built in. The category of electronic total station includes models that serve virtually all survey needs, and less expensive models have been designed specifically for use in construction. These instruments are very efficient when used alone, and the addition of a data collector, replacing hand-written notes, is the next step in a fully electronic system. The data collector reduces mistakes in the field and can perform some useful calculation functions quickly and easily. The final component is the office computer to which the data collector can transfer the information. Through the use of specialized surveying software, maps are generated by a plotter or printer (Fig. 4).

Data electronic. Miniaturization of electronic components made it possible to combine an electronic total station with the data collector. Small removable modules or cards are used to store the data, which can subsequently be transferred to the office computer. These models provide a one-piece instrument, eliminating cables and facilitating transport. Some models have full alphanumeric keyboards so that descriptions and notes can be added to the measured data. Special units have been designed that use custom programs.

Fig. 4. Electronic total station. (*Pentax Corp.*)

Motorized electronic. This type of electronic total station has two characteristics that distinguish it from the standard type. The manual alignment mechanisms are replaced by motors, and the optical telescope is replaced by an electronic type. The instrument can align itself to the prism reflector, thus eliminating the instrument operator. Data recording can be initiated remotely by the reflector operator, making the process a one-person operation. This reduction in personnel is the primary advantage of this system.

Global Positioning System. The U.S. Department of Defense installed a satellite system for navigation and for establishing the position of planes, ships, vehicles, and so forth. This system uses special receivers and sophisticated software to calculate the longitude and latitude of the receiver. It was discovered early in the program that the distance between two nonmoving receivers could be determined very accurately and that the distance between receivers could be many miles apart. This technology has become the standard for highly accurate control surveys, but it is not in general use because of the expense of the precision receivers, the time required for each setup, and the sophistication of the process.

Mapping of city features such as systems for supplying water, sewage collection, and electricity does not require the precision of normal surveying. Low-cost, low-precision receivers are used for this purpose, and cities are able to have cost-effective multilayer databases that contain all of the city structure located below and above the ground. *See* GEOGRAPHIC INFORMATION SYSTEMS; SATELLITE NAVIGATION SYSTEMS; SURVEYING. Kerry W. Kemper

Bibliography. J. M. Anderson and E. M. Mikhail, *Surveying: Theory and Practice*, 7th ed., 1997; C. D. Burnside, *Electromagnetic Distance Measurement*, 3d ed., 1991; M. A. R. Cooper, *Modern Theodolites and Levels*, 2d ed., 1982; F. Deumlich, *Surveying Instruments*, 1982; B. Hofmann-Wellenhof and H. Lichtenegger, *Global Positioning System: Theory and Practice*, 5th ed., 2001; A. Leick, *GPS Satellite Surveying*, 3d ed., 2003; J. M. Rueger, *Electronic Distance Measurement*, 4th ed., 1996.

Susceptance

The imaginary part of the admittance of an alternating-current circuit.

The admittance, Y, of an alternating current circuit is a complex number given by Eq. (1). The imaginary part, B, is the susceptance. The units of susceptance like those of admittance are called siemens or mhos. Susceptance may be either positive or negative. For example, the admittance of a capacitor C at frequency ω is given by Eq. (2), and so B is positive. For an inductor L, the admittance is given by Eq. (3), and so B is negative.

$$Y = G + jB \tag{1}$$

$$Y = jC\omega = jB \tag{2}$$

$$Y = -\frac{j}{L\omega} = jB \tag{3}$$

In general, the susceptance of a circuit may depend on the resistors as well as the capacitors and inductors. For example, the circuit in the **illustration** has impedance given by Eq. (4) and admittance given by Eq. (5), so that the susceptance, given by Eq. (6), depends on the resistor R as well as the inductor L.

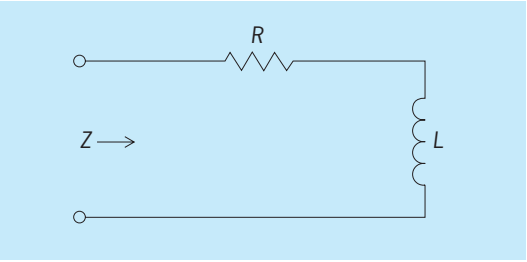

Circuit with a resistor and inductor in series.

$$Z = R + jL\omega \tag{4}$$

$$Y = \frac{1}{R + jL\omega} \tag{5}$$

$$B = \frac{-L\omega}{R^2 + L^2\omega^2} \tag{6}$$

See ADMITTANCE; ALTERNATING-CURRENT CIRCUIT THEORY; ELECTRICAL IMPEDANCE. J. O. Scanlan

Suspension

A system of small solid particles dispersed in a liquid. Suspensions are a type of colloidal dispersion. By classical definition, suspension particles have diameters (or, if not spherical, the largest dimension) between 1 nanometer and 1 micrometer. However, for practical purposes the principles of colloid science can be usefully applied to suspensions for which the particles are smaller, down to about 0.1 nm (nanoparticle suspensions), and also to those for which the largest particle dimensions are tens or even hundreds of micrometers. The particles in a suspension are large enough that they do not behave like the atoms and molecules of classical chemistry. For example, suspensions do not generally behave like true solutions, and may have undetectable freezing-point depressions. On the other hand, the particles are small enough that they do not behave like the macroscopic particles of classical physics. For example, sedimentation of a suspension's particles may occur extremely slowly in apparent violation of Stokes' law. *See* COLLOID; PARTICULATES; SEDIMENTATION (INDUSTRY).

In most suspensions, the solid particles are dispersed in an aqueous liquid phase (that is, water containing dissolved substances), which may be denoted solids-in-water or S/W. In some suspensions, however, the liquid phase is an oil (usually a hydrocarbon liquid), and the suspension is denoted solids-in-oil or S/O. The **table** shows some examples.

Preparation. Suspensions can be formed by one of two basic methods. In the nucleation process, either a gas is made to condense to form solid particles or particles are precipitated from a solution. Once the fine particles have been created, they can grow to larger sizes as additional material condenses or precipitates onto the surfaces of existing particles (Ostwald ripening). The second basic suspension preparation method involves breaking-up large particles into progressively smaller ones (comminution). This process requires shearing forces, which can be provided by various devices, including cutting, rolling, crushing, or grinding mills. When the solid particles are already quite small, other devices such as propeller-style mixers or ultrasound generators can be used.

If a small amount of powdered fine solid is added to a cup of tap water and vigorously agitated in a blender, the fine solids will become quite well dispersed in the water and a suspension is made. This suspension may or may not be very stable, however.

Examples of suspension occurrence

Application area	Examples	Suspension types*
Environment	Sediment particles in oceans, lakes, rivers	S/W
	Muds, quicksand	S/W
	Rock fragments in lava	S/L
	Solid silicates in magma	S/L
	Solid particles in raindrops	S/W
	Biocolloids (viruses, bacteria) in rivers, ponds	S/W
Mineral processing	Mineral process slurries (pulps)	S/W
	Mineral flotation froths	S/W
	Oil flotation and oil-assisted froths	S/W, S/O
	Mineral tailings slurries and ponds	S/W
Petroleum production	Migrating fine solids in an oil reservoir	S/W
	Drill-cutting slurries	S/W, S/O
	Dispersed asphaltenes in crude oils	S/O
	Oil-well-produced solids	S/W, S/O
	Oilfield tailings ponds, sludges	S/W, S/O
Manufacturing and materials	Cellulose fiber pulp slurries	S/W
	Deinking pulp slurries	S/W
	Pigment-containing paints	S/L
	India ink	S/W
Food products	Ice crystals in frozen ice cream	S/O
	Vegetable shortening	S/O
	Jellies	S/W
	Chocolate drinks	S/O
	Water ice (a semifrozen drink)	S/W
Agriculture	Pesticide, herbicide, and fungicide	S/W
Biology and medicine	Polymer-encapsulated drugs	S/W
	Biodegradable drug suspensions	S/W
	Diagnostic suspensions	S/W
Personal care	Exfoliating scrubs	S/W
	Facial masks	S/W
	Lipsticks and lip balms/glosses	S/O

*Solid, liquid, water, and oil.

In the absence of turbulence and if the particles are not electrically charged, they will fall (settle) out if their density is greater than that of the liquid. It is principally the small size that keeps suspended sediment particles in a river from rapidly settling out. A fine sand particle of 100-μm diameter will settle 1 m (3 ft) in water in about 2 minutes, whereas a 10-nm colloidal particle will take about 20 years (in the absence of turbulence).

In addition, the particles may not remain separated from each other. When Brownian motion, stirring, or settling cause particles to encounter each other, they may stick together, possibly touching at one or more points. This is called aggregation. In suspensions, there are two principal kinds of aggregation: coagulation and flocculation. Coagulation refers to the formation of compact aggregates. Flocculation refers to the formation of a loose network of particles. Coagulated particles tend to settle quickly and form compact, dense sediments, while flocculated particles tend to settle slowly and form loose, high-volume sediments. Which kind of behavior occurs depends mainly on the properties of the particles' surfaces. *See* BROWNIAN MOVEMENT.

In order for dispersed particles to be reasonably stable and not aggregate or settle quickly, suspension particles need to be stabilized. The two most common means of stabilizing suspended particles are by creating electric charges or adsorbing polymers on the particles' surfaces. Some particles gain a surface electric charge due to reactions that occur when they are dispersed into water. Another way to create particle surface charges is by adsorption of chemicals such as ionic surfactants or polymers. Electrically charged particles repel each other, which can dramatically slow down the rates of both aggregation and sedimentation, making a suspension reasonably stable. *See* ADSORPTION; POLYMER; SURFACTANT.

The stability of a suspension can also be enhanced by adsorbing certain kinds of polymers onto particle surfaces, without involving electric charges. In this case, the adsorbed polymer molecules extend outward from the particle surfaces into solution, making it difficult for particles to approach each other, and therefore providing a barrier to aggregation (steric stabilization). Properly formulated suspensions can remain stable for months or even years.

Properties. Although suspensions can be prepared that will remain stable for weeks, months, or years, no suspensions are truly stable. If the particles are dispersed in a liquid of lower density, they will eventually fall (sediment) due to gravity. Similarly, if the particles are dispersed in a liquid of greater density,

the particles will eventually rise (cream) to the surface.

By dispersing fine particles in a liquid, a very small amount of solid material can be made to expose a very large amount of surface area. This can create a very large number of sites at which chemical reactions can occur. As a result, nanoparticle suspensions have potential applications in high-surface-area catalysis, nanosized pigments and/or fillers in paints and other kinds of coatings, and ceramics. *See* NANOPARTICLES.

In a very dilute suspension of small particles, the individual particles will not encounter each other very often. However, this changes as the volume occupied by the particles increases above a few percent. As particles begin to encounter each other more and more frequently, there is a crowding effect, and it becomes more difficult to make the suspension flow. That is, the suspension's viscosity increases. If the particle surfaces are electrically charged and/or have adsorbed surfactant or polymer molecules, suspension viscosity increases even more. This can have a great impact on suspension stability since suspended particles will sediment out more slowly as the viscosity increases. *See* VISCOSITY.

Another consequence of dispersed particles encountering each other in suspension is that some of the particles will begin to aggregate. Clay mineral particles are of special interest because they can have very different particle dimensions. Their ratio of particle diameter to thickness can range from 10:1 to 250:1, making them appear to resemble sheets of paper more than spheres or polyhedra. For such particles, flocculation produces aggregates of particles that are oriented edge-to-edge and/or edge-to-face, in a "house of cards" fashion. Such particle structuring can increase to the point where essentially all particles in the suspension are networked and the suspension gels. The particular ways in which particles interact with each other in concentrated suspensions can cause some unusual properties. Some suspensions, like ketchup, flow more easily the faster they flow (pseudoplastic). Other suspensions, like the silicone-based polymer Silly Putty®, actually flow less easily the faster they are made to flow (dilatant). Still other suspensions, like grease and toothpaste, do not flow at all until a sufficiently high shearing force is applied (yield stress). *See* CLAY MINERALS; GEL; RHEOLOGY.

Characterization. Suspensions may be characterized in a number of ways.

1. The nature of the suspension (that is, whether the particles are dispersed in water, an aqueous solution, or an oil) can often be quickly determined by electrical conductivity.

2. Particle size and particle size distribution can be determined by techniques such as optical or electron microscopy and light scattering. Large-sized particles ($>5\ \mu$m) can be evaluated using sieves.

3. The nature of the solids themselves is usually determined by separating them out, followed by mineralogical analysis using a technique such as x-ray diffraction. A variety of spectroscopic techniques can be used to study the nature and properties of the particle surfaces.

4. Stability of a suspension against sedimentation is usually determined using standardized bottle or column tests, in which suspension samples are allowed to stand for a specified time with or without centrifuging and then are examined. *See* LIGHT-SCATTERING TECHNIQUES; X-RAY DIFFRACTION.

Occurrence. Suspensions of colloid-sized particles are important because they feature prominently in both desirable and undesirable contexts and in a wide variety of practical disciplines, products, and industrial processes (see table). The problems associated with suspensions are usually interdisciplinary in nature, and a broad scientific base is required to understand them completely.

Suspensions commonly occur in the environment. Examples range from the suspended sediments in rivers, lakes, and oceans to bogs and quicksand. Suspensions are commonly used in many industries, such as pulp and papermaking. They are created at an early stage of processes used to separate valuable minerals by froth flotation. Suspensions are also quite common in the petroleum industry, and may occur in reservoirs, drilling fluids, production fluids, process plant streams, and tailings ponds. In these kinds of applications, it is probably the norm rather than the exception for suspensions to contain not just solid particles and water but also emulsified oil and even dispersed gas bubbles.

Some familiar suspensions include those occurring in foods like batters, puddings, and sauces, in pharmaceuticals like cough syrups and laxatives, and in household and industrial products like inks, paints, and "liquid" waxes. Examples in personal care products include lipstick and lip balms, which are usually concentrated suspensions of solid oils in a liquid oil or in a mixture of liquid oils. The properties of such lipsticks can be adjusted by dispersing additional kinds of particles into them, such as adding solid titanium dioxide or zinc oxide to provide sunscreening, or by adding bismuth oxychloride (synthetic pearl) to create a frosted effect, or by adding mica or silica particles to create shimmer or glitter effects.

Suspension destabilization. If a suspension is stabilized by having electric charges at particle surfaces, a simple way to destabilize the suspension is to reduce the electric repulsion forces caused by those charges. This can be done in one of two ways. First, simply adding a soluble salt will produce a screening effect that reduces the strength and penetrating power of the electric repulsion forces, allowing the particles to aggregate and settle. Although almost any soluble salt will do, the most effective charge screening will be produced by salts having multicharged (multivalent) ions that are of opposite charge to that of the particle surfaces. For example, suppose a suspension of negatively charged particles can be made to aggregate by the addition of 1 mole/liter sodium chloride (NaCl). The same state of aggregation could be produced by the addition of only

about 0.018 mole/L calcium chloride ($CaCl_2$) or an even smaller amount, about 0.002 mole/L aluminum chloride ($AlCl_3$). Second, the particle surface charges can be directly counteracted by adding species of opposite charge that can adsorb directly onto the particle surfaces. For negatively charged particle surfaces, such species could be heavy-metal ions (such as copper ions) or positively charged surfactants or polymers.

Just as the stability of a suspension can be enhanced by adding polymers that adsorb onto particle surfaces (causing steric stabilization; see above), suspension stability can also be reduced by the adding large polymers that can act as bridging agents by adsorbing onto more than one particle (bridging flocculation). The bridging action causes the formation of large clusters of particles, called flocs, which can rapidly settle out and then be filtered out easily. Selecting an appropriate polymer can involve considerable testing due to the large number of factors than can influence suspension stability. Bridging flocculation also requires addition of just the right amount of polymer, since too little polymer will be ineffective and too great an addition can cause the polymers to adsorb onto individual particles instead of bridging across several particles (this can even make the suspension more stable than it was before treatment).

Once a suspension is destabilized, the particles can usually be easily separated by sedimentation, centrifugation, or filtration. A combination of simple salt addition (such as alum) for charge screening and polymer addition for bridging, followed by settling and filtration, forms the basis of many municipal water-treatment plant processes. *See* SURFACE AND INTERFACIAL CHEMISTRY; WATER SUPPLY ENGINEERING.

Laurier L. Schramm

Bibliography. R. K. Iler, *The Chemistry of Silica: Solubility, Polymerization, Colloid and Surface Properties and Biochemistry*, Wiley, 1979; L. L. Schramm, *Emulsions, Foams, and Suspensions: Fundamentals and Applications*, Wiley-VCH, Weinheim, 2005; L. L. Schramm (ed.), *Suspensions: Fundamentals and Applications in the Petroleum Industry*, American Chemical Society, 1996; H. van Olphen, *An Introduction to Clay Colloid Chemistry*, 2d ed., Wiley-Interscience, 1977; S. Yariv and H. Cross, *Geochemistry of Colloid Systems for Earth Scientists*, Springer-Verlag, Berlin, 1979.

Swamp, marsh, and bog

Wet flatlands, where mesophytic vegetation is areally more important than open water, are commonly developed in filled lakes, glacial pits and potholes (see **illus.**), or poorly drained coastal plains or floodplains. Swamp is a term usually applied to a wetland where trees and shrubs are an important part of the vegetative association, and bog implies lack of solid foundation. Some bogs consist of a thick zone of vegetation floating on water.

Unique plant associations characterize wetlands in various climates and exhibit marked zonation characteristics around the edge in response to different thicknesses of the saturated zone above the firm base of soil material. Coastal marshes covered with vegetation adapted to saline water are common on all continents. Presumably many of these had their origin in recent inundation due to post-Pleistocene rise in sea level.

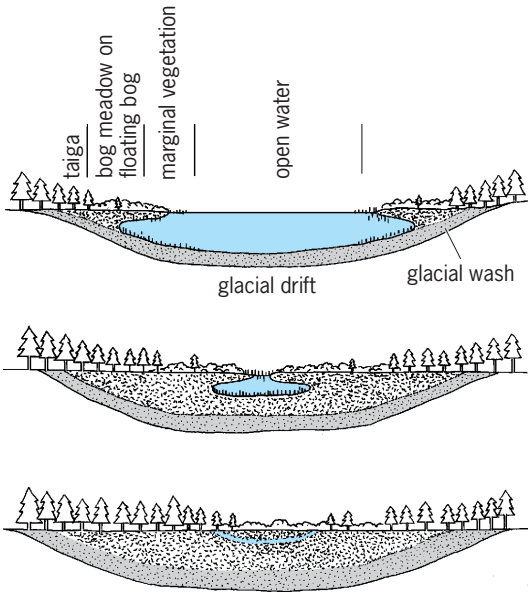

Cross-sectional diagram representing the progressive filling by vegetation of a pit lake in recently glaciated terrain.

The total area covered by these physiographic features is not accurately known, but particularly in glaciated regions many hundreds of square miles are covered by marsh. *See* GLACIATED TERRAIN; MANGROVE.

Luna B. Leopold

Sweat gland

A coiled, tubular gland found in mammals. There are two kinds, merocrine (or eccrine) and apocrine. The latter are generally associated with hair follicles (see **illus.**). Merocrine glands are distributed extensively over the body in the human, whereas the apocrine variety is restricted to the scalp, nipples, axilla, external auditory meatus, external genitals, and perianal areas. Apocrine sweat glands are more numerous in mammals, with the exception of the chimpanzee and human, in which the merocrine variety predominates. The mammary glands probably represent modified apocrine sweat glands which grow inward and increase in complexity. In association with adipose tissue, they eventually form pendant structures, the mammae, which project outward from the general contour of the skin's surface. The secretion process is apocrine with a considerable portion of the cell being discharged. The discharged portion of such a gland cell disintegrates to free fat droplets and albuminous substances. A mammary gland is complex and represents an association of lobes. Each

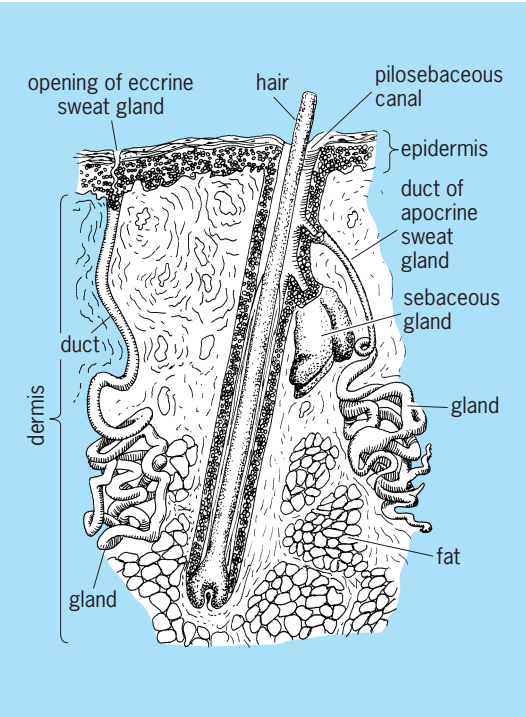

Human skin showing structure of both eccrine and apocrine sweat glands.

lobe contains a compound alveolar (acinous) gland with a separate lactiferous duct which opens on the nipple in the human. The glands of Moll associated with the eyelashes are relatively large modified apocrine glands as are the ceruminous or wax glands in the external auditory meatus. The anal sacs of the skunk presumably are apocrine glands modified by the addition of muscle fibers from the levator ani muscle which enables the pungent contents to be ejected with force. *See* EPITHELIUM; GLAND; LACTATION; MAMMARY GLAND.

Olin E. Nelsen

Sweetgum

The tree *Liquidambar styraciflua*, also called redgum, a deciduous tree of the southeastern United States. It is found northward as far as southwestern Connecticut, and also grows in Central America. The tree is commonly 80 to 120 ft (24 to 37 m) in height and $1^1/_2$ to 3 ft (0.45 to 0.9 m) in diameter, but individual trees may exceed these dimensions. Sweetgum is readily distinguished by its five-lobed, or star-shaped, leaves (see **illus.**) and by the corky wings or ridges usually developed on the twigs. The erect trunk is a dark gray, but the branches are lighter in color. In winter the persistent, spiny seedballs are an excellent diagnostic feature. The strong, close-grained wood is light brown tinged with red, and has a satiny luster and attractive grain. Because of its tendency to warp, it was long considered to be of inferior quality, but technical processing has largely overcome these difficulties. It is sometimes marketed as satin or Circassian walnut, or as hazelwood. *See* HAMAMELIDALES.

The annual commercial production of sweetgum comes mainly from the South Atlantic and Gulf Coastal Plains and from the lower Mississippi Valley. It is used for furniture, interior trim, railroad ties, cigar boxes, crates, flooring, barrels, woodenware,

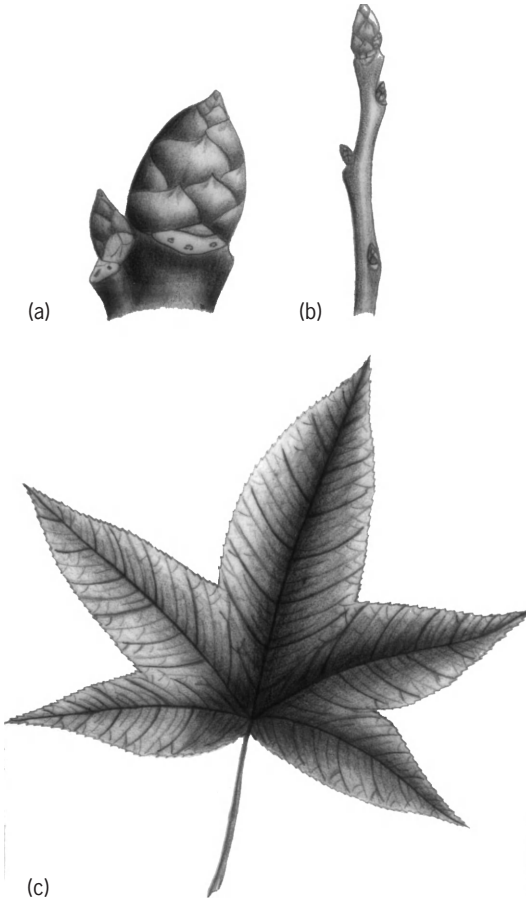

Sweetgum (*Liquidambar styraciflua*): (*a*) terminal bud, (*b*) twig, and (*c*) the distinctive five-lobed leaf.

and wood pulp, and it is one of the most important materials for plywood manufacture. Sweetgum is one of the most desirable ornamental trees, chiefly because of its brilliant autumn coloration. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Swim bladder

A gas-filled sac found in the body cavities of most bony fishes (Osteichthyes). The swim bladder has various functions in different fishes, acting as a float which gives the fish buoyancy, as a lung, as a hearing aid, and as a sound-producing organ. In many fishes it serves two or three of these functions, and in the African and Asiatic knifefishes (Notopteridae) it may serve all four. The swim bladder contains the same gases that make up air, but often in different proportions. *See* OSTEICHTHYES.

Respiratory function. The most ancient bony fish probably had a single lung which gave rise in the course of evolution to the swim bladders of modern fishes. The lung is retained as such in some primitive modern bony fishes, such as the bowfin (*Amia*), the garpike (*Lepisosteus*), and the lungfishes. The African and South American lungfishes (*Protopterus* and *Lepidosiren*) have paired lungs. Fish with lungs use them to breathe air in essentially the same way as land-living vertebrates, though they use their gills to obtain oxygen from water as well. Most teleost swim bladders have lost the respiratory function, but a few retain it or have regained it. *See* RESPIRATORY SYSTEM.

Buoyancy. The gas in the swim bladder naturally makes a fish more buoyant, whether or not its main function is to give buoyancy. Most fish without swim bladders are a good deal denser than water, but many with swim bladders are almost exactly the same density as the water they live in. Most fish need swim bladders occupying about 7% of the whole volume of the body to make them the same density as freshwater. Rather smaller swim bladders suffice in seawater, because it is denser.

The main advantage of being the same density as the water is that it saves energy. It enables the fish to hover almost motionless in midwater, whereas a fish which is denser than water cannot rest in midwater but must either keep swimming or allow itself to sink. Fish which are denser than water also need more energy to swim at a given speed than similar fish of the same density as water. This is because they need hydrodynamic lift to support themselves in the water (they obtain this lift by holding their fins extended like airplane wings or by similar means). This increases the resistance to their progress through the water, because hydrodynamic lift cannot be obtained without induced drag.

It has been calculated that a fish which had no swim bladder and which swam most of the time would use up to 20% more energy than a similar fish with similar habits with a swim bladder which made it the same density as the water. This puts the fish without a swim bladder at a severe disadvantage, and it is easy to see why most teleosts which spend much time swimming have swim bladders. It is hard to see why some mackerels and tunas do not have swim bladders, although they swim all the time.

Many teleosts which live and feed on and near the bottom have rudimentary swim bladders or no swim bladders at all. Flatfishes (Pleuronectiformes), for instance, have no swim bladder. A fish with the same density as water has no weight in water, so there is no friction to keep it in position when it rests on the bottom. This may outweigh the advantages of having a swim bladder.

Secretion and resorption of gases. Pressure under water increases with depth, so a fish must increase the quantity of gas in its swim bladder as it swims deeper if it is to keep the volume of the swim bladder and the density of the body constant. Similarly, a fish must remove gas from its swim bladder when it swims nearer the surface if its density is not to change. The more primitive teleosts such as the trout (*Salmo*, **Fig. 1**) retain the tube connecting the swim bladder to the gut which served as a windpipe for the lung of their ancestors. They can spit gas out of their swim bladders whenever they have to but obviously can take in air only through this tube when they are at the surface. They also can add gas slowly to their swim bladders by releasing dissolved gases from their blood. More advanced teleosts, such as the perch (*Perca*), have lost the tube and can only change the quantity of gas in their swim bladders by secreting gases from the blood or resorbing them into the blood. These are slow processes.

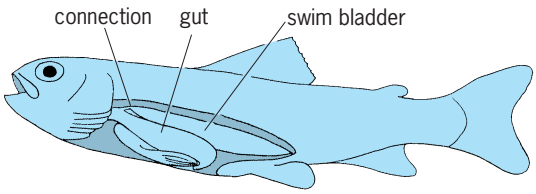

Fig. 1. Young trout (*Salmo trutta*), showing the swim bladder. The trout is one of the primitive teleosts which have the swim bladder connected to the gut.

Most fish which have been investigated take 4 to 48 h to refill their swim bladders by secretion after the gas has been removed experimentally. Secretion and resorption can therefore only keep pace with very slow changes of depth. The oceanic lantern fishes (Myctophidae), which swim hundreds of meters up toward the surface at dusk and down again at dawn, cannot secrete and resorb gases fast enough to keep the volumes of their swim bladders constant. Herrings (*Clupea*) and other fishes which make more modest daily changes of depth probably cannot compensate even for them.

Even if a fish makes no adjustments for depth changes, it still needs to be able to add gas to its swim bladder. This is because gases dissolved in the water tend to be in equilibrium with the atmosphere, so that their total partial pressure is about 1 atm (10^2 kilopascals). Gases in the swim bladder are always above atmospheric pressure, so long as the fish is submerged, because of the head of water above the fish. Therefore, gas tends to diffuse from the swim bladder to the water, mainly via the blood. The swim bladder would get gradually smaller if the lost gases were not replaced.

Gas gland. Many teleosts have a conspicuous red spot on the inside of the swim bladder wall from which gas is secreted. It is known as the gas gland. The bundle of blood vessels running to it is known as the rete mirabile. In the rete, capillaries carrying blood to the gland and capillaries carrying blood from it run parallel, mixed together (**Fig. 2**).

When the blood reaches the gas gland, lactic acid is secreted into it. This raises the partial pressure of the dissolved gases, especially of the oxygen. Therefore, when the blood returns to the rete, some of the oxygen and other gases diffuse across to the arterial capillaries instead of being carried away to the veins.

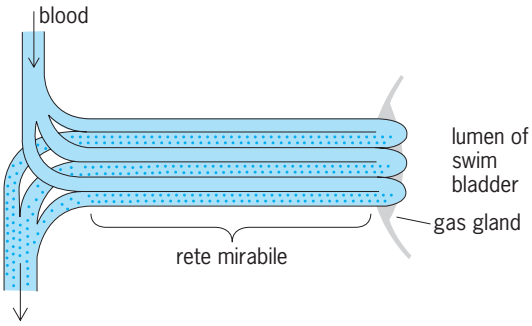

Fig. 2. Diagram showing the blood supply to the gas gland. Blood traveling away from the gland is stippled. There are far more capillaries in the rete than are shown in this diagram.

Thus high concentrations of gases get built up at the gas gland end of the rete. When the partial pressure of a gas in the blood comes to exceed its partial pressure in the swim bladder, it diffuses into the swim bladder. This theory of the mechanism of secretion was developed by Werner Kuhn and colleagues at the University of Basel. It has been largely confirmed by experiments on eels (*Anguilla*), though some doubt remains. Kuhn's calculations show that the theory provides a plausible explanation even for the secretion of gases at great depths in the sea, where the partial pressures of gases in the blood must be raised to enormous values before they will escape into the swim bladder.

Oval. Many teleosts have a specialized organ for resorbing gases from the swim bladder, as well as a gas gland for secreting them. The commonest type of resorbent organ is the oval. This is a pocket in the swim bladder wall, richly supplied with blood vessels. It can be closed by a ring of muscle around its opening. Resorption is a passive process because the pressure of the gases in the swim bladder always exceeds the total partial pressure of the dissolved gases in the blood. So long as the oval is open, gases from the swim bladder will diffuse into the blood in its blood vessels.

Sound production and reception. A swim bladder can vibrate, alternately swelling and contracting. Any disturbance will set it vibrating at its resonant frequency, just as a blow sets a gong vibrating at its resonant frequency. It can therefore be used to produce sounds. Various teleosts, including such different species as the toadfish (*Opsanus*) and the piranha (*Serrasalmus*), have special muscles which set their swim bladders vibrating and produce grunting or hooting noises. These sounds are probably used in courtship and in aggressive behavior. *See* UNDERWATER SOUND.

Sound vibrations are amplified near a swim bladder because gas is more compressible than water and responds with bigger volume changes to fluctuations of pressure. In some teleosts, such as herrings, an extension of the swim bladder reaches forward to the ears, giving them the full benefit of the amplifying effect of the swim bladder. In the minnows, catfishes, and their relatives, the same effect is achieved with tiny bones, the Weberian ossicles, which connect the swim bladder to the ear. The degree of amplification must depend on how near the frequency of the sound is to the resonant frequency of the swim bladder, but it has been calculated that sound at the resonant frequency may be amplified as much as a hundred times. It has been shown by experiment that catfish (*Ictalurus*) can hear much better than fish without Weberian ossicles and that their good hearing is spoiled if the swim bladder is deflated. *See* PHONORECEPTION.

R. McNeill Alexander

Bibliography. R. McN. Alexander, *Function Design in Fishes*, 3d ed., 1974.

Swine production

The science of breeding, raising, and marketing swine. This agricultural enterprise is usually located in proximity to sources of high-energy feedstuffs; in the United States, the geographical distribution of swine production is closely related to maize and sorghum production. Swine, as nonruminants, utilize large quantities of concentrate feeds; it is estimated that 30% of the maize grain produced is fed to swine. Swine can utilize only limited quantities of roughage in their diet.

The United States produces about one-fourth of the world's hogs. The swine production industry has changed from the status of a supplement to farm income to that of intensified units in which swine production is a major enterprise. The total number of swine producers decreased by 80% from 1950 to 1980, but the volume per unit is increasing sufficiently to maintain the total volume of production. More than 90% of all hogs are from independent producers. The others are produced by the intensified methods of integrated enterprises in confinement facilities which require highly trained biological managers capable of applying the principles of breeding, nutrition, physiology, environmental control, and economics.

Swine products. The primary products of swine are pork, lard, hides, and innumerable pharmaceutical by-products. Pork and lard supply about 15% of the total calories consumed as food in the United States. Pork is more successfully cured and stored than many other meats, and it is estimated that about 60% of the swine carcass is cured by various methods.

Breeding. There are at least 85 recognized swine breeds in the world, and perhaps more than 220 "varieties" recognized as breeds in certain geographical locations. Breeders of purebred, hybrid, and crossbred swine provide the seed stock for commercial hog production, and are entrusted with the responsibility of changing the carcass so that it better meets the preferences of consumers. The pure breeds of major importance in the United States include Berkshire, Chester White, Duroc, Hampshire, Landrace, Poland China, spotted Poland China, and Yorkshire. An example of a meat-type market-weight pig is shown in the **illustration**. In Great Britain and Europe, the Landrace and Yorkshire breeds are most prevalent.

Most commercial producers of swine practice a system of crossbreeding, mating individuals from two, three, or four different breeds, since crossbreeding increases the vigor and feeding efficiency of the offspring. Crossbred dams mated to a boar of a third breed farrow and wean larger and heavier litters than purebred dams. In practical swine operations, rotation breeding is accomplished by using male "seed" stock from different breeds and retaining female stock sired by the boar. *See* BREEDING (ANIMAL).

Reproduction. American and European breeds reach sexual maturity between the ages of 4 and 7 months, though males usually exhibit sexual maturity at a slightly earlier age than females. Boars are not usually used for breeding until they are at least $7^1/_2$ months of age. Females are usually 7–8 months of age and weigh at least 220 lb (100 kg). Some native breeds and varieties, such as those in China, may reach sexual maturity by $3^1/_2$ months of age.

Estrus, or the time of sexual excitement in the sow, lasts 48–72 h and occurs approximately every 21 days with ovulation occurring in the latter part of the period. Female swine shed an average of 15–18 ova at each estrous period, although only 60–70% survive the gestation period of 114 days. With gilts (swine that have yet to produce a litter) the number of ova ovulated may be increased by feeding them a high-energy ration for 2–3 weeks before breeding. Once a gilt is bred, however, embryonic mortality is reduced by limiting energy intake during gestation. Maternal weight gain of 70 lb (32 kg) for gilts and 44 lb (20 kg) for sows (post–first litter females) during gestation is sufficient to produce a healthy litter of pigs and to form a nutrient reserve for lactation.

Litter size is affected by many factors, and averages about 9.8 pigs per sow in the United States breeds. Certain breeds in China and other Asian countries may average 14–16. The number of pigs per litter at birth usually increases up to the third litter, or when the sow is $2^1/_2$–3 years of age.

The majority of the sows in the United States are bred by natural service. Use of artificial insemination is increasing, however, and allows wider distribution of genetic material to the industry. Embryo transfer is especially effective for introducing new genetic material into a herd and eliminates the risks of possible disease introduction when intact animals are brought into a herd.

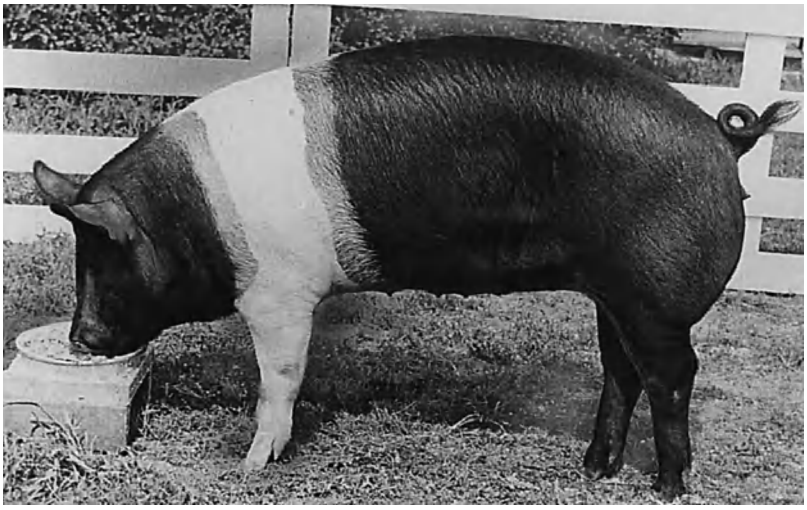

Typical example of a modern meat-type hog.

Nutrition. Cost of feed represents 50–75% of the total cost of production in commercial swine units. Diets used in feeding programs in modern swine production units are scientifically formulated to ensure nutrition that is adequate for maximum rate and efficiency of weight gain and for quality of product. In highly mechanized units, mixing and distribution to feeders of diets are computer-controlled.

The most critical nutritional period is during the first few weeks of the pig's life. Needs during the other periods of the life cycle—gestation, lactation, growing, and finishing—are easily met by fortifying natural feedstuffs. Water, energy (derived from carbohydrates, fats, and proteins), amino acids (protein), minerals, and vitamins are essential.

Water serves as a medium for digestion, absorption, transportation, and excretion of other nutrients. Dietary requirements decrease from about 12% of the liveweight for young swine to 5% for mature swine (about 2 kg of water per kilogram of feed consumed).

Most of the energy need is supplied by dietary carbohydrate. Fat usually plays a minor role, but supplemental fat is frequently in diets for very young pigs and for gestating (day 100 to parturition) and lactating sows. In general, carbohydrates such as starch, dextrins, disaccharides (except lactose), and monosaccharides are equally metabolizable by a weanling pig. Lactose is an excellent energy source

TABLE 1. Energy needs of swine*

Live weight, lb (kg)	Daily gain, lb (kg)	Daily feed, lb (kg)	Daily metabolizable energy, kcal (kJ)
2–11 (1–5)	0.44 (0.20)	0.55 (0.25)	900 (3800)
11–22 (5–10)	0.66 (0.30)	1.10 (0.50)	1700 (7100)
22–44 (10–20)	1.10 (0.50)	2.20 (1.00)	3160 (13,200)
44–77 (20–35)	1.32 (0.60)	3.30 (1.50)	4740 (19,800)
77–132 (35–60)	1.54 (0.70)	4.40 (2.00)	6320 (25,400)
132–220 (60–100)	1.76 (0.80)	6.60 (3.00)	9480 (39,700)

*Daily needs will vary due to environmental and management effects.

TABLE 2. Essential amino acids for growing-finishing swine

Amino acid	Content, %	
	Maize	Soybean meal*
Arginine†	0.52	3.40
Histidine	0.19	1.10
Isoleucine	0.37	2.50
Leucine	1.00	3.40
Lysine	0.22	2.90
Methionine‡	0.17	0.64
Phenylalanine§	0.44	2.20
Threonine	0.34	1.70
Tryptophan	0.09	0.70
Valine	0.42	2.40

*Remainder after the oil removed from soybeans by the solvent extraction method.
†Is not essential for satisfactory reproduction.
‡Cystine can supply up to 56% of the pig's methionine requirement.
§Tyrosine can satisfy 50% of the requirement for phenylalanine.

for the baby pig, but may cause diarrhea in older pigs because of a developmental shift with increasing age in intestinal enzyme activities. Energy requirements are usually expressed as a daily need for total feed or digestible nutrients. Feed intake increases as nutrient requirements for maintenance and weight gain increase (**Table 1**).

Amino acids required for maintenance, growth, gestation, and lactation, are chemically linked as proteins in feedstuffs. An essential amino acid is one that swine cannot synthesize at a sufficiently rapid rate to permit normal growth and activity. Ten of the amino acids in feedstuffs normally fed are classified as essential for growing pigs (**Table 2**). Lysine and tryptophan are of most practical importance since most cereal grains (maize, sorghum, wheat) have low levels. Soybean meal is the most common dietary supplement used to correct the amino acid deficiencies in cereal grains.

Weanling pigs have special dietary requirements (**Table 3**). Cereal grains and soybean meal lack adequate quantities of vitamins and minerals; hence, practical diets normally are supplemented accordingly.

Calcium, phosphorus, sodium, and chloride are normally added as ground limestone, dicalcium phosphate or defluorinated rock phosphate, and trace mineralized salt. The salt serves as a carrier for trace minerals, including iron, copper, manganese, selenium, iodine, and zinc. The trace minerals may not always be dietary essentials since they may be obtained from extradietary sources. Supplemental vitamins are most economically provided as synthetically produced vitamin supplements (**Table 4**).

Growth promotants (antibiotics and other antimicrobials) are regularly used in pig diets. Usually, the younger the pig, the greater the growth rate response. The amounts used in diets for finishing pigs and gestating and lactating sows should reflect the general health status of the herd. Supplements containing antimicrobials (such as chlortetracycline, oxytetracycline, penicillin, tylosin) are usually added to provide 10–20 ppm in the diet; copper sulfate is also often used in the diet as an antimicrobial. The specific action of these feed additives has not been completely delineated, but in general, herd health has been improved and the pigs perform more nearly to their genetic potential.

Most pig diets are mixtures (Table 4), and usually are self-fed to growing-finishing pigs. This ensures a "balanced" intake of nutrients by the pig. Gestating females are limit-fed to minimize excess body weight gain. Lactating sows are usually fed liberal amounts to provide adequate nutrition for full milk production. *See* ANIMAL FEEDS.

Management. The aim of swine management is to minimize adverse environmental factors and provided maximum opportunity for survival and growth. Mortality of piglets may reach 25% or higher within the first three days after birth, usually due to crushing by the sow, enteric infections, or starvation. Modern climatically regulated farrowing quarters frequently have farrowing crates to help protect the newborn.

A newborn piglet is incapable of maintaining body temperature, and is therefore dependent on supplemental heat from the environment. Other biological management practices for the newborn piglet include clipping needle teeth, applying antiseptic to the navel, and providing iron to compensate for its severe deficiency in milk.

Since the pig has very limited ability to dissipate body heat by sweating, finishing pigs or breeding stock must be protected during periods of high temperature. Shades and water mists have been used in the past, but modern confinement buildings for

TABLE 3. Mineral and vitamin needs for weanling pigs (33–77 lb or 15–35 kg)*

Mineral	Amount	Vitamin	Amount
Calcium, %	0.60	Vitamin A, IU/lb (IU/kg)	1500 (3300)
Phosphorus, %	0.50	Vitamin D, IU/lb (IU/kg)	150 (330)
Sodium chloride, %	0.35	Vitamin E, IU/lb (IU/kg)	5 (11)
Copper, ppm	4	Vitamin K, ppm	2.2
Iron, ppm	80	Riboflavin, ppm	2.2
Iodine, ppm	0.14	Niacin, ppm	17.8
Manganese, ppm	3	Pantothenic acid, ppm	11
Zinc, ppm	60	Choline, ppm	890
Selenium, ppm	0.15	Vitamin B_{12}, ppb	15

*IU = International Unit; ppm = parts per million; ppb = parts per billion.

TABLE 4. Typical ration for growing pigs

Ingredient	Amount, %
Yellow corn	76.45
Soybean oil meal	21.00
Dicalcium phosphate	1.25
Ground limestone	0.75
Trace mineralized salt	0.35
Vitamin supplement*	0.10
Antibiotic supplement	0.10

*Supplies per kilogram of ration: 3333 IU vitamin A; 333 IU vitamin D; 22 IU vitamin E; 1.1 mg riboflavin; 5.5 mg pantothenic acid; 16.5 mg niacin; 110 mg choline; 17.8 microgram vitamin B_{12}.

finishing hogs and breeding stock minimize environmental extremes. The reduction of temperature extremes encourages more rapid and more efficient gains, particularly in finishing swine.

Diseases. Swine are particularly susceptible to diseases and parasites. Although cholera was once a major problem, it has been eliminated from the United States and some other countries. Of major concern in recent years is African swine fever, which is similar in effect to cholera, but there is as yet no adequately effective vaccine or curative. Erysipelas, dysentery, leptospirosis, atrophic rhinitis, pseudorabies, and transmissible gastroenteritis are some of the other important diseases. Several diseases can be prevented by use of appropriate vaccines or other medication, but specific control has not been realized for many diseases.

Swine are especially subject to infestations with ascarids, which are parasites primarily of the intestinal tract that impair metabolism and growth. Following ingestion of the embryonated ascarid eggs by the pig, the larvae burrow through the intestinal wall and migrate through the body causing considerable damage to the liver and lungs. Passing from the lungs to the mouth, the larvae are ingested and develop into mature ascarids in the intestine. Such parasites are particularly prominent in tropical areas.

The use of confinement facilities, particularly those with slotted floors, reduces the spread of disease and parasites and allows rapid removal of excreta from the pens, thereby reducing opportunities for contact with the many vectors of the diseases. Excellent sanitation is part of any sound biological management system.

Marketing. In the United States, commercially produced pigs should attain a market weight of 198–220 lb (90–100 kg) at $5^1/_2$–6 months of age. In European countries where restricted-feed programs are used, pigs of the same size will be considerably older. About 87% of the market pigs in the United States are barrows (male pigs that were castrated at a young age) and gilts (a female that has not borne young), and the remainder are sows, boars, and stags (males castrated after becoming sexually mature). The market hogs move from farm to the slaughter plants through terminal public markets, country buying stations, and auction markets or by direct sale to the packer.

Market grades of slaughter barrows and gilts have been developed according to a scheme which places a premium on the four major lean cuts: ham, loin, picnic, and Boston butt. In addition, market grades penalize excess fat in the carcass. Barrow and gilt carcasses are graded as U.S. No. 1, U.S. No. 2, U.S. No. 3, U.S. No. 4, and utility. The utility grade is characterized by a low degree of finish. In 1980, 96% were graded 1 and 2 while in 1968 it was 50%. Thus, the swine producers, by careful selection, breeding, feeding, and management, have changed to a type of pig that provides products more acceptable to the consumers. Aldon H. Jensen

Bibliography. S. Baxter, *Intensive Pig Production*, 1984; J. L. Howard, *Current Veterinary Therapy: Food Animal Practice*, 1981; W. G. Pond and J. H. Maner, *Swine Production and Nutrition*, 1984; R. N. Van Arsdall and K. E. Nelson, *U.S. Hog Industry*, USDA Agr. Econ. Rep. 511, 1984.

Switched capacitor circuit

A module consisting of a capacitor with two metal oxide semiconductor (MOS) switches connected as shown in **Fig. 1***a*. These elements in the module are easily realized as an integrated circuit on a silicon chip by using MOS technology. The switched capacitor module is approximately equivalent to a resistor, as shown in Fig. 1*b*. The fact that resistors are relatively difficult to implement gives the switched capacitor a great advantage in integrated-circuit applications requiring resistors. Some of the advantages are that the cost is significantly reduced,

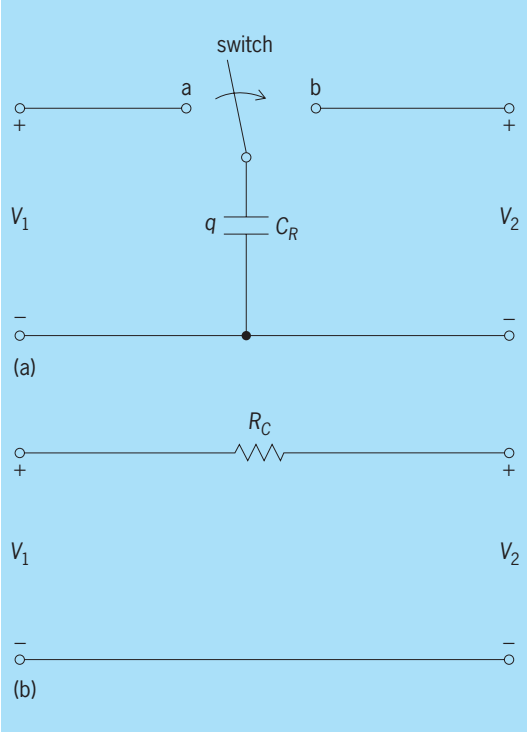

Fig. 1. Switched capacitor. (*a*) Basic circuit. (*b*) Equivalent resistive circuit.

the chip area needed is reduced, and precision is increased. Although the switched capacitor can be used for any analog circuit realization such as analog-to-digital or digital-to-analog converters, the most notable application has been to voice-frequency filtering. *See* ANALOG-TO-DIGITAL CONVERTER; DIGITAL-TO-ANALOG CONVERTER.

Although a switch has long been used as an element in circuits and systems, it was not until the late 1970s that its potential and practicality in integrated-circuit design was realized.

Integrator circuits. A conventional *RC* (resistance-capacitance) integrator circuit is shown in **Fig. 2***a*. The output voltage is given by Eq. (1), where $1/s$ indicates the operation of integration (s is the differential operator), showing that integrator performance depends on R_1. In MOS integrated circuits, the value of R_1 cannot be controlled better than 20% by using standard fabrication techniques. In addition, considerable chip space is required to realize resistors in the megohm range. In contrast, the switched capacitor realizations shown in Fig. 2*b* and *c* depend on the ratio of capacitors which can be controlled with great accuracy. For example, the output voltage of the differential integrator shown in Fig. 2*c* is given by Eq. (2), where f_c is the clock frequency. With C_1 and C_2 in the range of 1 picofarad and f_c at 100 kHz, this circuit has an equivalent resistance of 10 megohms, and the gain of the circuit is about 10^4. The silicon chip area required to implement the capacitors is about 0.01 mm^2. If a resistor is used in place of the switched capacitors, an area at least 100 times larger would be required.

$$V_{out} = \frac{1}{R_1 C_2}\frac{1}{s}V_{in} \tag{1}$$

$$V_{out} = \frac{C_1 f_c}{C_2 s}(V_2 - V_1) \tag{2}$$

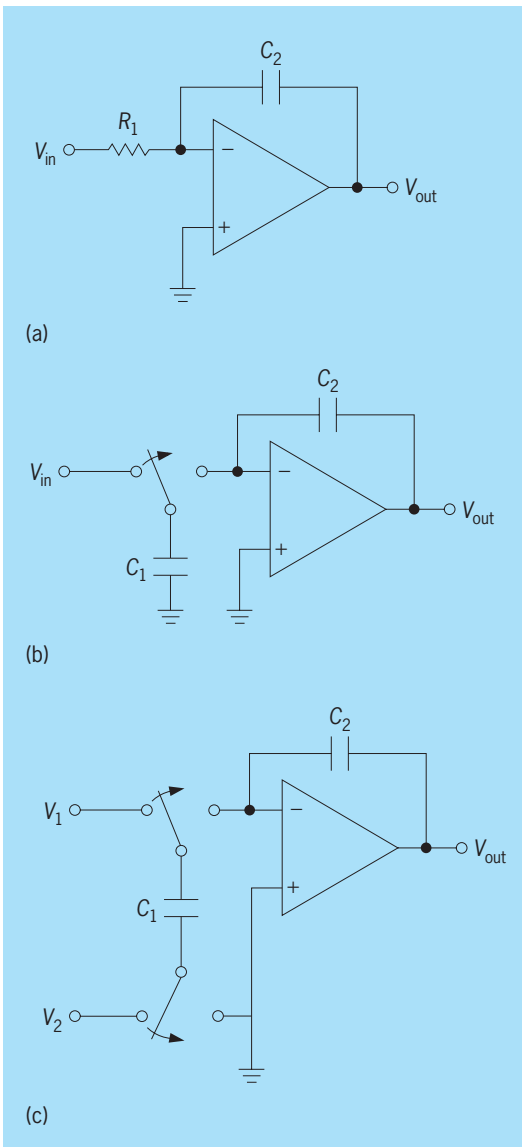

Fig. 2. Integrator circuits. (*a*) Conventional *RC* integrator circuit. (*b*) Single-input switched-capacitor integrator. (*c*) Switched-capacitor differential integrator circuit.

Equivalent resistance. Returning to the switched-capacitor circuit of Fig. 1*a*, the operation may be visualized as follows. With the switch in position *a*, the capacitor C_R is charged to the voltage V_1. The switch is then thrown to position *b*, and the capacitor discharged at voltage V_2. The amount of charge transferred is then $q = C(V_2 - V_1)$. If the switch is thrown back and forth at a clock frequency f_c, the average current will be $C(V_2 - V_1)f_c$. The size of an equivalent resistor to give the same value of current is given by Eq. (3). From this equation, it is seen that

$$R_C = \frac{1}{Cf_c} \tag{3}$$

with $C = 1$ pF and $f_c = 100$ kHz, the value of 10 megohms used previously is obtained.

The accuracy of the equivalence between the switched capacitor and the resistor depends on the relative size of the clock frequency f_c and the frequencies in the signal being processed. If the switching frequency is much larger than the signal frequencies of interest, the equivalence is excellent and the time sampling of the signal can be ignored in a first-order analysis, such that the switched capacitor is a direct replacement for a conventional resistor. If this is not the case, then sampled-data techniques in terms of a z-transform variable must be used for accuracy.

Analog operations. The switch that has been used in describing the switched capacitor is actually realized by an MOS transistor to which a pulse of voltage at the clock frequency is applied to produce the off and on conditions of the switch. This periodically operating switch is used for a number of analog operations, such as addition, subtraction, inversion, and integration. These operations are essential in the construction of analog filters, as well as in other applications of switched capacitors. These operations may be explained in terms of the circuits of Fig. 2. In Fig. 2*a* and *b* the analog operation of integration is accomplished. In addition, these circuits are of the inverting type, meaning that a sign reversal is

accomplished in addition to integration. The sign reversal of a voltage can be accomplished directly by using switches and a capacitor, as seen in Fig. 2*c*. Assume that V_1 is grounded or $V_1 = 0$. The operation of the switches is such that the voltage applied to the MOS operational amplifier is the negative of V_2. With the switch operating from left to right, V_2 with respect to ground is reversed. With V_1 not grounded, the circuit of Fig. 2*c* is a differential integrator, meaning that the output voltage is a function of the voltage difference, $V_2 - V_1$. In conventional active-filter design, these analog operations are accomplished by means of additional stages incorporating operational amplifiers. In switched-capacitor design, these analog operations are implemented with switches. *See* AMPLIFIER; ANALOG COMPUTER; TRANSISTOR.

Filter design. Although there are many strategies for filter design, the discussion will be restricted to the case of filters based on the passive *LC* (inductance-capacitance) ladder with resistive terminations at both ends. Extensive tables are available giving element values to achieve various forms of frequency response, such as Butterworth, Chebyshev, and Cauer (elliptic). All tables are given in terms of a normalized termination of 1 ohm, and a normalized frequency of $\omega_o = 1$ rad/s and for the low-pass case. It is standard procedure to make use of frequency transformations to realize high-pass, band-pass, band-elimination, and similar kinds of responses, and to use frequency and magnitude scaling to give practical element values. The passive ladder structure with double terminations is chosen because it has low sensitivity of changes in transmission with changes in element sizes.

Starting with the low-sensitivity, low-pass ladder structure, a frequency transformation is first accomplished. From these steps, a structural simulation is then carried out by replacing the actual filter by its signal flow graph representation. The flow graph is chosen so that most of the operations required are integration. The elements in the flow graph are then simulated by circuits like those shown in Fig. 2*c*.

An example of filter design is shown in **Fig. 3**. The ladder network shown in Fig. 3*a* is known as the low-pass prototype. In the usual case, $R_1 = R_2 = 1$ ohm. The elements C_1 and L_2 are determined from tables, depending on the form of frequency response required. Going from Fig. 3*a* to *b* accomplishes a low-pass to band-pass transformation in which all element values are determined from those in Fig. 3*a*, and the specification of the center frequency and bandwidth of the band-pass case. In Fig. 3*c* the filter of Fig. 3*b* is represented by its flow graph, in which the lines and arcs with arrows indicate the structure of the circuit of Fig. 3*b*, and the associated symbols represent the impedance or admittance. To this structural simulation of the ladder filter, an element simulation is next applied. All elements of a form such as $1/RC_A$s are realized by using the integrator of Fig. 2*c* with differences of voltages accomplished by the switched capacitors. The final result, shown in Fig. 3*d*, is then implemented as an integrated circuit containing only switched capacitors, ordinary capacitors, and operational amplifiers. The chip area required to realize a filter of modest order might be 100 mils (2.5 mm) on each side. *See* ELECTRIC FILTER; INTEGRATED CIRCUITS.

M. E. Van Valkenburg

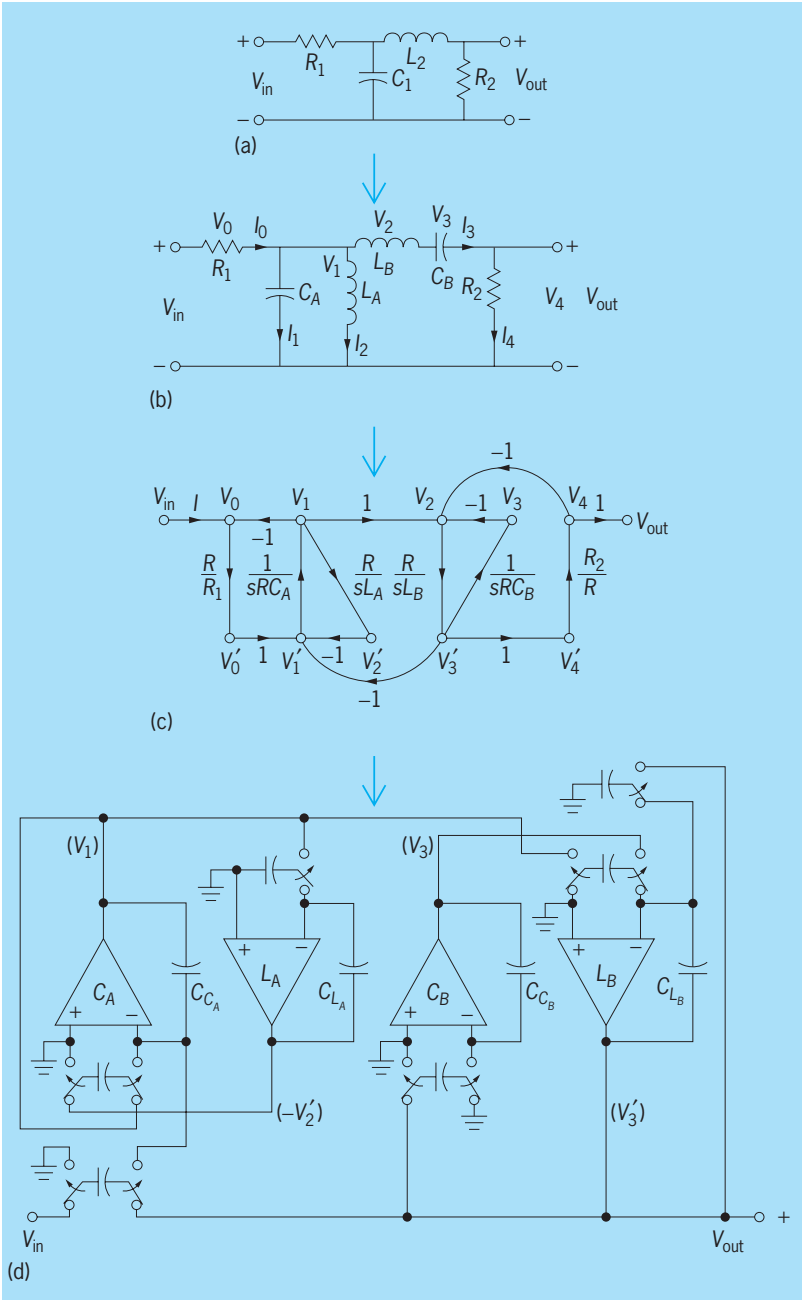

Fig. 3. Steps in the realization of a switched-capacitor filter. (*a*) Low-pass prototype filter. (*b*) Corresponding band-pass filter. (*c*) Signal flow graph representation of the circuit of *b*. (*d*) Final switched-capacitor band-pass filter.

Bibliography. P. E. Allen and E. S. Sinencio, *Switched Capacitor Circuits*, 1984; H. Baher, *Selective Linear-Phase Switched-Capacitor and Digital Filters*, 1993; M. Ghausi and K. Laker, *Modern Filter Design: Active RC and Switched Capacitor*, 1981; G. M. Moschytz (ed.), *MOS Switched-Capacitor Filters: Analysis and Design*, 1984; R. Ubehauen and A. Cichocki, *MOS Switched-Capacitor and Continuous Time Integrated Circuits and Systems*, 1989.

Switching circuit

A constituent electric circuit of a switching or digital data-processing system which receives, stores, or manipulates information in coded form to accomplish the specified objectives of the system. Examples include digital computers, dial telephone systems, and automatic accounting and inventory systems. *See* DIGITAL COMPUTER; SWITCHING SYSTEMS (COMMUNICATIONS); SWITCHING THEORY.

Physically, switching circuits consist of conducting paths interconnecting discrete-valued electrical devices. The most generally used switching circuit devices are two-valued or binary, such as switches and relays in which manual or electromagnetic actuation opens and closes electric contacts; vacuum and gas-filled electronic tubes, and semiconductor rectifiers and transistors, which do or do not conduct current; and magnetic structures, which can be saturated in either of two directions.

The electrical conditions controlling these switching circuit devices are also generally two-valued or binary, such as open versus closed path, full voltage versus no voltage, large current versus small current, and high resistance versus low resistance. Such two-valued electrical conditions, as applied to the input of a switching circuit, represent either (1) a combination of events or situations which exist or do not exist; (2) a sequence of events or situations which occur in a certain order; or (3) both combinations and sequences of events or situations. The switching circuit responds to such inputs by delivering at its output, also in two-valued terms, new information which is functionally related to the input information.

Electronic switching circuits are characterized by these functional relationships between input and output information, as well as by such attributes as response time, power consumption, packaging densities, and other performance parameters.

Functional characteristics. Functional characteristics of switching circuits are defined by the logical operation and memory capabilities of the discrete devices from which they are assembled, as well as by the means used to interconnect the devices.

For example, switching circuits embody such logical relationships as output X is to exist only if input A and B occur simultaneously; and output Y is to exist if either input A or input B occurs. The factor of memory, in turn, enables a switching circuit to hold or retain a given state after the condition that produced the state has passed.

Even in large and complex switching systems, the majority of circuit requirements can be met by a relatively small number of types of circuits, each of which performs one or a limited number of somewhat distinct functions. These functional circuits, some examples of which are described below, are the basic building blocks of a switching system.

Basic combinational circuits. A combinational switching circuit is one in which a particular set of input conditions always establishes the same output, irrespective of the history of the circuit. An example involving a simple combinational circuit is the controlling of the entrance-hall light of a residence by three up-down wall switches located in three different rooms. Analysis of this problem shows that the circuit must meet the following simple requirements. If any one or all three wall switches are down, the hall lamp must light; if any one or all three switches are up, the lamp must be dark. An obvious (but not the most efficient) circuit meeting these requirements is shown in **Fig. 1**. In this problem the circuit inputs are, of course, the manual switch settings, and the circuit output is the control of the light.

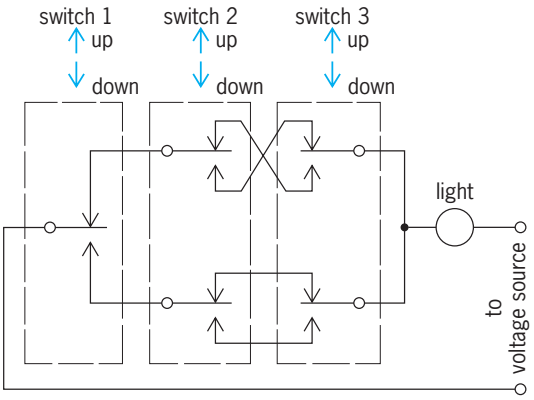

Fig. 1. Elementary combinational switching circuit.

In electronic switching circuits, so-called gates are used to perform logical functions equivalent to these series-parallel networks of switch contacts. In this sense, an electronic gate is an elementary combinational circuit. Gates do not function by physically inserting or removing metallic conduction paths between contacts of manually operated switches or remotely controlled relays. Instead, they function by control of voltage or current levels at their output.

The most commonly encountered gates are the AND and the OR gates. The AND gate produces an output only if all its inputs are concurrently present; an OR gate produces an output if any one or any combination of its inputs is present. **Figure 2** shows both an AND gate and an OR gate, using rectifier or diode elements. *See* LOGIC CIRCUITS.

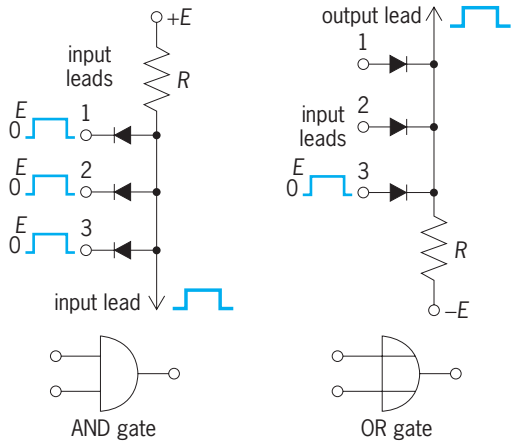

Fig. 2. Typical switching gates using crystal diodes.

In the AND gate the rectifiers are so oriented that current from a positive voltage source E passes through the relatively large resistance R and then through the low forward resistance of any one of the rectifiers to ground in the circuits controlling the gate. Thus, in the inactive state of the gate, the output lead is at or near ground potential. If all three input leads of this gate concurrently receive a positive voltage pulse of magnitude E, the rectifiers approach open circuit, and the output lead will be raised from near ground to a positive potential for the duration of the input pulse. In other words, input leads 1 and 2 and 3 must all receive the positive pulse to obtain the positive output voltage.

In the OR gate the rectifiers are reversed so that current flows from ground in the input circuits through the low forward resistance of any rectifier and then through the relatively large resistance R to the negative voltage source $-E$. Thus, in the inactive state of the gate the output lead is at or near ground potential. If, however, a relatively high positive voltage pulse is applied to input leads 1 or 2 or 3, the remaining two diodes are cut off and the output is raised to a positive potential for the duration of the input pulse.

Gates may, of course, be constructed with other electronic devices, such as tubes, transistors, and magnetic cores.

Basic sequential circuits. A sequential switching circuit is one whose output depends not only upon the present state of its input, but also on what its input conditions have been in the past. Sequential circuits, therefore, require memory elements.

By way of illustration, consider the following simple sequential circuit problem. When a telephone customer lifts the handset, a lamp is to light in front of a switchboard operator. When an operator answers, the light should go out to avoid other operators also answering. After satisfying the customer's request for a connection, the operator withdraws. The light, however, should not relight now, even though the conditions existing at this time are seemingly identical with those at the start; that is, the customer has the handset lifted and no operator is on the line. A sequential relay circuit meeting these simple requirements is shown in **Fig. 3**. In this circuit, when the hand is lifted, the handset off-hook switch connects

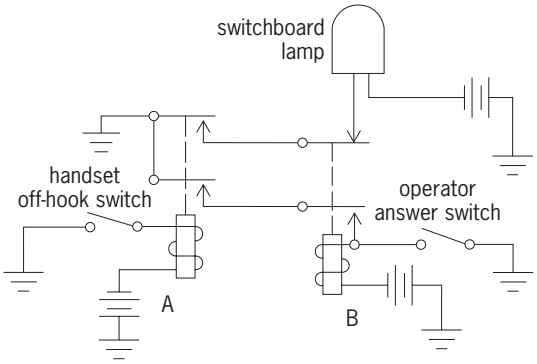

Fig. 3. Elementary sequential relay switching circuit.

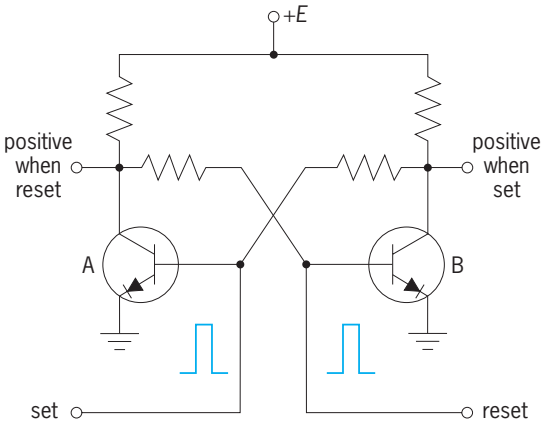

Fig. 4. Transistor switching memory element (flip-flop).

a ground input to relay A which operates and lights the switchboard lamp. When the operator answers, another ground input operates relay B and this relay puts out the light. A holding circuit on relay B keeps relay B operated until the handset off-hook switch is again opened and relay A is deenergized. Relay B "remembers" that the operator has answered and prevents the relighting of the lamp when the operator withdraws. It is, therefore, the memory element of the circuit.

A typical electronic memory element used in sequential circuits is a simple circuit called a flip-flop. A flip-flop consists of two amplifiers connected so that the output of one amplifier is the input of the other. A voltage pulse will set the flip-flop into one of two states, and that state remains until another voltage pulse resets the flip-flop or returns it to its original state. It can therefore be used to remember that an event has taken place.

Figure 4 is an *npn*-transistor flip-flop. When set, transistor A is conducting and transistor B is cut off. When reset, transistor B is conducting and transistor A is cut off. A positive output voltage with respect to ground may be obtained from either transistor to indicate the condition of the flip-flop. *See* TRANSISTOR.

Relays, flip-flops, and similar memory elements provide static, or fixed, memory; they hold the stored information indefinitely, or until they are told to "forget," commonly called "resetting." In contrast, a delay line provides transient memory. A delay line has the property that an electrical signal applied to its input is delayed on its way to the output.

Selecting circuits. A selecting circuit receives the identity (called the address) of a particular item and selects that item from among a number of similar ones. The selectable items are often represented by terminals or leads. Selection usually involves marking the specified terminal or lead by applying to it some electrical condition, such as a voltage or current pulse, or a steady-state dc signal. By means of this electrical condition, the selected circuit is alerted, sized, or controlled.

An electronic selecting circuit using AND gates is the matrix type (**Fig. 5**). In this type of circuit an input signal appears on one of the horizontal input

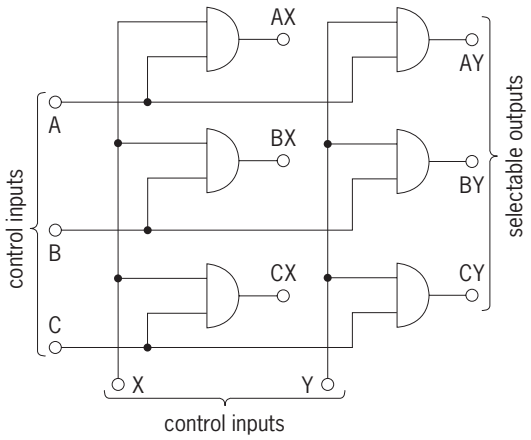

Fig. 5. Matrix selecting circuit using AND gates.

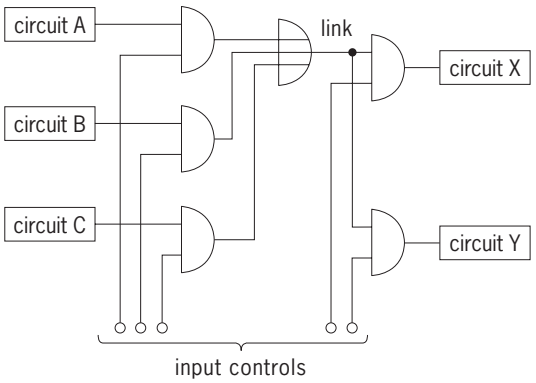

Fig. 6. Connecting circuit using AND and OR gates.

leads and concurrently on one of the vertical input leads. The selected output is at the cross point of these two leads.

Connecting circuits. A switching system is an aggregate of functional circuit units, some of which must sometimes be directly coupled to each other to interchange information. **Figure 6** shows a simple electronic connecting circuit using AND and OR gates. In this arrangement a communication path is provided over a single link from any one of the three functional circuits A, B, C, to either the X or Y circuit by an external control circuit activating the appropriate pair of AND gates. To provide a multilead link, or to provide for other simultaneous interconnections, additional AND gates would, of course, be required. The OR gate maintains separation of the inputs at the common junction point.

Lockout circuits. In switching systems, situations often arise where several similar circuit units are ready at the same instant to request collaboration with another type of functional circuit. Mutual interference among the requesting circuits is prevented by the lockout circuit (sometimes referred to as hunting or finding circuits). In response to concurrent inputs from a number of external circuits, a lockout circuit provides an output indication corresponding to one, and only one, of these circuits at any time.

Figure 7 shows a typical electronic lockout circuit using cold-cathode gas-filled tubes. The external circuits furnish positive potential on the input leads to the control gaps of the tubes as indications of service requests. The operation of the circuit is based on the dynamic negative-resistance characteristics of gas tubes. If such tubes are provided with a common impedance in their conduction paths (the cathode impedance in this circuit), simultaneous input signals will result in the ionization of only one tube. Once the control gap of a tube is ionized, conduction current starts flowing in its main gap and this current through the common impedance instantaneously reduces the voltage across all the other tubes below the value needed to ionize them. This reduced voltage is, however, adequate to keep the single ionized tube in the conducting state until its conduction path is opened. The identity of the particular ionized tube is derived from the anode resistance individual to each tube; the output lead whose potential has been lowered by this resistance represents the circuit whose request has been granted.

Translating circuits. Switching systems process information in coded form; the information is generally in the form of numbers. Numerical codes are many and varied, each with its own characteristics and more or less distinct advantages for different switching circuit situations. Therefore, one of the common functional circuits in switching systems is the translating circuit, which translates information received in one code into the same information expressed in another code. These translating circuits are combinational circuits; a given input signal combination representing a code to be translated always produces the same output signals, which represent the desired code.

Figure 8 is an example of a magnetic-core translating circuit that translates from binary code (1,2,4) to a one out of eight code (0,1,2, . . . ,6,7). The circuit has three flip-flops which are set or reset (not set) according to the binary input code combination. The translating elements are eight magnetic cores, each with five windings, and are represented in

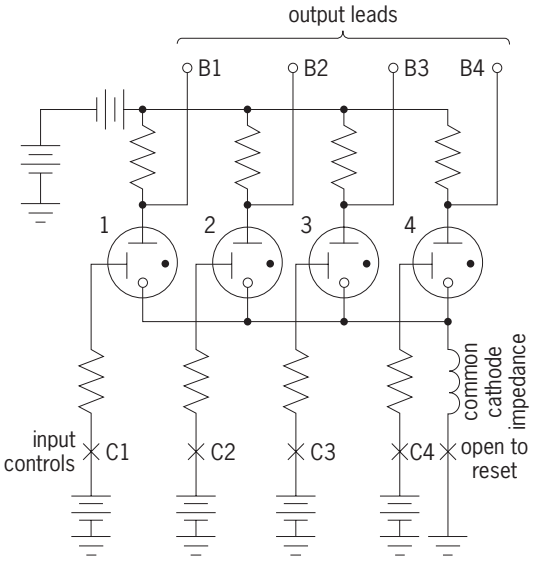

Fig. 7. Lockout circuit using cold-cathode gas tubes.

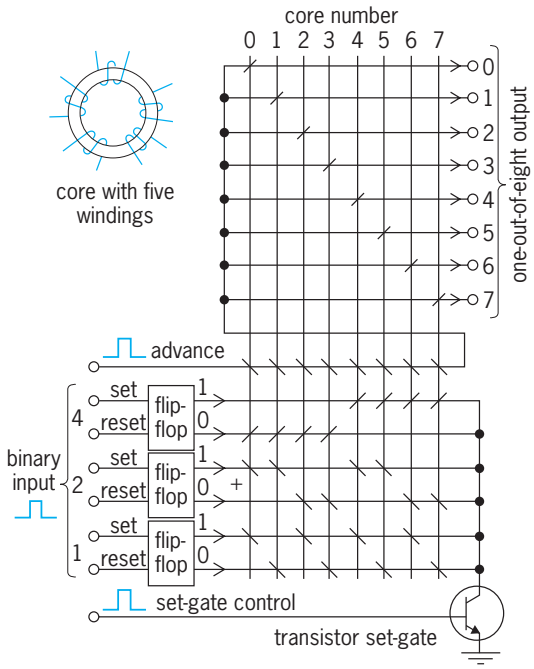

Output digit desired	Flip-flop 4		Flip-flop 2		Flip-flop 1	
	Set	Reset	Set	Reset	Set	Reset
0		x		x		x
1		x		x	x	
2		x	x			x
3		x	x		x	
4	x			x		x
5	x			x	x	
6	x		x			x
7	x		x		x	

Fig. 8. Code register and translating circuit using magnetic cores, each of which has five windings.

Fig. 8 by a vertical line. Each short, slanting line segment represents a separate winding on a core. These slanting lines also symbolize a mirror action; an input current pulse coming from a flip-flop sets a core if it is reflected upward by the mirror, and prevents setting or resets the core if reflected downward. Once set, the subsequent resetting of a core induces a current which flows upward in the vertical line (in a direction opposite to the resetting current) and is reflected to the left or to the right by each mirror symbol.

With this explanation of the symbolism, the circuit works as follows. The input is binary; that is, it consists of a positive voltage pulse to each of the three flip-flops either on its set or on its reset input lead, according to the **table**. (By adding the numerical designations to those flip-flops which are set in a particular combination, the value of the output digit is determined.)

While the flip-flops are being set, their output current is prevented from flowing into the core windings by the transistor set-gate which is normally nonconducting. Shortly after the binary input combination is recorded in the flip-flops, this set-gate is pulsed for a moment into its conducting state. During this moment, output current will flow from each flip-flop in its "1" output lead (if the flip-flop has been set) or in its "0" output lead (if the flip-flop has been reset). As Fig. 8 shows, the output current of flip-flop 4 is always used to set the cores; that is, the current in the "0" output lead of this flip-flop is used to magnetize the first four cores in the set direction, or the current in its "1" output lead is used to magnetize the last four cores in the set direction. In contrast, the output currents from flip-flops 2 and 1 are always used to magnetize the cores in the opposite or reset direction. Initially, all cores are in the reset condition, and cores that receive both set and reset currents simultaneously will not change this initial condition. An analysis of Fig. 8 will therefore show that, for any desired digit, one and only one of the eight cores will be set by the flip-flops in combination. For instance, if output 3 is desired, the current from flip-flop 4 tends to set cores 0, 1, 2, and 3, but cores 0, 1, and 2 are prevented from being set by the output current from either or both flip-flops 2 and 1. When the translated code is needed, the current pulse on the advance lead resets the single previously set core, and consequently an induced output current pulse appears on the appropriate output lead. (The rectifiers in the input and output portions of the circuit prevent unwanted reverse current.)

Register circuits. Information received by a switching system is not always used immediately. It must be stored in register circuits for future use.

In a register circuit the coded information to be stored is applied as input and retained by memory elements of the circuit, and when needed, the registered information is taken as output in the same code or in a different code. Figure 8 embodies a register function as well as a translating function. Register circuits are devised with a great variety of memory elements and have capacities to store from a few to millions of information bits.

A frequently encountered form of register circuit is the shift register. This type of register has the ability to shift its stored digital information internally to positions representing higher or lower numerical values in the code employed. For example, in decimal code registration a digit may be shifted from the units to the tens position. An obvious use of such registers is in digital computers when, for example, partial multiplication products have to be lined up for addition.

Counting circuits. One of the most frequently encountered circuits in switching systems is the counting circuit whose function, in general, is to detect and count repeated current or voltage pulses which represent incoming information.

Performance characteristics. In theory, functionally equivalent switching circuits can be designed by using any available switching technology; that is, mechanical manual or relay-operated switches, or a number of electronic devices including a wide variety of semiconductors, and vacuum or gas-filled tubes. The selection of a particular technology for use in computer systems is often determined by the time required to perform logical operations. The

Digital integrated circuit families

Logic type	Gate delay (t_d), ns	Power dissipation (P_D), mW/gate	Density, gates/mm^2 (gates/in.^{2})
I. Bipolar—use bipolar junction transistors as basic switching elements			
A. Transistor-transistor logic (TTL)			
1. Standard small-scale integration/medium-scale-integration (SSI/MSI) TTL	5–10	10	20–80 (13,000–50,000)
2. Low-power Schottky (LPS)	5–10	2	20–80 (13,000–50,000)
3. Advanced LPS for large-scale integration (LSI)*	1–2	1–2	200–600 (130,000–400,000)
B. Emitter-coupled logic (ECL)			
1. Standard SSI/MSI ECL	0.5–1	25	20–50 (13,000–32,000)
2. Advanced ECL for LSI*	0.7–1	10	100–200 (65,000–130,000)
3. Current-mode logic (CML)†	0.5–1	0.2–1	200–300 (130,000–200,000)
C. Integrated injection logic (I^2L)			
1. Standard LSI I^2L*	2–5	0.05	200–400 (130,000–250,000)
2. Integrated Schottky logic (ISL)†	2	0.3	500–1000 (320,000–650,000)
3. Schottky transistor logic (STL)†	2	0.3	500–1000 (320,000–650,000)
4. Implanted advanced composed technology (impact)†	1–2	0.1	800 (500,000)
II. Metal oxide semiconductor (MOS)—use MOS field-effect transistors (MOSFETs) as basic switching elements			
A. *P*-channel MOS (PMOS)			
Standard LSI PMOS	50	1	100–200 (65,000–130,000)
B. *N*-channel MOS (NMOS)			
1. Standard LSI NMOS	10–20	1–2	200–400 (130,000–250,000)
2. Very large-scale-integration (VLSI) NMOS	0.5–1	0.5	1500–3000 (1,000,000–2,000,000)
3. Exotic structures			
a. Vertical channel (VMOS)	0.5–1	0.5	500–1000 (320,000–650,000)
b. Double-diffused (DMOS)	0.5–1	0.5	500–1000 (320,000–650,000)
C. Complementary MOS (CMOS)			
1. Standard SSI/MSI CMOS	20–50	0.01–0.1	50–100 (32,000–65,000)
2. LSI*/VLSI CMOS	0.5–1	0.01–0.5	500–2000 (320,000–1,300,000)
3. Silicon-on-sapphire (CMOS/SOS)	0.5–1	0.01–0.5	750–2500 (500,000–1,600,000)

*Uses advanced fabrication technology with feature sizes below 5 μm.
† Feature sizes of 2.5 μm or smaller.

maximum number of logical operations per second which a computer switching circuit can execute is determined by both the time required by the switching device to change state (gate delay, t_d) and the time it takes signals to propagate between switching circuits, the latter being proportional to the distance between the circuits. Device gate delay is dependent upon the active element's principle of operation, and is normally inversely proportional to gate power dissipation (P_D) since larger current supplied to the switching devices reduces the effects of junction storage time, capacitance, and parasitic elements. The product of t_d and P_D is a popular figure of merit for comparing different device technologies.

Circuit density. One of the objectives of large-scale integrated-circuit (LSI) and very large-scale integrated-circuit (VLSI) design is to incorporate fast device technologies into compact packages which reduce, to the greatest extent possible, propagation delay between circuits. Thus, circuit density, in terms of the number of gates per square millimeter, is another important parameter affecting overall logic speed. Thin-film fabrication techniques for both active and passive components result in dense packaging with higher reliability, finer adjustment of component values, and simpler manufacturing processes. The table lists the major semiconductor switching circuit technologies, along with their acronyms and important performance characteristics. *See* INTEGRATED CIRCUITS.

Josephson junctions. Since semiconductors produce heat, dense circuit packaging is hampered by the need to remove heat. The Josephson junction avoids this problem by using superconducting circuits, that is, circuits cooled to near absolute zero where the electrical resistance that generates heat virtually vanishes. Josephson found that a superconducting current between two wires separated by an ultrathin insulating film can be controlled by a magnetic field applied to the insulator. While the heat dissipated is very small (5 nanowatts per gate), gate delays of less than 0.01 nonosecond in miniaturized devices a few micrometers in length and width can be achieved. Implementation problems, including the need to provide cryotats for low operating temperatures, have thus far precluded commercially viable products. *See* JOSEPHSON EFFECT; SUPERCONDUCTING DEVICES.

Magnetic bubble memory. Another advanced technology is based on the fact that the polarization of tiny cylindrical magnetized areas, contained in a thin film of magnetic material, can be used to represent data 1's or 0's. Such devices are called magnetic bubble memories, and when used for data storage, the principal advantages are extremely high storage density

(7000 bits/mm^2 or 4.5×10^6 bits/in.^{2} and the fact that they are nonvolatile (data are not lost when power is shut off). The memory access time of 10 milliseconds is faster than a disk drive but slower than semiconductor memories. Power consumption is extremely low, being 10 microwatts per bit during reading or writing operations. Joseph A. Pecar

Bibliography. W. Anacker et al., Special issue on Josephson technologies, *IBM J. Res. Develop.*, vol. 24, no. 2, March 1980; K. L. Chopra and I. Kaur, *Thin Film Device Application*, 1983; A. Friedman and P. R. Memon, *Theory and Design of Switching Circuits*, 1975; S. Fulton, *Logic and Switching Circuits*, 1994; M. P. Mitchell, *Switching Circuits for Engineers*, 3d ed., 1975; J. Watson, *Analog and Switching Circuit Design*, 2d ed., 1989.

Switching systems (communications)

The assemblies of switching and control devices provided so that any station in a communications system may be connected as desired with any other station. A telecommunications network consists of transmission systems, switching systems, and stations. Transmission systems carry messages from an originating station to one or more distant stations. They are engineered and installed in sufficient quantities to provide a quality of service commensurate with the cost and expected benefits. To enable the transmission facilities to be shared, stations are connected to and reached through switching system nodes that are part of most telecommunications networks. Switching systems act under built-in control to direct messages toward their ultimate destination or address.

Most switching systems, known as central or end offices in the public network and as private branch exchanges (PBXs) when applied to business needs, are used to serve stations. These switching systems are at nodes that are strategically and centrally located with respect to the community of interest of the served stations. With improvements in technology, it has become practical to distribute switching nodes closer to stations. In some cases to serve stations within a premise, switching is distributed to take place at the stations themselves. A smaller number of systems serve as tandem (intermediate) switching offices for large urban areas or toll (long-distance) offices for interurban switching. These end and intermediate office functions are sometimes combined in the same switching system. *See* PRIVATE BRANCH EXCHANGE.

There are many types of telecommunication services. The principal ones are voice, data (record), picture (still), and video (motion pictures). For each service there is a different balance between the relative investment in transmission, switching, and station (terminal) facilities. This article deals primarily with systems for the switching of voice. Some data services use the voice network. Since separate networks for data services are available, a brief discussion will be given of switching exclusively for voice and data services.

The basic form of switching for current services utilizes circuit switching where a path in space or time is allocated to each message or call. The characteristics of telephone messages are quite different from data and imaging, which are either bursty or long and one-way. For these, packet switching has been introduced, where the message are subdivided into uniform-size arrays which need not be received in real time. *See* DATA COMMUNICATIONS; INTEGRATED SERVICES DIGITAL NETWORK (ISDN); PACKET SWITCHING.

Switching System Fundamentals

Telecommunications switching systems generally perform three basic functions: they transmit signals over the connection or over separate channels to convey the identity of the called (and sometimes the calling) address (for example, the telephone number), and alert (ring) the called station; they establish connections through a switching network for conversational use during the entire call; and they process the signal information to control and supervise the establishment and disconnection of the switching network connection.

In some data or message switching when real-time communication is not needed, the switching network is replaced by a temporary memory for the storage of messages. This type of switching is known as store-and-forward switching.

Signaling and control. The control of circuit switching systems is accomplished remotely by a specific form of data communications known as signaling. Switching systems are connected with one another by telecommunication channels known as trunks. They are connected with the served stations or terminals by lines.

In some switching systems the signals for a call directly control the switching devices over the same path for which transmission is established. For most modern switching systems the signals for identifying or addressing the called station are received by a central control that processes calls on a time-shared basis. Central controls receive and interpret signals, select and establish communication paths, and prepare signals for transmission. These signals include addresses for use at succeeding nodes or for alerting (ringing) the called station.

Most electronic controls are designed to process calls not only by complex logic but also by logic tables or a program of instructions stored in bulk electronic memory. The tabular technique is known as action translator (AT). The electronic memory is now the most accepted technique and is known as stored program control (SPC). Either type of control may be distributed among the switching devices rather than residing centrally. Microprocessors on integrated circuit chips are a popular form of distributed stored program control. *See* COMPUTER STORAGE TECHNOLOGY; INTEGRATED CIRCUITS; MICROPROCESSOR.

Common channel signaling (CCS) comprises a network of separate data communication paths used for transmitting all signaling information between offices. It became practical as a result of processor

control. To reduce the number of data channels between all switching nodes, a signaling network of signal switching nodes is introduced. The switching nodes, known as signal transfer points (STPs), are fully interconnected with each other and the switching offices they serve. All links and signal transfer points are duplicated to ensure reliable operation. Each stored-program-control toll switching system connects to the two signal transfer points in its region.

Numbering plan. Telecommunications networks are bound together by a single address coding plan which provides for uniquely identifying every station and terminal. The North American system is based upon decimal digits. (When letters are on the dial, the telephone system actually recognizes the numerals associated with the letters.)

A telecommunications central office customarily has the nominal maximum capacity to serve 10,000 main stations, using the number series 0000–9999. When there are more than 10,000 main stations, more than one central office is provided, sometimes in more than one building or wire center. Each office is given a separate three-digit designation or code. The minimum requirement is for an adequate number of digits or characters in each number to address each main station in the dialing area. When a call reaches the called office, the called telephone station is determined from the last four numerals (the main station code).

A seven-digit numbering plan has adequate capacity for only a small portion of the telephones in North America. Hence, a geographical area, such as a state or a Canadian province, is selected as a numbering plan area (NPA), within which there are no duplications of seven-digit numbers. The more populous states, which have large numbers of central office codes, are divided into two or more numbering plan areas.

Each numbering plan area is given a three-digit NPA code. Examples are 803 for the state of South Carolina and 415 for the portion of California that includes San Francisco. With this plan the equipment uses, first, the NPA code to determine which area is desired; second, the central office code to select the office in that area; and third, the main telephone number to determine the particular telephone being called. To enable the reuse of area codes as central office codes, a 1 is dialed ahead of each 10-digit number.

Area and central office codes have generally represented geographical areas or exchanges for which uniform rate tariffs are enforced. Non-geographical codes are used, such as 700, 800, and 900, for universal service distinctions. Code numbers for special services, such as 411 for directory assistance and 911 for emergency calls, are used.

In the United States, long-distance services are provided by networks of several carriers (interLATA carriers, discussed below). Callers may presubscribe to a particular carrier or may precede the telephone number by a five-digit prefix code, such as 10722, that designates the carrier whose services are desired on a particular call. These are known as carrier identification codes.

When calling, digits, known as prefixes, may precede the telephone number to represent the service desired on a call. Typically, these are to request the services of an operator (0, 00), to choose a particular telecommunication company (8, 9, or 10XXX, where X equals a digit other than 0), or to call outside of North America (01, 011). To distinguish calls not requiring the selection of services, a 1 is frequently used as a prefix.

Switching fabrics. Space and time division are the two basic techniques used in establishing connections. When an individual conductor path is established through a switch for the duration of a call, the system is known as space division. When the transmitted speech signals are sampled and the samples multiplexed in time so that high-speed electronic devices may be used simultaneously by several calls, the switch is known as time division. Space-division switching has been employed since the early manual switchboards in which the connectives were cords, plugs, and jacks.

Most switching is now automatic. Operators are required only for ancillary functions that cannot yet be economically automated. Such systems use the most modern techniques. In the United States, cord, plug, and jack switchboards have almost disappeared, having been replaced by cordless consoles. Calls are distributed automatically to operators seated at cordless consoles or computer keyboards with keys and lamps to permit them to serve calls requiring judgment. The positions may be located many miles from the stored-program-control system through which the connection is established or processed.

The switching fabric frequently comprises two primary-secondary arrangements: first, the line link (LL) frames on which the telephone lines appear and, second, the trunk link (TL) frames on which the trunks appear (**Fig. 1**). A switching entity may grow to a maximum of 60 line link and 30 trunk link frames. Each line link frame is interconnected with every trunk link frame by a network of links called junctors. Each line link frame has a basic capacity for 290 telephone lines and may be supplemented in 50-line increments to a maximum of 590 lines. The

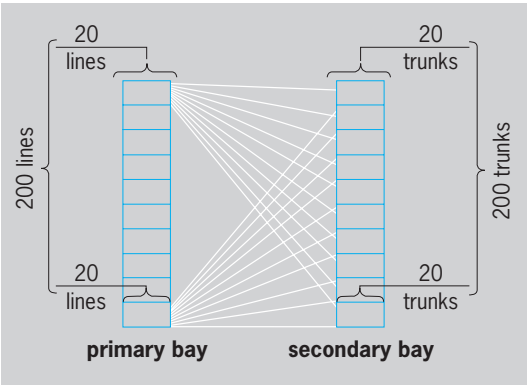

Fig. 1. Primary-secondary link arrangement.

size used in a particular office depends upon the calling rate and holding time of the assigned lines.

Systems use automatic message accounting (AMA) to make call records for billing purposes. Originally the AMA record was made on perforated tape. Now minicomputers located in the office or reached over data links record the AMA data on magnetic tape. On calls within the local area where a message-unit basis of charging applies, frequently only the calling telephone and the number of message units are recorded. For toll calls, the calling and called numbers, answering and disconnect times (from which the length of conversation is computed), and other data are recorded. Automatic data processors at accounting centers convert the recorded information into a form for printing the bill statement.

Centralized automatic message accounting (CAMA) has been applied to the offices used for tandem or toll switching functions. The calling station is identified automatically in the local office, and this number is sent along with the called number over the trunk to the tandem office. Where local offices are not arranged to identify the calling station, an operator is momentarily added to the connection to ask the calling party for his or her number and to key it into the system.

Electronic Switching

In the United States, 10,000 local and 600 toll stored-program-control switching systems have been placed in service, involving more than 120,000,000 lines (as of 1993). Stored program control has become the principal type of control for all types of new switching systems throughout the world, including toll, private branch, data, and Telex systems. About 2000 offices were earlier placed in service, principally in the United Kingdom and France, using electronic logic and memory controls based on the action translator principle (table look-up and proprietary control circuitry).

Service features. Two types of data are stored in the memories of electronic switching systems. One type is the data associated with the progress of the call, such as the dialed address of the called line. Another type, known as the translation data, contains infrequently changing information, such as the type of service subscribed to by the calling line and the information required for routing calls to called numbers. These translation data, like the program, are stored in a memory which is easily read but protected to avoid accidental erasure. This information may be readily changed, however, to meet service needs. The flexibility of a stored program also aids in the administration and maintenance of the service so that system faults may be located quickly.

The availability of large memories and the ease of changing the program residing in them has led to the development and deployment in most stored-program-control systems of many new and optional services such as abbreviated dialing, call waiting, call forwarding, and three-way calling. The wide introduction of common-channel signaling into the network broadens the range of stored-program-control service offerings. These services take advantage of the rapid exchange of information about calls between the originating, intermediate, and terminating offices before connections are established. As a result, calls may be directed within the entire network to suit the individual users and subscribers whose needs are likely to be nationwide rather than residing in a single central office or private branch exchange.

One of the most impressive results of employing electronics in switching is the space savings. Even though the early systems that were installed used mostly discrete semiconductor components, the space savings were as high as 60% when compared with electromechanical switching. Additional space savings have been achieved as integrated circuits are applied.

No. 1A electronic switching system. The high speed of electronics enables the systems to be designed so that all calls in progress are processed and supervised by the same control equipment. The No. 1A electronic switching system widely used in the United States for large local and small toll offices has a capacity as high as 345,000 local calls per hour and may serve as many as 125,000 lines (**Fig. 2**). High-speed peripheral buses connect the stored-program-control processor to the peripheral circuits of the system that access lines, trunks, and the switching network control.

The stored program control comprises a general-purpose assemblage of semiconductor circuits that are structured to interpret the instructions used in programs for the processing of the calls and for the maintenance of the system. These instructions are stored in a memory subsystem as coded programs that are read in sequences that determine actions to be taken.

Portions of the call-processing functions are used repetitively for each call. The programs are relatively fixed and remain in the system memory while the call information is stored for relatively short periods. The No. 1A ESS is provided with integrated circuit (IC) memory for call data, translation, and program storage. Integrated circuit chips are used in a random-access storage. A typical program requires more than 500,000 words of 32 bits each. All control circuits are duplicated to ensure service continuity should a component fail, or to enable the system to grow while in service. Memory redundancy is also provided. Copies of the call processing as well as additional, less frequently used maintenance, administrative, and operations (MAO) programs are stored on magnetic disks. The periphery includes data links to centralized MAO facilities and for common-channel signaling, as well as local maintenance consoles and tape drive units.

Space-division switching fabrics. Initially, most stored-program-control systems employed space-division fabrics. Many unique matrix arrays of devices, both metallic and nonmetallic, have been developed and used in commercial electronic switching systems.

For some electronically controlled switching systems, metallic contacts are used in the space-division

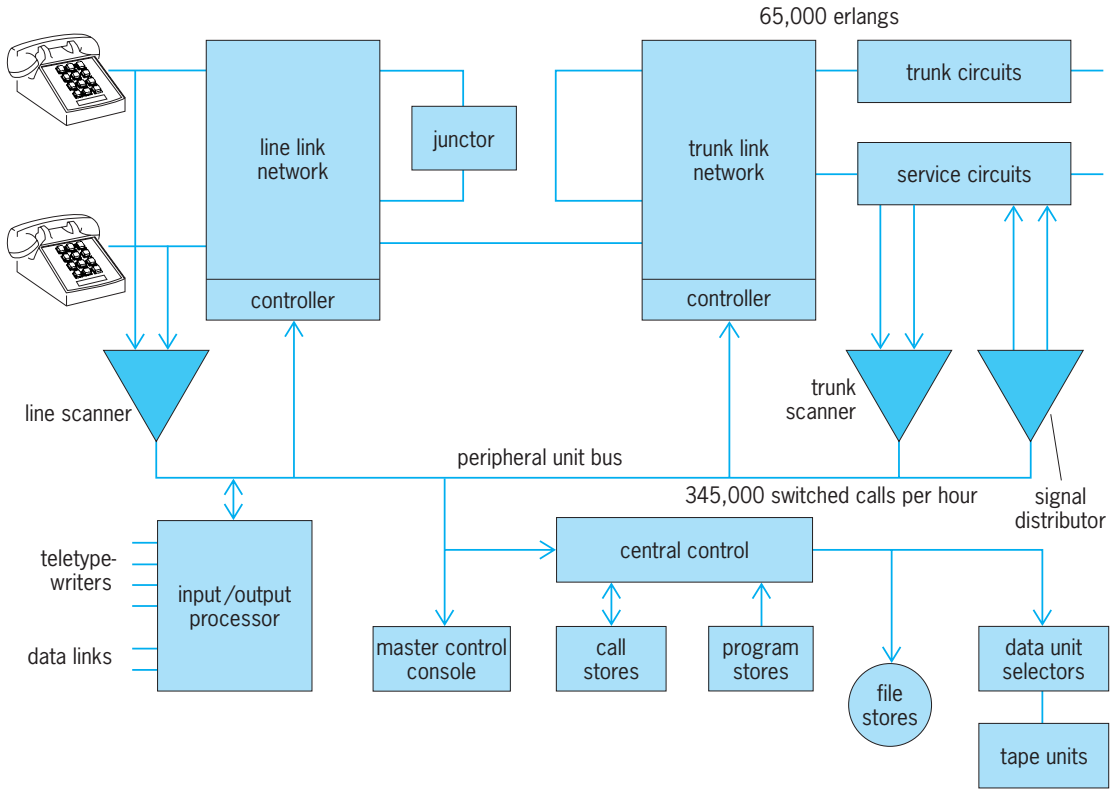

Fig. 2. Block diagram of No. 1A electronic switching system. Junctor is an intraoffice trunk circuit.

fabric. Many of these contacts are magnetic reeds sealed in glass with an inert gas. They latch magnetically when activated by a short pulse of one polarity and release with a pulse of opposite polarity. When the contacts are made of hard magnetic material, such as those employed in the No. 1A ESS, they are known as remreeds. Other systems use miniature crossbar or similar coordinate switches. In all of these space-division fabrics, the switching device is slow in comparison to the speed of the electronic controls. The network controller provides the buffering between high-speed stored program or action translator controls and the slower space division networks.

Electronics in the form of semiconductor or integrated circuit devices are used as crosspoints or gates in the electronic switching fabrics. Generally their use has been confined to smaller systems or PBXs (less than 2000 terminals) or in combination with metallic crosspoints for larger systems.

Semiconductor crosspoints may be used in fabric configurations in the same manner as reed switches. They are designed to be bistable and held actuated over the established speech path. Therefore, these devices combine both transmission and control characteristics. Each crosspoint may use a pair of bistable devices as the two conductors of the connection, or a single device with one wire and a ground return. Some modern semiconductor crosspoint devices are designed to function in the same high-energy environment as metallic devices.

Time-division switching fabrics. Time-division switching is practical only with high-speed electronic techniques. It is used in switching fabrics as well as in the control portions of systems. For time division, analog speech signals are sampled at a rate at least twice the highest frequency to be transmitted. Typically for voice this is 8000 times per second.

Pulse modulation. In a switching system the samples may be pulses of varying amplitude that are analogs to the electrical signals representing the voice. This is known as pulse-amplitude modulation (PAM).

More robust forms of pulse modulation use digital or on-off signals that can be readily encoded from the amplitude pulse, sent over a transmission medium, and periodically reformed to eliminate most of the impairments of transmission and switching. These types of digitally sampled transmission are known by the form of coding employed; the most popular is referred to as pulse-code modulation (PCM). In one type of pulse-code modulation, each sample is coded into one of approximately 256 amplitudes and represented by eight binary (on-off) pulses. The eight pulses representing each speech channel may be placed in sequence or time-multiplexed into groups, typically of 24 or 32 channels, so that the line and repeaters may be used more efficiently. *See* PULSE MODULATION.

Switching techniques. Time-multiplexed coded voice signals reaching a switching system are switched by using two techniques. Assuming signals arrive at the switching fabric on different multiplexed lines, they first need to be synchronized with respect to multiplex channel identification. This usually requires some form of time delay or buffering. This establishes uniform channel periods or time slots. Within

the switch there may be more time slots than in the lines delivering the signals.

One necessary switching fabric function uses further time buffering by placing successive digitized channel samples in a memory in one order and removing them as indicated by the switching selection requirements of each call. This function is known as time slot interchange (TSI).

The other switching fabric function provides for interchanging channels between time-multiplexed lines. Generally this function is carried out by using a high-speed space-division fabric of one or more stages. The space-division fabric control acts to change the actuated crosspoints between time slots so that the successive channels of input lines may be switched to corresponding channels in the same or other lines. Space-division stages operating at time slot rates are time-multiplexed switches (TMS).

Time-division switching fabrics are based upon the use of successive time slot interchanges, called T (for time) stages, and time-multiplexed switches, called S (for space) stages. Typical systems are said to employ T-S-T or S-T-S types of networks. Lines or trunks not reaching the switching system by time-multiplexed facilities must first be multiplexed with similar inputs as part of the time-division switching-fabric function.

Since the coded signals represent only one direction of transmission, the switching-fabric function in a circuit switch is duplicated by reciprocity to provide for the other direction of transmission. Also, since the elements of the network are active and may be used for hundreds of simultaneous calls, redundancy of the network and its controls, similar to that used in the call-processing portions of the system, is usually part of the system architecture.

Digital (PCM) toll systems. Time-division switching is a natural adjunct to digital time-division transmission where the coding is performed for purposes of multiplexing. Pulse-code-modulation transmission was initially economical on interoffice trunk routes from 10 to 50 mi (16 to 80 km). This made time-division digital switching attractive where such trunk facilities were found, namely, for tandem and toll applications.

As an example of a stored-program central time-division digital switching system developed for this application, the No. 4 ESS is the switching system of greatest capability. It has a capacity of 100,000 trunks, and its stored program control is capable of switching 670,000 calls per hour.

The switching fabric for the No. 4 ESS consists of both time-slot-interchange and time-multiplex-switching stages; the latter is reconfigured at the rate of 1,024,000 times per second (128 time slots for each frame of 8000 samples per second). A sample progresses through both the memory (time-slot-interchange) and four-space-division time-multiplexed switch stages for one direction of a typical time-slot connection, with the network control memory being read out cyclically at a rate of 8000 times per second (**Fig. 3**).

The No. 4 ESS serves three types of transmission channels: analog metallic, analog multiplex carrier, and digital multiplex carriers (**Fig. 4**). The digital

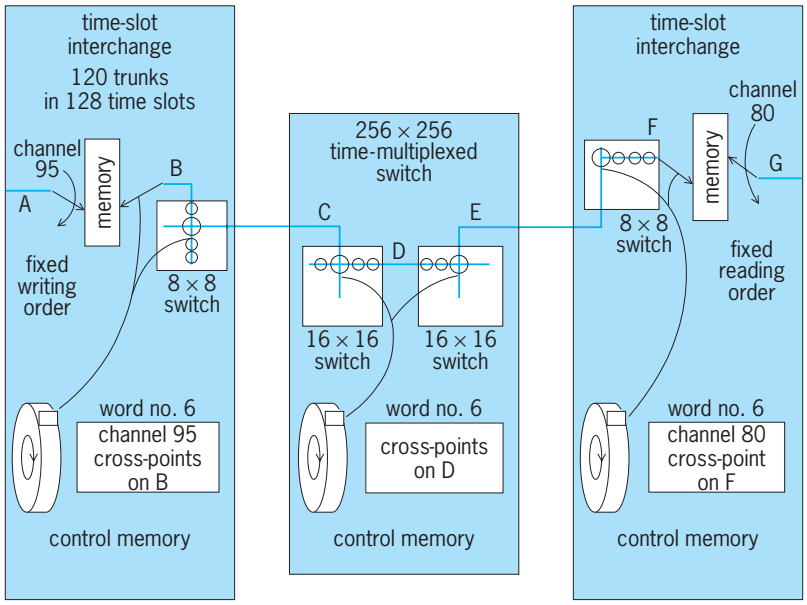

Fig. 3. Diagram of digital time-division switching fabric showing a typical connection. The network has 128 time slots throughout, and the sample shown is in time slot number 6 throughout the network.

time-division switching network routes digital signals from incoming trunks to the desired outgoing trunks. The audio signals on the analog channels are converted by the LT-2 connectors or D4 channel banks into pulse-code-modulation digital samples. The digital interface frame (DIF) processes the digital signals into the format required for the switching network and removes the signaling information. For analog and digital channels, signaling information is detected by the equivalent of trunk circuits in the digital interface frame except where common channel signaling is now being used. Echo suppression on a digital basis is inserted ahead of the switching network.

All operations are directed and supervised by the stored-program-control processor and are aided in routine tasks by the signal processors built into the digital interface frames. The signal processors provide the scanner and distributor functions for a portion of the trunks and, in turn, pass the significant information content in a more compact form to the central control. The use of peripheral or distributed processors enables the central processor to devote its attention to the more critical decisions in the processing of calls, thereby providing for greater call-attempt capacity. System units synchronize their operations under the control of a system clock.

Digital (PCM) local systems. Systems have been designed to terminate ordinary telephone lines carrying analog voice signals directly on analog-to-digital conversion circuits. To use time-division-multiplex switching, voice signals are digitized, assigned time slots, and then multiplexed. The line interface circuit, colloquially known as BORSCHT, provides the system with these capabilities, as well as those normally expected within local space-division switching systems. BORSCHT is an acronym for Battery (to feed the analog telephone transmitter), Overvoltage

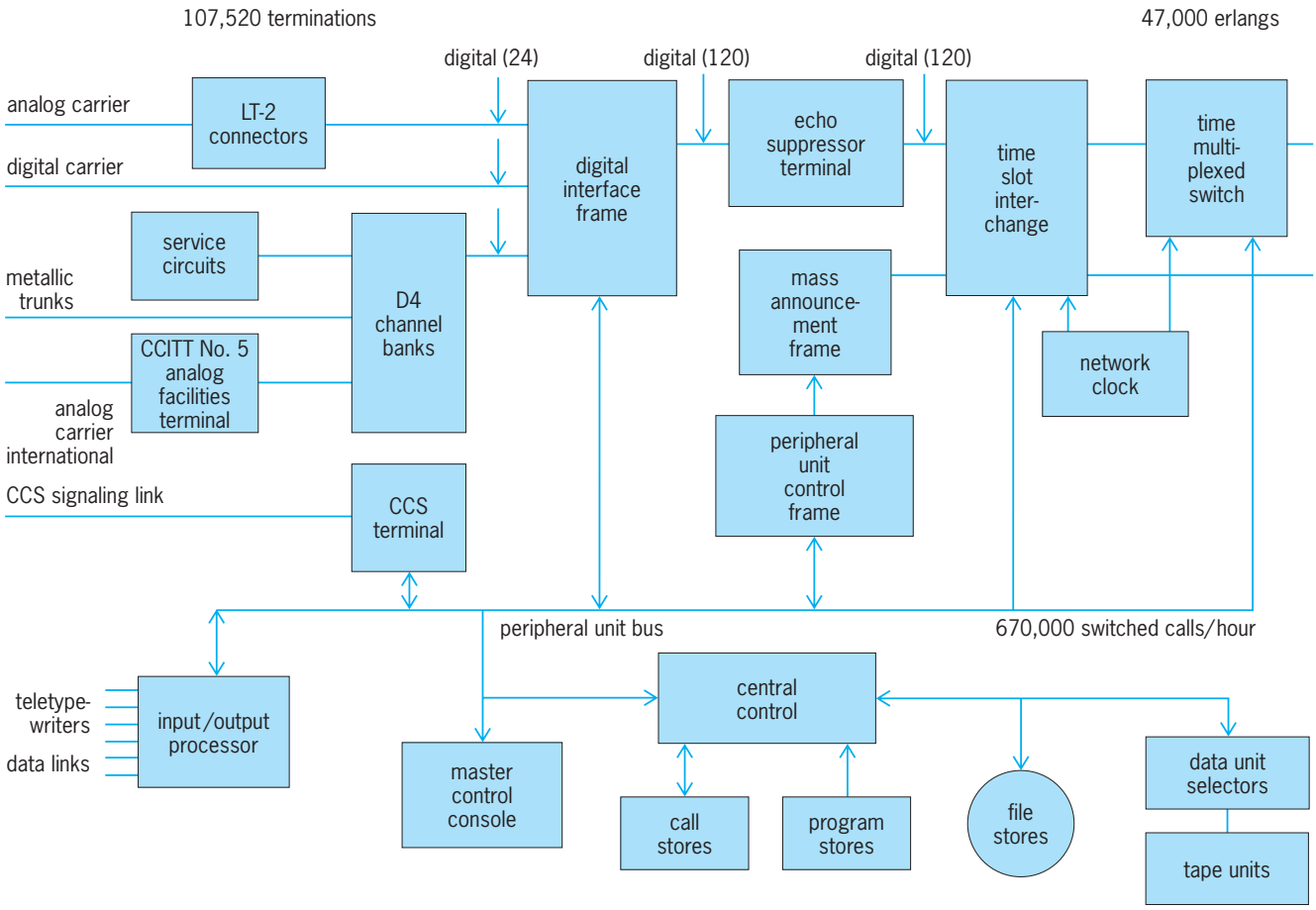

Fig. 4. Block diagram of No. 4 electronic switching system. CCITT = Comité Consultatif International de Téléphonique et Télégraphique.

protection, Ringing, Supervision, digital Coding, Hybrid to separate directions of transmission (a requirement for digital time-division-multiplex systems), and Test access (to reach both the in-office and outside plants).

To provide large call-attempt capacity, particularly for toll system applications, additional control processors are used, usually by providing several identical processors that share the load. These are called multiprocessor systems. Also, microprocessors are associated with the periphery of systems to distribute some of the stored-program-control functions. For smaller systems, some basic control functions are built into logic circuits rather than being stored as programs in the processor memory. These are action translator systems with call sequences determined by the use of memory on a look-up-table basis.

Distributed-control TDM systems. In more advanced local digital switching systems, not only is the interface circuit highly integrated, but line and trunk terminations are grouped into modules that permit the application of distributed control techniques. Some or all of the routine call-processing functions are performed by programs stored in microprocessors associated with the modules. In addition, the first switching stage, usually a time stage, is included in the module. The module therefore includes all of the basic switching functions of the network, interface, and control (**Fig. 5**). Systems with varying degrees of distributed control and network are known as distributed switching systems. The close association of these functions in a unit leads to further opportunities for the use of very large-scale integrated (VLSI) circuits.

Distributed switching has also renewed interest in separating the modules from the time-multiplex-switching and central control system core, which together are sometimes called the host. The modules may be located away from the system core, closer

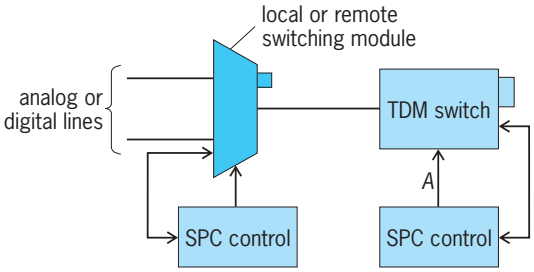

Fig. 5. Second-generation local time-division digital switching system. If the TDM switch is self-directing, the path labeled *A* is omitted.

to the customers. Not only are line facilities saved, but they use digital transmission to connect these modules to the host. Typically these remote units serve from a hundred to several hundred lines. For the larger units, the microprocessor controls used in the switch module are sufficiently autonomous to enable them to function on intralocal calls should the trunk linking the module with the host be disrupted. When remote units provide for some operation independent of the host central office, they are known as remote switching modules, units, or systems.

Changes in the system core have also begun to appear as a result of the application of distributed switching and large-scale circuit integration (LSI). These changes affect both the network and the control. The network employs both time-multiplex switching and time-slot interchange techniques, interchanging channel information between multiplexed links to and from the modules and between time slots. Large-scale integrated circuit chips can be designed with sufficient intelligence that the modules themselves may control the establishment of paths through the switching network. This technique is known as self-selection since the switching actions in the chip are controlled by signals arriving over the links.

Alternative switching techniques. Switching is gaining its own technology with integrated circuits that include complete time slot interchange, line interface (BORSCHT), and other circuits. As a result, many different system architectures and devices are being introduced around the world. Many systems are of the stored-program type with high-capacity random-access memory chips, such as those with 256,000 and 10^6 bits.

Maintenance and administrative functions have been centralized in computers that serve several central offices. Communication with these centers is two-way, with some information for special calls being stored at the central location.

The time-division digital switching technique is being used in newly installed systems as it becomes economically competitive with space division. Distributed switching is applicable to space division as well as time division. These applications may also use powerful low-cost microprocessors for call processing in switching modules remote from the host. A remote switch employing a nonmetallic switching network has been deployed in the United States.

Another technique for the switching of digital signals is the use of existing metallic space-division switching systems. Most modern switching systems will pass 64,000 bit/s digitized voice signals, highlighting the fact that a digital switching system may employ time or space division. But there is interfacing synergy between time-division-multiplex transmission and time-division switching systems. Furthermore, a time-division system provides an opportunity to transmit digital signals from nonvoice sources, such as computers and terminals. This capability gives rise to the concept of networks integrating digital multiplex transmission, switching, and services.

Data switching. For the switching of data and other digitized information, packet switching can be used. Digitizing the voice at the telephones results in end-to-end integrated services digital networking completely replacing the analog network.

Time-division digital switches serve over 65% of the access lines in the United States. As described above, they are limited to individual connections of 64,000 bits/s. By setting up multiple connections, up to 24, greater digital throughput is possible for certain applications.

For future telecommunications networks, a mixture of packet and circuit switching is envisioned with the message information, whether voice, data, or image, digitized and packaged into small cells of 53 bytes. These cells constitute the principal transmitted signals in a technique known as ATM. As the demand for broadband integrated services digital networks (ISDN) builds, ATM switching fabrics will be added to existing public switches. In the meantime, small ATM switches with a number of ports will appear to couple local-area networks, wide-area networks, and other networks serving relatively small groups of terminals on private data networks. *See* LOCAL-AREA NETWORKS; WIDE-AREA NETWORKS.

Toll Calling

In the United States, telephone service is provided by companies serving geographical areas known as exchanges. The local telephone companies are known as local exchange carriers. In much of the country the local exchange carriers serve one or more contiguous exchanges known as local access transport areas (LATAs). Service between LATAs is generally provided by many interLATA carriers (IXCs). There are several methods by which calls are extended to these carriers. A popular arrangement for serving these toll calls, both inward and outward, is known as equal access. Interconnection between the interLATA carriers and the local exchange carriers is usually through an access tandem. The call is switched to the interLATA carrier of the caller's choice either by presubscription or by dialing the prefix 10XXX, where XXX represents identification of the desired carrier. First, the calling line identification number is forwarded through the tandem to the interLATA carrier. The called number is then forwarded from the originating office to the interLATA carrier. Both numbers are sent by using alternating-current trunk signaling. In some cases the interLATA carrier has trunks directly to or from end offices.

The No. 4 ESS is one of the most commonly used systems for the switching of most interLATA toll calls, as well as for some intraLATA calls (**Fig. 6**). There are two methods of placing calls from the calling telephones. In one, the call is routed from a local office through an access tandem to the nearest toll office. Generally a 1 is dialed ahead of the area code on such calls. In the other, a call is routed from the local office through an operator system, when a customer dials 0 before the called number or makes a long-distance call from a public telephone.

The interLATA carriers use the calling number to

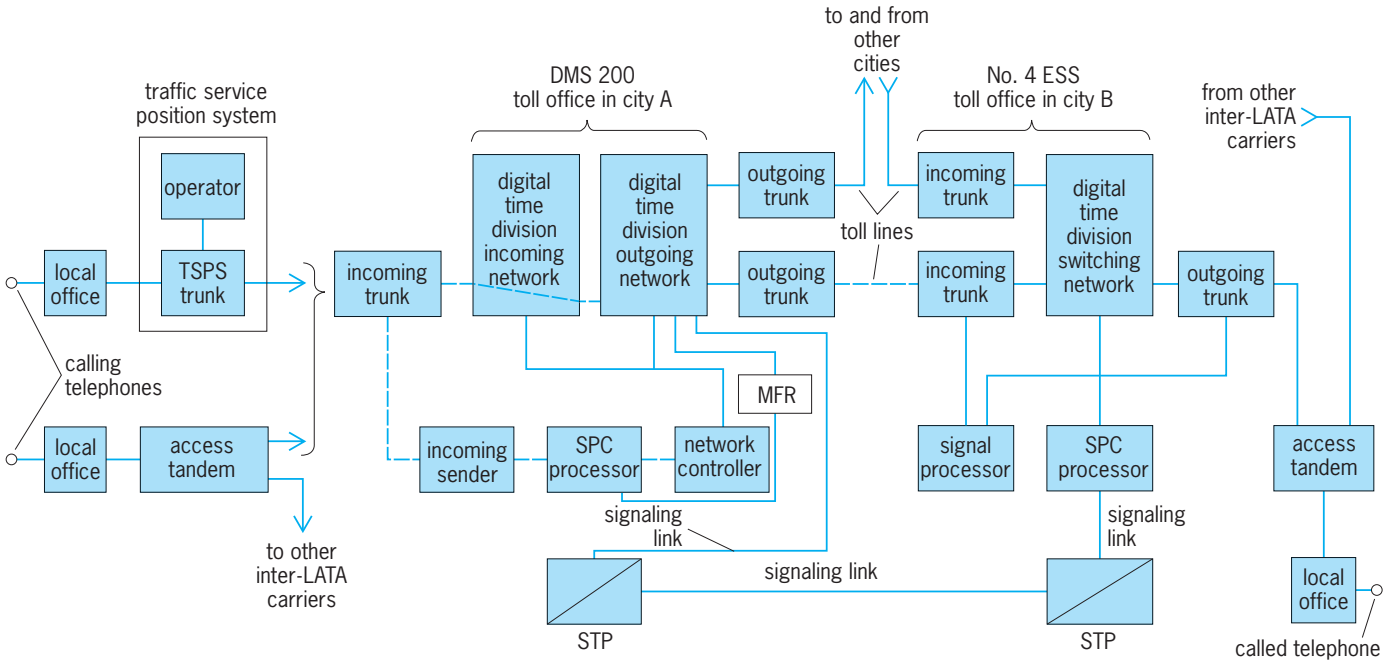

Fig. 6. Diagram of typical toll-call path between distant cities A and B, where the toll office in the city B uses No. 4 ESS and the toll office in city A uses the DMS 200 switching system. Abbreviations are explained in text.

determine if the caller is their customer and use the information for billing purposes. If the number is acceptable, the call is routed from the first interLATA carrier switching office to an office that reaches into the terminating LATA.

Calls arriving at the first interLATA carrier office are connected to a multifrequency receiver (MFR), and signals representing the calling and called numbers are sent from the access tandem.

In most interLATA carrier networks, common channel signaling is used between offices to send information about the call at 56 kilobits per second, including the identity of the outgoing trunk, to the two signal transfer points serving city A. One of these signal transfer points will transfer this information to a signal transfer point serving city B. This signal transfer point will then contact the No. 4 ESS office at which the trunk terminates. The stored program control in the office at B determines the routing and completes the connection to the local office, which in turn sets up the connection to the called telephone. The common-channel-signaling process includes a check that the transmission path selected for the call is viable.

The stored-program-control processor checks the calling number and determines the routing of the call. If all of the direct paths are busy, the interLATA carriers may have alternate routes planned that may carry the calls through intermediate interLATA carrier switches.

An important feature of modern switching systems is their ability to pick an alternate route if the first choice route between cities A and B has no idle toll lines. For example, a route from city A to city C to city B may be selected if no direct trunk to city B is available. Several such alternate routes may be examined, and a call may be routed through several switching points before reaching the terminating toll center. As many as 16 alternate routes may be preplanned. Some alternate routing arrangements are prioritized according to the time of day. The stored-program-control processor then passes information through the network controller to set up a connection from the incoming trunk to outgoing trunks through the incoming and outgoing networks.

Alternate routing ensures good service and economical trunking since, if one route is either busy momentarily or out of service due to an equipment or cable failure, the network is engineered to ensure that there is a good possibility that a toll line in some other route is available. The switching systems in complex networks employing alternate routing must be capable of sending pulses of the required type to operate the switches in the various switching centers. They can also delete digits or add digits to the called telephone number as required to operate the various switches.

Intelligent Networking

Large data storage subsystems, known as service or network control points, are used to verify information provided on credit card calling or for advancing calls, such as 800, to regular addresses in the network. This includes information for routing calls to a toll carrier of choice by the subscriber to the 800 service. The common channel signaling networks for these regional and long-distance carriers have been interconnected, shortening the time from the completion of dialing to the start of ringing on calls independent of their distance.

The information stored at the control points has become very sophisticated, particularly as it applies

to call processing. Calls may be directed differently depending upon time, date, and point of origination. Verbal prompts and voice recognition have been added to request and obtain further information from callers before further call processing. *See* SPEECH RECOGNITION.

The generic name given to these and other network enhancements of this type is intelligent networking. Since the databases are at centralized points, changes in the software in each central office switch may be avoided. Furthermore, the software of these switches is being modified with triggers that, for certain calls from particular stations, will forward the calls to the control points or intelligent peripherals. Intelligent peripherals enable third parties to devise and provide enhanced services to customers who may not have the capability of implementing new services on their own.

Untethered Switched Services

Modern mobile radio service has a considerable dependency on switching. The territory served by a radio carrier is divided into cells of varying geographical size, from microcells that might serve the floors of a business or domicile to cells several miles across that provide a space diversity for serving low-power radios.

Switching systems reach each radio cell site that detects, by signal strength, when a vehicle is about to move from one cell to another. The switching system then selects a frequency and land line for the communication to continue on another channel in a different cell without interruption. This is known as cellular mobile radio service and is used not only for voice but also facsimile and other forms of telecommunication. *See* MOBILE COMMUNICATIONS; TELEPHONE; TELEPHONE SERVICE. Amos E. Joel, Jr.

Bibliography. R. J. Chapuis and A. E. Joel, Jr., *Electronics, Computers, and Telephone Switching*, 1990; A. E. Joel, Jr., *Electronic Switching: Digital Central Offices of the World*, 1982; F. Mazda, *Switching Systems and Applications*, 1996; J. C. McDonald, *Fundamentals of Digital Switching*, 2d ed., 1990; *Proceedings of the International Switching Symposium*, Florence, Italy, May 1984, AEI Publication, Milan, 1984; *Proceedings of the International Switching Symposium*, Stockholm, May 1990.

Switching theory

The theory of circuits made up of ideal digital devices. Included are the theory of circuits and networks for telephone switching, digital computing, digital control, and data processing.

Switching theory generally is concerned with circuits made of devices or elements that can be in two or more discrete conditions or states. Examples of such devices are switches or relay contacts, which can be opened or closed, rectifying diodes, which can be either forward- or back-biased, solid-state elements (such as transistors), which can be saturated or cut off, and magnetic elements which can be magnetized to saturation in either of two directions. Switching theory establishes an ideal representation of the digital circuit, examines the properties of the representation, then interprets these as properties of the circuit. Switching theory is not concerned with the physical phenomena of action or stability in a particular condition or with the details of transition from one state to another. It takes these as established and proceeds to examine more or less complex combinations of digital devices whose properties are assumed to be ideal.

The bulk of switching theory is concerned with circuits that are made of binary (two-valued) devices, since these are most common. Switching theory can be based in part on mathematical logic. *See* BOOLEAN ALGEBRA.

A switching circuit whose outputs are determined only by the concurrent inputs is called a combinational circuit (or logic circuit). A circuit in which outputs at one time may be affected by inputs at a previous time is called a sequential circuit.

Combinational circuits. A rule by which the outputs of a combinational circuit can be determined from its inputs is called a switching function. Since the variables are discrete, a switching function may be expressed in tabular form as a truth table, or may be indicated by a diagram or geometric pattern. If the function and variables are binary, the symbols of 1 and 0 are commonly used to represent the two values. The function may then be represented by a boolean algebraic expression. The two values of a switching function can represent closed and open circuits, as for switches or relay contacts, or high and low or plus and minus voltages, as in electronic circuits.

The simplest combinational switching functions are the NOT function, the AND function, and the OR function. The NOT function is designated by the prime in boolean algebra; $Y = X'$ means that Y is closed (high, plus) when X is open (low, minus), and vice versa. The AND function is designated by the boolean product; $Z = X \cdot Y$ means that Z is closed (high, plus) only if both X and Y are closed. The OR function is designated by the boolean sum; $Z = X + Y$ means that Z is closed if either X or Y or both are closed. All other combinational switching functions can be made by combining these elementary building blocks.

For example, **Fig. 1** shows a switching circuit with three switches, or contacts, X, Y, and Z', each of which can be either open or closed. These can be thought of as input variables. The circuit as a whole will be open or closed depending upon the individual positions of X, Y, and Z'. Its condition can be designated by W, an output variable. Let 0 represent the open condition, and 1 the closed condition. The table in Fig. 1 represents the switching function of the circuit. The boolean expression for this function is $W = Z'(X + Y)$. To interpret this expression the rules of simple boolean algebra must be used: $0 + 0 = 0$ $0 \cdot 0 = 0$ $0' = 1$ $0 + 1 = 1$ $0 \cdot 1 = 0$ $1' = 0$ $1 + 0 = 1$ $1 \cdot 0 = 1 + 1 = 1$ $1 \cdot 1 = 1$.

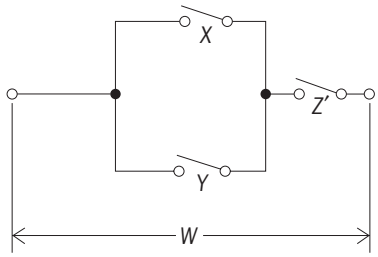

X	Y	Z′	W
0	0	1	0
0	0	0	0
0	1	1	1
0	1	0	0
1	0	1	1
1	0	0	0
1	1	1	1
1	1	0	0

Fig. 1. Combinational circuit. $W = Z'(X + Y)$. X and Y are normally open contacts. Z' is a normally closed contact.

Switching theory establishes a number of methods for analysis and synthesis of combinational circuits. A significant problem is minimization, that is, given a switching function, to synthesize the simplest circuit which will realize it. A problem of some theoretical difficulty is that of realizability, that is, given a statement of specifications, to determine whether a switching circuit exists which satisfies them.

Analysis of a series-parallel combination of switches or relay contacts can be carried out by a direct application of boolean algebra. Variables or terms corresponding to contacts, or combinations in parallel, are added, and those in series are multiplied. The values are interpreted according to the rules of boolean algebra. Similar methods can be applied to combinational circuits which employ various solid state switching elements, such as transistors. Circuits that are not series-parallel can be dealt with by an extension of the boolean method, by the use of matrices with discrete-valued elements, or by a number of special methods.

A switching function can be simply synthesized as a series-parallel combination of contacts by giving boolean symbols circuit interpretations explained previously. Electronic logic circuits can be synthesized in a similar fashion. This approach will lead to a method for embodying any switching function expressed in boolean terms. The boolean expression of a function given in tabular or diagrammatic form is easily obtained.

Synthesizing the minimal circuit, or minimization, is more difficult since for every switching function there are many possible circuits. Where the number of variables is small, the minimization problem can often be reduced to one that has already been solved. Tables of minimal or nearly minimal solutions for relay circuits and solid-state circuits are available for circuits with one output and as many as four inputs. Harvard chart methods and Karnaugh map methods utilize geometrical relationships to explore systematically functions with one output and as many as six inputs.

As the number of variables increases, the possible number of functions rapidly becomes large. For example, there are more than 10^{19} different functions of six binary variables. No completely general and practical design methods have been discovered. However, a growing array of special methods for synthesis and minimization is available.

Sequential circuits. Since the outputs of sequential circuits depend on past, as well as present, inputs, they must contain means for remembering or storing the effect of past inputs, such as locking relays, flip-flops, delay lines, or solid-state memories. A device with two stable states can remember one binary digit, or bit. The amount of memory in a circuit can be measured either in bits or in internal states. An internal state of a circuit is a particular configuration of its internal memory devices. The number of internal states is equal to 2^n, where n represents its number of bits. Binary counters and shift registers are examples of sequential circuits.

It is possible to represent a sequential circuit as a combinational circuit with feedback. Thus, the combinational circuit of **Fig. 2** becomes a sequential circuit with two bits of memory if two of its outputs are connected to two of its inputs. Any such closed loop must contain gain and some delay; sometimes additional delay is inserted.

If the combinational circuit and the delays in Fig. 2 are completely specified, the internal description of the circuit is known and its behavior can be analyzed. If the switching function of the combinational circuit is such that $m_1 = M_1$ and $m_2 = M_2$ for a given set of inputs, no change can occur as a result of the action of the memory loops and the circuit is stable; otherwise, it is unstable. If it is unstable, the inputs must cause a transition to a new state, which in turn may

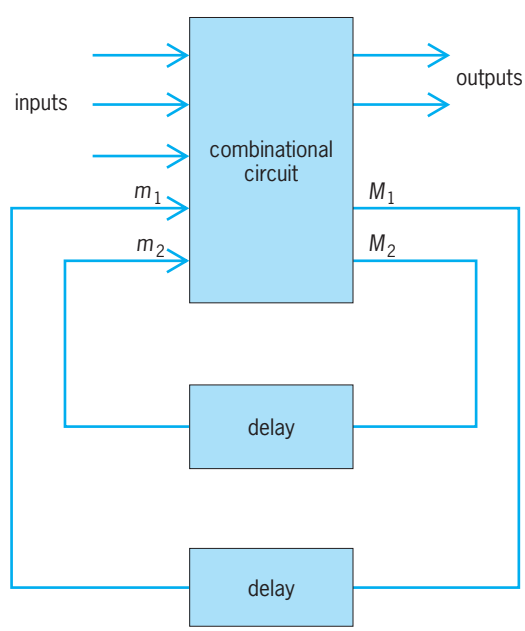

Fig. 2. Sequential circuit with two memory loops.

be stable or unstable. If no stable state is reached, the circuit is said to buzz. If the state to which a circuit may pass depends on which of two or more memory loops acts first, the circuit is said to have a race condition, and its performance may be ambiguous. This difficulty does not occur in circuits in which changes are caused or timed by repetitive clock pulses. Such circuits are called synchronous. Circuits which make transitions at the natural internal rate are known as asynchronous, and these asynchronous circuits must be designed with greater care.

To proceed from external circuit requirements to an internal description of a sequential circuit requires art and skill, as well as knowledge of switching theory. *See* DIGITAL COMPUTER; LOGIC CIRCUITS; SAMPLED-DATA CONTROL SYSTEM; SWITCHING CIRCUIT; SWITCHING SYSTEMS (COMMUNICATIONS).

Willard D. Lewis

Bibliography. E. D. Fabrioius, *Modern Digital Design and Switching Theory*, 1992; J. Y. Hui, *Switching and Traffic Theory for Integrated Broadband Networks*, 1990; Z. Kohavi, *Switching and Finite Automata Theory*, 2d ed., 1978; S. Muroga, *Logic Design and Switching Theory*, 1979, reprint 1990.

Sycamore

American sycamore (*Platanus occidentalis*), a member of the plane tree family, known also as American plane tree, buttonball, or buttonwood, and ranging from southern Maine to Nebraska and south into Texas and northern Florida. Ordinarily this tree is 60–120 ft (18–37 m) in height and has a trunk which is 2–5 ft (0.6–1.5 m) in diameter. Individuals 140 ft (43 m) tall and 14 ft (4.3 m) in diameter have been recorded. It has the most massive trunk of any American hardwood. Characteristic are the white patches which are exposed when outer layers of the bark slough off; the simple, large, lobed leaves whose stalks completely cover the conical winter buds; and the spherical fruit heads that are always borne singly in the American species and persist throughout the winter (**illus.** *a*). The tough, coarse-grained wood is difficult to work, but is useful for butchers' blocks, saddle trees, vehicles, tobacco and cigar boxes, crates, and slack cooperage. The trees are usually scattered through the forest in moist soil. *See* ROSALES.

Sycamores. (*a*) American sycamore (*Platanus occidentalis*), bud, leaf with fruit ball, and budding twig. (*b*) London plane tree (*P. acerifolia*), leaf with fruit balls.

London plane (*P. acerifolia*) is supposedly a hybrid of *P. occidentalis* and *P. orientalis*, and it is one of the most desirable trees for planting in crowded cities because of its resistance to injury from gases, smoke, dust, and drought. It can be recognized by the usually three-lobed leaves resembling those of a maple, and by the fruit balls borne in groups of two or three (illus. *b*). The sycamore of Europe is usually understood to be a maple, *Acer pseudoplatanus*, known also as the sycamore maple. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Sycettida

An order of the subclass Calcaronea in the class Calcarea. This order comprises a rather diverse group of calcareous sponges, and includes the families Sycettidae, Heteropiidae, Grantiidae, Amphoriscidae, and Lelapiidae. Choanocytes with apical nuclei are limited to flagellated chambers and never occur lining the general spongocoel. The family Sycettidae resembles the contrasting order Leucosoleniida in lacking the true dermal membrane or cortex possessed by the other five families. The most massive skeleton (of bundles of modified triradiate spicules) is found in the family Lelapiidae, which superficially resembles the unrelated pharetronid sponges (abundant in the Permian and Triassic periods). The canal system is never asconoid (as in Leucosoleniida), but can be either syconoid or leuconoid. Leuconoid species seem to be derived from syconoid forms. Genera include *Grantia*, *Sycon*, *Sycetta*, *Heteropia*, *Amphoriscus*, and *Leucilla*. *See* CALCAREA; LEUCOSOLENIIDA.

W. D. Russell-Hunter

Syenite

A phaneritic (visibly crystalline) plutonic rock with granular texture composed largely of alkali feldspar (orthoclase, microcline, usually perthitic) with subordinate plagioclase (oligoclase) and dark-colored (mafic) minerals (biotite, amphibole, and pyroxene). If sodic plagioclase (oligoclase or andesine) exceeds the quantity of alkali feldspar, the rock is called monzonite. Monzonites are generally light to medium gray, but syenites are found in a wide variety of colors (gray, green, pink, red), some of which make the material ideal for use as ornamental stone.

Composition. Syenites may be classed as normal (calc-alkali) syenites or alkali syenites. In the latter the alkali feldspar and mafics are soda-rich. Intergrowths of potash and soda feldspar (perthite) are of various types and are strikingly developed. Feldspar grains usually show fair crystal outlines (subhedral) or may be irregular (anhedral). Many are highly interlocking. Sanidine occurs in some of the finer-grained varieties and in syenite porphyry. Normal syenites generally carry crystals of soda plagioclase

(oligoclase) which are usually subhedral and may be zoned (with calcic cores and sodic rims). Some alkali syenites contain discrete grains of albite. Plagioclase of monzonites may be as calcic as sodic andesine.

Black flakes (microscopically brown) of biotite mica and irregular to stubby prisms of green hornblende are characteristic of normal syenite. Diopsidic augite is the most common pyroxene and frequently forms cores within hornblende crystals. In alkali syenite the mafic minerals show wide variation. Biotite is deeply colored and iron-rich. Amphiboles are soda-rich (arfvedsonite, hastingsite, or riebechite) and are commonly zoned. Diopsidic and titanium-rich augite crystals are commonly encased by shells of aegerine-augite and aegerite.

Minor constituents may include quartz which is usually interstitial. When present in amounts between 5 and 10%, the rock is called quartz syenite; in excess of this amount, the rock becomes granite. Small amounts of feldspathoid (nepheline, sodalite, or leucite) may be present; but if in excess of 10%, the rock becomes a feldspathoidal syenite (nepheline syenite).

Accessory minerals include zircon, sphene, apatite, magnetite, and ilmenite. Accessories in special varieties of syenite include iron-rich olivine, corundum, fluorite, spinel, and garnet.

Texture. The texture of syenite is most commonly even-grained. Very coarse or pegmatitic textures are local. In some syenites numerous, relatively large crystals (phenocrysts) of alkali feldspar give the rock a porphyritic texture. These may be of early or late generation and may range from euhedral (well-formed crystals) to anhedral. They are particularly abundant in finer-grained varieties and in syenite porphyries.

Structure. A variety of directive structures may be present. Banding and parallel wavy streaks (schlieren) of different minerals are seen in some syenites; flow structures due to clustering and parallel orientation of elongate minerals may be present. Euhedral tabular feldspar crystals in parallel arrangement give the rock a distinctive appearance. In some cases these directive features represent effects of magma flow; in others they represent vestigial bedding or foliation in metasomatic or metamorphic rocks.

Occurrence and origin. Syenite is an uncommon plutonic rock and usually occurs in relatively small bodies (dikes, sills, stocks, and small irregular plutons). Normal syenite may be associated with monzonite, quartz syenite, and granite, whereas alkali syenites are associated with alkali granites or feldspathoidal rocks.

Many syenites have crystallized directly from syenitic magma (rock melt); others may have formed by reaction between magma of nonsyenitic composition and abundant contaminating rock fragments. Still others may have formed metasomatically as alkali-rich emanations, perhaps escaping from deeply buried magmas, have permeated rocks of special composition, and have replaced them with abundant alkali feldspar. *See* IGNEOUS ROCKS; MAGMA; METAMORPHISM; METASOMATISM.

Carleton A. Chapman

Symbiotic star

A double star system in the late stage of stellar evolution. A symbiotic star is a binary system and not a single star. The symbiotic phase represents a brief span in the life of the binary. Symbiotic stars are rare objects, and their distances are as a rule many hundreds of parsecs (1 parsec = 3.26 light-years). Thus, on ordinary photographic plates of the sky, symbiotic stars appear pointlike and are not resolved into two individual stars. On closer inspection, symbiotics are always associated with a nebular environment. The "near-official" list of symbiotic stars contains 188 safe entries; 15 of them are extragalactic, of which 1 lies in the dwarf galaxy Draco, 6 lie in the Small Magellanic Cloud, and 8 lie in the Large Magellanic Cloud. The list contains in addition 30 suspected candidates. *See* LOCAL GROUP; MAGELLANIC CLOUDS.

Spectra and variability. Symbiotic stars became targets of research in the early 1930s. Their spectra display titanium oxide absorption bands, indicative of a cool star, together with emission lines of singly ionized helium or doubly ionized oxygen and occasionally even many times ionized iron, which point to a hot nebula. The spectra indicate the presence of a cool M-type star with a surface temperature below 4000 K, and a hot nebula quite similar to a planetary nebula. The light output is variable on time scales of months and years, and the light curve can show abrupt changes, with increases that imply changes of at least a factor 2—but much more on other occasions—in the total energy output on time scales of days and months. *See* ASTRONOMICAL SPECTROSCOPY; LIGHT CURVES; NEBULA; PLANETARY NEBULA; VARIABLE STAR.

Evidence of binary systems. In 1981 the first astrophysical colloquium dedicated to symbiotic stars took place. Of the then more than 100 known symbiotics, only 5 were proven binaries. The proof came from periodic eclipses and from periodic Doppler shifts in wavelength that showed that the emitting object was moving to and from the observer. Binary periods between one and several years were found. During the following years the evidence accumulated that all symbiotics are double stars and that indeed their binary periods are between one year and many dozens of years. Their orbits are sufficiently wide for the two stars not to be in direct contact, and both stars lie safely within their Roche lobe. Thus, for the two stars stellar evolution proceeds over the whole main-sequence lifetime practically uninfluenced by the partner. That changes dramatically in the later stages of their evolution. *See* DOPPLER EFFECT; ECLIPSING VARIABLE STARS.

Dust. Systematic observations in the infrared led to a division into two classes: D-types that show infrared emission indicative of astronomical dust, and S-types that emit infrared radiation typical of red

giant atmospheres. D-types always seem to contain Mira variables, and their binary periods are probably longer than 10 years. The S-types are relatively dust-free and have binary periods shorter than 20 years. In the D-types, dust emits thermal radiation at average temperatures of typically 1000 K. Due to the wide binary separation in these systems, the cool star has been permitted to evolve to a state with substantial dust production in its envelope. *See* GIANT STAR; INFRARED ASTRONOMY; MIRA.

Late-stage stellar evolution. The gas temperature in the symbiotic nebula is approximately 15,000 K. This is much below the temperature that is needed to collisionally ionize the atoms in the nebula to the observed degree. The nebular gas is obviously radiatively ionized. This cannot be due to the cool star; it must be the action of a small, very hot star. The combination of observations and model calculations shows the cool star to be a red giant. Typical radii are 100±50 solar radii, or up to 300 solar radii for Miras, and temperatures of 3500±500 K, or approximately 2500 K for Miras. For the hot star, temperatures between 50,000 and 200,000 K are found. The star's size is that of a white dwarf or of a central star in a planetary nebula, thus between ~0.01 solar radius (approximately the size of the Earth) and ~0.1 solar radius; higher values are found for particularly active phases. Only in a few cases is it possible to determine stellar masses. They are approximately 2.5±1.5 solar masses for the cool giants and 0.6±0.2 solar mass for the hot star. The low mass of the white dwarf makes it unlikely that by further accretion this star could later explode as a supernova of type Ia; however, that issue is not resolved.

The symbiotic nebula has, as a rule, the same chemical composition as expected from red giants. This is taken as evidence that the red giant suffers considerable mass loss. Values between 10^{-8} and 10^{-7} solar mass per year have been found. Wind velocities are between 10 and 30 km/s (16 and 18 mi/s) for the red giant, and ~1000 km/s (600 mi/s) for the white dwarf. This constellation of properties situates symbiotics in the late stage of stellar evolution. They began as stars with masses of the order of 5 solar masses. The more massive of them has gone faster through the main sequence, and when it arrived in the red giant region, after having converted hydrogen in its core to helium, it shed most of its mass through a stellar wind. That material can occasionally still be detected in the wider environment. The star has become a hot white dwarf. The originally less massive star has retained its mass and has now entered the red giant phase, with hydrogen burning in a shell, and a large mass loss. This phase may last on the order of 10^7–10^8 years. At the end of it, a planetary nebula is likely to be created for a few ten thousand years. *See* STELLAR EVOLUTION; WHITE DWARF STAR.

Interactive binary. A fraction of the mass lost by the red giant is captured by the white dwarf. This provides new hydrogen-rich fuel on top of the white dwarf's carbon-oxygen core. The accretion proceeds via some kind of accretion disk, the details of which are not known. It liberates gravitational energy

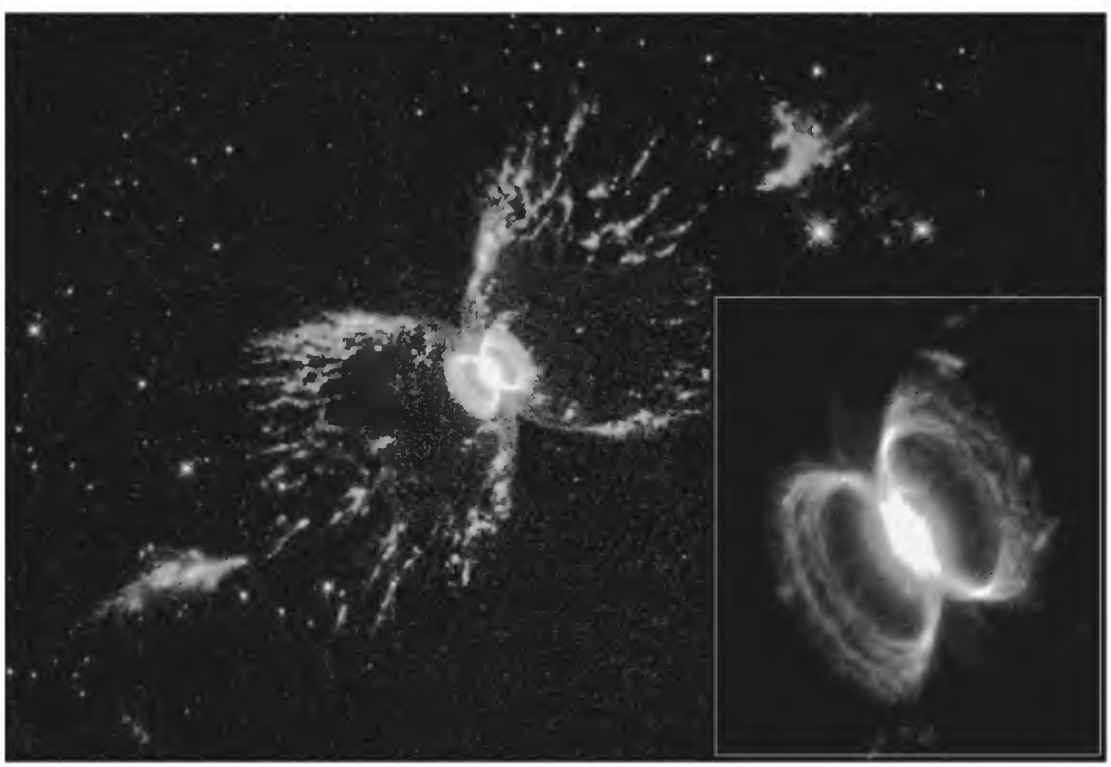

Fig. 1. ***Hubble Space Telescope*** **image of the Southern Crab Nebula, about 4400 parsecs (14,000 light-years) away. It contains a symbiotic system with a Mira-type star and a white dwarf, too close together to be visible as separate objects. The inset shows a bright nebula embedded in the center of a larger one. Both structures are probably due to a symbiotic nova outburst thousands of years ago. (*Courtesy of R. Corradi, M. Livio, U. Munari, H. Schwarz; NASA*)**

which can convert into radiative energy and be responsible for some of the irregular luminosity variations. The hot radiation from the white dwarf ionizes a fraction of the wind and the outer atmosphere of the red giant. This influences the dust production in the red giant.

Symbiotic nova. When the white dwarf has accumulated a critical mass of hydrogen, thermonuclear reactions on its surface lead to an outburst of energy lasting for over 100 years (**Fig. 1**). The energy production during that outburst can reach many thousand times that of the Sun. The mechanism is similar to a nova explosion, except that the nova has a higher peak energy output but a shorter duration. The total energy output of a symbiotic nova is comparable to that of a classical nova. Whereas in symbiotic novae the outburst on the white dwarf is not disturbed by the red giant, in classical novae the two stars are closer to each other, which has repercussions on the outburst.Whether some symbiotics in this way can accumulate sufficient mass to explode as a supernova of type Ia is still an open question. *See* CATACLYSMIC VARIABLE; NOVA; SUPERNOVA.

Nebular environment. Observations have shown traces of the mass formerly lost by the hot star. However, as noted above, the bulk of the matter now detected as an ionized nebula is due to the present mass loss of the red giant in its stellar wind. In its active phase the white dwarf may have a stellar wind of its own. In that case the nebular environment will be strongly structured by the collision of the two winds. This leads to shock zones with temperatures of several million kelvins. The origin of bipolar gaseous jets (**Fig. 2**) observed in several symbiotics is not yet known. Observational evidence indicates that the white dwarf possesses a strong magnetic field which could play a major role.

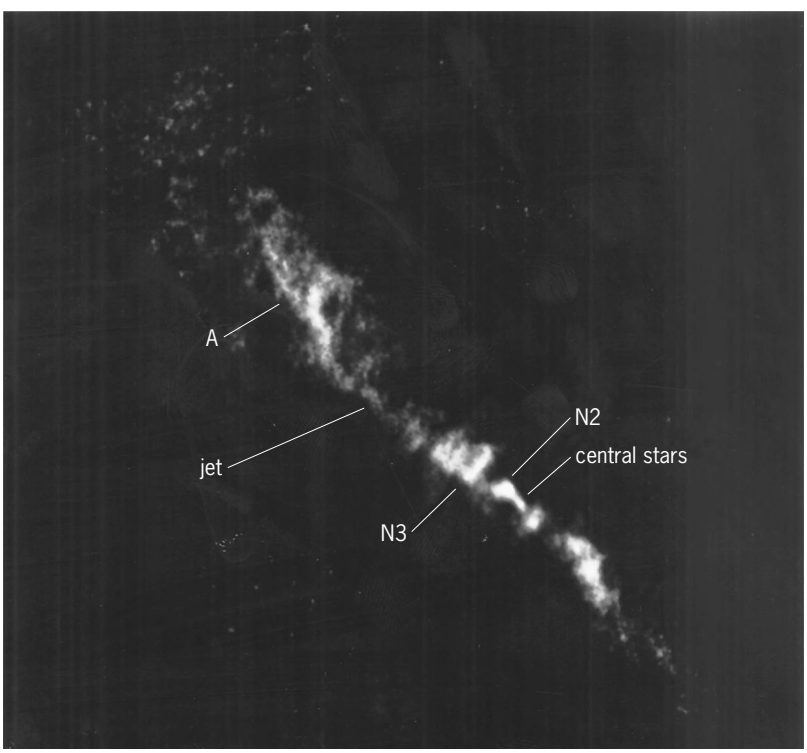

Fig. 2. ***Hubble Space Telescope*** **image of R Aquarii, the closest symbiotic system, about 200 parsecs (650 light-years) away. Its highly collimated bipolar jets are ejected from the central stars, which are too close together to be seen as separate objects. The jets encounter dense regions, such as N2, N3, and A, where they create shock waves, inducing strong radiation emission. (*Courtesy of F. Parece, European Southern Observatory; NASA*)**

Observations. The simultaneous presence of cool dust, a cool star, a hot star, and a hot nebula distributes the spectral information from x-rays through the far-ultraviolet, the visual range, and the infrared, up to radio wavelengths. In addition to ground-based observatories, several artificial satellites—in particular the *International Ultraviolet Explorer* (*IUE*), the *Hubble Space Telescope* (*HST*), the *Roentgen-Satellite* (*ROSAT*), the *Advanced Satellite for Cosmology and Astrophysics* (*ASCA*), the *Far Ultraviolet Spectroscopic Explorer* (*FUSE*), as well as *XMM-Newton*, the *Chandra X-ray Observatory, INTEGRAL, SWIFT*, and the *Infrared Space Observatory* (*ISO*)—have contributed spectra and pictures. Amateur astronomers, following daily light variations, have provided valuable contributions. The discovery of strong Rayleigh and Raman scattering within the symbiotic system, which leads to polarization of the scattered light, has helped to determine the binary movement and the highly complex structure of the symbiotic nebula. *See* BINARY STAR; CHANDRA X-RAY OBSERVATORY; GAMMA-RAY ASTRONOMY; HUBBLE SPACE TELESCOPE; POLARIMETRY; RAMAN EFFECT; SATELLITE (ASTRONOMY); SCATTERING OF ELECTROMAGNETIC RADIATION; STAR; ULTRAVIOLET ASTRONOMY; X-RAY ASTRONOMY.

Harry Nussbaumer

Bibliography. K. Belczynski et al., A catalogue of symbiotic stars, *Astron. Astrophys. Suppl.*, 146: 407–435, 2000; R. L. M. Corradi, J. Mikolajewska, and T. J. Mahoney (eds.), *Symbiotic Stars Probing Stellar Evolution*, ASP Conf. Proc., vol. 303, Astronomical Society of the Pacific, San Francisco, 2003; S. J. Kenyon, *The Symbiotic Stars*, Cambridge University Press, 1986; J. Mikolajewska (ed.), *Physical Processes in Symbiotic Binaries and Related Systems* (Proceedings of an international conference held in 1996), Copernicus Foundation for Polish Astronomy, Warsaw, 1997; J. Mikolajewska et al. (eds.), *The Symbiotic Phenomenon* (Proceedings of the IAU Colloquium held in 1987), Kluwer Academic, 1988.

Symbolic computing

The manipulation of symbols, representing variables, functions, and other mathematical objects, and combinations of these symbols, representing formulas, equations, and expressions, according to mathematical rules.

Giving the solution to an equation using symbols instead of numbers is referred to as symbolic manipulation or symbolic computing. Often, symbolic computing refers to computations or calculations that involve the manipulation of mathematical symbols to obtain solutions. For example, the differentiation and integration of functions in calculus, manipulation of

matrices in linear algebra, or determination of the function that solves a differential equation are now commonly referred to as symbolic computations.

Consider the equation

$$3x - 5 = 0 \tag{1}$$

whose exact solution is $x = 5/3$ or $x = 1^2/_3$ or $x = 1.66\overline{6}$, where the digit 6 repeats forever. All of these solutions are called exact solutions to the equation. If we write $x \approx 1.666$, we say 1.666 is an approximate numerical solution to our equation. Determining exact decimal digit solutions to an equation is not always possible. Consider the equation

$$x^2 - 2 = 0 \tag{2}$$

whose exact solution is $x = \pm\sqrt{2}$. Since $\sqrt{2}$ is a symbol for representing a number, this is referred to as the exact solution in symbolic form. An exact decimal digit representation for $\sqrt{2}$ is not possible, since there is no repeating pattern in the infinite number of digits for $\sqrt{2}$.

Now consider the general form for each of the above two equations. The general form for the first equation would be $ax+b = 0$, where a and b would represent real numbers. In particular, letting $a = 3$ and $b = 5$ in this equation leads to the first equation. The solution to this equation in symbolic form is $x = -b/a$ as long as $a \neq 0$. The second equation is a specific example of the general quadratic equation $ax^2+bx+c = 0$, whose solution in symbolic form is

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{3}$$

One would let $a = 1$, $b = 0$, and $c = -2$ to obtain the second equation. *See* CALCULUS; LINEAR ALGEBRA.

Software packages. In the past, symbolic computation was done by the human mind with paper and pencil or in some similar manner. Now computers are able to do symbolic manipulations, commonly using software packages or hand-held graphing calculators. Some common software packages that can do symbolic computing are Mathematica®, Maple®, and MATLAB®. Texas Instruments and Casio are two of the most common hand-held calculators that have symbolic computing features. Microsoft Office's Excel allows Maple as an add-in so that symbolic computing can be done. Software programming environments that can do symbolic computing are usually called symbolic computing packages or computer algebra systems. *See* CALCULATORS.

Symbolic computing software packages are able do algebra and calculus in any dimension—that is, they have the ability to do scalar or vector calculus (**Fig. 1**). They can do linear algebra operations. They can solve many differential equations and perform transformations on objects, like differential equations. Many packages have components that do abstract analysis. Symbolic computing packages can work with real or complex numbers and usually have some routine for dealing with complex arithmetic and its subtleties. Symbolic computing packages are capable of doing most university undergraduate mathematics and many areas of high-level research mathematics. Most of the standard algorithms used in mathematical and scientific research are incorporated in the package. Sometimes the algorithm is stored in a library inside the software package and has to be loaded into the package from a directory. *See* ALGORITHM; MATHEMATICAL SOFTWARE.

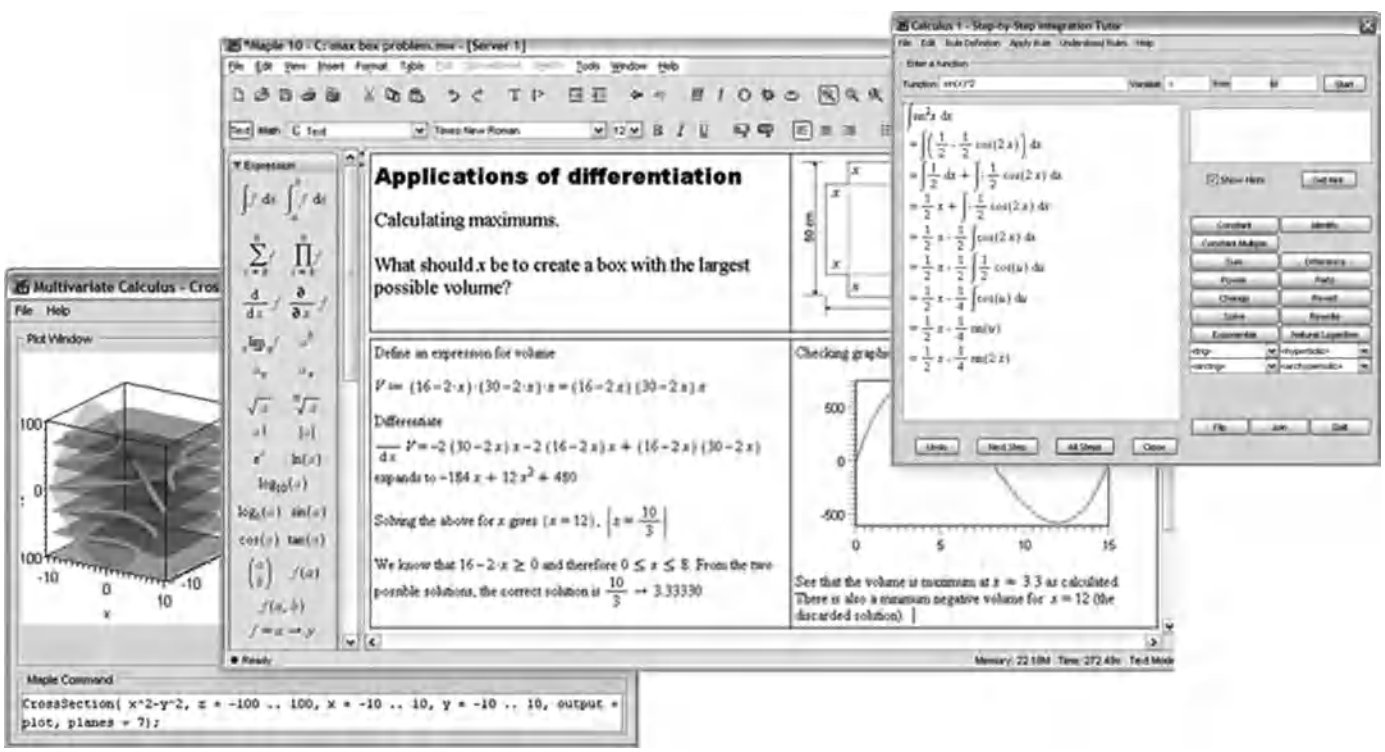

Fig. 1. Example of a calculus software package. (*Courtesy of Maplesoft*)

Symbolic computing packages are equipped with algorithms and tables that can replace a symbol for a number with an accurate or exact decimal digit numerical value. For example, one can ask the package to give the first 200 decimal digits for $\sqrt{2}$. If the package gives the answer as a symbolic expression, one can replace the parameters in the expression with numerical values or ask the package to do it and then ask it to give numerical values to the symbolic expression to as many decimal digits as desired.

The symbolic formulas given by the software package as an answer can usually be visualized in a variety of formats. These packages are able to rotate graphics and view graphics from different viewpoints and in different formats. Graphics frames can be built and then animated in sequence to appear as a movie. This allows visual scientific analysis and discovery. The software packages can also visualize data sets of numerical values in various formats. The graphics for both of these cases can be done in one, two, or three dimensions with different coloring schemes and plotting symbols depending on the user's preferences (**Fig. 2**).

In most cases, the graphics generated by these software packages is of publishable quality. In fact, most symbolic computing packages have features that allow their graphics to be exported in various formats for implementation in a variety of word processors. Some packages will even translate their symbolic output into the format needed by specialized scientific word processors.

Most symbolic computing packages come with a high-level language that the user can program. These programs can usually be translated by the software package to a high-level language like C or Fortran. These C or Fortran programs can then be run on state-of-the-art computers to improve the speed of calculations, the efficiency of calculating, and the robustness (accuracy) of results. The data run by these codes can then be inputted into the software package where they can be visualized. The data can be viewed in a variety of formats. *See* PROGRAMMING LANGUAGES.

Applications. Symbolic computing software packages allow a user to do symbolic manipulations in seconds. These same manipulations can take minutes, hours, or even days for humans to do. Some of the manipulations are not humanly possible. The fact that symbolic computation packages can do manipulations so quickly allows teachers to show students in the classroom environment symbolic computations that used to take several classroom sessions or could not be demonstrated at all in the classroom because

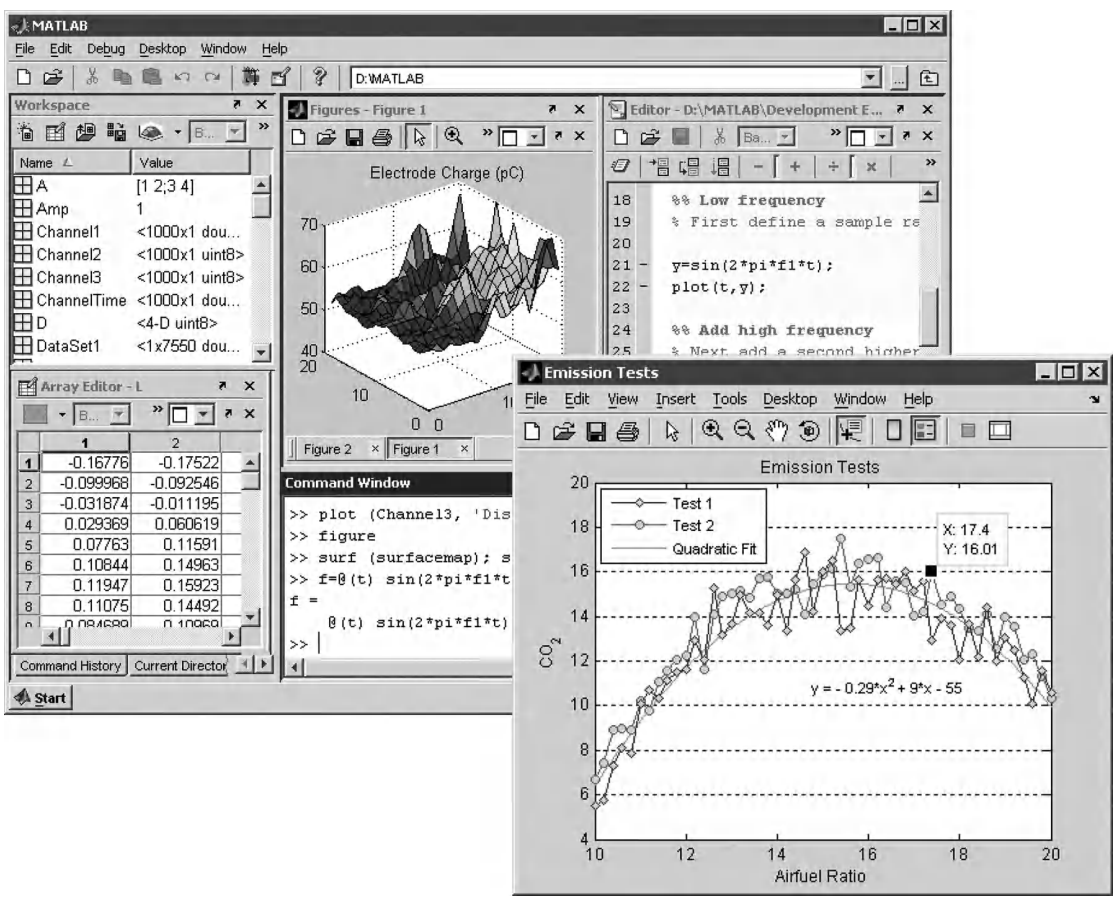

Fig. 2. MATLAB development environment, which allows users to develop algorithms, interactively analyze data, view data files, and manage projects, showing line plots of multiple engine emission test results with a curve fitted to the raw data. (*Courtesy of The MathWorks, Inc.*)

of time considerations. Symbolic computing packages are giving teachers much more flexibility and power in the classroom.

These software packages can also do professional-grade research. They are used to improve and enhance engineering design in many industries, such as electronics, automotive, and aviation. Many software packages have the ability to be run in parallel. This feature allows the increase in the speed of calculations and the speed of visualization formats.

Almost all symbolic computing packages have a new-user learning environment. This tool allows the new user to go through a set of lessons to learn how to use the symbolic computing package and its components. The learning environment usually shows the user how to build sessions, called worksheets, and write programs in the computing packages language that can be used not only by the user but also by other users. The learning environment usually describes how to add features contained in the software package's libraries. Many textbooks are available for learning Maple, Mathematica, or MATLAB. There are also many mathematical, scientific, and engineering textbooks that use them as part of the text and have exercises for which the reader must use one of these packages.

Worksheets. One strength of symbolic computing packages is their ability to allow the user to build worksheets or programs that other users can access as a tool or as a component. This allows teachers, for example, to build a fully documented worksheet. The teacher can then have students work through the worksheet to learn some mathematical or scientific concept. The worksheet can contain all the steps in sequence for the student to go through one at a time. Symbolic computing packages allow the user to put text in the worksheet in textbook format so that it is easy for others to see how to use the worksheet. The packages allow mathematical symbols to be typed in textbook format in the text. The author of the worksheet can use this feature to detail the mathematics that the worksheet can perform. As another example, an engineer could build a worksheet in the symbolic computation package to use as a tool and then make this worksheet available to all the other engineers in the firm as a component to the symbolic computing package.

Most symbolic computing packages also allow the user to make personal worksheets available to everyone on the Internet. Each software package has a specific format for doing this. The Web sites of symbolic computing software package vendors offer worksheets that can be run over the Internet. Most of these are scientific sessions and usually have some sophisticated visualization that displays the symbolic and/or numerical calculation. *See* INTERNET.

James Sochacki

Bibliography. W. Gander and J. Hrebícek (eds.), *Solving Problems in Scientific Computing Using Maple and MATLAB*, 4th ed., 2004; S. Wolfram, *The Mathematica Book*, 5th ed., 2003.

Symmetrodonta

A group of extinct mammals that range from the Late Triassic to Late Cretaceous. Their fossil remains have been found in Mesozoic deposits worldwide. These small insectivorous or carnivorous mammals are the size of a shrew or mouse. They are considered to be distant relatives of the more derived therian mammals, including the extinct eupantotheres, such as dryolestids and the living placental and marsupial mammals.

Symmetrodonts are characterized by a distinctive feature: the main cusps of their cheek teeth are arranged in a symmetrical triangle. The stylocone, paracone, and metacone of the upper tooth form a triangle known as the trigon, while the protoconid, paraconid, and metaconid of the lower tooth form a triangle known as the trigonid. The triangular upper and lower cheek teeth fill in the gaps between the adjacent teeth of the opposite tooth row, and are specialized for crushing insects or slicing worms. However, symmetrodonts lack a basinlike heel in the lower teeth (known as the talonid basin) that would allow the grinding of the ingested food in the more derived living therian mammals and their kin. Symmetrodonts are also distinguished from more derived therians by lacking the angle on the mandible.

A newly discovered symmetrodont skeleton from China has a straight cochlea, an auditory structure in the ear that is less derived than those in living therians. Its shoulder girdle was mobile and more derived than those of living monotreme mammals. However, symmetrodonts retain a primitive interclavicle bone that no longer exists as a separate bone in the adults of living therian mammals. Its forelimb had a more splayed posture than those of the living therians.

Symmetrodonts are classified as the order Symmetrodonta of the class Mammalia. Three families are notable among the diverse symmetrodonts. The Kuehneotheriidae, found in the Upper Triassic to the Lower Jurassic deposits of Europe and India, are considered to be close to the ancestry of all other therian mammals. Kuehneotheriids have primitive feature characters in the mandible. The Amphidontidae are from the Upper Jurassic deposits of North America and England. The Spalacotheriidae, which are characterized by the highly acute triangular cheek teeth, are the most diverse symmetrodont family and have been found on most continents (except Antarctica) in the Upper Jurassic to the Upper Cretaceous sediments.

Zhexi Luo

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1987; R. L. Cifelli and S. K. Madsen, Spalacotheriid symmetrodonts (Mammalia) from the medial Cretaceous (upper Albian or lower Cenomanian) Mussentuchit local fauna, Cedar Mountain Formation, Utah, USA, *Geodiversitas*, 21(2):167–214, 1999; P. Ensom and D. Sigogneau-Russell, New symmetrodonts (Mammalia, Theria) from the Purbeck Limestone Group, Lower Cretaceous, southern England, *Cretaceous Res.*, 21:767–779, 2000; Y. Hu et al., A new

symmetrodont mammal from China and its implications for mammalian evolution, *Nature*, 390:137–142, 1997; M. C. McKenna and S. K. Bell, *Classification of Mammals Above the Species Level*, 1997; D. Sigogneau-Russell and P. Ensom, Thereuodon (Theria, Symmetrodonta) from the Lower Cretaceous of North Africa and Europe, and a brief review of symmetrodonts, *Cretaceous Res.*, 19:445–470, 1998.

Symmetry breaking

A deviation from exact symmetry. According to modern physical theory the fundamental laws of physics possess a very high degree of symmetry. Several deep insights into nature arise in understanding why specific physical systems, or even the universe as a whole, exhibit less symmetry than the laws themselves.

Spontaneous symmetry breaking. This mechanism occurs in quite diverse circumstances and brings a cluster of important observable consequences whenever it occurs. The most symmetrical solutions of the fundamental equations governing a given system may be unstable, so that in practice the system is found to be in a less symmetrical, but stable, state. When this occurs, the symmetry is said to have been broken spontaneously.

For example, the laws of physics are unchanged by any translation in space, but a crystalline lattice is unchanged only by special classes of translations. A crystal does retain a large amount of symmetry, for it is unchanged by those finite translations, but this falls far short of the full symmetry of the underlying laws. *See* CRYSTAL STRUCTURE.

Another example is provided by ferromagnetic materials. The spins of electrons within such materials are preferentially aligned in some particular direction, the axis of the poles of the magnet. The laws of physics governing the interactions among these spins are unchanged by any rotation in space, but the aligned configuration of spins has less symmetry. Indeed, it is left unchanged only by rotations about the polar axis. *See* FERROMAGNETISM.

In both these examples, the loss of symmetry is associated with the appearance of order. This is a general characteristic of spontaneous symmetry breaking.

These examples can provide some understanding of why spontaneous symmetry breaking is so common. For concreteness, ferromagnets will be discussed. The laws of physics operating between two spins attach a particularly low energy, and therefore special stability, to the configuration wherein both point in the same direction. This in itself does not break the rotation symmetry, because the spins can be rotated together by any amount, without changing the energy. If a third spin is added, it is found to be particularly stable when it is aligned with the previous two, and so forth. In this way, a system of many spins aligned in some particular direction is built up, for which rotation symmetry is effectively lost. For while small disturbances from the outside world, or even quantum fluctuations, will cause a system of a few spins to rotate as a whole, with many spins this becomes increasingly difficult, and eventually a definite stable direction is fixed, breaking the symmetry.

Consequences. As mentioned above, there is a cluster of important observable consequences associated with spontaneous symmetry breaking.

Nambu-Goldstone bosons. This is a class of low-energy excitations associated with gentle variations of the order. In the case of ferromagnets, an overall rotation of all the spins at once does not change the energy. If the direction of alignment is changed slowly in space, each spin will be nearly aligned with its neighbors, and the overall cost in energy will be small. Thus, there is a class of excitations of the ferromagnet, the magnons, that exist as a consequence of the spontaneous symmetry breaking, and that have very low energy. Similarly, in the case of crystals, phonons are associated with gentle distortions of the lattice structure. *See* MAGNON; PHONON.

Phase transitions. At high temperatures the energy gained by assuming an ordered structure is increasingly outweighed by the entropy loss associated with the constraints it imposes, and at some point it will no longer be favorable to have spontaneous symmetry breaking in thermal equilibrium. Changes from broken symmetry to unbroken symmetry are marked by phase transitions. For a magnet, the transition occurs (by definition) at the Curie temperature, above which the material no longer displays a magnetic moment. For a crystal, it is melting into a liquid or sublimation into a gas. *See* CURIE TEMPERATURE; ENTROPY; PHASE TRANSITIONS; THERMODYNAMIC PRINCIPLES.

Defects. These are imperfections in the ordering. The most familiar examples are domain walls in magnets. On either side of the domain wall the spins are aligned, but the direction of alignment changes in crossing the wall. *See* CRYSTAL DEFECTS; DOMAIN (ELECTRICITY AND MAGNETISM).

Systems with long-range forces. In systems with long-range forces as well as spontaneous symmetry breaking, it need no longer be true that gradual changes require only a small input of energy, because even distant regions interact significantly. Thus, the Nambu-Goldstone bosons no longer have very low energies, and they are not easily excited. Conversely, the system will exhibit a special rigidity, with strong correlations between distant points. These ideas are central to modern theories of superconductivity and of particle physics (the Higgs mechanism). *See* ELECTROWEAK INTERACTION; STANDARD MODEL; SUPERCONDUCTIVITY.

Other forms. There are several other sources of symmetry breaking in nature, besides spontaneous symmetry breaking.

Frozen accidents. Sometimes the processes tending to enforce an equilibrium configuration that has some symmetry are so slow that, for practical purposes, this configuration is never achieved. Then so-called frozen accidents occurring during the formation of the system are locked in, and generally ruin its

potential symmetry. Examples are provided by glasses and by the handedness of biological molecules. *See* GLASS; MOLECULAR ISOMERISM; STEREOCHEMISTRY.

Asymptotic symmetries. Sometimes symmetries are imperfect, but are in a well-defined sense as good as they can be, subject to limitations imposed by fundamental laws of physics. Under resizing symmetry, also known as scale invariance, physical laws operate in the same way if the sizes of all bodies are multiplied by a common factor. In quantum chromodynamics, resizing symmetry becomes more nearly valid as smaller subatomic objects are considered. Perfect resizing symmetry could be true only in a physically trivial theory, but quantum chromodynamics comes as close as a nontrivial theory can. In inflationary universe models, the universe is as homogeneous as it can be, subject to irreducible quantum fluctuations in the original state. *See* INFLATIONARY UNIVERSE COSMOLOGY; QUANTUM CHROMODYNAMICS.

Symmetry limited by partial vision. A symmetry may appear to be broken to an observer with only a partial view of the system under consideration. For example, a magnet just above its Curie temperature has large regions where nonzero average alignments persist for long times, and only by averaging over extremely large regions or very long times are averages to zero obtained.

In quantum mechanics, a highly symmetric wave function may describe, when applied to any particular set of experiments, much less symmetry. For example, the wave function describing a photon emitted in atomic decay may be completely symmetric under rotations. Nevertheless, any particular measurement of the photon finds it moving in some definite direction, breaking the symmetry. According to one interpretation of quantum mechanics, the wave function potentially describes many worlds—that is, many possible results of macroscopic experiments—and the limited vision of occupants of any one world prevents them from seeing the complete symmetry. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; SYMMETRY LAWS (PHYSICS). Frank Wilczek

Bibliography. P. Anderson, *Basic Notions of Condensed Matter Physics*, 1984; M. Mezard, G. Parisi, and M. Virasoro, *Spin Glass Theory and Beyond*, 1987.

Symmetry laws (physics)

The physical laws which are expressions of symmetries. The term symmetry, as it is used in mathematics and the exact sciences, refers to a special property of bodies or of physical laws, namely that they are left unchanged by transformations which, in general, might have changed them. For example, the geometric form of a sphere is not changed by any rotation of the sphere around its center, and so a sphere can be said to be symmetric under rotations. Symmetry can be very powerful in constraining form. Indeed, referring to the same example, the only sort of surface which is symmetric under arbitrary rotations is a sphere.

The concept that physical laws exhibit symmetry is more subtle. A naive formulation would be that a physical law exhibits symmetry if there is some transformation of the universe that might have changed the form of the law but in reality does not. This formulation is much too rigid, however, since the comparison of different universes is generally not feasible or desirable. A more fruitful definition of the symmetry of physical law exploits a profound feature of successful descriptions of nature: locality. Locality is the principle that the behavior of a given system is only slightly affected by the behavior of other bodies far removed from it in space or time. Because of locality, it is possible to define symmetry by using transformations that do not involve the universe as a whole but only a suitably isolated portion of it. Thus the statement that the laws of physics are symmetric under rotations means that (say) astronauts in space would not be able to orient themselves—to determine a preferred direction—by experiments internal to their space station. They could do this only by referring to weak effects from distant objects, such as the light of distant stars or the small residual gravity of Earth. Any experiment that is not sensitive to these small effects will give the same result, whether it is done before or after a rotation of the space station and its contents as a whole.

Symmetries of space and time. Perhaps the most basic and profound symmetries of physical laws are symmetry under translation in time and under translations in space.

The statement that fundamental physical laws are symmetric under translation in time is equivalent to the statement that these laws (unlike human laws) do not change or evolve. Time-translation symmetry is supposed to apply, fundamentally, to simple isolated systems. Large complicated systems, and in particular the universe as a whole, do of course age and evolve. Thus in constructing the big-bang model of cosmology, it is assumed that the properties of individual electrons or protons do not change in time, although of course the state of the universe as a whole, according to the model, has changed quite drastically.

More precise and controlled tests of time-translation symmetry are based on observations of the spectral lines in light from distant stars, which was emitted long ago. Such observations indicate that the atoms in distant stars long ago obeyed, with great accuracy, the same laws as the atoms of the same elements on Earth do today. Other tests involve searches for anomalies in the Moon's orbit, which could indicate changes in the strength of gravity; and geochemical evidence involving naturally occurring nuclear reactions early in the Earth's history, whose character is highly sensitive to the powers of the fundamental interactions.

The statement that fundamental laws are symmetric under translations in space is another way of formulating the homogeneity of space. It is the

statement that the laws are the same throughout the universe. It says that the astronauts in the previous example cannot infer their location by local experiments within their space station. To do this, they must look to weak effects from distant objects. The power of this symmetry is that it makes it possible to infer, from observations in laboratories on Earth, the behavior of matter anywhere in the universe. Used in conjunction with symmetry under time translation, it is the foundation of astrophysics, and indeed of all scientific predictions about the behavior of matter at places and times besides here and now.

The symmetry of physical law under rotations, mentioned above, embodies the isotropy of space. It is very accurately tested by measurements of radiation from atoms, which reveal no preferred direction.

In the mathematical formulation of dynamics, there is an intimate connection between symmetries and conservation laws. Symmetry under time translation implies conservation of energy; symmetry under spatial translations implies conservation of momentum; and symmetry under rotation implies conservation of angular momentum. *See* ANGULAR MOMENTUM; CONSERVATION LAWS (PHYSICS); CONSERVATION OF ENERGY; CONSERVATION OF MOMENTUM.

The fundamental postulate of the special theory of relativity, that the laws of physics take the same form for observers moving with respect to one another at a fixed velocity, is clearly another statement about the symmetry of physical law. The idea that physical laws should be unchanged by such transformations was already clearly discussed by Galileo, who illustrated it by an observer's inability to infer motion while on a calm sea voyage in an enclosed cabin. The novelty of Einstein's theory arises from combining this velocity symmetry with a second postulate, deduced from experiments, that the speed of light is a universal constant and must take the same value for both stationary and uniformly moving observers. To accommodate both postulates, it was necessary to recognize that what the moving observer identifies as the interval of time between two events is a mixture of the distance and time intervals identified by the stationary observer. In the mathematical formulation of the theory, the space-time transformations which implement velocity symmetry (the Lorentz transformations) are closely analogous to the transformations among purely spatial coordinates which implement ordinary rotation symmetry. *See* GALILEAN TRANSFORMATIONS; LORENTZ TRANSFORMATIONS; RELATIVITY.

The existence of objects with particular locations invalidates complete translation symmetry, but for crystals a large and useful residue remains. A very small displacement of a crystal lattice does not lead to an identical object, but a finite translation, which takes the lattice into itself, does. Whereas complete translation symmetry implies the conservation of momentum (as mentioned above), the partial symmetry characteristic of a lattice gives a weaker conservation law. Specifically, the crystal symmetry allows a discrete set of possible changes in momentum for electromagnetic radiation that scatters from it. However, insofar as the crystal does not change its state, it provides a time-translation invariant background, and the energy of the radiation (that is, its frequency and wavelength) cannot change. Thus x-rays scattering from a crystal are deflected in a discrete set of directions, determined by the underlying crystal symmetry. This is the principle underlying the determination of crystal structures by x-ray diffraction. *See* CRYSTAL STRUCTURE; X-RAY DIFFRACTION.

Discrete symmetries. Before 1956, it was believed that all physical laws obeyed an additional set of fundamental symmetries, denoted P, C, and T, for parity, charge conjugation, and time reversal, respectively. At that time, subtle experiments involving the particles known as K mesons led to the suggestion that P might be violated in the weak interactions, and violations were indeed observed. This discovery eventually led to questioning—and in some cases overthrow—of other cherished symmetry principles.

Parity, P, roughly speaking, transforms objects into the shapes of their mirror images. If P were a symmetry, the apparent behavior of the images of objects reflected in a mirror would also be the actual behavior of corresponding real objects. This abstract concept can be made concrete by referring to a specific experimental arrangement, which tests the validity of P symmetry. A K^+ meson is observed to decay into a positron and a neutrino. The positron has the property of spin, which for present purposes may be visualized as an intrinsic rotation of the positron about an axis passing through its center. It is observed that positrons emitted in K^+ decay are almost always right-handed in the sense that their rotation is in the direction the fingers of a right hand curl if the thumb points in the direction of motion. *See* SPIN (QUANTUM MECHANICS).

If this whole situation were reflected in a mirror, the image positrons would appear to be left-handed. But the real positrons were right-handed. Hence the mirror-image particles do not behave in the same way as real particles. Thus P is not a valid symmetry of physical law. *See* PARITY (QUANTUM MECHANICS).

Charge conjugation, C, changes particles into their antiparticles. It is a purely internal transformation; that is, it does not involve space and time. If the laws of physics were symmetric under charge conjugation, the result of an experiment involving antiparticles could be inferred from the corresponding experiment involving particles. Thus, from the fact that K^+ decays into right-handed positrons and neutrinos, it would be inferred from C symmetry that K^- decays into right-handed electrons and antineutrinos at the same rate. In fact, the decay of K^- into electrons and antineutrinos does occur at very nearly the predicted rate, but the electrons are left-handed. Thus C is not a valid symmetry of physical law.

Remarkably, by combining the transformations P and C, thus simultaneously performing a mirror reflection in space and a change of particles into antiparticles, a result is obtained, CP, which is much

more nearly a valid symmetry than either of its components separately. However, in 1964 it was discovered experimentally that even *CP* is not quite a valid symmetry.

Although the preceding discussion has emphasized the failure of *P*, *C*, and *CP* to be precise symmetries of physical law, both the strong force responsible for nuclear structure and reactions and the electromagnetic force responsible for atomic structure and chemistry do obey these symmetries. Only the weak force, responsible for beta radioactivity and some relatively slow decays of exotic elementary particles, violates them. Thus these symmetries, while approximate, are quite useful and powerful in nuclear and atomic physics. Furthermore, the violation of these symmetries within the weak interactions is intricately patterned and structured. Elucidating this pattern was a key step in arriving at the modern theory of these interactions. It was found, in a sense that can be made precise, that *P* and *C* symmetry violation by the weak interaction is maximal. Indeed, at a fundamental level, only left-handed quarks and leptons, and right-handed antiquarks and antileptons, participate in this interaction. *See* ELECTROWEAK INTERACTION; FUNDAMENTAL INTERACTIONS; LEPTON; QUARKS; WEAK NUCLEAR INTERACTIONS.

The operation of time-reversal symmetry, *T*, involves changing the direction of motion of all particles. For example, it relates reactions of the type $A + B \rightarrow C + D$ to their reverse $C + D \rightarrow A + B$. No direct violation of *T* has been detected. A powerful test of *T* symmetry is obtained by studying the response of neutrons to an electric field. If the spin axes of neutrons were to align with the direction of an applied electric field, *T* symmetry would be violated, since reversing the direction of motion changes the direction of spin. So far, no such effect has been found. *See* TIME REVERSAL INVARIANCE.

Time-reversal symmetry, even if valid, applies in a straightforward way only to elementary processes. It does not, for example, contradict the one-way character of the second law of thermodynamics, which states that entropy can only increase with time. The consequences of *T* for macroscopic systems are built up from its microscopic foundation, by using the locality of physical interactions. They include the Onsager reciprocity relations, which relate different transport processes. *See* THERMODYNAMIC PRINCIPLES; TIME, ARROW OF.

Fundamental principles of quantum field theory suggest that the combined operation *PCT*, which involves simultaneously reflecting space, changing particles into antiparticles, and reversing the direction of time, must be a symmetry of physical law. Existing evidence is consistent with this prediction. *See* CPT THEOREM.

Internal symmetry. Internal symmetries, like *C*, do not involve transformations in space-time but change one type of particle into another. An important, although approximate, symmetry of this kind is isospin or i-spin symmetry. It is observed experimentally that the strong interactions of the proton and neutron are essentially the same. This regularity explains, for example, why the most stable light nuclei contain equal numbers of protons and neutrons. Such combinations are the most symmetric, and therefore it is understandable that they are the most energetically favorable. *See* I-SPIN.

Now, the most long-range component of the strong interaction can be considered as being generated by the exchange of π mesons or pions. Charged π^+ and π^- mesons are relatively long-lived, and were discovered soon after pions were proposed. However, it was shown theoretically that these two mesons were not enough to allow a proper symmetry transformation that transformed neutrons into protons and vice versa, and also transformed π mesons among themselves, in such a way as to leave the strong interaction invariant. A third, electrically neutral π^0 meson was therefore proposed, and a particle with the predicted properties was duly found. *See* MESON.

This was the prototype of several other successful predictions of the existence and properties of new particles, based on postulates of internal symmetries. Perhaps the most notable was the prediction of the mass and properties of the Ω^- baryon, based on an extension of the symmetry group SU(2) of isospin to a larger approximate SU(3) symmetry acting on strange particles as well. These symmetries were an important hint that the fundamental strong interactions are at some level universal, that is, act on all quarks in the same way, and thus paved the way toward modern quantum chromodynamics, which does implement such universality. *See* BARYON; GROUP THEORY; UNITARY SYMMETRY.

Much simpler mathematically than SU(2) internal symmetry, but quite profound physically, is U(1) internal symmetry. It illustrates the general relationship between symmetry and conservation laws alluded to above, for it is uniquely related to the conservation of some quantum number. For concreteness, the important case of the electric charge quantum number will be considered. The action of the U(1) internal symmetry transformation with parameter λ is to multiply the wave function of a state of electric charge q by the factor $e^{i\lambda q}$. An amplitude between two states with electric charges q_r and q_s will therefore be multiplied by a factor $e^{i\lambda(q_s-q_r)}$. Since the physical predictions of quantum mechanics depend on such amplitudes, these predictions will be unchanged only if the phase factors multiplying all nonvanishing amplitudes are trivial. This will be true, in turn, only if the amplitudes between states of unequal charge vanish; that is, if charge-changing amplitudes are forbidden, which is just a backhanded way of expressing the conservation of charge. Conversely, given an additive conservation law, the same argument could be reversed to construct an internal U(1) symmetry. Thus symmetry under an internal U(1) transformation is equivalent to the existence of an additive conservation law. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; QUANTUM NUMBERS.

Localization of symmetry. The concept of local gauge invariance, which is central to the standard model of fundamental particle interactions, and in

a slightly different form to general relativity, may be approached as a generalization of the U(1) internal symmetry transformation, where a parameter λ independent of space and time appears. Although it is mathematically consistent to have such a parameter, it goes against the spirit of locality, according to which each point in space-time has a certain independence. There is therefore reason to consider the possibility of a more general symmetry, involving a space-time-dependent transformation in which the wave function is multiplied by $e^{i\lambda(x,t)q(x,t)}$, where $q(x,t)$ is the density of charge at the space-time point (x,t). These transformations are much more general than those discussed above, and invariance under them leads to much more powerful and specific consequences.

It might be expected that a space-time-dependent symmetry requires the amount of charge at every point in space to be conserved separately; that is, charge can never move. If, in addition, symmetry is demanded under velocity transformations, as required by special relativity, then the charge must simply vanish everywhere. Although this is not valid for electromagnetic charge, it does crudely describe the behavior of color charges in the strong interaction. Indeed, the fundamental quarks and gluons, which carry color charges, cannot exist as free particles—they are said to be confined—essentially as a result of local color symmetry. *See* GLUONS; QUANTUM CHROMODYNAMICS.

For electromagnetism, the situation is more subtle. Roughly speaking, charged particles can move, consistent with the space-time-dependent symmetry transformation specified above, if they leave behind a compensating disturbance in the electromagnetic field. Indeed, if this very restrictive and demanding idea is implemented, the required interactions of matter with the electromagnetic field are predicted precisely. Thus, the theory of the electromagnetic field—Maxwell's equations and quantum electrodynamics—can be said to be the unique ideal embodiment of the abstract concept of a space-time-dependent symmetry, that is, of local gauge symmetry. *See* MAXWELL'S EQUATIONS; QUANTUM ELECTRODYNAMICS.

As mentioned above, it is similarly true that the modern theory of the strong interaction arises from implementing a conceptually simple generalization of this symmetry. A mathematical discussion of isospin symmetry involves symmetry transformations that are rotations in a two-dimensional space. The approximate conservation laws and relations among interactions required for invariance under this symmetry follow from a generalization of U(1) symmetry, with parameter λ, where λ is allowed to become a fixed 2×2 (hermitian) matrix. By analogy, it is interesting to consider allowing this λ-matrix to depend on space and time. This sort of theory was first proposed in 1954, but had serious difficulties. These were resolved by later theoretical work, especially in the early 1970s, that led directly to the modern standard model. *See* STANDARD MODEL.

Likewise, general relativity is closely associated with localization of symmetries, namely the symmetries of space-time translation, rotation, and velocity. The construction of a theory invariant under this wide symmetry group involves new issues, and in particular the specification of a preferred space-time field for the measure of distance, the metric $g_{\mu\nu}$. Nevertheless, in spite of important technical differences between gravity and the other interactions, the forms of all the known fundamental interactions can be said to be dictated by local symmetry principles. *See* GAUGE THEORY.

Guidance to physical law. As mentioned above, symmetry arguments have played an important, if implicit, role in the discussion of physical law since the beginnings of modern science. The conscious use of symmetry as a guide to physical law is, however, a twentieth-century phenomenon. Symmetry provided guidance toward the current synthesis, the standard model, discussed above. Today symmetry dominates much thinking about possible new physics beyond the standard model.

Modern gauge theories of the strong, weak, and electromagnetic interactions have very similar mathematical structures, suggesting that their symmetries might be synthesized into a larger encompassing symmetry. The idea is that the larger "unified" symmetry is spontaneously broken, so that while it is not valid for experiments done at the comparatively low energies which are practically accessible, it becomes more accurate for experiments at high energy or in describing events during the early moments of the big bang. *See* GRAND UNIFICATION THEORIES; SYMMETRY BREAKING.

Supersymmetry postulates invariance under transformations which change the spin of elementary particles. As for unifying gauge symmetries, it is fruitful to consider the possibility that supersymmetry, although not accurately valid, is spontaneously broken. There is striking semiquantitative evidence that the observed strong, weak, and electromagnetic interactions derive from spontaneous breakdown of a theory that is not only unified but also supersymmetric. Finally, superstring theory promises to incorporate these ideas in a framework that also allows for unification with the gravitational interaction. *See* ELEMENTARY PARTICLE; SUPERSTRING THEORY; SUPERSYMMETRY.

Frank Wilczek

Bibliography. S. Dimopoulos, S. Raby, and F. Wilczek, Unification of couplings, *Phys. Today*, 44(10):25–33, October 1991; J. E. Lannutti and P. K. Williams (eds.), *Current Trends in the Theory of Fields: A Symposium in Honor of P. A. M. Dirac*, 1978; J. Sakurai, *Invariance Principles and Elementary Particles*, 1969; E. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra*, 1959.

Symmorphosis

A theory of structural design of biological organisms that postulates that structure is quantitatively matched to functional demand as a result of regulated

morphogenesis during growth and maintenance. Symmorphosis is a theory of economic design. In biological organisms, all functions depend on structural design, specifically on the morphometric characteristics of the organs. In general terms, the larger the structure the greater the functional capacity. A central postulate of symmorphosis, as a theory of economic design, is that differences in the functional demand on an organ require quantitative adjustments of its structural design parameters in order to match functional capacity to (maximal) functional demand.

Many functional systems comprise a sequence of steps, each supported by a different set of structures. For example, to support cell energetics oxygen must be supplied to the cells; this involves a chain of events from gas exchange in the lung, through cardiovascular transport of blood, to oxidative metabolism on the membranes of the cells' mitochondria. Symmorphosis postulates that quantitative design of each step is matched to overall functional demand. There should be enough structure but not too much, as it would be wasteful to design one step vastly in excess of the others.

The notion that animals, and humans, should be designed economically follows from common sense, but it is also supported by many observations. Thus blood vessel architecture ensures blood flow distribution with minimal energy loss; bone structure is patterned according to stress distribution and also quantitatively adapted to total stress; with training, athletes can specifically adjust the structure of their muscles and of their cardiovascular system to higher functional demands, and these modifications are soon reversed when training is stopped. These examples suggest optimization of structural design based on a cost-to-benefit relationship: the larger an organ, the more it will cost in terms of construction, maintenance, and load on the body. However, these examples are anecdotal, and one can clearly come up with similar anecdotal evidence that proves the opposite.

The validity of the concept of economic design as an important principle of structure-function relationship in biology can be tested by the hypothesis of symmorphosis which predicts that if the functional requirements on an organ system vary, structural design should vary in parallel. This hypothesis has been tested mainly on the mammalian respiratory system, where large natural variations in functional capacity are observed: the maximal rate of oxidative metabolism is much larger in small animals than in large animals (allometric variation) and in athletic species compared to more sedentary species (adaptive variation). The general result is that the hypothesis of symmorphosis is acceptable for most steps of the functional cascade, but the lung is a case where a certain, though not excessive, redundancy in design is noted.

The theory of symmorphosis postulates ideal adaptation of structural design to functional capacity. This is, however, hardly a reasonable assumption, because good engineering design of complex systems requires some redundancies as safety factors in view of imperfections in functional performance and variable boundary conditions. By using the concept of symmorphosis as a hypothesis, such deviations from idealized economic design can be detected.

Ewald R. Weibel

Bibliography. E. R. Weibel, C. R. Taylor, and H. Hoppeler, The concept of symmorphosis: A testable hypothesis of structure-function relationship, *Proc. Nat. Acad. Sci. USA*, 88:10357–10361, 1991; E. R. Weibel, C. R. Taylor, and H. Hoppeler, Variations in function and design: Testing symmorphosis in the respiratory system, *Resp. Physiol*, 87:325–348, 1992.

Sympathetic nervous system

The portion of the autonomic nervous system concerned with nonvolitional preparation of the organism for emergency situations. *See* AUTONOMIC NERVOUS SYSTEM.

The sympathetic nervous system is best understood in mammals. This system consists of two neuron chains from the thoracic and lumbar regions of the spinal cord to viscera and blood vessels. The first or preganglionic neuron has its cell body in the spinal cord and sends its axon to synapse with a postganglionic sympathetic neuron, which lies either in a chain of sympathetic ganglia paralleling the spinal cord or in a sympathetic ganglion near the base of the large blood vessels vascularizing the alimentary

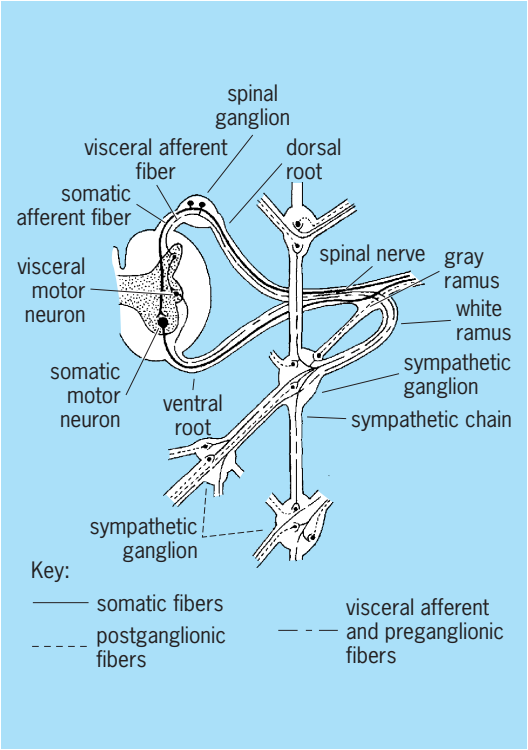

Fig. 1. Visceral reflex arc and the sympathetic chain. (*After B. A. Houssay et al., Human Physiology, 2d ed., McGraw-Hill, 1955*)

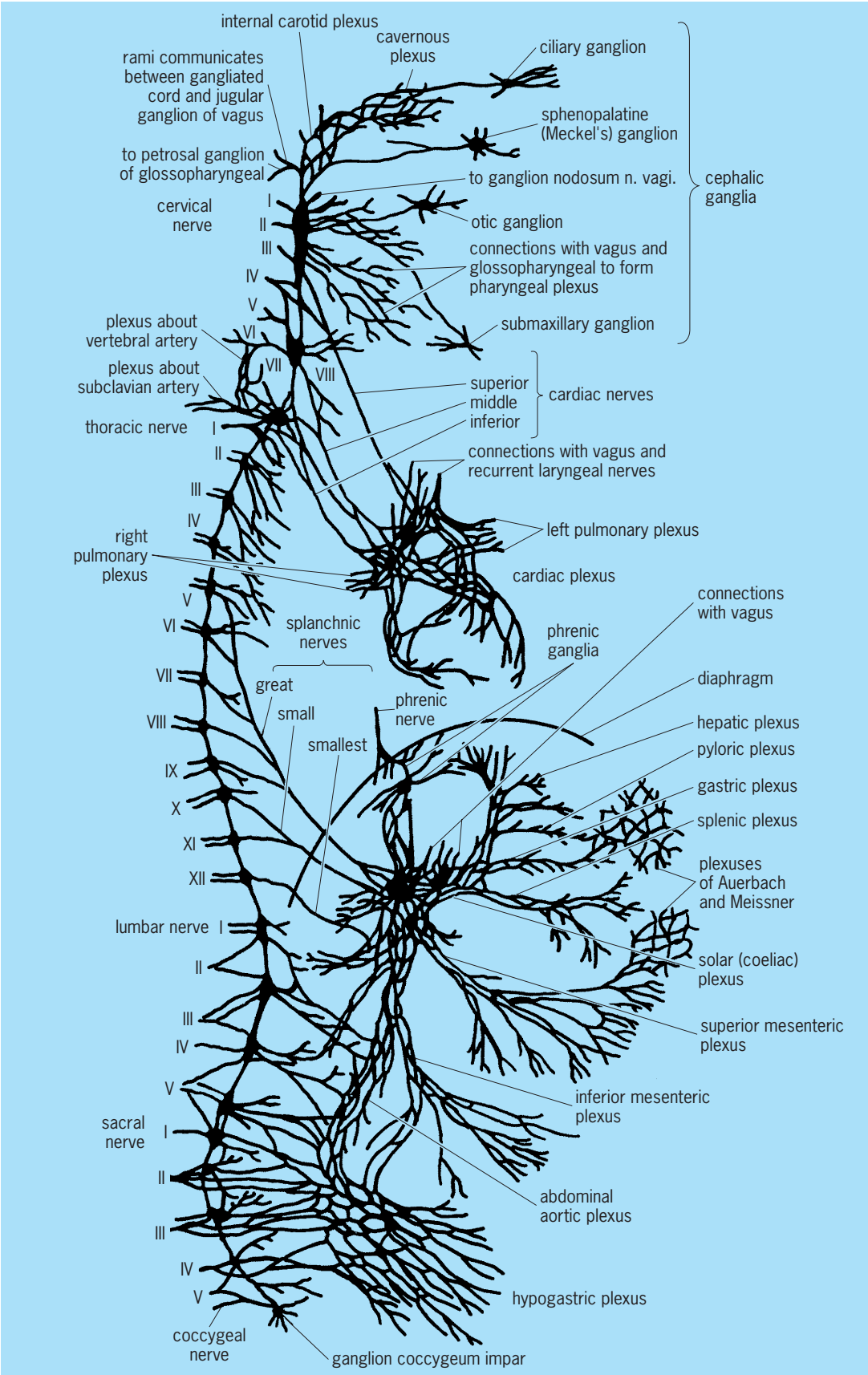

Fig. 2. Diagram showing the chain of ganglia and the great plexuses of the sympathetic system. (*After J. P. Schaeffer, ed., Morris' Human Anatomy, 11th ed., Blakiston-McGraw-Hill, 1953*)

viscera. The postganglionic axons are longer than the preganglionic axons and extend to glands or smooth muscles of viscera and blood vessels (**Fig. 1**). Sensory visceral nerve fibers innervate blood vessels and viscera and carry sensory information to the spinal cord, thus providing a visceral reflex arc (**Fig. 2**).

Among nonmammalian vertebrates, the sympathetic nervous system of birds and reptiles is similar to that in mammals, but it is somewhat different in anamnoites. In amphibians and bony fish the sympathetic neurons arise from the entire length of the cord. In cartilaginous fish (such as sharks) and agnathans (such as lampreys) the sympathetic nervous system is still not well understood and apparently not discretely separated. *See* NERVOUS SYSTEM (VERTEBRATE); PARASYMPATHETIC NERVOUS SYSTEM.

Douglas B. Webster

Sympathetic vibration

The driving of a mechanical or acoustical system at its resonant frequency by energy from an adjacent system vibrating at this same frequency. Examples include the vibration of wall panels by sounds issuing from a loudspeaker, vibration of machinery components at specific frequencies as the speed of a motor increases, and the use of tuned air resonators under the bars of a xylophone to enhance the acoustic output. Increasing the damping of a vibrating system will decrease the amplitude of its sympathetic vibration but at the same time widen the band of frequencies over which it will partake of sympathetic vibration. *See* MECHANICAL VIBRATION; RESONANCE (ACOUSTICS AND MECHANICS); VIBRATION.

Lawrence E. Kinsler

Symphyla

A class of the Myriapoda. The symphylans, like the pauropods, are tiny, pale, centipedelike creatures that inhabit humus or soil, or live under debris; in general, they live wherever there is sufficient moisture to preclude excessive water loss. They are similar to the Pauropoda and Diplopoda in being progoneate (having reproductive tracts open at anterior end of body) and anamorphic (gradually increasing in complexity of form and function). Each of their mandibles, like those of millipedes, but unlike the simple pauropod mandible, bears a movable gnathal lobe; at the same time their two pairs of maxillae are more reminiscent of the chilopods and lower insects than of the singly maxillate millipedes and pauropods. Additional signal characteristics include the following: The antennae are unbranched and simple; there is 1 pair of spiracles arising in the head and opening into tracheae; there are 12 pairs of legs, 1 pair per body segment; most of the legs have peculiar basal eversible vesicles with associated styli: the tergites number at least 15 and do not form diplotergites; there is a prominent pair of terminal spinnerets.

As is the case for the Pauropoda, little is known about symphylan biology. It is established that they feed upon decaying material as well as upon living plants; their role as greenhouse pests is widely appreciated by agriculturalists. Presumably all symphylans hatch with a reduced number of legs, six or seven in the species investigated, and thereafter undergo molts not only until the adult complement is gained but throughout life, which means 4–5 years in some forms.

The class consists of three families to which not more than 60 species have been assigned.

Ralph E. Crabill

Bibliography. G. Eisenbeis and W. Wichard, Symphyla, pp. 170–179 in *Atlas on the Biology of Soil Arthropods*, Springer-Verlag, Berlin, 1987; G. O. Poinar and C. A. Edwards, First description of a fossil symphylan, *Scutigerella dominicana* sp. n. (Scutigerellidae, Symphyla), in Dominican amber, *Experientia*, 51:391–393, 1995; U. Scheller, A new troglobitic species of Hanseniella Bagnall (Symphyla: Scutigerellidae) from Tasmania, *Austral. J. Entomol.*, 35:203–207, 1996; U. Scheller and J. A. Adis, Pictorial key for the Symphylan families and genera of the Neotropical region south of Central Mexico (Myriapoda, Symphyla), *Stud. Neotrop. Fauna Environ.*, 31:57–61, 1996.

Synapsida

The clade that includes all living and extinct species that share a closer relationship with modern mammals than with modern reptiles. The evolutionary history of synapsids is usually described in terms of three major radiations. The oldest synapsids are Middle Pennsylvanian in age (approximately 320 million years ago [Ma]) and belong to the earliest, "pelycosaur" radiation. Pelycosaur-grade synapsids are known primarily from Texas, Oklahoma, and surrounding states, although several fossils have been discovered in Europe. By the Middle Permian (260 Ma), a group of advanced synapsids, the therapsids, had achieved nearly complete distribution across the supercontinent of Pangea. As with their pelycosaur forebears, therapsids diversified into both carnivorous and herbivorous species, some growing up to 3–4 m (10–13 ft) in length.

Therapsid fossils are best known from Russia and South Africa, although finds in Antarctica, Brazil, China, India, and North America demonstrate their cosmopolitan nature. At the end of the Permian Period (251 Ma), a major mass extinction killed off numerous therapsid groups and left others with far fewer representatives. One group of therapsids that did survive, however, was the cynodonts. This group is of particular importance because it includes the ancient ancestors of mammals. The final radiation of synapsids is the mammals, which first appear in the fossil record in the Late Triassic and continue to the present. Traditional taxonomies group the

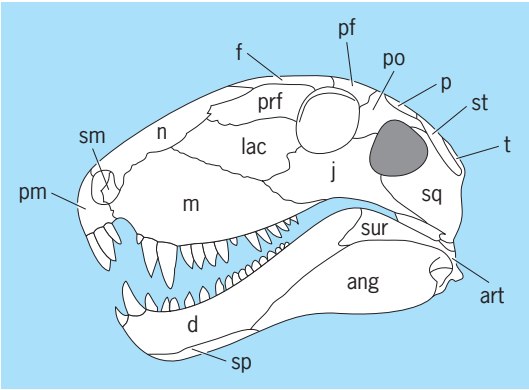

Skull of the pelycosaur-grade synapsid *Dimetrodon* in lateral view. The shaded region represents the lateral temporal fenestra, a diagnostic feature of all synapsids, including modern mammals. Abbreviations: ang, angular; art, articular; d, dentary; f, frontal; j, jugal; lac, lacrimal; m, maxilla; n, nasal; p, parietal; pf, postfrontal; pm, premaxilla; po, postorbital; prf, prefrontal; sm, septomaxilla; sp, splenial; sq, squamosal; st, supratemporal; sur, surangular; t, tabular. (*Reconstruction modified from B. S. Rubidge and C. A. Sidor, Evolutionary patterns among Permo/Triassic therapsids, Annu. Rev. Ecol. Systemat., 32:449–480, 2001*)

pelycosaurs and therapsids together as the "mammal-like reptiles," emphasizing that they possess some, but not all, of the features of mammals. All synapsid skulls, however, are characterized by a single temporal opening, bordered in early forms by the jugal, squamosal, and postorbital bones (see **illustration**). *See* ANIMAL EVOLUTION; MAMMALIA; PELYCOSAURIA; REPTILIA; THERAPSIDA. Christian A. Sidor

Bibliography. M. J. Benton, *Vertebrate Palaeontology*, 2005; R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1988.

Synaptic transmission

The physiologic mechanisms by which one nerve cell (neuron) influences the activity of an anatomically adjacent neuron with which it is functionally coupled. Brain function depends on interactions of nerve cells with each other and with the gland cells and muscle cells they innervate. The interactions take place at specific sites of contact between cells known as synapses. The synapse is the smallest and most fundamental information-processing unit in the nervous system. By means of different patterns of synaptic connections between neurons, synaptic circuits are constructed during development to carry out the different functional operations of the nervous system. In order to understand nervous function, it is first necessary to understand the mechanism of transmission at a single synapse, and then the varieties of synapses out of which synaptic circuits are formed. *See* NEURON.

Electrical synapses. The simplest type of synapse is the electrical synapse (**Fig. 1***a*), which consists of an area of unusually close contact between two cells packed with channels that span the two membranes and the cleft between them. Because the cleft persists at these sites, electrical synapses are also known as gap junctions. Electrical and metabolic communication between two cells is established by the components of the gap junctions. Six subunits of the polypeptide connexin form a half-channel in each membrane, and the two half-channels are joined end to end across the intercellular gap to form the complete channel. A variety of influences, including calcium ions (Ca^{2+}), pH, membrane potential, neurotransmitters, and phosphorylating enzymes, may act on the channels to regulate their conductance in one direction (rectification) or both directions. For example, the channels are closed by high concentrations of calcium ions or by low pH.

Electrical synapses are present throughout the animal kingdom in a variety of functional settings. In vertebrates, they are numerous in the central nervous systems of fish as well as in certain nuclei of the mammalian brain, in regions where rapid transmission and synchronization of activity is important. Electrical synapses also interconnect glial cells in the brain. *See* BIOPOTENTIALS AND IONIC CURRENTS.

Chemical synapses. The most prevalent type of junction between nerve cells is the chemical synapse (Fig. 1*b*). At chemical synapses there are no gap junctions. Instead, neurotransmitters are released from the presynaptic cell, diffuse across the synaptic cleft, and bind to receptors on the postsynaptic cell. Chemical synapses are found only between nerve cells or between nerve cells and the gland cells and muscle cells that they innervate.

Neuromuscular junction. The neuromuscular junction is a prototypical synapse. The mechanisms of transmission at a single synapse have historically been most closely studied at the neuromuscular junction, that is, the junction between the axon terminals of a motoneuron and a muscle fiber (**Fig. 2**). Muscle, like nerve cells, is electrically excitable and contracts in response to depolarization. The mechanisms responsible for neurotransmission at this synapse are representative of chemical synapses in general. Three basic elements constitute the synapse: a presynaptic process (in this case, the motoneuron axon terminal) containing synaptic vesicles; an end plate (a specialized site of contact between the cells); and a postsynaptic process (in this case, the muscle cell). In motoneurons the nerve terminal is large and contains many individual active zones, each consisting of a cluster of presynaptic vesicles along a dense barlike structure. Opposite each of these is a fold of the postsynaptic membrane containing receptors for the transmitter substance released from the presynaptic terminal.

The chemical substance that serves as the transmitter at the vertebrate neuromuscular junction is acetylcholine (ACh). Within the nerve terminal, acetylcholine is concentrated within small spherical vesicles (Fig. 2). At rest, these vesicles undergo exocytosis at low rates, releasing their acetylcholine in quantal packets to diffuse across the cleft and bind to and activate the postsynaptic receptors. Each quantum gives rise to a small depolarization of the postsynaptic membrane. These miniature end-plate

potentials are rapid, lasting only some 10 milliseconds, and small in amplitude, only some 500 microvolts, below the threshold for effecting any response in the muscle. They represent the resting secretory activity of the nerve terminal.

When an organism wants to move its muscles, electrical impulses known as action potentials are generated in the motoneurons. These are conducted along the axon and invade the terminal, causing a large depolarization of the presynaptic membrane. This opens special voltage-gated channels for calcium ions, which enter the terminal and bind to special proteins, causing exocytosis of one or two vesicles at many active zones simultaneously. Calcium ions are thus the necessary and crucial link between the electrical signals in the presynaptic neuron and the chemical signals sent to the postsynaptic neuron. An action potential causes as many as 100 of the 500 or so active zones at the neuromuscular junction of the frog to release a quantum of acetylcholine. The combined action of this acetylcholine on postsynaptic receptors sets up a large postsynaptic depolarization known as the end-plate potential, which has a time course of 20–30 milliseconds and an amplitude of 40–50 millivolts. This exceeds the threshold for generating an impulse in the surrounding membrane, and causes the muscle to contract. The action of acetylcholine at its receptor is terminated by an enzyme, acetylcholinesterase, which is present in the synaptic cleft and hydrolyzes the acetylcholine to acetate and choline.

In order for neurotransmission to continue after a round of exocytosis, synaptic vesicles have to be regenerated. Vesicles are rapidly and efficiently reformed in the nerve terminal by endocytosis, the process by which cells internalize extra plasma membrane. The machinery responsible for this endocytosis must sort proteins to include in the reforming synaptic vesicle membrane, generate the inward curvature of the nascent vesicle, and finally pinch the neck of the bud to release the newly formed vesicle into the nerve terminal. Specific neurotransmitter transporters within the membrane then fill the synaptic vesicle with appropriate neurotransmitter. The regenerated vesicle either returns to the plasma membrane where it rejoins the releasable pool of vesicles, or remains in the nerve terminal as part of a reserve pool. *See* ACETYLCHOLINE; MUSCLE.

Molecular mechanisms. Synaptic transmission requires coordinated interactions among many proteins found in the sending (presynaptic) and receiving (postsynaptic) compartments of the synapse. In the synaptic terminal, these include a special category of proteins known as SNAREs (synaptosome-associated protein receptors), which are small membrane-anchored proteins that are essential for membrane fusion and exocytosis. Specific SNAREs are found on synaptic vesicles (synaptobrevin or VAMP, molecular weight 18,000) and the presynaptic plasma membrane (syntaxin, molecular weight 35,000, and SNAP-25, molecular weight 25,000). SNAREs form tight alpha-helical complexes with each other that are thought to pull the membranes

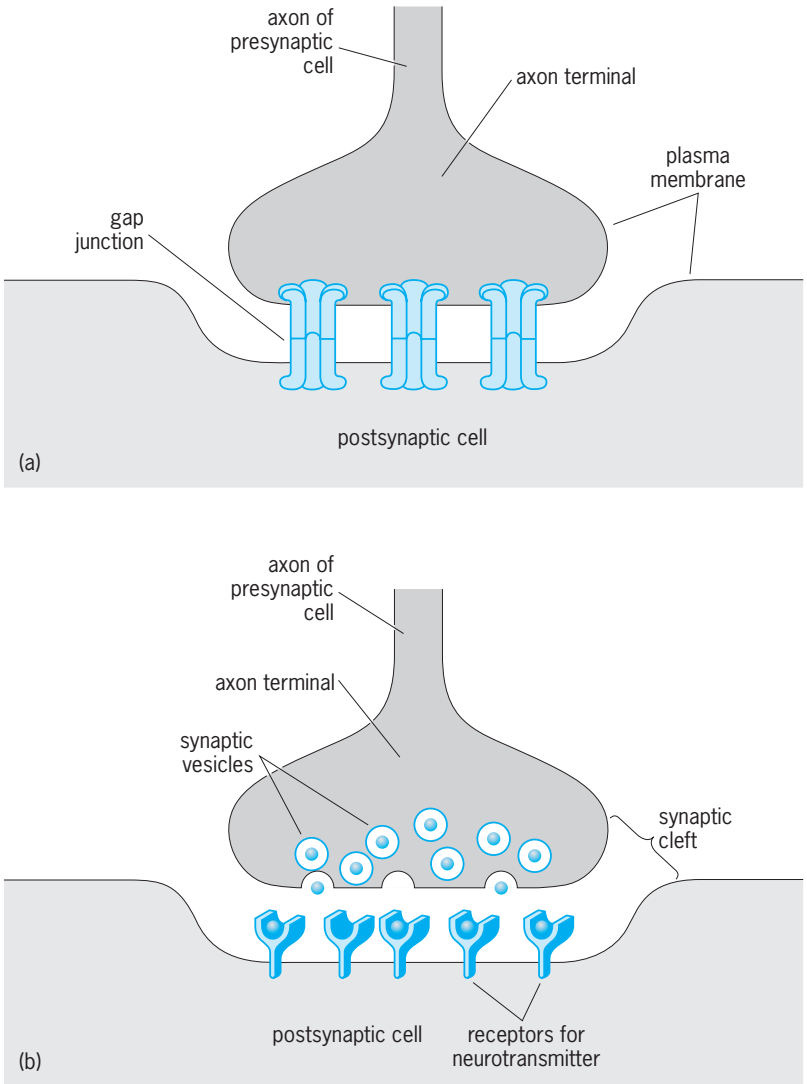

Fig. 1. Types of synapses. (*a*) An electrical synapse, showing the plasma membranes of the presynaptic and postsynaptic cells linked by gap junctions. (*b*) A chemical synapse, showing neurotransmitters released by the presynaptic cell diffusing across the synaptic cleft and binding to receptors on the postsynaptic membrane.

in which they are anchored close together and either facilitate or directly promote fusion, and hence exocytosis. A large number of other proteins, including a calcium-responsive trigger protein known as synaptotagmin, are also important in regulating vesicle exocytosis. Additional sets of proteins are responsible for endocytosis and regeneration of synaptic vesicles.

In the postsynaptic membrane, the target of the axon terminal's chemical signal is a neurotransmitter receptor; in the case of vertebrate muscle, this is the acetylcholine receptor, a membrane protein composed of five separate polypeptide chains or subunits. There are two alpha subunits, and beta, gamma, and delta subunits, totaling a molecular weight of approximately 260,000. The five subunits span the membrane and form a ring around a central Pore, which acts as a channel for the passage of ions across the membrane. When activated, the channel is cation-selective, allowing sodium (Na^+) and

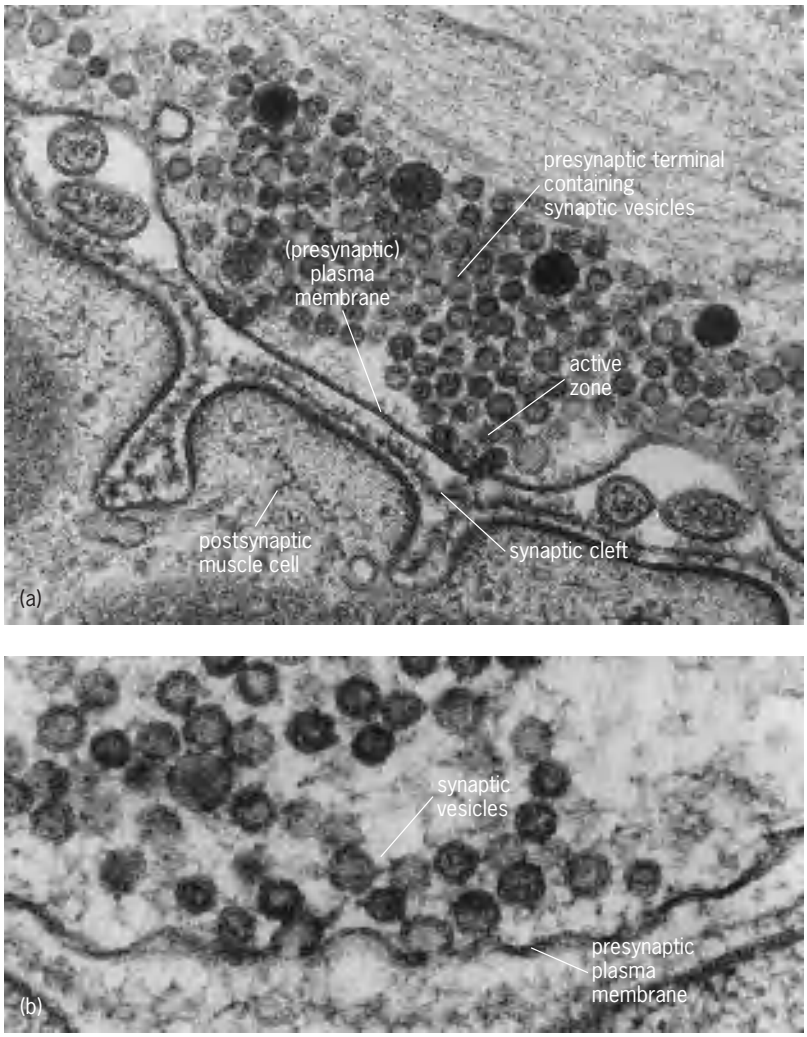

Fig. 2. Ultrastructure of exocytosis at the neuromuscular junction. (*a*) Thin-section electron micrograph of a neuromuscular junction from resting frog sartorius muscle. Vesicles are clustered in the presynaptic terminal at one of many active zones; some are docked at the plasma membrane (arrows). The postsynaptic muscle cell is seen across the synaptic cleft from the nerve terminal. (*b*) A similar presynaptic terminal frozen shortly (5 milliseconds) after electrical stimulation. Vesicles are seen in the process of fusing with the plasma membrane. (*Reproduced from J. Cell Biol., 88:564–580, 1981; copyright by The Rockefeller University Press*)

potassium (K^+) ions to pass through; under normal conditions this depolarizes the membrane potential toward zero.

Activation of the receptor occurs when two molecules of acetylcholine bind to each of the alpha subunits, causing an allosteric conformational change in the protein structure, which increases the ionic conductance of the channel. The acetylcholine receptor channel opens abruptly, has a relatively constant open conductance, and closes abruptly. These properties suggest that the protein jumps directly from being open to being closed in response to the presence or absence of acetylcholine.

Central synapses. The neuromuscular junction is an example of a large terminal with multiple synaptic sites. Most terminals, however, are small and contain essentially one active zone. In the central nervous system the presynaptic process containing synaptic vesicles is most often an axon terminal and the postsynaptic process a dendrite, making an axodendritic synapse, but other relationships are also seen. The effect of transmitter on a postsynaptic cell is either excitatory or inhibitory, meaning that it either depolarizes or hyperpolarizes the membrane. Whether a transmitter has an excitatory or inhibitory effect on a cell is determined by the type of ion able to pass through the cell's receptor channels. Most central synapses consist of two principal types.

Amino acid transmitters. Type 1 central synapses are characterized by small round presynaptic vesicles and a densification of the two apposed membranes that is asymmetrical, being much heavier postsynaptically. This type of synapse commonly releases an amino acid transmitter, such as glutamate, whose action produces an excitatory postsynaptic potential. At low levels of activity the glutamate binds to the glutamate receptor and activates a relatively small conductance increase for sodium and potassium ions. Glutamate is the transmitter at many excitatory synapses throughout the central nervous system.

GABA transmitters. The other type of central synapse, type 2, is characterized by presynaptic vesicles with varied shapes and a symmetrical densification of the apposed membranes. This type is usually associated with inhibitory synaptic actions. The most common inhibitory transmitter is gamma-amino butyric acid (GABA). Most inhibitory interneurons in different regions of the brain make these kinds of synapses on relay neurons in those regions. The GABA receptor is a complex channel-forming protein with several types of binding sites and several conductance states.

Other transmitters. Other types of central synapses utilize different transmitter substances, and there may be several types of receptor molecules for each of these substances. These can be distinguished from each other by the fact that they bind different agonistic substances, or are blocked by different antagonists. In the case of acetylcholine, at the neuromuscular junction the acetylcholine receptor binds nicotine, and is therefore called the nicotonic acetylcholine receptor, whereas at other synapses another type of acetylcholine receptor binds muscarine, and is known as the muscarinic acetylcholine receptor. At muscarinic synapses, second messengers phosphorylate a separate channel protein, bringing about a conformational change that activates a conductance that leads to an excitatory or inhibitory postsynaptic potential. This response has a slow time course, on the order of tenths of a second to several seconds. Muscarinic receptors are found especially at motoneuron synapses on gland cells and smooth muscle cells, and on specific populations of neurons in the forebrain. At these sites, the synapse may be much less clearly defined morphologically, and the transmitter may diffuse over long distances to act on receptor molecules.

Other classes of synaptic transmitter substances include the biogenic amines, such as the catecholamines, norepinephrine and dopamine, and

indoleamine 5-hydroxytryptamine (also known as serotonin). *See* DOPAMINE; NORADRENERGIC SYSTEM; SEROTONIN.

A final type of transmitter substance consists of a vast array of neuropeptides. These vary in size from simple dipeptides, such as carnosine, to complex molecules of 20 or more amino acids, which include some kinds of hormones. They are a diverse group and include substances that stimulate the release of hormones (such as thyrotropin-releasing hormone); those that act at synapses in pain pathways in the brain (the endogenous morphinelike substances, enkephalins and endorphins); and many of still undetermined functions (for example, cholecystokinin and vasoactive intestinal peptide). The release mechanisms and receptors for these peptides are still largely unknown, nor is it known whether they are confined to the sites of morphologically defined synapses. Peptides may act not only directly on receptors but also indirectly, modifying the state of a receptor in its response to other transmitter substances, and they may do this in an activity-dependent manner. Most of the known responses to peptides have long latencies of onset and long durations of action, lasting for several minutes or more. In view of the complexity and slow time course of many of their effects, these peptides are often referred to as neuromodulators. *See* ENDORPHINS; HORMONE.

G-proteins. At central synapses, rapid responses to transmitters are most commonly the result of direct synaptic transmission, in which the receptor itself is the ion channel. Such receptors (for example, the nicotinic acetylcholine receptor) are referred to as ionotropic, and typically have effects within milliseconds of being activated. There are, however, other receptor molecules for each of the neurotransmitters, and many of these are not themselves ion channels. These receptors (for example, the muscarinic acetylcholine receptor) are known as metabotropic receptors, and they affect neurotransmission indirectly via a set of intermediary proteins called G-proteins. G-proteins are so named because they bind to guanine nucleotides and change their molecular state depending on whether guanosine triphosphate (GTP) or guanosine diphosphate (GDP) is bound to the G-protein's alpha subunit. Activated G-proteins have a variety of effects on synaptic processes. Some are mediated by direct interactions with ion channels. However, many other effects are mediated by activation of cellular second messenger systems involving messengers such as calcium and cyclic adenosine monophosphate (cyclic AMP). The time courses for effects caused by activation of G protein-coupled receptors is much longer than that of ionotropic channels (milliseconds), reflecting the lifetime of the activated G-protein subunits (seconds) and second messengers (seconds to minutes). Such longer-lasting signals greatly increase the complexity of chemical neurotransmission and synaptic modulation.

Activity-dependent changes. The synapse is a dynamic structure whose function is very dependent on its activity state. At the neuromuscular junction, following vesicle release, the active zone undergoes periods of depression and potentiation that are increased in intensity by high-frequency stimulation. The number of receptors appears to be regulated by a mechanism in which a small percentage of receptors, when bound by acetylcholine, become desensitized; that is, they are unable to bind acetylcholine again. Increased synaptic activity causes an increase in the numbers of desensitized receptors, leading to decreases in the total number of receptors, a process known as down regulation. In this way the synapse is constantly adjusted for its information load.

At glutamatergic synapses, high levels of input activity bring about a different transmission state. The buildup of postsynaptic depolarization relieves the normal block of a specialized glutamate receptor channel by external magnesium ions, permitting influx of calcium ions into the postsynaptic process. Since the conductance of the channel is dependent on the depolarization state of the membrane, it is said to have a voltage-gated property, in addition to being ligand-gated by its transmitter. The calcium ion acts as a second messenger to bring about a long-lasting increase in synaptic efficacy, a process known as long-term potentiation. The conjunction of increased pre- and postsynaptic activity to give an increase in synaptic efficacy is called Hebbian, and is believed to be the type of plasticity mechanism involved in learning and memory. *See* LEARNING MECHANISMS; MEMORY.

Drug actions. The synapse is one of the primary targets of drug actions. The first example to be identified was the arrow poison curare, which blocks nicotinic neuromuscular transmission by binding to acetylcholine receptor sites. Many toxic agents have their actions on specific types of receptors; organic fluorophosphates, for example, are widely used pesticides that bind to and inactivate acetylcholinesterase. Most psychoactive drugs exert their effects at the synaptic level. Minor tranquilizers, such as the benzodiazapine antianxiety drugs, block GABA receptors among their actions. Antipsychotic drugs, on the other hand, appear to have in common the fact that they block dopamine transmission. This led to the hypothesis that schizophrenia is the result of increased activity at central dopamine synapses. Another theory has held that mental depression is due to a deficiency of catecholamine transmitters, especially norepinephrine; in this theory, mania would be due to increased activity at these synapses. However, all these mental syndromes are much more complex, and most of the transmitter systems are involved to varying extents. *See* NERVOUS SYSTEM (VERTEBRATE); NEUROBIOLOGY.

Gordon M. Shepherd; Phyllis I. Hanson

Bibliography. J. G. Nicholls et al., *From Neuron to Brain*, 4th ed., 2000; G. M. Shepherd, *Neurobiology*, 3d ed., 1994; G. Siegel et al., *Basic Neurochemistry: Molecular, Cellular and Medical Aspects*, 5th ed., 1994; M. J. Zigmond et al., *Fundamental Neuroscience*, 1999.

Synbranchiformes

An order of eel-like fishes in the class Actinopterygii, commonly called swamp eels. Synbranchiforms are among the most highly specialized fishes. They are characterized by a usually scaleless eel-like body lacking pelvic fins with pectoral fins present or absent; gill openings entirely ventral and often confluent across the breast; premaxillae lacking ascending processes and nonprotrusible; ectopterygoid bones enlarged; endopterygoids reduced or absent; swim bladder present or absent; and gills poorly developed, with respiration in some species accomplished in part by highly vascularized buccopharyngeal pouches. The order has no fossil record and comprises three families, 15 genera, and about 99 species, all but three of which are freshwater inhabitants. Synbrachiformes are thought to form a monophyletic group with the smegmamorphs. *See* ACTINOPTERYGII.

Family Synbranchidae (swamp eels). Swamp eels (**Fig. 1**) are characterized by a lack of pectoral fins, except in the early development of some species; vestigial dorsal and anal fins; a caudal fin, if present, that is small or vestigial; scales usually absent; small eyes with some species blind; widely separated nostrils; a gill opening that is a small slit or pore; absence of a swim bladder and ribs. They are the only teleosts to possess an amphistylic jaw suspension (the upper jaw is attached directly to the skull). They have a maximum total length of 150 cm (59 in.). Most species are protogynous hermaphrodites (first females, then males). Most species can breathe air, with some of them highly evolved in this capacity by having lunglike suprabranchial pouches. Swamp eels are primarily freshwater (all but three species), living in swamps, in caves (with blind species found in the Yucatan Peninsula of Mexico and Liberia), and in sluggish waters of tropical and subtropical western Africa, Asia, the Indo-Australian Archipelago, and Mexico to South America. The species *Monopterus albus* was introduced into inland waters of the United States. Swamp eels are capable of overland excursions, and some species can live out of water for extended periods by burrowing in the substrate at the onset of a dry season. The family comprises four genera and 17 species.

Families Chaudhuriidae and Mastacembelidae. These two families have a physoclistic swim bladder (that is, the air bladder is not connected with the digestive tract by an open duct); the pectoral girdle is attached to the vertebral column by ligaments; the posttemporal bone is absent; and the dorsal and anal fins are continuous with a small caudal fin. *See* SWIM BLADDER.

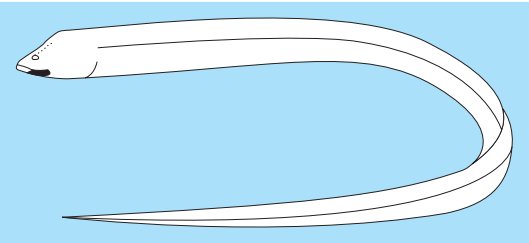

Fig. 1. Example from the Synbranchidae (swamp eels).

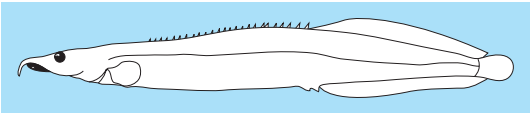

Fig. 2. Example from the Mastacembelidae (spiny eels).

Members of the Chaudhuriidae (earthworm eels) are distinguished by a lack of dorsal and anal fin spines; a usually naked body; absence of a lateral line, a rostral appendage, and endopterygoid and epineural bones; presence of a basisphenoid bone; and a maximum length of about 8 cm (3.2 in.). The six genera and nine species inhabit freshwaters of northeastern India and through Thailand to Korea, including parts of Malaysia and Borneo.

The Mastacembelidae (spiny eels) (**Fig. 2**) differ from other synbranchiforms by having from 9 to 42 individual (that is, isolated, not attached to each other by membranes) spines preceding the soft dorsal fin, which has from 52 to 131 rays and a fleshy rostral appendage. They are further distinguished by a body covered with small scales (naked body in perhaps three species); no basisphenoid bone; and maximum length of 90 cm (35 in.). In addition to occupying a wide variety of habitats, some species can remain buried in the substrate for long periods of time in order to escape dry periods. The five genera and 73 species occupy the freshwaters of tropical Africa and through Syria to the Malay Archipelago, China, and Korea. Herbert Boschung

Bibliography. R. Britz, S. Favorito, and G. D. Johnson, The hyopalatine arch of a 25 mm larva of *Synbranchus* and homology of the single pterygoid in the Synbranchidae (Teleostei: Synbranchiformes), *Proc. Biol. Soc. Wash.*, 116(2):337–340, 2003; P. T. Fuller, L. G. Nico, and J. D. Williams, Nonindigenous fishes introduced into inland waters of the United States, *Amer. Fish. Soc. Spec. Publ.*, no. 27, Bethesda, MD, 1999; G. D. Johnson and C. Patterson, Percomorph phylogeny: A survey of acanthomorphs and a new proposal, *Bull. Mar. Sci.*, 52(1):554–626, 1993; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Syncarida

A unique superorder of malacostracan Crustacea, noted particularly for the total lack of a carapace or carapace shield. Two orders are recognized, Anaspidacea and Bathynellacea. The Stygocarididae, at one time considered a distinct order, are now assigned familial rank within the Anaspidacea.

The syncarid body plan is simple, consisting of a cephalon (head and sometimes one thoracic somite), thorax of seven or eight somites, and an abdomen of six somites and telson or five somites and pleotelson (fused sixth somite-telson). Eyes, with apposition

optics, may be stalked or sessile, or sometimes absent. There are six to eight thoracic appendages (thoracopods), usually two to five abdominal appendages (pleopods), and uropods. Pleopods may be entirely lacking, or the first one or two pairs may be modified as copulatory appendages. Sexes are separate. Young hatch as immature adults and pass through several molts before reaching maturity. Studies have suggested that bathynellids, at least, arose by neoteny from a zoealike malacostracan ancestor similar to the Penaeidae.

Extant syncarids are specially adapted to freshwater, often interstitial environments. Whereas the Anaspidacea are restricted to the Southern Hemisphere (Australia, Tasmania, New Zealand, South America), bathynellids have been found worldwide, except in Antarctica. *See* BATHYNELLACEA; CRUSTACEA; MALACOSTRACA. Patsy A. McLaughlin

Bibliography. L. Botosaneanu (ed.), *Stygofauna munidi*, 1986; P. A. McLaughlin, *Comparative Morphology of Recent Crustacea*, 1980; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 1982; H. K. Schminke, Adaptation of Bathynellacea (Crustacea, Syncarida) to life in the interstitial ("Zoea Theory"), *Int. Rev. Ges. Hydrobiol.*, 66(4):575–637, 1981.

Synchronization

The process of maintaining one operation in step with another. The commonest example is the electric clock, whose motor rotates at some integral multiple or submultiple of the speed of the alternator in the power station. In television, synchronization is essential in order that the electron beams of receiver picture tubes will be at exactly the same spot on the screen at each instant as is the beam in the television camera tube at the transmitter. Synchronism in television is achieved by transmitting a synchronizing pulse at the end of each scanning line, to make all receivers move simultaneously to the start of the next line. A similar vertical synchronizing pulse is transmitted when the camera beam reaches the bottom of the picture, to make all beams go back to the top for the start of the next field. *See* OSCILLOSCOPE; TELEVISION. John Markus

Synchronous converter

A synchronous machine used to convert alternating current (ac) to direct current (dc), or vice versa. The ac-to-dc converter has been superseded by solid-state converters using silicon controlled rectifiers (SCRs) or power MOSFETs (for reasons of efficiency, lower maintenance costs, and less trouble) or by motor-generator sets. Converters are no longer manufactured, but there are converters still in use. *See* POWER INTEGRATED CIRCUITS; SEMICONDUCTOR RECTIFIER.

The synchronous converter has an armature similar to those used in dc machines, with connections to a commutator and brushes on the dc end and to slip rings on the ac end. The field is similar to that of a dc machine and may have shunt and series windings on the main poles, interpoles with series windings to assist commutation, and amortisseur windings to aid starting and to prevent hunting. The machine is thus a combination of a synchronous motor and a dc generator, having all the essential elements of each but with no duplication of parts.

As compared with a motor-generator set consisting of a synchronous motor driving a dc generator, the synchronous converter has the following advantages: (1) a single magnetic circuit comprising field poles, yoke, armature core, and teeth; (2) a single armature in whose conductors the ac and dc currents tend to cancel, thereby reducing greatly the copper losses; (3) a single pair of bearings and a smaller and shorter shaft; (4) higher efficiency, lower first cost, and lighter weight. There are, however, attendant disadvantages: (1) The ratio of dc to ac voltage is fixed and cannot be changed materially by adjusting the field excitation. (2) The machine is quite sensitive to sudden load changes and may hunt or flashover its commutator. (3) The heating in the armature is not the same for all conductors. (4) Power-factor correction by field adjustment is quite limited. *See* DIRECT-CURRENT GENERATOR; SYNCHRONOUS MOTOR. Loyal V. Bewley

Synchronous motor

An alternating-current (ac) motor which operates at a fixed synchronous speed proportional to the frequency of the applied ac power supply. A synchronous machine can be operated as either a generator or a motor (and sometimes as a source of reactive power, fulfilling the function of a capacitor), depending only on its applied shaft torque (whether positive, negative, or zero) and its excitation. There is no fundamental difference in the theory, design, or construction of a machine intended for any of these roles, although certain design features are stressed for each of them. In use, the machine may change its role from instant to instant. As a result, it is preferable to set up a common general theory for synchronous generators, motors, and capacitors. The distinction between generator and motor is merely a difference in the direction of the currents and the sign of the torque angles. *See* ALTERNATING-CURRENT GENERATOR; ALTERNATING-CURRENT MOTOR.

Basic theory. A single-phase, two-pole synchronous machine is shown in **Fig. 1**. The coil is on the pole axis at time $t = 0$, and the sinusoidally distributed flux Φ linked with the coil at any instant is given by Eq. (1), where ωt is the angular displacement of the coil and Φ_{max} is the maximum value of the flux. This flux will induce in a coil of N turns an instantaneous voltage e, given by Eq. (2).

$$\phi = \Phi_{max} \cos \omega t \tag{1}$$

$$e = -N\frac{d\phi}{dt} = \omega N \Phi_{max} \sin \omega t = E_{max} \sin \omega t \tag{2}$$

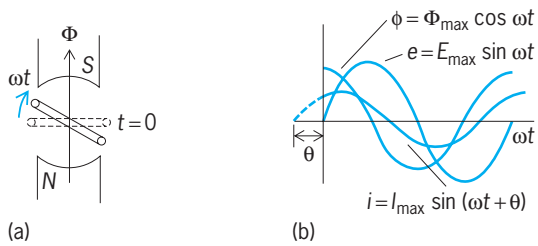

Fig. 1. Single-phase, two-pole synchronous machine. (*a*) Configuration of coil and poles. (*b*) Variation of flux Φ linking coil, voltage *e*, and current *i* with time.

The effective (rms) value E of this voltage is given by Eq. (3).

$$E = \frac{E_{\max}}{\sqrt{2}} = \sqrt{2}\,\pi f N \Phi_{\max} = 4.44 f N \Phi_{\max} \quad (3)$$

Armature reaction. Assume that the impedance of the coil and its external circuit of resistance R_t and reactance X is given by Eq. (4). There will flow a current,

$$Z = R_t \pm jX = Z\underline{/\pm\theta} \quad (4)$$

with a value given by Eq. (5), in which the phase

$$\mathbf{I} = \frac{\mathbf{E}}{\mathbf{Z}} = \frac{E}{Z}\underline{/\mp\theta} \quad (5)$$

angle ϕ is taken positive for a leading current. This current will develop a sinusoidal space distribution of armature reaction as in Eq. (6). If this single-phase

$$A = 0.8NI_{\max}\sin(\omega t + \theta) \quad (6)$$

mmf is expressed as a space vector and resolved into direct (in line with the pole axis) A_d and quadrature A_q components, it is given by Eq. (7).

$$\begin{aligned}\mathbf{A} &= A_d + jA_q \\ &= 0.4NI_{\max}\{[\sin\theta + \sin(2\omega t + \theta)] \\ &\quad + j[\cos\theta - \cos(2\omega t + \theta)]\}\end{aligned} \quad (7)$$

Polyphase armature reaction. In a three-phase machine, assuming balanced currents, the phase currents are given by Eqs. (8). Upon writing Eq. (7) for ωt,

$$\begin{aligned} i_a &= I_{\max}\sin(\omega t + \theta) \\ i_b &= I_{\max}\sin(\omega t + \theta - 120^\circ) \\ i_c &= I_{\max}\sin(\omega t + \theta - 240^\circ)\end{aligned} \quad (8)$$

$\omega t - 120^\circ$, and $\omega t - 240^\circ$, respectively, and adding, Eq. (9) results for the polyphase armature reaction.

$$\mathbf{A} = A_d + jA_q = 1.2NI_{\max}(\sin\theta + j\cos\theta) \quad (9)$$

Equation (10) gives the three-phase power of the machine, and Eq. (11) gives the developed torque.

$$P = 3EI\cos\theta \quad (10)$$

$$T = \frac{P}{\omega} = \frac{3}{\omega}EI\cos\theta \quad (11)$$

The above equations constitute the essential description of the operation of a synchronous generator. The same equations apply for a motor if the currents are reversed, that is, by changing the sign of the current I. They may also be interpreted in the form of phasor diagrams, and show the two cases of a smooth-rotor and a salient-pole machine.

Smooth-rotor synchronous machine. In the smooth-rotor machine, the reluctance of the magnetic path is essentially the same in either the direct or quadrature axes. In **Fig. 2***a* let the flux Φ be selected as reference phasor and drawn vertically. Then comparing Eqs. (1) and (2) it is seen that the induced voltage E_f lags the flux by 90°. By Eq. (5) the current I lags the voltage by an angle θ for an inductive circuit, and by Eq. (9) causes a constant mmf of armature reaction A in phase with the current. This armature reaction causes a flux ϕ_a, stationary in space with respect to the field poles, which in turn induces a voltage E_a lagging it by 90°. The two induced voltages E_f (due to the field flux Φ) and E_a (due to the armature reaction flux ϕ_a) combine vectorially to give the resultant voltage E'. But the terminal voltage V is less than E' by the resistance and reactance drops, $R\mathbf{I}$ and $jx_l\mathbf{I}$ in the winding, and Eq. (12) applies.

$$\mathbf{V} = \mathbf{E}' - (R + jx_l)\mathbf{I} \quad (12)$$

The leakage reactance drop $jx_l\mathbf{I}$ lags the current by 90° as does the armature reaction voltage E_a. If a fictitious reactance of armature reaction x_a is introduced to account for E_a, it is obvious that Eq. (12) may be rewritten to give Eqs. (13), in which

$$\begin{aligned}\mathbf{V} &= E_f - jx_a\mathbf{I} - (R + jx_l)\mathbf{I} \\ &= E_f - R\mathbf{I} - j(x_a + x_l)\mathbf{I} \\ &= E_f - (R + jX_s)\mathbf{I}\end{aligned} \quad (13)$$

$X_s = x_a + x_l$ is called the synchronous reactance of the machine.

Salient-pole synchronous machine. In a similar fashion the phasor diagram for a salient-pole machine, Fig. 2*b*, may be set up. Here the effects of saliency result in proportionately different armature reaction fluxes in the direct and quadrature axes, thereby necessitating corresponding direct, X_d, and quadrature, X_q, components of the synchronous reactance. The angle δ in Fig. 2 is called the torque angle. It is the angle between the field-induced voltage E_f and the terminal voltage V and is positive when E_f is ahead of V.

The foregoing equations and phasor diagrams were established for a generator. A motor may be regarded as a generator in which the power component of the current is reversed 180°, that is, becomes an input instead of an output current. The motor vector diagram is shown in **Fig. 3**. Here the torque angle δ is reversed, since V is ahead of E_f in a motor (it was behind in the generator). Therefore a motor differs from a generator in two essential respects: (1) The currents are reversed, and (2) the torque angle has changed sign. As a result the power input, Eq. (10), for a motor is negative, or has become a power output, and the torque is reversed in sign.

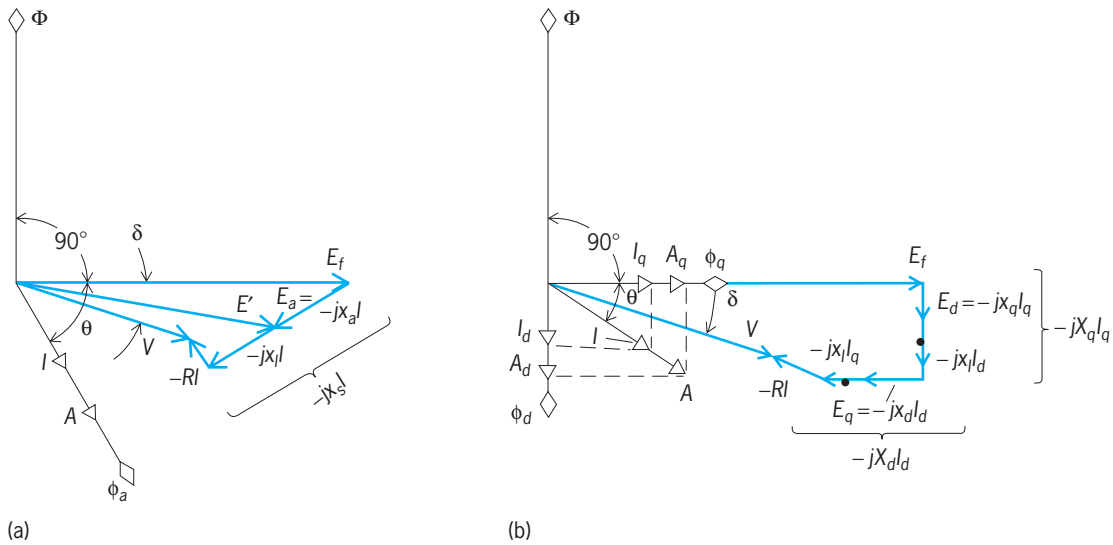

Fig. 2. Vector diagrams of synchronous generators. (*a*) Smooth-rotor machine. (*b*) Salient-pole machine.

When the current **I** is 90° out of phase with the terminal voltage V, the torque angle δ is nearly zero, being just sufficient to account for the power lost in the resistance.

Therefore, a synchronous machine is a generator, motor, or capacitor, depending on whether its torque angle δ is positive, negative, or zero. For these conditions the output current is, respectively, at an angle in the first or fourth quadrant, second or third quadrant, or essentially ±90° with respect to the terminal voltage at 0°. For any given power input or output, the machine can be made to operate at either leading or lagging power factor by changing the magnitude of the field current producing the flux Φ.

Synchronous capacitor. A synchronous capacitor can be made to draw a leading current and to behave like a capacitance by overexciting its field. Or, it will draw a lagging current on underexcitation. This characteristic thus presents the possibility of power-factor correction of a power system by adjusting the field excitation. A machine so employed at the end of a transmission line permits a wide range of voltage regulation for the line. One used in a factory permits the power factor of the load to be corrected. Of course, a synchronous motor can also be used for power-factor correction, but since it must also carry the load current, its power-factor correction capabilities are more limited than for the synchronous condenser. *See* STATIC VAR COMPENSATOR.

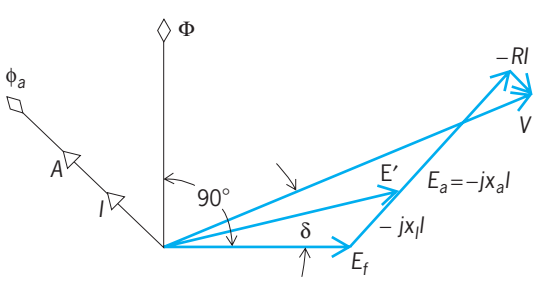

Fig. 3. Phasor diagram of synchronous motor.

Power equations. The power output P_o and reactance power output Q_o of a round-rotor synchronous machine are given by Eqs. (14), in which $\tan\alpha = R/X_s$ and $\mathbf{Z}_s = R + jX_s = Z_s < 90° - \alpha$.

$$P_o = \frac{VE_f}{Z_s}\sin(\delta+\alpha) - \frac{RV^2}{Z_s^2}$$
$$Q_o = \frac{VE_f}{Z_s}\cos(\delta+\alpha) + \frac{X_s V^2}{Z_s^2} \tag{14}$$

For a round-rotor motor, both the torque angle δ and the electrical power output are negative.

For a salient-pole machine, neglecting resistance, Z_s is equal to the direct-axis reactance X_d, α is zero, and P_o is given by Eq. (15). Thus the power or torque

$$P_o = \frac{VE_f}{X_d}\sin\delta + V^2\frac{X_d - X_q}{2X_dX_q}\sin 2\delta \tag{15}$$

depends essentially on the product of the terminal and induced voltages and sine of the torque angle δ; but in the case of the salient-pole machine there is also a second harmonic term which is independent of the excitation voltage E_f. This term, the so-called reluctance power, vanishes for nonsaliency when $X_d = X_q$. The small synchronous motors used in some electric clocks and other low-torque applications depend solely on this reluctance torque.

Excitation characteristics. The so-called V curves of a synchronous motor are curves of armature current plotted against field current with power output as parameter. Usually a second set of curves with input power factor (pf) as parameter is superimposed on the same plot. Such curves (**Fig. 4**), where armature current is plotted against generated voltage, can be determined from design calculations or from test; they yield a considerable amount of data on the performance of the motor. Thus, given any two of the four variables E_f, I, pf, mechanical power P_m, the remaining two may be easily determined, as well as the conditions of maximum power,

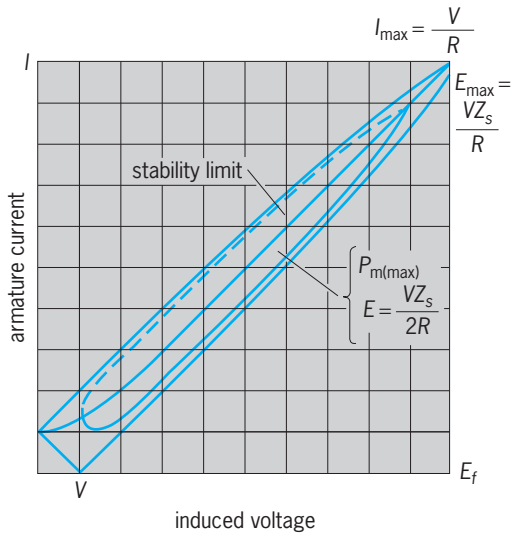

Fig. 4. V curves (armature current versus induced voltage) of synchronous motor.

constant pf, minimum excitation, stability limit, and so forth.

Losses and efficiency. The losses in a synchronous motor comprise the copper losses in the field, armature, and amortisseur windings; the exciter and rheostat losses of the excitation system; the core loss due to hysteresis and eddy currents in the armature core and teeth and in the pole face; the stray loss due to skin effect in conductors; and the mechanical losses due to windage and friction. The efficiency of the motor is then given by Eq. (16).

$$\text{Eff} = \frac{\text{output}}{\text{input}} = \frac{\text{output}}{\text{output} + \text{losses}} \qquad (16)$$

Mechanical oscillations. A synchronous motor subjected to sudden changes of load, or when driving a load having a variable torque (for example, a reciprocating compressor), may oscillate about its mean synchronous speed. Under these conditions the torque angle δ does not remain fixed, but varies. As a result the four separate torques expressed in Eq. (17)

$$\begin{pmatrix}\text{Synchronous}\\ \text{motor torque}\\ \text{Eq. (15)}\end{pmatrix} + \begin{pmatrix}\text{induction motor}\\ \text{torque of}\\ \text{amortisseur}\end{pmatrix} = \begin{pmatrix}\text{torque to}\\ \text{overcome}\\ \text{inertia}\end{pmatrix} + \begin{pmatrix}\text{torque}\\ \text{required}\\ \text{by the load}\end{pmatrix} \qquad (17)$$

act on the machine rotor. The possibility exists that cumulative oscillations will build up and cause the motor to fall out of step.

Starting of synchronous motors. Synchronous motors are provided with an amortisseur (squirrel-cage) winding embedded in the face of the field poles. This winding serves the double purpose of starting the motor and limiting the oscillations or hunting. During starting, the field winding is either closed through a resistance, short-circuited, or opened at several points to avoid dangerous induced voltages. The amortisseur winding acts exactly as the squirrel-cage winding in an induction motor and accelerates the motor to nearly synchronous speed. When near synchronous speed, the field is excited, and the synchronous torque pulls the motor into synchronism. During starting, Eq. (17) applies, since all four types of torque may be present. Of course, up to the instant when the field is excited the portion of the synchronous motor torque depending on E_f does not exist, although the reluctance torque will be active.

Other methods of starting have been used. If the exciter is direct-connected and a dc source of power is available, it may be used to start the synchronous motor. In the so-called supersynchronous motor the stator is able to rotate in bearings of its own, and is provided with a brake band. For stator allowed to come up to nearly synchronous speed by virtue of the amortisseur windings; the field is then excited and the stator brought to synchronous speed, the rotor remaining stationary. Then as the brake band is tightened, the torque on the rotor causes it to accelerate while the speed of the stator correspondingly slackens; finally the stator comes to rest and is locked by the brake band. In this way maximum synchronous motor torque is made available for acceleration of the load.

For other types of synchronous motors *see* HYSTERESIS MOTOR; RELUCTANCE MOTOR.

Richard T. Smith

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; A. E. Fitzgerald, C. Kingsley, Jr., and S. D. Umans, *Electric Machinery*, 6th ed., 2003; L. W. Matsch and J. D. Morgan, *Electromagnetic and Electromechanical Machines*, 3d ed., 1986.

Synchrotron radiation

Electromagnetic radiation emitted by relativistic charged particles following a curved path in magnetic or electric fields. Synchrotron radiation is generally produced in laboratories by a beam of relativistic (moving near the speed of light) electrons or positrons in a generally circular path in a specialized synchrotron called a storage ring. A storage ring produces electromagnetic radiation with high flux, brightness, and coherent power levels. Synchrotron radiation is used for a wide variety of basic and applied research in biology, chemistry, and physics, as well as for applications in medicine and technology. *See* ELECTROMAGNETIC RADIATION; PARTICLE ACCELERATOR; RELATIVISTIC ELECTRODYNAMICS.

Electron storage rings provide radiation over a very broad range of photon energies or wavelengths. Radiation is generally produced in the infrared, visible, near-ultraviolet, vacuum-ultraviolet, soft x-ray, and hard x-ray parts of the electromagnetic spectrum, depending on the energy of the electrons in the storage ring and the strength of the magnetic field used to produce the radiation. Lasers are a much

Storage ring synchrotron radiation facilities*

Country	City	Facility	Energy, GeV
Armenia	Yerevan	CANDLE (Center for the Advancement of Natural Discoveries using Light Emission)	3.2‡
Australia	Melbourne	Australian Synchrotron	3†
Brazil	Campinas	LNLS (Laboratorio Nacional de Luz Sincrotron)	1.35
Canada	Saskatoon	CLS (Canadian Light Source)	2.9†
China	Beijing	BSRF (Beijing Synchrotron Radiation Facility)	1.5–2.8
China	Hefei	NSRL (National Synchrotron Radiation Laboratory)	0.8
China	Shanghai	SSRF (Shanghai Synchrotron Radiation Facility)	3.5‡
China, Republic of (Taiwan)	Hsinchu	NSRRC (National Synchrotron Radiation Research Center)	1.3-1.5†
Denmark	Aarhus	ASTRID (Aarhus Storage Ring in Denmark)	0.6
France	Grenoble	ESRF (European Synchrotron Radiation Facility)	6†
France	Gif-sur-Yvette	SOLEIL	2.75†
Germany	Berlin	BESSY II (Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung)	1.7–1.9†
Germany	Bonn	ELSA (Electron Stretcher Accelerator)	1.5–3.5
Germany	Dortmund	DELTA (Dortmund Electron Test Accelerator)	1.5
Germany	Hamburg	DORIS (originally Double Ring Store) [at HASYLAB (Hamburger Synchrotronstrahlungslabor)]	4.5
Germany	Karlsruhe	PETRA (originally Positron-Electron Tandem Ring Accelerator Facility) [at HASYLAB]	7–14
Germany	Karlsruhe	ANKA (Angstromquelle Karlsruhe)	2.5†
India	Indore	INDUS I	0.45
India	Indore	INDUS II	2.5‡
Italy	Frascati	DAFNE	0.5
Italy	Trieste	ELETTRA	2–2.4†
Japan	Nishi Harima	SPRing-8	8†
Japan	Nishi Harima	New SUBARU	1.5†
Japan	Okasaki	UVSOR (Ultraviolet Synchrotron Orbital Radiation Facility)	0.75
Japan	Sendai	TSRF (Tohoku Synchrotron Radiation Facility)	1.5
Japan	Tsukuba	Photon Factory	2.5
Japan	Tsukuba	PF-AR (Photon Factory Accelerator Ring)	6.5
Jordan	Allaan	SESAME (Synchrotron-light for Experimental Science and Applications in the Middle East)	2.5‡
Korea	Pohang	PLS (Pohang Light Source)	2†
Russia	Dubna	DELSY (Dubna Electron Synchrotron)	0.95
Russia	Novosibirsk	VEPP-2	0.7
Russia	Novosibirsk	VEPP-3	2.2
Russia	Novosibirsk	VEPP-4	5–7
Singapore	Singapore	SSLS (Singapore Synchrotron Light Source): HELIOS 2	0.7
Spain	Barcelona	ALBA	3‡
Sweden	Lund	MAX I	0.55
Sweden	Lund	MAX II	1.5†
Switzerland	Villigen	SLS (Swiss Light Source)	2.4†
Thailand	Nakhon Ratchasima	NSRC (National Synchrotron Research Center)	1–1.3
Ukraine	Kiev	ISI-800	1‡
United Kingdom	South Oxfordshire	Diamond	3†
United States	Argonne, IL	APS (Advanced Photon Source)	7†
United States	Baton Rouge, LA	CAMD (Center for Advanced Microstructures & Devices)	1.3–1.5†
United States	Berkeley, CA	ALS (Advanced Light Source)	1.9†
United States	Gaithersburg, MD	SURF III (Synchrotron Ultraviolet Radiation Facility)	0.386
United States	Ithaca, NY	CHESS (Cornell High-Energy Synchrotron Source)	5.5
United States	Stanford, CA	SPEAR 3 (Stanford Synchrotron Radiation Center)	3†
United States	Stoughton, WI	SRC-Aladdin (Synchrotron Radiation Center)	0.8–1
United States	Upton, NY	NSLS (National Synchrotron Light Source)	0.8
United States	Upton, NY	NSLS II (National Synchrotron Light Source)	2.5–2.8‡

*Partial list of facilities as of June 2006, compiled from www.lightsources.org and other references.
†More recently constructed source, usually with performance of a "third-generation" facility.
‡Planned or under construction as of June 2006.

brighter source of radiation in the visible than are storage rings, so synchrotron radiation is only rarely used in the visible part of the spectrum. However, synchrotron radiation is particularly useful for producing x-ray beams.

Numerous storage ring sources of synchrotron radiation are in operation, construction, or design in many countries (see **table**). The flux [photons/(second, unit bandwidth)], brightness (or brilliance) [flux/(unit source size, unit solid angle)], and coherent power (important for imaging applications and proportional to brightness) available for experiments—particularly in the vacuum-ultraviolet, soft x-ray, and hard x-ray parts of the spectrum—are steadily increasing as higher-performance storage rings are constructed. A beam from a modern storage ring has flux, brightness, and coherent power many orders of magnitude higher than is available from other sources (**Fig. 1**).

Experimental Facilities

Synchrotron radiation has many features (natural collimation, high intensity and brightness, broad spectral bandwidth, high polarization, pulsed time

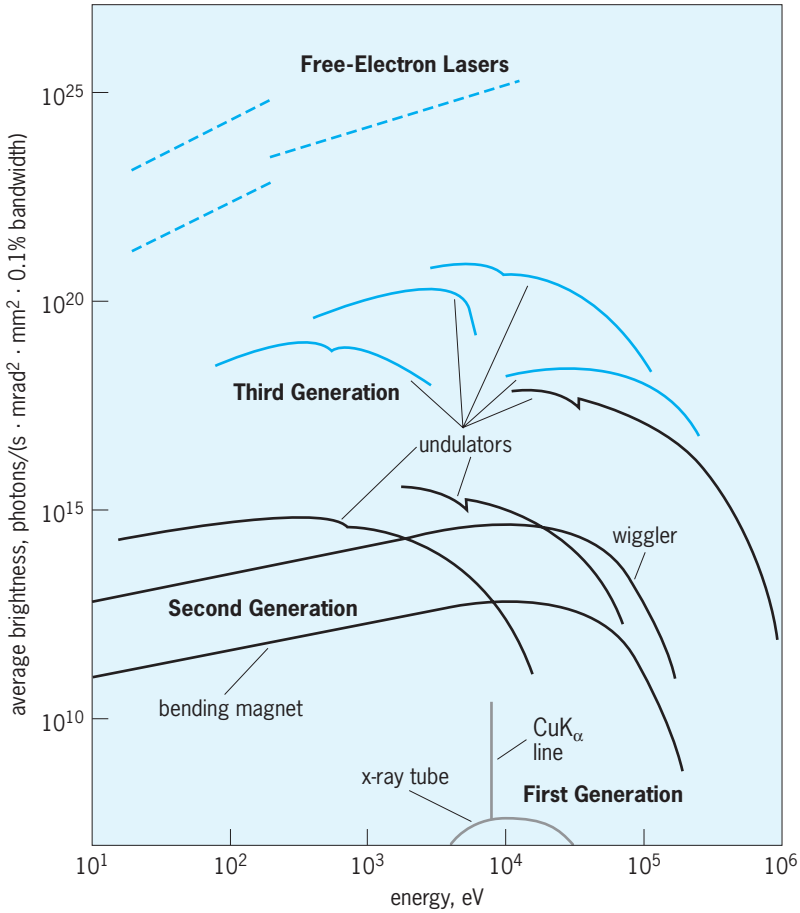

Fig. 1. Representative average brightness of synchrotron radiation sources and free-electron lasers as a function of photon energy. Present and extrapolated performances are shown.

structure, small source size, and high-vacuum environment) that make it ideal for a wide variety of applications in experimental science and technology. Synchrotron radiation was predicted theoretically early in the twentieth century, and first observed in 1947 at General Electric Laboratories. Very powerful sources of synchrotron radiation in the ultraviolet and x-ray parts of the spectrum became available when high-energy physicists began operating electron synchrotrons in the 1950s. Although synchrotrons produce large amounts of radiation, their cyclic nature results in pulse-to-pulse variations, rendering difficult their use in many types of research. By contrast, the electron-positron storage rings developed for colliding-beam experiments starting in the 1960s offered a constant electron-beam intensity and much better stability. Beam lines to bring synchrotron radiation from a storage ring to users were constructed on both synchrotrons and storage rings to allow the radiation produced in the bending magnets of these machines to leave the ring vacuum system and reach experimental stations. In most cases the research programs were pursued on a parasitic basis, secondary to the high-energy physics programs. These parasitic light sources are now referred to as first-generation light sources.

The brightness of x-ray sources has continued to increase since synchrotron radiation surpassed conventional x-ray sources (**Fig. 2**). The proportional rate of increase of the brightness of synchrotron-radiation sources is greater than that of computer processing speed.

Second-generation light sources. Storage rings dedicated as synchrotron light sources have been completed in many countries around the world (see table). They are called second-generation light sources to distinguish them from the first-generation synchrotrons that were built for research in high-energy physics.

Special arrays of magnets may be inserted into the straight sections between storage-ring bending magnets to produce beams with extended spectral range or with higher flux and brightness than is possible with the ring bending magnets. These devices, called wiggler and undulator magnets, utilize periodic transverse magnetic fields to produce transverse oscillations of the electron beam with no net deflection or displacement. They provide orders-of-magnitude increase in flux and brightness over storage-ring bending magnets (Fig. 1), again opening up new research opportunities.

Third-generation light sources. Third-generation light sources are storage rings with many straight sections for wiggler and undulator insertion device sources and with a smaller transverse size and angular divergence of the circulating electron beam. The product of the transverse size and divergence is called the emittance. The lower the electron-beam emittance, the higher the photon-beam brightness and coherent power level. With smaller horizontal emittances and with straight sections that can accommodate longer undulators, third-generation

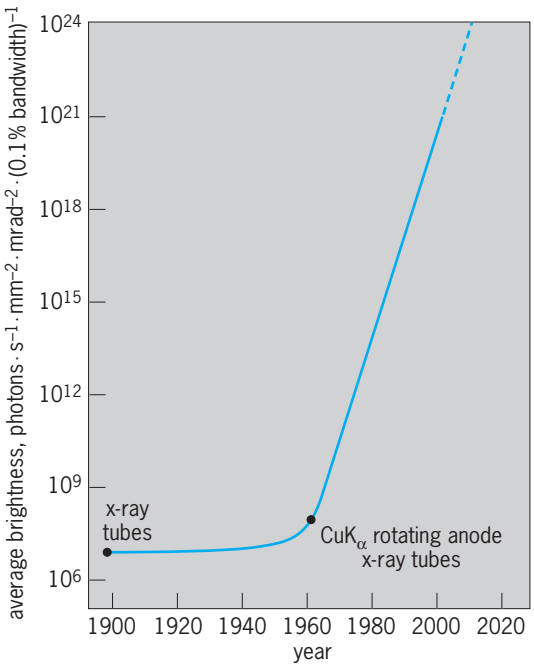

Fig. 2. Average brightness of x-ray sources from discovery of x-rays to the present (extrapolated to future).

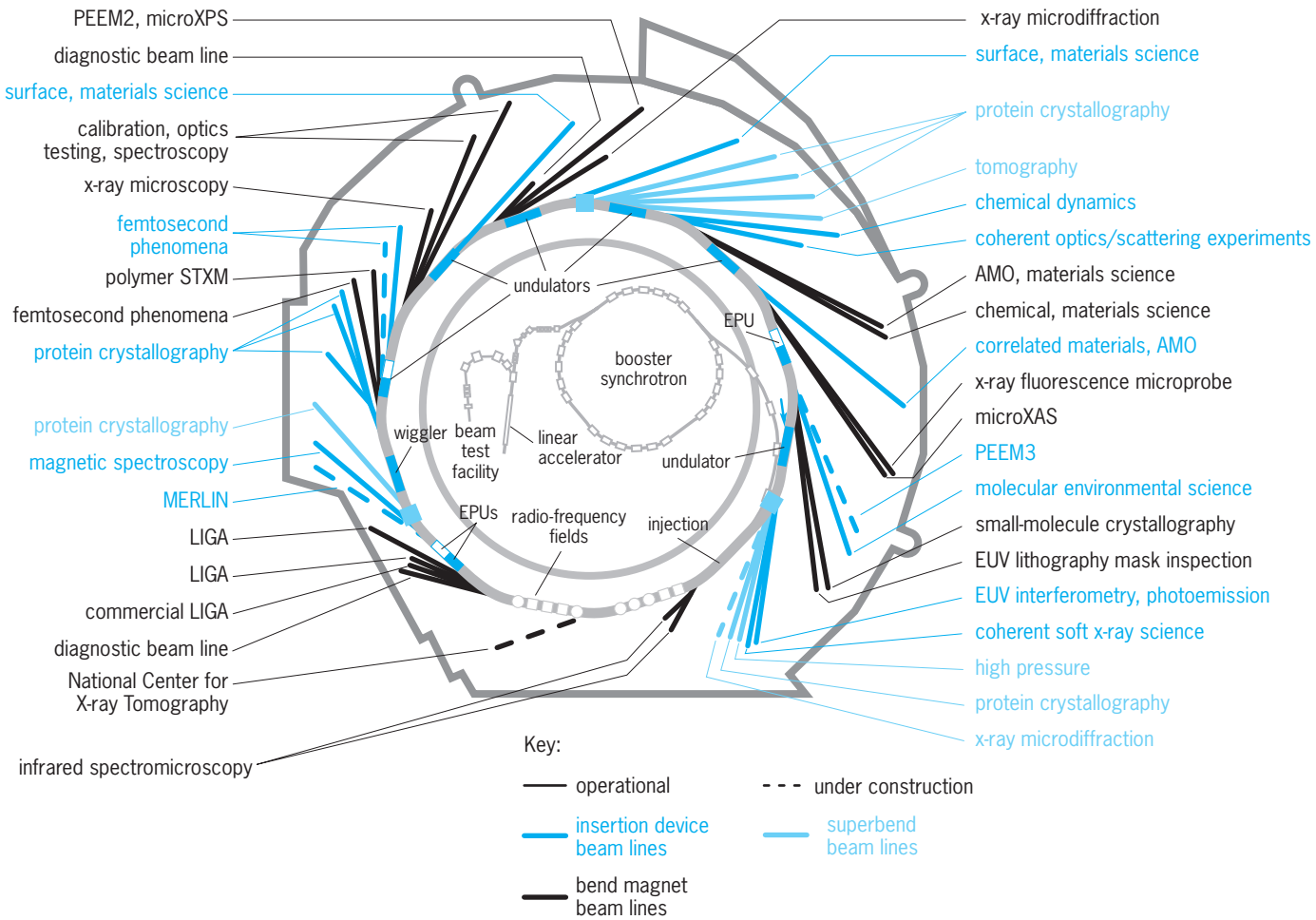

Fig. 3. Layout of the 1.9-GeV Advanced Light Source at Lawrence Berkeley Laboratory, a low-energy, third-generation synchrotron radiation source, as of 2006. Applications of experimental stations on beam lines are indicated. EPU = elliptically polarizing undulator; PEEM = photoemission electron microscope; XPS = x-ray photoelectron spectroscopy; STXM = scanning transmission x-ray microscopy; LIGA = Lithographie, Galvanoformung, und Abformung (lithography, electroforming, and molding); AMO = atomic, molecular, and optical physics; XAS = x-ray absorption spectroscopy; EUV = extreme ultraviolet.

rings provide two or more orders of magnitude higher brightness and coherent power level than earlier sources. One consequence of the extraordinary brightness (brilliance) of these light sources is that the x-ray beam is partially coherent, allowing research in interferometry and speckle spectroscopy, among other uses. *See* COHERENCE; INTERFEROMETRY; SPECKLE.

A considerable number of third-generation rings are in operation (see table). Third-generation light sources with a low-energy electron beam (typically 1–2 GeV) [**Fig. 3**] are generally optimized to produce high-brightness radiation in the vacuum ultraviolet (VUV) and soft (low-energy) x-ray spectral range, up to photon energies of about 2–3 keV. (Lower-brightness light extends to above 10 keV, and the use of superconducting bending magnets can extend the photon energy range to considerably higher values.) Third-generation light sources with a high-energy electron beam (6–8 GeV) can produce harder (higher-energy) x-rays, with energies of tens of kiloelectronvolts and above. Some of the newest storage rings that are intermediate in size, with beam energies up to 3 GeV, have even lower emittances than the first third-generation machines, and are equipped with advanced insertion-device technology with which they can also reach high brightness in the hard x-ray region.

Free-electron lasers. The next logical step in the evolution of synchrotron radiation sources will be a fully coherent source, namely an x-ray laser. A fully coherent source will have a peak brightness 10 orders of magnitude greater than existing third-generation light sources. To produce a source this brilliant requires a very small and well-collimated electron beam; the best way to achieve a particle beam with the required characteristics is not with a storage ring but with a linear accelerator (linac). The beam accelerated by the linac will pass through a very long undulator (approximately 100 m or 300 ft in length) and generate x-ray beams of very high brightness through a process called self-amplified stimulated emission (SASE). Along with a 10,000-fold increase in average x-ray brightness over third-generation sources, fourth-generation free-electron laser sources will have subpicosecond pulse

durations, resulting in x-ray beams with unparalleled instantaneous power levels. Several free-electron lasers are being built to produce vaccum ultraviolet and x-ray beams, such as the LCLS (Linac Coherent Light Source) at Stanford University. The unique properties of x-ray free-electron lasers, such as the instantaneous or peak brightness of a single ultrashort pulse, will surely open new fields of scientific research. *See* LASER.

Properties of Synchrotron Radiation

The radiation produced by an electron in circular motion at low energy (speed much less than the speed of light) is weak and rather nondirectional. At relativistic energies (speed close to the speed of light), the radiated power increases markedly, and the emission pattern is folded forward into a cone with a half-opening angle in radians given approximately by mc^2/E, where mc^2 is the rest-mass energy of the electron (0.51 MeV) and E is the total energy of the electron. Thus, at electron energies of the order of 1 GeV, much of the very strong radiation produced is confined to a forward cone with an instantaneous opening angle of about 1 mrad (0.06°). At higher electron energies this cone is even smaller. The large amount of radiation produced combined with the natural collimation gives synchrotron radiation its intrinsic high brightness. Brightness is further enhanced by the small cross-sectional area of the electron beam, which is as low as 0.01 mm^2 in the third-generation storage rings.

Bending magnet sources. As an electron moves in an arc through a bending magnet (**Fig. 4***a*), the instantaneous forward cone sweeps out a horizontal fan of continuum, polarized radiation. The small vertical opening angle, $2/\gamma$, is preserved, where γ is the ratio of the moving mass of the electron to its rest mass, but the horizontal angle is determined by the acceptance of the beam pipe or optical elements (for example, mirrors, gratings, or crystals). By accepting horizontal angles larger than $2/\gamma$, more flux can be utilized. The brightness, however, is not increased.

Wiggler magnet sources. The electron beam in a wiggler magnet executes transverse oscillations with an angular excursion much larger than $1/\gamma$, meaning excursions through angles larger than the instantaneous radiation. The result is an incoherent sum of radiation from each bend of the electrons. A fan of radiation is produced (Fig. 4*b*), which is similar to that produced by a bending magnet but with an enhancement of intensity due to the accumulation of radiation from each pole. The magnetic field can be higher than the storage-ring bending magnet field, resulting in a spectrum that extends to higher photon energy. Superconducting wigglers with fields up to 7.5 teslas have been used to extend the spectrum deeper into the hard x-ray region. Wigglers (and undulators) can be electromagnets or permanent magnets (**Fig. 5**). Permanent magnets are commonly used, particularly for devices with short periods and many poles, but superconducting electromagnets are useful when the highest magnetic fields are desired. *See* MAGNET.

Undulator magnet sources. An undulator magnet is designed to cause the electron beam to execute transverse oscillations with an angular excursion of the order of $1/\gamma$. The electron-beam angular deflection of the same order as the natural emission angle of synchrotron radiation results in a coherent superposition of radiation from each deflection of the electron beam. The radiation emerging from an undulator is concentrated into a narrow, foward-directed cone with the smallest possible opening angle (Fig. 4*c*). Brightness is increased over that of bend-magnet radiation because of the collimation of the radiation in both horizontal and vertical planes (Fig. 1). Undulators thus produce the brightest radiation possible from an electron beam. Furthermore, the spectral distribution of radiation from an undulator is qualitatively different from that of a bending magnet or wiggler. Nearly all undulators used in third-generation synchrotron light sources are fabricated from permanent magnets; however, superconducting devices are becoming popular when higher photon energies are desired.

Developments in undulator technology have considerably improved their performance. An example is an undulator in which the magnetic material is inside (rather than outside) the vacuum chamber, allowing considerably higher magnetic fields at the electron beam than were previously obtainable. Other new types of undulators can produce variable polarization. For example, devices with four arrays of magnets (two above and two below the beam) can produce circular, helical, or linear oscillations of the

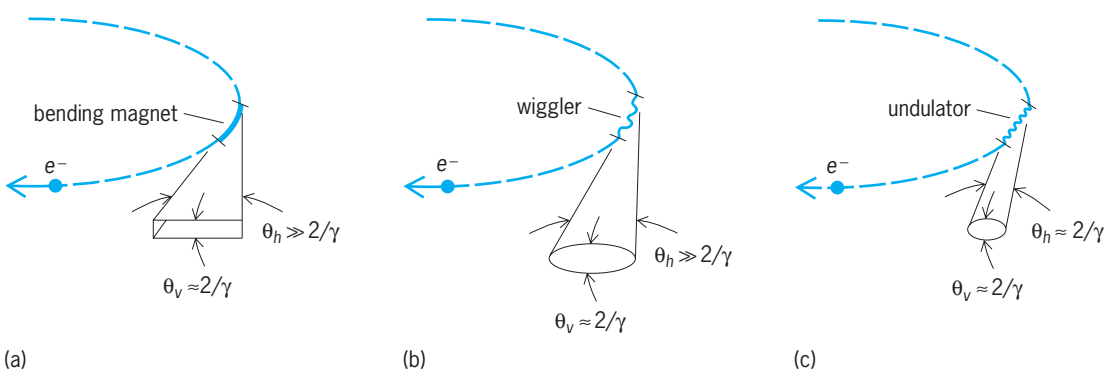

Fig. 4. Spatial characteristics of photon beams from (*a*) bending magnets, (*b*) wigglers, and (*c*) undulators.

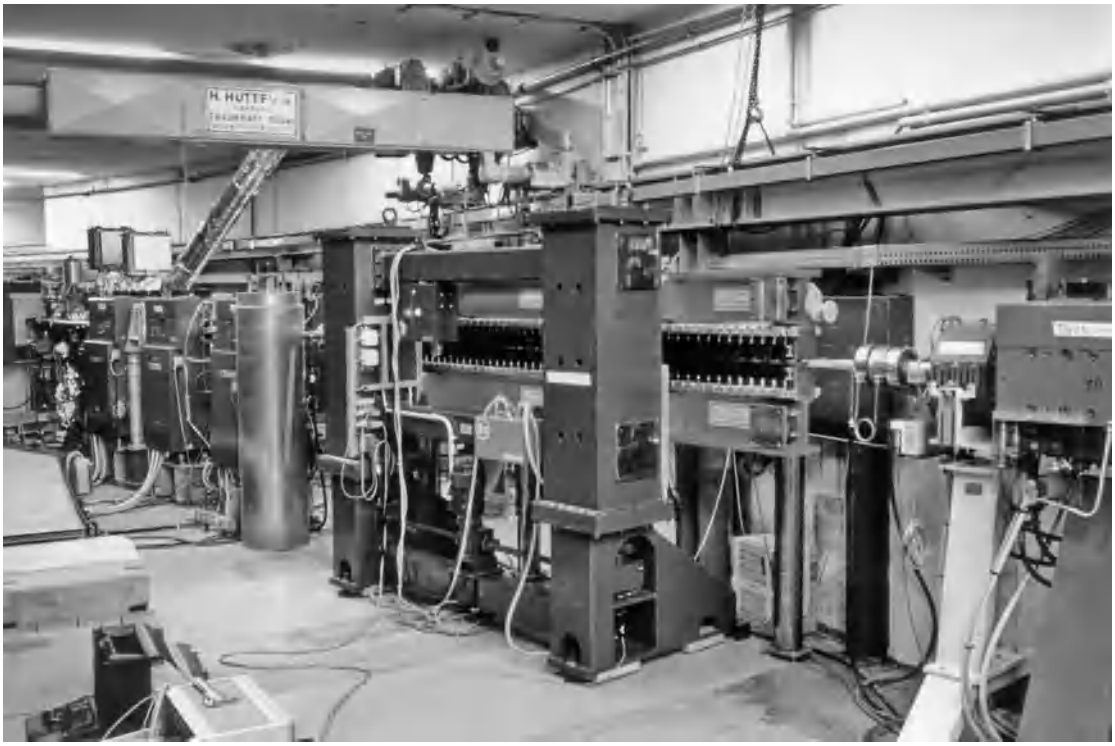

Fig. 5. Open wiggler based on arrays of permanent magnets in the storage ring DORIS at the Hamburger Synchrotronstrahlungslabor (HASYLAB) at DESY (Deutches Elektronen-Synchroton). The spectrum of emitted synchrotron radiation can be varied by varying the distance between the two magnet arrays. The massive structure is required to maintain precision alignment between the jaws. (*DESY*)

electrons, depending on the relative phases of the fields; the radiation produced by this type of undulator has controllable polarization, including linear polarization in any plane and elliptical or circular polarization.

Spectrum. The main spectral features of the radiation produced by bending magnets, wigglers, and undulators are compared in **Figs. 6**, **7**, and **8**. Figures 6 and 7 show the spectrum (flux and brightness) of bending magnets of the Advanced Light Source at the Lawrence Berkeley Laboratory and of a wiggler operating in that ring. Notably higher flux is provided by the wiggler. The flux spectra are characterized by a single parameter—the critical energy ε_c, given in kiloelectronvolts by $2.2E^3/R$ or $0.665BE^2$, where E is the electron energy in gigaelectronvolts, R is the radius of curvature in meters, and B is the magnetic flux density in teslas. Half the power is radiated above the critical energy and half below. Typically, useful flux is available at energies up to four or five times the critical energy. The smooth, continuous nature of the spectra of bending magnets and wigglers over a broad range of photon energy makes it possible to select the particular energy of interest (for example, the absorption edge energy of a particular element) or to study a process such as x-ray absorption or photoemission as a function of photon energy. In use, the broadband radiation is monochromatized by a grating or crystal so that the bandwidth ($\Delta\varepsilon/\varepsilon$, where ε is the energy of the radiation and $\Delta\varepsilon$ is the energy width of the radiation) is 10^{-3} or smaller. Thus, highly monochromatic radiation at any energy within the range of the synchrotron radiation source is readily available.

The spectrum of radiation from an undulator consists of one or more peaks in energy. Interference effects from the coherent superposition of radiation from each deflection of the electron beam produce a spectrum with peak frequencies that are integral multiples of a fundamental frequency. The spectrum is characterized by the undulator strength parameter, K.

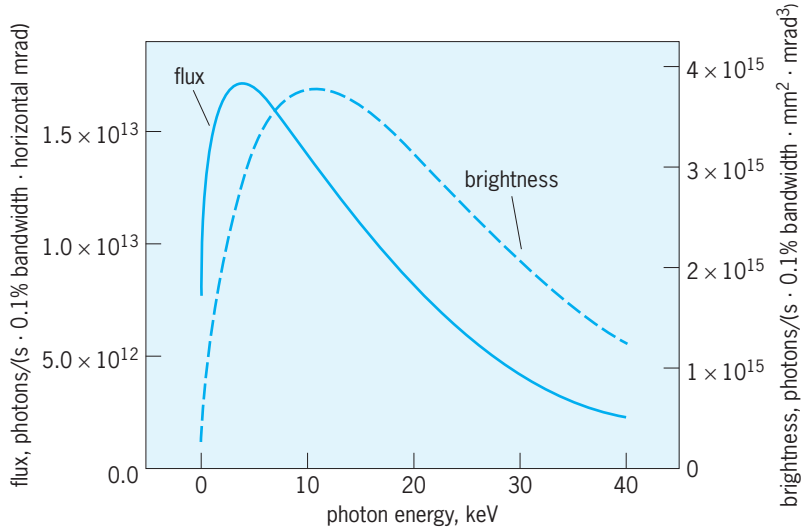

Fig. 6. Photon flux and brightness as functions of photon energy from superconducting bending magnets at the Advanced Light Source.

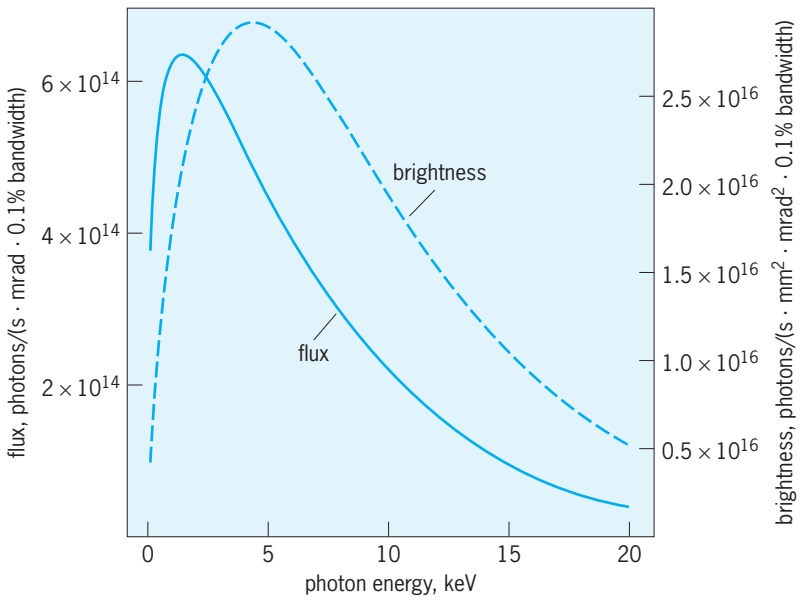

Fig. 7. Photon flux and brightness as functions of photon energy from a wiggler at the Advanced Light Source.

The spectrum from an undulator depends on the value of K. For values of K much less than 1, only the fundamental peak is important. For K approximately equal to 1, the power in the fundamental is a maximum and the first few harmonics have appreciable intensity. For K greater than 1, the wavelength of the fundamental becomes longer and more harmonics appear. For K much greater than 1, the fundamental has very long wavelength and there are many closely spaced harmonics. In this limit, the device is a wiggler, with the envelope of the spectrum approaching the continuous spectrum of that device. Figure 8 shows the spectral brightness of the radiation from an undulator in the Advanced Light Source (ALS), and shows the extremely high brightness offered by the undulator at discrete photon energies that can be varied over a range of values. The spectrum is the locus of the discrete spectral peaks as the undulator magnetic field is varied by changing the gap distance between the magnet arrays.

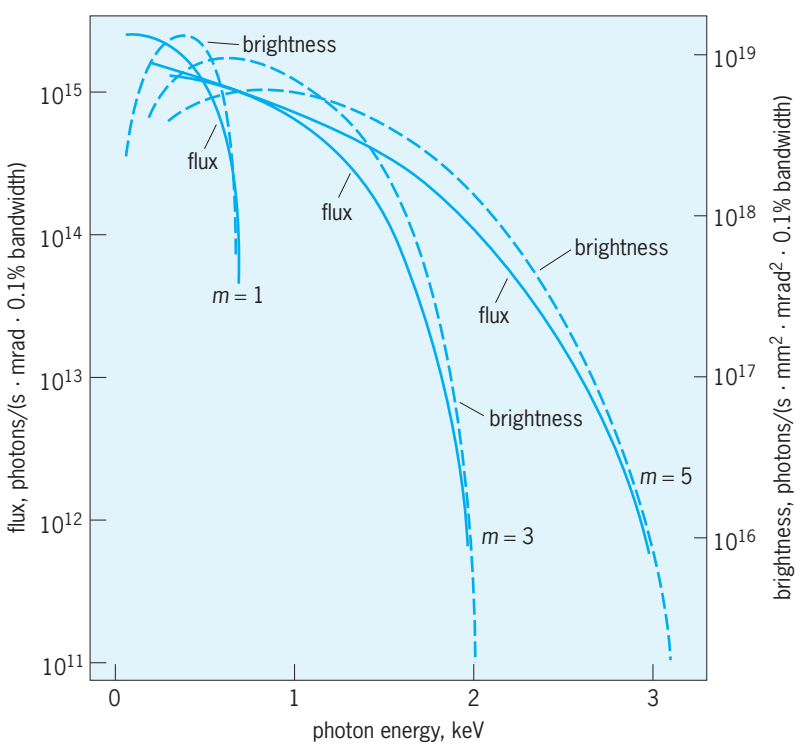

Fig. 8. Photon flux and brightness for an undulator at the Advanced Light Source. The letter *m* indicates the harmonic produced by the undulator. (Only the first few odd harmonics are useful.)

Polarization. The radiation from bending magnets and from wigglers and undulators with vertical magnetic fields is nearly 100% linearly polarized in the plane of the electron orbit, with the electric vector parallel to the electron acceleration. Out of the plane of the orbit, bending magnets produce elliptically polarized radiation. Undulators that produce circular, helical, or linear electron oscillations produce radiation whose polarization is controllable, allowing linear polarization in any plane and elliptical or circular polarization. Circular and helical polarization are used in many kinds of studies, especially of magnetic materials. *See* POLARIZATION OF WAVES.

Pulsed time structure. Radiation from a synchrotron storage ring is produced in pulses as each electron bunch sweeps by the observation point, due to the bunching of the electron beam by the radio-frequency accelerating system of the ring. The pulse duration can be as short as about 50 picoseconds and, in large rings, the interval between pulses can be 1 microsecond or more. This pulse structure facilitates study of time-dependent phenomena and is also used to enhance the data collection rate in some experiments (for example, by measuring photoelectron velocities using time-of-flight techniques).

Although femtosecond x-ray pulses will be most thoroughly exploited at free-electron lasers, it is now possible to produce 100-fs x-ray pulses (that is, pulses shorter than the electron bunch length), in lower fluxes in conventional storage rings. Femtosecond lasers can interact with electrons in a storage ring to produce femtosecond pulses of x-rays. Some ideas for manipulating the electron beam also may make it possible to produce pulses of around 1 ps duration with high flux, but such pulses have not been demonstrated so far. Ultrashort pulses are useful in many interesting studies, for example, investigations of the dynamics of phase changes in matter and the time evolution of electronic and magnetic processes.

Brightness. Brightness is a different property than photon flux, or intensity, which is measured in photons per second. Brightness is directed intensity and includes the source size and the solid angle of the radiation. In this regard, the comparison of a lightbulb with a laser pointer is instructive: The lightbulb radiates photons in most directions, while a laser pointer directs photons into a small spot and a narrow cone. Furthermore, a lightbulb produces a broad spectrum of radiation, while a laser pointer produces radiation in a very narrow range of wavelengths (for example, the red beam from a helium-neon laser). A conventional x-ray tube is analogous to a lightbulb, while

radiation from an undulator is analogous to the beam from a laser pointer.

Brightness is a conserved quantity and cannot be increased by an optical system such as a lens. Brightness is thus a property related to the radiation source, such as the low-emittance beam of electrons in a third-generation storage ring that is coupled with an undulator to produce the radiation.

A high-brightness beam can be focused to a small spot. This property is very important in spectroscopy with high spectral resolution, as high flux can be passed through a narrow slit in a monochromator. High brightness is also important in x-ray microscopy, such as in a scanning x-ray microscope, where spatial resolution is limited by the size of the photon beam. Finally, as noted above, a high-brightness beam can be partially coherent, allowing such applications as interferometry and speckle spectroscopy.

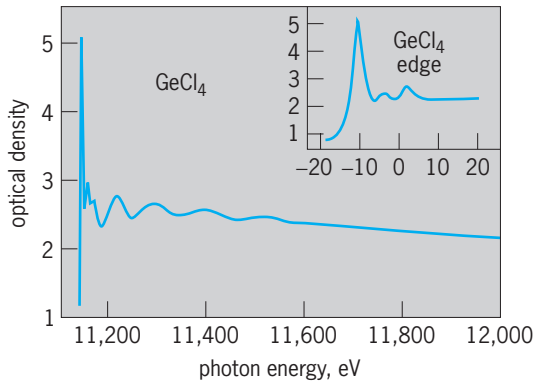

Fig. 9. X-ray absorption spectrum above the germanium K-edge in $GeCl_4$ vapor, showing oscillatory structure known as EXAFS. The inset shows the edge region with a shifted energy scale, displaying XANES spectrum. (*After B. M. Kincaid and P. Eisenberger, Synchrotron radiation studies of the K-edge photo-absorption spectra of Kr, Br_2, and $GeCl_4$: A comparison of theory and experiment, Phys. Rev. Lett., 34:1361–1364, 1975*)

Research Applications

Many of the conventional techniques and processes employing ultraviolet and x-radiation in research and technology have experienced considerable improvement as a result of one or more of the special properties of synchrotron radiation described above. For example, the wavelengths corresponding to the photon energies typically exploited in synchrotron radiation span length scales ranging from the atomic level to biological cells, a range that well suits advanced research in materials science, physical and chemical sciences, metrology, geosciences, environmental sciences, biosciences, medical sciences, and pharmaceutical sciences. The ability to probe and image small areas makes synchrotron radiation particularly suitable for the new field of nanoscience. The availability of synchrotron radiation has already made possible major scientific and technological advances.

The fundamental parameters that characterize the physical world (energy, momentum, position, and time) lend themselves to three broad categories of synchrotron experimental measurement techniques: spectroscopy, scattering, and imaging. Each of these techniques can be performed in a time-resolved mode by exploiting the short pulse lengths of synchrotron radiation.

Spectroscopy. Spectroscopy techniques are used to study the energies of particles that are emitted or absorbed by samples that are exposed to the light-source beam and are commonly used to determine the characteristics of chemical bonding and electron motion.

X-ray absorption spectroscopy (XANES and EXAFS). A plot of x-ray absorption versus photon energy (**Fig. 9**) shows a steep rise in the absorption at the absorption-edge energy, where the incident photon is just able to excite core electrons into empty states. X-ray absorption near-edge structure (XANES, also known as NEXAFS for near-edge x-ray absorption fine structure) is due to transitions into localized atomic levels which are modified strongly by the absorbing atom's surroundings. At higher energies, the observed oscillatory structure of the absorption, known as the extended x-ray absorption fine structure (EXAFS), is due to interference between the outgoing photoelectron waves and the electron waves backscattered from atoms adjacent to the absorbing atoms. This interference modulates the probability of exciting the electron. Because (1) XANES and EXAFS are significant modifications of the atomic absorption and (2) the absorption-edge energies of different atomic species are usually quite different, their analysis provides considerable information about the average atomic environment of each atomic species in complex, polyatomic solids, liquids, and gases.

The highly intense, continuous synchrotron radiation spectrum from multi-gigaelectronvolt storage rings is ideally suited for absorption measurements as a function of photon energy. Consequently, it has been used extensively for structural studies of amorphous materials, heterogeneous catalysts, and noncrystalline metalloproteins whose atomic arrangements are not ordinarily determinable by other structural techniques. These structural studies have led, in turn, to increased understanding of the properties of these materials. EXAFS patterns have also been measured in fractions of a second, so that time-resolved studies of structural changes may be anticipated.

Similar modulations of the x-ray fluorescence yield, the Auger electron yield, and the desorbed atom yield versus photon energy are caused by the same interference phenomenon. The fluorescence yield is used extensively to study the atomic environments of extremely dilute (less than 100 parts per million) constituents of complex systems such as alloys, semiconductors, toxic waste, and proteins in solution. The Auger electron yield is used for the determination of the atomic environments of atoms at the surfaces of materials. The desorbed atom yield is used to determine the surroundings of atoms adjacent to specific atoms on a surface. For example, it is used to determine the environments of those metal atoms adjacent to oxygen atoms on a partially

oxidized metal surface. The high-brightness beams achieved with insertion devices in second- and third-generation sources have led to x-ray absorption spectroscopy microscopy in which micrometer-size areas are examined. *See* AUGER EFFECT; DESORPTION; SPECTROSCOPY; SURFACE PHYSICS; X-RAY FLUORESCENCE ANALYSIS.

Photoemission spectroscopy. Energy and angular distributions of electrons ejected from solids by photons of varying incident energy provide basic information about band structure and core levels of both the surface and bulk of the sample. The extremely high intensity and tunability of synchrotron radiation in the ultraviolet and soft x-ray regions of the spectrum have made possible high-resolution studies which delineate the band structures of crystalline solids, as well as electronic states in highly correlated solids, such as high-temperature superconductors, in which band theory is inadequate. In addition, electrons of energy between 30 and 100 eV have such a large probability for interaction with other atoms that they have an escape depth of only about 0.5 nanometer in most solids. As a result, the detected electrons reveal the properties of surface electronic energy levels of clean materials and the extrinsic surface states of chemisorbed material, opening a route to the study of the phenomena of oxidation, corrosion, and catalysis. Such studies have led to major revisions of models of the metal–semiconductor interface which determines many properties of solid-state electron devices. Again, the tunability, when coupled with the high intensity of synchrotron radiation, makes possible the separate delineation of surface and bulk electronic states. High-brightness beams have led to photoelectron microscopy whereby electronic states in micrometer-size areas are examined. *See* BAND THEORY OF SOLIDS; SEMICONDUCTOR; SUPERCONDUCTIVITY; SURFACE PHYSICS.

In addition, synchrotron radiation facilitates photoemission studies in two ways: the high-vacuum environment of the storage ring minimizes contamination of clean surfaces prepared in place; and the sharply pulsed time structure of the radiation makes it possible to measure photoelectron energies by time of flight. *See* ELECTRON SPECTROSCOPY; PHOTOEMISSION.

X-ray fluorescence trace-element analysis. The high intensity, tunability, high collimation, and high linear polarization of synchrotron radiation make it valuable for trace-element analysis through x-ray fluorescence. Such trace-element analysis is used to determine very low concentrations of impurities introduced in semiconductor processing procedures, with the goal of further decreasing those concentrations (**Fig. 10**).

Scattering. Scattering or diffraction techniques make use of the patterns of light produced when x-rays are deflected by the closely spaced lattice of atoms in solids and are commonly used to determine the structures of crystals and large molecules such as proteins.

X-ray diffraction and scattering. Synchrotron radiation from multi-gigaelectronvolt storage rings is characterized by its high intensity, natural collimation, tunability, and unique polarization and time structure. These properties have contributed to its profound impact in the field of x-ray diffraction and scattering. The high intensity and natural collimation of the

Fig. 10. Worker checking silicon wafers for minute impurities at the Stanford Synchrotron Radiation Laboratory (SSRL). (*Photo by Peter Ginter*)

x-rays have made measurements of the atomic arrangements on surfaces and at interfaces of liquids and solids a routine technique. Detailed studies have been made of the structural changes that occur during two-dimensional phase transitions (melting and freezing) as well as changes in these processes as the system under study goes from a two-dimensional system to a three-dimensional system through the addition of individual atomic layers. *See* PHASE TRANSITIONS.

Anomalous (or resonant) scattering has been utilized in wide-angle diffraction experiments and small-angle x-ray scattering. Close to an atom's absorption edge, the atom's scattering cross section (atomic scattering factor) can vary rapidly as a function of incident x-ray energy. This anomalous scattering is used to alter the contribution of that atomic species to the overall scattering amplitude and thereby obtain additional information about the overall structure of the sample. The tunability of the highly monochromatized synchrotron radiation makes such experiments possible. *See* SCATTERING OF ELECTROMAGNETIC RADIATION.

The use of polarized x-rays makes possible studies of the magnetic properties of materials. The interaction of x-rays and the spin and orbital angular momentum of the electron is very weak, but in some cases it can be enhanced by many orders of magnitude by tuning the incident x-ray energy to an absorption edge of one of the constituent elements in the sample. This resonant scattering technique not only allows the bulk magnetic properties of materials to be studied by x-rays but also permits surface and interface magnetic properties to be explored. *See* ELECTRON SPIN; MAGNETISM.

The high brightness and flux of synchrotron radiation is routinely exploited to perform time-resolved scattering and diffraction experiments. Measurements on a millisecond time scale have been made to monitor the structure of muscles during contraction in an effort to understand the associated molecular rearrangements that occur during the contraction process. Even shorter time scales can be attained by using the pulsed nature of the source. Nanosecond time-resolved studies have been made of melting and recrystallization of semiconductors irradiated with short-pulse lasers, and the structural dynamics of macromolecules undergoing laser photolysis have been studied using synchrotron radiation. *See* LASER PHOTOCHEMISTRY.

All the experiments described so far deal with elastic scattering, in which the scattered x-ray has the same energy as the incident photon. Elastic scattering is not the only process that can occur in a sample. The incident x-ray can give up (or gain) some energy in the scattering process by exchanging energy with excitations in the sample. Energies associated with electronic excitations in samples are typically several hundred millielectronvolts, while vibrational excitations (for example, phonons in crystals) have energies of many millielectronvolts. One of the challenges of performing inelastic x-ray scattering experiments is to prepare monochromators and analyzers (to separate the inelastically scattered x-rays from elastically scattered ones) with energy resolutions, $\Delta E/E$, of a few times 10^{-7} (several millielectronvolts per 10^4 eV). The combination of very high collimation and high spectral brightness (number of photons per second per unit energy width, ΔE) of undulators on third-generation sources has allowed such experiments to be performed. Phonon dispersion curves of several materials have been mapped out. Using resonant techniques, the vibrations of specific atomic species can be isolated in complex systems and their partial densities of states can be measured. The partial densities of states related to the iron atoms in myoglobin and hemoglobin have been determined in this manner. *See* LATTICE VIBRATIONS; PHONON; X-RAY DIFFRACTION.

Structural molecular biology. Structural molecular biology is growing at an extraordinary pace, due in no small part to the impact of synchrotron radiation research. In addition to basic research, a significant contemporary use of macromolecular crystallography is so-called structure-based drug design, which enables pharmaceutical companies to develop potential drug candidates with the knowledge of where and how they bind to their targets in cellular structures (**Fig. 11**). Structural analysis of macromolecules (specifically proteins and nucleic acids) can, in many cases, be performed with considerably better resolution using synchrotron radiation as compared to laboratory x-ray sources. In particular, there has been progress in overcoming two major limitations of conventional x-ray diffraction research, associated with radiation damage and small crystal size. Frequently the limitation on the number of Bragg peaks whose intensity can be accurately measured is the x-ray damage of the crystal itself. However, it has been observed that a given dose of x-rays does

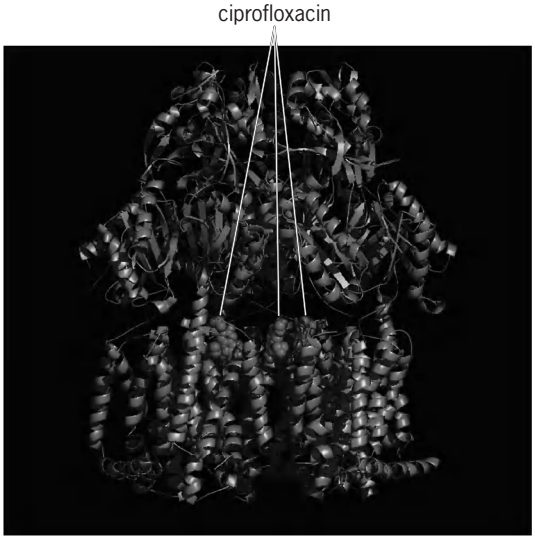

Fig. 11. Ciprofloxacin, an antibiotic used to treat a variety of bacterial infections including inhaled anthrax, bound in the large central cavity of *Escherichia coli* AcrB, a bacterial protein complex that repels a wide range of antibiotics. (*From Lawrence Berkeley Laboratory, Research News: Secrets of Drug Resistance Revealed, May 9, 2003*, http://www.lbl.gov/Science-Articles/Archive/PBD-drug-resistance.html)

less damage if applied over a short time; that is, more intense x-ray beams allow the data to be collected before the crystal is destroyed. As a result, synchrotron radiation makes possible the determination of many more measurable peaks before the crystal is damaged by the probing radiation. Crystals held at cryogenic temperatures have been found to be even more resistant to radiation. The combined effects of brief x-ray exposure and cryogenic temperatures allow many more macromolecular crystals to be candidates for crystal structural analysis. The quality of data from those crystals has improved significantly, with advanced analysis techniques producing resolution better than 0.1 nm. Structure of protein crystals whose unit cells have sides of over 100 nm are now being solved.

Anomalous scattering may be used to help solve the so-called phase problem in crystallography. Using the tunability of the synchrotron radiation, data sets can be collected at several wavelengths around the absorption edge of a strongly absorbing atom (an atom with a relatively high atomic number) contained within the macromolecule. This technique of multiwavelength anomalous dispersion (MAD) eliminates the necessity of growing crystals with three different isomorphic chemical forms [the multiple isomorphous replacement (MIR) technique], an important advantage since it is frequently impossible to grow crystals in three different forms. This technique was more fully realized when it was found possible to incorporate selenium in place of sulfur (selenomethionine in place of methionine) in samples grown in certain strains of *Escherichia coli*.

It is often difficult, if not impossible, to grow the large single crystals of macromolecules that were needed for structural analysis with laboratory x-ray sources. The high intensity of synchrotron radiation allows structures to be determined from crystals only tens of micrometers on a side.

Crystallographers are no longer confined to static observations of protein structures. The availability of extremely intense, multiwavelength, pulsed synchrotron x-ray sources has reduced exposure times enough to capture "movie sequences" of fundamental molecular processes. This approach has been applied to studies of myoglobin, the iron-based molecule responsible for oxygen transport in muscles. The entire photolysis, relaxation, and rebinding processes occur in less than 5 milliseconds at room temperature. Researchers are now extending this approach to several other light-sensitive signaling systems that are chemically and biologically diverse, and are developing new techniques that will enhance the time resolution from the nanosecond range, first to a few hundred picoseconds and perhaps ultimately to femtoseconds. *See* MOLECULAR BIOLOGY; X-RAY CRYSTALLOGRAPHY.

Lensless imaging. In a big step beyond the periodic objects studied by crystallography, it is possible to reconstruct images of nonperiodic samples using x-ray diffraction imaging (XDI) or "lensless imaging." Like crystallography, x-ray diffraction imaging is based on the analysis of diffraction patterns, but it uses iterative algorithms to extract the phase information needed to reconstruct the object and requires that the diffraction intensity be zero outside the object's boundary. The better this boundary is known, the faster the iterations converge to an accurate image. The boundary can be supplied by other imaging techniques or by more sophisticated iterative computation methods. Two-dimensional images of objects as complex as yeast cells with resolutions of a few tens of nanometers have been achieved, as have three-dimensional images of simpler, solid test objects, such as collections of gold balls. It is anticipated that a three-dimensional resolution of 10 nm will ultimately be possible for life science samples, where radiation damage is an issue, and 2 nm for solids.

Imaging. Imaging techniques use the light-source beam to obtain pictures with fine spatial resolution of the samples under study, and are used in diverse research areas such as cell biology, lithography, infrared microscopy, radiology, and x-ray tomography. There are two basic experimental methods, scanning and full-field imaging. In either case, the contrast in the images must have some physical cause. This contrast can be due to changes in absorption (amplitude contrast) or refractive index of the sample (phase contrast) caused by, for example, variable elemental or chemical composition or magnetic properties. The tunability of synchrotron radiation is absolutely essential for the creation of contrast mechanisms. Tomography can be combined with x-ray microscopy for full three-dimensional images, a technique that is especially useful in the life sciences in imaging cellular structures. Newer "lensless" imaging schemes based on reconstruction of scattering patterns made with coherent x-ray beams are now coming into use. Related to imaging is spatially resolved spectroscopy and diffraction, which allow the construction of images based on structural features other than density, such as an absorption feature, local lattice strain, and many others. Multidimensional data sets that contain, in the case of spectroscopy, a complete spectrum for each point (to the resolution of the technique being used) in the image can then be analyzed in multiple ways.

Owing to their somewhat different properties, soft and hard x-rays differ in their imaging applications. The wavelengths of soft x-ray photons (1–15 nm) are very well matched to the creation of nanoscopes capable of probing the interior structure of biological cells and inorganic mesoscopic systems. Problems addressed by soft x-ray imaging techniques include cell biology, nanomagnetism, environmental science, and soft matter and polymers. Hard x-ray imaging nondestructively visualizes samples, and frequently their internal or hidden components. It is applicable to nearly all fields of science from the life sciences to engineering and archeology. Electron microscopes will always have better spatial resolution, but they will be limited in the range of samples that they can study. The uniqueness of x-ray imaging resides in its deeper penetration, enabling the study, for example, of buried interfaces and wet biological samples.

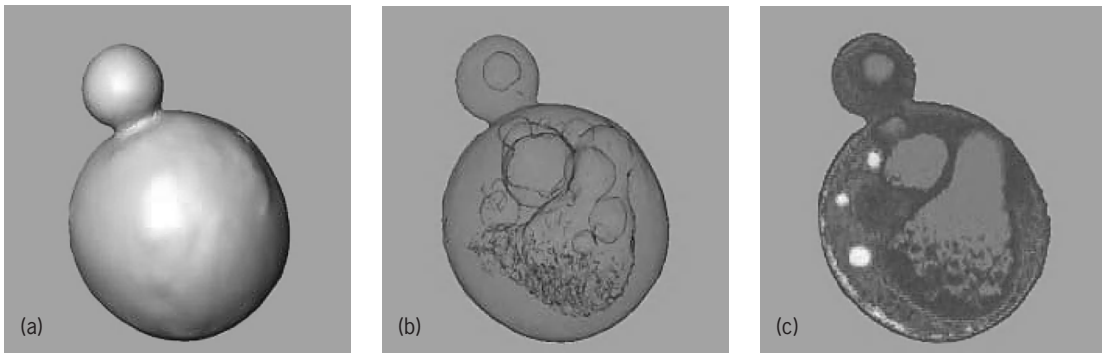

Fig. 12. Reconstructed data of a budding yeast, using different volume-analysis algorithms. (*a*) Opaque surface. (*b*) Transparent surface showing internal vesicles. (*c*) Volume-rendered thick-slice section. (*From C. A. Larabell and M. A. Le Gros, X-ray tomography generates 3-D reconstructions of the yeast, Saccharomyces cerevisiae, at 60-nm resolution, Mol. Biol. Cell, 15:957–962, 2004*)

Ampitude contrast. Literally from the moment they were discovered in 1895, x-rays have been used to image visibly opaque objects. This type of imaging, familiar from medical x-rays, is based on the contrast that arises from the attenuation of the amplitude of the incident x-ray wave by the different constituents of the sample. The focusing required for imaging in the x-ray region is difficult, but several advances in x-ray optics, coupled with the very high brightness of third-generation sources, have led to important advances in this area, and there are now several approaches in which x-ray beams with dimensions of 15 nm and higher can be produced for both scanning and full-field imaging techniques.

In scanning imaging, a very small illuminated spot is created on the sample using focusing devices. Fresnel zone plates and specially curved mirrors perform the same function for x-rays that lenses do for visible light. The image is then built up by "raster scanning" the sample through the illuminated spot. The species detected can be photons (transmitted or fluorescent) or photoelectrons. With the use of two zone plates (a condenser and an objective), it is possible to create a full-field image. In photoemission electron microscopy, a smallish spot on the sample is illuminated and the emitted photoelectrons are passed through an electron microscope column to produce a magnified full-field image. *See* DIFFRACTION; ELECTRON; MICROSCOPE.

Imaging methods can then be employed as probes, using any of the techniques previously discussed to carry out spatially resolved spectroscopy, microdiffraction, and microfluorescence analysis. Two-dimensional distributions of strains, chemical elements, and valence states of a particular element can be mapped out. These techniques allow the determination of the strains in individual grains in a metal, the elemental distributions of metals in cells treated with anticancer drugs, and the oxidation states in heavy metals that accumulate in the root systems of plants being considered for bioremedial applications. X-ray microscopes are fast becoming one of the most sought-after instruments at synchrotron radiation facilities. *See* X-RAY MICROSCOPE.

X-ray tomography. Imaging at multiple angles can also be used to reconstruct three-dimensional pictures of objects with submicrometer resolution using the same techniques as are used on larger scales for medical purposes in computerized tomography (CT) scans. In the life sciences, x-ray tomography is the first high-throughput imaging technology that generates images of whole, hydrated cells at better than 60-nm resolution (**Fig. 12**), thereby bridging the middle area between light (200 nm) and electron microscopy (0.3 nm). High-resolution images provide contrast between cellular structures and allow for discernment of individual structures. After data collection, tomographic techniques are used to reconstruct the original information into quantifiable three-dimensional views of the entire cell. Through the use of computer algorithms, the researchers then process the reconstructed data to create made-to-order images of whole cells and their internal structures. In the physical sciences, three-dimensional maps of the concentrations of individual atomic species in complex, polyatomic materials can be constructed. Applications range from the study of petroleum distributions in soils to the analysis of osteoporosis development and treatment. *See* COMPUTERIZED TOMOGRAPHY.

Phase contrast. When extremely bright (that is, partially coherent) x-ray beams are employed for the purposes of imaging, contrast mechanisms other than absorption may be used, in particular phase-contrast imaging. Phase-contrast imaging does not rely on attenuation of the beam for image contrast, but on the variation of the refractive index of the constituent components. The variation of refractive index results in a distortion of the phase fronts of a (partially) coherent incident beam as it traverses the sample. Since the x-rays are coherent, the direct and scattered waves can interfere with each other, resulting in a modulation of the amplitude, and hence intensity, of the x-ray beam. Phase-contrast imaging is particularly useful for imaging thin biological samples where the sample is composed primarily of low-atomic-number atoms, resulting in weak absorption contrast. **Figure 13** shows an image of an ant's head taken with phase-contrast imaging techniques.

Phase-contrast imaging is related to another type of imaging phenomenon in which the interference pattern from the combined incident and scattered

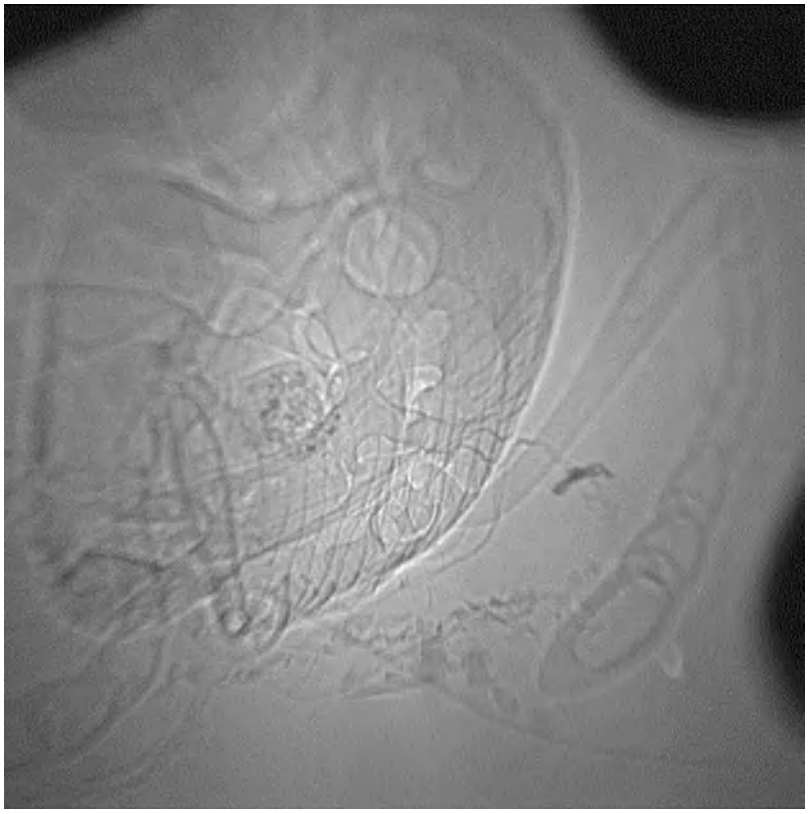

Fig. 13. Phase-contrast image of an ant's head, acquired at one of the undulator beam lines at the Advanced Photon Source, Argonne National Laboratory.

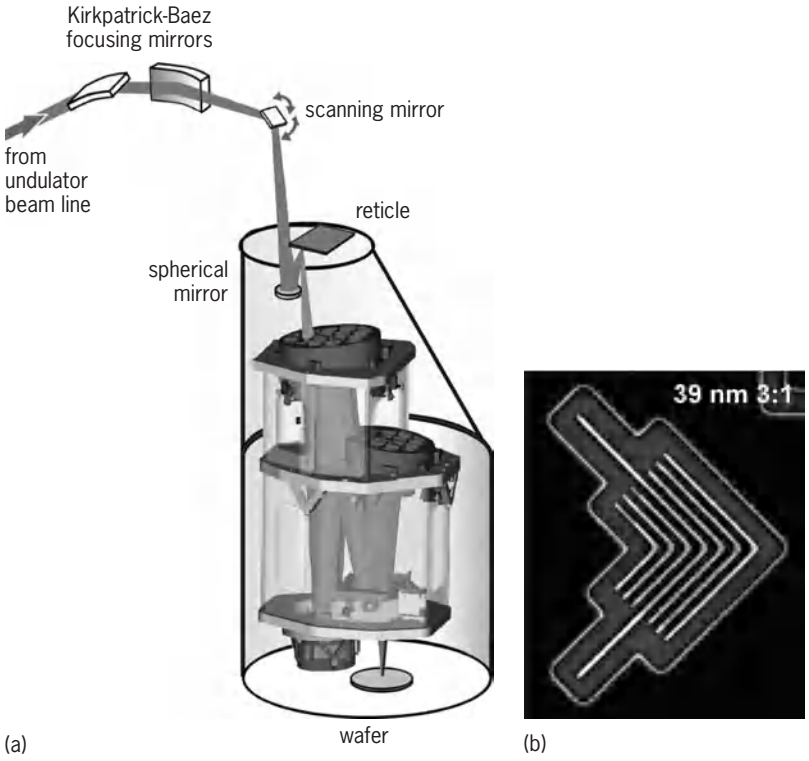

Fig. 14. Extreme ultraviolet (EUV) lithography. (*a*) Setup of equipment. The scanning mirror directs focused EUV light from the Kirkpatrick-Baez mirrors to the mask-carrying reticle with the desired degree of spatial coherence and illumination pattern. The optic images the EUV light reflected from the reticle onto a resist-covered silicon wafer. (*b*) Printed elbow test pattern with 39-nm-wide lines and 3:1 line-to-spacing ratio. (*From P.P. Naulleau et al., Static microfield printing at the Advanced Light Source with the ETS Set-2 optic, Proc. SPIE, 4688:64–71, 2002*)

waves is recorded, namely holography. Capturing a three-dimensional image with spatial resolution of a fraction of a nanometer is the ultimate goal of the x-ray imaging scientist. Third-generation sources are on the verge of achieving this capability, and the next generation of sources may capture three-dimensional images of cells and other extremely small objects with a single 100-Fs flash of x-rays from a free-electron laser. *See* HOLOGRAPHY; PHASE-CONTRAST MICROSCOPE.

Imaging of humans. Coronary arteries are difficult to image with normal radiography techniques because the heart and other surrounding tissue absorb much more radiation than do the arteries. In order to achieve contrast, it is necessary to introduce into the arteries a dye containing elements such as iodine that heavily absorb x-rays. In synchrotron radiation angiography, a much lower iodine concentration is needed, making it possible to use less hazardous catheterization procedures to introduce the iodine. The contrast is achieved by taking images in rapid succession at two different x-ray wavelengths, one just above and one just below the iodine absorption-edge wavelength. Subtraction of one image from the other yields an image of iodine-containing arteries. *See* MEDICAL IMAGING; RADIOGRAPHY.

In conventional mammograms, differences in tissue densities and composition show up as contrasting areas due to x-ray absorption, allowing doctors to see tumors or changes in tissue. However, differences between healthy and cancerous tissues are very small, and scattering of x-rays can lead to blurring and even lower contrast. A mammography technique called diffraction-enhanced imaging (DEI) uses ultrabright x-rays and provides a dramatic contrast between normal tissues and tumors. The DEI method uses a single-energy beam of x-rays instead of the broad-energy beam used in conventional imaging, and an analyzer crystal that can differentiate between x-rays that are traveling much less than one ten-thousandth of a degree apart, thereby reducing scatter, and helps visualize low-contrast areas that otherwise would be lost. In addition to mammography, potential applications of DEI include other low-contrast tissues and organs such as kidneys, and the nondestructive testing of materials.

X-ray lithography. A technique used in the art world for many centuries, lithography has been adopted, and adapted with phenomenal success, by high-technology industries. In microchip manufacturing, a silicon wafer is coated with a thin layer of photosensitive material called a resist. An image of a mask containing the desired pattern is projected onto the resist. The exposed (or unexposed) parts of the resist are etched away and, with further processing, the desired circuit is built up. The same basic process can be used in the manufacture of small mechanical components known as microelectromechanical structures (MEMS). In MEMS, features are not always as fine as in microchip circuits, but they must be sufficiently robust to withstand stresses due to fluid flow (in miniature chemical processing systems) or to mechanical motion (in gears, motors, and the like).

Work at synchrotron light sources focuses primarily on the exposures of the resists for MEMS, for optical devices such as zone plates, and for testing components and processes for microchip fabrication. The features that make synchrotron radiation attractive for this purpose are (1) high intensity, which provides for short exposures and therefore high throughput; (2) high collimation, which makes possible high resolution; and for MEMS (3) the ability to obtain high aspect ratios. In addition, hard x-rays (with energies greater than 12 keV) can penetrate deep into the photosensitive resist, making structures with high aspect ratios (deeper than they are wide). **Figure 14** shows an example of extreme ultraviolet lithography. Undulator light from a synchrotron is used to expose a test pattern onto a silicon wafer coated with resist, resulting in a test pattern with 39-nm lines. *See* INTEGRATED CIRCUITS; MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS); MICROLITHOGRAPHY.

Alfred S. Schlachter; Arthur L Robinson;
Arthur Bienenstock; Dennis Mills;
Gopal Shenoy; Herman Winick

Bibliography. M. Altarelli, F. Schlachter, and J. Cross, Making ultrabright x-rays, *Sci. Amer.*, 279(6):66–73, December 1998; F. Ciocci (ed.), *Insertion Devices for Synchrotron Radiation and Free Electron Laser*, World Scientific, 2000; P. J. Duke, *Synchrotron Radiation: Production and Properties*, Oxford, 2000; P. Eisenberger, *Materials Science Using Synchrotron Radiation*, Elsevier, 2000; M. R. Howells, J. Kirz, and D. Sayre, X-ray microscopes, *Sci. Amer.*, 264(2):88–94, February 1991; E. E. Koch, T. Sasaki, and H. Winick (eds.), *Handbook on Synchrotron Radiation*, vols. 1–4, Elsevier, 1983–1991; T.-K. Sham (ed.), *Chemical Applications of Synchrotron Radiation*, World Scientific, 2002; H. Winick (ed.), *Synchrotron Radiation Sources: A Primer*, World Scientific, 1995.

Syncline

In its simplest form, a geologic structure marked by the folding of originally horizontal rock layers into a systematically curved, concave upward profile geometry (**illus.** *a*). A syncline is convex in the direction of the oldest beds in the folded sequence, concave in the direction of the youngest beds. Although typically upright, a syncline may be overturned, recumbent, or upside down (illus. *d*). Synclines occur in all sizes, from microscopic to regional. Profile forms may be curved smoothly (illus. *a*) to sharply angular (illus. *b*). Fold tightness of a syncline, as measured by the angle at which the limbs of the syncline join, may be so gentle that the fold is barely discernible, to so tight that the limbs are virtually parallel to one another (illus. *c*). The orientation of the axis of folding is horizontal to shallowly plunging, but synclines may plunge as steeply as vertical.

Synclines are products of the layer-parallel compression that arises commonly during mountain building. Stresses parallel to or at an inclined angle to rock layering can achieve shortening of the layers through buckling and bending. True bending is accommodated often by flexural slip, a mechanism in which the rock layers, like slick pages in a book or magazine, slip past one another along bedding-plane weaknesses. The final profile form of the fold reflects the mechanical properties of the rock sequence under the temperature-pressure conditions of folding, and the percentage of shortening required by the deformation. *See* ANTICLINE; FOLD AND FOLD SYSTEMS.

George H. Davis

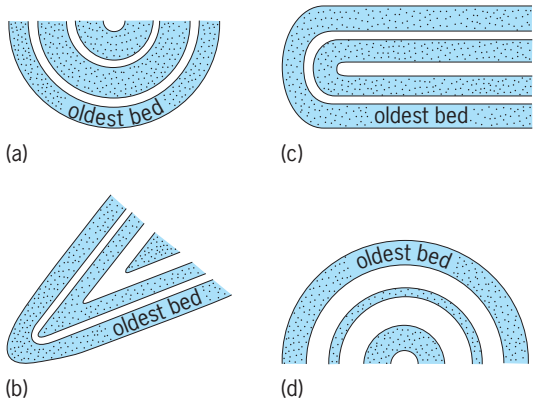

Varieties of synclines as seen in profile view. (*a*) Upright syncline with smoothly curved limbs. (*b*) Overturned, sharply angular syncline with planar limbs. (*c*) Recumbent, isoclinal syncline with parallel limbs. (*d*) Upside-down syncline, sometimes called an antiformal syncline.

Bibliography. G. H. Davis, *Structural Geology*, 2d ed., 1996; S. Mitra and G. W. Fisher, *Structural Geology of Fold and Thrust Belts*, 1992; N. J. Price and J. W. Cosgrove, *Analysis of Geological Structures*, 1991.

Synthetic aperture radar (SAR)

Radar, airborne or satellite-borne, that uses special signal processing to produce high-resolution images of the surface of the Earth (or another object) while traversing a considerable flight path. The technique is somewhat like using an antenna as wide as the flight path traversed, that being the large "synthetic aperture," which would form a very narrow beam. Synthetic aperture radar is extremely valuable in both military and civil remote-sensing applications, providing surface mapping regardless of darkness or weather conditions that hamper other methods.

Basic theory. Resolution is the quality of separating multiple objects clearly. In radar imaging, fine resolution is desired in both the down-range and cross-range dimensions. In radar using pulses, down-range resolution is achieved by using broad-bandwidth pulses, the equivalent of very narrow pulses, allowing the radar to sense separate echoes from objects very closely spaced in range. This technique is called pulse compression; resolution of a few nanoseconds (for example, 5 ns $= 5 \times 10^{-9}$ s gives about 0.75 m or 2.5 ft resolution) is readily achieved in modern radar.

Cross-range resolution is much more difficult to

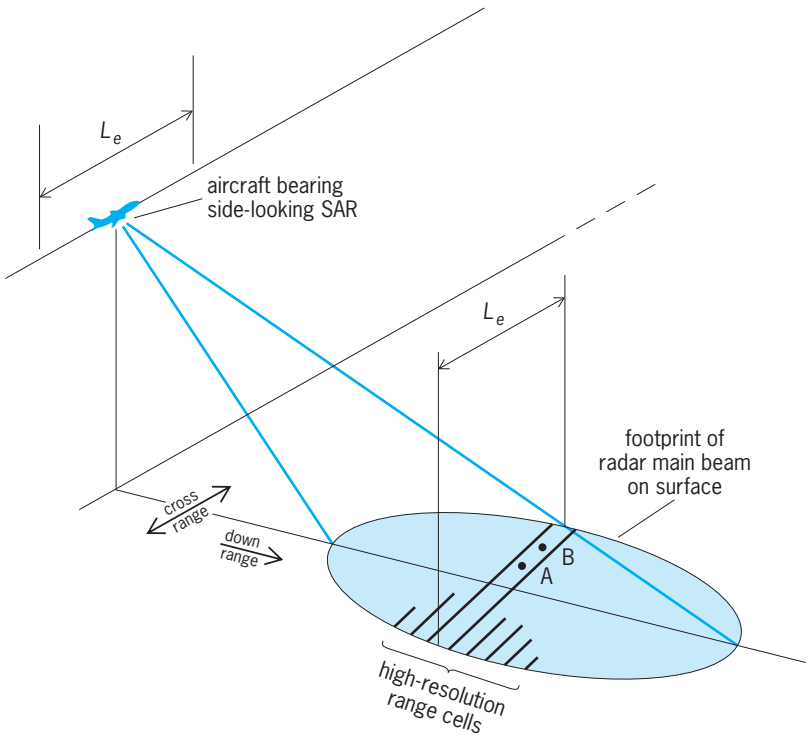

Fig. 1. The basic idea of synthetic aperture radar (SAR); a side-looking case is illustrated. Two example scatterers, A and B, are shown in the ground scene. L_e = maximum flight path length for effective SAR processing.

achieve. Generally, the width of the radar's main beam determines the cross-range, or lateral, resolution. For example, a 3° beam width resolves targets at a range of 185 km (100 nautical miles) only if they are separated laterally by more than 100 m (330 ft), not nearly enough resolution for quality imaging.

However, surface objects produce changing Doppler shifts as an airborne radar flies by. In side-looking radar (**Fig. 1**), even distant objects actually go from decreasing in range very slightly to increasing in range, producing a Doppler-time function (**Fig. 2**). If the radar can sustain high-quality Doppler processing for as long as the "footprint" of the beam illuminates the scene, these Doppler histories will reveal the lateral placement of objects. In fact, if such processing can be so sustained, the cross-range resolution possible is one-half the physical width of the actual antenna being used, a few feet perhaps. Furthermore, this resolution is independent of range, quite unlike angle-based lateral resolution in conventional radar. *See* DOPPLER EFFECT.

Many synthetic aperture radars use other than just a fixed side-looking beam. Spotlighting involves steering the beam to sustain illumination for a longer time or to illuminate a designated scene at some other angle. The principles remain unchanged: fine resolution in both down-range and cross-range dimensions (achieved by pulse compression and Doppler processing, respectively) permits imaging with picture cells (pixels) of remarkably fine resolution (**Fig. 3**). Many synthetic aperture radars today achieve pixels of less than 1 m (3.3 ft) square.

Focusing. Echoes from objects as the beam footprint passes over them have range-dependent rates of change of Doppler shift (Fig. 2), changing from positive to negative rapidly for objects close to the flight path and more slowly for objects at greater range. Conventional Doppler filtering, essential for clutter suppression in ordinary airborne radar, separates one fixed Doppler shift from others. Such processing is inadequate in synthetic aperture radar because of the need to compensate for the changing Doppler shift associated with the range of interest. Using such a reference in the signal processing results in good images being formed in swaths, regions relatively well focused around the reference range. Echoes from other ranges do not produce sharp images, unless processed separately with appropriate references. Synthetic aperture radars that form images in this range-dependent way are said to be focused.

Focusing must also compensate for variations in the radar's flight path that would introduce Doppler shifts not related to the object's position. While measurement of the flight-path deviations could be attempted (as with very sensitive instruments in the aircraft), automatic focusing is used now in most synthetic aperture radars. Autofocusing derives a reference signal (various techniques exist) from the scene itself that compensates for flight-path irregularities.

Waveform limitations. Radars determine Doppler by sensing the progressive phase change from pulse to pulse in a coherent dwell of many pulses. (Coherent means simply that both the amplitude and the phase of the returns are used in the processing.) Resolution in Doppler depends upon the length of time of the coherent processing. In synthetic aperture radar, that length of time is limited by the illumination time of the scene, clearly the width of the beam's footprint at the range of interest divided by the aircraft speed (again, for a fixed-beam side-looking radar). It would appear that if the antenna were arbitrarily small in width, the beam would be very broad indeed and the processing time would

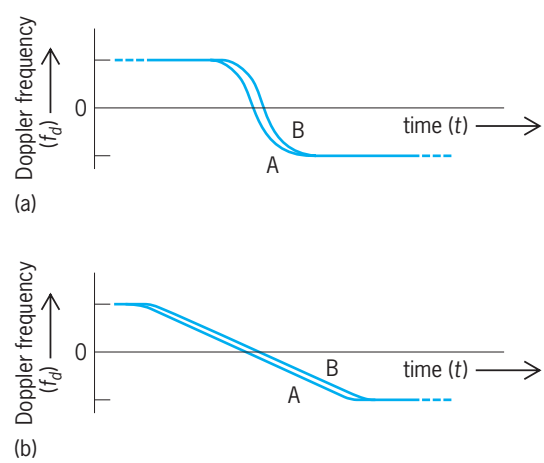

Fig. 2. The range-dependent Doppler-time function of surface objects (reference objects A and B of Fig. 1), as the synthetic aperture radar passes by, for (*a*) A and B close to the flight path and (*b*) A and B at greater range.

be great, giving very fine resolution in cross-range. While those relationships are true, a sampling problem arises related to the beam's footprint. The pulse repetition frequency (prf) must be high enough that no Doppler ambiguity occurs in the range of Doppler shifts being experienced across the beam's footprint. The wider the beam, the higher the prf must be. On the other hand, the beam's down-range footprint suggests that the prf must be low enough that no range ambiguity exists in that dimension. This latter limit on the prf can be eased somewhat by allowing poorly focused overlap (range-ambiguous echoing) from nearer and farther objects, perhaps with acceptable image degradation. *See* COHERENCE.

Continuing research in synthetic aperture radar processing addresses methods to overcome these waveform limitations: Forming multiple beams in cross-range and processing in separate beam-related channels may permit lower prfs; forming adaptive nulls in the down-range antenna pattern to mitigate ambiguous range return may permit higher ones. In any case, the bounding of prf imposed by the beam's footprint precludes use of an arbitrarily small antenna to achieve arbitrarily fine cross-range resolution.

Interferometric SAR. Modern synthetic aperture radar is able to form topological or three-dimensional images using interferometry. A synthetic aperture radar system using two radars, or at least two receiving antennas, spaced vertically by many wavelengths, can determine the relative height of each pixel quite accurately by interpreting the phase difference of the returns at the two antennas. In flat terrain, the down-range pattern of constructive and destructive interference seen at the two antennas would produce regularly spaced isophase contours in the interferogram covering the scene. Changes in terrain elevation distort these contours, and appropriate image processing can produce accurate topological data. The United States' Shuttle Imaging Radar (SIR) used this technique in the mid-1990s with a single radar making two passes over the same scene with different orbit altitudes, and in February 2000 using two antennas separated by 60 m (197 ft) for single-pass operation. *See* INTERFEROMETRY; SPACE SHUTTLE.

Polarimetric SAR. The polarization of an electromagnetic wave describes the spatial orientation of the electric field component. How an object reflects the radar signal in both the transmitted polarization and the orthogonal one is very significant in classifying that object. A polarimetric radar makes such measurements. Many synthetic aperture radars are polarimetric, permitting classification of the terrain (the kind of vegetation, distinctions between natural and artificial objects, and so forth). Adaptive systems can estimate the polarimetric state of uninteresting returns (the clutter) and reduce sensitivity to it, increasing the contrast between such background and pixels containing more interesting objects. Furthermore, if knowledge of the polarimetric character of a target of particular interest is available, sensitivity to its presence may be increased by polarimetric weighting techniques. Many synthetic aperture radars are multifrequency (operating simultaneously in many bands) and polarimetric. *See* POLARIZATION OF WAVES.

Fig. 3. Synthetic aperture radar image of an office building in Lincoln, Massachusetts. (*Lincoln Laboratory, Massachusetts Institute of Technology*)

Inverse SAR. Another technique for imaging targets is closely related to synthetic aperture radar, but is applicable to even fixed land-based radars. As an aircraft, for example, approaches a radar equipped with synthetic-aperture-radar-like Doppler processing in highly resolved range cells, the slight rotation of the target itself relative to the radar provides a meaningful Doppler spectrum in each range cell, indicative of the lateral position of the scatterers in that cell. Autofocusing is used to compensate for the translation of the target and for flight instabilities, leaving just the Doppler shift due to the target's orderly change of aspect. As in synthetic aperture radar, a few seconds of coherent processing is typically sufficient to achieve the Doppler resolution necessary for the desired cross-range resolution. At microwave wavelengths, only a few degrees of aspect change are necessary to produce discernible Dopplers. Since the radar is stationary and the target moving, this technique is called inverse synthetic aperture radar (ISAR). **Figure 4** shows images of an aircraft obtained in experiments with a C-band ISAR in China, where different methods of autofocusing were being explored. Ships, with their somewhat regular local rotations (rolling, pitching), have been imaged successfully with ISAR.

Other considerations. A limitation in image quality comes from the convenient but false assumption that

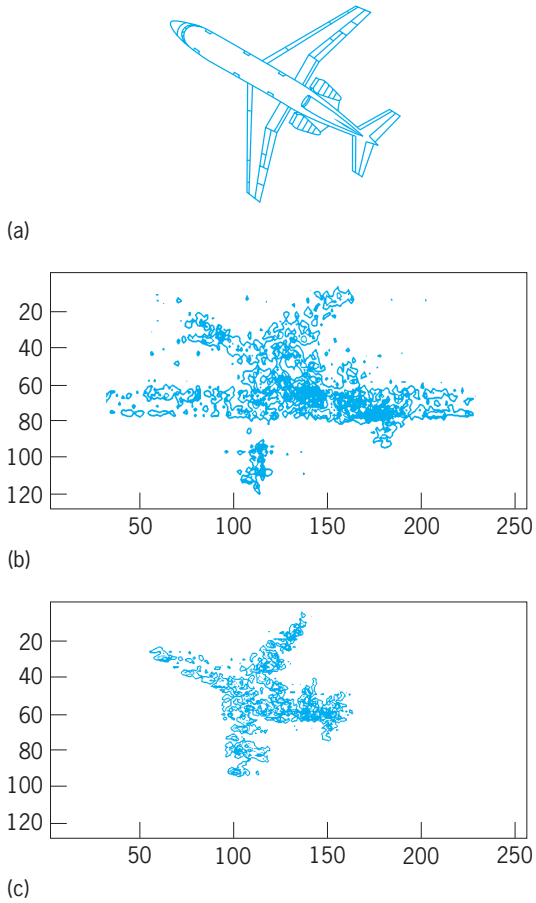

Fig. 4. Example of inverse synthetic aperture radar (ISAR). (*a*) Outline of YAK-42 aircraft. (*b*, *c*) ISAR images of the aircraft for two different focusing methods. (*Xidian University of China*)

the scatterers involved are all ideal point scatterers. Research continues toward techniques to accommodate better the return from structural elements actually found in real targets.

Another problem is that moving surface targets generally produce Doppler-time profiles different from those of stationary objects, causing their images to be misplaced and poorly focused unless compensation is made for the particular radial velocity involved. Such refinements are the subject of research.

Also being developed are techniques to access images efficiently, using various preview strategies. Having the ability to map continent-sized areas with submeter resolution introduces the problem of selecting what is of interest to image in a specific application.

SAR and ISAR radars must have good coherent processing capability, with high signal quality and phase coherence being maintained over seconds of time, much longer than that required of ordinary radar. The further inclusion of polarimetric, interferometric, and multiband capability adds to the development challenges in this fascinating radar application. *See* RADAR; REMOTE SENSING. Robert T. Hill

Bibliography. J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar: Systems & Signal Processing*, John Wiley, 1991; S. A. Hovanessian, *Introduction to Synthetic Array and Imaging Radars*, Artech House, 1980; M. I. Skolnik (ed.), *Radar Handbook*, 2d ed., McGraw-Hill, 1990; D. R. Wehner, *High Resolution Radar*, 2d ed., Artech House, 1995.

Synthetic fuel

A gaseous, liquid, or solid fuel, also known as synfuel, that does not occur naturally. Synthetic fuel can be made from tar sand, coal, or oil shale. Included in the category are various fuel gases, such as substitute natural gas and synthesis gas. *See* COAL; OIL SAND; OIL SHALE.

Synthetic crude oil (syncrude) is a complex mixture of hydrocarbons, somewhat similar to petroleum, obtained from the bituminous material that is found in tar sand formations, from coal (liquefaction), from synthesis gas (a mixture of carbon monoxide and hydrogen), or from oil shale. Synthetic crude oil generally differs in composition from petroleum; for example, syncrude from coal usually contains more aromatic hydrocarbons than petroleum. Accordingly, the hydrogen-to-carbon atomic ratios vary widely (**Fig. 1**). Gaseous fuels can be produced from sources other than petroleum and natural gas. In addition, synthetic crude oil varies in composition and character, depending upon the source (**Table 1**). *See* PETROLEUM.

Synthetic crude oil from tar sand. Tar sand is sand saturated with a highly viscous crude hydrocarbon

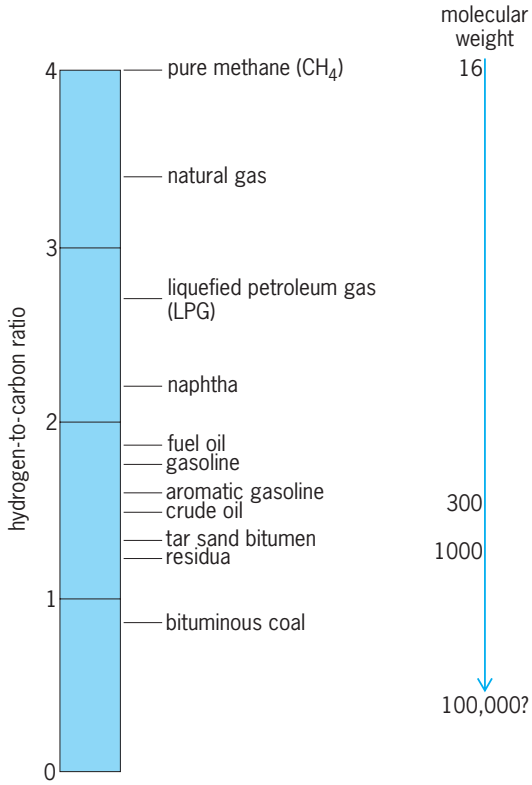

Fig. 1. Simplified comparison of the atomic hydrogen/carbon ratios and molecular weights of various fuels and fuel sources.

TABLE 1. Comparison of synthetic crude oil from various sources

Source	Tar sand bitumen	Coal	Oil shale
Gravity; °API	32	12–20	15–25
Viscosity,			
centistoke @ 100°F	6	15–25	<10
Composition, wt %			
Carbon	86	88	84
Hydrogen	13	10	11
Nitrogen	<0.1	0.3	2.0
Sulfur	<0.2	<0.1	2.0
Nickel, ppm	0	0	2
Vanadium, ppm	0	0	1
Carbon residue, wt %	0	5	5

that is not recoverable in its natural state by ordinary production methods, including enhanced oil (tertiary) recovery methods. Technically, tar sand should be called bituminous sand since the hydrocarbon is bitumen (soluble in carbon disulfide). High-grade tar sands contain about 18% by weight of bitumen that is almost equivalent in consistency to an atmospheric petroleum residuum and can be recovered by using technologies such as solvent extraction, thermal retorting, in situ combustion, and steam injection methods. In petroleum refining and in coal processing, the term "tar" is reserved to mean the cracked residue of a thermal cracking process. The term "oil sand" is used in allusion to the synthetic crude oil that can be manufactured from the bitumen. *See* BITUMEN; PETROLEUM.

A major source of synthetic crude oil is the tar sand deposits that occur in the Athabasca region of northeastern Alberta (Canada). The tar sand reserves in the United States are too deep for surface mining and require in situ treatment as the means of bitumen recovery. Some of the surface deposits are worked to produce asphalt for highway application and other minor uses. *See* ASPHALT AND ASPHALTITE.

Tar sand is a mixture of sand, water, and bitumen, and the sand component is predominantly quartz in the form of rounded or subangular particles, each of which is wet with a film of water. Surrounding the wetted sand grains and somewhat filling the void among them is a film of bitumen. The balance of the void volume is filled with connate water and sometimes a small volume of gas. *See* OIL FIELD WATERS.

In terms of bitumen separation and recovery, the hot-water process is, to date, the only successful commercial process used in North America. The process exploits the linear and the nonlinear variation of bitumen density and water density, respectively, with temperature so that the bitumen that is heavier than water at room temperature becomes lighter than water at 80°C (180°F).

In the hot-water extraction process, the tar sand is introduced into a conditioning drum in which it is heated and mixed with water to form a pulp. The pulp enters the separation cell through a central feed well and distributor. The bulk of the sand settles in the cell and is removed from the bottom, but the majority of the bitumen floats to the surface and is removed as froth. The froth from the hot-water process may be mixed with a hydrocarbon diluent, such as coker naphtha, and centrifuged.

The quality of tar sand bitumen is poor compared to that of conventional crude oil and heavy oil (**Table 2**) and has a lower proportion of distillable constituents (**Table 3**). The high carbon residue of bitumen dictates that considerable amounts of coke will be produced during thermal refining. Upgrading and refining bitumen requires a different approach to that used for upgrading heavy oil. In addition, the distance that the bitumen must be shipped to the refinery and in what form as well as product quality must be taken into account when designing a bitumen refinery. *See* COKE.

To upgrade tar and bitumen to synthetic crude oil, coking is the process of choice, and bitumen is

TABLE 2. Comparison of synthetic crude oil from bitumen with conventional petroleum and with heavy oil

Property	Athabasca bitumen (Canada)	Cold Lake heavy oil (Canada)	Lloydminster heavy oil (Canada)	Conventional crude oil (Canada)
Gravity, °API	8	12	14	35
Viscosity				
Centipoise @ 100°F (38°C)	500,000	2,000	500	10
Pour point, °F	50		5	0
Elemental analysis, wt %				
Carbon	83	84	83	86
Hydrogen	10.6	11	12	13.5
Sulfur	4.8	4.4	3.6	0.1
Nitrogen	0.4	0.4	0.4	0.2
Oxygen	1	0.2	1	0.2
Fractional composition, wt %				
Asphaltenes	19	12	12	5
Resins	32	28	17	10
Aromatics	30	35	24	25
Saturates	19	25	47	60
Metals, ppm				
Vanadium	250	190	100	10
Nickel	100	70	40	5
Carbon residue, wt %	14	11	10	5
Heating value, Btu/lb	17,500	18,000	18,200	19,500

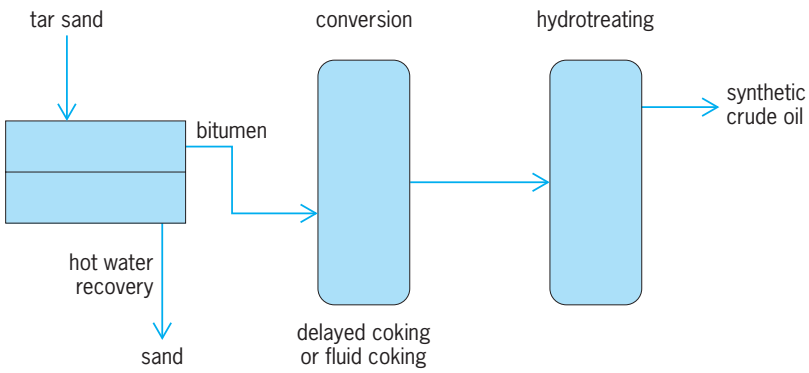

Fig. 2. Production of synthetic crude oil from tar and bitumen.

currently converted commercially by delayed coking and by fluid coking (**Fig. 2**). In each case, the charge is converted to distillate oils, coke, and light gases. The coke fraction and product gases can be used for plant fuel. The coker distillate is a partially upgraded material in itself and is a suitable feed for hydrodesulfurization to produce a low-sulfur synthetic crude oil. *See* COKING (PETROLEUM).

Delayed coking is a semibatch process in which feed bitumen is heated before being fed to coking drums that provide sufficient residence time for the cracking reactions to occur. Bitumen conversion to liquids is on the order of 75% by volume, with fluid coking giving a generally higher yield of liquids compared to delayed coking. The remainder appears as coke (approximately 15% by weight) and gases.

Fluid coking is a continuous process employing two vessels with fluid coke. It provides a better yield of overhead (distillate) products than delayed coking. Feed oil flows to the reactor vessel, where cracking and formation of coke occur; coke is combusted in the burner. Fluid transfer lines between these vessels provide the coke circulation necessary for heat balance. The proportion of coke burned is just sufficient to satisfy heat losses and provide the heat for the cracking reactions. *See* CRACKING.

In the fluid coking process, whole bitumen (or topped bitumen) is preheated and sprayed into the reactor, where it is thermally cracked in the fluidized coke bed at temperatures of 510–540°C (950–1000°F) to produce light products and coke. The coke is deposited on the fluidized coke particles, while the light products pass overhead to a scrubbing section in which any high-boiling products are condensed and recombined with the reactor fresh feed. The uncondensed scrubber overhead passes into a fractionator in which liquid products of suitable boiling ranges for downstream hydrotreating are withdrawn. Cracked reactor gases contain butanes and lower-molecular-weight hydrocarbon gases that pass overhead to a gas recovery section. The propane material ultimately flows to the refinery gas system, and the condensed butane and butenes may (subject to vapor pressure limitations) be combined with the synthetic crude. The heat necessary to vaporize the feed and to provide the heat of reaction is supplied by hot coke that is circulated back to the reactor from the coke heater. Excess coke that has formed from the fresh feed and deposited on hot circulating coke in the fluidized reactor bed is withdrawn (after steam stripping) from the bottom of the reactor. *See* FLUIDIZED-BED COMBUSTION.

Sulfur is distributed throughout the boiling range of the delayed coker distillate, as with distillates from direct coking. Nitrogen is more heavily concentrated in the higher-boiling fractions but is present in most of the distillate fractions. Raw coker naphtha contains significant quantities of olefins and diolefins that must be saturated by downstream hydrotreating. The gas oil has a high aromatic content typical of coker gas oils. *See* NITROGEN; SULFUR.

The primary liquid product from the coking process is then hydrotreated (secondary conversion or refining) to remove sulfur and nitrogen (as

TABLE 3. Distillation data for bitumen and conventional crude oil

Cut point		Cumulative wt % distilled		
°C	°F	Athabasca bitumen (Canada)	P. R. Spring bitumen (U.S.)	Leduc conventional crude oil (Canada)
200	390	3	1	35
225	435	5	2	40
250	480	7	3	45
275	525	9	4	51
300	570	14	5	
325	615	26	7	
350	660	18	8	
375	705	22	10	
400	750	26	13	
425	795	29	16	
450	840	33	20	
475	885	37	23	
500	930	40	25	
525	975	43	29	
538	1000	45	35	
538+	1000+	55	65	

hydrogen sulfide and ammonia, respectively) and to hydrogenate the unsaturated sites exposed by the conversion process. It may be necessary to use separate hydrotreaters for light distillates and medium-to-heavy fractions; for example, the heavier fractions require higher hydrogen partial pressures and higher operating temperatures to achieve the desired degree of sulfur and nitrogen removal. Commercial applications have therefore been based on the separate treatment of two or three distillate fractions at the appropriate severity to achieve the required product quality and process efficiency. *See* HYDROCRACKING; PETROLEUM PROCESSING AND REFINING.

Hydrotreating is generally carried out in downflow reactors containing a fixed bed of cobalt-molybdate catalysts. The reactor effluents are stripped of the produced hydrogen sulfide and ammonia. Any light ends are sent to the fuel gas system, and the liquid products are recombined to form synthetic crude oil. *See* CATALYSIS.

Finishing and stabilization (hydrodesulfurization and saturation) of the liquid products is achieved by hydrotreating two or three separate liquid streams. This is necessary because of the variation in conditions and catalysts required for treatment of a naphtha fraction relative to the conditions required for treatment of gas oil. It is more efficient to treat the liquid product streams separately and then to blend the finished liquids to a synthetic crude oil. In order to take advantage of optimum operating conditions for various distillate fractions, the coker distillate is treated as separate fractions, such as naphtha, kerosene, and gas oil or naphtha and mixed gas oils. The hydrotreated fractions are combined to form synthetic crude oil that is then shipped by pipeline to a refinery. The upgraded or synthetic crude oil has properties that are quite different from the original feedstock (**Table 4**) and are closer to the properties of a conventional high-API-gravity crude oil (Tables 2 and 3). *See* PETROLEUM PRODUCTS.

TABLE 4. Properties of synthetic crude oil from Athabasca bitumen

Property	Bitumen	Synthetic crude oil
Gravity, °API	8.0	32.0
Sulfur, wt %	4.8	0.2
Nitrogen, wt %	0.4	0.1
Viscosity, centipoise @ 100°F	500,000	10
Distillation profile, wt %		
0°C, 30°F	0	5
30°C, 85°F	0	30
220°C, 430°F	3	60
345°C, 650°F	17	90
550°C, 1020°F	45	100

Synthetic fuel from coal. There are several routes (**Fig. 3**) by which synthetic fuels can be prepared from coal. Gasification can yield clean gases for combustion or synthesis gas that has a controlled ratio of hydrogen to carbon monoxide. Catalytic conversion of synthesis gas to liquids (indirect liquefaction) can be carried out in fixed- and fluidized-bed reactors and in dilute-phase systems. The processes also require the addition of hydrogen to upgrade coal into finished products to accompany the molecular weight reduction that occurs during liquids production. *See* COAL.

Coal liquefaction. The chemical objectives of coal liquefaction are (1) to reduce the effect of weak bonds and thus separate fairly large units of coal structure into smaller units, (2) to bring about the

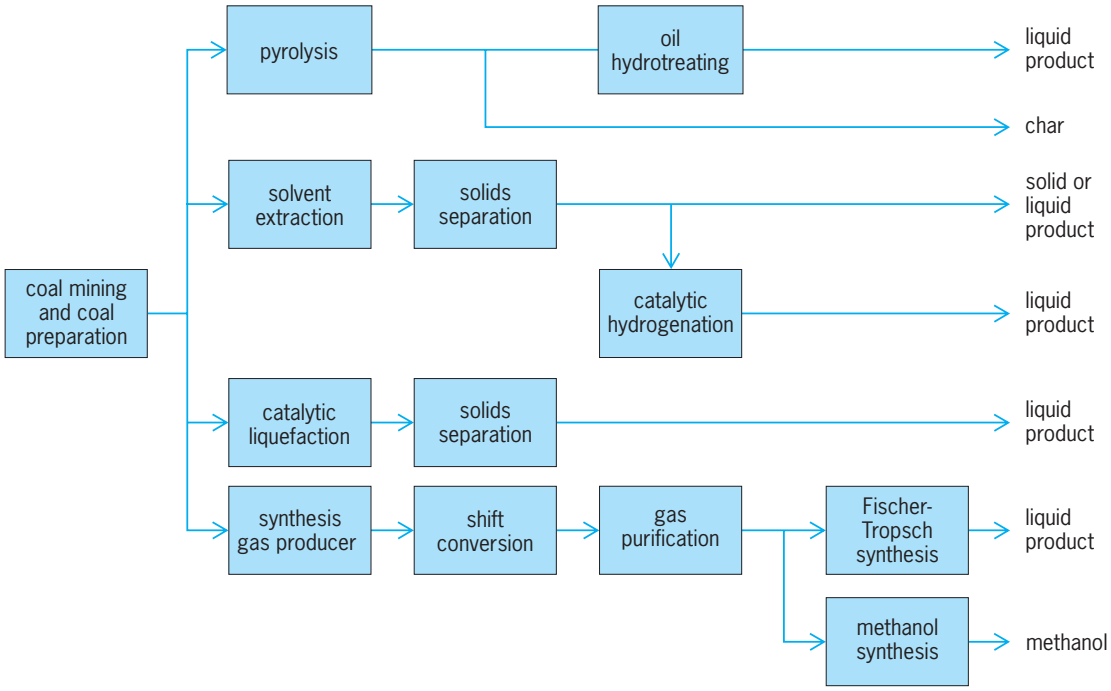

Fig. 3. Simplified schematic of the production of synthetic crude oil and syngas from coal.

decomposition of chemical bonds within the coal to form smaller fragments, and (3) to increase the hydrogen-to-carbon atomic ratio to produce a low-sulfur, ash-free liquid product that is comparable to crude oil, gasoline, or even heavy oil and bitumen.

Coal liquefaction is accomplished by four methods: (1) direct catalytic hydrogenation, (2) solvent extraction, (3) pyrolysis, and (4) indirect catalytic hydrogenation (of carbon monoxide). For example, hydrogenation of coal can be achieved by mixing coal and the catalyst with a coal-derived recycle oil (an oil that is recycled at some point in the process); and the slurry is pumped into a high-pressure system where hydrogen is present and operating conditions are on the order of 400–480°C (752–896°F) and 1000–8000 psi (6.9–55 MPa). Catalysts containing iron, molybdenum, cobalt, nickel, and tungsten are effective for such a process. *See* COAL LIQUEFACTION.

The concept of catalytic liquefaction uses a catalyst to add hydrogen to the coal. These processes usually require a liquid medium with the catalyst dispersed throughout, or may even employ a fixed-bed reactor. On the other hand, the catalyst may also be dispersed with the coal, whereupon the combined coal-catalyst system can be injected into the reactor. *See* HYDROGENATION.

In the catalytic liquefaction processes, coal is brought into contact with a catalyst in the presence of hydrogen. Many processes of this type have the advantage of eliminating the need for a hydrogen donor solvent (and the subsequent rehydrogenation of the spent solvent), but there is still the need for an adequate supply of hydrogen. The nature of the process also virtually guarantees that the catalyst will be deactivated by the mineral matter in the coal as well as by coke lay-down during the process. *See* HYDROGEN.

In order to achieve the direct hydrogenation of the coal, the catalyst and the coal must be in intimate contact, but if this is not the case, process inefficiency is the general rule. There has been the tendency of late to achieve coal-catalyst contact by the use of a hydrogen donor solvent. Alternatively, a catalyst with a sufficiently high vapor pressure may be used so that catalyst deposition on the coal surface is achieved under process conditions.

The major features of catalytic liquefaction processes are (l) rapid heating to temperatures of the order of 450–600°C (840–1110°F), (2) short residence times, and (3) reactor effluent quenching. Instead of using a slurry-feed system, the coal is entrained in a rapidly moving stream of hydrogen and has a residence time of (usually) less than 20 s in the reactor at 500°C (930°F) and 2000 psi (13.8 MPa).

In solvent extraction processes, coal is mixed with a solvent that is capable of effecting the transfer of hydrogen from the solvent to the coal or even from gaseous hydrogen to the coal) at temperatures up to 500°C (930°F) and pressures up to 5000 psi (34.5 MPa). Processing by solvent extraction solubilizes and disperses coal in a solvent that transfers some of its hydrogen to the coal. Ash and insoluble coal are separated from the liquid product to recover recycle oil and product oil. The carbonaceous residue is reacted with steam in a gasifier to produce the hydrogen needed. Hydrogenation of the solvent is also practiced to improve the product quality. *See* SOLVENT EXTRACTION.

Solvent extraction processes generally use liquids (often derived from the feed coal) as donor solvents that are capable of donating hydrogen to the system under the conditions of the reaction. The overall result is an increase (relative to pyrolysis processes) in the amount of coal that is converted to lower-molecular-weight, that is, soluble, products. Reaction temperatures usually have an upper limit of 510°C (950°F); hydrogen may be supplied (under pressure) during the process, or it may be introduced by hydrogenation of the solvent prior to the process. Such hydrogen may be produced from any unreacted coal, from feed coal, or from by-product gases.

High-temperature solvent extraction processes of coal have been developed in three process configurations: (1) extraction in the absence of hydrogen but using a recycle solvent that has been hydrogenated in a separate process stage; (2) extraction in the presence of hydrogen with a recycle solvent that has not been previously hydrogenated; and (3) extraction in the presence of hydrogen using a hydrogenated recycle solvent.

The solvent extraction concept can also be achieved under milder conditions, but the product may be high-nitrogen solid or heavy oil with only low yields of light oils and gases. Severe conditions are more effective for sulfur and nitrogen removal to produce a "lighter" liquid product that is more amenable to downstream processing.

A novel approach to the solvent extraction process uses bitumen or heavy oil as the process solvent. In fact, the coprocessing of coal with a variety of petroleum-based feedstocks (such as heavy oils) has received attention. The concept is a variation of the solvent extraction process for coal liquefaction. Whether the coprocessing option is a means of producing more liquids or whether the coal should act as a scavenger for the metals and nitrogen species in the petroleum material is dependent upon the process conditions. What is certain is that special effort should be made to ensure the compatibility of feedstocks and products. Incompatibility can, at any stage of the liquefaction operation, lead to expensive shutdowns as well as to (in respect of the heavier feedstocks) the onset of coke formation.

Pyrolysis processes involve heating coal to temperatures in excess of 400°C (750°F), which results in the conversion of the coal to gases, liquids, and char (carbon). The char is hydrogen-deficient, thereby enabling inter- or intramolecular hydrogen transfer processes to be operative, resulting in relatively hydrogen-rich gases and liquids. Unfortunately, the char produced often amounts to more than 45% by weight of the feed coal and, therefore, such processes have often been considered to be an uneconomical or inefficient use of coal. *See* PYROLYSIS.

Pyrolysis (or carbonization) is perhaps the oldest technique for obtaining liquid products directly from

coal and involves heating the coal in the absence of air (or oxygen) to produce heavy and light oils, gases, and char. In the presence of hydrogen, the process is called hydrocarbonization. The composition and relative amounts of the products formed are dependent on the process parameters such as heating rate, pressure, coal type, coal (and product) residence time, coal particle size, and reactor configuration. A major disadvantage of this type of process is the large yields of char, reducing the yield of liquid products.

On the other hand, pyrolysis and hydrocarbonization processes are less complex than liquid-phase hydrogenation processes. The operating pressures for pyrolysis processes are usually less than 100 psi (690 kPa; more often between 5 and 25 psi or 34–172 kPa), but hydrocarbonization processes require hydrogen pressures of the order of 300–1000 psi (2.1–6.9 MPa). In both processes, the operating temperature can be as high as 600°C (1110°F).

Three types of pyrolysis reactors are of interest: mechanically agitated entrained-flow reactors and fluidized-bed reactors. The agitated reactor may be quite complex, but the entrained-flow reactor has the advantage of either down-flow or up-flow operation and can provide short residence times. In addition, the coal can be heated rapidly, leading to higher yields of liquid (and gaseous) products that may well exceed the volatile matter content of the coal as determined by the appropriate test. The short residence time also allows a high throughput of coal and the potential for small reactors.

Coal gasification. The other category of producing synthetic fuel (gas and liquid) from coal invokes the concept of the coal gasification followed by conversion of the gaseous products to liquid fuel. This is often referred to as indirect liquefaction of coal insofar as the coal is not converted directly into liquid products. *See* COAL GASIFICATION.

The indirect liquefaction of coal involves a two-stage conversion operation in which coal is first converted (by reaction with steam and oxygen) to a gaseous mixture that is composed primarily of carbon monoxide and hydrogen (syngas, or synthesis gas). The gas stream is subsequently purified (to remove sulfur, nitrogen, and any particulate matter), after which it is catalytically converted to a mixture of liquid hydrocarbon products. In addition to the production of fuel gas, coal may be gasified to prepare synthesis gas (carbon monoxide and hydrogen), which may be used to produce a variety of products, including ammonia, methanol, and liquid hydrocarbon fuels.

The synthesis of hydrocarbons from carbon monoxide and hydrogen (synthesis gas) is a procedure (the Fischer-Tropsch synthesis) for the indirect liquefaction of coal. This process is the only coal liquefaction scheme currently in use on a relatively large commercial scale; South Africa is currently using the Fischer-Tropsch process on a commercial scale. Notably, Germany produced roughly 156 million barrels of synthetic petroleum annually using the Fischer-Tropsch process during World War II. *See* FISCHER-TROPSCH PROCESS.

Coal is converted to gaseous products at temperatures in excess of 800°C (1470°F), and at moderate pressures, to produce synthesis gas:

$$C_{coal} + H_2O \rightarrow CO + H_2$$

The purified gas has a heating value of 325 Btu/ft^3 (12,135 kJ/m^3) if oxygen is used or about 100 Btu/ft^3 (3725 kJ/m^3) if air is used. In some processes, light hydrocarbons are formed by cracking of higher-molecular-weight compounds, and the heating value can be higher.

There are several processes systems for coal gasification, including the commercially available Lurgi, Koppers-Totzek, or Winkler processes. Many coal gasification units are in operation throughout the world to produce synthesis gas or fuel gas. Modern gasification systems prepare synthesis gas from coal or oil for methanol, ammonia, and oxo-alcohol production. The gasification may also be attained by means of gasification of the coal in place (underground gasification or in situ gasification). The exothermic nature of the process and the decrease in the total gas volume in going from reactants to products suggest the most suitable experimental conditions to use in order to maximize product yields. The process should be favored by high pressure and relatively low reaction temperature. In practice, the Fischer-Tropsch reaction is carried out at temperatures on the order of 200–350°C (390–660°F) and at pressures of 75–4000 psi (0.5–28 MPa); the hydrogen/carbon monoxide ratio is usually approximately 2.2:1 or 2.5:1. *See* COAL MINING.

Since up to three volumes of hydrogen may be required to achieve the next stage of the liquids production, the synthesis gas must then be converted (by means of the water-gas shift reaction) to the desired level of hydrogen:

$$CO + H_2O \rightarrow CO_2 + H_2$$

Next, the gaseous mix is purified and converted to a wide variety of hydrocarbons:

$$CO + (2n + 1)H_2 \rightarrow C_nH_{2n+2} + n\,H_2O$$

This method of the catalytic hydrogenation of carbon monoxide is also a method of producing synthetic liquid fuel from coal. Catalysts can be prepared from iron, nickel, cobalt, ruthenium, and zinc either in a metal salt or contained on a support. Each catalyst gives a different product distribution that is also a function of the method of preparation and pretreatment. Primary products are normally methane and higher-molecular-weight hydrocarbons, alcohols, and organic acids. Operating conditions for the Fischer-Tropsch synthesis are usually in the range of 300–500 psi (2–3.5 MPa) and 200–400°C (392–752°F). These reactions result primarily in low- and medium-boiling aliphatic compounds; present commercial objectives are focused on the conditions that result in the production of *n*-hydrocarbons as well as olefinic and oxygenated materials.

Although the efficiency of a two-stage process as compared with a single-stage liquefaction process has often been cited as a disadvantage of the indirect liquefaction of coal, the advantages of the indirect approach are that (1) the gasification stage can tolerate a higher degree of impurities (such as mineral matter) in the coal; (2) the feed to the liquefaction stage can be maintained at the desired hydrogen level regardless of the nature and quality of the feed coal; (3) a "clean" product is produced; (4) the method is adaptable to the use of the products from the underground gasification of coal; and (5) there is the added advantage that hydrogen from a variety of other hydrogen production methods may be used, thereby increasing the overall production of liquids from a given amount of coal.

The underground gasification of coal in place is also an option for coal conversion that does not require a mining operation. It is suitable for difficult-to-mine coal and seams that are below an economical thickness. The concept was patented in 1909 in the United States. In this process, air, oxygen, or a mixture of steam and air or of steam and oxygen is introduced into the coal seam through boreholes or shafts. The burned coal produces a gaseous product that is recovered through properly spaced boreholes or shafts.

Underground gasification affords the recovery of energy from unminable coal, greater recovery of the coal seam, and negligible impact on the health and safety of miners. In some systems, miners are needed only for the preparation of shafts and galleries. There are no spoil banks, slack piles, or acid mine drainage problems. On the other hand, problems can occur with control of the combustion, roof support, bed permeability, ground-water control, leakage to adjacent strata, and monitoring of the underground events.

Synthetic fuel from oil shale. Shale oil (synthetic crude oil from oil shale) is produced by the thermal processing of oil shale at high temperatures in a variety of retorts or even in situ (that is, without being mined). The ingredient in the shale that is decomposed, kerogen, is an insoluble material containing, in addition to carbon and hydrogen, considerable amounts of nitrogen, oxygen, and sulfur. The resources of oil shale in the United States are substantial, easily dwarfing the resources of petroleum. The conversion of oil shale to shale oil has received commercial setbacks due to fluctuations in the price of petroleum that have rendered many of the technologies uneconomical. *See* KEROGEN.

Crude shale oil, sometimes called retort oil, is the liquid oil condensed from the effluent in oil shale retorting (**Fig. 4**). Crude shale oil typically contains appreciable amounts of water and solids, and tends to form sediments. As a result, it must be upgraded to a synthetic crude oil before being suitable for pipelining or substitution for petroleum crude as a refinery feedstock. However, shale oils are sufficiently different from petroleum that processing shale oil presents some unusual problems.

Shale oils, especially those from the Green River in southern Wyoming, have particularly high nitrogen contents, typically 1.7–2.2 wt % compared to approximately 0.2–0.3 wt % for conventional (Table 2) petroleum. In many other shale oils (including those from eastern United States oil shale), nitrogen contents are lower than in the Green River shale oils but still higher than those typical of petroleum. Because retorted shale oils are produced by a thermal cracking process, olefin (—CH═CH—) and diolefin [—CH═CH—$(CH_2)_n$—CH═CH—, where $n \geq 0$] contents are high. It is the presence of these olefins and diolefins, in conjunction with high nitrogen contents, which gives crude shale oils their characteristic instability toward sediment formation. The sulfur content of shale oils varies widely, but is generally lower than that of high-sulfur petroleum crude oil and tar sand bitumen.

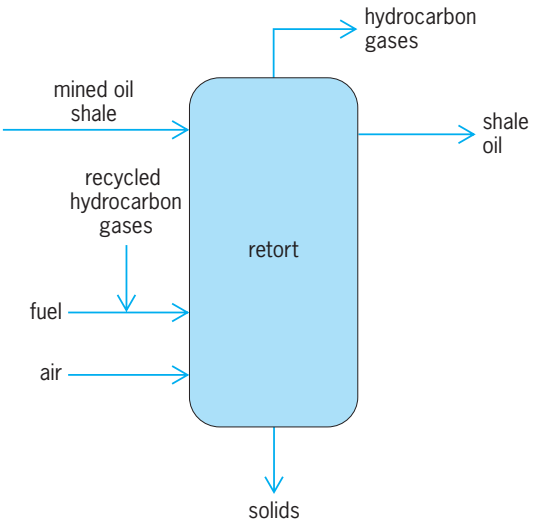

Fig. 4. Production of synthetic crude oil from oil shale.

The Green River shale oils contain appreciable amounts of aromatics, polar aromatics, and pentane-insolubles. Oxygen contents are higher than those typically found in petroleum but lower than those of crude coal liquids. Crude shale oil also contains appreciable amounts of soluble arsenic, iron, and nickel that cannot be removed by filtration. *See* ARSENIC; IRON; NICKEL.

Upgrading, or partial refining, to improve the properties of a crude shale oil may be carried out with different objectives, depending on the intended use for the product: (1) stabilization to produce a product that can be transported to a distant refinery by pipeline; (2) more complete upgrading to produce a premium refinery feedstock with low nitrogen, low sulfur, and essentially no residuum; (3) upgrading to produce chemical feedstock streams; and (4) complete refining of the crude shale oil or selected fractions to produce finished end products (such as gasoline, diesel, jet fuel).

The arsenic and iron in shale oil would poison and foul the supported catalysts used in hydrotreating. Because these materials are soluble, they cannot be removed by filtration. Several methods have been used specifically to remove arsenic and iron. Other

methods involve hydrotreating; these also lower sulfur, olefin, and diolefin contents and thereby make the upgraded product less prone to form gum. Shale oil is also rich in high-molecular-weight, waxy paraffinic material. Thermal cracking lowers molecular weight but yields straight-chain products of low octane number. Fluid catalytic cracking not only lowers molecular weight but also causes isomerization to produce branched products with higher octane numbers. As a result, the "cat cracked" shale naphtha is a more desirable feedstock for hydrotreating to make gasoline blend stock than is the naphtha from thermal cracking or coking of shale oil. *See* HYDROCRACKING; ISOMERIZATION. James G. Speight

Bibliography. J. G. Speight (ed.), *The Chemistry and Technology of Petroleum*, 3d ed., Marcel Dekker, 1999; J. G. Speight (ed.), *The Desulfurization of Heavy Oils and Residua*, 2d ed., Marcel Dekker, 2000; J. G. Speight (ed.), *Fuel Science and Technology Handbook*, Marcel Dekker, 1990; J. G. Speight, *Gas Processing: Environmental Aspects and Methods*, Butterworth-Heinemann, Oxford, England, 1993.

Syphilis

A sexually transmitted infection of humans caused by *Treponema pallidum* ssp. *pallidum*, a corkscrew-shaped motile bacterium (spirochete). Due to its narrow width, *T. pallidum* cannot be seen by light microscopy but can be observed with staining procedures (silver stain or immunofluorescence) and with dark-field, phase-contrast, or electron microscopy. *Treponema pallidum* cannot be continuously cultured in vitro by standard bacteriological methods. For antigen production and experimental studies, the organism is propagated by inoculation of laboratory rabbits. Limited multiplication of *T. pallidum* has been obtained in a tissue culture system at a temperature of 91–95°F (33–35°C) in an atmosphere containing 1–5% oxygen and in media containing reducing agents. The organism is very sensitive to environmental conditions and to physical and chemical agents. The complete genome sequence of the *T. pallidum* Nichols strain (1.13×10^6 base pairs) has been determined. Analysis of the nucleotide sequence of the small, circular treponemal chromosome indicates that *T. pallidum* lacks the genetic information for many of the metabolic activities found in other bacteria. Thus, this spirochete is dependent upon the host for most of its nutritional requirements. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); ELECTRON MICROSCOPE.

Syphilis is usually transmitted through direct sexual contact with active lesions and can also be transmitted by contact with infected blood and tissues. If untreated, syphilis progresses through various stages (primary, secondary, latent, and tertiary). Infection begins as an ulcer (chancre) and may eventually involve the cardiovascular and central nervous systems, bones, and joints. Congenital syphilis results from maternal transmission of *T. pallidum* across the placenta to the fetus. *See* SEXUALLY TRANSMITTED DISEASES.

Epidemiology. *Treponema pallidum* is an obligate parasite of humans and does not have a reservoir in animals or the environment. Syphilis has a worldwide distribution. Its incidence varies widely according to geographical location, socioeconomic status, and age group. Although syphilis is controlled in most developed countries, it remains a public health problem in many developing countries. Epidemics of syphilis recently occurred in the former Soviet Union due to conditions that have limited identification and treatment of infected individuals. Approximately 6000 cases of primary and secondary syphilis were reported in the United States in 2000. This represents a dramatic decrease from the post–World War II peak of 51,060 cases of primary and secondary syphilis reported in 1990. The number of reported cases of congenital syphilis also declined from 4438 cases in 1991 to approximately 300 cases in 2000. Various studies have shown that syphilis is a risk factor for infection with the human immunodeficiency virus (HIV) since syphilitic lesions may act as portals of entry for the virus. There is little natural immunity to syphilis infection or reinfection. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS).

Clinical manifestations. Syphilis has been called "the great imitator" because its diverse clinical manifestations can mimic many infectious and noninfectious diseases. *Treponema pallidum* enters the body through microscopic abrasions of skin or mucosal membranes, colonizes host cell surfaces, and begins to multiply extracellularly at the site of infection. Within a few hours, significant numbers of treponemes depart the local site and are carried to the regional lymph nodes. The organisms spread throughout the body via the circulation. The incubation period of syphilis ranges from 10 to 90 days. Clinical signs and symptoms of the primary stage include a painless chancre at the site of infection (genital, oral, or anal) and enlargement of regional lymph nodes. The chancre and lymph nodes contain treponemes, and the disease is communicable. Although the untreated chancre heals in a few weeks, presumably due to the host immune response, infectious treponemes are still present in the individual.

The secondary or disseminated stage of syphilis occurs within a few weeks of the healing of the primary lesion(s). It is characterized by low-grade fever, headache, generalized enlargement of lymph nodes, malaise, loss of appetite (anorexia), and a hyperpigmented skin rash that appears on the palms and soles and eventually spreads to other areas. Superficial sores (mucous patches) may occur in the mouth, vagina, or anus, and wartlike lesions (condylomata) may form in moist areas of the body. Severe hepatitis and meningitis are occasionally observed. Although secondary lesions disappear within a few weeks, the untreated person remains infectious.

The latent stage of syphilis represents conversion from acute to chronic disease. Relapses of secondary

syphilis may occur early in latency. After the first year of infection, the individual usually becomes clinically asymptomatic. However, treponemes continue to persist in infected tissues. Although transmission of syphilis from an infected pregnant woman to her fetus can occur during the latent stage, the absence of exposed lesions reduces the probability of transmission to a sexual partner. Approximately two-thirds of persons with untreated syphilis will remain in the latent stage for the remainder of their lives.

Tertiary (late) syphilis occurs in the remaining one-third of untreated cases within months to decades after the initiation of latency. Treponemes invade the central nervous system, cardiovascular system, eye, skin, bones, or internal organs where they produce damage due to a host delayed-type hypersensitivity reaction to treponemal antigens. Approximately 80% of syphilis mortalities are caused by cardiovascular involvement, while 20% are from neurologic involvement. Destructive, soft, tumorlike lesions (gummas) may develop in skin, bone, and other tissues. During the tertiary stage, sexual transmission of syphilis does not occur and congenital transmission is rare.

Congenital syphilis results from transplacental transmission of *T. pallidum* from an infected pregnant woman to her fetus. The extent of damage to the fetus depends on the disease stage at the time of transmission. Overwhelming infection of the fetus may result in miscarriage or stillbirth. Infants born with congenital syphilis may have skin lesions, mucoid nasal discharge (snuffles), long-bone malformation, anemia, or involvement of the central nervous system. Clinical manifestations of congenital syphilis may be delayed until puberty or adulthood. Signs of late congenital syphilis may include blindness, deafness, tooth malformation, bone involvement, soft lesions, neurosyphilis, and, rarely, cardiovascular disease. Late congenital syphilis is not infectious.

Diagnosis. The diagnosis of syphilis is aided by direct microscopic examination of specimens obtained from genital and skin lesions and the detection of antibodies that develop during infection. Microscopic examination of specimens from oral and anal lesions is not reliable due to the presence of spirochetal species that are members of the normal flora. Nontreponemal screening tests such as the Rapid Plasma Reagin (RPR) or the Venereal Disease Research Laboratory (VDRL) detect serum antibodies formed by the patient in response to cardiolipin, a phospholipid. These tests can yield false-negative reactions in early syphilis due to their limited sensitivity, and false-positive reactions can occur with patients who are pregnant or have autoimmune diseases. Treponemal tests such as the fluorescent treponemal antibody absorbed (FTA-ABS) test and the *Treponema pallidum* particle agglutination test are used to confirm reactive screening test (TPPA) results.

Recombinant-based *T. pallidum* proteins have been recently used in Europe as antigens in enzyme-linked immunosorbent assays (ELISA) and Western blot assays for the serodiagnosis of syphilis. These newer assays have advantages over certain serologic tests for syphilis, such as the FTA-ABS test and the TPPA test, since they do not require the use of native *T. pallidum* antigens and are compatible with automation. *See* ANTIBODY; ANTIGEN; FLUORESCENCE MICROSCOPE; SEROLOGY.

Treatment. Parenteral penicillin G is the preferred antibiotic for treatment of all stages of syphilis. Alternative antibiotics for syphilis treatment include erythromycin and tetracycline. However, tretracycline is contraindicated for pregnant women and young children. Although penicillin resistance has not emerged in *T. pallidum*, erythromycin resistance has been documented in a clinical isolate, and erythromycin treatment failures have been reported. *See* ANTIBIOTIC.

Prevention and control. The prevention and control of syphilis are dependent upon case reporting to health authorities, contact tracing, adequate treatment of cases and contacts, education of high-risk populations, use of condoms, prenatal screening of pregnant women, and screening of blood donors. There is currently no vaccine to prevent syphilis. However, it is anticipated that information obtained from the *T. pallidum* genome sequence will lead to further improvements in diagnostic tests for syphilis and to the eventual development of a vaccine that would prevent infection. Certain proteins encoded by the multigene *T. pallidum tpr* family are under investigation as potential vaccine candidates. *See* PUBLIC HEALTH.

Lola V. Stamm

Bibliography. C. M. Fraser et al., Complete genome sequence of *Treponema pallidum*, the syphilis spirochete, *Science*, 281:375–387, 1998; K. K. Holmes et al. (eds.), *Sexually Transmitted Diseases*, 3d ed., McGraw-Hill, New York, 1999; P. R. Murray et al. (eds.), *Manual of Clinical Microbiology*, 7th ed., ASM Press, Washington, DC, 1999; *1998 Guidelines for Treatment of Sexually Transmitted Diseases*, MMWR 47 (RR-1):28–48; L. V. Stamm and H. L. Bergen, A point mutation associated with bacterial macrolide resistance is present in both 23S rRNA genes of an erythromycin-resistant, *Treponema pallidum* clinical isolate, *Antimicrob. Agents Chemother.*, 44:806–807, 2000.

Systellommatophora

An order (or superorder) in the gastropod molluscan subclass Pulmonata containing three families of sluglike mollusks that lack any trace of a shell, have separate external male and female orifices, lack a mantle cavity, have a posterior anus and excretory pore, and bear eyes on the tops of two contractile, but not retractile, tentacles. The Rathouisiidae include about 30 species of carnivorous land slugs and range from South China and India to Queensland, Australia, in very wet tropical forests. They have slender, often brightly colored bodies. The Veronicellidae comprise about 200 species of slugs that feed on decaying plant matter. Their distribution is now

pantropical (including Florida and south Texas), owing to accidental horticultural introductions. They are minor garden and crop pests in several subtropical and tropical areas. They have long flattened bodies with usually dull colors and are very abundant nocturnally in garden and lawn areas. The Onchididae include perhaps 250 species of intertidal to high-subtidal dwellers, reaching greatest abundance in the rocky shore areas of the tropics. They range from a few inches to nearly 4 in. (100 mm) in length. Many species have dorsal eyespots or branchial plumes present, and can be confused with opisthobranch snails. *See* PULMONATA. G. Alan Solem

System design evaluation

A comprehensive and life-cycle assessment of the effectiveness and cost of competing system designs, in order to choose the best candidate. System design evaluation is essential within the systems engineering process. It should be embedded appropriately within the process and then pursued continuously as system design and development progress. This article focuses on (1) the placement of design evaluation within the systems engineering process, (2) the evaluation factors that underpin system design evaluation, and (3) the system selection decision.

Evaluation within systems engineering. The systems engineering process is supported by a morphology for linking technology to customer needs (**Fig. 1**). It is composed of three major activities: synthesis, analysis, and evaluation.

Synthesis. System synthesis is accomplished by combining top-down and bottom-up design activities. The focus should be on customer requirements, normally expressed in functional terms (Fig. 1, block 0). Key elements are the design team (block 1) interacting with traditional and computer-based tools for design synthesis (block 2). Existing components, parts, and subsystems (block 5) are integrated to synthesize candidate systems. The product is a candidate system design that should be evaluated against the need and compared with other candidates.

Analysis. Analysis of the candidate system design is a necessary but not sufficient ingredient in system design evaluation. It involves estimating and predicting design dependent parameter (DDP) values (Fig. 1, block 3), forecasting design independent parameter (DIP) values (block 5), and maintaining physical and economic databases. System analysis and operations research is a step on the way to system design evaluation, but adaptation of the models and techniques is required.

Evaluation. Evaluation of each candidate system design (Fig. 1, block 4) is accomplished after receiving design dependent parameter values for the candidate from block 3. It is the specific values for design dependent parameters that differentiate (or instance) candidate systems. Design independent parameter values from block 5 are externalities: They apply across all candidate systems. Each candidate is optimized in block 4 before being subjected to the design decision schema (block 6). It is here that the best candidate is sought (based on the customer's subjective evaluation). *See* SYSTEMS ENGINEERING.

Evaluation factors. To be comprehensive, system design evaluation must encompass both system effectiveness and life-cycle cost. To be rigorous, evaluation should be pursued with the aid of models and

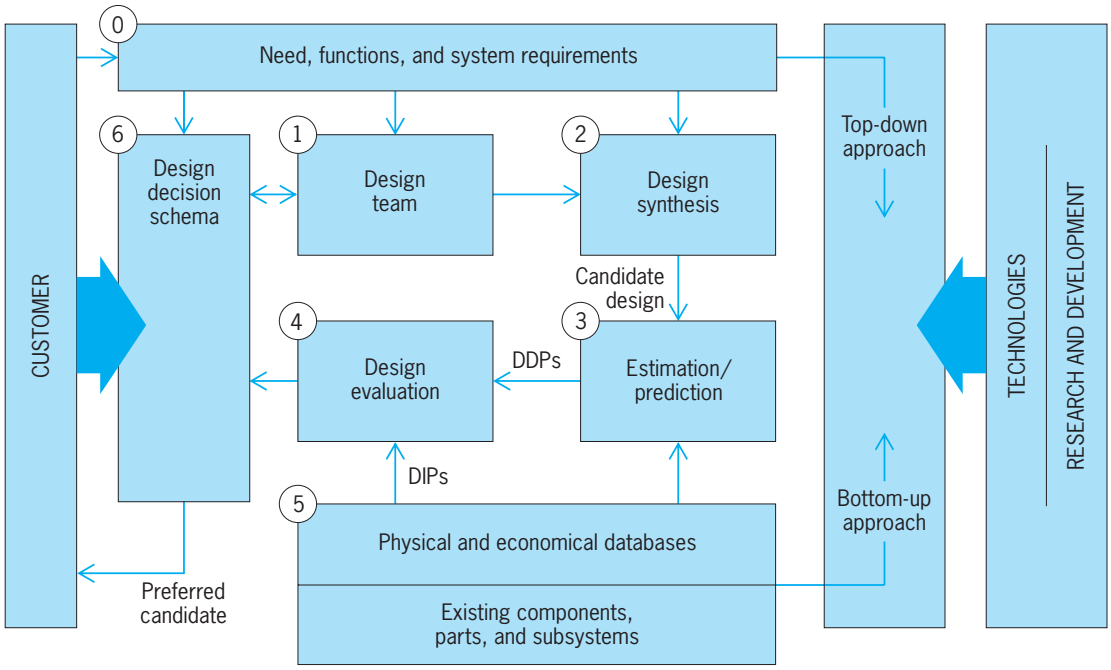

Fig. 1. Systems engineering schematic.

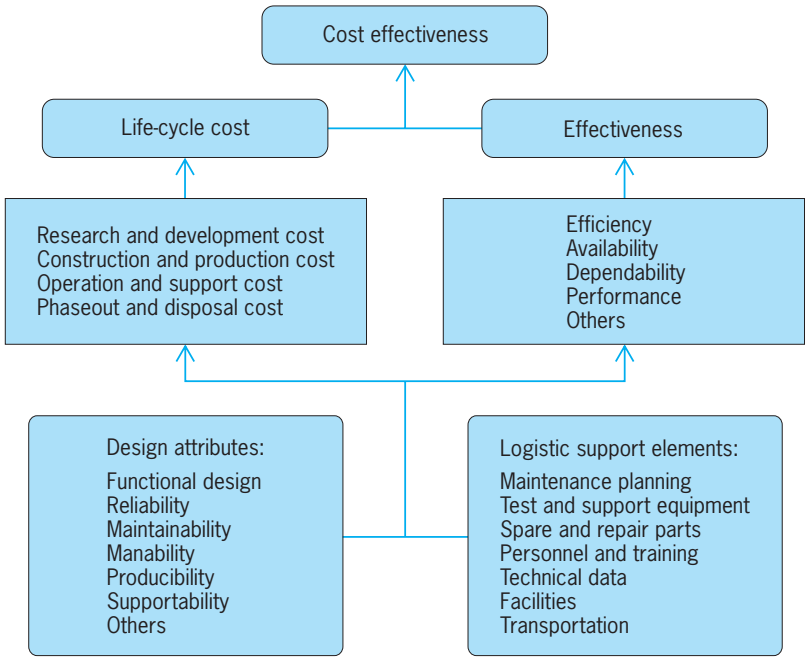

Fig. 2. Factors to be evaluated.

simulation. Even when expert opinion must be used in place of formal analytical approaches, a modeling structure may offer guidance. *See* MODEL THEORY; SIMULATION.

Cost and effectiveness. Life-cycle cost arises from the four life-cycle phases: research and development; construction and production; operation and support; phaseout and disposal (**Fig. 2**). Effectiveness is a multidimensional measure based on mission fulfillment in terms of a stated need. Mission fulfillment may be expressed by one or more criteria (effectiveness measures), depending on the type of system and the objectives it seeks to satisfy. Some common effectiveness measures are shown on the right side of Fig. 2.

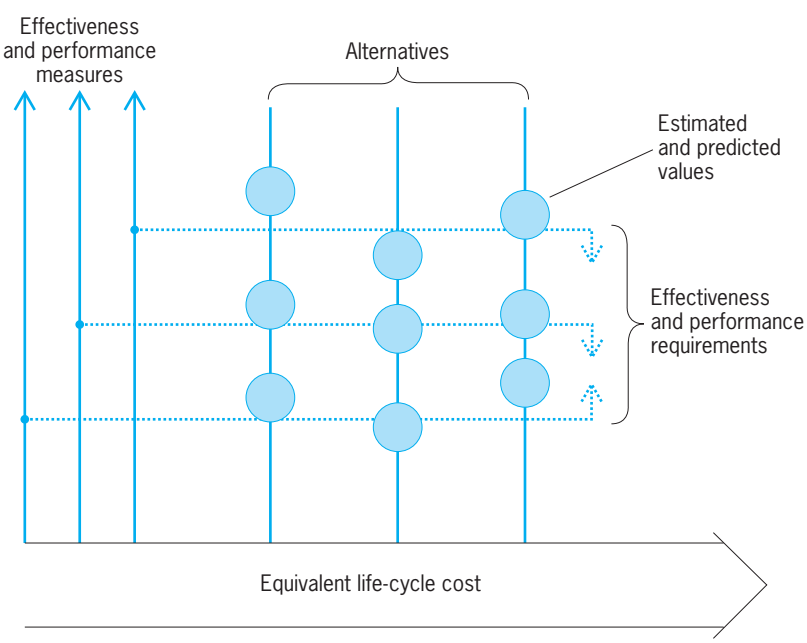

Fig. 3. Design evaluaton display.

Design dependent parameters are design characteristics inherent in the product or system (for example, reliability, maintainability, and disposability). They define the design space and are subject to manipulation by the design team during the process of seeking the best design. The objective is to consider and balance all applicable cost and effectiveness factors.

Evaluation mathematics. The evaluation process incorporates a paradigm based on normative models from the domains of operations research and systems analysis. This paradigm involves the identification and incorporation of design dependent parameters (Fig. 1, block 3) to mathematically link design characteristics with operational outcomes. Optimization and tradeoff decisions are facilitated during the systems engineering process with the aid of a design evaluation function (block 4) and a design evaluation display (block 6).

The design dependent parameter approach utilizes the design evaluation function, expressed as

$$E = f(X; Y_d, Y_i)$$

where E = a life-cycle complete evaluation measure (usually equivalent life-cycle cost incorporating the time value of money and inflation factors); X = design variables (such as number of deployed units, armor thickness, retirement age, number of repair channels, or pier spacing); Y_d = design dependent parameters (for example, design life, reliability, maintainability, weight, capacity, and disposability); Y_i = design independent parameters (for example, cost of money, labor rates, material cost per unit, energy cost per unit, and shortage cost penalty).

The design evaluation function, with its design dependent and design independent parameters, facilitates design optimization. It provides the mathematical basis for clarifying the true difference between alternatives (a design-based choice among a set of design dependent parameters, Y_d) and optimizing a given system (a search-based choice among a set of design variables, X).

System selection decision. The factors used to evaluate candidate system design must be aggregated to make them apparent. Accordingly, an evaluation process must be followed which combines evaluation factors and decision making. The candidate system design that complies best with the customer's requirement is the preferred choice and may be recommended for implementation.

System design evaluation must recognize and incorporate multiple criteria if it is to be viable as part of the systems engineering process. Multiple criteria considerations arise when both economic and noneconomic elements are present in the evaluation. In these common situations, design evaluation is facilitated by the use of a design evaluation display (DED), exhibiting both cost and effectiveness measures (**Fig. 3**). Life-cycle cost and one or

more effectiveness measures may be displayed simultaneously as an aid in decision making. Effectiveness requirements, or thresholds, are shown in fig. 3. These are useful to the decision maker in subjectively assessing the degree to which each candidate system satisfies effectiveness criteria. Life-cycle cost is an objective measure. The goal is to select the alternative with the lowest life-cycle cost that satisfies the effectiveness measures to an acceptable degree.

Formal trade-off between life-cycle cost and effectiveness measures would require the application of preference functions and utility theory. This is sometimes done, but in practice a simple weighting scheme is often used. In many applications it may be sufficient to have the decision maker or team decide by visually inspecting a design evaluation display.

The commitment to technology, system configuration, performance, and life-cycle cost is particularly important during the early stages of system design. A large information gap exists between this commitment and the system-specific knowledge available during conceptual and preliminary design. Evaluation approaches utilizing modeling and indirect experimentation may be used to help narrow this gap to reduce development risk. *See* DECISION SUPPORT SYSTEM; OPERATIONS RESEARCH; OPTIMIZATION; SYSTEMS ANALYSIS. Wolter J. Fabrycky

Bibliography. B. S. Blanchard and W. J. Fabrycky, *Systems Engineering and Analysis*, 3d ed., 1998; C. W. Churchman, R. L. Ackoff, and E. L. Arnoff, *Introduction to Operations Research*, 1957; W. J. Fabrycky and B. S. Blanchard, *Life-Cycle Cost and Economic Analysis*, 1991; G. A. Hazelrigg, *Systems Engineering*, 1993; A. P. Sage and W.B. Rouse, *Handbook of Systems Engineering and Management*, 1999.

System families

Large systems that are formed from a variety of component systems: custom systems, which are newly engineered from the "ground up;" existing commercial-off-the-shelf (COTS) systems that are subsequently tailored for a particular application; and existing or legacy systems. Such related terms as systems of systems (SOS), federations of systems (FOS), federated systems of systems (F-SOS), and coalitions of systems (COS) are often used to characterize these systems. These appellations capture important realities brought about by the fact that modern systems are not monolithic. Rather, they have five characteristics, initially summarized by Mark Maier, that make one of the system family designations appropriate:

1. *Operational independence of the individual systems*. A system of systems is composed of systems that are independent and useful in their own right. If a system of systems is disassembled, the constituent systems are capable of independently performing useful operations by themselves and independently of one another.

2. *Managerial independence of the systems.* The component systems generally operate independently in order to achieve the technological, human, and organizational purposes of the individual unit that operates the system. These component systems are generally individually acquired, serve an independently useful purpose, and often maintain a continuing operational existence that is quite independent of the larger system of systems.

3. *Geographic distribution*. Geographic dispersion of the constituent systems in a system of systems is often quite large. Often, these constituent systems can readily exchange only information and knowledge with one another, and not substantial quantities of physical mass or energy.

4. *Emergent behavior*. The system of systems performs functions and carries out purposes that do not reside uniquely in any of the constituent systems. These behaviors arise as a consequence of the formation of the entire system of systems and are not the behavior of any constituent system. The principal purposes supporting engineering of these individual systems and the composite system of systems are fulfilled by these emergent behaviors.

5. *Evolutionary and adaptive development*. A system of systems is never fully formed or complete. Development of these systems is evolutionary and adaptive over time, and structures, functions, and purposes are added, removed, and modified as experience of the community with the individual systems and the composite system grows and evolves.

Unfortunately, there is no universally accepted definition of these system families, and at present there are no definitive criteria that distinguish a system of systems from other systems. In a formal sense, almost anything could be regarded as a system of systems. A personal computer is a system. However, the monitor, microprocessor, disk drive, and random-access memory are also systems. Thus, should a personal or a mainframe computer be called a system of systems? Is the Internet a system of systems? Instinctively, they are quite different. The big difference in these two examples is that the former system may be massive in size but monolithic in purpose, while the latter is capable of supporting the myriad communications and commerce purposes of utilizing organizations and humans. There is also a distinction based on operational and managerial independence of the system components and in evolution and emergence possibilities. *See* INTERNET; MICROCOMPUTER.

Federations and coalitions of systems. Often, appropriate missions exist for relatively large systems of systems in which there is a very limited amount of centralized command-and-control authority. Instead, a coalition of partners has decentralized power and authority and potentially differing perspectives of situations. It is useful to term such a system a "federation of systems" and sometimes a "coalition of systems." The participation of the federation or

coalition of partners is based upon collaboration and coordination to meet the needs of the federation or coalition.

Innovation and change. Support for innovation and change of all types is a desirable characteristic of these system families. Innovation includes both technological innovation and organizational and human conceptual innovation. Accomplishing this requires continuous learning, a reasonable tolerance for errors, and experimental processes to accomplish both the needed learning and the needed change. The systems fielded in order to obtain these capabilities will not be monolithic structures in terms of either operations or acquisition. Rather, they will be systems of systems, coalitions of systems, or federations of systems that are integrated in accordance with appropriate architectural constructs in order to achieve the evolutionary, adaptive, and emergent cooperative effects that will be required to achieve human and organizational purposes and to take advantage of rapid changes in technology. They can potentially accommodate system life-cycle change, in which the life cycle associated with use of a system family evolves over time; system purpose change, in which the focus in use of the system emerges and evolves over time; and environment change, in terms of alterations in the external context supporting differing organizational and human information and knowledge needs, as well as in the technological products that comprise constituent systems.

Autonomy, heterogeneity, and dispersion. One can contrast system families and identify relationships between conventional systems, systems of systems, and federations or coalitions of systems with regard to three characteristics, as initially noted by A. J. Krygiel: autonomy, heterogeneity, and dispersion. A federation of systems, federated system of systems, or coalition of systems will generally have greater values of these characteristics than a (nonfederated) system of systems.

Many conventional systems are built for special purposes, as a mixture of commercial-off-the-shelf systems and custom developments of hardware and software. These constituents are generally provided by multiple contractors who are used to supporting a specific customer base and working under the leadership of a single vertical program management structure. For best operation, these systems should be managed as a system of systems, federation of systems, or coalition of systems.

A system of systems generally has achieved integration of the constituent systems across communities of contractors, and sometimes across multiple customer bases, and is generally managed by more horizontally organized program management structures, such as integrated product and process development (IPPD) teams. When the IPPD team effort is well coordinated, the team is generally well able to deal with conflict issues that arise due to business, political, and other potentially competing interests.

Federations of systems face the same dilemmas identified for systems of systems but are generally much more heterogeneous along transcultural and transnational socio-political dimensions, are often managed in an autonomous manner without great central authority and direction such that they satisfy the objectives and purpose of an individual unit in the federation, and often accommodate a much greater geographic dispersion of organizational units and systems. Thus, the delimiters between systems of systems and federations of systems or coalitions of systems, while generally subjective, are nonetheless principled.

The notions of autonomy, heterogeneity, and dispersion are not independent of one another. Increasing geographic dispersion will usually lead to greater autonomy and consequently will also increase heterogeneity. The Internet is perhaps the best example of a system that began under the aegis of a single sponsor, the U.S. Department of Defense, and has grown to become a federation of systems.

Federalism. This approach, initially studied by C. Handy as "new federalism," addresses the necessary considerations for the structuring of loosely coupled organizations in order to help them adapt to changes in the information and knowledge age. The application of federalist political principles to the management of systems of systems and federations of systems is an appealing way of obtaining a systems engineering ecology—in other words, a sustainable systems engineering approach that possesses adaptation, evolution, and emergence characteristics analogous to those in a natural ecology. This is so because many contemporary organizational alliances, including those for engineering and development, take the form of virtual organizations or virtual teams rather than the classical physical organizations and teams. The concept of federalism is particularly appropriate since it offers a well-recognized way to deal with the systems engineering management paradoxes of power and control such that the desired systems ecological balance is obtained. Generally, this is accomplished by making things big by keeping them small, a goal which, in turn, is accomplished by expanding the domain of an enterprise by instantiating multiple quasiautonomous units as opposed to acquiring mass by aggregation around a centralized command-and-control authority base; encouraging autonomy but within appropriate bounds set by process and architecture standards; and combining variety with shared purpose, and individuality with partnerships at national and global levels.

Federalism is based on five principles: subsidiarity, interdependence, a uniform and standardized way of doing business, separation of powers, and dual citizenship. Subsidiarity, the most important principle, suggests that power belongs to the lowest possible point within the engineering team of a federation of systems. Interdependence, or pluralism, requires that the autonomous development units or teams of a development federation of a federation of systems stick together because they need one another as much as they need management leadership and leadership authority. Having a uniform and

standardized way of doing business ensures interdependence within federated engineering organizations of a system of systems or federation of systems through agreement on basic rules of conduct, common traditions of communicating, and common units of measurement of progress and quality. Separation of powers requires that management, monitoring, and governing aspects of engineering programs and projects of federations of systems be viewed as separate functions to be accomplished by separate bodies whose membership may overlap. Finally, dual citizenship requires that every individual is a "citizen" in two communities: the local development group, professional group, or union, and the overall program of the federation of systems.

Systems engineering approaches. These systems-of-systems and federations-of-systems concepts have numerous implications for systems engineering and management.

Grand design approach. Contemporary organizations often treat the engineering of systems of systems or federations of systems with systems engineering protocols that are, at best, suitable only for monolithic systems. The archetype of such ill-advised protocols is the "grand design" life cycle, which is based on the waterfall model that came into prominence around 1970. A large number of problems have been encountered with grand design efforts to engineer a system. Today, the classic waterfall approach is suggested only in those rare cases where user and system-level requirements are crystal clear and unlikely to change at all during or after engineering the system, and where funding for the grand design is essentially guaranteed. This is rarely the case for major systems, especially those that are software-intensive, and would be the rarest of all cases for a system of systems or federation of systems. Changing user and organizational needs and changing technologies virtually guarantee that major systems cannot be developed using the grand design approach.

Incremental and evolutionary approaches. Two leading alternatives to the grand design approach for the engineering of systems were initially termed incremental and evolutionary, although the term "evolutionary" is now generally used to characterize both of these. In incremental development, the system is delivered in preplanned phases or increments, in which each delivered module is functionally useful. The overall system capability improves with the addition of successive modules. The desired system capability is planned to change from the beginning, as the result of "build N" being augmented and enhanced through the phased increment of "build $N+1$." This approach enables a well-functioning implementation to be delivered and fielded within a relatively short time and augmented through additional builds. It also allows time for system users to thoroughly implement and evaluate an initial system with limited functionality compared to the ultimately desired system. Generally, the notion of preplanning of future builds is strong in incremental development. As experience with the system at build N is gained, requirements changes for module $N+1$ may be more easily incorporated into this, and subsequent, builds.

Evolutionary life-cycle development is similar in approach to its incremental complement; however, future changes are not necessarily preplanned. This approach recognizes that it is impossible to initially predict and set forth engineering plans for the exact nature of these changes. The system is engineered at build $N+1$ through reengineering the system that existed at build N. Thus, a new functional system is delivered at each build, rather than obtaining build $N+1$ from build N by adding a new module. The enhancements to be made to obtain a future system are not determined in advance, as in the case of incremental builds. Evolutionary development approaches can be very effective in cases where user requirements are expected to shift dramatically over time, and where emerging and innovative technologies allow for major future improvements. They are especially useful for the engineering of unprecedented systems that involve substantial risk and allow potentially enhanced risk management. Evolutionary development may help program managers adjust to changing requirements and funding priority shifts over time, since new functionality introductions can be advanced or delayed in order to accommodate user requirements and funding changes. Open, flexible, and adaptable system architecture is central to the notion of evolutionary and emergent development. These are major elements in the contemporary U.S. Department of Defense Initiatives in evolutionary acquisition. *See* DISTRIBUTED SYSTEMS (CONTROL SYSTEMS); LARGE SYSTEMS CONTROL THEORY; REENGINEERING; RISK ASSESSMENT AND MANAGEMENT; SYSTEMS ARCHITECTURE; SYSTEMS ENGINEERING; SYSTEMS INTEGRATION.

Andrew P. Sage

Bibliography. P. G. Carlock and R. E. Fenton, Systems of systems (SoS) enterprise systems engineering for information-intensive organizations, *Sys. Eng.*, 4(4):242–261, 2001; P. Chen and J. Clothier, Advancing systems engineering for systems-of-systems challenges, *Sys. Eng.*, 6(3):170–183, 2003; C. Handy, Balancing corporate power: A new Federalist Paper, *Harvard Bus. Rev.*, 70(6):59–72, November-December 1992; C. Handy, Trust and the virtual organization, *Harvard Bus. Rev.*, 73(3):8–15, May/June 1995; A. J. Krygiel, *Behind the Wizard's Curtain: An Integration Environment for a System of Systems*, CCRP Publication Series, Vienna, VA, 1999; M. W. Maier, Architecting principles for systems-of-systems, *Sys. Eng.*, 1(4):267–284, 1998; J. Morganwalp and A. P. Sage, A system of systems focused enterprise architecture framework and an associated architecture development process, *Inform. Knowl. Sys. Manag.*, 3(4):87–105, 2003; A. P. Sage and C. D. Cuppan, On the systems engineering and management of systems of systems and federations of systems, *Inform. Knowl. Sys. Manag.*, 2(4):325–345, 2001; A. P. Sage and W. B. Rouse (eds.), *Handbook of Systems Engineering and Management*, Wiley, New York, 1999.

Systematics

The comparative analysis of living and fossil species, including their discovery, description, evolutionary relationships to other species, and patterns of geographic distribution. Because systematics involves crucial activities of describing and understanding the history of life, it is the primary science of biological diversity.

Systematics can be divided into four major fields. Taxonomy, often equated with systematics, is the discipline concerned with the discovery, description, and classification of organism groups, termed taxa (singular, taxon). Classification is the clustering of species into a hierarchical arrangement according to some criterion, usually an understanding of their relationships to other species. Phylogenetic analysis, an increasingly important aspect of systematics, is the discovery of the historical, evolutionary relationships among species; this pattern of relationships is termed a phylogeny. The fourth component of systematics is biogeography, the study of species' geographic distributions. Historical biogeography examines how species' distributions have changed over time in relationship to the history of landforms, ocean basins, and climate, as well as how those changes have contributed to the evolution of biotas (groups of species living together in communities and ecosystems).

Taxonomy. The habitats and ecosystems of the natural world are extremely complex in terms of their ecological and taxonomic diversity. From those environments, systematists have discovered and described approximately 1.4 million species, of which about 950,000 are insects, 250,000 are plants, and 45,000 are vertebrates. Estimates of the number of species left to be discovered range from 3 million to nearly 100 million. A major reason for this uncertainty involves the difficulty of sampling over the entire surface of the Earth many different kinds of organisms, especially very small organisms. Documenting biodiversity is among the most challenging scientific problems. For example, an enormous effort is needed to describe and classify these species, one that far outstrips available taxonomic expertise. Moreover, the lack of a global electronic information system that can store information about those species which have already been described as well as those that will be discovered in the future also hinders progress.

A fundamental problem underlying all systematic inquiry, and a major concern for taxonomic practice, is the nature of species. Debates about how species should be defined and how they might be recognized are among the most long-standing and intense in all of biology. Many authorities have proposed solutions to the species problem and offered a variety of definitions, but basically two primary criteria have been used to define species. By far the most commonly applied criterion is diagnosability: one or more populations that are recognizably distinct from other such populations and that cannot be subdivided into other diagnosable units constitute a species. The notion of diagnosability means that once comparisons have been made to establish patterns of variability within and among different populations, the systematist finds that some populations can be characterized as distinct because the organisms share one or more characters not found in other populations. These characters may be morphological, behavioral, physiological, or biochemical. This view of species has been the one most commonly used by systematists, especially those studying insects and other small organisms, which make up the majority of the world's species.

A second view of species, adopted frequently within vertebrate systematics, is the biological species concept: species are those populations or groups of populations that are reproductively isolated from other such populations. Thus, if an ancestral population is subdivided by some extrinsic ecological or geological barrier into two geographically isolated populations, given sufficient time one or both will likely differentiate, perhaps to the extent that if the barriers were to disappear and the populations came back into contact they would be incapable of interbreeding or of producing fertile offspring. This scenario is the classic allopatric speciation model, which is often closely linked to the biological species concept. *See* SPECIES CONCEPT.

In principle, application of the biological species concept requires that the putative species geographically overlap (that is, be sympatric with) its close relatives so that reproductive isolation can be assessed. In nature, however, many populations remain geographically separated (allopatric) from one another, and in these situations followers of the biological species concept attempt to infer the presence of reproductive isolation by evaluating the degree of morphological, behavioral, or ecological differences among the populations. The more differentiated the populations, it is thought, the greater the likelihood that they will not be capable of interbreeding upon becoming sympatric.

It is important to understand that species designations are hypotheses regardless of which definition might be used. In contrast to individual organisms, a species cannot be directly observed, and therefore new information may lead a systematist to reevaluate species limits.

For over 200 years naturalists have sought to document Earth's species diversity by mounting expeditions. These inventories have helped build the major natural history collections that contain worldwide some 2 billion specimens. Once specimens are collected and brought back to the museum or herbarium for study, they are cataloged and identified as to species. If they cannot be assigned to any known species, they are described as new and a type specimen is designated. The type serves as the reference specimen for the new name, and future workers will compare the type with newly collected specimens to see whether any of those are new. This comparison leads to revisionary monographs, which summarize all the species of a group, their characters, distributions, and relationships. Monographs

constitute a major vehicle for studies of phylogenetic relationships and for a summary classification. *See* TAXONOMY.

Phylogeny. Organisms have been evolving and diversifying on Earth for several billion years, and a major task of systematics is the generation of hypotheses about the historical pattern of this evolution, that is, about phylogeny. Determining phylogenetic relationships is essentially a discovery process in which comparisons are made among species, similarities and differences are assessed, and a hierarchical pattern of relationships is hypothesized based on these patterns of similarity.

For many years there were debates over how similarities were to be interpreted during the process of phylogenetic inference. One group of systematists advocated an approach called phenetics, in which overall similarity was quantified and relationships were then based on those measures. Another group proposed a method termed cladistics, in which observed similarities are partitioned into primitive or derived similarity. Only derived similarities—that is, characters modified from more primitive conditions—could then be used as evidence of relationships. For example, vertebrates commonly called reptiles were placed in their own formal classificatory group, Reptilia. Once comparisons among reptiles and other vertebrate groups were made, however, it was discovered that the characters of the Reptilia (such as scales, absence of hair or feathers) were primitive and that other, derived characters linked some reptiles with mammals and some with birds. Thus, theropod dinosaurs are known to be more closely related to birds than to many other so-called reptiles.

Phylogenetic analysis is of major importance to biology because an understanding of relationships provides the basis for making powerful predictions about the characters of organisms that have yet to be studied. Phylogenetic relationships, because they establish a historical hierarchy, constitute the most logical framework for synthesizing all that is known about species and groups of species. The picture of animal phylogeny is very incomplete, but the use of new characters that are revealed by technological advances such as electron microscopy and deoxyribonucleic acid (DNA) sequence contribute to the solution of many phylogenetic puzzles. *See* PHYLOGENY.

Classification. Biological classification is the grouping together of species on the basis of one or more criteria and then naming those groups. Present-day classification schemes are founded on the work of the eighteenth-century Swedish botanist, Carolus Linnaeus. He proposed a hierarchical series of taxonomic categories, such as kingdom, phylum, class, order, family, genus, and species. Thus, the tiger has the generic and specific name *Panthera tigris*, is a member of the family Felidae, which is in the order Carnivora, and carnivores and other mammals are in the class Mammalia.

Debates over which criteria are most appropriate for establishing hierarchical relationships have been closely tied to debates over methods of assessing phylogenetic relationships. Again, some have advocated constructing the hierarchy on the relative degree of overall similarity (phenetics), but the importance of basing classifications on the pattern of phylogenetic relationships is clearly recognized.

The Linnaean scheme of subordinate categorical ranking is intrinsically hierarchical, and so too is the historical pattern of life's evolution. To the extent that a classification scheme mirrors the hierarchical pattern of the best-corroborated hypotheses about group relationships, the groups of the classification can be said to be natural. Constructing classifications that reflect the genealogical relationships of taxa is critical for making the findings of systematics maximally useful within basic and applied biology. Natural classifications are highly predictive. For example, a newly discovered species of cat belonging to the genus *Felis* would be predicted to have characters in common with other species of the genus, some that are common with the other species in the family Felidae, and so on to include the higher taxonomic levels of Carnivora and Mammalia. *See* CLASSIFICATION, BIOLOGICAL.

Biogeography. A major component of systematic research involves documenting the distributions of species—both contemporaneous and fossil—and understanding why groups of species have such distributions and how the distributions have changed over time. The distributions of most species are very imperfectly known, mainly because they have been determined from a small number of specimens collected from a few sites. Even in well-known, highly visible groups such as birds, distributional limits are often uncertain. This uncertainty underscores the importance of an extensive inventory effort around the globe. Information about distributions is crucial for many aspects of basic and applied biology, from understanding how and why groups diversify over space and time, to identifying the geographic sources of agricultural pests, diseases, or other introduced exotics.

Distribution patterns of species can be explained by the interaction of a number of processes that take place over different spatiotemporal scales. Long-distance dispersal involves the idiosyncratic movement of relatively few individuals across a geographic or ecologic barrier; population dispersion involves the movement of many individuals of a population into new areas that are ecologically compatible; and vicariance is a process whereby a widespread population is divided into smaller components by the origin of a geographic or ecological barrier. *See* POPULATION DISPERSAL; POPULATION DISPERSION.

Prior to the application of phylogenetic methods in systematics, long-distance dispersal was often considered to be the primary mechanism explaining species distributions. This view grew out of a belief that continents, islands, and ocean basins were stable over time. With the realization that continents had shifted their positions, systematists began invoking dispersion of many groups from one region to another over land connections in order to

explain similarities in biotas. It also became clear to workers attempting to explain disjunct distributions among closely related groups that these geographic patterns could be explained by vicariance of a once widespread ancestral population. This line of thinking has led to the discipline of vicariance biogeography, which entails a search for congruent historical patterns in distributions across different groups of organisms and then seeks to explain that congruence by postulating a common pattern of vicariance.

Evidence suggests that the geographic history of species and biotas, whether terrestrial or marine, is intimately tied to the history of the areas in which they reside. When changes in climate or Earth history reduce or eliminate geographic or ecological barriers, many species of a biota expand their ranges together, and thus the biota becomes more cosmopolitan. The subsequent origin of barriers partitions this widespread biota into smaller, isolated areas within which populations differentiate to create areas of endemism. *See* BIOGEOGRAPHY.

Importance of systematics. Humans use tens of thousands of species in their daily lives for food, medicine, clothing, shelter, and many other aspects of commerce. At the same time, more than 40,000 mi^2 (100,000 km^2) of habitat are being lost each year. In order to conserve biodiversity and use what remains in a sustainable manner, resource managers must have access to reliable scientific information about species. Included in this information is knowledge about species systematics.

Systematists provide answers to four fundamental questions that are required by resource managers and users of biodiversity: What species is it? What are its characteristics? What are its relationships to other species? Where is it distributed? Correct identifications may be life-and-death issues or may have major economic implications: for example, distinguishing between pathogenic and nonpathogenic organisms, or between an indigenous, benign species and an introduced pest. Likewise, knowing the characteristics of organisms is fundamental for developing new genetic resources or designing a biocontrol program for an exotic agricultural pest. Predictive classifications, made possible by an understanding of relationships, help direct the search for new genetic resources to improve agricultural productivity, or lead to the discovery of new pharmacologically active compounds in close relatives of medicinally useful plants.

Finally, systematic data and interpretations underlie progress in all of biology. An understanding of relationships, in particular, is fundamental for interpreting comparative data across different kinds of organisms, whether those data be morphological, physiological, or biochemical. Joel Cracraft

Bibliography. D. R. Brooks and D. A. McLennan, *Phylogeny, Ecology, and Behavior: A Research Program in Comparative Biology*, 1991; N. Eldredge and J. Cracraft, *Phylogenetic Patterns and the Evolutionary Process*, 1980; P. L. Forey et al., *Cladistics: A Practical Course in Systematics*, 1992; G. J. Nelson and N. I. Platnick, *Systematics and Biogeography: Cladistics and Vicariance*, 1981; E. O. Wiley, *Phylogenetics: The Theory and Practice of Phylogenetic Systematics*, 1981.

Systems analysis

The application of mathematical methods to the study of complex human physical systems. A system is an arrangement or collection of objects that operate together for a common purpose. The objects may include machines (mechanical, electronic, or robotic), humans (individuals, organizations, or societal groups), and physical and biological entities. Everything excluded from a system is considered to be part of the system's environment. A system functions within its environment. Examples of systems include the solar system, a regional ecosystem, a nation's highway system, a corporation's production system, an area's hospital system, and a missile's guidance system. A system is analyzed so as to better understand the relationships and interactions between the objects that compose it and, where possible, to develop and test strategies for managing the system and for improving its outcomes.

Systems of physical processes, such as the weather system, an oil refinery plant, or an energy generation system, are often studied under the methodological headings of systems theory, systems science, or systems engineering. The procedures that constitute these methodologies have contributed greatly to the ability to organize and control very large and complex physical systems, with space exploration via crewless satellites being a prime example. These successes have caused investigators to attempt to apply systems engineering concepts to a wide range of systems that are characterized by having a very strong human component. In particular, a major thrust in systems research is directed toward the study of problems that arise in public policy (at all levels of government), and business and industrial decision making. The term "systems analysis" is reserved for the study of systems that include the human element and behavioral relationships between the system's human element and its physical and mechanical components, if any. Examples of public policy systems are the federal government's welfare system, a state's criminal justice system, a county's educational system, a city's public safety system, and an area's waste management system. Examples of industrial systems are a manufacturer's production distribution system and an oil company's exploration, production, refining, and marketing system. Examples with physical environmental components are the atmospheric system (in which the concerns are with air pollution and the greenhouse effect) and a water supply system (in which water quality and quantity are the attributes of concern). The direct transfer of systems engineering concepts to the study of a system in which the human element must be considered is restricted by limitations in the ability to comprehend and quantify human interactions. (Operations research, a related field of study, is directed toward the analysis of

components of such systems. Public policy analysis is the term used for a system study of a governmental problem area.) *See* DECISION THEORY; OPERATIONS RESEARCH; SYSTEMS ENGINEERING.

Process. Systems comprise interrelated objects, with the objects having a number of measurable attributes. A mathematical model of a system attempts to quantify the attributes and to relate the objects mathematically. The resultant model can then be used to study how the real-world system would behave as initial conditions, attribute values, and relationships are varied systematically. Thus, the mathematical model serves as a controlled means of studying the system in a descriptive sense. Most importantly, if the purpose of the systems analysis is to aid in the design of a new system or the modification of an ongoing system, the mathematical model serves as an experimental tool that can prescribe a particular alternative solution. Here it is assumed that some criterion—an objective (for example, cost) or other measure of effectiveness (for example, police response time)—can be stated that enables the analysis to differentiate between possible alternative designs. The process is often confounded by the fact that there are usually many criteria that impact the final decision. To a limited extent, multicriteria methods are available to address such situations. *See* MODEL THEORY.

The systems analysis process is an iterative one that cycles repeatedly through the following interrelated and somewhat indistinct phases: (1) problem statement, in which the system is defined in terms of its environment, goals, objectives, constraints, criteria, actors (decision makers, participants in the system, impacted constituency), and other objects and their attributes; (2) alternative designs, in which solutions are identified; (3) mathematical formulation, in which a mathematical description of the system is developed, tested, and validated; (4) evaluation of alternatives, in which the mathematical model is used to evaluate and rank the possible alternative designs by means of the criteria; and (5) selection and implementation of the most preferred solution. The process includes feedback loops in which the outcomes of each phase are reconsidered based on the analyses and outcomes of the other phases. For example, during the implementation phase, constraints may be uncovered that hinder the solution's implementation and thus cause the mathematical model to be reformulated. The analysis process continues until there is evidence that the mathematical structure is suitable; that is, it has enough validity to yield answers that are of value to the system designers or the decision maker.

Mathematical methods and examples. The mathematical foundations of systems analysis derive from engineering and its use of differential equations as applied in optimal control theory; from computer science and its theory of automata; and from applied mathematics (especially operations research), which includes optimization, statistical analysis, simulation, and the calculus of variations. The analytical procedures used in systems analysis encompass the full range of scientific and computational methodology. Systems analyses have been applied in such diverse scientific fields as biology, engineering, transportation, ecology, and energy, and in social areas such as welfare, housing, and health. *See* AUTOMATA THEORY; CALCULUS OF VARIATIONS; DIFFERENTIAL EQUATION; OPTIMIZATION; SIMULATION; STATISTICS.

Typical systems analyses and their mathematical bases include the following: (1) Energy flow in a marine ecosystem was modeled as a set of differential equations under steady-state assumptions. (2) A regional blood bank improved its operations after a study based on stochastic analyses of its supply and demand, and creation of a mathematical optimization model of its distributional activities. (3) The Georges Bank haddock fishery system used a combination of biological fish population models and econometric market models to study the financial operation of a fishing system over time. (4) The future structure of the United States energy system was studied with the aid of an integrated econometric linear-programming-based model of the total United States energy production and distribution system. (5) The economic management of an Alaskan salmon fishery was analyzed by a computer-based simulation. (6) Proposed changes to the U.S. Department of Agriculture's food stamp program were analyzed by a model based upon microanalytic simulation in which the actions of the population in general were studied by a representative sample of households. (7) The appropriate number and location of firehouses in Wilmington, Delaware, were determined by using travel-time analyses and optimization-based location models. (8) The entire Dutch water management system was analyzed to provide a basis for a new national water management policy. *See* LINEAR PROGRAMMING; STOCHASTIC PROCESS; SYSTEMS ECOLOGY.

Status. The form and substance of systems analysis have been codified and applied with much success in a number of important areas. The technique is, on the whole, an accepted one, but there are concerns about its application and its ability to serve the needs of society. The incorrect application of systems analysis as originally developed and practiced in the Department of Defense led to some critical and costly failures. This was due mainly to the fact that the engineering basis of systems analysis was applied to social systems, in which the relationships (political and bureaucratic) between the objects (people) were not well understood and the values of the parameters were not accurately known. These difficulties still exist, but current systems analyses now take them into account. As systems analysis has evolved, its practitioners have learned to differentiate between different types of systems.

Hard and soft systems. As originally developed, systems analysis studies have been applied to those areas that are "hard" in that they are well defined and well structured in terms of objectives and feasible alternative systems (for example, blood-bank design, and integrated production and inventory processes). The aim of hard systems analysis is to select

the best feasible alternative. In contrast, soft systems are concerned with problem areas that involve ill-defined and unstructured situations, especially those that have strong political, social, and human components. These generally involve public and private organizations (for example, design of a welfare system, and structure and impact of a corporate mission statement). The objectives of soft systems and the means to accomplish them are problematical and, in fact, a systemic view of the problem area is not assumed. The aim of soft systems analysis is to find a plan of action that accommodates the different interests of its human actors.

Large-scale systems. There is also need for further study of large-scale systems, which by definition are most complex. It is important to find ways to describe mathematically the systems that represent the totality of an industrial organization, the pollution concerns of a country and a continent, or the worldwide agricultural system. These are multicriteria problems with the solutions conflicting across criteria, individuals, and countries. The possibility that such systems may be studied in a computer-based laboratory is very promising. But this challenge must be approached cautiously, with the awareness that the methods and models employed are only abstractions to be used with due consideration of the goals of the individual and society. *See* LARGE SYSTEMS CONTROL THEORY; LINEAR SYSTEM ANALYSIS. Saul I. Gass

Bibliography. P. Checkland, *Systems Thinking, Systems Practice*, 1981; P. Checkland and J. Scholes, *Soft Systems Methodology in Action*, 1981; C. W. Churchman, *The Design of Inquiring Systems*, 1971; C. W. Churchman, *The Systems Approach*, 1968; P. Edwards, *Systems Analysis and Design*, 1993; K. E. Kendall and J. E. Kendall, *Systems Analysis and Design*, 3d ed., 1994; I. R. Hoos, *Systems Analysis in Public Policy*, 1972; M. D. Mesarovic and Y. Takahara, *Abstract Systems Theory*, 1989; H. J. Miser and E. S. Quade (eds.), *Handbook of Systems Analysis*, vol. 1: *Overview of Uses*, *Procedures*, *Applications*, *and Practice*, 1995, vol. 2: *Craft Issues and Procedural Choices*, *and Practices*, 1988, vol. 3: *Cases*, 1996; J. Moder and S. E. Elmaghraby, *Handbook of Operations Research*, 1978; E. S. Quade, *Analysis for Public Decisions*, 3d ed., 1989, L. A. Zadeh and C. A. Desoer, *Linear Systems Theory*, 1963, reprint 1979.

Systems architecture

The discipline that combines system elements which, working together, create unique structural and behavioral capabilities that none could produce alone. The degree to which well-designed systems-level architectures are critical to the success of large-scale projects—or the lack thereof to failure—has been dramatically demonstrated in recent years. The explosion of technological opportunities and customer demands has driven up the size, complexity, costs, and investment risks of such projects to levels feasible for only major companies and governments. Without sound systems architectures, these projects lack the firm foundation and robust structure on which to build. This conclusion holds whether the system is a satellite surveillance system, a crewed spacecraft, a stealth aircraft, a commercial airliner, a line of personal computers, or an advanced microprocessor.

Growth of large systems. One of the earliest major system developments began in the early 1900s with the change from 100-customer telephone companies to the Bell System, a million times larger and capable of providing quality services. Air service progressed from airplanes that could barely carry a few mailbags to DC-3's, the global jet air transportation system with its vast array of local and international airlines and airports. Simple 200-lb (100-kg) satellites were replaced by multiton spacecraft exploring the entire solar system, providing global communications and navigation, and helping to keep world peace through surveillance. The combination of global transportation and global communication forced regional markets to become global in less than a decade. Software programs, supported by 100,000-element microprocessors and million-line codes, revolutionized many fields of endeavor. Crewed space programs developed into major national enterprises. Nuclear power systems became major contributors to energy production in many countries. Programs of this size and cost rapidly became matters of social and political debate, necessarily complicating system design, engineering management, production planning, and operational constraints. *See* SYSTEMS ENGINEERING.

Concept of an architecture. In the 1950s, the technical literature began reflecting the importance of the architecture of communication, computer, aerospace, and industrial systems. The word "architecture" is commonly used to describe the underlying structure of networks, command-and-control systems, spacecraft, and computer hardware and software [By implication, if an architecture is not strong enough, it and the system it supports will collapse.] *See* COMPUTER ARCHITECTURE.

It was a small step to recognize the philosophical similarities between these architectures and those begun millennia ago and carried forth by successive civilizations. The field of classical architecture provides principles, insights, and approaches of proven value for the more recent engineering disciplines. Not the least of these is the concept of an architecture. Only the context has been changed.

Complexity and its consequences. A recent addition to these historical principles is the concept of systems, first propounded for feedback control applications in the late 1940s. Systems are collections of dissimilar elements which collectively produce results not achievable by the elements separately. Their added value comes from the relationships or interfaces among the elements. (For example, open-loop and closed-loop architectures perform very differently.) But this value comes at a price: a complexity potentially too great to be handled by standard rules or rational analysis alone. Complexity

comes from many sources: uncertainties in technological change; geopolitical upheaval; environmental concerns; public perceptions of cost and risk; an opponent's weapons of attack or defense; and uncertainties in funding and political support. Similar complexities had been treated by earlier architects, but the principles of systems extended the scope to more disciplines and technologies. *See* CONTROL SYSTEMS.

As projects became ever more complex and multidisciplinary, new structures were needed for projects to succeed. It was no longer sufficient simply to use known elements, even in new ways. A skyscraper is not built of wood and stone. A stealth aircraft is not a converted airliner. Distributed workstations are not sliced-up mainframes. Object-oriented software architectures are neither modified communications protocols nor table-driven interpreters.

Nor could analytic techniques be used to find optimal solutions. Indeed, given the disparate perspectives of different customers, suppliers, and government agencies, unique optimal solutions generally would not exist. Instead, many possibilities might be good enough, with the choice dependent more on ancillary constraints or on the criteria for success than on detailed analysis. Nor could subsystems be locally optimized; the system as a whole might suffer.

Conceptual phases. As increasingly complex systems were built and used, it became clear that success or failure had been determined very early in their projects. The industry adage that "all the serious mistakes (in software) are made in the first day" may be an overstatement, but in the early phases all the critical assumptions, constraints, choices, and priorities are made that will determine the end result. Unfortunately, no one knows in the beginning just what the final performance, cost, and schedule will be. But the decisions which will determine them will have been made. The later phases, assuming they are carried out well, contribute much less.

Systems-level architecture. It is no coincidence that the systems-level architecture is laid out in the early phases. Systems-level architecture specifies how system-level functions and requirements are gathered together in related groups. It indicates how the subsystems are partitioned, what the relationships between the subsystems are, what communication exists between the subsystems, and what parameters are critical. It makes possible the setting of specifications, the analysis of alternatives at the subsystem level, the beginnings of detailed cost modeling, and the outlines of a procurement strategy.

Design with incomplete information. Information critical for rational architectural decision making is seldom available at the systems-level architecture stage. The practice of systems architecture (architecting), then, must be more an art than a science, more dependent on insights and intuition than on quantitative data.

It is simply unrealistic to postulate that complex systems designs begin with iron-clad requirements which flow down to specifications, and hence to an optimum system. For example, there rarely is enough information early in the design stage for the client to decide on the relative priority of the requirements without having some idea of what the end system might be. Instead, provisional requirements and alternative system concepts have to be iterated until a satisfactory match is produced.

Joint participation. Unavoidably, successful systems architecting in the conceptual phase becomes a joint process in which both client and architect participate heavily. In the ideal situation, the client makes the value judgments (What is good enough? What is affordable? When is it needed?), and the architect makes the technical decisions (What can be built? How can it be done? How long will it take?). Failure of either party to make timely decisions has all the subsequent deleterious consequences that might be expected: performance deficiencies, cost and schedule overruns, emergency redesigns, software patches and restarts, disappointed sponsors and users, recriminations and even lawsuits.

Conceptual model. Systems-level architecture begins with a conceptual model, a top-level abstraction which attempts to discard features deemed not essential at the system level (for example, the color of paint on an aircraft). Physically, the model can be a small-scale display built out of balsa and paper, a block diagram of an advanced radar, a computer simulation of a spacecraft, or a bubble diagram of a software program. Such a model is an essential tool of communication between client, architect, and builder, each viewing it from a different perspective. As the system comes into being, the model is progressively refined. Secondary features are brought into play; subsystems are fleshed out; subsystem components critical to the system as a whole are identified and tracked (for example, an untried technology).

Acceptance criteria. As development nears completion, a hard truth becomes apparent. What is actually being built is determined by the acceptance criteria—not the desired functions, not the conceptual model, but what will pass the specified acceptance tests.

The builder's interest is in being paid for delivery of a system that has passed its final test, no more and no less. The buyer's interest is in receiving a system as originally conceived. The key to resolving this conflict is the early incorporation of acceptance criteria into the conceptual model of the systems-level architecture.

This failure to foresee and resolve such situations in the conceptual stage can easily lead to severe disappointment, recriminations, financial loss to the builder, delays in operation for the buyer, and costly redesign. Systems that have not been partitioned to minimize communications between the subsystems will fail when first assembled, seriously complicating system-level testing and diagnosis. Untestable or unpassable acceptance criteria will provoke argument and dissension. Systems that have not been designed to be tested will prove untestable. Systems that do

not have designed-in quality will face difficult trials as efforts are made to test them. System designs that do not solve the users' real problems will be rejected.

Recent developments. The most important addition to the field of systems architecture since 1990 is software systems architecture. Not only does it contribute greatly to the theory and practice of architecting, but it comes right at the time when "smart," or software-intensive, systems are displacing simpler, hardware-only predecessors. Smart systems are far more capable, less expensive to produce, and faster to reach the marketplace, and consequently are much more profitable than their predecessors. They are also critically dependent not only on their own software architectures but on the architecture of the systems they support. *See* SOFTWARE ENGINEERING.

Eberhardt Rechtin

Bibliography. M. W. Maier and E. Rechtin, *The Art of System Architecting*, 2d ed., CRC Press, 2000; G. Nadler, *The Planning and Design Approach*, John Wiley, 1981; E. Rechtin, *Systems Architecting: Creating and Building Complex Systems*, Prentice Hall, 1991; E. Rechtin, *Systems Architecting of Organizations: Why Eagles Can't Fly*, CRC Press, 2000; M. Shaw and D. Garlan, *Software Architecture: Perspectives on an Emerging Discipline*, Prentice Hall, 1996.

Systems ecology

The analysis of how ecosystem function is determined by the components of an ecosystem and how those components cycle, retain, or exchange energy and nutrients. Systems ecology typically involves the application of computer models that track the flow of energy and materials and predict the responses of systems to perturbations that range from fires to climate change to species extinctions. Systems ecology is closely related to mathematical ecology, with the major difference stemming from systems ecology's focus on energy and nutrient flow and its borrowing of ideas from engineering. *See* ECOLOGY; MATHEMATICAL ECOLOGY.

Historical overview. Systems ecology emerged as a mathematical framework for posing questions about how ecosystems work, shortly after the concept of ecosystems gained momentum. R. H. Lindeman launched systems ecology with his pioneering idea of depicting ecological systems, such as a lake, as compartments of feeding groups (producers, herbivores, carnivores) or trophic levels through which energy flowed. As ecologists began to track the flow of energy, they also began to track the flow of nutrients and started asking questions about efficiency, stability, and production, all of which bring to mind theories from engineering and the industrial design of production systems. At first systems ecology was in conflict with evolutionary ecology because systems ecologists often discussed how systems evolved toward particular properties (more efficiency and more stability), whereas evolutionary biologists argued that there was no grand design of ecosystems and that evolution was the result of natural selection acting on individuals. Ecosystems do not evolve. Systems ecology no longer discusses the evolution of ecosystem attributes, but it still focuses on emergent properties such as stability and resilience. *See* ECOLOGICAL COMMUNITIES; ECOLOGICAL ENERGETICS; ECOSYSTEM.

Ecosystem theory. Ecosystem theory draws heavily on applied mathematics and engineering studies to ask questions about system performance. For example, when designing a manufacturing or control system, systems engineers have recognized the value of redundancy, so that if one component fails the entire system does not fail. This seemingly abstract idea becomes practical when investigating the value of protecting biodiversity. If it is believed that many species are redundant, then an engineering perspective would suggest that losing species might not produce any noticeable changes in ecosystem functioning over a short time scale. Only after too many redundant species were lost would a pronounced change in system performance be noticed (**illus.** *a*). Alternatively, if species were not redundant, then each loss of a species might be expected to decrement ecosystem performance, resulting in a linear decline in productivity or stability as the number of species declined (illus. *b*). The scientific

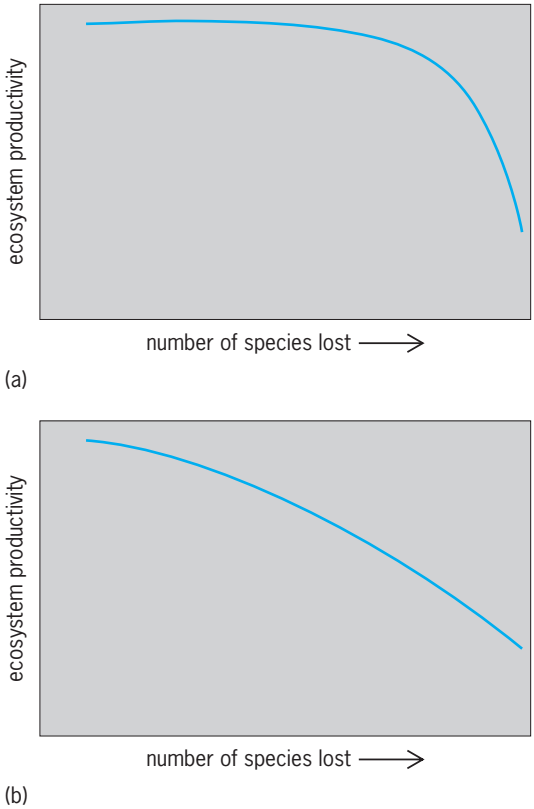

Decrease in system productivity with loss of species. (*a*) Under the assumption that most species are redundant. Initial losses do not matter because the remaining species are able to carry out the ecosystem function necessary for production. (*b*) Under the assumption that there are no redundant species. In this case, each species that is lost causes some decrement in system productivity.

challenge is to determine how ecosystems are constructed—are they clusters of functional groups, with many redundant species in each functional group, or are they collections of truly unique species, with each species performing a unique function? *See* BIODIVERSITY; SYSTEMS ENGINEERING; THEORETICAL ECOLOGY.

Large-scale studies and models. Systems ecology is an especially valuable approach for investigating systems so large and complicated that experiments are impossible, and even observations of the entire system are impractical. In these overwhelming settings, the only approach is to break down the research into measurements of components and then assemble a system model that pieces together all components. The world's largest ecosystem is the North Pacific Ocean gyre, which covers 20 million square kilometers of the Earth's surface. This huge ecosystem cannot be experimentally manipulated. It is even hard to measure aspects of this huge oceanic ecosystem. Nonetheless, by constructing food-web models that track the flow of energy and carbon from photosynthetic plankton to herbivores to higher trophic levels, systems ecologists have collaborated with oceanographers to discover an important pathway for carbon, which has implications for feedback loops between biological processes and global climate change. *See* FOOD WEB; PACIFIC OCEAN.

Specifically, it appears that there is a major downward flux of carbon into the bottom of the Pacific Ocean, where it becomes the basis of a bacterial community. The amount of carbon that is pumped downward is so substantial that variations in this pump could alter the amount of carbon dioxide in the atmosphere and, hence, the rate of global warming. The story is complicated because phytoplankton production, which "primes the pump," varies as a result of ocean circulation patterns and nutrient limitation. Thus there is a very complex system in which global climate change may alter patterns of ocean circulation and primary production, which in turn alters the carbon sink represented by the ocean, which then feeds back to influence the magnitude of carbon dioxide increases in the atmosphere and global climate change. These cycles, pumps, and feedback loops are not understood well enough to make firm predictions; however, it is clear that the tools of systems ecology are necessary for investigating such complex dynamics, in which no simple experiment or suite of observations is likely to produce unequivocal insight. *See* CARBON DIOXIDE; GLOBAL CLIMATE CHANGE; PHYTOPLANKTON.

Addressing environmental problems. An important contribution of ecosystem science is the recognition that there are critical ecosystem services such as cleansing of water, recycling of waste materials, production of food and fiber, and mitigation of pestilence and plagues. The United Nations and nonprofit organizations are combining to fund global assessments of these ecosystem services, with the idea of providing key information to policymakers regarding possible threats to environmental well-being. Systems ecology becomes an important tool in interpreting data about changing ecosystems and possible loss of ecosystem services, because it asks whether crossing certain thresholds can so fundamentally change a system that recovery is impossible. An example of such a threshold involves semiarid grazing systems in which the vegetation is composed of grasses and woody shrubs. Cattle eat grasses but not the woody plants. In the absence of cattle, many of these savannah systems are dominated by grasses with a scattering of shrubs. As the number of cattle on the land increases, the density of grass declines. At some point the cattle grazing pressure can be so severe that the grasses effectively disappear and the landscape becomes a barren shrubland. At that point, even if the cattle are taken off the land it may never recover, because the system has moved to a new equilibrium in which the shrubs are able to outcompete and prevent the grasses from returning. This system has been modeled, and the implications of this model for resource management are clear: overgrazing of certain systems may not be reversible and, hence, must be managed carefully to maintain the land's capacity for food production. *See* ECOLOGY, APPLIED.

Scaling laws. Variability changes as a function of the spatial scale and temporal scale over which a process is observed. The scaling of variability is important because variability drives ecological dynamics and influences human resource extraction. In addition, if scientists fail to appreciate the sensitivity of scale measurements, the results obtained can be seriously misleading. A result obtained at the scale of a square meter may disappear if measured at a scale of thousands of square kilometers. Spatial scale also commonly determines how tight the feedback loops are expected to be, with implications for environmental policy. For example, it is easy to get neighborhood groups to adopt environmentally rational zoning laws because the feedback loops between zoning laws and local greenspace, parks, and quality of life are tight. However, it is much harder to get the public excited about modifying their individual patterns of carbon consumption in ways that would reduce global warming, because the feedback loop between local behavior and total global carbon budget is very loose; the translation from local action to global effects is almost invisible when it comes to carbon fluxes.

Linking individual ecology, population ecology, and ecosystem dynamics. Systems ecology is one of the few theoretical tools that can simultaneously examine a system from the level of individuals all the way up to the level of ecosystem dynamics. An excellent example of this type of system analysis entails a detailed model of Florida's Big Cypress Natural Preserve, in which the trophic exchanges (measured as grams of carbon per square meter per year) among 68 principal taxa are inserted in a network model depicting all feeding relationships. The taxa in this cypress swamp ranged from phytoplankton to vines, cypress trees, snails, frogs and salamanders, woodpeckers, vultures, and all the way up the food chain to the Florida panther and American alligator. Without a model, the first guess would be that

any increase in a predator population would naturally decrease the population of its prey. However, by quantifying the details of carbon exchange among all of these interacting components, biologists have found that in some cases an increase in a predator population actually also increases its prey population.

A notable example of the positive effects of an increase in predators on prey populations involves the alligator and frogs and salamanders in the Cypress swamp. Alligators certainly eat large numbers of frogs and salamanders, but they also eat snakes, which are especially voracious predators of frogs and salamanders. The net result is that the more alligators in a region, the fewer snakes and the more frogs and salamanders. Software packages exist that can determine when these sorts of counterintuitive results are likely to emerge from complicated networks of interacting species. By using such a network system model, it is possible to predict the consequences of removing any species in an ecosystem. When such a network model is used to broadly explore the role of alligators, it becomes clear that alligators are a keystone species for the Cypress swamp—this means that any large change in alligator abundance would dramatically alter the Cypress swamp ecosystem. *See* POPULATION ECOLOGY.

Human impact. No ecosystem is immune to human impact. In fact, for many ecosystems, human activities are the dominant drivers of change. This means that studying ecosystems without also studying human systems (such as commerce, agriculture, and living patterns) is a sterile and disconnected academic pursuit. For that reason, systems ecology is now combining with the social sciences to model human and natural systems in a coherent framework. The principles and theory are similar—attention is paid to redundancy, attributes that might confer stability, and feedback loops. Models that couple patterns of human land use to biological properties of systems are a good example of how social science might be infused into systems ecology. Land-use change is the most important pressure altering biodiversity; land-use change also alters regional climate, hydrology, agricultural productivity, and even human exposure to diseases. Patterns of taxation, zoning laws, and highway systems interact with soil fertility and vegetation types to shape land-use changes. Consequently, the dynamics of ecosystems may be governed as much by zoning laws as by predator-prey relationships and grazing. But there is more to the melding of social sciences and systems ecology than simply noting that human systems alter ecological dynamics. The basic productivity of ecosystems may have constrained the type of human cultures that have evolved, and it prohibits the development of wealth if lands are so degraded that food production is impossible. Thus, ecological dynamics also alter human systems in profound ways. *See* BIOLOGICAL PRODUCTIVITY.

P. M. Kareiva

Bibliography. T. Allen and T. Starr, *Hierarchy*, University of Chicago Press, 1982; R. Kitching, *Systems Ecology: An Introduction to Ecological Modeling*, University of Queensland Press, 1983; S. Levin, *Fragile Dominion*, Perseus Publishing, Cambridge, MA, 1999; H. T. Odum, *Systems Ecology*, Wiley Interscience, New York, 1983; C. Walters, *Adaptive Management of Renewable Resources*, McGraw-Hill, New York, 1986.

Systems engineering

A management technology involving the interactions of science, an organization, and its environment as well as the information and knowledge bases that support each (**Fig. 1**). The functional definition of technology, as a fundamental human activity, is the organization, application, and delivery of scientific knowledge for the presumed enhancement of society. Thus, technology inherently involves a purposeful human extension of one or more natural processes. Management involves the interaction of the organization with the environment. Associated with this, to make an effective management technology, is the information and knowledge that enables understanding and action to effect change.

Characteristics. The purpose of systems engineering is to support organizations that desire improved performance. This improvement is generally obtained through the definition, development, and deployment of technological products, services, or processes that support functional objectives and fulfill needs. Thus, systems engineering is inherently associated with user organizations and humans in fulfillment of their objectives. The engineering of a system also involves the interaction with humans and organizations that are responsible for the physical implementation of systems. Systems engineers generally play an important role as brokers of information and knowledge in working with user enterprises and with the implementation specialists who accomplish the actual realization of physical systems. **Figure 2** illustrates these conceptual interactions. The enterprise will have various functional needs; the conceptual design parts of a systems engineering effort are concerned with expressing these needs in the form of a functional architecture. Systems engineers are also concerned with translation of this functional architecture into a physical architecture that describes the logical breakdown of the system into subsystems. Each subsystem should be as independent as possible and should be such that integrating them after implementation is straightforward. This physical architecture, or logical design description, of the system is next translated into an implementation architecture that provides guidance for implementation contractors in bringing about the various subsystems. This is also represented in Fig. 2.

Systems engineering has triple bases: a physical (natural) science basis, an organizational and social science basis, and an information science and knowledge basis. The natural science basis involves primarily matter and energy processing. The organizational and social science basis involves human, behavioral, economic, and enterprise concerns. The

information science and knowledge basis is more difficult to cope with, in many ways, than the other bases, since knowledge is not a truly fundamental quantity. In part, this derives from the structure and organization inherent in the natural sciences and in the organizational and social sciences. It also results from the uses to which information is to be put, and the experiential familiarity of information holders with the task at hand and the environment into which the task is embedded such as to enable interpretation of information, within an appropriate context, as knowledge. Thus, the presence of information in Fig. 1 is especially important.

There are at least three major drivers of new technologies: (1) The natural sciences must provide new discoveries that can be converted into technological innovations. (2) There must be a marketplace need for technological innovations. (3) There must be a need to ensure sustainable development.

These three suggest consideration of a technical system, an enterprise system, and a knowledge system. Management of the environment for each is needed. Systems engineers often act as brokers of knowledge in enterprises that have needs for support and implementation specialists able to deliver innovative technologies and services that provide this support. **Figure 3** illustrates these interrelations. It indicates that systems engineering knowledge comprises (1) knowledge perspectives, which represent the view that is held relative to future directions in the technological area under consideration; (2) knowledge principles, which generally represent formal problem-solving approaches to knowledge, generally employed in new situations or unstructured environments; and (3) knowledge practices, which represent the accumulated wisdom and experiences that have led to the development of standard operating policies for well-structured problems.

These three components act together and are associated with learning to enable continual improvement in performance over time. On the basis of the appropriate use of these knowledge types, systems engineers are able to accomplish the technological system design and management system design that lead to an innovative product or service. Knowledge perspectives enable the forecasting of the need for innovation. Innovation results when new knowledge principles are applied to produce new and different products and services, and associated knowledge practices, that fulfill a societal need.

Systems engineering may also be defined as management technology to assist and support policy making, planning, decision making, and associated resource allocation or action deployment. It accomplishes this by quantitative and qualitative formulation, analysis, and interpretation of the impacts of action alternatives upon the needs perspectives, the institutional perspectives, and the value perspectives of clients to a systems engineering study. In fact, all of systems engineering can be thought of as consisting of formulation, analysis, and interpretation activities. Practitioners may exercise these in a formal sense or in an "as if" or experientially based intuitive

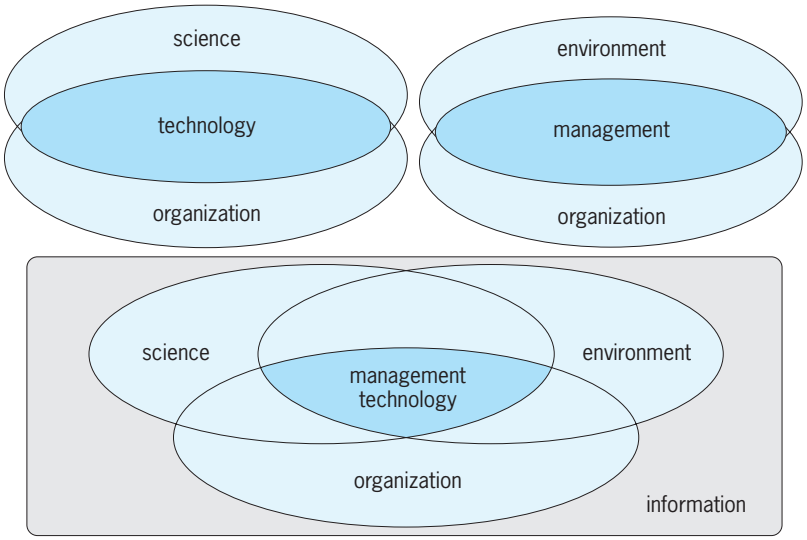

Fig. 1. Systems engineering as a management technology.

sense. Each essential phase of a systems engineering effort—definition, development, and deployment—is associated with formulation, analysis, and interpretation. These enable systems engineers to define the needs for a system, develop the system, and deploy it in an operational setting and provide for maintenance over time. These are the components of a framework for systems engineering (**Fig. 4**).

In order to resolve large-scale and complex problems or manage large systems, systems engineers must be able to deal with contemporary issues that involve:

1. Many considerations and interrelations
2. Many different and perhaps controversial value judgments
3. Knowledge from several disciplines
4. Risks and uncertainties involving future events which are difficult to predict
5. Fragmented decision-making structures
6. Needs perspectives and value perspectives as well as technology perspectives
7. Resolution of issues at the level of institutions and values as well as the level of symptoms

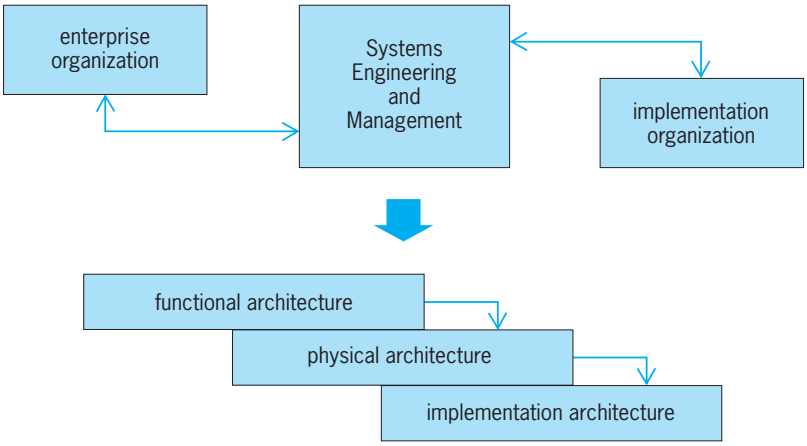

Fig. 2. Systems engineering as a broker of knowledge to enable specification of system architectures.

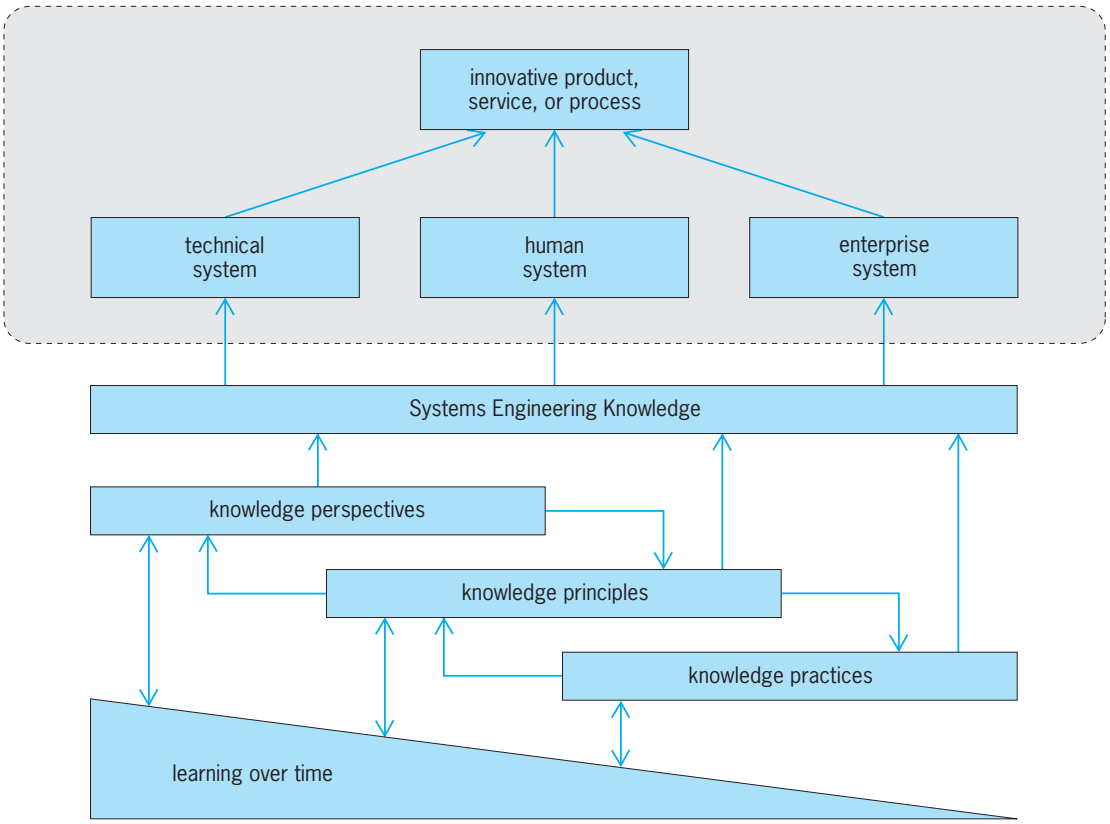

Fig. 3. Systems engineering knowledge and results of its effective use.

In addition, there are potential pitfalls to be avoided:

1. Overreliance upon specific methods or technologies advocated by some group
2. Consideration of perceived problems and issues only at the level of symptoms, and the development and deployment of "solutions" that address only symptoms
3. Failure to develop and apply appropriate lifecycle processes that allow formulation, analysis, and interpretation of impacts in terms of institutional and value considerations
4. Failure to involve the client
5. Failure to consider the effects of poor information processing
6. Failure to identify a sufficiently robust set of options for development
7. Failure to make and properly utilize reactive, interactive, and proactive measurements to guide the systems engineering efforts
8. Failure to identify risks associated with system costs and benefits, or effectiveness
9. Failure to properly design the system for effective user interaction
10. Failure to consider implications of strategies adopted in one of the three life cycles [research, design, testing, and evaluation (RDT&E); acquisition; and planning and marketing] on the other two life cycles
11. Failure to address quality issues comprehensively through all phases of all life cycles
12. Failure to properly integrate a new system with heritage or legacy systems
13. Failure to obtain top-level enterprise management support for the effort
14. Failure to consider sustainability issues

Emergence. Throughout history, the development of more sophisticated tools has invariably been associated with a decrease in our dependence on human physical energy. Generally, this is accomplished by control of nonhuman sources of energy in an automated fashion. The industrial revolution represented a major thrust in this direction.

In most cases, a new tool or machine makes it possible to perform a familiar task in a new and different way, typically with enhanced efficiency and effectiveness and sometimes with increased explicability. Sometimes a new tool makes it possible to do something entirely new.

Profound societal changes have often been brought about by new tools. In the 1850s about 70% of the labor force in the United States was employed in agriculture. Now, less than 3% is so employed, but is able to produce sufficient food for the entire country, generally with large surpluses. Occasionally, tools have produced undesired side effects. Pollution due to chemical plants and potential depletion of fossil fuel sources are examples. There is the potential for even worse side effects, such as following operational errors in nuclear power plants.

Concerns associated with the design of tools for efficient and effective use have always been addressed,

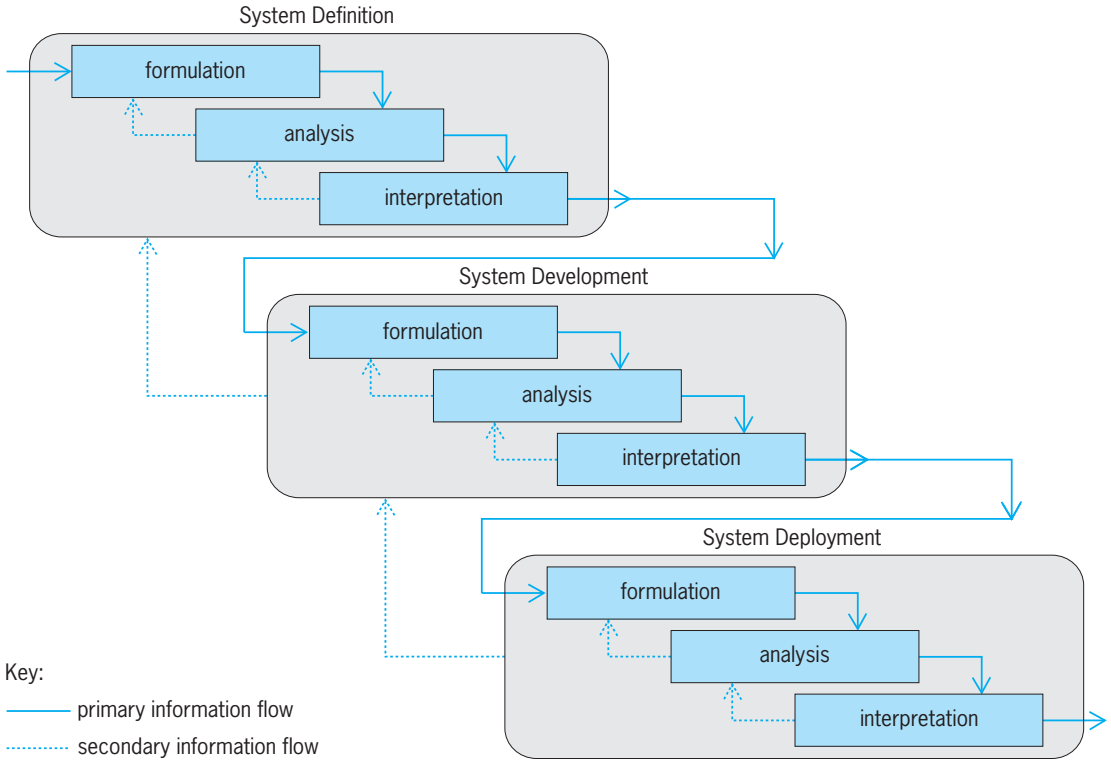

Fig. 4. A systems engineering framework comprising three phases and three steps per phase.

but often on an implicit and trial-and-error basis. In the past, when tool designers were also tool users, the designs were often good initially or soon evolved into good designs. But when physical tools, machines, and systems became so complex that they required a team of designers, a host of new problems emerged. To cope, a number of methodologies associated with systems engineering have evolved. Through these, it has been possible to decompose large design issues into smaller component ones to design the subsystems, and then to collect these subsystems into a complete system. However, simply connecting together the individual subsystems often does not result in an effective and efficient system. This has led to the realization that systems integration engineering and systems management are necessary throughout an entire system life cycle. Thus, contemporary systems engineering focuses on tools, methods, and metrics, as well as on the engineering of life-cycle processes that enable appropriate use of these tools to produce trustworthy systems. There is also a focus on systems management to enable the wise determination of appropriate processes. *See* SYSTEMS INTEGRATION.

Computers, information, and knowledge technologies. Computers are fundamentally different from the usual combination of motors, gears, pulleys, and other physical components that have assisted humans in performing physical tasks. While a computer can assist in performing functions associated with physical tasks, such as computing the optimal trajectory for an aircraft to take between two locations with minimum energy consumption and cost, it can also assist humans in cognitive tasks such as planning, resource allocation, and decision making. The availability of digital computers, and associated communications and networking capabilities, has led to the information technology revolution. In turn, this has led to an increased focus on the use of knowledge and intellectual capital as a major driver of productivity in the developed nations.

The digital computer is an information machine, a knowledge machine, and a cognitive support machine. It has led to the growth of a new engineering field which involves information and knowledge management. This professional area is concerned with efforts whose structure, function, and purpose are associated with the acquisition, representation, storage, transmission, and use of data that are of value as information, and the ability to associate this information with context so as to enable it to become knowledge. This has led to a fundamental change in the way that systems engineering is accomplished. *See* COMPUTER; COMPUTER-AIDED DESIGN AND MANUFACTURING; COMPUTER-AIDED ENGINEERING; COMPUTER-INTEGRATED MANUFACTURING; DATA COMMUNICATIONS; DATABASE MANAGEMENT SYSTEM; DIGITAL COMPUTER.

Requirements. To achieve a high degree of functionality, the requirements of a system must be appropriately defined. The system is often an assemblage of subsystems that must be integrated into existing systems. Systems engineers need to develop a system concept, or architecture, that results in a design that is capable of being efficiently and effectively produced, used, maintained, retrofitted, and modified throughout an extended life cycle. The system life cycle begins with the need for conceptualization

and identification, and proceeds through specification of system requirements and architectural structures, production or manufacturing, to ultimate system installation and evaluation, to ultimate operational implementation.

Many difficulties are often associated with the production of functional reliable, and trustworthy systems of large scale:

1. Large systems are expensive.
2. System capability is less than expected.
3. System deliveries are quite late.
4. Large system cost overruns occur.
5. Large system maintenance is complex and error-prone.
6. Large system documentation is inadequate.
7. Large systems are cumbersome to use, and system design for human interaction is generally lacking.
8. Individual subsystems cannot be integrated.
9. Large systems cannot be transitioned to a new environment or modified to meet evolving needs.
10. Large system performance is unreliable.
11. Large systems do not perform according to specifications.
12. System requirements do not adequately capture user needs.

A plethora of potential difficulties have led to these observations, including inconsistent, incomplete, and otherwise imperfect system requirements specifications; system requirements that do not provide for change as user needs evolve; and poorly defined management structures for product design and delivery. These same studies generally show that the major problems associated with the production of trustworthy systems have more to do with the organization and management of complexity than with the underlying technologies.

Defunctions. Systems engineering may be viewed as the design, production, and maintenance of functional, reliable, and trustworthy systems within cost and time constraints. There is an embedded hierarchy of performance levels for systems engineering efforts. Discussion of systems engineering is assisted by definitions of the discipline in terms of structure, function, and purpose.

Structure. The structural definition of systems engineering states that practitioners are concerned with a framework for problem resolution that, from a formal perspective at least, consists of three steps: issue formulation, issue analysis, and issue interpretation. Regardless of the characterization of the design process, and regardless of the type of process or system that is being designed, all characterizations necessarily involve (1) formulation of the issue, in which the needs and objectives of a client group are identified, and potentially acceptable alternatives or options are identified or generated; (2) analysis of the alternatives, in which the impacts of the identified design options are evaluated; and (3) interpretation and selection, in which options are compared by evaluating the impacts of the alternatives. The needs and objectives of the client group are used as a basis for evaluation. The most acceptable alternative is selected for implementation or further study in the systems engineering effort.

Function. The functional definition of systems engineering states that practitioners will be concerned with the various tools and techniques that enable us to design systems. Often, these will be systems science and operations research tools. It also tells us that systems engineers will be concerned with a combination of these tools to enable the establishment of an appropriate life cycle in engineering a system. Finally, the definition states that systems engineers will establish an appropriate setting. The term "systems management" will be used to refer to the cognitive tasks in producing a useful process from a systems methodology and design study, combining systems science and operations research to resolve issues and to engineer systems.

Purpose. The purpose of systems engineering is the organization and management of information and knowledge to assist clients who desire to develop policies for management, direction, control, and regulation. These activities relate to forecasting, planning, development, production, and operation of total systems to maintain overall integrity as related to trustworthiness in performance and reliability. The model of the steps of the fine structure of the systems process, shown in Fig. 4, is based upon this conceptualization.

Systems design model. Figure 4 illustrates a formal rational model of how design is accomplished. There is the need for much iteration between steps when it is discovered that improvements in an earlier step are needed in order to obtain a quality result at a later step. Also, this description does not emphasize the key role of information throughout all of systems engineering. This framework is not a complete view of the realities of systems engineering. It is correct in an "as if" manner. It is a morphological box that consists of a number of phases and steps. There is also a third-dimensional variable, effort level, and others could be identified as well.

A number of questions may be posed with respect to formulation, analysis, and interpretation that indicate the role of values in every portion of a systems engineering effort. Issue formulation questions are:

1. What is the problem? the needs? the constraints? the alterables?
2. How do the client and the analyst bound the issue?
3. What objectives are to be fulfilled?
4. What alternative options are appropriate?
5. How are the alternatives described?
6. What alternative state of nature scenarios are relevant to the issue?

Analysis questions of importance are:

7. How are pertinent state variables selected?
8. How is the issue formulation disaggregated for analysis?
9. What generic outcomes or impacts are relevant?
10. How are outcomes and impacts described across various societal sectors?
11. How are uncertainties described?

12. How are ambiguities and other information perfections described?

13. How are questions of planning period and planning horizon dealt with?

Interpretation concerns with respect to value influence are:

14. How are values and attributes disaggregated and structured?

15. Do value and attribute structuring and associated formal elicitation augment or replace experience and intuition?

16. How are flawed judgment heuristics and cognitive information processing biases dealt with?

17. Are value perspectives altered by the phase of the systems engineering effort being undertaken?

Finally, how is total issue resolution time divided between formulation, analysis, and interpretation? This is important because the allocation of resources to various systems engineering activities reflects the value perspectives of the analyst and the client. All of this has strong implications for a set of guidelines for professional practice of systems engineering.

There must also be an awareness of appropriate methods, tools, metrics, and techniques of systems science and operations research; and appropriate human judgment at the cognitive process level (systems management). Figure 4 and **Figure 5** show representations of the complete, simplified, systems engineering process described here. We will consider the three primary phases (definition, development, deployment) and three fundamental steps within each (formulation, analysis, interpretation).

Systems engineering efforts are concerned with the technical direction and management of systems design, production, and maintenance. The management technology of systems engineering aims to ensure that correct systems are engineered, which requires considerable emphasis on the front end of the systems life cycle. It also requires attention to ensure not only that the engineered system satisfies the technological specifications (verification) but that it performs so as to satisfy user needs (validation).

To support these ends, there must be emphasis on the accurate definition of a system, what it should do, and how people should interact with it before the system is produced and implemented. In turn, this requires emphasis upon conformance to system requirements specifications, and the development of standards to ensure compatibility and integratibility of system products. Such areas as documentation and communication are important. Thus, the need for the technical direction and management technology efforts that constitute systems engineering is apparent.

Life-cycle methodology. Figure 5 illustrates a typical sequence of seven phases for a systems engineering effort. These may be thought of as being sequenced in an iterative, "once-through," manner. In actual practice, the phases must be sequenced in an incremental or evolutionary manner so as to produce various versions of a product. Within each phase it is necessary to have a number of steps as represented in Fig. 4. Another representation is the systems engineering process framework, as shown in **Fig. 6**, comprising 21 steps, three for each of the seven phases. In practice, many more than 21 activity steps are associated with a systems engineering effort. One dichotomy of systems engineering practice is that, especially for large efforts, it is desirable to consider each of these activity boxes as

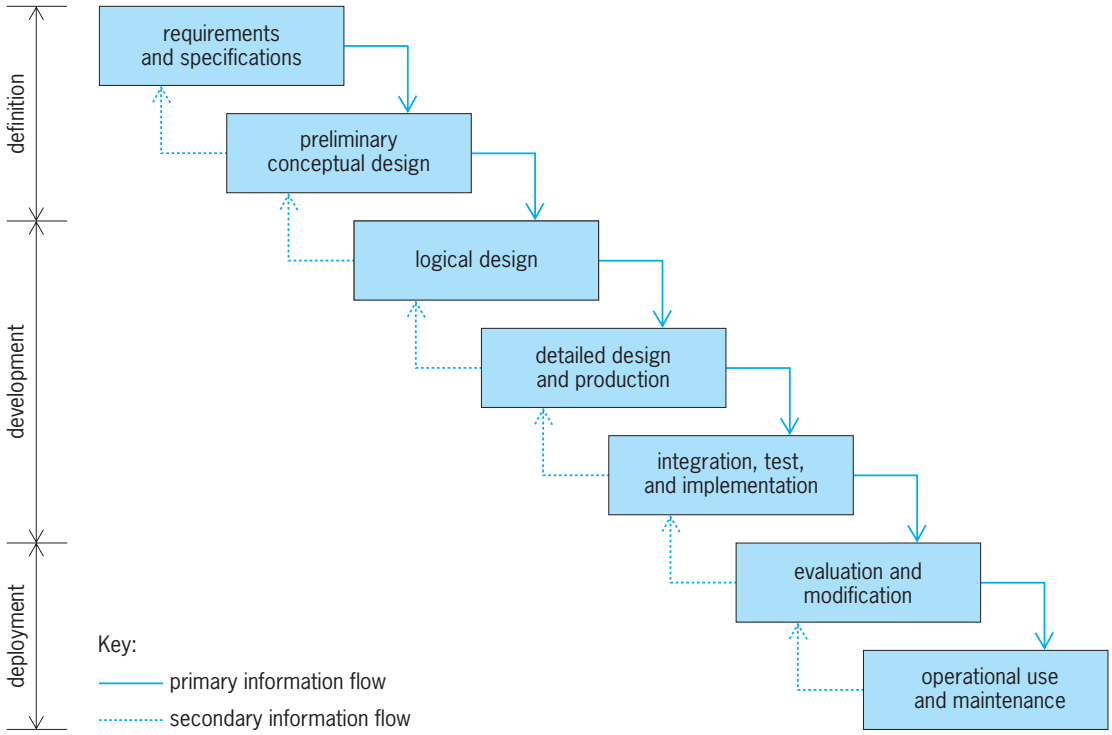

Fig. 5. One of several possible life-cycle models for systems engineering.

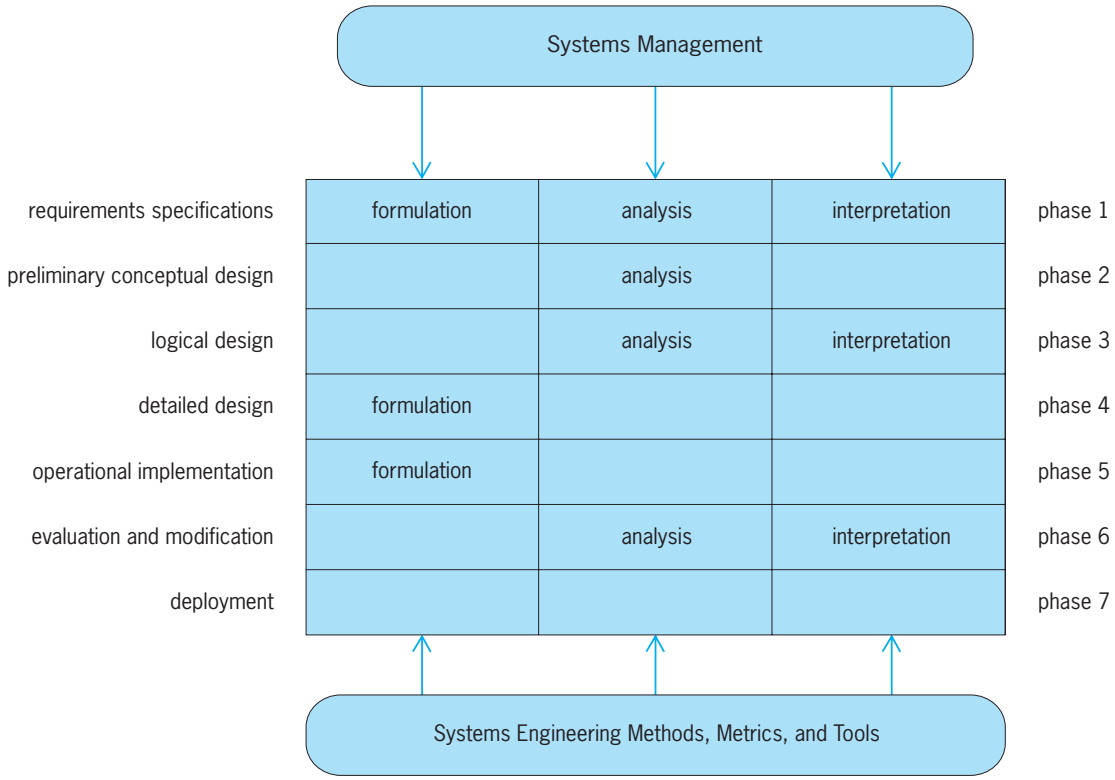

Fig. 6. An alternative representation of a two-dimensional systems engineering framework.

if it can be conducted in a manner dependent only on results from the previous activities. However, the whole is generally greater than the sum of the parts, and this independence rarely occurs. This leads to the need to implement a number of systems project management approaches, such as configuration management and control, that attempt to assure the needed integration throughout the systems engineering process. *See* ENGINEERING DESIGN; SOFTWARE ENGINEERING.

Life-cycle phases. In Fig. 5, the requirements specification phase of systems engineering has as its goal the identification of client needs, activities, and objectives to be achieved by implementation of the resulting design as a product, process, or system. Initially, the user requirements are identified, and these are translated to functional requirements, or specifications for the system. The effort in this phase should result in the description of preliminary conceptual design considerations that are appropriate for the next phase. It is necessary to translate operational deployment needs into requirements specifications in order that these needs be addressed by the system design efforts. Thus, the information concerning the requirements specifications is affected by each of the other phases of the systems engineering process.

From the requirements specifications phase, a clear definition of design issues should result so that it becomes possible to decide whether to undertake preliminary conceptual design. If the requirements specifications effort indicates that client needs can be accommodated in a satisfactory manner, then documentation is typically prepared concerning specifications for the preliminary conceptual design phase. Initial specifications for the following three phases are typically also prepared, and a concept design team is selected to implement the next phase.

Preliminary conceptual design typically includes or results in the functional architecture of the system. This leads to logical design and architectural specification of the physical architecture for the ultimate system (product, service, or process). Arriving at the functional architecture of the system may lead to the development of a prototype that is responsive to the requirements specifications. Preliminary concept design according to the requirements specifications should be obtained. Rapid prototyping of the functional architecture of the conceptual design is desirable for many applications. *See* PROTOTYPE.

The desired product of the next phase, logical design, in Fig. 5 is a set of detailed architectural specifications for a physical system architecture that should result in a useful product, process, or system. There should be a sufficient degree of user confidence that a useful product will result from detailed design, in terms of realization of an implementation architecture, or the entire effort should be redone or possibly abandoned. Another product of this phase is a refined set of specifications for the integration of the resulting system, both in terms of subsystems and in terms of legacy systems. In this third phase of effort, these specifications are translated into detailed representations in logical form, in terms of a physical architecture, such that detailed system development may occur.

A product, service, or process (system) is produced in the fourth phase of design. This is not the final design, but the result of implementation of the physical architecture and the functional architecture

resulting from the last two phases. Establishment of an implementation architecture is needed in this phase. In the fifth phase, the implementation architecture is realized in the form of a system in an operational setting. The actual details of implementation may be the responsibility of the systems engineering organization or of another organization so charged.

Evaluation of the implementation—in terms of the resulting product, process, or system—is achieved in the sixth phase of this process. Preliminary evaluation criteria are obtained as a part of requirements specifications and modified during the following two phases. The evaluation effort must be adapted to other phases such that it becomes an integral part of the overall process. Generally, the critical issues for evaluation are adaptations of the elements in the requirements specifications phase. A set of evaluation test requirements and tests are evolved from the objectives and needs determined in requirements specifications. These should be such that each objective measure and critical evaluation issue component can be measured with at least one evaluation test instrument. *See* SYSTEM DESIGN EVALUATION.

If it is determined, perhaps through an operational evaluation, that the system cannot meet user needs, the effort reverts to an earlier phase and effort continues. An important by-product of evaluation is determination of ultimate performance limitations for an operationally realizable system. Often, operational evaluation is the only realistic way to establish meaningful information concerning functional effectiveness of the result of a systems engineering effort. Successful evaluation is dependent upon explicit development of a plan for evaluation before initiation of the evaluation effort.

The last phase of the systems engineering effort concerns final acceptance and operational deployment. The process could continue until such extended life-cycle phases as maintenance, retrofit, and reengineering are identified. Alternatively, the operational deployment phase can be considered to include these activities. *See* REENGINEERING.

Systems management. Many ingredients are associated with the development of trustworthy systems, including:

1. Systems engineering processes, including process development life cycle and configuration management
2. Process risk, operational-level quality assurance and evaluation, and product risk and development standards
3. Metrics for quality assurance, and process and product evaluation
4. Metrics for cost estimation, and product cost and operational effectiveness evaluation
5. Strategic quality assurance and management, or total quality management
6. Organizational cultures, leadership, and process maturity
7. Reengineering at the levels of systems management, organizational processes and product lines, and product

A number of related issues concern enterprise management and systems integration, economic systems analysis, cognitive ergonomics, and system assessment and evaluation.

One of the first efforts in systems management is to identify process life cycle for production of a trustworthy system. The precise life cycle that is followed will depend on client needs. It will also depend on such environmental factors as the presence of existing subsystems into which a new system must be integrated, and the presence of existing software modules that may be retrofitted and reused as part of the new system. This need for system integration brings about a host of systems management issues and, in many cases, legal issues that are much larger in scale than those associated with program development. Similarly, the development of appropriate system-level architectures is very important. Efficiency and effectiveness in systems architecting have great influence on the ease with which systems can be integrated and maintained and, therefore, of the extent to which an operational system is viewed as trustworthy and of high quality. *See* SYSTEMS ARCHITECTURE.

Following the identification of an appropriate systemic process development life cycle, configuration management plans are identified. This step involves using the life cycle and defining a specific development process for the set of tasks at hand. Metrics are needed for effectiveness. Examples are the metrics of cost analysis, or cost estimation, for systems engineering such as cost and economic estimation for software and information technology–based systems; and of effectiveness analysis or estimation of software productivity indices. The development of a product is coupled with the process needs. These metrics must form a part of a management approach for process, and ultimately product, improvement if substantial progress is to be made in such important areas as software testing and evaluation. *See* ACTIVITY-BASED COSTING; SOFTWARE TESTING AND INSPECTION.

Conclusion. Much contemporary thought concerning innovation, productivity, and quality can be cast into a systems engineering framework. This framework can be valuably applied to systems engineering in general and information technology and software engineering in particular. The information technology revolution provides the necessary tool base that, together with knowledge management–enabled systems engineering and systems management, allows the needed process-level improvements for the development of systems of all types. The large number of ingredients necessary to accomplish needed change fit well within a systems engineering framework. Systems engineering constructs are useful not just for managing big systems engineering projects according to requirements, but for creative management of the organization itself. *See* INFORMATION SYSTEMS ENGINEERING; LARGE SYSTEMS CONTROL THEORY; QUALITY CONTROL; SYSTEMS ANALYSIS.

Andrew P. Sage

Bibliography. J. N. Martin, *Systems Engineering Guidebook: A Process for Developing Systems and Products*, CRC Press, Boca Raton, FL, 1997; E. Rechtin and M. W. Maier, *The Art of Systems*

Architecting, CRC Press, Boca Raton, FL, 1997; A. P. Sage, *Systems Engineering*, John Wiley, New York, 1992; A. P. Sage, *Systems Management for Information Technology and Software Engineering*, John Wiley, New York, 1995; A. P. Sage and W. B. Rouse (eds.), *Handbook of Systems Engineering and Management*, John Wiley, New York, 1999.

Systems integration

A professional activity that utilizes processes and procedures from systems engineering, systems management, and product development for the purpose of developing large-scale complex systems. These complex systems involve hardware and software and may be based on existing or legacy systems coupled with totally new requirements to add significant functionality through integration of new systems or subsystems. Systems integration generally involves combining products of several contractors to produce the working system. Systems integration applications range from creation of complex inventory tracking systems to designing flight simulation models and reengineering large logistics systems.

Life-cycle activities. Application of systems integration processes and procedures generally is in accordance with a set of life-cycle phases for systems engineering. Minimally, these systems engineering life-cycle phases are requirements definition, development and integration of the system, and deployment together with associated operations and maintenance. These three phases include such necessary efforts as feasibility analysis, program and project plans, logical and physical design, design compatibility and interoperability tests, reviews and evaluations, and graceful system retirement.

Systems integration activities require in-depth knowledge of pertinent methodologies, experience in the implementation and management of large-scale integration programs, and the ability to combine and utilize the results of existing system details and constraints on new requirements. A primary need is to work with existing systems and resources that may have been in place for years and have undergone a number of previous modifications to determine how the integration program should be managed to assure that the needs of the user are met. This will enable informed insights concerning such subjects as maintenance of the existing systems in the face of implementation of additional system requirements, and the transition of the existing system to the new one over a significant interval of time.

Systems integration relies on a number of activities that are interactive and iterative. The systems integration life cycle generally envisions the utilization of computer-assisted systems (or software) engineering (CASE) tools to develop, analyze, and implement large-scale systems integration programs. The purposes for CASE tools are to enhance productivity, improve quality of the product, and increase the functionality of the product in operation. *See* SOFTWARE ENGINEERING.

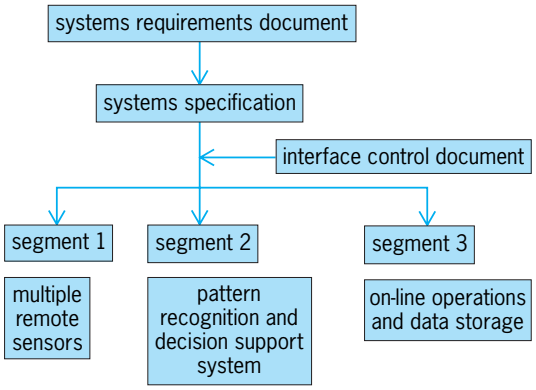

Fig. 1. Systems integration activity, showing the architecture for the system.

As an example, a prime contractor might be responsible for the development of a new data-fusion system involving an existing pattern-recognition system, modification of data storage and on-line operations, and improvement of decision support (**Fig. 1**). The associated activities might proceed as follows. Requirements specifications and feasibility analyses for the system are generally the responsibility of the prime contractor. Traceability matrices and an audit trail are established to track and account for all aspects of the new and improved system. System management parameters are fixed, and project planning is completed. Logical and physical designs are developed. Subcontract specifications and requests for proposal are prepared for the new hardware and software to be procured. Interface design and compatibility and interoperability tests are developed. Subcontractors are selected and are monitored by the prime contractor for contract compliance, quality control, and on-time and within-budget delivery. As component parts and subsystems are delivered, systems integration is accomplished by the prime contractor, generally with the assistance of subcontractors. Interface design, hardware and software compatibility, and interoperability testing are accomplished. Reviews and evaluations with the customer are continued throughout the process. System delivery is accomplished, and actual system operation is completed to customer satisfaction.

Primary uses. Systems integration is essential to the design and development of systems of all types, and especially systems that are information technology- and software-intensive, that automate key operations for business and government. It is required for major procurements for the military services and for private businesses.

Systems integration involves management of the entire life-cycle process from requirements definition to system development and ultimate deployment. It also includes implementation and training. The key personnel required for systems integration are program managers, communications specialists, software specialists, functional specialists, and hardware specialists. The most critical person is the program manager. Program management furnishes planning, controlling, and operating functions and applies these to large-scale complex systems

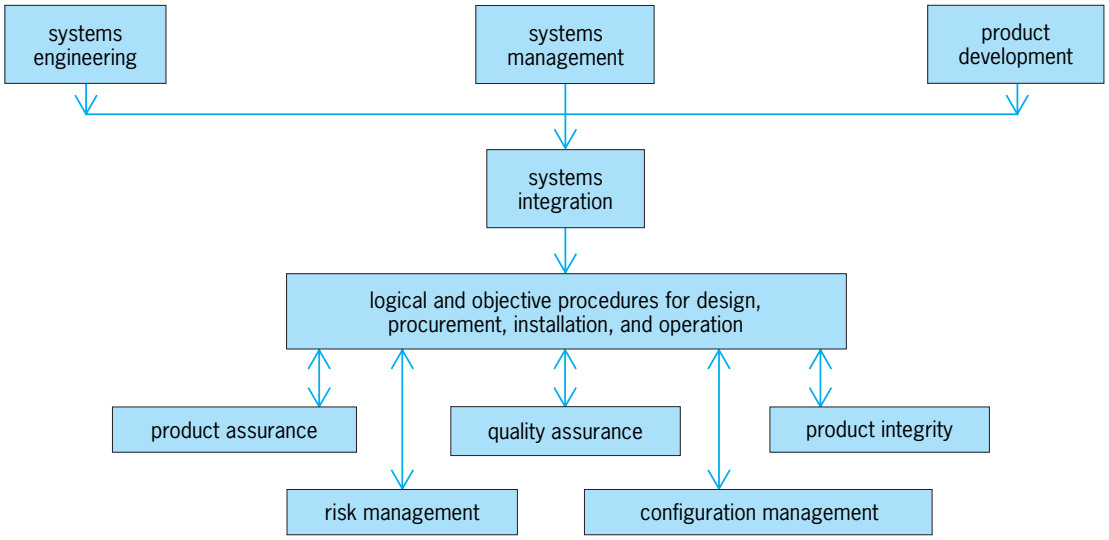

Fig. 2. Tactical and strategic perspective of systems integration, showing essential elements.

programs that include hardware, software, facilities, personnel, procedures, and training needed to meld existing systems and new technologies.

In the above example, systems integration required the coordination of preexisting or coexisting system components with newly developed or modified systems components. This resulted in significant growth of the existing system, merger of two other systems, and the combination of existing system fragments with commercial off-the-shelf components.

For the sample system, this included development and application of a strategic plan, assignment to subcontractors of all terms and conditions from the prime contract, assurance of full compliance with all requirements and constraints, development of an audit trail for complete traceability, and tracking of procurements and schedules. This process assured system growth, future modification, and expansion capabilities to meet initially unrecognized user needs. Systems integration processes and procedures include cost-control mechanisms and work breakdown schedules. They also include attempts to avoid bureaucratic procedures that increase costs to the client and delay deployment of the system (**Fig. 2**).

Application of processes and procedures. The application of systems integration to large-scale complex systems follows standard systems engineering lifecycle activities. The successful application of the systems integration approaches requires (1) completely and correctly defined client needs; (2) complete traceability; (3) consistent system performance criteria; (4) trade-off assessments for alternative solutions, rated as to efficacy and cost; and (5) operational test and evaluation of the integrated system to verify that it satisfies requirements.

The application of systems integration approaches requires that client needs are completely and correctly defined. Elicitation of client needs is one of the most difficult tasks in the development of large-scale complex systems. In most systems integration procurements, specifications are provided by the client and are to be interpreted and analyzed by the development team. This process requires significant interaction between client and developer to assure that the right system is being constructed and that it carries out the desired functions at the performance level needed. Generally, it is highly desirable to maintain close contact between the client and development team throughout the entire life cycle of the program to assure client satisfaction. Thus, requirements and specifications management and the maintenance of a complete audit trail of all actions taken in the program are essential for successful delivery and operation of the completed system.

Procurement criteria, based on function and performance need, must be established prior to consideration of trade-offs, alternatives, or options. This is a complex activity that involves understanding of the economic impact of the procurement, usually without full information on the current state of the art, as practiced by potential subcontractors. Selection criteria must be developed on a functional basis and not by reference to existing commercial configurations. An audit trial that leads from the functional performance requirements of the client to the final procurement choice must be maintained.

Operational test and evaluation criteria must be developed as early as feasible, preferably during requirements specifications, so as to assure that performance requirements will be met. These criteria are to be applied to each of the development phases and most especially to hardware and software procured from outside vendors.

Advantages. Systems integration approaches enable early capture of design and implementation needs. The interactions and interfaces across existing system fragments and new requirements are especially critical. It is necessary that interface and intermodule interactions and relationships across components and subsystems that bring together new and existing equipment and software be articulated. The systems integration approach supports this through application of both a top-down and a

bottom-up design philosophy; full compliance with audit-trail needs, system-level quality assurance, and risk assessment and evaluation; and definition and documentation of all aspects of the program. It also provides a framework that incorporates appropriate systems management application to all aspects of the program. One of the principal advantages of this approach is that it disaggregates large and complex issues and problems into well-defined sequences of simpler problems and issues that are easier to understand, manage, and build. Systems integration provides a suitable methodology that encompasses the entire integration program from requirements through design to construction, test, and finally deployment and maintenance; and supports problem understanding, risk management, and communication between all parties at all stages of development. It is generally necessary to develop plans for systems integration early in the systems engineering life cycle. Needs for systems integration should be established during the definition phase, and the systems architecture for integration is established early in the system development phase. *See* INFORMATION SYSTEMS ENGINEERING; RISK ASSESSMENT AND MANAGEMENT; SYSTEMS ANALYSIS; SYSTEMS ENGINEERING.

James D. Palmer

Bibliography. R. A. Beutel, *Contracting for Computer Systems Integration*, Michie Company, Charlottesville, VA, 1991; Computer Science and Telecommunications Board, *Systems Integration: Keeping the U.S. Industry Competitive*, National Academy Press, Washington, DC, 1991; J. O. Grady, *Systems Integration*, CRC Press, Boca Raton, FL, 1994; A. P. Sage, *Systems Engineering*, Wiley, 1992; A. P. Sage, *Systems Management for Information Technology and Software Engineering*, Wiley, New York, 1995; A. P. Sage and C. L. Lynch, Systems integration and architecting: An overview of principles, practices, and perspectives, *Sys. Eng.*, 1(3):176–227, November 1998; A. P. Sage and J. D. Palmer, *Software Systems Engineering*, Wiley, New York, 1990; A. P. Sage and W. B. Rouse (eds.), *Handbook of Systems Engineering and Management*, Wiley, New York, 1999; J. Wyzalek (ed.), *Enterprise Systems Integration*, CRC Press, Boca Raton, FL, 2000.

Syzygy

The alignment of three celestial objects within a solar system. Syzygy is most often used to refer to the alignment of the Sun, Earth, and Moon at the time of new or full moon. Although syzygy is strongly associated with these two lunar phases in many minds, it must be emphasized that the alignment of any three celestial objects within the solar system (or within any other system of objects in orbit about a star) constitutes syzygy. Alignments need not be perfect in order for syzygy to occur: because the orbital planes for any three bodies in the solar system rarely coincide, the geometric centers of three objects that are in syzygy almost never lie along the same line. *See* PHASE (ASTRONOMY); SOLAR SYSTEM.

In general, syzygy occurs whenever an observer on one of the three objects would see the other two objects either in opposition or in conjunction. Opposition occurs when two objects appear 180° apart in the sky as viewed from a third object. Conjunction occurs when two objects appear near one another in the sky as seen from a third object.

In the **illustration**, e-e' is the orbital plane of the Earth (E), p-p' is the orbital plane of an outer planet (P) that is tilted at an angle θ to the orbital plane of the Earth, and m-m' is the orbital plane of the Moon (M), which is tilted at an angle α to the Earth's orbital plane. Syzygy occurs for the Earth, Sun, and the planet when the planet is either at P_1 (conjunction) or P_2 (opposition). The Sun, Earth, and Moon are in syzygy when the Moon is either at M_1 (new moon) or M_2 (full moon). If the Moon is at M_2 when the planet is at P_2, then the Earth, Moon, and the planet are in syzygy. (As seen from the Earth, the Moon and the planet would be near one another in the sky; that is, the Moon and the planet would be in conjunction.) The Earth, Moon, and the planet will be in syzygy any time that the Moon and the planet are seen in the same direction from the Earth: the Moon need not be full and the planet need not be in opposition. The maximum angle for θ is about 17° (Pluto); the value for α is slightly more than 5°. There is no accepted upper limit for θ; if a new planet were discovered with an orbital plane tilted 25° from that of the Earth, syzygy would still occur at P_1 and P_2. *See* MOON; PLANET.

Solar and lunar eclipses are dramatic results of syzygy. During a solar eclipse, when the Moon is in its new phase, the alignment of the Sun, Earth, and Moon is so nearly perfect that the Moon's shadow falls on the Earth; during a lunar eclipse, which occurs at the time of the full moon, the Moon passes through the Earth's shadow. *See* ECLIPSE.

An occultation is another type of eclipse that can occur during syzygy. For an Earth-based observer, an occultation occurs when the Moon is seen to pass in front of a planet or other member of the solar system. The occultation of a star by the Moon does not qualify as syzygy, since the star is far beyond the limits of the solar system. *See* OCCULTATION.

Harold P. Coyle

Bibliography. V. Illingworth, *The Facts on File Dictionary of Astronomy*, 4th ed., 2000; J. Mitton, *Cambridge Dictionary of Astronomy*, 2001; J. M. Pasachoff and A. Filippenko, *The Cosmos: Astronomy in the New Millennium*, 3d ed., 2007.

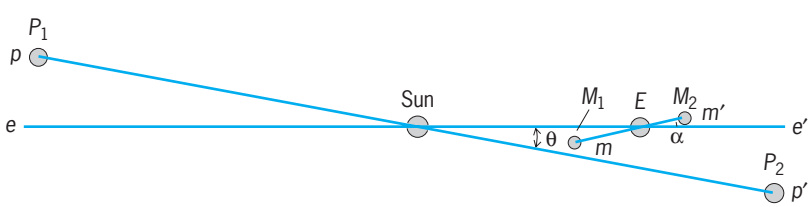

Diagram, not to scale, of the Sun, the Earth, the Moon, and a planet, illustrating syzygy. Symbols are explained in text.

T

T Tauri star

A young low-mass star characterized by variability, the presence of hydrogen emission lines, and association with dark or bright nebulae. T Tauri stars are named after a variable star in the constellation Taurus that exhibits particularly strong hydrogen emission. They were originally believed to be ordinary field stars passing through and interacting with a star-forming nebula. Their association with these regions was soon discovered to be more than coincidental, however, and they are now identified with the earliest phase of young stellar evolution, in which a star emerges from its natal molecular cloud to be detectable at visible wavelengths. T Tauri stars with very strong emission lines are designated classical T Tauri stars; their counterparts with reduced hydrogen emission are known as weak-line T Tauri stars. In addition to the original defining properties, T Tauri stars are generally accompanied by excess x-ray, ultraviolet, infrared, and millimeter-wave emission that arises from an accreting circumstellar disk of dust and gas. *See* MOLECULAR CLOUD.

Young age. Independent lines of evidence confirm the extreme youth of T Tauri stars. They are typically located next to dark clouds of dust and gas in aggregates (T associations) with stellar densities that are too low to have withstood disruption by galactic tides for more than 10^7 years. They exhibit high variability, as expected for objects that have not yet achieved equilibrium. They have luminosities and surface temperatures that place them above the zero-age main sequence on the Hertzsprung-Russell diagram, and have values that are reproduced by theoretical simulations of young stars prior to the onset of nuclear fusion of hydrogen in their inner cores. Evolutionary models indicate stellar ages between 10^5 and 10^7 years. Finally, T Tauri stars exhibit strong lithium absorption lines. The presence of lithium in low-mass stars with turbulent envelopes is an indicator of youth, since the lithium is mixed down to interior regions where it is rapidly depleted by nuclear reactions in more mature stars. *See* HERTZSPRUNG-RUSSELL DIAGRAM.

Origin of emission lines. The presence of high-velocity gas is implied by Doppler broadening of strong emission lines in the spectra of classical T Tauri stars. These features were originally attributed entirely to gas ejected from the star, partly on the basis of observations of bipolar jets in spectrophotometric images at very high spatial resolution. Theoretical simulations indicated that the jets are fueled by coupling of accreting disk material to a rotating stellar magnetosphere. Material leaves the disk plane close to the star and travels along magnetic field lines in a "funnel flow." Although most material falls onto the stellar photosphere, some is flung out of the system along open field lines to appear as bipolar jets. Classical T Tauri stars thus resemble more exotic examples of disk-jet systems, such as accreting white dwarf stars, pulsars, and black holes at the center of active galactic nuclei. More recent comparisons of observed spectral features with simulations from this theoretical framework have led to the conclusion that the high-velocity wings of spectral lines result mostly from the gas that is falling directly onto the star, although some features may still result from outflow. *See* BLACK HOLE; DOPPLER EFFECT; PULSAR; WHITE DWARF STAR.

Properties of circumstellar disks. The existence of circumstellar dust disks around classical T Tauri stars was first established by observations of long-wavelength emission in excess of stellar photospheric values. The distribution of dust radiation from infrared to millimeter wavelengths coincides with that of a disk surrounding the star with temperatures that decrease with disk radius. Simulations of the emission indicate the disks have masses from 0.01 to 10% of the mass of the Sun, sizes of about 100 astronomical units (1.5×10^{10} km or 1×10^{10} mi), and constituent dust grains that are substantially larger than for the interstellar medium. These

conclusions are supported by imaging at high spatial resolution. Aperture-synthesis images of carbon monoxide (CO) emission at millimeter wavelengths display the Doppler signature of a disk that is rotating in accord with Kepler's laws. Scattered light images obtained with the *Hubble Space Telescope* and ground-based telescopes reveal the outline of disk surfaces. Silhouetted disks, or proplyds (protoplanetary disks), also appear around stars located in front of bright nebulae. *See* HUBBLE SPACE TELESCOPE; KEPLER'S LAWS.

The most sensitive infrared observations to date indicate that circumstellar disks are detected around virtually all classical T Tauri stars. Far fewer are observed around weak-line T Tauri stars. This is as expected by theoretical interpretations that link emission line strength and the accretion of gas from a disk onto the star. The low detection rate may indicate the presence of disks that have evolved toward the formation of a planetary system, or it may belie the existence of young stars with little or no circumstellar material to begin with. Small amounts of dust from the collisions of large boulders and planetesimals in "debris disks" with very little molecular gas are detected around some stars in the solar neighborhood. Disks with these properties are currently difficult to observe at the distance of nearby star-forming regions, but a handful of detections around weak-line T Tauri stars indicates that at least some are surrounded by transitional planetesimal disks. Taken as a whole, the properties, rate of occurrence, and inferred evolution for disks around classical and weak-line T Tauri stars provide strong evidence that planet formation is a common by-product of the star-forming process. *See* EXTRASOLAR PLANETS; INFRARED ASTRONOMY; PROTOSTAR; SOLAR SYSTEM; STELLAR EVOLUTION. David Koerner

Bibliography. I. Appenzeller and R. Mundt, T Tauri stars, *Astron. Astrophys. Rev.*, 1:291–334, 1989; C. Bertout, T Tauri stars: Wild as dust, *Annu. Rev. Astron. Astrophys.*, 27:351–395, 1989; G. H. Herbig, The properties and problems of T Tauri stars and related objects, *Adv. Astron. Astrophys.*, 1:47–103, 1962; B. Reipurth, D. Jewitt, and K. Keil (eds.), *Protostars and Planets V*, University of Arizona Press, Tucson, 2006.

Tabulata

One of two principal orders of extinct Paleozoic corals. The Tabulata appeared in the Lower Ordovician and reached their acme in the Middle Devonian before being severely affected by the Late Devonian extinction event. Their subsequent rediversification was limited, and they became extinct at the end of the Permian. Tabulates were closely related to the other principal Paleozoic coral order, the Rugosa, but neither was ancestral to the post-Paleozoic Scleractinia, which evolved from a different group of anemones. Some Cambrian corals have been claimed to be Tabulata, but are more appropriately assigned to a new order.

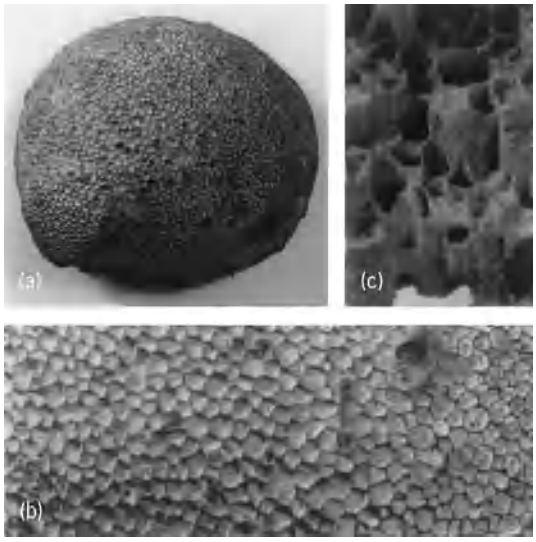

Fig. 1. Structure of the widespread tabulate coral *Favosites*. (*a*) Massive domal colony of *F. multipora* (Silurian, England), ×0.5. (*b*) Detail of surface showing polygonal corallites, ×1.5. (*c*) Oblique view of corallites in *Favosites* sp. (Silurian, Ontario). Note the short septal spines on walls, mural pores (bottom right), and tabulae (upper right, center and left), ×2.

Tabulate corals were exclusively colonial. Polyps secreted slender calcitic (calcium carbonate) tubes (corallites, ranging 0.5–20 mm in diameter, but predominantly 1–3 mm in diameter), polygonal in cross section when in contact, or cylindrical when surrounded by colonial skeletal material (coenenchyme) or not in contact (**Figs. 1** and **2**). The corallites are almost always partitioned by flat or curved, complete or incomplete plates (tabulae). Septal structures, as spines, or less commonly thin plates radially arranged in corallites, often number 12

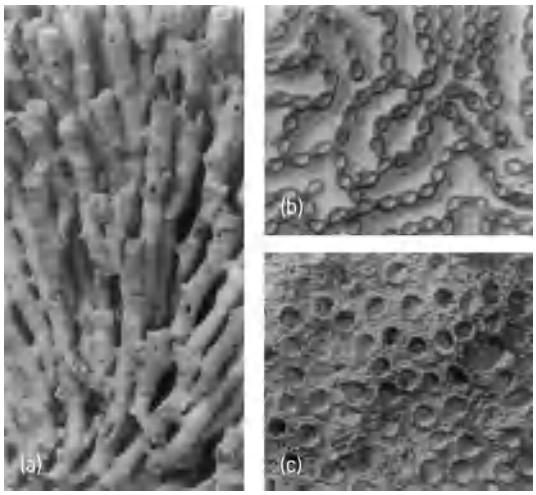

Fig. 2. Common tabulate corals. (*a*) *Syringopora geniculata* (Carboniferous, England), a fasciculate colony with interconnecting tubules (center-left). (*b*) *Halysites catenularius* (Silurian, England), a cateniform colony (chain coral). (*c*) *Heliolites megastoma* (Silurian, England), a massive colony with corallites set in colonial tissue (coenenchyme). All ×1.5.

when well developed. However, these structures are usually weakly developed or completely absent. The Tabulata is divided into six suborders—Lichenariina, Sarcinulina, Favositina, Halysitina, Heliolitina, and Auloporina—based mainly on the structural arrangement of corallites in the colony and the presence or absence of communication (mural pores or connecting tubules) between adjacent corallites. Exceptionally preserved polyps have been found in *Favosites* sp. from the Silurian of eastern Canada. Some groups formerly assigned to the tabulate corals, such as the Chaetetina, are now known to be calcareous sponges.

Tabulate coral colonies could take on a range of external forms. Shape was determined by the interaction of internal controls on colonial growth with prevailing environmental conditions. Solid masses of corallites, either close-packed or set in coenenchyme, are massive (Figs. 1*a* and 2*c*). Such colonies vary from thin, spreading sheets, through domal and bulbous forms, to columnar and branching forms (in which each branch is composed of a mass of corallites). Species vary in the range of forms they could adopt. Another group is characterized by corallites arranged like posts in a fence, giving a chainlike appearance (cateniform) in surface view (Fig. 2*b*). Others consist of short, discrete corallites ramifying across a host skeletal surface or the substrate, with only the aperture turned upward, or are bushy (fasciculate), with or without interconnecting tubules between adjacent corallites (Fig. 2*a*). Most have an external wall, the holotheca, which limits their ability to cement themselves to hard surfaces. Colonies are generally 30–600 mm (1.2–24 in.) in diameter, but massive colonies may reach 2 m (3.3 ft) and some fasciculate colonies may be even larger.

Tabulate corals were most abundant and diverse in temperate to warm shelf environments, particularly in biostromes and bioherms. Smaller, laminar to domal massive colonies inhabited deeper, cooler waters. They were an important component of true reefs, particularly in back-reef, reef-flat and fore-reef environments, but their limited ability of secure attachment to hard surfaces restricted their contribution to reef framework. *See* ANTHOZOA; HEXACORALLIA; RUGOSA; SCLERACTINIA. Colin Scrutton

Bibliography. D. Hill, Rugosa and Tabulata, in C. Teichert (ed.), *Treatise on Invertebrate Paleontology*, pt. F: *Coelenterata*, suppl. 1, vols. 1, 2, 1981; W. A. Oliver, Jr., Origins and relationships of Paleozoic coral groups and the origin of the Scleractinia, in G. D. Stanley, Jr. (ed.), Paleobiology and biology of corals, *Paleontol. Soc. Pap.*, 1:107–134, 1996; C. T. Scrutton, Corals and other Cnidaria, in R. C. Selley, L. R. M. Cocks, and I. R. Plimer (eds.), *Encyclopedia of Geology*, vol. 2, pp. 321–334, Elsevier, Oxford, 2004; C. T. Scrutton, The Palaeozoic corals, *Proc. Yorkshire Geol. Soc.*, 51:177–208, 1997, and 52:1–57, 1998; J. E. Sorauf, Biocrystallization models and skeletal structure of Phanerozoic corals, in G. D. Stanley, Jr. (ed.), Paleobiology and biology of corals, *Paleontol. Soc. Pap.*, 1:159–185, 1996.

Tacan

A member of the rho-theta family of air navigation systems which define an aircraft's position by its distance and bearing to a single beacon. Such systems inherently answer the navigator's questions of "in what direction" and "how far," without additional computation, and are of particular value when the beacon has to be placed on a ship, oil-drilling rig, or small island. The main weakness of such systems is that bearing errors cause spatial error to increase with distance from the beacon. Major attention is therefore given to the reduction of bearing errors.

Principles of operation. Tacan, first proposed in 1947, is the major survivor of a number of military proposals following World War II. It allows the distance-measuring equipment (DME) to provide bearing service also, without the large antennas or site errors characteristic of the civil very high-frequency omnidirectional range (VOR). Range and accuracy are the same as DME (300 mi or 480 km, and 0.1 mi or 0.16 km, respectively), with a bearing accuracy of 1°. As in DME, operation is on 252 channels, spaced 1 MHz apart, 962–1213 MHz.

To provide the added bearing service, the DME transponder is first arranged to operate at constant duty cycle. This means that the number of output pulses is held constant, whether the beacon is being interrogated by one or a hundred aircraft. When few interrogations are present, the balance is made up of so-called squitter pulses (in early sets these were obtained by merely allowing receiver gain to increase until noise pulses reached a preassigned level).

The total output of the transponder is amplitude-modulated by the rotating directional antenna system (**Fig. 1**). At the center of this system is the central radiator connected to the DME transponder, just as in the conventional DME. However, rotating around this radiation at 15 revolutions per second

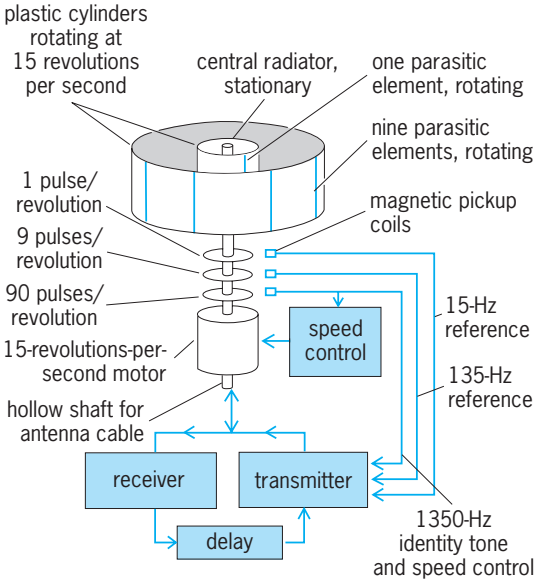

Fig. 1. Tacan transponder with rotating directional antenna system.

are two concentric dielectric cylinders. The inside one, about 6 in. (15 cm) in diameter, contains a single parasitic reflector which imparts a 15-Hz amplitude modulation to the DME replies, and the outside one, about 33 in. (84 cm) in diameter, contains nine parasitic elements which impart a 135-Hz amplitude modulation. On the same rotating shaft are mounted reference pulse generators which additionally modulate the transmitter with coded pulses, once per revolution for the 15-Hz signal (called the north reference burst) and nine times per revolution for the 135-Hz signal (called the auxiliary reference bursts). The composite radiated signal therefore looks like **Fig. 2**. To compensate for the reduced signal amplitude in the "valleys," Tacan beacons are usually of 5 kW power or more.

In the airborne receiver the 15- and 135-Hz sine waves are detected and filtered and compared with the decoded reference bursts to provide a two-speed or fine-coarse bearing display whose accuracy and site freedom are superior to those of conventional VOR, yet with a ground antenna system which is small enough to mount on a ship's mast. This is of particular interest to aircraft based on aircraft carriers, but is also of value in portable transponders.

Site freedom. Tacan's site freedom derives from three factors: (1) Its horizontal antenna aperture is about three wavelengths, as compared with a half wavelength of the standard VOR. (The Doppler VOR, with five wavelengths, is better in this respect, but requires a 150-ft- or 45-m-diameter antenna counterpoise.) (2) Because of Tacan's higher frequency, vertical antenna aperture of up to 12 elements is feasible, leading to reduced radiation into the ground and consequently less site error. Such vertical apertures have been found to be impractical at lower frequencies. (3) Whatever error remains is then divided by 9 by the airborne coarse-fine instrumentation.

Improvements. Major improvements in Tacan for both airborne and ground equipment include the use of all-solid-state components, even at the 5-kW radio-frequency level; replacement of analog signal-processing circuits by digital, primarily to reduce the need for adjustments during manufacture and maintenance; fixed ground equipment monitored from remote service centers to reduce maintenance calls; and reduction in sizes and weights along with increased reliability. Thus, a Vortac station has 5 cabinets, compared to 17 in 1980, and requires a service call once in 3 months instead of weekly.

Installation. About 30,000 military aircraft, chiefly in the NATO countries, are Tacan-equipped. These receivers, of course, also receive distance-only service from civil DME transponders.

In the United States and other countries having a common civil and military airways system, VOR, DME, and Tacan are combined into the Vortac system (**Fig. 3**). Here the VOR and Tacan stations are colocated, one above the other. Civil aircraft obtain their DME service from the Tacan transponder. There are about 750 such stations in the United States.

It might be asked why the VOR is necessary if Tacan provides the same service. The reasons are that

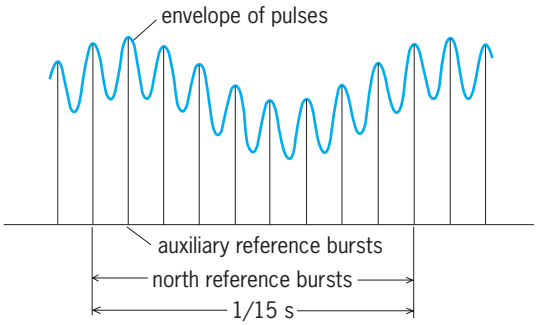

Fig. 2. Composite radiated signal of Tacan transponder. Spaces between reference burst are filled with 2700 random DME replies per second.

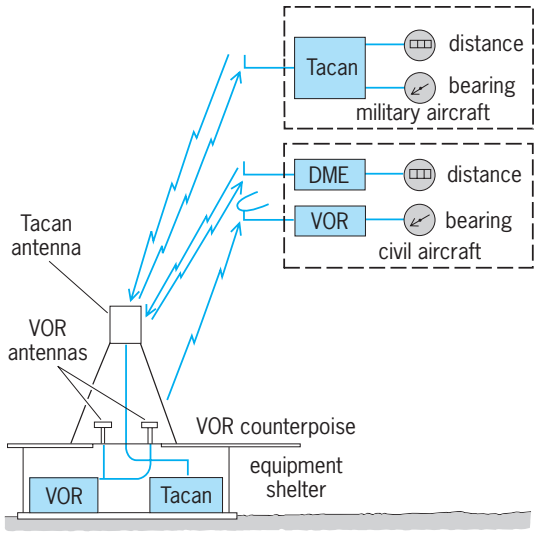

Fig. 3. Vortac system.

VOR preceded Tacan by about a decade and that the VOR receiver, by itself, not only is common to the instrument landing system (ILS) localizer, but also has inherently low cost. There are about 100,000 general-aviation aircraft in the world which carry VOR receivers, but no DME or Tacan. *See* DISTANCE-MEASURING EQUIPMENT; DOPPLER VOR; ELECTRONIC NAVIGATION SYSTEMS; INSTRUMENT LANDING SYSTEM (ILS); RHO-THETA SYSTEM; VOR (VHF OMNIDIRECTIONAL RANGE).

Sven H. Dodington

Bibliography. B. E. Bjerede, A unified signal processor for Tacan navigation sets, *Navigation*, 23:119–127, 1976; M. Kayton and W. Fried, *Avionics Navigation Systems*, 2d ed., Wiley, 1997; A. H. Lang, Second-generation Vortac equipment, *Elect. Commun.*, 58(3):256–262, 1984; P. C. Sandretto, *Electronic Avigation Engineering*, 1958; G. J. Sonnenberg, *Radar and Electronic Navigation*, 6th ed., 1988.

Tachometer

An instrument that measures angular speed, as that of a rotating shaft. The measurement may be in revolutions over an independently measured time interval,

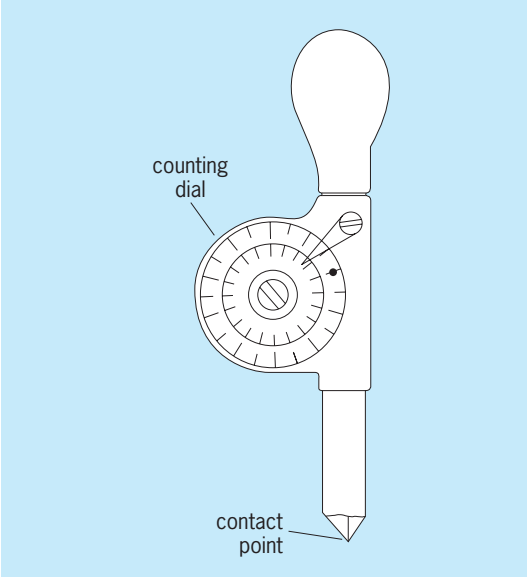

Fig. 1. Revolution counter, the simplest form of tachometer. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

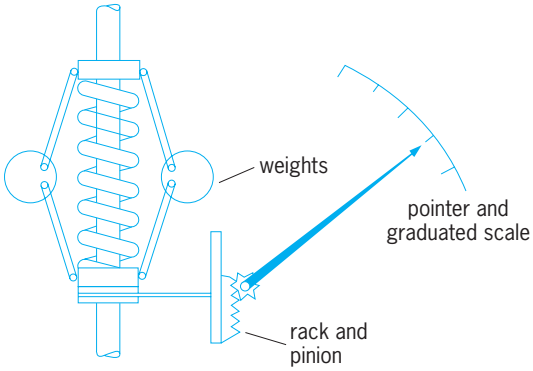

Fig. 2. Centrifugal-force tachometer indicates instantaneous speed. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

as in a revolution counter, or it may be directly in revolutions per minute. The instrument may also indicate the average speed over a time interval or the instantaneous speed. Tachometers are used for direct measurement of angular speed and as elements of control systems to furnish a signal as a function of angular speed.

Revolution counter. The simplest form of tachometer is the revolution counter shown in **Fig. 1**. Held in contact with the rotating shaft, it counts the number of shaft rotations. It is used in conjunction with a timer to obtain average speed over the elapsed period of time. Accuracy is best for uniform speeds measured over long time periods.

Chronometric tachometer. This tachometer counts the revolutions over a fixed interval of time and presents the measurement directly in terms of speed. The result is the average speed for the time interval. The shorter the time interval, the more nearly the instantaneous speed will be approached.

Centrifugal tachometer. The centrifugal force developed by a rotating mass is a function of speed. **Figure 2** shows the essential parts of such a meter. This form indicates instantaneous speed. It is capable of accuracy on the order of ±1% of full-scale value.

Vibrating-reed tachometer. The vibrating-reed tachometer consists of a group of reeds of different length. The lowest natural frequency of vibration of each reed is a function of its length, mass, and cross-sectional dimensions. Observation of the reed that is vibrating forms the means of measuring the frequency of vibration (**Fig. 3**). The instrument is brought into mechanical contact with the device whose mechanical frequency, oscillation, or rotation is to be measured by touching it to the bearing support, case, or frame. The reeds may be adjusted to an accuracy of ±0.3% and are usually guaranteed to ±0.5%.

Impulse tachometer. Tachometers falling into this group may be classified as capacitor charging-current type, inductor type, or interrupted dc type.

Capacitor charging-current tachometer. In the instrument shown in **Fig. 4**, the charging current of a capacitor is utilized. The pickup head usually contains a reversing switch, operated from a spindle, which reverses twice with each revolution. The indicator

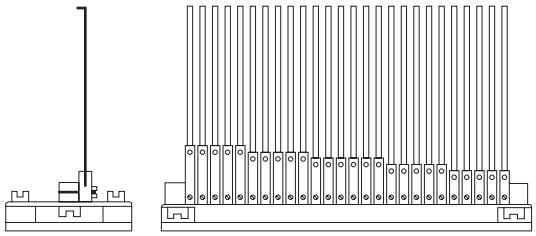

Fig. 3. Vibrating-reed tachometer consists of reeds of varying lengths. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, 2d ed., McGraw-Hill, 1974*)

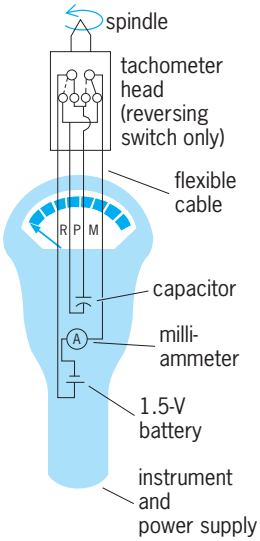

Fig. 4. Capacitor-type impulse tachometer. (*After G. K. McMillan and D. M. Considine, eds., Process/Industrial Instruments and Controls Handbook, 5th ed., McGraw-Hill, 1999*)

responds to the average value of these pulses. The indication is proportional to time rate of these pulses and therefore to the time rate of the spindle revolutions. The battery voltage must be steady and the circuit must be adjusted to the actual value of this voltage. With a steady voltage, this device is capable of good accuracy.

Inductor tachometer. With this instrument the rotating member, which could be a piece of soft iron or laminated iron, causes the magnetic flux of a circuit, containing a magnet and pickup coil, to rise and fall. The rise in flux produces a pulse of one polarity and the fall of flux produces a pulse of the opposite polarity. This is then rectified for a permanent-magnet movable-coil instrument. The pulses may be produced by a lump of magnetic material passing close to a coil having a permanent-magnet core. **Figure 5** shows a form of this tachometer in which the soft-iron piece, or magnet, rotates close to the pickup coil.

Interrupted dc tachometer. The interrupted dc of an ignition-circuit primary, whether battery-excited or magneto-excited, provides a frequency of pulses which can be used to measure speed. **Figure 6** shows a frequency-responsive circuit, in which current from the battery is interrupted by the contactor and excites the ignition coil. The voltage drops from these pulses excite the saturable-core transformer to produce current, which is rectified for the average-reading dc instrument.

Eddy-current tachometer. This form, also known as drag type, is widely used for automobile speedometers and for measuring aircraft-engine speed. It contains a permanent magnet, rotated by the shaft

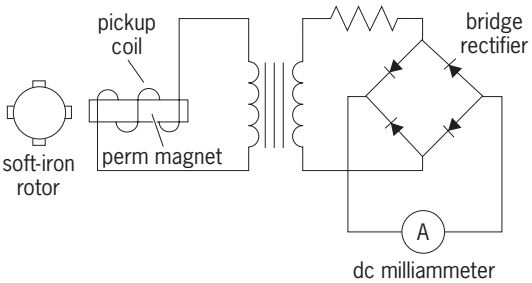

Fig. 5. Circuit for an inductor form of tachometer.

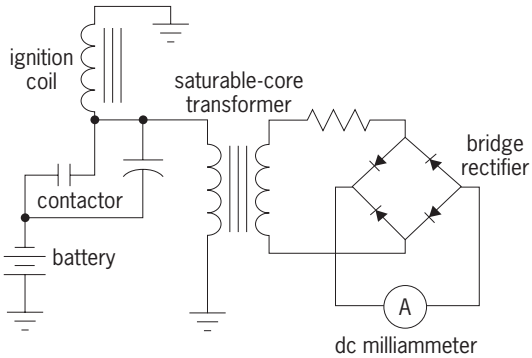

Fig. 6. Schematic circuit of an interrupted dc pulse type of tachometer, as used with the ignition circuit of internal combustion engine.

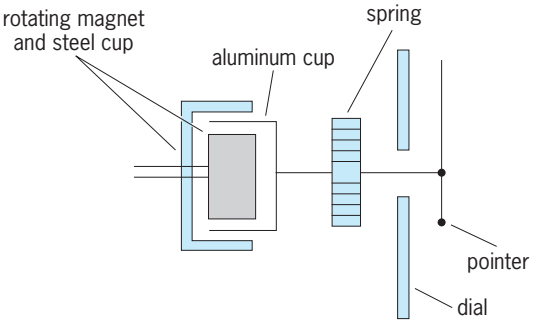

Fig. 7. Drag-type, or eddy-current, tachometer. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, 2d ed., McGraw-Hill, 1974*)

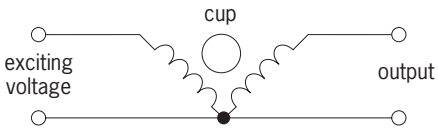

Fig. 8. Schematic circuit and components of a drag-cup tachometer generator.

whose speed is to be measured. Close to the revolving magnet is an aluminum disk or cup, mounted to a staff with a pointer, pivoted and free to turn against a spring (**Fig. 7**). The pointer is associated with a calibrated scale. As the permanent magnet revolves, eddy currents are produced in the disk or cup. The magnetic fields caused by these eddy currents produce a torque, which acts in a direction to resist the turning magnet field. The cup or disk will then turn against the spring, in the direction of the rotating-magnet field, and will turn (or be dragged) until the torque developed by the eddy currents equals that of the spring. An accuracy of 10 revolutions/min in a full scale of 3000 revolutions/min may be obtained. The movable member may be revolved as much as 1080°, affording high resolution in indication.

Drag-cup generators, which also use eddy currents, are used in control systems. They have two stationary windings, positioned so as to have zero coupling, and a nonmagnetic metal cup, which is revolved by the source whose speed is to be measured (**Fig. 8**). One of these windings is used for excitation, inducing eddy currents in the rotating cup.

These eddy currents in the cup produce a field which induces in the other winding an emf proportional to the speed of the rotating cup and at the same frequency as the exciting source. These generators usually have a high degree of linearity, low electrical noise, and low starting and running torque. The output emf and energy are also low.

Velocity-head tachometer. With this type of tachometer the device whose speed is to be measured drives a pump or blower, producing a fluid flow, which is converted to a pressure. **Figure 9** shows a form of this tachometer that not only indicates the speed but produces a tape recording. This form has been used on railroad locomotives. The tape is moved by the drive shaft whose speed is being measured.

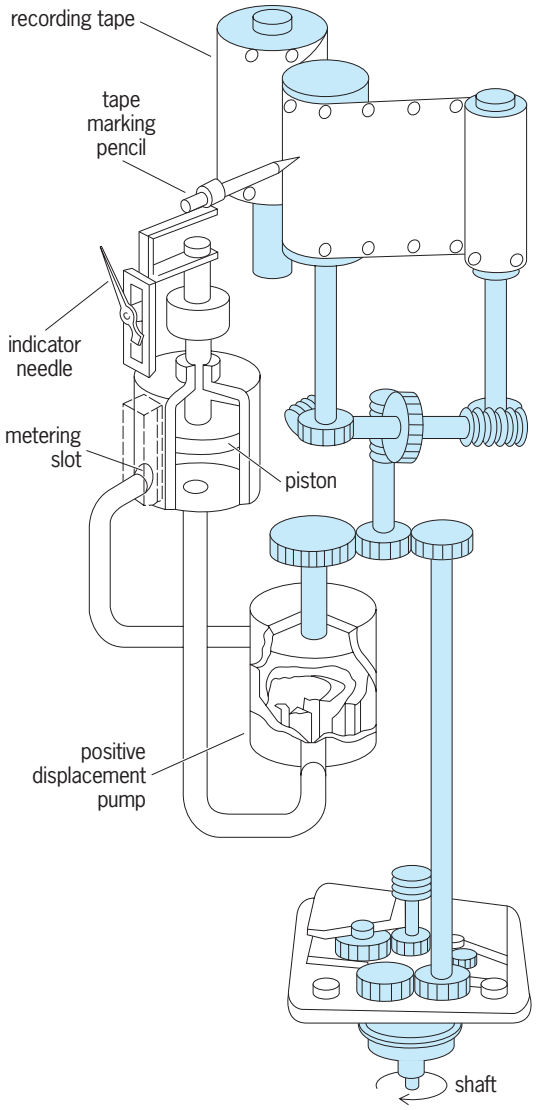

Fig. 9. Sectional view of velocity-head-type tachometer. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

Electronic tachometer. Circuits including vacuum tubes or transistors are sometimes needed to amplify weak pulses or to shape pulse waves. These pulses may be produced by photoelectric, magnetic, or inductor transducers. These shaped waves are applied to frequency-responsive circuits, are rectified, and then are applied to a dc milliammeter.

Electronic counters, also known as EPUT (events per unit of time) meters, are designed to measure frequency, whether sine-wave or pulse. The input pulse waves, clipped and shaped by a discriminating circuit, are allowed to pass through a gating circuit to the displaying decade counters. The open time of the gate is controlled by a crystal oscillator or other suitable time base. The displaying counters are responsive to the number of pulses passed by the gate in a definite time. The accuracy, established by the time-base generator, can be better than 0.01%. Since the display counters or readout may be read to ± 1 count, the number of pulses applied to the input should be sufficiently high to realize its inherent accuracy. *See* FREQUENCY COUNTER.

Electric-generator tachometer. A widely used and flexible form of electric tachometer comprises a combination of electric generator and indicator. There are two principal forms: a dc generator with a dc voltmeter, and an ac generator with an ac voltmeter (or dc voltmeter with rectifier). In either case, the emf developed is proportional to the shaft speed.

The ac generator form may include a circuit that is responsive to frequency and affected only slightly by voltage, giving greater accuracy.

Direct-current generator. The dc tachometer is a small permanent-magnet generator with an output of 2–10 V per 1000 revolutions/min. A high-resistance voltmeter, calibrated in revolutions per minute, indicates the speed.

The dc generator assumes a polarity dependent upon direction of rotation. When used with a zero-center indicator, the instrument will indicate the direction of rotation. With standardizing, an accuracy in the order of 0.25–0.1% may be realized. *See* DIRECT-CURRENT GENERATOR.

The low starting and running torque makes it useful for measuring wind velocity. Besides measuring speed, the dc tachometer is used as a stabilizing component in velocity servomechanisms. *See* CONTROL SYSTEMS.

Alternating-current generator. The ac tachometer generator can be constructed with a stationary winding and a revolving permanent-magnet field. Both generated voltage and frequency are proportional to speed of rotation. The voltage may be rectified and applied to a permanent-magnet moving-coil instrument calibrated in revolutions per minute. *See* ALTERNATING-CURRENT GENERATOR.

By the addition of a saturable-core transformer to the instrument circuit, as shown in **Fig. 10**, the instrument indication becomes frequency-responsive and only slightly affected by voltage, affording greater accuracy.

The difference of two speeds may be measured or indicated by connecting the outputs of two speed-measuring circuits to a differential bridge as shown in **Fig. 11**. The individual speeds may also be indicated.

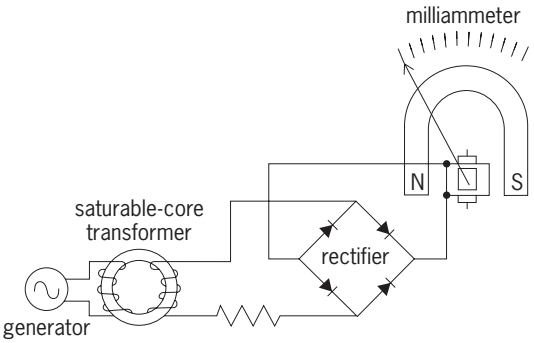

Fig. 10. Schematic circuit of an ac tachometer to which a saturable-core transformer has been added to increase accuracy. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

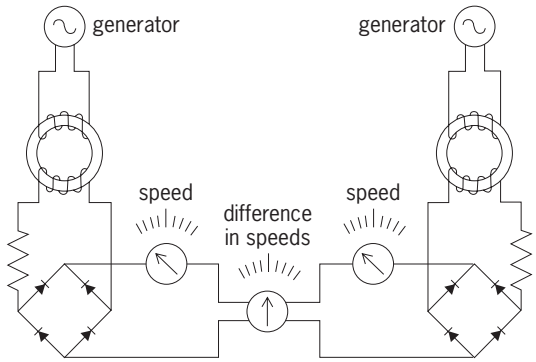

Fig. 11. System employing two ac electric-generator tachometers for indicating individual and differential speeds. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

This difference of speed indication is independent of the actual speeds.

If one of the generators is replaced by a fixed frequency, the other speed may be measured in reference to this frequency, or the scale may be calibrated in terms of the speed. This circuit also affords the means for suppressing as much as 90% of a speed range and allowing the top 10% to occupy the entire scale. Alfred H. Wolferz

Bibliography. D. M. Considine and S. D. Ross (eds.), *Handbook of Applied Instrumentation*, 1964; reprint 1982; G. K. McMillan and D. M. Considine (eds.), *Process/Industrial Instruments and Controls Handbook*, 5th ed., 1999; P. H. Sydenham (ed.), *Handbook of Measurement Science*, vol. 2: *Practical Fundamentals*, 1983.

Tachyon

A hypothetical faster-than-light particle consistent with the special theory of relativity. According to this theory, a free particle has an energy E and a momentum $\mathbf{p}$ which form a Lorentz four-vector. The length of this vector is a scalar, having the same value in all inertial reference frames. One writes Eq. (1), where

$$E^2 - c^2p^2 = m^2c^4 \tag{1}$$

c is the speed of light and the parameter m^2 is a property of the particle, independent of its momentum and energy. Three cases may be considered: m^2 may be positive, zero, or negative. The case $m^2 > 0$ applies for atoms, nuclei, and the macroscopic objects of everyday experience. The positive root m is called the restmass. If $m^2 = 0$, the particle is called massless. A few of these are known: the electron neutrino, the muon neutrino, the tau neutrino, the photon, and the graviton. The third case, $m^2 < 0$, was studied originally by S. Tanaka and by O. M. P. Bilaniuk, V. K. Deshpande, and E. C. G. Sudarshan. Further contributions were made by G. Feinberg, who gave the name tachyons (after a Greek word for swift) to the particles with $m^2 < 0$. Whether such particles exist is an interesting speculation, but there has been no experimental evidence for them.

In general, the particle speed is given by Eq. (2).

$$v = \frac{cp}{E}c \tag{2}$$

If $m^2 > 0$, Eq. (1) implies $E > cp$ and Eq. (2) gives $v < c$. If $m^2 = 0$, then $E = cp$ and $v = c$. In case $m^2 < 0$, one finds $E < cp$ and $v > c$. Tachyons exist only at faster-than-light speeds.

The quantities p or E can be eliminated from Eqs. (1) and (2) to get expressions for E and p in terms of the speed. In case $m^2 > 0$, the familiar results are given by Eqs. (3), where the radical signs

$$E = \frac{mc^2}{\sqrt{1 - \dfrac{v^2}{c^2}}} \qquad p = \frac{mv}{\sqrt{1 - \dfrac{v^2}{c^2}}} \tag{3}$$

imply the positive roots. The way to make the analogous formulas for tachyons is to introduce the positive number μ such that $m^2 = -\mu^2$ and Eq. (4) holds. Then Eqs. (2) and (4) give Eqs. (5). It is seen that for

$$E^2 - c^2p^2 = -\mu^2c^4 \tag{4}$$

$$E = \frac{\mu c^2}{\sqrt{\dfrac{v^2}{c^2} - 1}} \qquad p = \frac{\mu v}{\sqrt{\dfrac{v^2}{c^2} - 1}} \tag{5}$$

ordinary particles as v increases, E increases, but to speed them up to $v = c$ would involve an infinite amount of energy. In contrast for tachyons, as v decreases, E increases, but to slow them down to $v = c$ would involve an infinite amount of energy.

For an ordinary free particle with $m^2 > 0$, there is always a special reference frame, called the rest frame, in which $\mathbf{p} = 0$. From Eq. (1) it is seen that the energy there has the minimum value mc^2. No such special frame exists for massless particles. For tachyons also the situation is quite different. For a free tachyon a special frame can be found in which $E = 0$ and, according to Eq. (4), p has the minimum value μc. Since the tachyons may exist at zero energy, there is no energy obstacle in creating them in elementary particle reactions.

According to electromagnetic theory, a charged particle moving at a speed greater than the speed of light in a medium emits light, the Cerenkov radiation. If charged tachyons existed, they would spontaneously radiate light even in a vacuum. *See* CERENKOV RADIATION.

Attempts to detect tachyons have been made by looking for the Cerenkov radiation (T. Alväger and M. N. Kreisler) and by analyzing for negative m^2 values in elementary particle reactions (C. Baltay, R. Linsker, N. K. Yeh, and G. Feinberg). The conclusions were negative, and present indications are that this type of particle does not exist. *See* ELEMENTARY PARTICLE; RELATIVITY. Roland H. Good, Jr.

Bibliography. O. M. P. Bilaniuk, V. K. Deshpande, and E. C. G. Sudarshan, Meta relativity, *Amer. J. Phys.*, 30:718–723, 1962; G. Feinberg, Particles that go faster than light, *Sci. Amer.*, 222(2):68–73, February 1970.

Taconite

The name given to the siliceous iron formation from which the high-grade iron ores of the Lake Superior district have been derived. It consists chiefly of fine-grained silica mixed with magnetite and hematite. As the richer iron ores approach exhaustion in the United States, taconite becomes more important as a source of iron. To recover the ore mineral in a usable form for the production of iron, taconite must be finely ground, and the magnetite or hematite concentrated by a magnetic or other process. Finally, the concentrate must be agglomerated into chunks of size and strength suitable for the blast furnace. *See* IRON METALLURGY; ORE AND MINERAL DEPOSITS; ORE DRESSING.

Cornelius S. Hurlbut, Jr.

Taeniodonta

An extinct order of quadrupedal eutherian land mammals known from the early Cenozoic deposits of the Rocky Mountain intermontane basins of western North America and, based on a single tooth of *Ectoganus gliriformis*, from early Cenozoic rocks in South Carolina. The nine known genera of taeniodonts are classified into two families: (1) the medium-size (5–15 kg or 11–33 lb), relatively primitive, omnivorous conoryctids (*Schochia, Onychodectes*, and *Conoryctella*, early Paleocene; *Huerfanodon* and *Conoryctes*, middle Paleocene); and (2) the larger (15–110 kg or 33–243 lb) and more advanced stylinodontids (*Wortmania*, early Paleocene; *Psittacotherium*, middle Paleocene; *Ectoganus*, late Paleocene to early Eocene; and *Stylinodon*, middle Eocene).

Conoryctids developed enlarged canines, but the lower jaws were unspecialized and the cheek teeth were low-crowned, enamel-enclosed, and cuspidate. Stylinodontid taeniodonts developed deep massive jaws, peglike teeth that were ever-growing, large curved canines bearing enamel bands, and large laterally compressed and recurved claws on the front paws. The most derived taeniodonts, such as *Stylinodon mirus*, were apparently active diggers, rooters, and grubbers living in warm temperate to subtropical environments. In terms of modern analogs, an advanced stylinodontid may be thought of as an aardvark with the head of a pig.

Taeniodonts were never common members of the faunas in which they occurred; they reached their maximum species diversity during the early and middle Paleocene and then persisted at low diversity and absolute numbers into the middle Eocene. *See* EUTHERIA.

Robert M. Schoch

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, W. H. Freeman, 1988; C. M. Janis, K. M. Scott, and L. L. Jacobs (eds.), *Evolution of Tertiary Mammals of North America*, vol. 1: *Terrestrial Carnivores, Ungulates, and Ungulatelike Mammals*, Cambridge University Press, 1998; R. M. Schoch, *Systematics, Functional Morphology and Macroevolution of the Extinct Mammalian Order Taeniodonta*, Yale Univ. Peabody Mus. Nat. Hist. Bull. 42, 1986.

Taiga

A zone of forest vegetation encircling the Northern Hemisphere between the arctic-subarctic tundras in the north and the steppes, hardwood forests, and prairies in the south. The chief characteristic of the taiga is the prevalence of forests dominated by conifers. The taiga varies considerably in tree species from one major geographical region to another, and within regions there are distinct latitudinal subzones. The dominant trees are particular species of spruce, pine, fir, and larch. Other conifers, such as hemlock, white cedar, and juniper, occur locally, and the broad-leaved deciduous trees, birch and poplar, are common associates in the southern taiga regions. Taiga is a Siberian word, equivalent to "boreal forest." *See* FOREST AND FORESTRY; TUNDRA.

Climate. The northern and southern boundaries of the taiga are determined by climatic factors, of which temperature is most important. However, aridity controls the forest-steppe boundary in central Canada and western Siberia. In North America there is a broad coincidence between the northern and southern limits of the taiga and the mean summer and winter positions of the arctic air mass. In the taiga the average temperature in the warmest month, July, is greater than 50°F (10°C), distinguishing it from the forest-tundra and tundra to the north; however, less than four of the summer months have averages above 50°F (10°C), in contrast to the summers of the deciduous forest farther south, which are longer and warmer. Taiga winters are long, snowy, and cold—the coldest month has an average temperature below 32°F (0°C). Permafrost occurs in the northern taiga. It is important to note that climate is as significant as vegetation in defining taiga. Thus, many of the world's conifer forests, such as those of the American Pacific Northwest, are excluded from the taiga by their high precipitation and mild winters.

Subzones. The taiga can be divided into three subzones (see **illus.**) in most of the regions which it occupies; these divisions are recognized mainly by the particular structure of the forests rather than by changes in tree composition. These subdivisions are the northern taiga, the middle taiga, and the southern taiga.

Northern taiga. This subzone is characterized on moderately drained uplands by open-canopy forests, dominated in Alaska, Canada, and Europe by spruce and in Siberia by spruce, larch, and pine. The well-spaced trees and low ground vegetation, usually rich lichen carpets and low heathy shrubs, yield a beautiful parkland landscape; this is exemplified in North America by the taiga of Labrador and Northern Quebec. This subzone is seldom reached by roads and railways, in part because the trees seldom exceed 30 ft (9 m) in height and have limited commercial value. These forests are important as winter range of Barren Ground caribou, but in many parts of the

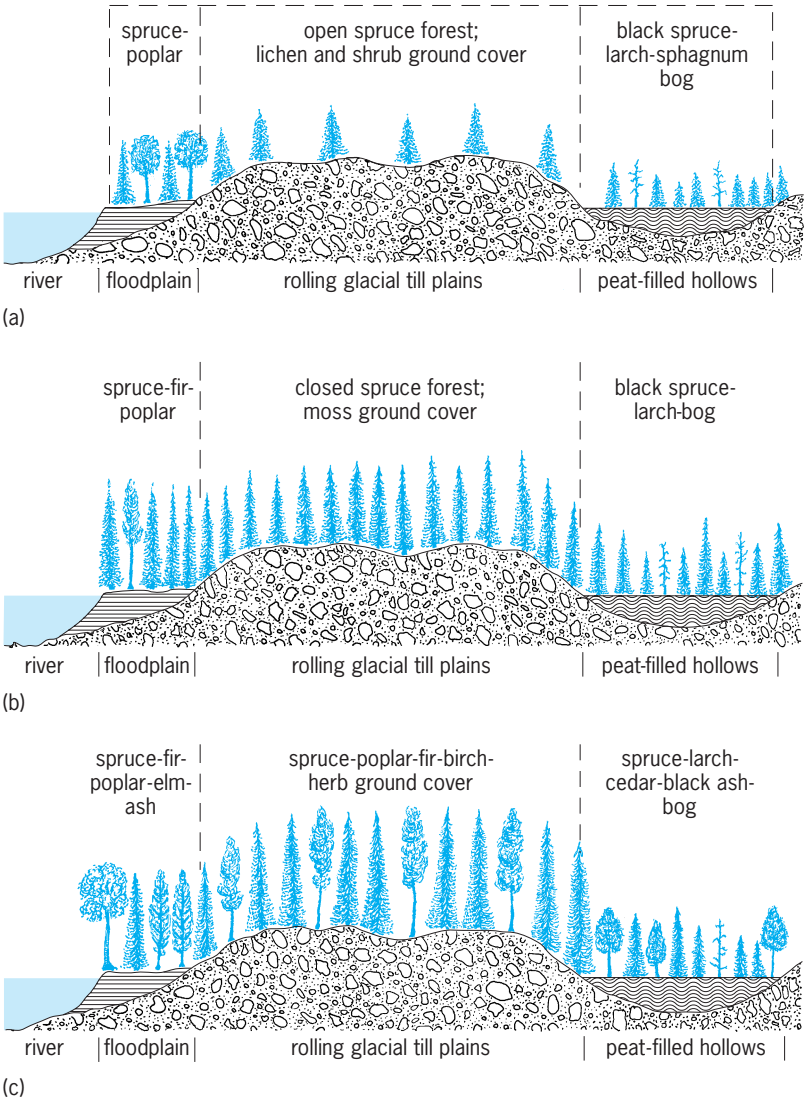

Schematic profile of the three main subzones of the North American taiga, showing the main forest assemblages on three of the more important landform types. (*a*) Northern taiga. (*b*) Middle taiga. (*c*) Southern taiga.

drier interior of North America their area has been decreased by fire, started both by lightning and by humans.

Middle taiga. This subzone is a broad belt of closed-canopy evergreen forests on uplands. The dark, somber continuity is broken only where fires, common in the drier interiors of the continents, have given temporary advantage to the rapid colonizers, pine, paper birch, and aspen poplar. The deeply shaded interior of mature white and black spruce forests in the middle taiga of Alaska and Canada permits the growth of few herbs and shrubs; the ground is mantled by a dense carpet of mosses. Here, as elsewhere in the taiga, depressions are filled by peat bogs, dominated in North America by black spruce and in Eurasia by pine. Everywhere there is a thick carpet of sphagnum moss associated with such heath shrubs as bog cranberry, Labrador tea, and leatherleaf. Alluvial sites bear a well-grown forest yielding merchantable timber, with fir, white spruce, and black poplar as the chief trees.

Southern taiga. This subzone is characterized on moderately drained soils throughout the Northern Hemisphere by well-grown trees (mature specimens up to 95 ft or 29 m) of spruce, fir, pine, birch, and poplar. These trees are represented by different species in North America, Europe, and Siberia. Of the three taiga zones, this has been exploited and disturbed to the greatest extent by humans, and relatively few extensive, mature, and virgin stands remain. In northwestern Europe this subzone has been subject to intensive silviculture for several decades and yields forests rich in timber and pulpwood. In Alaska and Canada the forests have a much shorter history of forest management, but they yield rich resources for the forest industries. *See* FOREST MANAGEMENT.

Fauna. In addition to caribou, the taiga forms the core area for the natural ranges of black bear, moose, wolverine, marten, timber wolf, fox, mink, otter, muskrat, and beaver. The southern fringes of the taiga are used for recreation. J. C. Ritchie

Bibliography. B. V. Barnes and S. H. Spurr, *Forest Ecology*, 4th ed., 1998; A. Bryson, *Geographical Bulletin*, vol. 8, 1966; *The New Oxford Atlas*, 1975.

Tail assembly

An assembly at the rear of an airplane, consisting of the tail cone, the horizontal tail, and one or more vertical tails.

Normal configuration. The tail assembly, or empennage, of an airplane is normally composed of a vertical tail and a horizontal tail attached to the rear, or tail cone, of the airplane's fuselage. The vertical tail is composed of the vertical stabilizer and the rudder (**illus.** *a*). The vertical stabilizer is attached rigidly to the fuselage and is intended to provide stability about a vertical axis through the airplane's center of gravity. The rudder is attached by hinges to the rear of the vertical stabilizer and can rotate from side to side in response to pilot control input. It also contributes to stability, but its main function is to provide a yawing moment about the airplane's vertical (yaw) axis, thereby causing the airplane to yaw (turn) to the left or right. *See* AIRCRAFT RUDDER; FLIGHT CONTROLS; FUSELAGE; STABILIZER (AIRCRAFT).

The horizontal tail, similar to the vertical tail, is composed of the horizontal stabilizer and the elevator. The horizontal stabilizer is fixed rigidly to the fuselage and provides stability about a horizontal axis directed along the wing and through the center of gravity, and known as the pitch axis. The elevator is hinged to the rear of the horizontal stabilizer and rotates up and down as the pilot moves the control column fore and aft. The elevator also contributes to stability about the pitch axis, but its main purpose is to provide a pitching moment about the pitch axis, which causes the airplane to nose up or down. *See* ELEVATOR (AIRCRAFT).

Airplanes that operate at supersonic speeds usually have the horizontal tail swept back and in one

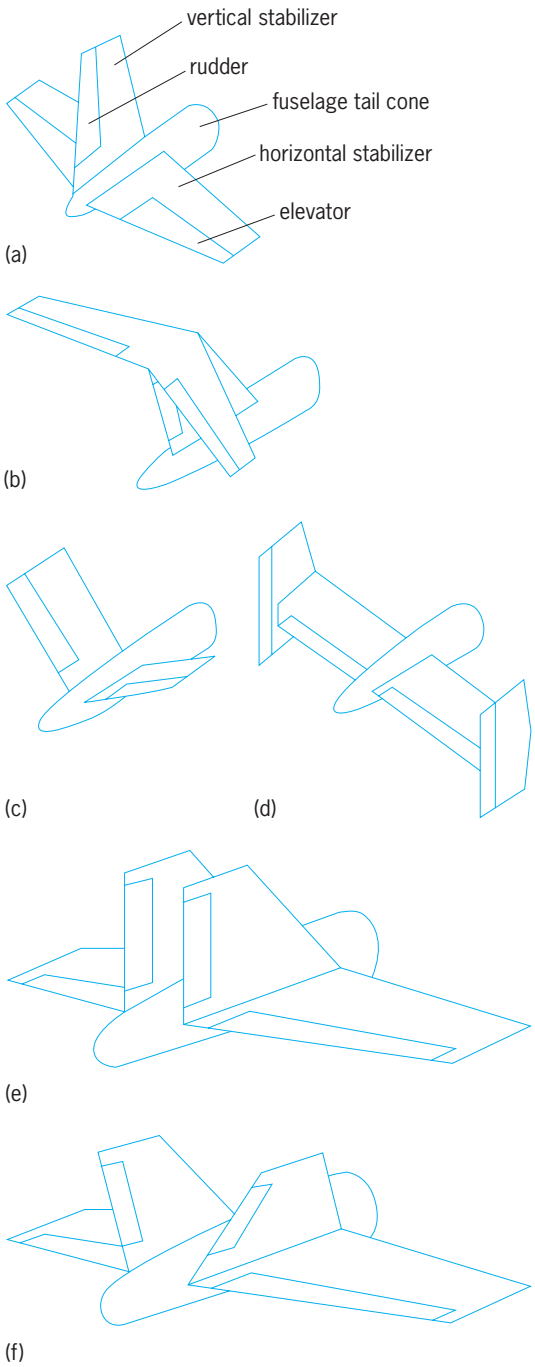

Tail assembly configurations. (*a*) Normal configuration. (*b*) T-tail. (*c*) V-tail. (*d*) Twin-tail. (*e*) Uncanted twin vertical tails. (*f*) Canted twin tails.

piece that is movable and is controlled by the motion of the pilot's control stick. Such a surface is frequently called a stabilator. *See* SUPERSONIC FLIGHT.

Other configurations. Many airplanes employ empennages that depart from the normal configuration.

T-tail. In this configuration (illus. *b*), the horizontal tail is attached to the top of the vertical tail in order to raise it above the wake from the wing. Most jet transports employ this type of tail. The T-tail is an extreme example of the more general case where the horizontal tail may be placed vertically at any location along the vertical tail. Also, when the horizontal tail is attached to the fuselage, its location may be ahead of, in line with, or behind the vertical tail. The horizontal placement of the horizontal tail is frequently done to improve the spin-recovery characteristics of the airplane.

V-tail. This tail (illus. *c*) has been used sparingly in the past and its best-recognized application is in the Beechcraft Bonanza, a single-engine, piston-powered, four-place airplane. The single tail, shaped like a V, provides aerodynamic forces about both the yaw and pitch axes, thus providing the functions of both the horizontal and vertical tail surfaces. Pitch control is accomplished by control surfaces which, when moved up or down together, provide a pitching moment in a manner similar to the conventional elevator. If these control surfaces are moved differentially, that is, one up and the other down, then a yawing moment results similar to that produced by a rudder.

Multiple vertical tails. It is not uncommon for airplanes to have more than one vertical tail. An empennage with two vertical tails is referred to as a twin tail (illus. *d*). The purpose of having two or more vertical tails may be aerodynamic, but it can also be to avoid the use of one large vertical tail. The Lockheed Constellation, one of the first four-engine, propeller-driven transports of the 1950s, is reputed to have resorted to three vertical tails because the height of a single tail would have exceeded the hangar doors then in use by the airline that first ordered it.

Many modern fighters use two vertical tails mounted near the center just outboard of the fuselage. These tails may be oriented vertically (Fig. 1*e*), as in the case of the F-15 fighter, or they may be canted outward at the top (Fig. 1*f*), as used on the F-14 and F-18 fighters. An appreciable amount of cant is employed in the F-18 tails for the purpose of enhancing the stability and control of the airplane at angles of attack in excess of 50°. *See* AIRFRAME; AIRPLANE; MILITARY AIRCRAFT. Barnes W. McCormick, Jr.

Bibliography. J. D. Anderson, Jr., *Introduction to Flight*, 4th ed., 1999; B. W. McCormick, *Aerodynamics, Aeronautics and Flight Mechanics*, 2d ed., 1994; L. M. Nicolai, *Fundamentals of Aircraft Design*, 1975; D. Stinton, *The Design of the Airplane*, 1983; D. B. Thurston, *Design for Flying*, 2d ed., 1994; E. Torenbeek, *Synthesis of Subsonic Airplane Design*, 1982, reprint 1986.

Takakiales

An order of liverworts in the subclass Jungermanniidae, consisting of a single genus and two species: *Takakia lepidozioides* with the smallest chromosome number, 4, known in bryophytes, and *T. ceratophylla* (5 chromosomes).

Some authors put the Takakiales in the Calobryales owing to branching from a prostrate branched stem, lack of rhizoids, copious mucilage secretion, and massive frequently scattered archegonia that lack protective envelopes. However, sporophytes have

never been seen, and so it is difficult to demonstrate any meaningful relationship.

The members of the Takakiales consist of a very small gametophyte made up of a branched system of prostrate, leafless stolons and erect, radially organized branches with terete appendages. The "leafy" branches are simple or forked and have a weak central strand of narrow cells enclosed by larger cells. The leafy appendages are small and scalelike below, larger and crowded above, and variable in arrangement. They are dissected to the base into two to four segments consisting of one to several central cells and an epidermis of somewhat smaller cells. The leaf cells lack oil bodies. Mucilage is secreted by simple filaments on leafy branches and by branched filaments clustered on both leafy and stoloniform branches. The archegonia are scattered and not enclosed by protective structures, but both antheridia and sporophytes are unknown. *See* BRYOPHYTA; CALOBRYALES; HEPATICOPSIDA. Howard Crum

Talc

A hydrated magnesium layer silicate (phyllosilicate) with composition close to $Mg_3Si_4O_{10}(OH)_2$. Limited substitution of aluminum (Al) or titanium (Ti) for silicon (Si) and of iron (Fe), manganese (Mn), or aluminum for magnesium (Mg) is possible. Steatite is massive talc aggregate that contains less than 1.5% calcium oxide (CaO), 1.5% iron(II) oxide + iron(III) oxide ($FeO + Fe_2O_3$), and 4% aluminum oxide (Al_2O_3). Talc commonly is white, but it may appear pale green or grayish depending on the amount of minor impurities. It usually occurs in coarse to fine platy or fibrous aggregates that often have a more or less parallel arrangement. Shreds and plates are often bent; some tabular crystals exhibit perfect basal cleavage, yielding flexible, slightly elastic lamellae. Talc has a greasy feel and pearly luster, and has been used as one of the hardness standards for rock-forming minerals with the value 1 on the Mohs scale. Because talc is soft, it can be scratched by fingernails.

Talc frequently occurs in magnesium-rich metamorphosed serpentinites and siliceous dolomites. Talc-rich rocks include massive soapstones, massive steatite, and foliated talc schists. In regional and thermal metamorphic terranes, talc is associated with magnesite, dolomite, tremolite, or chlorite in medium-grade rocks, and with enstatite, olivine, or anthophyllite in high-grade rocks. Progressive changes of talc-bearing assemblages with increasing grade of metamorphism toward granitic plutons have been well documented in the Alps and in California. In the upper amphibolite facies, talc dehydrates to enstatite and quartz at temperatures above 1380°F (750°C). At lower temperatures in the greenschist facies, both tremolite and chlorite may be converted to talc by carbon dioxide (CO_2) metasomatism, but at still lower temperatures talc is unstable in the presence of calcium oxide and carbon dioxide, and is replaced by dolomite or magnesite.

Talc is also often a hydrothermal mineral formed at the expense of serpentine and tremolite in shear zones of serpentinites, as lenticular veins, or as actinolite–chlorite–talc rinds around some blocks enclosed within serpentinite. Steatization is commonly associated with serpentinization; the conversion of serpentine to talc may occur by the addition of silica and removal of magnesium, or by the addition of carbon dioxide to form magnesite + talc. Although not abundant in sediments, authigenic talc has been identified in evaporites and in unmetamorphosed carbonates.

Hydrothermal experiments and geologic observations suggest that talc is stable over a wide range of temperatures. In the presence of quartz, talc may form from serpentine at $T < 660°F$ (350°C). It is stable up to 1380°F (750°C), where it dehydrates to enstatite or anthophyllite. The stability field of talc becomes very restricted in environments of low activity of silicon dioxide (SiO_2). A decrease in the activity of silicon dioxide in hydrothermal fluid causes talc to be replaced by serpentine (at low temperatures) or olivine (at high temperatures). Occurrence of talc in serpentinitized ultramafic rocks suggests introduction of silicon dioxide or carbon dioxide during hydrothermal alteration. Talc-bearing assemblages can be used to estimate temperature and fluid composition for thermal or regional metamorphism of siliceous dolomite and serpentinite.

Talc is a good insulating material. It has been commonly used in industry as a raw material for ceramics, paints, plastics, cosmetics, papers, rubber, and many other applications. *See* SILICATE MINERALS. J. G. Liou

Bibliography. S. W. Bailey (ed.), *Hydrous Phyllosilicates (Exclusive of Micas), Reviews in Mineralogy*, vol. 19, 1998; W. A. Deer, R. A. Howie, and J. Zussman, *The Rock-Forming Minerals*, vol. 3: *Sheet Silicates*, 1966; J. J. Hemley et al., Mineral equilibria in the MgO-SiO_2-H_2O system: I. Talc-chrysotile-forsterite-brucite stability relations, *Amer. J. Sci.*, 277:322–351, 1977; K. L. Kimball, F. S. Spear, and H. J. B. Dick, High temperature alteration of abyssal ultramafics from the Islas Orcadas Fracture zone, South Atlantic, *Contrib. Mineral. Petrol.*, 91:307–320, 1985.

Tall oil

A by-product from the pulping of pine wood by the kraft (sulfate) process. In the kraft process the wood is digested under pressure with sodium hydroxide and sodium sulfide. (The alternate process name, sulfate, is derived from the sodium sulfate used to replace sulfur lost in the process.) The volatilized gases are condensed to yield sulfate turpentine. During the pulping the alkaline liquor saponifies fats and converts the fatty and resin acids to sodium salts. Concentration of the pulping solution (black liquor) prior to recovery of the inorganic pulping chemicals allows the insoluble soaps to be skimmed from the surface. Acidification of the skimmed soap yields crude tall oil.

Crude tall oil from southern pines contains 40–60% resin acids (rosin), 40–55% fatty acids, and 5–10% neutral constituents. Abietic and dehydroabietic acids comprise over 60% of the resin acids, while oleic and linoleic acids predominate in the fatty acid fraction. The neutral components of tall oil are principally sterols (sitosterol) and diterpene alcohols and aldehydes. During heating of tall oil, esterification takes place, reducing the acid number.

Tall oil is refined by vacuum distillation. Fractionation of 1.1 tons (1 metric ton) of tall oil yields approximately 660 lb (300 kg) of fatty acids, 770 lb (350 kg) of rosin, and 770 lb (350 kg) of intermediate (distilled tall oil), head, and pitch fractions. United States fractionation capacity is over 1×10^6 tons (9×10^5 metric tons) per year.

Fatty acids from tall oil distillation may contain as much as 10–40% resin acids (distilled tall oil) or as little as 0.5%. About 35% of United States fatty acid production and more than half of the unsaturated fatty acid part come from tall oil. Major uses of tall oil fatty acids as chemical raw materials are in coatings, resins, inks, adhesives, and soaps and detergents, and as flotation agents. Tall oil is an important source of rosin in the United States. *See* PINE TERPENE; ROSIN; WOOD CHEMICALS. Irving S. Goldstein

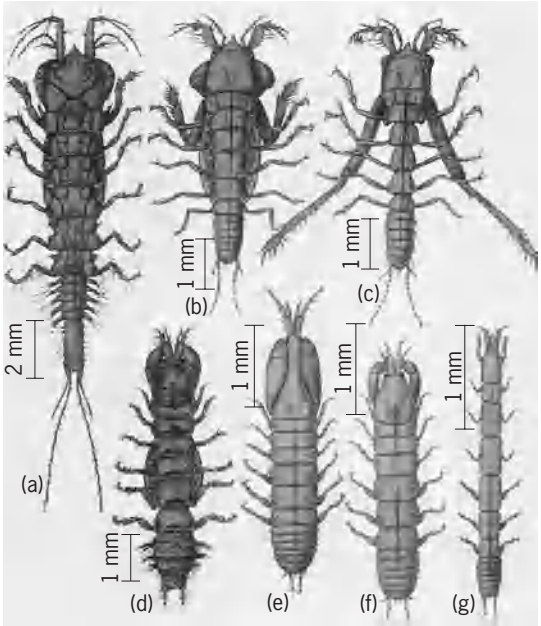

Tanaidacea. (*a*) *Apseudes spinosus*, female; (*b*) *Sphyrapus anomalus*, female; (*c*) *Sphyrapus anomalus*, male; (*d*) *Tanais cavolini*, female; (*e*) *Heterotanais oerstedi*, male; (*f*) *Heterotanais oerstedi*, female; (*g*) *Leptagnathia filiformis*, female.

Tanaidacea

An order of the eumalacostracans of the superorder Percarida, derived from the genus *Tanais*. These animals have a worldwide distribution and with few exceptions are marine. They occur from the shore down to abyssal depths. They are free-living and benthonic. The order is divided into 2 suborders with 5 families, 44 genera, and about 350 species. The body is linear, more or less cylindrical or dorsoventrally depressed (see **illus.**). Thoracid segments 1 and 2 are fused with the head, forming a carapace enclosing a respiratory chamber on each side, and the last abdominal segment is fused with the telson. The left mandible has a lacinia mobilis, while this structure may be present or absent on the right mandible. Eight pairs of thoracic legs are present, of which the first pair are maxillipeds, the second pair chelipeds (the first pair of pereiopods), and the following six pairs pereiopods. The third pair in *Monokonophora*, as a rule, is fossorial, and the chelipeds and first pair of pereiopods usually have vestigial exopodites. Pleopods may be present or absent, and the uropods are filiform.

The nervous system consists of a brain, a subesophageal mass, and a ventral chain. Eye lobes are present and sessile, with or without visual elements. Antennulae, especially in the males, have aesthetes. Sense hairs or sense spines are found on the segments and legs.

In the reproductive organs, the gonads are double. The oviducts open laterally at the base of the fifth pair of pereiopods. The vasa deferentia have a common vesicula seminalis, which is ventromedian on the last thoracic segment. Hermaphroditism, protandry, and protogyny can occur, and sex dimorphism is common. The antennulae are always different in the two sexes. Differences can occur also in the shape of the head, the mouthparts, the chelipeds, the second pair of pereiopods, and the pleopods.

The alimentary canal consists of a ventral mouth, a stomach with a complicated filter and masticatory apparatus, a syncytial midgut, and a terminal anus. As a rule, there are two pairs of hepatopancreas. The excretory organs are a single pair of maxillary glands.

The females produce several broods, each preceded by a molting, by which the exterior morphology may undergo considerable alterations. The eggs develop in an incubatory pouch formed by one or four pairs of oostegites in which the embryos are dorsally flexed. The newly hatched larva lacks the last pair of pereiopods and the pleopods. The young probably undergo four larval (the manca) stages. *See* PERACARIDA; SEXUAL DIMORPHISM. Karl Lang

Bibliography. K. Lang, Neotanaidae nov. fam., with some remarks on the phylogeny of the Tanaidacea, *Arkiv Zool.*, 9:469–475, 1956; K. Lang, The postmarsupial development of the Tanaidacea, *Arkiv Zool.*, 4:409–422, 1953; K. Lang, Protogynie bei zwei Tanaidaceen-Arten, *Arkiv Zool.*, 11:535–540, 1958; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; R. Siewing, Besteht eine engere Verwandschaft zwischen Isopoden und Amphipoden?, *Zool. Anz.*, 147:166–180, 1951; R. Siewing, Morphologische Untersuchungen an Tanaidaceen und Lophogastriden, *Z. Wiss. Zool.*, 157:333–426, 1954; T. Wolff, Crustacea Tanaidacea from depths exceeding 6000 meters, *Galathea Rep.*, 2:187–241, 1956.

Tangent

A term describing a relationship of two figures (usually of the same dimension) in the neighborhood of a common point. The figures are tangent at a point P if they touch at P but do not intersect in a sufficiently small neighborhood of P. To be more precise, if P denotes a point of a curve C (see **illus.**), a line L is a

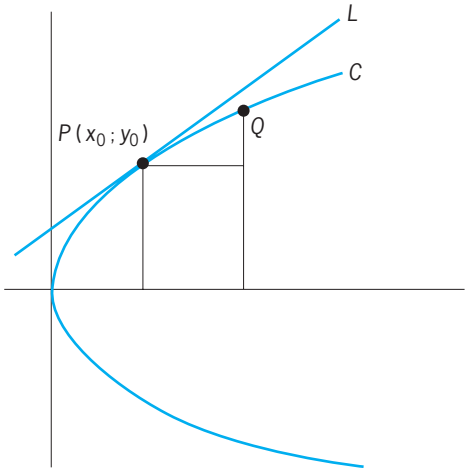

Line *L* tangent to curve *C* at *P*.

tangent to C at P provided L is the limit of lines joining P to a variable point Q of C, as Q approaches P along C (that is, for Q sufficiently close to P, the line PQ is arbitrarily close to L). If curve C has equation $y = f(x)$ and point P on C has coordinates (x_0, y_0), it is shown in the calculus that the slope of the line tangent to C at P is the value of the derivative $f'(x)$ for $x = x_0$. *See* CALCULUS.

Leonard M. Blumenthal

Tangerine

A name applied to certain varieties of a variable group of loose-skinned citrus fruits belonging to the species *Citrus reticulata*. Although mandarin and tangerine are often used interchangeably to designate the whole group, tangerine is applied more strictly to those varieties (cultivars) having deep-orange or scarlet rinds, whereas the term mandarin is more properly used to include all members of this quite variable group of citrus fruits. *See* MANDARIN.

The two principal tangerine varieties are the Dancy, grown mainly in the United States, and the Algerian (Clementine), grown extensively in the Mediterranean countries but also in California, Arizona, and Texas. Robinson and Murcott tangerines are also important in Florida. Trees are medium-sized with rounded tops bearing leaves pointed at the tip, rounded at the base, and having narrowly winged petioles. The flowers are white and small. The fruits are deep orange, loose-skinned, and small to medium-sized, and possess small seeds with green cotyledons. Tangerines are easily peeled and eaten out of hand as fresh fruit. About one-third of the crop is utilized in juice, sherbets, and canned sections.

A number of tangerine hybrids, the result of controlled breeding programs by the U.S. Department of Agriculture and California experiment stations, are delicious fruits. Important among these are the tangelos, which are hybrids of Dancy tangerine by grapefruit. *See* BREEDING (PLANT).

Frank E. Gardner; C. Jack Hearn

Tantalum

A chemical element, symbol Ta, atomic number 73, and atomic weight 180.948. It is a member of the vanadium group of the periodic table and is in the 5*d* transitional series. Oxidation states of IV, III, and II are also known. *See* PERIODIC TABLE; TRANSITION ELEMENTS.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Tantalum metal is used in the manufacture of capacitors for electronic equipment, including citizen band radios, smoke detectors, heart pacemakers, and automobiles. It is also used for heat-transfer surfaces in chemical production equipment, especially where extraordinarily corrosive conditions exist. Its chemical inertness has led to dental and surgical applications. Tantalum forms alloys with a large number of metals. Of special importance is ferrotantalum, which is added to austenitic steels to reduce intergranular corrosion.

The metal is quite inert to acid attack except by hydrofluoric acid. It is very slowly oxidized in alkaline solutions. The halogens and oxygen react with it on heating to form the oxidation-state-V halides and oxide. At high temperature it absorbs hydrogen and combines with nitrogen, phosphorus, arsenic, antimony, silicon, carbon, and boron. Tantalum also forms compounds by direct reaction with sulfur, selenium, and tellurium at elevated temperatures.

Edwin M. Larsen

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; R. N. Crockett, *Niobium (Columbium) and Tantalum: International Strategic Minerals Inventory Summary Report*, 1994; P. Moeller, P. Cerny, and S. Saupe (eds.), *Lanthanides, Tantalum, and Niobium*, 1989.

Tantulocarida

A subclass of the Crustacea. The Tantulocarida was initially proposed as a class of the Crustacea to accommodate two species formerly assigned to the Copepoda or Cirripedia, but now, with four additional genera that have been described, it is ranked as a subclass. Tantulocarids are minute ectoparasites, less than 0.01 in. (0.3 mm) in length, that infest deep-sea copepods, isopods, tanaids, and ostracods (see **illus.**). As a result of their parasitic mode of life, adult females have lost all resemblance to crustaceans; males are free-living but nonfeeding. Infection of the host occurs at the tantulus larval stage, which is believed to occur immediately after the larva is released. The head, or cephalon, of the tantulus larva is covered by a dorsal shield that may protrude anteriorly to form a rostrum. No cephalic appendages are present, although some of the specialized structures found in tantulocarids may possibly be derived from them. The mechanism of attachment, the mouth tube, is located on the ventral surface with a distal mouth opening, and is surrounded by an oral disk. The walls of the mouth tube consist, at least in part, of a pair of chitinous bars, each with a basal swelling. A second tubular structure, the longitudinal organ, projects into the mouth tube, reaching almost to the oral opening. A structure distinguished by its striated appearance, and referred to as the striated organ, lies above, in proximity to the dorsal surface of the head. A prominent, elongate cephalic stylet is located behind the oral region.

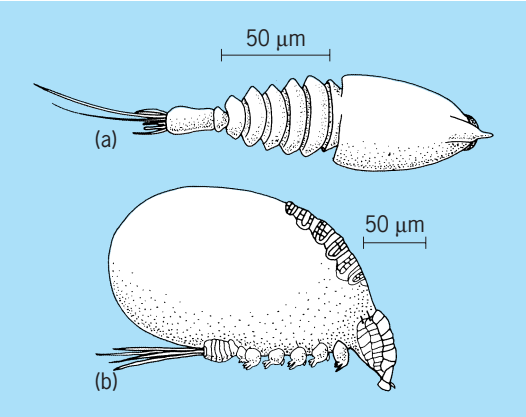

Tantulocarida. (*a*) *Deoterthron aselloticola*. (*b*) *Basipodella harpacticola*. (*After G. A. Boxshall and R. J. Lincoln, Tantulocarida, a new class of Crustacea ectoparastic on other crustaceans, J. Crust. Biol., 3:1–16, 1983*)

Posterior to the cephalon, the trunk is made up of six free somites, each with well-developed tergal plates and a pair of appendages (thoracopods). Each of the first five pairs consists of a basal segment with median, spined endite and a pair of rami. The sixth pair of thoracopods is much narrower; rami and endites are lacking. The abdomen varies in segmentation from two in *Deoterthron* and *Microdajus* to seven in *Stygotantulus*, and it terminates in a pair of small caudal rami.

Development of the maturing larva can give rise to either the adult male or female. In the former, the tantulus swells between the fifth and sixth thoracic somites or posterior to the sixth. Although the head and remainder of the thorax remain unchanged, a major reorganization of the body occurs within the trunk sac. From an undifferentiated cellular mass, the body of the male develops. When mature, the male possesses a cephalothorax; four free thoracic somites, each with a pair of thoracopods; and an unsegmented abdomen terminated by a pair of long caudal setae. In the development of the female, the tantulus swells behind the cephalic shield, the larval trunk is sloughed off, and the female body, still attached to the host by the head, becomes a simple saclike structure.

Phylogenetic relations of the Tantulocarida are not clearly understood. Although tantulocarids lack the antennules and antennae characteristic of Crustacea, two clusters of aesthetascs, or sensory hairs, in adult males are probably antennulary in origin. The basic tagmosis exhibited and male gonopore position suggest an affinity with some cirriped taxa. Some researchers include the Tantulocarida as a subclass of the Maxillopoda; others accept it as a distinct class. *See* CIRRIPEDIA; COPEPODA; CRUSTACEA; MAXILLOPODA.

Patsy A. McLaughlin

Bibliography. G. A. Boxshall and R. Huys, New tantulocarid, *Stygotantulus stocki*, parasitic on harpacticoid copepods, with an analysis of the phylogenetic relationships within the Maxillopoda, *J. Crust. Biol.*, 9:126–140, 1989; G. A. Boxshall and R. J. Lincoln, The life cycle of the Tantulocarida (Crustacea), *Phil. Trans. Roy. Soc. Lond.*, ser. B, 315:267–303, 1987; G. A. Boxshall and R. J. Lincoln, Tantulocarida, a new class of Crustacea ectoparastic on other crustaceans, *J. Crust. Biol.*, 3:1–16, 1983.

Taper pin

A tapered self-holding pin used to connect parts. Standard taper pins have a diametral taper $^1/_4$ in. in 12 in. (0.6 cm in 30 cm) and are driven in holes drilled and reamed to fit. The pins are made of soft steel or are cyanide-hardened. They are sometimes used to connect a hub or collar to a shaft (see **illus.**). Taper pins are frequently used to maintain the location of

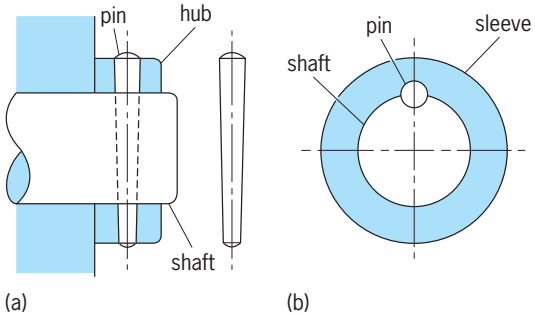

Taper pins. (*a*) Connecting hub or collar to shaft. (*b*) Connecting sleeve to shaft.

one surface with respect to another. A disadvantage of the taper pin is that the holes must be drilled and reamed after assembly of the connected parts; hence they are not interchangeable. Paul H. Black

Taphonomy

A subdiscipline of paleobiology that investigates the processes of preservation and their influence on information in the fossil record. Coined by J. A. Efremov in 1940, the term "taphonomy" involves all processes that affected the organism during its life, its transferral from the living world (biosphere) to the geological realm (lithosphere), and all physical and chemical interactions from the time of burial until collection. Besides the conspicuous characteristics of the preserved organism that can be seen easily—either morphological (external) and/or anatomical (internal) features—there are often less prominent details that record the fossil's history. Taphonomists are forensic scientists. By analyzing preserved details, paleontologists can understand an organism's mode of death or disarticulation; the biological processes that may have modified the remains before burial, including their use by hominids; the response of the organism or one of its parts to transport by animals, sediment, water, or wind; its residency time in a depositional setting before final entombment; and the alterations of tissues or skeletal parts within a wide range of chemical settings. The processes of fossilization appear to be environmentally site-specific, resulting in a mosaic of preservational traits in terrestrial and marine environments. Few fossil assemblages are exactly identical with regard to formative processes, but general patterns exist. An understanding of taphonomic assemblage features within its environmental context allows for a more accurate interpretation of the fossil record.

Early investigations focused on vertebrates, with a surge of studies encompassing all groups of organisms after 1970, and subsequent applications to forensics and archeology. Comparisons of the consistency of processes and assemblage attributes through geologic time have been made both in the rock record and in modern analog settings. These results have influenced the way in which paleontologists collect, analyze, and interpret their data. Before it was realized that all fossilized material in an assemblage offered information about preburial interactions, paleontologists expended their efforts on collecting the best-preserved (museum-quality) and most diagnostic fossils. Interpretations based upon these data sets often were heavily biased because only a small part of the fossil record was used. However, paleontologists have recognized that the inherent taphonomic biases in a fossil assemblage may not provide a complete and wholly accurate picture but, when taken into consideration, provide the most accurate means of assessing the record of life. *See* TRACE FOSSILS.

A very small percentage of organisms that have lived on Earth have had the potential to become fossils, because most biomass produced on a yearly basis is either ingested by organisms higher on the food chain or microbially degraded and recycled back into the environment. Ultimately, nearly all biomass is recycled (see **illustration**). The possibility of fossilization exists only when biomass accumulates under geologically rapid and site-specific physiochemical conditions. *See* BIOMASS.

Stages. Once fossilized, organic remains successfully have passed through several taphonomic stages, including necrology, biostratinomy, and diagenesis. However, remains may be affected at any stage, effectively removing them from the fossilization process.

Necrology is the death or loss of a part of an organism. Death is a prerequisite for most animals, but not

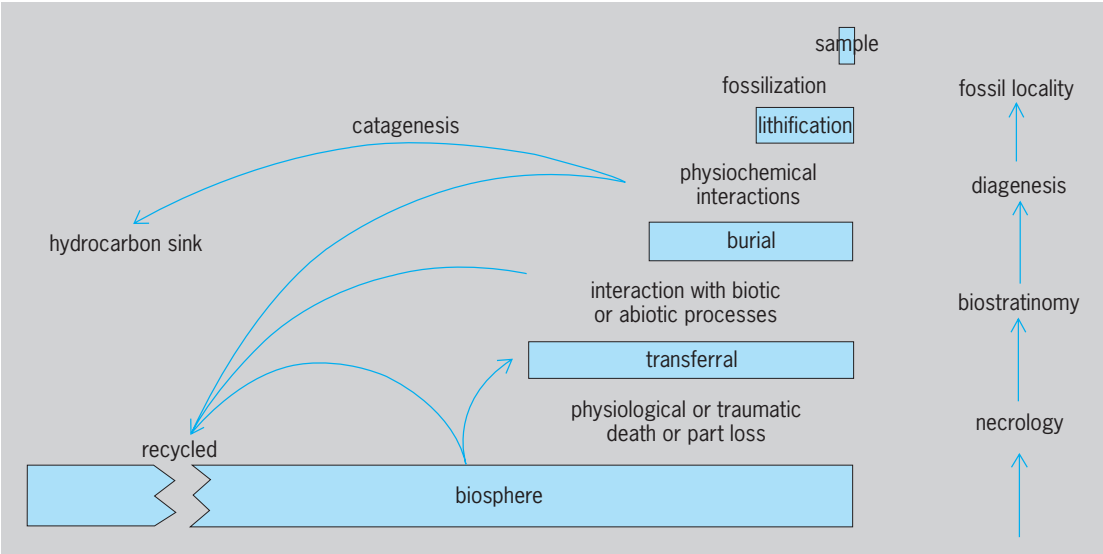

Subdisciplines of taphonomy and their relationships. (*After A. K. Behrensmeyer and S. M. Kidwell, Taphonomy's contribution to paleobiology, Paleobiology, 11:105–119, 1985; and A. Traverse, ed., Sedimentation of Organic Particles, Cambridge University Press, 1994*)

for plants that shed many different parts during their life cycle. Necrology may be induced by physiological means (old age, disease, or temporal or climatic shedding) or traumatic means (sudden catastrophic death or part loss in response to natural disruptions such as violent storms, volcaniclastic ash fall, mudflows, or flood events). The potential for any organic remains to successfully pass into the fossil record is dependent on the original biochemical composition. Organic remains of more resistant biochemicals or mineralized hard parts have a higher probability of fossilization than those of easily degraded compounds.

Biostratinomy focuses on the processes and interactions that follow necrology, until final burial. Biologic interactions include the effects of scavenging on carcasses, the use of discarded parts as domiciles, or borings in resistant structural parts. Abiotic processes include mechanical and physical alteration or breakdown under different transport conditions (fragmentation and rounding in river channels), orientation or concentration of resistant parts under varying hydrological regimes, or reexposure and reworking of previously buried remains in response to changing geological circumstances. This stage ends when organic remains are entombed within sediment, effectively isolating the material from the effects of biological degradation. The integrity of the buried part and resistance to diagenetic degradation will ultimately determine its residency time and potential for long-term burial and fossilization.

Diagenesis affects organic remains following burial and involves the physical (compaction) and chemical (cementation, recrystallization) changes in the sediment before, during, and after lithification (conversion of soft, unconsolidated sediments into hard rock). Mineralized skeletal elements of animals originally are composed of calcium carbonate (shell), calcium phosphate (bone and Precambrian shell), or silica (plankton test); resistant skeletal elements of plants are composed of biomacromolecules such as algaenans (microalgal cell walls), lignins (wood), cutins and cutans (cuticles), and sporopollenin (pollen and spores). Inorganic skeletal elements may be altered texturally and replaced mineralogically during diagenesis, whereas organic molecules undergo devolatization (loss of gases and water) and carbon concentration. Reactivity of pore waters with the organics soon after burial may result in pyritization (sulfur-reducing bacterial activity) or authigenic cementation (iron carbonate precipitation). Cellular voids may be infilled with crystallized hydrous minerals, resulting in the permineralization and preservation of internal features or cast replicas. If organics are subjected to high temperatures and pressures generated under deep burial conditions, catagenesis may result in the production of hydrocarbon fuels. *See* ORGANIC GEOCHEMISTRY; PALYNOLOGY.

Fidelity. The fidelity of any fossil assemblage is important because it provides information about the quantitative faithfulness of the record of body plans, age classes, species richness and abundances, trophic structures, and other aspects of the original biological community. Laboratory and field studies of modern communities have been used deductively to assess fossil assemblages in both terrestrial and marine realms. Information from live:dead studies, where the composition of a recent death assemblage is compared to the living community, shows a range of quantitative relationships. There is now a clear sense of how the construction and life habits of organisms and their postmortem environment in several systems combine to determine how well the death assemblage resembles the community from which it was sourced. There is tremendous variance in fidelity even among organisms that have a high probability of preservation, and this is related to the durability of their hard parts.

Fossil record and time resolution. The completeness of the stratigraphic record is restricted in time and space because of the assortment of Earth processes operating globally throughout geologic history. Taphonomic processes operating within this framework bias fossil-assemblage formation to specific depositional settings. By their nature, such assemblages generally reflect contributions from one primary systematic group: plants, invertebrates, or vertebrates. Fossils within these assemblages provide paleontological data concerning ontogenetic completeness (spectra of developmental stages within species), taxonomic completeness (the proportion of a higher taxon's component species), and ecosystem completeness (the share of an ecosystem's component species) in the fossil record at the time of accumulation.

The time represented within a fossil accumulation varies between depositional sites. Taphonomic assemblage characteristics allow paleontologists to estimate the relative or absolute time interval that has elapsed from biosphere to lithosphere transfer, which in turn provides a means of evaluating assemblage fidelity. Assemblages may represent the spectrum from the instantaneous picture (such as volcaniclastic ash burial) to accumulations that span thousands of years [for example, chenier plains (continuous ridges of beach material built upon swampy deposits) of shell debris]. Accumulations that represent short time intervals provide the highest-fidelity record, whereas time-averaged deposits represent contributions into a single bed or assemblage from successive populations. These latter assemblages preserve a mixture of noncontemporaneous remains that introduce another set of biases that must be recognized before credible interpretations can be formulated. Various statistical techniques have been devised to assess quantitative data with the prospect of deciphering patterns within time-averaged accumulations.

Relationship to paleoecology. Historical ecosystem responses to changing intrinsic short-term (local) or extrinsic long-term (global) factors are recorded in the fossil record at various resolutions. From these records it is possible to evaluate paleocommunity structure and composition and how these change through time in response to environmental

disruption, especially those associated with climate change. In turn, the dynamics associated with biodiversification of Earth, from individual assemblages to biomes, can be determined and rates of evolutionary events understood. These retrodictive (infering a past state from present data) patterns can be used as models to predict ecological reorganization under changing global conditions into the future. *See* FOSSIL; PALEOBOTANY; PALEOECOLOGY; PALEONTOLOGY.

Robert A. Gastaldo

Bibliography. R. A. Allison and D. E. G. Briggs (eds.), *Taphonomy: Releasing the Data Locked in the Fossil Record*, 1991; A. K. Behrensmeyer et al., Taphonomy and paleobiology, *Paleobiology*, 26(4)[suppl.]:103–144, 2000; A. K. Behrensmeyer et al., *Terrestrial Ecosystems Through Time*, 1992; D. J. Bottjer et al. (eds.), *Exceptional Fossil Preservation*, 2002; C. E. Brett and S. E. Speyer, *Comparative Taphonomy: Pattern and Processes in Fossil Preservation*, 2005; S. K. Donovan (ed.), *The Processes of Fossilization*, 1991; C. T. Gee and R. A. Gastaldo (eds.), Plant Taphonomy Special Issue, *PALAIOS*, vol. 20, no. 5, 2005; J. P. Huntley and S. Stallibrass (eds.), *Taphonomy and Interpretation*, 2000; S. M. Kidwell and A. K. Behrensmeyer, *Taphonomic Approaches to Time Resolution in Fossil Assemblages: Short Courses in Paleontology*, vol. 6, 1993; V. A. Krasilov, *Paleoecology of Terrestrial Plants: Basic Principles and Techniques*, 1975; R. L. Lyman et al. (eds.), *Vertebrate Taphonomy*, 2001; R. E. Martin, *Taphonomy: A Process Approach*, 1999; A. Traverse (ed.), *Sedimentation of Organic Particles*, 1994.

Tapir

An odd-toed ungulate of the family Tapiridae. These animals, with four species, have a discontinuous distribution in South America and Asia and consequently have a theoretical importance biologically.

Tapirs are nocturnal, timid animals; they spend the daytime in dense thickets. They are generally silent and solitary except during the rutting season. Their diet consists of fruits and leaves, and they can cause considerable damage to cultivated crops. The tapir has a dental formula of I 3/3 C 1/1 Pm 3/3 M 3/3 for a total of 40 teeth. The forefoot has four toes, the hindfoot three. These animals have a prehensile trunklike muzzle, a short tail, and small eyes. The female gives birth to a single young after a gestation period of 13 months. The male leaves the female to care for the young alone.

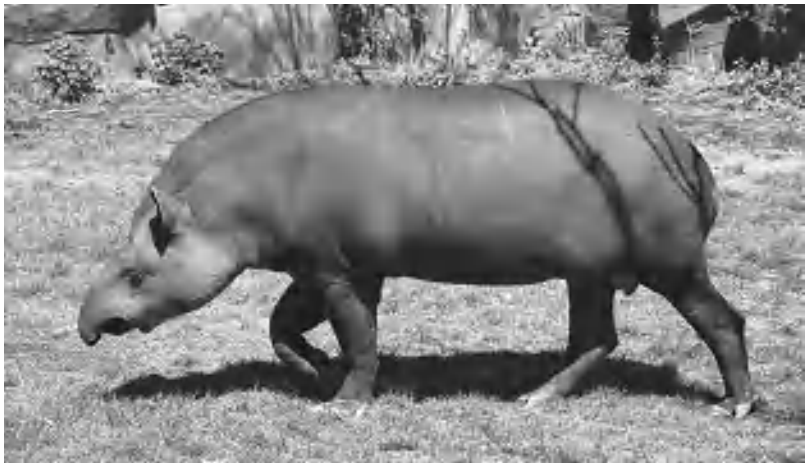

Lowland tapir (*Tapirus terrestris*). (*Courtesy of Brent Huffman/Ultimate Ungulate Images*)

South American species differ from the Malayan species in the coloring and markings of the adults. The most common species is the Brazilian tapir (*Tapirus terrestris*) [see **illustration**]. This animal is found in the lowlands of northern South America and prefers marshy areas, where it can browse on water plants and other vegetation; it is a good swimmer. It requires a large intake of water. This species stands about 3 ft (0.9 m) high and is about 6 ft (1.8 m) long, the adult attaining a maximum weight of 650 lb (295 kg). The life-span is about 30 years.

The Malayan tapir (*T. indicus*) is common in the dense forests of Sumatra and Malaya, where it browses during the night on vegetation. The adults may weigh up to 400 lb (180 kg) and reach a length of 6 ft (1.8 m). Vision is relatively poor, but the olfactory and hearing senses are good. This species spends considerable time in the water and is a good swimmer. The coat of the New World species is dark, whereas that of the adult Malayan tapir is white from the shoulder to rump region and black elsewhere. *See* MAMMALIA; PERISSODACTYLA.

Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999.

Tardigrada

A phylum of microscopic, bilaterally symmetrical invertebrates which are generally less than 1 mm in length. About 400 species are known. Commonly called water bears, bear animalcules, or urslets, they are worldwide in distribution and are found in all habitats.

Anatomy. The tardigrade body (**Fig. 1***a*) consists of an anterior prostomium and five segments. The mouth is located in the prostomium in a centroterminal position. A soft, nonchitinous cuticle surrounds the body and lines the fore- and hindgut. The cuticle may be smooth or sculptured and forms innervated cephalic appendages and spines on the trunk and legs. Four pairs of ventrolateral legs arise from the trunk and terminate in claws or other modified structures. A pair of oral glands and stylets are present which are partly cuticular and partly chitinous. Both the pharynx and esophagus are muscular structures. The digestive tract is tubular and more or less lobed due to the presence of diverticular dilations. In the heterotardigrades, separate anal and genital openings occur, while in the eutardigrades there is a single anogenital opening, the cloaca. The sexes are separate and the gonads are unpaired dorsal sacs with paired gonoducts in the male and in theory also in the female.

Histologically, the musculature of the body, legs, and mouth is of the smooth type.

The brain is a voluminous supraesophageal ganglion connected to a subesophageal ganglion by a pair of connectives. The brain has two or three inner and two outer lobules. Each of the latter, in most species except Arthrotardigrada, has one eyespot. This structure is an ocellus which is cup-shaped, and consists of one retinal and one pigmented cell which may be black, red, or pigment free. The ventral nerve cord, with four ventral ganglia, is a continuation of the subesophageal ganglion. Food storage cells float in the spacious body cavity, the coelom, which lacks a parietal or visceral peritoneum in the adult. Circulatory and respiratory structures are lacking.

These animals exhibit the phenomenon known as cell constancy. The number of epidermal cells is the same in all species of a genus. The pharynx has 27 epithelial and 24 myoepithelial cells. This number is constant in all genera except *Milnesium*, which has 24 and 39, respectively.

The lumen of the pharynx is triradial. Body muscles are metameric and comprise dorsal, ventral (Fig. 1*b*), and lateral groups. Each muscle is either a single fibrillar, uninucleated cell or a chain of such structures. The number of myocytes is constant in all individuals of the same species. The muscles retract and bend the body, while during relaxation the pressure of the fluid in the body cavity extends and stretches the animal. Storage cells of newly hatched *Echiniscoides* are constant in number and appressed to the epidermis; older cells multiply when dissociated from the epidermis.

Embryology. These animals lay eggs and development is direct. Embryonic development lasts 3–40 days, varying according to the species and surrounding temperature. Fertilization, as known in the Eutardigrada, may be external while eggs are laid in the old cuticle during molting, or internal within the ovary. Internal fertilization occurs after ejaculation of sperm into the female cloaca. Parthenogenesis has been observed in two species.

The oocytes mature with abortive oocytes serving as nurse cells. The ovarian endothelium produces the shell (chorion), which is smooth or has processes (ornamentations) that are of taxonomic importance (**Fig. 2***a*). The diploid number of chromosomes is 10–14 for species of Eutardigrada. Liberated eggs vary in diameter from 60 to 200 micrometers. The number of eggs in a single oviposit varies from 3 to 30. Cleavage is total and nearly equal, and is irregular because of asynchronous division of the blastomeres.

Gastrulation, by coeloblastic delamination, occurs at the 80–100-cell stage. Primordial germ cells are recognizable in the primary entoderm. The enterocoelous origin of the entomesoderm results in the formation of five pairs of coelomic sacs (Fig. 2*b*). The four anterior pairs give rise to muscles and storage cells, while the posterior pair unites to form the gonad from which gonoducts develop. The nervous system originates from the epidermis by delamination. The secondary entoderm becomes the midgut.

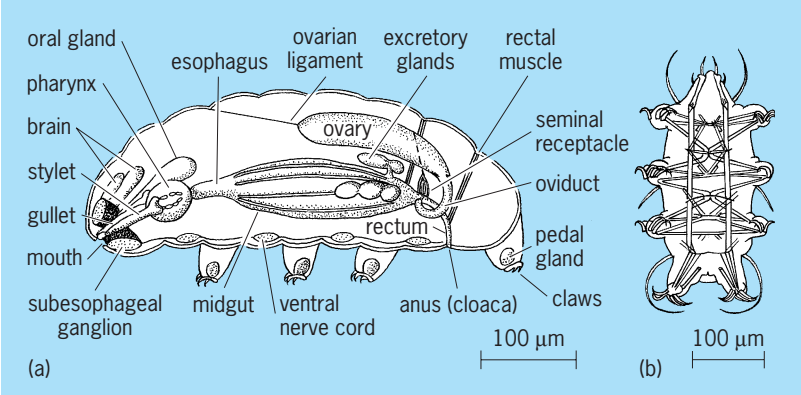

Fig. 1. Tardigrade morphology. (*a*) *Macrobiotus*, principal organs. (*b*) Muscles, ventral view.

Newly hatched tardigrades measure about 51 μm, while the largest adult is about 1200 μm. The very young animals swell as a result of water intake. Fewer claws and appendages are characteristic of the juvenile. These animals ordinarily grow to three to five times their initial size, and the immature tardigrades may lay eggs.

Molting. During their active life of 18 months, tardigrades molt about 12 times. They are unable to feed in the 5–10 days of molting; the buccal cavity requires at least 5 days for renewal (**Fig. 3**). Molting begins when the animal passes to the simplex stage, which is characterized by expulsion of the buccal

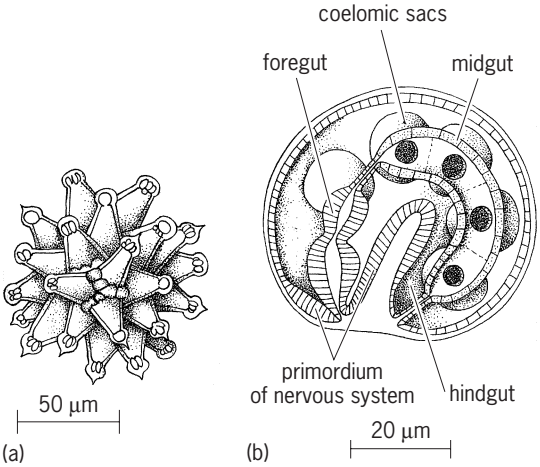

Fig. 2. Embryology of tardigrades. (*a*) Egg with processes. (*b*) Development of coelomic sacs.

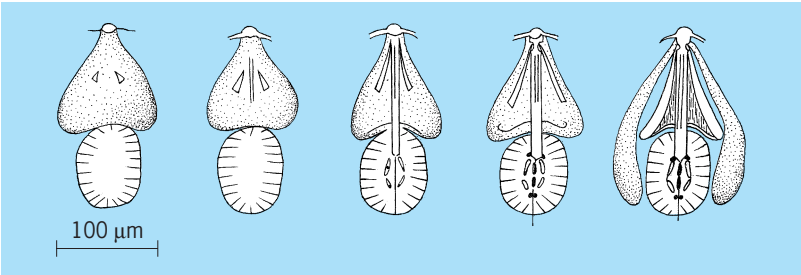

Fig. 3. Tardigrade buccal apparatus in 5 consecutive days after molting.

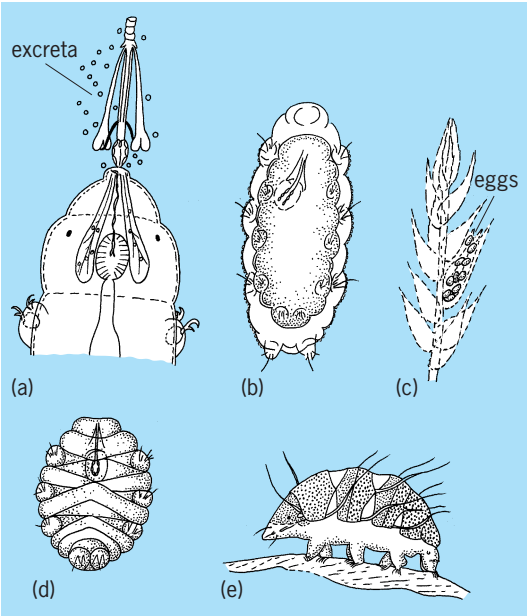

Fig. 4. Active and inactive forms of Tardigrada. (*a*) Expulsion of the buccal apparatus and excreta. (*b*) Cyst. (*c*) Eggs fastened to a moss leaf. (*d*) Barrel. (*e*) Scutechiniscid in the walking position.

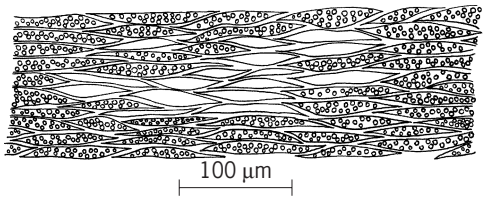

Fig. 5. Sixteen moss cells emptied by feeding action of tardigrade.

cuticular parts (**Fig. 4*a***). During molting, the epidermis secretes a new cuticle, and epidermal foot glands renew the toes or claws. The oral glands shorten and then regenerate stylets and bearers. In *Diphascon*, epithelia of the gullet, pharynx, and rectum replace corresponding cuticles. Under certain conditions, molting tardigrades, at least in the Eutardigrada, may remain in the old cuticle while the new one thickens to form a cyst (Fig. 4*b*). Hunger and perhaps unusual warmth seem to elicit encystment. Histolysis within a cyst appears improbable. It is possible that consumption of stored food causes the tardigrade to leave its cyst upon formation of the third cuticle. Molting is completed when the newly formed stylets pierce the old cuticle and allow the animal to escape. Duration of active life, without encystment, for as long as 30 months has been recorded for these animals.

Physiology. Tardigrada live as active forms, without encystment, only when surrounded by a pellicle of water. They are mainly herbivorous, and feed by piercing the wall of plant cells with their stylets. They ingest the contents of these cells (**Fig. 5**) by means of pumping pharynx. As a result, many ingested chloroplasts are fragmented. The foregut is acid in reaction (pH 4.4–5.2), while the midgut is alkaline (8.4–8.7). Digestion occurs only in the cells of the midgut (**Fig. 6**) which absorb the green-colored droplets containing plant matter. These cells contain digestive ferments. Fats, starch, and glycogen have been demonstrated in the storage cells.

Sometimes tardigrades, especially *Milnesium*, pierce the cuticle of rotifers, nematodes, other tardigrades, and nauplii to ingest their soft parts. Active animals can fast as long as 5 weeks. In active life in the Heterotardigrada, defecation occurs during molting with deposition of waste in the old cuticle (Fig. 4). In the Eutardigrada, it is independent of molting. Excretion occurs from the intestinal cells into the lumen of the gut, also from the oral glands at the beginning of a molt, and from the epidermis (part of pigments) during the formation of the new cuticle.

When the surrounding medium dries up, most tardigrades continue to live as inactive or anabiotic barrel-shaped structures called cysts without any protective cover (Fig. 4*d*). Desiccation begins when there is a loss of oxygen from the water. The animal responds by contraction and loss of body water. Dried eggs also survive. Moistened animals usually revive, but anabiosis and revival cannot be repeated indefinitely. Experimentally, the limit of a survival was 14 cycles of desiccation-hydration. Anabiosis can last for as long as $6^1/_2$ years, but minimal manifestations of muscular activity were observed when barrels which had remained dry for 120 years were moistened. By computing maximum periods of desiccation plus a total active life of 30 months, it is found that a tardigrade might live for 67 years, hence the name *Macrobiotus*. The limit of life during anabiosis does not depend on the amount of stored food, since those

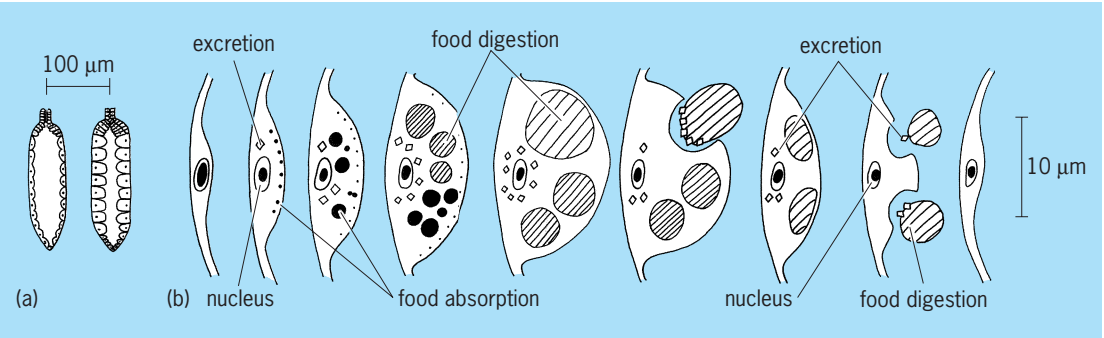

Fig. 6. Digestive activities in tardigrade. (*a*) Gut before and after absorption. (*b*) Digestive processes in gut cells.

with a reserve food supply do not differ in anabiotic capacity from those without one. These facts imply that the cause of death must be endogenous.

A 100-year-old discussion on anabiotic animals was characterized by two opposing views. One argument maintained that anaerobiosis in these animals was actually a very slow metabolism or "vita minima." The other claimed that life stopped without death supervening, and the organism was compared to a wound clock which had stopped ticking. A. Pigón and B. Weglarska settled the matter in 1953 by establishing a vita minima. They found that cysts were resistant to 198°F (92°C) heat in the air for 1 h. The cysts withstood cold at temperatures to −314°F (−192°C) for 20 months, and to −458°F (−272°C) for $8^1/_2$ h. In liquid air or helium used to study low-temperature effects, cysts are not completely dry, but contain water within and between their cells.

Tardigrades contract to the characteristic barrel shape when dissolved oxygen decreases in the water (Fig. 4*d*). They cannot maintain this state in water, however, and within 48 h they become maximally distended and immobile. Such asphyctic tardigrades may continue to live for 5 days and can revive if oxygen and food are made available. Animals emerging from anabiosis pass into a state of asphyxy, since the cells have lost their osmotic capacity. The arthrotardigrade, *Batillipes*, does not endure desiccation. For *Echiniscoides sigismundi*, about 42% of a population withstood dryness for 10 days, but none survived after 4 weeks. All were found to be alive in rainwater after 3 days, and in an asphyctic state.

Movement in the tardigrades is varied. Scutechiniscidae walk with legs which are nearly perpendicular, under the trunk (Fig. 4*e*), while *Tetrakentron* crawls. No species swims. *Echiniscoides* shows positive phototaxis, while the negative response of *Macrobiotus dispar* is a directed reaction. Eggs are frequently laid with positive thigmotaxis either between the leaflet and stalk of moss or in the shells of water fleas (Fig. 4*c*).

Ecology. Most species are widely distributed. Dissemination may be by wind, birds, and certain terrestrial animals which transport tardigrade eggs and barrels. Other moss dwellers, such as rhizopods, rotifers, and nematodes, are more easily transported than tardigrades. The tropics have few species, and Scutechiniscidae are rare on the Antarctic continent.

Echiniscoides sigismundi is the most euryokous species since it is worldwide in distribution and found both in littoral algae and at altitudes up to 3000 ft (1000 m) in the eastern Congo. In this environment tardigrade populations probably survive the dry periods. Population density varies considerably. In one sample of water containing *Enteromorpha*, 1 ml yielded 60 specimens.

A few scutechiniscids and several eutardigrades are limnetic organisms. The limnetic fauna is not sharply delimited from the terrestrial fauna, so that more limnetic than marine species are found, usually among algae, aquatic mosses, sand, and mud. Sandy beaches of soft-water lakes contain more specimens that do those of hard-water lakes. Some populations of permanently aquatic habitats cannot survive desiccation. Most species are eurythermal, having the ability to live through a wide range of temperature conditions. They are active between near-freezing temperature and 77–86°F (25–30°C). One known species is continuously exposed to temperatures of 104°F (40°C).

Most tardigrades are terrestrial. They are found among lichens, liverworts, densely growing soft-leaved mosses, and also in rather hard-leaved Pottiacea and Grimmiacea. They are rarer in some ferns, lycopods, and phanerogams such as *Sedum*, *Saxifraga*, and *Haastia*. The soil among fallen needles and leaves will yield tardigrades. As many as 22,000 animals have been recovered from 0.04 oz (1 g) of air-dried moss. An average population produces 780 lb (354 kg) of humus per hectare yearly, chiefly by means of fecal decomposition.

Fewer species are found in permanently wet or damp mosses than in land mosses. Oxygen supply is better in the land mosses, and desiccation excludes many rival organisms.

Tardigrades and their eggs are preyed upon chiefly by Amoebozoa, especially *Difflugia*. Nematodes attack the eutardigrades. Common pathogens of tardigrades are Phycomycetes (*Macrobiotophthora vimariensis*) and Microsporidia (*Pleistophora*). *See* CELL CONSTANCY; CLEAVAGE (DEVELOPMENTAL BIOLOGY); EUTARDIGRADA; HETEROTARDIGRADA; POLYCHAETA.

Eveline Marcus

Bibliography. K. W. Cooper, The first fossil tardigrade, *Psyche*, 71(2):41–48, 1964; E. Marcus, Tardigrada, in H. G. Bronn (ed.), *Klassen und Ordnungen des Tierreichs*, 1929; E. Marcus, Tardigrada, in F. Schulze and W. Kükenthal (eds.), *Das Tierreich*, pt. 66, 1936.

Tarragon

A herb of the genus *Artemesia* in the aster family (Asteraceae) that is used as a spice. Tarragon is often divided into French and Russian types. French tarragon can be distinguished by its highly aromatic leaves, low seed set or fertility, and compact growth habit. Due to its long history of vegetative propagation to preserve its fine scent characteristics, French tarragon has all but lost its ability to form viable seeds. Russian tarragon does produce seeds, but the plant is much lower in overall oil content and is not as fine-scented.

Tarragon is a perennial that grows to a height of 2–3 ft (60–120 cm) in one season, and then dies back to the ground with frost. In spring the plant sprouts from underground rhizomes and resumes its growth. The leaves contain 1.5–3% volatile oil with an odor similar to anise and chervil.

Tarragon requires special cultivation techniques because of its sterility. Root division, or more properly rhizome division, is the predominant method of propagation. This entails digging up the tarragon plants, then separating the rhizomes and replanting a larger area. Rooting stem tips and tissue culture

techniques have also been developed with some success.

Tarragon is susceptible to a number of diseases, including wilts and dodder (a parasitic plant). Nematodes are a particular problem because they inhabit the rhizomes and are often passed on to new plantings. Weed control is usually accomplished by hand, but some chemical control is practiced outside of the United States. *See* ASTERALES; SPICE AND FLAVORING.

Seth Kirby

Bibliography. L. H. Bailey, *Manual of Cultivated Plants*, rev. ed., 1975; F. Rosengarten, *The Book of Spices*, 1969.

Tartaric acid

Any of the stereoisomeric forms of 2,3-dihydroxybutanedioic acid: L(+), D(−), and meso [(**1**), (**2**), and (**3**), respectively]. L(+)-Tartaric acid is present

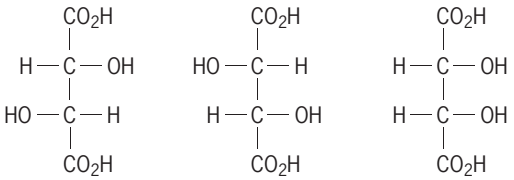

(**1**) (**2**) (**3**)

in the juice of various fruits and is produced from grape juice as a by-product of the wine industry. The monopotassium salt precipitates in wine vats, and L(+)-tartaric acid is recovered from this residue. On heating in alkaline solution, the L(+) acid is converted to the racemic mixture of (**1**) and (**2**), plus a small amount of the meso acid (**3**).

Several systems have been used to designate the stereochemical configuration of the tartaric acid isomers, as shown in the **table**. The early names were based on the sign of optical rotation; the carbohydrate convention or the nomenclature of the International Union of Pure and Applied Chemistry (IUPAC system) are now preferred.

Tartaric acid has played a central role in the discovery of several landmark stereochemical phenomena. In 1848, L. Pasteur isolated enantiomers (**1**) and (**2**) by mechanical separation of hemihedral crystals of the racemic mixture. He also used tartaric acid and its salts to demonstrate a distinction between the meso isomer (**3**) and the racemic mixture (**1**) + (**2**), and between enantiomers and diastereoisomers in general. The difference in properties between (**1**) [or (**2**)] and the meso form (**3**) was later a key in establishing the relative configuration of the pentose and hexose sugars.

In 1951 the absolute configuration of L(+)-tartaric acid was determined to be structure (**1**) by anomalous x-ray scattering of the potassium rubidium salt. This experiment established the correctness of the arbitrary assignment of the configuration of the hexose sugars and of all other compounds whose configuration had been correlated with that of tartaric acid. *See* STEREOCHEMISTRY.

Both L(+)- and D(−)-tartaric acid and the esters are inexpensive compounds and are used as chiral auxiliary reagents in the oxidation of alkenes to enantiomerically pure epoxides. This method employs a hydroperoxide oxidant, titanium alkoxide catalyst, and L(+)- or D(−)-tartrate, and involves chirality transfer from the tartrate to the product. *See* ASYMMETRIC SYNTHESIS; EPOXIDE.

Tartaric acid has some use as an acidulant in foods and also as a chelating agent. Potassium hydrogen tartrate (cream of tartar) is an ingredient of baking powder. The potassium sodium salt, commonly called Rochelle salt, was the first compound used as a piezoelectric crystal. *See* CHELATION; PIEZOELECTRICITY.

James A. Moore

Taste

Taste, or gustation, is one of the senses used to detect the chemical makeup of ingested food—that is, to establish its palatability and nutritional composition. Flavor is a complex amalgam of taste, olfaction, and other sensations, including those generated by mechanoreceptor and thermoreceptor sensory cells in the oral cavity. Olfactory (smell) sensory cells of the nose are particularly important in the perception of flavor. Taste sensory cells respond principally to the water-soluble chemical stimuli present in food, whereas olfactory sensory cells respond to volatile (airborne) compounds. *See* SENSATION.

Anatomy. The sensory organs of gustation are termed taste buds. In humans and most other mammals, taste buds are located on the tongue in the fungiform, foliate, and circumvallate papillae and in adjacent structures of the throat. There are

Stereoisomers of tartaric acid†

	(**1**)	(**2**)	(**3**)	(**1**) + (**2**)
Nomenclature				
Early usage	*d*-	*l*-	meso-	racemic acid
Carbohydrate convention	L(+)-	D(−)-	*meso*-	*rac*-
IUPAC	*2R, 3R*	2*S*, 3*S*	2*R*, 3*S*-	—
Chemical Abstracts	*R*-(R*R*)	*S*-(R*R*)	R*, S*	*R*, R*(±)
Properties				
Melting point °C (°F)	168–170 (334–338)	168–170 (334–338)	159 (318)	206 (403)
Optical rotation, [α]	+12°	−12°	0°	0°

†Structures (**1**)–(**3**) correspond to those displayed in the text.

approximately 5000 taste buds in humans, although this number varies tremendously from person to person. Taste buds are goblet-shaped clusters of 50 to 100 long slender cells. Microvilli protrude from the apical (upper) end of sensory cells into shallow taste pores. Taste pores open onto the tongue surface and provide access to the sensory cells. Taste cells are modified epithelial cells that develop at the base of the taste bud, differentiate into functional sensory cells, and ultimately die. This cycle of differentiation, decline, and replacement, which lasts from 8 to 10 days, continues as long as the nerve supply to the taste buds remains intact. When the taste nerves are cut, taste buds disappear. Taste buds reappear when the peripheral taste nerve fibers regenerate and reinnervate the tongue epithelium. Individual sensory nerve fibers branch profusely within taste buds and make contacts (synapses) with taste bud sensory cells. Taste buds also contain supporting and developing taste cells (**Fig. 1**). *See* TONGUE.

The anterior two-thirds of the tongue is innervated by a sensory branch of the facial nerve (cranial nerve VII), the posterior tongue by the glossopharyngeal nerve (cranial nerve IX), and the throat and larynx by branches of the vagus nerve (cranial nerve X). These nerves carry information from touch, temperature, and pain sensory cells, as well as from taste buds. The taste fibers from the anterior part of the tongue branch from the lingual nerve in a smaller nerve called the chorda tympani (sensory branch of the facial nerve). The chorda tympani nerve traverses the eardrum en route to the brainstem. When the chorda tympani is damaged, as in the surgical removal of the tympanum, taste sensitivity is lost on

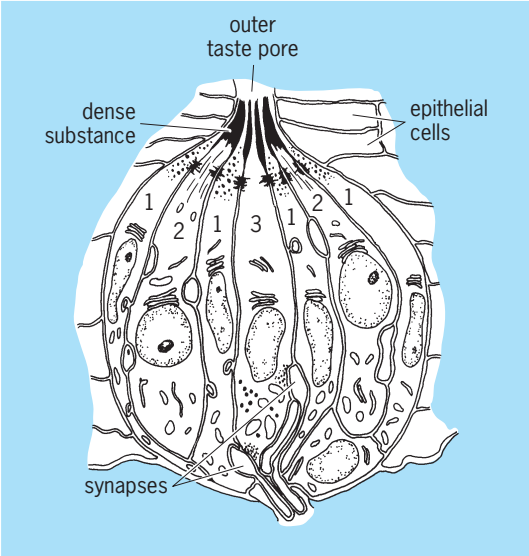

Fig. 1. Drawing of an electron microscopic section through a rabbit taste bud. A taste receptor cell (no. 3) is shown making two synaptic contacts with sensory nerve fibers. Several cells (no. 2) that make numerous physical contacts with nerve fibers also may be involved in impulse transmission. Darker cells (no. 1) are presumed to be supporting cells. (*After R. G. Murray, The ultrastructure of taste buds, in I. Friedmann, ed., The Ultrastructure of Sensory Organs, pp. 3–81, 1973*)

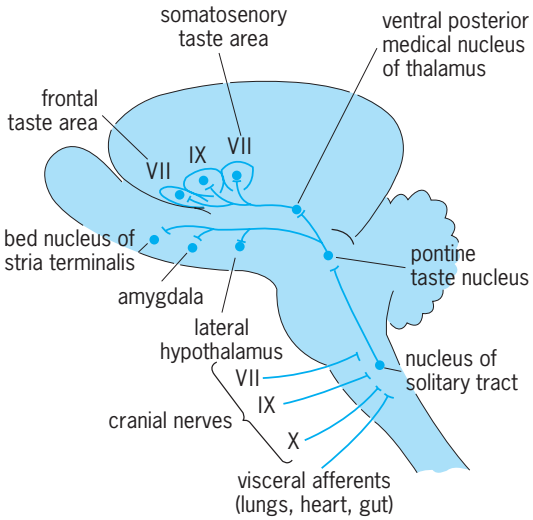

Fig. 2. Schematic drawing of the two neural pathways for taste in the rodent brain. (*After R. Norgren, Taste pathways to hypothalamus and amygdala, J. Comp. Neurol., 166:17–30, 1976*)

the anterior two-thirds of the tongue on the same side as the surgery because the nerve is interrupted.

The cell bodies (somata) of taste nerve fibers (primary sensory afferent fibers) are clustered in several small ganglia nestled within the cranium but outside the brain itself. In the brain, taste fibers from the lingual, glossopharyngeal, and vagus nerves converge in the solitary tract and its nucleus in the medulla oblongata. These fibers terminate on neurons in a region contiguous to the area where touch and temperature sensory nerve fibers from the tongue also end. Second-order fibers from these neurons ascend to a cluster of small neurons located in the ventrobasal thalamus. (In rodents, another relay site—the pontine taste nucleus, or parabrachial nucleus—is interposed between the nucleus of the solitary tract and the thalamus; **Fig. 2**.) From the thalamus, neurons project to the lateral sensory cortex to two areas, one a mixed tactile-taste zone and the other largely a taste area, the primary gustatory cortex. This thalamocortical system is concerned largely with discrimination among different tastes. Outputs from the primary gustatory cortex terminate in the orbitofrontal cortex, where axons from the olfactory sensory cortex and other sensory cortical areas converge. This area of the brain appears to be dedicated to processing more complex perceptions such as flavor, where information from several sensory modalities is integrated. A second projection from the pontine taste area travels more ventrally, sending collaterals to the hypothalamus and other limbic structures to terminate ultimately in the central nucleus of the amygdala. This ventral pathway is presumed to mediate taste preferences, aversions, and hedonic responses (pleasantness or unpleasantness) of taste (Fig. 2). *See* BRAIN; NERVOUS SYSTEM (VERTEBRATE).

Qualities. The basic taste qualities experienced by humans include sweet, salty, sour, and bitter. (In some species, pure water also strongly stimulates taste bud cells.) A fifth taste, umami, is now

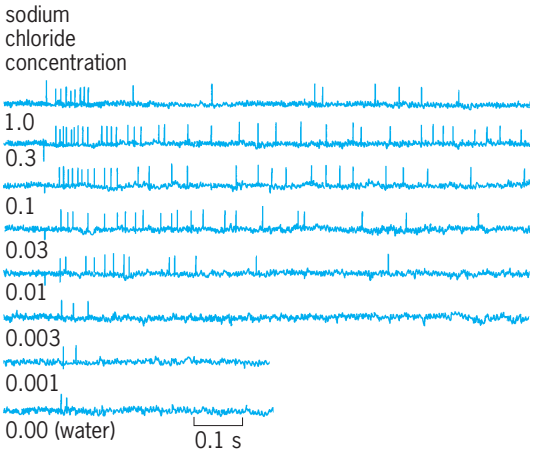

Fig. 3. Typical oscillographic record of impulses in a single taste sensory nerve fiber of a rat. (*After C. Pfaffmann, Gustatory nerve impulses in rat, cat and rabbit, J. Neurophysiol., 18:432–433, 1955*)

recognized by many as distinct from the other qualities. Umami is a Japanese term roughly translated as "good taste" and is approximated by the English term "savory." It refers to the taste of certain amino acids such as glutamate (as in monosodium glutamate) and certain monophosphate nucleotides. These compounds occur naturally in protein-rich foods, including meat, fish, cheese, and certain vegetables.

The middorsum (middle top portion) of the tongue surface is insensitive to all tastes. Only small differences, if any, exist for the taste qualities between different parts of the tongue (contrary to the "tongue maps" of taste sensitivities that are commonly published). Electrophysiological records from individual taste bud sensory cells show increased depolarization with an increased concentration of an effective stimulus. This leads to the release of neurotransmitters at synapses on the afferent taste nerves (Fig. 1) and an increase in frequency of nerve impulses therein (**Fig. 3**). No simple direct relationship exists between chemical stimuli and a particular taste quality except, perhaps, for sourness (acidity). Sourness is due to H^+ ions. The taste qualities of inorganic salts are complex, and sweet and bitter tastes are elicited by a wide variety of diverse chemicals.

Sensitivity. Specific sensitivity to particular chemicals is attributed to ion channels and to molecular receptors on the surface of sensory cells in taste buds, particularly on the apical membrane. Taste bud sensory cells are believed to have a spectrum of chemical sensitivities that differs from cell to cell. Some taste bud sensory cells probably respond to only one of the basic taste qualities, whereas others may be less discriminating and respond to multiple chemicals. This suggests that taste bud sensory cells possess more than one type of molecular receptor on their membranes. Alternatively, the surface membrane receptors themselves may be somewhat broadly sensitive to chemical stimuli. The sensitivity of individual taste bud sensory cells and their taste receptors is currently a topic of intense investigation.

More is known about the chemical sensitivity of sensory nerve fibers that innervate taste buds (chorda tympani, glossopharyngeal, and vagal). These fibers represent the sensory information output from taste buds. A fraction of the fibers are activated when the tongue is stimulated with a specific chemical or closely related compounds such as sucrose and fructose, and thus are narrowly receptive to chemical stimulation. However, even fibers with a broad sensitivity to diverse chemicals usually respond best (that is, give the highest frequency of action potentials) to a single type of chemical stimulus. The perception of taste quality (sweet, sour, salty, bitter, umami) is probably coded by the concurrent activation of some sensory fibers that are highly selective for certain chemical compounds as well as by other fibers that are less selective. That is, the pattern of activity within a heterogeneous population of sensory nerve fibers elicited by a particular taste stimulus determines taste perception (**Fig. 4**). No single homogeneous population of nerve fibers carries the information for a given taste quality. Various statistical procedures (such as factor analysis and multidimensional scaling) have been used to analyze these patterns of activity in taste fibers and have provided quantitative evidence for the existence of sweet, sour, salty, bitter, and umami taste qualities. *See* CHEMORECEPTION.

Sour. Studies by different investigators of human sour taste (acidity) reveal a wide range of thresholds for detection. For instance, hydrochloric acid (HCl)

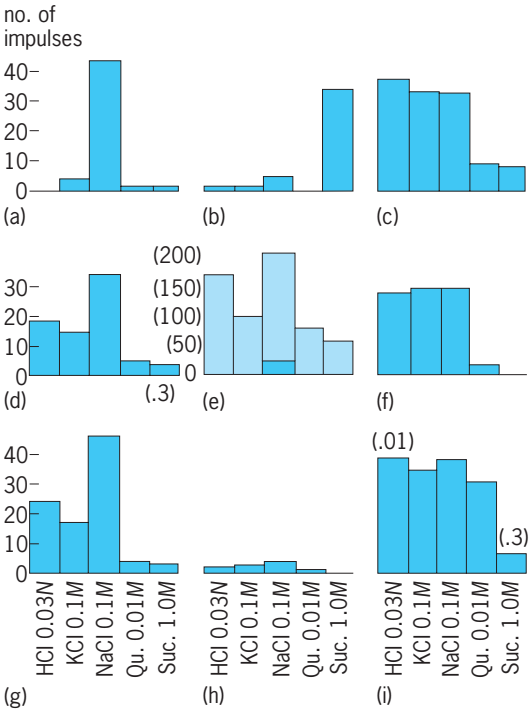

Fig. 4. Bar diagram of the different sensitivity patterns in (*a–i*) each of nine single sensory fibers from a rat. Dark color indicates number of impulses during first second of discharge to each of five test solutions shown along base: HCl (hydrochloric acid); KCl (potassium chloride); NaCl (sodium chloride); Qu (quinine); Suc (sucrose). Light color on fiber *e* shows relative amount of neural activity in total nerve (sum of all fibers) to same test solutions. (*After C. Pfaffmann, Gustatory nerve impulses in rat, cat and rabbit, J. Neurophysiol., 18:432–433, 1955*)

can be tasted at a concentration of about 0.0005 molar (*M*). Sourness increases with increases in hydrogen ion concentration, but weak organic acids are more sour than would be predicted from their hydrogen ion concentration. Increasing carbon chain length in the aliphatic acid series, for example, appears to enhance the potency of taste stimulation.

A number of explanations for how acid-sensitive taste sensory cells respond to hydrogen ions have been put forth. Possible mechanisms include the following:

1. Penetration of positively charged hydrogen ions (protons, H^+) through specific ion channels such as epithelial sodium channels (ENaC) in the apical membrane. The influx of H^+ would depolarize and excite the taste sensory cell. Taste bud sensory cells express ENaC channels, and pharmacological agents that block these channels reduce sour sensitivity in certain species, including humans.

2. Activation of cation-selective ion channel receptors by protons. Mammalian degenerin-1 (MDEG1) is one such receptor that has been identified in taste bud sensory cells in mice. Activation of MDEG1 by H^+ would allow the influx of cations such as sodium (Na^+), again depolarizing and exciting the taste sensory cell.

3. Block of potassium-selective ion channels in the apical membrane by H^+. Potassium ions (K^+) tend to leave the cell through these channels, keeping the sensory cell at a negative resting potential. Any agents, such as protons, that block these channels will result in depolarization and excitement of the taste sensory cell, as has been demonstrated in amphibian taste bud sensory cells.

4. Possible stimulation of specialized ion channels called hyperpolarization-activated, cyclic-nucleotide-gated (HCN) or pacemaker channels by H^+. Such channels have been identified in taste bud cells in some species. Their stimulation would lead to depolarization of the cell.

Perhaps some combination of all the above events underlies sourness, and sour taste mechanisms may differ from species to species. The final answer is not yet achieved. *See* BIOPOTENTIALS AND IONIC CURRENTS.

Salt. Most salts, including sodium chloride (NaCl), elicit other qualities, such as bitter or sour, in addition to salty. The dominant taste elicited by table salt is salty. Low-molecular-weight salts are predominantly salty, many higher-molecular-weight salts are bitter, and the salts of lead and beryllium are sweet. The median human threshold for detecting sodium chloride is approximately 0.01 *M*, but a wide range of values has been reported. Both the anion (for example, Cl^-) and cation (for example, Na^+) contribute to saltiness and to stimulus potency. According to one theory, salt-sensitive taste bud cells are stimulated by the influx of Na^+ through ENaC channels in their cell membrane. ENaC channels are found in other tissues in the body that are involved in Na^+ transport, such as renal tubules. Blocking ENaC channels with pharmacologic agents has been reported to reduce salt taste responses in some species. (However, the effects of blocking ENaC channels on human salt taste are controversial at present.) Further, H^+ also passes through ENaC channels, and agents that block ENaC channels alter sour taste. These observations raise the conundrum that the same mechanism (ENaC channels) appears to underlie both salt and sour tastes. However, there is no complete explanation yet for these observations. *See* SALT (FOOD).

Sweet. Sweet taste is associated largely with organic compounds such as sugars, alcohols, glycols, and sugar derivatives, with the exception of certain lead and beryllium salts. Sucrose thresholds for humans have a median value of 0.1 *M*, whereas the synthetic sweetener saccharin is 700 times more potent. The discovery that L-aspartyl-L-phenylalanine methyl ester (aspartame) is extremely sweet has led to further study of other dipeptides, but not all are sweet; some are sour, bitter, umami, or even tasteless. Some proteins have been discovered that elicit sweet taste with thresholds for humans in the 10^{-7} *M* concentration range. Slight changes in spatial arrangement render a sweet molecule tasteless or even bitter (**Fig. 5**). *See* ASPARTAME.

Studies of the specific blockage of sweet sensitivity by gymnemic acid, which does not affect salt, sour, or bitter, point to the existence of specific sweet receptors on the surface of taste bud sensory cells. One theory attributes the sweet taste to substances with a molecular system AH-B, where A and B are electronegative atoms separated by a distance of 0.3 nanometer. The AH moiety is a proton donor and the B moiety a proton acceptor. Sweet taste receptors are assumed to possess a complementary

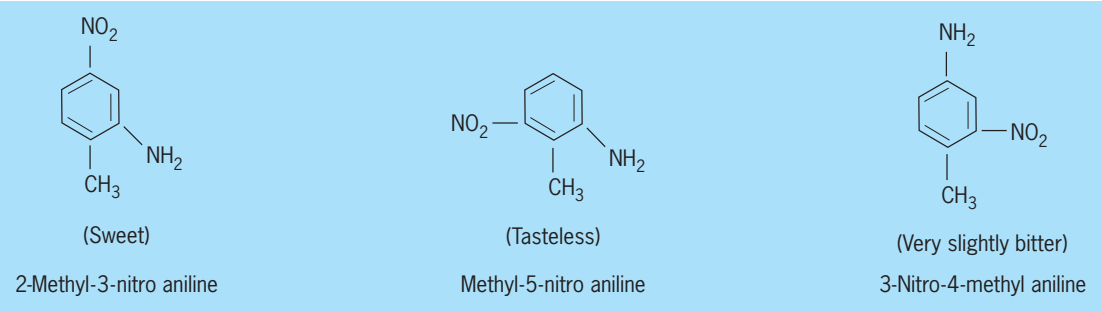

Fig. 5. Taste quality changes when spatial arrangements within the molecule change. (*After S. S. Stevens, ed., Handbook of Experimental Psychology, Wiley, 1951*)

AH-B system with which the stimulus system can interact, forming two intersystem hydrogen bonds. More recent formulations add a third molecular feature. Two candidate receptors for sweet taste have recently been cloned from human and rodent taste buds. These receptors (T3Rs) are G protein–coupled receptors, meaning that stimulation of the sweet receptor activates a cascade of intracellular enzymatic reactions involving G proteins. One such G protein that is associated with sweet taste is gustducin, which is found in taste bud cells and in some related cells. The net result of activating the sweet receptors on sweet-sensitive taste bud cells appears to be the block of K^+ channels, which depolarizes and excites taste bud sensory cells.

Bitter. Bitter is elicited by many chemical compounds and may be found in association with sweet and other taste qualities. An increase in the molecular weight of inorganic salts or in the length of a carbon chain of organic molecules may be associated with increased bitterness. Typical of substances with the bitter taste are the alkaloids such as quinine, caffeine, and strychnine, which are often toxic. A median threshold value for quinine has been cited at 0.000008 *M*. Taste blindness is an inherited inability to taste the bitterness of specific compounds such as phenyl thiocarbamide (PTC) and substances with the thiocarbamide group (N—C═S). About one-third of Caucasians are nontasters of PTC.

Two mechanisms have been proposed to explain how bitter-sensitive taste bud cells respond to stimulation. For example, certain bitter compounds directly block K^+ channels, a mechanism shared by sour and sweet sensing cells, and bitter taste receptors have been cloned and identified in human and rodent taste bud cells. Bitter taste receptors (T2Rs), like the candidate sweet taste receptors (T3Rs), are also G protein–coupled receptors. Unlike the candidate sweet receptors that have been identified to date, a large family of genes (about 30) encode bitter taste receptors. The G protein gustducin is also believed to interact with bitter receptors and activate an intracellular signaling cascade. Mice lacking gustducin show reduced taste responses to sweet and bitter chemicals. Thus, the determining factor for whether a given taste bud sensory cell responds to bitter or sweet is which receptor or receptors are expressed on the apical surface.

Umami. Umami, or savory taste, is stimulated by L-glutamate and certain nucleotide monophosphates such as 5′-inosine monophosphate and 5′-guanosine monophosphate. The threshold for human perception of L-glutamate is about 0.0008 *M*. Humans and other mammals find L-glutamate a preferred taste. Consequently, the natural glutamate content of prepared foods is often increased to enhance palatability by adding glutamate in the form of monosodium glutamate (MSG) or as protein hydrolysates. An important characteristic of umami is the synergistic interaction between glutamate and nucleotide monophosphates. Low concentrations of glutamate and nucleotide monophosphates interact to generate a robust savory taste.

A candidate G protein–coupled receptor for umami has been identified. It has been termed taste-mGluR4 and is a variant of the synaptic glutamate receptor mGluR4. The G proteins and enzymes stimulated by taste-mGluR4 have not yet been identified.

Adaptation. Prolonged exposure of the tongue to a taste stimulus decreases the response to that stimulus. This phenomenon, seen as a decrease in nerve activity or in the perceived intensity of the stimulus, is called adaptation. Adaptation by flowing taste solutions continuously over the tongue may lead to a rise in threshold or even complete disappearance of the particular taste sensation. Not only does that taste sensation disappear, but water and all concentrations of the stimulus at or below that used for adaptation show a contrasting taste that is often quite intense. Sodium chloride and sucrose adaptation tends to produce a bitter subadapting taste; adaptation to HCl and quinine produces a sweet subadapting taste. A salty taste is less common but may follow adaptation to compound sour-bittersweet substances such as urea. That bitter or sour taste can be masked by sweetening agents is well attested by the use of sweetners in coffee or in sour lemonade.

Nutrition and taste. The ability of an organism to select nutritious or necessary ingredients of the diet by taste can be demonstrated with the self-selection technique, in which an animal is given free choice among individual containers with necessary nutrients in pure form. After certain physiological stresses, like glandular imbalances, the selection of the various nutritive agents often shows compensatory changes. A severely salt-deficient rat (adrenalectomized) may show a significant increase in the intake of sodium chloride sufficient to counteract the usually fatal outcome in the absence of salt replacement therapy. Similar effects have been noted in children; but in adults, food habits, cultural conditioning, learning, and other complex psychological factors play significant enough roles in food acceptance to override the physiological factors controlling behavior. Nonetheless, food palatability, as determined by taste, olfaction, and other sensory effects, has such a profound effect on food acceptance that a substantial applied science of flavor technology has developed. *See* CHEMICAL SENSES.

Carl Pfaffmann; Nirupa Chaudhari; Stephen Roper

Bibliography. T. E. Finger, W. L. Silver, and D. Restrepo, *Neurobiology of Taste and Smell*, 2d ed., Wiley-Liss, 2000; T. A. Gilbertson, S. Damak, and R. F. Margolskee, The molecular physiology of taste transduction, *Curr. Opin. Neurobiol.*, 10(4):519–527, 2000; B. Lindemann, Taste reception, *Physiol. Rev.*, 76(3):718–766, 1996; S. D. Roper, Gustatory and Olfactory Sensory Transduction, Ch. 49 in N. Sperkelakis (ed.), *Cell Physiology Sourcebook: A Molecular Approach*, 3d ed., Academic Press, San Diego, 2001; S. A. Simon and S. D. Roper, *Mechanisms of Taste Transduction*,

CRC Press, Boca Raton, FL, 1993; D. V. Smith and R. F. Margolskee, Making sense of taste, *Sci. Amer.*, 284(3):32–39, 2001.

Taurus

The Bull, a large zodiacal northern constellation prominent in the evening winter sky (see **illustration**). Taurus is usually said to be the form assumed by Zeus in Greek mythology in order to carry off the princess Europa. *See* ZODIAC.

The head of the bull is marked by the Hyades, a large V-shaped open cluster, against which is projected the red giant star Aldebaran, though it is only half as far away. From the V near Aldebaran, two bright stars farther out mark the points of Taurus's long horns. Near one of those stars, Zeta, is the Crab Nebula (M1, the first object in Messier's eighteenth-century catalogue, though it had been previously discovered by John Bevis), the remnant of a supernova whose light reached Earth in AD 1054, and which is detectable all across the spectrum of gamma rays, x-rays, ultraviolet, light, infrared, and radio waves. *See* ALDEBARAN; CRAB NEBULA; HYADES; MESSIER CATALOG.

The Pleiades, the seven sisters of Greek mythology, are a star cluster riding on Taurus's back. Only six are visible to the unaided eye, though binoculars or telescopes show dozens or hundreds more stars in the cluster. The Pleiades (M45, the 45th object in Messier's catalogue) have the shape of a small dipper, and long exposures show dust surrounding many of the stars, which are hot and therefore bluish. The dust reflects the starlight toward us. *See* PLEIADES.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Tautomerism

The reversible interconversion of structural isomers of organic chemical compounds. Such interconversions usually involve transfer of a proton (prototropy), but anionotropic (allylic, Wagner-Meerwein) rearrangements may be reversible and so be classed as tautomeric interconversions.

Lactam-lactim tautomerism. A cyclic system containing the grouping —CONH—is called a lactam, and the isomeric form, —COH═N—, a lactim. These terms have been extended to include the same structures in open-chain compounds when considering the shift of the hydrogen from nitrogen to oxygen.

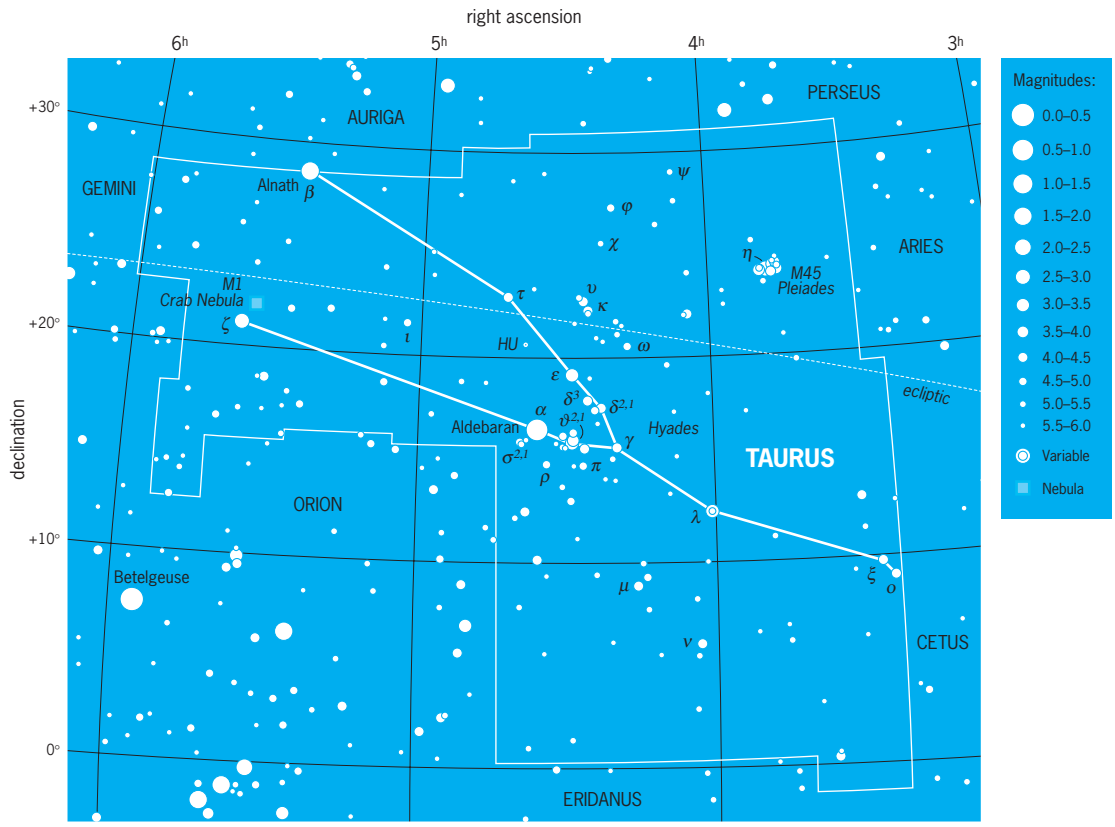

Modern boundaries of the constellation Taurus, the Bull. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

Isatin (**1**) appears to react in either the lactam (**1**) or the lactim (**2**) structure. Thus, there is a

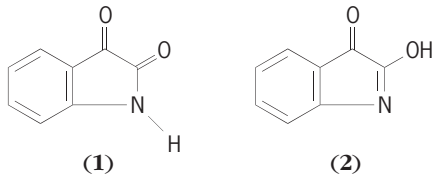

(**1**) (**2**)

precedent for proving the existence of lactim-lactam tautomerism—that chemical behavior may be inferred from the structure of a reaction product. Spectroscopic techniques are more reliable, and it is often possible to determine from the absorption spectrum whether a given substance has either one structure or both. *See* LACTAM.

Keto-enol tautomerism. The molecular grouping

may in certain substances exist partly or wholly as

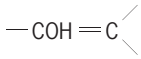

The former constitutes the keto form and the latter the enol form. Kurt Meyer first studied the keto ($CH_3COCH_2CO_2C_2H_5$) and enol ($CH_3OHC{=}CHCO_2C_2H_5$) forms of ethyl acetoacetate, and recognized them respectively by reactions specific for the carbonyl group and the carbon-carbon double bond. Both forms may be obtained in relatively pure condition, the former by freezing it out of the mixture and the latter by slowly distilling the mixture in quartz apparatus. However, each is slowly converted into the equilibrium mixture of the two. Extensive chemical and spectroscopic studies showed that the enol content of such an equilibrium mixture is a function of the physical state of any given substance. The gas phase or solution in a nonpolar solvent (hexane) favors the enol form, whereas more polar solvents (chloroform, alcohols) repress its formation.

The existence of an enol in an acyclic system requires that a second carbonyl group, or its equivalent, for example,

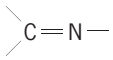

be attached to the same

as an aldehyde or ketone carbonyl. Thus, ethyl acetoacetate tautomerized demonstrably, but ethyl malonate ($C_2H_5O_2CCH_2CO_2C_2H_5$) does not. Occasionally an enol form exists, these requirements notwithstanding. For example, ethyl pyruvate is partially enolized ($CH_3COCO_2C_2H_5 \rightleftharpoons CH_2{=}COHCO_2C_2H_5$), and α-hydroxy ketones (or aldehydes) exhibit the characteristics of the tautomeric enediols, for example, benzoin [reaction (1)].

$$C_6H_5COCHOHC_6H_5 \rightleftharpoons C_6H_5OHC{=}COHC_6H_5 \qquad (1)$$

Where the enol form includes an aromatic ring such as phenol, the existence of the keto form is often not demonstrable, although in some substances such as 4-nitrosophenol (**3**) and 4-hydroxypyridine (**4**) there may be either chemical

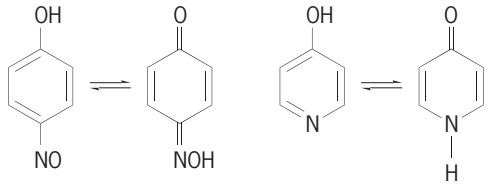

(**3**) (**4**)

or spectroscopic evidence for both forms. Closely related to keto-enol tautomerism is the prototropic interconversion of nitro and aci forms of aliphatic nitro compounds such as nitromethane (**5**).

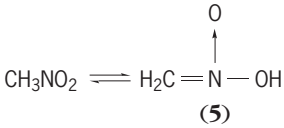

(**5**)

Ring-chain tautomerism. An acyclic hydroxyaldehyde may exist in equilibrium with its cyclic hemiacetal. The failure of glucose to form a normal acetal with an alcohol and the production of two isomeric glucosides instead led to the postulate that carbohydrates exist principally as inner or cyclic hemiacetals in equilibrium with only enough free aldehyde to permit typical aldehyde reactions with reagents that either oxidize the carbonyl group or form derivatives that effectively remove it from the equilibrium [reaction (2)]. The glycosides are formed by the elimina-

$$CH_2OHCHOHCHOHCHOHCHOHCHO \rightleftharpoons CH_2OH\,\underbrace{CHCHOHCHOHCHOHCHOH}_{O} \qquad (2)$$

tion of water between the hydroxyl (OH) derived by hemiacetal formation and an alcohol, with two structures being possible, and the hydroxyl lying above or below the hetero ring. *See* GLUCOSE; GLYCOSIDE.

In general, tautomeric forms exist in substances possessing functional groups that can interact additively and that are so placed that intramolecular reaction leads to a stable cyclic system. The cyclic form usually predominates (especially if it contains five or six members).

Certain alkenic acids are tautomeric with their lactones [reaction (3)], and this is called lacto-enoic tau-

$$RCH{=}CHCH_2CO_2H \rightleftharpoons \underbrace{RCHCH_2CO}_{O} \qquad (3)$$

tomerism. However, lacto-enoic tautomerism not involving a prototropic shift has also been observed [reaction (4) Ar = aromatic group].

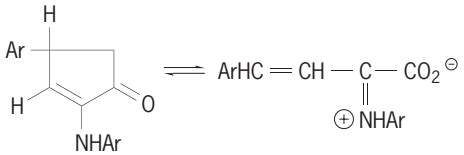

(4)

Still another type of ring-chain isomerism not involving a prototropic shift is demonstrable by the reactions of phthaloyl chloride [reaction (5)].

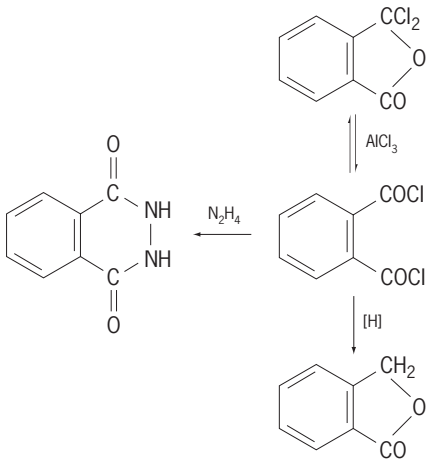

(5)

The latter type of ring-chain tautomerism is closely related to anionotropic rearrangements such as the allylic [reaction (6)] and Wagner-Meerwein [reaction (7)], which may thus be considered examples

$$RCHClHC{=}CH_2 \rightleftharpoons RHC{=}CHCH_2Cl \qquad (6)$$

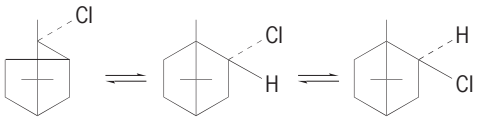

(7)

of anionotropic tautomerism. *See* MOLECULAR ISOMERISM. Wyman R. Vaughan

Valence tautomerism. The term valence tautomerism applies generally to relatively low-energy processes in which concerted, synchronous σ- or π-electron shifts occur over several atoms. Neither diradical nor dipolar species intervene in such processes, nor do atom or functional group migrations occur in conjunction with these electron shifts. The phenomenon of valence tautomerism is distinct from and should not be confused with resonance. In the former case, individual tautomers represent energy minima that interconvert via a transition structure that is an energy maximum on the potential surface that leads from one tautomer to the other. When this concerted electron-shift process occurs in systems for which the energy barrier to interconversion of valence tautomers is relatively high, this process frequently is referred to as valence isomerization. *See* CHEMICAL BONDING. Alan P. Marchano

Bibliography. Z. Rappoport (ed.), *The Chemistry of Enols*, 1990; R. E. Valters, and W. Flitsch, *Ring-Chain Tautomerism*, 1985.

Taxis

A mechanism of orientation by means of which an animal moves in a direction related to a source of stimulation. There exists a widely accepted terminology in which the nature of the stimulus is indicated by a prefix such as phototaxis, chemotaxis, geotaxis (gravity), thigmotaxis (contact), rheotaxis (water current), and anemotaxis (air current). The directions toward or away from the stimulus are expressed as positive or negative, respectively. Finally, the sensory and locomotory mechanisms by means of which the orientation is achieved are denoted by a second type of prefix forming a compound noun with taxis. Positive phototropotaxis thus describes a mechanism by means of which an animal carries out a directed movement toward a source of light along a path which permits the animal's paired eyes to receive equal intensities of light throughout the movement. The following are examples of various types of taxes.

Klinotaxis. A well-analyzed case of this type of taxis is the way in which a fly maggot moves away from the light immediately before pupating in a dark and sheltered place. When such a maggot is exposed to a horizontal beam of light above the substrate, it moves along a fairly straight path away from the light in the direction of the beam. In doing so it waves its front end from side to side, exposing alternately the left and right lateral aspects of its front end to the light shining from behind. As long as the light intensity falling on the light-sensitive surfaces on either side remains equal in subsequent exposures, the animal follows a straight path away from the light source. This may be explained by the hypothesis that the extent of the swing toward one side is a function of the light intensity falling on the other. If the animal starts its course from a position at an angle with the light beam, the differential light intensities falling on its anterior flanks automatically steer it into a path curving into line with the beam. The prefix "klino" in this case denotes that "exploratory" side-to-side bending of the body brings about orientation by directed, though waving movement, related in its direction to the stimulus.

Tropotaxis. Tropotaxis is a term closely related to Jacques Loeb's original notion of tropism, which has now become restricted to the description of orientation phenomena in sessile organisms. The essential point is the unwavering turn of the organism into the stimulus direction by means of innate reflex mechanisms linking bilaterally symmetrical receptors with the organs of locomotion. Paths toward or away from a single source are straight. In the case of two or more sources, they run along the resultant of incident intensities. After unilateral receptor loss, tropotactic steering leads, in a uniform field of illumination, to continued "circus movement" away from or toward the injured side in the case of positive or negative taxis, respectively.

Telotaxis. Whereas bilateral intensity balance on receptors is essential in klinotaxis and tropotaxis, orientation in telotaxis occurs, as it were, by the orientation of one or the other of two bilaterally symmetrical receptors toward the stimulus source. Of a number of simultaneously offered stimuli, all but one may be "ignored" at any given instant by means of built-in switch mechanisms based on inhibition or

block. Unilateral sense organ loss does not lead to circus movement.

A special case of telotaxis is the light-compass reaction. In this the animal moves at a temporarily fixed angle with respect to the stimulus direction. The angle can be changed at will. The term pharotaxis has been suggested to describe an orientation toward a progressively changing direction of stimulus, such as the change of distribution of polarization of the light of the sky during the course of the day (navigation by honeybees and other arthropods).

The description of taxes given so far is based on concepts formulated by Alfred Kühn and after him by G. Fraenkel and D. Gunn. They apply largely to mechanisms of orientation found in small-brain animals which rely in their basic orientation on a relatively small number of innate responses. These are characterized by their stereotyped and rather inflexible nature. O. Koehler and with him a number of students of animal behavior have attempted to make the taxis concept part of the more complex and plastic behavior of vertebrates, including humans.

Thus Koehler's definition of a taxis is as follows: Every purely reflexive and every voluntary act begins with certain postures and movements which orient the animal's head, limbs, and body with reference to the direction of the eliciting stimulus or of the final goal of the movement. These intentional postures or movements of postural adjustment are to be considered the taxis components of the animal's act. Even in humans, the turning of eyes or head into the direction of an object of interest or desire is held to fall within this definition. *See* POSTURAL EQUILIBRIUM; PLANT MOVEMENTS.

Otto E. Lowenstein

Taxodonta

An order of bivalves in subclass Lamellibranchia whose hinge dentition is characterized by a series of numerous similar alternating teeth and sockets. Typical genera of so-called ark shells are *Arca*, *Barbatia*, and *Anadara*, all with heavy coarse or matted periostracum, and mostly with strong radial ribs on globose or purse-shaped shells. The filibranch gills are enlarged to form the four double lamellae of the filter-feeding organs as in all true lamellibranchs, and the labial palps form complex sorting surfaces.

Ark-shell bivalves are probably distantly related to the true marine mussels (Mytilacea), with which they share certain features of gill ciliation and byssal organization. Confusion arises because many primitive protobranch bivalves (totally unrelated and now placed in the subclass Protobranchia), using palp proboscides in feeding and lacking lamellate gills, also have taxodont hinge dentition. Formerly the name Taxodonta was used to designate a subclass of bivalves consisting of a mixed group of lamellibranch ark shells and primitive protobranchs. *See* BIVALVIA; LAMELLIBRANCHIA; MOLLUSCA; PROTOBRANCHIA.

W. D. Russell-Hunter

Taxonomic categories

Any one of a number of formal ranks used for organisms in a Linnaean classification. Biological classifications are orderly arrangements of organisms in which the order specifies some relationship. Taxonomic classifications are usually hierarchical and comprise nested groups of organisms. The actual groups are termed taxa. In the hierarchy, a higher taxon may include one or more lower taxa, and as a result the relationships among taxa are expressed as a divergent hierarchy that is formally represented by tree diagrams. In Linnaean classifications, taxonomic categories are devices that provide structure to the hierarchy of taxa without the use of tree diagrams. By agreement, there is a hierarchy of categorical ranks for each major group of organisms, beginning with the categories of highest rank and ending with categories of lowest rank, and while it is not necessary to use all the available categories, they must be used in the correct order.

Categories. Categories commonly used in botanical and zoological classifications are listed below, from highest to lowest rank:

Botanical categories	*Zoological categories*
Divisio	Phylum
Classis	Class
Ordo	Order
Familia	Family
Genus	Genus
Species	Species

Conceptually, the hierarchy of categories is different than the hierarchy of taxa. For example, the taxon Cnidaria, which is ranked as a phylum, includes the classes Anthozoa (anemones), Scyphozoa (jellyfishes), and Hydrozoa (hydras). Cnidaria is a particular and concrete group that is composed of parts. Anthozoa is part of, and included in, Cnidaria. However, categorical ranks are quite different. The category "class'' is not part of, nor included in, the category "phylum.'' Rather, the category "class'' is a shelf in the hierarchy, a roadmark of relative position. There are many animal taxa ranked as classes, but there is only one "class'' in the Linnaean hierarchy. This is an important strength of the system because it provides a way to navigate through a classification while keeping track of relative hierarchical levels with only a few ranks for a great number of organisms.

When Linnaeus invented his categories, there were only class, order, family, genus, and species. These were sufficient to serve the needs of biological diversity in the late eighteenth century, but were quite insufficient to classify the increasing number of species discovered since 1758. As a result, additional categorical levels have been, and continue to be, created. These categories may use prefixes, such as super- and sub-, as well as new basic levels such as tribe. An example of a modern expanded botanical hierarchy of ranks between family and species is:

Familia
Subfamilia
Tribus
Subtribus
Genus
Subgenus
Sectio
Subsectio
Series
Subseries
Species

Conceptual issues. Linnaean categories are the traditional devices used to navigate the hierarchy of taxa. But categories are only conventions, and alternative logical systems are frequently advocated. Taxonomic categories could be dispensed with, and relative position could be indicated by using only indentation. Taxa indented at the same level would have equivalent places in the hierarchy. But this method is cumbersome, and especially so when the classification occupies many pages and there are many subordinate taxa separating two closely related taxa of the same rank. Actual physical measurements would be needed using some arbitrary standard. Alternatively, a numbering scheme for ranks might be substituted. Taxa with equivalent codes would have the same position in the classification. While logical, this system is cumbersome. The Linnaean system of ranks seems more familiar, even to those who are not taxonomists.

Exactly what a categorical rank is supposed to indicate is also controversial. Evolutionary taxonomists adjust the ranks of taxa to reflect subjective judgments as to the significance of one taxon compared to another. For example, birds are accorded the rank of class while their closest living relatives, crocodiles, are considered an order in another class (Reptilia). Thus, the classifications of evolutionary taxonomists do not always reflect common ancestry relationships. Phylogenetic systematists (cladists) rank taxa strictly by their common ancestry relationships. As a result, sister taxa, that is, taxa of the same age or origin that uniquely share a common ancestor, always have the same rank.

The last controversy concerns the comparability of taxa of similar rank. Sister taxa are biologically comparable, but taxa in different groups that are ranked at the same level have no necessary biological equivalency. In preevolutionary concepts, rank could be correlated with increasing perfection. Thus, a family of butterflies might be comparable to a family of buttercups. Modern taxonomists acknowledge that categorical rank is a poor to misleading criterion on which to base comparative biology and that strict comparisons must either be restricted to sister groups or to groups where the phylogeny can be consulted to determine relative comparability. *See* SYSTEMATICS; CLASSIFICATION, BIOLOGICAL; PLANT TAXONOMY; SPECIES CONCEPT; TAXONOMIC CATEGORIES; ZOOLOGICAL NOMENCLATURE. Edward O. Wiley

Bibliography. R. E. Blackwelder, *Taxonomy: A Text and Reference Book*, 1967; P. H. Davis and V. H. Heywood, *Principles of Angiosperm Taxonomy*, 1963; E. Mayr and P. D. Ashlock, *Principles of Systematic Zoology*, 2d ed., 1991; E. O. Wiley, *Phylogenetics, The Theory and Practice of Phyolgenetic Systematics*, 1981.

Taxonomy

The human activity of naming organisms and organizing these names according to a given criterion. In biology, taxonomy has evolved into formal rules and systems to classify organisms. It is part of the broader discipline of systematics, a discipline that covers all of comparative biology, such as comparative morphology, genomics, and biogeography. Systematists who develop formal biological classifications and publish taxonomic revisions and keys are practicing taxonomy.

Taxonomy is one of the oldest activities of humans, originating with language. Such "folk" taxonomies survive today, even in developed cultures, in the form of common names for plants and animals. Many of these folk taxonomies are organized around the function of the organism or the importance of the organism to the society. For example, we might recognize groups of predators and prey, those who hunt us and those we hunt. Among plants, we might recognize edible and poisonous plants. Peoples in hunter-gatherer societies are quite good at recognizing the same species of plants and animals recognized by taxonomists. The dividing line between such folk taxonomies and scientific taxonomies can be thin, but in general, scientific taxonomies are attempts to organize diversity in a manner that humans think exists in nature in order to reflect some general scientific principle. Ultimately, this takes the form of a classification, an attempt to organize the world around us using names. Taxonomies are not restricted to biology; astronomers place stars into classes that reflect their mass and brightness, two characteristics that predict how any individual star formed and how it will behave before its eventual death. The classification has predictive power that reflects what we think we know about nature. In a similar manner, modern biologists have developed taxonomies that reflect evolutionary descent, which opens the door to all sorts of comparative predictions ranging from the ecologies of species to the potential for some species to be pathogens. *See* ANIMAL SYSTEMATICS; CLASSIFICATION, BIOLOGICAL; PLANT TAXONOMY.

History. The ancient Greeks applied science as an exercise in explaining the world around them in natural rather than supernatural terms, and so the first scientific taxonomies were Greek. Aristotle (384–322 BCE) classified animals with similar characteristics. For example, he recognized animals with blood, what we now call vertebrates, and grouped them into five genera: fishes, whales, mammals (quadrupeds who gave birth to live young), reptiles and amphibians (quadrupeds that lay eggs), and birds. He also recognized animals without blood (invertebrates) and distinguished between shelled

animals (such as clams), cephalopods such as octopi and squids, insects, crustaceans, and plant-like animals such as jellyfishes. The larger groups were called genera and the smaller were called species.

Aristotle's work was amazingly detailed, even if he did not recognize whales as mammals. A student of Aristotle, Theophrastus (ca. 370-285 BCE), produced one of the first plant classifications, distinguishing plants on the basis of their form (trees, scrubs, herbs, etc.) and did fundamental work on plant morphology. Because plants were important in medicine, this work was continued, and by the first century CE the Roman physician Dioscorides had cataloged some 600 medicinal plants. His work was copied for physicians for more than a millennium. Except for the work of Albertus Magnus (1193–1280), who recognized vascular and nonvascular plants as well as monocots and dicots, little progress was made in plant taxonomy until the dawn of the Renaissance. Aristotle's work preserved by Arab culture through the translation of Averroes of Cordoba (Ibn Rushd; 1126–1198) was the accepted work on animals until the sixteenth century.

World exploration by Europeans coupled with the rise of their science brought renewed interest in animal and plant diversity as newly discovered specimens flooded European centers of learning. By 1555 the Swedish naturalist Konrad Gessner published the first of his *Historiae Animalium* (1555–1558), considered by many as the foundation for modern zoology. Andrea Cesalpino, an Italian physician, published *De plantis libri XVI* in 1583. His classification of plants by their seed types, fruits, and form rather than providing a simple list of medicinal plants laid the foundation for modern plant classification. Between these early works and the standardization of taxonomic nomenclature were important works on plants and animals by such workers as John Ray (1627–1705). Modern taxonomy began to emerge. The simple "genus" and "species" dichotomy was replaced with something resembling modern categories by Augustus Rivinus (1652–1723) and Joseph Pitton de Tournefort (1656–1709), who used more categories (class, order, section) and a consistent generic name along with a modifier (a phrase) that later evolved into Linné's binomial nomenclature. Before these works, plants in different "higher genera" frequently had the same name; to Ray apples were *Malus* and peaches were *Malus Persica*. This evolution of taxonomy reached its first peak with the work of Carl Linné (Linnaeus), the Swedish botanist who introduced the idea of a consistent binomial for each species of plant and animal and adopted or invented higher categories to classify larger groups, producing what become known as the Linnaean hierarchy.

The primary purposes of early classifications, such as those of Linné, were first, the classifications organized diversity and second, they provided a means of identification. In this second role, many early taxonomies worked more like what we recognize as keys, dichotomous choices of characters designed to identify a specimen. Thus, early classifications were built around a limited number of characters thought to be important in identifying organisms or showing where unidentified organisms might fit given the characters used. Also, the form of the classification made a place for any newly discovered organisms. Many kinds of taxonomies were proposed during the eighteenth and nineteenth centuries, each built around a concept thought to produce a natural classification and thus an enduring taxonomy. For example, the Quinterians thought that groups came in fives and attempted to classify around this principle. Some, such as Richard Owen (England), built taxonomy around the concept of idealized types (archetypes) that defined major groups. Darwin provided a different concept; natural classifications could be built around the principles of evolutionary descent and common ancestry. This concept was taken to its logical end by the German entomologist Willi Hennig in the mid-twentieth century and emerged as the dominant paradigm of biological classification used to form modern taxonomies.

Taxa and categories. Most modern taxonomies are hierarchical classifications of groups within groups, resulting in a biological classification. In addition, modern taxonomies include detailed descriptions of the organisms and a history of the various names by which these organisms have been called. Such works are usually referred to as taxonomic revisions.

In classifications and taxonomic revisions, each group of plants or animals is given a name. The group itself is a taxon (examples: *Homo sapiens*, Vertebrata). The place of the group within the taxonomy is denoted by ranking. In Linnaean taxonomies, rank is denoted by placing the taxon name with a category rank (example: class Vertebrata). Each category has a place on the hierarchy relative to other categories. Each taxon is unique, but the categories are not unique. For example, there are over 400 different families of bony fishes. This relative ranking is set by convention for each system of taxonomy. There are several systems of taxonomy, including systems for bacteria (prokaryotes) and viruses. The two most familiar are for plants and animals. Some of the common categorical ranks and examples of taxa are provided below for humans and pine trees:

Humans
- Domain Eukarya
  - Kingdom Animalia
    - Phylum Chordata
      - Class Synapsida
        - Order Primates
          - Family Hominidae
            - Genus *Homo*
              - Species *Homo sapiens*

Pine trees
- Domain Eukarya
  - Kingdom Plantae
    - Division Pinophyta
      - Class Pinopsida
        - Order Pinales (=Coniferales)

Family Pinaceae
Genus *Pinus*
Species *Pinus ponderosa*

See TAXON; TAXONOMIC CATEGORIES.

Modern taxonomy in practice. There may be 12–20 million different species of organisms alive today. Each one requires a name so that scientists can communicate about the species. Recognition of species and their classification is fundamental to fields as diverse as conservation biology and medicine, and the status of a particular population of organisms as a species can have international consequences as nations attempt to conserve and use their biological resources. One major goal of taxonomy is naming species. Each species has a two-part name, a genus name and a specific epitaph. *Homo sapiens* (humans) and *Pinus ponderosa* (ponderosa pine trees) are examples. Species names are always "set apart" from the rest of the text, usually by italics (*Homo sapiens*). The second major goal of taxonomy is to organize these species into larger groups and give these groups names. Each group of species is given a single name (example: Mammalia). Groups of species are frequently referred to as supraspecific taxa or higher taxa, with "higher" meaning that their rank category is higher than the rank of species. *See* SPECIES CONCEPT.

From the time of Linnaeus until the end of the twentieth century, there was a dialog among systematists as to what might constitute the most general kind of classification. Intense debates were held in the literature and in scientific meetings from the 1950s to 1990s among proponents of three distinct points of view. The Pheneticists asserted that general classifications should be based on some measure of overall similarity (counting differences and similarities) in order to be repeatable. The Phylogeneticists (Cladists) asserted that general classifications should be based on common ancestry. The Evolutionary Taxonomists grabbed the middle ground and asserted that some groups could be recognized based on common ancestry while others should be based on overall similarity (especially overall genetic similarity). By the end of the twentieth century, it became apparent that the Phylogeneticists had won the battle, primarily because they had demonstrated that it was possible to reconstruct evolutionary histories of organisms (phylogenies) objectively, thus blunting the objections of the Pheneticists who claimed that reconstructing phylogenies was too problematic to be objective. They also demonstrated that the hybrid systems of the Evolutionary Taxonomists (the middle ground) resulted in loss of information or even misleading information about common ancestry relationships and similarity. Thus, in the twenty-first century, taxonomy is returning to Darwin's assertion that taxonomies should be, whenever possible, based on common ancestry relationships. This paradigm unites the diverse fields of systematics, bringing taxonomies in line with phylogenies when phylogenies are available. Taxon names become meaningful in an evolutionary sense because the organisms classified in the taxon share a common ancestral species only among themselves. Such a taxonomy is said to be logically consistent with the evolutionary history of the group.

The effects of this paradigm shift can be profound for taxonomy. Consider birds. Old-style taxonomies (and many current school texts) recognize birds as a taxon removed from reptiles and rank both groups as classes (class Reptilia, class Aves). Modern classifications recognize birds as an order (or even suborder) within a larger group, Archosauria, that includes crocodiles (superorder Archosauria, order Crocodilia, order Aves). The old-style classification is one that emphasizes the similarities between crocodiles and lizards and the obviously distinctive nature of birds. The modern classification recognizes that birds share a common ancestor with crocodiles (as well as with dinosaurs and other archosaurs).

The key to understanding why modern biologists prefer the modern classification is the predictive power of the taxonomy: the ability of the classification to predict and explain aspects of both similarity and difference in a manner that is superior to the old-style classification. Crocodiles and birds have certain bones that are hollow; lizards do not. The old-style classification predicts that these hollow bones evolved independently in crocodiles and birds. The modern classification predicts that hollow bones evolved in the common ancestor of crocodiles and birds—an observation consistent with the accepted phylogeny where such hollow bones are interpreted as evolutionary innovations of a common ancestor shared by all archosaurs, crocodiles, dinosaurs, and birds, but not in a common ancestor these archosaurs share with lizards. The old-style classification predicts that a scale-covered body is an evolutionary innovation of the common ancestor of lizards and crocodiles. The modern classification predicts that while a scale-covered body was, indeed, a homology of the common ancestor of lizards and crocodiles, it evolved at a much earlier date and the ancestor is shared with birds as well. The old-style classification predicts that parental care evolved independently in crocodiles and birds, but it is no surprise to modern taxonomists that dinosaurs practiced parental care, given that crocodiles do so. Such care is best explained as a behavioral evolutionary innovation that originated in the common ancestor of all archosaurs.

Taxonomic formalities. Taxonomy as an activity that gives names to groups of organisms has evolved over the last two centuries. Although binomial nomenclature of Linné was adopted by 1800, his set of rules for naming plants fell into disuse. Each taxonomist created his own set of naming rules, and this produced chaos as different names were applied to the same species or group by different authors. This prompted both the zoological and botanical communities to formulate rules for naming in different countries. But this hardly solved the problem: Americans might have a set of rules that differed from Germans, hardly the way to foster international understanding. By the beginning of the twentieth

century, international rules governing the formation and use of names were adopted that were meant to be universal in application, irrespective of borders or language. These rules were formalized by the adoption of Codes of Nomenclature, first for animals and plants, and much later for prokaryotes, cultivated plants, and viruses.

All codes are independent of each other, so some have slightly different rules. In general, each Code of Nomenclature performs two basic functions. First, within the scope of the code, it specifies what names are to be used and how this is determined. In general, disputes (two names for the same taxon) are arbitrated by declaring that the oldest published name is to be used. This is the Law of Priority. Younger names, called junior synonyms, are common because of a lack of communication or because of taxonomic disputes or failure to appreciate sexual or ontogenetic differences. In addition, each Code specifies how names are to be formed for taxa placed at specific categorical ranks. This naming principle requires, for example, that a taxon ranked at the family level have an ending of "-idea" (zoology) or "-aceae" (botany, Prokaryota). The intention of the person who names a species is signaled by that person pointing to a type specimen or series of type specimens. Such type specimens are meant to leave an objective record of the intention of the original describer; if no type specimen is designated, a later worker can designate a type specimen but cannot arbitrarily change the name. This is termed the Principle of Typification and can extend at various levels above the rank of species by designating type species, type genera, etc., as required by the Codes. These rules of nomenclature can be quite complex, leading to a group of systematists who specialize in understanding and interpreting the various Codes. Disputes can be brought to an international commission that has the power to make decisions to accept or change names as needed.

The various Codes of Nomenclature have survived and prospered for several reasons. (1) The names of species (*Homo sapiens*) and higher (more inclusive) taxa such as *Homo*, Mammalia, and Canidae are formed in Latin, a holdover of a time that all scholarship in Western society was written in Latin. As a "dead language," it is a convenient neutral language for taxonomists of different cultures and native languages. (2) Designation of hierarchical relationships through the use of ranks (genus, family, class) provides a reusable system of ranks. But perhaps the greatest strength is the fact that (3) the Codes are neutral as to biological meanings of the taxa. The Codes specifically state that their purposes are directed toward the use and formation of names, not the biological meaning or significance of the names. They leave it up to the biological community to make that determination. For example, to a phylogeneticist the name *Homo* is used to name a group of bipedal mammals that share a common ancestor unique to them (*Homo* is a unique evolutionary unit or monophyletic group), while to a pheneticist *Homo* might be used to name a group of bipedal mammals that are more similar to each other than to another group of bipedal mammals, like *Australopithecus*. Each generation of biologists can use or reject names based on their ideas of what constitutes a good taxonomy, but must work within the Codes; the Codes only govern which names are available for use. This has protected the Codes from changing ideas as to the nature of biological diversity while ensuring that names are used in a consistent manner that can be understood by everyone. *See* BACTERIAL TAXONOMY; PLANT NOMENCLATURE; TYPE METHOD; ZOOLOGICAL NOMENCLATURE.

Phylogenies: basic to modern taxonomy. The triumph of phylogenetic systematics and its attendant methods of discovering the evolutionary relationships among species have revolutionized taxonomy and biological classification. The principle insight provided by Hennig and his predecessors (such as the German systematist Walter Zimmerman) was that different homologies have different value in reconstructing phylogeny. Homologies are similarities in structure, behavior, and other characteristics that have a single evolutionary origin and thus have potential to act as "markers" of common ancestry. In contrast, homoplasies are similarities that have two to several evolutionary origins. For example, having legs is a homologous similarity marking the origin of tetrapod vertebrates such as ourselves. In contrast, homothermy, the ability to maintain a constant body temperature, is a homoplasous similarity observed in birds and mammals and does not have a single origin, given our current understanding of vertebrate relationships. One might think that all homologies are useful in reconstructing common ancestry relationships, and this is true, but not all are useful at the same time. For example, humans and frogs have multiple digits on their legs while living horses have a single digit. Horses and humans have hair, but frogs lack hair. Just given this, we might say that humans are as similar to frogs as they are to horses. But having more than one digit, while each is homologous with the other, is a character that appeared very early in the evolution of land vertebrates and is a characteristic of the common ancestor not only of frogs and humans but also of horses. In contrast, having hair evolved much later and is characteristic of the common ancestor of horses and humans but not of the ancestors of all three. Thus, we can conclude that horses and humans shared a more recent common ancestor and thus are more closely related. The presence of hair is a marker for all mammals sharing a common ancestor not shared with frogs. In short, all homologies can be used to reconstruct phylogenies but not at the same level of the hierarchy.

A large number of systematists now devote their research to reconstructing phylogenies. There are detailed protocols for determining which homologies are relevant to particular levels of the Tree of Life and which similarities are actually homoplasies and not homologies. Critical to such research are the 300 years of taxonomic research that have resulted in our present classifications. In some cases, the "phylogenetic signal" was so strong that clades (common

ancestry groups) recognized today were recognized by those who came before the rise of modern methods. After all, Aristotle himself recognized animals with backbones. In other cases, the clades were more subtle, or preconceived ideas about the distinctiveness of a group caused the classification to be different from what we now think are the common ancestry relationships (birds and crocodiles; great apes and humans). In these cases, the taxonomies must be changed to reflect the phylogenies. In other cases, there are no phylogenies, and we must rely on the present taxonomies. *See* PHYLOGENY.

Challenges to Linnaean taxonomy. Can a system of nomenclature invested by an eighteenth-century Swedish botanist survive the discovery of 20 million species of organisms and the attendant phylogeny of their relationships? This question is now being debated. Some have suggested that it cannot survive and have formulated an alternative system that dispenses with the current Codes and substitutes their own, the PhyloCode. Others suggest that the Linnaean system can be modified to suit the future needs of taxonomy. Given that major organizations that use taxonomy have adopted the Linnaean system, it is doubtful that any alternative system will soon replace it. For example, GenBank, the major repository of all genetic sequences, is organized around the system of Linnaean names and ranks, adhering to the present Codes. However, just as the debate over phenetics, evolutionary taxonomy, and phylogenetics invigorated systematic and taxonomic research, the debate over retaining the Linnaean system or replacing it is also healthy in stimulating debate over the form and function of modern taxonomy. E. O. Wiley

Bibliography. J. Cracraft and M. J. Donoghue (eds.), *Assembling the Tree of Life*, Oxford University Press, New York, 2004; W. Hennig, *Phylogenetic Systematics*, University of Illinois Press, Urbana, 1966; J.-W. Wägele, *Foundations of Phylogenetic Systematics*, Verlag Dr. Friedrich Pfeil, Munich, 2005; E. O. Wiley, *Phylogenetics: The Theory and Practice of Phylogenetic Systematics*, Wiley, New York, 1981.

Tea

A small tree (in cultivation, constant pruning makes it a shrub 3–4 ft or 0.9–1.2 m tall); a preparation of its leaves dried and cured by various processes; and a beverage made from these leaves. The plant, *Camellia sinensis* (or *Thea sinensis*), is an evergreen tree of the Theaceae family, native to southeastern Asia, and does best in a warm climate where the rainfall averages 90–200 in. (2200–5000 mm). The slower growth at higher altitudes improves the flavor. China, Japan, Taiwan, India, Sri Lanka, and Indonesia are among the leading tea-producing countries, with China contributing about one-half of the world's supply. According to Chinese folklore, the Chinese emperor Shen-mung discovered the use of tea about 2500 B.C. Tea leaves contain caffeine, various tannins, aromatic substances attributed to an essential oil, and other materials of a minor nature, including proteins, gums, and sugars. The tannins provide the astringency, the caffeine the stimulating properties.

Flowering branch of tea plant (*Thea sinensis*).

The tea plant normally produces a full crop of leaves (flush) about every 40 days. In China and Japan the leaves are plucked when the flush has mostly completed its growth, but in India and Indonesia the leaves are picked every week or two, resulting in most of the leaves being gathered at the best time. The leaves are named, beginning with the topmost, as pekoe tip, orange pekoe, pekoe, first souchong, second souchong, first congou, and second congou.

In its natural state the tea plant grows to a height of 15–30 ft (4.5–9 m). The leaves vary from $1^1/_2$ to 10 in. (4 to 25 cm) long and are thick, smooth, and leathery. The white, fragrant flowers are produced in the axils of the leaves (see **illus.**). The fruits are dark brown capsules. *See* THEALES. Earl L. Core

Diseases. Blister blight disease of tea was recognized in the midnineteenth century, but despite sporadic occurrence throughout the world, it was not considered a serious threat until it devastated plantings in Sri Lanka (Ceylon) and India during 1946. About 25,000 acres (10,000 hectares) of tea in Sumatra were destroyed 3 years later, and more recent reports of 30% losses attest to the destructive capability of the disease. Most serious losses attributed to blister blight have occurred in southeastern Asia.

The disease is almost entirely foliar, although young, green stems and buds may be attacked. Mostly round, translucent, sometimes pinkish spots lighter in color than surrounding leaf tissues enlarge to form noticeably concave depressions on the upper leaf surfaces. As seen from below, these convex "blister" surfaces colored light brown to white soon rupture to expose thousands of spores (with mucilaginous coverings) of the causal fungus *Exobasidium vexans*. These relatively fragile basidiospores, killed by exposure to 95°F (35°C) or 30 min of sunlight, are readily airborne and may germinate within 2 h at 80% relative humidity, but can remain dormant for as long

as 5 days. Infection follows germination within 24 h and, after 3–10 days incubation and 6–9 days for secondary lesion development, sporulation may occur again after 10–20 days. Production of basidiospores follows a diurnal cycle, with greatest concentration of spores in the air between midnight and 4 a.m. Disease development is retarded by low temperatures and the more intense solar insolation of higher altitudes, but frequent interruptions of full sunshine and prolonged periods of rainy weather at such elevations favor the disease.

High rainfall conditions are common to many tea-growing areas, and copper compounds (copper oxide or copper oxychloride) used either as sprays or dusts have been found most effective for control of blister blight. Disease forecasting, based on hours of sunshine, has been attempted experimentally.

Oil spot of tea is a condition known to occur only in Sri Lanka at elevations above 5000 ft (1500 m). There is presumptive evidence that the spotting symptoms seen on young leaves prior to defoliation result from a toxin produced by an unidentified fungus cultured from discolored wood below affected shoots. Early attempts at control with fungicides have been ineffective.

Curled leaves, shortened internodes, and zigzag growth of stems are often associated with phloem necrosis disease of tea. The only diagnostic symptom of the complex syndrome is necrosis of the phloem, seen as irregular or blotchy spotting on the inner phloem surface of bark removed by peeling. Known only from Sri Lanka, where incidence has ranged to 70%, phloem necrosis is almost completely restricted to elevations above 4000 ft (1200 m), and is more severe above 5500 ft (1650 m). The disease has been transmitted by grafting, but is not transmitted through soil. Plants placed at 82°F (28°C) have produced new shoots without symptoms, but the only practical control at present is to rogue diseased plants and replace them with healthy plants.

Chlorotic or brown leaves on dead branches or entire bushes are the obvious symptoms of collar and branch canker caused by *Phomopsis theae*. Bark cankers that usually develop on young plants result from infection through pruning or inadvertent injury to stems or branches. Application of fungicide to injuries has been recommended for control.

Symptoms of mottle scab (*Elsinoe theae*) include irregular to angular, dark lesions found only on mature leaves. Lesions caused by white scab (*E. leucospila*), also dark in color, occur on both surfaces of young leaves and later lighten to nearly white, enlarge to 0.04–0.12 in. (1–3 mm) in diameter, and become circular in shape. Both diseases are relatively uncommon.

Root diseases of tea, caused by *Poria hypobrunnea* (red root), *Rosellinia necatrix* (white root rot), *Helicobasidium compactum* (purple root rot), *Armillariella mellea* (Armillaria root rot), and *Ustulina zonata*, often remain undetected until the infection has spread to adjacent plants. No treatment adequate to rejuvenate infected plants has been devised, and control presently consists of roguing infected plants, fumigating of the planting site, and replanting with healthy stocks. *See* PLANT PATHOLOGY.

A. A. Cook

Tea production. Tea production technology depends on whether black, green, or semifermented tea is to be manufactured. The characteristic appearance and flavor of each arise more from different processing and manufacturing methods than from type of leaf.

Black tea. This tea requires quick transportation of the green leaf from garden to factory, where it is carefully spread on withering tats. The leaf loses moisture and becomes flaccid within 24 h. Circulation of conditioned air during withering shortens the time substantially. The next process, called rolling, ruptures the leaf cells and releases their juices. This is accomplished in a brass roller box, equipped with an adjustable pressure cap and an open base. The box is filled with withered leaf and rotated eccentrically over a brass table fitted with curved battens. During a 30-min period the leaf cells rupture under the pressure and the rubbing action. In a second rolling period the leaf is twisted and coated with juices, leading to balling. The leaf changes to a bright copper color, and a characteristic tea aroma develops. A roll breaker is subsequently employed to disintegrate the balls for fermentation.

In fermentation, enzyme action and oxidation occur. The rolled leaf is spread to a 2–3-in. (5–7.5-cm) depth on fermenting beds and held under high humidity at 75–80°F (24–27°C). Since the body, or strength, of the tea depends on this fermentation, careful attention to its progress is required. Approximately $3^1/_2$ h are needed for the rolling and fermentation operations.

Firing or drying of fermented tea leaf conventionally is carried out by passing it through heated ovenlike chambers. Temperatures of 170–180°F (77–82°C) arrest fermentation and develop the familiar blackish color. Drying is practically completed in one pass, but further heat is applied to produce the "case hardening" that protects quality. The leaf is sieved, graded, and packed into chests lined with aluminum foil. Each chest contains about 100 lb (45 kg) of finished tea.

Green tea. Green tea manufacture requires that the plucked leaves be steamed as quickly as possible. Such processing at 160°F (71°C) makes them soft and pliable and, by inactivating the natural enzymes, prevents fermentation. After steaming, the leaf is alternately rolled and dried until it becomes too stiff for further manipulation. For export, green tea is refired in pans with mechanical stirring to produce a luster.

Oolong teas. Oolong teas are midway between black and green teas in that they are semifermented. After a short sun-withering in the garden, the leaf is gently rolled in the plucker's hands, so a slight fermentation is initiated. After a short period this leaf is sent to the factory to be fired and packed for shipment.

Instant tea. This type of tea is obtained by spray-drying of a black tea extract, with or without the admixture of maltodextrins. The technology is

patterned closely after that followed in manufacturing of instant coffee. Concentrated tea extract also may be combined with a heavy sugar syrup and marketed as a liquid. Solubility of instant tea powder can be varied by manipulations in processing. *See* FOOD ENGINEERING. John H. Nair

Bibliography. A. Stella, *The Book of Tea*, 1992.

Technetium

A chemical element, Tc, atomic number 43, discovered by Carlo Perrier and Emilio Segrè in 1937. They separated and isolated it from molybdenum (Mo; atomic number 42), which had been bombarded with deuterons in a cyclotron. Technetium does not occur naturally. *See* ATOMIC NUMBER; ELEMENT (CHEMISTRY).

In the periodic table, technetium is located in the middle of the second-row transition series, situated between manganese and rhenium. Because of the lanthanide contraction, the chemistry of technetium is much more like that of rhenium, its third row congener, than it is like that of manganese. Its location in the center of the periodic table gives technetium a rich and diverse chemistry. Oxidation states −1 to +7 are known, and complexes with a wide variety of coordination numbers and geometries have been reported. *See* LANTHANIDE CONTRACTION; PERIODIC TABLE; RHENIUM; TRANSITION ELEMENTS.

1												13	14	15	16	17	18
1 H	2																2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

The most readily available chemical form of technetium is the ion pertechnetate (TcO_4^-), the starting point for all of its chemistry. In its higher oxidation states (+4 to +7), technetium is dominated by oxo chemistry, which is dominated by complexes containing one or two multiply bonded oxygen (oxo) groups.

Since about 1979, inorganic chemists have been very interested in understanding and developing the fundamental chemistry of technetium, utilizing the isotope ^{99g}Tc. Most of this interest has arisen from the utility of another isotope of technetium, ^{99m}Tc (where m designates metastable), to diagnostic nuclear medicine. In fact, ^{99m}Tc in some chemical form is used in about 90% of all diagnostic scans performed in hospitals in the United States. The nuclear properties of ^{99m}Tc make it an ideal radionuclide for diagnostic imaging. This technetium isotope has a 6-h half-life, emits a 140-keV gamma ray which is ideal for detection by the gamma cameras used in hospitals, and emits no alpha or beta particles. Silvia S. Jurisson

Bibliography. R. Alberto, *Technetium, Rhenium: Topics in Current Chemistry*, vol. 176, Springer-Verlag, 1995; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. Schwochau, *Technetium: Chemistry and Radiopharmaceutical Applications*, Wiley-VCH, 2000.

Technology

Systematic knowledge and action, usually of industrial processes but applicable to any recurrent activity. Technology is closely related to science and to engineering. Science deals with humans' understanding of the real world about them—the inherent properties of space, matter, energy, and their interactions. Engineering is the application of objective knowledge to the creation of plans, designs, and means for achieving desired objectives. Technology deals with the tools and techniques for carrying out the plans.

For example, certain manufactured parts may need to be thoroughly clean. The technological approach is to use more detergent and softener in the wash water, to use more wash cycles, to rinse and rerinse, and to blow the parts dry with a stronger, warmer air blast. Often such refinements provide an adequate action. However, if they do not suffice, the basic technique may need to be changed. Thus, in this example, science might contribute the knowledge that ultrasonically produced cavitation counteracts surface tension between immiscible liquids and adhesion between clinging dirt and the surface to be cleaned, and thereby produces emulsions. Engineering could then plan an ultrasonic generator and a conveyor to carry the parts through a bath tank in which the ultrasonic energy could clean them. The scientist may use ultrasonic techniques to determine properties of materials. The engineer may design other types of devices that employ ultrasonics to perform other functions. These specialists enlarge their knowledge of ultrasonics and their skill in using this technique not for its own sake but for its value in their work.

The technologist is the specialist who carries out the technique for the purpose of accomplishing a specified function, and extends knowledge and skill of ultrasonic cleaning by refinement and perfection of the technique for use on various materials soiled in different ways. Technological advances improve and extend the application to cleaning other parts under other conditions. *See* AERONAUTICAL ENGINEERING; CHEMICAL ENGINEERING; CIVIL ENGINEERING; ELECTRICAL ENGINEERING; ENGINEERING; ENGINEERING, SOCIAL IMPLICATIONS OF; ENGINEERING DESIGN; FOOD ENGINEERING; HUMAN-FACTORS ENGINEERING; MECHANICAL ENGINEERING; NUCLEAR ENGINEERING; PRODUCTION ENGINEERING; SCIENCE; SPACE TECHNOLOGY. Robert S. Sherwood; Harold B. Maynard

Tektite

A member of one of several groups of objects that are composed almost entirely of natural glass formed from the melting and rapid cooling of terrestrial rocks by the energy accompanying impacts of large extraterrestrial bodies. Tektites are dark brown to green, show laminar to highly contorted flow structure on weathered surfaces and in thin slices, are brittle with excellent conchoidal fracture, and occur in masses ranging to as much as tens of kilograms but are mostly much smaller to microscopic in size. The shapes of tektites are those of common fluid splash and rotational forms including drops, spheres, and dumbbells, unless they have been abraded together with surface gravels. A few tektites have shapes that are caused by two different heating events: the impact that melted the parent rock to form the glass, and a second event apparently due to reentry aerodynamic heating.

Occurrences, groups, and ages. With the discovery of tektites in the Soviet Union, there are now five major groups known: (1) North American, 3.4×10^7 years old, found in Texas (bediasites) and the Georgia Coastal Plains, with a single specimen reported from Martha's Vineyard, Massachussetts; (2) Czechoslovakian (moldavites), 1.5×10^7 years old, found both in Bohemia (green and transparent) and in Moravia (brown and turbid); (3) Ivory Coast, 1.3×10^7 years old; (4) Russian (irgizites; see **illus.**), 1.1×10^7 years old, found in the Northern Aral Region; and (5) Australasian, 700,000 years old, occurring notably in Australia, the Philippines, Belitung, Thailand, and numerous other localities. The North American, Ivory Coast, and Australasian tekites also occur as microtektites in oceanic sediment cores near the areas of their land occurrences. In the land occurrences, virtually all of the tektites are found mixed with surface gravels and recent sediments that are younger than their formation ages.

Tektites, in general, are not rare objects. Millions of tektites have been recovered from the Australasian occurrence. However, some of the more pleasingly colored specimens from Bohemia and Georgia are prized by collectors and command substantial prices, either as specimens or potential gem material.

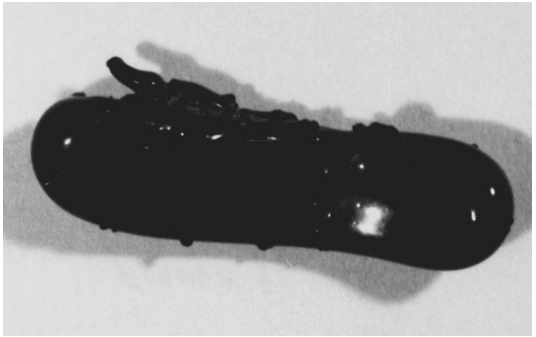

Dumbbell-shaped tektite (irgizite) from the Zhamanshin meteoritic impact crater, Northern Aral Region, Russia. The glass is so dark brown that it appears black in reflected light.

Three of the tektite groups are associated with known large impact craters. The irgizites occur in and immediately around the Zhamanshin meteoritic impact crater [10–15 km (6–9 mi) in diameter] in the Northern Aral Region of Russia. Both the Ivory Coast tektites and the moldavites occur close to probable source impact craters of identical age, the Bosumtwi Crater, Ghana [10.5 km (6.5 mi) in diameter] and the Ries Crater, Germany [24 km (15 mi) in diameter], respectively. The source craters for the North American and Australasian tektites have not yet been identified positively.

Composition. The chemical compositions of tektites differ from those of ordinary terrestrial rocks principally in that they contain less water and have a greater ratio of ferrous to ferric iron, both of which are almost certainly a result of their very high-temperature history. The extreme ranges of major element compositions, expressed as oxides in weight percent, are as follows: SiO_2, 48–85 wt %; Al_2O_3, 8–18 wt %; FeO, 1.4–11 wt %; MgO, 0.4–28 wt %; CaO, 0.3–10 wt %; Na_2O, 0.3–3.9 wt %; K_2O, 1.3–3.8 wt %; TiO_2, 0.3–1.1 wt %. However, portions of the foregoing ranges are based on analyses of glass particles from oceanic sediment cores that may or may not be tektites, as this identification is difficult for these very small particles. If other glass occurrences from impact craters were included as tektites, such as splash form glass from the Wabar Craters in Saudi Arabia, and the Lonar Crater in India, the ranges would expand significantly. As indicated by the wide range of chemical compositions, there are corresponding wide ranges of properties such as specific gravity and refractive index.

Inclusions. Spherical vesicles ranging in size from microscopic to as much as several centimeters are common in most tektites, as are small lechatelierite (silica glass) particles. The presence of coesite, a high-pressure polymorph of silica, in some tektites from Southeast Asia is additional evidence of the impact genesis of tektites. Coesite is known to form at the Earth's surface only from the very high pressures of transient shock waves caused by large hypervelocity meteorite impacts and cratering events. Baddeleyite, monoclinic ZrO_2, is present in some tektites as a high-temperature decomposition product of the mineral zircon. Meteoritic nickel-iron also has been observed in a few tektites, which is further evidence of the origin of tektites as meteoritic impact melts. Such inclusions of coesite, baddeleyite, and nickel-iron are common in the glassy fusion products in and around a number of terrestrial impact craters. *See* COESITE; METEORITE.

Historical perspective. Tektites have been recognized as unusual objects for more than a millennium, and they were the objects of intensive scientific research from the turn of the century until a few years before the return of lunar samples by the Apollo missions. This activity was generated by the possibility that tektites originated from the Moon as secondary ejecta from impact craters on the lunar surface. This

point of view was effectively advocated by a number of scientists until the recognition of the associated impact craters with some tektite groups on the Earth. Later analyses of lunar samples demonstrated that they are not suitable parent rocks for tektites. Unfortunately, the discovery of the Russian tektites, in a clear relation to an impact crater on the Earth, came too late to influence these arguments significantly.

Unanswered questions. The source crater of the Australasian tektites has not yet been identified, although this is the youngest and most widespread tektite group. Also, present understanding of the mechanics and sequence of events in very large impact crater formation does not permit specifying unambiguously the mode by which some tektites escape the Earth's atmosphere to reenter and form the remelted layer as a result of aerodynamic heating. Elbert A. King

Bibliography. G. B. Dalrymple, *P40 SAr: P39 SAr Age Spectra and Total-Fusion Ages of Tektites from Cretaceous-Tertiary Boundary Sedimentary Rocks in the Beloc Formation, Haiti*, U.S. Geological Survey, 1993; E. A. King, The origin of tektites: A brief review, *Amer. Sci.*, 65:212–218, 1977; E. A. King, *Space Geology: An Introduction*, 1976.

Telecast

A television broadcast, involving the transmission of the picture and sound portions of the program by separate transmitters at assigned carrier frequencies within the channel assigned to a television station. A telecast is intended for reception by the general public, just as is a radio broadcast. The picture may be either in black and white or in full color, using amplitude modulation, while the sound portion (in the United States) uses frequency modulation. The channels assigned for telecasts by the Federal Communications Commission, each 6 megahertz (MHz) wide, cover frequencies as follows: 54–72 MHz (channels 2 through 4), 76–88 MHz (channels 5 and 6), 174–216 MHz (channels 7 through 13), and 470–890 MHz (channels 14 through 83). *See* RADIO SPECTRUM ALLOCATION; TELEVISION; TELEVISION NETWORKS; TELEVISION RECEIVER; TELEVISION STANDARDS; TELEVISION TRANSMITTER. John Markus

Telecommunications civil defense system

When any possible emergency threatens the security or socioeconomic structure of the United States or any political subdivision, an appropriate warning must be made in order for a reaction to occur within an appropriate period of time. Any such reactions cannot operate without a means of communication to prewarn of impending disaster or attack and to assist the civil defense and military effort after a disaster has occurred. To meet these requirements in an integrated manner, the Federal Emergency Management Agency (FEMA) and the National Communications System (NCS), working with the telecommunications industry, provide warning centers and private-line telecommunications systems to carry vital emergency information and alerts to both the military and the general public.

Federal Emergency Management Agency. The Federal Emergency Management Agency has primary responsibility for financial and technical support of the two basic communications systems that meet this need during a crisis situation: the National Warning System (NAWAS) and the Emergency Broadcast System (EBS).

National Warning System. Operating 24 h a day, the National Warning System is a voice-grade service that consists of over 65,000 channel miles (100,000 km) connecting 2300 warning points in the United States with a national primary warning center located with the North American Air Defense (NORAD) Command. Should the primary warning center be rendered inoperative, alternate centers can cover all warning points that were dependent on the inoperative center. Similarly, modern switching facilities are available to replace any portion of the network that may become inoperative. The warning centers are located at key defense installations so that full advantage can be taken of all available information.

This warning system (originally called the Civil Defense Warning System) has been in existence since the late 1950s. The system is designed to bypass, insofar as practicable, major cities that are likely targets. Warning information can be sent to all points on the network in approximately 15 s. From the 2300 warning points on the network, warning information is relayed through state and local systems to more than 5000 local points in an average time of 7 min.

The National Warning System network consists of a control circuit and various warning (ring-down) circuits. The control circuit is used to pass administrative information and to coordinate operation. This circuit passes both voice and tone signals to activate switching devices that control national and regional circuit configurations. A selective signaling arrangement with predetermined codes is used to set up the various regional configurations. State circuits interconnect with regional circuits through a unique arrangement that permits state or local extensions to receive voice communication directly from regional circuits. Operation of a switching device at the regional-state interface transfers control of the state circuit to the state operator for two-way voice operation with other state and local-extension users. Many of the local users are fire stations, police, sheriffs, and other civilian agencies with an emergency control mission. Most regional, state, and local-extension National Warning System stations consist of a four-wire telephone instrument, a push-to-talk handset, a loudspeaker, and a bell which operates independent of the telephone instrument switchhook.

Emergency Broadcast System. The mission of the Emergency Broadcast System is to provide the President of the United States with a means of communicating with the general public through the use of

nongovernment broadcast stations during the period preceding, during, and following an enemy attack, natural disaster, or other national emergency. The Broadcast System consists of four multipoint private-line services routed to 31 locations at 2400 bits per second. Either of the two point-of-origin, send-receive stations (a primary location and its alternate) can transmit the activation, termination, and National Industry Advisory Council (NIAC) order messages to all other stations. Each point-of-origin send-receive station is equipped with a line controller unit, a 4420 terminal, diagnostic modems, and floppy disk storage units. All other receive-only stations are equipped with 2024 modems and a Model 43 buffered terminal which is programmed to provide an answer with the depression of a key. Activation of a switch at these locations and at the point of origin allows voice confirmation and emergency coordination.

To activate the national-level Emergency Broadcast System, the President gives an order to the White House Communications Agency duty officer, who notifies one of the two origination points by telephone or radio. At the origination point, the systems controller or a communications center assistant releases a message over the EBS Net. This message, containing the Emergency Activation notice and National Industry Advisory Council order number, reaches key offices of the broadcast networks, press agencies, and American Telephone and Telegraph offices. Calls among the White House Communications Agency, origination points, and industry can be authenticated. Authentication lists are furnished by the National Security Agency. The broadcast networks send the Emergency Activation Notice to their affiliated stations by internal network alerting systems. The press agencies, after confirmation, transmit a teletype message over their Radio Wire Teletype Networks. Broadcast stations transmit the Federal Communications Commission attention signal, and an emergency action broadcast announcement is made over the air.

National Communications System. The National Communications System was established in 1963 as a confederation in which federal departments and agencies participate with their telecommunications assets to provide essential communications under all conditions ranging from normal day-to-day situations to national and international crises, including nuclear attack. The principal assets of the National Communications System comprise the telecommunications networks of 22 federal agencies including the Federal Emergency Management Agency and the Department of Defense. The Secretary of Defense serves as the executive agent for the National Communications System with the director of the Defense Communications Agency designated as the National Communications System manager. The National Communications System manager performs various functions to support the mission of the National Communications System. Among these are directing a National Communications System/Defense Communications Agency Operations Center (NCS/DCAOC) and the National Coordinating Center (NCC).

National Communications System/Defense Communications Agency Operations Center. The Defense Communications Agency systems control officers and their staff at the center provide the National Communications System a 24-h daily point of contact for status on essential government networks and communications services. The center is connected to commercial telecommunications carriers' network operations centers, which provide status on a 24-h basis to the center regarding any significant failures or threatening conditions that might affect any Defense Communications service provided over their facilities.

Of the systems provided by commercial carriers, two of the most vital to the nation's defense are the Strategic Air Command's Primary Alert System (PAS) and the NORAD Alert System (NAS). The Primary Alert System is a voice alerting network specifically designed to serve the Strategic Air Command. Tests are conducted frequently to assure the viability of this system. Presently made up of two networks, the NORAD Alert System consists of the Primary Voice Alert System (PVAS), which connects principal military locations, and the Weapons Alert System (WAS) which is mainly used to alert fighter interceptor squadrons. For utmost in service reliability, hardened cable facilities, diversity, and redundancy are used in these systems. In addition, constant upgrades are made to the systems as new concepts and technology become available, which will further enhance reliability and survivability.

National Coordinating Center. Another major function that is under the National Communications System manager is the operation of the National Coordinating Center. This center was established in January 1984 and is the vehicle by which the federal government and the telecommunications industry jointly respond to National Security and Emergency Preparedness (NSEP) telecommunications requirements. Representatives from the telecommunications industry and federal government staff the National Coordinating Center. They provide for a rapid exchange of information and expedite responses during crisis situations.

The National Coordinating Center has the capability to suggest responses to a broad spectrum of emergency or crisis situations. The extent of its authority is governed by the nature of the emergency and the following specific legal conditions: after Section 606 of the Communications Act of 1934 is invoked; during disasters declared by the President; and during any emergency or crisis defined in the National Security Emergency Preparedness Procedure Manual distributed by the Defense Communications Agency. Activation of the National Coordinating Center is brought about by one of these legal conditions and any one of the following operational situations: (1) requests for assistance or coordination during an emergency are received from a telecommunications industry; (2) the federal government agency requesting an emergency expedite declares that existing procedures will not result in sufficient service;

or, (3) the scope of the National Security and Emergency Preparedness service or support requirements involves coordination across several states, regions, companies, or organizations. In a wartime situation where the assets of the telecommunications industry become nationalized, the representative in the National Coordinating Center would be responsible for carrying out directives, related to their specific company, consistent with federal government authority.

James P. McQuillan

Teleconferencing

Broadly, the various ways and means by which people communicate with one another over some distance. In a narrow sense, a teleconference is a two-way, interactive meeting, between relatively small groups of people (approximately 1 to 10 at each end), who usually use permanent teleconferencing facilities. A teleconference involves audio communication between the locations, but may also involve video, graphics, or facsimile.

Teleseminar. A teleseminar is utilized for educational purposes; it is primarily one-way communication to many destinations from one source. A teleseminar almost always uses audio communication, and may also use video and some form of graphics. The source location usually has just lecturers, and the destination locations are usually classrooms or meeting rooms with no more than 20–30 people in each location. There are typically 2–10 receiving locations. A means is provided for the receiving locations to ask questions of the instructor via microphones or telephone handsets. The facilities are usually built or converted especially for teleseminars. Many large corporations utilize teleseminars for continuing education purposes to keep a widely distributed work force equally informed on the latest topics.

More sophisticated training operations utilize viewer response systems. These small desktop devices let students respond to questions presented by the lecturer. At each classroom location, all such devices connect to a computer, which immediately tabulates the responses to the question and sends them back over a separate circuit to the lecturer. Such systems can be used to give tests or to provide feedback from students to teachers to see if they are really learning the material.

Telemeeting. A telemeeting is often called an ad hoc teleconference, with the ad hoc referring to places, times, participants, and purpose. A telemeeting is similar to a teleseminar in that it is primarily a one-way communication, usually staged or prepared by video program professionals. It may be set up to order, using temporary equipment or circuits, and typically utilizing hotel facilities. Events include stockholders' meetings, new product introductions, sales meetings, promotions, and press conferences, and thus involve large numbers of people at all locations. Audio and video are almost exclusively used to communicate and, as with a teleseminar, a means is usually provided at the receiving locations for questions to be asked of those at the transmitting location. Corporations that have many locations spread out across the country often lease satellite transponders and use them for telemeetings to stay in touch with their employees, disseminating information and answering questions.

Computer conferencing. Computer conferencing is a method for people to communicate by using computers. The medium is quite flexible, as it can be used between just two people, between one and many people, or among many people. Basically, computer conferencing involves typing a message on a computer terminal and transmitting it to one or more destinations electronically.

Networks such as the Internet or corporate "intranets" (wide area networks) or building networks (local area networks, or LANs) are required to accomplish computer conferencing between many users; the other extreme, a simple data modem and telephone circuit, can allow two people at a time to conference. Networks allow both real-time and non-real-time conferencing. Non-real-time conferencing is like an electronic mailbox, commonly called e-mail, whereby, for example, a paper is distributed to certain people via computer with a request for comments; each person could transmit comments via computer to the originator. Real-time conferencing is fairly common today between many users; it occurs via chat rooms hosted by Internet service providers, schools, companies, and many other organizations. Another form of real-time conferencing between two or more people is exemplified by the AOL instant messaging type of service. Computer conferencing is widely used within offices that contain networks of personal computers (LAN). Telecommunications links can extend such networks, and therefore computer conferencing, outside of one building to a number of geographically diverse locations (WANs). The building of the Internet has greatly facilitated computer conferencing. The spread of personal computers into households and the growth of on-line services, which have enabled computers to connect to the Internet, have stimulated other forms of computer conferencing, such as bulletin boards, list servers, and the sharing of digitized music and pictures. *See* DATA COMMUNICATIONS; ELECTRONIC MAIL; LOCAL-AREA NETWORKS; MODEM; WIDE-AREA NETWORKS.

Audio communication. As a minimum, teleconferencing audio systems utilize a device such as a speakerphone to allow hands-free operation. More typically, microphones are either clipped to people or attached to a table or ceiling, and loudspeakers are mounted in a wall, cabinet, or ceiling. Depending on the circuits utilized, the audio performance can vary greatly in quality with corresponding variations in cost and complexity. Scrambling is used in some systems to assure communications privacy. *See* COMMUNICATIONS SCRAMBLING; LOUDSPEAKER; MICROPHONE; TELEPHONE.

Video and videographic communication. Video systems use one or more cameras as image sources, and

television receivers, monitors, or large-screen projectors as a means of display. Cameras may be used to view people, objects, chalkboards or easels, slides, transparencies, or text on paper; video tape players and computers are also used as sources. Again, as a function of the communications equipment and circuits used, the nature and quality of the video images may range from that of broadcast television to that of slow scan or freeze frame. Slow scan paints a new picture on the screen in 10 to 60 s, but nevertheless succeeds in providing a perception of presence of people and good-quality videographics. High-speed digital compression codecs are also used to transmit images, which are near full motion, over telephone trunks. Audio and graphic information may also be sent through these codecs, thus allowing one circuit to be used for all forms of communication. *See* MAGNETIC RECORDING; TELEPHONE SERVICE; TELEVISION; TELEVISION CAMERA; TELEVISION RECEIVER; VIDEOTELEPHONY.

Multimedia conferencing. Video codecs have been reduced in size to fit inside personal computers. Low-resolution video cameras have been reduced to smaller than a computer mouse. Many personal computers contain high-fidelity sound cards that drive stereo speakers and allow microphone inputs. Computer modems provide the capability to store and dial telephone numbers. Personal-computer-based software has been written to integrate all of these functions, thereby turning a personal computer into a personal teleconferencing terminal. Such software typically divides a computer screen into quadrants, two of which are allocated for incoming and outgoing video and two of which are allocated for applications that are shared and worked on jointly during the teleconference; these applications may be computer based, such as word processing or spread sheets, or may be video graphics or facsimile. *See* MICROCOMPUTER; MULTIMEDIA TECHNOLOGY; SOFTWARE.

Graphic communications. Numerous devices are available that allow nonvideo graphics to be used in conferencing. These include electronic blackboards and writing tablets which digitize, transmit, and reproduce an exact replica of the source at the destination. Computers are also used to generate graphic material for conferences. The material can be created before a meeting, and with the computer connected into the communications system, the material can be transmitted to a similar computer and displayed at the distant end. Devices are available which allow the computer's screen to be projected onto a large screen, suitable for viewing by a conference room full of people. Computer graphics can thereby directly replace artist-generated graphics that must be turned into overhead transparencies. *See* COMPUTER GRAPHICS.

Facsimile communications. Equipment of this type basically consists of office facsimile units that are used in teleconferences to send standard-size pages of textual material between sites, either over telephone circuits or through codecs over high-speed circuits. There are more elaborate versions of the equipment available which also allow the transmitted page to be displayed via a monitor or projector at both the transmitting and receiving locations. *See* FACSIMILE.

Popular usage. Currently, the most popular form of teleconferencing by far is the audio conference. Using plain speaker-telephones, special speaker phones, corporate private branch exchange (PBX) systems, or special dial-in, conferencing services, most business people use this form of teleconference regularly. Probably next in popularity is the Teleseminar, used for "distance learning," or formal education and training. After that, the chat room and instant messenger service are the ones frequently used. The Internet has become a popular medium to facilitate all of these forms of teleconferencing, and will become more popular as high-speed connections such as digital subscriber line (DSL), cable modem, and satellite become more prevalent. *See* COMMUNICATIONS SATELLITE; INTERNET; PRIVATE BRANCH EXCHANGE.

John J. Bleiweis

Bibliography. J. L. Creighton and J. W. R. Adams, *Cybermeeting: How To Link People and Technology in Your Organization*, 1997, 2003; K. Kelleher and T. B. Cross, *Teleconferencing: Linking People Together*, 1985; B. C. Lindberg, *Future of Communications Teleconferencing—Audio, Graphics, Data and Video*, 1998; E. Margulies, *Audio Teleconferencing—The Complete Handbook*, September 1997.

Telegraphy

A method of communication employing electrical signaling impulses produced and received manually or by machines. Telegraph signals are transmitted over open wire or cable land lines, submarine cables, or radio. Telegraphy as a communication technique uses essentially a narrow frequency band and a transmission rate adapted to machine operations. *See* ELECTRICAL COMMUNICATIONS.

Early equipment devised by Samuel F. B. Morse consisted of a mechanical transmitter and receiver or register. Operators soon learned to handle messages faster by using simple manual keys and audible sounders. Subsequently, telegraph transmission and reception became mechanized. Telegraphy may also be used in other ways.

Telegraph facilities for use by the general public to transmit messages both domestically and internationally are provided by communication companies and government administrations. Special telegraph facilities include those for news services, distribution of market prices of securities and commodities, and private lines between such points as the factories and offices of a company for the exchange of messages, orders, payroll data, and inventories. Fire and police alarms are a special form of telegraphy. The armed forces have extensive fixed and mobile telegraph systems.

Telegraph codes. For manual operation, the code consists of short dot and long dash signals (**Fig. 1**). The original Morse code also used various-length

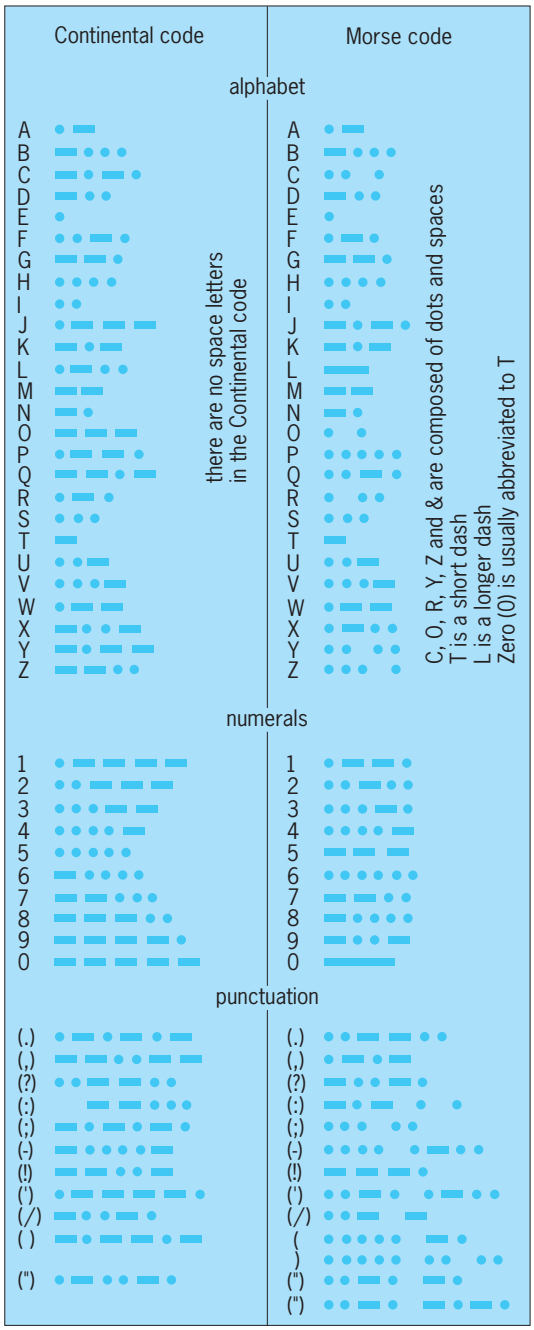

Fig. 1. Continental code is commonly used for telegraph communication. Morse code continues in use on a few land lines in United States and Canada.

spaces; the Continental code avoids them. On submarine cables, the dots and dashes are of equal length for most efficient use of transmission characteristics and are distinguished by being of opposite electrical polarity.

Most automatic printing telegraph circuits, including American cable operation, use a code of five equally spaced signals or units per letter or other symbol perforated into a paper tape (**Fig. 2**). The presence or absence of current or current reversals during these intervals constitutes the distinguishing feature. When a perforated tape is used, it usually is driven by a sprocket running in holes between the second and third code unit. Machines translate automatic teleprinter code to cable code, cable code to automatic code, or, for special applications, make other translations. For stock quotation systems and teletypesetter operation, a six-unit code is used to provide control of machine action. For data transmission or machine control, a seven- or eight-unit code may be used.

Telegraph circuits and equipment. A single circuit provides transmission in only one direction at a time. For transmission in both directions simultaneously, a duplex circuit is used. Multiplex (frequency-division) apparatus provides two, three, or more channels operable in both directions simultaneously over a single circuit. Carrier-current and frequency division techniques enable several circuits, each comprising one or more communication channels, to operate through the same wide-band wire, cable, or radio facility. *See* ELECTRICAL COMMUNICATIONS; MULTIPLEXING AND MULTIPLE ACCESS.

In automatic transmission, an operator at a manual keyboard, operated like a typewriter, perforates a tape. The tape is fed through an electromechanical or photocell tape reader that drives the tape at a rapid and uniform rate and transmits the electrical code. Interconnecting wire lines or other communication channels carry these code pulses to the receiver. The received impulses may automatically actuate a reperforator to produce a duplicate punched tape for retransmission or later transcription; the impulses may actuate a teletypewriter (also called a teleprinter) to retype the original message; or they may actuate other terminal equipment, such as an accounting machine.

Alternatively, the equipment at both terminals may be teletypewriters, in which keyboard and printing mechanisms are combined in one machine. Such machines use a five-unit code but operate without perforated tape. Business firms that originate and receive numerous messages have such equipment installed at their offices for direct service. Also widely used are facsimile instruments that transmit or reproduce a typed or handwritten message as a picture on electrosensitive paper. *See* FACSIMILE; TELETYPEWRITER.

Grosvenor Hotchkiss; William R. Webster

Message service. Individual circuits are established as needed from a nationwide network of central and branch telegraph offices and interconnecting wires, coaxial cables, and radio relay stations. The total traffic capacity of each of these interconnecting facilities may be divided by frequency-division multiplex or time-division multiplex into individual telegraph channels. Such networks of national telegraph companies and government administrations are interconnected for worldwide service.

Automatic switching. To achieve speed and accuracy in handling messages en route, direct circuits may be set up from the originating keyboard to the distant teleprinter. This is done in telex services. This mode of system operation is known as circuit switching. In previous times, if direct channels were not immediately available, or if the volume of traffic required a storage interval, the message

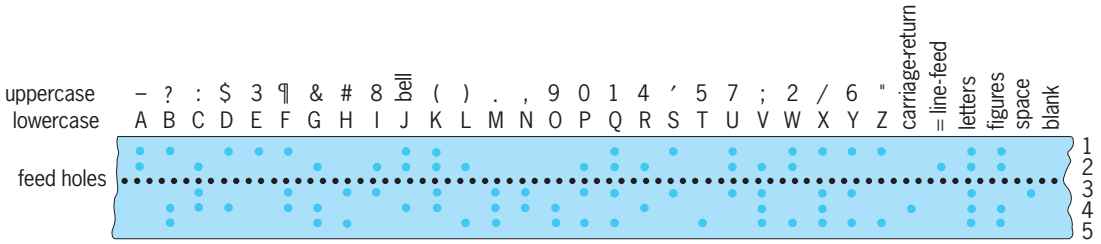

Fig. 2. Code for automatic telegraphy as it appears on a punched tape. (*Western Union Telegraph Co.*)

could be transferred by means of perforated tape at switching centers. This mode is known as message switching or store-and-forward switching. Message switching has the advantage of using the channel capacities of trunk lines more fully than do direct connections. This function is now performed by computers at central sites.

An example of message switching is the route followed by a telegram in previous times from Providence, Rhode Island, to Walla Walla, Washington. In the Western Union telegraph system, the United States had 11 area switching centers, all directly and continuously interconnected. The message from Providence went to the area switching center at Boston, Massachusetts, preceded by the director signal code PR perforated by the Providence operator. An automatic sensor at Boston received this code (meaning Portland, Oregon), actuated equipment to find an idle reperforator in the Portland trunk group, and released the fully prepared telegram from Providence into the Boston storage for Portland. After the message was completely stored in Boston, its reperforator signaled its availability to a seeker switch, the function of which was to connect reperforators with idle trunks. In this way the message was transmitted to Portland, where it was again perforated and punched on tape. The message then appeared before the Portland operator who, seeing the Walla Walla address, routed it by push button to its destination, where it was received on a teleprinter.

The reperforator switching network just described is based upon electromechanical implementation of the route-selection and transmission functions. Modern general-purpose digital computers have replaced this electromechanical switching equipment. Suitably programmed computers perform the switching function with more speed and versatility. When they are not busy routing messages, they are used to do billing, generate traffic statistics, or perform any other off-line task related to the communication operation. *See* DATA COMMUNICATIONS; DIGITAL COMPUTER. W. E. Girardin; William R. Webster

Overseas communications. In overseas communications, telegraph messages or telegrams are transmitted over high-frequency radio (3–30 MHz), transoceanic submarine cables, or satellite communication channels. A decreasing number of messages are carried on high-frequency radio. These types of overseas channels are also used for private-line service, which is referred to as leased-channel service, and customer-to-customer teletypewriter service, which is known as Telex in the international service. *See* COMMUNICATIONS SATELLITE; RADIO-WAVE PROPAGATION; SUBMARINE CABLE. Eugene D. Becken; William R. Webster

Bibliography. N. S. Biswas, *Principles of Telegraphy*, 1964, reprint 1968; R. L. Freeman, *Telecommunications Transmission Handbook*, 4th ed., 1998; J. Lehnert, *Introduction to Telegraph Engineering*, 1977; J. D. Ralphs, J. E. Flood, and C. J. Hughes (eds.), *Principles and Practice of Multi-Frequency Telegraphy*, 1985; D. A. Tugal and O. Tugal, *Data Transmission*, 2d ed., 1989; L. Wiesner, *Telegraph and Data Transmission over Shortwave Radio*, 3d ed., 1984.

Telemetering

The branch of engineering, also called telemetry, which is concerned with collection of measurement data at a distant or inconvenient location, and display of the data at a convenient location. One example of a complex telemetering system is used to measure temperature, pressure, and electrical systems on board a space vehicle in flight, radio the data to a station on Earth, and present the measurements to one or several users in a useful format. A simpler form of telemetering would be measuring temperature, fuel level, battery voltage, and speed in a car, and then via a hard-wire link displaying that information on the instrument panel. Telemetering involves movement of data over great distances, as in the above example, or over just a few meters, as in monitoring activity on the rotating shaft of a gas turbine. It may involve less than 10 measurement points or more than 10,000.

Telemetering involves a number of separate functions: (1) generating an electrical variable which is proportional to each of several physical measurements; (2) converting each electrical variable to a proportional voltage in a common range; (3) combining all measurements into a common stream; (4) moving the combined measurements to the desired receiving location, as by radio link; (5) separating the measurements and identifying each one; (6) processing selected measurements to aid in mission analysis; (7) displaying selected measurements in a useful form for analysis; and (8) storing all measurements for future analysis.

Aerospace systems. The largest category is commonly called aerospace telemetry, used in testing developmental aircraft and in monitoring low-orbit space vehicles. Other applications include missile

and rocket testing, automobile testing, and testing of other moving vehicles. Because of the broad range of applications for this type of system, military test ranges in the United States, through the guidance of the Range Commanders Council-Telemetry Group (RCC-TG), maintain ever-evolving Inter-Range Instrumentation Group (IRIG) standards. These standards include the IRIG 106, which deals with transmitters, modulation techniques, data formatting, data recording, and other related topics; as well as the IRIG 118 family of standards, which define test procedures for certifying telemetry equipment such as radio-frequency (RF) receivers. These documents are the standard for aerospace telemetry in the United States, and have become the de facto standard for aerospace telemetry manufacturers and users throughout most of the world.

Transducers. An aerospace telemetry system (**Fig. 1**) may send a very small number of measurements or as many as several thousand. It may send just a few measurement points per second or more than a million. In general, transducers accept various types of physical measurement variables, which may include temperature, pressure, strain, vibration, acceleration, shock, liquid or gas flow, liquid level, torque, and radiation. Each physical variable is converted into an electrical variable (current, voltage, resistance, or capacitance, for example) whose magnitude is proportional to the magnitude of the physical measurement.

Signal conditioners. For each electrical variable, a signal conditioner generates an output voltage with a standard range (−5 V to +5 V, for example), where the output represents the magnitude of the related physical measurement variable and follows it; the lowest expected physical magnitude may be represented by a −5-V output, the highest by a +5-V output, and other magnitudes in direct proportion.

Data multiplexer. All conditioned voltages are applied to a data mutliplexer, which combines all measurement voltages into a common link for movement to the receiving station. The commonly used methods of data combining are frequency-division multiplexing (where measurements are placed on individual frequency-modulated subcarriers located at different frequencies for transmission) and time-division multiplexing (where measurement voltages are sampled or commutated and placed in a specific time sequence for transmission). Since measurements occupy unique frequency or time slots, they do not interfere with each other and can be separated and identified at the receiving station. *See* MULTIPLEXING AND MULTIPLE ACCESS.

Pulse-code modulation. Most time-division multiplexing systems use pulse-code modulation (PCM), where each sampled measurement point is converted into a binary number, typically with 10 or 12 bits of resolution, and shifted out one bit at a time in a continuous data stream. In pulse-code modulation, a unique pattern signals the beginning of each data sequence or frame. *See* PULSE MODULATION.

Generally, where pulse-code modulation is used as the modulation scheme, the system not only accepts encoded analog data but also accepts information in digital formats. The most common source of such information is an on-board digital computer, but other sources may include on-off occurrences or other types of data. Generally, a system is programmed to sample each measurement point five or more times for each cycle of the highest expected data frequency.

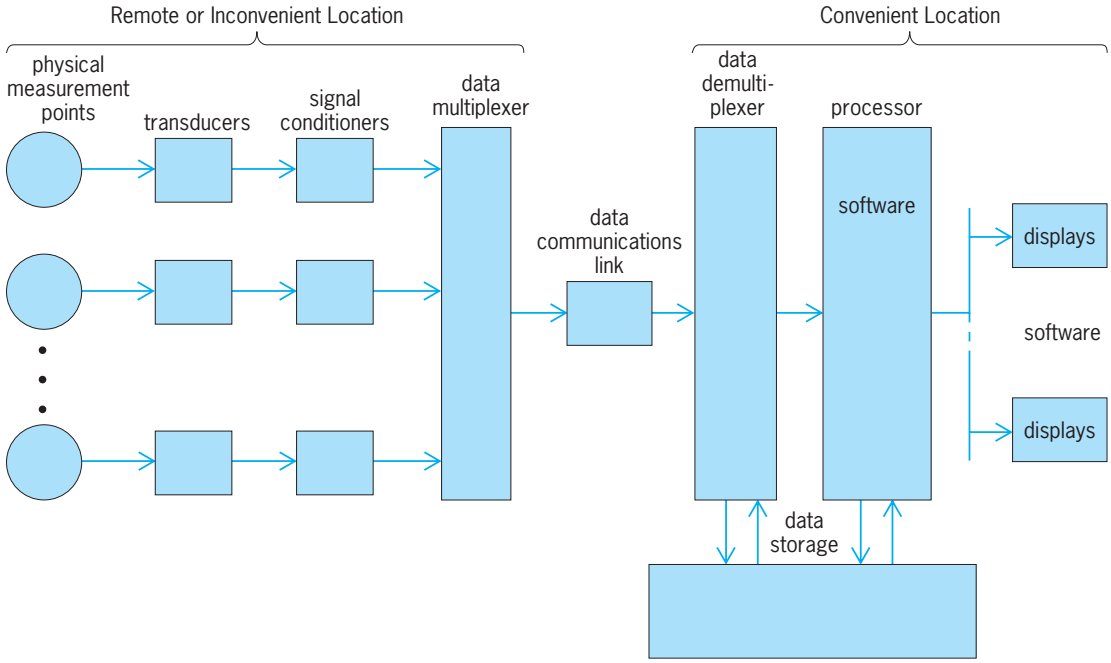

Fig. 1. Functions of a generic aerospace telemetry system.

Modulating a frequency-modulated (FM) carrier with non-return-to-zero (NRZ) PCM has been the standard for a number of years. With trends toward higher bit rates and less allowable bandwidth for telemetry users, alternate, more efficient modulation techniques have been developed. IRIG 106 now specifies Feher Patented Quadrature Phase Shift Keying-B (FQPSK-B), Feher Patented Quadrature Phase Shift Keying-JR (FQPSK-JR), Single Offset Quadrature Phase Shift Keying (SOQPSK), and Advanced Range Telemetry Continuous Phase Modulation (ARTM CPM). A 5-megabit-per-second (Mbps) data stream, for example, would require a transmission bandwidth of 12.5 MHz using NRZPCM, whereas it would require only 6.5 MHz using FQPSK-B, FQPSK-JR, or SOQPSK, and only 5 MHz using ARTM CPM. *See* FREQUENCY MODULATION; PHASE MODULATION.

Data communications link. The data communications link moves frequency-division-multiplexed or time-division-multiplexed data from the remote or inconvenient location to the convenient location by one of several methods. The most common link uses radio, where a transmitter's output is frequency-modulated or phase-modulated by the multiplexed data. Most aerospace users throughout the world operate in the 1435–1540-MHz band, the 2200–2290-MHz band, or the 2310–2390-Mhz band. While telemetry signals may be transmitted within these bands, telemetry is the primary user only through 1435-1525 MHz, 2200-2290 MHz, and 2360-2390 MHz. Telemetry signals may be transmitted in the other portions of these bands on a noninterfering basis. Alternatively, the data communications link may use fiber optics, telephone, or other communications media.

Where radio is used, generally the electrical power at the data collection site is limited, so the radio transmission power must be as low as practical (typically 2–5 W on a small missile, and 5–10 W on an airplane). To compensate for this relatively low radiated power, often a high-gain parabolic dish antenna is used at the receiving station. Further, when the vehicle under test is moving, it is necessary to steer the receiving antenna automatically or manually to follow the vehicle and thus receive data continuously as the test progresses. *See* DATA COMMUNICATIONS; RADIO SPECTRUM ALLOCATION.

Data demultiplexer. At the receiving station, a data demultiplexer identifies each measurement from the data communications link. For frequency-division multiplexes, band-pass filters are tuned to select the subcarriers, and each one is demodulated to yield measurement voltages for all channels. (The equipment is called a discriminator.) For time-division multiplexes, a frame synchronizer recognizes the unique pattern which signals the beginning of each data sequence, and counts samples to identify all measurements. (The equipment is called a decommutator.) *See* ELECTRIC FILTER; FREQUENCY-MODULATION DETECTOR.

Examination, processing, and display. In almost every aerospace telemetry system, data are routed to a processor for examination, processing, and display. Examination may consist of comparing each measurement value with operator-defined upper and lower limits and generating an alarm when any measured value goes outside those limits. Processing usually includes converting each received measurement value into a new number which denotes its actual physical value, such as degrees Celsius. Often it involves also frequency analysis or other unique calculations. Measurement displays may consist of numerical or graphical representations as selected by the individual users.

In many systems the major processing load is in a separate hardware device known as a preprocessor. This high-speed special-purpose device is under control of the computer with its unique telemetry software to ensure operation with minimum operator error. It is equipped to perform the 100 or more processing functions which are used most often in telemetry system applications, thus relieving the system computer of routine time-consuming examination and processing. With the advent of personal computers with clock speeds over 2 GHz, a lot of functions formerly performed in hardware are now being performed in software. The use of software allows for changing processing algorithms quickly as the need arises.

The system's processing power can be contained in one device, but in medium and large systems it is usually distributed between a general-purpose computer and one or more computer-powered workstations, each controlled independently to meet the needs of its user. Interconnection of the computer and display stations is by an industry-standard local-area network, such that display sites can be added or removed easily as requirements change. *See* LOCAL-AREA NETWORKS.

System software is built on a standard operating system, with special telemetry applications programs to handle the data rates and versatile processing requirements. A system uses two special databases: one defines all characteristics of the incoming data format, such as the modulation type, data rate, and unique synchronization pattern; and the other defines virtually all aspects of each measurement being received and processed, such as the measurement name (such as temperature), units of measurement (such as degrees Celsius), relationship between the physical units of measurement and the related telemetry codes being received, and other details which may be necessary for high-speed processing and display. The software also performs system diagnostics and reports on the operating condition of each component whenever electrical power is applied or on special command. *See* OPERATING SYSTEM; SOFTWARE.

Data storage. To archive data for further analysis, most systems include data storage, possibly consisting of a wide-band magnetic tape recorder at the data demultiplexer section and of a magnetic disk recorder at the processor section. When data are stored, they are annotated with the time of day so that they can be analyzed with respect to time. *See* COMPUTER STORAGE TECHNOLOGY.

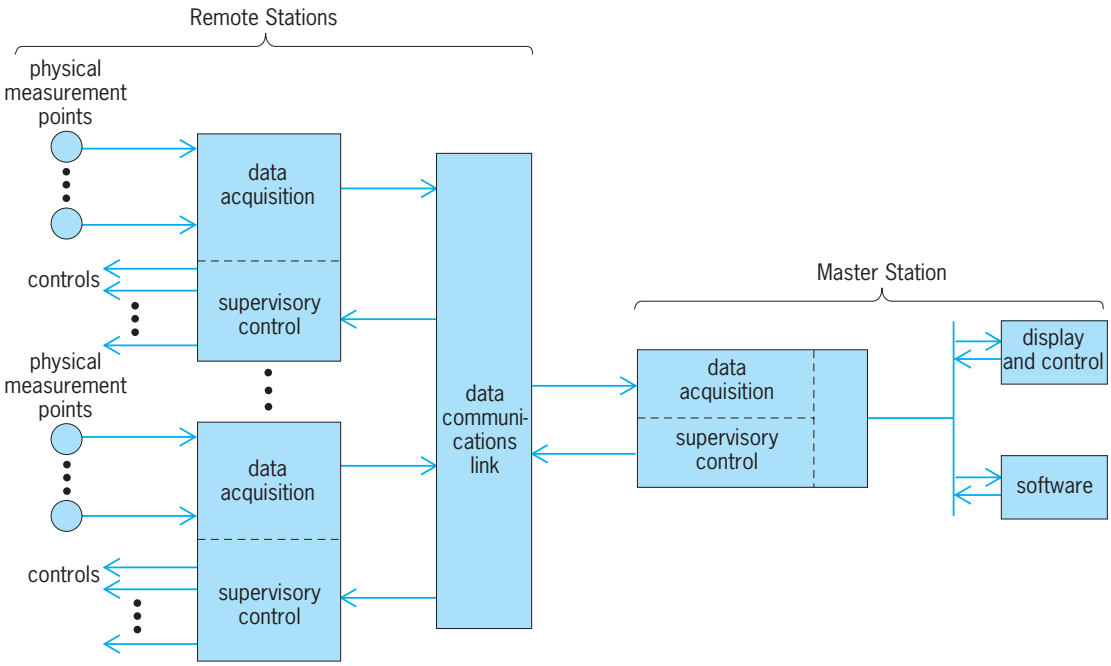

Fig. 2. Functions of a supervisory control and data acquisition (SCADA) system.

Industrial systems. The version of telemetry commonly used in an industrial application includes supervisory control of remote stations as well as data acquisition from those stations over a bidirectional communications link. The generic term is supervisory control and data acquisition (SCADA); the technology is normally used in electrical power generation and distribution, water distribution, and other wide-area industrial applications. There is no authoritative SCADA standards committee; nevertheless, these systems have many characteristics in common. *See* ELECTRIC POWER SYSTEMS; WATER SUPPLY ENGINEERING.

The number of remote stations in a SCADA system (**Fig. 2**) may be one, several, or even 100 or more. The number of data measurements per station varies over a wide range, as does the number of unique commands which can be delivered. Data or command transfer rates are quite low as compared to aerospace telemetry; generally they are 10–1000 per second.

In a SCADA system, data acquisition at a remote station is quite similar to the same function in an aerospace telemetry system, except that measurements are usually encoded into American Standard Code for Information Interchange (ASCII) characters for transfer, rather than frequency-division multiplex or time-division multiplex formats. These ASCII characters can be transferred as individual messages or grouped as a sequential message.

Data acquisition equipment at the master station brings these characters in and prepares them for the display stations, where each message is processed and displayed for the user in meaningful units for analysis. A user at any display station can initiate control messages for any remote station. These messages are relayed through the supervisory control equipment at the master station, and received by the equipment at the appropriate remote station to close or open switches, move rotary devices, initiate data acquisition, or perform other supervisory tasks. Since ASCII characters are compatible with a wide range of data communications equipment, the two-way data communications link can be commercial-grade telephone line, electric utility line, or low speed radio link. *See* REMOTE-CONTROL SYSTEM.

Each SCADA system is powered by one or more computers with versatile software. In many systems, the remote stations are computer powered also. SCADA technology is quite similar to aerospace telemetry technology in many areas. The major differences are in formatting and in the use of two-way communications.

One SCADA application involves monitoring wind direction and velocity as indicated by anemometers located on the approach and departure paths near an airport, so that air-traffic controllers can make pilots aware of dangerous differences in wind direction and velocity, known as wind shear. This type of system is operated over a radio communications link, with the appropriate anemometers being interrogated as their measurements are needed by the computer for wind analysis. *See* AERONAUTICAL METEOROLOGY; WIND MEASUREMENT.

Somewhat similar systems are used in oceanographic data collection and analysis, where instrumented buoys send water temperature and other data on command. *See* INSTRUMENTED BUOYS.

Space systems. Because two-way communication with a complex and distant Earth synchronous satellite or other spacecraft presents a unique challenge, a technology called packet telemetry is in widespread use for these applications. Here, messages between the Earth station and the spacecraft are formed

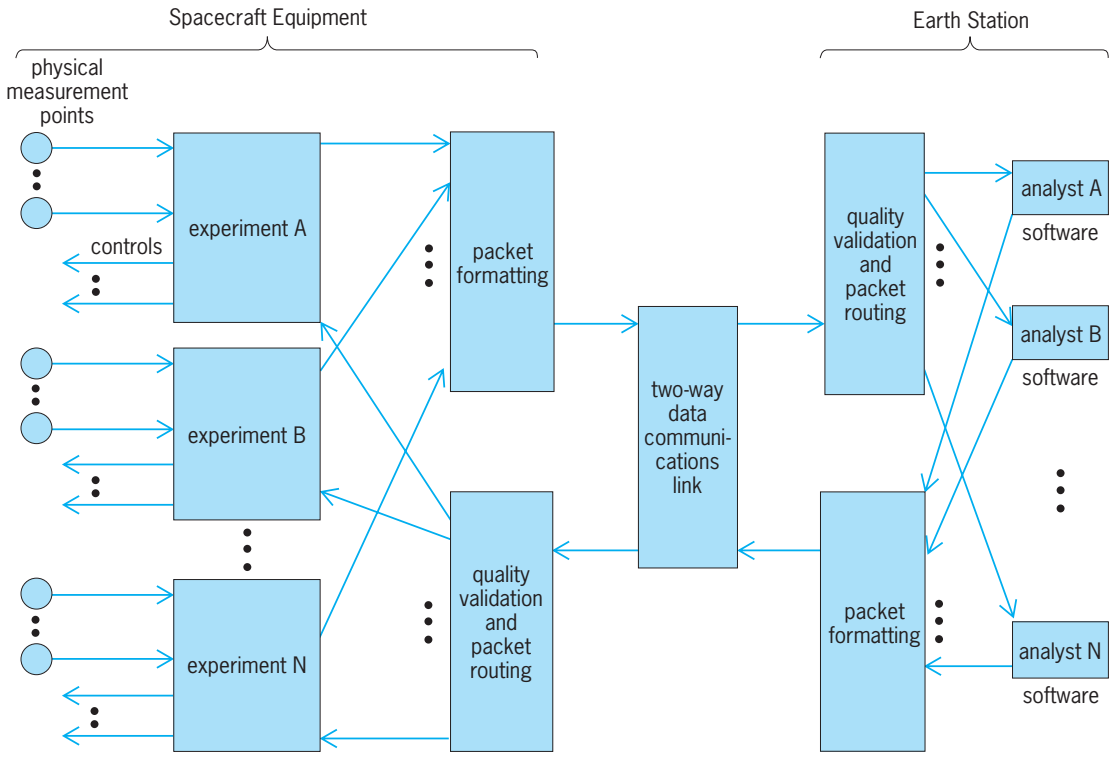

Fig. 3. Functions of a spacecraft telemetry and command system, using packet telemetry.

into groups of measurements or commands called packets to facilitate routing and indentification at each end of the link (**Fig. 3**). Each packet begins with a definitive preamble and ends with an error-correcting code for data quality validation. Packet technology is defined by an international committee, the Consultative Committee for Space Data Systems (CCSDA). *See* PACKET SWITCHING.

A spacecraft may have several on-board devices (labeled experiment A, experiment B, and so forth), each of which can accept commands from and send data to the Earth station. Several analysts are at the Earth station. Typically analyst A is concerned with experiment A, analyst B with experiment B, and so forth. When analyst A sends a command, it is put into a standard packet and sent to the spacecraft via a two-way communications link. The quality of the received packet is validated, and it is routed to the proper experiment for execution. Responses generated by the command are sent through the spacecraft's packet formatter and communications link to the Earth station. After quality validation, the response is routed to the proper analyst for processing and display. The packet technique provides optimum efficiency in use of the two-way communications link, as each analyst has access to the related experiment, subject only to occasional delays in link access.

Much of the technology in a space telemetering system is similar to that used in aerospace or industrial systems. The major differences relate to formatting, validation, and routing of packets.

Spacecraft telemetry must operate with limited electrical power. To overcome such limitations, a spacecraft may have a steerable antenna, typically 1–5 m (3–15 ft) in diameter, on board to boost the effective radio power. Often it sends data by phase-modulated rather than frequency-modulated radio, and typically it operates at a relatively low data rate in order to get the most efficiency in transmission. The ground-based station is likely to have an extremely large high-gain parabolic dish antenna, possibly as much as 30 m (100 ft) in diameter. *See* SPACE COMMUNICATIONS; SPACECRAFT GROUND INSTRUMENTATION.

Other systems. The three major families of telemetering systems (aerospace, industrial, and space) each use equipment in which simplicity is less important than standardization. Often a user is willing to use equipment which has higher performance than the application requires but which is readily available from one or more manufacturers. Certain special applications fall outside the three families, however, and some of those users have chosen to build special equipment which is optimum for their needs.

One unique example is the telemetering system used to monitor the condition of individuals in a cardiac care unit of a hospital. A single channel data-acquisition unit sends heartbeat data from a given individual over a low-powered radio transmitter (operating in the 460–470-MHz band) to a central receiving station nearby. There, a specialist receives and monitors the heartbeats of all individuals simultaneously and is able to respond quickly if a problem occurs.

A special multichannel medical telemetry system is used on some ambulances to monitor and radio vital signs from a person being transported to a

hospital, so that medical staff can prepare to treat the specific condition which caused the emergency.

Another unique system is possibly the oldest user of radio telemetry, the radiosonde. A data collection and transmission system is lifted by a balloon to measure and transmit pressure, temperature, and humidity measurements from various altitudes as an aid to weather prediction. *See* BALLOON; METEOROLOGICAL INSTRUMENTATION.

A system using telemetry multiplexing and recording technology is utilized on most operational commercial aircraft for recording selected data measurement points. In case of crash, the recording medium from this flight data recorder can be recovered from the crash scene and analyzed to help determine the cause. O. J. Strock; Tomas C. Chavez

Bibliography. F. Carden, R. Jedlicka, and R. Henry, *Telemetry Systems Engineering*, 2002; S. Horan, *Introduction to PCM Telemetering Systems*, 2d ed., 2002; International Telemetering Conference, *Proceedings*, annually; E. M Rueger and O. J. Strock, *Telemetry System Architecture*, 3d ed., 1995.

Teleostei

The largest, youngest (first appearing in the Upper Triassic), and most successful group of the actinopterygians (rayfin fishes), and a sister group of the Amiidae (bow fins). The 23,600 species of teleosts make up more than half of all recognized species of living vertebrates, and over 96% of all living fish species. There are 494 families, of which 69 are extinct leaving 425 extant families, 43% of which have no fossil record. However, the fish beds at Monte Bolca in Verona, Italy (lowermost Eocene), provide the earliest records for a large number of teleost higher taxa. *See* ACTINOPTERYGII; AMIIFORMES.

Morphology. Much of the evidence for teleost monophyly and relationships comes from the caudal skeleton and concomitant acquisition of a homocercal tail (upper and lower lobes symmetrical). This type of tail primitively results from an ontogenetic fusion of centra (bodies of vertebra) and the possession of paired bracing bones located bilaterally along the dorsal region of the caudal skeleton, derived ontogenetically from the neural arches (urolneurals) of the ural (tail) centra. The presence of uroneurals is a synapomorphy (homology) for all teleosts and is recognizable in the fossil groups, as is another synapomorphy—the articulation of the first two hypurals with the same vertebral centrum (ural centrum 1, which results from an ontogenetic fusion of two centra). Other characters of the teleosts include the mobile premaxilla bone, the extension of the posterior myodome (the eye muscle canal) into the basioccipital bone, and the development of the swim bladder.

Phylogeny. By the end of the Jurassic, several of the taxa with Recent representatives had evolved; these include the Osteoglossomorpha (bonytongues; *Arapaima*, mormyrids and allies; 217 species), Elopomorpha (tarpons, eels, halosaurs; 801 species), and Clupeomorpha (herrings; 360 species), but it was in the succeeding Cretaceous Period that the great radiation of the group occurred, resulting in some 20 or more orders (see **illus.**). *See* GEOLOGIC TIME SCALE; OSTEICHTHYES.

Since the Late Cretaceous, the teleosts have been by far the largest and most diverse group of vertebrates, occupying a plethora of habitats from the deepest ocean abyss to the highest mountain lakes—all of which makes the unraveling of their evolutionary history a daunting task. Consequently, ichthyologists are still actively involved in reconstructing teleost phylogeny, particularly that of the euteleosteans.

Classification. The Osteoglossomorpha, Elopomorpha, and Clupeomorpha are now generally regarded as successive clades (groups) above the level of the fossil, paraphyletic pholidophorids. The Clupeomorpha is considered to be the sister group of the Ostariophysi.

Clupeomorpha and Ostariophysi. The Clupeomorpha includes almost 80 genera and some 360 Recent species in three main families: Engraulidae (anchovies), Dussumieridae (round herrings), Clupeidae (herrings), as well as Pristigasteridae. The Ostariophysi make up nearly 75% of the fresh-water fishes of the world (over 6000 species) and include the Cypriniformes (carps, loaches, and relatives), Siluriformes (catfishes), Gymnotiformes (knife fishes, electric eel), and *Chanos*. They are characterized by the development of a bony chain of ossicles and associated structures connecting the swim bladder with the inner ear, the so-called Weberian apparatus. Interestingly, in the clupeomorphs, there is a direct connection between the diverticulum of the swim bladder and the labyrinth of the ear, with the diverticulum passing through an opening in the exoccipital bone and terminating in two distinct vesicles (one in the prootic, the other in the pterotic). Thus, in both ostariophysans and clupeomorphs, stimuli are transmitted from the swim bladder to the utriculus. However, the nature of the two otophysic (ear–swim bladder) connections are not regarded as homologous. *See* CYPRINIFORMES; OSTEOGLOSSIFORMES; PHOLIDOPHORIFORMES.

It has been shown that the Clupeomorpha plus Ostariophysi shared an ancestry with all the remaining teleosts. In other words, the Clupeomorpha plus Ostariophysi are the sister group of the Euteleostei.

Euteleostei. The Euteleostei are by far the largest teleost taxon with more than 22,000 species arranged in some 340 families. They comprise two major lineages, the Protacanthopterygii and the Neognathi (see illus.).

The Protacanthopterygii (often regarded as lower euteleosteans) include four groups: the salmonoids (salmon and allies; 66 species) and the osmeroids (smelts, salangrids, *Lepidogalaxius*; 72 species), which together are the sister group of the alepocephaloids (slickheads; 60+ species) plus the argentinoids (argentines or herring smelts; 60+ species). The Neognathi, on the other hand, are

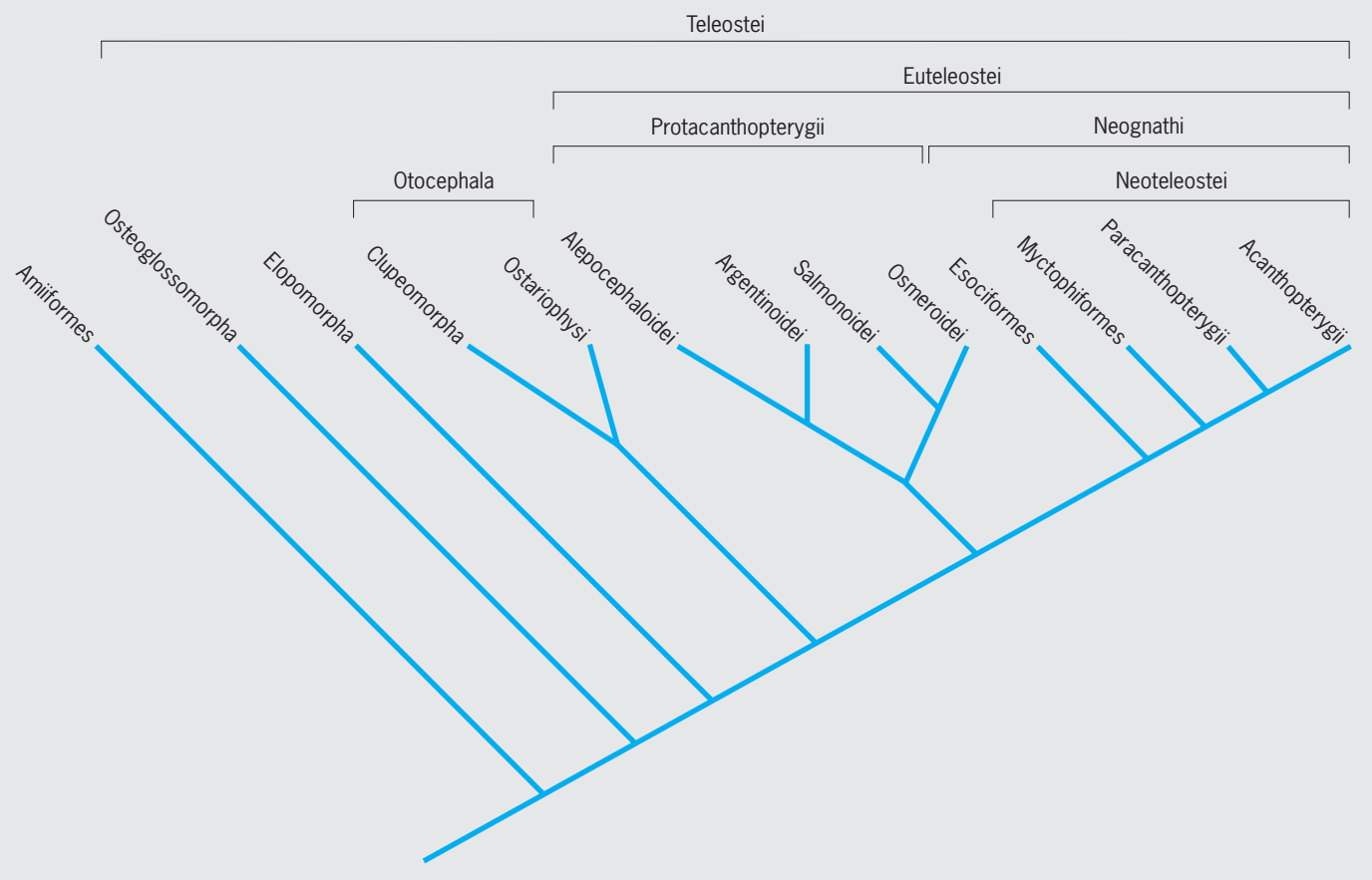

Phylogeny of the Teleostei.

made up of the esocoids (pike and mudminnows; 10+ species) plus the sister group Neoteleostei (see illus.). *See* SALMONIFORMES.

The Neoteleostei, with more than 15,319 species, comprise four main clades, the Stomiiformes, Aulopiformes, Myctophiformes, and Acanthomorpha (paracanthopterygians plus acanthopterygians); the Myctophiformes are the sister group of the Acanthomorpha. The basal clade of the Neoteleostei is the Stomiiformes, a morphologically diverse group (321 species) of deep-water oceanic fishes with unique photophores living down to 1000 m. In contrast, the Aulopiformes (219 species) are a morphologically diverse group of benthic (deep-water) and pelagic (surface-water) fishes that range in habitat from the abyss to estuaries.

The Neoteleostei include three groups of very advanced fishes, the Myctophiformes (242 species), paracanthopterygians (1212 species), and Acanthopterygians (13,414 species). The paracanthopterygians, like the Aulopiformes, contain various deep-sea groups, including the Lophiiformes (angler fishes; 297 species), and the Gadiformes with about 482 species, many of which are important commercial fishes that constitute over one-quarter of the world's marine fish catch. Other paracanthopterygian groups include the fresh-water Percopsiformes (trout and pirate perches and cave fishes; 9 species) and the Batrachoidiformes (coastal, benthic toadfishes; 69 species). *See* BATRACHOIDIFORMES; GADIFORMES; LOPHIIFORMES; PERCOPSIFORMES.

The acanthopterygians are the largest subgroup of the Euteleostei, distributed among 12 orders and 218 families (whose interrelationships are mostly unknown). The Perciformes form the largest of these orders with 9293 species (many members of which bear spines on their fin). Of the more advanced members, the Scorpaeniformes (mail-cheeked fishes), with 25 families and about 1271 species, are second to the Perciformes in species richness, whereas the Pleuronectiformes (flat fishes) with 570 species and the Tetradontiformes (trigger fish and puffer fish) with 339 species are a distant third and fourth. Members of a lower clade include the Gasterosteiformes (sticklebacks, sea horses, pipe fishes, and so on; 257 species) and the Atherinomorpha (rainbow fishes, silver sides, blue eyes; 285 species). *See* GASTEROSTEIFORMES; PERCIFORMES; PLEURONECTIFORMES.

Biology. An important characteristic of the Teleostei is the presence of a swim bladder, a gas-filled hollow organ lying between the gut and kidneys, which develops as a diverticulum from the dorsal wall of the esophagus. In the basal groups, the swim bladder usually remains in contact with the esophagus via the pneumatic duct. In more advanced groups, such as the myctophoids, this connection may be lost, and a new structure—the

gas gland—develops. Thus in the myctophoids, the principal function of the swim bladder is hydrostatic and gas can be secreted into or removed from it. Consequently, myctophoids can undergo diurnal migrations of several hundred meters. *See* SWIM BLADDER.

In some osteoglossomorphs, clupeomorphs, and euteleosts there is a direct otophysic connection between the labyrinth of the ear and the swim bladder, whereas in ostariophysans there is an indirect connection via a paired chain of bones. There can be little doubt that these Weberian ossicles allow the swim bladder to function as a manometer or pressure receptor, since the behavior of the loaches is used to predict weather changes. In the Clupeomorpha there is also a connection between the swim bladder and the lateral line, and it has been shown that these swim bladder–ear connections in the herrings are important in shoaling behavior. Finally, the swim bladder can also function as a resonator when beaten by the ribs, as in certain toadfishes. Brian Gardiner

Bibliography. J. S. Nelson, *Fishes of the World*, 3d ed., Wiley, New York, 1994; J. R. Paxton and W. N. Eschmeyer (eds.), *Encyclopaedia of Fishes*, 2d ed., Academic Press, San Diego, 1998; M. L. J. Stiassny, L. R. Parenti, and G. D. Johnson (eds.), *Interrelationships of Fishes*, Academic Press, San Diego, 1996.

Teleostomi

A monophyletic grade of craniates comprising Acanthodii, Actinopterygii (ray-finned fishes), and Sarcopterygii (coelacanths, lungfishes, and tetrapods). [Monophyletic refers to any form evolved from a single interbreeding population. Craniates are vertebrates distinguished by a cranium.] The Acanthodii (all fossil species) is the sister group to the actinopterygians and sarcopterygians, a relationship based on otoliths and certain details of the vertebral column and associated elements. The term Teleostomi, coined by C. L. Bonaparte in 1836, has changed in meaning, as have Acanthodii and Actinopterygii, which originated with E. D. Cope in 1871. The term Osteichthyes, first used by T. H. Huxley in 1880, is not used by J. S. Nelson in a formal taxonomic sense; however, the term is conveniently used for the vast number of fishes with a bony skeleton. The term Sarcopterygii was used by A. S. Romer to include the lobed-finned fishes, that is, crossopterygians and dipnoans; however, E. O. Wiley, and D. E. Rosen and collaborators include the tetrapods in Sarcopterygii to form a monophyletic group. The grade Teleostomi contains about 53,633 extant valid species, of which 26,891 are actinopterygians and 26,742 are sarcopterygians. *See* ACANTHODII; ACTINOPTERYGII; OSTEICHTHYES; SARCOPTERYGII. Herbert Boschung

Bibliography. J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; D. E. Rosen et al., Lungfishes, tetrapods, paleontology, and plesiomorphy, *Bull. Amer. Mus. Nat. Hist.*, 167(4):159–276, 1981; E. O. Wiley, Ventral gill arch muscles and the interrelationships of gnathostomes, with a new classification of the Vertebrata, *Zool. J. Linn. Soc.*, 67:149–179, 1979.

Telephone

An instrument containing a transmitter for converting the acoustic signals of a person's voice to electrical signals, a receiver for reconverting electrical signals to acoustic signals, and associated signaling devices (the dial and alerter) for communicating with other persons using similar instruments connected to a network. The term "telephone" also refers to the complicated system of transmission paths and switching points, called the Public Switched Network (PSN), connected to this instrument. This article discusses only the instrument (also called a set); for a discussion of the telephone system *see* TELEPHONE SERVICE.

Telephone set design underwent a steady evolution for a century from the first instruments of A. G. Bell and T. A. Edison. During this period a standard interface was maintained between the telephone instrument and the telephone system network. A basic telephone instrument provides plain old telephone service (POTS). The interface between the instrument and the network is two copper wire conductors brought into the subscriber's premises. These are called ring and tip. Sometimes a third conductor (ground) and a fourth (sleeve) are added to provide additional capabilities such as selective ringing on party lines and coin collection on coin telephones. The invention of the transistor and developments in digital electronic circuits have allowed substantial changes in this connection method, which are exploited in wireless and some business telephones, but the copper tip and ring arrangement is still used for basic telephones. Bell's original invention, the creation of an electrical analog to the acoustic signal of a person's voice, continues to be the basis for the speech-related circuitry of a POTS telephone.

A POTS or "standard desk" telephone set contains eight basic parts: transmitter, receiver, antisidetone network, dial, alerter, switchhook, connecting cords (wires), and chassis. Added to these may be other feature-oriented hardware (and computer programs) such as electronic memory (for dialing preprogrammed numbers), loudspeakers and microphones (for hands-free use), a voice synthesis unit, a voice recorder (for storing messages), and a video screen (for displaying text, graphs, or an image of the person at the other end of the connection). The set's complexity, depending on its intended usage, can approach that of a small computer. The telephone's original function, that of permitting voice communication with another person, may therefore be enhanced to permit communication with computers, automatic reporting of emergencies, and use of written or graphic information.

Transmitter. The transmitter is a transducer that converts acoustic energy into electric energy. E. Berliner's invention of the carbon transmitter was

the key to practical telephony because it amplified the power of the speech signal, making it possible to communicate over distances of many miles. In the carbon transmitter, sound pressure on the diaphragm varies the pressure on the carbon. Changes in pressure on the carbon granules vary the resistance of the transmitter, causing the magnitude of the direct current supplied by the central office to change in proportion to the sound, a process called modulation. *See* MODULATION.

Current designs are based on the charged electret (a condenser microphone) or on electrodynamic principles. Both the electret and electrodynamic transmitters use transistors to provide needed power gain; they introduce less distortion than the carbon transmitter. *See* ELECTRET TRANSDUCER; MICROPHONE; TRANSDUCER; TRANSISTOR.

Transmitters in a basic set have a frequency-response range intentionally limited to 250–5000 Hz to improve voice clarity on the Public Switched Network. Sets that use another network for transmission may have a wider frequency response. Other components in the Public Switched Network may have even less bandwidth than the transmitter provides. The frequency response of the transmitter rises uniformly to a broad maximum in the region of 2500 Hz. Even though normal human hearing has a much broader frequency response, speech heard on the telephone resembles closely that heard by a listener a few feet from the person speaking. Confusion is possible between "f" and "s" sounds, but the listener usually subconsciously resolves this by the context of the speech. *See* HEARING (HUMAN).

The heart of an electret transmitter (see **illustration**) is an electrical capacitor formed by the metal on the diaphragm, a conductive coating on top of the metalized lead frame, and the plastic and air between the metal layers. The diaphragm is made of a special plastic that can be given a permanent electrostatic charge (analogous to the magnetization of a permanent magnet). As sound waves entering the sound port cause pressure changes, the diaphragm moves closer to and farther away from the metalized lead frame. This changes the value of capacitance and produces a varying electric voltage which is the analog of the impinging sound-pressure wave. The signal is amplified by the internal amplifier chip to a level that is suitable for transmission on the telephone network.

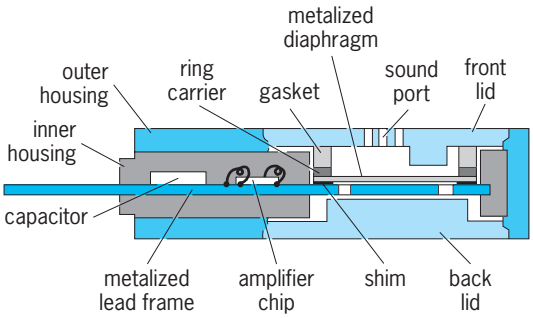

Cross section of an electret transmitter.

Receiver. The receiver transducer operates on the relatively low power used in the telephone circuit; it converts electric energy back into acoustic energy. As in the transmitter, careful design of the relationship of the acoustical and electrical elements produces a desired response-frequency characteristic. Usually a rising response over the frequency range 350–3500 Hz is desirable to match the transmission characteristics of the traditional telephone network, optimize intelligibility, and minimize noise. However, sets intended for use with networks other than the traditional network may have a wider and flatter frequency response.

Historically, there were two common types of receiver units with fixed-coil windings, the ring armature receiver and the bipolar receiver. Fixed-coil receivers were designed for close coupling to the ear and were placed in a cuplike enclosure shaped to fit closely to the ear. Moving-coil designs, similar to loudspeakers, have replaced these in most telephones. This change can provide hands-free (speakerphone) capability in an ordinary handset but, because an ear-shaped cup is often not used, results in poorer performance in noisy environments. *See* LOUDSPEAKER.

Anti-sidetone network. The anti-sidetone network has two basic functions: to reduce sidetone (explained below), and sometimes to provide equalization (reduce the variation in loudness that is caused by varying lengths of cable between the central office and different subscribers). However, the incorporation of volume controls in many sets makes equalization unnecessary. This network may also provide more efficient coupling of the transducers to the telephone line if it incorporates a transformer.

Sidetone is the sound of the speaker's voice reaching his or her own ear. Normally this occurs through the air. However, when one is using a telephone handset, the acoustic sidetone path to one ear is replaced by an electric sidetone path through the telephone's transmitter, network, and receiver. Since the transmitter amplifies power impinging on its diaphragm, this new path may result in a louder sound in the ear coupled to the receiver than would exist if the receiver were not being used. This sound is unpleasant and can also cause the telephone user to talk more softly, thereby making it harder for the person at the far end to understand the conversation. Anti-sidetone networks may consist entirely of passive electric components such as transformers, capacitors, and resistors; or may include active components such as transistors embedded in integrated circuits. *See* ELECTRIC FILTER; INTEGRATED CIRCUITS.

Dial. The dial was originally designed to permit selection of a particular far-end station by operation of switches in the central office. The all-mechanical rotary dial used for many years contains a pair of electric contacts that periodically interrupt the power flowing from the central office. The resulting direct-current pulses are generated at the rate of 10 per second. Dual-tone multifrequency (DTMF) codes are used for the same purpose in Touch Tone telephones.

In this system, two audio tones selected from a 3 × 4 matrix are used to represent each digit. The tones are selected by a corresponding 3 × 4 matrix of pushbuttons. Each time a button is pushed, the two tones are sent simultaneously. The tone pairs are selected so that they are unlikely to occur in natural speech, thereby avoiding the inadvertent simulation of digits by speech during the normal dialing interval. These tones have an advantage over the low-frequency direct-current pulses in that they can be transmitted over the transmission network to the far-end station (as is the electrical analog of the voice signal), whereas low-frequency pulses normally will not carry beyond the local central office. Used this way, the multifrequency tone signals can convey information to a so-called intelligent telephone such as an answering machine or to a computer at the far end of the connection. *See* SWITCHING SYSTEMS (COMMUNICATIONS).

Alerter. The alerter, also called the ringer, which normally is powered by a 20-Hz 88-V alternating current, serves to alert the subscriber to an incoming call. In some older party-line systems, selective ringing of different sets on the common loop (one pair of wires connecting several different subscribers to the central office) was accomplished by using several different frequencies, one for each subscriber or by using the ground conductor mentioned previously. Mechanical ringers consist of a clapper that periodically strikes a gong, the bell. Both the electrical circuit that powers the clapper and the mechanical elements of the clapper are made resonant at the right frequencies to maximize sound output and minimize sensitivity to signals of other frequencies. A nonlinear magnetic circuit serves to present a higher impedance to speech signals on the telephone line than to the ringing signal. Electronic (or tone) ringers make use of electrical oscillator design techniques to produce their sound. They also contain nonlinear circuitry to emulate the characteristics of the mechanical ringer's nonlinear magnetic circuit. The frequency selectivity of the ringer helps prevent ringing when dial pulse or noise signals are present on the telephone line. *See* OSCILLATOR.

Switchhook. The switchhook contains a set of electrical contacts that interrupt the flow of current from the central office whenever the set is "on hook." The closing of these contacts when the handset is lifted signals the central office that the set's user wants to initiate or answer a call.

Connecting cords. Telephone sets may include two cords: a line cord that connects the instrument to tip and ring, and a handset cord that connects the telephone handset to the chassis.

Line cords. Basic telephone sets require two-conductor line cords, while more complicated sets, with capability of two or more lines and other special services, use four, six, eight, or more conductors in their cords. Specialized service requirements may require twisted-pair construction to reduce crosstalk problems between multiple-voice circuits, or between data and voice circuits. Line cords are typically constructed with individually insulated stranded copper conductors covered with an outer jacket. Material selection for both the conductor insulation and the outer jacket is critical in determining mechanical strength, chemical resistance, protection against voltage hazards, appearance, and cost. *See* CROSSTALK; DIELECTRIC MATERIALS; ELECTRICAL INSULATION.

Handset cords. Handset cords typically use four conductors. Two conductors connect to the telephone receiver, and the other two to the telephone transmitter. Telephone handset cords need to be extremely flexible and to withstand a great deal of physical abuse. The handset cord conductors are of a special construction, called tinsel, that provides superior mechanical strength, flexibility, and fatigue properties. A tinsel conductor consists of a nylon or polyester thread, or center core, with three or four thin copper-alloy ribbons wrapped helically around it. These thin, springlike conductive ribbons are extremely flexible and are able to withstand the pulling, bending, and twisting loads to which handset cords are subjected. During manufacture, handset cords are coiled, or made retractile, by wrapping the cord around a heated rod called a mandrel. After cooling, the cord retains its coiled shape.

Modular connectors. Usually, line cords are connected to the telephone network with standardized modular plugs and jacks (connectors). This replaces the former practice of hard-wiring instruments to their network connection. The mating surfaces in the plug and jack are usually gold-plated to ensure good electrical contact. A flexible locking tab on the modular jack secures the modular plug and is depressed to disconnect the cord. This system permits the telephone to be moved easily from one location to another. Modular connectors are also used on the handset cord, making cord replacement easy, an important consideration since cords are subjected to heavy wear and abuse.

Chassis. The chassis provides a mechanical support to hold all the other parts together. It must be capable of withstanding considerable abuse as telephones are often inadvertently dropped. The chassis must also present the components in a way that is convenient for the user of the telephone. The dial in desk sets, for example, should be mounted at an angle easy to use.

Special features. Many attempts were made in early telephones to add memory to the telephone dial. This feature was never popular, however, until digital circuits were used instead of mechanical devices to store telephone numbers. A display, included in some sets, allows the user to see what telephone number has been stored. This same display can also give the time and date when not showing telephone numbers or can display the number of the calling party. Another popular feature is "hands-free" telephony, which substitutes a microphone and loudspeaker for a handset-mounted transmitter and receiver.

Communication with machines. Applications of telephony in which one of the two parties communicating is a machine have become common. One

example is telephone banking, a service using the DTMF dial to send numerical information such as account numbers, passwords, transaction codes, and dollar amounts from a human subscriber to the service provider's computer on the other end of the connection. Another example is the automated receptionist function where the caller is asked to dial a digit corresponding to the service desired (make a doctor appointment, speak to a nurse, fill a prescription, and so forth).

Features for hearing-impaired users. Special features are provided for hearing-impaired users. These include small amplifiers for either the transmitter or the receiver that are built into the handset, devices analogous to ringers that provide visual signals, and receivers having a leaky magnetic field that can be picked up by specially equipped hearing aids. For the totally deaf, a machine with a keyboard and alphanumeric display may be used in place of an acoustic instrument.

Cordless telephones. Cordless telephones use a two-way radio link between one or more handsets and a base unit; this replaces the handset cord. The base unit is connected through a normal line cord to the telephone network. Typically, the dial, alerter, and switchhook functions are built into the handset, providing full calling and answering capability on a portable basis within a radius of several hundred feet of the base unit. The base unit normally requires a connection to alternating-current power, as well as to the telephone network, while the handset uses batteries that are recharged when the handset is replaced on the base unit. Allowable radio power between the handset and the base unit is severely limited by law, resulting in a telephone suitable for use around a residence but not suitable for longer distances. *See* RADIO SPECTRUM ALLOCATIONS.

Wireless telephones. The cellular radio system, a form of mobile radio telephony, serves the need for a telephone with which the user can wander over a large geographical area while making a call. Cellular telephones use a two-way radio link to replace the line cord; thus tip and ring are not brought to the telephone. The telephone communicates with one of a number of base stations spread in cells throughout the service area. As the telephone user changes location, the link is automatically switched from cell to cell to maintain a good connection. The transmitted radio power for a cellular telephone is larger than that for a cordless telephone, resulting in the larger service range. Cellular telephones operate on different radio frequencies from those used by cordless telephones. Both analog and digital radio links are used; analog links are being replaced by digital to provide more channels and longer operational time on a battery charge. Several different protocols are used by providers between the telephone and base station, so the telephone used must be chosen to be compatible with the service provider's network. Also, enhanced features, such as text messaging and transmission of photographs or video, are available in some protocols and require telephones having the necessary hardware and software to support them. *See* MOBILE COMMUNICATIONS. Richard M. Rickert

Business terminals. Business telephones designed for use on a customer's premises may connect to an on-premises switching system known as a private branch exchange (PBX), or a central office switch. They generally offer the ability to support multiple-line appearances without requiring more than a single port on the switch. Multiple-line appearances provide the ability to receive or originate additional voice or data calls at the terminal while still being engaged in a primary voice call, and the ability to bring additional parties onto a primary call (forming a conference call) or to transfer the primary call to a second party. *See* MICROCOMPUTER; PRIVATE BRANCH EXCHANGE; TELECONFERENCING. Peter S. Warwick; Wesley L. Shanks

Digital transmission and VoIP. Digital transmission of speech is used in long-distance telephony because digital signals can be regenerated precisely and thus are much less subject to degradation as a function of increasing transmission distance than are analog signals. The cost of converting the analog signal back and forth from the digital format usually makes this mode of operation uneconomic for local loops (those within a single central office). However, when all or a major part of the information flow from a set to a central point is already in digital format, as in the wireless telephones and business terminals discussed above, it is economical to digitize the speech at the set. In this arrangement, all the information flowing from the instrument (including the voice signals) is converted to the form of binary digits before leaving the set. In this case the local loop and serving central office or private branch exchange must also be digital. These offices are part of the existing Public Switch Network. Since an all-digital set cannot communicate directly with standard analog instruments, equipment is needed at some central point in the transmission system to convert the digital bit stream to an analog signal, and vice versa. *See* PULSE MODULATION.

The wide availability of high-speed connections to the Internet has made possible the introduction of a new type of telephone design, network connections, and usage patterns referred to as voice over the Internet Protocol (VoIP). VoIP telephones may retain some analog circuits near the transmitter and receiver interfaces, but the interface from the set to the outside world is digital and does not involve a tip or ring. VoIP calls are not carried on the Public Switched Network. Special interfaces are needed to complete a call from a telephone using VoIP technology and one connected to the Public Switched Network. These interfaces are provided by the company selling VoIP service. Sometimes VoIP service is provided by using a standard telephone as described in this article and a converter between the telephone and a high-speed Internet connection. *See* VOICE OVER IP. Richard M. Rickert

Bibliography. J. C. Bellamy, *Digital Telephony*, 3d ed., 2000; J. Carr, S. Winder, and S. Bigelow, *Understanding Telephone Electronics*, 4th ed., 2001;

T. Dean, *Guide to Telecommunications Technology*, 2002; D. G. Fink and D. Christiansen, *Electronics Engineers' Handbook*, 4th ed., 1997; Hill Associates Inc., *Telecommunications: A Beginner's Guide*, 2001; H. Newton, *Newton's Telecom Dictionary*, 22d ed., 2006; A. M. Noll, *Introduction to Telephones and Telephone Systems*, 3d ed., 1999.

Telephone service

The technology of providing many types of communications services via networks that transmit voice, data, image facsimile, and video by using both analog and digital encoding formats.

Component sectors and core technologies. Telephone services involve three distinct sectors of components: (1) customer premises equipment (CPE), such as telephones, fax machines, personal and mainframe computers, and systems private branch exchanges (PBXs); (2) transmission systems, such as copper wires, coaxial cables, fiber-optic cables, satellites, point-to-point microwave routes, and wireless radio links, plus their associated components; and (3) switching systems that often can access associated databases, which can add new intelligent controls for the network's users. *See* COAXIAL CABLE; COMMUNICATIONS CABLE; COMMUNICATIONS SATELLITE; COMPUTER; DIGITAL COMPUTER; MICROWAVE; OPTICAL COMMUNICATIONS; OPTICAL FIBERS; RADIO; SWITCHING SYSTEMS (COMMUNICATIONS); TELEPHONE.

Advances in these components are driven by three core technologies: microelectronics, especially microprocessors; photonics, particularly fiber optics and lasers; and software, which is especially important as a means of introducing, operating, and maintaining telephone services. *See* LASER; MICROPROCESSOR; SOFTWARE.

Controlled by systems integration techniques, the interaction of these three component sectors and these three core technologies has expanded both the capacities and capabilities of the domestic and international communications systems. At the same time, the synergy of the core technologies has reduced significantly the costs of installation and operation of telecommunications systems. This means more sophisticated telephone services can be made more affordable to a broader range of customers. *See* SYSTEMS INTEGRATION.

Private networks. Most telephone service customers share the network facilities with other customers (also known as subscribers). However, some customers, especially large businesses and government agencies, lease private lines that are available only to them. Numerous private networks have been formed within the facilities of the public switched networks. A growing number of these systems employ transmission and switching concepts that create temporary virtual networks as needed, using whatever network facilities are available at the moment. Some of these network applications are discussed below.

A commonly used private network is based on a small on-site switchboard called the private branch exchange (PBX), which serves facilities ranging from a small law office to a multibuilding manufacturing corporation. Although earlier private branch exchange switchboards, such as in hospitals and hotels, required one or more operators, modern versions of private branch exchanges are completely automatic. Calls can be made among telephone stations in the network without the circuit going outside the facility. Some phone companies offer a substitute service, called centrex, in which a section of the central office switch serves the customer like a private branch exchange. *See* PRIVATE BRANCH EXCHANGE.

One widely used private branch exchange feature is voice mail. In essence, it is a centralized answering machine in which a prerecorded message asks the caller to leave a spoken message. The person assigned to that number can listen to the voice-mail messages either upon return to the office or hotel room or by calling from a remote location. Each private branch exchange phone number has an assigned "mailbox" in which message recordings are stored.

Voice and video. Communication, by definition, involves the exchange of information. In telephone systems, this is done by establishing and maintaining a connection between two or more communications terminal devices such as telephones, fax machines, computers or computer modems, video transmitters and receivers, teletypewriters or telex machines, scanners, and voice-mail systems.

A connection between subscribers may be local, regional, transcontinental, or international, and may last for only a few seconds or for years, depending on the needs of the subscribers. A simultaneous two-way connection is known as a duplex channel or circuit, the capacity of which is measured in the range of frequencies, or hertz (Hz), needed to transmit the information. This capacity is called the circuit's bandwidth.

For a simple analog voice-grade phone line, the bandwidth is specified as 4000 Hz, of which about 300–400 Hz are used for supervisory signaling (dial tone, line-busy tone, network-busy tone, and so forth). The distinctive narrow-range sound spectrum of a telephone voice is the result of transmitting only about one-fourth of the actual frequency range used in speaking. Providing the full range of sound would require far more transmission capacity or retrofitting existing networks with more complex components, increasing the cost of service.

In digital systems, the analog speech sounds are encoded into digital signals by a sampling process, using a method known as pulse-code modulation (PCM). This is based on two concepts: first, a continuously varying signal wave, formed by a rapidly changing series of different frequencies (that is, the sounds of speech), can be sliced into a series of samples that can be reconstructed to reproduce the signal wave; and second, the samples can be adequately approximated by discrete numbers. *See* DIGITAL-TO-ANALOG CONVERTER; PULSE MODULATION.

To achieve a natural speech sound (within the voice-channel spectrum), the analog 4000-Hz signal must be sampled digitally at a relatively high rate (8000 times per second, using eight bits to describe the amplitude of the waveform at each sample), resulting in transmission at a nominal rate 64,000 bits per second (64 kb/s). However, advances in coding techniques have lowered the number of bits needed to maintain the quality of telephone voice transmission. Of course, the digitally encoded speech signals must be decoded into analog sounds before a human can hear them. *See* DATA COMPRESSION.

Use of a telephone channel for ordinary analog voice conversations, often referred to as POTS (plain old telephone service), continues to be the dominant mode of service in both domestic and international networks. Although most telephone communication is based on speech (or, increasingly, data exchanges between desktop or portable computer terminals), limited-motion video transmission is now possible on ordinary analog phone lines. There are even videophones designed for encrypted transmission between compatible units to prevent unauthorized interceptions. In addition, digital videophones, employing integrated services digital network (ISDN) technology, serve a variety of businesses. ISDN systems have been particularly successful in Europe, where they have been used to "leapfrog" aging analog systems. Full-motion video currently requires more bandwidth than is available on an ordinary analog phone line. *See* INTEGRATED SERVICES DIGITAL NETWORK (ISDN); VIDEOTELEPHONY.

Significant numbers of telephone customers have special physical needs. They are physically challenged by deafness, blindness, or speech impairment. Special devices are designed to overcome such problems. These include the TTY (teletypewriter) or TDD (telecommunications device for the deaf) systems that display text on small screens, as well as volume controls on receivers, large-button or Braille-button keypads on telephone sets, and electronic larynxes.

Infrastructure. Made over a web of circuits known as a network, a connection often involves several different telephone companies or carriers. In the United States, there are more than 1300 local exchange companies (LECs) providing switched local service, plus more than 500 large and small interexchange companies (IECs), that provide switched long-distance services.

In addition, wireless mobile telephone services are provided on a city-by-city basis by cellular systems operated by two different companies in each metropolitan area. Each cellular system is connected to the wire networks of the local exchange companies and interexchange companies. *See* MOBILE COMMUNICATIONS.

Local exchange company operations of wired systems are divided geographically into 164 local access and transport areas (LATAs), each containing a number of cities and towns. Within a LATA, the telephone company operates many local switches, installed in facilities known as central offices or exchanges.

Transmission. Each central office or exchange switch in connected to local telephone customer premises by a system of twisted-pair wires, coaxial cables, and fiber-optic lines called the loop plant. Direct current (dc) electricity that carries signals through the wires (or powers the lasers and photodetectors in the glass-fiber lines) is provided by a large 48-V stationary battery in the central office which is constantly recharged to maintain its power output. If a utility power outage occurs, the battery keeps the transmission system and the office operating for a number of hours, depending on the battery's size.

The transmission is usually analog to and from the customer site, especially for residences. However, in many systems a collection point between the customer and the switch, known as a subscriber loop carrier (SLC), provides a conversion interface to and from one or more digital cable facilities called T1 carriers (each with 24 channels) leading to the switch. In turn, most switches are interconnected by digital fiber-optic, coaxial, or copper-wire cables or by microwave transmission systems either directly or via intermediate facilities called tandem switches between local switches (**Fig. 1**) or between a group of central offices and a long-distance toll switch. *See* TELEPHONE SYSTEMS CONSTRUCTION.

Domestic dialing. This matrix of local, regional, and national networks is said to be transparent to telephone users, since the connections between different networks usually are made automatically by the caller simply dialing the number. For domestic calls within the United States and Canada, the telephone number for a local call consists of seven digits. The first three digits identify the exchange or central-office switch serving the customer's

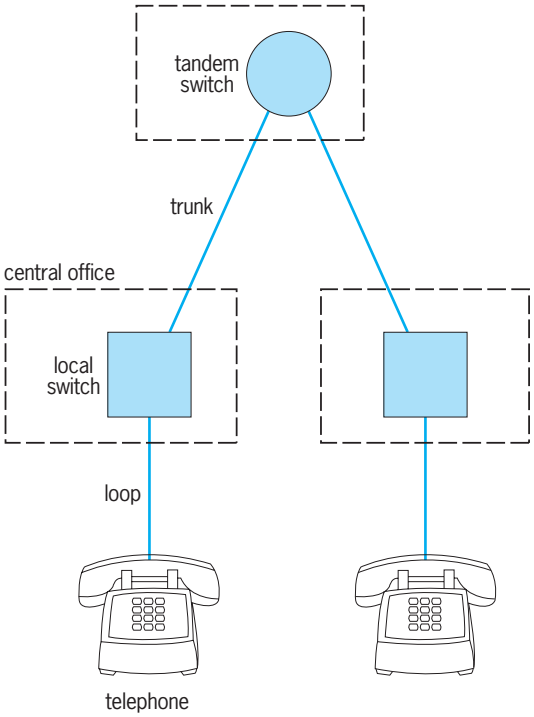

Fig. 1. Elements of the local telephone network.

premises, and the next four digits identify the line number.

If a call is placed beyond the exchange or switch that serves a given phone line, the dialed number often must be proceded by 1 for intra-LATA toll calls (involving switches outside the local area but still inside the LATA), or a 1 plus a three-digit area code and the seven-digit number for a long distance call between LATAs. By the mid-1990s, the reservoir of 152 area codes was depleted; the restriction on the second digit (formerly limited to either 0 or 1) was removed, expanding the feasible selection to 792 area codes.

Although most long-distance calls are dialed directly, using an automatic switching technology called direct distance dialing (DDD), some may require operator assistance (for example, for collect calls charged to the number being called, or for calls to be charged to a telephone credit card). In these cases, the prefix 0 is dialed instead of 1, followed by the area code and number being called. An operator answers the call and processes it from that point on, or a recording instructs the caller to enter the card number and a personal identification number.

International dialing. For international direct-distance dialing (IDDD), the caller usually must dial about 15 digits, starting with 011 to reach an overseas circuit, followed by a country code as well as an area code and telephone number. The dialed numbers are processed automatically through an international gateway switching machine.

In 1963, an international commission developed a numbering plan for all member countries of the United Nations. Each nation was assigned a country code (of one, two, or three digits), followed by the national number of the subscriber's access line.

Switching and transmission designs. Telephone switching machines used analog technology from 1889 to 1974, when the first all-digital switch was introduced, starting with long-distance or toll service and expanded later to central offices or local exchanges. Digital switching machines are comparable to digital computers in their components and functions. Unlike the analog electromechanical switches of the past, digital switches have no moving parts and can operate at much faster speeds. In addition, digital switches can be modified easily by use of operations systems—special software programs loaded into the switch's computer memory to provide new services or perform operational tasks such as billing, collecting and formatting traffic data from switches, monitoring the status of transmission and switching facilities, testing trunk lines between end offices, and identifying loop troubles.

Digital transmission. Transmission systems also evolved from analog technology to digital technology. The first digital transmission system preceded the first digital switch by more than a decade, but use of such systems was limited to trunks between switches or other relatively short routes. The introduction of high-capacity fiber-optic transmission systems in the 1980s made the present-day digital network feasible. The first long-distance fiber (or light-wave) system linked Boston, New York City, Philadelphia, and Washington, D.C., in 1981. The first transatlantic fiber-optic submarine cable, which also introduced undersea cable branching for multiple landing sites, was activated in 1988, followed in 1989 by the first transpacific fiber-optic cable system. *See* SUBMARINE CABLE.

Although fiber-optic cables are much thinner than copper cables, they have far more capacity in terms of channels. The last copper transatlantic cable (installed in 1983 and known as TAT-7) was rated for 4000 simultaneous analog conversations, while the first transatlantic fiber cable (TAT-8) was rated for 40,000 simultaneous digital conversations. Subsequent fiber cables were rated for 80,000 digital conversations (TAT-9 and TAT-10).

The submarine light-wave cables provided the first true high-speed digital telephone circuits across the Atlantic and Pacific oceans, but they also encountered a disparity between the digital protocols of North American (where pulse-code modulation encoding is done with a mu-law algorithm) and Europe, Africa, South America, and most Asian countries (where such coding uses an incompatible A-law algorithm). By international agreement, the mu-law countries such as the United States are responsible for providing gateway transcoding facilities. The digital signals transmitted via the submarine cables are coded in A-law pulse-code-modulation signals.

Digital hierarchies. In North America, digital switching and transmission are conducted within a hierarchy of multiplexing levels. The single digital telephone line, rated at 64 kilobits per second (kb/s), is known as a digital signal 0 (DS-0) level. The lowest digital network transmission level is DS-1, equivalent to 24 voice channels multiplexed by time-division multiplexing (TDM) to operate at 1.544 megabits per second (Mb/s). DS-1 is the most common digital service to customer premises. Two interim levels, now seldom used, are followed by DS-3, perhaps the most widely used high-speed level. DS-3 operates at 44.736 Mb/s, often rounded out to 45 Mb/s, the highest digital signal rate conventionally provided to customer premises by the telephone network. Of course, groups of DS-3 trunks can be connected to a customer facility if needed. Outside the United States, other digital multiplexing schemes are used. This results, for example, in a DS-1 level having 30 voice channels rather than 24 channels. *See* MULTIPLEXING AND MULTIPLE ACCESS.

A second network hierarchy appeared during the early 1990s and was gradually implemented in the world's industrial nations. Known in North America as SONET (synchronized optical network) and in the rest of the world as SDH (synchronized digital hierarchy), these standards move voice, data, and video information over a fiber network at any of eight digital transmission rates. These range from OC-1 at 51.84 Mb/s to OC-48 at 2.488 gigabits per second (Gb/s), and the hierarchy can be extended to more than 13 Gb/s. The OC designation stands for optical carrier.

This concept enhances the power of the public

network by increasing the amount of information that can be carried, using international standards for digital transmission. SONET/SDH also brings network intelligence closer to end users by enlarging the capacity of the overhead channel embedded in the circuit's signal. This channel carries operations, administration, maintenance, and provisioning (OAM&P) information accessible by the user, in addition to the customer's data (called the payload portion of the signal). The built-in capabilities of microprocessors or firmware in network elements constitute the means to make the network self-aware, self-adapting, and self-managing. Problems are detected and corrected automatically at their root, a process sometimes known as a self-healing network.

In the standard digital hierarchy, customer networking applications are supported up to the DS-3 (45-Mb/s) transmission rate. Advanced technologies and services, such as the asynchronous transfer mode (ATM) and the broadband integrated services digital network (BISDN), define rates as high as 155 Mb/s and 600 Mb/s, equivalent to SONET's OC-3 and OC-12 levels. Such high-speed rates help to improve the performance of certain applications such as multimedia and distributed data processing that can strain the resources of conventional digital systems operating at lower bit rates. *See* MULTIMEDIA TECHNOLOGY.

Asynchronous transfer mode. This technology is based on a technique called cell-oriented switching and multiplexing. This is a packet-switching concept, but the packets are much shorter (always 53 bytes) and faster (for better response times) than are the packets in such applications as the global Internet and associated service networks, based on the X.25 standard, which can have up to 4096 bytes of data in one packet. The asynchronous transfer mode cell-relay system moves cells through the network at speeds measured in megabits per second instead of kilobits. With its tremendous speed and capacity, asynchronous transfer mode technology permits simultaneous switching of data, video, voice, and image signals over cell-relay networks. *See* DATA COMMUNICATIONS; PACKET SWITCHING.

Standards. Uniform technical standards have been adopted worldwide. For example, facsimile or fax machines using a global standard known as Group 3 can communicate with each other. *See* FACSIMILE.

Within the United States, technical standards for telephone services are promulgated by the American National Standards Institute (ANSI). The Federal Communications Commission (FCC) regulates all interstate and foreign radio and wire communications services originating in the United States.

Globally, the United Nations oversees the International Telecommunications Union (ITU). In 1992, the ITU was reorganized into three groups that develop and recommend standards: the Radio Communications Sector (formerly the Consultative Committee for International Radio or CCIR), the Telecommunications Standardization Sector (formerly the Consultative Committee for International Telephony and Telegraphy, or CCITT), and the Telecommunications Development Sector (BDT). Their standards are only

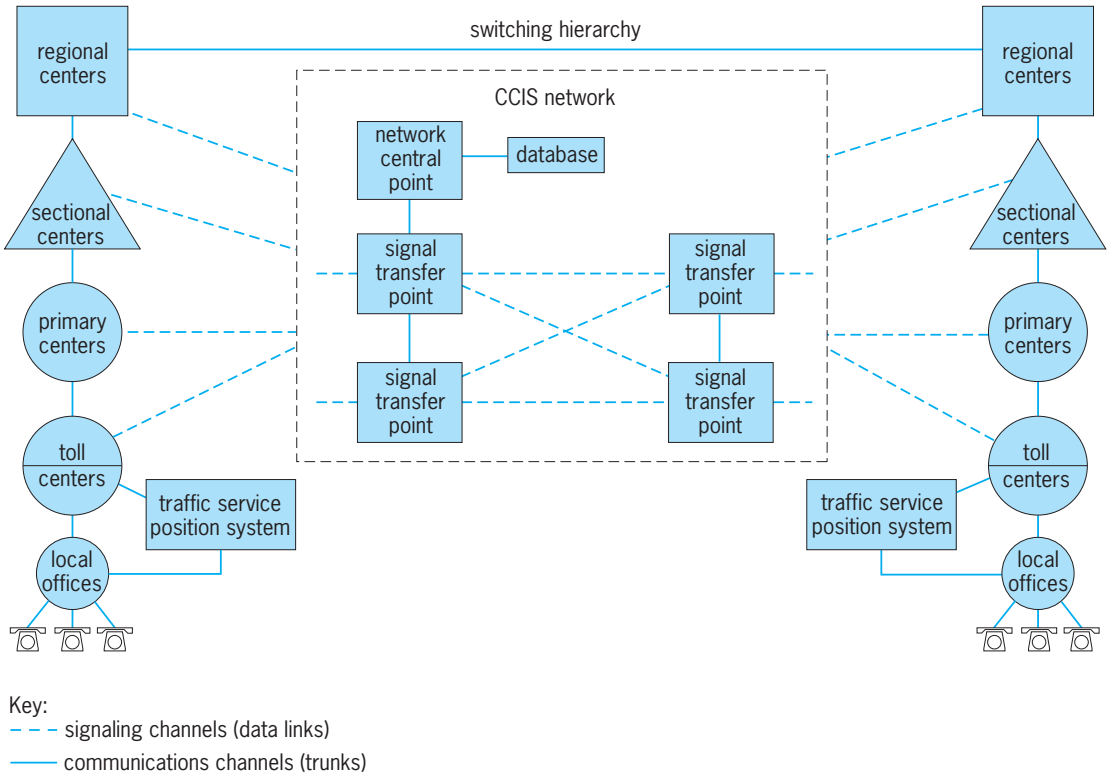

Fig. 2. National long-distance telephone network of the early 1980s, which introduced common-channel interoffice signaling (CCIS) to speed call completions, using 14 special packet switches called signal transfer points. Regional, sectional, primary, and toll centers are part of toll switching system. The traffic service position system is an operator services system. (*AT&T*)

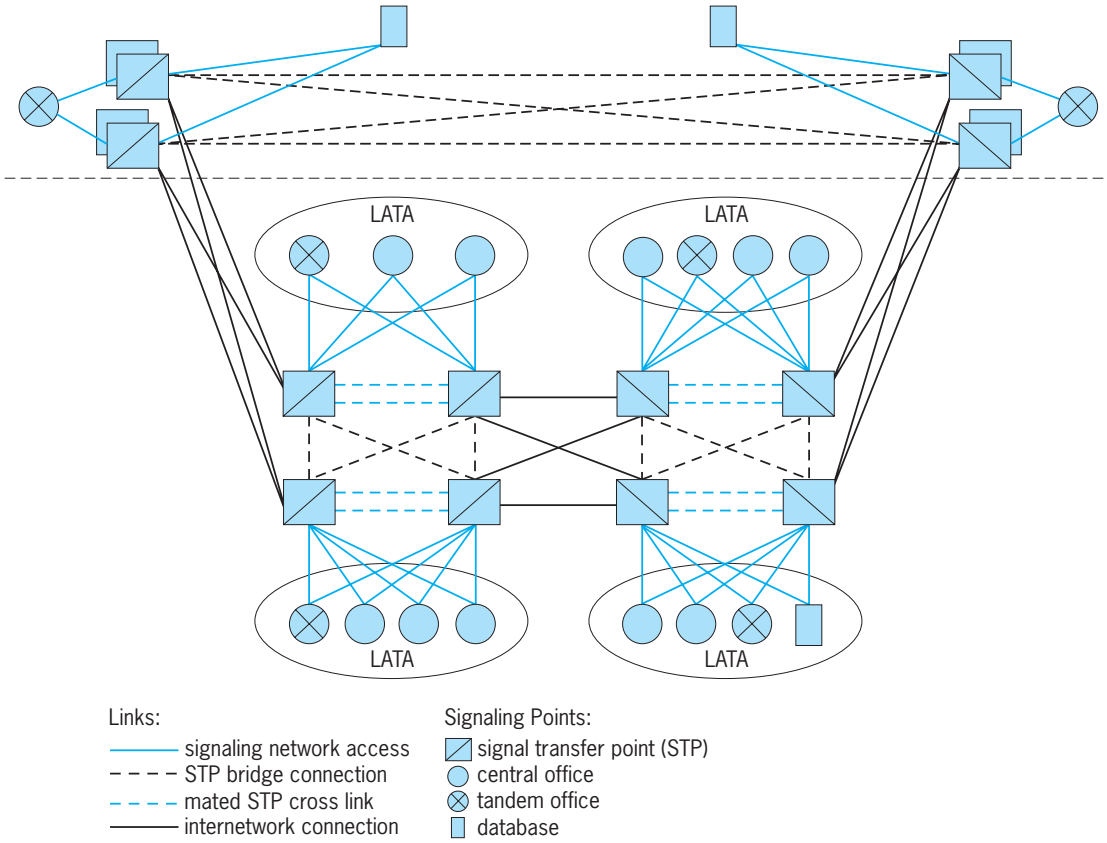

Fig. 3. National long-distance telephone network of the early 1990s, which uses common-channel Signaling System 7 technology. Signal transfer points serve local switches as well as the long-distance network with out-of-band access to various data banks, and intelligent networks can be set up. (*AT&T*)

recommendations, but most nations accept them verbatim, while others, including the United States, make minor modifications.

Out-of-band signaling systems. When the voice or communications information is encoded in a stream of bits, it becomes possible to use designated bits instead of analog tones for the supervisory signaling system. This simplifies the local switching equipment and allows the introduction of sophisticated data into the signaling, which significantly expands the potential range of telephone services.

Operations information now flows between switches as packet-switched signaling data via a connection that is independent of the channel being used for voice or data communications (a technique known as out-of-band signaling). Typically, a channel between two switching systems is known as a trunk. A trunk group between switches carries a number of channels whose combined signaling data are transmitted over a separate common channel as a signaling link operated at 56 kb/s. This technique is called common-channel interoffice signaling (CCIS).

SS6. In 1976, the first United States version of CCIS was introduced as a modification of an international standard, the CCITT No. 6 signaling system. The out-of-band signaling system was deployed in the long-distance (toll) network only, linking digital as well as analog switches via a network of packet-data switches called signal transfer points (STPs; **Fig. 2**). Also known as Signaling System 6 (SS6), it helped to make possible numerous customer-controlled services, such as conferencing, call storage, and call forwarding.

One key advantage of common-channel signaling technology was more efficient use of the national network; it typically reduced the routing time for a direct-dialed coast-to-coast call from around 10 s to only a second or two. The technology also enables the virtual elimination of a major means of fraudulent use of the network, since the common-channel signaling system is inaccessible to customers.

SS7. A major advance in common-channel signaling was introduced with the CCITT's Signaling System 7 (SS7), which operates at up to 64 kb/s and carries more than 10 times as much information as SS6. In addition, Signaling System 7 is used in local exchange service areas as well as in domestic and international toll networks (**Fig. 3**). The first trials took place in 1989; deployment in the United States long-distance networks soon followed, with local exchange companies deploying it at a somewhat slower rate.

In addition to increasing network call setup efficiency, Signaling System 7 enables and enhances services such as ISDN applications; automatic call distributor (ACD); and local-area signaling services (LASS), known as caller identification (**Fig. 4**). Other innovations include 800 or free-phone service, in

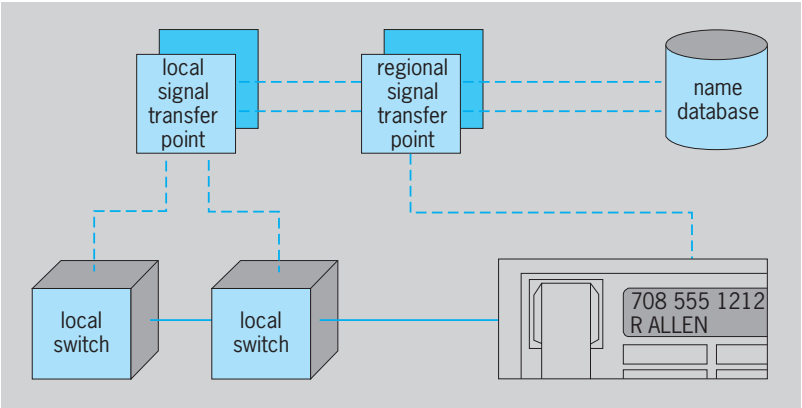

Fig. 4. Caller identification service in Signaling System 7, whereby call waiting identifies the caller's name and directory number for waiting calls. Broken lines indicate SS7 links. (*AT&T*)

which customers can place free long-distance calls to businesses and government offices; 900 services, in which customers pay for services from businesses; and enhanced 911 calls, in which customer names and addresses are automatically displayed in police, fire, or ambulance centers when calls are placed for emergency services.

The Signaling System 7 software also can include numerous other custom calling services for residential and business customers. Call waiting sends a special audio signal to a telephone that is being used, alerting the called party that another caller is waiting; the system allows the called party to place the first call on hold and answer the second call without hanging up on the first call. Selective call waiting gives the user a distinctive tone when a call is received from a number on a specified list of callers.

Other custom calling features include call forwarding, which transfers incoming calls to another number specified by the customer for a temporary period; three-way calling, which lets the customer set up conference calls without operator assistance; and speed calling, which enables the customer to program 8 or 30 frequently called numbers into the memory of the local exchange switch, so the customer need only press two or three buttons to reach these numbers.

Intelligent networks. The greatest potential for Signaling System 7 use is in emerging intelligent networks, both domestic and international. The intent is to provide customers with much greater control over a variety of network functions, yet to protect the network against misuse or disruption. Intelligent networks are evolving from the expanding use of digital switching and transmission, starting with the toll networks. The complexity of intelligent networks mandates the use of networked computers, programmed with advanced software.

An intelligent network allows the customer to setup and use a virtual private network as needed, and be charged only for the time that network is being used. A conventional private leased network, by contrast, reserves dedicated circuits on a full-time basis and charges for them whether they are used or not. A virtual private network enables a business to simultaneously reap the benefits of a dedicated network and the shared public-switched network by drawing on the infrastructure of intelligent networking (**Fig. 5**).

The introduction in 1964 of customer dialing for international calls is at the root of the intelligent network concept. The international direct distance dialing service has been expanded and made more flexible by implementing the SS7 together with new databases (such as 800 or free-phone services, emergency hot lines, capture of calling-party names and numbers, screening of incoming calls, automatic call-up and on-screen display of client order-billing records, computer-based voice messaging store-and-forward systems, and automated calling-card services. The databases are fundamentally software programs that interact with voice or data signals, either for business applications or for network operations and maintenance. *See* DATABASE MANAGEMENT SYSTEM.

For example, a retail chain can link its stores in the United States to factories operated by its vendors in Asia and Central and South America. Constant communication between computers at the sales points and the factory sources can provide savings in inventory space, better control of labor costs, and a quicker response to market conditions. Stores can be added to or dropped from the system as required, and factories can be canceled or appended as needed, by using the virtual network concept.

Domestic intelligent networks employ computers that function as signal transfer points. These signal transfer points route intelligent network messages (such as identifying 800 number destinations and

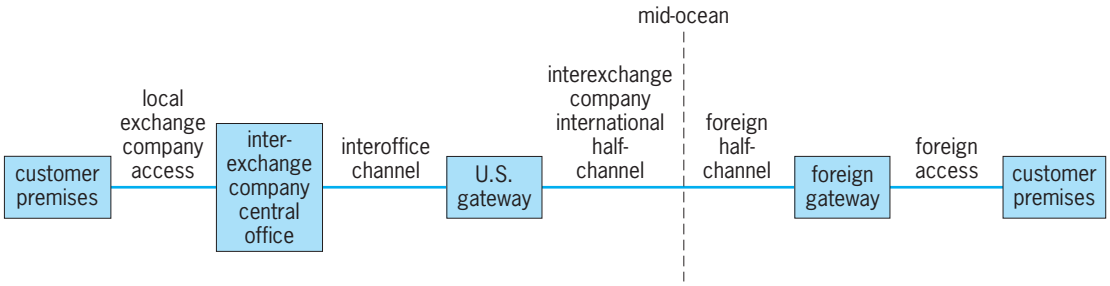

Fig. 5. Virtual private network with international links that must involve service providers of both the United States and other countries. (*AT&T*)

call screening) between switches and advanced network databases. They operate in pairs, with each signal transfer point in a pair capable of picking up the entire load if failure of the other occurs. They communicate constantly, checking on each other's health as well as finding clear paths for calls and providing their status to central operations-maintenance centers. Most major carriers, both local and long distance, use such network operations centers (NOCs) to monitor the health of their networks.

Government and business networks frequently use global services to control costs and to supplement the functionality of their private networks. Data calls, for example, can be funneled through packet-data switches that communicate with packet switches in other countries, so a data message can be transmitted through the telephone network over a variety of available channels at lower cost than if the data were sent over a single leased channel. This enables global credit-card validation, for example, on a real-time basis, with the card company computer located on one continent and the sales point on another.

Private networks are sometimes called overlay networks, because they are superimposed on the public switched telephone networks and share the same facilities. These networks can be enormous in scale, whether domestic or global.

Government networks. The United States government also employs private networks. The original Federal Telecommunications System (FTS) was activated in 1964, and in 1989 was replaced by a new digital voice-data network for government facilities. The FTS 2000 system supplies six communications services: switched voice service for transmission of voice and analog data at up to 4.8 kb/s; switched data service for circuit-switched digital data communications at 56 kb/s and 64 kb/s; switched digital integrated service for digital transmission of voice, data, image, and video signals at transmission rates of up to 1.544 Mb/s (also furnishing the integrated access lines and user-network interfaces needed for ISDN applications); packet-switched service for transmission of data in packet format, including electronic mail service; video transmission service for teleconferencing in both compressed (near-full-motion) and wideband (full-motion) video modes; and dedicated transmission service for point-to-point, dedicated voice and data lines at rates up to 45 Mb/s.

The Department of Defense sponsors the Defense Commercial Telecommunications Network (DCTN), a digital system designed to serve users in the Department of Defense and other federal agencies at about 150 locations. Another Department of Defense system, the Automated Voice Network (AUTOVON), was activated in April 1964, becoming the first worldwide switched network for private telephone and data transmission. AUTOVON is an analog network, divided into the continental United States portion and the overseas portion.

Certain federal government agencies require a high degree of protection against eavesdropping on telephone communications. For this purpose, a special encrypted telephone system called the STU III (Secure Telephone Unit Model III) was developed during the 1980s. It can be used as an ordinary voice telephone, or it can be activated for sending and receiving encrypted voice and data transmission (to and from another such terminal) by turning a special key in the side of the unit. Modified versions of these telephones are available for civilian use. *See* CRYPTOGRAPHY.

Internet. One federal government experiment that began in 1969 became unexpectedly popular in the scientific community and has proliferated into a global system of interconnected networks, now known as the Internet. The first component was the ARPANET, developed by the Defense Advanced Research Projects Agency (DARPA) to provide data communications links between major existing computer facilities and users in academic, industrial, and government research laboratories. ARPANET operates with 56-kb/s digital links using packet switching, and it served as a test for the development of advanced network protocols such as the TCP-IP (Transmission Control Protocol–Internet Protocol). During the 1980s, as desktop personal computers and workstations gained in use, ARPANET was expanded to include members of the NSFnet (sponsored by the National Science Foundation) as well as other data networks. Eventually, several thousand smaller data networks were linked into a global, but informal, Internet.

Telephone circuits form the transmission media for users of the Internet. The circuits are leased from various carriers by the network sponsors, who link into the public network through gateway packet switches. Most of the residential or small-business end users are equipped with desktop or portable computers that employ a modem (modulator-demodulator) to convert the computer's digital signals to analog tones acceptable for transmission over the analog loop lines, and to convert incoming analog tones to digital data accepted by the computer. Most messages are in the form of e-mail (electronic text), but graphics and still images also can be transmitted. Services provided over the Internet are made more accessible through browsers that use a protocol called hypertext linking that allows users to scan the World Wide Web. *See* ELECTRONIC MAIL; INTERNET; WORLD WIDE WEB.

Wireless. Another telephone technology that has gained popularity at unexpected speed is the wireless services.

Cordless telephones. In residences and offices, wireless transmission is handled by cordless phones, which are actually small radiotelephone receiver-transmitter systems. A base unit is plugged into the nearest telephone wall jack for linking to the loop facilities, and also is plugged into an electrical outlet for the necessary alternating current used to power its internal transmitter-receiver as well as to recharge the battery in the portable handset. The range of a cordless telephone is relatively limited and is usually less than 0.5 mi (0.8 km).

Originally, cordless phones were analog, and many were not protected against signal interception, so unauthorized listeners might hear one side of the conversation. More advanced models are based on digital technology and are far more secure during transmission.

Cellular service. A major innovation in wireless technology resulted from work on improving the service to mobile telephones in automobiles (advanced mobile phone service, or AMPS), which was introduced in 1983. This service was called cellular telecommunications because it uses a system of low-power radiotelephone transceivers serving a series of interlinked cells throughout a city or along a heavily traveled highway. As the car carrying the phone moves from one cell to another, the cellular system senses the change in location and transfers, or hands off, the call from one antenna site to another.

Although originally envisioned as an improved service for cars, the cellular service has grown dramatically, propelled more by the use of small, hand-held cellular telephones than by vehicular units. In addition to the convenience offered to telephone users, hand-held cellular phones represent a major advantage for emerging countries that lack the wired infrastructure that was previously required to establish communications.

The rapid growth of cellular-service subscribers has led to congestion in some cellular systems. The solution is to introduce digital cellular systems, which can hande up to 10 times more calls in the same frequency range. However, in countries such as the United States, which already have an analog infrastructure, the FCC has mandated that any digital system must be compatible with the existing analog system. The initial digital designs, based on a technology called time-division multiple access (TDMA), provide a three-times growth factor. Another digital technology, based on code-division multiple access (CDMA), a spread-spectrum technique, could increase the growth factor to 10 or even 20. *See* RADIO SPECTRUM ALLOCATION; SPREAD SPECTRUM COMMUNICATION.

Personal communications service (PCS). Another evolving concept, this is a sort of cellular system, in which a pocket-size telephone is carried by the user. A series of small transmitter-receiver antennas operating at lower power than cellular antennas are installed throughout a city or community (mounted on lampposts and building walls, for example). All the antennas of the personal communications network (PCN) are linked to a master telephone switch that is connected to the main telephone network.

Services to ships and aircraft. Perhaps the oldest form of mobile radiotelephone service is the long-range high-frequency (HF) service centered in Florida, New Jersey, and California to furnish ship-to-shore links. Initiated in 1929, this service uses various radio channels in the bands between 2 and 23 MHz, providing communications links between commercial and private vessels as well as aircraft. The route and progress of a ship or plane must be known in advance, so that the radio frequencies and antennas most suitable to its location can be used. Modern systems employ geostationary satellites operated by the International Maritime Satellite (INMARSAT) organization. That system has eight satellites covering the Atlantic, Pacific, and Indian oceans and is responsible for mobile communications on land, sea, and air.

Services to remote areas. Not all satellite communications links serve mobile telephones. In remote areas, digital packet-switching transceivers are now used to serve as many as 120 telephones each. The conversations are relayed via geostationary satellite to and from metropolitan switching centers linked by cable to the global telephone networks. Similar systems were used in the Persian Gulf War of 1990–1991 to link troops in the Saudi Arabian desert with their families in the United States, and to reestablish the Kuwait telephone links with the rest of the world until the regular facilities could be reconstructed. *See* ELECTRICAL COMMUNICATIONS. John H. Davis

Bibliography. J. Bellamy, *Digital Telephony*, 3d ed., 2000; R. J. Chapuis and A. E. Joel, Jr., *Electronics, Computers and Telephone Switching*, 1990; B. B. Lee, M. Kang, and J. Lee, *Broadband Telecommunications Technology*, 2d ed., 1996; D. E. McDysan and D. L. Spohn, *ATM: Theory and Application*, 1995; A. M. Noll, *Introduction to Telephones and Telephone Systems*, 3d ed., 1999; M. Sexton and A. Reid, *Transmission Networking: SONET and the Synchronous Digital Hierarchy*, 1992; G. A. Silver and M. L. Silver, *Data Communications for Business*, 3d ed., 1994; T. H. Wu, *Fiber Network Service Survivability*, 1992.

Telephone systems construction

Construction of the physical components of the telephone system from the central office to the end user. The telephone system can be roughly divided into three major components: outside plant, switching equipment, and interoffice facilities. The outside plant can be thought of as the local cables that bring the phone service to the subscribers' premises. The switching equipment resides in a building called a central office and serves to route the call to the correct location. The interoffice facilities connect one central office to another. When a call is initiated, the outside plant carries the signal to the local central office. The switching equipment in the local central office then determines the location to which the call needs to be routed. The call is then routed over the interoffice facilities to the central office serving the person being called, where the switching equipment then connects the call to the local outside plant. This article describes the outside plant portion of the telephone system.

Outside plant consists of cables and their supporting facilities. This cable can be aerial (supported by poles), underground in conduit, or buried directly in the ground. There are two main types of cable, metallic and optical.

Metallic cable. Metallic cable consists of usually copper, but at times aluminum, conductors twisted

together in pairs. These conductors are commonly 22, 24, or 26 AWG (American Wire Gauge). A POTS (plain old telephone service) circuit uses a single pair of conductors. Additional services may require more than one pair of wires.

Initial cables consisted of copper conductors, insulated with wood pulp. These conductors were twisted together in pairs, and then the pairs were twisted together in groups. The entire bundle was encased in a lead sheath to create a telephone cable. This type of cable was referred to as lead or pulp cable. The number of individual pairs in a cable depended on the application for which the cable was designed, but could range up to 4200 pairs. In order to facilitate identification of individual pairs, the pulp insulation was color-coded. These cables presented a number of problems: Lead was found to be a hazardous material; the lead sheath was prone to cracking when flexed, especially in cold weather; the pulp insulation wicked moisture easily; and the color coding on the conductors would fade with time

Modern telecommunications cable consists of copper conductors coated with a plastic color-coded insulation. There are a number of different plastic insulations available, depending on the application, manufacturer, and local practice. The pairs are then twisted together and grouped into 25-pair complements, which are further twisted together to make a cable. The cable is then covered in a protective sheath, also called a turnplate, made of aluminum or steel (see **illustration**). The turnplate provides mechanical support to the cable and also provides electromagnetic shielding to the conductors in the cable. Sometimes additional armoring materials are added to the sheath, typically for resistance to abrasion or to prevent damage from burrowing animals in buried applications. The turnplate is then covered in a sheath made of polyethylene. Cable of this type is referred to as PIC (pronounced "pick") for plastic insulated cable. The center of the cable can be left as is, and is referred to as air core; or it can be flooded with a moisture-repelling compound, and is referred to as jelly-filled cable. Because of the sticky nature of this compound, people who work with this type of cable often call it "icky-PIC."

Fiber cable. The second major type of cable deployed in the outside plant is optical fiber. Optical fiber consists of a glass core, coated with a glass cladding, and then coated with plastic. Light is transmitted along the core of the fiber. The interface between the core and the cladding causes the light to reflect back into the core, and in effect "bounce" back and forth as it travels along the core. Two major types of fiber are used in telephony, single-mode and multimode. A single-mode fiber has a core that ranges 8.3–10 micrometers in diameter. A multimode fiber has a core that is 50–100 μm in diameter. In a single-mode fiber, there is only enough "room" for the light to take one path along the core. In a multimode fiber, the light can reflect in a number of ways and take multiple paths along the fiber. Single-mode fiber provides higher transmission speeds over longer distances; however, it is more expensive and requires a more precise light source. *See* OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

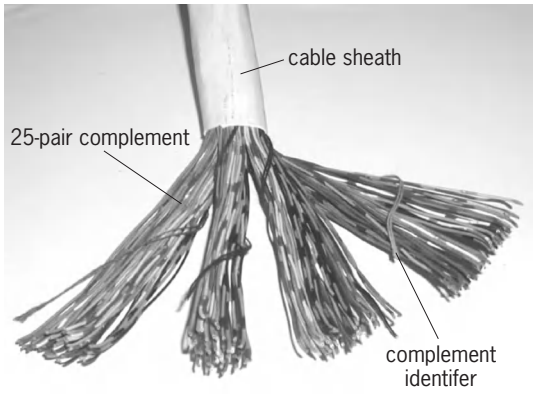

A 100-pair plastic insulated cable. (*From J. Clayton, McGraw-Hill Illustrated Telecom Dictionary, 4th ed., McGraw-Hill, 2002*)

Cables are made up of multiple fibers. In a loose tube configuration, individual fibers are placed in plastic tubes. These tubes are larger then the fibers contained within, and may sometimes be flooded with a water-repelling compound. This oversized tube serves to prevent stresses due to handling the cable from being transferred to the delicate fibers. The tubes are gathered together around a central strength member, either metal or plastic. The whole bundle is then wrapped with an aramid yarn for strength and covered with a plastic sheath. In buffered cables, the individual fibers are covered with a protective plastic coating, and directly covered in the supporting yarns and sheath. Fibers can also be ganged together in a flat ribbon consisting of typically 6, 12, or 24 fibers held together by a plastic ribbon material. These fibers are then "stacked" and placed in a single tube, around which the cable is constructed. Each method has different applications. The cables vary in degrees of cost, toughness, ability to access individual fibers, and ease of splicing. As with metallic cable, fiber cable is also available with an armored sheath. *See* COMMUNICATIONS CABLE.

Aerial plant. Aerial plant consists of telephone cable that is suspended from structures, typically telephone poles. The majority of telephone poles are made of wood. Typical woods used are southern pine, western red cedar, and Douglas-fir. Other woods are used as well, including jack pine, red pine, and eastern cedar. The choice of wood depends mainly on the location of the manufacturer's wood lots. Utility poles are usually treated with chemicals to increase their rot resistance; the treatment method used depends on the manufacturer and local practice. While the majority of poles are wood, alternative products are available, including fiberglass, concrete, metal, and laminated wood poles. All of these materials have certain benefits and drawbacks, and are usually used only in specific situations.

A telephone pole is set into a hole in the ground. The depth of the hole depends on the size of the pole and how much weight it is designed to carry, but typically ranges 5–10 ft (1.5–3 m). In the past, these holes were dug with hand tools, and the poles were carried and tipped up into the holes by hand with the assistance of long poles with sharp tips called pike poles. Currently, the majority of poles are placed using a digger derrick machine. This machine consists of a truck with a powered auger that is used to dig the hole, and a powered boom with a winch and clamping jaws that is used to set the pole. However, in some situations such as in a backyard, the digger derrick cannot be maneuvered into position, and the poles must still be carried and set by hand.

The telephone poles are connected by a steel strand that is anchored to each pole. Cable is placed by hanging rollers, called blocks, on the strand. The cable is then pulled through the blocks. There are a variety of methods for performing this step. The difference lies in where the cable reel is placed and how the truck is moved. The decision on what method is used depends on the size of the job, the physical layout of the area, and personal preference.

Once the cable is suspended in the blocks, a machine called a cable lasher is placed over the cable and strand. Pulled along the cable, it wraps a wire around the cable and strand, lashing the cable snugly to the strand, and it also pushes the cable blocks along the strand ahead of it.

Underground plant. Underground plant consists of facilities that are placed in conduit and vaults. Vaults are typically constructed of concrete. They either can be precast and set into an excavation by crane or can be cast in place using forms. Precast vaults are available in a variety of sizes, and cast-in-place ones can be any size desired. However, a typical telecommunications vault is approximately 7 ft (2.5 m) deep, 6 ft (2 m) wide, and 12 ft (4 m) long. Most telecommunication vaults have two lids and many have four, depending on size. The metal cover and frame sit on top of a cement or brick chimney that extends down to the roof of the vault. The height of this chimney can be adjusted to accommodate changes in road height.

Vaults are connected by conduit, which is also known as ducts. At present, conduit is typically 4-in. (10-cm) polyvinyl chloride (PVC) pipe. In the past a variety of materials, including clay tile, wood, cement, and steel, were used as duct. A trench is excavated between vaults, and a number of conduits are placed there before the trench is backfilled. Conduit can also be run from a vault to a building basement, or to the base of a pole where the cable can emerge, run up the pole, and continue as an aerial cable.

When the conduit system is constructed, a pulling line is installed in each duct for the installation of cable. If the pulling line is missing, the duct can be rodded from vault to vault using a series of stiff steel rods and a new rope can be pulled in. Rodding can also be used to remove obstructions that may occur from a crack in a duct.

Once a rope is in place, it is used to pull in a cable. The cable is prepared with a special lubricant, and it is winched through the duct. A variety of rollers and shoes are available to prevent the cable from kinking and binding as it is pulled down into the vault and into the duct. Once in place, the cable is racked onto metal framework along the sides of the vault.

Buried plant. Buried plant consists of cables that are buried directly in the ground. This method is much cheaper than constructing a conduit system, but lacks flexibility. In a conduit system, if a cable needs to be replaced or upgraded, it can be pulled out and a new one pulled in. With direct-buried cable, the entire trench must be reexcavated. Direct-buried cable is also susceptible to damage from rocks. Cable should be bedded in sand or some other suitable fill, but rocks can work their way to the cable and eventually puncture the sheath. Direct-buried cables are usually used in limited situations in garden apartments, planned communities, and complexes. Such cables are usually made specifically for being buried and are armored and jelly-filled.

Splicing. Once cable is placed in the outside plant, it must be connected by splicing. There are a number of methods used to join copper conductors. In the past, conductors were bared, twisted together, and soldered. In today's outside plant, special connectors are used. Some connectors permanently join together a single conductor. If such a connector needs to be removed, it must be cut out and a new connector used. Other connectors are designed to be pluggable. Pluggable connectors make use of the fact that cables are built in 25-pair complements. Pluggable connectors are designed to take all of the conductors in a 25-pair complement and terminate the ends of the conductors into a large plug module. The modules can then be snapped together and apart to reconfigure the cable as needed. All modern connectors are insulation displacement connectors (IDC). The wire does not need to be stripped. The connector is applied with a pliers-like tool and pierces the insulation to make contact with the conductor inside.

Fiber cable is also spliced in a number of ways. Individual fibers can be joined using either mechanical or fusion methods. With mechanical methods, the ends of the fiber are cleaned and cleaved to create a flat clean end, and then placed in a device that holds them precisely together. There are also pluggable connectors that can be applied to the ends of fibers, enabling them to be connected and disconnected. These are often used when connecting fibers to equipment. The fusion method involves use of an electric arc. The ends of the fibers are prepared by cleaning and cleaving, and then placed into special holders. These holders are placed into a fusion splicing machine, which precisely aligns the ends of the fiber and applies a brief high-voltage arc. This arc fuses the ends of the two fibers together. Machines are available that can splice entire ribbons of fiber together at once. Mechanical splicing requires a much lower investment in tools, but costs more per splice and produces an inferior splice that causes more loss in the fiber. Tools for fusion splicing are

quite expensive, but the cost per splice is much less and the quality of the splice is higher.

Once a splice is created, a splice closure needs to be placed to protect it. In the past, splices were encased in muslin and then covered in molten lead in a technique known as wiping. Modern splice closures are plastic cases that snap around the open cable. They are available in a number of sizes and manufacturing methods.

Air pressure. If moisture gets into a copper cable, it will begin to act as a path for electricity to flow between conductors. At first, this will create noise or static on a POTS line or degrade a data circuit. If the cable becomes wet enough, the conductors in the cable will short and the circuit stops working. Telecommunications cables are pressurized with air to help prevent this failure.

An air system consists of dryers, air pipes, and transducers. An air dryer provides a source of pressurized dry air. Most air dryers are located in a central office, but on longer cable runs they can be located in the field as well. The compressed dry air flows from the dryer to a variety of manifolds and pipe panels, which break the single air feed from the dryer into a number of smaller air pipes for each cable. Transducers are mounted on the pipe panel to monitor air pressure and flow.

The air pipes are run next to the cables, and at regular intervals are tapped into the cable via a valve installed through the cable sheath. This valve provides pressurized air to the air core of a cable. Any breach in a cable's sheath that would let water in instead forces air out. Thus, this pressure difference protects the cable against water.

Pressure and flow transducers are also placed at regular intervals along the cable run. These transducers are monitored. An increase in flow or a decrease in pressure serves as an early warning that a cable may be damaged.

Cables placed underground in a conduit system are pressurized, as are some sections of cables in the air. When a cable changes from pressurized to nonpressurized, a plug of epoxy is poured in the cable to stop the airflow. When splice cases are placed on pressurized sections of cables, the cases must also be airtight.

Additional equipment. Besides the basic construction of the outside plant portion of a telephone network, there are many additional pieces of equipment. Various electronics can be placed in the field. They can serve to extend the distance a signal can travel, increase the number of voice circuits that can be carried on a particular copper pair, or provide higher-speed and -bandwidth data circuits. These electronics can be placed in small above-ground buildings called huts; in buried vaults called controlled environment vaults (CEV) that have heating, cooling, lighting, and dehumidification; or in above-ground mounted cabinets. New equipment is always being developed, and as it is deployed old equipment must be removed, due to either space limitations or incompatibility with the new equipment.

David F. Dolch

Bibliography. J. Clayton, *McGraw-Hill Illustrated Telecom Dictionary*, 4th ed., 2002; R. A. Meyers, *Encyclopedia of Telecommunications*, 1989; A. M. Noll, *Introduction to Telephones and Telephone Systems*, 3d ed., 1999.

Telescope

An instrument used to collect, measure, or analyze electromagnetic radiation from distant objects. A telescope overcomes the limitations of the eye by increasing the ability to see faint objects and discern fine details. In addition, when used in conjunction with modern detectors, a telescope can "see" light that is otherwise invisible. The wavelength of the light of interest can have a profound effect on the design of a telescope. *See* ELECTROMAGNETIC RADIATION; LIGHT.

Parameters. The utility of a telescope depends on its ability to collect large quantities of light and to resolve fine details. The brightness of an image is proportional to the area of the light-gathering element, which is proportional to the square of that element's aperture. The brightness also depends on the area over which the image is spread. This area is inversely proportional to the square of the focal length (f) of the lens. The brightness of the image therefore depends on the square of the f/ratio, just as in an ordinary camera. The resolving power of a telescope depends on the diameter of the aperture and the wavelength observed; the larger the diameter, the smaller the detail that can be resolved. An 11-cm-aperture (4.5-in.) telescope will resolve approximately 1-arcsecond details. On Palomar Mountain, California, the 5-m (200-in.) Hale telescope has an aperture that is roughly 1000 times the diameter of the dark-adapted eye. Thus, it gathers 1 million times more light than the eye. Theoretically, it can resolve details as small as 0.02 arcsecond, but because it must look at celestial objects through the Earth's turbulent atmosphere, it cannot resolve details as fine as the theoretical limit.

In a telescope, the angular separation A of two points on the object is magnified to the angle A' as seen by the eye. The magnification (A'/A) is numerically equal to the ratio of the focal length of the main optical element, f_1, to the focal length of the eyepiece, f_2; that is, the magnification is f_1/f_2. This is true whether the main optical element is a lens, a mirror, or a combination of lenses and mirrors. *See* FOCAL LENGTH; LENS (OPTICS).

Limitations. For many applications, the Earth's atmosphere limits the effectiveness of larger telescopes. The most obvious deleterious effect is image scintillation and motion, collectively known as poor seeing. Atmospheric turbulence produces an extremely rapid motion of the image resulting in a smearing. On the very best nights at ideal observing sites, the image of a star will be spread out over a 0.25-arcsecond seeing disk; on an average night, the seeing disk may be between 0.5 and 2.0 arcseconds. It has been demonstrated that most of the air currents

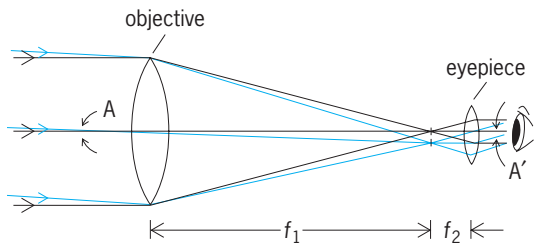

Fig. 1. Simplified optical diagram of a refracting telescope.

that cause poor seeing occur within the observatory buildings themselves. Substantial improvements in seeing have been achieved by modern design of observatory structures.

The upper atmosphere glows faintly because of the constant influx of charged particles from the Sun. This airglow adds a background exposure or fog to photographic plates that depends on the length of the exposure and the speed (*f*/ratio) of the telescope. The combination of the finite size of the seeing disk of stars and the presence of airglow limits the telescope's ability to see faint objects. One solution is placing a large telescope in orbit above the atmosphere. In practice, the effects of air and light pollution outweigh those of airglow at most observatories in the United States. *See* AIRGLOW.

Optical telescopes. There are basically three types of optical systems in use in astronomical telescopes: refracting systems whose main optical elements are lenses which focus light by refraction; reflecting systems, whose main imaging elements are mirrors which focus light by reflection; and catadioptric systems, whose main elements are a combination of a lens and a mirror. The most notable example of the last type is the Schmidt camera. In each case, the main optical element, or objective, collects the light from a distant object and focuses it into an image that can then be examined by some means.

Refracting telescopes. Small refracting telescopes are used in binoculars, cameras, gunsights, galvanometers, periscopes, surveying instruments, rangefinders, astronomical telescopes, and a great variety of other devices. Parallel or nearly parallel light from the distant object enters from the left, and the objective lens forms an inverted image of it (**Fig. 1**). The inverted image is viewed with the aid of a second lens, called the eyepiece. The eyepiece is adjusted (focused) to form a parallel bundle of rays so that the image of the object may be viewed by the eye without strain. The objective lens is typically compound; that is, it is made up of two or more pieces of glass, of different types, designed to correct for aberrations such as chromatic aberration. To construct a visual refractor, a lens is placed beyond the images formed by the objective and viewed with the eye. To construct a photographic refractor or simply a camera, a photographic plate is placed at the position of the image (**Fig. 2**).

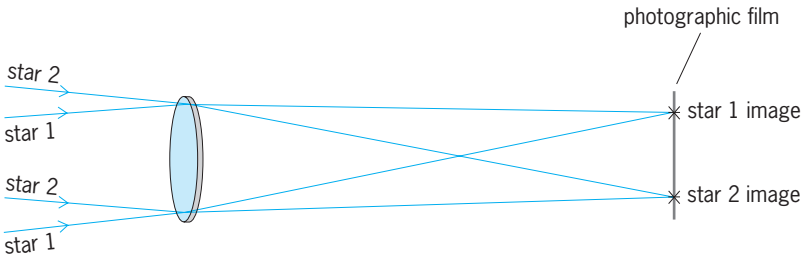

Fig. 2. Refracting optical system used to photograph a star field.

Generally, refracting telescopes are used in applications where great magnification is required, namely, in planetary studies and in astrometry, the measurement of star positions and motions. However, this practice is changing, and the traditional roles of refractors are being carried out effectively by a few reflecting telescopes, in part because of effective limitations on the size of refracting telescopes.

A refractor lens must be relatively thin to avoid excessive absorption of light in the glass. On the other hand, the lens can be supported only around its edge and thus is subject to sagging distortions that change as the telescope is pointed from the horizon to the zenith; thus its thickness must be great enough to give it mechanical rigidity. An effective compromise between these two demands is extremely difficult, making larger refractors unfeasible.

The largest refracting telescope is the 1-m (40-in.) telescope—built over a century ago—at Yerkes Observatory. This size is about the limit for optical glass lenses.

Reflecting telescopes. The principal optical element, or objective, of a reflecting telescope is a mirror. The mirror forms an image of a celestial object (**Fig. 3**) which is then examined with an eyepiece, photographed, or studied in some other manner.

Reflecting telescopes generally do not suffer from the size limitations of refracting telescopes. The mirrors in these telescopes can be as thick as necessary and can be supported by mechanisms that prevent sagging and thus inhibit excessive distortion. In addition, mirror materials having vanishingly small expansion coefficients, together with ribbing techniques that allow rapid equalization of thermal gradients in a mirror, have eliminated the major thermal problems plaguing telescope mirrors. Some advanced reflecting telescopes use segmented mirrors, composed of many separate pieces.

By using a second mirror (and even a third one, in some telescopes), the optical path in a reflector can be folded back on itself (**Fig. 4***a*), permitting a long focal length to be attained with an instrument housed in a short tube. A short tube can be held by a smaller mounting system and can be housed in a smaller dome than a long-tube refractor.

A variety of optical arrangements are possible in large reflecting telescopes, including the prime focus, the newtonian focus, the Cassegrain focus, and the coudé focus.

The newtonian focus is probably most widely used by amateur astronomers in reflectors having apertures on the order of 6 in. (15 cm; Fig. 3*b*). A flat mirror placed at 45° to the optical axis of the primary mirror diverts the focused beam to the side of the telescope, the image being formed by the paraboloidal primary mirror alone. An eyepiece,

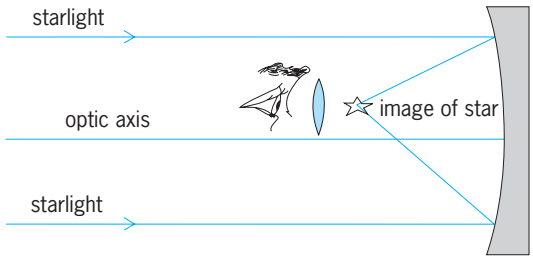

Fig. 3. Viewing a star with a reflecting telescope. In this configuration, the observer may block the mirror unless it is a very large telescope.

camera, or other accessories can be attached to the side of the telescope tube to study the image. In the largest telescopes, with apertures over 100 in. (2.5 m), provision is not usually made for a newtonian focus. Instead, an observing cage is placed inside the tube, where the observer can take accessories to observe the image formed by the primary mirror. This prime focus is identical to the newtonian focus optically, since the newtonian flat does nothing more than divert the light beam. The modern reflectors have fast primary mirrors, so that the focal ratio at the prime focus is *f*/2.5 to *f*/6. Lower focal ratios permit shorter exposures on extended objects such as comets and nebulae.

A Cassegrain system consists of a primary mirror with a hole bored through its center, and a convex secondary mirror that reflects the light beam back through the central hole to be observed behind the primary mirror (Fig. 3*a*). Since the secondary mirror is convex, it decreases the convergence of the light beam and increases the focal length of the system as a whole. The higher focal ratios (*f*/8 to *f*/13) of Cassegrain systems permit the astronomer to observe extended objects, like planets, at higher spatial resolution and to isolate individual stars from their neighbors for detailed studies.

The classical Cassegrain system consists of a paraboloidal primary mirror and a hyperboloidal secondary mirror. The newtonian focus, prime focus, and Cassegrain focus are not affected by spherical aberration. However, all of the systems are plagued by coma, an optical aberration of the paraboloidal primary. Coma causes a point source off the center of the field of view to be spread out into a comet-shaped image. To correct for the effect of coma, a corrector lens is often used in front of the photographic plate. The design of prime-focus corrector lenses is a major consideration in large telescope design, since the corrector lens itself can introduce additional aberrations. *See* ABERRATION (OPTICS); ASTRONOMICAL PHOTOGRAPHY; MIRROR OPTICS.

To avoid complicated corrector lenses at the Cassegrain focus, the Ritchey-Chrétien system, an alternate design with both a hyperboloidal primary and a hyperboloidal secondary, is used in modern telescopes. This arrangement is not affected by coma or spherical aberration, so it has a wider field of view than the classical Cassegrain.

Reflecting telescopes are often used in applications where great light-gathering capabilities are required. Since 1990, a significant number of large astronomical telescopes have been built with mirrors over 8 m (315 in.) in diameter.

Catadioptric telescopes. Catadioptric telescopes combine both mirrors and lenses. This combination is generally used to image wide fields. Catadioptric telescopes with apertures of 15–18 cm (5–6 in.) are popular with amateur astronomers. One variant of the catadioptric system is the Schmidt design.

The Schmidt is an optical system used almost exclusively for photographic applications such as sky surveys, monitoring of galaxies for supernova explosions, and studies of comet tails. The primary mirror of a Schmidt camera has a spherical shape, and therefore suffers from spherical aberration. To correct for this problem, the light passes through a thin corrector plate as it enters the tube. Schmidt camera correcting plates are among the largest lenses made for astronomical applications. The special features of this system combine to produce good images over a far larger angular field than can be obtained in a Cassegrain. *See* SCHMIDT CAMERA.

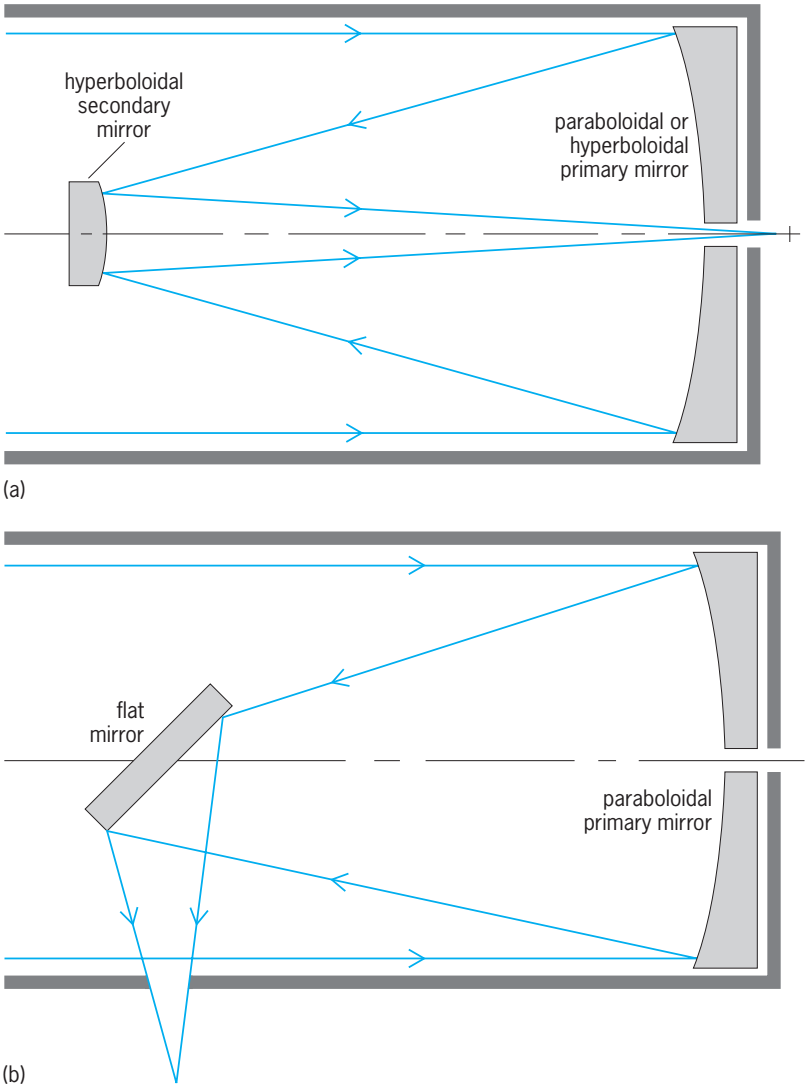

Fig. 4. Diagrams of reflecting telescopes. (*a*) Cassegrain telescope, with either classical or Ritchey-Chrétien optics. (*b*) Newtonian telescope.

Tools. Astronomers seldom use large telescopes for visual observations. Instead, they record their data for future study. Modern developments in photoelectric imaging devices are supplanting photographic techniques for many applications. The great advantages of detectors such as charge-coupled devices is their high sensitivity, and the images can be read out in a computer-compatible format for immediate analysis. *See* CHARGE-COUPLED DEVICES.

Light received from most astronomical objects is made up of radiation of all wavelengths. The spectral characteristics of this radiation may be extracted by special instruments called spectrographs. Wide field coverage is not critical in spectroscopy of stellar objects, so spectrographs are mounted at the Cassegrain and coudé foci. In the coudé configuration, flat mirrors divert the focused beam so that the focal point is fixed in position, regardless of where the telescope points in the sky. Thus a spectrograph or other instrument that is too heavy or too delicate to be mounted on the moving telescope tube can be placed at the coudé focus. *See* ASTRONOMICAL SPECTROSCOPY.

Photoelectric imaging devices may be used in conjunction with spectrographs to record spectral information. Photoelectric detectors are useful tools for classifying stars, monitoring variable stars, and quantitatively measuring the light flux from any astronomical object. Photometry carried out with different filters yields basic information about the source with shorter observing time than that required for a complete spectroscopic analysis. *See* PHOTOMETRY.

Collectors of radiant energy. As collectors of radiation from a specific direction, telescopes may be classified as focusing and nonfocusing. Nonfocusing telescopes are used for radiation with energies of x-rays and above. Focusing telescopes, intended for nonvisible wavelengths, are similar to optical ones, but they differ in the details of construction.

Solar telescopes. Solar instrumentation differs from that designed to study other celestial objects, since the Sun emits great amounts of light energy. In one design (**Fig. 5**), a large heliostar reflects sunlight down the fixed telescope tube to a spectrograph. This spectrograph is evacuated to avoid problems that would be created by hot air currents. In an alternative design, the heliostat feeds light into a vertical telescope that is evacuated. *See* SUN.

Radio telescopes. Radio telescopes utilize mirrors of very large size that, because of the long wavelength for which they are used, may consist of only an open wire mesh. Because of the limitation by diffraction, arrays of large reflectors have been constructed. The output of each mirror of the array is combined in a process called aperture synthesis to yield a resolution roughly equivalent to that provided by a telescope the size of the array. An example of this type is the Very Large Array (VLA) near Socorro, New Mexico. *See* RADIO TELESCOPE.

Infrared telescopes. In infrared telescopes the secondary mirror (that near *A* in illus. *b*) is caused to

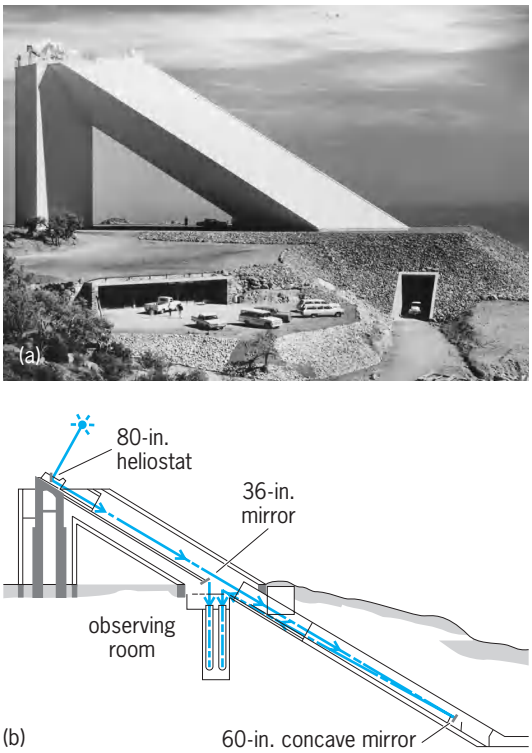

Fig. 5. Solar telescope configuration. (*a*) McMath 1.5-m (60-in.) solar telescope, Kitt Peak, Arizona. (*b*) Diagram of the optical elements in the telescope. 1 in. = 2.5 cm.

oscillate rotationally about an axis through a diameter. This motion causes an infrared detector at *B* to see alternately the sky and the sky plus the desired object. The signals received at these two mirror positions are subtracted and, as a consequence, the large background radiation received both from the atmosphere and from the telescope is canceled. However, because of the random nature of thermal radiation, the fluctuations of the background emission are not canceled. Thus infrared telescopes are additionally designed to reduce the telescope background radiation and its fluctuations. Two large infrared telescopes are the 3.8-m (150-in.) United Kingdom Infrared Telescope and the 3-m (120-in.) NASA Infrared Telescope Facility on Mauna Kea. *See* INFRARED ASTRONOMY.

Ultraviolet telescopes. Ultraviolet telescopes have special mirror coatings with high ultraviolet reflectivity. Since the atmosphere is not transparent below 300 nanometers, ultraviolet telescopes are usually flown above the atmosphere either in rockets or in orbiting spacecraft. An example is the *International Ultraviolet Explorer*. *See* ULTRAVIOLET ASTRONOMY.

X-ray telescopes. X-ray telescopes must be used above the atmosphere and are flown in rockets or satellites. The focusing type uses an unusual optical design in which the reflection from the surfaces occurs at nearly grazing incidence. This is the only way of achieving reflective optics for x-rays.

Nonfocusing x-ray telescopes use opaque heavy-metal (lead) channels or tubes in front of an x-ray detector to confine the directional sensitivity of the detector. The detectors may be proportional

counters or scintillation detectors. *See* X-RAY TELESCOPE.

Gamma-ray telescopes. Gamma-ray telescopes use coincidence and anticoincidence circuits with scintillation or semiconductor detectors to obtain directional discrimination. With coincidence counting, two or more detectors in a line must give a simultaneous detection for a gamma ray to be counted. Other detectors are often used to surround the telescope to reduce the unwanted background arising from undesired particles. A simultaneous count received in one or more of these shielding detectors nullifies (by anticoincidence) the detection otherwise registered in the coincidence detectors. Thus gamma rays that trigger only the coincidence circuits are detected. Since many gamma rays are produced within the atmosphere by other particles, the telescopes are usually flown in balloons, rockets, or satellites. However, at very high energies (more than 100 GeV), ground-based techniques have been used either to detect the Cerenkov light from the shower of electrons produced when a gamma ray hits the atmosphere or to detect directly the particles that penetrate to the ground. *See* GAMMA-RAY ASTRONOMY.

Cosmic-ray telescopes. Cosmic-ray telescopes are used to detect primary protons or heavier-element nuclei or to detect the products produced when these particles interact with the atmosphere. In its simplest form a cosmic-ray telescope may consist of nuclear track emulsions borne aloft in balloons or spacecraft. A very large cosmic-ray telescope deep in a mine in Utah detects penetrating mesons in the 1–100-teraelectronvolt range. This telescope uses Cerenkov detectors in combination with plane parallel arrays of cylindrical spark-tube counters. The Cerenkov counter detects the presence of a high-energy particle and triggers the acoustic sensing of sparks in the counters. A computer is used to analyze the spark data and determine the direction of the incoming meson. The arrival of high-energy cosmic rays appears to be isotropic, most likely because of scattering by magnetic fields of the galaxy. *See* CERENKOV RADIATION; COSMIC RAYS.

Neutrino telescopes. Although neutrinos are not electromagnetic radiation, devices to collect astrophysically generated neutrinos are, by analogy, often called neutrino telescopes. Neutrinos from several astrophysical sources have been recorded by omnidirectional detectors operating in deep mines. Neutrinos from the Sun have been observed since the 1970s by a detector in a mine in South Dakota. Several underground detectors registered neutrinos from supernova 1987A in February 1987. *See* NEUTRINO; SOLAR NEUTRINOS; SUPERNOVA.

The newest major neutrino detector is the Sudbury Neutrino Observatory (SNO), which was built 2100 m (6800 ft) belowground in a mine near Sudbury, Ontario, and is operated by the Queens University. SNO is designed to detect neutrinos produced by fusion reactions in the Sun and in supernovae in the Milky Way Galaxy. SNO consists of a 12-m-diameter (40-ft) acrylic sphere, filled with 1000 tons of heavy water (D_2O). When neutrinos react with the heavy water in the sphere, it produces flashes of Cerenkov radiation. This light is then detected by an array of 9600 photomultiplier tubes surrounding the heavy-water vessel. The observatory, located in the deep mine, is shielded from cosmic rays by the intervening layers of rock. The detector laboratory was designed to reduce background radiation signals, which might obscure the very weak neutrino signal. *See* NEUTRINO ASTRONOMY.

Notable telescopes. The definition of a large telescope depends on its type. For a refracting telescope to be considered a large telescope, its objective lens must be larger than about 0.6 m (24 in.), whereas a reflecting telescope will have to exceed 79 in. (2 m) to be considered large (see **table**). A few telescopes that are notable for their large size or innovative design will be discussed. *See* ASTRONOMICAL OBSERVATORY.

The 1.02-m (40-in.) refractor at the Yerkes Observatory has been a major contributor to astronomy. The 40-in. objective lens has a focal length of 19 m (62 ft) and is housed in a 23-m (90-ft) dome.

The 5-m (200-in.) Hale telescope at Palomar Mountain, California, was completed in 1950. The primary mirror is 5 m in diameter with a 1.02-m (40-in.) hole in the center. Its focal length is 14 m (660 in.) and it has a paraboloidal figure. The focal ratio of the prime focus is *f*/3.3, and of the Cassegrain focus *f*/16.

The 4-m (158-in.) Mayall reflector at the Kitt Peak National Observatory was dedicated in 1973. The 158-in. mirror has an *f*/2.7 hyperboloidal shape and is made from a 61-cm-thick (24-in.) fused quartz disk which is supported in an advanced-design mirror cell. The prime focus has a field of view six times greater than that of the Hale reflector. An identical telescope was subsequently installed at Cerro Tololo Inter-American Observatory, in Chile.

The mirrors for these traditional large telescopes were all produced using the same general methodology. A large, thick glass mirror blank was first cast; then the top surface of the mirror was laboriously ground and polished to the requisite shape. In the process, great quantities of glass were removed from the blank. The remaining glass helps to give the mirror the required structural rigidity as it is tilted at different angles as the telescope is pointed to different parts of the sky.

The practical and economical limit to the size of traditional mirror designs was nearly reached by the 6-m (236–in.) telescope in the Caucasus Mountains, Russia. Newer telescopes have been designed and built that use either a number of mirrors mounted such that the light collection by them is brought to a common focus, or lightweight mirrors in computer-controlled mounts.

In a multi-mirror telescope (MMT), all mirrors are of similar shape, but additional optical elements are introduced to bring the light to a common focus. An example was the MMT on Mount Hopkins in Arizona, in which six 1.8-m (72-in.) mirrors were combined to collect as much light as a single 4.5-m (177-in.)

Large telescopes of the world

Mirror diameter m	Mirror diameter in.	Observatory	Year completed
		Some of the largest reflecting telescopes	
10.4	404	Gran Telescopio Canarias (GTC), La Palma, Canary Islands	2006
10.0 × 2	400 × 2	Keck Telescopes, Mauna Kea, Hawaii	1993–1996
10.0	394	South African Large Telescope (SALT), Sutherland, South Africa	2005
9.0	360	Hobby-Eberly Telescope, McDonald Observatory, Fort Davis, Texas	1997
8.4 × 2	331 × 2	Large Binocular Telescope (LBT), Mount Graham, Arizona	2005
8.2	319	Subaru Telescope, National Astronomical Observatory, Japan; Mauna Kea, Hawaii	2000
8.2 × 4	320 × 4	Very Large Telescope (VLT), ESO Paranal Observatory, Cerro Paranal, Chile	1998–2000
8.1	319	Gemini North (Gillett), National Optical Astronomy Observatory, Mauna Kea, Hawaii	1999
8.1	319	Gemini South, National Optical Astronomy Observatory, Cerro Pachon, Chile	2000
6.5	256	Multi-Mirror Telescope, Mount Hopkins, Arizona	1999
6.5 × 2	256 × 2	Magellan Telescopes, Cerro Las Campanas, Chile	2000–2001
6.0	236	Special Astrophysical Observatory, Zelenchukskaya, Caucasus, Russia	1976
5.1	200	California Institute of Technology/Palomar Observatory, Palomar Mountain, California	1950
4.2	165	William Herschel Telescope, La Palma, Canary Islands	1987
4.1	161	Southern Astrophysical Rearch (SOAR) Telescope, Cerro Pachon, Chile	2006
3.8	147	Mayall Telescope Kitt Peak National Observatory, Arizona	1973
4.0	158	Cerro Tololo Inter-American Observatory, Chile	1976
4.0	158	Visible and Infrared Survey Telescope for Astronomy (VISTA), ESO Paranal Observatory, Cerro Paranal, Chile	2007
3.9	153	Anglo-Australian Telescope, Siding Spring Observatory, Australia	1975
3.8	150	United Kingdom Infrared Telescope, Mauna Kea Observatory, Hawaii	1978
3.6	144	Canada-France-Hawaii Telescope, Mauna Kea Observatory, Hawaii	1979
3.6	142	Cerro La Silla European Southern Observatory, Chile	1976
3.6	141	New Technology Telescope, European Southern Observatory, Chile	1989
3.5	138	Calar Alto Observatory, Calar Alto, Spain	1984
3.5	138	Astrophysics Research Consortium, Apache Point, New Mexico	1993
3.5	138	Wisconsin-Indiana-Yale-NOAO (WIYN), Kitt Peak, Arizona	1994
3.6	140	Telescopio Nazionale Galileo (TNG), La Palma, Canary Islands	1998
3.2	126	NASA Infrared Telescope, Mauna Kea Observatory, Hawaii	1979
3.0	120	Lick Observatory, Mount Hamilton, California	1959
2.7	107	McDonald Observatory, Fort Davis, Texas	1968
2.6	102	Crimean Astrophysical Observatory, Ukraine	1960
2.6	102	Byurakan Observatory, Yerevan, Armenia	1976
2.6	102	VLT Survey Telescope, ESO Paranal Observatory, Cerro Paranal, Chile	2006
2.5	100	Nordic Optical Telescope, La Palma, Canary Islands	1989
2.5	100	Mount Wilson and Las Campanas, Mount Wilson, California	1917
2.5	100	Carnegie Southern Observatory, Cerro Las Campanas, Chile	1976
2.5	96	Isaac Newton Telescope, La Palma, Canary Islands	1984
2.5	98	Sloan Telescope, Apache Pont, New Mexico	1998
2.4	94	University of Michigan-Dartmouth College-Massachusetts Institute of Technology, Kitt Peak, Arizona	1986
2.4	94	Hubble Space Telescope, low Earth orbit	1990
		Largest refracting telescopes	
1.02	40	Yerkes Observatory, Williams Bay, Wisconsin	1897
0.91	36	Lick Observatory, Mount Hamilton, California	1888
0.83	33	Observatoire de Paris, Meudon, France	1893
0.80	32	Astrophysikalisches Observatory, Potsdam, Germany	1899
0.76	30	Allegheny Observatory, Pittsburgh, Pennsylvania	1914

mirror. The multi-mirror design for this particular telescope was abandoned in favor of a single mirror design. Segmented mirror telescopes (SMT) have also been built.

New generation telescopes. The Keck Telescopes on Mauna Kea, Hawaii, completed in 1993–1996, are the largest of the segmented mirror telescopes to be put into operation (**Fig. 6**). The telescopes are a fairly traditional design. However, their primary mirrors are made up of 36 individual hexagonal segments mosaiced together to form single 10-m (386-in.) mirrors. Electronic sensors built into the edges of the segments monitor their relative positions, and feed the results to a computer-controlled actuator system which maintains the relative positions of the segments to 5 nm (0.2×10^{-6} in.). The mosaic has four times the light-gathering power of the 5-m (200-in.) Hale Telescope, though its weight is the same as the 200-in. mirror. The two Keck telescopes are 85 m (280 ft) apart.

Work is under way to combine the light of the two 10-m telescopes to produce an optical interferometer. This requires a number of 2-m (80-in.) telescopes in addition to the Keck giants. By combining the light from the large telescopes and the 2-m telescopes, the interferometer will be capable of finding Jupiter-size planets orbiting nearby stars. *See* EXTRASOLAR PLANETS; INTERFEROMETRY.

In 1989, the European Southern Observatory put into operation their New Technology Telescope. The 3.58-m (141-in.) mirror was produced by a technique known as spin-casting, where molten glass is poured into a rotating mold. As the glass cools and solidifies, the surface of the relatively thin mirror takes

on a shape that is close to the desired one, reducing substantially the need for grinding away excess glass. In operation, computers monitor the mirror's shape, and direct a support system that will maintain the proper shape.

The Hobby-Eberly Telescope (HET) is a 9-m (360-in.) telescope built as an international collaboration between the University of Texas at Austin, Pennsylvania State University, and Stanford University in the United States and two universities in Germany. It uses a 91-segment spherical main mirror that is tipped at a constant 35° from the zenith. To move from observation to observation, the telescope moves only in azimuth. During a single observation, the telescope remains fixed in azimuth and objects are tracked by moving a spherical aberration corrector. Since the primary mirror remains at a fixed angle with respect to gravity, significant cost savings were realized in the mirror support system. The Hobby-Eberly Telescope was designed for astronomical spectroscopy. It began operations in 1998. The Southern African Large Telescope (SALT), completed in 2005, is based on the HET design.

Worldwide efforts are under way on a new generation of large, ground-based telescopes, using both the spin-casting method and the segmented method to produce large mirrors. The Gemini project of the National Optical Astronomy Observatory has built twin 8.1-m (319-in.) telescopes, one on Mauna Kea, Hawaii, and the other on Cerro Pachon in Chile. The Gemini North telescope (now called Gillett; **Fig. 7**) was dedicated in June 1999. Gemini South, in the high dry mountains of central Chile, went into operation in 2000. Each telescope is on an alt-azimuth mounting, requiring precise computer control of altitude, azimuth, and image rotation in order to track astronomical sources.

The Very Large Telescope (VLT), operated by the European Southern Observatory on Cerro Paranal, Chile, consists of four 8-m (315-in.) "unit" telescopes with spin-cast mirrors and several 1.8-m (71-in.) auxiliary telescopes. The light from the telescopes is combined to give the equivalent light-gathering power of a 16-m (630-in.) telescope. The last of the four unit telescopes began collecting scientific data in September 2000. In 2001, the telescopes of the VLT system were successfully used together as the VLT Interferometer (VLTI). This system has successfully measured the angular diameters of several bright stars. Cerro Paranal is in the Atacama Desert, where it almost never rains (the average precipitation is less than a centimeter per year). The clear, dry skies are ideal for infrared observations.

Very close to the Keck telescopes on the summit of Mauna Kea is the Subaru Telsecope, an 8.2-m. (319-in.) optical-infrared telescope that achieved first light in 1999 (**Fig. 8**). The Subaru Telescope is the main telescope of the National Astronomical Observatory of Japan. Its main mirror is a single 22.8-ton piece of low-expansion glass. The observatory dome of the Subaru was designed to minimize local atmospheric turbulence. The 8.2-m mirror is supported by an active system, consisting of 261 actuators, that maintains its surface accuracy with high precision.

Fig. 6. Keck Telescope, the world's largest optical-infrared telescope. The primary mirror, 10 m (386 in.) in diameter, is viewed through the W. M. Keck Observatory's open aperture. (*R. Ressmeyer-Starlight; California Association for Research in Astronomy*)

The Large Binocular Telescope (LBT), built by an international collaboration, has two 8.4-m (331-in.) telescopes on a common binocular mount. The pair has the light-gathering power of an 11.8-m (465-in.) telescope and the resolution of a 22.8-m (900-in.) instrument. The Large Binocular Telescope, located on Mount Graham in Arizona, achieved first light in October 2005.

Space-based telescopes. A number of space-based telescopes have been highly successful. The *International Ultraviolet Explorer* (*IUE*), launched into geosynchronous orbit in 1978, is not a giant telescope; its mirror is only 45 cm (18 in.) in diameter. The satellite, positioned over the Atlantic Ocean, was operated in real time for almost 19 years to obtain high-dispersion ultraviolet spectra of solar system objects, stars, nebulae, and extragalactic objects.

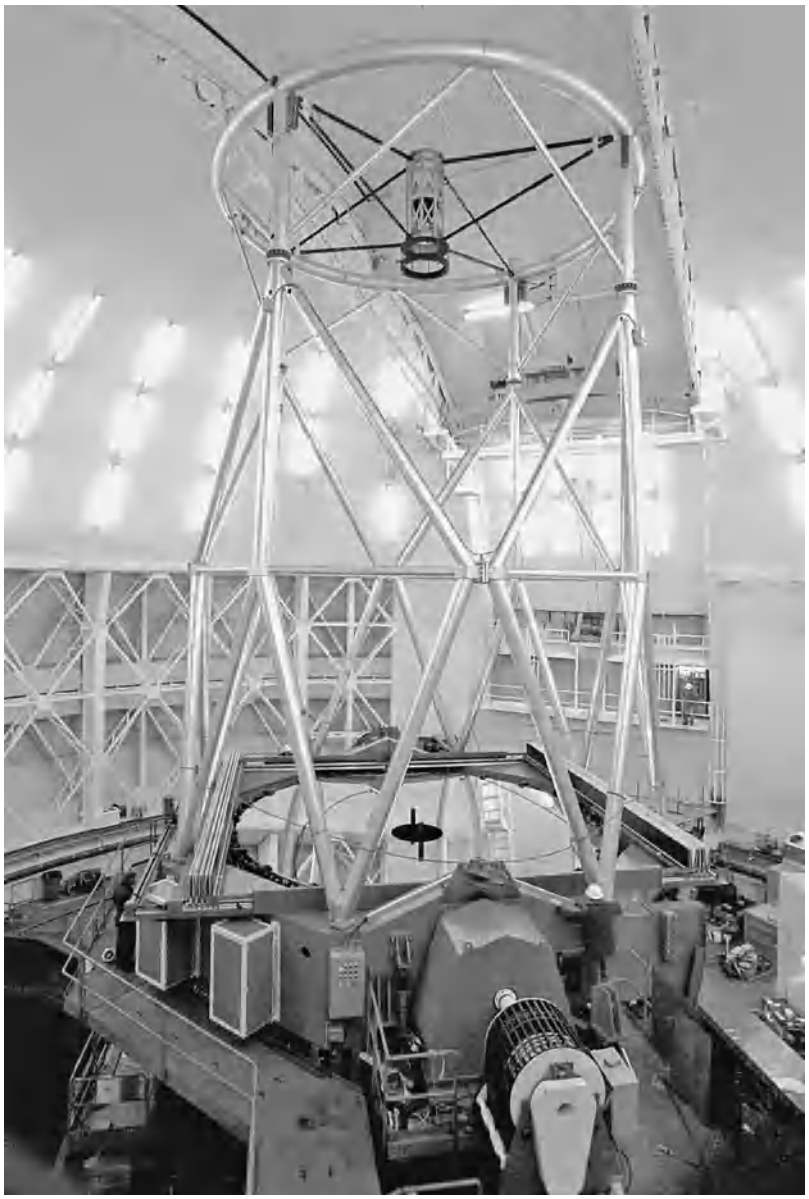

Fig. 7. Gemini North (Gillett) telescope in its dome. (*Copyright Gemini Observatory/AURA/NOAO/NSF, all rights reserved*)

The ability of large telescopes to resolve fine detail is limited by a number of factors. The ultimate limitation set by diffraction of light is rarely or never achieved by large telescopes. Instead, the finest resolved detail is limited by atmospheric inhomogeneities which set a limit of about 0.2 arcsecond. Distortion due to the mirror's own weight causes additional problems in astronomical telescopes. The Earth-orbiting *Hubble Space Telescope* (*HST*), with an aperture of 2.4 m (94 in.), was designed to eliminate these problems. The telescope operates in ultraviolet as well as visible light, resulting in a great improvement in resolution not only by the elimination of the aforementioned terrestrial effects but by the reduced blurring by diffraction in the ultraviolet. *See* DIFFRACTION; RESOLVING POWER (OPTICS).

The *Hubble Space Telescope* was launched into Earth orbit by the space shuttle in 1990. It has a Ritchie-Chrétien reflector which feeds light into a number of scientific instruments. Soon after the telescope was launched, it was discovered that the optical system was plagued with spherical aberration, which severely limited its spatial resolution. In December 1993, space-shuttle astronauts serviced and repaired the telescope, adding what amounts to eyeglasses for the scientific instruments. Since that first servicing mission, the *Hubble Space Telescope* has exceeded its prelaunch specifications for spatial resolution. A second servicing mission occurred in 1997, in which two new scientific instruments replaced first-generation instruments. One of the new instruments, the Near Infrared Camera Multi-Object Spectrography, extended the *Hubble Space Telescope's* capabilities into the infrared. A third servicing mission was flown in 1999 and a forrth in 2002. *See* HUBBLE SPACE TELESCOPE.

The *Einstein Observatory* (*High Energy Astronomy Observatory 2*), launched in 1978, and *ROSAT*, launched in 1990, carried telescopes that imaged soft x-rays. The *Chandra X-Ray Observatory* (one of NASA's Great Observatories) was launched in 1999. This space observatory (a follow-on to the *Einstein Observatory*) was named after the Indian-American scientist Subrahmanyan Chandrasekhar. The *Chandra X-Ray Observatory* is provides significantly better resolution than earlier x-ray

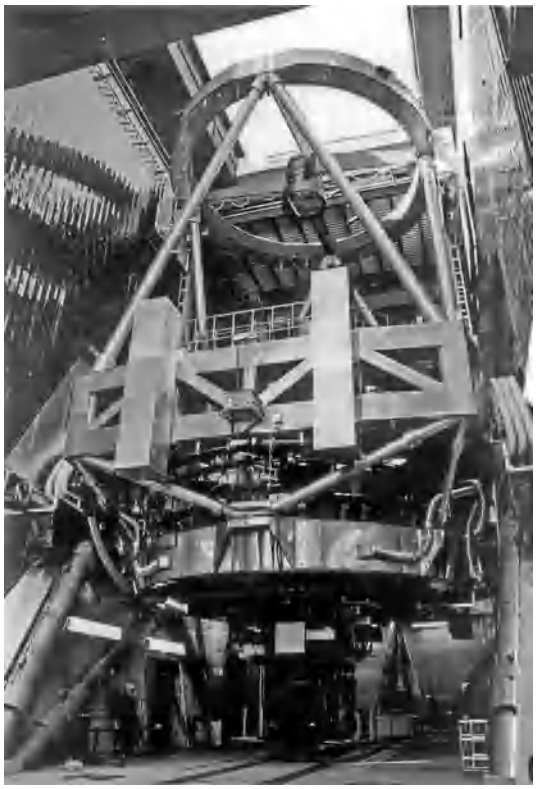

Fig. 8. Subaru Telescope of the National Astronomical Observatory of Japan (NAOJ). (© *The Subaru Telescope, National Astronomical Observatory of Japan*)

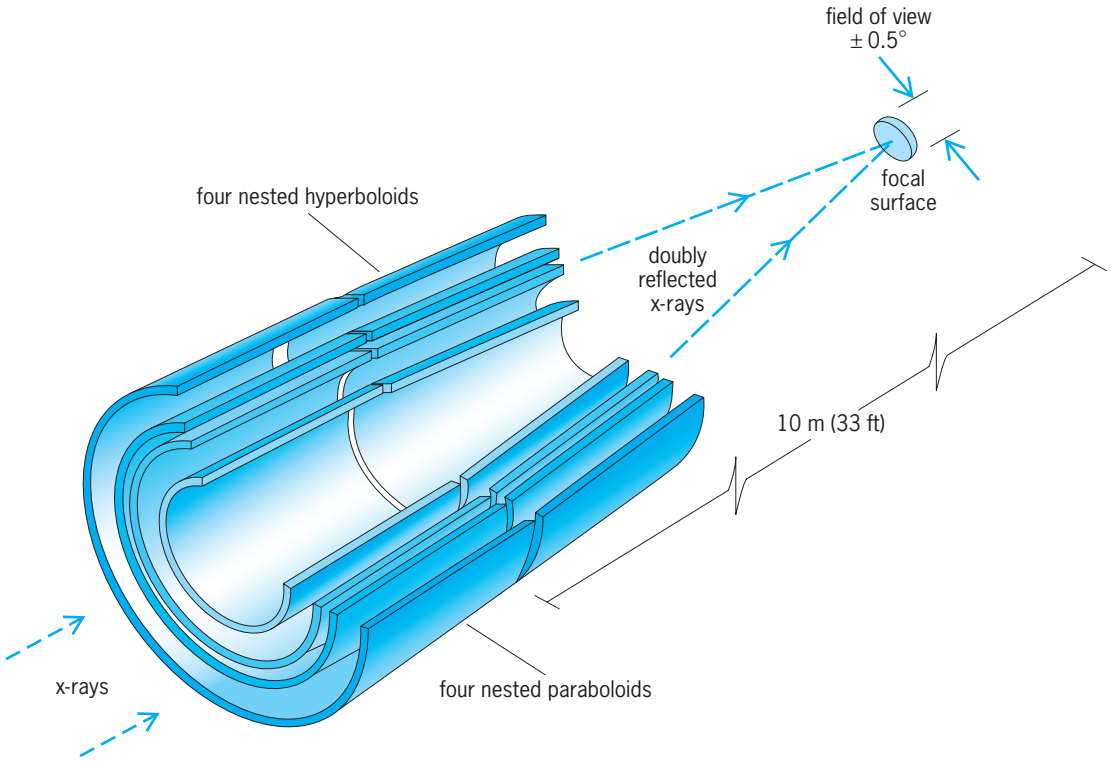

Fig. 9. The four nested Wolter Type I telescopes that are the x-ray focusing optics of the *Chandra X-Ray Observatory*. The mirror elements are 0.8 m (2.4 ft) long and 0.6–1.2 m (2.4 ft) in diameter. (*Raytheon Optical Systems, TRW Inc., and Chandra X-Ray Center, Smithsonian Astrophysical Observatory*)

observatories. The optics of the *Chandra X-Ray Observatory* consist of four nested, barrel-shaped telescopes called Wolter Type I telescopes (**Fig. 9**). These telescopes operate such that x-rays will be reflected if they strike a mirror at grazing angles, while the rays would simply pass through the mirror if they struck it at nearly right angles. The four nested telescopes provide a large cross section to collect adequate quantities of x-rays, and provide about 0.5-arcsecond resolution. The telescopes are made of Zerodur glass, coated with iridium. *See* CHANDRA X-RAY OBSERVATORY; X-RAY ASTRONOMY.

The *James Webb Space Telescope* (JWST) will have a deployable 6.5-m-diameter (252-in.) mirror, and will be placed in an orbit approximately 1.5 km (0.93 mi) from the Earth. It will be optimized to study infrared radiation. Current plans are for a launch in the 2013 time frame. *See* SATELLITE (ASTRONOMY); SCIENTIFIC SATELLITES.

Robert D. Chapman; William M. Sinton

Bibliography. Keck VLT interferometers break new ground, *Sky Telesc.*, 106(4):24, October 2003; P. R. Lawson, Optical interferometry comes of age, *Sky Telesc.*, 105(5):30–39, May 2003; L. Mertz, *Excursions in Astronomical Optics*, Springer-Verlag, New York, 1996; P. Moore, *Eyes on the Universe: The Story of the Telescope*, Springer-Verlag, London, 1997; H. S. Stockman, *Next Generation Space Telescope: Visiting a Time When Galaxies Were Young*, AURA, 1997; Subaru sees first light, *Sky Telesc.*, 97(5):18, May 1999.

Telestacea

An order of the cnidarian subclass Alcyonaria (Octocorallia). Telestacea are typified by *Telesto*, which forms an erect branching colony by lateral budding from the body wall of an elongated primary or axial polyp. The stolon is bandlike or membranous. Sclerites are scattered singly, partly fused, or entirely fused to form a rigid tube. *See* OCTOCORALLIA (ALCYONARIA).

Kenji Atoda

Teletypewriter

An electronic-mechanical device, also called a teleprinter, for preparing, storing, transmitting, and receiving messages over a telegraph or data communication circuit. The term teletypewriter was used first for teleprinters utilized in teletypewriter exchange (TWX) service in the early 1930s. It was later applied to terminals used in other services, and became interchangeable with the terms teleprinter and printing telegraph terminal. A generic term, it should be used instead of Teletype, which is a registered trademark of the AT&T Teletype Corporation. Teletypewriters were invented by several individuals in the early 1900s, initially being adaptations of the commercial typewriters of the day. Relays, magnets, and solenoids were added to provide telegraph signal generation from the operation of the typewriter keyboards, and caused the typebaskets

or typewheels to print automatically on sheets of paper.

When greater ruggedness was required, the inventors devised more sturdy mechanisms. Unreliable machinery and open-wire telegraph circuits caused many errors to occur, necessitating repetition of an errored message to get good printed copies. The use of punched tape helped somewhat, with operators fixing their perceived keyboarding errors before transmission, and expedited the repetition process by allowing reruns of the punched tape. The next generation of teletypewriters printed on narrow paper tape that was gummed on the back. Each line of a message was moistened from a sponge or wick, cut off, and pasted onto a telegram form, with errors being corrected by pasteovers (a repetition of the error portion of the message, in correct form). Although this was labor-intensive, with low wage scales the tape printers were commercially successful. Later, the increasing cost of labor led to the return to page printers exclusively. *See* TELEGRAPHY.

Commercially obsolete machines can be used by deaf persons who cannot use regular telephones. Several thousand teletypewriters have been equipped with modems or acoustic couplers for attachment to telephone handsets or lines. A visible alarm lamp is provided to signal the deaf person whenever a call is being received. Many of the sets are equipped with automatic answering facilities, including an answer-back message transmitter that identifies the terminal, so that messages can be received in the owner's absence. If a regular telephone happens to connect to one of these terminals, the person making the call will hear a whistle as the station answers. An improved instrument for deaf people is a portable teletypewriter with a built-in modem which can couple acoustically to any standard telephone handset, allowing use anywhere. These teletypewriters have compact keyboards, and some are equipped with a display that shows one line of characters, allowing conversational visible communication. Others may have a very compact printing mechanism that uses a narrow roll of paper, similar to that used in some calculators; other designs feature a small CRT display. Ransom D. Slayton

Bibliography. *Teletypewriter Circuits and Equipment: Fundamentals*, 1991.

Television

The electrical transmission and reception of transient visual images. Like motion pictures, television consists of a series of successive images which are registered in the brain as a continuous picture because of the persistence of vision. Each visual image impressed on the eye persists for a fraction of a second. At the television transmitter, minute portions of a scene are sampled individually for brightness and color, and the information for each portion is transmitted consecutively. At the receiver, each portion is synchronized and reproduced in its proper position and with correct brightness and color to reproduce the original scene.

Fundamental Principles

The technology of television is based on the conversion of light rays from still or moving scenes and pictures into electronic signals for transmission or storage, and subsequent reconversion into visual images on a screen. A similar function is provided in the production of motion picture film. However, whereas film records the brightness variations of a complete scene on a single frame in an exposure no longer than a fraction of a second, the elements of a television picture must be scanned one piece at a time. In the television system, a scene is dissected into a frame composed of a mosaic of picture elements or pixels. (A pixel is defined as the smallest area of a television image that can be transmitted within the parameters of the system.) This process is accomplished by (1) analyzing the image with a photoelectric device in a sequence of horizontal scans from the top to the bottom of the image to produce an electric signal in which the brightness and color values of the individual picture elements are represented as voltage levels of a video waveform; (2) transmitting the values of the picture elements in sequence as voltage levels of a video signal; and (3) reproducing the image of the original scene in a video signal display of parallel scanning lines on a viewing screen.

Scanning lines and fields. The image pattern of electrical charges on a camera pickup element (corresponding to the brightness levels of a scene) are converted to a video signal in a sequential order of picture elements in the scanning process. At the end of each horizontal line sweep, the video signal is blanked while the beam returns rapidly to the left side of the scene to start scanning the next line. This process continues until the image has been scanned from top to bottom to complete one field scan. *See* TELEVISION CAMERA.

After completion of this first field scan, at the midpoint of the last line, the beam again is blanked as it returns to the top center of the target where the process is repeated to provide a second field scan. The spot size of the beam as it impinges upon the target must be sufficiently fine to leave unscanned areas between lines for the second scan. The pattern of scanning lines covering the area of the target, or the screen of a picture display, is called a raster.

Because of the half-line offset for the start of the beam return to the top of the raster and for the start of the second field, the lines of the second field lie in between the lines of the first field. Thus, the lines of the two are interlaced. The two interlaced fields constitute a single television frame. **Figure 1** shows a frame scan with interlacing of the lines of two fields.

Reproduction of the camera image on a cathode-ray tube (CRT) is accomplished by an identical operation, with the scanning beam modulated in density by the video signal applied to an element of the electron gun. This control voltage to the CRT varies the brightness of each picture element on the phosphor

screen. Alternatively, the image may be displayed on a flat-panel display device. *See* CATHODE-RAY TUBE; FLAT-PANEL DISPLAY DEVICE; PICTURE TUBE.

Blanking of the scanning beam during the return trace (retrace) is provided for in the video signal by a "blacker-than-black" pulse waveform. In addition, in most receivers and monitors another blanking pulse is generated from the horizontal and vertical scanning circuits and applied to the CRT electron gun to ensure a black screen during scanning retrace.

The interlaced scanning format, standardized for monochrome and compatible color, was chosen primarily for two partially related and equally important reasons: (1) to eliminate viewer perception of the intermittent presentation of images, known as flicker, and (2) to reduce the video bandwidth requirements for an acceptable flicker threshold level.

The standard adopted by the Federal Communications Commission (FCC) for monochrome television in the United States specified a system of 525 lines per frame, transmitted at a frame rate of 30 Hz, with each frame composed of two interlaced fields of horizontal lines. Initially in the development of television transmission standards, the 60-Hz power line waveform was chosen as a convenient reference for vertical scan. Furthermore, in the event of coupling of power line hum into the video signal or scanning-deflection circuits, the visible effects would be stationary and less objectionable than moving hum bars or distortion of horizontal-scanning geometry. In the United Kingdom and much of Europe, a 50-Hz interlaced system was chosen for many of the same reasons. With improvements in television receivers, the power line reference was replaced with a stable crystal oscillator.

The initial 525-line monochrome standard was retained for color in the recommendations of the National Television System Committee (NTSC) for compatible color television in the early 1950s. The NTSC system, adopted in 1953 by the FCC, specified a scanning system of 525 horizontal lines per frame, with each frame consisting of two interlaced fields of 262.5 lines at a field rate of 59.94 Hz. Forty-two of the 525 lines in each frame are blanked as black picture signals and reserved for transmission of the vertical scanning synchronizing signal. This results in 483 visible lines of picture information. *See* TELEVISION SCANNING.

Synchronizing video signals. In monochrome television transmission, two basic synchronizing signals are provided to control the timing of picture-scanning deflection: (1) horizontal sync pulses at the line rate, and (2) vertical sync pulses at the field rate in the form of an interval of wide horizontal sync pulses at the field rate. Included in this interval are equalizing pulses at twice the line rate to preserve interlace in each frame between the even and odd fields (offset by one-half line).

In color transmissions, a third synchronizing signal is added during horizontal scan blanking to provide a frequency and phase reference for signal encoding circuits in cameras and decoding circuits in receivers. These synchronizing and reference signals are combined with the picture video signal to form a composite video waveform (**Fig. 2**).

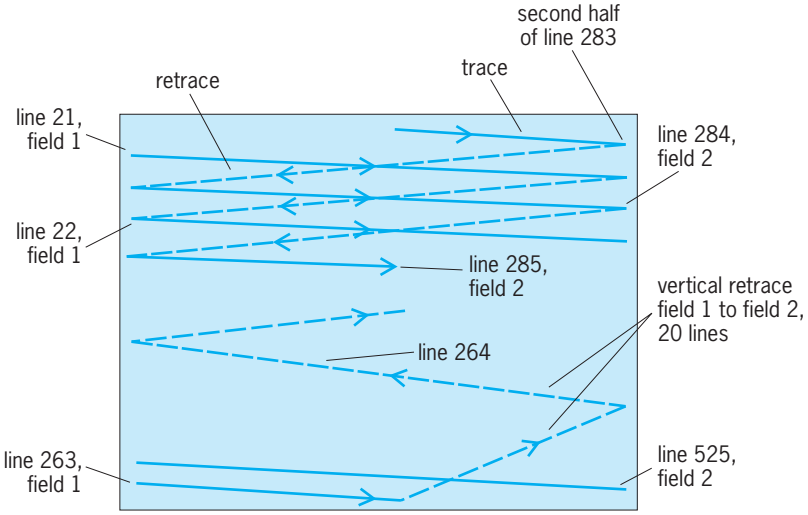

Fig. 1. Interlace scanning pattern (raster) of the television image.

The receiver scanning and color-decoding circuits must follow the frequency and phase of the synchronizing signals to produce a stable and geometrically accurate image of the proper color hue and saturation. Any change in timing of successive vertical scans can impair the interlace of even and odd fields in a frame. Small errors in horizontal scan timing of lines in a field can result in a loss of resolution in vertical line structures. Periodic errors over several lines that may be out of the range of the horizontal-scan automatic frequency control circuit in the receiver will be evident as jagged vertical lines.

Industry standards. Three primary color transmission standards are in use today:

1. NTSC (National Television Systems Committee)—used in the United States, Canada, Mexico, Central America, some of South America, and Japan. In addition, NTSC is used in various countries or possessions heavily influenced by the United States.

2. PAL (Phase Alternate each Line)—used in the United Kingdom, most countries and possessions influenced by England, most European countries, and China. Variation exists in PAL systems.

3. SECAM (SEquential Color with [Avec] Memory)—used in France, countries and possessions influenced by France, the nations of the former Soviet Union, other former Soviet Bloc nations, including East Germany, and other areas influenced by Russia.

The three standards are incompatible for a variety of technical reasons, but standard converter devices are commercially available.

Transmission frequencies. The band of frequencies assigned to a television station for the transmission of synchronized picture and sound signals is called a television channel. In the United States a television channel is 6 MHz wide, with the visual carrier frequency 1.25 MHz above the lower edge of the band and the aural carrier 0.25 MHz below the upper edge of the band.

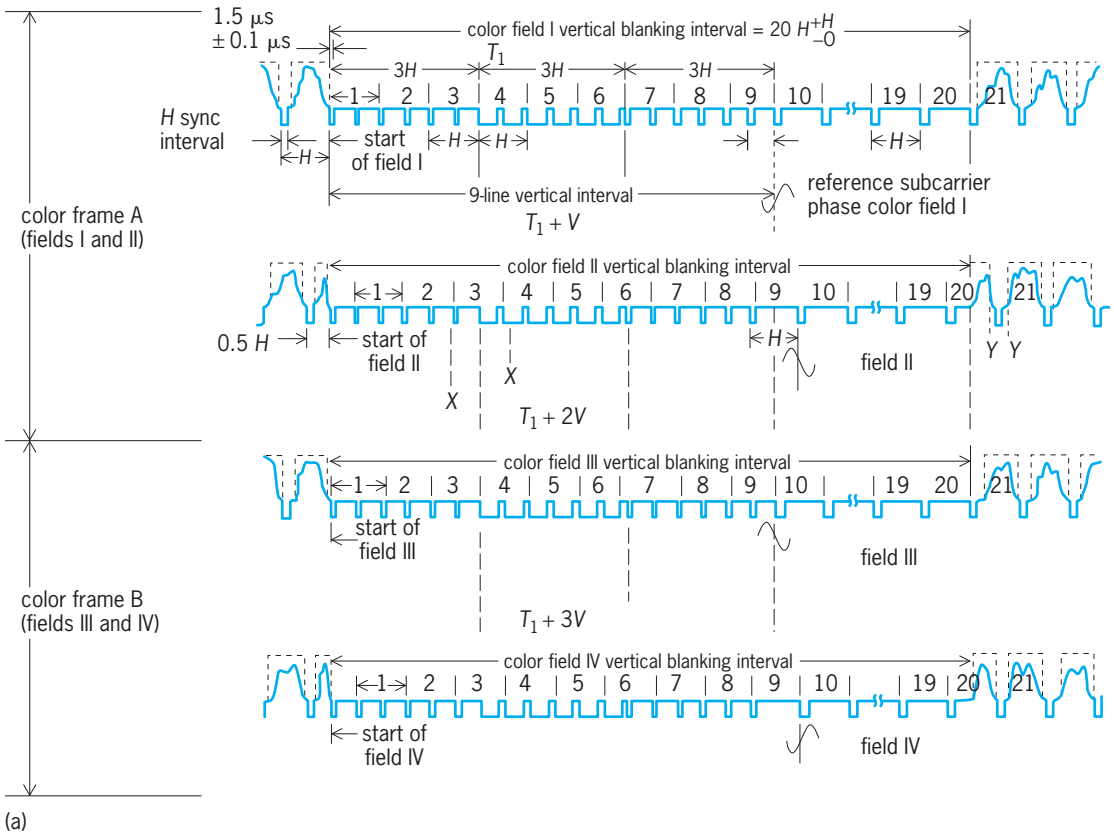

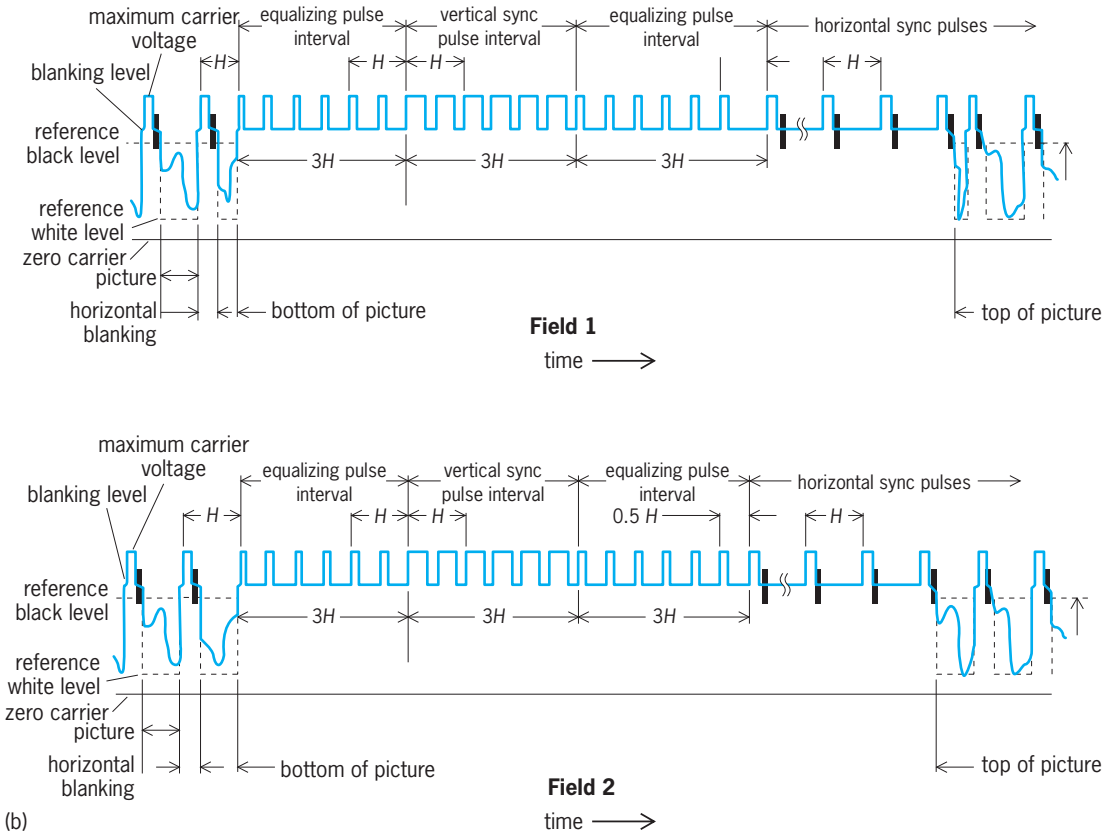

Fig. 2. NTSC color television waveform. (*a*) Principal components. (*b*) Detail of picture elements. (*Electronic Industries Association*)

TABLE 1. Television channels in the United States

Channel number	Frequency band, MHz	Channel number	Frequency band, MHz
2	54–60	36	602–608
3	60–66	37	608–614
4	66–72	38	614–620
5	76–82	39	620–626
6	82–88	40	626–632
7	174–180	41	632–638
8	180–186	42	638–644
9	186–192	43	644–650
10	192–198	44	650–656
11	198–204	45	656–662
12	204–210	46	662–668
13	210–216	47	668–674
14	470–476	48	674–680
15	476–482	49	680–686
16	482–488	50	686–692
17	488–494	51	692–698
18	494–500	52	698–704
19	500–506	53	704–710
20	506–512	54	710–716
21	512–518	55	716–722
22	518–524	56	722–728
23	524–530	57	728–734
24	530–536	58	734–740
25	536–542	59	740–746
26	542–548	60	746–752
27	548–554	61	752–758
28	554–560	62	758–764
29	560–566	63	764–770
30	566–572	64	770–776
31	572–578	65	776–782
32	578–584	66	782–788
33	584–590	67	788–794
34	590–596	68	794–800
35	596–602	69	800–806

Television channels in the United States are identified by numbers, starting with channel 2. The frequency originally assigned to channel 1 was later reassigned to other uses. These channels are in three frequency bands (**Table 1**). Channels 2–6 occupy 54 to 88 MHz, channels 7–13 are 174 to 216 MHz, and channels 14–69 are 470 to 806 MHz. The first two groups of channels fall in the very high frequency (VHF) band; the channels in the last group are in the ultrahigh-frequency (UHF) band. When digital television, whose stations operate at different radio frequencies, is widely implemented, the current radio frequencies used by the analog television stations will be used for other transmision services. *See* RADIO SPECTRUM ALLOCATIONS.

Grades of service. The FCC has established two grades of television service for the United States. Grade A service provides relatively high freedom from interference from other television stations and also good freedom from artificial and receiver noise. It specifies that picture quality acceptable to the average observer is expected to be available at least 90% of the time at the best 70% of all receiver locations at the outer geographical limits of the service area. Grade B service may be more vulnerable to interference and noise. It specifies that service equal to that of grade A is available to only 50% of all receiver locations at the service area limit.

Sound transmission. In the United States, the sound portion of the program is transmitted by frequency modulation (FM) at a carrier frequency 4.5 MHz above the picture carrier. Maximum frequency deviation (bandwidth) of the sound signals is 25 kHz. Additional sound channels can be provided using narrowband FM subcarriers located near the primary audio channel. *See* FREQUENCY MODULATION.

The normal frequency response is altered in the transmitter to emphasize the higher audio frequencies with respect to the lower frequencies. This preemphasis is accomplished by a circuit that causes the audio response to increase with frequency. A corresponding circuit is used in the receiver to produce an equal and opposite decrease of response to higher audio frequencies. By so doing, noise produced in the receiver is attenuated without the overall audio-frequency response being affected. *See* TELEVISION RECEIVER; TELEVISION STANDARDS; TELEVISION TRANSMITTER.

Digital Television

Digital television (DTV) has ushered in a new era in television broadcasting. The impact of DTV is more significant than simply moving from an analog system to a digital system. DTV permits a level of flexibility wholly unattainable with analog broadcasting. An important element of this flexibility is the ability to expand system functions by building upon the technical foundations specified in the basis standards.

With NTSC, and its PAL and SECAM counterparts, the video, audio, and some limited data information are conveyed by modulating a radio-frequency (RF) carrier in such a way that a receiver of relatively simple design can decode and reassemble the various elements of the signal to produce a program consisting of video and audio, and perhaps related data (for example, closed captioning). As such, a complete program is transmitted by the broadcaster that is essentially in finished form. In the DTV system, however, additional levels of processing are required after the receiver demodulates the RF signal. The receiver processes the digital bit stream extracted from the received signal to yield a collection of program elements (video, audio, or data) that match the services the consumer has selected. This selection is made using system and service information that is also transmitted. Audio and video are delivered in digitally compressed form and must be decoded for presentation. Audio may be monophonic, stereo, or multichannel. Data may supplement the main video-audio program (for example, closed captioning, descriptive text, or commentary), or it may be a standalone service (for example, a stock or news ticker). *See* CLOSED CAPTION TELEVISION.

The nature of the digital television system is such that it is possible to provide new features that build upon the infrastructure within the broadcast plant and the receiver. One of the major enabling developments of digital television, in fact, is the integration of significant processing power in the receiving device itself. Historically, in the design of any broadcast system—be it radio or television—the goal has always been to concentrate technical sophistication (when needed) at the transmission end and thereby

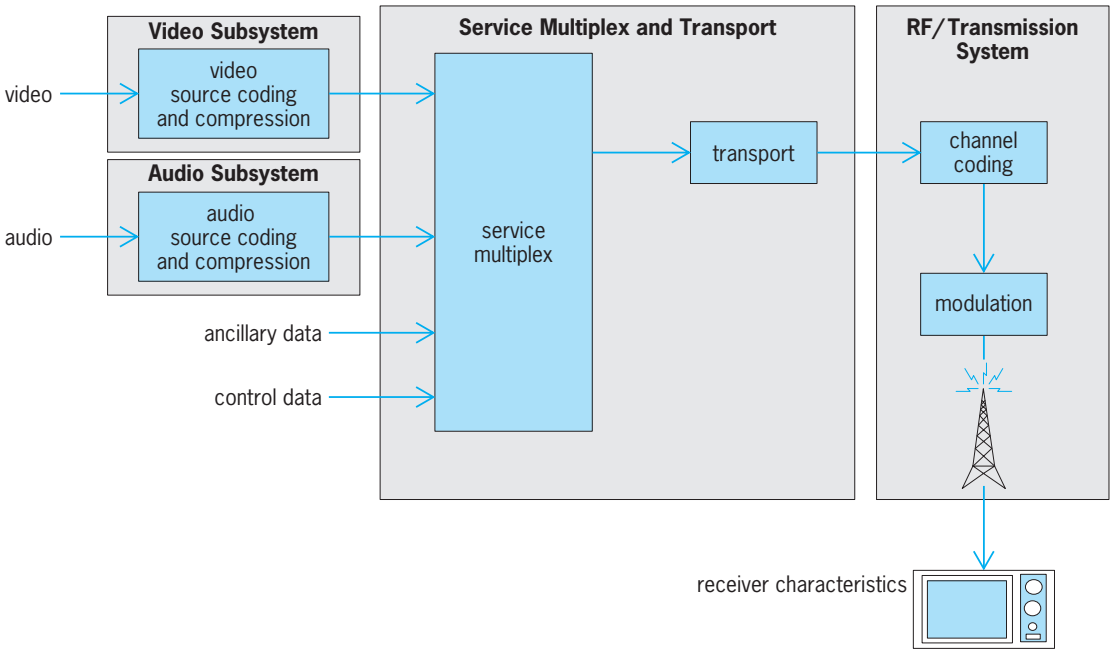

Fig. 3. ITU-R Digital terrestrial television broadcasting model of the International Telecommunications Union, Radiocommunication Sector (ITU-R). (*From Advanced Television Systems Committee, ATSC Digital Television Standard, Doc. A/53B, 2001*)

facilitate simpler receivers. Because there are far more receivers than transmitters, this approach has obvious business advantages. While this trend continues to be true, the complexity of the transmitted bit stream and compression of the audio and video components require a significant amount of processing power in the receiver, which is practical because of the enormous advancements made in computing technology. Once a receiver reaches a certain level of sophistication (and market success), additional processing power is essentially "free."

The development of digital television occurred over a period of many years on several continents. Three major systems emerged:

1. The Advanced Television Systems Committee (ATSC), a North American–based DTV standard organization, developed the ATSC terrestrial DTV series of standards.

2. The Digital Video Broadcasting Project (DVB), a European-based standard organization, developed the DVB series of DTV standards, which were standardized by the European Telecommunication Standard Institute (ETSI).

3. The Integrated Service Digital Broadcasting (ISDB) standards, a series of DTV standards, were developed and standardized by the Association of Radio Industries and Business (ARIB) in Japan.

The remainder of this article focuses on the ATSC DTV system.

ATSC DTV system overview. The Digital Television Standard describes a system designed to transmit high-quality video and audio and ancillary data over a single 6-MHz channel. The system can deliver about 19 megabits per second in a 6-MHz terrestrial broadcasting channel and about 38 Mbits/s in a 6-MHz cable television channel. This means that encoding high-definition video essence at raw data rates that are typically greater than 1 gigabit per second requires a bit rate reduction by about a factor of 50. To achieve this bit rate reduction, the system is designed to be efficient in utilizing available channel capacity by exploiting complex video and audio compression technology. The compression scheme optimizes the throughput of the transmission channel by representing the video, audio, and data sources with as few bits as possible while preserving the level of quality required for the given application. *See* DATA COMPRESSION.

A basic block diagram representation of the ATSC DTV system is shown in **Fig. 3**. According to this model, the digital television system can be seen to consist of three subsystems: source coding and compression, service multiplex and transport, and RF/transmission. The RF/transmission subsystems described in the Digital Television Standard are designed specifically for terrestrial and cable applications. The structure is such that the video, audio, and service multiplex and transport subsystems are useful in other applications.

Source coding and compression refers to the bit-rate reduction methods (data compression) appropriate for application to the video, audio, and ancillary digital data streams. (The term ancillary data encompasses control data; conditional-access (private) data; data associated with the program audio and video services, such as closed captioning; and independent program services.) The purpose of the coder is to minimize the number of bits needed to represent the audio and video information.

The modulation subsystem offers two modes: the terrestrial broadcast mode, 8-VSB (8-level vestigal sideband); and the high-data-rate mode, 16-VSB

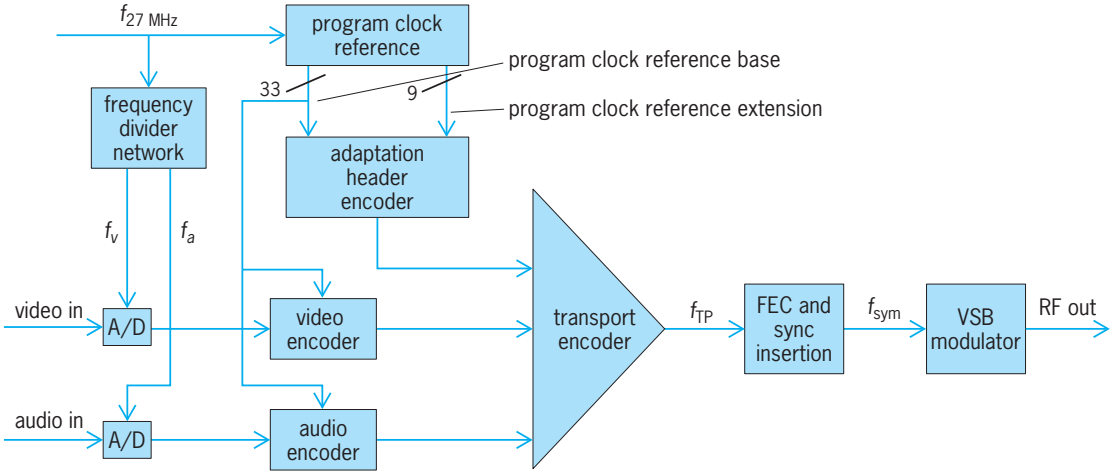

Fig. 4. High-level view of the DTV encoding system, illustrating the relationship and derivation of various clock frequencies within the encoder and the use of forward error correction (FEC). A/D = analog-to-digital. VSB = vestigial sideband. (*From Advanced Television Systems Committee, ATSC Digital Television Standard, Doc. A/53B, 2001*)

(16-level vestigial sideband). *See* AMPLITUDE MODULATION.

Figure 4 gives a high-level view of the encoding equipment. This view is not intended to be complete, but is used to illustrate the relationship and derivation of various clock frequencies within the encoder and the use of forward error correction (FEC) in the system.

The DTV video-compression algorithm conforms to the Main Profile syntax of the standard ISO/IEC 13818-2 (MPEG-2). The allowable parameters are bounded by the upper limits specified for the Main Profile/High Level in this standard. **Table 2** lists the allowed compression formats under the ATSC DTV Standard.

ATSC system block diagram. A basic block diagram representation of the ATSC DTV system is shown in **Fig. 5**. According to this model, the digital television system consists of four major elements, three within the broadcast plant plus the receiver.

Application encoders and decoders. The application encoders and application decoders, as used in Fig. 5, apply bit-rate reduction methods, also known as data compression, appropriate for application to the video, audio, and ancillary digital data streams. The purpose of compression is to minimize the number of bits needed to represent the audio and video information. The DTV system employs the MPEG-2 video stream syntax for the coding of video and the ATSC standard Digital Audio Compression (AC-3) for the coding of audio.

As noted above, ancillary data is a broad term that includes control data and data associated with the program audio and video services. As standards were developed to define how to transport and process data, it became clear that different forms of data served very different purposes, and different standards were needed for metadata and essence. (essence = content − metadata). In this context, (video) essence is the image itself without any ancillary data, and metadata is information about the data essence. Including system information as ancillary data is not strictly proper, as some such data are needed to reassemble the audio, video, and data services. Data delivered as a separate payload can provide independent services as well as data elements related to an audio- or video-based service.

Transport packetization and multiplexing. Transport packetization and multiplexing refers to the means of dividing each bit stream into "packets" of information, the means of uniquely identifying each packet or packet type, and the appropriate methods of

TABLE 2. ATSC DTV compression format constraints

Vertical size value	Horizontal size value	Aspect ratio	Frame-rate code*	Scanning sequence
1080†	1920	16:9, square pixels	1,2,4,5	Progressive
			4,5	Interlaced
720	1280	16:9, square pixels	1,2,4,5,7,8	Progressive
480	704	4:3, 16:9	1,2,4,5,7,8	Progressive
			4,5	Interlaced
	640	4:3, square pixels	1,2,4,5,7,8	Progressive
			4,5	Interlaced

*Frame-rate code: 1 = 23.976 Hz, 2 = 24 Hz, 4 = 29.97 Hz, 5 = 30 Hz, 7 = 59.94 Hz, 8 = 60 Hz.
†1088 lines actually are coded in order to satisfy the MPEG-2 requirement that the coded vertical size be a multiple of 16 (progressive scan) or 32 (interlaced scan).
source: After Advanced Television Systems Committee, *ATSC Digital Television Standard*, Doc. A/53B, 2001.

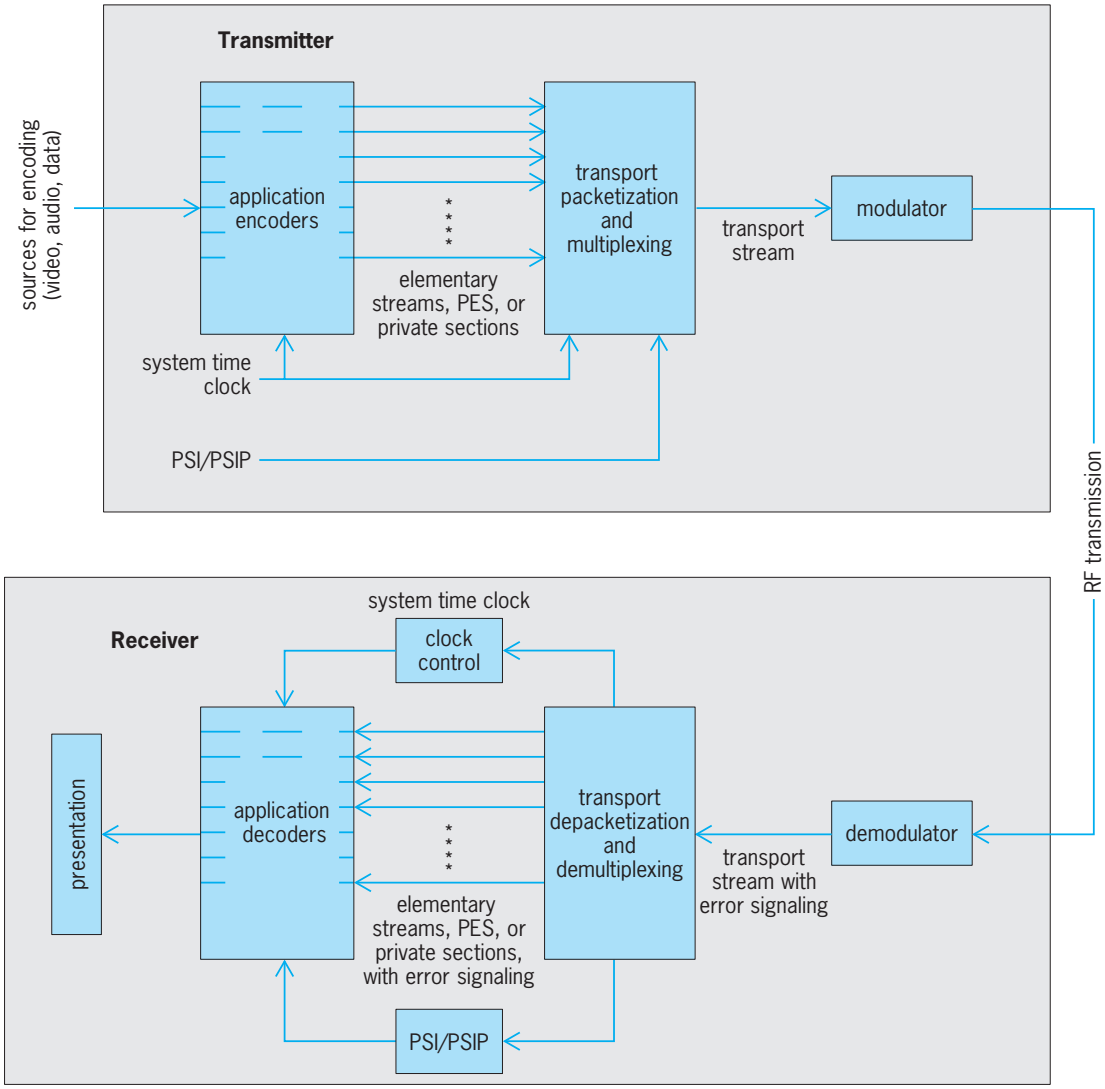

Fig. 5. Sample organization of functionality in a transmitter-receiver pair for a single DTV program. PSIP = Progam and System Information Protocol, PES = packetized elementary stream, PSI = program-specific information. (*From Advanced Television Systems Committee, Guide to the Use of the Digital Television Standard, Doc. A/54A, 2003*)

interleaving or multiplexing video bit stream packets, audio bit stream packets, and data bit stream packets into a single transport mechanism. (The inverse operations of transport depacketization and demultiplexing are carried out in the receiver.) The structure and relationships of these essence bit streams are carried in service information bit streams, also multiplexed into the single transport mechanism. In developing the transport mechanism, interoperability among digital media—such as terrestrial broadcasting, cable distribution, satellite distribution, recording media, and computer interfaces—was a prime consideration. The DTV system employs the MPEG-2 transport stream syntax for the packetization and multiplexing of video, audio, and data signals for digital broadcasting systems. The MPEG-2 transport stream syntax was developed for applications where channel bandwidth or recording media capacity is limited and the requirement for an efficient transport mechanism is paramount.

RF transmission. RF transmission refers to channel coding and modulation. The channel coder takes the digital bit stream and adds information that can be used by the receiver to reconstruct the data from the received signal which, due to transmission impairments, may not accurately represent the actual transmitted signal. The modulation (or physical layer) uses the digital bit stream information to modulate a carrier for the transmitted signal. As noted above, the modulation subsystem offers two modes: 8-VSB and 16-VSB.

Receiver. The ATSC receiver must recover the bits representing the original video, audio, and other data from the modulated signal. In particular, the receiver must tune to the selected 6-MHz channel; reject adjacent channels and other sources of interference; demodulate (equalize as necessary) the received signal, applying error correction to produce a transport bit stream; identify the elements of the bit stream using a transport layer processor; select each desired element and send it to its appropriate processor; decode and synchronize each element; and present the programming. *See* DEMODULATION.

Noise, interference, and multipath are elements of the terrestrial transmission path, and the receiver circuits are expected to deal with these impairments.

Innovations in equalization, automatic gain control, interference cancellation, and carrier and timing recovery create product performance differentiation and improve signal coverage. *See* AUTOMATIC GAIN CONTROL (AGC); EQUALIZER.

High-definition television. The resolution of the displayed picture is the most basic attribute of any video production system. Generally speaking, a high-definition television (HDTV) image has approximately twice as much luminance definition horizontally and vertically as the 525-line NTSC system or the 625-line PAL and SECAM systems. The total number of luminance picture elements (pixels) in the image, therefore, is four times as great. The wider aspect ratio of the HDTV system adds even more visual information. The HDTV image is 25% wider than the conventional video image for a given image height; the ratio of image width to height in HDTV systems is 16:9, or 1.78. The conventional video image has a 4:3 aspect ratio.

As a result of these attributes, the HDTV image may be viewed more closely than is customary in conventional television systems. Full visual resolution of the detail of conventional television is available when the image is viewed at a distance equal to about six or seven times the height of the display. The HDTV image may be viewed from a distance of about three times picture height for the full detail of the scene to be resolved.

Definition. Although there is no single, universal definition for HDTV, it is generally accepted to encompass several elements that are described by various consumer, broadcast, and regulatory groups. HDTV offers the potential for approximately twice the horizontal and twice the vertical resolution of current (NTSC) television. When combined with a wide-screen format (16:9 aspect ratio), this can result in considerably more visual information than conventional television. HDTV consumer television sets can have 720 or 1080 active vertical scanning lines, and are capable of decoding the transmitted 720 × 1280 and 1080 × 1920 ATSC formats and displaying them as a 16:9 aspect ratio image. These two high-definition formats can potentially provide over eight times as much picture information as delivered over broadcast NTSC. While these high-definition transmission formats will be supported by such sets, the actual delivered resolution may vary by broadcaster, by product, and by program. HDTV is normally accompanied by digital surround-sound capability.

Production versus transmission systems. Bandwidth is perhaps the most basic factor that separates production HDTV systems from transmission-oriented systems for broadcasting. A closed-circuit system does not suffer the same restraints imposed upon a video image that must be transported by radio-frequency means from an origination center to consumers. This distinction has led to the development of widely varied systems for production and transmission applications. *See* CLOSED-CIRCUIT TELEVISION.

Format development. In the staging of motion picture films intended for theatrical distribution, no provision generally is made for the limitations of conventional video displays. Instead, the full screen, in formats such as CinemaScope with wide aspect ratios, is used by directors for maximum dramatic and sensory impact. Consequently, cropping of essential information often may be encountered on the video screen. This problem is particularly acute in wide-screen features where cropping of the sides of the film frame is necessary to produce a print for video transmission. This objective is met in one of the following ways:

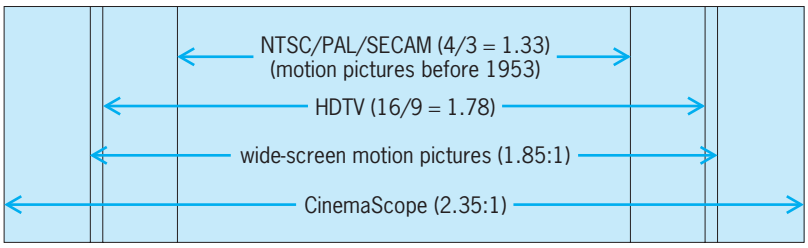

Fig. 6. Common image aspect ratio formats. (*After K. B. Benson and D. G. Fink: HDTV: Advanced Television for the 1990s, McGraw-Hill, New York, 1990*)

1. Letterbox transmission, with blank areas above and below the wide-screen frame. Audiences in North America and Japan have not generally accepted this presentation format, primarily because of the reduced size of the picture images and the distraction of the blank screen areas.

2. Printing the full frame height and cropping equal portions of the left and right sides to provide a 4:3 aspect ratio. This process frequently is not ideal because, depending upon the scene, important visual elements may be eliminated.

3. Programming the horizontal placement of a 4:3 aperture to follow the essential picture information. Called pan and scan, this process is used in producing a print or in making a film-to-tape transfer for video viewing. Editorial judgment is required for determining the scanning cues for horizontal positioning and, if panning is used, the rate of horizontal movement. This is an expensive and laborious procedure and, at best, it compromises the artistic judgments made by the director and the cinematographer in staging and shooting and by the film editor in postproduction.

One reason for moving to a 16:9 format is to take advantage of consumer acceptance of the 16:9 aspect ratio commonly found in motion picture films. Actually, motion pictures are produced in several formats, including 4:3 (1.33); 2.35, used for 35mm anamorphic CinemaScope film; and 2.2 in a 70mm format Still, the 16:9 aspect ratio generally is supported by the motion picture industry. **Figure 6** illustrates some common aspect ratios. *See* CINEMATOGRAPHY.

Jerry C. Whitaker

Bibliography. J. Boston, *DTV Survival Guide*, McGraw-Hill, New York, 2000; M. Robin and M. Poulin, *Digital Television Fundamentals*, 2d ed., McGraw-Hill, New York, 2000; P. Symes, *Video Compression Demystified*, McGraw-Hill, New York, 2001; J. C. Whitaker (ed.), *NAB Engineering Handbook*, 9th ed., National Association of Broadcasters, Washington, DC, 1997; J. C. Whitaker, *Standard Handbook of Broadcast Engineering*, McGraw-Hill, New York, 2005; J. C. Whitaker, *Standard Handbook of Video and Television Engineering*, 4th ed., McGraw-Hill, New York, 2003.

Television camera

An electrooptical system used to pick up and convert a visual image or scene into an electrical signal called video. The video may be transmitted by cable or wireless means to a suitable receiver or monitor some distance from the actual scene. It may also be recorded on a video tape recorder for playback at a later time.

A television camera may fall within one of several categories: studio (**Fig. 1**), portable, or telecine. It may also be one of several highly specialized cameras used for remote viewing of inaccessible places, such as the ocean bottom or the interior of nuclear power reactors. The camera may be capable of producing color or monochrome (black and white) pictures. Most modern cameras are entirely solid-state, including the light-sensitive element, which is composed of semiconductors called charge-coupled devices (CCDs). Inexpensive or special-purpose cameras, however, may use one or more vacuum tubes, called vidicons, with a light-sensitive surface in lieu of the charge-coupled devices. *See* CHARGE-COUPLED DEVICES.

Television cameras intended for industrial, consumer, or broadcast portable use are usually one piece, with all elements of the camera system contained in one assembly. Such cameras may be combined with a detachable or built-in videocassette recorder to form a camera recorder or camcorder (**Fig. 2**). A broadcast-quality studio camera, on the other hand, usually consists of a separate head and camera control unit (CCU) connected by a multiple-conductor cable.

In such a two-piece camera, the head unit consists of the optical system with lens, the picture pickup devices, and a minimum of electronics necessary to generate and amplify the signals from the pickup devices. It will also have a small, built-in television screen that functions as a viewfinder. The camera control unit may contain electronics and controls that allow a skilled operator to adjust brightness (luminance), color (balance, saturation, and hue), and, for cameras that employ pickup tubes, certain correction circuits (registration, gamma, and aperture) that improve the picture. A special-purpose oscilloscope, called a waveform monitor, is generally provided to the operator of the camera control unit so that accurate voltage levels may be set. Modern cameras provide for automatic as well as manual setup of some or all of the above adjustments. Some cameras have a built-in microprocessor that enables complete setup of all parameters by simply pushing a button. Triax cameras utilize a small-diameter triaxial cable (three concentric conductors) to connect the camera head to the camera control unit, instead of heavier, multiple-conductor cable with as many as 81 conductors. Triax can be used because all of the normal signals are multiplexed onto the three wires. The one-piece cameras generally incorporate similar functions with somewhat lower levels of performance and flexibility, owing to their smaller size. *See* MICROPROCESSOR.

Fig. 1. Television studio camera. (*Thomson Multimedia*)

Fig. 2. Televison camcorder. (*Thomson Multimedia*)

Essential elements. Every camera shares certain essential elements: an optical system, one or more picture pickup devices, preamplifiers, scanning circuits, blanking and synchronizing circuits, video processing circuits, and control circuits. Color cameras also include some kind of color-encoding circuit.

Optical system. The optical system consists minimally of a fixed-focal-length lens placed directly in front of a pickup device. Provision must be made to focus the image on the focal plane of the pickup device. Usually the pickup device, complete with its mounting and scanning components, is adjusted with respect to the lens, while the lens is set at the distant extreme of its focusing range. This establishes the correct back-focus so that a sharp image can be obtained by adjusting only the lens-focusing ring as camera-to-subject distances vary.

All but the most rudimentary cameras replace the fixed-focal-length lens with a zoom-type lens that allows a smooth, continuous transition from wide-angle to telephoto focal lengths while the image remains focused. Modern consumer cameras and camcorders almost always provide an autofocus feature, based on an infrared rangefinder circuit, so that subjects at various distances from the camera will

remain in focus without operator intervention. Some advanced professional cameras, and many consumer-grade cameras, offer an image stabilization feature that makes steady pictures attainable even if the camera cannot be held steady. Most cameras provide for the insertion of various filters into the optical path in order to correct for the color temperature of the lighting or to create special optical effects. *See* COLOR FILTER; GYROSCOPE; RANGEFINDER (OPTICS).

Some portable cameras, especially those intended to capture fast-moving subjects, are equipped with either mechanical or electronic shutters, which can be used to increase the effective frame rate. This significantly reduces the blurring of the image that would occur at the normal video vertical field rate [59.94 Hz for the NTSC (National Television Systems Committee) color system] when, for example, the camera is shooting a fast-moving object. Shutters also cause a loss of sensitivity, which increases with the shutter speed.

Color cameras must split the incoming light into suitable primary colors (usually red, green, and blue). On very inexpensive consumer cameras, this is generally done by means of a stripe filter (with thin vertical, alternating stripes of red, green, and blue) placed in front of a single pickup device. Somewhat better performance is achieved in industrial-grade cameras by using two or more pickup devices with a relay lens and dichroic mirror system placed between the objective lens and the pickup devices (**Fig. 3***a*). Dichroic mirrors reflect light of one color while passing all others; that is, a red dichroic mirror reflects red and passes all other colors. The highest-performance (and most expensive) broadcast-quality cameras utilize three pickup devices and a glass prism with dichroic materials coated onto the glass (Fig. 3*b*).

The advantages of the costlier prism system include ruggedness and, most importantly, the elimination of air-to-glass surfaces except at the entry and exit points. This reduces image deterioration due to misalignment of the mirrors and dust accumulation on the surfaces. The prism also makes a good mount for charge-coupled-device pickup devices, which can be aligned and permanently bonded to the prism by the manufacturer. *See* DICHROISM.

Pickup device. The picture pickup device used in most modern cameras is a type of highly integrated solid-state circuit called a charge-coupled device. Charge-coupled devices consist of a large number of photodiodes (light-sensitive semiconductor junctions) aligned in a matrix so that each diode's output voltage can be related to a particular point in the picture that is focused on the array (**Fig. 4**). As photons of light strike the junctions, small voltages are developed; these voltages are amplified and processed many times each second by circuitry attached to the charge-coupled device. This circuitry constructs an electrical signal that represents the total image striking the charge-coupled device. Charge-coupled device pickups are very resistant to shock, have precise image geometry, and should have nearly indefinite life. They also offer freedom from microphonic noise and image burn-in, which are commonly found in pickup tubes. *See* PHOTODIODE.

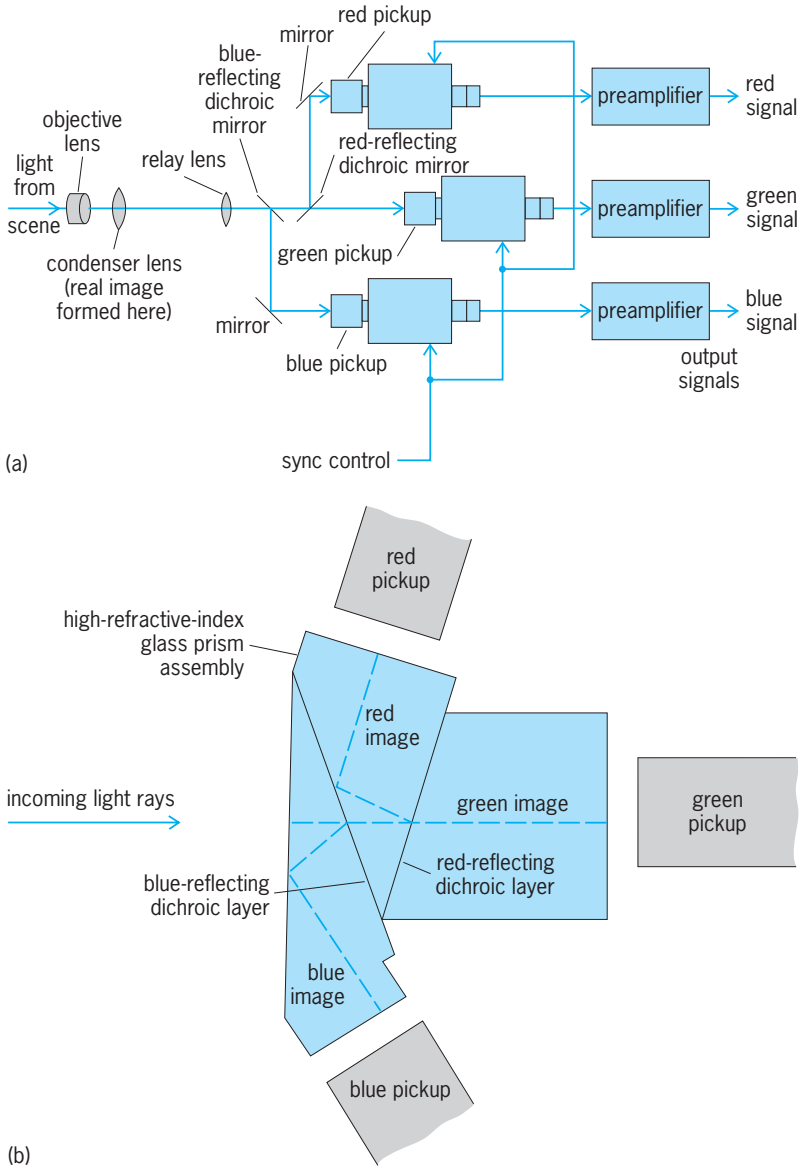

Fig. 3. Color television camera optical systems. (*a*) Dichroic mirror system (*after G. W. Bartlett, ed., NAB Engineering Handbook, 6th ed., National Association of Broadcasters, 1975*). (*b*) Dichroic prism system.

In those older cameras that still use vacuum pickup tubes, the tube's light-sensitive target is scanned in a geometric raster by an electron beam, resulting in an output voltage that varies in proportion to the amount of light striking each point on the target. Monochrome cameras and less sophisticated, consumer-type color cameras may utilize vidicon pickup tubes. High-quality broadcast cameras may use various types of tubes, such as lead oxide target tubes.

Monochrome cameras invariably have only one pickup device; color cameras may have one, two, or three pickup devices. Modern broadcast color cameras employ three pickup devices, one for each of three primary colors used to derive the full-color spectrum. Red, green, and blue are used in additive

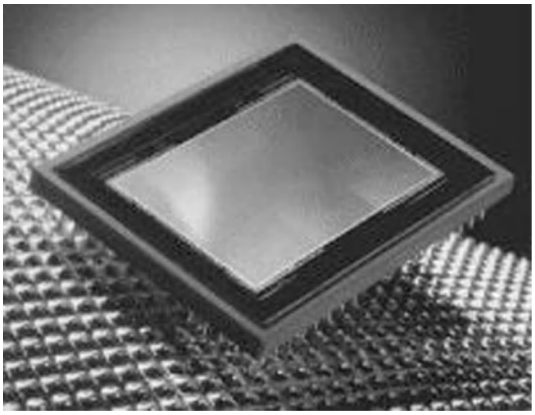

Fig. 4. Charge-coupled-device (CCD) image sensor array. (*Photo courtesy of Philips Semiconductors*)

color systems. While subtractive color systems have been devised that use white (full-spectrum luminance), red, and blue, most modern color cameras use the additive scheme. Each pickup device is scanned in synchronism so that separate red, green, and blue representations of the scene being viewed are always being generated within the camera. *See* COLOR; TELEVISION CAMERA TUBE.

Preamplifier. The pickup device's output is fed to a preamplifier, which helps to maintain a high signal-to-noise ratio and also provides a level sufficient to operate an electronic viewfinder at the camera. *See* PREAMPLIFIER.

Scanning, blanking, and synchronizing circuits. A major difference between charge-coupled devices and pickup tubes is the complexity of required scanning circuits.

The electron beam that scans the photosensitive target of the pickup tube is caused to sweep by means of the action of magnetic fields impressed on the tube by using deflection coils. One coil controls horizontal deflection while the other controls vertical deflection. The coils are often assembled into an integrated yoke assembly. Yokes must be made with great precision and are often computer-matched so that all pickup tubes in a multiple-tube camera scan with precisely the same geometry; otherwise misregistration and color fringing will result. In a mixed-field tube, one of the coils is replaced by electrostatic deflection plates. Horizontal and vertical drive circuits cause the deflection yokes to deflect the electron beam so that it scans the photosensitive target in the pickup tube according to a definite pattern, called a raster. These drive circuits are synchronized by pulses from a synchronizing generator, which also provides horizontal and vertical blanking pulses to suppress the pickup tube output during the retrace interval. If the pickup tube output was not suppressed, there would be objectionable lines through the camera's picture output. *See* TELEVISION SCANNING.

Charge-coupled-device cameras, on the other hand, do not require complicated deflection or retrace blanking circuits. By their very nature, they are scanned by using computer-type addressing techniques. Pulse counters, divider chains, and memory provide the correct sequential readout of the charge-coupled-device array so that a picture is recovered from the individual photodiode outputs.

Before the camera output can be viewed on a conventional video monitor, horizontal and vertical synchronizing (sync) and blanking pulses must be added. In two-piece cameras this is ordinarily done in the camera control unit.

Video processing and control circuits. Other functions that are necessary to obtain high-quality pictures include gamma correction, aperture correction, registration, and color balance. Gamma correction is required because the pickup devices do not respond linearly to increasing light levels. It allows the camera to capture detail in the dark areas of high-contrast scenes, essentially by "stretching" the video levels in those areas. Aperture correction provides several benefits mainly related to an even overall response to scenes with more or less detail. It also helps to improve the signal-to-noise ratio of the camera's output video. Registration must be adjusted on multiple-tube cameras to ensure that the separate red, blue, and green images are precisely aligned on one another; charge-coupled-device cameras are usually registered once, at the factory. Color balance must be properly set on color cameras and must be consistent from dark scenes to bright scenes, or there will be an objectionable tint to the camera output. This is referred to as black balance and white balance. The former is set up with all light blocked from the camera's lens, while the latter is set up by using a so-called white card under the actual lighting conditions of the scene. Color balance is frequently an automatic push-button feature on modern cameras.

All color cameras have a color encoder which combines the three primary colors into a composite signal that must conform to one of several standards used in various countries. In the additive color NTSC system used in the United States, the red (R), green (G), and blue (B) signals are matrixed by a combining network and amplified to provide three new signals, known as M, I, and Q. The M signal contains luminance information, while I and Q contain the color, or chrominance, information from the televised scene. Component M consists of 30% red plus 59% green plus 11% blue. The I signal consists of 60% red minus 28% green minus 32% blue. Component Q is made up of 21% red minus 52% green plus 31% blue. It is possible to subtract voltage levels by inverting the phase and summing, and this is exactly what is done in the color encoder (**Fig. 5**). I and Q then modulate a 3.58-MHz carrier in a two-phase (90° quadrature), balanced modulation system. The resulting amplitude-modulated sidebands are added to the M signal along with the synchronizing and blanking signals already mentioned. A short reference burst of unmodulated 3.58-MHz carrier is also added before the start of each television line. The resultant signal, which is known as composite NTSC video, contains all necessary information to recreate the original color scene on a color monitor or receiver. The M signal is used by black and white

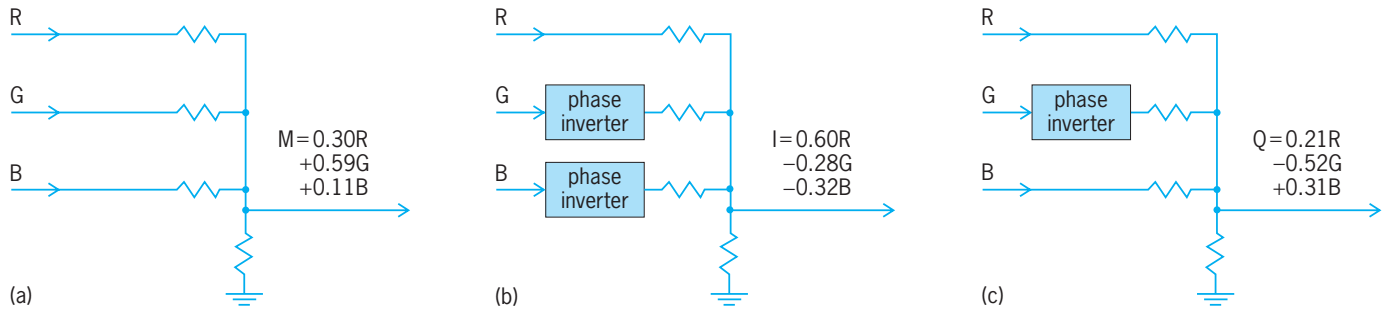

Fig. 5. Color matrix encoding diagrams showing derivation of (*a*) M, (*b*) I, and (*c*) Q signals. (*After G. W. Bartlett, ed., NAB Engineering Handbook, 6th ed., National Association of Broadcasters, 1975*)

monitors or receivers, which are insensitive to I and Q, to create an image in shades of gray. Color monitors and receivers have special circuits to process the I and Q signals back into the primary colors. In a black and white scene the red, green, and blue signals are equal, and therefore I and Q become zero and a black and white picture is produced on both monochrome and color television monitors. Most modern cameras rely heavily on digital integrated circuits and digital signal processing to achieve excellent performance in a compact form factor. *See* INTEGRATED CIRCUITS.

Typical configurations. Studio cameras (Fig. 1) are equipped with several ancillary systems to enhance their operation. An electronic viewfinder (actually a small television monitor) shows the camera operator what the camera is seeing, making it possible to frame and focus the picture. The tally system consists of one or more red lights that illuminate when the camera's picture is "on the line" so that production and on-camera personnel know which camera is active. Generally an intercom system is built into the camera so that the director can communicate with the camera operator. The camera itself may be mounted upon a tripod, but more often it is on a dolly and pedestal, which allows the camera to be moved around on the studio floor and raised or lowered as desired. A pan head permits the camera to be rotated to the left or right and furnishes the actual mounting plate for the camera. The lens zoom and focus controls are mounted on a panning handle convenient to the operator. A common accessory on studio cameras is the videoprompter. This is a television monitor on which the program script may be displayed and read by the on-camera personnel.

Some studio cameras are now being equipped with robotic control systems. These systems can pan, tilt, zoom, and focus the camera. More advanced systems can raise or lower the camera pedestal height, and some can move the camera around the studio floor under the control of a remote operator. The cameras can even be placed entirely under the control of a computer that makes all of the adjustments necessary to obtain the desired camera shots. *See* REMOTE-CONTROL SYSTEM; ROBOTICS.

Telecine cameras are used in conjunction with film or slide projectors to televise motion pictures and still images. Many of the usual controls are automatic so as to require less operator attention

The film projectors used in television may utilize either a constant-rate pulldown or 3-2 intermittent mechanism to translate the motion picture's 24 frames per second to the 30 frames-per-second television frame rate. The constant-rate pulldown has a shutter which is closed before the film is rapidly pulled. The shutter opens 120 times per second so that the camera sees each film frame in a 5:4 ratio. The 3-2 pulldown moves the film intermittently, so that one film frame is held in the film gate for two television fields ($^2/_{60}$ s) and the next film frame is held in the film gate for three television fields ($^3/_{60}$ s). Hence two film frames are displayed in $^5/_{60}$ s, and 24 film frames are shown in 1 s (30 television frames).

Generally, one telecine camera serves several projectors through the use of an optical multiplexer. The optical multiplexer may use stationary, half-silvered mirrors or movable, front-silvered mirrors to direct the image from each projector into the camera. Movable mirrors offer superior light-transmission efficiency and are preferable in color systems to minimize light losses. Telecine camera systems have largely been supplanted by flying-spot or linear-array charge-coupled-device scanners, which convert film images line-by-line directly into a video raster more precisely than the older telecine.

Portable cameras usually combine all of the basic elements into one package and may be used for a multitude of purposes. They have found their way into electronic news gathering for broadcast television, and into electronic field production, where they can be used for production of broadcast programs, commercials, and educational programs. The units often have built-in microphones, videocassette recorders, and batteries for completely self-contained operation (Fig. 2). One person can easily handle these compact and lightweight camcorders. The most compact units fit in the palm of the hand. They have become so popular and inexpensive that they have almost completely supplanted film-based home movie cameras.

HDTV cameras. Cameras used in high-definition television (HDTV) are fundamentally similar in appearance and operation to previous cameras. In fact, some modern cameras are switchable to produce either a conventional output or an HDTV output. The conventional output has a 4:3 aspect ratio raster

and the scan rates match the 525-horizontal-line, 59.94-Hz-vertical-field-rate NTSC standard in the United States. When switched to HDTV mode, the aspect ratio becomes 16:9 and the horizontal scan rate is usually increased to either 720 progressively scanned lines or 1080 interlace-scanned lines with a 60-Hz vertical field rate. Various manufacturers employ proprietary algorithms to derive 16:9 images from a 4:3 charge-coupled-device array and vice versa. *See* TELEVISION; TELEVISION RECEIVER; TELEVISION STANDARDS; TELEVISION STUDIO; TELEVISION TRANSMITTER.

Earl F. Arbuckle, III

Bibliography. T. D. Burrows et al., *Television Production*, 6th ed., 1995; P. Hodges, *Video Camera Operator's Handbook*, 1994; A. C. Luther, *Video Camera Technology*, 1998; National Association of Broadcasters, *NAB Engineering Handbook*, 8th ed., 1992; J. C. Whitaker and K. B. Benson (eds.), *Standard Handbook of Video and Television Engineering*, 3d ed., 2000; H. Zettl, *Video Basics*, 3d ed., 2000.

Television camera tube

An electron tube having a light-sensitive receptor that converts an optical image into an electrical television video signal. The tube is used in a television camera to generate a train of electrical pulses representing the light intensities present in an optical image focused on the tube. Each point of this image is interrogated in its proper turn by the tube, and an electrical impulse corresponding to the amount of light at that point of the optical image is generated by the tube. This signal represents the video or picture portion of a television signal. Television camera tubes are designed for broadcast television to pick up live programs, indoors or outdoors, as well as to reproduce motion pictures and other filmed material. *See* TELEVISION CAMERA.

The tubes are also used extensively in closed-circuit cameras for surveillance, and in training studios, schools, video tape recorder cameras, and military special-purpose cameras. Special versions are designed to work with intensifier tubes which increase the effective sensitivity so that the cameras can operate at very low light levels. These are used for nighttime surveillance work, television astronomy, and viewing low-intensity x-ray fluoroscope images in medical x-rays and in baggage inspection units in airports. In general, three tubes are used in color television cameras. A class of tubes has built-in stripe color filters which allow a single tube to develop a complete color picture, although with somewhat reduced fidelity compared to the multitube cameras. Although the television camera tube is sensitive primarily to visible light, special tubes are sensitive to radiant energy in the infrared and the ultraviolet.

Charge-coupled devices are a new generation of solid-state electronic image sensors which can produce a television signal from an optical image. These can operate independently or can be incorporated into an intensifier-type vacuum tube to achieve enhanced sensitivity. **Figure 1** shows contemporary television camera tubes. *See* CHARGE-COUPLED DEVICES.

Image orthicon. The image orthicon made broadcast television practical. It was used for more than 20 years as the primary studio and field camera tube for black and white and color television programming because of its high sensitivity and its ability to handle a wide range of scene contrast and to operate at very low light levels. It is one of the most complicated camera tubes. It is an outgrowth of the earlier multiplier orthicon and image iconoscope, and was made possible by the invention of the "two-sided" storage target. The image orthicon is divided into an image section, a scanning section, and a multiplier section (**Fig. 2**), within a single vacuum envelope.

Image section. A light image is focused on the photoemissive layer, which is a continuous film inside the tube faceplate. Electrons absorb the energy and leave the surface in numbers proportional to the intensity of the illumination at each point. These photoelectrons flow in essentially parallel streams through the image section. The magnetic field focuses each to a sharp focus at the target plane. *See* PHOTOEMISSION.

The two-sided target consists of a fine wire mesh screen placed several thousandths of an inch from a glass membrane less than 0.0002 in. (5.0 micrometers) thick. Most of the photoelectrons pass through the target mesh and hit the front surface of the target glass. Each photoelectron knocks several additional electrons from the target glass surface, producing a positive charge at the impact point. The secondary electrons are collected by the target mesh, which is held at a slightly more positive voltage. *See* SECONDARY EMISSION.

Scanning section. The positive charge pattern is stored on the front side of the target glass. A beam of low-velocity electrons generated by an electron gun is made to scan the rear surface of the glass by varying magnetic fields produced by the deflecting coils. As the beam moves across the glass, it deposits electrons wherever positive charges are built up on the image side. The glass resistance is controlled so that charges can move from one face to the other before the scanning beam returns to the same spot; yet the glass is of high enough resistance to inhibit lateral movement of the charges. When the scanning beam has deposited enough electrons at each point to neutralize the charge on the glass and reduce it to the voltage of the electron gun cathode, the remaining electrons return toward the electron gun. When the beam scans an uncharged (dark) area, the full beam is returned. When the beam scans a highly charged (bright) area, most of the beam is deposited and little returns. The variations in the return beam current constitute the television picture information, at low intensity. The return beam is amplified about 1000 times in the electron multiplier section of the tube and then is taken out at the anode of the multiplier as a video signal current. *See* CATHODE-RAY TUBE.

Fig. 1. Typical television tubes in modern use. From left to right: image orthicon, lead oxide vidicon, industrial-type vidicon, miniature vidicon, silicon intensifier vidicon, and charge-coupled device. (*RCA Corp.*)

Multiplier section. The electron multiplier is of unique construction, although it operates like the multiplier used in a multiplier phototube. It consists of a flat first-dynode structure and a series of pinwheel multipliers. When the return beam strikes the first dynode, a shower of secondary electrons cascades through the pinwheels, which are maintained at progressively higher voltages, where repeated secondary emission multiplies their number. The final group of electrons is collected by the anode and forms the video signal current. *See* PHOTOMULTIPLIER.

Image isocon. The image isocon is a further development of the image orthicon. The excess primary electrons in the scanning beam of the image isocon are returned from the target in two components: the scattered electrons and those that are reflected (**Fig. 3**). The image isocon works on the principle of separating out these two components of the return scanning beam and utilizing the scattered electron component, which has the highest signal compared to the random noise in the beam current. The separation section (**Fig. 4**) directs only the scattered beam into the electron multiplier. This improves the signal-to-noise ratio of the output signal and allows the camera utilizing the image isocon to operate at very low light levels in such fields as astronomy and intensification of x-ray fluoroscopic images.

Photoconductive tubes. These types have a photoconductor as the light-sensitive portion. A photoconductor is a material that absorbs light and transfers the energy of the photons of light to electrons in the material. This frees some of the electrons and allows them to move through the material, and thereby changes the electrical conductivity of the material where the light is absorbed. The electron tube is designed to detect this change in electrical conductivity and develop a television signal. *See* PHOTOCONDUCTIVITY.

The name vidicon was applied to the first photoconductive camera tube developed by RCA. It is

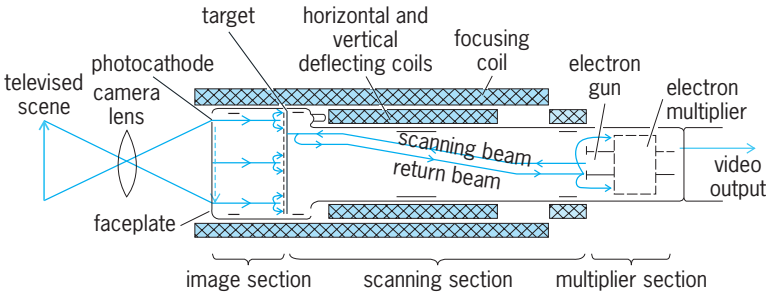

Fig. 2. Image orthicon and its associated deflecting and focusing coils. (*After D. G. Fink, ed., Television Engineering Handbook, McGraw-Hill, 1957*)

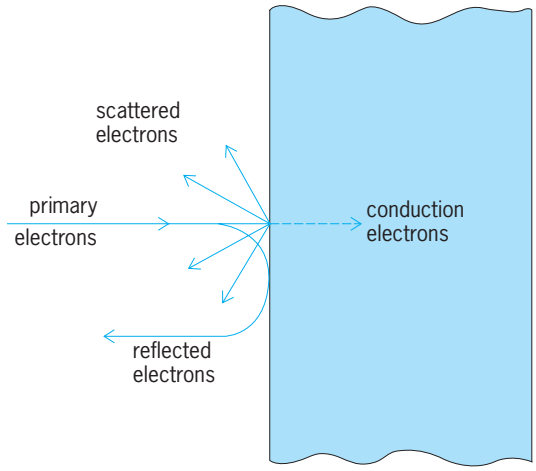

Fig. 3. Formation of the two components in the return beam of the image isocon: the scattered and the reflected electrons.

loosely applied to all photoconductive camera tubes, although some manufacturers adopt their own brand names to identify the manufacturer or the type of photoconductive material used.

The vidicon tube is a small tube that was first developed as a closed-circuit or industrial surveillance television camera tube. The development of new photoconductors has improved its performance to the point where it is now utilized in one form or another in most television cameras. Its small size and simplicity of operation make it well suited for use in systems to be operated by relatively unskilled people.

The vidicon is a simply constructed storage type of camera tube (**Fig. 5**). The signal output is developed directly from the target of the tube and is generated by a low-velocity scanning beam from an electron gun. The target generally consists of a transparent signal electrode deposited on the faceplate of the tube and a thin layer of photoconductive material, which is deposited over the electrode. The photoconductive layer serves two purposes. It is the light-sensitive element, and it forms the storage surface for the electrical charge pattern that corresponds to the light image falling on the signal electrode.

The photoconductor has a fairly high resistance when in the dark. Light falling on the material excites additional electrons into a conducting state, lowering the resistance of the photoconductive material at the point of illumination. A positive voltage is applied to one side of the photoconductive layer by means of the signal electrode. On the other side, the scanning beam deposits sufficient electrons at low velocity to establish a zero voltage. In the interval between successive scans of a particular spot, the light lowers the resistance in relation to its intensity. Current then flows through the surface at this point, and the back surface builds up a positive voltage until the beam returns to scan the point. The signal output current is generated when the beam deposits electrons on these positively charged areas. An equal number of electrons flow out of the signal electrode and through a load resistor, developing a signal voltage that is fed directly to a low-noise video signal amplifier.

A fine-mesh screen stretched across the tube near the target causes the electron scanning beam to decelerate uniformly at all points and approach the target in a perpendicular manner. The beam is brought to a sharp focus on the target by the longitudinal magnetic field of the focusing coil and the proper voltage for the focusing electrode. The beam scans the target under the influence of the varying magnetic fields of the deflecting coils.

Photoconductor properties determine to a large extent the performance of the different types of vidicon tubes. The first and still widely used photoconductor is porous antimony trisulfide. The

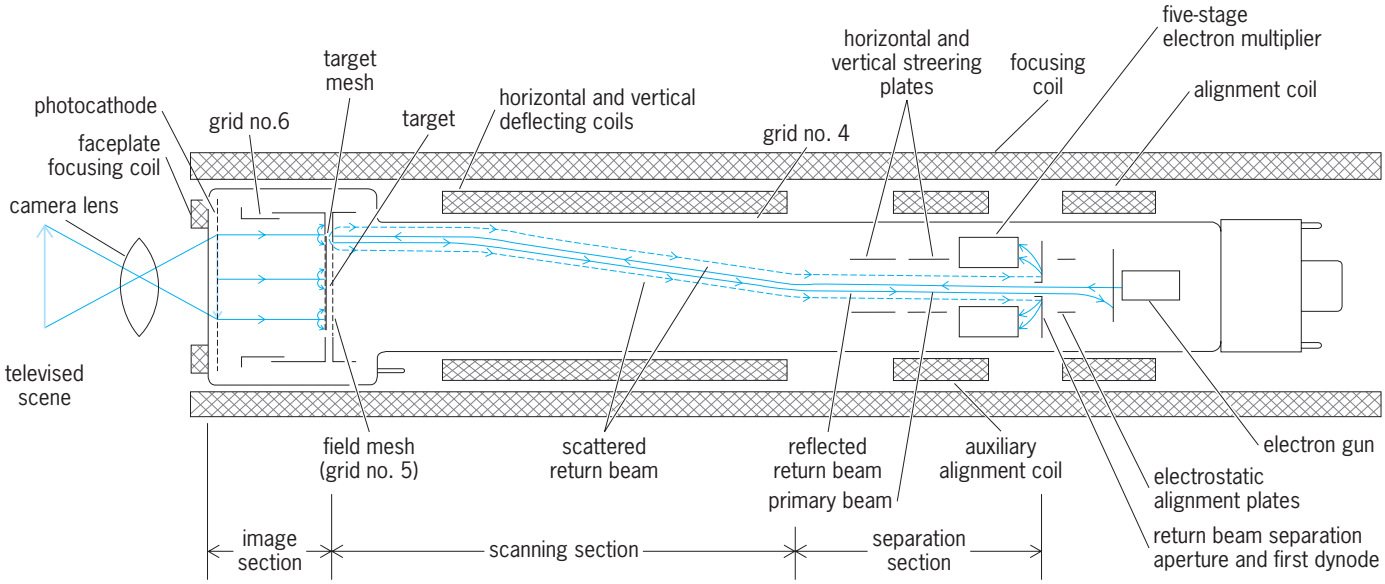

Fig. 4. Image isocon and 175 associated deflectron and focusing components. The separation section isolates the scattered and reflected return beam components. (*RCA Corp.*)

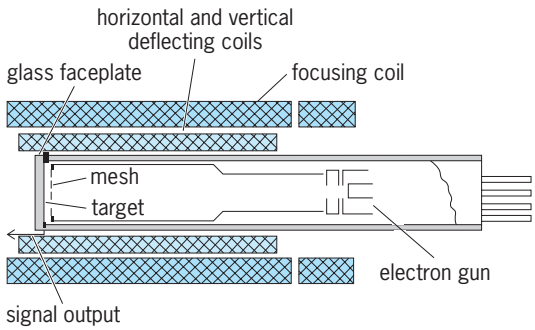

Fig. 5. Cross section of a vidicon tube and its associated deflection and focusing coils.

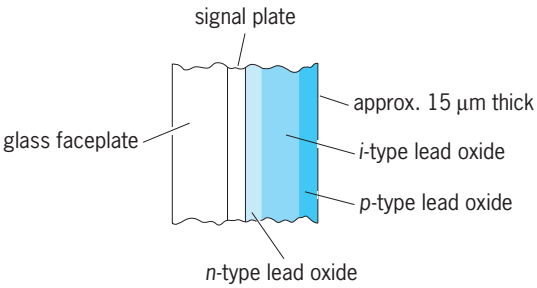

Fig. 6. Cross-sectional view of lead oxide barrier-layer type of photoconductor.

latest photoconductors are the lead oxide, selenium-arsenic-tellurium, cadmium selenide, zinc-cadmium telluride, and silicon diode arrays. All of these either improve the sensitivity or the speed of response (ability to capture motion without "smearing") or both. The first two are barrier-layer types that operate like large-area reversed-bias junctions. **Figure 6** illustrates the configuration of the barrier-layer or reversed-bias junction photoconductors; *n* refers to good electron conductivity, *i* refers to good electron and hole conductivity, and *p* refers to good hole conductivity. A positive voltage is applied to the *n* side through the signal plate, and a negative voltage is applied to the *p* side by the scanning beam electrons. This reverse-biases the junctions between the *i* and *p* and the *i* and *n* sections. This produces low dark current in the absence of light, a very desirable characteristic. Light is absorbed in the bulk of the material, where it produces charge carriers that provide a positive charge image on the side opposite the faceplate. The silicon diode tube is also a reversed-biased junction type, but consists of an array of hundreds of thousands of individual diodes on a wafer of silicon (**Fig. 7**). The photocarriers are generated in the silicon wafer, but are collected and stored on the diode cells (**Fig. 8**). *See* SEMICONDUCTOR.

Most color television cameras utilize the lead oxide or selenium-arsenic-tellurium tubes. Industrial and scientific-industrial cameras utilize the other types.

Silicon intensifier. The silicon intensifier camera tube utilizes a silicon diode target, but bombards it with a focused image of high-velocity electrons. These electrons are emitted by a photoemitter on the inside of the window on the front of the image section (**Fig. 9**). A fiber-optics window is utilized so that the emitting surface can be curved to produce good uniformity of focus of the high-energy electrons on the silicon diode target. Each high-energy electron can free thousands of electron carriers in the silicon wafer (compared to one carrier per photon of light on a silicon diode vidicon). This high amplification allows the camera to operate at light levels below that of the dark-adapted eye. With such a camera tube, it is possible to "see" the individual photons of light that compose a low-level optical image. The silicon intensifier tube is utilized for nighttime surveillance and other extremely low-light-level television uses in industrial, scientific, and military applications. It can be operated over a very wide range of light levels by varying the image section voltage. This changes the amplification over a range of more than 1000 to 1.

Solid-state imagers. These are solid-state devices in which the optical image is projected onto a large-scale integrated-circuit device which detects the light image and develops a television picture signal. Typical of these is the charge-coupled-device imager. The term charge-coupled device refers to the action of the device which detects, stores, and then reads out an accumulated electrical charge representing the light on each portion of the image. The charge-coupled device transfers the individual charges to the

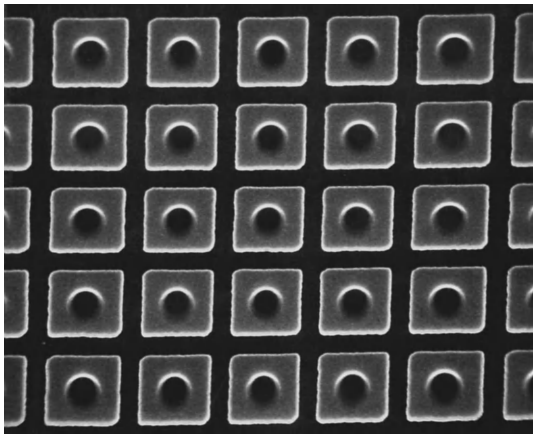

Fig. 7. Small section of the scanned side of a silicon diode target. Diodes are circles, and beam landing pads are squares. (*RCA Corp.*)

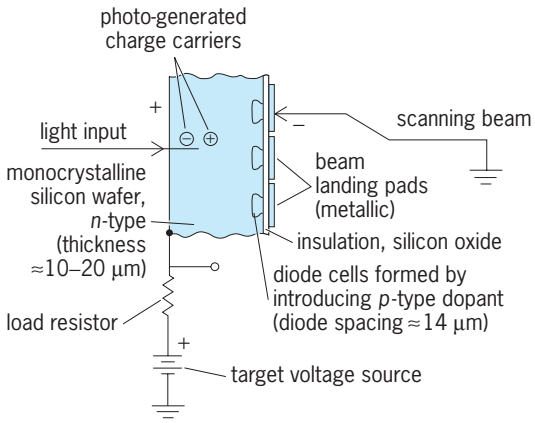

Fig. 8. Cross section of silicon diode target.

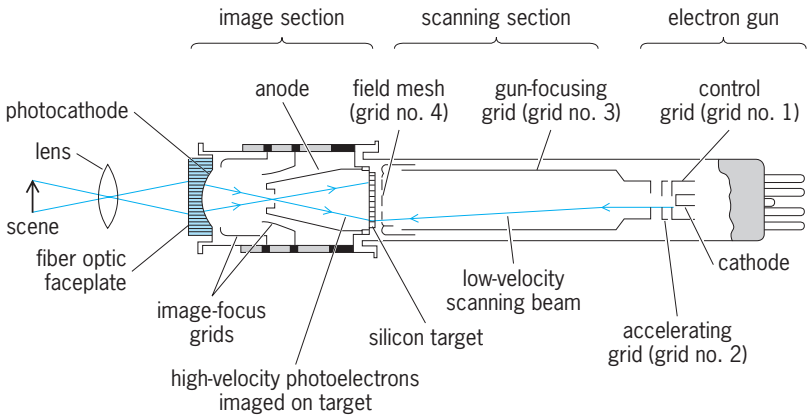

Fig. 9. Silicon intensifier vidicon. (*RCA Corp.*)

output in the proper television scanning sequence to constitute a television video signal. The device detects light by absorbing it in a photoconductive substrate, such as silicon. The charge carriers generated by the light are accumulated in isolated wells on the surface of the silicon that are formed by voltages applied to an array of electrodes on top of an oxide insulator formed on the surface of the silicon. These wells are actually small MOS (metal-oxide-semiconductor) capacitors (**Fig. 10***a*). Charges are transferred through the structure by varying the voltages on the metal electrodes. For example, if electrodes A and B are made more negative and C is made positive, the charges will move laterally from point 1 to point 2 (Fig. 10*b*). *See* INTEGRATED CIRCUITS.

A practical charge-coupled-device imager employing these principles consists of a structure that forms several hundred thousand individual wells or pixels, and transfers the charges accumulated out to an output amplifier in the proper sequence.

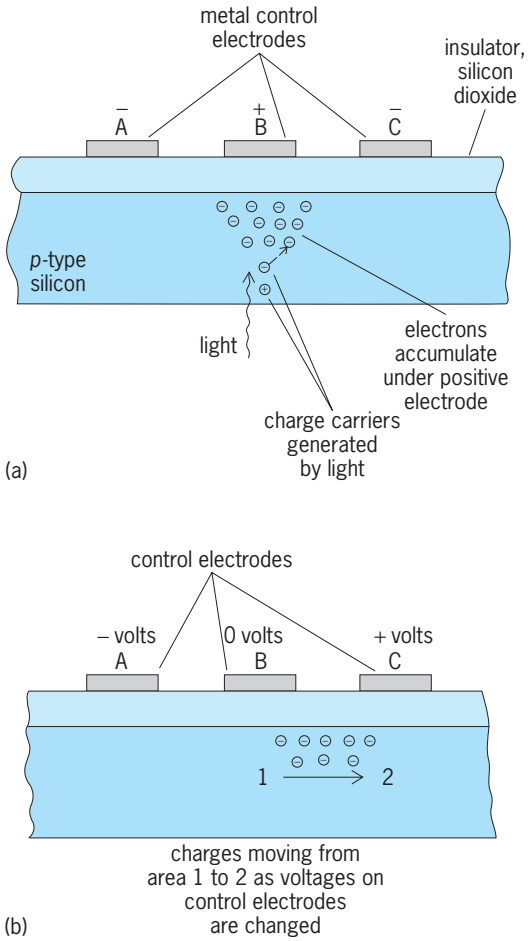

Fig. 10. Charge-coupled device. (*a*) Accumulation of an electron charge in a pixel element. (*b*) Movement of accumulated charge through the silicon by changing the voltages on the electrodes A, B, and C.

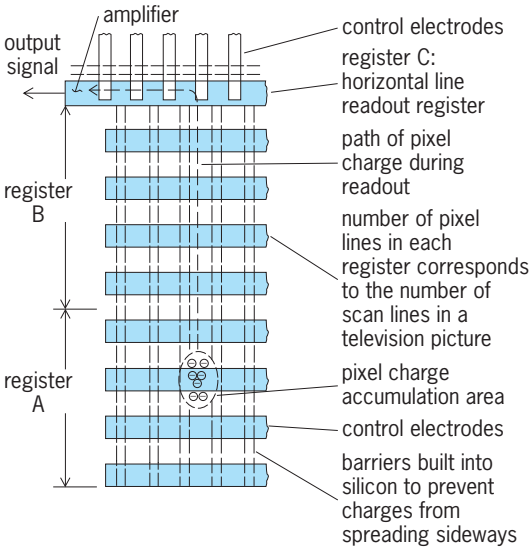

Fig. 11. One type of charge-coupled-device imager. Register A accumulates the pixel charges produced by photoconductivity generated by the light image. The B register stores the lines of pixel charges and transfers each line in turn into register C. Register C reads out the charges laterally as shown into the amplifier.

An example is shown in **Fig. 11**. Here the charges are accumulated by light exposure for the time it takes to complete a single television picture, or approximately $^1/_{\sqrt{60}}$ s. Then all of the charges are rapidly transferred line by line upward into the storage register. In the storage area all charges are then moved upward one scan line at the end of each television line interval. The upper line of charges is moved into the horizontal readout register. Then the pixel charges are moved to the left through the horizontal readout register to the output amplifier. During this readout period of the lines in the storage area, a new group of image charges are being accumulated in the imaging area, and the complete sequence is repeated every $^1/_{60}$ s.

Investigation has been undertaken of many versions of solid-state imagers using different light-sensitive materials and charge storage and readout methods.

A charge-coupled-device imager has been incorporated in an image intensifier tube. In this device the charge carriers are generated by a focused image of high-energy electrons in the charge-coupled-device imager in the same manner as in the silicon intensifier tube. This greatly enhances the sensitivity of

the device and allows it to operate at very low light levels. *See* LIGHT AMPLIFIER. Robert G. Neuhauser

Bibliography. K. B. Benson and J. Whitaker, *Television Engineering Handbook: Featuring HDTV Systems*, rev. ed., 1992; K. Jackson and B. Townsend, *Television and Video Engineer's Reference Book*, 1991; B. Kazan (ed.), *Advances in Image Pickup and Display*, 6 vols., 1974–1983.

Television networks

Arrangements of communications channels, suitable for transmission of video and accompanying audio signals, which link together groups of television broadcasting stations or closed-circuit television users in different cities so that programs originating at one point can be fed simultaneously to all others.

In the United States, television network service is furnished by the long-distance or local-exchange carriers (hereafter identified as telephone companies) and satellite carriers, as well as by broadcaster-owned networks. The facilities, when provided by the telephone companies, consist of intercity channels, which interconnect the principal long-distance telephone offices in various cities, and local channels, which connect the telephone offices with the broadcasters' studios or other user locations in each city. In the terminating offices, the intercity and local channels are brought together in a television operating center (TOC) where means are provided for testing, monitoring, and connecting the channels in various patterns as required for service. *See* TELEPHONE SERVICE.

Use. The principal users of television network facilities are broadcast network organizations (broadcast networks). These may be national, regional, or local in scope, and may be regular or occasional in nature and may be commercial or noncommercial. In the United States, special regulations apply to the practices of commercial broadcast networks which offer an interconnected program service on a regular basis for 15 or more hours per week to at least 25 affiliated television licensees in 10 or more states. Four major broadcast networks currently meet this definition: ABC, CBS, Fox, and NBC. In addition, there are at least two other emerging national broadcast networks, and two national broadcast networks in the Spanish language. These regular national broadcast networks provide programming to affiliates on a daily basis at prearranged times. There are also occasional and specialized broadcast networks that provide coverage of local and regional news and of sports events, and meet other programming needs. Finally, there is PBS, a national noncommercial broadcast network that supplies programming to stations specifically allocated for educational purposes.

Broadcast networks typically consist of a programming entity which simultaneously feeds content over the network interconnection facilities to commonly owned and operated (O&O) stations, as well to other stations (affiliates) that have a contractual relationship (affiliation agreement) with the broadcast network. Commercial broadcast networks typically pay compensation to affiliates for carriage of the broadcast network's programming, which includes commercial advertisements (spots) by the network. This benefits the broadcast networks by spreading the cost of programming for their O&O stations over a much larger base, as well as generating revenue through the sale of advertising on a national scale; it benefits the affiliates by reducing program acquisition costs, providing a source of direct revenue, and increasing the value of spots retained by the affiliate within and adjacent to the broadcast network's programming.

In the United States, the broadcast networks have long-term arrangements to provide their primary feed to affiliates. They typically transmit their main programming to O&O stations and affiliates over two separate sets of communications channels. One feed is provided to those stations in the Eastern and Central Time Zones; a separate feed is provided to stations in the Pacific Time Zone to eliminate the need to tape-delay programming, and to permit programs to be seen at the same clock time by populations in the east and west. Some Mountain Time Zone stations are fed with the Eastern feed, others with the Pacific feed. In addition to these primary feeds, the broadcast networks provide a separate feed of programming in a digital format to be used by O&O stations and affiliates in the operation of their digital television (DTV) stations. These can be in either the high definition (HDTV) or standard definition (SDTV) formats.

In addition to their regular interconnection channels, the broadcast networks frequently order "occasional" channels to pick up programming from other than the usual originating points. An example is the origination of sports and special events programming from remote points, as well as the reception of material from diverse points to be included in news programming.

History. The first intercity television transmission in the United States occurred in 1927. Coaxial cable was installed in 1936 to permit the transmission of television images and sound between New York and Philadelphia. By 1940, coaxial cable was being used to provide programming for commercial broadcast.

During this period, United States broadcasters supported a private radio relay system linking television stations, while the telephone companies proposed use of coaxial cables and microwave links. This was resolved when federal regulators supported the common carrier approach favored by the telephone companies.

Following World War II, the telephone companies in the United States began to provide facilities in many major cities for television interconnection. Most major eastern cities were linked by the end of 1947. In 1949, a circuit linked the east with Chicago, first shared by all networks, and subsequently with each obtaining a discrete circuit. By August 1951, the transcontinental line was completed, enabling the broadcast networks to broadcast simultaneously from coast to coast.

The national grid grew explosively during the 1950s, so that by the end of the decade every significant community in the continental United States was capable of receiving a network feed. Stations in Hawaii and Alaska, however, continued to rely on "bicycling" of film, kineoscope transcriptions, and video tape from west coast stations until the late 1970s.

As time passed, the use of coaxial cable for intercity transmissions greatly diminished, replaced by a grid of microwave facilities that relay television and other communications transmissions. By the mid-1970s, communications satellites began to be used more and more for occasional channels to "back-feed" programming to the television networks, and the broadcast networks subsequently began to use communications satellites for the primary program feed to their O&O stations and affiliates. Now fiber-optic rings are being used for local television interconnection, and fiber-optic is being used for some long-haul video traffic.

Transmission requirements. United States television networks transmit color signals employing the National Television Systems Committee (NTSC) standards of monochrome information similar to that of an ordinary black-and-white picture in the frequency band from 30 Hz to about 4 MHz, plus color information transmitted by a color carrier at approximately 3.6 MHz which is both amplitude- and phase-modulated and whose upper sideband extends to 4.2 MHz. Two additional frequency-modulated (FM) subcarriers at 5.8 and 6.4 MHz are used to provide the simultaneous transmission of the audio portion of the signal over terrestrial systems. When transmission is over satellite, the subcarriers usually are at 6.2 and 6.8 MHz. Networks in other countries employ the PAL (phase alternation by line) or SECAM (sequential color with memory) standards. Some network have begun to transmit signals intended for broadcast by digital stations under the Advanced Television Standards Committee (ATSC) standards. *See* FREQUENCY MODULATION; TELEVISION STANDARDS.

Satisfactory transmission of these signals requires that the video channels pass the broad range of frequencies from 30 Hz to 6.5 MHz with negligible distortion. Performance standards (NTC7) were set by the National Transmission Committee, a joint effort of the major United States broadcast networks and the telephone companies. The channels used are designed to have attenuation and delay characteristics which are essentially flat over this range, and to produce a signal-to-noise ratio of 50 dB or better (peak power signal to root-mean-square noise) from pickup to ultimate user over a distance of up to 4000 mi (6500 km) or over a satellite transmission link. Differential distortion—that is, nonlinearity of output compared with input over the whole range of levels to be transmitted—must also be kept to a minimum. Otherwise, the color carrier could vary in amplitude or phase as the monochrome signal changes from black to white. *See* DISTORTION (ELECTRONIC CIRCUITS).

Interexchange carrier channels. The facilities for intercity video channels include terrestrial microwave networks, synchronous satellite systems, and fiber-optic systems. The terrestrial microwave networks are FM radio systems, usually in the frequency band of 3.7–4.2 GHz, with some links in the 6-GHz and 11-GHz bands and higher. Depending on service requirements and design, these microwave systems provide up to 12 wide-band channels in each direction. Relay stations are spaced an average of 25–30 mi (40–50 km) apart. Automatic switching is provided at section terminals that transfer service to a protection channel in case of equipment failure or severe fading on a regular channel. Satellite systems are used to provide both domestic intercity transmission and transmission to and from points located overseas. Fiber-optic links are used for local loops and for point-to-point transmissions over long distances. *See* COMMUNICATIONS SATELLITE; MICROWAVE.

Local channels. The local channels, which connect the telephone company office with the broadcasters' studios and transmitters and with closed-circuit users' locations, are largely composed of cable and fiber-optic facilities. To minimize low-frequency interference, cable channels make use of special balanced video cable pairs, generally 16-gage wire insulated with a low-loss dielectric. Each video pair is shielded by copper tapes or woven copper mesh. The attenuation of a 16-gage video pair is about 19 dB/mi (11 dB/km) at 4.5 MHz. Video amplifiers are spaced at intervals up to 4.5 mi (7.2 km) to compensate for attenuation. *See* COMMUNICATIONS CABLE.

Use of fiber-optic systems for local video distribution is growing in popularity, as fiber is a very effective means of point-to-point communications, but it is not yet economically viable for point-to-multipoint communications. The fiber-optic transmissions can be either FM or digitally modulated. The use of FM overcomes signal-to-noise-ratio and intermodulation distortion, removing the problems of signal quality from the optical to the electrical domain. The video baseband, with multiplexed FM subcarriers containing audio and data signals, is frequency-modulated on a carrier to produce a signal with a fixed amplitude excursion, the value of which modulates the fiber's light source over a linear portion of its range. Currently, fiber-optic systems using digital modulation often use dense wavelength division multiplexing (DWDM), which permits multiple independent signals to be transmitted over a single fiber by using different optical wavelengths. *See* MODULATION; OPTICAL COMMUNICATIONS; OPTICAL FIBERS; PULSE MODULATION.

Local channels may also be provided by fixed or portable microwave radio facilities. These can use either FM or digital modulation. Economic considerations usually determine whether radio, fiber, or cable should be used, but the decision may be influenced by the need for a speedy installation, such as for one-time news or sports events.

Television network control centers. In network operations, frequent rearrangement of the channel connections is required to change the point of origin

or to feed successive programs to different groups of broadcasting stations. These rearrangements are accomplished by means of the switching systems in network operating centers. Switches may be made either on a time schedule or on voice cue. Scheduled switching, when ordered, is done between the various 30-min segments into which the broadcasters divide their daily programming. Switching on cue is required for programs, such as athletic events, with indefinite completion times, or to rearrange a network in midprogram so that different commercials can be delivered to various geographical regions.

Although most television programs originate in the studios of the broadcasting companies, many originate at other locations, such as sports stadiums. Currently, most "backhaul" feeds from fixed locations hosting recurrent televised events, such as sports stadiums, are transmitted by fiber-optic facilities. Local feeds from other remote locations use microwave relay, and feeds to national program services are typically relayed directly by satellite to the network's main control center. From the control center the program is fed to the permanent network.

Fiber-optic distribution. There is also a growing use of fiber-optic distribution of television signals, particularly in backhaul and other point-to-point transmissions. Signal deterioration due to transmission-medium attenuation and noise addition is completely removed by regeneration, the process of detecting the digital signal, performing error correction, and retransmitting the signal to the next regeneration station. In this way, error-free transmission to any distant location is possible.

Fixed satellite networks. Transmission by satellite is widely used by broadcast networks, by cable television distributors, and in closed-circuit applications. The broad-frequency band-pass of the satellite transponder and the fact that the satellite transmitting antenna can be designed to provide transmission to a broad geographic area make satellites an attractive tool for providing point-to-multipoint video services. Satellite television service is configured so that one of several Earth stations sends the program (uplinks the program) to a satellite transponder. The signal is amplified, translated (converted from the uplink to the downlink frequency), and transmitted back to Earth. The satellite transmitting antenna pattern can be constructed so as to cover a large geographical area such as the contiguous United States, and on some satellites certain transponders serve offshore points such as Alaska, Hawaii, and Puerto Rico. It is this point-to-multipoint mode (that is, one station transmitting to many receiving stations) of satellite transmission that makes satellite distribution so attractive to program distributors. In a typical television network application, a master network transmitting Earth station can control many receiving Earth stations. Control may be extended to the receiving locations by a terrestrial data line or over the satellite on a low-power subcarrier along with the television signal. Regionalization of a satellite network, such as for football or baseball telecasting, where different games are to be seen in different cities, requires separate transponders for each event and close control of each receiving location.

Use of satellites for news backhaul. The ability to uplink to satellites from virtually anywhere in the world with portable equipment has made news reporting almost instantaneous. News backhaul usually uses Ku-band satellites, with uplink transmissions to the satellite at 14–14.5 GHz, and downlinks to receive stations at 11.7–12.2 GHz. The major portion of news backhaul signals are delivered to the editing point by frequency-modulation of the video, multiplexed with FM subcarriers of the related audio. This occupies about 27 MHz of bandwidth per program. Digital video compression and transmission enables more news backhaul to be completed with portable equipment, using lower satellite power and reduced transponder bandwidth.

Direct broadcast satellites. Geostationary satellite antenna footprints can be made to cover large areas of the Earth, thus making them ideal for point-to-multipoint transmissions. Direct broadcast satellite systems have been placed into operation in many developed nations. Their implementation in the United States is typical of other systems.

After the cable television industry in the United States began in the late 1970s to use the relatively low-powered (10–17 watts per channel) C-band fixed service satellites to expand the availability of program sources, individuals without access to cable saw satellite reception as a means of expanding their own programming options. The television receive-only (TVRO) industry began to provide consumers with receive antennas, typically using parabolic dishes of 10–12-ft (3–4-m) diameter, and associated electronic equipment. Program suppliers responded to this by using encoders to scramble their signals, permitting only authorized users access to decoding devices. The U.S. Congress subsequently passed legislation that required program suppliers to give consumers access to scrambled satellite programming, and integrated receiver decoders (IRDs) were marketed that permitted consumers to subscribe to and receive satellite programming services.

The Federal Communications Commission (FCC) allocated a portion of the Ku band, from 17.3 to 17.8 GHz, and downlink frequencies from 12.2 to 12.7 GHz to a new Broadcast Satellite Service in 1982. Space stations in that service are known as direct broadcast satellites (DBS). These satellites typically have 32 transponders of 120 W output each, although they can be reconfigured to operate with a lesser number of higher-power, 240-W transponders. These satellites receive and retransmit digital signals, and the amount of transponder power output has a direct bearing on the bit rate that can be handled by the satellite. *See* DIRECT BROADCASTING SATELLITE SYSTEMS.

In order to be competitive with terrestrial cable systems, the DBS system must provide consumers with far more than 32 program services. Therefore, the DBS providers have chosen to operate digital

systems using significant amounts of compression. This permits the 32 transponders to carry approximately 200 separate program services, often with a picture quality surpassing terrestrial television. To do this, the systems make use of the MPEG2 compression standard for video transmission, and Layer II of the MPEG1 compression standard for audio. The compression technology permits adaptive use of the available spectrum. Thus, programming containing little motion or change from scene to scene is more highly compressed than material requiring greater detail. This permits spectrum and transponder power conservation while preserving picture quality. *See* DATA COMPRESSION.

There are two principal DBS licensees operating in the United States: DirecTV and EchoStar, which does business as the Dish Network. The DirecTV operations center is in Castle Rock, Colorado, and the EchoStar operating center is in Cheyenne, Wyoming.

Earth stations at the operations centers receive programming materials from multiple sources. In addition, the centers contain banks of program origination equipment so that pay-per-view movies and other programming can be originated directly from the center. The source materials are input to video and audio compressors. The outputs, in the form of packetized elementary streams (PES), are input to a multiplexer. Several separate services are transformed into a single bitstream for each transponder. Forward error correction (FEC) is then applied to the bitstream from the multiplexer, and then input into a quadrature phase-shift keying (QPSK) modulator. The bitstream from the modulator is then frequency-upconverted to the appropriate transponder frequency and transmitted to the satellite. *See* INFORMATION THEORY.

The satellite transponds the signal and retransmits it to the Earth, where it is received by a small receive dish, typically an 18-in. (50-cm) reflector antenna. The antenna utilizes an offset low-noise block (LNB) amplifier. The low-noise block downconverts the signal to the L band (950–1450 MHz) and sends it via coaxial cable to the settop integrated receiver decoder. The integrated receiver decoder selects the appropriate transponder, demodulates the QPSK waveform, removes the forward error correction, demodulates it to extract the selected service from the bitstream, stores the signal in a buffer stage for further processing, and then decompresses the audio and video for reception by the television receiver.

During its initial years of operation, the DBS operators provided primarily nonbroadcast programming services to their subscribers. Subscribers wishing to receive local broadcast signals either had to install off-air antennas or subscribe to cable service in addition to their satellite subscriptions. Since the passage of the Satellite Home Viewer Improvement Act (SHVIA) in 1999, satellite television companies have begun to offer local programming to many areas throughout the United States, and as new satellites with spot-beam technology replace those currently in orbit, greater portions of the United States will be able to receive local signals by DBS.

Cable distribution by satellite. The C band (5925–6425 GHz up, and 3.7–4.2 GHz down) is used for cable television–satellite television networking. Cable television systems receive programming by satellite, and this programming is then sold to subscribers. Because of the premium value of these signals, they are frequently scrambled so they can be viewed only through appropriate equipment.

While all satellite television networks once used FM transmission, many operators now use digital video compression to transmit 4–10 signals over a single satellite transponder using quadrature phase-shift keying (QPSK) modulation. This changes the instantaneous amplitude of the two orthogonal carriers (I and Q) at a rate of about 38 Mbps. The received data require good forward error correction for perfect decompression to video and audio. Picture quality at four to six compressed video programs per satellite transponder is as good, or better, than FM-transmitted signals.

Cable delivery methods. Over 96% of television households in the United States have access to cable television, and most of them receive all their television programming in analog format via cable. Programming usually includes the local broadcast stations and satellite services.

Because of the closed nature of cable television systems, all the frequencies between 50 and 750 MHz and higher can be used for distribution. The analog signals received at cable television headends are combined side by side on a frequency-division multiplex (FDM) basis for carriage in the cable systems. Digital satellite signals are decompressed in integrated receiver decoders, and the programs are passed through to the subscribers as analog signals (**Fig. 1**). *See* MULTIPLEXING AND MULTIPLE ACCESS.

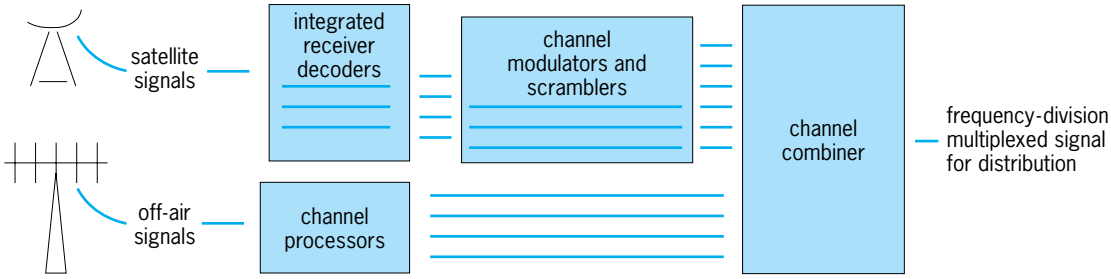

Fig. 1. Typical cable television headend. The processing used in creating the frequency-division-multiplex signal is shown. (*General Instrument Corp.*)

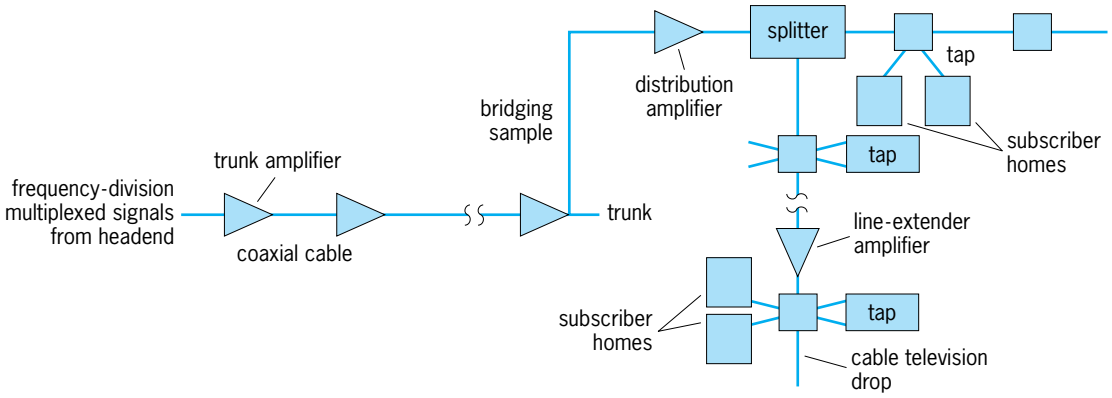

Fig. 2. Coaxial cable television system. (*General Instrument Corp.*)

Since many satellite-delivered programs are premium services, scrambling is applied to the cable signals and descramblers are leased to subscribers. Those descramblers often are necessary for distoration-free viewing of the available channels on the cable system.

Coaxial cable systems. Cable television systems used to have very long cascades of amplifiers, interconnected by coaxial cable. The loss of a coaxial cable increases in proportion to the square root of the frequency. Thus, the loss doubles when the frequency quadruples. For example, the loss at 200 MHz is twice that at 50 MHz, and the loss at 800 MHz is twice that at 200 MHz. *See* COAXIAL CABLE.

Amplification at higher frequencies is also more difficult to accomplish. When large numbers of amplifiers are cascaded, the flatness of the frequency response is critical because individual amplifier flatness deviations add to each other. Older cable television systems often had amplifier cascades of 40 or more. When an amplifier operates in a cable television system, the amplitude at which the signals can be carried is limited by the amplifier output capability, the noise contributed by the amplifier, and the number of television channels being carried. *See* AMPLIFIER.

A number of methods of improving the reach and quality of cable television signals are employed. One technique is to build the cable television facility in two parts. The trunk, with very large diameter cable to minimize loss, carries the signals most of the distance. The distribution system, built in parallel, bridges some of the trunk signal and, using less costly smaller-diameter cable, feeds the signals along the streets to the homes.

The amplifiers used in the trunk are very linear and often share the same housing as the bridger amplifier. Line-extender amplifiers are used when the taps feeding individual homes reduce the signal level to the point of requiring amplification. Distribution amplifiers generally have high power, lower flatness performance, and lower cost (**Fig. 2**).

In addition to larger coaxial cable, supertrunks are sometimes used in which FM signals are carried in a frequency-division-multiplex frequency plan and demodulated in a subheadend. This technique is expensive, but less than creating a second headend.

These techniques are sometimes supplemented by the use of amplitude-modulated links. Here the frequency-division multiplex cable television signals are sent by microwave radio to remote receivers acting as remote headends for normal cable television distribution. These systems are also beneficial when rivers and lakes block the routing of coaxial cable. They have problems of cost, licensing procedures, and rain fade, since they operate at frequencies near the Ku band.

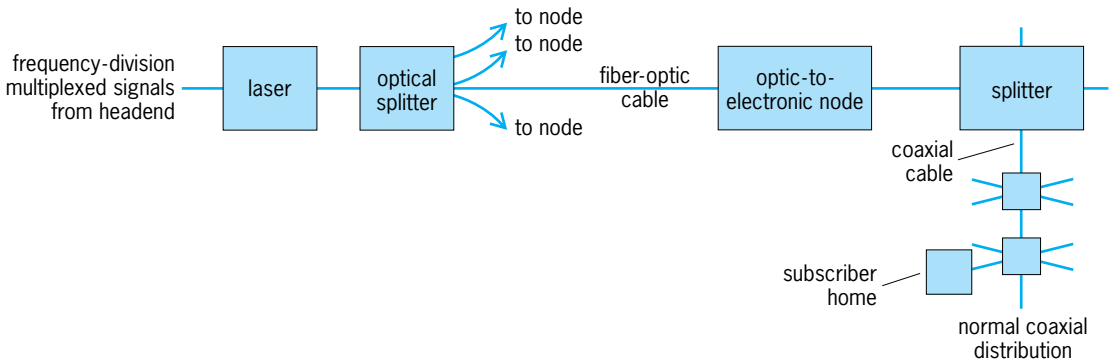

Fig. 3. Hybrid fiber-coaxial television system. (*General Instrument Corp.*)

Fiber optics in cable television. Most modern cable systems use fiber optics to replace the coaxial trunk cable completely, while imposing only the noise and distortion equal to one trunk amplifier. The wide use of fiber optics by cable systems contrasts with broadcast networks. The small optical-fiber light path requires lasers generating coherent light to launch significant light into the fiber. The modulation of these lasers is linear in that television signals are modulated by using frequency-division multiplexing onto a laser and then into the fiber-optic cable. *See* LASER.

The fiber-optic cables use two of the lowest optical frequencies capable of supporting low-loss single-mode transmission. Losses increase as the optical frequency is increased, just as happens in coaxial cable. The two widely used wavelengths are around 1310 and 1550 nanometers. Between them is a quite lossy water absorption line.

The optical-fiber cable used in cable television systems is either step-index or graded-index, both with almost equal performance. Both fibers have a loss of about 0.56 dB per mile (0.35 dB per kilometer) at 1310 nm, and 0.35 dB per mile (0.22 dB per kilometer) at 1550 nm. This is very low compared to 1 dB per 100 ft (33 dB per kilometer) for 1-in.-diameter coaxial cable at 750 MHz.

Distributed-feedback lasers operating at 1310 nm are generally used, with power outputs up to about 12 milliwatts. This power is seldom used for a single fiber-optic run, but is split to feed a number of cable television nodes. Each node is a local distribution system delivering cable television to 500–2000 subscribers. Usually, the area is limited by the number of customers or a maximum of four coaxial cable amplifiers in cascade. The desire is to design cable television distribution systems that will permit a reasonable number of customers to be served with all the planned services, including telephone, television, interactive services, multimedia, and two-way data services.

Hybrid fiber-coaxial (HFC) cable television systems require about 1 mW of received optical power in order to maintain a good carrier-to-noise ratio at the node where the optical power is converted to electronic power. For 1 mW out, a higher power input is required. For a 6.25-mi (10-km) span from the headend to the distribution node, about 4 dB of optical loss is incurred. This means that the power at the input must be 4 dB above 1 mW or 2.5 mW, permitting dividing the original 12 mW to supply four such paths (**Fig. 3**). *See* CABLE TELEVISION SYSTEM; TELEVISION. Joseph B. Glaab; Clifford M. Harrington

Bibliography. W. Cicoria et al., *Modern Cable Television Technology*: *Video, Voice, and Data Communications*, Morgan Kaufman Publishers, New York, 2d ed., 2004; D. Roddy, *Satellite Communications*, 4th ed., McGraw-Hill Telecom Engineering, New York, 2006; J. Whitaker and B. Benson, *Standard Handbook of Video and Television Engineering*, 4th ed., McGraw-Hill, New York, 2003; E. Williams et al., *National Association of Broadcasters Engineering Handbook*, 10th ed., Elsevier, 2007.

Television receiver

The equipment used to receive the transmitted modulated radio-frequency signals and produce synchronized visual images and sound for entertainment or educational purposes. The radio-frequency portion operates on the superheterodyne principle. *See* MODULATION; RADIO RECEIVER.

The first television receivers to be mass-produced were monochrome; that is, they provided pictures in black and white only. Later, color receivers, which produce pictures in full color as well as black and white, became available. Some television receivers now can receive stereophonic sound or alternate language in accordance with multichannel television sound standards. For basic discussion of a television system. *See* TELEVISION; TELEVISION STANDARDS.

Early television receivers used vacuum-tube technology. Present-day receivers use solid-state technology with many functions integrated on a few chips. The only function still primarily implemented by using vacuum-tube technology is the display by the cathode-ray tube (CRT). *See* CATHODE-RAY TUBE; INTEGRATED CIRCUITS; VACUUM TUBE.

Monochrome receivers. Figure 1 shows a block diagram of a conventional monochrome television receiver.

Antenna and transmission line. Since most broadcast television transmissions in the United States are horizontally polarized (some are circularly polarized), the most basic type of television-receiving antenna is the horizontally mounted half-wave dipole. Because the stations serving a given area may operate on widely different frequencies, however, the dipole dimensions must be a compromise that permits reasonable performance on all the desired channels. More complex antennas combine several dipole elements of various lengths, and passive reflectors may be used to achieve some degree of horizontal directivity, which increases the amplitude of the receiver signal and reduces interference from other stations. Highly directive antennas are frequently mounted on remotely controlled rotators so that they can be pointed in the direction providing the best reception of the desired signal. The most common types of transmission line between the antenna and receiver are 300-ohm "twin-lead," employing polyethylene as a dielectric spacer between two uniformly spaced, unshielded wires. Also 75-ohm coaxial cable is used. *See* ANTENNA (ELECTROMAGNETISM); COAXIAL CABLE; POLARIZATION OF WAVES; TRANSMISSION LINES; YAGI-UDA ANTENNA.

Tuner. The tuner of a television receiver selects the desired channel and converts the frequencies received to lower frequencies within the passband of the intermediate-frequency amplifier. For very high-frequency (VHF) reception older tuners have 12 discrete switch positions, corresponding to channels 2–13. For ultrahigh-frequency (UHF) reception, continuous tuning is employed in older tuners. Most receivers now use a frequency synthesizer circuit for generating the local oscillator frequency, and may be stabilized with an automatic frequency control.

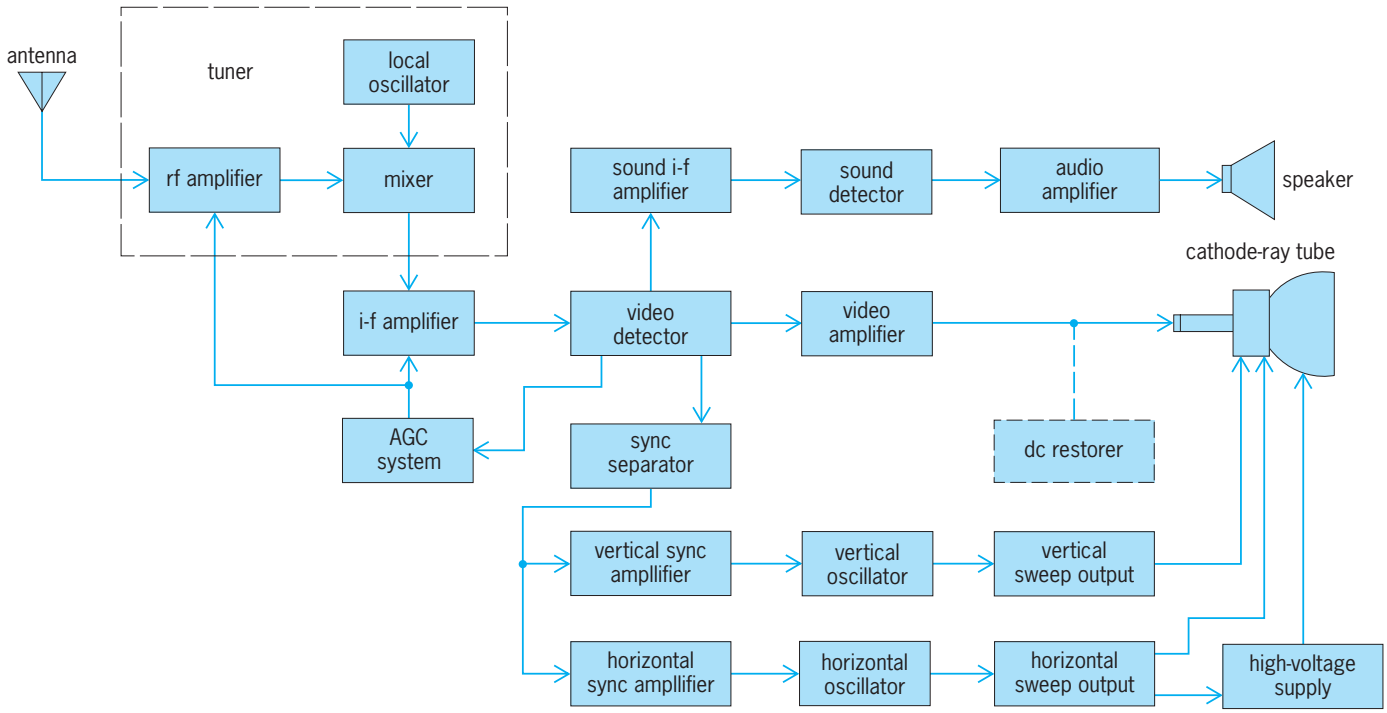

Fig. 1. Block diagram of a typical monochrome television receiver.

Nearly all VHF tuners employ a radio-frequency (rf) amplifier, a mixer, and local-oscillator circuits arranged as shown in Fig. 1. Some receivers are provided with additional channel-tuning capability for use with a community antenna television (CATV) system. *See* AMPLIFIER; AUTOMATIC FREQUENCY CONTROL (AFC); CLOSED-CIRCUIT TELEVISION; OSCILLATOR; RADIO-FREQUENCY AMPLIFIER.

The received signal and the local oscillator signal are applied to the mixer. Difference frequencies, representing the picture and sound carriers, are produced and remain essentially constant as the radio-frequency amplifier, mixer, and oscillator circuits are tuned to the different channels. Known as intermediate frequencies (41.25 MHz for sound and 45.75 MHz for picture), they are available for further amplification.

Such performance characteristics as noise factor, gain, bandwidth, and oscillator radiation must be optimized in the design of the tuner.

Intermediate-frequency amplifier. The output from the tuner is applied to the intermediate-frequency (i-f) amplifier. The gain of this amplifier is essentially constant from 43 to 45 MHz. Above 45 MHz the response decreases such that at 45.75 MHz, the picture carrier frequency, it is 50%. This slope is required to compensate for the vestigial sideband transmitted signal.

Below 43 MHz the response decreases until at 41.25 MHz, the sound carrier frequency, it is 5–10% of the flat response. This minimizes cross modulation between picture and sound carriers. Fixed tuned trap circuits are used to produce sharp cutoffs at the lower and upper limits of the intermediate-frequency passband. Sufficient selectivity is provided to minimize interference from signals originating in adjacent television channels.

Separation of video and audio. The output of the intermediate-frequency amplifier consists of two modulated radio-frequency signals. One of these, which is amplitude-modulated, provides a varying signal corresponding to the black and white portions of the picture, a blanking signal to render the return trace invisible on the picture tube, horizontal sync pulses to initiate the retrace of the beam at the end of each line, and vertical sync pulses to initiate the retrace of the beam at the end of each picture field. The other signal is frequency-modulated and contains the transmitted sound information. *See* AMPLITUDE MODULATION; FREQUENCY MODULATION.

There are two types of detector which may be used for video modulation detection: the envelope detector and the synchronous detector. In the envelope detector the two radio-frequency signals are applied to a diode which produces a rectified output that follows the instantaneous peak value of the amplitude-modulated picture carrier. *See* AMPLITUDE-MODULATION DETECTOR; DIODE.

The synchronous detector is a highly linear device and is used in the better television receivers where high-quality results are desired for color, teletext, alphanumeric and graphical data digital signals, and multichannel sound. The band-pass filter characteristic has much less attenuation for the sound carrier and upper video-frequency components. The synchronous detector is basically an analog multiplier.

The polarity of the video detector output depends upon the design of the video amplifier and method of picture-tube drive. Usually, maximum picture carrier (sync-pulse modulation) produces a negative output voltage.

Coincidentally a 4.5-MHz signal results from the

heterodyne beat of the picture and sound carriers. This signal contains the frequency-modulated sound information, which can be further amplified and detected in the sound channel. This is known as the intercarrier sound (ICS) system.

Following the detector is a video amplifier. An output level of about 100 volts is ordinarily sufficient to assure full drive of the picture tube over its modulation range. A 4.5-MHz trap is included in the video amplifier circuit to prevent the appearance of the intercarrier sound signal on the picture tube. The sound reproduction system is discussed below. *See* VIDEO AMPLIFIER.

Automatic gain control. Since television receivers, like radio receivers, may be subjected to widely varying incoming signal strengths, some form of automatic gain control (AGC) is necessary. Circuits for this function provide a nearly constant carrier signal level to the video detector. *See* AUTOMATIC GAIN CONTROL (AGC).

Sync separator circuits. Picture synchronizing information is obtained from the video signal by means of sync separation circuits. In addition, these circuits must separate this information from noise and interference during the reception of weak signals, particularly if impulse noise is present. In general, sync separation circuits perform the following functions: (1) separation, by means of amplitude clipping, of the sync information from the picture information; (2) separation of the desired horizontal and vertical timing information by means of frequency selection; and (3) rejection of noise signals that are higher in amplitude than sync pulses by amplitude limiting or gating (noise suicide) circuits. *See* ELECTRICAL NOISE.

Sweep systems. Two independent systems are employed in the vertical and horizontal sweep circuits. Each employs a timing generator, generally of the oscillatory type, controlled by the synchronizing information obtained from the sync separators. The oscillators are followed by drive and waveform-shaping circuits. These are followed by power amplifier stages capable of providing the currents required by the deflection coils of the yoke for picture-tube beam deflection. Substantially different techniques are required for vertical and horizontal scanning.

Vertical deflection. Generally the vertical oscillator is of the relaxation type operating at approximately 60 Hz. Its frequency is accurately controlled by a signal obtained from the sync separator. The output waveform of the sync separator consists of a train of pulses representing the horizontal and vertical synchronizing pulses. When these are passed through a low-pass filter or integrating circuit, a sawtooth-shaped voltage wave representing vertical sync is obtained. This is used to synchronize the vertical oscillator. A frequency control in the vertical oscillator circuit is so adjusted that its free-running frequency is slightly lower than the synchronizing signal frequency. For good interlace it is necessary that no horizontal frequency components be included in the vertical synchronizing voltage. *See* ELECTRIC FILTER; TELEVISION SCANNING.

The vertical output stage is generally operated as a class A amplifier. The yoke is transformer-coupled to the plate of the output tube to match the yoke impedance to the output tube impedance. Since the yoke impedance is partly resistive and partly inductive, the voltage waveform across it is the sum of a sawtooth and a rectangular pulse. The current through the yoke has essentially a sawtooth waveform, but each sawtooth has a symmetrical S shape to take care of picture-tube faceplate geometry and result in a linear scan. *See* POWER AMPLIFIER; TRANSFORMER.

Horizontal deflection. A more complex system is required for horizontal scanning. There are several basic reasons for this: (1) Horizontal sync pulses are of much shorter duration than are vertical sync pulses; (2) some form of automatic frequency control (AFC) of the horizontal oscillator is required to average the incoming horizontal sync information and retain accurate phase; and (3) considerably greater power output is required to generate the deflecting yoke fields as well as the high voltage, of 10–20 kV, for the picture tube. *See* AUTOMATIC FREQUENCY CONTROL (AFC).

The horizontal oscillator is generally of the Hartley or Colpitts type. The frequency of oscillation is determined both by a time-constant control and by a bias voltage derived from an AFC circuit. The AFC circuit may be a phase comparator, in which the pulses from the sync separator are compared to the oscillator output signal. The output of the comparator is a voltage proportional to the phase departure of the two signals. Many receivers lock the horizontal and vertical scanning waveforms together with a countdown circuit, thus eliminating the vertical oscillator. In this case the countdown begins at a harmonic of the horizontal scan rate.

The desired current waveform in the horizontal windings of the deflection yoke is a line-frequency sawtooth, possibly modified by the addition of a small amount of S curvature to compensate for picture-tube face geometry. Energy from the horizontal drive, or output tube, is normally supplied to the yoke (through the horizontal output transformer) only during approximately the last half of each sawtooth period. At the conclusion of the sawtooth period, the horizontal driver is cut off, and the energy stored in the form of current through the yoke causes an oscillation in the self-resonant circuit consisting of the yoke, horizontal output transformer, and the associated capacitances. This oscillation is permitted to continue for only one half-cycle, during which time the current through the yoke reverses in polarity and attains a negative value almost equal to the original positive value. The self-resonant frequency of the horizontal output circuit must be high enough to permit the full current reversal to be accomplished within the horizontal blanking interval. The oscillation is stopped after the first half-cycle by the action of a damper (normally a diode), which

controls the release of the energy stored in the yoke in such a way that the current follows the desired sawtooth waveform. In approximately the middle of the sawtooth period, the damper becomes nonconductive, and the horizontal driver takes over the task of supplying the energy required for the next cycle.

High-voltage supply. Since the impedance of the yoke at horizontal scan frequency is primarily inductive, the voltage across the horizontal deflection windings is essentially constant during active scan. During the retrace period, however, the high rate of current change causes the generation of a high-voltage pulse having a shape similar to that of a half sine wave and a duration equal to the retrace period. It is common practice to employ a stepup winding on the horizontal output transformer to raise this so-called kickback pulse to a still higher voltage level, which is commonly about 18 kV for black-and-white displays, and to pass it through a simple rectifier and filter to serve as the high-voltage supply for the cathode-ray tube.

Picture tubes. The display device for a monochrome television receiver is a cathode-ray tube, consisting of an evacuated bulb containing an electron gun and a phosphor screen, which emits light when excited by an electron beam. The intensity of the electron beam is controlled by the video signal, which is applied either to the grid or to the cathode of the electron gun. The position of the electron beam is controlled by electromagnetic fields produced by the deflection yoke placed around the neck of the tube. *See* PICTURE TUBE.

Controls. Certain controls are available to the user for adjustment of the receiver. These are the audio volume, channel selector, and brightness and contrast controls. In some receivers fine-tuning and horizontal hold and vertical hold controls are also available. Other controls, normally mounted on the rear of the chassis or under a removable panel, may include height, width, and linearity controls.

The on-off switch for the receiver is frequently mounted on the same shaft as the audio volume control, which controls the gain of the audio channel. The channel selector adjusts the tuner's selective circuits for optimum performance at the desired channel, and fine tuning is a vernier control for the frequency of the local oscillator. Brightness is usually a manual adjustment of the bias on the electron gun in the picture tube. The contrast control adjusts the level of the video signal by some means such as a potentiometer in one of the video amplifier stages.

The horizontal and vertical hold controls adjust the free-running frequencies of the horizontal and vertical oscillators to achieve the most reliable synchronization with the incoming signal. In some cases, the controls may actually consist of variable resistors in the circuits of the respective oscillators.

Vertical linearity is generally controlled by a variable resistance in the circuit of the vertical output stage, and picture height may be controlled by a variable resistor in the circuit of the vertical oscillator. Horizontal linearity may be controlled by a variable inductor placed between the damper and the source of voltage.

Color receivers. Television receivers designed to produce images in full color are necessarily more complex than those designed to produce monochrome images only, because additional information must be handled to produce color. In monochrome systems, the video signal controls only the luminance of the various areas of the image. In color systems, it is necessary to control both the luminance and the chrominance of the picture elements.

The chrominance of a color refers to those attributes which cause it to differ from a neutral (white or gray) color of the same luminance. While chrominance can be expressed in a great variety of ways, it is always necessary to employ at least two variables to express the full range of chrominance that can be perceived by the human eye. In qualitative terms, chrominance may be regarded as those properties of a color that control the psychological sensations of hue and saturation. For color television purposes, chrominance is most frequently expressed quantitatively in terms of the amounts of two hypothetical, zero-luminance primary colors (usually designated I and Q), which must be added to or subtracted from a neutral color of a given luminance to produce the color in question. *See* COLOR.

As a practical matter, color television receivers produce full-color images as additive combinations of red, green, and blue primary-color images, and it is necessary to process the luminance and chrominance information in a color signal in such a way as to make it usable by a practical reproducing device.

Color signal. Color television broadcasts in the United States employ signal specifications that are fully compatible with those used for monochrome, making it possible for color programs to be received on monochrome receivers and monochrome programs to be received on color receivers. (Color pictures are produced, of course, only when color programs are viewed through color receivers—in all other cases, the images are in black and white only.) Compatibility is achieved by encoding the color information at the transmitting end of a color television system in such a way that the transmitted signal consists essentially of a normal monochrome signal (conveying luminance information) supplemented by an additional modulated wave conveying chrominance information. **Figure 2** shows the major components of a color television signal. Although it is added directly to the monochrome signal component before transmission, the color subcarrier signal does not cause objectionable interference, because of the use of the frequency interlace technique. Because the chrominance information involves two variables, the modulated subcarrier signal varies in both amplitude and phase, and it is necessary to employ synchronous detectors to recover the two variables. A phase reference for the special local oscillator, which provides the synchronized carriers in each

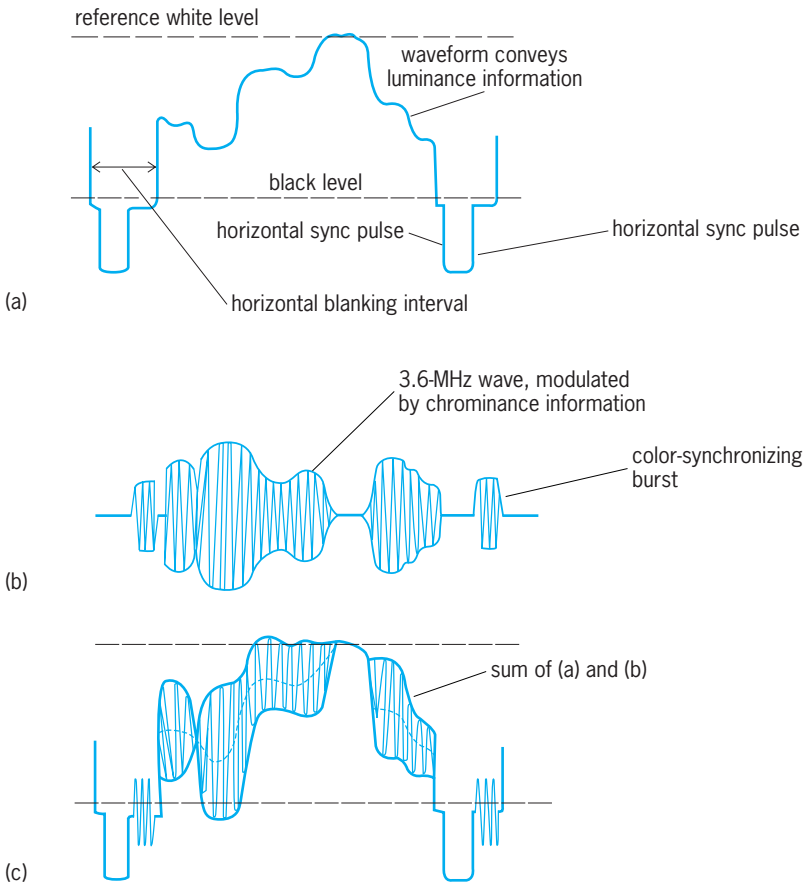

Fig. 2. Waveform sketches of major components of color television signal. (*a*) Normal monochrome signal. (*b*) Color subcarrier signal. (*c*) Complete color signal.

color receiver, is transmitted in the form of so-called color synchronizing bursts. These are short samples of unmodulated subcarrier transmitted during the horizontal blanking periods after the horizontal sync pulses.

Performance standards and requirements. A simplified block diagram for a color television receiver is shown in **Fig. 3**. Many of the circuits in a color receiver are the same in principle as the corresponding circuits in a monochrome receiver, but all circuits handling the complete color signal must be designed for high performance standards. Because the chrominance information is received in the form of sidebands occupying the upper portion of the video spectrum (centered on approximately 3.6 MHz), it is necessary that the antenna, tuner, intermediate-frequency amplifier, and video detector be designed to handle the full 4-MHz bandwidth provided in the broadcast transmission standards if degradation of the color information is to be avoided. Because the color subcarrier signal is simply added to the normal monochrome signal before transmission, it is necessary that all stages handling the complete signal be linear, so as to avoid intermodulation or distortion of the various signal components. The deflection circuits for a color receiver are similar in principle to those used in monochrome receivers, although the output stages are normally designed for a higher power level because of the greater deflection requirements for color cathode-ray tubes. Color displays use 22 to 30 kV for the high-voltage supply.

Color decoding circuits. Special decoding circuits are necessary in a color receiver to process the luminance and chrominance information in a color signal so that it can be used for the control of a practical color cathode-ray tube utilizing red, green, and blue primary colors. The major features of the most common approach to the color decoding circuits are shown within the broken lines in Fig. 3.

The video amplifier shown at the bottom handles the monochrome portion of the signal and is designed to provide attenuation in the vicinity of 3.6 MHz to block the passage of chrominance information. The chrominance information is recovered from the modulated subcarrier signal through a band-pass filter (centered at 3.6 MHz) and a pair of synchronous demodulators, in which the modulated wave is heterodyned against fixed carriers of two different phases but of the same frequency. In the most rigorous type of color decoding circuit, the chrominance components recovered from the modulated subcarrier signal are the same I and Q originally used to produce the modulated wave, but it is possible to use almost any two phase positions (not necessarily 90° apart) to recover any two independent combinations of the original I and Q signals. The bandwidths of the signals produced by the demodulators are normally adjusted somewhere between 0.5 and 1.5 MHz, and delay compensation may be required to keep all three signal components in time coincidence. The matrix circuit is essentially a linear cross-mixing network for combining the M, I, and Q signals in the proper proportions to produce red, green, and blue signals. If signals other than I and Q are produced by the chrominance demodulators, it is necessary only to design the matrix circuit so that it has slightly different mixing constants.

The synchronous carriers required for the demodulation of the chrominance information are provided by a subcarrier regenerator, which is usually a burst-controlled oscillator operating at the subcarrier frequency. Control information for the subcarrier regenerator is obtained from a burst separator, which is a gate circuit turned on only during the horizontal blanking periods by pulses derived from the horizontal deflection system. The separated bursts are compared with the output of the local subcarrier oscillator in a phase detector. If an error exists, a correction voltage is developed, which may be applied to restore the subcarrier oscillator to the proper frequency and phase. For good noise immunity, a time constant is normally provided so that control information is averaged over at least several line periods.

Color cathode-ray tube and convergence circuits. The great majority of color television receivers employ the shadow-mask color cathode-ray tube in which color images are produced in the form of closely intermingled red, green, and blue dots. The primary-color phosphor dots are excited by three separate electron beams, which are prevented from striking dots of the wrong color by the shadowing effect of an

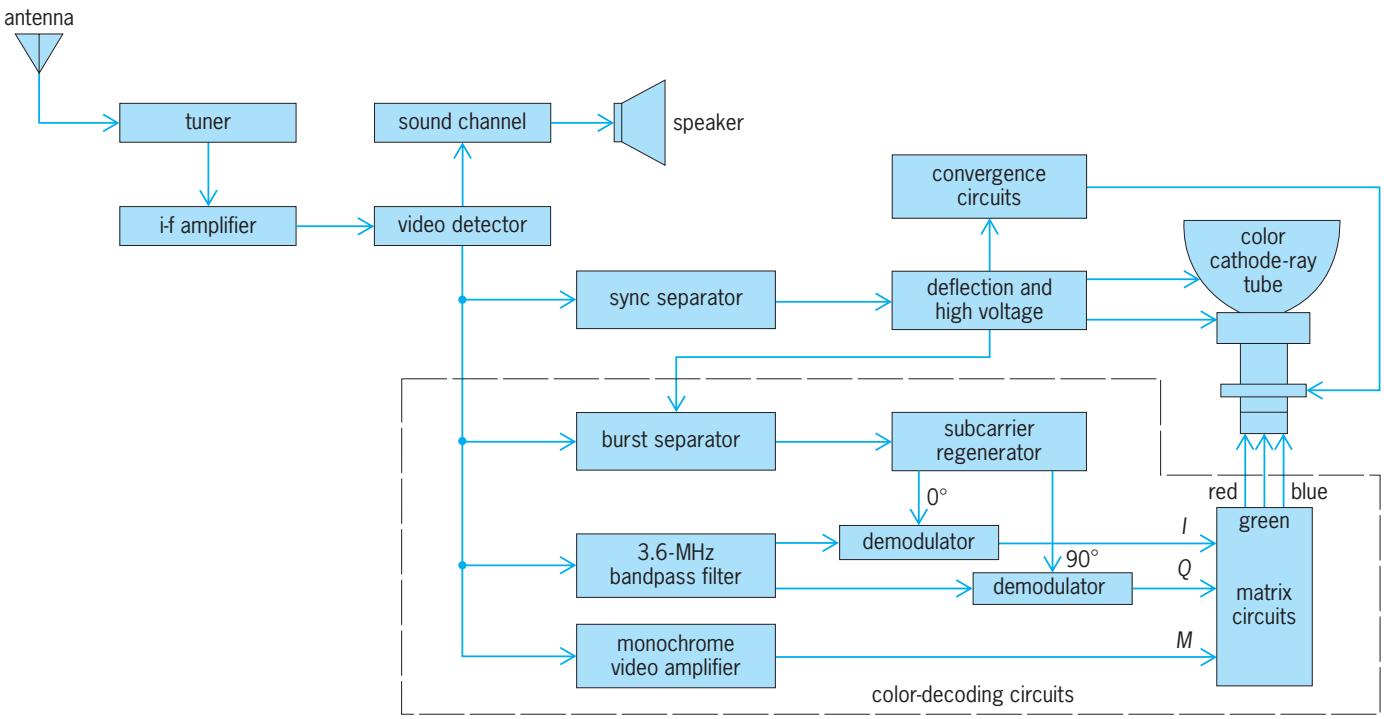

Fig. 3. Simplified block diagram of a color television receiver.

aperture mask located about $^1/_2$ in. (1.25 cm) behind the special phosphor screen. The beams in such a cathode-ray tube are deflected simultaneously by the fields produced by a single deflection yoke placed conventionally around the neck of the tube. New cathode-ray-tube designs and deflection yokes are self-converging and do not require auxiliary convergence deflection.

Color controls. In addition to the same controls required for monochrome receivers, color receivers normally have controls for convergence, hue, and saturation. The convergence controls, considered servicing adjustments only, adjust the relative amplitudes and phases of the signal components that are added together to form the proper waveforms for the convergence yoke. The hue control usually adjusts the phase of the burst-controlled oscillator and alters all the colors in the image in a systematic manner comparable to the effect achieved when a color circle diagram is rotated in one direction or the other. The proper setting for the hue control is normally determined by observing skin tones of persons on the television screen. The saturation control, frequently labeled chroma or simply color, adjusts the gain of the chrominance circuits relative to the monochrome channel and controls the saturation or vividness of the reproduced colors. When this control is set too low, the colors are all pale or pastel, and when it is reduced to zero, the picture is seen in black and white only.

Sound reproduction. The typical circuit for recovery of the sound from the aural transmitter consists of a band-pass filter to extract the frequency-modulated 4.5-MHz carrier from the video detector, amplification and amplitude limiting, frequency-modulation detection with a frequency discriminator, deemphasis with a resistance-capacitance (*RC*) low-pass filter having a time constant of 75 microseconds, followed by volume control, a power amplifier, and a loudspeaker. The result is single-channel monophonic sound. *See* LOUDSPEAKER.

Multichannel television sound. Standards have been adopted for the broadcasting of multichannel television sound (MTS) consisting of a monophonic service, a stereophonic service, and a simultaneous second-audio-program (SAP) service. A frequency spectrum plot of the multichannel sound modulation impressed upon the aural frequency-modulation transmitter is shown in **Fig. 4**. The main channel signal consists of the normal audio-frequency components up to 15 kHz, and in the case of stereophonic transmissions is made up of the arithmetic sum of the left and right audio signals ($L + R$), thus making for a compatible result for the monophonic listener. Stereophonic audio information in the form of the arithmetic difference between the left and right audio signals ($L - R$) is amplitude-compressed in

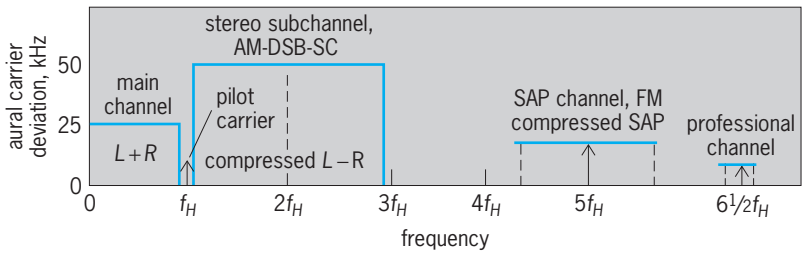

Fig. 4. Composite multichannel television sound signal baseband frequency spectrum. f_H is the horizontal scanning frequency, 15.734 kHz.

accordance with a specific algorithm and impressed as double-sideband, suppressed-carrier, amplitude modulation (AM-DSB-SC) on a carrier at twice the picture horizontal scanning frequency (31.47 kHz). A pilot carrier is sent at the picture horizontal scanning frequency (15.73 kHz) to synchronize the receiver stereo demodulator. Additionally, a second audio program (SAP; a second language, for example) is frequency-modulated on a carrier centered at five times the picture horizontal scanning frequency (78.67 kHz). The SAP audio is monophonic and is compressed with the same algorithm as the stereophonic difference ($L - R$) audio. A professional channel centered at a frequency of about 6.5 times the horizontal scanning frequency may be present and is reserved for use by the broadcaster. The sum of all the components using defined ratios is the composite multichannel baseband signal which frequency-modulates the aural carrier. *See* STEREOPHONIC RADIO TRANSMISSION; STEREOPHONIC SOUND.

Multichannel sound receiver. The MTS receiver design is still evolving, but many receivers make direct use of the 4.5-MHz frequency-modulation intercarrier available at the video detector to feed a tuned amplifier which has a wider bandwidth than that used in the monophonic receiver, a limiter, and a wideband discriminator (much wider than the monophonic equivalent) to reproduce the MTS composite baseband signal as described above. The main channel is deemphasized as in the monophonic case to recover $L + R$ audio. The stereophonic subcarrier at 31.47 kHz is synchronously demodulated with a reference carrier derived from the second harmonic of the pilot at 15.73 kHz to recover the compressed $L - R$ audio. A complementary expander using the inverse of the compression algorithm provides the correct $L - R$ audio signal which is matrixed with the $L + R$ audio to make the original left (L) and right (R) audio signals. A stereophonic amplifier and tone, balance, and volume controls complete the system together with appropriate loudspeakers. The SAP channel requires a separate FM subcarrier demodulator centered at 78.67 kHz to recover the compressed SAP audio. When the SAP audio is chosen for listening, a switch routes the compressed SAP audio to the $L - R$ expander, and disconnects both $L + R$ and $L - R$ audio signals. The compression-expansion (companding) scheme provides an effective way to reduce noise and other interferences encountered in the transmission path between transmitter and receiver, such as when the received signal is weak. The main monophonic channel noise performance is adequate without a companding scheme. *See* SOUND-REPRODUCING SYSTEMS; VOLUME CONTROL SYSTEMS.

Carl G. Eilers

Bibliography. C. G. Eilers, TV multichannel sound: The BTSC system, *IEEE Trans. Consumer Electr.*, CE-31:1-7, 1985; B. Grob and C. Herndon, *Basic Television and Video Systems*, 6th ed., 1998; M. S. Kiver and M. Kaufman, *Television Electronics: Theory and Servicing*, 8th ed., 1983; S. R. Prentiss, *Modern Television: Service and Repair*, 1989; L. B. Tyler, M. F. Davis, and W. A. Allen, A companding system for multichannel TV sound, *IEEE Trans. Consumer Electr.*, CE-30:633–640, 1984.

Television scanning

The process used to convert a three-dimensional image intensity into a one-dimensional television signal waveform. The image information captured by a television camera conveys color intensity (in terms of red, green, and blue primary colors) at each spatial location, with horizontal and vertical coordinates, and at each time instance. Thus, the image intensity is multidimensional, since it involves two spatial dimensions and time. It needs to be converted to a unidimensional signal so that processing, storage, communications, and display can take place.

First, the television scene is sampled many times per second in order to create a sequence of images (called frames). Then, within each frame, sampling is done vertically to create scan lines. Scanning proceeds sequentially, left to right and top to bottom. In a television camera, an electron beam scans across an electrically photosensitive target upon which the image is focused. At the other end of the television chain, with raster scanned displays, an electronic beam scans and lights up the picture elements in proportion to the light intensity. While it is convenient to think of all the samples of a single frame occurring at a single time (similar to the simultaneous exposure of a single frame for film), the scanning in a camera and in a display results in every sample corresponding to a different instance in time, and successive lines occur later in time. *See* TELEVISION CAMERA TUBE; TELEVISION RECEIVER.

Progressive and interlace scanning. There are two types of scanning approaches: progressive (also called sequential) and interlaced. In progressive scanning, the television scene is first sampled in time to create frames, and within each frame all the raster lines are scanned from top to bottom. Therefore, all the vertically adjacent scan lines are also temporally adjacent and are highly correlated even in the presence of rapid motion in the scene. Film can be thought of as naturally progressively scanned, since all the lines were originally exposed simultaneously, so the correlation between adjacent lines is guaranteed. Almost all computer displays (except some low-end computers) are sequentially scanned. *See* ELECTRONIC DISPLAY.

In interlaced scanning, all the odd-numbered lines in the entire frame are scanned first, and then the even numbered lines (see **illus.**). This process produces two distinct images per frame, representing two distinct samples of the image sequence at different points in time. The set of odd-numbered lines constitute the odd field, and the even-numbered lines make up the even field. All current television systems use interlaced scanning. One principal benefit of interlaced scanning is to reduce the scan rate (or the bandwidth). This is done with a relatively high field

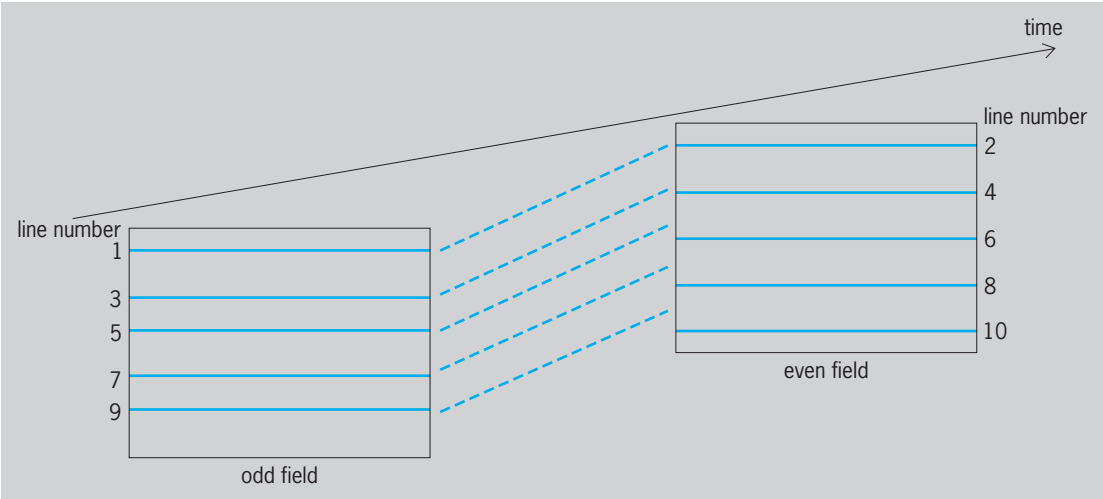

Interlaced scanning. A television frame is divided into an odd field (containing odd-numbered scan lines) and an even field (containing even-numbered scan lines).

rate (a lower field rate would cause flicker), while maintaining a high total number of scan lines in a frame (lower number of lines per frame would reduce resolution on static images). Interlace cleverly preserves the high-detail visual information and, at the same time, avoids visible large-area flicker at the display due to temporal postfiltering by the human eye.

The North American Television Standard (NTSC) has 15,735 scan lines per second or 525 lines per frame, since there are 29.97 frames per second. For each scan line, a small period of time (16–18% of total line time), called blanking or retrace, is allocated to return the scanning beam to the left edge of the next scan line. European systems (PAL and SECAM) have 625 lines per frame, but only 25 frames per second. The larger number of lines results in better vertical resolution, whereas the larger number of frames in NTSC results in better motion rendition. *See* TELEVISION STANDARDS.

High-definition television. While there is no worldwide standard yet, all versions of high-definition television (HDTV) have approximately twice the horizontal and vertical resolution of conventional systems. In addition, high-definition television is digital, where the television scan lines are also sampled horizontally in time and digitized. Such sampling produces an array of pixels (picture elements). High-definition television has an array of approximately 1000 lines and as many as 2000 pixels per line. If the height-to-width ratio of the television raster is equal to the number of scan lines divided by the number of samples per line, the array is referred to as having square pixels; that is, the electron beam is spaced equally in the horizontal and vertical directions. This facilitates digital image processing as well as computer synthesis of images.

An agreement (not a standard) on high-definition formats has been reached and is in use, enabling the transition from analog to digital television. In early 2001, there were more than 150 stations broadcasting high-definition television. The format has 1920 active pixels per line and 1080 active (out of a total of 1125) lines in a frame. The frames may be interlaced or progressive, and the frame rate is 29.97 Hz. Progressive frames at a rate of 23.976 Hz are also permitted to accommodate film material. An alternative progressive-only format that provides additional temporal resolution at the expense of some spatial resolution has also been approved and is currently in use. This format has 1280 active pixels per line and 720 active (out of a total of 750) lines per frame, and the frame rates permitted are 23.976 Hz, 29.97 Hz, or 59.94 Hz. Both high-definition formats have square pixels and a 16 × 9 aspect ratio. *See* DATA COMPRESSION; IMAGE PROCESSING; TELEVISION.

Arun N. Netravali

Bibliography. A. N. Netravali and B. G. Haskell, *Digital Pictures*, 1994; Y. Ninomiya et al., An HDTV broadcasting system utilizing a bandwidth compression technique—MUSE, *IEEE Trans. Broadcasting*, BC-33:130, 1987; C. A. Poynton, *Digital Video and HDTV*: *Pixels*, *Pictures*, *and Perception*, 2001.

Television standards

The accepted criteria for a television system, including the image aspect ratio, number of lines per frame, type of scanning, original video signal bandwidth, transmission format and bandwidth, reception, demodulation, decoding, and sound system. The implementation of high-definition television (HDTV), where the image resolution and audio fidelity are significantly higher than for conventional television, has required new standards. *See* TELEVISION.

Early television. Black-and-white television dates back to the mid-1930s. In the United States the screen has 525 interlace lines in one frame. Within the 525 interlaced lines are 483 active lines of video information. Interlacing implies that the odd-numbered lines are transmitted in the first field in 1/60 second, followed by the even numbered lines in the second field in 1/60 s. Thus there are two fields per frame, and

30 frames/s, or 60 fields/s—just enough to eliminate any discernible flicker in the displayed image.

As an image is scanned and then later displayed, the intensity defines whether the picture element (pixel) is white, when the intensity is low, or black, when the intensity is high, or gray, when the intensity takes on a middle value. This intensity video signal is called the luminance. It takes 63.5 microseconds to traverse one line of an image. This time interval determines the line frequency as $1/(63.5\ \mu s) = 15.75$ kilohertz. The conventional television image has an aspect ratio (ratio of the image width to the image height) of 4:3. Using these television system parameters, the bandwidth of the video signal is approximately 4.2 megahertz. The resolution of a television image is the number of image lines that can be distinguished in the vertical and horizontal directions. The vertical resolution is the number of active lines, 483, although the slight misalignment of the raster scan reduces this by 30% to 340 lines. The horizontal resolution using the video signal bandwidth and the line frequency is approximately 450 lines. *See* BANDWIDTH REQUIREMENTS (COMMUNICATIONS); ELECTRONIC DISPLAY; TELEVISION SCANNING.

A complete television signal is created by including a monaural sound signal above the video signal using frequency modulation (FM) with a sound subcarrier of 4.5 MHz. The bandwidth of the resulting sound signal is approximately 80 kHz. The complete signal is then transmitted using vestigial sideband (VSB) amplitude modulation (AM) with a large carrier. The VSB-AM signal includes the entire upper sideband (4.5 MHz) and 1.25 MHz of the lower sideband. The monochrome television signal occupies a transmission bandwidth of 6 MHz including a frequency guard band of 250 kHz. The television carrier frequencies are assigned by the federal communications commission (FCC) in the very high frequency (VHF) ranges (54–72 MHz, 76–88 MHz, and 174–216 MHz), as well as in the ultrahigh frequency (UHF) range (470–806 MHz). After transmission, the VSB signal is demodulated, and the luminance and sound signals are recovered. *See* AMPLITUDE MODULATION; FREQUENCY MODULATION; SINGLE SIDEBAND; TELEVISION TRANSMITTER.

Color television standards. In the early 1950s, the National Television Systems Committee (NTSC) was formed to set standards for a color television signal (in the United States) that would be fully compatible with the existing monochrome signal. Any color can be formed as a linear combination, that is, a weighted sum, of red (R), green (G), and blue (B). The NTSC standard started with three image signals—R, G, and B—and matrixed these three primary color image signals as linear combinations into one luminance signal (the conventional black-and-white video signal, often called the Y-signal) and two chrominance signals that control hue and saturation. The chrominance signals are termed the in-phase (I) signal and the quadrature (Q) signal. Although the luminance signal retained the original 4.2-MHz bandwidth, the characteristics of human color perception allowed the I-signal to be limited to 1.5 MHz and the Q-signal to only 0.5 MHz.

The power spectrum of the luminance video signal, that is, the distribution of signal power versus frequency, displays distinct lines, or concentrations of power, at multiples of the horizontal line frequency (15.75 kHz). To avoid interference between the luminance signal and the chrominance signals, the color signals are multiplexed with a color subcarrier whose frequency is exactly midway between the 227th and the 228th harmonic of the horizontal line frequency. Using a color subcarrier of 3.58 MHz allows the spectral lines of the chrominance signal to be interleaved between the high-frequency (and low-amplitude) spectral lines of the luminance signal. To retain the original video signal bandwidth, the I-signal is amplitude-modulated with the cosine of the color subcarrier and then filtered to 4.2 MHz, while the Q-signal is amplitude-modulated with the sine of the color subcarrier.

The resulting composite color television signal contains the luminance signal plus the interleaved quadrature amplitude-modulated (QAM) chrominance signals with a 3.58-MHz color subcarrier plus the FM sound signal. A black-and-white television receiver can recover the luminance (Y-signal) and the sound, while a color television can also recover the chrominance (I- and Q-signals). Thus compatibility is achieved. The color television receiver converts the Y-, I-, Q-signals back to R-, G-, B-signals via a decoding matrix, which is just another set of linear combinations, before displaying the color image. Further developments of the sound signal resulted in stereo audio similar to commercial FM radio stations that use a left plus right (L + R) primary signal and a frequency-multiplexed left minus right (L − R) secondary signal. The total audio signal is then applied to the television FM generator. Modern color television receivers recover L + R along with L − R, and produce stereo sound by adding and subtracting the primary (L + R) and the secondary (L − R) audio signals. *See* TELEVISION RECEIVER.

While NTSC color television standards are used in North America, South America, and Japan, the European International Radio Consultative Committee (CCIR) system is used in England, Germany, Italy, and Spain. The color system employed with CCIR television is called Phase Alternate Line (PAL). A modified CCIR television standard is used in France and Russia, where the color system is called SECAM (Sequential Couleur à Mémoire). A list of some of the television standards is given in **Table 1**. These standards use interlaced scanning, so the field rate is twice the frame rate.

HDTV standards. The next generation of television, HDTV, is not compatible with the previous television systems. HDTV relies on digital technology, making it more amenable with computer displays, while taking full advantage of the power and efficiency of digital signal processing (DSP). In 1987, the Federal Communications Commission (FCC) chartered an advisory committee to recommend initial HDTV standards for the United States. This advisory committee

TABLE 1. Broadcast color television standards

Characteristics	NTSC	CCIR (PAL)	CCIR (SECAM)
Video bandwidth, MHz	4.2	5.0	5.0
Channel bandwidth, MHz	6.0	7.0	8.0
Total lines/frame	525	625	625
Frame rate, Hz	30	25	25
Line rate, kHz	15.75	15.625	15.625
Aspect ratio	4:3	4:3	4:3

was transformed into a group of eight organizations called the Grand Alliance (GA) in 1993 as the number of proposed HDTV schemes was reduced to four digital television systems. The desired characteristics of the HDTV system that would replace the NTSC system was expected to have a resolution that would approach the quality of a 35-mm film, that is, approximately twice the horizontal and twice the vertical resolution of conventional television, with a widescreen aspect ratio of 16:9. The target HDTV system standard was required to avoid interlace scanning artifacts, as well as chrominance artifacts and deliver digital multichannel audio. The end result was a 1996 FCC digital television (DTV) standard, and HDTV broadcasting commenced in the United States in 1998. The successor to the Grand Alliance is the Advanced Television Systems Committee (ATSC), an international organization that has established voluntary technical standards and recommended practices for HDTV.

The most challenging aspect of the DTV standard was to achieve the desired resolution with digital video and digital audio, while maintaining a transmission channel bandwidth of 6 MHz for terrestrial broadcast channels. This was achieved by using powerful digital compression algorithms to reduce the raw data rate from 1.5 Gbps (gigabits per second) to approximately 20 Mbps (megabits per second). The United States DTV standard that was adopted actually consists of multiple standards with different resolutions and frame rates. Details of the two HDTV formats are provided below along with the data compression and transmission specifications. *See* DATA COMPRESSION.

The transmitter of an HDTV system consists of a set of operations. It starts with the scanned luminance and chrominance video signals, and the wideband, multichannel audio signals, and ends with the radio-frequency (RF) signal that is broadcast to HDTV receivers. The main HDTV transmitter operations are video/audio coding and compression, data multiplexing into packets, data scrambling, channel coding (for error detection and correction), synchronization multiplexing, and digital modulation (for broadcast transmission). The HDTV receiver reverses the operations of the transmitter.

The two video formats included in the DTV standard use a 16:9 aspect ratio, but the horizontal and vertical resolutions differ, and several scanning frame rates are allowed. These are given in **Table 2**, where "active lines/frame" indicates the vertical resolution, and "pixels/scan line" designates the horizontal resolution. Progressive scanning is simply traversing each line of the image from top to bottom to complete one frame.

The HDTV video signal bandwidth specification is 24.9 MHz, and the audio signal bandwidth is 20 kHz, where five such audio signals can be encoded to provide multichannel Dolby digital surround sound. The necessary data compression is achieved for the video signal by employing a Motion Picture Experts Group (MPEG) encoding technique, called MPEG-2. This technique blocks the image into groups of 8×8 pixels within each frame, and executes a discrete cosine transform (DCT) on the 64 pixels resulting in a new set of coefficients, many of which are insignificant and need not be transmitted. In addition, frame-to-frame redundancy is removed by motion estimation and motion compensation via predictive coders.

Data multiplexing combines the digitized, compressed video and audio with ancillary data for such things as closed captions, resulting in transport packets that are scrambled to ensure random binary data. The data rate, with the synchronization signal included, is 19.39 Mbps. Channel coding consists of a Reed-Solomon (R-S) forward error correcting (FEC) code that increases the data rate to 21.5 Mbps, and rate 2/3 trellis coded modulation (TCM) that further increases the data rate by 50% to 32.3 Mbps. The inclusion of these channel coding operations guarantees that errors can be detected and corrected by the HDTV receiver. *See* DATA COMMUNICATIONS; INFORMATION THEORY; MULTIPLEXING AND MULTIPLE ACCESS.

To ensure that the broadcast signal fits into the specified 6-MHz transmission bandwidth, groups of 3 bits are converted to one of 8 levels, resulting in a digital signal that has a symbol or baud rate that is one-third of the 32.3 Mbps data rate, that is, 10.76 Mbps. This 8-level baseband signal is symbol-shaped to avoid intersymbol interference (ISI) using a sine-rolloff filter with a rolloff factor, $r = 0.0575$. The resulting baseband signal is amplitude-modulated using 8-level VSB, with a pilot

TABLE 2. HDTV image formats

Characteristics	HDTV 1	HDTV 2
Active lines/frame	1080	720
Pixels/scan line	1920	1280
Frame rate (Hz) and scan type	24, progressive 30, progressive 30, interlace	24, progressive 30, progressive 60, progressive

carrier signal, whose shaped upper sideband (5.69 MHz) is included, along with 5.4% of the lower sideband (0.31 MHz), such that the 6-MHz transmission bandwidth is maintained. *See* AMPLITUDE MODULATION; MODULATION.

In Europe, a different transmission standard has been adopted, in lieu of 8-VSB for digital television. The European system is referred to as the Digital Video Broadcasting (DVB) standard. It uses coded orthogonal frequency division multiplexing (COFDM) as a modulation scheme.

The FCC expects everyone to be using new HDTV receivers by the year 2011, at which time NTSC broadcasting will cease, and all NTSC color television receivers will need to be replaced or modified with some type of converter to be able to decode and display HDTV images. Joseph L. LoCicero

Bibliography. *Guide to the Use of ATSC Digital Television Standard*, Advanced Television Systems Committee, Washington, DC, Document A/54, October 4, 1995; G. N. Patchett, *Color TV Systems*, 1969; J. C. Whitaker, *HDTV: The Revolution in Digital Video*, 1999.

Television studio

A facility designed for the production of television programs, which may be broadcast live concurrently with the production or recorded for later broadcast. A television studio consists of the studio room, wherein the actual program takes place, and various support rooms, which include the control room, the equipment room, and the property room.

Studio room. The studio room is where the program action occurs and is analogous to a theatrical stage (**Fig. 1**). Studio rooms may be of almost any size, depending on use, but invariably provide certain facilities, such as a flexible lighting and scenery system, one or more cameras, one or more microphones for sound pickup, and a communications system to allow coordination during the program. The larger studios may provide seating for an audience.

Fig. 1. Television studio room. (*KDVR, Fox-31, Denver, CO*)

Most studios use a lighting grid suspended from a high ceiling that allows flexible placement of the various lighting fixtures. The grid is a steel framework that has electrical power outlets at regular intervals throughout its structure and allows the fixtures to be moved into any desired position. Portions of the grid may be lowered, or a catwalk may be provided for adjustments. Cabling to a lighting switch panel connects each electrical outlet. Jumper cables or multiple-pole switches are employed to assign each outlet to a particular dimmer circuit. Several outlets may be assigned to each dimmer, so that the lighting operator may control a number of lights with one control. Each dimmer control circuit may vary the brightness of its associated lights from full off to full on, usually by means of silicon controlled rectifiers grouped together in a convenient location and remotely controlled from the lighting control panel. Each fixture on the lighting grid may be equipped with various accessories, such as "barn-doors" (deflectors), scrims (diffusers), and gels (color filters), to provide the effects desired by the lighting director. *See* SEMICONDUCTOR RECTIFIER.

Automated lighting systems employ computers which have the ability to store in memory the various lighting calls used in a production. Once the desired sequence is programmed during rehearsal, it may be recreated flawlessly during the actual performance. An added benefit is that complicated lighting changes, requiring quick or simultaneous actions, are easily within the capability of the computer. *See* COMPUTER.

Studios may be equipped with various forms of scenery or sets. Often a fixed set may be used repeatedly for a series production or news programs where a familiar setting is required. Other productions may call for rapid changes of scenery, and the studio ceiling grid is used for hoisting undesired scenic elements out of camera view. The cyclorama, a scenic device often found in the studio, is essentially a curved wall which joins the floor so that no distinct corners are formed, the on-camera effect of which is the illusion of infinite space. This allows many visual effects to be created by use of special lighting. Powerful computers now allow the creation of virtual sets, which are combined with camera video to create the illusion of a large and complex physical set. Merely changing the computer program can redesign the set. News programs take particular advantage of this technology to enhance the viewing experience.

The well-designed television studio makes provision for plugging cameras, microphones, and intercom headsets into wall outlets placed in appropriate locations. It provides monitoring facilities for both video and audio, and has a studio announcing system so that the director can simultaneously address the entire production team from the control room.

Special sound-absorbing materials usually line the walls and ceilings to prevent unwanted acoustic reflections within the studio and to keep outside noises from intruding. The studio floor must be level and smooth, so that camera and microphone boom dollies can be rolled about the room noiselessly and without creating any jiggle in the camera's picture. The entire studio room must be provided with an air-conditioning system capable of noiselessly removing the considerable heat given off by the lighting system. *See* MICROPHONE; TELEVISION CAMERA.

Control room. The studio control room is the nerve center of the television production facility (**Fig. 2**). The room is often at the rear of the studio room in an acoustically isolated booth. Visual contact between the control room and the studio may be provided through soundproof windows, but this is not necessary. The program director, technical director, and sound engineer each have a station within the control room equipped with intercom headsets so that they can communicate quickly and efficiently.

The control room usually has a bank of video monitors with screens which display the output of each camera, videotape recorder, or special-effects generator, as well as a previous monitor which shows the director what the next shot will look like and a line monitor which shows the scene currently on the air. The technical director sits at a video switcher which allows him or her to put various cameras or other video sources on the air in response to orders from the program director. The video switcher is a device which can provide a variety of transitions between different sources, such as dissolves, fades, cuts, and wipes. It can also provide chroma key effects which can make the studio appear to be part of the prerecorded background scene. Modern switchers may incorporate computers to assist the technical director in complicated transitions or special effects.

The sound engineer usually sits behind the technical director, often in a sound-isolated booth. The sound engineer operates an audio mixer console which has every microphone used in the studio connected to a separate input. Other sources, such as turntables, audio tape recorders, tape cartridge machines, compact disk players and audio hard disk recorders (or audio servers), may also be connected to the audio mixer. The sound engineer controls the levels in accordance with the program director's instructions and ensures that program dialog can be heard above background sound effects. Equipment is often available that can alter the sound passing through it by means of equalization, echo chambers, and a wide variety of digital effects. The sound engineer is responsible for keeping the overall audio level within a standard range. *See* SOUND-RECORDING STUDIO.

Equipment room. The equipment room may be located some distance from the studio and control rooms. It contains the electronic equipment associated with the cameras, the video switcher, and special-effects generators. It may contain such equipment for several studios in large facilities. Video engineers operate the camera control units and other equipment by using waveform monitors and vectorscopes to ensure that video levels are correct. Therefore the technical director in the control room may select various video sources without large deviations in signal level or picture quality appearing on the air. The equipment room may also provide video- and audio-processing equipment to make final corrections to the picture and sound before the program is sent to a tape recorder, video hard disk recorder (also called a video server), the station transmitter, or a network of stations.

Fig. 2. Studio control room. (*KDVR, Fox-31, Denver, CO*)

Property room. The property (or prop) room is usually located adjacent to the studio. It is generally a large open area with shelves and bins suitable for storage of the scenic elements and props, such as chairs and tables, used in the production of television shows.

Earl F. Arbuckle, III

Bibliography. T. D. Burrows et al., *Television Production*, 6th ed., 1995; W. E. McCavitt, *Television Studio Operations*, 1980; National Association of Broadcasters, *NAB Engineering Handbook*, 8th ed., 1992; P. B. Seel and A. E. Grant, *Broadcast Technology Update: Production and Transmission*, 1997; J. Watkinson, *An Introduction to Digital Video*, 2d ed., 2001; J. C. Whitaker and K. B. Benson (eds.), *Standard Handbook of Video and Television Engineering*, 3d ed., 2000; A. Wurtzel et al, *Television Production*, 4th ed., 1995; H. Zettl, *Television Production Handbook*, 7th ed., 1999.

Television transmitter

An electronic device that converts audio and video signals into modulated radio-frequency (rf) energy which can be radiated from an antenna and received by a television receiver. The term can also refer to the entire television transmitting plant, consisting of the transmitter proper, associated visual and aural input

and monitoring equipment, transmission line, the antenna with its tower or other support structure, and the building in which the equipment is housed. In the United States, both analog NTSC (National Television Systems Committee) and digital 8-VSB transmitters are in service. The digital transmitters are used for what is termed high-definition television (HDTV).

An analog television transmitter can be thought of as two separate transmitters integrated into a common cabinet (**Fig. 1***a*). Video information is transmitted via a visual transmitter, while audio information is transmitted via an aural transmitter. Because video and audio have different characteristics, the two transmitters differ in terms of bandwidth, modulation technique, and output power level. Nevertheless, a common transmitting antenna is generally used, and the two transmitters feed this antenna via an rf diplexer or combiner. Some modern analog UHF transmitters actually combine the visual and aural signals at a very low level and then employ common high-power amplification systems. This requires careful engineering to avoid significant intermodulation products.

Fig. 1. Television transmitters. (*a*) Analog solid-state transmitter (*Harris Broadcast*). (*b*) Digital inductive-output tube (IOT) transmitter (*Thomcast Communications*).

A digital transmitter (Fig. 1*b*) accepts a single encoded digital bit stream that may contain video, audio, and data. In the United States, the digital terrestrial transmission standard is known as 8-VSB, which is an eight-level, vestigial sideband format. The FCC has mandated that all U.S. television stations convert to digital and terminate analog transmissions by 2006, although many view this as an optimistic timetable for the changeover.

Television stations are licensed to operate on a particular channel, but since it takes a very wide bandwidth to transmit a television picture, these channels are allocated over a broad range of frequencies. Channels 2 through 6 are low-band very-high-frequency (VHF) channels, while channels 7 through 13 are high-band VHF channels. Channels 14 through 69 are ultrahigh-frequency (UHF) channels. Each channel is 6 MHz wide. Because of the wide range of frequencies, television transmitters are designed to work in only one of the foregoing groups, and employ specific circuits which are most efficient for the channels involved. Nevertheless, every television transmitter, regardless of operating frequency, transmits a standard television signal in conformity with the regulations of the country in which it is operated. *See* RADIO SPECTRUM ALLOCATION.

Signal characteristics. In the United States the Federal Communications Commission (FCC) specifies the standard television signal in its rules This specification enables receiver manufacturers to market receivers that are compatible with all television transmitters.

For the NTSC standard, the FCC requires that a visual transmitter produce an amplitude-modulated (AM) carrier with an upper sideband extending to 4.2 MHz above the carrier and a lower sideband extending to only 0.75 MHz below the carrier. The lower sideband is restricted to this narrow bandwidth in order to conserve valuable frequency spectrum. Since both sidebands contain the same information, only one is required to transmit a picture. This is known as vestigial sideband transmission (**Fig. 2**). *See* AMPLITUDE MODULATION; SINGLE SIDEBAND.

FCC rules provide that the aural transmitter be frequency-modulated (FM) and that its carrier frequency must be 4.5 MHz above the visual carrier. This standard spacing between carriers allows use of a simplified sound receiver in most television sets. In the aural transmitter 100% modulation is defined as equal to 25 kilohertz deviation, and the transmitter must be capable of faithfully passing audio-modulating frequencies from at least 50 Hz to 15 kHz. *See* FREQUENCY MODULATION; TELEVISION STANDARDS.

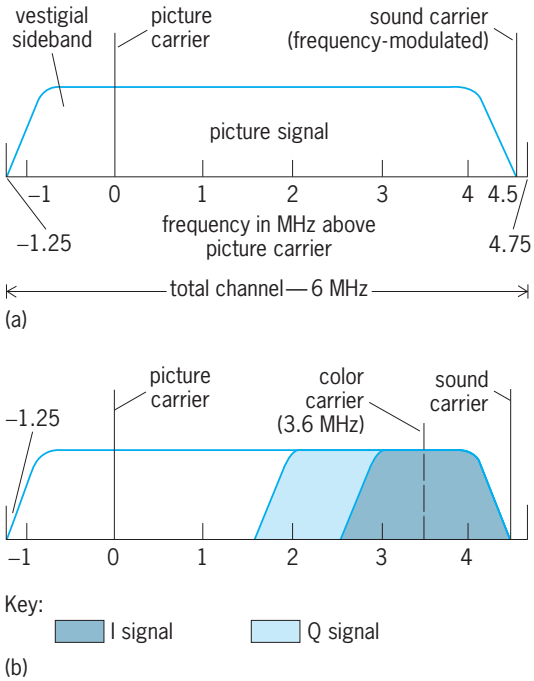

Fig. 2. Diagram of television channel showing portions occupied by color and monochrome signal components. (*a*) Monochrome television channel. (*b*) Color television channel. (*After G. W. Bartlett, ed., NAB Engineering Handbook, 6th ed., National Association of Broadcasters, 1975*)

Transmitter power. Television broadcast stations are limited to a specific effective radiated power (ERP) by the FCC. ERP is defined as the transmitter output power multiplied by transmitting antenna gain and an efficiency factor (less than 100%) due to the losses of the transmission line components between the transmitter and antenna.

To provide a consistent signal strength at the receiver, the FCC allows UHF stations to operate with more power (a maximum of 5 MW peak visual ERP in analog service, or 1 MW average ERP in digital service) than high-band VHF stations, and high-band VHF stations may operate with greater power (316 kW maximum) than low-band VHF stations (100 kW maximum). The maximum power that any station may utilize is reduced proportionately if its antenna height exceeds 1000 ft (305 m) above average terrain. Furthermore, due to its narrower bandwidth and other factors, the aural FM signal tends to carry better than the visual signal, so it is restricted to between 10 and 20% of the visual power by the FCC.

Because transmitting antennas usually have a power gain greater than unity, television transmitters need only provide a fraction of the ERP. Modern transmitters are rated at 10 W to 30 kW for the low-band VHF, and 10 W to 75 kW for high-band VHF; UHF transmitters are manufactured with outputs ranging from 100 W to more than 300 kW.

The lowest-power transmitters are generally configured as television translators, which are used to relay the signal of a high-power primary station into areas where terrain or other factors prevent viewers from being able to receive the primary station. The translator accepts the input on the primary channel and shifts its output to another channel so that one does not interfere with the other.

Transmitting antennas. Antenna gain is a function of design and the number of sections employed. Gain usually increases with physical size, due to increased radiating area, and may be increased further by stacking identical elements vertically; VHF stations usually operate with antenna gains of less than 5, while UHF stations commonly use gains of up to 50 in order to generate the much greater effective radiated power allowed by the FCC. The same antenna may be used for analog or digital transmission, sometimes simultaneously as in the case of multichannel, broadband antennas.

The horizontal radiation pattern of most television transmitting antennas is circular, providing equal radiated signal strength to all points of the compass. Higher-gain antennas achieve greater power in the direction of the horizon by reducing the power radiated at vertical angles above and below the horizon. Since this could result in weaker signals at some receivers close to the transmitter, beam tilt and null

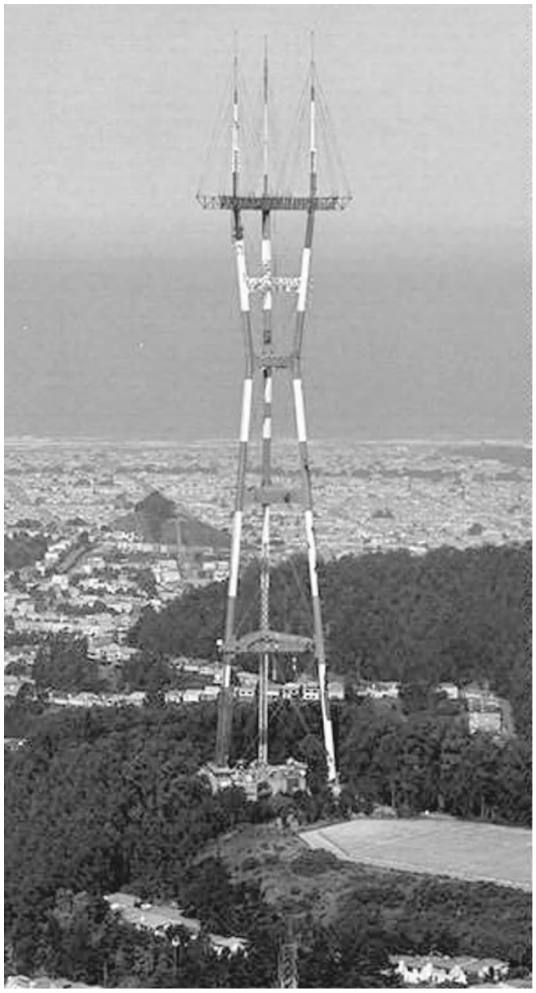

Fig. 3. Sutro Tower, with most of the television transmitters in San Francisco. (*KGO-TV, San Francisco*)

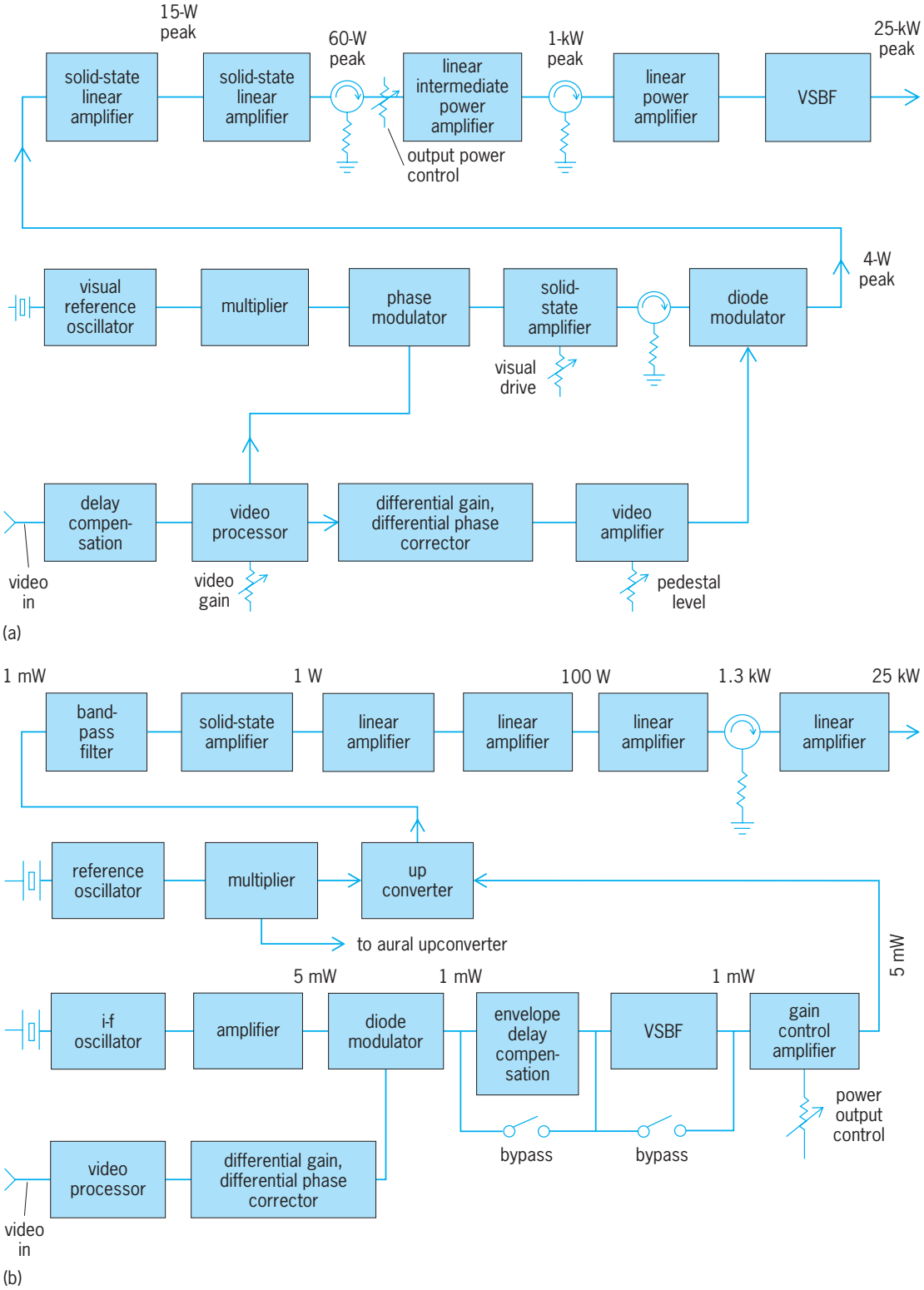

Fig. 4. Visual transmitters with modulation and vestigial-sideband filter (VSBF). (*a*) Transmitter with filtering of high-power signal at final frequency. (*b*) Transmitter with filtering of low-level, intermediate-frequency signal. (*After G. W. Bartlett, ed., NAB Engineering Handbook, 6th ed., National Association of Broadcasters, 1975*)

fill are often used to lower the angle of maximum radiated power.

Until about 1978, all television stations employed horizontally polarized antennas so that their signals would suffer less interference from impulse noise sources, such as automobile ignition systems, which tend to be vertically polarized. Some broadcasters then began utilizing circularly polarized antennas in the belief that television sets with "rabbit-ear" receiving antennas would obtain a better signal.

Because television signals travel in a "line of sight," transmitting antennas are usually placed as high as possible above ground with respect to the surrounding service area. Such locations minimize signal blockage or ghosting due to tall buildings and hills. It is also desirable to locate all of the transmitting antennas serving a given locality in the same place. This allows viewers to orient their receiving antennas in one direction for the best reception from all of the stations. In Los Angeles, for example, advantage is taken of a 5700-ft (1740-m) summit, Mount Wilson, to locate most of the transmitting antennas on individual 200- to 500-ft (60- to 150-m) towers. In New York City all of the transmitting antennas were mounted on a common 365-ft (111-m) mast atop the 1366-ft (416-m) World Trade Center's North Tower before the building was destroyed in a terrorist attack on September 11, 2001. In San Francisco, most of the television transmitters are located on the 977-ft (300-m) Sutro Tower (**Fig. 3**). In areas with neither mountains nor tall buildings, extremely tall towers have been erected, some of which exceed 2000 ft (600 m). Where many broadcasters are colocated, the impact of radio-frequency radiation (RFR) has become an issue. Federal guidelines require that human exposure to such radiation be limited to specific levels. Controversy exists as to the real impact of such exposure. *See* ANTENNA (ELECTROMAGNETISM).

Transmitter designs. There are two broad classes of VHF analog visual television transmitter design philosophy. The classical approach modulates the visual carrier at a moderate power level, amplifies the carrier to rated output power by means of high-power linear amplifiers, and then filters this high-power carrier to obtain the required vestigial-sideband signal (**Fig. 4***a*). The more contemporary approach, used by nearly all transmitter manufacturers, employs modulation at a very low power level of an intermediate-frequency (i-f) signal. The required vestigial-sideband filtering is imposed on this low-level signal, generally by means of a highly stable surface-acoustic-wave filter, whereupon the signal is upconverted to the carrier frequency and amplified by linear amplifiers to rated output power (Fig. 4*b*). *See* AMPLIFIER; SURFACE-ACOUSTIC-WAVE DEVICES.

Generally, modern analog and digital UHF transmitters employ large power tubes, called inductive output tubes (IOTs), to produce the large amounts of rf carrier power required. **Figure 5** shows such a tube in its circuit assembly.

Every analog transmitter contains a visual and aural exciter. This element determines the operating frequency of the visual and aural carriers and must be extremely stable, since the FCC requires that they be kept within 1 kHz of the assigned frequencies. The aural exciter produces the required frequency modulation by varying the frequency of the rf carrier oscillator at an audio rate (**Fig. 6**). Digital transmitters handle audio as simply an indistinguishable component of the digital bit stream applied to the digital exciter.

Regardless of the visual modulation approach taken, certain parameters must be kept within FCC tolerances. Suitable circuits must compensate nonlinearity of the rf power amplifier stages. Certain time delays, within the transmitter, called envelope delays, must be the complement of those found in the home television receiver, so that colors in the picture are properly superimposed. Flat frequency response over adequate bandwidth is necessary to ensure that the picture has good detail and accurate color rendition. The FCC, however, does not permit excessively wide response, so that a low-pass filter that attenuates video above 4.75 MHz by at least 20 dB is commonly inserted in the video input circuit. *See* ELECTRIC FILTER.

Analog transmission of color requires that the color subcarrier (3.58 MHz above the visual carrier) not be affected by changes in the luminance level of the picture. A change in color saturation (chrominance level) brought about by a change in luminance level is termed differential gain. A change in color hue (subcarrier phase) brought about by a change in luminance level is termed differential phase. Both parameters are to be minimized.

The FCC also strictly regulates the radiated rf carrier. Harmonics, or multiples, of the carrier frequency must be attenuated at least 60 dB below the peak power level. Harmonic filters are placed on the output of both the visual and aural transmitters, ahead of any aural-visual diplexers, to ensure

Fig. 5. Ultrahigh-frequency (UHF) inductive output tube (IOT) in circuit assembly. (*Marconi Applied Technology*)

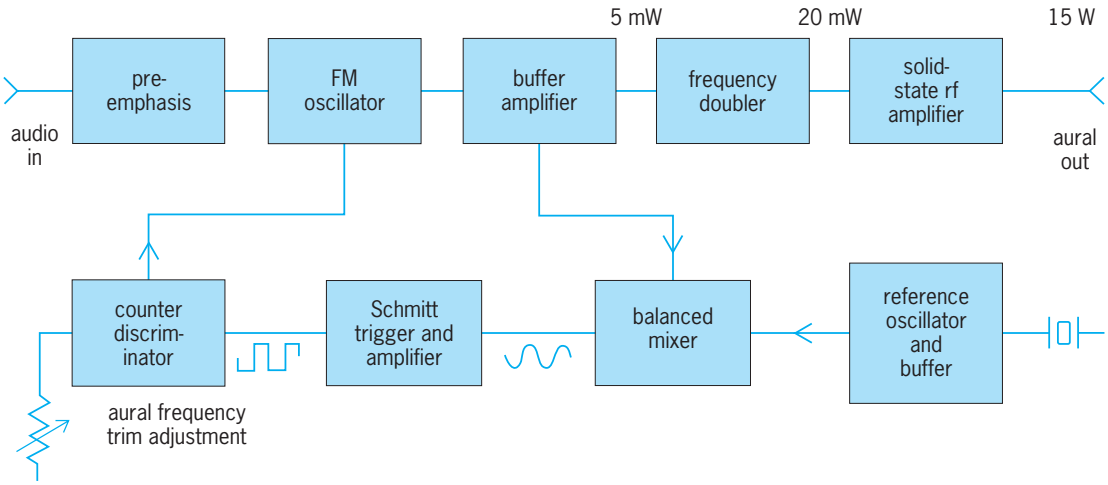

Fig. 6. Typical aural exciter. (*After G. W. Bartlett, ed., NAB Engineering Handbook, 6th ed., National Association of Broadcasters, 1975*)

compliance. In those transmitters which employ high-level vestigial-sideband filtering, a device known as a filterplexer often combines the functions of sideband filtering and aural-visual diplexing. Such filters are also used on the output of digital transmiters. *See* TELEVISION; TELEVISION CAMERA; TELEVISION CAMERA TUBE; TELEVISION NETWORKS; TELEVISION RECEIVER; TELEVISION SCANNING; TELEVISION STUDIO. Earl F. Arbuckle, III

Bibliography. R. Blair, *Digital Techniques in Broadcasting Transmission*, 2d ed., 2002; D. Christiansen (ed.), *Electronics Engineers' Handbook*, 4th ed., 1997; P. Dambacher, *Digital Terrestrial Television Broadcasting: Designs, Systems, and Operation*, 1998; P. Hodges, *An Introduction to Video Measurement*, 2d ed., 2001; P. B. Seel and A. E. Grant, *Broadcast Technology Update: Production and Transmission*, 1997; M. Silbergleid and M. J. Pescatore, *The Guide to Digital Television*, 3d ed., 2000; J. C. Whitaker and K. B. Benson (eds.), *Standard Handbook of Video and Television Engineering*, 4th ed., 2003; E. Williams et al., *National Association of Broadcasters Engineering Handbook*, 10th ed., Elsevier, 2007.

Tellurium

A chemical element, Te, atomic number 52, and atomic weight 127.60. There are eight stable isotopes of natural tellurium. Tellurium makes up approximately 10^{-9} % of the Earth's igneous rock. It is found as the free element, sometimes associated with selenium. It is more often found as the telluride sylvanite (graphic tellurium), nagyagite (black tellurium), hessite, tetradymite, altaite, coloradoite, and other silver-gold tellurides, as well as the oxide, tellurium ocher. *See* PERIODIC TABLE.

There are two important allotropic modifications of elemental tellurium, the crystalline and the amorphous forms. The crystalline form has a silver-white color and metallic appearance. This form melts at 841.6°F (449.8°C) and boils at 2534°F (1390°C). It has a specific gravity of 6.25, and a hardness of 2.5 on Mohs scale. The amorphous form (brown) has a specific gravity of 6.015. Tellurium burns in air with a blue flame, forming tellurium dioxide, TeO_2. It reacts with halogens, but not sulfur or selenium, and forms, among other products, both the dinegative telluride anion (Te^{2-}), which resembles selenide, and the tetrapositive tellurium cation (Te^{4+}) which resembles platinum(IV).

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Tellurium is used primarily as an additive to steel to increase its ductility, as a brightener in electroplating baths, as an additive to catalysts for the cracking of petroleum, as a coloring material for glasses, and as an additive to lead to increase its strength and corrosion resistance. *See* SELENIUM. Stanley Kirschner

Bibliography. P. W. Atkins et al., *Inorganic Chemistry*, 4th ed., 2006; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; D. Sangeeta, *Inorganic Materials Chemistry Desk Reference*, 1997.

Telosporea

A class of the subphylum Sporozoa. These protozoa are divided into two subclasses, the Gregarinia and Coccidia. All members of the group are either intra- or extracellular parasites, and the life cycles have

both sexual and asexual phases. The spores lack a polar capsule and develop from an oocyst. The sporozoite is the usual infective stage which initiates the asexual phase in the life cycle. *See* COCCIDIA; GREGARINIA; PROTOZOA; SPOROZOA.

Elery R. Becker; Norman Levine

Temnopleuroida

An order of Echinacea with a camarodont lantern, smooth or sculptured test, tubercles imperforate or perforate (and usually crenulate), ambulacral plates of diademoid or echinoid type, and branchial slits which are usually shallow. The order includes a long phylogenetic series in which the original characters change considerably so that a concise but exact diagnosis is not possible. Following is an evolutionary summary of the three included families. (1) The Glyphocyphidae, known only from the Cretaceous and Eocene and probably ancestral to the other two families, were small forms with a sculptured test, perforate crenulate tubercles, and diademoid ambulacral plates. Their sculptured test links them with (2) the Temnopleuridae whose tubercles, however, are imperforate, though usually crenulate. This family arose in the Cretaceous and abounds today, especially in the tropics, on strandlines. Most of the order, so far, had shallow branchial slits, but transitional forms link them with (3) the Toxopneustidae, Tertiary and extant forms where the slits are deep and the sculpture tends to vanish. At the same time, the tubercles become imperforate and noncrenulate and the ambulacral plates change to the echinoid type. *See* ECHINACEA; ECHINODERMATA; ECHINOIDA; ECHINOIDEA.

Howard B. Fell

Temnospondyli

One of the largest recognized groups of early amphibians with about 180 described genera. The Temnospondyli first appear in the fossil record in the lower Carboniferous and reach substantial diversity in the later Carboniferous and Permian. The end-Permian extinction event wiped out several families, but the surviving lineages rediversified and temnospondyls remained numerous through the Triassic before declining in the later Mesozoic, the last survivor (*Koolasuchus*) being known from the mid-Cretaceous of Australia. *See* GEOLOGIC TIME SCALE.

General morphology. Earlier temnospondyls were superficially salamander-like with large flat heads, no necks, short limbs, and an undulating gait. In the context of other early amphibians, they are recognized by the large openings across the roof of the mouth (palatal vacuities) and by the semicircular embayments behind the cheek region for froglike large eardrums. They had short straight ribs which combined with the large skulls, suggest that they breathed by buccal pumping, ramming air into the lungs, and not using costal respiration. The individual vertebrae were of multipart construction with the body (centrum) made up of three separate bones, namely one large intercentrum anteriorly and two smaller pleurocentra posteriorly. The skull was massively constructed and survives intact in the fossil record even when the rest of the skeleton is lost.

Types and diversity. Carboniferous and Permian temnospondyls include a wide range of morphological types. As well as amphibious salamander-like forms, they evolved into larger superficially crocodile-like animals both amphibious and terrestrial (*Eryops*) [see **illustration**]. Others were specialized aquatic forms, retaining lateral-line canals to a large size and filling eel-like niches. Most were 30 cm (1 ft) to 1 m (3.3 ft) in length, with a few forms growing to 2 m (6 ft). Most of the Carboniferous and early Permian records are from North America and Europe (equatorial Euramerica in the late Paleozoic), but from the mid-Permian onward the group is found in all continents.

After the end-Permian extinction event, one subgroup of temnospondyls, the Stereospondyli, produced a new adaptive radiation of crocodile-like forms, mostly amphibious or aquatic and growing to 5 m (16.5 ft) [*Mastodonsaurus*]. In some of these

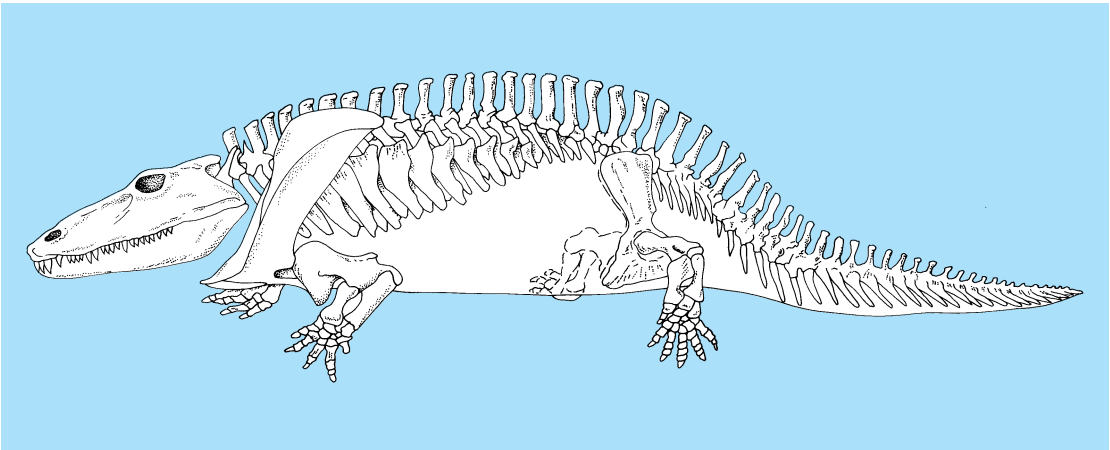

Common Permian temnospondyl *Eryops*; length is approximately 6 ft (2 m). (*After American Museum of Natural History*)

families, the vertebrae become further modified, with the main body of the vertebrae being reduced to a single large intercentrum. Many of the families forming this radiation were global in distribution. They include the Triassic trematosaurids, which are sometimes found in marine assemblages. At the end of the Triassic most temnospondyls vanish from the fossil record, but a few aquatic forms persisted in East Asia and Australia through the Jurassic and Early Cretaceous.

Relation to modern amphibians. The origin of the modern amphibian groups is controversial, but many workers consider the Temnospondyli to be the group from which they evolved. Two Carboniferous-Permian families—the Amphibamidae and the Branchiosauridae—comprise small temnospondyls which show resemblances to frogs and salamanders. The lower Permian amphibamid *Doleserpeton* has pedicellate bicuspid teeth like those of modern amphibians, and the branchiosaurid *Apateon* shows a pattern of skull bone development very like that in modern salamanders. *See* AMPHIBIA; ANTHRACOSAURIA; ICHTHYOSTEGA; LABYRINTHODONTIA; LISSAMPHIBIA. Andrew Milner

Bibliography. J. A. Clack, *Gaining Ground*, Cambridge University Press, 2002; H. Heatwole and R. L. Carroll (eds.), *Amphibian Biology*, vol. 4, Surrey Beatty, Chipping Norton, NSW, Australia, 2000.

Temperature

A concept related to the flow of heat from one object or region of space to another. The term refers not only to the senses of hot and cold but to numerical scales and thermometers as well. Fundamental to the concept are the absolute scale and absolute zero and the relation of absolute temperatures to atomic and molecular motions.

Numbers for temperatures, such as 100°C and −15°F, have been used for only about 300 years. By the seventeenth century, science had developed to the point that, to fully describe the properties of matter, a numerical, quantitative scale of temperature differences was needed. For example, in 1756 Joseph Black in Scotland discovered that ice does not change temperature when it melts. Almost all substances so behave; also, the melting temperature depends on the purity of the substance. Thus one reason for devising a thermometer (literally, a meter for temperature) was that with it the composition of matter could be studied.

Temperature measurements are useful for studying molecular motions in material. **Figure 1** shows how the temperature of 1 g of H_2O, starting as ultracold ice, changes as heat is added at the constant rate of 1 cal/s (4.18 joules/s), assuming that no heat is lost to the surroundings. The graph plateaus illustrate Black's discovery that the temperature is constant during a phase change of solid to liquid or liquid to gas. *See* BOILING POINT; MELTING POINT.

Thermometers do not measure a special physical quantity. They measure length (as of a mercury column) or pressure or volume (with the gas thermometer at the National Bureau of Standards) or electrical voltage (with a thermocouple). The basic fact is that, if a mercury column has the same length when touching two different, separated objects, when the objects are placed in contact no heat will flow from one to the other. *See* THERMOMETER.

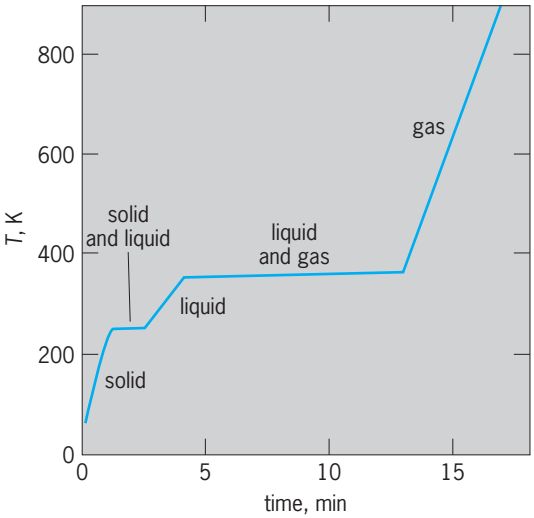

Fig. 1. Temperature of 1 g of H_2O, starting at 0 K, with a constant heat input of 1 cal/s (4.18 J/s). °F = (K × 1.8) − 459.67.

Empirical scales. The numbers on the thermometer scales are merely historical choices; they are not scientifically fundamental. The most widely used scales are the Fahrenheit (°F) and the Celsius (°C). The Centigrade scale with 0° assigned to ice water (ice point) and 100° assigned to water boiling under one atmosphere pressure (steam point) was formerly used, but it has been succeeded by the Celsius scale, defined in a different way than the Centigrade scale. However, on the Celsius scale the temperatures of the ice and steam points differ by only a few hundredths of a degree from 0° and 100°, respectively. **Figure 2** shows how the Celsius and Fahrenheit scales compare and how they fit onto the absolute scales. *See* ICE POINT.

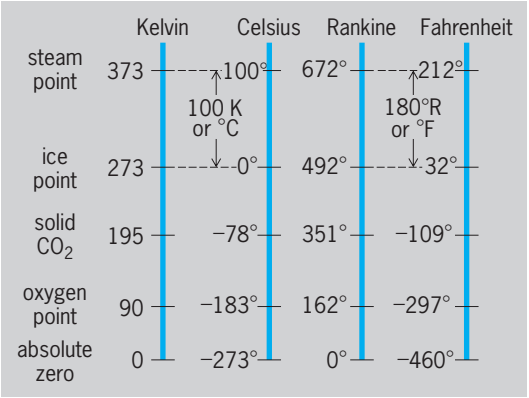

Fig. 2. Comparisons of Kelvin, Celsius, Rankine, and Fahrenheit temperature scales. Temperatures are rounded off to nearest degree. (*After M. W. Zemansky, Temperatures Very Low and Very High, Van Nostrand, 1964*)

These scales have one common value: $-40°C = -40°F$. This fact can be used to change a temperature from one scale to the other. Given a temperature in °C or in °F, add 40, multiply by 95 if converting from °C to °F or by $^5/_9$ if from °F to °C, then subtract 40. Example: Normal human body temperature is 98.6°F. To convert to °C, $98.6 + 40 = 138.6$; $138.6 \times {}^5/_9 = 77.0$; $77.0 - 40 = 37.0°C$.

Absolute temperature scale. In 1848 William Thomson (Lord Kelvin), following ideas of Sadi Carnot, stated the concept of an absolute scale of temperature in terms of measuring amounts of heat flowing between objects. Most important, Kelvin conceived of a body which would not give up any heat and which was at an absolute zero of temperature. Experiments have shown that absolute zero corresponds to −273.15°C or −459.7°F. Two absolute scales, shown in Fig. 2, are the Kelvin (K) and the Rankine (°R). *See* ABSOLUTE ZERO.

Interest in temperature and heat flow was stimulated in the early nineteenth century by efforts to improve the efficiency of steam engines. Out of this came the concept of a Carnot engine. This is not a real machine, but an imagined, ideal, frictionless system. A Carnot engine takes in heat Q_b from a higher temperature source at T_b (kelvins), does work W, and exhausts heat Q_l into a lower temperature source T_l. Two important deductions are: (1) The efficiency of the engine is $W/Q_b = 1 - T_l/T_b$. For example, a Carnot engine operated between the boiling point ($T_b = 373$ K) and the ice point ($T_l = 273$ K) of water has an efficiency of 0.268. A real engine would have an efficiency less than this, but the concept is nevertheless of great importance in engineering. (2) The ratio of temperatures equals the ratio of heats, namely, $T_b/T_l = Q_b/Q_l$. Lord Kelvin suggested that this be the basis for the absolute temperature scale. One special system (water at its triple point with ice, liquid, and water vapor present) is defined to have a particular value of absolute temperature (273.16 K). Any unknown temperature T_u, can be measured by operating a Carnot engine between 273.16 K and T_u, measuring the heats $Q_{273.16}$ and Q_u absorbed and rejected, and calculating $T_u = 273.16$ K $(Q_u/Q_{273.16})$. *See* CARNOT CYCLE; THERMODYNAMIC PRINCIPLES; TRIPLE POINT.

In practice, absolute temperatures are not measured this way. Instead low-density helium gas and dilute paramagnetic crystals, the most nearly ideal of real materials, allow measurement of temperatures virtually identical with those defined by a Carnot process. The advantages are that gas pressures and volumes and magnetic fields and magnetizations can be measured more conveniently and accurately than heat flows. *See* GAS THERMOMETRY.

The measurement of a single temperature with a gas or magnetic thermometer is a major scientific event done at a national standards laboratory. Only a few temperatures have been measured, including the freezing point of gold (1337.91 K or 1948.57°F), and the boiling points of sulfur (717.85 K or 832.46°F), oxygen (90.18 K or −297.35°F), and helium (4.22 K or −452.07°F). Various other types of thermometers (platinum, carbon, and doped germanium resistors; thermocouples) are calibrated at these temperatures and used to measure intermediate temperatures. *See* TEMPERATURE MEASUREMENT.

Kinetic temperature. An important aspect of the absolute temperature scale is its relation to the motions of atoms and molecules, whether vibrations as in solids and liquids, or straight path flights with collisions as in gases. There are two important facts here: (1) There is a definite distribution of motions. For example, in a gas, even though the motions are chaotic and a particular molecule changes velocity after each collision with another molecule, at any instant a definite number of molecules have a particular velocity. It cannot be said that the gas is at a definite temperature unless the molecules have this definite distribution of velocities, although different small portions of the gas may have definite, though different, temperatures. The same idea holds for the distribution of vibration frequencies in solids and liquids. (2) A body has a minimum amount of motion energy. It was supposed in the nineteenth century that this minimum was zero energy, but modern theories and experiments show that the minimum is greater than zero. A body in its lowest energy state cannot give out heat and is at absolute zero.

A system may have several degrees of freedom. The molecules of a gas, besides having straight-line motions, may rotate and vibrate and their electrons may be in different energy levels. When a system is in equilibrium, the energies stored in these different degrees of freedom are related to a common absolute temperature. In several cases, one can measure something related to a particular degree of freedom of a system. Then if the system is in equilibrium, its absolute temperature can be inferred.

Examples of this are measurements of temperatures in the Sun's corona and in the remote regions of the Milky Way Galaxy. The Sun's corona is so hot that atoms lose several electrons, that is, are multiply ionized. This greatly affects the wavelengths of light that these atoms emit. From the measured wavelengths that have been identified, the corona's temperature has been estimated at 2×10^6 K (3.6×10^6 °F).

Besides light, atoms and molecules can emit radio waves. The straight-line motions of atoms toward or away from Earth-based radio telescopes slightly shift the received wavelengths. From these Doppler shifts, temperatures from 1 to 100 K (−458 to −280°F) have been found in the vast hydrogen atom clouds in the galaxy spiral arms. Rotations of OH molecules influence the populations of energy levels and thus affect the intensities of emitted radio signals. From these also, deep-space absolute temperatures are inferred. *See* KINETIC THEORY OF MATTER.

Negative absolute temperatures. Negative Celsius and Fahrenheit temperatures are readily accepted because 0° on these scales is arbitrarily set above absolute zero (Fig. 2). But the idea that a system could be at, say −50 K, was introduced only in the early 1950s. The concept applies only to systems with a finite number of energy levels, that is, those that can

store a finite amount of energy. Thus the translational energy of a gas or the vibration energy of a crystal cannot be at negative absolute temperatures. However, the energies of electron and nuclear magnetic moments (spins) in a magnetic field do have upper limits. Normally there are more spins in lower energy levels, in which case they are at positive absolute temperature. With special techniques one can arrange equal numbers of spins in the energy levels. Then the temperature is infinite kelvins. This does not mean infinite heat can be extracted from such a system. One can go further and cause there to be more spins in the higher energy levels than in the lower, and this situation is described as being at a negative absolute temperature.

The spin system is not in equilibrium with the crystal lattice vibrations, which remain at positive temperatures. But at temperatures around a few kelvins, the spins need minutes or hours to exchange heat with the lattice. Negative absolute temperatures are hotter than positive temperatures; this is reasonable since heat flows from the negative temperature spins to the positive temperature lattice. *See* NEGATIVE TEMPERATURE.

Extreme temperatures. For reasons involving both pure and applied science, researchers endeavor to achieve extremes of both low and high temperatures. At very low temperatures, phenomena that are well understood become frozen out and new predicted and unpredicted effects are sought. Experimenters have reported holding 4 lb (2 kg) of copper at about 50 microkelvins (9×10^{-5} °F above absolute zero) for a couple of days. Achieving and measuring such temperatures involve using the nuclear magnetic spin system. In the procedure to reach the 50-μK lattice temperature, the copper spins were brought down to 50 nanokelvins (9×10^{-8} °F above absolute zero). One object of such research is to see if nuclear spins will spontaneously align as electron spins do in ferromagnetic materials like iron. *See* LOW-TEMPERATURE PHYSICS; LOW-TEMPERATURE THERMOMETRY.

The highest equilibrium temperatures reached on Earth are around 10^7 K (2×10^7 °F) in experiments to achieve fusion of hydrogen nuclei. These temperatures have been sustained in very low-density gases like the Sun's corona for a few seconds. At these temperatures hydrogen nuclei have speeds of about 4.3×10^6 ft/s (1.3×10^6 m/s). Such speeds are necessary if the positively charged nuclei are to overcome their electric repulsion force. When the nuclei get close enough together, nuclear forces attract them to fuse and there is a net energy release. The goal of the research is to convert this energy into conventional electrical energy. *See* NUCLEAR FUSION.

Roland A. Hultsch

Bibliography. T. D. McGee, *Principles and Methods of Temperature Measurements*, 1988; B. W. Magum and G. T. Furukawa, *Guidelines for Realizing the Temperature Scale of 1990 (ITS-90)*, NIST Tech. Note 1265, 1990; K. Mendelssohn, *The Quest for Absolute Zero*, 2d ed., 1977; T. J. Quinn, *Temperature*, 2d ed., 1991; J. F. Schooley (ed.), *Temperature: Its Measurement and Control in Science and Industry*, vol. 5, 1982; M. W. Zemansky, *Temperatures Very Low and Very High*, 1964, reprint 1981.

Temperature adaptation

The ability of animals to survive and function at widely different temperatures is a result of specific physiological adaptations. Temperature is an all-pervasive attribute of the environment that limits the activity, distribution, and survival of animals. Ocean temperatures range from 28 to 86°F (−2 to 30°C), but considerably warmer temperatures are found near deep-sea hydrothermal vents (up to 662°F or 350°C), backwaters of desert streams (109°F or 43°C), and geothermal springs (194–212°F or 90–100°C). Air temperatures range from −94°F (−70°C) in polar regions to 176°F (80°C) at the desert surface. Although some bacteria and blue-green algae live at temperatures up to 230°F (110°C), life processes are generally restricted to the temperatures between 32 and 113°F (0 and 45°C), and most animals live within an even narrower range. Limits for reproduction and development are generally narrower than those for survival of adults.

Changes in temperature influence biological systems, both by determining the rate of chemical reactions and by specifying equilibria (**Fig. 1**). Because temperature exerts a greater effect upon the percentage of molecules that possess sufficient energy to react (that is, to exceed the activation energy) than upon the average kinetic energy of the system, modest reductions in temperature (for example, from 77 to 59°F or from 25 to 15°C, corresponding to only a 3% reduction in average kinetic energy) produce a marked depression (two- to threefold) in reaction rate. In addition, temperature specifies the equilibria between the formation and disruption of the noncovalent (electrostatic, hydrophobic, and hydrogen-bonding) interactions that stabilize both the higher levels of protein structure

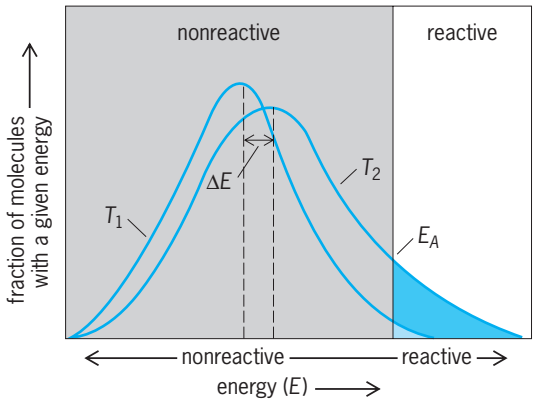

Fig. 1. Energy distribution profiles for a population of molecules at two different temperatures $T_2 > T_1$. The peaks in the two curves represent the average energy contents, and ΔE is the temperature-induced change in average energy content; E_A is the activation energy, and the shaded sections under the two curves show the change in the proportion of molecules with $E > E_A$.

and macromolecular aggregations such as biological membranes. Maintenance of an appropriate structural flexibility is a requirement for both enzyme catalysis and membrane function, yet cold temperatures constrain while warm temperatures relax the conformational flexibility of both proteins and membrane lipids, thereby perturbing biological function.

Animals are classified into two broad groups depending on the factors that determine body temperature. For ectotherms, body temperature is determined by sources of heat external to the body; levels of resting metabolism (and heat production) are low, and mechanisms for retaining heat are limited. Such animals are frequently termed poikilothermic or cold-blooded, because the body temperature often conforms to the temperature of the environment. In contrast, endotherms produce more metabolic heat and possess specialized mechanisms for heat retention. Therefore, body temperature is elevated above ambient temperature; some endotherms (termed homeotherms or warm-blooded animals) maintain a relatively constant body temperature. There is no natural taxonomic division between ecto- and endotherms. Most invertebrates, fish, amphibians, and reptiles are ectotherms, while true homeothermy is restricted to birds and mammals. However, flying insects commonly elevate the temperature of their thoracic musculature prior to and during flight (to 96°F or 36°C), and several species of tuna retain metabolic heat in their locomotory musculature via a vascular countercurrent heat exchanger. *See* THERMOREGULATION.

Endotherms. Homeotherms, by defending a constant body temperature, circumvent the problems associated with the maintenance of physiological function under varying body temperatures. The ability to regulate body temperature is dependent upon a suite of physiological adaptations involving the management of heat production, the distribution of heat within the body, and the exchange of heat with the environment (**Fig. 2**). When stressed by cold, homeotherms maintain body temperature by both shivering and nonshivering thermogenesis, that is, the production of heat by processes that do not involve muscle contraction. Both processes are strictly controlled, and the regulated heat production is just sufficient to maintain body temperature. Nonshivering thermogenesis is of particular interest, because the principal site of heat production occurs in brown adipose tissue, a tissue whose sole function is thermogenesis. Brown adipose tissue is a particularly important source of heat in young or cold-acclimated mammals and in arousing hibernators. Conversely, when stressed by heat, homeotherms lose heat to their environment by the evaporation of water from some body surface. In cattle, horses, and humans, high rates of evaporation are accomplished by sweating, whereas in species that do not sweat (dogs and cats), panting occurs. Insulation, in the form of subcutaneous adipose tissue and external pelage, is an adaptation to reduce the cost of thermoregulation in cold environments. Circulatory adaptations permitting the redistribution of blood flow are also important to thermoregulation. When body temperature rises, large volumes of blood are shunted through skin capillaries to increase heat transfer to the environment; conversely, peripheral blood flow is reduced in response to hypothermia in order to curtail heat loss. *See* ADIPOSE TISSUE; HIBERNATION AND ESTIVATION.

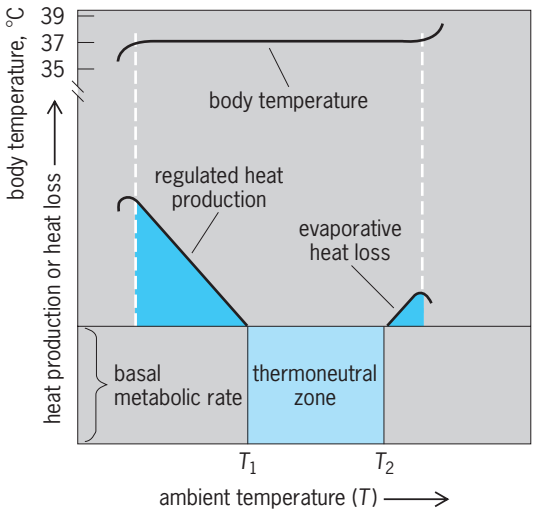

Fig. 2. Relationship between heat production and heat loss as a function of ambient temperature for a homeotherm. Thermoregulation can be achieved at minimal metabolic cost in the thermoneutral zone (defined by T_1 and T_2). Broken lines define the limits of ambient temperature over which a homeotherm can successfully defend (or maintain) its body temperature. $°F = (°C \times 1.8) + 32$.

Ectotherms. Many ectotherms are essentially isothermal with their environment; even the muscles of actively swimming fish are within 1.8°F (1°C) of the water temperature. Consequently, geographic, seasonal, or diurnal fluctuations in temperature pose particular problems for the maintenance of physiological function in these animals. Yet, although metabolic rates are depressed by acute exposure to cold, many ectotherms (but not all—some ectotherms become torpid at low temperature as a means of conserving energy) remain active in the cold and exhibit similar levels of activity at seasonal extremes of temperature. Similarly, arctic and antarctic ectotherms perform as effectively near 32°F (0°C) as tropical species do at 86°F (30°C). Maintenance of similar rates of activity at widely different body temperatures is a clear indication that ectotherms can adapt to temperature extremes. At the cellular level, thermal adaptations are most commonly reflected in the lipid composition of cell membranes and the catalytic properties of enzymes.

Lipid-mediated adaptations. Membranes perform many vital cell functions, including regulating the exchange of material between the cell and its environment; storing energy in the form of transmembrane ion gradients; providing a matrix in which many metabolic processes are organized; and controlling the flow of information between the cell

and its surroundings by generating intracellular messengers in response to extracellular signals. Yet, the physical properties of phospholipids—the primary structural elements of biological membranes—are markedly temperature-dependent. With cooling, the acyl domain of phospholipids is transformed from a fluid to a gel phase. Such phase transitions significantly perturb membrane function, for in the gel phase the membrane is rigid, passive permeability is reduced, and the activity of membrane-associated enzymes declines. Conversely, at elevated temperatures membranes become hyperfluid and leaky to ions (loss of potassium from muscle cells is a contributing factor in heat death). *See* CELL MEMBRANES; OSMOREGULATORY MECHANISMS.

Ectotherms overcome these problems by restructuring their membranes so that lipids of appropriate physical properties are matched to the prevailing ambient temperature. As temperature drops, high-melting lipids are replaced by lower-melting ones; consequently, membranes remain fluid at cold temperatures. The melting point of membrane lipids is lowered primarily by increasing the degree of acyl chain unsaturation, which introduces a kink into the acyl chain and prevents close packing at low temperatures. Two metabolic adjustments contribute to this restructuring process: increased activities of acyl chain desaturases at cold temperatures, and the operation of a deacylation–reacylation cycle, which permits the acyl chain composition to be altered independently of the rest of the phospholipid molecule. In addition, phospholipids with small, as opposed to bulky, head groups also increase in abundance at low temperatures, and the resulting rise in the ratio of conically to cylindrically shaped lipids may offset the direct effects of temperature change upon lipid packing (that is, it may disrupt packing at low temperature). Differences in lipid composition between polar and tropical species resemble those between seasonally adapted individuals of temperate species.

As a consequence of lipid restructuring, membrane fluidity is relatively constant when compared at the respective growth temperatures to which an animal has become adapted, even though varying markedly with acute changes in temperature. This phenomenon is known as homeoviscous adaptation. *See* LIPID.

Protein-mediated adaptations. In addition to the homeoviscous regulation of membrane fluidity, ectoderms display other adaptations that permit function over a broad temperature range. These range from evolutionally fixed differences in the structure and function of specific proteins to seasonal adjustments in the rates and patterns of energy metabolism.

Enzyme structure and function vary interspecifically in a manner consistent with the conservation of catalytic rates and regulatory properties at different temperatures. The catalytic efficiency of enzymes is inversely correlated with habitat or cell temperature. For example, lactate dehydrogenase from an antarctic fish produces nearly twice as much product per minute per mole of enzyme as does the homologous

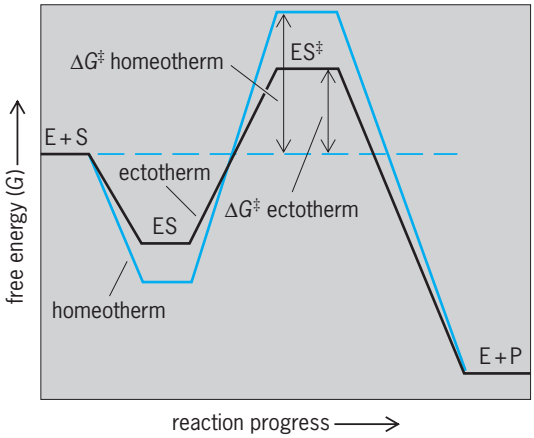

Fig. 3. Energy profiles for the reactions catalyzed by homologous enzymes from an ectotherm and a homeotherm. Note that the activation free-energy barrier is higher for the homeothermic than the ectothermic enzyme because of the tighter binding of substrate by the former. E = free enzyme; S = free substrate; P = free product; ES = enzyme substrate complex, not in the activated state; ES‡ = enzyme substrate complex in the activated state; ΔG‡ = free energy of activation. (*After P. W. Hochachka and G. N. Somero, Biochemical Adaptation, Princeton University Press, 1984*)

enzyme from rabbit muscle when compared at 41°F (5°C). Such increased catalytic efficiencies typical of ectothermic enzymes are a reflection of a lowered activation energy. It has been postulated that differences in catalytic efficiency between homologous enzymes of ectotherms and endotherms reflect varying amounts of weak bond formation between the enzyme and substrate during the activation step of catalysis. The disruption of relatively few weak bonds in the formation of products by the enzymes of cold-adapted ectotherms would keep the free energy of the reaction low and the rate relatively high, thus providing an important mechanism for compensation of metabolic rates at low temperatures (**Fig. 3**). This hypothesis is substantiated by the observation that enzymes of homeotherms bind substrates as strongly at 98.6°F (37°C) as do those of antarctic species at 32°F (0°C). Such interspecific differences in enzyme function are presumed to reflect genetically fixed differences in primary structure.

In contrast to the evolutionary tailoring of enzymes to specific thermal environments, there is little evidence that seasonal acclimatization results in the production of environment-specific isozymes best suited to function at either warm or cold temperatures. Instead, seasonal acclimation generally results in altered levels of enzyme activity, which are presumed to reflect temperature-dependent differences in cellular enzyme content. In contrast to qualitative changes in the efficiency of an enzyme, this quantitative adaptation simply involves altered amounts of an identical enzyme at seasonal extremes of temperature. Acyl-chain desaturating enzymes are induced in cold-acclimated ectotherms; in addition, the activities of enzymes of aerobic (mitochondrial) metabolism are generally elevated at cold temperatures, providing an explanation for the thermal

compensation of metabolic rate. *See* CATALYSIS; COLD HARDINESS (PLANT); ENZYME. Jeffrey R. Hazel

Bibliography. R. C. Aloia, C. C. Curtain, and L. M. Gordon (eds.), *Physiological Regulation of Membrane Fluidity*, 1988; A. R. Cossins and K. Bowler, *Temperature Biology of Animals*, 1987; J. R. Hazel and C. L. Prosser, Molecular mechanisms of temperature compensation of poikilotherms, *Physiol. Rev.*, 54(3):620–677, 1974; P. W. Hochachka and G. N. Somero, *Biochemical Adaptation*, 1984; C. L. Prosser, *Adaptational Biology: Molecules to Organisms*, 1986.

Temperature inversion

The increase of air temperature with height; an atmospheric layer in which the upper portion is warmer than the lower. Such an increase is opposite, or inverse, to the usual decrease of temperature with height, or lapse rate, in the troposphere of about 3.3°F/1000 ft (6.5°C/km) and somewhat less on mountain slopes. However, above the tropopause, temperature increases with height throughout the stratosphere, decreases in the mesosphere, and increases again in the thermosphere. Thus inversion conditions prevail throughout much of the atmosphere much or all of the time, and are not unusual or abnormal. *See* AIR TEMPERATURE; ATMOSPHERE.

Inversions are created by radiative cooling of a lower layer, by subsidence heating of an upper layer, or by advection of warm air over cooler air or of cool air under warmer air. Outgoing radiation, especially at night, cools the Earth's surface, which in turn cools the lowermost air layers, creating a nocturnal surface inversion a few centimeters to several hundred meters thick. Over polar snowfields, inversions may be a kilometer or more thick, with differences of 54°F (30°C) or more. Solar warming of a dust layer can create an inversion below it, and radiative cooling of a dust layer or cloud top can create an inversion above it. Sinking air warms at the dry adiabatic lapse of 5°F/1000 ft (10°C/km), and can create a layer warmer than that below the subsiding air. Air blown onto cool water from warmer land or onto snow-covered land from warmer water can cause a pronounced inversion that persists as long as the flow continues. Warm air advected above a colder layer, especially one trapped in a valley, may create an intense and persistent inversion.

Inversions effectively suppress vertical air movement, so that smokes and other atmospheric contaminants cannot rise out of the lower layer of air. California smog is trapped under an extensive subsidence inversion; surface radiation inversions, intensified by warm air advection aloft, can create serious pollution problems in valleys throughout the world; radiation and subsidence inversions, when horizontal air motion is sluggish, create widespread pollution potential, especially in autumn over North America and Europe. *See* AIR POLLUTION; SMOG.

Arnold Court

Temperature measurement

Measurement of the hotness of a body relative to a standard scale. The fundamental scale of temperature is the thermodynamic scale, which can be derived from any equation expressing the second law of thermodynamics. Efforts to approximate the thermodynamic scale as closely as possible depend on relating measurements of temperature-dependent physical properties of systems to thermodynamic relations expressed by statistical thermodynamic equations, thus in general linking temperature to the average kinetic energy of the measured system. Temperature-measuring devices, thermometers, are systems with properties that change with temperature in a simple, predictable, reproducible manner. *See* TEMPERATURE; THERMODYNAMIC PRINCIPLES.

Temperature scale. In the establishment of a useful standard scale, assigned temperature values of thermodynamic equilibrium fixed points are agreed upon by an international body (the General Conference of Weights and Measures), which updates the scale about once every 20 years. Thermometers for interpolating between fixed points and methods for realizing the fixed points are prescribed, providing a scheme for calibrating thermometers used in science and industry.

The scale now in use is the International Temperature Scale of 1990 (ITS-90). Its unit is the kelvin, K, arbitrarily defined to be 1/273.16 of the thermodynamic temperature T of the triple point of water (the state in which the liquid, solid, and vapor phases coexist). The scale extends upward from 0.65 K. For temperatures above 273.15 K, it is common to use International Celsius Temperatures, t_{90} (rather than International Kelvin Temperatures, T_{90}), having the unit degree Celsius, with symbol °C. The degree Celsius has the same magnitude as the kelvin. Temperatures, t_{90} are defined as $t_{90}/^\circ\mathrm{C} = T_{90}/\mathrm{K} - 273.15$, that is, as differences from the ice-point temperature at 273.15 K. The ice point is the state in which the liquid and solid phases of water coexist at a pressure of 1 atm (101,325 pascals). [The Fahrenheit scale, with symbol °F, still in common use in the United States, is given by $t_F/^\circ\mathrm{F} = t_{90}/^\circ\mathrm{C} \times 1.8) + 32$, or $t_F/^\circ\mathrm{F} = (T_{90}/\mathrm{K} \times 1.8) - 459.67$.] The ITS-90 is defined by 17 fixed points (**Table 1**). Between 0.65 and 5.0 K, the ITS-90 is defined in terms of the vapor pressure of ^{3}He and ^{4}He; between 3.0 and 24.5561 K, by interpolating constant-volume gas thermometry (using either ^{3}He or ^{4}He); between 13.8033 K and 961.78°C, by platinum resistance thermometry; and above 961.78°C, by radiation thermometry. There are overlapping ranges of vapor pressure and interpolating constant-volume gas thermometry, and of interpolating constant-volume gas thermometry and platinum resistance thermometry, with the different definitions having equal status. For radiation thermometry, based on Planck's radiation formula, the silver, gold, or copper point may be used as the reference temperature. Below 0.65 K, no internationally agreed upon scale exists, but the scale may be extended into this range in the future. At present,

TABLE 1. Defining fixed points of the International Temperature Scale of 1990 (ITS-90)

Equilibrium state*	Temperature T_{90}, K	Temperature t_{90}, °C
Vapor pressure equation of helium	3 to 5	−270.15 to −268.15
Triple point of equilibrium hydrogen†	13.8033	−259.3467
Vapor pressure point of equilibrium hydrogen† (or constant volume gas thermometer point of helium)	≈17	≈−256.15
Vapor pressure point of equilibrium hydrogen† (or constant volume gas thermometer point of helium)	≈20.3	≈−252.85
Triple point of neon	24.5561	− 248.5939
Triple point of oxygen	54.3584	−218.7916
Triple point of argon	83.8058	−189.3442
Triple point of mercury	234.3156	−38.8344
Triple point of water	273.16	0.01
Melting point of gallium	302.9146	29.7646
Freezing point of indium	429.7485	156.5985
Freezing point of tin	505.078	231.928
Freezing point of zinc	692.677	419.527
Freezing point of aluminum	933.473	660.323
Freezing point of silver	1234.93	961.78
Freezing point of gold	1337.33	1064.18
Freezing point of copper	1357.77	1084.62

*The triple point is the equilibrium temperature at which the solid, liquid, and vapor phases coexist. The freezing point and the melting point are the equilibrium temperatures at which the solid and liquid phases coexist under a pressure of 101,325 Pa, 1 standard atmosphere. The isotopic composition is that naturally occurring.

† Equilibrium hydrogen is hydrogen with the equilibrium distribution of its ortho and para states at the corresponding temperatures. Normal hydrogen at room temperature contains 25% para and 75% ortho hydrogen.

temperatures below 0.65 K are determined by magnetic thermometry, nuclear orientation thermometry, and noise thermometry. In some cases, a ^{3}He melting-curve thermometer is used. *See* LOW-TEMPERATURE THERMOMETRY.

Primary thermometers. These are devices which relate the thermodynamic temperature to statistical mechanical formulations incorporating the Boltzmann constant k_B or the Boltzmann factor $\exp(-E_i/k_BT)$, where E_i is the energy of the ith state of the system (**Table 2**). However, the fixed points of at and below 419.527°C of the ITS-90 are all based on one or more types of gas thermometry, with those above 419.527°C being determined by spectral radiation pyrometry referenced to gas thermometry at a temperature near 460°C. *See* GAS THERMOMETRY.

Secondary thermometers. These are used as reference standards in the laboratory because primary thermometers are often too cumbersome. Since this is especially true for realization of the ITS-90 below the silver point, it is necessary to establish standard or secondary thermometers referenced to one or more fixed points for interpolation between fixed points. Over the range 13.8033 K to 961.78°C, the platinum resistance thermometer with specified characteristics is the prescribed standard; above 961.78°C, optical pyrometry and spectroscopic methods are used. In the low-temperature range, the rhodium-iron resistance thermometer is an excellent candidate for interpolation device between about an 0.3 and 35 K, to overlap the platinum thermometer, the vapor-pressure scales of the helium isotopes, and the interpolating constant-volume gas thermometer. The germanium resistance thermometer also is a suitable interpolating device between about 0.05 and 30 K. *See* PYROMETER; THERMOCOUPLE; THERMOMETER.

Lower-order thermometers. These are used for most practical purposes and, when fairly high accuracy is required, can usually be calibrated against reference standards maintained at national standards laboratories, such as the U.S. National Institute of Standards and Technology, or against portable reference devices (sealed freezing or melting point cells). Examples of these practical thermometers include wide varieties of resistance thermometers (employing pure metals or semiconductors, the latter including carbon thermometers, germanium thermometers, and thermistors), thermoelectric thermometers (thermocouples), liquid-in-glass thermometers, vapor-pressure thermometers, magnetic thermometers, and capacitance thermometers. To obtain the highest accuracy when a calibrated thermometer is put into use, the conditions must reproduce those established during the calibration. Special care must be taken to ensure that the thermometer is in good thermal contact with the body whose temperature is to be measured; that sufficient time is allowed before measurement is made for the thermometer to equilibrate its temperature with that of the body; and that extraneous heat leaks are eliminated. *See* THERMISTOR.

Temperature indicators. Often it is useful to ascertain when or if a certain temperature has been reached or exceeded. Such qualitative information can be acquired by the use of temperature indicators, which are either irreversible or reversible. Examples of irreversible indicators are pyrometric ceramic cones, which can be formulated to soften and bend over a specific range of temperatures, and metallic

TABLE 2. Primary thermometry methods

Method	Approximate useful range of T, K	Principal measured variables	Relation of measured variables to T	Remarks
Gas thermometry	1.3–950	Pressure P and volume V	Ideal gas law plus correction: $PV \propto k_B T$ plus corrections	Careful determination of corrections necessary, but capable of high accuracy
Acoustic interferometry	1.5–3000	Speed of sound W	$W^2 \propto k_B T$ plus corrections	
Magnetic thermometry 1. Electron paramagnetism 2. Nuclear paramagnetism	 0.001–35 0.000001–1	Magnetic susceptibility	Curie's law plus corrections: $\chi \propto 1/k_B T$ plus corrections	
Gamma-ray anisotropy or nuclear orientation thermometry	0.01–1	Spatial distribution of gamma-ray emission	Spatial distribution related to Boltzmann factor for nuclear spin states	Useful standard for $T < 1$ K
Thermal electric noise thermometry 1. Josephson junction point contact 2. Conventional amplifier	 0.001–1 4–1400	Mean square voltage fluctuation $\overline{V}^2$	Nyquist's law: $\overline{V}^2 \propto k_B T$	Other sources of noise serious problem for $T > 4$ K
Radiation thermometry (visual, photoelectric, or, photodiode)	500–50,000	Spectral intensity J at wavelength λ	Planck's radiation law, related to Boltzmann factor for radiation quanta	Needs blackbody conditions or well-defined emittance
Infrared spectroscopy	100–1500	Intensity I of rotational lines of light molecules	Boltzmann factor for rotational levels related to I	Also Doppler line broadening ($\propto\sqrt{k_B T}$) useful; principal applications to plasmas and astrophysical observations; proper sampling, lack of equilibrium, atmospheric absorption often problems
Ultraviolet and x-ray spectroscopy	5000–2,000,000	Emission spectra from ionized atoms—H,He, Fe, Ca, and so on	Boltzmann factor for electron states related to band structure and line density	

pellets or notched rods, which melt or deform at known temperatures. Examples of reversible indicators are certain paints or crayons and liquid-crystal materials, which change color, sometimes exceedingly sharply, at predeterminable temperatures. *See* LIQUID CRYSTALS. B. W. Mangum

Bibliography. R. P. Benedict, *Fundamentals of Temperature, Pressure and Flow Measurements*, 3d ed., 1984; J. R. Leigh, *Temperature Measurement and Control*, 1988; T. D. McGee, *Principles and Methods of Temperature Measurement*, 1988; T. J. Quinn, *Temperature*, 2d ed., 1990; F. Schooley, *Thermometry*, 1988.

Tempering

The reheating of previously quenched alloy to a predetermined temperature below the critical range, holding the alloy for a specified time at that temperature, and then cooling it at a controlled rate, usually by immediate rapid quenching, to room temperature. Tempering has a long history, beginning as an empirical process to impart toughness. Consequently, the term is broadly applied to any process that toughens a material. *See* ALLOY.

In alloys, if the composition is such that cooling produces a supersaturated solid solution, the resulting material is brittle. Heating the alloy to a temperature only high enough to allow the excess solute to precipitate out and then rapidly cooling the saturated solution fast enough to prevent further precipitation of grain growth result in a microstructure combining hardness and toughness. *See* SOLID SOLUTION.

With steel, precipitation of carbide out of the supersaturated unstable martensite or interstitial solution of carbon in iron results in improved ductility. To be most successful, the tempering must be carried out by slow heating to avoid steep temperature gradients, stress relief being one of the objectives. Properties produced by tempering depend on the temperature to which the steel is raised and on its alloy composition. For example, if hardness is to be retained, molybdenum or tungsten is used in the alloy.

For tool steel, to obtain great hardness, the finished or nearly finished tool may be heated to a critical temperature dependent on the particular alloy. While at this temperature below the transformation range, the steel is refined; that is, it forms a solid solution with small grain size. After being held at the refining temperature long enough to reach thermal equilibrium throughout, the steel is rapidly cooled or quenched so as to pass quickly through temperature regions in which grain growth can occur. For steel

with 0.83% carbon the refining temperature is slightly above 1290°F (700°C); the faster this alloy is cooled the harder it becomes. *See* HEAT TREATMENT (METALLURGY). Frank H. Rockett

Bibliography. American Society for Testing and Materials, *Temper Embrittlement of Alloy Steels*, 1972; L. Coudurier, D. W. Hopkins, and I. Wilkomirsky, *Fundamentals of Metallurgical Processes*, 1985; J. V. Woodworth, *Steel Hardening, Tempering, and Forging Made Easy*, 1991.

Temporary structure (engineering)

A structure erected to aid in the construction of a permanent project. Temporary structures are used to facilitate the construction of buildings, bridges, tunnels, and other above- and below-ground facilities by providing access, support, and protection for the facility under construction, as well as assuring the safety of the workers and the public. Temporary structures either are dismantled and removed when the permanent works become self-supporting or completed, or are incorporated into the finished work. Temporary structures are also used in inspection, repair, and maintenance work.

There are many types of temporary structures, including cofferdams; earth-retaining structures; tunneling supports; underpinning; diaphragm/slurry walls; roadway decking; construction ramps, runways, and platforms; scaffolding; shoring; falsework; concrete formwork; bracing and guying; site protection structures such as sidewalk bridges, boards and nets for protection against falling objects, barricades and fences, and signs; and all sorts of unique structures that are specially conceived, designed, and erected to aid in a specific construction operation.

These temporary works have a primary influence on the quality, safety, speed, and profitability of all construction projects. More failures occur during construction than during the lifetimes of structures, and most of those construction failures involve temporary structures. However, codes and standards do not provide the same scrutiny as they do for permanent structures. Typical design and construction techniques and some industry practices are well established, but responsibilities and liabilities remain complex and present many contractual and legal pitfalls.

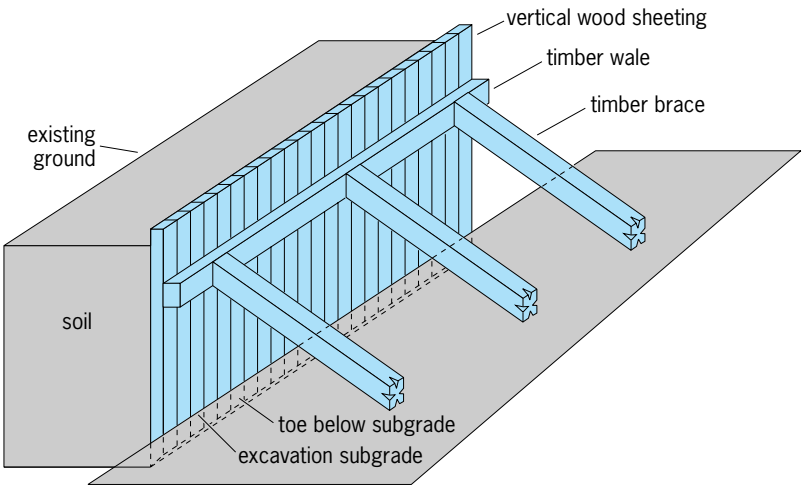

Fig. 1. Excavation sheeting and bracing. (*After R. T. Ratay, Handbook of Temporary Structures in Construction, 2d ed., McGraw-Hill, 1996*)

Cofferdams. Cofferdams are temporary enclosure walls designed to keep water and soil out of the work area, thus allowing the construction of the permanent facility in a dry environment. They enable the construction of bridge piers, intake structures, pump houses, locks, and dams below the surrounding water level by encircling the work area with an impervious wall. After the cofferdam is constructed, the area within it is pumped out to provide the dry construction site. Pumping usually continues at a moderate rate during the construction activities to remove any seepage from the surrounding body of water. *See* COFFERDAM.

Earth-retaining structures. The primary purpose of temporary earth-retaining structures is to permit the safe digging of a hole or trench in the ground and to provide a safe work area below the ground level. Another purpose is the mitigation of subsidence of the surrounding ground and settlement of nearby structures, pavements, and utilities. There are several common types of temporary earth retaining structures, including vertical wood sheeting and bracing, steel sheet piling, and soldier piles and horizontal sheeting. *See* RETAINING WALL.

Vertical wood sheeting and bracing. This type is the oldest and, under certain conditions, the most economical method of retaining earth. Its most common uses are in trench excavations, where the sheeting is made up of vertical planks, 2 to 3 in. (5 to 7.5 cm) thick, that are lowered or driven down and held in position against the soil by one or more lines of heavy horizontal timber wales (construction members), which in turn are pressed against the sheeting by horizontal or inclined bracing struts (**Fig. 1**). Successive levels of sheeting and bracing are installed as the excavation progresses downward.

Steel sheet piling. This type consists of a row of interlocking sheet piles, installed prior to excavation by driving the sheet piles one after another into the soil to a depth several feet below the bottom of the future excavation so as to provide embedment for the sheet piling to act as a cantilever retaining wall. Variations on the cantilever type are the braced and tied-back systems where, in addition to the anchorage by embedment at the bottom, the sheet piling is supported at one or more levels along its height. *See* CANTILEVER.

Soldier piles and horizontal sheeting. This system includes vertical steel beams, spaced 6 to 10 ft (1.8 to 3 m) apart, that act as vertical cantilevers supporting the sheeting panels that hold back the soil. The system is installed by driving the soldier piles (also known as soldier beams) to a depth several feet below the bottom of the future excavation, so as to provide embedment for their cantilever action, and then placing sheeting planks to span horizontally between soldiers as the excavation progresses. Variations on the cantilever type are the braced and tied-back systems where, in addition to the anchorage by embedment

at their bottoms, the solider piles are supported at one or more levels along their heights.

Tunneling supports. The basic types of temporary supports used in the construction of tunnels and shafts include steel liner plates; for soft ground, steel ribs and lagging; and for rock tunneling, rock bolting and grouting. There are a number of industry standards with many variations to suit given conditions. In most cases, temporary supports used during construction of tunnels and shafts remain in place after construction and become parts of the permanent supports. *See* TUNNEL.

Underpinning. Underpinning is a temporary or permanent support installed below an existing foundation to provide greater depth, a different load path, or increased bearing capacity, necessitated by some desired alteration of a foundation. The purpose is to relieve the foundation load at a particular location, transfer it to a temporary support, and transmit the load to the soil at a lower elevation. Common underpinning methods use pits and piers or driven or augered piles. Typically, a pit is dug below a part of the foundation, and a pier or pile is installed to support the foundation at that location. The load is transferred to the new support by drypacking or wedging between the bottom of the existing foundation and the top of the new pier. Excavation sheeting and concrete formwork, when used for constructing the piers, may be left or removed. The process is repeated under an adjacent area until the entire length of the existing foundation is underpinned.

Diaphragm/slurry walls. This term refers to a reinforced concrete wall constructed below ground level and utilizing the slurry method of trench stabilization. A trench with the width of the intended thickness of the concrete wall is excavated. As the digging progresses, the excavated material is immediately and continually replaced by a bentonite slurry mix. The lateral pressure created by the slurry prevents the walls of the trench from collapsing. After the trench is excavated to its final elevation, caged reinforcing steel is lowered through the slurry. The concrete, placed by the tremie method, fills the trench from the bottom up, displacing the lighter-weight bentonite slurry. Once the concrete cures and the wall attains its intended strength, the soil is excavated from one side, making the wall a temporary, and then later a permanent, retaining or foundation wall.

Roadway decking. Temporary roadway decking is the support and travel surface constructed (within the limits of streets and highways) to allow traffic over open-cut or belowground work areas. Once excavation within the roadway is completed, a support structure is constructed on which the surface decking is placed (**Fig. 2**). An exception is the simplest form of temporary decking: a very heavy steel plate placed over a relatively small excavation, set directly on the pavement without structural supports. The most common material for decking is heavy timber beams, similar in appearance to railroad ties, followed by heavy precast concrete slabs.

Construction ramps, runways, and platforms. Construction ramps, runways, and platforms provide vehicular access to and within a construction site. A construction ramp is utilized to provide access between the foundation or basement and the street level, or between two levels of the project. A runway is a strip of level roadway that provides a track for wheeled vehicles; it is usually a narrow extension into the construction site. A platform is a raised horizontal surface used by cranes, vehicles, equipment, and workers for convenient positioning. These facilities are constructed of earth, timber, steel, precast concrete, or combinations thereof.

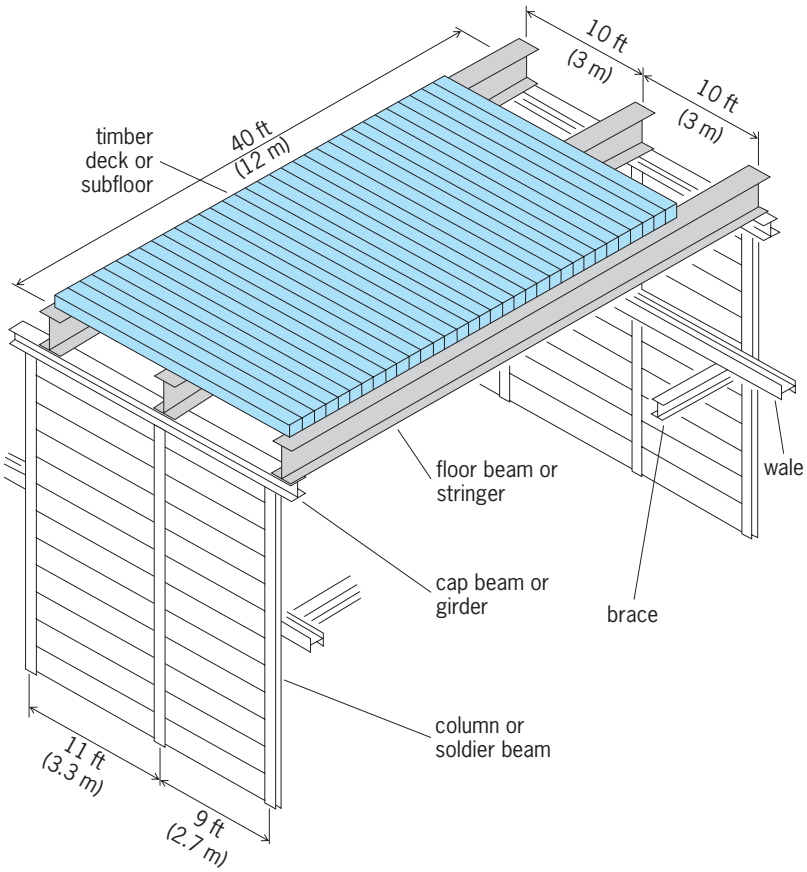

Fig. 2. Typical roadway decking with steel beams and timber decking. (*After R. T. Ratay, Handbook of Temporary Structures in Construction, 2d ed., McGraw-Hill, 1996*)

Scaffolding. Scaffolding is a structure used to provide access to a location of work that is too high for workers to reach. Scaffolds are made of aluminum or steel tubes assembled into structural systems, with wood or aluminum planking as the working platform. (In some Asian countries bamboo rods tied by lashings are used instead of metal tubes.) The two common types are the tube-and-coupler and the sectional scaffolds that can be assembled to hundreds of feet in height.

Tube-and-coupler. These scaffolds are assembled at the time and place of their use from five basic structural elements: the vertical posts, which rise from the ground; the horizontal bearers, which attach to the posts and support the working platforms; the runners, which attach to the posts directly below the bearers and provide longitudinal connections along the length of the scaffold; the diagonal bracing connected to the posts to provide stability in the

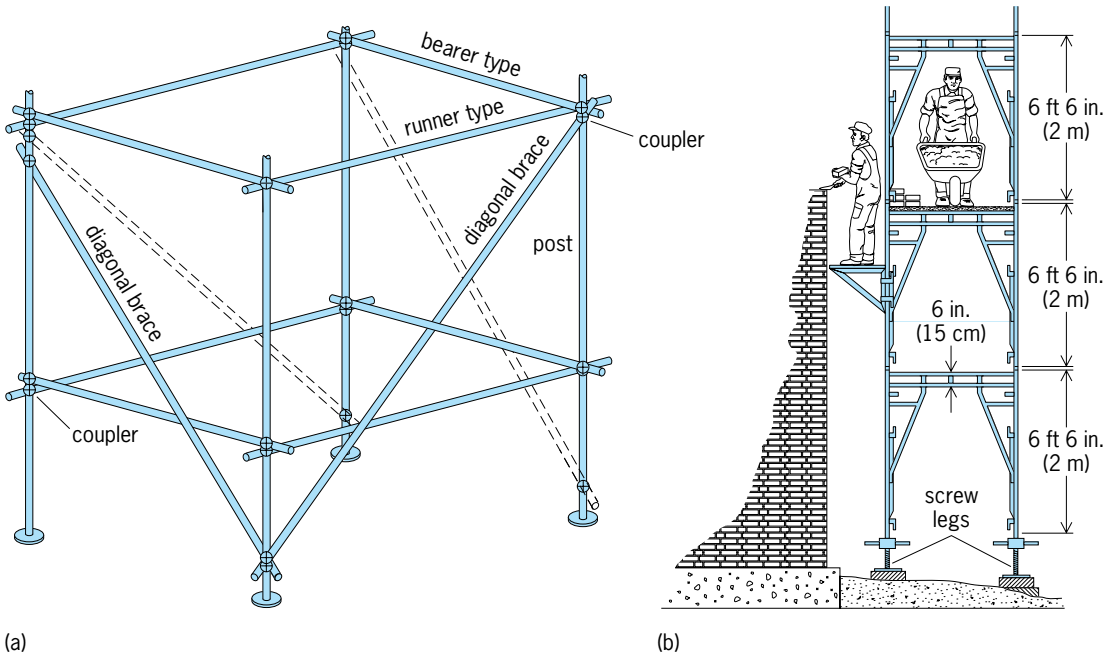

Fig. 3. Scaffolds. (*a*) Basic assembly of tube-and-coupler type. (*b*) Sectional walk-through type. (*After R. T. Ratay, Handbook of Temporary Structures in Construction, 2d ed., McGraw-Hill, 1996*)

longitudinal direction; and the couplers through which the parts connect (**Fig. 3***a*). Stability in the short direction is usually provided by anchoring against the permanent structure.

Sectional. These scaffolds have two basic components: prefabricated sectional frames that contain both posts and bearers (Fig. 3*b*), and vertical x-bracing that connects the frames and provides stability along the length of the scaffolding. Frames are stacked one on top of another up to the desired height. Some stability in the short direction is provided by the rigid frames themselves, but safety regulations require external bracing or guying above certain heights. Important additional parts of a stationary scaffold are the screw legs with base plates for solid founding and proper leveling of the assembly.

In addition to stationary ground-based scaffolds, there are rolling scaffold towers on casters, movable suspended scaffolds, and other special-purpose arrangements.

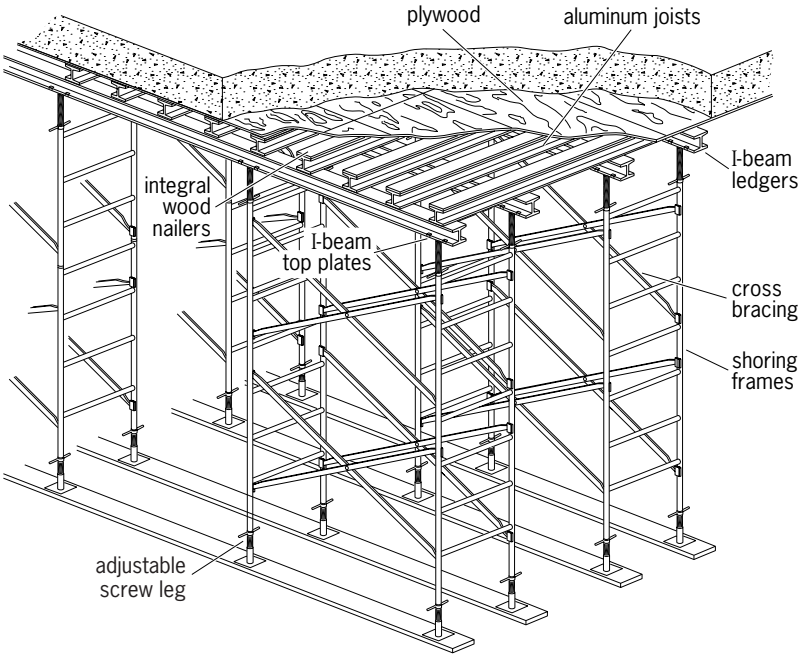

Fig. 4. Shoring towers used as falsework. (*After R. T. Ratay, Handbook of Temporary Structures in Construction, 2d ed., McGraw-Hill, 1996*)

Shoring. Shoring is a structure assembled for the purpose of temporarily supporting a load, usually during some construction, repair, or maintenance work. It may consist of a simple block, a single post, a single shoring tower, or an extensive system of columns, beams, braces, and towers. The traditional shoring material is wood, but manufactured shoring components made of tubes of structural steel are the most common. So-called shoring towers are used extensively because of their adaptability to load and size requirements. These towers are assembled at the site from two or more prefabricated steel shoring frames that are connected with x-bracing for lateral stability. Units of braced frames are stacked one on top of another to the desired height. Important additional components are the base (for solid founding and proper alignment of the shoring) and the head (for proper engagement of the supported load). Steel-tube shoring may look like scaffolding but it is different. While the function of scaffolding is to support the relatively light weights of workers and materials, shoring usually is used for heavy loads.

Falsework. Shoring that is used to support the formwork for fresh concrete or other materials is known as falsework (**Fig. 4**). Its purpose is to hold

the formwork in place until the concrete attains enough strength and stiffness to be self-supporting without excessive deformation. Falsework must be well founded, well braced, and tight against the formwork to prevent movement. After the removal of the formwork, the falsework is usually replaced by reshoring to provide continued support to the hardened concrete for an additional time period. Reshoring components are usually single wood or steel pipe posts.

Concrete formwork. Formwork is the mold into which the fresh concrete for a floor, beam, column, pier, and so forth is placed and remains until hardened. The formwork thus contains and temporarily supports the concrete and defines its permanent shape. Although the materials most frequently used for such applications are lumber and plywood because of their economy, availability, and workability in the job site, more durable materials such as aluminum, steel, or reinforced plastic are common for reusable forms. Forms may be custom-built at the site or constructed from prefabricated modular units. Tying the forms together, anchoring them, and supporting the framework is critical to maintain required shape and to resist movement due to very high liquid-concrete pressures. Premature or careless removal of the forms can create permanent damage to the built structure. *See* CONCRETE.

Bracing and guying. Temporary bracing and guying of a structure during construction provide lateral stability (resistance to horizontal forces). A brace is usually a rigid strut working in compression; a guy is a taut cable working in tension. In many instances, jacks are attached to braces, and turnbuckles to guy cables, to allow tightening or loosening in order to facilitate installation, alignment, and removal. Temporary bracing and guying may be installed either within or external to the permanent structure (**Fig. 5**), and remain until the construction of the permanent structure advances to the point which it can provide its own lateral stability. Thereafter the braces and guys are usually removed, but they may remain incorporated into the permanent structure. Because braces and guys are normally at an incline, the axial forces within them have horizontal and vertical components that must be included in calculating the construction design loads on the permanent structure.

Site protection structure. Hazards to life and property on and around a construction site are significant. Government regulations impose strict rules for safety at construction sites. A number of temporary structures are employed to warn and protect the workers and the public. Sidewalk bridges (shields) are strong platforms built on shoring posts or frames to protect the public from falling construction debris while maintaining pedestrian traffic on the sidewalk. Fall protection platforms and nets, used to catch falling objects and protect workers, people, and property below, are wooden planks and outstretched nets installed on outriggers near the top level of the construction; they are moved up periodically as the construction progresses. Barricades and fences protect people, equipment, and materials; delineate areas of the site; and organize the flow of traffic. Signs provide identification, information, and warning. Special-purpose temporary structures are employed at the adjacent areas to provide dust protection, blast protection, noise abatement, heat protection, and wind enclosures and to facilitate other construction-related activities.

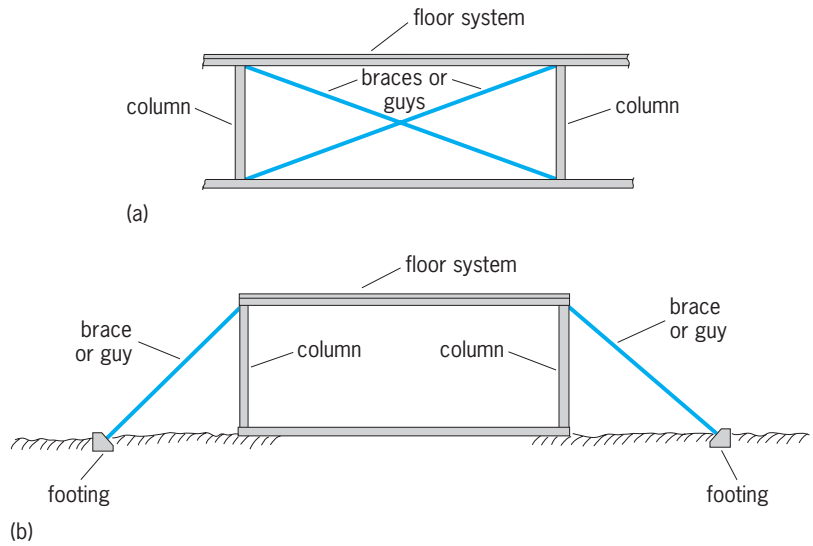

Fig. 5. Bracing and guying for lateral stability during construction. (*a*) Internally opposed braces or guys. (*b*) Externally opposed braces or guys. (*After R. T. Ratay, Handbook of Temporary Structures in Construction, 2d ed., McGraw-Hill, 1996*)

Robert T. Ratay

Bibliography. R. T. Ratay, *Forensic Structural Engineering Handbook*, McGraw-Hill, 2000; R. T. Ratay, *Handbook of Temporary Structures in Construction*, 2d ed., 1996.

Tendon

A cord connecting a muscle to another structure, often a bone. A tendon is a passive material, lengthening when the tension increases and shortening when it decreases. This characteristic contrasts with the active behavior of muscle. *See* MUSCLE.

Away from its muscle, a tendon is a compact cord. It can slide, in the direction of its length, freely through the body. At the muscle, it spreads into thin sheets called aponeuroses, which lie over and sometimes within the muscle belly. The large surface area of the aponeuroses allows the attachment of muscle fibers with a total cross-sectional area that is typically 50 times that of the tendon. The angle of attachment is usually 5–10°.

A tendon lengthens when subjected to a longitudinal tensile force. Stiffness increases with stress, reaching a constant value once the crimp has been straightened out. A tendon is a good spring in the sense that about 93% of the energy required in stretching is returned on release. Being constructed of partially independent, crimped fibers, tendons

can bend and buckle easily and without damage. The main exception to this flexibility is in the legs of some birds (for example, turkeys) where the tendons are calcified and thus are stiffer, especially against bending.

Structure. Tendons are living tissues that contain cells. In adult tendons, the cells occupy only a very small proportion of the volume and have a negligible effect on the mechanical properties. Like other connective tissues, tendon depends on the protein collagen for its strength and rigidity. The arrangement of the long, thin collagenous fibers is essentially longitudinal, but incorporates a characteristic waviness known as crimp. The fibers lie within a matrix of aqueous gel. Thus, tendon is a fiber-reinforced composite (like fiberglass), but its collagen is much less stiff than the glass and its matrix is very much less stiff than the resin. *See* COLLAGEN.

Where tendons go around corners, the matrix on the inside of the bend may be unusually stiff. This specialization makes such parts of tendon somewhat like cartilage. A more extreme variation is the incorporation of hard, calcified lumps, called sesamoid bones (for example, the kneecap), within a tendon.

Function. Tendons transmit force. They allow the force from the muscle to be applied in a restricted region. For example, the main muscles of the fingers are in the forearm, with tendons to the fingertips. If the hand had to accommodate these muscles, it would be too plump to be functional. Tendon extension can also be significant in the movement of a joint. For example, the tendon which flexes a human thumb joint is about 7 in. (170 mm) long. The maximum force from its muscle stretches this tendon about 0.1 in. (2.9 mm), which corresponds to rotation of the joint through an angle of about 21°. *See* JOINT (ANATOMY).

Some tendons save energy by acting as springs. In humans, the Achilles tendon reduces the energy needed for running by about 35%. This tendon is stretched during the first half of each step, storing energy which is then returned during takeoff. This elastic energy transfer involves little energy loss, whereas the equivalent work done by muscles would require metabolic energy in both stages.

When a joint is misplaced, its ligaments may be overloaded and torn (for example, a sprained ankle). Tendons, in general, are much less at risk. Since they are in series with muscles, the maximum force which can be applied to a tendon is limited by the force its muscle can exert. Only those tendons which serve as springs are thin enough for the stress to reach far into the linear region, and are at risk from an excessive tensile load. For this reason, the Achilles tendon is the main human tendon prone to rupture. *See* CONNECTIVE TISSUE; MUSCULAR SYSTEM; SKELETAL SYSTEM. Robert F. Ker

Bibliography. R. McN. Alexander, *Exploring Biomechanics: Animals in Motion*, 1992; V. C. Mow and W. C. Hayes (eds.), *Basic Orthopaedic Biomechanics*, 2d ed., 1997.

Tenrec

An insectivorous mammal indigenous to Madagascar. There are 30 species in 10 genera. These animals are nocturnally active and feed on insects, worms, and mollusks. All tenrecs are essentially primitive unspecialized mammals, with poor vision (see **illustration**). The digits are clawed, and the first digit is not opposable to the others. Some species, such as the tailless or common tenrec (*Tenrec ecaudatus*), become dormant during the hot dry season. The largest tenrecs are 12–16 in. (30–40 cm) long. The body of the tenrec is covered with a mixture of hair, spines, and bristles; the tail is rudimentary; and the toes may be separate or webbed, depending on the species. These animals exude a strong, offensive odor. The dental formula is I 2/3 C 1/1 Pm 3/3 M 4/3, for a total of 40 teeth. Tenrecs are adapted to climbing and swimming. The female is prolific, with litters of 12–20 young, and has 22 mammary glands to feed the offspring. *See* DENTITION; SCENT GLAND.

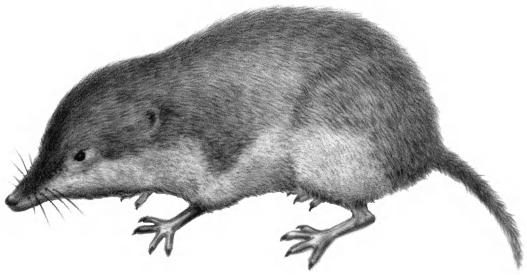

The tenrec is a small insectivorous mammal.

The hedgehog or spring tenrec (*Setifer tenrec*) has stiff hairs that tend to be erected when the animal is disturbed. Some species may curl up into a ball when disturbed, as does the hedgehog. The rice tenrecs (*Oryzorictes talpoides* and *O. tetradactylus*) are fossorial, molelike species which do considerable damage to crops when they burrow for insects. *Microgale*, with 19 species, is the largest tenrec genus. *See* INSECTIVORA; MAMMALIA. Charles B. Curtin

Bibliography. R. M. Nowak, *Walker's Mammals of the World*, Johns Hopkins University Press, 1999.

Tensor analysis

The systematic study of tensors which led to an extension and generalization of vectors, begun in 1900 by two Italian mathematicians, G. Ricci and T. Levi-Civita, following G. F. B. Riemann's proposal concerning a generalization of Euclidean geometry. The principal aim of the tensor calculus (absolute differential calculus) is to construct relationships which are generally covariant in the sense that these relationships or laws remain valid in all coordinate systems. The differential equations for the geodesics in a Riemannian space are covariant expressions; they yield a description of the geodesics which is valid

for all coordinate systems. On the other hand, Newton's equations of motion require a preferred coordinate system for their description, namely, one for which force is proportional to acceleration (an inertial frame of reference). Thus Albert Einstein was led to a study of Riemannian geometry and the tensor calculus in order to construct the general theory of relativity.

Arithmetic, or vector, n-space. The coordinates of a point in a three-dimensional Euclidean space are given by the triple of numbers (x,y,z) or (x^1,x^2,x^3), with $x = x^1$, $y = x^2$, $y = x^2$, $z = x^3$. The superscripts in x^i, $i = 1, 2, 3$, are simply labels which enable one to distinguish and order the various elements in the triple of numbers. The totality of all number triples of the form (x^1,x^2,x^3) with the x^i, $i = 1, 2, 3$, real, yields the arithmetic 3-space, designated as V_3. A simple generalization yields the arithmetic n-space, V_n. This space or manifold consists of all n-tuples of the form $(x^1,x^2, \ldots, x^n)$, the x^i, $i = 1, 2, \ldots, n$, taken as real numbers. Now a space of n dimensions is defined as any set of objects which can be put in 1:1 reciprocal correspondence with the arithmetic n-space. The 1:1 correspondence between the elements or points of the n-space can be chosen in many ways, and in general, the choice depends on the nature of the physical problem. In the special theory of relativity, an event is specified by the three space coordinates and the time, so that each event corresponds to a four-tuple $(x^1,x^2,x^3, x^4 = ct)$.

Let a point P of an n-space correspond to the n-tuple $(x^1,x^2, \ldots, x^n)$. Now consider the n equations, (1), and assume that one can solve for each x^i, yielding Eq. (2). It is assumed that Eqs. (1) and

$$y^i = y^i(x^1, x^2, \ldots, x^n) \quad i = 1, 2, \ldots, n \tag{1}$$

$$x^i = x^i(y^1, y^2, \ldots, y^n) \quad i = 1, 2, \ldots, n \tag{2}$$

(2) are single valued, and that the partial derivatives $\partial y^i/\partial x^i$, $\partial x^i/\partial y^i$, where $i, j = 1, 2, \ldots, n$, exist and are continuous. The n-space of which P is an element can also be put into 1:1 correspondence with the arithmetic n-space $(y^1,y^2, \ldots, y^n)$. Thus, one has a new description of every point in the n space. The points or elements of the n-space have not changed; one merely has a new description (coordinate system) of these points. Thus, Eq. (1) is called a transformation of coordinates.

The algebra of tensors is simplified by use of Einstein's summation convention. From the calculus, Eq. (1) yields Eq. (3). The index α occurs just

$$\frac{\partial y^i}{\partial y^j} = \sum_{\alpha=1}^{n} \frac{\partial y^i}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial y^j} = \frac{\partial y^i}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial y^j} \tag{3}$$

twice in the last expression of Eq. (3), thus indicating a summation over this index. In a space of three dimensions, the expression $S = a_{\alpha\beta}x^\alpha x^\beta$ requires a double summation because the indices α and β occur exactly twice. Thus,

$$S = \sum_{\beta=1}^{3}\sum_{\alpha=1}^{3} a_{\alpha\beta}x^\alpha x^\beta$$

and Einstein found it convenient to remove the summation signs. A further convenience occurs if the Kronecker delta is introduced. Thus, δ^i_j, $i,j = 1, 2, \ldots, n$, is a set of quantities whose numerical values are given by Eq. (4).

$$\delta^i_j = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} \qquad i, j = 1, 2, \ldots, n \tag{4}$$

In particular

$$\delta^1_1 = \delta^2_2 = \cdots = \delta^n_n = 1$$
$$\delta^1_2 = \delta^1_3 = \cdots = \delta^{n-1}_n = 0$$

and for an n-space

$$\delta^\alpha_\alpha = \delta^1_1 + \delta^2_2 + \cdots + \delta^n_n = n$$

Equation (3) may be written

$$\delta^i_j = \frac{\partial y^i}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial y^i}$$

This expression yields

$$\left|\frac{\partial y^i}{\partial x^j}\right| \left|\frac{\partial x^\alpha}{\partial y^\beta}\right| = 1$$

applying the rule for the product of two determinants, with

$$\left|\frac{\partial y^i}{\partial x^j}\right|$$

the Jacobian of the transformation in Eq. (1), and the

$$\left|\frac{\partial x^i}{\partial y^j}\right|$$

Jacobian of the inverse transformation in Eq. (2).

Contravariant vectors. In an n-space the locus of elements given by Eq. (5) represents a space curve Γ with λ a parameter. The n-tuple in notation (6), des-

$$x^i = x^i(\lambda) \quad i = 1, 2, \ldots, n \quad \lambda_0 \leq \lambda \leq \lambda_1 \tag{5}$$

$$\left(\frac{dx^1}{d\lambda}, \frac{dx^2}{d\lambda}, \ldots, \frac{dx^n}{d\lambda}\right) \tag{6}$$

ignated by $dx^i/d\lambda$, is defined to be a tangent element to the space curve, Eq. (5). Under an allowable coordinate transformation, Eq. (1), the space curve Γ can be represented by Eq. (7). The components of the

$$y^i = y^i\left(x^1(\lambda), x^2(\lambda), \ldots, x^n(\lambda)\right) = y^i(\lambda) \tag{7}$$

$$i = 1, 2, \ldots, n$$

tangent vector for the $y = (y^1,y^2, \ldots, y^n)$ coordinate

system are given by notation (8), which represent the elements of the n-tuple, notation (6).

$$\frac{dy^i}{d\lambda} \qquad i = 1,\ 2,\ \ldots, n \tag{8}$$

Certainly the x-coordinate system is no more important than the y-coordinate system as a description of the space curve Γ. One cannot say that $dx^i/d\lambda$ is the tangent vector; nor can one say that $dy^i/d\lambda$ is the tangent vector. If one considers all allowable coordinate systems, one obtains the entire class of tangent elements, each element (n-tuple) claiming to be the tangent vector for its particular coordinate system. It is the abstract collection of all these tangent elements that is specified as the tangent vector. To find what relationship exists between the components of the tangent vector when described by two different coordinate systems, Eq. (7) can be used to give Eq. (9).

$$\frac{dy^i}{d\lambda} = \frac{\partial y^i}{\partial x^\alpha}\frac{dx^\alpha}{d\lambda} \qquad i = 1,\ 2,\ \ldots, n \tag{9}$$

Note that at any point each component $dy^i/d\lambda$ depends linearly on every component of $dx^i/d\lambda$. Moreover, Eq. (10) holds.

$$\frac{dx^i}{d\lambda} = \frac{\partial x^i}{\partial y^\alpha}\frac{dy^\alpha}{d\lambda} \qquad i = 1,\ 2,\ \ldots, n \tag{10}$$

One can now form the following generalization: Any set of numbers $A^i(x^1,x^2,\ \ldots,\ x^n)$, $i = 1,\ 2,\ \ldots,\ n$, which transform according to the law shown by Eq. (11) under the coordinate transformation

$$\overline{A}^i(\overline{x}^1, \overline{x}^2,\ \ldots, \overline{x}^n) = A^\alpha(x^1, x^2,\ \ldots, x^n)\frac{\partial \overline{x}^i}{\partial x^\alpha} \tag{11}$$

$$i = 1,\ 2,\ \ldots, n$$

$\overline{x}^i = \overline{x}(\overline{x}^1,\ x^2,\ \ldots,\ x^n)$, $i = 1,\ 2,\ \ldots,\ n$, is said to be a contravariant vector. The vector is not just the set of components in any coordinate system, but is rather the abstract element which is represented in each coordinate system x by the set of numbers $A^i(x)$. The $\overline{A}^i(\overline{x})$ depend linearly on the $A^\alpha(x)$; therefore the sum and difference of two vectors in an n-space are also vectors.

One immediately sees that the law of transformation for a contravariant vector is transitive. Let

$$\overline{A}^i = A^\alpha \frac{\partial \overline{\overline{x}}^i}{\partial x^\alpha} \qquad \overline{\overline{A}}^i = \overline{A}^\alpha \frac{\partial \overline{\overline{x}}^i}{\partial \overline{x}^\alpha}$$

so that

$$\overline{\overline{A}}^i = \overline{A}^\beta \frac{\partial \overline{\overline{x}}^i}{\partial \overline{x}^\beta} = A^\alpha \frac{\partial \overline{x}^\beta}{\partial x^\alpha}\frac{\partial \overline{\overline{x}}^i}{\partial \overline{x}^\beta} = A^\alpha \frac{\partial \overline{\overline{x}}^i}{\partial x^\alpha}$$

which proves the statement.

If the components of a contravariant vector are known in one coordinate system, the components are known in all allowable coordinate systems from Eq. (11). A coordinate transformation does not yield a new vector; it merely changes the components of the same vector. Thus, a vector is said to be invariant under a coordinate transformation. Any object which is not changed under a coordinate transformation is called an invariant.

If $\phi(x^1,x^2,\ \ldots,\ x^n)$ is a scalar point function $\phi(x^1,x^2,\ldots,x^n) = \phi(x^1(\overline{x}),x^2(\overline{x}),\ \ldots,\ x^n(\overline{x})) = \overline{\phi}(\overline{x}^2, \overline{x}^2,\ \ldots,\ \overline{x}^n)$, then Eq. (12) can be derived.

$$\frac{\partial \overline{\phi}}{\partial \overline{x}^i} = \frac{\partial \phi}{\partial x^\alpha}\frac{\partial x^\alpha}{\partial \overline{x}^i} \tag{12}$$

The n-tuple

$$\left(\frac{\partial \phi}{\partial x^1}, \frac{\partial \phi}{\partial x^2}, \ldots, \frac{\partial \phi}{\partial x^n}\right)$$

yields the components of a covariant vector, called the gradient of ϕ.

More generally Eq. (13) holds, if one says that the

$$\overline{A}_i(\overline{x}^1, \overline{x}^2,\ \ldots, \overline{x}^n) = A(x^1, x^2, \ldots, x^n)\frac{\partial x^\alpha}{\partial \overline{x}^i} \tag{13}$$

$$i = 1,\ 2,\ \ldots, n$$

$A_\alpha(x)$ are the components of a covariant vector.

The law of transformation, Eq. (13), differs from that of Eq. (11) because in general,

$$\frac{\partial x^\alpha}{\partial \overline{x}^i} \neq \frac{\partial \overline{x}^i}{\partial x^\alpha}$$

However, for the group of orthogonal transformations $\overline{x}^i = a^i_\alpha x^\alpha$ such that $\mathbf{A} = \|a^i_j\|$ satisfies $\mathbf{A}^T = \mathbf{A}^{-1}$, one can show that $\partial \overline{x}^i/\partial x^\alpha = \partial x^\alpha/\partial \overline{x}^i$, and for this reason no distinction occurs between contravariant and covariant vectors in elementary vector analysis.

An important scalar invariant can be formed from a contravariant vector A^i and a covariant vector B_i. From their laws of transformation it follows that Eq. (14) holds.

$$\overline{A}^i\overline{B}_i = A^\alpha B_\beta \frac{\partial \overline{x}^i}{\partial x^\alpha}\frac{\partial x^\beta}{\partial \overline{x}^i} = A^\alpha B_\beta \delta_\alpha{}^\beta$$

$$= A^\alpha B_\alpha = A^i B_i \tag{14}$$

Thus, the expression

$$A^i B_i = \sum_{i=1}^{n} A^i B_i$$

is invariant both in form and in its numerical value under a transformation of coordinates. One calls $A^\alpha B^\alpha$ the scalar, or inner, product, of the two vectors A^i, B_i.

Tensors. The contravariant and covariant vectors discussed above are special cases of differential invariants called tensors. The components of a tensor are of the form

$$T^{a_1 a_2\ \cdots\ a_r}_{b_1 b_2\ \cdots\ b_s}(x^1, x^2,\ \ldots,\ x^n)$$

and indices $a_1, a_2,\ \ldots,\ a_r, b_1, b_2,\ \ldots,\ b_s$ run through the integers $1, 2, \ldots, n$. The components transform according to the rule shown by Eq. (15).

$$\overline{T}^{a_1 a_2 \ldots a_r}_{b_1 b_2 \ldots b_s} = \left|\frac{\delta x}{\delta \overline{x}}\right|^N T^{\alpha_1 \alpha_2 \ldots \alpha_r}_{\beta_1 \beta_2 \ldots \beta_s} \frac{\partial \overline{x}^{\alpha 1}}{\partial x^{\alpha 1}} \cdots \frac{\partial \overline{x}^{ar}}{\partial x^{\alpha r}}\frac{\partial x^{\beta 1}}{\partial \overline{x}^{b1}} \cdots \frac{\partial x^{\beta s}}{\partial \overline{x}^{bs}} \tag{15}$$

The exponent N of the Jacobian

$$\left|\frac{\partial x}{\partial \overline{x}}\right|$$

is called the weight of the tensor field. For $N = 0$ the tensor field is absolute; otherwise it is of weight N. A tensor density occurs for $N = 1$. The number of indices is $r + s$, the rank of the tensor. Vectors are tensors of rank one. The tensor of Eq. (15) is contravariant of order r and covariant of order s.

An important property of tensors is immediately evident from Eq. (15). If every component of a tensor vanishes in one coordinate system, the components vanish in all coordinate systems.

Two tensors are said to be of the same kind if they have the same order in their contravariant and covariant indices and if they are of the same weight. Additional tensors can be constructed as follows:

1. The sum and difference of two tensors of the same kind yields a new tensor of the same kind. This is apparent from the linear property of Eq. (15).

2. The product of two tensors is a new tensor. This can be shown for a special case. Let

$$\overline{T}^a_b = \left|\frac{\partial x}{\partial \overline{x}}\right| T^\alpha_\beta \frac{\partial x^\beta}{\partial \overline{x}^b}\frac{\partial \overline{x}^a}{\partial x^\alpha} \qquad \overline{S}^c = \left|\frac{\partial x}{\partial \overline{x}}\right|^3 S^\gamma \frac{\partial \overline{x}^c}{\partial x^\gamma}$$

Then

$$\overline{W}^{ac}_{\ \ b} \equiv \overline{T}^a_b \overline{S}^c = \left|\frac{\partial x}{\partial \overline{x}}\right|^4 (T^\alpha_\beta S^\gamma)\frac{\partial \overline{x}^a}{\partial x^\alpha}\frac{\partial \overline{x}^c}{\partial x^\gamma}\frac{\partial x^\beta}{\partial \overline{x}^b}$$

$$= \left|\frac{\partial x}{\partial \overline{x}}\right|^4 W^{\alpha\gamma}_{\ \ \beta}\frac{\partial \overline{x}^a}{\partial x^\alpha}\frac{\partial \overline{x}^c}{\partial x^\gamma}\frac{\partial x^\beta}{\partial \overline{x}^b}$$

To illustrate the concept of contraction, consider the absolute tensor

$$\overline{A}^{ij}_{kl} = A^{\alpha\beta}_{\sigma\tau}\frac{\partial \overline{x}^i}{\partial x^\alpha}\frac{\partial \overline{x}^j}{\partial x^\beta}\frac{\partial x^\sigma}{\delta \overline{x}^k}\frac{\partial x^\tau}{\partial \overline{x}^l}$$

Replacing k by i and summing yields

$$\begin{aligned}\overline{B}^j_l \equiv \overline{A}^{ij}_{il} &\equiv A^{\alpha\beta}_{\sigma\tau}\frac{\partial \overline{x}^i}{\partial x^\alpha}\frac{\partial \overline{x}^j}{\partial x^\beta}\frac{\partial x^\sigma}{\partial \overline{x}^i}\frac{\partial x^\tau}{\partial \overline{x}^l}\\ &= A^{\alpha\beta}_{\sigma\tau}\frac{\partial x^\sigma}{\partial x^\alpha}\frac{\partial \overline{x}^j}{\partial x^\beta}\frac{\partial x^\tau}{\partial \overline{x}^l}\\ &= A^{\alpha\beta}_{\sigma\tau}\delta^\sigma_\alpha\frac{\partial \overline{x}^j}{\partial x^\beta}\frac{\partial x^\tau}{\partial \overline{x}^l}\\ &= A^{\sigma\beta}_{\sigma\tau}\frac{\partial \overline{x}^j}{\partial x^\beta}\frac{\partial x^\tau}{\partial \overline{x}^l} = B^\beta_\tau\frac{\partial \overline{x}^j}{\partial x^\beta}\frac{\partial x^\tau}{\partial \overline{x}^l}\end{aligned}$$

so that $B^\beta_\tau \equiv A^{\sigma\beta}_{\sigma\tau}$ is a mixed tensor. In general, one equates a certain covariant index with a contravariant index, sums on this repeated index, and obtains a new tensor whose rank is two less than that of the original tensor. This process of producing a new tensor is called the method of contraction.

A few examples of tensors can be listed.

1. The Kronecker delta is a mixed absolute tensor because

$$\delta^i_j\frac{\partial x^j}{\partial \overline{x}^\alpha}\frac{\partial \overline{x}^\beta}{\partial x^i} = \frac{\partial x^i}{\partial \overline{x}^\alpha}\frac{\partial \overline{x}^\beta}{\partial x^i} = \frac{\partial \overline{x}^\beta}{\partial \overline{x}^\alpha} = \overline{\delta}^\beta_\alpha$$

2. If A_i and B_i are absolute covariant vectors, then

$$\overline{C}_{ij} \equiv \overline{A}_i\overline{B}_j = A_\alpha B_\beta\frac{\partial x^\alpha}{\partial \overline{x}^i}\frac{\partial x^\beta}{\partial \overline{x}^j} = C_{\alpha\beta}\frac{\partial x^\alpha}{\partial \overline{x}^i}\frac{\partial x^\beta}{\partial \overline{x}^j}$$

so that $C_{\alpha\beta} \equiv A_\alpha B_\beta$ are the components of an absolute covariant tensor of rank two.

3. Let ϕ_i be the components of an absolute covariant vector. From $\overline{\phi}_i = \phi_\alpha(\partial x^\alpha/\partial \overline{x}^i)$, it follows that

$$\begin{aligned}\overline{F}_{ij} &\equiv \frac{\partial \overline{\phi}_i}{\partial \overline{x}^j} - \frac{\partial \overline{\phi}_j}{\partial \overline{x}^i} = \left(\frac{\partial \phi_\alpha}{\partial x^\beta} - \frac{\partial \phi_\beta}{\partial x^\alpha}\right)\frac{\partial x^\alpha}{\partial \overline{x}^i}\frac{\partial x^\beta}{\partial \overline{x}^j}\\ &= F_{\alpha\beta}\frac{\partial x^\alpha}{\partial \overline{x}^i}\frac{\partial x^\beta}{\partial \overline{x}^j}\end{aligned}$$

so that

$$F_{\alpha\beta} \equiv \frac{\partial \phi_\alpha}{\partial x^\beta} - \frac{\partial \phi_\beta}{\partial x^\alpha}$$

are the components of an absolute covariant tensor. Note that $F_{\alpha\beta} = -F_{\beta\alpha}$, so that the tensor is skew symmetric. If the ϕ_i, $i = 1, 2, 3, 4$, are the components of the electromagnetic vector potential, the F_{ij} are those of the electromagnetic field tensor.

Line element of Riemannian geometry. A simple generalization of the euclidean metric, or line element, $ds^2 = dx^2 + dy^2 + dz^2$, led Riemann to consider an n-space such that a metric is imposed on the space which yields the invariant distance ds between two nearby points whose coordinates differ by dx^i. The line element in a riemannian n-space is given by Eq. (16).

$$ds^2 = g_{\alpha\beta}(x^1, x^2, \ldots, x^n)\, dx^\alpha dx^\beta \qquad (16)$$

$$g_{\alpha\beta} = g_{\beta\alpha}$$

Under a coordinate transformation, $x^\alpha = x^\alpha(\overline{x})$, Eq. (16) becomes Eq. (17).

$$ds^2 = g_{\alpha\beta}\frac{\partial x^\alpha}{\partial \overline{x}^\sigma}\frac{\partial x^\beta}{\partial \overline{x}^\tau}\, d\overline{x}^\sigma\, dx^\tau = \overline{g}_{\sigma\tau}\, d\overline{x}^\sigma d\overline{x}^\tau \qquad (17)$$

$$\overline{g}_{\sigma\tau} = g_{\alpha\beta}\frac{\partial x^\alpha}{\partial \overline{x}^\sigma}\frac{\partial x^\beta}{\partial \overline{x}^\tau}$$

Hence the $g_{ij}(x)$ are the components of an absolute covariant tensor of rank two. One can show that the $g^{ij}(x)$ defined by $g^{ij}g_{jk} = \delta^i_k$ are the components of an absolute contravariant tensor, provided $|g_{ij}| \neq 0$. Furthermore, it can be shown that

$$|\overline{g}_{ij}| = |g_{ij}|\left|\frac{\partial x}{\partial \overline{x}}\right|^2$$

Thus, if A^i are the components of an absolute contravariant vector, then $|g_{\alpha\beta}|^{1/2}A^i$ are the components of a contravariant vector density.

Geodesics in a Riemannian space. If a space curve in a Riemannian n-space is given by $x^i = x^i(\lambda)$, $\lambda_0 \leqq \lambda \leqq \lambda_1$, the distance along the curve between the end points is given by Eq. (18).

$$L = \int_{\lambda_0}^{\lambda_1}\left(g_{\alpha\beta}\frac{dx^\alpha}{d\lambda}\frac{dx^\beta}{d\lambda}\right)^{1/2} d\lambda \qquad (18)$$

The geodesic path between two fixed points of a Riemannian space is that curve joining the two points which yields a minimum value for the integral

of Eq. (18). The same path is obtained if one minimizes, as shown in notation (19).

$$\int_{P_0}^{P_1} \left(g_{\alpha\beta} \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} \right) ds \tag{19}$$

Applying the Euler-Lagrange equations of the calculus of variations,

$$\frac{d}{ds}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = \frac{\partial L}{\partial x^i}$$

with $L = g_{\alpha\beta}\dot{x}^\alpha \dot{x}^\beta$, yields the differential equations of the geodesics, Eq. (20), with Eq. (21) defining a term.

$$\frac{d^2x^i}{ds^2} + \Gamma^i_{jk}(x)\frac{dx^j}{ds}\frac{dx^k}{ds} = 0 \tag{20}$$

$$\Gamma^i_{jk} = \frac{1}{2} g^{i\sigma}\left(\frac{\partial g_{\sigma j}}{\partial x^k} + \frac{\partial g_{k\sigma}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^\sigma}\right) \tag{21}$$

The important elements Γ^i_{jk} are called the Christoffel symbols of the second kind. The elements $\{i,jk\} = g_{i\sigma}\Gamma^\sigma_{jk}$ are the Christoffel symbols of the first kind. Under a coordinate transformation, the Christoffel symbols transform according to the rule expressed in Eq. (22).

$$\overline{\Gamma}^i_{jk}(\overline{x}) = \Gamma^\alpha_{\beta\gamma}(x)\frac{\partial x^\beta}{\partial x^j}\frac{\partial x^\gamma}{\partial \overline{x}^k}\frac{\partial \overline{x}^i}{\partial x^\alpha} + \frac{\partial^2 x^\alpha}{\partial \overline{x}^j \partial \overline{x}^k}\frac{\partial \overline{x}^i}{\partial x^\alpha} \tag{22}$$

Thus, the Christoffel symbols are not the components of a mixed tensor. However, if the $\Gamma^i_{jk}(x)$ are given in an x-coordinate system, their values can be computed in an $\overline{x}$-coordinate system from Eq. (22).

Covariant differentiation. From Eq. (11) follows Eq. (23), so that $\partial A^\alpha/\partial x^\beta$ are not the components

$$\frac{\partial \overline{A}^i}{\partial \overline{x}^j} = \frac{\partial A^\alpha}{\partial x^\beta}\frac{\partial \overline{x}^i}{\partial x^\alpha}\frac{\partial x^\beta}{\partial \overline{x}^j} + A^\alpha \frac{\partial^2 \overline{x}^i}{\partial x^\beta \partial x^\alpha}\frac{\partial x^\beta}{\partial \overline{x}^j} \tag{23}$$

of a mixed tensor. This is not surprising; in a euclidean space using coordinates other than rectangular coordinates, the unit vectors change directions from point to point. A differentiation process which yields a new tensor can be introduced by combining Eqs. (22) and (23). It can be shown, Eq. (24), that

$$A^i_{,j} \equiv \frac{\partial A^i}{\partial x^j} + A^i_\alpha \Gamma^i_{\alpha j} \tag{24}$$

the terms are the components of a mixed tensor. The term $A^i{}_j$ is called the covariant derivative of A^i (the comma denoting covariant differentiation).

In a euclidean space using rectangular coordinates, the $\Gamma^i{}_{jk}$ vanish from Eq. (21), and the covariant derivative reduces to the ordinary partial derivative.

$$ds^2 = \sum_{i=1}^{n} (dx^i)^2$$

The intrinsic derivative of A^i along a path $x^i = x^i(s)$, with s arc length, is given by Eq. (25).

$$\frac{\delta A^i}{\delta s} \equiv A^i_{,j}\frac{dx^j}{ds} = \frac{dA^i}{ds} + A^\alpha \Gamma^i_{\alpha j}\frac{dx^j}{ds} \tag{25}$$

The covariant derivative of the general tensor given by Eq. (15) is Eq. (26).

$$\begin{aligned} T^{\alpha_1\alpha_2\cdots\alpha_r}_{\beta_1\beta_2\cdots\beta_s,j} &= \frac{\partial T^{\alpha_1\alpha_2\cdots\alpha_r}_{\beta_1\beta_2\cdots\beta_s}}{\partial x^j} + T^{\mu\alpha_2\cdots\alpha_r}_{\beta_1\beta_2\cdots\beta_s}\Gamma^{\alpha_1}_{\mu j} \\ &+ \cdots + T^{\alpha_1\alpha_2\cdots\alpha_{r-1}\mu}_{\beta_1\beta_2\cdots\beta_s}\Gamma^{\alpha_r}_{\mu j} - T^{\alpha_1\alpha_2\cdots\alpha_r}_{\mu\beta_2\cdots\beta_s}\Gamma^{\mu}_{\beta_1 j} \\ &- \cdots - T^{\alpha_1\alpha_2\cdots\alpha_r}_{\beta_1\beta_2\cdots\beta_{s-1}\mu}\Gamma^{\mu}_{\beta_s j} - NT^{\alpha_1\alpha_2\cdots\alpha_r}_{\beta_1\beta_2\cdots\beta_s}\Gamma^{\mu}_{\mu j} \end{aligned} \tag{26}$$

A few examples of covariant differentiation now follow:

1. Let A_i be an absolute covariant vector. The curl of A_i is defined by Eq. (27).

$$\begin{aligned} \operatorname{curl} A_i &= A_{i,j} - A_{j,i} \\ &= \frac{\partial A_i}{\partial x^j} - A_\alpha \Gamma^\alpha_{ij} - \frac{\partial A_j}{\partial x^i} + A_\alpha \Gamma^\alpha_{ji} \\ &= \frac{\partial A_i}{\partial x^j} - \frac{\partial A_j}{\partial x^i} \end{aligned} \tag{27}$$

2. The divergence of A^i is defined by Eq. (28) with

$$\begin{aligned} \operatorname{div} A^i &= A^\alpha_{,\alpha} = \frac{\partial A^\alpha}{\partial x^\alpha} + A^\alpha \Gamma^\beta_{\alpha\beta} \\ &= \frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^\alpha}(= \sqrt{|g|}\, A^\alpha) \end{aligned} \tag{28}$$

$|g| = |g|_{ij}|$. The proof of Eq. (28) is omitted.

3. If ϕ is an absolute scalar, the gradient of ϕ is defined by $\phi_{,i} = \partial\phi/\partial x^i$. The associated vector of $\phi_{,i}$ is $g^{\alpha\beta}\phi_{,\beta} = g^{\alpha\beta}(\partial\phi/\partial x^\beta) = A^\alpha$. Applying Eq. (28) to A^α yields the Laplacian of ϕ given by Eq. (29).

$$\operatorname{Lap}\phi = \frac{1}{\sqrt{|g|}}\frac{\partial}{\partial x^\alpha}\left(\sqrt{|g|}g^{\alpha\beta}\frac{\partial\phi}{\partial x^\beta}\right) \tag{29}$$

See CALCULUS OF VECTORS; RIEMANNIAN GEOMETRY.

Harry Lass

Bibliography. R. Abraham, J. E. Marsden, and T. Ratiu, *Manifolds, Tensor Analysis and Applications*, 1983, reprint 1988; C. T. Dodson and T. Poston, *Tensor Geometry: The Geometric Viewpoint and Its Uses*, 2d ed., 1991; A. W. Joshi, *Matrices and Tensors in Physics*, 3d ed., 1995; D. C. Kay, *Schaum's Outline of Tensor Calculus,* 1988; J. L. Synge and A. Schild, *Tensor Calculus*, 1969, reprint 1978; R. Wasserman, *Tensors and Manifolds: With Applications to Mechanics and Relativity*, 1992; E. C. Young, *Vector and Tensor Analysis*, 2d ed., 1992.

Terbium

Element number 65, terbium, Tb, is a very rare metallic element of the rare-earth group. Its atomic weight is 158.924, and the stable isotope ^{159}Tb makes up

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

100% of the naturally occurring element. *See* PERIODIC TABLE.

The common oxide, Tb_4O_7, is brown and is obtained when its salts are ignited in air. Its salts are all trivalent and white in color and, when dissolved, give colorless solutions. The higher oxides slowly decompose when treated with dilute acid to give the trivalent ions in solution. Although the metal is attacked readily at high temperatures by air, the attack is extremely slow at room temperatures. The metal has a Néel point at about 229 K and a Curie point at about 220 K. For properties of the metal *see* RARE-EARTH ELEMENTS

Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

Terebratulida

An order of articulated brachiopods consisting of a group of sessile, suspension-feeding, marine, benthic, epifaunal bivalves with representatives occurring from the Early Devonian Era. It is most diverse and abundant group of living brachiopods, which probably exhibits maximum diversity in present-day seas.

The shells are biconvex and usually smooth, although some show radial ribbed ornamentation. The valves articulate about a hinge structure and posterior edges of the valves are not coincident with the hinge axis (nonstrophic condition). The shell is calcareous and punctate. A fleshy pedicle usually attaches the animal to the substrate (**illus.** *a*), but in thecidine brachiopods the ventral valve is cemented to the substrate. Terebratulids have a pair of adductor muscles and a pair of diductor muscles, and a dorsal pair and a ventral pair of adjustor muscles.

The tentacular feeding organ (lophophore) occupies the mantle cavity as a looped structure (the ptycholophe) in the smaller forms or as a looped and coiled structure (the plectolophe) in the larger forms. In the smaller forms the lophophore is supported by a calcareous ridge, but in the larger forms by a calcareous loop. Both of these structures emanate from the dorsal (brachial) valve. The digestive system consists of a mouth, pharynx, esophagus, stomach, digestive diverticula, pylorus, and blind intestine. Digestive diverticula usually occur as a pair of posterior lobes which open to the stomach through one of three pairs of ducts.

The excretory system consists of a pair of ciliated funnels (metanephridia) which during spawning function as gonoducts and allow the discharge of gametes from the coelom into the mantle cavity. Some solid waste may also be ejected through the nephridiopores, enmeshed in mucus, while the main

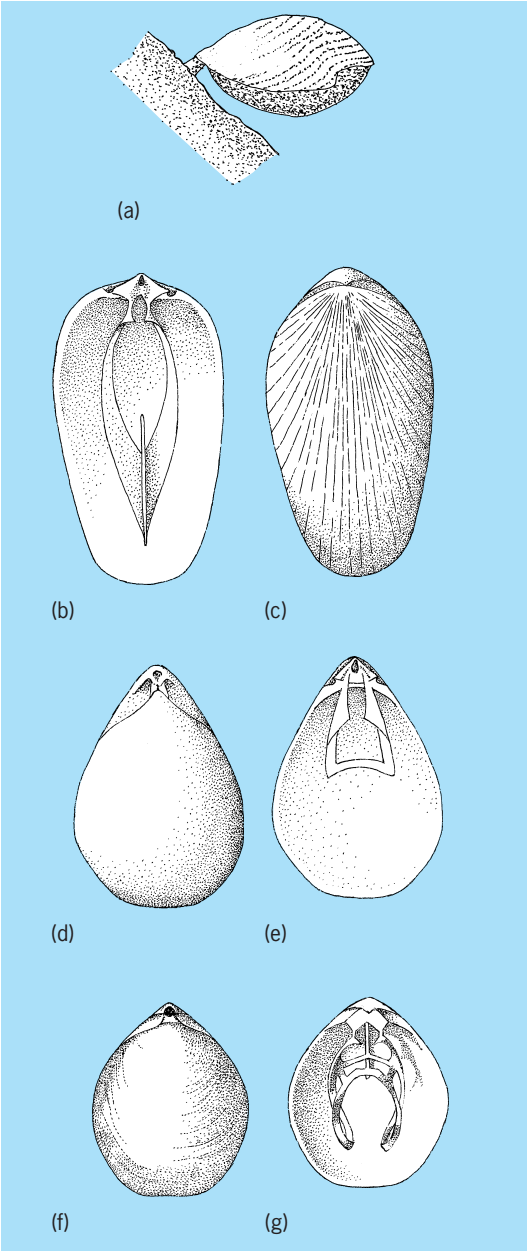

Terebratulida. (*a*) Living *Calloria*, attached to the substrate by the pedicle. (*b, c*) *Rensselaeria* (Centronellida), dorsal view and brachial interior; (*d, e*) *Cranaena* (Terebratulidina), dorsal view and brachial interior; (*f, g*) *Laqueus* (Terebratellidina), dorsal view and brachial interior. (*After R. C. Moore, ed., Treatise on Invertebrate Paleontology, pt. H, Geological Society of America and University of Kansas Press, 1965*)

excretory product, ammonia, is voided through the tissues of the mantle and lophophore.

Terebratulids possess an open circulatory system of blood vessels and coelomic canals, containing fluid which is coagulable and has free cellular inclusions consisting of blood cells and coelomocytes. They have a central nervous system containing unsheathed nerves, but no differentiated sense organs. Members of this order are known to be gonochoristic or hermaphroditic and produce lecithotrophic larvae, which in some species are brooded between the tentacles of the lophophore or in specialized brood chambers.

Three suborders are recognized: Centronellida, which have a primitive calcareous lophophore loop support of oval form (illus. *b* and *c*); Terebratulidina, which have a short calcareous lophophore loop support (illus. *d* and *e*); and Terebratellidina, which have a long calcareous lophophore loop support (illus. *f* and *g*). *See* RHYNCHONELLIFORMEA; BRACHIOPODA.

Mark A. James

Bibliography. M. A. James et al., Biology of living brachiopods, *Adv. Mar. Biol.*, 28:175–387, 1992.

Terpene

A class of natural products having a structural relationship (**1***a*) to isoprene (**1***b*). Over 5000 structurally determined terpenes are known; many of these have also been synthesized in the laboratory. Historically terpenes have been isolated from green plants, but new compounds structurally related to isoprene continue to be isolated from other sources as well, so the class is also referred to as terpenoids, reflecting the biochemical origin without specification of the natural source.

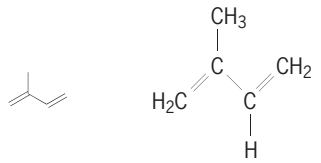

(**1***a*) (**1***b*)

Classification. Terpenes are classified according to the number of isoprene units of which they are composed, as is shown below.

5	hemi-	25	ses-
10	mono-	30	tri-
15	sesqui-	40	tetra-
20	di-	$(5)_n$	poly-

Although they may be named according to the systematic nomenclature and numbering systems set by the International Union of Pure and Applied Chemistry for all organic compounds, it is often easier to refer to terpenes by their common names, which usually reflect the botanical or zoological name of their source.

Biogenesis. The function of terpenes in plants and other organisms is not clear, although they sometimes possess toxic properties linked to the protection of the species. Terpenes are products of secondary metabolism [reaction (1)]; the key building block, isopentenylpyrophosphate (**2**), arises from mevalonic acid (**3**) via hydroxymethyl glutarate (**4**).

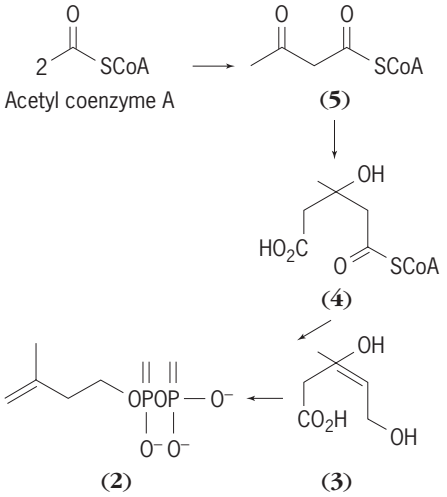

(1)

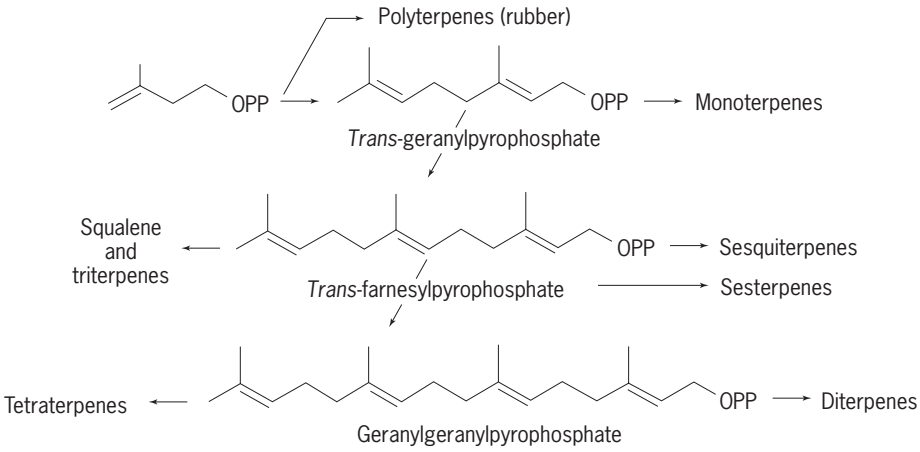

(2)

The starting point of this metabolic pathway is believed to be the condensation of two molecules of acetic acid to form acetoacetyl coenzyme A (**5**). Other biogenetic pathways available for the production of hydroxymethyl glutarate (**4**) depend on the particular organism.

Isopentenylprophosphate (**2**) has two reactive sites and can polymerize in a variety of ways to form larger terpenes. Its head-to-tail dimerization is the most commonly encountered process, although head-to-head and tail-to-tail linkages are also found. The multitude of structural types in terpenes arises from intramolecular rearrangements of several basic skeletons formed by the cyclizations of the linear precursors shown in reactions (2), where OPP represents oxygen-alkylated pyrophosphate. *See* POLYMER.

Monoterpenes. Acyclic or cyclic C_{10} hydrocarbons and their oxygenated derivatives are known as monoterpenes. The most common structural types are biogenetic derivatives of geraniol (**6**), the main constituent of geranium oil. Menthol (**7**) is the chief constituent of peppermint oil; limonene (**8**) composes over 90% of lemon oil; oil of rosemary contains α-pinene (**9**); camphor (**10**) is the main component of sage oil; and other interesting monoterpenes are iridoids, substances which have been isolated from ants, such as iridodial (**11**).

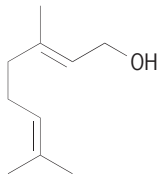

(**6**)

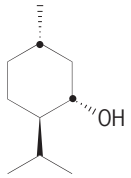

(**7**)

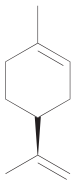

(**8**)

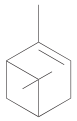

(**9**)

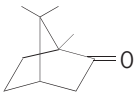

(**10**)

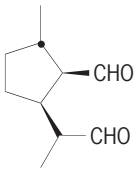

(**11**)

Loganin (**12**; Me = CH_3 and Glu = β-glucose) is the biogenetic precursor of some alkaloids. Its combination with tryptophane, for example, commences the biogenesis of yohimbine alkaloids. Cyclic monoterpenes can be easily oxidized to fully aromatic ring systems such as *p*-cymene (**13**), the constituent of ajowan oil.

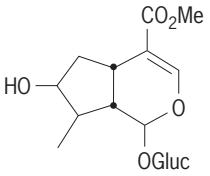

(**12**)

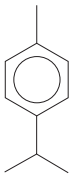

(**13**)

Monoterpenes are widely used in the flavor and perfume industries because of their attractive odors, low molecular weights, and high volatilities. Most are synthesized rather than extracted from plant sources.

Monoterpenes have been well studied, and most of them have been prepared in the laboratory. *See* CAMPHOR; MENTHOL; PINE TERPENE.

Sesquiterpenes. These are C_{15} hydrocarbons or their oxygenated analogs. The major categories are mentioned below. They arise from the cyclization of farnesylpyrophosphate and subsequent rearrangements of the resulting carbonium ions (**14** and **15**), as in reaction (3). Almost all known sesquiterpenes

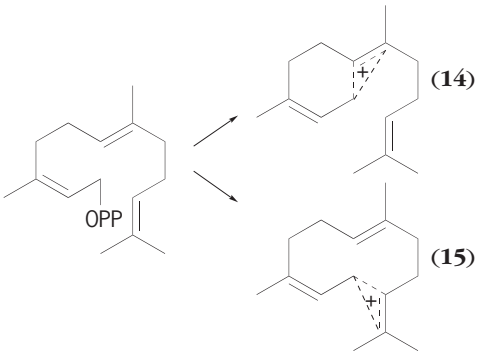

(3)

can be derived from these two cations.

Many sesquiterpenes are important constituents of the characteristic aromas of plant products, and many others have interesting or useful physiological properties.

Farnesanes and eudesmanes. These are mono- and bicylic terpenes derived from the alkylative cyclization of farnesylprophosphate. They can be exemplified by the structures of abscisic acid (**16**), a plant growth regulator, zingiberine (**17**), the constituent of ginger oil, cadinene (**18**) from the oil cubebs, and α-santonine (**19**) from plants of the *Artemisia*.

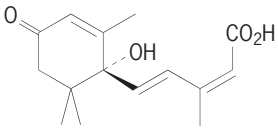

(**16**)

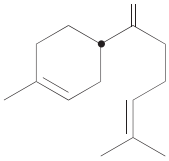

(**17**)

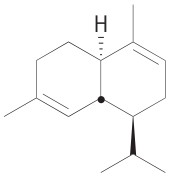

(**18**)

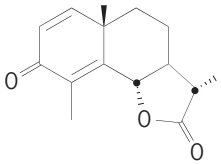

(**19**)

Acoranes, cedranes, chamigrane. These sesquiterpenes are characterized by spirocyclic skeletons. They are obtained from various species of wood, such as the Alaskan cedar. Examples are α-cedrane (**20**), β-acorenol (**21**), and β-chamigrene (**22**).

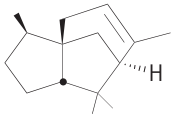

(**20**)

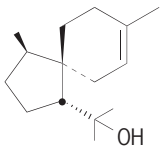

(**21**)

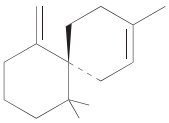

(**22**)

Caryophyllanes, illudanes, humulenes. These are biogenetically related macrocyclic terpenes; they are exemplified by caryophyllene (**23**) and humulene (**24**), which occur in hops. Related to illudanes is hirsutic acid (**25**).

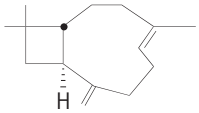

(**23**)

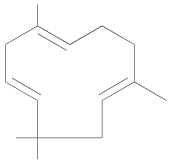

(**24**)

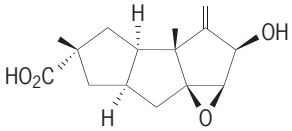

(**25**)

Himachalenes and longifolanes. These contain fused six- and seven-membered rings, and are derived from decalin skeletons via rearrangements; longifolene (**26**) and β-himachalene (**27**) are examples.

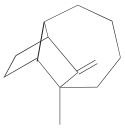

(**26**)

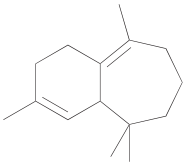

(**27**)

Perhydroazulenes. The fastest-growing group of identified sesquiterpenenes are the perhydroazulenes. The structures of these compounds were formerly determined by their oxidation to derivatives of the blue hydrocarbon azulene (**28**). A representative member of this class is quaiol (**29**) from quaiacum wood oil.

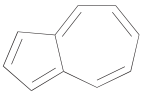

(**28**)

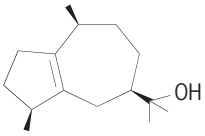

(**29**)

Quaianes and quaianolides. These constitute the largest single group of sesquiterpenes. New species continue to be isolated from fungi, marine organisms, or plants. The sesquiterpene lactones or quaianolides have remarkable cytoxic properties and are represented by a general structure (**30**).

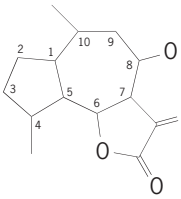

(**30**)

Germacranes. These are macrocyclic terpene lactones, exemplified by eupassopin (**31**). The quaianes, quaianolides, and germacranes have been shown to include a number of compounds with promising antitumor activity.

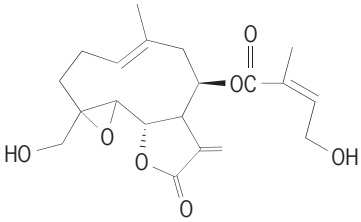

(**31**)

Diterpenes. These are C_{20} hydrocarbons, and their oxidized derivatives are composed of four isoprene units. They are synthesized by organisms from geranylgeranalpyrophosphate and are utilized in protective coatings of higher plants. Extraction of resins of various coniferous species yields a number of diterpenes. Their skeletons do not exhibit the variety shown by sesquiterpenoids. Abietic acid derivatives are the principal constituent of rosin, a resin obtained mainly from pine with a variety of industrial uses. Diterpene frameworks are found in many complex alkaloids of the delphinium or atisine types. Nitrogen is incorporated into the diterpene nucleus via a carboxylic acid function.

Mild antibiotic activity is associated with some diterpenes. The gibberellins are plant growth promoters. The best-known open-chain diterpene is vitamin A (**32**). The structure of other diterpenes

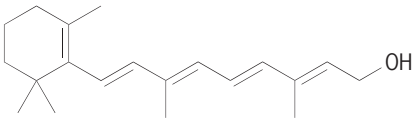

(**32**)

varies from bicyclic to tetracyclic skeletons.

Labdanes and clerodanes. These are bicyclic resin acids isolated as the "bitter principles" of bark, roots, and stems. Agathic acid (**33**), marubiin (**34**) from horehound, and clerodane (**35**) are representatives.

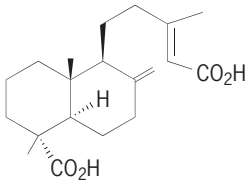

(**33**)

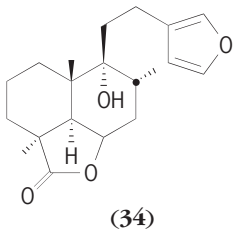

(**34**)

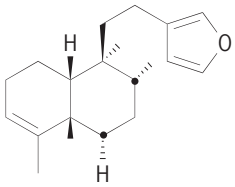

(**35**)

Tricyclic and tetracyclic diterpenes are typified by the structures of abietic acid (**36**; actually an artifact of isolation and not a natural product) from pine

resin, and phyllocladene (**37**), a typical diterpene constitutent of many essential oils.

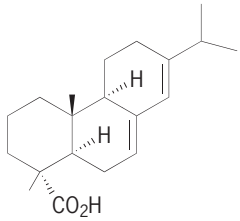

(**36**)

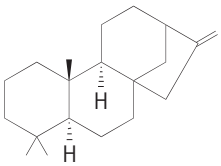

(**37**)

Gibberellins. The metabolites of a rice fungus, known as gibberellins, are exemplified by the structure of gibberellic acid (**38**).

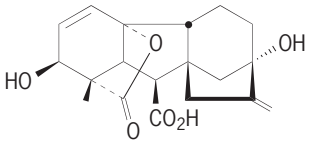

(**38**)

See GIBBERELLIN.

Sestquiterpenes. These are C_{25} compounds derived from geranylfarnesol pyrophosphate. They can be isolated from protective coatings and waxes of insects or from various fungal sources. The first-known member of the sesquiterpenes was isolated in 1965. Some examples are ophiobolin A (**39**), retigeranic acid (**40**), and gastardic acid (**41**).

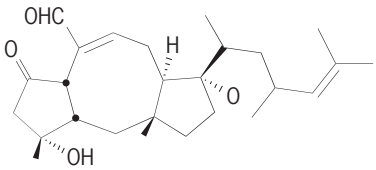

(**39**)

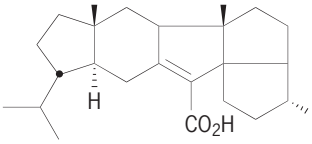

(**40**)

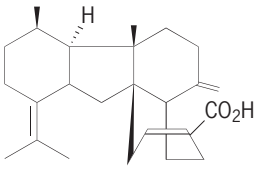

(**41**)

Triterpenes. These have 30 carbon atoms and are composed of six isoprene units. They form the largest group of terpenes. *See* TRITERPENE.

Tetraterpenes. These are C_{40} or higher terpenes. They have a large number of polyene units and therefore do not possess the stereochemical and structural complexity of lower terpenes. Biogenetically, tetraterpenes arise by the dimerization of C_{20} units in a tail-to-tail fashion, with concomitant cyclizations restricted to the termini. Most known tetraterpenes are carotenoid pigments, exemplified by α-carotene (**42**). Carotenes and various oxygenated analogs,

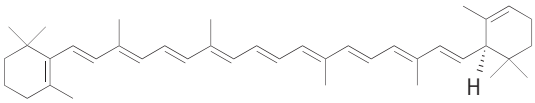

(**42**)

found in the chloroplasts of all green plants and in some algae, serve as accessory pigments for photosynthesis. Vitamin A may be regarded as a degraded carotene. *See* CAROTENOID; VITAMIN A.

Polyterpenes. These are various natural rubbers from *Hevea*, *Guayule*, and other plants and usually consist of 500–5000 cis-linked isoprene units; gutta-percha contains 100 trans-linked isoprene residues. These polymers arise by a genetic error characteristic of only about 1% of the plant kingdom, whereby species possess the ability to channel all of their mevalonate biogenesis into polymerization. *See* ESSENTIAL OILS; ROSIN; RUBBER. Tomas Hudlicky

Bibliography. K. F. Albizati et al., *Synthesis of Marine Natural Products*, vol. 1: *Terpenoids*, 1992; Chemical Society, London, *Terpenoids and Steroids*, vols. 1–8, Specialist Periodical Reports, 1978; J. S. Glasby, *Encyclopedia of the Terpenoids*, 1983; J. B. Harborne and F. A. Tomas-Barberan, *Ecological Chemistry and Biochemistry of Plant Terpenoids*, 1991; T.-L. Ho, *Carbocyclization: Implications in Terpene Synthesis,* 1988.

Terracing (agriculture)

A method of shaping land to control erosion on slopes of rolling land used for cropping and other purposes. In early practice the land was shaped into a series of nearly level benches or steplike formations. Modern practice in terracing, however, consists of the construction of low-graded channels or levees to carry the excess rainfall from the land at nonerosive velocities. The physical principle involved is that, when water is spread in a shallow stream, its flow is retarded by the roughness of the bottom of the channel and its carrying, or erosive, power is reduced. Since direct impact of rainfall on bare land churns up the soil and the stirring effect keeps it in suspension in overland flow and rills, terracing does not prevent sheet erosion. It serves only to prevent destruction of agricultural land by gullying and must be supplemented by other erosion-control practices, such as grass rotation, cover crops, mulching, contour farming, strip cropping, and increased organic matter content.

In areas of low rainfall and absorbent soils, nearly level terraces are used to retain the runoff and conserve soil moisture, thus preventing wind erosion of the soil. *See* EROSION; SOIL CONSERVATION.

Types of terraces. The two major types of terraces are the bench and the broadbase (see **illus.**).

Bench. This is essentially a steep-land terrace and consists of an almost vertical retaining wall, called a riser, or a steep vegetative slope to hold the nearly level surface of the soil for cultivation, orchards, vineyards, or landscaping. The bench terrace has

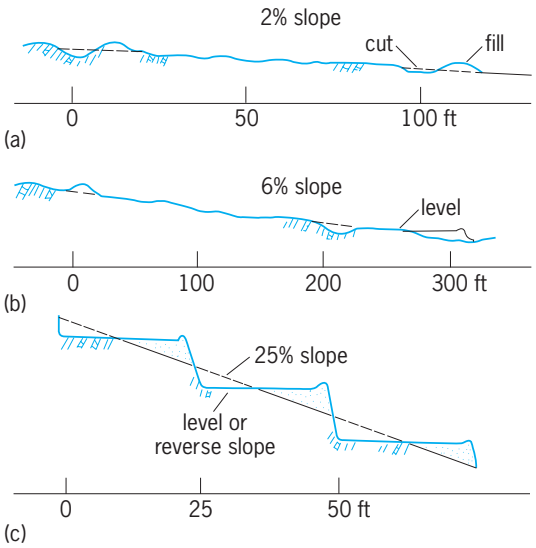

Types of terraces. (*a*) Broadbase. (*b*) Conservation bench. (*c*) Bench. 1 ft = 0.3 m. (*After Soil and Water Conservation Engineering, 2d ed., The Ferguson Foundation Agricultural Engineering Series, John Wiley and Sons, 1966*)

been in use all over the world for the past several thousand years, particularly in Europe, Australia, and Asia. Bench terraces are adaptable to slopes of 25–30% and are costly to construct. For the most part, bench terraces have been abandoned for general farming because they are not adaptable to the efficient use of modern-day farm equipment.

Modern mechanized farming greatly increased the need for erosion control; at the same time, however, modern machinery furnished the power for increased terracing. The early United States practice of constructing hillside ditches across the slopes of fields to prevent up- and downhill gullying was followed by the development of the more easily controlled levee or ridge. This was called a narrow-base ridge terrace and was, in some cases, vegetated and developed into a bench terrace. By plowing to this narrow levee and maintaining the drainage channel, its base was widened until it became known as a broadbase terrace.

Broadbase. This terrace has the distinguishing characteristic of farmability; that is, crops can be grown on this terrace and worked with modern-day machinery. These terraces are constructed either to remove or retain water and, based on their primary function, are classified either as graded or level.

A graded broadbase terrace is constructed from the upper side only with a variable or uniform grade in the channel to remove water at an acceptable rate. This is sometimes called the Nichols or channel terrace.

A level broadbase terrace is constructed on the contour and from both sides. It is recommended in areas where the soil is permeable to prevent overtopping. This terrace is sometimes called the Mangum or ridge terrace.

Another terrace in use today in semiarid regions where maximum moisture conservation is needed is the Zingg conservation bench terrace. It consists of an earthen embankment and a very broad, flat channel resembling a level bench. The runoff store in the channel comes from the sloping area above the bench which extends upslope to the next terrace.

The design of any type terrace must take into consideration factors such as soil characteristics, the cropping system to be used, soil management practices, climatic conditions, and others. All influence the effectiveness of the terrace system.

Terrace outlets. Since terrace channels concentrate rainfall on hillsides, outlets are a major feature of any successful terrace system. There are many different schemes for outlet construction. Masonry structures, such as storm sewers, concrete flumes, or drop inlets, may be used on steep land. Vegetation, such as grass or other thick-growing crops, may be used on gentle slopes. Modern practice in outlet construction calls for the use of natural channels with careful shaping and vegetating before large concentrations of water are turned into them.

Parallel terraces. Parallel terraces have been a major change in the layout of the terracing system. The object of parallel terracing is to facilitate the use of modern farm machinery by the elimination of point rows. Point rows in a field cause an excessive loss of productive time because of the turning of equipment. They also cause the destruction of crops and terraces by making it necessary for machinery to pass over them in their course of turning. Parallel terraces are made possible by a relatively small amount of land forming, such as moving land from high to low spots in the field and smoothing. This modification allows the efficient use of 4-, 6-, and 8-row cultivating equipment.

Construction equipment. Terracing equipment may be classified according to the method by which it moves the soil: lift and roll, throw, scrape and push, and carry.

The ideal terracing machine will meet certain requirements, such as (1) displacement of soil laterally to the desired position in the ridge; (2) placing the topsoil on or near the top of the ridge; (3) high work speed; (4) operation on slopes up to 15–20% and (5) low initial and operating cost. *See* AGRICULTURAL SOIL AND CROP PRACTICES; LAND DRAINAGE (AGRICULTURE).

Charles B. Ogburn

Terrain areas

Subdivisions of the continental surfaces distinguished from one another on the basis of the form, roughness, and surface composition of the land. These areas of distinctive landforms are the product of various combinations and sequences of events involving both deformation of the Earth's crust and surficial erosion and deposition by water, ice, gravity, and wind. The pattern of landform differences is strongly reflected in the arrangement of such other features of the natural environment as climate, soils, and vegetation. These regional associations must be carefully reckoned with by anyone planning activities as diverse as agriculture, transportation, city development, and military operations.

The **illustration** distinguishes among eight

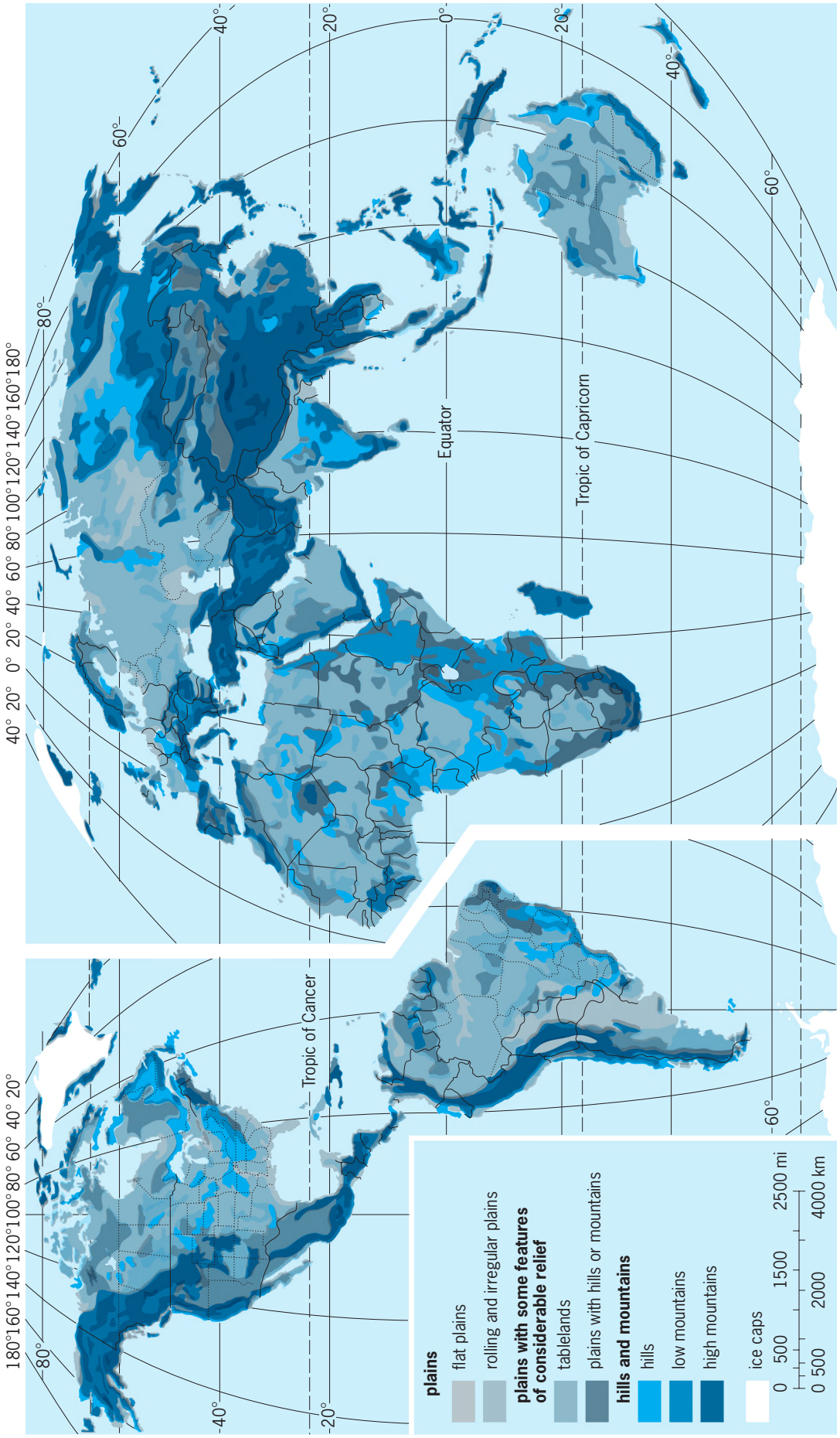

Distribution of terrain classes over the Earth. The subdivisions are intended to bring out only the most striking contrasts among land surfaces. Percentage of area occupied by each class is given in the text. (*After G. T. Trewartha, A. H. Robinson, and E. H. Hammond, Elements of Geography, 5th ed., McGraw-Hill, 1967*)

classes of terrain, on the basis of steepness of slopes, local relief (the maximum local difference in elevation), cross-sectional form of valleys and divides, and nature of the surface material. Approximate definitions of terms used and percentage figures indicating the fraction of the world's land area occupied by each class are as follows: (1) flat plains: nearly level land, slight relief, 4%; (2) rolling and irregular plains: mostly gently sloping, low relief, 30%; (3) tablelands: upland plains broken at intervals by deep valleys or escarpments, moderate to high relief, 5%; (4) plains with hills or mountains: plains surmounted at intervals by hills or mountains of limited extent, 15%; (5) hills: mostly moderate to steeply sloping land of low to moderate relief, 8%; (6) low mountains: mostly steeply sloping, high relief, 14%; (7) high mountains: mostly steeply sloping, very high relief, 13%; and (8) ice caps: surface material, glacier ice, 11%.

The continents differ considerably. Australia, the smoothest continent, has only one-fifth of its area occupied by hill and mountain terrain as against one-third of North America and more than one-half of Eurasia. Antarctica is largely ice covered; the only other great ice cap is on Greenland.

North America, South America, and Eurasia are alike in that most of their major mountain systems are linked together in extensive cordilleran belts. These form a broken ring about the Pacific basin, with an additional arm extending westward across southern Asia and Europe. The principal plains of Eurasia and the Americas lie on the Atlantic and Arctic sides of the cordilleras, but are in part separated from the Atlantic by lesser areas of rough terrain.

Most of Africa and Australia, together with the eastern uplands of South America and the peninsulas of Arabia and India, show great similarity to one another. They lack true cordilleran belts, and are composed largely of upland plains and tablelands, locally surmounted by groups of hills and mountains, and in many places descending to the sea in rough, dissected escarpments. *See* CONTINENT; CORDILLERAN BELT; HILL AND MOUNTAIN TERRAIN; PLAINS. Edwin H. Hammond

Bibliography. *Hammond World Atlas*, 4th ed., 2002; T. L. McKnight and D. Hess, *Physical Geography: A Landscape Appreciation*, 8th ed., 2004; A. H. Strahler and A. Strahler, *Introducing Physical Geography*, 4th ed., 2005.

Terrestrial coordinate system

The perpendicular intersection of two curves or two lines, one relatively horizontal and the other relatively vertical, is the basis for finding and describing terrestrial location. The Earth's graticule, consisting of an imaginary grid of east-to-west-bearing lines of latitude and north-to-south bearing lines of longitude, is derived from the Earth's shape and rotation, and is rooted in spherical geometry. The development of latitude and longitude likely originated during the classical period in Greece in the second century B.C. Hipparchus of Rhodes is thought to have been instrumental in its development. Plane coordinate systems, equivalent to horizontal X and vertical Y coordinates, are based upon cartesian geometry and differ from the graticule in that they have no natural origin or beginning for their grids.

Latitude and longitude. The Earth, which is essentially a sphere, rotates about an axis that defines the geographic North and South poles. The poles serve as the reference points on which the system of latitude and longitude is based (see **illus.**). *See* LATITUDE AND LONGITUDE.

Latitude is arc distance (angular difference) from the Equator and is defined by a system of parallels, or lines that run east to west, each fully encompassing the Earth. The Equator is the parallel that bisects the Earth into the Northern and Southern hemispheres, and lies a constant 90° arc distance from both poles. As the only parallel to bisect the Earth, the Equator is considered a great circle. All other parallels are small circles (do not bisect the Earth), and are labeled by their arc distance north or south from the Equator and by the hemisphere in which they fall. Parallels are numbered from 0° at the Equator to 90° at the poles. For example, 42°S describes the parallel 42 degrees arc distance from the Equator in the Southern Hemisphere. For increased location precision, degrees of latitude and longitude are further subdivided into minutes ($1^\circ = 60'$) and seconds ($1' = 60''$). *See* EQUATOR; GREAT CIRCLE, TERRESTRIAL.

Longitude is defined by a set of imaginary curves extending between the two poles, spanning the Earth. These curves, called meridians, always point to true geographical north (or south) and converge at the poles. In the present-day system of longitude, meridians are numbered by degrees east or west of the beginning meridian, called the Prime Meridian or the Greenwich Meridian, which passes through the Royal Observatory in Greenwich, England. The Prime Meridian was selected at an international conference in Washington, DC, in 1884 and assigned a longitude of 0°.

Since the Earth is a fundamentally a sphere, its circumference describes a circle containing 360°, the arc distance through which the Earth rotates in 24 hours. The arc distance from the Prime Meridian describes the location of any meridian (see illus.). The 180° meridian is commonly referred to as the International Dateline. Together, the Prime Meridian and the International Dateline describe a great circle that bisects the Earth, as do all other meridian circles. The west half of the Earth, located between the Prime Meridian and the International Dateline, comprises the Western Hemisphere, and the east half on the opposite side forms the Eastern Hemisphere. Meridians within the Western Hemisphere are labeled with a W, and meridians within the Eastern Hemisphere are labeled with an E. A complete description of longitude includes an angular measurement and a hemispheric label. For example, 78° W is the meridian 78° west of the Prime Meridian. Neither the 0° meridian (Prime) nor the 180° meridian (Dateline) is given a hemispheric suffix because they divide the two hemispheres, and therefore do not belong to either one.

Interaction of latitude and longitude. As parallels approach the poles, their circumferences shorten, while meridians converge as they approach the poles. Consequently, the length of a degree of latitude is shorter the closer it is to a pole. At the Equator, the length of a degree of longitude is greatest, and is equal to 69.17 mi (110 km). At 50° a degree of longitude equals 44.5 mi (71.7 km), and at 80° a degree of longitude equals 12.0 mi (19.4 km). The length of one degree of longitude at given latitude equals the cosine of the latitude times the length of a degree of longitude at the Equator (69.17 mi). In comparison, one degree of latitude is approximately 69 mi everywhere on the Earth because all meridians are of equal length.

The complete description of a location using the graticule describes a point where a parallel intersects a meridian. For example, the description 42°S, 78° W provides sufficient information to find that particular location on a map of the Earth.

Determining latitude and longitude. The empirical determination of latitude and longitude has a long history. For thousands of years it was common navigational knowledge that latitude can be measured through the observation of fixed stellar bodies, such as the Sun or North Star, using a sextant or any device that measures arc distance from the Earth's horizon to the star. No such means for measuring longitude existed until an accurate chronometer (clock) was invented in 1774. Before this, historical navigation maps had great discrepancies between the accuracy of latitude and longitude, where N-S locations showed good accuracy, but E-W locations were often in error. Unlike latitude, longitude has no natural starting point on the Earth. Longitude must be determined with respect to the difference in local time—the actual solar time at any location—between arbitrarily chosen Prime Meridian and the observation position, knowing that the Earth rotates at a relatively constant speed (1 degree of longitude in 4 minutes). *See* CELESTIAL NAVIGATION; CHRONOMETER; NAVIGATION; SEXTANT; TIME.

In the twentieth century, radio signals from Greenwich indicating the exact time at the Prime Meridian increased the accuracy of longitude determination. Today, precise measurement of latitude, longitude, and altitude is made possible through GPS (Global Positioning System) technology. GPS includes a constellation of 24 satellites orbiting the Earth which continually transmit location information and time to hand-held receivers on the Earth. Some GPS receivers are capable of determining Earth positions within centimeters. *See* SATELLITE NAVIGATION SYSTEMS.

Plane coordinates. Coordinate system alternatives to the graticule evolved in the early twentieth century because of the complexity of using spherical geometry in determining latitude, longitude, and direction. Plane (two-dimensional) or cartesian coordinate systems presume, with a small amount of error, that a relatively nonspherical Earth exists in smaller areas. Plane coordinates are superimposed upon these small areas, with coordinates being determined by the equivalent of a grid composed of a number of parallel vertical lines (X) and a complementary set of parallel horizontal lines (Y). All plane coordinate systems select a false origin for the grid because, unlike the graticule, there is no natural starting point to choose as an origin.

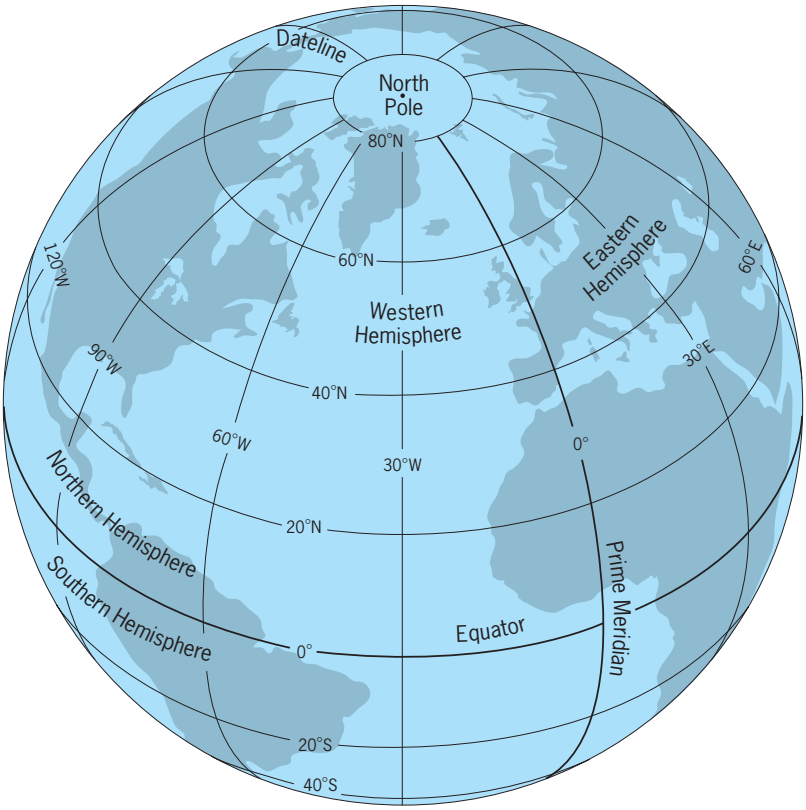

Earth's graticule. Meridians of longitude run from north to south, but are measured east or west of the Prime Meridian. Parallels run from east to west, but are measured north or south of the Equator.

State plane coordinate system (SPC). This system, used only in the United States, partitions each state into zones, each of which has its own coordinate system that slightly overlaps adjacent zones to facilitate connections to adjacent zones. The number of zones designated in each state is determined by the size of the state. Illinois, for example, has two zones, whereas Texas has five. Zone boundaries follow, with exceptions made for political subdivisions and natural features, either meridians or parallels depending on the shape of the state. All measurements are made in feet from a zone's false origin.

Universal Transverse Mercator (UTM) system. The UTM system is a worldwide coordinate system in which locations are expressed using metric units. The basis for the UTM system is the Universal Transverse Mercator map projection. This projection becomes vastly distorted in polar areas above 80°, and for this reason the UTM system is confined to extend from 84°N to 80°S. The UTM system partitions the Earth into 60 north-south elongated zones, each having a width of 6° of longitude. UTM measurements (in meters) are made in reference to a false origin located at the western edge of the zone and the Equator in the Northern Hemisphere or 90°S in the Southern Hemisphere. *See* MAP PROJECTIONS.

A number of other coordinate systems are in use

today. Foremost among these are the U.S. Public Land Survey System, the Universal Polar Stereographic (UPS) system, and the World Geographic Reference (GEOREF) system. Stephen Lavin

Bibliography. L. M. Bugayevskiy and J. P. Snyder, *Map Projections*, 1995; J. Campbell, *Map Use and Analysis*, 4th ed., 2000; A. H. Robinson et al., *Elements of Cartography*, 6th ed., 1995; N. J. W. Thrower, *Maps and Civilization*, 2d ed., 1999.

Terrestrial ecosystem

A community of organisms and their environment that occurs on the land masses of continents and islands. Terrestrial ecosystems are distinguished from aquatic ecosystems by the lower availability of water and the consequent importance of water as a limiting factor. Terrestrial ecosystems are characterized by greater temperature fluctuations on both a diurnal and seasonal basis than occur in aquatic ecosystems in similar climates, because water has a high specific heat, a high heat of vaporization, and a high heat of fusion compared with the atmosphere, all of which tend to ameliorate thermal fluctuations. The availability of light is greater in terrestrial ecosystems than in aquatic ecosystems because the atmosphere is more transparent than water. Gases are more available in terrestrial ecosystems than in aquatic ecosystems. Those gases include carbon dioxide that serves as a substrate for photosynthesis, oxygen that serves as a substrate in aerobic respiration, and nitrogen that serves as a substrate for nitrogen fixation. Terrestrial environments are segmented into a subterranean portion from which most water and ions are obtained, and an atmospheric portion from which gases are obtained and where the physical energy of light is transformed into the organic energy of carbon-carbon bonds through the process of photosynthesis.

Extent. The Earth has an estimated surface area of 197,272,000 mi^2 (510,934,000 km^2) of which about 29%, or 57,200,000 mi^2 (148,100,000 km^2), is occupied by land. About 1,540,000 mi^2 (4,000,000 km^2) of the land surface is occupied by fresh-water ecosystems of lakes, streams, and marshes; so terrestrial ecosystems occupy 55,660,000 mi^2 (144,150,000 km^2), or 28.2%, of Earth's surface.

Earth and the other planets are believed to be about 4.5 billion years old; the earliest fossil life forms are about 3.5 billion years old. The first terrestrial organisms appeared in the Silurian Period, about 425 million years ago. Therefore, terrestrial ecosystems have an age no greater than 9.4% of the total age of the Earth and 12.1% of the duration of life on Earth.

Principal organisms. Although they are comparatively recent in the history of life and occupy a much smaller portion of Earth's surface than marine ecosystems, terrestrial ecosystems have been a major site of adaptive radiation of both plants and animals. Major plant taxa in terrestrial ecosystems are members of the division Magnoliophyta (flowering plants), of which there are about 225,000 species in the class Magnoliopsida (dicots) and 50,000 species in the class Liliopsida (monocots), and the division Pinophyta (conifers), of which there are about 500 species. Members of the division Bryophyta (mosses and liverworts), of which there are about 24,000 species, are also important in some terrestrial ecosystems. Major animal taxa in terrestrial ecosystems include the classes Insecta (insects) with about 900,000 species, Aves (birds) with 8500 species, and Mammalia (mammals) with approximately 4100 species. *See* SYSTEMATICS; PLANT TAXONOMY; TAXONOMY.

Organisms in terrestrial ecosystems have adaptations that allow them to obtain water when the entire body is no longer bathed in that fluid, means of transporting the water from limited sites of acquisition to the rest of the body, and means of preventing the evaporation of water from body surfaces. They also have traits that provide body support in the atmosphere, a much less buoyant medium than water, and other traits that render them capable of withstanding the extremes of temperature, wind, and humidity that characterize terrestrial ecosystems. Finally, the organisms in terrestrial ecosystems have evolved many methods of transporting gametes in environments where fluid flow is much less effective as a transport medium.

Energy and chemical flow. The organisms in terrestrial ecosystems are integrated into a functional unit by specific, dynamic relationships due to the coupled processes of energy and chemical flow. Those relationships can be summarized by schematic diagrams of trophic webs, which place organisms according to their feeding relationships. The base of the food web is occupied by green plants, which are the only organisms capable of utilizing the energy of the Sun and inorganic nutrients obtained from the soil to produce organic molecules. The total living mass present in an ecosystem at any given time is referred to as biomass. The change in biomass through time is referred to as net productivity. Productivity due to the green plants and microorganisms is termed primary productivity while that due to animals is secondary productivity. Productivity will always diminish from stage to stage in the trophic web due to the laws of thermodynamics. This limitation is summarized by the equation $P_g = P_n + R$, where P_g is gross productivity, or total energy entering a trophic level, P_n is the net productivity of that trophic level, and R is the respiratory cost of maintaining that trophic level. For green plants, P_g is the energy initially captured in photosynthesis, R is the cost of maintaining the organisms, and P_n is the energy remaining to produce a biomass increment. Since it is only this biomass increment that is available to higher trophic levels, net productivity must always diminish at progressively higher trophic levels. *See* BIOLOGICAL PRODUCTIVITY; BIOMASS; FOOD WEB; PHOTOSYNTHESIS.

Terrestrial food webs can be broken into two segments based on the status of the plant material that enters them. Grazing food webs are associated with the consumption of living plant material by herbivores. Detritus food webs are associated with the consumption of dead plant material by detritivores.

Area and productivity of the various terrestrial ecosystems*

Type	Area occupied, 10^6 km^2	Mean net primary productivity, g/(m^2)(year)	Total annual productivity, 10^9 tons/year
Tundra	8.0	140	1.1
Desert	42.0	40	1.7
Temperate grassland	9.0	600	5.4
Woodland and shrubland	8.5	700	6.0
Cultivated land	14.0	650	9.1
Boreal forest	12.0	800	9.6
Savanna	15.0	900	13.5
Temperate forest	12.0	1250	14.9
Tropical forest	24.5	2000	49.4

*1 km^2 = 0.386 mi^2; 1 g/m^2 = 0.029 oz/yd^2; 1 ton = 0.9 metric ton.
source: After R. H. Whittaker, *Communities and Ecosystems,* 2d ed., Macmillan, 1975.

The relative importance of those two types of food webs varies considerably in different types of terrestrial ecosystems. Grazing food webs are more important in grasslands, where over half of net primary productivity may be consumed by herbivores. Detritus food webs are more important in forests, where less than 5% of net primary productivity may be consumed by herbivores. *See* SOIL ECOLOGY.

Energy flow is unidirectional in ecosystems, with a portion of that energy dissipated at each trophic level. As energy is dissipated, nutrients are released back into the environment. Nutrient flow, therefore, is cyclic. Energy flow and nutrient flow are impossible to separate functionally since they are intimately coupled. Energy flow will depend upon the rate at which nutrients are recycled, and nutrient flow will depend upon the availability of energy. *See* BIOGEOCHEMISTRY; ECOLOGICAL ENERGETICS.

Seasonality. Because of the marked seasonality of most of the terrestrial ecosystems, due to either temperature fluctuations in temperate locations or rainfall fluctuations in tropical locations, there are corresponding fluctuations in net primary productivity. The plants may sometimes be able to complete their life cycle in a brief time period, as do annual plants in the desert, or they may undergo periods of dormancy, as do deciduous trees in temperate forests. Similarly, animals either may hibernate during unfavorable periods or may migrate to other locations where conditions are more favorable. Many birds, for example, migrate from temperate latitudes during summer to tropical latitudes during winter in the temperate zones. Others may even migrate between the respective summer seasons in north and south temperate zones. Those animals that remain active in terrestrial ecosystems during periods of low primary productivity must have adaptations that allow them to subsist during unfavorable periods, often storing fat during productive periods and depleting those reserves during unfavorable periods.

Types. There is one type of extensive terrestrial ecosystem due solely to human activities and eight types that are natural ecosystems. Those natural ecosystems reflect the variation of precipitation and temperature over Earth's surface (see **table**). The smallest land areas are occupied by tundra and temperate grassland ecosystems, and the largest land area is occupied by tropical forest. The most productive ecosystems are temperate and tropical forests, and the least productive are deserts and tundras. Cultivated lands, which together with grasslands and savannas utilized for grazing are referred to as agroecosystems, are of intermediate extent and productivity. Because of both their areal extent and their high average productivity, tropical forests are the most productive of all terrestrial ecosystems, contributing 45% of total estimated net primary productivity on land. Due to their importance in productivity, tropical forests are believed to play an important role in the global carbon budget. There is increasing concern that deforestation of tropical ecosystems associated with agricultural development may contribute to increasing atmospheric carbon dioxide concentrations that may lead to climatic change. *See* DESERT; ECOLOGICAL COMMUNITIES; ECOSYSTEM; FOREST AND FORESTRY; GRASSLAND ECOSYSTEM; SAVANNA; TUNDRA.

S. J. McNaughton

Bibliography. W. B. Clapham, Jr., *Natural Ecosystems*, 2d ed., 1983; E. J. Kormondy, *Concepts of Ecology*, 3d ed., 1984; E. P. Odum, *Basic Ecology*, 1983; R. E. Ricklefs, *Ecology*, 3d ed., 1989; R. H. Young, S. Cousins, and D. Green (eds.), *Landscape Ecology and Geographical Information Systems*, 1993.

Terrestrial radiation

Electromagnetic radiation emitted from the Earth and its atmosphere. Terrestrial radiation, also called thermal infrared radiation or outgoing longwave radiation, is determined by the temperature and composition of the Earth's atmosphere and surface. The atmosphere is composed of two groups of gases, one with a nearly permanent concentration, consisting principally of nitrogen (N_2) and oxygen (O_2) molecules, and another with variable concentrations of other gases. Although considered to be permanent constituents, carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), and carbon monoxide (CO) have been observed to increase in association with anthropogenic activities. One of the principal variable gases is water (H_2O) vapor, the major compound that modulates the hydrological cycle involving evaporation, cloud formation, and precipitation. Water vapor concentration decreases rapidly with

latitude, almost following an exponential function. Ozone (O_3) concentration also varies significantly with space and time and occurs principally at altitudes of about 15–30 km (10–20 mi). A significant variable gas is a mixture of chlorofluorocarbons (CFCs) produced by industrial activities. *See* ATMOSPHERE; ATMOSPHERIC GENERAL CIRCULATION; HYDROLOGY; RADIATIVE TRANSFER.

The atmosphere also contains aerosol particles ranging in size from about 10^{-3} to 20 micrometers that are known to be produced by natural processes as well as by human activity. Aerosol concentrations generally decrease rapidly with height. Some aerosols promote condensation and ice nuclei, upon which cloud particles may form. Clouds are global in nature and regularly cover more than 50% of the Earth. There are various types of clouds. Some clouds, such as high-level cirrus in the tropics and low-level stratus in the Arctic and near coastal areas, are climatologically persistent. Some clouds generate precipitation. *See* AEROSOL; CLOUD; CLOUD PHYSICS.

The temperature structure of the Earth and the atmosphere is a result of numerous physical, chemical, and dynamic processes. In a one-dimensional context, the temperature structure is determined by the balance between radiative and convective processes. From the surface to a height of about 10 km (6 mi), the temperature decreases at a typical rate of about 6.5°C (12°F) per kilometer; this region is referred to as the troposphere, where major weather events occur. Variability of the global temperature field is modulated by the transport of various types of energy by the winds, the radiative energy emitted from the Sun and received by the Earth, and the radiative energy emitted from the Earth and the atmosphere. *See* HEAT BALANCE, TERRESTRIAL ATMOSPHERIC; SOLAR RADIATION; SUN; TROPOSPHERE; WIND.

Emissivities. The Earth's surface emits electromagnetic radiation according to the laws that govern a blackbody or a graybody. A blackbody absorbs the maximum radiation and at the same time emits that same amount of radiation so that thermodynamic equilibrium is achieved, defining a uniform temperature. The rate at which emission takes place is a function of temperature and wavenumber (cm^{-1}) or wavelength (μm) according to Planck's law. A graybody is characterized by incomplete absorption and emission and is said to have emissivity less than unity. The thermal infrared emissivities from water and land surfaces are normally 90–95%. It is usually assumed that the Earth's surfaces are approximately black in the analysis of infrared radiative transfer. Exceptions include snow and some sand surfaces whose emissivities are wavelength-dependent and could be less than 90%. *See* BLACKBODY; GRAYBODY; HEAT RADIATION; PLANCK'S CONSTANT.

Absorption and emission of radiation by atmospheric molecules are more complex and require a fundamental understanding of quantum mechanics. A molecule, composed of atoms, can rotate, or revolve, about an axis through its center of gravity, and therefore has rotational energy. The atoms of the molecule are bounded by certain forces like springs in which the individual atoms can vibrate about their equilibrium positions relative to one another. The molecule therefore also has vibrational energy. It is also possible for the energy of a molecule to change due to a change in the energy state of the electrons of which it is composed. Thus the molecule has electronic energy. These three forms of energy are said to be quantized, and assume only discrete values. Absorption and emission of radiation take place when the molecules undergo transitions from one energy state to another, and these transitions are generally governed by selection rules in quantum mechanics. Rotational energy changes are relatively small, and many of the rotational energy levels are populated at terrestrial temperatures. Changes in vibrational energy are much larger than the minimum changes in rotational energy. Thus, vibrational transitions never occur alone, but are coupled with simultaneous rotational transitions, producing a group of absorption lines known as the vibrational–rotational band in the thermal infrared spectrum of the Earth's atmosphere associated with atmospheric gases. These lines have finite widths and overlap one another. The line shape is determined by molecular collision in the lower atmosphere and by random thermal motion in the upper atmosphere. The width of the line is primarily dependent upon pressure but is also affected by temperature. *See* INFRARED SPECTROSCOPY; QUANTUM MECHANICS; SELECTION RULES (PHYSICS).

The water molecule is composed of two hydrogen atoms and one oxygen atom that form an isosceles triangle that is obtuse, referred to as an asymmetric top configuration. This configuration produces the pure water rotational band ranging 0–1000 cm^{-1}, which is important in the generation of tropospheric cooling. The band located at 1594.78 cm^{-1} (6.25 μm) is the vibrational–rotational band of water that has been used for remote sensing of its concentration from satellites. Water vapor also exhibits less selective absorption in the region 800–1200 cm^{-1}, the so-called thermal infrared window. Carbon dioxide has a linear symmetrical configuration, with the carbon atom in the middle and oxygen atoms on each side. The perpendicular vibration of oxygen atoms coupled with rotational transitions produces the 15-μm CO_2 band, known for causing greenhouse warming. Ozone has an asymmetric top configuration similar to water and exhibits absorption in the 9.6-μm region. Methane and nitrous oxide also show strong absorption bands in the 7–8-μm region. Other gases with absorption bands in the thermal infrared region are sulfur dioxide, ammonia, and chlorofluorocarbons (CFC). The above gases are referred to as greenhouse gases because of their ability to absorb or trap the radiation emitted by the Earth and the atmosphere. *See* CARBON DIOXIDE; GREENHOUSE EFFECT; METHANE; WATER.

Clouds are global in nature and can interact with terrestrial radiation. Clouds are composed of waterdroplets or ice crystals and can both reflect and transmit the radiation emitted from the surface and the

atmosphere while emitting infrared radiation according to the temperature structure within them. If the cloud as a whole is a blackbody, it will behave just like the Earth's surface. In this case, radiation from below and above the cloud would not be able to penetrate the cloud. The emitted radiance at the cloud top or bottom is governed by Planck's law. Most clouds that are composed of waterdroplets are black clouds because of high number concentrations and small droplet sizes. However, clouds that are composed of ice crystals with sizes ranging from a few micrometers to 1000 μm, such as cirrus clouds, are generally nonblack. Determination of the radiative properties of cirrus clouds in conjunction with remote sensing and climate studies is a subject of contemporary research.

Radiation measurement. **Figure 1** displays the principal absorbing gases that are identified in an observed infrared spectrum. This spectrum was obtained from a scanning high-resolution interferometer sounder, an instrument that measures the emitted thermal infrared radiation of 3.3–18 μm from the NASA high-flying ER-2 aircraft at about 20 km (12 mi) with a spectral resolution on the order of 0.01 cm^{-1}. The spectrum is presented in terms of the brightness temperature (K) as a function of wavenumber (cm^{-1}) and wavelength (μm). Identified in Fig. 1 are the 15-μm CO_2 band, 9.6-μm O_3 band, 6.25-μm H_2O vibrational–rotational band, CH_4 and N_2O in the 7–8-μm region, two CFC bands, a part of the H_2O rotational band, and the thermal infrared window of 800–1200 cm^{-1} discussed previously. In addition to these bands in the thermal infrared, the 4.3-μm CO_2 and the spectral bands for CO, N_2O, and CH_4 at wavelengths less than 5 μm are shown. Note that because the temperature structure of the Earth–atmosphere system mentioned above, its emission (terrestrial radiation) is primarily confined to wavelengths longer than about 5 μm, while the radiation from the Sun is largely in the shorter wavelength region from about 0.2 to 5 μm, referred to as solar or shortwave radiation.

K. N. Liou

Meteorological satellites. Terrestrial radiation originating from the Earth–atmosphere–ocean system, as well as solar radiation reflected and scattered back to space, is measured on a daily basis by meteorological satellites. Instruments on meteorological satellites measure visible, ultraviolet, infrared, and microwave radiation. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; ELECTROMAGNETIC RADIATION; METEOROLOGICAL SATELLITES; REFLECTION OF ELECTROMAGNETIC RADIATION; SCATTERING OF ELECTROMAGNETIC RADIATION.

Each spectral region provides meteorologists and other Earth system scientists with information about atmospheric ozone, water vapor, temperature, aerosols, clouds, precipitation, lightning, and many other parameters. Measuring atmospheric radiation allows the detection of sea and land temperature, snow and ice cover, and winds at the surface of the

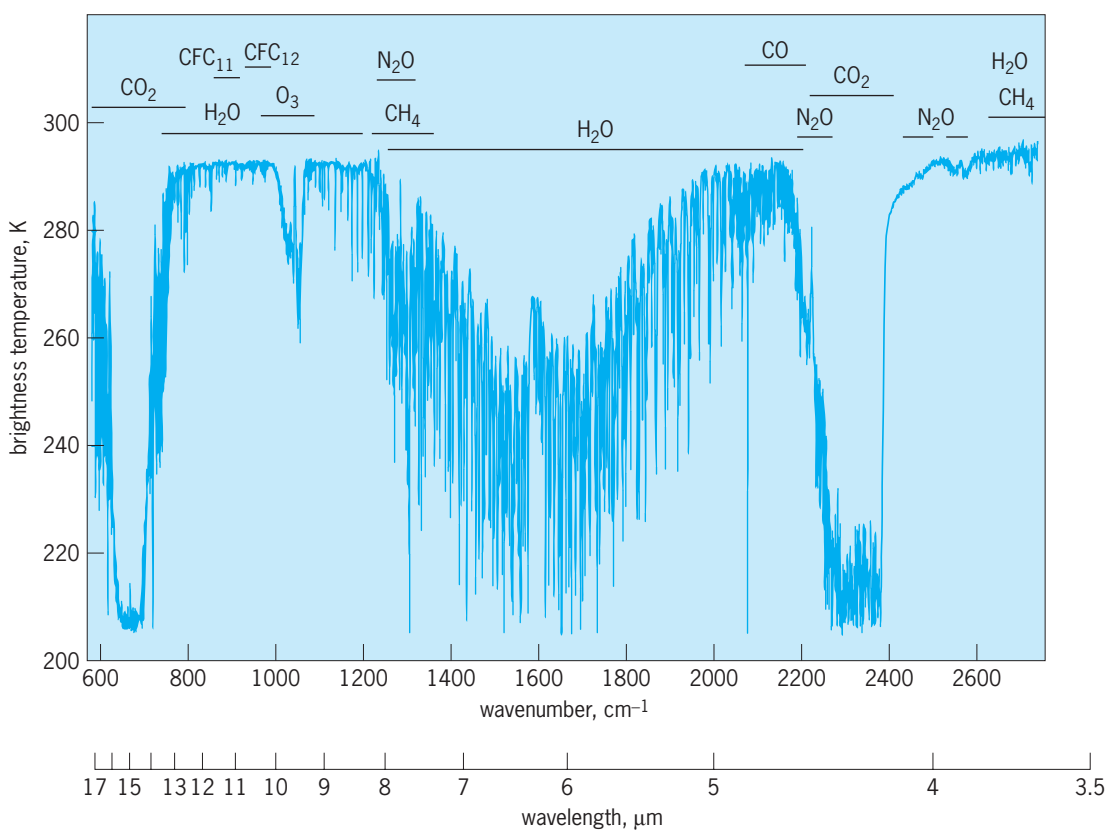

Fig. 1. Observed infrared spectrum illustrating the absorption gases and their spectral location. This spectrum was obtained from the scanning high-resolution interferometer sounder (S-HIS), which measured the emitted thermal radiation between 3.3 and 18 μm, onboard the NASA ER-2 aircraft over the Gulf of Mexico southeast of Louisiana on April 1, 2001. (*Data taken from K. N. Liou, 2002; originally provided by A. Huang and D. Tobin*)

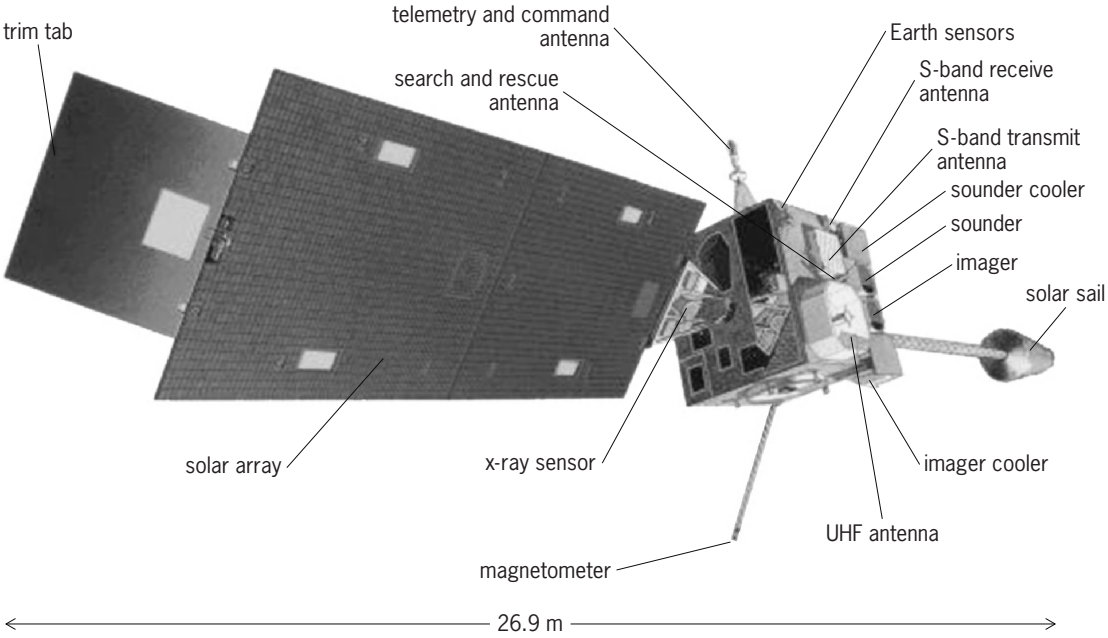

Fig. 2. Sketch of a *Geostationary Operational Environmental Satellite* (*GOES*) of the late 1990s.

ocean. By tracking the movement of clouds and other atmospheric features, such as aerosols and water vapor, it is possible to obtain estimates of winds above the surface. *See* SATELLITE METEOROLOGY.

Scientists use physically based algorithms to process the satellite-observed spectral radiation into either images of atmospheric and surface features or desired estimates of weather parameters. For example, the physical relationship between the absorption and reemission of infrared radiation is used to estimate atmospheric temperature profiles over vast reaches of ocean areas. *See* REMOTE SENSING.

Orbits of meteorological satellites are generally selected as one of two types. Geostationary (geosynchronous) orbits are in the equatorial plane and have the advantage of obtaining continuous, time-lapse observations of a portion of Earth as large as the Americas (**Fig. 2**). Other, near-polar orbits are Sun-synchronous and are designed to measure all areas of Earth at nearly the same local times each day.

Since meteorological satellites were first flown by NASA in the 1960s, a system of research and operational weather forecast-related satellites has been developed in the United States by the National Oceanic and Atmospheric Administration, and by the weather services of Japan, India, Russia, and the European Union. Today, worldwide weather satellite coverage is providing input to computer forecast models and to National Weather Service and television weather forecasters on an hourly basis. *See* WEATHER FORECASTING AND PREDICTION.

The scientific discoveries in weather and climate credited to observations from satellites include (1) measurement of the energy output from the Sun and the radiation budget of the Earth—warmer and darker than was believed before satellite observations; (2) the organization of tropical weather systems and waves; (3) the existence of mesoscale convective clusters—major precipitation-producing weather systems over midlatitude continents; (4) development and growth of severe thunderstorms on "gust fronts" outflowing from other storms; and (5) variation of the warm core inner structure of tropical storms and associated spiral bands.

Thomas H. Vonder Haar

Bibliography. S. Q. Kidder and T. H. Vonder Haar, *Satellite Meteorology: An Introduction*, Academic Press, 1995; K. N. Liou, *An Introduction to Atmospheric Radiation*, 2d ed., Academic Press, 2002.

Territoriality

Behavior patterns in which an animal actively defends a space or some other resource. One major advantage of territoriality is that it gives the territory holder exclusive access to the defended resource, which is generally associated with feeding, breeding, or shelter from predators or climatic forces.

Feeding territories involve defense of an area where food is found, or defense of the food itself. Breeding, or reproductive territories, may involve defense of a breeding site where access to mates is increased, or defense of offspring, such as in a nest. Shelter territories usually involve a hole or other structure where an animal can rest or hide safely, a place where a cold-blooded animal basks in the sun to increase its body temperature, a refuge site from extreme temperatures or humidities, or a physical barrier to wind or water currents. Feeding and breeding territories can be mobile, such as when an animal defends a newly obtained food source or a temporarily receptive mate. Stationary territories often serve multiple functions and include access to food, a place to rear young, and a refuge site from predators and the elements.

Territoriality can be understood in terms of the benefits and costs accrued to territory holders.

Benefits include time saved by foraging in a known area, energy acquired through feeding on territorial resources, reduction in time spent on the lookout for predators, or increase in number of mates attracted and offspring raised. Costs usually involve time and energy expended in patrolling and defending the territorial site, and increased risk of being captured by a predator when engaged in territorial defense.

Because territories usually include resources that are in limited supply, active defense is often necessary. Such defense frequently involves a graded series of behaviors called displays that include threatening gestures such as vocalizations, spreading of wings or gill covers, lifting and presentation of claws, head bobbing, tail and body beating, and finally, direct attack. Direct confrontation can usually be avoided by advertising the location of a territory in a way that allows potential intruders to recognize the boundaries and avoid interactions with the defender. Such advertising may involve odors that are spread with metabolic by-products, such as urine or feces in dogs, cats, or beavers, or produced specifically as territory markers, as in ants. Longer-lasting territorial marks can involve visual signals such as scrapes and rubs, as in deer and bear. *See* CHEMICAL ECOLOGY.

Individual size of combatants often determines the outcome of behavioral interactions for an unused territory. Once such a territory is established, however, a territory holder frequently can defend and repel individual intruders, even when the intruder is larger. This advantage can be overcome by invading as part of a group, with the many members of the group serving as "multiple lures." The defender can repel only a small number of intruders, while other members feed on the territorial resources.

In reproductive territories, the advantage accrued to a successful territory holder can be measured by counting the number of matings that different territory holders obtain. Frequently the largest, strongest, or most experienced male will occupy the best territory and mate with the most females. Such territories may often contain a place for egg deposition, such as in many birds, frogs, and fishes. In other organisms, such as antelope, grouse, and bower birds, many males may compete for adjoining display sites that do not contain obvious resources. Here the central location in the display ground, or lek, is often the preferred one, and its occupant obtains more matings than other males. Females then go off somewhere else to bear young or lay eggs. *See* ETHOLOGY; POPULATION ECOLOGY; REPRODUCTIVE BEHAVIOR.

Gene S. Helfman

Bibliography. J. Alcock, *Animal Behavior: An Evolutionary Approach*, 7th ed., 2001; J. W. Grier and T. Burk, *Biology of Animal Behavior*, 2d ed., 1992; J. R. Krebs and N. B. Davies, *An Introduction to Behavioral Ecology*, 3d ed., 1993.

Tertiary

The older major subdivision (period) of the Cenozoic Era, extending from the Cretaceous (top of the Mesozoic Era) to the beginning of the Quaternary

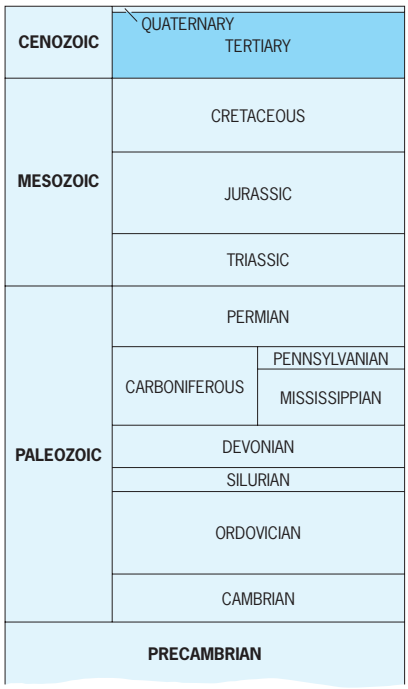

(younger Cenozoic Period). The term Tertiary corresponds to all the rocks and fossils formed during this period. Although the International Commission on Stratigraphy uses the terms Paleogene and Neogene (pre-Quaternary part) instead, Tertiary is still widely used in the geologic literature. Typical sedimentary rocks include widespread limestones, sandstones, mudstones, marls, and conglomerates deposited in both marine and terrestrial environments; igneous rocks include extrusive and intrusive volcanics as well as rocks formed deep in the Earth's crust (plutonic). *See* CRETACEOUS; FOSSIL; ROCK.

The Tertiary Period is characterized by a rapid expansion and diversification of marine and terrestrial life. In the marine realm, a major radiation of oceanic microplankton occurred following the terminal Cretaceous extinction events. This had its counterpart on land in the rapid diversification of multituberculates, marsupials, and insectivores—holdovers from the Mesozoic—and primates, rodents, and carnivores, among others, in the ecologic space vacated by the demise of the dinosaurs and other terrestrial forms. Shrubs and grasses and other flowering plants diversified in the middle Tertiary, as did marine mammals such as cetaceans (whales), which returned to the sea in the Eocene Epoch. The pinnipeds (walruses, sea lions, and seals) are derived from land carnivores, or fissipeds, and originated in the Neogene temperate waters of the North Atlantic and North Pacific. Indeed, the great diversification on land and in the sea of birds and, particularly, mammals has led to the informal designation of the Tertiary as the Age of Mammals (**Fig. 1**) in textbooks on historical geology.

Geography. The modern configuration of continents and oceans developed during the Cenozoic Era as a result of the continuing process known as plate tectonics. Mountain-building events (orogenies) and uplifts of large segments of the Earth's crust

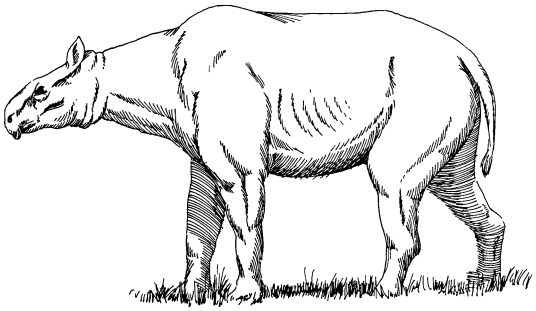

Fig. 1. ***Baluchitherium*, the largest land mammal known, from the Tertiary (Oligocene Epoch) of Asia. It was 18 ft (5.4 m) high at the shoulders. (*After R. A. Stirton*, *Time*, *Life and Man*, *Wiley*, *1959*)**

(epeirogenies) alternated with fluctuating transgressions and regressions of the seas over land. This resulted in a complex alternation of marine and terrestrial sediments and their contained records of the passage of life (fossils). Some modern inland seas (for example, Lake Baikal and the Caspian Sea) are remnants of once more extensive widespread epeiric (shallow) seaways of the early Tertiary.

The middle to late Tertiary Alpine-Himalayan orogeny and the late Tertiary Cascadian orogeny led to the east-west and north-south mountain ranges, respectively, which are located in Eurasia and western North America. *See* CORDILLERAN BELT; MOUNTAIN SYSTEMS; OROGENY; PLATE TECTONICS.

Rocks. Tertiary sedimentary rocks occur as a relatively thin veneer of marine rocks on the margins of continents around the world. In the petroliferous province of the Gulf of Mexico, Tertiary rocks attain thicknesses in excess of 30,000–40,000 (9000–12,000 m); whereas in the more tectonically active borderlands around the Pacific Ocean, such as the Santa Barbara–Ventura Basin of California, and the flanks of the uplifted Himalayan-Alpine chain of Eurasia, thicknesses in excess of 50,000 ft (15,000 m) have been recorded. Terrestrial (nonmarine) strata are generally thinner, are more patchy in distribution, and occur predominantly in the internal basins of the continents (for example, the Basin and Range Province of North America and the Tarim Depression of Asia). Major Tertiary volcanic provinces include those of the Deccan region of India, the basaltic plateaus of Greenland and Iceland, and the Columbia Plateau of the northwest United States.

Stratigraphy and history. Although the ancient Greeks recognized the shells of mollusks far inland from the Aegean Sea as fossil marine organisms, as did Leonardo da Vinci some 2000 years later, it was not until the era of enlightenment in the eighteenth century that the first attempt was made to place the Earth's rock record into a historical context. The term Tertiary is derived from Giovanni Arduino, who in 1759 formulated a threefold subdivision of the Earth's rock record in Primary, Secondary, and Tertiary. While the first two terms have long since disappeared from geologic hagiography, the term Tertiary persists in modern scientific literature. In its more modern sense the term Tertiary is defined by its usage in 1810 by the French Scientists Alexandre

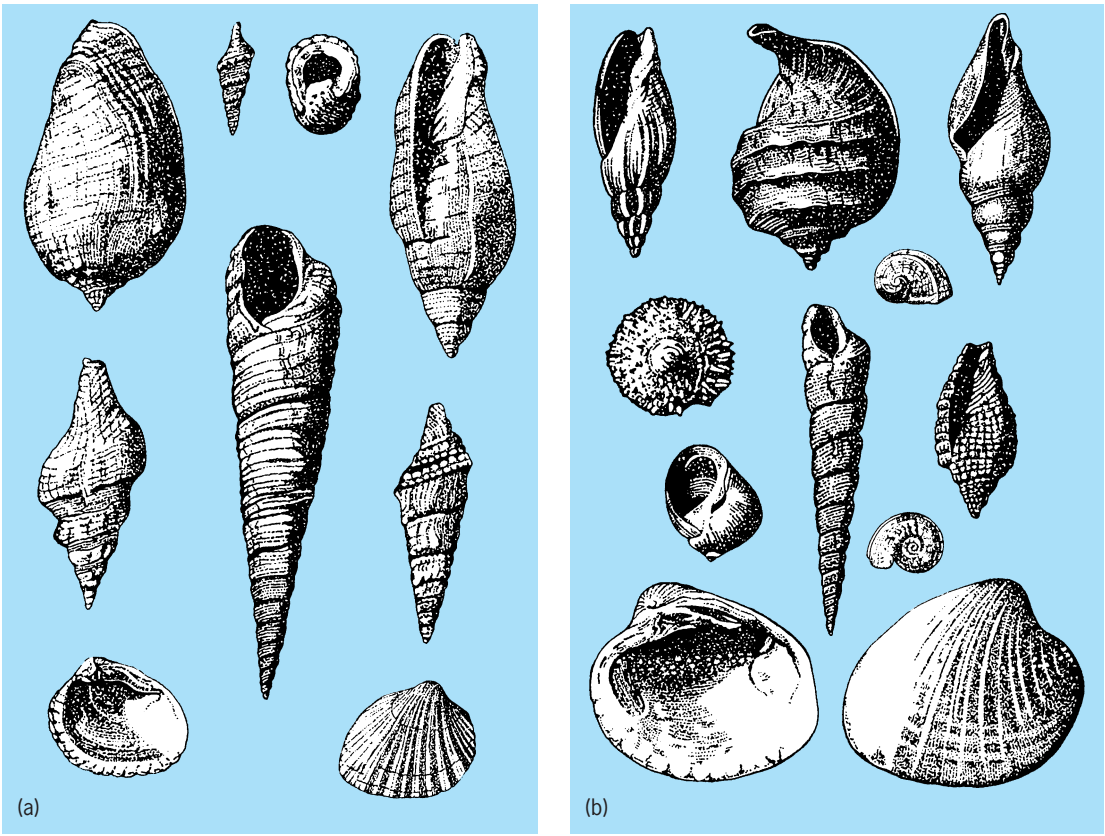

Fig. 2. Mollusk fossils used by Lyell to zone the Tertiary. (*a*) Miocene. (*b*) Eocene. (*After C. Lyell*, *Principles of Geology*, *1833*)

Brogniart and Georges Cuvier for all the rock formations in the Paris Basin that lay above the Cretaceous chalk sequence. Although many subdivisions of the Tertiary exist that developed in the succeeding two centuries, only five major time-rock units generally are recognized. In 1833, Charles Lyell made the first systematic hierarchical subdivision of the Tertiary Period based upon the observations of his Parisian colleague M. Deshayes and other contemporary European conchologists that the percentages of living species in the fossil record increased as the Tertiary stratigraphic record ascended (**Fig. 2**). Lyell's Tertiary subdivisions include, in ascending order, the Eocene, Miocene, Older Pliocene, and Newer Pliocene. The last term was subsequently (1839) changed to Pleistocene. Heinrich Ernst von Beyrich later defined the term Oligocene for rocks exposed in the North German Basin and the Rhine Basin that had been previously allocated to a part of either the Eocene or Miocene by Lyell. The paleobotanist W. P. Schimper added the term Paleocene in 1874 based on the oldest Tertiary terrestrial strata exposed in the east Paris Basin. *See* CENOZOIC; EOCENE; HOLOCENE; MIOCENE; OLIGOCENE; PALEOBOTANY; PALEOCENE; PALEONTOLOGY; PLEISTOCENE; PLIOCENE; STRATIGRAPHY. W. A. Berggren

Bibliography. R. H. Dott, Jr., and D. R. Prothero, *Evolution of the Earth*, 7th ed., 2003; B. M. Funnell and W. R. Riedel, *The Micropaleontology of Oceans*, 1971; S. J. Gould, *Time's Arrow, Time's Cycle*, 1987; H. L. Levin, *The Earth Through Time*, 7th ed., 2003.

Testis

The organ of sperm production. In addition, the testis (testicle) is an organ of endocrine secretion in which male hormones (androgens) are elaborated. In the higher vertebrates (reptiles, birds, and mammals), the testes are paired and either ovoid or elongated in shape. In mammals, the testes are usually ovoid or round. In many species (for example, humans) they are suspended in a pouch (scrotum) outside the main body cavity; in other species they are found in such a pouch only at the reproductive season; in still others the testicles are permanently located in the abdomen (for example, in whales and bats).

Histology

Within a firm and thick capsule of connective tissue, the tunica albuginea, the testis contains a varying number of thin (0.005–0.017 in. or 0.12–0.45 mm, depending upon the species) but very long seminiferous tubules which are the sites of sperm formation. Essentially, these tubules are simple loops which open with both their limbs into a network of fine, slitlike canals, the rete testis. From this the sperm drains through a few narrow ducts, the ductuli efferentes, into the epididymis, where sperm mature and are stored.

The seminiferous tubules constitute most of the testis, and in different species vary greatly in complexity. Often they are extremely coiled and winding. Sometimes they branch and interconnect as in man. Their total length in man is about 700 ft (210 m), in the bull about 3 mi (4.8 km). Each tubule is surrounded by a layer of thin cells (**Fig. 1**) which is contractile and enables the tubules to wriggle slowly. The spaces between tubules are filled with connective tissue, blood vessels, an extensive network of very thin-walled lymph vessels, and secretory cells, the interstitial cells or cells of Leydig, which secrete male hormone. The interstitial cells vary in different species and different periods in life. In the human fetus they are extremely abundant and functional; later they regress and change into inconspicuous connective tissue cells. At puberty they reappear and enter the main productive phase of their life.

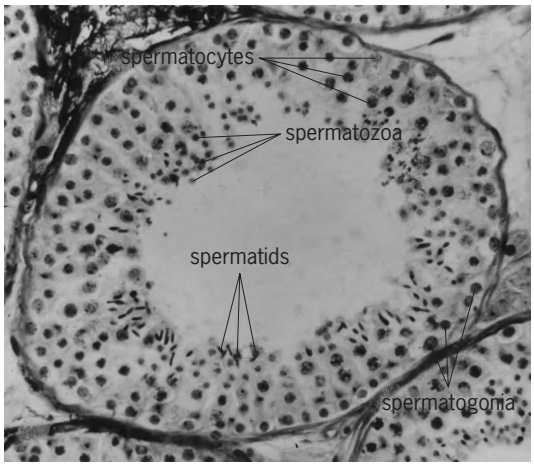

Fig. 1. Human seminiferous tubules in cross section. A multitude of stages of germ cells is present.

Sperm Formation

The sperm cells, spermatozoa, develop in the wall of the seminiferous tubules, either periodically, as in most vertebrates, or continually, as in humans. Most of the cells in the tubules are potential spermatozoa (spermatogenic or germ cells). Nursing cells (Sertoli cells) are interspersed at regular intervals between them (Fig. 1). The Sertoli cells support and surround the developing spermatogenic cells and provide a specialized environment, which is absolutely necessary for normal sperm development. *See* SPERM CELL.

Four cell types representing characteristic stages in the development of sperm cells have been given specific names. The youngest ones are called spermatogonia or, at their earliest stages, stem cells. They lie at the periphery of the tubule and show no resemblance to mature spermatozoa. Once a reproductive phase has begun, the spermatogonia undergo periodic divisions. A few of them stop dividing after one or two mitoses and revert to stem cells. These are the progenitors of future generations of spermatozoa. Most spermatogonia, however, go through additional divisions and transform into another cell type, the spermatocyte, in which the nucleus is slowly preparing for the maturation divisions (meiosis). The small cells resulting from these divisions are called

spermatids. They are haploid; that is, their chromosome number is reduced by one-half. Usually four spermatids arise from one spermatocyte. The subsequent transformation of the spermatid into a spermatozoon is called spermiogenesis. This is an extremely complex process which involves the development of a distinctive head structure containing the highly condensed nucleus, head cap structures such as the acrosome that help the sperm to attach and enter the egg, and a tail which allows the sperm to move in a swimming motion. During condensation of the sperm nucleus, there is considerable reorganization of the chromosomes, with the appearance of many unique nuclear proteins, such as the testis-specific histone Hlt, and the gradual replacement of histones by another group of proteins that are unique to the sperm nucleus, the protamines. These specific nuclear proteins facilitate very dense packing of the nuclear material in the sperm head, although at the expense of the genes ability to remain active. During spermiogenesis, groups of spermatids become attached to each nursing cell and often indent its membrane deeply. In the final stage the spermatozoa retract toward the lumen and are released; they are still hardly motile and usually they are not fertile until they have spent some time within the epididymis.

The duration of spermatogenesis is known with fair accuracy for several species. Usually it is about 5–7 weeks, in human males 75 days, from the first spermatogonial division to the release of the spermatozoa from the testis. About one-fourth of this time is spent in the spermatogonial stage of cell multiplication, a little more than one-third in the long preparation for, and the rapid completion of, the maturation divisions, and about one-third in spermatogenesis. Spermatogenesis in the testis is the result of a balance between proliferation and differentiation, and cell degeneration or apoptosis. Apoptosis of the spermatogenic cells is largely hormonally controlled, and specifically directed apoptosis occurs in conditions of testicular damage due to environmental insults such as heat, radiation, or chemical toxicants. Recovery of spermatogenesis is possible provided the stem cells are not depleted by these processes.

In nearly all mammals, spermatogenesis occurs in rigid patterns which are essentially similar. Large groups of cells filling sections of tubules up to several centimeters in length develop synchronously. Generations follow each other at definite intervals so that in any section of a tubule, characteristic associations of cells are present. In addition, along the tubule, groups of cells which are in synchronous development border on other groups in immediately preceding or following stages. This constitutes the spermatogenic wave. In humans and apes, the spermatogenic wave is less well defined, with a spiral arrangement of several short asynchronous waves of development running along the tubule. As a result, each cross section of a tubule reveals a multitude of stages of germ cells coexisting in random array (Fig. 1).

In lower vertebrates, many patterns and arrangements of spermatogenesis are found, sometimes in gonadal sacs which open directly into the abdominal cavity from which they are released through abdominal pores (cyclostomes), sometimes in tubules or compartments which have only temporary connections with outleading ducts (selachians). In all lower vertebrates a multitude of spermatozoa develops in the presence of definite nursing cells. In most cases, large groups of cells, often whole compartments (in insects and amphibians), develop synchronously. *See* SPERMATOGENESIS.

Mark P. Hedger; Edward C. Roosen-Runge

Physiology

The functions of the testis are dependent on the secretion of gonadotropic hormones, the release of which from the pituitary gland is regulated by the central nervous system.

In some lower forms, for example, many crustaceans, the hormone-producing "androgenic gland" is anatomically distinct from the gametogenic organ or testis proper. In mammals, male-hormone production resides in the Leydig cells, located in the intertubular tissue of the testes. Destruction of the seminiferous epithelium by x-irradiation, heat, or cadmium poisoning does not abolish the endocrine function of the testis.

Formation and release of testosterone. The principal androgenic hormone released by the testis into the bloodstream is testosterone. The testis is able to utilize simple molecules, such as acetic acid, to form cholesterol and to convert this sterol to Δ^5-pregnenolone. The major pathways leading to the formation of testosterone from pregnenolone are shown in **Fig. 2**. One pathway involves progesterone and Δ^4-androstene-3,17-dione as intermediates; androstenedione often accompanies testosterone in the venous blood of the testis. Alternatively, removal of the side chain at carbon-17 may precede oxidation of the hydroxyl group at carbon-3, so that additional Δ^5-3β-hydroxysteroid intermediates, such as dehydroepiandrosterone and Δ^5-androstenediol, are formed; the latter is itself a fairly potent androgen. In humans, production of testosterone occurs mainly through the Δ^5 pathway (pregnenolone to androstenediol), with the Δ^4 pathway through progesterone contributing relatively little to androgen production. This appears to be due to the relatively poor affinity of the human steroidogenic enzyme for the Δ^4 intermediates. In some species, however, the Δ^4 pathway is more important. Both testosterone and androstenedione may be further metabolized into estrogens in the testis. This conversion takes place in the tubular tissue, although interstitial cells also contribute. The production of estrogens in the male varies quite widely among species, from relatively low in humans to very high, for example in stallions and boars. It is now apparent, however, that estrogens are important in the development and proper function of the ducts which drain the testis (the rete testis and ductuli efferentes), even in species with relatively low levels of estrogens. The testis of the stallion is one of the richest natural sources of estrogen. Administration of the luteinizing hormone (LH) of the pituitary to human males leads to increased urinary excretion of estrogens.

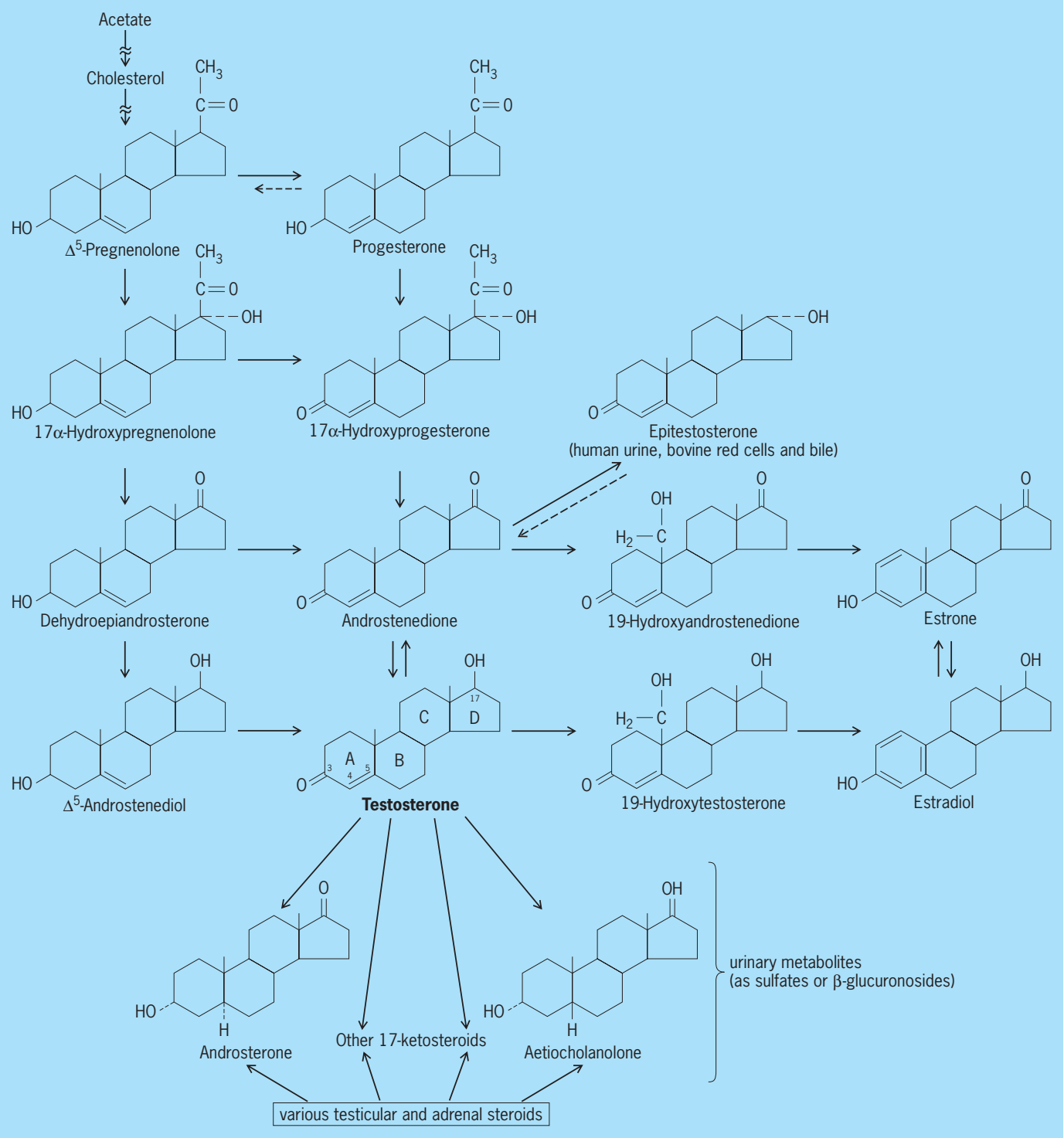

Fig. 2. Major pathways for the biosynthesis of testosterone in the testis and subsequent metabolism of the hormone.

Testosterone synthesis is normally limited by the rate of pituitary gonadotrophin secretion: administration of the luteinizing hormone or of chorionic gonadotrophin results in increased testosterone synthesis and release within minutes. These hormones also stimulate growth and multiplication of Leydig cells. Hypophysectomy leads to cessation of androgen formation.

The bulk of the hormone produced enters the spermatic venous blood, where it becomes associated with albumin and with a specific testosterone-binding globulin on its way to remote target organs. Some of the hormone, however, is released directly into the testicular interstitial fluid and lymph and diffuses across the basement membrane into the tubules; in the tubular fluid it is carried to the epididymis. This local diffusion of androgen assists sperm maturation and storage in this organ.

Metabolism and excretion of testosterone. Testosterone is rapidly inactivated in the liver; some metabolism also occurs in target organs, in red blood cells, and in the kidneys. In man, the final products

are chiefly 17-ketosteroids with a saturated A-ring, such as androsterone and aetiocholanolone (Fig. 2); these are excreted as sulfates or glucuronosides in the urine. Similar products are formed from adrenal steroids and from testicular steroids other than testosterone, so that urinary 17-ketosteroid excretion does not accurately reflect male hormone secretion. In the bull, 90% of the testosterone metabolites, such as epitestosterone (Fig. 2), are excreted via the bile into the feces. In some animal types, both urinary and biliary excretion occur, with partial resorption of steroid metabolites from the gut (enterohepatic circulation).

Development of function and puberty. The enzymes required for the synthesis of testosterone first appear in the testis of the early fetus, but the Leydig cells regress before or shortly after birth and resume their development only toward the onset of puberty. Androgens are present in the blood in small amounts even in infancy, but the testicular secretion increases markedly and changes its character shortly before puberty: The secretion of the prepubertal testis of the bull, rat, guinea pig, and possibly of boys differs from that of the adult in that androstenedione rather than testosterone is the predominant steroid formed; androstenedione is a very feeble androgen but, like testosterone, it stimulates body growth. The testis of the immature rat is further distinguished by the presence of reducing enzymes which convert androstenedione and testosterone to compounds of greater hormonal activity; these enzymes disappear from the testis with the onset of puberty.

The brain plays an important part in initiating puberty. The hypothalamic centers that stimulate pituitary gonadotrophin release are subject to negative feedback exerted by the steroid hormones present in the blood. In the immature animal, this feedback inhibition becomes operative at much lower levels of circulating testicular hormones than in the adult. Thus an upward resetting of the hypothalamic control center, or "gonadostat," appears to be a crucial event in the initiation of puberty. Tumors involving the posterior hypothalamus or the pineal gland are often associated with precocious puberty. Precocious development of genitalia and secondary sex characters, unaccompanied by spermatogenesis, may be due to Leydig cell or adrenal tumors. Much variation in the onset of puberty, however, is normal.

In most mammals, including humans, testicular development is dependent upon the presence of a Y chromosome. The testis-determining region of the Y chromosome has been identified, and is called SRY (for sex-determining region on the Y chromosome). The product of this gene is a protein, which binds to and activates several other genes that are necessary for testicular development. Absence of the *SRY* gene usually excludes the appearance of a testis or development of male characteristics. However, other genes that also play a pivotal role in gonadal development are still being discovered, including *DAX*, which induces ovarian development. This means that creation of the testis involves the sequential activation or inactivation of a number of genes within the developing gonad.

Biological actions of testosterone and development of the testis. At the ambisexual stage of embryonic development, the testis promotes the growth of the paired Wolffian ducts and their differentiation into the epididymis, vasa deferentia, and seminal vesicles; the fetal testis also causes masculinization of the urogenital sinus, fusion of the labioscrotal folds in the midline, and development of the genital tubercle into a phallus. All these actions can be mimicked by androgenic steroids in the organism or in organ cultures, and prevented by administration of antiandrogenic drugs. The Müllerian ducts are suppressed by the fetal testes through the release of a substance, Müllerian-inhibiting hormone or anti-Müllerian hormone (MIH or AMH), which spreads by local diffusion. This hormone is a single-chain protein, produced by the fetal Sertoli cells. Recent research also indicates that MIH may be involved in testicular descent and spermatogenic development in early postnatal life. Castration of male fetuses in mammals leads to persistence of these ducts, which then form a uterus, and to the development of external genitalia of female type. Similar abnormal development is observed in man in cases of gonadal dysgenesis due to a genetic defect, such as the absence of the Y chromosome (45/XO constitution) in Turner's syndrome. In some lower forms, however, the Y chromosome is not essential for the expression of male characters (*Drosophila*) or even for male fertility (mouse). Humans with more than one X chromosome (for example, 47/XXY) have male external genitalia, but the penis may be imperfectly formed (hypospadic) and the testes are small and produce no sperm (Klinefelter syndrome); Leydig cells are abundant, but may not produce normal amounts of androgen. Many forms of male infertility have been recognized to have a genetic cause due to more precise defects of genes involved in testicular function. Deletions of several specific regions of the Y chromosome are associated with spermatogenic failure, and disorders of the androgen receptor gene (located on the X chromosome in humans) can lead to problems ranging from lack of testicular development to arrested spermatogenesis. These disorders can be accompanied by other complications, such as neural degeneration.

In the newborn rat, the hormonal secretion of the testis has been shown to modify permanently the function of brain centers that govern the pattern of gonadotrophin release by the pituitary gland in adult life. As a result, a steady tonic secretion of luteinizing hormone is established, instead of the cyclic surges characteristic of the mature female. The male hormone also acts on other brain centers to impose male-type mating behavior and aggressiveness. However, in man early rearing and environmental factors are important in shaping sexual behavior patterns.

Testosterone promotes the retention of dietary nitrogen and its incorporation into muscle protein. This effect is responsible for the greater muscular

development of the male and contributes to the prepubertal growth spurt of boys. Therapeutically, the anabolic action of the hormone is utilized for treating patients suffering from wasting diseases, with testosterone derivatives of lesser virilizing activity used. Low doses of testosterone enhance linear growth, but higher doses accelerate epiphyseal closure in long bones. Toward puberty, increased secretion of testosterone stimulates the growth of the penis, scrotum, and male accessory glands responsible for the formation of the seminal plasma, for example, the prostate and seminal vesicles. The hormone brings about the appearance of secondary sex characters, such as the male-type distribution of hair and body fat and lowered pitch of voice in man, the growth of the comb and wattles in birds, the clasping pads of amphibians, or the dorsal spine of certain fishes. Substances that produce such effects are referred to as androgens. The thermoregulatory function of the scrotum is testosterone-dependent. Testosterone stimulates the synthesis of ribonucleic acid and protein, and of specific seminal constituents such as fructose and citric acid, in the male accessory glands. Castration leads to atrophy of these glands. Some forms of prostatic cancer regress when deprived of androgen.

Testosterone is essential for efficient spermatogenesis, but high doses of testosterone suppress spermatogenesis by inhibiting pituitary gonadotropin secretion. Protein hormones from the testis also regulate gonadotropin secretion. Follicle-stimulating hormone (FSH), which is important for establishment of the spermatogenic process at puberty and regulates Sertoli cell activity throughout life, is controlled by a protein hormone called inhibin, produced by the Sertoli cell.

Unlike the ovary, the testis remains functional throughout life, with ongoing spermatogenic development. However, the efficiency of spermatogenesis falls away, and androgen levels begin to fall due to a declining Leydig cell activity. These events can lead to reduced fertility, and androgen insufficiency problems in later life in some men, which may require medical treatment.

Control of spermatogenesis. Hypophysectomy abolishes spermatogenesis. Full restoration of spermatogenesis usually requires the replacement of both gonadotropins (LH and FSH), although large doses of testosterone and other steroids, for example, dehydroepiandrosterone and pregnenolone, temporarily maintain spermatogenesis in hypophysectomized rats. Only FSH of human or primate source is effective in humans. In addition to hormonal control, it is now recognized that many protein growth factors, or cytokines, are involved in integrating the control of testicular function. Cytokines such as interleukin-1, which is responsible for many of the manifestations of inflammatory diseases, is produced by the Sertoli cell and is involved in regulation of spermatogonial development. Other cytokines are involved in testicular development, and integration of the function of the separate cells of the testis. Disturbances in cytokine production during inflammation and disease often leads to disruption of testicular function and fertility.

Sympathetic nerves supply vasomotor fibers to the testis and motor fibers to the smooth muscles of the epididymis. Administration of the catecholamine hormones of the adrenal medulla or of serotonin causes intense vasoconstriction of the testicular vessels and damage to spermatogenesis.

Sperm transport from the testis. A fluid forms in the seminiferous tubules, about 2.4 in.^{3} (40 ml) per testis per day in the ram, which is continually transferred through the rete testis to the epididymis, carrying in it a dilute suspension of immature spermatozoa. This fluid is produced by the Sertoli cells. Almost 99% of the testicular fluid is resorbed in the head of the epididymis. As it leaves the rete, this medium is practically devoid of glucose, the chief energy source of the testis, but contains some lactate, inositol, and strikingly high concentrations of glutamic acid. The role of these substances is not well understood. Glutamine seems to favor the growth and differentiation of testes in tissue and organ culture, and polyglutamic acid, present in high concentration in the reproductive tract of the cock, has a motility-inhibiting, yet life-extending, action on sperm.

Sperm transport is probably assisted by ciliary movement in the efferent ducts and by contractile cells surrounding the tubules. The posterior hypophyseal hormone oxytocin induces contractions of the tubules and epididymis.

During sexual inactivity spermatozoa are voided continually in the urine (marsupials, ram) or by spontaneous seminal discharges, for example, in rats subjected to mechanical restraints.

Blood–testis barrier. The tubular basement membrane and the adjacent layer of Sertoli cells present a barrier to the penetration of certain substances which readily pass across most capillary membranes, for example, inulin, *p*-aminohyppuric acid, and the larger plasma proteins; others, such as Rb^+, are admitted only slowly. This blood-testis barrier is important for maintaining the specialized environment

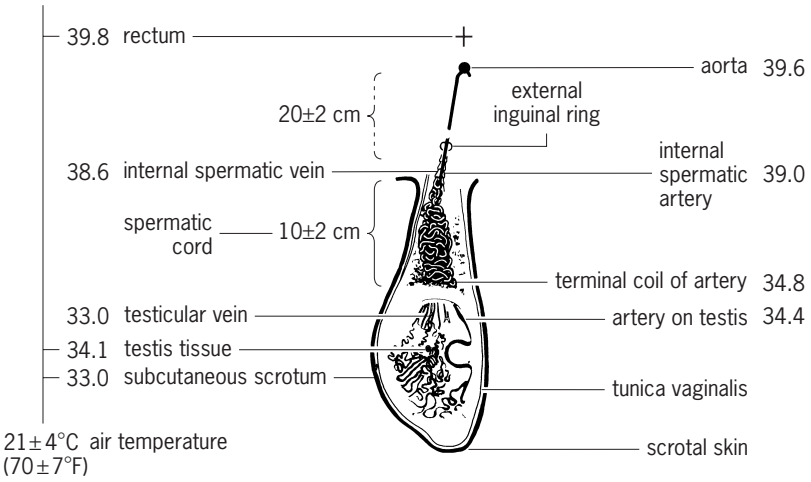

Fig. 3. Relation of testicular temperature to scrotal and deep-body (rectal) temperature, and temperature gradients along the testicular and spermatic vessels in the ram. (*After G. M. H. Waites and G. R. Moule, J. Reprod. Fert., 2:213, 1961*)

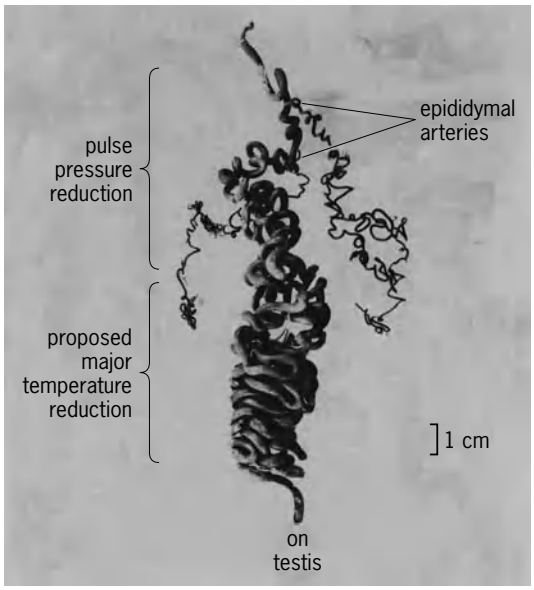

Fig. 4. Latex cast of internal spermatic artery supplying right testis of a ram. Course of the vessel from the external inguinal ring to testis is shown with the arterial coils in their natural disposition. (*From G. M. H. Waites and G. R. Moule, J. Reprod. Fert., 1:223, 1960*)

necessary for spermatogenesis, and also to sequester the developing germ cells from the immune system. Being normally sequestered from the circulation, the spermatogenic cells are not recognized as "self" by the antibody-forming tissues. Thus, when testicular tissue is injected into an animal or when the testis is damaged by inflammation, autoimmune reactions may be induced against the seminiferous tissue and cause its destruction.

Temperature control of testis. In most mammals the testes descend, either permanently or for the duration of the breeding season, into an extraabdominal pouch, the scrotum, where its temperature is maintained 5–13°F (3–7°C) below the animal's deep-body temperature. If this temperature gradient is eliminated by failure of the testis to descend (cryptorchidism) or by severe environmental heat stress, spermatogenesis is impaired and eventually abolished. The reasons for this are not understood. Glucose and oxygen supply to the tubules may become inadequate during scrotal heating. The pachytene stage of the prolonged meiotic prophase of the spermatocytes seems to be particularly sensitive to elevated temperature, but dividing spermatocytes and early stages of spermiogenesis are also adversely affected; spermatogonia are somewhat more resistant. The requirement for low testicular temperature is not universal: Birds, whose body temperature is generally higher than that of mammals, have intraabdominal testes. The testes remain in the abdomen in Monotremata (such as the duckbilled platypus), some Insectivora (such as Madagascar "hedgehogs"), Edentata (such as armadillos), Cetacea (such as whales), Proboscidae (African elephants), Hyracoidea (such as rock hyrax), and Sirenia (such as sea cows).

Testicular descent is accelerated by the luteinizing hormone and testosterone, but may be mechanically obstructed in the presence of teratomatous growth or tumor.

The scrotal skin is thin, devoid of subcutaneous fat, and richly supplied with small blood vessels. It contains heat-sensitive nerve endings of the superficial perineal nerves whose firing rate increases as scrotal temperature rises (above 95°F or 35°C in the ram). Heat loss is effected by increasing the scrotal surface through relaxation of the tunica dartos muscle; by increased scrotal blood flow, including the opening up of arteriovenous shunts; by more frequent synchronous discharges of the large scrotal apocrine sweat glands, triggered by efferent adrenergic fibers in the external spermatic nerves; and by reflex thermal polypnea (panting). The scrotal cooling effect is transmitted to the deeper parts of the gland by the precooling of the arterial inflow through countercurrent heat exchange with the returning venous blood (**Fig. 3**): In most mammals the spermatic artery is extremely long and coiled (**Fig. 4**), thin-walled and closely applied to the veins of the pampiniform plexus. This arrangement also serves to attenuate the arterial pulse (**Fig. 5**) and to utilize its energy to aid venous return. In marsupials, the spermatic artery branches into a leash of more than 100 vessels surrounded by parallel veins. In species with intra-abdominal testes, the spermatic vessels are

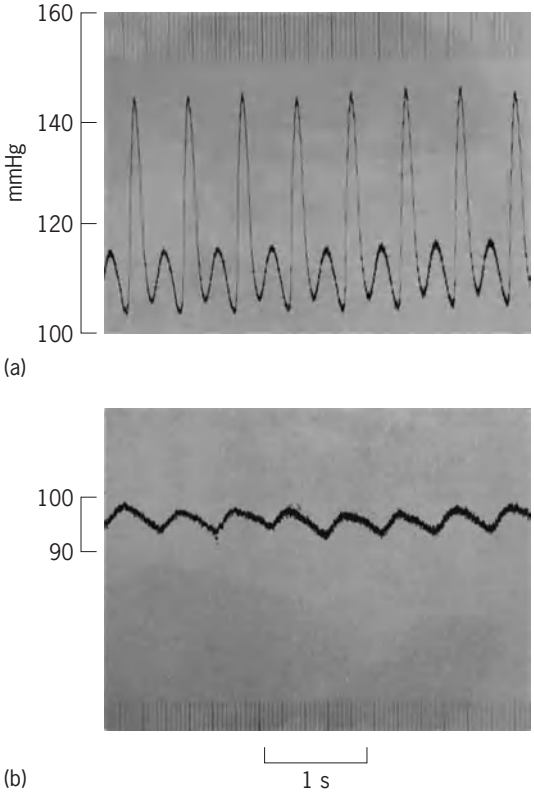

Fig. 5. Record of pressure changes in internal spermatic artery of a dog. (*a*) Above vascular cone. (*b*) On testis. Note attenuation of arterial pulse during its passage through coiled part of artery. 1 mmHg = 133 pascals. (*From G. M. H. Waites and B. P. Setchell, The Gonads, Appleton-Century-Crofts, 1969*)

generally short and straight. *See* REPRODUCTIVE SYSTEM. Mark P. Hedger; Hans R. Lindner

Bibliography. D. M. de Kretser (ed.), *Cellular and Molecular Mechanisms in Male Reproduction*, 1993; C. Desjardins and L. L. Ewing (eds.), *Cell and Molecular Biology of the Testis*, 1993; A. Gorbman et al., *Comparative Endocrinology*, 2d ed., 1987; E. Knobil et al. (eds.), *The Physiology of Reproduction*, 1988; K. McElreavey (ed.), *The Genetic Basis of Male Infertility*, 2000; J. D. Wilson, *Williams Textbook of Endocrinology*, 9th ed., 1998.

Tetanus

An infectious disease, also known as lockjaw, caused by the exotoxin of *Clostridium tetani*. The organism may be isolated from fertile soil and the intestinal tract or fecal material of humans and other animals. Infection commonly follows dirt contamination of deep wounds or other injured tissues such as war wounds or crash injuries, burns, or septic abortion in which tissue necrosis is present. Any deep puncture wound that is contaminated by soil or feces is suspect. Tetanus neonatorum is frequent in developing countries if the cut surface of the umbilical cord is not properly treated. Use of understerilized syringe needles by drug users is also recognized as a cause of tetanus. *See* TOXIN.

The causal organism is a strictly anaerobic, slender bacillus with a spherical terminal spore swelling the vegetative cell. Identification of the organism, however, should not rest on microscopic evidence but should be confirmed by neutralization tests with known tetanus antitoxin of the toxin produced by the pure culture. This species produces two toxic substances: a hemolysin known as tetanolysin, which is probably not involved in the disease, and the potent lethal toxin known as tetanospasmin, which has a strong affinity for the cells of the central nervous system. The neurotoxin has been isolated in crystalline form. This organism does not ferment carbohydrates but depends on the fermentation of amino acids for energy. *See* TOXIN-ANTITOXIN REACTION.

The diagnosis of tetanus is on clinical symptoms, and the incubation period is usually 5–10 days. The disease is characterized by convulsive tonic contraction of voluntary muscles. Therapy of suspected or diagnosed cases includes antibiotics (penicillin, tetracyclines) combined with injections of human antitoxin; equine antitoxin is to be avoided since the individual may be hypersensitive to horse serum. Prevention of tetanus rests on the proper, prompt surgical care of contaminated wounds and prophylactic use of antitoxin if the individual has not been protected by active immunization with toxoid. Babies aged 1–3 months should be actively immunized with DPT (diphtheria, pertussis, tetanus) toxoid; this immunization should be repeated at 10-year intervals. *See* IMMUNOLOGY. Leland S. McClung

Bibliography. S. M. Finegold, *Anaerobic Infections in Humans*, 1989; L. S. McClung, *The Anaerobic Bacteria: Their Activities in Nature and Disease*, 7 vols., 1982; G. L. Mandell, R. G. Douglas, and J. E. Bennett (eds.), *Principles and Practice of Infectious Diseases*, 5th ed., 2000; L. L. Simpson (ed.), *Anaerobic Bacteriology: Clinical and Laboratory Practice*, 3d ed., 1977; A. J. Zehnder (ed.), *Botulinum Necrotoxin and Tetanus Toxin*, 1989.

Tetraodontiformes

An order of the class Actinopterygii, also known as the Plectognathi. This most highly derived order of actinopterygian fishes is noted for its reduction in skeletal elements. Lost are the parietal, nasal, and infraorbital bones of the skull. Usually there are no pleural ribs; vertebrae are as few as 16 and no more than 30; and the maxillae are united with the premaxillae. The gill opening is a short vertical slit in front of the pectoral fin. Scales are usually modified to form spines, shields, or plates; a lateral line may be present or absent; and a swim bladder is present, except in the family Molidae. The 357 extant species are in three suborders, nine families (see **illustration**), and about 101 genera. *See* ACTINOPTERYGII; OSTEICHTHYES; SWIM BLADDER.

Suborder Triacanthodoidei. The single family, Triacanthodidae (spikefishes), is characterized by a deep and moderately compressed body, which is covered with a thick skin and tiny scales; each scale bears spicules, resulting in the skin having a shagreen-like appearance (that is, rough-looking skin covered with small close-set tubercles); the caudal fin is rounded to truncate; there are usually six dorsal fin spines and 12–18 dorsal fin rays; 11–16 anal rays; and a caudal fin with 12 principal rays. Each pelvic fin consists of one stout spine and one or two soft rays; the pelvic spines can be locked into place. There are 20 vertebrae. Some species have an extremely long snout; for example, in *Halinochirurgus akacki* the snout is about equal in length to the entire body posterior to the eye. Species of spikefishes range in total length from about 5.3 to 26 cm (2 to 10 in.). They are benthic in deep waters of the tropical and subtropical Atlantic and Indo-Pacific oceans.

Suborder Balistoidei. Balistoidei is also known as Sclerodermi. In the following four families, the frontal bones extend far anterior to the articulation between the lateral ethmoid and ethmoid bones.

Triacanthidae (triplespines). This family, formerly placed in the suborder Triacanthodoidei, is superficially similar to the Triacanthodidae but can be distinguished from them by their deeply forked caudal fin vs. round or truncate caudal fin; dorsal fin rays 19–26 vs. 12–18; anal fin rays 13–22 vs. 11–16; and maximum length about 28 cm (11 in.). The family comprises four genera and seven species, limited to shallow waters of the Indo-Pacific Ocean.

Balistidae (triggerfishes). This family is characterized by a deep and moderately compressed body covered with a very thick, tough skin bearing rectilinear scale plates. The scales above the pectoral fin base are usually enlarged and slightly separated, forming a

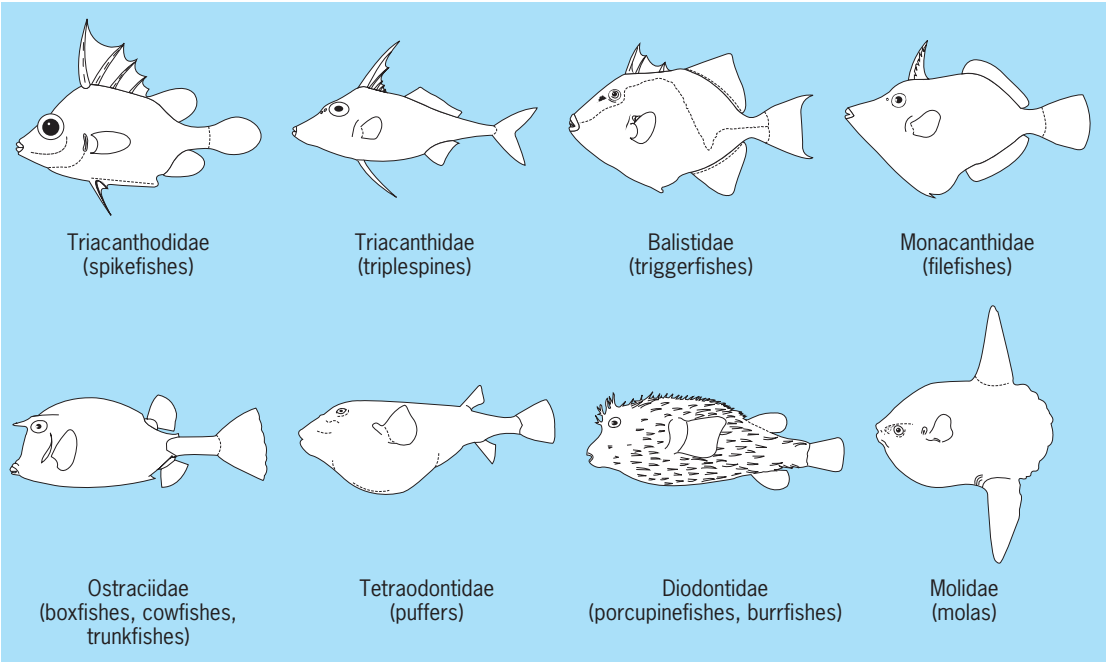

Representations of eight families (Triodontidae is not pictured) of the Tetraodontiformes. (*Courtesy of J. S. Nelson, Fishes of the World, 4th ed., Wiley, 2006*)

flexible tympanum. There is a small and usually terminal mouth, and a row of eight strong outer teeth in both the upper and lower jaws. There are three dorsal fin spines, the first of which can be locked in an upright position by the second; pelvic fins and spines are rudimentary or absent. There are 18 vertebrae. Eleven genera and about 40 species occur in the Atlantic, Indian, and Pacific oceans, either pelagic open water or benthic around rocky or coral reefs.

Monacanthidae (filefishes). Filefishes are characterized by a deep and compressed body covered with tiny scales, which are prickly or velvety to touch. There are usually two dorsal spines; the second is usually much smaller or absent. The upper jaw usually has three teeth in the outer series and two in the inner on each premaxillary, developed for nibbling. There are 19–31 vertebrae. Most species feed on a wide variety of benthic invertebrates, but some specialize on corals or zooplankton. Filefishes lay demersal (living at or near the bottom of the sea) eggs in a site prepared and guarded by the male or both parents. Some of the subtropical species release eggs in open water. Filefishes are generally small, and some reach a length of only 3.6 cm (1.4 in.); however, the largest member of the family, *Aluterus scriptus*, attains a length of 110 cm (43.3 in.). The family comprises approximately 102 marine species, which occur in the Atlantic, Indian, and Pacific oceans; about half of them occur in Australian waters.

Ostraciidae (boxfishes, cowfishes, trunkfishes). The Ostraciidae are characterized by the entire body (except the caudal peduncle) being encased in an immovable bony carapace consisting of fused hexagonal plates; no pelvic skeleton; no spinous dorsal; dorsal and anal fins each with 9–13 rays; nonprotrusible upper jaw; usually 18 vertebrae; and maximum length of 11–55 cm (4.3–21.6 in.), depending on the species. Some species secrete ostracitoxin, poisonous to other fishes and to some extent other ostracians. The habitat is usually seagrass and coral, where they feed on worms and other small sessile invertebrates. Some species use strong jets of water from the mouth to dislodge prey from the substrate (an interesting adaptation for a fish with a small nonprotrusible mouth and few moving body parts). Movement is slow and accomplished primarily by a sculling motion of the slender caudal peduncle. The family comprises two subfamilies, Aracaninae (a ventral ridge more or less developed and 11 principal caudal rays) and Ostraciinae (no ventral ridge and 10 principal caudal rays). Some authors recognize the subfamilies as separate families. There are about 33 species, some with beautiful bright colors, in the tropical and temperate waters of the Atlantic, Indian, and Pacific oceans.

Suborder Tetraodontoidei. These fishes do not have true teeth; instead the upper and lower jaws have sharp cutting edges that form a beak similar to that seen in the parrotfishes; there may be two, three, or four such "teeth," depending on the presence or absence of sutures; and the posttemporal bones are absent, as are usually the urohyal, pelvis, and pelvic fins.

Triodontidae (threetooth puffers). The single species of this family has three fused teeth (a median suture in the upper jaw, none in the lower); pelvis present; dorsal and anal fins each usually with 11 rays; caudal fin with 12 principal rays and deeply forked; ribs and epipleurals present; large belly flap; and maximum total length about 54 cm (21 in.). It is a reef associate of the Indo-West Pacific to depths of 50–300 m (164–984 ft).

Tetraodontidae (puffers). Members of this family are identified by a robust body, naked or with short prickles on sides and belly; some species with small fleshy appendages (lappets) on the sides; jaw teeth fused but separated by a median suture in each jaw, giving rise to four fused teeth; dorsal and anal fins far posteriorly, each consisting of 7–18 dorsal soft rays; ribs and epipleurals lacking; caudal fin with 10 principal rays and moderately forked to rounded; and maximum size of 3.5–120 cm (1.4–47.2 in.) total length, depending on the species. Puffers are capable of greatly inflating themselves with water or air when agitated. Some puffers contain the potential lethal poison tetrodotoxin, especially in the viscera and in the gonads of some during spawning season. They are chiefly marine, usually inhabiting shallow inshore waters of tropical and temperate waters of the Atlantic, Indian, and Pacific oceans; however, many species inhabit brackish water, but only 12 of about 162 species of puffers are limited to freshwater. *See* TOXIN.

Diodontidae (porcupinefishes, burrfishes). Members of this family are also capable of inflating the body with water or air, but differ from the puffers in having strong sharp spines and two fused teeth in the jaws, which form a powerful parrotlike beak. The spines of burrfishes are permanently erect, whereas those of the porcupinefishes are erect when the body is inflated. The family is further identified by opposite premaxillaries and dentaries being fused at midline. Maximum size is 27–91 cm (10.6–35.8 in.) total length. The family is represented by 19 species in tropical and temperate waters of the Atlantic, Indian, and Pacific oceans. Adults occupy inshore waters, whereas eggs and young are pelagic. Adults feed primarily on hard-shelled invertebrates crushed by the beak.

Molidae (molas). Molas are characterized by a moderately compressed body and thick leathery skin; tiny mouth and jaws with two fused beaklike teeth; no caudal peduncle or true caudal fin; posterior end of body reduced to a leathery flap (clavus) or a pseudocaudal fin formed by posteriorly migrated dorsal and anal fin rays; no pelvic fins; no lateral line; no swim bladder; and 16–18 vertebrae. Molas are strong swimmers, using the sculling motion of the powerful dorsal and anal fins. The ocean sunfish (*Mola mola*) is probably the most fecund fish with an estimated 300 million eggs. The young differ markedly from adults in having a spiny globular body. Adults feed largely on jellyfishes and other soft-bodied pelagic invertebrates and grow to 337 cm (11 ft) in total length and 1500 kg (3300 lb) in weight. Distribution is worldwide in tropical to temperate seas. Herbert Boschung

Bibliography. T. Abe, Balistidae, in W. Fischer and P. J. P. Whitehead (eds.), *FAO Species Identification Sheets for Fishery Purposes: Eastern Indian Ocean (Fishing Area 57) and Western Central Pacific (Fishing Area 71)*, vol. 1, FAO, Rome, 1974; K. Matsuura, Balistidae: Triggerfishes, pp. 3911–3928 in K. E. Carpenter and V. Niem (eds.), *FAO Species Identification Guide for Fishery Purposes: The Living Marine Resources of the Western Central Pacific*, vol. 6, FAO, Rome, 2001; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006; C. R. Robins and G. C. Ray, *A Field Guide to Atlantic Coast Fishes of North America*, Houghton Mifflin, Boston, 1986; R. Santini and C. Tyler, A phylogeny of the families of fossil and extant tetraodontiform fishes (Acanthomorpha, Tetraodontiformes), Upper Cretaceous to Recent, *Zool. J. Linn. Soc.*, 139:565–617, 2003; M. M. Smith and P. C. Heemstra, Balistidae, pp. 876–882 in M. M. Smith and P. C. Heemstra (eds.), *Smith's Sea Fishes*, Springer-Verlag, Berlin, 1986; E. Tortonese, Molidae, pp. 1348–1350 in P. J. P. Whitehead et al. (eds.), *Fishes of the North-eastern Atlantic and the Mediterranean*, vol. 3, UNESCO, Paris, 1986; J. C. Tyler, Balistidae, Diodontidae, Monacanthidae, Ostraciidae, in W. Fischer (ed.), *FAO Species Identification Sheets for Fishery Purposes: West Atlantic (Fishing Area 31)*, vols. 1–5, FAO, Rome, 1978; J. C. Tyler, *Osteology, Phylogeny, and Higher Classification of the Fishes of the Order Plectognathi (Tetraodontiformes)*, NOAA Tech. Rep. NMFS Circ. 434, October 1980.

Tetraphididae

A subclass of the mosses (class Bryopsida) consisting of two families and three genera, especially characterized by growth from protonematal flaps, three-ranked leaves, and peristomes of four teeth made up of whole cells (rather than thickened parts of cells). The Tetraphididae include small acrocarpous mosses with peristome teeth in fours (see **illus.**). The plants grow from buds produced on leaflike protonematal flaps. They are erect and simple or merely forked, with oblong-ovate leaves in three rows. The leaves usually have a single costa that ends near the apex. The cells are short and smooth. Brood bodies are generally present in terminal clusters with or without a cuplike whorl of differentiated leaves. The setae are elongate, and the capsules are cylindric to

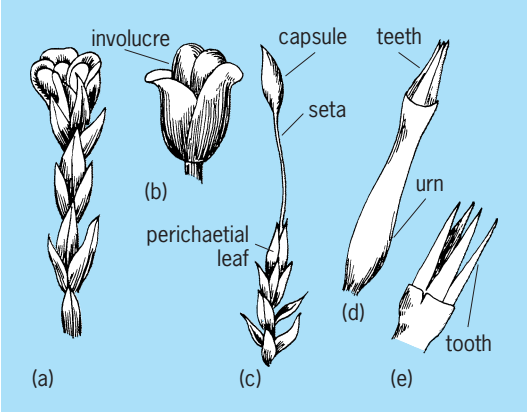

***Tetraphis pellucida*. (*a*) Gemmiferous branch; (*b*) involucre with gemmae (*after W. H. Welch, Mosses of Indiana, Ind. Dept. Conserv., 1957*). (*c*) Sporophyte; (*d*) capsule with peristome (*after H. S. Conrad, How To Know the Mosses, Jaques, 1944*). (*e*) Enlarged peristome (*after W. H. Welch, Mosses of Indiana, Ind. Dept. Conserv., 1957*).**

ovoid-cylindric and erect, with a distinct suboral rim of dark cells. The operculum is conic or subulate-pointed, and the peristome teeth are wedge-shaped, consisting of elongate cells thickened all around and derived from the entire tissue internal to the operculum. The calyptrae are mitrate and plicate. Chromosome numbers are 7 and 8. *See* BRYOPHYTA; BRYOPSIDA.

Howard Crum

Tetraphyllidea

An order of tapeworms of the subclass Cestoda. All species are intestinal parasites of elasmobranch fishes and are small in size, usually less than 2 in. (5 cm) in length. An outstanding feature of the order is the variation in the structure of the holdfast organ or scolex (see **illus.**). All species are segmented, and

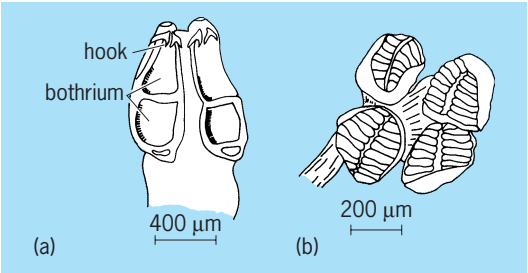

Tetraphyllidea. (*a*) *Acanthobothrium* sp., scolex. (*b*) *Rhinebothrium* sp., scolex.

segments are usually shed from the body while sexually immature; these develop to sexual maturity as independent units in the host's intestine. Segment anatomy is very similar to that of Proteocephaloides. A complete life cycle is not known, but larval forms have been found in a variety of invertebrates and bony fishes. *See* EUCESTODA; PROTEOCEPHALOIDEA.

Clark P. Read

Tetrapoda

The superclass of the subphylum Vertebrata whose members typically possess limbs in contrast to the other superclass, the Pisces (fishes), whose members have fins. *See* PISCES (ZOOLOGY).

The animals making up the Tetrapoda typically live part or all of their lives on land, whereas the members of the Pisces live in water. The classes of the Tetrapoda are Amphibia (frogs and toads, salamanders, and caecilians), Reptilia (snakes and lizards, turtles, and crocodiles and their kin), Aves (birds), and Mammalia (mammals). The term Tetrapoda comes from Greek words meaning "four feet," but there are tetrapods that have only two limbs or none at all, such as some amphibians and reptiles. These forms have, however, evolved from four-footed ancestors.

The division of the vertebrates into the superclasses Tetrapoda and Pisces is in some respects merely a classification of convenience, and the use of characters other than limbs and fins can result in a different separation. For example, the eggs of reptiles and birds possess an embryonic membrane called the amnion that permits development of the embryo in relatively dry situations. The same membrane also surrounds a developing mammal, and these animals may be classified as the Amniota. Amphibians and other lower vertebrates do not have the amnion; their eggs survive only in water or under very moist conditions. These animals may be classified as the Anamnia. *See* AMNIOTA; AMPHIBIA; ANAMNIA; AVES; MAMMALIA; REPTILIA.

Richard G. Zweifel

Teuthoidea

An order of the class Cephalopoda (subclass Coleoidea) commonly known as squids. They are characterized by 10 appendages (eight arms and two longer tentacles) around the mouth; an elongate, tapered, usually streamlined body; an internal, rod- or bladelike chitinous shell (gladius); and fins on the body. The two tentacles are strongly elastic, contractile, but not retractile into pockets as in cuttlefishes (Sepioidea). Two rows of suckers (infrequently four or six rows) occur on the arms on muscular stalks, with sucker rings that are chitinous, smooth, toothed, or modified as clawlike hooks. The muscular tentacles have terminal clubs with two rows, usually four, ranging up to many rows of suckers (and/or hooks in some families). Adults of the family Octopoteuthidae and genera *Gonatopsis* and *Lepidoteuthis* characteristically lose their tentacles. *See* SEPIOIDEA.

The Teuthoidea are divided into two suborders. The Myopsida, the nearshore, shallow-water squids, have a transparent skin (cornea) covering the two eyes, with a minute pore anteriorly, arms and tentacular clubs with suckers only, never hooks, and a single gonoduct in females, not paired. The Oegopsida, the oceanic squids, have no cornea over the eyes and no anterior pore, arms and tentacular clubs with suckers (and/or hooks in many families), and paired gonoducts in females (some exceptions).

Squids inhabit a wide variety of marine habitats, depending on the species, from very shallow grass flats, mangrove roots, lagoons, bays, and along coasts (myopsids) to the open ocean from the surface of the sea (*Ommastrephes*) to nearly 9600 ft (3000 m) in the deep sea (other oegopsids such as *Bathyteuthis*, *Neoteuthis*, and *Grimalditeuthis*). Most species require full oceanic salinities, 32–37‰ (parts per thousand), but some, like *Lollinguncula brevis*, can tolerate salinities as low as 8.5‰ in the warm-water estuaries and bays of the Atlantic coasts of North and South America and the Caribbean. While some species of squids live near the bottom where they feed, spawn, or rest, principally the myopsids, most species live near the surface or in the middepths away from the bottom. Even the deepest-living species of oegopsids live far above the abyssal ocean bottom.

Teuthoids are the masters of animal jet propulsion; they suck water into the mantle cavity, then forcibly eject it through the funnel, driving the squid

tail-first through the water. The propulsion system is controlled from the well-developed brain by a complex of nerves concentrated in a cluster in the mantle, chief of which is the giant axon, the largest single nerve fiber in the animal kingdom. The highly developed brain and nervous system also support large, acute eyes that rival those of fishes, the ability to rapidly change color and color patterns through thousands of pigment cells (chromatophores), and the capacity to produce light (bioluminescence) from specialized organs (photophores).

The sexes are separate, and the male transfers sperm to the female, often following dramatic and colorful courtship displays, in cylindrical packets (spermatophores), frequently via a specially modified arm, the hectocotylus. The fertilized eggs are laid in gelatinous fingerlike masses attached to the bottom (myopsids) or in gelatinous sausages or balls, or as individual eggs, in the open sea (oegopsids). Embryos undergo direct development and upon hatching are planktonic young that look much like adults; no dramatic larval stages occur in squids, as they do in most other marine mollusks and invertebrates. Adult teuthoids can be from about 1 in. (25 mm) total length (myopsid, *Pickfordiateuthis pulchella*) to 60 ft (18 m) total length, the giant squid (oegopsid, *Architeuthis* species). Squids are preyed upon by fishes, toothed whales, seals, and sea birds, and they account for around 10^6 metric tons captured annually in fisheries around the world for nutritious human food. Scientists study the biology and behavior of squids, as well as their giant axons for neurophysiology, pharmacology, and biophysics. *See* CEPHALOPODA; COLEOIDEA; SQUID. Clyde F. E. Roper

Bibliography. K. N. Nesis, *Cephalopods of the World*, transl. from Russian, 1987; C. F. E. Roper, M. J. Sweeney, and C. E. Nauen, *Cephalopods of the World*, FAO Fisher. Synop. 125, vol. 3, 1984; E. R. Truman and M. R. Clarke (eds.), *The Mollusca*, vol. 11: *Form and Function*, 1988.

Textile

A material made mainly of natural or synthetic fibers. Modern textile products may be prepared from a number of combinations of fibers, yarns, films, sheets, foams, furs, or leather. They are found in apparel, household and commercial furnishings, vehicles, and industrial products. Materials made solely from plastic sheet or film, leather, fur, or metal are not usually considered to be textiles. *See* LEATHER AND FUR PROCESSING; MANUFACTURED FIBER; NATURAL FIBER; PLASTICS PROCESSING.

The term fabric may be defined as a thin, flexible material made of any combination of cloth, fiber, or polymer (film, sheet, or foams); cloth as a thin, flexible material made from yarns; yarn as a continuous strand of fibers; and fiber as a fine, rodlike object in which the length is greater than 100 times the diameter. The bulk of textile products are made from cloth, and this discussion centers on the most common forms of cloth and their manufacture.

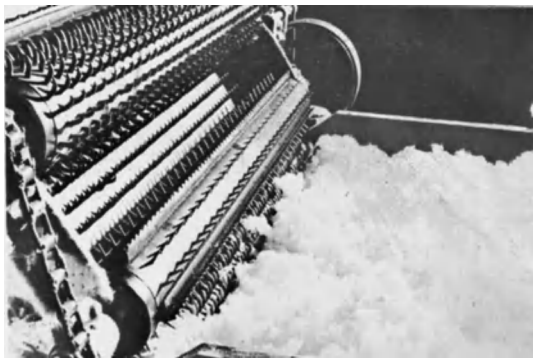

Fig. 1. Cotton going through opening machine where fibers are loosened and straightened. (*Pepperell Mfg.*)

The natural progression from raw material to finished product requires the cultivation or manufacture of fibers; the twisting of fibers into yarns (spinning); the interlacing (weaving) or interlooping (knitting) of yarns into cloth; and the finishing of cloth prior to sale.

Spinning processes. The ease with which a fiber can be spun into yarn is dependent upon its flexibility, strength, surface friction, and length. Exceedingly stiff fibers or weak fibers break during spinning. Fibers which are very smooth and slick or fibers which are very short do not hold together. To varying degrees, the common natural fibers (wool, cotton, and linen) have the proper combinations of the above properties. The synthetic fibers are textured prior to use in order to improve their spinning properties by simulating the convolutions of the natural fibers. Natural and synthetic filament fibers, because of their great length, need not be twisted to make useful yarns.

The properties of a yarn are influenced by the kind and quality of fiber, the amount of processing necessary to produce the required fineness, and the degree of twist. The purpose of the yarn determines the amount and kind of processing. The yarn number (yarn count) is an indication of the size of a yarn—the higher the number, the finer the yarn. The degree of twist is measured in turns per inch (tpi) and is varied from three to six times the square root of the yarn number for optimum performance.

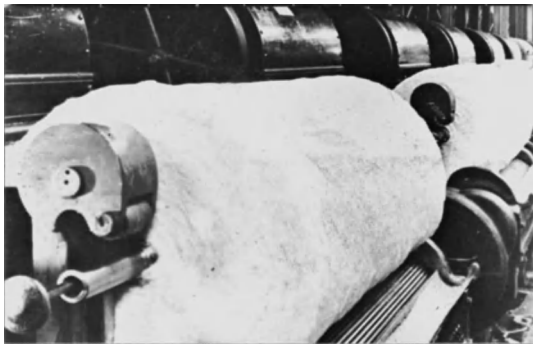

Fig. 2. Cotton lap from the picker room where the dust, leaves, twigs, and other foreign matter have been removed. (*Pepperell Mfg.*)

Fig. 3. Sliver leaving carding machine, where cotton is further cleaned and disentangled. (*Pepperell Mfg.*)

Fig. 4. After carding, the slivers are doubled to increase the density of the future cotton yarn. (*Pepperell Mfg.*)

The conversion of staple fiber into yarn requires the following steps: picking (sorting, cleaning, and blending), carding and combing (separating and aligning), drawing (reblending), drafting (reblended fibers are drawn out into a long strand), and spinning (drafted fibers are further attenuated and twisted into yarn).

Picking. Raw fibers are delivered to the spinster from a number of suppliers. They must be sorted, trash must be removed, and fibers must be blended in order to assure quality and uniformity. **Figure 1** shows tufts of raw cotton being fed into an opener where the spiked teeth will pull the tufts apart and separate the fiber from its impurities. The product of this step is the picker lap (**Fig. 2**).

Carding. The picker lap is still too tangled and coarse to be suitable for spinning. Fibers are aligned in the carding process, in which the lap is unrolled and drawn on a revolving cylinder covered with very fine wire brushes. A moving belt covered with wire brushes lies atop the cylinder. The motion of the belt and cylinder pulls the fibers, disentangles them, and arranges them in parallel in the form of a thin web. The web is drawn through a funnel-shaped tube that molds it into a round ropelike mass, about the thickness of a broomstick, called sliver (**Fig. 3**). Card sliver produces yarns serviceable for inexpensive fabrics. After carding, several slivers are combined to form a relatively narrow lap of compacted staple fibers. The compactness of these fibers permits this stock to be drawn out to a sliver of smaller diameter without falling apart (**Fig. 4**).

Combing. When the fiber is intended for fine yarns, the sliver is put through an additional alignment. In this operation, fine-toothed combs continue straightening the fibers until they are arranged with such a high degree of parallelism that the short fibers are combed out and separated from the long fibers. The comb sliver is made of the longest fibers and produces a smoother and more even yarn. The combined process, therefore, is identified with better-quality goods.

Drawing and drafting. The combining of several slivers in the drawing process eliminates irregularities which would affect yarn quality. The drawing frame has several pairs of rollers, each advanced set of which revolves at a progressively higher speed. This action pulls the fibers lengthwise over each other and produces a longer, thinner sliver. After several stages of drawing, the condensed sliver is given a slight twist and wound onto bobbins. The product is called roving.

Spinning. The roving is placed on the spinning frame (**Fig. 5**) and passed through another set of drafting rolls. The highly attenuated yarn is fed onto a high-speed spindle by a guide (traveler) which rotates on a ring surrounding the spindle. The traveler rotates at a speed slightly slower than that of the spindle. This difference in speed causes twist to be inserted in the yarn. At the same time, the traveler is raised and lowered along the length of the spindle to form the yarn into a neat package.

It should be noted that modern spinning plants are equipped with integrated machines which perform many of these operations within a single unit; for example, opening, blending, and carding are now combined in one operation. In addition, the new open-end spinning systems can convert card sliver into yarn without the necessity of separate drawing and drafting steps.

Weaving. The process of weaving allows a set of yarns running in the machine direction (warp) to be interlaced with another set of yarns running across

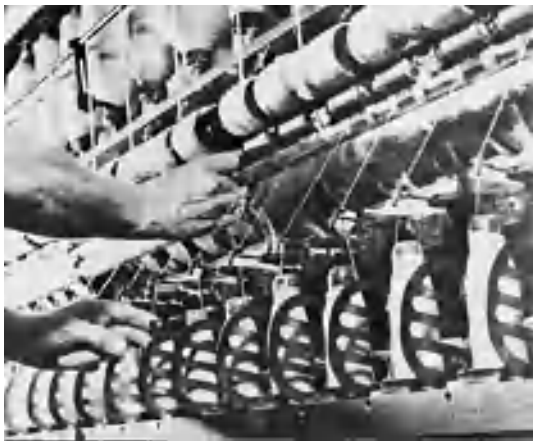

Fig. 5. On spinning frame, roving passes from the top through a series of rollers that draw out the cotton into thread. The thread is twisted as it is being wound onto the bobbin. (*Pepperell Mfg.*)

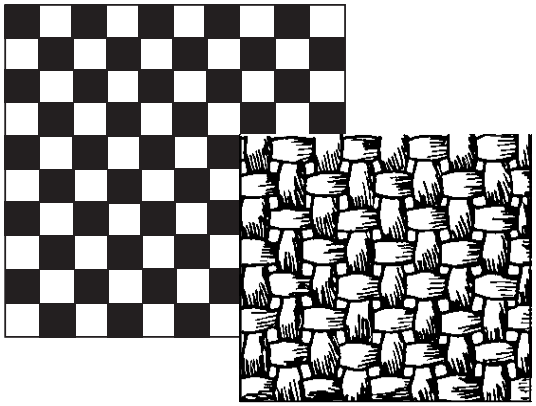

Fig. 6. Construction design for plain weave; filling yarns pass under and over alternate warp yarns, as shown at right. When fabric is closely constructed, there is no distinct pattern. (*After M. D. Potter and B. P. Corbman, Fiber to Fabric, 3d ed., McGraw-Hill, 1959*)

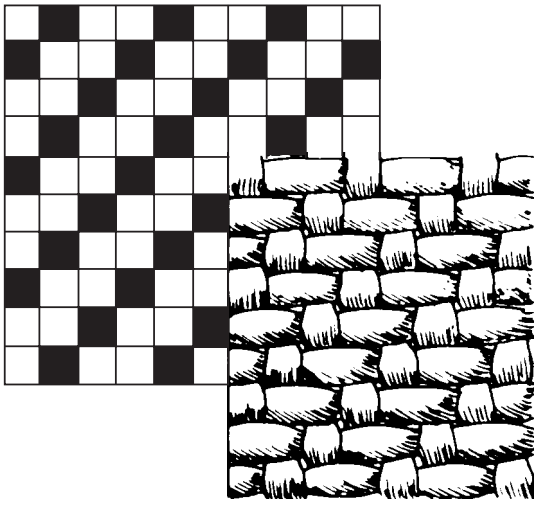

Fig. 7. Three-shaft twill. Two warp yarns are interlaced with one filling yarn. (*After M. D. Potter and B. P. Corbman, Fiber to Fabric, 3d ed., McGraw-Hill, 1959*)

the machine (filling or weft). The weaving process involves four functions: shedding (raising the warp yarns by means of the appropriate harnesses); picking (inserting the weft yarn); battening (pushing the weft into the cloth with a reed); and taking up and letting off (winding the woven cloth onto the cloth beam and releasing more warp yarn from the warp beam).

Mechanical systems. From the earliest handlooms to the most innovative weaving machines, weaving systems have six basic components: a warp beam, heddles, harnesses, a means for inserting the filling yarns (shuttle, rapier, or fluid jet), a reed, and a takeup roll. The warp beam is a large metal spool capable of holding about 4500 separate yarns, each about 500 yards (450 m) long; it holds the warp yarns and feeds them into the cloth. Heddles are thin, flat metal rods pierced by a central slot, and are used for separating and guiding the warp yarns. Harnesses are frames which can be used to raise and lower groups of heddles. The reed is a large, comblike device that packs the filling yarns into the cloth so that they do not slip. The takeup roll looks much like a warp beam, and is used to hold the woven cloth. The warp yarns are raised preparatory to inserting the weft yarn.

In a process called beaming, the warp yarns are taken from packages mounted on a large rack (creel), passed through a comb which separates and guides them, passed through a bath of starch or other viscous material (sizing) which applies a protective lubricant, and wound onto the warp beam. The positioning of the yarns in the heddles is determined by the particular weave pattern. The simplest weaves require only two harnesses, while more complex patterns may take up to 40. The warp yarns are fixed to the takeup roll after passing through the reed. The filling yarns are wound on small spindles (quills) which can be mounted in the shuttle. Most looms are equipped with automatic devices which change quills when they have delivered their yarn.

Classification of weaves. The manner in which groups of yarns are interlaced determines the pattern of the weave and, to a great extent, the properties of the cloth. The three basic weaves are the plain, twill, and satin (**Figs. 6–8**) and their respective variations. Important constructions are also obtained from more elaborate weaves such as dobby, swivel, lappet, gauze, double cloth, and pile. Jacquard looms can produce cloth of extremely intricate patterns.

Knitting. Knit cloth is produced by interlocking one or more yarns through a series of loops. The lengthwise columns of loops are known as the wales, and the crosswise rows of loops are called courses. Filling (weft) knits (**Fig. 9**) are those in which the

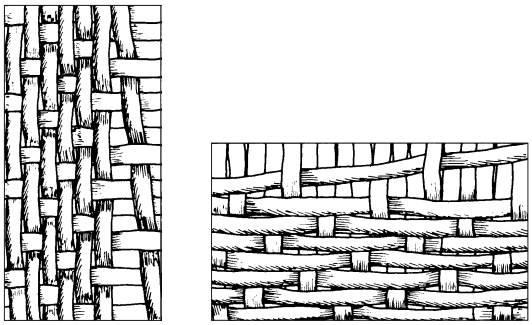

(a) (b)

Fig. 8. Long floats typical of satin and sateen weaves: (*a*) five-shaft satin construction with floats in warp direction; (*b*) eight-shaft sateen construction with floats in the filling direction. (*After M. D. Potter and B. P. Corbman, Fiber to Fabric, 3d ed., McGraw-Hill, 1959*)

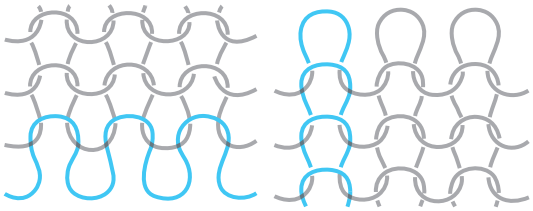

(a) (b)

Fig. 9. Interlocking yarns of (*a*) course and (*b*) wale in a jersey knit cloth. (*After B. P. Corbman, Fiber to Fabric, 5th ed., McGraw-Hill, 1975*)

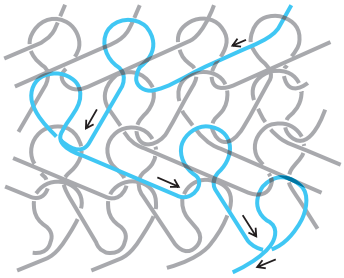

Fig. 10. Single-warp (one-bar) tricot knit. (*After B. P. Corbman, Fiber to Fabric, 5th ed., McGraw-Hill, 1975*)

courses are composed of continuous yarns, while in warp knits (**Fig. 10**) the wale yarns are continuous.

Weft. This type of knitting is accomplished on linear (flat knits) or cylindrical (circular knits) machines in which latch needles are raised and lowered by the action of cams (**Fig. 11**). As the needle rises, the hooked end loops over the yarn. As the needle descends, the yarn is trapped within the crook by the previously formed loop. At the bottom of its motion, the earlier loop slips off the needle, while the new loop is held within the closed needle. At the next traverse of the cam, the process is repeated.

Warp. This type of knitting differs from weft knitting in that each needle is supplied with its own yarn, taken from a warp beam similar to that used in weaving. A yarn guide lays the yarn across the ascending needle, and returns to position as the needle descends. The yarn, which is fixed at one end by the previous loop and at the other by the yarn guide, is trapped within the crook of the spring beard needle as the needle descends. A depresser bar closes the hook, allowing the old loop to slip off while the new loop is held.

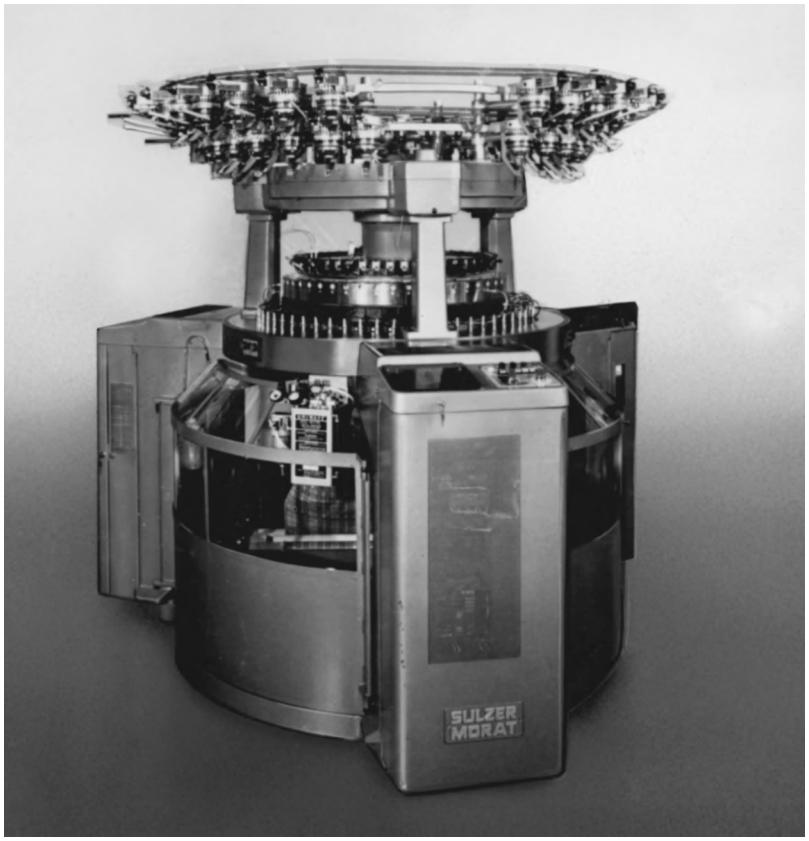

Fig. 11. Circular Jacquard knitting machine. (*Sulzer Co.*)

Classification of knits. In weft knitting, four stitches may be formed: knit, in which the loop is drawn from the back to the front of the cloth; purl, in which the loop is drawn from the front to the back of the cloth; miss, in which no loop is formed; and tuck, in which two courses on one wale are looped over a third. From these stitches all of the patterns of knit and double knit cloths can be made. The basic knit cloths are the jersey, made from only knit or only purl stitches; the rib, made from alternating wales of knit and purl stitches; and the purl, made from alternating courses of knit and purl stitches.

Simple warp knitting machines make an open loop, a closed loop, or no loop. Variations in the arrangement of these elements provide the different types of patterned fabrics. The simplest of the warp knits is the tricot, in which parallel wales of closed loops are interlocked in a zigzag pattern. More intricate forms of warp knitting may be used to produce fibers which range from simple, coarse netting to intricate, sophisticated lace.

Finishing. Newly constructed knit or woven fabric is not suitable for sale. It must pass through various finishing processes to make it suitable for its intended purpose. Finishing enhances the appearance of fabric and also adds to its serviceability. Finishes can be solely mechanical, solely chemical, or a combination of the two. Those finishes, such as scouring and bleaching, which simply prepare the fabric for further use are known as general finishes. Functional finishes, such as durable press treatments, impart special characteristics to the cloth. For discussions of important finishing operations *see* BLEACHING; DYEING; TEXTILE CHEMISTRY; TEXTILE PRINTING.

Performance. The choice of a fabric is determined by the requirements of appearance, comfort, ease of maintenance, durability, and cost. Appearance is affected by light reflection, pattern, and drape. Comfort is sensitive to the diffusion of air and moisture through the cloth, texture, elasticity, and stretch. Ease of maintenance includes soil and stain resistance and removal, wrinkle resistance and recovery, and dimensional stability. Durability is measured by strength, abrasion resistance, and the ability to withstand attack by sunlight, atmospheric fumes, insects, and microorganisms. These factors are all affected, to different degrees, by the choice of fiber, yarn construction, fabric type, and finish. In general, woven cloths provide greater durability than knit cloths, while knit fabrics tend to be more comfortable.

The density of a woven construction is measured by the yarn count (thread count), which is defined as the number of warp and the number of filling yarns in 1 in.^{2} (6.5 cm^2) of cloth. The density of a knit cloth is measured by the gauge, which is defined as the number of stitches in $1^1/_2$ in. (3.8 cm). Heavy, compact constructions are more durable than

lightweight, open fabrics. Usually the more intricate the pattern, the more costly the material. Ira Block

Bibliography. B. P. Corbman, *Textiles: Fiber to Fabric*, 6th ed., 1983; M. L. Joseph, *Introductory Textile Science*, 5th ed., 1993; M. L. Joseph and A. Gieseking-Williams, *Illustrated Guide to Textiles*, 4th ed., 1986; D. S. Lyle, *Modern Textiles*, 2d ed., 1982; B. F. Smith and I. Block, *Textiles in Perspective*, 1996; P. G. Tortora, *Understanding Textiles*, 6th ed., 2000.

Textile chemistry

The applied science of textile materials, consisting of the application of the principles of the many basic fields of chemistry to the understanding of textile materials and to their functional and esthetic modification into useful and desirable items. Textile chemistry includes the study of the chemical and physical properties of the various textile fibers, both natural and synthetic. Included also is the application of the principles of surface chemistry to the many cleaning operations and modifications of the textile material. Finally, there is the application of organic chemistry to the synthesis and formulation of the chemical products necessary for achieving these objectives.

Being primarily an applied form of chemistry, textile chemistry draws many of its concepts from the more basic and theoretical fields. At the same time, it is a highly specialized field which requires the blending of theory and knowledge with empirical information. The textile chemist must be able to relate knowledge of the organic structures of both the fibers and the chemicals being used to specific chemical, physical, and esthetic properties. This combination of the theoretical and the practical makes possible the development of the thousands of different textile chemicals necessary for the production of finished articles of commerce. *See* TEXTILE.

Fibers. The study of textile chemistry begins with the knowledge of the textile fibers themselves. These are normally divided into two groups: natural and manufactured, those produced from substances of natural or synthetic origin. The natural fibers include the protein fibers (silk and wool), as well as the cellulose and cellulose-related fibers (cotton, linen, ramie, and jute), plus certain other minor vegetable fibers. The manufactured fibers of natural origin are derived by chemical and physical modification of the naturally occurring fibrous materials, particularly cellulose; the important members of this group are rayon, acetate, and triacetate. The manufactured fibers of synthetic origin, also known as synthetic fibers, are produced by polymerization of synthetic chemicals; the starting point for these fibers is not a naturally occurring polymer. This group includes the polyamides (nylons), polyacrylics, and polyesters. The synthetic fibers share many common properties, but also differ greatly in certain ways because of their varying chemical natures. *See* MANUFACTURED FIBER; NATURAL FIBER; POLYACRYLONITRILE RESINS; POLYAMIDE RESINS; POLYESTER RESINS.

Chemicals. The enormous number of chemicals used in textile processing may be divided broadly into two categories: those intended to remain on the fiber, and those intended to wet or clean the fiber or otherwise function in some related operation. The former includes primarily dyes and finishes, which will be covered in the discussion of the appropriate operations. The latter group consists mainly of surface-active agents, commonly known as surfactants. These chemicals are characterized by the general structure R—Y, in which R, the major portion of the molecule, is oil-soluble, and Y is extremely water-soluble. The balance between these two sections determines the properties and uses of the surfactant. Typically, R is a long hydrocarbon chain. However, Y may be charged positively or negatively, or be uncharged. Some common Y groups are:

Anionic: carboxylate, $—COO^-$
sulfate, $—O—SO_3^-$
sulfonate, $—C—SO_3^-$
phosphate, $—PO_4^-$
Cationic: quaternary ammonium, $—N^+(CH_3)_3$
Nonionic: ethyoxylate, $(—CH_2—CH_2—O)_n$

Thus, probably the oldest surfactant is soap, sodium stearate, $CH_3(CH_2)_{15}CH_2—COO^-Na^+$. Choice of a particular combination of structures will depend on the specific use intended for the chemical, and will be illustrated in discussing the textile processes. *See* DETERGENT; QUATERNARY AMMONIUM SALTS; SOAP; SURFACTANT.

Processes. All textile fibers go through a variety of physical and chemical processes designed to produce useful articles of commerce. Although the order may vary in certain instances, there is a general sequence of physical and chemical operations, beginning with the fiber itself and ending with the finished fabric ready for use by its ultimate consumer. The common order of operations is: aligning the fiber and spinning the yarn; weaving or knitting the yarn into the fabric; preparation; fabric coloring or printing; and fabric finishing.

The major physical operations are, of course, spinning the fibers into yarns, followed by weaving or knitting these yarns into fabrics. Although a certain percentage of textile materials reaches the ultimate consumer in the form of yarn, unquestionably the major usage is in woven, knitted, or nonwoven fabrics. Spinning, weaving, and knitting are almost completely physical operations, although chemical products such as lubricants and sizings composed of starch or other polymers are used to protect the fibers and yarns from damage. These materials must eventually be removed, and thus knowledge of their composition is important to the textile chemist. In addition to these physical operations, textile fibers, either in that form or in the form of yarn or fabric, are usually subjected to a variety of chemical operations; the study of these operations constitutes a major portion of the field of textile chemistry.

Preparation. Preparation is a term applied to a group of essentially wet chemical processes having

as their object the removal of all foreign matter from the fabric. This results in a clean, absorbent substrate, ready for the subsequent coloring and finishing operations.

Since the object of preparation is the removal of foreign matter from the fiber, and since the nature of these impurities is dependent on the fiber itself, the operations constituting preparation depend primarily on the fibers being handled. Synthetic fibers contain little or no natural impurities, so that the only materials that normally must be removed are the oils and lubricants or water-soluble sizes needed to facilitate earlier processing. This is generally accomplished by washing with water and a mild detergent capable of emulsifying the oils and waxes. A typical material for this use would be a nonionic with a relatively low degree of ethoxylation. On the other hand, natural fibers contain relatively high amounts of natural impurities, and in addition frequently are sized with materials presenting difficulties in removal. In the case of cotton, prolonged hot treatment with alkali, usually sodium hydroxide, and strong detergent is necessary to break down and remove the naturally occurring impurities. Sizing, which is usually based on starch, must be broken down by treatment with enzymes or peroxide, substances which attack the starch molecule and convert it to more soluble materials. After these treatments, strenuous washing is necessary to remove the residue from the cotton fiber. Sulfates and sulfonates are frequently used in these operations, but the phosphate surfactants are generally preferred, because of their superior functioning in hot alkaline solutions, and their ability to disperse and suspend solid soil. *See* STARCH; SULFONATION AND SULFATION.

Special scours are necessary for cleaning such materials as wool and silk: these are actually processed in relatively small volume in the United States. The protein fibers are very sensitive to alkali and strong detergents; they are usually washed with mild soap or sulfated alcohols. Agents having a protective action on the fiber, such as highly sulfated castor oil, are frequently included. Wool may be carbonized, that is, treated with strong sulfuric acid to destroy vegetable matter adhering to the fiber.

After other impurities are removed from the fiber, it is usually desirable to remove any coloring material. This process, known as bleaching, is necessary to some degree with certain synthetics; it is essential with the natural fibers if white or light-colored fabrics are to be prepared. Several chemicals destroy colored impurities. Sodium hypochlorite and sodium chlorite were used in the past, and are still used to some extent. However, by far the major bleaching agent in use is hydrogen peroxide, which is efficient in color removal, while still being considered relatively controllable and safe for use. *See* BLEACHING; HYDROGEN PEROXIDE.

The preparation of blends of fibers frequently presents some difficulty, but in general it may be stated that the blend is prepared in a manner similar to that component which in itself is most difficult to clean. Thus, a polyester-cotton blend would be prepared in a manner relatively similar to a 100% cotton fabric.

Mercerization. Mercerization is a special process applied only to cotton. The fabric or yarn is treated with a strong sodium hydroxide solution, usually about 20% strength, while being held under tension. This process causes chemical and physical changes within the fiber itself, resulting in a substantial increase in luster and smoothness of the fabric, plus important improvements in dye affinity, stabilization, tensile strength, and chemical reactivity. It is necessary in this process that the caustic solution penetrate quickly and uniformly, and equally important that after reaction has taken place, the caustic be removed rapidly and thoroughly. Special surface-active agents have been developed to offer maximum effectiveness in highly concentrated sodium hydroxide solutions. Originally these were based on cresylic acid, the mixed isomer of cresol, but increasingly strict pollution controls have led to the development of noncresylic penetrants. These are typically sulfate esters of relatively short-chain alcohols, such as hexyl or octyl: $CH_3(CH_2)_nCH_2—O—SO_3Na$ ($n = 4$ or 6).

Coloring. Although many textiles reach the consumer in their natural color or as a bleached white, most textiles are colored in one way or another. Coloring may be accomplished either by dyeing or printing, and the coloring materials may be either dyes or pigments.

Dyeing. In its simplest form, dyeing consists of immersing the fabric or yarn in a solution of a dyestuff. This dyestuff in one way or another is attracted to the fiber, and travels from the solution into the interior of the fiber. By contrast, a pigment is soluble neither in the solution nor in the fiber itself; it is simply deposited on the surface of the fabric, where it is held, or glued, by an insoluble binder. There are some dyeing procedures which appear to omit some of the normal steps, and there are also certain procedures which partly involve working with the dye in pigment form. However, the important distinction is that a dye exists in solution or ultrafine dispersion in the solvent, usually water, and is deposited or dissolved within the interior of the fiber; a pigment is never soluble during any stage of the process, and is deposited and bound to the surface of the fiber only. *See* DYEING.

Printing. Dyeing essentially consists of immersing the entire fabric in the solution, so that the whole fabric becomes colored. Printing may be considered localized dyeing. In printing, a thickened solution of dyestuff or pigment is used. This thickened solution, or paste, is applied to specific areas of the fabric by means such as engraved rollers or partially porous screens. Application of steam or heat then causes the dyestuff to migrate from the dried paste into the interior of the fiber, but only in those specific areas where it has been originally applied. Printing makes possible extremely complicated, multicolored designs, without resort to difficult and unusual weaving or knitting procedures. *See* TEXTILE PRINTING.

Dyestuffs. The choice of dyestuff depends on the physical and chemical nature of the fiber. Fibers

such as wool and nylon contain positively charged NH groups, and thus the dyes used contain negative groups such as the sulfonate group (SO_3). On the other hand, acrylic fibers have negative sulfonate groups present in the form of copolymers, and are thus processed with dyes carrying a positive charge. Polyester offers no such ionic sites, and thus dyes used for polyester contain no charge, but are in the form of ultrafine particles, which form solid solutions in the fiber. Cotton contains only hydroxyl groups as dye sites, and thus the dyes used may be fiber-reactive, forming covalent bonds, or direct, forming hydrogen bonds with these groups. Other cotton dyes, such as vats and naphthols, are instead converted to insoluble derivatives in the fiber. *See* DYE.

Other dyeing chemicals. Many textile chemicals are used in the dyeing process, besides the actual dyestuffs themselves. It is necessary that the dyestuff penetrate as fully as possible into the individual fibers, and be as uniformly distributed as possible. Certain chemicals aid the penetration of the dyestuff; others retard the penetration, and are used to permit the color to build up at a gradual rate, thus ensuring uniformity. Another type of textile chemical used in dyeing is the migrating agent, which aids the individual dyestuff molecules in moving from one site to another within the fiber, thus promoting uniformity in a different way.

Penetrants may be surfactants of appropriate type or, in the case of polyester or certain other fibers, may be swelling agents. Retarders and migrating agents may be of two types. Certain chemicals have structures similar to, but simpler than, the dyestuffs. These smaller molecules occupy the available dye sites before the dyestuff molecule, but eventually are replaced by it.

The other method of retardation is the use of either a material of reverse charge or a nonionic, or both, to form an unstable complex with the dyestuff molecule.

Finally, certain dyestuffs are after treated with chemicals designed to increase their resistance to being washed out of the fiber. The textile chemist must therefore have a knowledge of the fiber, the dyestuff, and the dyeing auxiliaries. Many of the same factors apply also to printing. In addition, a knowledge of the different types of thickening agents available is also essential. These agents may be derived from natural sources, such as starch or gum, or may be a synthetic material such as a polyacrylate. Proper choice of thickener is absolutely essential for effective printing.

Finishing. Finishing includes a group of mechanical and chemical operations which give the fabric its ultimate feel and performance characteristics. The fabric may be compressed, to minimize shrinkage, or the surface may be polished or roughened; these are all essentially mechanical operations. On the other hand, many desirable characteristics may be imparted to the fabric through the application of various chemical agents at this point.

Chemical operations. Softeners are used to give a desirable hand or feel to the fabric. These chemicals are generally long fatty chains, with solubilizing groups which may be cationic, nonionic, or occasionally anionic in character. They are similar to many of the surfactants mentioned earlier, but are constructed so as to contain a higher proportion of fatty material in the molecule. Thus poly(ethylene glycol) monostearate is a typical nonionic softener. Similarly, dimethyl tallow ammonium chloride is widely used as a cationic softener.

Other softeners are emulsions formed of oil, wax, or polyethylene, using nonionic or cationic surfactants as emulsifiers. By proper selection of softener, it is possible to make a fabric feel slick, or limp, or lofty. Conversely, certain types of polymeric material such as poly(vinyl acetate) or polymerized urea-formaldehyde resins are used to impart a stiff or crisp hand to a fabric. *See* POLYVINYL RESINS; UREA-FORMALDEHYDE RESINS.

It is in finishing that the so-called proof finishes are applied, including fire-retardant and water-repellent finishes. A fire-retardant finish is a chemical or mixture containing a high proportion of phosophorus, nitrogen, chlorine, antimony, or bromine. For example, the combination of tetrakishydroxymethyl phosphonium chloride (THPC) with trimethylol melamine (TMM) is effective in producing fire-retardant fabrics. A truly waterproof fabric may be made by coating with rubber or vinyl, but water-repellent fabrics are produced by treating with hydrophobic materials such as waxes, silicones, or metallic soaps.

In the case of cellulosic fabrics, various cross-linking agents are applied to react with the cellulose chains and impart resistance to shrinking and creasing. The most widely used cross-linking agents are *N*,*N*-dimethylol carbamate and dimethylol dihydroxyl ethylene urea. These materials form covalent links between adjacent cellulose chains.

Other types of highly specialized treatments, such as antistatic, antibacterial, or soil-repellent finishes, may be applied to fit the fabric to a particular use.

Mechanical operations. Finishes are usually applied by an operation known as padding, which consists of drawing fabric through a water solution of the chemical being applied, followed by passage through a set of squeeze rolls which removes the excess liquid. A wetting agent, or penetrant, is frequently included in this padding solution to aid penetration into the fabrics. The amount of chemical deposited within the fabric is thus determined first by the concentration of the chemical in the solution, and second by the amount of pressure applied by the squeeze rolls, and thus the amount of solution allowed to remain in the goods. Typically, a fabric may still retain 50–100% of its own weight in solution, and a considerable amount of water must remain to be evaporated.

Evaporation of unneeded water is an energy-intensive, and thus expensive, procedure, and attempts have been made to devise means of reducing the amount of water to be evaporated. In one approach, the concentration of chemical in the solution is greatly increased, and this solution is

sprayed on the fabric. Unfortunately, serious difficulties have been encountered in maintaining uniform distribution. Another procedure offering considerable promise is foam finishing; in this technique a concentrated solution containing the chemical to be applied, as well as a foaming agent, is converted into a foam of specific density and wetness by mechanical beating and injection of air. In effect, air is substituted for the major part of the water as a diluent for the chemicals. *See* FOAM.

Techniques involving use of organic solvents in place of water have been demonstrated, as well as other more novel finishing techniques. It is safe to assume that one area of textile chemical research to become increasingly important will be development of low-energy processing techniques.

Garment processing. In traditional processing, the operations of preparing, dyeing, and finishing are applied to the fabric; this fabric is later cut and sewed into garments. In garment processing, the fabric may be converted into garments at any time; the remaining operations are performed on the garments in large laundry washing machines.

Dyeing by this method is not nearly as uniform as a classical yarn or fabric dyeing. However, casual styles have been developed and accepted for these unevenly dyed garments. The main advantage is the reduction of inventories; garments are colored in response to actual sales.

A variety of this form of processing is the creation of partially worn-out garments. As a full load of garments is tumbled in a washing machine, chemicals such as potassium permanganate or sodium hypochlorite are used to destroy color irregularly. Furthermore, rocks, pumice stones, or cellulose-attacking enzymes are introduced to weaken and fray the garment, creating an overall worn and used appearance. In fact, the garment actually is somewhat weakened so that the result is poorer quality.

David H. Abrahams

Bibliography. S. V. Kulkarni, *Textile Dyeing Operations: Chemistry, Equipment, Procedures, and Environmental Aspects*, 1986; M. Lewin, *Handbook of Fiber Chemistry*, 2d ed., 1998; E. R. Trotman, *Dyeing and Chemical Technology of Textile Fibers*, 6th ed., 1991; T. L. Vigo, *Textile Processing and Properties*, 1994.

Textile microbiology

That branch of industrial microbiology concerned with textile materials. Most of the microorganisms on textiles—the fungi, actinomycetes, and bacteria—originate from air, soil, and water. Some of the microorganisms are harmful to either the fibers or the consumer. They may decompose the cellulose or protein in the fiber or affect the consumer's health. Since the minimum moisture content for microorganism development is 7%, dry storage is an effective prevention measure. Some of the microorganisms are useful, for example in the retting process, in which fibers are liberated from the stalks of such fiber plants as flax, hemp, and jute.

Cotton. Microbial attack of cotton can occur at any time from the opening of the boll. The fiber or fabric may be degraded, resulting in a decrease in length of the fiber and strength of the cloth, uneven dyeability, darkening, or formation of colored spots. A large variety of fungi can be active in this process (**Fig. 1**). Representatives of the genera *Chaetomium* and *Myrothecium* have the highest cellulose-decomposing activity; representatives of *Alternaria*, *Cladosporium*, *Fusarium*, and *Diplodia* are active in field cotton; *Aspergillus*, *Penicillium*, and *Stachybotrys* are active in stored cotton and fabrics (mildewing). *See* ASCOMYCOTA.

The number of contaminating bacteria varies from about 1 million to 100 million per gram of raw cotton. Among these are the cellulose-decomposing bacteria (**Fig. 2**). They all belong to the Myxobacteriales and have been found on field cotton where they probably play an important role in degradation. *Aerobacter cloacae* has also been found on raw cotton. This organism is of interest because it causes a respiratory disease.

Cotton is examined for microbial damage by a series of tests. In the Congo red test, microscopic examination is made of the cotton fibers after a dye, Congo red, has been applied. The undamaged cotton fibers appear pink, and the damaged cotton fibers appear red and may be cracked. In the swelling test,

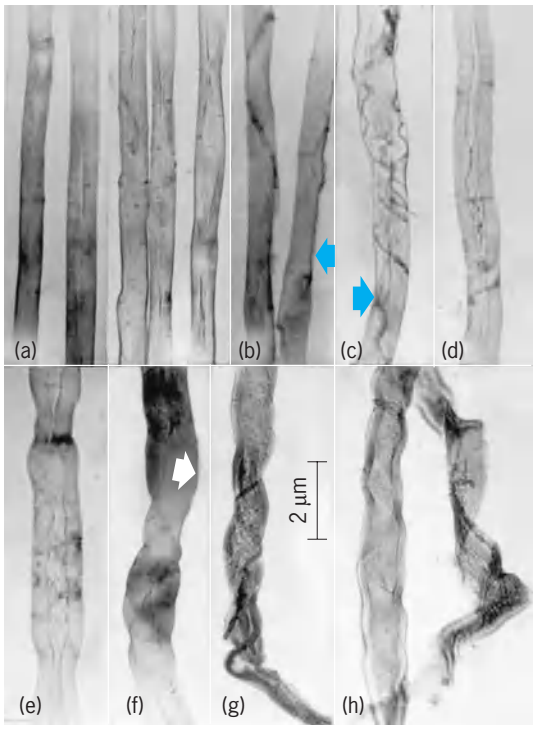

Fig. 1. Stages of deterioration of cotton fiber under influence of fungi. (*a*) Normal fibers. (*b*, *c*) Fungal hyphae growing on the outside of the fiber (shown by arrows). (*d*) Fungal hyphae growing in the lumen of the fiber. (*e*, *f*) Fibers showing excessive and irregular swelling. (*e*) Cavitomic stage. (*f*) Mildewed fiber with cuticle damaged and loosened (shown by arrow). (*g*, *h*) Final stages of deterioration. Note fungal hyphae in spiral around fiber *g*. (*From A. N. J. Heyn, Textile Ind., vol. 120, 1956*)

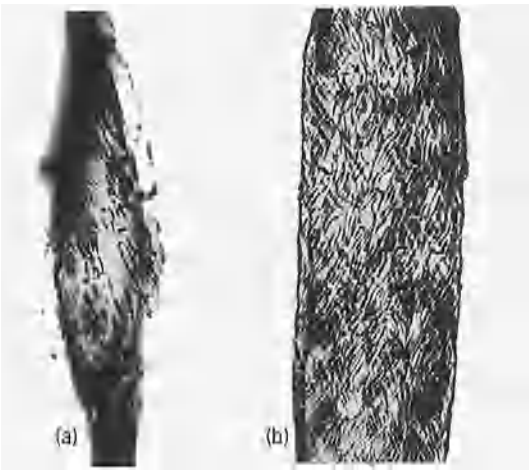

Fig. 2. Two growth stages of cellulose bacteria (*Sporocytophaga myxococcoides*) on cotton fiber. (*a*) Initial stage. (*b*) Final stage. (*From A. N. J. Heyn, Textile Res. J., 27:591–602, 1957*)

cotton attacked by microbes has a higher degree of swelling than normal cotton. Other testing methods determine the pH of an aqueous extract of the cotton and the fluorescence of the cotton fibers in ultraviolet light.

Wool. The number of bacteria and molds on raw wool has been reported as 1.2 million per gram and it may increase to 65 million in wet, scoured wool. *Achromobacter liquefaciens*, a nonspore-former, and various sporeforming bacteria, such as *Bacillus subtilis*, cause the greatest damage to wool. In the degradation process the fiber scales are broken down and hydrolysis of the intercellular substances of the cortex takes place. If woolen fabrics are heated to 140°F (60°C), putrefaction caused by growth of *B. subtilis* and *Agarbacterium mesentericum* may develop and produce an uneven dyeability.

Fungi may develop on stored wool under humid conditions and when degradation products of fibrous proteins are present. Microbial damage of wool can be tested, for example, by the brilliant blue fluorescence of the degreased sample under ultraviolet light, and by the absence of partial decomposition of the scales.

Retting. This is a microbial process used for liberating fiber bundles from the stalks of fiber plants. In principle, retting consists of a breaking down of pectic substances between the cell walls (middle lamellae) of the individual cells of the tissue surrounding the bundles. As a result, the bundles become separated from the surrounding tissue and can then easily be extracted mechanically.

In water retting, the stalks are immersed in cold or warm, slowly renewed water, for from 4 days to several weeks. The water may be from rivers or constructed containers. The active organism is *Clostridium felsineum* and related types, which break down the pectin to a mixture of organic acids (chiefly acetic and butyric), alcohols (butanol, ethanol, and methanol), carbon dioxide (CO_2), and hydrogen (H_2). Its optimum activity at 91.4°F (33°C) is the basis for the warm-water retting process. Related species may be active in cold-water retting. In dew retting the stems are spread out in moist meadows; here the pectin decomposition is accomplished by molds and aerobic bacteria with the formation of CO_2 and H_2O. A retting process has been found to cause the liberation of fibers from the fruit husk of the coconut during soaking. *See* INDUSTRIAL MICROBIOLOGY; TEXTILE CHEMISTRY.

A. N. J. Heyn

Textile printing

The localized application of color on fabrics. In printing textiles, a thick paste of dye or pigment is applied to the fabric by appropriate mechanical means to form a design. The color is then fixed or transferred from the paste to the fiber itself, maintaining the sharpness and integrity of the design. In a multicolor design, each color must be applied separately and in proper position relative to all other colors. Printing is one of the most complex of all textile operations, because of the number of variables and the need for a high degree of precision, particularly since there is no way to correct a bad print. *See* DYE; DYEING.

Methods. A design may be applied in three major ways: raising the design in relief on a flat surface (block printing); cutting the design below a flat surface (intaglio or engraved printing); and cutting the design through a flat metal or paper sheet (stencil or screen printing). All three methods have been used for hand printing and reciprocating printing machines. In addition, these methods have been converted into rotary action by replacing blocks or plates with cylinders.

Block printing still exists to a very limited degree, and cylinders with a raised design are still occasionally used in wallpaper printing. However, this method is rarely found in textile applications, except in printing carpets.

The engraved plate has been replaced by the engraved cylinder, or roller. The design is etched into copper rollers, which are then generally chromium-plated. This method is still in wide use, particularly when large quantities of a relatively simple design are required.

The stencil was first developed into the flatbed screen, in which a silk gauze is stretched over a frame. Parts of the screen are blocked out, where no color is to penetrate, and paste is forced through the open parts of the screen, either by hand or mechanically, by a moving squeegee. The silk gauze was later replaced by a fine metal mesh that permitted this screen to be rolled into a cylinder, so that the mechanical action could be rotary instead of reciprocating. This resulted in rotary screen printing, actually a variation of flatbed printing.

Another method of printing utilizes individual computer-controlled nozzles for each color. The nozzles are used to paint a design on the fabric.

Paste and thickener. Each printing method requires a paste with special characteristics, frequently referred to as flow characteristics. For instance, a roller

printing operation requires a paste that tends to flow from the engraved portions of the roll onto the fabric; screen printing operations require a paste that flows only under the direct pressure of the squeegee. For this reason, the choice of thickener is dependent not only on the type of dye-stuff, but on the type of printing machine on which the printing is to be done, and frequently also on the type of fixation to be utilized.

Most natural thickening agents are based on combinations of starch and gums such as guar, locust, or alginate, chemically modified so as to be soluble in cold water. The synthetic thickening agents used are generally extremely high-molecular-weight polymers capable of developing a very high viscosity at a relatively low concentration.

In a few instances, thickening is accomplished by forming a thick emulsion, either oil in water, or water in oil. This last method is falling into disfavor because of the relatively large amount of solvent evaporated into the air during drying.

Sequence of operations. The first step is the preparation of print paste, which is made by dissolving the dyestuff and combining it with a solution of the appropriate thickening agent. Humectants and other materials necessary to aid in dye-stuff fixation are added at this point, as well as other auxiliaries, such as defoamers, which may be necessary to aid in printing. The fabric is then printed by any of the standard methods. The goods are then dried in order to retain a sharp printed mark and to facilitate handling between printing and subsequent processing operations.

The next operation, steaming, may be likened to a dyeing operation. Before steaming, the bulk of the dyestuff is held in a dried film of thickening agent. During the steaming operation, the printed areas absorb moisture and form a very concentrated dyebath, from which dyeing of the fiber takes place. The thickening agent prevents the dyestuff from spreading outside the area originally printed, because the printed areas act as a concentrated dyebath that exists more in the form of a gel than a solution and restricts any tendency to bleed. Excessive moisture can cause bleeding, and insufficient moisture can prevent proper dyestuff fixation. Steaming is generally done with atmospheric steam, although certain types of dyestuffs, such as disperse dyes, can be fixed by using superheated steam or even dry heat. In a few instances, acetic or formic acid is added to the steam to provide the acid atmosphere necessary to fix certain classes of dyes.

Flash aging is a special fixation technique used for vat dyes. The dyes are printed in the insoluble oxidized state by using a thickener which is very insoluble in alkali. The dried print is run through a bath containing alkali and reducing agent, and then directly into a steamer, where reduction and color transfer take place.

Printed goods are generally washed thoroughly to remove thickening agent, chemicals, and unfixed dyestuff. This washing must be carefully done to prevent staining of the uncolored portions of the fabric. Drying of the washed goods is the final operation of printing.

Special types of printing. There are several special printing processes.

Pigment printing. This process is actually one of the most widely used processes in textile printing. Instead of a dyestuff which can actually penetrate the fiber, an insoluble pigment is used as a coloring agent. A resinous binder added to the print paste serves to hold the pigment in place on the finished goods. The print is fixed by a simple dry heat application, and normally no washing is necessary after drying. This is generally the lowest-cost printing process, but the color is frequently less permanent, and the fabric generally tends to have a harsh feel.

Discharge printing. In this process, a color-destroying agent such as sodium hydrosulfite is printed onto dyed fabric, resulting in a white design appearing on a colored background.

Resist printing. This process is somewhat the reverse of discharge printing. The goods are printed with a material which will resist dyestuff. After printing, the goods are immersed in a dyebath, leaving the printed areas undyed. *See* TEXTILE CHEMISTRY.

David H. Abrahams

Bibliography. J. Fish, *Designing And Printing Textiles*, 2005; L. W. C. Miles, *Textile Printing*, 2d rev. ed., 2003; T. L. Vigo, *Textile Processing and Properties*, 1994; D. R. Waring and G. Hallas (eds.), *The Chemistry and Application of Dyes*, 1989; G. Wasike Namwamba, *Digital Textile Printing*, 2005.

Thaliacea

A small class of pelagic Tunicata especially abundant in warmer seas. This class of animals contains three orders: the Salpida, Doliolida, and Pyrosomida. Oral and atrial apertures occur at opposite ends of the body. Members of the orders Salpida and Doliolida are transparent forms, partly or wholly ringed by muscular bands (**Fig. 1**). The contractions of these

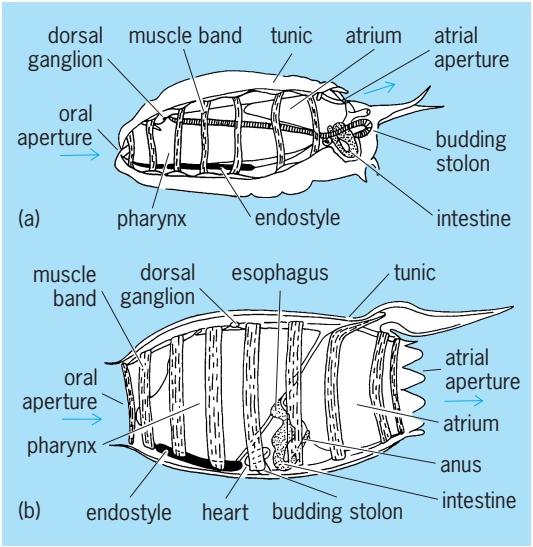

Fig. 1. Representatives of two orders of the class Thaliacea. (*a*) Salp, *Thalia democratica* (Salpida), solitary asexual form. (*b*) Doliolid, *Doliolum* (Doliolida), solitary asexual form. Arrows show direction of current.

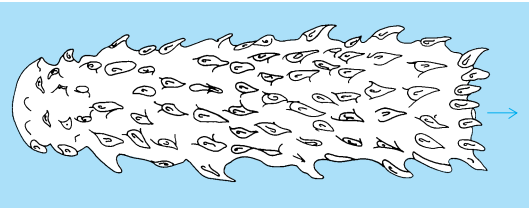

Fig. 2. Colony of *Pyrosoma atlanticum* (Pyrosomida).

bands produce currents used in propulsion, feeding, and respiration. The life cycle involves an alternation between solitary, asexual oozooids, which reproduce by budding from a complex stolon, and colonial, sexually reproducing blastozooids. The order Pyrosomida includes species which form tubular swimming colonies (**Fig. 2**) and which are often highly luminescent. *Salpa*, *Doliolum*, and *Pyrosoma* are familiar genera. *See* BIOLUMINESCENCE; TUNICATA (UROCHORDATA). Donald P. Abbott

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; H. Thompson, *Pelagic Tunicates of Australia*, Commonwealth Council for Scientific and Industrial Research, Australia, 1948.

Thallium

A chemical element, TI, atomic number 81, relative atomic weight of 204.37. The valence electron notation corresponding to its ground state term is $6s^2 6p^1$, which accounts for the maximum oxidation state of III in its compounds. Compounds of oxidation state I and apparent oxidation state II are also known. *See* PERIODIC TABLE.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Thallium occurs in the Earth's crust to the extent of 0.00006%, mainly as a minor constituent in iron, copper, sulfide, and selenide ores. Minerals of thallium are considered rare. Thallium compounds are extremely toxic to humans and other forms of life.

The insolubility of thallium(I) chloride, bromide, and iodide permits their preparation by direct precipitation from aqueous solution; the fluoride, on the other hand, is water-soluble. Thallium(I) chloride resembles silver chloride in its photosensitivity.

Thallium(I) oxide is a black powder which reacts with water to give a solution from which yellow thallium hydroxide can be crystallized. The hydroxide is a strong base and will take up carbon dioxide from the atmosphere.

Thallium also forms organometallic compounds of the following general classes, R_3TI, R_2TIX, and $RTIX_2$, where R may be an alkyl or aryl group and X a halogen. *See* ORGANOMETALLIC COMPOUND. Edwin M. Larsen

Bibliography. P. W. Atkins et al., *Inorganic Chemistry*, 4th ed., 2006; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004.

Thallobionta

One of the two commonly recognized subkingdoms of plants. In contrast to the more closely knit subkingdom Embryobionta, the Thallobionta (also called Thallophyta) are diverse in pigmentation, food reserves, cell-wall structure, and flagellar structure. They still form a natural group, however, in that they are all probably derived from ancestors which would be referred to the Thallobionta, without the intervention of any ancestors which would have to be referred to other groups. The Thallobionta are here considered to include seven divisions, the Rhodophyta, Chlorophyta, Euglenophyta, Pyrrophyta, Chrysophyta, Phaeophyta, and Fungi. All these groups have both modern and fossil representatives.

The Thallobionta are united more by the absence of certain specialized tissues or organs than by positive resemblances. They do not have the multicellular sex organs commonly found in most divisions of Embryobionta. Gametes, when produced at all, are borne in unicellular gametangia which may or may not be differentiated into oogonia (which bear eggs) and antheridia (which bear sperms). Many of the Thallobionta are unicellular, and those which are multicellular seldom have much differentiation of tissues. None of them has tissues comparable to the xylem (a water-conducting tissue) found in most divisions of the Embryobionta, and only some of the brown algae (division Phaeophyta) have tissues comparable to the phloem (a food-conducting tissue) found in most divisions of the Embryobionta.

A large proportion of the Thallobionta are aquatic, and those which grow on dry land seldom reach appreciable size. The most obvious exceptions to this generality are provided by some of the fungi, such as mushrooms and puffballs, and even here the visible fruiting body is merely the aerial reproductive structure of a plant which has a much more diffuse mycelium (mass of filamentous vegetative strands) permeating the substrate. The Thallobionta thus consist of all those plants which have not developed the special features that mark the progressive adaptation of the Embryobionta to life on dry land. *See* CHLOROPHYCOTA; CHRYSOPHYCEAE; EMBRYOBIONTA; EUGLENOPHYCEAE; FUNGI; PHAEOPHYCEAE; PLANT KINGDOM; RHODOPHYCEAE. Arthur Cronquist

Bibliography. A. Cronquist, *Basic Botany*, 1973.

Theales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Dilleniidae of the class Magnoliopsida (dicotyledons). The order consists of 18 families and nearly 3500 species. The largest families are the Clusiaceae, sometimes called Guttiferae (about 1200 species), Theaceae (about 600 species), Dipterocarpaceae (about 600 species), and Ochnaceae (about 400 species). They are mostly woody plants, less often herbaceous, with simple or occasionally compound leaves. The perianth and stamens of the flowers are attached directly to the receptacle; the calyx is arranged in a tight spiral; and the petals are usually separate from each other (see **illus.**). The stamens are numerous and initiated in centrifugal sequence or, less often, are few and cyclic; the pollen is nearly always binucleate. The gynoecium consists of two to many, more or less united carpels. The tea plant (*Thea sinensis*), *Camellia* (in the Theaceae), and St.-John's-wort (*Hypericum*, in the Clusiaceae) are familiar members of the Theales. *See* DILLENIIDAE; FLOWER; MAGNOLIOPHYTA; PLANT KINGDOM; TEA. Arthur Cronquist; T. M. Barkley

Franklin tree (*Franklinia alatamaha*) flowers, a characteristic member of the family Theaceae in the order Theales, named in honor of Benjamin Franklin. A rare native of the southeastern United States, it has not been seen in the wild since 1790, but it is frequently planted in the United States and Europe. (*Photograph by A. W. Ambler, from National Audubon Society*)

Thecanephria

An order of Pogonophora, a group of elongate, tentaculated, tube-dwelling, sedentary, nonparasitic marine worms lacking a digestive system. In this order the "coelomic" space in the anterior tentacular region is horseshoe-shaped, and the excretory (osmoregulatory) portion of its ducts come close together medially near the median, adneural blood vessel. The spermatophores are generally broad and leaf-shaped. The tube is often rigid anteriorly. The species in this order are usually large (0.04–0.12 in. or 1–3 mm in diameter) and multitentaculate.

The order includes four families: Polybrachiidae (with seven genera), Sclerolinidae (one genus), Lamellisabellidae (two genera), and Spirobrachiidae (one genus). *See* POGONOPHORA. Edward B. Cutler

Thelodontida

Sometimes called Coelolepida, this is an extinct group of Agnatha or jawless vertebrates known from the Lower Silurian to Middle Devonian of Europe, Asia, Australia, and North America. Since they had no hard skeleton, their structure and relationships are poorly known. Well-preserved specimens of *Logania* (see **illus.**) suggest a relationship to Heterostraci in their widely spaced, lateral eyes, nearly terminal mouth, pectoral flaps at the posterior end of the head, and hypocercal tail with a downwardly directed main lobe. *Phlebolepis* resembles Anaspida in general form but lacks the dorsal nostril of that group. *See* JAWLESS VERTEBRATES; HETEROSTRACI; OSTRACODERM; PTERASPIDOMORPHA. Robert H. Denison

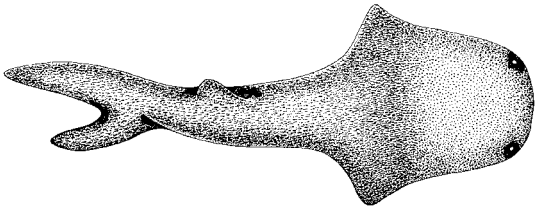

***Logania scotica*, dorsal, showing the eyes and dorsal fin. The tail, in profile, shows the caudal fin. (*After R. H. Traquair, Geological survey in the Upper Silurian rocks of Scotland, Trans. Roy. Soc. Edinburgh, vol. 40, 1905*)**

Theorem

A proposition arrived at by the methods of logical deduction from a set of basic postulates or axioms accepted as primitive and therefore not subject to deductive proof. So long as a theorem is part of a purely formal system, it is not meaningful to speak about the "truth" of a theorem but only about its "correctness." It becomes true when it, or its consequences, can be shown to be in accord with observable facts. The classic example of a system of theorems is afforded by Euclid's system of geometry, which is now recognized to have only a purely formal character, although it was formerly considered to be meaningful to ask whether Euclid's geometry was "true." *See* LOGIC. Percy W. Bridgman; Henry Margenau

Theoretical ecology

The use of verbal models, analytical models, or simulation models to explain patterns, suggest experiments, or make predictions in ecology. Because ecological systems are idiosyncratic, extremely complex, and variable, ecological theory faces special

challenges. Unlike physics or genetics, which use fundamental laws of gravity or of inheritance, ecology has no widely accepted first-principle laws. Instead, different theories must be invoked for different questions, and the theoretical approaches are enormously varied. Indeed, a central problem in ecological theory is determining what type of model to use and what to leave out of a model. The traditional approaches have relied on analytical models based on differential or difference equations; but recently the use of computer simulation has greatly increased with advances in computational power and ease of use. *See* ECOLOGICAL MODELING; ECOLOGY; ECOLOGY, APPLIED; SIMULATION.

Levels of ecological theory. The nature of ecological theory varies depending on the level of ecological organization on which the theory focuses. The primary levels of ecological organization are (1) physiological and biomechanical, (2) evolutionary (especially applied to behavior), (3) population, and (4) community.

Physiological and biomechanical theory. At the physiological and biomechanical level, the goals of ecological theory are to understand why particular structures are present and how they work. The approaches of fluid dynamics and even civil engineering have been applied to understanding the structures of organisms, ranging from structures that allow marine organisms to feed, to physical constraints on the stems of plants. A simple example of this type of analysis is the demonstration that there are lower limits to the body size of a mammal. The ratio of surface area to volume increases as an animal gets smaller. If the animal is too small, it cannot maintain a constant temperature. This argument can be made quantitative, and the predictions of the smallest body size agree well with the size distribution of existing mammal species.

Behavioral and evolutionary theory. At the behavioral level, the goals of ecological theory are to explain and predict the different choices that individual organisms make. Underlying much of this theory is an assumption of optimality: the theories assume that evolution produces an optimal behavior, and they attempt to determine the characteristics of the optimal behavior so it can be compared with observed behavior. One area with well-developed theory is foraging behavior (where and how animals choose to feed). One basic formulation assumes that resources (such as seeds for birds or flowers for pollinators) are present in different patches, and it seeks to understand how an animal should divide its searching time among patches. In the simplest case, an animal should leave any patch when the rate of return (rate of food consumption) in the patch is less than the mean rate of return over all patches. This prediction has been upheld by observations of many bird species. This theory is limited, however, because it ignores any interactions among the foraging individuals. If the fitness consequences of a behavioral action depend on which conspecifics (individuals or populations of a given species) an animal interacts with, a more sophisticated approach is needed. *See* BEHAVIORAL ECOLOGY.

In the 1970s, John Maynard Smith introduced the use of game theory to understand the evolution of behaviors that are apparently not optimal for an individual but may instead be better for a group—say, ritualized fighting. For example, in some African ungulates, one animal will back down when a pair engages in a fight. At first sight, this appears to contradict the principle of optimality at the level of the individual—the animal that gives up (a "dove") would seem to clearly have some chance of winning the contest. Conversely, an individual that fights until it dies or wins (a "hawk") is clearly worse off than a dove on the occasions that it encounters another hawk and dies in battle. Thus it can be shown that in a population of all hawks, doves are favored because they do not suffer from having to fight and win. Conversely, in a population of all doves, hawks are favored since they win every contest. Thus, evolution should lead to a population in which some individuals are hawks and some are doves. Another possibility is that individuals are sometimes hawks and doves, as the model does not distinguish who plays each strategy, but focuses only on the fraction of time the strategy is used. Note that the optimal situation for the population as a whole (which would be all doves and thus no battle casualties) is not the solution that is predicted to be reached by evolution (which should always include some hawks, since hawks are favored when they are rare). *See* OPTIMIZATION.

Determining the precise outcome in this example involves the use of concepts from game theory (originally developed by J. von Neumann and J. Nash to look at questions from economics) to find what has been called the evolutionary stable strategy. In this example, the evolutionary stable strategy is the proportion of hawks and doves at which the fitnesses of both are equal and any increase in the frequency of hawks reduces the fitness of hawks and any increase in the frequency of doves reduces the fitness of doves. Expressing the result this way clearly shows that frequency dependence drives the outcome. This approach is powerful and has been used to understand a wide range of ecologically important behaviors, from seed dispersal to the production of alarm calls. *See* GAME THEORY.

Population theory for single species. The population level has the longest history of ecological theory and perhaps the broadest application. The simplest models of single-species populations ignore differences among individuals and assume that the birth rates and death rates are proportional to the number of individuals in the population. If this is the case, the rate of growth is exponential (like compound interest), a result that goes back at least as far as Malthus's work in the 1700s. As Malthus recognized, this result produces a dilemma: exponential growth cannot continue unabated. Thus, one of the central goals of population ecology theory is to determine the forces and ecological factors that prevent exponential growth and to understand the consequences for the dynamics of ecological populations. *See* ECOLOGICAL METHODS; POPULATION ECOLOGY.

Density dependence. For a single species, intraspecific effects on the rate of growth are called density dependence. One primary density-dependent factor is food supply limitation; as the density of a species goes up, it increases competition for limited resources, causing the birth rate or survival rate to go down. Other density-dependent factors include the number of territories or nesting sites, incidence of cannibalism, and availability of space, light, or nutrients for plants. If density dependence acts without delay, the populations will tend toward an equilibrium, explaining the persistence of species at a stable level. This model is usually expressed in mathematical terms with a differential equation called the logistic equation (1),

$$dN/dt = rN(1 - N/K) \qquad (1)$$

where N is the size of the population in numbers of biomass at time t, r is the intrinsic rate of increase of the population, and K is the carrying capacity. The population approaches the stable equilibrium, $N = K$. *See* MATHEMATICAL ECOLOGY.

This simple theoretical approach has many extensions. A simple modification of the logistic model theorizes that for many species, such as small mammals like mice or voles and forest insects, density dependence acts in a delayed fashion and can produce cycles. Delayed density dependence would arise if a population consumes its food supply, which would then reduce the food available for the next generation. Cannibalism would have the same dynamical effect.

If the population is assumed to have discrete, nonoverlapping generations, the resulting dynamics can be much more complex, as was first elucidated by Robert May. One model of this form is called the Ricker equation (2), where N_t is the population size

$$N_{t+1} = N_t e^{r[t - (N_t/k)]} \qquad (2)$$

in generation t, r is the intrinsic rate of increase for the population, and K is the carrying capacity. The exact form of the equation is not important, but the property of overcompensation is—the total number of individuals the following year goes down at high population levels, not just the number per capita (the number next year divided by the number this year). In this case, not only can cycles result, but also a pattern of population change called chaos.

Chaos. The study of chaos has been a recent focus of theoretical ecology. A system is defined as chaotic if it exhibits sensitive density dependence on initial conditions for all possible initial conditions, and differences in initial conditions grow exponentially. Theoretical models have clearly shown that chaos is a plausible possibility for natural populations. Several empirical studies have attempted to determine if chaos is in fact present by using statistical techniques to look at time series from a variety of natural populations. The results of these empirical studies are still being debated. *See* CHAOS.

Maximum sustainable yield. Modifications and extensions of the theoretical approaches like the logistic model have also been used to guide the management of renewable natural resources. Here, the most basic concept is that of the maximum sustainable yield, which is the greatest level of harvest at which a population can continue to persist. In simple models, this is found by maximizing the growth rate of the populations as a function of the population level. The growth rate of the population at this level is then the maximum sustainable yield. In the logistic model, the maximum sustainable yield occurs when the population is at half the carrying capacity. Not surprisingly, allowing for random environmental fluctuations in population levels greatly changes the determination of maximum sustainable yield.

Stochasticity. Inclusion of random population fluctuations is important in many theories of population dynamics, especially in applied areas. Ecologists distinguish demographic stochasticity (fluctuations due to the randomness of births and deaths in small populations) from environmental stochasticity (fluctuations due to environmental randomness such as weather). Once stochasticity is included, all populations will eventually go extinct if there is a ceiling to the population level, so the focus changes from looking for a population equilibrium to calculating times to extinction or probability of extinction over a fixed time. Using the concept of minimum viable population size, conservation biologists have emphasized the role of stochasticity in predicting the probability that a population will persist over a fixed time. *See* STOCHASTIC PROCESS.

Metapopulation. Another important theoretical approach to modeling the dynamics of a single species, tracing back to the work of Richard Levins, focuses on the concept of a metapopulation. Ecologists have observed that many species have a population structure in which there are relatively small subpopulations, each of which can go extinct, and new subpopulations are founded by colonizers from other subpopulations. The simplest model describing this situation would focus only on the proportion of habitat or patches that are occupied by the species of interest. The rate of change for this fraction of occupied patches is given by the balance between the per-patch extinction rate and the colonization rate. This model has the same mathematical form as the logistic equation and thus predicts (if the colonization rate is high enough relative to the extinction rate) that the fraction of occupied habitat approaches an equilibrium. The metapopulation approach has been used to understand the effects of human-caused habitat loss and fragmentation on ecological species and communities.

Population dynamics of interacting species. Interactions between two species have been very fruitful areas of theoretical ecology, as ecologists attempt to understand communities by first considering the simplest communities—those with two species. In the study of these interactions, the concept of stability is paramount. The basis behind many of the investigations is that natural communities correspond to stable equilibria of the models, so the conditions for stable equilibria to exist should be determined. An equilibrium is stable if solutions that start near

the equilibrium approach the equilibrium. A simple analogy would be a marble in a bowl: even though you roll the marble around in the bowl, it will always approach the stable equilibrium, which is the marble at the bottom of the bowl. Although this reliance on stable equilibria has been called into question in recent years, it still is a fruitful way to proceed in many cases.

The interactions between two species are often classified as competition (two species that share common resources), predation (one species eating another), or mutualism (two species that help each other). Additional categories are host-parasitoid interactions (a specialized predatory-prey interaction among insects in which the parasitoid lays its eggs in the larvae of the host) and disease or parasitic interactions.

Competition. Competition theory, beginning with the classic work of A. J. Lotka, V. Volterra, and G. F. Gause in the 1920s and 1930s, is perhaps the most influential ecological theory. The central theme of this work is that coexistence between competitors requires differences between the two species. More recent work suggests that coexistence of plants requires that they have different limiting nutrients. *See* PLANT-ANIMAL INTERACTIONS.

Predator-prey. The study of interaction between predator and prey also goes back to the work of Lotka, Volterra, and Gause. Here, a central finding is that this kind of interaction can produce cycles in population numbers because there is essentially a form of delayed density dependence—increases in prey numbers lead to increases in predator numbers, which reduces the number of prey, which leads to a reduction in the number of predators, which causes prey populations to rebound (and so on).

Diseases. Diseases are another kind of interaction between species. Theory describing the dynamics of diseases was developed in the 1920s, when biologists and mathematicians attempted to predict and describe how diseases such as malaria spread. A disease will increase in prevalence only if its reproductive number (the number of new infections caused by a single individual) is greater than one. This number depends on the probability of transmission of the disease and the size of the host population. Using this concept, the fraction of a population that needs to be immunized to cause a disease to die out (as with polio or smallpox) can be determined. Because good records are available for childhood diseases such as measles, disease dynamics has played a central role in recent attempts to make quantitative matches between ecological theory and natural systems. *See* EPIDEMIOLOGY.

Community theory. The primary goal of ecological theory at the community level is to understand diversity at local and regional scales. Recent work has emphasized that a great deal of diversity in communities may depend on trade-offs. For example, a trade-off between competitive prowess and colonization ability is capable of explaining why so many plants persist in North American prairies. Plant species that are competitively superior can hold onto a piece of land they currently occupy, but when they die, species that are good colonizers (but poor competitors) arrive early and may persist for long enough that they reproduce before being displaced. Thus instead of only one competitively superior species persisting in a prairie, a variety of species can find opportunities for reproduction in prairies because of the trade-off between competitive prowess and dispersal. Similarly, trade-off between competitive prowess and resistance to grazing or physical disturbance may explain the diversity of algae in the rocky intertidal shoreline communities. Some species thrive as colonizers of newly open intertidal space following a storm, whereas other species owe their success to their ability to displace those early colonizers. Diversity arising from such trade-offs typically produces a patchwork of species, with each species having lived in its current position for different periods of time.

Another major concept in community theory is the role of disturbance. Understanding how disturbances (such as fires, hurricanes, or wind storms) impact communities is crucial because humans typically alter disturbance. Humans may think that they are benefiting nature by stopping wildfires, when in fact the opposite is true—the fires could be essential for maintaining diversity. *See* BIODIVERSITY; ECOLOGICAL COMMUNITIES.

A related issue at the community level is to understand the effects of changes at one trophic level on numbers at another trophic level. The central theme here is the outgrowth of a question raised in the 1960s—why is the Earth "green" (with plants and vegetation)? Presumably, consumers of plants should increase in number until they reduce their food supply. One proposed solution to this paradox is that predators hold the number of herbivores low enough that there remains a large supply of green plant material. If this scenario is correct, the addition of nutrients will lead to increased plant growth, which would be reflected by increases in the number of herbivore predators. Conversely, removal of top predators that feed on the predators of herbivores should lead to a decrease in the herbivores and subsequent increase in the density of plants—an explanation for the increase in phytoplankton in overfished lakes. This notion of trophic cascades is a central theme in the study of food webs. *See* FOOD WEB; TROPHIC ECOLOGY.

Alan Hastings

Bibliography. T. Case, *An Illustrated Guide to Theoretical Ecology*, Oxford University Press, New York, 2000; A. Hastings, *Population Ecology: Concepts and Models*, Springer-Verlag, New York, 1997; R. M. May (ed.), *Theoretical Ecology*, Sinauer Associates, Sunderland, MA, 1981; J. Maynard Smith, *Evolution and the Theory of Games*, Cambridge University Press, 1982.

Theoretical physics

The description of natural phenomena in mathematical form. It is impossible to separate theoretical physics from experimental physics, since a complete

understanding of nature can be obtained only by the application of both theory and experiment. *See* PHYSICS.

Purposes. There are two main purposes of theoretical physics: the discovery of the fundamental laws of nature and the derivation of conclusions from these fundamental laws.

Discovery of fundamental laws. Physicists aim to reduce the number of laws to a minimum to have as far as possible a unified theory. When the laws are known, it is possible from any given initial conditions of a physical system to derive the subsequent events in the system. Sometimes, especially in quantum theory, only the probability of various events can be predicted. *See* DETERMINISM.

Conclusions from fundamental laws. The conclusions to be derived from the fundamental laws of nature may be of several different types.

1. Conclusions may be derived in order to test a given theory, particularly a new theory. An example is the derivation of the spectrum of the hydrogen atom from quantum mechanics; the verification of the predictions by accurate measurements is a good test of quantum mechanics. On rather rare occasions experiment has been found to contradict the predictions of an existing theory, and this has then led to the discovery of important new physical laws. An example is the Michelson-Morley experiment on the constancy of the velocity of light, an experiment which led to special relativity theory. *See* ATOMIC STRUCTURE AND SPECTRA; LIGHT; RELATIVITY.

2. Theory may be required for experiments designed to determine physical constants. Most fundamental physical constants cannot be accurately measured directly. Elaborate theories may be required to deduce the constant from indirect experiments. An example is the Millikan oil-drop determination of the electron charge, which requires the knowledge of the motion of small droplets in air as deduced from hydrodynamic theory. *See* FUNDAMENTAL CONSTANTS.

Another example is the precise determination of the value of the magnetic moment of the electron by Hans Dehmelt, to an accuracy of 1 part in 3×10^{10}. By using an elaborate theory, this can be used (1) to establish the correctness of quantum electrodynamics and (2) to determine the value of the fine structure constant α to be given by Eq. (1), where h is Planck's

$$\frac{1}{\alpha} = \frac{hc}{2\pi e^2} = 137.035991 \qquad (1)$$

constant, c is the speed of light, and e is the charge of the electron. *See* QUANTUM ELECTRODYNAMICS.

3. Predictions of physical phenomena may be made in order to gain understanding of the structure of the physical world. In this category fall theories of the structure of the atom leading to an understanding of the periodic system of elements, or of the structure of the nucleus in which various models are tested (for example, shell model or collective model). In the same category fall applications of theoretical physics to other sciences, for example, to chemistry (theory of the chemical bond and of the rate of chemical reactions), astronomy (theory of planetary motion, internal constitution, and energy production of stars), or biology.

4. Engineering applications may be drawn from fundamental laws. All of engineering may be considered an application of physics, and much of it is an application of mathematical physics, such as elasticity theory, aerodynamics, electricity, and magnetism. The generation and propagation of radio waves of all frequencies are examples of application of theoretical physics to direct practice. *See* AERODYNAMICS; ELASTICITY; ELECTRICITY; MAGNETISM; RADIO-WAVE PROPAGATION.

Content. Apart from the classification of the fields of theoretical physics according to purpose, a classification can also be made according to content. Here one may perhaps distinguish three classification principles: type of force, scale of physical phenomena, and type of phenomena.

Type of force. At present four different types of force are known to physics. The best-understood type of force is electricity and magnetism. Here the fundamental laws, Maxwell's equations, are completely known. Corrections due to quantum theory exist but can be calculated. For practical purposes electromagnetic fields can be calculated with confidence and precision, from dc fields to the shortest gamma-ray wavelength. *See* ELECTROMAGNETIC RADIATION; FORCE; MAXWELL'S EQUATIONS.

The second type of force, known for the longest time, is the gravitational force. For practical purposes, Newton's inverse-square law is usually sufficient. The most complete theory of gravitation, however, is Einstein's general theory of relativity, which has great beauty and has now been well established by the propagation of radio waves from satellites on paths going close to the surface of the Sun. *See* GRAVITATION.

The weak force of nuclear physics is responsible for radioactive beta decay. Well over 100 nuclei are known to undergo beta decay. The process is essentially

$$\text{neutron} \rightarrow \text{proton} + \text{electron} + \text{neutrino}$$

The theory was given by Enrico Fermi in 1934, and fits observation. However, quantum theory requires that the theory be renormalizable, and a theory in which one particle turns into three is not renormalizable. It was therefore postulated that the process goes in two steps:

$$\text{neutron} \rightarrow \text{proton} + W^-$$
$$W^- \rightarrow \text{electron} + \text{neutrino}$$

The W^- particle has been discovered experimentally and has a mass about 80 times that of a proton. This large mass means that the W^- is not actually formed in beta decay but is only "virtual" in the sense of quantum theory. *See* INTERMEDIATE VECTOR BOSON; RADIOACTIVITY; RENORMALIZATION; WEAK NUCLEAR INTERACTIONS.

There is also a W^+ particle, and a neutral counterpart, Z^0, of mass about 100 protons. The mass

difference between W and Z is of theoretical interest. Steven Weinberg, Sheldon L. Glashow, and Abdus Salam predicted the existence of the W and Z particles and showed that they are related to the electromagnetic force, even though in most beta decays the weak force has a strength of only about 10^{-20} times the electromagnetic one. *See* ELECTROWEAK INTERACTION; STANDARD MODEL.

The last force is the strong force which holds atomic nuclei together. According to experiment, each nucleon contains three subunits called quarks, having charges of $+2/3\ e$ (up quarks) and $-1/3\ e$ (down quarks), where e is the charge of the proton. The quarks interact by interchanging gluons. Neither quarks nor gluons exist outside the nucleon. In addition to charge, they carry another property, called color. The theory of their interaction is called quantum chromodynamics (QCD). The theory is quite complicated, but its coupling constant has been determined [Eq. (2)] and is much bigger than the fine

$$\frac{2\pi g^2}{hc} = 0.117 \qquad (2)$$

structure constant α [Eq. (1)]. *See* COLOR (QUANTUM MECHANICS); GLUONS; QUANTUM CHROMODYNAMICS; QUARKS.

Calculation of nuclear properties from quantum chromodynamics requires elaborate computer programs, and only a beginning has been made. The size of the nucleon has been determined and is in accord with experiment.

For calculation of nuclear properties, a semiempirical approach has been used. The two-nucleon interaction has many components whose magnitude is derived from two-nucleon scattering experiments. An additional interaction between three nucleons is needed to account for the binding energy of the triton, ^{3}H. With this, the binding of ^{4}He can be calculated and agrees with experiment, without invoking an additional force between four nucleons. The structure of nuclei up to ^{16}O has been calculated with this scheme. *See* FUNDAMENTAL INTERACTIONS; NUCLEAR STRUCTURE; STRONG NUCLEAR INTERACTIONS; TRITON.

Scale of physical phenomena. The motion of bodies on the scale of everyday life can be described by the classical mechanics of Isaac Newton. Phenomena in very small dimensions, especially inside atoms or atomic nuclei, must be described by quantum mechanics. The latter theory contains Newton's mechanics as a special case. *See* CLASSICAL MECHANICS.

The description of physical phenomena is also different according to the velocities of the bodies involved. When the velocity is a substantial fraction of that of light, the special theory of relativity must be used to describe the motion. (The special theory of relativity has hardly anything in common with general relativity theory except the name, and is established beyond doubt by an enormous number of experiments.) Newton's classical mechanics again is a special case of the mechanics of special relativity. *See* RELATIVISTIC ELECTRODYNAMICS; RELATIVISTIC MECHANICS.

Special relativity and quantum mechanics are examples of the development of physical theory. Neither of them has made classical mechanics wrong or obsolete, but they have extended classical mechanics into domains which were outside of human experience until 1900. When a physical law is discovered, it can be expected to hold as long as the general conditions are not radically changed from those holding in the experiments from which the law was originally derived; for example, classical mechanics holds for objects of not too small size moving with moderate velocities. Classical mechanics, not quantum theory, is still valid for microbes of 1 micrometer size. Relativity is negligible for the speed of a spacecraft. Going to still smaller sizes, no limit has been found as yet on the validity of quantum mechanics. In a completely new area (for example, where there is very small size or very high speed) it cannot be expected a priori that the same laws will continue to hold; but if the laws do change under the new conditions, this does not invalidate the old laws in the domain for which they were originally formulated, except for minor corrections. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

The most general theory of motion now known is quantum field theory, which combines both quantum mechanics and relativity theory and at the same time embodies the observed fact that particles can be created and annihilated. This theory may thus be called a unified field theory. Attempts have also been made toward other unification, in particular to unify the theories of two types of forces, gravitational and electromagnetic; this is commonly called unification field theory in the literature. In the 1930s and 1940s, these attempts were not successful. In the 1980s, by the addition of quantum field theory, a beginning toward unification may have been made. *See* GRAND UNIFICATION THEORIES; QUANTUM FIELD THEORY; SUPERGRAVITY.

Type of phenomena. The most customary classification of theoretical physics is according to the type of phenomena described. The following are the main fields under this heading:

1. Mechanics is the theory of motion of bodies under given forces. It is normally understood to involve classical mechanics only, and includes particle mechanics and mechanics of rigid bodies. In particle mechanics, celestial mechanics is an important subdivision; this includes planetary motion, the motion of artificial satellites, and the complicated motions resulting when three bodies interact (the classical three-body problem). The field of rigid-body mechanics includes the complicated theory of gyroscopic motion with and without external fields of force. *See* CELESTIAL MECHANICS; GYROSCOPE; MECHANICS; RIGID-BODY DYNAMICS.

2. Continuum mechanics is the theory of motion of bodies, taking into account their internal properties. One branch of this is the theory of elasticity, which is basic for structural engineering design. Another branch is hydro- and aerodynamics. Here a number of problems can be solved approximately by potential theory, but most of modern aerodynamics

requires a more physical approach. Knowledge of physical properties, such as those given by the equation of state of a gas, is essential; these properties can be explained only on a molecular scale. Acoustics is a classical branch of continuum mechanics. A combination of aerodynamics and electrodynamics is required for the modern field of magnetohydrodynamics. *See* ACOUSTICS; CLASSICAL FIELD THEORY; HYDRODYNAMICS; MAGNETOHYDRODYNAMICS; POTENTIALS.

3. Heat presents a problem that can be treated on a phenomenological level by thermodynamics, which is the basis of heat engineering, as well as of the theory of chemical equilibrium. On the molecular level, heat is described by statistical mechanics, which may be considered the physical foundation of thermodynamics. Beyond this, statistical mechanics permits the calculation of the properties of bulk substances (gases, liquids, and solids) in terms of their atomic properties. *See* HEAT; STATISTICAL MECHANICS; THERMODYNAMIC PRINCIPLES.

4. Electrodynamics is well understood. Subdivisions are electrostatics; the theory of stationary currents (the basis of electrical generating machinery); the theory of oscillating electrical circuits (the basis of the technology of ordinary radio); the theory of electromagnetic waves, including their propagation in air as well as in waveguides and similar devices (the basis of radar); and finally the electromagnetic theory of light. *See* ALTERNATING CURRENT; DIRECT CURRENT; ELECTRODYNAMICS; ELECTROMAGNETIC WAVE TRANSMISSION; ELECTROSTATICS.

5. Optics is customarily treated as a special field, although, strictly speaking, it is a branch of electrodynamics. Geometrical optics and the theory of diffraction phenomena are two of the principal topics. Emission and absorption of light can be understood only on the basis of atomic physics. The same is true of dispersion, that is, the behavior of the refractive index as a function of frequency. *See* OPTICS.

6. Atomic physics includes the theory of the structure of the atom; the motion of the electrons in the atom; the periodic system; the energy levels and spectral lines of atoms and molecules; the behavior of atoms and molecules in external fields; and collisions of atoms with each other, with electrons, and with other particles. Atomic physics is the basis of the calculation of properties of matter in bulk and of the emission and absorption of light. Related is the theory of molecular structure, which is the basis of theoretical chemistry. Collisions between molecules explain the rate of chemical reactions. *See* ATOMIC PHYSICS; ATOMIC STRUCTURE AND SPECTRA; MOLECULAR PHYSICS; MOLECULAR STRUCTURE AND SPECTRA.

7. Nuclear and particle physics includes the theory of nuclear forces and of the structure of atomic nuclei. A complete theory would predict all energy levels of any nucleus and thus the electromagnetic radiations which can be emitted by the nucleus. The topic also includes the theory of nuclear reactions, which is the basis of the technology of nuclear reactors. In an effort to understand the origin of nuclear forces, theoretical physicists have investigated the production and properties of mesons and the so-called strange particles. Radioactive decay, and particularly beta decay, is another branch of nuclear physics involving weak rather than strong nuclear forces. High-energy nuclear physics aims at understanding the properties of particles—nucleons as well as unstable particles of various kinds. *See* ELEMENTARY PARTICLE; MATHEMATICAL PHYSICS; NUCLEAR PHYSICS. Hans A. Bethe

Bibliography. H. Frauenfelder and E. M. Henley, *Subatomic Physics*, 2d ed., Prentice Hall, 1991; L. D. Landau and E. M. Lifshitz, *A Shorter Course of Theoretical Physics*, 2 vols., 1992.

Therapsida

An order of Reptilia, subclass Synapsida, often called advanced mammallike reptiles, that flourished from the middle Permian through the Late Triassic. The group is highly diverse and subdivided into six suborders. Two of these, Eotitanosuchia and Dinocephalia, include relatively primitive mid-Permian carnivores and herbivores. A third, the Dicynodontia, made up of small to large herbivores, was abundant in the late Permian. Dicynodonts were associated with two carnivorous suborders, the Therocephalia and Gorgonopsia, which are morphologically intermediate between Eotitanosuchia and the cynodonts. Although these five developing lines are distinct, the skulls and skeletons in each became increasingly mammal-like.

The trend continued among the highly diverse Cynodontia. This suborder includes a variety of carnivores, omnivores, and herbivores. The most highly derived herbivorous cynodonts were the tritylodonts of the Late Triassic and Early Jurassic. They were very mammallike and, although now they are placed among the therapsids, when first found they were classified as mammals.

During the evolution of some lines of the carnivorous cynodonts, the medioposterior bones of the lower jaw, including the articular which formed part of the jaw joint, became successively reduced, accompanied by increase in the relative size of the tooth-bearing bone, the dentary (see **illus.**). This trend was most fully expressed among highly derived cynodonts of the infraorder Chinquodontoidea, including the family Trithelodontidae (=Ictidosauria). In the most specialized forms the skull–jaw joint is formed partly by the "old reptilian joint" between the articular and quadrate bones and partly by the "mammalian joint" between the dentary and squamosal bones. The quadrate and articular eventually became free of the skull and jaw and, along with the single reptilian ossicle, the stapes, went to form the three ossicles of the middle ear of mammals. Coincident in time with the existence of the chinquodontoids and the tritylodonts was the appearance of the first fossils classified as mammals.

The late Triassic Therapsida existed along with the rapidly expanding Archosauria, including the diversifying dinosaurs. The two groups for the most part

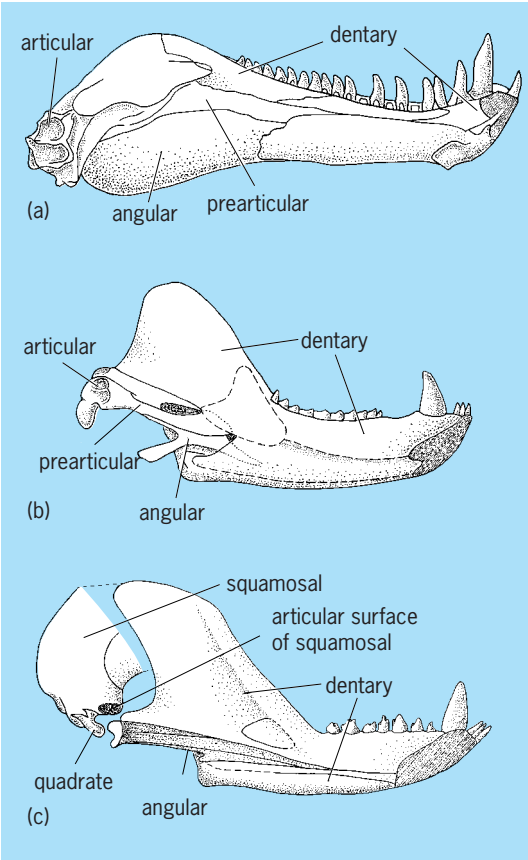

Therapsid lower jaws. (*a*) Sphenacodont pelycosaur (*after A. S. Romer and L. I. Price, Review of the Pelycosauria, Geol. Soc. Amer. Spec. Pap. 28, 1940*). (*b*) Intermediate cynodont therapsid (*after E. F. Allin, Evolution of the mammalian ear, J. Morphol., vol. 148, 1975*). (*c*) Cinquodontid therapsid [showing posterior part of the skull] (*after A. S. Romer, The Chañares (Argentine) Triassic Reptilian Fauna, Museum of Comparative Zoology, Harvard University, Breviora 344, 1970*)

inhabited different areas. Climatic changes during the Triassic rather than direct competition largely accounted for the decline of the therapsids and the rapid expansion of the Archosauria. Only the very mammallike therapsids, the herbivorous tritylodonts, and the minute derived first mammals survived into the Early Jurassic. *See* ARCHOSAURIA; MAMMALIA; REPTILIA; SYNAPSIDA. Everett C. Olson

Bibliography. M. J. Burton, *Vertebrate Paleontology*, 1990; R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1987.

Theria

One of the four subclasses of the class Mammalia, including all living mammals except the monotremes. The Theria were by far the most successful of the several mammalian stocks that arose from the mammallike reptiles in the Triassic. The subclass is divided into three infraclasses: Pantotheria (no living survivors), Metatheria (marsupials), and Eutheria (placentals). These were not strictly contemporaneous; the Pantotheria arose directly from mammallike reptiles, and the Metatheria and Eutheria in turn arose from pantotherelike forms during the Jurassic or Cretaceous, many millions of years later. Therian mammals are characterized by the distinctive structural history of the molar teeth. The fossil record shows that all the extremely varied therian molar types were derived from a common tribosphenic type in which three main cusps, arranged in a triangle on the upper molar, are opposed to a reversed triangle and basinlike heel on the lower molar. *See* MAMMALIA. D. Dwight Davis; Frederick S. Szalay

Thermal analysis

A group of analytical techniques developed to continuously monitor physical or chemical changes of a sample which occur as the temperature of a sample is increased or decreased. Thermogravimetry, differential thermal analysis, and differential scanning calorimetry are the three principal thermoanalytical methods. Commercial instruments are available for all types of thermal analysis. Most instruments use modern computer technology to control the instrument while collecting and analyzing data. *See* ANALYTICAL CHEMISTRY; COMPUTER.

The occurrence of physical or chemical changes upon heating a sample may be explained from either a kinetic or thermodynamic viewpoint. Kinetically the rate of a process may be increased by raising the temperature as shown by the Arrhenius equation (1),

$$\text{Rate} = A\epsilon^{-E_a/RT} \tag{1}$$

where A, E_a, and R represent the preexponential factor, activation energy, and the gas law constant, respectively. At some point, the rate becomes significant and readily observable. Similarly, an increase in temperature can change the Gibbs free energy [Eq. (2), where ΔG° is the Gibbs free energy, ΔH°

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \tag{2}$$

is the reaction enthalpy, and ΔS° is the entropy change for the process] to a more favorable (that is, more negative) value. In particular, ΔG° will become more negative if ΔS° is positive and the temperature is increased. In many cases a combination of these factors causes the observed physiochemical process. *See* CHEMICAL THERMODYNAMICS; KINETICS (CLASSICAL MECHANICS).

Instrumentation for thermal analysis requires sensitive and robust temperature sensors and mechanisms for heating samples. Results are best with small (milligram) sample sizes and low heating rates. Sample and reference holders must be chemically inert and allow rapid thermal equilibration. Samples may be studied in air or other atmospheres, inert or reactive, at reduced, ambient, or elevated pressures. Most samples are solids, although an increasing number of liquid samples are being analyzed. References are used only in differential methods. Exact experimental conditions are chosen to enhance the process under study and ensure reproducibility. Selectable experimental parameters are heating

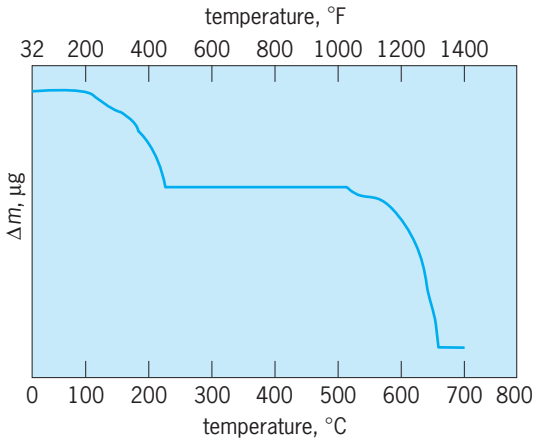

Fig. 1. Typical thermogravimetric curve for calcium nitrate tetrahydrate, $Ca(NO_3)_2 \cdot 4H_2O$, illustrating mass loss.

(cooling) rate, temperature range, atmospheric composition and pressure, and sample preparation. Each of these may affect the thermogram produced.

Thermograms are plots of the measured property, such as mass in thermogravimetry or differential heat flow in differential scanning calorimetry, versus the sample temperature (**Fig. 1**). Under rigorously controlled conditions, thermograms uniquely represent the system under study. Melting, crystallization, decomposition, oxidation, adsorption, absorption, desorption, polymerization reactions, and heat-capacity changes are observable on thermograms. A unique feature of thermograms is that they may also reflect a sample's history or method of preparation such as the cooling rate used to prepare a polymer, or ultraviolet or nuclear radiation exposure and damage.

In the study of reaction kinetics, isothermal analyses are often performed. This procedure changes the sample temperature rapidly from ambient to a desired value; the thermogram is the appropriate physiochemical property plotted as a function of time.

Thermogravimetry. This method involves measuring the changes in mass of a substance, typically a solid, as it is heated. Specially designed thermobalances are required to continuously monitor sample mass during the heating process. Modern balances have a capacity of 1–1500 milligrams and can accurately detect mass changes of 0.1 microgram.

Any type of physiochemical process which involves a change in sample mass may be observed by using thermogravimetry. Mass losses are observed for dehydration, decomposition, desorption, vaporization, sublimation, pyrolysis, and chemical reactions with gaseous products. Mass increases are noted with adsorption, absorption, and chemical reactions of the sample with the atmosphere in the oven, such as the oxidation of metals.

Quantitative gravimetric analyses may be performed due to the precise measure of the mass change obtained. Rates of mass change have been used to evaluate the kinetics of a process and to estimate activation energies. Fine details of these thermograms may also be used to deduce reaction intermediates and reaction mechanisms (Fig. 1).

Primary applications of thermogravimetry are to deduce stabilities of compounds and mixtures of elevated temperatures and to determine appropriate drying temperatures for compounds and mixtures. Evaluation of polymers, food products, and pharmaceuticals is a major application of thermogravimetry.

Differential thermal analysis. This method involves the monitoring of the temperature difference T_D between a sample and inert reference material (such as aluminum oxide) as they are simultaneously heated, or cooled, at a predetermined rate. Multijunction thermocouples and thermistors are the most common temperature sensors used for this purpose; they are arranged in an oven (**Fig. 2**). As enthalpic changes occur, T_D will be positive if the process is exothermic and negative if it is endothermic. Endotherms are plotted as negative deviations from the baseline (**Fig. 3**).

Differential thermal analysis thermograms are affected by instrumental factors such as furnace design, sample-holder material and geometry, thermocouple size and placement, and instrument response and furnace atmosphere. Sample characteristics that affect heat flow such as particle size, thermal conductivity, heat capacity, packing density, and the amount of sample can affect the appearance of thermograms. Under rigorously specified conditions, thermograms will uniquely represent the physicochemical characteristics of the sample.

More physical and chemical processes may be observed using differential thermal analysis as compared to thermogravimetry. Endothermic physical processes include crystalline transitions, fusion, vaporization, sublimation, desorption, and adsorption. Endothermic physical processes include crystalline

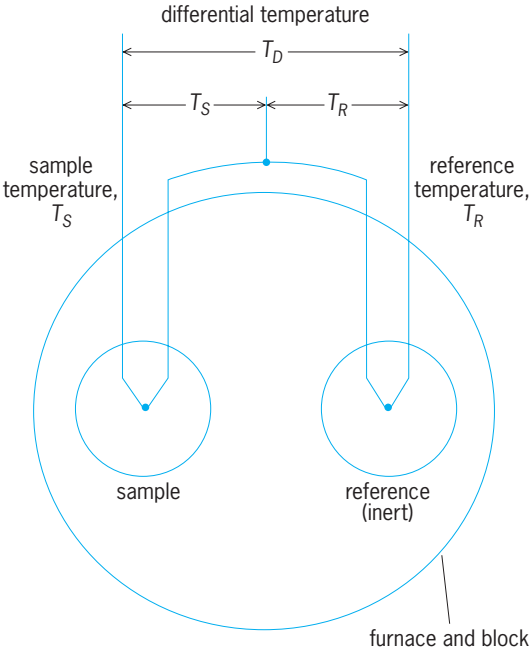

Fig. 2. Diagram of heating and temperature sensing units for differential thermal analysis. The oven heats the sample and the reference, and thermocouples determine ΔT (change in temperature) as well as the oven temperature.

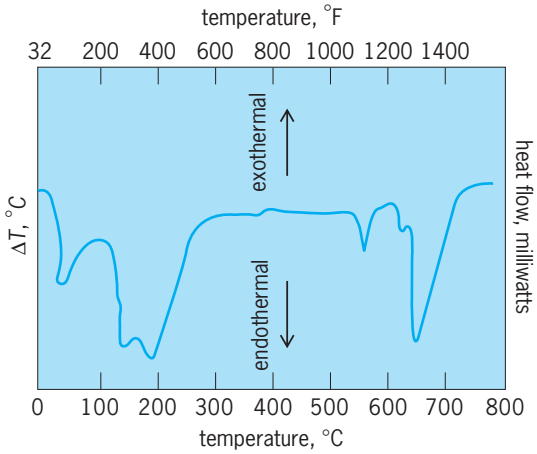

Fig. 3. Typical thermogravimetric curve for calcium nitrate tetrahydrate, $Ca(NO_3)_2 \cdot 4H_2O$, for differential thermal analysis when the *y* axis is ΔT, and for differential scanning calorimetry when the *y* axis has units of heat flow in milliwatts.

transitions, fusion, vaporization, sublimation, desorption, and adsorption. Endothermic chemical processes include dehydration, decomposition, gaseous reduction, redox reactions, and solid-state metathesis. Exothermic processes include adsorption, chemisorption, decomposition, oxidation, redox reactions, and solid-state metathesis reactions. Both solids and liquids can be studied by differential thermal analysis. Hermetically sealed capsules are often used for liquids and some solids. Other samples are studied in open or crimped pans.

Analytical applications of this technique include the identification, characterization, and quantitation of a wide variety of materials, including polymers, pharmaceuticals, metals, clays, minerals, and inorganic and organic compounds. Characteristic thermograms can be used to determine purity, heats of reaction, thermal stability, phase diagrams, catalytic properties, and radiation damage.

Differential scanning calorimetry. In this method, a sample and a reference are individually heated at a predetermined rate by separately controlled resistance heaters (**Fig. 4**). Enthalpic (heat-generating or -absorbing) processes are detected as differences in electrical energy supplied to either the sample or the reference material to maintain this heating rate (Fig. 3). This difference in electrical energy, in milliwatts per second, of the heat flow into or out of the sample is due to the occurrence of a physical or chemical process. Modulated differential scanning calorimetry is a new method that superimposes a sine wave on the heating ramp. A significant increase in sensitivity is often observed with modulated differential scanning calorimetry.

Analytical uses of differential scanning calorimetry are very similar to those of differential thermal analysis. Usually one calibration standard is sufficient to calibrate the entire operating range of the instrument. Differential scanning calorimetry instruments are highly sensitive and may measure heat flows as small as 1 nanowatt. Differential scanning calorimetry is very useful in determining heat capacities of substances over large temperature ranges. Such evaluation has become important in polymer and biochemical studies. Small (approximately 1–10 mg) samples are used in most cases, although some instruments have been developed which use up to 1 ml of a liquid sample. *See* CALORIMETRY.

Quantitative analysis in both differential thermal analysis and differential scanning calorimetry involves relationship (3) between the peak area A,

$$A = k(m)(\Delta H) \tag{3}$$

sample mass m, and enthalpy ΔH of the transition. The proportionality constant k varies with temperature in differential thermal analysis. The best results are obtained when the instrument is calibrated (k is determined) with a standard that has an enthalpic peak at a temperature close to the sample's peak. In differential scanning calorimetry the value of k varies less with temperature but still should be determined as with differential thermal analysis. A number of reliable calibration standards are readily available.

A significant property measurable by both differential thermal analysis and differential scanning calorimetry is the glass transition temperature T_g of polymers. Polymers often soften over a wide temperature range as glass does, rather than melt at a sharply defined temperature as smaller compounds do. The T_g of a polymer is an important physical property. *See* POLYMER.

Other methods. Thermomechanical analysis, thermoluminescence, emanation thermal analysis, evolved-gas analysis, thermomagnetic analysis, differential thermal rheology, differential photocalorimetry, dynamic mechanical thermal analysis, simultaneous thermal analysis (thermogravimetry and

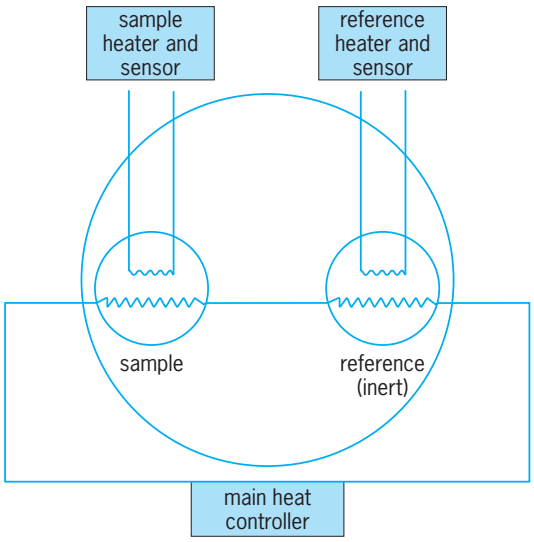

Fig. 4. Diagram of heating and temperature sensing units for differential scanning calorimetry. The main heater heats the sample and the reference equally and at a rate slightly less than the preset heating rate. Individual sample and reference heaters add heat to maintain the preset heating rate. The difference in energy to individual heaters is the heat flow plotted in the thermogram.

differential thermal analysis together), and microthermal analysis are additional thermal analytical methods. Thermomechanical analysis can be used to evaluate the physical stability of structural or electronic components. Thermoluminescence is often used for authenticating ancient objects. Emanation and evolved-gas thermal analysis are used to determine gaseous products, often from destructive oxidation. These methods are often combined with mass spectrometry or gas chromatography/mass spectrometry. Thermomagnetic analysis can be used to determine Curie points of metals. Differential thermal rheology measures the deformation of matter with temperature increases. Differential photo calorimetry monitors heat generated by light-induced chemical processes. Dynamic mechanical thermal analysis instruments can monitor cantilever bending, tensile strength, shear, compression, and creep as a function of temperature. Simultaneous analysis allows differential thermal analysis and thermogravimetry to be performed on the same sample at the same time with a considerable savings in operator time. Microthermal analysis places a heater and temperature sensor at the tip of an atomic force microscope to measure microscopic thermal features. *See* RHEOLOGY; SCANNING ELECTRON MICROSCOPE.

In some instances the first derivative of the thermogram is the best format for qualitative assessment of samples. For instance, the first-derivative plots show small features, such as shoulders on peaks, more clearly. Derviative curves are produced by digital calculations on the computer stored data. Virtually every method discussed above has been adapted to differential methods. Integration of peak areas for quantitative work is now done by special algorithms and computer programs. Computer-driven calculations provide consistent data treatment unattainable with manual analysis of thermograms. *See* CHEMICAL DYNAMICS; PHASE EQUILIBRIUM; THERMAL EXPANSION; THERMOCHEMISTRY; TRANSITION POINT.

Neil D. Jespersen

Bibliography. J. Cazes (ed.), *Ewing's Analytical Instrumentation Handbook*, 3d ed., 2004; P. J. Haines (ed.), *Principles of Thermal Analysis and Calorimetry*, 2002; G. Hohne, W. Hemminger, and H. J. Flemmersheim (eds.), *Differential Scanning Calorimetry: An Introduction*, Springer, Berlin, 2d ed., 2003; B. Wunderlich, *Thermal Analysis of Polymeric Materials*, 2005.

Thermal conduction in solids

Thermal conduction in a solid is generally measured by stating the thermal conductivity K, which is the ratio of the steady-state heat flow (heat transfer per unit area per unit time) along a long rod to the temperature gradient along the rod. Thermal conductivity varies widely among different types of solids, and depends markedly on temperature and on the purity and physical state of the solids, particularly at low temperatures.

From the kinetic theory of gases the thermal conductivity can be written as in Eq. (1), where S is the

$$K = (\text{constant})\, Svl \tag{1}$$

specific heat per unit volume, v is the average particle velocity, and l is the mean free path. In solids, thermal conduction results from conduction by lattice vibrations and from conduction by electrons. In insulating materials, the conduction is by lattice waves; in pure metals, the lattice contribution is negligible and the heat conduction is primarily due to electrons. In many alloys, impure metals, and semiconductors, both conduction mechanisms contribute. *See* CONDUCTION (HEAT); KINETIC THEORY OF MATTER; LATTICE VIBRATIONS; SPECIFIC HEAT OF SOLIDS.

Insulating solids. In insulators, heat is transported by lattice vibrations. The vibrations of the crystal lattice can be resolved into traveling elastic waves (lattice waves). Each lattice wave can be specified by a wave vector **k** and by a mode of polarization (either longitudinal or transverse). The energy of the lattice vibrations is taken as the energy of lattice waves of frequency ν, where ν is of the order of k_BT/h; k_B is Boltzmann's constant, T is the absolute temperature, and h is Planck's constant. The energy of each wave is described as consisting of an integral number of phonons, each of energy $h\nu$ and of momentum $hk_B/2\pi$. *See* PHONON.

At equilibrium, the phonons are distributed isotropically in momentum space; when a temperature gradient is established along the crystal, an asymmetry in the phonon distribution is set up. An interchange of energy among the lattice waves, which can result from the anharmonicity of the lattice forces and from lattice imperfections, tends to restore equilibrium; the rate at which equilibrium is restored is associated with a finite mean free path l for the

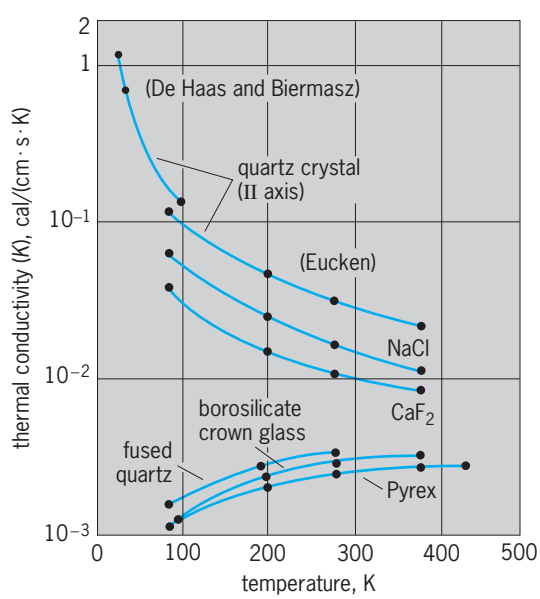

Fig. 1. Thermal conductivity of some insulating and some amorphous materials as a function of temperature. °F = (K × 1.8) − 459.67. (*After C. Kittel, Introduction to Solid State Physics, 3d ed., Wiley, 1966*)

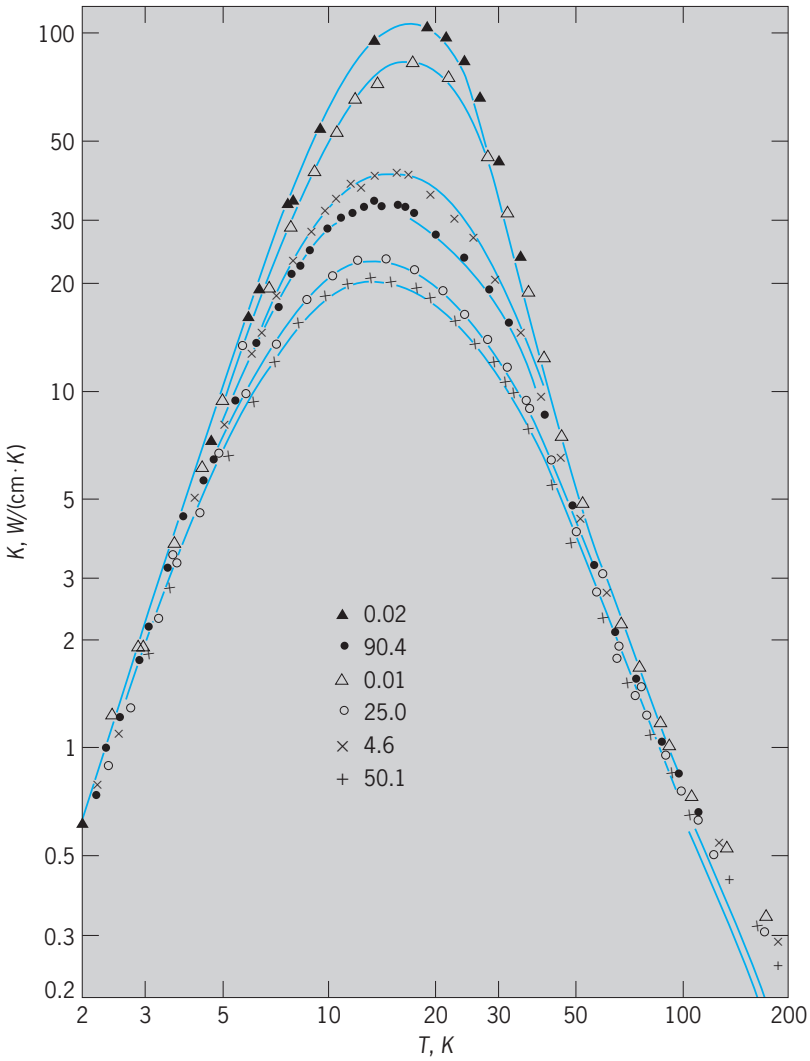

Fig. 2. Thermal conductivity of lithium fluoride as a function of temperature for various percentage concentrations of ^{6}Li. In the peak the highest value is approximately five times that of the lowest peak. °F = (K × 1.8) − 459.67. (*After R. Berman, Heat conductivity of non-metallic crystals at low temperatures, Cryogenics, 5:297–305, 1965*)

phonons. At low temperatures K varies markedly with temperature and is dependent on the variation of l with phonon frequency ν. **Figure 1** shows the conductivity of some insulating materials.

The form of l at low temperatures is closely related to the type of phonon scattering and hence to the type or types of imperfections present in the particular solid. The thermal conductivity at low temperatures is a sensitive function of these imperfections. For example, in **Fig. 2**, the effect of the isotope of lithium ^{6}Li on the conductivity of lithium fluoride crystals is shown; here the differences in conductivity are due principally to the mass difference between ^{6}Li and ^{7}Li.

The introduction of molecular impurities into dielectric crystals produces resonance scattering, as shown in **Fig. 3** for the potassium chloride crystals containing cyanide ions; here the center of the resonance (the dip in the thermal conductivity curves) corresponds to the frequency of the particular phonons carrying heat at that temperature. The thermal conductivity of solid helium has been measured at low temperatures, and is shown in **Fig. 4**; for comparison, the thermal conductivities of very pure lithium fluoride and sapphire are given. At less than 1 K (2°F above absolute zero), the conductivity of solid helium increases very rapidly with temperature and is characterized as Poiseuille flow, in which the phonons are described as flowing as a viscous fluid. *See* FLUID MECHANICS; HYDRODYNAMICS.

Glass and polycrystalline materials. For amorphous materials, glasses, and polycrystalline materials, the mean free path for the phonons is generally small, of the order of the distance between atoms, and is independent of temperature. The thermal conductivity for these materials is much less than that for the insulating materials above; for example, quartz (fused) glass can be compared with crystal quartz in Fig. 1, and sintered alumina (aluminum oxide) in **Fig. 5** can be compared with sapphire (single-crystal aluminum oxide) in Fig. 4.

Metals, alloys, and semiconductors. The thermal conductivity of a metal can be written as the sum of an electronic component K_e and a lattice component

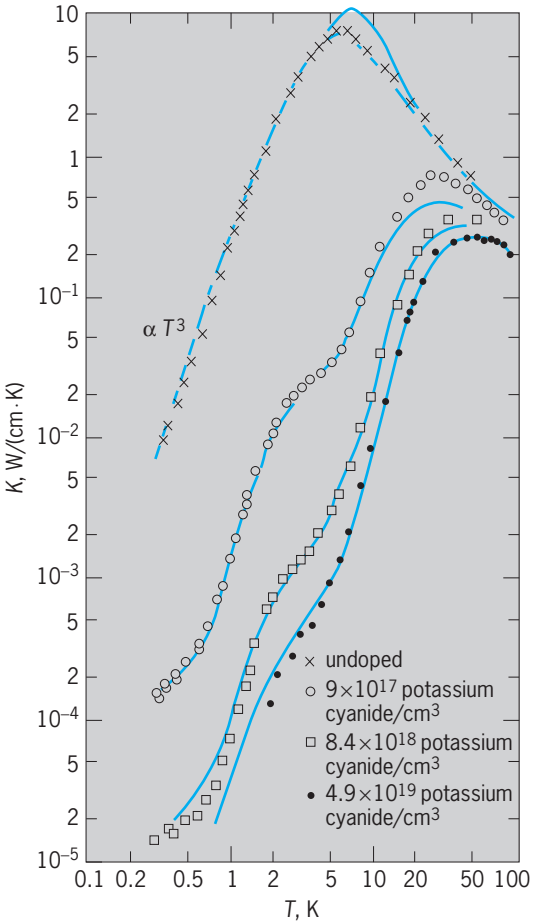

Fig. 3. Thermal conductivity of pure potassium chloride and of potassium chloride containing cyanide ions as a function of temperature. °F = (K × 1. 8) − 459.67. (*After R. Berman, Heat conductivity of non-metallic crystals at low temperatures, Cryogenics, 5:297–305, 1965*)

K_g; K_e and K_g are each limited by various scattering mechanisms. In real crystals, as compared to the ideal crystal with a perfect lattice, the electrons undergo scattering caused by the thermal vibrations of the lattice and by imperfections (such as impurities, point defects, dislocations, and grain boundaries).

In pure metals, the electronic component accounts for nearly all the heat conducted, while the lattice component, in most cases, is negligible. The electronic thermal conductivity is related to the electrical conductivity through the mean free path. In certain temperature regions the value of the mean free path for both thermal and electrical conduction can be assumed to be the same; for these cases, the Wiedemann-Franz law is applicable, as in Eq. (2),

$$\frac{K_e}{\sigma T} = L = \frac{\pi k_B^2}{3e^2} \qquad (2)$$

where L (the Lorentz constant) $= 2.45 \times 10^{-8}$ (watt)(ohm)(K^{-2}) $= 7.5 \times 10^{-9}$ (watt)(ohm)(°F^{-2}), σ is the electrical conductivity, and e is the electronic charge. Here the free-electron theory of metals is assumed, in which the conduction electrons form a gas which obeys Fermi-Dirac statistics. *See* FREE-ELECTRON THEORY OF METALS.

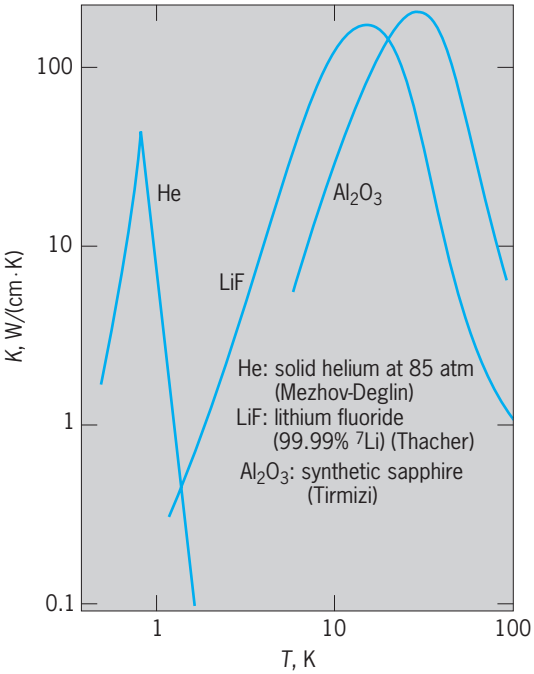

Fig. 4. Thermal conductivity of solid helium as a function of temperature. Thermal conductivities of very pure lithium fluoride and sapphire are given for comparison. °F = (K × 1.8) − 459.67. (*After R. Berman, Heat conductivity of non-metallic crystals at low temperatures, Cryogenics, 5:297–305, 1965*)

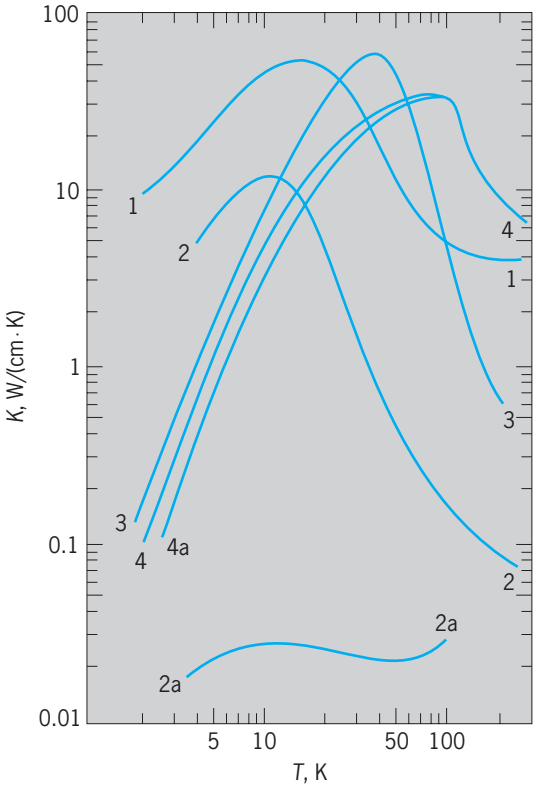

Fig. 5. Thermal conductivity of some insulators as a function of temperature. Curve 1, pure copper (for comparison). Curve 2, quartz crystal. Curve 2*a*, quartz crystal after intense neutron irradiation. Curve 3, aluminum oxide (sintered alumina). Curve 4, diamond. Curve 4*a*, diamond, reduced diameter. °F = (K × 1.8) − 459.67. (*After R. Berman, Heat conductivity of non-metallic crystals at low temperatures, Cryogenics, 5:297–305, 1965*)

The electrical resistivity ρ (the reciprocal of the electrical conductivity σ) can be separated into two parts, one called the residual resistivity ρ_0, which results from the elastic scattering of electrons by imperfections and which is independent of temperature, and the other called the ideal resistivity $\rho_i(T)$, which results from scattering due to lattice vibrations and which is temperature-dependent. Thus the electrical resistivity is written $\rho = 1/\sigma = \rho_0 + \rho_i(T)$. The thermal resistivity for metals W (the reciprocal of the thermal conductivity K_e) can be written analogously as the sum of two terms, one, W_0, due to scattering by imperfections, and the other, $W_i(T)$, due to scattering by lattice vibrations. Thus the thermal conductivity is written $W = 1/K_e = W_0 + W_i(T)$. In **Fig. 6** the thermal conductivities of three samples of gold are given; since the purity of the gold differs, the samples differ in both ρ_0 and W_0. In the low-temperature region, the thermal conductivity is proportional to T. At higher temperatures, the thermal resistivity due to lattice vibrations exceeds W_0, and the curves for the three different samples converge.

In the case of alloys, W_0 is much larger than in pure metals and a lattice thermal conductivity must also be included. The thermal conductivities of several alloys of gold are also shown in Fig. 6. It is possible from such curves to obtain K_g as a function of T by also measuring ρ; ρ_0 can be determined from W_0. In well-annealed alloys, K_g is proportional to T_2 at low temperatures; at higher temperatures, K_g decreases with T.

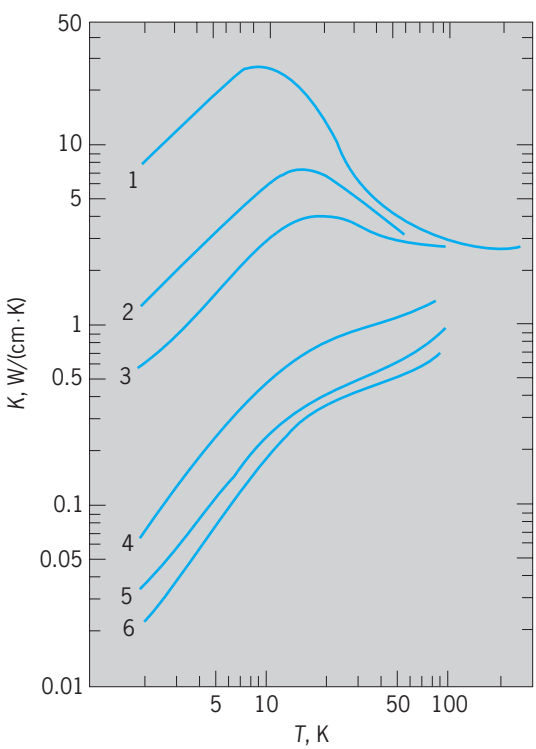

Fig. 6. Thermal conductivity of gold and some alloys as a function of temperature. Curve 1, 99.999% gold, annealed. Curve 2, same, cold-drawn. Curve 3, 99.9% gold, cold-drawn. Curve 4, gold with 0.7% platinum. Curve 5, gold with 1.7% platinum. Curve 6, gold with 0.2% chromium. °F = (K × 1.8) − 459.67. (*After F. Seitz and D. Turnbull, eds., Solid State Physics, vol. 7, 1958*)

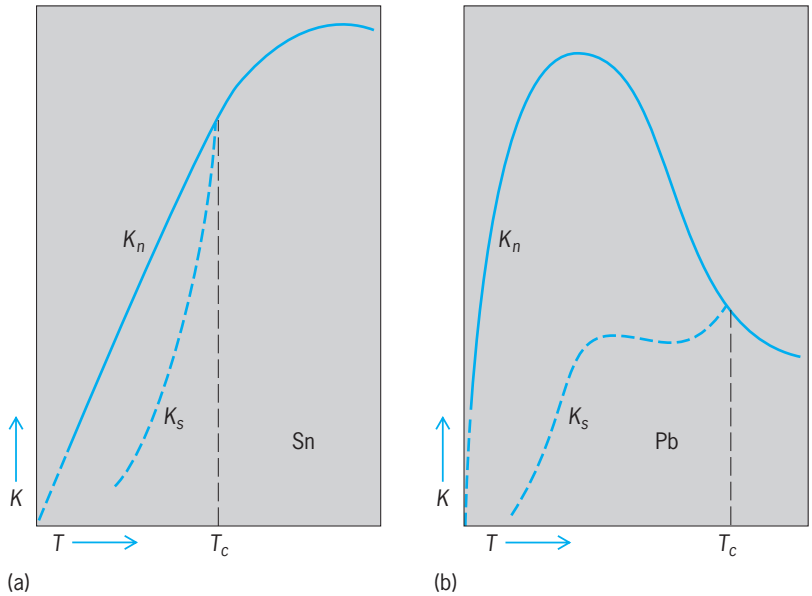

Fig. 7. Temperature dependences of thermal conductivity of (*a*) tin and (*b*) lead in the normal and superconducting states. (*After F. Seitz and D. Turnbull, eds., Solid State Physics, vol. 12, 1961*)

In the case of semiconductors, such as high-purity germanium and silicon, phonons are primarily responsible for the thermal conduction. The form of the thermal conductivity versus temperature curves is similar to that for single crystals of dielectric materials.

Superconductors. In superconductors at temperatures below the critical temperature T_c, the electronic conduction is reduced; at sufficiently low temperatures, the thermal conductivity becomes entirely due to lattice waves and is similar to the form of the thermal conductivity of an insulating material. In **Fig. 7** the temperature dependences of the thermal conductivity of tin and lead are given; here K_n is the thermal conductivity with the material in the normal state, and K_s is the thermal conductivity with the material in the superconducting state. *See* SUPERCONDUCTIVITY. Kathryn A. McCarthy

Bibliography. R. Berman, *Thermal Conduction in Solids*, 1976; C. M. Bhandari and D. M. Rowe, *Thermal Conduction in Semiconductors*, 1988; G. E. Childs, L. J. Erick, and R. L. Powell, *Thermal Conductivity of Solids at Room Temperature and Below: A Review and Compilation of the Literature*, National Bureau of Standards Monogr. 131, 1973; Y. Godovsky and V. P. Privalko, *Thermal and Electrical Conductivity of Polymer Materials*, 1995; P. G. Klemens and T. K. Chu, *Thermal Conductivity*, 1976.

Thermal converters

Devices consisting of a conductor heated by an electric current, with one or more hot junctions of a thermocouple attached to it, so that the output emf

responds to the temperature rise, and hence the current. Thermal converters are used with external resistors for alternating-current (ac) and voltage measurements over wide ranges and generally form the basis for calibration of ac voltmeters and the ac ranges of instruments providing known voltages and currents.

Basic form. In the most common form, the conductor is a thin straight wire less than 0.4 in. (1 cm) long, in an evacuated glass bulb, with a single thermocouple junction fastened to the midpoint by a tiny electrically insulating bead. Thermal inertia keeps the temperature of the heater wire constant at frequencies above a few hertz, so that the constant-output emf is a true measure of the root-mean-square (rms) heating value of the current. The reactance of the short wire is so small that the emf can be independent of frequency up to 10 MHz or more. An emf of 10 mV can be obtained at a rated current less than 5 mA, so that resistors of reasonable power dissipation, in series or in shunt with the heater, can provide voltage ranges up to 1000 V and current ranges up to 20 A. However, the flow of heat energy cannot be controlled precisely, so the temperature, and hence the emf, generally changes with time and other factors. Thus an ordinary thermocouple instrument, consisting of a thermal converter and a millivoltmeter to measure the emf, is accurate only to about 1–3%. *See* THERMOCOUPLE.

AC-DC transfer instrument. To overcome this, a thermal converter is normally used as an ac-dc transfer instrument (ac-dc comparator) to measure an unknown alternating current or voltage by comparison with a known nearly equal dc quantity (see **illus.**). By replacing the millivoltmeter with an adjustable, stable, opposing voltage V_b in series with a microvoltmeter D, very small changes in emf can be detected. The switch S is connected to the unknown ac voltage V_{ac}, and V_b is adjusted for a null (zero) reading of D. Then S is immediately connected to the dc voltage V_{dc}, which is adjusted to give a null again, without changing V_b. Thus $V_{ac} = V_{dc}(1 + d)$, where d is the ac-dc difference of the transfer instrument, which can be as small as a few parts per million (ppm).

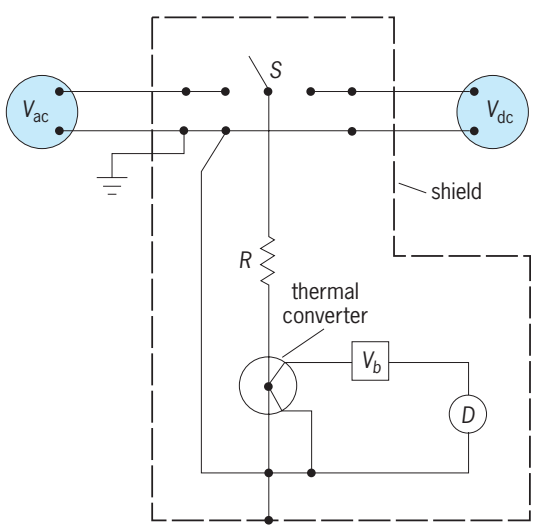

Basic circuit for ac-dc transfer measurements of ac voltages.

In many commercial instruments, all of the components are conveniently packaged in the shield, shown with a broken line, and several ranges are available by taps on R. Accuracies of 0.001% are attainable at audio frequencies. For the highest accuracy, the average of the two directions of V_{dc} should be taken as the reference, because of slight imperfections of the thermal converter, and the ac measurement should be repeated, to guard against small drifts. The resistances of the two pairs of leads from V_{ac} and V_{dc} to S should be equal, and the leads should be shielded.

Multijunction thermal converter. A converter with many thermocouples attached uniformly along a bifilar heater can have intrinsic values of d less than 0.1 ppm from thermoelectric effects. Such multijunction thermal converters (MJTCs) are the primary ac-dc standards in several national metrology laboratories. MJTCs can be made from wire heaters and thermocouples or by using microfabrication techniques to produce thin-film heaters and thermocouples on silicon substrates. The heaters and thermocouple hot junctions are formed on a thin, thermally isolating membrane such as SiO_2/Si_3N_4 (not directly on the silicon substrate), and the cold junctions are formed on the silicon substrate to provide a stable temperature reference.

Solid-state converter. This type of converter uses a transistor, sensitive to temperature, in place of a thermocouple. Dual heaters and transistors are connected in a feedback circuits so that an output voltage proportional to the input current is obtained. *See* TRANSISTOR.

MJTCs and solid-state converters are also used in a number of more complex instruments. In some instruments the heater is switched at regular intervals between the unknown ac input and a dc current that is automatically controlled to keep the output constant. Other instruments have dual multijunction thermal converters, with two sets of thermocouples connected to each other in opposition and connected to an amplifier that supplies that second heater. In these feedback circuits the output voltage is equal to the voltage on the first heater. *See* FEEDBACK CIRCUIT.

Calibration. The values of d (which depend on range and frequency) are best determined by comparison with ac-dc transfer instruments of known characteristics. These, in turn, are evaluated in metrology laboratories, such as the National Institute of Standards and Technology in the United States, where ac-dc differences have been studied very carefully, because the ultimate accuracy of almost all alternating-current and voltage measurements depends upon thermal converters. *See* ELECTRICAL MEASUREMENTS. F. L. Hermach; Joseph R. Kinard

Bibliography. F. L. Hermach, AC-DC comparators for audio-frequency current and voltage measurements of high accuracy, *IEEE Trans. Instrum. Meas.*, IM-25:489–494, 1976; J. R. Kinard et al., Development

of thin-film multijunction thermal converters at NIST, *IEEE Trans. Instrum. Meas.*, 46(2):347–351, April 1997; E. S. Williams, *The Practical Uses of AC-DC Transfer Instruments*, NBS Tech. Note 1166, 1982.

Thermal ecology

The examination of the independent and interactive biotic and abiotic components of naturally heated environments. Geothermal habitats are present from sea level to the tops of volcanoes and occur as fumaroles, geysers, and hot springs. Hot springs typically possess source pools with overflow, or thermal, streams (rheotherms) or without such streams (limnotherms). Hot-spring habitats have existed since life began on Earth, permitting the gradual introduction and evolution of species and communities adapted to each other and to high temperatures. Other geothermal habitats do not have distinct communities.

Hot-spring pools and streams, typified by temperatures higher than the mean annual temperature of the air at the same locality and by benthic mats of various colors, are found on all continents except Antarctica. They are located in regions of geologic activity where meteoric water circulates deep enough to become heated. The greatest densities occur in Yellowstone National Park (northwest United States), Iceland, and New Zealand. Source waters range from 40°C (104°F) to boiling (around 100°C or 212°F depending on elevation), and may even be superheated at the point of emergence. Few hot springs have pH 5–6; most are either acidic (pH 2–4) or alkaline (pH 7–9).

Alkaline hot springs. Alkaline springs have been studied in the greatest detail. The main results of these studies are described below.

Physicochemical properties. Heated subterranean water becomes highly charged with dissolved materials due to its increased dissolution capabilities. Hot-spring waters have 10–15 times more dissolved inorganic solutes than nonthermal waters; their nutrient levels are similar to eutrophic lakes. The major solutes of nonacid springs include calcium, chlorine, potassium, magnesium, sodium, silica, and sulfates. Toxic elements, such as arsenic and mercury, may render the water unpotable. *See* EUTROPHICATION.

While source waters remain relatively constant, the abiotic properties do change with distance from source. Thus, unidirectional gradients of heat, hydrogen sulfide, carbon dioxide, nitrates, and pH exist along these streams (**Fig. 1**). The rate of change per unit distance depends on stream characteristics, meteorological conditions, and biotic use or release of solutes. In Fig. 1, region A is inhabited by patches of bacteria attached to the substrate. Region B begins at the 73°C (163°F) isotherm with the first appearance of continuous benthic mat. Above 70°C (158°F) this mat contains only *Synechococcus lividus*; below 70°C (158°F) it is joined by *Chloroflexus aurantiacus*. Region C represents an area of increasing species diversity. At temperatures in the upper 50s

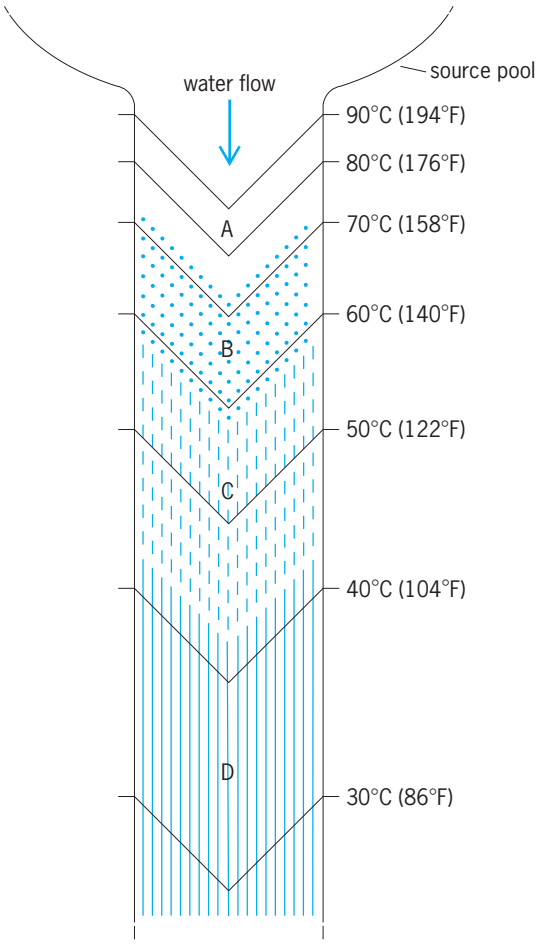

Fig. 1. Distribution of organisms and temperatures along an alkaline thermal spring. Isotherms are shown for 10°C (18°F) intervals. The distinctive V shape results from more rapid heat loss at edges than in main current; heat loss is approximately exponential with distance from source. Four regions (A–D) of differing biological character are illustrated; these are described in the text.

[58°C (136°F) here] other cyanobacteria join the mat, and at lower temperatures eukaryotes are present (see **table**). Region D often is influenced by the grazing of herbivores. While most upstream (B and C) organisms are present here as well, *C. aurantiacus* is absent, and the mat may be dominated by different species (for example, *Calothrix* spp.).

Communities. Cyanobacteria (blue-green algae) and other auto- and heterotrophic bacteria form the benthic mats of alkaline hot springs from 30 to 73°C (86 to 163°F). Mat coloration results from variations of surface pigments due to species distributions, temperature, and light history. Only prokaryotes are found above 60°C (140°F); eukaryotes are not common above 40–45°C (104–113°F; see table).

Source pools, even at boiling temperatures, may contain bacteria. These chemoautotrophs oxidize sulfides, but have not been cultured or identified. Yellow patches of organoheterotrophic bacteria *Thermus aquaticus* are present from 73 to 79°C (163 to 174°F).

As water temperature decreases downstream, overlapping populations of microorganisms appear

Upper temperature limit and physiological activity of representative hot-spring organisms

Organism	Taxonomic category	Physiological activity*	Maximum temperature, °C (°F)
	Alkaline hot springs		
Thermothrix thioparus	Bacterium	CA	80 (176)
Thermus aquaticus	Bacterium	OH	79 (174)
Synechococcus lividus	Cyanobacterium	PA	73 (163)
Chloroflexus aurantiacus	Photosynthetic bacterium	PH	70 (158)
Phormidium laminosum	Cyanobacterium	PA	60 (140)
Mastigocladus laminosus	Cyanobacterium	PA	58 (136)
Oscillatoria terebriformis	Cyanobacterium	PA	54 (129)
Potamocypris sp.	Ostracod	OH	48 (118)
Achnanthes exigua	Alga (diatom)	PA	44 (111)
Oscillatoria princeps	Cyanobacterium	PA	~42 (108)
Paracoenia turbida	Brine fly	OH	36–38 (97–100)
	Acid hot springs		
Sulfolobus acidocaldarius	Bacterium	CA	89 (192)
Bacillus acidocaldarius	Bacterium	OH	70 (158)
Cyanidium caldarium	Alga (red)	PA	56 (133)
Dactylaria gallopava	Fungus (deuteromycete)	OH	52 (126)
Euglena mutabilis	Alga (euglenoid)	PA	35–40 (95–104)
Zygogonium sp.	Alga (green)	PA	~30 (86)

*Autotrophs (A) obtain cell carbon through reduction of CO_2; heterotrophs (H) rely on organic molecules. Cellular energy may be derived from sunlight [photo- (P)], from oxidation of reduced inorganic molecules [chemo- (C)], or from oxidation of organic molecules [organo- (O)]. For example, photosynthetic plants are photoautotrophs (PA), and most animals are organoheterotrophs (OH). The designated activity may not be the only one possible for that organism.

as components of the mat community at temperatures representing their upper limit (Fig. 1; table). These distinctive thermal mats begin at 73°C (163°F) with *S. lividus*. Below 50°C (122°F) taxonomic and trophic diversity increases rapidly with decreasing temperatures. Protozoa, ostracods, and other invertebrates become components of the systems as herbivores at these lower temperatures. By the time the water has cooled to around 40°C (104°F), multispecies food chains and webs exist. *See* FOOD WEB.

The upstream and downstream distribution of a species may also depend on requirements for and tolerance of other abiotic factors (such as hydrogen sulfide concentration) and biotic factors (such as competition, grazing, and commensalism).

Mats below 40–50°C (104–122°F) often are thin (1–4 mm or 0.04–0.16 in.) and cutrose and show little or no subsurface stratification. Higher-temperature mats are thicker (3–100 mm or 0.12–3.9 in.) and gelatinous, and usually have distinct layers (**Fig. 2**). Species composition along the vertical aspect of gelatinous mats is determined by light and oxygen.

Light-dependent microorganisms are restricted to the upper regions of the gelatinous mats. Cyanobacterial photosynthesis oxygenates the upper layer diurnally and secretes photosynthate used for growth of aerobes and facultative anaerobes in this region. During nocturnal periods, low levels of atmospheric oxygen diffuse to the mat surface; the oxygen is consumed by microbial respiration, however, rendering the upper layers anoxic at night.

The undermat layer (Fig. 2), principally composed of *C. aurantiacus*, assimilates cyanobacterial photosynthate for photoheterotrophic growth during the day and for aerobic respiration and growth at night when it migrates toward the surface to utilize atmospheric oxygen. Alternation between surface growth of cyanobacteria and this nocturnal migration yields a lamellar mat structure similar in appearance to algal stromatolites. Cyanobacterial cells covered by migrations and continued surface growth become moribund and die, thereby contributing organic material for growth of anaerobes in the region below light penetration where anoxic conditions persist. Decomposition primarily occurs in the bottom layers.

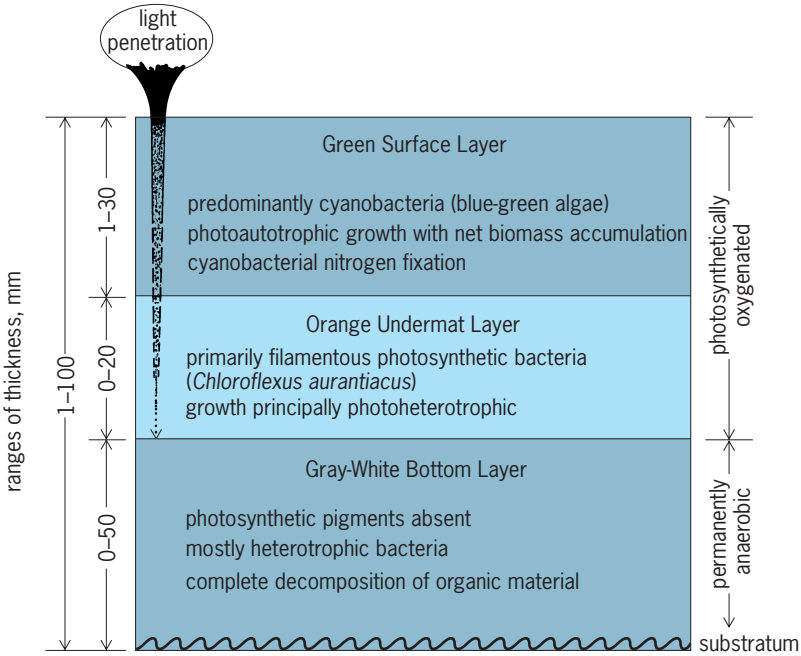

Fig. 2. Properties and distribution of activities within mature gelatinous mats along the vertical axis. The mats vary in thickness and composition with stream temperature. Above 70°C (158°F) only the surface layer is present; below about 45°C (113°F) the orange undermat area may be absent. Light reduction (due to self-shading) and extent of diffusion of oxygen from surface cyanobacterial photosynthesis depend on mat thickness. 1 mm = 0.04 in.

Cyanobacteria colonize barren thermal habitats; heterotrophic bacteria are limited by availability of dissolved organic compounds. Maximum cyanobacterial diversity is achieved within 3–4 weeks in thermal areas. Since the dominants of young mats are similar to those of mature mats, few successional patterns for microbial species are recognized. In some streams, *Chloroflexus* precedes *Synechococcus* and provides an attachment surface for the cyanobacterium; in low-nitrogen streams, nitrogen-fixing species (such as *Mastigocladus laminosus* and *Calothrix* spp.) are most active during early successional stages.

Cyanobacterial biomass accumulates rapidly, and within 2 weeks stable subsurface habitats are available to anaerobes. Some springs are as productive as eutrophic lakes. More than 97% of the community's biomass is prokaryotic. The standing crop biomass of mature mats ranges from 200 to 500 g/m^2 with peak values above 50°C (122°F). The undisturbed mats accumulate biomass and thicken until decomposition equals production. Grazers, such as the ostracod *Potamocypris*, may significantly decrease the standing crop, while others, such as ephydrid fly larvae, may stimulate productivity through nutrient turnover.

Adaptations of thermal biota. Many hot-spring organisms remain within their temperature range along the thermal gradient by movement. When suddenly exposed to high temperature shifts, insect larvae and the ostracod *Potamocypris* sp. drift downstream in the current to cooler temperatures. The ostracod returns upstream by positive thermotaxis until encountering its maximum (50°C or 122°F), when it reverses direction (thermophobic response). *Thermacarus nevadensis* moves upstream by positive rheotaxis. *Oscillatoria terebriformis* maintains its position along thermal and light-intensity gradients by thermo- and phototactic movements, respectively.

The temperature optimum for metabolic activity of nonmotile microorganisms usually is at or near the habitat temperature. This physiological adaptation ensures maximal activity where the organism is growing, and is known for cyanobacterial photosynthesis and nitrogenase activity and for incorporation of organic molecules by bacteria. This is even true for carbon dioxide fixation by the chemoautotrophs at boiling temperatures. Cyanobacterial cells at the mat surface are adapted to high light intensities, while those within the mat are adapted to low light.

Physiological adaptations have led to the selection of genetic strains in some species. Four stenothermal strains (ecotypes) of *S. lividus* with distinct cardinal (maximum, minimum, optimum) temperatures for growth comprise sequentially the upper *Synechococcus* layer from 50 to 73°C (122 to 163°F). Light-intensity ecotypes of the cyanobacterium *Plectonema notatum* also exist. In summer, the high-intensity strains dominate the mat, while low-intensity strains persist in subsurface layers; the latter strains dominate the mat during winter. The life cycles of some invertebrates are genetically adapted to their thermal niche. *Paracoenia bisetosa* adults oviposit on mat surfaces surrounded by lethal-temperature water. The larvae decimate mat structure through feeding activities, but prior to the reentry of hot water many of the adults emerge from pupae, and the remaining larvae and pupae are killed.

Distribution and dispersal of species. The microflora of thermal habitats is widespread within geographic regions, while the fauna is more restricted. Thermal-spring species seem as sensitive to unfavorable conditions encountered during dispersal as their mesophilic counterparts; additionally, thermophilic species must survive extended exposure to suboptimal temperatures.

Acid hot springs. Acid hot springs are less common and less well studied than the alkaline variety. Their main characteristics are described below.

Physicochemical properties. Most of these springs are acidic due to abiotic subterranean or surface biotic oxidation of hydrogen sulfide to sulfuric acid. Consequently, sulfate levels are high (often around 1 g/liter). Concentrations of metallic ions of aluminum, copper, iron, manganese, magnesium, and zinc are higher than in alkaline springs, as is ammonium (NH_4^+).

Biota. Compared to the alkaline streams, acid hot springs have less species diversity. No microorganisms live in boiling acid source waters, and cyanobacteria do not grow at pH < 4. *Cyanidium caldarium*, a eukaryote with red algal characteristics, forms emerald green mats (1–20 mm thick) from 35 to 55°C (95 to 131°F). These mats are not stratified, but filaments of the fungus *Dactylaria gallopava* and rod-shaped bacteria (such as *Bacillus* spp. and *Thiobacillus* spp.) are interspersed within them. These heterotrophs rely on *Cyanidium* photosynthate for growth.

Few microorganisms inhabit acid thermal streams above 55°C (131°F). *Bacillus acidocaldarius* occurs up to 65°C (149°F), and *Sulfolobus acidocaldarius*, a chemoautotroph, is found at temperatures up to 89°C (192°F). At less than 40°C (104°F) other photosynthetic organisms such as *Chlorella* sp. and *Euglena mutabilis* occur. A purple mat of *Zygogonium* sp. is present around 30°C (86°F). This mat comprises a community including the above algae, herbivorous brine flies, *Ephydra thermophila*, and carnivorous midges, *Bezzia setulosa*. *See* ECOLOGY.

Conrad Wickstrom

Bibliography. T. D. Brock, *Thermophilic Microorganisms and Life at High Temperatures*, 1978; C. Edwards (ed.), *Microbiology of Extreme Environments*, 1990; R. A. Herbert and G. A. Codd (eds.), *Microbes in Extreme Environments*, 1986; J. K. Kristjansson, *Thermophilic Bacteria*, 1991.

Thermal expansion

Solids, liquids, and gases all exhibit dimensional changes for changes in temperature while pressure is held constant. The molecular mechanisms at work

and the methods of data presentation are quite different for the three cases and are therefore discussed separately in this article.

Expansion of solids. The temperature coefficient of linear expansion α_l is defined by Eq. (1), where l is

$$\alpha_l = \frac{1}{l}\left(\frac{\partial l}{\partial t}\right)_{p=\text{const}} \tag{1}$$

the length of the specimen, t is the temperature, and p is the pressure. For each solid there is a Debye characteristic temperature Θ, below which α_l is strongly dependent upon temperature and above which α_l is practically constant. Many common substances are near or above Θ at room temperature and follow approximate equation (2), where l_0 is the length at 32°F

$$l = l_0(1 + \alpha_l t) \tag{2}$$

(0°C) and t is the difference between the temperature in °F and 32°F (the temperature in °C). The total change in length from absolute zero to the melting point has a range of approximately 2% for most substances. Typical room temperature values of a_l are given in **Table 1**.

Linear, harmonic vibration of the atoms in a solid cannot account for changes in volume, hence this must result from nonlinearity of the thermally excited vibration. The theory of E. Grüneisen takes this into account and shows the coefficient of expansion to be proportional to the constant-volume specific heat of the solid. At low temperatures (small amplitude vibration), the coefficient of expansion approaches zero.

Pure crystals may have different values of α_l along different axes, but substances such as structural steel have many crystals randomly oriented and are almost free from this effect. At certain temperatures, crystalline substances may change in lattice arrangement, and a sudden change of volume occurs at constant temperature, making α_l momentarily infinity. *See* LATTICE VIBRATIONS; SPECIFIC HEAT OF SOLIDS; THERMOCOUPLE.

TABLE 1. Temperature coefficients of linear expansion for typical substances at room temperature

Substance	Coefficient of linear expansion per °F (°C) × 10^6
Aluminum, commercial	13 (24)
Copper	9 (17)
Diamond	0.6 (1)
Glass, commercial	6 (11)
Glass, Pyrex	2 (3)
Granite	4.6 (8.3)
Ice	28 (50)
Iron	7 (12)
Invar alloy	0.5 (0.9)
Quartz, crystalline	3 (5)
Quartz, fused	0.3 (0.5)
Oak, along fiber	3 (5)
Oak, across fiber	30 (54)
Rubber, hard	44 (80)

Expansion of gases. So-called perfect gases follow the relation in Eq. (3), where p is absolute pressure, v

$$\frac{pv}{T} = \frac{R}{\text{molecular weight}} \tag{3}$$

is specific volume, T is absolute temperature, and R is a constant. The magnitude of R, the so-called gas constant, is 1544 ft-lb/(°R)(lb-mole) in the English system, or 8.3144 joules/(K)(g-mole) in the metric system. Real gases often follow this equation closely; for example, **Table 2** shows values of R at atmospheric pressure and 0°C. *See* GAS CONSTANT.

TABLE 2. Values for gas constant R at atmospheric pressure* and 32°F (0°C)

Gas	R
Air	1545
Hydrogen	1546
Nitrogen	1543
Oxygen	1544
Methane	1539

*1 atm = 10^2 kPa.

The coefficient of cubic expansion α_v is defined by Eq. (4), and for a perfect gas this is found to be $1/T$.

$$\alpha v = \frac{1}{v}\left(\frac{\partial v}{\partial t}\right)_{p=\text{const}} \tag{4}$$

The behavior of real gases is largely accounted for by van der Waals' equation (5), where $a$and b are

$$p = \frac{RT}{v - b} - \frac{a}{v^2} \tag{5}$$

constant for a given gas. When the specific volume is large, the effects of these constants are unimportant, and the real gas behaves as a perfect gas. In the regions where a and b have a dominant effect it is usually found desirable to use experimentally determined graphs or charts of properties. *See* GAS; KINETIC THEORY OF MATTER.

Expansion of liquids. For liquids, α_v is somewhat a function of pressure but is largely determined by temperature. Though α_v may often be taken as constant over a sizable range of temperature (as in the liquid expansion thermometer), generally some variation must be accounted for. For example, water contracts with temperature rise from 32 to 39°F (0 to 4°C), above which it expands at an increasing rate, as shown by the data in **Table 3**, which were taken at atmospheric pressure. One approach to this variation is to evaluate the constants α, β, and γ in Eq. (6),

$$v = v_0(1 + \alpha t + \beta t^2 + \gamma t^3) \tag{6}$$

where v_0 is the volume at 0°C (32°F), and v is the volume at temperature t in °C. Typical values of the coefficients appear in **Table 4**.

TABLE 3. Behavior of water at different temperatures

t, °F (°C)	Volume expansion, in.^{3}/oz (ml/g)
14 (−10)	1.73321 (1.00186)
32 (0)	1.73022 (1.00013)
39 (4)	1.72999 (1.00000)
50 (10)	1.73046 (1.00027)
212 (100)	1.742 (1.007)

TABLE 4. Coefficients of volume expansion of gases

Liquid	$\alpha \times 10^3$	$\beta \times 10^6$	$\gamma \times 10^8$
Ethyl alcohol (99.3% by volume)	1.012	2.20	
500 atm*	0.866		
3000 atm*	0.524		
Carbon tetrachloride	1.184	0.899	1.351
Mercury	1.182	0.0078	
Petroleum	0.8994	1.396	
Water	−0.06427	8.5053	−6.7900

*1 atm = 10^2 kPa.

Thermal stresses. When a homogeneous body is subject to constant boundary loads and is raised uniformly in temperature, the stress pattern in it will not change unless its elastic properties change. In general, stresses arise if (1) the body is made up of substances having different coefficients of expansion, (2) changes of boundary dimensions are restrained, or (3) temperature distribution is not uniform. A simple example of the first case is shown in the **illustration**, where the aluminum bar, if heated, would tend to expand faster than the iron bars, thereby putting the iron in tension and the aluminum in compression. Considering the aluminum alone, its change of length would be restrained, and therefore stresses would arise in it. If the aluminum bar were replaced by an iron one and if this one alone were heated, again the other bars would be in tension and the center one in compression. More complex stress patterns may arise in continuous bodies; for example, if the bars were joined along the sides rather than at the ends, shear stresses would arise in the seams. In iron, 360 lb/in.^{2} (2.48 megapascals) tensile stress would produce the same elongation as would a temperature rise of 1.8°F (1°C).

Since one source of temperature variation is the gradient necessary for heat transfer, thermal conductivity and heat capacity may both play a role in determining the stress pattern. *See* CONDUCTION (HEAT); HEAT CAPACITY; THERMAL CONDUCTION IN SOLIDS; THERMAL STRESS; THERMOMETER. Ralph A. Burton

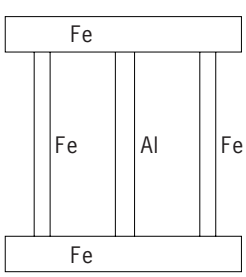

Body composed of substances having different coefficients of expansion. Thermal stresses would arise if it were subjected to heat.

Thermal hysteresis

A phenomenon in which a physical quantity depends not only on the temperature but also on the preceding thermal history. It is usual to compare the behavior of the physical quantity while heating and the behavior while cooling through the same temperature range. The **illustration** shows the thermal hysteresis which has been observed in the behavior of the dielectric constant of single crystals of barium titanate. On heating, the dielectric constant was observed to follow the path *ABCD* and so on, cooling the path *DCEFG*. *See* FERROELECTRICS; PERMITTIVITY.

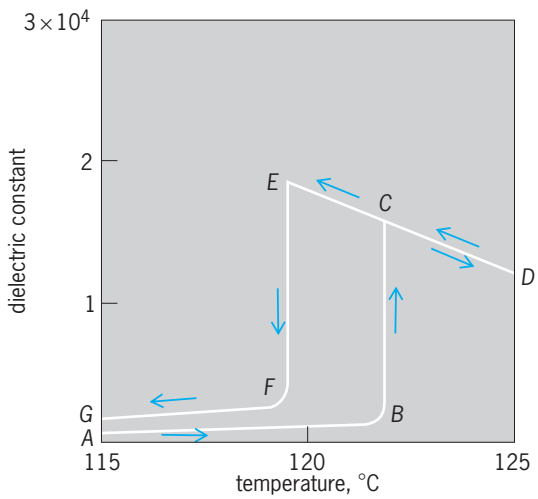

Plot of dielectric constant versus temperature for a single crystal of barium titanate. °F = (°C × 1.8) + 32. (*After M. E. Drougard and D. R. Young, Phys. Rev., 95:1152–1153, 1954*)

Perhaps the most common example of thermal hysteresis involves a phase change such as solidification from the liquid phase. In many cases these liquids can be dramatically supercooled. Elaborate precautions to eliminate impurities and outside disturbances can be instrumental in supercooling 110 to 150°F (60 to 80°C). On raising the temperature after freezing, however, the system follows a completely different path, with melting coming at the prescribed temperature for the phase change.

Solidification (or appropriate phase change) occurs by a nucleation mechanism, while melting (the reversal of phase) does not. The viable nucleus only forms below the melting temperature, being held up by the competing demands of surface and bulk free energy. Once formed, the nucleus grows rapidly to bring on the sudden precipitation of the solid phase. This rapid release of latent heat and the catastrophic nature of the solidification process illustrate the basic irreversibility

and hysteresis of thermally induced phase transformations. *See* CRYSTAL GROWTH; NUCLEATION; PHASE TRANSITIONS. H. B. Huntington; R. K. MacCrone

Thermal neutrons

Neutrons whose energy distribution is governed primarily by the kinetic energy distribution of molecules of the material in which the neutrons are found.

The molecules of the material usually have a kinetic energy distribution very close to a Maxwell-Boltzmann distribution. This distribution shows a peak at an energy equal to half the product of the temperature and the Boltzmann constant. At high energies it decreases exponentially, and at low energies it is proportional to the square root of the energy. When the material is large and very weakly absorbing, the neutron energy distribution closely approaches this maxwellian. *See* BOLTZMANN STATISTICS; KINETIC THEORY OF MATTER.

Deviations occur for the following reasons:

1. In nuclear reactors, neutrons are born at high energies and moderated by collisions with the material. These neutrons add a high-energy tail to the neutron energy distribution, which is inversely proportional to the neutron energy. Particularly when absorption is strong, the tail reaches down in energy, resulting in an increase in the peak energy of the neutron distribution relative to that in the moderator. It is common to describe this change in the peak by attributing a higher temperature to the neutrons. *See* REACTOR PHYSICS.

2. If an absorbing atom in the system has strong resonance absorption in the thermal energy range, it can create a depression in the neutron distribution around the peak resonant energy. *See* NEUTRON SPECTROMETRY.

3. If the loss of thermal neutrons is dominated by leakage rather than absorption, the neutron spectrum at low energies peaks at a lower temperature than otherwise. This phenomenon, called diffusion cooling, actually lowers the neutron temperature relative to that of the material.

The most common way of generating thermal neutrons is to allow neutrons from a source—reactor, accelerator, or spontaneous fission neutron emitter—to diffuse outward through a large block or tank of very weakly absorbing moderator. Neutrons in such a thermal column have very small tails to their spectral distribution and a temperature very close to that of the thermal-column medium.

Neutrons are sometimes scattered from cryogenic moderating materials in order to maximize the number of very low energy neutrons available for experiment. Here, although the spectrum is not quite maxwellian, it is still characterized by a temperature corresponding to the spectral peak. *See* NEUTRON.

Bernard I. Spinrad

Bibliography. G. I. Bell and S. Glasstone, *Nuclear Reactor Theory*, 1970, reprint 1979; J. R. Lamarsh, *Introduction to Nuclear Reactor Theory*, 1966; M. M. R. Williams, *The Slowing Down and Thermalization of Neutrons*, 1966.

Thermal stress

Mechanical stress induced in a body when some or all of its parts are not free to expand or contract in response to changes in temperature. In most continuous bodies, thermal expansion or contraction cannot occur freely in all directions because of geometry, external constraints, or the existence of temperature gradients, and so stresses are produced. Such stresses caused by a temperature change are known as thermal stresses.

Structures subject to stress. Problems of thermal stress arise in many practical design problems, such as those encountered in the design of steam and gas turbines, diesel engines, jet engines, rocket motors, and nuclear reactors. The high aerodynamic heating rates associated with high-speed flight present even more severe thermal-stress problems for the design of spacecraft and missiles. These may be further complicated by transient heating phenomena, a complex geometry, and changes in the material properties at high temperature, as well as by the different materials employed as structure.

The study of the thermal-stress problem is not restricted to the classical case of finding the elastic thermal stresses produced by a given temperature distribution in a structure, but touches on all phases of structural design. One must also include in the analysis temperature distribution, elastic and inelastic deformation, and the behavior of materials at elevated temperature. In addition, thermal buckling, fatigue, shock and aeroelastic effects at high temperatures must be considered.

The method of solving these problems is to formulate the complete problem and then to simplify it by making assumptions based on the physical situation or on experimental data. From this simplified formulation, one obtains an analytic solution that identifies the basic parameters of the problem and then constructs charts showing how the stresses vary with these parameters. The solution is then refined by examining the simplifying assumptions and obtaining correction factors. In most cases, this procedure yields results sufficiently accurate for use in the design of structures for high-temperature applications.

Procedure at elevated temperatures. To design for thermal stress in aircraft, missiles, and spacecraft, it is necessary to know either the expected mission profiles or the flight histories of speed, altitude, angle of attack, and bank angle. From these predicted performance profiles for various vehicles, the most severe combinations of applied, thermal, and allowable stresses are selected to be used in the design.

At elevated temperatures, the designer is faced with two factors which tend to increase the weight of the structure. First, the strength and moduli of most materials decrease as temperature rises. Second, the thermal stresses may act in concert with the

applied stresses to induce a larger deflection than is estimated for the applied stresses alone. To obtain the minimum possible structural weight, both of these problems must be considered. For flights of short duration and with low heating rates, insulation and radiative cooling may help to keep the temperature of structural material low, and may reduce the effects of thermal stresses. For longer flights with low heating rates, insulation and cooling techniques can also minimize temperature in structure. For short-duration flights with high heating rates, such as those of ballistic missiles and low L/D (lift-to-drag ratio) spacecraft reentering the Earth's atmosphere, ablation cooling is the most effective means of reducing interior structural temperatures. *See* ATMOSPHERIC ENTRY.

A mechanical design that allows sufficient deflection for thermal expansion, such as a corrugated skin structure in spacecraft, can relieve some thermal stress, but maintaining structural stiffness is not always a simple problem. Corrugated webs, ribs, and clip attachments may relieve some thermal stresses and still maintain stiffness; however, expansion joints for the corrugated structure may be difficult to design. In addition, heavy spar caps or thick skin with integral stiffeners can produce high thermal stresses. One possible way to relieve thermal stress is to design structural members in the inelastic portion of the stress-strain relation to provide the deflection and to absorb the thermal expansion. This procedure involves the concept of strain design rather than stress design, but it is possible in many cases to design for the applied loads without regard to thermal stresses, and then to add the thermal strains without obtaining appreciable permanent set. In other cases, a design which allows for some permanent set may be feasible. However, if the structure is exposed to numerous temperature cycles, thermal fatigue must be considered in such designs. *See* STRESS AND STRAIN.

Allowable stresses. Combined thermal and applied stresses at elevated temperatures may occur in either the elastic or inelastic range of the stress-strain curve for structural materials. If they are in the inelastic range, the mechanical properties of the structural material must be determined for various temperatures. To design the structure to support the applied and thermal stresses, it is also necessary to know the allowable stresses of the material under various loading and temperature conditions. A honeycomb-sandwich structure can increase the strength of a material at elevated temperature for a given weight.

Creep. The creep rate of most materials increases substantially with temperature and stress. Materials used in aircraft and missile structures creep at elevated temperatures, so that the structure may undergo large deformations if the load is applied for long periods. To keep these deformations within permissible limits, it may be necessary to reduce the applied stresses on the structure. *See* CREEP (MATERIALS).

Thermal fatigue. At room temperature, fatigue limits the number of stress cycles a material can withstand before it ruptures. At high temperatures, thermal fatigue is caused by both the stress and the temperature cycles, and can result in rupture and deformation arising from creep. When a structure is restrained, it is possible at sufficiently high temperatures to produce thermal stresses in the inelastic region of the stress-strain curve that exceed yield stress and result in plastic flow or rupture.

Thermal shock. If a body is subjected to sudden heating or cooling, so that local large thermal stresses are induced, thermal shock may result. Such sudden heating may occur when a body enters the atmosphere from space. For a thin structural member, such as a plate or a shell, thermal stress can be either normal or bending or a combination of both. In general, if the temperature distributions are symmetrical with respect to axes of symmetry in the symmetrical cross section, the thermal stress induced will not be of the bending type. However, when the temperature of the structure is nonuniform, bending stresses may be induced by an unsymmetrical temperature distribution, by an unsymmetrical cross section, by the use of different materials in the structure, or by the variation of physical properties with respect to temperature. Brittle and ductile materials react differently to such thermal stresses. Because the thermal stress arises from the strain due to thermal expansion, brittle materials, which can endure little strain before rupture, may fail under the thermal shock. *See* AEROELASTICITY; SPACECRAFT STRUCTURE.

Shih-Yuan Chen

Bibliography. B. A. Boley and J. H. Weiner, *Theory of Thermal Stresses,* 1960, reprint 1997; D. Burgreen, *Elements of Thermal Stress Analysis*, 1971, reprint 1987; R. B. Hetnarski (ed.), *Thermal Stresses,* 3 vols., 1986, 1987, 1989.

Thermal wind

The difference in the geostrophic wind between two heights in the atmosphere over a given position on Earth. It approximates the variation of the actual winds with height for large-scale and slowly changing motions of the atmosphere. Such structure in the wind field is of fundamental importance to the description of the atmosphere and to processes causing its day-to-day changes. The thermal wind embodies a basic relationship between vertical fluctuations of the horizontal wind and horizontal temperature gradients in the atmosphere. This relationship arises from the combination of the geostrophic wind law, the hydrostatic equation, and the gas law.

The geostrophic wind law applies directly to steady, straight, and unaccelerated horizontal motion and is a good approximation for large-scale and slowly changing motions in the atmosphere. The law specifies that the wind direction is perpendicular to the horizontal pressure-gradient force with a direction such that lower pressure is to the left and higher pressure to the right in the Northern Hemisphere (the opposite sense applies in the Southern Hemisphere). Wind speed is proportional to the horizontal pressure-gradient force per unit mass (the speed

also depends on the Coriolis parameter, which varies only with latitude). This law, thus, relates horizontal motions to horizontal variations of pressure in the atmosphere.

The hydrostatic equation combined with the gas law relates the atmospheric pressure and temperature fields. The relationship is accurate for most atmospheric situations but not for small-scale and rapidly changing conditions such as in turbulence and thunderstorms. The equation gives the change of pressure in the vertical direction as a function of pressure and temperature. The key conclusion is that at a given level in the atmosphere the pressure change (decrease) with height is more rapid in cold air than in warm air.

The thermal wind relationship is obtained by combining the geostrophic wind law and hydrostatic equation. The geostrophic wind law provides that the vertical change in horizontal wind depends upon the vertical change in horizontal pressure-gradient force per unit mass. In turn, the hydrostatic equation shows that the vertical change in this pressure-gradient force depends primarily on the horizontal temperature gradient at the level in question. Thus, the vector of wind change with height is in a direction essentially perpendicular to the temperature gradient and parallel to the isotherms of temperature on a horizontal surface with the cold air to the left and warm air to the right in the Northern Hemisphere (the opposite sense applies in the Southern Hemisphere). The magnitude of the change is proportional to the temperature gradient. The thermal wind relationship explains the tendency for westerly winds to increase with height in the middle latitude troposphere, reflecting the general fact that tropospheric temperatures are cooler in the polar regions than in the tropics. *See* ATMOSPHERIC GENERAL CIRCULATION; CORIOLIS ACCELERATION; GEOSTROPHIC WIND; HYDROSTATICS; TROPOSPHERE.

David D. Houghton

Bibliography. D. G. Andrews, J. R. Holton, and C. B. Conway (eds.), *Middle Atmosphere Dynamics*, 1987; D. L. Hartmann, *Global Physical Climatology,* 1994; J. R. Holton, *An Introduction to Dynamic Meteorology*, 3d ed., 1992; D. D. Houghton (ed.), *Handbook of Applied Meteorology*, 1985; C. Riegel, *Fundamentals of Atmospheric Dynamics and Thermodynamics*, 1992; G. T. Trewartha and L. H. Horn, *Introduction to Climate*, 5th ed., 1980.

Thermionic emission

The emission of electrons into vacuum by a heated electronic conductor. In its broadest meaning, thermionic emission includes the emission of ions, but since this process is quite different from that normally understood by the term, it will not be discussed here. Thermionic emitters are used as cathodes in electron tubes and hence are of great technical and scientific importance. Although in principle all conductors are thermionic emitters, only a few materials satisfy the requirements set by practical applications. Of the metals, tungsten is an important practical thermionic emitter; in most electron tubes, however, the oxide-coated cathode is used to great advantage.

Richardson-Dushman equation. The thermionic emission of a material may be measured by using the material as the cathode in a vacuum tube and collecting the emitted electrons on a positive anode. If the anode is sufficiently positive relative to the cathode, space charge (a concentration of electrons near the cathode) can be avoided and all electrons emitted can be collected; the saturation thermionic current is then measured. Actually, the emission current increases slightly with increasing field strength at the cathode, and in order to obtain the true saturation current one should extrapolate to zero applied field. *See* SCHOTTKY EFFECT.

The emission current density J increases rapidly with increasing temperature; this is illustrated by the following approximate values for tungsten:

T (K)	1000	2000	2500	3000
J (A/cm^2)	10^{-15}	10^{-3}	0.3	15

The temperature dependence of J is given by Eq. (1), the Richardson-Dushman (or Richardson)

$$J = AT^2 e^{-(\phi/kT)} \tag{1}$$

equation. Here A is a constant, k is Boltzmann's constant ($=1.38 \times 10^{-23}$ joule per degree), and ϕ is the work function of the emitter. The work function has the dimensions of energy and is a few electronvolts for thermionic emitters.

The temperature dependence of J is essentially determined by the exponential factor, since its temperature dependence predominates strongly over that of the factor T^2. Both A and ϕ may be obtained

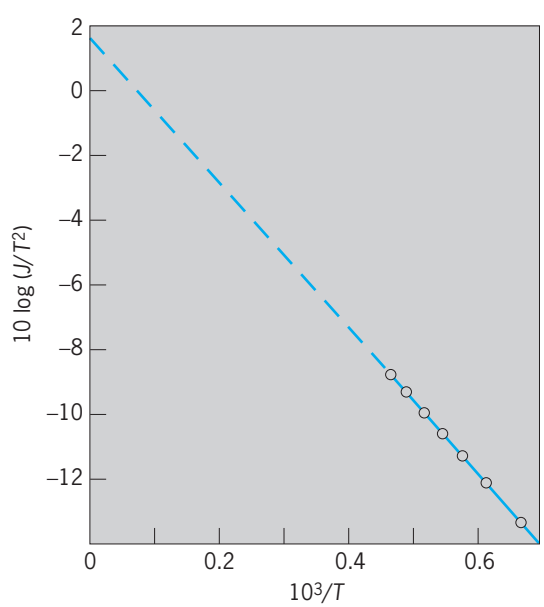

Fig. 1. Richardson-Dushman plot for tungsten, an important thermionic emitter. (*After G. Herrmann and S. Wagener, The Oxide-Coated Cathode, vol. 2, Chapman and Hall, 1951*)

experimentally by plotting the logarithm of J/T^2 versus $1/T$, as for tungsten in **Fig. 1**. The Richardson-Dushman formula can be derived for metals and semiconductors on the basis of relatively simple physical models.

Metals. According to quantum theory the electrons in a free atom occupy a set of discrete energy levels. When atoms are brought together to form a solid, these energy levels broaden into energy bands; the broadening is a result of the perturbing fields produced by neighboring atoms on the electrons and is most pronounced for the outer or valence electrons. In a metal, the perturbing influence on the valence electrons is so strong that they can no longer be associated with particular atoms but must be considered as moving freely throughout the crystal. These so-called free, or conduction, electrons are responsible for the high electrical and thermal conductivity of metals and also for the thermionic emission. *See* BAND THEORY OF SOLIDS; FREE-ELECTRON THEORY OF METALS.

The free electrons may be assumed to move in an approximately constant potential as indicated in **Fig. 2**. The bottom of the box corresponds to the energy of a conduction electron at rest in the metal; the "vacuum" level represents the energy of an electron at rest in free space. According to quantum mechanics, the electrons in this model can assume only particular states of motion which correspond to a set of very closely spaced energy levels. The probability for a given state to be occupied depends on the energy E of the state and on the absolute temperature T in accordance with Eq. (2), the so-called

$$F(E) = \frac{1}{1 + \exp\,[(E - E_F)/kT]} \qquad (2)$$

Fermi-Dirac distribution function. The quantity E_F is called the Fermi energy; it is determined by the number of electrons per unit volume in the metal and is of the order of a few electronvolts. Since kT at room temperature ($T = 300$ K) is only about 0. 025 eV, $E_F > kT$ for all temperatures below the melting point of metals. Note that for $T = 0$, $F(E) = 1$ for $E < E_F$, and $F(E) = 0$ for $E > E_F$. Hence, at absolute zero all energy levels up to E_F are occupied by electrons, whereas those above E_F are empty. For temperatures different from zero, some electrons have energies larger than E_F and the thermionic emission is due to those electrons in the "tail" of the Fermi distribution for which the energy lies above the vacuum level in Fig. 1. Note that when $E = E_F$, $F(E) = 0.5$; that is, the Fermi energy corresponds to those states for which the probability of being occupied is equal to 0.5.

When these ideas are put in a quantitative form, one arrives at the Richardson-Dushman equation with the specific value $A = 120$ A/cm^2 (if one takes into account reflection of electrons against the surface potential barrier, the theoretical value of A is <120 A/cm^2).

Experiments by M. N. Nichols in 1940 and by G. F. Smith in 1954 on single crystals of tungsten have shown that experimental values for A and ϕ depend

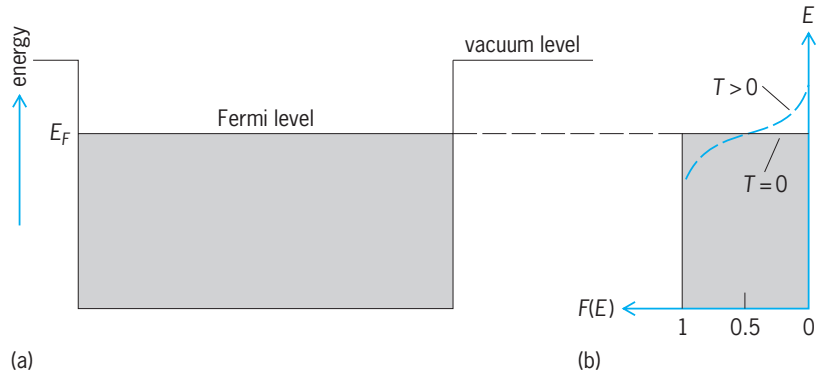

Fig. 2. Free electrons are assumed to move in approximately constant potential. (*a*) Occupation of electron states between the bottom of the conduction band and the Fermi level of a metal is indicated for $T = 0$ by the shaded area. (*b*) Fermi distribution function is represented schematically for $T = 0$ and for $T > 0$.

on the crystallographic plane from which the emission is measured; values for A (in A/cm^2) and ϕ (in electronvolts) for two crystallographic directions are given in the **table**. For polycrystalline metals, the experimental values for A and ϕ are thus average values for the particular specimen.

Experimental values for single crystals of tungsten

	Nichols		Smith	
Direction	A	ϕ	A	ϕ
(111)	35	4.39	52	4.38
(100)	117	4.56	105	4.52

Semiconductors. For semiconductors, the thermionic emission is also due to the escape of electrons which have energies above the vacuum level. The theory leads to the Richardson-Dushman formula, as it does for metals. The work function measures again the difference between the Fermi level of the semiconductor and the vacuum level. *See* SEMICONDUCTOR.

A. J. Dekker

Thermionic power generator

A device for converting heat into electricity through the use of thermionic emission and no working fluid other than electric charges. An elementary thermionic generator, or thermionic converter, consists of a hot metal surface (emitter) separated from a cooler electrode (collector) by an insulator seal (**Fig. 1**). The interelectrode gap is usually a fraction of a millimeter in width. The hermetic enclosure contains a small amount of an easily ionizable gas, such as cesium vapor maintained by a liquid-cesium reservoir. In some experimental devices, the enclosure may be evacuated.

Electrons evaporated from the emitter cross the interelectrode gap, condense on the collector, and are returned to the emitter via the external electrical load circuit. The thermionic generator is essentially a heat engine utilizing an electron gas as the

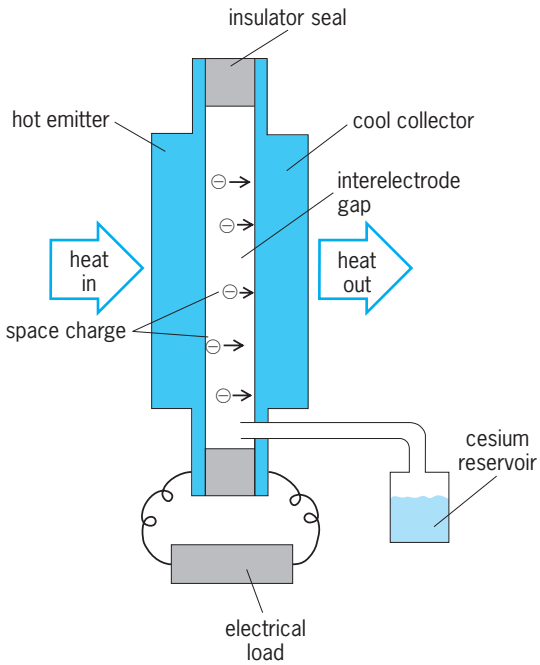

Fig. 1. Diagram of thermionic converter.

working fluid. The temperature difference between the emitter and the collector drives the electron current.

Thermionic generators are characterized by high operating temperatures, typically emitter temperatures between 1600 and 2500 K (2420 and 4040°F) and collector temperatures ranging from 800 to 1100 K (980 to 1520°F); low output voltage (approximately 0.5 V per converter); high current density (around 5–10 A/cm^2); and high conversion efficiency (about 10–15%). These characteristics, especially the relatively high heat-rejection temperature, make the thermionic generator attractive for producing electric power in space applications with nuclear-reactor or radioisotope energy sources. The high electrode temperatures make thermionic generators also attractive as topping units for steam power plants, and for the cogeneration of electricity in combination with heat for intermediate-temperature industrial processes. Topping units increase the overall system efficiency. *See* COGENERATION; NUCLEAR BATTERY.

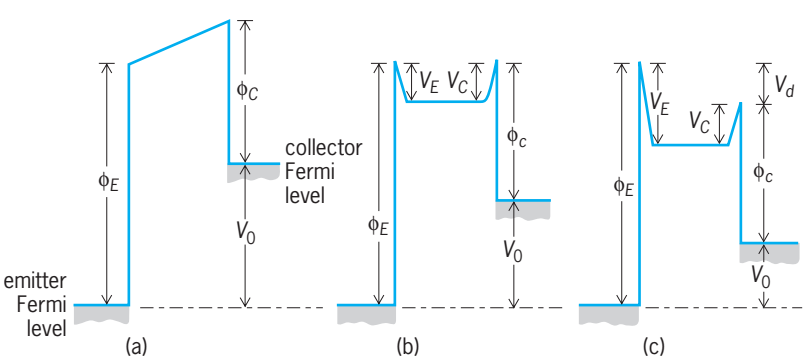

Fig. 2. Motive diagrams of a thermionic converter operating in (*a*) ideal, (*b*) unignited, and (*c*) ignited modes.They give the spatial variation of the energy of an electron from emitter to collector. Here, ϕ_E and ϕ_C are emitter and collector work functions, V_0 is the output voltage, V_E and V_C are voltages across sheath regions at emitter and collector, and V_d is the arc drop.

Vacuum devices. Because the energy of the electrons in the collector is greater than that in the emitter (**Fig. 2***a*), the collected electrons perform work as they flow back to the emitter through the electrical load. The output voltage V_0 is given by the difference in the Fermi levels between the emitter and the collector. The Fermi level is a characteristic energy of electrons in a material. It corresponds to the highest energy of the electrons in the material at zero temperature. The difference between the motive just outside a surface and the Fermi level is a property of the surface and is called the work function. The emitter work function ϕ_E is given by the difference of the motive of a point just outside the emitter and the emitter Fermi level. Likewise, the collector work function ϕ_C is given by the difference between the motive of a point just outside the collector and the collector Fermi level. *See* FREE-ELECTRON THEORY OF METALS; WORK FUNCTION (ELECTRONICS).

The electron emission from the emitter or collector electrodes is given by the Richardson equation (1), where J is the current density in amperes per square centimeter, ϕ is the work function of the electrode in electronvolts, and T is the electrode temperature in kelvins. *See* THERMIONIC EMISSION.

$$J = 120T^2 e^{-(11604\phi/T)} \tag{1}$$

The output power density of the converter is the product of V_0 times the load current density J that flows through the converter. Usually J is substantially less than the emitter current density J_E. In the ideal mode (Fig. 2*a*), J is given by Eq. (2) for $V_0 < \phi_E - \phi_C$;

$$J = J_E e^{-[11604(V_0+\phi_C-\phi_E)/T_E]} \tag{2}$$

for $V_0 > \phi_E - \phi_C$, $J = J_E$. In Eq. (2), T_E is the emitter temperature in kelvins, and it is assumed that the collector temperature T_C is low enough that back emission from the collector is negligible. The output voltage V_0 can be varied between zero and its open circuit value by changing the load resistance, giving rise to the ideal-mode current density-voltage characteristic (**Fig. 3**).

The ideal mode (Fig. 2*a*) corresponds to an evacuated interelectrode spacing of 0.01 mm or less. Small thermionic diodes with such spacings have been built. However, the difficulty of maintaining the necessary close spacing over large areas at high temperatures makes the vacuum diode mostly of academic interest.

Cesium-vapor devices. The problems of extremely close spacing can be circumvented by introducing low-pressure cesium vapor between the emitter and the collector. The cesium serves two functions. First, it adsorbs on the electrode surfaces and reduces their work functions. Each work function is reduced to a value that depends on the ratio of the electrode temperature and the cesium reservoir temperature. Second, the cesium supplies positive ions to neutralize

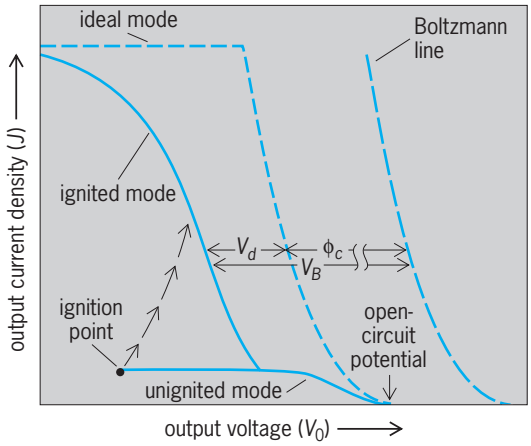

Fig. 3. Current density-voltage characteristics of a thermionic converter. V_B is the barrier index; other symbols are as in Fig. 2.

the electron space charge. The ions are produced by surface ionization and electron-impact ionization in the interelectrode gap. Surface ionization occurs when the cesium vapor contacts the hot emitter. The ionization efficiency is high when the emitter work function is comparable to the 3.89-eV ionization potential of cesium. The addition of barium vapor to the cesium vapor may lead to better converter performance.

Unignited mode. A thermionic converter in which the predominant source of cesium ions is surface ionization operates at low cesium pressure, less than about 10 pascals or 0.1 torr, and is said to operate in the unignited, lower, extinguished, or Knudsen mode (Figs. 2*b* and 3). Sheath regions at the emitter and collector couple the neutral plasma (which has essentially equal ion and electron concentrations) to the adjacent electrodes. Unfortunately, in the unignited mode, a converter provides useful power densities at emitter temperatures above 2500 K (4040°F), a region for which there are few practical energy sources.

Ignited mode. In order to obtain practical power densities at more moderate emitter temperatures, say 1600 to 2000 K (2420 to 3140°F), the cesium pressure must be increased to the order of 10^2 Pa (1 torr). Then cesium adsorption on the emitter reduces its work function so that the current density is greater than 5 A/cm^2, and the converter is said to operate in the ignited mode (Figs. 2*c* and 3). To obtain this mode, not only is the cesium pressure increased but the output voltage is reduced along the unignited-mode characteristic curve until the ignition point is reached and electron impact ionization in the interelectrode space results in a discharge which triggers the thermionic converter into the ignited mode. It is possible to operate back and forth along the ignited-mode curve. However, if the output potential is increased to the intersection with the unignited mode, the discharge is extinguished, and the ignited mode can only be reestablished by reducing the output voltage along the unignited-mode curve to the ignition point.

The electrons are scattered many times by the cesium as they cross to the collector, resulting in resistive and current losses in the plasma. There is an additional loss due to the energy required to ionize the cesium atoms. The sum of these losses is usually called the arc drop and is denoted V_d. It represents the loss in output voltage that must be supplied to the plasma in order to operate the converter in the ignited mode. Thus far, this mode has been the only practical means of operating a thermionic generator. For good performance, it is clear that V_d and ϕ_C must be small. *See* ELECTRICAL CONDUCTION IN GASES.

Barrier index. The barrier index V_B is defined by Eq. (3). This index serves as an inverse figure of merit of the thermionic converter performance; the lower the value of V_B, the higher the converter performance. Reductions in V_B can be translated into higher efficiency at a given temperature or a lower emitter temperature at a given efficiency. Present converters have V_B values between 1.9 and 2.2 eV. Operationally, V_B is defined as the minimum potential difference between the ignited-mode characteristic and the Boltzmann line. The Boltzmann line represents the ideal current density-voltage characteristic assuming zero arc drop, zero collector work function, and zero collector temperature.

$$V_B = V_d + \phi_C \tag{3}$$

Elias P. Gyftopoulos; George N. Hatsopoulos

Bibliography. S. W. Angrist, *Direct Energy Conversion,* 4th ed., 1982; F. G. Baksht et al., *Thermionic Converters and Low-Temperature Plasma*, 1978; R. Decher, *Direct Energy Conversion: Fundamentals of Electric Power Production*, 1997; G. N. Hatsopoulos and E. P. Gyftopoulos, *Thermionic Energy Conversion*, vol. 1, 1974, vol. 2, 1979; *Proceedings of the 11th Symposium on Space Power and Propulsion*, Albuquerque, 1994; N. S. Rasor, *Applied Atomic Collision Physics*, vol. 5, 1982.

Thermionic tube

An electron tube that relies upon thermally emitted electrons from a heated cathode for tube current. Thermionic emission of electrons means emission by heat. In practical form an electrode, called the cathode because it forms the negative electrode of the tube, is heated until it emits electrons. The cathode may be either a directly heated filament or an indirectly heated surface. With a filamentary cathode, heating current is passed through the wire, which either emits electrons directly or is covered with a material that readily emits electrons. Some typical filament structures are shown in **Fig. 1**. Filaments of tungsten or thoriated tungsten are commonly used in high-power transmitting tubes where their ruggedness and ability to withstand high voltages are essential. Oxide-coated filaments are used in a few small high-voltage rectifier tubes.

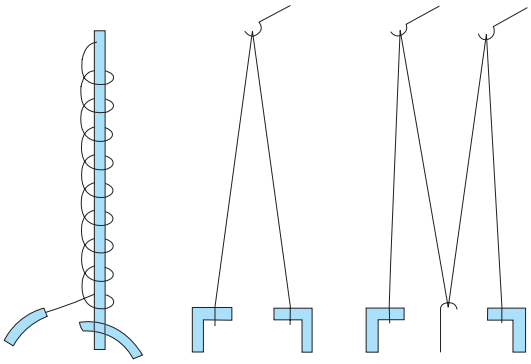

Fig. 1. Typical filamentary cathode structures for thermionic tubes.

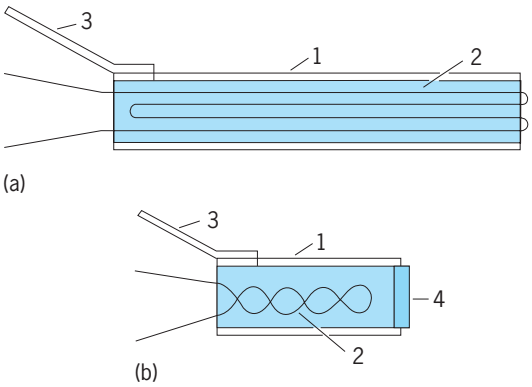

Fig. 2. Cathodes. (*a*) Receiving-tube cathode; 1, cathode sleeve, oxide-coated on exterior; 2, folded heater, insulated with refractory oxide; 3, cathode tab, for electrical connection. (*b*) Kinescope cathode: 1, cathode sleeve; 2, heater, insulated with refractory oxide; 3, cathode tab, for electrical connection; 4, emitting "button," oxide-coated on right surface.

Indirectly heated cathodes have a filament, commonly called the heater, located within the cathode electrode to bring the surface of the cathode to emitting temperature. Some common forms are shown in **Fig. 2**. They are usually coated with barium-strontium oxide, on the periphery in receiving tubes and on the end in kinescopes. Because the emitting surface carries no heating current, there is no voltage drop along the surface. Hence such cathodes are usually known as equipotential cathodes. The high emission capability, the equipotential surface, and the favorable geometry of these cathodes make possible the close-spaced tube structures that lead to the high transconductances required in modern applications. Hence, oxide-coated equipotential cathodes are used in almost all receiving and medium-power transmitting tubes. They are also used in some high-power pulsed transmitting tubes, where the remarkable ability of the oxide cathode to emit very high current densities (tens of amperes per square centimeter, for microsecond periods at low repetition rates) is exploited. The majority of all vacuum tubes are thermionic tubes. It is possible to make so-called cold-cathode tubes, but they tend to be unstable in vacuum and find their main application in gas tubes, not vacuum tubes. *See* ELECTRON TUBE; GAS TUBE; THERMIONIC EMISSION; VACUUM TUBE.

Leon S. Nergaard

Thermistor

An electrical resistor with a relatively large temperature coefficient of resistance. Thermistors are useful for measuring temperature and gas flow or wind velocity. Often they are employed as bolometer elements to measure radio-frequency, microwave, and optical power. They also are used as electrical circuit components for temperature compensation, voltage regulation, circuit protection, time delay, and volume control. A common type of thermistor is a semiconducting ceramic composed of a mixture of several metal oxides. Metal electrodes or wires are attached to the ceramic material so that the thermistor resistance can be measured conveniently. The temperature coefficient of resistance is negative for these thermistors. Other types can have either negative or positive temperature coefficients. *See* ELECTRICAL RESISTIVITY; VOLTAGE REGULATOR; VOLUME CONTROL SYSTEMS.

At room temperature the resistance of a thermistor may typically change by several percent for a variation of 1°C of temperature, but the resistance does not change linearly with temperature. The temperature coefficient of resistance of a thermistor is approximately equal to a constant divided by the square of the temperature in kelvins. The constant is equal to several thousand kelvins and is specified for a given thermistor and the temperature range of intended use.

The electrical and thermal properties of a thermistor depend upon the material composition, the physical dimensions, and the environment provided by the thermistor enclosure. Thermistors range in form from small beads and flakes less than 25 micrometers (10^{-3} in.) thick to disks, rods, and washers with centimeter dimensions. The small beads are often coated with glass to prevent changes in composition or encased in glass probes or cartridges to prevent damage. Beads are available with room-temperature resistances ranging from less than 100 Ω to tens of megohms, and with time constants that can be less than a second. Large disks and washers have a similar resistance range and can have time constants of minutes. *See* TIME CONSTANT.

Temperature measurement. An instrument can be made to measure temperature by connecting a battery, a thermistor, and a current meter in series. The current indicated by the meter is determined by the thermistor resistance and therefore is an indication of the temperature. A thermistor across the terminals of an ohmmeter or a Wheatstone bridge will also provide an indication of the temperature. By careful design, a thermistor can be used to measure temperature changes of less than 10^{-3} °C. *See* TEMPERATURE MEASUREMENT; THERMOMETER.

Velocity or flow measurement. Wind velocity or gas flow can be measured by increasing the current in

the thermistor of a temperature-measuring instrument so that the thermistor temperature is considerably higher than the ambient temperature. The thermistor temperature then will depend on how fast the thermistor can dissipate the heat and therefore how fast the air is flowing over the thermistor surface. If the ambient temperature is varying, a second flow meter with the thermistor shielded from air currents can be used to determine the effect of temperature variations alone. *See* FLOW MEASUREMENT; VELOCIMETER.

Level measurements. The same type of instrument used for measuring gas flow can be used to sense liquid level. The thermistor is placed above the surface of the liquid. If the liquid level rises to contact the thermistor, heat will be dissipated rapidly from the thermistor and the thermistor temperature will decrease, giving a decrease in the electric current. With great care, thermistors can be made sufficiently sensitive to detect objects very close to, but not touching, the thermistor. *See* LEVEL MEASUREMENT.

Radio-frequency power measurement. One of the most accurate techniques for measuring radio-frequency and microwave power is to use thermistors to measure a substituted change in direct-current power. Initially, direct-current power is dissipated in the thermistor, and the amount of power needed to achieve a specified resistance, and therefore temperature, is recorded. The thermistor is then used to dissipate both radio-frequency and direct-current power. In order to achieve the same resistance, less direct-current power is needed, and this decrease in direct-current power is approximately equal to the radio-frequency power dissipated by the thermistor. The technique is very effective because direct-current power can be measured very accurately.

The direct-current portion of the circuit is a bridge (see **illustration**). The radio-frequency signal is incident from the left side of the illustration through a coaxial transmission line with a characteristic impedance of R_0. Capacitors are used to isolate the radio-frequency and direct-current portions of the circuit except for two identical thermistors, which are part of both portions. The two thermistors in series constitute one arm of the bridge circuit. The operational amplifier provides feedback that keeps the voltage difference between its input terminals (+ and − in the illustration) very small. This is accomplished when the resistor ratios on the left and right are identical. Since the two resistors on the right have the same value R_1, the combined thermistor resistance will be made equal to $4R_0$. Thus, the individual thermistors have a resistance of $2R_0$. The thermistors appear in parallel with the radio-frequency circuit, and their combined resistance matches the characteristic impedance of the transmission line, which maximizes the absorbed radio-frequency power. The voltage V across the thermistor pair is recorded by the voltmeter without any radio-frequency input. The radio-frequency signal is then turned on and the voltage recorded again. The power dissipated in the thermistor pair is $V^2/4R_0$. The difference in power dissipated by the thermistor pair in the two cases is calculated, and this is approximately equal to the radio-frequency power. *See* BRIDGE CIRCUIT; TRANSMISSION LINES.

There are small differences between the radio-frequency power and the change in direct-current power due to radio-frequency losses in the transmission line, differences in the current distributions, and reactances in the circuit. This difference is characterized by an effective efficiency that is measured with other techniques. Coaxial thermistor sensors, also known as bolometer mounts, typically have effective efficiencies of greater than 95%. Similar devices using single thermistors are also available with rectangular waveguides. *See* BOLOMETER; WAVEGUIDE.

For radio-frequency measurements, because the thermistor resistance is held at a fixed value, its nonlinear characteristics do not affect the measurement, and the radio-frequency power measurement is very linear over most of its operating range, which is generally up to about 10 milliwatts. A slight nonlinearity does arise at higher power levels because the two thermistors are not exactly the same. The bridge will keep their series resistance equal to $4R_0$, but their parallel resistance will no longer

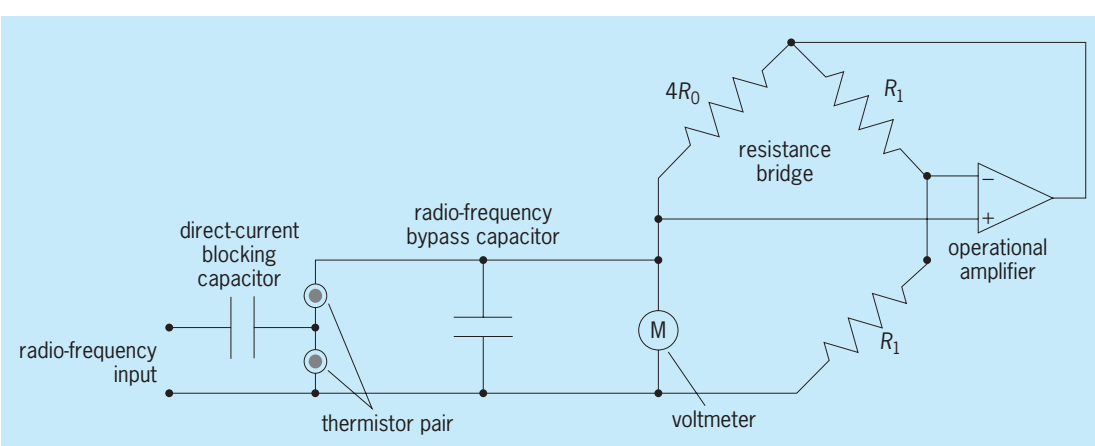

Bridge circuit for measuring radio-frequency power by use of a pair of thermistors.

match the transmission line impedance of R_0. *See* ELECTRIC POWER MEASUREMENT; MICROWAVE POWER MEASUREMENT. Thomas P. Crowley

Bibliography. A. Fantom, *Radio Frequency and Microwave Power Measurement*, Peter Peregrinus, London, 1990; C. J. Kaiser, *The Resistor Handbook*, 2d ed., Saddleman Press, 1994; E. D. Macklen, *Thermistors*, Electrochemical Publications, 1979; F. Zandman, P.-R. Simon, and J. Szwarc, *Resistor Theory and Technology*, Vishay Intertechnology, Malvern, PA, 2001.

Thermoacoustics

The study of phenomena that involve both thermodynamics and acoustics. A sound wave in a gas is usually regarded as consisting of coupled pressure and displacement oscillations, but temperature oscillations accompany the pressure oscillations. When there are spatial gradients in the oscillating temperature, oscillating heat flow also occurs. The combination of these four oscillations produces a rich variety of thermoacoustic effects. *See* ACOUSTICS; OSCILLATION; SOUND; THERMODYNAMIC PRINCIPLES.

Although the oscillating heat transfer at solid boundaries does contribute significantly to the dissipation of sound in enclosures such as buildings, thermoacoustic effects are usually too small to be obviously noticeable in everyday life. For example, the amplitude of the temperature oscillation in conversational levels of sound is only about 0.0001°C (0.0002°F). However, thermoacoustic effects in intense sound waves inside suitable cavities can be harnessed to produce extremely powerful pulsating combustion, thermoacoustic refrigerators, and thermoacoustic engines. Gas compressibility and inertia, heat transport, and thermal expansion and contraction are important in all of these systems.

Pulsating combustion. Oscillations can occur whenever combustion takes place in a cavity. In industrial equipment and residential appliances, these oscillations are sometimes encouraged in order to stir or pump the combustion ingredients, while in rocket engines such oscillations must usually be suppressed because they can damage the rocket structure. The oscillations occur spontaneously if the combustion progresses more rapidly or efficiently during the compression phase of the pressure oscillation than during the rarefaction (expansion) phase—the Rayleigh criterion. *See* COMBUSTION; GAS DYNAMICS.

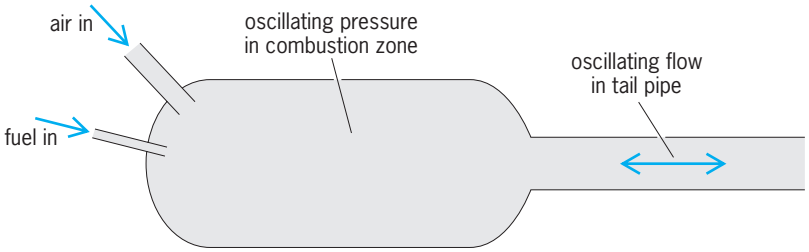

Fig. 1. Acoustic resonance in a pulse combustor. The mass of gas in the tail pipe bounces against the "spring" of compressibility in the combustion zone, forming a resonant oscillator. The oscillations are driven by periodic, pressure-enhanced combustion.

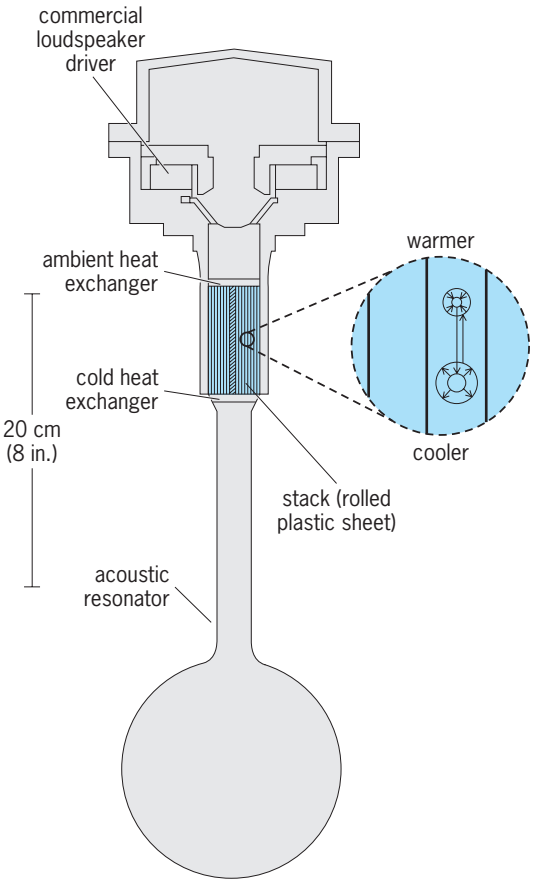

Fig. 2. An early standing-wave thermoacoustic refrigerator that cooled to −60°C (−76°F). Heat is carried up the temperature gradient in the stack. At the right is a magnified view of the oscillating motion of a typical parcel of gas. The volume of the parcel depends on its pressure and temperature. (*After T. J. Hofler, Thermoacoustic Refrigerator Design and Performance, Ph.D. thesis, University of California at San Diego, 1996*)

The geometry of the cavity determines the oscillation frequency, just as the length of an organ pipe determines its pitch (**Fig. 1**). Stored resonance energy shifts back and forth between kinetic energy of moving gas in the tail pipe and compressive energy of pressurized gas in the combustion zone, much as the energy in a pendulum shifts back and forth between kinetic energy and gravitational potential energy. At one time of the oscillation, the air is motionless and the combustion zone is pressurized above atmospheric pressure. Combustion is enhanced at this time, and all the resonance energy is stored in the compressibility of the pressurized air. A quarter cycle later, after the pressure has accelerated the air in the tail pipe rightward, the air velocity in the tail pipe reaches its largest (rightward) value when the pressure in the combustion zone reaches its ambient value, so all the resonance energy is stored kinetically in the moving air. Next, inertia of the moving air keeps it moving rightward, pulling air out of the combustion zone and lowering the pressure there, so that the air soon comes to rest with the combustion

zone below ambient pressure, shifting the resonance energy back to compressibility and suppressing the combustion. This low pressure then pulls air back into the tail pipe, so that after another quarter cycle the air is moving leftward most rapidly, storing kinetic energy, and the combustion zone has been refilled to ambient pressure. Finally, inertia keeps the air moving leftward, pressurizing the combustion zone, so that the system comes to rest again at the starting condition. *See* INERTIA.

Thermoacoustic refrigerators. Thermoacoustic refrigerators use acoustic power to pump heat from a low temperature to ambient temperature (**Fig. 2**). The heat-pumping mechanism takes place in the pores of a structure called a stack. As a typical parcel of the gas oscillates along a pore, it experiences changes in temperature. Most of the temperature change comes from adiabatic compression and expansion of the gas by the sound pressure, and the rest is a consequence of the local temperature of the solid wall of the pore. A thermodynamic cycle results from the coupled pressure, temperature, position, and heat oscillations. At the topmost position of the parcel, it dumps heat to the pore wall, since the parcel temperature was raised above the local pore temperature by adiabatic compression by the acoustic wave. Similarly, at its lowermost position, the parcel absorbs heat from the pore, since the parcel temperature is below the local stack temperature. Thus, each parcel moves a little heat a small distance from the bottom toward the top along the stack, a small distance up the temperature gradient, during each cycle of the acoustic wave. The overall effect, much as in a bucket brigade, is the net transport of heat from the cold heat exchanger to room temperature. *See* ADIABATIC PROCESS; SOUND PRESSURE; THERMODYNAMIC CYCLE; THERMODYNAMIC PROCESSES.

The extrema in pressure and gas motion are in phase in this thermoacoustic refrigerator and in the pulsating combustion discussed above. A wave with such time phasing is called a standing wave.

Thermoacoustic engines. While standing-wave thermoacoustic systems have matured only recently, Stirling engines and refrigerators have a long, rich history. Recently, new insights have resulted from applying thermoacoustics to Stirling systems, treating them as traveling-wave thermoacoustic systems in which the extrema in pressure and gas motion are approximately 90° out of phase in time. In the thermoacoustic-Stirling engine, the thermodynamic cycle is accomplished in a traveling-wave acoustic network, and acoustic power is produced from heat with an efficiency of 30% (**Fig. 3**). The wave circulating around the torus causes the gas in the regenerator to experience oscillating pressure and oscillating motion, with traveling-wave time phasing. Excellent thermal contact between the typical parcel of gas and the wall of its pore ensures that the gas temperature always matches the local solid temperature. Hence, downward motion of the parcel causes thermal expansion, and upward motion causes thermal contraction. The time phasing between this motion and the oscillating pressure ensures that the thermal expansion occurs at high pressure and the thermal contraction occurs at low pressure, so that the parcel does net work during each cycle. Since this

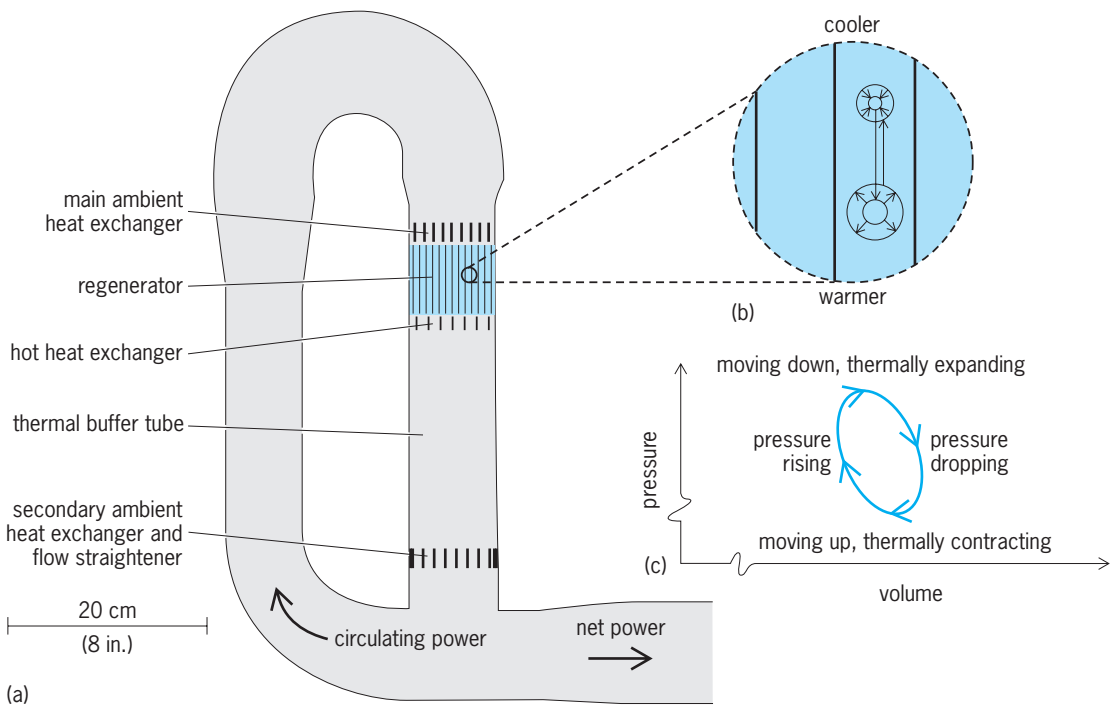

Fig. 3. Thermoacoustic-Stirling hybrid engine. (*a*) A traveling wave runs clockwise around the toroidal loop. Its acoustic power is amplified in the regenerator, so that net power up to 1 kW (1.3 horsepower) can be extracted at the tee and delivered to a load. (*b*) Magnified view of the oscillating motion of a typical parcel of gas. (*c*) Thermodynamic cycle of the typical parcel of gas. (*After S. Backhaus and G. Swift, Thermoacoustic-Stirling heat engine: Detailed study, J. Acous. Soc. Amer., 107:3148–3166, 2000*)

net work is produced at the acoustic frequency, it increases the acoustic power of the wave passing through the regenerator, so net power is available for extraction elsewhere. *See* STIRLING ENGINE; TORUS; WAVE (PHYSICS).

Current developments. Typical engines and refrigerators have crankshaft-coupled pistons or rotating turbines, whereas thermoacoustic systems often have no moving parts or flexing moving parts (as in a loudspeaker), with no sliding seals. Such systems appear attractive because of their elegance, reliability, and low cost. Some product-oriented research and development is under way in pulsating combustion and thermoacoustics, and many academic groups are studying nonlinear behavior at a fundamental level. *See* NONLINEAR ACOUSTICS. Greg Swift

Bibliography. S. Backhaus and G. Swift, Thermoacoustic-Stirling heat engine: Detailed study, *J. Acous. Soc. Amer.*, 107:3148–3166, 2000; T. J. Hofler, *Thermoacoustic Refrigerator Design and Performance*, Ph.D. thesis, University of California at San Diego, 1986; L. E. Kinsler et al., *Fundamentals of Acoustics*, 4th ed., 1999; A. J. Organ, *Thermodynamics and Gas Dynamics of the Stirling Cycle Machine*, 1992; J. W. S. Rayleigh, *The Theory of Sound*, vol.2, 1976; G. Walker, *Cryocoolers*, 1983.

Thermochemistry

A branch of physical chemistry concerned with the absorption or evolution of heat that accompanies chemical reactions. Closely related topics are the latent heat associated with a change in phase (crystal, liquid, gas), the chemical composition of reacting systems at equilibrium, and the electrical potentials of galvanic cells. Thermodynamics provides the link among these phenomena.

A knowledge of such heat effects is important to the chemical engineer for the design and operation of chemical reactors, the determination of the heating values of fuels, the design and operation of refrigerators, the selection of heat storage systems, and the assessment of chemical hazards. Thermochemical information is used by the physiologist and biochemist to study the energetics of living organisms and to determine the calorific values of foods. Thermochemical data give the chemist an insight to the energies of, and interactions among, molecules. *See* CHEMICAL EQUILIBRIUM; CHEMICAL REACTOR; PHASE EQUILIBRIUM; POTENTIALS; REFRIGERATOR.

Thermodynamic principles. The first law of thermodynamics expresses the principle of conservation of energy. When a closed system changes from an initial state to a final state, its internal energy, U, changes by the amount shown in Eq. (1), where q is the heat

$$U(\text{final}) - U(\text{initial}) = \Delta U = q + w \qquad (1)$$

energy transferred to the system from the outside and w is the work done on the system by external forces. A positive sign of q, w, or ΔU means energy is transferred to the system, while a negative sign means energy is removed. Some authors use the opposite sign convention for w. *See* CONSERVATION OF ENERGY; THERMODYNAMIC PRINCIPLES.

Internal energy. Internal energy (U) is a variable of state. This means that its value depends only on the state of the system and not on its previous history. The value of ΔU, the change in internal energy, depends only on the initial and final states. Such states are identified by chemical composition, physical phase, temperature, pressure, and sometimes other relevant variables. The values of q and w, however, depend both on the states and on the way the transformation is brought about. Their sum must always satisfy Eq. (1). The first law governs any system and any combination of states. If a chemical reaction occurs during the transformation, the initial and final states will have different compositions. Then q is called the heat of reaction.

If the change in states takes place with no work done, then $w = 0$ and $\Delta U = q$. The most common example is a process which takes place irreversibly at constant volume. However, in the laboratory, chemical reactions are usually conducted at constant pressure. In this case, $w = -P\Delta V$, where P is the pressure and ΔV is the change in volume for the process. Under such conditions, the value for ΔU is given by Eq. (2).

$$\Delta U = q - P\Delta V \qquad (2)$$

See INTERNAL ENERGY.

Enthalpy. The property H is called enthalpy. It is defined in general as $H = U + PV$. When Eq. (2) is solved for the term q, the result gives the value for ΔH [Eq. (3)], the change in enthalpy.

$$q = \Delta U + P\Delta V = \Delta H \qquad (3)$$

See ENTHALPY.

If the change in states is brought about reversibly, then w is algebraically a minimum and q is a maximum. In this case, $q = T\Delta S$, where ΔS is the change of entropy of the system and T is the temperature. Heats of reaction are seldom measured directly under such conditions, however.

If q is positive for the irreversible process (energy transferred to the system) the reaction is called endothermic, and if q is negative (heat given off) the reaction is exothermic.

Change in enthalpy (ΔH) and change in internal energy (ΔU) for a chemical reaction are reported as a certain quantity of energy for the number of moles indicated in the balanced chemical equation. Equation (3) provides the relationship between the two quantities. If all reactants and products are liquids or solids, the difference is negligible (except at very high pressures). If gases are involved, the difference is significant only when the number of moles of product gases differs from the number of moles of reactant gases.

A chemical transformation may take place in a series of steps, each corresponding to a certain reaction. Thus, for example, at 1 bar (10^5 pascals)

and 25°C (77°F), the reaction steps shown as (*4a–d*) took place. (The subscript *r* is the symbol for chemical reaction.) The net result of all of these steps is shown in reaction (*4e*).

$$CH_4(gas) \rightarrow C(graphite) + 2H_2(gas) \quad (4a)$$

$$\Delta_r H = 74.52 \Delta_r U = 72.04$$

$$C(graphite) + {}^1/_2O_2(gas) \rightarrow CO(gas) \quad (4b)$$

$$\Delta_r H = -110.525 \Delta_r U = -111.764$$

$$CO(gas) + {}^1/_2O_2(gas) \rightarrow CO_2(gas) \quad (4c)$$

$$\Delta_r H = -282.984 \quad \Delta_r U = -281.744$$

$$2H_2(gas) + O_2(gas) \rightarrow 2H_2O(liquid) \quad (4d)$$

$$\Delta_r H = -571.66 \quad \Delta_r U = -564.22$$

$$CH_4(gas) + 2O_2(gas) \rightarrow CO_2(gas) + 2H_2O(liquid) \quad (4e)$$

$$\Delta_r H = -890.64 \quad \Delta_r U = -885.69$$

The values of $\Delta_r H$ and $\Delta_r U$ are in kilojoules for the reaction specified. Since both U and H are variables of state, values of ΔU and ΔH are sums of the corresponding values for the intermediate steps for any path which leads from the initial to the final state. This result is called Hess's law, but it is simply a consequence of the first law of thermodynamics.

The change of any property, symbolized by X, associated with a chemical reaction may be equated to a sum of terms, one for each reactant and each product, by Eq. (5), where the $X(i)$ are the correspond-

$$\Delta_r X = \Sigma \nu_i X(i) \quad (5)$$

ing properties of reactants and products, i; and ν_i are the coefficients in the balanced chemical equation (positive for products, negative for reactants). The enthalpy of formation, sometimes called heat of formation, of a compound is the change in enthalpy for a reaction in which the compound is synthesized from its component elements. These quantities are seldom measured directly but are calculated from enthalpies of other reactions through the application of Hess's law. The enthalpy change for any reaction can be calculated from the enthalpies of formation of the reactants and products by the substitution of $\Delta_f H$ for X in the right side of Eq. (5) [the subscript f is the symbol for formation from elements]. *See* ENTHALPY.

Heats of reaction. Enthalpies of reactions may be obtained from several types of measurement. Three classes are recognized. The first-law heat of reaction is measured directly in a calorimeter. The second-law heat of reaction is calculated from the effect of temperature change on the equilibrium constant by the use of the Van't Hoff equation. The third-law heat of reaction is calculated by $\Delta_r H = \Delta_r G + T\Delta_r S$. $\Delta_r G$ is the change in Gibbs energy for the reaction calculated from an equilibrium constant or electrical cell potential, and $\Delta_r S$ is the corresponding change in entropy calculated from measured heat capacities and the third law of thermodynamics. In principle, all three methods give the same result for any particular reaction. The ease of measurement and the attainable accuracy may vary for different situations.

Enthalpies of reactions are slowly changing functions of temperature. The enthalpies of a reaction at the two temperatures T_1 and T_2 may be related to the difference in enthalpies for each component of the reaction between the two temperatures by Eq. (6).

$$\Delta_r H(T_2) = \Delta_r H(T_1) + \Delta_r[H(i, T_2) - H(i, T_1)] \quad (6)$$

The quantities in the last term of Eq. (6) may be calculated from the corresponding heat capacities by Eq. (7).

$$H(i,T_2) - H(i,T_1) = \int_{T_1}^{T_2} C_p(i)\, dT \quad (7)$$

See CHEMICAL THERMODYNAMICS; ENTROPY; FREE ENERGY.

Calorimetric measurements. A calorimeter is an instrument for measuring the heat added to or removed from a process. There are many designs, but the following parts can generally be identified: the vessel in which the process is confined, the thermometer which measures its temperature, and the surrounding environment called the jacket. The heat associated with the process is calculated by Eq. (8), where

$$q = C[(T(\text{final}) - T(\text{initial})] - q_{ex} - w \quad (8)$$

T is the temperature. The quantity C, the energy equivalent of the calorimeter, is obtained from a separate calibration experiment. The work transferred to the process, w, is generally in the form of an electric current (as supplied to a heater, for example) or as mechanical work (as supplied to a stirrer, for example) and can be calculated from appropriate auxiliary measurements. The quantity q_{ex} is the heat exchanged between the container and its jacket during the experiment. It is calculated from the temperature gradients in the system and the measured thermal conductivities of its parts.

Two principal types of calorimeters are used to measure heats of chemical reactions. In a batch calorimeter, known quantities of reactants are placed in the vessel and the initial temperature is measured. The reaction is allowed to occur and then the final equilibrium temperature is measured. If necessary, the final contents are analyzed to determine the amount of reaction which occurred.

In a flow calorimeter, the reactants are directed to the reaction vessel in two or more steady streams. The reaction takes place quickly and the products emerge in a steady stream. The rate of heat production is calculated from the temperatures, flow velocities, and heat capacities of the incoming and outgoing streams, and the rates of work production and heat transfer to the jacket. Dividing this result by the rate of reaction gives the heat of reaction.

The combustion of a substance in oxygen is often studied in a specially designed reaction calorimeter. The heats of combustion of liquid or solid samples are usually measured in a batch-type calorimeter. The

vessel is a strong steel alloy bomb which is placed in a container of water fitted with appropriate thermometers and stirrers. The sample is placed in the bomb in an atmosphere of oxygen at high pressure (around 30 atm or 3 megapascals). It is then ignited by an electrical fuse. Heats of combustion of gases are usually measured in a type of flow calorimeter called a flame calorimeter. Heats of combustion can be measured to accuracies of 1 part in 10,000. The primary limiting factor in such measurements is the purity of the samples. Heats of combustion in fluorine have also been measured. *See* CALORIMETRY.

Units and symbols. The International Union of Pure and Applied Chemistry (IUPAC) Commission on Thermodynamics has recommended the general symbol $\Delta_\alpha^\beta X$ to represent the change of any property of a system when it changes from an initial state (denoted by α) to a final state (denoted by β). In addition, certain commonly observed processes are given special symbols of the type $\Delta_b X$, where the subscript b represents the process. Other symbols are vap, vaporization of a liquid; sub, sublimation of a solid; fus, fusion of a solid; mix, mixing without reaction; r, chemical reaction; f, formation from elements; and c, combustion. The subscript m placed after the property symbol indicates 1 mole of substance. Additional specifications can be placed within parentheses. For example, $\Delta_{\text{vap}}H_m(H_2O, 298.15\text{ K})$ symbolizes the heat of vaporization of 1 mole of water at 298.15 K, $\Delta_r G(1000\text{ K})$ the change in Gibbs energy for a chemical reaction at 1000 K, and $\Delta_c H_m(C_4H_{10}, \text{g}, 300\text{ K})$ the enthalpy of combustion of 1 mole of butane gas at 300 K.

Thermochemical quantities are usually reported and tabulated for substances in their standard states. The standard state of a solid is the thermodynamically stable crystal, of a liquid the liquid, and of a gas the hypothetical ideal gas, all at unit pressure. For the past century the pressure unit has been the atmosphere. It has been suggested that the bar is more suitable for this role as it is more compatible with the International System (SI) of units. Standard states can be defined for any temperature, but 25°C (298.15 K) has been traditional. In Customary units, this standard state is given at 77°F (536.67°R). The ideal gas is a hypothetical state, but its properties can readily be calculated from the equation of state of the real gas. The internal energy and enthalpy of an ideal gas are independent of its pressure.

The standard state for the solvent in a solution is the pure liquid. The standard states for the solutes are hypothetical ideal solutions at unit concentrations. Concentrations are usually expressed as molalities or mole fractions. Properties of real solutions can be related to those of the hypothetical ideal solutions by appropriate auxiliary data. Standard states for individual ions in solution are defined with the help of the additional conventions that the enthalpy and Gibbs energies of formation of the hydrogen ion are zero. A degree symbol (°) designates a property of a standard state. The standard-state concept promotes compactness and explicitness for the tabulation of data.

In the past, thermochemical quantities usually have been given in units of calories. A calorie is defined as the amount of heat needed to raise the temperature of 1 gram of water 1°C. However, since this depends on the initial temperature of the water, various calories have been defined, for example, the 15° calorie, the 20° calorie, and the mean calorie (average from 0 to 100°C). In addition, a number of dry calories have been defined. Those still used are the thermochemical calorie (exactly 4.184 joules) and the International Steam Table calorie (exactly 4.1868 J).

Thermochemical quantities have also been reported in terms of British thermal units (Btu). This unit is the amount of heat required to raise the temperature of 1 lb of water 1°F. A proliferation of Btu's similar to that for calories has occurred. The Btu in common use is the International Steam Table Btu (1055.056 J).

The SI rules do not recognize either the calorie or the Btu. The energy unit is the absolute joule (J). Most modern literature uses this unit. *See* PHYSICAL MEASUREMENT.

Sources of data. Original reports of measured values of thermochemical quantities are widely scattered among the world's scientific literature. A number of compilations of enthalpy of formation $\Delta_f H^\circ$, Gibbs energy of formation $\Delta_f G^\circ$, absolute entropy S°, and heat capacity at constant pressure C_p° of pure compounds at 298.15 K have been published during the past century. Some of them contain data at other temperatures and values for mixtures and ions as well.

A unique example is the International Critical Tables which appeared as a series of seven volumes between 1926 and 1930. The series was the result of an international cooperation among scientists to collect all reliable physical and chemical properties of materials available at that time. Volume V contains thermochemical data. More recent compilations have been made for inorganic compounds. The Landolt-Börnstein Tables, which have undergone a series of revisions since the 1890s, contain extensive thermodynamic data. Other compilations which are regularly updated by supplements and revisions are the JANAF Tables for low-molecular-weight inorganic compounds and the Thermodynamics Research Center publications for organic and some nonmetallic inorganic compounds.

A consequence of Hess's law is that thermochemical values such as heats of formation, combustion, reaction, and phase transition at a fixed temperature are interrelated through a system of linear algebraic equations. Hundreds, or even thousands, of such equations are available for even limited sets of compounds. They usually form an overdetermined set. The compiler has the job of selecting values of heats of formation which best fit the experimental data with consideration of the assigned uncertainties. If data at different temperatures and second- and third-law heats of reactions are included, the system of equations becomes nonlinear. Formerly the selection was made manually by a series of iterations; more recently computer programs have been written to help

in the task of data management and equation solving. The whole process must be repeated to incorporate new data.

To promote internal consistency among thermochemical compilations, a division of the Committee on Data for Science and Technology (CODATA) has recommended certain values of key properties. These represent a basic starting point for most other compilations.

Heating values. The heating value (also called calorific value) of a fuel is the heat of combustion (with a positive sign) of a certain quantity of fuel expressed in some units when burned under given conditions. The price for wholesale commodity transfer is based on the heating value, rather than mass or volume. Engineers use heating values to carry out heat balance calculations for furnaces, engines, and chemical processes.

Many specific definitions of heating value have been issued by trade and standards organizations around the world. An organization may give different definitions for solid, liquid, and gaseous fuels. Some of these have found their way into long-term sales contracts and legal systems of many countries. Therefore the term heating value does not have a universally recognized quantitative meaning.

In many English-speaking countries, heating values have been reported in Btu per pound at 15°F. For a gross heating value the water produced by the combustion is assumed to be liquid. For a net heating value it is assumed to be a gas. A precise definition requires additional specifications such as the nature of other products (for example, those formed from nitrogen or sulfur if present), the amount of water in the fuel (dry or wet basis), and whether the fuel is burned in air or oxygen. If the fuel is a gas, the state, real or ideal, must be indicated. If it is in a real state, the pressure must also be specified.

Commercial calorimeters are available for the measurement of heating values, but accurate measurements are difficult to make and require skilled technicians. Many fuels are complicated mixtures whose composition is not completely known. If the composition is known, the heating value can be calculated from the standard-state enthalpies of combustion of the pure components. This may require a large amount of auxiliary data such as heat capacities, heats of mixing, equation of state of pure and mixed systems, and values of unit conversions.

Organizations concerned with definitions of heating values are the U.S. National Bureau of Standards and its counterpart in other countries, American Society for Testing and Materials, Gas Processors Association, International Standards Organization, and Groupe International des Importateurs de Gaz Natural Liquifie (GIIGNAL). *See* FOSSIL FUEL

. Randolph C. Wilhoit

Bibliography. P. W. Atkins, *Physical Chemistry*, 6th ed., 1998; H. Brodowsky and H. J. Schaller, *Thermochemistry of Alloys*, 1989; *Bulletin of Chemical Thermodynamics*, Thermochemistry Inc., Stillwater, Oklahoma; J. D. Cox, D. P. Wagman, and V. A. Mevdev (eds.), *CODATA Key Values for Thermodynamics*, 1989; M. Frenkel (ed.), *Thermochemistry and Equilibria of Organic Compounds*, 1993; *JANAF Thermochemical Tables*, Dow Chemical Co., accumulated supplements; I. Mills et al., *Quantities, Units and Symbols in Physical Chemistry*, 2d ed., 1993; *Physical and Thermodynamic Properties of Pure Compounds*, TRC Hydrocarbon Project, and *Selected Values of Properties of Chemical Compounds*, TRC Data Project, Thermodynamics Research Center, Texas A&M University, semiannual supplements; A report of IUPAC Commission 1.2 on Thermodynamics, *J. Chem. Thermodyn.*, 14:805–815, 1982.

Thermocouple

A device in which the temperature difference between the ends of a pair of dissimilar metal wires is deduced from a measurement of the difference in the thermoelectric potentials developed along the wires. The presence of a temperature gradient in a metal or alloy leads to an electric potential gradient being set up along the temperature gradient. This thermoelectric potential gradient is proportional to the temperature gradient and varies from metal to metal. It is the fact that the thermoelectric emf is different in different metals and alloys for the same temperature gradient that allows the effect to be used for the measurement of temperature.

Circuit. The basic circuit of a thermocouple is shown in the **illustration**. The thermocouple wires, made of different metals or alloys A and B, are joined together at one end H, called the hot (or measuring) junction, at a temperature T_1. The other ends, C_A and C_B (the cold or reference junctions), are maintained at a constant reference temperature T_0, usually but not necessarily 32°F (0°C). From the cold junctions, wires, usually of copper, lead to a voltmeter V at room temperature T_r. Due to the thermoelectric potential gradients being different along the wires A and B, there exists a potential difference between C_A and C_B. This can be measured by the voltmeter, provided that C_A and C_B are at the same temperature and that the lead wires between C_A and V and C_B and V are identical (or that V is at the temperature T_0, which is unusual).

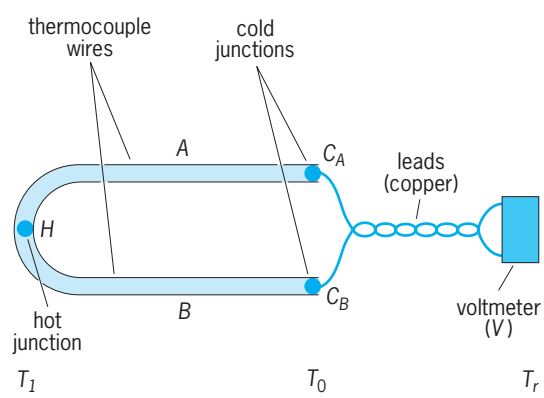

Basic circuit of a thermocouple.

Such a thermocouple will produce a thermoelectric emf between C_A and C_B which depends only upon the temperature difference $T_1 - T_0$. If, however, the wires *A* or *B* are not physically or chemically homogeneous, this is no longer the case and the thermoelectric emf will also depend upon the shape of the temperature profile along the wires between *H* and *C*. Herein lies the principal limitation in the use of thermocouples for accurate temperature measurement. It is impossible, in practice, to obtain perfectly homogeneous wires and, moreover, as soon as one end is heated whatever homogeneity had previously existed is degraded to some extent. These problems limit the accuracy with which temperature measurements can be made by means of even the very best thermocouples to about 0.9°F (0.5°C) in the range from 32 to 1800°F (0 to 1000°C).

The thermoelectric emf in a thermocouple is developed in the temperature gradient and is in no way a junction phenomenon. Indeed, to avoid the effects of inhomogeneities near the junctions it is always advisable to try to arrange for both the hot and cold junctions to be in regions of uniform temperature. Under these conditions the way in which the junctions are made, whether it be by soldering, by welding, or simply by twisting the wires together, is of no consequence. *See* THERMOELECTRICITY.

Types. A large number of pure metal and alloy combinations have been studied as thermocouples, and the seven most widely used are listed in the **table**. The letter designations were originally introduced by the Instrument Society of America and have now gained worldwide acceptance.

The thermocouples in the table together cover the temperature range from about −420°F (−250°C or 20 K) to about 3300°F (1800°C). The most accurate and reproducible are the platinum/rhodium thermocouples, types R and S, while the most widely used industrial thermocouples are probably types K, T, and E. Each has its own special advantages and applications. For very low temperatures, below −420°F (−250°C or 20 K), thermocouples made from copper-nickel and gold-iron alloys have been developed. These make use of the Kondo effect and give useful thermoelectric emf's down to 1 K (−458°F or −272°C). For temperatures above the maximum range of type B platinum/rhodium thermocouples (about 3300°F or 1800°C), recourse has to be made to thermocouples using wires of alloys of tungsten and rhenium. These can be used up to 4900°F (2700°C) but must be very carefully protected from oxidation. For industrial applications these and many other thermocouples are manufactured in the so-called mineral-insulated (MI) version. The thermocouple wire, the refractory metal oxide insulator (magnesium oxide or beryllium oxide, for example), and the hermetically sealed metal sheath are assembled and sealed during manufacture. Such mineral-insulated thermocouples are thus to a very large extent protected from chemical contamination and mechanical damage. *See* KONDO EFFECT; LOW-TEMPERATURE THERMOMETRY.

Letter designations and compositions for standardized thermocouples*

Type designation	Materials
B	Platinum-30% rhodium/platinum-6% rhodium
E	Nickel-chromium alloy/a copper-nickel alloy
J	Iron/another slightly different copper-nickel alloy
K	Nickel-chromium alloy/nickel-aluminum alloy
R	Platinum-13% rhodium/platinum
S	Platinum-10% rhodium/platinum
T	Copper/a copper-nickel alloy

*After T. J. Quinn, *Temperature*, Academic Press, 1983.

Calibration. For all of the thermocouples listed in the table, international agreement has been obtained on standard reference tables of emf versus temperature. These are given in the International Electrotechnic Commission Standard, IEC 584, and are identical to those given in ASTM E-230-77 and many other national standards. Calibration of a thermocouple is best carried out by measuring the difference between the thermoelectric emf given by the thermocouple and that predicted by the standard reference table. Having measured this difference at a small number of fixed points, a complete calibration table is easily obtained by interpolating these differences from the standard reference table between the measured fixed points.

Pressure and magnetic field corrections. Due to the detailed mechanism of the thermoelectric effects, the magnitude of the thermoelectric emf given by a thermocouple is also affected by pressure and by the presence of magnetic fields. At very high pressures, for example, a type S thermocouple at 1800°F (1000°C) and at a pressure of 4 gigapascals (approximately 40 kilobars) is subject to a correction of about 36°F (20°C). For use in magnetic fields of 8 teslas and above, type E thermocouples are recommended as being one of the least affected combinations.

Extension and compensating wires. In many large industrial applications, the hot and cold junctions are widely separated, but the measured thermoelectric emf may be almost entirely developed in the first several feet (few meters) of wire. The remaining length of wire serves mainly to transmit the emf to the measuring system. The thermoelectric properties of this long length of wire near room temperature, or at least at temperatures below 212°F (100°C), are very much less critical than those of that part of the wire in the steep temperature gradient. Considerable economies can, therefore, be gained by using, in this less critical section, not the high-specification thermocouple wire, but another, cheaper wire whose thermoelectric properties are a reasonable match over the temperature range, say from 68 to 212°F (20 to 100°C). If these wires are made to the same nominal composition of the thermocouple wire but not to the same high specification, they are called extension wires. If, on the other hand, they are of a different

composition but one chosen to match the thermoelectric properties over this restricted temperature range, they are known as compensating wires. *See* TEMPERATURE MEASUREMENT. T. J. Quinn

Bibliography. T. W. Kerlin, *Practical Thermocouple Thermometry*, 1999; T. D. McGee, *Principles and Methods of Temperature Measurement*, 1988; R. M. Parks (ed.), *Manual on the Use of Thermocouples in Temperature Measurement*, American Society for Testing and Measurement, MNL 12, 1993; D. D. Pollock, *Thermocouples: Theory and Practice*, 1991; T. J. Quinn, *Temperature*, 2d ed., 1991.

Thermodynamic cycle

A procedure or arrangement in which one form of energy, such as heat at an elevated temperature from combustion of a fuel, is in part converted to another form, such as mechanical energy on a shaft, and the remainder is rejected to a lower-temperature sink as low-grade heat.

Common features of cycles. A thermodynamic cycle requires, in addition to the supply of incoming energy, (1) a working substance, usually a gas or vapor; (2) a mechanism in which the processes or phases can be carried through sequentially; and (3) a thermodynamic sink to which the residual heat can be rejected. The cycle itself is a repetitive series of operations.

There is a basic pattern of processes common to power-producing cycles. There is a compression process wherein the working substance undergoes an increase in pressure and therefore density. There is an addition of thermal energy from a source such as a fossil fuel, a fissile fuel, or solar radiation. There is an expansion process during which work is done by the system on the surroundings. There is a rejection process where thermal energy is transferred to the surroundings. The algebraic sum of the energy additions and abstractions is such that some of the thermal energy is converted into mechanical work. *See* HEAT.

A steam cycle that embraces a boiler, a prime mover, a condenser, and a feed pump is typical of the cyclic arrangement in which the thermodynamic fluid, steam, is used over and over again. An alternative procedure, after the net work flows from the system, is to employ a change of mass within the system boundaries, the spent working substance being replaced by a fresh charge that is ready to repeat the cyclic events. The automotive engine and the gas turbine illustrate this arrangement of the cyclic processes, called an open cycle because new mass enters the system boundaries and the spent exhaust leaves it.

The basic processes of the cycle, either open or closed, are heat addition, heat rejection, expansion, and compression. These processes are always present in a cycle even though there may be differences in working substance, the individual processes, pressure ranges, temperature ranges, mechanisms, and heat transfer arrangements.

Air-standard cycle. It is convenient to study the various power cycles by using an ideal system such as the air-standard cycle. This is an ideal, frictionless mechanism enveloping the system, with a permanent unit charge of air behaving in accordance with the perfect gas relationships.

The unit air charge is assumed to have an initial state at the start of the cycle to be analyzed. Each process is assumed to be perfectly reversible, and all effects between the system and the surroundings are described as either a heat transfer or a mechanical work term. At the end of a series of processes, the state of the system is the same as it was initially. Because no chemical changes take place within the system, the same unit air charge is conceivably capable of going through the cyclic processes repeatedly.

Whereas this air-standard cycle is an idealization of an actual cycle, it provides an amenable method for the introductory evaluation of any power cycle. Its analysis defines the upper limits of performance toward which the actual cycle performance may approach. It defines trends, if not absolute values, for both ideal and actual cycles. The air-standard cycle can be used to examine such cycles as the Carnot and those applicable to the automobile engine, the diesel engine, the gas turbine, and the jet engine.

Cyclic standards. Many cyclic arrangements, using various combinations of phases but all seeking to convert heat into work, have been proposed by many investigators whose names are attached to their proposals, for example, the Diesel, Otto, Rankine, Brayton, Stirling, Ericsson, and Atkinson cycles (see **illus.**). All proposals are not equally efficient in the conversion of heat into work. However, they may offer other advantages which have led to their practical development for various applications. Nevertheless, there is one overriding limitation on efficiency. It is set by the dictates of the Carnot cycle, which states that no thermodynamic cycle can be projected whose thermal efficiency exceeds that of the Carnot cycle between specified temperature levels for the heat source and the heat sink. Many cycles may approach and even equal this limit, but none can exceed it. This is the uniqueness of the Carnot principle and is basic to the second law

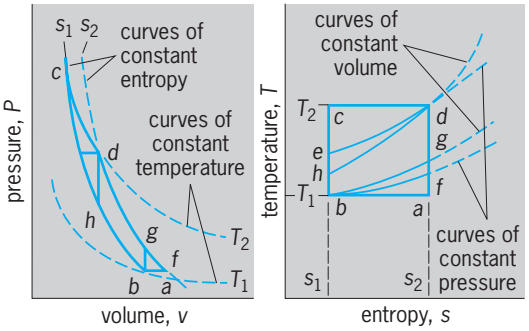

Comparison of principal thermodynamic cycles. Cycles are, in the order of decreasing efficiency, Carnot cycle (*a-b-c-d-a*), Brayton cycle (*b-e-d-f-b*), Diesel cycle (*b-e-d-g-b*), Otto cycle (*b-h-d-g-b*).

of thermodynamics on the conversion of heat into work. *See* BRAYTON CYCLE; CARNOT CYCLE; DIESEL CYCLE; OTTO CYCLE; STIRLING ENGINE; THERMODYNAMIC PROCESSES. Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; R. T. Balmer, *Thermodynamics*, 1989; I. Granet, *Thermodynamics and Heat Power*, 5th ed., 1995; W. C. Reynolds and H. C. Perkins, *Engineering Thermodynamics*, 2d ed., 1977; M. W. Zemansky, *Heat and Thermodynamics*, 6th ed., 1981.

Thermodynamic principles

Laws governing the transformation of energy. Thermodynamics is the science of the transformation of energy. It differs from the dynamics of Newton by taking into account the concept of temperature, which is outside the scope of classical mechanics. In practice, thermodynamics is useful for assessing the efficiencies of heat engines (devices that transform heat into work) and refrigerators (devices that use external sources of work to transfer heat from a hot system to cooler sinks), and for discussing the spontaneity of chemical reactions (their tendency to occur naturally) and the work that they can be used to generate.

The subject of thermodynamics is founded on four generalizations of experience, which are called the laws of thermodynamics. Each law embodies a particular constraint on the properties of the world. The connection between phenomenological thermodynamics and the properties of the constituent particles of a system is established by statistical thermodynamics, also called statistical mechanics. Classical thermodynamics consists of a collection of mathematical relations between observables, and as such is independent of any underlying model of matter (in terms, for instance, of atoms). However, interpretations in terms of the statistical behavior of large assemblies of particles greatly enriches the understanding of the relations established by thermodynamics, and a full description of nature should use explanations that move effortlessly between the two modes of discourse. *See* STATISTICAL MECHANICS.

There is a handful of very primitive concepts in thermodynamics. The first is the distinction between the system, which is the assembly of interest, and the surroundings, which is everything else. The surroundings is where observations on the system are carried out and attempts are made to infer its properties from those measurements. The system and the surroundings jointly constitute the universe and are distinguished by the boundary that separates them. If the boundary is impervious to the penetration of matter and energy, the system is classified as isolated. If energy but not matter can pass through it, the system is said to be closed. If both energy and matter can penetrate the boundary, the system is said to be open. Another primitive concept of thermodynamics is work. By work is meant a process by which a weight may be raised in the surroundings. Work is the link between mechanics and thermodynamics. *See* WORK.

None of these primitive concepts introduces the properties that are traditionally regarded as central to thermodynamics, namely temperature, energy, heat, and entropy. These concepts are introduced by the laws and are based on the foundations that these primitive concepts provide.

Zeroth law of thermodynamics. The zeroth law of thermodynamics establishes the existence of a property called temperature. This law is based on the observation that if a system A is in thermal equilibrium with a system B (that is, no change in the properties of B take places when the two are in contact), and if system B is in thermal equilibrium with a system C, then it is invariably the case that A will be found to be in equilibrium with C if the two systems are placed in mutual contact. This law suggests that a numerical scale can be established for the common property, and if A, B, and C have the same numerical values of this property, then they will be in mutual thermal equilibrium if they were placed in contact. This property is now called the temperature.

In thermodynamics it is appropriate to report temperatures on a natural scale, where 0 is ascribed to the lowest attainable temperature. Temperatures on this thermodynamic temperature scale are denoted T and are commonly reported in kelvins. The relation between the Kelvin scale and the Celsius scale (θ) is $T(\mathrm{K}) = \theta(^\circ\mathrm{C}) + 273.15$. *See* TEMPERATURE.

First law of thermodynamics. The first law of thermodynamics establishes the existence of a property called the internal energy of a system. It also brings into the discussion the concept of heat.

The first law is based on the observation that a change in the state of a system can be brought about by a variety of techniques. Indeed, if attention is confined to an adiabatic system, one that is thermally insulated from its surroundings, then the work of J. P. Joule shows that same change of state is brought about by a given quantity of work regardless of the manner in which the work is done (**Fig. 1**). This observation suggests that, just as the height through which a mountaineer climbs can be calculated from the difference in altitudes regardless of the path the climber takes between two fixed points, so the work, w, can be calculated from the difference between the final and initial properties of a system. The relevant property is called the internal energy, U. However, if the transformation of the system is taken along a path that is not adiabatic, a different quantity of work may be required. The difference between the work of adiabatic change and the work of nonadiabatic change is called heat, q. In general, Eq. (1) is

$$\Delta U = w + q \tag{1}$$

satisfied, where ΔU is the change in internal energy between the final and initial states of the system. *See* ADIABATIC PROCESS; ENERGY; HEAT.

The down-to-earth implication of this slightly casuistical argument is that there are two modes of

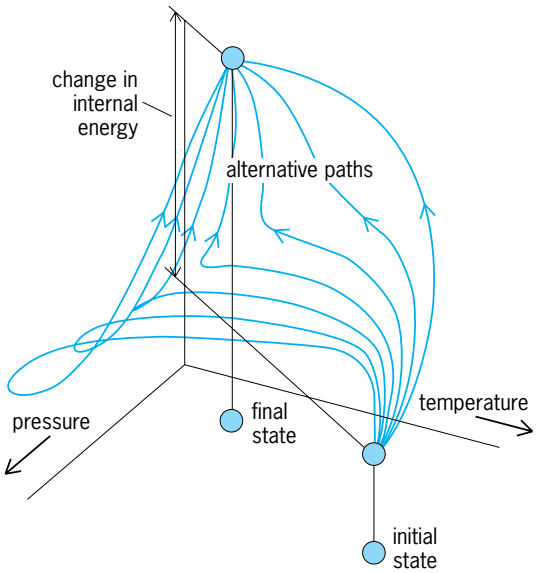

Fig. 1. The observation that the same work is done whatever the adiabatic path between two specified states of a system implies that each state has a characteristic property such that the difference in its values is equal to the adiabatic work required to pass between them. This characteristic property, a so-called state property, is called the internal energy of the system.

transferring energy between a system and its surroundings. One is by doing work; the other is by heating the system. Work and heat are modes of transferring energy. They are not forms of energy in their own right. Work is a mode of transfer that is equivalent (if not the case in actuality) to raising a weight in the surroundings. Heat is a mode of transfer that arises from a difference in temperature between the system and its surroundings. What is commonly called heat is more correctly called the thermal motion of the molecules of a system.

The molecular interpretation of thermodynamics adds insight to the operational distinction between work and heat. Work is a transfer of energy that stimulates (or is caused by) organized molecular motion in the surroundings. Thus, the raising of a weight by a system corresponds to the organized, unidirectional motion of the atoms of the weight. In contrast, heat is the transfer of energy that stimulates (or is caused by) chaotic molecular motion in the surroundings. Thus, the emergence of energy as heat into the surroundings is the chaotic, tumbling-out of stored energy.

The first law of thermodynamics states that the internal energy of an isolated system is conserved. That is, for a system to which no energy can be transferred by the agency of work or of heat, the internal energy remains constant. This law is a cousin of the law of the conservation of energy in mechanics, but it is richer, for it implies the equivalence of heat and work for bringing about changes in the internal energy of a system (and heat is foreign to classical mechanics). The first law is a statement of permission: no change may occur in an isolated system unless that change corresponds to a constant internal energy.

It is common in thermodynamics to switch attention from the change in internal energy of a system to the change in enthalpy, ΔH, of the system. The change in internal energy and the change in enthalpy of a system subjected to constant pressure p are related by Eq. (2), where ΔV is the change in volume of

$$\Delta H = \Delta U + p\Delta V \tag{2}$$

the system that accompanies the change of interest. The interpretation of ΔH is that it is equal to the energy that may be obtained as heat when the process occurs. This interpretation follows from the fact that the term $p\Delta V$ takes into account the work of driving back the surroundings that must take place during the process, and that is therefore not available for providing heat. Enthalpy changes are widely used in thermochemistry, the branch of physical chemistry concerned with the heat productions and requirements of chemical reactions. *See* ENTHALPY; THERMOCHEMISTRY.

The enthalpy itself is defined in terms of the internal energy by Eq. (3). Two important quantities in

$$H = U + pV \tag{3}$$

thermodynamics are the heat capacities at constant volume C_V and constant pressure C_p. These quantities are defined as the slope of the variation of internal energy and enthalpy, respectively, with respect to temperature. The two quantities differ on account of the work that must be done to change the volume of the system when the constraint is that of constant pressure, and then less energy is available for raising the temperature of the system. For a perfect gas, they are related by Eq. (4), where n is the amount of

$$C_p - C_V = nR \tag{4}$$

substance and R is the gas constant. *See* GAS; HEAT CAPACITY.

Second law of thermodynamics. The second law of thermodynamics deals with the distinction between spontaneous and nonspontaneous processes. A process is spontaneous if it occurs without needing to be driven. In other words, spontaneous changes are natural changes, like the cooling of hot metal and the free expansion of a gas. Many conceivable changes occur with the conservation of energy globally, and hence are not in conflict with the first law; but many of those changes turn out to be nonspontaneous, and hence occur only if they are driven.

The second law was formulated by Lord Kelvin and by R. Clausius in a manner relating to observation: "no cyclic engine operates without a heat sink" and "heat does not transfer spontaneously from a cool to a hotter body," respectively (**Fig. 2**). The two statements are logically equivalent in the sense that failure of one implies failure of the other. However, both may be absorbed into a single statement: the entropy of an isolated system increases when a spontaneous change occurs. The property of entropy is introduced to formulate the law quantitatively in exactly the same way that the properties

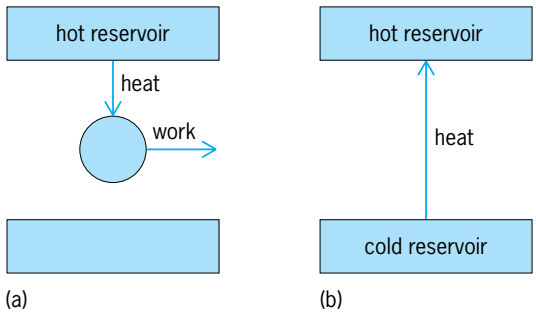

Fig. 2. Representation of the statements of the second law of thermodynamics by (*a*) Lord Kelvin and (*b*) R. Clausius. In each case, the law states that the device cannot operate as shown.

of temperature and internal energy are introduced to render the zeroth and first laws quantitative and precise.

The entropy, S, of a system is a measure of the quality of the energy that it stores. The formal definition is based on Eq. (5), where dS is the change in

$$dS = \frac{dq_{\text{reversible}}}{T} \tag{5}$$

entropy of a system, dq is the energy transferred to the system as heat, T is the temperature, and the subscript "reversible" signifies that the transfer must be carried out reversibly (without entropy production other than in the system). When a given quantity of energy is transferred as heat, the change in entropy is large if the transfer occurs at a low temperature and small if the temperature is high.

This somewhat austere definition of entropy is greatly illuminated by L. Boltzmann's interpretation of entropy as a measure of the disorder of a system. The connection can be appreciated qualitatively at least by noting that if the temperature is high, the transfer of a given quantity of energy as heat stimulates a relatively small additional disorder in the thermal motion of the molecules of a system; in contrast, if the temperature is low, the same transfer could stimulate a relatively large additional disorder. This connection between entropy and disorder is justified by a more detailed analysis, and in general it is safe to interpret the entropy of a system or its surroundings as a measure of the disorder present.

The formal statement of the second law of thermodynamics is that the entropy of an isolated system increases in the course of a spontaneous change. The illumination of the law brought about by the association of entropy and disorder is that in an isolated system (into which technology cannot penetrate) the only changes that may occur are those in which there is no increase in order. Thus, energy and matter tend to disperse in disorder (that is, entropy tends to increase), and this dispersal is the driving force of spontaneous change. *See* TIME, ARROW OF.

This collapse into chaos need not be uniform. The isolated system need not be homogeneous, and there may be an increase in order in one part so long as there is a compensating increase in disorder in another part. Thus, in thermodynamics, collapse into disorder in one region of the universe can result in the emergence of order in another region. The criterion for the emergence of order is that the decrease in entropy associated with it is canceled by a greater increase in entropy elsewhere. *See* ENTROPY.

Third law of thermodynamics. The practical significance of the second law is that it limits the extent to which the internal energy may be extracted from a system as work. In short, in order for a process to generate work, it must be spontaneous. (There is no point in using a process to produce work if that process itself needs to be driven; it is then no more than a gear wheel.) Therefore, any work-producing process must be accompanied by an increase in entropy. If a quantity of energy were to be withdrawn from a hot source and converted entirely into work, there would be a decrease in the entropy of the hot source, and no compensating increase elsewhere. Therefore, such a process is not spontaneous and cannot be used to generate work. For the process to be spontaneous, it is necessary to discard some energy as heat in a sink of lower temperature. In other words, nature in effect exacts a tax on the extraction of energy as work. There is therefore a fundamental limit on the efficiency of engines that convert heat into work (**Fig. 3**).

The quantitative limit on the efficiency, ϵ, which is defined as the work produced divided by the heat absorbed from the hot source, was first derived by S. Carnot. He found that, regardless of the details of the construction of the engine, the maximum efficiency (that is, the work obtained after payment of the minimum allowable tax to ensure spontaneity) is given by Eq. (6), where T_{hot} is the temperature of the

$$\epsilon = 1 - \frac{T_{\text{cold}}}{T_{\text{hot}}} \tag{6}$$

hot source and T_{cold} is the temperature of the cold sink. The greatest efficiencies are obtained with the coldest sinks and the hottest sources, and these are

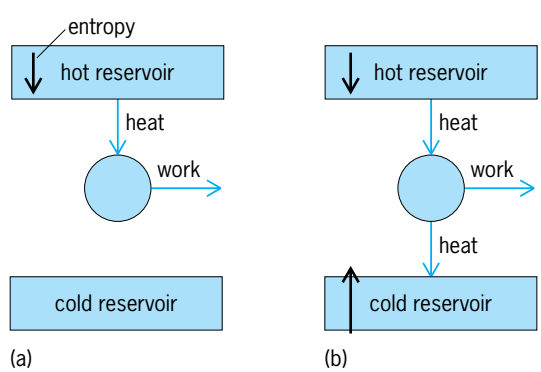

Fig. 3. Criterion for ability of a device to produce work. (*a*) This device cannot produce work because there is a decrease in entropy of the hot reservoir as energy leaves it as heat, but no compensating increase in entropy of the cold sink; overall, there is a decrease in entropy. (*b*) So long as a certain quantity of energy is discarded as heat into the cold reservoir, the overall change in entropy may be positive, and the engine can produce work spontaneously.

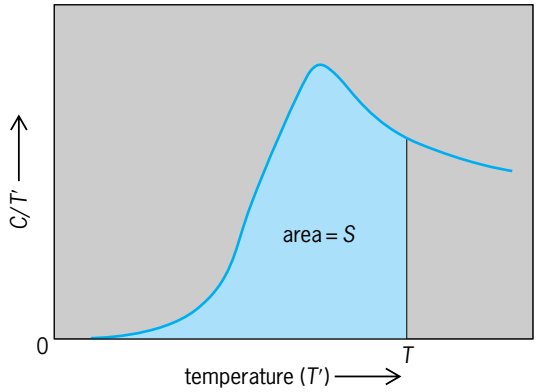

Fig. 4. Representation of the experimental determination of entropy. Measurements are made of the heat capacity, *C*, down to as low a temperature, *T′*, as possible, and *C/T′* is plotted against *T′*. The area under the curve up to the temperature of interest, *T*, is equal to the entropy, *S*, of the system; this procedure supposes that $S = 0$ at $T = 0$.

the design requirements of modern power plants. *See* CARNOT CYCLE.

Perfect efficiency ($\epsilon = 1$) would be obtained if the cold sink were at absolute zero ($T_{\text{cold}} = 0$). However, the third law of thermodynamics, which is another summary of observations, asserts that absolute zero is unattainable in a finite number of steps for any process. Therefore, heat can never be completely converted into work in a heat engine. The implication of the third law in this form is that the entropy change accompanying any process approaches zero as the temperature approaches zero. That implication in turn implies that all substances tend toward the same entropy as the temperature is reduced to zero. It is therefore sensible to take the entropy of all perfect crystalline substances (substances in which there is no residual disorder arising from the location of atoms) as equal to zero. A common short statement of the third law is therefore that all perfect crystalline substances have zero entropy at absolute zero ($T = 0$). This statement is consistent with the interpretation of entropy as a measure of disorder, since at absolute zero all thermal motion has been quenched. *See* ABSOLUTE ZERO.

In practice, the entropy of a sample of a substance is measured by determining its heat capacity, C, at all temperatures between zero and the temperature of interest, T, and evaluating the integral given in Eq. (7). Graphically, C/T' is plotted against T', and

$$S = \int_0^T \frac{C}{T'}\, dT' \tag{7}$$

the area under the curve up to the temperature of interest is equal to the entropy (**Fig. 4**). In practice, measurements of the heat capacity are made down to as low a temperature as possible, and certain approximations are generally carried out in order to extrapolate these measurements down to absolute zero. A polynomial is fitted to the data, and the integration in Eq. (7) is performed analytically. If there are phase transitions below the temperature of interest, a contribution from each such transition is added, equal to the enthalpy change of the transition divided by the temperature at which it occurs. Such determinations show that the entropy of a substance increases as it changes from a solid to a liquid to a gas. *See* PHASE TRANSITIONS.

Gibbs free energy. One of the most important derived quantities in thermodynamics is the Gibbs energy, G, which is widely called the free energy. It is defined by Eq. (8), where H is the enthalpy of the

$$G = H - TS \tag{8}$$

system, T is its thermodynamic temperature, and S is its entropy. The principal significance of G is that a change in G is a measure of the energy of the system that is free to do work other than simply driving back the walls of the system as the process occurs. For instance, it is a measure of the electrical work that may be extracted when a chemical reaction takes place, or the work of constructing a protein that a biochemical process may achieve.

The Gibbs energy can be developed in two different ways. First, it is quite easy to show from formal thermodynamics that Eq. (9) is valid. That is, the

$$\Delta G = -T\Delta S(\text{total}) \tag{9}$$

change, ΔG, in the Gibbs energy is proportional to the total change, $\Delta S(\text{total})$, in the entropy of the system and its surroundings. The negative sign in Eq. (9) indicates that an increase in the total entropy corresponds to a decrease in the Gibbs energy (**Fig. 5**). Because a spontaneous change occurs in the direction of the increase in total entropy, it follows that another way of expressing the signpost of spontaneous change is that it occurs in the direction of decreasing Gibbs energy. To this extent, the Gibbs energy is no more than a disguised version of the total entropy.

However, the Gibbs energy is much more than that, for (as discussed above) it shows how much

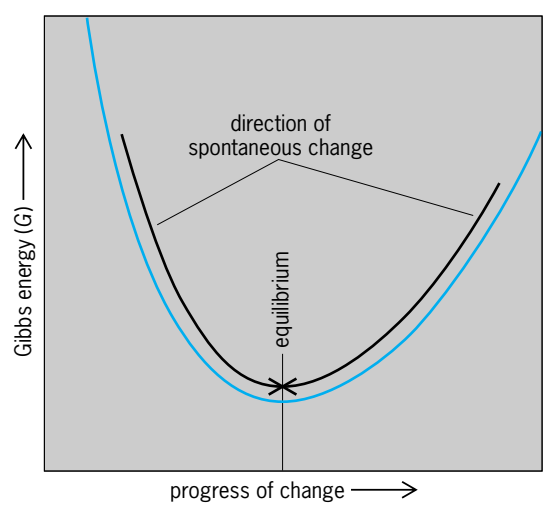

Fig. 5. Gibbs energy of a system. A spontaneous change (under conditions of constant temperature and pressure) corresponds to a decrease in Gibbs energy. This is a disguised form of the identification of the spontaneous direction of change with an increase in the total entropy of the system and its surroundings.

nonexpansion work may be extracted from a process. Broadly speaking, because a change in Gibbs energy at constant temperature can be expressed as $\Delta G = \Delta H - T\Delta S$, the latter term represents the tax exacted by nature to ensure that overall a process is spontaneous. Whereas ΔH measures the energy that may be extracted as heat, some energy may need to be discarded into the surroundings to ensure that overall there is an increase in entropy when the process occurs, and that quantity ($T\Delta S$) is then no longer available for doing work. This is the origin of the name free energy for the Gibbs energy, for it represents that energy stored by the system that is free to be extracted as work.

For a chemical reaction, the standard reaction Gibbs energy is calculated from the differences of the standard Gibbs energies of formation of each substance, the change in Gibbs energy accompanying the formation of the substance from its elements under standard conditions (a pressure of 1 bar or 10^5 pascals). The standard reaction Gibbs energy is the principal thermodynamic function for applications in chemistry.

The Gibbs energy is at the center of attention in chemical thermodynamics. It is deployed by introducing a related quantity called the chemical potential. The chemical potential is defined as the slope of the graph showing how the total Gibbs energy of a system varies as the composition of one of its components is increased. The slope of the graph varies with composition, so the chemical potential also varies with composition. Broadly speaking, the chemical potential can be interpreted as a measure of the potential of a substance to undergo chemical change: if its chemical potential is high, then it has a high potential for bringing about chemical change. Thus, a gas at high pressure has a higher chemical potential than one at low pressure, and a substance at high temperature has a higher chemical potential than the same substance at a lower temperature.

An implication of the second law is that the chemical potential of a substance must be the same throughout any phase in which it occurs and the same in all phases that are at equilibrium in a system. These requirements lead, by a subtle argument, to one of the most celebrated conclusions in chemical thermodynamics, the Gibbs phase rule, Eq. (10).

$$F = C - P + 2 \tag{10}$$

In this expression, C is the number of components in the system (essentially the number of chemically distinct species), P is the number of phases that are in equilibrium with one another, and F is the variance, the number of variables that may be changed without changing the number of phases that are in equilibrium with one another. (In some formulations, C is denoted $S - M$, where S is the number of substances and M is the number of reactions which relate them and which are at equilibrium.) The phase rule is particularly important for discussing the structure of phase diagrams, which are charts showing the range of temperature and composition over which various

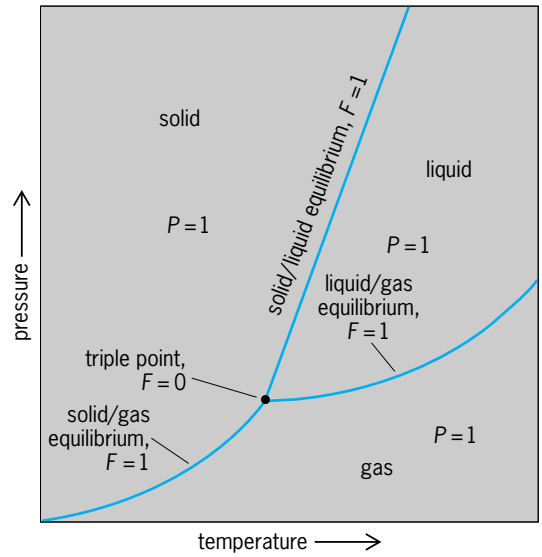

Fig. 6. Simple phase diagram for a one-component system ($C = 1$) and its interpretation in terms of the phase rule. The lines ($F = 1$) show the conditions under which two phases ($P = 2$) are in mutual equilibrium; the triple point ($F = 0$) is the unique set of conditions under which three phases ($P = 3$) can mutually coexist in equilibrium. The regions separated by the phase boundaries show the conditions where the specified phase is thermodynamically the most stable.

phases of a system are in equilibrium (**Fig. 6**). *See* PHASE EQUILIBRIUM; PHASE RULE.

In systems in which chemical reactions can occur, the chemical potentials of the reactants and products can be used to determine the composition at which the reaction mixture has reached a state of dynamic equilibrium, with no remaining tendency for spontaneous change. In chemistry, the state of chemical equilibrium is normally expressed in terms of the equilibrium constant, K, which is defined in terms of the concentrations or partial pressures of the participating species. In general, if the products dominate in the reaction mixture at equilibrium, then the equilibrium constant is greater than 1, and if the reactants dominate, then it is less than 1. Manipulations of standard thermodynamic relations show that the standard reaction Gibbs energy of any reaction is proportional to the negative of the logarithm of the equilibrium constant. *See* CHEMICAL EQUILIBRIUM; CHEMICAL THERMODYNAMICS; FREE ENERGY.

Thermodynamics of irreversible processes. The thermodynamic study of irreversible processes centers on the rate of entropy production and the relation between the fluxes and the forces that give rise to them. These fluxes include diffusion (the flux of matter), thermal conduction (the flux of energy of thermal motion), and electrical conduction (the flux of electric charge). In each case, the flux arises from a generalized potential difference of some kind. Thus, diffusion is related to a concentration gradient, and thermal conduction is related to a temperature gradient. In each case, the rate of change of entropy arising from the flux is proportional to both the flux and

the gradient giving rise to the flux. Thus, a high flux of matter down a steep concentration gradient results in a rapid change in entropy. *See* CONDUCTION (ELECTRICITY); CONDUCTION (HEAT); DIFFUSION.

An important observation is that the fluxes, J_i, and potentials, X_j, are not independent of one another. Thus, a temperature gradient can result in diffusion, and a concentration gradient can result in a flux of energy. The general relation between flux and potential is therefore given by Eq. (11), where the L_{ij}

$$J_i = \sum_j L_{ij} X_j \tag{11}$$

are called the phenomenological coefficients. It was shown by L. Onsager that for conditions not far from equilibrium, $L_{ij} = L_{ji}$. This Onsager reciprocity relation implies that there is a symmetry in the ability of a potential X_j to give to a flux J_i and of a potential X_j to give rise to a flux J_j. *See* THERMODYNAMIC PROCESSES; THERMOELECTRICITY. P. W. Atkins

Bibliography. P. W. Atkins, *Physical Chemistry*, 6th ed., 1998; W. T. Grandy, *Foundations of Statistical Mechanics*, 2 vols., 1988; J. Keizer, *Statistical Thermodynamics of Nonequilibrium Processes*, 1987; I. Müller, *Thermodynamics*, 1985; R. E. Sonntag, C. Borgnakke, and G. J. Van Wylen, *Fundamentals of Thermodynamics*, 5th ed., 1997; K. Wark, *Thermodynamics*, 6th ed., 2000; S. E. Wood and R. Battino, *Thermodynamics of Chemical Systems*, 1990.

Thermodynamic processes

Changes of any property of an aggregation of matter and energy, accompanied by thermal effects. The participants in a process are first identified as a system to be studied; the boundaries of the system are established; the initial state of the system is determined; the path of the changing states is laid out; and, finally, supplementary data are stated to establish the thermodynamic process. These steps will be explained in the following paragraphs. At all times it must be remembered that the only processes which are allowed are those compatible with the first and second laws of thermodynamics: Energy is neither created nor destroyed and the entropy of the system plus its surroundings always increases.

A system and its boundaries. To evaluate the results of a process, it is necessary to know the participants that undergo the process, and their mass and energy. A region, or a system, is selected for study, and its contents determined. This region may have both mass and energy entering or leaving during a particular change of conditions, and these mass and energy transfers may result in changes both within the system and within the surroundings which envelop the system.

As the system undergoes a particular change of condition, such as a balloon collapsing due to the escape of gas or a liquid solution brought to a boil in a nuclear reactor, the transfers of mass and energy which occur can be evaluated at the boundaries of the arbitrarily defined system under analysis.

A question that immediately arises is whether a system such as a tank of compressed air should have boundaries which include or exclude the metal walls of the tank. The answer depends upon the aim of the analysis. If its aim is to establish a relationship among the physical properties of the gas, such as to determine how the pressure of the gas varies with the gas temperature at a constant volume, then only the behavior of the gas is involved; the metal walls do not belong within the system. However, if the problem is to determine how much externally applied heat would be required to raise the temperature of the enclosed gas a given amount, then the specific heat of the metal walls, as well as that of the gas, must be considered, and the system boundaries should include the walls through which the heat flows to reach the gaseous contents. In the laboratory, regardless of where the system boundaries are taken, the walls will always play a role and must be reckoned with.

State of a system. To establish the exact path of a process, the initial state of the system must be determined, specifying the values of variables such as temperature, pressure, volume, and quantity of material. If a number of chemicals are present in the system, the number of variables needed is usually equal to the number of independently variable substances present plus two such as temperature and pressure; exceptions to this rule occur in variable electric or magnetic fields and in some other well-defined cases. Thus, the number of properties required to specify the state of a system depends upon the complexity of the system. Whenever a system changes from one state to another, a process occurs.

Whenever an unbalance occurs in an intensive property such as temperature, pressure, or density, either within the system or between the system and its surroundings, the force of the unbalance can initiate a process that causes a change of state. Examples are the unequal molecular concentration of different gases within a single rigid enclosure, a difference of temperature across the system boundary, a difference of pressure normal to a nonrigid system boundary, or a difference of electrical potential across an electrically conducting system boundary. The direction of the change of state caused by the unbalanced force is such as to reduce the unbalanced driving potential. Rates of changes of state tend to decelerate as this driving potential is decreased.

Equilibrium. The decelerating rate of change implies that all states move toward new conditions of equilibrium. When there are no longer any balanced forces acting within the boundaries of a system or between the system and its surroundings, then no mechanical changes can take place, and the system is said to be in mechanical equilibrium. A system in mechanical equilibrium, such as a mixture of hydrogen and oxygen, under certain conditions might undergo a chemical change. However, if there is no net change in the chemical constituents, then the

mixture is said to be in chemical as well as in mechanical equilibrium.

If all parts of a system in chemical and mechanical equilibrium attain a uniform temperature and if, in addition, the system and its surroundings either are at the same temperature or are separated by a thermally nonconducting boundary, then the system has also reached a condition of thermal equilibrium.

Whenever a system is in mechanical, chemical, and thermal equilibrium, so that no mechanical, chemical, or thermal changes can occur, the system is in thermodynamic equilibrium. The state of equilibrium is at a point where the tendency of the system to minimize its energy is balanced by the tendency toward a condition of maximum randomness. In thermodynamics, the state of a system can be defined only when it is in equilibrium. The static state on a macroscopic level is nevertheless underlaid by rapid molecular changes; thermodynamic equilibrium is a condition where the forward and reverse rates of the various changes are all equal to one another. In general, those systems considered in thermodynamics can include not only mixtures of material substances but also mixtures of matter and all forms of energy. For example, one could consider the equilibrium between a gas of charged particles and electromagnetic radiation contained in an oven.

Process path. If under the influence of an unbalanced intensive factor the state of a system is altered, then the change of state of the system is described in terms of the end states or difference between the initial and final properties.

The path of a change of state is the locus of the whole series of states through which the system passes when going from an initial to a final state. For example, suppose a gas expands to twice its volume and that its initial and final temperatures are the same. Various paths connect these initial and final states: isothermal expansion, with temperature held constant at all times, or adiabatic expansion which results in cooling followed by heating back to the initial temperature while holding volume fixed.

Each of these paths can be altered by making the gas do varying amounts of work by pushing out a piston during the expansion, so that an extremely large number of paths can be followed even for such a simple example. The detailed path must be specified if the heat or work is to be a known quantity; however, changes in the thermodynamic properties depend only on the initial and final states and not upon the path.

There are several corollaries from the above descriptions of systems, boundaries, states, and processes. First, all thermodynamic properties are identical for identical states. Second, the change in a property between initial and final states is independent of path or processes. The third corollary is that a quantity whose change is fixed by the end states and is independent of the path is a point function or a property. However, it must be remembered that by the second law of thermodynamics not all states are available (possible final states) from a given initial state and not all conceivable paths are possible in going toward an available state.

Pressure-volume-temperature diagram. Whereas the state of a system is a point function, the change of state of a system, or a process, is a path function. Various processes or methods of change of a system from one state to another may be depicted graphically as a path on a plot using thermodynamic properties as coordinates.

The variable properties most frequently and conveniently measured are pressure, volume, and temperature. If any two of these are held fixed (independent variables), the third is determined (dependent variable). To depict the relationship among these physical properties of the particular working substance, these three variables may be used as the coordinates of a three-dimensional space. The resulting surface is a graphic presentation of the equation of state for this working substance, and all possible equilibrium states of the substance lie on this *P-V-T* surface. The *P-V-T* surface may be extensive enough to include all three phases of the working substance: solid, liquid, and vapor.

Because a *P-V-T* surface represents all equilibrium conditions of the working substance, any line on the surface represents a possible reversible process, or a succession of equilibrium states.

The portion of the *P-V-T* surface shown in **Fig. 1** typifies most real substances; it is characterized by contraction of the substance on freezing. Going from the liquid surface to the liquid-solid surface onto the solid surface involves a decrease in both temperature and volume. Water is one of the few exceptions to this condition; it expands upon freezing, and its resultant *P-V-T* surface is somewhat modified where the solid and liquid phases abut.

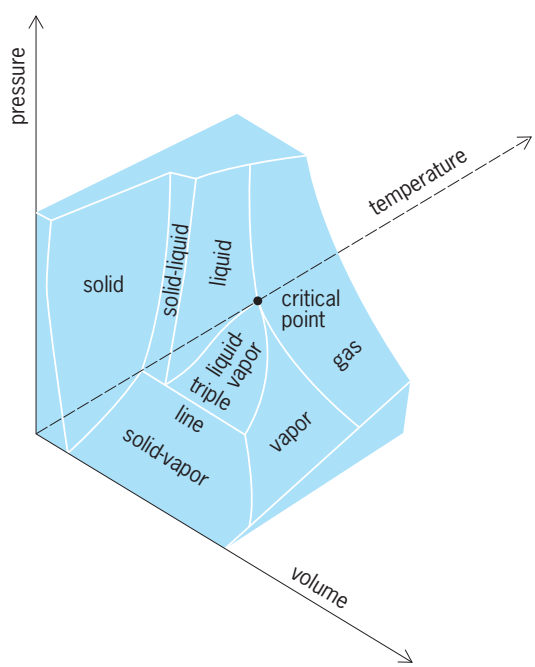

Fig. 1. Portion of pressure-volume-temperature (*P-V-T*) surface for a typical substance.

Gibbs' phase rule is defined in Eq. (1). Here f is the

$$f = c - p + 2 \tag{1}$$

degree of freedom; this integer states the number of intensive properties (such as temperature, pressure, and mole fractions or chemical potentials of the components) which can be varied independently of each other and thereby fix the particular equilibrium state of the system (see discussion under "Temperature-entropy diagram" below). Also, p indicates the number of phases (gas, liquid, or solid) and c the number of component substances in the system. Consider a one-component system (a pure substance) which is either in the liquid, gaseous, or solid phase. In equilibrium the system has two degrees of freedom; that is, two independent thermodynamic properties must be chosen to specify the state. Among the thermodynamic properties of a substance which can be quantitatively evaluated are the pressure, temperature, specific volume, internal energy, enthalpy, and entropy. From among these properties, any two may be selected. If these two prove to be independent of each other, when the values of these two properties are fixed, the state is determined and the values of all the other properties are also fixed. A one-component system with two phases in equilibrium (such as liquid in equilibrium with its vapor in a closed vessel) has $f = 1$; that is, only one intensive property can be independently specified. Also, a one-component system with three phases in equilibrium has no degree of freedom. Examination of Fig. 1 shows that the three surfaces (solid-liquid, solid-vapor, and liquid-vapor) are generated by lines parallel to the volume axis. Moving the system along such lines (constant pressure and temperature) involves a heat exchange and a change in the relative proportion of the two phases. Note that there is an entropy increment associated with this change.

One can project the three-dimensional surface onto the P-T plane as in **Fig. 2.** The triple point is

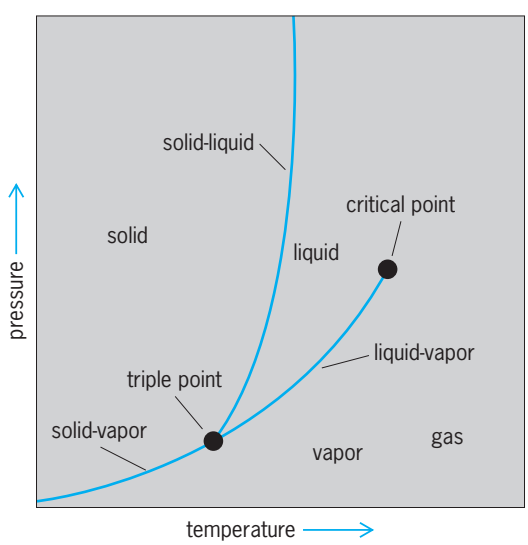

Fig. 2. Portion of equilibrium surface projected on pressure temperature (*P-T*) plane.

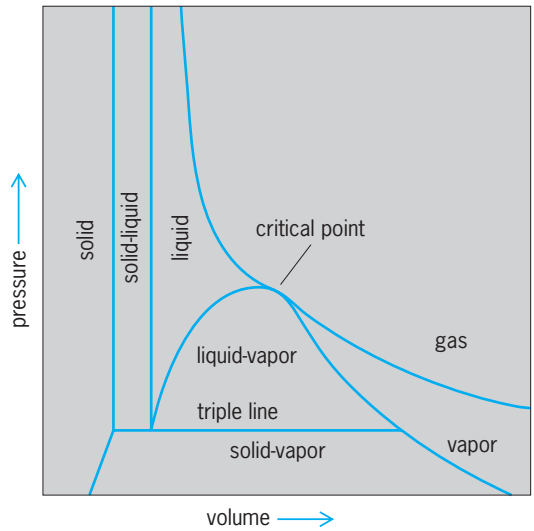

Fig. 3. Portion of equilibrium surface projected on pressure-volume (*P-V*) plane.

the point where the three phases are in equilibrium. When the temperature exceeds the critical temperature (at the critical point), only the gaseous phase is possible. The gas is called a vapor when it can coexist with another phase (at temperatures below the critical point). The P-T diagram for water would have the solid-liquid curve going upward from the triple point to the left (contrary to the ordinary substance pictured in Fig. 2). Then the property so well known to ice skaters would be evident. As the solid-liquid line is crossed from the low-pressure side to the high-pressure side, the water changes from solid to liquid: ice melts upon application of pressure.

Work of a process. The three-dimensional surface can also be projected onto the P-V plane to get **Fig. 3**. This plot has a special significance. The area under any reversible path on this plane represents the work done during the process. The fact that this P-V area represents useful work can be demonstrated by the following example.

Let a gas undergo an infinitesimal expansion in a cylinder equipped with a frictionless piston, and let this expansion perform useful work on the surroundings. The work done during this infinitesimal expansion is the force multiplied by the distance through which it acts, as in Eq. (2), wherein dW is

$$dW = F\,dl \tag{2}$$

an infinitesimally small work quantity, F is the force, and dl is the infinitesimal distance through which F acts.

But force F is equal to the pressure P of the fluid times the area A of the piston, or PA. However, the product of the area of the piston times the infinitesimal displacement is really the infinitesimal volume swept by the piston, or $A\,dl = dV$, with dV equal to an infinitesimal volume. Thus Eq. (3) is valid. The

$$dW = PA\,dl = P\,dV \tag{3}$$

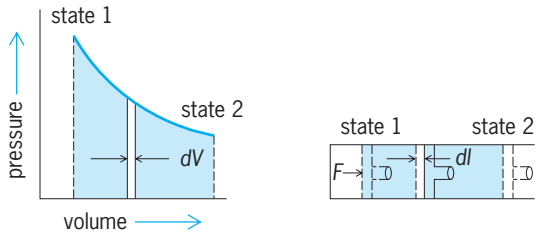

Fig. 4. Area under path in *P-V* plane is work done by expanding gas against piston.

work term is found by integration, as in Eq. (4).

$$ {}_1W_2 = \int_1^2 P\,dV \tag{4} $$

Figure 4 shows that the integral represents the area under the path described by the expansion from state 1 to state 2 on the *P-V* plane. Thus, the area on the *P-V* plane represents work done during this expansion process.

Temperature-entropy diagram. Energy quantities may be depicted as the product of two factors: an intensive property and an extensive one. Examples of intensive properties are pressure, temperature, and magnetic field; extensive ones are volume, magnetization, and mass. Thus, in differential form, work has been presented as the product of a pressure exerted against an area which sweeps through an infinitesimal volume, as in Eq. (5). Note that as a gas

$$ dW = P\,dV \tag{5} $$

expands, it is doing work on its environment. However, a number of different kinds of work are known. For example, one could have work on polarization of a dielectric, of magnetization, of stretching a wire, or of making new surface area. In all cases, the infinitesimal work is given by Eq. (6), where X is a generalized

$$ dW = X\,dx \tag{6} $$

applied force which is an intensive quantity such as voltage, magnetic field, or surface tension; and dx is a generalized displacement of the system and is thus extensive. Examples of dx include changes in electric polarization, magnetization, length of a stretched wire, or surface area.

By extending this approach, one can depict transferred heat as the product of an intensive property, temperature, and a distributed or extensive property, defined as entropy, for which the symbol is S. *See* ENTROPY.

If an infinitesimal quantity of heat dQ is transferred during a reversible process, this process may be expressed mathematically as in Eq. (7), with T being

$$ dQ = T\,dS \tag{7} $$

the absolute temperature and dS the infinitesimal entropy quantity.

Furthermore, a plot of the change of state of the system undergoing this reversible heat transfer can be drawn on a plane in which the coordinates are absolute temperature and entropy (**Fig. 5**). The total heat transferred during this process equals the area between this plotted line and the horizontal axis.

Reversible processes. Not all energy contained in or associated with a mass can be converted into useful work. Under ideal conditions only a fraction of the total energy present can be converted into work. The ideal conversions which retain the maximum available useful energy are reversible processes.

Characteristics of a reversible process are that the working substance is always in thermodynamic equilibrium and the process involves no dissipative effects such as viscosity, friction, inelasticity, electrical resistance, or magnetic hysteresis. Thus, reversible processes proceed quasistatically so that the system passes through a series of states of thermodynamic equilibrium, both internally and with its surroundings. This series of states may be traversed just as well in one direction as in the other.

If there are no dissipative effects, all useful work done by the system during a process in one direction can be returned to the system during the reverse process. When such a process is reversed so that the system returns to its starting state, it must leave an effect on the surroundings since, by the second law of thermodynamics, in energy conversion processes the form of energy is always degraded. Part of the energy of the system (including heat source) is transferred as heat from a higher temperature to a lower temperature. The energy rejected to a lower-temperature heat sink cannot be recovered. To return the system (including heat source and sink) to its original state, then, requires more energy than the useful work done by the system during a process in one direction. Of course, if the process were purely a mechanical one with no thermal effects, then both the surroundings and system could be returned to their initial states. *See* CARNOT CYCLE; THERMODYNAMIC CYCLE.

It is impossible to satisfy the conditions of a quasistatic process with no dissipative effects; a reversible process is an ideal abstraction which is not realizable in practice but is useful for theoretical calculations. An ideal reversible engine operating between hotter and cooler bodies at the temperatures

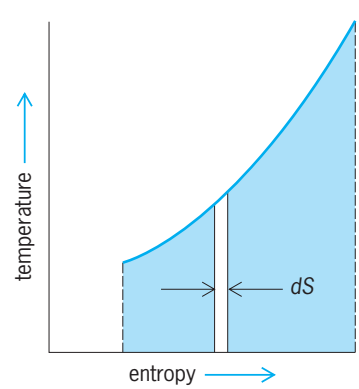

Fig. 5. Heat transferred during a reversible process is area under path in temperature-entropy (*T-S*) plane.

T_1 and T_2, respectively, can put out $(T_1 - T)/T_1$ of the transferred heat energy as useful work.

There are four reversible processes wherein one of the common thermodynamic parameters is kept constant. The general reversible process for a closed or nonflow system is described as a polytropic process. *See* ISENTROPIC PROCESS; ISOBARIC PROCESS; ISOMETRIC PROCESS; ISOTHERMAL PROCESS; POLYTROPIC PROCESS.

Irreversible processes. Actual changes of a system deviate from the idealized situation of a quasistatic process devoid of dissipative effects. The extent of the deviation from ideality is correspondingly the extent of the irreversibility of the process.

Real expansions take place in finite time, not infinitely slowly, and these expansions occur with friction of rubbing parts, turbulence of the fluid, pressure waves sweeping across and rebounding through the cylinder, and finite temperature gradients driving the transferred heat. These dissipative effects, the kind of effects that make a pendulum or yo-yo slow down and stop, also make the work output of actual irreversible expansions less than the maximum ideal work of a corresponding reversible process. For a reversible process, as stated earlier, the entropy change is given by $dS = dQ/T$. For an irreversible process even more entropy is produced (turbulence and loss of information) and there is the inequality $dS > dQ/T$. Philip E. Bloomfield; William A. Steele

Bibliography. H. A. Bent, *The Second Law*, 1965; Y. A. Cengel and M. A. Boles, *Thermodynamics: An Engineering Approach*, 3d ed., 1998; J. P. Holman, *Thermodynamics,* 4th ed., 1988; M. Mott-Smith, *The Concept of Energy Simply Explained*, 1934; F. W. Sears and G. L. Salinger, *Thermodynamics, the Kinetic Theory of Gases and Statistical Mechanics,* 3d ed., 1975; K. Wark, *Thermodynamics,* 6th ed., 1999.

Thermoelectric power generator

A solid-state heat engine which employs the electron gas as a working fluid. It directly converts heat energy into electrical energy using the Seebeck effect. This phenomenon can be demonstrated using a thermocouple which comprises two legs (thermoelements) of dissimilar conducting materials joined at one end to form a junction. If this junction is maintained at a temperature which differs from ambient, a voltage is generated across the open ends of the thermoelements. When the circuit is completed with a load, a current flows in the circuit and power is generated. In practice the thermocouples are fabricated generally from *n*- and *p*-type semiconductors, and several hundred are connected electrically in series to form a module which is the active component of a thermoelectric generator. Provided a temperature difference is maintained across the device, it will generate electrical power. Heat is provided from a variety of sources depending on the application, and they include burning fossil fuels in terrestrial and military applications, decaying long-life isotopes in medical and deep-space applications, and waste heat. The performance of the thermoelectric generator, in terms of efficiency, output power, and economic viability, depends upon its temperature regime of operation; the materials used in the module construction; its electrical, thermal, and geometrical design; and the generator load. The power output spectrum of thermoelectric generators spans 14 orders of magnitude and ranges from nanowatt generators fabricated using integrated circuit technology to the nuclear reactor–powered 100-kW SP-100 generator intended to provide electrical power to orbiting space stations. *See* NUCLEAR BATTERY; RADIOACTIVITY AND RADIATION APPLICATIONS; SEEBECK EFFECT; SPACE POWER SYSTEMS; SPACE PROBE; THERMOCOUPLE; THERMOELECTRICITY.

Generating parameters. A thermoelectric generator is a heat engine, and like all heat engines it obeys the laws of thermodynamics. The efficiency of an ideal thermoelectric generator, one in which there are no heat losses, is defined as the ratio of the electrical power delivered to the load to the heat absorbed at the hot junction. *See* EFFICIENCY; THERMODYNAMIC PRINCIPLES; THERMODYNAMIC PROCESSES.

Expressions for the important parameters in thermoelectric generation can readily be derived by considering the simplest generator consisting of a single thermocouple with thermoelements fabricated from *n*- and *p*-type semiconductors respectively (**Fig. 1**). The efficiency of the generator is given by Eq. (1). If it

$$\phi = \frac{\text{energy supplied to the load}}{\text{heat energy absorbed at hot junction}} \qquad (1)$$

is assumed that the electrical conductivities, thermal conductivities, and Seebeck coefficients of the thermoelements *a* and *b* are constant within an arm, and that the contact resistances at the hot and cold junctions are negligible compared with the sum of the arm resistances, then the efficiency can be expressed

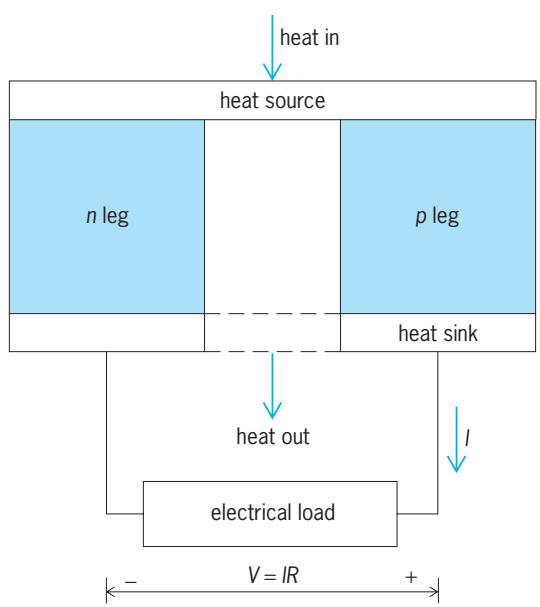

Fig. 1. Diagram of single-couple thermoelectric generator.

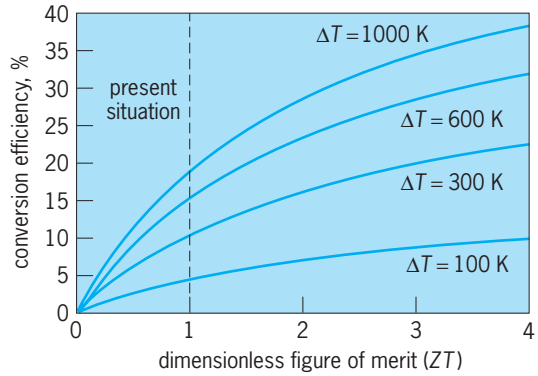

Fig. 2. Conversion efficiency as a function of the dimensionless figure of merit, *ZT*, for temperature differences of operation, ΔT, with cold junction Z_2 at 300 K (80°F). (For temperature differences, 1 K = 1.8°F.)

as Eq. (2), where I is the current, α_{ab} is the total See-

$$\phi = \frac{I^2 R_L}{\alpha_{ab} I T_1 + \lambda'(T_1 - T_2) - \frac{1}{2} I^2 R} \tag{2}$$

beck coefficient of a and b, λ' is the thermal conductance of a and b in parallel, R is the series resistance of a and b, R_L is the load resistance, and T_1 and T_2 are the absolute temperatures of the hot and cold junctions. In thermoelectric materials, the electrical conductivity, thermal conductivity, and seebeck coefficient change with temperature, and in actual application this is taken into account. Appropriate allowances can also be made for contact resistances. However, the simple expression for the efficiency, Eq. (2), can still be employed with an acceptable degree of accuracy. Assuming average values for these parameters provides results which are within 10% of the true value. *See* CONDUCTION (ELECTRICITY); CONDUCTION (HEAT); SEMICONDUCTOR.

Efficiency is clearly a function of the ratio of the load resistance to the sum of the generator arm resistances. At maximum power output, it can be shown that the efficiency is given in Eq. (3), where Z_c is the figure of merit of the couple, given by Eq. (4). The maximum efficiency is given by Eq. (5), where n_c,

$$\phi_p = \frac{T_1 - T_2}{\frac{3T_1}{2} + \frac{T_2}{2} + \frac{4}{Z_c}} \tag{3}$$

$$Z_c = \frac{\alpha_{ab}^2}{R\lambda'} \tag{4}$$

$$\phi_{\max} = \eta_c \gamma \tag{5}$$

the Carnot efficiency, is given by Eq. (6), and γ is given by Eqs. (7) and (8). The maximum efficiency

$$\eta_c = \frac{T_1 - T_2}{T_1} \tag{6}$$

$$\gamma = \frac{\sqrt{1 + Z_c \overline{T}} - 1}{\sqrt{1 + Z_c \overline{T}} + \frac{T_2}{T_1}} \tag{7}$$

$$\overline{T} = \frac{T_1 + T_2}{2} \tag{8}$$

is thus the product of the Carnot efficiency, which is clearly less than unity, and γ, which embodies the parameters of the materials. *See* CARNOT CYCLE.

If the geometries of a and b are matched to minimize heat absorption, then the figure of merit is given by Eq. (9), where σ_a and σ_b are the electrical con-

$$Z_c = \frac{\alpha_{ab}^2}{\left[\left(\frac{\lambda_a}{\sigma_a}\right)^{1/2} + \left(\frac{\lambda_b}{\sigma_b}\right)^{1/2}\right]^2} \tag{9}$$

ductivities a and b, and λ_a and λ_b are their thermal conductivities. In practice the two arms of the junction have similar material constants, in which case the concept of a figure of merit for a material is employed and is given by Eq. (10). The quantity Z varies

$$Z = \frac{\alpha^2 \sigma}{\lambda} \tag{10}$$

with temperature, and a more meaningful parameter is the dimensionless figure of merit ZT, where T is the absolute temperature at which the performance of the material is considered; $\alpha^2\sigma$ is referred to as the electrical power factor.

Improving conversion efficiency. Historically the use of thermoelectric generators has been restricted to specialized applications where combinations of their desirable properties, such as the absence of moving parts, reliability, and silent operation, has outweighed their low overall conversion efficiency, typically 5%. In these applications, fuel cost or weight is a major consideration, and improving the conversion efficiency is the main research target. The effect of ZT on the generating conversion efficiency for different temperature differences of operation is

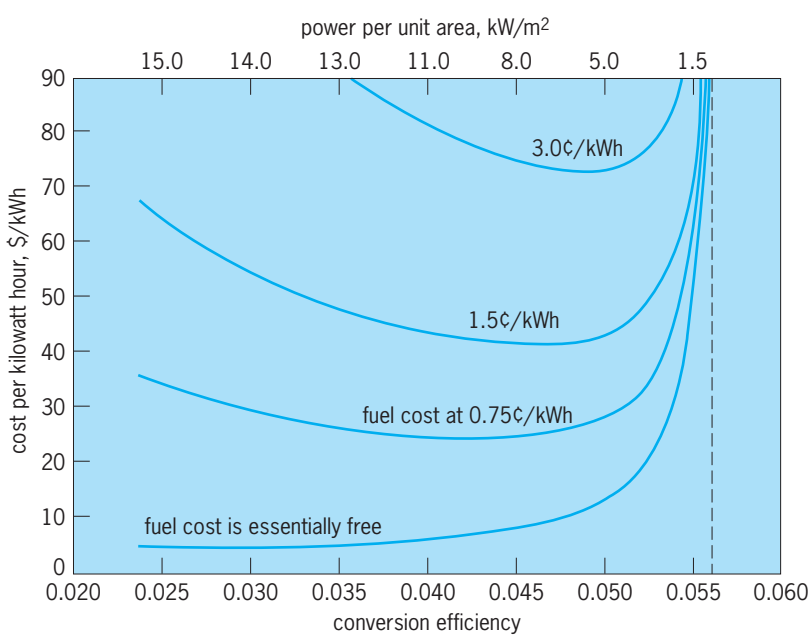

Fig. 3. Cost per killowatthour as a function of conversion efficiency and power per unit area for different fuel costs. The calculation is based on a typical 127-pair thermocouple module operating at ΔT = 120 K (216°F) for 15 years.

displayed in **Fig. 2**. Understandably, improving the figure of merit has been regarded as the most important factor in increasing the conversion efficiency of a thermoelectric generator.

Cost reduction. Consideration has been given to applications when the fuel cost is low or essentially free as with waste heat. The economic viability is measured by the cost per kilowatthour of electricity C. In the case of a thermoelectric generator, it can be shown that C is related to the generators fabrication cost C_g and the running cost (such as fuel) C_r by Eq. (11), where p and Δt are the power and operat-

$$C = \frac{C_g}{p\Delta t} + \frac{C_r}{\phi} \tag{11}$$

ing period. The relative importance of C_g and C_r for the fuel cost is shown in **Fig. 3**. When the fuel is relatively expensive, the module should be optimized to achieve high conversion efficiency. However, when the fuel cost is low or essentially free as in waste heat recovery, then the cost per watt is mainly determined by the power per unit area and the operating period.

Materials. Selection of thermocouple material depends upon the generator's temperature regime of operation. The figures of merit of established thermoelectric materials reach maxima at different temperatures, and semiconductor compounds or alloys based on bismuth telluride, lead telluride, and silicon germanium cover the temperature ranges up to 150°C (300°F), 650°C (1200°F), and 1000°C (1830°F) respectively, with the best materials capable of generating electrical power with an efficiency of around 20%. Material research is focused on improving the figure of merit and to a lesser, though an increasing, extent, the electrical power factor.

Improving the figure of merit. Although no theoretical upper boundary to the dimensionless figure of merit ZT exists, phenomenological evidence points to the existence of a barrier around $ZT = 2$. Nevertheless, improvement in the figure of merit is being sought primarily by reducing the lattice thermal conductivity. The bulk materials being researched include glasslike crystals such as the filled skutterudites within whose atomic cage weakly bound atoms or molecules rattle and should conduct heat like a glass but electricity like a crystal. The quest for improving the figure of merit by reducing the thermal conductivity has moved to exotic low-dimensional structures such as superlattices, quantum wires, and quantum dots, where in theory the reduced dimensions and the presence of interfaces give rise to a thermal conductivity which is lower than the bulk counterpart. A ZT close to 2 has been reported but as yet has not been confirmed independently. *See* ARTIFICIALLY LAYERED STRUCTURES; MESOSCOPIC PHYSICS; QUANTIZED ELECTRONIC STRUCTURE (QUEST); THERMAL CONDUCTION IN SOLIDS.

Improving the power factor. The emergence of thermoelectrics as a technology for application in waste heat recovery has resulted in a successful search for materials with high electrical power factors and cheap materials. The rare-earth ytterbium-aluminum compound $YbAl_3$ has a power factor almost three times that of bismuth telluride, the established material for low-temperature application, while magnesium tin (MgSn) has almost the same performance as lead telluride but is available at less than a quarter of the cost.

Applications. Thermoelectric generators continue to find novel applications, such as miniature thermocouple arrays to power a wristwatch using body heat and the utilization of automobile exhaust heat to thermoelectrically supplement the electrical power supply. Thermoelectric generation is an established technology and in some applications has become economically competitive, particularly in waste heat recovery.

D. M. Rowe

Bibliography. D. M. Rowe (ed.), *CRC Handbook of Thermoelectrics*, CRC Press, 1995; D. M. Rowe and C. M. Bhandari (eds.), *Modern Thermoelectrics*, Prentice Hall, 1983.

Thermoelectricity

The direct conversion of heat into electrical energy, or the reverse, in solid or liquid conductors by means of three interrelated phenomena—the Seebeck effect, the Peltier effect, and the Thomson effect—including the influence of magnetic fields upon each. The Seebeck effect concerns the electromotive force (emf) generated in a circuit composed of two different conductors whose junctions are maintained at different temperatures. The Peltier effect refers to the reversible heat generated at the junction between two different conductors when a current passes through the junction. The Thomson effect involves the reversible generation of heat in a single current-carrying conductor along which a temperature gradient is maintained. Specifically excluded from the definition of thermoelectricity are the phenomena of Joule heating and thermionic emission. *See* ELECTROMOTIVE FORCE (EMF); JOULE'S LAW; THERMIONIC EMISSION.

The three thermoelectric effects are described in terms of three coefficients: the absolute thermoelectric power (or thermopower) S, the Peltier coefficient Π, and the Thomson coefficient μ, each of which is defined for a homogeneous conductor at a given temperature. These coefficients are connected by the Kelvin relations, which convert complete knowledge of one into complete knowledge of all three. It is therefore necessary to measure only one of the three coefficients; usually the thermopower S is chosen. The combination of electrical resistivity, thermal conductivity, and thermopower (more generally, these quantities as tensors) is sufficient to provide a complete description of the electronic transport properties of conductors for which the electric current and heat current are linear functions of both the applied electric field and the temperature gradient. *See* ELECTRICAL CONDUCTIVITY OF METALS.

Thermoelectric effects have significant applications in both science and technology and show

promise of more importance in the future. Studies of thermoelectricity in metals and semiconductors yield information about electronic structure and about the interactions between electrons and both lattice vibrations and impurities. Practical applications include the measurement of temperature, generation of power, cooling, and heating. Thermocouples are widely used for temperature measurement, providing both accuracy and sensitivity. Research has been undertaken concerning the direct thermoelectric generation of electricity using the heat produced by nuclear reactions or generated at automobile exhausts. Cooling units using the Peltier effect have been constructed in sizes up to those of home refrigerators. Development of thermoelectric heating has also been undertaken.

Seebeck Effect

In 1821, T. J. Seebeck discovered that when two different conductors are joined into a loop and a temperature difference is maintained between the two junctions, an emf is generated. Such a loop is called a thermocouple, and the emf generated is called a thermoelectric (or Seebeck) emf.

Measurements. The magnitude of the emf generated by a thermocouple is standardly measured using the system shown in **Fig. 1**. Here the contact points between conductors A and B are called junctions. Each junction is maintained at a well-controlled temperature (either T_1 or T_0) by immersion in a bath or connection to a heat reservoir. These baths or reservoirs are indicated by the white squares. From each junction, conductor A is brought to a measuring device M, usually a potentiometer. When the potentiometer is balanced, no current flows, thereby allowing direct measurement of the open-circuit emf, undiminished by resistive losses and unperturbed by spurious effects arising from Joule heating or from Peltier heating and cooling at the junctions. This open-circuit emf is the thermoelectric emf. *See* POTENTIOMETER.

Equations. According to the experimentally established law of Magnus, for homogeneous conductors A and B the thermoelectric emf depends only upon the temperatures of the two junctions and not upon either the shapes of the samples or the detailed forms of the temperature distributions along them. This emf can thus be symbolized $E_{AB}(T_0,T_1)$. According to both theory and experiment, if one of the conductors, say B, is a superconductor in its superconducting state, it makes no contribution to E_{AB} (except for very small effects near the superconducting transition temperature T_c, discussed below). That is, when B is superconducting, $E_{AB}(T_0,T_1)$ is determined solely by conductor A, and can be written as $E_A(T_0,T_1)$. *See* SUPERCONDUCTIVITY.

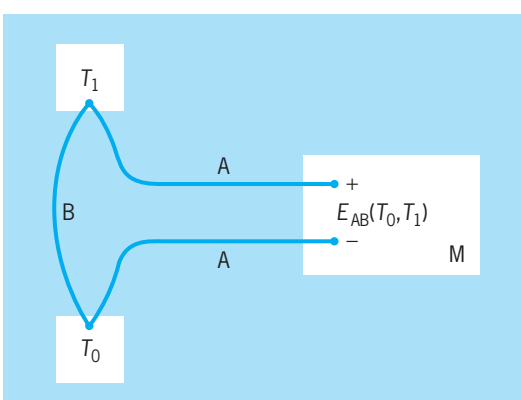

Fig. 1. Diagram of apparatus usually used for measuring thermoelectric (Seebeck) emf $E_{AB}(T_0,T_1)$. M is an instrument for measuring potential.

It is convenient to express this emf in terms of a property that depends only upon a single temperature. Such a property is the absolute thermoelectric power (or, simply, thermopower) $S_A(T)$, defined so that Eq. (1) is valid.

$$E_A(T_0, T_1) = \int_{T_0}^{T_1} S_A(T)\, dT \tag{1}$$

If $E_A(T, T+\Delta T)$ is known—for example, from measurements involving a superconductor—$S_A(T)$ can be determined from Eq. (2). If Eq. (1) is valid for any

$$S_A(T) = \lim_{\Delta T \to 0} \frac{E_A(T, T+\Delta T)}{\Delta T} \tag{2}$$

homogeneous conductor, then it ought to apply to both sides of the thermocouple shown in Fig. 1. Indeed, it has been verified experimentally that the emf $E_{AB}(T_0,T_1)$ produced by a thermocouple is just the difference between the emf's, calculated using Eq. (1), produced by its two arms. This result can be derived as follows. Employing the usual sign convention, to calculate $E_{AB}(T_0,T_1)$, begin at the cooler bath, T_0, integrate $S_A(T)\, dT$ along conductor A up to the warmer bath, T_1, and then return to T_0 along conductor B by integrating $S_B(T)\, dT$. This circular excursion produces $E_{AB}(T_0,T_1)$, given by Eq. (3). Inverting the last integral in Eq. (3) gives Eq. (4), which from Eq. (1) can be rewritten as Eq. (5). Alternatively, combining the two integrals in Eq. (4) gives Eq. (6). Defining S_{AB} according to Eq. (7) then yields Eq. (8).

$$E_{AB}(T_0, T_1) = \int_{T_0}^{T_1} S_A(T)\, dT + \int_{T_1}^{T_0} S_B(T)\, dT \tag{3}$$

$$E_{AB}(T_0, T_1) = \int_{T_0}^{T_1} S_A(T)\, dT - \int_{T_0}^{T_1} S_B(T)\, dT \tag{4}$$

$$E_{AB}(T_0, T_1) = E_A(T_0, T_1) - E_B(T_0, T_1) \tag{5}$$

$$E_{AB}(T_0, T_1) = \int_{T_0}^{T_1} [S_A(T)] - S_B(T)]\, dT \tag{6}$$

$$S_{AB}(T) = S_A(T) - S_B(T) \tag{7}$$

$$E_{AB}(T_0, T_1) = \int_{T_0}^{T_1} S_{AB}(T)\, dT \tag{8}$$

Equation (6) shows that $E_{AB}(T_0,T_1)$ can be calculated for a given thermocouple whenever the thermopowers $S_A(T)$ and $S_B(T)$ are known for its two constitutents over the temperature range T_0 to T_1.

By convention, the signs of $S_A(T)$ and $S_B(T)$ are chosen so that, if the temperature difference $T_1 - T_0$ is taken small enough so that $S_A(T)$ and $S_B(T)$ can be presumed constant, then $S_A(T) > S_B(T)$ when the emf $E_{AB}(T_0,T_1)$ has the polarity indicated in Fig. 1.

Results of equations. These equations lead directly to the following experimentally and theoretically verified results.

Uniform temperature. In a circuit kept at a uniform temperature throughout, $E = 0$, even though the circuit may consist of a number of different conductors. This result follows directly from Eq. (8), since $dT = 0$ everywhere throughout the circuit. It follows also from thermodynamic reasoning. If E did not equal 0, the circuit could drive an electrical motor and make it do work. But the only source of energy would be heat from the surroundings which, by assumption, are at the same uniform temperature as the circuit. Thus, a contradiction with the second law of thermodynamics would result. *See* CHEMICAL THERMODYNAMICS.

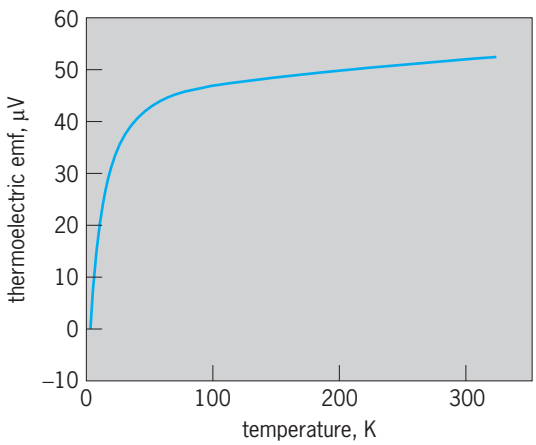

Fig. 2. Thermoelectric emf of a thermocouple formed from pure annealed and pure cold-worked copper. The cold junction reference temperature is 4.2 K (−452°F). °F = (K × 1.8) − 459.67. (*After R. H. Kropschot and F. J. Blatt, Thermoelectric power of cold-rolled pure copper, Phys. Rev., 116:617–620, 1959*)

Homogeneous conductor. A circuit composed of a single, homogeneous conductor cannot produce a thermoelectric emf. This follows from Eq. (6) when $S_B(T)$ is set equal to $S_A(T)$. It is important to emphasize that in this context "homogeneous" means perfectly uniform throughout. A sample made of an isotropic material can be inhomogeneous either because of small variations in chemical composition or because of strain. **Figure 2** shows the thermoelectric emf generated by a thermocouple in which one arm is a cold-rolled copper (Cu) sample, and the other arm is the same material after annealing at an elevated temperature to remove the effects of the strain introduced by the cold-rolling. **Figure 3** shows how the addition of impurities can change the thermopower of a pure metal. An additional effect can occur in a noncubic material. As illustrated in **Fig. 4**, a thermocouple formed from two samples cut in different

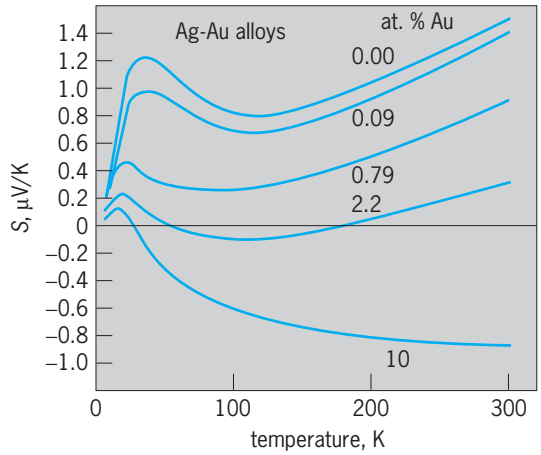

Fig. 3. Thermopower *S* from 0 to 300 K for pure silver (Ag) and a series of dilute silver-gold (Ag-Au) alloys, °F = (K × 1.8) − 459.67. (*After R. S. Crisp and J. Rungis, Thermoelectric power and thermal conductivity in the silver-gold alloy system from 3–300°K, Phil. Mag., 22:217–236, 1970*)

orientations from a noncubic single crystal can generate a thermoelectric emf even if each sample is highly homogeneous. *See* CRYSTAL STRUCTURE.

If material B is superconducting, so that $S_B = 0$, Eq. (5) reduces to $E_{AB}(T_0,T_1) = E_A(T_0,T_1)$, as assumed above.

Source of emf. Finally, Eq. (6) makes clear that the source of the thermoelectric emf in a thermocouple lies in the bodies of the two materials of which it is composed, rather than at the junctions. This serves to emphasize that thermoelectric emf's are not related to the contact potential or Volta effect, which is a potential difference across the junction between two different metals arising from the difference between their Fermi energies. The contact potential is present even in the absence of temperature gradients or electric currents. *See* CONTACT POTENTIAL DIFFERENCE.

Peltier Effect

In 1834, C. A. Peltier discovered that when an electric current passes through two different conductors

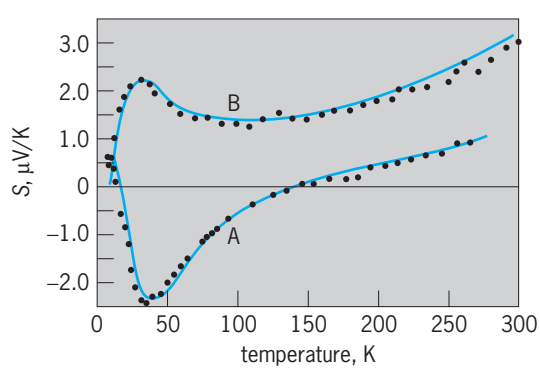

Fig. 4. Thermopower *S* of zinc parallel (A) and perpendicular (B) to the hexagonal axis. °F = (K × 1.8) − 459.67. (*After V. A. Rowe and P. A. Schroeder, Thermopower of Mg, Cd and Zn between 1.2° and 300°K, J. Phys. Chem. Sol., 31:1–8, 1970*)

connected in a loop, one of the two junctions between the conductors cools and the other warms. If the direction of the current is reversed, the effect also reverses: the first junction warms and the second cools. In 1853, Quintus Icilius showed that the rate of heat output or intake at each junction is directly proportional to the current i. The Peltier coefficient Π_{AB} is defined as the heat generated per second per unit current flow through the junction between materials A and B. By convention, Π_{AB} is taken to be positive when cooling occurs at the junction through which current flows from conductor A to conductor B. Quintus Icilius's result guarantees that the Peltier coefficient is independent of the magnitude of the current i. Additional experiments have shown that it is also independent of the shapes of the conductors. It therefore depends only upon the two materials and the temperature of the junction, and can be written as $\Pi_{AB}(T)$ or, alternatively, as $\Pi_A(T) - \Pi_B(T)$, where $\Pi_A(T)$ and $\Pi_A(T)$ are the Peltier coefficients for materials A and B, respectively. The second form emphasizes that the Peltier coefficient is a bulk property which can be defined for a single conductor.

Because of the difficulty of measuring heat input or output from a junction, as well as complications resulting from the simultaneous presence of Joule heating and the Thomson effect, $\Pi_{AB}(T)$ has rarely been quantitatively measured. Rather its value is usually determined from the Kelvin relations, using experimental values for S_{AB}. The Peltier effect does, however, form the basis for thermoelectric heating and cooling.

Thomson Effect and Kelvin Relations

When an electric current passes through a conductor that is maintained at a constant temperature, heat is generated at a rate proportional to the square of the current. This phenomenon is called Joule heat, and its magnitude for any given material is determined by the electrical resistivity of the material. In 1854, William Thomson (Lord Kelvin), in an attempt to explain discrepancies between experimental results and a relationship between Π_{AB} and S_{AB} that he had derived from thermodynamic analysis of a thermocouple, postulated the existence of an additional reversible generation of heat when a temperature gradient is applied to a current-carrying conductor. This heat, called Thomson heat, is proportional to the product of the current and the temperature gradient. It is reversible in the sense that the conductor changes from a generator of Thomson heat to an absorber of Thomson heat when the direction of either the current or the temperature gradient (but not both at once) is reversed. By contrast, Joule heating is irreversible in that heat is generated for both directions of current flow.

The magnitude of Thomson heat generated (or absorbed) is determined by the Thomson coefficient μ. Using reasoning based upon equilibrium thermodynamics, Thomson derived results equivalent to Eqs. (9) and (10), called the Kelvin (or Kelvin-Onsager) relations.

$$\frac{\Pi A}{T} = S_A \tag{9}$$

$$\frac{\mu A}{T} = \frac{dS_A}{dT} \tag{10}$$

Here μ_A is the Thomson coefficient, defined as the heat generated per second per unit current flow per unit temperature gradient when current flows through conductor A in the presence of a temperature gradient. Equation (10) can be integrated to give Eq. (11), in which the third law of thermodynamics

$$S_A = \int_0^T \frac{\mu_A(T')}{T'}\, dT \tag{11}$$

has been evoked to set $S_A(0) = 0$. Using Eq. (11), $S_A(T)$ can be determined from measurements on a single conductor. In practice, however, accurate measurements of μ_A are very difficult to make; therefore, they have been carried out for only a few metals—notably lead (Pb), platinum (Pt), and tungsten (W)—which then serve as standards for determining $S_B(T)$ by using measurements of $S_{AB}(T)$ in conjunction with Eq. (7).

Long after the Thomson heat was observed and the Kelvin relations were verified experimentally, debate raged over the validity of the derivation employed by Thomson. However, the theory of irreversible processes, developed by L. Onsager in 1931, and by others, yields the same equations and thus provides them with a firm foundation.

Thermopowers of Metals and Semiconductors

Since the Kelvin relations provide recipes for calculating any two of the thermoelectric coefficients, S, Π, and μ, from the third, only one of the three coefficients need be measured to determine the thermoelectric properties of any given material. Although there are circumstances under which one of the other two coefficients may be preferred, because of ease and accuracy of measurement, it is usually the thermopower S that is measured.

Reference materials. Because S must be measured using a thermocouple, the quantity measured experimentally is $S_A - S_B$, the difference between the thermopowers of the two conductors constituting the thermocouple. Only when one of the arms of the thermocouple is superconducting and therefore has zero thermopower can the absolute thermopower of the other be directly measured. At temperatures up to about 18 K (−427°F) superconducting niobium-tin (Nb_3Sn) wire can be used for conductor B, thereby permitting direct determination of S_A. Direct measurements have been extended to just above 90 K (−298°F) using the $YBa_2Cu_3O_7$ high-temperature superconductors, and should be extendable to above 120 K (−244°F) using still higher-temperature superconductors. For even higher temperatures, a standard thermoelectric material is needed. For historical reasons, the reference material up to about room temperature has been chosen to be Pb. Until the mid-1970s, the thermopower

TABLE 1. Absolute thermoelectric power of *S* of pure lead between 0 and 350 K*

T, K†	S_{Pb}, μV/K
0.000	0.000
5.000	0.000
7.500	−0.220
8.000	−0.256
8.500	−0.298
9.000	−0.343
10.000	−0.433
11.000	−0.517
12.000	−0.593
14.000	−0.707
16.000	−0.770
18.000	−0.786
20.000	−0.779
30.000	−0.657
40.000	−0.575
50.000	−0.537
60.000	−0.527
70.000	−0.531
80.000	−0.544
90.000	−0.562
100.000	−0.583
110.000	−0.606
130.000	−0.656
150.000	−0.708
170.000	−0.760
190.000	−0.810
210.000	−0.858
230.000	−0.904
250.000	−0.948
270.000	−0.989
290.000	−1.028
310.000	−1.065
330.000	−1.101
350.000	−1.136

*After R. B. Roberts, The absolute scale of thermoelectricity, *Phil. Mag.*, 36:91–107, 1977.
†°C = K − 273.15; °F = (K × 1.8) − 459.67.

of Pb was calculated from Thomson coefficient measurements made in the early 1930s, and all published values of S were ultimately traceable to those measurements. In 1977, new measurements of the Thomson coefficient of Pb were made up to 350 K (170°F). The revised values of S_{Pb} are listed in **Table 1**. Above 50 K (−370°F) they differ from the old values by 0.25–0.3 microvolt/K; older values of S must be corrected for these differences if accuracy is important. Measurements in 1985 of the Thomson coefficients of Pt and W allow these two metals to be used as references up to 1600 K (2900°F) or 1800 K (3250°F), respectively.

Temperature variation. Figure 5 shows the variation with temperature of the thermopowers of four different pure metals. The data for gold (Au), aluminum (Al), and Pt are typical of those for most simple metals and for some transition metals as well. The thermopower S consists of a slowly varying portion that increases approximately linearly with absolute temperature, upon which is superimposed a "hump" at lower temperatures. In analyzing these results, S is written as the sum of two terms, as in Eq. (12), where

$$S = S_d + S_g \tag{12}$$

S_d, called the electron-diffusion component, is the slowly varying portion, and S_g, called the phonon-drag component, is the hump. For some transition metals, on the other hand, the behavior of S is more complex as illustrated by the data for rhodium (Rh) in Fig. 5. **Figure 6** shows comparative data for a sample p-type semiconductor. The separation of S into S_d and S_g is still valid, but at high temperatures S_d now varies more weakly than linearly with temperature. Note also the different ordinate scales in Fig. 5 (μV/K) and Fig. 6 (mV/K)—the thermopower of a semiconductor can be a thousand times larger than that of a metal.

Theory. When a small temperature difference ΔT is established across a conductor, heat is carried from

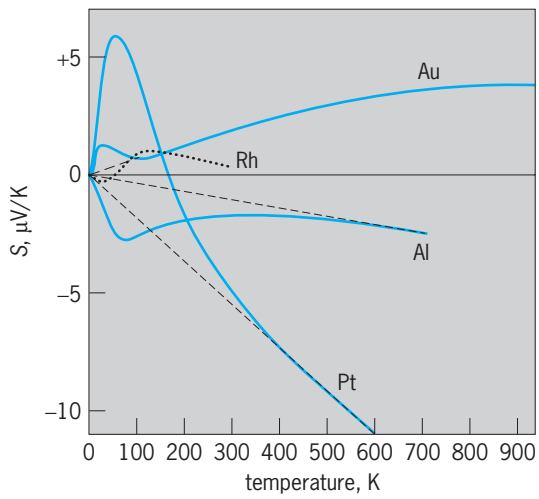

Fig. 5. Thermopower *S* of the metals gold (Au), aluminum (Al), platinum (Pt), and rhodium (Rh) as a function of temperature. The differences between the solid curves for Pt, Al, and Au and the broken lines indicate the magnitude of the phonon-drag component *S*. °F = (K × 1.8) − 459.67.

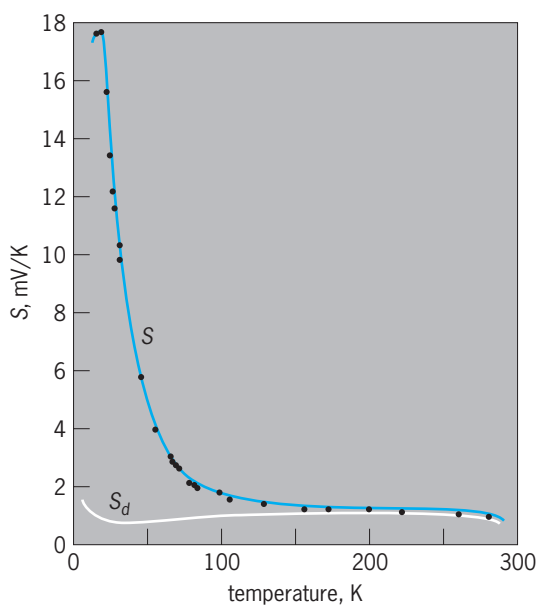

Fig. 6. Thermopower *S* of *p*-type germanium (1.5 × 10^{14} acceptors per cubic centimeter) and calculated value for the electron-diffusion thermopower S_d. (*After C. Harring, The role of low-frequency phonons in thermoelectricity and thermal conductivity, Proc. Int. Coll. 1956, Garmisch-Partenkirchen, Vieweg. Braunschweig, 1958*)

its hot end to its cold end by the flow of both electrons and phonons (quantized lattice vibrations). If the electron current is constrained to be zero—for example by the insertion of a high-resistance measuring device in series with the conductor—the electrons will redistribute themselves in space so as to produce an emf along the conductor. This is the thermoelectric emf. If the phonon flow could somehow be turned off, this emf would be just $S_d \Delta T$. However, the phonon flow cannot be turned off, and as the phonons move down the sample, they interact with the electrons and "drag" them along. This process produces an additional contribution to the emf, $S_g \Delta T$. *See* CONDUCTION (HEAT); LATTICE VIBRATIONS; PHONON; THERMAL CONDUCTION IN SOLIDS.

Source of S_d. The conduction electrons in a metal are those having energies near the Fermi energy η. Only these electrons are important for thermoelectricity. As illustrated in **Fig. 7**, in a metal, the energy distribution of these electrons varies with the temperature. At high temperatures, the metal has more high-energy electrons and fewer low-energy ones than when it is at low temperatures. This means that if a temperature gradient is established along a metal sample, the total number of electrons will remain constant, but the hot end will have more high-energy electrons than the cold end, and the cold end will have more low-energy electrons than the hot end. The high-energy electrons will diffuse toward the cold end, and the low-energy electrons will diffuse toward the hot end. However, in general, the diffusion rate is a function of electron energy, and thus a net electron current will result. This current will cause electrons to pile up at one end of the metal (usually the cold end) and thereby produce an emf that opposes the further flow of electrons. When the emf becomes large enough, the current will be reduced to zero. This equilibrium emf is the thermoelectric emf arising from electron diffusion. Essentially the same argument applies to semiconductors, except that in that case the electrons (or holes) are those just above (or just below) the band gap. *See* FREE-ELECTRON THEORY OF METALS; HOLE STATES IN SOLIDS; SEMICONDUCTOR.

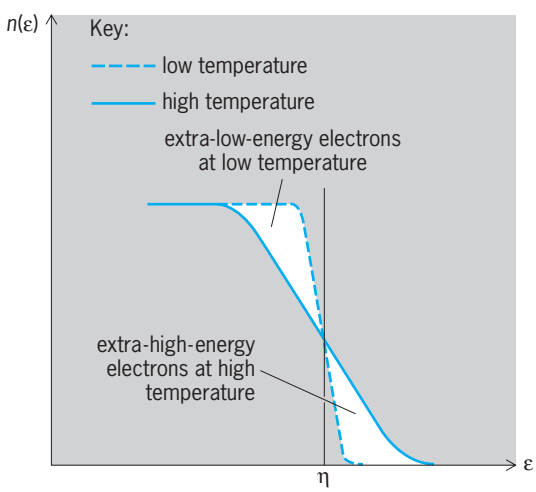

Fig. 7. Variation with energy ϵ of the number of conduction electrons $n(\varepsilon)$ in a metal in the vicinity of the Fermi energy η for two different temperatures. A small variation of η with temperature has been neglected.

TABLE 2. Comparision between theoretical values of *S* and experimental data

Metal	Thermopower (*S*), μV/K: Theoretical values at 0°C according to Eq. (13)	Thermopower (*S*), μV/K: Experimental data at approximately 0°C
Lithium (Li)	−2	+11
Sodium (Na)	−3	−6
Potassium (K)	−5	−12
Copper (Cu)	−1.5	+1.4
Gold (Au)	−2	+1.7
Aluminum (Al)	−0.7	−1.7

S_d for a metal. For a completely free-electron metal, S_d should be given by Eq. (13), where k is Boltzmann's constant, e is the charge on an electron, T is the absolute (Kelvin) temperature, and η is the Fermi energy of the metal. According to Eq. (13), S_d should be negative—since e is a negative quantity—and should increase linearly with T. In **Table 2** the predictions of Eq. (13) are compared with experiment for a number of the most free-electron-like metals. Equation (13) correctly predicts the general size of S_d, but often misses the actual value by a factor of 2 or more and in several cases predicts the wrong sign. To understand the thermopowers of real metals, it is necessary to use a more sophisticated model that takes into account interactions between the electrons in the metal and the crystal lattice, as well as scattering of the electrons by impurities and phonons. The proper generalization of Eq. (13) is Eq. (14), where

$$S_d = \frac{\pi^2 k}{2e}\left(\frac{kT}{\eta}\right) \tag{13}$$

$$S_d = \frac{\pi^2 k^2 T}{3e}\left[\frac{\delta \ln \sigma(\varepsilon)}{\delta \varepsilon}\right]_{\varepsilon=\eta} \tag{14}$$

$\sigma(\varepsilon)$ is a generalized energy-dependent conductivity defined so that $\sigma(\eta)$ is the experimental electrical conductivity of the metal, and the logarithmic derivative with respect to the energy ε is to be evaluated at $\varepsilon = \eta$. For free electrons, Eq. (14) reduces to Eq. (13). But more generally, Eq. (14) is able to account for all of the deviations of experiment from Eq. (13). If the logarithmic derivative is negative, S_d will be positive; S_d will differ in size from Eq. (13) if the logarithmic derivative does not have the free-electron value $(3/2)\eta^{-1}$; and $S_d(T)$ will deviate from a linear dependence on T if the logarithmic derivative is temperature-dependent.

In metals, research on S_d has attacked such diverse topics as understanding changes due to alloying with both magnetic and nonmagnetic impurities,

strain, application of pressure or magnetic fields, and changes near phase transitions. In some cases the changes can be dramatic. **Figure 8** shows that the addition of very small amounts of the magnetic impurity iron (Fe) can produce enormous changes in S_d for copper (Cu) at low temperatures. Sample 1 (in which the deviation of the thermopower from zero is too small to be seen with the chosen scale) is most representative of pure Cu because the Fe is present as an oxide and is thus not in "magnetic form." **Figure 9** shows that at low temperatures application of a magnetic field H to Al can cause S_d to change sign. [To obtain a temperature-independent quantity, S_d has been divided by the absolute temperature T. To remove the effects of varying impurity concentrations, H has been divided by $\rho(4.2\text{ K})nec$, where $\rho(4.2\text{ K})$ is the sample resistivity at 4.2 K, n is the number of electrons per unit volume in the metal, and c is the speed of light.] **Figure 10** illustrates the significant changes that can occur in S when a metal melts.

Substantial effort has been devoted to the study of thermoelectricity in liquid metals and liquid metal alloys. There has also been considerably interest in the thermopower of quasi-one-dimensional conductors (**Fig. 11**), in amorphous metals (also called metallic glasses or metglasses), in many-body effects in thermoelectricity, and in thermoelectric effects in inhomogeneous superconductors, such as loops consisting of two different superconducting materials. Thermoelectric effects in superconductors are much smaller than those in normal metals and are generally visible only very near T_c. Their study gives insight into nonequilibrium processes in superconductors. *See* METALLIC GLASSES.

S_d for a semiconductor. Equation (13) is appropriate for a free-electron gas that obeys Fermi-Dirac statistics. The conduction electrons in a metal obey these statistics. However, there are so few conduction electrons in a semiconductor that, to a good approximation, they can be treated as though they obey different statistics—Maxwell-Boltzmann statistics. For free electrons obeying these statistics, S_d is given by Eq. (15), which predicts that S_d should be

$$S_d = \frac{3k}{2e} \tag{15}$$

temperature-independent and have the value $S_d = -130$ times; 10^{-6} V/K. For a p-type extrinsic semiconductor, in which the carriers are approximated as free holes, the sign of S_d will be reversed to positive. Examination of the data of Fig. 6 shows that S_d is nearly independent of temperature but is considerably larger than predicted by Eq. (15). Again, a complete understanding of the thermopowers of semiconductors requires the generalization of Eq. (15). The appropriate generalizations are different for single-band and multiband semiconductors, the latter being considerably more complicated. For a single-band (extrinsic) semiconductor, the generalization is relatively straightforward and yields predictions for S_d which, in agreement with experiment, vary slowly with temperature and are several times larger than the prediction of Eq. (15). [The white curve for S_d in Fig. 6 is calculated from this generalization.] Experimental interest in the thermopower of semiconductors concerns topics similar to those for metals. In addition, the large magnitudes of the thermopowers of semiconductors continue to spur

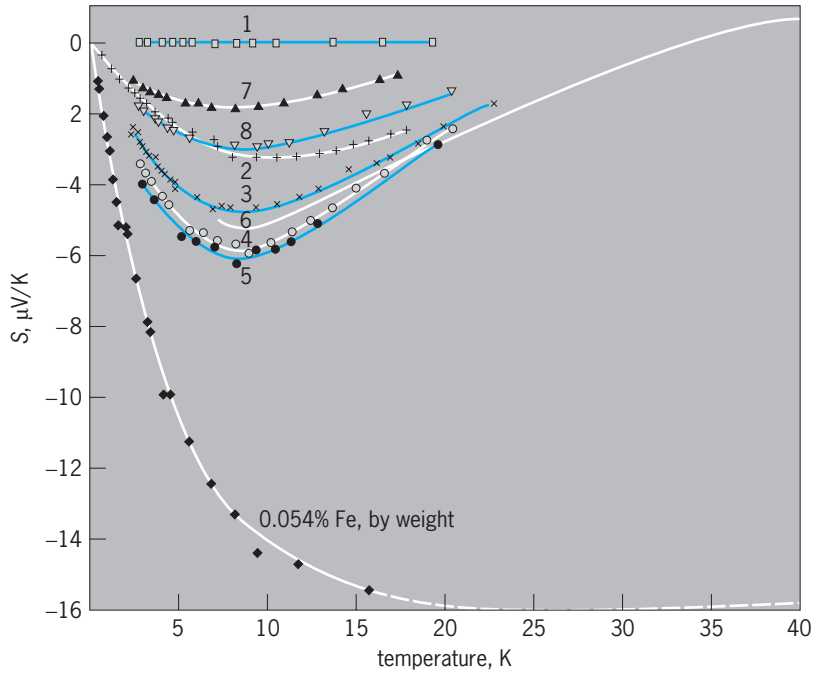

Fig. 8. Low-temperature thermopowers of various samples of copper containing very small concentrations of iron (Fe). Specific compositions of samples 1–8 are unknown. °F = (K × 1.8) − 459.67. (*After A. V. Gold et al., The thermoelectric power of pure copper, Phil. Mag., 5:765–786, 1960*)

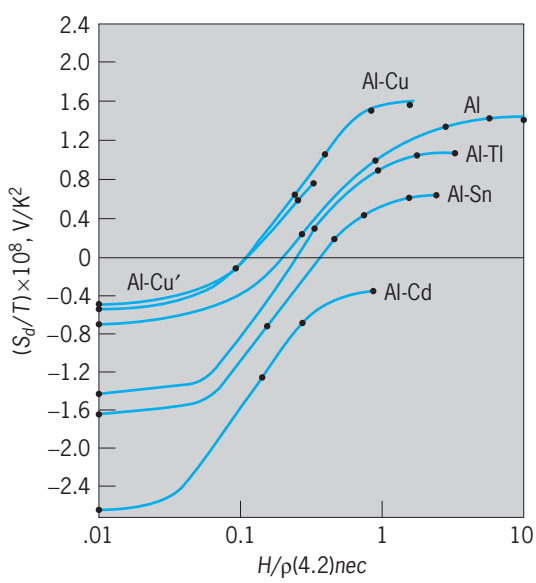

Fig. 9. Variation with magnetic field H of the low-temperature electron-diffusion thermopower S_d of aluminum (Al) and various dilute aluminum-based alloys. Sample labeled Al-Cu is a second sample of Al-Cu. (*After R. S. Averback, C. H. Stephan, and J. Bass, Magnetic field dependence of the thermopower of dilute aluminum alloys, J. Low Temp. Phys., 12:319–346, 1973*)

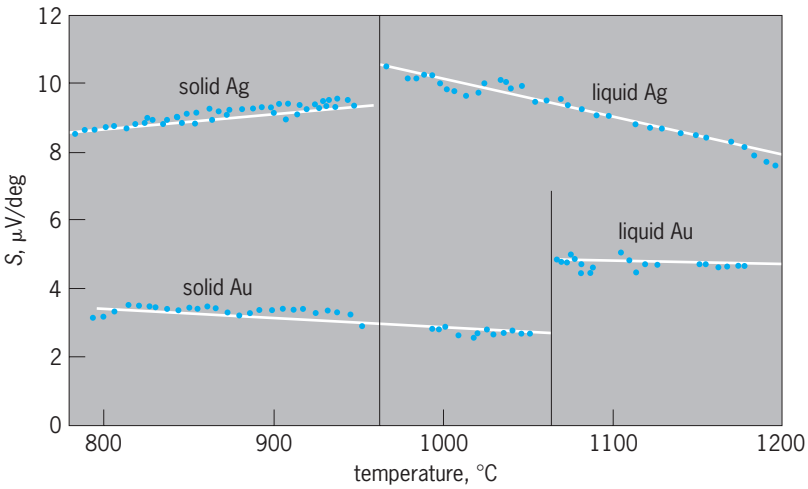

Fig. 10. Changes in the thermopowers of gold (Au) and silver (Ag) upon melting. °F = (°C × 1.8) + 32. (*After R. A. Howe and J. E. Enderby, The thermoelectric power of liquid Ag-Au, Phil. Mag., 16:467–476, 1967*)

efforts to develop materials better suited for electric power generation and thermoelectric cooling. *See* BAND THEORY OF SOLIDS; BOLTZMANN STATISTICS; FERMI-DIRAC STATISTICS.

Source of S_g. Unlike the behavior of S_d, which is determined in both metals and semiconductors primarily by the properties of the charge carriers, the behavior of S_g is determined in both cases primarily by the properties of the phonons. At low temperatures, phonons scatter mainly from electrons or impurities rather than from other phonons. The initial increase in S_g with increasing temperature at the very lowest temperatures in Figs. 5 and 6 results from an increasing number of phonons becoming available to drag the electrons along. However, as the temperature increases, the phonons begin to scatter more frequently from each other. Eventually, phonon-phonon scattering becomes dominant, the electrons are no longer dragged along, and S_g falls off in magnitude with increasing temperature. Interest in phonon drag is associated with such questions as whether it is the sole source of the humps in Figs. 5 and 6, how it changes as impurities are added, and how it is affected by a magnetic field.

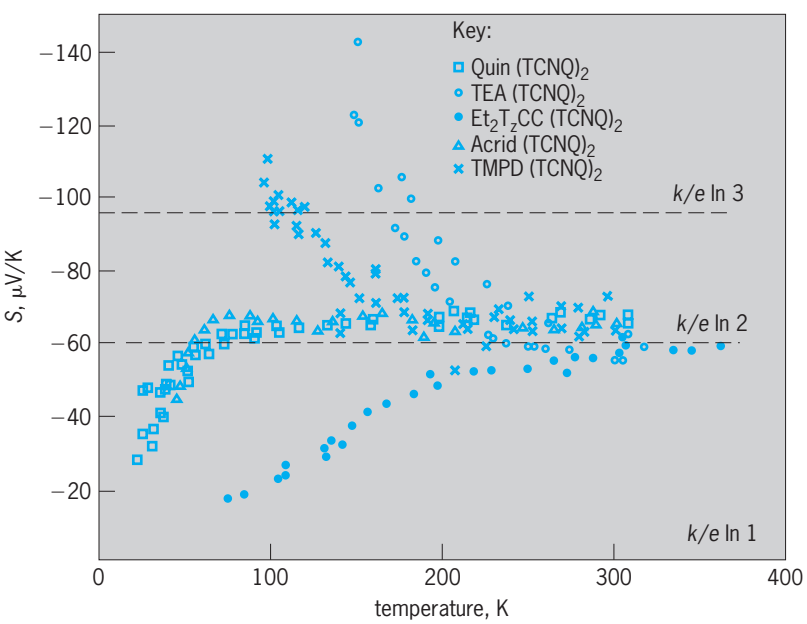

Fig. 11. Thermopower of highly conducting salts of the form (Donor)$^+$ (TCNQ)$_2^-$. By 300 K (80°F) all of the thermopowers are very close to the "entropy per carrier" of (k/e) ln 2 = −60 μV/K, where k is Boltzmann's constant and e is the electron charge. °F = (°K × 1.8) − 459.67. (*After F. J. Blatt and P. A. Schroeder, eds., Thermoelectricity in Metallic Conductors, Plenum Press, 1978*)

Applications

The most important practical application of thermoelectric phenomena to date is in the accurate measurement of temperature. The phenomenon involved is the Seebeck effect. Of less importance so far are the direct generation of electrical power by application of heat (also involving the Seebeck effect) and thermoelectric cooling and heating (involving the Peltier effect).

A basic system suitable for all four applications is illustrated schematically in **Fig. 12**. Several thermocouples are connected together in series to form a thermopile, a device with increased output (for power generation or cooling and heating) or sensitivity (for temperature measurement) relative to a single thermocouple. The junctions forming one end of the thermopile are all at the same low temperature T_L, and the junctions forming the other end are at the high temperature T_H. The thermopile is connected to a device D that differs for each application. For temperature measurement, the temperature T_L is fixed, for example, by means of a bath; the temperature T_H becomes the running temperature T that is to be measured; and the device is a potentiometer for measuring the thermoelectric emf generated by the thermopile. For power generation, the temperature T_L is fixed by connection to a heat sink; the temperature T_H is fixed at a value determined by the output of the heat source and the thermal conductance of the thermopile; and the device is whatever is to be run by the electricity that is generated. For heating or cooling, the device is a current generator that passes current through the thermopile. If the current flows in the proper direction, the junctions at T_H will heat up and those at T_L will cool down. If T_H is fixed by connection to a heat sink, thermoelectric cooling will be provided at T_L. Alternatively, if T_L is fixed, thermoelectric heating will be provided at T_H. If the heat sink is room-temperature, such a system has the advantage that at any given location it can be converted from a cooler to a heater merely by reversing the direction of the current.

Temperature measurement. In principle, any material property that varies with temperature can serve as the basis for a thermometer. In practice, the two properties most often used for precision thermometry are electrical resistance and thermoelectric emf. Thermocouples are widely employed to measure temperature in both scientific research and industrial processes. In the United States alone,

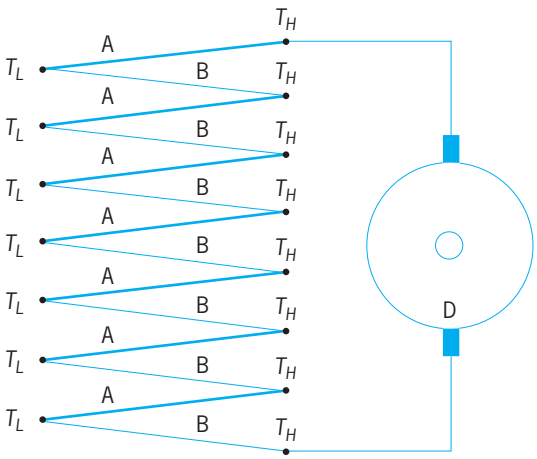

Fig. 12. Thermopile, a battery of thermocouples connected in series; *D* is a device appropriate to the particular application.

several hundred tons of thermocouple materials are produced annually. *See* TEMPERATURE MEASUREMENT; THERMOCOUPLE.

Construction of instruments. In spite of their smaller thermopowers, metals are usually preferred to semiconductors for precision temperature measurements because they are cheaper, are easier to fabricate into convenient forms such as thin wires, and have more reproducible thermoelectric properties. With modern potentometric systems, standard metallic thermocouples provide temperature sensitivity adequate for most needs—small fractions of a degree Celsius are routinely obtained. If greater sensitivity is required, several thermocouples can be connected in series to form a thermopile (Fig. 12). A 10-element thermopile provides a temperature sensitivity 10 times as great as that of each of its constituent thermocouples. However, the effects of any inhomogeneities are also enhanced 10 times.

The thermocouple system standardly used to measure temperature is shown in **Fig. 13**. It consists of wires of three metals, A, B, and C, where C is usually the metal Cu. The junction between the wires of metals A and B is located at the temperature to be measured T. Each of these two wires is joined to a wire of metal C at the reference temperature T_0. The other ends of the two wires of metal C are connected to the potentiometer at room temperature T_r. Integrating the appropriate thermopowers around the circuit of Fig. 13 yields the total thermoelectric emf E in terms of the separate emf's generated by each of the four pieces of wire, as given in Eq. (16).

$$E = E_A(T_0, T) - E_B(T_0, T) + E_{C1}(T_0, T_r) - E_{C2}(T_0, T_r) \qquad (16)$$

If the two wires C1 and C2 have identical thermoelectric characteristics, the last two terms in Eq. (16) cancel, and, with the use of Eq. (5), Eq. (17) results.

$$E = E_A(T_0, T) - E_B(T_0, T) = E_{AB}(T_0, T) \qquad (17)$$

That is, two matched pieces of metal C produce no contribution to the thermoelectric emf of the circuit shown in Fig. 13, provided their ends are maintained at exactly the same two temperatures. This means that it is not necessary to use either of the sometimes expensive metals making up the thermocouple to go from the reference-temperature bath to the potentiometer. That part of the circuit can be constructed of any uniform, homogeneous metal. Copper is often used because it is inexpensive, is available in adequate purity to ensure uniform, homogeneous samples when handled with care, can be obtained in a wide variety of wire diameters, and can be either spotwelded or soldered to the ends of the thermocouple wires. Special low-thermal emf alloys are available for making solder connections to Cu in thermocouple circuits.

Choice of materials. Characteristics that make a thermocouple suitable as a general-purpose thermometer include adequate sensitivity over a wide temperature range, stability against physical and chemical change under various conditions of use and over extended periods of time, availability in a selection of wire diameters, and moderate cost. No single pair of thermocouple materials satisfies all needs. Platinum versus platinum–10% rhodium can be used up to 1700°C (3100°F). A thermocouple combining the two alloys chromel and alumel gives greater sensitivity and an emf that is closely linear with temperature, but cannot be used to as high a temperature. A combination of Cu versus the alloy constantan also has high sensitivity above room temperature and adequate sensitivity down to as low as 15 K (−433°F). For temperatures of 4 K (−452°F) or lower, special gold-cobalt alloys versus Cu or gold-iron alloys versus chromel are used.

Thermocouple tables. To use a thermocouple composed of metals A and B as a thermometer, it is necessary to know how $E_{AB}(T_0,T)$ varies with temperature T for some reference temperature T_0. According to Eq. (6), $E_{AB}(T_0,T_1)$ can be determined for

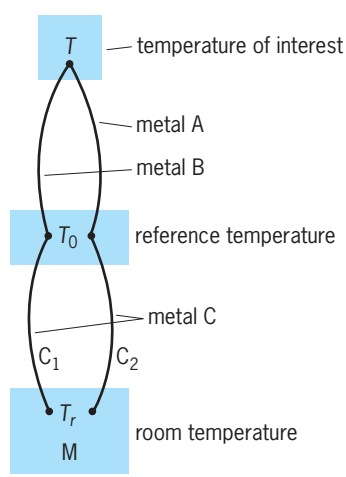

Fig. 13. Thermocouple system standardly used to measure temperature; M is a measuring device, usually a potentiometer, which is at room temperature.

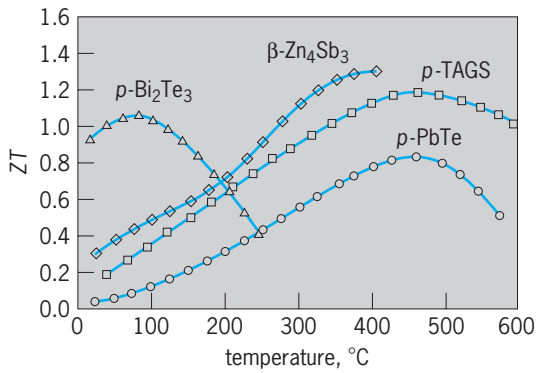

Fig. 14. Representative values of the dimensionless figure of merit *ZT* as a function of temperature for *p*-type β-Zn_4Sb_3 (diamonds). These values are to be compared with those for state-of-the-art *p*-type thermoelectric materials: PbTe- and Bi_2Te_3-based alloys and TAGS (Te-Ag-Ge-Sb) alloys. °F = (°C × 1.8) + 32. (*From T. Caillat, J.-P. Fleurial, and A. Borschevsky, A low thermal conductivity compound for thermoelectric applications: β-Zn_4Sb_3, 15th IEEE International Conference on Thermoelectrics, p. 151, 1996*)

any two temperatures T_0 and T_1 if both $S_A(T)$ and $S_B(T)$ are known for all temperatures between T_0 and T_1. Knowledge of $S_A(T)$ and $S_B(T)$ allows construction of a table of values for $E_{AB}(T_0,T)$ using any arbitrary reference temperature T_0. Such tables are available for the thermocouples mentioned above, and for some others as well, usually with a reference temperature of 0°C (32°F). A table of $E_{AB}(T_0,T)$ for one reference temperature T_0 can be converted into a table for any other reference temperature T_2 merely by subtracting a constant value $E_{AB}(T_0,T_2)$ from each entry in the table to give Eq. (18). Here $E_{AB}(T_0,T_2)$

$$E_{AB}(T_2, T) = E_{AB}(T_0, T) - E_{AB}(T_0, T_2) \qquad (18)$$

is a positive quantity when T_2 is greater than T_0 and when $S_{AB}(T)$ is positive between T_0 and T_2.

Other uses. Thermoelectric power generators, heaters, or coolers made from even the best presently available materials have the disadvantages of relatively low efficiencies and concomitant high cost per unit of output. Their use has therefore been largely restricted to situations in which these disadvantages are outweighed by such advantages as small size, low maintenance due to lack of moving parts, quiet and vibration-free performance, light weight, and long life.

Figure of merit. A measure of the utility of a given thermoelectric material for power generation, cooling, or heating at an absolute temperature T is provided by the dimensionless figure of merit ZT given by Eq. (19). Here S is the thermopower of the material, σ is its electrical conductivity, and κ is its thermal conductivity. The largest values for ZT are obtained in semimetals and highly doped semiconductors, which are the materials normally used in practical thermoelectric devices. As illustrated in **Fig. 14**, for most materials ZT varies with temperature, going through maxima at different temperatures. Thus, combining available materials into thermocouples often results in values of ZT too small to be competitive over a wide enough temperature range to be useful. The best available values of $ZT \approx 1$ yield conversion efficiencies of only a few percent. Values of $ZT > 2$ over a wide enough temperature range could make thermoelectrics competitive for some uses described below, and values of $ZT \geq 4$ over wide temperature ranges in materials with high-temperature stability and affordable cost might revolutionize heating, cooling, and power generation. *See* THERMOELECTRIC POWER GENERATOR.

$$ZT = \frac{S^2 \sigma T}{\kappa} \qquad (19)$$

For a long time, little progress was made in increasing ZT beyond the values for the established BiTe- and PbTe-based thermoelectric materials shown in Fig. 14, although no rigorous theoretical limit on the value of ZT is known. However, research on thermoelectric materials has accelerated due to the discovery of values of $ZT \geq 1$ in β-Zn_4Sb_3 (Fig. 14) and ternary filled skutterudites (**Fig. 15**) of the form LnT_4Pn_{12} (Ln = rare earth or Th; T = Fe, Ru, Os, Co, Rh, Ir; Pn = P, As, Sb), as well as new ideas that might lead to larger ZT.

The quantity of importance in a thermoelectric device is the figure of merit of the thermocouple rather than the separate figures of merit of its constituents. Although at least one constituent should have a high figure of merit, two constituents with high figures of merit do not guarantee that the figure of merit of the thermocouple will be high. For example, if the thermopowers of the two constituents are identical, the figure of merit of the couple will be zero.

Just as the figures of merit for single materials vary with temperature, so do the figures of merit for thermocouples formed from two such materials. This means that one thermocouple can be better than another in one temperature range but less efficient in another. To take maximum advantage of the different properties of different couples, thermocouples are often cascaded (**Fig. 16**). Cascading produces power generation in stages, the higher temperature

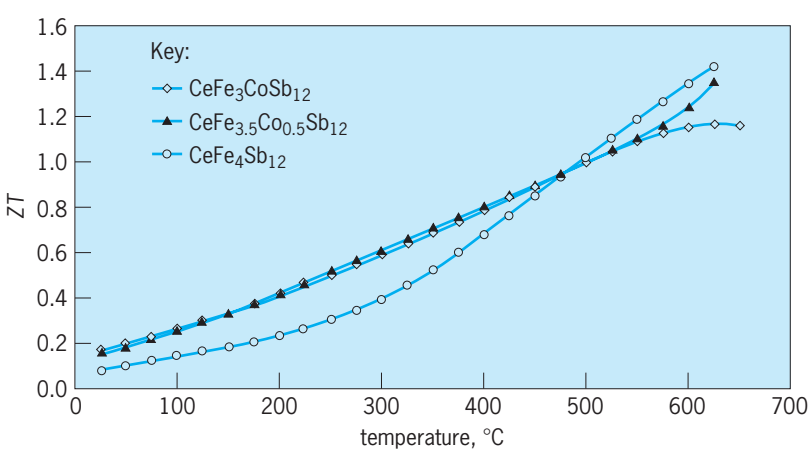

Fig. 15. Dimensionless figure of merit *ZT* as a function of temperature for cerium (Ce)-filled skutterudite samples with different compositions. °F = (°C × 1.8) + 32. (*From J.-P. Fleurial et al., High figure of merit in Ce-filled skutterudites, 15th IEEE International Conference on Thermoelectrics, p. 91, 1996*)

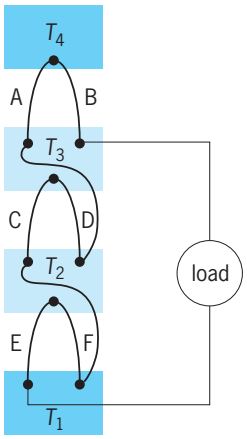

Fig. 16. Three-level cascade consisting of three different thermocouples (A versus B, C versus D, and E versus F) at four temperatures (*T*).

of each stage being determined by the heat rejected from the stage above. Thus, the highest and lowest temperatures T_4 and T_1 are fixed by connection to external reservoirs, whereas the middle temperatures T_3 and T_2 are determined by the properties of the materials. By cascading, a series of thermocouples can be used simultaneously in the temperature ranges where their figures of merit are highest. Cascaded thermocouple systems have achieved conversion efficiencies as high as 10–15%.

Thermoelectric generators. A thermoelectric generator requires a heat source and a thermocouple. Kerosine lamps and firewood have been used as heat sources to produce a few watts of electricity in remote locations, and systems using motor fuel burners have produced 100 W. Test systems using sunlight have also been constructed. Radioactive sources, especially strontium-90 (^{90}Sr) and plutonium-238 (^{238}Pu), have provided the heat to activate small rugged thermoelectric batteries for use in lighthouses, in navigation buoys, at isolated weather stations or oil platforms, and in spaceships. Such small nuclear batteries operate pacemakers implanted in humans and data transmission systems for deep-space planetary probes. Thermopiles based on silicon-germanium (SiGe) alloys and powered by plutonium-238 supplied more than 100 W of power at an efficiency of over 6% to the *Voyager 1* and *Voyager 2* spacecraft for the 12 years (1977–1989) of their missions to the outer planets, and are expected to remain electrically active until at least the year 2015. Nuclear-powered batteries for medical use must be designed to remain intact following the maximum credible accident. Capabilities such as retention of integrity after crushing by 1 ton, or impact at 50 m/s (112 mi/h), or salt-water corrosion for centuries, or cremation at temperatures up to 1300°C (2400°F) for half an hour are required. Investigations have been made of the feasibility of thermoelectric generation using the copious heat generated by nuclear reactors, or the heat generated in the exhaust system of automobiles. Such uses would require the development of more efficient thermoelectric materials able to operate for a long time at the high temperatures that are encountered. *See* NUCLEAR BATTERY; NUCLEAR POWER; RADIOACTIVITY; SPACE POWER SYSTEMS.

Peltier cooling. With available values of *ZT*, thermoelectric refrigerators suitable for use in homes or automobiles are more expensive and less efficient than standard vapor-compression-cycle refrigerators. Their use is thus largely restricted to situations in which lower maintenance, increased life, absence of vibration, or quiet performance are essential, or to situations (as in space vehicles or artificial satellites) in which the compressor type of refrigerator is impractical. A number are in use in hotels and other large facilities. A typical unit of about 50-liter (13-gallon) capacity requires a dc power input of 40 W, has a refrigerative capacity of 20 kcal/h (23 W), and a time to cool of 4–5 h. Larger *ZT* would also make thermoelectricity competitive for cooling of high-power electronic devices. *See* REFRIGERATION.

For lower temperatures, the proper choice of thermoelectric materials and the use of cascading can result in a reduction in temperature at the coldest junctions of as much as 150°C (270°F). Temperature drops of 100°C (180°F) have been obtained in single crystals of the semimetal bismuth through use of the thermomagnetic Ettingshausen effect. A large enough *ZT* down to temperatures of −200°C (−328°F) could allow widespread use of new high-temperature superconductors in electronic devices. *See* THERMOMAGNETIC EFFECTS.

Small cooling units with capacities of 10 W or less have been developed for miscellaneous applications such as cold traps for vacuum systems, cooling controls for thermocouple reference junctions, cooling devices for scientific equipment such as infrared detectors, cold stages on microscopes or on microtomes used for sectioning cooled tissues, and electronic equipment coolers. However, major commercial success of thermoelectric refrigeration requires thermocouple materials with higher values of *ZT*.

Thermoelectric heating. A thermoelectric heater referenced to room temperature is nothing more than a thermoelectric refrigerator with the current reversed. No large heaters have been marketed. However, various small household convenience devices have been developed, such as a baby-bottle cooler-warmer that cools the bottle until just before feeding time and then automatically switches to a heating cycle to warm it, and a thermoelectric hostess cart. *See* ELECTRICITY.

Jack Bass

Bibliography. F. J. Blatt et al., *Thermoelectric Power of Metals*, Plenum Press, New York, 1976; F. J. Blatt and P. A. Schroeder (eds.), *Thermoelectricity in Metallic Conductors*, Plenum Press, New York, 1978; F. J. DiSalvo, Thermoelectric cooling and power generation, *Science*, 285:703–706, 1999; D. M. Rowe (ed.), *CRC Handbook of Thermoelectrics*, CRC Press, Boca Raton, FL, 1995; *Proceedings of the 15th IEEE International Conference on Thermoelectrics*, IEEE Catalog no. 96TH8169, 1996; T. M. Tritt et al. (eds.), Thermoelectric Materials 1998—The Next

Generation Materials for Small-Scale Refrigeration and Power Generation Applications, *MRS Soc. Proc.*, vol. 545, 1998.

Thermoluminescence

The emission of light when certain solids are warmed, generally to a temperature lower than that needed to provoke visible incandescence. Two characteristics of thermoluminescence distinguish it from incandescence. First, the intensity of thermoluminescent emission does not remain constant at constant temperature, but decreases with time and eventually ceases altogether. Second, the spectrum of the thermoluminescence is highly dependent on the composition of the material and is only slightly affected by the temperature of heating. If a thermoluminescent material emits both thermoluminescence and incandescent light at some temperature of observation, the transient light emission is the thermoluminescence and the remaining steady-state emission is the incandescence. The transient nature of the thermoluminescent emission suggests that heating merely triggers the release of stored energy previously imparted to the material. Supporting this interpretation is the fact that after the thermoluminescence has been reduced to zero by heating, the sample can be made thermoluminescent again by exposure to one of a number of energy sources: x-rays and gamma rays, electron beams, nuclear particles, ultraviolet light, and, in some cases, even shortwave visible light (violet and blue). A thermoluminescent material, therefore, has a memory of its earlier exposure to an energizing source, and this memory is utilized in a number of applications. Many natural minerals are thermoluminescent, but the most efficient materials of this type are specially formulated synthetic solids (phosphors). *See* LUMINESCENCE.

Mechanism. In addition to special sites capable of emitting light (luminescent centers), thermoluminescent phosphors have centers that can trap electrons or holes when these are produced in the solid by ionizing radiation. The luminescent center itself is often the hole trap, and the electron is trapped at another center, although the reverse situation can also occur. In the former case, if the temperature is low and the energy required to release an electron from a trap (the trap depth) is large, electrons will remain trapped and no luminescence will occur. If, however, the temperature of the phosphor is progressively raised, electrons will receive increasing amounts of thermal energy and will have an increased probability of escape from the traps. Freed electrons may then go over to luminescent centers and recombine with holes trapped at or near these centers. The energy liberated by the recombination can excite the luminescent centers, causing them to emit light. *See* HOLE STATES IN SOLIDS; TRAPS IN SOLIDS.

Glow curves. A plot of luminescence intensity versus temperature (or time, if a constant heating rate is employed) is called a glow curve. The initial rise of intensity with temperature is due to the increased number of electrons escaping from traps as the temperature is raised. When all traps are of the same depth, the intensity peaks at some temperature due to the decrease in population of the trapped electrons, and it finally drops to zero when all the traps are emptied. The exact shape of the glow curve depends on the heating rate. The higher the heating rate, the brighter is the thermoluminescent glow, but the shorter its duration. The total number of light quanta emitted (the light sum) is generally independent of the heating rate, however.

When the glow curve consists of a single peak, corresponding to a single species of trap, the trap depth E (in joules or other energy units) is given to a good approximation by the equation below, where

$$E = \frac{1.51\, kT^*T'}{(T^* - T')}$$

k is Boltzmann's constant, T^* is the temperature (in kelvins) of the phosphor at the peak of the curve, and T' is the temperature on the low-temperature side at which the emission is one-half its peak value. The **illustration** shows single-peaked glow curves of this type for several zinc sulfide phosphors, all of which thermoluminesce green due to the presence of copper (the activator) in the luminescent center. Traps of different depths are produced by different trivalent impurities (coactivators). In thermoluminescent materials containing more than one type of trapping center, the glow curve comprises a corresponding number of peaks, which often may be resolved and analyzed as described above. Thermoluminescence may thus be used to obtain information about the properties of traps in solids.

Applications. Radiation dosimeters based on thermoluminescence are widely used for monitoring integrated radiation exposure in nuclear power plants, hospitals, and other installations where high-energy radiations are likely to be encountered. The key elements of the dosimeters, thermoluminescent phosphors with deep traps, can store some of the energy

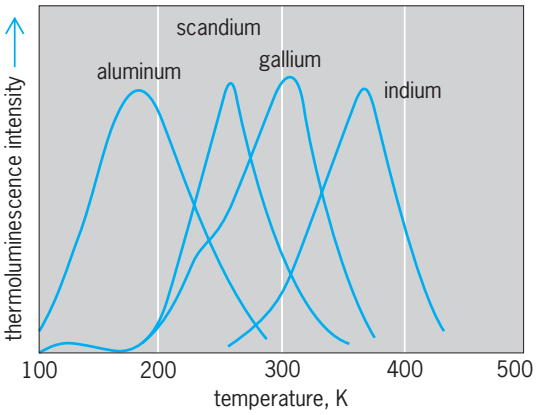

Glow curves for the green luminescence of several zinc sulfide phosphors, each of which contains traces of copper and different trivalent ions. Luminescent center is due to presence of copper, the activator in each case. Traps are put in by various trivalent coactivators as shown. °F = (K × 1.8) − 459.67.

absorbed from these radiations for very long periods of time at normal temperatures and release it as luminescence on demand when appropriately heated. The brightness (or light sum) of the luminescence is a measure of the original radiation dose. *See* DOSIMETER.

Thermoluminescence induced by ionizing radiation also offers a method for archeological and possibly geological dating. Archeological dating is based on the principle that the high-temperature firing during the preparation of a ceramic object empties all the electron and hole traps in the material. It is further assumed that during the burial of a potsherd it has been exposed to a constant rate of irradiation from its self-contained isotopes, radioactive minerals in the ground, and cosmic rays. The dose accrued by the archeological specimen is determined by measuring its glow curve and calibrating the observed thermoluminescent emission against a portion of the sample that has been given a known dose of appropriate radiation. The annual dose to the potsherd from its own constituents and the surrounding environment must be estimated by separate measurements of their radioactivities. From these data the specimen age can be derived. *See* DATING METHODS.

Although the basic principle is thus quite simple, thermoluminescent dating involves a number of complex considerations, in particular, the determination of the annual dose. A number of different procedures have been developed, however, to deal with the various complications.

Thermoluminescence has also been used to evaluate radiation doses received by survivors of atomic bomb radiation in Hiroshima and Nagasaki, Japan, decades after the events, by using bricks and tiles from buildings that had been exposed to the radiation. James H. Schulman

Bibliography. M. J. Aitken, *Thermoluminescence Dating*, 1985; G. Lalos (ed.), *Calibration Handbook: Ionizing Radiation Measuring Instruments*, 1984; S. W. S. McKeever, *Thermoluminescence of Solids*, 1988; W. C. Roesch (ed.), *Final Report on U.S.-Japan Reassessment of Atomic Bomb Radiation Dosimetry in Hiroshima and Nagasaki*, vol. 1, 1986; D. R. Vij, *Thermoluminescent Materials*, 1993.

Thermomagnetic effects

Electrical or thermal phenomena occurring when a conductor or semiconductor which is carrying a thermal current (that is, is in a temperature gradient) is placed in a magnetic field. *See* SEMICONDUCTOR.

Let the temperature gradient be transverse to the magnetic field H_z, for example, along x. Then the following transverse-transverse effects are observed:

1. Ettingshausen-Nernst effect, an electric field along y, as in Eq. (1), where Q is known as the Ettingshausen-Nernst coefficient. This coefficient is related to the Ettingshausen coefficient P by Eq. (2),

$$E_y = Q\frac{\partial T}{\partial x}H_z \tag{1}$$

$$P = QT\sigma \tag{2}$$

where σ is the thermal conductivity in a transverse magnetic field. This relation was discovered by P. W. Bridgman; it has been shown to be an example of the Onsager reciprocity relations of irreversible thermodynamics. *See* CHEMICAL THERMODYNAMICS; GALVANOMAGNETIC EFFECTS; THERMOELECTRICITY.

2. Righi-Leduc effect, a temperature gradient along y, as in Eq. (3), where S is known as the Righi-Leduc coefficient.

$$\frac{\partial T}{\partial y} = S\frac{\partial T}{\partial x}H_z \tag{3}$$

Also, the following transverse-longitudinal effects are observed:

3. An electric potential change along x, amounting to a change of thermoelectric power.

4. A temperature gradient change along x, amounting to a change of thermal resistance.

Let the temperature gradient be along H. Then changes in thermoelectric power and in thermal conductivity are observed in the direction of H.

For related phenomena *see* HALL EFFECT; MAGNETORESISTANCE. Elihu Abrahams; Frederic Keffer

Thermometer

An instrument that measures temperature. Although this broad definition includes all temperature-measuring devices, they are not all called thermometers. Other names have been generally adopted. For a discussion of two such devices *see* PYROMETER; THERMOCOUPLE. For a general discussion of temperature measurement *see* TEMPERATURE MEASUREMENT.

A variety of techniques and designs are used in instruments known as thermometers. Some of these depend on the expansion of a liquid or metal for the indicating means. Others employ the change in pressure of a gas to detect the temperature. Still others use the change in electrical resistance which occurs with temperature changes.

Liquid-in-glass thermometer. This thermometer consists of a liquid-filled glass bulb and a connecting partially filled capillary tube. When the temperature of the thermometer increases, the differential expansion between the glass and the liquid causes the liquid to rise in the capillary. In **Fig. 1***a* the graduations are etched on the glass stem. The thermometer in Fig. 1*b* has a separate graduated scale similar to that of the common household thermometer. A variety of liquids, such as mercury, alcohol, toluene, and pentane, and a number of different glasses are used in thermometer construction, so that various designs cover diverse ranges between about −300 and +1200°F (−185 and +650°C). The range covered by the ordinary mercury-in-glass thermometers is from about −38 to +500°F (−39 to +260°C), and with a special glass and an inert gas in the capillary tube the

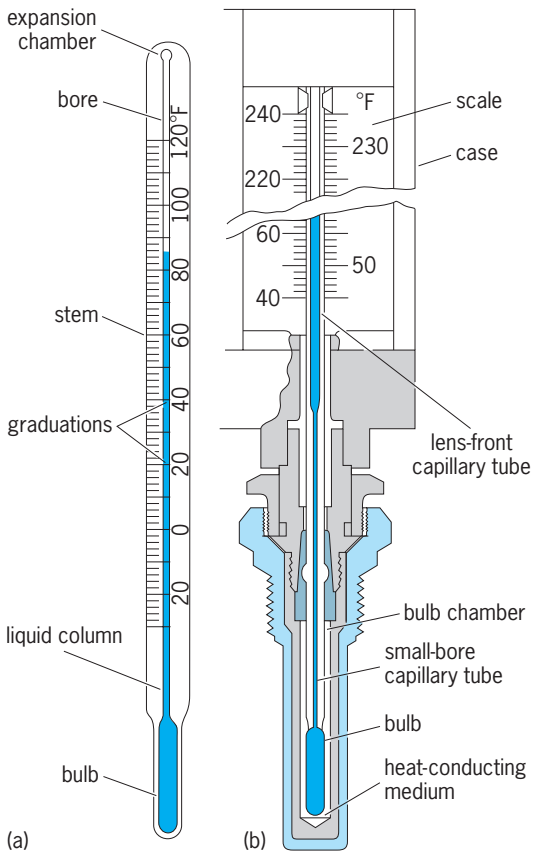

Fig. 1. Liquid-in-glass thermometers. (*a*) Etched-stem clinical thermometer. (*b*) Graduated-scale industrial thermometer. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

upper limit may be raised to about 1100°F (600°C). The usual ranges covered with the other liquids are: ethyl alcohol, −110 to +160°F (−79 to +71°C); and pentane, −330 to +85°F (−201 to +29°C).

Expansion and contraction chambers are sometimes provided at each end of the capillary to permit overrange and underrange use of the thermometer without loss of accuracy. When the entire thermometer is not subjected to the same temperature, an error occurs unless the thermometer is calibrated for these conditions. Many thermometers have an emergent stem and are calibrated for this type of service.

Maximum-registration thermometers, such as a fever thermometer, allow expansion of the liquid with increasing temperature but maintain the liquid column as the temperature decreases, thereby indicating the maximum value attained. Minimum registering thermometers are also available. If the liquid filling is metallic, electric contacts can be mounted in the stem wall to complete a circuit when temperature reaches a specified value. These are used in alarm and control systems. Some thermometers destined for rugged service are enclosed in a metal sleeve or armor.

Accuracies and speeds of response vary widely with the designs, ranges, and installations. For example, the Beckmann thermometer can be easily read to an accuracy (stated in terms of maximum error) of 0.002°F (0.001°C) and the fever thermometer to 0.1°F (0.05°C). Industrial thermometers and armored thermometers are seldom more accurate than 1°F (0.5°C) as they are used. Mercury and glass thermometers have time constants as low as 0.1 s in well-stirred water, but industrial thermometers and all thermometers installed in wells may have time constants as long as 1 min.

Bimetallic thermometer. In this thermometer the differential expansion of thin dissimilar metals, bonded together into a narrow strip and coiled into the shape of a helix or spiral, is used to actuate a pointer (**Fig. 2**). In some designs the pointer is replaced with low-voltage contacts to control, through relays, operations which depend upon temperature, such as furnace controls. The designs of cases range from those for light laboratory service to those for heavy industrial use. Range spans are seldom shorter than 50°F (28°C) or longer than 400°F (220°C), with a maximum upper temperature limit for continuous service of 800°F (425°C) and a minimum of −300°F (−185°C). The shorter range spans are used near room temperatures, and accuracies in the neighborhood of 1°F can be achieved. At high and low temperatures the accuracy is seldom better than 5°F. The time constants of these thermometers are greater but of the same order of magnitude as the liquid-in-glass thermometer.

Filled-system thermometer. This type of thermometer, shown schematically in **Fig. 3**, has a bourdon

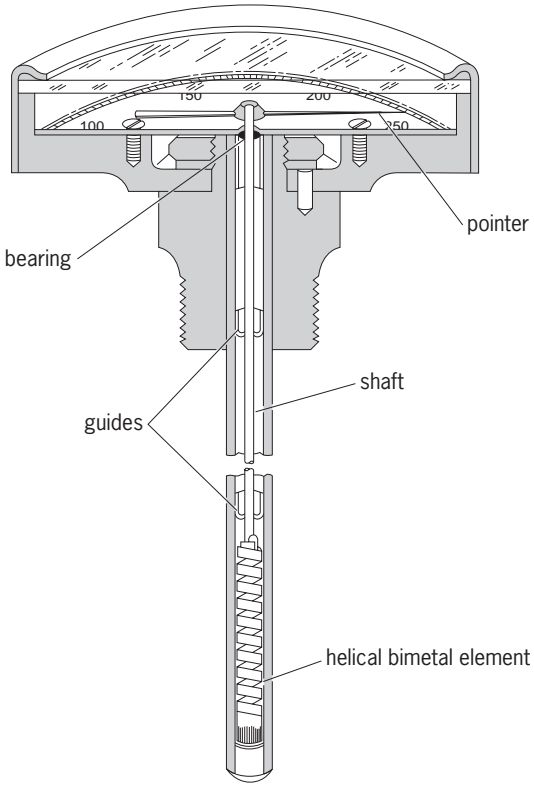

Fig. 2. Bimetallic thermometer. (*Weston Instruments, Division of Daystrom, Inc.*)

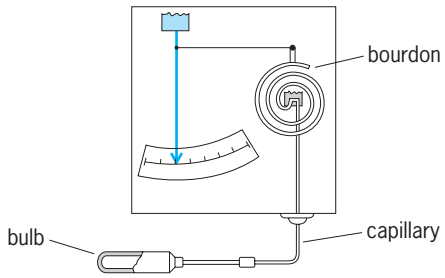

Fig. 3. Filled-system thermometer. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

tube connected by a capillary tube to a hollow bulb. When the system is designed for and filled with a gas (usually nitrogen or helium) the pressure in the system substantially follows the gas law, and a temperature indication is obtained from the bourdon tube. The temperature-pressure-motion relationship is nearly linear. Atmospheric pressure effects are minimized by filling the system to a high pressure. When the system is designed for and filled with a liquid, the volume change of the liquid actuates the bourdon tube. When mercury or its alloys are used as a filling medium, the temperature-volume-motion relationship is substantially linear. When hydrocarbon liquids are used, the liquid compressibility is appreciable, and the temperature-motion relationship is not so linear.

Since the fluids (liquid or gas) are homogeneous and extend to the bourdon tube, temperature changes on the capillary and on the bourdon tube will cause errors. These are made small by minimizing the volume in the capillary and bourdon tube and by providing ambient-temperature compensation. This compensation can be a duplicate system without a bulb to subtract the effect of the error; it can be a bimetallic compensator on the bourdon tube alone; or in the case of the mercury system a special capillary may be threaded with an invar wire compensator. The gas system has a relatively large bulb and a long range span (about 200°F minimum at room temperatures, 400°F near 1000°F; about 110°C minimum at room temperatures, 240°C near 540°C); the span may extend to a lower limit of about −400°F (−240°C) and an upper limit of about 1200°F (650°C). Hydrocarbon liquid systems have small bulbs and short range spans (as low as 25°F or 14°C); the span may extend to a lower limit of about −125°F (−87°C) and an upper limit of about 600°F (315°C). Mercury systems have somewhat larger bulbs (because of mercury's low temperature coefficient of expansion) and longer range spans and are used at temperatures between −40°F (−40°C) and 1200°F (650°C). Normally, accuracies of 1% of the range span are obtained from these instruments, but this is achieved only by proper selection with full knowledge of application conditions.

Vapor-pressure thermal system. This filled-system thermometer utilizes the vapor pressure of certain stable liquids to measure temperature, as shown in **Fig. 4**. The useful portion of any liquid-vapor pressure curve is between approximately 15 psia (100 kilopascals absolute) and the critical pressure, that is, the vapor pressure at the critical temperature, which is the highest temperature for a particular liquid-vapor system. Thus, when propane is used, the critical temperature is 206°F (97°C) and the vapor pressure is 617 psia (4.254 megapascals). A nonlinear relationship exists between the temperature and the vapor pressure, so the motion of the bourdon tube is greater at the upper end of the vapor-pressure curve. Therefore, these thermal systems are normally used near the upper end of their range, and an accuracy of 1% or better can be expected. Vapor-pressure systems are designed so that the active liquid-vapor interface occurs in the bulb, and the effective temperature occurs at this interface. There is no error due to ambient temperature changes on the capillary, and only the temperature effect on the metal bourdon tube produces an error at this point. The bourdon-tube error is normally small and may be compensated (bimetallic) if it must be reduced.

When the bulb and the bourdon tube are not at the same level a hydrostatic error occurs, but this is easily removed by zero setting. The effect of atmospheric pressure variations is minimized by using only the elevated portion of the vapor-pressure curve of the various liquids. Range spans vary widely, but near room temperature the useful portion of the span is about 120°F (67°C), and at elevated temperatures it is 200°F (110°C). Few vapor-pressure systems are used below 0°F (−18°C) and above 650°F (343°C).

The greatest advantage of the filled-system thermometer is its ability to provide a good, low-cost, temperature indication or record at a convenient point reasonably remote (up to 200 ft or 60 m) from the temperature being measured. The bourdon tube is powerful enough to operate sensitive detectors, the output of which can be amplified pneumatically, electrically, or hydraulically for control purposes. The particular characteristics of each class of thermal

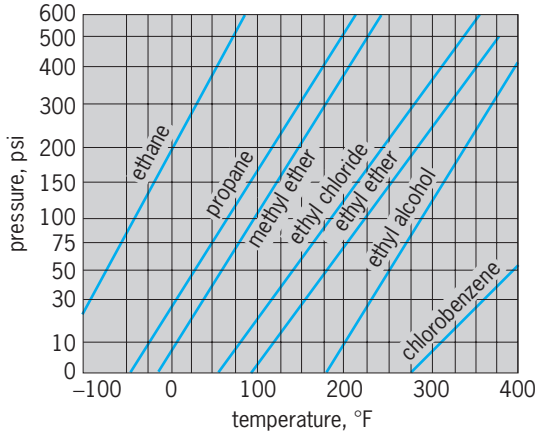

Fig. 4. Vapor-pressure versus temperature curves when various stable elements are used as thermal system fills. 1 psi = 6.895 kPa. °C = (°F − 32)/1.8. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

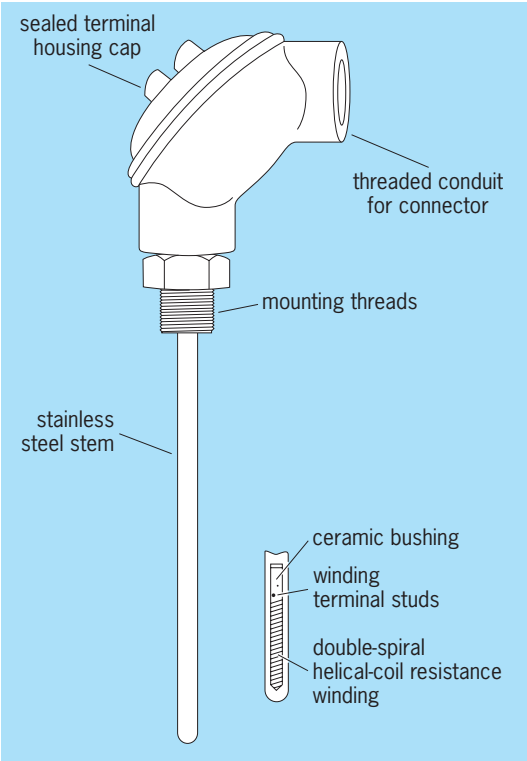

Fig. 5. Industrial-type resistance thermometer. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

system determine which will give the best service in various applications.

Resistance thermometer. In this type of thermometer the change in resistance of conductors or semiconductors with temperature change is used to measure temperature. Usually, the temperature-sensitive resistance element is incorporated in a bridge network which has a reasonably constant power supply.

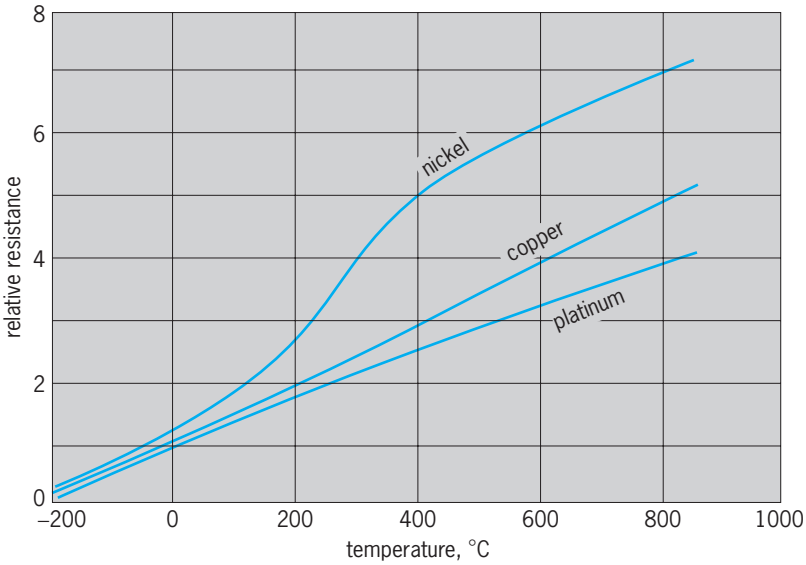

Fig. 6. Typical relative-resistance curves of several metals used in resistance thermometers. Relative resistance is the ratio of the resistance at the temperature of the metal to the resistance at 32°F (0°C). °F = (°C × 1.8) + 32. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

Although a deflection circuit is occasionally used, almost all instruments of this class use a null-balance system, in which the resistance change is balanced and measured by adjusting at least one other resistance in the bridge. All of the resistors in the bridge, except the measuring resistance, have low temperature coefficients, and the entire bridge circuit is designed to be insensitive to ambient temperature effects. The power supply to the resistance thermometer may be either direct or alternating current, the former preferred for precision measurements and the latter preferred when a servo system is used to rebalance the bridge. **Figure 5** shows an industrial resistance thermometer.

Metals commonly used as the sensitive element in resistance thermometers are platinum, nickel, and copper, and the change in resistance per degree Celsius is illustrated in **Fig. 6**. These are the most satisfactory metals, since they are stable, have a reasonable temperature coefficient of resistance, and can be drawn into fine homogeneous wires with a high resistance per unit length. For the resistance element, fine wire with a diameter of about 2.5 mils (64 micrometers) is usually wound into a small-diameter helix which is wound or otherwise mounted onto a mica support. Platinum wire elements can be used satisfactorily between −432 and 1650°F (−258 and 900°C), nickel between −238 and 572°F (−150 and 300°C), and copper between −328 and 248°F (−200 and 120°C).

A bare-wire element can be used in a clean, noncorrosive, nonconducting gas flowing at a low velocity, and its rate of response to temperature changes will be very rapid, not exceeding a few seconds. Usually, however, the element is mounted in a protecting well or tubing, from which it is electrically insulated, resulting in a time constant of a minute or more. The temperature-resistance relation of elements is determined by calibrations at the ice point, steam point, and sulfur point (832.28°F or 444.6°C). With calibrated industrial-grade resistance thermometers the uncertainty in temperature values may be about ±0.5°F (±0.3°C), while with a good, well-calibrated, and well-maintained laboratory instrument the uncertainty may be in the range of ±0.02°F (±0.01°C).

Since resistance thermometers carry a current, a self-heating error occurs. By keeping currents small and thermal conductivities high, this effect may be made negligible on most applications. In dc thermometry, thermal emfs must be carefully avoided in the circuitry. In ac thermometry, the circuitry must minimize inductive and capacitive disturbances.

Thermistor. This device is made of a solid semiconductor with a high temperature coefficient of resistance. The thermistor has a high resistance (**Fig. 7**), in comparison with metallic resistors, and is used as one element in a resistance bridge. Since thermistors are more sensitive to temperature changes than metallic resistors, accurate readings of small changes are possible. Thermistors are ceramic recrystallized mixtures of oxides of various metals and are usually in the form of small beads or disks with metallic leads.

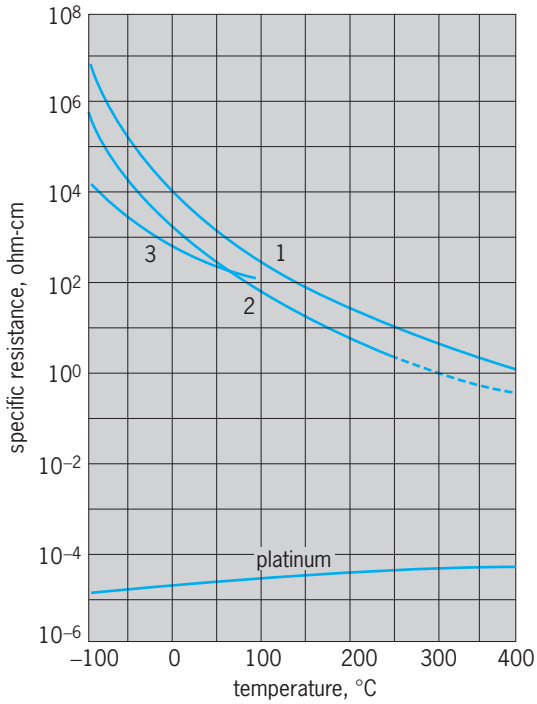

Fig. 7. Resistance-temperature properties characteristic of some typical thermometers. °F = (°C × 1.8) + 32. (*After D. M. Considine, ed., Process Instruments and Controls Handbook, McGraw-Hill, 1957*)

Thermistors are not as stable as metallic resistances, but certain compositions with good protection and care may change less than 1% per year. In general, thermistors are used between 212 and 750°F (100 and 400°C). They drift or deteriorate at higher temperatures, and at low temperatures their resistance tends to become excessive. *See* THERMISTOR.

Howard S. Bean

Bibliography. R. P. Benedict, *Fundamentals of Temperature, Pressure and Flow Measurements*, 3d ed., 1984; B. G. Liptak, *Temperature Measurement*, 1993; T. D. McGee, *Principles and Methods of Temperature Measurement*, 1988; G. K. McMillan and D. M. Considine (eds.), *Process/Industrial Instruments and Controls Handbook*, 5th ed., 1999; L. Michalski, K. Eckersdorf, and J. McGhee, *Temperature Measurement*, 2d ed., 2001; J. F. Schooley (ed.), *Temperature: Its Measurement and Control in Science and Industry*, vol. 5, 1982; J. F. Schooley, *Thermometry*, 1986.

Thermonuclear reaction

A nuclear fusion reaction which occurs between various nuclei of the light elements when they are constituents of a gas at very high temperatures. Thermonuclear reactions, the source of energy generation in the Sun and the stable stars, are utilized in the fusion bomb. *See* HYDROGEN BOMB; NUCLEAR FUSION; STELLAR EVOLUTION; SUN.

Thermonuclear reactions occur most readily between isotopes of hydrogen (deuterium and tritium) and less readily among a few other nuclei of higher atomic number. At the temperatures and densities required to produce an appreciable rate of thermonuclear reactions, all matter is completely ionized; that is, it exists only in the plasma state. Thermonuclear fusion reactions may then occur within such an ionized gas when the agitation energy of the stripped nuclei is sufficient to overcome their mutual electrostatic repulsions, allowing the colliding nuclei to approach each other closely enough to react. For this reason, reactions tend to occur much more readily between energy-rich nuclei of low atomic number (small charge) and particularly between those nuclei of the hot gas which have the greatest relative kinetic energy. This latter fact leads to the result that, at the lower fringe of temperatures where thermonuclear reactions may take place, the rate of reactions varies exceedingly rapidly with temperature. *See* PLASMA (PHYSICS).

The reaction rate may be calculated as follows: Consider a hot gas composed of a mixture of two energy-rich nuclei, for example, tritons and deuterons. The rate of reactions will be proportional to the rate of mutual collisions between the nuclei. This will in turn be proportional to the product of their individual particle densities. It will also be proportional to their mutual reaction cross section σ and relative velocity υ. Thus Eq. (1) gives the rate of reaction. The quantity $(\sigma\upsilon)_{12}$ indicates an average value of σ and υ obtained by integration of these quantities over the velocity distribution of the nuclei (usually assumed to be maxwellian). Since the total density $n = n_1 + n_2$, if the relative proportions of n_1 and n_2 are maintained, R_{12} varies as the square of the total particle density.

$$R_{12} = n_1 n_2 \langle \sigma \upsilon \rangle_{12} \qquad \text{reactions/(cm}^3\text{)(s)} \qquad (1)$$

The thermonuclear energy release per unit volume is proportional to the reaction rate and the energy release per reaction, as in Eq. (2). If this energy release, on the average, exceeds the energy losses from the system, the reaction can become self-perpetuating. *See* CARBON-NITROGEN-OXYGEN CYCLES; KINETIC THEORY OF MATTER; MAGNETOHYDRODYNAMICS; NUCLEAR REACTION; PINCH EFFECT; PROTON-PROTON CHAIN.

$$P_{12} = R_{12} W_{12} \qquad \text{ergs/(cm}^3\text{)(s)} \qquad (2)$$

Richard F. Post

Bibliography. R. O. Dendy, *Plasma Physics: An Introductory Course*, 2d ed., 1994; S. Glasstone and R. H. Lovberg, *Controlled Thermonuclear Reactions*, 1960, reprint 1975; K. Miyamoto, *Plasma Physics for Nuclear Fusion*, rev. ed., 1989; K. Nishikawa and M. Wakatani, *Plasma Physics: Basic Theory with Fusion Applications*, 3d ed., 2000.

Thermoregulation

The processes by which many animals actively maintain the temperature of part or all of their body within a specified range in order to stabilize or optimize temperature-sensitive physiological processes.

Thermoregulation is evident when the temperature of part or all of the body of a free-living animal is consistently different from, or less variable than, the temperature of a collection of inanimate objects of identical external properties scattered randomly around the same habitat. Body temperatures of normally active animals may range from 32 to 115°F (0 to 46°C) or more, but the tolerable range for any one species is much narrower.

Animals are commonly classified as warm-blooded or cold-blooded. When the temperature of the environment varies, the body temperature of a warm-blooded or homeothermic animal remains high and constant, while that of a cold-blooded or poikilothermic animal is low and variable. However, supposedly cold-blooded reptiles and insects, when active, may regulate body temperatures within 2–4°F (1–2°C) of a species-specific value in the 93–108°F (34–42°C) range typical of birds and mammals. Supposedly warm-blooded mammals and birds may allow their temperature to drop to 37–68°F (5–20°C) during hibernation or torpor. Further, optimal temperature varies with organ, time of day, and circumstance. In mammals, the testes are too hot to produce sperm at the same temperature required for normal brain function. A temporary high temperature (hyperthermia) improves muscle power output, digestion, and defense against many bacterial and some parasitic infections. Most animals have daily body temperature cycles of 2–9°F (1–5°C) or more. The cycles appear important; for example, desert iguanas will die in a few days if they are held continuously at the temperature preferred during activity. Animals may conserve energy with daily or seasonal periods of very low body temperature (hypothermia). Thus, this classification is often misleading.

A better classification is based on the principal source of heat used for thermoregulation. Endotherms (birds, mammals) use heat generated from food energy. Ectotherms (invertebrates, fish, amphibians, reptiles) use heat from environmental sources. Ecological roles are related to this classification. For example, terrestrial ectotherms have an advantage in warm climates, particularly where food is scarce, as 30–40% of energy intake is available for growth and reproduction. Endotherms can continue to be active when endotherms must seek shelter from cold, but only 1–3% of energy intake is available for growth and reproduction. Therefore, the high rate of heat loss in water means that only highly active animals that are large or well insulated can be aquatic endotherms. This classification also has limitations, however. Endotherms routinely use external heat sources to minimize the food cost of thermoregulation, and some ectotherms use food energy for thermoregulation. Social insects are ectotherms as individuals, but may be endotherms as colonies. Thus, it is best to focus on mechanisms widely used for thermoregulation. *See* PHYSIOLOGICAL ECOLOGY (ANIMAL).

Thermal environment. An organism and its thermal environment are a functional unit. A useful starting point for thermoregulation is operative temperature, indicated by a thermometer with external properties identical to those of the animal (for example, a taxidermic mount on a hollow metal core with a thermometer inside), located at the same point as the animal and with the same orientation to sun and wind. Such a thermometer combines air temperature and solar radiation in proportions determined by wind and external properties of the animal such as size, shape, color, fur, or feathers. At a fixed air temperature such as 68°F (20°C), operative temperature is increased by sunlight and decreased by wind. Sunlight usually heats darker-colored objects more than lighter ones. Wind cools small objects more effectively than large ones. Thus, small or light-colored animals are most sensitive to air temperature and wind, while large or dark-colored animals are relatively more sensitive to sun. Similarly, operative temperature may vary over the body. Appendages that are smaller than the torso are more sensitive to wind. The back and flanks may receive more solar radiation than the ventral side, which is usually shaded. In contrast, aquatic or burrowing animals are affected uniformly and almost exclusively by water or soil temperature. *See* CONVECTION (HEAT); HEAT RADIATION.

Passive thermoregulation. Variation in the body temperature of an animal may be noticeably reduced relative to typical animals when it or its nest has a particular color, shape, or size. The goal of thermoregulation is thus aided or achieved, but the means are passive because they do not involve a physiological or behavioral response to thermal conditions. For example, in an often-sunny environment, a light-colored animal may experience a cooler and less variable environment than would a darker one. The compass termites of Australia construct their nests as long, narrow, high structures with the long axis oriented within ±20° of a north-south line. Thus, a large area is heated by sunlight in the morning and evening when the air is cool and sunlight weak, while a much smaller area absorbs sunlight in the heat of the day. A fairly stable daytime colony temperature of 91–95°F (33–35°C) results. Large animals such as rhinoceroses and camels do not need to sweat much in the African sun. They passively store the solar heat in their massive bodies by allowing body temperature to increase 11–13°F (6–7°C). This heat is lost during the cool nights. A small increase in the temperature of a wet-skinned amphibian automatically results in a large increase in evaporation that prevents further rise in body temperature.

Behavior. Behavior is the most obviously active form of thermoregulation. Most animals are mobile, sensitive to their environment, and capable of complex behaviors. The simplest thermoregulatory behavior consists of moving to a favorable location. Fish can detect and respond to changes of 0.09°F (0.05°C) in water temperature. Terrestrial animals move short distances between sun and shade or wind and shelter to secure a favorable operative temperature. Therefore, some reptiles can regulate their temperature within 2–4°F (1–2°C) of a preferred value.

Because of heat storage, only the average operative temperature is important. For example, the antelope ground squirrel forages at lethal operative temperatures on the desert surface by allowing body temperature to rise 9°F (5°C). It then returns to a cool burrow, releases the stored heat, and repeats the cycle.

Operative temperature may be altered by changing posture in one place. Lizards face the sun to minimize the area exposed to solar heating, or orient broadside to maximize it. The sun also heats the surface that supports the animal, which then heats the boundary layer of relatively still air near the ground. Lizards basking in the sun in high mountains flatten their bodies onto the supporting surface to immerse themselves in warm, still, air. This behavior allows them to obtain an operative temperature up to 54°F (30°C) warmer than the temperature of free air. Conversely, gray gulls escape the stifling boundary layer on the Atacama desert in Chile, for example, by standing over their nests in the heat of the day. Some reptiles and amphibians also expand or contract pigmented cells in their skin to increase or decrease solar heating. Roadrunners and grebes increase solar heating by facing away from the sun and adjusting their rump feathers to expose patches of black skin. Many birds decrease solar heating by raising long feathers on their back to form a parasol that shades the body. Some ground squirrels use their tail as a similar sunshade.

Evaporative cooling. Evaporation is an effective means of cooling the body. Evaporation from the respiratory mucous membranes is the most common mechanism. Nose breathing minimizes heat and water loss. Evaporation from the mucous membranes cools the nose during inhalation. During exhalation, water vapor condenses on the cool nasal membranes and is recovered. Evaporation can be greatly increased by exhaling from the mouth to prevent condensation. Additional increases in evaporation result from panting, that is, rapid breathing at the resonant frequency of the respiratory system. Evaporation from the eyes and the mucous membranes of the mouth and tongue is another source of cooling. Water is also evaporated from the skin of all animals, and can be varied for thermoregulation. Some desert frogs control evaporation by spreading an oily material on the skin. Reptiles, birds, and mammals have relatively impermeable skins, but evaporation can be increased by various means. The sweat glands in the skin of mammals are particularly effective and are one of the few purely thermoregulatory organs known. *See* SKIN.

Insulation. Changes in circulation can be used to regulate heat flow. Galápagos marine iguanas increase circulation to warm their bodies rapidly while basking on shore, and minimize circulation to slow cooling while foraging in the ocean. Countercurrent heat exchange is used to regulate heat flow to particular parts of the body while maintaining oxygen supply. Vessels carrying warm blood are in intimate contact with vessels carrying cool blood in the opposite direction. Large vessels may be divided into intermingled masses of small vessels to maximize heat exchange, forming an organ called a rete. However, retes can be bypassed by alternative circulation paths to regulate heat flow. Many animals living in warm environments have a rete that regulates brain temperature by cooling the arterial blood supply to the brain with blood draining from the nasal membranes, eyes, ears, or horns. Similar retes regulate heat loss from body to the legs in birds and mammals, from flight muscles to the abdomen of bumblebees, and from swimming muscles to the water in large fish such as tunas and sharks. *See* CARDIOVASCULAR SYSTEM; COUNTERCURRENT EXCHANGE (BIOLOGY).

Heat exchange with the environment is limited by the fur of mammals, feathers of birds, and fur-like scales or setae of insects. Erection or compression of this insulation varies heat flow. Insulation thickness varies over the body to exploit variations in local operative temperature. Thermal windows are thinly insulated areas that are either shaded (abdomen of mammals, axilla of birds and mammals) or of small size (ears, face, legs) so that solar heating is minimized. Adjustments to circulation with retes and a change in posture vary heat flow to thermal windows. Conversely, the bison and vicuña present thickly furred areas to the wind in cold, windy conditions.

Metabolic heat. The oxidation of foodstuffs within the metabolic pathways of the body releases as much heat as if it were burned. Basal metabolism is the energy use rate of a fasting animal at rest. Activity, digestion, and thermoregulation increase metabolism above the basal rate. Endothermy is the utilization of metabolic heat for thermoregulation. Birds and mammals are typical examples, but significant endothermy also occurs in large salt-water fish, large reptiles, and large flying insects. Well-insulated animals can tolerate temperatures of −76°F (−60°C) indefinitely by using endothermy, but small or poorly insulated animals may use endothermy successfully only under mild conditions or for short periods.

The energy use of an idealized endotherm varies with standard operative temperature, that is, temperature which is adjusted for the effect of wind on the rate of metabolic heat loss (see **illus.**). In the thermal neutral zone, heat production is constant, and adjustments to insulation and evaporative cooling regulate body temperature. In the cold stress zone, insulation is at its maximum, and additional heat is produced for thermoregulation. The rate of increase of heat production as temperature falls is greater for poorly insulated animals than for well-insulated animals. Thus, a well-insulated animal has a relatively wider thermal neutral zone as well as an ability to tolerate lower temperatures. In the heat stress zone, evaporative cooling is increased by means such as panting that also increase heat production. Evaporative cooling increases nearly exponentially with temperature. If possible, endotherms use behavior to obtain temperatures in the thermal neutral zone unless the benefits of activity at stressful temperatures outweigh the costs.

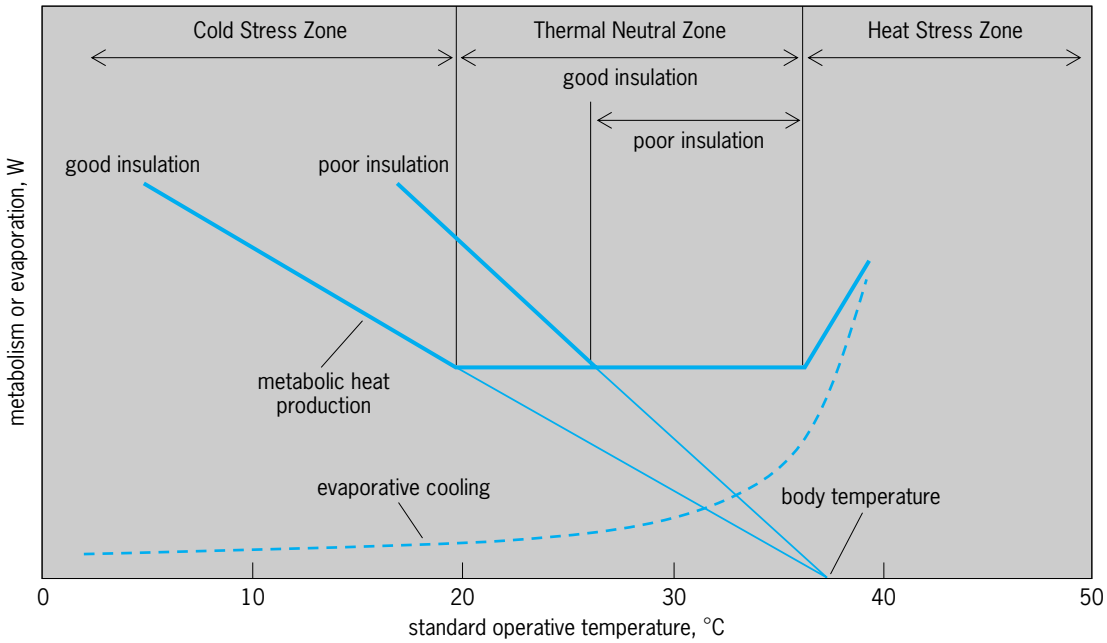

Metabolic heat production and evaporative cooling of a hypothetical endotherm in response to standard operative temperature. The line relating heat production and temperature in the cold stress zone extrapolates to zero at body temperature.

Endotherms regulate only the temperature of the body core, that is, the brain, heart, and lungs. The heat production of these metabolically active organs is often supplemented with heat produced in muscles. Heat produced as a by-product of activity may substitute partially for thermoregulatory heat production, and imposes no thermoregulatory energy cost. Sharks and tuna use elaborate retes to conserve heat generated by swimming and so regulate body temperature at 86±9°F (30±5°C) in 9–86°F (5–30°C) water. In contrast, shivering produces heat only for thermoregulation and results in an extra cost. Large flying insects often cannot leave the ground until shivering of the flight muscles raises temperature to 95–104°F (35–40°C). Some animals have specialized heater organs for nonshivering thermogenesis, which is more efficient than shivering. Brown adipose tissue is a fatty tissue with a high density of mitochondria. It is found in the thorax of mammals, especially newborns and hibernators, and it warms the body core efficiently. Billfish and mackerel use modified eye muscles to generate large amounts of heat. Blood leaving these organs passes through a rete to warm the blood supply to the brain and the retina of the eye. However, many species, apparently including all birds, lack nonshivering thermogenesis. *See* METABOLISM.

Social insects. Massed colonies of some social insects thermoregulate as an endothermic superorganism. The combined metabolism of the million or more individuals and their fungus gardens in a colony of the African termite *Macrotermes bellicosus* heats air in the nest. The heated air rises up a system of chimneys and drives circulation in an elaborate system of ducts. The result is a colony temperature of 86±4°F (30±2°C) with outside air temperatures ranging 59–100°F (15–38°C). Bumblebees are adapted to colder climates and gather fibrous materials to insulate their nests. Colony metabolism can then regulate blood temperatures at 88±4°F (31±2°C) against Arctic summer air temperatures of 41–52°F (5–11°C). Honeybees have an even more complex thermoregulatory system.

Neural control. It was once believed that the preoptic area and hypothalamus of the brain was the master thermostat for vertebrates. However, the variety of mechanisms used in thermoregulation indicates a corresponding complexity in neural control. Temperature sensors distributed over the skin respond nearly immediately to changes in the environment and provide the major input. Nearly all parts of the central nervous system also respond to local thermal stimulation. These peripheral and central thermal inputs are integrated at a series of centers beginning in the spinal cord. This series clearly extends to the cerebral cortex, as a learning period is required before behavioral thermoregulation reaches maximum precision. Various components respond to the rate of temperature change as well as the difference between preferred and actual temperature. The neuroendocrine system then regulates metabolic heat production, the sympathetic nervous system controls blood flow, and the cerebral cortex controls behavioral thermoregulation. *See* ENDOCRINE SYSTEM (VERTEBRATE); HOMEOSTASIS; NERVOUS SYSTEM (VERTEBRATE).

George Bakken

Bibliography. D. M. Gates, *Biophysical Ecology*, 1980; J. E. Heath (ed.), Thermoregulation in vertebrates, *Annu. Rev. Physiol.*, vol. 48, 1986;

B. Heinrich, *The Hot Blooded Insects*, 1993; P. C. Withers, *Comparative Animal Physiology*, 1992.

Thermosbaenacea

A small order of the crustacean superorder Peracarida. In thermosbaenaceans, the carapace, which may cover part of the cephalic region and one to several thoracic somites, is fused only to the first thoracic somite. The carapace of females is expanded to provide a dorsally positioned brood pouch where embryos hatch as subadults (see **illus.**). Eyes are reduced or absent. The abdomen consists of six somites and a telson; however, in *Thermosbaena*, at least, the telson and sixth somite are fused to form a pleotelson. The first pair of thoracic appendages are modified as maxillipeds, and may be sexually dimorphic; the remaining five to seven pairs provide the animals with locomotion. Sexes are separate.

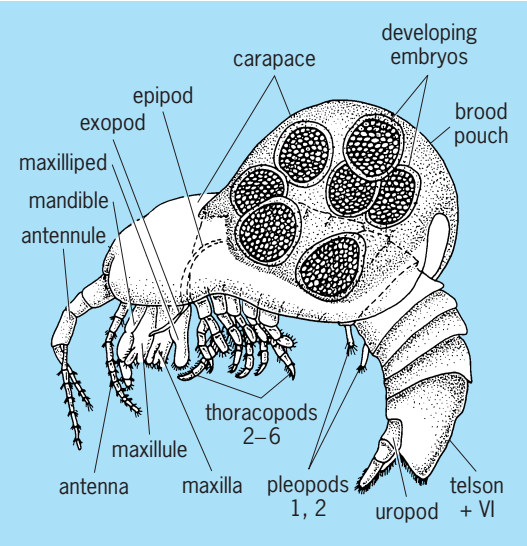

An ovigerous female, *Thermosbaena* (Thermosbaenacea). (*After D. Barker, A study of Thermosbaena mirabilis (Malacostraca, Peracardia) and its reproduction, Quart. J. Micro. Sci., 103:261–286, 1962*)

Thermosbaenaceans are bottom dwellers; however, they can swim, and they are oriented ventral side up. They feed on detritus scraped from the substrate. Thermosbaenaceans have been found principally in thermal, but occasionally in cool, fresh- or brackish-water, lakes, springs, and interstitial coastal areas, and also in cave pools, in a geographic band stretching from the Mediterranean to the Caribbean and Gulf of Mexico. *See* PERACARIDA.

Patsy A. McLaughlin

Bibliography. T. E. Bowman and T. M. Iliffe, *Tulumella unidens*, a new genus and species of thermosbaenacean crustacean from the Yucatán Peninsula, Mexico, *Proc. Biol. Soc. Wash.*, 101(1):221–226, 1988; P. A. McLaughlin, *Comparative Morphology of Recent Crustacea*, 1980; F. R. Schram, *Crustacea*, 1986.

Thermosphere

A rarefied portion of the atmosphere, lying in a spherical shell between 50 and 300 mi (80 and 500 km) above the Earth's surface, where the temperature increases dramatically with altitude. Many satellites orbit in the upper thermosphere, and the drag on these satellites exerted by the atmosphere eventually brings them down to burn up in the lower atmosphere.

The thermosphere responds to the variable outputs of the Sun, the ultraviolet radiation at wavelengths less than 200 nanometers, and the solar wind plasma that flows outward from the Sun and interacts with the Earth's geomagnetic field. This interaction energizes the plasma, accelerates charged particles into the thermosphere, and produces the aurora borealis and aurora australis, nearly circular-shaped regions of luminosity that surround the magnetic north and south poles respectively. Embedded within the thermosphere is the ionosphere, a weakly ionized plasma. *See* IONOSPHERE; MAGNETOSPHERE; PLASMA (PHYSICS); SOLAR WIND.

Global mean structure. The atmosphere below about 50 mi (80 km) is almost entirely molecular: about 78% molecular nitrogen (N_2), 21% molecular oxygen (O_2), and 1% other minor gas constituents, such as ozone (O_3), carbon dioxide (CO_2), and argon (Ar). Above 50 mi (80 km), in the thermosphere, these molecular species are subjected to intense solar ultraviolet radiation and photodissociation that gradually turns the molecular species into the atomic species oxygen (O), nitrogen (N), and hydrogen (H). Solar radiation at wavelengths less than 130 nm can also ionize both molecular and atomic species and turn the neutral molecular and atomic species into a plasma at altitudes above 600 mi (1000 km). Thus there are various upper-atmosphere regimes (**Fig. 1**). Up to above 60 mi (100 km), atmospheric turbulence

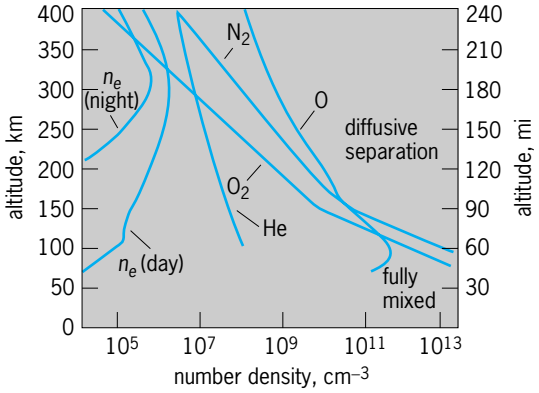

Fig. 1. Distribution of some important constituents in the thermosphere: atomic oxygen (O), molecular nitrogen (N_2), molecular oxygen (O_2), helium (He), and ionospheric electrons (n_e) for both daytime and nighttime conditions.

keeps the atmosphere well mixed, with the molecular concentrations dominating in the lower atmosphere. Above 60 mi (100 km), solar ultraviolet radiation most strongly dissociates molecular oxygen, and there is less mixing from atmospheric turbulence. The result is a transition area where molecular diffusion dominates and atmospheric species settle according to their molecular and atomic weights. Above 60 mi (100 km), atomic oxygen is the dominant species; above about 300 mi (500 km), helium (He) is dominant, and eventually atomic hydrogen, from the photodissociation of water vapor, becomes the dominant species at altitudes above 600 mi (1000 km).

The number of particles per unit volume (density values) of ionospheric electrons that peak near (180–240 mi) (300–400 km) depends upon the intensity of solar ultraviolet radiation (Fig. 1). The ionospheric densities are greatest during the day, and the lower ionosphere can chemically recombine at night. The dominant ion in the upper ionosphere is O^+, eventually supplemented by H^+ at high altitudes.

About 60% of the solar ultraviolet energy absorbed in the thermosphere and ionosphere heats the ambient neutral gas and ionospheric plasma; 20% is radiated out of the thermosphere as airglow from excited atoms and molecules; and 20% is stored as chemical energy of the dissociated oxygen and nitrogen molecules, which is released later when recombination of the atomic species occur. Most of the neutral gas heating that establishes the basic temperature structure of the thermosphere is derived from excess energy released by the products of ion-neutral and neutral chemical reactions that occur in the thermosphere and ionosphere. This heating is so strong in the rarefied upper atmosphere that heating rates range from about 50 K (90°F) per day near 60 mi (100 km), in the lower thermosphere, to 1000 K (1800°F) per day near 300 mi (500 km), in the upper thermosphere. *See* AIRGLOW; ULTRAVIOLET RADIATION.

Atomic and molecular gases in the thermosphere neither effectively radiate heat to space by infrared radiation nor provide the cooling rates needed to balance the intense heating rates. The average vertical temperature profile is, therefore, determined by a balance of local solar heating by the downward conduction of molecular thermal product to the region of minimum temperature near 50 mi (80 km; **Fig. 2**). There, energy is radiated to space by optically active infrared constituents such as carbon dioxide and ozone, which cannot exist at higher altitudes in the thermosphere because of intense solar photodissociation of these species. Infrared radiation from atomic species, such as atomic oxygen, is small in the thermosphere. *See* INFRARED RADIATION.

For heat to be conducted downward within the thermosphere, the temperature of the thermosphere must increase with altitude. In fact, the global mean temperature increases from about 200 K (−100°F) near 80 km (50 mi) to 700–1400 K (800–2100°F) above 180 mi (300 km), depending upon the intensity of solar ultraviolet radiation reaching the Earth. Above 180 mi (300 km), molecular thermal conduction occurs so fast that vertical temperature differences are largely eliminated; the isothermal temperature in the upper thermosphere is called the exosphere temperature. Above about 300 mi (500 km), in the exosphere, collisions between atmospheric gases are so rare that individual gas particles can either escape the Earth's gravitational pull or can obtain ballistic trajectories in their transit.

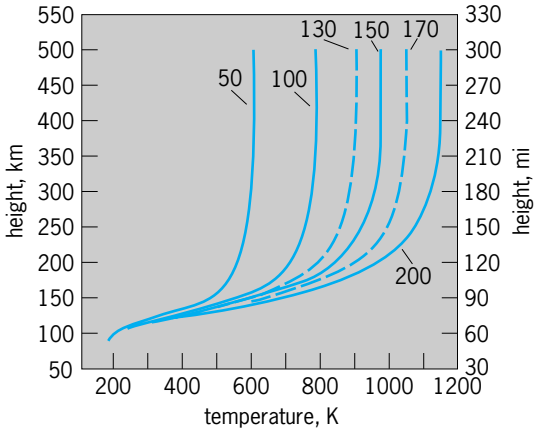

Fig. 2. Global mean-temperature profile in the thermosphere for various levels of solar activity. Labels on curves correspond to F10.7 values, which represent the level of the 10.7-cm radio emission from the Sun in units of 10^{-22} W m^{-2} Hz^{-1}. The value F10.7 = 50 is a hypothetical calculation, if the Sun should drop to this value. An F10.7 of about 70 represents a solar-cycle minimum, whereas an F10.7 of 200 represents solar cycle maximum activity. °F = (K × 1.8) − 459.67.

The temperature profile is strongly dependent upon the intensity of solar ultraviolet radiation, which varies with time and correlates with the sunspot cycle (approximately 11 years). During sunspot minimum the solar ultraviolet output produces an exospheric temperature of 500–700 K (440–800°F). At sunspot maximum the solar ultraviolet output can be as much as 10 times stronger at short wavelengths (10–30 nm) and 2–3 times stronger at longer ultraviolet wavelengths (100–150 nm), producing global exospheric temperatures of 1200–1500 K (1700–2200°F).

The chemical composition and densities of the thermosphere and ionosphere also vary with the solar ultraviolet output and vertical temperature profile. The intense solar radiation at sunspot maximum dissociates more molecular oxygen and nitrogen, thereby increasing the magnitude of the oxygen and nitrogen chemical cycles. It also produces larger electron and ion number densities within the ionosphere and drives other chemical cycles. In addition, the increased heating causes the atmosphere to expand outward from the Earth. Satellites orbiting at a given altitude thereby experience variations in atmospheric drag as the atmosphere expands and contracts in response to changing solar ultraviolet radiative output. *See* SUN.

Thermospheric circulation. As the Earth rotates, absorption of solar energy in the thermosphere

undergoes a daily variation. Dayside heating causes the atmosphere to expand and the loss of heat at night contracts it. This heating pattern creates pressure differences that drive a global circulation, transporting heat from the warm dayside to the cool nightside.

The meteorology of the thermosphere differs from the familiar weather experienced in the lower atmosphere. The fluid motions are governed by the same system of equations that is used by meteorologists to study the lower atmosphere; however, there are notable differences between some of the physical and chemical processes in the upper and lower atmosphere. In the troposphere temperature decreases with altitude, while in the thermosphere temperature increases with altitude. For this reason the thermosphere is more stable than the troposphere. Also, as the air in the thermosphere becomes rarefied, turbulence ceases; and the motions are strongly influenced by molecular diffusion, which provides an additional stability. Viscosity primarily transfers momentum between various altitudes, smoothing out vertical gradients in wind velocity, just as molecular thermal conductivity produces an isothermal vertical temperature profile. Above about 300 km (180 mi), the atmospheric layers are coupled so strongly by viscosity that the wind velocity becomes nearly constant with altitude.

Another important force that acts in the thermosphere is an ion drag force produced when the neutral gases in the thermosphere collide with the plasma of the ionosphere. Above about 75 mi (125 km), where the ion gyrofrequency (the frequency at which an ion spirals around a magnetic field line) is greater than the ion-neutral collision frequency, ions are essentially locked to the Earth's geomagnetic field lines and can move across them only in response to an electric field. A neutral-gas wind flowing through the relatively immobile ions experiences a collisional drag that is largest at the peak of the ionospheric layer, providing the main resistance that balances the driving pressure force within the upper thermosphere. In the lower thermosphere, however, the ion drag force is not as effective. The pressure force is primarily balanced by the Coriolis force, and winds flow perpendicular to pressure gradients, as they do in the lower atmosphere.

Thus, meteorology in the thermosphere is considerably different from that in the lower atmosphere. The flow is stable and viscous. Because of the stability and viscosity, there are no instabilities that lead to the development of high-pressure regions and low-pressure storms, such as those that exist in the troposphere. The thermosphere, however, can easily support the transmission of waves such as the diurnal and semi-diurnal tides, gravity waves generated by plasama processes, auroral particle inputs, and the Joule dissipation of ionospheric current systems. Joule dissipation is electrical resistive heating by a current flowing through, and electrons and ions colliding with, neutral particles, thereby transferring directed motion into heat. A crude visualization of the flow in the thermosphere is a wiggling bowl of viscous jelly responding to the variable forcings of solar ultraviolet radiation, auroral inputs, and complex neutral-gas-plasma interactions. *See* TROPOSPHERE.

Auroral influences. In addition to solar radiative heating driving the basic thermospheric circulation, there is an important heat and momentum source at high magnetic latitudes associated with auroral activity. The aurora is driven by the interactions of the solar-wind plasma with the Earth's geomagnetic field. Complex interactions between the solar wind and magnetosphere energize the plasma, causing some of it to be transferred into the ionosphere and thermosphere. Energetic electrons and ions bombard the atmosphere in the auroral ovals in both magnetic hemispheres. The electrons are sufficiently energetic to ionize atoms and dissociate molecules and produce copious amounts of emissions that can be seen from the ground and by satellites from space. The electrons flowing into the ionosphere and thermosphere carry an electric current of about a million amperes. During intense geomagnetic or auroral storms, the electric currents and energetic particles may deposit as much energy at high latitudes as that from the absorption of solar ultraviolet radiation. Typically, the global energy input during very quiet auroral activity is about 10^{10} W. This increases to 10^{11} W during moderate activity and up to 10^{12} W for short periods during intense auroral storms. *See* GEOMAGNETIC VARIATIONS.

The thermosphere responds to this energy input in a number of ways. The strong and rapid changes in energy input generate a variety of atmospheric waves that can transport energy from high to low latitudes. In addition, the more sustained energy inputs generate a mean circulation that flows from high to low latitudes; they can be large enough to compete with the circulation generated by solar radiative energy absorption that is basically from low to high latitudes. In the mean thermospheric circulation for equinox and solstice conditions, there is competition between the solar radiative forced circulation and auroral forced circulation for the three levels of geomagnetic activity (**Fig. 3**). During geomagnetic quiet conditions the circulation is from the Equator toward both poles during equinox and from the summer to winter pole for solstice. During moderate auroral activity the energy and momentum additions at magnetic conjugate high latitudes generate a pole-to-Equator component that is superimposed upon the solar-driven circulation. During an auroral storm the pole-to-Equator circulation driven by auroral heating dominates. Thus, it is evident that the mean motions in the thermosphere are in a constant state of agitation, depending upon the amount and duration of high-latitude heating from auroral processes. *See* AURORA; EQUINOX; MIDDLE-ATMOSPHERE DYNAMICS; SOLSTICE.

Global change. The Earth's atmosphere is subject to global change because of the release of trace gases, such as carbon dioxide and methane, by human activity or natural causes. These greenhouse gases have

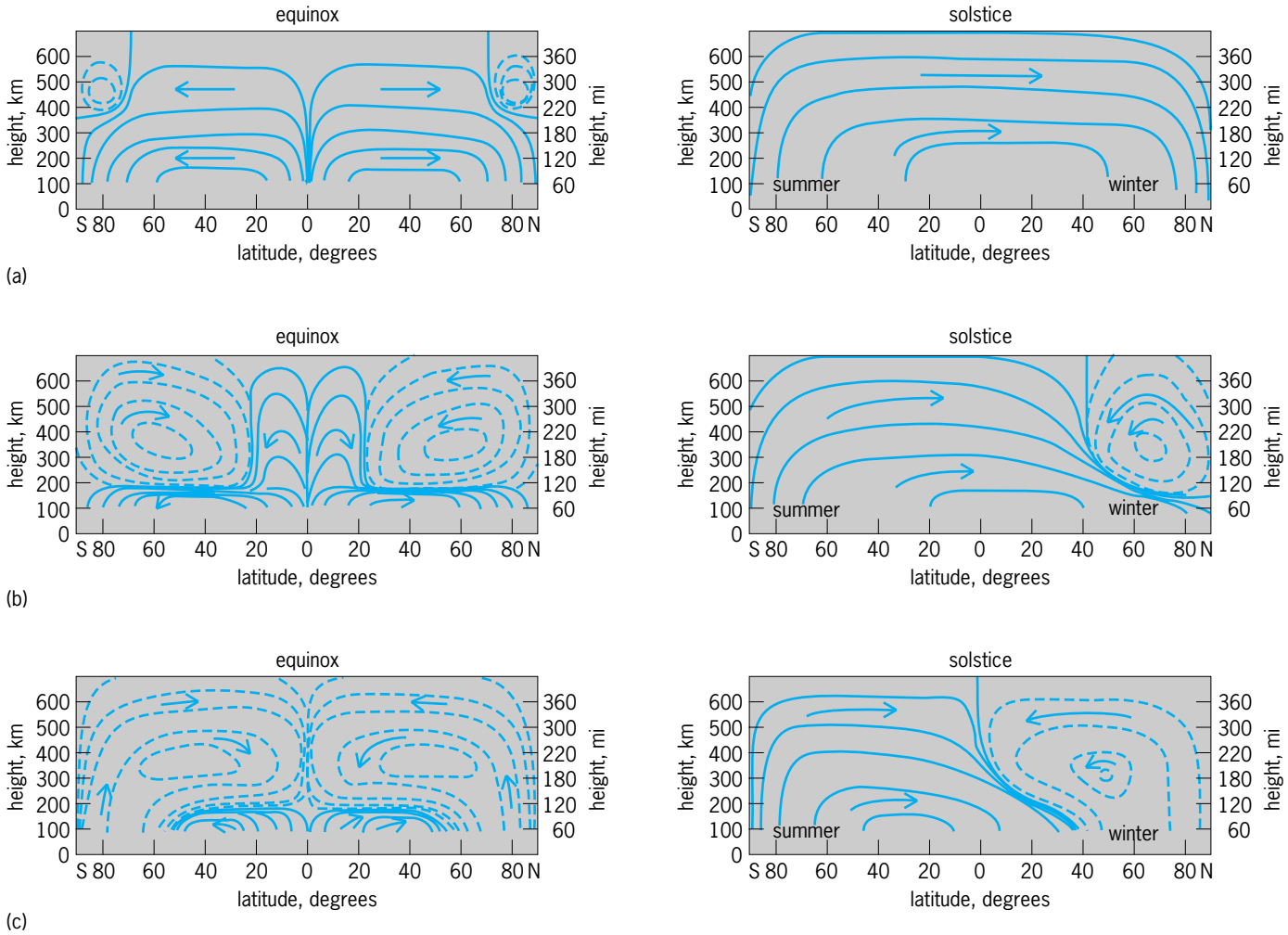

Fig. 3. Zonal mean meridional circulation in the Earth's thermosphere during equinox and solstice for various levels of auroral activity: (*a*) quiet auroral activity where circulation is primarily driven by solar ultraviolet heating; (*b*) average auroral activity where heating is 10^{11} W; and (*c*) during geomagnetic storms where the heating is about 10^{12} W. The contours schematically illustrate the mass flow, and the arrows indicate the direction of the motion.

been projected to cause changes in the lower atmosphere structure, where the troposphere is expected to warm a few degrees and the stratosphere to cool by 10–20 K (18–36°F) in response to a doubling of certain trace gases. Model calculations have been made that suggest that global change effects will also occur in the upper mesosphere of the atmosphere, primarily due to the increased carbon dioxide cooling to space. With a doubling of the concentrations of carbon dioxide and methane (from present-day conditions) that is expected to occur by the end of the twenty-first century, the mesosphere is predicted to cool by about 10 K (18°F) and the thermosphere to cool by about 50 K (90°F). In the lower atmosphere, increased carbon dioxide traps radiation and the atmosphere warms. But as the radiation transfers upward, it eventually finds space and it is no longer trapped. Therefore, increased concentrations of carbon dioxide allow more energy to escape to space. The Earth in steady state is in balance between solar ultraviolet radiation and visible energy absorbed and infrared carbon dioxide cooling to space. This cooling will cause the atmosphere to contract, reducing by 50% the drag on satellites at a given altitude and causing a redistribution of certain minor species. The peak of the ionospheric layer is predicted to drop by 12 mi (20 km), with a redistribution of the present-day plasma. The overall consequences of these effects is not known at the present time; however, it is likely that thermal tides, thermosphere-ionosphere interactions, and the thermosphere and ionosphere response to solar and auroral variability will all change considerably toward the end of the twenty-first century. *See* ATMOSPHERE; CLIMATE MODELING; MESOSPHERE; STRATOSPHERE. R. G. Roble

Bibliography. S.-I. Akasofu and Y. Kamide (eds.), *The Solar Wind and the Earth*, 1987; Geophysics Study Committee, National Academy of Sciences, *The Upper Atmosphere and Magnetosphere: Studies in Geophysics*, 1977; S. Kato, *Dynamics of the Upper Atmosphere*, 1980; M. C. Kelly, *The Earth's Ionosphere: Plasma Physics and Electrodynamics*, 1989; M. H. Rees, *Physics and Chemistry of the Upper Atmosphere*, Cambridge, 1989.

Thermostat

An instrument which directly or indirectly controls one or more sources of heating and cooling to maintain a desired temperature. To perform this function a thermostat must have a sensing element and a transducer. The sensing element measures changes in the temperature and produces a desired effect on the transducer. The transducer converts the effect produced by the sensing element into a suitable control of the device or devices which affect the temperature. The most commonly used principles for sensing changes in temperature are (1) unequal rate of expansion of two dissimilar metals bonded together (bimetals), (2) unequal expansion of two dissimilar metals (rod and tube), (3) liquid expansion (sealed diaphragm and remote bulb or sealed bellows with or without a remote bulb), (4) saturation pressure of a liquid-vapor system (bellows), and (5) temperature-sensitive resistance element.

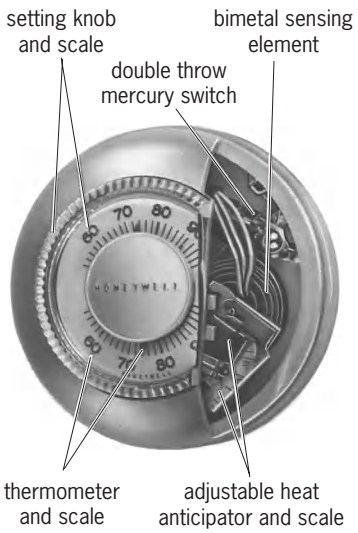

Fig. 1. Typical heat-cool thermostat. (*Honeywell Inc.*)

The most commonly used transducers are a switch that makes or breaks an electric circuit, a potentiometer with a wiper that is moved by the sensing element, an electronic amplifier, and a pneumatic actuator.

The most common thermostat application is for room temperature control. **Figure 1** shows a typical on-off heating-cooling room thermostat. In a typical application the thermostat controls a gas valve, oil burner control, electric heat control, cooling compressor control, or damper actuator.

To reduce room temperature swings, high-performance on-off thermostats commonly include a means for heat anticipation. The temperature swing becomes excessive if thermostats without heat anticipation are used because of the switch differential (the temperature change required to go from the break to the make of the switch), the time lag of the sensing element (due to the mass of the thermostat) in sensing a change in room temperature, and the inability of the heating system to respond immediately to a signal from the thermostat.

To reduce this swing, a heater element (heat anticipator) is energized during the on period. This causes the thermostat to break prematurely. **Figure 2** shows a comparison of the room temperature variations when a thermostat with and without heat anticipation is used.

The same anticipation action can be obtained on cooling thermostats by energizing a heater (cool anticipator) during the off period of the thermostat. Room thermostats may be used to provide a variety of control functions, such as heat only; heat-cool; day-night, in which the night temperature is controlled at a lower level; and multistage, in which there may be one or more stages of heating, or one or more stages of cooling, or a combination of heating and cooling stages.

Thermostats are also used extensively in safety and limit applications. Thermostats are generally of the following types: insertion types that are mounted on ducts with the sensing element extending into a duct, immersion types that control a liquid in a pipe or tank with the sensing element extending into the liquid, and surface types in which the sensing element is mounted on a pipe or similar surface.

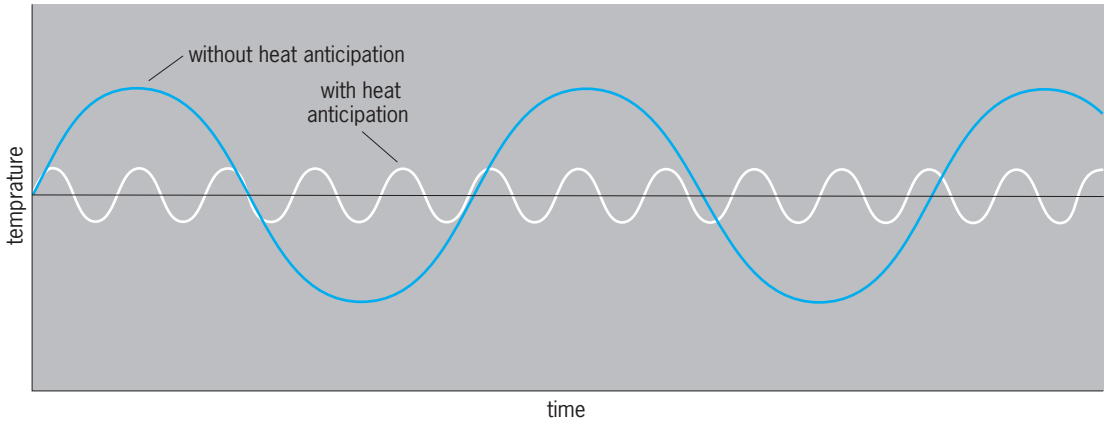

Fig. 2. Comparison of temperature variations using a timed on-off thermostat with and without heat anticipation.

See COMFORT HEATING; OIL BURNER.

Nathaniel Robbins, Jr.

Bibliography. V. C. Miles, *Thermostatic Control*, 2d ed., 1974; J. Olivieri, *How to Design Heating-Cooling Comfort Systems*, 1987; R. K. Schneider, *HVAC Control Systems*, 2d ed., 1988; J. Trost, *Efficient Buildings: Heating and Cooling*, 1987; G. Vacuumschmelze, *Thermostat Metals*, 1984.

Thermotherapy

The treatment of disease by the local or general application of heat. The following discussion is limited to the local application of heat as an adjunct to therapeutic management of musculoskeletal and joint diseases. The most commonly used methods for this form of treatment include hot packs, hydrotherapy, radiant heat, shortwave diathermy, microwave diathermy, ultrasound, and laser therapy.

Biophysics. The reason so many different methods are employed is that each modality heats selectively different anatomical structures, and thus the modality selected for a given treatment is based on the temperature distribution produced in the tissues. For vigorous heat application to a given site, the location of the peak temperature produced by the modality must coincide with the site so that maximally tolerated tissue temperatures can be obtained there without burning elsewhere. Customarily, the modalities are divided into those that heat superficial tissues and those that heat deep-seated tissues.

Heating superficial tissues. Hot packs, hydrotherapy, and radiant heat are used to heat superficial tissues. The heat is transferred from a hot pack or hot water, primarily through conduction and convection, whereas with application of radiant heat, the photons, or electromagnetic radiation, emitted from the lamp are converted into heat by absorption. However, since absorption occurs in the most superficial tissues, this form of therapy is classified as superficial heating. The photons used—from the yellow and red bands of the visible spectrum and the infrared band within the invisible range—have a relatively long wavelength and low energy content. *See* ELECTROMAGNETIC RADIATION.

Heating deep-seated tissues. Photons of ultraviolet radiation, x-rays, and radium penetrate deeper into the tissues and produce photochemical reactions long before the temperature increases significantly. Other forms of energy used for heating deep-seated tissues include shortwave, microwave, and ultrasound: shortwave diathermy induces a high-frequency (27-MHz) electromagnetic current in the tissues; microwaves operate at frequencies of 915 and 2456 MHz; and ultrasonic waves represent an acoustic vibration of 1-MHz frequency.

In the case of shortwaves, the electrical properties of the tissue determine the distribution and absorption of current and thus the heating pattern. Shortwaves may heat joints covered by very little soft tissue; superficial musculature; or, if applied with internal electrodes, the pelvic organs. Microwaves are reflected at tissue interfaces and are absorbed; microwave application generally heats selectively tissues with high water content. Ultrasound is reflected at interfaces between tissues of different acoustic impedance, which is the product of density times sound velocity. Ultrasound selectively heats these interfaces because (1) the reflected wave is superimposed upon the incoming wave and increases the available energy for absorption and (2) a significant part of the longitudinal compression wave of ultrasound is converted into shear waves at the interface. The shear waves are locally absorbed rapidly. Finally, the interface between a medium with low coefficient of absorption (such as soft tissue) and a medium with high coefficient (such as bone) is heated selectively, because most of the energy is absorbed in the superficial layers of the medium with high coefficient. By using these principles, ultrasound selectively heats deep-seated joints. *See* MICROWAVE; ULTRASONICS.

Therapeutic effects of heat. The therapeutic effects produced by heating selectively by ultrasound include an increase in the extensibility of collagen tissues. Disease or injury, such as arthritis, burns, scarring, or long-term immobilization in a cast, may cause shortening of collagen tissue producing severe limitation of the range of motion at a joint. This application of heat (mostly ultrasound) is often used in conjunction with physical therapy, such as stretching and other joint mobilization techniques. It also has been shown that the stiffness associated with rheumatoid arthritis can be alleviated by heat application.

Heat applied by using shortwaves and microwaves may reduce muscle spasms secondary to musculoskeletal pathology, such as when a slipped intervertebral disc impinges upon a nerve root. (Interestingly, cooling the area may have the same physiological effect.) Selective heating of muscle with microwave radiation has been used to accelerate absorption of hematomas (a collection or clot of blood in the muscle) and to prepare for stretching of the contracted and stiffened muscle. Unlike microwaves, ultrasound is not absorbed significantly in normal homogeneous muscle.

Heat therapy in the form of hyperthermia has been used as an effective adjunct to cancer therapy in combination with ionizing radiation in the form of x-rays or radium therapy.

Laser therapy. The most commonly used laser in physical therapy is the helium-neon laser, which produces photons of a wavelength of 632.3 nanometers at lower intensities than those applied in surgical procedures. At this wavelength the depth of penetration and the heating effect are similar to infrared light. However, the major difference between laser light and diffuse light of the same wavelength is that laser is a columnated beam of photons of the same frequency, with the wavelength in phase so that any desirable intensity can be easily produced. *See* LASER PHOTOBIOLOGY.

Side effects. All the therapies described are valuable adjuncts to therapy with drugs and other approaches. However, the energy output of all these

modalities is high enough to produce tissue destruction if not used with great care. Examples of other significant side effects are the malfunction of pacemakers as a result of shortwave or microwave application, the overheating of metal implants and the tissues surrounding them, and the production of destructive lesions in the eye when using diathermy. Ultrasound may produce cavitation (gas bubbles) in the fluid media of the eye, with possible destruction of the retina and blood vessels. High-intensity lasers may burn, destroy, or produce the type of destructive effects associated with ionizing radiation.

Justus F. Lehmann

Bibliography. F. J. Kottke and J. F. Lehmann (eds.), *Krusen's Handbook of Physical Medicine and Rehabilitation*, 4th ed., 1990; J. F. Lehmann (ed.), *Therapeutic Heat and Cold*, 4th ed., 1990.

Thévenin's theorem (electric networks)

A theorem from electric circuit theory. It is also known as the Helmholtz or Helmholtz-Thévenin theorem, since H. Helmholtz stated it in an earlier form prior to M. L. Thévenin. Closely related is the Norton theorem, which will also be discussed. Laplace transform notation will be used. *See* LAPLACE TRANSFORM.

Thévenin's theorem states that at a pair of terminals a network composed of lumped, linear circuit elements may, for purposes of analysis of external circuit or terminal behavior, be replaced by a voltage source $V(s)$ in series with a single impedance $Z(s)$. The source $V(s)$ is the Laplace transform of the voltage across the pair of terminals when they are open-circuited; $Z(s)$ is the transform impedance at the two terminals with all independent sources set to zero (**Fig. 1**). The Thévenin equivalent may also be found experimentally.

Norton's theorem states that a second equivalent network consists of a current source $I(s)$ in parallel with an impedance $Z(s)$. The impedance $Z(s)$ is identical with the Thévenin impedance, and $I(s)$ is the Laplace transform of the current between the two terminals when they are short-circuited (**Fig. 2**).

Thévenin's and Norton's equivalent networks are related by the equation $V(s) = Z(s) \cdot I(s)$. This may be seen by comparing Figs. 1*b* and 2*b*. In Fig. 1*b*, if terminals A and B are short-circuited, a current $I(s) = V(s)/Z(s)$ will flow; this is also true in Fig. 2*b*. Similarly the open-circuit voltage in Fig. 2*b* is $V(s) = Z(s) \cdot I(s)$. *See* ALTERNATING-CURRENT CIRCUIT THEORY.

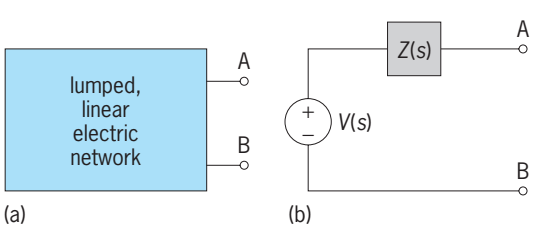

Fig. 1. Network and its Thévenin equivalent. (*a*) Original network. (*b*) Thévenin equivalent circuit.

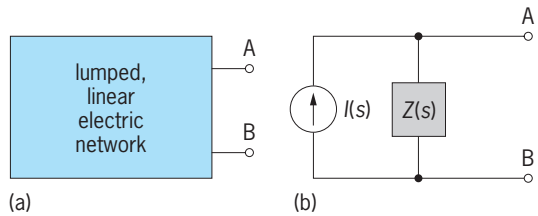

Fig. 2. Network and its Norton equivalent. (*a*) Original network. (*b*) Norton equivalent circuit.

These theorems are useful for the study of the behavior of a load connected to a (possibly complex) system that is supplying electric power to that load. The system may be a power distribution system, such as in a home or office, in which case the load may be lights or appliances. The system may be an electronic amplifier, in which case the load may be a loudspeaker. However, the theorem is of no value in studying the internal system behavior, because the behavior of the equivalent network is very different from that of the original.

Examples. Two examples will be used to show how Thévenin and Norton equivalent networks may be calculated from the original network and then used for some typical calculations.

Power distribution circuit. Suppose a simplified power distribution circuit contains the elements shown in **Fig. 3***a*. Sinusoidal steady-state operation is assumed. In this circuit the voltage across A, B is given by Eq. (1), and with the source set to zero, the impedance at A, B is given by Eq. (2). Thus the

$$\begin{aligned} V_{AB} &= \frac{Z_3}{Z_1 + Z_2 + Z_3} \cdot V_s \\ &= \frac{-j122.6}{0.55 + j3.5 - j122.6} \cdot 240\underline{/0^\circ} \\ &= 247.1\underline{/-0.26^\circ}\ V \end{aligned} \tag{1}$$

$$\begin{aligned} Z_{TH} &= Z_4 + \frac{Z_3(Z_1 + Z_2)}{Z_1 + Z_2 + Z_3} \\ &= 0.35 + \frac{(-j122.6)(0.55 + j3.5)}{0.55 + j3.5 - j122.6} \\ &= 3.72\underline{/75.47^\circ}\ \text{ohms} \end{aligned} \tag{2}$$

Thévenin equivalent is given in Fig. 3*b*, and the Norton equivalent in Fig. 3*c*.

When a load $Z_L = 18.3 + j2.1 = 18.42\underline{/6.55^\circ}$ is connected, at A, B, the current through the load is given by Eq. (3), and the power delivered to the load

$$\begin{aligned} I_L &= \frac{V_{TH}}{Z_{TH} + Z_L} = \frac{247.1\underline{/-0.26^\circ}}{3.72\underline{/75.47^\circ} + 18.42\underline{/6.55^\circ}} \\ &= (12.32)^2\underline{/-16.77^\circ} \end{aligned} \tag{3}$$

is $(12.32)^2(18.3) = 2.776$ kW. Other loads are handled in a similar fashion.

Amplifier. As a second example, suppose that the circuit of **Fig. 4***a* is a simplified model of an electronic amplifier, and that a load (loudspeaker) is to be connected at A, B. The circuit is driven by a current source $I(s)$. For analysis, the voltages $V_1(s)$ and

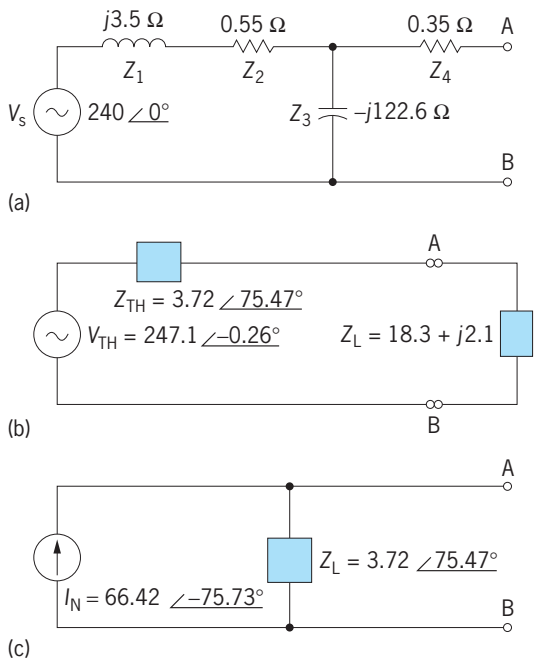

Fig. 3. Power distribution circuit and its Thévenin and Norton equivalents. (*a*) Original circuit. (*b*) Thévenin equivalent circuit with its load Z_L connected. (*c*) Norton equivalent circuit.

$V_2(s)$ are the transforms of the voltages on their respective nodes, and become the dependent variables in the analysis. The dependent or controlled source (g_mV_1) models the amplification. Two Kirchhoff current law equations, (4) and (5), may be used to find the Thévenin voltage, which is also V_2. Solution of this pair of equations gives Eq. (6).

$$I(s) = V_1\left(\frac{1}{R_1} + \frac{1}{R_2}\right) - V_2\left(\frac{1}{R_2}\right) \tag{4}$$

$$-g_mV_1 = -V_1\left(\frac{1}{R_2}\right) + V_2\left(\frac{1}{R_2} + \frac{1}{R_3}\right) \tag{5}$$

$$V_2(s) = \frac{(-g_m + 1/R_2)\cdot I(s)}{\dfrac{1}{R_1R_2} + \dfrac{1}{R_1R_3} + \dfrac{1}{R_2R_3} + \dfrac{g_m}{R_2}} \tag{6}$$

See KIRCHHOFF'S LAWS OF ELECTRIC CIRCUITS.

To find the Thévenin impedance, $I(s)$ must be set to zero, which leaves an infinite impedance in the branch, and an auxiliary current source $I_2(s)$ must be added between A and B (Fig. 4*b*). A new set of equations, (7) and (8), is written, and solved for

$$0 = V_1\left(\frac{1}{R_1} + \frac{1}{R_2}\right) - V_2\left(\frac{1}{R_2}\right) \tag{7}$$

$$-g_mV_1 + I_2(s) = -V_1\left(\frac{1}{R_2}\right) + V_2\left(\frac{1}{R_2} + \frac{1}{R_3}\right) \tag{8}$$

the ratio $V_2(s)/I_2(s)$, which is the desired Thévenin impedance; the controlled source must not be set to zero.

Solution of Eqs. (7) and (8) gives, after simplification, Eq. (9).

$$Z_{TH}(s) = \frac{V_2(s)}{I_2(s)} = \frac{R_3(R_1 + R_2)}{R_1 + R_2 + R_3 + g_mR_1R_3} \tag{9}$$

To make a numerical example in a purely resistive circuit, let $R_1 = 2.0\ k\Omega$, $R_2 = 8.2\ k\Omega$, $R_3 = 400\ \Omega$, $g_m = 510 \times 10^3$ S (siemens), and $I = 4.0$ mA. Substitution of these into the equations gives a Thévenin voltage $V_{TH} = -31.96$ V, and a Thévenin impedance $Z_{TH} = 9.75\ \Omega$. This is shown in Fig. 4*c*, and the corresponding Norton equivalent is shown in Fig. 4*d*, where $I_N = -31.96/9.75 = -3.28$ A. In both cases the negative signs lead to a polarity reversal, which is reflected in Fig. 4*c* and *d*.

Suppose an 8.0-ohm speaker is connected at A, B. A current of 31.96/(8.0 + 9.75) = 1.80 A will flow, giving a power of $(1.80)^2(8.0) = 25.95$ W.

Proof. To prove this theorem, consider a general network with two accessible terminals, as in Figs. 1*a* and 2*a*, to which an auxiliary voltage source V and an impedance $Z_L(s)$ have been added (**Fig. 5**). Let this source be such as to cause $I_L = 0$ when switch S is closed. By superposition, Eq. (10) is valid. This

$$I_L = 0 = \frac{V_{TH}}{Z_{TH} + Z_L} - \frac{V}{Z_{TH} + Z_L} \tag{10}$$

shows that $V_{TH} = V$, and that the current that flows

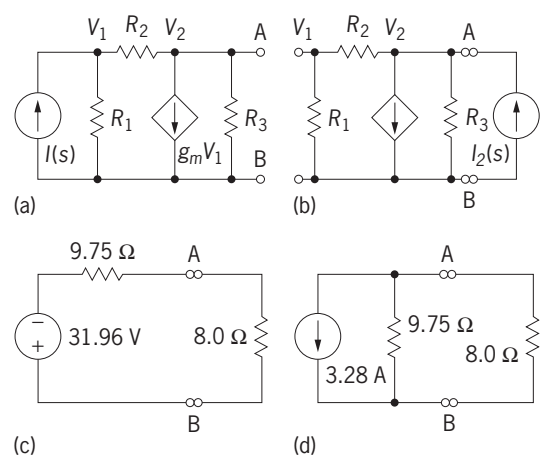

Fig. 4. Amplifier and its Thévenin and Norton equivalents. Numerical values are given in the text. (*a*) Original circuit. (*b*) Circuit constructed to find Thévenin impedance. (*c*) Thévenin equivalent circuit. (*d*) Norton equivalent circuit.

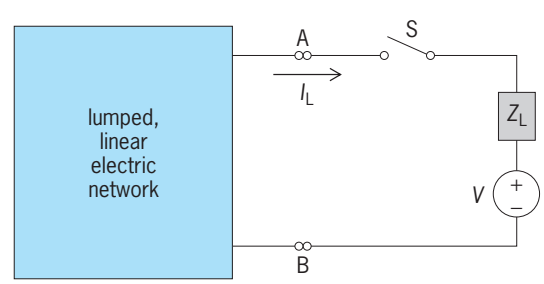

Fig. 5. Circuit constructed to demonstrate Thévenin's theorem.

when V is removed is given by Eq. (11).

$$I_L = \frac{V_{TH}}{Z_{TH} + Z_L} \tag{11}$$

See NETWORK THEORY; SUPERPOSITION THEOREM (ELECTRIC NETWORKS). Edwin C. Jones, Jr.

Bibliography. R. DeCarlo and P. -M. Lin, *Linear Circuit Analysis*, 2d ed., 2001; J. D. Irwin and C.-H. Wu, *Basic Engineering Circuit Analysis*, 7th ed., 2001; M. Reed and R. Rohrer, *Applied Introductory Circuit Analysis*, 1999.

Thiamine

A water-soluble vitamin, also known as vitamin B_1 or aneurin, found in many foods; pork, liver, and whole grains are particularly good sources. Loss of the vitamin during production of flour and polished rice has led to the need for enrichment of these staples. The structural formula is shown below.

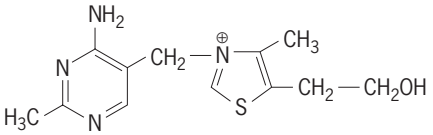

Chemistry. Thiamine is heat-labile, and considerable amounts are destroyed during cooking. It is unstable in alkaline solutions but relatively stable in acid solutions. It acts like a weak base and can be absorbed on basic ion-exchange materials such as decalso and fuller's earth, a property used to concentrate it so that it can be detected at the levels needed for analysis. Biological and microbiological methods for its estimation are available but are seldom used. A chemical assay based on the production of thiochrome, a fluorescent product obtained from alkaline oxidation of thiamine, is often used. The red blood cell transketolase assay, which is dependent upon the level of the coenzyme of thiamine, is the most commonly used functional assay for thiamine status.

Biochemistry. Thiamine functions as the coenzyme thiamine pyrophosphate (TPP) in enzyme systems that catalyze decarboxylations of α-keto acids and ketolations (condensation reactions of ketols) in sugar phosphates. Thiamine pyrophosphate–containing enzymes located in mitochondria catalyze the decarboxylation of pyruvate, α-ketoglutarate, and branched-chain α-ketoacids to eventually form acetyl-coenzyme A (CoA), succinyl-CoA, and amino acids (such as leucine, isoleucine, and valine), respectively (see **illustration**). These activities in higher organisms, including humans, are in association with other enzyme subunits that constitute so-called dehydrogenase complexes. The coenzyme TPP is also involved in the cytosol, where it associates with transketolase to effect interconversions of pentose to heptulose in the hexose monophosphate shunt, which is an alternative to the conventional glycolytic pathway of anaerobic glucose metabolism. Participation of thiamine as TPP in metabolic pathways is shown in the illustration.

A recently characterized lyase that is required to shorten 3-methyl-branched fatty acids by α-oxidation also requires TPP. *See* CARBOHYDRATE METABOLISM; CITRIC ACID CYCLE; COENZYME; ENZYME.

Deficiency. Thiamine deficiency is known as beriberi in humans. Other species manifest the deficiency with polyneuritic conditions (which are characterized by degenerative or inflammatory lessions of several nerves simultaneously, usually symmetrical). Muscle and nerve tissues are affected by the deficiency, and poor growth is observed. People with beriberi are irritable, depressed, and weak; they often die of cardiac failure. Wernicke-Korsakoff syndrome, often observed in chronic alcoholics, is a cerebral beriberi characterized by brain lesions, liver disease, and partial paralysis, particularly of the motor nerves of the eye. As is the case in other vitamin B deficiencies, the deficiency of thiamine is commonly accompanied by insufficiencies of some of the other vitamins. In relatively rare cases of inborn errors of metabolism, pharmacologic doses of thiamine (5 to 10 mg/day) are required to prevent lactic acidosis due to low activity of the liver pyruvate dehydrogenase complex and ketoaciduria (the presence of keto acids in the urine) due to low activity of the branched-chain α-keto-acid dehydrogenase complex. *See* MALNUTRITION.

Dietary requirements. A dietary source of thiamine is required by all nonruminant animals that have been studied. Thiamine is the most poorly stored of the B vitamins. Individuals eating vitamin-deficient diets are likely to develop beriberi symptoms first. Approximately 5 mg of thiamine can be absorbed per day by normal adults. Excess thiamine given by mouth or parenterally is well tolerated but usually lost through excretion in the urine and feces. Thiamine requirements are related to caloric intake. More thiamine is required in high-carbohydrate than in high-fat diets. Some foods, particularly raw fish, contain enzymes that destroy thiamine. More thiamine is needed in altered physical states, such as hyperthyroidism, pregnancy, and lactation. Thiamine requirements of humans are primarily estimated by means of urinary excretion data and the red cell transketolase assay. The dietary allowances recommended in the United States by the Food and Nutrition Board of the Institute of Medicine are from 1.1 to 1.2 mg of thiamine per day for adults, with an additional increase to 1.4 mg/day for pregnant or breast-feeding women. *See* VITAMIN.

Stanley N. Gershoff; Donald B. McCormick

Manufacture. Industrial synthesis of thiamine is accomplished by linking chloromethylpyrimidine with 4-methyl-5-(β-hydroxyethyl)-thiazole to give thiamine. Another way to produce thiamine on an industrial scale is to convert 4-amino-5-cyanopyrimidine into the thioformyl-aminomethyl derivative via catalytic hydrogenation and reaction with sodium

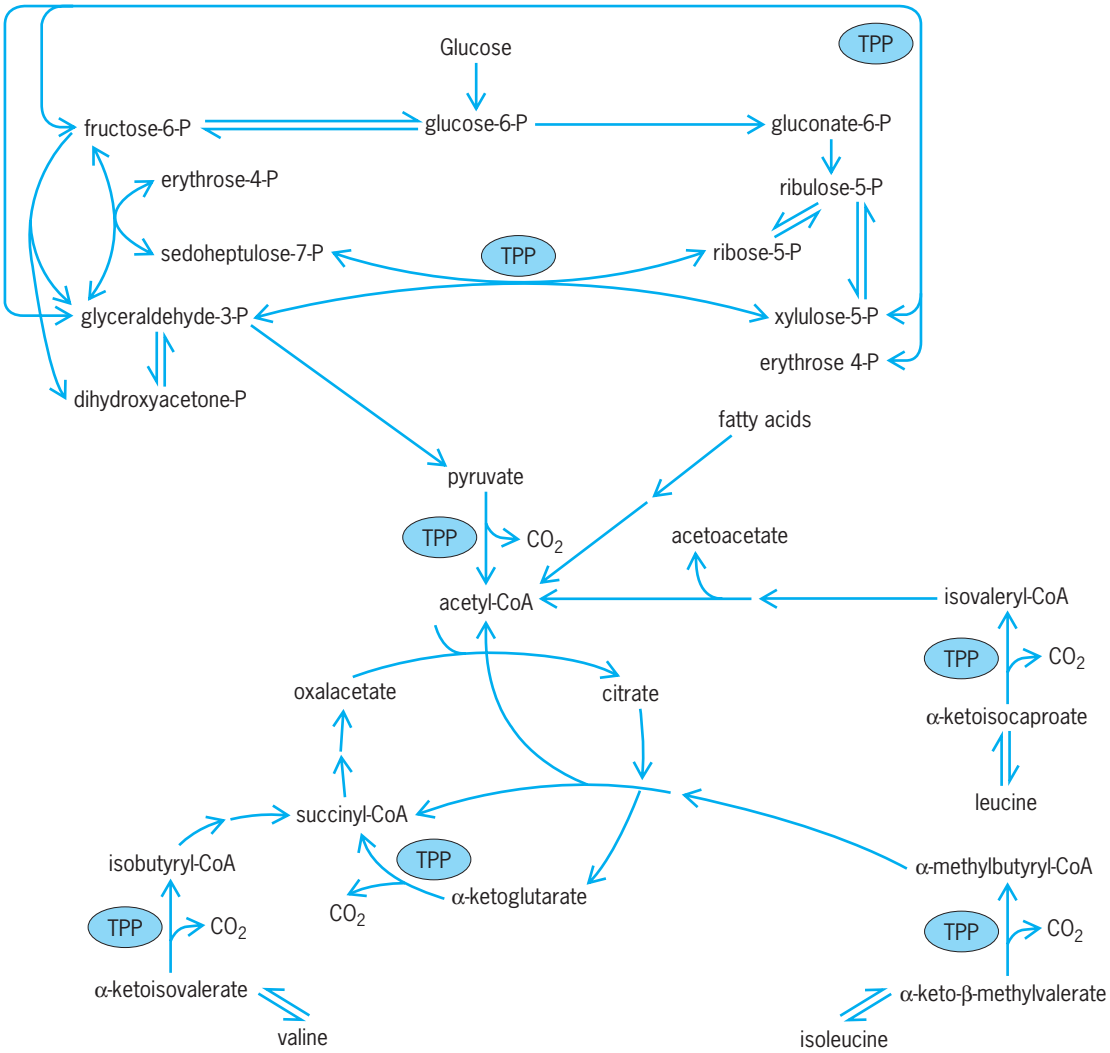

Metabolic pathways involving thiamine pyrophosphate (TPP). CoA, coenzyme A; P, phosphate.

dithioformate. The resulting compound is then treated with 1-acetoxy-3-chloro-4-pentanone to form the thiazole ring in situ connected to the pyrimidine ring via a methylene bridge (U.S. Patents 2,193,858 and 2,218,350.) Fernand de Montmollin

Bibliography. C. J. Bates, Thiamin, in B. A. Bowman and R. M. Russell (eds.), *Present Knowledge in Nutrition*, 8th ed., pp. 184–190, ILSI Press, Washington, D.C., 2001; Food and Nutrition Board, Institute of Medicine, Thiamin, in *Dietary Reference Intakes: Thiamin, Riboflavin, Niacin, Vitamin B_6, Vitamin B_{12}, Pantothenic acid, Biotin, and Choline*, pp. 58–86, National Academy Press, Washington, D.C., 1998; J. Higdon, Thiamin, in *An Evidence-Based Approach to Vitamins and Minerals: Health Benefits and Intake Recommendations*, pp. 33–38, Thieme, New York, 2003; D. Lonsdale, *A Nutritionist's Guide to the Clinical Use of Vitamin B-1*, 1988; V. Tanphaichitr, Thiamin, in M. E. Shils et al., (eds.), *Modern Nutrition in Health and Disease*, 9th ed., pp. 381–389, Williams and Wilkins, Baltimore, 1999.

Thick-film sensor

A sensor that is based on a thick-film circuit. Thick-film circuits are formed by the deposition of layers of special pastes onto an insulating substrate. The pastes are usually referred to as inks, although there is little resemblance to conventional ink. The printed pattern is fired in a manner akin to the production of pottery, to produce electrical pathways of a controlled resistance. Parts of a thick-film circuit can be made sensitive to strain or temperature. The thick-film pattern can include mounting positions for the insertion of conventional silicon devices, in which case the assembly is known as a thick-film hybrid. The process is relatively cheap, especially if large numbers of devices are produced, and the use of hybrid construction allows the sensor housing to include sophisticated signal conditioning circuits. These factors indicate that thick-film technology is likely to play an increasing role in sensor design.

The three main categories of thick-film inks are

conductors, dielectrics (insulators), and resistors. Conductors are used for interconnections, such as the wiring of bridge circuits. Dielectrics are used for coating conducting surfaces (such as steel) prior to laying down thick-film patterns, for constructing thick-film capacitors, and for insulating crossover points, where one conducting path traverses another. Resistor inks are the most interesting from the point of view of sensor design, since many thick-film materials are markedly piezoresistive.

Fabrication process. The main constituents of a thick-film ink are the binder (a glass frit), the vehicle (an organic solvent), and the active elements (metallic alloys or oxides). After printing, each layer of a thick-film pattern is dried to remove the organic solvents (the vehicle), which give the ink its viscosity. Drying also improves the adhesion properties, bonding the ink to its substrate and rendering the pattern immune to smudging. This stage is usually performed in a conventional oven at 100–150°C (212–302°F).

A final high-temperature firing is required to remove any remaining solvent and to sinter the binder and the active elements. During the firing cycle a thick-film pattern is raised to 500–1000°C (932–1832°F). The glass frit melts, wets the substrate, and forms a continuous matrix, which holds the functional elements. The heating and cooling gradients, the peak temperature, and the dwell time determine the firing profile. This has a critical effect on the production of a thick-film circuit, since it allows the electrical characteristics of the inks to be modified. Resistor materials are especially sensitive to the firing profile, and the resistor layer is usually therefore the last to be fired. However, the need for passivation of a circuit often necessitates covering it with a dielectric layer. To avoid changing the resistor values, a low-melting-point dielectric is often used for the final layer. *See* SINTERING.

The need for high-temperature firing can cause problems if thick-film materials are to be applied to previously heat-treated components. The temperatures used can adversely affect, for example, the properties of toughened or hardened steels.

Thick-film circuits and sensors are created by screen printing. This is essentially a stencil process, in which the printing ink is forced through the open areas of a mesh-reinforced screen onto the surface of a substrate. The screen stencils are formed by photolithography. In this process a photosensitive mesh-filling material is exposed to ultraviolet light through a mask depicting the required pattern. The image is photographically developed, and those parts of the pattern which have not been fixed are subsequently washed away.

The use of thick-film technology was originally introduced as a means of miniaturizing circuits without incurring the expense associated with fabrication in silicon. It was soon noted that thick-film materials had temperature- and stress-dependent properties. Although this was awkward from the point of view of circuit fabrication, it has since been turned to good account in sensor design. The piezoresistive properties of thick-film resistor inks can be used to form strain sensors. This approach is commonly used to manufacture pressure sensors and is exploited to produce accelerometers. *See* ACCELEROMETER; PRESSURE TRANSDUCER; STRAIN GAGE.

Piezoresistive sensors. The basis of most strain-based thick-film transducers is the piezoresistive effect. A piezoresistive material is one in which a change in electrical resistance occurs in response to changes in the applied stress.

Piezoresistive sensors are formed by placing stress-sensitive resistors on highly stressed parts of a suitable mechanical structure. The piezoresistive transducers are usually attached to cantilevers, or other beam configurations, and are connected in a Wheatstone bridge circuit. The beam may carry a seismic mass to form an accelerometer or may deform in response to an externally applied force. The stress variations in the transducer are converted into an electrical output, which is proportional to strain, by the piezoresistive effect. *See* WHEATSTONE BRIDGE.

Piezoresistive devices are relatively easy to construct, provide a low-frequency response extending down to dc (0 frequency), and work well over a relatively large temperature range (−50 to +150°C or −58 to 302°F). Another valuable feature is their ability to include signal-processing and communication functions within the sensor package at little extra cost.

The drawbacks of piezoresistive devices are that the output signal level is moderate (typically 100 mV full scale for a 10-V bridge excitation), the sensitivity can be temperature-dependent, and the usable bandwidth is not as large as that which may be obtained from a piezoelectric transducer.

Analysis of piezoresistance. If a rectilinear resistor has length l, width w, thickness t, and bulk resistivity ρ, its resistance R is given by Eq. (1). The gauge factor or strain sensitivity is defined as k in Eq. (2),

$$R = \frac{\rho l}{wt} \tag{1}$$

$$k = \frac{dR/R}{\varepsilon} \tag{2}$$

where ε is the relative change in length of the resistor (the strain) due to a stress, σ, applied to the substrate parallel to its length. The **illustration** shows the consequences of the applied stress. The length increases by an amount dl, while the width and thickness decrease by amounts dw and dt determined by Poisson's ratio v: $dw = vw\varepsilon$ and $dt = vt\varepsilon$. *See* ELASTICITY.

The original cross section is given by Eq. (3). Owing to the strain, the new cross-sectional area is given

$$A = wt \tag{3}$$

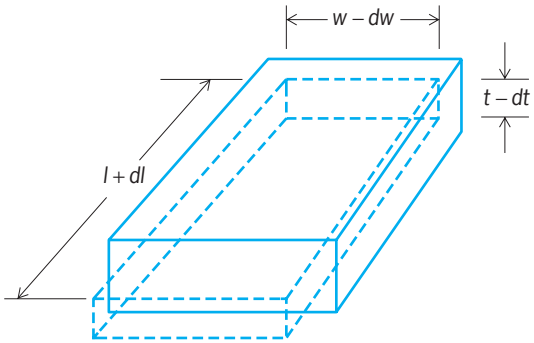

Changes in the dimensions of a rectangular resistor (length *l*, width *w*, and thickness *t*) due to a stress applied parallel to the length.

by Eq. (4). The term $v^2wt\varepsilon$ is very small compared

$$A' = (w - dw)(t - dt) = wt - 2vwt\varepsilon + v^2wt\varepsilon^2 \tag{4}$$

with the other two terms in Eq. (4) and can be neglected. Therefore, the change in cross-sectional area can be written as Eq. (5), given Eq. (6). Differentiating Eq. (1) gives Eq. (7), and hence the gauge factor k is given by Eq. (8).

$$A' - A = dA = -2v\varepsilon A \tag{5}$$

$$\frac{dA}{A} = -2v\varepsilon \tag{6}$$

$$\frac{dR}{R} = \frac{d\rho}{\rho} + \frac{dl}{l} - \frac{dA}{A} \tag{7}$$

$$k = \frac{d\rho/\rho}{\varepsilon} + (l + 2v) \tag{8}$$

Typically v will be between 0.2 and 0.3. Equation (8) therefore shows that the longitudinal gauge factor is a function of changes in both longitudinal resistivity and geometry. In conventional foil or wire strain gauges, the piezoresistive effects are negligible, and the variations in resistance are mainly a function of dimensional changes. For a foil gauge, k is approximately 2. For piezoresistive strain gauges, the first term in Eq. (8) is significant, and higher gauge factors (typically around 10) can be achieved, giving enhanced sensitivity. However, the resistivity of most piezoresistive materials is strongly temperature-dependent, and as a result piezoresistive strain gauges generally have a higher thermal sensitivity than other types.

Temperature sensors. The linear temperature coefficient of resistance possessed by certain platinum-containing conductive inks has allowed thermistors to be printed onto suitable substrates using thick-film fabrication techniques. Thick-film thermistors are very inexpensive and physically small, and have the further advantage of being more intimately bonded to the substrate than a discrete component. It has been shown that thick-film thermistors can have as good, if not better, performance than a comparable discrete component. *See* THERMISTOR.

Chemical sensors. Thick-film materials have been used for a number of chemical sensing applications, including the measurement of gas and liquid composition, acidity (pH), and humidity. It is difficult to define a comprehensive scheme for the classification of thick-film chemical sensors, since there are so many variants. However, a classification based on two categories seems to cover most devices: impedance-based transducers, in which the measurand causes a variation of resistance, capacitance, and so forth; and electrochemical systems, in which the sensed quantity causes a change in electrochemical potential or current. *See* ELECTRICAL IMPEDANCE; ELECTROCHEMISTRY.

Impedance-based chemical sensors. Thick-film gas sensors can be fabricated from a printable paste of semiconducting metal-oxide powder, inorganic additives, and organic binders. The paste is printed over an alumina substrate containing metal film electrodes and a back-heating resistor; the paste is then fired in an infrared or thermal belt furnace. The most frequently used material for this type of sensor is tin oxide (stannic oxide, SnO_2). *See* TIN.

Humidity sensors. Several thick-film humidity sensors have been reported in which capacitive effects are used. A common approach is to fabricate the sensor in the form of a pair of closely spaced interdigitated electrodes, which are screen-printed and fired onto an alumina substrate. The dielectric layer to be tested is then applied on top of this structure, and capacitance changes of up to 3 nanofarads are observed for porous coatings. *See* CAPACITANCE; TRANSDUCER.

John D. Turner

Bibliography. J. D. Turner and M. Hill, *Instrumentation for Engineers and Scientists*, Oxford University Press, 1999; N. M. White and J. D. Turner, Thick-film sensors: Past, present and future, *Meas. Sci. Technol.*, 8(1):1–20, January 1997.

Thickening

The production of a concentrated slurry from a dilute suspension of solid particles in a liquid. In practice, a thickener also usually generates a clear liquid; therefore clarification is generally a concurrent process. Thickening and clarification are outcomes of sedimentation, and both are representative of a group of industrial processes termed mechanical separations. *See* CLARIFICATION; SEDIMENTATION (INDUSTRY).

Operation. Although thickening may be carried out either batchwise or continuously, the latter method is more common. Special means are used to move the concentrated slurry to the discharge outlet in the mechanically agitated continuous thickener (**Fig. 1**). This equipment consists of a settling tank fitted with slow-moving rakes driven by a vertical control shaft. The tank may have a flat bottom or a shallow-cone bottom. Continuous thickeners may be large; diameters range from 20 to 300 ft (6 to 90 m) and the depth from 8 to 12 ft (2.5 to 4 m). Small tanks may be made of wood or steel, intermediate ones

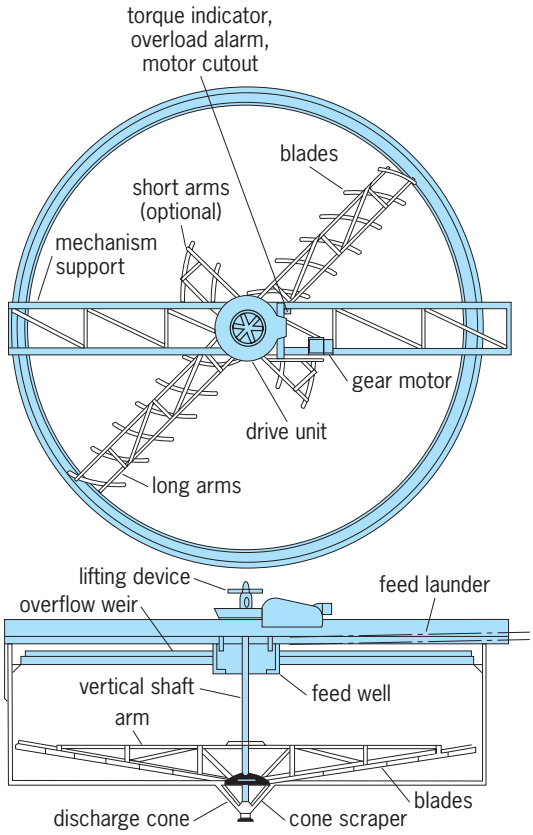

Fig. 1. Top and side views of mechanically agitated, continuous thickener. (*After W. L. McCabe, J. C. Smith, and P. Harriott, Unit Operations of Chemical Engineering, 4th ed., McGraw-Hill, 1985*)

of steel, and large units of concrete. In large thickeners, the rakes may rotate only once every 30 min.

In operation, dilute feed pulp is admitted continuously through a launder to a central well immersed to a depth of 2 to 3 ft (0.5 to 1 m) below the surface of the liquid in the tank. The liquid from the feed moves radially to the wall of the tank and overflows across a weir that bounds the periphery of the tank. As the liquid flows radially from the feed well, the solids it carries settle, so that the liquid is clear when it reaches the overflow weir. The solids settle to the bottom of the tank and form a concentrated slurry. The rakes, without repulping the solids into the liquid, gently agitate the solids, break up the flocs to aid the process of consolidation, and move the thickened solids to the discharge in the center of the tank bottom. From the discharge, the thickened slurry flows to the suction side of a sludge pump.

For a continuous thickener, the concentration of solids in the liquid decreases from the top to the bottom of the tank. There are two main zones, the free-settling and compression, which are separated by an interface (**Figs. 2** and **3**). The upper zone, in which clarification is accomplished, is free from solid in its top layers and supplies the clarified liquid overflow. Within the clarification zone, the solid concentration varies from zero to a low value at the interface between the zones. In the clarification zone, the solid particles are sufficiently far apart that free settling takes place. The bottom zone, in which thickening is accomplished, contains most of the inventory of solid in the tank. Here the concentration of solid changes rapidly from that in the clarification zone to that of the thickened slurry leaving the thickener. The process occurring in this zone is essentially compression.

To attain satisfactory capacities, the feed to a thickener is often flocculated. The performance of a given plant operating on a specified feed slurry depends largely on the major dimensions of the tank. To obtain a clear overflow, the upward velocity of the liquid in the clarification zone must be less than the minimum terminal settling velocity of the smallest particles. Then the capacity of the unit to handle clarified liquid is proportional to the horizontal cross-sectional area and, therefore, to the square of the diameter of the tank. The solid concentration in the underflow, and hence the degree of thickening achieved, depends on the time allowed for action in the compression zone. Once the feed rate of dilute slurry is fixed, the time for compression is proportional to the height of the compression zone. Thus

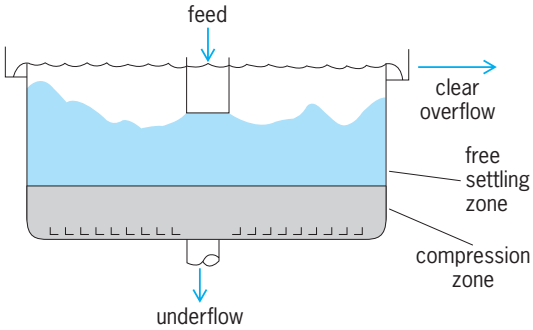

Fig. 2. Zones in continuous thickener. (*After W. L. McCabe and J. C. Smith, Unit Operations of Chemical Engineering, 2d ed., McGraw-Hill, 1967*)

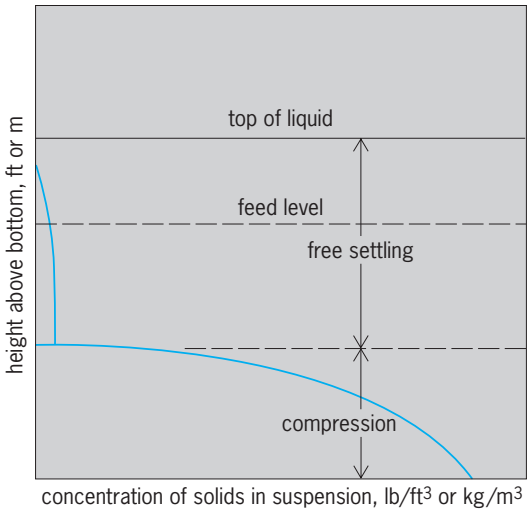

Fig. 3. Variation of solids concentration with the height in the tank of a continuous thickener. (*After W. L. McCabe, J. C. Smith, and P. Harriott, Unit Operations of Chemical Engineering, 4th ed., McGraw-Hill, 1985*)

the performance of the unit as a thickener is a function of tank depth.

Types. Thickeners are especially useful when large volumes of dilute slurries must be treated, as in manufacture of cement, production of magnesium from seawater, treatment of sewage, purification of water, treatment of coal, and dressing of metallurgical ores.

Several variants of the circular, continuous thickener are widely used. One type is fitted with multitray thickeners where flows are in series or parallel. This design reduces base area and also provides either more depth for consolidation of the thickened discharge or more overflow area for clarification. Another type, the National Coal Board deep-cone thickener, is used for processing coal and metallurgical ores. In the Swedish Lamella thickener, the liquid flows upward between inclined plates stacked closely together; the feed enters the stack from a side feed box. In theory the effective settling area is the sum of the horizontal projected areas of all the plates, but it is only about 50% efficient. The sludge can be further consolidated by vibration or raking. *See* MECHANICAL SEPARATION TECHNIQUES.

Vincent W. Uhl

Bibliography. M. C. Bustos et al., *Sedimentation and Thickening: Phenomenological Foundation and Mathematical Theory*, 1999; D. S. Jones, *Elements of Chemical Process Engineering*, 1996; W. McCabe, J. Smith, and P. Harriott, *Unit Operations of Chemical Engineering*, 7th ed., 2004; R. H. Perry and D. W. Green (eds.), *Perry's Chemical Engineers' Handbook*, 7th ed., 1997; J. F. Richardson, J. H. Harker, and J. Backhurst, *Chemical Engineering*, vol. 2, 5th ed., 2002; R. K. Sinnott, *Chemical Engineering Design*, vol. 6, 4th ed, 2005.

Thinner

A material used in paints and varnishes to adjust the consistency for application. Thinners are usually solvents for the vehicle used in the coating and are expected to evaporate after application. Because their only function is to make the application simple, it is important that their cost be low. Water is used as a thinner in emulsion paints and in certain water-soluble paints such as watercolors and calcimines.

Petroleum fractions are most commonly used for oil and resin coatings. The fraction boiling between 300 and 400°F (149 and 204°C), called mineral spirits, is most widely used. A lower-boiling and faster-evaporating solvent is called VM&P (varnish makers' and painters') naphtha. Still faster-evaporating materials are called petroleum ether, lacquer diluent, or rubber solvent. Stronger solvents contain substantial amounts of aromatic hydrocarbons and may be derived from petroleum or coal tar. These may be essentially pure materials, such as toluene or xylene, or mixtures designed to have the solvency and evaporation characteristics desired.

Since numerous coating resins are not sufficiently soluble in hydrocarbons, other materials or mixtures must be used. These include alcohols such as denatured ethyl or isopropyl alcohols for shellac, esters such as amyl acetate for nitrocellulose, and ketones and other compounds for acrylic and vinyl resins. Chlorinated hydrocarbons are used for some materials which are otherwise hard to dissolve, but toxicity limits their usefulness.

The selection of a thinner for a coating formulation depends upon the resins used, the application and curing conditions, and the effects desired. For example, fast-drying solvents will reduce the temperature of the surface, and under humid conditions they may cause moisture to condense on the surface, producing the phenomenon known as blushing. Vinyl, epoxy, and chlorinated rubber, among other types of coatings, commonly employ a mixture of thinners known as a solvent blend. These contain materials having different evaporation rates designed to hold the film open to avoid solvent entrapment and subsequent blistering, or to hasten or retard the drying rate as affected by atmospheric temperature, humidity, and wind conditions.

Historically, the thinner used for conventional paints was turpentine, but because of newer and cheaper solvents it has largely disappeared from paint manufacturing, although it is still used to some extent for thinning paints on the job. *See* PAINT AND COATINGS; SOLVENT; SURFACE COATING; VARNISH.

C. R. Martinson; C. W. Sisler

Thiocyanate

A compound containing the —SCN group, typically a salt or ester of thiocyanic acid (HSCN). Thiocyanates are bonded through the sulfur(s) and have the structure R—S—C≡N. They are isomeric with the isothiocyanates, R—N=C≡S, which are the sulfur analogs of isocyanates (—NCO). The thiocyanates may be viewed as structural analogs of the cyanates (—OCN), where the oxygen (O) atom is replaced by a sulfur atom.

Thiocyanic acid is an unstable gas at room temperature and a yellow solid below 5°C (41°F). It is produced by the reaction of barium thiocyanates and sulfuric acid. The thiocyanates are stable to air and oxidation, and form a large number of metal complexes and organic compounds. For example, a test for the presence of ferric iron [Fe(III)] or thiocynate is the formation of the blood red complex $[Fe(SCN)_6]^{3-}$. The nucleophilic displacement of alkyl halides to produce alkyl thiocyanates is an example of the organic chemistry of the thiocyanate anion [reaction (1)]. Furthermore, like the cyanates, thio-

$$RCH_2Cl + SCN^- \longrightarrow RSCH_2SCN + Cl^- \qquad (1)$$

cyanates can be converted to thiocarbamates [reaction (2)].

$$RSCN + H_2O \longrightarrow RSCONH_2 \qquad (2)$$

The principal commercial derivatives of thiocyanic acid are ammonium and sodium thiocyanates.

Thiocyanates and isothiocyanates have been used as insecticides and herbicides. Specifically, ammonium thiocyanate is used as an intermediate in the manufacture of herbicides and as a stabilizing agent in photography. Sodium and potassium thiocyanates are used in the manufacture of textiles and the preparation of organic thiocyanates.

In living systems, thiocyanates are the product of the detoxification of cyanide ion (CN−) by the action of 3-mercaptopyruvate sulfur transferase. In addition, thiocyanates can interfere with thyroxine synthesis in the thyroid gland and are part of a class known as goitrogenic compounds. *See* CYANIDE; SULFUR; THYROXINE.

Thomas J. Meade

Thiophene

An organic heterocyclic compound containing a diunsaturated ring of four carbon atoms and one sulfur atom. *See* HETEROCYCLIC COMPOUNDS.

Thiophene (**1**), methylthiophenes, and other

(**1**)

alkylthiophenes are found in relatively small amounts in coal tar and petroleum. Thiophene accompanies benzene in the fractional distillation of coal tar. Purification of coal-tar benzene is effected by treatment with concentrated sulfuric acid, which selectively forms water-soluble thiophene-sulfonic acid. Alternatively, treatment with aluminum chloride selectively polymerizes the thiophene in the benzene to nonvolatile materials. 2.5-Dithienylthiophene (**2**)

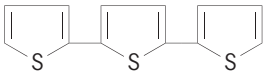

(**2**)

has been found in the marigold plant. Biotin, a water-soluble vitamin, is a tetrahydrothiophene derivative.

Properties. The parent compound (**1**) is nearly insoluble in water (forming 0.02–0.04% solutions at 20°C or 68°F), mp −38.2°C (−36.8°F), bp 84.2°C (183. 6°F), n_D^{20} 1.5287, and specific gravity (20/4) 1.0644. Thiophene has a resonance energy of 29–31 kcal/mole (121–130 kilojoules/mole), is stable to heat, and undergoes electrophilic substitutions (nitration, sulfonation, acetylation, halogenation, chloromethylation, and mercuration). Thus, thiophene is an aromatic compound. Generally, electrophilic substitutions occur with greater ease than with benzene, but less readily than with furan or pyrrole. The entering group favors the α position. Thiophenes are stable to alkali and other nucleophilic agents, and are relatively resistant to disruption by acid. *See* AROMATIC HYDROCARBON.

Most oxidative processes (nitric acid, ozone, hydrogen peroxide) involving the nucleus have not proved useful in opening the thiophene ring. Peracetic or perbenzoic acid oxidizes thiophenes such as 3,4-dimethylthiophene (**3**) to the corresponding sulfones (**4**), which behave more as butadiene derivatives than as thiophenes [reaction (1)].

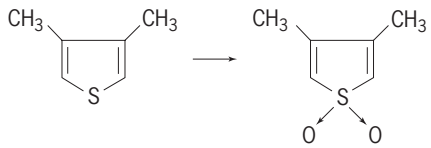

(1)

(**3**) (**4**)

Sodium in liquid ammonia and methanol converts thiophene to a mixture of dihydro and acyclic products. Raney nickel strips sulfur from thiophenes in a ring-opening reaction (2), converting (**5**) to (**6**).

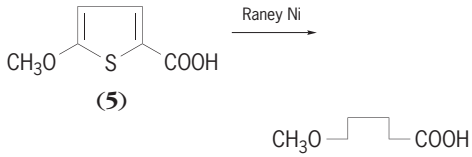

(2)

(**5**) (**6**)

Catalytic hydrogenation over molybdenum or cobalt sulfide catalysts at high temperature and pressure, as well as over platinum or palladium catalysts in massive amounts, saturates the ring.

Bromine and chlorine react readily with thiophenes, which undergo both substitution and addition reactions. Control of conditions as well as the possibility of dehydrohalogenation by alkali of the products first formed furnishes halogenated thiophenes in practical preparations. Iodination of thiophene in the presence of mercuric oxide or iodination of mercurated thiophenes gives iodinated derivatives.

Thiophene undergoes the Diels-Alder reaction with the more active dienophiles, such as acetylenedicarboxylic ester, to form benzene derivatives by extrusion of sulfur [reaction (3)].

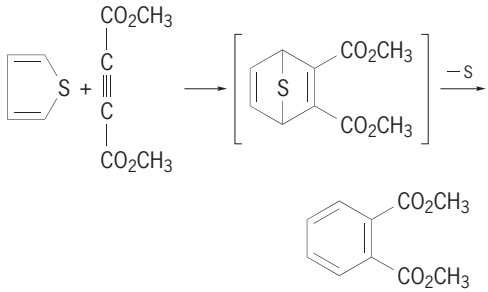

(3)

Preparation. The thiophene ring system is formed by cyclization of 1,4-dicarbonyl compounds in the presence of phosphorus sulfides (for example, 2,5-hexadione gives 2,5-dimethylthiophene; 4-oxo-3-ethylpentanoic acid gives 2-methyl-3-ethylthiophene), or by cyclization of hydrocarbons with sulfur or sulfur compounds at elevated temperatures [for example, the reaction of 2-methylbutadiene with sulfur at 320–420°C (610–790°F) gives 3-methylthiophene; the reaction of ethylbenzene with sulfur in a bimolecular process gives 2,4-diphenylthiophene].

The commercial production of thiophene (**1**) from readily available butane or butadiene awaits only a large-scale demand. A laboratory synthesis converts sodium succinate to thiophene by heating with phosphorus sulfide.

Alkylthiophenes are prepared by ring synthesis, by alkylation of thienylmagnesium halides with sulfate or sulfonate esters, or by reduction of thiophene ketones. 2-Vinylthiophene, potentially of interest as a polymerizable monomer, can be prepared by reducing 2-acetylthiophene to methyl-2-thienylcarbinol, and dehydrating.

Thiophene aldehydes are prepared by treatment of the thiophene with hexamethylenetetramine (Sommelet process), or with the *N*-methylformanilide-phosphorus oxychloride reagent pair. Friedel-Crafts acylation, often with mild catalysts, gives thiophene ketones in good yields. Thiophene carboxylic acids result from the silver oxide oxidation of thiophene aldehydes, the haloform oxidation of acetylthiophene, and the carbonation of thiophenemetal derivatives. Thiophene aldehydes, ketones, and acids show normal chemical behavior, similar to the corresponding benzene derivatives. *See* ORGANOSULFUR COMPOUND.

Walter J. Gensler; Martin Stiles

Bibliography. D. Barton and W. D. Ollis (eds.), *Comprehensive Organic Chemistry*, vol. 4, 1979; S. Gronowitz (ed.), *Chemistry of Heterocyclic Compounds: Thiophene and Its Derivatives*, vol. 44, 4 pts., 2d ed., 1992; A. R. Katritzkey (ed.), *Advances in Heterocyclic Chemistry*, vol. 1, 1963; E. Lukevics and A. E. Skorova, *Thiophene Derivatives of Group IV B Elements*, 1982.

Thiosulfate

A salt containing the negative ion $S_2O_3^{2-}$. This species is an important reducing agent and may be viewed as a structural analog of the sulfate ion (SO_4^{2-}) where one of the oxygen (O) atoms has been replaced by a sulfur (S) atom. The sulfur atoms of the thiosulfate ion are not equivalent. Thiosulfate is tetrahedral, and the central sulfur is in the formal oxidation state 6+ and the terminal sulfur is in the formal oxidation state 2−.

Commercial production of thiosulfates involves the boiling of elemental sulfur in an alkaline solution containing sulfite ions (SO_3^{2-}) [reaction (1)].

$$S + SO_3^{2-} \longrightarrow S_2O_3^{2-} \qquad (1)$$

Alternatively, sulfides may react with sulfur dioxide (SO_2), sulfite, or bisulfite (HSO_3^-) to produce thiosulfates. Thiosulfuric acid ($H_2S_2O_3$) is unstable and readily decomposes to elemental sulfur and sulfurous acid (H_2SO_3), while the sodium salt ($Na_2S_2O_3$) is stable.

Principal uses of thiosulfates include agricultural, photographic, and analytical applications. Ammonium thiosulfate [$(NH_4)_2S_2O_3$] is exploited for both the nitrogen and sulfur content, and it is combined with other nitrogen fertilizers such as urea. Thiosulfate ion is an excellent complexing agent for silver ions (bound through sulfur). The sodium salt and the ammonium salt are well known as the fixing agent "hypo" used in photography. The aqueous (aq) thiosulfate ion functions as a scavenger for unreacted solid (s) silver bromide on exposed film and therefore prevents further reaction with light [reaction (2)]. A

$$AgBr(s) + 2S_2O_3^{2-}(aq) \longrightarrow [Ag(S_2O_3)_2]^{3-}(aq) + Br^-(aq) \qquad (2)$$

well-known and important volumetric analysis is the rapid oxidation of thiosulfate by iodine to produce tetrathionate [reaction (3)].

$$2S_2O_3^{2-} + I_2 \longrightarrow S_4O_6^{2-} + 2I^- \qquad (3)$$

In nature, thiosulfate is converted into hydrogen sulfide (H_2S) via enzymatic reduction. Hydrogen sulfide, in turn, is converted into the thiol group of cysteine by the reaction with *O*-acetylserine. *See* COORDINATION COMPLEXES; OXIDATION-REDUCTION; PHOTOGRAPHIC MATERIALS; SULFUR.

Thomas J. Meade

Thirst and sodium appetite

The sensations caused by dehydration, the continuing loss of fluid through the skin and lungs and in the urine and feces while there is no water intake into the body. Thirst becomes more and more insistent as dehydration worsens. Water and electrolytes are needed to replace losses, and an adequate intake of sodium as well as water is important for maintaining blood volume. Herbivores and human vegetarians, whose diets lack sodium, have a natural appetite for sodium; however, severe sodium deficiency in carnivorous animals and humans can result in the development of a well-marked sodium appetite as well. Water intake varies considerably between individuals and depends on climate, custom, and diet. Reproduction affects drinking behavior; fluid intake increases during pregnancy and especially during lactation. Normally, the amounts of water drunk and taken in food are more than enough to maintain hydration of the body, and the usual mixed diet provides all the electrolytes required.

Drinking behavior. The passage of water through the mouth satisfies a basic urge to drink whether or not the body needs water. An indication of the importance of this urge is the fact that rats infused with water at rates far in excess of their requirements, by routes that bypass the mouth and pharynx, continue to drink. Responding to this urge when the body is in a healthy state is entirely beneficial because the water provides for body fluid homeostasis by the kidney and any surplus is excreted (although in kidney disease, normal drinking behavior may result in an inappropriately large and possibly harmful intake of fluid). Even more demanding than the basic urge to drink are the sensations of thirst and sodium appetite aroused by fluid deficits in the body, which lead to primary or regulatory drinking. The urgency of these

sensations, especially of thirst, indicates how vital drinking is for survival.

Classification. The three categories of drinking behavior are as follows:

Primary, regulatory, or deficit-induced drinking. A deficit of fluid in one or both of the major fluid compartments of the body is the signal to increase drinking.

1. Cellular dehydration, detected by osmoreceptors, causes thirst and vasopressin release.
2. Hypovolemia, detected by volume receptors in the heart and large veins and the arterial baroreceptors, causes immediate thirst, a delayed increase in sodium appetite, activation of the renin-angiotensin system, and increased mineralocorticoid and vasopressin secretion.

Secondary or nonregulatory drinking. Drinking occurs in the absence of a fluid deficit. Drinking associated with feeding, for instance, is important in the day-to-day control of body water. It ensures a regular supply of water so that renal regulation of body fluids can take place. Secondary drinking is distinguished by the following factors:

1. Mealtime or food-associated.
2. Fail-safe, normal.
3. Oropharyngeal cues, dry-mouth.
4. Schedule-induced.
5. Palatability-induced NaCl intake.

Clinical disturbances in drinking behavior. Increases or decreases in amounts drunk in disease may result from normal or abnormal functioning of mechanisms of thirst or sodium appetite:

1. Symptomatic: mechanisms of thirst and sodium appetite respond normally to excessive fluid loss caused by disease.
2. Pathologic: thirst mechanisms are activated excessively (hyperdipsia, polydipsia) or inadequately (hypodipsia) because of abnormalities of the mechanisms.

Physiology and ontogeny. Recent research on the physiology of drinking behavior has yielded extensive information on the mechanisms, neural substrates, and neuropharmacology of the different types of drinking behavior. Many mechanisms, some involving hormones, come into play to cause drinking, depending on the particular circumstances. The role of angiotensin II as a hormone and a central neurotransmitter or neuromodulator of thirst and sodium appetite is becoming clearer. The importance of the hormones of sodium deficiency, of pregnancy and lactation, and of the stress hormones of the hypothalamo-pituitary-adrenocortical axis in sodium appetite has been established. Information on the ontogeny of drinking behavior is now available. In much of this work, the newer specific receptor blockers—and more recently still, antisense oligonucleotides and gene-knockout animal models—have added new insights to theories based on classical behavioral and physiologic methods.

Oropharyngeal versus systemic factors. Theories of thirst in which dryness of the mouth and throat plays a central role are quite old. However, during the nineteenth century, Dupuytren, Latta, Magendie, and Bernard, among others, showed that thirst could be relieved by intravenous fluids and that the passage of water through the oropharynx was neither necessary nor sufficient for relief. Thirst was considered to be a sensation of general origin caused by lack of water in the body. Nevertheless, the dry-mouth theory received strong support from Cannon in the early twentieth century, and dryness of the mouth and throat (such as might be caused by lecturing, singing, mouth breathing, a spicy meal, or apprehension) certainly can cause increased drinking in the absence of any systemic need for water. However, since the early 1950s when B. Andersson and colleagues established that there is a thirst center in the hypothalamus, and as the significance of cellular dehydration and hypovolemia (low blood volume) as causes of thirst and sodium appetite began to be better understood, the dry-mouth theory has become less influential. It is now recognized that there are many other causes of increased drinking behavior, including cases in which there is an obvious systemic need for water as well as cases in which there does not appear to be an immediate need.

Cellular dehydration. H. Wettendorff first proposed the cellular dehydration theory of thirst at the beginning of the twentieth century. In 1937, A. Gilman showed that dogs drank more water after intravenous administration of hyperosmotic saline (a substance which dehydrates cells) than after the same osmotic load of hyperosmotic urea (a substance which penetrates cells), providing crucial supporting evidence for Wettendorff's theory. Observations of other species using a variety of osmotic challenges have extended these findings and have established that hyperosmotic solutions of solutes that are excluded from cells cause more drinking than equiosmolar amounts of solutes that penetrate cells. Thus, the osmotic shift of water out of the cells caused by the excluded solutes provides the critical stimulus to drinking. Continuing water loss in the absence of intake is perhaps a more significant cause of cellular dehydration than administration of an osmotic load, but the same mechanisms apply. *See* OSMOREGULATORY MECHANISMS; OSMOSIS.

Sharing in the overall cellular dehydration, whether caused by osmotic loads or water deprivation, are osmoreceptors which initiate the responses of thirst and renal conservation of water. There has also been considerable discussion about whether or not a sodium-sensitive system in the walls of the cerebral ventricles might account for some of the drinking induced by cellular dehydration, especially in herbivores. Osmoreceptors are mainly located in the hypothalamus. Injection of minute quantities of hyperosmotic solutions of dehydrating substances in the lateral preoptic area and adjacent zona incerta causes water-replete animals to drink, whereas injection of water, or bilateral destruction of the region, prevents the animal with generalized cellular dehydration from drinking. The nervous tissue in the hypothalamus surrounding the anterior third cerebral ventricle and, in particular, the vascular organ of the lamina terminalis also respond to osmotic

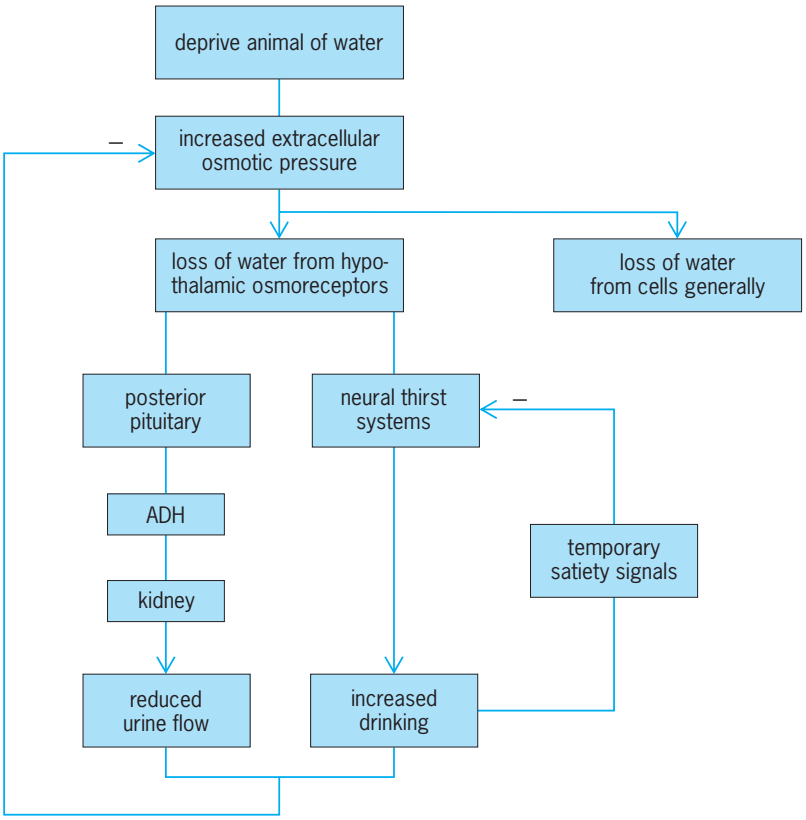

Fig. 1. Regulation of cellular water. ADH = antidiuretic hormone.

stimuli. Osmoreceptors initiating thirst work in conjunction with osmoreceptors initiating antidiuretic hormone (ADH) release to restore the cellular water to its prehydration level (**Fig. 1**). In addition to reducing urine loss, ADH may lower the threshold to the onset of drinking in response to cellular dehydration and other thirst stimuli. The cellular dehydration system is very sensitive, responding to changes in effective osmolality of 1–2%.

Hypovolemia. The cells of the body are bathed by sodium-rich extracellular fluid that corresponds to the aquatic environment of the unicellular organism. Thirst has been known since very early times to accompany hemorrhage, severe diarrhea (such as in cholera), sodium deficiency, and other diseases in which the brunt of fluid loss is borne by the extracellular fluid. Extracellular fluid consists of two main components of similar composition, blood plasma and interstitial fluid, which is the fluid outside the vasculature and in immediate contact with the cells. Loss of sodium is inevitably accompanied by loss of water, resulting in hypovolemia with thirst followed by a delayed increase in sodium appetite. If not corrected, continuing severe sodium loss eventually leads to circulatory collapse.

Among experimental procedures that have been used to cause hypovolemic drinking are (1) removing blood; (2) causing sodium deficiency by dieting, sweating, peritoneal dialysis, and other means; (3) causing extracellular fluid to accumulate outside the vasculature in the peritoneal cavity or subcutaneously by the technique of hyperoncotic dialysis; and (4) mimicking the effects of severe dehydration on the circulation by interfering with the flow of blood back to the heart by obstructing the abdominal inferior vena cava. All of these experimental methods cause thirst and a delayed increase in sodium appetite and either overhydrate the cellular compartment or do not alter its water content. In 1936, R. A. McCance described how three human subjects made sodium-deficient by dieting and sweating experienced a complex sensation of thirst and craving for salt. Initially, the developing sodium deficit and accompanying loss of water led to progressive hypovolemia with little change in osmotic pressure. Later, volume was preserved at the expense of osmotic pressure, which fell. Changes in drinking behavior were caused by hypovolemia, not by cellular dehydration, because the cells are usually overhydrated in sodium deficiency.

Stretch receptors in the walls of blood vessels entering and leaving the heart and in the heart itself are thought to initiate hypovolemic drinking. Volume receptors in the venoatrial junctions and receptors that register atrial and ventricular pressure respond to the underfilling of the circulation with a reduction in inhibitory nerve impulses to the thirst centers, which results in increased drinking. Angiotensin II and other hormones (such as aldosterone and ADH) are also involved in this response. Arterial baroreceptors function in much the same way as the volume receptors on the low-pressure side of the circulation, exerting continuous inhibitory tone on thirst neurons. A fall in blood pressure causes increased drinking, whereas an acute rise in blood pressure inhibits drinking. (It should be noted that the sustained rise in arterial blood pressure of established hypertension does not have any permanent inhibitory effect on drinking, presumably because the baroreceptors are reset to the higher arterial pressure.) The anterior third cerebral ventricle region, which is implicated in angiotensin-induced drinking, plays a crucial role in hypovolemic drinking, body fluid homeostasis, and blood pressure control. Some of the pathways involved are illustrated in **Fig. 2**.

Renin-angiotensin systems and drinking. It is believed that drinking caused by hypovolemic stimuli partly depends on the kidneys because research shows that (1) the amounts of water drunk after such stimuli are reduced by prior nephrectomy but not by ligation of both ureters and (2) extracts of kidney injected into water-replete animals cause drinking. The renal thirst factor is the proteolytic enzyme renin, which is secreted into the circulation by the juxtaglomerular cells of the kidney in response to hypovolemia. Renin cleaves an inactive decapeptide, angiotensin I, from angiotensinogen, an α_2-globulin that is synthesized in the liver and released into the circulation. Angiotensin I is converted to the physiologically active but short-lived octapeptide angiotensin II during the passage of blood through the lungs. Angiotensin II is an exceptionally powerful stimulus of drinking behavior in many

mammals, birds, reptiles, and bony fish when administered systemically or into the brain. Increased activation of the renin-angiotensin system may sometimes account for pathologically increased thirst in humans. Injection of angiotensin II into sensitive limbic structures in the anterior hypothalamus and the anterior third cerebral ventricle region of the brain of a rat causes an almost immediate increase in water intake, often followed by a slower increase in sodium intake. Angiotensin II also produces (1) a rise in arterial blood pressure by causing vasoconstriction and hypertrophy of vascular smooth muscle, release of norepinephrine from sympathetic nerve endings, and secretion of adrenomedullary hormones; and (2) water and sodium retention by causing release of ADH from the posterior pituitary and stimulation of renal tubulular transport of sodium through direct action on the kidney and indirectly through increased aldosterone secretion from the adrenal cortex. (Other less well-defined stimulating actions of angiotensin are on cell growth, membrane function, protein synthesis, prostaglandin release, learning, and memory.) *See* ALDOSTERONE; KIDNEY.

Renin, synthetic renin substrate, and angiotensin I are also effective stimuli to increase drinking, their action being mediated through local generation of angiotensin II. Angiotensin peptides that do not rely on renal renin for their production are produced in many tissues, including nervous tissue. The anterior third cerebral ventricle region, the vascular organ of the lamina terminalis, the median preoptic nucleus, and subfornical organ—which are particularly important in body fluid homeostasis and blood pressure control—are also well provided with angiotensinergic nerve terminals and receptors. The median preoptic nucleus and some other angiotensin-sensitive tissues lie inside the blood-brain barrier and presumably are not accessible to circulating angiotensin but are accessible to peptide released from nerve endings. There are also regions in the brain where production of angiotensin peptides takes place at a distance from receptors. This suggests that angiotensin produced in the brain may act both as a paracrine agent producing volume effects some distance from the point of release (because it diffuses into the extracellular fluid) and as a conventional point-to-point neurotransmitter. The role of cerebral renin and its relation to the more abundant and better-known renal renin remain uncertain.

Genetic aspects of renin-angiotensin systems are beginning to be investigated, and the genes encoding the proteins and receptors are being identified. Use of antagonists of the various stages of the renin-angiotensin cascade and exploitation of the more recently introduced selective antagonist subtypes and antisense oligonucleotides (a deoxyribonucleic acid or ribonucleic acid sequence with two or more covalently linked nucleotides) to block angiotensin synthesis or receptors have been invaluable in helping to understand the physiology of angiotensininduced drinking. At least three angiotensin receptor subtypes have been identified—AT_1, AT_2, and AT_4—

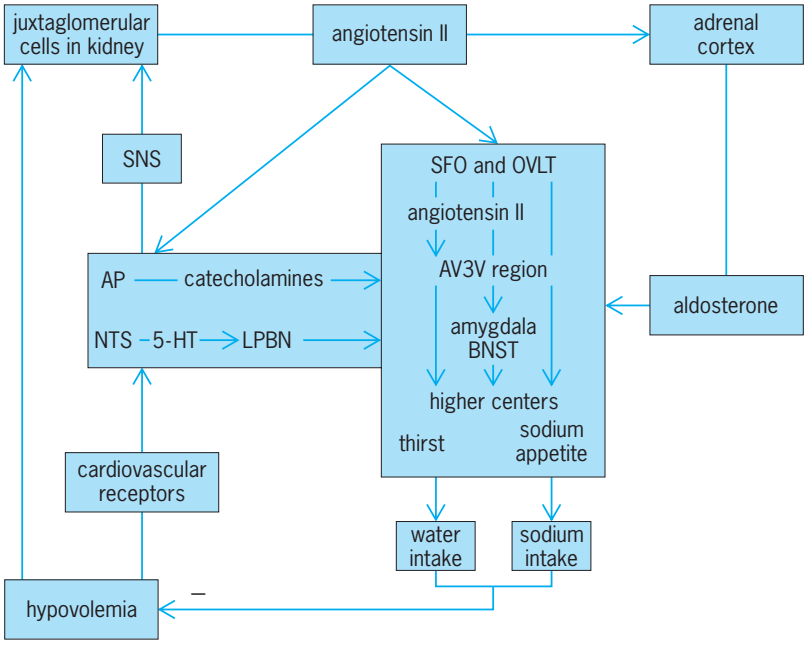

Fig. 2. Hypovolemic thirst and sodium appetite are caused by altered sensory information from an underfilled circulation, reinforced by increases in circulating angiotensin II and mineralocorticoids. Angiotensin peptides generated from components of a cerebral renin-angiotensin system are also involved. (The central nervous system pathways are based on A. K. Johnson and R. L. Thunhorst, 1997.) AP, area postrema; AV3V, anteroventral third ventricle; BNST, bed nucleus of the stria terminalis; 5-HT, 5-hydroxytryptamine or serotonin; LPBN, lateral parabrachial nucleus; OVLT, vascular organ of the lamina terminalis; SFO, subfornical organ; SNS, sympathetic nervous system.

but most known functions of angiotensin, including angiotensin-induced drinking, are associated with AT_1 receptors. However, the AT_2 receptors are the predominant receptors in the fetal brain. Although their function is uncertain, AT_1 and AT_2 receptors may have opposing effects on apoptosis (cell death), vasoconstriction, myoendothelial proliferation, and possibly on drinking behavior. Mutant mice lacking the gene encoding the AT_2 receptor show impaired drinking when water-deprived and an increased pressor response to intracarotid infusion of angiotensin II, indicating that the receptor could be involved in these functions. The potential of approaches such as the use of antisense oligonucleotides and gene-knockout animal models is considerable, though few results are yet available. *See* GENE; OLIGONUCLEOTIDE.

An integrated response to hypovolemia involves both neural and hormonal mechanisms (Fig. 2). The reduced sensory discharge from the cardiovascular stretch receptors resulting from underfilling of the circulation directly activates hypothalamic and limbic drinking systems. The same sensory signals lead to reflex release of renin through sympathetic nerves to the kidney. The resulting increases in circulating angiotensin II act on the subfornical organ, vascular organ of the lamina terminalis, and area postrema, which lie outside the blood-brain barrier, and contribute to the increases in thirst and sodium appetite and hemodynamic responses to hypovolemia by sensitizing hypothalamic and limbic drinking systems to the altered sensory information

from cardiovascular receptors. Angiotensin peptides generated inside the blood-brain barrier may also be involved, but their functional relation to the effects of increases in circulating angiotensin II is unknown. Increases in aldosterone secretion also contribute to the increases in sodium appetite. The renal renin-angiotensin system may have a more important emergency role in the arousal of thirst and sodium appetite in circulatory collapse (for example, in severe hemorrhage or adrenal insufficiency) than in the more modest day-to-day variations in extracellular fluid volume. *See* BRAIN; SYMPATHETIC NERVOUS SYSTEM.

Neuropharmacology of drinking. Many substances released by neurons, and in some cases by neuroglial cells, affect drinking behavior when injected into the brain and may interact with the brain and modify angiotensin-induced drinking. They can act as neurotransmitters and produce effects localized to the postsynaptic membrane or presynaptic endings close to where they are released, they may diffuse into the extracellular fluid and exert paracrine or volume effects on nervous structures some distance from their source, or they may be released into the bloodstream and function as hormones. The time course of action varies, and there may be long-term trophic actions. Substances may stimulate or inhibit drinking, or both, depending on the species and the conditions of the experiment. Acetylcholine is a particularly powerful stimulus to drink in rats, and no inhibitory effects on drinking have been described. Histamine also seems to be mainly stimulatory. However, a lengthening list of neuroactive substances, including norepinephrine, serotonin, nitric oxide, opioids, bombesin-like peptides, tachykinins, and neuropeptide Y, may either stimulate or inhibit drinking with varying degrees of effectiveness, depending on the species or the site of injection in the brain. Natriuretic peptides, prostaglandins, and gamma-amino butyric acid seem to be exclusively inhibitory. *See* ACETYLCHOLINE; NEUROBIOLOGY; SYNAPTIC TRANSMISSION.

Many hormones also affect water or sodium intake. Relaxin stimulates water intake, and ADH (or vasopressin) lowers the threshold to thirst in some species. Vasopressin injected into the third cerebral ventricle may stimulate water intake, suggesting a possible role for vasopressinergic neurons. Increased sodium appetite in pregnancy and lactation depends partly on the conjoint action of progesterone, estrogen, adrenocorticotrophic hormone (ACTH), cortisol, corticosterone, prolactin, and oxytocin. Aldosterone and other mineralocorticoids, the stress hormones of the hypothalamo-pituitary-adrenocortical axis, corticotrophin, ACTH, and the glucocorticoids also stimulate sodium intake. *See* ENDOCRINE MECHANISMS; NEUROHYPOPHYSIS HORMONE.

The effect of many of these substances on drinking behavior shows both species and anatomical diversity. Serotonin stimulates drinking in pigeons, but inhibits drinking in rats when injected into the lateral parabrachial nucleus on the pathway in the hindbrain that responds to overfilling of the circulation (Fig. 2). Tachykinins stimulate drinking in birds but inhibit thirst and sodium appetite in rats. Oxytocin released from the posterior pituitary into the bloodstream contributes to increased sodium appetite caused by the conjoint action of reproductive hormones, but as a neurotransmitter or paracrine agent released from oxytocinergic neurons in the paraventricular nucleus, it may inhibit sodium appetite in circumstances in which water is more urgently and immediately need. The multiplicity of effects of many of these substances makes it impossible to generalize on their role in natural thirst, but none of these substances seems to be as consistent and as universal a stimulus of increased thirst and sodium appetite as angiotensin.

Overview. Mechanisms that ensure a continuing intake of water and sodium are vital. There is a powerful urge to drink independent of need so that in temperate climates and under stable conditions of activity and diet, the body's need for water is fully met. In times of good health, such intake is beneficial, but it may result in overhydration during illness. The role of sodium is less clear. In most western societies, sodium intake in the diet exceeds need, which may be a factor in the increased incidence of hypertension. On the other hand, an adequate intake of sodium is essential, and palatability and increased sodium appetite in sodium deficiency ensure this. When the body lacks water or sodium, it is imperative that physiologic responses restore the normal contents. Renal conservation can slow the rate of fluid loss, but the emergency mechanisms of thirst and sodium appetite in response to deficits are far more important for survival. James T. Fitzsimons

Bibliography. D. A. Denton, *The Hunger for Salt*, Springer-Verlag, Berlin, 1982; J. T. Fitzsimons, Physiology and pathophysiology of thirst and sodium appetite, in D. W. Seldin and G. Giebisch (eds.), *The Kidney*: *Physiology and Pathophysiology*, 3d ed., Lippincott, Williams and Wilkins, Philadelphia, 2000; A. K. Johnson and R. L. Thunhorst, The neuroendocrinology of thirst and salt appetite: Visceral sensory signals and mechanisms of central integration, *Frontiers Neuroendocrinol.*, 18:292–353, 1997; E. M. Stricker, *Handbook of Behavioral Neurobiology*, vol. 10: *Neurobiology of Food and Fluid Intake*, Plenum Press, New York, 1990.

Thomson effect

A phenomenon discovered in 1854 by William Thomson (Lord Kelvin). He found that there occurs a reversible transverse heat flow into or out of a conductor of a particular metal, the direction depending upon whether a longitudinal electric current flows from colder to warmer metal or from warmer to colder. Any temperature gradient previously existing in the conductor is thus modified if a current is

turned on. The Thomson effect does not occur in a current-carrying conductor which is initially at uniform temperature.

From these observations it may be shown that for copper there is a heat output where positive charge flows down a temperature gradient and a heat input where positive charge flows up a temperature gradient; whereas for iron the reverse is true. All metals may be divided into two classes with respect to the direction of the Thomson effect. These flows of heat require that a distributed seat of electromotive force act at all points in the conductor. The total Thomson emf along the length of a conductor is given by

$$\int_{T_1}^{T2} \sigma \, dT$$

where σ is the Thomson coefficient for the metal in question, and T_1 and T_2 are the temperatures at the two ends of the conductor. With the discovery of the Thomson effect a complete thermodynamical theory of thermoelectricity became possible. *See* THERMOELECTRICITY. John W. Stewart

Thoracica

The major order of the crustacean subclass Cirripedia. The adult animals are permanently attached. The mantle is usually reinforced by calcareous plates (**Fig. 1**). Six pairs of biramous cirri are present, and the abdomen is absent or represented by caudal appendages. Antennules are present in the adult, and cement glands are strongly developed. Most species are hermaphroditic.

Thoracica are subdivided into three suborders: Lepadomorpha, stalked or goose barnacles; Balanomorpha, the common acorn barnacles; and Verrucomorpha, a rare group of asymmetric barnacles. The stalked barnacles are attached by a peduncle, and the body is enclosed in a bivalved fold or mantle, the capitulum, which is typically strengthened by calcareous plates. In acorn barnacles the body is enclosed in a strong shell of four, six, or eight plates rigidly united or fused together, and the mantle opening is protected by a pair of movable opercular valves, each formed of two plates, the tergum and the scutum (Fig. 1*c*, *d*).

Barnacles feed by sweeping the fan of cirri (**Fig. 2**) through the water and straining out minute organisms. The mouth is furnished with a large upper lip or labrum and three pairs of mouth appendages: mandibles, maxillulae, and maxillae. The alimentary canal is divided into a fore-, mid-, and hindgut, with the midgut having the associated digestive diverticula. The anus is terminal. Excretory organs open on the maxillae. The nervous system is typically crustacean, though it is shortened.

The paired ovaries lie in the stalk or in the mantle in the stalkless acorn barnacles, with the oviduct opening at the base of cirrus I. The testes lie in the thorax, and the paired seminal vesicles run backward to the base of the penis. Cross-fertilization is usual, with the sperm being deposited within the mantle by the elongated penis of an adjacent barnacle. Self-fertilization can occur. The eggs are laid as two flat coherent masses into the mantle space, where they hatch into nauplius larvae. The nauplii are planktonic and after passing through five further similar floating naupliar stages, they change into the very different cypris, which ceases to be pelagic and seeks a suitable substrate for attachment and metamorphosis into the adult. Growth is rapid, especially in warm waters, and the young barnacle may be sexually mature in 3 weeks, though occasionally fertilization may not occur until 1 or even 2 years after settlement. Barnacles may live

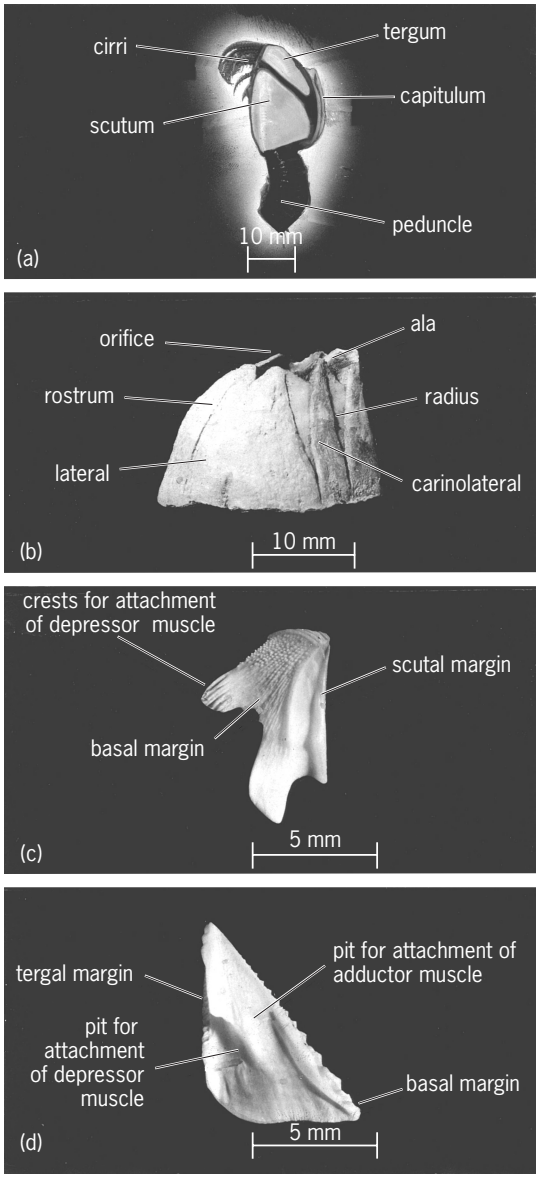

Fig. 1. Morphology of representative Thoracica. (*a*) *Lepas anatifera* (*from D. P. Henry, The Cirripedia of Puget Sound with a key to the species, Univ. Wash. Publ. Oceanogr., 4(1):1–48, 1940*). (*b*) *Balanus eburneus*, lateral view of shell, (*c*) inner view of tergum, and (*d*) inner view of scutum (*from D. P. Henry, American Waters, Friday Harbor Symposium in Marine Biology, University of Washington Press, 1959*).

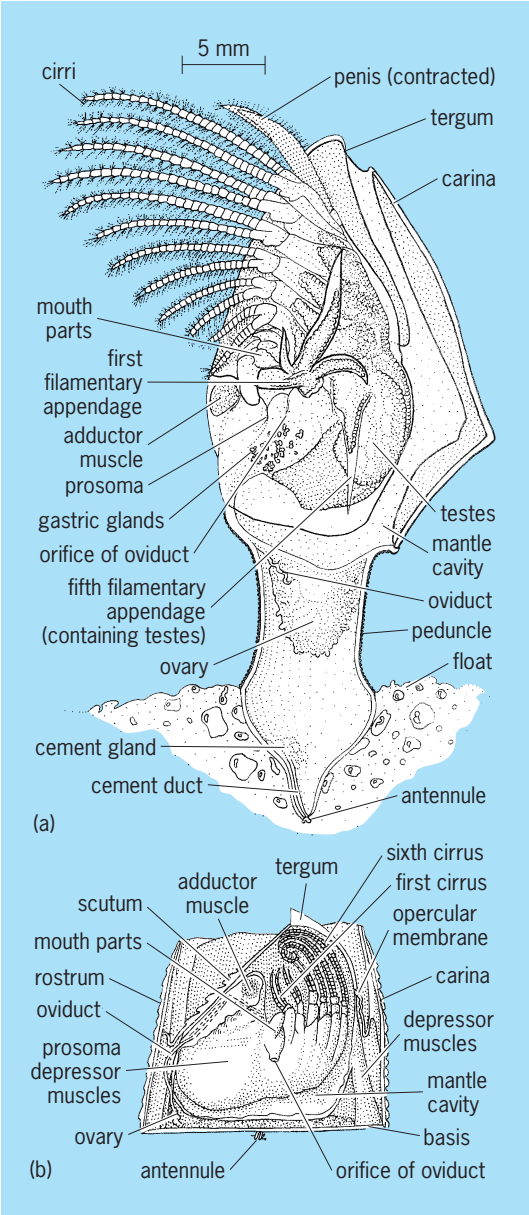

Fig. 2. Thoracica, internal anatomy. (*a*) *Lepas fascicularis*, with right side of capitulum, peduncle, and float removed. (*b*) *Balanus*, with right side of wall removed. (*After R. W. Hegner, Invertebrate Zoology, Macmillan, 1933*)

only months or occasionally as long as 10 years in some slow-growing and deep-water species. *See* CIRRIPEDIA; LEPADOMORPHA.

H. G. Stubbings

Thorite

A mineral, thorium silicate, in which the element thorium was discovered in 1828. Thorite is tetragonal in crystallization and has a crystal structure identical with that of the nesosilicate zircon, $ZrSiO_4$. The idealized chemical formula of thorite is $ThSiO_4$. All natural material departs widely from this composition owing to the partial substitution of uranium, rare earths, calcium, and iron for thorium. Structurally, thorite usually has completely lost its crystallinity because of radiation damage from the contained uranium and thorium (metamict state). The specific gravity ranges between about 4.3 and 5.4. The hardness on Mohs scale is about $4^1/_2$. The color commonly is brownish yellow to brownish black and black. Thorogummite is a chemical variant of thorite with the same crystal structure and very similar properties. It is deficient in silica and contains small amounts of OH in substitution for oxygen.

Thorite occurs chiefly in pegmatites. It also occurs as an accessory mineral in black sands and other detrital deposits derived from granitic or gneissic terrains. Vein deposits containing thorite and thorogummite associated with barite and fluorite occur in the Wet Mountains, Custer and Fremont counties, Colorado. Similar deposits occur in the Lemhi Pass district in Idaho and Montana. A vein deposit of monazite containing thorium is mined at Steenkampskraal near Van Rhynsdorp, Cape Province, South Africa. *See* METAMICT STATE; RADIOACTIVE MINERALS; SILICATE MINERALS; THORIUM.

Clifford Frondel

Bibliography. C. Frondel, *Systematic Mineralogy of Uranium and Thorium*, USGS Bull. 1064, 1958; J. W. Frondel, M. Fleischer, and R. S. Jones, *Glossary of Uranium and Thorium-Bearing Minerals*, USGS Bull. 1250, 1967; E. W. Heinrich, *Mineralogy and Geology of Radioactive Raw Materials*, 1958.

Thorium

A chemical element, Th, atomic number 90. Thorium is a member of the actinide series of elements. It is radioactive with a half-life of about 1.4×10^{10} years. *See* PERIODIC TABLE.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Thorium oxide compounds are used in the production of incandescent gas mantles. Thorium oxide has also been incorporated in tungsten metal, which is used for electric light filaments. It is employed in catalysts for the promotion of certain organic chemical reactions and has special uses as a high-temperature ceramic material. The metal or its oxide is employed in some electronic tubes, photocells,

and special welding electrodes. Thorium has important applications as an alloying agent in some structural metals. Perhaps the major use for thorium metal, outside the nuclear field, is in magnesium technology. Thorium can be converted in a nuclear reactor to uranium-233, an atomic fuel. The energy available from the world's supply of thorium has been estimated as greater than the energy available from all of the world's uranium, coal, and oil combined.

Monazite, the most common and commercially most important thorium-bearing mineral, is widely distributed in nature. Monazite is chiefly obtained as a sand, which is separated from other sands by physical or mechanical means. *See* MONAZITE.

Thorium has an atomic weight of 232. The temperature at which pure thorium melts is not known with certainty; it is thought to be about 1750°C (3182°F). Good-quality thorium metal is relatively soft and ductile. It can be shaped readily by any of the ordinary metal-forming operations. The massive metal is silvery in color, but it tarnishes on long exposure to the atmosphere; finely divided thorium has a tendency to be pyrophoric in air.

All of the nonmetallic elements, except the rare gases, form binary compounds with thorium. With minor exceptions, thorium exhibits a valence of 4+ in all of its salts. Chemically, it has some resemblance to zirconium and hafnium. The most common soluble compound of thorium is the nitrate which, as generally prepared, appears to have the formula $Th(NO_3)_4 \cdot 4H_2O$. The common oxide of thorium is ThO_2, thoria. Thorium combines with halogens to form a variety of salts. Thorium sulfate can be obtained in the anhydrous form or as a number of hydrates. Thorium carbonates, phosphates, iodates, chlorates, chromates, molybdates, and other inorganic salts of thorium are well known. Thorium also forms salts with many organic acids, of which the water-insoluble oxalate, $Th(C_2O_4)_2 \cdot 6H_2O$, is important in preparing pure compounds of thorium. *See* ACTINIDE ELEMENTS; RADIOACTIVITY.

Harley A. Wilhelm

Bibliography. P. W. Atkins et al., *Inorganic Chemistry*, 4th ed., 2006; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; S. Cotton, *Lanthanide and Actinide Chemistry*, 2d ed., 2006.

Throat

The region that includes the pharynx, the larynx, and related structures. Both the nasal passages and the oral cavity open into the pharynx, which also contains the openings of the Eustachian tubes from the ears (see **illus.**). The lower (inferior) portion of the pharynx leads into the esophagus and the trachea or windpipe. The rather funnel-shaped pharynx is suspended from the base of the skull and the jaws; it is surrounded by three constrictor muscles that function primarily in swallowing. *See* EAR (VERTEBRATE); PHARYNX.

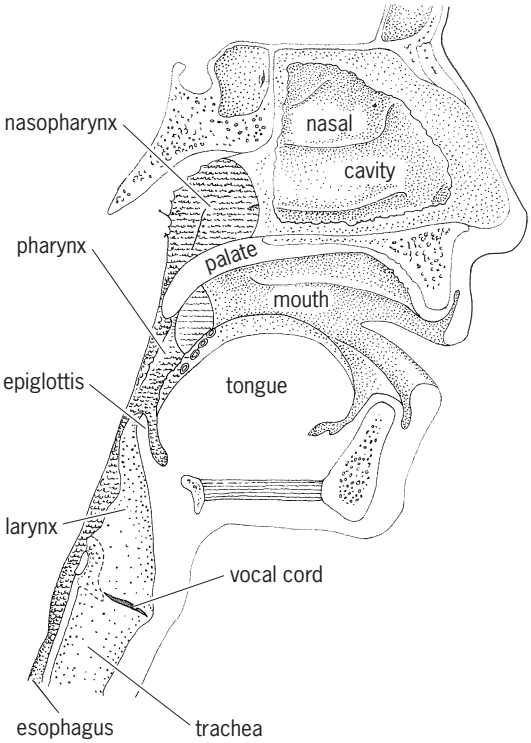

Sagittal section of a human head showing the structure of the pharynx and associated structures. (*After J. C. B. Grant, A Method of Anatomy, Williams and Wilkins, 1952*)

The larynx, or voice box, is marked externally by the shield-shaped thyroid cartilage which forms the Adam's apple. The larynx contains the vocal cords that act as sphincters for air regulation and permit phonation. The lower end of the larynx is continuous with the trachea, a tube composed of cartilaginous rings and supporting tissues. *See* LARYNX.

The term throat is also used in a general sense to denote the front (ventral side) of the neck.

Thomas S. Parsons

Thrombosis

The process of forming a thrombus, which is a solid mass or plug in the living heart or vessels composed of the constituents of the blood. Thrombosis usually occurs in a diseased blood vessel, as a result of arteriosclerosis. The consequences of thrombosis include local obstruction causing both tissue death and hemorrhage. Thrombosis is a significant factor in the death of an individual affected by arteriosclerotic cardiovascular disease, malignancy, and infection. *See* HEMORRHAGE; INFARCTION.

Thrombus formation. Thrombosis is usually initiated by vascular damage and consequent platelet adhesion and clumping. The vascular endothelium plays an important role in protecting against thrombus formation. Normal endothelium releases nitric oxide, a vasodilator. Prostacyclin, a platelet aggregation inhibitor, releases heparinlike anticoagulants and protein inhibitors such as thrombomodulin.

Fig. 1. Organized thrombus in an artery with recanalization.

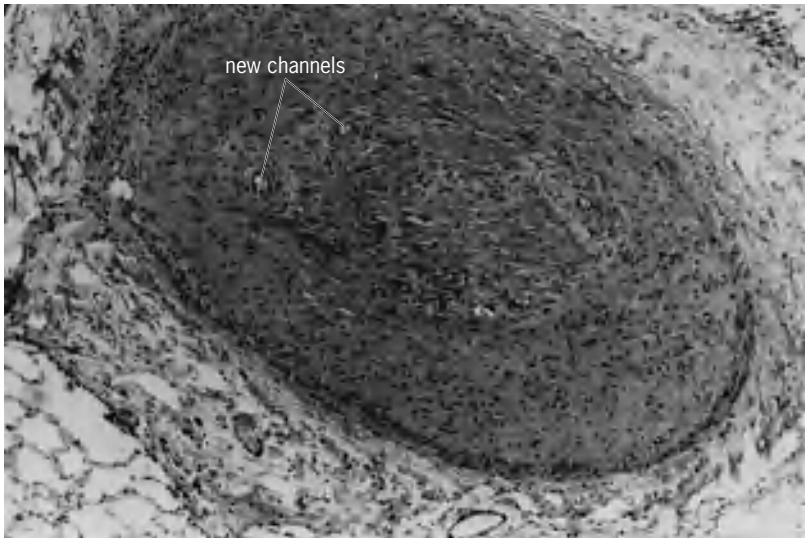

Fig. 2. Organized thrombus in an artery with scar formation; the artery is completely occluded by the fibrotic mass.

Injury to tissues, adequate concentrations of the coagulation elements, and stasis of the blood may all play a role in furthering the process of thrombus formation. The process is initiated when platelets specifically adhere to the subendothelial collagen at the points of damage to the endothelium. At the same time that the platelets begin to aggregate and release products that will further promote thrombus formation, the protein factors of the blood, which help to form the insoluble meshwork of the thrombus, become activated. This latter process is known as blood coagulation. The proteins of the coagulation system, through a series of cascading reactions (the intrinsic and extrinsic pathways), eventually reach a final common pathway to form fibrin, the insoluble protein that forms the scaffolding of the thrombus. As blood flows by the thrombus, more platelets and fibrin are deposited. Red blood cells and white blood cells become entrapped in the thrombus and are integrated into its structure. The thrombus thus consists of alternating zones of platelets and columns of fibrin with irregular layers of red cells and leukocytes.

Location of thrombi. Thrombi may form in arteries, in veins, on heart valves, or on the surfaces of the chambers of the heart. In arteries, which are vessels with relatively rapid flow of blood, the thrombus is predominantly composed of fibrin and platelets, a so-called white thrombus. In veins, vessels with less rapid flow, the thrombus is made up of greater portions of red cells, a so-called red thrombus. In capillaries, a thrombus usually consists of only platelets and fibrin.

Fate of thrombi. Once a thrombus forms, it can have one of four fates. (1) It may be digested, destroyed, and removed by proteolytic enzymes of the plasminogen–plasmin system. (2) If factors favor continued thrombus formation, the thrombus may continue to propagate itself and eventually occlude the vessel. This propagation usually takes place downstream to the site of original thrombus. (3) It may give rise to an embolus. Fresh portions of the thrombus at its outer surface tend to be friable and may break off, giving rise to thromboemboli. These emboli may cause tissue damage at sites distant from the origin of the thrombus. For example, lung infarction may develop from venous emboli, usually arising in the legs, and stroke may occur from thrombi that form in the heart and travel to the brain. (4) Thrombi may undergo a process known as organization. In this process, fibroblasts and capillaries proliferate and grow into the thrombus at its attachment to the vessel wall. The mass may be converted to a mass of vascularized connective tissue; the new channels may reestablish blood flow through the vessel (**Fig. 1**) to some degree. Organization helps stabilize the thrombus, and it may result in incorporation of a contracted fibrous mass into the vessel wall (**Fig. 2**). *See* EMBOLISM.

Prevention and treatment. Maintaining good blood flow (especially in the veins) helps prevent thrombosis. Exercise, support hose, and proper leg elevation when sitting will decrease the tendency of venous stasis in the legs. Treating hypertension and hypercholesterolemia retards atherosclerosis, which is a major cause of arterial thrombosis. Agents that interfere with platelet function, such as aspirin and fish oils, may help avoid thrombotic episodes. Anticoagulants prevent the formation of fibrin and may also be used to prevent thrombosis, especially in the veins or on prosthetic heart valves. If treatment can be given in the early stages of thrombosis, fibrinolytic therapy, utilizing agents that will help form plasmin, can minimize the tissue damage caused by thrombosis. *See* ARTERIOSCLEROSIS; CIRCULATION DISORDERS; PHLEBITIS.

Irwin Nash; Romeo A. Vidone

Bibliography. R. Colman et al., *Hemostasis and Thrombosis*, 4th ed., 2000; J. M. Kissane, *Anderson's Pathology*, 9th ed., 1989; J. L. Robbins et al., *Basic Pathology*, 6th ed., 2000; W. Williams et al., *Hematology*, 5th ed., 1996.

Thrust

The force that propels an aerospace vehicle or marine craft. Thrust is a vector quantity. Its magnitude is usually given in newtons (N) in International System (SI) units or pounds-force (lbf) in U.S. Customary Units. A newton is defined as 1 kilogram mass times an acceleration of 1 meter per second squared. One newton equals approximately 0.2248 lbf. *See* FORCE; UNITS OF MEASUREMENT.

The thrust power of a vehicle is the thrust times the velocity of the vehicle. It is expressed in joules (J) per second or watts (W) in SI units. In U.S. Customary Units thrust power is expressed in foot-pounds per second, which can be converted to horsepower by dividing by 550. *See* TURBINE ENGINE SUBSYSTEMS; POWER; RECIPROCATING AIRCRAFT ENGINE; ROCKET; TURBOJET. J. Preston Layton

Thulium

A chemical element, Tm, atomic number 69, atomic weight 168.934. It is a rare metallic element belonging to the rare-earth group. The stable isotope ^{169}Tm makes up 100% of the naturally occurring element. *See* PERIODIC TABLE.

1												13	14	15	16	17	18
1 **H**	2																2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

The salts of thulium possess a pale green color and the solutions have a slight greenish tint. The metal has a high vapor pressure at the melting point. When ^{169}Tm is irradiated in a nuclear reactor, $_{170}$Tm is formed. The isotope then emits strongly an 84-keV x-ray, and this material is useful in making small portable x-ray units for medical use. *See* RARE-EARTH ELEMENTS. Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

Thunder

The acoustic radiation produced by thermal lightning channel processes. The lightning return stroke is a high surge of electric current (~20,000 A) that occurs when the lightning flash makes contact with the Earth. The current surge has a very short duration, depositing approximately 95% of its electrical energy during the first 20 microseconds with its peak power occurring at 2 μs. Spectroscopic studies have shown that the lightning channel is heated to temperatures in the 20,000–30,000 K (36,000–54,000°F) range by this process. The lightning channel at this time has a diameter of approximately 1 cm (0.4 in.), and the pressure of the hot channel exceeds 10 atm (10^6 pascals). The hot, high-pressure channel expands supersonically and reaches a radius of 5 cm (2 in.) within the 20-μs period during which it is being heated. The channel continues expanding and forms a shock wave as it pushes against the surrounding air. Because of the momentum gained in expanding, the shock wave overshoots, causing the pressure in the core of the channel to go below atmospheric pressure temporarily. The outward-propagating wave separates from the core of the channel, forming an N-shaped wave that eventually decays into an acoustic wavelet. *See* SHOCK WAVE; STORM ELECTRICITY.

If the lightning channel were a perfectly straight line, the thunder produced by a single return stroke would consist of a single acoustic wavelet; the sound would be similar to that of a passing supersonic aircraft. However, real lightning channels are very crooked or tortuous. The formation of a single acoustic wavelet as described above applies, therefore, not to the channel as a whole but to the many, approximately straight segments of the lightning channel that locally behave as line sources of shock waves. The sound that is eventually heard or detected, thunder, is the sum of many individual acoustic pulses, each a remnant of a shock wave, that have propagated to the point of observation from the generating channel segments. The first sounds arrive from the nearest part of the lightning channel and the last sounds from the most distant parts.

To explain the amplitude variability in thunder, it is necessary to examine the radiation pattern from a channel segment. Laboratory experiments with sparks have shown that 80% of the acoustic energy from a line source is confined to a zone within ±30° of a plane perpendicular to and bisecting the line source. This means that the collection of thunder pulses from a long section of the channel perpendicular to the observer's line of sight (line of sound, actually) will all have higher amplitudes than the collection of pulses arriving from a long channel section more nearly parallel to the line of sound. Hence, the claps of high-amplitude thunder and the rumbles of low-amplitude thunder are each associated with the orientation of lightning channel segments with respect to the observer's location. It follows from this reasoning that the sound of thunder is unique to the shape of each lightning channel and to the location of the observer in relation to the lightning channel.

Another complicating aspect of thunder is the temporal structure of lightning. The description of thunder given above represents acoustics of a

single stroke of lightning. A lightning flash may produce several return strokes. In addition to the return strokes, there are lightning leader processes that also produce acoustic signals; however, they are thought to be much less energetic. For a complete description of thunder, the acoustic signals of all of these sources must finally be added together.

A further complicating aspect is the propagation path of the thunder pulses in the atmosphere. An acoustic ray is the path followed by a narrow beam of acoustic signals. In the atmosphere, which has temperature gradients and wind shears, the acoustic rays are bent into curved paths. Because of the decrease of air temperature with height, rays are most frequently bent upward. Therefore, observers on the surface may not hear all of the thunder from a visible lightning flash. Specifically, the higher the source of the sound, the farther it can be heard. Frequently, the thunder that is heard originates in the cloud and not in the visible channel. On some occasions, the observer may hear no thunder at all; this is more frequent at night when lightning can be seen over long distances and thunder can be heard only over a limited range (~10 km or 6 mi). *See* LIGHTNING; THUNDERSTORM. Arthur A. Few

Bibliography. A. A. Few, Thunder, *Sci. Amer.*, 233(1):80–90, 1975; National Research Council Staff (eds.), *The Earth's Electrical Environment*, 1986; H. S. Ribner and D. Roy, Acoustics of thunder: A quasilinear model for tortuous lightning, *J. Acous. Soc. Amer.*, 72:1911–1925, 1982; M. A. Uman, *The Lightning Discharge*, 1987, reprint 2001.

Thunderstorm

A convective storm accompanied by lightning and thunder and a variety of weather such as locally heavy rainshowers, hail, high winds, sudden temperature changes, and occasionally tornadoes. The characteristic cloud is the cumulonimbus or thunderhead, a towering cloud, generally with an anvil-shaped top. A host of accessory clouds, some attached and some detached from the main cloud, are often observed in conjunction with cumulonimbus. The height of a cumulonimbus base above the ground ranges from 1000 to over 10,000 ft (300 to 3000 m), depending on the relative humidity of air near the Earth's surface. Tops usually reach 30,000–60,000 ft (9000–18,000 m), with the taller storms occurring in the tropics or during summer in midlatitudes. Thunderstorms travel at speeds from near zero to 70 mi/h (30 m/s). In many tropical and temperate regions, thunderstorms furnish much of the annual rainfall. *See* LIGHTNING; THUNDER.

Development. Thunderstorms are manifestations of convective overturning of deep layers in the atmosphere and occur in environments in which the decrease of temperature with height (lapse rate) is sufficiently large to be conditionally unstable and the air at low levels is moist. In such an atmosphere, a rising air parcel, given sufficient lift, becomes saturated and cools less rapidly than it would if it remained unsaturated because the released latent heat of condensation partly counteracts the expansional cooling. The rising parcel reaches levels where it is warmer (by perhaps as much as 18°F or 10°C over continents) and less dense than its surroundings, and buoyancy forces accelerate the parcel upward. The convection may be initiated by a combination of surface heating, cooling of the upper layers of the troposphere, forced ascent of moist low-level air over rising terrain or at fronts and other boundaries (for example, drylines, cold-air outflow boundaries from other thunderstorms) or in gravity waves, and forced lifting of air by upper-air weather disturbances. The rising parcel is decelerated and its vertical ascent arrested at altitudes where the lapse rate is stable, and the parcel becomes denser than its environment. The forecasting of thunderstorms thus hinges on the identification of regions where the lapse rate is unstable, low-level air parcels contain adequate moisture, and surface heating or uplift of the air is expected to be sufficient to initiate convection. *See* FRONT.

Occurrence. Thunderstorms are most frequent in the tropics, and rare poleward of 60° latitude. In the United States, the Florida peninsula has the maximum activity with 60 thunderstorm days (days on which thunder is heard at a given observation station) per year. Thunderstorms occur at all hours of day and night, but are most common during late afternoon because of the diurnal influence of surface heating. The weak nighttime maximum of thunderstorms in the Mississippi Valley of the central United States is still a topic of debate.

Structure. Radar is used to detect thunderstorms at ranges up to 250 mi (400 km) from the observing site. Much of present-day knowledge of thunderstorm structure has been deduced from radar studies, supplemented by visual observations from the ground and satellites, and in-place measurements from aircraft, surface observing stations, and weather balloons. *See* METEOROLOGICAL INSTRUMENTATION; RADAR METEOROLOGY; SATELLITE METEOROLOGY.

Thunderstorms occur in isolation, in chaotic patterns over wide areas, in the eye walls and spiral bands of hurricanes, in clusters within large-scale weather systems, and in squall lines perhaps several hundred miles long. An individual thunderstorm typically covers a surface area of 10–400 mi^2 (25–1000 km^2) and consists of one or more distinct cells, each of which is several miles across, lasts about an hour, and undergoes a characteristic life cycle. In the cumulus or growing stage, a cell consists primarily of updrafts (vertical speeds of 20–110 mi/h or 10–50 m/s) with precipitation suspended aloft; in the mature stage, updrafts and downdrafts coexist and heavy rain falls to the ground; in the dissipating stage, a cell contains weakly subsiding air and only light precipitation. During the mature stage, downdrafts may reach 35 mi/h (15 m/s). The downdraft air is denser than its surroundings due to evaporational cooling, which occurs as clear air is entrained into the cloud from outside, and is forced downward by gravitational pull and by the drag of falling

precipitation. The downflowing air spreads outward in all directions as it nears the surface, and forms a cold, gusty wind that is directed away from the precipitation area. This advancing cold air may provide the necessary lift in neighboring warm moist air for the formation of new updraft cells. Intense, narrow downdrafts (sometimes from innocuous-looking cumulonimbi) produce locally damaging, divergent outflow winds as they impact the ground. These small-scale flow features, known as microbursts, contain large wind shears that are extremely hazardous to low-flying aircraft. *See* HURRICANE; PRECIPITATION (METEOROLOGY); SQUALL LINE.

In an environment where the winds increase and veer with height, and midlevel air is dry enough to provide the potential for strong downdrafts, a thunderstorm may become organized so as to maintain a nearly steady state for hours. In such a strong vertical shear of the horizontal wind, the updraft is tilted so that precipitation falls out of the updraft instead of through it, and updraft and downdraft can coexist for several hours in the configuration shown in **Fig. 1**. A long-lived storm in a sheared environment may consist of a single intense cell (supercell) or of many cells with an organized growth of new cells on one side of the storm (generally, the southwest in the Northern Hemisphere) and decay of old cells on the opposite flank. *See* WIND.

Severe storms. Thunderstorms are considered severe when they produce winds greater than 58 mi/h (26 m/s or 50 knots), hail larger than $\frac{3}{4}$ in. (19 mm) in diameter, or tornadoes. While thunderstorms are generally beneficial because of their needed rains (except for occasional flash floods), severe storms have the capacity of inflicting utter devastation over narrow swaths of the countryside. The greatest frequency of severe storms occurs in the Great Plains region of the United States during the spring, but only a small percent of the thunderstorms are severe. Severe storms are most frequently supercells which form in environments with high convective instability and moderate-to-large vertical wind shears. The supercell may be an isolated storm or part of a squall line.

Since severe storms constitute a hazard to aircraft, their internal dynamics has been deduced largely from radar measurements. Doppler radar is specialized to measure the velocity of radar targets parallel to the radar beam, in addition to the intensity of precipitation. Doppler radar studies and analysis of surface pressure falls have shown that large hail, high winds, and tornadoes often develop from a rotating thunderstorm cell known as a mesocyclone. Large hail, high winds, and weak tornadoes may form from nonrotating (on broad scale) multicellular storms, but are less likely. Maximum tangential winds around the typical mesocyclone are roughly 50 mi/h (20 m/s) and are located in a circular band which is 1–3 mi (1.5–4.5 km) in radius. A surface pressure deficit of several millibars exists at the mesocyclone center. In one case, a pressure drop of 34 mbar (3.4 kilopascals) was measured. Identification of a mesocyclone signature on radar has been

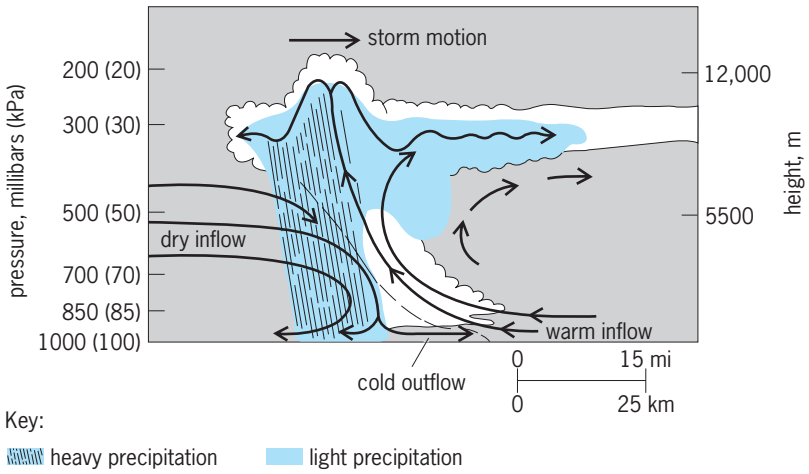

Fig. 1. Cloud boundaries and simplified circulation (arrows denote flow) of a typical mature thunderstorm in winds which blow from left to right and increase with height. Vertical scale has been exaggerated fivefold compared with the horizontal scale. 1 m = 3.3 ft; 1 km = 0.6 mi.

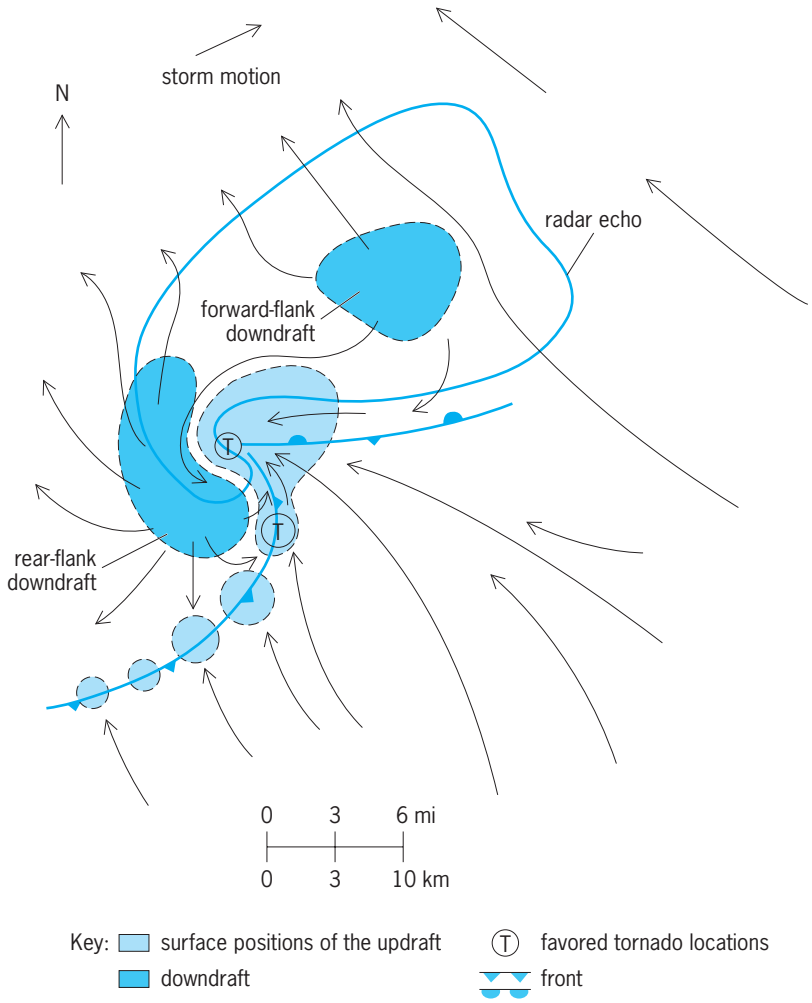

Fig. 2. Schematic surface plan view of a tornadic thunderstorm. The gust front is the boundary between unmodified warm, moist, inflowing air and rain-cooled, outflowing air. Arrows depict flow streamlines relative to the storm. The northern T at the mesocyclone center (wave apex) indicates where a major tornado is most likely. The southern T indicates the favored place for new mesocycline and tornado development. For storms in the Southern Hemisphere, transpose north and south. (*After R. Davies-Jones, Tornado dynamics, in E. Kessler, ed., Thunderstorm Morphology and Dynamics, 2d ed., University of Oklahoma Press, 1986*)

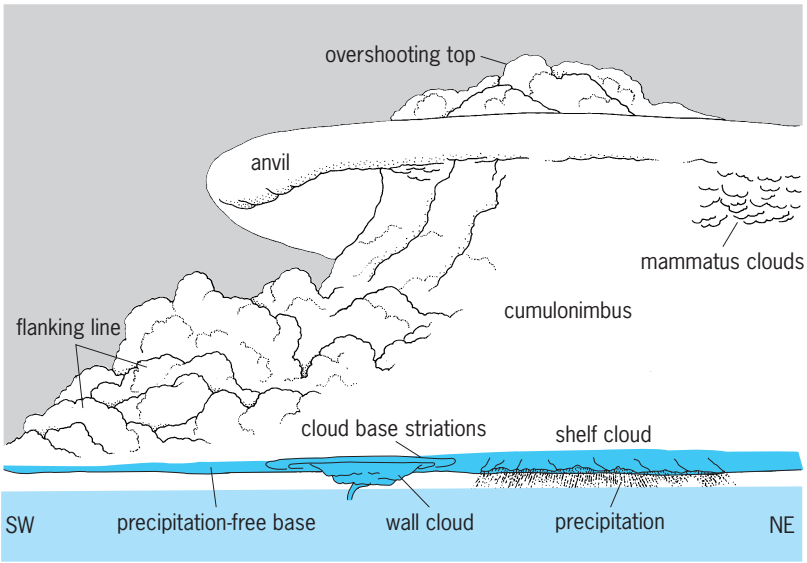

Fig. 3. Composite view of a typical tornado producing cumulonimbus as seen from a southeasterly direction. Horizontal scale is compressed, and all the features shown could not be seen from a single location. (*NOAA picture by C. Doswell and B. Dirham*)

used to issue severe weather warnings. The structure of a supercell storm is shown in **Fig. 2**.

On conventional radar displays, hook-shaped appendages to echoes are also good indications of mesocyclones, but unfortunately a large percentage of tornadic storms never exhibit such a hook. A mesocyclone sometimes is recognizable visually by rotation of a wall cloud, a discrete and distinct lowering of the cumulonimbus base (**Fig. 3**). The wall cloud is often seen visually to be rotating as an entity. The wall cloud is frequently the seat of intense vertical motions at low levels. The initial rotation of the mesocyclone at midlevels of the storm stems from the vertical shear of the winds flowing into the storm at low levels. An example of this effect is what happens to an initially vertical line drawn by a skywriter; changing wind speed (direction) with height causes the line to turn about a horizontal axis perpendicular (parallel) to the wind. The physically relevant winds are the ones relative to to the storm since it is the updraft that converts horizontal vorticity (spin) into vertical vorticity. In the extreme case when the storm-relative winds are constant in speed but veer (turn anticyclonically) with height, air parcels flowing into the storm have streamwise vorticity (that is, they spin about their direction of motion). As the parcels flow into the updraft, their spin axes are tipped upward, like a propeller on an aircraft entering a climb, and so the updraft rotates as a whole.

If the storm-relative winds back instead of veer with height, the vorticity is antistreamwise, resulting in anticyclonically rotating updrafts. In storm environments (on a rotating Earth), ground-relative winds generally veer and increase with height owing to effects of friction and flow of warmer air into the region. Generally the storm motion is such that the storm-relative winds also veer. Thus the majority of updrafts rotate cyclonically.

When the storm-relative winds increase with height without changing direction, air parcels spin like wheels normal to their direction of travel in the storm's reference frame. As the parcels flow into the updraft, their spin axes are tilted toward the vertical, resulting in cyclonic and anticyclonic spin on the right and left sides, respectively. Thus, the two halves of the updraft rotate in different directions in this case. In strong shear, the initial storm splits into two supercells that move to the right and left of the original storm motion. The cyclonic and anticyclonic halves of the initial updraft become the updrafts of the supercells. These updrafts continue to rotate as a whole after the split because net streamwise and antistreamwise vorticity is present in the inflows of the right- and left-moving supercells, respectively, owing to the deviate motions.

Once vertical vorticity has been generated, it can be amplified by the vertical stretching of air parcels in the lower part of the updraft where the flow converges horizontally. This process is analogous to a spinning ice skater. The development of rotation at the ground is a more complicated process, involving thermal generation of vorticity and the storm's downdraft as well as its updraft. *See* TORNADO; VORTEX.

Attempts have been made to modify thunderstorms to increase areal rainfall and suppress hail. The results of such experiments have been inconclusive. *See* HAIL; SQUALL; STORM DETECTION; WEATHER MODIFICATION.

For further discussion of storm electricity *see* ATMOSPHERIC ELECTRICITY; STORM ELECTRICITY.

Robert Davies-Jones

Bibliography. D. Atlas (ed.), *Severe Local Storms*, Amer. Meteorol. Soc., Meteorol. Monogr., vol. 5, no. 27, 1963; E. Kessler (ed.), *Instruments and Techniques for Thunderstorm Observation and Analysis*, 2d ed., 1988; E. Kessler (ed.), *The Thunderstorm in Human Affairs*, 2d ed., 1983; E. Kessler (ed.), *Thunderstorm Morphology and Dynamics*, 2d ed., 1992; F. H. Ludlam, *Clouds and Storms*, 1980.

Thyme

Any of a large and diverse group of plants in the genus *Thymus* utilized for their essential oil and leaves in both cooking and medicine.

Hundreds of different forms, or ecotypes, of thyme are found in the Mediterranean area, where thyme occurs as a wild plant. *Thymus vulgaris*, generally considered to be the true thyme, is the most widely used and cultivated species. Both "French" and "German" thyme are varieties of this species. Other species, such as *T. serpyllum* (mother-of-thyme) and *T. herba-barona* (caraway-scented), are grown or used in much the same way as *T. vulgaris*, but they do not have the same flavor or scent characteristics. Most types of thyme are low-growing perennials that do not exceed 15 in. (38 cm) in height. Typically, small smooth-edged leaves to 0.4 in. (1 cm) long are closely spaced on stems that become woody with age. Depending on soil type and climate, thyme may

live 10 years or longer, though in culture thyme is replanted every 5 years or less.

Cultivation of thyme is similar to that of rosemary. Much of the thyme seed available is harvested from wild plants. Seeds from a large number of forms are often mixed together. Though seed may be suitable or even desirable for small plantings, it is often unsuitable for large-scale cultivation. Rooted cuttings are usually employed for farm-size plantings, thus avoiding the variability of seed. Once established, thyme requires little care, since it has few pests or associated diseases. Verticillium wilt and *Rhizoctonia solani* have however, been reported to infect thyme.

Wild European thyme, the source of much imported material, is usually harvested only once a year, while cultivated plants in the United States are harvested mechanically up to three times a year. As with most herbs, both stems and leaves are harvested and then dehydrated. Dried stems and leaves are separated mechanically. Thyme oil is extracted from fresh material.

Thyme is a widely used herb, both alone and in blends such as "fines herbes." Thyme oil is used for flavoring medicines and has strong bactericidal properties. *See* SPICE AND FLAVORING. Seth Kirby

Thymosin

A polypeptide hormone synthesized and secreted by the endodermally derived reticular cells of the thymus gland.

Action. Thymosin exerts its actions in several loci: (1) in the thymus gland, either on precursor stem cells derived from fetal liver or from bone marrow, or on immature thymocytes, and (2) in peripheral sites, on either thymic-derived lymphoid cells or on precursor stem cells. The precursor stem cells, which are immunologically incompetent whether in the thymus or in peripheral sites, have been designated as predetermined T cells or T_0 cells, and mature through stages termed T_1 and T_2, each reflecting varying degrees of immunological competence. Thymosin promotes or accelerates the maturation of T_0 cells to T_1 cells as well as to the final stage of a T_2. In addition to this maturation influence, the hormone also increases the number of total lymphoid cells by accelerating the rate of proliferation of both immature and mature lymphocytes. Thus, both the numbers and state of immunological maturity of the lymphocyte, one of the prime classes of cells contributing to host immunity, are under the influence of thymosin. It is evident, therefore, that the thymus gland and its hormone, thymosin, play an essential role in the development and maintenance of host immunological competence. *See* IMMUNITY.

Host immunological competence. This phenomenon can be broadly divided into two types, namely humoral immunity and cell-mediated immunity. In the former, T cells participate cooperatively with lymphoid cells derived from bone marrow (B cells) in the synthesis of humoral circulating antibody as a response to the presence of a foreign soluble or insoluble antigen. In cell-mediated immunity, the immunologically competent T cell functions in phenomena based upon cell-mediated immunological responsivity of the host, for example, ability to recognize and reject foreign tissue or organ transplants or grafts, resistance to infections, particularly of a viral or fungal nature, and responsivity to foreign allergens, that is, normal or adequate responsivity in the delayed hypersensitivity reaction. It is also recognized that additional classes of T cells may exert other specialized functions. Thus, the ability to recognize "self," as contrasted with "nonself" or cells of foreign origin, resides apparently in a class of specialized T cells that have been termed suppressor T cells. *See* CELLULAR IMMUNOLOGY.

The actions of thymosin are the basis of its roles, and that of the thymus, in the regulation of the structure and functioning of host lymphoid tissue, including parameters of immunological competence.

Bioassay. There are several assay methods for thymosin; two are discussed here. One is done with cultures of sheep erythrocytes and mouse spleen cells; the other is done by injecting lymphoid cells into animals.

Spontaneous rosette cell-forming assay. This assay for thymosin activity is based upon the adherence of sheep erythrocytes to mouse spleen cells because of recognition by the spleen cells of a foreign antigen on the sheep erythrocytes. This results in four to six sheep erythrocytes adhering to the perimeter of each individual lymphoid cell, giving a "rosette" appearance to the latter. The numbers of spontaneous rosette-forming cells (SRFCs) formed in cell cultures by spleen cells of normal mice are inhibited by a concentration of azathioprine approximately one-tenth of that required to decrease the number of SRFCs in spleen cells of adult thymectomized mice. However, incubation of the latter cells with thymosin in cultures restores the sensitivity of these cells to the inhibitory action of azathioprine. Thus, it is possible to assay the potency of a particular thymosin preparation by titrating to the minimum quantity of thymosin required to produce a 50% inhibition of SRFCs numbers in the presence of a quantity of azathioprine that is inhibitory for spleen cells from a normal mouse. This assay is highly sensitive and has been utilized to detect picogram quantities of thymosin as well as to assess blood levels of circulating thymosin and the half-life (rate of disappearance from the blood) of intravenously injected thymosin.

GVH assay. One assay for thymosin is based upon the graft-versus-host (GVH) reaction. In this assay, lymphoid cells, usually spleen cells, of an animal (for example, mouse or rat) are injected into an animal of the same species but of differing histocompatibility and immunologically suppressed. This will result in an "attack" by the injected cells upon the spleen of the recipient animal, inducing a marked proliferative response with a significant splenic enlargement (splenomegaly) evident within 5 to 7 days after the injection. The magnitude of the splenomegaly is dependent upon the number of spleen cells

administered and the status of the immunological competence of the injected animal. Thus, injection of allogenic spleen cells into a mouse rendered immunologically incompetent will not cause splenomegaly in the recipient animal. In contrast, administration of thymosin to the immunologically incompetent animal will restore to a degree, dependent upon the amount and period of thymosin pretreatment, such competence and the animal will be able to respond to allogenic spleen cell injection with a resultant splenomegaly, or GVH reaction.

Purified preparations. Utilizing the above assays to assess biological activity, it has been possible, by classical methods of protein fractionation and purification, to obtain thymosin from fresh calf thymus glands in a purified, physically and chemically homogeneous form. A highly purified preparation of thymosin has also been obtained from human thymic tissue, and thymosinlike activity has been demonstrated in thymic tissue of the mouse, rat, guinea pig, and hog, as well as in human, mouse, and pig serum.

Physical and chemical studies of the homogeneous preparation of thymosin from calf thymus glands have established that the hormone belongs to the class of polypeptide hormones. Purified bovine thymosin has a molecular weight of 12,500 ± 200, consists of 108 amino acid residues, and is free of carbohydrate and lipid. No unusual amino acids are present in the molecule. Both calf and human thymosin are weakly antigenic and, when administered to rabbits under appropriate experimental conditions, stimulate antibody production. The antibody-containing serum has been utilized as an experimental tool for counteracting endogenous circulating thymosin, thereby producing an immunologically depressed or incompetent animal. Antibody-containing serum has also been used for the development of a radioimmune assay for assessing circulating blood levels of thymosin. *See* PROTEIN; RADIOIMMUNOASSAY.

Blood levels. The radioimmune assay and the spontaneous rosette cell-forming assay applied to human serum have revealed that thymosin blood levels are maximal at birth, with the concentration maintained at this level until approximately age 20. At this time, thymosin concentration in the blood begins to decline slowly and is nondetectable beyond 45 years of age. Preliminary data for serum thymosin concentrations in persons suspected of having a lower than normal immunological responsivity indicate that blood thymosin levels may be lower than normal in such individuals when compared with healthy individuals of the same age group.

Verifying hormone activity. One of the basic tenets for designating a substance as a hormone of an endocrine gland, such as the thymus, requires evidence that administration of the preparation to animals deprived of the gland in question will prevent the development of deleterious consequences, that is, the hormone can function in the absence of the gland. To provide such evidence for the thymus, a variety of laboratory animals were subjected to one of two experimental procedures: surgical removal of the thymus in the early neonatal period (within 24 to 48 h following birth); or thymectomy of the adult animal followed by involution or destruction of the major lymphoid organs either by exposure to whole-body x-radiation or by injection of an immunosuppressive drug such as cyclophosphamide or Imuran. *See* HORMONE.

Either of these experimental approaches results primarily in a loss of host immunological competence. This is reflected in failure of normal growth and development, a depressed or total loss of capacity to respond to foreign antigens by synthesizing antibodies, the inability to reject a foreign organ or tissue transplant, and a lack of resistance to infectious agents. Thus, the experimental animal exhibits depressed ability to function in the two major divisions of immunity: humoral immunity and cell-mediated immunity. *See* IMMUNOGLOBULIN; TRANSPLANTATION BIOLOGY.

Thymosin has been shown to function in the absence of the thymus gland. The effects of neonatal thymectomy on immunological capability can be prevented from occurring by treatment of the thymectomized animal with daily injections of thymosin preparations. These replacement therapeutic studies with thymosin, and other experimental approaches, have established that the thymus is an endocrine gland and that thymosin is a hormone, synthesized and secreted by the thymus and exerting its influence both within the thymus and in extrathymic loci, notably in other lymphoid organs (discrete lymph nodes) as well as the spleen. **Tables 1** and **2** list some of the demonstrated biological activities of thymosin in two test systems.

Immunodeficiency disorders. At least 21 well-recognized antibody and cellular immunodeficiencies have been described in humans. Certain of these disorders, characterized as primary immunological deficiency diseases, are directly related to the absence or hypoplasia of the thymus gland at birth. In others,

TABLE 1. Biological activities of thymosin in cell culture models

Activity	System
Conversion of precursor cells to immunologically competent lymphocytes	Rosette assay Cytotoxicity assay: expression of σ and TL antigens Responsivity to mitogens Mixed lymphocyte interaction Primary and secondary antibody response Conversion of cultured bone marrow cells into cells reactive in the graft versus host assay T and B cell cooperation (incubation in cell cultures, cells tested in animals)
Inhibition	Autosensitization of human lymphocytes Autologous rosette formation

TABLE 2. Biological activities of thymosin in animal models

Model	Activity
Normal mice	Lymphocytopoiesis Enhanced rate of allograft rejection Enhanced resistance to progressive growth of Moloney virus-induced sarcoma Enhanced mixed lymphocyte reaction (thymosin injected into animal, cells tested in culture) Enhanced lymphoid cell response to mitogens (thymosin injected into animal, cells tested in culture) Enhanced antibody synthesis
Germ-free mice	Lymphocytopoiesis
Adrenalectomized mice	Enhanced lymphocytopoiesis
Neonatally thymectomized mice	Increased survival and rate of growth Lymphocytopoiesis Restoration of ability to reject skin allograft Enhancement of mixed lymphocyte interaction Stimulation of antibody synthesis
Adult thymectomized mice	Restoration of sensitivity of spleen cells to azathioprine Inhibition of autologous rosette cell formation
Athymic "nude" mice	Lymphocytopoiesis Restoration of response of lymphoid cells to mitogen Reduction of allogeneic and xenogeneic tumor growth rate
Immunosuppressed mice	Enhancement of ability to reject skin allograft
NZB mice*	Delay of appearance of abnormal thymocyte differentiation, with loss of suppressor function

*Animals developing spontaneous autoimmune disease.

immune disorders may be associated with thymic malfunction. Studies indicate that aberrant or perhaps inadequate production of thymosin, or other thymic factors, may be a significant etiological factor in a variety of primary immune disorders, such as autoimmune diseases, in certain malignancies, as well as in resistance to oncogenic viruses and to many fungal and mycobacterial pathogens.

In selected numbers of immune disorders in which the deficiency is one of cell-mediated immunity, as for example, the DiGeorge syndrome, the number of circulating lymphocytes may be normal, or only slightly lower than normal. However, the maturation of the lymphoid cells has been arrested due to a failure of normal thymic development, including probable lack of functioning of cells responsible for the synthesis and secretion of thymosin and, perhaps, other thymic factors. Studies in selected immunodeficient individuals injected with a partially purified preparation of calf thymosin suggest that the hormone may induce the maturation of lymphoid cells of an immature or early T type to T_1 and T_2 cells with resultant enhancement of host immunological competence, as reflected, for example, in responsivity in tests for delayed hypersensitivity. *See* ENDOCRINE SYSTEM (VERTEBRATE); IMMUNOLOGICAL DEFICIENCY; THYMUS GLAND.

Abraham White

Bibliography. J. Kuby, *Immunology*, 3d ed., 1997; D. O. Norris, *Vertebrate Endocrinology*, 3d ed., 1996; E. R. Stiehm and V. A. Fulginiti, *Immunological Disorders in Infants and Children*, 5th ed., 2001; D. Van Bekkum and A. Kruisbeek (eds.), *Biological Activity of Thymic Hormones*, 1975; A. White and A. L. Goldstein, The endocrine role of the thymus, and its hormone, thymosin, in the regulation of the growth and maturation of host immunological competence, *Advan. Metabol. Dis.*, 8:361–376, 1975.

Thymus gland

An important central lymphoid organ in the neck or upper thorax of all vertebrates from elasmobranchs to mammals. The most primitive representatives of vertebrates which have been shown to possess this organ are the cyclostomes *Eptatretus stoutii* (California hagfish) and *Petromyzon marinus* (sea lamprey).

Embryologically, the thymus gland arises as an endodermal outgrowth from the pharyngeal portion of the alimentary canal and is invested by the surrounding mesodermal tissue, which gives rise to its connective tissue elements and blood vessels. The organ is later invaded by additional mesodermal blood-borne stem cells from the blood islands or bone marrow, and these differentiate into the precursor cells for the thymus-derived lymphocytes (thymocytes and T lymphocytes). The thymus gland is most prominent during early life. In many laboratory species of mammals and in humans it reaches its greatest relative weight at the time of birth, but its absolute weight continues to increase until the onset of puberty. Then it undergoes an involution and progressively decreases in size throughout adult life. The degree of involution varies greatly from species to species and appears to be correlated with the hormonal status of the animal.

Lymphocytes. The thymic stem cells generate a large population of small lymphocytes (thymocytes) through a series of mitotic divisions. Simultaneously these dividing lymphocytes show evidence of cellular differentiation within the special thymic environment. During this division and maturation phase the developing thymocytes undergo an intrathymic migration from the peripheral cortical area to the medullary core of the organ. Some thymocytes degenerate within the organ, but many enter the circulating blood and lymph systems at various stages of maturity. A small percentage of the T lymphocyte population (5–10%) within the thymus is antigenically competent and capable of recognizing antigenic determinants on foreign cells or substances. Some of the T lymphocytes have the capacity to lyse the foreign tissue cells, while others are involved in recognizing the "foreignness" of the antigens and assisting a second subpopulation of bone-marrow-derived lymphocytes (B lymphocytes) to respond to the antigen by producing a specific antibody. These two types of immunocompetent T lymphocytes are called killer cells and helper cells, respectively. They are involved in both tissue transplantation and

humoral antibody responses. These cells become a part of the memory cell components of the lymphoid tissue and recirculate between the blood and lymph systems. On the other hand, the vast majority of the thymic lymphocytes are immunologically incompetent (90–95%). Some thymocytes are thought to give rise to the smaller pool of immunocompetent T lymphocytes, but many emigrate into the circulating blood. Their extrathymic fate and function are presently unknown.

Comparative anatomy and embryology. Throughout life in most fishes and amphibians, and during development in other forms, the thymus is closely associated with the epithelial lining (endoderm) of the gill cavities, or the corresponding pharyngeal pouches of nonaquatic forms. *See* PHARYNX.

Fishes. The lamprey has a rudimentary thymus which is composed of a focus of lymphoid cells associated with the lining of the pharyngeal pouches. Elasmobranch fishes, sharks, and rays have a well-developed thymus which demonstrates a definitive cortex and medullary arrangement. These primitive vertebrates are also the first to show lymphoid aggregations in other areas of the body and immunological specificity toward foreign antigens.

Amphibians. The thymus occupies essentially the same position as in fishes with respect to the pharyngeal pouches. The tissue may be somewhat further removed from the pharyngeal mucous membrane and fuse into a bilateral organ which is composed of one or two masses in the neck.

Reptiles and birds. Thymus glands in these species usually occur as a bilateral chain of organs in the neck dorsolateral to the esophagus and medial to the jugular vein at approximately the level of the thyroid gland, or more rostrally. The individual anlage is derived mainly from the third and fourth pharyngeal pouches and may separate into several masses which correspond roughly to the spaces between the segmental nerves of the neck.

Mammals. The thymus gland is derived from the caudal pharyngeal pouches, III and IV. During development the thymuses become dissociated from the pharynx and appear to migrate both caudally and ventromedially. In many species the right and left glands approximate each other in the midline to form a single bilobed organ in the lower cervical or anterior thoracic region, or both. In humans and rodents the gland is principally situated in the anterior mediastinum between the sternum and the pericardium. Because of the migration pattern of this organ during development, inconstant accessory thymuses may be found in other locations between the pharynx and the thorax.

Histology of mammalian thymus. Each lobe is covered by a connective tissue capsule which subdivides the lobe into numerous lobules by trabeculalike extensions into the gland. These connective tissue septa provide a pathway for blood vessels to enter the gland. Beneath the connective tissue capsule and septa are the parenchymal components of the thymus. The parenchyma consist of a peripheral cortical region which is adjacent to the capsule and septa and a centrally positioned medulla. In humans each lobule is 0.02–0.08 in. (0.5–2 mm) in diameter and continuous with adjacent lobules by parenchymal extensions (see **illus.**).

The thymic mass is composed principally of two types of cells, reticular cells and lymphocytes. A small number of macrophages are also present within the gland. The reticular cells have a histological appearance similar to those found in other lymphatic organs. They possess a relatively large reticulated, pale-staining nucleus with one or more visible nucleoli. The nucleus is surrounded by a scant acidophilic cytoplasm which contains a small number of subcellular organelles. The thymic reticular cells interconnect with each other in a spongelike framework by extensions of their cytoplasm. They differ from the mesodermally derived reticular cells of the other lymphomyeloid tissues in that they are derived from endoderm, lack the ability to produce reticular fibers, and are nonphagocytic. The proposed

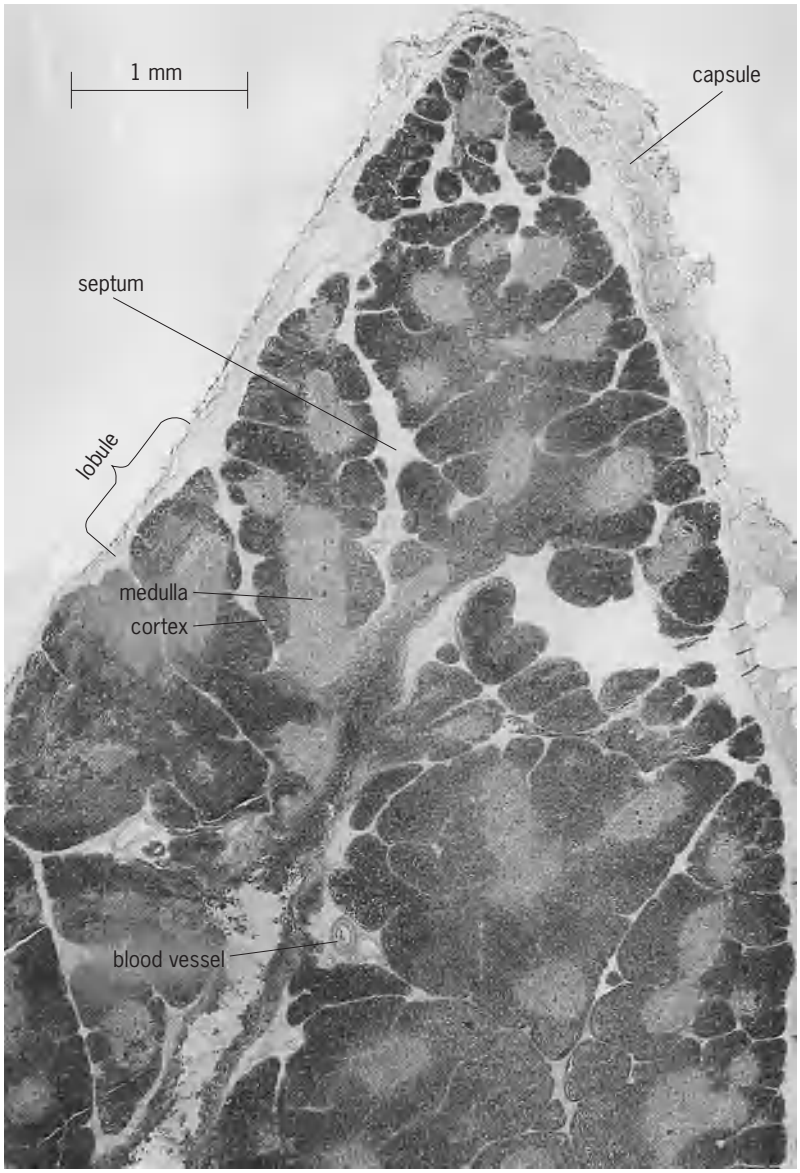

Thymus of a 3-month-old human infant. Histological section through a portion of one lobe.

functions of reticular cells are to support the lymphocyte population within the thymus, assist in the formation of the thymic-blood barrier, and possibly produce a thymic hormone.

The thymus lymphocytes include large, medium, and small forms. The larger lymphocytes are concentrated mainly in the subcapsular and septal regions of the cortex and possess a large round nucleus, 9 micrometers in diameter. These cells are mitotically active in producing smaller lymphocytes and usually show abundant amounts of strongly basophilic cytoplasm. The small lymphocytes (thymocytes) of the cortex superficially resemble the typical small lymphocytes of other lymphatic organs and circulating blood. They make up the vast majority of the cells found within the cortical region and possess a densely stained nucleus, 4–5 μm in diameter, with a slight amount of surrounding cytoplasm. Thymocytes differ from the small lymphocytes in the peripheral blood and other lymphatic tissue in that they are easily destroyed by corticosteroids and x-radiation; they are less well provided with intracellular organelles (nucleoli, mitochondria, and endoplasmic reticulum); they are immunologically incompetent; and they possess a unique array of surface antigens which allows them to be distinguished from the immunocompetent T lymphocytes and the bone-marrow-derived lymphocytes.

The medullary region contains fewer small lymphocytes relative to the number of reticular cells. This cellular organization produces a less compact appearance, and more large blood vessels are evident. The reduced cellular density of the medulla contrasts markedly with the densely populated cortex. The interface between these two distinct regions of the gland is termed the corticomedullary junction. The medulla contains the majority of the immunocompetent T lymphocytes, which are believed to arise from a population of maturing thymocytes. The majority of thymus macrophages are also present in the medulla and may be of blood-borne origin. In addition, a real histological peculiarity of the thymus exists within the medulla: Hassall's bodies, which are concentrically layered, flattened cells, probably of reticular origin. The overall size of these bodies ranges from 30 to 100 μm in diameter. The central cells in these structures tend to degenerate and hyalinize in a manner reminiscent of stratified epithelium and epithelial pearls of other organs. They are not present in the thymus at birth but do arise early in life during the development of immunological maturity. Their origin and function are unknown, but they serve as a histological marker for thymic tissue even during stages of severe involution.

The thymic blood supply and vascular organization have been extensively studied and provide some insight into possible mechanisms of thymic function. Arterioles leave the connective tissue septa and enter the parenchyma at the corticomedullary junction. They ramify within the medulla and supply a capillary network in the cortex. Blood in the cortex returns to the postcapillary venules, which are also located at the corticomedullary junction and in the medulla. The capillaries within the cortex are modified to prevent the movement of high-molecular-weight macromolecules into the thymic parenchyma. This specialized capillary structure is the basis for the thymic blood barrier, which functions to prevent antigenic material from penetrating the cortical region of the thymus. The blood vessels within the medulla are, however, permeable to antigens and to other macromolecules as well.

Physiology. Lymphopoiesis within the body may be induced from a stem cell source or an antigenic stimulation of immunocompetent small lymphocytes. The production of lymphocytes from stem cell sources occurs almost entirely within the central lymphoid tissue: bone marrow and thymus. Antigenically induced lymphopoiesis can occur at any site within the body into which the necessary cells can migrate and accumulate in the presence of the inducing antigen. These sites are most commonly found within the spleen and lymph nodes.

Lymphopoiesis. The lymphopoiesis within the thymic cortex is unique in that the lymphocytes are being produced and differentiating in an antigen-free environment. This specialized environment results from the presence of the thymic blood barrier within the cortical region of the thymus. As the cortical thymocytes mature, they undergo an intrathymic migration from the peripheral region of the cortex to the corticomedullary junction. This migration takes 2–4 days. The fate of the mature thymocytes is not entirely clear. Many of these cells undergo a premitotic death within the thymus and form visible structures known as tingible bodies. These are dark-staining nuclear concentrations principally found in the cortex, and each contains a double complement of DNA. Many of the remaining thymocytes emigrate from the thymus into the peripheral blood. It is believed that their route of emigration is via the venules and lymphatics present at the corticomedullary junction. This leaves a small percentage (5–10%) of the thymocytes which undergo an intrathymic maturation to give rise to that immunocompetent T-cell population which is found mainly in the medulla. The immunologically incompetent thymocytes can be distinguished from the immunologically competent T lymphocytes by the presence of specific surface antigens. The thymocytes that emigrate into the circulating blood may also be capable of transforming into mature T lymphocytes in the peripheral lymphoid tissue. *See* HEMATOPOIESIS.

Immunology. The T lymphocytes both within and outside the thymus display a functional heterogeneity and are involved in a wide range of immunological functions. They are responsible for tissue transplantation immunity and the rejection of homografts, for delayed hypersensitivity responses, and for immune responses against many microorganisms and viruses. In the tissue transplantation and delayed hypersensitivity responses, these cells interact with foreign cells and produce cell destruction by lysing the foreign or altered cells. An independent subpopulation of T lymphocytes functions as a group of

antigen-recognizing cells and assists the B lymphocytes in the production of humoral antibodies. The T lymphocytes are also responsible for much of the long-term immunological memory which allows higher animals to respond in a more efficient manner to recurring exposure to antigens. These cells make up the major portion of the population of small lymphocytes which constantly recirculate through the major lymphatic channels. *See* CELLULAR IMMUNOLOGY; TRANSPLANTATION BIOLOGY.

Thymectomy. Much knowledge about the function of the thymus has been derived from experiments on animals in which the thymus has been surgically or chemically removed. Such thymectomized animals show profound defects in their abilities to populate peripheral lymphoid tissues with lymphocytes and induce immune responses. If the thymectomy is performed in the prenatal or early postnatal periods, the thymectomized animals show severe lymphopenia and virtually total incompetence in their immunological reactivity to thymus-dependent antigens. The population of lymphocytes in the thoracic duct system of neonatally thymectomized animals may be only 2–3% of that found in normal, nonthymectomized control animals. The lymph nodes and spleens of these neonatally thymectomized animals are also poorly populated with lymphocytes. If the thymectomy is performed in an adult animal, little immediate change in population or function is noted. Peripheral lymphatic organs and pools of lymphocytes show normal populations of T lymphocytes. Immunological activity is also normal. With time, however, the T lymphocyte population begins to decrease and is no longer replaced in the absence of the thymus. Adult thymectomized animals will eventually show the same effects as those seen in neonatally thymectomized animals, particularly in responses to newly encountered antigenic stimuli.

Involution. The functional activity of the thymus is most intensive from the late fetal stages to puberty in most mammals. In humans the thymus begins to involute after puberty but is still an active and functional organ throughout most of adult life. The process of involution decreases the parenchymal elements of the thymus and gradually replaces these elements with fat. During involution, thymocytes are first to disappear, followed by the T lymphocytes and reticular cells. Histologically, the cortex becomes indistinct as the thymocytes are depleted, and finally the medullary region may be distinguishable only by the presence of keratinized Hassall's bodies. Premature involution is produced in times of stress, mainly caused by increased levels of circulating corticosteroids. Therapeutic dosages of corticosteroids may deplete a normal thymus of its lymphocytic population within 48 h. Repopulation of the thymus is then initiated by blood-borne stem cells from the bone marrow as the circulating steroid levels again approach normal concentrations.

Hormones. In addition to lymphopoiesis, production of a hormone has been proposed as a second major thymic function. There is much speculative evidence for the presence of a thymus hormone, but a great deal of controversy still surrounds its specific nature and action. In fact, more than one hormone has been proposed. It has been suggested that the endodermally derived reticular cells of the thymus are the sources of these hormones and that they act to produce a maturing influence on the differentiating thymocytes. Their influence is not only intrathymic but systemic, as suggested by studies involving thymic extract injections and transplants with thymuses enclosed in millipore chambers. What role these putative hormones may play in thymic function remains to be determined through further investigation. The thymic response to pituitary growth hormone and adrenal corticosteroids, together with its endocrinelike developmental history, suggests that some hormonal activity is plausible. *See* ENDOCRINE SYSTEM (VERTEBRATE).

Clinical disorders. While the preceding discussion emphasizes the role of the thymus in providing the animal with an efficient means of protection against certain diseases, a defect in thymic function may underlie some of the etiology for other diseases. Myasthenia gravis has been shown to correlate closely with thymic disorders. Thymoma and the presence of thymic germinal follicles are commonly found in individuals with myasthenia gravis. Some evidence for a common antibody against thymic epithelial cells and muscle cells suggests that myasthenia gravis may be an autoimmune disease. Thymomas have also been associated with anemias and other suspected autoimmune disorders. These diseases appear to be enhanced when an immunological deficiency is present. It has been postulated that clones of cells that could react with self-antigen escape from a suppressed stage and proliferate. These lymphocytes then react with the various organ systems which contain these antigens. It has also been shown that immune deficiencies can enhance the incidence of carcinogenesis (cancer). Thymic disorders and involution are associated with increased carcinogenesis. On the other hand, some leukemias require an active thymus to sustain the disease. This may be due to the requirement for some hormone factors or to the intrathymic growth of a viral agent that could promote the leukemia. In these cases a thymectomy may be an asset in the treatment of the disease even though it results in an immunological deficiency with time. Thymocytes are also affected in acquired immune deficiency syndrome (AIDS). A human immune deficiency virus (HIV) can infect thymocytes and lead to their destruction and a subsequent immunodeficiency. *See* AUTOIMMUNITY; LEUKEMIA; MYASTHENIA GRAVIS; ONCOLOGY.

Bursa of Fabricius. The development of the bursa of Fabricius as a gut-associated lymphoepithelial organ in birds appears to be a unique step in evolution. The bursa serves as a second central lymphoid organ and, together with the thymus, plays an important role in the development of immunological competency in birds. Mammals lack a definitive bursa, and attempts to define a bursal equivalent have thus far failed. Some scientists feel that the gut-associated lymphoid tissue may represent this avian structure, but it is equally probable that the

bone marrow lymphocyte population contains the bursal equivalent in mammals. These cells would be the source of the B lymphocyte population. *See* LYMPHATIC SYSTEM.

Charles E. Slonecker

Bibliography. P. S. Amenta, *Histology and Human Anatomy*, 6th ed., 1991; N. A. Byron and J. R. Hobbs (eds.), *Thymic Factor Therapy*, 1984; J. A. Goss, Y. Nakafusa, and W. M. Flye, *The Thymus and Cellular Immunity*, 1993; M. W. Hess, *Experimental Thymectomy*, 1968; L. C. Junqueira et al., *Basic Histology*, 9th ed., 1998; M. D. Kendall (ed.), *The Thymus Gland*, 1982; D. Metcalf, *The Thymus*, 1966.

Thyrocalcitonin

A hormone, the only known secretory product of the parenchymal or C cells of the mammalian thyroid and of the ultimobranchial glands of lower forms. The hormone has been isolated and characterized from porcine and human thyroid tissue as well as from fish ultimobranchial tissue. In all cases it is a polypeptide of 3800 molecular weight characterized by a single polypeptide chain. The amino acid sequence of the porcine hormone is

H_2N-Cys-Ser-Asn-Leu-Ser-Thy-Cys-Val-Leu-Ser-Ala-
└──────────────────┘
Tyr-Trp-Arg-Asn-Leu-Asn-Asn-Phe-His-Arg-Phe-Ser-
Gly-Met-Gly-Phe-Gly-Pro-Glu-Thr-Pro-$CONH_2$

The sequences of both the human and fish hormone have also been determined and are different from each other and from the porcine hormone except in the *N*-terminal seven amino acids. *See* ULTIMOBRANCHIAL BODIES.

Physiological activity. In conjunction with the parathyroid hormone, thyrocalcitonin is of prime importance in regulating calcium and phosphate metabolism. Its major function is to protect the organism from the dangerous consequences of elevated blood calcium. Its sole known effect is that of inhibiting the resorption of bone. It thus produces a fall in the concentration of calcium and phosphate in the blood plasma because these two minerals are the major constituents of bone mineral and are released into the bloodstream in ionic form when bone is resorbed. The inhibition of bone resoption also leads to a decreased excretion in the urine of hydroxyproline-containing peptides, which are end products in the breakdown of bone matrix, collagen. *See* BONE; CALCIUM METABOLISM; PARATHYROID GLAND; PARATHYROID HORMONE.

Thyrocalcitonin also causes an increased excretion of phosphate in the urine under certain circumstances, but a question remains as to whether this is a direct effect of the hormone upon the kidney or an indirect consequence of the fall in blood calcium which occurs when the hormone inhibits bone resorption. *See* PHOSPHATE METABOLISM.

Biochemical mechanisms. Practically nothing is known about the biochemical basis for the hormonal inhibition of bone resorption, although it has been shown that thyrocalcitonin inhibits the resorption of bone grown in tissue culture when this resorption is increased by the addition of parathyroid hormone, vitamin D, or vitamin A. *See* VITAMIN A; VITAMIN D.

Repeated administration of the hormone to young animals leads to an increased bone mass, but it is not known whether this is due solely to the fact that the hormone has inhibited bone resorption while bone formation has continued unchanged, or whether it has also enhanced bone formation.

Age factors. One of the most unusual features of this hormone's effect is the considerable decrease in sensitivity with age. With many hormones, a given dose, on a unit per weight or surface area basis, is one-half to one-third less effective in adults as compared to young animals. However, in the case of thyrocalcitonin, adult animals are considerably less sensitive, by a factor of one-fiftieth or more, than young growing animals. There is a good correlation between this decreased sensitivity and the fact that the rate of bone remodeling, both formation and resorption, is very high in young growing animals but decreases greatly in adult life.

Control mechanisms. Another unusual aspect of thyrocalcitonin physiology is the fact that the content of thyrocalcitonin in thyroid tissue increases greatly when there is no stimulus for its release, that is, when the blood calcium is low for any reason. In most other endocrine organs there appears to be a close relationship between the rate of synthesis of the hormone and the rate of its secretion, so that the content of hormone increases with increased secretion rates and decreases with decreased secretion rates.

Thus the only known stimulus to thyrocalcitonin secretion is a rise in blood calcium. The increased secretion, so induced, leads to a fall in blood calcium so that a negative feedback relationship exists between thyrocalcitonin secretion, bone resorption, and blood calcium.

Therapeutic value. There are no known diseases which have been clearly recognized as being due to chronic overproduction or underproduction of this hormone. Nevertheless, it has been found effective in the treatment of two known human diseases: idiopathic hypercalcemia of infancy, and hypercalcemia in adults caused by overproduction of parathyroid hormone. In both conditions, the concentration of calcium in the blood is greater than normal, and this leads to serious consequences, particularly calcification of the kidneys, which may eventually cause sufficient damage to the kidneys to cause death. The administration of thyrocalcitonin to people with these diseases leads to a prompt fall of the plasma calcium to normal levels.

It is hoped that this hormone will also be useful in several other disorders of bone metabolism, but because of the extremely small amount present in thyroid glands, and the difficulty of purifying it from these glands, only limited amounts have been available for medical use. *See* THYROID GLAND.

Howard Rasmussen

Bibliography. D. O. Norris, *Vertebrate Endocrinology*, 3d ed., 1996.

Thyroid gland

An endocrine gland found in all vertebrates that produces, stores, and secretes the thyroid hormones. The primary function of the thyroid, in warm-blooded vertebrates at least, is to regulate the rate of metabolism. In humans, the gland is located in front of, and on either side of, the trachea (**Fig. 1**). Thyrocalcitonin, a hormone of the thyroid gland, assists in regulating serum calcium by reducing its levels. The thyroid gland is capable of accumulating inorganic iodides and uniting them with the amino acid tyrosine to produce iodinated proteins. This activity is regulated by thyrotropic hormone from the anterior lobe of the pituitary gland. Microscopically (**Fig. 2**), the tissue consists of thyroid follicles composed of a single layer of secretory cuboidal epithelium which secretes hormone as a colloid into a blind lumen. The height of the secretory cells and the amount of colloid vary with the functional state of the gland. A main site of action of thyroid may be in the mitochondria, where energy-rich phosphate bonds are formed.

The thyroid receives an exceptionally rich blood supply. Postganglionic fibers from the cervical ganglia and vagus enter with the blood vessels and form extensive plexuses around the smaller arteries. This innervation does not appear to be essential for normal thyroid function except as it controls the rate of blood flow through the gland.

Comparative Anatomy

The thyroid is usually a well-encapsulated single gland, often with two distinct lobes connected by a narrow isthmus. However, there are morphological variations among the various vertebrate classes.

Fishes. In most fishes the thyroid is represented by diffuse masses of follicles scattered along the large arteries which enter the gills ventrally, or tucked between muscles in the pharyngeal floor. In elasmobranchs the follicles are aggregated in a single mass. In the remaining fishes (except lampreys) and in higher vertebrates a pair of thyroid masses, often connected by a median strand (isthmus), usually occurs. Minute accessory thyroid masses are common.

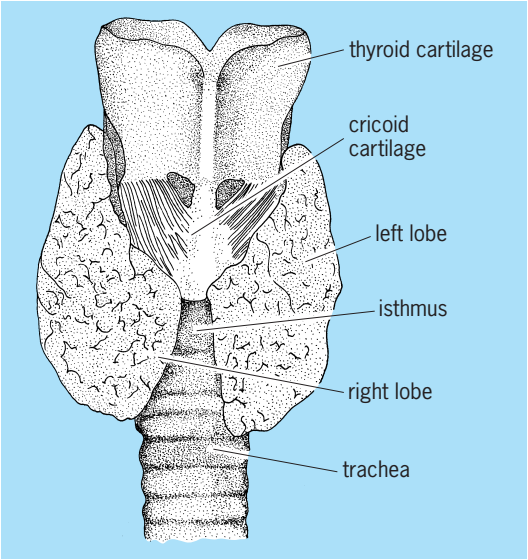

Fig. 1. Ventral view of human thyroid gland shown in relation to trachea and larynx. (*After C. K. Weichert, Elements of Chordate Anatomy, 3d ed., McGraw-Hill, 1967*)

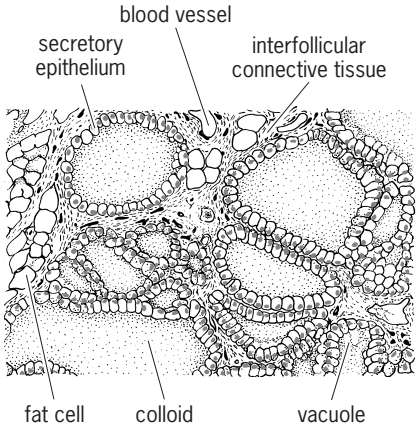

Fig. 2. Microscopic view of histologic features of the normal thyroid gland of the rat. (*After C. D. Turner, General Endocrinology, 4th ed., Saunders, 1966*)

Amphibians. In amphibians the thyroid glands lie under cover of certain muscles of the buccal or pharyngeal floor near the caudal angle of the jaws (anurans), or at the base of the branchial arches (urodeles).

Amniotes. In amniotes the two oval or elongated glands lie against the trachea, immediately below the larynx in mammals (Fig. 1), partway down the trachea in lizards, still farther caudad in other reptiles, or just above the bifurcation of the bronchi in birds. In most mature reptiles the gland is unpaired. In humans the thyroid glands are attached to the thyroid (shield-shaped) cartilage of the larynx.

Phylogeny

A clue to the phylogenetic history of the thyroid is found in lampreys. In marine species an elongated rod of cells capable of selectively absorbing iodine-rich substances is located in the pharyngeal floor between the second and fifth gill pouches. In larval brook lampreys the iodine-capturing cells are part of a complicated subpharyngeal gland (endostyle) which evaginates from the embryonic pharyngeal floor. At metamorphosis the gland loses its connection with the pharynx and remains as isolated thyroid masses underneath the pharynx. Embryonic origin of the thyroid of higher vertebrates as a pharyngeal outpocketing probably represents a recapitulation of the phylogeny of the gland. Although the embryonic pharyngeal connection is usually lost, a duct remains patent in some elasmobranchs. Even in humans remnants of the duct may persist. George C. Kent

Embryonic Origin

In the lamprey, the thyroid arises from some of the cells that line the larval endostyle, a groove lying in the floor of the pharynx. In all other vertebrates it originates from the same region, usually arising as a groove or pit at the level of the first pair of gill pouches (**Fig. 3**). In some fishes and amphibians

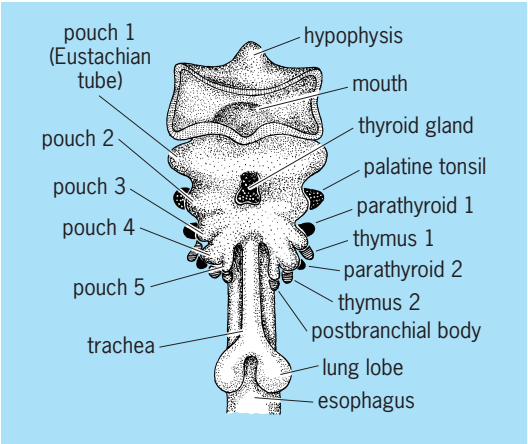

Fig. 3. Ventral view of pharyngeal region of a human embryo showing the pharyngeal pouches and their glandular derivatives; semidiagrammatic. (*After H. V. Neal and H. W. Rand, Chordate Anatomy, Blakiston, 1939*)

the cells appear as a solid bud rather than a hollow structure. In further development the bud or pit separates from the pharynx and migrates backward to lie ultimately in the throat region below the ventral aorta or the trachea. During this migration the cells of the primordium multiply and arrange themselves into elongate cords or flattened plates. The cords finally break up into discontinuous groups of cells, the early follicles, which with the onset of secretion become blind sacs. The definitive follicles consist of a single epithelial cell layer surrounding a cavity, the lumen, filled with a fluid, the colloid. The follicles are bound together by connective tissue (Fig. 2), and the whole gland becomes covered by a well-defined capsule. Frequently, small groups of functional follicles are left behind along the embryonic route of migration. These are called accessory thyroids or thyroid "rests." In some bony fishes, in which the thyroid is not encapsulated, follicles may wander rather far from the pharyngeal area to such unusual regions as the kidney, spleen, eye, or brain.

Various vertebrate groups differ as to the time when the thyroid first shows evidence of secretory activity and when its hormones begin to affect development. In the frog the first organized follicles are seen quite early, when the tadpole has completed only about 10% of larval life; in the chick they appear later, when about 30% of the period of incubation has passed; in the human, at approximately halfway through intrauterine development; and in the rat, after 80–90% of intrauterine life. The onset of thyroid function, determined by the point at which the gland is able to take up radioiodine, precedes the actual appearance of colloid, but there is evidence that organic combination of iodine, and therefore formation of true hormone, is closely correlated with the time of origin of colloid. W. Gardner Lynn

Development, Differentiation, and Morphogenesis

The thyroid gland as a distinct histologic entity is found only in vertebrates. The origin of vertebrates from the lower forms is obscure, with the closest relatives being the protochordates, tunicates, and amphioxus. Even in these groups, however, evidence of thyroid evolution from prevertebrate ancestry is inconclusive. In forms below vertebrates, there is no significant thyroid hormone formation, although iodotyrosines (iodinated amino acids; of unknown function) are found. Conversion from the sessile form of the jellyfish to the medusa or free-floating form depends partly on iodine concentration and suggests some form of control by iodotyrosine.

The role of thyroid hormones in lower vertebrates, therefore, is of interest. The thyroid hormones thyroxine (T4) and triiodothyronine are synthesized in the thyroid gland, but most of triiodothyronine is formed by the peripheral deiodination of thyroxine. Triiodothyronine is considered to be the active hormone, with thyroxine being considered a prohormone. Triiodothyronine plays an important role in thermal regulation and in the control of postembryonic development. *See* THERMOREGULATION.

Developmental effects are found most strikingly during the metamorphosis of a tadpole into a frog; metamorphosis will not occur in the absence of thyroid hormones. Interestingly, the deiodinase enzyme responsible for converting thyroxine to triiodothyronine is detectable in the skin of premetamorphic tadpoles, and enzyme activity significantly increases as the tadpoles go through spontaneous metamorphic climax. The major tissues responsible for producing triiodothyronine from thyroxine in amphibians are gut and skin. Significant deiodinase activity is present in tail tissue once the fin starts to be resorbed.

Thyroid hormones and retinoic acid. Development depends on morphogenetic signals that determine changes in gene expression in particular cells. The role played by thyroid hormones in development became clearer after it was discovered that retinoic acid is a major factor in morphogenesis and the retinoic acid receptor is a member of the thyroid receptor family. Furthermore, the retinoic acid receptor can activate gene expression through a thyroid hormone response element. The implication is that thyroid hormones and retinoic acid, acting through their respective receptors, control overlapping gene networks involved in the regulation of vertebrate development, differentiation, and morphogenesis.

The relationship between the retinoid receptor and the thyroid hormone receptor is surprising because retinoids and thyroid hormones bear little resemblance to one another. However, this probably reflects a common mode of action by which they elicit their particular regulatory effects. Thus, the interaction of retinoic acid with its receptor would induce a cascade of regulatory events that result from activation of specific sets of genes by the hormone receptor complex. The demonstration that the retinoic acid receptor is part of the steroid receptor superfamily suggests that mechanisms controlling morphogenesis may be more universal than previously suspected.

Metamorphosis. Thyroid hormones are essential for amphibian metamorphosis. Extensive structural changes occur during the transition of the larval tadpole to the adult frog (**Fig. 4**): Resorption of gills and tail occurs with development of lung and limb buds. Profound biochemical changes also occur.

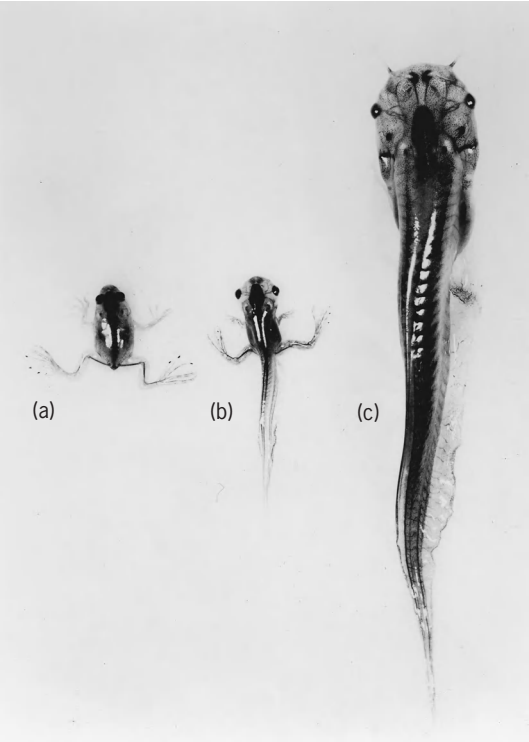

Fig. 4. Effect of thyroid hormone on metamorphosis of the amphibian *Xenopus laevis* (South African clawed toad). (*a*) Newly metamorphosed toad. (*b*) Metamorphosing tadpole. (*c*) Giant tadpole showing no metamorphosing signs after thyroidal hormone synthesis was blocked by immersion in perchlorate, an antithyroid drug. (*From J. A. Dodd and A. J. Matty, Comparative aspects of thyroid function, in R. Pitt-Rivers and W. R. Trotter, eds., The Thyroid Gland, Butterworth, 1964*)

Metamorphosis has been divided into three stages: premetamorphosis, prometamorphosis, and metamorphic climax. Premetamorphosis is characterized by rapid body growth without differentiation and without thyroid hormone. Prometamorphosis is characterized by diminishing growth rate, the beginnings of differentiation, and increasing thyroid hormone concentrations. The metamorphic climax is associated with growth cessation, maximal rates of differentiation, and a thyroid hormone surge that eventually drops to undetectably low levels. Although the tadpole is very sensitive to thyroid hormone effects, the frog seems not very sensitive.

Tremendous diversity exists within the metamorphic process. Some organs are stimulated under the influence of thyroid hormones while others resorb. Despite the similarity between thyroid hormones and steroid receptors, gonadal development in amphibians is not influenced by thyroid hormones. If early metamorphosis is evidenced by exogenous administration of thyroid hormones, precocious sexual maturity does not occur.

The sequencing and spacing of metamorphosis are determined by the secretory activity of the thyroid and by differences in the rate of response of affected tissues. For instance, hindlimb growth begins early and requires weeks or months for completion. In contrast, tail resorption and loss of mouthparts in the tadpole occur later in metamorphosis and more rapidly.

The various tissue and organ responses differ among groups of amphibians. Anurans lose their tails at metamorphosis while urodeles do not. Rapid growth of the hindlimbs occurs in the anurans but not in the urodeles. The characteristic differences in the mode of response are genetically determined. Tissues of very young larvae are not capable of manifesting any metamorphic response even when exposed to high concentrations of thyroid hormones. Sensitivity to thyroid hormones appears at the stage of development when the external gills become covered by the opercular folds. The concept of a time window for the action of thyroid hormones occurs with higher species as well.

Neoteny. Some amphibians, such as the Mexican axolotl (*Ambystoma mexicanum*), fail to metamorphose. Instead, they mature and reproduce but retain a larval form; this is known as neoteny. However, treatment with thyroid hormones will induce these salamanders to undergo metamorphosis. Neoteny is characterized by the cessation of metamorphosis at the larval stage, coupled with an increase in size of the larvae and maturity of their gonad systems.

The hypothalamus is essential for amphibian metamorphosis. The preoptic nucleus of the hypothalamus may be the location of the control center for thyroid-stimulating hormones (TSHs). Also, an important fraction of the brain thyrotropin-releasing hormone (TRH) is located in this area. In birds and mammals, TRH stimulates thyrotrophic cells in the pituitary to produce TSH; TRH is also present in the hypothalamus of amphibians, including the axolotl, and may control thyroid activity in amphibians.

While thyroxine is indispensable for metamorphosis of amphibian larvae, the pituitary hormone prolactin is the growth-promoting hormone. An antagonistic interaction exists between prolactin and thyroxine, but it does not appear to be mediated through the thyroid gland.

Fishes. Extraordinarily high concentrations of triiodothyronine have been found in the ammocoete stage of the lamprey, a cyclostome. Nuclear triiodothyronine receptors in the liver of these animals appear to resemble mammalian receptors. Despite the presence of thyroid hormones and thyroid hormone receptors, the function of thyroid hormones in fishes has remained elusive. Little convincing direct evidence has been found to indicate that thyroid hormones are involved in growth and development of fishes. Thyroid hormones characteristically function in coordination with other hormones and growth factors. If the coordinate signal is missing, thyroid hormones may not produce any effects.

Exogenous thyroid hormones do induce metamorphosis in flounder larvae. The antithyroid drug, thiourea, blocks the effect of thyroid hormones. The role of thyroid hormones in metamorphosing larvae, at least in flounder, is comparable to its role in amphibians.

Development in homeotherms. Thyroid hormone also plays an essential role in the ontogenesis of higher vertebrates. As in amphibians, it does this at a

well-defined stage of development. The precise stage at which thyroid hormones influence development in a given species is related to duration of gestation and degree of maturity at birth.

In the rat the diminished growth rate characteristic of a deficiency of thyroid hormones (hypothyroidism) does not occur until the neonate is 8 days old. Clinical recognition of hypothyroidism in humans at birth is extremely difficult. Apparently, the major effects of thyroid hormones in both the rat and the human become manifest after a surge of thyroid hormones in a manner analogous to the amphibian metamorphic climax.

The manifestation of neonatal hypothyroidism in humans is most noticeably expressed in marked skeletal retardation with profound disturbances in the central nervous system. Thyroid hormone deficiency involves the maturation of individual cells as well as the retardation of general body growth. Although deficiency of thyroid hormones in the adult can be easily corrected, abnormalities in the neonate can be prevented only if the hormones are replaced at specific times. Thyroid hormones appear to function by an increase in general growth as well as by acceleration of differentiation of specific cell types. Thyroid hormones are essential for the pituitary gland to produce growth hormone and perhaps other growth factors as well. *See* ADENOHYPOPHYSIS HORMONE.

Gerard N. Burrow

Physiology

The thyroid gland synthesizes, stores, and secretes the iodine-containing hormones, thyroxine and triiodothyronine. Thyroid hormones regulate metabolic rate in warm-blooded animals and are essential for normal growth and development.

Several technical advances have contributed greatly to the understanding of thyroid physiology. One is the use of radioactive isotopes of iodine, which has enabled investigators to study the fate of iodine in the body. Another is the development of the technique of radioimmunoassay, which has afforded a means of quantifying the minute quantities of hormones and other compounds in the blood and tissues. *See* AUTORADIOGRAPHY; IODINE; RADIOIMMUNOASSAY.

Metabolism of hormones. Upon entering the blood, both thyroxine and triiodothyronine become bound to specific plasma proteins. The interaction of the thyroid hormones with these so-called transport proteins conforms to a reversible binding equilibrium in which almost all the hormone is bound and only a very small proportion is unbound or free. Only the unbound hormone is available to the tissues for metabolism and induction of hormone action. The binding of triiodothyronine to the transport proteins is much weaker than that of thyroxine. A major route of metabolism of the thyroid hormones is by way of deiodination, in which the iodine atoms—four in the case of thyroxine and three in the case of triiodothyronine—are removed progressively from the compound. In the case of thyroxine, the initial monodeiodination can occur either in the lower ring, resulting in the formation of triiodothyronine, or in the upper ring, yielding reverse triiodothyronine (**Fig. 5**). This process accounts for about 80% of the triiodothyronine produced, the remainder being secreted by the thyroid gland. It also accounts for all the reverse triiodothyronine generated. Since triiodothyronine is about three times more active than thyroxine, whereas reverse triiodothyronine is inactive, the relationship between the initial upper- or lower-ring monodeiodination of thyroxine may represent a level of regulatory control of hormone action in the tissues. Of the iodide liberated by deiodination, part is reaccumulated by the thyroid gland and utilized in hormone synthesis, and the remainder is excreted by the kidneys. Other routes of metabolism of the thyroid hormones include conjugation with glucuronate and sulfate in the liver followed by their excretion into the bile and modification of the side chain to form other derivatives, such as tetra- and triiodothyroacetic acids. *See* THYROXINE.

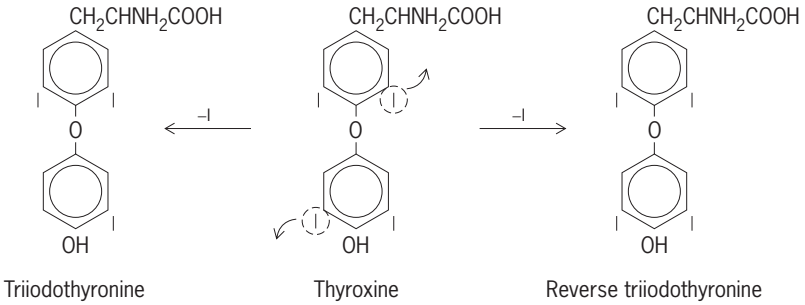

Fig. 5. Monodeiodination of thyroxine yielding either triiodothyronine or reverse triiodothyronine.

Regulation of thyroid function. The function of the thyroid gland is regulated by the thyroid-stimulating hormone of the anterior lobe of the pituitary gland. The secretion of TSH is regulated by a negative feedback mechanism in the pituitary, under hypothalamic control, which is sensitive to the concentration of unbound thyroxine or triiodothyronine in the perfusing blood. A decrease in the concentration of hormone in the blood stimulates the secretion of TSH, which tends to increase the function and size of the thyroid gland, thereby restoring to normal the concentration of thyroid hormones. Conversely, an increase in the concentration of hormone in the blood decreases the secretion of TSH and thereby tends to decrease the function and size of the gland. The threshold of feedback control of the secretion of TSH is set by the hypothalamic peptide, TSH-releasing hormone. In addition to regulation by the hypothalamus-pituitary complex, the thyroid gland possesses an intrinsic regulatory mechanism whereby the glandular content of organic iodine, present as iodinated compounds in thyroglobulin, determines in an inverse manner the activity of the iodide-trapping mechanism and the sensitivity of the gland to TSH. *See* ENDOCRINE MECHANISMS.

When the intrinsically normal thyroid gland is deprived of TSH stimulation as a result of hypothalamic or pituitary disease, atrophy of the thyroid gland and hypothyroidism ensue.

Physiological actions. Two cardinal actions of the thyroid hormones are their stimulation of the basal metabolic rate (BMR) in warm-blooded animals and their influence on the growth and development of tissues. The metabolism of carbohydrates, proteins, and fats is influenced by the thyroid gland; too much hormone in the circulation intensifies symptoms of diabetes. *See* DIABETES.

A deficiency of thyroid hormones (hypothyroidism), resulting from either a qualitative or a quantitative deficiency of thyroid tissue, produces certain distinctive changes in humans or experimental animals. The decrease in basal metabolic rate is reflected in a decrease in oxygen consumption in the whole animal and in isolated tissue preparations. Body temperature may be subnormal. Cardiac output, heart rate, and respiratory rate are decreased. There is a slowing of mental activity and depression of neuromuscular excitability. In humans, the skin is cool, dry, and coarse, and accumulation in the dermis of a mucinous material gives it a puffy appearance (myxedema). When hypothyroidism begins in early life and treatment is delayed, severe retardation of mental development and growth results (cretinism). If begun in time, treatment with thyroid hormones reverses virtually all the manifestations of the hypothyroid state.

An excess of thyroid hormones (hyperthyroidism) results from toxic goiter in humans, or its effects can be produced experimentally by the administration of large quantities of thyroid hormones. In general, the effects are the converse of those that occur in hypothyroidism. The basal metabolic rate is increased. Cardiac output, heart rate, and respiratory rate are increased, and neuromuscular excitability is enhanced. There is increased sweating, weight loss despite an increased appetite, and muscle weakness. Protrusion of the eyeballs (exophthalmos) occurs in some patients with toxic goiter; the reason for this is not known, but it is not due to the excess of thyroid hormones per se, since it cannot be produced by their administration.

Thyroid hormones do not appear to stimulate the metabolic rate in cold-blooded animals.

Kenneth A. Woeber

Biochemistry

The size of the normal thyroid is subject to more variation than other organs in the body, fluctuating with age, reproductive state, diet, and external environment. The average weight in the adult human is from 0.08 to 1.4 oz (25 to 40 g).

The thyroid has the greatest ability to trap iodine, binding one-third to one-quarter of the total amount of this element in the body. The gland takes up iodine and fixes it extremely rapidly. As the first step in the synthesis of the thyroid hormone, iodine is extracted from the circulation, and in the gland it combines with the amino acid tyrosine. The colloid that fills the follicles of the gland is a protein globulin, known as thyroglobulin. Thyroglobulin is a glycoprotein of molecular weight 669,000. It is stable in salt solutions and ranges between pH 5 and slightly alkaline pH, and dissociates into subunits under a variety of conditions. After iodination, thyroglobulin contains approximately twelve molecules of monoiodotyrosine plus diiodotyrosine, two of thyroxine, one-third of triiodothyronine, and traces of other tyrosine derivatives (Fig. 5). It cannot pass through the cell membrane into the circulation in this form, however, because of the large size of the molecule, so the thyroglobulin is broken down by an enzyme system into its constituents, one of which is thyroxine. Thyroxine is the form in which the hormone enters the circulation. The thyroxine may be degraded or changed by the tissues to still another compound, L-triiodothyronine, which is biologically more active than thyroxine itself. *See* METABOLISM; THYROID GLAND DISORDERS.

Choh Hao Li

Bibliography. R. Ekholm, *Control of the Thyroid Gland: Regulation of Its Normal Function and Growth*, 1989; M. Greer, *The Thyroid Gland*, 1990; S. H. Ingbar and L. E. Braverman (eds.), *Werner's The Thyroid: A Fundamental and Clinical Text*, 5th ed., 1986; J. Tepperman, *Metabolic and Endocrine Physiology*, 4th ed., 1980.

Thyroid gland disorders

Disorders of the thyroid gland may be classified according to anatomical and functional characteristics. Those thyroid disorders that are primarily anatomical include goiter and neoplasia; those that are primarily functional result in either hyperthyroidism or hypothyroidism.

Thyroid gland enlargement. Thyroid gland enlargement, or goiter, is the most common disorder. Its classification is based upon both the anatomy and function of the gland (see **table**). An enlarged but normally functioning thyroid gland is termed a nontoxic goiter. This condition affects hundreds of millions of people throughout the world in areas where the diet is deficient in iodine. In other areas it may be caused by subtle disorders in the biosynthesis of thyroid hormone. In both cases, there is compensatory enlargement of the gland that can be diffuse and symmetrical or can produce a multinodular goiter. When goiter is present in more than 10% of the population it is termed endemic and can represent a major public health issue. If the condition is untreated, the thyroid gland can increase 10–20 times its normal size. Dietary iodine supplementation is effective in preventing endemic goiter. Surgical removal of the enlarged gland is rarely required.

The multinodular goiter can grow independently from pituitary gland control and produce excess thyroid hormone, causing hyperthyroidism. However, hyperthyroidism is most often the result of Graves' disease. Thyroid enlargement can also be caused by Hashimoto's thyroiditis, in which the individual's immune system develops abnormal antibodies that react with proteins in the thyroid gland. This autoimmunity can make the gland enlarge or become underactive. *See* AUTOIMMUNITY.

Neoplasia. Tumors of the thyroid account for a small fraction of human neoplasms and an even smaller fraction of deaths due to cancer. The vast

Diseases of the thyroid gland and their effects on thyroid size and hormone production

Disease	Thyroid size	Hormone production
Multinodular goiter	Increased	Normal or increased
Thyroid adenoma	Normal or increased	Normal or increased
Thyroid carcinoma	Increased	Normal
Hashimoto's thyroiditis	Increased	Normal or decreased
Graves' disease	Increased	Increased
Hypothyroidism	Normal, increased, or decreased	Decreased

majority of thyroid neoplasms are follicular adenomas, which are benign; however, some thyroid neoplasms are malignant. Malignant tumors generally arise from thyroid follicular epithelium (follicular, papillary, and anaplastic carcinoma), but may also derive from parafollicular cells (medullary thyroid carcinoma) or from lymphocytes within the gland (lymphoma). Rarely, tumors arising elsewhere in the body can metastasize to the thyroid gland. Risk factors for the development of thyroid neoplasia include exposure of the thyroid to ionizing radiation (tumors of follicular epithelium), a family history of thyroid cancer (medullary carcinoma of the thyroid), and lymphocytic or Hashimoto's thyroiditis (lymphoma of the thyroid). Most thyroid tumors are detected as a solitary nodule when the thyroid is palpated. Relative to normal thyroid tissue, these nodules are usually hypofunctional and thus appear inactive on isotope scans of the thyroid gland. Aspiration (removal of fluid by suction) of these nodules for diagnostic cytologic evaluation has become standard practice. If malignant cells are detected in this manner, the tumor is surgically removed and radioative iodine is usually administered to ablate any residual thyroid tissue. The survival of patients 5 years after diagnosis is excellent (approximately 95%). *See* ONCOLOGY; RADIOACTIVE TRACER.

Hyperthyroidism. Hyperthyroidism is the clinical condition that results from excessive levels of the circulating thyroid hormones thyroxine and triiodothyronine, which are secreted by the thyroid gland. Signs and symptoms include weight loss, tachycardia (increased heart rate), heat intolerance, sweating, and tremor. Graves' disease, the most common form of hyperthyroidism, is mediated by an abnormal antibody directed to the thyroid-stimulating hormone (TSH) receptor on the surface of the thyroid cell, which stimulates secretion of thyroid hormone. Unique to Graves' disease is the associated protrusion of the eyes (exophthalmos). As noted above, single or multiple thyroid nodules can also cause hyperthyroidism when they produce excessive levels of thyroid hormone. Acute or subacute inflammation of the thyroid (thyroiditis) causes hyperthyroidism when thyroid hormone is released by the inflamed gland into the bloodstream.

In all cases, the goal of therapy is to promptly reduce thyroid hormone levels to normal. This is most rapidly accomplished by the administration of drugs that impair thyroid hormone biosynthesis and the release of preformed thyroid hormone from the gland. Other possible treatments include administration of radioactive iodine or surgical excision of the thyroid gland.

Hypothyroidism. Hypothyroidism is the clinical state that results from subnormal levels of circulating thyroid hormones. Manifestations in infancy and childhood include growth retardation and reduced intelligence; in adults, cold intolerance, dry skin, weight gain, constipation, and fatigue predominate. Individuals with hypothyroidism often have a slow pulse (bradycardia), puffy dry skin, thin hair, and delayed reflexes. In its most extreme form, hypothyroidism can lead to coma and death if untreated.

The vast majority of cases are due to thyroid gland failure as a result of chronic inflammation (Hashimoto's thyroiditis). When the thyroid begins to fail, the pituitary gland senses a reduction of thyroid hormone in the circulation and responds by secreting thyrotropin. The resultant rise in serum thyrotropin is the biochemical hallmark of primary hypothyroidism. Hypothyroidism is treated by the oral administration of thyroxine. *See* ADENOHYPOPHYSIS HORMONE; ENDOCRINE MECHANISMS; PITUITARY GLAND; THYROID GLAND; THYROID HORMONES. Leslie J. DeGroot; David A. Ehrmann

Bibliography. R. H. Cobin and D. K. Sirota (eds.), *Malignant Tumors of the Thyroid: Clinical Concepts and Controversies*, 1992; M. Greer, *The Thyroid Gland*, 1990; M. H. Wheeler and J. H. Lazarus, *Thyroid Disease*, 1993.

Thyroid hormones

Any of the chemical messengers produced by the thyroid gland. For many years it was generally believed that the thyroid gland produced only two closely related hormones, thyroxine and triiodothyronine. However, in 1961 a new hormonal principle, calcitonin, was discovered by Harold Copp. He originally thought that this substance was secreted by the parathyroid glands, but in 1963 it was established that the source of this hormone in mammals was the thyroid gland, and it was renamed thyrocalcitonin. In 1967 it was isolated in pure form and in 1968, synthesized. It is a polypeptide containing a single chain of 33 amino acids, in contrast to thyroxine and triiodothyronine which are iodinated thyronines. *See* HORMONE.

Thyroid gland. The thyroid gland is in reality two separate endocrine organs composed of two different cell types: acinar cells, manufacturing both thyroxine and triiodothyronine; and interstitial, parenchymal, or C cells, producing thyrocalcitonin. These two cell types differ in location, embryologic origin, histochemical characteristics, and control of the secretion of their respective hormones. The acinar cells are derived embryologically from the fourth branchial pouch; the parenchymal cells, from the fifth branchial pouch. In higher animals the parenchymal cells are enveloped by the lateral lobes

of the thyroid during subsequent development; in birds, reptiles, fish, and amphibia, they persist as separate organs, the ultimobranchial bodies which, rather than the thyroid gland, are the source of thyrocalcitonin. *See* THYROID GLAND.

Biochemistry and biosynthesis. Thyroxine (abbreviated T4) and triiodothyronine (T3) have very similar chemical structures (shown below) and are iodine-

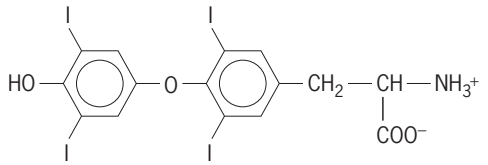

Thyroxine

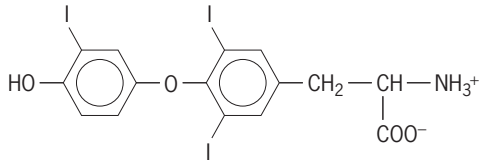

Triiodothyronine

containing derivatives of the amino acid tyrosine. [Thyrocalcitonin, with its entirely different function from T3 and T4, is discussed separately. *See* THYROCALCITONIN.] Because these two thyroid hormones are the only important biological compounds in higher animals that contain iodine and because radioisotopes of iodine are readily available to researchers, the biosynthesis, secretion, and distribution throughout the body of thyroid hormones has been extensively investigated and is well understood. *See* IODINE.

Ordinarily, cells manufacture chemical compounds of this small size using biochemical pathways. In a biochemical pathway, common metabolic intermediates (such as amino acids) are chemically modified in a series of enzyme-catalyzed steps until the desired compound is produced. Thyroxine and triiodothyronine, by contrast, are synthesized in a completely different and unusual manner. The acinar cells of the thyroid gland produce a large protein called thryoglobin, which contains approximately 600 amino acid residues. Iodine is chemically attached to approximately 30 of this protein's tyrosine residues, forming 2,5-diiodotyrosine residues. These iodinated tyrosine residues are then chemically linked to one another to form thyroxine-like and triiodothyronine-like residues within the thryoglobin protein. Lastly, the modified thyroglobin protein is broken down into many tiny fragments, most of which are degraded. Among the resulting fragments are the desired products, thyroxine and triiodothyronine. Five or six active hormone molecules are derived from a single thyroglobin protein.

The biosynthesis of these hormones is tightly regulated. The thyroid-stimulating hormone (TSH), released by the pituitary gland, controls how much thyroxine and triiodothyronine are produced and released into the bloodstream. Furthermore, the amount of iodine in one's diet influences the rate at which these hormones can be produced and released. *See* ADENOHYPOPHYSIS HORMONE; PITUITARY GLAND.

Mode of action. Thyroxine and triiodothryonine are nonpolar molecules that are not readily soluble in the aqueous environments of the bloodstream and cellular interiors. Consequently, upon release into the bloodstream from the thyroid gland, these hormones bind to a protein found in the blood called thyroxine-binding protein. The thyroxine-binding protein transports these hormones through the bloodstream to the hormone's target cells in the body. Another protein in the blood, serum albumin, can also transport these hormone molecules. *See* ALBUMIN.

Thyroid hormones are taken up by nearly all cells of the body, but the rate of uptake differs from cell type to cell type. Liver cells take up these hormones very rapidly, but many cells of the brain take up these hormones at nearly undetectable levels. Because of their nonpolar character, these hormones readily pass through the membranes of their target cells, where they bind to another thyroid hormone–binding protein in the cell's cytoplasm. This protein–hormone complex does not itself enter the cell's nucleus; instead it is believed to function in maintaining a reservoir of thryroid hormone within the cell.

In order to act upon the target cell, the thyroid hormone molecules leave the cytoplasmic binding protein and pass directly through the nuclear membrane. Once inside the nucleus, these hormone molecules bind to the thyroid hormone receptor protein, which is structurally related to the steroid hormone receptor proteins. The thyroid hormone receptor protein is bound to the chromosomal DNA at defined positions. When this protein binds to the hormone, it activates transcription of nearby genes that begins the production of a variety of metabolic enzymes.

It appears that the majority of the effects of thyroid hormones on cells comes from their ability to enter the cell's nucleus and stimulate the transcription of certain genes. However, there is evidence that these hormones also enter the mitochondria, the energy-producing organelles of the cell, and directly stimulate oxidative phosphorylation—a process through which adenosine triphosphate (ATP) is formed.

Physiological activities and abnormalities. The maintenance of a normal level of thyroxine is critically important for normal growth and development as well as for proper bodily function in the adult. The hormone is also critically important in amphibian metamorphosis. Its absence leads to delayed or arrested development. It is one of the few hormones with general effects upon all tissues. Its lack leads to a decrease in the general metabolism of all cells, characteristically measured as a decrease in nucleic acid and protein synthesis, and a slowing down of all major metabolic processes. These widespread and profound changes have been the subjects of intense investigative interest. *See* THYROXINE.

Two abnormalities have been described that involve either too little or too much thyroid hormone in the bloodstream. Hypothyroidism is an abnormality in which the levels of thyroid hormone

are too low. Individuals with hypothyroidism have a slowed metabolism and are generally lethargic and obese. Hyperthyroidism is an abnormality in which the levels of thyroid hormone are too high. Individuals with hyperthyroidism show the opposite effects, namely increased oxygen consumption, a speeding up of the metabolic processes, and hyperactivity. *See* THYROID GLAND DISORDERS.

M. Todd Washington; Howard Rasmussen

Bibliography. W. Green (ed.), *The Thyroid*, 1986, reprint 2001; J. E. Griffin and S. R. Ojeda (eds.), *Textbook of Endocrine Physiology*, 4th ed., 2000; D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry*, 4th ed., 2004; D. Voet and J. G. Viet, *Biochemistry*, 3d ed., 2004.

Thysanoptera

An order of small, slender insects, commonly called thrips, having exopterygote development, sucking mouthparts, and highly modified wings. The order is a relatively small one, but individuals are often very numerous in favorable environments.

The mouthparts are conical and used for scraping, piercing, and sucking; flowers, leaves, and even twigs are attacked, resulting in severe injury to the epidermal cells of the plant. Vegetables and fruit trees are especially subject to damage by these insects. A very few species of thrips are predacious, feeding on mites and aphids.

The wings of these insects are exceptionally narrow, with few or no veins, and are bordered by long hairs. The tarsi terminate in an inflatable membranous bladder, which has remarkable adhesive properties.

The eggs of thrips are laid on the surface of twigs (suborder Tubulifera) or in small cuts made by the ovipositor (suborder Terebrantia). There are usually four nymphal stages, the last of these being quiescent and pupalike. There are from one to several generations produced in a single year. *See* INSECTA.

Frank M. Carpenter

Bibliography. N. F. Johnson and C. A. Triplehorn (eds.), *Borror and DeLong's Introduction to the Study of Insects*, 7th ed., 2004.

Thysanura (Zygentoma)

An order of wingless insects with soft, fusiform bodies 0.12–0.8 in. (3–20 mm) long, often covered with flat scales forming diverse patterns. The mouthparts are free with dicondylous mandibles used for scraping and chewing. Antennae are long and threadlike (filiform) with muscles present only in the scape and pedicel. Visual organs may be a cluster of simple ommatidia (cylinder-shaped units of a compound eye) or lacking altogether. The abdomen terminates in three "tails": a pair of lateral cerci and a median caudal filament (telson). Females have well-developed ovipositors; males have a penis and often one or rarely two pairs of parameres.

Classification. The order Thysanura (also called Zygentoma) includes about 400 described species which most taxonomists group into three families. The Lepidotrichidae, known as forest silverfish, are well represented in Oligocene fossils but were thought to be extinct until 1959, when living specimens were discovered in northwestern California. The single extant species lives in decaying bark and rotten wood. Members of the family Nicoletiidae are typically found underground in caves, mammal burrows, or the nests of ants or termites. Some species mimic ants and steal their food. The family Lepismatidae is worldwide in distribution. It encompasses more than 300 species of fast-running insects that feed as scavengers on algae, lichens, starchy vegetable matter, fungal hyphae (mycelium), and woodland debris. This family includes all of the common species found in human dwellings, notably the silverfish, *Lepisma saccharina*, and the firebrat, *Thermobia domestica*. Silverfish (sometimes known as fishmoths) are silvery gray, active at night, and often regarded as nuisance pests. Firebrats are usually tan or brown. These heat-tolerant insects prefer warm domestic habitats like kitchens, attics, and boiler rooms. They have been reported as serious pests in bakeries. Both species can cause extensive damage to household goods by feeding on stored food, cardboard packaging, wallpaper paste, book bindings, and the starch sizing of textiles.

Reproduction and development. Mating behavior of males and females involves external (indirect) fertilization. Males package sperm in desiccation-resistant spermatophores which they attach to the substrate. Insemination occurs when a female finds a spermatophore and draws it into her genital opening. Silverfish have an elaborate courtship ritual to ensure efficient exchange of sperm. The male spins a silken thread between the substrate and a vertical object. He deposits a spermatophore beneath this thread and then coaxes a female to walk under it. When her cerci contact the silk, she picks up the spermatophore with her genital opening. Sperm are released into her reproductive system, and then she ejects the empty spermatophore and eats it.

Thysanurans may be long-lived. A 3-year life-span is probably typical, and firebrats have been known to live up to 6 years. Development is ametabolous: other than an increase in size, there is no change in physical appearance from immatures (called young) to adults. These insects molt frequently, even as adults, and may complete more than 40 instars (stages between molts). This unusual adaptation could help reduce the risk of infection by parasitic fungi. *See* APTERYGOTA; INSECTA.

John Meyer

Bibliography. L. M. Ferguson, Insecta: Microcoryphia and Thysanura, pp. 935–949 in D. Dindal (ed.), *Soil Biology Guide*, Wiley, New York, 1990; N. F. Johnson and C. A. Triplehorn, (eds.), *Borror and DeLong's Introduction to the Study of Insects*, 7th ed., 2004; J. A. Wallwork, *Ecology of Soil Animals*, McGraw-Hill, New York, 1970; P. Wygodzinsky, A review of the silverfish (Lepismatidae, Thysanura) of the United States and Caribbean area, *Amer. Mus. Novit.*, 2481:1–26, 1972.

Tick paralysis

A loss of muscle function or sensation in humans or certain animals following the prolonged feeding of female ticks. Paralysis, of Landry's type, usually begins in the legs and spreads upward to involve the arms and other parts of the body. Evidence suggests that paralysis is due to a neurotoxin formed by the feeding ticks rather than the result of infection with microorganisms. *See* IXODIDES; TOXIN.

Present information indicates that (1) in endemic areas, resistance to tick paralysis is found in older animals as well as in certain animal species, (2) some animals which have recovered are not immune, (3) only occasional female, but not male, ticks may induce the disease under favorable host or environmental conditions, the specific requirements being as yet unknown, (4) paralysis of adult persons and domestic animals as large as a 1000-lb (450-kg) bull has resulted from only one partially engorged tick usually but not necessarily attached about the head or upper spine, (5) death may ensue if respiratory centers are reached by the ascending paralysis before the offending tick completes feeding or is removed, and (6) in the United States recovery is prompt, a matter of hours, when the ticks are removed. It is highly important, therefore, that the disease not be confused with infantile paralysis, that it be properly and promptly diagnosed, and that search for the tick be instituted immediately.

The disease has been reported in North America, Australia, South Africa, and occasionally in Europe, and is caused by appropriate species of indigenous ticks. In Australia, *Ixodes holocyclus* causes frequent cases in dogs, and occasionally in humans, and paralysis has been known to progress even after removal of ticks; serum from recovered dogs has been shown to have some curative properties and was at one time produced for treatment. *Ixodes cubicundus* is associated with the disease in South Africa.

Since 1903 over a hundred human cases and many outbreaks in cattle, sheep, and even domesticated bison have been recorded in the northwestern United States and southern British Columbia, due to attacks by *Dermacentor andersoni*. Several human fatalities, one in Idaho in June 1958, and some losses of stock have occurred when deticking was delayed. April to June are the months of most prevalence. Incidence is highest in children of 1–5 years of age, with more than twice as many girls affected as boys, presumably because their longer hair conceals feeding ticks. However, the sex ratio is reversed among the fewer cases in adults because of difference in exposure. Young to yearling stock are most prone to the disease in sporadic years for reasons still unknown.

The related American dog tick (*D. variabilis*) has paralyzed persons and dogs in the southeastern United States, and a few cases have been associated with the lone star tick (*Amblyomma americanum*).

The female tick requires 4–5 days of feeding before initial symptoms appear, and the disease progresses rapidly during the next 2–4 days. Experimental, fatal paralysis has been produced by *D. andersoni* females in woodchucks, ground squirrels, wood rats, hamsters, guinea pigs, dogs, and lambs. Signs of the disease occur within a few hours of transfer of partially fed females, but not males, to fresh animals. Isolation of the toxic principle has been claimed only by G. H. Kaire from the Australian tick (*I. holocyclus*) by chromatographic methods. Cornelius B. Philip

Bibliography. B. Fifaz, T. Petney, and I. Horak (eds.), *Tick Vector Biology: Medical and Veternarian Aspects*, 1993.

Tidal bore

A part of a tidal rise in a river which is so rapid that water advances as a wall often several feet high. The phenomenon is favored by a substantial tidal range and a channel which shoals and narrows rapidly upstream, but the conditions are so critical that it is not common. A shoaling channel steepens the tidal curve. If the curve becomes vertical or nearly so, a bore results (**Fig. 1**). A narrowing channel increases the tidal range. Since the tidal range is greatest at spring tides, some rivers exhibit bores only then. Although the bore is a very striking feature, Fig. 1 shows that the tide continues to rise after the passage of the bore and that this subsequent rise may be greater. Bores may be eliminated by changing channel depth or shape. *See* RIVER TIDES; TIDE.

In North America three bores have been observed: at the head of the Bay of Fundy (**Fig. 2**), at the head of the Gulf of California, and at the head of Cook Inlet, Alaska. The largest known bore occurs in the Tsientang Kiang, China. At spring tides this bore is a

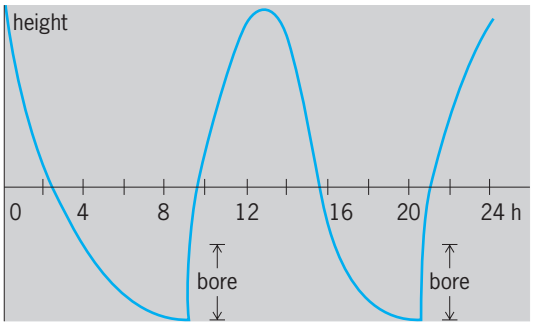

Fig. 1. Tidal curve of a river with a tidal bore.

Fig. 2. Tidal bore of the Petitcodiac River, Bay of Fundy, New Brunswick, Canada. Rise of water is about 4 ft (1. 2 m). (*New Brunswick Travel Bureau*)

wall of water 15 ft (4.5 m) high moving upstream at 25 ft/s (7.5 m/s). *See* OPEN CHANNEL. Blair Kinsman

Tidal datum

A reference elevation of the sea surface from which vertical measurements are made, such as depths of the ocean and heights of the land. The intersection of the elevation of a tidal datum with the sloping shore forms a line used as a horizontal boundary. In turn, this line is also a reference from which horizontal measurements are made for the construction of additional coastal and marine boundaries.

Since the sea surface moves up and down from infinitely small amounts to hundreds of feet over periods of less than a second to millions of years, it is necessary to stop the vertical motion in order to have a practical reference. This is accomplished by hydraulic filtering, numerical averaging, and segment definition of the record obtained from a tide gage affixed to the adjacent shore. Waves of periods up through wind waves are effectively damped by a restricting hole in the measurement well. Recorded hourly heights are averaged to determine the mean of the higher (or only) high tide of each tidal day (24.84 h), all the high tides, all the hourly heights, all the low tides, and the lower (or only) low tide. The length of the averaging segment is a specific 19 year, which averages all the tidal cycles through the regression of the Moon's nodes and the metonic cycle. [The metonic cycle is a time period of 235 lunar months (19 years); after this period the phases of the Moon occur on the same days of the same months.] But most of all, the 19-year segment is meaningful in terms of measurement capability, averaging meteorological events, and for engineering and legal interests. However, the 19-year segment must be specified and updated because of sea-level changes occurring over decades. The present sea-level epoch is 1960 through 1978. It has been updated about every 20 years.

The legally recognized tidal datums of the United States are Mean Higher High Water (MHHW), Mean High Water (MHW), and Mean Lower Low Water (MLLW). Depths of the ocean, especially in harbors and coastal regions, are measured from the tidal datum of Mean Lower Low Water, called Chart Datum, and printed on nautical charts. This lowest of the ordinary tidal datums was chosen as a safety factor for early mariners who might not have been in possession of tide predictions. Later, as predictions became easily available, the predictions at any time could be added to the printed depths for harbor access and safety.

Mean Sea Level (MSL), obtained from averaging all the hourly heights, is used to monitor and study apparent secular trends in sea level.

Over the decades, land elevations have been based on a geodetic network supported by the tidal datums of several groupings of tide stations. They are now computed within the network of the North American (Canada, United States, Mexico) Vertical Datum of 1988. This network is attached to the tidal datum at only one tide station, Father Point, Quebec, Canada.

Tidal datums are legal entities. Because of variations in gravity, semistationary meteorological conditions, semipermanent ocean currents, changes in tidal characteristics, ocean density differences, and so forth, the sea surface (at any datum elevation) does not conform to a mathematically defined spheroid. Therefore, a series of tide stations along the marine and Great Lakes coastlines of the United States is necessary in order to provide vertical continuity over horizontal distances. *See* GEODESY; TIDE.

Steacy D. Hicks

Bibliography. G. W. Groves, Numerical filters for discrimination against tidal periodicities, *Trans. Amer. Geophys. Union*, 36(6):1073–1084, 1955; S. D. Hicks, *Tide and Current Glossary*, NOAA, 1989; E. Lisitzin and J. Pattullo, The principal factors influencing the seasonal oscillations of sea level, *J. Geophys. Res.*, 66:845–853, 1961; H. A. Marmer, *Tidal Datum Planes*, 5th ed., USCGS Spec. Publ. 135, 1951; J. Pattullo et al., The seasonal oscillations in sea level, *J. Mar. Res.*, 14:88–156, 1955; J. R. Rossiter, Note on methods of determining monthly and annual values of mean water level, *Int. Hydrogr. Rev.*, May 1958.

Tidal power

Tidal-electric power is obtained by using the recurring rise and fall of coastal waters in response to the gravitational forces of the Sun and the Moon. Marginal marine basins are enclosed with dams, making it possible to create differences in the water level between the ocean and the basins. The oscillatory flow of water filling or emptying the basins is used to drive hydraulic turbines that propel electric generators.

Electric power could be developed in the world's coastal regions having tides of sufficient range, although even if fully developed this would amount to only a small percentage of the world's potential water (hydroelectric) power. Nevertheless, tidal-electric power could become locally important, particularly because it produces no air or thermal pollution, consumes no exhaustible resource, and produces relatively minor impacts on the environment.

The use of ocean tides for power purposes dates back to the tidal mills in Europe during the Middle Ages and to those in America during colonial times. The first tidal development producing electric power in operation, the Rance development in northwestern France, was completed in 1967. It has an installed capacity of 240,000 kW in 24 units and is capable of producing about 500×10^6 kWh annually.

Tidal range is measured as the difference in level between the successive high and low waters. Although there are variations at certain locations in the intervals between successive high tides, at most places the tides reach the highest levels at intervals of about 12 h 25 min. The tidal ranges vary from

day to day. The highest tides, known as spring tides, occur twice monthly near the time of the new moon and the full moon when the Sun and Moon are in line with the Earth. The lowest tides, known as neap tides, occur midway between the spring tides when the Sun and Moon are at right angles with the Earth. The highest spring tides occur near the time of the equinoxes in the spring and fall of the year. Except for variations caused by meteorological changes, the tides are predictable and follow similar patterns from year to year.

Large tidal ranges occur when the oscillation of the ocean tides is amplified by relatively shallow bays, inlets, or estuaries. There are limited locations where the tidal ranges are sufficiently large to be considered favorable for power development. The largest tidal ranges in the world, reaching a maximum of over 50 ft (15 m), are said to occur in the Bay of Fundy in Canada. Other locations with large maximum tidal ranges are the Severn Estuary in Britain, 45 ft (14 m); the Rance Estuary in France, 40 ft (12 m); Cook Inlet in Alaska, 33 ft (10 m); and the Gulf of California in Mexico, 30 ft (9 m). Large tidal ranges also occur at locations in Argentina, India, Korea, Australia, and on the northern coast of Russia. *See* ELECTRIC POWER GENERATION; TIDE.

George G. Adkins

Bibliography. American Society of Civil Engineers, *Civil Engineering Guidelines for Planning and Designing Hydroelectric Developments*, vol. 5: *Pumped Storage and Tidal Power*, 1989; American Society of Civil Engineers, *Developments in Tidal Energy*, 1990; American Society of Civil Engineers, *Tidal Power: Symposium Proceedings*, 1987; G. Boyle, *Renewable Energy*, 2d ed., 2004.

Tidalites

Sediments deposited by tidal processes. Until recently, "tidalites" referred to sediments deposited by tidal processes in both the intertidal zone (between normal low- and high-tide levels) and shallow,

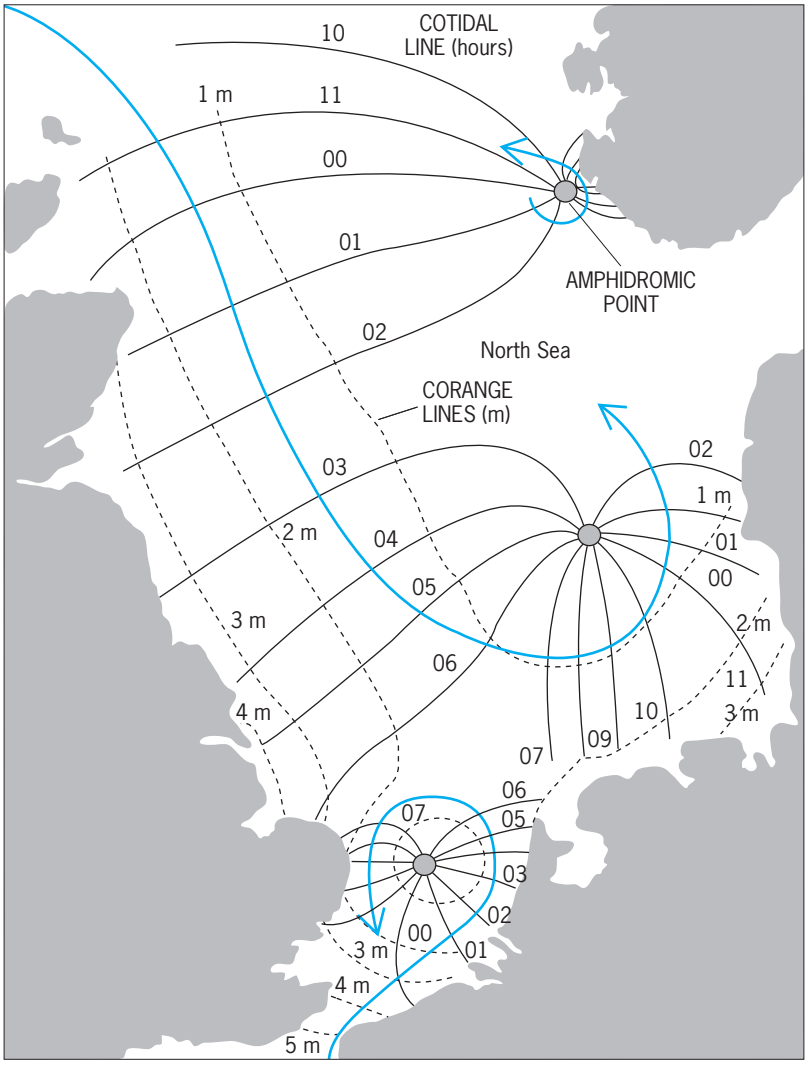

Fig. 1. North Sea amphidromic tidal system. Corange lines indicate equal tidal range. Cotidal lines show times of high water. Arrows show rotation directions of the tidal waves. (*Modified from R. W. Dalrymple, Tidal Depositional Systems, in R. G. Walker and N. P. James, eds., Facies Models Response to Sea Level Changes, pp. 195–218, Geological Association of Canada, 1992*)

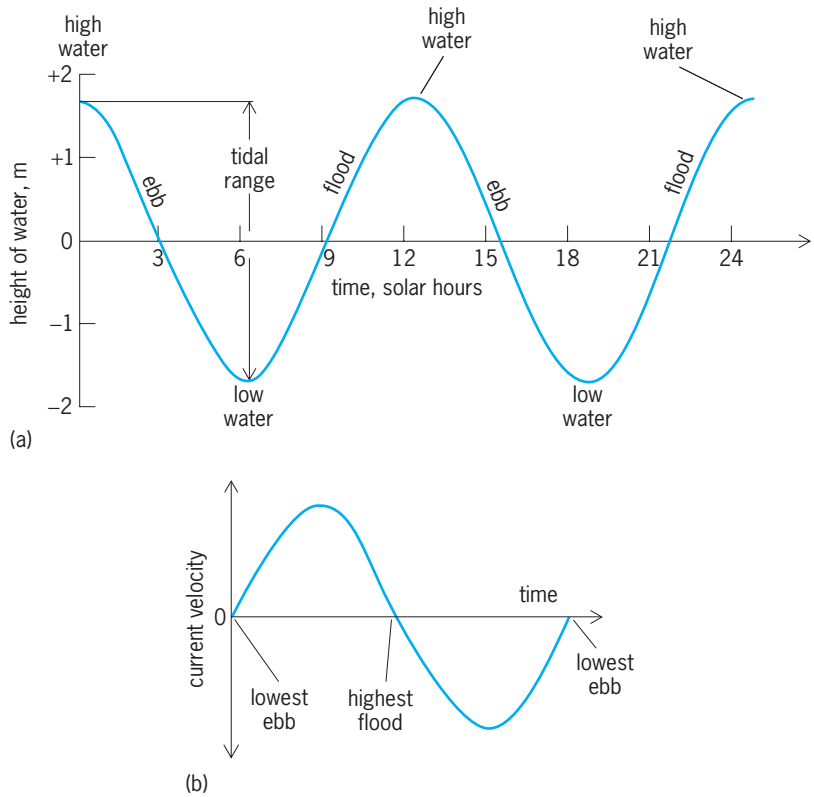

Fig. 2. Flood–ebb cycle. Idealized (*a*) semidiurnal tidal cycle and (*b*) time–current velocity curve. (*Modified from S. D. Nio and C. S. Yang, Recognition of tidally influenced facies and environments, Short Course Note Ser. 1, International Geoservices BV, Leiderdorp, Netherlands, 1989*)

subtidal (permanently submerged), tide-dominated environments less than 200 m (660 ft) deep. Tidalites are now known also to occur within supratidal environments (above normal high tide and flooded only during storms or very high spring tides) and submarine canyons at depths much greater than 200 m. Common usage has drifted toward describing tidalites as ripple- and dune-scale features rather than more composite deposits such as large linear sand ridges of tidal origin present on continental shelves or point bars associated with migrating tidal channels. Both of these larger-scale features, however, would be composed of tidalites.

Recognition criteria. By identifying tidalites in either the modern or the ancient geological record, geologists are implying that they know that the sediments were deposited by tidal processes rather than by storms or waves. Tidalites are not always easy to identify with certainty, especially in the rock record. In order to do so, it is necessary to understand the basic tidal cycles that can influence sedimentation.

Tidal theory. Tides are generated by the combined gravitational forces of the Moon and Sun on the Earth's oceans. Some sources are misleading in suggesting that the tidal forces from the Moon and Sun, in combination with centrifugal forces associated with the spin of the Earth, produce oceanic bulges on opposite sides of the Earth. While it is true that the combined gravitational forces of the Moon and, to a lesser extent, the Sun produce tides on the Earth, the Earth does not spin through two oceanic bulges that form on opposite sides of the Earth. This conceptual model has little bearing on real-world tides. Rather, water within each of the Earth's ocean basins is forced to rotate as discrete waves about a series of fixed (amphidromic) points (**Fig. 1**). For a fixed point along an ocean coastline, a tidal system is referred to as diurnal if it experiences the passing of the resultant tidal wave once every 24 h 50 min. The tidal system is semidiurnal if the resultant tidal wave passes the fixed point twice during the same time.

In the open ocean, the motion of a tidal wave is largely expressed as a vertical movement of water masses. In shallow basins along the coast, water movements are more horizontal, with tides moving in and out of estuaries and embayments, resulting in a change in water level as the tidal wave passes. The daily or semidaily rise in tides is called the flood tide, and the fall is referred to as the ebb tide (**Fig. 2***a*). Tidal currents are maximized between flood and ebb tides and minimized at highest flood or lowest ebb tides (Fig. 2b). The difference between the high tide and the low tide is called the tidal range.

The intensity or height of the daily or twice-daily tides can vary in a number of ways. Cyclic semimonthly changes in daily tidal heights associated with neap–spring tidal cycles are the most pronounced of these. Spring tides occur twice a month when the tidal range is greatest, and neap tides occur twice a month when tidal range is least. Neap–spring

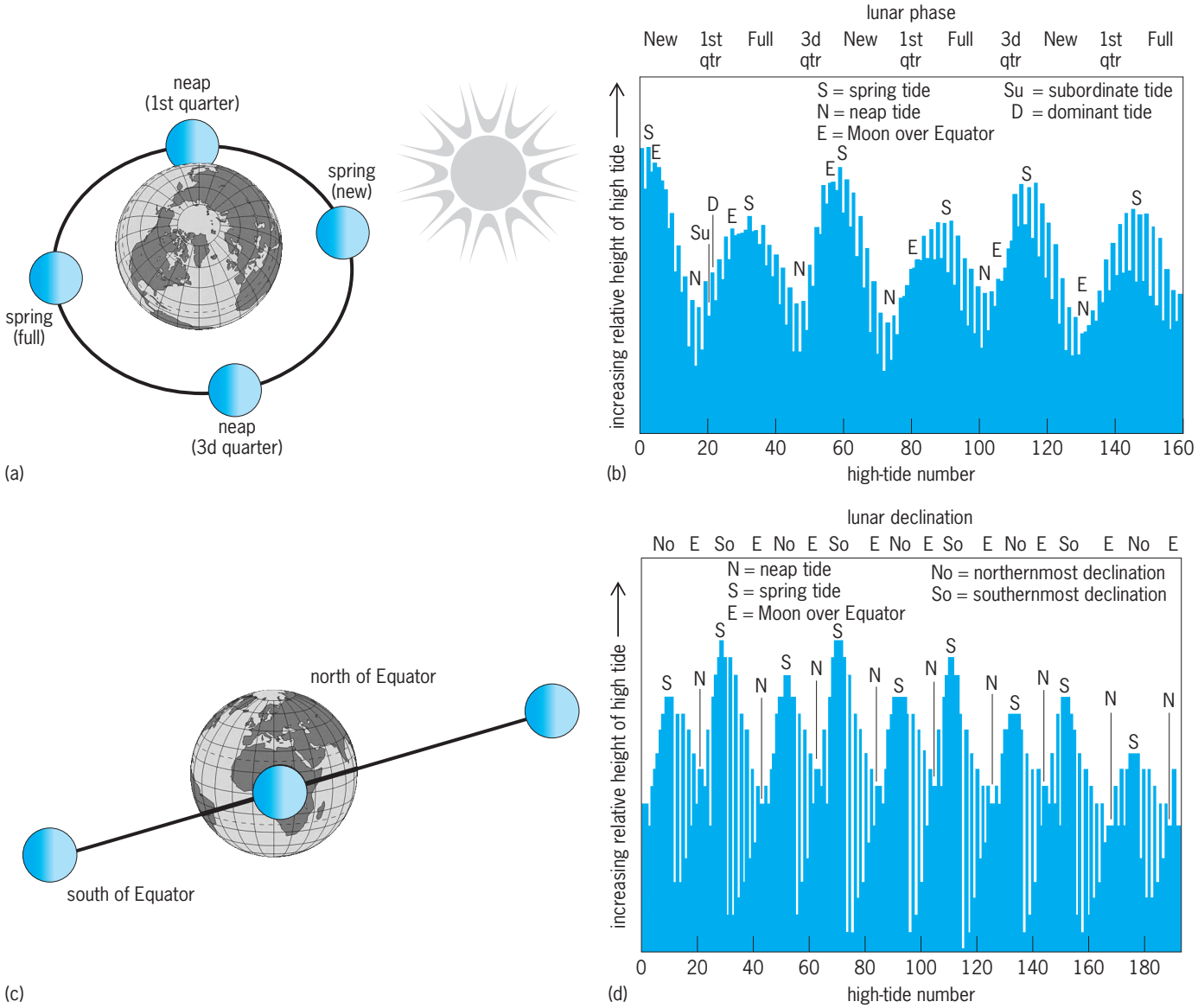

Fig. 3. Idealized models of origin of neap–spring tidal cycles: (*a*) Synodic month. (*b*) A segment of the 1991 predicted high tides from Kwajalein Atoll, Pacific. (*c*) Tropical month. (*d*) A segment of the 1994 predicted high tides from Barito River estuary, Borneo. (*Modified from E. P. Kvale, K. H. Sowder, and B. T. Hill, Modern and ancient tides: Poster and explanatory notes, Society for Sedimentary Geology, Tulsa, OK, and Indiana Geological Survey, Bloomington, IN, 1998*)

cycles can be generated in two ways. The familiar neap–spring cycle is related to the phase changes of the Moon. Spring tides occur every 14.76 days when the Earth, Moon, and Sun are nearly aligned at new or full moon (**Fig. 3***a*). Neap tides occur when the Sun and Moon are aligned at right angles from the Earth at first- and third-quarter phases of the Moon. The result is that spring tides are higher than neap tides (Fig. 3*b*). The time from new moon to new moon is called the synodic month, which has a modern period of 29.53 days. This type of neap–spring cycle is referred to as synodically driven, and it dominates the coastlines of western Europe and the eastern coastline of North America.

A second type of neap–spring cycle is less well known but no less common, and is related to the orbit of the Moon around the Earth. The Moon's orbital plane is inclined relative to the Earth's equatorial plane. The period of the variation in lunar declination relative to the Earth's Equator is called the tropical month, and is the time the Moon takes to complete one orbit, moving from its maximum northerly declination to its maximum southerly declination and return (Fig. 3*c*). In this type of neap–spring cycle, the tidal force depends on the position of the Moon relative to the Earth's Equator. The tide-raising force at a given location is greater when the Moon is at its maximum declination every 13.66 days. These periods correspond to the generation of spring tides (Fig. 3*d*). The neap tides occur when the Moon is over the Earth's Equator. The current length of the tropical month is 27.32 days, and neap–spring cycles in phase with the tropical month are referred to as tropically driven. These types of neap-spring cycles dominate coastlines in the Gulf of Mexico and many areas in the Pacific.

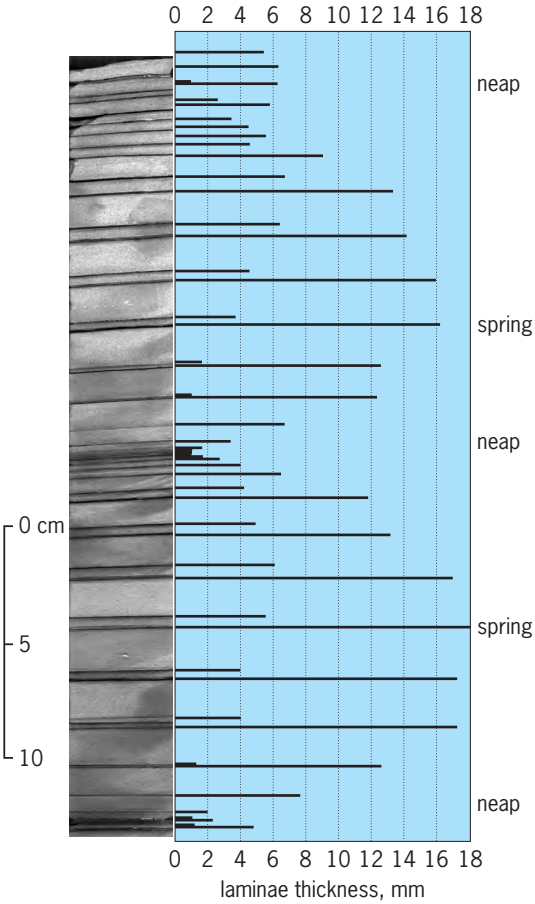

Fig. 4. The core shows small-scale tidalites from the Hindostan Whetstone beds, Mansfield Formation, Indiana. The chart shows thicknesses of layers as measured between dark clay-rich bands. The interval shows approximately one synodic month of deposition.

Besides generating neap–spring cycles in many parts of the world, the changing position of the Moon relative to the Earth's Equator through the tropical month causes the diurnal inequality of the tides in semidiurnal tidal systems. In tidal systems that experience two high tides and two low tides per day, the tropical monthly cycle results in the morning high tide being greater or lesser than the evening high tide. The diurnal inequality is reduced to zero when the Moon is over the Equator, resulting in the morning tide and the evening tide being of equal magnitude (Fig. 3*b* and *d*).

Other tidal cycles besides those mentioned above can influence sedimentation and have been documented in the geologic record. These include monthly, semiannual, and multiyear tidal cycles. *See* EARTH ROTATION AND ORBITAL MOTION; MOON; TIDE.

Examples of tidalites. How the various tidal cycles manifest themselves in the geologic record and how geologists can identify their influence on sedimentation has been studied for nearly 75 years. To recognize tidalites in the geologic record, geologists look for evidence of one or more of the following:

1. Sediment deposited by reversing currents (that is, flood–ebb cycles).

2. A stacked sequence of sediments that show a recurring change from sediments transported (and deposited) by currents at maximum current velocity to sediments deposited from suspension at minimum current velocity (Fig. 2*b*).

3. Stacked packages of sediments in which each package shows evidence of subaerial exposure superimposed on sediments deposited in subaqueous settings (sediments transported and deposited during flood tides and exposed during low ebb tide).

4. A sequence of sediment packages in which the thickness or accretion of successive packages of sediments varies in a systematic way, suggesting diurnal, semidiurnal, and/or neap–spring tidal cycles.

An example of a small-scale tidalite can be found in the Mansfield Formation (Pennsylvanian Period) in Orange County, Indiana (**Fig. 4**). The sample shown is from a rock core taken through this interval. The lighter-colored layers are siltstone, and the thin dark layers are finer-grained claystone. The regular and repeating change in deposition from siltstone to claystone indicates systematic current velocity fluctuations related to the tidal cycle over a 12-h period (see item 2 above). The thick–thin pairing of the lighter bands of siltstone suggests the influence of the semidiurnal inequality over a 24-h period (see item 4). In addition, the regular and systematic overall thickening and thinning of the siltstone layers, as shown in the bar chart next to the core in Fig. 4, suggests that neap–spring tidal cycles controlled the thicknesses of the silt layers. The higher spring tides resulted in thicker accumulations of silt than the lower neap tides. *See* PENNSYLVANIAN; SEDIMENTARY ROCKS.

An example of a large-scale tidalite can be found in the Jurassic Sundance Formation of northern Wyoming (**Fig. 5**). This tidalite is the remnant of a migrating subtidal dune or sandwave. The preserved inclined beds of the avalanche face (I) indicate the migration direction of the dune from right to left (Fig. 5*a*). The evidence for tidal influence, however, lies within the inclined, less resistant, and more recessed lighter-colored bands (examples marked by arrows in Fig. 5*a*). In this interval (Fig. 5*b*), one sees evidence of (1) cessation of dune migration and a reversal of current direction from flood tide to ebb tide with small ripples migrating up the avalanche face (II); (2) a mud drape (III) resulting from fine-grained materials settling out of suspension when the current velocities reached zero as the tide reversed (Fig. 2); and (3) a reactivation of the migrating dune above the mud drape (IV) as current velocity increased during the next flood tide. The right-to-left migration of the dune was also controlled by the neap–spring cycle, with greater migration (interval between lighter-colored bands) occurring during spring tides and lesser migration during neap tides (Fig. 5*a*). In the example shown, the neap tide deposits are centered on line N. *See* JURASSIC.

Geologic record. Deposits of tidalites are known from every geologic period from the modern back into the Precambrian and from depositional environments with water chemistries ranging from fresh to hypersaline. Studies of tidalites are important

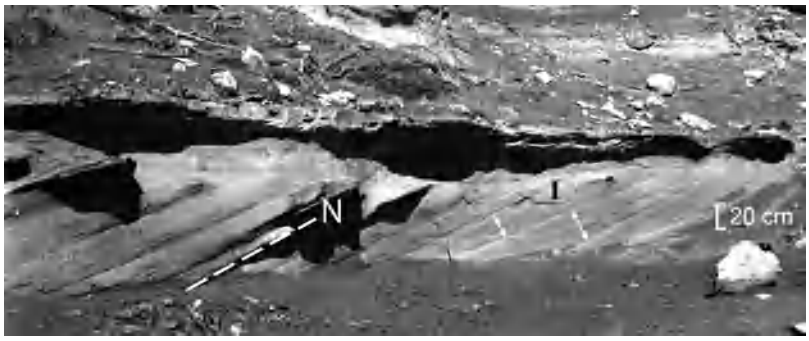

(a)

(b)

Fig. 5. Photographs from Sundance Formation of northern Wyoming. (*a*) Example of large-scale tidalites. (*b*) Closeup of inclined light-colored band showing evidence of current reversals.

because geologists have used these features not only to interpret the original depositional settings of the deposits but also to calculate ancient Earth–Moon distances, interpret paleoclimates existent during deposition, and calculate sedimentation rates. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; GEOLOGIC TIME SCALE; MARINE SEDIMENTS; SEDIMENTOLOGY.

Erik P. Kvale

Bibliography. C. Alexander, R. Davis, and V. Henry (eds.), *Tidalites: Processes and Products*, Geological Society Publishing, 1998; D. E. Cartwright, *Tides: A Scientific History*, Cambridge University Press, 1998; G. deV. Klein, A sedimentary model for determining paleotidal range, *Geol. Soc. Amer. Bull.*, 82:2585–2592, 1971; G. deV. Klein, Determination of paleotidal range in clastic sedimentary rocks, *XXIV International Geological Congress*, 6:397–405, 1972; D. T. Pugh, *Tides, Surges and Mean Sea Level*, Wiley, 1987; H. G. Reading (ed.), *Sedimentary Environments: Processes, Facies and Stratigraphy*, 3d ed., Blackwell Science, Cambridge, MA, 1996.

Tide

Stresses exerted in a body by the gravitational action of another, and related phenomena resulting from these stresses. Every body in the universe raises tides, to some extent, on every other. This article deals only with tides on the Earth, since these are fundamentally the same as tides on all bodies. Sometimes variations of sea level, whatever their origin, are referred to as tides.

Introduction. The tide-generating forces arise from the gravitational action of Sun and Moon, the effect of the Moon being about twice as effective as that of the Sun in producing tides. The tidal effects of all other bodies on the Earth are negligible. The tidal forces act to generate stresses in all parts of the Earth and give rise to relative movements of the matter of the solid Earth, ocean, and atmosphere. The Earth's rotation gives these movements an alternating character having principal periodicities of 12.42 and 12.00 h, corresponding to half the mean lunar and solar day, respectively.

In the ocean, the tidal forces act to generate alternating tidal currents and displacements of the sea surface. These phenomena are important to shipping and have been studied extensively. The main object of tidal studies has been to predict the tidal elevation or current at a given seaport or other place in the ocean at any given time.

The prediction problem may be attacked in two ways. Since the relative motions of Earth, Moon, and Sun are known precisely, it is possible to specify the tidal forces over the Earth at any past or future time with great precision. It should be possible to relate tidal elevations and currents at any point in the oceans to these forces, making use of classical mechanics and hydrodynamics. Such a theoretical approach to tidal prediction has not yet yielded any great success, owing in great part to the complicated shape of the ocean basins. However, use of numerical-hydrodynamical models (such as the work of K. T. Bogdanov, N. Grijalva, W. Hansen, M. C. Henderschott, and C. L. Pekeris) has yielded some satisfactory results and undoubtedly will have practical importance.

The other approach, which consists of making use of past observations of the tide at a certain place to predict the tide for the same place, has yielded practical results. The method cannot be used for a location where there have been no previous observations. In the harmonic method the frequencies of the many tidal constituents are derived from knowledge of the movements of Earth, Moon, and Sun. The amplitude and epoch of each constituent are determined from the tidal observations. The actual tide can then be synthesized by summing up an adequate number of harmonic constituents. The method might loosely be thought of as extrapolation.

A "convolution" method of tidal analysis and prediction has been proposed by W. H. Munk and D. E. Cartwright. In this method past observations at a place are used to determine a numerical operator which, when applied to the known tide-producing forces, will calculate the resulting tide.

In the following discussion only the lunar effect is considered, and it is understood that analogous statements apply to the solar effect.

Tide-generating force. If the Moon attracted every point within the Earth with equal force, there would

be no tide. It is the small difference in direction and magnitude of the lunar attractive force, from one point of the Earth's mass to another, which gives rise to the tidal stresses.

According to Newton's laws, the Moon attracts every particle of the Earth with a force directed toward the center of the Moon, with magnitude proportional to the inverse square of the distance between the Moon's center and the particle. At point A in **Fig. 1**, the Moon is in the zenith and at point B the Moon is at nadir. It is evident that the upward force of the Moon's attraction at A is greater than the downward force at B because of its closer proximity to the Moon. Such differential forces are responsible for stresses in all parts of the Earth. The Moon's gravitational pull on the Earth can be expressed as the vector sum of a constant force, equal to the Moon's attraction on the Earth's center, and a small deviation which varies from point to point in the Earth (**Fig. 2**). This small deviation is referred to as the tide-generating force. The larger constant force is balanced completely by acceleration (centrifugal force) of the Earth in its orbital motion around the center of mass of the Earth-Moon system, and plays no part in tidal phenomena. *See* GRAVITATION.

The tide-generating force is proportional to the mass of the disturbing body (Moon) and to the inverse cube of its distance. This inverse cube law accounts for the fact that the Moon is 2.17 times as important, insofar as tides are concerned, as the Sun, although the latter's direct gravitational pull on the Earth, which is governed by an inverse-square law, is about 180 times the Moon's pull.

The tide-generating force, as illustrated in Fig. 2, can be expressed as the gradient of the tide-generating potential, Eq. (1), where λ is the zenith

$$\psi = \frac{3}{2}\frac{\gamma M r^2}{c^3}\left(\frac{1}{3} - \cos^2\lambda\right) \qquad (1)$$

distance of the Moon, r is distance from the Earth's center, c is distance between the centers of Earth and Moon, γ is the gravitational constant, and M is the mass of the Moon. In this expression, terms containing higher powers of the smaller number r/c have been neglected. As ψ depends only on the space variables r and λ, it is symmetrical about the Earth-Moon axis.

It helps one visualize the form of the tide-generating potential to consider how a hypothetical "inertialess" ocean covering the whole Earth would respond to the tidal forces. In order to be in equilibrium with the tidal forces, the surface must assume the shape of an equipotential surface as determined by both the Earth's own gravity and the tide-generating force. The elevation of the surface is given approximately by Eq. (2), where ψ is evaluated

$$\bar{\zeta} = -\frac{\psi}{g} + \text{const} \qquad (2)$$

at the Earth's surface and g is the acceleration of the Earth's gravity. The elevation ζ of this hypothetical ocean is known as the equilibrium tide. Knowledge of the equilibrium tide over the entire Earth determines completely the tide-generating potential (and hence the tidal forces) at all points within the Earth as well as on its surface. Therefore, when the equilibrium tide is mentioned, it shall be understood that reference to the tide-generating force is also being made.

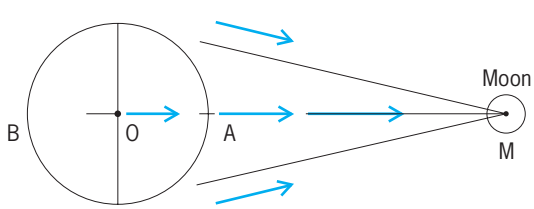

Fig. 1. Schematic diagram of the lunar gravitational force on different points in the Earth.

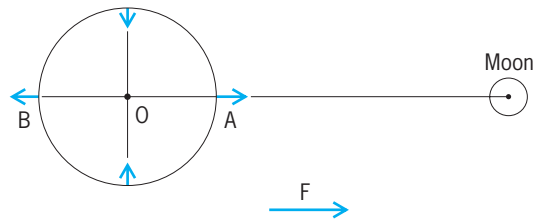

Fig. 2. Schematic diagram of the tide-generating force on different points in the Earth. The vector sum of this tide-generating force and the constant force *F* (which does not vary from point to point) produce the force field indicated in Fig 1. Force *F* is compensated by the centrifugal force of the Earth in its orbital motion.

Harmonic development of the tide. The equilibrium tide as determined from relations (1) and (2) has the form of a prolate spheroid (football-shaped) whose major axis coincides with the Earth-Moon axis. The Earth rotates relative to this equilibrium tidal form so that the nature of the (equilibrium) tidal variation with time at a particular point on the Earth's surface is not immediately obvious. To analyze the character of this variation, it is convenient to express the zenith angle of the Moon in terms of the geographical coordinates θ, ϕ of a point on the Earth's surface (θ is colatitude, ϕ is east longitude) and the declination D and west hour angle reckoned from Greenwich α of the Moon. When this is done, the equilibrium tide can be expressed as the sum of the three terms in Eq. (3), where a is the Earth's radius.

$$\begin{aligned}\bar{\zeta} = &\frac{3}{4}\frac{\gamma M}{g}\frac{a^2}{c^3}[(3\sin^2 D - 1)(\cos^2\theta - {}^1/_3) \\ &+ \sin 2D \sin 2\theta \cos(\alpha + \phi) \\ &+ \cos^2 D \sin^2\theta \cos 2(\alpha + \phi)] \qquad (3)\end{aligned}$$

The first term represents a partial tide which is symmetrical about the Earth's axis, as it is independent of longitude. The only time variation results from the slowly varying lunar declination and distance from Earth. This tide is called the long-period tide. Its actual geographical shape is that of a spheroid whose axis coincides with the Earth's axis and whose oblateness slowly but continuously varies.

The second term of Eq. (3) represents a partial tide having, at any instant, maximum elevations at 45°N and 45°S on opposite sides of the Earth, and two minimum elevations lying at similar, alternate positions on the same great circle passing through the poles. Because of the factor cos $(\alpha + \phi)$ the tide rotates in a westerly direction relative to the Earth, and any geographical position experiences a complete oscillation in a lunar day, the time taken for α to increase by the amount 2π. Consequently, this partial tide is called the diurnal tide. Because of the factor sin $2D$, the diurnal equilibrium tide is zero at the instant the Moon crosses the Equator; because of the factor sin 2θ, there is no diurnal equilibrium tidal fluctuation at the Equator or at the poles.

The third term of Eq. (3) is a partial tide having, at any instant, two maximum elevations on the Equator at opposite ends of the Earth, separated alternately by two minima also on the Equator. This whole form also rotates westward relative to the Earth, making a complete revolution in a lunar day. But any geographic position on the Earth will experience two cycles during this time because of the factor cos $2(\alpha + \phi)$. Consequently, this tide is called the semidiurnal tide. Because of the factor $\sin^2 \theta$, there is no semidiurnal equilibrium tidal fluctuation at the poles, while the fluctuation is strongest at the Equator.

It has been found very convenient to consider the equilibrium tide as the sum of a number of terms, called constituents, which have a simple geographical shape and vary harmonically in time. This is the basis of the harmonic development of the tide. A great number of tidal phenomena can be adequately described by a linear law; that is, the effect of each harmonic constituent can be superimposed on the effects of the others. Herein is the great advantage of the harmonic method in dealing with tidal problems. The three terms of Eq. (3) do not vary with time in a purely harmonic manner. The parameters c and D themselves vary, and the rapidly increasing α does not do so at a constant rate owing to ellipticity and other irregularities of the Moon's orbit. Actually, each of the three partial tides can be separated into an entire species of harmonic constituents. The constituents of any one of the three species have the same geographical shape, but different periods, amplitudes, and epochs.

The solar tide is developed in the same way. As before, the three species of constituents arise: long-period, diurnal, and semidiurnal. The equilibrium tide at any place is the sum of both the lunar and solar tides. When the Sun and Moon are nearly in the same apparent position in the sky (new Moon) or are nearly at opposite positions (full Moon), the lunar and solar effects reinforce each other. This condition is called the spring tide. During the spring tide the principal lunar and solar constituents are in phase. At quadrature the solar effect somewhat cancels the lunar effect, the principal lunar and solar constituents being out of phase. This condition is known as the neap tide.

The entire equilibrium tide can now be expressed by Eq. (4), where $H = 3\gamma Ma^2/gc^3 = 54$ cm, and

$$\bar{\zeta} = H\Big[{}^1/_2(1 - 3\cos^2\theta)\sum_L f_i C_i \cos A_i + \sin 2\theta \sum_D f_i C_i \cos(A_i + \phi) + \sin^2\theta \sum_S f_i C_i \cos(A_i + 2\phi)\Big] \qquad (4)$$

$1/c$ represents the mean (in time) value of $1/c$. Each term in the above series represents a constituent. Terms of higher powers of the Moon's parallax (a/c) are not included in Eq. (4) because of their different latitude dependence, but they are of relatively small importance. The subscripts L, D, and S indicate summation over the long-period, diurnal, and simidiurnal constituents, respectively. The C's are the constituent coefficients and are constant for each constituent. They account for the relative strength of all lunar and solar constituents. In a purely harmonic development, such as carried out by A. T. Doodson in 1921, the A parts of the arguments increase linearly with time, and the node factors f are all unity. In George Darwin's "almost harmonic" development of 1882, the constituents undergo a slow change in amplitude and epoch with the 19-year nodal cycle of the Moon. The node factors f take this slow variation into account. The A's increase almost linearly with time. Tables in U.S. Coast and Geodetic Survey Spec. Publ. 98 enable one to compute the phase of the argument of any of Darwin's constituents at any time, and values of the node factors for each year are given.

In spite of the many advantages of the purely harmonic development, Darwin's method is still used by most agencies engaged in tidal work. In Darwin's classification, each constituent is represented by a symbol with a numerical subscript, 0, 1, or 2, which designates whether the constituent is long-period, diurnal, or semidiurnal. Some of the most important of Darwin's constituents are listed in the **table**.

The periods of all the semidiurnal constituents are grouped about 12 h, and the diurnal periods about 24 h. This results from the fact that the Earth rotates much faster than the revolution of the Moon about the Earth or of the Earth about the Sun. The principal lunar semidiurnal constituent M_2 beats against the others giving rise to a modulated semidiurnal

Darwin's constituents

Constituent	Speed, deg/h	Coefficient
Long-period		
Mf, lunar fortnightly	1.098	0.157
Ssa, solar semiannual	0.082	0.073
Diurnal		
K_1, lunisolar	15.041	0.530
O_1, larger lunar	13.943	0.377
P_1, larger solar	14.959	0.176
Semidiurnal		
M_2, principal lunar	28.984	0.908
S_2, principal solar	30.000	0.423
N_2, larger lunar elliptic	28.440	0.176
K_2, lunisolar	30.082	0.115

waveform whose amplitude varies with the Moon's phase (the spring-neap effect), distance, and so on. Similarly, the amplitude of the modulated diurnal wave varies with the varying lunar declination, solar declination, and lunar phase. For example, the spring tide at full Moon or new Moon is manifested by constituents M_2 and S_2 being in phase, thus reinforcing each other. During the neap tide when the Moon is at quadrature, the constituents M_2 and S_2 are out of phase, and tend to cancel each other. The other variations in the intensity of the tide are similarly reflected in the "beating" of other groups of constituents.

Tides in the ocean. The tide in the ocean deviates markedly from the equilibrium tide, which is not surprising if one recalls that the equilibrium tide is based on neglect of the inertial forces. These forces are appreciable unless the periods of all free oscillations in the ocean are small compared with those of the tidal forces. Actually, there are free oscillations in the ocean (ordinary gravity seiches) having periods of the order of a large fraction of a day, and there may be others (planetary modes) having periods of the order of several days. For the long-period constituents the observed tide should behave like the equilibrium tide, but this is difficult to show because of their small amplitude in the presence of relatively large meteorological effects.

At most places in the ocean and along the coasts, sea level rises and falls in a regular manner. The highest level usually occurs twice in any lunar day, the times bearing a constant relationship with the Moon's meridional passage. The time between the Moon's meridional passage and the next high tide is called the lunitidal interval. The difference in level between successive high and low tides, called the range of the tide, is generally greatest near the time of full or new Moon, and smallest near the times of quadrature. This results from the spring-neap variation in the equilibrium tide. Tide range usually exhibits a secondary variation, being greater near the time of perigee (when the Moon is closest to the Earth) and smaller at apogee (when it is farthest away).

The above situation is observed at places where the tide is predominantly semidiurnal. At many other places, it is observed that one of the two maxima in any lunar day is higher than the other. This effect is known as the diurnal inequality and represents the presence of an appreciable diurnal variation. At these places, the tide is said to be of the "mixed" type. At a few places, the diurnal tide actually predominates, there generally being only one high and low tide during the lunar day.

Both observation and theory indicate that the ocean tide can generally be considered linear. As a result of this fact, the effect in the ocean of each constituent of the series in Eq. (4) can be considered by itself. Each equilibrium constituent causes a reaction in the ocean. The tide in the ocean is the sum total of all the reactions of the individual constituents. Furthermore, each constituent of the ocean tide is harmonic (sinusoidal) in time. If the amplitude of an equilibrium constituent varies with the nodal cycle of the Moon, the amplitude of the oceanic constituent varies proportionately.

As a consequence of the above, the tidal elevation in the ocean can be expressed by Eq. (5),

$$\zeta = \sum f_i h_i \cos (A_i - G_i) \tag{5}$$

where $h_i(\theta,\phi)$ is called the amplitude and $G_i(\theta,\phi)$ the Greenwich epoch of each constituent. The summation in Eq. (5) extends over all constituents of all species. The f's and the A's have the same meaning as in Eq. (4) for the equilibrium tide and are determined from astronomic data.

To specify completely the tidal elevation over the entire surface of the ocean for all time, one would need ocean-wide charts of $h(\theta,\phi)$, called corange charts, and of $G(\theta,\phi)$, called cotidal charts, for each important constituent. Construction of these charts would solve the ultimate problem in tidal prediction. Many attempts have been made to construct cotidal charts. These attempts have been based on a little theory and far too few observations.

Figures 3 and **4** show Dietrich's cotidal chart for M_2. Each curve passes through points having high water at the same time, time being indicated as phase of the M_2 equilibrium argument. A characteristic feature of cotidal charts is the occurrence of points through which all cotidal curves pass. These are called amphidromic points. Here the amplitude of the constituent under consideration must be zero. The existence of such amphidromic points has been borne out by theoretical studies of tides in ocean basins of simple geometric shape. The mechanism which gives rise to amphidromic points is intimately related to the rotation of the Earth and the Coriolis force.

The amplitude of a constituent, $h(\theta,\phi)$, is generally high in some large regions of the oceans and low in others, but in addition there are small-scale erratic variations, at least along the coastline. Perhaps this is partly an illusion caused by the placement of some tide gages near the open coast and the placement of others up rivers and estuaries. It is well known that the phase and amplitude of the tide change rapidly as the tidal wave progresses up a river. *See* RIVER TIDES.

The range of the ocean tide varies between wide limits. The highest range is encountered in the Bay of Fundy, where values exceeding 50 ft (15 m) have been observed. In some places in the Mediterranean, South Pacific, and Arctic, the tidal range never exceeds 2 ft (0.6 m).

The tide may be considerably different in small adjacent seas than in the nearby ocean, and here resonance phenomena frequently occur. The periods of free oscillation of a body of water are determined by their boundary and depth configurations. If one of these free periods is near that of a large tidal constituent, the latter may be amplified considerably in the small sea. The large tidal range in the Bay of Fundy is an example of this effect. Here the resonance period is nearly 12 h, and it is the semidiurnal constituents that are large. The diurnal constituents are

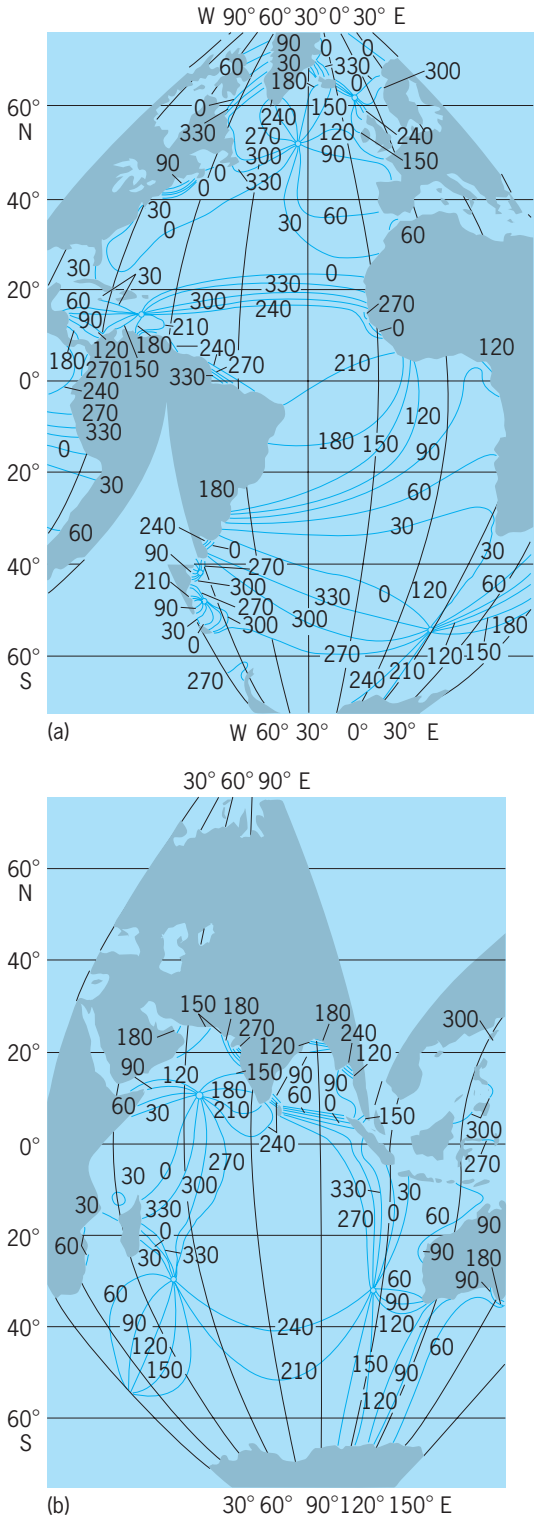

Fig. 3. Cotidal chart for M_2. (*a*) Atlantic Ocean. (*b*) Indian Ocean. (*After G. Dietrich, Veroeff. Inst. Meeresk., n.s. A, Geogr.- naturwiss. Reihe, no. 41, 1944*)

not extremely greater in the Bay of Fundy than in the nearby ocean. *See* SEICHE.

In lakes and other completely enclosed bodies of water the periods of free oscillation are usually much smaller than those of the tidal constituents. Therefore the tide in these places obeys the principles of statics. Since there is no tidal variation in the total volume of water in lakes the mean surface elevation does not change with the tide. The surface slope is determined by the slope of the equilibrium tide, and the related changes in elevation are usually very small, of the order of a fraction of a millimeter for small lakes.

Tidal currents. The south and east components of the tidal current can be developed in the same way as the tidal elevation since they also depend linearly on the tidal forces. Consequently, the same analysis and prediction methods can be used. Expressions similar to Eq. (5) represent the current components, each constituent having its own amplitude and phase at each geographic point. It should be emphasized that the current speed or direction cannot be developed in this way since these are not linearly related to the tidal forces.

Only in special cases are the two tidal current components exactly in or out of phase, and so the tidal current in the ocean is generally rotatory. A drogue or other floating object describes a trajectory similar in form to a Lissajous figure. In a narrow channel only the component along its axis is of interest. Where shipping is important through such a channel or port entrance, current predictions, as well as tidal height predictions, are sometimes prepared.

Owing to the rotation of the Earth, there is a gyroscopic, or Coriolis, force acting perpendicularly to the motion of any water particle in motion. In the Northern Hemisphere this force is to the right of the current vector. The horizontal, or tractive, component of the tidal force generally rotates in the clockwise sense in the Northern Hemisphere. As a result of both these influences, the tidal currents in the open ocean generally rotate in the clockwise sense in the Northern Hemisphere, and in the counterclockwise sense in the Southern Hemisphere. There are exceptions, however, and the complete dynamics should be taken into account. *See* CORIOLIS ACCELERATION.

The variation of the tidal current with depth is not well known. It is generally agreed that the current would be constant from top to bottom were it not for stratification of the water and bottom friction. The variation of velocity with depth due to the stratification of the water is associated with internal wave motion. Serial observations made from anchored or drifting ships have disclosed prominent tidal periodicities in the vertical thermal structure of the water.

Dynamics of ocean tide. The theoretical methods for studying tidal dynamics in the oceans were put forth by Laplace in the eighteenth century. The following assumptions are introduced: (1) The water is homogeneous; (2) vertical displacements and velocities of the water particles are small in comparison to the horizontal displacements and velocities; (3) the water pressure at any point in the water is given adequately by the hydrostatic law, that is, it is equal to the head of water above the given point; (4) all dissipative forces are neglected; (5) the ocean basins are assumed rigid (as if there were no bodily tide), and the gravitational potential of the

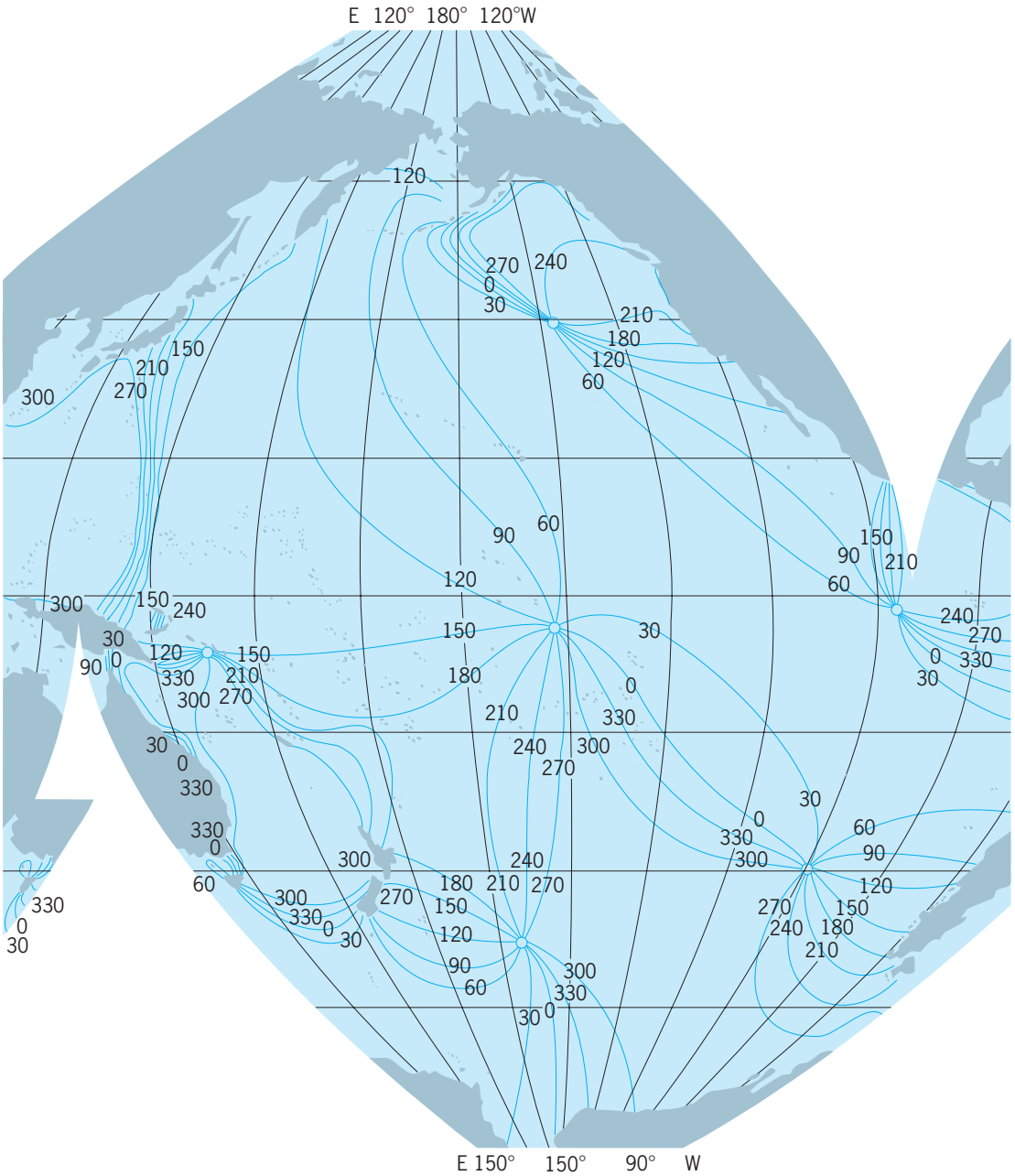

Fig. 4. Pacific Ocean cotidal chart for M_2. (*After G. Dietrich, Veroeff. Inst. Meeresk., n.s. A, Geogr.-naturwiss. Reihe, no. 41, 1944*)

tidally displaced masses is neglected; and (6) the tidal elevation is small compared with the water depth.

If assumptions (1) and (3) are valid, it can readily be shown that the tidal currents are uniform with depth. This is a conclusion which is not in complete harmony with observations, and there are internal wave modes thus left out of Laplace's theory. Nevertheless the main features of the tide are probably contained in the equations.

The water motion in the oceans is, in theory, determined by knowledge of the shape of the ocean basins and the tide-generating force (or equilibrium tide) at every point in the oceans for all time. The theory makes use of two relations: (1) the equation of continuity, which states that the rate of change of water mass in any vertical column in the ocean is equal to the rate at which water is flowing into the column; and (2) the equations of motion, which state that the total acceleration of a water "particle" (relative to an inertial system, thus taking into account the rotation of the Earth) is equal to the total force per unit mass acting on that particle. Under the above assumptions, the equation of continuity takes the form of Eq. (6), where $d(\theta,\phi)$ is the water

$$\frac{\partial \zeta}{\partial t} = -\frac{1}{a \sin\theta}\left[\frac{\partial}{\partial \theta}(ud \sin\theta) + \frac{\partial}{\partial \phi}(vd)\right] \quad (6)$$

depth. The equations of motion in the southward

and eastward directions, respectively, are given by Eqs. (7), where ω designates the angular rate of ro-

$$\begin{aligned} \frac{\partial u}{\partial t} - 2\omega v \cos\theta &= \frac{g}{a}\frac{\partial}{\partial\theta}(\zeta - \overline{\zeta}) \\ \frac{\partial v}{\partial t} + 2\omega u \cos\theta &= -\frac{g}{a}\csc\theta\frac{\partial}{\partial\phi}(\zeta - \overline{\zeta}) \end{aligned} \tag{7}$$

tation of the Earth, and u and v the south and east components of the tidal current. All other quantities are as previously defined.

It is probable that exact mathematical solutions to Eqs. (6) and (7), taking even approximately into account the complicated shape of the ocean basins, will never be obtained. However, the equations have certain features which serve to give us some insight into the nature of ocean tides. For instance, it is evident that if many equilibrium tides are acting simultaneously on the ocean, then the ocean tide will be the sum of the individual reactions. This linearity results directly from assumption (6). In certain shallow regions of the ocean the tides are noticeably distorted, as would be expected if assumption (6) were violated. This distortion is usually considered as resulting from the presence of so-called shallow-water constituents having frequencies equal to harmonics and to beat frequencies of the equilibrium constituents. These must be considered, at some places, or there will be large discrepancies between prediction and observation. Certain mathematical solutions to Eqs. (6) and (7) have been obtained for hypothetical ocean basins of simple geometric shape. Laplace solved them for an ocean of constant depth covering the entire Earth. Several solutions have been obtained for an ocean of constant depth bounded by two meridians. The result of one of the solutions obtained by J. Proudman and A. Doodson is shown in **Fig. 5**, which represents a cotidal chart of the K_2 tide in an ocean of depth 14,520 ft (4426 m) bounded by meridians 70° apart. The K_2 tide was calculated because of mathematical simplifications, but the M_2 tide should be quite similar. Comparison of Fig. 5 with the Atlantic Ocean in Fig. 3 discloses no striking similarities except for the general occurrence of amphidromic systems.

Bodily tide. The solid part of the Earth suffers periodic deformation resulting from the tide-generating forces just as the oceans do. *See* EARTH TIDES.

The gravest known modes of free oscillation of the solid Earth have periods of the order of an hour, much shorter than those of the principal tidal constituents. Therefore, the principles of statics can be used to describe the bodily tide, in contrast to tides in the oceans and atmosphere, where the inertial effect is important.

Associated with the bodily tide are periodic changes in gravity, manifesting themselves as (1) a variation of the vertical, or plumb line, with respect to any solid structure embedded in the Earth's crust; and (2) a variation in the magnitude of the acceleration of gravity at any point. These effects arise from the gravitational attraction of the tidally displaced matter of the Earth (solid, ocean, and atmosphere)

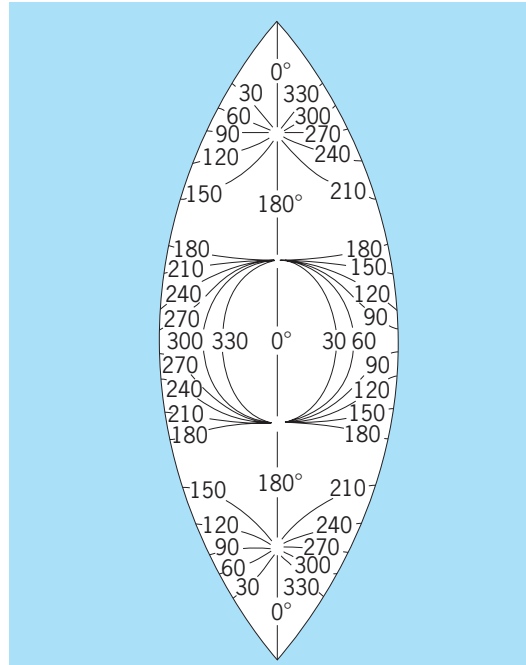

Fig. 5. Cotidal chart for K_2 in a hypothetical ocean of constant depth bounded by meridians 70° apart. (*After A. T. Doodson and H. D. Warburg, Admiralty Manual of Tides, London, 1941*)

as well as directly from the tide-generating forces. The magnitude of the former factor is of the order of several tens of microgals (1 gal = 1 cm/s^2).

Atmospheric tides. Since air, as other matter, is subject to gravitational influence, there are tides in the atmosphere possessing many features of similarity with those in the ocean. One of the characteristics of these tides is a small oscillatory variation in the atmospheric pressure at any place. This fluctuation of pressure, as in the case of the ocean tide, may be considered as the sum of the usual tidal constituents, and standard tidal analysis and prediction methods may be used. The principal lunar semidiurnal constituent M_2 of the pressure variation has been determined for a number of places, and found to have an amplitude of the order of 0.03 millibar (3 pascals). The dynamical theory of these tides has been the subject of considerable study. The equations which have been considered have the same general form as those for ocean tides. The S_2 constituent shows a much larger oscillation with an amplitude of the order of 1 millibar (10^2 Pa), but here diurnal heating dominates the gravitational effects. If diurnal heating were the whole story one would expect an even larger S_1 effect, and the fact that S_2 is larger is attributed to an atmospheric resonance near 12 h. *See* ATMOSPHERIC TIDES; UPPER-ATMOSPHERE DYNAMICS.

Tidal analysis and prediction. The distribution in space and time of the tidal forces within the Earth is precisely known from astronomic data. The effects of these forces on the oceans cannot, by present methods, be described in detail on a worldwide basis because of the difficult nature of the dynamical relationships and the complicated shape of the ocean basins. Practical prediction methods make use

of past observations at the place under consideration.

The procedure is the same for prediction of any tidal variable—such as the atmospheric pressure, component displacements of the solid Earth, components of the tidal current, and so on—which depends linearly on the tidal forces. In the harmonic method the frequencies, or periods, of the tidal constituents are determined by the astronomic data, and the harmonic constants (amplitudes and epochs) are obtained from the observations. Equation (5) then represents the tide at all past and future times for the place under consideration, where the values of b are the amplitudes of whatever tidal variable is being predicted. In this discussion the sea-level elevation will be used as an example, since it is the variable for which predictions are most commonly made. The procedure is basically the same for each constituent, but is most easily described for the series of constituents, $S_1, S_2, S_3, \ldots$, whose periods are submultiples of 24 h.

Suppose that the tidal elevation at 1:00 A.M. is averaged for all the days of the tide record, and similarly for 2:00 A.M., 3:00 A.M., and for each hour of the day. The 24 values thus obtained represent the average diurnal variation during the entire record. Any constituent whose period is not a submultiple of 24 h will contribute very little to the average of all the 1:00 A.M. values since its phase will be different from one day to the next, and its average value at 1:00 A.M. will be very close to zero for a long record. The same is true for each hour of the day, and so its average diurnal variation is small. The longer the record the freer will be the average diurnal oscillation from the effects of the other constituents. The diurnal oscillation is then analyzed by the well-known methods of harmonic analysis to determine the amplitudes and phases of all the harmonics of the 24-h oscillation. *See* FOURIER SERIES AND TRANSFORMS.

The same procedure is used for each other constituent; that is, the tide record is divided into consecutive constituent days, each equal to the period (or double the period in the case of the semidiurnal constituents) of the constituent. If the tide record is tabulated each solar hour, there is a slight complication because the constituent hours do not coincide with the solar hours. This difficulty is overcome by substituting the tabulated value nearest the required time and later compensating the consistent error introduced by an augmenting factor.

Since the record length is always finite, the harmonic constants of a constituent determined by this method are somewhat contaminated by the effects of other constituents. A first-order correction of these effects can be made by an elimination procedure. In general, it is more efficient to take the record length equal to the synodic (beat) period of two or more of the principal constituents. The longer the record, the better. Standard analyses consist of 29 days, 58 days, 369 days, and so on.

It is not practical to determine the harmonic constants of the lesser constituents in this way if errors or uncertainties of the data are of the same order of magnitude as their amplitudes. If tidal oscillations in the oceans were far from resonance then the amplitude H of each constituent should be expected to be approximately proportional to its theoretical coefficient C, and the local epochs G all to be near the same value. For the semidiurnal constituent X, Eqs. (8) should hold. Here X is referred to M_2 for

$$\frac{H(X)}{C(X)} = \frac{H(M_2)}{C(M_2)} \qquad G(X) = G(M_2) \qquad (8)$$

the reason that the latter is one of the principal constituents whose harmonic constants can be determined with best accuracy. Any other important constituent could be used. Inferring the harmonic constants of the lesser constituents by means of Eqs. (8) is sometimes preferable to direct means. It should be borne in mind that a constituent of one species cannot be inferred from one of another species because their equilibrium counterparts have different geographic shapes and no general relationship such as Eqs. (8) exists.

Once the harmonic constants are determined, the tide is synthesized according to Eq. (5), usually with the help of a special tide-predicting machine, although any means of computation could be used. Usually only the times and heights of high and low water are published in the predictions.

Tidal friction. The dissipation of energy by the tide is important in the study of planetary motion because it is a mechanism whereby angular momentum can be transferred from one type of motion to another. An appreciable amount of tidal dissipation takes place in the ocean, and possibly also in the solid Earth. In 1952 Sir Harold Jeffreys estimated that about half the tidal energy present in the ocean at any time is dissipated each day. A large part of this dissipation takes place by friction of tidal currents along the bottom of shallow seas and shelves and along the coasts. The rate of dissipation is so large that there should be a noticeable effect on the tide in the oceans.

If the planet's speed of rotation is greater than its satellite's speed of revolution about it, as is the case in the Earth-Moon system, then tidal dissipation always tends to decelerate the planet's rotation, with the satellite's speed of revolution changing to conserve angular momentum of the entire system. The Moon's attraction on the irregularly shaped tidal bulge on the Earth exerts on it a decelerating torque. Thus tidal friction tends to increase the length of day, to increase the distance between Earth and Moon, and to increase the lunar month, but these increases are infinitesimal. The day may have lengthened by 1 s during the last 120,000 years because of tidal friction and other factors.

Gordon W. Groves

Bibliography. D. E. Cartwright, *Tides: A Scientific History*, 1998; P. Crean, T. S. Murty, and J. A. Stronach, *Mathematical Modeling of Tides and Estuarine Circulation*, 1988; H. Lamb, *Hydrodynamics*, 6th ed., 1945; G. I. Marchuk and B. A. Kagan, *Dynamics of Ocean Tides*, 1989; D. T. Pugh, *Tides, Surges and Mean Sea Level*, 1987.

Tie rod

A rod or bar, usually circular in cross section, used in structural parts of machines to tie together or brace connected members, or, in moving parts of machines or mechanisms, used to connect arms or parts to transmit motion. In the first use the rod ends are usually a threaded fastening, while in the latter they are usually forged into an eye for a pin connection.

In steering systems of automotive vehicles, the rod connects the arms of steering knuckles of each wheel. The connection between the rod and arms is a ball and socket joint. *See* AUTOMOTIVE STEERING.

In pressure piping, large forces are produced between connected parts. The pipes or parts are constrained by tie rods that may be rectangular in cross section, with pinned ends. *See* PIPELINE.

Paul H. Black

Tile

As a structural material, a burned clay product in which the coring exceeds 25% of the gross volume; as a facing material, any thin, usually flat, square product. Structural tile used for load bearing may or may not be glazed; it may be cored horizontally or vertically. Two principal grades are manufactured: one for exposed masonry construction, and the other for unexposed construction. Among the forms of exposure is frost; tile for unexposed construction where temperatures drop below freezing is placed within the vapor barrier or otherwise projected by a facing in contrast to roof tile.

Structural tile with a ceramic glaze is used for facing. The same clay material that is molded and fired into structural tile is also made into pipe, glazed for sewer lines, or unglazed for drain tile.

As a facing, clay products are formed into thin flat, curved, or embossed pieces, which are then glazed and burned. Commonly used on surfaces that are subject to water splash or that require frequent cleaning, such vitreous glazed wall tile is fireproof. Unglazed tile is laid as bathroom floor. By extension, any material formed into a size comparable to clay tile is called tile. Among the materials formed into tile are asphalt, cork, linoleum, vinyl, and porcelain. *See* CLAY, COMMERCIAL.

Frank H. Rockett

Till

Sediment deposited directly from glacier ice. Till is characteristically nonsorted and nonstratified and is deposited by lodgment or melt-out beneath a glacier or by melt-out on the surface of a glacier. The texture of till varies greatly (see **illus.**), and while all tills are characterized by a wide range of particle sizes, some are predominantly fine-grained (clayey or silty), while others are medium-grained (silty or sandy) or coarse-grained (gravelly or stony). Till contains a variety of rock and mineral fragments which reflect the source material over which the glacier flowed. The particles in the deposit usually show a preferred orientation related to the nature and direction of the ice flow. The overall character of the till reflects the source material, position and distance of transport, nature and position of deposition, and postdepositional changes.

Exposure of glacial till at the Black Rocks near Llandudno, Wales. Heterogeneous debris, ranging in size from large boulders to fine powder, displays no sorting or stratification. (*Photograph by K. F. Mather*)

Material released through melting on the surface of a glacier undergoes secondary modification by melt water or through mass movement. The former gives rise to various types of stratified glaciofluvial and glaciolacustrine deposits; the latter often gives rise to viscous debris (mud) flows, the deposits of which are lithologically similar to till. Such deposits have been considered a type of till and are called flow till. However, on a strict genetic basis, these deposits are no longer till; resedimentation has taken place, and they are debris flow deposits.

Till is a common surficial deposit in middle- and high-latitude areas that were glaciated during the Quaternary Period, and is the parent material for some of the best agricultural soils in the world. *See* GLACIATED TERRAIN.

W. Hilton Johnson

Bibliography. J. Ehlers et al. (eds.), *Glacial Deposits in North-East Europe*, 1995; D. E. Lawson, *Sedimentological Analysis of the Western Terminus Region of the Matanuska Glacier, Alaska*, Cold Reg. Res. Eng. Lab. Rep. 79–9, Hanover, NH, 1979; R. F. Legget (ed.), *Glacial Till*, Roy. Soc. Can. Spec. Pub. 12, 1976.

Tillodontia

An extinct order of early Cenozoic (about 65 to 40 million years ago) quadrupedal eutherian land mammals, represented by nine known genera, from the late Paleocene to middle Eocene of North America (*Esthonyx* [*Azygonyx*], *Megalesthonyx*, *Trogosus*, and *Tillodon*), early Paleocene to late Eocene of China (*Lofochaius*, *Meiostylodon*, *Adapidium*, *Trogosus* [*Kuanchuanius*]), middle Eocene of Pakistan (*Basalina*), and the early Eocene of Europe (*Plesiesthonyx*). *Anchippodus*, the first named tillodont genus (1868), based on a single left

lower molar found in middle Eocene rocks of New Jersey, is a nomen dubium (it is not certain which species of tillodont is represented by this tooth) and may represent either *Trogosus* or *Tillodon*. An indeterminate tillodont incisor fragment found on Ellesmere Island demonstrates that tillodonts lived in what is now the Canadian High Arctic during the Eocene under much warmer conditions than exist there today.

The tillodonts left no known descendants and were probably most closely related to the extinct order Pantodonta (another group of extinct ungulate-like mammals from the Paleocene and Eocene that in turn may be related to the early eutherian mammals known as arctocyonids).

Tillodonts were medium- to large-sized mammals (their skulls range in length from 5 to 37 cm or 2 to 15 in.) that probably fed primarily on roots and tubers in warm temperate to subtropical habitats. Tillodonts were most common in the early Eocene faunas of North America. They developed large second incisors that became rodentlike, relatively long snouts, massive skeletons, and moderately large claws. In some respects the adaptive morphology of the tillodonts converged on that of the taeniodonts, and members of the two groups may have competed for similar resources. In the past, some tillodont specimens have been misidentified as taeniodonts; the Pakistan tillodont *Basalina*, for instance, was originally referred to the Taeniodonta. *See* ARCHAIC UNGULATE; MAMMALIA; TOOTH. Robert M. Schoch

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, W. H. Freeman, 1988; C. L. Gazin, *The Tillodontia: An Early Tertiary Order of Mammals*, Smithsonian Miscellaneous Collections, 1953; C. M. Janis, K. M. Scott, and L. L. Jacobs (eds.), *Tillodontia, Evolution of Tertiary Mammals of North America*, vol. 1: *Terrestrial Carnivores, Ungulates, and Ungulatelike Mammals*, Cambridge University Press, 1988.

Time

The dimension of the physical universe that orders the sequence of events at a given place; also, a designated instant in this sequence, such as the time of day, technically known as an epoch, or sometimes as an instant.

Measurement. Time measurement consists of counting the repetitions of any recurring phenomenon and possibly subdividing the interval between repetitions. Two aspects to be considered in the measurement of time are frequency, or the rate at which the recurring phenomena occur, and epoch, or the designation to be applied to each instant.

A determination of time is equivalent to the establishment of an epoch or the correction that should be applied to the reading of a clock at a specified epoch. A time interval may be measured as the duration between two known epochs or by counting from an arbitrary starting point, as is done with a stopwatch. Time units are the intervals between successive recurrences of phenomena, such as the period of rotation of the Earth or a specified number of periods of radiation derived from an atomic energy–level transition. Other units are arbitrary multiples and subdivisions of these intervals, such as the hour being 1/24 of a day, and the minute being 1/60 of an hour. *See* DAY; MONTH; TIME-INTERVAL MEASUREMENT; YEAR.

Bases. Several phenomena are used as bases with which to determine time. The phenomenon traditionally used has been the rotation of the Earth, where the counting is by days. Days are measured by observing the meridian passages of the Sun or stars and are subdivided with the aid of precision clocks. The day, however, is subject to variations in duration because of the variable rotation rate of the Earth. Thus, when a more uniform time scale is required, other bases for time must be used.

Sidereal time. The angle measured along the celestial equator between the observer's local meridian and the vernal equinox is the measure of sidereal time. In practice, a conventionally adopted mathematical expression provides this time as a function of civil time. It is reckoned from 0 to 24 hours, each hour being subdivided into 60 sidereal minutes and the minutes into 60 sidereal seconds. Sidereal clocks are used for convenience in many astronomical observatories because a star or other object outside the solar system comes to the same place in the sky at virtually the same sidereal time.

Solar time. The angle measured along the celestial equator between the observer's local meridian and the Sun is the apparent solar time. The only true indicator of local apparent solar time is a sundial. Mean solar time has been devised to eliminate the irregularities in apparent solar time that arise from the inclination of the Earth's orbit to the plane of the Sun's motion and the varying speed of the Earth in its orbit. In practice it is defined by a conventionally adopted mathematical expression. Intervals of sidereal time can be converted into intervals of mean solar time by dividing by 1.002 737 909 35. Both sidereal and solar time depend on the rotation of the Earth for their time base. *See* EQUATION OF TIME.

Universal Time (UT). Historically, the mean solar time determined for the meridian of 0° longitude using astronomical observations was referred to as UT1. Currently UT1 is used only as an angle expressed in time units that depends on the Earth's rotation with respect to the celestial reference system. It is defined by a conventional mathematical expression and continuing astronomical observations. These are made at a number of observatories around the world. The International Earth Rotation and Reference System Service (IERS) receives these data and provides daily values of the difference between UT1 and civil time. *See* EARTH ROTATION AND ORBITAL MOTION.

Because the Earth has a nonuniform rate of rotation and a uniform time scale is required for many timing applications, a different definition of a second was adopted in 1967. The international agreement calls for the second to be defined as 9,192,631,770 periods of the radiation derived from an energy-level

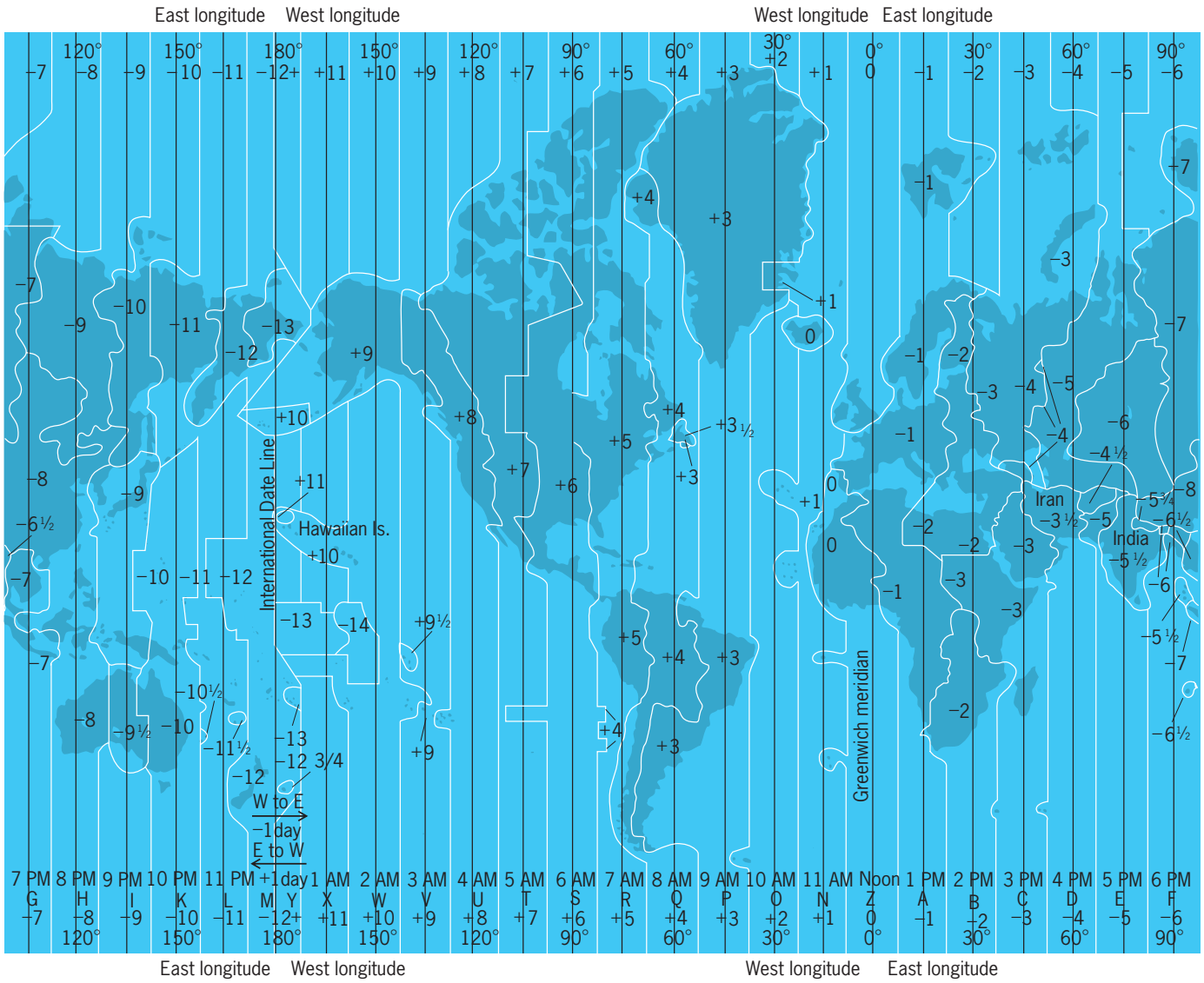

Division of the world into 24 time zones, progressively differing from Greenwich by 1 hour. Some countries use half-hour intervals or fractional hours. Numerical designations indicate number of hours by which zone time must be increased or decreased to obtain Coordinated Universal Time. Longitudes of standard meridians, letter designations, and the times in the zones when it is noon at Greenwich are also shown. (*Updated with data from www.worldtimezone.com and other Web sites*)

transition in the cesium atom. This second is referred to as the international or SI (International System) second and is independent of astronomical observations. International Atomic Time (TAI) is maintained by the International Bureau of Weights and Measures (BIPM) from data contributed by time-keeping laboratories around the world. *See* ATOMIC TIME.

Coordinated Universal Time (UTC) uses the SI second as its time base. However, the designation of the epoch may be changed at certain times so that UTC does not differ from UT1 by more than 0.9 s. UTC forms the basis for civil time in most countries and may sometimes be referred to unofficially as Greenwich Mean Time. The adjustments to UTC to bring this time scale into closer accord with UT1 consist of the insertion or deletion of integral seconds. These "leap seconds" may be applied preferably at 23 h 59 m 59 s of June 30 or December 31 of each year according to decisions made by the IERS. UTC differs from TAI by an integral number of atomic seconds.

Dynamical time. Dynamical time is based on the apparent orbital motion of the Sun, Moon, and planets. It is the time inferred in the ephemerides of the positions of these objects, and from its inception in 1952 until 1984 was referred to as Ephemeris Time. Barycentric Dynamical Time (TDB) refers to ephemerides that have been computed by using the barycenter of the solar system as a reference. Terrestrial Dynamical Time (TDT) is the practical realization of dynamical time and is defined as being equal to TAI + 32.184 seconds. In 1991, the International Astronomical Union recommended that TDT be renamed Terrestrial Time (TT), that Geocentric Coordinate Time (TCG) be the time coordinate for the geocenter, and that Barycentric Coordinate Time (TCB) be the time coordinate for the barycenter of the solar system. These times are related by the appropriate relativistic transformations. *See* DYNAMICAL TIME.

Civil and standard times. Because rotational time scales are local angular measures, at any instant they vary from place to place on the Earth. When the mean solar time is 12 noon at Greenwich, the mean solar time for all places west of Greenwich is earlier than noon and for all places east of Greenwich later than noon, the difference being 1 hour for each 15° of longitude. Thus, at the same instant at short distances east of the 180th meridian the mean solar time is 12:01 A.M., and at a short distance west of the same meridian it is 11:59 P.M. of the same day. Thus persons traveling westward around the Earth must advance their time 1 day, and those traveling eastward must retard their time 1 day in order to be in agreement with their neighbors when they return home. The International Date Line is the name given to a line where the change of date is made. It follows approximately the 180th meridian but avoids inhabited land. To avoid the inconvenience of the continuous change of mean solar time with longitude, zone time or civil time is generally used. The Earth is divided into 24 time zones, each approximately 15° wide and centered on standard longitudes of 0°, 15°, 30°, and so on (see **illustration**). Within each of these zones the time kept is related to the mean solar time of the standard meridian. *See* INTERNATIONAL DATE LINE.

Zone time is reckoned from 0 to 24 hours for most official purposes, the time in hours and minutes being expressed by a four-figure group followed by the zone designation. For example, "1009 zone plus five" refers to the zone 75° west of Greenwich, where zone time must be increased by 5 hours to obtain UTC. The various zones are sometimes designated by letters, especially the Greenwich zone which is Z, "1509 Z" meaning 1509 UTC. The zone centered on the 180th meridian is divided into two parts, the one east of the date line being designated plus 12 and the other minus 12. The time July 2,2400 is identical with July 3,0000.

In civil life the designations A.M. and P.M. are often used, usually with punctuation between hours and minutes. Thus 1009 may be written as 10:09 A.M. and 1509 as 3:09 P.M. The designations for noon and midnight, however, are often confused, and it is better to write 12:00 noon and July 2–3, 12:00 midnight, in order to avoid ambiguity. In some occupations where time is of special importance, there is a rule against using 12:00 at all, 11:59 or 12:01 being substituted. The time 1 minute after midnight is 12:01 A.M. and 1 minute after noon is 12:01 P.M.

The illustration shows the designations of the various time zones, the longitudes of the standard meridians, and the letter designations and the times in the various zones when it is noon at Greenwich. In the United States the boundaries of the time zones are fixed by the Department of Transportation. Frequently the actual boundaries depart considerably from the meridians exactly midway between the standard meridians. Ships at sea and transoceanic planes usually use UTC for navigation and communication, but for regulating daily activities onboard they use any convenient approximation to zone time, avoiding frequent changes during daylight hours.

Many countries, including the United States, advance their time 1 hour, particularly during the summer months, into "daylight saving time." For example, 6 A.M. is redesignated as 7 A.M. Such a practice effectively transfers an hour of little-used early morning light to the evening.

Time scales are coordinated internationally by the BIPM. Most countries maintain local time standards to provide accurate time within their borders by radio, telephone, and TV services. These national time scales are often intercompared by using the Global Positioning System (GPS) or time signals transferred by artificial Earth satellites. *See* SATELLITE NAVIGATION SYSTEMS.

Dennis D. McCarthy

Bibliography. D. W. Allan, N. Ashby, and C. C. Hodge, *The Science of Timekeeping*, Hewlett Packard Appl. Note 1289, 1997; D. D. McCarthy, Astronomical time, *Proc. IEEE*, 79:915–920, 1991; R. A. Nelson et al., The leap second: Its history and possible future, *Metrologia*, 38:509–529, 2001; G. M. R. Winkler, Timekeeping and its applications, in L. Marton (ed.), *Advances in Electronics and Electron Physics*, vol. 44, pp. 33–39,1997.

Time, arrow of

The uniform and unique direction associated with the apparent inevitable flow of time into the future. There appears to be a fundamental asymmetry in the universe. Yet herein lies a paradox, for all the laws of physics, whether they are the equations of classical mechanics, classical electromagnetism, general relativity, or quantum mechanics, are time-reversible in the sense that they admit solutions in either direction of time. This reversibility raises the question of how these fundamentally time-symmetrical equations can result in the perceived asymmetry of temporally ordered events.

Fundamental time asymmetries. The symmetry breaking of temporal order has not yet been fully explained. There are certain indications that an intrinsic asymmetry exists in temporal evolution. Thus it may be that the fundamental laws of physics are not really time-symmetric and that the currently known laws are only symmetrized approximations to the truth. Indeed, the decay of the K^0 meson is not time-reversible. However, it is not at all clear how such a rare and exotic instance of time asymmetry could emerge into the world of essentially macroscopic, electromagnetic phenomena as an everyday observable. *See* TIME REVERSAL INVARIANCE.

Another, more ubiquitous example of a fundamentally time-asymmetric process is the expansion of the universe. It has been speculated that this expansion is the true basis of time asymmetry, and that a resolution of the problem will come when quantum theory (the best available set of equations for mechanics) and general relativity (which deals with the structure of space-time) are combined into a single global theory. The existence of a single dimension

of time is also related to the question of time asymmetry. Because time has only one dimension, there is no analog of spatial rotation, and an entity that is evolving forward in time cannot reverse its direction in time in the same way that it can rotate in space and retrace its steps. Possibly, the inception of the universe at the big bang resulted in progress along time in a particular direction (the one to which the name "forward" is given), and it is now impossible to rotate evolution into the opposite direction. Even if the universe were to collapse again, there is no need to suppose that events will run in reverse, for the possibility is open for a final singularity to be far more elaborate than the initial singularity; so a cosmic asymmetry may exist even in a closed universe (**Fig. 1**). *See* BIG BANG THEORY; COSMOLOGY.

Statistical arguments. An alternative point of view is that time's arrow would exist even in the absence of these fundamental asymmetries. This view is essentially based on the statistical interpretation of the second law of thermodynamics, which identifies the direction of spontaneous change with the increase in disorder of the universe. The first serious attempt to relate the increase in entropy, the statistical measure of disorder, to the underlying time-symmetrical physical laws was made by L. Boltzmann. In his $\mathcal{H}$ theorem, he purported to demonstrate that a property which he called $\mathcal{H}$, and which is an integral over the positions and momenta of particles in a system, invariably decreased with time. Then, by identifying $\mathcal{H}$ with the negative of the entropy, he considered that he had proven that entropy increases with time, and thereby demonstrated the presence of time's arrow in statistically large assemblies of particles. However, Boltzmann's proof was invalid: he had imposed the equivalent of time asymmetry at one stage in the argument. The crucial step was to suppose that although the positions and momenta of two particles were not correlated before they collided, they became correlated by virtue of the collision. Thus, the collision imposed temporal asymmetry. *See* ENTROPY; STATISTICAL MECHANICS; THERMODYNAMIC PRINCIPLES.

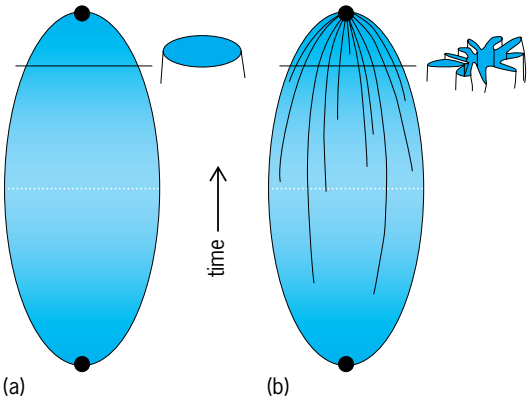

Fig. 1. Depiction of the possible course of evolution of a closed universe (one that starts with a big bang and ends in a big crunch). (*a*) Time-symmetric universe in which the final singularity is no more complex than the initial singularity. (*b*) Time-asymmetric universe in which the final universe is vastly more complex than the initial singularity.

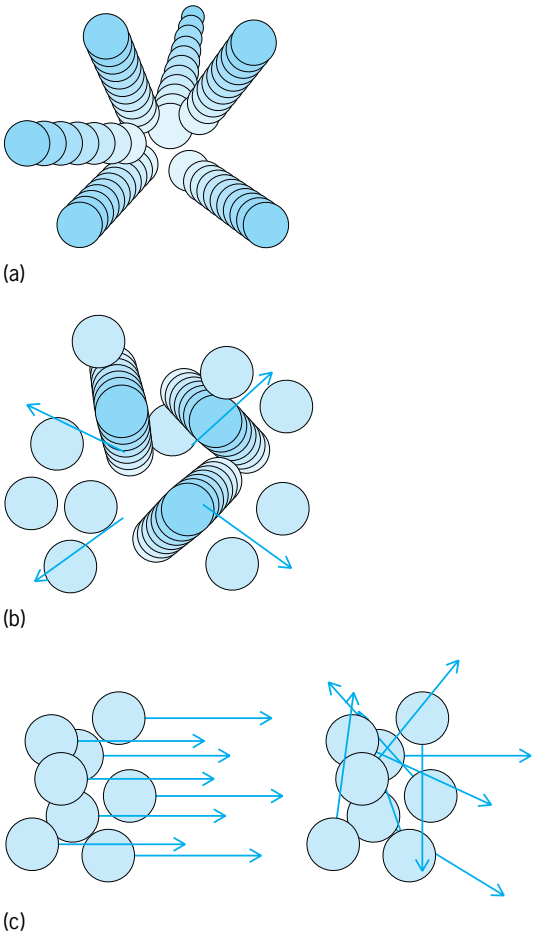

Fig. 2. The three basic irreversibilities of nature which ensure that events are irreversible: (*a*) dispersion of matter, (*b*) dispersion of energy, and (*c*) disorganization of orderly motion.

Consciousness of time asymmetry. It is appropriate to distinguish human consciousness of time asymmetry from an objective asymmetry in the evolution of events. Human consciousness of the unidirectional flow of time stems from the accumulation of memories, which are stored in an as yet unknown format in the brain. However, there can be no doubt that the mechanism of recording is chemical, and that memory is a neurochemical process. Therefore, any irreversibility of chemical reactions in the brain will result in storage of a memory, and hence in building a personality. The effective irreversibility of chemical reactions, neurochemical or otherwise, is well understood. Personality can evolve only as memories are added to previously existing memories, and this accumulation of memory is at the root of human consciousness of the passage of time. If time sometimes stood still and sometimes reversed its direction (there is not the slightest evidence for either event, and they would be difficult to reconcile with special relativity), then there would be no way of detecting the interruptions to time's forward flow, as all physiological changes would be suspended or reversed before resuming again.

Dispersal of energy and matter. The apparent irreversibility of phenomena is understood in terms of the dispersal of energy and matter that accompanies chemical change and the extreme unlikelihood that that dispersal will run in reverse spontaneously (**Fig. 2**). Thus, when a gas is released into a larger volume it spreads throughout the container because of the chaotic motion of its molecules. There is a vanishingly small probability that all the molecules will accumulate simultaneously and spontaneously back into the initial region that they occupied. They might accumulate there, but there is a serious chance of this happening only after such long intervals that the system will have been changed beyond recognition by mundane changes, astronomical disasters, the expansion of the universe, or baryon decay. The same is true of another basic irreversibility, that is, the flow of energy from high temperature to low. This takes place by a similar dispersal mechanism of the thermal motion of molecules, and it is no more reversible than is the dispersal of particles. These basic irreversibilities underpin chemical irreversibility, and hence lie beneath the irreversibility of all material change.

There may be fundamental reasons relating to the structure of space-time that account for the perceived asymmetry of time despite the current formulation of basic physical laws. Alternatively, even a time-symmetrical universe will have a statistical behavior in which configurations of molecules and localizations of energy have significant probabilities of recurring only after enormously long time intervals. Indeed, such time intervals are longer than the times required for the ceaseless expansion of the universe and the evolution of its component particles. Time's arrow is destined, either by the nature of space-time or the statistics of large assemblies, to fly into the future. P. W. Atkins

Bibliography. P. Coveney and R. Highfield, *The Arrow of Time*, 1990; R. Flood and M. Lockwood (eds.), *The Nature of Time*, 1988; J. J. Halliwell, J. Perez-Mercader, and W. Zurek, *The Physical Origins of Time Asymmetry*, 1994; P. Horwich, *Asymmetries in Time*, 1987; G. J. Whitrow, *The Natural Philosophy of Time*, 1980.

Time constant

A characteristic time that governs the approach of an exponential function to a steady-state value. When a physical quantity is varying as a decreasing exponential function of time as in Eq. (1), or as an increasing exponential function as in Eq. (2) [see **illus.**], the approach to the steady-state value achieved after a long time is governed by a characteristic time T as given in Eq. (3). This time T is called the time constant.

$$f(t) = e^{-kt} \tag{1}$$

$$f(t) = 1 - e^{-kt} \tag{2}$$

$$t = \frac{1}{k} = T \tag{3}$$

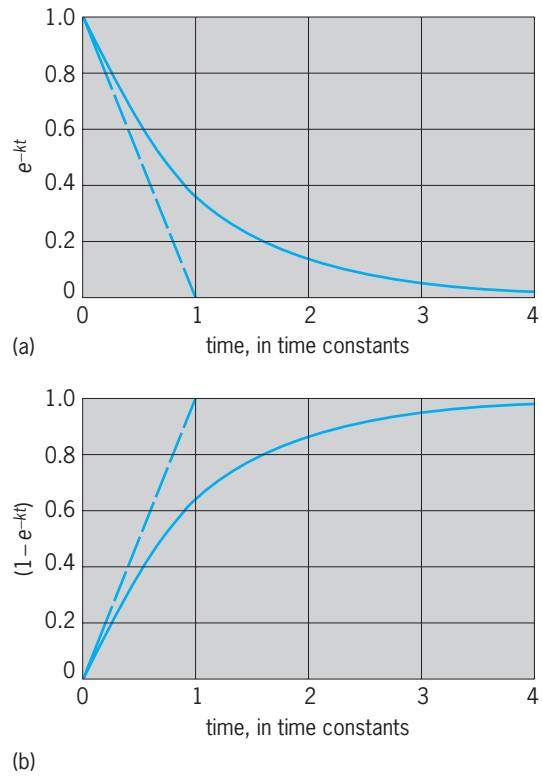

Universal time-constant curve indicated (*a*) for the decreasing function and (*b*) for the increasing function.

When time t is zero, $f(t)$ in Eq. (1) has the magnitude 1, and when t equals T the magnitude is $1/e$. Here e is the transcendental number whose value is approximately 2.71828, and the change in magnitude is $1 - (1/e) = 0.63212$. The function has moved 63.2% of the way to its final value. The same factor also holds for Eq. (2). *See* E (MATHEMATICS).

The initial rate of change of both the increasing and decreasing functions is equal to the maximum amplitude of the function divided by the time constant. Parts *a* and *b* of the illustration are universal in that the plotted function is of unit height and the time scale is given in terms of time constants. To use these curves for a specific problem, the values in the ordinate axis are multiplied by the maximum amplitude of the quantity occurring in the problem, and the values in the abscissa axis are multiplied by the numerical value of the corresponding time constant.

The concept of time constant is useful when evaluating the presence of transient phenomena. The relative amplitude of a transient after an elapsed time of a certain number of time constants is readily computed:

Elapsed time, time constants	*Transient completed, %*
1	63.2
2	86.5
3	95.0
4	98.2
5	99.3
10	99.996

Usually a transient can be considered as being over after a period of 4–5 time constants.

For electric circuits, the coefficient k and thus the time constant T is determined from the parameters of the circuit. For a circuit with resistance R and capacitance C, the time constant T is the product RC. When the circuit consists of inductance L and resistance R, the time constant is L/R. *See* ELECTRIC TRANSIENT.

The concept of time constant can be applied to the transient envelope of an ac signal; however, it is more common to describe the change in amplitude in terms of logarithmic decrement. For further discussion *See* DAMPING. Robert L. Ramey

Time-interval measurement

A determination of the duration between two instants of time (epochs). Time intervals are measured with high precision with a digital display counter. An electronic oscillator generates pulses; the count begins with a start signal and ends with a second signal. For an oscillator frequency of 100 MHz, for example, a direct reading is correct to 10 nanoseconds (1 ns = 10^{-9} s). Two atomic clocks, however, can be compared in epoch to 1 picosecond (1 ps = 10^{-12} s) by electronic interpolation. *See* ATOMIC CLOCK; DIGITAL COUNTER; OSCILLATOR; OSCILLOSCOPE.

Rapid motions can be studied at short intervals by means of a large variety of high-speed cameras, including stroboscopic, rotating film-drum, rotating mirror, streak, and image converter cameras. In one camera a helium turbine rotates a mirror at 20,000 revolutions per second to form 130 frames at 25,000,000 frames per second for 5.2 microseconds (1 μs = 10^{-6} s). The framing interval is 40 ns. An electronic streak camera can separate two pulses 1 ps apart. *See* PHOTOGRAPHY; STROBOSCOPIC PHOTOGRAPHY.

Ultrashort laser pulses are used to study rapid processes caused by the interaction of photons with an atom or molecule. The duration of interaction is $\tau = L/c$, where L is the pulse length and c the speed of light. One technique splits the pulse into two pulses. One pulse excites a reaction, and the other, optically delayed, probes the reaction. The probe pulse can be split into several pulses, staggered in time, with an echelon. Differences in path length give differences in time. Pulses as short as three wavelengths of 620-nm light, with $\tau = 6$ femtoseconds (1 fs = 10^{-15} s), have been formed. *See* LASER; LASER PHOTOCHEMISTRY; OPTICAL PULSES; ULTRAFAST MOLECULAR PROCESSES.

Radioactive decay is used to measure long time intervals, to about 5×10^9 years, concerning human history, the Earth, and the solar system. *See* GEOCHRONOMETRY; RADIOCARBON DATING. William Markowitz

Bibliography. J. C. Diels and W. Rudolp, *Ultrashort Laser Pulse Phenomena*, 2d ed., 2006; Eastman Kodak Co., *Encyclopedia of Practical Photography*, 1978; H. E. Edgerton, *Electronic Flash, Strobe*, 3d ed., 1987; L. Stroebel and R. D. Zakia (eds.), *The Focal Encyclopedia of Photography*, 3d ed., 1996.

Time-of-flight spectrometers

A general class of instruments in which the speed of a particle is determined directly by measuring the time that it takes to travel a measured distance. By knowing the particle's mass, its energy can be calculated. If the particles are uncharged (for example, neutrons), difficulties arise because standard methods of measurement (such as deflection in electric and magnetic fields) are not possible. The time-of-flight method is a powerful alternative, suitable for both uncharged and charged particles, that involves the measurement of the time t that a particle takes to travel a distance l. If the rest mass of the particle is m_0, its kinetic energy E_T can be calculated from its measured speed, $v = l/t$, using the equation below, where c is the speed of light.

$$E_T = m_0c^2\left\{\left[1-\left(\frac{v}{c}\right)^2\right]^{-1/2}-1\right\}$$

$$\approx \frac{m_0v^2}{2} \qquad \text{if } v$$

Some idea of the time scales involved in measuring the energies of nuclear particles can be gained by noting that a slow neutron of kinetic energy $E_T =$ 1 eV takes 72.3 microseconds to travel 1 m. Its flight time along a 10-m path (typical of those found in practice) is therefore 723 μs, whereas a 4-MeV neutron takes only 361.5 nanoseconds.

The time intervals are best measured by counting the number of oscillations of a stable oscillator that occur between the instants that the particle begins and ends its journey (see **illus.**). Oscillators operating at 100 MHz are in common use. If the particles from a pulsed source have different energies, those with the highest energies arrive at the detector first. Digital information from the "gated" oscillator consists of a series of pulses whose number $N(t)$ is proportional to the time-of-flight t. These pulses can be counted and stored in an on-line computer that provides many thousands of sequential "time channels," $t_0, t_0 + \Delta t, t_0 + 2\Delta t, t_0 + 3\Delta t, \ldots$, where t_0 is the time at which the particles are produced and Δt is the period of the oscillator. To store an event in channel $N(t)$, the contents of memory address $N(t)$ are updated by "adding 1."

Time-of-flight spectrometers have been used for energy measurements of uncharged and charged elementary particles, electrons, atoms, and molecules. The popularity of these instruments is due to the broad energy range that can be covered, their high resolution ($\Delta E_T/E_T \approx 2\Delta t/t$, where ΔE_T and Δt are the uncertainties in the energy and time measurements, respectively), their adaptability for studying different kinds of particles, and their relative

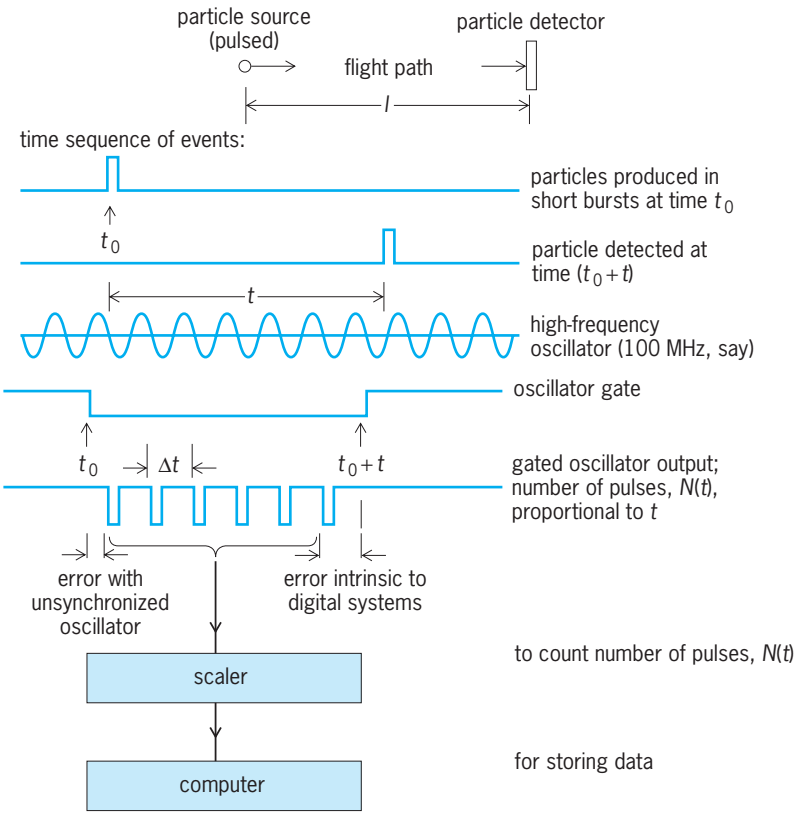

Schematic diagram of a time-of-flight spectrometer.

simplicity. *See* MASS SPECTROSCOPE; NEUTRON SPECTROMETRY.

Frank W. K. Firk

Bibliography. R. J. Cotter (ed.), *Time-of-Flight Mass Spectroscopy*, 1994; E. W. Schlag (ed.), *Time-of-Flight Mass Spectroscopy and Its Applications*, 1994.

Time-projection chamber

An advanced particle detector for the study of ultrahigh-energy collisions of positrons and electrons that was developed originally at the Lawrence Berkeley Laboratory. The underlying physics of the scattering process can be studied through precise measurements of the momenta, directions, particle species, and correlations of the collision products. The time-projection chamber (TPC) provides a unique combination of capabilities for these studies and other problems in elementary particle physics by offering particle identification over a wide momentum range, and by offering high resolution of intrinsically three-dimensional spatial information for accurate reconstruction of events.

The time-projection chamber concept is based on the maximum utilization of ionization information, which is deposited by high-energy charged particles traversing a gas. The ionization trail, a precise image of the particle trajectory, also contains information about the particle velocity. A strong, uniform magnetic field and a uniform electric field are generated within the time-projection chamber active volume in an exactly parallel orientation. The parallel configuration of the fields permits electrons, products of the ionization processes, to drift through the time-projection chamber gas over great distances without distortion; the parallel configuration offers a further advantage in that the diffusion of the electrons during drift can be greatly suppressed by the magnetic field, thus preserving the quality of track information. In practice, the track images are drifted on the order of 100 cm (40 in.) or more, yet with measurement precision typically better than ±0.02 cm.

At the end of the drift volume the ionization electrons are multiplied by an avalanche process on an array of several hundred wires acting as proportional amplifiers. A highly segmented cathode plane just behind the wire array detects the avalanches and provides two-dimensional spatial coordinates in the plane of the array. The drift time provides the trajectory coordinate perpendicular to the plane of the array, hence suggesting the name time-projection chamber. The ionization density, measured precisely by the wire plane signals, offers the means to determine the particle velocity with resolution sufficient to establish the particle mass by a comparison of velocity and momentum.

Several large time-projection chambers are in operation or under construction at the premier storage ring facilities in the United States, Europe, and Japan. *See* ELECTRICAL BREAKDOWN; PARTICLE DETECTOR.

David R. Nygren

Bibliography. J. A. Macdonald (ed.), *The Time Projection Chamber*, AIP Conf. Proc. 108, 1984; J. N. Marx and D. R. Nygren, The time projection chamber, *Phys. Today*, 31:46–53, October 1978.

Time reversal invariance

A symmetry of the fundamental (microscopic) equations of motion of a system; if it holds, the time reversal of any motion of the system is also a motion of the system. To date, only two phenomena have shown evidence (at least indirect) for violation of time reversal invariance. One is the violation of *CP* invariance observed in the decays of the neutral mesons K_L and B^0, $\bar{B}^0$. The other is the baryon asymmetry of the universe.

Time reversal invariance is not evident from casual observation of everyday phenomena. If a movie is taken of a phenomenon, the corresponding time-reversed motion can be exhibited by running the movie backward. The result is usually strange. For instance, water in the ground is not ordinarily observed to collect itself into drops and shoot up into the air. However, if the system is sufficiently well observed, the direction of time is not obvious. For instance, a movie which showed the motion of the planets (in which each of the objects that make up the system can be observed individually), would look just as right run backward or forward. The apparent irreversibility of everyday phenomena results from the combination of imprecise observation and starting from an improbable situation (a state of low entropy, to use the terminology of statistical mechanics). *See* ENTROPY; STATISTICAL MECHANICS.

The known fundamental equations of motion are all time-reversal invariant. For instance, suppose that $\vec{r}_j(t)$ and $\vec{p}_j(t)$, the coordinates and momenta of charged particles, and $\vec{E}(\vec{r}, t)$ and $\vec{B}(\vec{r}, t)$, the electric and magnetic fields, satisfy the equations of motion of the particles and the fields (Newton's and Maxwell's equations); that is, these functions of time describe a motion of the system. Then the functions given in Eqs. (1), which describe the time reversal of the original motion, also satisfy the equations of motion. Thus the system is time-reversible. *See* MAXWELL'S EQUATIONS; NEWTON'S LAWS OF MOTION.

$$\begin{aligned} \vec{r}_{jREV}(t) &= \vec{r}_j(-t) \\ \vec{p}_{jREV}(t) &= -\vec{p}_j(-t) \\ \vec{E}_{REV}(\vec{r}, t) &= \vec{E}(\vec{r}, -t) \\ \vec{B}_{REV}(\vec{r}, t) &= -\vec{B}(\vec{r}, -t) \end{aligned} \tag{1}$$

Quantum mechanics. In quantum mechanics, if the hamiltonian operator H is independent of time (energy-conserving) and real, and if the wave function $\psi(t)$ is a solution of the Schrödinger equation (2), where $\hbar$ is Planck's constant divided by 2π and ∂_t is partial differentiation with respect to time, then the function given in Eq. (3),

$$i\hbar\partial_t\psi = H\psi \tag{2}$$

$$\psi_{REV}(t) = \psi^*(-t) \tag{3}$$

where the asterisk indicates complex conjugation, is also a solution of the Schrödinger equation. The function ψ_{REV} is the wave function of the time reversal of the motion described by ψ; hence the Schrödinger equation with a constant and real Hamiltonian is time reversal–invariant. In the time-reversed motion, other kinds of amplitudes, for example, transition amplitudes, are also complex-conjugated. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; SCHRÖDINGER'S WAVE EQUATION.

Tests. If time reversal invariance holds, no particle (a physical system with a definite mass and spin) can have an electric dipole moment $\vec{d}$, that is, an interaction energy of the form $-\vec{d}\cdot\vec{E}$, where $\vec{E}$ is the applied electric field. This is because the only intrinsic vector quantity of such a system (that is, the only vector quantity with nonvanishing expectation value in the rest frame) is the spin $\vec{S}$; thus an electric dipole moment $\vec{d}$ would have to be a multiple of $\vec{S}$, resulting in an interaction energy of the form $c\vec{S}\cdot\vec{E}$, where c is a constant. This energy changes sign under time reversal, because $\vec{S}$ does but $\vec{E}$ does not. (Since spin is an axial vector, an electric dipole moment of a particle would also violate space inversion symmetry or parity.) Although a polar body, for example, a water (H_2O) molecule, has an electric dipole moment, this shows up only in transition matrix elements between its eigenstates of energy (and spin). No particle has been observed to have an electric dipole moment; for instance, the present experimental upper limit on the electric moments of the electron and the neutron are approximately 10^{-27} cm times e and 10^{-25} cm times e, respectively, where e is the charge of the proton. *See* DIPOLE MOMENT; ELECTRON; NEUTRON; PARITY (QUANTUM MECHANICS); POLAR MOLECULE; SPIN (QUANTUM MECHANICS).

Another test of time reversal invariance is to compare the cross sections for reactions which are inverse to one another, for example, the reactions $^{16}O + d \leftrightarrow {}^{14}N + \alpha$. The present experimental upper limit on the relative size of the time reversal invariance–violating amplitude of such reactions is approximately 3×10^{-3}; unfortunately, this is far larger than any expected violation. *See* NUCLEAR REACTION.

A class of tests involves looking at the relative phases of amplitudes. For instance, if a nuclear transition emits both electric quadrupole and magnetic dipole electromagnetic radiation (gamma rays), certain interference terms in the angular distribution of the radiation cannot occur, because of the relative phase of the two amplitudes imposed by time reversibility. Experiments looking for such effects put an upper limit of approximately 10^{-3} on the relative size of a time reversal invariance–violating amplitude. *See* MULTIPOLE RADIATION.

Evidence for violation. A consequence of the *CPT* theorem is that violation of *T* (time reversal

invariance) is equivalent to violation of *CP*, that is, invariance of the fundamental equations under the combined operations of charge conjugation *C* and space inversion *P*. Hence *CP* violation observed in the decay of the long-lived neutral *K* meson (K_L) and in the decay of the neutral *B* mesons is evidence for *T* violation. *See* CPT THEOREM; MESON.

The K^0 and $\bar{K}^0$ mesons differ only in their strangeness, +1 and −1, respectively. The charged current weak interaction (the exchange of a charged weak boson) changes quark flavors, including strangeness, and so in second order (that is, acting twice) it can turn a K^0 into a $\bar{K}^0$ or vice versa. Consequently a K^0 or $\bar{K}^0$ is not a mass eigenstate but a coherent mixture of two mass eigenstates; conversely, the two mass (and decay rate) eigenstates are coherent mixtures of K^0 and $\bar{K}^0$. If *CP* were conserved, these mass eigenstates would be *CP* eigenstates, the more rapidly decaying one (called K_S, *S* for short-lived) having $CP = +1$, the *CP* value of the dominant decay mode of a neutral *K*, namely two pions, and the other (called K_L, *L* for long-lived) having $CP = -1$, thus unable to decay to two pions. But in fact the K_L has a small but nonvanishing branching ratio, ∼0.3%, for decay into two pions; this was the first observation of *CP* violation. Other evidences of *CP* violation are that the K_L has unequal branching ratios for decay into the *CP* conjugate states $\pi^+ e \bar{\nu}_e$ (a decay mode of $\bar{K}^0$) and $\pi^- \bar{e} \nu_e$ (a decay mode of K^0), and that the probabilities at short decay times for an original K^0 to decay to $\pi^+ e \bar{\nu}_e$ and for an original K^0 to $\pi^- \bar{e} \nu_e$, are unequal. *See* ELECTROWEAK INTERACTION; FLAVOR; QUARKS; STANDARD MODEL; WEAK NUCLEAR INTERACTIONS.

Similarly, B^0 and $\bar{B}^0$ mesons are mixtures of mass eigenstates; but no one decay mode dominates, so the lifetimes of the two mass eigenstates are not very different, and there is no simple way of experimentally distinguishing them. The observed *CP* violation is equivalent to saying that a $\bar{B}^0$ is about twice as likely to decay to $J/\psi\ K_S$ as is a B^0. This is not directly observable because *B* factories create (*B* and *B*) mesons in equal numbers, using the resonance reaction $e\bar{e} \rightarrow \Upsilon(4S) \rightarrow B^0\bar{B}^0$, which creates B^0 and $\bar{B}^0$ mesons in equal numbers. But if one of the two mesons created together is observed to decay into a negatively (positively) charged lepton, *e* or μ, plus other particles, it was a $\bar{B}^0(B^0)$. Hence at that time the other meson was a $B^0(\bar{B}^0)$, and by comparing its observed decay rate into $J/\psi\ K_S$ as a function of time (relative to the time of the semileptonic decay) to the theoretical formula, a *CP* violation parameter is deduced. Qualitatively described: In events in which one of the members of the $B^0\bar{B}^0$ pair decays semileptonically and the other decays into the hadronic mode $J/\psi\ K_S$, the *CP* violation is the observation that if the lepton of the semileptonic decay is positive, that decay is more likely to precede the $J/\psi\ K_S$ decay, whereas if negative, it is more likely to follow. *See* PARTICLE ACCELERATOR; UPSILON PARTICLES.

According to the standard model, *CP* violation arises from the CKM (Cabibbo-Kobayashi-Maskawa) quark-mixing matrix. When a down-type quark (*d*, *s*, or *b*) absorbs a W^+ or emits a W^- boson, it becomes a linear combination of up-type quarks (*u*, *c*, *t*); the coefficients of these three linear combinations is called the CKM matrix. (The same matrix also describes the linear combinations of down-type quarks resulting from absorption of a W^- or emission of a W^+ by an up-type quark.) It is a 3×3 unitary matrix with determinant 1. The unitarity of the CKM matrix expresses a universality of the weak interaction. For instance, a strange quark, *s*, becomes a *u*, *c*, or *t* quark with relative probabilities of 5%, 95%, and 0.2%, respectively. Because of energy conservation, in the decay of a strange particle such as K^+, the *c* or the *t* could only be virtual. The *u* could be real, that is, a constituent of a decay product; the coefficient 0.22 gives a K^+ decay rate about 5% of that predicted from the strength of the weak interaction as seen in ordinary β decay, where the quark transition is $|d> \rightarrow 0.975|u> + \cdots$. In other words, the CKM matrix gives the relative couplings of the charged weak boson to quarks. In general (that is, for a generic CKM matrix), all the couplings cannot be made real by any choice of the phases of the quark states; this results in violation of *T*. (The original quark-mixing matrix was a 4×4 matrix, proposed by N. Cabibbo, before the discovery of the third weak doublet, *t,b*; a 4×4 quark-mixing matrix never implies *T* violation. M. Kobayashi and T. Maskawa noted that the extension to a 3×3 matrix could give *T* violation in a natural way.) The CKM matrix consistently describes all weak-decay observations, including the *CP* violation seen in neutral *B* decay. *See* MATRIX THEORY.

An indirect but very prominent evidence for *T* violation is the baryon asymmetry of the universe, that is, the fact that ordinary matter contains baryons, not antibaryons. Put more quantitatively: The observed fact that there is roughly 1 baryon for every 10^9 cosmic blackbody photons means that early in the history of the universe, when the value of kT (the product of Boltzmann's constant and the thermodynamic temperature) was larger than 1 GeV (and hence the number of both baryons and antibaryons was roughly the same as the number of photons), there was an excess of baryons over antibaryons of the order of one part in 10^9. It is conceivable that this rather small baryon-number asymmetry developed in an originally baryon-number symmetric universe. This would require interactions, acting in the early universe, which violate both baryon number and *T*. The existence of the former would not be surprising, since particle interactions that do not conserve baryon number are always present in grand unified theories—gauge theories in which the strong and electroweak interactions, and likewise quarks and leptons, are unified. *See* COSMIC BACKGROUND RADIATION; COSMOLOGY; ELEMENTARY PARTICLE; GRAND UNIFICATION THEORIES; SYMMETRY LAWS (PHYSICS).

Charles J. Goebel

Bibliography. I. I. Bigi and A. I. Sanda, *CP Violation*, Cambridge University Press, 2000; W. M. Gibson and B. R. Pollard, *Symmetry Principles in Elementary Particle Physics*, 1976, paper 1980; E. M.

Henley, Parity and time-reversal invariance in nuclear physics, *Annu. Rev. Nucl. Sci.*, 19:367–432, 1969; I. B. Khriplovich and S. K. Lamoreaux, *CP Violation without Strangeness*, Springer-Verlag, 1997; K. Kleinknecht, *Uncovering CP Violation*, Springer-Verlag, 2003; D. Park, *Introduction to the Quantum Theory*, 3d ed., McGraw-Hill, 1992, reprint, Dover, 2005; R. G. Sachs, *The Physics of Time Reversal*, University of Chicago Press, 1987; M. Skalsey et al. (eds.), *Time Reversal*, AIP Press, 1993; L. Wolfenstein (ed.), *CP Violation*, Elsevier Science, 1990.

Time-reversed signal processing

A means for improving the performance of remote sensing and communication systems that rely on electromagnetic- or acoustic-wave propagation but must contend with wave reflections, diffraction, and scattering. In particular, for applications of remote sensing—including radar, sonar, biomedical imaging, and nondestructive evaluation—the main intent is the detection, localization, and identification of distant objects or features that either scatter or generate electromagnetic or acoustic waves. Although the disparity in objectives between military-radar and ultrasonic-imaging systems may be considerable, they have in common that random wave scattering and diffraction tend to limit the accuracy and confidence with which such systems can be used and the distances over which such systems can operate. For radar and sonar, turbulence and wave motions in the ocean or atmosphere, or rough terrestrial or ocean surfaces, may cause such random wave scattering and diffraction. In ultrasonic remote sensing, the random scattering and diffraction may be caused by grain structure within metals or by variations between and within tissues. *See* REMOTE SENSING.

Applications. Exploitation of time reversal in wave propagation problems dates back to the early 1960s, when similar concepts arose in the fields of geoacoustics, radio waves, and underwater sound propagation. In 1965, Antares Parvulescu and Clarence Clay published a study in which acoustic signals were recorded with a single transducer 20 nautical miles (37 km) from a sound source in water that was approximately 1 nautical mile (1.85 km) deep. The recorded signals were time-reversed, and retransmitted through the ocean. The signal received after the second transmission was much clearer and far less distorted than the signal received after the first transmission. Since that time, the sophistication of transmitters, receivers, and signal processing algorithms has allowed time reversal concepts to permeate nearly every application of remote sensing that relies on wave propagation to or from an object or feature of interest. Components for aircraft engines have been inspected with time-reversing ultrasonic arrays. The U.S. Navy is considering time-reversal concepts for new active sonar systems and underwater communication links. Time-reversal concepts are now being applied in biomedical diagnostic ultrasound to increase the clarity of images and in therapeutic ultrasound to better target tumors and kidney stones. Feasibility studies are getting underway for radar systems that incorporate time reversal to enhance operating distances and detection of targets under foliage, and the future of time-reversed signal processing is likely to include structural monitoring of buildings and machines. *See* ACOUSTIC SIGNAL PROCESSING; BIOMEDICAL ULTRASONICS; NONDESTRUCTIVE EVALUATION; RADAR; SONAR; ULTRASONICS; UNDERWATER SOUND.

Time-reversal process. The basic process of time reversal can be described by four steps (**Fig. 1**). First, waves are generated or scattered by an object or feature of interest and travel forward through the environment to an antenna or array of transducers. The paths that waves follow between the object and the array may be complicated and unknown. Second, the array records the signal in the usual manner. These recordings will include signal distortion from echoes, scattering, and diffraction in the environment. Third, the signal recorded by each element in the array is retransmitted from that element with the direction of time inverted; the end of the signal is transmitted first, and the start of the signal is transmitted last. In the final step, these array-transmitted time-reversed waves travel backward through the environment, retracing their paths to converge at the location where they originated. Although specific applications of time reversal typically involve more steps and greater processing of the array-received signals, all are based, directly or indirectly, on these simple steps.

When time reversal is working properly, the waves that return to the location of their origin focus tightly and are undistorted, even though they may have passed through a complicated environment that generates echoes and causes random scattering and diffraction. In fact, the size of the focal region may be much smaller when the environment is complicated than when it is uniform in all directions. Both the tight focusing and distortion-removal capabilities of time reversal are of interest in remote sensing.

Distortion removal and focusing. The time-reversed waves are able to remove distortion and focus well when the transducer array is many wavelengths long, the absorption of wave energy by the environment is weak, background noise is low, and the environment changes little (or not at all) between forward and backward wave travel. The process by which time reversal accomplishes distortion removal may be understood by considering an environment containing three travel paths—fast, middle, and slow—with different travel times between the object and the transducer array. In this case, a signal that starts at the object will be received at the array as three possibly overlapping signals. This phenomenon is called multipath distortion, and it occurs when there are well-defined wave-travel paths and when there is wave scattering from particles, turbulence, or other fluctuations between the object and the transducer array. In the three-path environment, the time-reversed broadcast will launch the signal on the slow

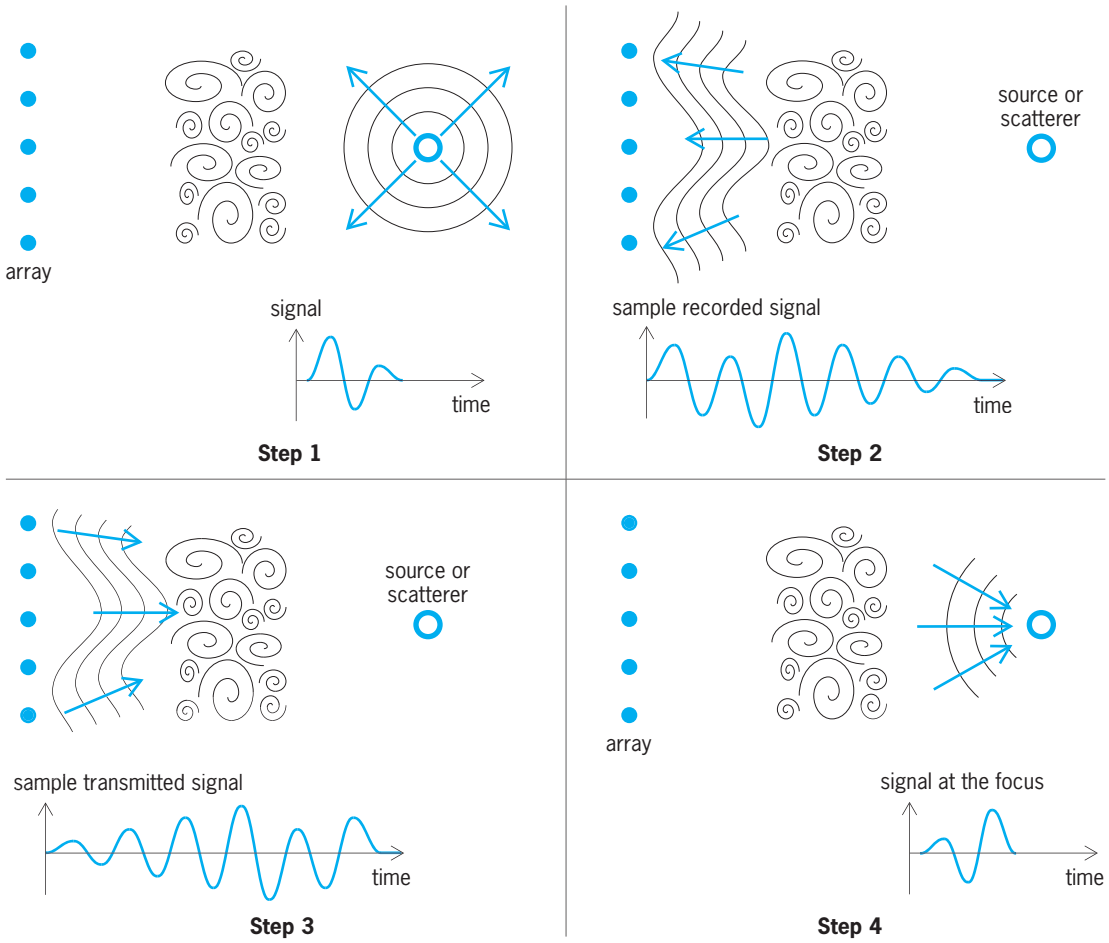

Fig. 1. Focusing and distortion compensation of waves that travel through an unknown random medium with time-reversed signal processing. Step 1: signal generation. Step 2: ordinary recording. Step 3: time-reverse playback. Step 4: focus formation.

path first, the middle path second, and the fast path last. Here, the timing of the signal launches will exactly undo the multipath distortion, and the signal will arrive undistorted back at its place of origin.

The superior focusing characteristics of time-reversed waves come from their ability to exploit reflection, scattering, and diffraction within the environment. The range of angles through which waves converge determines how tightly the converging waves focus. If all the time-reversed waves come from the same direction, the focus is larger than if the waves converge from above, from below, and from either side. When the environment is uniform, there is only one travel path between the object (a source or scatterer) and each element of the transducer array, so the focus size of the time-reversed waves is determined by the array's angular aperture. When the environment produces echoes, random scattering, and diffraction, there may be many travel paths between the object and the transducer array, and some of these paths can be used by the time-reversed waves to increase the range of convergence angles occurring during backward travel to decrease the focus size compared to what would occur in a uniform environment (**Fig. 2**). However, when the environment changes between forward and backward wave travel, the advantages of time reversal may be degraded or even lost.

Techniques. There are several signal processing techniques that have been used to convert time reversal from an alluring oddity of wave physics into a useful remote sensing tool. These techniques generally fall into two categories: those that require an active transducer array that can transmit and receive waves, and those that merely require a passive receiving array.

Active techniques. The most popular of the active techniques, referred to by the French acronym DORT (décomposition de l'operateur de retournement temporel; decomposition of the time-reversal operator), was developed by Claire Prada and Mathias Fink for nondestructive evaluation and biomedical ultrasound applications. This method can be used to detect and separately illuminate distinct scatterers. In its simplest implementation, DORT requires the measurement of the object-scattered signal at every array element when the object is separately illuminated by each array element. This carefully constructed measurement matrix can then be mathematically analyzed to determine the number or scattering objects, and the directions toward the objects in an environment with some random scattering and

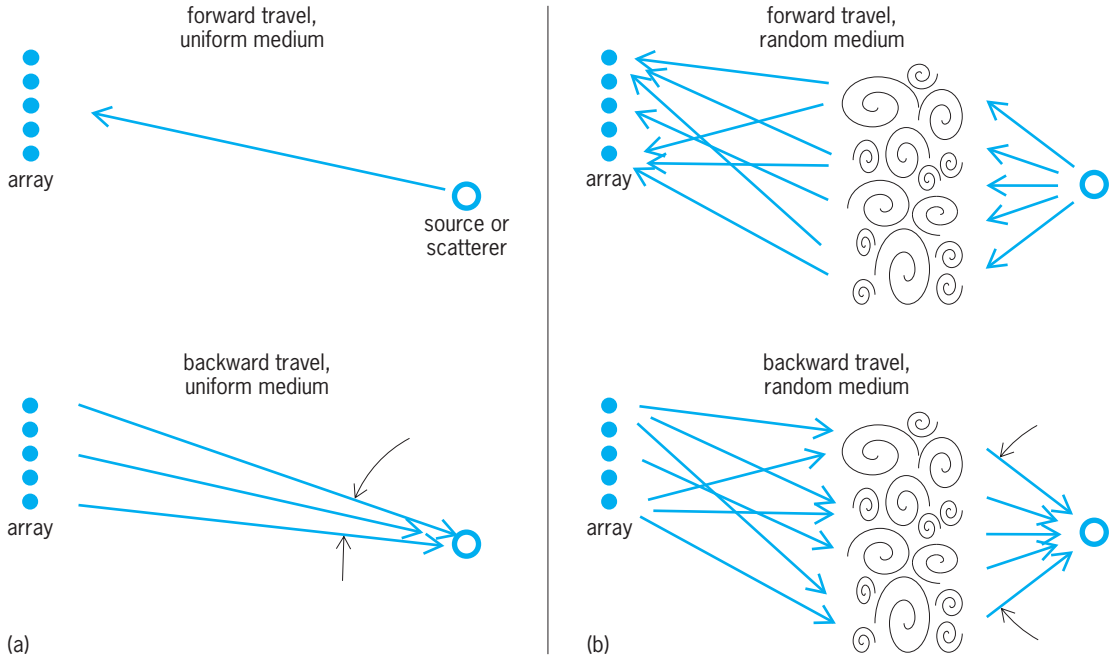

Fig. 2. Time-reversed focusing. (*a*) Without random wave scattering. (*b*) With random wave scattering.

diffraction but no strong echoes. In such an environment, an image of the scattering objects may be formed using DORT.

A second active technique may be combined with DORT and involves identifying and selectively time-reversing segments of the received signal that may contain scattered wave energy from the object or feature of interest. When the correct signal segment is time-reversed and transmitted, the waves that return to the transducer array will concentrate or peak at one particular time when an item of interest is present. When the rebroadcast signal segment corresponds to an ordinary section of the medium, the returning waves will not concentrate at any particular time. This technique has been used to detect the presence of internal flaws in titanium, an important aerospace material.

Other active array techniques have been proposed for underwater acoustic communication, security barriers, and reverberation reduction in active sonar systems.

Passive techniques. The passive array signal processing techniques based on time reversal are also important. The best known of these is matched-field processing (MFP). Here, the array-received signals are delivered to a computer program that can simulate the backward wave travel. The computed location where the backward traveling waves converge is the presumed location of the source or scatterer. Thus, this technique can be used by the computer operator to locate remote sources or scatterers based on the array reception alone; a transmission by the array is unnecessary. Although matched-field processing can theoretically locate or image objects with subwavelength accuracy, this technique cannot be implemented successfully without enough environmental information to ensure the accuracy of the computer model. *See* MATCHED-FIELD PROCESSING.

There are also passive time-reversed signal processing techniques that do not require any knowledge of the environment. One of these, passive phase conjugation (PPC), may be used for underwater communication between a cooperating remote source and a receiving array. If the source wishes to send a coded information stream to the array, it first sends a single pulse to characterize the multipath distortion of the environment. When the source starts its coded-information broadcast, the receiving array uses the measurements of the distorted single-pulse signal to unravel the distortion put into the coded message by the environment. Similar methods are used by modems to correct for variations in telephone lines. Another promising passive array technique, artificial time reversal, is similar to PPC but does not require the initial single-pulse broadcast.

David R. Dowling

Bibliography. J. Berryman et al., Statistically stable ultrasonic imaging in random media, *J. Acous. Soc. Amer.*, 112:1509–1522, 2002; D. R. Jackson and D. R. Dowling, Phase-conjugation in underwater acoustics, *J. Acous. Soc. Amer.*, 89:171–181, 1991; E. Kerbrat et al., Imaging in the presence of grain noise using the decomposition of the time reversal operator, *J. Acous. Soc. Amer.*, 113:1230–1240, 2003; W. A. Kuperman et al., Phase-conjugation in the ocean: Experimental demonstration of an acoustic time reversal mirror, *J. Acous. Soc. Amer.*, 103: 25–40, 1998; N. Mordant, C. Prada, and M. Fink, Highly resolved detection and selective focusing in a waveguide using the D.O.R.T. method, *J. Acous. Soc. Amer.*, 105:2634–2642, 1999; A. Parvulescu and C. S. Clay, Reproducibility of signal transmissions in the ocean, *Radio Electr. Eng.*, 29:223–228, 1965.

Timothy

A plant, *Phleum pratense*, of the order Cyperales, long the most important hay grass for the cooler temperate humid regions. It is easily established and managed, produces seed abundantly, and grows well in mixtures with alfalfa and clover. It is a short-lived perennial, makes a loose sod, and has moderately leafy stems 2–4 ft (0.6–1.2 m) tall and a dense cylindrical inflorescence (see **illus.**). Timothy responds to fertile soils with high yield and nutritive content. Cutting promptly after heading improves the feed quality. Timothy-legume mixtures still predominate in hay and pasture seedings for crop rotations in the northern half of the United States, but orchard grass and bromegrass have increasingly replaced timothy in such mixtures in many areas. *See* ALFALFA; CLOVER; GRASS CROPS; INFLORESCENCE. Howard B. Sprague

Timothy (*Phleum pratense*).

Tin

A chemical element, symbol Sn, atomic number 50, atomic weight 118.69. Tin forms tin(II) or stannous (Sn^{2+}), and tin(IV) or stannic (Sn^{4+}) compounds, as well as complex salts of the stannite (M_2SnX_4) and stannate (M_2SnX_6) types. *See* PERIODIC TABLE.

Tin melts at a low temperature, is highly fluid when molten, and has a high boiling point. It is soft and pliable and is corrosion-resistant to many media. An important use of tin has been for tin-coated steel containers (tin cans) used for preserving foods and beverages. Other important uses are solder alloys, bearing metals, bronzes, pewter, and miscellaneous industrial alloys. Tin chemicals, both inorganic and organic, find extensive use in the electroplating, ceramic, plastic, and agricultural industries.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

The most important tin-bearing mineral is cassiterite, SnO_2. No high-grade deposits of this mineral are known. The bulk of the world's tin ore is obtained from low-grade alluvial deposits. *See* CASSITERITE.

Two allotropic forms of tin exist: white (β) and gray (α) tin. Tin reacts with both strong acids and strong bases, but it is relatively resistant to solutions that are nearly neutral. In a wide variety of corrosive conditions, hydrogen gas is not evolved from tin and the rate of corrosion becomes controlled by the supply of oxygen or other oxidizing agents. In their absence, corrosion is negligible. A thin film of stannic oxide forms on tin upon exposure to air and provides surface protection. Salts that have an acid reaction in solution, such as aluminum chloride and ferric chloride, attack tin in the presence of oxidizers or air. Most nonaqueous liquids, such as oils, alcohols, or chlorinated hydrocarbons, have slight or no obvious effect on tin. Tin metal and the simple inorganic salts of tin are nontoxic. Some forms of organotin compounds, on the other hand, are toxic. Some important physical constants for tin are shown in the **table**.

Stannous oxide, SnO, is a blue-black, crystalline product which is soluble in common acids and strong alkalies. It is used in making stannous salts for plating and glass manufacture. Stannic oxide, SnO_2, is a white powder, insoluble in acids and alkalies. It is an excellent glaze opacifier, a component of pink, yellow, and maroon ceramic stains and of dielectric and refractory bodies. It is an important polishing agent for marble and decorative stones.

Stannous chloride, $SnCl_2$, is the major ingredient in the acid electrotinning electrolyte and is an intermediate for tin chemicals. Stannic chloride, $SnCl_4$, in the pentahydrate form is a white solid. It is used in the preparation of organotin compounds and chemicals to weight silk and to stabilize perfume and colors in soap. Stannous fluoride, SnF_2, a white water-soluble compound, is a toothpaste additive.

Organotin compounds are those compounds in which at least one tin-carbon bond exists, the tin usually being present in the + IV oxidation state. Organotin compounds that find applications in industry are the compounds with the general formula R_4Sn, R_3SnX, R_2SnX_2, and $RSnX_3$. R is an organic group, often methyl, butyl, octyl, or phenyl, while X is an inorganic substituent, commonly chloride, fluoride,

Properties of tin*

Property	Value
Melting point	231.9°C (449.4°F)
Boiling point	2270°C (4118°F)
Specific gravity,	
α-form (gray tin)	5.77
β form (white tin)	7.29
Liquid at melting point	6.97
Transformation temperature, °C	13.2
Specific heat, cal/g,	
white tin at 25°C	0.053
gray tin at 10°C	0.049
Latent heat of fusion, cal/g	14.2
Latent heat of vaporization, cal/g	520 ± 20
Heat of transformation, cal/g	4.2
Thermal conductivity, cal/(cm)(cm^2)°C(s), white tin at °C	0.150
Coefficient of linear expansion at 0°C	19.9×10^{-6}
Shrinkage on solidification, %	2.8
Resistivity of white tin, microhms/cm^3	
at 0°C	11.0
at 100°C	15.5
Brinell hardness, 10 kg/(5 mm)(180 s)	
at 20°C	3.9
at 220°C	0.7
Tensile strength as cast, lb/in.^{2}	
at 15°C	2100
at 200°C	650
at −40°C	2900
at −120°C	12,700

*1 cal = 4.184 J, 1 lb/in.^{2} = 6.895 kPa, °F = (°C × 1.8) + 32.

oxide, hydroxide, carboxylate, or thiolate. *See* TIN ALLOYS. Joseph B. Long

Bibliography. P. W. Atkins et al., *Inorganic Chemistry*, 4th ed., 2006; F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., 1999; D. R. Lide, *CRC Handbook of Chemistry and Physics*, 85th ed., 2004.

Tin alloys

Solid solutions of tin and some other metal or metals. Alloys cover a wide composition range and many applications because tin alloys readily with nearly all metals. *See* ALLOY.

Soft solders constitute one of the most widely used and indispensable series of tin-containing alloys. Common solder is an alloy of tin and lead, usually containing 20–70% tin. It is made easily by melting the two metals together. With 63% tin, a eutectic alloy melting sharply at 361°F (169°C) is formed. This is much used in the electrical industry. A more general-purpose solder, containing equal parts of tin and lead, has a melting range of 56°F (31°C). With less tin, the melting range is increased further, and wiping joints such as plumbers make can be produced. Lead-free solders for special uses include tin containing up to 5% of either silver or antimony for use at temperatures somewhat higher than those for tin-lead solders, and tin-zinc base solders often used in soldering aluminum. *See* SOLDERING.

Bronzes are among the most ancient of alloys and still form an important group of structural metals. Of the true copper-tin bronzes, up to 10% tin is used in wrought phosphor bronzes, and from 5 to 10% tin in the most common cast bronzes. Many brasses, which are basically copper-zinc alloys, contain 0.75–1.0% tin for additional corrosion resistance in such wrought alloys as Admiralty Metal and Naval brass, and up to 4% tin in cast leaded brasses. Among special cast bronzes are bell metal, historically 20–24% tin for best tonal quality, and speculum, a white bronze containing 33% tin that gained fame for high reflectivity before glass mirrors were invented. *See* BRONZE; COPPER ALLOYS.

Babbitt or bearing metal for forming or lining a sleeve bearing is one of the most useful tin alloys. It is tin containing 4–8% each of copper and antimony to give compressive strength and a structure desired for good bearing properties. An advantage of this alloy is the ease with which castings can be made or bearing shells relined with simple equipment and under emergency conditions. Aluminum-tin alloys are used in bearing applications that require higher loads than can be handled with conventional babbitt alloys. *See* ANTIFRICTION BEARING.

Pewter is an easily formed tin-base alloy that originally contained considerable lead. Thus, because Colonial pewter darkened and because of potential toxicity effects, its use was discouraged. Modern pewter is lead-free. The most favorable composition, Britannia Metal, contains about 7% antimony and 2% copper. This has desired hardness and luster retention, yet it can be readily cast, spun, and hammered. *See* PEWTER.

Type metals are lead-base alloys containing 3–15% tin and a somewhat larger proportion of antimony. As with most tin-bearing alloys, these are used and remelted repeatedly with little loss of constituents. Tin adds fluidity, reduces brittleness, and gives a structure that reproduces fine detail.

Flake and nodular gray iron castings are improved by adding 0.1% tin to give a fully pearlitic matrix with attendant higher hardness, heat stability, and improved strength and machinability.

Tin is commonly an ingredient in costume jewelry, consisting of pewterlike alloys and bearing-metal compositions often cast in rubber molds; in die castings hardened with antimony and copper for applications requiring close tolerances, thin walls, and bearing or nontoxic properties; and in low-melting alloys for safety appliances. The most common dental amalgam for filling teeth contains 12% tin. *See* TIN; TIN METALLURGY. Bruce W. Gonser

Bibliography. American Society for Testing and Materials, *Annual Book of ASTM Standards*, vol. 02.04: *Nonferrous Metals*, 1993; B. T. K. Barry and C. G. Thwaites, *Tin and Its Alloys and Compounds*, 1983; *Conference on Tin Consumption*, International Tin Council, London, 1972.

Tin metallurgy

The extraction of tin from its ores and its subsequent refining and preparation for use. Most tin concentrates are primarily cassiterite (SnO_2), the naturally

occurring oxide of tin. These are comparatively easy to reduce by using carbon at high temperatures. However, this operation differs from the smelting of most common metals because retreatment of the slag is necessary to obtain efficient metal recovery. *See* CASSITERITE.

In primary smelting, carbon monoxide (CO) formed during heat-up reacts with the solid cassiterite particles to produce tin (Sn) and carbon dioxide (CO_2). As the temperature increases, silica (present in nearly all concentrates) also reacts under reducing conditions with the SnO_2 to give stannous silicate. Iron, also present as an impurity in all concentrates, reacts with the silica to form ferrous silicate ($FeSiO_3$). These silicates fuse with the added fluxes to form a liquid slag, at which point unreacted carbon from the fuel becomes the predominant reductant in reducing both stannous silicate to tin and ferrous silicate to iron. The metallic iron then reduces tin from stannous silicate, as shown in the reaction below.

$$SnSiO_3 + Fe \rightleftharpoons FeSiO_3 + Sn$$

Primary smelting can be effected in a reverberatory, rotary, or electric furnace with the choice being more dependent on economic than technical considerations. In the Far East, for example, reverberatory furnaces fired with anthracite coal are widely used. Both Malaysia and Singapore have added electric furnaces to improve smelting efficiencies. Indonesia and Singapore use slow-speed rotary furnaces. Reverberatory and rotary furnaces are also used in Indonesia. On the other hand, the smelters in Zaire and Rwanda as well as those in South Africa, which are far away from coal sources, use electric furnaces because of the availability of electric power. In the case of Bolivia which has complex concentrates that may range as low as 15% tin from lode mining, roasting may be needed as a pretreatment before smelting in order to remove such undesirable impurities as sulfur and arsenic plus some lead, antimony, and bismuth. *See* ELECTRIC FURNACE.

One of the greatest contributions to modern tin smelting has been the fuming of tin slags. Stimulated by the need for better metal recoveries, this process relies on the formation and volatilization of tin as stannic oxide (SnO_2) in a type of blast furnace. The process requires the addition of pyrites to the tin-rich slag, where it reacts to produce $FeSiO_3$ and stannous sulfide (SnS). The SnS vapor oxidizes to SnO_2 and is carried out in the furnace exhaust gases, from which it is collected and recycled.

Fuming is an alternative to roasting in the smelting of low-grade concentrates (5–25% tin). A tin oxide dust, free of iron, is obtained which is fed back to a conventional smelting furnace.

The crude tin from slags and smelted concentrates is further refined by heat treatment (that is, liquidation or boiling) or sometimes electrolytic processes. In liquidation, tin is heated on a sloping hearth to just above its melting point. The tin runs into a poling kettle, while metals with higher melting points remain in the dross. Most of the iron is removed in this manner as well as part of the arsenic, antimony, and copper. Lead and bismuth remain. In the final refining step, the molten tin is agitated with steam, compressed air, or poles of green wood which produce steam. The remaining traces of impurities form a scum which is removed and recirculated through the smelting cycle. The pure tin is cast in iron molds to ingots of about 99 lb (45 kg). Purity is guaranteed to exceed 99.8%.

Iron, copper, arsenic, and antimony can be readily removed by the above processes or variations on these. However, for removing large amounts of lead or bismuth, electrolysis or a vacuum-refining process is used.

Secondary tin from metal scrap amounts to about one-quarter of the total tin consumed in the United States. Most of this comes from tin-bearing alloys, and secondary smelters rework them into alloys and chemicals. However, additional tin of high purity is recovered from the detinning of tinplate scrap. *See* ELECTROCHEMICAL PROCESS; ELECTROMETALLURGY; HEAT TREATMENT (METALLURGY); PYROMETALLURGY, NONFERROUS; TIN.

Daniel Maykuth

Bibliography. American Society for Testing and Materials, *Annual Book of ASTM Standards*, vol. 2.04: *Nonferrous Metals*, 1993; T. S. Mackey, Review of recent developments in tin—1981, *J. Metals*, pp. 72–75, April 1982; P. A. Wright, *Extractive Metallurgy of Tin*, 1982.

Tintinnida

An order of the Spirotrichia whose members are conical or trumpet-shaped pelagic forms bearing shells (loricae). These protozoa are planktonic ciliates

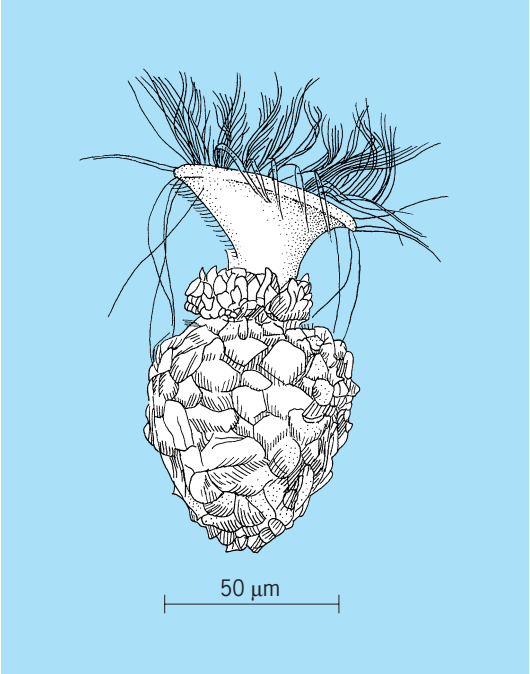

Fig. 1. *Tintinnopsis*, a living specimen shown protruding from its lorica, or shell.

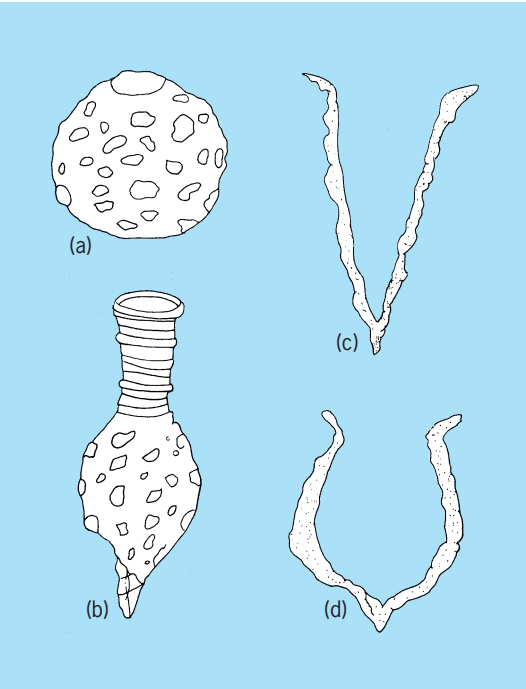

Fig. 2. Fossil and modern Tintinnida. (*a*) *Tintinnopsis*, Jurassic to Recent. (*b*) *Codonellopsis*, Recent. (*c*) *Amphorellina* (section), Lower Cretaceous. (*d*) *Calpionella* (section), Recent.

and are especially abundant in oceans, notably the Pacific. The lorica is composed of a resistant organic compound in which various foreign mineral grains are embedded; its shape may range from trumpet- or bell-form to cylindrical or subspherical, and its size from 50 to 200 micrometers. The exact structure, often quite elaborate, and the dimensions of the lorica are so recognizably different among the hundreds of known genera that the taxonomic arrangement of forms within the order has been based solely on characteristics and properties of this secreted "house." The adoral zone of membranelles (**Fig. 1**) is prominent, while the other ciliature is greatly reduced.

Fossil tintinnids, representing practically the only fossilized species of ciliate protozoa known to science, are identified on the basis of the shape of the lorica in cross section as seen in randomly oriented thin sections of the rocks in which they are found. Twelve genera of fossil tintinnids have been described from limestones and cherts of the Jurassic and Cretaceous.

Common genera are illustrated. Figure 1 shows a present-day member of a species of the genus *Tintinnopsis*. **Figure 2** shows drawings of loricae, whole and in section, including some fossil forms. *See* CILIOPHORA; PROTOZOA; SPIROTRICHIA.

John O. Corliss; Daniel J. Jones

Tire

A continuous pneumatic rubber and fabric cushion encircling and fitting onto the rim of a wheel. Sizes range from only a few inches in diameter up to 12 ft (3.7 m) and 12,500 lb (5700 kg).

In modern tire building, rubber (both natural and synthetic) and fabric remain the basic ingredients. Chemicals are compounded into the rubber to help it withstand wear, heat, and aging and to produce desired changes in its characteristics. Fabric (rayon, nylon, or polyester) is used to give the tire body strength and resilience. In belted tires, additional layers of fabric (rayon, fiber glass, finely drawn steel, or aramid) are placed just under the tread rubber to increase mileage and handling. Steel wire is used in the bead that holds the tire to the rim.

A tire is made up of two basic parts: the tread, or road-contacting part, which must provide traction and resist wear and abrasion, and the body or carcass, consisting of rubberized fabric that gives the tire strength and flexibility. In compounding the rubber, large amounts of carbon black are mixed with it to improve abrasion resistance. Other substances, such as sulfur, are added to enable satisfactory processing and vulcanization. *See* RUBBER.

Manufacture. The basic part of a tire-building machine is a collapsible cylinder shaped like a wide drum that turns under power and is controlled by the tire builder. First, plies of rubberized fabric are wrapped around the drum. The beads then are placed in position, and the sides of the plies are wrapped around them. If the tire is of belted bias or radial construction, belts of rubberized material are centered on the plies. Other narrow strips of material are placed near the beads for further strength when the tire flexes and rubs against the rim. Next, the tread and sidewall rubber is wrapped around the drum over the fabric. All of the components, which can number more than 20 in radial auto tires, are pressed together with rollers. The drum is collapsed and the tire is removed, ready to be molded and vulcanized. At this point the tire looks like a barrel with both ends open.

The tire still needs shaping and curing (or vulcanizing) to gain its final shape and strength characteristics. Besides molding the tread design and equalizing the stresses within the tire body, the vulcanization changes the rubber compound into a tough, highly elastic material and bonds the parts of a tire into one integral unit. When the tire emerges from the curing press, the building process is complete.

Types. There are three types of tires: bias-ply, radial, and belted bias (see **illus.**). For bias tires, cords in the plies extend diagonally across the tire

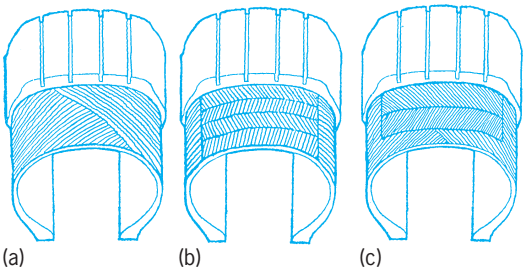

Tire construction. (*a*) Bias-ply. (*b*) Radial. (*c*) Belted bias. (*Goodyear Tire and Rubber Co.*)

from bead to bead. The cords run in opposite directions in each successive ply, resulting in a crisscross pattern. For radial tires, cords in the plies extend transversely from bead to bead, substantially perpendicular to the direction of travel. Belts are placed circumferentially around the tire. For belted bias tires, plies are placed in a manner similar to that used in the bias-ply tire, with belts of material placed circumferentially around the tire between the plies and the tread rubber.

Developments. Although most tire improvements appear gradually, a number of important developments have marked great advancements.

Fabric. Rayon was introduced in the late 1930s as a replacement for cotton. Nylon followed in the late 1940s, and it remains the basic material in truck, earthmover, and aircraft tires. Polyester, combining the best features of rayon and nylon, was first used in auto tires in the early 1960s; it became the most used tire cord and is in virtually all auto tires.

Tubeless tires. Prior to the mid-1950s all tires had to have inner tubes to contain the air pressure. The development of the tubeless tire brought increased puncture resistance and less heat buildup.

Belted tires. The belted bias tire was developed in the late 1960s to increase tread life and tire performance. Then in the 1970s the radial tire, long popular in Europe, won acceptance in the United States and became the most popular form of auto tire construction. Fiber glass, steel, and finally aramid were developed as materials for the belts.

Rubber. Early tires were totally dependent on natural rubber, which was often poor in quality. During World War II synthetic rubber was developed and now accounts for about 80% of the rubber used by the tire industry. Other compounds have allowed for greatly improved traction on ice, even without metal studs. Compound development also has led to lower rolling resistance, improved gasoline mileage, and longer tread life. *See* RUBBER. David B. Harrison

Bibliography. J. C. Dixon, *Tires, Suspension, and Handling*, 2d ed., 1996; H. B. Pacejka, *Tire and Vehicle Dynamics*, 2d ed., 2005; Society Of Automotive Engineers, *Tire and Wheel Technology*, 1999.

Tissue

An aggregation of cells more or less similar morphologically and functionally. The animal body is composed of four primary tissues, namely, epithelium, connective tissue (including bone, cartilage, and blood), muscle, and nervous tissue. The process of differentiation and maturation of tissues is called histogenesis. *See* CONNECTIVE TISSUE; EPITHELIUM; PLANT TISSUE SYSTEMS. Charles B. Curtin

Tissue culture

The branch of biology in which tissues or cells of higher animals and plants are grown artificially in a controlled environment. Such studies were undertaken in the hope that the behavior of various body components could be studied and their potentialities more readily analyzed under the simpler and more readily manipulated conditions possible in the test tube.

Study of the growth and interaction of animal cells with physical and chemical environments outside the body began about 1900. During the first decade it was demonstrated that cell multiplication from chick tissue transplanted into glass vessels could be maintained indefinitely, if suitable physical conditions for cell attachment to a solid substrate were provided, and if the necessary complex nutrient medium was replenished as fast as it was depleted by the cells' activities. From early beginnings, tissue culture has developed in many directions (**Fig. 1**).

Early methods of tissue culture were successful in promoting cell multiplication only when large numbers of cells were seeded together in a community. Growth of such large populations permitted many kinds of important experimentation on the multiplication process, but it also left many questions unsolved. For example, it was impossible to determine which fraction of the cells of any population had retained the ability to multiply. It also was difficult to determine the specific conditions which the individual cell requires in order to be able to initiate its reproductive process. A major advance along these lines was made by a group of scientists who succeeded in providing conditions permitting growth of single cells when individually sealed in capillary tubes. These cells were later grown into huge populations so that it was demonstrated that at least some cells which had originated in the mammalian body maintain their ability to multiply indefinitely in isolation, just like independent bacteria.

In early tissue culture, growth of cells was successful only when they were attached to a solid substrate

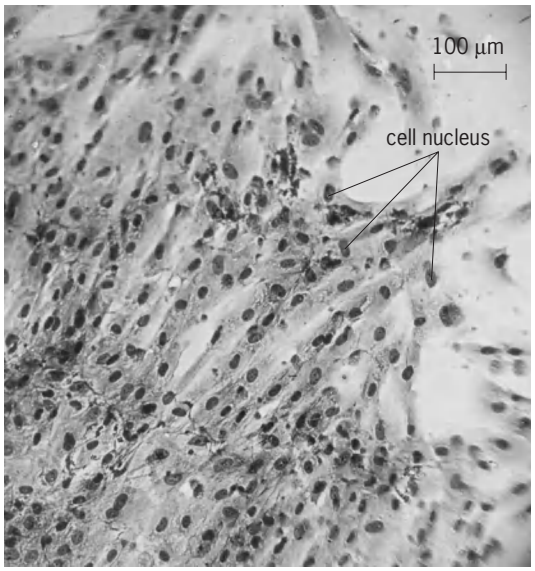

Fig. 1. Human spleen cells grown in a glass vessel containing a nutrient medium.

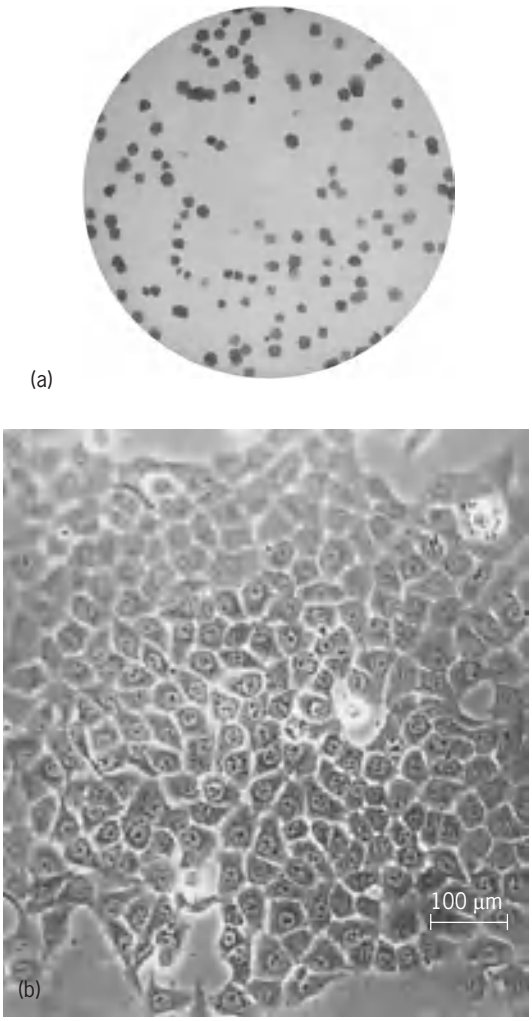

Fig. 2. Colonies developed from single human cancer (HeLa) cells. (*a*) Colonies grown on a glass dish. (*b*) Photomicrograph of a typical colony.

like glass or cellophane. In 1954 it became possible to grow cells in liquid suspension as well, a technique that permits many new operations, such as continuous farming of cells in the same vessel for indefinite periods. In addition, it became possible to simplify greatly the medium required for cell multiplication so as to eliminate, at least in some cases, the need for animal serum. Definition of the chemical requirements for mammalian cell growth in test tubes proceeded in a variety of laboratories and reached the stage wherein massive populations can be grown for long periods in a completely molecularly defined medium. Single cells can be reliably grown into discrete colonies in a medium containing completely defined, small-molecular weight constituents and a purified protein fraction obtained from blood. The completely defined small-molecular weight constituents are salts, amino acids, glucose, choline and inositol, and the vitamins biotin, pantothenic acid, folic acid, niacinamide, pyridoxine, riboflavin, thiamine, and B_{12}. Such advances have been tremendously important in establishing the specific nutrient requirements for different types of mammalian cells, and for elucidation of the metabolic pathways taken by the different molecules, in both healthy and diseased subjects.

Quantitation. In research undertaken to permit quantitative measurement of cell growth, means were found by which animal cells grown in tissue culture could be dispersed singly, then added to a glass dish, under conditions wherein every single cell would reproduce in isolation to form a discrete colony (**Fig. 2**). This aim was, at first, successfully achieved for cancer cells and then was also achieved for cells from normal human and animal tissues. This method of "plating" single mammalian cells made possible many more precise kinds of experiments. The effects of different physical and chemical agents on growth of cells could now be measured with much greater accuracy, since the number of cells able to reproduce under the required conditions could be precisely determined by a simple colony count. It also became possible to measure accurately and conveniently the growth rate of such single cells in different media (**Fig. 3**).

Study of hereditary mechanisms. Development of single-cell techniques afforded tissue-culture means for study of hereditary mechanisms in animal cells.

Cells grown in tissue culture by older techniques have demonstrated changes in their chromosome numbers and structures occurring with the passage of time. Since these bodies contain the genes which determine the hereditary potentialities of the cells, the genetic constitution of such cells was uncertain. Inferences drawn from the behavior of such cultures and then applied to interpretation of functions and potentialities of cells in the body, where chromosomal integrity is rarely altered, were often of doubtful significance. Various investigators then turned attention to these problems and developed new methods for study of the chromosomal constitution of mammalian cells in tissue culture.

One of the results of such advances in technique was the production of methods for regulating cell growth through extended periods of cultivation, so that the chromosomal integrity was maintained as reliably as it is in the body. It thus became possible to make many kinds of biochemical and genetic studies

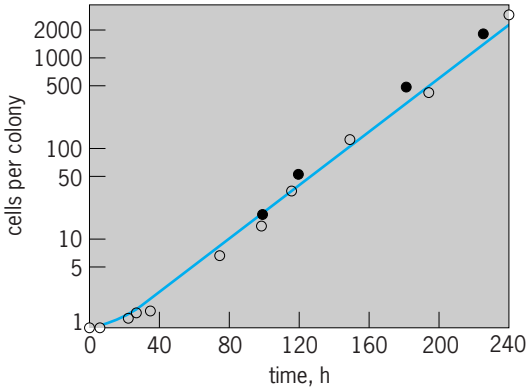

Fig. 3. Typical growth curve of single cells plated in a complete, nutrient medium. Cells begin to reproduce after initial delay of about 18 h and continue to double every 20 h as long as medium is not exhausted.

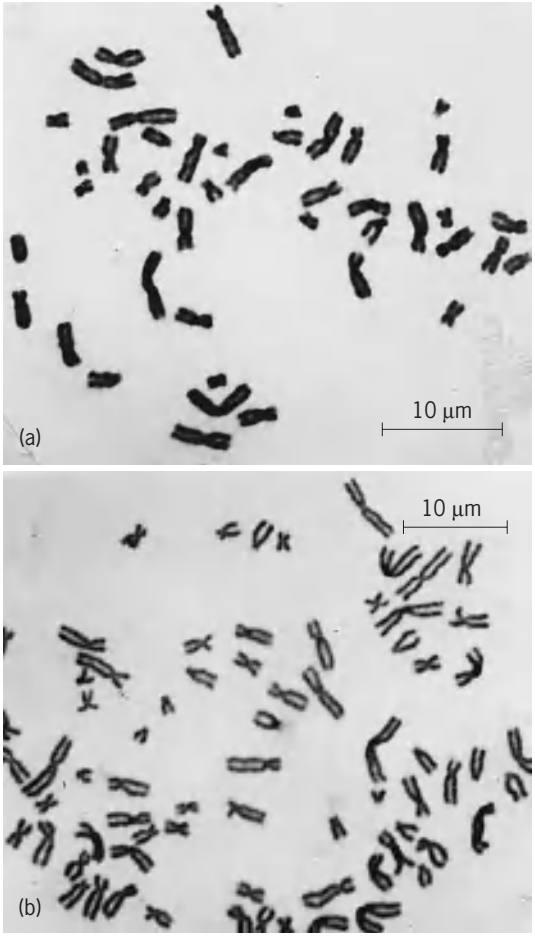

Fig. 4. Chromosome constitution. (*a*) Cells taken from normal human male and grown in tissue culture; 46 chromosomes, constant in number. (*b*) Human cancer; the chromosome number and structure are changed and variable from cell to cell.

on such cells with considerably more confidence in the applicability of results to an understanding of cell functions in the intact animal.

Advances in methods of taking cell specimens for culture reached the point where it became possible to obtain with ease a cell sample from a minute amount of skin taken from any individual and to cultivate these cells stably in the test tube.

Another fundamental result arising from chromosomal studies in tissue culture was the demonstration that the chromosome constitution of humans, which had been accepted for approximately 30 years, was in error. In 1956, tissue-culture studies demonstrated that the normal cells of humans contain 46 instead of 48 chromosomes. These results were verified in many laboratories by study of cells drawn from a variety of tissues of many human subjects. However, the occasional occurrence of a chromosome number other than 46 has been demonstrated in human subjects with genetic disease.

Diseases like ovarian dysgenesis (Turner's syndrome) have been found to be accompanied by a chromosome number of 45 instead of the normal 46. The missing one is a sex chromosome, so that such individuals have the X O sex chromosomal constitution. Many hitherto mysterious human diseases are now known to be due to specific chromosomal abnormalities. Thus, Down syndrome, a birth defect which results in a subnormal intelligence and certain physical signs, is now known to be caused by an extra chromosome of pair number 21. *See* DOWN SYNDROME; HUMAN GENETICS.

The universally used system for identification and classification of the human chromosomes was devised in 1959 by an international group and is known as the Denver System. Each one of the human chromosomes was identified and its structure delineated. In contrast to a human cell with its normal complement of chromosomes, the karyotype of human cancer cells has been studied and many of these have been found to possess abnormally high

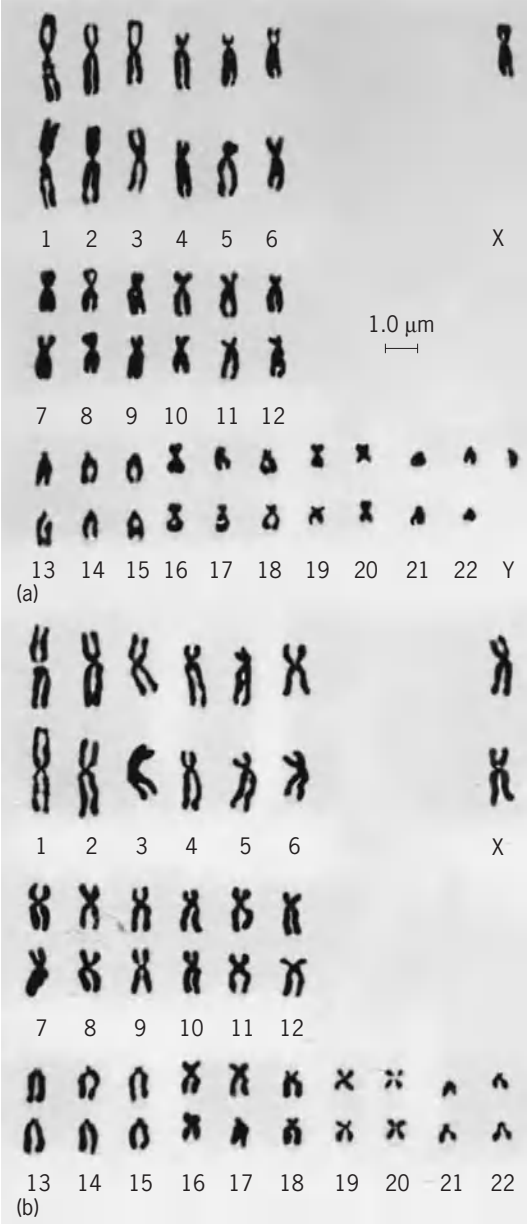

Fig. 5. The normal human chromosome constitution. (*a*) Male cells. (*b*) Female cells. Newer methods of specimen preparation reveal characteristic chromosome bands which also aid in identification of the chromosome.

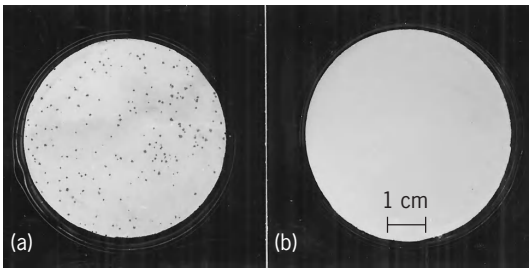

Fig. 6. Behavior of different mutant cells from the same cell population. Both cell types are identical in appearance and general behavior. Two hundred cells of each kind were placed in identical dishes containing the same nutrient medium. (*a*) The original cell reproduced and developed approximately 200 colonies. (*b*) The mutant cell produced no colonies whatever. However, if the amino acid glycine is added to the mutant cell, it produces colonies exactly like the unmutated cell type.

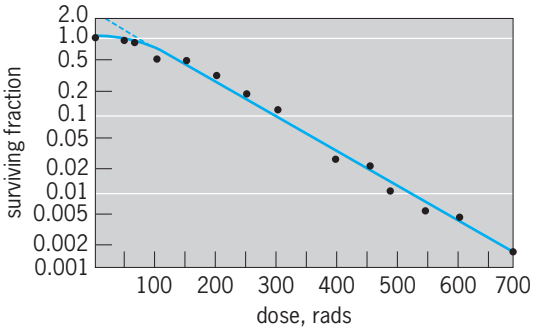

Fig. 7. Survival of reproductive capacity of cells from culture of a human cancer. The average lethal dose obtained from this curve is 96 rads.

chromosome numbers (**Fig. 4**). A karyotype is an idiogram of the chromosome complement characteristic of any individual or group of related organisms. The normal karyotype of human somatic cells is shown in **Fig. 5**. The Y chromosome, possessed only by males, is less than one-third the size of the X chromosome, which occurs singly in male cells but is doubled in female cells. All other chromosomes have been characterized by their sizes, by the position of their centromeres, and by the position of characteristic bands.

In studies of the genetic biochemistry of mammalian cells, methods were developed for the production and isolation of mutant cell cultures, whose biochemistry could be studied in the same fashion that was so successful in elucidating the molecular biology of bacteria. Mutants of cell populations were obtained which differ in their requirements for growth in the test tube (**Fig. 6**), and their underlying biochemistry has been shown. This research makes it possible to measure accurately the power of various physical and chemical agents to produce mutations in mammalian cells. Mutational rates have been measured for several mutagens. *See* MUTAGENS AND CARCINOGENS.

There are methods for locating the genes on their chromosomes, and powerful techniques for rapid and accurate measurement of the enzyme contents of cells with different genetic constitutions. With these tools the links between genes, enzymes, and specific developmental processes in different tissues are demonstrable in mammals.

Radiation studies. Investigations of the effect of high-energy irradiation on single human and other animal cells have accurately defined the dose required to prevent colony formation by reproduction of single cells (**Fig. 7**). This dose is approximately 100 rads for most mammalian cells. Before these studies it had been generally considered that the human cell is many times more resistant to radiation than this figure implies. These measurements made possible interpretation of the effect of high-energy radiation on the human body in a manner that satisfactorily explains many previously obscure aspects of radiation pathology. Study of the mechanisms by which radiation interferes with reproductive processes in mammalian cells has shown that most damage is confined to the cellular genetic apparatus and particularly to the chromosomes. **Figure 8** presents a series of pictures showing normal human chromosomes and those from human cells irradiated with various doses of x-rays. *See* RADIATION INJURY (BIOLOGY).

Use in virus studies. Tissue-culture studies opened a new era in investigations of mammalian viruses. These techniques made it possible to prepare large quantities of viruses for the work of virologists and immunologists. Thus the preparation of vaccines was speeded, as well as the performance of a wide variety of physical, chemical, and biological studies on these infectious agents.

Tissue-culture studies also opened whole new vistas of understanding of the virus-cell interaction. It became possible to adapt to tissue culture the plaque technique, a method in which single virus particles are recognized by the round area of cell destruction they produce (**Fig. 9**); thus an accurate determination can be made of the number of virus particles in the original suspension. This technique allows a much more precise measurement of the effects of viruses under different kinds of controlled physical and chemical situations. Since all the virus progeny of a single plaque area have descended from the same individual, it becomes more readily possible to carry out controlled genetic investigations on viruses. *See* VIRUS.

Mammalian cell molecular biology. Molecular biology was formed from the amalgamation of two previously distinct scientific fields, biochemistry and genetics, and it achieved an explosion of understanding of the simplest living cells. Molecular biological studies with bacteria furnished the first blueprint for the workings of a living cell: The genetic substance is a group of specific molecules of known chemical structure, deoxyribonucleic acid (DNA); these molecules are capable of self-replication; and these molecules which constitute the genes regulate the biosynthesis of the cell proteins from moment to moment throughout the life of the cell. Therefore, since DNA is responsible for construction of the cell machinery, it is ultimately responsible for the whole range of biochemical activity exhibited by any cell.

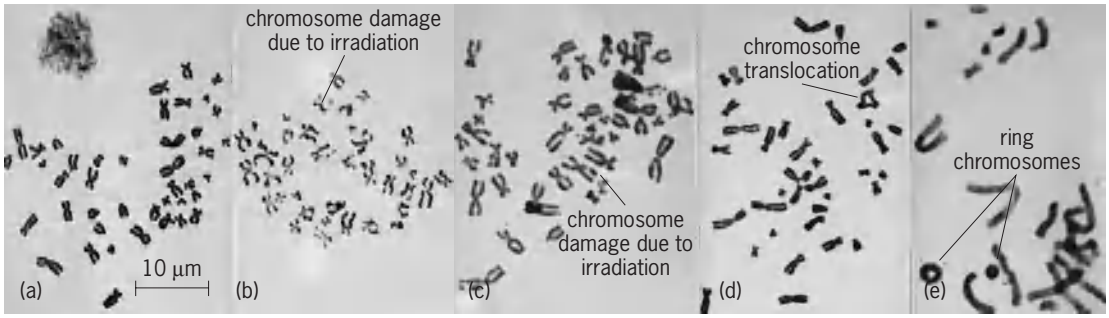

Fig. 8. Typical chromosomal lesions produced in human cells grown in tissue culture and irradiated with various doses of x-rays. (*a*) Normal human chromosome complement, unirradiated cell. (*b*) Chromosomes of cell after irradiation with 50 rads. (*c*) Chromosomes of cell irradiated with 75 rads. (*d*) Chromosomes of cell irradiated with 75 rads. Translocations have appeared as a result of abnormal restitution of the multiple breaks. (*e*) Development of ring chromosomes as a result of irradiation of a normal human cell with 150 rads.

See DEOXYRIBONUCLEIC ACID (DNA); MOLECULAR BIOLOGY.

An attempt to apply a similar approach to humans was not possible before 1956, because human genetics was fragmentary and weak. This situation existed for two reasons. The reproductive period of the simple bacterial cells is 20 min, while the generation time in humans is almost 25 years. Of equal significance was the fact that it is not possible in humans to make those matings which would be most illuminating genetically. Perhaps the greatest triumph of tissue culture in this century was the development of single-cell plating methods, which made possible study of human and mammalian genetics in the test tube by the methods of cloning which had been demonstrated to be so powerful with the simple bacterial cells. *See* GENETIC ENGINEERING.

In addition to study at the chromosome level, the human genetic structure has been examined at the single-gene level. Large numbers of genes have been identified and their protein products isolated. These developments made possible the study of biochemical pathways in human cells with great precision, and have illuminated previously mysterious diseases in which particular biochemical steps are defective. *See* MOLECULAR PATHOLOGY.

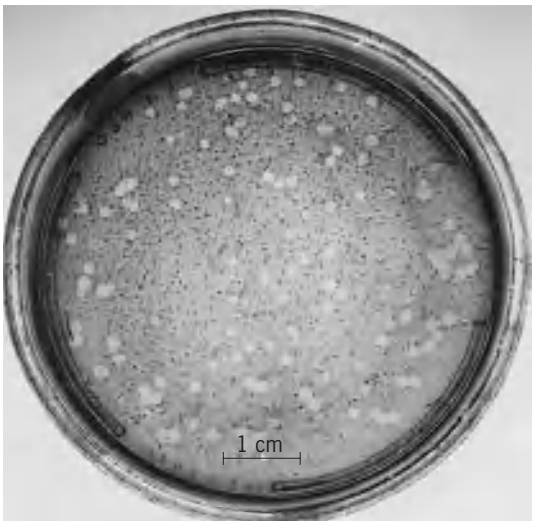

Fig. 9. Plaques or areas of cell destruction produced by the action of the virus of Newcastle disease on chick fibroblasts cultivated in a glass dish.

Another important development was the process of cell hybridization, which in effect makes possible the mating of cells in the test tube. Examining the progeny of such cellular mating processes uncovers details of human genetic structure and function. This procedure made possible determination of dominance and recessiveness in human genes in a very simple and definitive fashion. In addition, it made possible mapping of the human genes on their chromosomes, which is necessary in order to tell which genes are closely linked and which are far apart. Perhaps of even greater importance, however, is the need to understand the mechanism of gene regulation and how it is affected by gene position. Mammalian cells differ from simple bacterial cells in possessing the enormously complex property of differentiation in which different genetic regions are active in cells of the different tissues, permitting each cell grouping to act in its own highly characteristic fashion. In the simple bacterial cells, contiguous genes are often turned on and off together. Elucidation of mechanisms of this kind will provide understanding of the differentiation process and of the many different diseases which are due to defective gene regulation either in the course of earlier embryonic development or later in life. There is strong evidence to indicate that at least some aspects of the aging process and various degenerative diseases are due to failures in gene regulation. *See* CELL DIFFERENTIATION; CELL SENESCENCE; GENETIC MAPPING; SOMATIC CELL GENETICS.

One of the most important developments in this respect has been gene mapping at every level from that of parts of the DNA molecule to the entire chromosome. In addition to providing the fundamental data which should eventually make gene regulation understandable, these techniques provide new methods of diagnosis of genetic disease. These methods are applicable even very early in pregnancy, so that the presence of severe genetic disability can often be determined in time to terminate the pregnancy if the prospective parents choose. The procedure could make possible prevention of many human genetic tragedies.

Mutant cells with respect to regulatory processes involved in critical diseases like those dealing with cholesterol metabolism have been studied both in individuals and in the test tube. Methods were developed for isolating and identifying the proteins which are the characteristic gene products of mammalian cells. Receptor sites which are the targets for the action of hormones have been identified in a variety of tissue culture studies. Along with these developments in genetic biochemistry, tissue culture has permitted elucidation of cell structures which were previously unsuspected. *See* CELL (BIOLOGY).

There are also tissue culture methods for monitoring environmental agents for their ability to cause gene and chromosomal mutations, which underlie human genetic disease. It is now presumed that the same kinds of defects are also responsible for a large proportion of cancers. The use of these methods may also make it possible in a single operation to develop protection against agents which are responsible for a great deal of human genetic disease and cancer.

Other techniques. Many advances were carried out by cinema photomicrography, which portrayed the surprising variety of movements of different kinds of cells growing in tissue culture. It was shown that cells taken from a whole animal and dispersed individually have the ability to reaggregate in tissue culture to form tissues very much like those of their original site. *See* CELL ADHESION.

There are other techniques for mapping with great precision the point at which various biochemical steps occur in the life cycle of a cell. In this way, the specific events of the cell's reproductive cycle appear to be capable of delineation and the point in the life cycle at which various drugs and other agents exert their effects can now be determined. *See* CELL CYCLE.

Cells in tissue culture can be made cancerous by treatment with certain viruses, x-irradiation, or carcinogenic hydrocarbons. They usually acquire specific changes in their patterns of biochemical activity, and they lose the capacity to inhibit one another's reproduction as occurs when noncancerous cells are crowded together in a tissue culture. *See* MUTAGENS AND CARCINOGENS.

A number of techniques produce specific differentiation of cells in tissue culture so that they can carry out the biochemical steps characteristic of various organs of the body. Various hormones are effective in causing nondifferentiated cells to synthesize specific enzymes like those characteristic of normal cells. Under the influence of feeder layers, layers of cells whose reproductive power has been suppressed by x-irradiation, cells from muscle or bone marrow will multiply and, at the same time, develop differentiated functions characteristic of their tissue of origin.

While this discussion has been confined to animal tissue culture, mention should be made that there have been many parallel developments in the field of plant tissue culture. Theodore T. Puck

Bibliography. R. F. Beers, Jr., and E. G. Bassett (eds.), *Cell Fusion: Gene Transfer and Transformation*, 1984; L. Goldstein and D. M. Prescott (eds.), *Cell Biology: A Comprehensive Treatise*, vol. 4, 1981; B. M. Martin, *Tissue Culture Techniques: An Introduction*, 1994; T. T. Puck, *The Mammalian Cell as a Microorganism*, 1972; A. E. Sowers (ed.), *Cell Fusion*, 1987; T. A. Springer (ed.), *Hybridoma Technology in the Biosciences and Medicine*, 1985.

Tissue typing

A procedure involving a test or a series of tests to determine the compatibility of tissues from a prospective donor and a recipient prior to transplantation. The immunological response of a recipient to a transplant from a donor is directed against many cell-surface histocompatibility antigens controlled by genes at many different loci. However, one of these loci, the major histocompatibility complex (MHC), has the greatest genetic complexity and controls antigens that evoke the strongest immunological response. The MHC is a cluster of closely linked gene loci and is conserved in all vertebrate species, including humans. The human MHC is known as the HLA system, which stands for the first (A) Human Leukocyte blood group system discovered. The HLA complex was first described as a gene locus on chromosome 6 that controls the allograft rejection response, which is the rejection of a graft from a donor by a genetically dissimilar recipient of the same species. It has since been found, however, that these genes are also physiologically important in the regulation of the immune response to highly foreign antigens such as bacterial or viral antigens, as well as to self-antigens. *See* CELLULAR IMMUNOLOGY; HISTOCOMPATIBILITY.

The success of transplantation is greatly dependent on the degree of histocompatibility (identity) between the donor and recipient, which is determined by the HLA complex. When the donor and recipient have a low degree of histocompatibility, the organ is said to be mismatched, and the recipient mounts an immune response against the donor antigen. The intensity of this response depends on the haplotype combination of HLA locus alleles, since certain alloantigens are more immunogenic than others. Additionally, previous sensitization of a potential transplant recipient to HLA antigens through pregnancy, blood transfusion, or a prior transplant may result in preexisting anti-HLA antibodies. If these preformed antibodies are specific for the donor HLA antigens, hyperacute rejection of transplanted tissue can occur. Therefore, histocompatibility testing plays an important role in the selection of donors and recipients for organ transplantation. By laboratory testing, the degree of antigenic similarity between the donor and the recipient and the degree of preexisting recipient sensitization to donor antigens can be determined. This is known as cross-matching. The success of transplantation depends largely on the degree of MHC compatibility between donor and recipient. and an absence in the recipient of preformed antibodies to donor antigens.

Phenotyping of HLA-A, -B, and -C (ABC typing) of an individual is determined by reacting that individual's lymphocytes with a large panel of antisera directed against specific HLA antigens. These antisera are generally obtained from women who have had two or more pregnancies and are selected so that, collectively, all significant HLA antigens are likely to be encountered. The procedure is known as complement-mediated cytotoxicity assay. The person's lymphocytes are incubated with the different antisera and complement is added. Killing of the cells being tested indicates that they express the HLA determinants recognized by the particular antiserum being used. Killing of potential donor lymphocytes in the complement-mediated cytotoxicity assay indicates the presence of antibodies specifically directed against HLA antigens, and is a contraindication to transplantation of tissue from that donor. *See* COMPLEMENT; HYPERSENSITIVITY; IMMUNOASSAY.

Assistance in tissue typing is being sought with additional techniques. These include the use of primed lymphocytes (lymphocytes that have been stimulated previously by contact with lymphocytes bearing target HLA specificities), monoclonal antibodies directed against selected specificities, and direct analysis of deoxyribonucleic acid (DNA) either by means of suitable probes and the polymerase chain reaction or by limited enzymatic digestion followed by analysis of the restriction fragment length polymorphism.

In addition to its important role in organ transplantation, determination of the HLA phenotype is useful in paternity testing, forensic medicine, and the investigation of HLA-disease associations. Although the immune system is carefully regulated, it can react against the individual's own tissues (autoimmunity). Several autoimmune disorders, such as ankylosing spondylitis, rheumatoid arthritis, and diabetes, have been linked to a high degree with certain HLA antigens. *See* AUTOIMMUNITY; TRANSPLANTATION BIOLOGY.

M. Wayne Flye; T. Mohanakumar

Bibliography. E. Albert et al., Nomenclature for factors of the HLA system—1977, *Tissue Antigens*, 11:81–86, 1978; E. Albert, M. P. Bauer, and W. R. Mays (eds.), *Histocompatibility Testing, 1984*, 1985; P. Dyer and D. Middleton (eds.), *Histocompatibility Testing: A Practical Approach*, 1993; M. W. Flye (ed.), *Principles of Organ Transplantation*, 1989; G. Opelz, Correlation of HLA matching with kidney graft survival in patients with or without cyclosporine treatment, *Transplantation,* 40:240–243, 1985; R. Patel and P. I. Terasaki, Significance of a positive crossmatch test in kidney transplantation, *N. Engl. J. Med.*, 288:735–736, 1969.

Titanite

A calcium, titanium silicate, $CaTiOSiO_4$, of high titanium content. Titanite is also known as sphene.

Titanite is an orthosilicate (nesosilicate) in which silicate $[(SiO_4)^{-4}]$ tetrahedra do not share any apical oxygens with adjacent tetrahedra but are cross-linked by chains of octahedrally (sixfold) coordinated titanium ions (Ti^{4+}) and calcium ions (Ca^{2+}) that are coordinated to seven oxygens. This yields a monoclinic structure with the space group C2/c. It has a hardness of 5–5$^1/_2$ on the Mohs scale, a distinct cleavage, a specific gravity of 3.4–3.55, and an adamantine to resinous luster. It commonly occurs as distinct wedge-shaped crystals that are usually brown in hand specimens. Titanite may also be gray, green, yellow, or black. *See* COORDINATION CHEMISTRY; CRYSTAL STRUCTURE.

Titanite is a common accessory mineral in many igneous and metamorphic rocks. It may be the principal titanium-bearing silicate mineral, especially in intermediate and alkali-rich intrusive igneous rocks such as nepheline syenites. It occurs in abundance in the Magnet Cove, igneous complex in Arkansas and in the intrusive alkalic-rocks of the Kola Penninsula, Russia. It is also common in metamorphosed mafic rocks (that is, metabasalts) and metamorphosed impure limestones and dolostones. In metamorphic rocks, titanite's stability at high temperature is limited by reactions resulting in the formation of rutile (TiO_2) as the dominant titanium-bearing mineral. Titanite grains may also occur in detrital sediments and sedimentary rocks.

The composition of titanite may diverge from pure $CaTiSiO_4$ because of a variety of chemical substitutions. Calcium ions (Ca^{2+}) can be partially replaced by strontium ions (Sr^{2+}) and rare-earth ions such as thorium (Th^{4+}) and uranium (U^{4+}). Aluminum ions (Al^{3+}), ferric iron (Fe^{3+}), and ferrous iron (Fe^{2+}) may substitute for titanium ions (Ti^{4+}), whereas oxygen ions (O^{2-}) may be replaced by hydroxyl ions (OH^-), fluoride ions (F^-), and chloride ions (Cl^-). Because titanite commonly contains radioactive elements, it has been used for both uranium-lead and fission track methods of dating. *See* DATING METHODS; IGNEOUS ROCKS; METAMORPHIC ROCKS; SILICATE MINERALS; TITANIUM.

John C. Drake

Bibliography. W. A. Deer, R. A. Howie, and J. Zussman, *Rock Forming Minerals,* vol. 1: *Ortho- and Ring Silicates*, 1962; *Geology and Resources of Titanium*, U.S. Geol. Survey Prof. Pap. 959-E, 1976; J. A. Hunt and D. M. Kerrick, The stability of sphene: Experimental redetermination and geological implications, *Geochim. Cosmochim. Acta*, 41:279–288, 1977; C. Klein and C. S. Hurlbut, Jr., *Manual of Mineralogy*, 21st ed., rev. 1999; P. H. Ribbe (ed.), *Orthosilicates*, 1982.

Titanium

A chemical element, Ti, atomic number 22, and atomic weight 47.90. It occurs in the fourth group of the periodic table, and its chemistry shows many similarities to that of silicon and zirconium. On the other hand, as a first-row transition element, titanium has an aqueous solution chemistry, especially of the lower oxidation states, showing some resemblances to that of vanadium and chromium. *See* PERIODIC TABLE; TRANSITION ELEMENTS.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

The catalytic activity of titanium complexes forms thebasis of the well-known Ziegler process for the polymerization of ethylene. This type of polymerization is of great industrial interest since, with its use, high-molecular-weight polymers can be formed. In some cases, desirable special properties can be obtained by forming isotactic polymers, or polymers in which there is a uniform stereochemical relationship along the chain. *See* POLYOLEFIN RESINS.

The dioxide of titanium, TiO_2, occurs most commonly in a black or brown tetragonal form known as rutile. Less prominent naturally occurring forms are anatase and brookite (rhombohedral). Both rutile and anatase are white when pure. The dioxide may be fused with other metal oxides to yield titanates, for example, K_2TiO_3, $ZnTiO_3$, $PbTiO_3$, and $BaTiO_3$. The black basic oxide, $FeTiO_3$, occurs naturally as the mineral ilmenite; this is a principal commercial source of titanium.

Titanium dioxide is widely used as a white pigment for exterior paints because of its chemical inertness, superior covering power, opacity to damaging ultraviolet light, and self-cleaning ability. The dioxide has also been used as a whitening or opacifying agent in numerous situations, for example as a filler in paper, a coloring agent for rubber and leather products, a pigment in ink, and a component of ceramics. It has found important use as an opacifying agent in porcelain enamels, giving a finish coat of great brilliance, hardness, and acid resistance. Rutile has also been found as brilliant, diamondlike crystals, and some artificial production of it in this form has been achieved. Because of its high dielectric constant, it has found some use in dielectrics.

The alkaline-earth titanates show some remarkable properties. The dielectric constants range from 13 for $MgTiO_3$ to several thousand for solid solutions of $SrTiO_3$ in $BaTiO_3$. Barium titanate itself has a dielectric constant of 10,000 near 120°C (248°F), its Curie point; it has a low dielectric hysteresis. These properties are associated with a stable polarized state of the material analogous to the magnetic condition of a permanent magnet, and such substances are known as ferroelectrics. In addition to the ability to retain a charged condition, barium titanate is piezoelectric and may be used as a transducer for the interconversion of sound and electrical energy. Ceramic transducers containing barium titanate compare favorably with Rochelle salt and quartz, with respect to thermal stability in the first case, and with respect to the strength of the effect and the ability to form the ceramic in various shapes, in the second case. The compound has been used both as a generator for ultrasonic vibrations and as a sound detector. *See* PIEZOELECTRICITY. Arthur W. Adamson

In addition to important uses in applications such as structural materials, pigments, and industrial catalysis, titanium has a rich coordination chemistry. The formal oxidation of titanium in molecules and ions ranges from −II to +IV. The lower oxidation states of −II and −I occur only in a few complexes containing strongly electron-withdrawing carbon monoxide ligands.

The lower oxidation states of titanium are all strongly reducing. Thus, unless specific precautions are taken, titanium complexes are typically oxidized rapidly to the +IV state. Moreover, many titanium complexes are extremely susceptible to hydrolysis. Consequently, the handling of titanium complexes normally requires oxygen- and water-free conditions. *See* COORDINATION CHEMISTRY. L. Kieth Woo

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; M. F. Lappert, *Comprehensive Organometallic Chemistry II: Scandium, Yttrium, Lanthanides and Actinides, and Titanium, Zirconium, and Hafnium*, vol. 4, 1995; G. Lütjering and J. C. Williams, *Titanium: Engineering Materials and Processess*, 2003.

Titanium metallurgy

The winning of metallic titanium (Ti) from its ores followed by alloying and processing into forms and shapes that can be used for structural purposes.

History. In 1791 a British clergyman, William Gregor, published his observations and experiments with black sands found on Cornish beaches containing an unknown element. A few years later an Austrian chemist, M. H. Klaproth, identified this as an oxide of a new element which he named titanium.

The first metallic titanium was produced in the United States by M. A. Hunter at Rensselaer Polytechnic Institute and his associates at the General Electric Company. Titanium tetrachloride, $TiCl_4$, was reacted with sodium to produce metal.

In 1932 Wilhelm Kroll, a native of Luxembourg and the recognized father of the modern titanium industry, manufactured metallic titanium by combining titanium tetrachloride with calcium, and he made a few pieces of wire, sheet, and rod. By late 1940 Kroll had switched to reacting titanium tetrachloride with magnesium under an argon atmosphere, the basis for the first practical commercial process for producing titanium metal. By the mid-1950s a number of Japanese and American companies were producing many thousands of tons of metal using the Kroll method, a practice which still dominates the industry.

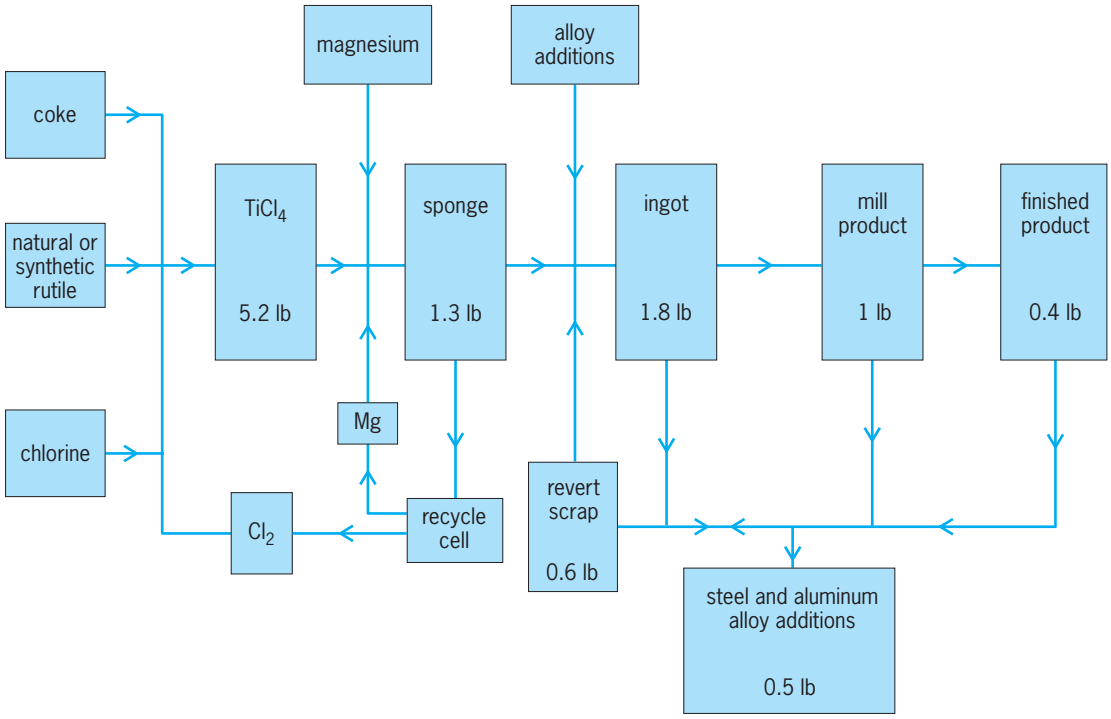

Fig. 1. Material cycle for titanium mill product manufacture. 1 lb = 0.45 kg.

In response to the unique material requirements of light weight and temperature and corrosion resistance for the evolving gas turbine engine, the titanium metal industry emerged in 1950. Titanium's unique properties—density half that of steel, excellent strength retention to 1000°F (538°C), and atmospheric corrosion immunity superior to that of other metals—made it an ideal construction material for both the engines and airframes of the newly developing jet airplanes. *See* AIRCRAFT ENGINE; AIRFRAME.

In the 1960s, discovery of titanium's excellent corrosive performance opened up a host of new applications in the chemical process industry. The first successful examples of industrial applications of titanium included reboilers in nitric acid concentrators, wet chlorine gas coolers in chlor-alkali plants, and chlorine dioxide bleach equipment in pulp/paper plants.

Natural occurrence. A stimulus for the initial interest in metallic titanium was the fact that it is the fourth most abundant structural element on the Earth's crust, exceeded only by aluminum, iron, and magnesium. Titanium deposits are widely scattered throughout the Earth's surface. Two forms dominate: rutile, essentially pure TiO_2 (95%), which usually occurs as black particles in beach sands; and the more abundant ilmenite, a titaniferrous ore, $FeTiO_3$ (50–65% Ti-bearing), which occurs in both alluvial and volcanic formations. *See* ILMENITE.

Of all the titanium minerals mined, only 3–5% are used to produce metal. The remainder is processed to titanium oxide (TiO_2) for use in the pigment industry, which utilizes either a sulfate or chlorination process to recover and purify it. Because of environmental concerns, the chlorination process is the preferred technology. This process requires a feedstock containing greater than 85% TiO_2, thus precluding the direct use of ilmenite. Because of the dwindling supplies of rutile, numerous plants are under construction for processes that can upgrade the more abundant ilmenite ore to greater than 90% TiO_2.

Winning. All commercial titanium metal is produced from titanium tetrachloride ($TiCl_4$), an intermediate compound produced during the chlorination process for titantium oxide pigment. The process (**Fig. 1**) involves chlorination of ore concentrates; reacting TiO_2 with chlorine gas (Cl_2) and coke (carbon; C) in a fluidized-bed reactor forms impure titanium tetrachloride as in reaction (1). For the production of acceptable metal, purification of the raw tetrachloride is required to remove other metal chlorides that would contaminate the virgin titanium. These critical purification steps involve distillation and precipitation of chlorides of vanadium, iron, zirconium, silicon, and magnesium, all of which occur with titanium in the ore.

$$TiO_2 + 2C + 2Cl_2 \rightarrow TiCl_4 + 2CO \qquad (1)$$

The purified titanium tetrachloride is delivered as a liquid to the reactor vessel. In these vessels, constructed of carbon or stainless steel, the titanium tetrachloride is reacted with either magnesium (Mg) or sodium (Na), as in reactions (2) and (3), to form

$$TiCl_4 + 2Mg \rightarrow Ti + 2MgCl_2 \qquad (2)$$

$$TiCl_4 + 4Na \rightarrow Ti + 4NaCl \qquad (3)$$

the pure metal called sponge, because of its porous cellular form. To avoid contamination by oxygen or nitrogen, the reaction is carried out in an argon atmosphere.

The magnesium chloride ($MgCl_2$) or sodium chloride (NaCl) can be recycled to obtain both the metal fraction, magnesium [reaction (4)] or sodium [reaction (5)], and chlorine. The recycling process is a

$$MgCl_2 \rightarrow Mg + Cl_2 \quad (4)$$

$$2NaCl \rightarrow 2Na + Cl_2 \quad (5)$$

conventional production method.

The sponge, removed from the reactor pot by boring, is cleaned by acid leaching. In Russia and Japan, these excess reactants are removed by vacuum distillation.

Consolidation. One of the earliest challenges in the production of ductile titanium metal and its alloys was the development of an economical and technically acceptable method to consolidate the titanium sponge without contamination and embrittlement from atmospheric oxygen and nitrogen. Further complicating the problem was the highly reactive nature of molten titanium metal and its propensity for dissolving the mold materials into which it is cast.

As a result of these considerations, the consumable-electrode arc furnace eventually evolved. A mass of sponge, alloy additions, and scrap are mixed, then compressed into compacts and welded together to form a sponge electrode. This is melted by an electric arc into a water-cooled copper crucible in a vacuum or an atmosphere of purified argon. The arc progressively consumes the sponge electrode to form an ingot. No reaction occurs between the cool copper wall and the molten titanium; the vacuum or inert atmosphere prevents contamination of the molten metal. To attain commercially acceptable uniformity, the first ingot is remelted at least once and sometimes twice in a similar consumable-electrode furnace. Ingots up to 30,000 lb (13,600 kg) are routinely produced by using this consolidation method.

The cold-hearth melting process consists of feeding raw materials into a water-cooled crucible shaped like a bathtub and completely enclosed within a vacuum or inert-gas environment. Either electron beams or plasma gas (argon, helium, or a mixture of both) is used as the heat source. The molten metal flows over the lip into a mold of the appropriate shape. The process offers the advantage of casting rectangular shapes that can be processed easily on plate mills (rolling mills for producing flat products). Also, the process can be designed to remove, by density separation, harmful high-density inclusions.

Forming. The conversion of the titanium ingot into mill products, such as forging billet, plate, sheet, and tubing, is accomplished for the most part on conventional metalworking equipment. Mills designed to roll and shape stainless or alloy steel are used with only slight modifications. For this reason titanium and its structural alloys are produced in most of the same forms and shapes as stainless steel. *See* METAL FORMING; STAINLESS STEEL.

Shaping. Fabricating titanium mill products into finished parts is performed on conventional metalworking equipment with only a few exceptions. During any heating operation it is necessary to minimize the contaminating and embrittling effects of oxygen, nitrogen, and hydrogen. Close control of furnace temperatures and environments for heating prior to forging, forming, or heat treating are critical. During welding operations, the molten and hot metal must be protected from the atmosphere; otherwise a brittle weld will result. However, shielding techniques using argon or helium gas are routinely practiced. *See* EMBRITTLEMENT.

Basic metallurgy. Titanium is a relatively light, silvery-gray metal with a specific gravity of 0.163 lb/in.^{3} (4.51 g/cm^3). Pure titanium has a high melting point, 3035°F (1668°C). Titanium has a lower coefficient of expansion and lower thermal conductivity than either steel or aluminum alloys, and is not magnetic. Its modulus of elasticity, a measure of stiffness, is 1.6×10^7 lb/in.^{2} (1.1×10^{11} pascals), midway between that of steel and aluminum.

Titanium is allotropic. Up to 1625°F (774°C), titanium atoms arrange themselves in a hexagonal close-packed crystal array known as alpha (**Fig. 2***a*). When titanium is heated above the transition temperature (beta transus) of 1625°F (774°C), the atoms rearrange themselves into a body-centered cubic structure known as beta (Fig. 2*b*). The addition of other metals to a titanium base will favor one or the other of the two crystallographic forms. Some common titanium alloy additions are as follows:

Alpha stabilizers	*Beta stabilizers*	*Neutral*
Aluminum	Vanadium	Zirconium
Oxygen	Tantalum	Tin
Nitrogen	Molybdenum	
Carbon	Chromium	
	Iron	
	Nickel	

Whether a particular element favors the alpha or beta phase will raise or lower the beta transus temperature. Aluminum, for example, favors (stabilizes) the

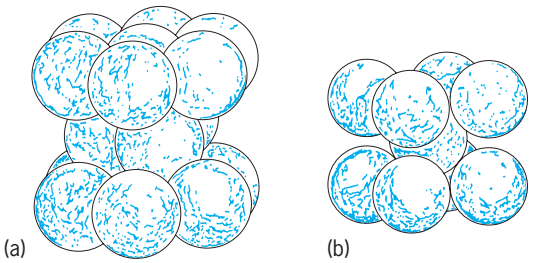

Fig. 2. Crystallographic forms of titanium. (*a*) Hexagonal close-packed alpha phase. (*b*) Body-centered cubic beta phase.

TABLE 1. Titanium corrosion resistance in corrosive environments*

Environment	Concentration, wt %	Temperature, °F (°C)	Corrosion rate, mil/yr† (mm/yr)
Acetic	All	212 (100)	<0.05 (0.001)
Aniline hydrochloride	5–20	122–212 (50–100)	<0.05 (0.001)
Benzene + HCl, no Cl	Vapor + liquid	349 (176)	0.2 (0.005)
Carbon tetrachloride	100	Boiling	<5.0 (0.1)
Chromic acid	25	212 (100)	<0.05 (0.001)
Ethanol	95	Boiling	0.05 (0.01)
Formaldehyde	37	Boiling	<5 (0.1)
Formic acid	10	212 (100)	<0.05 (0.001)
Hydrochloric acid	5	212 (100)	>250 (6.4)
Nitric acid	15	212 (100)	<0.05 (0.001)
	40	212 (100)	>25 (0.64)
Oxalic acid	25	212 (100)	<1000 (25)
Phosphoric acid	10	212 (100)	450 (11)
Sulfuric	5	Boiling	>2000 (50)
Terephthalic acid	75	Boiling	<0.05 (0.001)

*After S. R. Seagle and D. E. Thomas, Status of titanium technology, *Chem. Eng. Prog.*, p. 63, June 1986.
†1 mil = 0.001 in.

alpha structure, raising the temperature at which the alpha transforms to beta. Iron favors the beta phase; therefore, iron depresses the beta transus. *See* CRYSTAL STRUCTURE.

Pure titanium is soft, weak, and extremely ductile. However, through appropriate additions of other elements, the titanium metal base is converted to an engineering material having unique characteristics, including high strength and stiffness, corrosion resistance, and usable ductility. The type and quantity of alloy addition determine the mechanical and, to some extent, the physical properties.

Alloys. Titanium alloys are classified into three groups depending on the phases present: alpha, beta, or a combination of the two, alpha-beta alloys. *See* ALLOY.

Alpha alloys. The hexagonal-structure compositions generally possess the highest strength at elevated temperatures (600–1300°F or 316–705°C) and the best weldability. However, in general, these alloys have the lowest room-temperature strength and are not heat-treatable.

Commercially pure titanium alloys, containing minor amounts of oxygen and nitrogen as strengtheners, are alpha alloys, and they find wide use in applications where maximum corrosion resistance and ease of fabrication are necessary. A higher-strength alpha alloy containing 5% aluminum and 2.5% tin is used in aircraft applications.

Alpha-beta alloys. As a group, alpha-beta alloys have higher strength and respond to heat treatment, but are less formable than alpha alloys. This class of titanium alloy accounts for more than half of all titanium metal products used.

The alpha-beta alloys vary widely in their composition, and therefore in their general characteristics, strength, heat treatability, and ductility. By far, the most common alpha-beta alloy is 6% aluminum plus 4% vanadium, the basic titanium alloy for jet engines and airframes.

Through heat treatment, the tensile strength of the 6% aluminum plus 4% vanadium alloy can be varied from 120,000 to 180,000 lb/in.^{2} (827 to 1240 megapascals).

Beta alloys. Titanium can be made to exist entirely in the beta phase at room temperature by adding alloys which inhibit the beta-to-alpha transformation. The alloys are extremely ductile at room temperature and can be heat-treated easily to strengths exceeding 200,000 lb/in.^{2} (1380 MPa). *See* HEAT TREATMENT (METALLURGY).

Corrosion resistance. Titanium's corrosion resistance is based on its highly reactive metal surface. This protective and tenacious oxide film provides excellent corrosion resistance in chloride brines, especially those of a neutral or oxidizing character. In addition, titanium demonstrates good corrosion resistance in most organic media. **Table 1** shows this corrosion resistance in several industrially important environments.

Unfortunately, titanium possesses poor resistance to reducing mineral acids such as sulfuric or hydrochloric acid. However, impurities common to many commercial environments act as powerful inhibitors. In reducing acids, impurities such as iron, copper, chromium, cobalt, and nickel behave as oxidizers, altering the oxidation-reduction characteristics. The environment becomes oxidizing, protecting the oxide surface. This has resulted in the use of titanium in several industries such as metal finishing and hydrometallurgy. Hydrometallurgical applications are particularly attractive, since copper and nickel ores are routinely digested by using hot sulfuric acid in titanium autoclaves. *See* HYDROMETALLURGY.

Additions of noble metal alloys as low as 500 parts per million improve the corrosion resistance of titanium. Platinum and palladium additions lower corrosion rates in reducing acids by several orders of magnitude. This is incorporated in a popular alloy produced as ASTM (American Society for Testing and Materials) Grade 7 and includes 1500 ppm palladium.

A coating of a platinum-group metal oxide attached to a titanium substrate results in a

TABLE 2. Typical titanium applications for various markets

Market	Application
Gas turbine engines	Compressor blades, disks, ducts, cases
Airframes	Landing gear beams, wing structure, hydraulic tubing
General chemical	Heat exchangers, condensers, mixers, piping
Organic/petrochemical	Strippers, reboilers, condensers, reactors
Power plants	Surface condensers, turbine blades
Electrolysis	Anodes for chlorine, chlorate, manganese dioxide; cathodes for copper, manganese; cathodic protection of bridges
Pulp/paper	Bleach tanks, wet chlorine systems, drum washers
Water technology	Flash desalination, heat exchangers for desalination
Metal recovery	Plating of chromium, nickel, silver, gold, zinc and galvanizing; hydrometallurgy of copper
Energy extraction	Logging tools, seals, springs, tubulars for sour gas and geothermal power
Medical	Prosthetic implants, instruments
Environmental	Flue gas desulfurization, wet air oxidation of waste, incinerator stacks, nuclear waste
Marine	Heat exchangers, piping systems, ball valves, sonar masts
High-performance vehicles	Racing valves, springs, retainers, connecting rods

dimensionally stable anode with low chlorine overvoltage. Compared to mercury or graphite, this anode represents a significant improvement in the chlor-alkali industry. The success of the dimensionally stable anode was possible only because of the titanium substrate's exceptional corrosion resistance even under extremely oxidizing chloride conditions. An extension of this technology has been applied to concrete structures to prevent the steel reinforcement bars from corroding.

Titanium has low thermal conductivity but performs well in heat-transfer applications. Several titanium characteristics are combined to produce high-performance, competitive heat exchangers; they include zero corrosion allowance, high strength with low density, near immunity to process upsets, unlimited flow rates, and no corrosion film. The use of high-ductility ASTM Grades 1 or 2 in heat exchangers can in many cases produce a unit one-tenth the size of comparable graphite or copper-nickel units. *See* CORROSION; HEAT EXCHANGER.

Titanium industry. Titanium sponge plants are operating in five countries: Russia, the United States, Japan, the United Kingdom, and the People's Republic of China. Melting and processing capabilities are present in most countries having an aircraft industry.

The United States titanium market has been growing at an annual rate of 4% since 1972, and represents approximately 50% of the total world market excluding Russia. The capacity for producing titanium in Russia is very large, about equal to the world capacity. The United States market is 80% aerospace and 20% nonaerospace. In the remaining non-Russian world, it is 50% for each of these applications.

Because of the decreasing price of titanium mill products and the confidence gained through evaluations conducted during previous years, increasing quantities of titanium mill products are being used outside the aerospace sectors. Primary areas of such applications are chemical processing equipment; dimensionally stable anodes for the production of chlorine; and tubing for electrical-power-generating-plant surface condensers and heat exchangers used in oil refineries, desalination plants, and pollution control devices. Typical applications for titanium markets are shown in **Table 2**.

A significant quantity of titanium metal is used as an alloy addition in steel, aluminum, and nickel-base alloys as a strengthener, a grain refiner, and an improver of corrosion resistance. These applications have become an important adjunct of the titanium metal-producing industry. *See* METALLURGY; TITANIUM. Ward W. Minkler; Stanley R. Seagle

Bibliography. P. Crowson (ed.), *Minerals Handbook: 1994–95*, 1994; M. J. Donachie, Jr., *Titanium: A Technical Guide*, ASM International, 1988; D. Eylon, *Titanium for Energy and Industrial Applications*, Metallurgical Society AIME, 1987; F. H. Froes and I. Caplan, *Tritium Science and Technology*,1992; Institute of Metals, London, *Designing with Titanium*, 1986; R. W. Schutz, *Process Industries Corrosion*, 1986; S. R. Seagle and D. E. Thomas, Status of titanium technology, *Chem. Eng. Prog.*, pp. 63–68, June 1986; Titanium Development Ass., *Titanium 1986: Products and Applications*, 1986; U.S. National Materials Advisory Board, *Titanium: Past, Present and Future*, NMAB-392, 1983.

Titanium oxides

Chemical compounds of the metal titanium and oxygen. The most commonly found and used titanium oxides are the titanium dioxides, TiO_2; but other oxides are known including the sesquioxide Ti_2O_3, the monoxide TiO, and nonstoichiometric phases TiO_x, with x taking values between 0.7 and 1.3. Titanium dioxide exists in three common crystalline forms under ambient conditions: rutile, anatase (also known as octahedrite), and brookite (see **table**). Each polymorph contains titanium atoms surrounded by a distorted octahedron of oxygen atoms, and each form differs in the way in which the octahedral units are linked by various combinations of edge and corner sharing to give extended network structures (see **illustration**). Rutile is considered to be the most stable form of TiO_2, since at

Some physical properties of TiO_2 polymorphs

	Rutile	Anatase	Brookite
Crystal system	Tetragonal	Tetragonal	Orthorhombic
Space group	$P4_2/mmm$	$I4_1/amd$	$Pbca$
Unit cell	$a = b = 4.5845$ Å $c = 2.9533$ Å	$a = b = 3.7842$ Å $c = 9.5146$ Å	$a = 9.184$ Å $b = 5.447$ Å $c = 5.145$ Å
Density	4.2743 g cm^{-3}	3.895 g cm^{-3}	4.123 g cm^{-3}
Refractive index	2.60–2.90	2.49–2.55	2.58–2.70

temperatures above 500°C (930°F) both anatase and brookite are converted into rutile. *See* OXIDE; RUTILE; TITANIUM.

TiO_2 production. Both rutile and anatase are found in nature as minerals. There are also other more complex titanium-containing oxides, such as ilmenite ($FeTiO_3$) and leucoxene ($TiO_2 \cdot xFeO \cdot yH_2O$), from which the titanium dioxides may be extracted. These titanium ores are widespread throughout the world but are most readily exploited in Australia, the United States, India, and South Africa. Most of the titanium dioxide for commercial applications is extracted from titanium ores using one of two processes. In the chloride process, titanium ore is chlorinated using chorine gas and charcoal to yield volatile $TiCl_4$, which is separated by distillation from impurities and then oxidized by flame treatment to yield pure TiO_2. In the sulfate process, titanium ore is dissolved in sulfuric acid to produce a suspension of titanium oxyhydroxide from which TiO_2 is produced after filtration and firing. *See* ILMENITE.

Applications. The most widespread use of titanium dioxides is in the area of white pigments. Both anatase and rutile have exceptionally high refractive indices and therefore scatter light very effectively, offering highly effective opacity or hiding power as well as imparting whiteness and brightness to products. Rutile is most widely used since it has superior properties over anatase (a higher refractive index and a higher density, therefore offering greater opacity per gram), and has major pigmentary applications in coatings, paints, paper, inks, plastics, and rubbers, as well as in foods, cosmetics, toothpastes, and ultraviolet (UV) protection products. The particle size of TiO_2 pigments is an important property to control since the most effective light-scattering properties are exhibited by submicrometer crystals. Therefore, an important step in the manufacture of TiO_2 pigments is grinding or milling of powders to give the desired particle characteristics. TiO_2 pigments are not degraded during paint and plastic manufacture, giving them significant advantages over organic white pigments. Titanium dioxide pigments are considered to have almost no health hazards. They are nontoxic by ingestion, inhalation, or skin contact, and they are noncorrosive and noncombustible. For these reasons, TiO_2 white pigments are without commercial competition. Around 4 million metric tons of titanium dioxide pigment are produced worldwide annually. *See* PIGMENT (MATERIAL).

Photocatalysts. Titanium dioxides have been widely investigated as photocatalysts for a variety of applications, including catalyst materials for the production of energy by the decomposition of water into oxygen and hydrogen for solar energy conversion and the degradation of organic pollutants. These applications rely on the fact that TiO_2, both as anatase and rutile forms, is a semiconducting material with a band gap appropriate for the absorption of ultraviolet radiation by the excitation of valence electrons. This light-absorption process results in the formation of electron holes in the valence band and free electrons in the conduction band of TiO_2. Upon reaction with water, this excited form produces reactive hydroxyl and superoxide radicals that are able to initiate the destruction of organic pollutants into carbon dioxide and water. This process has been widely researched as a means of cleaning contaminated water from pollutants such as cyanide. One striking application of TiO_2 films is as coatings for self-cleaning

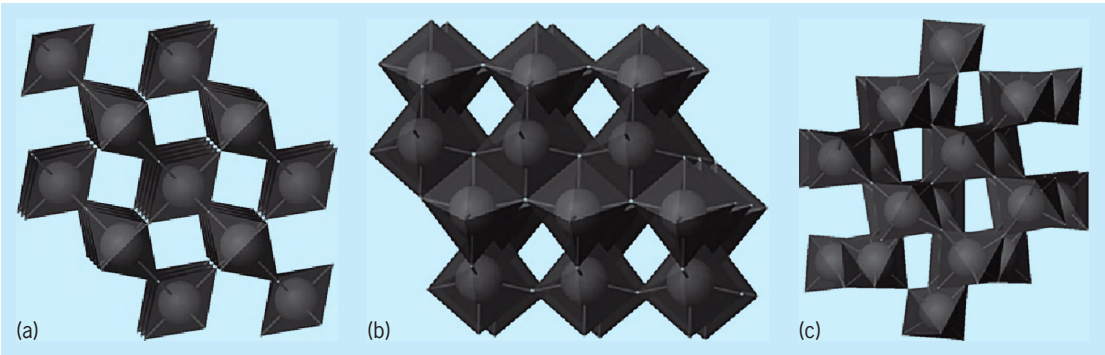

Representations of the crystal structures of the three polymorphs of titanium dioxide, TiO_2: (*a*) rutile, (*b*) anatase, and (*c*) brookite. The octahedral units represent titanium atoms surrounded by six oxygen atoms.

windows and tiles, where the destruction of organic deposits such as bacteria or odorous chemicals occurs upon the action of sunlight, and coupled with the hydrophilic nature of TiO_2 surfaces, washing is easily achieved by water alone. *See* BAND THEORY OF SOLIDS; ELECTRON-HOLE RECOMBINATION; FREE RADICAL.

Rechargeable batteries. One further use of more complex titanium oxide materials lies in rechargeable battery materials. Lithium/manganese/titanium rechargeable batteries are compact and were developed for watches and as backup power supplies for electronic devices such as pagers and timers. The batteries employ lithium–manganese complex oxide as the cathode material, and lithium–titanium oxide as the anode material. These batteries provide a capacity that is more than 10 times that of capacitors of the same size. *See* BATTERY; CAPACITANCE.

Richard Walton

Bibliography. J. H. Braun, Titanium dioxide: A review, *J. Coat. Technol.*, 69:59–72, May 1997; O. Carp, C. L. Huisman, and A. Reller, Photoinduced reactivity of titanium dioxide, *Prog. Solid State Chem.*, 32:33–177, 2004; A. Fujishima, T. N. Rao, and D. A. Tryk, Titanium dioxide photocatalysts, *J. Photochem. Photobiol. C: Photochem. Rev.*, 1:1–21, 2000; A. B. G. Lansdown and A. Taylor, Zinc and titanium oxides: Promising UV-absorbers but what effect do they have on intact skin?, *Int. J. Cosmetic Sci.*, 19:167–172, 1997.

Titration

A quantitative analytical process that is basically volumetric. However, in high-precision titrimetry the titrant solution is sometimes delivered from a weight buret, so that the volumetric aspect is indirect. Generally, a standard solution, that is, one containing a known concentration of substance X (titrant), is progressively added to a measured volume of a solution of a substance Y (titrand) that will react with the titrant. The addition is continued until the end point is reached. Ideally, this is the same as the equivalence point, at which an excess of neither X nor Y remains. If the stoichiometry or exact ratio in which X and Y react is known, it is possible to calculate the amount of Y in the unknown solution.

The normal requirements for the performance of a titration are: a standard titrant solution; calibrated volumetric apparatus, including burets, pipets, and volumetric flasks; and some means of detecting the end point.

Sometimes the standard titrant solution is prepared by the accurate weighing out of a primary standard, a substance that is stable and readily available in a high state of purity, such as potassium dichromate, anhydrous sodium carbonate, potassium hydrogen phthalate, or silver (which dissolves easily in nitric acid). The weighed material is dissolved, usually in water, and the solution is made up to a fixed volume in a volumetric flask of suitable capacity. The final solution must be homogeneous, a state normally achieved by ample shaking. If better-than-routine accuracy is required, the temperature must be known and constant.

Many valuable titrant solutions cannot be readily and accurately prepared by the direct method. For example, sodium hydroxide solution, much used for titrating acidic substances, is made from a solid that rapidly picks up moisture and so forth from the air. However, a solution of approximately the desired concentration can be prepared. This solution is then standardized by titrating (or being titrated with) another suitable standard solution. Sometimes a known weight of a suitable primary standard is placed directly in the titration vessel, dissolved in a suitable but unmeasured volume of solvent, and then titrated with the solution to be standardized. For example, sodium hydroxide solution is commonly standardized by titrating a known weight of potassium hydrogen phthalate. Analogous procedures are used for other titrant solutions that are made from substances that do not possess primary-standard properties.

Skillful operation may allow the titration to be stopped almost exactly at the end point. However, it is easy to overshoot, or add too much titrant. Provided that the total volume of titrant solution is noted, the titration can often be saved by the process of back titration. In the case mentioned above, a small overshoot of sodium hydroxide could be measured by back titration with standard hydrochloric acid solution. The process of back titration is sometimes used deliberately, especially for slow titration reactions. A known but excessive amount of titrant A is added to the substance to be determined. Then, after a suitable delay to allow for completion of the reaction, the remaining A is back-titrated with a reagent B that reacts rapidly with A. *See* CONCENTRATION SCALES; STOICHIOMETRY.

Classification by Chemical Reaction

For the purposes of titrimetry, chemical reactions can be placed in three general categories: acid-base or neutralization, combination, and oxidation-reduction.

Acid-base reactions. These titrations involve neutralization of an acid by titration with a base, or vice versa. However, the process is often nonspecific; in the titration of a mixture of nitric and hydrochloric acids, only the total acidity can be found without recourse to additional measurements. This arises because the only real reaction involved in these aqueous systems is the formation of water, reaction (1).

$$H_3O^+ + OH^- \rightarrow 2H_2O \qquad (1)$$

A salt derived from a strong base and a very weak acid can often be titrated just as if it were a base. For example, solutions of hydrochloric acid, or of other strong acids, are often standardized by titrating known weights of primary standard sodium carbonate. Sodium hydroxide solutions are often standardized by the use of primary standard potassium

hydrogen phthalate, which is a so-called acid salt. *See* ACID AND BASE.

Combination reactions. In titrimetry, attention is usually focused upon the combination of an ion in the titrant with one of the opposite sign in the titrand solution. Sometimes the combination may involve more than two species, some of which may be nonionic. The combination may result in precipitation. A classic example is the determination of chloride ion by titration with silver nitrate solution, when silver chloride is precipitated. Although a reaction may be known to yield a precipitate that has insignificant solubility and constant composition, this does not guarantee that the reaction will be useful titrimetrically. Any practical titration also requires that the reaction involved must not be unduly slow.

The same limitation also applies to complex-formation titrations, where precipitation may be absent or merely incidental. Except in a few special cases, complex-formation titrations were of little importance until the discovery of ethylenediaminetetraacetic acid (EDTA) and related compounds. These titrants not only are powerful complexing agents that combine with very many cations, but also form a single type of complex. This is in marked contrast to complexing agents such as ammonia or cyanide ion, which may yield a mixture of complexes.

With ubiquity comes lack of specificity. This difficulty with the EDTA family of titrants can sometimes be overcome by pH control and by technique masking (competitive complexation, whereby some species are prevented from taking part in the titration reaction). Then it may be possible to determine separately some or all of several metal ions in a mixture. *See* COORDINATION COMPLEXES; PRECIPITATION (CHEMISTRY).

Oxidation-reduction reactions. In so-called redox titrations the titrant is usually an oxidizing agent, and is used to determine a substance that can be oxidized and hence can act as a reducing agent. Because titrants that have usefully strong reducing properties are themselves attacked by oxygen in the air, the reverse procedure, although possible, is less usual. *See* OXIDATION-REDUCTION.

Coulometric Titration

Faraday's laws of electrolysis indicate that the extent of an electrochemical reaction is proportional to the total amount of electricity that is passed through the system. The passage of a uniform current for a measured period of time can be used to generate a known amount of a product such as a titrant. This fact is the basis of the technique known as coulometric titration. An obvious requirement is that generation shall proceed with a fixed, preferably 100%, current efficiency. The uniform current is then analogous to the concentration of an ordinary titrant solution, while the total time of passage is analogous to the volume of such a solution that would be needed to reach the end point.

Coulometric titration has several attractive features. Unless the sample to be titrated must be measured by volume, no volumetric glassware is needed. Nor are standard solutions; the titrant is generated during the process. This means that titrants of low stability, such as the strong reducing agent chromium(II) ion, can be employed routinely. The titration is started or stopped by the mere closing or opening of a switch, which can be remote from the titration vessel. This feature is obviously useful when titrating highly radioactive materials. The process can be controlled by an operator who is safely screened from the materials in the titration system. *See* ELECTROLYSIS.

Classification by End-Point Techniques

The precision and accuracy with which the end point can be detected is a vital factor in all titrations. Because of its simplicity and versatility, chemical indication is quite common, especially in acid-base titrimetry. That certain natural pigments such as litmus in an acidic solution change color when the solution is made basic has been known for centuries. Thus an acid solution that contains a small amount of litmus can be titrated with sodium hydroxide solution until the initial red color has just changed to blue. Litmus and other natural pigments, which are usually mixtures of various compounds, have been supplanted by synthetic indicators. These not only have sharper color transitions, but also can be made to suit particular applications.

Indicators. An acid-base indicator is a weak acid or a weak base that changes color when it is transformed from the molecular to the ionized form, or vice versa. The color change is normally intense, so that only a low concentration of indicator is needed. Phenolphthalein is an example of an indicator that acts as a weak monobasic acid according to reaction (2).

$$\underset{\substack{\text{Acidic form}\\ \text{(colorless)}}}{\mathrm{HIn}} \rightleftharpoons \underset{\substack{\text{Basic form}\\ \text{(red)}}}{\mathrm{H^+ + In^-}} \tag{2}$$

Methyl orange acts as a weak base according to reaction (3).

$$\underset{\substack{\text{Basic form}\\ \text{(yellow)}}}{\mathrm{InNR_2 + H^+}} \rightleftharpoons \underset{\substack{\text{Acidic form}\\ \text{(red)}}}{\mathrm{InNR_2H^+}} \tag{3}$$

The working range, or visual color change, of a typical acid-base indicator is spread over a hundredfold (~2 pH units) change in hydrogen ion concentration. Available indicators have individual working ranges that together cover the entire range of hydrogen ion concentration (10^{-1} to $10^{-13}M$, or pH 1 to 13) likely to be encountered in general acid-base titration. For example, the working ranges of phenolphthalein and of methyl orange are pH 8.0–9.8 and pH 3.0–4.4, respectively.

The success of a titration may hinge upon a suitable choice of indicator. **Figure 1** shows the titration curves of two monobasic acids (a strong acid, such as hydrochloric acid, and a moderately weak acid,

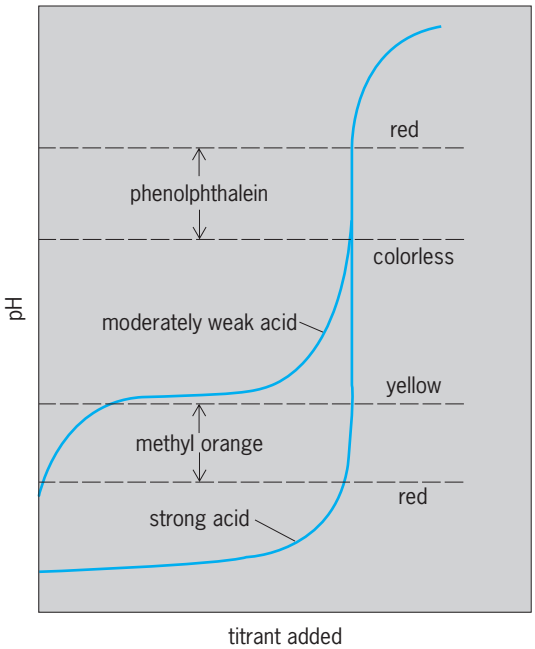

Fig. 1. Titration curves of two monobasic acids (0.1 *M*), with sodium hydroxide as the titrant.

such as acetic acid), each at a concentration of approximately 0.1 *M*, with the strong base sodium hydroxide as titrant. The respective working ranges of methyl orange and of phenolphthalein are also indicated. The curve for the strong acid is almost vertical in the approximate region pH 4–10. This means that a very small addition of titrant when the pH is approximately 4 causes methyl orange to change from orange to yellow. If phenolphthalein is used in place of methyl orange, the colorless indicator begins its change to the red form when the pH is approximately 8. Theoretically the end point with phenolphthalein requires a little more titrant than that required to reach the methyl orange end point. However, under routine conditions, the difference may be less than calibration-plus-operator error.

In the titration of the weak acid with methyl orange as the indicator, the pH rises to the lower limit of the working range before much titrant has been added. The curve rises slowly within this working range, so that the color change red through orange to yellow is both gradual and complete before all of the acid has been titrated. However, the weak acid curve eventually merges into the strong acid curve. A titration indicated by phenolphthalein therefore yields a sharp end point.

Similar reasoning applies to the titration of a base with an acid solution. If both species are strong, any one of a number of common indicators suffices. The successful titration of a weak base such as ammonia with a strong acid requires an indicator like methyl orange, which has a working range that is low on the pH scale. In the titration of a weak acid with a weak base, or the reverse, no simple indicator is suitable because the titration curve is not steep anywhere near the expected end point. The standardization of a solution of ammonia with standard acetic acid solution would be done indirectly. *See* ACID-BASE INDICATOR; HYDROGEN ION; PH.

In the complexometric titration of a cation, M^{n+}, such as that of magnesium, with EDTA or a similar agent, the concentration of M^{n+} falls most rapidly in the immediate region of the end point. The response resembles that shown by the curve of a strong acid in Fig. 1, but pM^{n+} replaces pH. The indicator is itself a suitable complexing agent that changes color when it is combined with M^{n+}. This is the situation during most of the titration. At or very near the expected end point, the titrant reacts with the small amount of cation in the indicator complex. The color then changes to that of the free agent.

When a solution of an oxidizing agent is used to titrate a substance that is readily oxidized, the electrochemical potential rises most rapidly in the region of the end point. The titration curve is generally similar to that shown for the weak acid in Fig. 1, but potential replaces pH. The chemical indicator is here a substance that undergoes a marked change in color when a suitable potential is reached. Unlike the indicators commonly used in acid-base or complex-formation titrations, oxidation-reduction indicators are often irreversible. Thus, if an excess of oxidizing titrant has been added, the indicator color may not revert to the original if back titration is subsequently attempted.

The important oxidizing titrant potassium permanganate forms an intensely purple solution. Its reduction products are essentially colorless. This titrant is thus self-indicating, provided that the titrand and its oxidation products have little or no color. The end point is taken as the first appearance of a pale but permanent pink color.

Sometimes no suitable chemical indicator can be found for a desired titration. Possibly the concentrations involved may be so low that chemical indication functions poorly. Other situations might be the need for high precision or for the automatic arrest of the titration. Recourse is then made to some physical method of end-point detection.

Potentiometric titration. If a pH meter is used, acid-base titration curves like those shown in Fig. 1 can be plotted or recorded. The meter and its associated electrodes are first standardized by use of a buffer solution of known pH. The electrodes are then immersed in the well-stirred solution to be titrated, and the titration is begun. For routine purposes, interest is in rapid and reasonably close end-point location. Titration is first carried out quite quickly until the meter shows signs of rapid response, and then is slowed so that it can be stopped at the pH jump or fall that marks the end point. Obviously, potentiometric titration can be used for a highly colored titrand solution, in which the response of a color-change indicator could not be seen.

There are two approaches to higher precision. In the first the pH at which the desired end point occurs is determined, and the titration of the actual sample is then arrested exactly at, or very close to, this pH value. The other approach is to stop at the point of steepest slope of the titration curve. This can

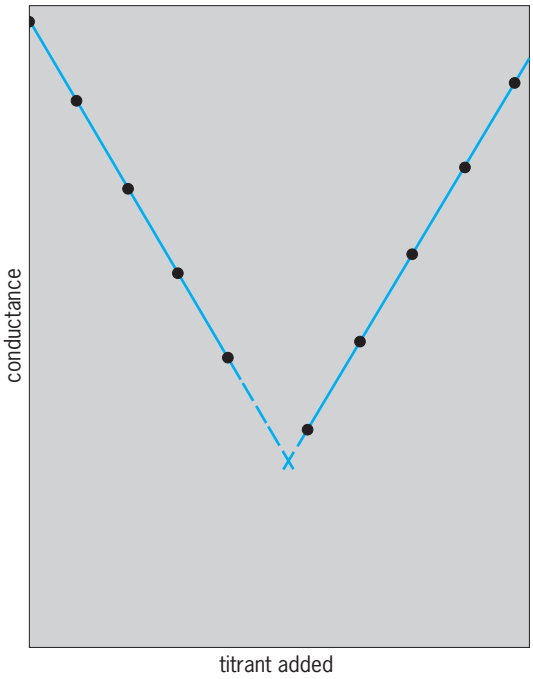

Fig. 2. Conductometric titration curves of a dilute hydrochloric acid solution with sodium hydroxide as the titrant.

be done by plotting (or, with suitable instrumentation, sensing) the first derivative ($\Delta pH/\Delta V$) or the second derivative ($\Delta^2 pH/\Delta V^2$) against the volume V of titrant added. In theory this method is applicable to any potentiometric titration, without the need to predetermine the end-point conditions.

By suitable choice of electrodes, these potentiometric methods can also be applied to combination titrations and to oxidation-reduction titrations. The advent of modern ion-selective electrodes has greatly extended the scope of potentiometric titration and other branches of titrimetry. *See* ELECTRODE POTENTIAL.

Conductometric titration. Several instrumental techniques give rise to titration curves that are essentially linear. Generally, the end point is found graphically from readings taken at a number of points that fall on appropriate linear portions of the curve. One advantage is that poor response in the actual vicinity of the end point does not cause difficulties. Another is good precision, due to the averaging effect of the many individual readings.

The underlying principles of conductometric titration are that the solvent and any molecular species in solution exhibit only negligible conductance; that the conductance of a dilute solution rises as the concentration of ions is increased; and that at a given concentration the hydrogen ion and the hydroxyl ion are much better conductors than any of the other ions.

One example is that of a dilute hydrochloric acid solution titrated with one of sodium hydroxide. The initial conductance, high because the hydrogen ion concentration is high, falls as the titrant is added (**Fig. 2**). When neutralization is complete, the conductance is that of a solution of sodium chloride. The change in conductance is due to progressive replacement of hydrogen ion by sodium ion, which is a poorer conductor. Further addition of titrant causes the conductance to rise again, principally because the highly conducting hydroxyl ion now remains in the system, instead of undergoing mutual destruction with the hydrogen ion. Addition of titrant naturally dilutes the titrand solution, causing distortion of the linear branches of the titration curve. This diluting effect can be compensated for by a simple arithmetic correction or can be rendered negligible by use of a concentrated titrant solution.

Conductometric titration is sometimes successful when chemical indication fails. A typical case is the titration of acetic acid with ammonia (**Fig. 3**). Both reactants are weak electrolytes, but ammonium acetate, the salt produced in the titration, is a strong electrolyte. The rising linear branch is due to the buildup of the ions of this salt. In the region beyond the end point, the addition of more titrant causes very little change in conductance. The concentration of ammonium ion then present forces the added titrant to remain almost entirely in the molecular state. Because of the effects of hydrolysis, the graph is decidedly rounded in the region of the end point. Although an old technique, conductometric titration still finds extensive use in studies of nonaqueous systems.

Certain precipitations can be satisfactorily performed conductometrically. However, conductometry is generally of little use in oxidation-reduction titrations. In fact the technique is difficult to use if the system contains appreciable concentrations of electrolytes other than those involved in the actual

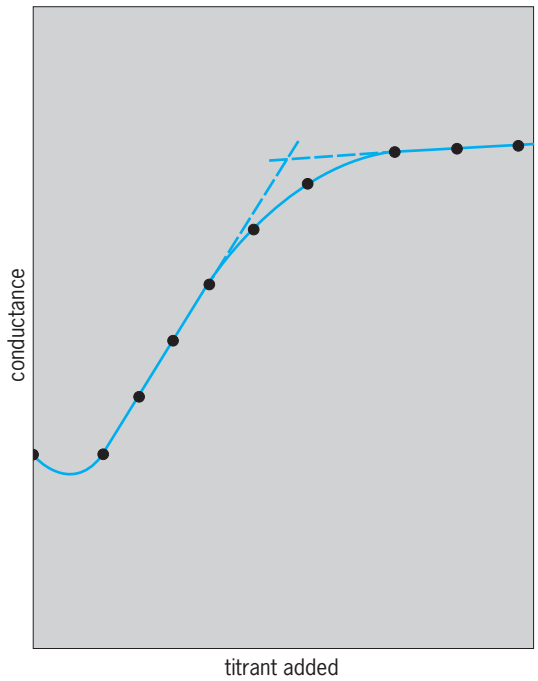

Fig. 3. Conductometric titration curve of acetic acid with ammonia as the titrant.

titration. The observed conductance is a function of the total ionic content, while the precision with which the end point can be determined depends upon the relative change in conductance. For example, the angle between the branches in Fig. 2 is less acute if the titrand solution contains sodium nitrate as well as hydrochloric acid. *See* ELECTROLYTIC CONDUCTANCE.

Conductometric titrations are normally carried out with the aid of audio-frequency alternating current. By use of radio frequencies, the titration can be monitored in a cell that has the electrodes on the outside, so that they are not in contact with the titrand solution. However, the titrand curves are usually less simple than those encountered in normal conductometric titration.

Spectrophotometric titration. The spectrophotometer is an optical device that responds only to radiation within a selected very narrow band of wavelengths in the visual, ultraviolet, or infrared regions of the spectrum. The response can be made both quantitatively and linearly related to the concentration of a species that absorbs radiation within this band. Titrations at wavelengths within the visual region are by far the most common. An example is the titration of iron(II) in dilute sulfuric acid solution with a standard solution of potassium permanganate. The spectrophotometer is adjusted to measure the absorbance of the highly colored titrant. Very little absorbance is observed until the end point is reached; iron(II) and the reaction products are essentially colorless and are not "seen" by the spectrophotometer. However, the absorbance rises as titration is continued beyond the end point because the titrant is no longer being destroyed. The titration curve is shown in **Fig. 4**.

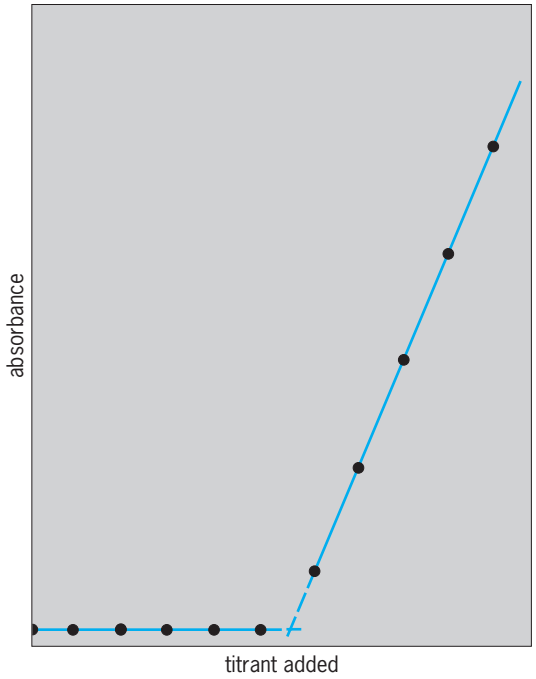

Fig. 4. Spectrophotometric titration curve of iron(II) in dilute sulfuric acid solution with potassium permanganate as the titrant. The curve shows the response when the titrant absorbs beyond the end point.

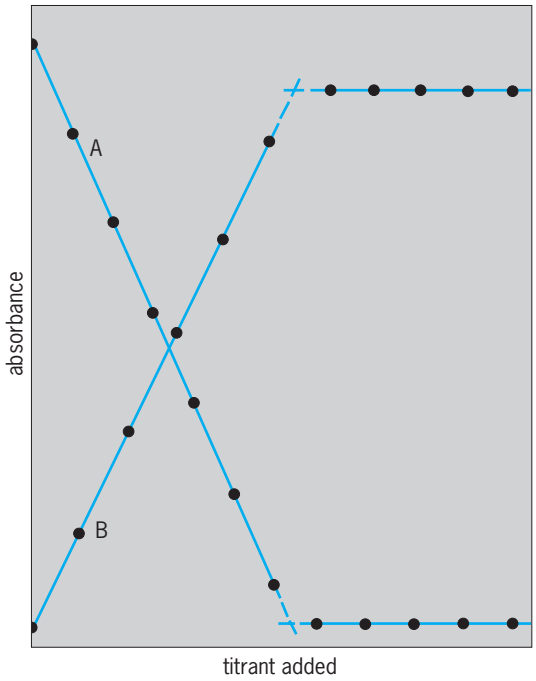

Fig. 5. Spectrophotometric titration curves: A, when the titrand absorbs but the products do not, and B, when neither of the reactants absorbs but the products do.

If the titrant does not absorb at the chosen wavelength, the curve is horizontal in the region beyond the end point. If the titrand absorbs but the products do not, the curve has the form of curve A in **Fig. 5**. Curve B shows the response when neither of the reactants absorb, but the products are strong absorbers.

In some cases, systems that involve no absorbing species can be handled by the addition of an optical indicator. This is a substance that undergoes no change in absorbance until the titrant is present in slight excess over that required for the main reaction.

Amperometric titration. By the use of a dropping-mercury or other suitable microelectrode, it is possible to find a region of applied electromotive force (emf) in which the current is proportional to the concentration of one or both of the reactants in a titration. **Figure 6** shows the separate current-voltage curves, at approximately similar concentrations, of two species X and Y that can react to form nonelectroactive products. For example, X and Y may form an essentially insoluble precipitate. Provided that conditions are suitable, the currents may be anodic or cathodic.

Suppose that the applied emf is fixed within the range *pq* and that Y is being titrated with X. The current remains at the small residual value until near the end point and rises as the titration is carried past the end point. The general shape of the titration curve is thus similar to that shown in Fig. 4. Titration at a fixed emf within the range *rs* gives a V-shaped curve, because both titrant and titrand are electroactive under these conditions.

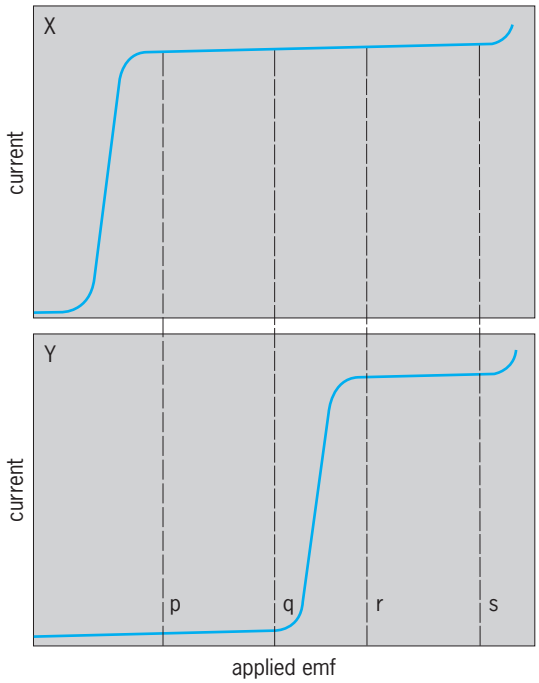

Fig. 6. Current-voltage curves at similar concentrations for two species (X and Y) that can react to form nonelectro active products.

Amperometric titration has been applied to all classes of reactions. Foreign ions may be present (and are often deliberately added to suppress migration currents), provided that they neither interfere with the titration reaction nor are electroactive under the chosen conditions.

Biamperometric titration is a closely related technique. An emf that is usually small is applied across two identical microelectrodes, usually of platinum, that dip into the titrand solution. This arrangement, which involves no liquid-liquid junctions, is obviously valuable in nonaqueous titrations, but also finds much use in aqueous titrimetry. The titration curves are nonlinear, so that direct observation, rather than graphing, is normally used to find the end point. *See* POLAROGRAPHIC ANALYSIS.

Thermometric or enthalpimetric titration. Many chemical reactions proceed with the evolution of heat. If one of these is used as the basis of a titration, the temperature first rises progressively and then remains unchanged as the titration is continued past the end point. If the reaction is endothermic, the temperature falls instead of rising. Thermometric titration is obviously applicable to all classes of reactions.

The total temperature change may be small, so that transfer of heat between the titration vessel and its surroundings must be minimized. Temperature changes are usually measured by means of a thermistor, which has both small heat capacity and a high negative coefficient of electrical resistance. The lower specific heats of solvents other than water indicate that the thermometric titration is particularly suited to nonaqueous titration. *See* THERMOCHEMISTRY.

Nonaqueous titration. Water, the cheapest solvent, may react with, or may not dissolve, the substances to be used in a titration. Recourse must then be made to nonaqueous titrimetry. However, the main use of this technique is to perform titrations that give poor or no end points in water. Although applicable in principle to all classes of reactions, acid-base applications have greatly exceeded all others.

When an acid AH is dissolved in an amphiprotic solvent SH, and equilibrium is established as shown in reaction (4), if SH is strongly basic (protophilic),

$$AH + SH \rightleftharpoons A^- + SH_2^+ \qquad (4)$$

the equilibrium is forced to the right. All acids, whether weak or strong in water, tend to behave as strong acids. This is termed the leveling effect. Compounds such as phenol are too weakly acidic to be titrated in water. However, the titration can be performed in a solvent such as ethylenediamine. A common titrant is tetrabutylammonium hydroxide in the same solvent.

Analogously, a strongly acidic (protogenic) solvent exerts a leveling effect on basic solutes. Pyridine and other bases that are very weak in water can be nicely titrated in anhydrous acetic acid. The usual titrant is perchloric acid in the same solvent. By suitable choice of solvent, the apparent strengths of individuals in a mixture of acids or bases can sometimes be spread out, so that a succession of separate end points is obtained.

The coefficient of thermal expansion of a typical nonaqueous solvent is several times greater than that of water. Titrations should therefore be carried out at, or very near to, the temperature at which the titrant was standardized. Nonaqueous titrations in which the solvent is a molten salt or salt mixture are also possible.

Automatic titration. Automation is particularly valuable in routine titrations, which are usually performed repeatedly. One approach is to record the titration curve and to interpret it later. This requires a buret of the syringe or pump type that has a constant rate of delivery and is coupled to the recorder to provide the volume-of-titrant axis. Another method is to stop titrant addition or generation automatically at, or very near to, the end point. Although a constant-delivery device is desirable, an ordinary buret with an electromagnetically controlled valve is often utilized.

The second method is common in automatic potentiometric titration. Overshoot must be prevented if precise and accurate results are needed in minimum time. Titrant addition, rapid during most of the process, is automatically slowed down when the potential is approaching the end-point value. Microcomputer control permits such refinements as the continuous adjustment of the titrant flow rate during the titration. In some cases it is possible to automate an entire analysis, from the measurement of the sample to the final washout of the titration vessel and the printout of the result of the analysis. *See* ANALYTICAL CHEMISTRY.

John T. Stock

Bibliography. G. D. Christian, *Analytical Chemistry*, 5th ed., 1993; G. W. Ewing, *Instrumental Methods of Chemical Analysis*, 5th ed., 1985; S. B. Khoo, *Analytical Chemistry*, 1990; E. Scholz, *Karl Fischer Titration*, 1984.

Tobacco

The plant genus *Nicotiana*, certain species in the genus, and dried leaves of these plants are all called tobacco. Most often tobacco means a leaf product containing 1–3% of the alkaloid nicotine, which produces a narcotic effect when smoked, chewed, or snuffed. The plant *N. rustica* provides tobacco in parts of Europe, but the tobacco of world commerce is *N. tabacum* (**Fig. 1**).

Tobacco is American in origin. Columbus found West Indians smoking it in a hollow forked stick. Historians do not know who first took tobacco to Europe, but most of them credit Jean Nicot in 1561. *Nicotiana* and nicotine incorporate his name. *See* SOLANALES.

Characteristics. This solanaceous annual, found only in cultivation, probably began as an amphidiploid or fertile hybrid between two species of *Nicotiana* native to Bolivia. It has 24 chromosome pairs, an erect, thick stem 2–9 ft (0.6–2.7 m) high, and alternate, sessile, oval, or lanceolate leaves. The flowers are produced in a panicle. They have a tubular corolla 3.5–5.5 cm long widening midway to a throat and ending in a five-pointed flare. There are five anthers which shed sticky pollen just as the flower opens. Most flowers are self-pollinated. The fruit is a two-section capsule with minute seeds (about 15,000 in a gram). *See* GENETICS; INFLORESCENCE.

Cultivation, harvesting, and curing. The local soil type and climate within a country, state, or even county determine the type of tobacco grown there.

Fig. 1. ***Nicotiana tabacum*, Connecticut cigar wrapper type. (*Connecticut Agriculture Experiment Station*)**

Generally all tobacco does best in a warm, even climate on light, well-drained, carefully fertilized soils that are clean cultivated and receive moisture weekly from rain or irrigation. Special seedbeds, either open or covered with cloth, glass, or plastic, provide transplants for fields. Plants are usually topped (blossom removed) to expand upper leaves. Harvest proceeds by cutting the whole stalk (stalk-cut) or picking leaves successively as they ripen (primed). Primed leaves are supported on strings, wires, or sticks.

Types of tobacco

Type	Major production areas	Use	Curing method
Flue-cured (bright)	United States (Virginia, North Carolina, South Carolina, Georgia-Florida border); Canada (Ontario); Rhodesia; Europe; China; Japan; Australia; India	Cigarettes	Primed leaves heat-cured to a bright yellow color
Light air-cured (Burley, Maryland)	United States (Kentucky, Maryland, Tennessee); Canada; Europe	Blending in cigarettes	Whole plant air-dried in ventilated sheds
Oriental (Turkish)	Greece (Macedonia); Turkey (Samsun, Smyrna)	Blending for aroma in cigarettes	Primed leaves cured in sun on strings
Cigar filler	Cuba*; Philippines; Puerto Rico; United States (Pennsylvania, Ohio); Dominican Republic; Central and South America	Central bulk of cigars	Whole plant air-dried in ventilated sheds
Cigar binder†	United States (Connecticut River Valley, Wisconsin, Pennsylvania)	Binding the filler into shape	Whole plant air-dried in ventilated sheds
Cigar wrapper	Under cloth shade: United States (Connecticut River Valley, Georgia-Florida border); Cuba; Italy; new in Canada and Central and South America Unshaded: Sumatra	Outer leaf or wrapper of cigar	Primed leaves heat- and air-dried to a uniform golden brown or a bright green (Candella)
Dark air-cured and fire-cured	United States (Kentucky, Tennessee, Virginia)	Pipe, chewing, snuff	Whole plant air-dried in ventilated sheds

*Embargoed in the United States since 1962.
† Acreage in Connecticut River Valley dropped 90% after 1956 after replacement by processed sheet reconstituted from pulverized tobacco.

Curing (drying) is done in ventilated barns with natural or artificial heat. In some areas, machine-harvested leaves are packed in special frames for curing in heat-regulated, forced-air chambers. Drying time ($^1/_2$–6 weeks) and temperature (70–170°F or 21–77°C) influence the amount and kind of changes that occur in proteins, carbohydrates, organic acids, alkaloids, and enzymes in the leaf. Before use in cigars, cigarettes, pipes, chewing tobacco, or snuff, cured leaves are fermented by storing them 6 weeks to 2 years at about 15% moisture and 80–110°F (27–43°C). The methods used for harvesting, curing, and fermenting depend on the type of tobacco, intended use, and local custom. A flat product made from ground-up tobacco plus binding agents is often blended with tobacco leaf.

Economic importance. Tobacco is economically important in 66 countries and is grown to some extend in all but a few countries. Principal tobacco types, major world production areas, uses, and curing methods are shown in the **table**.

Gordon S. Taylor

Diseases. Diseases cause production problems in all countries where tobacco is grown. Annual losses of tobacco products worldwide average 10% or more; in addition, production costs are invariably increased. About 25 principal pathogens attack the crop from seed sowing through marketing.

Root knot, caused by the nematode *Meloidogyne* sp., is the most important disease. It occurs throughout the warmer countries, particularly in sandy soils. Diseased plants are stunted and have wilted leaves that yellow prematurely. Knots or galls form on the roots (**Fig. 2**). Infected plants become more susceptible to other diseases such as black shank and bacterial wilt.

Lesion nematodes (*Pratylenchus* sp.) are also important root pathogens that kill the feeder roots, consequently depleting the root systems and resulting in stunted, weakened plants.

Black shank, caused by the soil-inhabiting fungus *Phytophthora parasitica*, also is a warm-weather disease. Tiny swimming spores (zoospores) formed by the fungus penetrate and kill the roots, the bases of the stems turn black, and the plants quickly die (**Fig. 3**). Water may carry the zoospores long distances to begin new outbreaks.

Fig. 2. Tobacco plant severely infected with root knot. (*USDA*)

Fig. 3. Black shank–infected tobacco plant with stalk split lengthwise to show typical decay and disking of the pith. Most of the roots are already dead. (*North Carolina State University*)

Fig. 4. Influence of (*a*) high (86°F or 30°C) and (*b*) low (68°F or 20°C) soil temperature on recovery of plants in black root rot–infested soil. The roots at 30°C were like those at 20°C at the beginning of the experiment. (*USDA*)

Black root rot, caused by *Thielaviopsis basicola*, a fungus that attacks many species of plants, is widespread in cooler soils. The fungus forms black lesions on the roots, which then decay.

Fig. 5. Bacterial wilt: young plant showing the one-sided wilting of green leaves characteristic of the disease. (*USDA*)

Severely infected plants have only a few black stubby roots attached to the base of the stem (**Fig. 4**).

Blue mold, caused by the downy mildew *Peronospora hyoscyami*, attacks the leaves and stems in both the seed bed and field during cool, wet weather. Small plants are killed in a few days. Tiny airborne spores, produced in great numbers, may be blown hundreds of kilometers to start new epidemics.

Pseudomonas solanacearum, which causes bacterial wilt, attacks wounded roots of plants grown in sandy soils of warm climates. Great numbers of the rapidly producing bacteria produce toxins and also interfere with water uptake and conduction. Leaves wilt (**Fig. 5**) and the plant dies.

Wildfire (angular leaf spot), caused by *Pseudomonas tabaci*, is the most widespread leaf disease, caused by a bacterium that enters the leaf through the stomates. The pathogen thrives in water-congested tissue and causes circular, yellow-green spots on the leaves. Often the spots fuse to form irregular dead areas that may destroy the entire leaf (**Fig. 6**).

Mosaic occurs worldwide and is the principal viral disease of tobacco. The virus is spread by contact. Rubbing a diseased leaf and then a healthy one will produce infection. Leaves of diseased plants are mottled with intermingled light- and dark-green areas (**Fig. 7**).

Cucumber mosaic, veinbanding (caused by potato virus Y), and etch viruses infect many wild and cultivated plants. They are spread by leaf-feeding aphids (*Myzus* spp.) that carry the virus on their stylets from infected to healthy plants. Symptoms caused by different viruses are often similar and vary with the cultivar, virus strain, and age of the plant when infected. Moreover, plants may be infected with more than one virus. Symptoms range from a mild mottling or mosaic, to leaf distortion, scalding, necrosis, and stunting of the plants. The tomato spotted wilt virus, widespread in Europe, deforms and stunts plants. It is transmitted by thrips. Control is difficult.

Fig. 6. Mature tobacco leaf showing wildfire symptoms. The yellow halo around the younger lesions is a typical symptom. The larger coalescing lesions often show no halo. (*USDA*)

Fig. 7. Mottled leaves of a young plant infected with tobacco mosaic virus. The distortion seen on the leaf at the right commonly occurs on plants infected when they have four to six leaves. (*North Carolina State University*)

Weather fleck is caused by air pollution. Ozone is the principal damaging agent. Numerous small brown spots form on affected leaves. Such leaves are worthless as cigar wrappers.

Drought, hail, drowning, frost, heat, excessive sunlight, herbicide injury, nutrient deficiencies, and toxicities, alone or in combination with disease, frequently decrease both yield and quality by stunting the plants, causing premature yellowing or death of leaf tissue, or upsetting the proper carbon/nitrogen ratio necessary for desirable smoke aroma and flavor.

No single disease control measure will control any disease satisfactorily. Efficient control requires a continuous integrated management program including: disease-resistant cultivars, healthy transplants; rotation with other crops; effective pesticides; destruction of plant debris by disking or plowing immediately after harvest; and weed and insect control. *See* PLANT PATHOLOGY. G. B. Lucas

Bibliography. G. B. Lucas, *Diseases of Tobacco*, 1975; H. D. Shew and G. B. Lucas (eds.), *Compendium of Tobacco Diseases*, 1991.

Toggle

Any of a wide variety of mechanisms, many used to open or close electrical contacts abruptly and all characterized by the control of a large force by a small one. The basic action of a toggle mechanism is shown in **illus.** *a*. When $\alpha = 90^\circ$, forces P and Q are independent of each other. Again, when $\alpha = 0^\circ$ the forces are isolated, force Q being sustained entirely by the frame, and force P serving only to hold the link in position. At $\alpha = 45^\circ$ from the symmetry $|P| = |Q|$, the mechanism serves to transfer the direction of forces to achieve equilibrium. If frame, connecting rod, and slider blocks with their pivots are sufficiently strong to support the maximum force encountered and if the coefficients of friction at the pin and slider joints are negligible, at intermediate positions the forces are related by the equation below.

$$P = Q \tan \alpha$$

The essential characteristic of the mechanism arises from the tangent relation between forces. As $\alpha \to 0$, a small force at P can overcome a large force at Q. *See* COUPLING.

The mechanism finds application in such machines as printing presses, embossing machines, friction brakes and clutches, and rock crushers. In such devices, the frame sustains the output force, the drive supplying only a small force through a sufficient distance to bring the linkage into position. Arranged to drive slightly past dead center, the mechanism serves as a clamp that can be released by a small reverse force at P, yet will hold against a large force at Q. Such action finds application in workpiece clamps, tool holders, and electric switches and circuit breakers. A spring is generally added to assist the release once the mechanism has been returned through dead center.

Because the simple configuration of illus. *a* requires low-friction sliders, it is impractical. A more useful structure replaces the vertical slider with a second link pinned to the frame (illus. *b*), in which case input P sets up forces in both links. A further modification (illus. *c*) replaces the other slider with a link. *See* FOUR-BAR LINKAGE; LINKAGE (MECHANISM). Frank H. Rockett

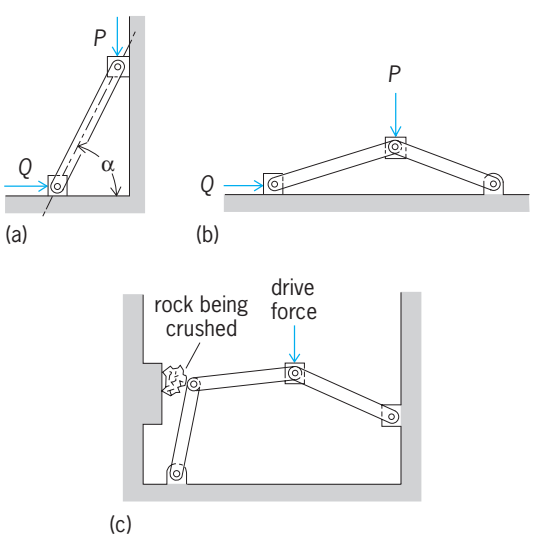

Toggle mechanism. (*a*) Simple structure. (*b*) Traditional configuration. (*c*) Typical application.

Tolerance

Amount of variation permitted or "tolerated" in the size of a machine part. Manufacturing variables make it impossible to produce a part of exact dimensions; hence the designer must be satisfied with manufactured parts that are between a maximum size and a minimum size. Tolerance is the difference between maximum and minimum limits of a basic dimension. For instance, in a shaft and hole fit, when the hole is a minimum size and the shaft is a maximum, the clearance will be the smallest, and when the hole is the maximum size and the shaft the minimum, the clearance will be the largest (see **illus.**).

If the initial dimension placed on the drawing represents the size of the part that would be used if it could be made exactly to size, then a consideration of the operating conditions of the pair of matting surfaces shows that a variation in one direction from the ideal would be more dangerous than a variation in the opposite direction. The dimensional tolerance should be in the less dangerous direction. This method of stating tolerances is called unilateral tolerance and has largely displaced bilateral tolerance, in which variations are given from a basic line in plus and minus values.

As an example, for a 1.5-in. shaft and hole for a free fit the standard allowance is 0.002 in. and the

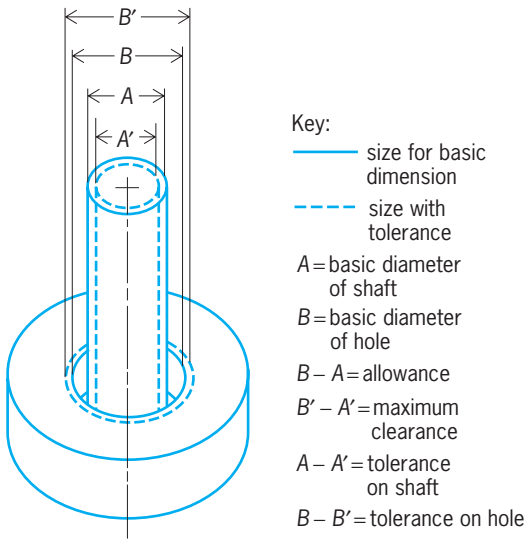

Maximum and minimum limits of the basic dimensions of a machine part.

tolerance for hole and shaft is 0.001 in.:

$$\begin{aligned}\text{Max. shaft diameter} &= \text{nominal size} - \text{allowance}\\ &= 1.500 - 0.002\\ &= 1.498 \text{ in.}\end{aligned}$$

In the unilateral method for stating tolerance, the shaft diameter is $1.498^{+0.000}_{-0.001}$, or between 1.498 and 1.497 in. The diameter of the hole is $1.500^{+0.001}_{-0.000}$, or between 1.500 and 1.501 in. The maximum clearance is 1.501 minus 1.497, or 0.004 in., and the minimum clearance is 0.002 in. Paul H. Black

Bibliography. D. A. Madsen, *Geometric Dimensioning and Tolerancing*, rev. ed., 1988; E. Oberg and F. D. Jones, *Machinery's Handbook*, 23d ed., 1988.

Tomato

An important vegetable, belonging to the genus *Lycopersicon*, especially *L. esculentum*, that is grown for its edible fruit. *Lycopersicon* species are native to South America, especially Peru, and the Galápagos Islands. The tomato was first domesticated in Mexico, and received its name from the Aztec word *xitomate*.

The tomato was introduced to Europe in the mid-sixteenth century, and was prominently featured in the early herbals. It was grown for the beauty of its fruit but was not often eaten, except in Italy and Spain, because many considered it to be poisonous like its relative the deadly nightshade. The tomato has since become one of the most popular vegetables in the world.

Although native to the New World, the tomato was introduced to America from Europe. It has been grown in the United States since colonial days, but it became an important vegetable there only in the past century.

Classification and structure. The genus *Lycopersicon* is a member of the Solanaceae, the nightshade family. *Lycopersicon esculentum*, the familiar tomato, can be hybridized with each of the eight other species of *Lycopersicon*. It can be easily crossed with *L. pimpinellifolium*, the red currant tomato. Crosses with the more distantly related, green-fruited *Lycopersicon* species are more difficult; usually the tomato must be the maternal parent, embryo culture is sometimes required, and the hybrids have varying degrees of fertility. Despite these obstacles, tomato breeders have transferred many genes, particularly for disease resistance, from wild *Lycopersicon* species to the tomato. *See* BREEDING (PLANT).

Lycopersicon is allied to *Solanum*, a diverse genus that includes potato, eggplant, and over 2000 other species. Two *Solanum* species, *S. pennellii* and *S. lycopersicoides*, have been crossed with the tomato.

The tomato, and all other species with which it can be crossed, has 12 pairs of chromosomes. There is good homology between chromosomes of each species. Tomato chromosomes differ in length, centromere position, and heterochromatin distribution, and each of the chromosomes of the tomato can be distinguished cytologically.

The tomato is a herbaceous perennial, but is usually grown as an annual in temperate regions since it is killed by frost. Originally it had an indeterminate plant habit, continuously producing three nodes between each inflorescence. Many modern varieties, however, are determinate; they have the self-pruning (*sp*) gene, and therefore have fewer than three nodes between inflorescences with the stem terminating in an inflorescence.

The compound leaves are usually alternate and odd-pinnate. The branches are procumbent or partly erect, and the weak stem is sometimes supported by a stake in cultivation.

The inflorescence is a monochasial cyme of 4 to 12 perfect and hypogynous flowers (**Fig. 1**). Primitive tomatoes have the solanaceous trait of five flower parts, but modern tomato varieties often have more than five yellow petals and green sepals. The five anthers are joined around the pistil in *Lycopersicon*, one of the key distinctions from the closely related *Solanum* genus. Wild *Lycopersicon* species, which are self-incompatible and therefore are obligate cross pollinators, have their style exserted beyond the anther cone. Cultivated tomatoes are self-fertile, and their style length is more similar to the anther length, a characteristic that favors self-pollination.

The fruit is a berry with 2 to 12 locules containing many seeds. Most tomato varieties have red fruit, due to the red carotenoid lycopene. Different single genes are known to produce various shades of yellow, orange, or green fruit. There is no basis for the common belief that yellow-fruited tomatoes are low in acidity. Many greenhouse varieties have pink fruit, due to a single gene (*y*) that prevents formation of the yellow pigment in the epidermis of the fruit.

Distribution and economic importance. The tomato is the most important processed vegetable, constituting over 23 lb (10.4 kg) of the 54 lb (24.3 kg) of processed vegetables the average American consumes each year.

Over three-quarters of the acres of processing tomatoes grown in the United States are in California. Other leading states, in order of acreage of processing tomatoes, are Ohio, Indiana, and New Jersey. Florida is the leading state for fresh market tomatoes, followed by California, South Carolina, Alabama, and Texas. Sizable amounts of fresh market tomatoes are imported to the United States from Mexico in winter months. Tomatoes are grown in every state, including Alaska. Although most tomato varieties prefer a long, warm season, very early varieties have been bred for northern areas.

Much of the world's production of tomatoes is in the United States. Russia has the largest acreage of tomatoes, but its production is second to the United States. Other leading countries, in order of acreage, are Italy, Egypt, Mexico, Turkey, and Spain. Tomatoes are often grown in greenhouses in northern Europe, especially in the Netherlands and England. The greenhouse tomato production in the United States is centered in Cleveland, Ohio.

Cultural practices. Tomatoes prefer warm weather. Cool temperature, 10°C (50°F) and below, delays seed germination, inhibits vegetative development, reduces fruit set, and impairs fruit ripening. High temperature, above 35°C (95°F), reduces fruit set and inhibits development of normal fruit color. Tomato plants cannot tolerate frost, and home gardeners often cover plants to protect them.

The tomato plant is day-neutral, flowering when grown with either short or long days. This makes it possible to grow tomatoes outdoors during the short days of winter in frost-free areas, such as southern Florida and the Imperial Valley of California, as well as in more northern areas during the long days of summer.

In California, seeds for most commercial tomatoes are planted directly in the field. Transplants are used in shorter-season areas, with plants grown for about 6 weeks in a local greenhouse or shipped from a southern state where they are grown in the field. Home gardeners often grow their own transplants in a warm, sunny location, then harden them in a cold frame before transplanting.

A starter solution, soluble fertilizer applied when transplanting, is often beneficial. Tomatoes prefer fertile soil, but can be grown on a variety of soils with proper fertilization. Heavy nitrogen fertilization does not make plants vegetative, as is often believed, but rather delays maturity and promotes rank vine growth. Nitrogen fertilizer and irrigation are often withheld late in the season when tomatoes are to be mechanically harvested, since this increases the proportion of ripe fruit in a single harvest. Phosphorus is often applied in a band near the seed when direct-seeding.

Level, uniform fields with few stones are best for mechanical harvesting. Heavy soils should be avoided when direct-seeding because of the danger of the soil crusting before the seedlings emerge. A soil pH of 6.0 to 6.5 is recommended.

Tomatoes respond to irrigation in many areas. A uniform supply of soil moisture is needed to prevent excessive losses from cracking and blossom-end rot. Good drainage is essential for high yields.

Fig. 1. Inflorescence of tomato (*Lycopersicon esculentum*) with open flowers and flower buds in different stages of development.

Better fruit set is often obtained with greenhouse tomatoes if the flowers are vibrated to ensure pollination. This is usually not necessary in the field, since the wind suffices.

Harvesting. Tomatoes for processing are harvested when red ripe and are soon sent to a nearby cannery. Tomatoes for fresh market, however, are often harvested at an earlier stage of maturity when they are still firm and better able to tolerate shipment to distant markets. Mature green fruits will ripen normally, but more immature fruits should not be harvested since their quality will be inferior.

Ethylene, a gas that is the natural ripening hormone in tomatoes and other fruit, is sometimes applied to stimulate green fruit to ripen after harvest. Ethephon (2-chloroethylphosphonic acid), a growth regulator that liberates ethylene, has been extensively used commercially to promote ripening of tomatoes.

Most tomatoes for fresh market are harvested by hand, but almost all of the processing tomatoes in California are harvested mechanically. Mechanical harvesting is less common in other areas, but is increasing. The harvester cuts the vine off at ground level, elevates it on an inclined chain, then shakes the fruit off and discards the vine. Workers on the machine remove most of the immature and defective fruit, and the rest are conveyed to a bulk container. Since only a single harvest is possible with mechanical harvesting and bruising is possible, varieties and cultural practices are used that favor (1) a maximum proportion of simultaneously ripe fruits, (2) the ability of fruits to remain in sound condition for a prolonged period after ripening, (3) tight attachment of fruits in order to prevent them from falling off when the plant is lifted by the machine, and (4) resilient

fruits that are not easily bruised. *See* AGRICULTURAL MACHINERY.

Culinary and biological uses. The tomato is highly esteemed as a source of vitamin C. Tomato varieties differ in ascorbic acid content, and many tomato breeders are striving to increase the content of vitamin C. Most tomato varieties have about 20–25 mg ascorbic acid per 100 g, and one medium-sized tomato provides about half of the required daily allowance of vitamin C for adults. *See* ASCORBIC ACID.

Tomatoes are also a significant source of vitamin A. Red tomatoes have about 1000 international units (IU) of vitamin A per 100 g. Some orange-fruited varieties which have the *B* gene, such as Caro Red and Caro Rich, have 10 times as much β-carotene (provitamin A) as red tomatoes, and a single fruit will provide more than the required daily allowance of vitamin A. *See* VITAMIN A.

Tomatoes are a good source of protein, but most of it is in the seeds. Tomato juice contains 19 amino acids, principally glutamic acid.

Most tomato varieties have 4.5 to 7.0% soluble solids, much of it as fructose or glucose. Citric acid is the predominant acid in tomato juice. The pH of tomatoes for processing should be below 4.5 to prevent spoilage, but the belief that modern hybrid varieties are lower in acidity than older varieties and more susceptible to botulism is unfounded. The ratio of sugar to acid is important for tomato flavor. The biochemistry and genetics of several of the aromatic organic compounds that contribute to tomato flavor have been determined.

The tomato has been a favored organism for genetic studies. Over a thousand genes are known for the tomato, and several hundred of these have been located on their respective chromosomes. Features of the tomato plant that make it very desirable for genetic and physiological studies include its high rate of natural self-pollination, the ease of controlled cross pollination, the large number of seeds per fruit and per plant, its ease of propagation by seed and cuttings and of culture in the field and the greenhouse, the relatively short time for each generation to reach maturity, the economic importance of the crop, and the wide assortment of available mutants, some of them with known effects on vitamin formation, carotenoid biosynthesis, hormones, mineral nutrition, water relations, and other physiological processes. *See* AGRICULTURAL SCIENCE (PLANT); FOOD MANUFACTURING; PLANT PHYSIOLOGY. R. W. Robinson

Germplasm. The ability of living organisms to pass on specific morphologic traits (form) and physiologic traits (function) to their offspring is due to the presence of germplasm, the self-reproducing deoxyribonucleic acid (DNA) found in the nuclei of their cells. There are large numbers of genes (discrete portions of the DNA) involved in such complex organisms, and after repeated cell divisions, significant numbers of minute changes occur in some of these genes. These changes may be expressed as variants, or mutants, of the parent plant. A small portion of these variants are better adapted than the parent to their environment; thus, they are "selected" in preference to the parent, or wild, type.

Fig. 2. Tomatoes, with genetic diversity, range from small-fruited wild types to large-fruited commercial types.

Like many other self-pollinated crop plants, the tomato is deficient in genetic diversity. As measured by molecular marker genes, the level of genetic variation prior to 1940 was vastly smaller in cultivated tomatoes than in related wild species. The same contrast was evident in genes for pest resistance, stress tolerance, fruit quality, and other economically desirable traits. Diversity among cultivated tomatoes was so depleted that little progress was made in breeding improved cultivars in the post-Mendelian period of 1900–1940. Recognition of this situation led tomato plant breeders to search for sources of increased diversity, such as spontaneous mutations (for example, determinate habit), induced mutations (for example, male sterility), and land races and related wild species (**Fig. 2**).

All wild species of *Lycopersicon* can be hybridized with *L. esculentum*, (with varying degrees of difficulty), permitting access to vast germplasm resources of *Lycopersicon*. Nearly all the major genes for disease resistance have been derived from land races and related wild species. Through experimental introgression, such enrichment of tomato germplasm has led to rapid improvement in tomato quality during the period from 1940 to the present, and has been partly responsible for up to a fivefold increase in yields per unit area compared with the preceding period. *See* AGRICULTURAL SCIENCE (PLANT); BREEDING (PLANT).

In the United States, tomato germplasm is maintained in two major public collections. The Tomato Genetics Resource Center at the University of California, Davis, has approximately 3400 accessions of wild species (9 species of *Lycopersicon* and 4 of related *Solanum*) and various genetic mutant and chromosomal variants. The U.S. Department of Agriculture/Agricultural Research Service (USDA/ARS) Plant Genetic Resources Unit in Geneva, New York, contains approximately 5000 items, mostly cultivars and other lines of *L. esculentum*, as well as representatives of the wild species. The main functions of each unit are to (1) acquire useful germplasm not already in the collection; (2) maintain adequate, viable seeds of each accession by procedures appropriate to

preserving their innate diversity; (3) evaluate each new accession in detail and repeat plantings of others for germination and phenotypic purity; (4) supply seed samples when requested by qualified investigators; (5) document all pertinent information in databases, which are also available on the World Wide Web. Both collections are backed up by samples in the National Seed Storage Laboratory at Fort Collins, Colorado, as insurance for long-term preservation.

Tomato breeders utilize these and other resources for developing improved cultivars. Progress to date has been accomplished largely through conventional breeding methods—hybridization (cross-mating of two genetically different plants) or backcrossing (first generation hybrid is crossed with parent generation) followed by pedigree selection of desired types. Recently, introgression of desired traits has been facilitated by the rapidly developing molecular technology of gene transposition (movement of DNA to new location in the chromosome); additional improvements are effected by transformation (movement and insertion of foreign DNA) and other manipulations. In the United States, F_1 (first-generation cross) hybrid cultivars constitute the majority of plantings for commercial production of fresh-market and processing tomatoes and for a large share of home gardens. Parent lines of such hybrids are cultivars or breeding lines derived as described above. *See* GENETIC ENGINEERING; SOMATIC CELL GENETICS.

This system of germplasm management and utilization facilitates the production of this nutritionally important crop for a steadily increasing worldwide market. Raymond L. Clark; Charles M. Rick

Diseases. Tomato plants are subject to well over 100 different diseases, most of which are caused by microorganisms (biotic agents), but some are caused by abiotic agents. The most frequent biotic agents are fungi, bacteria, viruses, and nematodes, while the most frequent abiotic agents are air pollutants, genetic disorders, and unbalanced mineral nutrition.

Some diseases of tomato occur wherever tomatoes are grown, whereas others may be restricted to certain regions. The absence of a disease in a region may be the result of the absence of the pathogen, resistance in the tomato cultivars, or unfavorable climate. A change in one or more of these factors may result in a sudden outbreak of disease. Some very common tomato diseases are discussed below.

Seedling damping-off. A wet rot of the roots, hypocotyl, or stem is characteristic of this disease. Death of seedlings usually occurs just before or just after seedling emergence from soils, or a few days after transplanting seedlings into soils. The three types of fungi that can cause this disease are *Rhizoctonia* spp., *Fusarium* spp., and *Pythium* spp. These fungi are all common inhabitants of most soils, but occur there in low populations. Prolonged wet weather or planting in soils that recently received green plant refuse increases the populations of the pathogens.

Root knot. Root knot is caused by a group of nematodes belonging to the genus *Meloidogyne*. The symptoms are conspicuous on roots because of abnormal growths called galls. The shoots of diseased plants appear pale green, are unvigorous, and sometimes wilt in hot weather. The nematodes occur naturally in soils in warm regions, and cause disease in many plants in addition to tomato. Some cultivars are resistant to root knot. *See* NEMATA (NEMATODA).

Wilts. Plants in the fruiting stage may wilt even though no obvious damage to roots has occurred and soil moisture is adequate for growth. Such wilting may be associated with a brown color in the stems. Three different microorganisms can cause this type of wilt, but each is different in its rapidity of wilt development. The fungus *Verticillium alboatrum*, most severe in cool weather, causes a slow-moving wilt with stunting, rather than killing, of the plant. The fungus *Fusarium oxysporum* f. sp. *lycopersici* causes a more rapidly moving wilt of the plant, which eventually dies. It is most severe in hot weather. The bacterium *Pseudomonas solanacearum* causes the most rapid wilt, and the plant usually dies soon after the first symptoms have been observed. All of these pathogens survive indefinitely in the soils. Resistant cultivars have been developed to all three wilt diseases.

Leaf spots and blights. Numerous necrotic spots and blights may develop on leaves and stems of tomato plants, some of which are very destructive and cause complete defoliation. The most damaging and common of these diseases are late blight and early blight. Late blight is caused by the fungus *Phytophthora infestans*. The first symptoms are greenish-black water-soaked spots. As the spots enlarge, a zone of white mold can be seen on the underside of leaves and at the margins of the spots when conditions are moist. Spread of the disease can be so rapid in wet cool weather that the entire plant will be killed. Early blight is caused by another fungus, *Alternaria solani*. As with late blight, spots may enlarge so that entire leaves are destroyed. The spots of early blight are different, however, in that conspicuous concentric rings occur in each lesion. A yellow color extends beyond the lesions as they age. Early blight is most severe on senescent leaves when plants bear enlarging fruit. High humidity and warm temperatures favor disease development.

Fruit rots. The two major rots of fruit are bacterial soft rot and anthracnose. Soft rot is caused by *Erwinia carotovora*. Breakdown of fruit, which can be very rapid at room temperatures, starts as a small water-soaked area, and continues until the fruit resembles a bag containing water. Insects with chewing larvae sometimes carry the bacterium from fruit to fruit in the field. As diseased fruit is harvested and washed, the bacterium is spread to healthy fruit. Many other fruits and vegetables are also susceptible to the soft-rot bacterium.

Anthracnose is a common disease of fruit that are allowed to grow close to the soil. The causal fungus, *Colletotrichum phomoides*, invades green fruit, but does not grow and cause rot until the fruit turn red.

In addition to the symptom of an expanding, sunken, circular rot, the signs of the fungus can be seen in the rotted portion, appearing as black areas, which may be arranged in concentric rings. In moist conditions, pink spore masses may appear on the surfaces of rotted areas. The fungus remains in soils on tomato debris from season to season, but since it only grows on tomatoes, rotation of tomatoes with other crops may be helpful to control the disease.

Mosaics. Several viruses cause mosaics of light green and dark green areas in leaves. These viruses also cause misshapen leaves, stunted growth, and sparse fruit set. Tobacco mosaic virus is probably the most common viral pathogen of tomato. The source of the virus is often tobacco products used by workers; their hands become contaminated with the virus and they transmit it from plant to plant during cultivation. Plants at all stages of growth are susceptible, though there are some resistant cultivars. Potato virus Y, another prevalent viral pathogen of tomato, produces symptoms very similar to those of tobacco mosaic. The virus differs, however, in that it is usually transmitted from plant to plant by aphids. Many solanaceous weeds are also susceptible to this virus, and they serve as reservoirs of the virus between crops. Two or more viruses may infect tomato plants at the same time; the multiple infections may cause more severe symptoms than those caused by either virus alone. *See* PLANT VIRUSES AND VIROIDS.

Abiotic diseases. Blossom-end rot of tomato fruit is caused by an imbalance in calcium in the fruit. Two to three weeks after fruit set, near the end of the fruit that is opposite the end attached to the plant, a water-soaked area may appear that rapidly enlarges and becomes sunken. Secondary fungi grow in the rotted area, which turns black. The calcium imbalances that cause this condition can themselves be caused by improper water content of soils or by the use of too much soluble fertilizer. Blotchy ripening of fruit, characterized by portions of the fruit turning yellow rather than red, is another abiotic problem of tomato fruit. The causes of this disorder are not clear, but certain cultivars are affected more than others. *See* PLANT PATHOLOGY. Robert E. Stall

Bibliography. J. G. Atherton and J. Rudich, *The Tomato Crop: A Scientific Basis for Improvement*, 1987; W. A. Gould, *Tomato Production, Processing, and Technology*, 1992; J. B. Jones et al. (eds.), *Compendium of Plant Diseases*, 1991; C. M. Rick, Tomato, in J. Smartt and N. W. Simmonds (eds.), *Evolution of Crop Plants*, Longman, London, 1995; C. M. Rick and R. T. Chetelat, Utilization of related wild species for tomato improvement, *Acta Hort.*, 412:21–38, 1995; C. M. Rick and J. I. Yoder, Classical and molecular genetics of tomato: Highlights and perspectives, *Annu. Rev. Genet.*, 22:281–300, 1988; S. D. Tanksley and S. R. McCouch, Seed banks and molecular maps: Unlocking genetic potential from the wild, *Science*, 277:1063–1066, 1997; U. S. Department of Agriculture, *Agricultural Statistics—1998*, 1998; J. Yoder, *Molecular Biology of Tomato: Fundamental Advances and Crop Improvement*, 1993.

Tommotian fauna

The first diverse assemblages of unquestionable animal fossils at the Proterozoic-Phanerozoic transition, which marks the change from a predominantly microbial biosphere to a modern type of biosphere abundant with multicellular life. The name derives from the Early Cambrian Tommotian Stage in Siberia, where the significance of this fauna of early skeletal fossils (often referred to as "small shelly fossils") was first realized. However, the concept goes beyond the geographical and temporal boundaries of the Tommotian Stage. It can be traced back to the low-diversity assemblages of skeletal animal fossils appearing near the end of the Neoproterozoic, continuing into the Cambrian, in Siberia first in moderate diversity in the Manykaian (Nemakit-Daldynian) Stage, then burgeoning in the Tommotian Stage itself. *See* CAMBRIAN; FOSSIL.

Traditionally, the fauna was considered to predate the earliest trilobites and was therefore often referred to as "pre-trilobitic"; however, recent stratigraphic studies suggest that at least some of the Tommotian Stage correlates with trilobite-carrying beds elsewhere. A number of the characteristic Tommotian taxa are now also known to continue into post-Tommotian strata. Thus, freed of its tight stratigraphic constraints, the Tommotian fauna is known to have radiated from all continents, but it is particularly diverse and abundant on the Siberian Platform in Russia, in Australia, and in the belt of phosphorite-rich deposits that extends from the South China Platform through Mongolia, Kazakhstan, the Himalayas, and Iran. *See* TRILOBITA.

The Tommotian fauna, being mainly restricted to animals with hard parts, does not yield as complete a picture of animal anatomy and diversity as the more well-known Cambrian Burgess Shale and Chengjiang (South China) faunas or the Upper Cambrian orsten fauna, all of which are characterized by frequent preservation of soft tissues. Nonetheless, it has been instrumental in forming our understanding of the Cambrian explosion (dramatic evolutionary radiation of animals beginning about 545 million years ago), because its preservation is generally not dependent on extraordinary conditions and is therefore less spotty. Also, fossilized embryos of Tommotian animals make it possible to understand the complete life cycles of some of the most basal members of the metazoan evolutionary tree. *See* ANIMAL EVOLUTION; BURGESS SHALE.

Faunal composition. Characteristic morphologic features of the Tommotian fauna include mineralized tubes, spicules, sclerites, and shells, often belonging to animals of unknown affinities. The minerals involved are opal (a hydrated gel of silica), apatite (calcium phosphate), and aragonite/calcite (calcium carbonate)—the same minerals that are common in animal skeletons today.

Tubes. The tubes represent many different kinds of animals that built protective sheaths reinforced with minerals. One of the earliest known such forms, the late Neoproterozoic *Cloudina*, had a multilayered

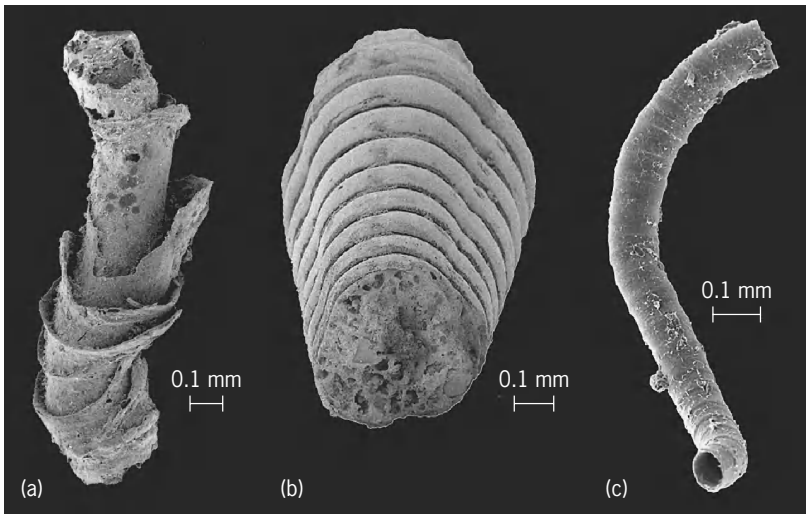

Fig. 1. Tube-dwelling animals of the Tommotian fauna. (*a*) *Cloudina*, one of the earliest animals with a mineralized skeleton. (*b*) *Aculeochrea*, an anabaritid showing the three-rayed symmetry typical of the group. (*c*) *Hyolithellus*, an animal reinforcing its tube with calcium phosphate.

cone-in-cone structure consisting of thin calcitic laminae (**Fig. 1***a*). Already in these early mineralized animal skeletons there is evidence of attacks from shell-boring predators, suggesting that exoskeletons initially arose as a response to predation. *Cloudina* is now known to be associated with several other forms of skeletal fossils. The triradially symmetrical tubes of anabaritids (Fig. 1*b*) partly overlap *Cloudina* in stratigraphic range and are a particularly characteristic component of the Proterozoic-Cambrian boundary beds. Anabaritids may be related to cnidarians, although triradial symmetry is not a characteristic feature of later cnidarians. Associated embryos show definitive cnidarian features, although the connection between these embryos and adult anabaritids is only implied from their joint occurrence. *See* CNIDARIA.

Other possible cnidarians are represented by forms such as *Olivooides*, which shows similarities with modern scyphozoans (jellyfish). A rich embryonic material of *Olivooides* shows that it developed directly within the egg, without going through a free planula stage (**Fig. 2**).

Other tube-dwelling forms are even more difficult to put in their proper phylogenetic place because of the lack of anatomically significant characters. These include the hyolithelminths (Fig. 1*c*) and other taxa that similarly reinforced their tubes with calcium phosphate, as well as forms that used calcium carbonate for the same purpose. There were also a number of animals that constructed their tubular sheaths of sclerotized organic matter or strengthened them with agglutinated sedimentary particles.

Spicules. Spicules (spikelike supporting structures) of various composition are widespread in the Animal Kingdom today. Although a number of mineralized spicules have been reported from the late Neoproterozoic, all of these reports are contested. Sponge spicules, however, are already considered a conspicuous component of the Tommotian fauna, and there are a number of spicules probably belonging to animals other than sponges.

The cup-shaped archaeocyathans (**Fig. 3**) long were of uncertain affinities and regarded by some as representing an extinct phylum. The presence of calcareous basal skeletons in several groups of recent sponges, however, has led to the now generally accepted idea that archaeocyathans are a group of

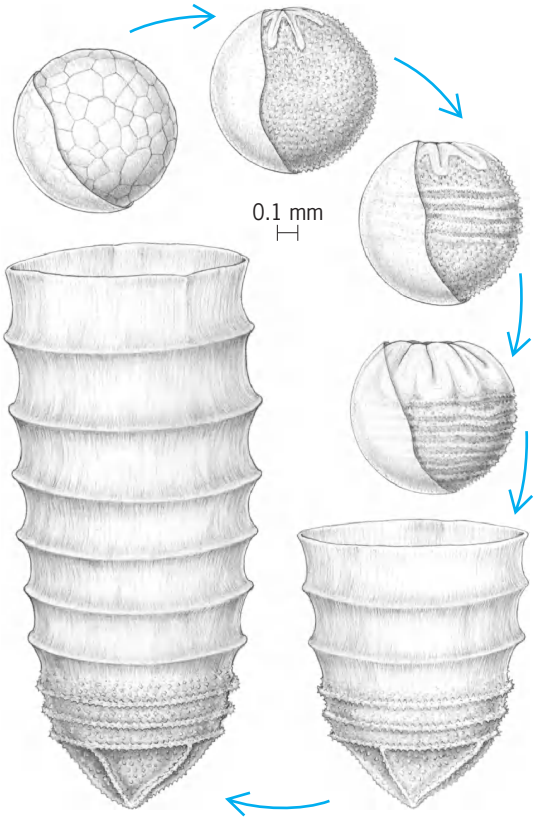

Fig. 2. Reconstruction of life cycle of the probably scyphozoan *Olivooides*, based on fossilized embryos and polyps.

Fig. 3. An archaeocyathan, interpreted to be a calcareous sponge.

calcified sponges that proliferated in the Early Cambrian and died out before the Ordovician. They are common in reefs but do not appear to have had the capacity of constructing reef frameworks. *See* ARCHAEOCYATHA.

Sclerites. A conspicuous element of Tommotian fossil assemblages is sclerites (hardened plates) belonging to composite exoskeletons, scleritomes. In many cases, the body shape of the bearer and the distribution of sclerites on the body are not known, but finds of complete scleritomes or even bodies in shale deposits give occasional and crucial insights. Thus the star-shaped composite sclerites of the chancelloriids are known to belong to a cactuslike animal that in its organization seems closest to sponges: sedentary, sac-shaped bodies with an apical orifice and no evidence of internal organs. These sclerites belong to a type called coelosclerites, consisting of a mineralized envelope around a space originally filled with soft tissue and showing no evidence of accretionary growth. The halkieriids are characterized by scale-shaped coelosclerites, and finds of complete specimens show the animal to have been slug-shaped, with two large anterior and posterior shell plates in addition to the sclerites (**Fig. 4***a*). Embryos of a segmented animal co-occurring with such sclerites have been proposed to belong to halkieriids. It is not clear whether coelosclerites are a convergent feature, independently evolved in several groups, or whether they were inherited from a common ancestor of the various groups of coeloscleritophorans.

Other sclerites were growing by stepwise accretion and often used calcium phosphate rather than calcium carbonate as shell mineral. The most widespread group of such phosphatic sclerite bearers are the tommotiids (Fig. 4*c*, *d*), of which so far no complete skeletons have been found. Tommotiids show a variety of sclerite shapes and ultrastructures and may in fact represent a polyphyletic assemblage of lineages that independently acquired a phosphatic scleritome. Other phosphatic sclerites include tooth- or hook-shaped objects as well as a variety of plate-like types, most of which are of unknown nature. The earliest known brachiopod is represented in the Tommotian fauna by a phosphatic shelled form.

Shells. Mollusk-like forms are a common component of the Tommotian fauna, although their phylogenetic placement continues to be contentious. A number of gastropodlike shells (Fig. 4*e*) occur, but it is not clear whether the soft parts did in fact possess the torsion characteristic of crown-group gastropods. Laterally flattened shells showing what appears to be an incipient dorsal hinge (Fig. 4*f*) have been interpreted as ancestral to the later rostroconchs and bivalves. A number of Tommotian taxa have been assigned to polyplacophorans, but all of these suggestions are controversial. Examples are the halkieriids (Fig. 4*a*) and siphogonuchitids (Fig. 4*b*), which have been interpreted alternatively as stem-group polyplacophorans, stem-group mollusks, and stem-group lophotrochozoans (a phylum grouping including mollusks, annelids, and brachiopods). *See* ANNELIDA; BIVALVIA; BRACHIOPODA; GASTROPODA; MOLLUSCA; POLYPLACOPHORA.

Evolutionary significance. The phylogenetic relationships of the Tommotian animals are very incompletely understood. Although some forms have with reasonable confidence been interpreted to belong to the crown group or stem group of a living phylum, others are more problematic. To some extent this is because many of them are incompletely known, perhaps only from disarticulated sclerites and no soft parts. More pertinent, however, may be the observation that the Tommotian fauna represents the first major adaptive radiation of animals and so is likely to include a number of short-lived lineages that do not share anatomical features to any greater extent than do living phyla.

The appearance in the fossil record of the Tommotian fauna near the base of the Cambrian has often been interpreted as being primarily a biomineralization event rather than a major radiation. This has then been attributed to extrinsic (for example, seawater chemistry) as well as intrinsic (for example, evolution of biomineralization pathways) factors. However, biomineralization not involving skeleton formation can be shown to be much older than the Cambrian. In addition, independent lines of evidence (trace fossils, soft-body preservation, organic microfossils) show the Cambrian explosion to have

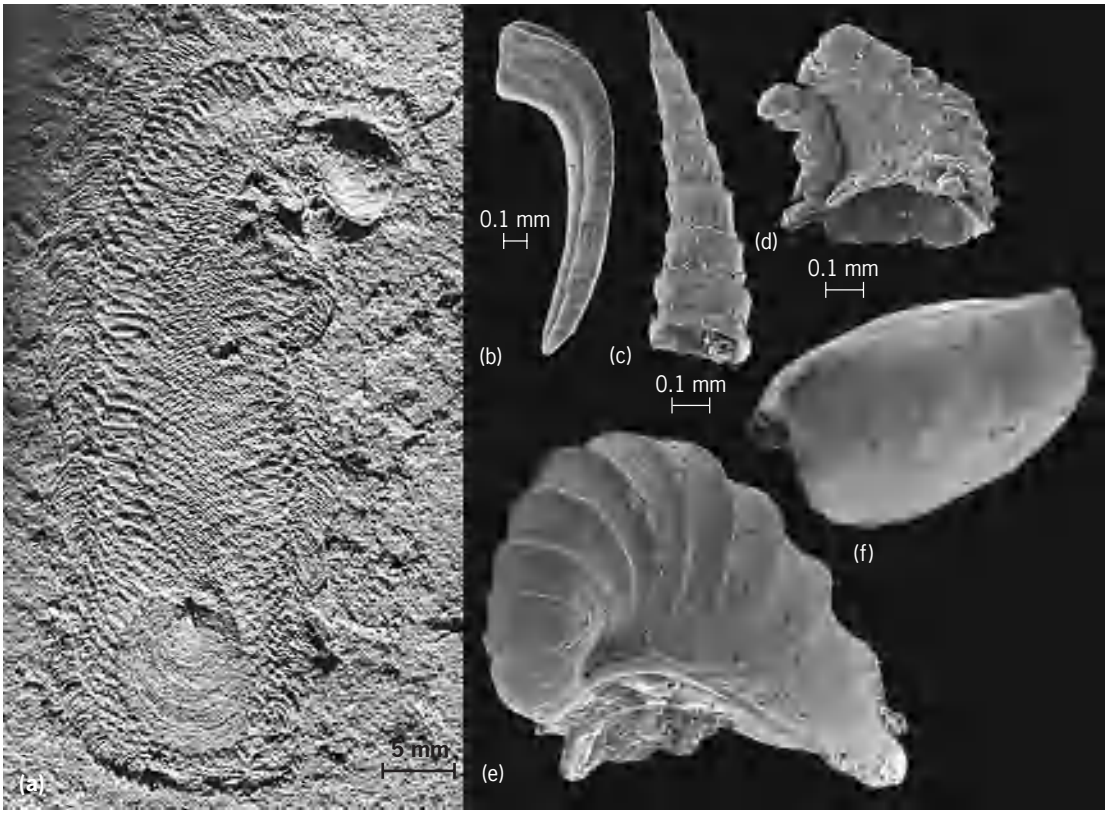

Fig. 4. Sclerite-bearing animals and probably mollusks. (*a*) *Halkieria*, a scale-and-plate-clad coeloscleritophoran (*reprinted with permission from M. S. Conway and J. S. Peel, Articulated halkieriids from the Lower Cambrian of north Greenland, Nature, 345:802–805, 1990, © Nature*). (*b*) *Siphogonuchites*, coeloscleritophoran, represented by a loose sclerite. (*c*) Sclerite of *Lapworthella*, a tommotiid. (*d*) Sclerite of *Eccentrotheca*, tommotiid. (*e*) *Archaeospira*, a possible early gastropod. (*f*) *Watsonella*, an early mollusk possibly representing an ancestral lineage to rostroconchs and bivalves (*reprinted with permission from A. P. Gubanov, A. V. Kouchinsky, and J. S. Peel, The first evolutionary-adaptive lineage within fossil molluscs, Lethaia, 32:155–157, 1999 © Taylor & Francis Group*).

involved nonskeletalized animals and nonanimals as well. Thus, the evolution of biomineralized skeletons assumes the same general significance as the evolution of various other kinds of animal tissues during this radiation event. Stefan Bengtson

Bibliography. S. Bengtson et al., Early Cambrian fossils from South Australia, *Mem. Ass. Australasian Palaeontol.*, 9:1–364, 1990; S. Bengtson and S. Conway Morris, Early radiation of biomineralizing phyla, in J. H. Lipps and P. W. Signor (eds.), *Origin and Early Evolution of the Metazoa*, Plenum, New York, 1992; E. Landing and K. E. Bartowski, Oldest shelly fossils from the Taconic allochthon and late Early Cambrian sea-levels in eastern Laurentia, *J. Paleontol.*, 70:741–761, 1996; V. V. Missarzhevskij, Drevnejshie skeletnye okamenelosti i stratigrafiya pogranichnykh tolshch dokembriya i kembriya [The oldest skeletal fossils and stratigraphy of the Precambrian-Cambrian boundary beds], *Trudy Geologicheskogo Instituta AN SSSR*, 443:1–237, 1989; Yi Qian and S. Bengtson, Palaeontology and biostratigraphy of the Early Cambrian Meishucunian Stage in Yunnan Province, South China, *Fossils and Strata*, 24:1–156, 1989; A. Yu. Rozanov, Problematica of the Early Cambrian, in A. Hoffman and M.H. Nitecki (eds.), *Problematic Fossil Taxa*, vol. 5, Oxford University Press, New York, 1986; A. Yu. Rozanov et al., Tommotskij yarus i problema nizhnej granitsy kembriya [The Tommotian Stage and the problem of the lower boundary of the Cambrian], *Trudy Geologicheskogo Instituta AN SSSR*, 206:1–380, 1969 [English transl. Amerind Publishing Co., New Delhi, 1981]; A. Zhuravlev and R. Riding (eds.), *Ecology of the Cambrian Explosion*, Columbia University Press, New York, 2000.

Ton of refrigeration

A rate of cooling that is equivalent to the removal of heat at 200 Btu/min (200 kilojoules/min), 12,000 Btu/h (13 megajoules/h), or 288,000 Btu/day (300 MJ/day). This unit of measure stems from the original use of ice for refrigeration. One pound of ice, in melting at 32°F (0°C), absorbs as latent heat approximately 144 Btu/lb (335 J/kg), and 1 ton (0.9 metric ton) of ice, in melting in 24 h, absorbs 288,000 Btu/day (300 MJ/day). In Europe, where the metric system is used, the equivalent cooling unit is the frigorie, which is a kilogram calorie, or 3.96 Btu. Thus 3000 frigories/h is approximately 1 ton of refrigeration. A standard ton of refrigeration is one developed at standard rating conditions of 5°F (−15°C) evaporator and 86°F (30°C) condenser temperatures, with 9°F (−13°C) liquid subcooling and 9°F (−13°C) suction superheat. *See* REFRIGERATION. Carl F. Kayan

Tone (music and acoustics)

Physically, a sound that is composed of discrete frequency (or sine-wave) components; psychologically, an auditory sensation that is characterized foremost by its pitch or pitches.

Physical meaning. The physical definition distinguishes a tone from a noise, wherein the components form a continuum of frequencies. Tones may be pure, consisting of a single frequency, or they may be complex. Complex tones, in turn, may be periodic or not periodic. Periodic complex tones repeat themselves at rapid regular intervals. They have frequency components that are harmonics—discrete frequencies that are integer multiples of a fundamental frequency. For example, the orchestra tunes to the tone of an oboe, a periodic tone consisting of a fundamental frequency of 440 hertz, a second harmonic component with a frequency of 880 Hz, a third harmonic at 1320 Hz, and so on. In general, musical instruments that generate continuous sounds—the bowed strings, the brasses, and the woodwinds—create such periodic tones. Tones that are not periodic (aperiodic) have frequency components that do not fit a harmonic series. Percussive instruments such as kettledrums and bells make such aperiodic tones. *See* HARMONIC (PERIODIC PHENOMENA); MUSICAL ACOUSTICS; MUSICAL INSTRUMENTS; PERIODIC MOTION.

Psychological meaning. Pitch is a sensation of highness or lowness that is the basic element of melody. Periodic complex tones tend to have a single pitch, which listeners will match by a pure tone having a frequency equal to the fundamental frequency of the periodic complex tone. Aperiodic complex tones tend to have multiple pitches. A second psychological attribute of complex tones is tone color or timbre. Tone color is often represented by descriptive adjectives. The adjectives may be linked to the physical spectrum. Thus, a tone with strong harmonics above 1000 Hz may be called "bright." A tone with no harmonics at all above 1000 Hz may be called "dull" or "stuffy." *See* PITCH; PSYCHOACOUSTICS; SOUND.

William M. Hartmann

Bibliography. D. Deutsch (ed.), *The Psychology of Music*, Academic Press, New York, 1998; W. M. Hartmann, *Signals, Sound, and Sensation*, Springer-Verlag, New York, 1997; B. C. J. Moore, *Introduction to the Psychology of Hearing*, Academic Press, London, 1997; T. D. Rossing, *The Science of Sound*, Addison-Wesley, Reading, MA, 1990.

Tongue

An organ located at the base of the oral cavity and found in all vertebrate animals. It is best developed in terrestrial vertebrates, where it takes on the functions of food procurement, food transport, and acquisition of chemosensory signals. The tongue generally is not a significant independent organ in fish, and it is secondarily reduced in organisms that feed aquatically, such as crocodilians and some turtles.

Within terrestrial vertebrates, there is considerable variability in the specific structure of the tongue, the degree of participation of the hyoid skeleton (that is, a complex of bones at the base of the tongue which supports the tongue and its muscles), and the mechanisms of movement. In birds the tongue is merely a thickened epithelium that overlies the hyoid apparatus. Movement is produced by moving various hyoid elements. In most amphibians, including both frogs and salamanders, the hyoid provides extensive support, but considerable intrinsic tongue musculature exists. In squamate reptiles (lizards and snakes) and mammals the tongue is largely independent of the hyoid apparatus and is composed entirely of muscle. The musculature is tightly packed in the tongue and is generally arranged in three mutually perpendicular planes. In the mammalian tongue the musculature is arranged into longitudinal, transverse, and vertical bundles. Organs composed entirely of muscle and lacking independent skeletal systems are widespread, and include elephant trunks and the tentacles and arms of cephalopod mollusks. Such organs have been termed muscular hydrostats.

In muscular hydrostats, muscle acts as both the effector of movement and as the support for that movement. The most important biomechanical feature of a muscular hydrostat is that its volume is constant, so that any decrease in one dimension will cause a compensatory increase in at least one other dimension. Tongue protrusion is primarily produced by elongation of the tongue, as well as movement of the tongue in space. Elongation occurs when muscle that decreases cross-sectional area contracts, a phenomenon found in the circular muscles of snakes and most lizards, or in transverse muscles in most mammals. The geometry of a muscular hydrostat may provide for amplification of muscle effect, or leverage, so that in muscular hydrostats that are protruded over long distances, the resting length of the organ is long in relation to its width. Bending requires simultaneous contraction of longitudinal muscle with another muscle, usually circular, transverse, or radial. In bending, one muscle acts to elongate the organ, while another resists this elongation at a specific spot, producing a bend. One of the primary advantages of a muscular hydrostat is that bending is not restricted to movement at joints, and the highly subdivided muscular and neural systems seen in mammalian tongues in particular produce movements that are remarkably specific, complex, and diverse.

While muscular-hydrostatic movements characterize the tongue of most mammals and many lizards and snakes, many of the most spectacular tongue projectors, such as chameleon lizards and plethodontid salamanders, do not use this mechanism in protrusion. In these animals the tongue is used as a prey capture mechanism and may be protruded with force for very long distances. These organisms have developed separate mechanisms in which the muscular tongue is projected ballistically from the body. In both, a muscle squeezes a process of the hyoid apparatus to generate the projectile force. In the

chameleon the tongue is projected off a hyoid base that remains in the mouth; in plethodontid salamanders the entire hyoid apparatus, which carries the tongue, is projected. *See* TASTE. Kathleen K. Smith

Tonsil

Localized aggregation of diffuse and nodular lymphoid tissue found in the region where the nasal and oral cavities open into the pharynx. The lymphoid tissue consists of small, closely packed round cells called lymphocytes supported in a specialized connective tissue framework called reticular tissue. When lymphocyte production is active, rounded and more densely packed clusters or nodules of these cells appear in the diffuse lymphoid tissue. The most active nodules possess lighter staining centers composed of somewhat larger, less densely packed lymphocytes showing evidence of cell division. Such areas are called germinal centers. In the tonsillar regions, the lymphoid tissue lies just beneath the lining epithelium. The tonsils are important sources of blood lymphocytes. They often become inflamed and enlarged, necessitating surgical removal.

Palatine tonsil. The two palatine (faucial) tonsils are almond-shaped bodies measuring about 1 × 0.5 in. (2 × 1 cm) and are embedded between folds of tissue connecting the pharynx and posterior part of the tongue with the soft palate (**Fig. 1**). These are the structures commonly known as the tonsils. The openings of 10 to 20 pits (crypts) which extend deep into the organ may be seen on the surface. The stratified squamous epithelium of this region covers the surface of the tonsil and lines the crypts (**Fig. 2**). A fibrous capsule separates the tonsil from the underlying muscle. Extensions from the capsule form supporting septa within the tonsil. Lymphoid tissue occupies all interstices between the capsule, septa, and epithelium. Crypts frequently become filled with detached epithelial cells, living and dead lymphocytes, and exuding fluids. Such sequestered masses form an excellent culture medium for the growth of certain bacteria and fungi. The protective quality of the crypt epithelium may be weakened by the passage of large numbers of lymphocytes through it.

Lingual tonsil. The lingual tonsil occupies the posterior part of the tongue surface. It is really a collection of 35–100 separate tonsillar units, each having a single crypt surrounded by lymphoid tissue. Each tonsil forms a smooth swelling about 0.08–0.1 in. (2–4 mm) in diameter. Since gland ducts open into these crypts, the contents are flushed out, and lingual tonsils rarely cause trouble. The epithelium is again stratified squamous, and a thin capsule is present around each unit.

Pharyngeal tonsil. The pharyngeal tonsil (called adenoids when enlarged) occupies the roof of the nasal part of the pharynx and is covered with pseudostratified ciliated columnar epithelium. The organ consists of a series of radiating folds leading forward from the region where the roof of the nasal pharynx

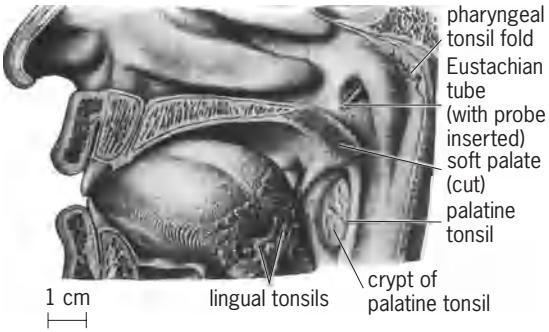

Fig. 1. Dissection showing the tonsil in relation to other structures in the pharyngeal region.

joins the posterior pharyngeal wall (**Fig. 3**). When cut at right angles to the folds, the intervening spaces resemble crypts. Septa are found in the folds, but a distinct capsule is lacking. Lymphoid tissue and lymphocytic invasion of the epithelium are similar to those of the other tonsils (**Fig. 4**). This tonsil may enlarge to block the nasal passage, forcing mouth breathing.

Development. Lymphocytic infiltration for the palatine tonsil begins at the site of the disappearing second pharyngeal pouch during the third fetal month. The pharyngeal and lingual tonsils appear during the fourth and fifth fetal months, respectively. Tonsils reach their maximum size during childhood and subsequently regress.

Function. All tonsils produce lymphocytes which are added to the circulating blood via the plexus

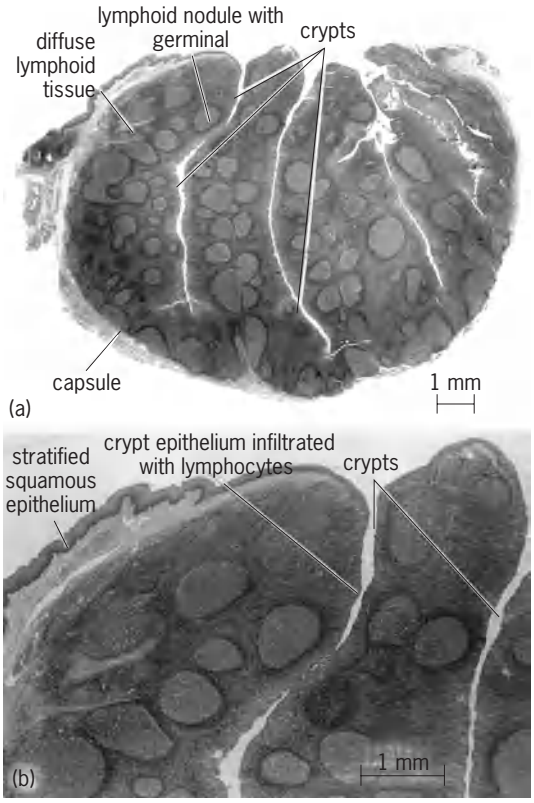

Fig. 2. Photomicrographs of a child's tonsil. (*a*) Vertical section. (*b*) Higher magnification of *a*.

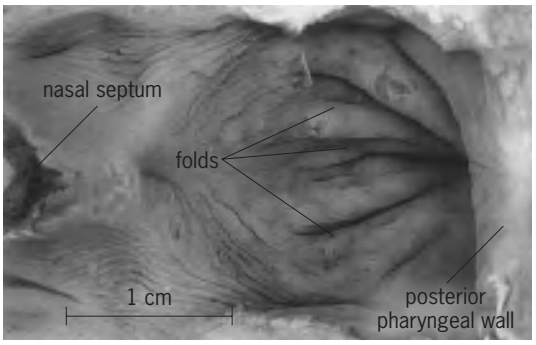

Fig. 3. Photograph of the roof of the nasal pharynx of a newborn baby showing the pharyngeal tonsillar folds.

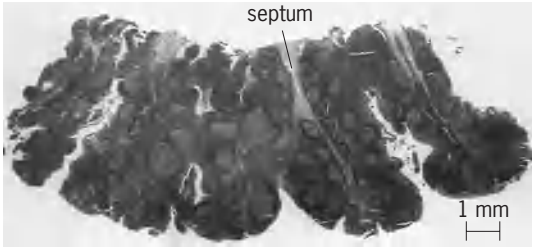

Fig. 4. Photomicrograph showing a section across the pharyngeal tonsil folds of a child.

of lymph capillaries which surrounds the lymphoid tissue. The flow of lymph is always away from the tonsillar sites. No other function has been firmly established. The three sets of tonsils along with lesser amounts of intervening lymphoid tissue form a complete ring around the upper reaches of the digestive and respiratory systems. Because of this strategic location, a protective function has been suggested. It is thought by some that protection, in response to entering bacteria, may be afforded through the production of antibodies. Evidence has shown that plasma cells (many of which are present in the tonsils), lymphocytes, or both are implicated in this process. *See* HEMATOPOIESIS; LYMPHATIC SYSTEM. Theodore Snook

Tonsillitis

An inflammation of the tonsil. Tonsillitis is a nonspecific term usually referring to bacterial or viral infection involving all or part of Waldeyer's ring, a collection of lymphatic tissue encircling the pharynx. It consists primarily of the tonsils (palatine tonsils), adenoids (pharyngeal tonsils), and lingual tonsils. The palatine tonsils are readily seen at the edges of the soft palate between the anterior tonsillar pillar (glossopalatinus muscle) and the posterior tonsillar pillar (pharyngopalatinus muscle). The pharyngeal tonsils (adenoids) cover the vault and posterior wall of the nasopharynx and cannot be seen unless the nasopharynx is examined with a mirror or the soft palate is retracted. The lingual tonsils are located at the base of the tongue.

The purpose of Waldeyer's ring, situated at the entrance to the respiratory tract and the alimentary tract, is presumed to be related to immunization. For example, it has been shown in the 7-day-old calf and piglet that there were only small numbers of bacteria in the lymphatic drainage of the head, intestines, and lungs but that there were large numbers in the tonsils. With increased age to adulthood, the number of bacteria decreased. This suggests that the tonsils may aid in the development of immunity to the bacteria of the animal's environment.

Virus infections. The viral flora of enlarged tonsils and adenoids have been studied. Fifty-eight percent of adenoids and 26% of enlarged tonsils contained adenovirus types I, II, or V. These latent viruses were most common in the first decade, and their incidence decreased sharply at puberty. The presence of these adenoviruses may be related to tonsillar hyperplasia, wherein there is an increase in the number of secondary lymph follicles and an increase in the size of the individual follicles.

Bacterial infections. Bacterial tonsillitis is caused by beta-hemolytic *Streptococcus* and possibly by *Staphylococcus aureus*. It involves the entire Waldeyer's ring, as well as the pharyngeal mucous membrane, and more correctly should be termed pharyngotonsillitis. Attacks of bacterial tonsillitis occur with varying frequency and severity in different individuals. They are most frequent in children younger than 6 years. Unfortunately, the appearance of tonsils between attacks does not correlate well with the number or severity of infections and gives no indication as to whether past attacks have been bacterial or viral. There is likewise little correlation between clinical history and microscopic findings. *See* STAPHYLOCOCCUS; STREPTOCOCCUS.

Complications. The complications of tonsillitis depend on which tonsil is involved. Recurrent adenoiditis with adenoid hypertrophy is frequently associated with recurrent otitis media, middle-ear fluid, and at times nasal obstruction with mouth breathing and snoring. Acute palatine tonsillitis may be complicated by peritonsillar abscess which may develop lateral to the tonsillar capsule.

Treatment. Streptococcal pharyngotonsillitis is treated with penicillin for at least 10 days in persons not allergic to the drug. Removal of the adenoids is considered when there is residual middle-ear fluid. Palatine tonsils must be removed after peritonsillar abscess, but otherwise their removal depends upon the frequency of recurrent attacks of bacterial pharyngotonsillitis in relation to the patient's age. *See* TONSIL. James A. Donaldson

Bibliography. D. D. DeWeese and W. H. Saunders, *Textbook of Otolaryngology*, 7th ed., 1987; P. M. Stell (ed.), *Scott Brown's Otolaryngology: Laryngology*, vol. 5, 5th ed., 1988.

Tooth

Any one of the structures found in the mouth of most vertebrates which, in their most primitive form, were conical and were usually used for seizing, cutting up, or chewing food, or for all three of these

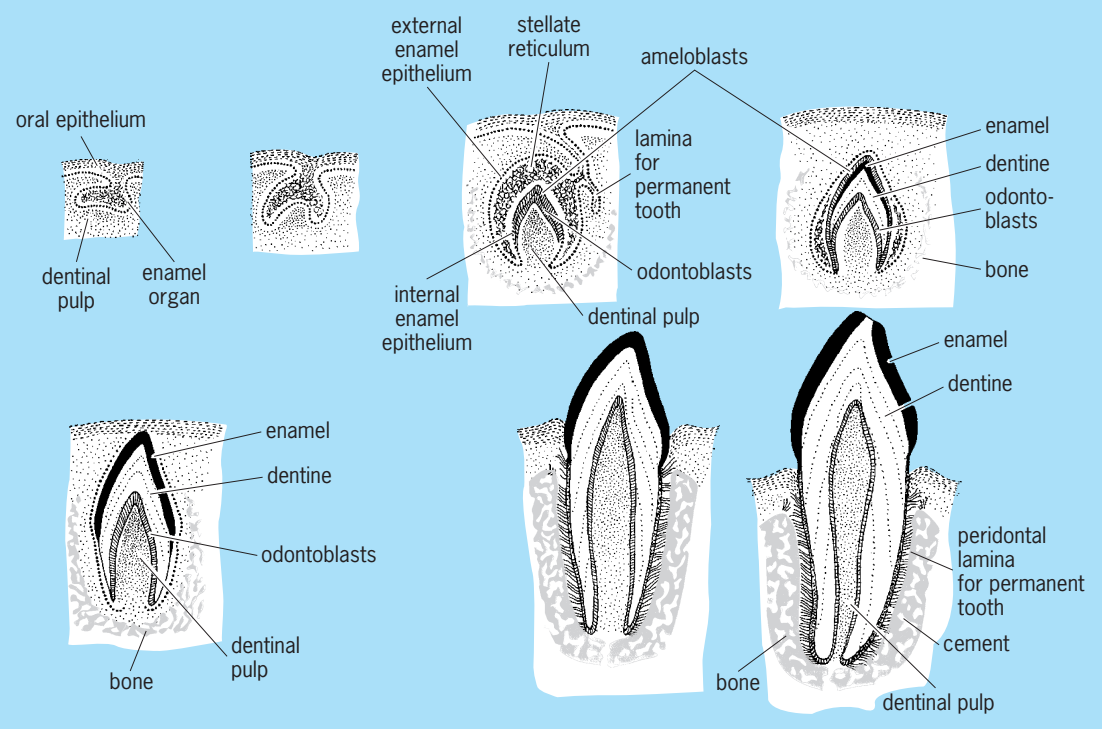

Fig. 1. Stages in the development of a tooth.

purposes. Although the true enamel of more advanced vertebrates is ectodermal, the remaining components of the teeth are mesodermal in their embryological origins (**Fig. 1**). The basic tissues that make up the vertebrate tooth are enamel, dentin, cementum, and pulp. Tooth replacement in vertebrates other than mammals may be explained by a rhythmical wave of impulses inducing the tooth germs, proceeding from the front to the back of the jaw. In their evolutionary origins teeth are derivatives of bony tubercles which developed on the outer part of the body of primitive agnathans, fishlike stem vertebrates, forming a protective shell around them. An easy confirmation of how the tubercles are thought to have developed into teeth can be seen in the identical embryonic development of the scales and teeth of young sharks.

Each tooth develops as a result of interaction between the thickening ectoderm (oral epithelium), which forms an enamel organ, and a mesodermal dental papilla. The ameloblast cells of the enamel organ secrete enamel, and the odontoblast cells of the dental papilla secrete the dentin. As the cells of the tooth bud proliferate rapidly, the tooth extends into the mesoderm, the outer layer becoming the external enamel epithelium, the inner one forming the internal enamel epithelium. The external and internal layers of enamel epithelium are separated by a mass of cells called the stellate reticulum. As the advanced tooth bud invaginates to form a cup-shaped structure around a mesodermal center, the inner epithelium cells form the ameloblast, which in turn first induces the formation of odontoblasts from the underlying mesodermal cells. As the odontoblasts begin to lay down dentin and the ameloblasts begin to secrete enamel, the two tissues are laid down in direct contact with each other.

Tissue components. Enamel is the hardest tissue in the body because of the very high concentration, about 96%, of mineral salts. The remaining 4% is water and organic matter. The enamel has no nerve supply, although it is nourished to a very slight degree from the dentin it surrounds. The fine, microscopic hexagonal rods (prisms) of apatite which make up the enamel form patterns that minimize breakage during mastication. The tiny rods are held together by a cementing substance.

Dentin, a very bonelike tissue, makes up the bulk of a tooth, consisting of 70% of such inorganic material as calcium and phosphorus, and 30% of water and organic matter, principally collagen. The rich nerve supply makes dentin a highly sensitive tissue. Although this sensitivity serves no obvious physiological function, it may be partly explained by the fact that in jawless, armored stem vertebrates the bony tubercles were not covered by enamel, and thus could function as a skinlike sensory system. As the tubercles became functional teeth, their sensory function was lost but the sensitivity remained.

Cement is a calcified tissue, a type of modified bone less hard than dentin, which fastens the roots of teeth to the alveolus, the bony socket into which the tooth is implanted. In mammalian herbivores, such as the rhinoceros, horse, and rabbit, which have evolved tall teeth, cementum is deposited between the enamel folds to strengthen the delicately folded enamel, as well as to form part of the occluding surface.

A miscellaneous tissue, consisting of nerves, fibrous tissue, lymph, and blood vessels, known as the pulp, occupies the cavity of the tooth surrounded by dentin. Lining the wall of the pulp cavity are odontoblasts which originally produced dentin tubules. If stimulated by decay or other external contact to the overlying dentin, these cells may produce secondary dentin. This is not as well organized in its structure as the first dentin, and it resembles bone.

Replacement. In most groups of vertebrates other than mammals, the individual teeth of the marginal dentition on the edge of the jaws are replaced throughout the life of the animals. Deep within the tissues of the tooth germs, teeth are constantly formed, grow, erupt, function, and are shed. Experiments on the tooth replacement of reptiles and amphibians resulted in an interesting hypothesis which seems to explain the previously unintelligible system of rhythmic replacement of teeth in nonmammalian vertebrates. Some contend that the spacing of waves of stimulation (which are genetically determined) in jaws, proceeding from front to back, results in tooth rudiments (or anlages) which begin development at rhythmically spaced time intervals. Spacing the beginning of development of tooth germs, assuming that the rate of development of teeth is even, results in the alternating appearance of erupting teeth. In **Fig. 2***a*, by the time the tooth-initiating impulse has traveled from position 1 to the position past 3, another anlage has been initiated at position 1. This means that a new tooth is started on its developmental path at a frequency slightly greater than two intervals between the induced tooth germs.

If the spacing of the tooth germs was exactly two intervals, the odd and even teeth would alternate exactly in the lower vertebrate jaw. The apparent irregular arrangement of the young, mature, and old teeth, both odds and evens, is explained by the assumption of the hypothesis that the spacing between the impulses is slightly greater than two intervals.

Evolution and phylogeny. The jawless fossil ostracoderms had no teeth as such, and presumably they were filter feeders. Teeth, along with jaws, first made their appearance among the placoderms. Although many sharks have simple conelike teeth, they usually have accessory cusps on the teeth, making their edges sharp and serrate. The teeth of various sharks and rays have been modified by being flattened into plates to crush mollusks, crustaceans, and other hard-shelled sea animals.

Both living and fossil lungfishes developed fan-shaped plates studded with ridges, or blunt, conical teeth. In the actinopterygian branch of bony fishes, relatively simple, conical teeth predominate, although a great variety of teleosts develop peculiar teeth of their own. Many teleosts have curved teeth which bend backward on a fibrous hinge, allowing the prey to enter the mouth but preventing its escape.

The probable direct ancestors of tetrapods, the crossopterygian bony fishes, as well as the earliest tetrapods, had conical teeth with groovelike folds of enamel cutting radially into the tooth all around.

In most of the reptiles simple conical teeth are predominant, although tooth plates were developed in the shellfish-eating placodonts and in several groups of herbivorous dinosaurs. Some lizards also evolved rather elaborate dentitions. Such reptiles are the ornithischian hadrosaur, the triceratopsian dinosaurs, and the recent iguanids. The mammallike reptiles evolved a great variety of complicated cheek teeth and a dentition which was differentiated into precanine and postcanine parts.

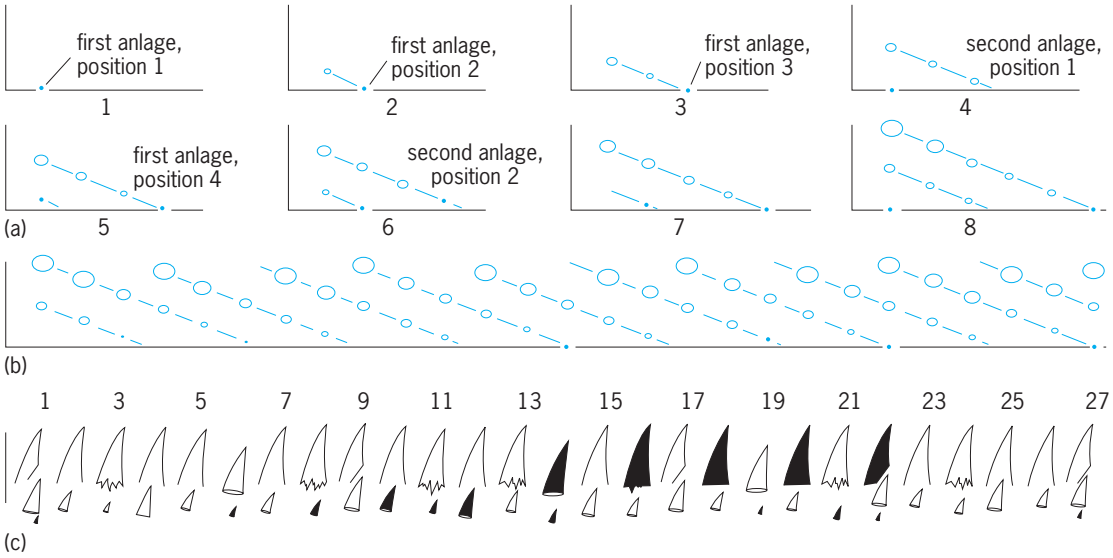

Fig. 2. Schema showing tooth replacement in lower vertebrates. (*a*, *b*) Horizontal lines represent the strand of tissue which connects the tooth germs in most vertebrates; vertical lines indicate the distance from the tooth germs to the surface of the jaw. In positions 1 to 8 of *a*, the dots on the horizontal lines represent the formation of tooth rudiments (anlages). As a result of this type of development, mature developing teeth in the jaw will be arranged as seen in *b*, and (*c*) they may appear from the side. In *c*, the white teeth are odd; the black are even.

Taxonomic studies. Teeth play a very important role in the study of mammalian evolution and the classification of living and fossil species. Because enamel is the hardest substance in the body, it is also most often preserved in rocks. Mammalian teeth often reflect adaptations to a particular kind of food and are therefore usually the most characteristic hard part of mammals. The crown of a tooth is genetically predetermined, and it is (at least in low-crowned teeth) fully formed before the teeth erupt. The only changes of the mammalian tooth crown, therefore, result from wear, breakage, or decay. Because it is very important that the complicated crown surfaces of the lower and upper molars occlude correctly, stringent genetic requirements determine occlusion.

Patterns of dentition. The dentition of therian mammals, at least primitively, consists of four different kinds of teeth. The incisors (I) are usually used for nipping and grasping; the canines (C) serve for stabbing or piercing; the premolars (Pm) grasp, slice, or function as additional molars; and the molars (M) do the chewing, cutting, and grinding of the food. Primitively the placentals have 40 teeth and the marsupials 50; the formulas below indicate their distribution in each side of the skull and jaw:

Primitive placental formula

$$\mathrm{I}\frac{1, 2, 3}{1, 2, 3} \qquad \mathrm{C}\frac{1}{1} \qquad \mathrm{Pm}\frac{1, 2, 3, 4}{1, 2, 3, 4} \qquad \mathrm{M}\frac{1, 2, 3}{1, 2, 3}$$

Primitive marsupial formula

$$\mathrm{I}\frac{1, 2, 3, 4, 5}{1, 2, 3, 4, 5} \qquad \mathrm{C}\frac{1}{1} \qquad \mathrm{Pm}\frac{1, 2, 3}{1, 2, 3} \qquad \mathrm{M}\frac{1, 2, 3, 4}{1, 2, 3, 4}$$

The numbers in the formulas depict the teeth in the jaw by their specific position in the series. For example, Pm_2 refers to a second lower premolar, whereas M^2 stands for a second upper molar. *See* MAMMALIA.

In therian mammals, probably because of the intricacies and vital importance of tooth occlusion, only part of the first (or "milk") dentition is replaced. This second, or permanent, dentition is made up of incisors, canines, and premolars; as a rule only one premolar is replaced in marsupials. Although the molars erupt late in development and are permanent, that is, not replaced, they are part of the first, or deciduous, dentition. *See* DENTITION. Frederick S. Szalay

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1987; R. Renner, *An Introduction to Dental Anatomy*, 1985; A. S. Romer, *Vertebrate Paleontology*, 3d ed., 1966; A. S. Romer and T. S. Parsons, *The Vertebrate Body*, 6th ed., 1986; J. Z. Young, *The Life of Vertebrates*, 1991.

Tooth disorders

Diseases and disturbances of the teeth and associated structures, including abnormal formation and growth of the teeth and jaws, tooth decay, inflammation of the tissues housing the roots of the teeth, and various diseases of the jaw bones. Tooth disorders are important because of their integral relationship to general body health, the essential role of the teeth in mastication and speech, and the contribution of normal dental function to the individual's psychological sense of well-being.

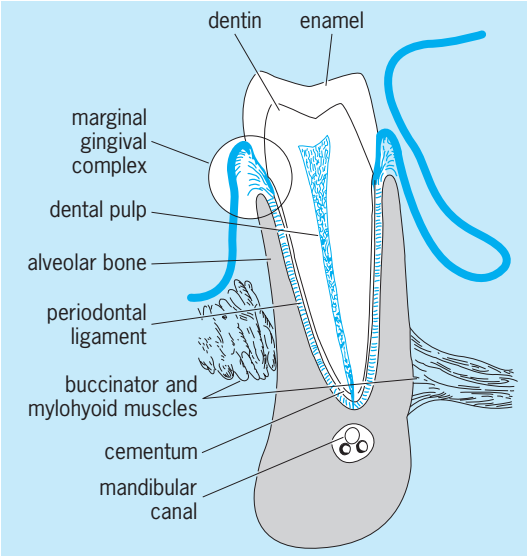

Structural features of the normal periodontal tissues.

Dentin, a calcified tissue containing viable cell processes, is the principal constituent of teeth. The dentin of the crown is covered by enamel, a nonviable tissue, and that of the root by cementum, a tissue closely resembling bone. The dental pulp contains nerves, blood vessels, and lymphatics which provide sensation and nutrition to the dentin. The root of the tooth is united to its body housing by the collagen fibers of the periodontal ligament (see **illus.**).

Abnormal formation. Defective formation of dentin and enamel are referred to respectively as dentinogenesis and amelogenesis imperfecta, and may be caused by febrile illness during the period of tooth formation, in hypophosphatasia, or by faulty calcium or phosphorus metabolism, such as occurs in rickets. In other cases, the causes remain unknown. Some or all of the teeth may fail to form completely, or extra or supernumerary teeth may be present. Abnormalities in the growth of the bones of the jaws and face with resulting faulty tooth position and malocclusion affect about 70% of all humans; in about 5%, these abnormalities are sufficiently severe to be disfiguring and handicapping.

Decay. In most societies, dental decay, or caries, is one of the most important and common tooth disorders. Decay rates are highest in individuals in the age range of 11–18 years who belong to low socioeconomic groups. Cavities arising in these individuals occur mostly in the developmental pits and fissures of the tooth crown and on the smooth surfaces between the teeth near the gum margin. Decay of the root surfaces becomes prevalent in individuals in the age range of 55–65 years. Decay occuring during the adolescent years is caused by a class of

microorganisms referred to as the cariogenic streptococci, of which *Streptococcus mutans* is the predominant member. However, a susceptible host and a cariogenic diet containing sucrose are also essential factors. Although the mechanism by which bacteria cause decay is not completely understood, most experts believe that cariogenic organisms, by using sucrose, and to a lesser extent other sugars, produce polymers which bind the organisms to the tooth surface and acids which cause demineralization resulting in cavity formation. Once the carious lesion penetrates the enamel and enters the dentin, the viable cell processes are affected and the tooth may become painful. At this stage, the organisms have relatively ready access to the pulp, where they may cause the formation of microabscesses. The associated inflammation and swelling cause severe pain and toothache. The pulp may become necrotic, causing formation of a periapical abscess, granuloma, or cyst in the alveolar bone at the root tip where the nerves and vessels enter the pulp. At this stage of disease, the tooth can be treated successfully by resolving the infection, removing the necrotic pulp, and obliterating the pulp chamber with inert filling material.

Fluoride administration, either in the drinking water at a concentration of 1 part per million (ppm), or by other routes at comparable levels, effectively reduces dental decay about 50–70%. This level of fluoride consumption appears to be without toxic or undesirable side effects, although concentrations in excess of 3–5 ppm may lead to browning or mottling of the enamel. The mechanism by which fluoride causes decreased decay rates is not understood, although some experts believe that the ion combines with hydroxyapatite, leading to the formation of more perfect crystals which are more resistant to acid attack. Fluoride may also enhance remineralization of beginning lesions.

Periodontal disease. Whereas dental decay is the principal cause of tooth loss in young individuals, inflammatory disease of the tissues surrounding the teeth, referred to as periodontal diseases, causes most of the tooth loss in adults. *See* PERIODONTAL DISEASE.

Jaw diseases. Diseases of the jaws may affect the teeth. The most common jaw disorders, other than abnormalities of growth and development discussed above, fall into four categories: (1) inflammation of the jawbone caused by infections such as osteomyelitis; (2) cysts associated with the teeth as well as those located in bone sutures of the jaws; (3) benign and malignant tumors of the jaws including osteoma, fibroma, sarcoma, tumors derived from tooth tissues, multiple myeloma, leukemia, and carcinoma; and (4) systemic diseases such as generalized skeletal abnormalities produced by endocrine dysfunction in which the jaws are affected. *See* DENTISTRY; TOOTH.

Roy C. Page

Bibliography. E. Newbrun, *Cariology*, 3d ed., 1989; E. Newbrun (ed.), *Fluorides and Dental Caries*, 3d ed., 1986; S. Schluger et al., *Periodontal Disease*, 2d ed., 1989.

Topaz

A mineral best known for its use as a gemstone. Crystals are usually colorless but may be red, yellow, green, blue, or brown. The wine-yellow variety is the one usually cut and most highly prized as a gem. Corundum of similar color sometimes goes under the name of Oriental topaz. Citrine, a yellow variety of quartz, is the most common substitute and may be sold as quartz topaz.

Topaz crystal habits. (*After C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., Wiley, 1993*)

Topaz is a nesosilicate with chemical composition $Al_2SiO_4(F,OH)_2$. The mineral crystallizes in the orthorhombic system and is commonly found in well-developed prismatic crystals with pyramidal terminations (see **illus.**). It has a perfect basal cleavage which enables it to be distinguished from minerals otherwise similar in appearance. Hardness is 8 on Mohs scale; specific gravity is 3.4–3.6.

Topaz is found in pegmatite dikes, particularly those carrying tin. It is also formed during the late stages of the solidification of rhyolite lavas. The minerals characteristically associated are tourmaline, cassiterite, fluorite, beryl, and apatite. It is also found as rolled pebbles in stream gravels. Fine yellow and blue crystals have come from Siberia and much of the wine-yellow gem material from Minas Gerais, Brazil. In the United States topaz has been found near Florissant, Colorado; in Thomas Range, Utah; in San Diego County, California; and near Topsham, Maine. *See* GEM; SILICATE MINERALS.

Cornelius S. Hurlbut, Jr.

Topographic surveying and mapping

The measurement of surface features and configuration of an area or a region, and the graphic expression of those features. Surveying is the art and science of measurement of points on, above, or under the surface of the Earth. Topographic maps show the natural and cultural features of a piece of land. The natural features include configuration (relief), hydrography, and vegetation. The cultural features include roads, buildings, bridges, political boundaries, and the sectional breakdown of the land. Topographic maps are used by a wide variety of people, such as engineers designing a new road; backpackers finding their way into remote areas; scientists describing soil or vegetation types, wildlife habitat, or hydrology; and military personnel planning field operations.

Presentation of information. Topographic map features can be expressed in a variety of ways. Maps that show natural and cultural features only in plan view are called planimetric maps, while maps that

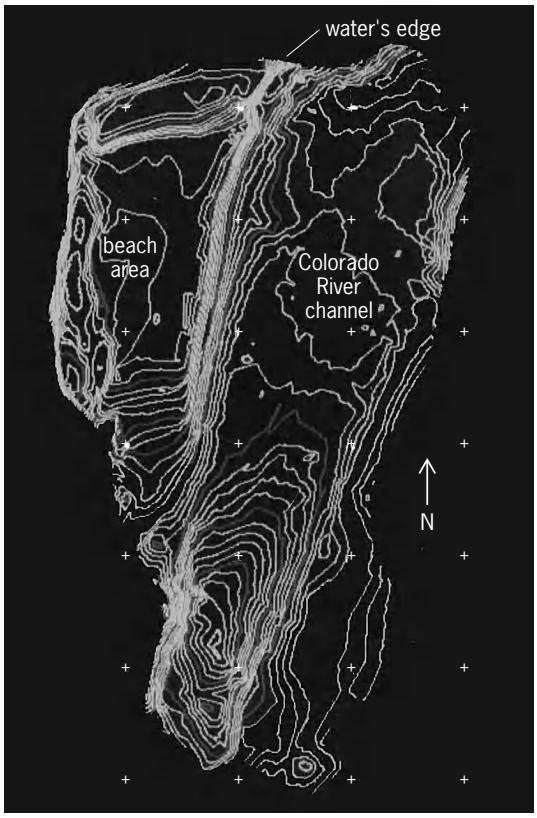

Fig. 1. Topographic and hydrographic map of the beach and river channel at River Mile 194 of the Colorado River in the Grand Canyon.

show relief are called hypsometric maps. Contour lines join points along a line of the same elevation across the ground. Contours show not only the elevation of the ground but also the geomorphic shape of features. Steepness of slope is determined by how close together the contour lines fall: the closer together the lines are, the steeper the slope. This is the most common way to show the physical characteristics of a piece of land (**Fig. 1**). *See* CONTOUR.

Hachures, also called vertical shading, are short lines drawn in the direction of slope: closely spaced and heavy for steep areas, farther apart and finer for more gentle slopes. Hachures give an impression of the shape of the land, but no indication of the actual elevation being represented. Shading and tinting are used in much the same way.

On maps that are used for navigation purposes, form lines are sometimes used. Form lines are like contour lines in that they follow lines of the same elevation, but they are not drawn with as high a degree of accuracy. This gives an impression of the shape of the features but does not show the actual elevation of the line itself.

A digital terrain model (DTM) is a computer-generated grid laid over the topographic information, which can then be rotated, tilted, and vertically exaggerated to give a three-dimensional view of the ground from different perspectives, including oblique representations. This technology is an excellent presentation tool: it utilizes the advances that have been made in computer mapping and drafting software.

Survey control. Prior to starting collection of data, a network of known horizontal and vertical control points must be established. This gives a reference location to the map, whether it be a local coordinate system or a state plane coordinate system, or geographically referenced to the rest of the world with latitude and longitude. The network also allows measurements made from several different locations in the same coordinate system to fit together into the same reference datum (the basis for the coordinate system).

Field methods. These include ground surveys, geographic positioning systems, and hydrographic surveys.

Ground surveys. Alidade and transit stadia surveys are performed, on occasion, but because of the increased speed and accuracy (centimeter-level measurements) most ground-based survey data are collected with a total station survey instrument and an electronic data collector. This information is transferred to a computer for conversion to coordinate locations and elevations. These data can then be imported into a modeling software package that generates contours at a specified interval by using description codes from the collected data to show breaks in the topography, and to draw in features such as roads, railway lines, streambeds, or structures. The information can then be plotted to create a hard copy of the maps at the desired scale to adequately display the information for the map's intended purpose. *See* SURVEYING INSTRUMENTS.

Geographic positioning system (GPS). This type of survey can be made to define topography. The Geographic Positioning System utilizes orbiting satellites operated by the U.S. Department of Defense. The satellites send out radio signals giving the precise location of the satellite and the time to a receiver on the ground. Several satellites are observed at the same time, and the position of the receiver on the ground is determined by correcting the signals sent by the satellites to a known point on the ground. The technology of the Geographic Positioning System provides virtually the same accuracy as a total station survey, but the receivers are too expensive to be in wide use. As the radio signals cannot pass through solid obstructions, this type of surveying is sometimes difficult, if not impossible, in cases of heavy tree canopy, tall buildings blocking the sky, or other features interfering with the signals.

Hydrographic surveys. The survey of a body of water, whether the ocean, a harbor, bay, lake, reservoir, or river, is a hydrographic survey. This type of survey is used to define the shoreline, depth of the water, shoals, rock outcrops, or other navigation hazards, as well as tidal changes and currents. Hydrographic surveys are used for dredging and maintenance of harbors, determining siltation of estuaries and bays, locating navigable channels, determining the sediment transport in rivers, and evaluating habitat for aquatic wildlife.

To produce a map of the bottom of a body of

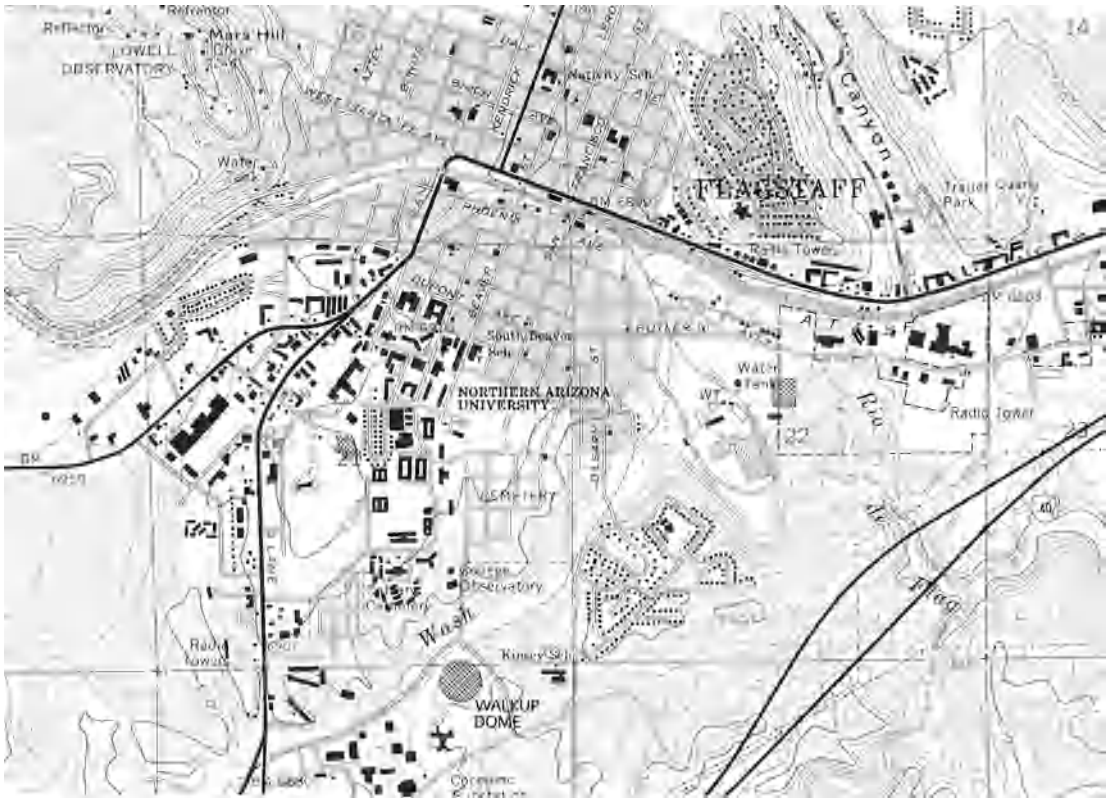

Fig. 2. Portion of USGS quadrangle sheet, showing topography and cultural information.

water requires combining the location of a point with the depth of the water at that point. Determining depth can be done in a variety of ways, from simply using a weight on the end of a graduated line dropped from a boat, to using a boat equipped with sonar to measure depth. After a control network has been established, the position of the boat can be recorded. A survey instrument set up on shore can be used to track the boat, with distance and angle measurements recorded and time-integrated with the recorded depths. This can be done either by hand recording or by using electronic devices that send the measurements from the shore to the boat by radio, where a computer program combines and stores the data. This process creates depth and position data for a great many points in a short amount of time. The position of the boat can also be determined by inertial guidance systems, distance measurements made from two or more radio beacons, or Global Positioning System locations. *See* HYDROGRAPHY; SONAR.

Photogrammetry and remote sensing. These techniques involve the art, science, and technology of using photography to obtain reliable measurements. The first aerial photography took place in 1858 to map the countryside. Emperor Napoleon III ordered aerial reconnaissance photos taken to prepare for the battle of Solferino in northern Italy in 1859. English, Russian and Germans used kites and rockets to bring single, multiple, and even gyro-stabilized cameras over their mapping areas. Photographs can be taken from airplanes, helicopters, and even satellites; thus the term remote sensing is applied to this technology.

By using pairs of photographs, a stereo image can be produced form which very accurate locations and elevations can be determined. This can then be transferred to a hard-copy map of a specific scale. The U.S. Geological Survey (USGS) has produced topographic maps of most of the United States showing topography, hydrology, vegetation, cultural features, and political boundaries (**Fig. 2**). The USGS has produced topographic maps of the Moon and Mars, and is mapping other planets and moons as well, by using radar and digital images from satellites near these bodies.

Although photogrammetry is most commonly used to map the ground, it can also be used to map objects that are difficult to study in other ways. The objects can be as different as the face of a dam, an astronomic radio reflector that deforms because of environmental conditions, or even the shapes of living organisms. Different types of applications use different types of collection techniques. Microwave, infrared, radar, ultraviolet, and multispectral data collection can be focused on different needs, such as thematic mapping, resource surveys, and topographic maps for foresters, engineers, geographers, geologists, surveyors, and others with special mapping needs. *See* AERIAL PHOTOGRAPH; PHOTOGRAMMETRY; REMOTE SENSING.

Plotting. Survey data are plotted primarily by mechanical devices, rather than being drafted by hand. This allows for very quick and accurate generation of maps as well as ease of making changes in scale, color, orientation, or location on the map page. These devices use ink pens, ink that is sprayed onto the medium, wax that is heated and applied to paper,

or coloring that is applied electrostatically. Some of the devices that use ink or wax are capable of producing several hundred colors, enabling the mapmaker to generate a highly customized product for a special purpose.

Maps need to show scale, contour interval, and a north arrow to illustrate orientation (magnetic, geodetic, polar, or grid north). A graphic scale or grid ticks with the coordinate location represented allows distance on the ground to be determined. A location map is also useful to provide the map reader with information on the general area that the large-scale map is detailing. *See* CARTOGRAPHY; MAP PROJECTIONS; MAP REPRODUCTION. Chris Brod

Bibliography. J. M. Anderson and E. M. Mikhail, *Surveying: Theory and Practice*, 7th ed., 1997; R. Graham, *Digital Aerial Survey: Theory and Practice*, 2002; E. M. Mikhail, J. S. Bethel, and J. C. McGlone, *Introduction to Modern Photogrammetry*, 2001; P. R. Wolf and R. C. Brinker, *Elementary Surveying*, 9th ed., 1997.

Topological dynamics

The study and application of topological transformation groups. Topological dynamics originated in the late-nineteenth-century investigations of the qualitative behavior of the solutions to the differential equations of classical mechanics by Henri Poincaré. The emergence of topological dynamics as a mathematical discipline occurred in the twentieth century with an abstract formulation of certain qualitative features of the classical systems by G. D. Birkhoff.

Topological transformation groups. A topological transformation group, or simply transformation group, is a triple (G, X, π), where G is a topological group called the phase group, X a topological space called the phase space, and π a continuous mapping of $G \times X$ onto X satisfying the homomorphism rule $\pi(gh, x) = \pi(g, \pi(h, x))$, $g, h \in G$, $x \in X$. The set $\{\pi(g, x) | g \in G\}$ is called the orbit through x. When G is the group $\mathbf{R}$ of real numbers, the transformation group is called a flow; when G is the group $\mathbf{Z}$ of integers, the transformation group is called a cascade. *See* ABSTRACT ALGEBRA; GROUP THEORY; TOPOLOGY.

Homomorphism and isomorphism. Let (G, X, π) and (G, X', π') be transformation groups. A continuous mapping φ of X onto X' is called a homomorphism if it is equivariant, that is, $\varphi(\pi(g, x)) = \pi'(g, \varphi(x))$. If φ is one-to-one with a continuous inverse, it is called an isomorphism.

Invariant measure. The Borel subsets of a topological space are the elements of the smallest σ-algebra of subsets of the space containing the open sets; a Borel measure is a measure on the Borel sets. A borel measure μ is invariant for a transformation group (G, X, π) if for all Borel sets A, $\mu(A_g) = \mu(A)$, where $A_g = \{\pi(g, x) | x \in \mathrm{A}\}$. The use of invariant measures in the study of dynamical systems is original with Poincaré.

Classical dynamical systems. A classical (autonomous) dynamical system is a first-order system of ordinary differential equations (1). The functions F_i,

$$\frac{dx_i}{dt} = F_i(x_1, \ldots, x_n) \qquad (1 \le i \le n) \qquad (1)$$

defined on an open set X in n-dimensional space ($\mathbf{R}^n$), are assumed to be such that (a) for each $y \in X$ there is a unique solution to Eq. (1), $x(t)$, $-\infty < t < \infty$, with $x(0) = y$; and (b) the solution in (a) varies continuously in y,t. A flow $(\mathbf{R}, X, \pi)$ may be associated to Eq. (1) by defining $\pi(t, y)$ to be the value at time t of the solution in (a). The homomorphism rule $\pi(s + t, y) = \pi(s, \pi(t, y))$ is a consequence of the uniqueness statement in Eq. (1). *See* DIFFERENTIAL EQUATION.

Hamiltonian systems. Let $X = U \times \mathbf{R}^m$, where U is an open subset of $\mathbf{R}^m$. Denote coordinates in X by $(q, p) = (q_1, \ldots, q_m, p_1, \ldots, p_m)$. A continuously differentiable (C^1) function H on X determines a hamiltonian system, given by Eqs. (2). In terms of the

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \qquad \frac{dp_i}{dt} = \frac{\partial H}{\partial q_i} \qquad (1 \le i \le m) \qquad (2)$$

Lagrange equations for a conservative system with m degrees of freedom, $q_1, \ldots, q_m$ represent position coordinates, $p_1, \ldots, p_m$ represent impulses or momenta, and $H = T + V$ is total energy (kinetic plus potential). H is called the hamiltonian of Eq. (2). *See* HAMILTON'S EQUATIONS OF MOTION; LAGRANGE'S EQUATIONS.

Liouville's theorem. Assume the functions F_i in Eq. (1) are C^1 on X. Lebesgue measure on X is an invariant measure for the associated flow if and only if the vector $F = (F_1, \ldots, F_n)$ has divergence 0, Eq. (3).

$$\operatorname{div} F = \sum_{j=1}^{n} \frac{\partial F_j}{\partial x_j} = 0 \qquad (3)$$

An immediate consequence of this theorem is that the flow associated to a hamiltonian system with a C^2 hamiltonian function preserves Lebesgue measure.

Poincaré recurrence theorem. This theorem, probably the first theorem in "ergodic theory," asserts that if a flow has a finite invariant measure μ, then with respect to μ almost every point $x \in X$ is Poisson-stable; for any neighborhood U of x there exist arbitrarily large (positive and negative) values of t such that $\pi(t,x) \in U$. Every Poisson-stable point belongs to the nonwandering set, the closed flow-invariant set of points x such that for any neighborhood U, $U \cap U_t \neq \phi$ for arbitrarily large values of t. In the case of Eq. (1), if X has finite volume and $\operatorname{div} F = 0$, it follows that every point belongs to the nonwandering set. *See* STATISTICAL MECHANICS.

Flows on manifolds. Let X be a differentiable manifold. A vector field F on X generates the flow $(\mathbf{R}, X, \pi)$ if $(d/dt)\pi(t, x) = F(\pi(t, x))$. That is, for every C_1 function f on X, $(d/dt)f(\pi(t, x))|_{t=0} = F(x)f(x)$. In the case of a classical system of Eq. (1), the vector field is given by Eq. (4). If X is a compact

$$F(x) = \sum_{j=1}^{n} F_j(x) \frac{\partial}{\partial x_j} \qquad (4)$$

manifold, every C_1 vector field on X generates a C_1 flow. *See* MANIFOLD (MATHEMATICS).

Integrals. Flows on manifolds occur naturally in the study of classical systems of Eq. (1). A C^1 function H is an integral of the system defined by Eq. (1) if its gradient is everywhere perpendicular to the vector field defined by Eq. (4); that is, Eq. (5) is satisfied.

$$\sum_{j=1}^{n} F_j(x)\,\frac{\partial H}{\partial x_j} = 0 \tag{5}$$

For example, the hamiltonian of a hamiltonian system defined by Eq. (2) is automatically an integral of that system. By the chain rule, if H is an integral, the level surface $H = c$ is flow-invariant. For almost all real numbers c, this level surface is empty or else an $(n-1)$-dimensional differentiable manifold (Sard's theorem), and in the latter case $(\mathbf{R},\{H = c\},\pi)$ is a flow on a manifold.

Geodesic flow. Let M be a compact riemannian manifold with tangent bundle TM, and let T_1M be the bundle of tangent vectors of unit length. If υ_p is a unit tangent vector at p, there exists a unique geodesic $x(t)$, $-\infty < t < \infty$, parametrized by arc length "in the direction υ_p" (so that negative arc length makes sense) with $x(0) = p$, $(dx/ds)|_{s=0} = \upsilon_p$. Define $\pi(t,\upsilon_p) = (dx/ds)|_{s=t}$. Since arc length is the parameter, $\pi(t,\upsilon_p)$ is a unit tangent vector [at $x(t)$], and $(\mathbf{R},T_1M,\pi)$ is a flow, called geodesic flow. *See* FIBER BUNDLE; RIEMANNIAN GEOMETRY.

Principle of least action. Let U be an open set in $\mathbf{R}^m$. Given a positive definite kinetic energy, defined by Eq. (6), and a potential energy $V(q)$, the Lagrange equations of motion are given by Eq. (7) where $L =$

$$T(q,\dot{q}) = {}^1/_2 \sum_{i,j=1}^{m} a_{ij}(q)\dot{q}_i\dot{q}_j \tag{6}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \qquad (1 \le i \le m) \tag{7}$$

$T - V$. [The impulses p_i in Eq. (2) are defined by Eq. (8).] The principle of least action of Maupertius-

$$p_i = \sum_{j=1}^{m} a_{ij}(q)\dot{q}_j \tag{8}$$

Euler-Lagrange-Jacobi states that among all the curves C on which the total energy $H = T + V$ is constant (h), the one joining a pair of points x_0, x_1 in U and representing a true motion of the system is the one which minimizes the "action integral"

$$_C\!\int 2T$$

Using Eq. (9), one can define a metric by Eq. (10).

$$T = h - V = (h - V)^{1/2}\left(\sum_{i,j=1}^{m} a_{ij}(q)\dot{q}_i\dot{q}_j\right)^{1/2} \tag{9}$$

$$ds^2 = 2(h - V(q))\sum_{i,j=1}^{m} a_{ij}(q)\,dq_i dq_j \tag{10}$$

The principle of least action now says that geodesics for ds^2 are true motions of the system. The speed of travel along a geodesic is $ds/dt = h - V$, but by a "change of clock" this speed can be made to be 1, and the geodesic flow results. *See* LEAST-ACTION PRINCIPLE.

Whitney sums. Let X be a differentiable manifold with tangent bundle TX. TX is the Whitney sum of continuous bundles A and B over X, $TX = A \oplus B$, if for each x the fibers of A and B at x are complementary subspaces of the tangent space at x. One can speak similarly of a Whitney sum of three or more bundles.

Hyperbolic structure. Let X be a compact riemannian manifold on which the length of a tangent vector is denoted $\|\upsilon\|$. A hyperbolic structure for a flow $(\mathbf{R}, X, \pi)$ is a decomposition of TX into a Whitney sum of three bundles, E, E^u, and E^s, each of which is flow-invariant (for the derivative of the flow) and such that (i) E is tangent to the flow (that is, to the orbits) and (ii) there exist constants $c < \infty$ and $\lambda > 0$ such that inequalities (11) and (12) are satisfied. E^s is

$$\|d\pi(t,x)v_x\| \;\le\; ce^{-\lambda t}\|v_x\| \tag{11}$$

$$(v_x \in E^s_x,\ t \ge 0)$$

$$\|d\pi(t,x)v_x\| \;\le\; ce^{\lambda t}\|v_x\| \tag{12}$$

$$(v_x \in E^u_x,\ t \le 0)$$

called the stable bundle and E^u the unstable bundle. The concept of a hyperbolic structure is also used for a diffeomorphism (or for the cascade generated by its repeated application); it differs from the hyperbolic structure of a flow only in the requirement that $TX = E^s \oplus E^u$, necessary because one cannot speak of a tangent to an orbit of a cascade.

Anosov flow. A differentiable flow on a manifold of dimension greater than 1 is said to be an Anosov flow if there exists a hyperbolic structure for the flow. The most important example of an Anosov flow is the geodesic flow on a compact riemannian manifold of negative sectional curvature.

Anosov diffeomorphism. A diffeomorphism is an Anosov diffeomorphism if there exists a hyperbolic structure for the diffeomorphism. The known examples are algebraic in nature. For example, let A be an $n \times n$ matrix with integer entries and determinant 1. As a linear transformation of $\mathbf{R}^n$, A sends the integer lattice $\mathbf{Z}^n$ to itself and therefore defines an automorphism of the quotient $\mathbf{R}^n/\mathbf{Z}^n = T^n$, the n-dimensional torus. If A has no eigenvalues of absolute value 1, the automorphism ("hyperbolic automorphism") is an Anosov diffeomorphism. J. Franks and A. Manning have shown that any Anosov diffeomorphism of the torus is isomorphic to some hyperbolic automorphism of the same torus; that is, the corresponding cascades are isomorphic.

Structural stability. A differentiable flow $(\mathbf{R}, X, \pi)$ on a compact manifold X is "structurally stable" if the phase portrait (or orbit structure) is insensitive to small perturbations in the equations (vector field) governing the flow. It is possible to give the set of all C^1 vector fields, which has a natural vector space structure, a norm (the C^1 norm) with respect to which it is a Banach space. A C^1 vector field F is structurally stable if there exists $\epsilon > 0$ such that whenever

G is a C^1 vector field with $\|F - G\| < \epsilon$, G and F generate flows with isomorphic orbit structure; that is, there is a homeomorphism of X which takes G orbits into F orbits, although "time" may not be preserved. The equations governing any physical system can in practice only be approximated, but when the "true" system is structurally stable, a sufficiently good approximation to it will have the same orbit structure. One can also define structural stability for diffeomorphisms (which also have a natural C^1 topology).

Anosov's theorem. Every Anosov flow or diffeomorphism is structurally stable. This implies the geodesic flow on a compact, negatively curved riemannian manifold is structurally stable; similarly, a hyperbolic automorphism of the torus is structurally stable.

Symbolic dynamics. Let Λ_r be a set with $r > 1$ elements, say $\Lambda_r = \{1,2,\ldots,r\}$. Provided with a natural metric, the set X_r of all bisequences $x = \{x_n\}^{\infty}{}_{n=-\infty}$ of elements from Λ_r is a compact metric space, and the left shift, σ, $(\sigma x)_n = x_{n+1}$, a homeomorphism of this space. This defines a cascade called the shift (on r symbols), written (σ, X_r) [rather than $(\mathbf{Z}, X_r, \pi)$]. Symbolic dynamics is the study of the shift, for its own sake and for application to other systems.

J. Hadamard (1898) set up a correspondence between certain symbolic sequences and geodesics on a negatively curved surface. Utilizing this correspondence, M. Morse (1922) proved the corresponding geodesic flow has minimal sets not consisting of a single closed orbit. (A minimal set is a nonempty closed invariant set containing no proper subset with the same property.) Generalizing this approach, Y. G. Sinai and R. Bowen have constructed an elaborate symbolic dynamics for flows and diffeomorphisms which are suitably "hyperbolic" (for example, Anosov). The theory has numerous applications, among them (a) a generalization of Morse's result to Anosov flows with smooth invariant measure, in particular the geodesic flow on a compact negatively curved manifold; (b) the statement that every minimal set for an Anosov flow is one-dimensional (that is, "small"); and (c) remarkable asymptotic formulas for the distribution of periodic orbits in Anosov flows (with application to the distribution of closed geodesics on a negatively curved manifold).

Ergodic theory. Ergodic theory is the study of measure-preserving transformations. Let X be a set, $\mathscr{B}$ a σ-field of subsets of X, and μ a measure on $\mathscr{B}$. A transformation T of X to itself is measure-preserving if $T^{-1}B \in \mathscr{B}$ and $\mu(T^{-1}B) = \mu(B)$ for all $B \in \mathscr{B}$. If T is one-to-one onto, and if both T and T^{-1} are measure-preserving, T is said to be an automorphism of $(X, \mathscr{B}, \mu)$. If G is a topological group, a representation of G on (X,B,μ) is a homomorphism $(g \to T^g)$ from G to the automorphism group of $(X, \mathscr{B}, \mu)$ with the property that $T^g x$ is measurable on the product space $G \times X$. Representations T^g and $T^g{}_0$ on $(X, \mathscr{B}, \mu)$ and $X_0, \mathscr{B}_0, \mu_0)$ are isomorphic if there exists a one-to-one equivariant map φ of X onto X_0 such that $\mu_0(\varphi B) = \mu(B)$ for all $B \in \mathscr{B}$, $\mu(\varphi^{-1}B_0) = \mu_0(B_0)$ for all $B_0 \in \mathscr{B}_0$.

Birkhoff ergodic theorem. Let T^t be a representation of $\mathbf{R}$ on a probability space $(X, \mathscr{B}, \mu)$. Birkhoff's ergodic theorem says that if f is integrable on X, then for almost all $x \in X$ the limit given by Eq. (13) exists;

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(T^t x)\, dt = \bar{f}(x) \tag{13}$$

furthermore, $\bar{f}$ is integrable, and Eq. (14) holds. A

$$\int_X f(x)\mu(dx) = \int_X \bar{f}(x)\mu(dx) \tag{14}$$

similar statement holds for a representation of $\mathbf{Z}$, the integral in Eq. (13) being replaced by a sum

$$\frac{1}{T} \sum_{t=0}^{T-1} f(T^t x)$$

Ergodicity and mixing. A group of automorphisms of a probability space is ergodic (or metrically transitive) if every measurable invariant set has measure 0 or 1. In the presence of ergodicity the function $\bar{f}$ in Eq. (13) is almost everywhere equal to

$$\int_X f(x)\mu(dx)$$

and hence "time averages equal space averages." Because of this interpretation, ergodicity is important for flows with invariant measure which are unstable, that is, sensitive to perturbations of the initial conditions (for example, Anosov flows).

An equivalent formulation of ergodicity (for $\mathbf{R}$) is the condition that Eq. (15) is satisfied for all A,B

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \mu(A \cap T^t B)\, dt = \mu(A)\, \mu(B) \tag{15}$$

$\in \mathscr{B}$. The condition for mixing is an "unaveraged" analog of Eq. (15) given by Eq. (16). Between the

$$\lim_{t\to\infty} \mu(A \cap T^t B) = \mu(A)\, \mu(B) \tag{16}$$

two notions is the notion of weak mixing, defined by Eq. (17). If A and $T^t B$ are independent sets, then by

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \left|\mu(A \cap T^t B) - \mu(A)\mu(B)\right| dt = 0 \tag{17}$$

definition $\mu(A \cap T^t B) = \mu(A)\mu(T^t B) = \mu(A)\mu(B)$. Ergodicity and the two kinds of mixing thus express, to varying degrees, the notion of asymptotic independence. D. V. Anosov and Sinai have proved an Anosov diffeomorphism with smooth invariant measure is mixing; an Anosov flow with smooth invariant measure is ergodic and, if it is weak-mixing, mixing; the geodesic flow on a compact, negatively curved riemannian manifold is mixing (in dimension two, a classical theorem of E. Hopf and G. A. Hedlund). Stronger statements have since been made.

Entropy. Let T be a measure-preserving transformation of a probability space (X, B, μ). Define $\eta(t)$, $0 \le t \le 1$, to be 0 if $t = 0$ and $-t \log t$ otherwise. If $P = \{A_1, \ldots, A_r\}$ is a partition of X into pairwise disjoint measurable sets, let P_N, $N \ge 1$, be the partition whose elements are sets of the form

$$A_{j_0} \cap T^{-1}A_{j_1} \cap \cdots \cap T^{-(N-1)}A_{j_{N-1}}$$

$$1 \le j_0, \ldots, j_{N-1} \le r$$

Define $$H(P_N) = \sum_{A \in P_N} \eta(\mu(A))$$

The limit $$\lim_{N\to\infty} \frac{H(P_N)}{N} = h(T, P)$$

exists and is called the entropy of T with respect to P. One has $0 \le h(T, P) \le \log r$; $h(T, P) = 0$ if and only if P is measurable with respect to the σ-algebra generated by $T^{-1}P, T^{-2}P, \ldots$. Thus entropy is a measure of "determinism." Finally $h(T)$, the entropy of T, is the supremum of $h(T, P)$ over all P. In some cases $h(T) = \infty$; but, for example, if T is a diffeomorphism of a compact differentiable manifold X, and if μ is smooth, then Kushnirenko's theorem states that $h(T) < \infty$. *See* ENTROPY.

Ornstein's theorem. If $p_1, \ldots, p_r$ are probabilities, it is possible to define a Borel probability measure μ on the space X_r of symbol sequences which governs the process of selecting elements from Λ_r independently and according to the given probabilities. The shift σ preserves μ, and is called a Bernoulli shift. A. N. Kolmogerov and Sinai showed that

$$h(\sigma) = \sum_{j=1}^{r} \eta(p_j)$$

since entropy is an isomorphism invariant, this settled the long-standing question of whether the two shift with probabilities $(1/2, 1/2)$ is isomorphic to the three shift with probabilities $(1/3, 1/3, 1/3)$. (They are not isomorphic because $\log 2 \neq \log 3$.) D. Ornstein (1969) established the deep result that entropy is a complete invariant among the Bernoulli shifts; two Bernoulli shifts with the same entropy are isomorphic. *See* PROBABILITY.

Bernoulli shifts as models. Ornstein's results have led to the measure theoretic classification of a number of cascades and flows. N. Friedman, Ornstein, Sinai, and R. Azencott showed that a C^2 Anosov diffeomorphism with smooth invariant measure is Bernoulli; Y. Katznelson showed that an ergodic automorphism of the n-dimensional torus is Bernoulli; Ornstein and B. Weiss showed that the geodesic flow on a compact riemannian manifold of negative curvature is Bernoulli (where a flow is defined to be Bernoulli if each of its time t maps is Bernoulli); and M. Ratner showed that an Anosov flow with smooth invariant measure is Bernoulli.

Minimal transformation groups. A transformation group is minimal if every orbit is dense in the phase space. (The concept is due to Birkhoff.) It is of interest to classify minimal transformation groups, one reason being that in an arbitrary compact phase space there exist nonempty closed invariant sets which are minimal. In what follows, X is a compact metric space with metric $d(x, y)$, (G, X, π) and $\pi(g, x)$ will be written (G, X) and gx, and transformation groups (arbitrary G) will be called flows.

Equicontinuous flows. A flow is equicontinuous if for each $\epsilon > 0$ there exists $\delta > 0$ such that $d(gx, gy) < \epsilon$, for all $g \in G$, whenever $d(x, y) < \delta$.

$$D(x, y) = \sup_g d(gx, gy)$$

is a compatible G invariant metric for X when (G, X) is equicontinuous. Since the group of isometries of (X, D) is, with a natural topology, compact, G maps homomorphically into a compact metric group acting on X. If (G, X) is minimal, the closure of the image of G in this group acts transitively on X. The study of minimal equicontinuous flows is therefore essentially the study of transitive compact point actions.

Distal flows. A point $x \in X$ is distal for (G, X) if

$$\inf_g d(gx, gy) = 0$$

implies $y = x$. This flow is distal if every point is distal. Every equicontinuous flow is distal, but there exist minimal distal flows which are not equicontinuous. (However, R. Ellis has shown that if G is finitely generated and X totally disconnected, distality implies equicontinuity. For example, a minimal distal subset of the shift is finite.) Minimal distal flows have been characterized by H. Furstenberg in terms of "isometric extensions." A homomorphism $\varphi:(G, X) \to (G, X')$ defines an isometric extension if there is a continuous function R on $\Delta = \{(x, y) \in X \times X | \varphi(x) = \varphi(y)\}$ such that (a) $R(gx, gy) = R(x, y)$ for all $g \in G$, and (b) for each $x' \in X'$, R defines a metric on $\varphi^{-1}x'$. Furstenberg's structure theorem states that a minimal distal flow can be "built up" by a transfinite sequence of isometric extensions and "inverse limits" beginning with the trivial flow (one-point phase space).

Point distal flows. If there exists a distal point with dense orbit, the flow (G, X) is point-distal. Point-distal flows are minimal and not necessarily distal. For example, there exist nonequicontinuous, hence nondistal, cascades with totally disconnected phase space. Furstenberg's structure theorem is generalized to point-distal flows using the notion of almost one-to-one extension, a homomorphism $\varphi:(G, X) \to (G, X')$ such that for some $x' \in X'$, $\varphi^{-1}x'$ is a single point. W. Veech and Ellis have shown that every point-distal flow has an almost one-to-one extension which can be built up from the one-point flow by a transfinite sequence of isometric extensions, almost one-to-one extensions, and inverse limits. A general structure theory for minimal flows has been established.

William A. Veech

Bibliography. E. Akin, *The General Topology of Dynamical Systems*, 1993; J. Alexander (ed.), *Dynamical Systems*, 1988; V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics*, 1968, reprint 1989; J. DeVries, *Elements of Topological Dynamics*, 1993; R. Mane, *Ergodic Theory and Differentiable Dynamics*, 1987; D. Ruelle, *Elements of Differentiable Dynamics and Bifurcation Theory*, 1989; C. L. Siegel and J. K. Moser, *Lectures on Celestial Mechanics*, 2d ed., 1994.

Topology

The branch of mathematics that studies the qualitative properties of spaces, as opposed to the more delicate and refined geometric or analytic properties. While there are earlier results that belong to the field, the beginning of the subject as a separate branch of mathematics dates to the work of H. Poincaré during 1895–1904. The ideas and results of topology have a central place in mathematics, with connections to almost all the other areas of the subject.

The difference between topological and geometric properties is illustrated by the example of a space with three separate pieces. The exact shapes of the pieces constitute a geometric property of the space, and the study of these shapes is in the domain of differential geometry, but the fact that the space has three separate pieces is a qualitative or topological property. As another example, if a round sphere is deformed to be pear-shaped (or even more irregularly shaped, like the surface of the Earth), then the geometric notions of distance, straight line, and angle are changed, but the topological properties of the surface are left unchanged. However, if a handle is added by cutting two holes in the sphere and connecting them by a curved pipe, then the topology of the surface is changed (**Fig. 1**).

Development. Four major areas of topology are algebraic topology, homotopy theory, general topology, and manifold theory. Algebraic topology, the first area of modern topology to be developed, is concerned with associating algebraic invariants to geometric spaces in order to measure higher-dimensional analogs of the number of pieces of a space or the number of handles of a surface. Algebraic topology has tremendous influence on other branches of mathematics, both direct (application of the invariants of algebraic topology to problems from other areas of mathematics and physics) and indirect (application in other contexts of ideas arising from algebraic topology).

As algebraic topology developed, it became clear that if one function could be continuously deformed to another (that is, if they were homotopic), then these two functions behaved in the same way as far as the invariants of algebraic topology were concerned.

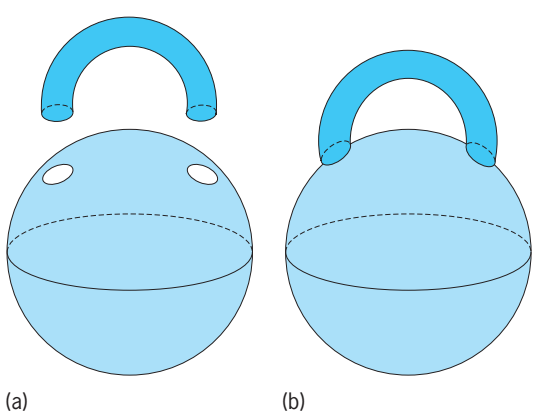

Fig. 1. Process of adding a handle to a 2-sphere. (*a*) Cutting of holes in sphere. (*b*) Connecting of holes by a curved pipe.

This led naturally to the study of invariants that remain unchanged as the maps are deformed by homotopies, that is, homotopy invariants. This study, which is an offshoot of algebraic topology, is called homotopy theory. Some of the most interesting homotopy invariants are the higher homotopy groups. These proved extremely difficult to compute, even for spaces as simple as the sphere, and are the subject of much investigation.

Early in the development of topology it was realized that the foundations of the subject needed attention. General or point-set topology studies the relationships between the basic topological properties that spaces may possess.

Before abstract topological spaces were defined, there were numerous examples of spaces arising from geometric and analytic problems. The most important of these is a class of spaces known as manifolds. Both because of their ubiquitous appearance throughout mathematics and because they possess extraordinarily rich topological properties, manifolds became one of the central objects of study in topology. The basic theme in manifold theory is to find sufficient algebraic invariants to classify, that is, to list comprehensively, all manifolds, and to give methods for evaluating these invariants in geometric cases. *See* MANIFOLD (MATHEMATICS).

General topology. The coordinate space of calculus, $\mathbf{R}^n$, is the space of all ordered n-tuples $(x_1, \ldots, x_n)$ of real numbers. On this space there is the pythagorean (or euclidean) distance function. If $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ are points of $\mathbf{R}^n$, then the distance from x to y is given by Eq. (1).

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2} \qquad (1)$$

This distance function indicates how the points of $\mathbf{R}^n$ are related to each other, and as discussed below, it gives $\mathbf{R}^n$ the structure of a topological space. An open subset of $\mathbf{R}^n$ is a subset $U \subset \mathbf{R}^n$ with the property that for each $x \in U$ there is a real number $\delta > 0$ such that all points y for which $d(x, y) < \delta$ are also contained in U. (The notation $x \in U$ means that x is a member of U.) The intersection of any finite number of open subsets of $\mathbf{R}^n$ is itself an open subset, and the union of any collection of open subsets of $\mathbf{R}^n$ is an open subset. *See* SET THEORY.

Topological spaces. This situation is vastly generalized by the notion of an abstract topological space. A topological space, or a space for short, is a pair consisting of a set S, called the underlying set of the space, and a collection of subsets $\{U_\alpha\}_{\alpha \in A}$ called the open subsets of the space. The collection of open subsets is required to satisfy four axioms:

1. The empty subset is an open subset of the space.
2. The entire set S is an open subset of the space.
3. If $U_1, \ldots, U_n$ is a finite collection of open subsets of the space, then the intersection $U_1 \cap \cdots \cap U_n$ is an open subset of the space.
4. If $\{U_\beta\}_{\beta \in B}$ is any collection of open subsets of the space, then the union $\cup_{\beta \in B} U_\beta$ is an open subset of the space.

The open subsets of a space X that contain a given point x are thought of as containing all the points of X close to x. Thus, a topology also describes how the points of a space cohere. The collection of open sets in $\mathbf{R}^n$ as defined above form a topology on $\mathbf{R}^n$.

A subspace A of X consists of a subset T of the set S underlying X with the induced topology; that is, a subset of A is open in A if and only if it is the intersection of an open set of X with T. The unit interval I is the subspace of $\mathbf{R}^1$ given by Eq. (2).

$$I = \{t \in \mathbf{R}^1 \,|\, 0 \leq t \leq 1\} \tag{2}$$

If X and Y are spaces, then a new space, denoted $X \times Y$, can be formed. The underlying set is simply the set of ordered pairs $\{(x,y)|x \in X, y \in Y\}$. The topology, called the product topology, is as follows: Any subset of $X \times Y$ of the form $U \times V$, where U is an open subset of X, and V is an open subset of Y, is open in $X \times Y$. More generally, any union of such sets is open in $X \times Y$.

Continuous maps and homeomorphisms. A set function f: $X \to Y$ is said to be continuous if, for every open subset U of Y, the subset of points of X that map by f to U, denoted $f^{-1}(U)$, is an open subset of X. Continuous functions are called maps. They preserve the notions of coherence in the spaces X and Y in the sense that close points in X are sent by f to close points in Y. This notion generalizes the usual notion of continuity from calculus for maps between subspaces of the real line. *See* CALCULUS.

A homeomorphism from X to Y is a continuous map f from X to Y, which is a bijection on the underlying sets and whose inverse function is also continuous. General topology is the study of topological spaces and continuous maps between them. In general topology, two spaces are considered equivalent if there is a homeomorphism between them.

An example of homeomorphic spaces is the unit circle $S^1 = \{(x,y) \in \mathbf{R}^2 | x^2 + y^2 = 1\}$ and the ellipse $\{(x,y) \in \mathbf{R}^2 | x^2 + (y^2/4) = 1\}$. The map $(x,y) \to (x, 2y)$ sends the circle to the ellipse. The map $(x, y) \to (x,y/2)$ sends the ellipse to the circle. Since these maps are continuous and inverses of each other, they are homeomorphisms. Hence, the circle and the ellipse are homeomorphic. The boundary of the unit square in the plane is made up of four straight lines meeting at right angles. It is homeomorphic to the unit circle. On the other hand, there is a continuous map from the half-open unit interval $J = \{t \in \mathbf{R}^1 | 0 \leq t < 1\}$ to S^1 given by $t \to (\cos(2\pi t), \sin(2\pi t))$. Though this map is a bijection, its inverse is not continuous since the inverse function sends (1,0) to 0 and sends points close to (1,0) with negative y coordinates to points near 1 instead of near 0. In fact, the circle and the half-open interval are not homeomorphic.

These examples indicate why topology is often described as the study of properties of spaces that remain constant under "deformation and bending but without tearing"; these words roughly describe a homeomorphism. Geometric properties such as size, curvature, straightness, and angles are completely lost when only properties invariant under homeomorphism are considered.

The circle and the 2-sphere $S^2 = \{(x, y, z) \in \mathbf{R}^3 | x^2 + y^2 + z^2 = 1\}$ are not homeomorphic. This result follows from one of the most celebrated theorems in general topology, called the invariance of domain theorem. It states that if U and V are nonempty open subsets of $\mathbf{R}^n$ and $\mathbf{R}^m$, respectively, and if U and V are homeomorphic, then $n = m$.

Spaces with special properties. Most topological spaces that arise from other parts of mathematics have many special properties. For example, a topological space is said to be Hausdorff if for every pair of distinct points x and y in the space there are open subsets U and V such that $x \in U$, $y \in V$, and $U \cap V =$ (that is, the intersection of U and V is empty). The space $\mathbf{R}^n$ and all its subspaces are Hausdorff. Another important property is compactness. A topological space is compact if for any collection of open subsets whose union is the entire space there is a finite subcollection whose union is also the entire space. An important theorem says that a subspace X of $\mathbf{R}^n$ is compact if and only if it satisfies the following two conditions:

1. X is bounded in the sense that there is a real number $\mathbf{R} > 0$ such that all points of X have distance at most $\mathbf{R}$ from the origin.
2. X is closed in the sense that $\mathbf{R}^n - X$ is an open subspace of $\mathbf{R}^n$.

Much of general topology focuses on compact Hausdorff spaces.

Another important class of spaces is the metric spaces. A metric space is a set S equipped with a distance function (or metric) d:$S \times S \to \mathbf{R}$ satisfying the following three axioms:

1. For all $x, y \in S$, $d(x, y) \geq 0$.
2. $d(x, y) = 0$ if and only if $x = y$.
3. $d(x, y) + d(y, z) \geq d(x, z)$.

The pythagorean distance in $\mathbf{R}^n$ satisfies these axioms. If (S, d) is a metric space, then it has an induced topology defined analogously to the topology on $\mathbf{R}^n$. Many theorems of general topology concern what sorts of topological properties are sufficient to ensure that a space admits a metric inducing the given topology.

Algebraic topology. Algebraic topology is the study of algebraic invariants associated to topological spaces. An algebraic invariant of a space is an algebraic object associated to the space that remains unchanged if the space is replaced by a homeomorphic space. By an algebraic object is meant either an algebraic structure, such as a group, ring, or field, or an element of an algebraic structure.

Homology. This is the first and most important invariant of algebraic topology. Only a vague description of the main idea behind homology will be given. The homology groups $H_i(X)$, $i = 1, 2, \ldots$, of a space X are abelian groups associated to the space that measure the higher-dimensional analog of the number of pieces, that is, the number of higher-dimensional holes in X. The nth-homology group $H_n(X)$ is constructed out of the n-dimensional cycles (a cycle is a geometric compact object with no boundary) in X that go around the n-dimensional holes. In order to form the homology group, it is necessary to divide out by (that is, to set equal to zero) those cycles that

do not surround any hole of X and hence are the boundary of a compact $(n+1)$-dimensional geometric object in X. *See* GROUP THEORY.

While the homology groups are difficult to define and understand, once defined they can be computed for any space that is presented in a reasonable geometric fashion. For example, the unit n-sphere for $n \geq 1$, which is the subset of $\mathbf{R}^{n+1}$ given by Eq. (3),

$$S^n = \{x \in \mathbf{R}^{n+1} \,|d(x,0) = 1\} \tag{3}$$

has a single n-dimensional hole, which is surrounded by the entire sphere, and no other holes of any dimension. Thus, the homology groups of $H_i(S^n)$ are zero for all $i > 0$ except for $i = n$, and $H_n(S^n)$ has a single generator corresponding to the sphere itself. In particular, the n-sphere and the m-sphere are not homeomorphic if $n \neq m$. This homological computation can be used to give another proof of the invariance of domain theorem discussed above. This was one of the first applications of homology theory and indicated its power.

Euler characteristic. This is one of the most frequently used invariants of algebraic topology. A space X has an Euler characteristic when it has only finitely many nonzero homology groups and each of these is finitely generated. The Euler characteristic $\chi(X)$ is given by Eq. (4), where $\beta_i(X)$ is the ith-Betti number of X, which is by definition the rank of the ith-homology group of X.

$$\chi(X) = \sum_i (-1)\beta_i(X) \tag{4}$$

Fundamental group. This invariant of algebraic topology, also called the Poincaré group, formed from the set of loops in X that begin and end at a preassigned point $x_0 \in X$, or are based at x_0. Two such loops γ_1 and γ_2 can be composed by simply tracing γ_1 followed by γ_2 to give a new loop based at x_0. Two loops are said to be equivalent if the first can be continuously deformed through a family of loops based at x_0 to the second. The composition yields a group structure on the set of equivalence classes of loops based at x_0. This group is the fundamental group of X based at x_0 and is denoted $\pi_1(X, x_0)$. A space is said to be simply connected if its fundamental group is the trivial group.

An important result computes the fundamental groups of the spheres: The sphere S^n is simply connected if $n > 1$; and the fundamental group of the circle is isomorphic to the group of integers and is generated by the loop going once around the circle.

Applications. In applying algebraic topology to problems from other areas of mathematics, it is necessary to compute the values of the invariants. To a large degree this is feasible for the homology groups and the fundamental group. Therefore, these invariants play very important roles in differential and algebraic geometry. *See* ALGEBRAIC GEOMETRY; DIFFERENTIAL GEOMETRY.

An important theorem for computing an invariant of algebraic topology by geometry is the Gauss-Bonnet theorem. It states that, if S is a compact riemannian surface, then the Euler characteristic $\chi(S)$ is given by Eq. (5), where $K(x)$ is the gaussian curvature of the surface and dvol is the volume element. *See* RIEMANNIAN GEOMETRY.

$$\chi(X) = \frac{1}{2\pi}\int_S K(x)\text{dvol} \tag{5}$$

The ideas of algebraic topology have had profound effect even in situations that are not directly topological. The homology of groups is an important tool in Galois theory and representation theory. Much progress in number theory has revolved around finding reasonable analogs to the invariants of algebraic topology. *See* NUMBER THEORY.

Homotopy theory. In algebraic topology the deformation of continuous maps is extremely important. The defnition of homotopy formalizes this notion. A homotopy of maps from X to Y is a continuous map given by expression (6), where I is the unit interval.

$$H: X \times I \to Y \tag{6}$$

Setting h_t equal to the restriction of H to the slice $X \times \{t\}$ gives a family of continuous maps from X to Y. The homotopy H is said to be a homotopy from h_0 to h_1, and h_0 and h_1 are said to be homotopic. The set of homotopy classes of maps from X to Y can then be introduced. The task of homotopy theory is to compute in some effective way the set of homotopy classes of maps from one space to another.

Not only do the invariants of algebraic topology associate an algebraic structure to each topological space, but they also associate to a continuous map from X to Y an algebraic map (a homeomorphism) from the structure associated to X to that associated to Y. For example, for any map $f: X \to Y$ there is an induced homeomorphism $f_*: H_n(X) \to H_n(Y)$ called the map induced on homology. Such invariants are called homotopy invariants if any time f is homotopic to g, the algebraic maps induced by f and g on the algebraic structures are equal. Both homology and the fundamental group are homotopy invariants.

One application of these ideas is the Lefschetz fixed-point formula. For any space X that belongs to a large class of compact spaces (including all manifolds, as defined below), any map $f: X \to X$ has a fixed point if the Lefschetz number $\lambda(f)$ defined by Eq. (7) is nonzero. Here, $T_i(f)$ is the trace of the map $f_*: H_i(X,\mathbf{Q}) \to H_i(X,\mathbf{Q})$ on homology with rational coefficients $\mathbf{Q}$. [The trace is defined as follows. A basis is chosen for the vector space $H_i(X,\mathbf{Q})$. In this basis the linear map f^* is given by a square matrix, and its trace is the sum of the diagonal entries. The trace does not depend on the basis chosen.] The Lefschetz fixed-point formula implies that any map of the disk to itself has a fixed point and that any map of the n-sphere to itself that is homotopic to the identity has a fixed point if n is even.

$$\lambda(f) = \sum (-1)^i T_i(f) \tag{7}$$

Other basic homotopy invariants from algebraic topology are the higher homotopy groups. The set of homotopy classes of maps from S^n to a space X, where the maps are required to send the point

$(1, 0, \ldots, 0) \in S^n$ to a base point $x_0 \in X$, is denoted $\pi_n(X, x_0)$. The fact that collapsing the equator of S^n to a point yields a space homeomorphic to two n-spheres glued together at a single point leads to a law of composition on this set that makes it a group. For $n = 1$, this group is simply another presentation of the fundamental group. For $n > 1$, these groups are called the higher homotopy groups. These are homotopy invariants. One of the first and most important theorems about these groups states that the homotopy group $\pi_n(X, x_0)$ is an abelian group for any $n \geq 2$ and any space X.

It is natural to try to compute the homotopy groups of spheres. The following results hold for a given point x_0 in S^n:

1. $\pi_k(S^n, x_0) = 0$ if $k < n$.
2. $\pi_n(S^n, x_0)$ is isomorphic to the group of integers.
3. $\pi_k(S^1, x_0) = 0$ for all $k \neq 1$.

These results show a close relationship between homology and homotopy. Surprisingly, it was shown that $\pi_3(S^2, x_0)$ is nontrivial; in fact it is isomorphic to the group of integers. This result established the computation of the homotopy groups of spheres as a major area of homotopy theory. Another important result, obtained in 1950, states that $\pi_k(S^n, x_0)$ is a finite group unless $k = n$ or n is even and $k = 2n - 1$; in the exceptional cases, $\pi_k(S^n, x_0)$ is isomorphic to the group of integers plus a finite group. It is also known that for every $n \geq 2$ the sphere S^n has infinitely many nonzero homotopy groups. While there are many other partial results in determining the homotopy groups of spheres, a solution of this problem still seems very distant.

Manifold theory. The spaces that arise most often outside of topology are manifolds. They have always been a focus of much of the work in topology. A topological n-dimensional manifold (or topological n-manifold) is a Hausdorff topological space that is a union of countably many open sets $\{U_i\}$, each of which is homeomorphic to an open subset of $\mathbf{R}^n$. The homeomorphisms $\varphi_i: U_i \to V_i \subset \mathbf{R}^n$ give local coordinates $(x_1, \ldots, x_n)$ on U_i. Thus, a manifold is covered by open sets that have local coordinates like the coordinates of $\mathbf{R}^n$. These pairs (U_i, φ_i) are called local coordinate charts. An equivalence between topological manifolds is homeomorphism.

Examples of manifolds. Euclidean space $\mathbf{R}^n$ is of course an n-manifold. The 2-sphere $S^2 \subset \mathbf{R}^3$ can be shown to be a 2-manifold, or surface, as follows: A stereographic projection from the north pole x of the sphere to the plane tangent to the south pole gives a homeomorphism from $S^2 - \{x\}$ to $\mathbf{R}^2$ (**Fig. 2**). Similarly, a stereographic projection from the south pole y gives a homeomorphism from $S^2 - \{y\}$ to $\mathbf{R}^2$. Thus, S^2 can be covered with two open sets, each of which is homeomorphic to $\mathbf{R}^2$. This proves that S^2 is a 2-manifold. Similarly, stereographic projections in higher dimensions can be used to show that S^n is an n-manifold.

The surface of a doughnut is an example of a 2-manifold; it is called the torus. More generally, the k-holed torus, also called the surface of genus k, is a 2-manifold (**Fig. 3**).

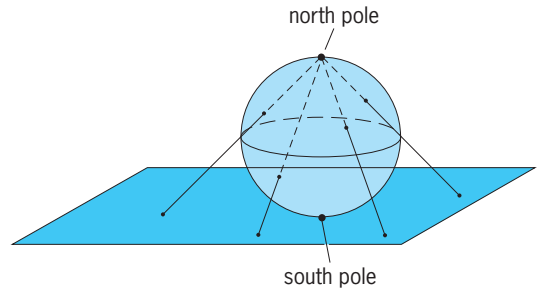

Fig. 2. Stereographic projection of a sphere from the north pole to the plane tangent to the south pole.

Smooth manifolds. The manifolds that arise from geometry and analysis carry an extra structure called a smooth (or differential) structure. A smooth structure on a topological manifold M is a collection of local coordinate charts, $\varphi_i: U_i \to V_i$, such that for every i and j the overlap function, which is the composition (8), is a smooth map (that is, an infinitely

$$\varphi_i(U_i \cap U_j) \xrightarrow{\varphi_i^{-1}} U_i \cap U_j \xrightarrow{\varphi_j} \varphi_j(U_i \cap U_j) \quad (8)$$

differentiable map). Technically speaking, two such systems of smoothly overlapping coordinate systems define the same smooth structure if their union is a smooth structure. A smooth manifold is a topological manifold with a smooth structure. The coordinates that were given above on S^2 have smooth overlap, so that S^2 has a smooth structure.

If M is a smooth manifold, then there is a well-defined notion of a smooth real-valued function on M, and of differentiation of such smooth functions. More generally, there is a natural notion of a smooth (or infinitely differentiable) map between smooth manifolds. A diffeomorphism or isomorphism between smooth manifolds is a smooth homeomorphism whose inverse is also a smooth map. Differential topology is the study of smooth structures on manifolds and smooth maps between smooth manifolds. In differential topology two manifolds are considered equivalent if there is a diffeomorphism between them.

Homology groups of manifolds. Some of the first results in algebraic topology concern the homology groups of compact manifolds. Poincaré proved a fundamental result, called Poincaré duality, which says, roughly, that the homology groups of a compact n-manifold in dimensions greater than $n/2$ can be computed from the groups in dimensions less than or equal to $n/2$.

The homology groups of compact surfaces have been completely determined: The first homology of the surface of genus g is a free abelian group of rank $2g$. In particular, the information about the genus

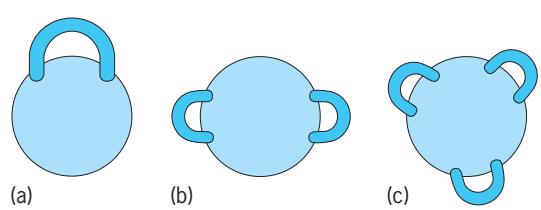

Fig. 3. Examples of 2-manifolds. (*a*) Surface of genus 1. (*b*) Surface of genus 2. (*c*) Surface of genus 3.

(or number of handles) can be recovered from the homology.

De Rham's theorem computes the homology groups of a smooth manifold from the calculus of the manifold. Many consequences flow from this result, implying that certain analytic or geometric quantities can be viewed as homology classes.

Classification of manifolds. One of the main problems in manifold theory is the classification of all manifolds (or smooth manifolds) up to homeomorphism (or diffeomorphism). There is also the auxiliary question of the difference (if any) between the two classification schemes; that is, whether every topological manifold admits a smooth structure, and, if so, how many nondiffeomorphic structures.

Results for surfaces were obtained during the nineteenth century. They can be formulated as follows: (1) Every topological surface admits a smooth structure and only one up to diffeomorphism; and (2) if S and S' are compact surfaces, then S and S' are homeomorphic if and only if their homology groups are isomorphic.

Based on this result, Poincaré asked whether or not the homology of a compact 3-manifold determined it up to homeomorphism, in particular, in the case of the 3-sphere. He answered this question in the negative by defining a new invariant, the fundamental group, and finding a 3-manifold Σ with the homology of S^3 but with a nontrivial fundamental group. Since the fundamental group of S^3 is the trivial group, Σ is not homeomorphic to S^3.

Then Poincaré refined his question to what is known as the Poincaré conjecture: if a 3-manifold has the homology of the 3-sphere and is simply connected, then is it homeomorphic to S^3? Despite much effort, this problem is unresolved.

S. Smale realized that higher-dimensional manifolds could be easier to study than 3-manifolds, and in 1961 he answered in the affirmative the generalization of Poincaré conjecture in all dimensions greater than or equal to 5. This was followed by considerable activity, starting with the classification of smooth structures on S^n for $n \geq 5$. While it has long been known that a complete classification is logically impossible in dimensions at least 4, by 1975 the classification for manifolds of dimension at least 5 had been reduced to problems in algebra and homotopy theory.

There are strong results about when a topological manifold admits a smooth structure. The answer is "not always," but the obstructions are well understood and are homotopy-theoretic in nature. A compact manifold of dimension at least 5 has only finitely many distinct smooth structures, and an upper bound for the number is given in terms of the algebraic topological invariants of the manifold.

By 1975 it appeared that the techniques applied in the higher dimensions could not be made to work in dimensions 3 and 4 because there was not enough room to maneuver. W. Thurston then introduced a vast program to study 3-manifolds. His idea was to cut up any 3-manifold into pieces that have especially nice metrics, with eight possible types. The most prevalent and interesting is the hyperbolic metric. Thurston's conjecture about the existence of these nice metrics vastly generalizes the Poincaré conjecture, which concerns the case of spherical metrics. Thurston has established his conjecture in many cases (but not, unfortunately, in the case of simply connected manifolds), lending much evidence that it is true.

In 1980 there were two major advances concerning 4-dimensional manifolds. M. Freedman showed that the high-dimensional techniques can be made to work for simply connected topological 4-manifolds. As a result, he obtained a complete classification of these manifolds, including an affirmative answer to the 4-dimensional version of Poincaré's conjecture. This completes the answer to Poincaré's conjecture in all dimensions except the one in which he originally made it.

S. Donaldson defined new algebraic invariants for smooth 4-manifolds and used them to show that some of the topological manifolds that Freedman had constructed did not admit smooth structures. He also used the invariants to show that certain well-known smooth manifolds, which were homeomorphic by Freedman's result, were not diffeomorphic. It is now known that many topological 4-manifolds admit infinitely many distinct smooth structures, whereas in all other dimensions the corresponding number is finite. In particular, $\mathbf{R}^4$ admits uncountably many distinct smooth structures, where in all other dimensions $\mathbf{R}^n$ has a unique smooth structure.

Donaldson's invariants are constructed using the moduli space of classical solutions (solutions to the Euler-Lagrange equations) of an action functional for a gauge theory. This action functional plays an important role in high-energy theoretical (quantum) physics. Other action functions and the properties of the quantum physical systems they describe can also be used to define manifold invariants, especially for three-manifolds and four-manifolds. Examples are the Seiberg-Witten invariants of four-manifolds (which are closely related to Donaldson's invariants), Chern-Simons invariants for three-manifolds, and the so-called quantum cohomology of a symplectic manifold. The deep interplay between mathematics and theoretical physics (particularly quantum field theory and string theory) is one of the primary sources of new insights in geometry and manifold theory. *See* ACTION; GAUGE THEORY; QUANTUM FIELD THEORY; SUPERSTRING THEORY; VARIATIONAL METHODS (PHYSICS).

John W. Morgan

Bibliography. G. L. Cain, *An Introduction to General Topology*, 1994; S. Eilenberg and N. Steenrod, *Foundations of Algebraic Topology*, 1952; D. S. Freed and K. K. Uhlenbeck, *Instantons and Four-Manifolds*, 2d ed., 1991; M. J. Greenberg and J. R. Harper, *Algebraic Topology: A First Course*, 1981; M. W. Hirsch, *Differential Topology*, 1976, reprint 1988; J. G. Hocking and G. S. Young, *Topology*, 1961, reprint 1988; R. C. Kirby and L. C. Siebenmann, *Foundational Essays on Topological Manifolds, Smoothings and Triangulations*, 1977; S. Lefschetz, *Topology*, 2d ed., 1990; E. Spanier, *Algebraic Topology*, 1966, reprint 1995.

Torbanite

A variety of coal that resembles a carbonaceous shale in outward appearance. It is fine-grained, brown to black, and tough, and breaks with a conchoidal or subconchoidal fracture. The name torbanite is derived from the initial discovery site of the material in 1850 at Torbane Hill, Linlithgowshire, Scotland. Torbanite is synonymous with boghead coal and is related to cannel coal. It is derived from colonial algae identified with the modern species of *Botryococcus braunii* Kütz and antecedent forms.

Major deposits of torbanite occur in Australia, Tasmania, New Zealand, Scotland, and South Africa. The South African deposit, which is in the Ermelo district of the Transvaal, yields 20–100 gal of oil per ton (80–400 liters per metric ton) on retorting. High-assay torbanite yields paraffinic oil, whereas low-assay material yields asphaltic oil. *See* COAL; SAPROPEL.

Irving A. Breger

Torch

A gas-mixing and burning tool that produces a hot flame for the welding or cutting of metal. The torch usually delivers acetylene and commercially pure oxygen producing a flame temperature of 5000–6000°F (2750–3300°C), sufficient to melt the metal locally. The torch thoroughly mixes the two gases and permits adjustment and regulation of the flame. Acetylene requires 2.5 times its volume of oxygen for complete combustion and, being an endothermic compound of carbon and hydrogen, can produce a higher flame temperature than other fuel gases. *See* ACETYLENE; WELDING AND CUTTING OF MATERIALS.

Torches are of two types: low-pressure and high-pressure. In a low-pressure, or injector, torch, acetylene enters a mixing chamber, where it meets a jet of high-pressure oxygen (see **illus.**). The amount of acetylene drawn into the flame is controlled by the velocity of this oxygen jet. In a high-pressure torch both gases are delivered under pressure. Heat developed at the work is controlled principally by the size of the nozzle or tip fitted to the torch. The larger the tip the greater the required gas pressure. Small flames are used with thin-gage metals; large flames are necessary for thick metal parts.

A welding torch mixes the fuel and gas internally and well ahead of the flame. For cutting, the torch delivers an additional jet of pure oxygen to the center of the flame. The oxyacetylene flame produced by the internally mixed gases raises the metal to its ignition temperature. The central oxygen jet oxidizes the metal, the oxide being blown away by the velocity of the gas jet to leave a narrow slit or kerf. In the case of iron, the oxides fuse at a lower temperature than the iron or steel so that the oxides form, melt, and blow away before the adjacent metal fuses. The temperature for the cutting action, once initiated, is maintained by the oxidization of the iron. Intricate shapes are accurately cut in low-carbon steel by torches automatically guided, such precision cutting being called flame machining.

Frank H. Rockett

Tornado

A violently rotating, tall, narrow column of air (that is, a vortex), typically about 300 ft (100 m) in diameter, that extends to the ground from a cumulonimbus cloud. The vast majority of tornadoes rotate cyclonically (counterclockwise in the Northern Hemisphere). Of all atmospheric storms, tornadoes are the most violent. *See* CLOUD; CYCLONE.

Visual appearance. Tornadoes are made visible by a generally sharp-edged, funnel-shaped cloud pendant from the cloud base, and a swirling cloud of dust and debris rising from the ground (**Fig. 1**). The funnel consists of small waterdroplets that form as moist

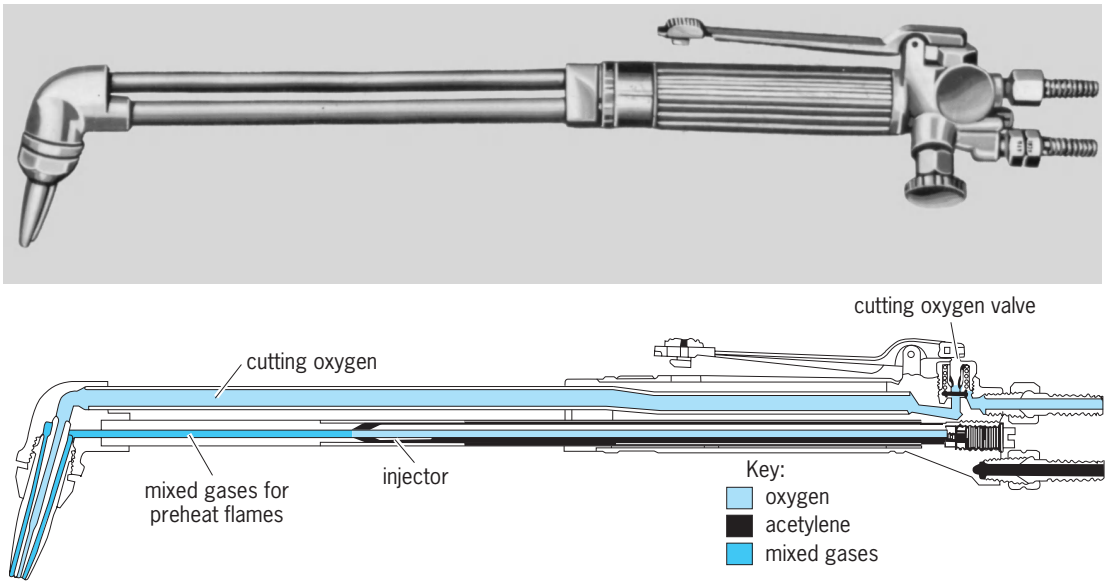

Low-pressure injector cutting torch. (*Linde Co.*)

Fig. 1. The Cordell, Oklahoma, tornado of May 22, 1981, in its decay stage. (*National Severe Storms Laboratory/ University of Mississippi Tornado Intercept Project*)

air entering the tornado's partial vacuum expands and cools. The condensation funnel may not extend all the way to the ground and may be obscured by dust. Many condensation funnels exist aloft without tangible signs that the vortex is in contact with the ground; these are known as funnel clouds. Tornado funnels assume various forms: a slender smooth rope, a cone (often truncated by the ground), a thick turbulent black cloud on the ground, or multiple funnels (vortices) that revolve around the axis of the overall tornado.

Many tornadoes evolve as follows: The tornado begins outside the precipitation region as a dust whirl on the ground and a short funnel pendant from a wall cloud on the southwest side of the thunderstorm; it intensifies as the funnel lengthens downward, and attains its greatest power as the funnel reaches its greatest width and is almost vertical; then it shrinks and becomes more tilted, and finally becomes contorted and ropelike as it decays. A downdraft and curtain of rain and large hail gradually spiral from the northeast cyclonically around the tornado, which often ends its life in rain. *See* HAIL; PRECIPITATION (METEOROLOGY); THUNDERSTORM.

Parent storm. Most tornadoes and practically all violent ones develop from a larger-scale circulation, the mesocyclone, which is 2–6 mi (3–9 km) in diameter and forms in a particularly virulent variety of thunderstorm, the supercell. The mesocyclone forms first at midaltitudes of the storm and in time develops at low levels and may extend to high altitudes as well. The tornado forms on the southwest side (Northern Hemisphere) of the storm's main updraft, close to the downdraft, after the development of the mesocyclone at low levels. Some supercells develop up to six mesocyclones and tornadoes repeatedly over great distances at roughly 45-min intervals. Tornadoes associated with supercells are generally of the stronger variety and have larger parent cyclones. Hurricanes during and after landfall may spawn numerous tornadoes from small supercells located in their rainbands. *See* HURRICANE.

Damage. Tornadoes are classified as weak, strong or violent, or from F0 to F5 on the Fujita (F) scale of damage intensity. Sixty-two percent of tornadoes are weak (F0 to F1). These tornadoes have maximum windspeeds less than about 110 mi/h (50 m/s) and inflict only minor damage, such as peeling back roofs, overturning mobile homes, and pushing cars into ditches. Thirty-six percent of tornadoes are strong (F2 to F3) with maximum windspeeds estimated to be 110–200 mi/h (50–90 m/s). Strong tornadoes extensively damage the roofs and walls of houses but leave some walls partially standing. They demolish mobile homes, and lift and throw cars. The remaining 2% are violent (F4 to F5), with windspeeds in excess of about 200 mi/h (90 m/s). They level houses to their foundations, strew heavy debris over hundreds of yards, and make missiles out of heavy objects such as roof sections, vehicles, utility poles, and large, nearly empty storage tanks. Engineers have concluded that structural damage is almost always due to wind-associated forces and missiles, not to the sudden reduction in atmospheric pressure. Typically, most debris is ejected out of a tornado along its direction of motion, so that the damage, apart from the characteristic well-defined long narrow path, appears to be from straight-line winds. However, ample signs of the circulation are usually evident in the overall debris distribution of large tornadoes. Multiple-vortex tornadoes that cross fields accumulate soil and vegetation in cycloidal rows, typically 6 in. (15 cm) high and 5 ft (1.5 m) wide, that are visible from the air. The marks are cycloidal because of the circular motion of an individual vortex about the tornado axis combined with the forward motion of the overall tornado. The rows are created by small debris being drawn into the bases of the vortices and left behind in lines of litter in the wakes of the vortices. The individual vortices occur at the locations of the strongest winds, lowest pressure, and most rapid pressure falls. Therefore, multivortex tornadoes passing through cities produce cycloidal swaths of more intense damage. *See* WIND.

Statistics. The typical tornado (described here for the Northern Hemisphere; north/south direction is the opposite for the Southern Hemisphere) is weak and short-lived (1–2 min), moves from the southwest at 30 mi/h (15 m/s), and inflicts damage to an area of about 1 mi × 150 ft (2 km × 50 m). In extreme cases, the length and width of the damage path may exceed 100 mi (150 km) and 2.5 mi (4 km), respectively; the lifetime may be over 1 h; and the translation speed may reach 68 mi/h (30 m/s). The majority of tornadoes (60%) approach from the southwest, and only 4% have a westward component of motion.

Climatology. Tornadoes occur most often at latitudes between 20° and 60°, and they are relatively

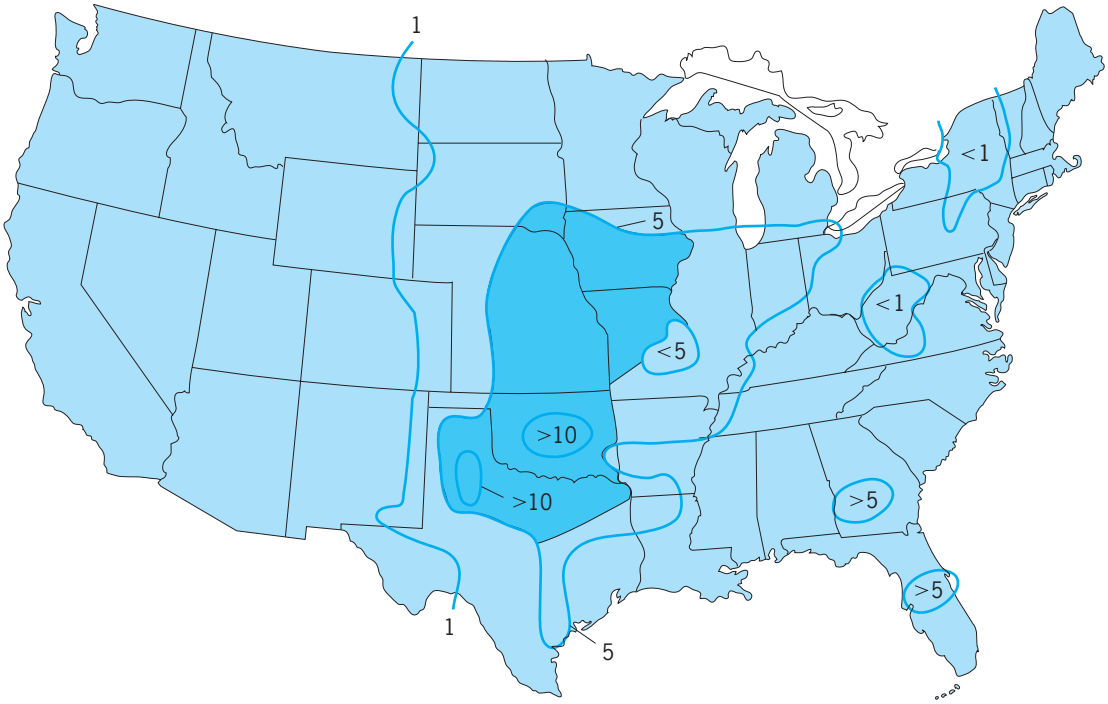

Fig. 2. Map showing average annual number of tornado occurrences per 10,000 mi^2 (26,000 km^2) from National Severe Storms Forecast Center data, 1953–1975. Shaded area shows the approximate location of so-called Tornado Alley. Contour lines indicate the number of tornadoes per 10,000 mi^2/year. West of the 1 contour line there is less than 1 per 10,000 mi^2 per year. (*After R. E. Peterson, ed., Proceedings of the Symposium on Tornadoes: Assessment of Knowledge and Implications for Man, Texas Technical University, 1976*)

frequent in the United States, Russia, Europe, Japan, India, South Africa, Argentina, New Zealand, and parts of Australia. Violent tornadoes are confined mainly to the United States, east of the Rocky Mountains, and (with less frequency) to the Bangladesh-Assam area. The world's highest frequency, 5–10 tornadoes annually per 10,000 mi^2 (26,000 km^2), occurs in the area known as Tornado Alley, which extends from Texas through Oklahoma and Kansas into Iowa (**Fig. 2**).

In winter in the United States, tornadoes are confined to the southern states, and generally have longer tracks because of stronger steering currents. The tornado belt generally shifts northward with the jet stream and the advance of tropical air through the central states and into the northern states during spring and summer. May is the most active month with 20% of the annual number of tornadoes; January is the least active with only 3%. Tornadoes occur at all hours but are most frequent between 1500 and 1900 local standard time because solar heating of the Earth's surface influences thunderstorm initiation.

Atmospheric conditions. Essentially, there are five atmospheric conditions that set the stage for widespread tornado development: (1) a surface-based layer, at least 3000 ft (1 km) deep, of warm, moist air, overlain by dry air at midlevels; (2) an inversion separating the two layers, preventing deep convection until the potential for explosive overturning is established; (3) rapid decrease of temperature with height above the inversion; (4) a combination of mechanisms, such as surface heating and lifting of the air mass by a front or upper-level disturbance, to eliminate the inversion locally; (5) pronounced vertical wind shear (variation of the horizontal wind with height). Specifically, storm-relative winds in the lowest 6000 ft (2 km) should exceed 20 knots (10 m/s) and veer (turn anticyclonically) with height at a rate of more than 10°/1000 ft (30°/km). (Storm-relative wind is the wind measured with respect to a coordinate system that moves with the storm.) Such conditions are prevalent in the vicinity of the jet stream and the low-level jet. Incidentally, the ground-relative winds often are quite similar to the storm-relative winds in strength and directional turning.

The first three conditions above indicate that the atmosphere is in a highly metastable state. There is a strong potential for thunderstorms with intense updrafts and downdrafts. The fourth condition is the existence of a trigger to release the instability and initiate the thunderstorms. The fifth is the ingredient for updraft rotation. When conditions warrant, the Storm Prediction Center issues tornado watches about 1 h in advance for rectangular areas of roughly 25,000 mi^2 (65,000 km^2) that indicate that the potential for tornadoes exists in the area for a 6-h period. *See* AIR MASS; FRONT; JET STREAM; TEMPERATURE INVERSION; THUNDERSTORM.

The reason that the world's worst tornadoes generally occur in the central United States and in Bangladesh-Assam is that both regions are located in the midlatitude belt of westerlies, with warm oceans equatorward and mountains to their west. During the spring, both regions are often in the transition zone between warm and cold air masses, so that they lie beneath the jet stream. Cyclones and upper-level disturbances generally develop in the vicinity of the jet stream and migrate along it. Ahead of cyclones

approaching from the west, warm, very moist air flows off the ocean into the region at low levels. At midlevels the air has been dried out by passage over the mountains to the west. Furthermore, the geography of both regions favors the development of low-level jets and associated strong low-level vertical wind shear.

Detection. Microbarographs, radars, satellites, and sferic detectors have all been used to identify potentially tornadic storms. Tornado warnings of about 1 h duration are issued by local weather service offices for specific countries on the basis of radar observations and reports from trained spotters; such warnings have contributed to a 50% decline in the annual numbers of deaths since 1950. At night, flashes from downed power lines may reveal the existence of a tornado. Conventional warning radars measure only the spatial distribution of the reflectivity of precipitation. Rotation is inferred by a hook-shaped appendage on the southwest side of the radar echo, observed sometimes at close range. The best detection device is a Doppler radar which measures the velocity components along the radar beam of raindrops and hailstones in addition to the reflectivity. Since 1988, a warning network of approximately 150 Doppler radars has been installed across the United States. Two types of vortex signature appear in Doppler velocity fields, the mesocyclone signature and the tornadic vortex signature (TVS). The TVS is the Doppler velocity field of the flow near the tornado, poorly resolved because the radar beam is broader than the tornado. It is generally seen only at close range because the width of the radar beam and the elevation of the radar horizon above the ground increase with range. At long range, "radar-indicated" tornadoes are based on detection of the much wider mesocyclone. For large tornadoes with signatures that first appear in the parent cloud, both signatures may be detected aloft 20 min or longer prior to touchdown, allowing warnings to be issued with considerable lead time. The TVS extends upward as well as downward, reaching heights of 40,000 ft (13 km) in rare cases. However, 50% of TVSs form as a column from the ground to more than 1 mi (2 km) and extend upward in time. Tornadoes associated with this type of signature form rapidly in 5–10 min and may occur with little advance warning. Untested detection instruments include an infrasonics sensor and seismic tornado detector. Home barometric devices based on detecting rapid pressure falls are inadequate because the warning may precede the tornado strike by only seconds. *See* DOPPLER RADAR; RADAR METEOROLOGY; SATELLITE METEOROLOGY; SFERICS; STORM DETECTION.

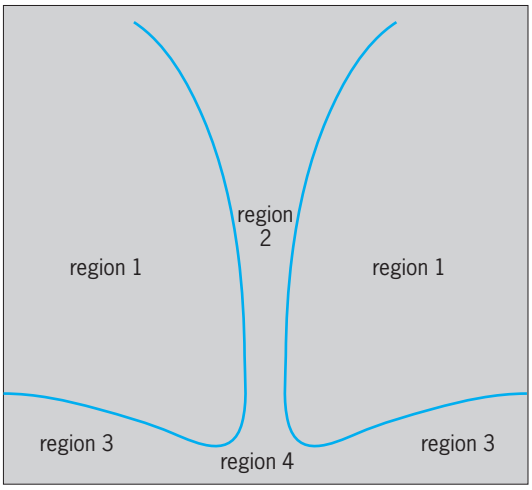

Fig. 3. Different flow regions of a tornado. (*After E. Kessler, ed., Thunderstorm Morphology and Dynamics, 2d ed., University of Oklahoma Press, 1986*)

Measurements. Winds and pressure deficits in tornadoes are not well known because tornadoes are difficult to intercept and extremely dangerous. Various estimates have been obtained from photogrammetric analyses of tornado movies, engineering analyses of structural failures and missiles generated by the vortex, analyses of the cycloidal marks left by multivortex tornadoes, chance readings from anemometers and microbarographs (when not destroyed) and attempts to introduce probes directly, Doppler radar measurements, and analysis of the size and shape of the funnel cloud. Portable Doppler radars driven to within a mile or two of large, violent tornadoes have recorded wind speeds of up to 280 mi/h (125 m/s). As yet, there is no evidence of wind speeds in excess of 300 mi/h (130 m/s). Substantial vertical velocities, up to 180 mi/h (80 m/s), have been documented. Radial inflow velocities may reach 110 mi/h (50 m/s) in narrow bands close to the ground. A surface pressure drop of 55 millibars (5.5 kilopascals) was recorded on a sensor placed in the path of a large, violent, multivortex tornado in 1995. This is probably not the maximum pressure deficit of the tornado because the measurement was neither in the center of the tornado nor in a satellite vortex. The temperature (also recorded) dropped a few degrees Celsius during the tornado owing to the low pressure. Temperature above the ground inside a tornado is unknown. *See* AIR PRESSURE.

Origins of rotation. The basic measure of local spin in a fluid is the vorticity vector, which is the curl or rotation of the wind vector and is aligned parallel to the spin axis in the direction of advance of a right-hand screw. The vorticity at a specified point and time is equal to twice the angular velocity of a minute spherical parcel of air that is centered on the point at that time. The strong vertical wind shear generally present prior to development of severe storms is associated with horizontal vorticity that is up to 100 times greater than the background vertical vorticity associated with the Earth's rotation and large-scale cyclones. Theory suggests that a thunderstorm updraft could develop a tornado directly from the background vertical vorticity in a few hours. However, tornadoes can be generated more rapidly from horizontal vorticity. Furthermore, three-dimensional computer models generate storms with rotating updrafts in suitably sheared environments with the Earth's rotation "switched off." Thus, it is widely believed that the mesocyclone forms as a result of the updraft tipping preexisting horizontal

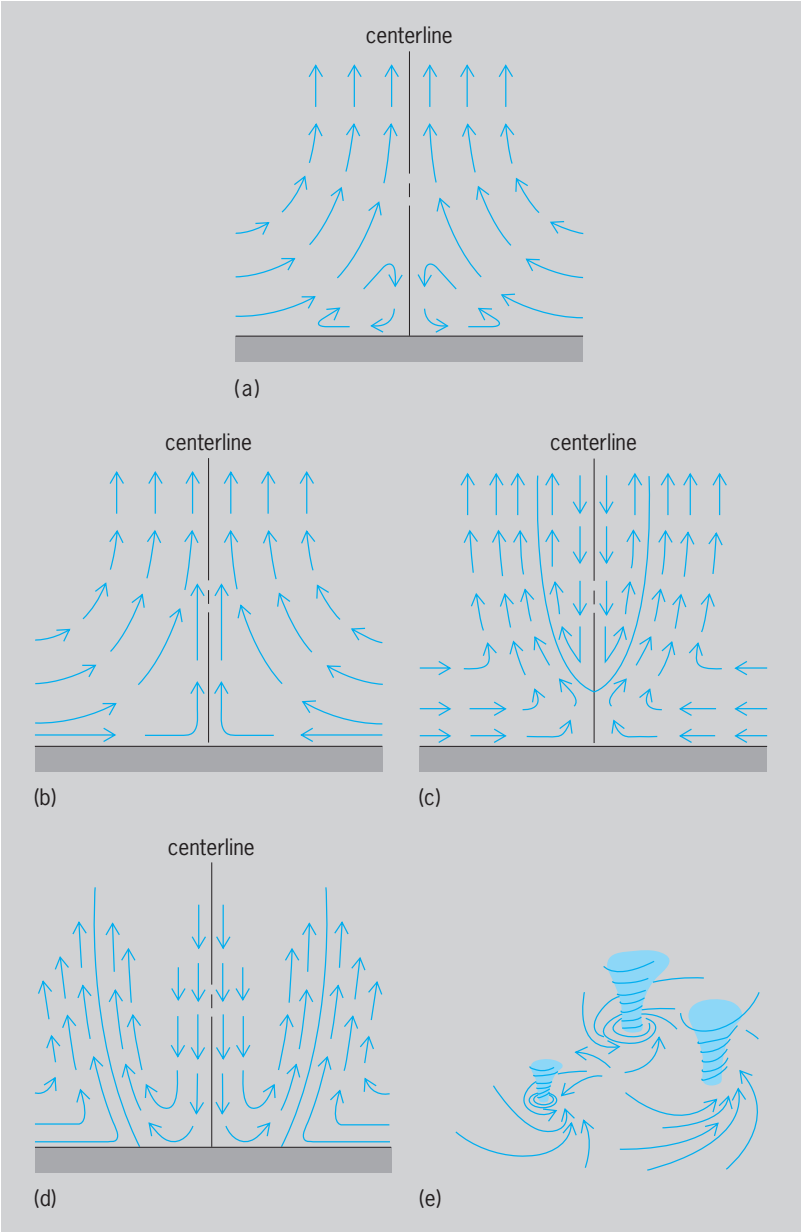

Fig. 4. Effect of increasing swirl ratio on the tornado flow. (*a*) Weak swirl—no tornado as flow in boundary layer separates and passes around corner region. (*b*) One-cell tornado (axial updraft). (*c*) Vortex breakdown (axial downdraft aloft does not reach ground). (*d*) Two-cell tornado with downdraft impinging on ground (core radius increases rapidly with increasing swirl ratio). (*e*) Multiple vortices (two to six with increasing swirl ratio). Only the flow in a radial-height cross section is depicted in *a–d*. (*After E. Kessler, ed., Thunderstorm Morphology and Dynamics, 2d ed., University of Oklahoma Press, 1986*)

vorticity toward the vertical. The updraft rotates as a whole when the vorticity is streamwise (that is, the spin axis is along the flow direction) relative to the storm. At this stage there is little rotation at very low levels because the vertical vorticity is being produced in rising air. *See* THUNDERSTORM; VORTEX.

Generation of rotation near the ground is a more complicated process. It does not happen in storm simulations if the evaporation of precipitation is switched off. In simulated storms, it originates in an initially dry airstream that flows through a precipitation region at the rear of the rotating updraft, is chilled by evaporating rain, subsides, and flows along the ground. The left side of this airstream is drawn up into the storm's main updraft, and a wall cloud (markedly lowered cloud base) forms in the process because the cool air is nearly saturated. The rotation is generated around the periphery of this downdraft in the form of a toroidal circulation (the coldest air in the center of the downdraft sinks faster than surrounding air). Each vortex ring is elongated in the downstream direction owing to the transport of vorticity by the horizontal wind, and tilted by the gradient of vertical wind since its back edge remains in the generation region in the downdraft and its leading edge moves downstream out of the downdraft. The vertical vorticity produced by the tilting is cyclonic on the left side of the airstream, and anticyclonic on the right side. The cyclonic side is drawn into the storm's rotating updraft, thus connecting the mesocyclonic vortex already present in the updraft to the ground. Horizontal contraction and vertical stretching of the lower part of this vortex by the updraft causes it to spin faster. Near the ground this spin-up is enhanced by radial inflow induced by friction, the same effect that causes accumulation of the tea leaves at the center of a stirred cup of tea. A cyclonic tornado may form simply as the result of continued uninterrupted contraction of this column for around 15 min. The process may be aborted prematurely by the lower part of the vortex being overtaken by cold sinking air, the same process that seems to cause tornadoes to decay. The anticyclonic side of the airstream is further away from the storm's updrafts and seldom is spun up into an anticyclonic tornado. Detailed analysis of data collected in field experiments such as VORTEX (Verification of the Origins of Rotation in Tornadoes Experiment), conducted in 1994–1995, which utilize exciting new mobile instrumentation, should provide information on tornado genesis.

A different mechanism applies to small tornadoes and waterspouts that develop in environments with insignificant vertical shear. These tend to develop along wind-shift lines, which become dynamically unstable and roll up into rows of vertical vortices, about 0.6 mi (1 km) in diameter. If one of the vortices happens to be located under the updraft of a cumulus congestus or cumulonimbus cloud, then a tornado may form as a result of the vertical stretching effect. These tornadoes develop when the temperature at low levels decreases strongly with height, and, unlike supercell tornadoes, occur early in the life of the parent cloud.

Energy source. The rate at which kinetic energy is produced in a tornado is estimated to be 10^3 MW in order of magnitude. To explain the high energy density of tornadoes, various exotic energy sources have been postulated in the past, including electrical heating from "rapid-fire" lightning discharges in a confined area. However, it is now believed that the tornado's energy is derived mostly from the buoyant potential energy that is released in the parent cloud by condensing water vapor.

Flow structure. Theory, laboratory simulations, computer models, and analysis of films reveal four basic flow regions in tornadoes (**Fig. 3**). The first (region 1) is the outer flow away from the ground

and the axis of rotation. This region consists of horizontally converging, rising air that approximately conserves its angular momentum. Consequently, the air spins faster as it approaches the axis. The core is region 2. It surrounds the axis, extending outward roughly to the radius of maximum winds. The core is approximately in solid-body rotation. The core flow is generally stable against radial displacements; and so it supports centrifugal waves, which are often seen moving up or down the funnel. This stability also prevents air from region 1 from being entrained into the core. The flow along the axis may be either upward or downward, and there may be stagnation points aloft separating axial upflow and downflow. In large tornadoes, the axial flow is probably downward all the way to the ground. Interaction with the ground creates a turbulent boundary layer, perhaps a few hundred meters deep, of reduced tangential velocities. In this third region, the outward centrifugal forces no longer can balance the inward pressure gradient forces (as they do in region 1). The resulting net inward force drives a strong radial inflow into region 4, the corner region at the foot of the vortex, where the strongest winds are located. Here, conservation of mass dictates a strong updraft into the core, either along the axis or surrounding a central downdraft. In the corner region, inflowing parcels overshoot the radius that they would reach in the absence of friction. Closer penetration to the axis would cause parcels to rotate faster if angular momentum losses were small enough. Conversely, loss of angular momentum would lead to slower rotation if the overshoot were negligible. The former happens to be the larger effect for all parcels except for those in the lowest few meters. Paradoxically friction causes more intense tornadoes!

Laboratory simulations reveal that the structure of turbulent vortices depends largely on a single flow parameter, the swirl ratio, which is roughly the ratio of the tangential and vertical velocities in the parent rotating updraft. As the swirl ratio increases, the core widens, and the flow changes (**Fig. 4**). At a moderate swirl ratio (Fig. 4*c*), the vortex breaks down from a high-speed rotating jet (with the wind more upward than tangential) to a slower broader vortex with downflow along the axis. The high wind speeds in the jet are associated dynamically with large pressure deficit on the axis. Below the breakdown point the flow is faster than the propagation speed of centrifugal waves so the waves are unable to propagate downward and modify the flow to a slower state. Tornadoes with this structure are narrow but strong, and have the capability of lofting heavy objects such as pickup trucks 100 ft (30 m) in the air. At large swirl ratio (Fig. 4*d*), the downflow along the axis reaches the surface, and parcels flowing into the tornado along the ground do not come close to the axis before rising. This type of tornado is often violent owing to large ambient angular momentum and low central pressure associated with a warm, light core, which arises from air descending along the axis and warming adiabatically. At still larger swirl ratios, transitions occur, first from a single vortex to two, then to three (Fig. 4*e*), up to six vortices. Extreme winds occur in these secondary vortices. *See* VORTEX.

Robert Davies-Jones

Bibliography. C. Church et al. (eds.), *The Tornado: Its Structure, Dynamics, Prediction, and Hazards*, 1993; R. Davies-Jones, Tornado formation and structure, *Sci. Amer.*, 1995; R. Davies-Jones, Tornadoes, *Sci. Amer.*, 273:34–41, 1995; E. Kessler (ed.), *Thunderstorm Morphology and Dynamics*, 2d ed., 1986; R. E. Peterson (ed.), *Proceedings of the Symposium on Tornadoes: Assessment of Knowledge and Implications for Man*, Texas Technical University, 1976; J. T. Snow, The tornado, *Sci. Amer.*, 250:86–97, 1984.

Torpediniformes

An order of batoid fishes occurring in the subclass Elasmobranchii and known as the torpedo electric rays and numbfishes. Typical members of Torpediniformes are identifiable by their flat, pancakelike body sector (disc); relatively robust tail sector; smooth skin; small or obsolete eyes; distinct caudal fin; and ovoviviparous development (that is, producing eggs that develop internally and hatch before or soon after extrusion). Of special interest is a pair of enlarged electric organs located on the disc lateral to the gill slits. These kidney-shaped organs can best be seen from the ventral side, although the columnar structures that compose them occupy the thickness of the disc. These electric organs, which may constitute 17% of the total body weight, deliver shocks up to 220 volts. The voltage depends on the species, its size, and physical condition. The electric organs are thought to be used primarily for feeding and defense. Electric rays are poor swimmers, depending primarily on the tail because the disc is rather inflexible and of little use in locomotion, although it is flexible enough for capturing prey. They spend most of their time partially buried in sand or mud and feed on a variety of invertebrates, including crustaceans, mollusks, and worms, as well as small fishes. Members of the genus *Torpedo* are reported to reach a length of 1.8 m (6 ft) and a weight of 44 kg (100 lb). Torpediniforms occur in intertidal waters to deep waters [1070 m or 3510 ft in the case of blind species (*Benthobatis*)] in temperate to tropical zones of all oceans.

The order comprises two families: Torpedinidae and Narcinidae.

Torpedinidae (torpedo electric rays). The disc is truncate anteriorly or emarginate; the rostrum is reduced; the jaws are extremely slender and lack labial cartilages. There are two genera, *Torpedo* (21 species) and *Hypnos* (1 species). *Torpedo* differs from *Hypnos* in having well-developed dorsal and caudal fins vs. very small dorsal and caudal fins, and in having worldwide distribution vs. distribution limited to off Australia.

Narcinidae (electric rays or numbfishes). The disc is rounded anteriorly; a rostrum is present; the jaws are stout and the labial cartilages strong. There are nine genera and 37 species in two subfamilies, Narcininae (numbfishes) with four genera and 26 species and Narkinae (sleeper rays) with five genera and

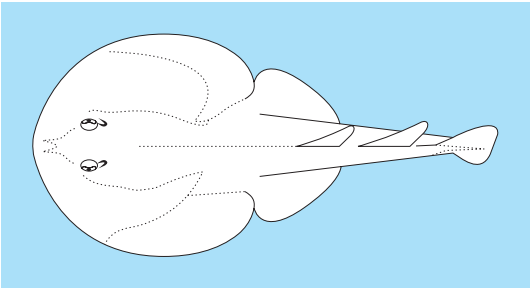

Example of a Narcinidae species. (*Courtesy of J. S. Nelson, Fishes of the World, 4th ed., Wiley, 2006*)

11 species. Narcininae (see **illustration**) differs from Narkinae in having long, strongly protractile jaws vs. short, weakly protractile jaws; a deep groove around the mouth and lips vs. shallow grooves; a broad rostrum vs. narrow; usually two dorsal fins vs. usually one dorsal fin; and species occur in Atlantic, Indian, and Pacific oceans vs. limited to Indo-West Pacific. *See* BATOIDEA; ELASMOBRANCHII; ELECTRIC ORGAN (BIOLOGY); PRISTIFORMES; RAJIFORMES; MYLIOBATIFORMES. Herbert Boschung

Bibliography. H. B. Bigelow and W. C. Schroeder, in J. Tee-Van et al. (eds.), *Fishes of the Western North Atlantic*, Sears Found. Mar. Res. Mem. 1, pt. 2, 1954; L. J. V. Compagno, Checklist of Chondrichthyes, pp. 503–547 in W. C. Hamlett (ed.), *Reproductive Biology and Phylogeny of Chondrichthyes: Sharks, Batoids and Chimaeras*, Science Publishers, Enfield, NH, 2005; J. D. McEachran and N. Aschliman, Phylogeny of Batoidea, pp. 79–113 in J. C. Carrier, J. A. Musick, and M. R. Heithaus (eds.), *Biology of Sharks and Their Relatives*, CRC Press, Boca Raton, FL, 2004; J. D. McEachran and M. R. de Carvalho, Batoid fishes, pp. 507–589 in K. E. Carpenter (ed.), *The Living Marine Resources of the Western Central Atlantic: FAO Species Identification Guide for Fishery Purposes*, vol. 1, FAO, Rome, 2002; J. D. McEachran, K. A. Dunn, and T. Miyake, Interrelationships of the batoid fishes (Chondrichthyes: Batoidea), pp. 63–84 in M. L. J. Stiassny, L. R. Parenti, and G. D. Johnson (eds.), *Interrelationships of Fishes*, Academic Press, San Diego, 1996; J. S. Nelson, *Fishes of the World*, 4th ed., Wiley, New York, 2006.

Torque

The product of a force and its perpendicular distance to a point of turning; also called the moment of the force. Torque produces torsion and tends to produce rotation. Torque arises from a force or forces acting tangentially to a cylinder or from any force or force system acting about a point. A couple, consisting of two equal, parallel, and oppositely directed forces, produces a torque or moment about the central point. A prime mover such as a turbine exerts a twisting effort on its output shaft, measured as torque. In structures, torque appears as the sum of moments of torsional shear forces acting on a transverse section of a shaft or beam. *See* COUPLE; TORSION. Nelson S. Fisk

Torque converter

A device for changing the torque-speed ratio or mechanical advantage between an input shaft and an output shaft. A pair of gears is a mechanical torque converter. A hydraulic torque converter, with which this article deals, is an automatically and continuously variable torque converter, in contrast to a gear shift, whose torque ratio is changed in steps by an external control. *See* AUTOMOTIVE TRANSMISSION; MECHANICAL ADVANTAGE.

Converter characteristics. A mechanical torque converter transmits power with only incidental losses; thus, the power, which is the product of torque T and rotational speed N, at input I is substantially equal to the power at output O of a mechanical torque converter, or $T_I N_I = kT_O N_O$, where k is the efficiency of the gear train. This equal-power characteristic is in contrast to that of a fluid coupling in which input and output torques are equal during steady-state operations. *See* FLUID COUPLING.

In a hydraulic torque converter, efficiency depends intimately on the angles at which the fluid enters and leaves the blades of the several parts. Because these angles change appreciably over the operating range, k varies, being by definition zero when the output is stalled, although output torque at stall may be three times engine torque for a single-stage converter and five times engine torque for a three-stage converter. Depending on its input absorption characteristics, the hydraulic torque converter tends to pull down the engine speed toward the speed at which the engine develops maximum torque when the load pulls down the converter output speed toward stall.

Converter power efficiency is highest (80–90%) at a design speed, usually 40–80% of maximum engine speed, and falls toward zero as shaft speed approaches engine speed. Because of this characteristic, the mode of operation may be modified to change from torque conversion to simple fluid coupling or to direct mechanical drive at high speed.

Hydraulic action. These characteristics are achieved by the exchange of momentum between the solid parts of the converter and the fluid (**Fig. 1**). A vaned impeller on the input shaft pumps the fluid from near the axis of rotation to the outer rim. Fluid momentum increases because of the greater

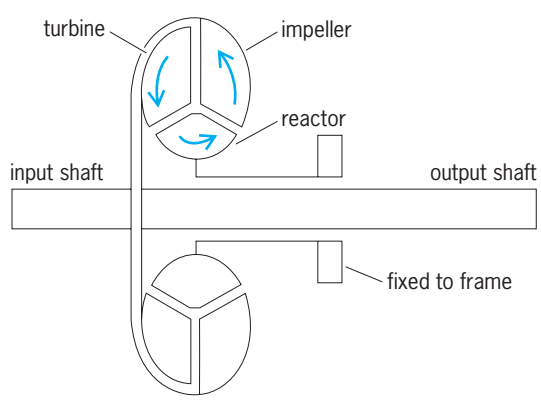

Fig. 1. Elementary hydraulic torque converter.

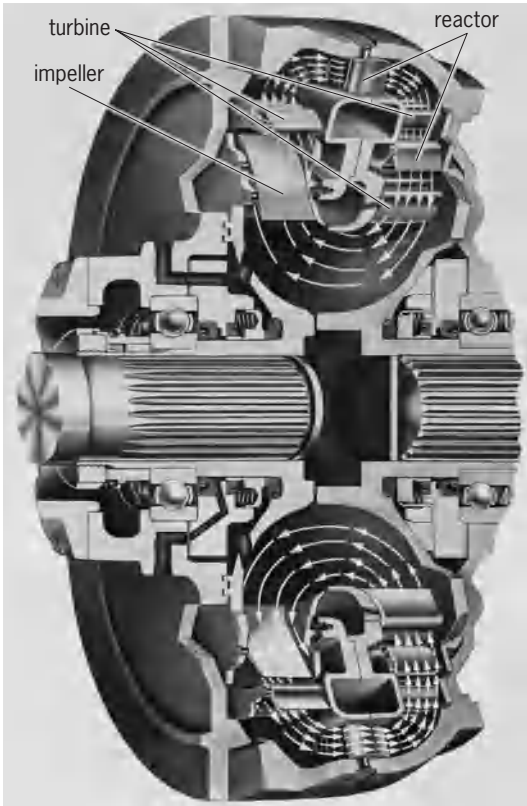

Fig. 2. Three-stage converter showing simplified fluid flow around torus. (*Twin Disc, Inc.*)

radius and the influence of the vanes. The high-energy fluid leaves the impeller and impinges on the blades of a turbine, giving up its energy to drive the turbine, which is connected to the output shaft. The fluid discharges from the turbine into a bladed reactor. The reactor blades are fixed to the frame; they deflect the fluid flow and redirect it into the impeller. This change in flow direction produced by the stationary reactor is equivalent to an increasing change in momentum which adds to the momentum imparted by the impeller to give a torque increase at the output of the converter (**Fig. 2**). *See* HYDRAULICS.

In a typical converter, as the output shaft comes up to the speed of the input shaft, efficiency decreases. Therefore, the reaction member may be mounted on a freewheel unit so that it rotates with the fluid at high speed ratio when torque multiplication is no longer possible. In addition, splitting the reaction member to give a four-element polyphase converter gives even more uniform efficiency. *See* SHAFTING.

Henry J. Wirry

Bibliography. G. Raczkowski, *Principles of Machine Dynamics*, 1979; H. F. Tucker, *Automatic Transmissions*, 1980.

Torricelli's theorem

The proposition that the speed of efflux of a liquid from an opening in a reservoir equals the speed that the liquid would acquire if allowed to fall from rest from the surface of the reservoir to the opening.

Torricelli, a student of Galileo, observed this relationship in 1643. In equation form, $v^2 = 2gh$, in which v is the speed of efflux, h the head (or elevation difference between reservoir surface and center line of opening if in a vertical plane), and g the acceleration due to gravity. (The equation is the same as that for a solid particle dropped a distance h in a vacuum.) The relationship can be derived from the energy equation for flow along a streamline, if energy losses are neglected.

An orifice (opening in the wall or bottom of a reservoir) is used as a flow-measuring device. From Torricelli's theorem, by solving for v and multiplying by the flow area, an expression for discharge Q, in volume per unit time, is obtained. In equation form, $Q = C_d A\sqrt{2gh}$, in which A is the area of opening and C_d is a dimensionless coefficient, determined experimentally, that corrects for contraction of the jet as it leaves the orifice and for energy loss due to viscosity. When h is measured, Q may be determined from the above formula. *See* BERNOULLI'S THEOREM; FLOW MEASUREMENT.

Victor L. Streeter

Torsion

A straining action produced by couples that act normal to the axis of a member. Torsion is identified by a twisting deformation.

In practice, torsion is often accompanied by bending or axial thrust as in the case of line shafting driving gears or pulleys, or propeller shafts for ship propulsion. Other important examples include springs and machine mechanisms usually having circular sections, either solid or tubular. Members with noncircular sections are of interest in special applications, such as structural members subjected to unsymmetrical bending loads that twist and buckle beams.

The shear properties of materials are determined by a torsion test. *See* SHEAR.

Cylindrical bars. The twist of a bar due to torque can be visualized as the accumulated rotational displacements of imaginary disks cut by transverse sections on which tangential forces operate. Shearing forces vary across the section and together furnish the internal resisting torque.

Torsional angle, designated θ, is the total relative rotation of the ends of a straight cylindrical bar of length L, when subjected to torque (**Fig. 1**).

Helical angle, designated ϕ, is the angular displacement of a longitudinal element, originally straight on

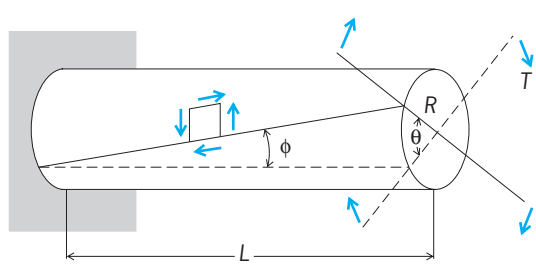

Fig. 1. Cylindrical bar in torsion.

the surface of the untwisted bar, which becomes helical after twisting (Fig. 1). Angle ϕ is the shear strain. For small twist, torsional and helical angles are related by geometry $\phi = R\theta/L$, where R is the radius of the bar.

Elastic shear stress. Within the elastic limit, shear stress S_s is found by Hooke's law, $S_s/\phi = E_s$, and is expressed in terms of the torsional angle as $S_s = (R/L)E_s\theta$, where E_s is the modulus of rigidity. *See* HOOKE'S LAW.

The shear stress varies linearly across the section, being maximum at the surface and zero at the center. For a circular section the maximum shear stress acting perpendicular to the radius at the extreme distance $R = D/2$ from the neutral axis is $S_{\max} = 16T/\pi D^3$, where T is the externally applied twisting moment.

Tangential shear stresses on the section are accompanied by longitudinal shear stresses along the bar. These complementary stresses induce tensile and compressive stresses, equal to the shear intensity, at 45° to the shear stresses. The longitudinal stresses are important in laminated materials, wood, or metals with seams. Brittle materials, low in tensile strength, fracture on a 45° helicoidal surface; ductile materials fracture on transverse sections after large twist.

Resisting torque equal to the applied torque is the moment of the internal shear forces about the neutral axis expressed in terms of the sectional dimensions and the stresses. A general expression for resisting torque is $T = S_{\max}J/R$, where J is the polar moment of inertia of the section. This relation is applicable to both solid and tubular circular sections which are differentiated by J. In terms of torque T, torsional angle θ is TL/E_sJ. Torsional angle per unit of length is a measure of torsional stiffness, which may limit the required dimensions of a shaft. In power transmission the torque associated with horsepower is found from hp $= TN/63{,}000$, where T is expressed in inch-pounds of moment and N is the rotation of the shaft in revolutions per minute.

Inelastic behavior in torsion. Strains exceeding the elastic limit are not completely recoverable after unloading and the behavior is inelastic. Torsional strains vary linearly from the center of the bar during elastic and plastic deformation, and the corresponding shear stresses reflect the stress-strain curve for the material (**Fig. 2**). After the extreme element reaches the yield point, continued twisting produces inelastic strains at increasing distances from the surface while the stress remains constant. When the action is fully plastic, the stresses are constant, equal to the yield point over the entire section. The fully plastic resisting torque is shown by Eq. (1), which is 1.33

$$T_p = \frac{4}{3}\frac{S_{\text{yp}}J}{R} \tag{1}$$

times that required to just produce surface yielding [S_{yp} = shear stress at yield point]. Torsional resistance increases because of strain hardening but is of interest only where large deformation can be tolerated. Elastic analysis is applicable to designs where

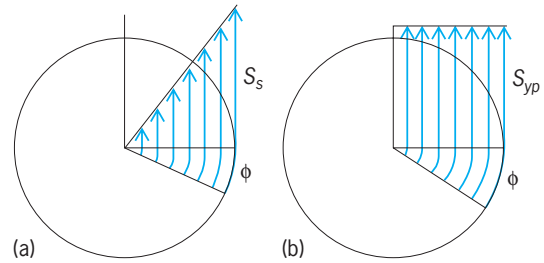

Fig. 2. Stress distribution. (*a*) Elastic torsional stress. (*b*) Fully plastic torsional stress.

permanent deformation must be avoided and where endurance (fatigue) properties limit the stresses.

Thin-walled tubes. Thin tubular members find application particularly in aircraft. Shear stresses are assumed uniform over the wall thickness when a thin-walled tube of any shape is subjected to torque at the ends. Shear force q per unit length of perimeter is constant.

Shear flow is the constant shear force q acting along the median line of the wall and is equal to the product of shear stress S times thickness t at any point; thus $q = St$ is constant. The concept of flow is drawn from the similarity of the expression for constant shear force with the constant quantity Q of a liquid passing variable sections of a channel having area A and velocity V, $Q = AV$. Resisting torque T is the summation of moments of shear forces on unit lengths ds of the wall perimeter about the center of rotation $T = 2Aq$, where A is the area enclosed by the centerline of the tube wall (**Fig. 3**). The stress at any point where thickness is t is $S = q/t = T/2At$. The torsional angle produced by applied torque T is found from Eq. (2), where S is the length of the perimeter

$$\theta = \frac{TL}{4A^2E}\int_0^S \frac{ds}{t} \tag{2}$$

and t is the variable thickness. For constant thickness, $\theta = TLS/4A_2Et$, where S is peripheral length of the centerline.

Solid noncircular sections. When a solid member with noncircular section is twisted, the sections become warped and the stresses do not vary linearly as in the case of circular sections. Evaluation of stresses and torsional twist requires the rigorous procedures

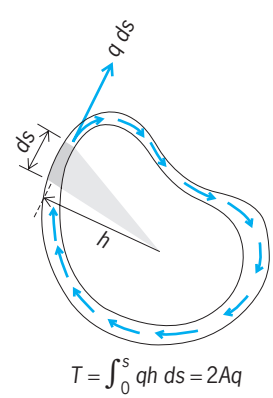

Fig. 3. Diagram showing the shear flow in a thin-walled tube.

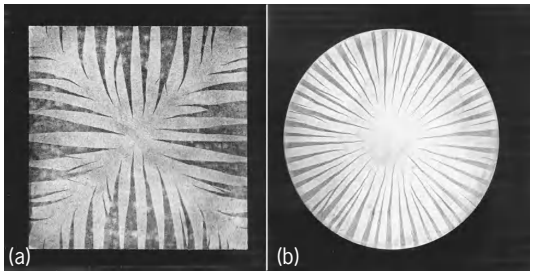

Fig. 4. Plastic strain in torsion. (*a*) Square bar. (*b*) Round bar.

of the theory of elasticity. If a grid is scribed on the surface of a square or rectangular bar and the bar is twisted, distortions of the grid indicate that maximum shear stress is at a boundary nearest the center. Contrary to theory applicable to circular sections, the stress is zero at the corners, which are the most remote elements. The location of maximum stress is indicated by points of initial plastic yielding as shown by the macrographs of a square and a round bar (**Fig. 4**). Sections were etched after yielding, thus differentiating the darker plastic zones. Formulas for maximum shear stress and torsional angle for common noncircular sections are presented in **Fig. 5**.

Helical springs subjected to axial loads involve all four possible straining actions: direct stress, transverse shear, bending, and torsional shear. For small obliquity of the coils, as in close-coiled springs, torsional shear is the most important action. When stresses and deflection are determined by formulas applicable to straight bars, a correction is necessary to account for the effect of curvature of the coils. *See* SPRING (MACHINES).

Membrane analogy. Shearing stresses in sections which cannot be conveniently analyzed mathematically are determined experimentally by membrane analogy. The analogy presented by Ludwig Prandtl (1903) is based on the similarity of the equilibrium equation for a membrane with pressure on one side and the differential equation for torsional stresses. In application, a thin membrane, such as a soap film, is placed over an opening in a plate having the same geometrical shape as the section under investigation. Slight air pressure on one side deflects the film, and micrometer measurements determine the contours of equal deflection. The slope at any point and the volume enclosed by the deflected membrane can be found from these measurements. If a bar having this section is twisted, the torsional shearing stress at any point is proportional to the slope of the membrane, the stress direction is tangent to the contour, and the torque is proportional to the volume enclosed by the deflected membrane.

The method is a valuable qualitative aid in locating points of maximum stress by visualizing or observing points of maximum slope of the deflected film. The high stress at a reentrant corner, such as at a fillet of a structural angle or channel section, is indicated by a steep slope of the film. John B. Scalzi

Bibliography. J. S. Golan, *Torsion Theories*, 1987; W. D. Pilkey, *Formulas for Stress, Strain, and Structural Matrices*, 1994.

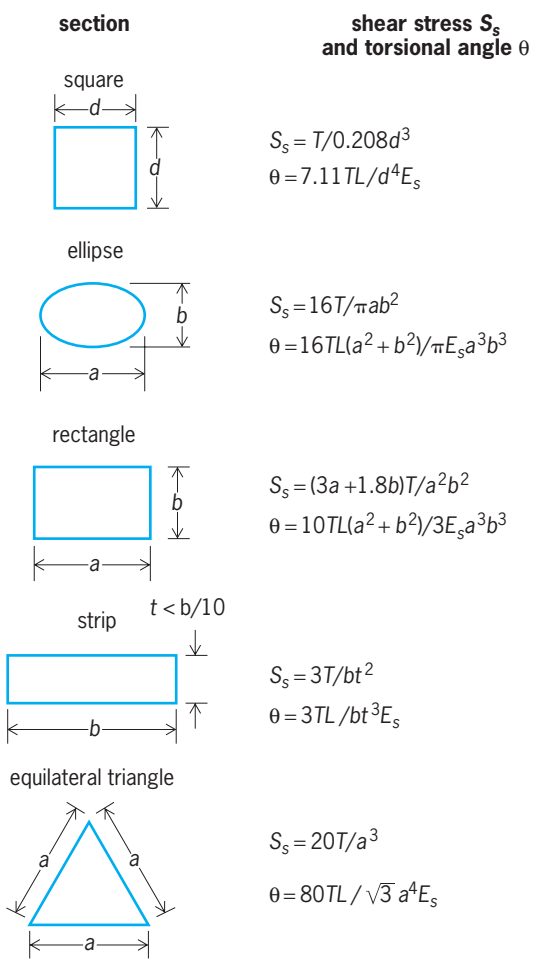

Fig. 5. Formulas for maximum shear stress and torsional angle for common noncircular sections.

Torsion bar

A spring flexed by twisting about its axis. Design of a torsion bar spring is primarily based on the relationships between the torque applied in twisting the spring, the angle through which the torsion bar twists, and the physical dimensions and material (modulus of elasticity in shear) from which the torsion bar is made. The **illustration** shows the elements of a simple torsion bar and the important dimensions involved in its design. Equation (1) relates

$$\theta = \frac{32Fal}{\pi D^4 G} \tag{1}$$

these dimensions. Here θ is angle of twist in radians, F is force in pounds, a is radius of arm of force in inches, l is length of torsion bar in inches, D is diameter of torsion bar in inches, and G is modulus of elasticity in shear in pounds per square inch. *See* TORQUE.

If the deflection or twist of the spring θ is large, force F must change direction if a is to remain constant. For this reason Eq. (1) is frequently written as

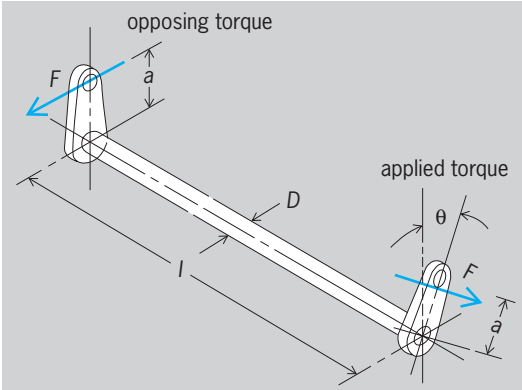

Diagram of a torsion bar.

Eq. (2), in which τ is torque in inch-pounds.

$$\theta = \frac{32\tau l}{\pi D^4 G} \tag{2}$$

Torsion bar springs are found in the spring suspension of truck and passenger car wheels, in production machines where space limitations are critical, and in high-speed mechanisms where inertia forces must be minimized. *See* SPRING (MACHINES); TORSION.

L. Sigfred Linderoth, Jr.

Torus

A surface obtained by rotating a circle about a line that lies in the plane of the circle but that has no points in common with the circle (see **illus.**). The

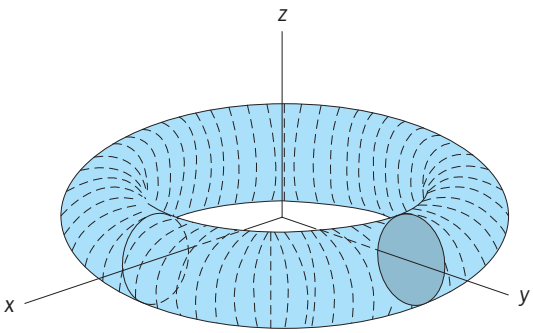

Diagram of a torus.

equations $x = u \cos v$, $y = u \sin v$, $z = [r^2 - (u - b)^2]^{1/2}$, $b > r > 0$, represent the upper half of the torus obtained by rotating about the z axis a circle of radius r whose center is the point $(b, 0, 0)$. The parameter u represents the distance of a point P of the torus from the z axis, and v is the angle of rotation. According to whether $b < u \leqq b + a$ or $b - a \leqq u < b$ or $u = b$, the corresponding point P is elliptic, hyperbolic, or parabolic, respectively, and the Gauss curvature of the surface at P is positive, negative, or zero. *See* DIFFERENTIAL GEOMETRY; MANIFOLD (MATHEMATICS); TOPOLOGY.

Leonard M. Blumenthal

Tourette's syndrome

A neurobehavioral disorder that is characterized by frequent, recurrent motor and vocal tics (described in 1885 by the French neurologist Gilles de la Tourette). The motor tics of Tourette's syndrome include brief, rapid, and darting movements of almost any muscle group, and can include eye blinking, eye rolling or deviations, nose wrinkling, facial grimacing, and head shaking. Some motor tics are more complex, are slow, and appear purposeful such as head turning, shoulder shrugging, touching, hopping, or twirling. Vocal tics are brief guttural sounds such as recurrent sniffing, throat clearing, coughing, and grunting or barking sounds. Complex vocal tics can be more meaningful and include expressions such as "No," "Stop," or "Some other time." Infrequently, persons with Tourette's syndrome repeat what they have said (palilalia), repeat what others have said (echolalia), or bark or grunt profane language (coprolalia). Tourette's syndrome has been described in nearly every country and ethnic group, with an estimated prevalence of one or two occurrences per 2000 people.

Symptoms. The motor and vocal tics begin in childhood. Within an individual the types and severity of tics will vary over time. The tics are often worse during adolescence and tend to improve during the twenties and thirties. The tic symptoms increase with stress and excitement and decrease with activities that require focused effort. While the motor and vocal tics are involuntary, they can be suppressed for brief periods of time, giving the false impression that the movements and sounds are voluntary. The characteristic waxing and waning nature of the tic movements can be confusing and often causes the disorder to be unrecognized and undiagnosed.

It is unusual for tic symptoms alone to be severe and incapacitating or to progress and worsen over the life of an individual. Often the impairment is related to associated behavioral problems, such as problems with attention, concentration, and impulsivity. Problems with recurrent, intrusive, and unwanted thoughts (obsessions) and behaviors or rituals (compulsions) can also cause significant impairment. Many other behavioral difficulties have been described in persons with Tourette's syndrome, but the scientific evidence for their association is not clear.

Inheritance pattern. There is a pattern of inheritance consistent with a single autosomal dominant gene whose expression is variable and dependent on the sex of the person. Symptoms can vary from transient tics to Tourette's syndrome and can include obsessive-compulsive symptoms. Males have the greatest risk of having some form of a tic disorder; females have a reduced risk of developing Tourette's syndrome but a greater risk than males of developing obsessive-compulsive symptoms. *See* HUMAN GENETICS; OBSESSIVE-COMPULSIVE DISORDER.

Brain involvement. The complexity of symptoms is likely related to the various brain regions implicated in the development of Tourette's syndrome.

The brain regions most likely to be involved are the basal ganglia, the frontal cortex, and the limbic system. The basal ganglia are responsible for movement control, the frontal cortex is responsible for the organization of thinking and feeling, and the limbic system is the considered the center of emotion in the brain. The dense interconnections of these brain regions may help explain why the symptoms of Tourette's syndrome can include uncontrolled motor movements and sounds in combination with unusual thoughts, feelings, and sensations. *See* BRAIN.

Treatment. Treatment can be targeted toward suppressing tics and the specific associated behavioral problems. Methods for tic suppression include medications such as haloperidol, fluphenazine, and pimozide—all major tranquilizers which exert their effect on the brain by blocking the neurotransmitter dopamine at the site of nerve-to-nerve connections. Treatments for the associated behavioral problems are dependent on the nature of the specific problem. Given that the overall impairment in persons with Tourette's syndrome is often related to the severity of the associated problems, the importance of treatment oriented toward these problems must be emphasized. *See* NERVOUS SYSTEM DISORDERS. John T. Walkup; Mark A. Riddle

Bibliography. T. N. Chase (ed.), *Tourette Syndrome: Genetics, Neurobiology, and Treatment*, 1992; R. Kurlan (ed.), *Handbook of Tourette's Syndrome and Related Tic and Behavioral Disorders*, 1993.

Tourmaline

A cyclosilicate mineral family with (BO_3) triangular groups and a complex chemical composition. The general formula can be written $XY_3Al_6(OH)_4(BO_3)_3(Si_6O_{18})$, in which X = Na, Ca, and Y = Al, Fe^{3+}, Li, Mg, Mn^{2+}. The more common tourmalines are dravite (X,Y = Na,Mg), schorl (X,Y = Na,Fe), uvite (X,Y = Ca,Mg), and elbaite (X,Y = Na,Li). Fluorine commonly substitutes in the hydroxyl position. Tourmaline is a hard ($7^1/_2$ on Mohs scale), varicolored mineral which can be an important semiprecious gemstone. *See* GEM.

Tourmaline is trigonal, space group R3m, with $a = 1.584$ nanometers and $c = 0.71$ nm for an elbaite. Other members have similar cell dimensions, the variation in size dictated by the size of the ions.

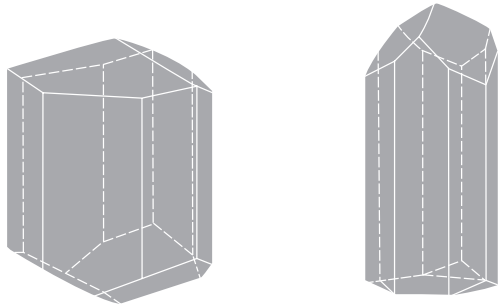

Tourmaline crystal habits. (*After C. S. Hurlbut, Jr., Dana's Manual of Mineralogy, 17th ed, Wiley, 1959*)

Crystals are typically trigonal prismatic (see **illus.**) and may occur as slender needles in elbaite to thick prismatic individual crystals of schorl which have been found in pegmatites. The tourmaline crystal is polar; thus, different forms are found at the opposite ends. Because of this polarity, tourmaline is piezoelectric; that is, if pressure is exerted at one end, opposite electrical charges will occur at opposite poles. It is also pyroelectric, with the electrical charges developed at the ends of the polar axis on a change in temperature. Because of its piezoelectric property, tourmaline can be cut into gages to measure transient pressures. *See* PIEZOELECTRICITY; PYROELECTRICITY.

Color in tourmalines is largely caused by incorporation of minor amounts of cations of the first transition series in the Y position and substitution for some of the octahedrally coordinated Al. For example, some Mn^{2+} (or Mn^{3+}) in the Y position in elbaite can color it pink to burgundy red, and this mineral is called rubellite. With $Fe^{2+} > Mn^{2+}$, tourmalines of the elbaite clan can be yellow to green, and with some Fe^{3+} present as well the gemstone is blue. These minerals are often called indicolite. The absence of transition-metal chromophores results in colorless achroite. Such colored, transparent flawless elbaites constitute the bulk of gemstones, and they are culled from granitic pegmatites. In these pegmatites, tourmalines occur in late-stage large chambers or pockets, characteristically associated with platy albite (cleavelandite); lepidolite; aquamarine beryl (gemmy bluish-green); morganite beryl (gemmy pink); and rarer accessory minerals such as lithiophilite, $LiMn(PO_4)$ and bertrandite, $Be_4(OH)_2(Si_3O_7)$. Magnificent gems have been obtained from certain late-stage pegmatites such as in San Diego County, California; Oxford County, Maine; Minas Gerais district, Brazil; Madagascar; and Afghanistan. In the same pegmatites, but at an earlier stage, occurs black schorl. Schorl, of no gem value, also occurs in regionally metamorphosed schists. Uvite and dravite typically occur in metamorphic rock assemblages such as regionally recrystallized marbles which have been intruded by granites. Their colors are characteristically dull green to brown and they are of no gem value. *See* PEGMATITE; SILICATE MINERALS. Paul Brian Moore

Bibliography. W. A. Deer, R. A. Howie, and J. Zussman, *Rock Forming Minerals: Disilicates and Ring Silicates*, vol. 1b, 1986.

Tower

A concrete, metal, or timber structure that is relatively high for its length and width. Towers are constructed for many purposes, including the support of electric power transmission lines, radio and television antennas, and rockets and missiles prior to launching.

Transmission towers. These towers are rectangular in plan and are not steadied by guy wires. A transmission tower is subjected to a number of forces: its own

weight, the pull of the cables at the top of the tower, the effect of wind and ice on the cable, and the effect of wind on the tower itself. Torsional forces on a tower caused by breakage of the cables in one span on one side of the tower must also be considered.

Radio and television towers. Such towers are either guyed or freestanding. Freestanding towers are usually rectangular in plan. In addition to their own weight, freestanding towers support the weight of the antenna and accessories and the weight of ice, unless a deicing circuit is installed. Wind forces must also be carefully considered.

Guyed towers are usually triangular in plan, with the main structural members, or legs, at the vertexes of the triangle. The legs are usually solid round steel bars. All members are galvanized and primed before erection. Television reception requires towers to be as high as possible to increase the area of coverage, and results in the use of the new high-strength steels to produce lighter-weight and taller towers. Field connections are made with galvanized bolts and lock-type nuts. The original design usually provides for increasing the height of the tower, when this is permitted by the Federal Communications Commission. *See* ANTENNA (ELECTROMAGNETISM); LAUNCH COMPLEX; TELEVISION TRANSMITTER; TRANSMISSION LINES.

Charles M. Antoni

Bibliography. D. A. Firmage, *Fundamental Theory of Structures*, 2d ed., 1980.

Towing tank

A tank of water used to determine the hydrodynamic performance of waterborne bodies such as ships and submarines, as well as torpedoes and other underwater forms. In the narrow sense, towing tanks are considered to be experimental facilities used to measure the forces, such as drag, on ship models and in turn to predict the performance of the full-scale prototype. In general, towing tanks are rectangular in planform with a uniform cross section. Different section shapes are used, ranging from rectangular to semicircular. The cross-section dimension may vary from about 8 to 52 ft (2.5 to 16 m) in width, from about 4 to 33 ft (1.5 to 10 m) in depth, and from under 100 ft (30 m) to almost 6560 ft (2000 m) in length; the size of the model varies in length from 4 to 30 ft (1.5 to 9 m).

Towing methods. The principal measurements made in a towing tank are force measurements, particularly drag or resistance of a towed ship model or other body. One of two principal systems for towing a model is used in most towing tanks. The simpler system consists of a gravity dynamometer and an endless cable attached to the model. A weight provides a constant towing force (**Fig. 1**). The time to traverse a fixed distance is measured when the model reaches a constant speed, thus establishing the speed-resistance relationship for the model. This dynamometer is simple and capable of high accuracy, but is limited to the measurement of the drag force of waterborne bodies. It is used in the smaller towing tanks in which the models are generally under 6 ft (2 m) in length.

In larger towing tanks the model is towed by a towing carriage mounted on rails at the side of the towing tank or suspended from an overhead track system (**Fig. 2**). Speed can be controlled and measured precisely on these carriages. Most carriages are equipped with a drag dynamometer as a permanent component (**Fig. 3**).

The dynamometer girder, mounted on the carriage frame, carries a long horizontal floating beam in pendulum fashion on two pairs of vertical arms terminating in flexible springs. A counterweight positioned at the upper end of a vertical swinging arm and mounted on the girder and attached to the floating beam maintains the beam in equilibrium at any position between the limit stops. The model resistance is transmitted as a horizontal force through the upper flexible link L_1 to the lower arm of the T-shaped balance, where it is balanced by the weight W. When the model resistance is not equal exactly to a unit weight W, the remainder is taken up (or applied) by the resiliency of the whole group of flexible spring supports; the exact amount of this auxiliary load or force is recorded on the drum by the link L_2 and the recording arm shown in Fig. 3.

In modern installations, the towing force is measured at the point where the towing link is connected to the towing post in the model. The measurements are usually made by measuring the strain in a flexure with strain gages, or by other means such as piezocrystals or inductance coils. The electronic signals are amplified and filtered and then sent to a computer for recording and analysis.

Law of similitude. Modern towing tank technology was established by William Froude in the 1870s, when he discovered the law of similitude for phenomena in which gravity is the predominating factor and established one of the essential principles of hydrodynamics for comparing model phenomena with the actual ship. There are three principal forces involved for a body moving through the water: inertia, gravity, and viscosity. The law of similitude requires that the ratio between inertia force and gravity force be the same for both the model and the prototype, and that the ratio between inertia force and viscous force be the same as well. *See* FROUDE NUMBER; REYNOLDS NUMBER.

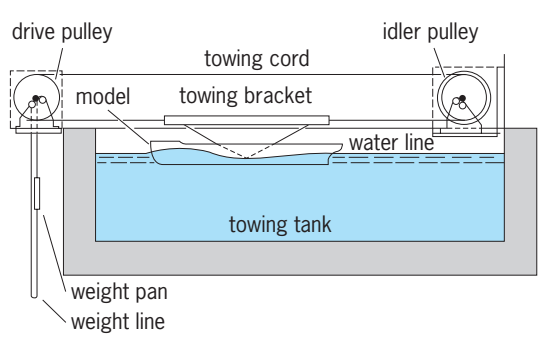

Fig. 1. Tank with model towed by falling weight.

Froude's law requires the validity of Eq. (1), where

$$\frac{V}{v} = \sqrt{\frac{L}{l}} = \sqrt{\lambda} \tag{1}$$

V and v represent the velocity of the prototype and model, respectively; L and l represent a characteristic dimension such as length of prototype and model, respectively; and λ represents the linear ratio of prototype to model.

Reynolds' law requires the validity of Eq. (2),

$$\frac{VL}{v_s} = \frac{vl}{v_m} \tag{2}$$

where v_s and v_m are the kinematic viscosity of the fluid for prototype and model, respectively. Because model towing tanks use fresh water and ships usually operate in seawater, v_s and v_m are essentially in the same order of magnitude. Thus, Froude's law requires a model velocity which is less than that of the prototype, whereas Reynolds' law requires a model velocity which is substantially higher than that of the prototype.

Froude overcame this difficulty by dividing the ship resistance into two parts: frictional resistance, and residuary resistance consisting mainly of wave-making resistance. Frictional resistance is primarily the effect of viscosity and is thus governed by Reynolds' law. Residuary resistance is due mainly to gravitational effect and thus is governed by Froude's law. Based on this distinction, Froude developed the technique of predicting ship resistance from the resistance test of a scaled model, a technique used to this day. It can be outlined as follows. A model geometrically similar to the prototype is made and the drag or resistance measured in a towing tank at the corresponding speeds as expressed in Eq. (1). Viscous or frictional resistance r_F of the model is calculated by assuming the resistance to be the same as that of a smooth flat plate of comparable area and length. Residuary resistance r_R, which is largely a gravitational effect, is then found by subtracting the frictional resistance from the total measured resistance r_T as in Eq. (3).

$$r_R = r_T - r_F \tag{3}$$

Residuary resistance of the ship R_R is calculated by the law of comparison at corresponding speeds as in Eq. (4). The frictional resistance of the ship

$$R_R = r_R \times \lambda^3 \tag{4}$$

R_F is calculated on the same assumptions used in calculating the viscous resistance of the model. The total resistance of the ship is finally found by adding the residuary and frictional resistance as in Eq. (5).

$$R_T = R_F + R_R \tag{5}$$

Besides the resistance test, there are two other important tests: the propeller open-water test and the self-propulsion test. From the measuring of the revolutions per minute, torque, and thrust of the model propeller, the shaft horsepower required to drive the prototype at designed speed can be predicted. *See* DIMENSIONLESS GROUPS.

Force measurements. Experimentation in modern towing tanks has been extended far beyond predicting the resistance and power of a moving body in still water. Towing tanks are used to measure any combination of forces and moments upon a waterborne or submerged body under steady-state conditions. A few of the many possible types of tests involving force measurements in a towing tank are (1) lift and

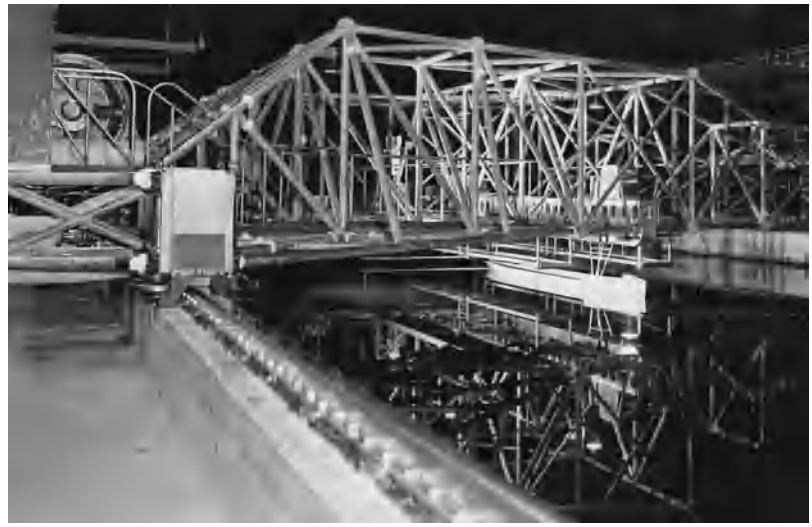

Fig. 2. Model in David Taylor Model Basin towed by carriage.

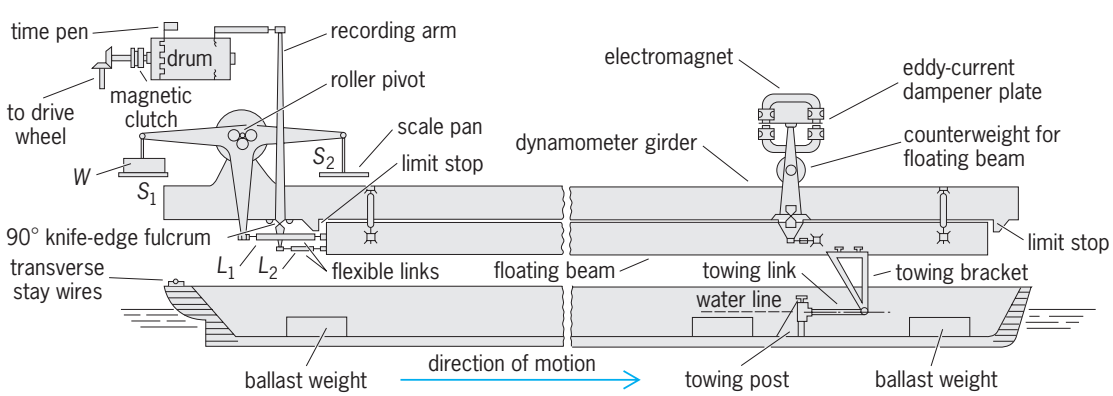

Fig. 3. Schematic arrangement of towing dynamometer.

drag of a planing surface; (2) lift, drag, and pitch moment of a submerged hydrofoil; (3) forces and moments on a submerged body towed at an angle of attack to obtain the static coefficients of a body for equations of motion; and (4) the turning moment on a ship's rudder.

Pressure and velocity measurement. A series of experiments has evolved from the measurement of pressures and velocities. The pressure measurements may be integrated over the surface to determine the force on the surface. Typical tests are the measurement of the pressure distribution on propeller blades, the duct of a ducted propeller, and the appendages of ship models. Various velocity-measuring devices have been developed, ranging from pitot tubes and hot-wire anemometers to sophisticated laser velocimetry. They are used to measure the velocity field at various locations around a model's hull, the most common of which is the plane in which the ship's propeller operates. *See* PITOT TUBE.

Lines of flow. A towing tank is also used to determine the lines of flow over portions of a hull; thus, appendages can be installed so that they will have minimum resistance and avoid or minimize the creation of cavitation. For surface ships, the conditions of comparable speed are maintained between model and prototype. Flow patterns on the hull are usually established by the emission of dyes upon a color-sensitive paint such as hydrogen sulfide upon a white lead-based paint. Flow patterns outside the boundary layer may be established by vanes or flags, which are free to pivot and orient themselves to the direction of flow. Typical tests are (1) establishment of the location of bilge keels; (2) orientation of shaft struts; and (3) determination of wave profile.

Wave experiments. Many towing tanks are provided with wavemakers, which extend the range of experimentation to the study of ship performance in head and following seas. At one end of the towing tank a wavemaker is installed which can generate waves of a more or less uniform profile with a predetermined height and length. At the other end of the tank a wave absorption beach is installed. Experiments are conducted on ship models at corresponding speeds in various head and following sea conditions, recognizing that the waves generated are much more regular than those encountered in the ocean. In addition to maintaining geometric similarity of the model and its prototype, the dynamic similarity of the system must be maintained. Measurements are made primarily of the motions of the body, particularly the rotational motions, pitch and roll, and the translational motions, heave (vertical) and surge (longitudinal). Some techniques involve free-running models with sensitive accelerometers. The carriage is used to carry the recording equipment and provide power to the model through flexible cables; these flexible cables do not exert any restraining force on the model. In some instances in connection with the measurement of motions, various components of force or moment may also be measured. *See* OCEAN WAVES.

Non-steady-state experiments. With the wide availability of electronic measuring instrumentation, the study of unsteady hydrodynamic phenomena has become increasingly important in towing tanks. Such fluctuating forces or pressures are measured as (1) pressure on a model's hull from propeller blades passing in proximity; (2) vibratory forces produced by a propeller operating in the variable velocity field behind a model; (3) route stability characteristics of a model from alternate course variations; and (4) forces and moments on a submerged body undergoing pure pitching and heaving motion. From the last measurement the coefficients for the equation of motion are obtained for a specific submarine configuration. This makes possible, through the use of the digital computer, the calculation of a submarine's motion without further experimentation and under a variety of conditions difficult to achieve in a towing tank. *See* DIGITAL COMPUTER.

Tank modifications. The towing tanks and the test equipment described have given rise to a number of special-purpose test facilities, which are modifications of the tanks described above. The most important of these are the maneuvering basin, the rotating-arm tank, seakeeping basins for both ships and offshore structures, ice tanks, and "vacuum" tanks for propulsion testing of large models. Maneuvering basins are used to measure the characteristics of ships pertaining to their maneuvering capability, such as turning radius, response to rudder deflections, and course stability. The rotating-arm tank is specialized in measuring the so-called rotary derivatives of a model. These derivatives link turning rate to forces and moments, and can be used for the determination of required rudder characteristics. Seakeeping basins are either towing tanks with a wave generation capability installed at one end, or a rectangular basin, with or without a carriage, with multiple wave flaps for the simulation of long-crested as well as short-crested irregular seas. The rectangular-basin type in particular has been instrumental in the explosive growth of the offshore oil industry. For fixed and floating offshore structures, the motion, anchoring forces, and internal forces are measured, including forces and moments so small that optical means must be sought to be able to measure them.

The most recent addition to the growing number of specialized towing tanks is the ice tank. These facilities may use actual ice or some other form of solid layer representing ice (in many cases a form of paraffin). They are used primarily to study the resistance characteristics of ships in ice, the effect of drift or pack ice on offshore structures, and the efficiency of ice breaking. *See* SHIP POWERING, MANEUVERING, AND SEAKEEPING; WATER TUNNEL (RESEARCH AND TESTING); SURFACE WAVES. Jacques B. Hadler

Bibliography. R. Bhattacharyya, *Dynamics of Marine Vessels*, 1978; D. A. Blank and A. E. Bock (eds.), *Introduction to Naval Engineering*, 2d ed., 1985; E. F. Gritzen, *Introduction to Naval Engineering*, 1980; E. Lewis (ed.), *Principles of Naval Architecture*, Society of Naval Architects and Marine Engineers, 3 vols., 1988; K. J. Rawson and E. C. Tupper, *Basic Ship Theory*, vol. 2, 4th ed., 1995.

Townsend discharge

A particular part of the voltage-current characteristic curve for a gaseous discharge device named for J. S. Townsend, who studied it about 1900. It is that part for low current where the discharge cannot be maintained by the field alone. Thus, if the agents producing the initial ionization were removed, conduction would cease.

In the lower end of this region, conduction is accomplished only by charges produced by external agents. As the electric field is increased, secondary ionization and more efficient collection of the primary ionization cause an increase in the current. After further increase in the field, the end of the Townsend region is reached. Any additional increase in the field causes a transition into a region where the discharge may be maintained by the field alone, whether it be glow, brush, or arc. *See* DARK CURRENT; ELECTRICAL CONDUCTION IN GASES; GLOW DISCHARGE.

Glenn H. Miller

Toxic shock syndrome

A serious, sometimes life-threatening disease usually caused by a toxin produced by some strains of the bacterium *Staphylococcus aureus*. The signs and symptoms are fever, abnormally low blood pressure, nausea and vomiting, diarrhea, muscle tenderness, and a reddish rash, followed by peeling of the skin.

Toxic shock syndrome was first reported in 1978 in seven pediatric patients. The disease, however, became prominent in 1980, when hundreds of cases were reported among young women without apparent staphylococcal infections. Epidemiologists observed that the illness occurred predominantly in young women who were menstruating and were using tampons, especially those that contained so-called superabsorbent synthetic materials. A toxin [toxic shock syndrome toxin number 1 (TSST-1)] that occurs in some strains of staphylococci has since been identified. These bacteria are known to proliferate in the presence of foreign particles in human infections, and it has been postulated that the tampons acted as foreign particles, allowing toxin-producing staphylococci to multiply in the vagina, where they usually compete with harmless bacteria.

Theoretically, millions of women were at risk, and yet the number of reported cases remained small. This suggests that those women who suffered from toxic shock syndrome were especially susceptible. Susceptibility may depend on their lacking antibodies to the toxin that occur in most adults. Several hundred cases of toxic shock syndrome not associated with menstruation have been reported. In these cases, which occurred in males as well as females, there was almost always an overt staphylococcal infection.

The toxin has been shown to occur in only about 1% of the staphylococcal strains studied. Moreover, there is some evidence that the syndrome may be caused also by other staphylococcal toxins, particularly enterotoxins. Cases of toxic shock syndrome that were caused by streptococci have been reported. A toxin distinct from TSST-1 appears involved. Persons with the symptoms of toxic shock syndrome should receive immediate medical care to reduce the chance of death. *See* STAPHYLOCOCCUS; TOXIN.

Jay O. Cohen

Bibliography. L. A. Cone et al., Clincial and bacterologic observations of a toxic shock-like syndrome due to *Streptococcus pyogenes, N. Engl. J. Med.*, 317:146–147, 1987; J. P. Davis et al., Toxic-shock syndrome: Epidemiologic features, recurrence, risk factors, and prevention, *N. Engl. J. Med.*, 303:1429–1435, 1980; A. L. Reingold et al., Toxic shock surveillance in the United States, 1980 to 1981, *Ann. Intern. Med.*, 96:875–880, 1982; K. N. Shands et al., Toxic-shock syndrome in menstruating women: Association with tampon use and *Staphylococcus aureus* and clinical features in 52 cases, *N. Engl. J. Med.*, 303:1436–1442, 1980; J. Todd et al., Toxic-shock syndrome associated with group-I staphylococci, *Lancet*, 2:1227–1229, 1978; W. Tyson et al., Atypical staphylococcal toxic shock syndrome: Two fatal cases, *Ped. Infect. Dis.*, 8:642–645, 1989.

Toxicology

The study of the adverse effects of chemical and physical agents on living organisms. Toxicology has also been referred to as the science of poisons. Since about 1970, there has been a substantial increase in the understanding of biological systems and the way in which chemicals can interfere with the proper functioning of such systems. At the same time, there has been a tremendous increase in the number and volume of chemicals that are in everyday use. Therefore, despite intensive research efforts, knowledge of fundamental toxicological principles and the specific effects of many individual compounds is incomplete. *See* POISON.

Factors that influence toxicity. The most important factor that influences the toxic effect of a specific chemical is the dose. All chemicals, including essential substances such as oxygen and water, produce toxic effects when administered in large enough doses. For a specific chemical, such as ethanol, the active ingredient of alcoholic beverages, the main question is the type of effects that this substance will produce and the doses at which these effects will occur. Another significant factor is the route of exposure. Living organisms may be exposed to a chemical by inhalation (into the lungs), ingestion (into the stomach), penetration through the skin, or, in special circumstances, injection into the body. In general, substances are absorbed into the body most efficiently through the lungs so that inhalation is often the most serious route of exposure. However, some chemicals do not vaporize very easily, so that other routes of exposure are more likely to occur and will be of greater toxicological significance.

A third factor is the fate of the chemical after the organism is exposed. The chemical may not be absorbed at all, limiting its possible adverse effects to the site of exposure, for example, a burn on the skin

or irritation of the throat. If it is absorbed, then it may travel throughout the body and has the potential to cause toxic effects at one or more sites remote from the site of entry. The remote sites where these adverse effects occur are called target organs.

A chemical that is absorbed and distributed throughout the body may be altered or metabolized, with the liver being the principal site of metabolism. There are two types of metabolic activity: useful chemicals are broken down into products (metabolites) that can be incorporated into the normal functioning of the organism; and unwanted chemicals are changed into forms that are more easily excreted from the body. The second type of activity can result in the formation of one or more new products, each of which may have greater or lesser toxicity than the chemical that was absorbed. When these products are less toxic, the process is known as detoxification. It serves to protect the organism from adverse toxic effects, but in some cases toxicity occurs as the dose is increased because the mechanisms for detoxification are overwhelmed.

The chemical and its metabolites can be excreted, stored, or transported in the organism and may, therefore, reach sites where toxic effects are induced. Substances that are rapidly excreted will have little opportunity to have adverse effects and thus are generally of low toxicity. Those that are excreted more slowly have the potential to cause long-term effects. Many substances are stored in the body, mainly in fat or bone, and thus can circulate throughout the organism for a long time.

Another significant variable is the time course of the exposure. A quantity of chemical administered at one time may have an effect even though the same quantity administered in small doses over time has no effect. If a chemical is completely excreted, then succeeding doses have no increased effect. However, if a residue remains, then it is possible for the second dose to add to the first and, if doses are repeated often enough, to lead to a level high enough to be toxic.

While the properties of a chemical are innate and unchangeable, those of the exposed organism may change over time. Individuals may be more susceptible at some life stages than others—for example, when they are very young or very old. The fetus is especially sensitive to chemical insult at certain developmental stages. In addition, previous exposures of the organism may make it more or less susceptible. There are also significant endogenous factors, such as the genetic makeup of the organism. Some species or individuals respond at very low doses, whereas others do not respond until very high levels are reached. The sex of the individual can also have an impact on susceptibility.

Classes of toxicity. In view of the importance of timing in producing adverse effects, toxicologists distinguish between two broad classes of toxicity, acute and chronic. Acute toxicity refers to effects that occur shortly after a single exposure or small number of closely spaced exposures. Chronic toxicity refers to delayed effects that occur after long-term repeated exposures.

Traditionally, the effect of most concern for acute toxicants (such as cyanide) is death. Consistent with this, acute toxicity is generally measured by using an assay to determine the lethal dose. In particular, rodents are given single doses and the number that have died 14 days later is recorded. The data are plotted for each dose, and the dose that is lethal for 50% of the animals (lethal dose 50 or LD_{50}) is used as the criterion for acute toxicity. *See* LETHAL DOSE 50.

Some synthetic chemicals, such as polychlorinated biphenyls (PCBs) and dichlorodiphenyltrichloroethane (DDT), exhibit their effects only after a number of repeated exposures and are considered chronic hazards, with cancer and reproductive effects being of greatest concern. To determine the dose at which chronic effects occur, rodents are exposed to daily doses of the chemical under study for long periods of time—from a few months to a lifetime. The highest dose at which no effects can be observed, the no observed effect level (NOEL), is used as a measure of chronic toxicity.

To study possible carcinogenic effects, rodents are exposed to very high doses over a lifetime in an attempt to detect low response in a small population of animals. Reproductive effects are studied by using special experimental protocols during which males and females are exposed during various periods of life, including the period of gestation for females. The experimental animals are utilized as surrogates for humans, and various techniques are employed to extrapolate results from rodents to humans. In many cases, the NOELs are divided by specific numbers, called safety factors, to estimate the levels below which effects are not likely to occur in humans. In other cases, particularly carcinogenicity, the high-dose results are extrapolated mathematically to low doses and, by using a scaling factor, to humans. These techniques have been designed to establish safe levels for humans.

Toxicology of metals. The basic principles of toxicity and toxicity assessment can be illustrated with the metals. These elements occur naturally, but the highest human exposures are usually of anthropogenic origin, especially when the metal is used in the workplace. Exposures can be due to other causes, however, such as burning coal or the use of pesticides to increase crop yields. The **table** lists metals that are toxicologically significant, but oversimplifies the situation since the types of toxic effects that occur often depend on a number of factors, such as the form of the metal (that is, inorganic or organic) and the specific type of chemical group to which the metal ion is bonded (for example, sulfate or phosphate). In addition, there are interactive effects so that deficiencies in levels of one metal may affect the toxicity of other metals.

The metals represent a very small fraction of the tens of thousands of chemicals in use, but there is much more information available about adverse effects in humans exposed to metals than about human toxicity resulting from most other chemicals in the environment. The metals that pose the greatest threats to public health are lead, mercury, and

Sources of exposure and toxicological properties of selected metals

Substance	Source of exposure	Acute effects in humans	Chronic effects in humans	Treatment
Arsenic	Occupational, pesticidal, food, drinking water, air	Fever, anorexia, skin lesions, multiorgan effects	Liver injury, peripheral vascular disease, cancer	Chelating agents for pulmonary and skin effects
Beryllium	Occupational, air (coal combustion)	Chemical pneumonia	Chronic pulmonary disease, cancer	Cease exposure
Cadmium	Occupational, food, air	Nausea and vomiting (ingestion), chemical pneumonia (inhalation)	Pulmonary disease, kidney damage, hypertension, bone fragility, cancer	Cease exposure
Chromium	Occupational, food	Respiratory irritation (inhalation)	Irritation, skin reactions, cancer	Cease exposure
Cobalt	Occupational, food	Vomiting, diarrhea, cardiomyopathy	Goiter, dermatitis, lung irritation	Cease exposure
Copper	Occupational, water	Nausea, vomiting, liver damage	Wilson's disease	Chelating agents
Gold	Medicinal	None reported	Dermatitis, kidney disease	Cease exposure
Lead	Food, water, air, lead paint, occupational	Vomiting, spasms	Neurological effects (especially in children), kidney damage, anemia	Chelating agents, removal of sources
Manganese	Occupational, food, air	Inflammation of the lung	Respiratory disease, neurological effects	L-Dopa
Mercury	Occupational, food, air	Bronchitis (inhalation); abdominal cramps, bloody diarrhea, and kidney damage (ingestion)	Neurological effects, kidney damage	Hemodialysis, chelating agents
Molybdenum	Occupational	Irritation	—	Cease exposure
Nickel	Occupational	Headache, nausea, vomiting, fever	Dermatitis, cancer	Cease exposure
Silver	Occupational	Gastrointestinal irritation	Argyria, kidney and lung damage	Cease exposure
Thallium	Occupational, pesticidal	Gastrointestinal irritation, neurological effects, cardiovascular effects	Hair loss, cataracts, neurological effects, kidney damage, loss of vision	Cease exposure
Zinc	Occupational, food, water, air	Abdominal distress, diarrhea (ingestion); fever and chills (inhalation)	Pulmonary effects	Cease exposure

arsenic, but toxic effects can be caused by other metals as well. *See* ENVIRONMENTAL TOXICOLOGY; MUTAGENS AND CARCINOGENS.

Lead. The most serious adverse effects of lead, mental retardation and learning problems, occur in young children subjected to chronic exposure, most often through ingestion of lead paints. Children with iron and calcium deficiencies absorb more lead and may suffer greater adverse effects. The principal effect of lead on adults is also neurological. The initial symptoms include nausea, vomiting, and joint pain, and at higher exposure levels may progress to toxic psychosis. Lead can also cause a variety of other effects, such as hypertension, anemia, kidney damage, and digestive problems. Lead acetate and lead phosphate are carcinogenic in animals, and it is possible that similar effects occur in humans. *See* LEAD.

Mercury. The toxicity of mercury is strongly dependent on its form. Mercury vapor causes bronchitis and pneumonitis, as well as damage to the central nervous system. Symptoms of the latter include excitability and tremors. Inorganic mercury poisoning is characterized by ulceration, bleeding, and necrosis of the gastrointestinal tract followed by severe renal effects. Chronic exposures may also result in kidney damage.

Organic mercury, particularly methylmercury, is the most important public health threat. Consumption of fish or grain containing methylmercury has led to significant epidemics in Japan, Iraq, and Pakistan, as well as lesser incidents in other countries. The major symptoms are neurological and include numbness, lack of coordination, speech impairment, and deafness. These reflect lesions in both the cerebellum and cerebral cortex. Organic mercury can also be toxic to the developing fetus. *See* MERCURY (ELEMENT).

Arsenic. The third metal most often implicated in human toxicity is arsenic. In this case, the valence form is critical to its toxicity. Trivalent arsenic, as in arsenic trioxide (As_2O_3), is usually more toxic to mammals than is pentavalent arsenic, as in arsenic pentoxide (As_2O_5).

The acute effects of high doses of arsenicals are broad and impact the upper respiratory tract and the cardiovascular, gastrointestinal, hematopoietic, and central nervous systems. A characteristic symptom resulting from ingestion is a pattern of skin abnormalities, including hyperpigmentation. Chronic, low-level exposures lead to more subtle effects on the gastrointestinal tract, nervous system, and skin. Continued intake can lead to degeneration of the liver and kidneys. Chronic exposure to arsenic is associated with cancers of the skin and lung and may be linked to cancers of internal organs. *See* ARSENIC.

Other metals. Toxic effects due to exposure to other metals may be of concern in situations where occupational or environmental exposure is high. These

other metals will be considered in alphabetical order.

1. *Beryllium*. Acute inhalation of high levels of beryllium causes severe inflammation of the upper respiratory tract and lungs. Exposure to lower doses can lead to more insidious disease, developing over weeks, with less severe pulmonary damage. A less toxic but more common effect is delayed-type allergic contact dermatitis. Chronic exposure can also lead to severe respiratory disease as well as to enlargement of organs such as the spleen and liver. Such effects may appear after a latency period of up to 20 years. Beryllium has been shown to cause cancer in laboratory animals, but this has not been confirmed in humans.

2. *Cadmium*. Exposure to high levels of cadmium over short periods of time leads to nausea, vomiting, and cramps. Inhalation of cadmium fumes can produce acute pneumonitis and pulmonary edema. Chronic exposure can result in kidney damage (renal tubular disease), chronic obstructive pulmonary disease, bone fragility, and emphysema. Ingested cadmium is carcinogenic in animals, but such carcinogenicity has not been confirmed in humans.

3. *Chromium*. Chromium intoxication depends on the valence form, with the hexavalent compounds being more toxic. These hexavalent compounds are irritating and corrosive but are slowly absorbed, so that chronic effects are localized, often limited to skin ulceration and gastrointestinal tract irritation. Chronic inhalation of chromium is associated with cancer of the lung in humans.

4. *Cobalt*. Cobalt is an essential nutrient and of low toxicity at usual levels. However, a number of fatalities due to cardiomyopathy occurred in people consuming beer containing a cobalt additive. Very high acute exposures cause vomiting and diarrhea, and high chronic exposures are linked to goiter. In addition, chronic occupational exposures may lead to respiratory irritation and dermatitis.

5. *Copper*. Copper is also an essential nutrient, and will not produce intoxication unless there is exposure to high levels. Acute symptoms of copper ingestion are nausea and vomiting due to vascular congestion of the gastrointestinal tract. Very high doses can lead to death from vascular collapse. In addition, liver necrosis may result in such cases. Chronic exposure to low levels of copper does not cause adverse effects except in rare individuals whose genetic makeup leads to excess copper accumulation (Wilson's disease).

6. *Gold*. Gold is toxic only after long exposures such as those associated with chronic intake of gold used medicinally. Effects include dermatitis, stomatitis, and injury to kidney tubules.

7. *Manganese*. Manganese toxicity has been seen only in miners and mill workers. Acute exposure can lead to a lung disease known as manganese pneumonitis. Chronic exposure can result in severe effects on the central nervous system, effects ascribed to interference with metabolism of amines. Symptoms include psychotic episodes, speech impairment, hypokinesia, and liver cirrhosis.

8. *Molybdenum*. Human poisoning due to molybdenum, an essential mineral, has rarely been observed and appears limited to irritation at the sites of exposure.

9. *Nickel*. The most serious effect of nickel exposure is lung and nasal cancer associated with chronic inhalation by refinery workers. Acute effects from ingestion of high doses of nickel include headache, nausea, and vomiting as well as pulmonary symptoms. Continued exposure can lead to lesions in lung and brain. The most common effect of low-level exposure in the general population is contact dermatitis.

10. *Silver*. The toxic effects of elemental silver are limited since it is poorly absorbed. Ingestion of very large amounts of silver can lead to acute gastroenteritis. Chronic occupational exposure can lead to argyria, a condition characterized by local or generalized changes in skin pigmentation. In rare instances, the respiratory tract and kidneys are also affected.

11. *Thallium*. The principal symptoms of acute thallium poisoning are gastroenteritis, headache, and rapid heartbeat followed by neurological effects such as paralysis and psychic disturbances. Lower levels of exposure can lead to lack of coordination, numbness, and tremor. Chronic effects include liver necrosis, nephritis, cataracts, neurological damage, and pulmonary edema. Hair loss is a characteristic symptom of long-term thallium exposure.

12. *Zinc*. The most significant toxic effect of zinc is zinc fume fever, fever and chills that can result from acute inhalation of zinc oxide fumes. In addition, ingestion of high doses of soluble salts of zinc can lead to acute gastroenteritis. Chronic inhalation of zinc can cause pulmonary irritation. *See* BERYLLIUM; CADMIUM; CHROMIUM; COBALT; COPPER; GOLD; MANGANESE; MOLYBDENUM; NICKEL; SILVER; THALLIUM; ZINC. Michael Kamrin; Robert W. Leader

Bibliography. L. Friberg et al. (eds.), *Handbook of the Toxicology of Metals*, 2d ed., 1986; M. A. Kamrin, *Toxicology: A Primer on Toxicology Principles and Applications*, 1988; C. D. Klaasen, M. O. Amdur, and J. Doull (eds.), *Casarett and Doulls' Toxicology*, 3d ed., 1986; F. C. Lu, *Basic Toxicology*, 1985.

Toxin

Properly, a poisonous protein, especially of bacterial origin. However, nonproteinaceous poisons, such as fungal aflatoxins and plant alkaloids, are often called toxins. *See* AFLATOXIN; ALKALOID.

Bacterial exotoxins. These proteins of disease-causing bacteria are usually secreted and have deleterious effects. Several hundred are known that account to a variable extent for the damage caused by bacterial diseases and infections. In some extreme cases a single toxin accounts for the principal symptoms of a disease, such as diphtheria, tetanus, and cholera. Bacteria that cause local infections with pus (such as *Streptococcus pyogenes*, *Staphylococcus aureus*, and *Clostridium perfringens*) often produce many toxins that affect the tissues around the

infection site or are distributed to remote organs by the blood. *See* CHOLERA; DIPHTHERIA; STAPHYLOCOCCUS; TETANUS.

Toxins may assist the parent bacteria to combat host defense systems, to increase the supply of certain nutrients such as iron, to invade cells or tissues, or to spread between hosts. Sometimes the damage suffered by the host organism has no obvious benefit to the bacteria. For example, botulinal neurotoxin in spoiled food may kill the person or animal that eats it long after the parent bacteria have died. In such situations it is assumed that the bacteria benefit from the toxin in some other habitat and that the damage to vertebrates is accidental. In complex situations the role played by a toxin in disease or in assisting the bacteria is best studied by comparing the courses of artificial infections established with genetically matched pairs of bacteria that differ only in the ability to synthesize a single toxin. *See* FOOD POISONING.

Intracellular action sites. Certain bacterial and plant toxins have the unusual ability to catalyze chemical reactions inside animal cells. Such toxins are always composed of two functionally distinct parts termed A and B, and they are often called A-B toxins. The B part binds to receptor molecules on the animal cell surface and positions the toxin upon the cell membrane. Subsequently, the enzymically active A portion of the toxin crosses the animal cell membrane and catalyzes some intracellular chemical reaction that disrupts the cell physiology or causes cell death. A and B may be adjacent regions of a single polypeptide chain (for example, in diphtheria toxin and *Pseudomonas aeruginosa* exotoxin A); two protein chains linked by disulfide bonds (for example, the plant toxins abrin and ricin); a complex of separate protein subunits encoded by different genes (for example, cholera toxin, pertussis toxin, and shigella toxin); or two unlinked proteins that only come together on the target cell's surface (for example, anthrax toxin and botulinal C2 toxin). The last type are also called binary toxins.

About one-half of the known A-B toxins catalyze a class of chemical reaction known as adenosine diphosphate (ADP) ribosylation whereby key cellular enzymes acquire a large chemical tag, that is, ADP-ribose, that disturbs their functions. Many of the target enzymes are proteins that bind guanosine triphosphate (GTP), called G proteins, and normally function in the regulation of cell physiology, shape, and growth.

Several of the A-B toxins increase the cell content of an intracellular messenger molecule, cyclic adenosine monophosphate (cyclic AMP). In some cases the toxins stimulate the activity of the target cell's enzyme adenylate cyclase which synthesizes cyclic AMP. In other cases the toxin itself is an adenylate cyclase. In phagocytes an increased cyclic AMP level translates into reduced migration to the site of infection, reduced phagocytosis, and reduced intracellular killing of bacteria.

When the intracellular reaction of an A-B toxin blocks protein synthesis, the target cell dies. Diphtheria toxin's fragment A inactivates a soluble protein key in protein synthesis: one molecule of fragment A is enough to kill an animal cell in this way. Several A-B toxins of plant origin, such as abrin, ricin, and viscumin, inactivate ribosomes. Other plant seeds, including wheat, barley, rye, and corn, contain A chains similar to the A chain of ricin, but no B chain. These A chains inactivate ribosomes in the test tube but have very limited abilities to enter cells and fortunately are not toxic. Much research has been directed at the formation of artificial hybrids containing the A portion of diphtheria toxin or one of the plant proteins, combined with a molecule that binds only to cancer cells such as a surface-specific antibody. In certain cases these hybrid immunotoxins are able to selectively kill tumor cells. *See* IMMUNOLOGIC CYTOTOXICITY.

Membrane damage. A large group of toxins (more than 200 are known) breach the normal barrier to free movement of molecules across cell membranes. In sufficient concentration such cytolytic toxins cause cytolysis, a process by which soluble molecules leak out of cells, but in lower concentration they may cause less obvious damage to the cell's plasma membrane or to its internal membranes. The cytolytic agents are commonly detected by their ability to lyse red blood cells. Some are enzymes that hydrolyze cell membrane phospholipids, such as *Clostridium perfringens* alpha toxin and *Staphylococcus aureus* beta toxin; others insert themselves into the membrane and form channels through which small molecules may pass, such as *Staphylococcus aureus* alpha and delta toxins, mellitin of bee venom, and streptolysin O and streptolysin S of *Streptococcus pyogenes*. *See* CELL MEMBRANES.

Neurotoxins. Tetanus and botulinal neurotoxins owe their extraordinary potencies to the fact that their target cells are neurons; both toxins block the transmission of nerve impulses across synapses. Tetanus toxin is transported by axons to the central nervous system where the blockage, mainly of inhibitory synapses, results in spastic paralysis, in which opposing muscles contract simultaneously. The botulinal neurotoxins principally paralyze neuromuscular junctions and cause flaccid paralysis, in which the muscles cannot contract. Snake and invertebrate venoms often contain less potent neurotoxins, such as the crotoxin of the Brazilian rattlesnake and the bungarotoxins produced by the Formosan banded krait. Also, many poisonous species contain low-molecular-weight (nonprotein) substances that have specific effects on the nervous system and are useful in brain research, such as tetrodotoxin of the puffer fish, black widow spider venom, and the algal saxitoxin of red tides.

Fever-producing toxins. Gram-negative bacteria, such as *Salmonella* and *Hemophilus*, have a toxic component in their cell walls known as endotoxin or lipopolysaccharide. Among other detrimental effects, endotoxins cause white blood cells to produce interleukin-1 (formerly known as endogenous pyrogen), a hormone responsible for fever, malaise,

Approximate median lethal doses of some toxins per kilogram of body weight

Toxin	Dose	Test animal
Tetanus	1 nanogram	Mouse, probably human
Botulinal neurotoxin	1 nanogram	Mouse, human
Shigella	1 nanogram	Monkey, human
	1 microgram	Mouse
Diphtheria	100 nanograms	Human
	1.6 milligrams	Mouse
Ricin	1 microgram	Human
Staphylococcus alpha	50 micrograms	Mouse
Neurotoxins of snakes and invertebrates	50 micrograms	Mouse

headache, muscle aches, and other nonspecific consequences of infection. The exotoxins of toxic shock syndrome and of scarlet fever induce interleukin-1 and also tumor necrosis factor, which has similar effects. *See* ENDOTOXIN; FEVER; SCARLET FEVER; TOXIC SHOCK SYNDROME.

Immunity to toxins. Toxoids are toxins that have been exposed to formaldehyde or other chemicals that destroy their toxicities without impairing immunogenicity. When injected into humans, toxoids elicit specific antibodies known as antitoxins that neutralize circulating toxins. Such immunization (vaccination) is very effective for systemic toxinoses, such as diphtheria and tetanus. It has resulted in the virtual eradication both of diphtheria and of its causative organism, *Corynebacterium diphtheriae*. Immunization with a toxoid is less successful in controlling diarrheal diseases caused by toxins produced within the intestine that are not exposed to antitoxins in the blood. *See* ANTIBODY; DIARRHEA; DIPHTHERIA; IMMUNITY; VACCINATION.

Relative toxicities. Toxicities vary with the route of administration and can differ greatly with the test animal. It may be difficult to predict human toxicities solely from measurements with other animals. For example, mice are many-thousand-fold less sensitive than humans to diphtheria toxin because their cells lack a surface component needed for the entry of fragment A (see **table**).

D. Michael Gill

Bibliography. R. J. Collier and D. A. Kaplan, Immunotoxins, *Sci. Amer.*, 251:56–64, 1984; D. M. Gill, Bacterial toxins: A table of lethal amounts, *Microbiol. Rev.*, 46:86–94, 1982; R. F. Keeler, N. Mandara, and A. T. Tu, *Natural Toxins: Toxicology, Chemistry and Safety*, 1992; A. I. Laskin and H. A. Lechevalier (eds.), *Handbook of Microbiology*, vol. 8, 1987; J. L. Middlebrook and R. B. Dorland, Bacterial toxins: Cellular mechanisms of action, *Microbiol. Rev.*, 48:199–221, 1984; C. A. Mims, *Pathogenesis of Infectious Disease*, 5th ed., 2001.

Toxin-antitoxin reaction

In serology, the combination of a toxic antigen with its corresponding antitoxin. If the antitoxin is derived from any species other than the horse, precipitation occurs over a wide range of reactant ratios, as in other antigen-antibody reactions. With horse antitoxin, flocculation occurs only if toxin and antitoxin are near equivalence, a twofold excess of either reactant giving soluble complexes. In most instances, the reaction results in partial or complete neutralization of the toxic activity of the antigen. *See* ANTIGEN; ANTITOXIN; NEUTRALIZATION REACTION (IMMUNOLOGY); TOXIN.

The Danysz reaction occurs when an exact equivalence of toxin is added to antitoxin, not in one portion, but in successive increments. The mixture will remain somewhat toxic, although it becomes neutral within a few days. This is the result of the variable combining proportions of antigen and antibody, which result in a binding of excess antibody by the initial increments of toxin giving soluble molecular complexes composed of one molecule of toxin and eight molecules of antibody (TA_8). As a result, some of the subsequent increments of toxin remain free. In time, the system rearranges to give the neutral equivalence composition, one molecule of toxin and two of antitoxin (TA_2). *See* ANTIBODY; SEROLOGY.

Henry P. Treffers

Toxoplasmea

A class of the subphylum Sporozoa. Inadequate information about the structure and life cycles of the members of this class has, for many years, resulted in the organisms being classified as parasites of unknown nature. Electron microscope studies have shown structural similarities between this group and the Telosporea, but characteristics of the Telosporea such as the presence of spores (oocysts) and sexual reproduction have only recently been reported for the Toxoplasmea. These findings will necessitate reclassification of the Toxoplasmea. *See* SPOROZOA; TELOSPOREA; TOXOPLASMIDA.

Morphology. The organisms are small and crescent-shaped. They move by body flection or gliding and have no flagella or pseudopodia. Characteristic structures are the two-layered pellicle and underlying longitudinal microtubules, micropyle, a conoid, paired organelles, and micronemes.

The most distinguishing characteristic of the Toxoplasmea is the unique means of reproduction. Although reproduction by binary fission has been reported for these organisms, electron microscope studies indicate that endodyogeny is the

sole method. Endodyogeny is an internal budding wherein two daughter cells are produced within a mother cell which is destroyed in the process.

Life cycle. Only two stages are known in most animals. One stage, the proliferative form or trophozoite, occurs singly or in groups within host cells. The other stage, the so-called cyst, consists of a large number of organisms which, with minor differences, are structurally similar to the proliferative forms.

The life cycle of one member of this group has recently been described. Although proliferative forms and cysts of *Toxoplasma gondii* are found in many animals, the complete life cycle has been found only in the cat. A resistant *Isospora*-like oocyst can transmit the organism from cat to cat. Asexual and sexual stages develop in the intestinal epithelium. Further study may elucidate similar life cycles for other members of the Toxoplasmea. *See* PROTOZOA.

Harley G. Sheffield

Bibliography. K. M. Adam et al., *Medical and Veterinary Protozoology*, rev. ed., 1980; J. N. Farmer, *The Protozoa: Introduction to Protozoology*, 1980; R. R. Kudo, *Protozoology*, 5th ed., 1977; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Toxoplasmida

An order of the class Toxoplasmea. Four genera, *Toxoplasma, Besnoitia, Sarcocystis*, and *Encephalitozoon*, make up the order. The organisms are parasites of vertebrates. *Toxoplasma* is often found encysted in nerve tissue, *Besnoitia* in connective tissue, and *Sarcocystis* in muscle. Very little is known about *Encephalitozoon*, but the parasite has been found in the brain of rabbits. *See* TOXOPLASMEA.

Only one or two stages of the life cycle of most members are known. A proliferative form may be found singly or in small groups within host cells. The other stage, the cyst, is found intra- or extracellularly.

Typical structures appear in the proliferative form of *T. gondii*. The organism is surrounded by a two-layered pellicle. A micropyle is located in the pellicle lateral to the nucleus. A cone-shaped structure, the conoid, is situated at the anterior end. Long clublike paired organelles extend posteriorly from the conoid area. The nucleus and adjacent Golgi zone are in the posterior half of the organism. Mitochondria, granules, vesicles, and ribosomes fill the cytoplasm.

Toxoplasma produces a serious, although rare, disease in humans. Toxoplasmosis occurs throughout the world. Serological studies indicate that many people have been exposed to the organism, although clinical disease is rare. Congenital toxoplasmosis may result in abortion or deformity, and ocular toxoplasmosis may destroy the eye.

While the mechanism of transmission of *Sarcocystis, Besnoitia*, and *Encephalitozoon* is unknown, that of *Toxoplasma* has been described. This parasite is a coccidian of cats, although proliferative forms and cysts have been found in many animals, including humans. The exogenous form is an *Isospora*-like oocyst which is passed in the feces. The oocyst is infective to cats, to most laboratory animals, and perhaps to humans. Probably the more common means of human infection is through ingestion of raw or poorly cooked meat.

Harley G. Sheffield

Bibliography. R. L. Barnes, *Invertebrate Zoology*, 6th ed., 1994; J. A. Pechenik, *Biology of the Invertebrates*, 4th ed., 2000.

Trace fossils

Fossilized evidence of animal behavior, also known as ichnofossils, biogenic sedimentary structures, bioerosion structures, or lebensspuren. The fossils include burrows, trails, and trackways created by animals in unconsolidated sediment, as well as borings, gnawings, raspings, and scrapings excavated by organisms in harder materials, such as rock, shell, bone, or wood. Some workers also consider coprolites (fossilized feces), regurgitation pellets, burrow excavation pellets, rhizoliths (plant root penetration structures), and algal stromatolites to be trace fossils. *See* STROMATOLITE.

Trace fossils are important in paleontology and paleoecology, because they are fossils that provide information about the presence of unpreserved soft-bodied members of the original communities, life habits of fossil organisms, evolution of certain behavior patterns through geologic time, and biostratigraphy of otherwise unfossiliferous deposits. Trace fossils also are useful in sedimentology and paleoenvironmental studies, because they are sedimentary structures that are preserved in place and are very rarely reworked and transported, as body fossils of animals and plants commonly are. This fact allows trace fossils to be regarded as reliable indicators of original conditions in the sedimentary environment. The production of trace fossils involves disruption of original stratification and sometimes results in alteration of sediment texture or composition.

Occurrence. Trace fossils occur in sedimentary deposits of all ages from the late Precambrian to the Recent. Host rocks include limestone, sandstone, siltstone, shale, coal, and other sedimentary rocks. These deposits represent sedimentation in a broad spectrum of settings, ranging from subaerial (such as eolian dunes and soil horizons) to subaqueous (such as rivers, lakes, swamps, tidal flats, beaches, continental shelves, and the deep-sea floor). *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS.

Preservation. Organisms may produce fossilizable traces on the sediment surface (epigenic structures) or within the sediment (endogenic structures). Trace fossils may be preserved in full three-dimensional relief (either wholly contained within a rock or weathered out as a separate piece) or in partial relief (either as a depression or as a raised structure on a bedding plane). Simply because a trace fossil is preserved on a bedding plane does not indicate that it originally was an epigenic trace.

Adolf Seilacher, a German paleontologist, proposed the following terminology to describe the

Fig. 1. Domichnial dwelling burrow with pelleted wall, probably created by a burrowing crustacean. *Ophiomorpha*, Cretaceous, North Dakota. (*From W. Häntzschel, in C. Teichert, ed., Treatise on Invertebrate Paleontology, pt. W, revised, University of Kansas Press, 1975*)

preservational mode of trace fossils: full relief (enclosed entirely within the sediment); positive epirelief (ridge, mound, or other raised structure on the upper surface of a bed); negative epirelief (groove, pit, or other indentation on the upper surface of a bed); positive hyporelief (raised structure on the sole of a bed); and negative hyporelief (indentation on the sole of a bed). Anders Martinsson, a Swedish paleontologist, proposed an alternative system to characterize the preservation of trace fossils: endichnia (enclosed entirely within the sediment); epichnia (exposed in relief on the upper surface of a bed); hypichnia (exposed in relief on the sole of a bed); and exichnia (entirely removed from the sediment in which it was produced).

Diagenetic alteration of sediment commonly enhances the preservation of trace fossils by differential cementation or selective mineralization. In some cases, trace fossils have been preferentially replaced by chert, dolomite, pyrite, glauconite, apatite, siderite, or other minerals. *See* DIAGENESIS.

Classification and nomenclature. The study of trace fossils is known as ichnology. The prefix "ichno-" (as in ichnofossil and ichnotaxonomy) and the suffix "-ichnia" (as in epichnia and hypichnia) commonly are employed to designate subjects relating to trace fossils. The suffix "-ichnus" commonly is attached to the ichnogenus name of many trace fossils (as in *Dimorphichnus* and *Teichichnus*).

In the nineteenth century, many trace fossils were mistakenly identified as fossil plants, so they were given the genus and species names of plants in accordance with established principles of Linnean taxonomy. Subsequent recognition that these fossils actually were biogenic sedimentary structures that represented animal activity in the sediment has not stopped the practice of assigning formal taxonomic names to trace fossils. Usually, geologists differentiate between trace fossils and body fossils by the terms "ichnogenus" and "ichnospecies" when speaking of trace fossils. Rules of the International Code of Zoological Nomenclature generally apply to trace fossils at the genus and species level. Although some workers have proposed higher taxonomic levels for trace fossils, such as ichnofamilies or ichnophyla, none of these have gained universal acceptance. *See* TAXONOMY.

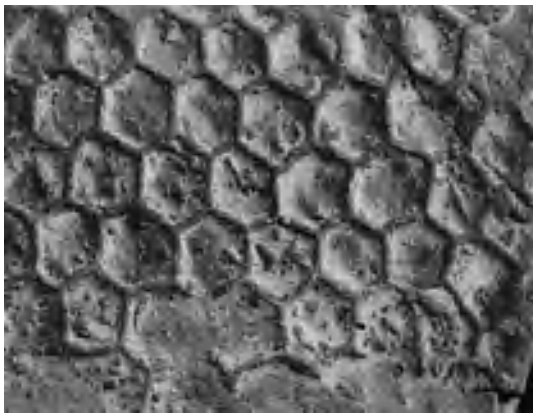

Fig. 3. Agrichnial farming trace of an unknown organism, composed of a hexagonal, meshlike network of tunnels. *Paleodictyon*, Tertiary, Poland. (*Photograph by W. Häntzschel*)

Fig. 2. Highly branched, fodinichnial mining burrow, possibly produced by a deposit-feeding crustacean or worm. (*a*) *Chondrites*, Tertiary, Austria. (*b*) *Chondrites*, Tertiary, Spain. (*From W. Häntzschel, in C. Teichert, ed., Treatise on Invertebrate Paleontology, pt. W, revised, University of Kansas Press, 1975*)

Fossil behavior. Trace fossils provide tangible information of the activities of ancient organisms, because they represent particular behavior patterns related to dwelling, feeding, locomotion, and resting. Ichnologists have established several behavioral categories of trace fossils, including the following eight groups that are most widely recognized today (**Figs. 1–10**): domichnia (dwelling structures, such as permanent burrows or agglutinated tubes); fodinichnia (burrows produced in the process of mining the sediment for food); agrichnia (burrows produced in order to farm or trap food inside the sediment); praedichnia (traces of predation); pascichnia (feeding trails, either within the sediment or on the sediment surface); repichnia (locomotion trails and trackways); fugichnia (escape traces, usually produced by an animal crawling out from beneath a rapidly deposited pile of sediment); and cubichnia (resting or nesting traces).

Fig. 4. Agrichnial farming traces of unknown organisms, including a double-spiral tunnel (*Spirorhaphe*) and a meshlike network of tunnels (*Paleodictyon*). Tertiary, Austria. (*Photograph by W. Häntzschel*)

Fig. 5. Loosely meandering, agrichnial farming trace of an unknown organism. *Cosmorhaphe*, Tertiary, Poland. (*From W. Häntzschel, in C. Teichert, ed., Treatise on Invertebrate Paleontology, pt. W, revised, University of Kansas Press, 1975*)

Environmental implications. Trace fossils are useful to geologists as indicators of ancient environments of deposition. Recurrent assemblages of trace fossils that represent certain environmental conditions, such as water depth, salinity, or character of the sea floor, are known as ichnofacies. Ichnofacies are named after common ichnogenera that exemplify this association. *See* FACIES (GEOLOGY).

Adolf Seilacher established a bathymetric zonation of universal ichnofacies representing unconsolidated sediments in marine settings, which can be found throughout the geologic column all over the world. The *Skolithos* ichnofacies typically represents nearshore, often intertidal, environments characterized by well-sorted clastic sediments that are dominated by primary sedimentary structures. Most of the trace fossils are domichnia, repichnia, and fugichnia. The *Cruziana* ichnofacies represents offshore settings, generally within wave base (that is, in water shallow enough for waves to move sediment grains on the sea floor). The trace fossils include an abundance of domichnia, repichnia, fugichnia, cubichnia, pascichnia, or fodinichnia. The *Zoophycos* ichnofacies usually represents quiet-water conditions far from shore, often on a submarine slope. The trace fossils are characterized by a low-diversity assemblage of fodinichnia. The *Nereites* ichnofacies represents fine-grained, distal turbidites that were deposited in relatively deep water. The trace fossils consist mainly of pascichnia and agrichnia. Another deep-water trace fossil association, simply referred to as the deep-sea ichnofacies, represents pelagic sedimentary environments, and it is characterized by a moderate-diversity assemblage of fodinichnia that were deeply emplaced within the sediment. *See* MARINE SEDIMENTS; TURBIDITE.

Other marine ichnofacies have been established for substrates that were not unconsolidated sediment. The *Glossifungites* ichnofacies represents very firm substrates (highly compacted but uncemented sediment). The *Trypanites* ichnofacies represents fully lithified substrates (cemented

Fig. 6. Praedichnial boring drilled in a bivalve shell by a carnivorous gastropod. *Oichnus*, Recent, Texas. (*Photograph by A. A. Ekdale*)

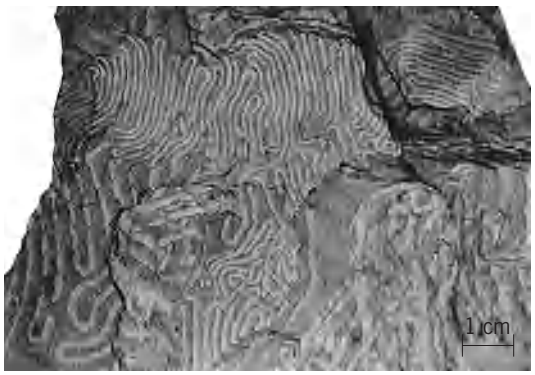

Fig. 7. Tightly meandering, pascichnial grazing trail, created by an unknown worm. *Helminthoida*, Tertiary, Austria. (*From W. Häntzschel, in C. Teichert, ed., Treatise on Invertebrate Paleontology, pt. W, revised, University of Kansas Press, 1975*)

Fig. 8. Repichnial trail of an arthropod, possibly a trilobite. *Cruziana*, Cambrian, Alberta. (*Photograph by J. P. A. Magwood*)

sedimentary rock and calcareous shell material). The *Teredolites* ichnofacies represents wood substrates. In all three cases, the dominant trace fossils are domichnia, namely borings, which were excavated by organisms with very special adaptations for penetrating the harder substrates.

Seilacher grouped all trace fossil associations in continental settings into one ichnofacies, the *Scoyenia* ichnofacies. Although it is widely recognized that several nonmarine ichnofacies exist in the geologic record, no precise delineation of these varied ichnofacies has achieved universal acceptance. Some workers have established various local ichnofacies that have yet to be accepted as having universal application.

Ichnofabric. The activities of burrowing and boring organisms can profoundly affect many aspects of the texture and internal structure of a sedimentary deposit, as sediment grains are sorted, modified, and redistributed by infaunal (living within the sediment below the sediment surface) animals. Sediment fabrics that result from bioturbation (vertical mixing of sediment by the burrowing activities of animals) and bioerosion activities are called ichnofabrics. In many situations, such as in pelagic carbonate deposits, continuous sedimentation and simultaneous bioturbation allow for the superimposition of different suites of organism traces, thus producing composite ichnofabrics (**Fig. 11**). Ichnofabric analysis can shed light on the paleoecology of the infaunal community, including trophic relationships and tiering structure. Ichnofabrics also allow the stability and firmness of the original substrate to be interpreted by examining the distinctness and degree of deformation of trace fossils. Interstitial oxygen conditions in the original sediment may be deciphered from the abundance and preservational modes of deep-tier burrows. Early diagenetic processes, including differential cementation and secondary mineralization, can enhance the preservation and visibility of ichnofabrics.

Sedimentologic implications. Trace fossils reflect the interplay among the three important sedimentologic processes of deposition, erosion, and burrowing of the sediment by organisms (causing disturbance or obliteration of primary stratification). Slow, continuous deposition, as occurs in the oxygen-rich water of offshore shelf and deep-sea environments, usually is accompanied by a total burrowing of the sediment. Numerous trace fossils, especially domichnia, fodinichnia, and pascichnia, characterize such situations. In contrast, slow, continuous deposition in eutrophic lakes or restricted marine basins that contain oxygen-depleted bottom water will yield laminated sediment without trace fossils, because the available oxygen is insufficient to support bottom-dwelling animals. Rapid, continuous deposition, as occurs in prograding beaches and laterally accreting point bars, commonly is reflected by sparse trace fossils (mainly fugichnia, cubichnia, and domichnia) superimposed on a sedimentary fabric of primary bed forms and primary sedimentary structures.

Discontinuous deposition usually results from the alternation of slow and rapid sedimentation events, as exemplified by turbidites, or from the alternation of rapid sedimentation and erosion events, as occurs during major storms along a marine coastline. In the former case, the fine-grained units of a turbidite represent a lengthy period of slow deposition, during which numerous kinds of organisms lived on and

Fig. 9. Unnamed escape structure (fugichnial trace) in sandstone. Cretaceous, Utah. (*Photograph by A. A. Ekdale*)

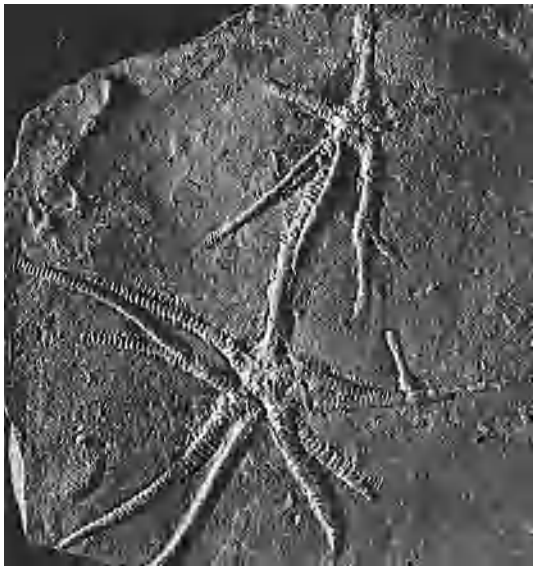

Fig. 10. Cubichnial resting trace of an ophiuroid brittle star. *Asteriacites*, Jurassic, Germany. (*Photograph by W. Häntzschel*)

Fig. 11. Complex, composite ichnofabric created by numerous successive phases of burrowing in a fine-grained pelagic carbonate deposit, which has been weakly cemented to become chalk. Upper Cretaceous, Denmark. (*Photograph by A. A. Ekdale and R. G. Bromley*)

in the sea floor, creating a wide variety of fossilizable traces (especially pascichnia, fodinichnia, and agrichnia). The coarser-grained unit of the same turbidite represents sudden deposition by a turbidity current, and this unit is characterized by domichnia of organisms that colonized the new sediment immediately after the turbidite event. Thus, turbidite sequences actually contain two separate generations of trace fossils: a predepositional trace fossil association in the fine-grained units and a postdepositional trace fossil association in the coarse-grained units.

In the latter case, laminated-to-burrowed sedimentary sequences represent the effects of storms, which erode the sea bottom and resuspend the sediment. As a storm subsides, sediment is redeposited, and the organism community is reestablished. The vertical transition from laminated sediment layers to burrowed sediment layers reflects a declining sedimentation rate, and the top of the sequence is bounded by an erosional unconformity that marks the next erosive storm event. *See* FOSSIL; PALEOECOLOGY; PALEONTOLOGY; SEDIMENTOLOGY.

A. A. Ekdale

Bibliography. R. G. Bromley, *Trace Fossils: Biology, Taphonomy and Applications.*, 2d ed., Chapman and Hall, 1996; A. A. Ekdale, *Palaeogeog. Palaeoclimatol. Palaeoecol.*, 50:63–81, 1985; A. A. Ekdale, *Palaios*, 3:464–472, 1988; A. A. Ekdale, R. G. Bromley, and S. G. Pemberton, *Ichnology: Trace Fossils in Sedimentology and Stratigraphy*, 1984; W. Häntzschel, in C. Teichert (ed.), *Treatise on Invertebrate Paleontology*, pt. W, revised, University of Kansas Press, 1975.

Trachylina

An order of jellyfish of the class Hydrozoa of the phylum Cnidaria. These jellyfish are of moderate size. They differ from other hydrozoan jellyfish in having balancing organs which develop partly from the digestive epithelium and in having only a small polyp stage or none at all. Many authorities now recognize three distinct orders of trachylines, Limnomedusae, Trachymedusae, and Narcomedusae, and in this case the older term Trachylina is abandoned.

Limnomedusae have a small polyp stage which produces other polyps and medusae by budding. The tentacles are hollow. The best-known Limnomedusae are *Gonionemus* and *Craspedacusta*. *Gonionemus* is widely used in zoology classes as a representative jellyfish and is described in most textbooks. It is often abundant in salt-water ponds and bays. *Craspedacusta* lives in fresh water and occurs in nearly all parts of the world.

Trachymedusae (see **illus.**) and Narcomedusae are jellyfish of the open seas. Their tentacles, unlike

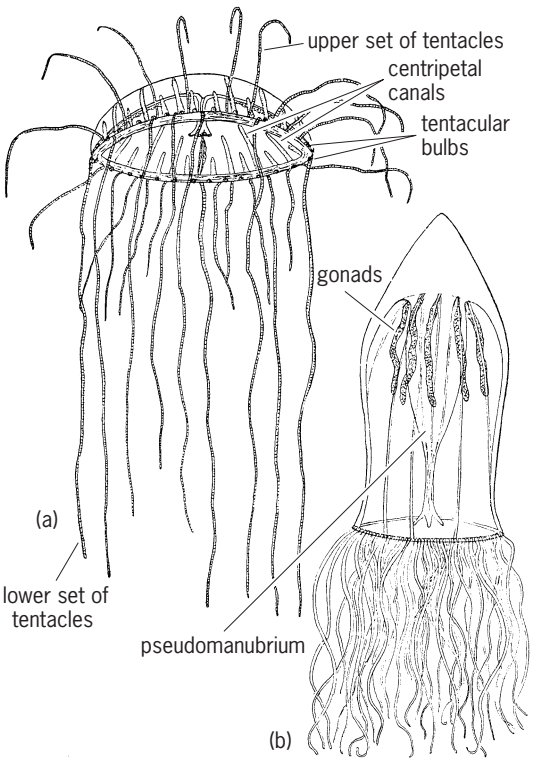

Trachymedusae, (*a*) *Olindias*, Bermuda. (*b*) *Aglantha*, Puget Sound. (*After L. H. Hyman, The Invertebrates, vol. 1, McGraw-Hill, 1940*)

those of Limnomedusae, have a solid core consisting of a single row of endodermal cells, and there is no polyp stage.

Narcomedusae differ from both Limnomedusae and Trachymedusae in having broad and often lobed stomachs; there are no peripheral canals, and tentacles are attached above the margin of the bell, which is heavily grooved and lobed. *See* HYDROZOA.

Sears Crowell

Bibliography. S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; F. S. Russell, *The Medusae of the British Isles*, 1954.

Trachyte

A light-colored, aphanitic (very finely crystalline) rock of volcanic origin, composed largely of alkali feldspar with minor amounts of dark-colored (mafic) minerals (biotite, hornblende, or pyroxene). If sodic plagioclase (oligoclase or andesine) exceeds the quantity of alkali feldspar, the rock is called latite. Trachyte and latite are chemically equivalent to syenite and monzonite, respectively. *See* SYENITE.

Texture. The extremely fine-grained texture and more or less glassy material are due to rapid cooling and solidification of the lava. Large crystals (phenocrysts) are commonly sprinkled liberally through the dense rock, giving it a porphyritic texture. These may be well formed and 1–2 in. (2.5–5 cm) wide. They appear as glassy crystals of sanidine, and in addition small mafic phenocrysts may be present. In latite the phenocrysts are largely plagioclase. As the quantity of glass increases, these porphyric rocks pass into vitrophyre; and as the abundance of phenocrysts increases, these rocks pass into trachyte porphyry. *See* PORPHYRY; VOLCANIC GLASS.

The detailed features of trachyte are best studied microscopically. Sanidine and orthoclase are dominant over oligoclase in normal (potash) trachyte. In alkali (soda) trachyte, both alkali feldspar and mafics are soda-rich.

Composition. Brown biotite mica is the common mafic. It occurs as flakes which may be more or less resorbed by the liquid in the late stages of solidification so that only patches of dusty iron oxide remain. Normal trachyte commonly carries somewhat corroded and resorbed hornblende or diopside. Alkali trachyte usually contains soda-rich amphibole (riebeckite, arfvedsonite, and barkevikite) or pyroxenes (aegirine-augite or aegirite). Zoned crystals with diopsidic cores and progressively more soda-rich margins are common.

Either free silica (quartz, tridymite, or cristobalite) or feldspathoids (leucite, nepheline, or sodalite) may be present in small amounts. With increase in free silica, the rock passes into rhyolite; and with increase in feldspathoids, it passes into phonolite. Accessory minerals as tiny grains and crystals are magmetite, ilmenite, apatite, zircon, and sphene.

Structure. Streaked, banded, and fluidal structures due to flowage of the solidifying lava are commonly visible in many trachytes and may be detected by a parallel arrangement of tabular feldspar phenocrysts. A distinctive microscopic feature is trachytic texture in which the tiny, lath-shaped sanidine crystals of the rock matrix are in parallel arrangement and closely packed. This rather uniform pattern is locally interrupted where the laths more or less deviate or wrap around the phenocrysts. Orthophyric texture is common where tiny feldspar crystals show a stumpy or square outline.

Occurrence and origin. Trachyte is not an abundant rock, but it is widespread. It occurs as flows, tuffs, or small intrusives (dikes and sills). It may be associated with alkali rhyolite, latite, or phonolite.

Trachyte is commonly considered to have been derived from a basaltic magma by differentiation, a process involving removal in large quantities of early formed crystals rich in iron, magnesium, and calcium. A factor of importance in the formation of some trachyte is contamination of the original magma by incorporation of foreign rock material. The chemical transformation of andesite to trachyte may have occurred (in the solid state) where calcium was removed and sodium added metasomatically. This may explain the origin of some keratophyres (a variety of soda-rich trachyte). *See* IGNEOUS ROCKS; MAGMA; SPILITE.

Carleton A. Chapman

Tractor

A wheeled, self-propelled vehicle for hauling other vehicles or equipment and for operating the towed implements; also, a crawler which runs on an endless, self-laid track and performs similar functions.

Farm tractor. A farm tractor is a multipurpose power unit. It has a drawbar for drawing tillage tools and a power takeoff device for driving implements or operating a belt pulley as shown in the **illustration**.

The acreage to be worked, type of crops grown, and the terrain all impose their requirements on tractor design. Accordingly, models vary in such details as power generated, weight, ground clearance, turning radius, and facilities for operating equipment. All models can, however, be grouped under four general types: four-wheel, row-crop or high-wheel, tricycle, and crawler.

Power delivered at the place where it is useful is what counts; therefore, tractors are rated by the horsepower they deliver at the drawbar and at the belt. On small models, the drawbar and belt horsepower may run as low as 10 (7.5 kW); on large models the drawbar horsepower runs as high as 132 (98 kW), while belt horsepower reaches about 144 (107 kW).

The major components are engine, clutch, and transmission. These components are intimately related and designed to work in conjunction with each other to accomplish specific work.

Unlike passenger-car engines, which are of the high-speed type, tractor engines are relatively low-speed; their maximum horsepower is generated at crankshaft speeds in the neighborhood of 2000 rpm. These engines have one, two, three, four, six, or eight

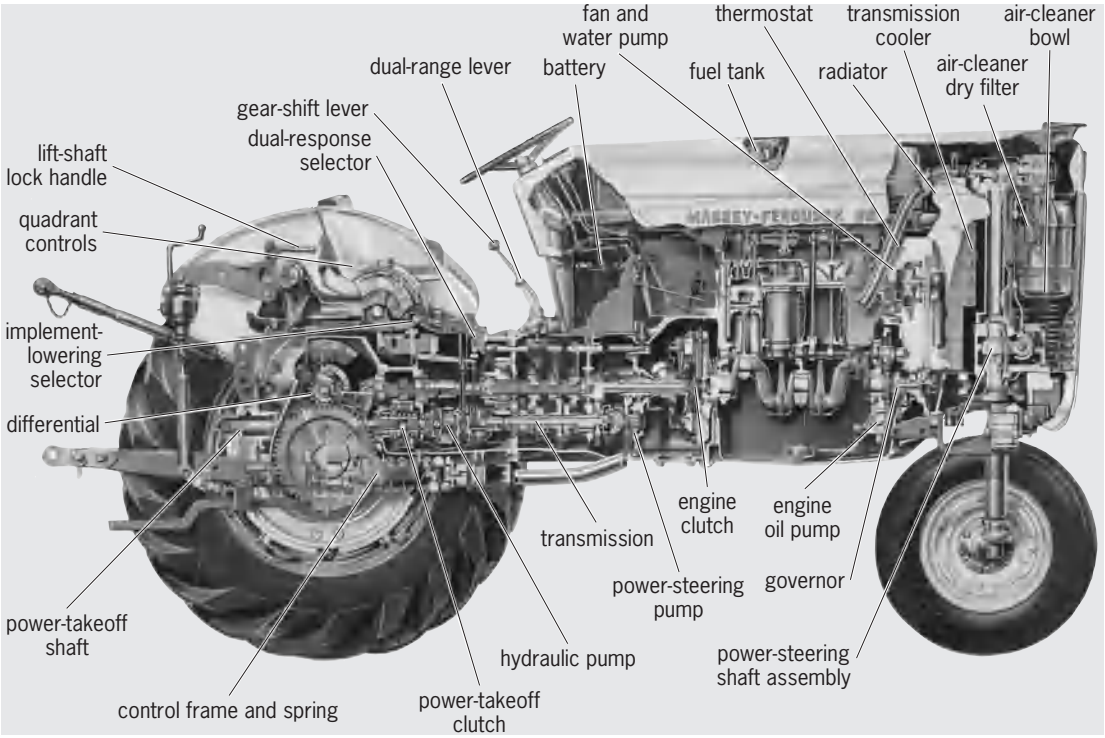

Four-wheel tractor showing major components. (*Massey-Ferguson, Inc.*)

cylinders and operate on gasoline, kerosine, liquid petroleum gas, or diesel fuel. They are of the spark-ignition or diesel type, operating on the four-stroke-cycle principle, and are cooled by water or air.

Power is transmitted to the rear wheels or to all four wheels. Drive to the front wheels is mechanical or hydrostatic, its purpose being to increase drawbar pull at the will of the operator. Transmissions have 3, 4, 5, 6, 8, 10, or 12 forward speeds and one or two reverse gears. Clutchless hydraulic transmissions are also used, making it possible to shift gears while in motion. *See* CLUTCH; TORQUE CONVERTER.

Vehicle speeds are low, ranging from slightly more than 1 mi/h(1.6 km/s) to about 18 mi/h (13 km/h) in high gear.

Between transmission and engine is a single- or multiple-disk clutch of the general automotive type, manually or hydraulically controlled.

Power is taken off the transmission by shaft for operating equipment. The power takeoff may be manually controlled or hydraulically engaged and disengaged independently of vehicle speed. It may be run at one or two set speeds, depending upon the nature of the work it is intended to do. With a pulley attached to the power takeoff, a similar choice of speeds is afforded for belt work.

Hydraulic systems are used for control of both rear- and front-mounted implements. The rear-mounted implement may be attached to the two arms of the drawbar and to a third installed arm to give a three-point hitch. An engine-driven hydraulic pump and cylinder built into the tractor provide the power to raise or lower the arms and thus to lift the implement to carry it, or lower it to the ground for work, at the will of the driver. The hydraulic mechanism may be so designed as to transfer weight from the front wheels to the rear wheels of the tractor as the load demands to give better traction when drawing multiple plows. Some systems are designed to control the depth of plow penetration automatically; others will disengage the clutch automatically if the mounted implement strikes an obstruction. The hydraulic cylinder may be attached to the implement, or may be part of it.

With the exception of crawlers, tractors are steered by turning a wheel. Hydraulic power devices are also used as manual assists.

A tractor has no frame and no springs. The supporting structure is a housing, or housings, for the transmission, clutch, hydraulic mechanism, drive shaft, and differential, to which are bolted the engine crankcase, or frame for carrying the engine, and the housing for the rear axle. The cushioning of the tractor load depends on the tires.

A typical front axle is pivoted at its center to accommodate for rough ground. Made in sections, it can be extended to change the width between wheels for straddling crop rows. The rear tread width can also be adjusted by assembling the wheel disk and rim in different positions.

Braking is by means of drum or disk brakes on the rear wheels, or by brakes on the differential. Both types permit the brakes to be applied independently on either side to assist in making sharp turns.

Crawler-type tractors in the smaller models are also used for farm work. In lieu of rubber-tired wheels, the crawler propels itself on two endless tracks made of hinged steel plates. Each track envelops a front and rear wheel; each is carried on a roller track frame which supports the weight of

the tractor. The track is made to move by engaging teeth on the periphery of the rear wheels. Dual clutches permit power to be applied independently to each track. Steering is accomplished by manipulating levers which control the clutches to send power to one track while keeping the other motionless. *See* AGRICULTURAL MACHINERY.

Industrial tractor. The basic design of an industrial tractor for hauling and for operating construction equipment departs little from that of a farm tractor, and differences in design of models fit the vehicle to its intended work.

Because high ground clearance is not needed for industrial work, the tractor is commonly built with a lower center of gravity and is capable of traveling a few miles per hour faster than a farm tractor. If its use is confined to hauling, it may not be equipped with hydraulic power. If it is to be used for operating a scraper, backhoe, or front-end loader, its structure may be heavier and more rugged. *See* BULK-HANDLING MACHINES.

Philip H. Smith

Tractrix

A plane curve for which the length of any tangent between the curve and a fixed line is constant c. If the x axis is the fixed line, its differential equation is Eq. (1). With ψ, the inclination of the tangent, as

$$\left(\frac{dy}{dx}\right)^2 = \frac{y^2}{c^2 - y^2} \tag{1}$$

parameter, this yields Eqs. (2). The tractrix has the

$$\begin{aligned} x &= c \log \tan {}^1/_2\,|\psi| + c\cos\psi \\ y &= c\sin\psi \end{aligned} \tag{2}$$

x axis as asymptote and cusps at $(0,\pm c)$. The arc AP (see **illus.**), measured from a cusp, is as in Eq. (3).

$$s = -c\ \log\sin\psi = c\ \log\frac{c}{y} \tag{3}$$

A taut string PQ, moving in a horizontal plane so that Q describes a straight line, will pull a heavy particle at P along a tractrix. The tractrix is an involute

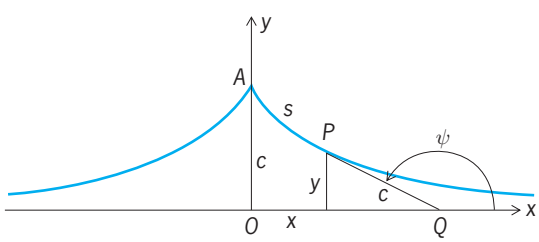

Tractrix (upper half).

of a catenary. A tractrix revolved about its asymptote generates a pseudospherical surface of constant negative gaussian curvature $-1/c^2$; its area is $4\pi c^2$, the same as a sphere of radius c. *See* CATENARY.

Louis Brand

Traffic-control systems

Systems that act to control the movement of people, goods, and vehicles in order to ensure their safe, orderly, and expeditious flow throughout the transportation system. Each of the five areas of transportation—roadways, airports and airways, railways, coastal and inland waterways, and pipelines—have unique systems of control.

Roadway traffic control. Roadway traffic-control systems are intended to improve safety, increase the operational efficiency and capacity of the roadway, and contribute to the traveler's comfort and convenience. They range from simple control at isolated intersections using signs and markings, to sophisticated traffic-control centers which have the ability to react to changes in the traffic environment. Traffic-control systems are used at roadway intersections, on highways and freeways, at ramp entrances to freeways, and in monitoring and controlling wider-area transportation networks. Intelligent roadway traffic control, now known as intelligent transportation systems (ITS), is a very sophisticated form of traffic control for roadway and other areas.

Traffic-control devices. Traffic-control devices are the primary means of communication with road users, including drivers, pedestrians, and bicyclists. Devices placed on, adjacent to, or above the roadway fulfill three general functions: guidance, regulation, and warning. Guidance devices provide information and direction, including mileage to destinations, points of interest, and route designations. Regulatory devices provide information concerning applicable traffic regulations and laws, including speed limits and turn prohibitions. Warning devices alert road users to potentially hazardous conditions. Traffic-control devices can also be thought of as active or passive. Passive devices, such as signs and markings, provide the same guidance at all times. Active devices, such as traffic signals and changeable message signs, provide guidance to drivers in response to changes in the roadway environment.

Traffic signs are the most common traffic-control devices. Signs are usually mounted above or beside the roadway and convey information by means of words or symbols. Specific colors and shapes have been reserved for each functional type of sign. In the United States, guidance signs are rectangular, with white letters and symbols on a green background. Warning signs are usually diamond shaped, with black markings on a yellow background. Regulatory signs, also rectangular, have black markings on a white background. The message or symbol upon the sign and the size and placement of the sign itself are carefully determined to provide consistent, understandable information to road users.

Changeable message signs present lighted, changeable messages to travelers on an as-needed basis. They are often used to provide drivers advanced warning of congestion, construction work zones, or other hazardous conditions. Changeable message signs may be permanently mounted beside

or above the roadway, or upon trailers or trucks for temporary use.

Traffic markings provide guidance to traffic, supplement information provided by signs, and channel vehicle movements without the use of barriers or other physical equipment. Typical roadway markings include stop bars, lane lines, channelizing lines, and other guidance markings. Markings may be made with reflectorized paint or with raised reflective markers.

Traffic islands are defined areas between traffic lanes for the control of vehicle movements or the accommodation of pedestrians. Traffic-control islands may be designated by paint, raised bars, curbs, guideposts, or other devices. Islands may be classified according to their principal function as pedestrian refuge islands, divisional islands, or traffic-channelizing islands. Pedestrian refuge islands provide a place of safety for pedestrians who cannot cross an entire roadway width at one time because of oncoming traffic or changing traffic signals. Traffic-divisional islands separate opposing- or same-direction traffic along the course of a highway. Divisional islands may be used to guide traffic around an obstruction in the roadway. When they are continuous, divisional islands are referred to as medians. Traffic-channelizing islands control and direct vehicle operators into the proper channels for their intended routes. Channelization is particularly useful in controlling the flow of traffic at roadways and intersections where there would otherwise be broad expanses of pavement.

Traffic signals include all power-operated devices (except signs) for regulating, directing, or warning motorists or pedestrians of traffic conditions. Although more expensive than signs, signals provide more positive guidance and increased flexibility. Traffic signals consist of illuminated, colored displays. The tricolor displays used at intersections are the most common; however, flashing yellow beacons are used to attract drivers' attention to warning signs, and red-green displays may be used on freeway ramps.

Traffic-control systems. Traffic conditions at specific locations, in conjunction with those in the wider area, determine the type and complexity of traffic control needed. The simplest roadway traffic-control systems consist of a combination of passive traffic-control devices, such as signs and markings. Increasingly complex traffic-control systems incorporate a variety of passive and active traffic-control devices, traffic monitoring technologies, and computer control systems.

Intersection traffic-control systems may employ several different control strategies. In the simplest system, signals operate in isolation and are pretimed, employing one or more preset signal patterns throughout the day. As the systems become more complex, signals along a route or in a network of intersections may be coordinated to improve traffic flow. All or part of the signal cycle at some locations may be demand actuated; for example, a left-turn arrow may be provided when sensors indicate that a vehicle is present in the turn lane.

Freeway traffic-control systems may incorporate a number of control strategies for both main lanes and high-occupancy-vehicle (HOV) lanes. These include ramp metering, lane controls, and in some cases, the use of changeable speed limits. Changeable message signs may be used to provide drivers with information on conditions, including congestion, incidents, construction, bad weather, and other hazardous situations.

Traffic monitoring systems provide information on traffic volume and speed on roadways. Inductive-loop detectors are the most common type of monitoring system. However, video-based systems for direct surveillance and those employing image analysis to extract traffic characteristics (speed and volume) are growing in use. In special circumstances, radar and sonic-based systems are also used. *See* IMAGE PROCESSING; INDUCTANCE; RADAR.

Intelligent transportation systems. Intelligent transportation systems, previously known as intelligent vehicle-highway systems (IVHS), involve the application of advanced technologies to surface transportation problems. A wide array of transportation areas are involved, including traffic and transportation management, travel demand management, advanced public transportation management, electronic payment, commercial vehicle operations, emergency services management, and advanced vehicle control and safety systems.

Advanced traffic management systems with traffic or transportation management centers (TMCs) operate in a number of urban areas. They provide integrated centers from which to efficiently manage transportation systems and mitigate the impact of accidents and other incidents. Traffic monitoring systems relay information on flow conditions to the traffic management center, where operators may manually adjust signal cycles in response to incidents, notify the appropriate emergency response personnel in case of incidents, and provide information to drivers through changeable message signs and highway advisory radio (HAR). In-vehicle navigation systems provide map and route guidance information. *See* HIGHWAY ENGINEERING.

Airport and airway traffic control. The federal government has designated airspace as either uncontrolled or controlled. In uncontrolled airspace, pilots may conduct flights without specific authorization. In controlled airspace, pilots may be required to maintain communications with the appropriate air-traffic control facility to receive authorization and instruction on traversing, taking off from, or landing, in that controlled area. Air-traffic control systems for controlled areas may be divided loosely into en route and terminal systems.

Terminal control. The level of control at terminals is determined by the designation of the airport and the type of service provided.

Runway lighting and approach lighting are provided for night operations and use during inclement weather when instrument flight rules (IFR) must be

followed. Colored lighting on runways and taxiways is used to guide ground movements. Runway approach lighting provides a visual reference for the glide path of the designated landing runway.

Instrument landing systems (ILS), used primarily during instrument flight rules conditions, provide electronic guidance to the pilot on the final approach to the runway. They provide both lateral (azimuth) and vertical (elevation) guidance to descending aircraft. *See* INSTRUMENT LANDING SYSTEM (ILS).

Microwave landing systems (MLS) are technically sophisticated instrument landing systems designed to overcome the limitations of instrument landing systems. The microwave landing system includes instrumentation for providing pilots with azimuth, elevation, and precision distance measurements to the end of the runway, and uses a time-reference scanning beam, allowing curved or multipath approaches to the active runway. The microwave landing system has seen very limited use in the United States and is being phased out. *See* MICROWAVE LANDING SYSTEM (MLS).

Precision-approach radar (PAR), located in the airport traffic-control tower, provides landing assistance to pilots. The pilot receives confirming radio information on lateral and vertical deviations from a predetermined glide path to the designated runway, and can then manually adjust the aircraft approach path as necessary. Precision-approach radar has been implemented only at military airports and a few very large civilian airports. *See* PRECISION APPROACH RADAR (PAR).

Airport surface detection equipment (ASDE) allows controllers to monitor all aircraft activity on the ground, including all runways, taxiways, and parking ramps. It is a relatively new development, available at only a few large airports in the United States. *See* AIRPORT SURFACE DETECTION EQUIPMENT.

En route control and guidance. Air route traffic control centers (ARTCCs) monitor and direct the movement of aircraft along the airways or to specific destinations. Each ARTCC controller is assigned a specific geographical area. Aircrafts are released as they enter adjoining areas. Each aircraft is tracked by using both radar and voice communications. Air route surveillance radar (ARSR), which is installed nationwide, allows controllers to precisely locate each aircraft along an airway. *See* SURVEILLANCE RADAR.

Air-traffic control radar beacon systems (ATCRBS) augment the ARSR. ATCRBS equipment includes an interrogator and transponder to provide more positive location information than the passive radar of the ARSR. The interrogator transmits a discretely coded signal requesting all transponders on that mode to respond. The controller views the coded aircraft target on the radar scope, including the plane's identification and altitude on an alphanumeric display.

En route guidance is provided by the very high-frequency multidirectional radio system, known as VORTAC. VORTAC is a combination of VOR (VHF Omnidirectional Range) and TACAN (Tactical Air Navigation). For altitudes of 24,000–60,000 ft (7315–18,228 m), airspace is under the direction of controllers using instrument flight rules, known as air positive control (APC). In heavily traveled corridors (for example, in the northeast), control is exercised above 18,000 ft (5486 m). *See* RHO-THETA SYSTEM; TACAN.

Inertial navigation systems are used by large commercial transports and some military aircraft for long-range travel, especially transoceanic flights. An on-board inertial device computes latitudes and longitudes, allowing pilots to pinpoint their locations along the airways and to provide controllers with accurate position reports. *See* INERTIAL GUIDANCE SYSTEM.

The Global Positioning System (GPS) is often used in aircraft tracking and navigation systems. Aircraft receiving corrections from the Differential Global Positioning System (DGPS) can pinpoint their positions to an accuracy of a few meters. The Global Positioning System has applications to both air and surface surveillance. *See* SATELLITE NAVIGATION SYSTEMS.

Advanced Air Traffic Management System. The Advanced Air Traffic Management System (ATMS) is an automated system that balances air-traffic departures with airspace system capabilities. Flight information and traffic counts are entered into a central flow-control facility, where they are compared to threshold capacity levels at various airports. Traffic managers are alerted to predicted congestion situations in specific areas of the national airspace. Managers can take action to mitigate airport and airway congestion and subsequent delay problems by directing aircraft to different routes and airports, or by holding aircraft on the ground. *See* AIR NAVIGATION; AIR-TRAFFIC CONTROL; AIR TRANSPORTATION.

Railroad traffic control. Railroads operate high-speed freight and passenger services essentially over an exclusive right of way. Railroads use both semaphore and light signals for traffic control. Semaphores convey visual messages to train operators according to predetermined rules indicating how the train is to be operated in specified areas.

Automatic block signaling prevents rear-end and head-on collisions on signal tracks. In this system, track sections are divided into blocks. Only one train is permitted to occupy a block at any time. Blocks are monitored by automatic circuitry that controls traffic signals, indicating the appropriate clear or stop signals to following or approaching trains. Similar block systems are used for subway systems.

The need to keep rail freight cars moving and to determine the location of individual cars has led to the development of an extensive traffic-control system both within terminals and en route. Centralized traffic-control systems may control hundreds of miles of track signals and switches. A dispatcher at a central location monitors the location of trains by means of visual displays of colored lights on a large track diagram, and can operate the switches and signals at key points from the central control console.

Automatic car identification systems electronically scan and record the information stored on an optically coded plate attached to the side of each freight

car. This information is transmitted to each railroad and to a centralized computer system called the Telerail Automated Information Network II (TRAIN II). This network connects individual railroads and permits a balancing of full and empty cars among the railroads and geographical regions. *See* RAILROAD CONTROL SYSTEMS; RAILROAD ENGINEERING.

Vessel traffic control. The volume of traffic on waterways is generally quite low. Marine aides, therefore, function more for informational, advisory, and guidance purposes than as positive traffic-control devices. Lighted or unlighted buoys indicate navigable areas in coastal waters and within waterways. Navigators accurately determine their positions by referencing sequentially numbered and colored-coded buoys to navigation charts. *See* BUOY; PILOTING.

Lightships and lighthouses with fog signals and radio beacons are placed as markers at prominent points during periods of limited visibility. Each station has a characteristic identification signal. Vessels can determine their positions in coastal waters by timing the interval between flashes of the fog light and the sounding of the horn signal. *See* LIGHTHOUSE.

Radar devices have become more common, even on smaller ships. They provide navigation and traffic control under conditions of poor visibility when it is difficulty to determine the location of the ship and to avoid other ships. Navigation systems often employ the Differential Global Positioning System.

The Vessel Traffic Service (VTS) is available in selected areas. Services may range from the provision of single advisory messages to extensive management of traffic communication and radar services. The Vessel Traffic Service is particularly useful in approaches to ports, access channels, and other areas with high traffic densities; for traffic carrying dangerous cargoes, in especially narrow channels; and in areas of environmental sensitivity. *See* MARINE NAVIGATION; NAVIGATION.

Pipeline traffic control. The 450,000 mi (750,000 km) of pipelines in the United States are a major part of the nation's transportation network, carrying about 25% of all intercity freight-ton mileage. The primary goods moved through pipelines are oil and oil by-products, natural gas, and fertilizers. Solid products, such as coal, may be moved through slurry pipelines.

The movement of goods in pipelines is controlled by systems of valves, pumps, and compressors. Pipe diameters generally range from 6 in. (15 cm) to 30–40 in. (75–100 cm). Pumping stations are generally positioned every 30–70 mi (50–110 km) along the pipeline, depending on the terrain and viscosity of the materials transported.

Slurry pipelines maintain a homogeneous mixture of a solid product and liquid to flow at approximately 3–5 mi/h (5–8 km/h). The speed is carefully calculated for the particular material being transported. If the material moves too fast, the walls of the pipe may be abraded. If it moves too slowly, the solids drop out of suspension. Large amounts of water are required for the operation of coal slurry pipelines.

The Trans-Alaska Pipeline System transports crude oil at about 7.5 mi/h (12 km/h), carrying 2×10^6 barrels/day (3×10^5 m^3/day). This requires 12 pumping stations along the 800-mi (1300-km) route. Much of the pipeline is above ground to avoid melting the permafrost. In those locations where below-ground installation was necessary because of the terrain or to permit animal crossing, refrigerant lines were located alongside the pipe to keep the warm moving oil (at about 135°F or 57°C) from melting the permafrost. *See* PIPELINE; TRANSPORTATION ENGINEERING. James Costantino; Donna C. Nelson

Bibliography. K. J. Button and D. A. Hensher (eds.), *Handbook of Transport Systems and Traffic Control*, 2001; J. D. Fricker and R. K. Whitford, *Fundamentals of Transportation Engineering: A Multimodal Approach*, 2003; N. J. Garber and L. A. Hoel, *Traffic and Highway Engineering*, 3d ed., 2001; S. Ghosh and T. Lee, *Intelligent Transportation Systems: New Principles and Architectures*, 2000; M. Kutz (ed.), *Handbook of Transportation Engineering*, 2003; J. M. Sussman, *Perspectives on Intelligent Transportation Systems*, 2005.

Trajectory

The curve described by a body moving through space, as a meteor moving through the atmosphere, a planet around the Sun, a projectile fired from a gun, or a rocket in flight. In general, the trajectory of a body in a gravitational field is a conic section—ellipse, hyperbola, or parabola—depending on the energy of the motion. The trajectory of a shell or rocket fired from the ground is a portion of an ellipse with the Earth's center as one focus; however, if the altitude reached is not great, the effect of gravity is essentially constant, and the parabola is a good approximation. *See* BALLISTICS. John P. Hagen

Tranquilizer

A psychopharmacologic drug that tends to calm overexcited individuals, producing a state of "tranquility." The term tranquilizer was originally applied to two groups of drugs. Members of one group, the antianxiety drugs, were called minor tranquilizers, and members of the second group, the antipsychotic drugs, were called major tranquilizers. Although drugs in both groups may have a calming effect in appropriate doses, it is now clear that the two groups are quite different.

Antianxiety drugs. This group of drugs, also known as anxiolytics, comprises three subgroups: propanediols, of which meprobamate (Miltown) is the best known; alcohol (ethanol) and the barbiturates, such as phenobarbital; benzodiazepines, of which diazepam (Valium) is the best known.

Like all drugs, antianxiety drugs have a range of effects at different doses, including—in the low-to-high order of doses at which the effects generally occur—(1) a period of transient excitement for some

persons; (2) a reduction in anxiety, the effect for which the drugs are prescribed; (3) antagonism of certain kinds of seizure activity in the brain, a so-called anticonvulsant activity; (4) a sedative or sleep-inducing effect; and (5) a generalized reduction in muscle tension, a muscle relaxant effect.

An important difference among the three groups of antianxiety drugs is based on this dose dependency: for the first two groups, the difference in doses that reduce anxiety and those that produce sedation is small; a major advantage of the third group is that this difference is large, so that they reduce anxiety with relatively less sedation. Furthermore, any sedation that may occur initially with a benzodiazepine tends to wane with repeated administration (an effect known as tolerance), whereas its antianxiety effect is less likely to do so. This phenomenon contributes to an increased separation of antianxiety and sedative effects. This differential, coupled with the relative lack of serious complications resulting from overdose, accounts for the enormous popularity of the benzodiazepines in the clinical management of anxiety.

The benzodiazepines modify brain activity by interacting with gamma-aminobutyric acid (GABA), a naturally occurring substance that inhibits the activity of brain cells by increasing their permeability to chloride ions; these ions decrease the likelihood that the cell will be responsive to input from other cells. The benzodiazepines increase the tendency for GABA to act at its own receptors, and thus have an indirect inhibitory role that is GABA-dependent. Ethanol and the barbiturates, on the other hand, act directly on the membrane to increase chloride permeability.

Gamma-amino butyric acid and changes in chloride permeability are also involved in the muscle relaxant, anticonvulsant, and sedative effects of these drugs. It is not clear, however, whether these interactions can account directly for their antianxiety effects. Meprobamate, for example, reduces anxiety but has not been shown to augment GABA; increased GABA activity may therefore be a sufficient but not necessary condition for anxiety reduction. On the other hand, another drug (valproic acid) can augment the activity of GABA but has not been shown to have clinical antianxiety effects; the benzodiazepines may therefore engage some process that acts in concert with GABA but that is not engaged by valproic acid. *See* AFFECTIVE DISORDERS.

Antipsychotic drugs. The control of psychotic symptoms by these drugs wrought a revolution in mental health care when they were introduced in the mid-1950s: chronic treatment with antipsychotics reversed, for the first time, the progressive increase in hospitalization of patients. Because of the symptom control that antipsychotics provide, it is possible to manage the majority of psychotic patients (mostly schizophrenics) on an outpatient basis. *See* SCHIZOPHRENIA.

There are a large number of antipsychotic drugs, all of which are equally useful in managing psychotic symptoms. However, the antipsychotics vary widely in potency, that is, the amount of drug that is required to produce effective symptom control. Furthermore, different side effects are associated with different potencies: a low-potency drug such as chlorpromazine (Thorazine), for example, is more likely to produce sedation and certain other side effects than is a high-potency drug such as haloperidol (Haldol); haloperidol is more likely to produce disturbances of movement than is chlorpromazine.

The major complication of treatment with antipsychotic drugs is the array of movement disturbances that they can induce. Most of these disturbances can be controlled by concurrent medication with other drugs. For example, one type of reaction mimics the signs of Parkinson's disease; it occurs in about 30% of patients, although only transiently in many cases. This type of reaction can be successfully managed by adjunctive treatment with one of the several drugs useful in the control of parkinsonian symptoms. On the other hand, another type of reaction that may occur in up to 25% of patients in the course of several years of treatment, tardive dyskinesia, has thus far proven to be refractory to adjunctive medication. Therefore, the patient may face a trade-off between disturbed movement and the debilitating symptoms of psychosis.

These disturbances are clearly related to a drug-induced attenuation in the activity of a substance called dopamine, which occurs naturally in various areas of the brain and is critically involved in the mediation of normal movement. It is less clear, however, as to whether the reduction in psychotic symptoms produced by these drugs is also due to attenuated dopamine activity. *See* PSYCHOPHARMACOLOGY; PSYCHOSIS.

Peter L. Carlton

Bibliography. R. J. Baldessarini, *Chemotherapy in Psychiatry*, 1985; H. Y. Meltzer (ed.), *Psychopharmacology: The Third Generation of Progress*, 1987.

Transamination

Transfer of an amino group from one carbon chain to another without the intermediate formation of ammonia. Typically, an α-amino acid serves as the donor of the amino group and is converted to an α-keto acid, and an α-keto acid serves as the acceptor of the amino group and is converted to an α-amino acid, as in reaction (1), where R_1 and R_2 represent two different side chains.

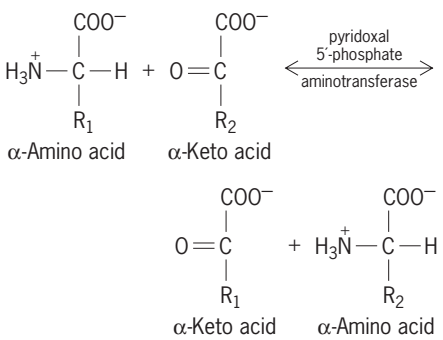

(1)

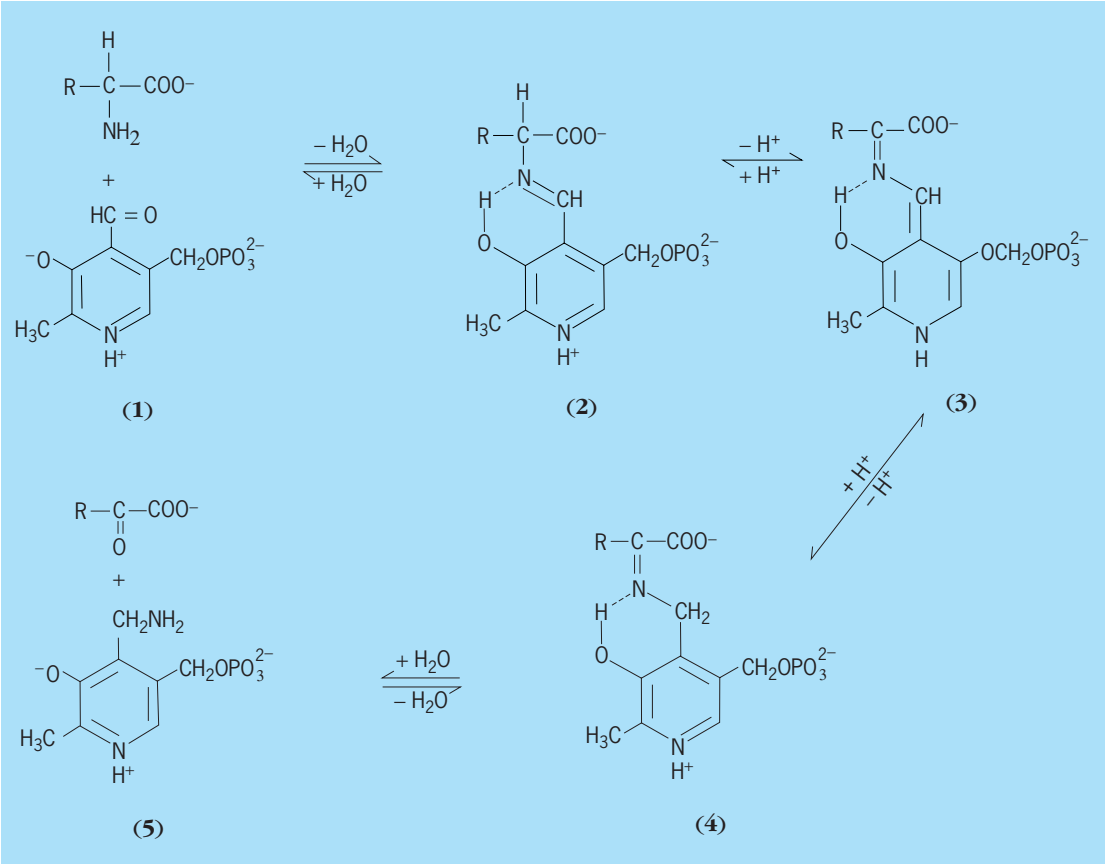

Fig. 1. Pathway of reactions for a transamination.

Enzyme-catalyzed reactions of this type were first discovered by A. E. Braunstein and M. G. Kritzman in 1937. Transamination reactions are catalyzed by enzymes called aminotransferases (or transaminases) that require pyridoxal 5′-phosphate as a cofactor. The pyridoxal 5′-phosphate functions alternately as an amino group acceptor [in the aldehyde form, PLP, as in reaction (2a)] and as an amino group donor [in the amine form, PMP, as in reaction (2b)], where

$$\text{Amino acid}_1 + \text{PLP} \rightarrow \text{keto acid}_1 + \text{PMP} \qquad (2a)$$

$$\text{Keto acid}_2 + \text{PMP} \rightarrow \text{amino acid}_2 + \text{PLP} \qquad (2b)$$

subscripts 1 and 2 indicate different amino (or keto) acids. **Figure 1** illustrates the chemistry of these reactions.

Most physiologically important aminotransferases have a preferred amino acid/keto acid substrate and utilize α-ketoglutarate/glutamate as the counter keto acid/amino acid. An example is aspartate aminotransferase (usually abbreviated AST for aspartate transaminase or SGOT for serum glutamate-oxaloacetate transaminase), which accepts aspartate or oxaloacetate as substrate and uses glutamate or α-ketoglutarate as cosubstrate (**Fig. 2**).

Amino acid metabolism. In human tissues, the amino acids alanine, aspartate, glutamate, tyrosine, serine, valine, isoleucine, and leucine are actively transaminated. Histidine, phenylalanine, methionine, cysteine, glutamine, asparagine, and glycine also may undergo transamination in human tissues; these amino acids are primarily metabolized by other types of reactions under normal physiological conditions, but intermediates in the degradation pathways of phenylalanine, cysteine, glutamine, asparagine, and glycine (that is, tyrosine formed from phenylalanine, cysteinesulfinate formed from cysteine, glutamate formed from glutamine, aspartate formed from asparagine, and serine formed from glycine) are actively transaminated. Threonine, lysine, tryptophan, proline, and arginine do not directly participate

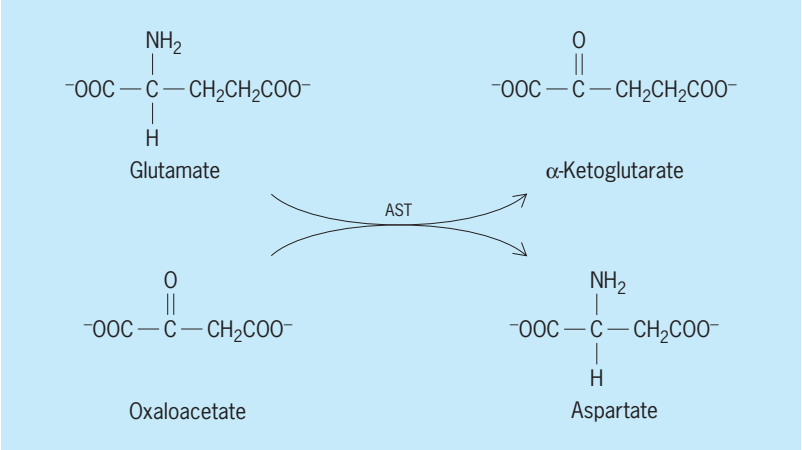

Fig. 2. Transamination reaction catalyzed by aspartate aminotransferase (AST).

in transamination reactions in mammalian tissues; nevertheless, intermediates in the degradation pathways of threonine (glycine), lysine (α-aminoadipate), tryptophan (3-hydroxykynurenine and alanine), and proline/arginine (ornithine and glutamate) do undergo transamination for removal of amino groups from the carbon chain. The interconversion of ornithine and glutamate γ-semialdehyde represents a more unusual type of transamination reaction in which one of the reactants is not an α-amino acid/α-keto acid couple. In this specific case, the δ-amino group of ornithine is transferred by ornithine δ-aminotransferase. *See* AMINO ACIDS.

Amino acid catabolism. Transamination reactions play a prominent role in the synthesis of the dispensable (or nonessential) amino acids in higher organisms. Amino groups are directly transferred to the keto acids pyruvate, oxaloacetate, and α-ketoglutarate to form alanine, aspartate, and glutamate, respectively. Transamination is also involved in the pathways of de novo synthesis of serine, glycine, glutamine, asparagine, proline, and arginine.

Aspartate. Because α-ketoglutarate is used so widely as the acceptor of amino groups in transamination reactions, the α-amino groups of numerous amino acids are funneled through glutamate in the process of amino acid catabolism. Aspartate aminotransferases are widespread in tissues, so amino groups are readily interchanged between glutamate and aspartate. Amino groups from various amino acids are funneled, via transamination reactions, to aspartate for entrance into the urea cycle. Aspartate is also an important donor of amino groups for purine nucleotide synthesis and of both carbon and nitrogen for pyrimidine nucleotide synthesis. The mitochondrial and cytosolic isozymes of AST play important roles in the movement of carbon and reducing equivalents across mitochondrial membranes by the malate-aspartate shuttle. During ureagenesis (urea synthesis) and gluconeogenesis (production of glucose from amino acid carbon chains or other nonglucose substrates such as lactate and glycerol) in the liver, aspartate carries carbon and reducing equivalents for hepatic gluconeogenesis as well as nitrogen for ureagenesis from the mitochondria to the cytosol. *See* LIVER.

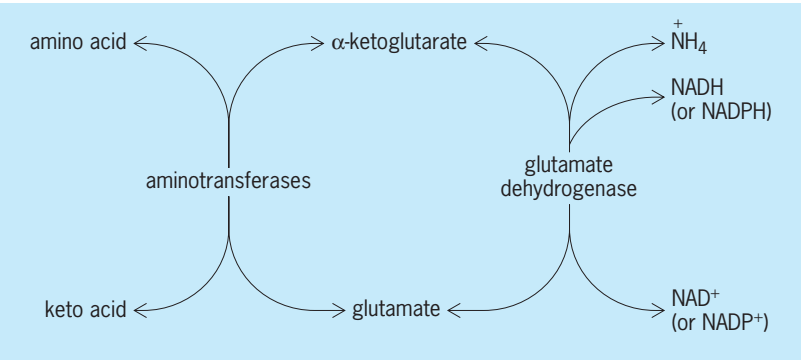

Fig. 3. Transamination reaction in which glutamate/α-ketoglutarate serves as cosubstrate coupled with the glutamate dehydrogenase reaction.

Alanine. Alanine aminotransferase (usually abbreviated ALT for alanine transaminase or SGPT for serum glutamate-pyruvate transaminase) catalyzes a reaction close to equilibrium in the liver, skeletal muscle, and the small intestine. Both alanine and glutamine are important carriers of amino groups derived from muscle amino acid catabolism to the liver for urea synthesis. Alanine is also released by the small intestine as a result of its catabolism of glutamine, glutamate, and aspartate. The liver metabolizes alanine by transamination, and uses the carbon chain for gluconeogenesis and the amino group for urea synthesis.

Glutamate. The movement of nitrogen between organic and inorganic pools is mediated by the glutamate/α-ketoglutarate couple and a reaction catalyzed by glutamate dehydrogenase, as shown in reaction (3).

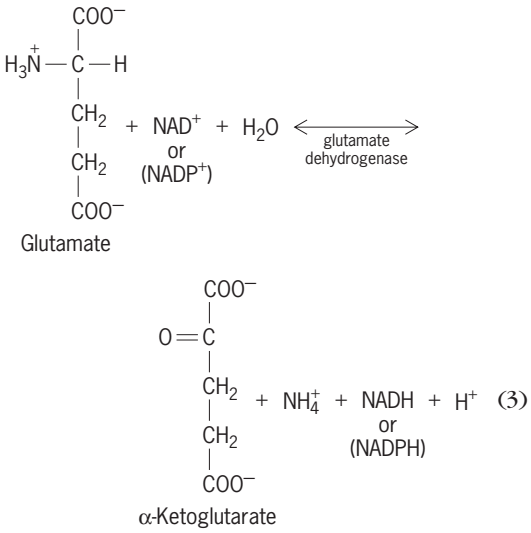

(3)

Glutamate dehydrogenase is a mitochondrial enzyme that catalyzes a near-equilibrium reaction in tissues with high activity (liver, kidney, brain), and it can operate to either incorporate ammonia into or release ammonia from the α-amino acid pool. The direction of flux depends on the provision and removal of reactants. The equilibrium nature of the glutamate dehydrogenase and transamination reactions in these tissues acts to maintain intracellular ammonia levels in a narrow range.

Transamination reactions in which the glutamate/α-ketoglutarate couple serve as cosubstrate coupled with the glutamate dehydrogenase reaction allow the α-amino group of most amino acids to be ultimately released as ammonia (**Fig. 3**). This ammonia may be incorporated into glutamine, which donates amide groups for various synthetic reactions, or incorporated into urea via carbamoyl phosphate synthesis. Ammonia/ammonium is also excreted in the urine as such and plays an important role in acid-base balance. *See* NITROGEN EXCRETION; URINE.

Aminotransferases as clinical indicators. Plasma or serum levels of ALT and AST are frequently used as clinical indicators of tissue damage. These enzymes leak out of damaged cells. Because alanine aminotransferase is highest in liver, a high plasma level of

this aminotransferase is usually an indicator of liver damage. High levels of aspartate aminotransferase are less specific because this enzyme can be released by damaged liver, muscle, heart, kidney, and brain tissue. *See* ENZYME; PROTEIN METABOLISM.

Martha H. Stipanuk

Bibliography. A. J. Cooper and A. Meister, An appreciation of Professor Alexander E. Braunstein: The discovery and scope of enzymatic transamination, *Biochimie*, 71:387–404, 1989; M. H. Stipanuk and M. Watford, Amino acid metabolism, in M. H. Stipanuk (ed.), *Biochemical and Physiological Aspects of Human Nutrition*, 2d ed., Saunders, Philadephia, 2006.

Transcription

The process that occurs within a living cell in which an enzyme makes a ribonucleic acid (RNA) copy of the deoxyribonucleic acid (DNA) that contains the genetic information. The resulting RNA is often used as a template to make proteins by the cellular protein-synthesizing (translation) machinery. Transcription is an essential step in the growth and differentiation of all cells and contributes to almost every aspect of the development of an organism. Because transcription is so fundamental to life, there is significant conservation of proteins and enzymatic steps involved in transcription in cells of bacterial, archaeal, and eukaryotic origin. An RNA polymerase catalyzes synthesis of RNA from DNA by binding to a sequence termed the promoter located adjacent to or upstream of the coding portion of the linked gene. Genetic and biochemical studies have defined distinct steps in the transcriptional process. The first step of the transcription cycle is initiation, which involves melting (strand separation) of the template DNA and formation of the first chemical bond that makes the RNA. Like a revving engine, the RNA polymerase carries out several rounds of synthesis of short RNA transcripts (termed abortive transcripts) until it becomes disengaged from the promoter (promoter escape) and moves along the DNA, synthesizing RNA at an estimated rate of 1500 bases per minute. This step of the transcription cycle is called elongation, which is followed by termination and re-initiation. Transcription is a highly regulated process; each step requires participation of accessory proteins and is subject to control by transcription factors that respond to intra- and extracellular signals. *See* DEOXYRIBONUCLEIC ACID (DNA); ENZYME; GENE; PROTEIN; RIBONUCLEIC ACID (RNA).

Bacterial transcription. In bacteria, a promoter is defined by two DNA sequence motifs that are located at 10 and 35 bases upstream of the transcriptional start site and are bound by a protein called the sigma factor. The sigma factor is a promoter selectivity factor; most gene promoters in *Escherichia coli* are bound by sigma 54, but additional sigma factors exist that direct transcription from a subset of bacterial genes. In addition to the sigma factor, the bacterial RNA polymerase comprises four subunits. In the presence of ribonucleotide triphosphates, the promoter-bound enzyme alone can initiate RNA synthesis. The transcription from certain genes is turned on and off in response to various signals. This involves proteins called transcriptional activators and repressors that associate with the promoter DNA and influence the activity of the RNA polymerase, resulting in increased or decreased rates of RNA synthesis. In bacteria, transcription factors are commonly partnered with another protein that acts as a sensor of environmental stimuli; this two-component system coordinately regulates RNA synthesis of multiple genes in response to a particular signal. Bacterial genes that encode proteins in a common pathway are transcribed as part of a single large RNA transcript from an operon (a group of distinct genes that are expressed and regulated as a unit) that is translated into individual proteins. In bacteria, transcription and translation are coupled, and therefore regulation of transcription ensures the coordinated synthesis of proteins. *See* BACTERIAL GENETICS; OPERON.

Eukaryotes. The large size of the eukaryotic genome and the complex programs of gene expression required to build multicellular organisms necessitated evolution of more elaborate transcriptional machineries and regulatory schemes, although the basic principles and mechanisms of RNA synthesis remain similar to those in bacteria. Eukaryotes encode three RNA polymerases that transcribe distinct classes of genes. Although the composition of each enzyme is more complex (12 subunits) than the bacterial polymerase, the largest subunits of each eukaryotic RNA polymerase, which carry the copying function, closely resemble the subunits of the bacterial enzyme. RNA polymerase I (Pol I) transcribes a single gene present in many copies that encodes the ribosomal RNA precursor. This takes place in a specialized structure within the cell nucleus called the nucleolus. RNA polymerase III (Pol III) transcribes genes coding for small RNAs such as transfer RNA. RNA polymerase II (Pol II) is responsible for transcribing mostly protein-encoding genes (~30,000 in mammals) and a variety of small noncoding RNA genes. Transcription by Pol II and Pol III takes place in the nucleus, and the RNA is transported to the cytoplasm where protein synthesis takes place. Thus, unlike bacteria, transcription in eukaryotes is not coupled to translation. *See* CELL NUCLEUS; CYTOPLASM.

Regulatory DNA. Eukaryotic gene promoters contain multiple DNA sequence motifs that serve as docking sites for the transcription machinery. To initiate transcription, proteins that specifically associate with the promoter region bind to DNA and in turn recruit other proteins in the transcription machinery that include RNA polymerase to form a stable protein–DNA complex called the pre-initiation complex. The promoter sequences and the factor requirement for each RNA polymerase are distinct; however, mechanistically they work in a similar manner. Each enzyme is brought to the promoter by contacting a protein complex that recognizes and binds

the promoter sequence. The TATA box-binding protein (TBP) serves a function similar to that of the bacterial sigma factor, and it is required for transcription by all three eukaryotic RNA polymerases. TBP was discovered as a transcription factor that binds to an adenine/thymine–rich sequence (TATA box) found in many protein-coding gene promoters transcribed by Pol II. It was later found to be a part of several distinct multiprotein complexes, each serving a specific function for each class of RNA polymerase by binding to promoter DNA. Further, TBP associates with transcription factor IIB (TFIIB) or B-related factor (BRF) to recruit Pol II and Pol III, respectively. The functional importance of these proteins in transcription is underscored by the fact that transcription in archaea requires two proteins related to TBP and TFIIB in addition to RNA polymerase. As in bacteria, the transcription machinery assembled at the eukaryotic gene promoter undergoes each step in the transcription cycle. The DNA template is melted, and the polymerase begins to catalyze the RNA-synthesizing reaction, which initially results in abortive transcripts. A chemical modification on the polymerase results in its dissociation from the promoter-bound transcription factors and transition to the elongation phase that requires a new set of proteins called elongation factors that associate with the enzyme complex. As the polymerase reaches the end of the gene, termination of transcription occurs, and the enzyme dissociates from the template to be recycled for the next round of transcription (re-initiation). It should be noted that the newly synthesized RNA must be appropriately modified and processed prior to being transported to the cytoplasm. Many proteins carry out distinct RNA-processing reactions, and it is now known that these proteins act in concert with the transcription machinery to produce mature RNA in a highly coordinated fashion. *See* NUCLEIC ACID.

Regulation by chromatin. A typical eukaryotic nucleus has a diameter of 6–10 micrometers and contains DNA of about 2 m (6.5 ft) in length. To fit inside the nucleus, the DNA is subject to an extraordinary degree of compaction. Histones are highly conserved eukaryotic proteins that serve to help package DNA. About 146 base pairs of DNA wrap around a core of eight histone proteins to form nucleosomes, which associate with additional proteins and undergo further compaction to form a chromatin fiber. Chromatin is inhibitory not only to transcription but to any cellular process involving DNA such as replication, repair, and recombination. In order to synthesize RNA or to "activate a gene," transcription machineries must contend with chromatin, a process not commonly found in bacteria. Eukaryotes make proteins that specialize in locally altering the structure of chromatin to permit transcription factors and RNA polymerase to bind to DNA and synthesize RNA. Two distinct classes of protein complexes participate in this process: chromatin-modifying complexes that contain enzymes that modify histones in a way to loosen the compacted structure of the nucleosomes; and chromatin-remodeling complexes that use the energy of adenosine triphosphate (ATP) hydrolysis to slide or displace histones from chromatin. In this way these protein complexes facilitate binding and assembly of the transcription initiation complex at the promoter and synthesis of RNA. *See* CHROMOSOME; NUCLEOPROTEIN; NUCLEOSOME.

Gene activation. How is a gene turned on from a transcriptionally silent state in chromatin? Current models based on well-studied genes propose the following scenario (see **illustration**). A transcriptional activator that is capable of recognizing its target DNA sequence when it is associated with chromatin binds to a regulatory region of a gene. The activators are sequence-specific DNA-binding transcription factors that when bound to DNA influence directly or indirectly the RNA polymerase transcription machinery to enhance transcription from the linked gene. The DNA-bound transcriptional activators initially make contact with chromatin-modifying

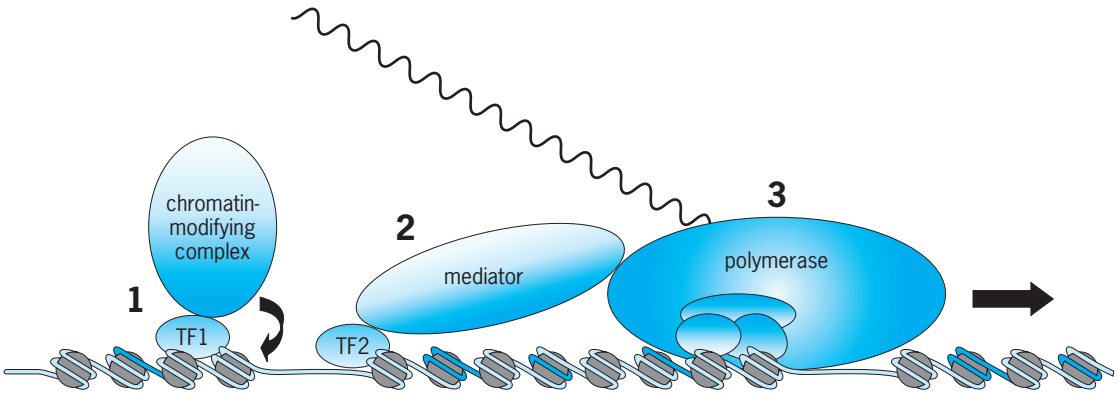

Transcriptional activation of a eukaryotic gene. Shown is a transcriptional regulatory region (promoter) of a gene; the DNA is wrapped around the nucleosomes. (Step 1) A DNA-binding transcription factor (TF1) binds to a sequence upstream of the transcriptional start site and recruits a chromatin-modifying or -remodeling complex, which alters the structure of the nearby chromatin (curved arrow). (2) This facilitates binding of another transcription factor (TF2) to its target site, which in turn recruits a transcriptional cofactor complex (mediator). (3) The mediator interacts with components of the basal transcription machinery, and a pre-initiation complex is formed at the promoter. Synthesis of RNA (wavy line) by the polymerase ensues (horizontal arrow). The relative importance of these steps and the factors involved in activation (or repression) is specific for each gene.

and/or -remodeling proteins and recruit them so that the nearby region in the chromatin may be altered. This facilitates recruitment of other transcription factors and ultimately the basal transcription machinery, including RNA polymerase to the promoter region to initiate RNA synthesis. Transcriptional activators also interact with intermediary factors or cofactors that do not contact DNA directly but associate with other proteins and components of the transcriptional machinery to assemble a specific protein–DNA complex at the promoter. The participation of dozens of transcription factors allows for combinatorial control and regulatory diversity that is critical for gene-specific regulation of transcription. Transcriptional repressors function to inhibit expression of certain genes or maintain them in a silent state. The mechanisms of repression may involve interfering with the function of activators or the basal machinery, or recruitment of co-repressors that modify histones in order to induce more compacted chromatin structure.

Disease. Modulation of transcription impacts on many areas of medicine, both as a cause and as a cure of disease. Mutations in transcription factors or in gene regulatory sequences have been implicated in numerous developmental and neurological disorders. An example is Rett syndrome, which involves mutations in a protein that regulates transcription. It is now becoming clear that these mutations subtly alter the global program of gene expression in a way that kills a small but vitally important subset of neurons in the brain. Cancer may be thought of as a transcriptional disease, and many cancer-causing proteins (oncoproteins and mutated tumor suppressors) function as transcription factors. The majority of human tumor cells carry mutations that directly affect the function of two critical transcription factors, p53 and the retinoblastoma protein. In normal cells, these proteins provide a protective role by preventing damaged cells from proliferating. They achieve this by blocking the transcription of genes needed for cell duplication and by activating genes that act as a brake to proliferation. Many pharmacological drugs interact directly with DNA-binding transcription factors, leading to activation or repression of specific target genes. One example is the steroid hormones, molecules that bind to a family of receptor proteins that serve as promoter-specific transcription factors. The enzyme complexes that alter chromatin structure are also important targets for anticancer drugs, and they work because rapidly proliferating tumor cells are more sensitive to the changes in gene expression that result from the drug treatment than normal cells. *See* MUTATION; ONCOGENES.

Summary. Transcription is a process essential to any living cell in which the genetic information encoded in the DNA is copied to RNA by the enzyme RNA polymerase. Studies show that the basic mechanism of transcription is essentially conserved among bacteria, archaea, and eukaryotes. However, eukaryotes possess a larger complement of transcription factors and more elaborate machineries and regulatory mechanisms to deal with transcription of a larger number of genes. To maintain the differentiated state, a mammalian cell may transcribe only one-half of its gene repertoire at any given time. The chromatin plays a critical role in keeping the remaining genes silent and permitting specific genes to be activated during processes such as differentiation or the cell cycle. More than 95% of the mammalian genome seems not to code for any genes, and yet they harbor DNA sequences that control transcription of many genes from a long distance. Future studies will be geared toward uncovering the vast amounts of regulatory information (instructions) encoded within eukaryotic genomes to help define the transcriptional regulatory networks responsible for cell differentiation and development of an organism.

Naoko Tanese; Angus C. Wilson

Bibliography. S. Hahn, Structure and mechanism of the RNA polymerase II transcription machinery, *Nat. Struc. Mol. Biol.*, 11(5):394–403, 2004; J. T. Kadonaga, Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors, *Cell*, 116(2):247–257, 2004; M. Levine and R. Tjian, Transcription regulation and animal diversity, *Nature*, 424(6945):147–151, 2003.

Transducer

A device that converts variations in one energy form into corresponding variations in another, usually electrical form. Measurement transducers or input transducers may exploit a wide range of physical, chemical, or biological effects to achieve transduction, and their design principles usually revolve around high sensitivity and minimum disturbance to the measurand, that is, the quantity to be measured. Output transducers or actuators are designed to achieve some end effect, for example, opening of a valve or deflection of a control surface on an aircraft. Actuators, therefore, normally operate at high power levels. The term sensor is often used instead of transducer, but strictly a sensor does not involve energy transformation; the term should be reserved for devices such as a thermistor, which is not energy-changing but simply changes its intrinsic electrical resistance in response to changes in temperature.

Vast numbers of transducers are in common use in the domestic environment, that is, in central heating systems, ovens, and refrigerators; in transport; in all forms of industry; and in medicine (see **table**). Extremely sophisticated devices are employed for the measurements needed in the pursuit of scientific research.

Both input and output transducers, together with the instrumentation to which they are connected, may be called upon to respond to both quasistatic (that is, slowly varying) signals or dynamic (that is, more rapidly varying) signals. This means that the transducer, together with its instrumentation system, must be designed to meet such a specification. Some prior knowledge is therefore required of the type of signal to be transduced, and the bandwidth

Principles of transducers and examples of their applications

Variable	Transduction principle	Example
Temperature	Thermoelectric	Thermocouple; thermopile
	Thermocapacitive	Change of capacitance (of, for example, ceramic capacitors)
	Thermoresistive	Thermistors; platinum resistance thermometer; semiconductor (silicon temperature sensors)
	Thermal expansion	Glass thermometer; bimetallic strip; quartz crystal (frequency change)
	Thermal radiation	Thermopile; thermistor; photovoltaic semiconductor; photoconductive semiconductor
	Thermochemical	Liquid crystals (temperature-sensitive chemical changes)
	Pyroelectric	Lead zirconate titanate (polarization change with temperature)
Displacement		
Low-mechanical-impedance types	Resistive	Potentiometer (rotational and linear)
	Capacitive	Change in capacitance
	Inductive	Change in inductance
	Transformer	Alternating-current-excited differential transformer
	Electromagnetic	Movement of coil through a magnetic field (no dc response)
	Optical	Moiré fringe; interferometry; reflection change
	Ultrasonic	Pulse transit time; phase shift
High-mechanical-impedance types	Piezoresistance	Very small displacements
	Mechanoresistance	Metal foil and wire strain gage; mercury gage
Velocity	Magnetic induction	Moving coil (gives output proportional to velocity)
	Doppler effect	Ultrasonic; optical (electromagnetic)
Acceleration	$f = ma$	Measurement of the force due to the movement of a known mass (sensing via piezoelectric, magnetostrictive, or mechanoresistive transducers)
Force	Elastic displacement	Primary: displacement of an elastic member Secondary: displacement transducer (as listed above)
	Piezoresistive	Semiconductor (for example, silicon)
	Piezoelectric	Barium titanate (ferroelectric ceramic, lacks dc response)
	Magnetostrictive	Induced voltage type (lacks dc response); variable inductance type
	Balance (feedback)	Electromagnetic force balance (null-type system)
Pressure	Elastic displacement	Diaphragm deflection (Bourdon tube, bellows, and so forth) measured with a displacement transducer
	Force balance	Manometer; electromagnetic force balance
Pressure (acoustic)	Capacitive	Moving plate (capacitor)
	Inductive	Moving coil
	Piezoelectric	Lead zirconate titanate
	Piezoresistive	Silicon diaphragm
Chemical	Gases; volatile chemicals	Piezoelectric; metal oxide; flame ionization; photionization
	Chemicals in solution	Absorption; fluorescence; luminescence
	Biochemical	Redox system; immunooptical; evanescent wave; enzyme-linked immunosystem
Atmosphere	Pollution monitoring	Light detection and ranging (LIDAR)
Flow	Pressure difference	Orifice type
	Mechanical	Rotameter: driving an electromagnetic emf-generating secondary transducer
	Thermal	Heat transport from aheated element (for example, self-heated thermistor)
	Electromagnetic	Electromagnetic flow (emf) generated by conducting liquid flowing through a magnetic field
	Ultrasonic	Pulse transit time; Doppler shift; cross correlation
	Optical	Laser Doppler shift
	Indicators	Dye dilution; thermal dilution; conductivity dilution
	Transit time	Marker transit time (bubbles, solids and so forth in carrying fluid)
Electromagnetic radiation	Gamma rays, x-rays, and so forth	Germanium or silicon detectors (cooled); silicon carbide detectors

of the transducer and instrumentation system must be suitably matched to this signal.

Transducers are often described in terms of their sensitivity to input signals; more properly this should be called responsivity. This is simply defined as the ratio of the output signal to the corresponding input signal. Once again, the responsivity of a transducer must be matched to the expected levels of signal to be transduced. *See* SENSITIVITY (ENGINEERING).

Displacement transducers. The measurement of displacement is very often part of the transduction process involved in the measurement of other quantities, such as acceleration, force, or pressure. Displacement transducers can be divided into two categories, low- and high-mechanical-impedance devices. The low-mechanical-impedance displacement transducer (large deflection, low force) is often based on the measurement of changes in electrical

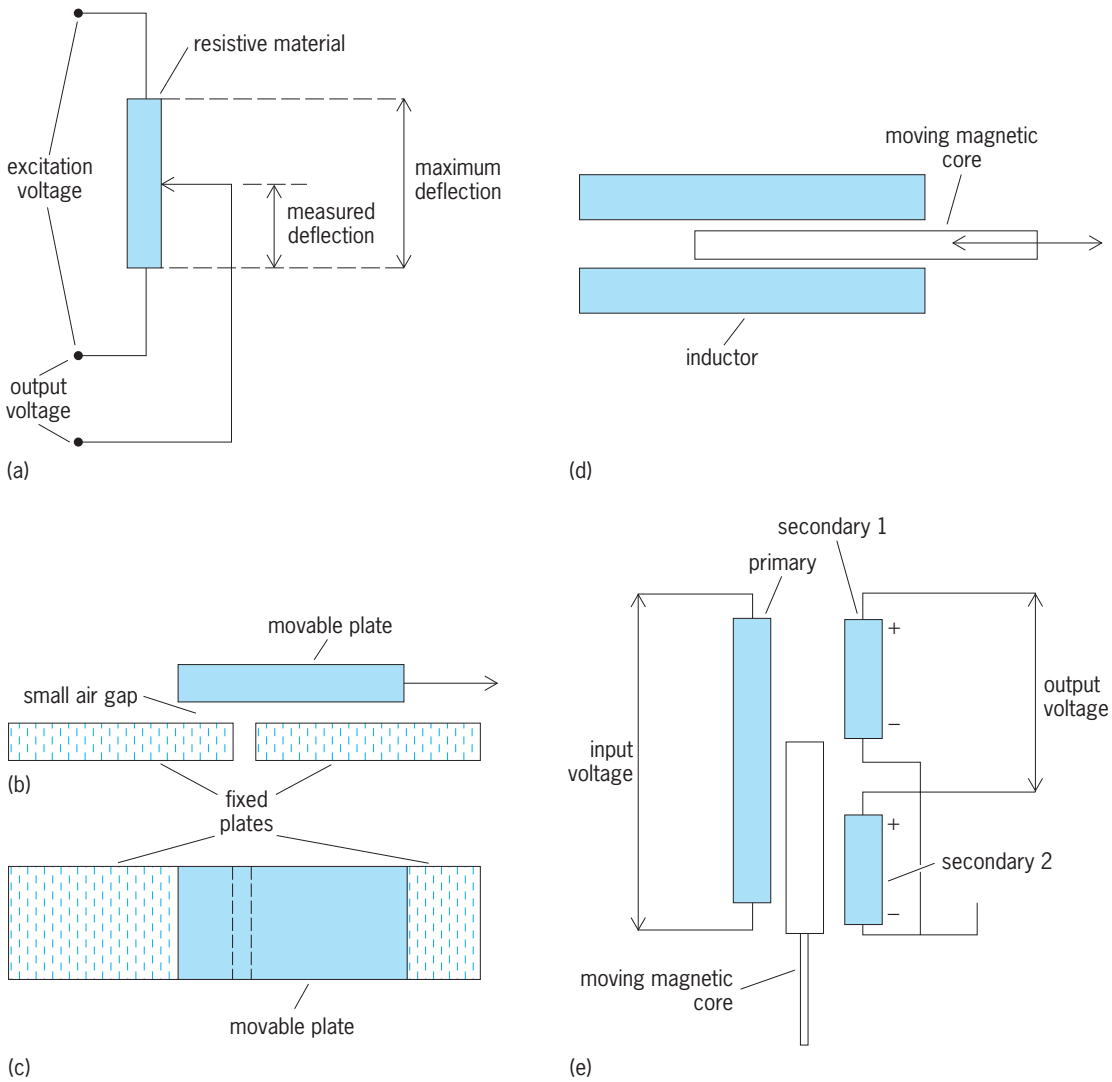

Fig. 1. Examples of displacement transducers. (*a*) Potentiometric transducer for linear deflection. Multiturn devices for rotational deflection measurement are also in common use. (*b*) Elevation and (*c*) plan view of capacitance transducer for linear deflection. Output voltage is from the moving plate with respect to the static plates. (*d*) Simple variable inductance for transducing deflection. (*e*) Circuit diagram of a linear variable differential transformer (LVDT) transducer.

resistance, capacitance, or inductance, and is usually employed for relatively large displacements, typically between a few millimeters (a fraction of an inch) up to several meters. For these devices, a restoring force is usually provided by a relatively soft spring in order not to disturb the measurand. Both linear (or translational) and rotational devices are readily available (**Fig. 1**). *See* CAPACITANCE; CAPACITANCE MEASUREMENT; INDUCTANCE MEASUREMENT; MECHANICAL IMPEDANCE; POTENTIOMETER; RESISTANCE MEASUREMENT; TRANSFORMER.

High-mechanical-impedance devices are associated with the measurement of very small displacements and might be based on, for example, the deformation of very stiff springs on which are mounted resistance strain gages. The materials used for the springs should have a linear stress–strain relationship over the range of stresses to be encountered. Other material properties, such as stability over the temperature range to be encountered, low creep, low relaxation, and low hysteresis, are also essential. Various specialty alloys have been developed for these purposes. Other high-mechanical-impedance transducers may be based on materials that are piezoresistive or piezoelectric. Piezoresistive materials such as heavily doped silicon can be used as the basis for strain gages that have very high sensitivities. *See* METAL, MECHANICAL PROPERTIES OF; PIEZOELECTRICITY; SPRING (MACHINES); STRAIN GAGE.

Force transducers. The measurement of force is very often accomplished by allowing an elastic member (spring or cantilever beam) to deflect and then measuring the deflection by using some form of displacement transducer. Two methods of achieving very high stiffness are available. One is to use an elastic member with an intrinsically small deflection for a given force. Very often, strain gages are employed to detect the small displacements involved (**Fig. 2**). The alternative method is to use feedback techniques so that the deflection caused by the

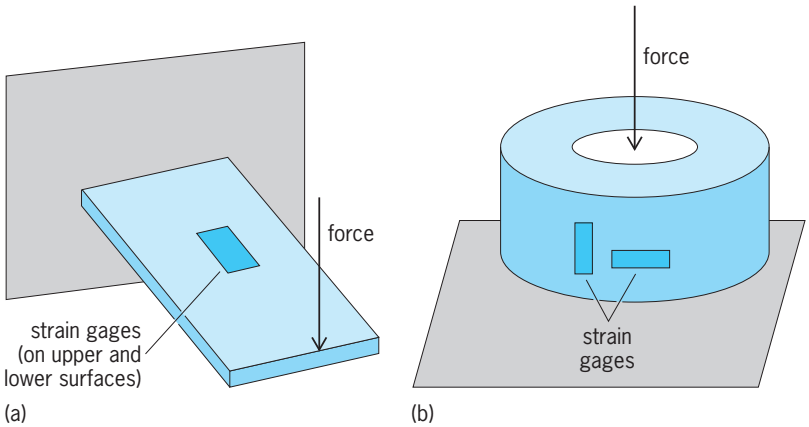

Fig. 2. Examples of force transducers. (*a*) Cantilever beam using strain gages to measure force. (*b*) Compression column plus strain gages to measure force.

application of the force generates a signal that restores the part of the transducer originally deflected to its quiescent position. Force balance systems, as these are often termed, have the advantage of good linearity. *See* LINEARITY.

Acceleration transducers. Transducers designed to measure acceleration are frequently based on the simple equation $f = ma$, where f is force, m is mass, and a is acceleration. Thus, if the force due to the movement of a known mass can be measured, it is possible to derive the acceleration. Very often, the measurement technique employed uses piezoelectric, magnetostrictive, or mechanoresistive materials (**Fig. 3**). Acceleration transducers or accelerometers are frequently employed for the measurement of vibration. *See* ACCELEROMETER; FORCE; MAGNETOSTRICTION.

Velocity transducers. Linear velocity transducers can be based on the magnetic induction effect or on the exploitation of the Doppler effect. In this latter case, the energy form used can be ultrasonic or electromagnetic. The velocity of rotating shafts can also be measured by an optical encoder with a suitable light source and detector. By choosing an appropriate pattern, the output data can be produced in binary form suitable for direct input to a computer system. This is one of the few examples of a digital transducer. (In general, the measured world is entirely analog and the transducers used to monitor it are also analog.) Rotational velocity transducers can also be based on the magnetic induction effect, using a coil rotating in the field of a permanent magnet and slip rings to detect the voltage proportional to velocity. Another approach is based on alternating currents: an exciting coil is fed with a constant-frequency signal and a detector coil picks up a voltage proportional to the rate of rotation of a cylinder, which is subject to eddy-current effects. *See* DOPPLER EFFECT; EDDY CURRENT; ELECTROMAGNETIC INDUCTION; SERVOMECHANISM.

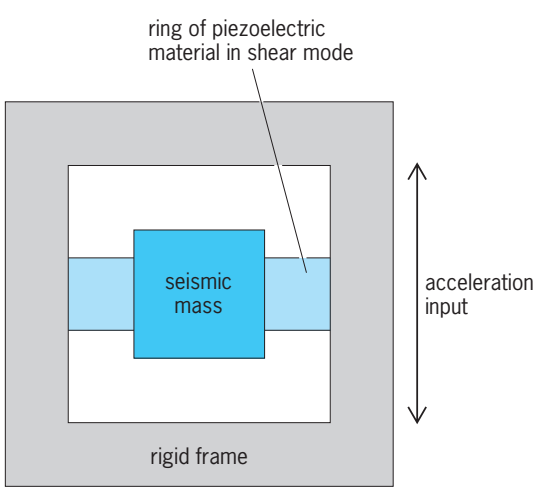

Fig. 3. Simplified diagram of a piezoelectric accelerometer.

Pressure transducers. Pressure measurement is often achieved by detecting the deflection of an elastic diaphragm or tube (for example, a Bourdon tube). The requirements in terms of material properties for such diaphragms or tubes are similar to those of restoring springs in high-mechanical-impedance displacement transducers. Semiconductor fabrication techniques have been applied to pressure transducers. A diaphragm of monocrystalline silicon possesses most of the elastic properties required, and a strain-gage bridge exploiting the piezoresistive properties of silicon can be fabricated on the diaphragm. This form of construction lends itself to microminiaturization. *See* PRESSURE MEASUREMENT; PRESSURE TRANSDUCER.

Temperature transducers. The most familiar instrument for measurement of temperature is the mercury-in-glass thermometer. Here, the measurement is really that of the expansion of the mercury as temperature increases. This expansion is magnified by employing a capillary tube in which to observe the changes in volume with temperature. This instrument has certain disadvantages as a transducer in that no electrical signal is produced. When such a signal is required, measurements may be based on either changes in the electrical resistance of certain materials or upon the thermoelectric effect.

Temperature transducers based on changes in the electrical resistance of platinum can be used over a very wide range (−200 to 900°C or −330 to 1650°F), and they are extremely stable and linear. They can be fabricated as a thin film on a ceramic substrate and are usually used with a bridge circuit. However, they produce a relatively small signal in response to temperature change (0.4 Ω/°C or 0.2 Ω/°F). *See* BRIDGE CIRCUIT.

A much higher output signal for a given temperature change is produced by devices known as thermistors, which are fabricated from semiconducting materials (for example, oxides of cobalt, nickel, and manganese). However, these are inherently nonlinear in terms of resistance change with temperature and have a relatively smaller useful range (−100 to 200°C or −150 to 390°F). They can be made fairly small and therefore have short thermal time constants, making it possible to follow rapid fluctuations in temperature. The nonlinear behavior of

thermistors can easily be overcome when they are operated in conjunction with a microprocessor. *See* MICROPROCESSOR; THERMISTOR.

Much smaller devices can be produced from thermocouples. These are transducers based on junctions between two dissimilar wires, one of which is held at a constant reference temperature and the other used to effect the measurement. The disadvantage of requiring a reference temperature can be simply overcome by electronic compensation. The disadvantage of the thermocouple is its very low output signal for a given temperature change (10–80 microvolts/°C or 5–40 μV/°F). However, a very wide temperature range is available. *See* THERMOCOUPLE.

A silicon diode has a temperature coefficient of about −2 mV/°C (−1 mV/°F) and can therefore be used as a temperature transducer, although in practice each *pn* junction needs individual calibration. However, silicon diodes have the advantage of being very inexpensive. Many other configurations of silicon-based temperature transducers have been developed using the properties of silicon, such as the temperature dependence of carrier transport, and sometimes employ numerous transistors on a single chip. *See* JUNCTION DIODE; TEMPERATURE MEASUREMENT; THERMOMETER.

Flow transducers. The measurement of flow is of considerable importance in industrial processes, and mass-flow measurement is often even more important since it may be used as the basis for metering valuable commodities (for example, oil and gas) for subsequent payment.

A wide variety of ingenious devices have been produced for flow measurement based on, for example, pitot tubes and orifice plates. The principle of both these devices is that of a pressure difference that can be transduced into a usable voltage signal. *See* FLOW MEASUREMENT.

Measurement of mass flow can be achieved by simultaneous measurement of both flow and density, but mass-flow transducers have been developed in which the liquid is made to flow through a cylinder that is forced into vibration. Differences in mass flow alter the vibrational properties of the cylinder, and these may be detected.

Chemical transducers. Transducers for a wide range of chemical species are available, but probably the most widely applied is the pH transducer for the measurement of hydrogen-ion concentration. The traditional method has relied on a glass membrane electrode used to make up an electrochemical cell. *See* HYDROGEN ION; ION-SELECTIVE MEMBRANES AND ELECTRODES; PH.

Measurements of the partial pressure of oxygen (pO_2) may be accomplished by the use of a Clark oxygen cell, which comprises a gas-permeable membrane controlling the rate of arrival of oxygen molecules at a noble-metal cathode that is held at 600–800 mV potential with respect to the anode. The ensuing reduction process gives rise to a cathode current from which oxygen concentration can be derived.

Other electrochemical transducers are based on similar principles and are used in such applications as voltametry, polarography, and amperometry. Chemical transduction is also possible by adsorbing a species onto a surface and detecting its presence by mass change, electrical property change, color change, and so on. *See* ELECTROCHEMICAL TECHNIQUES; POLAROGRAPHIC ANALYSIS; TITRATION.

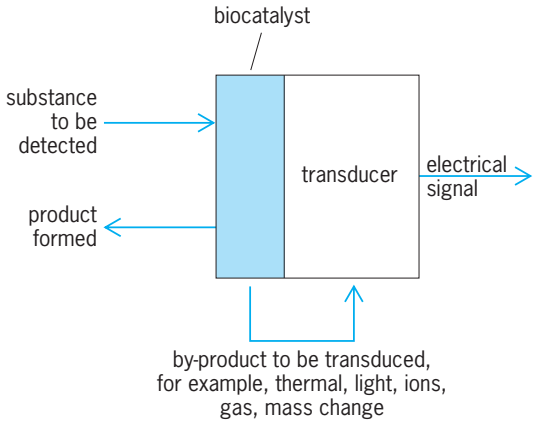

Fig. 4. Schematic diagram of a biosensor, showing the principle of operation.

Biological transducers. Measurements of the partial pressure of oxygen and the partial pressure of carbon dioxide (pCO_2) are also of particular importance in the context of blood gas analysis in medicine, and by using the Clark cell they can be performed without removing the blood from the body and noninvasively, that is, without puncturing the skin. However, it is necessary to raise the temperature of the surface of the skin onto which the Clark electrode is placed by a few degrees in order to promote permeation of oxygen (O_2) or carbon dioxide (CO_2) molecules to the cell membrane.

There have been remarkable advances in the area of biological transducers or biosensors (**Fig. 4**). Examples are the ion-selective field-effect transducer (ISFET), the insulated-gate field-effect transducer (IGFET), and the chemically sensitive field-effect transducer (CHEMFET). Many of these devices are still in the research stage, but it is almost certain that they will attain major importance in medicine, introducing a new approach to long-term patient monitoring in which a host of biochemical measurements are made at the bedside rather than in the pathology laboratory. The major advantage that this provides is speed of diagnosis and treatment. Similar devices will play a major role in measurements associated with biotechnology. Already, pH-sensitive ISFETs are widely used. The application of nanotechnology to biosensors is beginning to make possible a wide range of new applications in medical and biological research. *See* BIOELECTRONICS.

Advanced devices. Initially, the microelectronics revolution had little impact on transducer development. Efforts made to remedy this have yielded interesting results.

Solid-state and smart transducers. Considerable advantage can be obtained in many cases by placing the sensing process as close to the signal conditioning

as possible. This is particularly true of signals at very high frequencies or very low levels (because of the possibility of noise corruption in connecting cables). There have been proposals for implementing this, and silicon-based transducers are ideal for this application. These might make use of the piezoresistive, photovoltaic, or Hall effects for transduction. *See* HALL EFFECT; PHOTOVOLTAIC EFFECT.

A smart transducer or smart sensor is a device that not only undertakes measurement but also can adapt to the environment in which it is placed. Suchadaptation may range from simple changes in the characteristics of the transducer in response to changes in temperature, to more complex procedures such as adaptation of the transducer's performance to conform to overall system requirements. This may call for changes in linearity, sensitivity (or responsivity), stiffness, and so forth. In integrated transducers, much of the signal processing that might previously be done remotely is brought into the transducer packaging. This approach is most beneficial when high-frequency signals are involved, such as in an ultrasonic transducer array.

Systems of transducers. Systems have been designed that contain several transducers of various types that respond to differing parameters of the same measurand or display relatively broadband selectivity to the measurands. The cross-sensitivity between outputs of such transducers contains valuable information that may allow unique patterns of response to be obtained from such systems, and inherently poorly selective transducers to function as a group to yield highly selective information.

Fiber-optic transducers. The development of inexpensive fiber-optic materials for communications has led to an examination of the potential for using these devices as the basis for transduction. Two major types of devices have resulted: fiber-optic transducers for physical variables and similar devices devoted to chemical and biological determinations. The advantages of the all-optical transducer are its lack of susceptibility to electrical interference and its intrinsic safety. Small deformations of an optical-fiber waveguide cause a change in the light transmission of the fiber, and this has been exploited to produce force and pressure transducers. Alternatively, miniature transducers based on color chemistry can be fabricated at the end of a fiber and the color change can be sensed remotely. Devices of this type have been developed for measuring pH, the partial pressures of oxygen and carbon dioxide, and glucose. *See* FIBER-OPTIC SENSOR.

Sensors for extreme environments. Silicon carbide (also known as carborundum) is a semiconductor originally developed for high-power electronics applications. It displays a wide band gap compared with silicon which enables it to operate as a semiconductor at temperatures up to 1000°C (1800°F), whereas silicon is limited to a maximum of about 175°C (350°F). The binding energy between the carbon and silicon atoms in the lattice is much greater than that between silicon atoms, thus making silicon carbide radiation hard, leading to its use in radiation detectors. These properties also make the material useful for sensor applications in areas such as car exhaust and aerospace engine gas sensing and in the vents of volcanoes as a means of early warning of major activity. Silicon carbide has extremely low chemical activity and is thus potentially very useful for medical sensors and for the chemical engineering industry. The current challenge is to develop complex, inexpensive logic circuits for use with such sensors. Major progress has been made with wafers with few defects, and diameters up to 10 cm (4 in.) are available.

Micro-electro-mechanical systems. The most important recent technological development in the area of transducers, sensors, and actuators is micro-electro-mechanical systems (MEMS). There are a wide variety of MEMS devices, mostly fabricated in silicon. In general, MEMS-based transducers use conventional techniques to achieve transduction; for example, accelerometers are often based on a proof mass and cantilever principle with a geometry similar to that of Fig. 2*a*. At micrometer dimensions, care must be taken to ensure that a complete understanding of the mechanical behavior of such a transducer has been gained, since additional nonlinear behavior sometimes arises at these dimensions. Inexpensive silicon MEMS gyroscopes are in the second generation of commercial production, with emphasis on volume manufacture. *See* GYROSCOPE.

An injectable microsensor has been developed that is capable of transmitting signals from the motor cortex of the human brain using wireless technology. The development of such sensors is likely to be a major trend in the area of MEMS and indeed NEMS (nanometer electromechanical systems).

Considerable effort is directed at making MEMS devices that are compatible with complementary metal-oxide-semiconductor (CMOS) technology, and micromachined accelerometers giving 125 mV per *g* and a bandwidth of 2.5 kHz have been described. There is also interest in nanometer-scale devices using silicon-on-insulator (SOI) technology, giving rise to active devices working at below 1 V. MEMS-based microphones are capacitive microphones fully integrated with CMOS amplifiers and dc-to-dc converters. *See* INTEGRATED CIRCUITS; MICROPHONE; TRANSISTOR.

Microactuators using a comb drive or interdigitated electrode geometry are employed for accelerometers, gyroscopes, and micromirrors. Considerable activity is focused on the surface micromachining of micromirror hinges. Displacement sensors based on integrated optics and micro-opto-electro-mechanical systems (MOEMS) now play a major role in adaptive optical devices. *See* ADAPTIVE OPTICS; INTEGRATED OPTICS; MICRO-ELECTRO-MECHANICAL SYSTEMS (MOEMS).

Silicon micromachining is also used for constructing scanning confocal optical microscopes and, in a similar manner, modified atomic force microscopes have been described for potential use in data storage. These devices have a storage capacity of 50 times that of the CD-ROM. *See* CONFOCAL MICROSCOPY; SCANNING TUNNELING MICROSCOPE.

Gas-flow sensors have been described using air-foil plates, 30 μm thick, made of single-crystalline silicon. Other gas sensors have been micromachined to give rise to shear-stress sensors. *See* FLOW MEASUREMENT.

Micro-gas turbines based on hydrogen combustors micromachined from silicon have been manufactured. They were produced using the deep reactive ion etching (DRIE) technique, and a 150-W electrical power microcombustor has been demonstrated, where the device itself is only 0.066 cm^3 in volume. *See* GAS TURBINE.

The polymerase chain reaction (PCR) is one of the most widely used techniques for the replication of deoxyribonucleic acid (DNA) samples. PCR using micromachined structures gives rise to improved performance and much reduced reagent use. This is one of many developments leading to the so-called laboratory-on-a-chip. In addition, work on microvalve arrays has been described and, in a similar area, MEMS technology has been used to produce microneedles and microchannel-based electrical impedance spectrometers. *See* GENE AMPLIFICATION.

Considerable work is also being done on improving large-volume production techniques for MEMS-based devices. One approach is to attempt to integrate everything onto one device, as in the MEMS-based microphone. A second approach is to separate out the microelectronics, usually CMOS-based, from the microsensors and microactuators, which usually require materials that are not permitted in integrated-circuit production facilities. The two devices are then brought together, in one case using flip-chip bonding techniques. Improved production is also being sought through the use of microstamping techniques. Work is also going on to look at ways of sealing microtransducers and microsensors; in one approach, chemical vapor deposition techniques are used. *See* CONTROL SYSTEMS; CRYSTAL GROWTH; ELECTRONIC PACKAGING; INSTRUMENTATION AMPLIFIER; MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS); MICROSENSOR; PHYSICAL MEASUREMENT.

Peter A. Payne

Bibliography. R. C. Asher, *Ultrasonic Sensors for Chemical and Process Plant*, IOP Publishing, Bristol, 1997; E. O. Doebelin, *Measurement Systems: Application and Design*, 5th ed. McGraw-Hill, New York, 2003; P. Hauptmann, *Sensors: Principles and Applications*, Prentice Hall, Hemel Hempstead, UK, 1993; G. T. A. Kovacs, *Micromachined Transducers Sourcebook*, WCB/McGraw-Hill, Boston and London, 1998; H. K. P. Neubert, *Instrument Transducers: An Introduction to Their Performance and Design*, 2d ed., Oxford University Press, Oxford, 1975; S. Soloman, *Sensors Handbook*, McGraw-Hill, New York, 1998; T. Togawa, T. Tamura, and P. Å. Öberg, *Biomedical Transducers and Instruments*, CRC Press, Boca Raton, 1997; P. F. Turner, I. Karube, and G. S. Wilson (eds.), *Biosensors: Fundamentals and Applications*, Oxford University Press, New York, 1987; M. J. Usher and D. A. Keating, *Sensors and Transducers: Characteristics, Applications, Instrumentation, Interfacing*, 2d ed, Macmillan,London, 1996.

Transduction (bacteria)

A mechanism for the transfer of genetic material between cells. The material is transferred by virus particles called bacteriophages (in the case of bacteria), or phages. The transfer method differentiates transduction from transformation. In transformation the genetic material (deoxyribonucleic acid) is extracted from the cell by chemical means or released by lysis. *See* BACTERIOPHAGE; DEOXYRIBONUCLEIC ACID (DNA); TRANSFORMATION (BACTERIA).

Transduction has been demonstrated only in certain bacterial species, such as *Salmonella*, *Escherichia*, *Staphylococcus*, and *Haemophilus*. Even within these species it occurs only within certain strains. There is, however, no reason to assume that it is not more prevalent, and it may be that only technical limitations, such as the appropriate phage, prevent its demonstration in other forms.

The transduction mechanism has two features to distinguish it from the more usual mechanism of gene recombination, the sexual process. The most striking feature is the transfer of genetic material from cell to cell by viruses. The second distinguishing feature is the fact that only a small part of the total genetic material of any one bacterial cell is carried by any particular transducing particle. However, in general transduction, all of the genetic material is distributed among different particles.

Transducing bacteriophage. Transduction is not accomplished by all bacteriophages. It is done by some that are classified as "temperate." When such temperate bacteriophages infect sensitive bacteria, some of the bacteria respond by producing more bacteriophage particles. These bacteria donate the transducing material. Other bacteria respond to the infection by becoming more or less permanent carriers of the bacteriophage, in a kind of symbiotic relationship; these are called lysogenic bacteria. Bacteria in this latter class survive the infection, and it is among these that transduced cells are found. The proportion of bacteria in any culture that responds to infection in either manner can be influenced by the particular environment at the time of infection. Thus, when the goal is to obtain transducing activity, the culture is influenced to grow more bacteriophage. When the study of the transduction is desired, the culture is influenced to give more lysogenic responses. *See* LYSOGENY.

General transduction. General transduction is unrestricted as to the genetic property which can be transferred. The only limitation results from the fact that the frequency of cells transduced for any particular property is rather low. It is convenient to express the transducing activity of a preparation in terms of the number of transductions produced per particle capable of acting as an infective agent, or the number of transductions per viable viral particle. For any character, this number is of the order of one

transduction for every million particles. Since only relatively few bacteriophage particles can infect a single bacterium, the number of transduced bacteria for the studied property is not much larger than one in a million. However, with bacterial populations it is relatively simple to find bacteria of a unique kind even at such a low frequency. Bacteria can be grown to numbers a thousand times more than a million. If the property transduced can be disclosed by selection, there is no problem in finding the transduced cells.

Linked transduction. Although the general rule is that individual properties are transduced independently of each other and occasionally replace their genetic homologs, instances of simultaneous or linked transduction of two properties are found, indicating that the transducing fragments encompass more than one gene. Transduction has afforded investigators an opportunity to look carefully at small regions of genetic material. These studies have shown that very often genes controlling sequential biosynthetic steps are linked; that is, they are very close together on the bacterial chromosome. Transduction has also been of use in further demonstrating that the gene controlling an individual biosynthetic step can be subdivided by mutation and then reconstituted by genetic recombination. Thus many mutations, all affecting the synthesis of a specific enzyme, can be shown to lie at different points on the same gene. They can transduce each other to a nonmutant form. *See* BACTERIAL GENETICS.

Restricted transduction and oncogenesis. In addition to general transduction, certain phages carry out a more restricted kind of transduction. The restriction is that they carry only a specific section of bacterial genetic material; they transduce only a few genes. In this restricted case it has been shown that the transducing particle has exchanged a piece of its own genome for a piece of bacterial genome. The particular piece of bacterial genome is the one that the phage is linked to when it is carried by the bacteria.

Transduction thus affords investigators further insight into the nature of viruses and the fine structure of genetic material. *See* VIRUS.

Retroviruses also carry out specific or restricted transduction. It has long been known that these viruses can cause the formation of tumors (oncogenesis) in animals. (The first Rous sarcoma, a tumor virus that affects chickens, was discovered in 1905.) It is now known that these viruses exchange a small portion of their genome for a mutant cellular gene that has a role in gene regulation or replication. These viruses carrying mutant genes infect cells, causing them to be transformed into tumor cells. *See* ANIMAL VIRUS; RETROVIRUS; ROUS SARCOMA.

Transduction and evolution. Analysis of the DNA sequences of many different microorganisms has revealed that they are mosaics of different genomes, indicating that there was a great deal of horizontal gene transfer in their evolution. Transduction was probably the major vehicle for such exchanges, as many bacteria contain a myriad of pieces of phage genes.

Norton D. Zinder

Transfer cells

Plant cells characterized by the elaboration of an unlignified, secondary cell wall to form fingerlike projections or wall ingrowths which protrude into the cytoplasm of the cell (**Figs. 1** and **2**). These ingrowths are enveloped by plasma membrane (Fig. 2), forming a wall-membrane apparatus. The extent of wall ingrowth development varies from small, conical projections, through long, unbranched or branched structures, to Y-shaped flanges. In all cases their formation increases the surface-volume ratio of the cell, which implicates them in short-distance solute transport. The surface area of the plasma membrane may be increased up to 20 times that of a similar-sized cell lacking wall ingrowths. Most transfer cells elaborate numerous mitochondria (Fig. 2), many in close association with the wall ingrowths, suggesting that the transport process is energy-requiring. An enlarged, lobed nucleus (Fig. 2) and numerous endoplasmic reticulum cisternae and ribosomes are often characteristic of the cytoplasm of transfer cells. Plasmodesmata (cytoplasmic channels for intercellular transport) connect transfer cells to other transfer cells and to parenchyma cells (Fig. 2).

Location and function. The location of transfer cells within the plant provides circumstantial evidence for their involvement in solute transport. They are situated at many sites in the plant where secretion or absorption takes place. Transfer cells are most frequently associated with the conducting elements of the xylem and phloem (elements used for long-distance transport in the plant), where they may play a role in either loading or unloading solutes from these cells. Minor veins of angiosperm leaves have particularly well-developed phloem transfer cells, whereas stem nodes often possess numerous xylem transfer cells. With an increasing number of ultrastructural studies of plant roots, it is apparent that transfer cells are more prevalent in this organ than previously recognized. *See* PHLOEM; XYLEM.

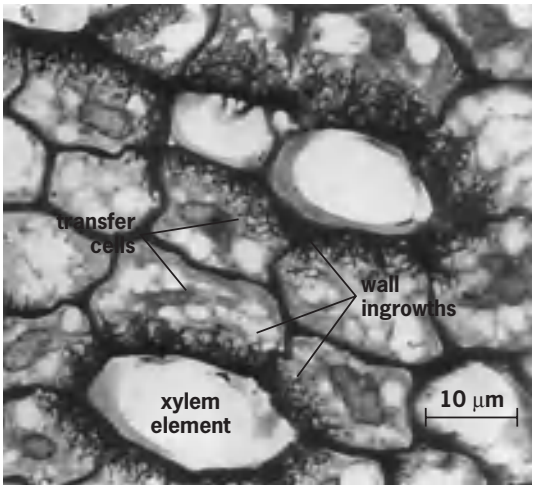

Fig. 1. Light micrograph of transfer cells adjacent to xylem elements in *Hieracium florentinum* rhizome. Long, branched wall ingrowths are evident.

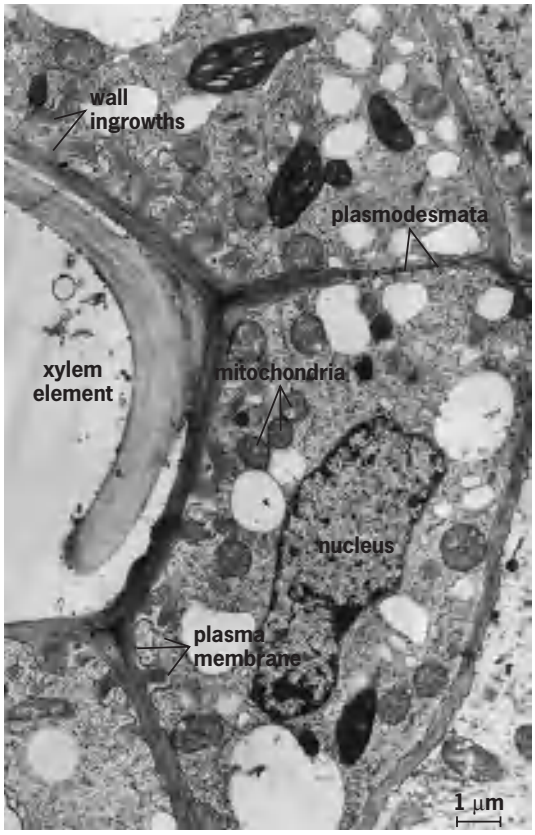

Fig. 2. Electron micrograph of transfer cells adjacent to a xylem element in *Hieracium florentinum* rhizome. Wall ingrowths enveloped by plasma membrane are located next to xylem element.

Four main categories of secretion or absorption involving transfer cells are generally recognized: (1) absorption of solutes from the external environment, with the epidermis of submerged leaves and the hydropotes ("water-drinking" glands) on the surface of water lily leaves being examples; (2) secretion of solutes to the external environment, as illustrated by various secretory glands and nectaries; (3) absorption of solutes from internal, extracytoplasmic compartments, as exemplified by transfer cells associated with tracheary elements of the xylem; and (4) secretion of solutes into internal, extracytoplasmic compartments, as shown by the secretion of materials from the tapetum into the locule of the anther in flowers. *See* SECRETORY STRUCTURES (PLANT).

Transport within the plant can occur from cell to cell either through plasmodesmata or through extracytoplasmic structures such as the cell wall and intracellular spaces. The former mode of transport is termed symplastic, while the latter is apoplastic. Transfer cells are thought to be involved in exchanges of solutes between these two systems. Wall ingrowths in transfer cells are porous enough to allow movement of lanthanum nitrate (a large molecule frequently used to show transport in the apoplast) and therefore presumably many other molecules of lower molecular weight. The plasma membrane provides the greatest resistance to solute movement. The occurrence of adenosine triphosphatases associated with the plasma membrane enveloping the wall ingrowths implies that active transport of solutes may be involved.

The exact mechanism by which transfer cells mediate exchange between the symplast and apoplast is, however, unknown. The early suggestion that these cells might increase absorption or secretion simply as a consequence of their increased surface area is probably too simplistic in some cases. Emphasis is being placed on the effects of wall ingrowth shape on transport processes. Wall ingrowths, by providing membrane-bound cytoplasmic clefts between them, may allow for the development of a standing osmotic gradient within the clefts. Primary carriers (porters) may then carry ions such as hydrogen (H^+), potassium (K^+), and chloride (Cl^-) across the membrane, thereby establishing electrical gradients. The presence of adenosinetriphosphatases on this membrane surface would be compatible with this energy-requiring step. Secondary porters driven by electrochemical gradients established by the primary porters may carry other ions or organic molecules across the membrane.

Regardless of the mechanism involved in transmembrane transport into or out of transfer cells, these cells have been shown to be capable of absorbing substances from the apoplast. Among several examples, hydropotes on water lily leaves have been shown by autoradiography to absorb sulfate ions from an external bathing medium, and transfer cells on leaves of carnivorous plants have been shown by autoradiographic and x-ray microanalytical techniques to absorb products from the digestion of insects. *See* ION TRANSPORT; PLANT TRANSPORT OF SOLUTES.

Differentiation. Although transfer cells are spectacular examples of cell differentiation within the plant, there is little firm evidence as to factors controlling their differentiation. In some cases, wall ingrowths are formed, or are increased in number, at the onset of solute transport, thus suggesting that solute presence might trigger parenchyma cells to form wall ingrowths. For example, in young pea leaves the appearance of wall ingrowths in phloem parenchyma cells coincides with export of photosynthates from the leaf. However, in other situations wall ingrowths develop during the normal ontogeny of the cell or tissue in question whether or not solutes are present for transport. An example of this was described in root nodule development in the pea. In effective nodules, wall ingrowths form in xylem parenchyma and in the pericycle prior to the onset of nitrogen fixation and the formation of nitrates for transport. Ineffective nodules, which are incapable of fixing nitrogen and therefore do not produce nitrates, still develop transfer cells in both the xylem parenchyma and pericycle. Hormonal or other morphogenetic stimuli may play a role in inducing wall ingrowth formation in parenchyma cells, but these factors are largely unknown. *See* PARENCHYMA.

Occurrence in plant groups. Transfer cells occur in diverse groups of plants, including algae, fungi, mosses, ferns, gymnosperms, and angiosperms.

Their development in presumably unrelated plant groups suggests that plants evolved this mechanism more than once to solve short-distance transport problems. The most highly evolved plant group, the angiosperms, probably has the highest frequency of species with transfer cells, but there are many angiosperm families which normally do not develop this cell type. However, the genetic information for transfer cell development may be present in species even though transport cells do not usually develop. This is shown most clearly in the roots of some species which do not normally develop transfer cells but do so when they become infected by nematodes. Root cortical cells in proximity to the head of the nematode enlarge, develop wall ingrowths, and become giant transfer cells. In some angiosperm families the presence or absence of transfer cells may be a useful taxonomic character when correlated with other morphologic characters. *See* CELL WALLS (PLANT); PLANT CELL. R. L. Peterson

Bibliography. E. G. Cutter, *Plant Anatomy*, pt. 1: *Cells and Tissues*, 2d ed., 1978; N. Harris and K. J. Oparka (eds.), *Plant Cell Biology: A Practical Approach*, 1994; J. S. Pate and B. E. S. Gunning, *Annu. Rev. Plant Physiol.*, 23:173–196, 1972.

Transform fault

One of the three fundamental types of boundaries between the mobile lithospheric plates that cover the surface of the Earth. Whereas spreading centers mark sites where crust is created between diverging plates, and subduction zones are where crust is destroyed between convergent plates, transform faults separate plates that are sliding past each other with neither creation nor destruction of crust. The primary tectonic feature of all transform faults is a strike-slip fault zone, a generally vertical fracture parallel to the relative motion between the two plates that it separates. Strike-slip fault zones are described as right-lateral if the far side is moving right relative to the near side (for example, the Queen Charlotte zone; **Fig. 1**), left-lateral if it is moving to the left (for example, the North Caribbean zone; Fig. 1). Not all such fault zones are plate-bounding transform faults. Small-scale strike-slip faulting is a common secondary feature of many subduction zones, especially where plate convergence is oblique, and of some spreading centers, especially those with propagating rifts; it also occurs locally deep in plate interiors. The distinguishing characteristic of a transform fault is that both ends extend to a junction with another type of plate boundary. At these junctions the divergent or convergent motion along the other boundaries is transformed into purely lateral slip. *See* EARTH CRUST; PLATE TECTONICS; SUBDUCTION ZONES.

Types. Transform faults are most readily classified by the types of plate boundary intersected at their ends, the variety of lithosphere (oceanic or continental) they separate, and by whether they are isolated or are part of a multifault system. The common oceanic type is the ridge-ridge transform, linking two literally offset axes of a spreading center (for example, Clipperton and Siqueiros transforms; Fig. 1). Also common are transform faults that link the end of a spreading center to a triple junction, the meeting place of three plates and three plate boundaries. For example, the Panama transform links a spreading axis on the mid-oceanic ridge between the diverging Cocos and Nazca plates to a Cocos-Nazca-Caribbean plate triple junction at the continental margin of Central America. *See* LITHOSPHERE; MID-OCEANIC RIDGE.

Other types are long trench-trench transforms at the northern and southern margins of the Caribbean plate, and the combined San Andreas/Gulf of California transform, which separates the North American and Pacific plates for 1500 mi (2400 km) between triple junctions at Cape Mendocino (California) and the mouth of the Gulf of California (Fig. 1). Strike-slip faulting in the Gulf of California (and on the northern Caribbean plate boundary) occurs along several parallel (en echelon) zones linked by short spreading centers, and the overall structure is more properly called a transform fault system; similar fault patterns are found at many ridge-ridge transforms. The San Andreas part of this plate boundary exhibits another type of transform fault system, one with several simultaneously active zones that overlap, rather than replacing each other in stepwise fashion; this pattern may be characteristic of wholly continental transforms.

Geology of oceanic transforms. The structure of ridge-ridge transforms on mid-oceanic ridges varies to some extent with their length, ranging 6–600 mi (10–1000 km), and with the rate of slip of their strike-slip faults 0.8–8 in. (20–200 mm) per year, but the structure depends mainly on the geologic history of changing plate motions. In the absence of such changes, where transform faults separate plates that have maintained the same motion for millions of years, the characteristic structure is a transform valley parallel to the direction of relative plate motion, and thereby transverse to the mid-oceanic ridge. The valley floor is occupied by a strike-slip fault zone, a band of shattered rock only 300–600 ft (100–200 m) wide, that is often marked by a groove or rift in the sea floor. There is some correlation of valley depth and transform length, and most ridge-ridge transforms longer than 60 mi (100 km) have valleys deeper than the 3300–6600-ft-deep (1000–2000-m) axial rift valleys typical of the crests of slow-spreading ridges. Slow-slipping transform valleys are generally deepest at their ends, at their orthogonal intersections with axial rift valleys, whereas the deepest parts of fast-slipping valleys are usually in their midsections, their ends being partly filled with lava that spills over from the intersecting spreading axes. *See* LAVA.

Transform valleys are structural troughs opened by a small component of extension across ridge-ridge transform faults. Where the faults are strictly parallel to relative plate motion, the origin of this extensional stress is probably thermal contraction of the young lithosphere accreted at the intersecting spreading axes. Larger components of extension, creating deeper transform valleys, can result from

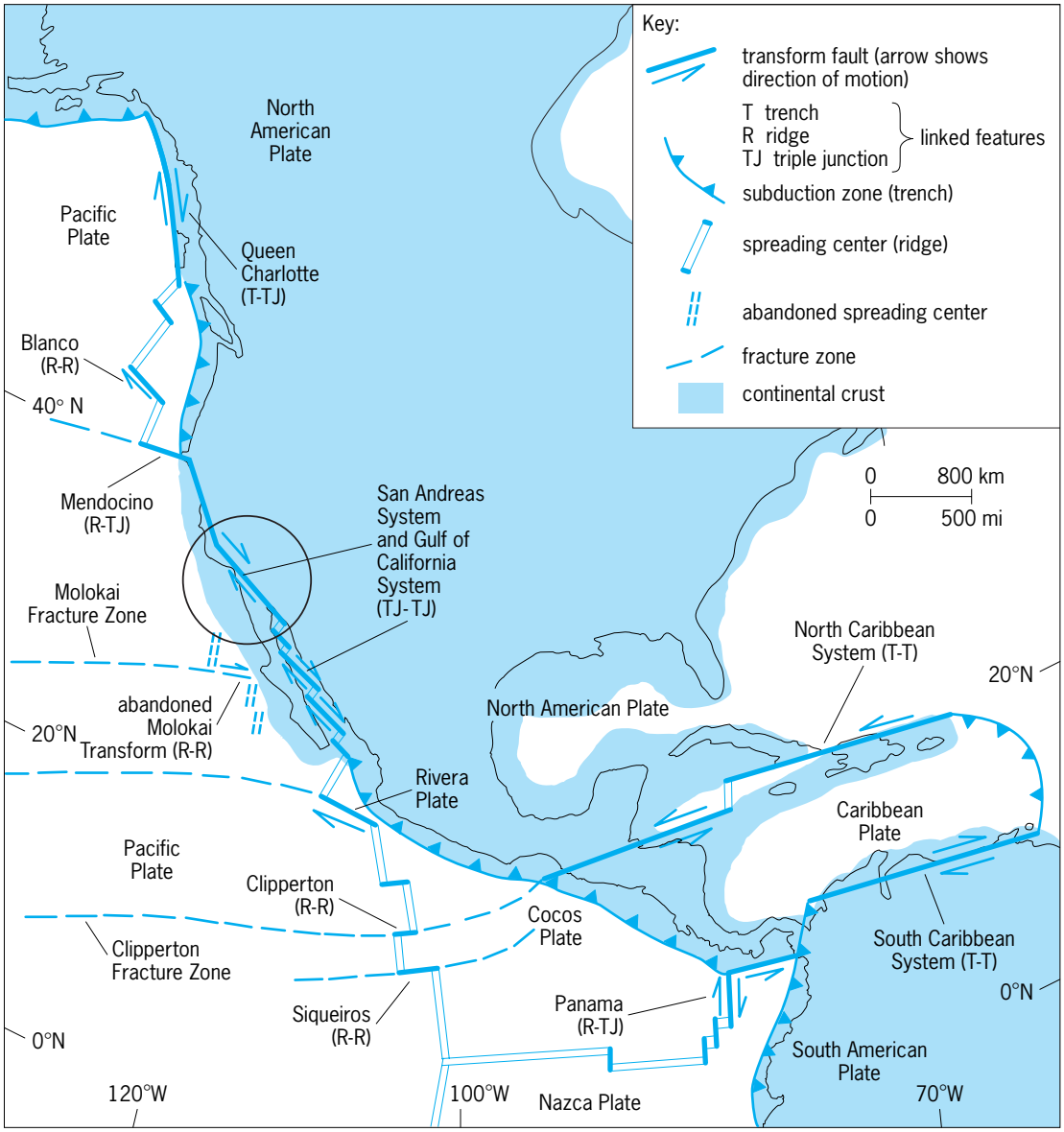

Fig. 1. Various types of transform fault mapped in parts of North and Central America. Area inside circle is shown in detail in Fig. 3.

small angular changes in the direction of plate motion; conversely, opposite changes can add a component of valley-closing compression. The motions of most oceanic plates are changing continuously, albeit slowly, affecting all the transform faults along their boundaries. A small rotation in the direction of plate motion has an opposite effect on the stresses at adjacent left-lateral and right-lateral transforms (**Fig. 2**). Such a motion has affected adjacent transform faults on part of the East Pacific Rise, where relative plate motion (spreading direction) has rotated anticlockwise by about 5° in the past 5 million years. The resulting convergence across the left-lateral Clipperton transform has closed its transform valley and thrust up a median ridge 2000 ft (600 m) high, with the strike-slip fault zone along its crest. The same rotation has caused extension at the right-lateral Siqueiros transform (Fig. 1). Some transforms react to a change of this sort by opening wider and deeper transform valleys, often accompanied by uplift of the valley margins by 0.6–1.8 mi (1–3 km) to form structures known as transverse ridges. In some cases, opening transform faults by adding a component of extension allows seawater to penetrate deep into the lithosphere, chemically altering the upper mantle to a low-density rock (serpentinite) that rises along the fault zone, forming median ridges similar in shape but quite different in origin to those at compressed transform faults. In a few examples, crustal divergence across the transform fault allows magma to leak out, building yet another type of median ridge. The Siqueiros transform shows a more common response to the change in plate motion. Instead of maintaining the same strike-slip fault zone, and adding an extensional component to its lateral motion, it developed a set of new fault zones, each parallel to the new plate motion, and slightly oblique to the overall trend of the transform fault system.

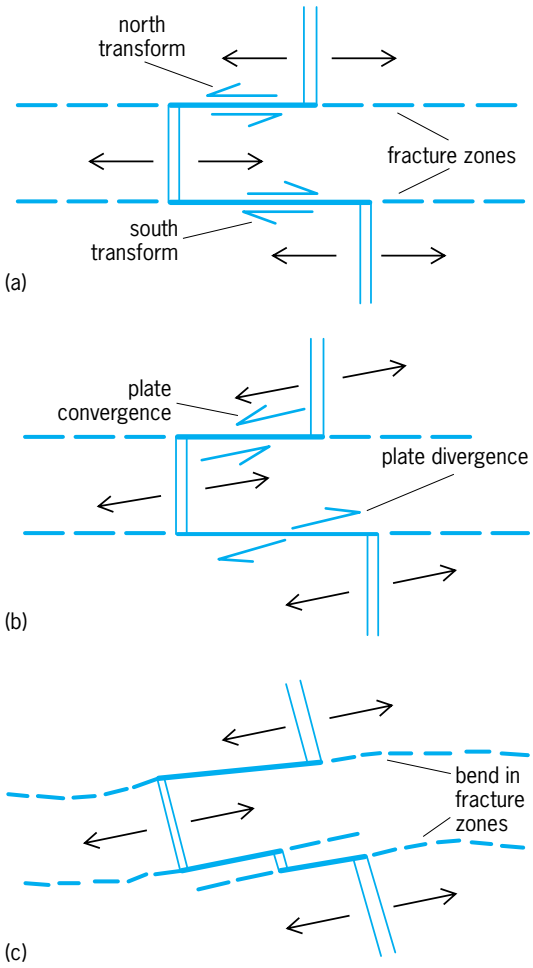

Fig. 2. The effect of changing plate motion on transform faults and their fracture zones. (*a*) East-west plate motion; offset spreading axes with a left-lateral North Transform and a right-lateral South Transform. (*b*) Change in motion; a small rotation of the direction of plate motion adds components of plate convergence to North Transform and divergence to South Transform. (*c*) Adjustment to change; fault is segmented into two fault zones aligned parallel to the new plate motion, which are linked by a short new spreading axis.

Magma does leak out within the fault system, but only at the short new spreading axes which link the parallel fault zones. *See* SERPENTINITE.

The geology of ridge-ridge transform faults is sensitive to the history of recent changes in the direction of plate motion, and the pattern of the fracture zones that they leave on the flanks of the mid-oceanic ridges provides a record of these changes. Fracture zones are bands of rough topography that extend down ridge flanks from the ends of transform faults. Their name is inherited from an early false interpretation that they are belts of strike-slip faulting across the flanks of mid-oceanic ridges. A foundation of the theories of sea-floor spreading and plate tectonics was the recognition that fracture zones are not active fault zones, merely the seams between crust that differs in age becuase it has spread different distances from laterally offset spreading axes. The lateral offsets occur at transform faults, so fracture zones are their inactive continuations. The azimuth of a fracture zone is parallel to the direction of plate motion at the time that the crust on its younger side spread off the risecrest, with fracture zone bends and kinks marking changes in direction. Mapping the trends of fracture zones is a principal method of investigating the past movements of lithospheric plates. *See* MARINE GEOLOGY.

Geology of continental transforms. Transform faults within the continental lithosphere fracture crust that is much thicker and less homogeneous than oceanic crust. Perhaps as a result, the fault zones tend to be less straight, with many local deviations from azimuths parallel to relative plate motion. Bends in the fault zones add components of extension or compression to the dominantly strike-slip motion, resulting in along-strike alternations of collapsed extensional basins and uplifted compressional ridges. A well-known compressional bend is the Big Bend of the San Andreas fault zone north of Los Angeles (**Fig. 3**), where oblique convergence of the Pacific and North American plates is raising the San Bernandino and San Gabriel mountains. Some of the sediment-filled basins formed along continental transform faults are important petroleum reservoirs, and secondary deformation on the margins of the fault zone commonly folds the sediment layers to form trapping structures for oil fields. *See* BASIN; PETROLEUM GEOLOGY.

Most continental transform fault systems have several belts of faulting, with complex spatial patterns of overlapping and splaying fault zones, and complex geologic histories, involving constant shifting of the share of the total interplate displacement among several zones, accompanied by the birth of new fault zones and the abandonment of others. In southern California, for example, motion on the San Andreas transform fault system is now concentrated on three narrow fault zones (Fig. 3) which differ in age and are accompanied by a multitude of less active subparallel zones, some of which may become dominant traces in the near geologic future. Very detailed geologic studies are needed to unravel the histories of continental transform fault systems, which do not leave fracture zone traces like their oceanic counterparts. The total lateral displacement between the two sides of a fault zone, commonly amounting to tens or hundreds of miles, can be estimated by recognizing the two displaced halves of preexisting geologic features that were split and separated by fault motion. On a much shorter time scale, recent displacements can be monitored by offsets in human-made features such as fence lines.

Some of the less active fault zones in the region of the San Andreas and Gulf of California systems are senescent rather than nascent transform faults. Until 6 million years ago, transform faulting was centered west of Baja California and west of the southern California coastline, and the inland shift of the Pacific–North America plate boundary has caused the almost complete cessation of faulting on the offshore San Benito and San Clemente fault zones.

Shearing continental margins. The now-inactive San Benito fault zone was representative of an

important class of transform faults that extend along continental margins, at or near the boundary between oceanic and continental lithosphere. A still-active example is the Queen Charlotte fault zone off the British Columbia coast (Fig. 1). Continental margins shaped by lateral shearing (transform faulting) have very steep continental slopes, but often with steps on the slope called marginal plateaus, crustal blocks that have subsided between the shifting fault zones of a transform fault system. The shifting is commonly away from the oceanic/continental boundary into adjacent weaker continental lithosphere; indeed, the San Andreas fault system can be considered a marginal shear zone that has shifted unusually far inland.

Most shearing margins of western North America were formerly, with an earlier arrangement of a lithospheric plates, convergent (subduction zone) margins. Shearing margins with active transform faults also play a role during the birth of ocean basins by continental rifting. Initial rifting, as in the split of North America from Africa about 200 million years ago, is commonly on laterally displaced fractures that develop into spreading axes linked by transform faults. These transform faults become part of the oceanic/continental boundary once continental separation has proceeded far enough for sea-floor spreading to occur, and eventually become ridge-ridge transforms once a risecrest develops in the new ocean basin. Many of the ridge-ridge transform faults on the Mid-Atlantic Ridge are inherited from fault zones that once formed shearing parts of the continental margin. Shearing margins occur on the boundaries of the very small, young ocean basins that have opened by the splitting of Baja California from mainland Mexico (for example, Guaymas Basin, with the Guaymas transform fault on its northeast side; Fig. 3). *See* CONTINENTAL MARGIN.

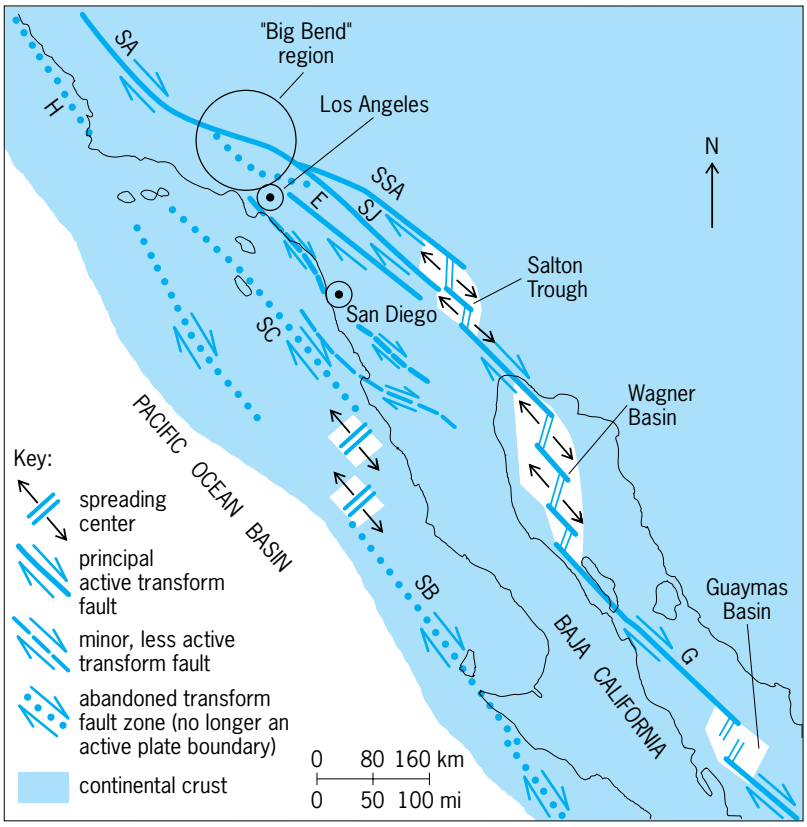

Fig. 3. Map showing parts of the San Andreas and Gulf of California fault systems. Many minor but still earthquake-generating strike-slip fault zones have been omitted. The abandoned offshore fault zones, relics of a time when the plate boundary was closer to the continental margin, still have a low level of residual earthquake activity. The individual fault zones are San Andreas (SA), Hosgri (H), South San Andreas (SSA), San Jacinto (SJ), Elsinore (E), San Clemente (SC), San Benito (SB), and Guaymas (G).

Earthquakes. Along a few strike-slip fault zones, lithospheric plates slide quietly and almost continuously past each other by the process called aseismic creep. Much more often, frictional resistance to the sliding in the brittle crust causes the accumulation of shear stresses that are episodically or periodically relieved by sudden shifts of crustal blocks, creating earthquakes. The largest lateral shifts (slips) of the ground surface along major continental transform faults have been associated with some of the largest earthquakes on record; in 1906 the Pacific plate alongside 270 mi (450 km) of the San Andreas Fault suddenly moved an average of 15 ft (4.5 m) northwest relative to the North American plate on the other side, and the resulting magnitude-8.2 earthquake destroyed much of San Francisco. The average slip in this single event was equivalent to about 150–250 years of Pacific–North American plate motion.

The maximum size of earthquake that a transform fault can generate is limited by the length of the fault, though generally, even in large earthquakes like the one in San Francisco in 1906, a fault does not fail along its entire length. The frequency of earthquakes is controlled by the average speed of relative plate motion across a transform fault plate boundary and, for fault systems with multiple overlapping fault zones, by the share of this motion that is carried by any individual fault. However, many local geologic and tectonic factors intrude to complicate estimates of how frequently a particular transform fault will produce earthquakes of any specified size or destructive power, and it is still more difficult to predict the exact timing of such an event. Extrapolation of the past record is probably the best method of estimating future magnitudes and frequencies, becuase many transform faults do seem to have a characteristic size of large earthquakes, and a consistent ratio of small to large events. In most cases, however, especially for remote oceanic transform faults, the record is too short and too incomplete to be of much practical use. *See* EARTHQUAKE; FAULT AND FAULT STRUCTURES; SEISMOLOGY.

Peter Lonsdale

Bibliography. W. G. Ernst (ed.), *The Geotectonic Development of California*, 1981; P. J. Fox and D. G. Gallo, The geology of North Atlantic transform plate boundaries and their aseismic extensions, *The Geology of North America*, vol. M, pp. 157–172, 1986; P. J. Fox and D. G. Gallo, Transforms of the eastern central Pacific, *The Geology of North America*, vol. N, pp. 111–124, 1989; J. T. Wilson, A new class of faults and their bearing on continental drift, *Nature*, 207:343–347, 1965.

Transformation (bacteria)

The addition of deoxyribonucleic acid (DNA) to living cells, thereby changing their genetic composition and properties. Such transformation was first reported by Fred Griffith in 1928 as occurring in pneumonia bacteria (pneumococci, or *Streptococcus pneumoniae*) growing in animals injected in the laboratory. Other investigators produced the same phenomenon in bacteria of the same species growing in the test tube. The process is looked upon as the transfer of a transforming principle, or substance, from donor bacteria to suitable recipient bacteria. The recipient bacteria are usually closely related to the donor strain. It is realized that the process may occur in natural conditions, for example, in a host animal infected with two parasitic strains, and indeed it might play a part in the rapid evolution of pathogenic bacteria.

There are several species of bacteria in which transformation has been achieved in the laboratory, including pneumococci, meningococci, *Haemophilus influenzae*, and *Bacillus subtilis*. There are several other species in which similar phenomena have been observed, and other strains of the same and other species in which it has not as yet been possible to achieve transformations.

Role of DNA. The most far-reaching clarification attained through laboratory investigation was the announcement by Oswald T. Avery and coworkers in 1944 that the purified agent (or transforming principle) inducing capsule-type transformation of pneumococcus had the properties of a pure DNA. It was later shown that other specific hereditary traits of these and other bacteria could be transformed by a similar process. In this way it is increasingly clear that the structures of deoxyribonucleic acids of different origins bear and transmit the imprint of their specific genetic origins. Supplementing this evidence, there have been signs that cell nuclei in general all contain DNA in constant, characteristic amounts, and that radiations and chemicals which can affect nucleic acids also provoke mutations in living organisms. A confirmation of the role of DNA was the evidence that the active components of certain viruses are the DNAs they carry. *See* ANIMAL VIRUS; DEOXYRIBONUCLEIC ACID (DNA); GENETICS; MUTATION; NUCLEIC ACID.

Nature of transformation. That bacterial transformation is true genetic transmission on a small scale, rather than controlled mutation, is demonstrated by the following characteristics: (1) A specific trait is introduced, coming always from donors bearing the trait. (2) The trait is transferred by determinant, gene-like material far less complex than whole cells or nuclei, and this material, DNA, is known to be present in gene-carrying chromosomes. (3) The trait is inherited by the progeny of the changed bacteria. (4) The progeny produce, when they grow, increased amounts of DNA carrying the specific property. (5) The traits are transferred as units exactly in the patterns in which they appear or in which they are induced by mutation. (6) The DNA transmits the full potentialities of the donor strain, whether these are in an expressed or in a latent state. (7) The traits are often attributable to the presence of a specific gene-determined enzyme protein. (8) Certain groups of determinants may occur "linked" within DNA molecules, just as genes may be linked, and if so, heat denaturation, radiation, or enzyme action will inactivate or separate them just to the extent that they can damage or break apart the DNA molecules. (9) Linked determinants, while transforming a new cell, may become exchanged (recombined) between themselves and their unmarked or unselective alternate forms in such a way that they bring about genetic variation, and in a pattern indicating the existence of larger organized genetic units. *See* BACTERIAL GENETICS; GENE.

Nonbacterial transformation. Through the application of a number of procedures prior to adding the DNA, transformation was extended first to many different bacterial species and then to eukaryotic cells. These procedures included adding high concentrations of calcium or dextran sulfate, or providing a large electric shock (electroporation). Today almost any cell type can be transformed. In some cases, tissues can be injected directly with naked DNA and transformed. However, unlike with bacteria, the naked DNA adds almost anywhere in the genome rather than recombining with its indigenous homolog. However, with special highly selective procedures, homologous recombination can be obtained. By treating embryonic stem cells and adding them to embryos that then go to term, specific and nonspecific transgenic animals can be obtained (for example, mice). *See* GENETIC ENGINEERING.

When the source of the DNA is some entity capable of independent replication, such as a virus or plasmid, the phenomenon is called transfection. If foreign DNA is then inserted into these entities, the result is recombinant DNA that can lead to transduction. *See* MOLECULAR BIOLOGY; TRANSDUCTION (BACTERIA).

Rollin D. Hotchkiss; Norton D. Zinder

Bibliography. O. T. Avery, C. M. MacLeod, and M. McCarty, *J. Exper. Med.*, 79:137–158, 1944.

Transformer

An electrical component used to transfer electric energy from one alternating-current (ac) circuit to another by magnetic coupling. Essentially it consists of two or more multiturn coils of insulated conducting material, so arranged that any magnetic flux linking one coil will link the others also. This configuration creates mutual inductances between the coils. The mutual magnetic field acts to transfer energy from one input coil (or primary winding) to the other coils, which are then referred to as secondary windings. Under steady-state conditions, only one winding can serve as a primary. *See* COUPLED CIRCUITS; INDUCTANCE.

The transformer accomplishes one or more of the following effects between two circuits: (1) an induced voltage of different magnitude, (2) an induced

current of different magnitude, (3) a difference in phase angle, (4) a difference in impedance level, and (5) a difference in voltage insulation level, either between the two circuits or to ground.

In an electric power system, transformers are used to perform a wide range of functions. Pole-type distribution transformers supply relatively small amounts of power to residences. Power transformers are used at generating stations to step up the generated voltage to high levels for transmission. The transmission voltages are then stepped down by transformers at the substations for local distribution. Instrument transformers are used to enable accurate measurents of voltages and currents. In other applications, audio- and video-frequency transformers must function over a broad band of frequencies. Radio-frequency transformers transfer energy in narrow frequency bands from one circuit to another. *See* INSTRUMENT TRANSFORMER.

Transformers are often classified according to the frequency for which they are designed. Power transformers are for power-frequency circuits, audio transformers for audio-frequency circuits, and so forth. Of course, many of the basic principles of operation apply to all.

Power Transformers

Systems for the transmission and distribution of alternating-current electrical energy would be impractical without transformers. Transformers are capable of stepping alternating voltages up or down with very little loss of power in the process. They thus allow engineers to design different parts of a power system to operate at different voltage levels. Electric generators may then be designed for voltages which make possible the most economical use of materials. These generators are connected through transformers to transmission lines operating at the much higher voltages required for high transmission efficiency. Transformers again step the voltage down to levels which permit the energy to be used safely by the ultimate consumer. Large ac generators, called alternators, are designed to operate at 20 to 25 kV. Power transmission voltages in the United States range from 115 to 750 kV, and utilization voltages range from 115/230 V for home use to 13,200 V in large industrial equipment. *See* ALTERNATING-CURRENT GENERATOR; ELECTRIC DISTRIBUTION SYSTEMS; ELECTRIC POWER SYSTEMS; TRANSMISSION LINES.

Power transformers, as a class, may be defined as those designed to operate at power-system frequencies: 60 Hz in the United States and Canada, and 50 Hz in much of the rest of the world. The largest power transformers connect generators to the power grid. Since a generator, together with its driving turbine and prime energy source, is called a generating unit, such transformers are called unit transformers. The classification "distribution transformers" refers to those supplying power to the ultimate consumers. They are designed for lower power and output-voltage ratings than the other transformers in the system.

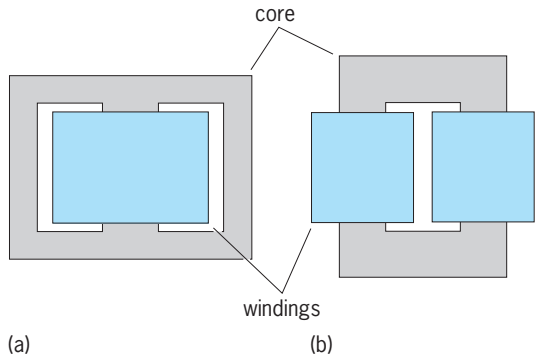

Fig. 1. Location of windings in single-phase cores. (*a*) Shell form. (*b*) Core form.

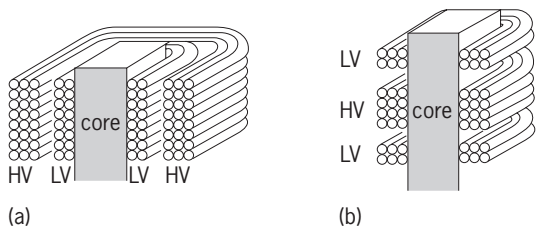

Fig. 2. Winding arrangements. (*a*) Concentric. (*b*) Interleaved.

Configuration. Any magnetic flux which arises from a current in one winding of a transformer but does not link (is not mutual to) the other windings is called the leakage flux of that winding. General-purpose power transformers are designed to maximize the mutual component of the flux of each winding by providing an iron-alloy core which links all windings. This core acts as a high-permeance path for the mutual magnetic flux. Practically speaking, all of the mutual flux flows in this core.

Typical configurations for single-phase transformers are shown in **Fig. 1**. The arrangement in Fig. 1*a* is called a shell-form transformer, while that in Fig. 1*b* is called a core-form transformer. Each of the rectangles labeled "windings" in this figure represents at least two coils. The coils may be concentric, or interleaved (**Fig. 2**). In the shell form, all of the windings are on the center leg. In the core form, half of the turns of the primary and half of those of the secondary are on each leg. The two halves of a given winding may be connected in series or in parallel. In the concentric arrangement, the low-voltage winding is closer to the core, so that less insulating material is required in the construction of the transformer.

Principles of operation. Transformers operate on the basis of two fundamental physical laws: Faraday's voltage law and Ampère's law. Faraday's law states that the voltage induced in a winding by a magnetic flux linking that winding is proportional to the number of turns and the time rate of change of the flux; that is, Eq. (1) holds, where e_i is the voltage induced

$$e_i = N_i \frac{d\phi_i}{dt} \qquad \text{volts} \tag{1}$$

in a coil of N_i turns which is threaded by a flux of ϕ_i

webers changing at a rate of $d\phi_i/dt$ webers per second. The ratio of the voltages induced in two windings of a transformer by the core flux is, then, given by Eq. (2). In other words, the voltages induced in the

$$\frac{e_1}{e_2} = \frac{N_1}{N_2}\frac{d\phi_{\text{core}}/dt}{d\phi_{\text{core}}/dt} = \frac{N_1}{N_2} \tag{2}$$

windings are proportional to the numbers of turns in the windings. This is the basic law of the transformer. A high-voltage winding will have many turns, and a low-voltage winding only a few. The N_1/N_2 ratio is usually called the turns ratio or transformation ratio and is designated by the symbol a, so that Eq. (3)

$$\frac{e_1}{e_2} = a \tag{3}$$

holds. *See* FARADAY'S LAW OF INDUCTION.

Since the flux must change to induce a voltage, steady-state voltages can be obtained only by a cyclically varying flux. This means that alternating voltages and fluxes are required for normal transformer operation, and that is the fundamental reason for ac operation of power systems. Devices operated on ac have lower losses when the voltages and fluxes are sinusoidal in form, and sinusoidal fluxes and terminal voltages will be assumed in this discussion. *See* ALTERNATING-CURRENT CIRCUIT THEORY.

Core excitation. **Figure 3** shows the elements of a two-winding, shell-form transformer. The center leg of the core carries the full mutual flux, and each of the outside legs carries half of it. Thus the cross-sectional area of each outside leg is half of that of the center leg.

The core is laminated; that is, it is made up of a stack of iron-alloy sheets. The laminations are insulated from each other by iron oxide or by some other coating. If this were not the case, the core would act like a closed, single-turn winding, and the voltage induced in it by the core flux would cause large circulating currents in the core, called eddy currents. These currents would cause excessive core heating, and result in a severe reduction in the efficiency of the transformer. There is some eddy-current loss even in laminated cores. *See* EDDY CURRENT.

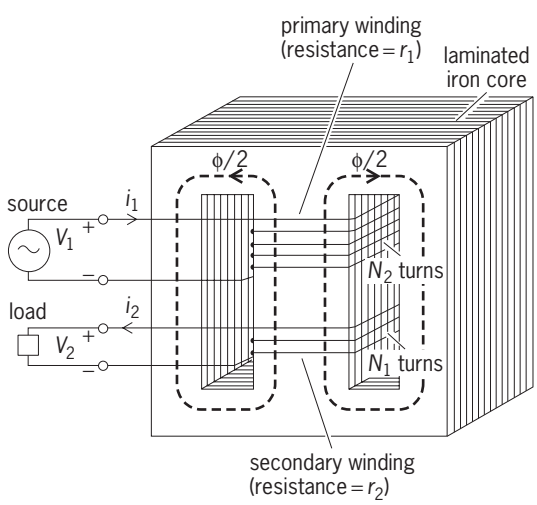

Fig. 3. Elements of a transformer.

The thickness chosen for the laminations is determined by the operating frequency. For 60 Hz, the thickness is usually 0.014 in. (0.36 mm). Lamination makes the core mechanically weak, so arrangements must be provided for its support. Cores for distribution transformers are often wound from iron-alloy strip, and the possibility of using magnetic glass in power transformer cores has been investigated.

The constant reorientation of the magnetic domains in the core iron as the flux alternates in direction results in hysteresis loss. The sum of the hysteresis and eddy-current losses is called the core loss. *See* CORE LOSS.

The concept of exciting current may be defined by considering the situation in which the secondary winding in Fig. 3 is open (the load is disconnected) and a sinusoidal voltage is applied to the terminals of the primary winding, given by Eq. (4). In this equa-

$$v_1 = \sqrt{2}\, V_1 \cos 2\pi f t \qquad \text{volts} \tag{4}$$

tion V_1 is the effective or root-mean-square (rms) voltage, the peak voltage is $2V_1$, f is the frequency in hertz, and $2\pi f$ is the angular frequency ω in radians per second. An alternating flux is set up in the core and a voltage e_1 is induced in the N_1 turns in the primary which is almost equal to v_1 and opposed to it. The small difference between v_1 and e_1 is due to the leakage flux and to the winding resistance. This difference is enough to allow a small current to flow, sufficient to set up the core flux and to supply the hysteresis and eddy-current losses of the core. This small current (1–5% of rated primary current) is called the exciting current, i_{ex}. Its root-mean-square phasor I_{ex} may be considered to be made up of two components, as in Eq. (5), where I_ϕ is the magne-

$$I_{\text{ex}} = I_\phi + I_{cl} \qquad \text{amperes} \tag{5}$$

tizing component, which lags the applied voltage by 90°; and I_{cl} is the core-loss component, which is in phase with the voltage. Even though the applied voltage and core flux are sinusoidal, the nonlinearity of the iron causes the exciting current to be nonsinusoidal; it contains a large third-harmonic component. (The concept of the phasor I_{ex} is actually an approximation; it represents a sinusoid having the same root-mean-square value as the actual nonsinusoidal waveform.)

The applied voltage v_1 and the voltage induced by the core flux are so nearly equal that Eq. (6) fol-

$$\sqrt{2}\, V_1 \cos 2\pi f t = N_1 \frac{d\phi_{\text{core}}}{dt} \tag{6}$$

lows from Eqs. (1) and (4). Integrating Eq. (6) gives Eq. (7), where ϕ_K, the constant of integration, is a

$$\phi_{\text{core}} = \frac{V_1}{N_1}\frac{\sqrt{2}}{2\pi f} \sin\, 2\pi f t + \phi_K \qquad \text{webers} \tag{7}$$

transient flux which decays after a few cycles, due to effects not included in Eq. (6). This transient is

responsible for the phenomenon of current inrush. The peak value of the core flux is, in the steady state, given by Eq. (8), and is seen to be proportional to

$$\phi_{\max} = \frac{V_1}{N_1 f}\frac{\sqrt{2}}{2\pi} \tag{8}$$

the volts per turn of the primary winding. Since the core flux density (teslas, or webers/m^2) is limited by saturation, the designer's choice of volts per turn of the windings will determine the cross-sectional area of the core, or vice versa.

Transformation of the secondary load. When an electrical load is connected to the terminals of the secondary winding, current flows through that winding and the load as a result of the voltage e_2 induced in the secondary by the alternating core flux. In Fig. 3, source current flows into the positive terminal of the primary, but out of the positive terminal of the secondary. Application of the right-hand rule to the two windings shows that the load current in the secondary results in a magnetomotive-force (MMF) field, which is opposed to the magnetomotive-force field of the primary. Now, Ampère's law in circuital form states that the magnetomotive force applied to a closed path is simply equal to the net current enclosed by that path. In Fig. 3, Ampère's law may be applied to either of the two $\phi/2$ paths. In either case, the net current enclosed is $N_1 i_1 - N_2 i_2$, and the magnetomotive force applied to the core around either path is given by Eq. (9).

$$F_{\text{core}} = N_1 i_1 - N_2 i_2 \qquad \text{amperes} \tag{9}$$

See AMPÈRE'S LAW.

Now the core flux is also given by the core permeance P times the applied magnetomotive force, Eq. (10). Combining Eqs. (9) and (10) gives Eq. (11), and solving Eq. (11) for i_1 gives Eq. (12). The term

$$\frac{\phi_{\text{core}}}{2} = \frac{P}{2}F_{\text{core}} \qquad \text{webers} \tag{10}$$

$$\frac{\phi}{P} = N_1 i_1 - N_2 i_2 \qquad \text{amperes} \tag{11}$$

$$i_1 = \frac{N_2}{N_1} i_2 + \frac{\phi_{\text{core}}}{PN_1} \qquad \text{amperes} \tag{12}$$

ϕ_{core}/PN_1 is the instantaneous value of the exciting current of Eq. (5). The large permeance of the core makes this term very small. Equation (12) may be rewritten as Eq. (13). Thus, with a load on the secondary, the primary current is equal to the secondary current transformed by the inverse of the turns ratio, plus the current required to excite the core. *See* PERMEANCE.

$$i_1 = \frac{i_2}{a} + i_{\text{ex}} \qquad \text{amperes} \tag{13}$$

Ideal transformer. An ideal transformer would require no exciting current. In other words, its core would have infinite permeance, and the primary current would be given by Eq. (14). The windings of an ideal transformer would have zero resistance, and would be wound so as to have no leakage flux (a physical impossibility). Then the induced voltages would exactly equal the terminal voltages, and Eq. (15) follows from Eqs. (3) and (14). Thus in

$$i_1 = \frac{i_2}{a} \qquad \text{amperes (ideal)} \tag{14}$$

$$v_1 i_1 = v_2 i_2 \qquad \text{volt-amperes} \tag{15}$$

an ideal transformer, the instantaneous input apparent power equals the instantaneous output apparent power. There are no internal losses, and the ideal transformer has an efficiency of 100%.

In terms of root-mean-square phasors, the voltages and currents are related by Eqs. (16). The complex

$$V_1 = aV_2 \qquad I_1 = \frac{I_2}{a} \tag{16}$$

load impedance, in ohms, is given by Eq. (17), but

$$Z_{L2} = \frac{V_2}{I_2} \tag{17}$$

the impedance seen by the generator connected to the primary is given by Eq. (18). Thus impedances

$$Z_{\text{in1}} = \frac{V_1}{I_1} = \frac{aV_2}{I_2/a} = a^2 Z_{L2} \tag{18}$$

are transformed by the square of the turns ratio. Similarly, if a load were applied to the number 1 winding, the impedance appearing at the number 2 terminals would be given by Eq. (19). Most power transformers are so nearly ideal that, for many engineering purposes, Eqs. (16), (18), and (19) may be used to obtain satisfactory numerical values with real transformers. *See* ELECTRICAL IMPEDANCE.

$$Z_{\text{in2}} = \frac{Z_{L1}}{a^2} \tag{19}$$

Circuit model. When the detailed performance of a transformer must be accounted for in a power system calculation, a circuit which models the performance of the actual transformer is included in the system circuit (**Fig. 4***a*). To simplify the problem solution, the primary and secondary series impedances are nearly always combined into one equivalent impedance. This is accomplished through the impedance-transforming property of the ideal transformer. If the secondary winding resistance and leakage reactance are transferred to the primary by the factor a^2, and are then added to the primary series impedance, the result is called the equivalent impedance referred to the primary, given in Eq. (20)

$$\begin{aligned} Z_{\text{eq1}} &= R_{\text{eq1}} + jX_{\text{eq1}} \\ &= r_1 + a^2 r_2 + j(x_1 + a^2 x_2) \end{aligned} \tag{20}$$

and diagrammed in Fig. 4*b*, where r_1 and r_2 are the resistances of the primary and secondary windings, and x_1 and x_2 are the reactances due to leakage fluxes. When the primary impedance is transferred and combined with the secondary impedance, the equivalent impedance referred to the secondary

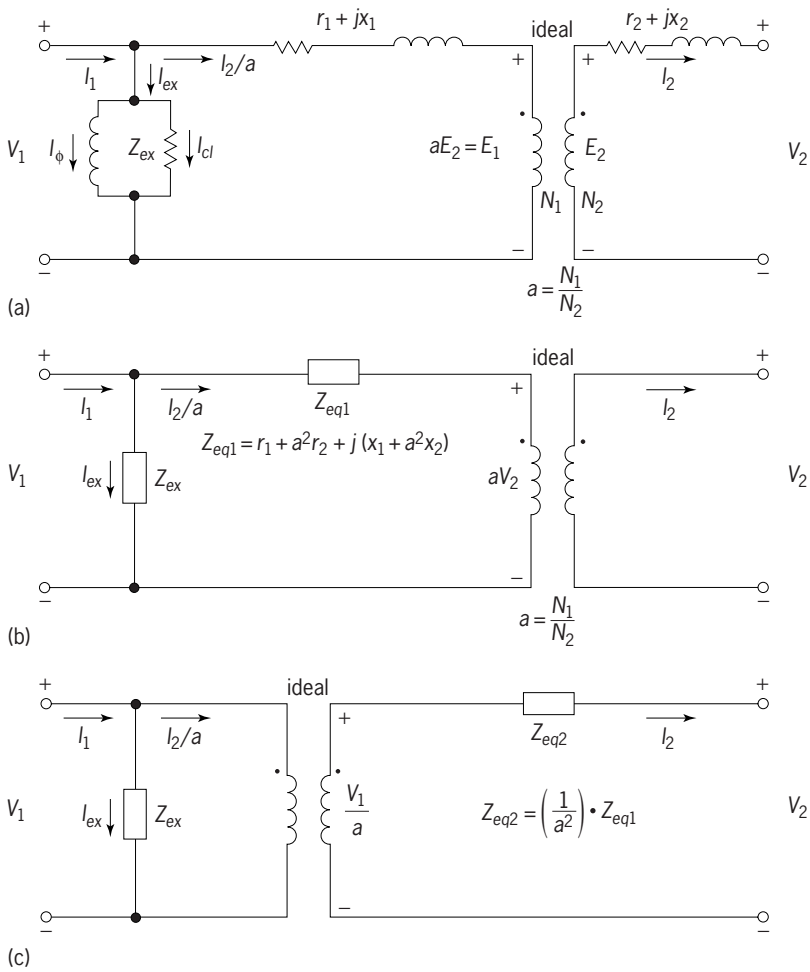

Fig. 4. Circuit models for performance calculations. (*a*) Transformer model showing separate primary and secondary impedances. r_1 and r_2 are resistances of primary and secondary windings, and x_1 and x_2 are reactances due to their leakage fluxes. (*b*) Secondary impedances transformed by ideal transformer and combined with primary impedance to form Z_{eq1}. (*c*) Transformer model in terms of the equivalent impedance referred to the secondary, Z_{eq2}.

results, as given in Eq. (21) and diagrammed in

$$Z_{eq2} = R_{eq2} + jX_{eq2} = \frac{r_1}{a^2} + r_2 + j\left(\frac{x_1}{a^2} + x_2\right)$$
$$= \frac{Z_{eq1}}{a^2} \tag{21}$$

Fig. 4*c*. The operator *j* in these expressions indicates that the voltages across the reactive elements of these impedances leads the voltages across the resistive elements by 90°. All voltages and currents in Fig. 4 are to be treated as ac phasor quantities. *See* REACTANCE.

The model makes it easy to calculate the primary voltage required to have rated secondary voltage at a given volt-ampere load and power factor. If the phase of the secondary voltage is set equal to 0°, as in Eq. (22), and if the power factor is cos θ, then the secondary current is given by Eq. (23), and from

$$V_2 = V_{2\text{rated}}\angle 0^\circ \tag{22}$$

$$I_2 = \frac{kVA \cdot 1000}{|V_{2\text{rated}}|}\angle\theta = |I_2|\angle\theta \tag{23}$$

Fig. 4*c*, the primary voltage is given by Eq. (24). The

$$V_1 = a[V_2 + |I_2|(\cos\theta \pm J\sin\theta) \cdot Z_{eq2}] \tag{24}$$

sign on the $j \sin\theta$ term is positive for leading and negative for lagging power factor, and Z_{eq2} is a complex number.

The model allows easy comparison between realistic and ideal transformers, as follows:

Realistic	*Ideal*
$V_1 = a(V_2 + I_2Z_{eq2})$	$V_1 = aV_2$
$I_1 = \frac{I_2}{a} + I_{ex}$	$I_1 = \frac{I_2}{a}$

The ratio of rated primary to secondary voltages shown on the transformer nameplate gives the actual turns ratio. This means that, if the secondary voltage is at rated value, the primary voltage will differ a little from its rated value.

Per unit impedance. The equivalent series impedance of a transformer is nearly always expressed as a fraction of the load impedance which would draw rated current at the rated voltage. This full-load impedance for a given winding is called the base impedance of that winding, as in Eq. (25). The

$$Z_{\text{base2}} = \frac{V_{2\text{rated}}}{I_{2\text{rated}}} \tag{25}$$

"per unit" impedance of a transformer is given in percent value by Eq. (26). The $R\%$ is considerably

$$Z\% = \frac{Z_{eq2}}{Z_{\text{base2}}} \cdot 100\%$$
$$= R\% + jX\% \tag{26}$$

smaller than $X\%$, so usually only the $X\%$ is shown on the transformer nameplate. Typical values of $R\%$ range from 0.3% in very large transformers to 1.0% in distribution transformers. It can be shown that $R\%$ is equal to the internal ohmic (I^2R) loss of the transformer at full load, expressed as a percent of the rated volt-amperes of the device.

The value of $X\%$ depends on the amount of leakage flux. High-voltage windings require more insulating material, and this results in more space between windings for leakage fluxes. Consequently a transformer with a very high-voltage winding will have a large value of $X\%$. Typical values range from 1.5% in distribution transformers to 15% in transformers with, say, a 345-kV winding.

Transforming three-phase power. Three-phase power may be transformed from one voltage level to another by sets of three single-phase transformers or by three-phase transformers. Three-phase transformers have three primary/secondary sets of windings on a single core. The arrangements for core and shell-form three-phase transformers are shown in **Fig. 5**. **Figure 6** shows a core-form, three-phase transformer removed from its tank.

When the core-form, three-phase transformer is operating with balanced voltages, the alternating fluxes of the three phases are equal in magnitude and 120° out of phase with each other. They thus add to

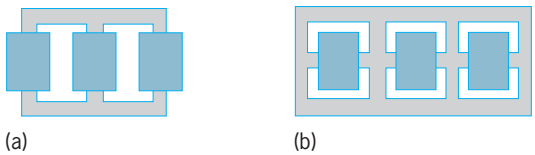

Fig. 5. Typical three-phase cores showing location of windings. (*a*) Core form. (*b*) Shell form.

zero at the center junctions of the core. When it is anticipated that the transformer will operate under unbalanced conditions, additional core legs are added at one or both ends of the core, to form "four-limbed" or "five-limbed" cores. These added legs carry the net flux resulting from the unbalance, which could otherwise flow in the steel transformer tank and cause overheating.

Three-phase transformers may be connected with their primary windings in wye or delta and their secondary windings in wye or delta, providing a great deal of flexibility. A wye-delta connection is illustrated in **Fig. 7**, having wye-connected windings with high-voltage line terminals a_1, b_1, and c_1, and delta-connected windings with low-voltage terminals a_2, b_2, and c_2. In wye-delta or delta-wye, there is always a 30° phase shift between the primary and secondary line voltages. It is standard practice in the United States to connect the windings in such a way that the voltages on low-voltage side (such as V_{a2b2}, the phasor voltage drop from line a_2 to line b_2 in Fig. 7) lag those on the high-voltage side (such as V_{a1b1}, the drop from line a_1 to line b_1) by 30°. There is some advantage in connecting the high-voltage windings in wye, because the neutral (with terminal N in Fig. 7) may be grounded, making the insulation requirements less severe. (In Fig. 7, V_{a1N}, the line a_1 to neutral phasor voltage on the high-voltage side, is the primary voltage of the transformer element whose secondary voltage is V_{a2b2}; thus V_{a2b2} is in phase with V_{a1N}.) It is desirable to have one set of windings in delta, because a delta tends to maintain balanced phase voltages and serves as a solution to certain third-harmonic problems associated with the wye connection. When a wye-wye connection is essential, a set of delta-connected "tertiary" windings is often provided, whose terminals may or may not be made accessible outside the transformer. Three-phase power transformation may also be accomplished by means of sets of two transformers, connected in T, open-delta, or open-wye-open-delta. These arrangements are usually used as distribution transformers, involving relatively small ratings, say, up to 500 kVA.

Efficiency. Efficiency is defined by Eq. (27). The

$$\eta = \frac{\text{output power}}{\text{input power}} \tag{27}$$

output power is equal to the input power, less the internal losses of the transformer. These losses include the ohmic (I^2R) loss in the windings, called copper loss, given by Eq. (28), and the core loss, called the

$$\begin{aligned}\text{Copper loss} &= |I_2|^2 R_{\text{eq2}} \text{ (single-phase)} \\ &= 3\,|I_2|^2 R_{\text{eq2}} \text{ (three-phase) watts}\end{aligned} \tag{28}$$

no load loss. The input power is thus the sum of the output power and the copper and core losses. Typical efficiency for a 20,000-kVA power transformer at full load is 99.4%, while that of a 5-kVA transformer is 94%.

Cooling. Transformer losses in the windings and core generate heat, which must be removed to prevent deterioration of the insulation and the magnetic properties of the core. Most power transformers are contained in a tank of oil. The oil is especially

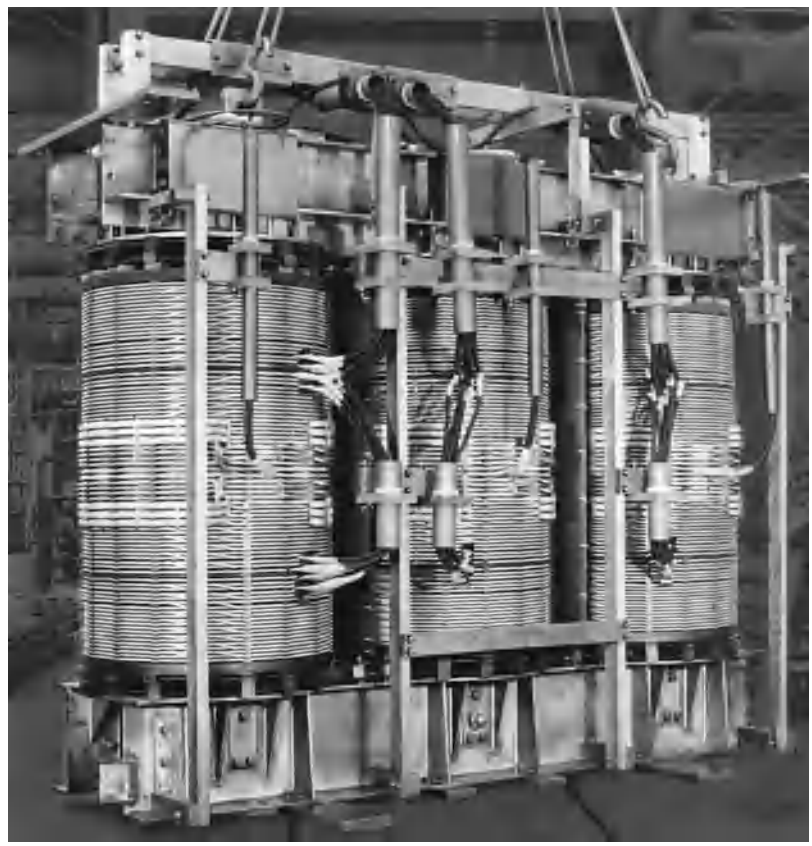

Fig. 6. Three-phase core and coils, rated at 50,000 kVA, 115,000 V.

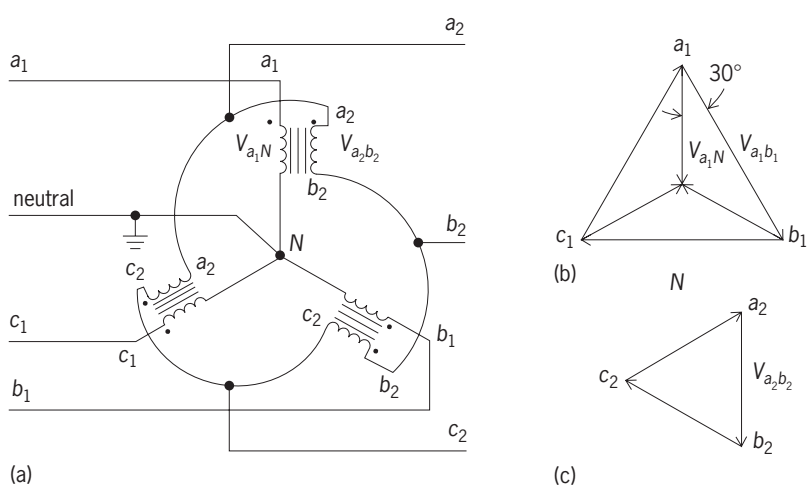

Fig. 7. Wye-delta connection of a three-phase transformer. (*a*) Wiring diagram. (*b*) Voltage-phasor diagram for the high-voltage (wye) side. (*c*) Voltage-phasor diagram for the low-voltage (delta) side.

formulated to provide good electrical insulation, and also serves to carry heat away from the core and windings by convection. In smaller transformers, the hot oil is cooled by flowing along the inside of the tank wall. Larger transformers have ribbed walls or are provided with fins or radiators to cool the oil. Very large transformers have pumps to circulate the oil through external radiators, which may be provided with fans. Transformers which are designed to operate in air are called "dry-type" transformers.

Bushings. Leads from the transformer windings are brought out of the tank through bushings inserted in openings in the tank wall or lid. When high voltages are involved, these bushings may be quite complex, consisting of concentric capacitors enclosed in a ceramic case filled with oil. The capacitors form a voltage divider which distributes the electric field between the transformer lead and the case in such a way as to minimize the probability of breakdown of the insulation. *See* CAPACITOR.

Figure 8 shows a cutaway 20-kVA, single-phase, shell-form, distribution transformer, designed for mounting on a utility pole. The tank is normally three-quarters filled with insulating, cooling oil. The transformer has one high-voltage bushing; the other terminal of the high-voltage winding is connected to the steel tank. A spark gap connects the high-voltage terminal to a lightning arrester, which protects the transformer insulation from overvoltage. Immediately above the core, inside the tank, is an overcurrent circuit breaker. Such a transformer is said to be completely self-protected (CSP). *See* CIRCUIT BREAKER; SURGE ARRESTER.

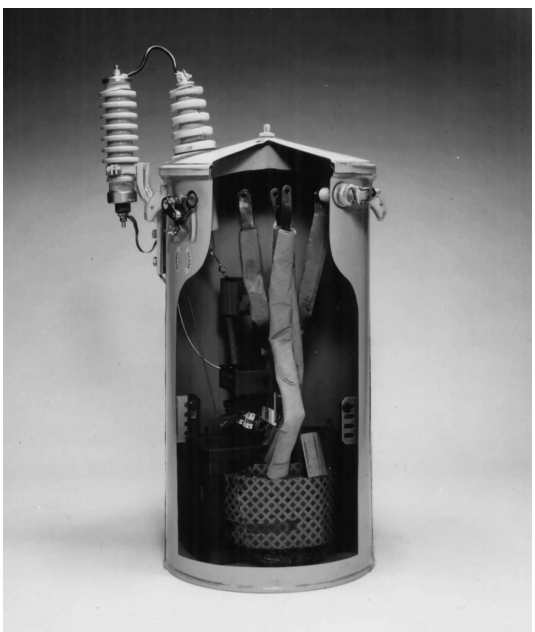

Fig. 8. Cutaway of single-phase, pole-type, completely self-protected (CSP) distribution transformer. (*Westinghouse Electric Corp.*)

George McPherson, Jr.

Audio- and Radio-Frequency Transformers

A second important application for transformers is in signal processing. Transformers are used to deliver a waveform of some kind from one circuit to a second (for example, the output of one transistor or integrated circuit to the input of another). Energy is delivered to the second circuit, and the manner of utilization of the energy is extremely important.

With signal processing, it is vital that the waveform of the signal be controlled. For this to be possible, the signal energy must be delivered to a load that is usually linear; often it should also be resistive. If the transformer is unloaded, the signal energy is used to induce magnetization in the transformer core, leading to harmonic and intermodulation distortion due to magnetic hysteresis, and frequency distortion due to the variation of the input reactance of the transformer primary. The resistive load must be placed on the secondary output to minimize the adverse effects of the nonlinear magnetizing current. In addition, the load impedance must be chosen so that it swamps out the magnetizing reactance effect and introduces the desired load impedance characteristics. It can be chosen to reduce the effects of variation of the impedance as a function of signal voltage developed on the secondary, since these variations can also introduce both harmonic and intermodulation distortion. As discussed below, the load may also be selected in a way that will limit the magnitude of the phase distortion introduced by a following amplifier and will assure stable operation. (This distortion often is called departure from the minimum-phase condition, and is a result of parasitic feedback paths in a circuit.) *See* DISTORTION (ELECTRONIC CIRCUITS).

Audio or video (broad-frequency-band) transformers are used to transfer complex signals containing energy at a large number of frequencies from one circuit to another. Radio-frequency (rf) and intermediate-frequency (i-f) transformers are used to transfer energy in narrow frequency bands from one circuit to another. Audio and video transformers are required to respond uniformly to signal voltages over a frequency range three to five or more decades wide (for example, from 10 to 100,000 Hz), and consequently must be designed so that very nearly all of the magnetic flux threading through one coil also passes through the other. These units are designed to have a coupling coefficient k, given in Eq. (29),

$$k = \frac{M}{\sqrt{L_1 L_2}} \tag{29}$$

nearly equal to 1. Here L_1 and L_2 are the primary and secondary inductances, respectively, and M is the mutual inductance (**Fig. 9**). The high coupling coefficient is obtained by the use of interleaved windings and a high-permeability iron core, which concentrates the flux. Typical values of k for highest-quality video transformers may be as high as 0.9998; that for power transformers need not be greater than 0.98.

All radio-frequency and intermediate-frequency

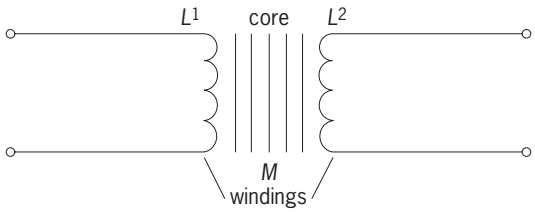

Fig. 9. Schematic of a transformer with symbols.

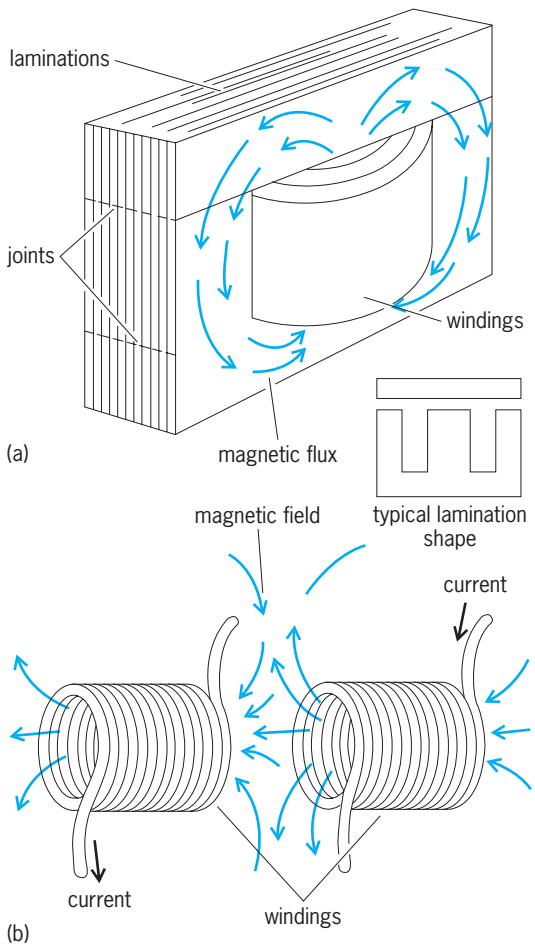

Fig. 10. Audio- and radio-frequency transformers. (*a*) Iron-core audio transformer. (*b*) Air-core radio-frequency transformer.

transformers are built from individual inductors whose magnetic fields are loosely coupled together, $k < 0.30$; one or more inductors are resonated with a capacitor to make efficient energy transfer possible near the resonant frequency. The structure of audio- and radio-frequency transformers is shown in **Fig. 10**. *See* RESONANCE (ALTERNATING-CURRENT CIRCUITS).

Audio and video transformers. Audio and video transformers have two resonances (caused by existing stray and circuit capacitances) just as many tuned transformers do. One resonance point is near the low-signal-frequency limit; the other is near the high limit. As the coefficient of coupling in a transformer is reduced appreciably below unity by removal of core material and separation of the windings, tuning capacitors are added to provide efficient transfer of energy. The two resonant frequencies combine to one when the coupling is reduced to the value known as critical coupling, then stay relatively fixed as the coupling is further reduced. It is possible for a single-tuned coupled circuit to be overcoupled, leading to a broad-tuned coupled circuit. This factor necessitates resistive loading on audio and video transformers. Overcoupling is also observed in the input coupled circuits commonly used in television receivers, and causes the susceptibility to spurious signals often noted in such receivers.

All transformers are devices for transferring energy from one circuit to another. The energy transferred is absorbed either in the circuits themselves or in an external load circuit. For this reason, proper termination is essential for achieving optimum behavior in circuits containing transformers.

Audio and video transformers have a minimum operating frequency at which the open-circuit reactance of the primary is approximately twice its effective loaded impedance. As with wide-band resistance-capacitance (*RC*) amplifiers, gain may be traded for bandwidth with transformer-coupled amplifiers. The reduction of the terminating resistance across the secondary of the transformer reduces the minimum operating frequency f_{min} and, in the presence of output capacitance, raises the maximum frequency f_{max}. The approximate values of the minimum and maximum frequencies and the resonant frequency f_r are given by Eqs. (30), where

$$f_{min} = \frac{R_c(N_1/N_2)^2}{\pi L_1}$$
$$f_{max} = \frac{1}{2\pi R_c C_2} \qquad (30)$$
$$f_r = \frac{1}{2\pi\sqrt{L_{22}C_2}}$$

$L_{22} = L_2 - (M^2/L_1) = L_2(1 - k^2)$. This L_{22} is the secondary inductance with the primary short-circuited; C_2 is the output capacitance, both external and internal, on the transformer; and R_c is the load resistance (**Fig. 11**). The resonant frequency f_r should be larger than f_{max} for best performance. *See* AMPLIFIER.

A transformer used to activate terminating circuitry is called an output transformer; one to activate an input circuit is an input transformer; and others are called interstage transformers.

Distortion. The distortion introduced into the amplified signal by a transformer is caused primarily by its hysteresis loss. This loss effect may be minimized by proper loading on the secondary. The load component of current then is large compared to the magnetizing current. In addition, a resistive load keeps

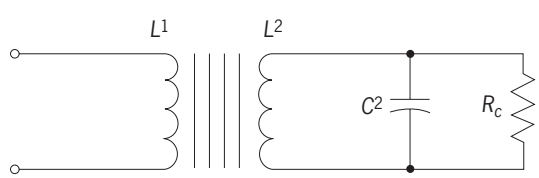

Fig. 11. Circuit of loaded transformer.

the amplification uniform as a function of frequency, and keeps the phase distortion to a minimum.

The magnetic core in an audio or video transformer is subject to two kinds of saturation, that due to applied direct current in the windings, and that due to excessively large signal currents. The direct current in the windings may make the hysteresis loop of the iron core nonsymmetrical, necessitating the use of a larger core having a built-in air gap. Both large signal amplitudes and low frequencies can cause signal saturation to occur in the core. *See* DISTORTION (ELECTRONIC CIRCUITS).

The development of ultracompact radios and other consumer electronics has caused critical problems for component designers. The problems are particularly difficult for units that employ bipolar transistors. The total volume of iron and copper severely influences the impedance and the coupling coefficient as well as the signal amplitude and the saturation characteristics of a transformer. The overall result is limitation of signal amplitude, frequency response, and power-handling capability.

RF and i-f transformers. These use two or more inductors, loosely coupled together, to limit the band of operating frequencies. Efficient transfer of energy is obtained by resonating one or more of the inductors. By using higher than critical coupling, a wider bandwidth than that from the individual tuned circuits is obtained, while the attenuation of side frequencies is as rapid as with the individual circuits isolated from one another.

The tuning of the primary, the secondary, or both may be accomplished either by the variation of the tuning capacitor or by an adjustable magnetic or conducting slug that varies the inductance of the inductor (**Fig. 12**).

Impedance. The operating impedance of a tuned circuit of a radio-frequency transformer is a function of its Q and its tuning capacitance. (Its Q measures the rate of dissipation of stored energy in a tuned circuit as it decays from a starting pulse.) In general, high-power circuits require a high capacitance for energy storage, and therefore have low values of impedance. In any application, the impedance level must be kept sufficiently small to prevent instability and oscillation. *See* Q (ELECTRICITY).

Transformer-coupled amplifiers. Control of voltage gain of transformer-coupled amplifiers is a crucial requirement for effective, stable, radio-frequency and intermediate-frequency amplifiers, both low-power and high-power types. Voltage gain is a function of tuned impedance and the transconductance of the active device, whether tube or transistor.

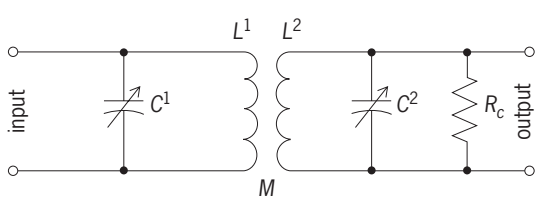

Fig. 12. Tuned radio-frequency transformer.

The tuned impedance Z_t of a tuned circuit depends on both the operating Q of the tuned circuit and its inductance-to-capacitance ratio. (Its operating frequency is a function of the product of inductance L and capacitance C.) Impedance of a capacitor is an inverse function of capacitance and frequency, according to Eq. (31).

$$Z_t \text{ (tuned)} = \frac{L}{CR} = \text{ function of } \frac{L}{C^{1/2}} \tag{31}$$

The voltage gain takes the form of the product of the transconductance and the tuned impedance, $g_m Z_t$. The transconductance can be expressed in terms of Eq. (32), where kappa κ is the

$$g_m = \kappa \left(\frac{q}{kT}\right) I_c \tag{32}$$

transconductance-per-unit-current efficiency, q is the electron charge, k is Boltzmann's constant, T is absolute temperature, and I_c is the instantaneous value of the output current for the device. The kappa for a bipolar transistor is approximately unity; for field-effect transistors and electron tubes, it normally is less than 0.02. The value of q/kT is approximately 39 siemens per ampere. A load impedance that gives an overall stage voltage gain of 10 or less in a circuit will usually operate in the proper mode and can be stabilized. Should the load impedance determined in this way prove to be too large, a radio-frequency transformer may be used to reduce the impedance level. The critical issue is that the input circuit to the amplifier look into a voltage source so that input loading will not degrade the operation of the previous amplifier. *See* CIRCUIT (ELECTRONICS); RADIO-FREQUENCY AMPLIFIER; SIGNAL PROCESSING.

Keats A. Pullen

Bibliography. W. M. Flanagan, *Handbook of Transformer Design and Applications*, 2d ed., 1993; I. M. Gottlieb, *Practical Transformer Handbook*, 1998; M. J. Heathcote, *J&P Transformer Book*, 12th ed., 1998; S. L. Herman and D. Singleton, *Delmar's Standard Guide to Transformers*, 1996; D. Horne (ed.), *Electrical Transformer Handbook*, vol. 2, 2005; R. W. Hurstled (ed.), *Electrical Transformer Handbook*, vol. 1, 2004; R. Lee, C. E. Carter, and L. Wilson, *Electronic Transformers and Circuits*, 3d ed., 1988; E. Lowdon, *Practical Transformer Design Handbook*, 2d ed., 1988; G. McPherson and R. D. Laramore, *An Introduction to Electrical Machines and Transformers*, 2d ed., 1990; A. J. Pansini, *Electrical Transformers and Power Equipment*, 3d ed., 1999; R. M. Del Vecchio et al., *Transformer Design Principles with Applications to Core-Form Power Transformers*, 2002.

Transfusion

The administration of blood or its components as a part of a medical treatment. There are certain fairly well-delineated indications for the use of some form of transfusion. Hemorrhage, severe burns, and certain forms of shock are perhaps the most important

conditions for which blood transfusion is utilized. Other disorders in which hemotherapy may be indicated include hemophilia, leukemia, certain anemias, and rare hereditary or familial disorders in which some portion of the blood is lacking or deficient. *See* HEMATOLOGIC DISORDERS.

Blood groups. In order for a recipient to accept a blood transfusion, the donor blood cells must be immunologically compatible with the recipient. That is, the recipient must recognize certain molecules (antigens) on donor blood cells as "self" and not foreign. The three main antigen systems on blood cells are the ABO, Rh, and HLA.

The ABO antigens are the most important for transfusion of red blood cells. A person may have A, B, both A and B, or neither A nor B antigens on their blood cells. Persons who do not have A antigen have anti-A antibody in their plasma, which would destroy red cells having A antigen; the same conditions apply to the B antigen. Therefore, red cells can be transfused from one person to another only if the recipient does not have antibodies against the antigens of the donor. Since O-type individuals have neither antigen, their blood can be given to people with types A, B, or AB blood.

The Rh antigens are the next most important factor in transfusion of red blood cells. Persons who lack one or more of the Rh antigens on their red cells (that is, they are Rh negative) do not have antibody against the Rh antigen. However, if an Rh negative person has been previously exposed to Rh by transfusion of Rh positive blood or if an Rh negative woman has been pregnant with an Rh positive infant, then that individual will develop antibodies against the Rh antigen. *See* RH INCOMPATIBILITY.

The human leukocyte antigen (HLA) complex is another factor to be considered in transfusions. The HLA antigens are found on platelets. If the donor HLA antigens are not the same as those in the recipients, the recipient may produce antibodies that destroy the donor platelets. *See* BLOOD GROUPS.

Blood collection and component preparation. Collection and preparation of blood follows certain general standards worldwide, but specifics may vary. The following discussion applies to the United States. Blood donors are screened to identify risks for the donor, such as any possible adverse effects of losing a unit of blood, as well as for the potential recipient. Blood donation is safe for healthy individuals. The blood is collected with sterile, disposable equipment so that there is no possibility of disease transmission to the donor. To determine whether the donor's blood is safe for recipients, donors are asked to complete a questionnaire about their past and present physical condition. In addition, after the blood donation the donor can notify the blood bank confidentially of any concerns regarding the safety of the blood for recipients. Donated blood is tested for blood groups, blood-group antibodies, and laboratory evidence of syphilis, hepatitis B, AIDS, human T-cell lymphotropic virus type I (HTLV-I, which is associated with adult T-cell leukemia), and hepatitis C. As a result, blood transfusions have become safer. However, persons who may have been exposed to AIDS should not donate blood, because in rare cases it has been found that blood may test negative for AIDS and yet still be capable of transmitting AIDS to recipients. This situation can arise because there is a period of time during which a recently infected individual has not yet made sufficient antibody to test positive.

Because of the concern that AIDS can be transmitted by blood transfusion, patients sometimes request donations from specific family members and friends. In general, such directed donations are statistically no safer than volunteer blood donation. More patients are donating their own blood (autologous blood) before elective surgery, for their own use during and after the surgery. Autologous blood is the safest blood for transfusion.

Generally, each donation consists of 1 pint (450 ml) of blood. The blood, which is collected in sterile plastic bags, can be separated into several components, such as red blood cells, plasma, and platelets (see **table**). Each component may be used as needed; often, these components are given to different recipients according to their specific needs. In addition, plasma from different donors can be pooled and made into plasma components such as Factor VIII or IX concentrate for hemophiliacs. *See* HEMOPHILIA.

Uses. Shortly before transfusion, the blood of the donor and recipient is tested once again to make sure that the blood groups are compatible. In an emergency, these tests are abbreviated, or type O red blood cells which can be transfused safely to any individual, are used. Transfusion of blood and blood components is essential to support many patients undergoing surgery, treatment for cancer, or organ transplantation, as well as premature infants. *See* TRANSPLANTATION BIOLOGY.

Most patients undergoing elective surgery do not need blood tranfusions. In some instances, the blood shed by the patient can be collected during surgery and given back to the patient. However, in emergency surgery, patients must rely on blood from volunteer donors. For the rare case in which a patient has lost a great deal of blood, massive transfusion may create special problems. Sufficient blood of the

Storage and indications for use of whole blood and common blood components

	Storage	Indications
Whole blood	35 days at 39°F (4°C)	Massive bleeding
Red blood cells	42 days at 39°F (4°C)	Inadequate red blood cells to deliver oxygen to tissues
Plasma	12 months at 0°F (−18°C)	Multiple coagulation deficiencies inadequate to stop or prevent bleeding
Platelets	5 days at room temperature	Inadequate platelets to stop or prevent bleeding

correct type may not be available. Also, during transfusion such a patient may develop abnormal bleeding, low body temperature, or shock and may require special blood components such as platelets and plasma to stop the bleeding.

Cancer patients frequently need blood transfusions. For example, since blood is produced in the bone marrow, patients with cancer of the bone marrow may fail to make enough of their own blood. Treatment with drugs and radiation often suppresses the production of blood in the bone marrow even further, so that transfusions must be given until the patient's bone marrow recovers. *See* CANCER (MEDICINE).

Premature infants often need blood transfusions, partly because blood drawn for laboratory tests constitutes a large percentage of their blood volume. They are, however, especially susceptible to certain diseases that are transmitted by transfusion, such as cytomegalovirus, and may need special blood.

Adverse effects. In addition to transmission of infections, adverse effects of blood transfusion are due to immune reactions between donor and recipient. Fever, the most frequent reaction, is caused by reaction of recipient antibody against donor white blood cells. Hives are due to allergic reactions to substances in donor plasma. Destruction of donor red cells (hemolytic reaction) occurs if the wrong type of blood is inadvertently given, or if recipient antibody is not detected prior to transfusion. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); BLOOD.

Pearl Toy

Bibliography. E. Rossi, T. Simon, and G. Moss (eds.), *Principles of Transfusion Medicine*, 2d ed., 1996; R. Walker (ed.), *Technical Manual*, 11th ed., 1993.

Transistor

A solid state device involved in amplifying small electrical signals and in processing of digital information. Transistors act as the key element in amplification, detection, and switching of electrical voltages and currents. They are the active electronic component in all electronic systems which convert battery power to signal power. Almost every type of transistor is produced in some form of semiconductor, often single-crystal materials, with silicon being the most prevalent. There are several different types of transistors, classified by how the internal mobile charges (electrons and holes) function. The main categories are bipolar junction transistors (BJTs) and field-effect transistors (FETs).

Single-crystal semiconductors, such as silicon from column 14 of the periodic table of chemical elements, can be produced with two different conduction species, majority and minority carriers. When made with, for example, 1 part per million of phosphorus (from column 15), the silicon is called *n*-type because it adds conduction electrons (negative charge) to form the majority carrier. When doped with boron (from column 13), it is called *p*-type because it has added positive mobile carriers called holes. For *n*-type doping, electrons are the majority carrier while holes become the minority carrier. For *p*-type doping holes are in larger numbers, hence are the majority carriers, while electrons are the minority carriers. All transistors are made up of regions of *n*-type and *p*-type semiconducting material. *See* SEMICONDUCTOR; SINGLE CRYSTAL.

The bipolar transistor has two conducting species, electrons and holes. Field-effect transistors can be called unipolar because their main conduction is by one carrier type, the majority carrier. Therefore, field-effect transistors are either *n*-channel (majority electrons) or *p*-channel (majority holes). For the bipolar transistor, there are two forms, n^+pn and p^+np, depending on which carrier is majority and which is the minority in a given region. As a result the bipolar transistor conducts by majority as well as by minority carriers. The n^+pn version is by far the most used as it has several distinct performance advantages, as does the *n*-channel for the field-effect transistors. (The n^+ indicates that the region is more heavily doped than the other two regions.)

Bipolar transistors. Bipolar transistors have additional categories: the homojunction for one type of semiconductor (all silicon), and heterojunction for more than one (particularly silicon and silicon-germanium, $Si/Si_{1-x}Ge_x/Si$). At present the silicon homojunction, usually called the BJT, is by far the most common. However, the highest performance (frequency and speed) is a result of the heterojunction bipolar transistor (HBT).

Bipolar transistors are manufactured in several different forms, each appropriate for a particular application. They are used at high frequencies, for switching circuits, in high-power applications, and under extreme environmental stress. The bipolar junction transistor may appear in discrete form as an individually encapsulated component, in monolithic form (made in and from a common material) in integrated circuits, or as a so-called chip in a thick-film or thin-film hybrid integrated circuit. In the *pn*-junction isolated integrated-circuit n^+pn bipolar transistor, an n^+ subcollector, or buried layer, serves as a low-resistance contact which is made on the top surface (**Fig. 1**). *See* INTEGRATED CIRCUITS; JUNCTION TRANSISTOR.

Basic operation. The n^+pn bipolar transistor (**Fig. 2***a*) has three differently doped regions and two junctions, the emitter-base junction and the collector-base junction, in a single crystal of silicon. It is

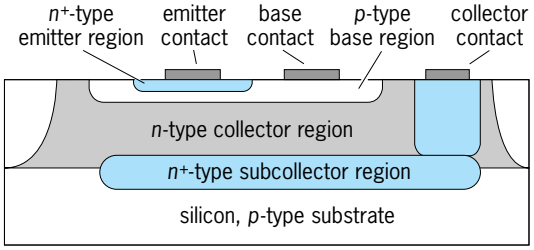

Fig. 1. Isolated n^+pn bipolar junction transistor for integrated-circuit operation.

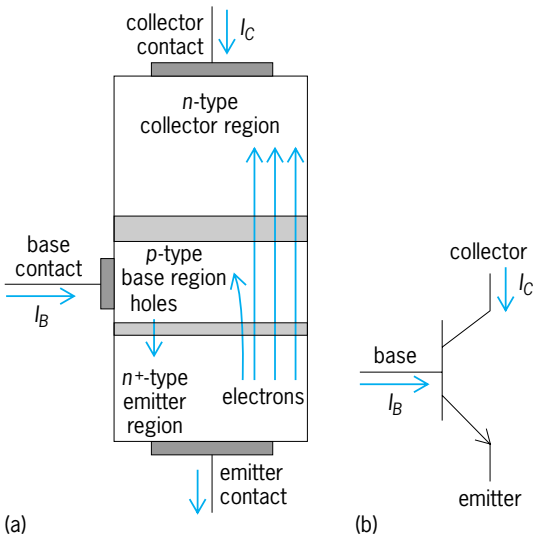

Fig. 2. Operation of an n^+pn bipolar junction transistor. (*a*) Conceptual cross section with carrier flows. (*b*) Circuit symbol.

possible to describe a large part of bipolar transistor operation by interpreting the device somewhat like a pair of back-to-back diodes. The emitter-base junction is usually forward-biased; that is, the voltage of the base with respect to the emitter, V_{BE}, is greater than 0. This voltage is small (less than 1 V), but the forward current across the junction is relatively large, just as in the forward-biased diode. The majority carriers in the emitter region (n^+-type) are electrons. These are injected across the emitter-base junction into the base region (*p*-type), which is quite thin, being on the order of 0.1–0.5 micrometer. With the base-collector junction reversed-biased (that is, the voltage of the collector with respect to the base, V_{CB}, is greater than 0), the holes that are present in the *p*-type base material do not cross it, as they are repelled by the electric field from the collector. However, the electrons that have been injected from the emitter into the very thin base region (where they become minority carriers) diffuse across the base-collector junction and are then collected under the influence of the positive collector potential (Fig. 2*a*). This electron current across the base-collector junction is almost as large as the electron current crossing from the emitter into the base region. In general, it runs from 99.5 to 99.99% of the emitter current, the small decrease being accounted for by electrons that are lost in the narrow base region. In terms of electron currents, for every 200 electrons flowing into the emitter from the external contact and crossing into the base region, perhaps one causes a current in the base lead, whereas 199 cross over into the collector and flow out into the external collector circuit. *See* DIFFUSION; JUNCTION DIODE; SEMICONDUCTOR DIODE.

The total current across any junction results from both hole and electron motion. In the emitter, electrons are injected from the n^+ region into the *p* region of the base, while holes are injected from the *p*-type base back into the emitter. These holes also cause current to flow into the base lead. The sum of these two currents crossing the junction determines the total emitter current. Even though holes and electrons flow in opposite directions, they add together as total current. However, in the emitter-base junction this hole current is usually several orders of magnitude less than the electron current, a consequence of much heavier doping in the emitter region compared to the base. The total base current is made up of two small components: the holes injected from the base to the emitter, and the electrons lost from the emitter due to recombination in the base. It is desirable for the base current to be as close to zero as possible, in order to obtain large current gains from base to collector.

Circuit symbol. The circuit symbol for an n^+pn bipolar transistor (Fig. 2*b*) has an arrowhead on the emitter lead whose direction indicates that electrons flowing in from the emitter to the base are the same as if positive charges were leaving that terminal. Hence, the arrowhead points out of the emitter.

Voltage-current characteristics. The input voltage-current characteristics of an n^+pn transistor (**Fig. 3***a*) are a family of curves of base current, I_B, versus base-emitter voltage, V_{BE}, for various values of collector-emitter voltage, V_{CE}. The output characteristics (Fig. 3*b*) are curves of collector current, I_C, versus collector-emitter voltage, V_{CE}, with the base current, I_B, as an independent parameter. These curves display two general characteristics. First, once the collector-emitter voltage is greater than 1 or 2 V, the collector current is relatively independent of this voltage; that is, each curve flattens out. Second, in this region the collector current is about 100 times the base current. The ratio of collector current to base current with the collector base voltage, V_{BE}, equal to zero is defined as the dc beta (β_{dc}) for the bipolar transistor.

The voltage-current characteristics emphasize the regions where the emitter-base junction is forward-biased (the base-emitter voltage, V_{BE}, is greater than 0) and the collector-base junction is reverse-biased (the collector-base voltage, V_{CB}, is greater than 0). These bias conditions define the active region for a bipolar transistor, which is mainly used for analog circuits. A quick glance at the input characteristics shows that the base current is greater than 0 under these conditions. When both junctions are reverse-biased, any currents that flow are several orders of magnitude smaller, and this is termed the cutoff region ("off" for digital circuits). The condition that both junctions are forward-biased defines the saturation region or the "on" region for digital circuits. Both the base-emitter voltage, V_{BE}, and the base-collector voltage, V_{BC}, are small positive values; and the collector-emitter voltage, V_{CE}, their difference, is also small and positive. The boundary between the active and saturation regions occurs where the collector-base voltage, V_{CB}, equals 0.

p^+np transistor. The p^+np transistor contains a narrow *n*-type base layer sandwiched between *p*-type emitter and collector regions. Forward bias on the emitter-base junction causes holes to be injected into

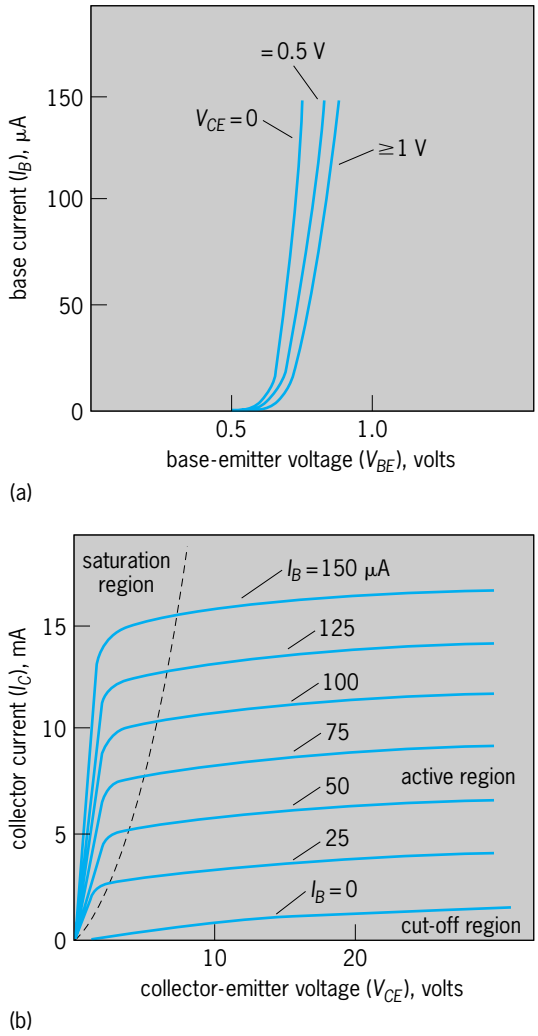

Fig. 3. Voltage-current characteristics of a typical low-power n^+pn bipolar transistor. (*a*) Input characteristics: base current (I_B) versus base-emitte voltage (V_{BE}) for various values of collector-emitter voltage (V_{CE}). When V_{CE} >1 V, the curves coincide. (*b*) Output characteristics: collector current (I_C) versus collector-emitter voltage (V_{CE}) for various values of base current (I_B). (*After W. H. Hayt, Jr., and G. W. Neudeck, Elecronic Circuit Analysis and Design, 2d ed., Wiley, reprint, 1995*)

the base region, most of which diffuse across the thin base and are collected by the reverse-biased base-collector junction. The total emitter current, I_E, is thus composed of the sum of this large hole current plus a much smaller electron current directed from the base toward the emitter.

Field-effect transistors. Majority-carrier field-effect transistors are classified as metal-oxide-semiconductor field-effect transistor (MOSFET), junction "gate" field-effect transistor (JFET), and metal "gate" on semiconductor field-effect transistor (MESFET) devices. MOSFETs are the most used in almost all computers and system applications. However, the MESFET has high-frequency applications in gallium arsenide (GaAs), and the silicon JFET has low-electrical noise performance for audio components and instruments. In general, the *n*-channel field-effect transistors are preferred because of larger electron mobilities, which translate into higher speed and frequency of operation.

MOSFETs. An *n*-channel MOSFET (**Fig. 4**) has a so-called source, which supplies electrons to the channel. These electrons travel through the channel and are removed by a drain electrode into the external circuit. A gate electrode is used to produce the channel or to remove the channel; hence it acts like a gate for the electrons, either providing a channel for them to flow from the source to the drain or blocking their flow (no channel). With a large enough voltage on the gate, the channel is formed, while at a low gate voltage it is not formed and blocks the electron flow to the drain. This type of MOSFET is called enhancement mode because the gate must have sufficiently large voltages to create a channel through which the electrons can flow. Another way of saying the same idea is that the device is normally "off" in an nonconducting state until the gate enhances the channel.

An *n*-channel MOSFET with a positive gate-source voltage, V_{GS}, and a small drain-source voltage, V_{DS}, has an electric field established across an insulating layer (Fig. 4). This field acts to repel positive carriers (holes) in the substrate and to attract negative carriers (electrons). As a result, a layer of substrate near the insulator becomes less *p*-type and its conductivity is reduced. As the gate-source voltage increases further, this surface region of the substrate eventually has more electrons than holes, and it inverts to *n*-type. Additional increases in gate voltage add more electrons to the channel and make it even more conductive. This *n*-channel (Fig. 4*a*) now conducts electrons from the n^+ source to the n^+ drain which has a positive voltage and attracts electrons. Between the *p*-type substrate and the *n*-type channel is a depletion (transition) region that serves to isolate the substrate from the channel, a process referred to as self-isolation. Since conduction is by electrons, the majority carrier, the MOSFET is a majority-carrier device.

The smallest value of the gate-source voltage, V_{GS}, that will produce a channel and a resultant value of drain current, I_D, greater than the few nanoamperes is called the threshold voltage, V_T, typically 0.2–2 V. V_T output voltage-current characteristics of the device are a family of curves of drain-current, I_D, versus drain-source voltage, V_{DS}, for several values of gate-source voltage, V_{GS} (**Fig. 5**). When the drain-source voltage is small (Fig. 5*a*), the device behaves as a voltage-controlled linear resistance. When the drain-source voltage becomes sufficiently large (Fig. 5*b*), the gate-to-drain voltage is less than the

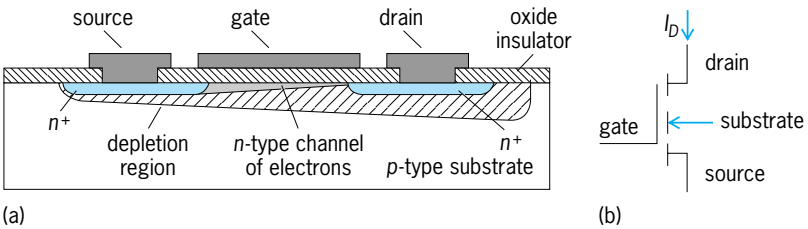

Fig. 4. An *n*-channel enhancement-mode metal-oxide-semiconductor field-effect transistor (MOSFET). (*a*) Cross section. (*b*) Standard circuit symbol.

threshold voltage, that is, Eq. (1) holds, and pinch-

$$V_{GD} = V_{GS} - V_{DS} \leq V_T \qquad (1)$$

off occurs at the drain end of the channel. Further increases in drain-source voltage do not lead to larger values of drain current, (that is, the current saturates), since the transistor is operating in the region beyond pinch-off. In this region of operation the MOSFET device behaves as a voltage-controlled current source.

The standard circuit symbol for the *n*-channel enhancement-mode MOSFET (Fig. 4*b*) shows the substrate as a separate connector. An arrow shows the direction from the *p* side (substrate) to the *n* side (channel) of the junction, while a segmented line indicates the enhancement mode; no channel is present until channel enhancement occurs at which point the gate-source voltage exceeds the threshold voltage.

The *p*-channel enhancement-mode MOSFET is the complement of the *n*-channel device. It has an *n*-type silicon substrate in which a *p*-type channel is induced (enhanced) by making the gate sufficiently negative that the gate-source voltage is less than the threshold voltage. The gate of a *p*-channel enhancement-mode MOSFET has an electric field between the gate and

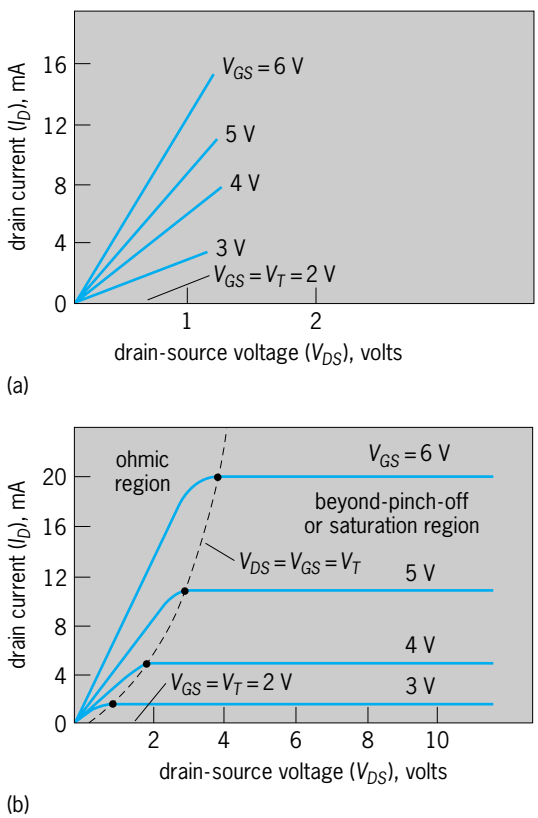

Fig. 5. Output characteristics of *n*-channel enhancement-mode MOSFET: drain current (I_D) versus drain-source voltage (V_{DS}) for various values of gate-source voltage (V_{GS}). (*a*) Small values of V_{DS}, where the device behaves as a voltage-controlled linear resistance. (*b*) Complete output characteristics. (*After W. H. Hayt, Jr., and G. W. Neudeck, Electronic Circuit Analysis and Design, 2d ed., reprint, Wiley, 1995*)

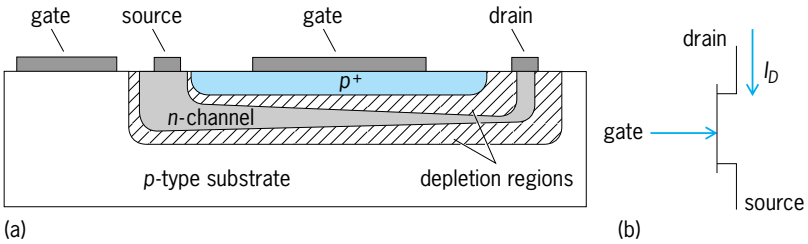

Fig. 6. An *n*-channel junction field-effect transistor (JFET). (*a*) Cross section. (*b*) Circuit symbol.

substrate which pushes out electrons, attracts holes, and eventually inverts the channel to *p* type. Now holes conduct between the p^+ source and drain electrodes.

JFETs. In the JFET (**Fig. 6***a*), a conducting majority-carrier *n* channel exists between the source and drain. When a negative voltage is applied to the p^+ gate, the depletion regions widen with reverse bias and begin to restrict the flow of electrons between the source and drain. At a large enough negative gate voltage (symbolized V_P), the channel pinches off. The standard circuit symbol (Fig. 6*b*) has a continuous bar since current flows with zero gate-source voltage, V_{GS}, at larger values of the drain source voltage, V_{DS}.

MESFETs. The MESFET is quite similar to the JFET in its mode of operation. A conduction channel is reduced and finally pinched off by a metal Schottky barrier placed directly on the semiconductor. Metal on gallium arsenide is extensively used for high-frequency communications because of the large mobility of electrons, good gain, and low noise characteristics. Its cross section is similar to that of the JFET (Fig. 6*a*), with a metal used as the gate. *See* SCHOTTKY BARRIER DIODE.

High-frequency transistors. High-frequency effects for the bipolar transistor are characterized by the emitter charging time (τ_e), the collector charging time (τ'_c), the minority-carrier transit time through the active base region (τ_b), and the base-collector depletion region transit time (τ_c). The emitter charging time equals the product of the emitter-base capacitance (proportional to the area of the emitter) and the thermal voltage divided by the dc current. The minority-carrier transit time through the active base region is approximately the square of the active width of the base region divided by twice the diffusion constant for the minority carriers that diffuse through the base. (The dependence on the active width indicates the need for a very thin base region.) The transit time through the collector-to-base depletion region equals the width or this region divided by twice a saturated velocity to which the carriers can accelerate. Thus, a short transit time requires a large saturated velocity or a small width, which means a small value of the collector-to-base voltage. The final term for the collector is its charging time, approximately the product of the collector contact resistance and the collector-base capacitance. A short charging time thus requires a small value of the former and a small collector area to reduce the latter.

Figures of merit. A figure of merit for the advanced bipolar transistor is the frequency, f_T, at which the short-circuit, current-signal gain is unity. This frequency equals the inverse of the sum of the four times discussed above, τ_e, τ_b, τ_c, and τ'_c. A large value of f_T indicates that the intrinsic device is fast.

A more circuit-oriented figure of merit is $f_{\max}$, the maximum frequency that gain can still be achieved in a circuit, given by Eq. (2). Here the external base

$$f_{\max} \cong \sqrt{\frac{f_T}{8\pi R_b C_{cb}}} \tag{2}$$

resistance, R_b, is important, as well as the base-collector area needed to reduce the collector base capacitance, C_{cb}.

Most very high speed logic circuits belong to the emitter-coupled logic (ECL) family of circuits or the current-mode logic (CML) family. The figure of merit for this type of circuit is given by Eq. (3), where R_c is

$$\tau_{cs} = 1.7\sqrt{\frac{(R_c + 2R_b)(3C_{cb} + C_{cs})}{2\pi f_{T\max}}} \tag{3}$$

the collector resistance, C_{cs} is the collector-substrate capacitance of the integrated bipolar transistor, and $f_{T\max}$ is the peak value of f_T when the collector current is varied. Again, this expression indicates the need for thin base regions, small emitter and collector areas, and low values of resistances contacting the device. *See* LOGIC CIRCUITS.

Structural improvements. The function of the subcollector in the integrated-circuit bipolar transistor (Fig. 1) is to reduce the collector resistance. Typical values of the current gain, that is, the dc beta, range from 80 to 300, and f_T ranges from 5 to 45 GHz with values of $f_{\max}$ up to 450 GHz. In an emitter-coupled logic circuit the transistor has a gate delay of as low as 15 picoseconds. An improvement to this structure is to reduce the sidewall components of capacitance with the local-oxidation-of-silicon (LOCOS) structure. In addition, a polysilicon-contacted emitter can be added to improve the dc beta, and the external base resistance can be reduced by increased base doping, somewhat similar to what is done in heterojunction bipolar transistors.

Single, self-aligned transistor. The single self-aligned bipolar transistor (SST) reduces the emitter area to 0.35×5 μm and has f_T values up to 20 GHz and a dc beta of 180. The use of a pedestal collector and double self-alignment improves the value of $f_{\max}$ and the emitter charging time by reducing the area and the external parasitic resistances. In all these cases the fabrication methods strive to reduce the area, hence the capacitances.

Heterojunction bipolar transistor. The heterojunction bipolar transitor is made from two different types of semiconductor material. The most promising is the silicon-germanium type. It is produced by epitaxially growing a narrow band-gap base region of heavily doped p-type $Si_{1-x}Ge_x$ on an n-type silicon collector and then capping it with an n^+ type silicon emitter. The silicon-germanium compound suppresses the base-injected holes (Fig. 2), and at the same time this allows the base to be doped very heavily to reduce the external base resistance. By grading the germanium content and the doping, f_T, values up to 32 GHz and $f_{\max}$ of 120 GHz have been achieved, with good values of beta. The circuit delay is about 20 ps. Other heterojunction bipolar transistors of interest include those using the compound semiconductors GaAlAs/GaAs, InGaP/GaAs, and InGaAs/InP. These devices have achieved f_T values of 37 GHz and $f_{\max}$ of 90 GHz with powers of 1–5 W. *See* SEMICONDUCTOR.

High-frequency field-effect transistors. The inability of the MOSFET to conduct large currents into capacitive loads has limited its use in extremely high-frequency circuits. However, because of its low power consumption it can be integrated into very dense circuits. The first requirement is that the channel length be small (approximately 1 μm), as it controls how fast the majority carrier can traverse between the source and drain. The carrier mobility must also be large; hence, electrons are preferred, as their mobility is typically two to three times larger than holes. A second requirement is for low values of source and drain resistance. In circuit applications a small value of capacitance between the gate and drain is necessary to reduce the total effective capacitance that is multiplied by the circuit voltage gain. Self-aligned gates, polysilicon, and channel lengths of less than 0.15 μm are used. Typical performance characteristics for a 0.5-μm gate length are an f_T of 10 GHz and an $f_{\max}$ of 15 GHz. In complementary metal-oxide-semiconductor (CMOS) circuits with gate lengths of 0.15 μm, gate delays as low as 21 ps per stage are possible.

The more advanced techniques use silicon-on-insulator (SOI) technology to further reduce the external parasitic capacitances around the source and drain. Other device structures include the high electron mobility transistor (HEMT), silicon-germanium MOSFETs, and combinations of bipolar transistors and MOSFETs (BiCMOS). Each technology has its particular advantages. The HEMT is produced from compound semiconductors and can yield an f_T of 300 GHz with gate delays of 25 ps. *See* MICROWAVE SOLID-STATE DEVICES.

Gerold W. Neudeck

Models. Whether the transistor is used in the design of small analog circuits or very large scale integrated circuits, its behavior has to be adequately understood by the designer. Analysis of the circuit is a prerequisite to its fabrication, thus pointing to the need for models. The higher levels of integration as well as of the cost of fabrication have increased the need for more accurate models and also their complexity. Circuit simulation programs have become rather commonplace and generally available for use on personal computers. The usefulness of such computer-aided design programs is directly influenced by the accuracy of the transistor models and their adequacy for the design application. In general the models can be categorized as large-signal (nonlinear) models used for dc or transient analysis, and

small-signal (linear) models used for ac or frequency-domain analysis.

Most large-signal models are represented by systems of equations relating currents and charges to terminal voltages. Different equations are typically used for different combinations of terminal voltages or regions of operation.

In many analog circuits, the signals are small enough that the nonlinear models can be replaced by linearized equivalent circuit models. Linear circuits are much less complicated to analyze than nonlinear ones. The hybrid-π configuration can be used for linear modeling of field-effect transistors of bipolar junction transistors. *See* AMPLIFIER; CIRCUIT (ELECTRONICS); ELECTRICAL MODEL.

Michael Artaki; Robert M. Fox

Bibliography. W. H. Hayt, Jr., and G. W. Neudeck, *Electronic Circuit Analysis and Design*, 2d ed., 1984, reprint 1995; Institute of Electrical and Electronics Engineers, *1994 International Electron Devices Meeting Technical Digest*, San Francisco, California, December 11–14 1994; G. Massobrio and P. Antognetti, *Semiconductor Device Modeling with SPICE*, 2d ed., 1998; G. W Neudeck, *The Bipolar Transistor*, 2d ed., 1989; G. W. Neudeck, *The P-N Junction Diode*, 2d ed., 1989.

Transit (astronomy)

The apparent passage of a celestial body across the apparent disk of a larger body, such as a planet across its parent star or of a satellite across its parent planet; also, the apparent passage of a celestial object or reference point across an adopted line of reference in a celestial coordinate system. Classically, the observed data were instants of internal and external tangency of the disks (contacts) at ingress and egress of the smaller body. In the modern era, data may also include the differential brightness of the two disks and the duration of any change of brightness.

Transits of Mercury and Venus. Mercury and Venus are the only planets whose orbits lie between the Earth and the Sun and thus can be seen from Earth to cross the disk of the Sun. The conditions are that the planet is in inferior conjunction at the same time that it passes one of the two nodes of its orbit, thus putting it essentially in a straight line between the Earth and the Sun. The first transit predicted and observed was that of Mercury on November 7, 1631. Venus transited exactly 1 month later. From then through the early twentieth century, transits of Mercury were observed for the purpose of getting precise positions of the planet to improve knowledge of its orbit, and transits of Venus to determine the solar parallax. The inherent errors of the observations, caused primarily by Earth's atmosphere, made analysis so difficult that there was great disagreement among astronomers as to their value.

Before these difficulties were known, Edmond Halley proposed the observation of Venus transits for determination of the solar parallax, one of the foremost objectives of astronomy. [The solar parallax is the angle, p, subtended by the Earth's equatorial radius, r, at the mean distance of the Sun from the Earth, a; $p = r/a$ when p is measured in radians. Measurement of the solar parallax determines the astronomical unit (AU) since 1 AU $= a = r/p$.] Dozens of international expeditions were sent to observe the four transits of Venus in the eighteenth and nineteenth centuries. These strained the bounds of contemporary methods of observation and analysis, and the value of the results was controversial. In hindsight, the value of the solar parallax derived from them is closer to the modern value than that obtained by any other method of the time. *See* ASTRONOMICAL UNIT; PARALLAX (ASTRONOMY).

In a century, there are 13 or 14 transits of Mercury. In the current era, the Earth crosses the line of nodes of Mercury's orbit each year on May 8 or 9 and November 10 or 11. If Mercury is in inferior conjunction within 3 days of the May crossing or 5 days of the November crossing, a transit will occur. November transits occur twice as often as May transits. The last two occurred on November 15, 1999, and May 7, 2003 (see **illus.**). The next one occurs on November 8–9, 2006.

The size, shape, and orientation of the orbit of Venus causes transits to be very rare—only 81 between the years −2000 and +4000. At present, inferior conjunction must occur within 2 days of June 7 or December 9 for a transit to occur. Transits usually occur in pairs separated by 8 years, with 105.5 or 121.5 years between pairs. The last one occurred on December 6, 1882; the next pair occurs on June 8, 2004, and June 6, 2012. *See* MERCURY (PLANET); PLANET; VENUS.

Transits of Jupiter's satellites. Transits of the galilean satellites of Jupiter occur at each of their inferior conjunctions with the exception of satellite IV, which occasionally passes clear of the planet's disk.

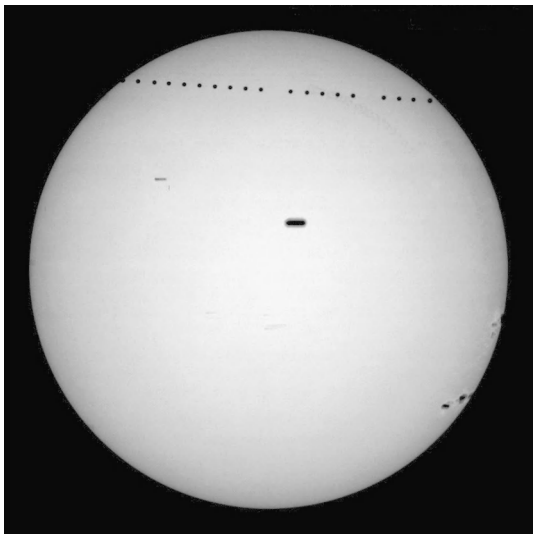

Transit of Mercury, as seen by the GONG instrument at Udaipur, India, on May 7, 2003, Mercury is at the top of the image. The dark spot near the middle is a sunspot, smeared from the compositing process. (*Photo courtesy of Cliff Turner, National Solar Observatory, Global Oscillation Network Group, National Science Foundation*)

They are difficult to observe and are used mainly to estimate the albedo (reflectivity) of the satellites relative to that of Jupiter. As each satellite passes in front of the planet, it casts its shadow on the planet's disk and causes the phenomenon of shadow-transit. *See* JUPITER; SATELLITE (ASTRONOMY).

Meridian transits. Until the close of the twentieth century, passages of stars and other celestial bodies across the local meridian were observed extensively for determining precise coordinates of the stars and planets, accurate time, or the position of the observer. The instrument commonly used is often called a transit circle. This type of observation has been almost completely superseded by interferometric methods from Earth's surface and orbiting satellites, and by other astrometric observations from spacecraft. *See* ASTRONOMICAL COORDINATE SYSTEMS; ASTRONOMICAL TRANSIT INSTRUMENT.

Transits of extrasolar planets. The search for planets around solarlike stars other than the Sun yielded the first positive results in 1995. More than 100 planets have been found since, and the number is rising rapidly. They have been found by techniques using astrometry and radial velocity. The most recently developed technique is to detect photometrically the minute decrease in brightness of a star as an orbiting planet crosses or transits its face. This can occur only if the planet's orbital plane lies nearly edge-on to the Earth. The photometric method can be used for fainter and more distant stars than the other techniques.

As a tool for verification of discovery by other methods, the first planetary companion detected this way was reported in 1999, orbiting the star HD 209458 in Pegasus. Even though that companion was larger than Jupiter, the technique is considered the most mature for detecting Earth-class extrasolar planets, that is, those that are 0.5–10 times the size of Earth. In December 2002 the first planet was initially discovered photometrically, orbiting the star OGLE-TR-56 in the next spiral arm of the Milky Way Galaxy. Within 6 months, two more planets were announced within the same group of candidate stars.

The observed data are the period of recurrence of the transit (frequency), the duration of the transit (expected to be 2–16 hours), and the decrease in brightness of the star, which may be as little as 1 part in 12,000. These values must be consistent for at least three occurrences to be considered confirmation of a single planet in transit. Otherwise there might be more than one, or a physical variability in the star occurring faster than the duration of the transit.

The star's spectral type is used to estimate its mass, diameter, and brightness. The frequency of transit gives the orbital period and, through Kepler's third law, the semimajor axis of the orbit. The fractional change in brightness gives the ratio of cross-sectional areas, or diameters, of the planet and star. From all this, the planet's temperature and density may be deduced. *See* KEPLER'S LAWS; SPECTRAL TYPE; STAR.

From statistical considerations of the orientation of orbits, 1% of solarlike stars with planets would show transits of inner planets. Very high precision photometry with sensitive charge-coupled-device (CCD) detectors and regular monitoring of hundreds of thousands of faint stars, using relatively small telescopes, makes Earth-based searching practical. At least three space missions planned for launch in 2005 or later [Kepler, FRESIP (Frequency of Earth-Sized Inner Planets), and COROT (Convection, Rotation, and Planetary Transits)] are proposed for continuous and simultaneous photometric monitoring of 5000–160,000 stars for up to 8 years. They are expected to find 50–480 or more Earth-class planets, and a smaller number of giant planets in transit. *See* CHARGE-COUPLED DEVICES.

Alan D. Fiala

Bibliography. Better size for the transiting exoplanet, *Sky Telesc.*, 101(1):29, January 2001; S. Dreizler et al., OGLE-TR-3: A possible new transiting planet, *Astron. Astrophys.*, 402:791–799, 2003; Extrasolar planet seen transiting its star, *Sky Telesc.*, 99(2):16–17, February 2000; J. Meeus, *Transits*, Willmann-Bell, Richmond, 1989; A new way to find planets, *Sky Telesc.*, 105(4):9, April 2003; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, University Science Books, Mill Valley, CA, 1992.

Transition elements

In broad definition, the elements of atomic numbers 21–31, 39–49, and 71–81, inclusive. The symbols of these elements, along with their atomic numbers and valence-shell electronic configurations, are given in **Fig. 1**. The elements are arranged in the order in which they appear in the long, or Bohr, form of the periodic table.

A more restricted classification of the transition elements, preferred by many chemists, is indicated by the heavy border drawn about the central portion of the table. All of the elements in this section of the table have one or more electrons present in an unfilled *d* subshell in at least one well-known oxidation state.

Chemical properties. In their compounds, the transition elements tend to exhibit multiple valency, the maximum valence increasing from 3+ at the beginning of a series (Sc, Y, Lu) to 8+ at the fifth member (Mn, Re). For the elements in any vertical column, the highest oxidation state is usually observed in the element at the bottom of the column. Thus the highest oxidation state of iron is 6+, whereas osmium attains an oxidation state of 8+.

One of the most characteristic features of the transition elements is the ease with which most of them form stable complex ions. Features which contribute to this ability are favorably high charge-to-radius ratios and the availability of unfilled *d* orbitals which may be used in bonding. Examples of such complexes include a very stable cyanide complex of aurous gold $Au(CN)_2^-$, of commercial importance for the recovery of the metal from low-grade ores; a similar complex $Ag(CN)_2^-$, useful in obtaining bright, firmly adherent deposits of silver by electroplating; and numerous ammonia complexes, of which the

21 Sc $3d4s^2$	22 Ti $3d^24s^2$	23 V $3d^34s^2$	24 Cr $3d^54s$	25 Mn $3d^54s^2$	26 Fe $3d^64s^2$	27 Co $3d^74s^2$	28 Ni $3d^84s^2$	29 Cu $3d^{10}4s$	30 Zn $3d^{10}4s^2$	31 Ga $3d^{10}4s^24p$
39 Y $4d5s^2$	40 Zr $4d^25s^2$	41 Nb $4d^45s$	42 Mo $4d^55s$	43 Tc $4d^65s$	44 Ru $4d^75s$	45 Rh $4d^85s$	46 Pd $4d^{10}$	47 Ag $4d^{10}5s$	48 Cd $4d^{10}5s^2$	49 In $4d^{10}5s^25p$
71 Lu $5d6s^2$	72 Hf $5d^26s^2$	73 Ta $5d^36s^2$	74 W $5d^46s^2$	75 Re $5d^56s^2$	76 Os $5d^66s^2$	77 Ir $5d^9$	78 Pt $5d^96s$	79 Au $5d^{10}6s$	80 Hg $5d^{10}6s^2$	81 Tl $5d^{10}6s^26p$

Fig. 1. Portion of periodic table showing transition elements, their atomic number, symbol, and electron configurations in valence shells; heavy border indicates the more restricted classification.

deep blue $Cu(NH_3)_4^{2+}$ is a representative example, widely used in colorimetric analyses for copper. Vitamin B_{12} is an example of a cobalt(III) complex that is important in nutrition; and hemin, the red pigment of blood, is an important iron(II) complex.

Most of the ions and compounds of the transition metals are colored, and many of them are paramagnetic; that is, when they are placed in a magnetic field, the magnetic flux within the compound is higher than that of the surrounding field. Both color and paramagnetism are related to the presence of unpaired electrons in the *d* subshell. Excitation of these relatively loosely bound electrons to higher energy states accounts for the absorption of light in the visible region of the spectrum, while the magnetic field associated with the electron spin is responsible for the magnetic behavior of the compounds. Study of the magnetic behavior of compounds and complex ions of the transition elements has contributed much to an understanding of chemical bonding in these elements, since utilization of the *d* electrons in bonding involves electron-pair formation, with consequent cancellation of the magnetic moments and altered magnetic properties. Because of their ability to accept electrons in unoccupied *d* orbitals, transition elements and their compounds frequently exhibit catalytic properties.

Many of the most important catalysts, such as nickel used in hydrogenation, are transition elements.

Broadly speaking, the properties of the transition elements are intermediate between those of the so-called representative elements, in which the subshells are completely occupied by electrons (alkali metals, halogen elements), and those of the inner or *f* transition elements, in which the subshell orbitals play a much less significant role in influencing chemical properties (rare-earth elements, actinide elements). *See* ATOMIC STRUCTURE AND SPECTRA; CATALYSIS; COORDINATION CHEMISTRY; COORDINATION COMPLEXES; MAGNETOCHEMISTRY.

The metals and their uses. All the transition elements are metals and, in general, are characterized by high densities, high melting points, and low vapor pressures. Included, for example, are tungsten and tantalum, melting at 3370 and 3030°C (6098 and 5486°F), respectively. At room temperature, the vapor pressure of tungsten is so low that it compares to only one gaseous atom in a volume of space equal to that of the known sidereal universe. In general, those properties related to strong cohesiveness or binding between the atoms in the metallic state, such as high density, extreme hardness, and high melting point, reach a broad maximum in the neighborhood of the fourth member of each series. Within a given subgroup, these same properties tend to increase with increasing atomic weight. Facility in the formation of metallic bonds is demonstrated also by the existence of a wide variety of alloys between different transition metals. The variation in some representative properties of the transition elements as a function of atomic number is shown in the **Fig. 2**.

The transition elements include most metals of major economic importance, such as the relatively abundant iron and nickel, on one hand, and the rarer coinage metals, copper, silver, and gold, on the other. Also included is the rare and relatively

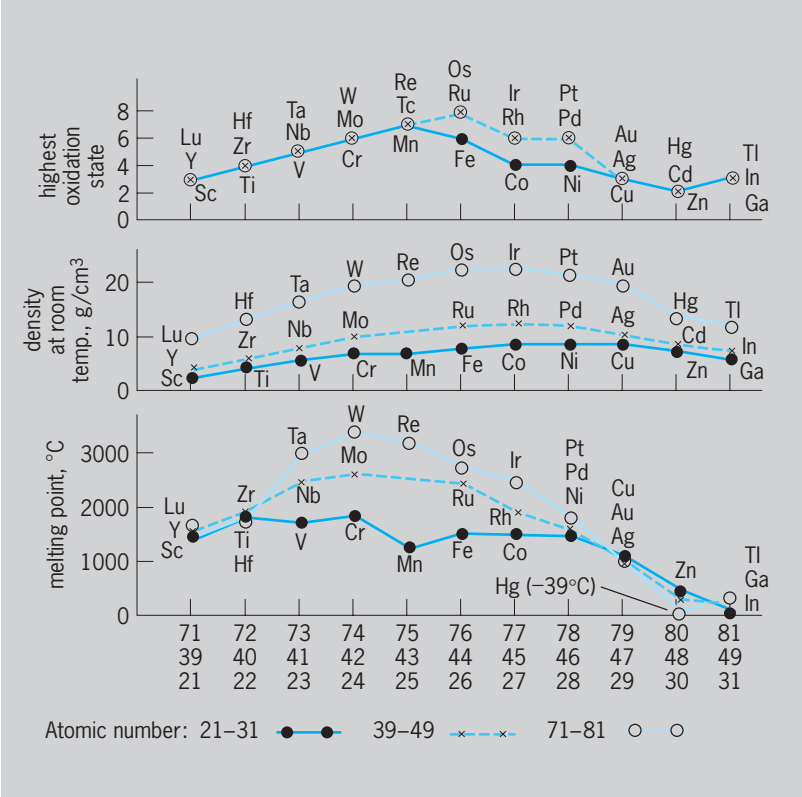

Fig. 2. Physical properties of transition elements as function of their atomic number.

unfamiliar element technetium, which is not found naturally in the terrestrial environment but is available in small amounts as a product of nuclear fission. Burris B. Cunningham

Bibliography. F. Basolo et al. (eds.), *Transition Metal Chemistry*, vol. 1, 1973, vol. 2, 1977; F. Basolo and J. L. Burmeister (eds.), *On Being Well-Coordinated: A Half-Century of Research on Transition Metal Complexes*, 2003; R. H. Crabtree, *The Organometallic Chemistry of the Transition Metals*, 3d ed., 2000; K. H. Whitmire, *The Interface of Main Group and Transition Metal Cluster Chemistry,* 1988.

Transition point

The point at which a substance changes from one state of aggregation to another. This general definition would include the melting point (transition from solid to liquid), boiling point (liquid to gas), or sublimation point (solid to gas); but in practice the term transition point is usually restricted to the transition from one solid phase to another, that is, the temperature (for a fixed pressure, usually 1 atm or 100 kilopascals) at which a substance changes from one crystal structure to another.

Some typical examples of transition points are:

$$\beta\text{-Fe (body-centered cubic)} \xrightarrow[(1664.3^\circ\text{F})]{\text{at 1180 K}} \gamma\text{-Fe (face-centered cubic)}$$

$$S_8 \text{ (rhombic)} \xrightarrow[(203.5^\circ\text{F})]{\text{at 369 K}} S_8 \text{ (monoclinic)}$$

$$CCl_4 \text{ (monoclinic)} \xrightarrow[(-53.8^\circ\text{F})]{225.5\text{ K}} CCl_4 \text{ (tetragonal)}$$

$$NH_4NO_3 \text{ (}\beta\text{-rhombic)} \xrightarrow[(90.2^\circ\text{F})]{\text{at 305.3 K}} NH_4NO_3 \text{ (}\alpha\text{-rhombic)}$$

$$NH_4NO_3 \text{ (}\alpha\text{-rhombic)} \xrightarrow[(183.6^\circ\text{F})]{\text{at 357.4 K}} NH_4NO_3 \text{ (trigonal)}$$

Another kind of transition point is the culmination of a gradual change (for example, the loss of ferromagnetism in iron or nickel) at the lambda point, or Curie point. This behavior is typical of second-order transitions. *See* BOILING POINT; MELTING POINT; PHASE EQUILIBRIUM; SUBLIMATION; TRIPLE POINT. Robert L. Scott

Transition radiation detectors

Detectors of energetic charged particles that make use of radiation emitted as the particle crosses boundaries between regions with different indices of refraction. An energetic charged particle moving through matter momentarily polarizes the material nearby. If the particle crosses a boundary where the index of refraction changes, the change in polarization gives rise to the emission of electromagnetic transition radiation. About one photon is emitted for every 100 boundaries crossed, for transitions between air and matter of ordinary density. Transition radiation is emitted even if the velocity of the particle is less than the light velocity of a given wavelength, in contrast to Cerenkov radiation. Consequently, this radiation can take place in the x-ray region of the spectrum where there is no Cerenkov radiation, because the index of refraction is less than one. *See* CERENKOV RADIATION; REFRACTION OF WAVES.

The radiation extends to frequencies greater than the plasma frequency by the factor γ = particle energy divided by particle mass. The production of x-rays requires γ equal to or greater than 1000. A threshold as high as this is difficult to achieve by other means. This fact has led to the application of this effect for the identification of high-energy particles. For example, electrons of about 10^9 eV will produce x-rays of a few kiloelectronvolts, while the threshold for pions is on the order of 10^{11} eV. The solid material should be of low atomic number, carbon or lighter, to minimize absorption of x-rays, which are emitted close to the particle direction. The material is often in the form of foils, which must be of the order of 0.0004 in. (0.01 mm) thick to avoid destructive interference of the radiation from the two surfaces, and similarly, the foil spacing is typically 0.004 in (0.1 mm). Random assemblies of fibers or foams are almost as effective as periodic arrays of foils. Effective electron detectors have been made with several hundred foils followed by a xenon proportional chamber for x-ray detection. *See* PARTICLE DETECTOR; PLASMA (PHYSICS); RELATIVISTIC MECHANICS. William J. Willis

Bibliography. X. Artru, G. B. Yodh, and G. Menessier, Practical theory of the multilayered transition radiation detector, *Phys. Rev. D,* 12:1289–1306, 1975; J. Cobb et al., Transition radiators for electron identification at the CERN ISR, *Nucl. Instrum. Meth.,* 140:413–427, 1977; G. M. Garibian, Transition radiation effects in particle energy losses, *Sov. Phys.—JETP,* 10:372–376, 1960.

Translucent medium

A medium which transmits rays of light so diffused that objects cannot be seen distinctly; that is, the medium is only partially transparent. Familiar examples are various forms of glass which admit considerable light but impede vision. Inasmuch as the term translucent seems to imply seeing, usage of the term is ordinarily limited to the visible region of the spectrum. M. G. Mellon

Transmission lines

A system of conductors suitable for conducting electric power or signals between two or more termini. Transmission lines take many forms in practice and have application in many disciplines. For example, they traverse the countryside, carrying telephone signals and electric power. The same transmission

lines, with similar functions, may be hidden above false ceilings in urban buildings. With the need to reliably and securely transmit ever larger amounts of data, the required frequency of operation has increased from the high-frequency microwave range to the still higher frequency of light. Optical fibers are installed in data-intensive buildings and form a nationwide network. Increasing demand also requires that transmission lines handle greater values of electric power.

Transmission lines can, in some cases, be analyzed by using a fairly simple model that consists of distributed linear electrical components. Models of this type, with some permutations, can also be used to describe wave propagation in integrated circuits and along nerve fibers in animals. The study of hollow metal waveguides or optical fibers is usually based upon an analysis starting from Maxwell's equations rather than employing transmission-line models. Fundamentals and definitions that can initially be obtained from a circuit model of a transmission line carry over to waveguides, where the analysis is more complicated.

Coaxial cables and strip lines. Two particular types of transmission lines for communication that have received considerable attention are the coaxial cable (**Fig. 1***a*) and the strip line (Fig. 1*b*). The coaxial cable is a flexible transmission line and typically is used to connect two electronic instruments together in the laboratory. In the coaxial cable, a dielectric separates a center conducting wire from a concentric conducting sleeve. The strip line is used in integrated circuits to connect, say, two transistor circuits together. A strip line also has a dielectric that separates the top conducting element from the base, which may be an electrical ground plane in the circuit. If the conductivity of the metal is sufficiently high and the dielectric is lossless, both of which are reasonable approximations in practice, then it is possible to assume that the time-varying electric and magnetic fields that are associated with the propagating wave are polarized in a plane that is transverse to the direction of the wave propagation. This is similar to plane-wave propagation in free space and is called the transverse electric and magnetic (TEM) mode of propagation. In that case, the electric field is polarized in a direction that is between the two conductors, and the magnetic field is in a direction that is perpendicular to the electric field and to the direction of propagation. Losses in the transmission

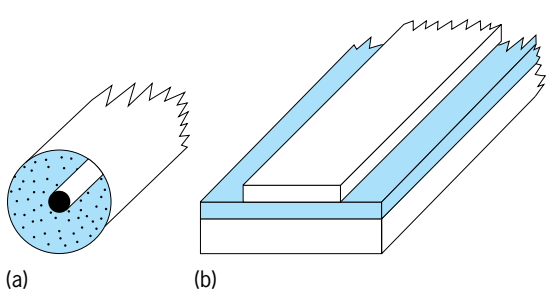

Fig. 1. Two common transmission lines for communication. (*a*) Coaxial cable. (*b*) Strip line.

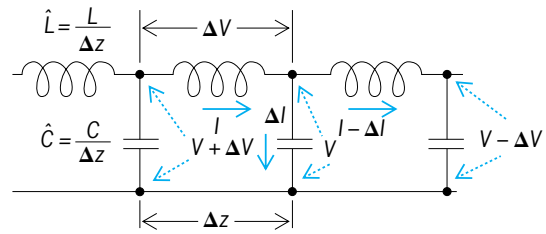

Fig. 2. Equivalent circuit that represents three sections of a lossless transmission line. Each section is of length Δz, and has an inductor (L) and capacitor (C). Voltages (V) and currents (I), and their differences (ΔV and ΔI) between successive sections are indicated.

line cause the signal to attenuate and become distorted as it propagates down the transmission line. *See* COAXIAL CABLE; INTEGRATED CIRCUITS.

Circuit model. These transmission lines can be most easily analyzed in terms of electrical circuit elements consisting of distributed linear inductors and capacitors. The values of these elements are in terms of the physical dimensions of the coaxial cable and the strip line, and the permittivity of the dielectric. Each of the elements is interpreted to be measured in terms of a unit length of the element. An equivalent circuit (**Fig. 2**) represents either the coaxial cable or the strip line as well as other transmission lines such as two parallel wires. *See* ALTERNATING-CURRENT CIRCUIT THEORY; CAPACITANCE; CAPACITOR; INDUCTANCE; INDUCTOR; PERMITTIVITY.

Losses are incorporated into the transmission-line model with the addition of a distributed resistance in series with the inductor and a distributed conductance in parallel with the capacitor. Additional distributed circuit elements can be incorporated into the model in order to describe additional effects. For example, the linear capacitors could be replaced with reverse-biased varactor diodes and the propagation of nonlinear solitons could be studied. *See* CONDUCTANCE; ELECTRICAL RESISTANCE; SOLITON; VARACTOR.

Telegraphist's equations. The circuit model (Fig. 2) can be analyzed by using basic laws from circuit theory in order to derive the wave equation for waves that propagate along the transmission line. The voltage difference ΔV between two adjacent nodes equals the voltage across the inductor. A current ΔI leaves a node and passes through the adjacent capacitor. Equations (1) and (2) describe these voltages

$$\Delta V(z,t) = \hat{L}\Delta z \frac{\partial I(z,t)}{\partial t} \Rightarrow \frac{\partial V(z,t)}{\partial z} = \hat{L}\frac{\partial I(z,t)}{\partial t} \tag{1}$$

$$\Delta I(z,t) = \hat{C}\Delta z \frac{\partial V(z,t)}{\partial t} \Rightarrow \frac{\partial I(z,t)}{\partial z} = \hat{C}\frac{\partial V(z,t)}{\partial t} \tag{2}$$

and currents, where Δz is the length of a section of the line, and $\hat{L}$ and $\hat{C}$ are the inductance and capacitance per unit length. In writing the partial differential equations that appear on the right of Eqs. (1) and (2), the limit of $\Delta z \to \partial z$ has been taken. The variables $V(z,t)$ and $I(z,t)$ depend on both

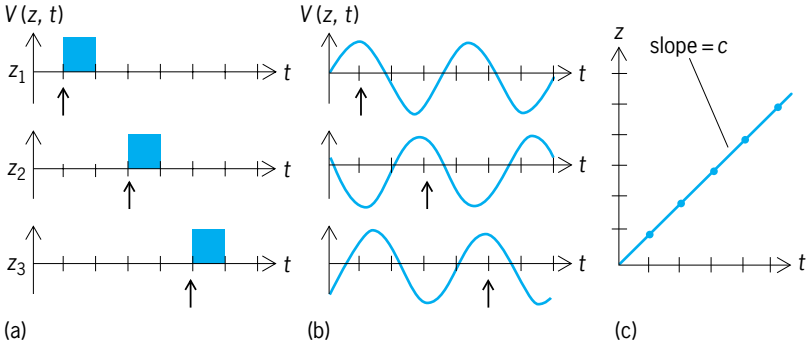

Fig. 3. Signal detection on an oscilloscope. (*a*) Oscilloscope pictures of voltage pulses at three locations, z_1, z_2, and z_3, on a transmission line, successively more distant from the signal generator. (*b*) Oscilloscope pictures of a periodic voltage sine wave at the same three locations on the transmission line. (*c*) Trajectory of propagating wave. The detecting probe position, z_j, is plotted as a function of the time of flight, t_j (indicated with arrows in parts *a* and *b*), of the wave. The slope of the line equals the velocity of propagation, *c*.

of the independent variables, position z and time t. This set of two linear coupled first-order partial differential equations is frequently called the telegraphist's equations.

Wave equation. Eliminating one of the dependent variables, say, the current $I(z,t)$, between Eqs. (1) and (2) results in a second-order partial differential equation (3), where $c = 1/\sqrt{\hat{L}\hat{C}}$. Equation (3) is a

$$\frac{\partial^2 V(z,t)}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 V(z,t)}{\partial t^2} = 0 \qquad (3)$$

standard form of the wave equation, and the parameter c is the velocity of propagation of the wave. This velocity equals the velocity of light in the dielectric that separates the two conductors. The most general solution for Eq. (3) is given by Eq. (4), where

$$V(z,t) = V_1(z - ct) + V_2(z + ct) \qquad (4)$$

$V_1(z - ct)$ is a wave that propagates to increasing values of the coordinate z, and $V_2(z + ct)$ is a wave that propagates to decreasing values of the coordinate z. This solution is predicated on the fact that time is increasing and that the functions V_1 and V_2 cannot be determined and sketched unless a numerical value appears within the parentheses. The actual shape of the propagating voltage wave which will be the same as the propagating current wave is determined by the excitation signal and the particular value of the number that appears within the parentheses. The signal could have the shape, for example, of a narrow pulse, of a step, or of a periodic sine wave. *See* WAVE EQUATION; WAVE MOTION.

Signal propagation. In a typical experimental setup, a signal generator is connected to one end of a transmission line. A trigger signal is available from the signal generator, and it is used to trigger an oscilloscope. The trigger signal causes the trace on the oscilloscope screen to have a known temporal relationship with the excited repetitive pulses from the signal generator or the phase of a periodic sine wave from the generator. The propagating signal is detected at various locations along the transmission line. Because of the finite velocity of propagation, c, there will be a nonzero time delay for the signal to propagate from one node to the next. *See* OSCILLOSCOPE.

If the signal generator excites a pulse, then a pulse will propagate on the transmission line (**Fig. 3***a*). As the distance from the signal generator increases, the response on the oscilloscope will appear later in time. If the signal generator excites a periodic sine wave signal, then the propagating wave will be a sine wave (Fig. 3*b*). From such experimental data, it is possible to obtain the velocity of propagation of the wave. This can be done by following the pulse or a point of constant phase of the periodic sine wave. By knowing the position and the time of flight, the trajectory of the wave's propagation can be given (Fig. 3*c*). The slope of the line is the velocity of propagation. In the case of the sinusoidal wave excitation, this velocity is called the phase velocity. *See* PHASE VELOCITY; WAVEFORM.

Equivalent signals for the current wave exist since the two components of the wave are coupled together via Eqs. (1) and (2). Both components of the wave have the same form; that is, a voltage pulse yields a current pulse, a voltage sine wave yields a current sine wave, and so forth. Once the voltage is determined, the current is also known. It is possible to launch a wave so that it will propagate in only one direction or in both directions.

Characteristic impedance. The ratio of the voltage $V(z,t)$ at a location z divided by the current $I(z,t)$ at the same location of the propagating wave is called the characteristic impedance of the transmission line. This parameter, $Z_c = \sqrt{\hat{L}\hat{C}}$, and it specifies several properties of the line. If the losses in the transmission line can be neglected, the characteristic impedance is a real number that corresponds to a resistance. The characteristic impedance of a transmission line depends upon the dielectric constant of the material inserted between the two conductors and the physical dimensions of the trasmission line. The transmission line is specified by its characteristic impedance, and the detailed equivalent circuit elements need not further be employed. A typical value for the characteristic impedance of a coaxial cable is $Z_c = 50\ \Omega$ and of two parallel wires, termed a twin lead, is $Z_c = 4$ or $8\ \Omega$. *See* ELECTRICAL IMPEDANCE.

Termination. A transmission line is terminated at one end with either a load impedance or another transmission line that may have a different characteristic impedance (**Fig. 4**). It is convenient to assume

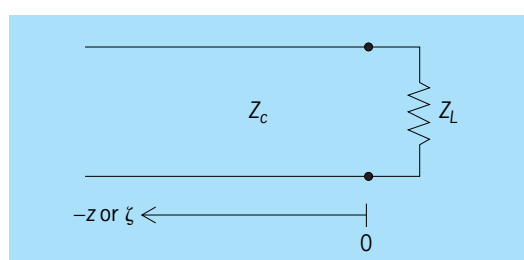

Fig. 4. Transmission line with characteristic impedance Z_c, terminated in a load impedance Z_L. A change of coordinate system (from z to ζ) is indicated.

that the excitation signal is located at $z = -\infty$, to define the location of the load impedance to be at $z = 0$, and to specify a time-harmonic wave excitation of the form $e^{+j\omega t}$. It is also common to redefine the coordinate system at this stage and assume that the coordinate $\zeta = -z$ increases to the left (Fig. 4). The voltage and current can be obtained from Eq. (1) as Eqs. (5) and (6), where $k = 2\pi/\lambda$ and λ are

$$V(\zeta,t) = Ae^{j(\omega t+k\zeta)} + Be^{j(\omega t-k\zeta)} \tag{5}$$

$$I(\zeta,t) = \frac{1}{Z_c}\left\{Ae^{j(\omega t+k\zeta)} - Be^{j(\omega t-k\zeta)}\right\} \tag{6}$$

the wave number and wavelength, respectively, A is the amplitude of the wave that propagates to increasing values of z (decreasing values of ζ), and B is the amplitude of the wave that propagates to decreasing values of z (increasing values of ζ). In this example, A would correspond to the amplitude of the wave launched from the signal generator and B would be the amplitude of the wave that is reflected by the impedance Z_L. The ratio of the voltage wave divided by the current wave, evaluated at $z = 0$, equals the load impedance Z_L. The choice of placing the load impedance at $z = \zeta = 0$ causes the exponential terms in Eqs. (5) and (6) to cancel, yielding Eq. (7).

$$Z_L = Z_c\frac{1+\dfrac{B}{A}}{1-\dfrac{B}{A}} \tag{7}$$

or

$$\mathscr{R} = \frac{B}{A} = \frac{Z_L - Z_c}{Z_L + Z_c}$$

The symbol $\mathscr{R}$ is called the reflection coefficient, and it can have values between -1 and $+1$. If $Z_L = Z_c$, then $\mathscr{R} = 0$; in that case, the transmission line is said to be matched and the reflected component is equal to zero. If $Z_L = 0$ (short circuit), then $\mathscr{R} = -1$; or if $Z_L = \infty$ (open circuit), then $\mathscr{R} = +1$. These values imply that a reflected pulse would have the opposite or the same polarity as the incident pulse. *See* REFLECTION AND TRANSMISSION COEFFICIENTS.

Junctions. If two transmission lines with different characteristic impedances are connected together, and if a voltage signal is incident upon the junction from the first line, a portion of this signal will be transmitted into the second line and a portion will be reflected back into the first line at the junction. The reflection coefficient is found from Eq. (7), where the load impedance is interpreted to be the characteristic impedance of the second line.

Standing waves. The sum of the incident and the reflected waves for the time-harmonic voltage can be written as Eq. (8), where the time dependence has

$$V(\zeta,t) = Ae^{j\omega t}[e^{jk\zeta} + \mathscr{R}\,e^{-jk\zeta}] \tag{8}$$

been separated from the other two terms. The sum of the two components within the square braces is called a standing wave. As noted above, if $Z_L = 0$ (short circuit), $\mathscr{R} = -1$, or if $Z_L = \infty$ (open circuit), $\mathscr{R} = +1$; in these two cases the voltage standing wave assumes a sinusoidal or cosinusoidal spatial variation. In general, the amplitude of the standing wave oscillates in time but is confined within an envelope (**Fig. 5**). The ratio of the maximum voltage to the minimum voltage of the standing wave, called the voltage standing-wave ratio (VSWR), is an easily measured parameter, given by Eq. (9).

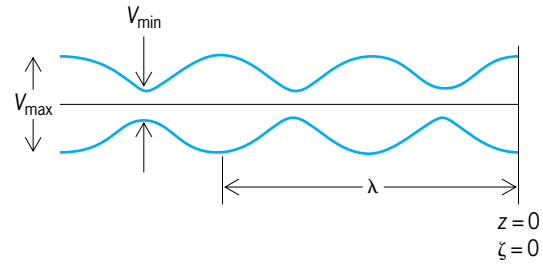

Fig. 5. Envelope of the voltage standing wave consisting of a wave propagating from $z = -\infty$ and being reflected at $z = 0$. The wavelength, λ, and the maximum and minimum voltages, V_{max} and V_{min}, are indicated.

$$\text{VSWR} = \frac{V_{max}}{V_{min}} = \frac{1+|\mathscr{R}|}{1-|\mathscr{R}|} \tag{9}$$

In practice, it is desirable to have the VSWR be as close to unity as possible. This implies that $\mathscr{R} \approx 0$ and the load impedance should be matched to the characteristic impedance of the transmission line, that is, $Z_L \approx Z_c$. Equation (9) can be solved for the reflection coefficient $|\mathscr{R}|$ in terms of the VSWR. Laboratory measurements of a standing voltage wave at several locations in space produce data from which the VSWR, hence from Eq. (9), $|\mathscr{R}|$ can be determined. The wavelength can be measured, and from the known value of the frequency, the phase velocity can be computed. The value of the wavelength and the location of the first minima yield the value of the normalized load impedance (Z_L/Z_c) by using Eq. (7). This allows the value of a complex load impedance to be computed from Eq. (7) since Z_c is known in practice. *See* WAVELENGTH MEASUREMENT.

Impedance matching. There are several techniques that can be employed to match transmission lines. These include the addition of circuit elements or sections of shorted transmission lines at critical locations in the main transmission line. Hence, the terminating end which consists of the actual load impedance and the additional elements will be matched to a microwave transmission line over a significant length of the line. The Smith chart is frequently used to provide graphical visualization of matching techniques. *See* IMPEDANCE MATCHING.

Fault location. Monitoring voltage pulse propagation, both the incident and the reflected pulses, can yield similar results. In addition, the measurement of the time of flight of a signal as it propagates along the transmission line and is reflected from a discontinuity can be used to find the location of a fault in a transmission line if the velocity of propagation is previously known. The technique is called time-domain reflectometry. This is particularly useful for finding faults in buried cables.

Attenuation. If the losses due to the distributed series resistance or shunt conductance terms are included in the model, it is found that the voltage and current signals attenuate as they propagate. This attenuation is exponential. In addition, the characteristic impedance becomes complex. *See* ATTENUATION (ELECTRICITY).

Waveguides. Up to this point, transmission lines have been described that could be modeled with distributed circuit elements. The derivation of the telegraphist's equations, (1) and (2), and the resulting wave equation, (3), directly followed. There are, however, several cases where it is better to directly start from Maxwell's equations in order to ascertain the characteristics of the electromagnetic waves that propagate in a certain direction. These waves are constrained or guided to propagate in a certain direction in hollow metal tubes called waveguides or a long dielectric slabs or optical fibers called dielectric waveguides. Transmission-line models for these structures do exist; hence all of the material that has been described already is applicable here.

The electromagnetic fields are determined from Maxwell's equations. In a vacuum, the wave equation (10) can be derived from these equations, where

$$\nabla^2\mathbf{E} - \frac{1}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} = 0 \tag{10}$$

$\mathbf{E}$ is the electric field, c is the speed of light, and ∇^2 is the laplacian operator. If the wave propagates in a waveguide (**Fig. 6**), in the direction of the z axis of a coordinate system, as given in Eq. (11), then Eq. (10) can be rewritten as Eq. (12). This is a vector wave

$$\mathbf{E}(x,y,z,t) = \mathbf{E}(x,y)e^{j(\omega t-\beta z)} \tag{11}$$

$$\nabla_t^2\mathbf{E} + \left(\frac{\omega^2}{c^2} - \beta^2\right)\mathbf{E} = 0 \tag{12}$$

equation in that the polarization of the electric field $\mathbf{E}$ is at this stage arbitrary. The notation ∇_t^2 indicates differentiation with respect to coordinates that are in the transverse plane, perpendicular to the direction of propagation. The wave equation (10) is similar to Eq. (3), and the description of waves with reference to transmission lines is also applicable here. *See* CALCULUS OF VECTORS; LAPLACIAN; MAXWELL'S EQUATIONS.

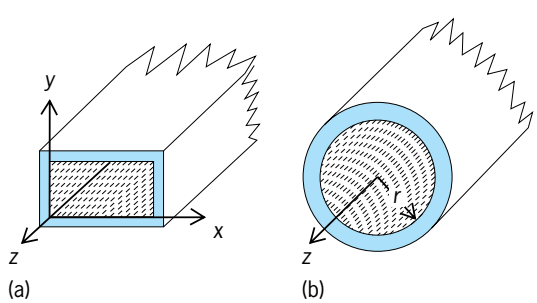

Fig. 6. Two waveguides, with conventional coordinates. (*a*) Rectangular waveguide. (*b*) Cylindrical waveguide.

Modes of propagation. Associated with the solution of Eq. (12), are boundary conditions that must be applied. Because of the high conductivity of the walls, the applicable boundary condition is that the component of the electric field that is tangent to the wall at the wall must be equal to zero. If the electric field in the transverse plane for a rectangular waveguide (Fig. 6*a*) is polarized in the y direction and is independent of the y coordinate, the solution of Eq. (12) is one in which the field depends sinusoidally on the x coordinate. Moreover, the field has an integral number of half-periods across the waveguide and has nodes at both walls. In this particular case, there will be no component of electric field in the direction of propagation, and this field configuration is called a transverse electric (TE) mode. Two integers m and n are further used to specify the mode of propagation within the waveguide. (In the case just discussed, $m = 0$, and n is the number of half-periods of the field across the waveguide.) The lowest-order mode is the $\mathrm{TE}_{nm} = \mathrm{TE}_{10}$ mode, with $n = 1$ and $m = 0$. Another family of transverse magnetic (TM) modes also exists. For a cylindrical waveguide (Fig. 6*b*), the radial distribution of the field will be in terms of Bessel functions. *See* BESSEL FUNCTIONS.

Dispersion curve. If the solution discussed above is substituted into Eq. (12), an algebraic relation, Eq. (13) is obtained that must be satisfied, where

$$\beta^2 = \frac{\omega^2}{c^2} - \left(\frac{n\pi}{a}\right)^2 \tag{13}$$

n is the number of half-periods of the field across the waveguide and a is the broad dimension of the waveguide. In order that the wave be able to propagate, the propagation constant β must be a real number. This implies that the frequency, $f = \omega/(2\pi)$, must be higher than a certain value $f_c = (nc/2a)$. This frequency, f_c, is called the cutoff frequency. When the frequency is reduced to this value, the wavelength of the wave equals $2a$ and the wave just bounces back and forth in the transverse direction with no energy propagating down the waveguide. There is a nonlinear relation between the propagation constant, β, and the angular frequency, ω. This relationship between β and ω is called a dispersion curve (**Fig. 7**). The propagation constant β asymptotically approaches a straight line as the angular frequency, ω, increases; the slope of the line equals the velocity of light, c.

Phase and group velocities. If two waves simultaneously propagate in the same waveguide and their frequencies are slightly above the cutoff frequency, there will be constructive and destructive interference between them. In the case that the excited signal is a pulse, the signal may become distorted as it propagates since individual frequency components of the pulse propagate with different velocities. This is called dispersion. A point of constant phase of an individual sine wave will propagate with the phase velocity, and the envelope of the wave will propagate with a different velocity that is called the group velocity. *See* DISTORTION (ELECTRONIC CIRCUITS).

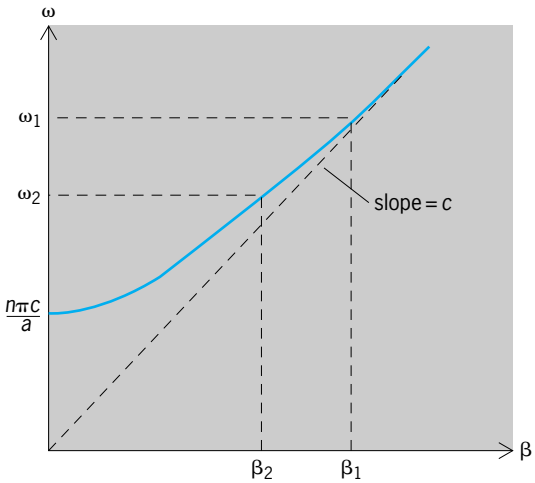

Fig. 7. Dispersion curve for a rectangular waveguide of width *a*, showing relationship between the propagation constant, β, and the angular frequency, ω. Here, *n* specifies the mode of propagation, and *c* is the speed of light.

The derivation of the group velocity can be carried out in a straightforward manner by considering two waves with slightly different frequencies, $\omega_1 \approx \omega_0 + \Delta\omega$ and $\omega_2 \approx \omega_0 - \Delta\omega$ (Fig. 7), propagating in the waveguide, where the intermediate frequency ω_0 is slightly above the cutoff frequency of a particular mode in the waveguide. There will be a unique propagation constant $\beta_1 = \beta_0 + \Delta\beta$ and $\beta_2 = \beta_0 - \Delta\beta$ associated with each frequency. The total electric field in the waveguide is just the linear superposition of the two electric field components in the waveguide. An individual wave will propagate with the phase velocity $v_\phi = \omega/\beta$. The amplitude modulation will propagate with a velocity that is called the group velocity $v_g = \Delta\omega/\Delta\beta$, which becomes $v_g = \partial\omega/\partial\beta$ in the limit that the frequency difference, $\Delta\omega$, approaches zero. In nondispersive media such as a vacuum, the two velocities are the same. They are, however, dramatically different in a waveguide. *See* GROUP VELOCITY.

Attenuation reduction. There are several techniques to match and reduce attenuation in hollow metal waveguides. Screws are frequently inserted in the center of a waveguide, or thin metal strips are placed across it. Losses are reduced by plating the inside of the waveguide with silver or gold, whose conductivity may be higher than the host waveguide.

Dielectric waveguides. Electromagnetic waves can also be guided at an interface between two dielectrics. The amplitude of the fields will exponentially decay to zero in the region external to the dielectric. The dispersion relation for this structure is determined by requiring that the tangential components of the electric and the magnetic fields be continuous at the interface. In certain materials, the inhomogeneous electric field may locally modify the dielectric and permit nonlinear waves called solitons to propagate in the dielectric. *See* ELECTROMAGNETIC WAVE TRANSMISSION; OPTICAL FIBERS; WAVEGUIDE.

Karl E. Lonngren

Power transmission lines. Electric power generating stations and distribution substations are connected by a network of power transmission lines, mostly overhead lines. Power transmitted is generally in the form of three-phase alternating current (ac) at 60 or 50 Hz. In a few instances, where a clear technical or economic advantage exists, direct-current (dc) systems may be used. As the distances over which the power must be transmitted become great and as the amount of power transmitted increases, the power lost in the transmission lines becomes an important component of the production cost of electricity, and it becomes advantageous to increase the transmission voltage. This basic consideration has led to electric power networks which use higher voltages for long-distance bulk power transfers, with several layers of underlying regional networks at progressively lower voltages which extend over shorter distances. The most common transmission voltages in use are 765, 500, 400, 220 kV, and so forth. Voltages below 69 kV are termed subtransmission or distribution voltages, and at these and lower voltages the networks may have fewer alternative supply paths (loops) or may be entirely radial in structure. *See* ALTERNATING CURRENT; DIRECT CURRENT; DIRECT-CURRENT TRANSMISSION; ELECTRIC DISTRIBUTION SYSTEMS.

Performance during normal conditions. The three phases of power transmission lines generally carry balanced voltages and currents; that is, the voltages (and currents) of the three phases are of equal magnitude and differ in phase angles by 120°. This balance is achieved by balancing the loads among the three phases at the distribution level, and by making sure that the impedances of the transmission lines, generators, and transformers are equal in all three phases. In the case of overhead transmission lines, phase impedance balancing can be achieved by transposing the phases, so that the relative position of each of the phases on the tower between the two ends of the transmission line averages out to be identical. Transposition of phases requires special towers which can accommodate the additional mechanical stresses, and consequently is often omitted at the highest voltage levels in the interest of economy.

Under balanced conditions, it is sufficient to analyze a single-phase circuit and assume that the performance of the other two phases is identical. For the purposes of engineering analysis of a power system in normal steady state, the equivalent circuit of each phase of a power transmission line is a π circuit (**Fig. 8**). The resistance and reactance (at the operational frequency), and the capacitance of the entire transmission line can be estimated from corresponding values per mile and then multiplying by the length of the line. This is the nominal-π representation, and is sufficiently accurate for lines of up to 100 mi (160 km) in length. For longer lines, an exact-π model, which uses the exact transmission-line equations in its derivation, is needed. For lines of several hundred miles, the difference between the nominal and exact π circuits is of the order of 5–10%. The resistance is quite small compared to the reactance, and for short lines the capacitance may be omitted. Transmission-line equivalent circuits

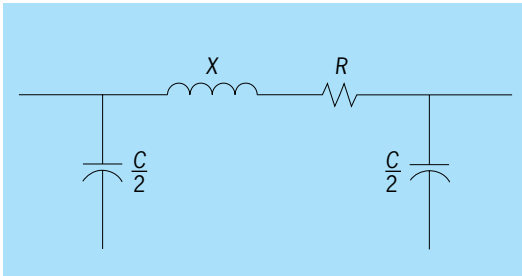

Fig. 8. Equivalent circuit of a power transmission line. The single-phase circuit is used to represent balanced three-phase lines. Here, *R* and *X* are the resistance and reactance (at the operational frequency), and *C* is the capacitance of the entire transmission line.

(Fig. 8) are used in calculation of power flows, voltage drops, power losses in transmission networks, electromechanical oscillations of connected generators and loads, and so forth. *See* REACTANCE.

The power-handling capability (transmission capacity) of a transmission line depends on many factors. The thermal capacity of the transmission-line conductors places a fundamental limit on power transfer capability. In many transmission lines, the power transfer capability is determined by a different phenomenon: transient stability. A power system must recover its normal operation after a fault is cleared by the protection system, and new power flow patterns are established in the remaining network. The synchronous generators supplying the network may go unstable if the transmission line which suffered the fault was carrying a large amount of power, or if the fault persisted for too long a duration. If the power carried by the transmission line is below a certain limit, the fault and its clearing would not lead to instability. The power transfer limit determined in this manner is known as the transient stability limit. Making reasonable assumptions about a power system, the transient stability limit of transmission lines can be shown to be a fairly well defined parameter.

Faults and protection systems. Faults on transmission lines are caused by short circuits between the phases, or between phases and the Earth. Faults may be caused by overvoltages caused by lightning, by the transients set up during swiching operations, or by equipment failure. The faulted line must be removed from service promptly, so that the remaining system can resume service normally. The fault removal is accomplished by the protection system installed at the two ends of each transmission line. It consists of circuit breakers, relays, and current and voltage transformers (transducers). Upon the occurrence of a fault, the relays, acting upon the changed conditions reflected in the outputs of the current and voltage transformers, make a determination if the fault is in the zone of protection for which the protection system is responsible. If the fault is within the zone, the relays will trip the circuit breakers, which will isolate the transmission line at its terminals from the rest of the network. At the highest operating voltages, the entire protection action may be completed in about 30–50 milliseconds after the occurrence of the fault. In most cases, the faults on the transmission lines are temporary, and the protection system will automatically attempt to reenergize the transmission line in order to test if the fault has disappeared. About 80% of the faults lead to successful reclosing, and are never perceived by the end user of electricity. *See* CIRCUIT BREAKER; ELECTRIC PROTECTIVE DEVICES; INSTRUMENT TRANSFORMER; RELAY.

Compensation. There are three types of compensation systems currently used in ac transmission systems: series capacitors, shunt reactors, and shunt capacitors. The power transfer capability limit imposed by transient stability can be relaxed if the series reactance X of the transmission line can be reduced. The obvious method for reducing the reactance is to put a capacitor in series with the transmission line. If the reactance of the capacitor (at the operating frequency) is X_c, the net reactance of the transmission line is reduced from X to $(X - X_c)$. The improvement in the power transfer capability of a transmission line is (approximately) inversely proportional to the reactance of the line. The capacitors used in this fashion are known as series compensation, and may be of a fixed size, or be switched in sections as needed, or could be made variable by using thyristors as control elements. *See* SEMICONDUCTOR RECTIFIER; STATIC VAR COMPENSATOR.

On long transmission lines, the capacitors in the equivalent circuit of the π-section representation lead to a voltage rise at the receiving end if the load on the transmission lines is dropped, or is reduced, for example at night. The voltage rise at light loads or no loads presents a serious operational problem to the connected equipment. This voltage rise can be reduced by connecting shunt reactors across the transmission line; these compensate the shunt capacitors of the π circuit. And finally, under heavy loading conditions, the drop in voltgage at the receiving end due to the line impedance may also be unacceptable. This can be corrected by using shunt capacitors across the transmission lines. *See* REACTOR (ELECTRICITY).

Environment. Power transmission lines are the most visible components of the electric power infrastructure. They impact the environment in several ways. The most obvious is the visual impact. Designs have been introduced to minimize the visual impact of transmission towers, and they are sited to avoid high ground or prominent ridges.

Transmission-line conductors have high electric field gradients at their surfaces, and under certain conditions will ionize and break down the surrounding air, causing what is generally referred to as corona discharge. Corona is a function of moisture and pollution in the air, as well as the smoothness of the conductor surface. It contributes to radio and television interference, and should be minimal in a well-designed and -maintained line. (Corona is often visible at night as a glow surrounding the power-line conductors.) *See* CORONA DISCHARGE.

Power-line currents may interfere with nearby telephone circuits, especially if there is

unbalance in the currents in the three phases. If significant ground currents exist under normal (or fault) conditions, they may flow through buried metallic objects in the ground, such as control cables and gas pipe lines. If ground currents persist, they may lead to galvanic corrosion of the metal structures. *See* COMMUNICATIONS SYSTEMS PROTECTION; CORROSION; INDUCTIVE COORDINATION.

Electric and magnetic fields exist in the neighborhood of power transmission lines. Although there has been considerable discussion of effects of power-frequency magnetic fields on human and animal health, no significant evidence for these effects has been found after a number of careful studies. *See* ELECTRIC POWER SYSTEMS; ELECTRIC POWER TRANSMISSION.

Arun G. Phadke

Bibliography. D. K. Cheng, *Field and Wave Electromagnetics*, 2d ed., 1989; R. E. Collin, *Foundations for Microwave Engineering*, 2d ed., 2000; A. Hirose and K. E. Lonngren, *Introduction to Wave Phenomena*, 1991, reprint 2003; S. Liao, *Microwave Devices and Circuits*, 3d ed., 1996; R. L. Liboff and G. C. Dalman, *Transmission Lines, Waveguides, and Smith Charts*, 1985; W. D. Stevenson, Jr., *Elements of Power System Analysis*, 4th ed., 1982.

Transmutation

The nuclear change of one element into another, either naturally, in radioactive elements, or artificially, by bombarding an element with electrons, deuterons, or alpha particles in particle accelerators or with neutrons in atomic piles.

Natural transmutation was first explained by Marie Curie about 1900 as the result of the decay of radioactive elements into others of lower atomic weight. Ernest Rutherford produced the first artificial transmutation (nitrogen into oxygen and hydrogen) in 1919. Artificial transmutation is the method of origin of the heavier, artificial transuranium elements, and also of hundreds of radioactive isotopes of most of the chemical elements in the periodic table. Practically all of these elements also have been artificially transmuted into neighboring elements under experimental conditions. *See* NUCLEAR REACTION; PERIODIC TABLE.

Frank H. Rockett

Transonic flight

In aerodynamics, flight of a vehicle at speeds near the speed of sound. When the velocity of an airplane approaches the speed of sound, roughly 660 mi/h (1060 km/h) at 35,000 ft (11 km) altitude, the flight characteristics become radically different from those at subsonic speeds. The drag increases greatly, the lift at a given attitude decreases, the moments acting on the airplane change abruptly, and the vehicle may shake or buffet. Such phenomena usually persist for flight velocities somewhat above the speed of sound. These flight characteristics, as well as the speeds at which they occur, are usually referred to as transonic. For configurations designed for subsonic flight these changes may occur at velocities of 70–110% of the speed of sound (Mach numbers of 0.7–1.1); for airplanes intended for transonic or supersonic flight they may be present only at Mach numbers of 0.95–1.05. *See* FLIGHT CHARACTERISTICS; MACH NUMBER.

Shock waves. The transonic flight characteristics result from the development of shock waves about the airplane. Because of the accelerations of airflow over the various surfaces, the local velocities become supersonic while the airplane itself is still subsonic. (The flight speed at which such local supersonic flows first occur is called the critical speed.) Shock waves are associated with deceleration of these local supersonic flows to subsonic flight velocities (**Fig. 1**). Such shock waves cause abrupt streamwise increases of pressure on the airplane surfaces. These gradients may cause a reversal and separation of the flow in the boundary layer on the wing surface in roughly the same manner as do similar pressure changes at subcritical speeds. When the wing carries lift, the shock-induced separation is particularly strong on the upper surfaces. *See* AERODYNAMIC FORCE; AERODYNAMIC WAVE DRAG; BOUNDARY-LAYER FLOW; SCHLIEREN PHOTOGRAPHY; SHOCK WAVE.

As for boundary-layer separation at lower speeds, the flow breakdown in this case leads to increases of drag, losses of lift, and changes of aerodynamic moments. The unsteady nature of the separated flow results in an irregular change of the aerodynamic forces acting on the airplane with resultant buffeting and shaking. As the Mach number is increased, the shock waves move aft so that at Mach numbers

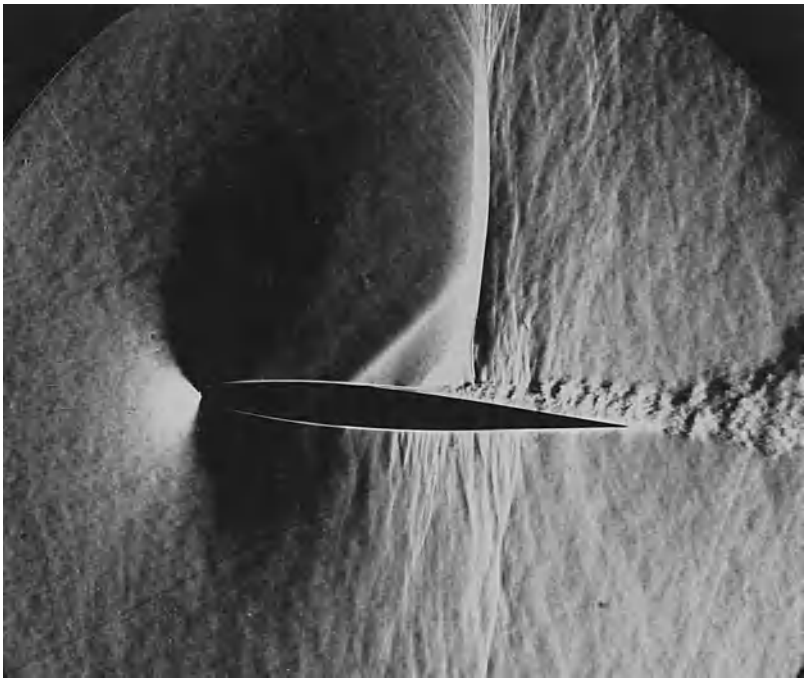

Fig. 1. Schlieren photograph of flow about airfoil section at low transonic speeds. Shock wave is the nearly vertical line near midchord.

of about 1.0 or greater they reach the trailing edges of the surfaces. With the shocks in these positions, the associated pressure gradients have relatively little effect on the boundary layer, and the shock-induced separation is greatly reduced.

Effects on flight characteristics. When the speed is increased to the higher transonic range, at and just above the speed of sound, the energy losses in the shock waves about an airplane may become large. As a result, the drag may increase to many times the subsonic value. At these speeds shock waves, in addition to those present near the aft parts of the surfaces, form ahead of the components. The various waves extend outward, interact, and merge to form two shock waves at a distance from the airplane. These two waves are relatively strong and extensive; they may extend outward for miles from the airplane. *See* SONIC BOOM.

As the speed is increased through the transonic range, the changes of the distribution of load on the wing, resulting first from boundary-layer separation and then from rearward movement of the shock wave, cause a marked rearward shift of the center of lift. This shift causes a nose-downward moment on the airplane that must be corrected by an increase of the negative lift on the usual tail to maintain trim. Also, the effectiveness of the usual flap-type elevator and aileron control surfaces used on subsonic airplanes decreases greatly at transonic speeds. Deflections of such flaps provide differences in the pressures on upper and lower parts of the main surface ahead of flap, as well as on the flap itself. At transonic speeds, the effect of the flap on the pressures on the main surface is greatly reduced because of the presence of local supersonic flows on this surface. In addition, the hinge moments required to deflect the control may be greatly increased at transonic speeds. *See* AILERON; ELEVATOR (AIRCRAFT); FLIGHT CONTROLS.

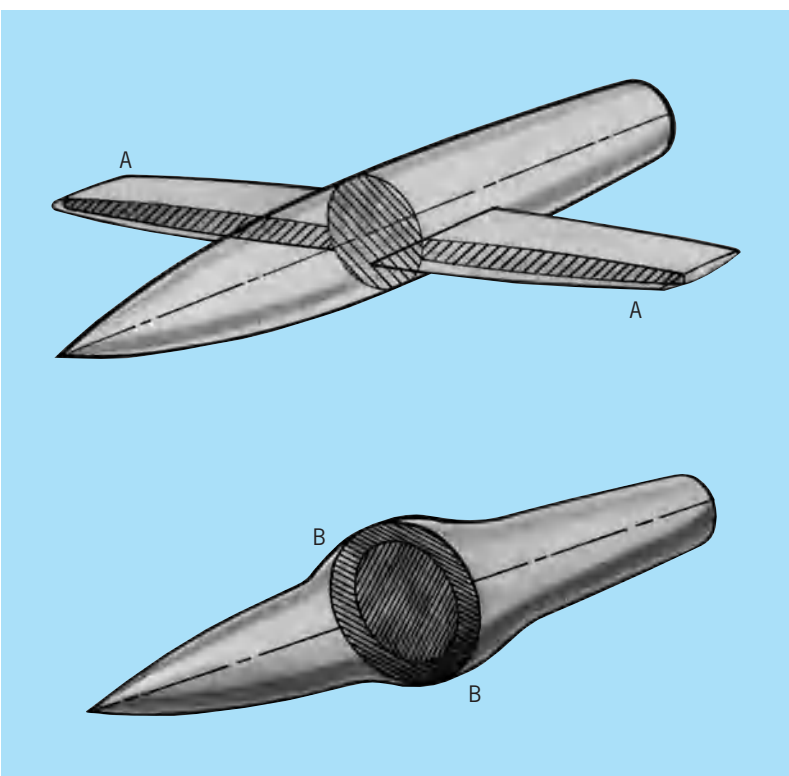

Fig. 2. Area rule comparison. The various normal cross-sectional areas for the body of revolution, such as at BB, are the same as those for the wing-fuselage combination at the corresponding longitudinal station, such as at AA. Therefore, the shock wave and resulting drag near the speed of sound are approximately the same for the two configurations.

Corrective means. Various means may be used to delay and reduce the adverse transonic characteristics.

Sweepback. The most effective means for improving the overall transonic characteristics is to mount the wings slanting backward. The action of such sweep may be understood by considering the airflow over a very long swept surface. Only the component of airflow normal to swept elements of this panel is effective in determining the nature of flow over the surface. Thus on such a swept surface, the onset of a shock wave, with the associated separation, is delayed until the reduced component of local velocity normal to the swept elements becomes supersonic. The use of sweep also greatly reduces the magnitude of the changes in the aerodynamic characteristics, once they occur. *See* WING.

The transonic characteristics are progressively delayed and reduced by increasing the sweep to relatively high values, but excessive sweep leads to a number of aerodynamic problems. Most significantly, a highly swept wing may have an abrupt nose-up moment at the higher lifts. This phenomenon, known as pitch-up, may result in excessive aerodynamic loads of stall. Pitch-up results from an initial separation of the boundary layer on the outboard part of the swept wing with an associated loss of lift for this region. Because this portion of a swept wing is aft of the center of gravity, the loss of lift causes a nose-up moment. Increasing the wing sweep also reduces the lift available for takeoff and landing.

Because of these limitations, most transport-type airplanes designed to fly at transonic speeds incorporate only moderate wing sweep. Usually the obliqueness of the midchord element for such airplanes is about 30°. Transonic and supersonic military airplanes may incorporate as much as 45° of midchord sweep.

According to simple sweep theory, sweepforward is as effective as sweepback in delaying the onset of adverse effects. However, a sweptforward wing would be structurally divergent; that is, a bending of the wing due to air loads would result in increases of the angle of attack of the outboard sections of the wing with resulting increases of the loads on these sections, which, of course, would then result in further bending.

Reduced thickness ratio. Substantial improvements in the adverse transonic characteristics are provided by reducing the thickness-to-chord ratios for the wing and tail surfaces. Such changes reduce the acceleration of the flow over the surfaces with a resulting delay in the onset of local supersonic flows and the associated shock wave. Also, the severity of the adverse longitudinal pressure gradients on the wing

surface is lessened so that boundary-layer separation is reduced. However, reductions of thickness are limited by the considerable increases in the weight of a structurally sound wing. Most high-speed transport wings have mean thickness-to-chord ratios of about 10%, whereas military airplanes may have thickness ratios as low as 3%.

Reduced aspect ratio. Reductions in the aspect ratio provide delays and reductions of the transonic changes similar to those provided by reductions in thickness ratio, although the magnitude of the effect is usually considerably less. More important, lower aspect ratios result in improvements of the wing structural characteristics, which allow the use of smaller thickness-to-chord ratios. However, the use of reduced aspect ratios increases the subcritical drag due to lift. Most high-speed transport wings have aspect ratios of about 7; transonic and supersonic military airplanes may have aspect ratios as low as 2.

Supercritical wing. The onset of the adverse transonic characteristics can be delayed substantially beyond the critical by the use of special speed streamwise section shapes (airfoils) for the wing. These shapes are called supercritical, and wings incorporating them are referred to as supercritical wings. *See* SUPERCRITICAL WING.

Area rule. For a nonlifting condition, the forms of the shock waves and consequently the drag are primarily a function of the longitudinal development of cross-sectional area, in section normal to the airstream, for the complete airplane. According to this relationship, called the area rule, the shock wave and the resulting drag near the speed of sound are approximately the same for two configurations that have the same total normal cross-sectional areas at corresponding longitudinal stations, even though their shapes may be very different (**Fig. 2**).

On the basis of the area rule, the transonic drag increment is reduced by shaping and arranging the airplane components so that area development for the airplane more nearly approaches the shape with the lowest drag (**Fig. 3**). The magnitude of the drag associated with such a shape is greatly reduced by increasing the overall length and reducing the maximum cross-sectional area. However, for practical reasons, the lengths and cross-sectional areas of airplanes must be limited to values corresponding to a body of revolution with a ratio of length to diameter of about 9.

The longitudinal developments of area for conventional subsonic airplanes differ greatly from the ideal shape, and as a result, the maximum transonic drag for such airplanes may be as much as 10 times subsonic drag values. The various wing features discussed above usually result in airplane area developments more nearly approaching the most satisfactory shape (Fig. 3), reducing the transonic drag to approximately three times the subsonic level.

The area developments for some transonic and supersonic airplanes have been made to approach the form for lowest drag by special shaping of the fuselage, reducing the transonic drag to as low as twice

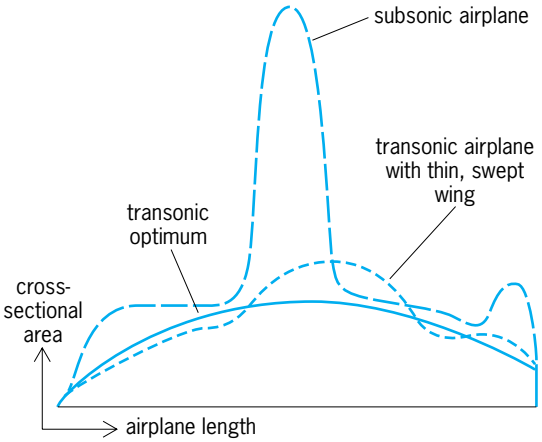

Fig. 3. Longitudinal area developments for various types of airplanes, designed to reduce drag.

the subsonic values. Such a shaping has been provided through the subtraction of fuselage volume in the region of the wing and tail, as well as by the addition of volume ahead of and behind these surfaces. The area development may also be improved by specially locating external bodies such as engine nacelles. *See* AIRPLANE; SUBSONIC FLIGHT; SUPERSONIC FLIGHT.

Richard T. Whitcomb

Bibliography. J. D. Anderson, *Fundamentals of Aerodynamics,* 3d ed., 2001; J. D. Cole and L. P. Cook, *Transonic Aerodynamics,* 1986; T. H. Moulden, *Fundamentals of Transonic Flow,* 1984, reprint 1991; H. J. Ramm, *Fluid Dynamics for the Study of Transonic Flow,* 1990.

Transplantation biology

The science of transferring a graft from one part of the body to another or from one individual, the donor, to another, the recipient. The graft may consist of an organ, tissue, or cells. If donor and recipient are the same individual, the graft is autologous. If donor and recipient are genetically identical (monozygotic), it is syngeneic. If donor and recipient are any other same-species individuals, the graft is allogeneic. If the donor and recipient are of different species, it is called xenogeneic.

Use in experimental biology. The transplantation of cells, tissues, or organs is not only a medical therapy but also an important research tool. For example, a defining criterion of the hormone-secreting, or endocrine, function of a tissue is the demonstration of its ability to reverse the effects of its removal after transplantation to an anatomically abnormal, or heterotopic, site. In embryology, grafting can reveal the interactions of cells of different types that lead to organogenesis. The manner in which genes determine coat coloration has been elucidated by studies involving the grafting of melanocytes (pigment-forming cells) of one genetic constitution into hair follicles of a different one. Primary chimeras, produced by combining cells from different embryos at a very early stage of development and allowing them

to develop in the uteri of surrogate mothers, are useful in the study of problems in embryology, developmental genetics, and immunology. Age chimeras, individuals bearing a functional organ or tissue graft that is chronologically much younger or older than its host, are useful in the study of aging. Studies on such individuals, usually mice, indicate that tissues have a finite life span. In cancer research, transplantation is widely employed to propagate malignant tissues in appropriate hosts for biochemical, immunological, and therapeutic investigations. In its extreme form, "transplantation" of a somatic cell nucleus from an appropriate adult donor to an enucleated recipient oocyte, which is then implanted in a surrogate mother, has been used to generate cloned mice, sheep, and pigs. *See* CHIMERA; MOSAICISM.

Clinical applications. There is remarkable evidence that successful transplantation surgery was performed by ancient Hindu *vaidya* some 2000 years ago. These Ayurvedic physicians reconstructed noses using pedicle flap autografts from the patient's own forehead. Saints Cosmas and Damian were said to have transplanted a human leg in Roman times, but transplantation progressed little until the twentieth century due to the problem of graft rejection. Large-scale use of one type of graft, blood transfusion, was developed during World War II as methods for cross-matching donor and recipient for ABO antigens were perfected. In the second half of the twentieth century, as understanding of the biology of transplantation antigens and the immunology of rejection advanced, the grafting of replacement organs and tissues to cure disease became a clinical reality. In the United States, nearly 90,000 kidney grafts had been performed by the turn of the century. Heart, lung, liver, and pancreas grafts have also become commonplace. Successful transplantation of islets of Langerhans for the treatment of diabetes mellitus has also been achieved and may replace pancreas grafts for this purpose. Hematopoietic stem cell transplantation is used in the treatment of aplastic anemia, leukemia, and immunodeficiency diseases, and its use in the treatment of autoimmune diseases is under investigation. These stem cells can be obtained from bone marrow, peripheral blood, or umbilical cord blood.

In theory, virtually any tissue or organ can be transplanted. The principal technical problems have been defined and, in general, overcome. Remaining major problems concern the safety of methods used to prevent graft rejection and the procurement of adequate numbers of donor organs.

Procurement, storage, and types of grafts. Living volunteers can donate one of a pair of organs, such as a kidney, only one of which is necessary for normal life. Volunteer donors may also be employed for large unpaired organs such as small bowel, liver, or pancreas, segments of which can be removed without impairment of function. Living donors can also provide tissues capable of regeneration; these include blood, bone marrow, and the superficial layers of the skin. In the case of a vital, unpaired organ, such as the heart, the use of cadaver donors is obligatory.

In practice, with the exception of blood and bone marrow, the great majority of transplanted organs are cadaveric in origin, a necessity that presents difficult logistic problems. Except for skin, corneas, and possibly lungs, most organs deteriorate rapidly after death. Clinical transplantation depends on the use of "brain-dead" donors whose vital functions are maintained with the aid of a respirator until organs can be removed. To minimize autolysis, organ grafts must be removed from donors as quickly as possible, generally within 30 min of termination of cardiorespiratory support. Recovered organs must be chilled immediately by immersion in, or perfusion with, an ice-cold physiological salt solution. Under these conditions, large organ grafts have a storage life of about 24–48 h. A few tissues, among them blood, bone marrow, embryos, spermatozoa, skin, and corneas, can be stored for months if frozen to very low temperatures. When impregnated with cryoprotective agents such as glycerol or dimethyl sulfoxide and then maintained at the temperature of liquid nitrogen (-196°C or -321°F), these cryopreserved tissues retain their viability on subsequent rewarming. The prospects for similar long-term preservation of larger organ grafts are still poor.

Certain types of grafts need not be viable, even at the time of transplantation. They may be stored frozen, in a freeze-dried or lyophilized state, or in chemical preservatives for long periods. Lyophilized bone allografts, for example, are used in orthopedic surgery to provide an inert scaffolding within which regeneration of new bone takes place through the activity of cells of the recipient. Such grafts are described as allostatic (as opposed to allovital). Xenogeneic heart valves from pigs that have been chemically treated and contain no living cells are routinely used to replace diseased human valves. *See* PROSTHESIS.

Autografts are used for an increasing number of purposes. Skin autografts are important in the treatment of full-thickness skin loss due to extensive burns or other injuries. Provided that the grafts comprise only the superficial levels of the skin, the donor sites reepithelialize spontaneously within a week or two. Autografts of bone and cartilage are used in orthopedic procedures, and a relatively unimportant blood vessel, nerve, or tendon can be sacrificed to repair or replace a more important one elsewhere in the body. The saphenous vein of the ankle is frequently transplanted to the heart to bypass coronary arteries obstructed by atherosclerosis. Autologous hematopoietic stem cell transplantation is sometimes used to restore blood cells to cancer patients who receive forms of chemotherapy that are lethal to their bone marrow. Islets of Langerhans can be removed from the pancreas of a person undergoing removal of a diseased exocrine pancreas; the recovered autologous islets can then be transplanted to the donor to prevent the diabetic state that would otherwise ensue.

Healing of grafts. In the case of large organs, the main blood vessels must be coupled at operation with those of the host to conserve their viability and

enable them to function. However, thin sheets of skin and small tissues, such as the parathyroid gland and islets of Langerhans, may be transplanted heterotopically with no attempt to preserve or restore the blood supply. Juxtaposed graft and host tissue soon become knit together through the activity of fibroblasts and other cells, and revascularization takes place naturally within 2–3 days as a consequence of penetration of graft tissue by regenerating vessels. This process involves some reutilization of the original vascular channels of the graft. *See* REGENERATIVE BIOLOGY.

Transplantation immunology. The most serious problem restricting the use of allografts is immunological. Because cells in the donor graft express on their surface a number of genetically determined transplantation antigens that are not present in the recipient, allografts provoke a defensive reaction analogous to that evoked by pathogenic microorganisms. As a consequence, after a transient initial period of apparent well-being, graft function progressively deteriorates and the donor tissue is eventually destroyed. The host response, known as allograft rejection, involves a large number of immunological agents, including cytotoxic antibodies and effector lymphocytes of various types. The modus operandi of the destructive process varies as a function of the type of graft involved as well as the degree of disparity between donor and recipient transplantation antigens. For example, the hyperacute rejection of kidneys that can occur within hours following transplantation is mediated by antibodies, whereas acute rejection is a lymphocyte-mediated process. In the absence of prophylactic treatment, rejection of most allografts is usually complete within about 2 weeks.

In the case of grafts exchanged between identical human twins, or between members of the same highly inbred or syngeneic strain of experimental animals, rejection does not occur because donor and recipient are genetically identical, and consequently no antigenic differences are involved. An exception to this rule is observed when the organ being replaced in the recipient was destroyed by an autoimmune disease process. For example, islets of Langerhans transplanted from a normal twin into a monozygotic sibling suffering from type 1 autoimmune diabetes mellitus are destroyed by recurrence of autoimmunity, an acquired immunological difference between the twins. *See* AUTOIMMUNITY; DIABETES.

There are a few special exemptions from rejection that apply to certain sites in the body or to certain types of graft. For example, the use of corneal allografts in restoring sight to individuals with corneal opacification succeeds because of the absence of blood vessels in the host tissue. Allogeneic islets of Langerhans transplanted into the testis survive in the absence of any antirejection therapy through a mechanism that is not well understood. Mammalian fetuses are natural, highly successful allografts in the maternal uterus in the sense that their tissues express foreign antigens inherited from the father. The reason for their success is still not completely understood. Contributory factors include (1) the absence of continuity between maternal and fetal blood circulations in the placenta; (2) failure of the layer of fetal tissue in direct contact with maternal tissue in the placenta (the trophoblast) to express transplantation antigens in an effective manner; (3) the local operation of nonspecific immunosuppressive factors, including hormones and suppressor cells; and (4) the absence of certain co-stimulatory signals required for lymphocyte activation. *See* PLACENTATION; PREGNANCY.

Overcoming graft rejection. Successful transplantation of allografts such as kidneys and hearts currently requires suppressing the recipient's immune response to the graft without seriously impairing the immunological defense against infection. Treatment of individuals with so-called immunosuppressive drugs and other agents prevents allograft rejection for prolonged periods, if not indefinitely. Under cover of nonspecific immunosuppression, the recipient's immune system appears to undergo an adaptation to the presence of the graft, allowing the dosage of the drugs to be reduced. However, in almost all successfully transplanted individuals, drug therapy at some low dose is required indefinitely.

Unfortunately, long-term treatment with drugs that nonspecifically suppress or disrupt the immune system does interfere with normal immune function and over time can predispose to the development of infection or cancer. The ideal solution would be to abrogate specifically the body's capacity to react against the foreign antigens of the graft, leaving its reactivity to other antigens completely unimpaired. Such a solution is likely to be based on the phenomenon of immunological tolerance—the survival of foreign (allogeneic or xenogeneic) tissue in normal recipients in the absence of immunosuppression. Analyses of the mechanisms of this phenomenon in transplanted laboratory animals suggest that it is possible to induce a state of graft tolerance through a variety of different strategies. A clinically applicable solution is not yet available, but recent advances suggest that "tolerance-based" transplantation may soon become a reality. *See* ACQUIRED IMMUNOLOGICAL TOLERANCE; IMMUNOSUPPRESSION.

Immunosuppressive agents. Whole-body irradiation was one of the first forms of immunosuppression employed in transplantation, but its use is now confined largely to stem cell transplantation. Most graft recipients receive immunosuppressive drugs, which suppress the proliferation of cells involved in the immune response. One of the first of these drugs was azathioprine, introduced in 1962. It is an antimetabolite that interferes with protein synthesis. Steroids, frequently administered in high doses to prevent or treat acute rejection, have many actions, including depression of nucleic acid and protein synthesis, suppression of host inflammatory response, and killing of lymphocytes.

Newer drugs target effector lymphocytes responsible for graft rejection. Cyclosporin, tacrolimus, sirolimus, and mycophenolate mofetil have dramatically improved results with human renal transplantation. Each interferes with the proliferation of

activated T lymphocytes. Monoclonal antibodies directed against surface antigens present on activated T cells, for example the interleukin-2 receptor or the CD3 antigen, now play an increasingly important role in the initial prevention of rejection and in the treatment of later rejection crises. These monoclonal reagents lead to the destruction or inactivation of T cells to which they bind. *See* MONOCLONAL ANTIBODIES.

Tissue typing. An individual's response against an allograft is directed against a large number of cell-surface transplantation antigens controlled by allelic genes at many different loci. However, in all species, one of these loci, the major histocompatibility complex (MHC), transcends all the other histocompatibility loci (the minor loci) in terms of its genetic complexity and the strength of the antigenic response it controls. In humans, the MHC, known as the HLA (human leukocyte antigen) complex, is on the sixth chromosome; its principal loci are designated A, B, C, DR, and DQ. The allelic products of the HLA genes can be detected by serology, polymerase chain reaction technology, or microcytotoxicity assays. The ABO red cell antigens are also important because they are expressed on all tissues. *See* BLOOD GROUPS; HISTOCOMPATIBILITY.

In kidney transplants between closely related family members, the degree of HLA antigen matching can be determined very precisely, and there is a very good correlation between the number of shared HLA antigens and the survival of the graft. With grafts from unrelated donors, HLA matching is more difficult and can delay transplantation, but it may be beneficial. The estimated 10-year rate of kidney graft survival was recently reported to be 52% for HLA-matched transplants, as compared with 37% for HLA-mismatched transplants. Donors and recipients are usually matched for both alleles at the HLA-A, -B, and -DR loci. HLA matching is not as clearly beneficial in the case of most other solid organ grafts, and no attempt is made to HLA-match heart, lung, liver, and pancreas grafts. With few exceptions, however, most donors and recipients are matched for the expression of ABO blood group antigens.

The early assumption that blood transfusions should not be given to transplant recipients for fear of sensitizing them has been disproved. Prior transfusion of these patients actually increased cadaver kidney survival rates when older immunosuppressive regimens were in use, but is no longer needed since the advent of newer drugs. The combination of donor-specific transfusion (a transfusion provided by the individual who is also the organ donor) with reagents that prevent signals required for T-cell activation is under investigation as a method of tolerance-based transplantation.

Bone marrow transplantation presents a unique problem in its requirements for HLA matching and for immunosuppression in advance of grafting. In addition to the possibility of rejection of the graft by the recipient, by virtue of immunologically competent cells still present in the recipient, bone marrow grafts can react against the transplantation antigens of their hosts. These are known as graft-versus-host reactions, and they can be fatal. *See* IMMUNOLOGY.

Xenotransplantation. As success rates for allograft transplantation to cure disease have increased, the supply of donor organs has been increasingly inadequate. Cadaveric donor organs are sufficient for approximately 10% of total current need. In certain special cases, for example skin and pancreatic islets, it may become possible to genetically engineer a suitable replacement tissue. Another possible source of replacement tissue is human fetal material, but ethical and legal barriers to its use may be insurmountable. A final alternative is the use of xenografts. Pigs, for example, could supply an unlimited number of organs similar in size and function to those required by humans. The immunological events that lead to the rejection of xenografts are, however, different from and less well understood than those responsible for allograft rejection. The small number of xenografts attempted to date have failed. In particular, xenografts are susceptible to hyperacute rejection by humans. This is due to the presence of certain glycoproteins in blood vessels of many species that are recognized by antibodies present in the blood of all humans. The potential use of xenografts in humans raises special ethical and psychological issues. It also raises the specter of introducing new diseases into the human population. The transfer of pig endogenous retroviruses to cultured human cells has been demonstrated. The specter of xenozoonotic epidemics has become a concern.

John P. Mordes; Rupert E. Billingham; Dale L. Greiner; Aldo A. Rossini

Bibliography. D. K. Cooper and R. P. Lanza, *Xeno: The Promise of Transplanting Animal Organs into Humans*, Oxford University Press, 2000; E. Fuchs and J. A. Segre, Stem cells: A new lease on life, *Cell*, 100:143–155, 2000; H. Gudmundsdottir and L. A. Turka, Transplantation tolerance: Mechanisms and strategies?, *Semin. Nephrol.*, 20:209–216, 2000; N. S. Hakim (ed.), *Introduction to Organ Transplantation*, World Scientific Publishing, River Edge, NJ, 1998; F. Locatelli, D. Rondelli, and G. R. Burgio, Tolerance and hematopoietic stem cell transplantation 50 years after Burnet's theory, *Exp. Hematol.*, 28:479–489, 2000; A. M. Marmont, New horizons in the treatment of autoimmune diseases: Immunoablation and stem cell transplantation, *Annu. Rev. Med.*, 51:115–134, 2000; J. L. Platt and T. Nagayasu, Current status of xenotransplantation, *Clin. Exp. Pharmacol. Physiol.*, 26:1026–1032, 1999; A. A. Rossini, D. L. Greiner, and J. P. Mordes, Induction of immunological tolerance for transplantation, *Physiol. Rev.*, 79:99–141, 1999.

Transport processes

The processes whereby mass, energy, or momentum are transported from one region of a material to another under the influence of composition, temperature, or velocity gradients. If a sample of a material in which the chemical composition, the temperature,

or the velocity vary from point to point is isolated from its surroundings, the transport processes act so as eventually to render these quantities uniform throughout the material. The nonuniform state required to generate these transport processes causes them to be known also as nonequilibrium processes. Associated with gradients of composition, temperature, and velocity in a material are the transport processes of diffusion, thermal conduction, and viscosity, respectively. For a large class of materials, the laws that govern the transport processes are quite simple.

Diffusion. **Figure 1** shows a sample of a material which is composed of two chemical species. The sample is stationary and has a uniform temperature throughout, but a composition difference is maintained across its two ends, and in this steady state the two species continuously migrate down their concentration gradients. The composition of the material may be expressed by means of the molar concentration of one species, c_1 (moles/m^3); then it is found that the number of moles of this species which cross unit area of the sample perpendicular to the z direction in unit time, known as the flux of mass (J_1), is given by Eq. (1), which is Fick's law of diffusion.

$$J_1 = -D\frac{dc_1}{dz} \tag{1}$$

The constant of proportionality, D, between the mass flux and the concentration gradient, which depends upon the nature of the material, its temperature, pressure, and composition, is known as the diffusion coefficient.

The phenomenon of diffusion occurs widely in nature, and it is frequently important in technological applications. For example, the transpiration of the leaves of plants, in which they absorb carbon dioxide from the atmosphere and give off water vapor, is controlled by a diffusion process. The rates of many chemical reactions in fluids that are promoted by catalysts may similarly be controlled by the diffusion of reactants to the active catalyst sites. *See* DIFFUSION.

Thermal conduction. In **Fig. 2**, a sample of a material is subjected to a steady temperature difference between two faces perpendicular to the z direction. Under these conditions, energy is continually transported from the hotter face to the colder, and the

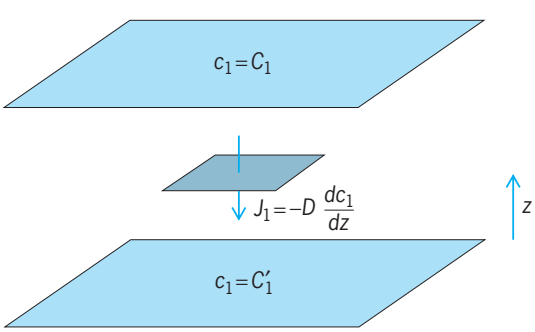

Fig. 1. Diffusion in a sample of material composed of two chemical species, c_1 and C'_1 represent the molar concentrations of one of the species in the two planes bounding the material.

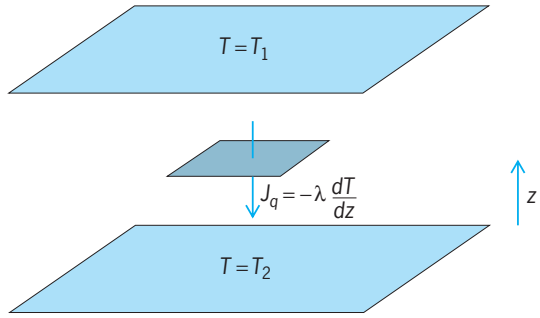

Fig. 2. Thermal conductivity in a sample of material subjected to a steady temperature difference.

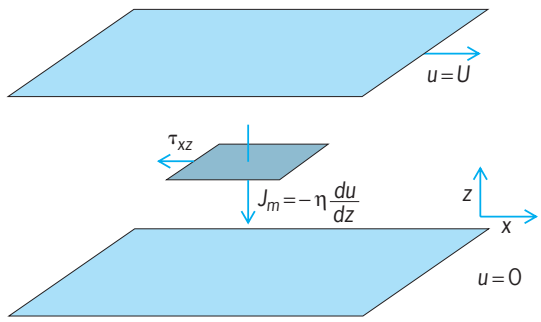

Fig. 3. Viscosity in a fluid whose upper surface is in contact with a solid boundary.

energy flux J_q in the z direction (the energy crossing unit area in unit time) is given by Fourier's law as Eq. (2). The constant of proportionality, λ, between

$$J_q = -\lambda\frac{dT}{dz} \tag{2}$$

the flux and the temperature gradient, dT/dz, is the thermal conductivity coefficient, which again depends upon the material as well as its temperature, pressure, and composition. *See* THERMAL CONDUCTION IN SOLIDS.

Viscosity. The phenomenon of viscosity is associated with the gradient of velocity in a material. Since it is difficult to maintain velocity gradients in solids, the phenomenon is readily observed only in fluids. Because velocity is a vector quantity, **Fig. 3** shows a fluid whose upper surface is in contact with a solid boundary which moves with a steady velocity U in the x direction only; the lower surface of the fluid is held stationary. As a result, various layers of the fluid in the z direction move with different x-direction velocities u. Associated with the motion in the x direction the fluid possesses a momentum, and the x-direction momentum is transported down the velocity or momentum gradient. The flux of x momentum in the z direction, J_m, is equivalent to a tangential shear stress τ_{xz} acting in the negative x direction on each layer of the fluid. This means that a tangential force must be applied to the upper plate to keep it in steady motion. Again the flux is proportional to the imposed gradient and is given by Eq. (3), which is Newton's law of viscosity. The pro-

$$J_m = \tau_{xz} = -\eta\frac{du}{dz} \tag{3}$$

portionality constant η is the viscosity coefficient for the material, and it too depends on the thermodynamic state of the material. The phenomenon of viscosity is revealed whenever a fluid flows near a solid boundary, and it is therefore of significance in almost every aspect of engineering. *See* VISCOSITY.

Thermal processes. Other, more subtle transport processes can occur. For example, in a mixture of two chemical species, the imposition of a temperature gradient leads not only to energy transport but also to a mass transport that causes a partial separation of the mixture. This phenomenon is known as thermal diffusion. Conversely, when diffusion takes place in an initially isothermal mixture as a result of a composition gradient, small temperature gradients can be observed in the material arising from an energy transport. This is the diffusion thermoeffect.

Transport coefficients. The coefficients D, λ, and η are known collectively as transport coefficients. The measurement of these coefficients for materials in solid, liquid, and gaseous phases has been the object of a considerable research effort for many years. The measurements can be carried out only rarely by directly implementing the situations envisaged in Figs. 1–3. This is because it is difficult to achieve the one-dimensional gradients of the quantities required when the sample is of a finite size. The exceptions to this are the measurement of thermal conductivity and diffusion in solids, where simple methods have proved effective. In fluids, the diffusion coefficient has most often been determined in a time-dependent experiment in which an initial concentration gradient in a mixture is allowed to decay in a closed vessel of known geometry. The approach to equilibrium, which is governed by the diffusion coefficient, is observed with a suitable concentration monitor.

The coefficient of thermal conductivity in fluids is also most accurately determined in a transient experiment. The fluid surrounds a thin, vertical wire which is suddenly heated by an electric current; the rate of the temperature rise of the wire is observed, and the thermal conductivity of the fluid is deduced from it.

The viscosity coefficient of gases and liquids is generally determined by one of two techniques. In the first, the fluid flows through a capillary tube of known geometry, and the pressure difference across the ends of the tube necessary to maintain a given rate of flow is determined. This pressure difference is then proportional to the viscosity coefficient of the fluid. In the second method, the damping of the torsional oscillations of a thin, solid, horizontal, circular disk is observed when the disk is suspended in the fluid. The measurement of the logarithmic decrement of the oscillations serves to determine the viscosity coefficient. *See* QUASIELASTIC LIGHT SCATTERING.

The results of measurements of the transport coefficients of materials are of importance since there are few technological activities that do not involve one or more of the transport processes. However, because the transport coefficients derive their values from the properties and behavior of the atoms and molecules that make up the material, they are also of more fundamental significance. In the particular case of gases at low density (near atmospheric pressure), the kinetic theory of gases has provided an almost complete description of their transport coefficients. In such gases the molecules are in continual random motion, undergoing collisions with each other and with the walls of a containing vessel. Thus, in a gas subject to, for example, a velocity gradient, the molecules that have an additional component of momentum, because they are in regions of high velocity in the gas, transport that momentum down the velocity gradient. Naturally, the rate of transport of momentum is influenced by the number of molecules transporting the momentum and their speed, as well as the likelihood and effect of collisions with other molecules. For the simplest molecular model, in which the molecules are seen as rigid elastic spheres of diameter σ, there is no force between molecules except when they touch. Thus, the dynamics and effects of the collisions are straightforward, and the transport coefficients for viscosity and thermal conductivity are given by Eqs. (4) and (5), in which m is the mass of a molecule, k is Boltzmann's constant, T the absolute temperature, and c_v the specific heat of the gas. Real molecules exert forces on each other over a distance, which are expressed in terms of an intermolecular pair potential such that the forces between molecules are repulsive at short range and attractive at long range. This means that the dynamics of molecular encounters are more complicated and the effects of the collisions on the transport of momentum and energy are altered. Within the kinetic theory of gases, these effects can be included; for systems of monatomic gases, the only modification to the formulas given in Eqs. (4) and (5) is effectively to introduce a temperature-dependent diameter $\sigma(T)$. The effective diameter is smaller at higher temperatures because the atoms approach each other more closely as their translational kinetic energy increases. For polyatomic molecules that possess internal energy, this single idea remains valid for the viscosity; but for the thermal conductivity, it is necessary to consider the transport of internal energy down the temperature gradient as well, and the corresponding theoretical result is more complicated. When the transport coefficients are evaluated for real molecules, the viscosity and thermal conductivity are found to increase with temperature approximately linearly, whereas the diffusion coefficient varies more nearly as the square of the temperature.

$$\eta = \frac{5}{16}\frac{(\pi mkT)^{1/2}}{\pi\sigma^2} \tag{4}$$

$$\lambda = \frac{25}{32}\frac{(\pi mkT)^{1/2}}{\pi\sigma^2}\frac{c_v}{m} \tag{5}$$

At low densities the viscosity and thermal conductivity of a gas are independent of pressure, while the diffusion coefficient is inversely proportional to

it. As the density of the material is increased toward that of a liquid and finally to that of a solid, significant changes in the molecular mechanism of the transport processes occur. The transport by free molecular motion becomes a smaller contribution as the volume available for such motion decreases. In addition, the attractive forces between molecules, which become increasingly significant, tend to inhibit molecular motion. Thus, on the one hand, the diffusion coefficient in condensed phases, which is still determined by molecular motion, is very much smaller (about 10^4 times) than that in low-density gases. On the other hand, the viscosity of liquids is very much greater than that in gases because the attractive forces between molecules make the relative motion of various layers in the fluid much more difficult to achieve. Because increasing the temperature of a liquid increases the average separation of the molecules as well as their energy, the diffusion coefficients for liquids increase rapidly with temperature and, for the same reasons, the viscosity decreases. In the solid the molecules acquire almost fixed positions, and the diffusion coefficient consequently becomes even smaller, whereas the viscosity is practically infinite. W. A. Wakeham

Bibliography. J. Kestin and W. A. Wakeham, *Transport Properties of Fluids: Thermal Conductivity, Viscosity and Diffusion Coefficients,* 1988; G. C. Maitland et al., *Intermolecular Forces: Their Origin and Determination,* 1987.

Transportation engineering

That branch of engineering related to the safe and efficient movements of people and goods. The primary modes of travel considered in transportation engineering include roadway, rail, air, water, and pipeline, along with nonmotorized pedestrian and bicycle travel. Special categories include urban and intermodal transportation.

Highway transportation. Engineering for highway transportation involves planning, construction, and operation of highway systems, urban streets, roads, and bridges, as well as parking facilities. Important aspects of highway engineering include (1) overall planning of routes, financing, environmental impact evaluation, and value engineering to compare alternatives; (2) traffic engineering, which plans for the volumes of traffic to be handled, the methods to accommodate these flows, the lighting and signing of highways, and general layout; (3) pavement and roadway engineering, which involves setting of alignments, planning the cuts and fills to construct the roadway, designing the base course and pavement, and selecting the drainage system; and (4) bridge engineering, which involves the design of highway bridges, retaining walls, tunnels, and other structures. *See* HIGHWAY ENGINEERING; TRAFFIC-CONTROL SYSTEMS; VALUE ENGINEERING.

Highway transportation engineers face the challenge of moving increasing volumes of traffic over existing routes while improving safety records. Therefore, a major initiative was launched in the United States to develop technology for a nationwide intelligent vehicle highway system (IVHS). In urban areas, such systems are being designed to feature increased use of vehicle detection systems, video cameras, variable message signs, and electronic toll collection with automatic vehicle identification. In more remote areas, sensors on highways are being designed to detect unsafe conditions such as slippery roads, relay the information via radio waves to a central computer, and transmit a warning via satellite to individual vehicles. Eventually, new cars will be instrumented so that the driver can communicate with municipalities and receive directions or other information. Posted speed limits will be changed automatically according to traffic flow.

In modern highway construction, particularly in urban areas, major emphasis is placed on developing regional transportation plans utilizing significant public input. Highways must meet the transportation needs of the public, adhere to strict environmental regulations, and satisfy concerns over traffic and esthetics, in order to gain wide support. Air pollution, noise problems, and destruction of wetlands are some concerns that must be carefully addressed.

Highway pavements must be maintained with surfaces of acceptable riding quality. This places severe demands on paving materials because of heavy traffic volumes, repetitive truck axle loads, weather conditions that may vary from severe freezing to hot summers, and the use of chemicals to aid in snow and ice removal. As a result of a national Strategic Highway Research Program (SHRP) in the United States, an improved asphalt paving system known as Superpave has been developed and has been implemented to improve pavement life substantially. *See* PAVEMENT.

Bridge engineers are now using new or improved materials to develop more cost-effective, longerlasting structures. A high-performance steel (HPS-70W) that provides improved toughness, weldability, and strength has been used to build several bridges. Also, a high-performance concrete (HPC) that provides higher strength and improved workability is being used. Bridge rehabilitation is a major area for infrastructure innovation. Estimates show that approximately one-third of the bridges in the United States are in urgent need of repair. Deterioration of the bridge deck often leads to deterioration of the underlying superstructure, and methods of constructing more durable decks have been developed. This includes the use of epoxy-coated steel reinforcing bars and high-performance concrete with microsilica for deck construction. Precast elements are sometimes effective in deck replacement, because disruption to traffic can be minimized. Bridges are susceptible to the forces generated by seismic events, as evidenced by the destruction resulting from the Loma Prieta and Northridge earthquakes in California. Bridge engineers have developed methods for retrofitting existing structures to resist earthquakes, as well as designing improved details for new construction. To increase load capacity and reliability, posttensioned cables can be added to beams and girders for longer

spans. For shorter spans of about 50 ft (15 m) or less, bridges can be replaced with large culverts constructed from prefabricated steel or aluminum components, or from precast concrete units. Other techniques include the use of lead-rubber seismic isolation devices to replace traditional bearings at supports, and constructing steel or fiberglass casings around concrete bridge columns to increase strength and ductility. *See* BRIDGE; EARTHQUAKE; STRUCTURAL MATERIALS.

Rail transportation. Engineering for railway transportation involves planning, construction, and operation of terminals, switchyards, loading/unloading facilities, trackage, bridges, tunnels, and traffic-control systems for freight and passenger service. For freight operations, there is an emphasis on developing more efficient systems for loading, unloading, shifting cars, and operating trains. Facilities include large marshaling yards where electronic equipment is used to control the movement of railroad cars. Also, there is a trend to developing more automated systems on trackage whereby signals and switches are set automatically by electronic devices. To accommodate transportation of containers, tunnels on older lines are being enlarged to provide for double-stack container cars.

Efforts for more efficient operations have begun to pay dividends. Reversing the declines of earlier years, rail freight volume has been steadily increasing since 1987. New facilities are being built to handle the increased volume, and tracks are being added at points of congestion. Although the total trackage in the United States today is only 113,000 mi (182,000 km), major railroads haul significant tonnage.

In Europe and Japan, rail passenger service has long been an important part of the overall transportation system. In the United States, passenger service is small compared to that of earlier years, but there is a renewed interest in rail passenger service in certain congested urban corridors. In Japan and Europe, high-speed systems have been developed. For example, the French have developed an Aerotrain that is propelled by a fan jet engine and floats on a cushion of low-pressure air on a guideway at 267 mi/h (430 km/h). A more conventional high-speed system introduced in 1994, the Eurostar, operates at service speeds of up to 190 mi/h (300 km/h). One development that has received limited attention in Germany, Japan, and the United States is a magnetic levitation system in which cars on a frictionless magnetic suspension are propelled along guideways by linear induction motors. A proposed line in Germany would be capable of speeds near 300 mi/h (480 km/h). *See* MAGNETIC LEVITATION.

Most high-speed rail systems require new alignments to accommodate the increased speed. However, in many areas it is difficult and expensive to construct new alignments. Thus, engineers are challenged to develop high-speed rail passenger cars that can operate on existing trackage. One approach has been to develop suspension systems that tilt the coaches on curves to enhance stability. In the mid-1990s the fastest trains in the United States were running about 100 mi/h (160 km/h). In 1998 the Transportation Research Board initiated a program to develop technologies to upgrade rail systems to higher speeds.

A major accomplishment in intercontinental rail transportation was the completion in late 1993 of the Channel Tunnel. Culminating nearly 8 years of design and construction, the twin tunnels facilitate a trip from London to Paris in 3 h. The tunnel was designed to handle a full range of services, including conventional and intermodal freight operations, motor car and truck carrying services, and high-passenger trains. *See* RAILROAD ENGINEERING; TUNNEL.

Air transportation. Engineering for air transportation encompasses the planning, design, and construction of terminals, runways, and navigation aids to provide for passenger and freight service. High-capacity, long-range, fuel-efficient aircraft, such as the 440-seat Boeing 777 with a range of 7200 mi (12,000 km) are desirable. Wider use of composites and the substitution of electronic controls for mechanical devices reduce weight to improve fuel economy. Smaller planes are more efficient for shorter runs. *See* AIR TRANSPORTATION; COMPOSITE MATERIAL.

Many airport terminals for smaller population centers are designed with little uncommitted space but with provisions for incremental expansion as traffic grows. In many larger population centers, state-of-the-art facilities and major improvements have been constructed, such as the terminals at Denver and Washington, D.C. (both National and Dulles). As air transportation continues to grow, increased emphasis is placed on terminal designs that facilitate the rapid movement of passengers and baggage. Extensive conveyor belt systems with electronic identifications have been installed in the newer facilities to speed movement of baggage from planes to common central areas. At the larger airports, moving sidewalks are commonly used to move passengers through widely dispersed gate areas. Also, computer-controlled shuttle cars without operators (people movers) are used to move people rapidly from central areas to gate areas. In some cases, rapid transit rail facilities or dedicated busways are available to carry passengers directly from the airport to central business districts. As airways continue to concentrate flights in large hub operations to increase efficiency, engineers and planners focus on developing improved facilities to speed the transfer of passengers and baggage. *See* AIRPORT ENGINEERING.

Air freight cargo facilities are increasingly being automated. An example is the cargo facility at New York's John F. Kennedy Airport, which features a computerized control system, a network of automated distribution conveyors, transfer vehicles, and stacking cranes. A large rack storage building with robotic cranes provides temporary storage and speeds loading and unloading of roll boxes and lightweight containers that hold groups of packages.

Runways must be of sufficient size to safely accommodate landings and takeoffs. The design of

pavements for runways and taxiways involves many of the same considerations as for highways, except that runways must be designed to reliably withstand the heavier wheel loads imposed by the aircraft. Navigation aids and other instrumentation, such as for the detection of destructive wind shears or approaching aircraft, are being developed and deployed to enhance safety. *See* AIR NAVIGATION; AIR-TRAFFIC CONTROL.

Water transportation. Engineering for water transportation entails the design and construction of a vast array of facilities such as canals, locks and dams, and port facilities. The transportation system ranges from shipping by barge and tugboat on inland waterways to shipping by oceangoing vessels. Although there is some transportation of passengers, such as on ferries and cruise ships, water transportation is largely devoted to freight.

One important inland waterway in North America is the St. Lawrence Seaway, which connects natural waterways, canals, and the Great Lakes. It permits shipping from the Atlantic Ocean to Duluth, Minnesota. Another is the Ohio-Missouri-Mississippi River System, which extends from Pittsburgh, Kansas City, and Minneapolis to the Gulf of Mexico. Many of the locks and dams in the latter system were constructed many years ago. Engineers are challenged to design and construct larger locks to serve larger barges. Improved navigation systems are needed to enhance safety. *See* CANAL; DAM; RIVER ENGINEERING.

Special facilities are required for loading and unloading oceangoing vessels. For example, supertankers require special offshore mooring where the oil can be rapidly pumped to shore through pipelines. Also, unit trains of 100 or more railcars, carrying bulk products such as grain or coal, require rapid unloading facilities for efficient operations. *See* HARBORS AND PORTS.

Pipeline transportation. Pipeline engineering embraces the design and construction of pipelines, pumping stations, and storage facilities. Pipelines are used to transport liquids such as water, gas, and petroleum products over great distances. Also, products such as pulverized coal and iron ore can be transported in a water slurry.

Water pipelines, probably the most common, can run from minimal sizes to diameters of 20 ft (6 m) or more for large penstocks. Pipelines are often underground, but may run aboveground, particularly in lightly populated areas. Either submerged lines or bridges are required at stream crossings. One aboveground line is the 798-mi (1280-km) 48-in.-diameter (1200-mm) Trans-Alaska pipeline, which was constructed under difficult conditions to transport oil from near Prudhoe Bay on the Arctic Ocean to the city of Valdez on the Gulf of Alaska.

Consideration must be given to route selection, determining the appropriate diameter and thickness of pipe, installation (trench construction, backfilling, and compaction), and durability. Installation under highways, rivers, and other difficult areas is sometimes done using microtunneling techniques to avoid disturbing the surface. With this method, the pipe is advanced underground by jacking from one side, using a laser-controlled guidance system to maintain pipe alignment. Design of pumping facilities requires study of power requirements for different types of material moved, standby facilities, and related considerations. *See* PIPELINE; STORAGE TANK.

Urban transportation. Engineering for urban transportation concerns the design and construction of light rail systems, subways, and people-movers, as well as facilities for traditional bus systems. To enhance public acceptance of new and expanded systems, increased use is being made of computer-aided design (CAD) to visualize alternatives for stations and facilities. Also, animated video systems are used for interactive visualization of plans. *See* COMPUTER-AIDED ENGINEERING.

As congestion in urban areas grows, increased emphasis is being placed on the construction and expansion of subways in major cities. Tunnel-boring machines and segmental liners are often used, along with various techniques to mobilize the strength of the surrounding soil or rock. Light-rail cars are used interchangeably aboveground on controlled right-of-ways or underground through tunnels. *See* SUBWAY ENGINEERING.

In some cities, high-occupancy vehicle lanes have been used to encourage car pooling. Typically these special lanes are reserved during rush hours for vehicles with three or more passengers.

Intermodal transportation. Intermodal transportation, often referred to as containerization, entails the use of special containers to ship goods by truck, rail, or ocean vessel. Engineers must design and construct intermodal facilities for efficient operations. The containers are fabricated from steel or aluminum, and they are designed to withstand the forces from handling. The ships are constructed with a cellular grid of compartments for containers below deck, and they can accommodate one or two layers on deck as well. Advantages include savings in labor costs, less pilferage, and lower insurance costs.

Seaports have special facilities for handling the containers, which arrive by rail or truck. For example, a modernization project in Boston, Massachusetts, provides a 13-acre (5.2-hectare) yard for container storage with 1950 ft (594 m) of continuous berth face for ships. Containers are handled by two special 30-ton-capacity (27-metric ton) gantry cranes that travel on rails with a 96-ft (29-m) gage, and provide an outreach of 115 ft (35 m). Such facilities, which are found throughout the major port cities, are designed to speed the transfer of freight. At major inland centers, intermodal facilities combine loading and handling facilities for containers from truck trailers and rail flatcars. *See* HOISTING MACHINES; MARINE CONTAINERS; MERCHANT SHIP.

Environmental considerations. The environment is a major consideration in planning, designing, and constructing transportation facilities. Extensive legislation at both federal and state levels has set forth requirements with the objective of protecting the health and welfare of the general public.

Beginning with the National Environmental Policy Act (NEPA) of 1970, over 25 major federal laws and countless amendments have been enacted. One provision of NEPA is the requirement for approved environmental impact statements (EIS) for federally funded projects. The EIS must document how the environment will be affected, including identification of any unavoidable adverse effects. Opportunities for public involvement in the process must be provided, and alternatives must be identified. To help control air quality, the Clean Air Act administered by the Environmental Protection Agency (EPA) requires review and approval of any new highway with an anticipated average daily traffic volume of more than 20,000 vehicles, or any new parking facility for 1000 cars or more. Approvals are also required for major modifications that increase capacity. The Noise Control Act limits noise levels from surface transportation, often requiring the construction of noise walls adjacent to highways in urban areas. To control water pollution, the Clean Water Act requires permits for discharges into streams and wetlands, both from construction activities and from storm water runoff of the completed facility. *See* AIR POLLUTION; WATER POLLUTION.

To avoid exposure of both workers and the public to lead poisoning, special precautions must be taken in the removal and containment of old, deteriorating paint systems from steel bridges. Regulations of the Occupational Safety and Health Administration (OSHA) must be met. Lead-based paint systems were used until the early 1980s, but new steel bridge construction uses either bare weathering steel or modern lead-free paint systems.

Because of the vast material quantities used to construct and maintain transportation systems, opportunities abound for the use of reclaimed and recycled resources. For example, in the production of new pavement, 20% or more of the content can be from old asphalt pavement that has been removed in repaving operations. Old concrete pavements can also be recycled, as aggregate for new concrete or for base layers. Industrial wastes such as fly ash and blast-furnace slag can be used in concrete. Many uses have been found for scrap tires, such as for erosion control devices, for safety devices (tire-sand inertial barriers), in chipped form to create a lightweight bridge approach embankment or to reduce earth pressure against abutment walls, and in the manufacture of rubber asphalt. *See* RECYCLING TECHNOLOGY.

Energy considerations. Engineers must help meet the challenge to reduce energy consumption by designing efficient transportation systems. About 97% of transportation energy in the United States is derived from oil. Indeed, the transportation system uses about 65% of all oil consumed by the United States. Although the fuel economy of vehicles has been improved significantly, this factor has been offset by a trend toward decreased vehicle occupancy rates. As a result, total usage of energy in highway passenger transportation is increasing. Despite significant gains that have been made in energy efficiency for commercial air travel and for rail freight, total energy consumption for transportation of all types continues to grow.

Efforts to curb energy use arise from a variety of concerns, including national security issues and environmental implications. Limiting carbon dioxide emissions to stem global warming is a significant consideration. In addition to fuel economy standards, legislative approaches to date have focused on promoting cleaner and more efficient alternative fuels. As an example, natural gas is now used to fuel certain commercial vehicles in some areas. However, because personal travel in light-duty vehicles consumes over 50% of transportation energy, effective efforts in energy conservation must be directed toward them. Further improvements in fuel economy will likely lead to lighter-weight vehicles with higher first costs. Efforts to relieve congestion in urban areas through incentives to make greater use of car pooling, such as special freeway lanes, and encouraging greater use of mass transit, deserve further emphasis.

Roger L. Brockenbrough

Bibliography. J. Banks, *Introduction to Transportation Engineering*, 2d ed., 2002; R. L. Brockenbrough and K. J. Boedecker, Jr., *Highway Engineering Handbook*, 2d ed., 2003; M. A. Chowdhury and A. W. Sadek, *Fundamentals of Intelligent Transportation Systems Planning*, 2003; Federal Highway Administration, *Manual on Uniform Traffic Control Devices*, 2003; C. Khisty and B. Lall, *Transportation Engineering: An Introduction*, 3d ed., 2003; R. Roess, E. Prassas, and W. McShane, *Traffic Engineering*, 3d ed., 2004; Transportation Research Board, *Highway Capacity Manual*, 2000; P. Wright and K. Dixon, *Highway Engineering*, 7th ed., 2004.

Transposons

Types of transposable elements which comprise large discrete segments of deoxyribonucleic acid (DNA) capable of moving from one chromosome site to a new location. In bacteria, the transposable elements can be grouped into two classes, the insertion sequences and the transposons. The ability of transposable elements to insert into plasmid or bacterial virus (bacteriophage) which is transmissible from one organism to another allows for their rapid spread. *See* BACTERIOPHAGE; PLASMID.

The insertion sequences were first identified by their ability to induce unusual mutations in the structural gene for a protein involved in sugar metabolism. These insertion sequences are relatively small (about 500–1500 nucleotide pairs) and can only be followed by their ability to induce these mutations. Most bacterial chromosomes contain several copies of such insertion sequence elements.

The transposons are larger segments of DNA (2000–10,000 base pairs) that encode several proteins, usually one or two required for the movement

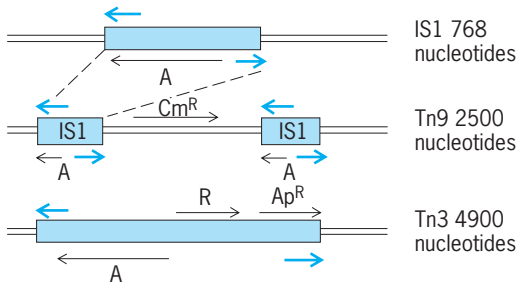

Structure of a typical transposable element. The colored arrows indicate the terminal inverted repeats characteristics of each element. Note that Tn9 is a composite transposon derived from directly repeated IS1 elements. The black arrows indicate genes for proteins involved in transposition (A and R) or antibiotic resistance. The Tn9 transposon is resistant to chloramphenicol (Cm), while the Tn3 element encodes resistance to ampicillin (Ap) and its derivative penicillin.

of the element and often an additional protein that imparts a selective advantage to the host containing a copy of that element. The structure of many transposons suggests they may have evolved from the simpler insertion sequence elements. For example, the transposon Tn9 contains two copies of the element IS1 flanking a region of unique DNA encoding resistance to the antibiotic chloramphenicol.

All transposable elements, both the simple insertion sequence elements and the more complex transposons, have a similar structure and genetic organization (see **illus.**). The ends of the element represent recognition sites and define the segment of DNA undergoing transposition. A short sequence present at one end of the element is repeated in an inverted fashion at the other end. These terminal inverted repeats are characteristic for each element. In the case of the composite transposons like Tn9, the inverted repeats present at the end of each IS1 element result in the entire transposon also having inverted repeats. One or more proteins essential in the recognition of the inverted repeat are encoded in the body of the element.

Members of a widespread group of transposons, the Tn3 family, all have a similar structure and appear to move by a similar mechanism. Transposase, one protein encoded by the element, promotes the formation of intermediates called cointegrates, in which the element has been duplicated by replication. A second element-encoded protein, resolvase, completes the process by converting the cointegrates into the end products of transposition, a transposon inserted into a new site. A third protein encoded by the Tn3 element imparts resistance to the antibiotic ampicillin.

Transposons are known that encode resistances to almost all antibiotics as well as many toxic metals and chemicals. In addition, some transposons have acquired the ability to direct the synthesis of proteins that metabolize carbohydrates, petroleum, and pesticides. Other transposable elements produce enterotoxins that cause travelers to become ill from drinking water contaminated with bacteria carrying the element. The broad spectrum of activities encoded by the transposable elements demonstrates the strong selective advantage that has accompanied their evolution.

The bacteriophage Mu (mutator) replicates itself in a mechanism that involves transposition into many sites in the host genome. In the process of high-frequency transposition, the bacteriophage often mutates genes in the host organism. Other phages have adopted transpositionlike events for special purposes.

Transposable elements are not restricted to prokaryotes. Yeast as well as higher eukaryotes have DNA segments that move and cause mutations. In fact, the earliest models suggesting the existence of transposable DNA segments were based on genetic work by B. McClintock in the 1930s with corn plants. The eukaryotic elements have much in common with their prokaryotic counterparts: the termini of the elements are composed of inverted repeats, and many of the larger elements are composed of two small insertion sequence–like regions flanking a unique central region. One class of eukaryotic virus, the ribonucleic acid (RNA) retrovirus, also has this structure and is thought to integrate into the host chromosome through a transposition-like mechanism. *See* ANTIBIOTIC; GENE; RETROVIRUS; VIRUS.

Randall Reed

Bibliography. D. Berg and M. Howe (eds.), *Mobile DNA,* 1989; N. Federoff, Controlling elements in maize, *Sci. Amer.*, 250(6):84–91, 1984; N. Kleckner, Transposable elements in prokaryotes, *Annu. Rev. Genet.*, 15:341–404, 1981; M. E. Lambert, J. F. McDonald, and I. B. Weinstein, *Eukaryotic Transposable Elements as Mutagenic Agents,* 1988; O. Nelson (ed.), *Plant Transposable Elements,* 1988; J. A. Shapiro (ed.), *Mobile Genetic Elements*, 1983.

Transuranium elements

Those synthetic elements with atomic numbers larger than that of uranium (atomic number 92). They are the members of the actinide series, from neptunium (atomic number 93) through lawrencium (atomic number 103), and the transactinide elements (with atomic numbers higher than 103). Of these elements, plutonium, an explosive ingredient for nuclear weapons and a fuel for nuclear power because it is fissionable, has been prepared on the largest (ton) scale, while some of the others have been produced in kilograms (neptunium, americium, curium) and in much smaller quantities (berkelium, californium, and einsteinium).

The concept of atomic weight as applied to naturally occurring elements is not applicable to the transuranium elements, since the isotopic composition of any given sample depends on its source. In most cases the use of the mass number of the longest-lived isotope in combination with an evaluation of its

availability has been adequate. Good choices at present are neptunium, 237; plutonium, 242; americium, 243; curium, 248; berkelium, 249; californium, 249; einsteinium, 254; fermium, 257; mendelevium, 258; nobelium, 259; lawrencium, 260; rutherfordium, 261; dubnium, 262; seaborgium, 266; bohrium 267; and hassium 269.

The actinide elements are chemically similar and have a strong chemical resemblance to the lanthanide, or rare-earth, elements (atomic numbers 57–71). The transactinide elements, with atomic numbers 104 to 118, should be placed in an expanded periodic table under the row of elements beginning with hafnium, number 72, and ending with radon, number 86. This arrangement allows prediction of the chemical properties of these elements and suggests that they will have an element-by-element chemical analogy with the elements that appear immediately above them in the periodic table. However, deviations from this analogy are expected and are observed in specific detailed chemical properties of transactinide elements. *See* ACTINIDE ELEMENTS; PERIODIC TABLE; RARE-EARTH ELEMENTS.

The transuranium elements up to and including fermium (atomic number 100) are produced in the largest quantity through the successive capture of neutrons in nuclear reactors. The yield decreases with increasing atomic number, and the heaviest to be produced in weighable quantity is einsteinium (number 99). Many additional isotopes are produced by bombardment of heavy target isotopes with charged atomic projectiles in accelerators; beyond fermium all elements are produced by bombardment with heavy ions. Brief descriptions of transuranium elements follow. They are listed according to increasing atomic number.

Neptunium. Neptunium (Np, atomic number 93, named after the planet Neptune) was the first transuranium element discovered. In 1940 E. M. McMillan and P. H. Abelson at the University of California, Berkeley, identified the isotope ^{239}Np (half-life 2.35 days), which was produced by the bombardment of uranium with neutrons according to reaction (1).

$$^{238}\mathrm{U}(n,\gamma)^{239}\mathrm{U} \longrightarrow {}^{239}\mathrm{Np} \qquad (1)$$

The element as ^{237}Np was first isolated as a pure compound, the oxide, in 1944 by L. G. Magnusson and T. J. La Chapelle. Neptunium in trace amounts is found in nature, and is produced in nuclear reactions in uranium ores caused by the neutrons present. Kilogram and larger quantities of ^{237}Np (half-life 2.14×10^6 years), used for chemical and physical investigations, are produced as a by-product of the production of plutonium in nuclear reactors. Isotopes from mass number 227 to 244 have been synthesized by various nuclear reactions. *See* NUCLEAR REACTION.

Neptunium displays five oxidation states in aqueous solution: Np^{3+} (pale purple), Np^{4+} (yellow-green), NpO_2^+ (green-blue), NpO_2^{2+} (pink), and NpO_5^{3-} (green). The ion NpO_2^+, unlike corresponding ions of uranium, plutonium, and americium, can exist in aqueous solution at moderately high concentrations. The element forms tri- and tetrahalides such as NpF_3, NpF_4, $NpCl_3$, $NpCl_4$, $NpBr_3$, NpI_3, as well as NpF_6 and oxides of various compositions such as those found in the uranium-oxygen system, including Np_3O_8 and NpO_2.

Neptunium metal has a silvery appearance, is chemically reactive, and melts at 637°C (1179°F); it has at least three crystalline forms between room temperature and its melting point. *See* NEPTUNIUM.

Plutonium. Plutonium (Pu, atomic number 94, named after the planet Pluto) in the form of ^{238}Pu was discovered in late 1940 and early 1941 by G. T. Seaborg, McMillan, J. W. Kennedy, and A. C. Wahl at the University of California, Berkeley. The element was produced in the bombardment of uranium with deuterons according to reaction (2).

$$^{238}\mathrm{U}(d,2n)^{238}\mathrm{Np} \xrightarrow[2.1\,\mathrm{days}]{\beta^-} {}^{238}\mathrm{Pu} \qquad (2)$$

The important isotope ^{239}Pu was discovered by Kennedy, Seaborg, E. Segrè, and Wahl in 1941. Because of its property of being fissionable with neutrons, plutonium-239 (half-life 24,400 years) is used as the explosive ingredient in nuclear weapons and is a key material in the development of nuclear energy for industrial purposes. 1 lb (0.45 kg) of plutonium produced is equivalent to about 10^7 kWh of heat energy; plutonium is produced in ton quantities in nuclear reactors. The alpha radioactivity and physiological behavior of this isotope make it one of the most dangerous poisons known, but means for handling it safely have been devised. Plutonium as ^{239}Pu was first isolated as a pure compound, the fluoride, in 1942 by B. B. Cunningham and L. B. Werner. Minute amounts of plutonium formed in much the same way as naturally occurring neptunium are present in nature. Much smaller quantities of the longer-lived isotope ^{244}Pu (half-life 8.3×10^7 years) have been found in nature; in this case it may represent the small fraction remaining from a primordial source or it may be caused by cosmic rays. Isotopes of mass number 232–246 are known. The longer-lived isotopes ^{242}Pu (half-life 390,000 years) and ^{244}Pu, produced in nuclear reactors, are more suitable than ^{239}Pu for chemical and physical investigation because of their longer half-lives and lower specific activities.

Plutonium has five oxidation states in aqueous solution: Pu^{3+} (blue to violet), Pu^{4+} (yellow-brown), PuO_2^+ (pink), PuO_2^{2+} (pink-orange), and PuO_5^{3-} (blue-green). The ions Pu^{4+} and PuO_2^+ undergo extensive disproportionation to the ions of higher and lower oxidation states. Four oxidation states (III, IV, V, and VI) can exist simultaneously at appreciable concentrations in equilibrium with each other, an unusual situation that leads to complicated solution phenomena.

Plutonium forms binary compounds with oxygen (PuO, PuO_2, and intermediate oxides of variable composition); with the halogens (PuF_3, PuF_4, PuF_6, $PuCl_3$, $PuBr_3$, PuI_3); with carbon, nitrogen, and silicon

(including PuC, PuN, $PuSi_2$); in addition, oxyhalides are well known (PuOCl, PuOBr, PuOI).

The metal is silvery in appearance, is chemically reactive, melts at 640°C (1184°F), and has six crystalline modifications between room temperature and its melting point. *See* PLUTONIUM.

Americium. Americium (Am, atomic number 95, named after the Americas) was the fourth transuranium element discovered. The element as ^{241}Am (half-life 433 years) was produced by the intense neutron bombardment of plutonium and was identified by Seaborg, R. A. James, L. O. Morgan, and A. Ghiorso in late 1944 and early 1945 at the wartime Metallurgical Laboratory at the University of Chicago. By using the isotope ^{241}Am, the element was first isolated as a pure compound, the hydroxide, in 1945 by B. B. Cunningham. Isotopes of mass numbers 237–247 have been prepared. Kilogram quantities of ^{241}Am are being produced in nuclear reactors. The less radioactive isotope ^{243}Am (half-life 7400 years), also produced in nuclear reactors, is more suitable for use in chemical and physical investigation.

Americium exists in four oxidation states in aqueous solution: Am^{3+} (light salmon), AmO_2^+ (light tan), AmO_2^{2+} (light tan), and a fluoride complex of the IV state (pink). The trivalent state is highly stable and difficult to oxidize. AmO_2^+, like plutonium, is unstable with respect to disproportionation into Am^{3+} and AmO_2^{2+}. The ion Am^{4+} may be stabilized in solution only in the presence of very high concentrations of fluoride ion, and tetravalent solid compounds are well known. Divalent americium has been prepared in solid compounds; this is consistent with the presence of seven 5*f* electrons in americium (enhanced stability of half-filled 5*f* electron shell) and is similar to the analogous lanthanide, europium, which can be reduced to the divalent state.

Americium dioxide, AmO_2, is the important oxide; Am_2O_3 and, as with previous actinide elements, oxides of variable composition between $AmO_{1.5}$ and AmO_2 are known. The halides AmF_2 (in CaF_2), AmF_3, AmF_4, $AmCl_2$ (in $SrCl_2$), $AmCl_2$, $AmBr_3$, AmI_2, and AmI_3 have also been prepared.

Metallic americium is silvery-white in appearance, is chemically reactive, and has a melting point of 1176°C (2149°F). It has two crystalline forms between room temperature and its melting point. *See* AMERICIUM.

Curium. The third transuranium element to be discovered, curium (Cm, atomic number 96, named after Pierre and Marie Curie), as the isotope ^{242}Cm, was identified by Seaborg, James, and Ghiorso in 1944 at the wartime Metallurgical Laboratory of the University of Chicago. This was produced by the helium-ion bombardment of ^{239}Pu in the University of California 60-in. (152-cm) cyclotron. Curium was first isolated, using the isotope ^{242}Cm, in the form of a pure compound, the hydroxide, in 1947 by L. B. Werner and I. Perlman. Isotopes of mass number 238–251 are known. Chemical investigations have been performed using ^{242}Cm (half-life 163 days) and ^{244}Cm (half-life 18 years), but the higher-mass isotopes ^{247}Cm and ^{248}Cm with much longer half-lives (1.6×10^7 and 3.5×10^5 years, respectively) are more satisfactory for this purpose; these are all produced by neutron irradiation in nuclear reactors.

Curium exists solely as Cm^{3+} (colorless to yellow) in the uncomplexed state in aqueous solution. This behavior is related to its position as the element in the actinide series in which the 5*f* electron shell is half filled; that is, it has the especially stable electronic configuration $5f^7$, analogous to its lanthanide homolog, gadolinium. A curium IV fluoride complex ion exists in aqueous solution. Solid compounds include Cm_2O_3, CmO_2 (and oxides of intermediate composition), CmF_3, CmF_4, $CmCl_3$, $CmBr_3$, and CmI_3.

The metal is silvery and shiny in appearance, is chemically reactive, melts at 1340°C (2444°F), and resembles americium metal in its two crystal modifications. *See* CURIUM.

Berkelium. Berkelium (Bk, atomic number 97, named after Berkeley, California) was produced and identified by S. G. Thompson, Ghiorso, and Seaborg in late 1949 at the University of California, Berkeley, and was the fifth transuranium element discovered. The isotope ^{243}Bk (half-life 4.6 h) was synthesized by helium-ion bombardment of ^{241}Am. The first isolation of berkelium in weighable amount, as ^{249}Bk (half-life 314 days), produced by neutron irradiation, was accomplished in 1958 by Thompson and Cunningham; this isotope, produced in nuclear reactors, is used in the chemical and physical investigation of berkelium. Isotopes of mass number 242–251 are known.

Berkelium exhibits two ionic oxidation states in aqueous solution: Bk^{3+} (yellow-green) and somewhat unstable Bk^{4+} (yellow), as might be expected by analogy with its rare-earth homolog, terbium. Solid compounds include Bk_2O_3, BkO_2 (and oxides of intermediate composition), BkF_3, BkF_4, $BkCl_3$, $BkBr_3$, and BkI_3.

Berkelium metal is chemically reactive, exists in two crystal structure modifications, and melts at 986°C (1807°F). *See* BERKELIUM.

Californium. The sixth transuranium element to be discovered, californium (Cf, atomic number 98, named after the state and University of California, Berkeley), in the form of the isotope ^{245}Cf (half-life 44 min), was first prepared by the helium-ion bombardment of microgram quantities of ^{242}Cm by Thompson, K. Street, Jr., Ghiorso, and Seaborg at Berkeley early in 1950. Cunningham and Thompson, at Berkeley, isolated californium in weighable quantities for the first time in 1958 using a mixture of the isotopes ^{249}Cf, ^{250}Cf, ^{251}Cf, and ^{252}Cf, produced by neutron irradiation. Isotopes of mass number 239–256 are known. The best isotope for the investigation of the chemical and physical properties of californium is ^{249}Cf (half-life 350 years), produced in pure form as the beta-particle decay product of ^{249}Bk.

Californium exists mainly as Cf^{3+} in aqueous solution (emerald green), but it is the first of the actinide elements in the second half of the series to exhibit the II state, which becomes progressively

more stable on proceeding through the heavier members of the series. It also exhibits the IV oxidation state in CfF_4 and CfO_2, which can be prepared under somewhat intensive oxidizing conditions. Solid compounds also include Cf_2O_3 (and higher intermediate oxides), CfF_3, $CfCl_3$, $CfBr_2$, $CfBr_3$, CfI_2, and CfI_3.

Californium metal is chemically reactive, is quite volatile, and can be distilled at temperature ranges of 1100–1200°C (2010–2190°F). It appears to exist in three different crystalline modifications between room temperature and its melting point, 900°C (1652°F). *See* CALIFORNIUM.

Einsteinium. The seventh transuranium element to be discovered, einsteinium (Es, atomic number 99, named after Albert Einstein), was found by Ghiorso and coworkers in the debris from the "Mike" thermonuclear explosion staged by the Los Alamos Scientific Laboratory in November 1952. Very heavy uranium isotopes were formed by the action of the intense neutron flux on the uranium in the device, and these decayed into isotopes of elements 99, 100, and other transuranium elements of lower atomic number. Chemical investigation of the debris in late 1952 by workers at the University of California Radiation Laboratory, Argonne National Laboratory, and Los Alamos Scientific Laboratory revealed the presence of element 99 as the isotope ^{253}Es. Einsteinium was isolated in a macroscopic (weighable) quantity for the first time in 1961 by Cunningham, J. C. Wallman, L. Phillips, and R. C. Gatti at Berkeley; they used the isotope ^{253}Es, produced in nuclear reactors, working with only a few hundredths of a microgram. The macroscopic property that they determined in this case was the magnetic susceptibility. Isotopes of mass number 243–256 have been synthesized. Einsteinium is the heaviest transuranium element to be isolated in weighable form. Most of the investigations have used the short-lived ^{253}Es (half-life 20.5 days) because of its greater availability, but the use of ^{254}Es (half-life 276 days) will increase as it becomes more available as the result of production in nuclear reactors.

Einsteinium exists in normal aqueous solution essentially as Es^{3+} (green), although Es^{2+} can be produced under strong reducing conditions. Solid compounds such as Es_2O_3, $EsCl_3$, EsOCl, $EsBr_2$, $EsBr_3$, EsI_2, and EsI_3 have been made.

Einsteinium metal is chemically reactive, is quite volatile, and melts at 860°C (1580°F); one crystal structure is known. *See* EINSTEINIUM.

Fermium. Fermium (Fm, atomic number 100, named after Enrico Fermi), the eighth transuranium element discovered, was isolated as the isotope ^{255}Fm (half-life 20 h) from the heavy elements formed in the "Mike" thermonuclear explosion. The element was discovered in early 1953 by Ghiorso and coworkers during the same investigation which resulted in the discovery of element 99. Fermium isotopes of mass number 242–259 have been prepared.

No isotope of fermium has yet been isolated in weighable amounts, and thus all the investigations of this element have been done with tracer quantities. The longest-lived isotope is ^{257}Fm (half-life about 100 days), whose production in high-neutron-flux reactors is extremely limited because of the very long sequence of neutron-capture reactions that is required.

Despite its very limited availability, fermium, in the form of the 3.24-h ^{254}Fm isotope, has been identified in the "metallic" zero-valent state in an atomic-beam magnetic resonance experiment. This established the electron structure of elemental fermium in the ground state as $5f^{12}7s^2$ (beyond the radon structure).

Fermium exists in normal aqueous solution almost exclusively as Fm^{3+}, but strong reducing conditions can produce Fm^{2+}, which has greater stability than Es^{2+} and less stability than Md^{2+}. *See* FERMIUM.

Mendelevium. Mendelevium (Md, atomic number 101, named after Dmitri Mendeleev), the ninth transuranium element discovered, was identified by Ghiorso, B. G. Harvey, G. R. Choppin, Thompson, and Seaborg at the University of California, Berkeley, in 1955. The element as ^{256}Md (half-life 1.5 h) was produced by the bombardment of extremely small amounts (approximately 10^9 atoms) of ^{253}Es with helium ions in the 60-in. (152-cm) cyclotron. The first identification of mendelevium was notable in that only one or two atoms per experiment were produced. (This served as the prototype for the discovery of all heavier transuranium elements, which have been first synthesized and identified on a one-atom-at-a-time basis.) Isotopes of mass numbers 247–259 are known. Although the isotope ^{258}Md (half-life 56 days) is sufficiently long-lived, it cannot be produced in nuclear reactors, and hence it will be very difficult and perhaps impossible to isolate it in weighable amount.

The chemical properties have been investigated on the tracer scale, and the element behaves in aqueous solution as a typical tripositive actinide ion; it can be reduced to the II state with moderately strong reducing agents. *See* MENDELEVIUM.

Nobelium. The discovery of nobelium (No, atomic number 102, named after Alfred Nobel), the tenth transuranium element to be discovered, has a complicated history. For the first time scientists from countries other than the United States embarked on serious efforts to compete in this field. The reported discovery of element 102 in 1957 by an international group of scientists working at the Nobel Institute for Physics in Stockholm, who suggested the name nobelium, has never been confirmed and must be considered to be erroneous. Working at the Kurchatov Institute of Atomic Energy in Moscow, G. N. Flerov and coworkers in 1958 reported a radioactivity that they thought might be attributed to element 102, but a wide range of half-lives was suggested and no chemistry was performed. As the result of more definitive work performed in 1958, Ghiorso, T. Sikkeland, J. R. Walton, and Seaborg reported an isotope of the element, produced by bombarding a mixture of curium isotopes with ^{12}C ions in the then-new Heavy Ion Linear Accelerator (HILAC) at Berkeley. They described a novel "double recoil" technique that permitted identification by chemical means, one atom at a time, of any daughter isotope of element 102

that might have been formed. The isotope ^{250}Fm was identified conclusively by this means, indicating that its parent should be the isotope of element 102 with mass number 254 produced by the reaction of ^{12}C ions with ^{246}Cm. However, another isotope of element 102, with half-life 3 s, also observed indirectly in 1958, and whose alpha particles were shown to have an energy of 8.3 MeV by Ghiorso and coworkers in 1959, was shown later by Flerov and coworkers (working at the Dubna Laboratory near Moscow) to be of an isotope of element 102 with mass number 252 rather than 254; in other words, two isotopes of element 102 were discovered by the Berkeley group in 1958, but the correct mass number assignments were not made until later. On the basis that they identified the atomic number correctly, the Berkeley scientists probably have the best claim to the discovery of element 102; they suggest the retention of nobelium as the name for this element.

All known isotopes (mass numbers 250–259) of nobelium are short-lived and are produced by the bombardment of lighter elements with charged particles (heavy ions); the longest-lived is ^{259}No with a half-life of 58 min. All of the chemical investigations have been, and presumably must continue to be, done on the tracer scale. These have demonstrated the existence of No^{3+} and No^{2+} in aqueous solutions, with the latter much more stable than the former. The stability of No^{2+} is consistent with the expected presence of the completed shell of fourteen 5*f* electrons in this ion. *See* NOBELIUM.

Lawrencium. Lawrencium (Lr, atomic number 103, named after Ernest O. Lawrence) was discovered in 1961 by Ghiorso, Sikkeland, A. E. Larsh, and R. M. Latimer using the HILAC at the University of California, Berkeley. A few micrograms of a mixture of ^{249}Cf, ^{250}Cf, ^{251}Cf, and ^{252}Cf (produced in a nuclear reactor) were bombarded with ^{10}B and ^{11}B ions to produce single atoms of an isotope of element 103 with a half-life measured as 8 s and decaying by the emission of alpha particles of 8.6 MeV energy. Ghiorso and coworkers suggested at that time that this radioactivity might be assigned the mass number 257. G. N. Flerov and coworkers have disputed this discovery on the basis that their later work suggests a greatly different half-life for the isotope with the mass number 257. Subsequent work by Ghiorso and coworkers proves that the correct assignment of mass number to the isotope discovered in 1961 is 258, and this later work gives 4 s as a better value for the half-life.

All known isotopes of lawrencium (mass numbers 253–260) are short-lived and are produced by bombardment of lighter elements with charged particles (heavy ions); chemical investigations have been, and presumably must be, performed on the tracer scale. Work with ^{260}Lr (half-life 3 min) has demonstrated that the normal oxidation state in aqueous solution is the III state, corresponding to the ion Lr^{3+}, as would be expected for the last member of the actinide series. *See* LAWRENCIUM.

Rutherfordium. Rutherfordium (Rf, atomic number 104, named after Lord Rutherford), the first transactinide element to be discovered, was probably first identified in a definitive manner by Ghiorso, M. Nurmia, J. Harris, K. Eskola, and P. Eskola in 1969 at Berkeley. Flerov and coworkers have suggested the name kurchatovium (named after Igor Kurchatov with symbol Ku) on the basis of an earlier claim to the discovery of this element. In 1964 they bombarded ^{242}Pu with ^{22}Ne ions in their cyclotron at the Joint Institute for Nuclear Research in Dubna and reported the production of an isotope, suggested to be ^{260}Ku, which was held to decay by spontaneous fission with a half-life of 0.3 s. After finding it impossible to confirm this observation, Ghiorso and coworkers reported definitive proof of the production of alpha-particle-emitting ^{257}Rf and ^{259}Rf (half-lives 4.5 and 3 s, respectively), demonstrated by the identification of the previously known ^{253}No and ^{255}No as decay products, by means of the bombardment of ^{249}Cf with ^{12}C and ^{13}C ions in the Berkeley HILAC.

All known isotopes of rutherfordium (mass numbers 253–262) are short-lived and are produced by bombardment of lighter elements with charged heavy-ion particles. The isotope ^{261}Rf (half-life 78 s) has made it possible, by means of rapid chemical experiments, to demonstrate that the normal oxidation state of rutherfordium in aqueous solution is the IV state corresponding to the ion Rf^{4+}. This is consistent with expectations for this first "transactinide" element which should be a homolog of hafnium, an element that is exclusively tetrapositive in aqueous solution. Gas chromatographic studies of volatile halides and halides oxides of rutherfordium demonstrate characteristics as expected for a group-4 element in the periodic table. However, some detailed chemical properties of rutherfordium compounds, studied in the aqueous phase and the gas phase, resemble more the behavior of its lighter homolog zirconium (atomic number 40) than hafnium (atomic number 72). *See* RUTHERFORDIUM.

Dubnium. Dubnium (Db, atomic number 105, named after the Dubna Laboratory), the second transactinide element to be discovered, was probably first identified in a definitive manner in 1970 by Ghiorso, Nurmia, K. Eskola, Harris, and P. Eskola at Berkeley. They reported the production of alpha-particle-emitting ^{260}Db (half-life 1.6 s), demonstrated through the identification of the previously known ^{256}Lr as the decay product, by bombardment of ^{249}Cf with ^{15}N ions in the Berkeley HILAC. Again the Berkeley claim to discovery is disputed by Flerov and coworkers, who earlier in 1970 reported the discovery of an isotope thought to be dubnium, decaying by the less definitive process of spontaneous fission, produced by the bombardment of ^{243}Am with ^{22}Ne ions in the Dubna cyclotron; in later work Flerov and coworkers may have also observed the alpha-particle-emitting isotope of dubnium reported by Ghiorso and coworkers.

The known isotopes of dubnium (mass numbers 256–263) are short-lived and are produced by bombardment of lighter elements with charged heavy-ion particles. Using rapid chemical techniques and the isotope ^{262}Db (half-life 40 s), it is possible to

study the chemical properties of dubnium. The results show that dubnium exhibits the V oxidation state like its homolog tantalum. A number of chemical studies demonstrate that, in specific chemical environments, dubnium behaves more like its lighter homolog niobium or sometimes like the pentavalent actinide element protactinium (atomic number 91). *See* DUBNIUM.

Seaborgium. The discovery of seaborgium (Sg, atomic number 106, named after Glenn T. Seaborg) took place in 1974 simultaneously as the result of experiments by Ghiorso and coworkers at Berkeley and Flerov, Y. T. Oganessian, and coworkers at Dubna. The Ghiorso group used the SuperHILAC (the rebuilt HILAC) to bombard a target of californium (the isotope ^{249}Cf) with ^{18}O ions. This resulted in the production and positive identification of the alpha-particle-emitting isotope ^{263}Sg, which decays with a half-life of 0.9 ± 0.2 s by the emission of alpha particles of a principal energy of 9.06 MeV. The definitive identification consisted of the establishment of the genetic link between the seaborgium alpha-particle-emitting isotope (^{263}Sg) and previously identified daughter (^{259}Rf) and granddaughter (^{255}No) nuclides, that is, the demonstration of the decay sequence: A total of 73 ^{263}Sg alpha particles and approximately the expected corresponding number of ^{259}Rf daughter and ^{255}No granddaughter alpha particles were recorded.

The Dubna group chose lead (atomic number 82) as their target because, they believed, its closed shells of protons and neutrons and consequent small relative mass leads to minimum excitation energy for the compound nucleus and therefore an enhancement in the cross section for the production of the desired product nuclide. They bombarded ^{207}Pb and ^{208}Pb with ^{54}Cr ions (atomic number 24) in their cyclotron to find a product that decays by the spontaneous fission mechanism (a total of 51 events), with the very short half-life of 7 ms, which they assign to the isotope ^{259}Sg. Later work at Dubna and the Gesellschaft für Schwerionenforschung (GSI) laboratory in Darmstadt, Germany, has shown that this assignment is not correct. ^{259}Sg is an alpha emitter of 0.48 s. In 1984 the isotopes ^{261}Sg and ^{260}Sg were discovered at GSI, Germany.

Long-lived isotopes ^{266}Sg (21 s) and ^{265}Sg (22 s) were discovered in 1994 at Dubna bombarding ^{248}Cm with ^{22}Ne. In 2000, spontaneous fissioning isotopes ^{258}Sg (2.9 ms) and ^{262}Sg (6.9 ms) were synthesized at GSI using ^{290}Bi- and ^{207}Pb-based reactions. Thirty years after discovery of the element seaborgium, 8 isotopes are known.

Fast chemical separations, performed with the Automated Rapid Chemistry Apparatus (ARCA) and On-Line Gas Chromatography Apparatus (OLGA) on a one-atom-at-a-time scale of ^{265}Sg, allowed investigating seaborgium in aqueous solution and in the gas phase. A series of experiments were done in international collaborations at GSI, Darmstadt. As expected from its projected position in the periodic table, seaborgium shows chemical properties similar to those extrapolated from its lighter group-6 homologs, molybdenum (atomic number 42) and tungsten (atomic number 74). Seaborgium is the heaviest element that has been studied in aqueous solution. *See* SEABORGIUM.

Bohrium. Bohrium (Bh, atomic number 107, named after Niels Bohr) was synthesized and identified by G. Münzenberg and coworkers at GSI-Darmstadt. The element has a long half-life in the millisecond range, indicating an unexpected relative stability in this region of high atomic numbers. In the discovery experiment, a target of ^{209}Bi [located in the region of closed shells ($N = 126$)] was used, which when bombarded with heavy ions, led to a compound nucleus of minimum excitation energy ($\sim$15 MeV), allowing for a cooling of the nucleus by emission of a single neutron. The element was identified by the alpha-particle-emitting isotope ^{262}Bh, and the genetic links with its known alpha-particle-emitting descendants was established, as was done by Ghiorso and coworkers in their discovery of seaborgium. In 1981, Münzenberg and coworkers observed six atoms of 5 ms (the time interval for its decay) ^{262}Bh, produced by the ^{209}Bi (^{54}Cr,n) reaction. *See* NUCLEAR STRUCTURE.

The experiment was repeated, and in 1988 about 40 decay chains corroborated the discovery of bohrium. The isotope ^{261}Bh was discovered in the same reaction by the 2-neutron emission channel.

Today the isotope ^{264}Bh, a member of the decay-chain of roentgenium, is known and two isotopes, ^{266}Bh and ^{267}Bh, synthesized in the reaction of ^{22}Ne with ^{249}Bk, have been discovered. ^{265}Bh with a half-life of 0.94 s was synthesized at the Institute of Modern Physics, Lanzhou, China, in 2004. All these isotopes manifest the trend towards larger half-lives in the region of seconds.

The one and so far only chemical separation and characterization of bohrium was done at the Paul-Scherrer Institute (PSI), Villigen, Switzerland, during gas chromatography of a chloride oxide compound of ^{267}Bh. As expected, bohrium showed chemical properties similar to those extrapolated from its lighter group-7 homologs, technetium (atomic number 43) and rhenium (atomic number 75). *See* BOHRIUM.

Hassium. Hassium (Hs, atomic number 108, named after the German state of Hessen from the Latin word Hassias) was synthesized in 1984 at GSI, Darmstadt, Hessen. A target nucleus of ^{208}Pb, an isotope with two closed nucleon shells ($Z = 82$, $N = 126$) was fused with ^{58}Fe to synthesize ^{265}Hs; see also bohrium for aspects of the hassium production mechanism. The element was identified by its links to known descendants. Three atoms of 1.8 ms ^{265}Hs were produced by the ^{208}Pb(^{58}Fe,n) reaction. The reaction was confirmed in many laboratories and serves as a calibration for Pb- and Bi-based reactions. Later the isotope ^{264}Hs was discovered by the same group in the reaction ^{207}Pb(^{58}Fe,n). It links the elements with even proton numbers via a bridge of an α-decay at ^{256}Rf to lighter elements and establishes a connection of absolute binding energies up to element hassium.

The isotope ^{267}Hs was discovered at Dubna in 1992. It was observed in a 5*n*-channel in the fusion of ^{34}S and ^{238}U. This isotope ^{269}Hs was confirmed at GSI, being found as a descendant in the α-decay chain of ^{271}Ds. The isotope ^{269}Hs was observed first in 1996 at GSI in a decay chain of ^{277}112 and has been reproduced since in the reaction ^{248}Cm(^{26}Mg,5*n*) by experiments at GSI, observing its decay after a chemical hassium separation. In the reaction of magnesium-26 with curium-248, the isotope ^{270}Hs was synthesized. This is the first isotope with the closed nuclear shell $N = 162$. With increasing neutron number, a trend toward larger half-lives in the few-seconds range is observed also for hassium isotopes.

Since 2001, studies of the formation and the behavior of hassium tetroxide were done one-atom-at-a-time with the isotope ^{269}Hs in three international collaborations at GSI. All the experiments showed chemical properties of hassium similar to those of osmium (atomic number 76), the lighter homolog in group-8 of the periodic table. *See* HASSIUM.

Meitnerium. Meitnerium (Mt, atomic number 109, named after Lise Meitner) was synthesized in 1982 at GSI. (For its production see also Bohrium above.) A target nucleus of ^{209}Bi, as done before for bohrium, was fused with ^{58}Fe to synthesize ^{266}Mt. Again the element was identified by its genetic links to known descendants. In 1982, one atom of 1.7 ms ^{266}Mt was produced by the ^{209}Bi(^{58}Fe,*n*) reaction. Since then, the discovery has been confirmed by about 10 more decay chains.

The isotopes ^{268}Mt and ^{270}Mt were observed as descendants in decay chains from heavier elements. ^{268}Mt was discovered at GSI, in 1994 in the discovery experiment of the element roentgenium. ^{270}Mt was reported from RIKEN, Japan, in 2004, together with the discovery of element 113 in the reaction ^{209}Bi(^{70}Zn,*n*). Half-lives stayed in the range below 1 second and the production cross section of 5×10^{-38} cm^2 is very small (the smallest cross section known, to date).

Meitnerium is expected to have chemical properties similar to those of iridium, its lighter homolog in group 9 of the periodic table. No chemical experiments have been done on meitnerium. *See* MEITNERIUM.

Darmstadtium. Darmstadtium (Ds, atomic number 110, named after Darmstadt, Germany, the location of the GSI laboratory) was synthesized in 1994 at GSI.

Darmstadtium should be a heavy homolog of the elements platinum, palladium, and nickel. It is the eighth element in the 6*d* shell.

Research for this element began in 1985. Experiments at Dubna, Russia; at GSI, and at Lawrence Berkeley Laboratory, Berkeley, California, failed to provide reliable evidence for a successful synthesis. However, at GSI on November 9, 1994, a decay chain was observed that proved the existence of the isotope ^{269}Ds (the isotope of darmstadtium with mass number 269). The isotopes were produced in a fusion reaction of a nickel-62 projectile with a lead-208 target nucleus. The fused system, with an excitation energy of 13 MeV, cooled down by emitting one neutron and forming ^{269}Ds, which by sequential alpha decays transformed to ^{265}Hs, ^{261}Sg, ^{257}Rf, and ^{253}No (nobelium-253). All these daughter isotopes were already known, and four decay chains observed in the following 12 days corroborated without any doubt the discovery of the element. **Illustration** *a*

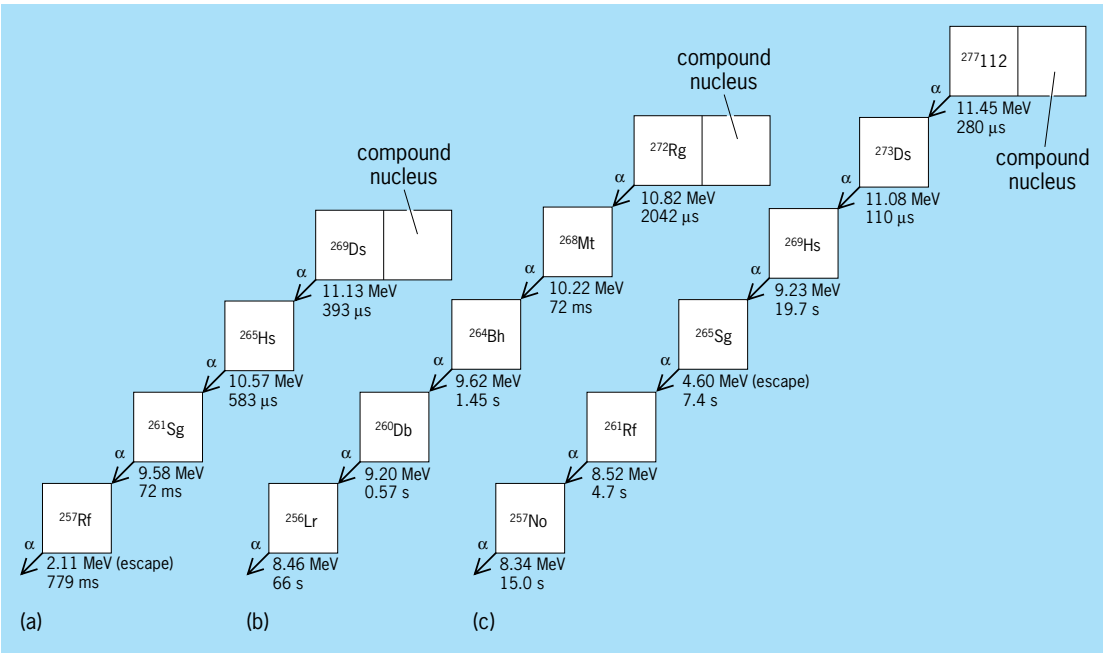

Decay chains that document the discoveries of new elements. The sequence of alpha decays is shown for each element. Numbers below boxes are alpha energies and correlation times. (*a*) Darmstadtium, produced in the reaction ^{62}Ni + ^{208}Pb → ^{269}Ds + 1*n*. (*b*) Roentgenium, produced in the reaction ^{64}Ni + ^{209}Bi → ^{272}Rg + 1*n*. (*c*) Element 112, produced in the reaction ^{70}Zn + ^{208}Pb → ^{277}112 + 1*n*.

shows the first decay chain observed, which ended in ^{257}Rf. The isotope ^{269}Ds has a half-life of 0.2 ms and is produced with a cross section of about 3×10^{-36} cm^2.

A second isotope, ^{271}Ds, was produced in a subsequent 12-day experiment by fusion of nickel-64 and lead-208. Nine atoms, with an excitation energy of 12 MeV, were produced. They were transformed by sequential alpha decay to the known isotopes ^{267}Hs, ^{263}Sg, ^{259}Rf, and ^{255}No (nobelium-255). The half-life of ^{271}Ds is 1.1 ms, and its production cross section amounts to 1.5×10^{-35} cm^2.

The methods used to produce element Ds were the same as those already used to synthesize the three preceding elements, Bh, Hs, and Mt. Improved beam intensity and quality, improvement of the detection efficiency, and a new detector system allowing nearly complete chain reconstruction made possible the discovery after an extensive search for the optimum bombarding energy. The total sensitivity for finding a new species was increased by a factor of 20.

Two additional isotopes, ^{270}Ds and ^{273}Ds, are known. ^{273}Ds was discovered as a descendant in the α-decay chain of ^{277}112 in 1996. In the fusion of ^{207}Pb and ^{64}Ni, the isotope ^{270}Ds with a half-life of 0.1 ms was synthesized at GSI in 2001. *See* DARMSTADTIUM.

Roentgenium. Roentgenium (Rg, atomic number 111, named after W. Roentgen) was synthesized in 1994 at GSI. Roentgenium should be a homolog of the elements gold, silver, and copper. It is the ninth element in the 6*d* shell.

The element was discovered on December 17, 1994, by detection of the isotope ^{272}Rg, which was produced by fusion of a nickel-64 projectile and a bismuth-209 target nucleus after the fused system was cooled by emission of one neutron. The optimum bombarding energy for producing ^{272}Rg corresponds to an excitation energy of 15 MeV for the fused system. Sequential alpha decays to ^{268}Mt, ^{264}Bh, ^{260}Db, and ^{256}Lr (lawrencium-256) allowed identification from the known decay properties of ^{260}Db and ^{256}Lr. In the decay chain in illus. *b*, the first three members are new isotopes. The isotope ^{272}Rg has a half-life of 1.5 ms, and is produced with a cross section of 3.5×10^{-36} cm^2. Altogether, three chains were observed during the 17 days of irradiation. The methods used to produce roentgenium were the same as those used in the discovery of darmstadtium. *See* ROENTGENIUM.

Element 112. Element 112 should be a heavy homolog of the elements mercury, cadmium, and zinc. It is expected to be the last element in the 6*d* shell. The element was discovered on February 9, 1996, at GSI by detection of the isotope ^{277}112, which was produced by fusion of a zinc-70 projectile and a lead-208 target nucleus following the cooling down of the fused system by emission of a single neutron. The fused system was observed at an excitation energy of 12 MeV. Sequential alpha decays to ^{273}Ds, ^{269}Hs, ^{265}Sg, ^{261}Rf, and ^{257}No (nobelium-257) allowed unambiguous identification by using the known decay properties of the last three members of the chain. In the decay chain in illus. *c*, the first three members are new isotopes. The isotope ^{277}112 has a half-life of 0.24 ms, and it is produced with a cross section of 0.5×10^{-36} cm^2. The new isotopes of Ds and Hs are of special interest. Their half-lives and alpha energies are very different, as is characteristic of a closed-shell crossing. At the neutron number $N = 162$, a closed shell was theoretically predicted, and this closed shell is verified in the decay chain observed. The isotope ^{269}Hs has a half-life of 9 s, which is long enough to allow studies on the chemistry of this element. The methods used to produce element 112 were the same as those used for the two preceding elements, Ds and Rg. The decay chain of the new element was observed in an irradiation time of about 3 weeks. The cross section measured is the smallest observed in the production of heavy elements.

The crossing of the neutron shell at $N = 162$ is an important achievement in the field of research on superheavy elements. The stabilization of superheavy elements is based on high fission barriers, which are due to corrections in the binding energies found near closed shells. The shell at $N = 162$ is the first such shell predicted, and is now verified. Next in line are the predicted shells at proton number $Z = 114$ and neutron number $N = 184$. *See* ELEMENT 112.

Peter J. Armbruster; M. Schädel; Glenn T. Seaborg

Bibliography. G. R. Choppin and J. Rydberg, *Nuclear Chemistry: Theory and Application*, 1980; R. Eichler et al., Chemical characterization of bohrium (element 107), *Nature*, 407:63–65, 2000; Ch. E. Düllmann et al., Chemical investigation of hassium (element 108), *Nature*, 418:859–862, 2002; V. I. Goldanski and S. M. Polikanov, *The Transuranium Elements*, 1973; S. Hofmann et al., The new element 111, *Z. Phys. A*, 350:281–282, 1995; S. Hofmann et al., The new element 112, *Z. Phys. A*, 354:229–230, 1996; S. Hofmann et al., Production and decay of ^{269}110, *Z. Phys. A*, 350:277–280, 1995; S. Hofmann, *On Beyond Uranium*, 2002; C. Keller, *The Chemistry of the Transuranium Elements*, 1971; Max Planck Society for the Advancement of Science, *Transurane-Transuranium Elements*, 1975; J. V. Kratz, Critical evaluation of the chemical properties of the transactinide elements, *Pure Appl. Chem.*, 75:103–138, 2003; M. Schädel, Aqueous chemistry of transactinides, *Radiochim. Acta*, 89:721–728, 2001; M. Schädel, *The Chemistry of Superheavy Elements*, 2003; M. Schädel et al., Chemical properties of element 106 (seaborgium), *Nature*, 388:55–57, 1997; G. T. Seaborg, The new elements, *Amer. Sci.*, 68:3, 1980; G. T. Seaborg, *Transuranium Elements: Products of Modern Alchemy*, 1978; G. T. Seaborg and W. D. Loveland, *The Elements Beyond Uranium*, 1990.

Traps in solids

Localized regions in a material that can capture and localize an electron or hole, thus preventing the electron or hole from moving through the material until

supplied with sufficient thermal or optical energy. Traps in solids are associated with imperfections in the material caused by either impurities or crystal defects. *See* BAND THEORY OF SOLIDS; CRYSTAL DEFECTS; HOLE STATES IN SOLIDS.

Imperfections that behave as traps are commonly distinguished from imperfections that behave as recombination centers. If the probability for a captured electron (or hole) at the imperfection to be thermally reexcited to the conduction (or valence) band before recombination with a free hole (or free electron) is greater than the probability for such recombination, then the imperfection is said to behave like an electron (or hole) trap. If the probability for a captured electron (or hole) at the imperfection to recombine with a free hole (or free electron) is greater than the probability for being thermally reexcited to the band, the imperfection is said to behave like a recombination center. In the equilibrium state in the dark, the occupancy of all imperfections is described in terms of a Fermi distribution centered on the equilibrium Fermi level. In the nonequilibrium condition of photoexcitation, the occupancy of a recombination center is determined by recombination kinetics involving the capture cross sections of the imperfection for free carriers. Since the occupancy of an electron (or hole) trap, however, is determined by a quasiequilibrium thermal exchange with the conduction (or valence) band and not by recombination in the steady state under photoexcitation as well as in the dark equilibrium state, the electron occupancy of a trap under steady-state photoexcitation can still be described in terms of a Fermi distribution, but now centered on the quasi (steady-state) Fermi level. It is possible for a specific chemical or structural imperfection in the material to behave like a trap under one set of conditions of temperature and light intensity, and as a recombination center under another.

Traps play a significant role in many phenomena involving photoconductivity and luminescence. In photoconductors, for example, the presence of traps decreases the sensitivity and increases the response time. Their effect is detectable through changes in the rise and decay transients of photoconductivity and luminescence, thermally stimulated conductivity and luminescence in which the traps are filled at a low temperature and then emptied by increasing the temperature in a controlled way, electron spin responance associated with trapped electrons with unpaired spins, and a variety of techniques involving the capacitance of a semiconductor junction such as photocapacitance and deep-level transient spectroscopy. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY; LUMINESCENCE; PHOTOCONDUCTIVITY; THERMOLUMINESCENCE.

Richard H. Bube

Bibliography. R. H. Bube, *Photoconductivity of Solids,* 1960, reprint 1978; R. H. Bube, *Photoelectronic Properties of Semiconductors,* 1992; S. W. S. McKeever, *Thermoluminescence of Solids,* 1985.

Trauma

Injury to tissue by physical or chemical means. Mechanical injury includes abrasions, contusions, lacerations, and incisions, as well as stab, puncture, and bullet wounds. Trauma to bones and joints results in fractures, dislocations, and sprains. Head injuries are often serious because of the complications of hemorrhage, skull fracture, or concussion.

Thermal, electrical, and chemical burns produce severe damage partly because they coagulate tissue and seal off restorative blood flow. Asphyxiation, including that caused by drowning, produces rapid damage to the brain and respiratory centers, as well as to other organs.

Frequent complications of trauma are shock, the state of collapse precipitated by peripheral circulatory failure, and also hemorrhage, infection, and improper healing. *See* SHOCK SYNDROME.

Edward G. Stuart; N. Karle Mottet

Bibliography. B. A. Landon and J. D. Goodall, *An Atlas of Trauma Management*: *The First Hour*, 1993.

Traveling-wave tube

A microwave electronic tube in which a beam of electrons interacts continuously with a wave that travels along a circuit, the interaction extending over a distance of many wavelengths. Traveling-wave tubes can provide amplification over exceedingly wide bandwidths. Typical bandwidths are 10–100% of the center frequency, with gains of 20–60 dB. Low-noise traveling-wave tube amplifiers serve as the inputs to sensitive radars or communications receivers. High-efficiency medium-power traveling-wave tubes are the principal final amplifiers used in communication satellites, the space shuttle communications transmitter, and deep-space planetary probes and landers. High-power traveling-wave amplifiers operate as the final stages of radars, wide-band radar countermeasure systems, and scatter communication transmitters. They are capable of delivering continuous-wave power levels in the kilowatt range and pulsed power levels exceeding a megawatt. *See* COMMUNICATIONS SATELLITE; ELECTRONIC WARFARE; RADAR; SPACE COMMUNICATIONS; SPACE PROBE.

Forward-wave amplifiers. In a forward-wave, traveling-wave tube amplifier (**Fig. 1**), a thermionic cathode produces the electron beam. An electron gun initially focuses the beam, and an additional focusing system retains the electron stream as a beam throughout the length of the tube until the beam is captured by the collector electrode. The microwave signal to be amplified enters the tube near the electron gun and propagates along a slow-wave circuit. The tube delivers amplified microwave energy into an external matched load connected to the end of the circuit near the collector. The slow-wave circuit serves to propagate the microwave energy along the tube at approximately

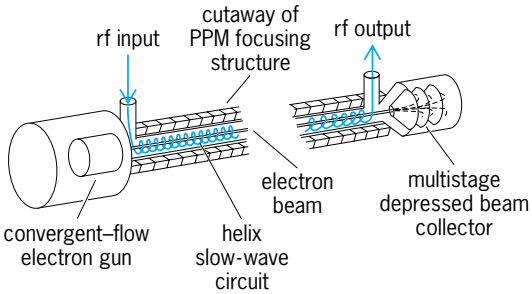

Fig. 1. Periodic-permanent-magnet (PPM) focused traveling-wave tube.

the same velocity as that of the electron beam. Interaction between beam and wave is continuous along the tube with contributions adding in phase.

Velocity and current modulations of the electron beam occur because the waves and the electrons travel in near synchronism and the amplification process takes place continuously. In the increasing electric field region of each cycle of the wave the electrons are accelerated slightly, and in the decreasing field regions the electrons are decelerated slightly. This leads to electron bunches forming with charge density in the bunches increasing with distance. The increasing charge density in the bunches induces in turn an electric field on the helix that continuously grows with distance. The electron beam is injected into the helix with a velocity slightly faster than the waves. In the ensuing interaction between fields and waves, the energy lost by the average deceleration of the electrons is the source of energy for the growing waves on the circuit. Because of the continuously distributed interaction, the power in the circuit wave grows exponentially with distance along the tube.

There are four basic elements of the traveling-wave tube: an electron gun, a means for focusing the electron beam, a slow-wave circuit, and an electron-beam collector.

Electron gun. Most modern traveling-wave tubes require electron beams, with current densities many times higher than can be achieved directly from a cathode surface. Thus, the gun design must draw the emitted electrons from the thermionic cathode and converge them into a much smaller cross-sectional area to achieve the required beam current density (**Fig. 2**).

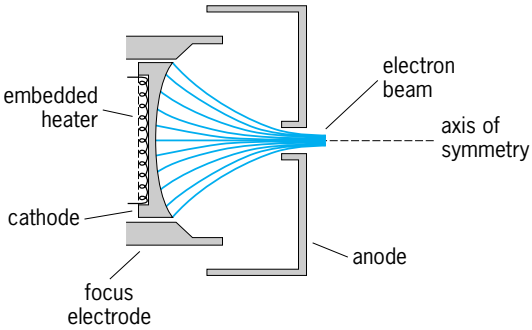

Fig. 2. Convergent-flow electron gun.

Cathode technology has continuously improved to provide high-current-density, long-life cathodes. These cathodes can produce an emitted current density of 1.0 A/cm^2 with a lifetime greater than 100,000 h. With convergent-flow electron guns and permanent-magnet focusing technology, a continuous beam current density of 75 A/cm^2 can be achieved.

Focusing methods for beam. There are two principal methods for focusing the electron beam through the length of the slow-wave circuit. The solenoid electromagnet and the periodic-permanent-magnet (PPM) focusing schemes. The former is rarely used except for special applications where consumption of power in the solenoid is acceptable. The periodic-permanent-magnet system has several advantages: (1) it consumes no power; (2) it concentrates the magnetic fields principally in the region of the electron beam and tends to cancel the fields external to the magnet system; and (3) as a result it can be made very compact and lightweight. It therefore has a great advantage for airborne and space-borne applications where weight and power consumption are important considerations.

The periodic-permanent-magnet structure can be thought of as a stack of magnetic lenses using ring magnets and disk pole pieces with alternate cells arranged in opposite polarity. This provides an axial magnetic field whose amplitude varies sinusoidally with distance. If the electron beam is introduced into this arrangement correctly, the beam can be focused over long distances with greater than 99.9% of the electron beam emerging from the end of the slow-wave circuit structure without being intercepted on it. The development of high-coercive-force permanent-magnet materials has allowed the achievement of high peak magnetic field strengths on the axis of the structure and as a result has allowed the achievement of excellent focusing of high-current-density electron beams. *See* MAGNETIC LENS.

Slow-wave circuit. In order for the electromagnetic signal wave to travel along the tube at a velocity approximately equal to the electron-beam velocity, the signal must be guided by a slow-wave circuit. Beam velocity requirements are typically 2–10% of the velocity of a free-space electromagnetic wave. The signal must be slowed down to this velocity at all wavelengths within the bandwidth of the tube. It is important that the axial electric field strength produced in the region of the electron beam be strong for a given power flow along the circuit. Certain circuits have this property in conjunction with a constant or slowly varying phase velocity as a function of frequency.

A helix is the simplest and best slow-wave circuit for low- and medium-power traveling-wave tubes and constitutes the most widely used structure (**Fig. 3***a*). It can be supported by ceramic rods within a metal vacuum envelope, and heat dissipated in the helix can be removed by conduction through

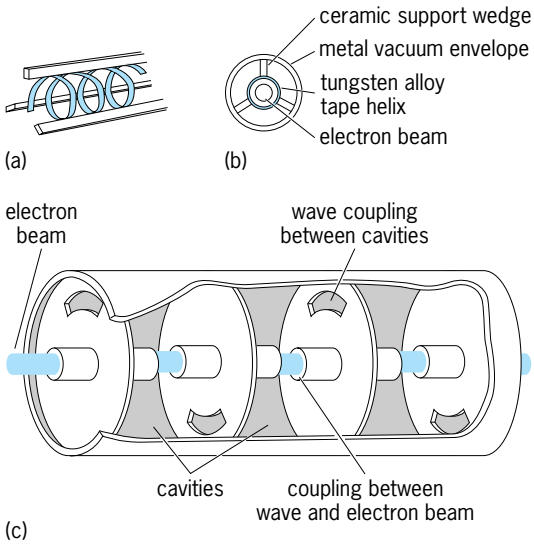

Fig. 3. Slow-wave circuits for traveling-wave tubes. (*a*) Helix circuit for medium-power tube. (*b*) Cross-sectional view of helix circuit. (*c*) Coupled-cavity circuit for high-power tube.

the rods and vacuum envelope into the surrounding magnet structure and subsequently into a heat sink.

Helix tubes have delivered up to 2500 W of continuous output power at frequencies as high as 8 GHz. For higher power output, coupled-cavity slow-wave circuits are generally used (Fig. 3*b*). This circuit has a narrower bandwidth than does a helix, 10% of center frequency being typical.

Microwave energy is usually coupled into and out of the helix through a window in the wall of the vacuum envelope of the tube. The window is usually ceramic and is designed to allow an impedance match from the external transmission system to the helix. This is usually accomplished with a direct coaxial connection, but higher-power devices often use a waveguide with a ceramic window and an internal coupling scheme to the helix or other slow-wave circuit. The coupling scheme must be impedance matched over the desired operating bandwidth to eliminate reflections. This requires very careful design. *See* COAXIAL CABLE; IMPEDANCE MATCHING; TRANSMISSION LINES; WAVEGUIDE.

To provide a stable amplifier, the input and output sections of the circuit must be isolated from each other. This is accomplished in helix-type tubes by applying lossy material to the ceramic support wedges near the center of the tube. This attenuation must have an extremely low reflection coefficient viewed from either direction not only so that stability is achieved but also so that fluctuations in gain versus frequency can meet exacting requirements. The attenuation completely absorbs the microwave signal at the downstream end of the input helix. The input and output sections of the helix are coupled by the electron beam that reexcites the growing wave in the output section due to the electron bunching that has occurred in the input section. In high-power coupled-cavity circuits, loss-impregnated ceramics are placed in selected cavities. When the lossy cavities are separated by sections of lossless circuit with 20 dB of gain, stability is good with only negligible degradation of efficiency.

Beam collector. In applications where efficiency is not important, the collector electrode and the slow-wave circuit are often connected to the same dc potential for power-supply simplicity.

A dramatic improvement in efficiency can be achieved by operating the collector electrodes at a potential that is lower than the helix. Thus, energy can be returned to the power supply as the electrons decelerate into the collector and are captured at a lower potential. The interaction that takes place between the electrons and the signal wave produces a spread in the electron velocities at the exit of the helix. Some electrons are speeded up and some are slowed down. The potential of the depressed collector can be lowered only near to the point where the slowest electrons reverse direction and return to the helix. Much effort has been expended in developing multistage collectors that allow electrons to be collected on a series of electrodes at successively lower potentials. The maximum depression and the greatest improvement in device efficiency are achieved in this manner. On the basis of the total power input to the device, overall tube efficiency of over 60% has been achieved. This technique is used extensively in satellite and deep-space applications where spacecraft power is limited and overall device efficiency is of utmost importance.

Backward-wave devices. A class of traveling-wave tubes exists called backward-wave devices in which energy on a slow-wave circuit flows in the direction opposite to the travel of electrons in the beam.

O-type oscillators. Although amplifiers are possible, the principle use of this technique is to create a voltage-tunable microwave oscillator. Typically it uses a hollow, linear electron beam and a helix circuit designed to emphasize the backward-wave fields. This represents the earliest type of voltage-tunable microwave oscillator. It is capable of generating power levels of 10–100 milliwatts with a tuning range of 2:1 in frequency. Its use has almost disappeared with the development of magnetically tuned microwave transistor oscillators using yttrium-iron-garnet (YIG) spherical resonators. *See* FERRIMAGNETIC GARNETS.

M-type oscillators. An M-type backward-wave oscillator is similar in principle to the O-type, except that focusing and interaction are through magnetic fields, as in magnetrons. Efficiency of M-type tubes is considerably higher than that of O-type tubes, typical efficiencies being 20–30%. Noise and spurious output power are also greater in the M-type tube. A continuous-wave output power of several hundred watts is typical. *See* MAGNETRON; MICROWAVE TUBE; OSCILLATOR.

Lester A. Roberts

Bibliography. A. S. Gilmour, Jr., *Microwave Tubes*, 1986; A. S. Gilmour, Jr., *Principles of Traveling Wave Tubes*, 1994; J. F. Gittens, *Power Traveling-Wave Tubes*, 1963.

Travertine

A rather dense, banded limestone (see **illus.**), sometimes moderately porous, that is formed by evaporation about springs, as is tufa, or in caves as stalac-

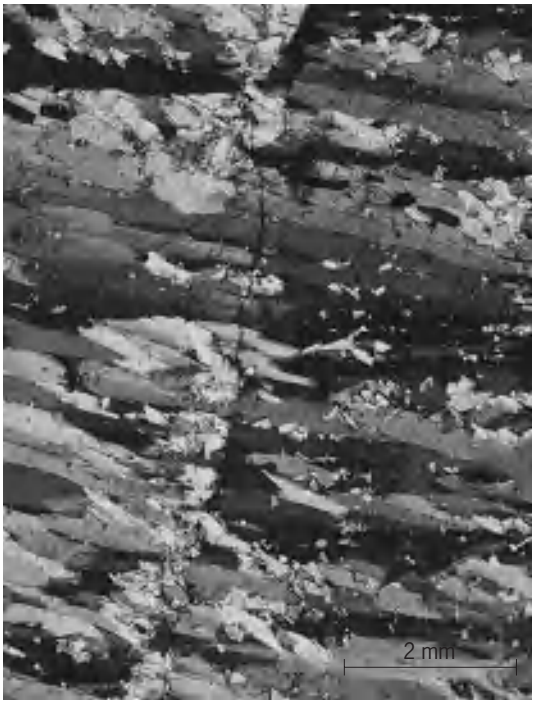

Travertine, Suisun, California. (*From E. W. Heinrich, Microscopic Petrography, McGraw-Hill, 1956*)

tites, stalagmites, or dripstone. Where travertine or tufa (calcareous sinter) is deposited by hot springs, it may be the result of the loss of carbon dioxide from the waters as pressure is released upon emerging at the surface; the release of carbon dioxide lowers the solubility of calcium carbonate, which precipitates. High rates of evaporation in hot-spring pools also lead to supersaturation. Travertine formed in caves is simply the result of complete evaporation of waters containing mainly calcium carbonate. *See* LIMESTONE; STALACTITES AND STALAGMITES; TUFA.

Raymond Siever

Tree

A perennial woody plant at least 20 ft (6 m) in height at maturity, having an erect stem or trunk and a well-defined crown or leaf canopy. However, no sharp lines can be drawn between trees, shrubs, and lianas (woody vines). For example, the strangler fig (*Ficus aurea*) is a climbing liana which may develop into a self-supporting tree if the host around which it twines is removed. Many large trees, such as paper birch (*Betula papyrifera*) or Alaska cedar (*Chamaecyparis nootkatensis*), become prostrate shrubs at the northern limits of their range in the boreal region or at their altitudinal limits near timberline.

Despite lack of agreement over an operational definition, the essence of the tree form is relatively large size, long life, and a slow approach to reproductive maturity. The difficulty of transporting water, nutrients, and storage products over long distances and high into the air against the force of gravity is a common problem of large treelike plants and one that is not shared by shrubs or herbs.

Age and size. Trees are the oldest and most massive living things on the Earth. Some of the bristlecone pine (*Pinus aristata*) in the mountains of California are approximately 4600 years old, though they are only small, gnarled trees clinging tenaciously to life in a rigorous alpine environment. These trees were 1000 years old in the biblical days of David and Solomon. Alerce (*Fitzroya cupressoides*) of Chile seem to be the second-oldest trees in the world; the most ancient was 3613 years old when it was cut in 1975. The redwoods (*Sequoia sempervirens*) reach an age of only about 2200 years but are the tallest trees in the world with a record height of 368 ft (112 m). The "big trees" (*Sequoiadendron giganteum*) are not as tall but are larger in diameter. The General Sherman tree, for example, has a diameter of 30 ft (9 m) at 4.5 ft (1.5 m) above the ground, weighs 6000 tons (5442 metric tons), and may be 3500 years old. Many other conifers of the Pacific Northwest frequently reach 200 ft (60 m) in height, and the eucalypts (especially *Eucalyptus regnans*) of Australia are nearly as large as the redwoods. In the eastern United States the oldest living trees are the bald cypress (*Taxodium distichum*), which may reach 1700 years of age. Most tree species, however, have a life-span of less than 200 years and heights under 100 ft (30 m). Girth or diameter is not a good indicator of age.

Zonation and distribution. With some exceptions (for example, the Joshua tree, *Yucca brevifolia*), trees are generally found in aggregates called forest stands. Because trees are the most noticeable component of the global flora, the forest cover is used to categorize world vegetation types. Forests are found from the tropics to the boreal regions and occupy one-half of the total land area. Precipitation and temperature are the major factors limiting growth of forests. In general, forests are restricted to areas where precipitation exceeds 25 in. (62.5 cm) per year. Trees are not found in the Arctic tundra and, except along water courses, they are excluded from the dry steppes of Russia, the Great Plains of the United States, and desert regions of the world. Timberline, the zone in alpine areas above which trees are not found, runs 13,000–14,000 ft (4000–4300 m) in the tropics, 10,000–12,000 ft (3000–4000 m) in the Sierra Nevada, 6000 ft (1800 m) in the Alps, 5000 ft (1500 m) in New England, and 1000–3000 ft (300–900 m) in southern Alaska. Timberline has fluctuated in concert with climate change; for example, in the Sierra Nevada region it has changed by as much as 230 ft (70 m) during the last 6300 years. *See* ALTITUDINAL VEGETATION ZONES.

As one travels from the boreal forest to the Equator, the variety of arborescent species increases from

a few forms in the subarctic to about 200 per acre in the tropics. Malaysia and Amazonia are the richest areas of the world in number of tree species. In the United States the southern Appalachians provided a refuge for migrating species during successive glacial periods and as a result have a diversity of tree forms. On the other hand, there is a relative paucity of species in the European flora. European forms became extinct when they were trapped between ice sheets advancing from the north and other glaciers originating in the Alps.

Unfortunately, the rate of deforestation is accelerating. Forest diversity, particularly in the tropics, is threatened by deforestation. On average, over 38 million acres (15 million hectares) of tropical forest was cleared and converted to other uses annually during 1981–1990.

Classification. Almost all existing trees belong to the seed plants (Spermatophyta). An exception are the giant tree ferns which were more prominent in the forests of the Devonian Period and today exist only in the moist tropical regions, where they grow to heights of 60 ft (18 m). The Spermatophyta are divided into the Pinophyta (gymnosperms) and the flowering plants, Magnoliophyta (angiosperms). The gymnosperms bear their seed naked on modified leaves, called scales, which are usually clustered into structures called cones—for example, pine cones. By contrast the seed of angiosperms is enclosed in a ripened ovary, the fruit. *See* MAGNOLIOPHYTA; PINOPHYTA; POLYPODIALES.

The orders Cycadales, Ginkgoales, and Pinales of the Pinophyta contain trees. *Ginkgo biloba*, the ancient maidenhair tree, is the single present-day member of the Ginkgoales. The Cycadales, characteristic of dry tropical areas, contain many species which are small trees. The Pinales, found throughout the world, supply much of the wood, paper, and building products of commerce. They populate at least one-third of all existing forest and include the pines (*Pinus*), hemlocks (*Tsuga*), cedars (*Cedrus*), spruces (*Picea*), firs (*Abies*), cypress (*Cupressus*), larches (*Larix*), Douglas-fir (*Pseudotsuga*), sequoia (*Sequoia*), and other important genera. The Pinales are known in the lumber trade as softwoods and are popularly thought of as evergreens, although some (for example, larch and bald cypress) shed their leaves in the winter. *See* CEDAR; CYCADALES; CYPRESS; DOUGLAS-FIR; FIR; HEMLOCK; LARCH; PAPER; PINALES; PINE; PINOPHYTA; SEQUOIA; SPRUCE.

In contrast to the major orders of gymnosperms which contain only trees, many angiosperm families are herbaceous and include trees only as an exception. Only a few are exclusively arborescent. The major classes of the angiosperms are the Liliopsida (monocotyledons) and the Magnoliopsida (dicotyledons). The angiosperm trees, commonly thought of as broad-leaved and known as hardwoods in the lumber market, are dicotyledons. Examples of important genera are the oaks (*Quercus*), elms (*Ulmus*), maples (*Acer*), and poplars (*Populus*). *See* ELM; LILIOPSIDA; MAGNOLIOPSIDA; MAPLE; OAK; POPLAR.

The Liliopsida contain few tree species, and these are never used for wood products, except in the round as posts. Examples of monocotyledonous families are the palms (Palmae), yucca (Liliaceae), bamboos (Bambusoideae), and bananas (Musaceae). *See* BAMBOO; BANANA.

Morphology and physiology. The morphology of a tree is similar to that of other higher plants. Its major organs are the stem, or trunk and branches; the leaves; the roots; and the reproductive structures. Almost the entire bulk of a tree is nonliving. Of the trunk, branches, and roots, only the tips and a thin layer of cells just under the bark are alive. Growth occurs only in these meristematic tissues. Meristematic cells are undifferentiated and capable of repeated division. *See* FLOWER; LATERAL MERISTEM; LEAF; PLANT GROWTH; ROOT (BOTANY); STEM.

Growth. Height is a result of growth only in apical meristems at the very tips of the twigs. A nail driven into a tree will always remain at the same height, and a branch which originates from a bud at a given height will never rise higher. The crown of a tree ascends as a tree ages only by the production of new branches at the top and by the death and abscission of lower, older branches as they become progressively more shaded. New growing points originate from the division of the apical meristem and appear as buds in the axils of leaves. *See* APICAL MERISTEM; BUD; PLANT GROWTH.

In the gymnosperms and the dicotyledonous angiosperms, growth in diameter occurs by division in only a single microscopic layer, three or four cells wide, which completely encircles and sheaths the tree. This lateral meristem is the cambium. It divides to produce xylem cells (wood) on the inside toward the core of the tree and phloem cells on the outside toward the bark. In trees of the temperate regions the growth of each year is seen in cross section as a ring. Because of this yearly increment of xylem elements, the tree structure develops as conical shells over shells (see **illus.**). *See* BARK; PHLOEM; XYLEM.

Xylem elements become rigid through the thickening and modification of their cell wall material. The tubelike xylem cells transport water and nutrients from the root through the stem to the leaves. In time the xylem toward the center of the trunk becomes impregnated with various mineral and metabolic products, and it is no longer capable of conduction. This nonfunctional xylem is called heartwood and is recognizable in some stems by its dark color. The light-colored, functional outer layer of the xylem is the sapwood. *See* WOOD ANATOMY.

The phloem tissue transports dissolved carbohydrates and other metabolic products manufactured by the leaves throughout the stem and the roots. Most of the phloem cells are thin-walled and are eventually crushed between the bark and the cambium by the pressures generated in growth. The outer bark is dead and inelastic but the inner bark contains patches of cork cambium which produce new bark. As a tree increases in circumference, the old outer bark splits and fissures develop, resulting in the rough appearance characteristic of the trunks of most large trees.

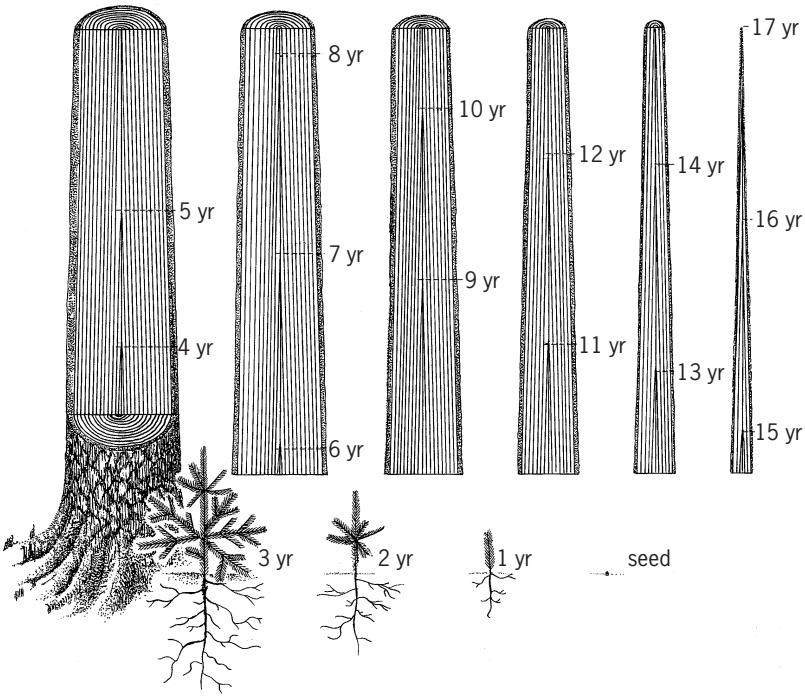

Schematic drawing of a 17-year-old coniferous tree showing the manner in which the trunk increases in thickness through the addition of annual increments. (*After A. J. Panshin et al., Textbook of Wood Technology, vol. 1, 2d ed., 1964*)

In the monocotyledons the lateral cambium does not encircle a central core, and the vascular or conducting tissue is organized in bundles scattered throughout the stem. The trunk is not wood as generally conceived although it does in fact have secondary xylem.

Phylogeny and evolution. Trees have a very ancient history on Earth, first being recorded in Devonian strata deposited 300 million years ago (Ma). These trees included fernlike plants which bore sporangia rather than seeds. The first confirmed seed plants, the seed ferns (pteridosperms), appeared in the Mississippian 250 Ma and are the probable ancestors of the gymnosperms. The seed ferns vanished after the Jurassic Period. During the Pennsylvanian, dense lowland forests of giant horsetails (*Equisetum*) and club mosses (*Lycopsida*), now extinct, produced the world's coal deposits. Later in this era the Cycadales and Ginkgoales were abundant, and the ancestors of the Pinales were also present. In the late Paleozoic, 190 Ma, early conifers were fossilized. The first recognizable angiosperms were trees and appeared in the Mesozoic Era, more than 150 Ma. Few angiosperms are represented in the Jurassic flora, although some of those present, such as sycamore (*Platanus*), tulip poplar (*Liriodendron*), and sweet gum (*Liquidambar*), are still recognizable today. During the Cretaceous Period, 100 Ma, the angiosperms increased greatly in numbers. Angiosperms undoubtedly evolved from ancient gymnosperms. All the vast number of herbaceous angiosperms known today have in turn evolved from the early flowering trees. *See* EQUISETALES; LEPIDODENDRALES; PTERIDOSPERMS; TREE FERNS.

The tree form with its upright stem 100–300 ft (30–90 m) in height may have been evolutionarily successful because of the advantage it conferred in competition for sunlight. On the other hand, the longer interval between generations may mean that tree populations cannot adapt to a changing environment through the genetic process of recombination and selection as rapidly as short-lived plants. Also, trees must be physiologically capable of surviving under the varying climatic conditions experienced from year to year and cannot afford to adapt to a specific set of conditions as can annual plants. However, all modern seed-bearing plants have evolved from tree ancestors. The trees themselves provide products necessary to humanity's continued well-being. These products range from medicines and other chemicals through paper to major construction materials. *See* DENDROLOGY; FOREST AND FORESTRY; FOREST TIMBER RESOURCES; PALEOBOTANY; PLANT EVOLUTION; PLANT PHYSIOLOGY; PLANT TAXONOMY; TREE DISEASES.

F. Thomas Ledig

Bibliography. R. R. Aicher, *Growth Stresses and Strains in Trees*, 1986; W. M. Harlow, E. S. Harrar, and F. M. White, *Textbook of Dendrology*, 8th ed., 1995; M. R. Sethuraj and A. S. Raghavendra (eds.), *Tree Crop Physiology*, 1987; M. H. Zimmerman and C. L. Brown, *Trees: Structure and Function*, 1985.

Tree diseases

Diseases of both shade and forest trees have the same pathogens, but the trees differ in value, esthetics, and utility. In forests, disease is significant only when large numbers of trees are seriously affected. Diseases with such visible symptoms as leaf spots may be alarming on shade trees but hardly noticed on forest trees. Shade trees with substantial rot may be ornamentals with high value, whereas these trees would be worthless in the forest. Emphasis on disease control for the same tree species thus requires a different approach, depending on location of the tree.

Diseases of forest trees. From seed to maturity, forest trees are subject to many diseases. Annual losses of net sawtimber growth from disease (45%) are greater than from insects and fire combined. Young, succulent seedlings, especially conifers, are killed by certain soil-inhabiting fungi (damping-off). Root systems of older seedlings may be destroyed by combinations of nematodes and such fungi as *Cylindrocladium, Sclerotium*, and *Fusarium*. Chemical treatment of seed or soil with formulations containing nematicides and fungicides, and cultural practices unfavorable to root pathogens help to avoid these diseases.

Roots rots are caused by such fungi as *Heterobasidion* (= *Fomes*) *annosus* (mostly in conifers) and *Armillariella* (= *Armillaria*) *mellea* (mostly in hardwoods). These fungi cause heart rot in the roots and stems of large trees and also invade and kill young, vigorous ones. Thinning pine plantations increases

infection by *H. annosus*, which invades fresh stumps and grows through root grafts to surrounding trees. Losses are minimized by strategic timing of thinning or by stump treatment with borax, urea, or the fungus *Peniophora gigantea*. In the southern United States thinnings are made in summer when temperatures are too high for spore development. *Armillariella mellea* invades and destroys roots and stems of trees predisposed by other stress factors. Both *H. annosus* and *A. mellea* occur worldwide in temperate zones.

Leaf diseases and wilt. In natural forests, leaf diseases are negligible, but in nurseries and plantations, fungal infections cause severe defoliation, retardation of height growth, or death. *Scirrhia acicola* causes brown spot needle blight and prevents early height growth of longleaf pine in the South; it defoliates Christmas tree plantations of Scotch pine (*Pinus sylvestris*) in northern states. Fungicides and prescribed burning are used successfully for control. In Christmas tree plantations, control is possible with maneb or chlorothalonil. *Dothistroma* needle blight is severe in plantations of Austrian pine (*P. nigra*) and ponderosa pine (*P. ponderosa*), and in plantations of Monterey pine (*P. radiata*) in Australia, New Zealand, Africa, and South America. In North America, *Lophodermium pinastri* killed millions of pine seedlings in the 1970s. Maneb and chlorothalonil are effective for control when new ascospores are being released.

Oak wilt is a systemic disease, with the entire tree affected through its water-conducting system. Trees of susceptible species may be killed in a few weeks to a year or more through plugging of the water-conducting vessels. The causal fungus, *Ceratocystis fagacearum*, spreads to nearby healthy trees by root grafts and to trees at longer distances by unrelated insects, including bark beetles of the Nitidulidae. The sporulating mats of the fungus develop between bark and wood, producing asexual and, sometimes, sexual spores, which are disseminated by insects attracted by aromatic odors of the fungus. Control is possible by eradicating infected trees and by disruption of root grafts by trenching or by chemicals.

Rust diseases and cankers. Stem rust diseases occur as cankers or galls on coniferous hosts and as minor lesions on other ones. A few, such as white pine blister rust (**Fig. 1**) and southern fusiform rust, are epidemic, lethal, and economically important. Others of less immediate importance (such as western gall rust) are capable of serious, widespread infection, requiring no other hosts to complete their life cycles. Most rust fungi require two unrelated hosts, such as the currant and white pine for blister rust. Eradication of hosts of secondary importance is cost-prohibitive, and resistant varieties are favored for control. Other control measures include pruning out early infections and spraying nursery trees with chemicals during periods favoring needle infection.

Except for stem rust, cankers of conifers are generally of minor importance. However, scleroderris canker is a new, serious threat to pine stands (especially hard pines) in Canada and the United States. The causal fungus (*Gremeniella abietina*) is favored by cold weather, and is epidemic and lethal in forest reproduction, nurseries, plantations, and young stands. Control in nurseries is possible by successive applications of chlorothalonil; pruning and sanitation are recommended for plantations.

Fig. 1. White pine blister rust (*Cronartium ribicola*) on white pine. The blisters are masses of aeciospores that are wind-disseminated for many miles to *Ribes* plants. (*Photograph by Robert Campbell*)

Stem infections by numerous fungi, resulting in localized death of cambium and inner bark, range from lesions killing small stems in a year (annual) to gross stem deformities (perennial), where cankers enlarge with stem growth. Chestnut blight, first known in the United States in 1904, destroyed the American chestnut as a commercial species (**Fig. 2**), and is an example of the annual lesion type. The less dramatic or devastating *Nectria* canker destroys stems of timber value, and is an example of the perennial lesion type. Large, living chestnut trees are rare in the United States, but disease-resistant roots maintain the chestnut as a shrub or small tree in native stands in the eastern hardwood forest. Resistant varieties of chestnut lack the superior qualities of the native American species. Hope for biological control of chestnut blight is in newly discovered hypovirulent strains of the blight fungus that prevent the killing of trees infected with virulent ones. The hypovirulent strains may be themselves diseased by a viruslike agent of double-stranded ribonucleic acid.

Tree decay. All tree species, including decay-resistant ones such as redwood, are subject to ultimate disintegration by fungi. Decay fungi (Hymenomycetes) are associated with nondecay fungi (Deuteromycetes) and bacteria. These microflora enter the tree through wounds, branch stubs, and roots, and are confined to limited zones of wood

Fig. 2. Chestnut blight in American chestnut tree (*USDA*). Insert is mycelial fan of chestnut blight caused by *Endothia parasitica*, advancing through bark of American chestnut. The tip of the fan on the left is surrounded by cortical tissue. The contents of the cortical cells back from the tip of the fan are discolored to a yellowish brown, as indicated by the darkened cells (*from J. S. Boyce, Forest Pathology, 3d ed., McGraw-Hill, 1961*).

by anatomical and wound-stimulated tissue barriers (**Fig. 3**). The extent of decay is limited by compartmentalization of decay in trees. Trees aged beyond maturity are most often invaded by wood-rotting fungi; losses can be minimized by avoiding wounds and by shortening cutting rotations. Losses from rot are especially serious in overmature coniferous stands in the western United States, Canada, and Alaska. *Fomes pini*, the cause of white pocket rot in most commercial conifers, and *Echinodontium tinctorium*, the cause of stringy brown butt rot in fir and hemlock, are two of the most destructive heart rot fungi. The processes of discoloration and decay in trees that reduce the quality of wood are initiated by wounds. If the tree response to wounding by chemical and physiological reactions is fast, compartmentalization will limit damage to a small discolored zone without decay. But if the tree response is slow, bacteria and nondecay fungi will predispose the wood to decay by the wood decay fungi. Soon after wounding, the cambium is stimulated to produce a special layer of cells that prevents invasion of newly formed tissue by internal microflora. Thus, new wood rings are free of defect. *See* WOOD DEGRADATION.

Fig. 3. Shelf fungus (*Fomes applanatus*) on a dead aspen tree. This wood-rotting fungus enters through roots and wounds and decays the heartwood and sapwood of both hardwoods and softwoods.

Dwarf mistletoe disease. Dwarf mistletoes are small seed plants that are obligate parasites on coniferous trees; they cause losses of more than 3×10^9 board feet of lumber annually. More than 20 species occur on millions of acres of western conifers, and on spruce in the eastern United States. Once limited by natural fires, they have increased in distribution and severity with better fire control. Dwarf mistletoe plants may be $^1/_2$ in. (13 mm) to several inches in height and are leafless and scalelike. They absorb nutrients and water from host plants through haustoria and sinkers in bark and wood. They depend on their hosts for most of their food and produce fruit with seeds that are forcibly discharged up to 30 ft (9 m). Limited seed dispersal distance results in circular infection patterns around source trees; spread is slow except for occasional transmission by birds.

Miscellaneous diseases. Another series of tree diseases is recognized for which there is no satisfactory explanation. Such diseases as birch dieback, little-leaf disease of shortleaf pine, pole blight of western white pine, and ohia decline in Hawaii represent complex combinations of common symptoms and multiple causal factors, defying positive diagnosis. One or more known pathogenic organisms that appear likely as potential first causes are later regarded only as secondary agents. Several adverse environmental factors cause stress, predisposing trees to colonization by weakly pathogenic fungi. These diseases reflect the degree to which human-induced change of environment may be regarded as a primary cause of tree disease. *See* REFORESTATION.

Disease of shade trees. Many shade trees are grown under conditions for which they are poorly adapted, and are subject to environmental stresses not common to forest trees. Both native and exotic trees planted out of natural habitats are predisposed to secondary pathogens following environmental stress of noninfectious origins. They are also susceptible to the same infectious diseases as forest trees. Appearance is more important than the wood produced, and individual value is higher per tree than for forest trees. Thus, disease control methods differ from those recommended for forest trees.

Fungal diseases. The most important and destructive shade tree disease known is Dutch elm disease, introduced from Europe to North America before 1930 (**Fig. 4**). Elms represented millions of prized shade trees in monoculture plantings in the urban

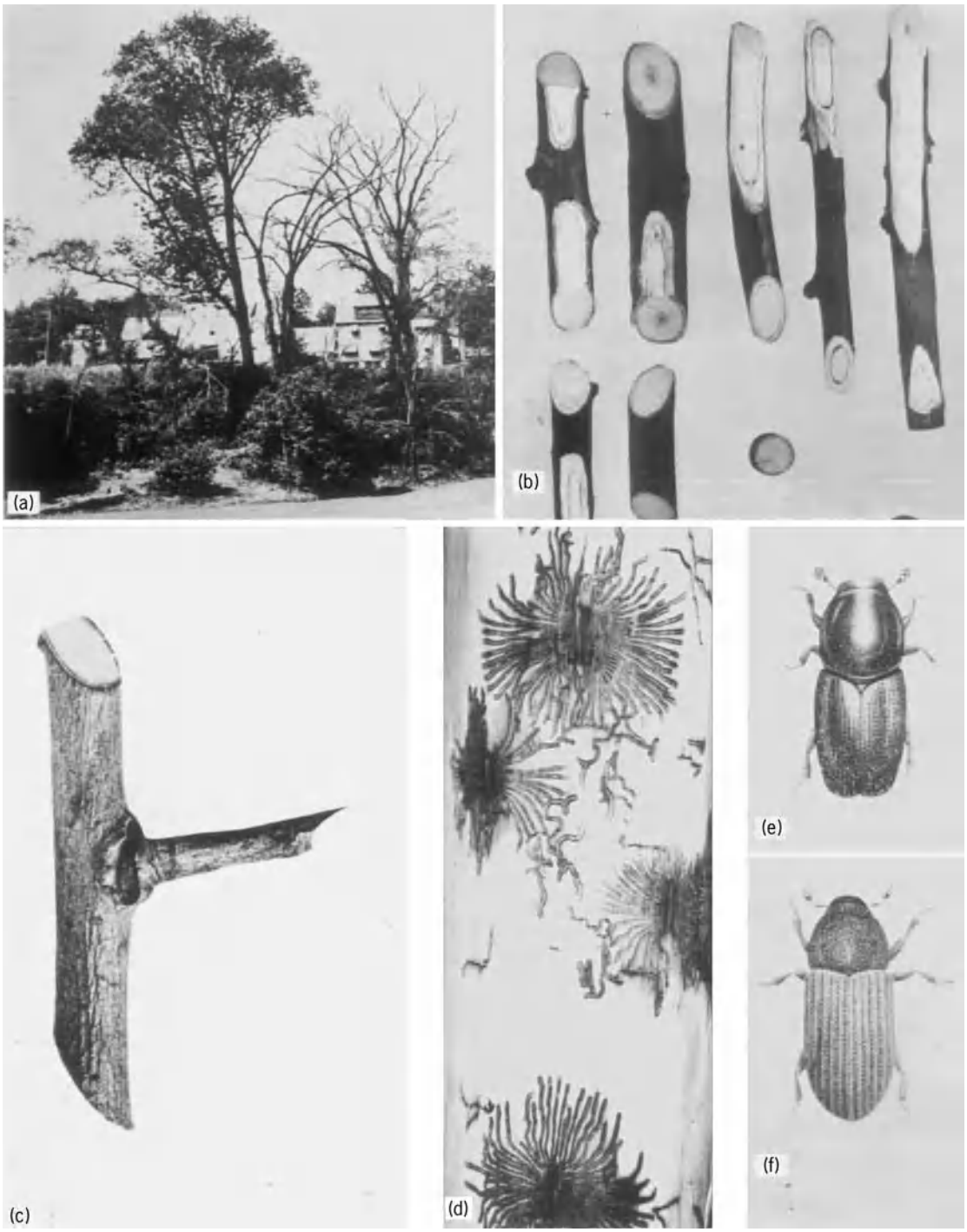

Fig. 4. Dutch elm disease. (*a*) Group of trees affected by the disease. (*b*) Discoloration in the sapwood of infected trees. (*c*) Feeding scar in small elm crotch made by an adult of the smaller European elm bark beetle (*Scolytus multistriatus*). (*d*) Brood galleries made by female beetles and larvae. (*e*) European elm bark beetle, the most important carrier of the Dutch elm disease. (*f*) Native elm bark beetle. (*a, c, d, USDA*; *b, e, f, Michigan State University*)

United States and Canada, and losses and control efforts cost millions of dollars annually. The causal fungus, *Ophiostoma ulmi* (*Ceratocytis ulmi*), is introduced to the water-conducting system of healthy elms by the smaller European elm bark beetle (*Scolytus multistriatus*) or the American elm bark beetle (*Hylurgopinus rufipes*). One or more new and more aggressive strains of the fungus have arisen since 1970. More devastating than the original ones, they are destroying the elms in North America and Europe that survived earlier epidemics. Effective means of prevention are sanitation (destroying diseased and dying and dead elm wood); insecticidal sprays (methoxychlor); disruption of root systems; and early pruning of new branch infections. Of much promise are resistant varieties of elm, systemic fungicides (such as solubilized benzimidazoles), and insect pheromones. Reports (subject to confirmation)

Fig. 5. Elm infected with bacterial wetwood. The bacterial ooze is coming from a wound.

on infectious agents (such as bacteria and viral particles) inside the Dutch elm disease fungus, as well as antifungal bacteria, offer hope for biological control.

Bacterial diseases. The most common bacterial disease of shade trees is wetwood (**Fig. 5**) of elm and certain other species. It is reported to be caused by a single bacterial species (*Erwinia nimipressuralis*), although causal associations of other bacteria are now suspected. The bacteria are normally present in the heartwood of mature elms and cause no disease unless they colonize sapwood by exterior wounds. Fermented sap under pressure bleeds from wounds and flows down the side of the tree. Sustained bleeding kills underlying cambial tissue. Internal gas pressure and forced spread of bacterial toxins inside living tissues of the tree can be reduced by strategic bleeding to avoid seepage into bark and cambium.

A second bacterial disease of elm, elm yellows (= elm phloem necrosis), is caused by a mycoplasma-like organism, considered to be a unique kind of bacterium. Elm yellows is as lethal as Dutch elm disease but is more limited in distribution. Foliage of infected trees turns yellow and falls off; infected trees die. Diagnostic symptoms are browning of inner bark near the base of the tree and wintergreen odor of affected tissue. The pathogen is carried by the elm leafhopper (*Scaphoideus luteolus*), which sucks phloem juice from leaf veins. Spread of disease also occurs through grafted root systems. Control measures include early destruction of infected trees, disruption of root systems, and insecticidal sprays. Injection with tetracycline and other antibiotics helps to slow the progress of the bacterium.

Leaf spot and heart rot. Leaf diseases of minor importance on forest trees often appear alarming or ominous on shade trees. These diseases are always present and become conspicuous following a wet spring. They are mostly cosmetic in impact, are rarely serious, and may be minimized by elimination of fallen leaves and timely fungicidal sprays in the early spring.

Heart rot is of minor importance in shade trees. Many trees with extensive heart rot remain healthy for decades. Nothing will stop this internal decay, but exposure to air may decrease its rate; sealing with cavity work may accelerate the decay.

Noninfectious disease. The most common and complex diseases of shade trees are diebacks and declines (such as maple decline). Many species show similar patterns of symptoms caused by multiple factors, but no single causal factor is known to cause any one of these diseases. Noninfectious agents of shade tree diseases are drought, soil compaction, mineral deficiency, soil pollution from waste or salt, air pollution, and so on. Trees affected experience chlorosis, premature fall coloration and abscission, tufting of new growth, dwarfing and sparseness of foliage, progressive death of terminal twigs and branches, and gradual to rapid decline in growth. Such trees are often infested with borers and bark beetles, and infected by branch canker (such as *Cytospora* and *Nectria* spp.) and root rot fungi (such as *Armillariella mellea*). Noninfectious stress predisposes trees to infectious disease that is caused by different kinds of weakly parasitic fungi as secondary pathogens. *See* PLANT PATHOLOGY.

Richard J. Campana

Tree ferns

Plants belonging to the families Cyatheaceae and Dicksoniaceae, whose members typically develop tall trunks crowned with leaves (fronds) which often reach some 20 ft (6 m) in length and 5–6 ft (1.5–1.8 m) in width (**Figs. 1** and **2**). Tree ferns reach their greatest development in the rainforests and cloud forests of the mountainous tropics. *See* RAINFOREST.

Distribution. Two families and 13 genera are recognized. The Cyatheaceae are represented by the genera *Lophosoria* (tropical America, 1 species); *Metaxya* (tropical America, 1 species); *Sphaeropteris* (tropical America, India and southeastern Asia to New Zealand, the Marquesas, and Pitcairn Island, about 120 species); *Alsophila* (pantropic, about 230 species); *Nephelea* (tropical America, about 30 species); *Trichipteris* (tropical America, about 90 species); *Cyathea* (tropical America, about 110 species); and *Cnemidaria* (tropical America, about 40 species). The other family, Dicksoniaceae, is represented by *Dicksonia* (tropics and southern subtropics in Malaysia, Australasia, America, Hawaii, St. Helena, about 25 species); *Cystodium* (Malaysia, 1 species); *Thyrsopteris* (Juan

Fig. 1. ***Cyathea arborea*** **in the Luquillo Forest, Puerto Rico. (*Courtesy of M. Canoso*)**

Fernández, 1 species); *Culcita* (tropical America, Azores, Malaysia, Australasia, about 7 species); and *Cibotium* (Southeast Asia, Malaysia, Hawaii, Central America, about 12 species).

Stems and leaves. Tree fern trunks may reach 65 ft (19.8 m) in height, as in *Alsophila australis* of Australia. Their diameters vary from 0.4 in. (1.0 cm) in *A. biformis* of New Guinea to 4 ft (1. 2 m) in the *Cibotium* trunks of Hawaii. The lower trunk is often densely covered with matted adventitious roots which greatly increase its diameter. Certain specimens branch near the base of the trunk or higher up; perhaps this branching occurs in response to injury.

Fig. 2. Upper part of the 30-ft (9-m) trunk of *Cyathea arborea* showing the crown of arching 7-ft (2-m) leaves (called fronds). (*Courtesy of T. Plowman*)

The degree of division of the leaves varies from simple in *A. sinuata* of Ceylon to four or five pinnate in *Lophosoria*. The leaflets (pinnae) are usually smaller toward the base of the leaf; when these basal leaflets are branched into threadlike divisions, they are called aphlebiae.

In most species the old leaves and petiole bases are eventually deciduous, leaving distinctive scars on the trunk. These scars are usually arranged in spirals whose spacing varies with the growth rate. In *Nephelea aureonitens* of Costa Rica, however, the scars are in rings at intervals of several inches. This pattern may relate to the reported occurrence of totally leafless resting periods in this species and certain others. The majority of species are evergreen, however, and produce a new flush of leaves during each wet season.

Scales and hairs. In the Dicksoniaceae and *Lophosoria* and *Metaxya*, simple septate hairs, often quite thick at the base, are found on the leaves and trunks. All species of *Sphaeropteris, Alsophila, Nephelea, Trichipteris, Cyathea*, and *Cnemidaria*, however, have scales, with hairs also on the upper and sometimes the lower leaf surface. Many species of Cyatheaceae are also characterized by variously developed spines on the leaf buds (croziers), petioles, and trunk.

Sori. The Dicksoniaceae have marginal sori terminal on the veins and protected by a bivalved indusium. The Cyatheaceae produce sori well away from the margin, usually seated at the forking of a vein or midway along a simple vein. The indusia may be absent, a small scale, a cup, or a completely enclosing membrane. Sporangia are relatively small and in most cases contain 64 trilete, variously ornamented spores. The indurated annulus is more or less oblique and uninterrupted by the attachment of the stalk.

Anatomy. The typical vascular system of both families is dictyostelic. The Cyatheaceae have accessory vascular strands in the pith and cortex. Fibrous sheaths around the vascular tissue and just inside the epidermis provide mechanical support. The xylem consists of scalariform tracheids and parenchyma, the phloem of sieve tubes and parenchyma. Numerous mucilage canals are embedded in the pith and cortex. *See* STEM.

Fossils. The earliest satisfactory fossils of this group are Jurassic and Cretaceous leaf compressions and trunk petrifactions. Cyatheo-Dicksoniaceous tree ferns had worldwide distribution during the Jurassic and Cretaceous, some occurring as far north as Greenland. In the Tertiary they began to disappear from the northern regions and became restricted to their present tropical distribution. *See* PALEOBOTANY.

Uses. The young leaf buds and pith have been used as food and fermented into alcohol. The woolly hairs of *Cibotium* are used to stanch wounds and stuff cushions. In southern China and Taiwan, apices of

Cibotium barometz are fashioned into "vegetable lambs" and sold to tourists. The trunks, highly resistant to termites and decay, are used in building and fencing. In New Zealand, the trunks are carved into ornamental objects called ponga ware. The adventitious roots are much sought for use in orchid culture, either as solid slabs or broken in potting mixtures.

Classification. Although the tree ferns have traditionally been classified in the Dicksoniaceae and Cyatheaceae, some studies unite them in the inclusive family Cyatheaceae with four subfamilies, and others maintain both families, considering each as an independent evolutionary line. Within the traditional Cyatheaceae, the former genera *Alsophila*, *Hemitelia*, and *Cyathea* have long been recognized as unnatural. Research indicates that the cellular detail of the petiole scales has very strong correlative value in distinguishing natural evolutionary lines. This is reflected in the modern classification of the scaly Cyatheaceae into the genera *Sphaeropteris, Alsophila, Nephelea, Trichipteris, Cyathea*, and *Cnemidaria*. *See* POLYPODIALES; TREE. Gerald J. Gastony

Bibliography. L. R. Atkinson and A G. Stokey, Comparative morphology of the gametophyte of homosporous ferns, *Phytomorphology*, 14(1):51–70, 1964; F. O. Bower, *The Ferns*, vol. 2, 1926; G. Dunk, *Ferns*, 1994; R. E. Holttum, Cyatheaceae, *Flora Malesiana*, ser. 2, 1(2):65–176, 1963; R. E. Holttum and U. Sen, Morphology and classification of the tree ferns, *Phytomorphology*, 11(4):406–420, 1961; W. R. Maxon, The tree ferns of North America, *The Smithsonian Report for 1911*, pp. 463–491, 1912; R. M. Tryon, The classification of the Cyatheacea, *Contributions from the Gray Herbarium of Harvard University*, vol. 200, 1970.

Tree growth

Trees grow to a larger size at maturity than other woody perennials, have a comparatively long period of development to maturity, and live a long time in the mature state. Trees, like other vascular plants, are made up of cells; growth is the result of adding more cells through cell division, and of the elongation and maturation of those cells into functional tissues. Cells and the tissues they compose differ in structure and function. Some tissues, such as the corky bark, mechanically protect the tree and insulate it against rapid temperature changes. Xylem tissues conduct water, minerals, and some hormones up the tree, and the phloem tissues conduct photosynthate sugar solutions and other organic molecules in most cases down the tree. That fraction of the xylem and phloem which conducts liquids is termed vascular tissue. There are also other nonconducting cells such as parenchyma within the xylem and phloem tissues, which is important in photosynthesis and food storage.

Tissues may also be compared by their origin. Primary tissues are derived from apical meristems of the stem and root, whereas secondary tissues arise from secondary cambia. Two important secondary cambia in trees are the vascular cambium which produces cells which become secondary xylem (the wood of commerce) to the inside and secondary phloem tissue to the outside, and the cork cambium or phellogen which produces corky bark tissues to the ouside and usually phelloderm tissue to the inside. Collectively, these secondary cambia account for most of the diameter growth in trees. *See* PLANT TISSUE SYSTEMS.

Photoperiod has a large influence on the growth of trees. Height growth rate, flowering, date of vegetative budset, and in some species leaf abscission are affected. Plant response is actually to length of the night rather than of the day. Thus, when it is said that growth is promoted by long photoperiods, in fact the plant senses short nights and so continues to grow rapidly. The glycoprotein-pigment complex phytochrome, located in the plasmalemma, is responsible for transducing the night length into the plant's physiological system. The exact nature of the biological action of the phytochrome system is not known. *See* PHOTOPERIODISM.

The photoperiod which affects growth or developmental responses is a function of the photoperiod of the location where the tree species evolved. Trees that have similar photoperiodic responses are termed photoperiodic ecotypes. Research has shown the heritable nature of photoperiodic response, since hybrids of trees from different photoperiodic ecotypes show intermediate responses.

Growth and the cell wall. In vascular plants, growth generally refers to the development of organized structures. Longitudinal growth depends upon cell division in apical meristems and upon subsequent elongation of these cells. Unlike animal cells, plant cells are encased in a rigid cell wall. Growth begins by breaking of chemical bonds in the cellulosic framework of the wall, followed immediately by insertion of new cellulose molecules into the wall. This type of cell-wall growth is termed intussusception and allows the primary cell wall to grow in area while retaining the same thickness. Primary walls are laid down by immature cells during the growth process, but as growth ends, secondary walls are formed in some types of cells (for example, the secondary xylem tracheids and vessels).

Primary cell walls have similar structure in a variety of higher plants. The composition and ultrastructure of secondary walls, however, vary considerably even from one cell type to another. In general, the secondary wall is considerably thicker than the primary wall and is laid down in three layers which differ in orientation of cellulose microfibrils. The secondary wall is formed by adding wall material to the inner surface of the enlarged cell. This is termed apposition and differs from the intussusception growth process of the primary cell wall. During early stages of secondary wall development, lignin precursors synthesized by the cell itself or diffused from the vascular cambium polymerize in spaces between the cellulose microfibrils of the primary and secondary walls to form the rigid, lignified, mature cell wall. *See* CELL WALLS (PLANT).

Meristems. Shoot and root apices, originating in the seed embryo, are major centers of organization in development of the tree. Cell divisions occur in distinct patterns in these apices. These patterns are the basis for development of tissues from which the mature axial structures of the shoot and root are formed. The shoot apex is the source of cells which develop in very predictable patterns into leaf primordia, axillary buds that can grow into lateral branches, and reproductive structures. The root apex is a less complex structure producing only cells which will develop into the internal tissues of the root. Lateral roots are not formed at the apex, but in pericycle tissues some distance behind the apex.

Plant cells are totipotent; that is, they have the total genetic information necessary to program for any cell type, organ, or the entire plant. Cells produced at apices can differentiate into cell types to form different tissues, depending on the chemical and physical environment to which they are exposed. The totipotency concept was proposed at the turn of the century. It was first demonstrated in the 1950s when single cells of carrot were cultured to whole plants. In trees, somatic embryos have been produced successfully in sweetgum, which is an angiosperm. Techniques for inducing somatic embryogenesis in conifers have been developed in Norway spruce (*Picea abies*), white spruce (*Picea glauca*), sugar pine (*Pinus lambertiana*), Douglas-fir (*Pseudotsuga menziesii*), and loblolly pine (*Pinus taeda*). These methods offer possibilities for the efficient vegetative propagation of these species.

Procambial strands differentiate, connecting new leaf and bud primordia to the continuous vascular system of the main stem or branch. In eastern cottonwood (*Populus deltoides*) it has been shown that substances transported from leaves through the mature vascular system affect the development of new leaf primordia and the primary vascular tissues serving them. Research on ponderosa pine (*Pinus ponderosa*) supports the hypothesis that the number of secondary needle fasicles determines the amount of the hormone auxin transported basipetially from the elongating shoot, and that the hormone concentration determines the amount of stored carbohydrates tranported into that shoot to support growth.

Apices and tissues subjacent to them are centers of organization in the plant because cell division and differentiation occur there. Biosynthetic patterns, transport and polar transport tissues, and polar transport gradients create biochemical environments necessary for specific differentiation of the new cells. *See* APICAL DOMINANCE; APICAL MERISTEM.

Lateral meristems and secondary vascular tissues. The vascular cambium is the source of cells that, as they enlarge, cause most of the diameter growth of trees. The vascular cambium is formed from procambial strands. Cell division in the vascular cambium adds cells to both the inside and the outside of the meristem. Cells just inside the cambium differentiate into secondary xylem tissue which primarily conducts water, whereas those outside the cambium develop into secondary phloem tissue (commonly called the inner bark) which conducts a solution of photosynthetically derived sugars. Studies of stem tissues and tissue callus cultures indicate that sucrose levels and levels of the plant hormones auxin (indoleacetic acid), gibberellic acid, and cytokinins control whether xylem or phloem cells develop from cambial derivatives. The sugar-conducting phloem tissue develops in an environment low in auxin, high in gibberellic acid, and high in sucrose, whereas the water-conducting xylem develops at higher auxin, lower gibberellic acid, and moderate sucrose levels. In many species, an early season predominance of xylem production has been reported, while in others primarily phloem is produced early in the season. The diameter of xylem is controlled by the level of auxin present. There is a gradient in auxin concentration from the top of a tree, where the level is high, to the roots, where the level is lower. In the presence of high auxin levels, cells differentiate rapidly; therefore, they do not attain as large a diameter as in locations having lower auxin levels. Thus, there is a gradient in the diameter of xylem from small near the top of the shoot to progressively larger down the shoot and in the roots. *See* AUXIN; GIBBERELLIN; PHLOEM; XYLEM.

Formation of vascular cambium in roots involves a specific balance of an auxin, a cytokinin, a cyclitol, and sucrose. Vascular cambial development obtained by feeding the above mixture to the cut end of excised roots in culture is similar to that found in intact plants.

The amount of annual radial growth in trees is a function of the number of cells produced by the vascular cambium. Fast-growing hemlock (*Tsuga canadensis*) trees have higher rates of cell division and a greater number of cambial initial cells per radial file than do slow-growing trees, and the difference in rate increases during the growing season. Rates of cell elongation, deposition of secondary cell wall material, and lysis of the cytoplasm to form the mature, dead, water-conducting tracheid or vessel do not vary with tree vigor. In the latter part of the growing season, increased levels of auxin, possibly due to modification of auxin transport by increased abscisic acid (ABA) levels in the area adjacent to the cambium where xylem cell maturation is occurring, cause a delay in the onset of lysis of the cytoplasm, resulting in a longer period of secondary cell wall formation. Cells which mature during this period have comparatively thick secondary cell walls and form the latewood.

Hormones in some cases affect the development of angiosperms and gymnosperms differently. Effects of auxin on incorporation of sugar into cell walls of pine (*Pinus sylvestris*) and cottonwood (*Populus deltoides*) are quite different; thus the effects of this hormone on angiosperm and gymnosperm wood formation could also differ. Evidence indicates a probable role of abscisic acid (ABA) in regulating winter dormancy of Douglas-fir and poplar. *See* LATERAL MERISTEM.

The shoot apex of the main stem of conifers and of some dicotyledons consistently outgrows the apices of lateral branches. This causes the conical, or sometimes even spirelike, appearance termed excurrent branching. In contrast, in most dicotyledons, lateral branch apices grow as fast as or faster than the apex of the main shoot, thus causing the main stem to eventually fork repeatedly and form a large spreading crown. This is termed decurrent growth form.

Shoot growth and dormancy. With few exceptions, periods of rapid stem elongation in trees alternate with periods of either very slow or no elongation even in environments favorable for continuous growth.

Tropical trees do not exhibit seasonal growth periodicity as is commonly observed in north temperate species. A few species grow continuously, forming new leaves and stems throughout the year. Intermittently growing tropical trees form true resting buds enclosed in bud scales at the end of each shoot growth period. Some species produce growth flushes only once per year while others have many, and periodicity does not necessarily follow a seasonal pattern. A third type of growth in tropical trees is termed manifold, referring to lack of synchrony between branches on the same tree during a given period.

At germination, seedling trees produce leaf primordia and elongate them coincidentally. This process is termed free growth. For loblolly pine, once free growth ends, the correlation between the subsequent growth of a genetic family to age one and the growth to age eight is much improved. This suggests a change in the genetic control of growth between the free-growth stage and the sequential growth that follows.

During sequential growth, leaf primordia are produced on the edges of the stem apex, and there is a period of delay until they elongate, which can be as short as several days and as long as over winter.

Trees of temperate regions all display intermittent growth. Periods of shoot growth vary with species and genotype, but always occur during periods of favorable temperature and adequate moisture. Shoot-extension patterns of these trees vary with the species but can be grouped into four basic types: (1) A single flush followed by formation of a resting bud that normally remains dormant through the winter (Douglas-fir, *Pseudotsuga menziesii*, and Georgia buckeye, *Aesculus georgiana*). (2) Recurrent growth flushes with bud formation at the end of each successive flush (loblolly pine, *Pinus taeda*, and slash pine, *Pinus elliottii*). (3) A flush of growth followed by shoot-tip abortion (beech, *Fagus grandifolia*, and black willow, *Salix nigra*). (4) A sustained flush of growth with extension of all primordia into leaves prior to bud scale elongation (sweetgum, *Liquidambar styraciflua*, and yellow poplar, *Liriodendron tulipifera*).

Under favorable moisture, temperature, photoperiod, and nutrient conditions, young trees of species that normally exhibit only one growth flush can have two or three growth flushes. This also occurs in older trees, but less frequently (**Fig. 1**). Staining techniques have been developed that allow each growth flush to be identified by marking the crown region that forms at the base of the bud. Such techniques are very useful in allowing retrospective growth analysis on trees that do not develop a whorl of branches at the end of each growth flush.

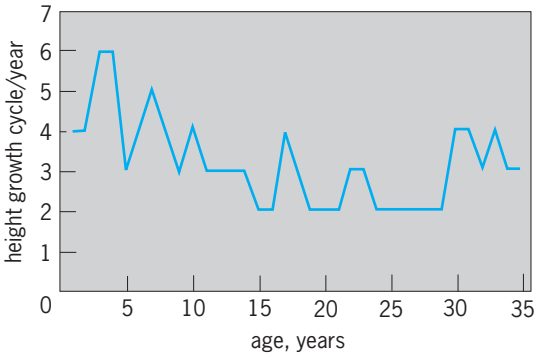

Fig. 1. Number of height growth cycles per year over the first 35 years in a loblolly pine tree. As many as four cycles can occur as late as age 33 in this tree.

In trees which normally exhibit only one growth flush per year, the end of the period of shoot extension usually occurs in mid or late summer, a considerable period of time prior to the onset of cool temperatures and frost danger of late fall and winter. After bud set, a period of rapid leaf initiation occurs at the apex, followed by a period of slower leaf initiation. The leaf primordia formed during this period will be elongated the following growing season. In early December (Douglas-fir), cell division in the developing bud ceases and the bud becomes dormant. The buds then require a substantial period of chilling before growth will resume under any environmental conditions. The amount of chilling hours (between 32 and 41°F or 0 and 5°C) required to break dormancy varies among species and between ecotypes within species (Douglas-fir approximately 2000 h and western hemlock, *Tsuga heterophylla*, 660 h).

Physiological processes controlling vegetative bud set, the onset of dormancy, and the effects of chilling on release from dormancy are under intensive study, but are not well understood. Early research indicated the accumulation of a chemical complex, termed beta-inhibitor, during onset of dormancy, and a reduction of this factor during chilling-induced release from dormancy. Subsequent analytical work indicated that ABA was present in the beta-inhibitor. Work on the Douglas-fir in Oregon showed that the concentration of ABA peaked in October and was lowest in the 3 months immediately prior to bud burst. In related studies the ratio of auxin to ABA was at its annual minimum when buds entered dormancy in December. These factors suggest a role for ABA and perhaps auxin in regulating the dormancy cycle in gymnosperms. Studies of hormonal control of dormancy in sycamore maple indicate that neither total ABA nor free or bound ABA concentrations in vegetative buds are correlated with emergence

from dormancy. This suggests that factors other than ABA in the beta-inhibitor complex change during the chilling period of dormancy conditioning prior to emergence from dormancy. If hormonal control of dormancy occurs, then the controlling factor is probably the balance between several compounds rather than simply a change in any one. Differences in mechanism could also exist between angiosperms and gymnosperms. Cell division begins in Douglas-fir in mid-March, a few weeks before budbreak occurs. At this time the chilling requirement has been fulfilled and the photoperiod is lengthening. Relationships between these events have not yet been experimentally defined, but it has been hypothesized that budbreak (extension of the overwintered primordial shoot in the bud) begins when soil temperatures warm to the point where the root system can supply the apical bud with necessary amounts of the hormone gibberellic acid. *See* ABSCISIC ACID; COLD HARDINESS (PLANT); DORMANCY; PLANT HORMONES.

Onset of reproductive competency. Trees usually grow more rapidly and longer when they are young than later when they have achieved a larger size. Research suggests that the cessation of vegetative growth early in the season allows the onset of reproductive competency.

It was suggested in the early 1960s that seasonal fluctuations in date of flushing could be one of the factors controlling the number of female strobili (cone primordia) differentiated on *Pinus banksiana* seedlings. It has been demonstrated that in *Pinus taeda* (loblolly pine), out-of-phase dormancy (induced by prolonging vegetative growth during the winter with long photoperiods and warm temperatures, then in early spring inducing vegetative budset by reducing photoperiod and temperature) will induce development of large numbers of female and male strobili at age 3 years. This finding led to the hypothesis that a critical factor in strobilus production (or, more generally, the onset of reproductive competency in trees) is the slowing of vegetative growth, to allow sufficient time for the comparatively slow process of initiation and differentiation of reproductive structures.

Growth hormones are also known to affect reproductive development. Whereas out-of-phase dormancy treatment on *Pinus taeda* induces production of both male and female strobili, treatment with gibberellin and water stress induced development of only female strobili. Gibberellin also promotes the development of mainly female strobili in most other members of the Pinacae. Such hormone treatment may act primarily by increasing the size of the vegetative bud, but it may also directly affect differentiation.

Treatments inducing earlier reproductive competency in trees are extremely valuable in the genetic improvement of commercial forest trees, as they allow shorter time periods for each breeding cycle. *See* FOREST GENETICS.

Root growth periodicity. Root apices exhibit growth periodicity, but do not appear to have the uniform and predictable dormancy cycles of the shoot. Root dormancy is preceded by cessation of root growth. In many conifers, superficial browning of roots progresses toward the root tip, culminating in formation of a layer of suberized and lignified cells—the metacutinization layer—around the root apex. Patterns vary substantially between species and locations; however, in general, root growth is at its peak for several weeks prior to shoot elongation, slows during shoot elongation, and resumes more rapid growth after shoot elongation has ceased (**Fig. 2**). Root grooth periodicity appears to occur at the individual apex level; that is, one apex can be growing while another is dormant. Sitka spruce (*Picea sitchensis*) roots respond individually to dormancy-inducing stimuli such as exogenous ABA, moisture stress, or nutrient stress. It is thus likely that the mechanism of dormancy resides in individual apices rather than in the shoot.

Correlative inhibition does occur within root systems such that if the main root apex becomes dormant or is destroyed, initiation and growth of lateral branches on that particular root are stimulated. Tree root gross morphology is heavily influenced by correlative inhibition. Initial morphology in trees under 10 years old in several species is taprooted (Douglas-fir, **Fig. 3***a*), whereas when the taproot loses viability, as by penetrating the water table or growing against an impermeable layer, lateral roots develop more rapidly (Fig. 3*b*), finally attaining the mature bell-shaped form (Fig. 3*c*). *See* ROOT (BOTANY).

Correlative relations in shoot and root development. The aboveground portion of the tree—the stem, branches, and leaves—has a geometrical structure well adapted to receiving incident sunlight and taking up carbon dioxide. Carbon dioxide is taken up through stomatal openings in leaves and, in the process, water is unavoidably lost. The belowground portion of the tree, the root system, has a geometrical structure adapted to anchorage and water and mineral nutrient uptake. While the shoot supplies the root with photosynthate, some vitamins such as thiamin, and perhaps the major portion of the hormone auxin, the root supplies the shoot with water, minerals, and some plant hormones, probably including cytokinins. Some studies suggest that some of the gibberellins translocated from the root affect shoot development.

The role of the root in water uptake is passive. The loss of water from the leaves by transpiration causes a tension in the water-conducting xylem stream which pulls water into the root and upward to the leaves. Severing the root and leaving the stem in water will increase water uptake since the resistance in the water conduction path of the root is eliminated. In nature, however, ramification of the root system through the soil is very important as it allows access to soil water. Deep root systems have been shown to be very important in delaying the onset of drought in trees subjected to dry weather. Soil water is usually available only within a few millimeters of the root, although it has been shown that fungal hyphae could extend the zone for mycorrhizal roots. *See* PLANT-WATER RELATIONS.

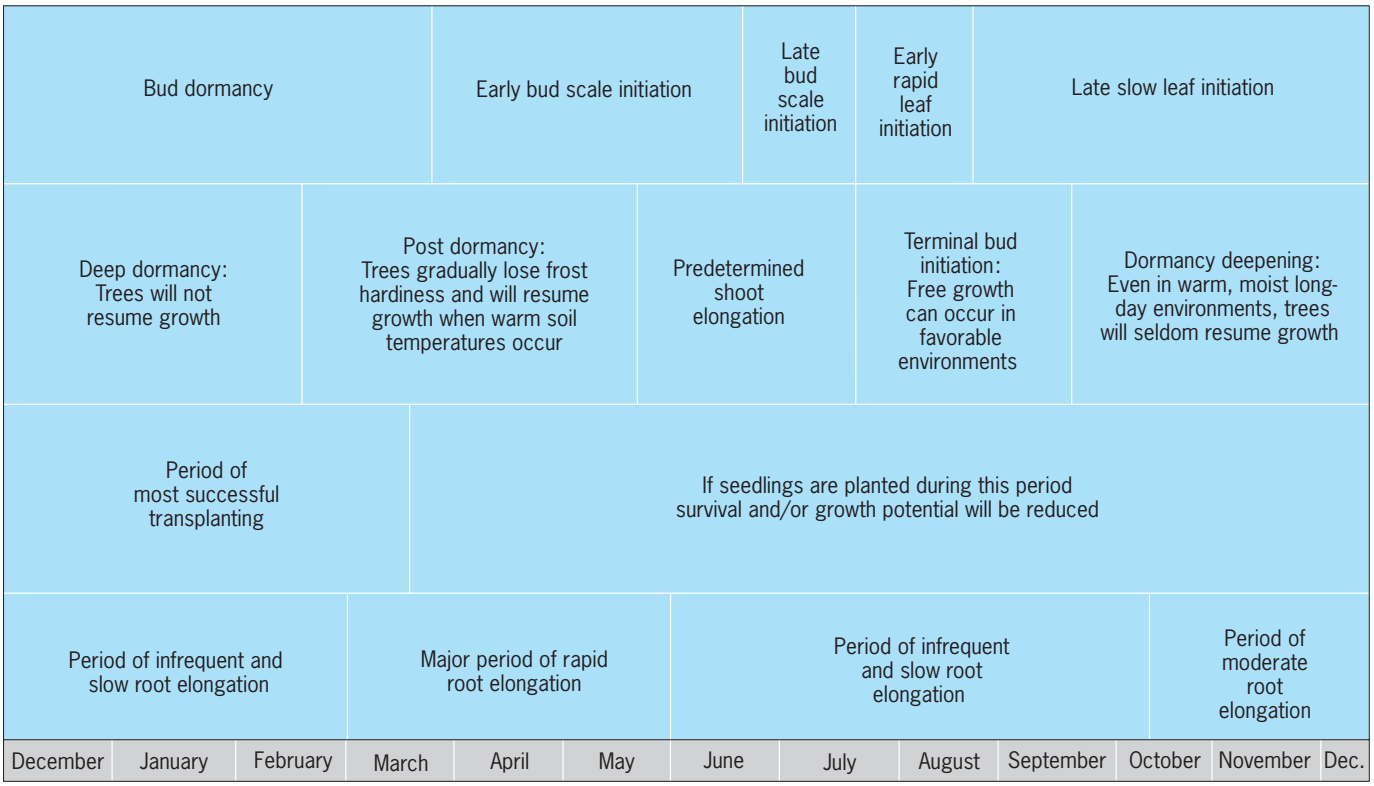

Fig. 2. Growth and dormancy cycle in Douglas-fir.

The interdependence between the shoot and root implies that they must grow in a predictable proportion to each other. The shoot-root dry weight ratio has been most often utilized to study this interdependent growth; however, directly functional parameters would be more descriptive. In many species, the amount of shoot dry matter increases more rapidly than the amount of root dry matter as seedlings grow into large trees. Root growth of peach trees can be limited by pot size. If root growth is limited, then a proportional reduction in shoot growth also occurs. Normal shoot growth can be obtained on such trees by applying the cytokinin 6-benzylamino-purine to the shoot.

Root grafting. It is often assumed that trees in a forest are distinct individuals that compete with one another for water and minerals belowground and light aboveground. However, for many years it has been observed that root grafting between trees of the same species is very common and that grafting between individuals of different species occurs but at a much lower frequency. When a root-grafted tree is cut and other trees of the union left intact, the stump of the severed tree will often remain alive and develop a ring of callus on the cut surface.

Occurrence of root grafts thus suggests that grafted root systems could act as a physiological unit, reducing the degree of belowground competition

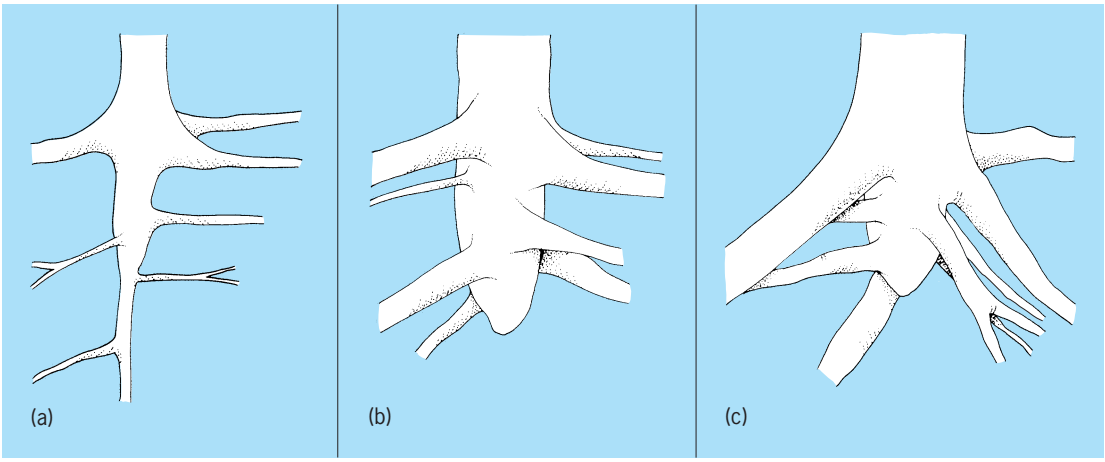

Fig. 3. Development of gross root system morphology in Douglas-fir trees. (*a*) Taprooted juvenile form. (*b*) Taproot dies. (*c*) Bell-shaped mature form.

within grafted groups of trees. In eastern white pine (*Pinus strobus*) the transport of organic substances occurs readily between grafted trees, but water and minerals, although transportable via grafts, might not be exchanged in physiologically significant volumes due to the nature of the graft union which requires water movement radially through xylem cell walls. Red pine (*Pinus resinosa*) root grafts excised from the tree and studied in the laboratory can readily transport water, requiring only a small gradient in xylem water tension to do so. It is apparent that root grafting and its effect on competitive interactions is a little-known area in which fundamental relationships are yet to be defined.

Mycorrhizal symbiosis. Mycorrhizae are symbiotic associations between plants and fungi in which short secondary lateral roots are invaded by specific fungi while the root is actively growing. Only the root cortical tissue is colonized; in ectomycorrhizae, hyphae grow over the surface of the root and between the cortical cells, whereas in endomycorrhizae, hyphae only grow directly into cortical cells. The apical meristem and vascular cylinder of the root remain fungus-free. Root cells retain a healthy appearance in mycorrhizal associations and develop no pathological symptoms. The role of mycorrhizae in ecosystem dynamics is not well described. It is known that mycorrhizal fungi obtain carbohydrates from the tree roots, and in return they increase nutrient absorption by the root, protect the root from some pathogens, and produce plant hormones. Mycorrhizal fungal hyphae can take up and transport water, but the magnitude of this function in nature is undefined. An important role of mycorrhizae is in uptake of phosphorus, which is relatively immobile in the soil. Fungal hyphae from mycorrhizal roots tremendously increase the volume of soil from which phosphorus can be obtained. It is apparent that the mineral nutrition of trees cannot be accurately described without consideration of the role of mycorrhizal fungi. *See* MYCORRHIZAE; PLANT MINERAL NUTRITION.

Certain mycorrhizal fungi allowed tree species to survive and grow well on coal mine spoils in eastern Pennsylvania that formerly could not be revegetated. This knowledge led to new approaches in revegetation of particularly adverse sites. *See* PLANT GROWTH.

William C. Carlson

Bibliography. P. J. Gregory, J. V. Lake, and D. A. Rose, *Root Development and Function: Effects of the Physical Environment*, 1987; L. R. Roberts, P. B. Gahan, and R. Aloni, *Vascular Differentiation and Plant Growth Regulators,* 1988; M. H. Zimmerman, *Xylem Structure and the Ascent of Sap,* 1983.

Tree physiology

The study of how trees grow and develop in terms of genetics; biochemistry; cellular, tissue, and organ functions; and interaction with environmental factors. While many physiological processes are similar in trees and other plants, particularly at the molecular and biochemical levels, trees possess unique physiologies that help determine their outward appearance. These physiological processes include carbon relations (photosynthesis, carbohydrate allocation), cold and drought resistance, water relations, and mineral nutrition.

Three characteristics of trees that define their physiology are longevity, height, and simultaneous reproductive and vegetative growth. Trees have physiological processes that are more adaptable than those in the more specialized annual and biennial plants. Height, exceeding 330 ft (100 m) in some species, allows trees to successfully compete for light, but at the same time this advantage creates transport and support problems. These problems were solved by the evolution of the woody stem which combines structure and function into a very strong transport system. Simultaneous vegetative and reproductive growth in adult trees causes significant competition for carbohydrates and nutrients, resulting in decreased vegetative growth. Trees accommodate both types of growth by having cyclical reproduction: one year many flowers and seeds are produced, followed by a year or more in which few or no flowers are produced.

Carbon relations. While biochemical processes of photosynthesis and carbon assimilation and allocation are the same in trees and other plants, the conditions under which these processes occur in trees are more variable and extreme. In evergreen species, photosynthesis can occur year round as long as the air temperature remains above freezing, while some deciduous species can photosynthesize in the bark of twigs and stem during the winter.

Carbon dioxide fixed into sugars moves through the tree in the phloem and xylem to tissues of high metabolism which vary with season and development. At the onset of growth in the spring, sugars are first mobilized from storage sites, primarily in the secondary xylem (wood) and phloem (inner bark) of the woody twigs, branches, stem, and roots. The sugars, stored as starch, are used to build new leaves and twigs, and if present, flowers. Once the new leaves expand, photosynthesis begins and sugars are produced, leading to additional leaf growth. Activation of the vascular cambium occurs at the same time, producing new secondary xylem and phloem. In late spring, the leaves begin photosynthesizing at their maximum rates, creating excess sugars which are translocated down the stem to support further branch, stem, and root growth. From midsummer through fall until leaf abscission (in deciduous trees) or until temperatures drop to freezing (in evergreen trees), sugars replenish the starch used in spring growth. Root growth may be stimulated at this time by sugar availability and warm soil temperatures. Throughout the winter, starch is used for maintenance respiration, but sparingly since low temperatures keep respiration rates low. *See* PHLOEM; PHOTOSYNTHESIS; XYLEM.

In adult trees, reproductive structures (flowers in angiosperms or strobili in gymnosperms) develop along with new leaves and represent large carbohydrate sinks. Sugars are preferentially utilized at

the expense of leaf, stem, and root growth. This reduces the leaf area produced, affecting the amount of sugars produced during that year, thereby reducing vegetative growth even further. The reproductive structures are present throughout the growing season until seed dispersal and continually utilize sugars that would normally go to stem and root growth.

Cold resistance. The perennial nature of trees requires them to withstand low temperatures during the winter. At higher latitudes and elevations, temperatures can reach −58°F (−50°C). Trees that evolved in these regions have physiological processes that protect their cells from damage to temperatures as low as −328°F (−200°C), whereas tropical species lacking these processes exhibit damage at above-freezing temperatures. Cell damage occurs from ice crystal formation within the cell, which disrupts the plasma membrane. Trees develop resistance to freezing through a process of physiological changes beginning in late summer. The degree of cold resistance within a tree depends on the tissue; buds and twigs typically tend to be more resistant than roots.

A tree goes through three sequential stages to become fully cold resistant. The process involves reduced cell hydration along wtih increased membrane permeability. The first stage is initiated by shortening days and results in shoot growth cessation, bud formation, and metabolic changes. Trees in this stage can survive temperatures down to 23°F (−5°C). The second stage requires freezing temperatures which alter cellular molecules. Starch breakdown is stimulated, causing sugar accumulation. Trees can survive temperatures as low as −13°F (−25°C) at this stage. The last stage occurs after exposure to very low temperatures (−22 to −58°F or −30 to −50°C), which increases soluble protein concentrations that bind cellular water, preventing ice crystallization. Trees can survive temperatures below −112°F (−80°C) in this stage. A few days of warmer temperatures, however, causes trees to revert to the second stage.

Water relations. Trees are likely to experience drought conditions during their lifetime. Unlike annual plants that survive drought as seeds, trees have evolved traits that allow them to avoid desiccation. These traits include using water stored in the stem, stomatal closure, and shedding of leaves to reduce transpirational area. All the leaves can be shed and the tree survives on stored starch. Another trait of some species is to produce a long tap root that reaches the water table, sometimes tens of meters from the soil surface.

On a daily basis, trees must supply water to the leaves for normal physiological function. If the water potential of the leaves drops too low, the stomata close, reducing photosynthesis. To maintain high water potential, trees use water stored in their stems during the morning which is recharged during the night. *See* PLANT-WATER RELATIONS.

Transport and support. Trees have evolved a means of combining long-distance transport between the roots and foliage with support through the production of secondary xylem (wood) by the vascular cambium. In older trees the stem represents 60–85% of the aboveground biomass (and about 40–50% of the total tree biomass). However, 90% of the wood consists of dead cells. These dead cells function in transport and support of the tree. As these cells develop and mature, they lay down thick secondary walls of cellulose and lignin that provide support, and then they die with the cell lumen becoming an empty tube. The interconnecting cells provide an efficient transport system, capable of moving 106 gal (400 liters) of water per day. The living cells in the wood (ray parenchyma) are the site of starch storage in woody stems and roots.

Further specialization within the secondary xylem results from variable sugar levels, and possibly hormonal levels, at the time of development. During spring growth when sugars are being used primarily for leaf growth, xylem cells form thinner secondary walls that have large lumens. These cells facilitate rapid water movement to the expanding foliage. Later in the year as more sugars are transported to the stem and roots, these cells form thick secondary walls and small lumens which function for support. *See* PLANT TRANSPORT OF SOLUTES.

Mineral nutrition. Mineral nutrients are required in proper concentrations and in relative proportions to one another by trees for normal growth and development. Nutrient deficiencies are similar in trees and other plants because of the functions of these nutrients in physiological processes. Tree nutrition is unique because trees require lower concentrations, and they are able to recycle nutrients within various tissues.

Trees adapt to areas which are low in nutrients by lowering physiological functions and slowing growth rates. In addition, trees allocate more carbohydrates to root production, allowing them to exploit large volumes of soil in search of limiting nutrients. Proliferation of fine roots at the organic matter–mineral soil interface where many nutrients are released from decomposing organic matter allows trees to recapture nutrients lost by leaf fall.

Internal cycling keeps nutrients from leaving the tree (then the nutrients could be lost to neighboring trees or competing vegetation) and accounts for a majority of the nutrients used annually for growth. Trees translocate nutrients out of leaves during senescence, prior to abscission. These nutrients can be translocated to newly expanding leaves or to storage sites located in either the twigs (deciduous trees) or younger foliage (evergreen trees). As an example, nitrogen is readily transported as amino acids during leaf senescence and, in deciduous trees, is stored as a special protein in twig bark. This process is triggered by shortening days in the fall. As buds enlarge in the spring, these proteins break down into their constituent amino acids and are translocated to newly developing leaves. *See* PLANT MINERAL NUTRITION; PLANT PHYSIOLOGY; TREE.

Jon D. Johnson

Bibliography. P. J. Kramer and T. T. Kozlowski, *Physiology of Woody Plants*, 1979; A. S. Raghavendra, *Physiology of Trees*, 1991.

Trematoda

A loose grouping of acoelomate, parasitic flatworms of the phylum Platyhelminthes formerly accorded class rank and containing the subclasses (or orders) Digenea, Monogenea, and Aspidobothria. Although neither the name Trematoda nor the classification scheme appears as valid as formerly in demonstrating helminth phylogenetic relationships (in light of recent information, covered in the Taxonomy section below), a discussion of the three older "trematode" groups just mentioned is profitable. These organisms commonly occur as adults in or on all vertebrate groups. They exhibit cephalization, bilateral symmetry, and well-developed anterior and ventral, or anterior and posterior, holdfast structures. Electron-microscope studies of some parasitic platyhelminths by D. L. Lee in 1966 and K. M. Lyons in 1970 indicate that the outer covering of these worms, once thought to be a secreted, nonliving cuticle, is probably a living epidermis. The mouth is anterior, and usually a blind, forked gut occurs, as well as three muscle layers. The excretory system consists of flame cells and collecting tubules. These animals are predominantly hermaphroditic and oviparous with operculated egg capsules. The life histories of the Digenea are complex, while those of the Monogenea and Aspidobothria are simple.

Taxonomy. Systematic studies are based mainly on comparative morphology of adults and larvae, although life history and ecological data are also used. In 1957 G. R. LaRue presented an interesting, detailed history of trematode systematics and a classification of digenetic trematodes based upon larval characteristics.

Also in 1957 B. E. Bychowsky, a recognized leader in parasite research, reiterating his ideas of 1937, commented extensively on trematode classification. J. Llewellyn in 1965 and H. W. Manter in 1969 also examined their systematics of the groups involved. H. W. Stunkard presented in 1962 a revised classification which was quite similar to that of L. H. Hyman in 1951:

Class Trematoda
- Order Monogenea
  - Suborder: Monopisthocotylea
    - Polyopisthocotylea
- Order: Aspidobothria
  - Digenea

It is noteworthy that Stunkard separated the aspidobothrids and the digeneids more distinctly as follows:

Class Trematoda
- Subclass Pectobothridia
  - Order: Monopisthocotylea
    - Polyopisthocotylea
- Subclass Malacobothridia
  - Order: Aspidobothrea
    - Digenea

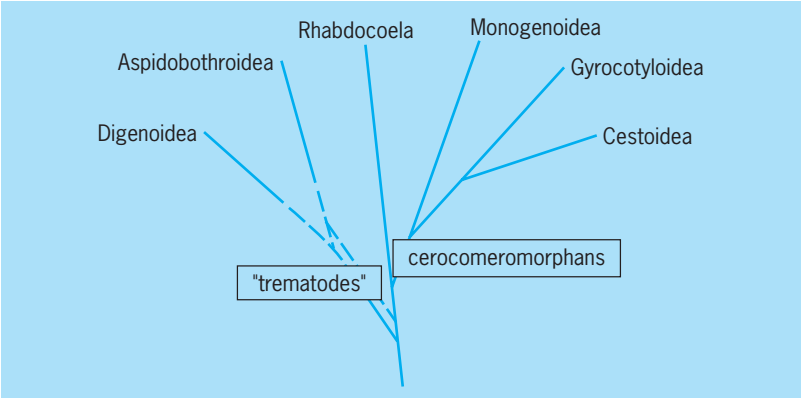

Fig. 1. Probable lines of evolution of Digenoidea, Aspidobothroidea, Monogenoidea, Gyrocotyloidea, and Cestoidea from rhabdocoel turbellarians as based on recent information.

In 1965 Llewellyn, basing some of his conclusions on Bychowsky's earlier works, recommended separation of the grouping Trematodes, which included the classes Digenoidea and Aspidobothroidea, from those platyhelminth parasites which generally have hooked larvae. Within the grouping Cercomeromorphae, he housed Monogenoidea, Gyrocotyloidea, Cestodaria, and Cestoda, which appear to have been accorded class ranking (**Fig. 1**).

Manter's classification encompasses many of the concepts of these later workers, differing somewhat in rank accorded Aspidogastrea (Aspidobothrea) and Digenea, which he called subclasses.

After extensive research in these laboratories the author is inclined to place some credence in the "cercomeromorph concept" and, keeping this in mind, to prefer the arrangement of Manter, rearranging the order so that the class Cestoidea are nearer the Gyrocotyloidea and Monogenoidea. Without more experience and evidence the author prefers not to address the class Mesozoa, a highly aberrant grouping of organisms. The position of Aspidobothroidea is less clear than some of the other groupings because of their general anatomical similarity to the Digenoidea and their monogeneid life histories. However, aspidobothreids seem sufficiently different from both groups to deserve equal status. The arrangement would appear as follows:

Class: Monogenoidea
- Gyrocotyloidea
- Cestoidea
- Subclass: Amphilinidea
  - Cestoda

Class: Digenoidea
- Aspidobothroidea

The current phylogenetic concept is represented in Fig. 1.

The Digenoidea and Aspidobothroidea are probably phylogenetically older than the Monogenoidea, since they are entoparasitic, parasitize both invertebrates and vertebrates, and have more complicated life histories. Since early mollusks were marine

forms, these trematodes are probably marine in origin. Monogenoidea probably originated in the sea also because a greater divergence of forms has taken place among marine species. All trematodes are strikingly similar to rhabdocoel turbellarians, from which they probably evolved. Indeed, some turbellarians have simple sucking holdfasts and tend toward parasitic habits. The major differences between free-living turbellarians and trematodes are the addition of complex holdfasts, development of a cuticle, loss of photoreceptors, and an increase in reproductive capacities with concomitant alterations in the reproductive system, all of which are modifications for the parasitic habit. *See* TURBELLARIA.

Morphology. Although the three classes share a common internal morphological plan, they differ considerably externally. The Digenoidea are commonly more elongate, with continuous outlines, while the Monogenoidea have enlarged anterior and posterior holdfasts and more irregular outlines, and the Aspidobothroidea are modified ventrally. Digeneids usually have oral and ventral suckers as adult holdfasts. Aspidobothreids have large, ventral, adhesive disks with many depressions, or alveoli. Monogenoidea have paired suckers, or adhesive glands, anteriorly and armed suckers or wedges, called opisthaptors, posteriorly. Though all classes have longitudinal, transverse, and oblique muscle layers, monogeneids usually have more complex arrangements of longitudinal muscle bands to operate the holdfasts. Digestive systems consist of an anterior mouth, muscular pharynx, short esophagus, and forked, blind gut, with the exceptions of *Udonella* and the Aspidobothroidea, which have single, median digestive sacs. In some digeneids and many monogeneids the main intestinal branches are much ramified. The nervous system has an enlarged anterior esophageal ganglion or brain. Dorsal nerve trunks usually lacking in Monogenoidea, as well as ventral, longitudinal, and transverse trunks, also occur. Photosensitive eyespots occur in the larvae of many Digenoidea and Monogenoidea, and in some adult monogeneids. Digenoidea usually have a single nephridiopore, while monogeneids possess two. Aspidobothroidea feature both conditions. Most trematodes possess turbellarianlike hermaphroditic reproductive systems with single testis or multiple testes, a single ovary, a common gonopore, and a protrusible, or eversible, copulatory organ. A few digeneids, the blood flukes, have separate sexes. All three groups may have homologous accessory ducts running from the female tracts to the outside, known as Laurer's canal in the Digenoidea and Aspidobothroidea; or they may have accessory tracts to the gut, like the genitointestinal canal of the Monogenoidea. Cross-fertilization and self-fertilization are possible and, depending on the availability of mates, both occur, but the former probably predominates. Oviparity is the rule, but a few monogeneids, such as *Gyrodactylus* and *Isancistrinae*, bear living young. The ciliated embryo, which usually does not occur in the Aspidobothroidea, is enclosed in an operculated egg capsule containing nourishing vitelline cells.

Life history. Egg production of most digeneids is high and, coupled with asexual multiplication of the larval stages, results in enormous numbers of young. Most probably succumb to the hazards of their complicated life cycles; therefore large numbers are necessary. The Monogenoidea produce fewer embryos (**Fig. 2**).

The complicated life cycles of the Digenoidea involve asexual reproduction in the first intermediate host, which is generally a mollusk, and sexual reproduction in the final hosts. Aspidobothroidea usually develop directly, and they find lodgment in the final molluscan host. Some aspidobothreids, however, have an intermediate molluscan, or decapod, host with turtles or fishes as the final host. Monogeneids possess ciliated larvae, which undergo simple metamorphosis into the juvenile form on a single host.

Physiology. Comparatively little is known of helminth physiology. Such studies are complicated by the parasitic habit, small size, and complex life histories of the worms. Helminth metabolism is not as simple as once believed, and involves great adaptability to varying oxygen tensions and pH ranges, among other factors. Many digeneids are capable of respiration under low-oxygen or anaerobic conditions. The physiological activities of larvae and juveniles are probably very different from those of the adults. Trematodes feed upon host tissues, body fluids, and exudates, and both extracellular and intracellular digestion takes place. Transcutaneous absorption of nutrients may occur. Efforts at culturing trematodes on artificial media have met with some success, and several investigators are pursuing transplantation studies.

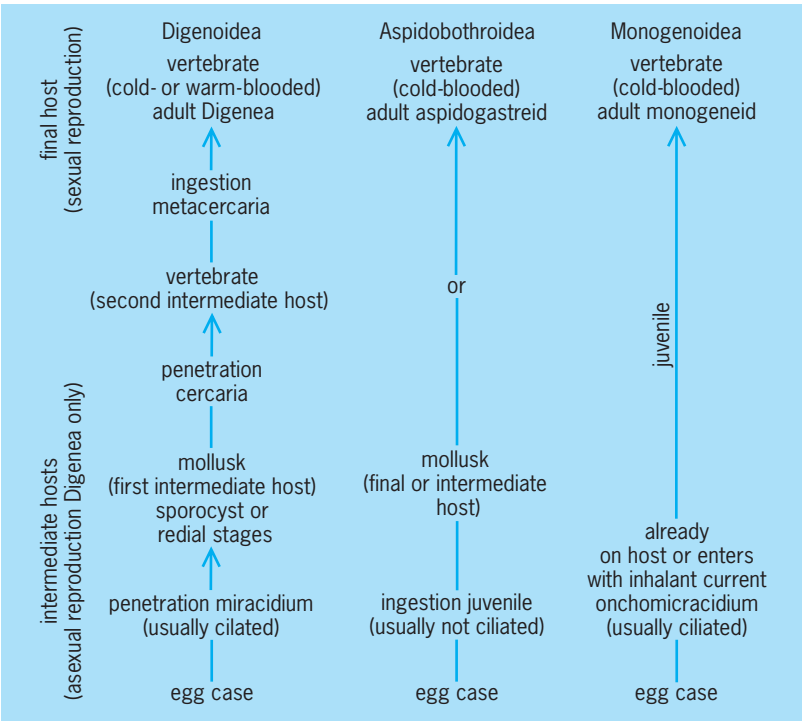

Fig. 2. Stylized diagrams of the life histories of three trematode classes.

Ecology. Trematodes parasitize a wide variety of invertebrates and vertebrates and occupy almost every available niche within these hosts. The adaptations demanded of the worms for survival are as varied as the characteristics of the microhabitats. Over the millions of years of coevolution of the hosts and their parasites, delicate balances have, for the most part, been attained, and under normal conditions it is probable that trematodes rarely demand more than the host can supply without undue strain. This has involved adjustments between the antigenic properties of the parasite and the defense reactions of the host. In many entoparasitic trematodes, delicate balances have developed between the protein digestive enzymes of the host and the resistant properties of the cuticle and other tissues of the parasite. In *Paragonimus* the cuticle actually appears to be digested by the host but is continually renewed by the parasite. Monogeneids do not have the same problems but have had to develop efficient adhesive mechanisms to keep from being swept from the surfaces of their hosts. This struggle has produced many interesting holdfast organs.

It seems axiomatic that the host must survive until the parasite can again gain access to another host or until the life cycle is completed. Those parasites which cause the least disruption of the host's activities are probably the oldest as well as the most successful. Immunities are sometimes developed by the hosts. Many trematodes seem to possess such rigid requirements and such responses to particular hosts that host specificity is a phenomenon of considerable significance. Monogeneids appear more host-specific than digeneids, and aspidobothreids seem less specific than both. Trematodes are of considerable veterinary and medical importance because under certain conditions they cause debility, even death. *See* ASPIDOGASTREA; DIGENEA; MONOGENEA; PLATYHELMINTHES. William J. Hargis, Jr.

Bibliography. J. G. Baer, *Ecology of Animal Parasites*, 1951; B. E. Bychowsky, *Monogenetic Trematodes, Their Classification and Phylogeny*, 1957; B. E. Bychowsky, Ontogeny and phylogeny of parasitic platyhelminths, *Izv. Akad. Nauk. SSSR, Ser. Biol.*, no. 4, 1353–1383, 1937; B. Dawes, *The Trematoda*, 1946; L. H. Hyman, *The Invertebrates*, vol. 2, 1951; A. V. Ivanov, *Udonella caligorum* Johnston, 1835, a representative of a new class of flatworms, *Zool. Zh.*, 31(2):175–178, 1952; G. R. LaRue, The classification of digenetic Trematoda: A review and a new system, *Exp. Parasitol.*, 6(3):306–349, 1957; D. L. Lee, The structure and composition of the helminth cuticle, *Adv. Parasitol.*, 4:187–254, 1966; J. Llewellyn, The evolution of parasitic platyhelminths, *Evolution of Parasites: 3d Symposium of the British Society for Parasitology*, pp. 47–78, 1965; K. M. Lyons, The fine structure and function of the adult epidermis of two skin parasitic monogeneans, *Entobdella soleae* and *Acanthocotyle elegans, Parasitology*, 60:39–52, 1970; H. W. Manter, Problems in systematics of trematode parasites, *Problems in Systematics of Parasites*, pp. 91–105, 1969; H. W. Stunkard, *Taeniocotyle* nom. nov. for *Macraspis* Olsson, 1869, preoccupied, and systematic position of the Aspidobothrea, *Biol. Bull.*, 122(1):137–148, 1962; S. Yamaguti, *The Digenetic Trematodes of Vertebrates*, in *Systema Helminthum*, vol. 1, pts. 1 and 2, 1958; S. Yamaguti, *Monogenea and Aspidocotylea*, in *Systema Helminthum*, vol. 4, 1963.

Tremolite

The name given to magnesium-rich monoclinic calcium amphibole $Ca_2Mg_5Si_8O_{22}(OH)_2$. The mineral is white to gray, but colorless in section, and optically negative. It usually exhibits long prismatic crystals with prominent (110) amphibole cleavage. Unlike other end-member compositions of the calcium amphibole group, very pure tremolite is found in nature. Substitution of Fe for Mg is common, but pure ferrotremolite, $Ca_2Fe_5Si_8O_{22}(OH)_2$, is rare. Intermediate compositions between tremolite and ferrotremolite are referred to as actinolites, and are green in color and encompass a large number of naturally occurring calcium amphiboles. The substitution of Na, Al, and Fe^{3+} ions into the amphibole structure is common in actinolites. The nature of these substitutions is complex and leads, under some conditions, to miscibility gaps between actinolites and aluminous calcium amphiboles called hornblendes. For the most part, however, the physical and chemical variations between actinolite and hornblende are continuous so that an arbitrary actinolite-hornblende division is necessary. Most of the amphibole classifications divide actinolite from the hornblende series on the basis of Al substitution for Si with actinolite containing less than 0.5 atom of Si replaced by Al per formula unit. *See* HORNBLENDE.

The basic building block for the tremolite-actinolite crystal structure is the silicon tetrahedron. In all amphiboles, double chains of SiO_4 tetrahedrons are formed through joining two or three tetrahedrons at their corners by consecutively sharing two or three oxygens in alternating fashion along the entire chain length. These double chains form two anionic layers: a nearly coplanar layer of oxygens at the base of the tetrahedrons, and a second layer of oxygens and associated OH and F along the apices of the tetrahedrons. Silicon, and to a lesser extent aluminum, atoms lie between these basal and apical anionic layers in fourfold coordination. The double chains are arranged so that along the *b* crystallographic direction the basal oxygen layer of one chain is approximately coplanar with the apical anionic layer of the adjacent chains. Along the *a* crystallographic direction, the apical oxygen layers of adjacent chains face each other, as do the basal oxygen layers. The *c* axis parallels the long axis of the double chains. The double-chain structure is held together by bonding apical oxygens of facing double chains to a cation strip of 5Mg and 2Ca which have six- to eightfold coordination. Tremolite is usually described in terms of the (001) face-centered cell (space group C2/m). Typical lattice parameters for tremolite are

$a = 9.83$; $b = 18.05$; and $c = 5.27$ (all in angstroms; 10 Å = 1 nanometer); and β angle = 104.5°. Actinolite, due to its higher Fe^{2+} content, has a larger cation strip and hence a proportionately larger b dimension.

Tremolite in pure form is a product of thermal and regional metamorphism of siliceous dolomites and marbles as shown by reaction (1). In similar rocks

$$\underset{\text{Dolomite}}{5CaMg(CO_3)_2} + \underset{\text{Quartz}}{8SiO_2} + H_2O \longrightarrow \underset{\text{Tremolite}}{Ca_2Mg_5Si_8O_{22}(OH)_2} + \underset{\text{Calcite}}{3CaCO_3} + 7CO_2 \quad (1)$$

at higher grades of metamorphism, in the presence of both calcite and quartz, tremolite breaks down to form diopside, as shown by reaction (2). Actinolites, owing to a more variable chemistry, are more

$$\underset{\text{Tremolite}}{Ca_2Mg_5Si_8O_{22}(OH)_2} + \underset{\text{Calcite}}{3CaCO_3} + \underset{\text{Quartz}}{2SiO_2} \longrightarrow \underset{\text{Diopside}}{5CaMgSi_2O_6} + 3CO_2 + H_2O \quad (2)$$

ubiquitous than tremolite in occurrence. They are most commonly found in regionally metamorphosed mafic igneous rocks such as basalts and are known to occur in a wide range of pressure conditions. Both tremolite and actinolite can form through the breakdown of olivine and pyroxenes in regionally metamorphosed ultrabasic rocks; associated minerals are talc, chlorite, and carbonates. Actinolitic amphiboles formed as a breakdown product of pyroxenes are referred to as uralite. Intergrown fibrous crystals of tremolite are known as nephrite, a form of jade widely used for centuries, in making of artifacts and jewelry. Highly fibrous tremolite is used in commercial asbestos. *See* AMPHIBOLE; ASBESTOS; DIOPSIDE; JADE; METAMORPHISM.

Barry L. Doolan

Trepostomata

An extinct order of bryozoans in the class Stenolaemata. Trepostomes possess generally robust colonies, composed of tightly packed, moderately complex, long, slender, tubular or prismatic zooecia, with solid calcareous zooecial walls. Colonies show a moderately gradual transition from endozone to exozone regions, and they are exclusively free-walled. *See* BRYOZOA; STENOLAEMATA.

Morphology. Trepostome colonies range from small and delicate to large and massive; they can be thin to thick encrusting sheets; tabular, nodular, hemispherical, or globular masses; or bushlike or frondlike erect growths. Most colonies are divisible into very distinct endozone and exozone regions, with that portion of each zooecium lying within the endozone relatively long compared with that within the exozone in most erect colonies. Trepostome colonies usually bear mesopores (or exilapores) but lack other polymorphs. There is no unallocated colonial skeleton (extrazooidal skeleton) secreted between zooecia where colonies were actively growing. However, secondary extrazooidal skeleton may cover over colony surfaces in older areas. Colony surfaces may be smooth, but more commonly there are regularly spaced maculae (spots), generated by clusters of polymorphs, enlarged zooecia, or even groups of normal zooecia. In many trepostomes, maculae generate small, conspicuous bumps (monticules).

Individual trepostome zooecia are long, straight to markedly curved, cylindrical to prismatic tubes. The walls of adjacent zooecia are fused, forming one continuous skeletal element in which a boundary between zooecia may or may not be visible. These walls are generally extensively calcified, nonporous, pronouncedly laminated in microstructure, thin in the endozone, and thick in the exozone. Commonly, colony surfaces have small to large spines projecting from the skeletal surface; these are seen as conspicuous rod-shaped structures (styles) in thin slices cut through the skeleton. Most zooecia are crossed inside by diaphragms; however, in some trepostomes they are crossed by hemiphragms, heterophragms, or cystiphragms. Apertures are round to polygonal in outline.

History and classification. Apparently exclusively marine, the trepostomes first appeared about the start of the Middle Ordovician; they apparently share a common ancestor with cystoporates based on similar ranges of colony forms, zooidal arrangements and shapes, and simple colony organization. The trepostomes immediately rose to dominance within the bryozoan fauna and are among the most common macrofossils of the early Paleozoic; in places, trepostomes contributed to construction of small reefs, as well as being abundant in level-bottom environments. They remained abundant through Silurian time, declined during the Devonian, and died out in the Late Triassic. Their assignment to family, genus, and species usually requires preparation of orientated thin sections or peels that clearly show details of colony interiors. *See* CYSTOPORATA.

Roger J. Cuffey; Frank K. McKinney

Bibliography. R. S. Boardman, Indications of polypides in feeding zooids and polymorphs in lower Paleozoic Trepostomata (Bryozoa), *J. Paleont.*, 73:803–815, 1999; R. S. Boardman, A. H. Cheetham, and A. J. Rowell (eds.), *Fossil Invertebrates*, Blackwell Scientific Publications, Palo Alto, CA, 1987; P. D. Taylor and G. P. Larwood, Major evolutionary radiations in the Bryozoa, in P. E. Taylor and G. P. Larwood (eds.), *Major Evolutionary Radiations*, pp. 209–233, Clarendon Press, Oxford, 1990.

Trestle

A succession of towers of steel, timber, or reinforced concrete supporting the horizontal beams of a roadway, bridge, or other structure. Little distinction can be made between a trestle and a viaduct, and the terms are used interchangeably by many engineers. A viaduct is defined as a long bridge consisting of a series of short concrete or masonry spans

supported on piers or towers, and is used to carry a road or railroad over a valley, gorge, another roadway, or across an arm of the sea. A viaduct may also be constructed of steel girders and towers. It is even more difficult to draw a distinction between a viaduct and a bridge than it is between a viaduct and a trestle. *See* BRIDGE.

A trestle or a viaduct usually consists of alternate tower spans and spans between towers. For low trestles the spans may be supported on bents, each composed of two columns adequately braced in a transverse direction. A pair of bents braced longitudinally forms a tower. The columns of one bent of the tower are supported on planed base plates or movable shoes to allow horizontal movement in the longitudinal direction of the trestle. Struts connect the column bases and force the movable shoes to slide. The width of the base of a bent is usually not less than one-third the height of the bent. This width is sufficient to prevent excessive uplift at windward columns when the trestle is unloaded. *See* STRUCTURE (ENGINEERING); TOWER. Charles M. Antoni

Triassic

The oldest period of the Mesozoic Era, encompassing an interval between about 248 and 206 million years ago (Ma). It was named in 1848 by F. A. von Alberti for the threefold division of rocks at its type locality in central Germany, where continental redbeds and evaporites of the older Buntsandstein and younger Keuper formations are separated by marine limestones and marls of the Muschelkalk formation. These carbonates were laid down in a shallow tongue of the Tethys seaway that extended from the Himalayas through the Middle East to the Pyrenees,

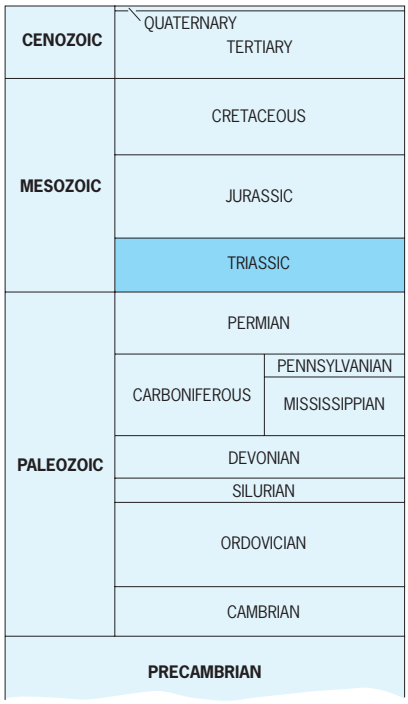

where more than 10,000 ft (3000 m) of carbonate were deposited. The German section was an unfortunate choice because it is atypical of other Triassic sections and carries a sparsely preserved fossil record. It was subsequently replaced by a marine carbonate sequence in the Alps as the standard for global Triassic reference and correlation. The North American standard marine section is in the western Cordilleras of British Columbia and the Sverdrup Basin of the Arctic. *See* MESOZOIC.

Major events. Triassic strata record profound paleontologic changes that reflect major physical changes in Earth history. Two of the five most catastrophic extinctions of the Phanerozoic Eon mark the beginning and end of the Triassic. More than 50% of all Permian families died out at the beginning of the period, including 85–90% of all marine species and 75% of land species; and more than 50% of all marine genera became extinct at the end of the period. The Triassic was also when many new families of plants and animals evolved, including the earliest known mammals.

As a very brief interval of geologic time (about 40 million years), the Triassic Period uniquely embraces both the final consolidation of Pangaea and the initial breakup of the landmass, which in the Middle Jurassic led to the opening of the Central Atlantic Ocean and formation of modern-day continental margins. The Triassic marks the beginning of a new Wilson cycle of ocean-basin opening through lithospheric extension and oceanic closing through subducting oceanic lithospheres along continental margins. The cycle was named for J. Tuzo Wilson, a pioneer of modern plate tectonic theory. The initial breakup of Pangaea occurred in the western Tethys (precursor of the Mediterranean Sea) between Baltica and Africa and in eastern Greenland between Baltica and the North American craton. Rifting then proceeded into the Central Atlantic, separating the North American and African cratons that led to the separation of Laurasia from Gonwanaland (**Fig. 1**). Rifting also occurred in Argentina, east Africa, and Australia. In the central Atlantic region, extensional tectonics was accompanied by a huge outpouring of continental flood basalts, forming the Central Atlantic magmatic province (CAMP), whose remnants are now found as feeder dikes and flood basalts on four circum-Atlantic continents, separated by thousands of miles of younger basalts of the oceanic crust (**Fig. 2***a*). *See* LITHOSPHERE; PLATE TECTONICS.

Final consolidation of Pangaea. The initial consolidation of Pangaea, which was marked by the formation of the Allegheny-Mauritanide-Variscan mountain chain in the middle Carboniferous (320 Ma), resulted from the collision of Gondwanaland and the combined Laurasia-Baltica-Siberian-Kazakhstania landmass (Fig. 1). Major plate accretion continued into the Middle-to-Late Triassic (230 ± 5 Ma), when southern China and Cimmeria (Asia Minor) were sutured to the northern margin of the Tethys seaway. Smaller terranes, called suspect or exotic, were also accreted to the western margin of North

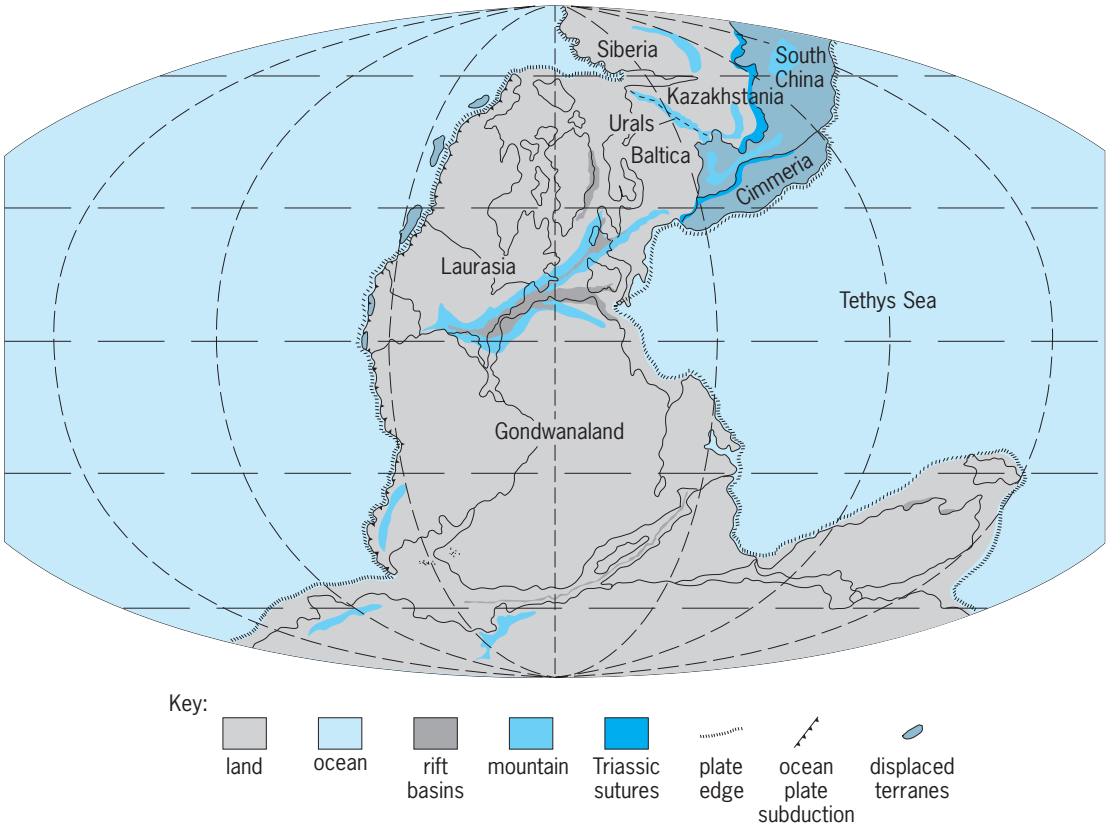

Fig. 1. Paleogeography of the Late Triassic Period: after the accretion of south China and Cimmeria (Turkey, Iran, and Tibet) to Laurasia; during the Incipient rifting of Pangaea in eastern North America and northwest Africa along the Allegheny-Mauritanide-Variscan orogeny; and concurrent with oceanic subduction and formation of deep-sea trenches and magmatic arc along the western plate boundary of North America. (*After R. K. Bambach, C. R. Scotese, and A. M. Ziegler, Before Pangea: The geographics of the Paleozoic world, Amer. Sci., 68:26–38, 1980*)

America at this time. Disconnected patches of Triassic strata occur from California through western British Columbia into Alaska, where they appear to be displaced island-arc terranes, microcontinents, and ocean-ridge segments, as inferred from paleomagnetic data in the lavas and by the exotic character of their Permian faunas. *See* PALEOGEOGRAPHY; PALEOMAGNETISM.

Pangaean supercontinent. The final phase of deformation produced a broadly convex continental plate that extended from the north to the south paleo poles, covered about 25% of the Earth's surface, and was surrounded by a global ocean called Panthalassa. It had a central arch standing about 1 mi (1.6 km) high and an average elevation of more than 4300 ft (1300 m) above the early Mesozoic sea level (Fig. 1). Because of its size, location, and pronounced orographic peaks that probably rivaled the Himalayas, the Pangaea landmass had a major impact on global climates. During the Middle Triassic, Florida lay about 5° south of the Equator, whereas Grand Banks (now off southeastern Newfoundland) was located about 20°N. Pangaea's climatic zones ranged from tropical savanna along its extensive coasts to arid and semiarid across its vast interior. *See* DESERT; SAVANNA; SUPERCONTINENT.

As the plate migrated north, transgressing about 10° of latitude between the Middle Triassic and Middle Jurassic, the plate was subjected to increased aridity as it moved under the influence of the subtropical high-pressure cell. Because of its large size, the landmass must have been subjected to monsoon circulation. Winters along the future central Atlantic probably were dominated by subtropical high-pressure cells bringing in cool dry air from aloft, whereas summers were dominated by equatorial low-pressure systems bringing in warm moist air from the Tethys seaway to the east. As moist air was uplifted almost 1.2 mi (2 km) over the Alleghenian-Variscan chain, it would have cooled adiabatically, yielding rainwater that fed major rivers (for example, the Congo River) flowing thousands of miles away from the axis of uplift across broad alluvial plains to the coastal regions of Alaska, Patagonia, India, and Siberia. *See* MONSOON METEOROLOGY.

With the onset of rifting in the Late Triassic and subsequent topographic changes, small ephemeral streams flowed into the rift valleys, creating huge lakes that may have been comparable in size to present Lake Tanganyika of the East African rift system. Where air masses descended into low-lying rift basins, along the Central Atlantic axis, they warmed adiabatically, causing evaporation and precipitation of evaporite minerals (for example, halite, gypsum, and anhydrite) in marginal epicontinental seas and in continental lacustrine basins. The Triassic and Lower

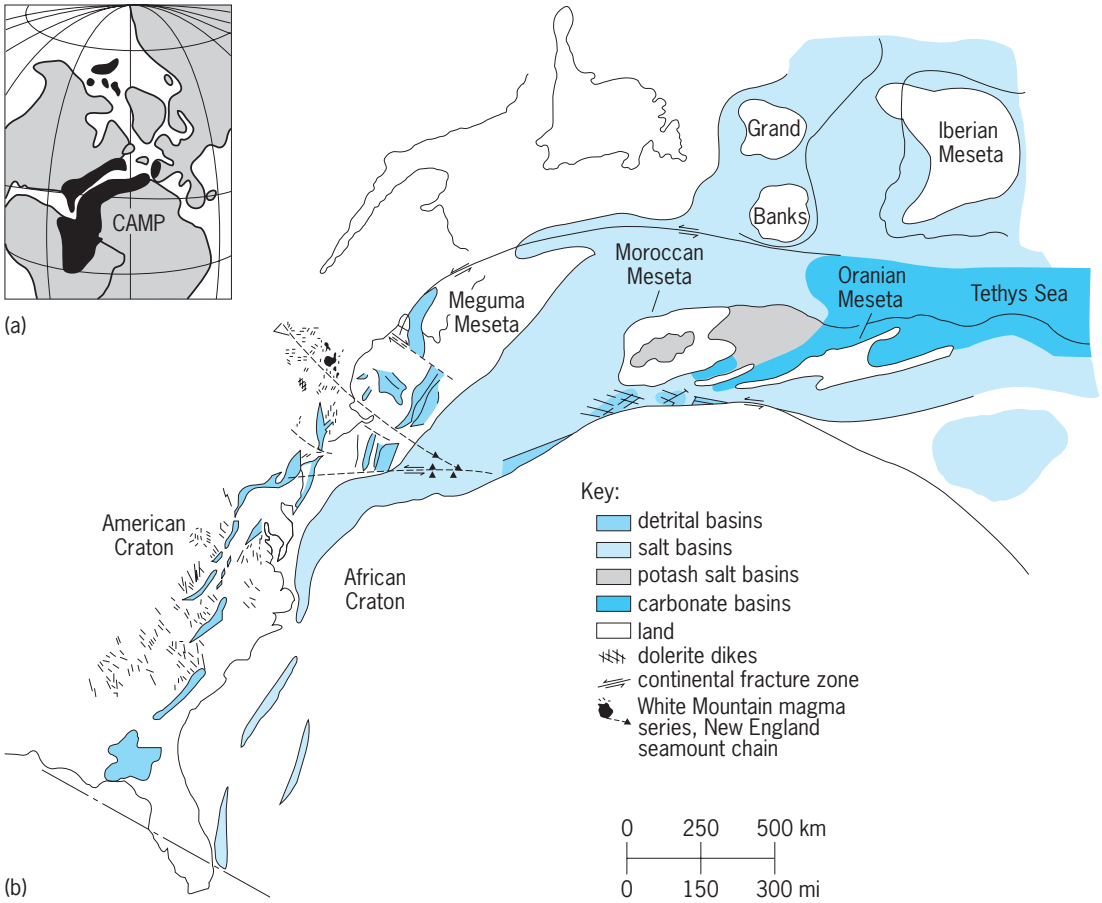

Fig. 2. (*a*) Map of CAMP. (*b*) Late Triassic–Early Jurassic reconstruction of eastern North America and northwest Africa, outlining rift basins, lithofacies, and tectonic elements, during breakup but prior to sea-floor spreading of the Central Atlantic in the Middle Jurassic. Arrows indicate the direction of intraplate movement along continental fracture zones. (*After W. Manspeizer, ed., Triassic-Jurassic Rifting: Continental Breakup and the Origin of the Atlantic Ocean and Passive Margins, Elsevier, 1988*)

Jurassic lake deposits show a pervasive cyclical pattern of wetting and drying, wherein lakes expanded and contracted with periodicities of 21,000, 42,000, 100,000, and 400,000 years. These intervals agree with the Holocene Milankovitch astronomical theory of climates that are related to small variations in the Earth's orbit and rotation. *See* BASIN; JURASSIC; RIFT VALLEY; SALINE EVAPORITES.

Crustal extension. The most important tectonic event in the Mesozoic Era was the rifting of the Pangaea craton, which began in the Late Triassic, culminating in the Middle Jurassic with the formation of the Central Atlantic ocean basin and the proto-Atlantic continental margins (Fig. 2*b*). Rifting began in the Tethys region in the Early Triassic, and progressed from western Europe and the Mediterranean into the Central Atlantic off Morocco and eastern North America by the Late Triassic. As crustal extension continued throughout the Triassic, the Tethys seaway spread farther westward and inland. Although marine palynomorphs from deep wells on Georges Bank indicate that epicontinental seas, from Tethys on the east or Arctic Canada (through eastern Greenland) on the north, transgressed the craton to the coast of Massachusetts in the Late Triassic, an ocean sea floor did not form in this region until the Middle Jurassic. By that time, rifting and sea-floor spreading extended into the Gulf of Mexico, separating North and South America. Africa and South America did not separate until the Early Cretaceous, when sea- floor spreading created the South Atlantic ocean basin, the great flood basalts of the Amazon and Karoo (Africa), and those of Transarctic and Tasmania record that Gondwanaland had begun to break up by the Triassic Period. *See* BASALT; CRETACEOUS; PALYNOLOGY.

Atlantic rift basins. Continental rift basins, passive continental margins, and ocean basins form in response to divergent stresses that extend the crust. Crustal extension, as it pertains to the Atlantic, embraces a major tectonic cycle marked by Late Triassic–Early Jurassic rifting and Middle Jurassic to Recent (Holocene) drifting. The rift stage, involving heating and stretching of the crust, was accompanied by uplift, faulting, basaltic igneous activity, and rapid filling of deep elongate rift basins. The drift stage, involving the slow cooling of the lithosphere over a broad region, was accompanied by thermal subsidence with concomitant marine transgression of the newly formed plate margin. The transition from rifting to drifting, accompanied by sea-floor spreading, is recorded by the postrift

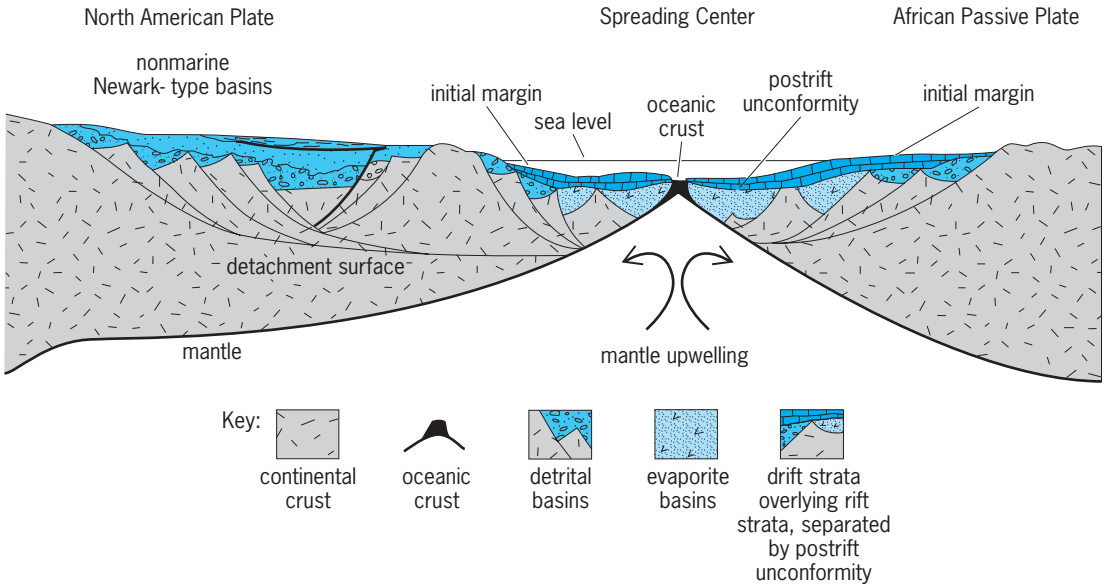

Fig. 3. Diagrammatic cross section of the Atlantic-type continental passive margine of North America and North Africa, taken at the beginning of the Middle Jurassic with the onset of sea-floor spreading that resulted from crustal thinning and mantie upwelling Note the setting of the Late Triassio-Early Jurassic continental and marine rift basins and their relation to the future passive margins, the postrift unconformity, and the overtying Middle Jurassic drift sequence. (*After W. Manapelzer, ed., Triassic-Jurassic Rifting: Continental Breakup and the Origin of the Atlantic Ocean and Passive Margins, pt. A, Elsevier, 1988*)

unconformity (**Fig. 3**). Late Triassic Proto-Atlantic rift basins occur in eastern North America, Greenland, the British Isles, north and central West Africa, and South America. *See* CONTINENTAL DRIFT; CONTINENTAL MARGIN; HOLOCENE; UNCONFORMITY.

Within the proto-Atlantic, off eastern North America and Morocco, lie about 50 northeast- to southwest-trending elongate rift basins, called the Newark rift basins, whose trend follows the fabric of the Alleghenian-Variscan orogen (Fig. 2). Some of these basins are exposed on the land, while others occur beneath the Coastal Plain and under the continental shelf. Almost all of them have developed along reactivated late Paleozoic thrust faults (Fig. 3). Seismic reflection surveys of both the onshore and offshore rift basins show that they are asymmetric half-grabens, bounded on one side by a system of major high-angle normal faults, and on the other side by a gently sloping basement with sedimentary overlap. These basins contain Late Triassic to Early Jurassic strata, which comprise the Newark Supergroup. At the end of the Triassic and into the Early Jurassic, the Newark strata of the Atlantic region were uplift, tilted, faulted, and intruded by tholeiitic sills and dikes. Subsequently, they were eroded and unconformably overlain by younger Jurassic post rift or drift strata. This episode of deformation, known as the Palisade disturbance, is most evident by the postrift unconformity in the offshore basins. *See* FAULT AND FAULT STRUCTURES; GRABEN.

Figure 2, showing a predrift paleogeographic reconstruction of the circum-Atlantic region, outlines the major Triassic basins and lithofacies. Two major basin types are recognized (Fig. 3): Newark-type detrital basins, which are exposed onshore as half-grabens and contain a thick (approximately 2.5–5 mi or 4–8 km) sequence of fluvial-lacustrine strata and border fanglomerates; and evaporite basins, which occur seaward of the string of detrital basins and contain a thick evaporite facies with interbeds of red mudstones and carbonates. As more than 3300 ft (1000 m) of salt was concentrated, these basins must have acted as huge evaporating pans. The Triassic-Jurassic systemic boundary, throughout the broad region of the Atlantic, typically is marked by tholeiitic lava flows and intrusives that are dated about 200 Ma (Early Jurassic), or only slightly older than the oldest dated crust of the Atlantic Ocean. *See* FACIES (GEOLOGY).

Central Atlantic magmatic province. The breakup of Pangaea was accompanied by the most extensive outpouring of continental basaltic lava known, covering an area estimated to be about 4 million mi^2 (10 million km^2). Basaltic remnants (flood basalts and feeder dikes) of this igneous province, named the Central Atlantic Magmatic Province (CAMP), are found on the rifted margins of four circum-Atlantic continents, particularly eastern North America, South America, western Africa, and southwestern Europe (Fig. 2*a*). Almost all of CAMP rocks are mafic tholeoiites that were intruded into or extruded onto clastic rocks in Newark-type rift basins. The Palisades Sill, along the west shore of the lower Hudson River in Northern New Jersey, is an example of this magmatic event. It is thought that the immensely thick and widespread seaward-dipping basaltic wedges, manifested by the East Coast magnetic anomaly, are linked to CAMP. *See* GEOMAGNETISM.

Recent multidisciplinary studies in stratigraphy, palynology, geochronology, paleomagnetism, and petrography indicate CAMP formed as a singular episodic event in Earth's history, occurring during

a very brief interval of geologic time, perhaps no longer than 4 million years. Importantly, this event occurred about 200 Ma and was contemporaneous with widespread mass extinction at the Triassic-Jurassic boundary. A causal relationship is postulated by many scientists to climate change that was forced by emission of huge quantities of volcanic gasses, estimated by researchers to be in the order of from $1–5 \times 10^{12}$ metric tons. Radical shifts in climate due to the ejection of aerosols into the atmosphere and destruction of environments by lava flows, ash falls, fires, and toxic pollution of soil and streams are suggested consequences of this event. A similar explanation has been offered for the mass extinctions that occurred at the end of the Paleozoic and Mesozoic Eras, with major volcanic eruptions of the Siberian Traps and Deccan Traps, respectively.

Western North America. Permian-to-Triassic consolidation of Pangaea in western North America led to the Sonoma orogeny (mountain building), which resulted from overthrusting and suturing of successive island-arc and microcontinent terranes to the western edge of the North American Plate. However, toward the end of the Triassic Period, as crustal extension was occurring in the Central Atlantic region, the plate moved westward, overriding the Pacific Plate along a reversed subduction zone. This created, for the remainder of the Mesozoic Era, an Andean-type plate edge with a subducting sea floor and associated deep-sea trench and magmatic arc. These effects can be studied in the Cordilleran mountain belt, from Alaska to California, where great thicknesses of volcanics and graywackes were derived from island arcs to the west, and in Idaho and eastern Nevada, where thick Lower Triassic marine limestones and sandstones were laid down adjacent to the rising Cordilleras on the west and interfinger with continental redbeds derived from the stable interior to the east. *See* CORDILLERAN BELT; LIMESTONE; OROGENY; REDBEDS; SANDSTONE.

As the epicontinental seas regressed westward, nonmarine fluvial, lacustrine, and windblown sands were deposited on the craton. Today many of these red, purple, ash-gray, and chocolate-colored beds are some of the most spectacular and colorful scenery in the American West. For example, the Painted Desert of Arizona, known for its petrified logs of conifer trees, was developed in the Chinle Formation, and the windblown sands of the Wingate and Navajo formations are exposed in the walls of Zion National Park in southern Utah. *See* PETRIFIED FORESTS.

Life. The Triassic is bracketed by two major biotic crises that terminated many groups of organisms. Triassic marine faunas can be distinguished from their predecessors by the absence of groups that flourished in the Permian, such as the fusulinid foraminiferans, the tabulate and rugose corals, the trepostome and cryptostome bryozoans, the productid and other brachiopod groups, the trilobites, and certain groups of echinoderms. Owing to a very low stand of sea level, Early Triassic marine faunas are not common, and show very small diversity except for ammonites. This is partly ecologic. The reef community, for example, is not known from the Early Triassic deposits; yet when it reappeared in mid-Triassic time, it contained sponges that were major members of Permian reefs, and that must have survived in settings that have not been found. *See* BRACHIOPODA; BRYOZOA; ECHINODERMATA; FORAMINIFERIDA; FUSULINACEA; PERMIAN; RUGOSA; TABULATA; TRILOBITA.

Triassic faunas are also distinguished from earlier ones by newly evolved groups of plants and animals. In marine communities, molluscan stocks proliferated vigorously. Bivalves diversified greatly and took over most of the niches previously occupied by brachiopods; ammonites proliferated rapidly from a few Permian survivors. The scleractinian (modern) corals appeared, as did the shell-crushing placodont reptiles and the ichthyosaurs. In continental faunas, various groups of reptiles appeared, including crocodiles and crocodilelike forms, the mammallike reptiles, and the first true mammals, as well as dinosaurs. *See* CEPHALOPODA; CROCODYLIA; DINOSAUR; MAMMALIA; MOLLUSCA; PLACODONTIA; SCLERACTINIA.

The Jurassic faunas lack numerous stocks lost in the Rhaeto-Liassic faunal crises. These include survivors of the Permian crises, such as the orthoceratid cephalopods and the conodonts. However, stocks that had flourished greatly in the Triassic also became extinct (phytosaurs, placodonts) or nearly extinct: ammonites were reduced to one or two surviving lineages, which then underwent no other great evolutionary surge in Jurassic time. Furthermore, new groups such as the plesiosaurs and pterosaurs appeared. Triassic land plants contain survivors of many Paleozoic stocks, but the gymnosperms became dominant and cycads appeared. The Permo-Triassic and Rhaeto-Liassic crises record a severe stressing of the biosphere, but the nature and origin of these stresses have not been established. *See* CONODONT; CYCADEOIDALES; EXTINCTION (BIOLOGY); INDEX FOSSIL; PALEOBOTANY; PALEOECOLOGY; PALEONTOLOGY; PINOPHYTA; PTEROSAURIA.

Warren Manspeizer

Bibliography. A. Hallam, The end-Triassic bivalve extinction event, *Paleogeo. Paleoclimatol. Paleoecol.*, vol. 35, pp. 1–44, 1981; W. E. Hames et al. (eds.), *The Central Atlantic Magmatic Province: Insights from Fragments of Pangea*, American Geophysical Union, Washington, D.C., 2003; G. D. Klein (ed.), *Pangea: Paleoclimate, Tectonics, and Sedimentation During Accretion, Zenith, and Breakup of a Supercontinent* Spec. Pap. No. 288, Geological Society of America, Boulder, 1994; P. M. Letournea and P. E. Olsen (eds.), *The Great Rift Valleys of Pangea in Eastern North America*, vol. 1: *Tectonics, Structure and Volcanism*, Columbia University Press, New York, 2003; W. Manspeizer (ed.), *Triassic-Jurassic Rifting: Continental Breakup and the Origin of the Atlantic Ocean and Passive Margins*, pt. A, Elsevier Science, Amsterdam, 1988; D. R. Prothero and R. H. Dott, *Evolution of the Earth*, 7th ed., McGraw-Hill, New York, 2004; S. M. Stanley, *Earth System History*, Freeman, New York, 1999.

Tribology

The science and technology of interactive surfaces in relative motion. It incorporates various scientific and technological disciplines such as surface chemistry, fluid mechanics, materials, lubricants, contact mechanics, bearings, and lubrication systems. It is customarily divided into three branches: friction, lubrication, and wear.

Friction. This phenomenon is encountered whenever there is relative motion between contacting surfaces, and it always opposes the motion. As no mechanically prepared surfaces are perfectly smooth, when the surfaces are first brought into contact under light load, they touch only along the asperities (real area of contact). The early theories attributed friction to the interlocking of asperities; however, it is now understood that the phenomenon is far more complicated. Although the theories of friction are new, its laws have changed little since Leonardo da Vinci's time: (1) The force of friction is proportional to the normal load pressing the surfaces together. (2) The force of friction is independent of the apparent area of contact. The coefficient of friction (the ratio of friction force to normal force) ranges, under normal, dry conditions, from small (for example, 0.04 for Teflon on steel) to large (for example, 1.1 for cast iron on cast iron). *See* FRICTION; IRON; STEEL.

Lubrication. When clean surfaces are brought into contact, their coefficient of friction decreases drastically if even a single molecular layer of a foreign substance (for example, an oxide) is introduced between the surfaces. For thicker lubricant films, the coefficient of friction can be quite small and no longer dependent on the properties of the surfaces but only on the bulk properties of the lubricant. Most common lubricants are liquids and gases, but solids such as molybdenum disulfide or graphite may be used. The machine elements that accommodate relative motion through the introduction of a lubricant between the contacting surfaces are named bearings. Bearings lubricated by liquids or gases operate either in the hydrostatic (externally pressurized) mode or in the hydrodynamic (self-acting) mode. Conformal hydrostatic or hydrodynamic bearings (for example, thrust or journal bearings) remain rigid, as the pressures are relatively small. In self-acting counterformal bearings (ball or roller bearings), the pressures are large (1–4 GPa), necessitating that both the elastic deformation of the surfaces and the pressure dependence of viscosity be taken into account during performance calculations; these bearings are said to operate in the elastohydrodynamic mode. *See* ANTIFRICTION BEARING; LUBRICATION; VISCOSITY.

Wear. This is the progressive loss of substance of one body because of rubbing by another body. There are many different types of wear, including sliding wear, abrasive wear, corrosion, and surface fatigue. The scientific study of wear is of recent origin and is made possible by the development of sensitive experimental equipment and techniques. The same type of complicated surface interactions that cause friction also cause wear; thus quantitative prediction of wear rates involves the same difficulties as that of friction. Actually, the situation is more severe: whereas the friction coefficient changes at most by one order of magnitude from one metal pair to another, the wear coefficient changes by several orders of magnitude depending on environmental and other conditions. *See* ABRASIVE; CORROSION; WEAR.

Andras Z. Szeri

Bibliography. B. Bhushan (ed.), *Fundamentals of Tribology and Bridging the Gap between the Macro and Micro/Nanoscales*, 2001; B. Bhushan (ed.), *Modern Tribology Handbook*, 2 vols., 2000; B. Bhushan, *Principles and Applications of Tribology*, 1999; K. C. Ludema, *Friction, Wear, Lubrication: A Textbook in Tribology*, 1996; E. Rabinowicz, *Friction and Wear of Materials*, 1995; A. Z. Szeri, *Fluid Film Lubrication: Theory and Design*, 1998; A. Z. Szeri, *Tribology: Friction, Lubrication and Wear*, 1980.

Trichasteropsida

A monospecific order of Asteroidea established for *Trichasteropsis wiessmanni*, the only Triassic asteroid known from articulated specimens. It is a small starfish with a relatively large disc and short arms. The skeleton is differentiated into marginals, actinals, and abactinals, the marginals comprising a single series of blocklike ossicles. The abactinal surface is composed of larger stellate plates and smaller rods and granules; a carinal row is present. Oral plates are large and well developed. Pedicellariae are not present. *Trichasteropsis* is the most primitive post-Paleozoic asteroid known and lies close to the latest common ancestor of all living asteroids. It comes from the Muschelkalk (Middle Triassic, Anisian-Ladinian) of Germany. *See* ASTEROIDEA; ECHINODERMATA.

Andrew B. Smith

Bibliography. D. B. Blake, *J. Nat. Hist.*, 21:481–528, 1987.

Trichomycetes

A polyphyletic class of Eumycota in the subdivision Zygomycotina, containing the orders Amoebidiales Asellariales, Eccrinales, and Harpellales. These orders are grouped together because they usually exist only as commensals in the mid- and hindgut of arthropods (*Amoebidium parasiticum* can be found on the outside of its hosts).

Asexual reproduction is accomplished by the formation of trichospores (Harpellales), arthrospores (Asellariales), sporgiospores (Eccrinales), or ameboid cells, cystospores, or rigid-walled spores (Amoebidiales). Sexual reproduction (zygospore formation) is known only in Harpellales, although conjugations have been observed in Asellariales.

The thallus of Amoebidiales is aseptate, but is regularly septate in the other three orders; septa with plugs in the lenticular cavities are produced by

Asellariales and Harpellales. Similar septa and plugs are formed by Dimargaritales and Kickxellales (Zygomycetes). *See* ZYGOMYCETES.

Classification is based on the type of reproduction; thallus branching pattern, complexity, and septation; and nature of the holdfast. Although most mycologists include Amoebidiales in the class Trichomycetes because of morphology and habitat, possibily the formation of ameboid spores renders the order unrelated to the other members. Eccrinales are treated either as the most advanced Trichomycetes or as nonrelated organisms. Because of similarities in septum and septal plug structure, and in general architecture of the asexual reproductive structures, Asellariales and Harpellales (Trichomycetes) and Dimargaritales and Kickxellales (Zygomycetes) may be more closely related to one another than to other orders in the two classes.

Trichomycetes occur worldwide and may be found anywhere a suitable host exists; they inhabit a more or less equatic environment (the gut); and the spores are discharged with feces. A few species have been cultured [*A. parasiticum* (Amoebidiales), some Harpellales], but most species are found in the host. Infection of a host is probably by the chance ingestion of a spore that lands on a food source.

More than 30 species of the Harpellales and one species of Amoebidiales (*A. parasiticum*) have been cultured. Most species are known only from their existence in the host. *See* EUMYCOTA; FUNGI; ZYGOMYCOTA. Gerald L. Benny

Bibliography. L. R. Batra, *Insect-Fungus Symbiosis*, 1979; D. H. Howard and J. D. Miller (eds.), *The Mycota*, vol. 6, 1996; R. W. Lichtwardt, *The Trichomycetes*, 1986; D. J. McLaughlin, E.G. McLaughlin, and P. A. Lemke (eds.), *The Mycota*, vol. 7, 1999; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Trichoptera

An aquatic order of the class Insecta commonly known as the caddis flies. The adults have two pairs of well-veined hairy wings, long antennae, and mouthparts capable of lapping only liquids (**illus.** *a*). The larvae are wormlike, with distinct heads, three pairs of legs on the thorax, and a pair of hook-bearing legs at the end of the body (illus. *b*). The pupae are delicate, with free appendages held close to the body, and have a pair of sharp mandibles, or jaws, which are used to cut and exit from the cocoon.

The adults live several weeks to several months, the females rapidly becoming mature. The crawl into water and lay eggs under stones and other objects. The eggs hatch into aquatic larvae. Some larvae construct a fixed retreat and some type of nest in cracks or crevices. Others build a portable case in which to live, and a few build neither, crawling about in moss and under stones. Most caddis-fly larvae are omnivorous, feeding on algae, other microorganisms, or other aquatic animals which are small enough for them to devour. A few forms are entirely predacious and live on other aquatic insect larvae. When full grown, the free-living and retreat-making larvae spin an oval cocoon under a rock or in a crevice and pupate in it. The case-makers anchor the case securely to some object in the water and pupate inside it. When mature, the pupa, using its jaws, cuts its way out of the cocoon (illus. *c* and *d*), swims to the surface, sometimes climbing out of the water on a stem or stone, and there the adult emerges from the pupal skin.

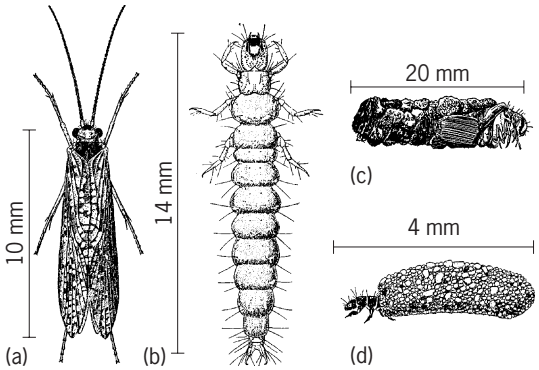

Life cycle of Trichoptera. (*a*) Adult and (*b*) the free-living larva of the widespread genus *Rhyacophila*. (*c*) Head and thorax of larva protruding from purselike case of a micro caddis fly, *Ochrotrichia*. (*d*) Head and thorax of larva protruding from case of a large caddis fly, *Limnephilus*. (*Illinois Natural History Survey*)

Except for a brackish-water species in New Zealand and a few moss-inhabiting species in Europe and North America, caddis flies occur only in fresh water. They abound in cold or running water relatively free from pollution. Altogether they compose a large and important segment of the biota of such habitats and of the fish feed economy.

The Trichoptera include about 10,000 described species, divided into 34 families, and occur in practically all parts of the world. The order probably arose over 2×10^8 years ago in early Mesozoic time, for fossils of typical trichopteran wings have been found in Late Triassic deposits. Many existing genera are probably of Cretaceous origin; hence, it seems almost certain that representatives of all the diverse family lines evolved during the middle part of the Mesozoic Era at the same time that the dinosaurs were proliferating. The Trichoptera were originally cool-adapted animals, as most of the primitive forms still are, but many warm-adapted lines have evolved. As a result, the caddis flies are found in arctic, temperate, and tropical habitats. *See* INSECTA. Herbert H. Ross

Bibliography. C. Betten, *The Caddis Flies or Trichoptera of New York State*, N. Y. State Mus. Bull. 292, 1934; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; H. H. Ross, *The Caddisflies, or Trichoptera, of Illinois*, Illinois Nat. Hist. Surv. Bull. 23, 1944; H. H. Ross, *The Evolution and Classification of the Mountain Caddisflies*, 1956; G. Ulmer, Trichoptera, *Genera Insectorum*, 60:1–259, 1907; R. L. Usinger (ed.), *Aquatic Insects of California*, 1956.

Trichroism

When certain optically anisotropic transparent crystals are subjected to white light, a cube of the material is found to transmit a different color through each of the three pairs of parallel faces. Such crystals are sometimes termed trichroic, and the phenomenon is called trichroism. This expression is used only rarely today since the colors in a particular crystal can appear quite different if the cube is cut with a different orientation with respect to the crystal axes. Accordingly, the term is frequently replaced by the more general term pleochroism. Even this term is being replaced by the phrase linear dichroism or circular dichroism to correspond with linear birefringence or circular birefringence. *See* BIREFRINGENCE; DICHROISM; PLEOCHROISM.

Cordierite is a typical trichroic crystal. In light with a vibration direction parallel to the X axis of the index ellipsoid, the crystal appears yellow. With the vibration direction parallel to the Y axis, the crystal is dark violet. In the Z direction the crystal is clear.

The phenomena of trichroism can be explained crudely as follows. Classically, one can consider an electron in a biaxial crystal as having three different force constants associated with a displacement directed along each of the principal axes. Linear polarized light traveling along the X axis with its electric vector parallel to the Y axis will displace the electron against the Y force constant and will experience a certain absorption and retardation. It will be unaffected by the force constants in the X and Z directions. Similarly, polarized light traveling in the Y direction will experience absorption and retardation. Unpolarized light will also be absorbed in a different fashion depending on the direction of propagation. In this case, light traveling in the X direction can be considered as composed of an equal mixture of light polarized parallel to the Y axis and the Z axis. The absorption will be intermediate between the two polarization directions. *See* CRYSTAL OPTICS; POLARIZED LIGHT.

Bruce H. Billings

Bibliography. E. E. Wahlstrom, *Optical Crystallography*, 5th ed., 1979; E. A. Wood, *Crystals and Light: An Introduction to Optical Crystallography*, 1977; A. Yariv and P. Yeh, *Optical Waves in Crystals*, 1983.

Tricladida

An order of the Turbellaria (of the phylum Platyhelminthes) known commonly as planaria, which are several millimeters to 20 in. (50 cm) or more in length. They have a diverticulated intestine with a single anterior branch and two posterior branches separated by a plicate pharynx or pharynges. Rhabdites are numerous and, except in cave planarians, two to many eyes are present. The much branched protonephridial tubules form a network with numerous nephridiopores on each side of the body. The female reproductive system includes a single pair of small anteriorly located ovaries, numerous minute yolk glands arranged in clusters along either side of the body, the common ducts, the female antrum, and usually one or more bursae. The male system has several to many testes, which are lateral in position and connected with a single sperm duct on each side. These ducts empty either directly or after fusion into the copulatory organ which lies in the male antrum. Following copulation and mutual insemination, capsules containing several fertilized eggs are attached to objects in the water and hatch in 2 or more weeks into young worms. Asexual reproduction by fission is common in forms such as the cosmopolitan *Dugesia tigrina* which has been much used in studies on regeneration. Fragmentation and regeneration is the usual method of reproduction in some land planarians such as *Bipalium kewense*, an exotic species which has become established through much of the southern United States. The marine planarian *Bdelloura candida* is a commensal on the horseshore crab. *See* TURBELLARIA.

E. Ruffin Jones

Bibliography. L. von Graff, Tricladida, in H. G. Bronn (ed.), *Klassen und Ordnungen des Tierrechs*, vol. 4, pt. 2, 1912–1917; R. C. Harrel, D. L. Bechler, and R. E. Ogren, First Texas record of *Geoplana arkalabamensis* (Turbellaria: Tricladida: Terricola) with a review of other land planarians reported in the state, *Tex. J. Sci.*, 46:45–49, 1994; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; L. Winsor, The biodiversity of terrestrial flatworms (Tricladida: Terricola) in Queensland: A preliminary report, *Mem. Mus. Victoria*, 56:575–579, 1997.

Trigger circuit

An electronic circuit that generates or modifies an existing waveform to produce a pulse of short time duration with a fast-rising leading edge. This waveform, or trigger, is normally used to initiate a change of state of some relaxation device, such as a multivibrator. The most important characteristic of the waveform generated by a trigger circuit is usually the fast leading edge. The exact shape of the falling portion of the waveform often is of secondary importance, although it is important that the total duration time is not too great. A pulse generator such as a blocking oscillator may also be used and identified as a trigger circuit if it generates sufficiently short pulses. *See* PULSE GENERATOR.

Peaking (differentiating) circuits. These circuits, which accent the higher-frequency components of a pulse waveform, cause sharp leading and trailing edges and are therefore used as trigger circuits. The simplest form of peaking circuits are the simple RC (resistance-capacitance) and RL (resistance-inductance) networks shown in **Fig. 1**. If a steep wavefront of amplitude V is applied to either of these circuits, the output will be a sudden rise followed by an exponential decay according to the equation $v_0 = V\epsilon^{-kt}$, where $k = 1/RC$ or R/L.

These circuits are often called differentiating

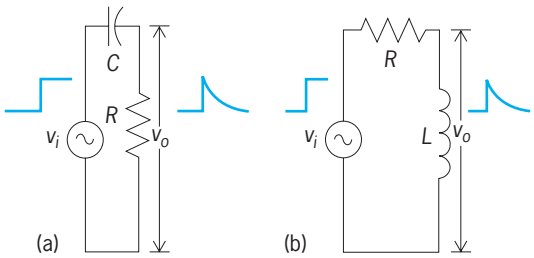

Fig. 1. Simple peaking circuits: (a) resistance-capacitance network; (b) resistance-inductance network.

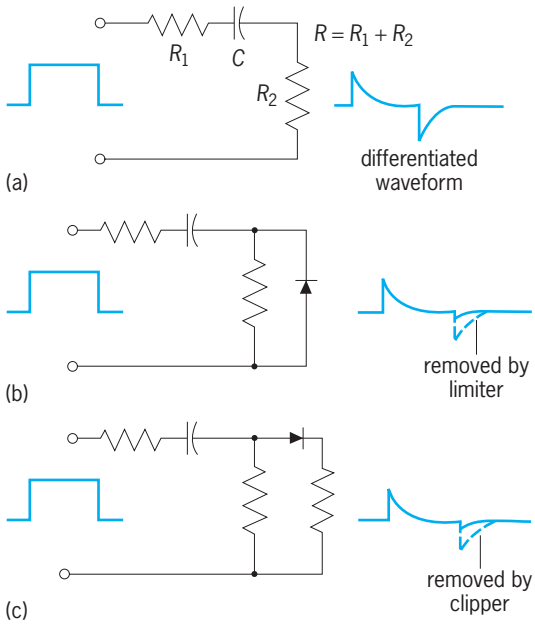

Fig. 2. Differentiated pulses. (*a*) Basic circuit. (*b*) Limiting unwanted portion. (*c*) Clipping unwanted portion.

circuits because the outputs are rough approximations of the derivative of the input waveforms, if the *RC* or *R/L* time constant is sufficiently small.

If the pulse is applied to the differentiating circuits, the resultant waveform shown in **Fig. 2** may be used as a trigger. It is sometimes necessary, however, to remove by limiting or clipping the undesired portion of the waveform to prevent circuits from responding to it.

The *RL* circuit of Fig. 1 cannot be considered in its simplest form when extremely fast rise times are required because of the distributed capacitance and small series resistance associated with the inductance. A more accurate representation of the circuit is that in **Fig. 3**. The response is limited as shown for a fixed value of L and C. The value for $k = 1$ is referred to as critical damping. A value of k slightly less than unity provides a pulse that is a suitable trigger for many applications.

Ringing circuits. A circuit of the form shown in Fig. 3 that is highly underdamped, or oscillatory ($k \ll 1$), and is supplied with a step or pulse input is often referred to as a ringing circuit. When used in the output of a field-effect or bipolar transistor as in **Fig. 4**, this circuit can be used as a trigger circuit. When the input pulse is applied, current in the output circuit is immediately cut off. Since the current in L cannot change instantaneously, it flows in the LC circuit in an oscillatory manner, gradually decaying because of the resistance in the circuit. However, if the diode is in the circuit, the circuit will be highly overdamped for the negative portion of the oscillatory waveform, and the oscillations will be damped out as shown.

If a transistor is used as the current source and operated near saturation, damping will take place when the transistor goes into saturation as shown in **Fig. 5**.

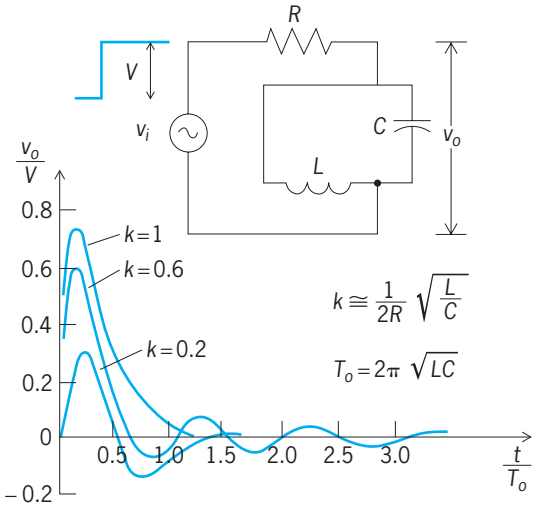

Fig. 3. Resistance-inductance-capacitance peaking circuit.

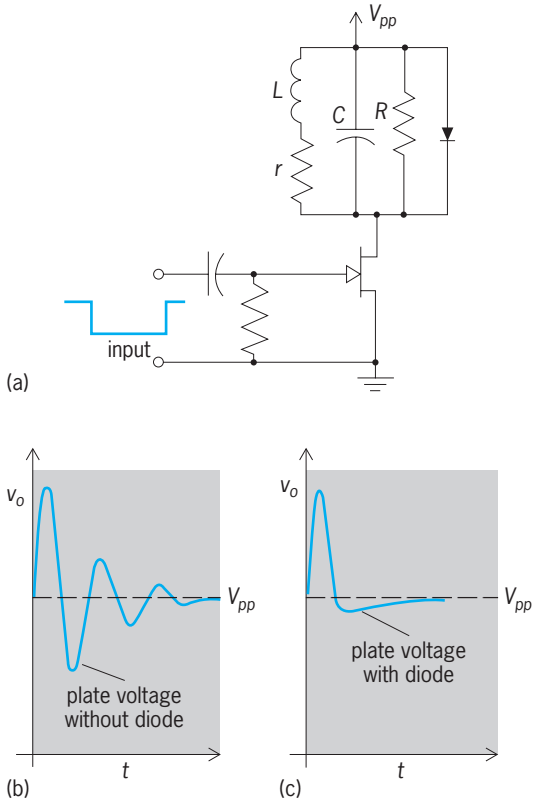

Fig. 4. Ringing circuit as trigger source. (*a*) Circuit diagram. (*b*) Drain-voltage waveform without diode limiter. (*c*) Drain-voltage waveform with diode limiter.

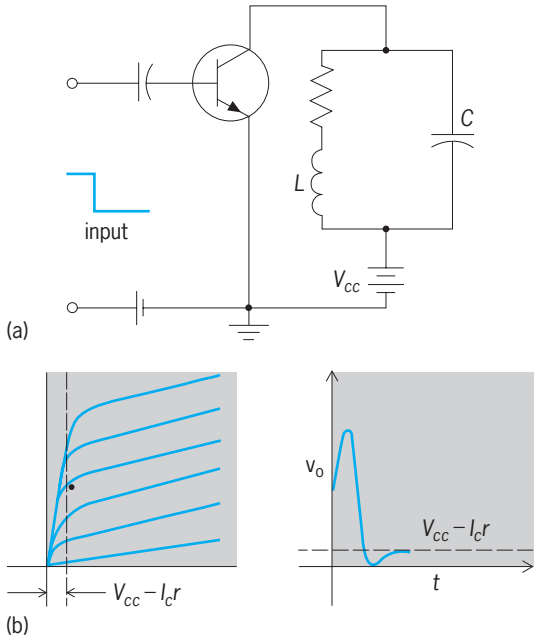

Fig. 5. Ringing circuit with transistor saturation damping: (*a*) circuit diagram; (*b*) waveforms.

The diode is not required. For other waveforms *See* WAVE-SHAPING CIRCUITS.

Glenn M. Glasford

Bibliography. D. Christiansen, *Electronics Engineers' Handbook*, 4th ed., 1996; J. Millman and A. Grabel, *Microelectronics*, 2d ed., 1987; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Triglyceride (triacylglycerol)

A simple lipid. Triglycerides are fatty acid triesters of the trihydroxy alcohol glycerol which are present in plant and animal tissues, particularly in the food storage depots, either as simple esters in which all the fatty acids are the same or as mixed esters in which the fatty acids are different. The triglycerides constitute the main component of natural fats and oils.

The generic formula of a triglyceride is shown below, where RCO_2H, $R'CO_2H$, and $R''CO_2H$ repre-

$$\begin{array}{l} CH_2-OOC-R \\ | \\ CH-OOC-R' \\ | \\ CH_2-OOC-R'' \end{array}$$

sent either the same or different fatty acids, such as butyric or caproic (short chain), palmitic or stearic (long chain), oleic, linoleic, or linolenic (unsaturated). Saponification with alkali releases glycerol and the alkali metal salts of the fatty acids (soaps). The triglycerides in the food storage depots represent a concentrated energy source, since oxidation provides more energy than an equivalent weight of protein or carbohydrate.

Animal and vegetable triglycerides contain predominantly even-chain-length fatty acids, with palmitic and oleic acids as the main components. Since n fatty acids may be esterified in $(n^3 + n^2)/2$ ways into glycerol, and since natural fats contain a variety of fatty acids, the number of component triglycerides of a relatively simple natural fat or oil may be high. Some pure simple and mixed triglycerides have been isolated from natural fats by fractional crystallizations at low temperatures, but in general physical methods are not yet available for the separation of naturally occurring mixtures. Several theories, such as those of even distribution and partial random distribution, have been advanced to account for the distribution of the fatty acids in the triglycerides. Many synthetic triglycerides have been prepared, and the study of the physical properties of these compounds has provided much useful information. Melting-point, x-ray-diffraction, and infrared-spectroscopy investigations have shown that triglycerides may exist in at least three polymorphic modifications. *See* MOLECULAR STRUCTURE AND SPECTRA.

The physical and chemical properties of fats and oils depend on the nature of the fatty acids present. Saturated fatty acids give higher-melting fats and represent the main constituents of solid fats, for example, lard and butter. Unsaturation lowers the melting point of fatty acids and fats. Thus, in the oil of plants, unsaturated fatty acids are present in large amounts, for example, oleic acid in olive oil and linoleic and linolenic acids in linseed soil. Oils are hydrogenated commercially to produce the proper consistency and melting point for use as edible fats. *See* CARBOXYLIC ACID; FAT AND OIL (FOOD); LIPID.

Roy H. Gigg; Herbert E. Carter

Bibliography. G. Fuller and W. D. Nes (eds.), *Ecology and Metabolism of Plant Lipids*, 1986; F. J. Mead et al., *Lipids: Chemistry, Biochemistry, and Nutrition*, 1986; P. Quinn and J. Harwood (eds.), *Plant Lipid Biochemistry, Structure, and Utilization*, 1991.

Trigonometry

The study of triangles and the trigonometric functions. One common use for trigonometry is to measure heights and distances that are awkward or impossible to measure by ordinary means. Surveyors use it to find heights of mountains and distances

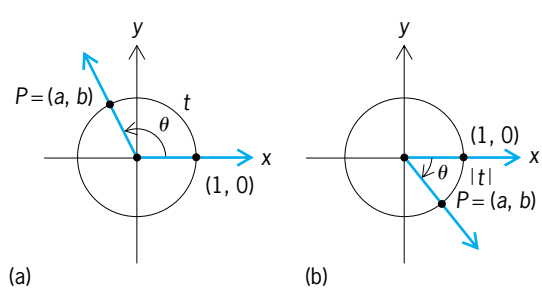

Fig. 1. Point *P* on the unit circle corresponding to $\theta = t$ radians. (*a*) $t \geq 0$: length of arc from (1,0) to *P* is *t* units. (*b*) $t < 0$: length of arc from (1,0) to *P* is $|t|$ units.

TABLE 1. Values of trigonometric functions at integral multiples of $\pi/4$ (90°)

θ, radians	θ	$\sin\theta$	$\cos\theta$	$\tan\theta$	$\csc\theta$	$\sec\theta$	$\cot\theta$
0	0°	0	1	0	Not defined	1	Not defined
$\pi/2$	90°	1	0	Not defined	1	Not defined	0
π	180°	0	−1	0	Not defined	−1	Not defined
$3\pi/2$	270°	−1	0	Not defined	−1	Not defined	0

across lakes and countries; engineers use it in the design of large structures and roads; astronomers use it in accurate measurements of the time and in locating the position of objects in the sky; and navigators on the sea and in the air use it to find latitudes, longitudes, and direction. Trigonometry has evolved from use by surveyors, engineers, and navigators to applications involving ocean tides, the rise and fall of food supplies in certain ecologies, brainwave patterns, the analysis of alternating-current electricity, and many other phenomena of a vibratory character.

Plane trigonometry. Plane trigonometry mostly deals with the relationships among the three sides and three angles of a triangle that lies in a plane.

A ray is that portion of a line that starts at a point on the line and extends indefinitely in one direction. The starting point of a ray is called its vertex. If two rays are drawn with a common vertex, they form an angle. One of the rays of an angle is called the initial side, and the other ray is the terminal side. The angle that is formed is identified by showing the direction and amount of rotation from the initial side to the terminal side. If the rotation is in the counterclockwise direction, the angle is positive; if the rotation is clockwise, the angle is negative.

The angle formed by rotating the initial side exactly once in the counterclockwise direction until it coincides with itself (1 revolution) is said to measure 360 degrees, written 360°. Thus, one degree, 1°, is 1/360 of a revolution. One-sixtieth of a degree is called a minute, written 1′.

By using a circle of radius r, an angle can be constructed whose vertex is at the center of this circle and whose rays subtend an arc on the circle whose length equals r. Such an angle measures 1 radian. For a circle of radius r, a central angle of θ radians subtends an arc whose length s is given by Eq. (1).

$$s = r\theta \tag{1}$$

Because a central angle of 1 revolution (360°) subtends an arc equal to the circumference of the circle ($2\pi r$), it follows that an angle of 1 revolution equals 2π radians; that is, 2π radians = 360°. *See* PLANE GEOMETRY; RADIAN MEASURE.

Trigonometric functions. A unit circle is a circle whose radius is one and whose center is at the origin of a rectangular system of coordinates. For the unit circle, Eq. (1) states that a central angle of θ radians subtends an arc whose length $s = \theta$. If t is any real number, let θ be the angle equal to t radians and P be the point on the unit circle that is also on the terminal side of θ. If $t \geq 0$, then the point P is reached by moving counterclockwise along the unit circle, starting at the point with coordinates (1, 0), for a length of arc equal to t units (**Fig. 1***a*). If $t < 0$, this point P is reached by moving clockwise along the unit circle beginning at (1,0), for a length of arc equal to $|t|$ units (Fig. 1*b*). Thus, to each real number t there corresponds a unique point $P = (a, b)$ on the unit circle. The coordinates of this point P are used to define the six trigonometric functions: If $\theta = t$ radians, the sine, cosine, tangent, cosecant, secant, and cotangent of θ, respectively abbreviated as $\sin\theta$, $\cos\theta$, $\tan\theta$, $\csc\theta$, $\sec\theta$, $\cot\theta$, are given by Eqs. (2), (3), and (4).

$$\sin\theta = b \qquad \cos\theta = a \tag{2}$$

$$\text{if } a \neq 0, \quad \tan\theta = b/a, \quad \sec\theta = 1/a \tag{3}$$

$$\text{if } b \neq 0, \quad \tan\theta = a/b, \quad \sec\theta = 1/b \tag{4}$$

See COORDINATE SYSTEMS.

TABLE 2. Values of trigonometric functions at integral multiples of $\pi/6$ (30°), $\pi/4$ (45°), and $\pi/3$ (60°)

θ, radians	θ	$\sin\theta$	$\cos\theta$	$\tan\theta$	$\csc\theta$	$\sec\theta$	$\cot\theta$
$\pi/6$	30°	1/2	$\sqrt{3}/2$	$\sqrt{3}/3$	2	$2\sqrt{3}/3$	$\sqrt{3}$
$\pi/4$	45°	$\sqrt{2}/2$	$\sqrt{2}/2$	1	$\sqrt{2}$	$\sqrt{2}$	1
$\pi/3$	60°	$\sqrt{3}/2$	1/2	$\sqrt{3}$	$2\sqrt{3}/3$	2	$\sqrt{3}/3$
$2\pi/3$	120°	$\sqrt{3}/2$	−1/2	$-\sqrt{3}$	$2\sqrt{3}/3$	−2	$-\sqrt{3}/3$
$3\pi/4$	135°	$\sqrt{2}/2$	$-\sqrt{2}/2$	−1	$\sqrt{2}$	$-\sqrt{2}$	−1
$5\pi/6$	150°	1/2	$-\sqrt{3}/2$	$-\sqrt{3}/3$	2	$-2\sqrt{3}/3$	$-\sqrt{3}$
$7\pi/6$	210°	−1/2	$-\sqrt{3}/2$	$\sqrt{3}/3$	−2	$-2\sqrt{3}/3$	$\sqrt{3}$
$5\pi/4$	225°	$-\sqrt{2}/2$	$-\sqrt{2}/2$	1	$-\sqrt{2}$	$-\sqrt{2}$	1
$4\pi/3$	240°	$-\sqrt{3}/2$	−1/2	$\sqrt{3}$	$-2\sqrt{3}/3$	−2	$\sqrt{3}/3$
$5\pi/3$	300°	$-\sqrt{3}/2$	1/2	$-\sqrt{3}$	$-2\sqrt{3}/3$	2	$-\sqrt{3}/3$
$7\pi/4$	315°	$-\sqrt{2}/2$	$\sqrt{2}/2$	−1	$-\sqrt{2}$	$\sqrt{2}$	−1
$11\pi/6$	330°	−1/2	$\sqrt{3}/2$	$-\sqrt{3}/3$	−2	$2\sqrt{3}/3$	$-\sqrt{3}$

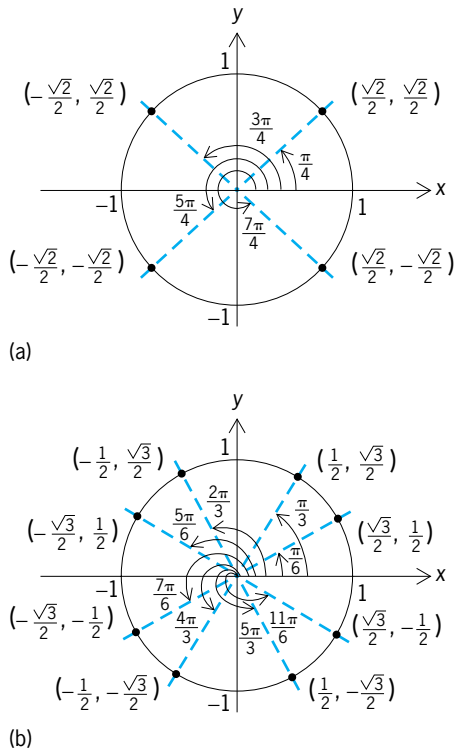

Fig. 2. Angles and coordinates of corresponding points on the unit circle. (*a*) Angles that are integral multiples of $\pi/4$ (45°). (*b*) Angles that are integral multiples of $\pi/6$ (30°) and $\pi/3$ (60°).

For example, for $\theta = 0$, the point (1,0) is on the terminal side of θ and is on the unit circle so that Eqs. (5) hold, with csc 0 and cot 0 not defined. The

$$\sin 0 = 0 \quad \cos 0 = 1 \quad \tan 0 = 0 \quad \sec 0 = 1 \tag{5}$$

trigonometric functions of angles that are integral multiples of $\pi/2$ (90°) are found similarly (**Table 1**).

The coordinates of points on the unit circle that are on the terminal sides of angles that are integral multiples of $\pi/6$ (30°), $\pi/4$ (45°), and $\pi/3$ (60°) can be found (**Fig. 2**). With Eqs. (2), (3), and (4), the trigonometric functions of these angles are obtained (**Table 2**).

It is not necessary to use a unit circle to define the trigonometric functions. This can be done by considering a point $P = (a,b)$ on the terminal side of the angle θ a distance r from the origin, so that P lies on the circle $x^2 + y^2 = r^2$ (**Fig. 3**). There is a corresponding point $P^* = (a^*,b^*)$ that is also on the terminal side of the angle θ but at unit distance from the origin, so that P^* lies on the unit circle. To P and P^* correspond the points $A = (a,0)$ and $A^* = (a^*,0)$, which lie at the feet of perpendiculars to the x axis from P and P^* respectively. Since the triangles OA^*P^* and OAP are similar, ratios of corresponding sides are equal and Eqs. (6), (7), and (8) hold.

$$\sin\theta = b^*/1 = b/r \qquad \cos\theta = a^*/1 = a/r \tag{6}$$

if a $\neq 0$,

$$\tan\theta = b^*/a^* = b/a \qquad \sec\theta = 1/a^* = r/a \tag{7}$$

if b $\neq 0$,

$$\cot\theta = a^*/b^* = a/b \qquad \csc\theta = 1/b^* = r/b \tag{8}$$

Properties of trigonometric functions. Based on Eqs. (2) and the above geometric construction (Fig. 1) for $\sin\theta$ and $\cos\theta$, θ can be any angle, so the domain of the sine and cosine functions is all real numbers. In Eqs. (3), if $a = 0$, the tangent and secant functions are not defined, so the domain of these functions is all real numbers, except odd multiples of $\pi/2$ (90°). In Eqs. (4), if $b = 0$, the cotangent and cosecant functions are not defined, so the domain of these functions is all real numbers, except multiples of π (180°). Also, since $|a| \leq 1$ and $|b| \leq 1$, the range of the sine and cosine functions is -1 to 1 inclusive. Since $|b| = |\sin\theta| \leq 1$ and $|a| = |\cos\theta| \leq 1$, it follows that $|\csc\theta| = 1/|b| \geq 1$ and $|\sec\theta| = 1/|a| \geq 1$. Thus the range of the secant and cosecant functions consists of all real numbers less than or equal to -1 or greater than or equal to 1. The range of both the tangent and cotangent functions consists of all real numbers.

Equations (2), (3), and (4) also reveal the reciprocal identities, given in Eqs. (9). Two other useful identities, given in Eqs. (10), also follow.

$$\csc\theta = 1/\sin\theta \quad \sec\theta = 1/\cos\theta \quad \cot\theta = 1/\tan\theta \tag{9}$$

$$\tan\theta = \sin\theta/\cos\theta \quad \cot\theta = \cos\theta/\sin\theta \tag{10}$$

Since (a,b) is on the unit circle, $a^2 + b^2 = 1$, and so $(\sin\theta)^2 + (\cos\theta)^2 = 1$. This is called a pythagorean

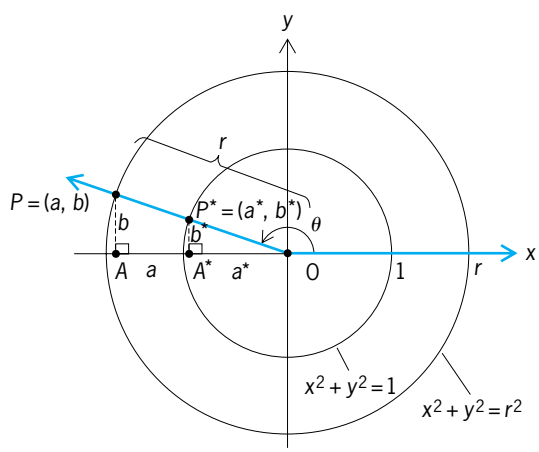

Fig. 3. Point *P* on a circle of radius *r* corresponding to angle θ, and corresponding point P^* on the unit circle.

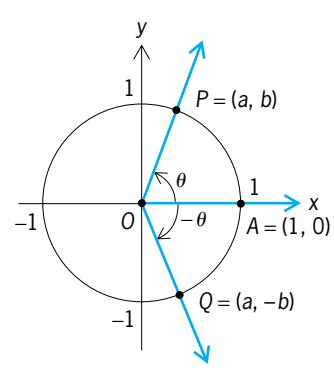

Fig. 4. Relationship between angles θ and $-\theta$, and coordinates of corresponding points on the unit circle.

TABLE 3. Signs of the trigonometric functions

Quadrant of point P	$\sin\theta$, $\csc\theta$	$\cos\theta$, $\sec\theta$	$\tan\theta$, $\cot\theta$
I	Positive	Positive	Positive
II	Positive	Negative	Negative
III	Negative	Negative	Positive
IV	Negative	Positive	Negative

identity and is written as Eq. (11).

$$\sin^2\theta + \cos^2\theta = 1 \tag{11}$$

As discussed above, for a given angle θ, measured in radians, there is a corresponding point $P = (a,b)$ on the unit circle. If 2π is added to θ (or any multiple k of 2π to θ), the point on the unit circle corresponding to the angle $\theta + 2\pi k$ is identical to P, and therefore the trigonometric functions of this angle are also the same. A function f is called periodic if there is a positive number p so that $f(\theta + p) = f(\theta)$ for all θ. The smallest number p for which this equation holds is called the period of f. The sine, cosine, secant, and cosecant functions are periodic with period 2π, whereas the tangent and cotangent functions have period π.

Since $\sin\theta = b$, $\cos\theta = a$, $\sin(-\theta) = -b$, and $\cos(-\theta) = a$ (**Fig. 4**), the even-odd properties given in Eqs. (12) and (13) follow.

$$\sin(-\theta) = -\sin\theta \quad \cos(-\theta) = \cos\theta \quad \tan(-\theta) = -\tan\theta \tag{12}$$

$$\csc(-\theta) = -\csc\theta \quad \sec(-\theta) = \sec\theta \quad \cot(-\theta) = -\cot\theta \tag{13}$$

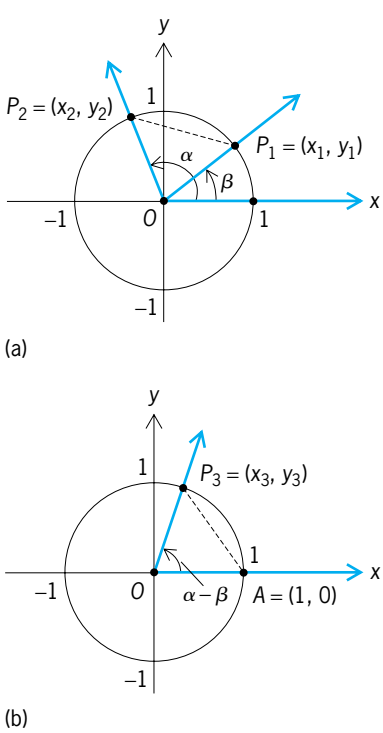

Fig. 5. Geometric construction used to prove the difference formula for cosines. (*a*) Distance between P_1 and P_2 equals (*b*) distance between A and P_3.

Signs of trigonometric functions. Once again, for an angle θ there is a corresponding point $P = (a,b)$ on the unit circle. If the quadrant in which the point P lies is known, then the signs of the trigonometric functions of θ can be determined (**Table 3**).

TABLE 4. Plot points for graph of the sine function

x	$y = \sin x$	(x,y)
0	0	(0,0)
$\pi/6$	1/2	$(\pi/6, 1/2)$
$\pi/2$	1	$(\pi/2, 1)$
$5\pi/6$	1/2	$(5\pi/6, 1/2)$
$\pi\gamma$	0	$(\pi, 0)$
$7\pi/6$	−1/2	$(7\pi/6, -1/2)$
$3\pi/2$	−1	$(3\pi/2, -1)$
$11\pi/6$	−1/2	$(11\pi/6, -1/2)$
2π	0	$(2\pi, 0)$

Sum and difference formulas. The sum and difference formulas for the cosine function are given in Eqs. (14) and (15). A geometric construction (**Fig. 5**)

$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta \tag{14}$$

$$\cos(\alpha - \beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta \tag{15}$$

is used to prove Eqs. (15). The points $A = (1,0)$, P_1, P_2, and P_3 lie on the unit circle, with P_1, P_2, and P_3 on the terminal sides of the angles β, α, and $\alpha - \beta$ respectively. Then the coordinates of P_1, P_2, and P_3 are given in Eqs. (16). The distance formula is used

$$P_1 = (\cos\beta, \sin\beta) \quad P_2 = (\cos\alpha, \sin\alpha) \quad P_3 = (\cos(\alpha - \beta), \sin(\alpha - \beta)) \tag{16}$$

to express the equality of the distances $|P_1, P_2|$ and $|AP_3|$. Then both sides of the resulting equation are squared, and Eq. (11) is used to arrive at the difference formula in Eq. (15). The fact that $\alpha + \beta = \alpha - (-\beta)$, and applying Eqs. (12) in Eq. (15), are used to obtain the sum formula in Eq. (14). *See* ANALYTIC GEOMETRY.

If $\alpha = \pi/2$ and $\beta = \theta$ are substituted in Eq. (15), the use of the values $\sin(\pi/2) = 1$ and $\cos(\pi/2) = 0$ (Table 1) yields Eq. (17). Likewise, if $\alpha = \theta - \pi/2$

$$\cos(\pi/2 - \theta) = \sin\theta \tag{17}$$

and $\beta = \pi/2$ are substituted in Eq. (14), then use of these same values and Eqs. (12) yields Eq. (18).

$$\sin(\pi/2 - \theta) = \cos\theta \tag{18}$$

If $\theta = \alpha + \beta$ is substituted in Eq. (17), then Eqs. (15), (17), and (18) can beused to obtain Eq. (19). The fact that $\alpha + \beta = \alpha - (-\beta)$ and Eqs. (12)

$$\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta \tag{19}$$

can then be used to obtain Eq. (20).

$$\sin(\alpha - \beta) = \sin\alpha\cos\beta - \cos\alpha\sin\beta \tag{20}$$

Product-to-sum formulas. The product-to-sum formulas are Eqs. (21), (22), and (23). Equation (21)

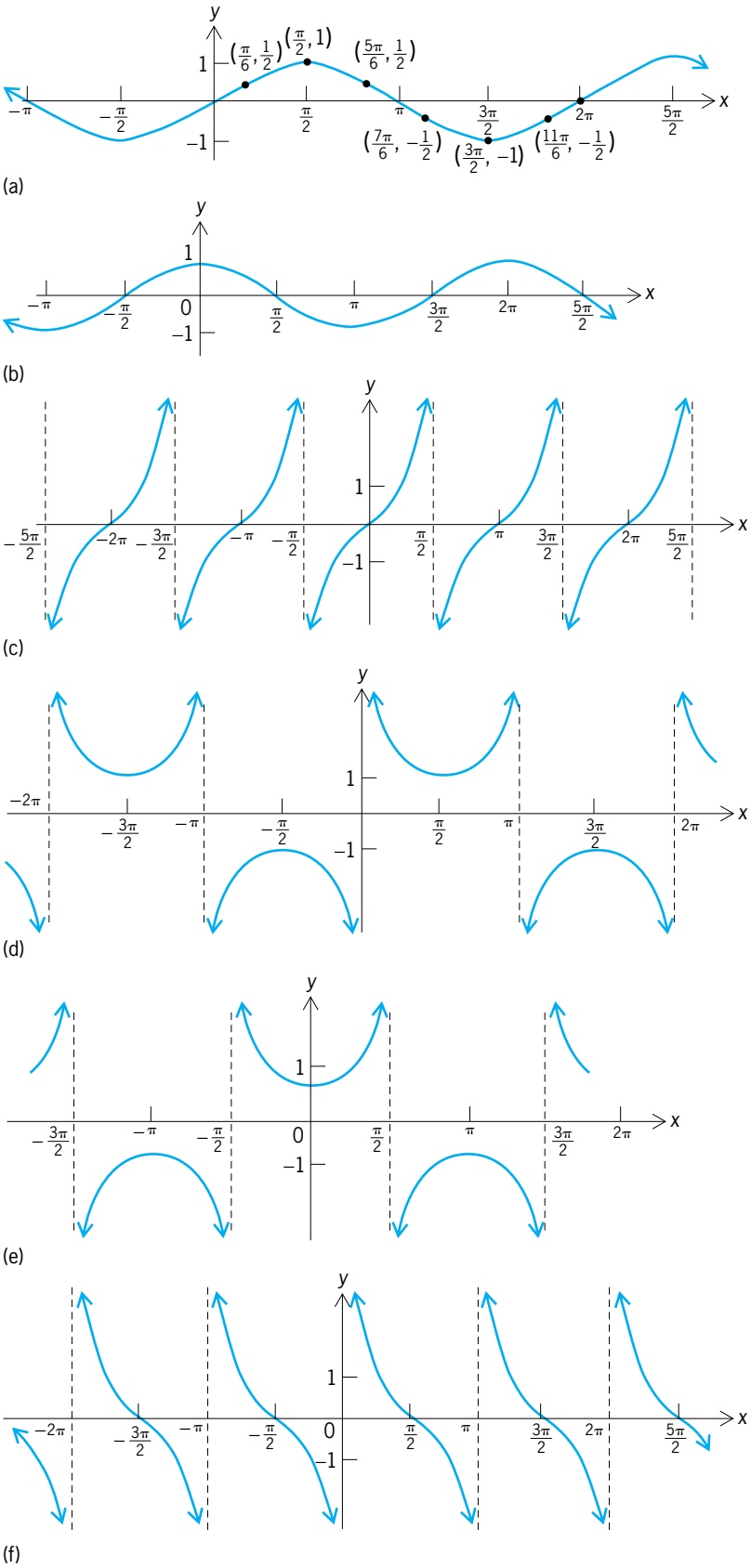

Fig. 6. Graphs of trigonometric functions. (*a*) $y = \sin x$. (*b*) $y = \cos x$. (*c*) $y = \tan x$, x not equal to odd multiples of $\pi/2$. (*d*) $y = \csc x$, x not equal to integral multiples of π. (*e*) $y = \sec x$, x not equal to odd multiples of $\pi/2$. (*f*) $y = \cot x$, x not equal to integral multiples of π.

TABLE 5. Plot points for graph of the tangent function

x	$y = \tan x$	(x,y)
$-\pi/3$	$-\sqrt{3} \approx -1.73$	$(-\pi/3, -\sqrt{3})$
$-\pi/4$	-1	$(-\pi/4, -1)$
$-\pi/6$	$-\sqrt{3}/3 \approx -0.58$	$(-\pi/6, -\sqrt{3}/3)$
0	0	(0,0)
$\pi/6$	$\sqrt{3}/3 \approx 0.58$	$(\pi/6, \sqrt{3}/3)$
$\pi/4$	1	$(\pi/4, 1)$
$\pi/3$	$\sqrt{3} \approx 1.73$	$(\pi/3, \sqrt{3})$

$$\sin\alpha \sin\beta = 1/2[\cos(\alpha - \beta) - \cos(\alpha + \beta)] \quad (21)$$

$$\cos\alpha \cos\beta = 1/2[\cos(\alpha - \beta) + \cos(\alpha + \beta)] \quad (22)$$

$$\sin\alpha \cos\beta = 1/2[\sin(\alpha + \beta) + \sin(\alpha - \beta)] \quad (23)$$

is obtained by subtracting Eq. (14) from Eq. (15); Eq. (22) is obtained by adding Eqs. (14) and (15); and Eq. (23) is obtained by adding Eqs. (19) and (20).

Graphs of trigonometric functions. Because the period of the sine function is 2π, once the graph of $y = \sin x$ is found on the interval $[0, 2\pi]$, the remainder of the graph will consist of repetitions of this portion. A series of plot points (x, y), with the angle x between 0 and 2π can easily be compiled (**Table 4**) from the values of the sine function already obtained (Tables 1 and 2). These points can be plotted and connected with a smooth curve. The complete graph of $y = \sin x$ is obtained by repetition (**Fig. 6***a*). From Eq. (19) it follows that $\sin(x + \pi/2) = \cos x$, and therefore the graph of $y = \cos x$ can be obtained by shifting the graph of $y = \sin x$ to the left by $\pi/2$ units (Fig. 6*b*).

Since the tangent function has period π, once the graph of $y = \tan x$ is found on the interval $(-\pi/2, \pi/2)$, the rest of the graph will consist of repetitions of that graph. Again, a series of plot points can easily be compiled (**Table 5**). If x is close to $\pi/2$ but remains less than $\pi/2$, then $\sin x$ will be close to 1, $\cos x$ will be close to 0, and $\sin x/\cos x = \tan x$ will be positive and large, approaching ∞. Similarly, if x is close to $-\pi/2$, but remains greater than $-\pi/2$, $\tan x$ will approach $-\infty$. The graph of $y = \tan x$ (Fig. 6*c*) has vertical asymptotes at $x = \ldots, -\pi/2, \pi/2, 3\pi/2, \ldots$. The graphs of $y = \csc x$, $y = \sec x$, and $y = \cot x$ are constructed in a similar fashion (Fig. 6*d, e, f*).

The graphs of $y = A \sin \omega x$ and $y = A \cos \omega x$ are called sinusoidal graphs. Since $|\sin x| \leq 1$, it follows

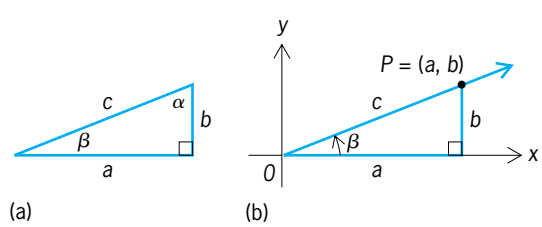

Fig. 7. Right triangle. (*a*) Labeling of sides and angles. (*b*) Relationship of sides to coordinates defining trigonometric functions.

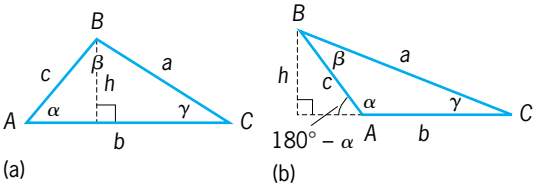

Fig. 8. Labeling of sides, angles, and vertices of oblique triangle, with altitude *h* used in proving the law of sines. (*a*) Angle α is acute. (*b*) Angle α is obtuse.

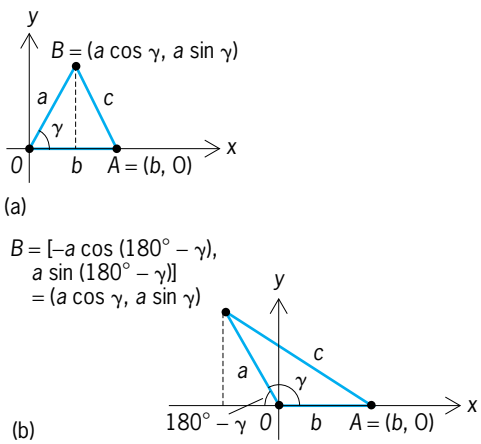

Fig. 9. Triangle and rectangular coordinates used in proving the laws of cosines. (*a*) Angle γ is acute. (*b*) Angle γ is obtuse.

that $-|A| \le A \sin \omega x \le |A|$, so the graph of $y = A \sin \omega x$ lies between $-|A|$ and $|A|$. The number $|A|$ is called the amplitude. The period of $y = A \sin \omega x$ is $2\pi/\omega$, so the graph of $y = A \sin \omega x$ repeats over any interval of length $2\pi/\omega$.

Inverse trigonometric functions. In the equation $x = \sin y$, if y is restricted so that $-\pi/2 \le y \le \pi/2$, then the solution of the equation for y is unique and is denoted by $y = \sin^{-1} x$ (read "y is the inverse sine of x"). Sometimes $y = \sin^{-1} x$ is written as $y = \text{Arcsin}\, x$. Thus, $y = \sin^{-1} x$ is a function whose domain is $-1 \le x \le 1$ and whose range is $-\pi/2 \le y \le \pi/2$. For example, $\sin^{-1} 1/2 = \pi/6$ and $\sin^{-1}(-1) = -\pi/2$.

Likewise in the equation $x = \cos y$, if y is restricted so that $0 \le y \le \pi$, then the solution of the equation for y is unique and is denoted by $y = \cos^{-1} x$ (read "y is the inverse cosine of x"). Thus, $y = \cos^{-1} x$ is a function whose domain is $-1 \le x \le 1$ and whose range is $0 \le y \le \pi$. Finally, in the equation $x = \tan y$, if y is restricted so that $-\pi/2 < y < \pi/2$, then the solution of the equation for y is unique and is denoted by $y = \tan^{-1} x$ (read "y is the inverse tangent of x"). Thus, $y = \tan^{-1} x$ is a function whose domain is $-\infty < x < \infty$ and whose range is $-\pi/2 < y < \pi/2$.

Solution of right triangles. The trigonometric functions can be expressed as ratios of the sides of a right triangle. Indeed, by Eqs. (6), (7), and (8), it follows that $\sin \beta = b/c$, $\cos \beta = a/c$, $\tan \beta = b/a$, and so on, where a and b are the sides adjacent to the right angle, c is the hypotenuse, and α and β are the angles opposite a and b respectively (**Fig. 7**).

If an angle and a side or else two sides of a right triangle are known, then the remaining angles and sides can be found. For any right triangle (Fig. 7*a*), Eqs. (24) hold.

$$c^2 = a^2 + b^2 \quad \alpha + \beta = 90^\circ \tag{24}$$

For example, to measure the height of a radio antenna, a surveyor walks 300 m from its base, and determines the angle of elevation to be 40°. To find the height b it is noted that $\tan 40^\circ = b/300$, so that $b = 300 \tan 40^\circ = 251.73$ m.

Other plane figures with lines as sides can often be solved by drawing perpendiculars to sides of the figure to divide it into right triangles and then solving the right triangles.

Solution of oblique triangles. If none of the angles of a right triangle is a right angle, the triangle is oblique. To solve such triangles, there are four possibilities to consider: (1) one side and two angles are given; (2) two sides and the angle opposite one of them are given; (3) two sides and the included angle are given; and (4) three sides are given. In all of the following discussion, the sides are labeled a, b, and c; the angles opposite these sides are α, β, and γ respectively; and the corresponding vertices are A, B, and C (**Fig. 8**).

Law of sines. The law of sines, Eq. (25), is used

$$\frac{\sin \alpha}{a} = \frac{\sin \beta}{b} = \frac{\sin \gamma}{c} \tag{25}$$

to solve possibilities (1) and (2). To prove the law of sines, an altitude h is drawn from vertex B (Fig. 8). The attitude forms two right triangles yielding Eqs. (26). Solving for h gives part of Eq. (25).

$$\sin \alpha = h/c \quad \sin \gamma = h/a \tag{26}$$

By drawing the altitude h', from vertex A, the rest of Eq. (25) is obtained.

Law of cosines. The law of cosines, used to solve possibilities (3) or (4), may be stated by three equivalent formulas, Eqs. (27), (28), and (29). To prove Eq. (27),

$$c^2 = a^2 + b^2 - 2ab \cos \gamma \tag{27}$$

$$b^2 = a^2 + c^2 - 2ac \cos \beta \tag{28}$$

$$a^2 = b^2 + c^2 - 2bc \cos \alpha \tag{29}$$

a rectangular coordinate system is used in which vertex C lies at the origin and side b lies on the x axis (**Fig. 9**). The coordinates of vertex B are ($a \cos \gamma$, $a \sin \gamma$). The distance formula is used to compute c^2, the square of the distance from $A = (b, 0)$ to B.

Infinite series representation. With the definition $n! = 1 \cdot 2 \cdot 3 \cdots n$, the sine and cosine functions may be represented by the two infinite series, given in Eqs. (30) and (31). These series converge for all x.

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots \tag{30}$$

$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots \tag{31}$$

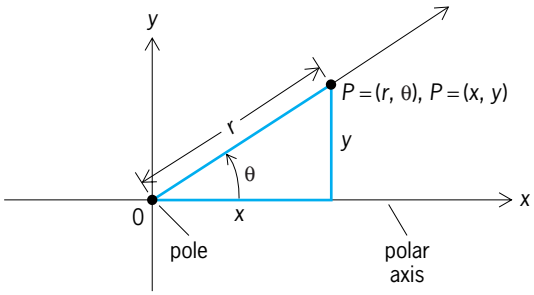

Fig. 10. Relationship between rectangular coordinates (*x*, *y*) and polar coordinates (*r*, *θ*).

Thus, to find a value of sin x or cos x, only as many terms of the series need to be used, as required to ensure required accuracy. The error that results will be less than the numerical value of the first term not used. For example, only the first two terms of the series in Eq. (30) are used to obtain $\sin 0.1 = 0.1 - (0.1)^3/6 = 0.099833$ (nearly). The first unused term, $(0.1)^5/120$ is less than 0.0000001, so the error in sin 0.1 due to only using the first two terms is less than 0.0000001. For five-decimal-place accuracy, only the first four terms of Eqs. (30) and (31) are required. *See* SERIES.

Polar coordinates. In a polar coordinate system, a point, called the pole, is selected, followed by a ray with vertex at the pole, called the polar axis. The pole coincides with the origin of a rectangular coordinate system, and the polar axis coincides with the positive x axis (**Fig. 10**). A point P may be represented by the ordered pair of numbers (r, θ), the polar coordinates of P. The relationship between rectangular coordinates (x, y) and polar coordinates (r, θ) is given by Eqs. (32). Many interesting curves are

$$x = r\cos\theta \quad y = r\sin\theta \tag{32}$$

defined by polar equations including limacons (with and without inner loops), lemniscates, roses, and spirals. *See* LEMNISCATE OF BERNOULLI; PLANE CURVE; ROSE CURVE; SPIRAL.

Complex numbers. A complex number is of the form $z = x + yi$, where i, the imaginary unit, is the number for which $i^2 = -1$. If (x, y) are the rectangular coordinates of a point, then the complex number $z = x + yi$ is in rectangular form. If (r, θ) are the polar coordinates of this point, then the polar form of z is given by Eq. (33). If $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$

$$\begin{aligned} z = x + yi &= r\cos\theta + ir\sin\theta i \\ &= r(\cos\theta + i\sin\theta) \end{aligned} \tag{33}$$

and $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$, then using Eqs. (33), (14), and (19), Eq. (34) follows.Using mathematical

$$z_1 z_2 = r_1 r_2[\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)] \tag{34}$$

induction, DeMoivre's theorem follows: If z is given by Eq. (33), then Eq. (35) is valid for $n \geq 1$, an integer.

$$z^n = r^n(\cos\ n\theta + i\sin\ n\theta) \tag{35}$$

Equation (35) also holds in n is a fraction, allowing the complex roots of a complex number to be found. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

Complex functions. The complex exponential function is defined by Eq. (36). If $z = i\theta$, then Eq. (37)

$$f(z) = e^z = e^x(\cos y + i\sin y) \tag{36}$$

$$e^{i\theta} = \cos\theta + i\sin\theta \tag{37}$$

holds. By using laws of exponents, DeMoivre's theorem, Eq. (38), follows. By substituting various values

$$e^{in\theta} = \cos n\theta + i\sin n\theta = (\cos\theta + i\sin\theta)^n \tag{38}$$

of n, this equation can be used to prove all the laws of trigonometry. *See* E (MATHEMATICS).

Equation (37) can be used to obtain a new trigonometry, in which the sine and cosine functions are defined by Eqs. (39). This theory has many applications in the theory of electricity. *See* ALTERNATING-CURRENT CIRCUIT THEORY.

$$\cos z = \frac{e^{iz} + e^{-iz}}{2} \qquad \sin z = \frac{e^{iz} - e^{-iz}}{2i} \tag{39}$$

Michael Sullivan

Bibliography. J. D. Baley and G. Sarell, *Trigonometry*, 3d ed., 1996; R. V. Churchill and J. W. Brown, *Complex Variables and Applications*, 6th ed., 1996; J. H. Clough-Smith, *An Introduction to Spherical Trigonometry*, 1987; L. Drooyan, W. Hadel, and C. C. Carico, *Trigonometry: An Analytical Approach*, 6th ed., 1990; E. R. Heineman and J. D. Tarwater, *Plane Trigonometry*, 7th ed., 1993; K. J. Smith, *Trigonometry for College Students*, 7th ed., 1998; M. Sullivan, *Trigonometry*, 4th ed., 1996.

Trihedron

A geometric figure bounded by three noncoplanar rays called edges that emanate from a common point called the vertex, and by the plane sectors called faces that are formed by each pair of edges (**Fig. 1**). A trihedron has three dihedrons formed by pairs of face planes, and three face angles formed by pairs of edges. A plane intersecting the edge of a dihedron cuts it into two trihedrons, whose trihedral angles are measures whose sum is the dihedral angle of the

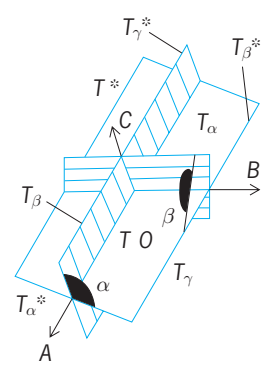

Fig. 1. Trihedron and trihedral angles.

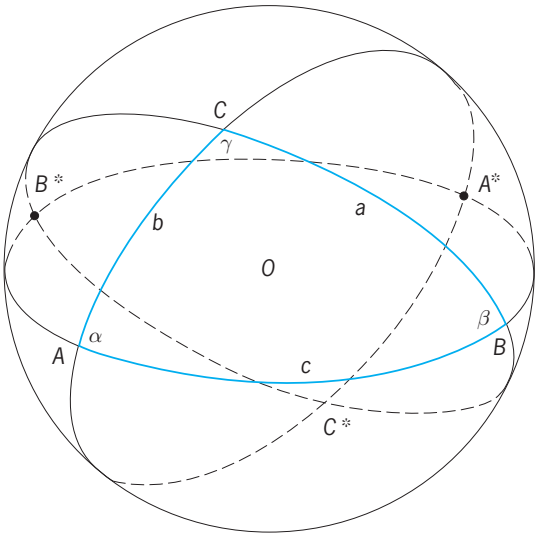

Fig. 2. Spherical triangle *ABC* formed by intersection of sphere with a trihedron. The center of the sphere is at the vertex of the trihedron.

dihedron. Three planes having a common point but not a common line cut space into eight associated trihedrons, of which opposite ones are congruent but not necessarily superposable. (One is the mirror image of the other.) If one of these eight trihedrons has dihedral angles α, β, and γ and trihedral angle σ, its three neighbors that each share one of its faces will have trihedral angles $\alpha - \sigma$, $\beta - \sigma$, and $\gamma - \sigma$. The sum of the four trihedral angles is 180°, so $2\sigma = \alpha + \beta + \gamma - 180^\circ$.

A sphere with its center at the vertex of a trihedron cuts the trihedron in a spherical triangle whose angles α, β, and γ are the dihedral angles of the trihedron and whose sides are measured by the face angles of the trihedron (**Fig. 2**). For a discussion of the relations between these angles and sides *see* TRIGONOMETRY

J. Sutherland Frame

Trilobita

A class of extinct Paleozoic arthropods, occurring in marine rocks of Early Cambrian through late Permian age. Their closest living relatives are the chelicerates, including spiders, mites, and horseshoe crabs (Xiphosura). About 3000 described genera make trilobites one of the most diverse and best-known fossil groups (**Fig. 1**). Species diversity peaked during the Late Cambrian and then declined more or less steadily until the Late Devonian mass extinction. Only four families survived to the Mississippian, and only one lasted until the group's Permian demise. Their dominance in most Cambrian marine settings is essential to biostratigraphic correlation of that system. *See* CAMBRIAN; CHELICERATA; DEVONIAN; PERMIAN.

Trilobites are typically represented in the fossil record by the mineralized portion of their exoskeleton, either as carcass or molt remains. The mineralized exoskeleton (**Fig. 2**) was confined mostly to the dorsal surface, curved under as a rimlike doublure (Figs. 1*b* and 2*b*); a single mineralized ventral plate, the hypostome, was suspended beneath the median region of the head (Fig. 2*b*). The mineralized exoskeleton was composed of low magnesian calcite and a minor component of organic material. Most of the ventral exoskeleton, including the appendages, was unmineralized.

Morphology of exoskeleton. The term "trilobite" refers to the longitudinal division of the body into an axial lobe and two lateral pleural regions (Fig. 2*c*); axial furrows separate the three divisions. The head shield, consisting of up to six fused segments and an anterior presegmental region, is called the cephalon. Its median (axial) lobe contains the glabella, typically convex and indented by a transverse occipital furrow and several pairs of lateral glabellar furrows (Fig. 2*a*). In primitive trilobites, segments of the palpebro-ocular (eye) lobes can be traced across the anterior region of the glabella (Fig. 1*a*).

Trilobites preserve the oldest known visual system in the history of life. Most had rigid compound eyes analogous to those of a housefly. The eyes are situated on the pleural field (genae, or cheeks). Most trilobites had a large number of small eye lenses that shared a single corneal covering (holochroal eye) [Fig. 1*e* and *f*.] The suborder Phacopina, a major Ordovician-through-Devonian group, had large separated lenses (schizochroal eye) [Fig. 1*i*].

In most trilobites, a facial suture, used for molting the exoskeleton, is developed on the dorsal side of the cephalon; it passes from the ventral side usually in front of the glabella, separates the visual surface of the eye from the palpebral lobe, and exits the cephalon in front of, through, or behind the genal angle. These different configurations of the suture are termed proparian (Fig. 1*e–h*), gonatoparian (Fig. 2*a*), and opisthoparian (Figs. 1*j* and 2*c*), respectively. The area between the axial furrow and facial suture is the fixigena (fixed cheek); together with the axial region of the cephalon (including the glabella), this single skeletal part (or sclerite) is the cranidium (Fig. 1*g*). The librigenae (free cheeks) represent the pleural areas outside the facial suture. Most trilobites had the anterior branches of the facial suture separated on the doublure by a rostral plate (Fig. 1*b* and *f*), although some had a median suture and others lost the ventral sutures and fused the doublure medially. The hypostome was rigidly sutured to the roof of the doublure in some groups (the conterminant condition; Fig. 2*b*), but in others it was free and supported by soft tissue (the natant condition).

Anterior wings on the lateral part of the hypostome bear processes that permitted its attachment to a stalk in the cephalic axial furrow (Fig. 2*b*). The thorax is composed of from 2 to more than 60 articulated segments (although typically 6 to 16), each consisting of an axial ring and pleural band. Articulation of the thorax, via processes and sockets on adjacent pleurae, allowed flexibility for enrollment (Fig. 1*i*). The pygidium is a posterior sclerite composed of one or more fused segments. Primitively, it is much smaller than the cephalon, but is enlarged in many

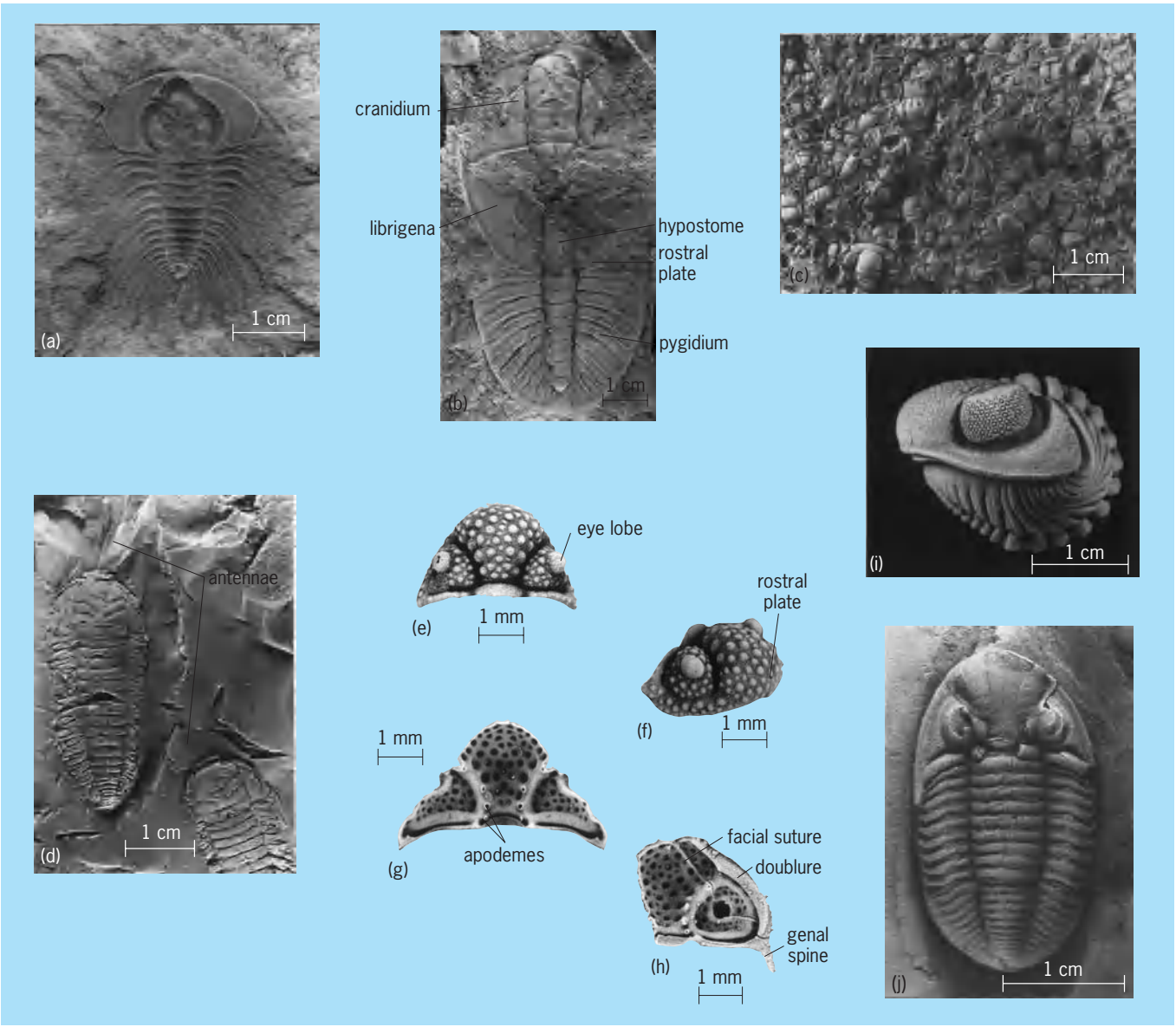

Fig. 1. Trilobite diversity and preservation. (*a*) Olenellid, Lower Cambrian (British Columbia). External mold of exoskeleton in shale. (*b*) *Ogygopsis*, Middle Cambrian (British Columbia). Molt assemblage, with cranidium and pygidium aligned, but librigenae, hypostome, and rostral plate inverted and rotated backward. Thorax is missing. Internal mold. (*c*) Slab of Upper Cambrian limestone (Sweden) with abundant disarticulated trilobite sclerites, mostly cranidia of *Olenus*. (*d*) *Triarthrus*, Upper Ordovician (New York). Exoskeletons replaced by pyrite, preserving antennae. (*e–h*) *Struszia*, Silurian (Northwest Territories), dorsal and anterolateral views of a cephalon, and ventral views of a cranidium and partial cranidium with attached librigena. Exoskeletons replaced by quartz; silicified fossils freed by dissolving limestone in acid. (*i*) *Phacops*, Devonian (Ohio). Enrolled exoskeleton, showing large lenses of schizochroal eye. (*j*) *Griffithides*, Mississippian (Indiana). Dorsal view of exoskeleton.

groups (Fig. 1*b*), and may bear spines along its margin. The cephalic doublure sometimes has notches or furrows that accommodated the pygidium and thoracic tips when the trilobite enrolled (Fig. 2*b*).

Appendages, preserved by pyrite or phosphate replacement or as films on shale, are well known for only a few trilobite species. A single pair of long, jointed antennae (Fig. 1*d*) projected forward from beneath the hypostome. Known Cambrian and Ordovician species have three pairs of postantennal cephalic appendages, while a Devonian example has four. In most cases, these show little structural differentiation from each other, or from postcephalic appendages on each segment along the length of the body (**Fig. 3**). The appendages are biramous, consisting of a jointed walking leg, or telopodite, and a filamentous exite, which attach toward the body axis to a spine-bearing coxa. Appendage-related musculature attached to the ventral exoskeleton at knoblike apodemes (Fig. 1*e* and *h*), just inward of the axial furrow. Enrollment and outstretching were achieved by flexor and extensor muscles; longitudinal, dorsoventral, and horizontal muscles have been observed, as well as a system of intersegmental bars. The exite (gill branch) functioned as a respiratory organ. The mouth opening was positioned above the rear margin of the hypostome and was directed posteriorly. The gut looped backward beneath the glabella, with

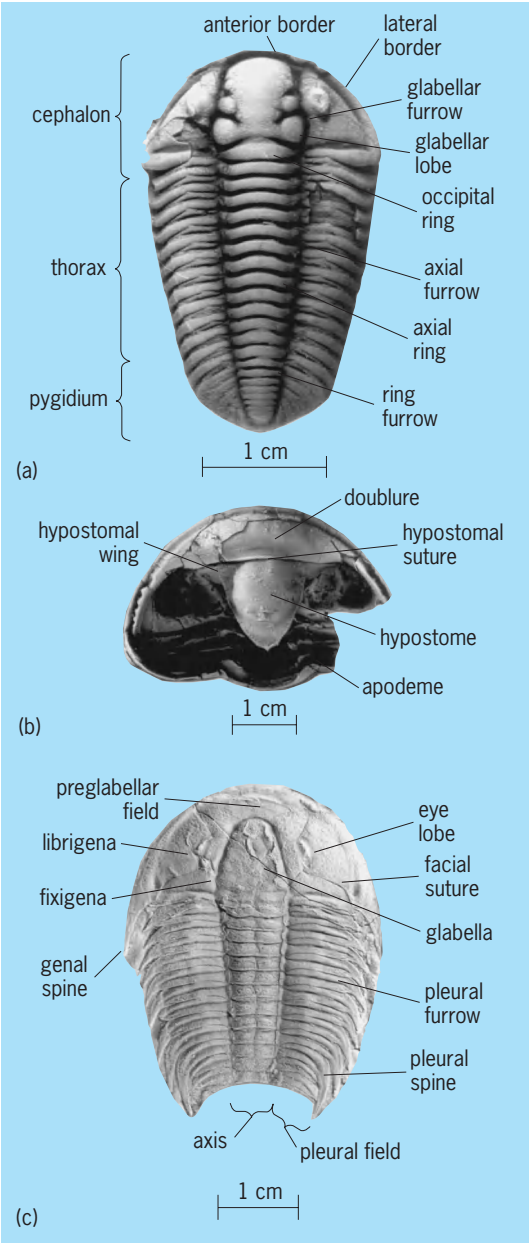

Fig. 2. Morphological features of Trilobita. (*a*) *Calymene*, Silurian (New York). Dorsal view of mineralized exoskeleton. (*b*) *Phacops*, Devonian (Ohio). Ventral (bottom) view of cephalon and anterior segments of thorax, with hypostome attached. (*c*) *Modocia*, Middle Cambrian (Utah). Dorsal view of cephalon and thorax preserved in shale; pygidium and last thoracic segment missing.

the digestive tract extending along the axis to a posterior anus.

Development and molting. Embryonic development of trilobites is unknown. Phosphatized arthropod eggs, which may be those of trilobites, have been discovered in Cambrian rocks. The term "protaspis" is applied to the earliest calcified larval stages, in which the cephalon and protopygidium are fused as an unjointed dorsal shield (**Fig. 4**). Several molts may occur within the protaspid period. The meraspid period is defined by articulation of the cephalon and transitory pygidium as separate sclerites; successive degrees are marked by the release of segments from the anterior part of the transitory pygidium to form the thorax. The holaspid has the complete adult complement of thoracic segments; development in this period is marked by continued increase in size and by changes in shape, but without further addition of segments to the thorax. Adult size ranges from 1.5 mm to 70 cm (0.06 to 28 in.); 2–5 cm (0.8–2 in.) is typical.

Trilobites show the typical arthropod solution to the problem of increasing size with a stiffened exoskeleton: they molted at regular intervals throughout the life cycle. In most species, this was effected by shedding the librigenae along the facial suture and shedding the hypostome. The soft-bodied animal emerged from the resulting gap. Several different molt strategies were employed by different trilobite groups, however, including shedding the entire cephalon, and inverting and rotating various skeletal elements (Fig. 1*b*). Molting results in the typical preservation of trilobite remains as disarticulated sclerites (Fig. 1*c*).

Ecology and macroevolution. Most trilobites were benthic deposit feeders or scavengers, living on the sediment-water interface or shallow-burrowing just beneath it. Some were evidently carnivores,

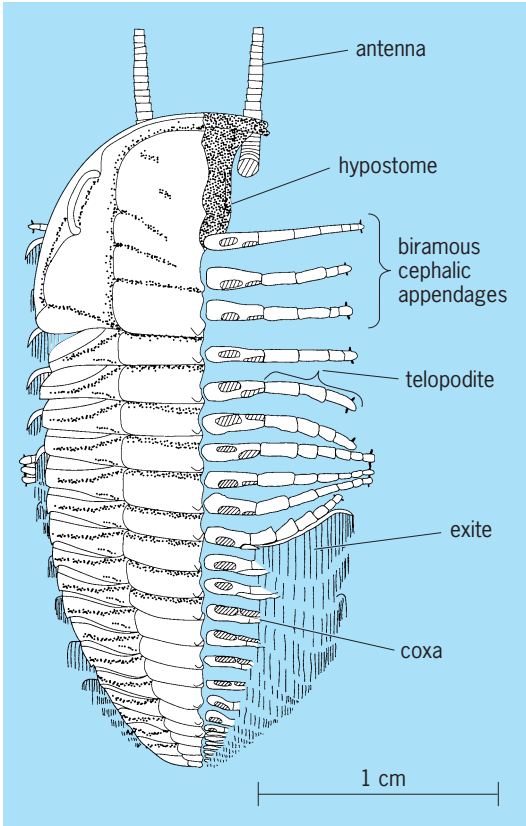

Fig. 3. *Triarthrus eatoni*, Upper Ordovician (New York). Reconstruction with dorsal exoskeleton removed on right side to show appendages. Antennae are incomplete (compare Fig. 1*d*). Exites of first nine postantennal appendages are removed to show structure of telopodite. The mouth was positioned above the posterior margin of the hypostome. (*After H. B. Whittington and J. E. Almond, Appendages and habits of the Upper Ordovician trilobite Triarthrus eatoni, Phil. Trans. Roy. Soc. Lond., B317:28, 1987*)

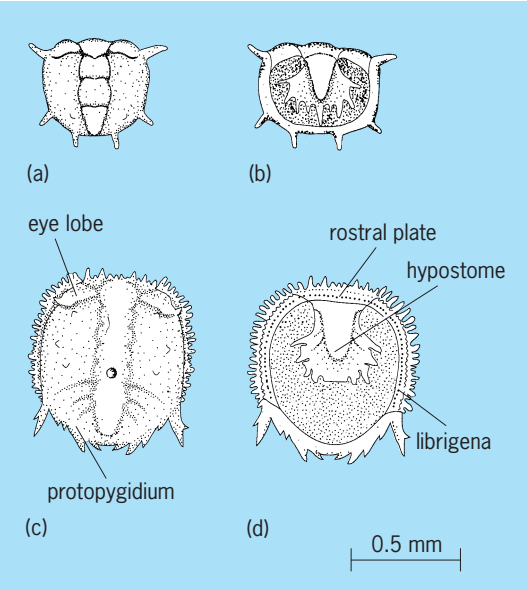

Fig. 4. ***Flexicalymene senaria*, Middle Ordovician (Virginia). Complete exoskeletons of protaspid larvae obtained from silicified residues. (*a*) Dorsal view and (*b*) ventral view of second of four protaspid instars for this species. (*c*) Dorsal view and (*d*) ventral view of fourth protaspid instar. Holaspides closely resemble the related genus *Calymene* (Fig. 2*a*). (*After B. D. E. Chatterton et al., Larvae and relationships of the Calymenina (Trilobita), J. Paleontol., 64:259, 1990*)**

equipped with sharp spines and processes projecting ventrally from their appendages. A few Cambrian and Ordovician groups acquired giant eyes coupled with narrow, streamlined bodies. The morphology and broad geographic and environmental ranges of these groups suggest they were active swimmers. Through their history, trilobites became adapted to all marine environments, from shallow high-energy shorefaces to deep-water, disaerobic habitats.

Trilobites are the most common marine fossils of the Cambrian Period, and their remains typically account for more than 90% of preserved Cambrian fossil assemblages. They were important through the Early Ordovician, but their numerical contribution to onshore communities was much reduced as a result of the Ordovician Radiation of marine life. This event saw filter-feeding organisms (for example, articulate brachiopods, bryozoans, crinoids) proliferate and rapidly evolve to dominate marine communities, a pattern that would last through the Paleozoic Era. Trilobites remained major components of deeper-water communities through the Silurian. Within-habitat, species diversity was generally constant in all environments from the Cambrian through the Silurian, despite their increasingly reduced relative abundance. This indicates that trilobites were largely unaffected by the major events of the Early Paleozoic and that their decline in importance was largely a function of increases in other groups.

Global trilobite diversity increased rapidly following the acquision of hard parts during the Cambrian Explosion, and it peaked during the Late Cambrian. Overall diversity gradually declined during the Ordovician, although a major subset of trilobite groups experienced an evolutionary burst during the Ordovician Radiation. The end-Ordovician mass extinction decimated the group, cutting their global diversity by about half. Surviving families were mainly those that had radiated during the Middle Ordovician. Global diversity continued to decline during the Silurian, although the most speciose trilobite faunas ever found occurred in this period. By the Devonian, trilobites were a relatively minor group, absent from many marine faunas, although still sometimes locally abundant. The Late Devonian mass extinction all but obliterated the trilobites, as only a handful of lineages survived. During the Late Paleozoic, trilobites were typically rare and confined to a limited number of facies. The last trilobites became extinct during the great end-Permian mass extinction. *See* CAMBRIAN; ORDOVICIAN; PALEOZOIC; SILURIAN.

Classification. Trilobita is usually assigned the ranking of class within Arthropoda. Affinities with Chelicerata are expressed by their grouping as Arachnata. Older classifications recognized a phylum or subphylum Trilobitomorpha, grouping Trilobita with an unnatural assortment of trilobite-, chelicerate-, or crustacean-like taxa lumped as Trilobitoidea. The soft-bodied Early-Middle Cambrian order Nectaspida is the closest relative (sister group) of the calcified Trilobita.

The high-level classification of trilobites remains controversial. Post-Cambrian groups (for example, orders Phacopida, Odontopleurida, Lichida, Proetida, Aulacopleurida) are well understood and are grouped into orders or suborders based on distinctive adult and larval morphologies. Cambrian trilobites are generally less well known (despite their abundance as fossils) and have tended to be classified in a small number of large unnatural orders such as Ptychopariida. A particular problem is a lack of understanding of the origins of post-Cambrian orders among Cambrian taxa, a phenomenon termed cryptogenesis. The result is that relationships between named orders of trilobites are essentially unknown. Recent progress has resulted from study of silicified life histories (Fig. 4), but inferring the high-level phylogeny of trilobites remains the cardinal problem in the paleobiology of the group.

A group of blind marine arthropods, the Agnostida, has traditionally been recognized as an order of trilobites. Agnostids share a calcified dorsal exoskeleton with Trilobita, but otherwise lack most diagnostic trilobite features, including a calcified protaspid stage, facial sutures, articulating thoracic segments, and a true transitory pygidium. The appendages of agnostids are also fundamentally unlike those of trilobites. Their affinities are currently debated, with some workers defending their position as ingroup trilobites and others considering the agnostids to be stem group Crustacea. *See* ARTHROPODA; TAXONOMY. Gregory D. Edgecombe; Jonathan Adrain

Bibliography. R. A. Fortey, Ontogeny, hypostome attachment and trilobite classification, *Palaeontology*, 33:529–576, 1990; R. A. Fortey, *Trilobite!*

Eyewitness to Evolution, 2000; R. L. Kaesler (ed.), *Treatise on Invertebrate Paleontology*, pt. O (rev.), vol. 1, 1997; H. B. Whittington, *Trilobites*, 1990.

Trimerophytopsida

Mid-Early-Devonian into Mid-Devonian vascular plants at a higher evolutionary level than Rhyniopsida. Branching was profuse and varied, dichotomous, pseudomonopodial, helical to subopposite and almost whorled, and often trifurcate (see **illus.**).

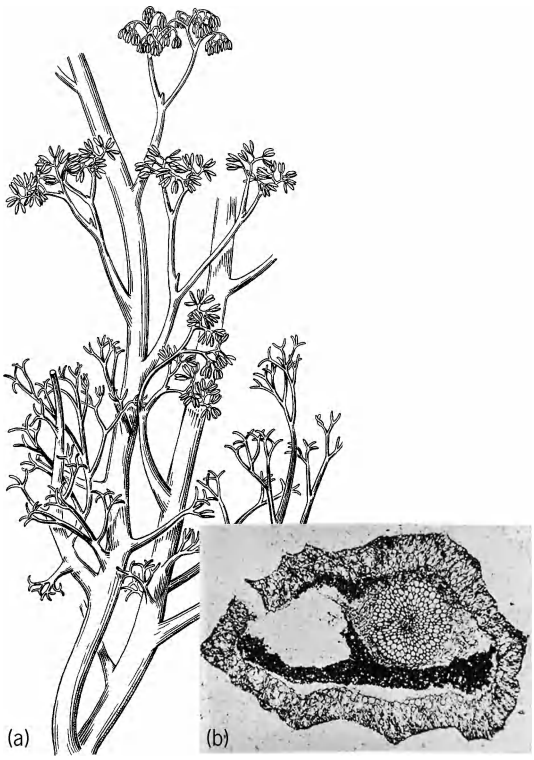

***Psilophyton dawsonii*. (*a*) Reconstruction of known portion of a plant showing leafless stem, lateral branch systems terminating in large clusters of fusiform sporangia that dehisce longitudinally, and vegetative branching (*after H. P. Banks, H. P. S. Leclercq, and F. M. Hueber, Anatomy and morphology of Psilophyton dawsonii sp. nov., from the Late Lower Devonian of Quebec (Gaspé), and Ontario, Canada, Paleontol. Amer., 8:73–77, 1975*). (*b*) Photomicrograph of a transverse section of one stem. Outer cortex and xylem only are petrified. Centrally located, smaller cells indicate the maturation of xylem was centrarch.**

Vegetative branches were often in a tight helix, terminated by tiny recurved branchlets simulating leaf precursors. The axes were leafless and glabrous or spiny. Fertile branches were trifurcate or dichotomized several times in three planes and terminated in loose or tight clusters of 32–256 fusiform sporangia that dehisced longitudinally. The number of sporangia varied, in part, by abortions or failure of dichotomies. Spores were uniform in *Apiculi retusispora* or *Retusotriletes*, depending on preservation. Xylem is known only in *Psilophyton*, and was centrarch. Tracheids were helical, circular-bordered and scalariform-bordered pitted, and peculiarly multiaperturate. *Trimerophyton* and *Pertica* are two other natural genera. Form genera include some species of *Dawsonites*, *Hostinella*, and possibly *Psilophytites* and *Psilodendrion*. Progymnosperms, ferns, and articulates are derived from this group. *See* EMBRYOBIONTA; PSILOTOPHYTA; RHYNIOPHYTA; RHYNIOPSIDA.

Harlan P. Banks

Triple point

A particular temperature and pressure at which three different phases of one substance can coexist in equilibrium. In common usage these three phases are normally solid, liquid, and gas, although triple points can also occur with two solid phases and one liquid phase, with two solid phases and one gas phase, or with three solid phases.

According to the Gibbs phase rule, a three-phase situation in a one-component system has no degrees of freedom (that is, it is invariant). Consequently, a triple point occurs at a unique temperature and pressure, because any change in either variable will result in the disappearance of at least one of the three phases. *See* PHASE EQUILIBRIUM.

Triple points are shown in the **illustration** of part of the phase diagram for water. Point *A* is the well-known triple point for Ice I (the ordinary low-pressure solid form) + liquid + water + water vapor at 0.01°C (273.16 K) and a pressure of 0.00603 atm (4.58 mmHg or 611 pascals). In 1954 the thermodynamic temperature scale (the absolute or Kelvin scale) was redefined by setting this triple-point temperature for water equal to exactly 273.16 K. Thus, the kelvin (K), the unit of thermodynamic temperature, is defined to be 1/273.16 of the thermodynamic temperature of this triple point.

Point *B*, at 251.1 K (−7.6°F) and 2047 atm (207.4 megapascals) pressure, is the triple point for liquid

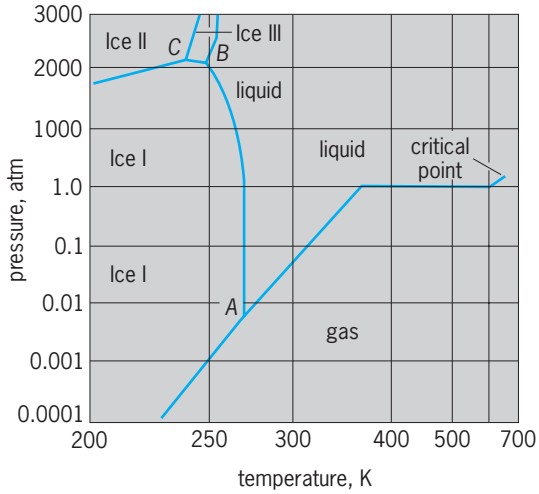

Phase diagram for water, showing gas, liquid, and several solid (ice) phases; triple points at *A, B,* and C. The pressure scale changes at 1 atm from logarithmic scale at low pressure to linear at high pressure. 1 atm = 100 kPa; °F = (K × 1.8) − 459.67.

water + Ice I + Ice III; and point *C*, at 238.4 K (−31°F) and 2100 atm (212.8 MPa) pressure, is the triple point for Ice I + Ice II + Ice III. At least four other triple points are known at higher pressures, involving other crystalline forms of ice.

For most substances the solid-liquid-vapor triple point has a pressure less than 1 atm (about 100 kilopascals); such substances then have a liquid-vapor transition at 1 atm (normal boiling point). However, if this triple point has a pressure above 1 atm, the substance passes directly from solid to vapor at 1 atm. *See* SUBLIMATION.

For a two-component system, the invariant point in a phase diagram is a quadruple point at which four phases coexist. The three-phase situation is then represented by a line in the three-dimensional pressure-temperature-composition diagram. *See* BOILING POINT; ICE POINT; MELTING POINT; TRANSITION POINT; VAPOR PRESSURE; WATER.

Robert L. Scott

Triplet state

A molecule exists in this electronic state when its total spin angular momentum quantum number S is equal to one. The triplet state is an important intermediate of organic chemistry. In addition to the wide range of triplet molecules available through photochemical excitation techniques, numerous molecules exist in stable triplet ground states, for example, oxygen molecules. Theoretical calculations, furthermore, make predictions concerning the spin multiplicities of the ground states of many prototype organic molecules such as cyclobutadiene, trimethylene methane, and methylene, and indicate that they will be triplets. *See* ATOMIC STRUCTURE AND SPECTRA; REACTIVE INTERMEDIATES; SPIN (QUANTUM MECHANICS).

Practical definition. A good working definition of a triplet state for the chemist is the following: A triplet is a paramagnetic even-electron species which possesses three distinct but energetically similar electronic states as a result of the magnetic interaction of two unpaired electron spins. The several important terms of this definition allow some insight as to the essential features of a triplet. First of all, a triplet is paramagnetic, and should thus display this property in a magnetic field. This paramagnetism serves as the basis for experimental magnetic susceptibility and electron spin resonance studies of the triplet state. However, one can imagine many paramagnetic odd-electron species which are not triplets, for example, nitric oxide. Thus, the criterion that a triplet must also be an even-electron species is apparent.

However, one can imagine paramagnetic, even-electron species which possess (1) only two distinct electronic states or (2) five or more electronic states. The former occurs when the paramagnetism results from two electrons which act as two independent odd electrons. For example, two carbon radicals separated by a long saturated chain will behave as two

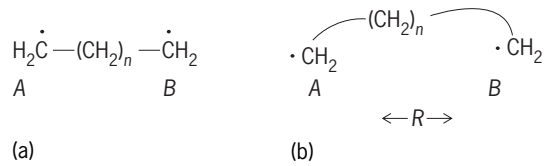

Fig. 1. Two carbon radicals connected by a long methylene chain. (*a*) Biradical state. (*b*) Triplet state.

doublet states if there is sufficient separation to prevent spin interactions. Five or more electronic states result when four or six parallel electronic spins interact (to yield quintet and septet states, respectively). *See* ELECTRON SPIN; PARAMAGNETISM.

One can now see that conceptual difficulties may arise in differentiating a biradical state (that is, a species possessing two independent odd-electron sites) from a triplet. Suppose two carbon radicals are separated by a long methylene chain as in **Fig. 1***a*. If the methylene chain is sufficiently long and the odd-electron centers are so far removed from one another that they do not interact (magnetically and electronically) with one another, then the system is a doublet of doublets, that is, two independent odd electrons or a true biradical. If the methylene chain should be folded (Fig. 1*b*) so that the odd electrons begin to interact (magnetically and electronically) with one another, then at some distance R between the —CH_2 groups the doublet or doublets will become a triplet state. This state will result from the fact that the spin of the electron on carbon A is no longer independent of the spin on carbon B. Since the spins are quantized, selection rule (1) applies, where S is

$$\text{Number of spin states} = 2\,|S| + 1 \qquad (1)$$

the sum of the spin quantum numbers for the two electrons. This means that either three spin states (if $S = 1$ or -1, that is, spins of both electrons on C_A and C_B are the same) or one spin state (if $S = 0$, that is, spin of the electron of C_A is paired with that of C_B) will result. The former describes a triplet state, and the latter a singlet.

This leads to a difficulty in terminology: The "triplet state" is not one state but three states even in the absence of an external magnetic field. Indeed, under favorable conditions transitions may be observed between triplet levels at zero external magnetic field. The effect of an external magnetic field is to further split the triplet levels and allow transitions between them to be more easily detected.

Properties. A triplet may result whenever a molecule possesses two electrons which are both orbitally unpaired and spin unpaired. As shown in **Fig. 2**, orbital unpairing of electrons results when a molecule absorbs a photon of visible or ultraviolet light. Direct formation of a triplet as a result of this photon absorption is a very improbable process since both the orbit and spin of the electron would have to change simultaneously. Thus, a singlet state is generally formed by absorption of light. However, quite often the lifetime of this singlet state is

sufficiently long to allow the spin of one of the two electrons to invert, thereby producing a triplet. The following discussion considers the ways in which such a species is unambiguously characteristic. *See* MOLECULAR ORBITAL THEORY.

The question to be answered is: What are the general properties to be expected of a molecule in the triplet state? Some of the more important physical properties are (1) paramagnetism; (2) absorption between triplet sublevels; (3) electronic absorption from the lowest triplet to upper triplets; (4) electronic emission from the lowest triplet to a lower singlet ground state (if the triplet level is not the ground state). The paramagnetism of the triplet results from the interaction of unpaired spins and the fact that an unpaired spin shows a paramagnetic effect (is attracted) in a magnetic field.

Absorption between triplet sublevels may be observed directly by the use of an electron spin resonance spectrometer. *See* ELECTRON PARAMAGNETIC RESONANCE (EPR) SPECTROSCOPY.

The triplet, like any other electronic state, may be excited to upper electronic states of the same spin as the result of light absorption. In favorable cases this may be observed by the method of flash spectroscopy. *See* PHOTOCHEMISTRY.

For most organic molecules the lowest triplet state is an excited electronic state and may emit light and pass to the ground singlet state. Since light absorption to form a triplet from a singlet is improbable, the symmetrically related emission of light from a triplet returning to a ground state is likewise improbable. Indeed, it takes the triplet states of some aromatic molecules an average of about 30 s to emit light. This phenomenon is known as phosphorescence and is to be contrasted with fluorescence, the emission of light from an excited singlet state returning to a singlet ground state, a process which often occurs in nanoseconds. *See* FLUORESCENCE; PHOSPHORESCENCE.

Although phosphorescence (long-lived emission) was the first method employed to study triplets, it is not a specific device for establishing whether a long-lived emission occurs from a triplet. For instance, examples are known for which the slow combination of positive and negative sites will generate excited molecules which emit light. In this case the combination reaction may be rate-determining for light emission.

Similarly, absorption from one triplet to another is not a specific method since the precise triplet-triplet absorption characteristics cannot be predicted accurately. It would thus remain to be proven that the absorbing species is indeed a triplet and not some other transient species.

Even paramagnetism is not an infallible probe for a triplet state since free radicals which are also paramagnetic are often produced by the absorption of light.

It appears that electron spin resonance (ESR) is probably the most powerful single method for establishing that a molecule is in its triplet state. The nature of the ESR signals may be predicted and fitted to theoretical relation (2), which describes the

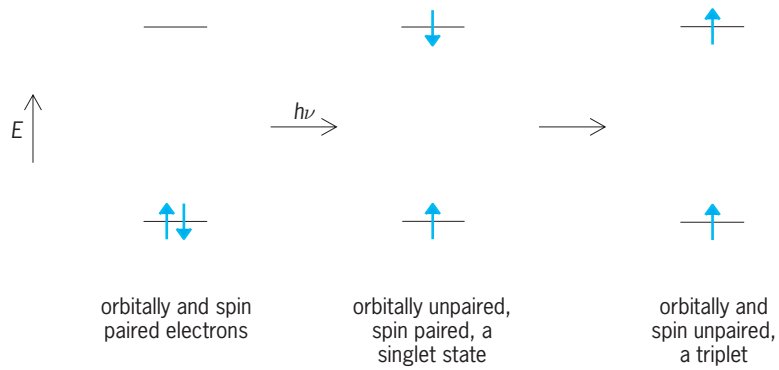

Fig. 2. Simple molecular orbital description of singlets and triplets.

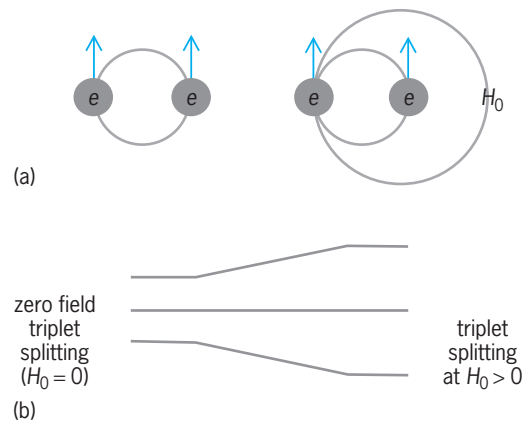

Fig. 3. Triplet state resulting from (*a*) internal spin-spin dipolar interaction (*D* + *E*) and (*b*) external interaction of electronic spin with magnetic field H_0.

$$H = g_0 H \cdot S + DS_z^2 + E(S_x^2 - S_y^2) \qquad (2)$$

magnetic spin interactions and expected absorptions. Here g_0 is the Landé g factor, D the dielectric constant, and E the electric field strength of the molecule. This particular equation is derived for the special case of molecules with a plane of symmetry and a symmetry axis perpendicular to that plane. However, the important general features of this equation are (1) the term $g_0 H{\cdot}S$ which describes the interaction of the external magnetic field H with the unpaired electron spin S; (2) the term $DS^2{}_z + E(S^2{}_x - S^2{}_y)$ which describes the spin-spin dipolar interactions along the x, y, and z axes of the molecule. These interactions are indicated in **Fig. 3**. Thus, from a study of the behavior of a triplet in a magnetic field, information on the electronic distribution in this excited state is obtained. In favorable cases, the nuclear geometry of the triplet may be derived.

Nicholas J. Turro

Bibliography. A. Devaquet et al., *Triplet States One*, 1975; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., vol. 8, 1993; N. J. Turro, *Modern Molecular Photochemistry*, 1981; P. J. Wagner et al.,

Triplet States: No. 3, 1976; U. P. Wild et al., *Triplet States Two*, 1975.

Tripylida

An order of nematodes in which the cephalic cuticle is simple and not duplicated; there is no helmet. The body cuticle is smooth or sometimes superficially annulated. Cepahalic sensilla follow the typical pattern in which one whorl is circumoral and the second whorl is often the combination of circlets two and three. The pouchlike amphids have apertures that are inconspicuous or transversally oval. The stoma is variable, being simple, collapsed, funnel shaped or cylindrical, and armed or unarmed. In most taxa the stoma is surrounded by esophageal tissue; that is, it is entirely esophastome. When the stoma is expanded, both the cheilostome and esophastome are evident. Esophagi are cylindrical-conoid. Esophageal glands open anterior to the nerve ring. Males generally have three supplementary organs, more in some taxa. A gubernaculum accompanies the spicules. Caudal glands are generally present.

There are two tripylid superfamilies, Tripyloidea and Ironoidea. The characteristically well-developed cuticular annulation of the Tripyloidea is only rarely seen in other Enoplida. These nematodes are commonly found in fresh water or very moist soils; however, some are found in brackish water and marine habitats. Intestinal contents indicate that their food consists primarily of small microfauna that often include nematodes and rotifers. The Ironoidea contains species (presumably carnivorous) occurring in both fresh-water and soil habitats. *See* NEMATA (NEMATODA).

Armand R. Maggenti

Triterpene

A hydrocarbon or its oxygenated analog containing 30 carbon atoms and composed of six isoprene units. Triterpenes form the largest group of terpenoids, but are classified into only a few major categories. Resins and saps contain triterpenes in the free state as well as in the form of esters and glycosides.

Biogenetically triterpenes arise by the cyclization of squalene (**1**) and subsequent skeletal rearrange-

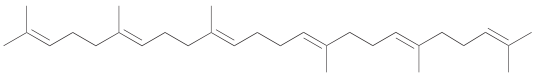

(**1**)

ments. Squalene can cyclize in five ways, leading to different stereochemical arrangements in the final triterpenoid structure. The conformation of a triterpenoid nucleus is determined in the initial folding of squalene into several chair or boat configurations. Apart from the linear squalene itself and some bicyclic, highly substituted skeletons, most triterpenes are either tetracyclic or pentacyclic compounds. The various structural classes are designated by the names of representative members. The lanosterol and euphol series comprise tetracyclic structures differing only in the stereochemical arrangements around the D ring. Lanesterol (**2**) occurs in sheep wool, and euphol (**3**) is obtained from Euphorbium

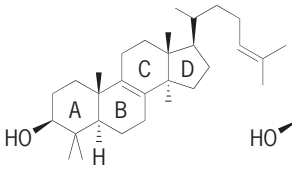

(**2**)

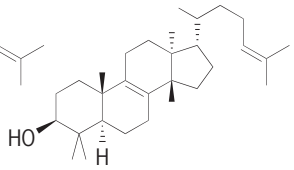

(**3**)

resin. The oleanane (β-amyrin, **4**, from grape seeds) and urasane (α-amyrin **5**, from Manila elemi resin)

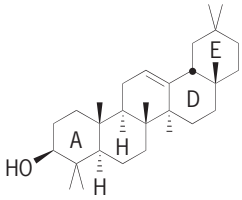

(**4**)

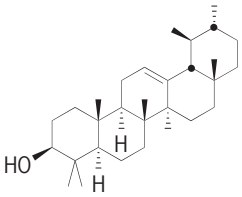

(**5**)

series of triterpenes are all pentacycles differing in the substitution pattern of methyl groups on ring E. Lupeol (**6**) from lupin seeds and hydroxyhopanone (**7**) from dammar resin are typical of

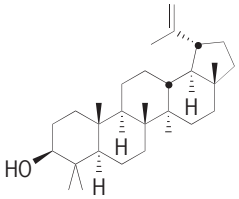

(**6**)

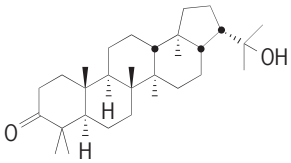

(**7**)

the lupane and hopane series. Hydroxyhopanone is one of the few triterpenoids that results from the cyclization of squalene without subsequent rearrangements.

Steroids and sterols are related to triterpenes. These important compounds are classed as nortriterpenes to indicate that they lack some of the 30 carbons of the triterpene skeleton. They have 27 to 29 carbons, and in most cases lack the geminal dimethyl group in ring A and one of the angular methyl groups in ring D of the lanosterol tetracyclic skeleton. Steroids are believed to arise from squalene via lanosterol, followed by the oxidative loss of methyl groups. Because of their significance in mammalian metabolism, steroids are usually treated as a separate class although their terpenoid origin is well understood. Cholesterol (**8**) and cholic acid (**9**)

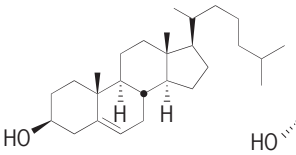

(**8**)

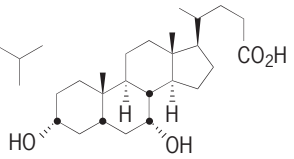

(**9**)

are representative examples of steroid structures. *See* SQUALENE; STEROID; TERPENE.

Tomas Hudlicky

Triticale

A cereal grass plant (× *Triticosecale*) obtained from hybridization of wheat (*Triticum*) with rye (*Secale cereale*). It is a crop plant with a small-seeded cereal grain that is used for human food and livestock feed. Worldwide, triticale is slowly gaining importance as a cereal grain. In 1998 it was estimated that about 2.9 million hectares (7.2 million acres) were planted each year. The European continent dominates triticale production with 70% of the total area. Countries with the greatest production (more than 100,000 ha, or 247,000 acres) are Poland, Russia, Germany, France, Australia, Brazil, and the United States.

Origin and types. Triticale was first developed in 1876, but not until the 1960s were types developed that were suitable for cultivation. Modern varieties are called secondary triticales because they were selected after interbreeding of various triticales, including primary types. In some triticale varieties, one or more rye chromosomes have been replaced by wheat chromosomes, giving secondary-substituted triticales, as contrasted to complete triticale having all seven rye chromosomes.

Triticale is produced by deliberate hybridization of either bread wheat [*Triticum aestivum*; diploid number of chromosomes ($2n$) = 42] or durum wheat (*T. turgidum* var. *durum*; $2n = 28$) with rye ($2n = 14$), followed by the doubling of the chromosome number of the hybrid plant (see **illus.**). The drug colchicine is commonly applied at a very low concentration to the seedling hybrid plants to cause the chromosome number to double. Hexaploid triticale (durum wheat × rye; $2n = 42$) is a more successful crop plant than octoploid triticale. The octoploid form ($2n = 56$) is produced by hybridization of bread wheat ($2n = 42$) with rye ($2n = 14$). A third type ($2n = 28$) from hybridizing diploid wheat (*T. monococcum*; $2n = 14$) with rye is of no economic consequence.

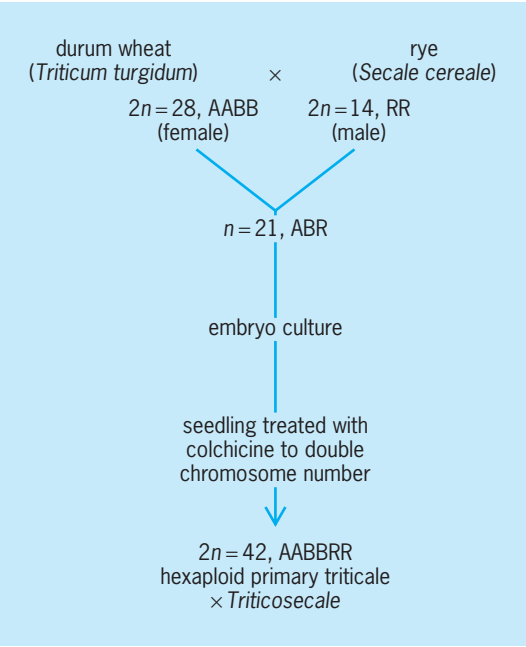

Flow diagram for hexaploid triticale development showing chromosome numbers and genome identifications. The octoploid form ($2n = 56$) is produced in a similar way, but starting with bread wheat (*Triticum aestivum*, $2n = 42$, AABBDD) as the female parent.

Adaptation. Triticale is grown from seeds sown in soil by using cultivation practices similar to those of wheat or rye. Both winter-hardy and nonhardy types exist, the latter used where winters are mild or for spring sowing. Triticale tends to have a greater ability than wheat to grow in adverse environments, such as saline or acid soils or under droughty conditions. While some triticales are resistant to attack by disease pathogens, there is great variation in resistance, just as in wheat. Ergot (caused by the fungus *Claviceps purpurea*) can infect triticale in some localities, probably to a higher degree than wheat. *See* ERGOT AND ERGOTISM.

Usage. Being a cereal grain, triticale can be used in food products made from wheat flour. Varieties tend to have large, somewhat irregularly shaped grains that produce a lower yield of milled flour than wheat. Bread and pastry products can be made very well with triticale flour. Experimental trials have shown triticale grain also to be acceptable as a starch source for beermaking.

As a livestock feed, triticale grain is a good source of carbohydrate and protein. Its protein is richer than wheat protein in lysine, one of the amino acids, making it especially desirable for the diets of monogastric animals, such as swine and birds, since these animals cannot synthesize lysine and so must obtain it in their diet. Triticale is also becoming more widely used as a fodder (forage) source for livestock feeding. Whole plants are harvested while still green, before the grain is mature. The fodder is fed mainly to cattle in one of three ways: immediately after harvest (green chop), after fermentation as silage, or after drying as hay. Triticale is also a pasture crop; sheep and cattle are allowed to graze on triticale during the early growth phase of the crop. However, triticale, or any other types of grain, infested with ergot bodies should not be eaten by humans or livestock.

Intense breeding and selection have made very rapid genetic improvements in triticale seed quality. Some triticale varieties now have seeds with hardness and shape similar to wheat. The gluten proteins of wheat are being incorporated into triticale so that triticale flour may be substituted for wheat in certain baked products. The agronomic advantages and improved end-use properties of the grain of triticale over wheat make triticale an attractive option for increasing global food production. *See* BREEDING (PLANT); RYE; WHEAT.

Calvin O. Qualset

Bibliography. M. Bernard and S. Bernard (eds.), *Genetics and Breeding of Triticale*, Institut National de la Recherche Agronomique, Paris, 1985; H. Guedes-Pinto, N. Darvey, and V. P. Carnide (eds.), *Triticale: Today and Tomorrow*, Kluwer Academic Publishers, Dordrecht, 1996; P. Juskiw (ed.), *4th International Triticale Symposium*, International Triticale Association, Red Deer, Alberta, Canada, 1998; A. Müntzing,

Advances in Plant Breeding, no. 10, supplement to *J. Plant Breed*. (Berlin), 1979; National Research Council, Board on Science and Technology for International Development, *Triticale: A Promising Addition to the World's Cereal Grains*, 1989.

Tritium

The heaviest isotope of the element hydrogen and the only one which is radioactive. Tritium occurs in very small amounts in nature but is generally prepared artificially by processes known as nuclear transmutations. It is widely used as a tracer in chemical and biological research and is a component of the so-called thermonuclear or hydrogen bomb. It is commonly represented by the symbol $^{3}{}_{1}H$, indicating that it has an atomic number of 1 and an atomic mass of 3, or by the special symbol T. For information about the other hydrogen isotopes. *See* DEUTERIUM; HYDROGEN.

Properties. Both molecular tritium, T_2, and its counterpart hydrogen, H_2, are gases under ordinary conditions. Because of the great difference in mass, many of the properties of tritium differ substantially from those of ordinary hydrogen.

Chemically, tritium behaves quite similarly to hydrogen. However, because of its larger mass, many of its reactions take place more slowly than do those of hydrogen. The ratio of reaction rates may be as large as 64:1. These differences in reactivity can give rise to serious errors of interpretation when tritium is used as a tracer for hydrogen.

The nucleus of the tritium atom, often called a triton and symbolized *t*, consists of a proton and two neutrons. It has a mass of 3.01700 atomic mass units (amu), a nuclear spin of 1/2, and a magnetic moment of 2.9788 nuclear magnetons. It undergoes radioactive decay by emission of a beta particle to leave a helium nucleus of mass 3. No gamma rays are emitted in this process. The half-life for the decay is 12.26 years. The most energetic of the beta particles emitted by tritium have the comparatively low energy of 18.6 keV; beta particles are completely stopped by 7 mm of air or by 0.01 mm of paper or similar material. The average energy of the beta particles is 5.69 keV.

When tritium is bombarded with deuterons of sufficient energy, a nuclear reaction known as fusion occurs and energy considerably greater than that of the bombarding particle is released. The reaction may be written as (1). This reaction is one of those which

$$^{3}_{1}H + ^{2}_{1}H \rightarrow ^{4}_{2}He + ^{1}_{0}n + 18\ \text{MeV} \qquad (1)$$

supply the energy of the thermonuclear bomb. It is also of major importance in the development of controlled thermonuclear reactors. Enormous quantities of tritium will be required if such reactors are perfected and brought into use as electric power generators.

Compounds. Very few compounds of pure tritium have been prepared and studied. Such compounds would undergo decomposition quite rapidly under the action of the tritium beta radiation. Tritium oxide, T_2O, has been prepared by oxidation of tritium gas with hot copper oxide or by passing an electric spark through a mixture of tritium and oxygen. Its melting point is 4.49°C (40.08°F), compared with 0°C (32°F) for ordinary water. Of much greater importance are compounds, especially organic compounds, in which a small fraction of the hydrogen atoms have been replaced by tritium. Such labeled compounds are employed in tracer studies, such as those indicated above. Tritium-labeled compounds may be prepared by ordinary synthetic chemical methods, such as the catalytic addition of tritium-hydrogen mixtures to unsaturated compounds. Tritium may be exchanged for hydrogen in the presence of a catalyst such as platinum or a strong acid. In recoil labeling, a mixture of an organic compound and a lithium salt are irradiated with neutrons in a nuclear reactor; some of the energetic tritons produced are incorporated into the organic compound.

Another important labeling procedure consists of the exposure of an organic compound to tritium gas in a sealed vessel; the tritium beta radiation facilitates the exchange of hydrogen in the compound with tritium in the gas. Some compounds of biological interest have been prepared by growing organisms in tritiated water.

Analysis. Because of its weak beta radiation, tritium is not readily measured by the ordinary Geiger-Müller counter. More efficacious is the introduction of tritium as a gas inside the counting tube. Alternatively, the ionization of a gas caused by the beta radiation may be measured in an ionization chamber, or the tritium compound may be dissolved in a suitable solvent containing a phosphor and the light pulses excited by the beta particles then may be counted with a scintillation counter. Tritium gas containing only small amounts of ordinary hydrogen may be analyzed with a mass spectrometer or by measuring the density of the gas. Because of the very short range of the tritium beta particle, autoradiography, the exposure of radioactive material to a photographic plate, is often used to locate precisely the position of tritium in biological material.

Natural occurrence. Before the start of thermonuclear weapons testing in 1954, rainwater contained approximately 1–10 atoms of tritium per 10^{18} atoms of hydrogen. Such tritium originates largely from the bombardment of nitrogen in the upper atmosphere with neutrons and protons from cosmic rays, as in reaction (2). Because the half-life of tritium is short

$$^{14}_{7}N + ^{1}_{0}n \rightarrow ^{3}_{1}H + ^{12}_{6}C \qquad (2)$$

in comparison with the time required for mixing of the ocean waters, the concentration of tritium in the ocean is much lower than in rainwater. Before 1954 the total amount of tritium on the Earth's surface was estimated at 1800 g (63 oz), of which about 11 g (0.39 oz) was in the atmosphere and 13 g (0.46 oz) in ground waters. Testing of thermonuclear weapons resulted in sharp rises in the tritium content of

rainwater to values as high as 500 atoms per 10^{18} atoms of hydrogen.

Preparation. Tritium was first produced in the laboratory by bombarding compounds of deuterium with high-energy deuterons, as in reaction (3). A

$$^{2}_{1}H + ^{2}_{1}H \rightarrow ^{3}_{1}H + ^{1}_{1}H \quad (3)$$

number of other nuclear reactions also give rise to tritium. The most important of these is the absorption of slow neutrons by the lithium isotope of mass 6, according to reaction (4). By irradiating enriched

$$^{6}_{3}Li + ^{1}_{0}n \rightarrow ^{3}_{1}H + ^{4}_{2}He \quad (4)$$

lithium-6, in the form of an alloy with magnesium or aluminum, with neutrons from a nuclear reactor, tritium may be prepared on a large scale.

Uses. As a result of its production for use in nuclear weapons, tritium became available in large quantities at very low cost. It is used in admixture with zinc sulfide in the production of luminous paints, which have largely replaced the radium formerly used on watch dials; such mixtures are also used to produce small, permanent light sources. Tritium adsorbed on metals is used in targets for the production of fast neutrons by bombardment with deuterons. Tritium has been much used in hydrological studies, since it is an ideal tracer for water movement. Some studies depend on natural tritium or that introduced by weapons testing; in order cases large amounts of tritium are deliberately added. Investigations include the distribution of groundwater in oil fields; the tracing of springs, rivers, and lakes; water seepage and loss from reservoirs; and the movement of glaciers.

Tritium has also been used as a tracer for hydrogen in the study of chemical reactions. The most widespread use of tritium has probably been in biological research, where it has been used both as a hydrogen tracer and as a molecular label in studies of metabolism, biosynthesis, and cytology. In particular, tritiated thymidine and other nucleotides and nucleosides have been extensively used in studies of the formation of DNA and RNA. *See* HEAVY WATER; NUCLEAR FUSION; RADIOACTIVE TRACER; RADIOCHEMISTRY; TRITON. Louis Kaplan

Bibliography. E. Buncel, *Tritium in Organic Chemistry*, vol. 4, 1978; F. Mannone, *Safety in Tritium Handling Technology*, 1993; G. Vasaru, *Tritium Isotope Separation*, 1993.

Triton

The nucleus of $_1H^3$ (tritium); it is the only known radioactive nuclide belonging to hydrogen. The triton is produced in nuclear reactors by neutron absorption in deuterium ($_1H^2 + {}_0n^1 \rightarrow {}_1H^3 + \gamma$), and decays by β^- emission to $_2He^3$ with a half-life of 12.4 years. The spin of the triton is $^1/_2$, its magnetic moment is 2.9788 nuclear magnetons, and its mass is 3.01700 atomic mass units. Much of the interest in producing $_1H^3$ arises from the fact that the fusion reaction $_1H^3 + {}_1H^1 \rightarrow {}_2He^4$ releases about 20 MeV of energy. Tritons are also used as projectiles in nuclear bombardment experiments. *See* NUCLEAR REACTION; TRITIUM. Henry E. Duckworth

Trochodendrales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the Eudicotyledon. The order consists of two families, the Trochodendraceae and Tetracentraceae, each with only a single species. The two species are often united into the single family Trochodendraceae. The group is of considerable botanical and evolutionary interest, as it is situated near the base of the advanced Eudicotyledon and links this larger group with more primitive flowering plants. Trochodendrales comprise trees of eastern and southeastern Asia with primitive (without vessels) wood. The flowers have a much reduced perianth with scarcely sealed carpels that are only slightly fused to each other. *See* PLANT KINGDOM; PLANT TAXONOMY. K. J. Sytsma

Trogoniformes

A small order of birds that contains only the family Trogonidae. Thirty-seven species are found throughout the tropics; two species reach the southern border of the United States. The trogons and quetzals are jay-sized birds with large heads, and tails that vary from medium length and squared to very elongated and tapered. The dorsal plumage of trogons and quetzals is predominantly metallic green, with blue, violet, red, black, or gray in a few. The ventral feathers are bright red, yellow, or orange. Despite their vivid coloration, the birds are inconspicuous when sitting quietly in the forest. Quetzals possess long tapered tails of upper covert feathers, not tail feathers. Trogon plumage is soft with lax feathers. Sexes are dissimilar in appearance, with the males being more brightly colored. The head is large and rounded, and the bill is small and weak. Legs are short and feet are weak, with the toes arranged in a heterodactyl fashion, with the first and second toes reverted, opposing the third and fourth toes. Flight is rapid, undulating, and brief; trogons rarely walk. The diet consists of fruit and small invertebrates, as well as insects caught in flight as the bird darts out from a perch.

Trogons are nonmigratory, arboreal, and sedentary, and they can remain on a perch for hours. The monogamous pairs nest in solitude in a hollow tree or termite nest. After the eggs have been incubated by both parents, the naked hatchlings remain in the nest and are cared for by both parents.

A few fossil trogons are known from the Oligocene and Miocene of Europe. The Eocene to Oligocene fossils have been assigned to the family Archaeotrogonidae (*Archaeotrogon*) of enigmatic affinities. They may be ancestral to the trogons, but they may also belong to a totally different group of birds.

The relationship of the trogons to other birds is still unresolved. Some authors argue for a close relationship to the Coraciiformes, but the evidence is inconclusive. Also difficult to explain is their current pantropical distribution. They may have dispersed over the Eocene land connection across the North Atlantic between Europe and eastern North America during a warmer climatic period. *See* CORACIIFORMES.

The quetzals are the most spectacular members of this brightly colored group of birds. The resplendent quetzal (*Pharomachrus mocinno*) of Central America is the national bird of Guatemala. *See* AVES.

Walter J. Bock

Bibliography. N. J. Collar, Order Trogoniformes, in J. del Hoyo et al. (eds.), *Handbook of the Birds of the World*, vol. 6, pp. 80–127, Lynx Edicions, 2001; P. A. Johnsgard, *Trogons and Quetzals of the World*, Smithsonian, 2000.

Trojan asteroids

Asteroids located near the equilateral lagrangian stability points of a Sun-planet system (see **illus.**). As shown by J. L. Lagrange in 1772, these are two of the five stable points in the circular, restricted, three-body system, the other three points being located along a line through the two most massive bodies in the system. In 1906 Max Wolf discovered an asteroid located near the lagrangian point preceding Jupiter in its orbit. Within a year, two more were found, one of which was located near the following lagrangian point. It was quickly decided to name these asteroids after participants in the Trojan War as given in Homer's *Iliad*. Hence the term "Trojan asteroid," originally referring to asteroids orbiting the Sun near one of Jupiter's equilateral lagrangian points. With the exception of (624) Hektor in the preceding "swarm" and (617) Patroclus in the following "swarm" (named before this convention was adopted), asteroids in Jupiter's preceding and following lagrangian swarms are named after Greek and Trojan participants, respectively.

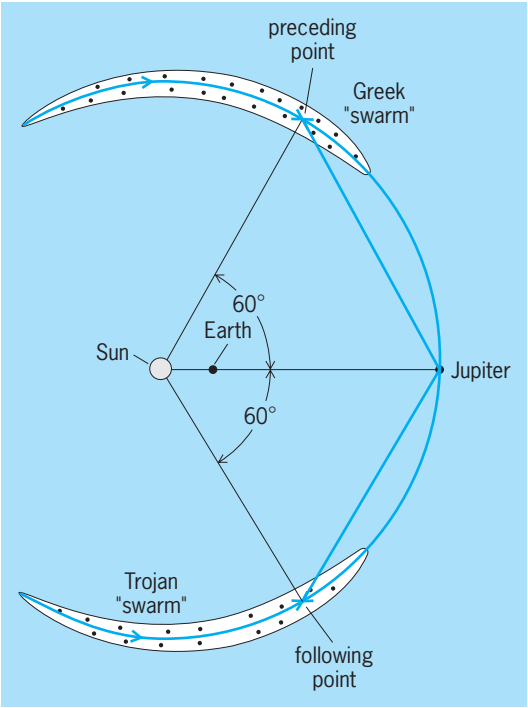

Lagrangian points and Trojan asteroids.

The term "Trojans" is sometimes used in a generic sense to refer to objects occupying the equilateral lagrangian points of other pairs of bodies. Small bodies have been found at the equilateral lagrangian points of Saturn's satellites Tethys and Dione: The satellites Calypso and Telesto orbit near the preceding and following points, respectively, of the Saturn-Tethys system, and the satellite Helene orbits near the preceding point of the Saturn-Dione system. Unsuccessful searches have been made for Trojans of the Earth, Saturn, and Uranus, as well as for the Earth-Moon system. Due to its eccentric orbit, and close proximity to Venus, it is considered unlikely that there are any stable regions for Mercury. It is believed that stable regions near the equilateral lagrangian points of Earth, Venus, Saturn, and Uranus exist but are considerably smaller than those of Jupiter. Indeed most of Jupiter's Trojans do not move in the plane of its orbit, but in orbits inclined by as much as 25° and at longitudes differing by up to 40° from the longitudes of the theoretical lagrangian points. *See* SATURN.

On June 20, 1990, D. H. Levy and H. E. Holt at Palomar Observatory discovered an asteroid, later named (5261) Eureka, occupying the following lagrangian point of the planet Mars. In 2001 the first Trojan of Neptune (2001 QR322) was discovered in the course of the Deep Ecliptic Survey, and as of September 2006 three additional Neptunian Trojans had been discovered. These Martian and Neptunian Trojans are the only confirmed non-Jupiter Trojans.

Trojan asteroids are classified using their osculating orbital elements, as opposed to their proper orbital elements. (Osculating orbital elements are the instantaneous orbital elements of the ellipse that the asteroid would follow if all bodies other than the Sun suddenly ceased to exist. Proper orbital elements are the orbital elements averaged over a long period of time according to a precisely specified procedure.) This classification technique works very well for Jupiter Trojans and for the four Neptunian Trojans discovered to date, but not so well for Martian Trojans. Thus, long-term integration of high-quality orbits will be necessary to eventually decide which of the six potential Martian Trojans that had been discovered as of December 2003 are actually Trojans. *See* ORBITAL MOTION.

Jupiter Trojans are dark objects reflecting only between 3 to 9% of the visual light they receive. The majority are compositionally similar to the most common type of outer main-belt asteroid, but some, perhaps as many as one-third, have no known analog among the asteroids or meteorites. *See* ASTEROID; METEORITE.

In September 2006 there were 2059 known Jupiter Trojans. Of these, 1139 were in the preceding swarm and 1920 in the following. Albedos (percent of visual light reflected) and diameters have been measured for 71 Jupiter Trojans; the mean albedo for objects in the two swarms is the same, the overall mean albedo being 5.5 ± 1.6%. Analysis of these data indicates that there are about 20 Trojan asteroids with diameters greater than 100 km (60 mi) and about 1750 with diameters exceeding 15 km (9 mi). Of these, 1000 ± 200 are in the preceding swarm and 750 ± 200 are in the following swarm. Approximately 20% of short-period comets are thought to originate from collisions among the Trojan asteroids. *See* ALBEDO; COMET. Edward F. Tedesco

Bibliography. B. Bottke et al. (eds.), *Asteroids III*, 2002; F. Marzari, P. Tricarico, and H. Scholl, Saturn Trojans: Stability regions in the phase space, *Astrophys. J.* 579:905–913, 2002; S. Tabachnik and N. W. Evans, Cartography for Martian Trojans. *Astrophys. J. Lett.*, 517:L63–L66, 1999; S. A. Tabachnik and N.W. Evans, Asteroids in the inner solar system—I. Existence, *Mon. Not. Roy. Astron. Soc.*, 319:63–79, (2000).

Trombidiformes

A suborder of the Acarina (also known as Prostigmata) commonly called the trombidiform mites, more closely related to the Sarcoptiformes than to the other suborders. They are usually distinguished by presence of a respiratory system opening at or near the base of the chelicerae. Other distinguishing characters are to be found in the tarsi, chelicerae, and genitalia (see **illus.**). These, although variable within the suborder, are distinct from those in all other groups. The Trombidiformes are probably the most heterogeneous group of mites, both morphologically and ecologically, varying from baglike forms with degenerate legs to the highly evolved, fully developed, parasitic forms. There are wormlike forms found in pores of their hosts and flattened types found under scales of lizards; some are parasites in the respiratory tracts of birds, others are free-living predators of other arthropods, and some are plant feeders. They are also variable in their life histories. Some are held within the brood sac of the mother until the siblings, or offspring of the same parents, have had an opportunity to copulate, while others hatch as larvae and pass through a series of molts before becoming sexually mature. Economically, this group contains two families of plant-feeding mites of great importance to agriculture: the Tetranychidae (spider mites) and the Eriophyidae (bud mites or gall mites). The Tarsonemidae, Eupodidae, and Tenuipalpidae are of less importance. The Tetranychidae cause damage by feeding on the leaves and weakening the tree, thus decreasing fruit production and even causing complete defoliation at times. Humans, through commerce, have disseminated some of the most important species throughout the world. The Eriophyidae, by feeding, cause weakening of trees and distortion of fruit and tree growth. A few are vectors of virus diseases of plants. In the United States one transmits streak-mosaic virus of wheat in the Middle West; another, peach mosaic virus in the Far West.

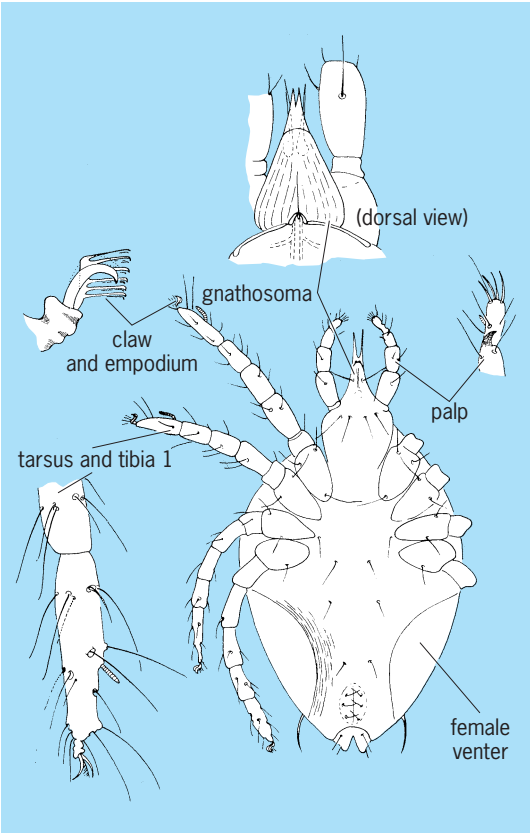

Trombidiform mite. (*Institute of Acarology, University of Maryland*)

While some of the Tarsonemini feed on plants, others feed on insects. Two species are of particular importance. The hay itch mite, *Pyemotes ventricosus*, normally lives on insect larvae, but when these are destroyed in the process of harvesting grain, the mites may cause a serious dermatitis on humans. The second species, *Acarapis woodi*, causes a disease of honeybees in many parts of the world, which so far has not been reported in the United States.

Medically, the Trombiculidae (chiggers, or red bugs) are important because the larval forms, which are parasites of vertebrates, can cause intense irritation to their hosts by their feeding. More seriously, some transmit a rickettsial disease, scrub typhus, to humans in the Far East and South Pacific regions. The nymphs and adults of the trombiculids are free-living and prey upon eggs of other arthropods. A few families related to the Trombiculidae are predators as nymphs and adults, but the larvae are parasites of arthropods.

The other families are either free-living and usually considered to be predacious, or are minor parasites of birds, reptiles, and mammals. Of these, the most important are the Demodicidae, or pore mites. *Demodex folliculorum* is frequently present in pores of humans, but it is seldom noticed and is of little

medical importance. Species on domesticated animals produce more apparent symptoms, and *Demodex canis* is at times fatal for dogs. A large and interesting group are the colorful Hydrachnellae which, with few exceptions, are found living in fresh water. A few are parasites of fresh-water mussels, and the larvae of others parasitize aquatic insects. Adults are predacious on small aquatic animals. *See* PLANT PATHOLOGY; PLANT VIRUSES AND VIROIDS.

Edward W. Baker

Trophic ecology

The study of the structure of feeding relationships among organisms in an ecosystem. Researchers focus on the interplay between feeding relationships and ecosystem attributes such as nutrient cycling, physical disturbance, or the rate of tissue production by plants and the accrual of detritus (dead organic material). Feeding or trophic relationships can be represented as a food web or as a food chain. Food webs depict trophic links between all species sampled in a habitat, whereas food chains simplify this complexity into linear arrays of interactions among trophic levels. Thus, trophic levels (for example, plants, herbivores, detritivores, and carnivores) are amalgamations of species that have similar feeding habits. (However, not all species consume prey on a single trophic level. Omnivores are species that feed on more than one trophic level.) *See* ECOLOGY; ECOSYSTEM; FOOD WEB.

The three fundamental questions in the field of trophic ecology are: (1) What is the relationship between the length of food chains and plant biomass (the total amount of plants at the bottom of the food chain)? (2) How do resource supply to producers (plants) and resource demand by predators determine the relative abundance of organisms at each trophic level in a food chain? (3) How long are real food chains, and what factors limit food chain length?

Effect of food chain length on plant biomass. A central theory in ecology is that "the world is green" because carnivores prevent herbivores from grazing green plant biomass to very low levels. Trophic structure (the number of trophic levels) determines trophic dynamics (as measured by the impact of herbivores on the abundance of plants). Indirect control of plant biomass by a top predator is called a trophic cascade. Cascades have been demonstrated to varying degrees in a wide variety of systems, including lakes, streams, subtidal kelp forests, coastal shrub habitats, old fields, grassland savannas, arctic tundra, shrublands, and below-ground soil communities. In many of these systems, the removal of a top predator has been shown to precipitate dramatic reductions in the abundance (or biomass) of species at lower trophic levels. Food chain theory predicts a green world when food chains have odd numbers of trophic levels, but a barren world (plants suppressed by herbivores) in systems with even numbers of trophic levels. The reduction and subsequent return of sea otters in coastal ecosystems provides a lucid example of cascading trophic effects in marine food chains and alternation between plant-dominated and plant-depleted subtidal habitats.

Sea otters once were abundant in coastal regions from northern Japan to central Baja California, Mexico, but were reduced to a number of widely scattered remnant populations by hunting. The reduction of sea otters likely resulted in dramatic changes in prey population dynamics in ecosystems previously occupied by sea otters. For example, in rocky subtidal habitats the abundance of sea urchins, a preferred prey item of otters, was much higher in habitats with depleted otter populations. In otter-free habitats (two trophic levels), urchin abundance was high enough to overgraze large benthic algae, including kelps. By contrast, in areas supporting remnant populations of otters (three trophic levels), urchin populations were limited to small individuals and

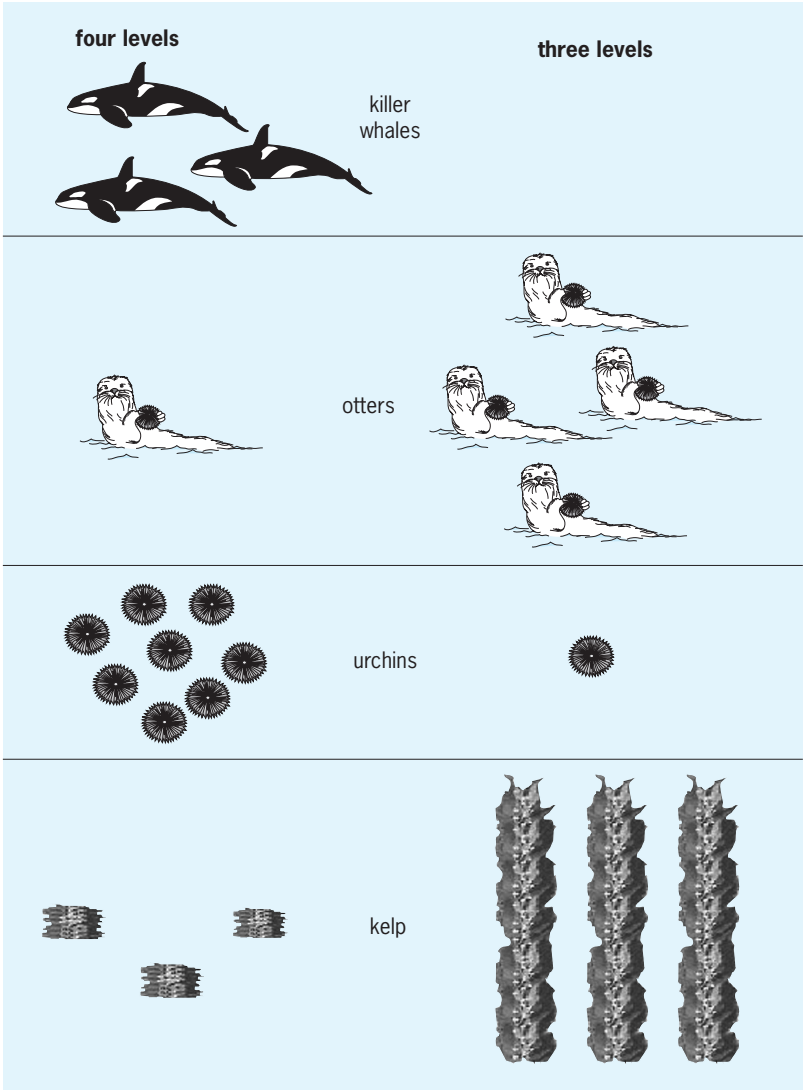

Food chain dynamics in subtidal kelp forests. Four-level system: in open coastal areas, killer whales decimate sea otter populations, releasing urchins which are capable of regulating macroalgae. Three-level system: in the absence of killer whales, otter populations increase and prevent urchins from razing kelp forests.

were often restricted to cryptic habitats. As a result, kelp biomass was much higher in these areas. Thus, odd and even food chains lead to green and barren subtidal worlds, respectively (see **illus.**). The result of the presence of otters in nearshore habitats is an increased abundance of kelps and other macroalgae which, in turn, provide habitat and food for a number of associated species. Otters thus play a critical role in structuring subtidal communities. Otters and other species whose effects on the abundance of other species in a food chain or food web are large compared with their relative abundance in the food web are called keystone species.

Recently, sea otters have declined in abundance in the Aleutian Islands for the first time since the regulation of commercial harvest (International Fur Seal Treaty). A convincing hypothesis for the cause of this decline is a shift in the foraging behavior of killer whales from the once abundant baleen whales and pinnipeds (seals, sea lions, and walruses) to sea otters. A comparison of otter abundance in coastal areas of Adak Island, Alaska, where otters were exposed to and protected from killer whale predation, revealed strong effects of these top predators on near-shore ecosystems. Comparison of urchin abundance, grazing rates, and kelp biomass during time periods before and after observed increases in killer whale predation on otters has revealed cascading effects of killer whales in this system. Prior to increased predation by killer whales, coastal subtidal habitats were characterized by low urchin and high kelp abundance. By contrast, these same habitats had much higher urchin densities and more barren substrates immediately following shifts in killer whale foraging and concomitant otter declines. The addition of killer whales (four trophic levels) led to a shift from green to barren worlds as a result of the otters' diminished control of urchin grazing. *See* MARINE ECOLOGY.

Supply and demand in food chain dynamics. Although predators often have strong indirect effects on plant biomass as a result of trophic cascades, both predation (a top-down force) and resource supply to producers (a bottom-up force) play strong roles in the regulation of plant biomass. The supply of inorganic nutrients (such as nitrogen and phosphorus) at the bottom of a food chain is an important determinant of the rate at which the plant trophic level produces tissue (primary production, or productivity) and, in some cases, of the total biomass of this trophic level. However, the degree to which nutrient supply enhances plant biomass accrual depends on how many herbivores are present (which in turn depends on how many trophic levels there are in the system). The relative importance of top-down (demand) versus bottom-up (supply) forces is well illustrated by lake systems, in which the supply of phosphorus (bottom-up force) and the presence of piscivorous (fish-eating) fish (top-down force) have significant effects on the standing stock of phytoplankton, the plant trophic level in lake water columns.

In small lakes of the Canadian Shield of North America, phytoplankton production is strongly dependent on phosphorus. In a classic experiment, a single lake was divided in half with a plastic barrier, and phosphorus, nitrogen, and carbon were added to one side of the lake while just nitrogen and carbon were added to the other (control) side. Blooms of algae turned the water green on the side to which phosphorus was added but not on the control. Side. This result was instrumental in convincing local and national governments in Canada and the United States to regulate phosphate release in sewage to prevent noxious, oxygen-depleting blooms of algae in lakes and other sources of drinking water. *See* FRESHWATER ECOSYSTEM; LAKE; PHYTOPLANKTON; ZOOPLANKTON.

Herbivores may be capable of counteracting nutrient-driven algal blooms, especially in relatively nutrient-poor (oligotrophic) lakes. In these lakes, the effects of top predators, such as largemouth bass, on phytoplankton are analogous to those of killer whales on kelp, and the response of phytoplankton to nutrient loading depends on the number of fish trophic levels in the system. Small lakes have up to two distinct trophic levels of fish species—those that eat zooplankton (zooplanktivores) and those that eat other fish (piscivores). In oligotrophic lakes with only zooplanktivores (that is, which have only three trophic levels), these fish deplete zooplankton which would otherwise graze on phytoplankton, thereby allowing phytoplankton biomass to increase with nutrient (phosphorus) loading from the bottom of the food chain. By contrast, if piscivores such as largemouth bass are added to these same lakes (which then have four-trophic levels), they eat the zooplanktivores, zooplanktivore abundance declines, and zooplankton recover and graze the once green lake until it is barren. In these four-trophic-level systems, phosphorus additions may increase the productivity but not the biomass of the plant trophic level. Thus, top-down control of phytoplankton biomass by predators is possible in oligotrophic lakes, depending on food chain length.

In more fertile (mesotrophic) lakes, piscivores may control zooplanktivore abundance, but the zooplankton (herbivores) are still not capable of keeping pace with increasing phytoplankton production across gradients of increasing nutrient loading. Top-down control attenuates between the third (zooplanktivore) and second (herbivore) trophic levels, and nutrient loading has positive effects on plant productivity and biomass.

In lakes with moderate-to-high nutrient loading, high productivity by species at all trophic levels may preclude control by consumers at the top. In this case, increased nutrient levels lead to increases in the biomass of not only phytoplankton but zooplankton and fish as well. Thus, although plant productivity may increase fairly predictably with nutrient loading, the biomass of the plant trophic level depends both on nutrient supply and trophic structure. Top-down control of plant biomass by top predators appears to be more important in nutrient-poor systems.

As nutrient supply increases, the relative influence of bottom-up control on plant biomass is increasingly important. *See* BIOLOGICAL PRODUCTIVITY; BIOMASS.

Determinants of food chain length. There are three major hypotheses for what determines the length of food chains in anture, based on energy, resilence, and ecosystem size.

Productivity and efficiency. From an energetic perspective, food chain length is limited by two factors: the total rate of plant or bacterial production, and the efficiency at which members of each trophic level assimilate this energy as it moves up the food chain. Productivity determines the total energy supply for a system; however, not all of this energy is incorporated by successively higher trophic levels. Inefficiency in transfer, in either consumption or assimilation, reduces the fraction of total available energy propagated between each trophic level. Typical transfer efficiencies range 5–15% for herbivores and carnivores (for example, zooplankton and bass). In three transfers, the total energy base of the food chain is reduced by more than 99%, limiting the energy available to an additional trophic level. Thus, given fixed transfer transfer efficiencies, additional trophic levels may be added only with increases in productivity at the base of the food chain. Several large-scale syntheses (on lake and arctic tundra island systems) suggest that food chain length does increase with plant productivity. Nevertheless, long food chains (more than five trophic levels) are extremely rare in nature—rare enough to suggest to some ecologists that these food chains may be inherently incapable of persisting. *See* BIOLOGICAL PRODUCTIVITY; ECOLOGICAL ENERGETICS.

Resilience. The idea that the upper limit to food chain length may be determined by the inherent instability of long chains derives from theoretical studies of food chains. In these studies, theoretical ecosystems are constructed as coupled differential equations describing the population dynamics of single-species trophic levels. To analyze food chain stability, disturbance is introduced by changing (reducing or increasing) the abundance at one trophic level. Resilience, an ecological metric for stability, is measured as the inverse of the time required for all trophic levels to return to their previous abundance levels. Longer food chains have consistently longer return times (lower resilience) than short chains (four or less trophic levels), suggesting that longer chains should be rare in nature because return times may be longer than the recurrence interval of disturbance. Although results from model food chains are consistent with field tests in small aquatic systems, the empirical mechanisms behind decreased stability of longer food chains at larger spatial scales are not as clearly developed as the logic of the productivity hypothesis. Moreover, experiments in rivers suggest that disturbance in some cases may act to lengthen food chains, suggesting that the effects of disturbance on food chain length may vary between ecosystem types. *See* SYSTEMS ECOLOGY; THEORETICAL ECOLOGY.

Ecosystem size. One final factor that may set constraints on food chain length is ecosystem size. Size can be defined relatively easily in habitats with discrete boundaries (for example, lakes and oceanic islands). It has been hypothesized that ecosystem size and productivity would interact to determine food chain length. This theory holds that the coverage (in area or volume) of a given level of productivity would provide a more comprehensive measure of the energy supply at the base of food chains and, thus, their potential length. Food chains should increase in length with increasing "productive space" rather than with increasing productivity alone. However, the size of ecosystems alone may provide an equally robust prediction of food chain length. Both body size and home range size may increase with trophic position. Small habitats are simply not large enough to support the home range or provide ample habitat for larger carnivorous species and so may limit the length of food chains. In lake systems, for instance, ecosystem size alone predicts more variability in food chain length than either productivity or productive space. In these systems, large piscivores are often found only in deep waters—habitats found only in larger lakes.

Bioaccumulation of contaminants in top predators. One application of trophic ecology has been the recognition and prevention of bioaccumulation of pesticides and heavy metals in top predators. An example of bioaccumulation is the near-demise of birds of prey such as the bald eagle, osprey, and peregrine falcon in the United States as a result of the formerly unrestricted application of DDT and other organochlorine pesticides to crop fields, lakes, and ponds to kill insect pests (crop herbivores and mosquitoes).

Although DDT is not lethal to animals at higher trophic levels, it is not excreted by them. Thus, birds and fish that consume large quantities of insects or other herbivores exposed to DDT concentrate the pesticide in their tissues. Birds of prey, in turn, further concentrate DDT because they rely heavily on fish or other birds as prey species (two trophic transfers of DDT). Peregrine falcons may be especially vulnerable to DDT magnification because in some areas these birds consume fish-eating seabirds (three trophic transfers of DDT). Birds with high pesticide levels tend to lay eggs with abnormally thin shells that crack during incubation. As a result of impaired reproduction, peregrine falcon populations showed dramatic declines throughout the world. Recognition of this problem led to a ban on the use of DDT and other stable pesticides in many industrialized countries. DDT was outlawed in the United States in 1972; this ban allowed the eventual recovery of peregrine falcons to over 80 breeding pairs in the United States. Unfortunately, DDT is still used in many developing nations. Although the effect that this chemical may have on humans is still unknown, its effect on raptors suggests that humans could also be vulnerable.

Many fish species common in markets around the world are top predators and may be reservoirs for

other common toxins such as mercury. Interestingly, fish from ecosystems with long food chains may pass along more mercury to humans than those from food chains with fewer trophic levels and, thus, fewer trophic transfers of this toxin. *See* ECOLOGY, APPLIED; INSECTICIDE; PESTICIDE. John L. Sabo; Leah R. Gerber

Bibliography. S. R. Carpenter and J. F. Kitchell (eds.), *The Trophic Cascade in Lakes*, Cambridge University Press, 1993; R. Carson, *Silent Spring*, Fawcett Crest, New York, 1962; S. L. Pimm, *Food Webs*, Chapman & Hall, New York, 1982; G. A. Polis and K. O. Winemiller (eds.), *Food Webs: Integration of Patterns and Dynamics*, Chapman & Hall, New York, 1996.

Tropic of Cancer

The parallel of latitude about $23^1/_2{}^\circ$ (23.45°) north of the Equator. The importance of this line lies in the fact that its degree of angle from the Equator is the same as the inclination of the Earth's axis from the vertical to the plane of the ecliptic. Because of this inclination of the axis and the revolution of the Earth in its orbit, the vertical overhead rays of the Sun may progress as far north as $23^1/_2{}^\circ$. At no place north of the Tropic of Cancer will the Sun, at noon, be 90° overhead.

On June 21, the summer solstice (Northern Hemisphere), the Sun is vertical above the Tropic of Cancer. On this same day the Sun is 47° above the horizon at noon at the Arctic Circle, and at the Tropic of Capricorn, only 43° above the horizon. The Tropic of Cancer is the northern boundary of the equatorial zone called the tropics, which lies between the Tropic of Cancer and Tropic of Capricorn. *See* LATITUDE AND LONGITUDE; MATHEMATICAL GEOGRAPHY; SOLSTICE. Van H. English

Tropic of Capricorn

The parallel of latitude approximately $23^1/_2{}^\circ$ (23.45°) south of the Equator. It was named for the constellation Capricornus (the goat), for astronomical reasons which no longer prevail.

Because the Earth, in its revolution around the Sun, has its axis inclined $23^1/_2{}^\circ$ from the vertical to the plane of the ecliptic, the Tropic of Capricorn marks the southern limit of the zenithal position of the Sun. Thus, on December 22 (Southern Hemisphere summer, but northern winter solstice) the Sun, at noon, is 90° above the horizon. On this same day, at noon, the Sun is 47° above the horizon at the Antarctic Circle, $66^1/_2{}^\circ$ at the Equator, and 43° at the Tropic of Cancer. Sun rays will just reach the horizon tangentially at the Arctic Circle.

The Tropic of Capricorn is the southern boundary of the equatorial zone referred to as the tropics, which lies between the Tropic of Capricorn and the Tropic of Cancer. *See* LATITUDE AND LONGITUDE; MATHEMATICAL GEOGRAPHY; SOLSTICE. Van H. English

Tropical meteorology

The study of atmospheric structure and behavior in the areas astride the Equator, roughly between 30° north and south latitude. The weather and climate of the tropics involve phenomena such as trade winds, hurricanes, intertropical convergence zones, jet streams, monsoons, and the El Niño Southern Oscillation. More energy is received from the Sun over the tropical latitudes than is lost to outer space (infrared radiation). The reverse is true at higher latitudes, poleward of 30°. The excess energy from the tropics is transported by winds to the higher latitudes, largely by vertical circulations that span roughly 30° in latitudinal extent. These circulations are known as Hadley cells, after George Hadley who first drew attention to the phenomenon in 1735. This type of circulation is an important ingredient of the tropical general circulation.

For the most part, the oceanic tropics (the islands) experience very little change of day-to-day weather except when severe events occur. Tropical weather can be more adverse during the summer seasons of the respective hemispheres. The near equatorial belt between 5°S and 5°N is nearly always free from hurricanes and typhoons: the active belt lies outside this region over the tropics. The land areas experience considerable heating of the Earth's surface, and the summer-to-winter contrasts are somewhat larger there. For instance, the land areas of northern India experience air temperatures as high as 108°F (42°C) in the summer (near the Earth's surface), while in the winter season the temperatures remain 72°F (22°C) for many days. The diurnal range of temperature is also quite large over land areas on clear days during the summer (32 °F or 18 °C) as compared to winter (18 °F or 10 °C).

Weather observations. Vast areas of the tropics are oceanic, and there is a general lack of surface and upper-air observations of temperature, wind, pressure, and humidity over most regions. The network for weather observations over the land areas of Africa and South America are rather limited in comparison to the rest of the inhabited tropical areas. Tropical meteorology has greatly benefited from the space-based observations by meteorological satellites, which provide day and night images of cloud cover from visible and infrared sensors and can resolve as high as a few kilometers over the Earth's surface. There are two types of satellites—the polar orbiter and the geostationary; the latter are over the Equator at a height of roughly 22,500 mi (36,000 km). Five geostationary satellites are in orbit for complete global tropical coverage, providing images and tropical air motions at the lower tropical troposphere (about 0.6 mi or 1 km above the Earth's surface) and at the upper troposphere (about 7 mi or 12 km above the Earth's surface). The air motions are estimated from cloud tracking from adjacent frames (some 30 min apart) of cloud images at high resolution. With satellite observations, it is possible to monitor tropical weather phenomena on a daily basis around the global belt.

See SATELLITE METEOROLOGY; WEATHER OBSERVATIONS.

Trade winds. The steady northeast surface winds over the oceans of the Northern Hemisphere between 5° and 20°N and southeast winds over the corresponding latitudes of the southern oceans constitute the trade winds. Trade winds have intensities of around 5–10 knots (2.5–5 m/s). They are the equatorial branches of the anticyclonic circulation (known as the subtropical high pressure). The steadiness of wind direction is quite high in the trades. On the equatorial side of the strongest trade winds (known as the cyclonic shear side), tropical depressions usually form over warm oceans. The trade winds of the Southern Hemisphere (Atlantic, Pacific, and Southern Indian oceans) exhibit more disturbance activity. Trade winds are not present north of the Equator over the Indian Ocean; this region experiences southwesterly monsoonal flows. Near the Equator, the trade winds carry moisture to rain bands known as the intertropical convergence zones. Air-sea interaction in the Atlantic is strongly affected by the steady and nonsteady components of the trade wind systems of the two hemispheres. In addition, the air-sea interaction varies strongly during passage of disturbances, such as hurricanes. The Bowen ratio (sensible heat flux/latent heat flux) over this region is generally much less than 1; the dominant transfer is that of the latent heat. In strong trades, fluxes of latent heat of the order of 275 watts/m^2 are of common occurrence. If a hurricanelike disturbance is present, then on the mesoscale, fluxes of the order of 500 watts/m^2 are noted from aircraft and model-based estimates. The more important area is that of air-sea coupling during the passage of hurricanes. A cold wake with ocean temperature anomalies of the order of 3–5°F (2–3°C) is often generated by the action of strong winds. The lifting of the thermocline and the resulting upwelling of colder waters gives rise to these cold wakes. Passage of one hurricane often leaves such a wake that affects the intensity of a second hurricane that follows it. *See* WIND.

Hurricanes. Hurricanes are also known as typhoons in the west Pacific and tropical cyclones in the Indian Ocean and south Pacific. If the wind speed exceeds 65 knots (33 m/s) in a tropical storm, the storm is labeled a hurricane. A hurricane usually forms over the tropical oceans, north or south of 5° latitude from the Equator. The strongest winds are found in the lowest levels above the ocean; however, the winds weaken with height very slowly in the troposphere and, therefore, a sizable strength of the vortex can be seen all the way up to 6 mi (10 km). The vertical extent of a hurricane extends up to 7 mi (12 km); around that level the cyclonic circulation (clockwise in the Southern Hemisphere and counterclockwise in the Northern Hemisphere) tapers off with an anticyclonic circulation, with outflowing air that extends to an altitude of 9 mi (15 km). The inflowing air is found in the planetary boundary layer, that is, mostly in the lowest mile over the ocean, and spirals into the storm in a few marked channels known as spiral rain bands that converge toward an eye wall of the hurricane, whose diameter is of the order of 6–12 mi (10–20 km). Hurricanes form in regions of warm tropical oceans (with temperatures generally above 81°F or 27°C), where the variation of wind with height over the troposphere is small. They possess a warm thermal core and usually form from incipient tropical weaker cyclonic weather systems. The source of energy that drives these storms is the evaporation from the oceans as a result of wind action. Great efforts have been placed on intense real-time monitoring of hurricanes using aircraft, radar, and satellites. A mix of these platforms has provided dropwindsonde winds, dual Doppler radar–based hydrometeor structure, and satellite-based indirect estimates of precipitation using onboard radar and microwave instrumentation. Numerical models have, on the mesoscale, improved the short-range predictive capability of tracks, intensity, and precipitation amounts. Global models, on the medium-range time frame, up to 6 days into the future, have developed sophisticated data assimilation systems to absorb all these diverse data sets. The global models have also shown marked improvement of skills in the above areas. *See* HURRICANE; TROPOSPHERE.

Intertropical convergence zones. These zones are located usually between 5 and 10°N latitude. They are usually oriented west to east and contain cloud clusters with rainfall of the order of 1.2–2 in. (30–50 mm) per day. The trade winds of the two hemispheres supply moisture to this precipitating system. Embedded within this line of cloud clusters are westward propagating tropical waves that usually move at a speed of roughly 300 mi (500 km) per day. Occasionally these waves and associated cloud systems move away from the equatorial latitudes, amplify, and form tropical depressions and eventually hurricanes. The intertropical convergence zone tends to form over the warm oceans and move slowly poleward to almost 10° from the Equator during the warmest summer months over the respective hemisphere. *See* CLOUD PHYSICS; PRECIPITATION (METEOROLOGY).

Jet streams. A number of fast-moving air currents, known as jets, are important elements of the tropical general circulation. With speeds in excess of 30 knots (15 m/s), they are found over several regions of the troposphere. The most prominent of these, the tropical easterly jet, is found near 5–10°N latitude over the Asian regions and extends to the west African Atlantic coast. The core of maximum winds is located roughly 8 mi (14 km) above sea level. It forms on the equatorward side of the Asian and west African summer monsoon. During this monsoon period, there also occurs a somewhat weaker but important lower tropospheric easterly jet stream over west Africa. The core of this jet is located generally close to 14°N latitude at around 2.4 mi (4 km) above the Sahara desert. Many air motions form on the equatorward side of this jet. These waves, known as African waves, have east-west scales of the order of 1800 mi (3000 km); they move westward out of Africa into the Atlantic Ocean. Some of these waves amplify and form hurricanes over the tropical Atlantic Ocean.

Another important jet, occurring during the winter seasons of the two hemispheres, is known as the subtropical westerly jet stream. It is most prominent during December, January, and February over the Northern Hemisphere and is located at an altitude of about 7 mi (12 km) above the Earth's surface, near 27.5°N latitude. This jet stream encircles the globe in a basic three-wave pattern that is relatively stationary. Strong winds of the order of 70 knots (35 m/s) are found near the southeastern United States coast and over the north African Mediterranean coast. The third wave is located off southern Japan, where winds as strong as 50 knots (100 m/s) are of common occurrence. The subtropical jet of the Southern Hemisphere is strongest during July and August. Its latitude is close to 25°S, and its core is located at the tropopause level near 7.2 mi (12 km). *See* ATMOSPHERIC GENERAL CIRCULATION; JET STREAM.

Monsoons. Basically the entire landmass from the west coast of Africa to Asia and extending to the date line experiences a phenomenon known as the monsoon. Monsoon circulations are driven by differential heating between relatively cold oceans and relatively warm landmasses. The west African region near 5–10°N latitude, the Indian region between 10 and 20°N latitude, and the east Asian and Chinese region between 10 and 40°N latitude experience the northern summer monsoon. An annual cycle of monsoon reveals itself as a north-south oscillation of the precipitating systems. These systems are located near the Indonesia-Borneo-northern Australian regions during the northern winter months, and they move northward toward the eastern foothills of the Himalaya by late June. Rainfall amounts in the active monsoon can exceed 120 in. (3000 mm) per month. The heaviest monsoon rainfall occurs at the foothills of the mountains, where rainfall amounts have been known to be as large as 100 in. (2500 mm) per month. It is not uncommon for single rainfall events, during the passage of a monsoon disturbance, to bring in amounts of order of 8 in. (200 mm) per day. Warm oceans are the source of moisture carried by monsoonal lower tropospheric flows. In the tropics the troposphere extends from the surface to about 8.4 mi (14 km). Over the northern summer monsoon this is a southwesterly flow, and in the northern winter it is generally a northeasterly flow. *See* MONSOON METEOROLOGY.

El Niño Southern Oscillation. Every 2–6 years the eastern equatorial Pacific Ocean experiences a rise in sea surface temperature of about 5–9°F (3–5°C). This phenomenon is known as El Niño, which is part of a larger cycle referred to as the El Niño Southern Oscillation (ENSO). The other extreme in the cycle is referred to as La Niña. El Niño has been known to affect global-scale weather.

A complete understanding of El Niño begins with how the ocean and the atmosphere operate in the equatorial Pacific region under normal conditions. Fluctuations in the atmosphere pressure occur between the west Pacific and Indonesia, and constitute the Southern Oscillation. The state of the pressure system of the Southern Oscillation is characterized by the Southern Oscillation Index, which is a measure of the pressure difference anomaly between Papeete, Tahiti, and Darwin, Australia. The difference between high pressure in the west Pacific and low pressure in Indonesia helps to drive the easterly trade winds along the Equator. Strong trade winds pile up the water in the west and actually cause the sea level to be an average of 16 in. (40 cm) higher in the west Pacific than in the east Pacific. In a simplified model, the Pacific Ocean can be considered as having two layers separated by a thermocline. The upper layer is warm and well mixed, while the lower layer is cold and well stratified. In the west Pacific basin the upper mixed layer is thicker than in the east Pacific. The difference in the thickness of the upper mixed layer leads to differences in the sea surface temperatures across the ocean. The water in the east Pacific is generally colder because of upwelling. Upwelling is a process caused by wind stress on the eastern boundary of the Pacific Ocean. The easterly winds tend to pull the water away from the coast. This water is replaced by cooler water from beneath the thermocline. Hence, the coastal waters of Ecuador and Peru are generally colder than the west Pacific. *See* WIND STRESS.

El Niño, considered the warm phase of the Southern Oscillation, is marked by a few prevalent conditions. First the Southern Oscillation Index decreases, inducing a westerly wind anomaly, that is, a weakening of the easterly trade winds. This forcing by the atmosphere causes a response by the ocean in the form of internal Kelvin waves. In the west Pacific Ocean a Kelvin wave is generated that propagates eastward. It takes 2–3 months for this wave to reach the coastline of South America. The Kelvin wave increases the thickness of the upper mixed layer in the central and eastern equatorial Pacific. The upwelling process now takes place in the warm mixed layer instead of bringing up cold water from below the thermocline. This causes the sea surface temperature to rise in the central and east Pacific. Tropical convective clouds and rain form over the anomalously warm oceanic region. Large areas of rising air are found in these regions of heavy rain, while sinking air is found surrounding the regions. As this warm pool of water moves eastward, it brings convection. These changes in the ocean and atmosphere in the equatorial Pacific have been shown to affect the global climate. *See* CLIMATE HISTORY; DROUGHT; EL NIÑO; MARITIME METEOROLOGY; METEOROLOGY; UPWELLING.

Biennial oscillation. The quasi-biennial oscillation (QBO) is a reversal of the west-to-east winds in the equatorial latitudes in the stratosphere. This reversal has a time scale of roughly 26 months. Vertical propagation of energy (and downward phase propagation) is a characteristic feature of this oscillation. A slow downward phase propagation with alternating westerlies and easterlies characterizes this system. It has its strongest amplitude near the 19-mi (30-km) level. It is generally seen clearly to about 9 mi (15 km). A tropospheric biennial oscillation (TBO) on the time scale of roughly 2 years has also drawn considerable

interest. It has been noted in the Asian monsoon longitude and the west Pacific Ocean. This time scale has been noted in the air-sea fluxes, the monsoon precipitation, and the elevation of pressure surfaces. Both the QBO and TBO are being studied intensely, and possible relationships among them are being investigated.

Weather prediction. Barotropic forecasts based on the principle of conservation of absolute vorticity was the centerpiece for tropical modeling in the early 1960s. Since then the progress has been steady. Multilevel complete physical/dynamical models at very high resolution are currently being used for prediction. Much thought has been given to the issue of how the effect of cumulus convection has to be included within the large-scale models. This is the area of cumulus parametrization. The coverage of data is a major issue. Large amount of satellite-based data sets provide cloud-tracked winds, water vapor–tracked winds, precipitation rates, and sea surface temperatures. Assimilation of these data sets has provided improvement in hurricane track and intensity forecasts and tropical precipitation forecasts. One area that has contributed to forecast skill improvement is ensemble forecasts. Here a large number of forecasts from the same start date provide a robust ensemble. Combining statistics and the ensemble of forecasts, it has been possible to make some major improvements in tropical forecasts. T. N. Krishnamurti

Bibliography. S. Ackerman and J. A. Knox, *Meteorology: Understanding the Atmosphere*, 2d ed., 2006; G. D. Atkinson, *Forecasters' Guide to Tropical Meteorology*, 2002; G. R. McGregor and S. Nieuwolt, *Tropical Climatology: An Introduction to the Climates of the Low Latitudes*, 2d ed., 1998; B. Wang, *The Asian Monsoon*, 2006.

Tropopause

The boundary between the troposphere and the stratosphere in the atmosphere. The tropopause is broadly defined as the lowest level above which the lapse rate (decrease) of temperature with height becomes less than 6°F/mi (2°C/km). In low latitudes the tropical tropopause is at a height of 9–10.5 mi or 15–17 km (∼−135°F or 180 K), and the polar tropopause between tropics and poles is at about 6 mi or 10 km (∼−63°F or 220 K). There is a well-marked "tropopause gap" or break where the tropical and polar tropopause overlap at 30–40° latitude. The break is in the region of the subtropical jet stream and is of major importance for the transfer of air and tracers (humidity, ozone, radioactivity) between stratosphere and troposphere. Tropopause breaks also occur in the neighborhood of polar jet streams. The height of the tropopause varies seasonally and also daily with the weather systems, being higher and colder over anticyclones than over depressions. The detailed vertical temperature structure is often complex, showing multiple or laminated tropopauses, and it is often difficult to decide on the precise height of the tropopause, particularly in winter at high latitudes. *See* AIR TEMPERATURE; ATMOSPHERE; RADIO-WAVE PROPAGATION; STRATOSPHERE; TROPOSPHERE. R. J. Murgatroyd

Troposphere

The lowest major layer of the atmosphere. The troposphere extends from the Earth's surface to a height of 6–10 mi (10–16 km), the base of the stratosphere. It contains about four-fifths of the mass of the whole atmosphere. *See* ATMOSPHERE.

On the average, the temperature decreases steadily with height throughout this layer, with a lapse rate of about 18°F/mi (6.5°C/km), although shallow inversions (temperature increases with height) and greater lapse rates occur, particularly in the boundary layer near the Earth's surface. Appreciable water-vapor contents and clouds are almost entirely confined to the troposphere. Hence it is the seat of all important weather processes and the region where interchange by evaporation and precipitation (rain, snow, and so forth) of water substance between the surface and the atmosphere takes place. *See* ATMOSPHERIC GENERAL CIRCULATION; CLIMATOLOGY; CLOUD PHYSICS; METEOROLOGY; WEATHER. R. J. Murgatroyd

Tropospheric scatter

A term applied to propagation of radio waves caused by irregularities in the refractive index of air. The phenomenon is predominant in the lower atmosphere; little or no scattering of importance occurs above the troposphere. Tropospheric scatter propagation provides very useful communication services but also causes harmful interference. For example, it limits the geographic separation required for frequency assignments to services such as television and frequency-modulation broadcasting, very high-frequency omnidirectional ranges, and microwave relays. It is used extensively throughout most of the world for long-distance point-to-point services, particularly where high information capacity and high reliability are required. Typical tropospheric scatter relay facilities (**Fig. 1**) are commonly 200–300 mi (320–480 km) apart. Some single hops in excess of 500 mi (800 km) are in regular use. High-capacity circuits carry 200–300 voice circuits simultaneously. *See* TROPOSPHERE.

Theory. To scatter is to spread at random over a surface or through a space or substance. Tropospheric scattering that tends to be coherent is more properly called forward scatter, reflection, refraction, focusing, diffraction by atmospheric inhomogeneities, or all of these. In the absence of the atmosphere, only diffraction by the earth would support propagation to distances well beyond the horizon. Atmospheric refractive index variations, which account for the remaining propagation mechanisms, are the result of variations in temperature, pressure, and gaseous constituents, the main variable being water vapor. The approximate formula for refractivity N, expressing changes in parts per million of the radio refractive

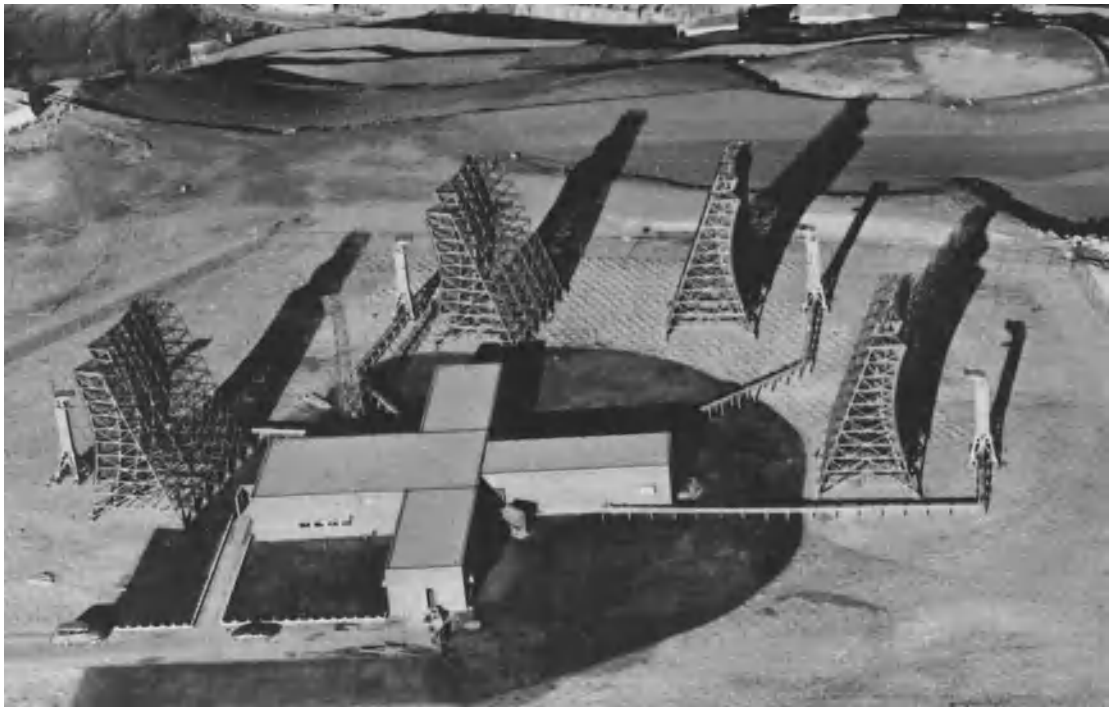

Fig. 1. Tropospheric scatter relay facility operated by the U.S. Air Force in Spain. High-powered ultra-high-frequency transmitters, sensitive receivers, and large parabolic antennas operating with frequency and space diversity make it possible to transmit many voice and teletype circuits simultaneously far beyond the horizon. (*Page Engineers, Inc.*)

index of air compared to that for a vacuum, is given by Eq. (1), where T is temperature in kelvins, P is

$$N = \frac{77.6}{T}\left(P + \frac{4810e}{T}\right) \tag{1}$$

total pressure in millibars, and e is partial water-vapor pressure in millibars.

Scattering from refractive index discontinuities, reflection from layers, and ducting from steep vertical gradients of refractive index account for nearly all of

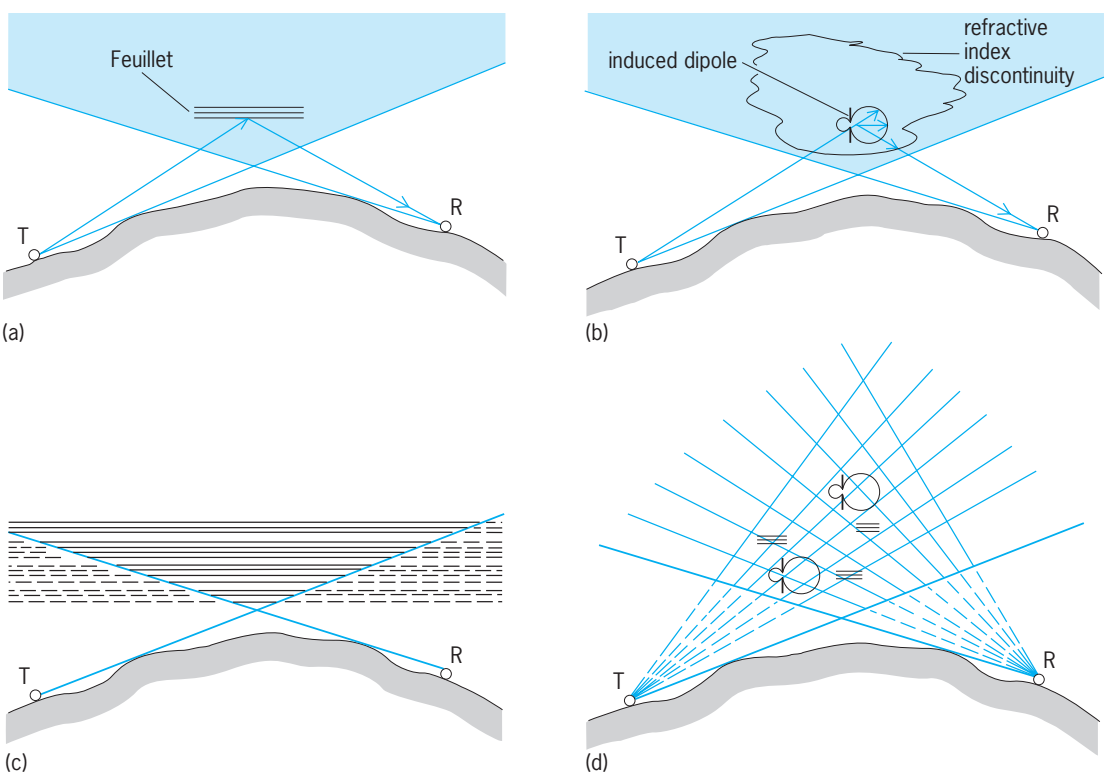

Fig. 2. Geometry for tropospheric forward scatter transmission from transmitter T to receiver R. (*a*) Feuillet. (*b*) Dipole induced in tropospheric layers. (*c*) Continuous scattering layers. (*d*) Limited scattering layers.

the radio energy at large distances beyond the horizon. These long-distance fields change with time. Rapid fading is caused by multipath components and is termed phase interference fading. The slower fading of 5-min or hourly medians from hour to hour and day to day throughout the year is called power fading and is caused by slowly changing atmospheric characteristics and by changes in the relative dominance of various propagation mechanisms.

The theory of tropospheric scattering is based on a combination of several models: The mechanism indicated by **Fig. 2***a* involves a tropospheric layer, or feuillet, which has a sufficiently abrupt change in refractive index, usually associated with fair weather conditions, to reflect a substantial amount of radio energy at the grazing angles and frequencies of interest. Layers can be either continuous as in Fig. 2*c*, or limited as in Fig. 2*d*. The most nearly specular reflections from atmospheric layers are usually observed between 30 and 200 MHz. At higher frequencies, where focusing, defocusing, and ducting are common and where extensive layers are not sufficiently abrupt or sufficiently numerous to provide strong reflections, a number of small, randomly oriented surfaces come into play. The radio wave is scattered forward by all the scattering subvolumes visible to both antennas, as indicated in Fig. 2*c* and *d*.

Data. Available long-term median forward scatter radio transmission loss data usually show attenuation in power inversely proportional to the wavelength cubed (λ^3). Long-term measurements on two or more frequencies rarely show power attenuation ratios outside the range λ^2 to λ^4. The λ^2 dependence is also characteristic of free-space propagation and of omnidirectional scattering by precipitation, and reflects the λ^2 dependence of the effective absorbing area of a receiving antenna. As a first approximation, Eq. (2) gives a simple formula for the long-term

$$L_{bm} = 46 + 30 \log_{10} \text{frequency (MHz)} + 20 \log_{10} \text{distance (miles)} + 0.1 \text{ distance (miles)} \quad (2)$$

median basic transmission loss L_{bm} for tropospheric scatter paths. L_{bm} is defined as the ratio, in decibels, of power delivered to an isotropic transmitting antenna to power from an isotropic receiving antenna, both antennas being lossless. *See* RADIO-WAVE PROPAGATION.

Robert S. Kirby

Bibliography. M. P. Hall, *Effects of the Troposphere on Radio Communication*, 1980; S. Shibuya, *A Basic Atlas of Radio-Wave Propagation*, 1986; M. Valkenburg, *Reference Data for Engineers: Radio, Electronic, Computer, and Communications*, 8th ed., 1996.

Truck

A motor vehicle ("lorry" in British English) carrying its load on its own wheels and primarily designed for the transportation of goods or cargo.

A truck is similar to a passenger car in many basic aspects, but truck construction is usually heavier throughout with strengthened chassis and suspension, and lower transmission and drive-axle ratios to cope with hilly terrain. Other common truck characteristics include cargo-carrying features such as rear doors or a tailgate, and a flat floor. However, there are many different kinds of trucks, often specially designed with unique features for performing a particular job, including catering trucks, cement trucks, dump trucks, fire trucks, flat-bed trucks, pickup trucks, refrigerated trucks, tank trucks, and walk-in van-bodied trucks. *See* AUTOMOBILE; BUS.

Types. A truck is rated by its gross vehicle weight (gvw), which is the combined weight of the vehicle and load. Trucks are classified as light-, medium-, or heavy-duty according to gross vehicle weight as follows:

Light-duty trucks
- Class 1: 0–6000 lb (0-2700 kg)
- Class 2: 6001-10,000 lb (2700-4500 kg)
- Class 3: 10,001-14,000 lb (4500-6300 kg)

Medium-duty trucks
- Class 4: 14,001-16,000 lb (6300-7200 kg)
- Class 5: 16,001-19,500 lb (7200-8775 kg)
- Class 6: 19,501-26,000 lb (8775-11,700 kg)

Heavy-duty trucks
- Class 7: 26,001-33,000 lb (11,700-14,850 kg)
- Class 8: 33,001 lb up (14,850 kg up)

Although a variety of models and designs are available in each category, there are two basic types of vehicles, the straight truck and the truck tractor. The straight truck has the engine and body mounted on the same chassis. The chassis includes the engine, frame, and other essential structural and mechanical parts, but not the body. The body is the structure or fixture especially provided to contain or support the goods or cargo to be transported.

The truck tractor is essentially a power unit that is the control and pulling vehicle for truck trailers such as full trailers or semitrailers. A full trailer has a front axle and one or more rear axles, and is constructed so that all its own weight and that of its load rests on its own wheels. A semitrailer has one or more axles at the rear, and is constructed so that the front end and a substantial part of its own weight and that of its load rests upon another vehicle. A retractable mechanism mounted on the front end of the semitrailer is lowered to support it when the pulling vehicle is disconnected. A full trailer may be drawn by a truck or behind a semitrailer.

Truck tractor. A truck tractor is a vehicle of short wheelbase for hauling semitrailers. It carries a swiveling mount, known as the fifth wheel, above the rear axle to support the front end of the semitrailer. If the tractor has two axles, the drive is through the rear axle. However, a three-axle tractor (one front and two rear) may drive through only one rear axle with one trailing, or through both rear axles.

The tractor-semitrailer combination permits the

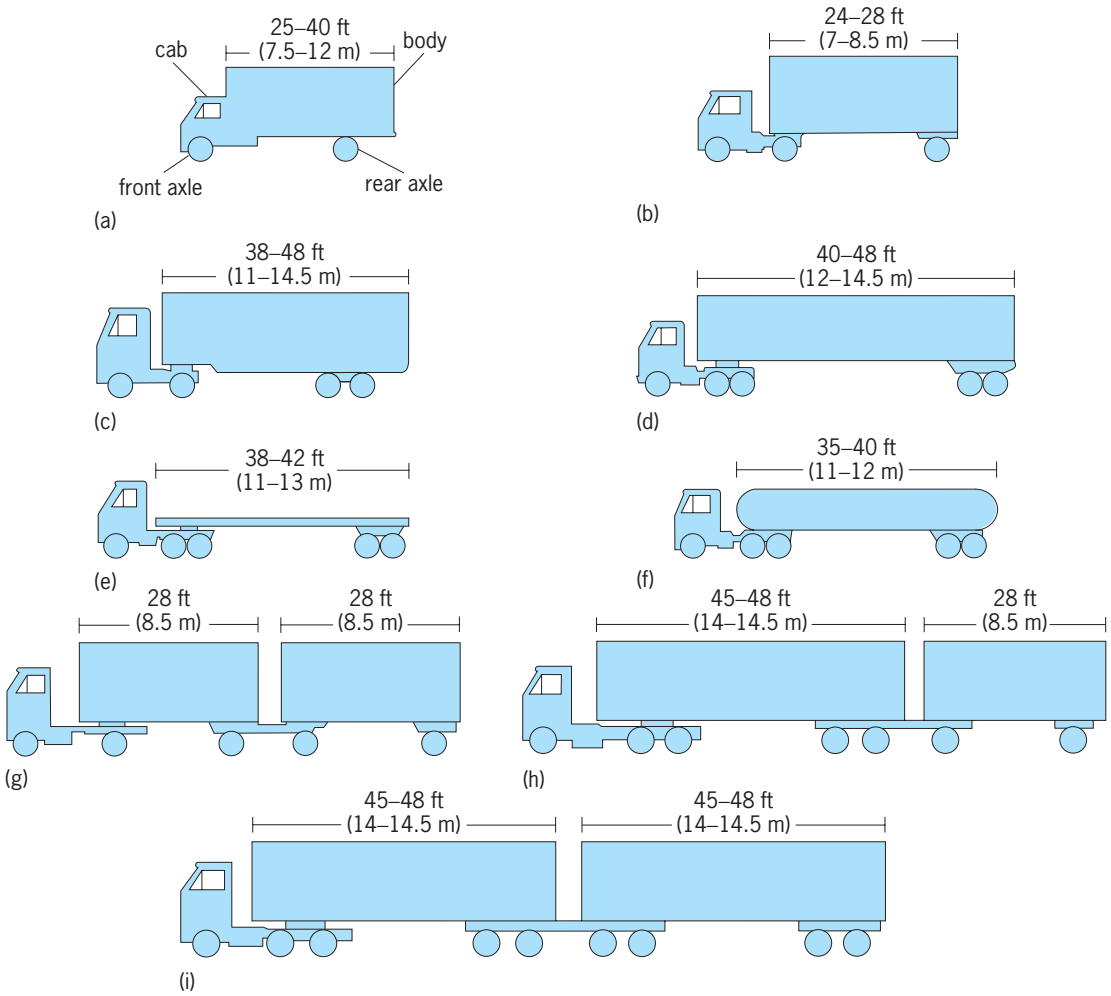

Fig. 1. Various types of truck, truck tractor, and trailer combinations. Lengths shown are typical. Other lengths are possible depending on the carrier's needs and state laws. For each type, the designation in brackets (see below) identifies the various combinations by axles: the first digit refers to the number of axles in the power unit; a second digit refers to a full trailer, while a second digit with prefix S refers to a semitrailer. For example, 3-S2-2 identifies a three-axle truck trailer pulling a two-axle semitrailer and a two-axle full trailer. (*a*) Straight truck [2]. (*b*) Three-axle tractor semitrailer [2-S1]. (*c*) Four-axle tractor semitrailer [2-S2]. (*d*) Five-axle tractor semitrailer [3-S2]. (*e*) Five-axle tractor flatbed trailer [3-S2]. (*f*) Five-axle tractor tank trailer [3-S2]. (*g*) Twin trailer or doubles [2-S1-2]. (*h*) Rocky Mountain doubles, operated only in certain states [3-S2-2]. (*i*) Turnpike doubles, operated only in certain states [3-S2-4].

use of longer bodies with greater carrying capacity and better maneuverability than is possible with a straight truck. The forward positioning of the cab, the short wheelbase of the tractor, and the multiplicity of axles provide maximum payloads and operating economy in the face of restriction on overall length imposed by some states, and regulations limiting the weight carried on a single axle. **Figure 1** shows various types of truck and of truck tractor and trailer combinations.

The standard Class 8 tractor-semitrailer combination, which has five axles and is often called an 18-wheeler, generally is limited to a gross combination weight (gcw) of 80,000 lb (36,300 kg) for operation on U.S. Interstate highways.

Cab and body. The cab is the part of the truck or tractor that encloses the driver and vehicle operating controls. It may be an integral part of the body, as in a van; or it may be a separate compartment alongside the engine, behind the engine, or over the engine. With the cab-over-engine design, the cab may be in fixed position or it may tilt forward for access to the engine. A cab with an interior or adjacent sleeping space is known as a sleeper or sleeper cab. The truck cab also serves many drivers as a mobile office. Its electronics may include a wireless computer, multifunction printer and fax machine, and global positioning system (GPS), which tracks vehicle location for security, navigation, and recommended routing that will be the most time-, cost-, and fuel-efficient. *See* SATELLITE NAVIGATION SYSTEMS.

At highway speeds, half the fuel burned in the engine is used to overcome the air resistance or aerodynamic drag of the cab, body, and trailer. Streamlined designs—for example, contoured windshields, hoods, fenders, and bumpers and devices such as air deflectors—reduce drag. This improves fuel economy and lowers the vehicle operating cost. *See* STREAMLINING.

Frame. The truck frame (**Fig. 2**) supports the load, power train, and steering mechanism while maintaining alignment of the components of the body

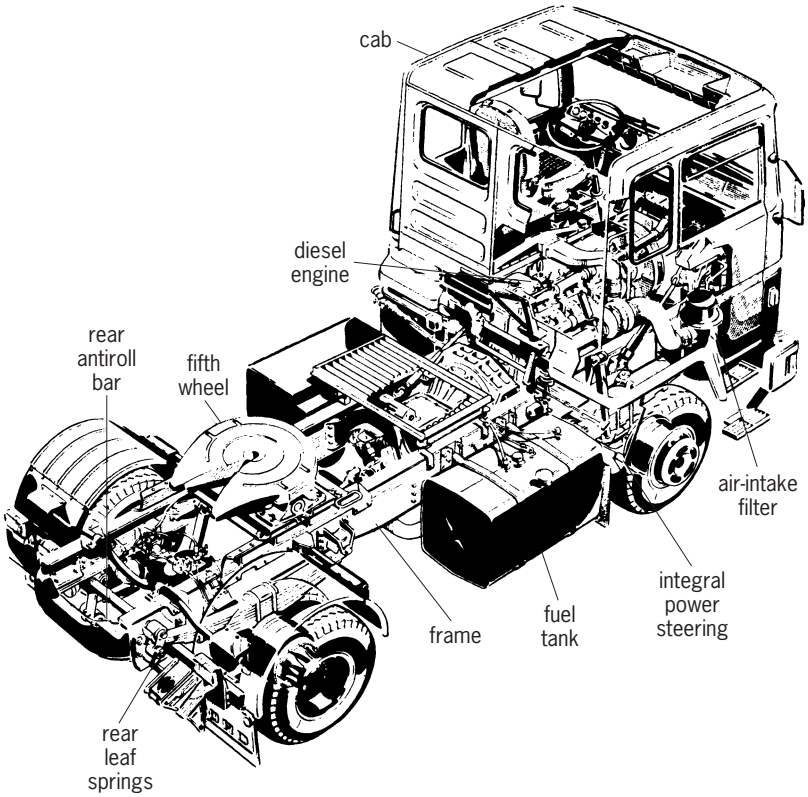

Fig. 2. Two-axle cab-over-engine truck tractor with fifth wheel mounted. (*Ford Motor Co.*)

and chassis. The load-bearing ability of a truck is determined by the strength of the frame, which is designed to handle the side, torsional, and vertical loads encountered in its load-rating category.

Frames are made of steel or aluminum alloy. Typical construction is of channel side rails held in place by cross-members, which resist buckling and frame twisting. In this ladder-type frame, the cross-members are usually riveted or welded to the side rails. Some cross-members are bolted in place to provide accessibility for service operations such as transmission removal. Most truck frames have a standard width of 34 in. (850 mm) between rails so specialty bodies from various body builders can be mounted. Integral body-and-frame construction is used in some light-duty trucks.

Engine. Trucks and truck tractors use inline, V-type, or pancake engines, which usually have 4, 6, 8, 10, or 12 cylinders, and use gasoline, compressed natural gas, liquid petroleum gas, or diesel fuel. The gasoline engines operate on the four-stroke cycle. The diesel engines are either two- or four-stroke. Electronic engine controls are used to help meet exhaust emission standards, reduce exhaust smoke, and improve fuel economy. *See* DIESEL FUEL; GASOLINE; LIQUEFIED NATURAL GAS (LNG); LIQUEFIED PETROLEUM GAS (LPG).

Most truck engines are liquid-cooled, and some are air-cooled. Brake horsepower for on-highway vehicles ranges from about 64 hp at 5500 revolutions/min for a gasoline engine in a light truck to 565 hp at 2000 revolutions/min for a diesel engine in a heavy-duty truck or tractor. Supercharging and turbocharging are used to develop more power from an engine of given size. *See* AUTOMOTIVE ENGINE; DIESEL ENGINE; ENGINE; INTERNAL COMBUSTION ENGINE; SUPERCHARGER; TURBOCHARGER.

Power train. The group of components that transmits power from the engine to the wheels is the drive train, or power train (**Fig. 3**). It includes the clutch, transmission, universal joints, drive shafts, and powered or drive axles. The clutch provides a means by which the driver can engage and disengage the transmission from the engine. A dry single-plate or two-plate friction clutch is normally used, with medium and heavy vehicles in severe start-stop operation having a wet clutch with two or more plates running in oil. *See* CLUTCH; UNIVERSAL JOINT.

Transmission. Manual, semiautomatic, and automatic transmissions are used in trucks and truck tractors. Because of the limited operating-speed range of most diesel engines, a relatively large number of forward gear ratios is required, especially in heavy-duty trucks and tractors. These may be provided, for example, by mounting a two-speed auxiliary transmission behind a five-speed manual transmission. This combination gives 10 forward speeds. When a two-speed rear drive axle is used, there are 20 possible forward speeds. *See* GEAR; GEAR TRAIN; PLANETARY GEAR TRAIN.

To reduce transmission length and weight, multiple countershaft designs are employed in some larger truck transmissions (**Fig. 4**). Two countershafts are spaced 180° apart on opposite sides of the mainshaft, or three countershafts are equally spaced around the mainshaft. Operation and function of the transmission are the same as with a single countershaft.

Two types of semiautomatic transmission are used, a straight mechanical one and a mechanical transmission with hydraulic torque converter. The semiautomatic allows the driver to choose any available gear ratio, while power shifting spares the driver the labor of gear changing. Declutching devices, which permit all forward shifts to be made without depressing the clutch pedal, also help the driver. With these, the clutch pedal is used only for standing starts. The automatics are similar to automobile automatic transmissions, with hydraulic torque converter and planetary gear trains providing four to six forward speeds.

Transfer case. Many transfer cases have two gear ratios and function as a combination two-speed auxiliary transmission and power divider. Transfer cases are used in military and other vehicles expected to engage in off-highway operation, and in on-highway vehicles requiring four-wheel or all-wheel drive. *See* AUTOMOTIVE TRANSMISSION.

Axle. Positioned transversely under the frame, an axle is a supporting member carrying the weight of a vehicle and its payload, and has mounted at either end the wheels on which the vehicle rolls. Drive axles transmit power from an input shaft to the wheels, forcing them to rotate. Nondriving or dead axles do not power the wheels but merely allow them to rotate freely. A steering function may be provided on either type by including means to pivot the wheels. Highway truck front axles (Fig. 3) are

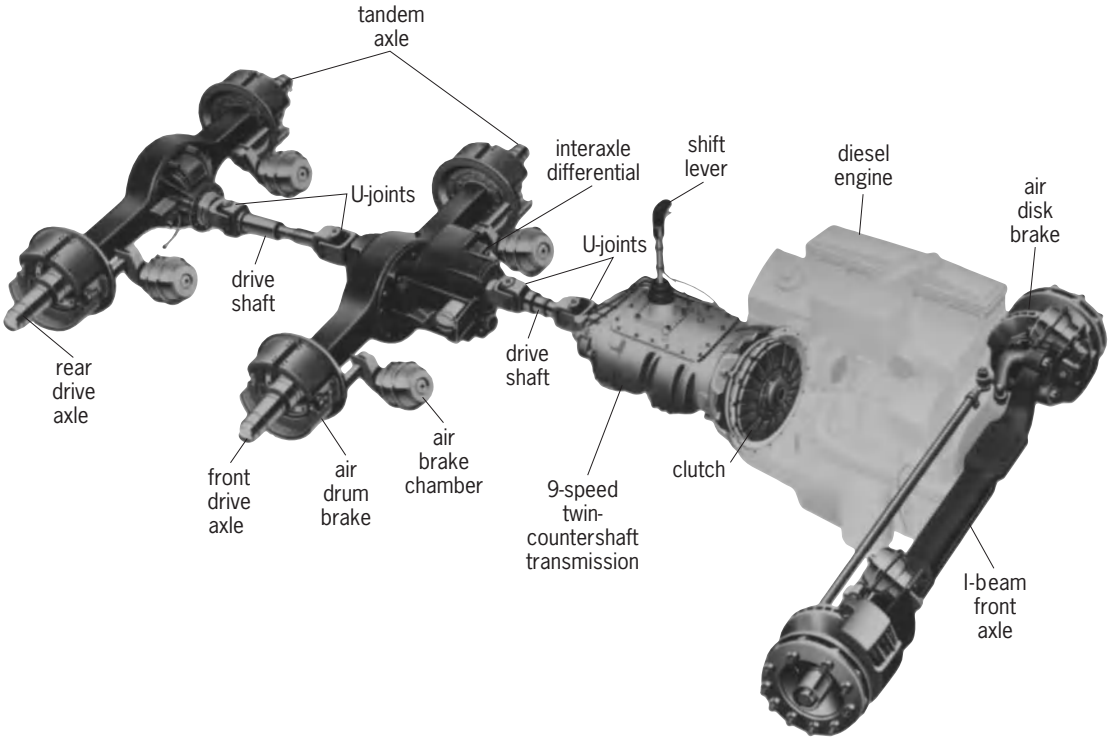

Fig. 3. Power train for an on-highway truck or truck tractor, with a nondriving front axle and a tandem axle at the rear. Both rear axles are driven. (*Rockwell International Corp.*)

typically the nondriving steer type, using a forged steel I-beam between pivot centers. Front axles for off-highway trucks are often steerable drive axles.

Drive axle. The truck or tractor may have a single driving axle at the rear for lighter loads, relatively short distances, and tight maneuvering. For heavier hauling, two rear axles (a tandem axle) or three rear axles (a triple or triaxle arrangement) may be used. In a tandem axle, the drive may be through one or both axles. When additional traction is needed, the vehicle may have all-wheel drive, in which the front axle and one or more rear axles are powered.

In a drive axle, power is transmitted from an input shaft at the center of the axle to a primary right-angle gear reduction, then to a differential mechanism integral with the gear reduction and through connecting axle shafts to the wheels. Drive axles may include a driver-controlled locking differential for additional traction during vehicle operation in mud, ice, and snow. Axles for off-highway trucks usually include additional reduction obtained with a planetary gear set at each wheel. Very large mining trucks may use a propulsion system adapted from the diesel-electric locomotive, with traction motors and planetary gear reduction at the driven wheels. *See* AUTOMOTIVE DRIVE AXLE; DIFFERENTIAL; LOCOMOTIVE.

Two-speed drive axles include an alternate second gear-reduction set, or power-flow path if of planetary design, thereby permitting operation in two speed ranges. Tandem- or triple-axle arrangements are powered by extending the input shaft through the first driving axle and coupling to the second (Fig. 3), and in like fashion from second to third driving axles.

An interaxle differential (Fig. 3) is used in tandem and triple arrangements to avoid internal torque buildup that could result from operation at slightly differing axle speeds, such as occurs with unmatched tires. Trucks used in off-highway service may operate satisfactorily without an interaxle differential, especially if operating in poor traction conditions. Axle lubricants are usually SAE-90 mineral oil modified with extreme-pressure additives because of

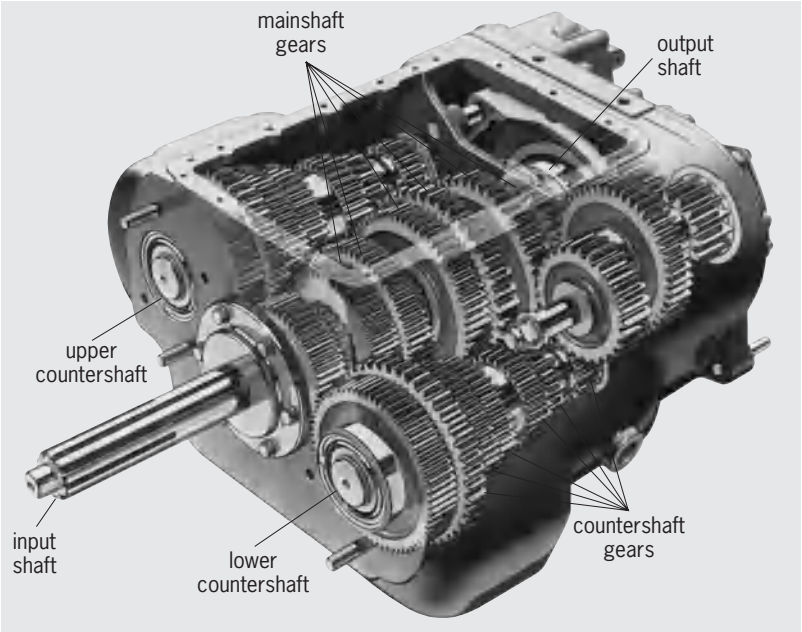

Fig. 4. Nine-speed twin-countershaft manual transmission. (*Rockwell International Corp.*)

the high specific loading and sliding tooth contact inherent in hypoid gear sets. *See* LUBRICANT.

Suspension. The suspension mounts the axle to the truck frame and reduces the shock and vibration transmitted to it. Leaf springs are widely used in truck suspension (Fig. 2). A front axle may be an I-beam suspended by leaf springs, or an independent front suspension using twin I-beams or control arms with coil springs. A variety of rear suspension systems are in use, including leaf, coil, rubber, and air springs. Air springs are employed especially when constant frame height and axle articulation, regardless of load, are important considerations. On tandem and triple-axle arrangements, the suspension includes the necessary beams and torque arms to distribute load to each axle. *See* AUTOMOTIVE SUSPENSION.

Steering. Trucks use Ackerman steering geometry, usually with a recirculating-ball steering gear. Power assist generally is provided by a hydraulic power-steering system or sometimes by air pressure in heavy-duty vehicles. Typical turning angle on nondrive steering axles is approximately 40°. Some heavy-duty trucks have a tandem front axle in which all four front wheels swing in and out for steering. For improved maneuverability, the rear axle on some trucks and trailers, such as fire and rescue vehicles, may also be a steering axle. *See* AUTOMOTIVE STEERING.

Brakes. Wheel brakes, commonly called service brakes or foundation brakes, and usually emergency or parking brakes are drum or disc brakes integral with or mounted on the axle (Fig. 3). Hydraulic brakes are used on light and medium-duty trucks, frequently with a hydraulic or vacuum brake booster for power assist. In heavy-duty trucks, air systems prevail because of their adaptability for tractor-trailer systems and their ability to operate at high temperatures that are generated in the brakes during severe service. Slack adjusters at each wheel brake automatically compensate for lining wear or are manually adjusted. Many newer trucks have an antilock braking system. *See* AIR BRAKE; AUTOMOTIVE BRAKE; BRAKE.

Truck brakes are not designed for continuous application. To prevent brake fade and possible failure during prolonged downhill braking, many heavy-duty trucks have an additional wear-free braking system called a retarder. It operates independently of the wheel brakes and may be a primary retarder located between the engine and transmission, or a secondary retarder located in the driveline between the transmission and the drive axle. Both types act to supplement the wheel brakes in dissipating the kinetic energy of the vehicle's motion.

Wheels and tires. Various combinations of wheels and tires are available for most trucks. Wheels and tires should be large enough for proper handling of the loads encountered. The total weight carried should not exceed the maximum rating of the tire and wheel. Trucks usually have steel or aluminum disk wheels or cast spoke wheels. Requirements for truck wheels include high fatigue strength and service life; lowest possible weight to increase payload; minimal unevenness, runout, and imbalance; and ease of assembly for tire mounting. A common truck and tractor wheel is a one-piece disk wheel with a rim having a 15° tapered-bead seat for use with a tubeless tire.

Tubeless radial tires are widely used to minimize rolling resistance and to improve fuel economy. Truck tires may be recapped to prolong carcass life. *See* TIRE.

Advances in truck technology. In light trucks, such as pickups, vans, and sport utility vehicles, many drivers expect the same comfort, convenience, and entertainment features as in an automobile, and most of these accessories are offered as factory- or dealer-installed options. In trucks of all sizes, electronics are incorporated into almost every system and component, including the transmission and brakes that previously could be controlled only by the driver.

To reduce engine fuel consumption and exhaust emissions, electronic devices and computer controls provide engine management and onboard diagnostics (**Fig. 5**). Rising fuel prices and the possibility of future fuel shortages are encouraging truck product development. Advances in truck technology include clean diesel technology, hybrid power sources, alternative fuels, lighter and stronger materials, and other methods of improving fuel efficiency, operation, and safety. *See* ELECTRIC VEHICLE.

Clean diesel technology. Most light-duty trucks have gasoline-fueled spark-ignition engines, although diesel engines in light-duty trucks have become more popular in recent years. Many medium-duty trucks and most heavy-duty trucks for highway use have diesel engines. However, diesel exhaust emissions of hydrocarbons (HC), carbon monoxide (CO), nitrogen oxides (NO_x), and soot or particulate matter (PM) are considered harmful to the

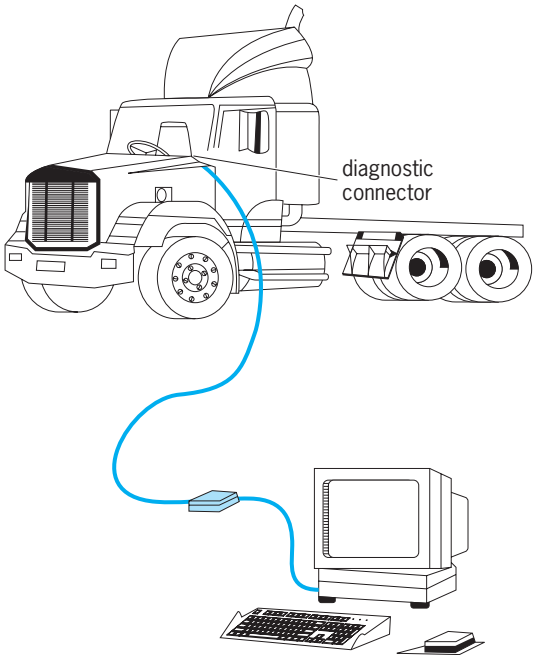

Fig. 5. Accessing the stored diagnostic trouble codes and other electronic data through the diagnostic connector in the cab of a heavy-duty truck.

environment and human health. In the presence of sunlight, nitrogen oxides react with hydrocarbons or carbon monoxide in the air to form ozone, smog, and acid rain. Small particles of particulate matter, if inhaled, may cause breathing and other health problems. *See* AIR POLLUTION.

The U.S. Environmental Protection Agency (EPA) issued new standards for the 2007 model year, requiring the petroleum industry to produce ultra-low-sulfur diesel fuel which is cleaner-burning and contains a maximum of 15 parts per million (ppm) sulfur by weight. This has been estimated as equivalent to one ounce of sulfur in a tank truck of diesel fuel. The new highway ultra-low-sulfur diesel fuel replaces most low-sulfur diesel fuel that contains up to a maximum of 500 ppm sulfur. By 2010, ultra-low-sulfur diesel fuel must also be used for off-highway vehicles.

Reducing the sulfur content of diesel fuel allows aftertreatment of the exhaust gas with particulate traps, which can be clogged by higher levels of sulfur. The combining of the three elements of (1) ultra-low-sulfur diesel fuel, (2) modifications of diesel engines to lower emissions of nitrogen oxides, and (3) advanced emission controls has gained the name "clean diesel technology." Application of clean diesel technology should result in diesel truck exhaust that has no smoke, no smell, and virtually no emissions.

Donald L. Anglin

Bibliography. *Bosch Automotive Handbook*, 2004; M. Ehsani et al., *Modern Electric, Hybrid Electric, and Fuel Cell Vehicles: Fundamentals, Theory, and Design*, CRC Press, 2005; Society of Automotive Engineers, *SAE Handbook*, 3 vols., annually; Society of Automotive Engineers, *Truck Systems Design Handbook*, vol. 2, 2002.

Truss

An assemblage of structural members joined at their ends to form a stable structural assembly. If all members lie in one plane, the truss is called a planar truss or a plane truss. If the members are located in three dimensions, the truss is called a space truss.

A plane truss is used like a beam, particularly for bridge and roof construction. A plane truss can support only weight or loads contained in the same plane as that containing the truss. A space truss is used like a plate or slab, particularly for long span roofs where the plan shape is square or rectangular, and is most efficient when the aspect ratio (the ratio of the length and width) does not vary above 1.5. A space truss can support weight and loads in any direction.

Because a truss can be made deeper than a beam with solid web and yet not weigh more, it is more economical for long spans and heavy loads, even though it costs more to fabricate. *See* BRIDGE; ROOF CONSTRUCTION.

The simplest truss is a triangle composed of three bars with ends pinned together. If small changes in the lengths of the bars are neglected, the relative positions of the joints do not change when loads are applied in the plane of the triangle at the apexes.

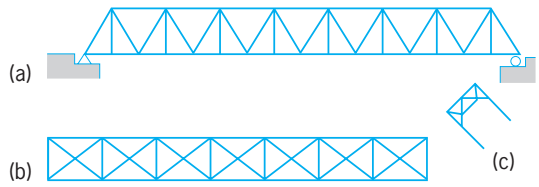

Fig. 1. Loads in through trusses are borne on the lower chord. (*a*) Through Warren truss with verticals. (*b*) Top-chord bracing as seen from above. (*c*) End-on view showing portal bracing.

Planar trusses. Such simple trusses as a triangle, perhaps with the addition of a vertical bar in the middle, are sometimes used to support peaked roofs of houses and other narrow structures. For longer spans, flat roofs, or bridges many triangles are combined to form a truss.

In metal trusses, connections may be riveted, bolted, welded, or pinned; in wood trusses, they may be bolted, nailed, or glued. Because of long spans, provision must be made to permit movement at one support due to loads and temperature changes; rollers, rockers, or sliding plates generally are used for this purpose.

The top members of a truss are called the upper chord; the bottom members, the lower chord; and the verticals and diagonals, the web members.

Framing to be carried by a truss usually is arranged so that it brings loads to bear on the truss at the intersections of a chord and web members. As a result, truss members are subjected only to direct axial stress—tension or compression—and can be made of less material than if they also had to resist bending stresses.

Roof trusses carry the weight of roof deck and framing and wind loads on the upper chord. They may also support a ceiling or other loads on the lower chord. On the other hand, bridge trusses may carry loads on either chord. Deck trusses support loads on the upper chord; through trusses, on the lower chord (**Fig. 1**).

To maintain stability of truss construction, bracing must be used normal to the planes of the trusses. Usually framing is inserted between the trusses. For roofs, trussed bracing should be placed in the plane of either the top chord or the bottom chord. For bridges, bracing must be inserted in the planes of both top and bottom chords, because of the greater need for stability under heavy moving loads. Vertical trusses are known as braced towers. Vertical trusses are used extensively in buildings to resist wind and earthquake loads.

Space trusses. Space trusses are essentially composed of two sets of plane trusses, with the top and bottom chords intersecting the edges at 90° (orthogonal; **Fig. 2**) or three sets of plane trusses with top and bottom chords intersecting the edges at acute angles (skewed; **Fig. 3**). Most space trusses have the individual plane trusses lying in an inclined plane so that the entire system is inherently stable. The inclination of the plane trusses results in

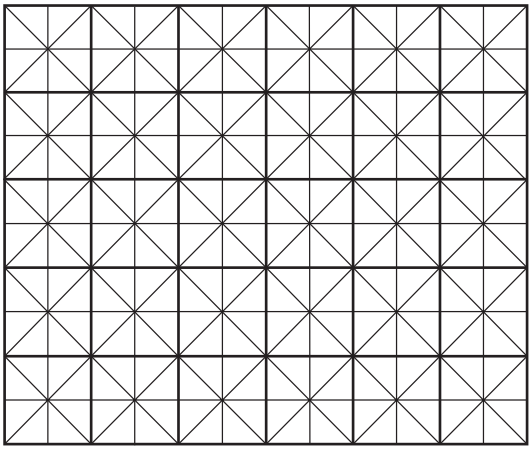

Fig. 2. Orthogonal truss.

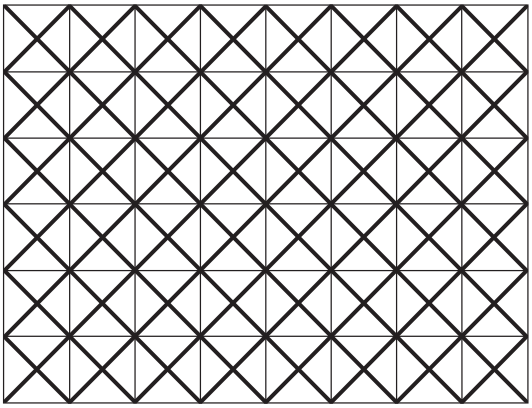

Fig. 3. Skewed space truss.

the basic element of the space truss being a pyramid (a tetrahedron). As a result, the top and bottom chord are offset one-half module. Curved space trusses are used to span long distances as part of an overall structural shape, in particular, a dome. One such space truss dome is the geodesic dome. *See* GEODESIC DOME.

Computing stresses. Primary axial stresses in truss members are computed on the assumption that connections at joints are made with frictionless pins. With loads applied at joints, each truss member or bar is subjected to pure tension or compression.

Since the bars change length under load, the angles of each triangle constituting the truss tend to change. But this change is resisted, since pins are not frictionless, and since rivets, bolts, or welds offer restraint. Consequently, members bend slightly, the bending moments creating secondary stresses.

Planar trusses of single-span configuration lend themselves to simple statical hand calculations. At a truss joint, the primary stresses and loads form a coplanar, concurrent force system in equilibrium. This force system satisfies two conditions: the sums of the horizontal and vertical components both equal zero. These equations are used in computing stresses by the method of joints: joints with two unknowns are selected in succession and the two equilibrium equations are applied to them to determine the stresses.

A section may be passed through the truss to cut three bars with unknown stresses. These, together with bars with known stresses that are cut and the loads on the part of the structure on either side of the section, constitute a coplanar, noncurrent force system in equilibrium. This system satisfies the two previous conditions; but in addition, the sum of the moments of the forces about any axis normal to the plane equals zero. With these three equations the three unknowns can be determined. However, the unknowns also can be found by the method of moments, in which two of the unknowns are eliminated by taking the moment axis at their point of intersection, and the third is found by equating the sum of the moments to zero. The method of shears is used to determine one force when the other two unknown forces are both normal to the shearing force, for example, for finding the stresses in the diagonals of parallel-chord Warren, Pratt, and Howe trusses.

If n is the number of joints in a truss, stresses can be found by the methods of joints, moments, or shears when the number of bars equals $2n - 3$. If a truss is composed of fewer bars, it is unstable; if of more bars, statically indeterminate.

Influence lines are useful in determining the stresses in bridge trusses, because the live load is a moving load. Influence lines can be drawn to show the variation in any function—stress, shear, moment, deflection—as a unit load moves along the truss.

Multiple-span plane trusses (defined as statically indeterminate or redundant) and space trusses require very complex and tedious hand calculations. Modern high-speed digital computers and readily available computer programs greatly facilitate the structural analysis and design of these structures. *See* COMPUTER; STRUCTURAL ANALYSIS.

Charles Thornton; I. Paul Lew

Trypanites

A simple cylindrical boring common in the fossil record for the past 540 million years. *Trypanites* is an ichnogenus, which is a formal name given to a distinct trace fossil (evidence of organism behavior in the fossil record as distinct from the remains or other representations of the organism's body). *Trypanites* is thus not a type of life but a structure built by a variety of wormlike animals over time. *See* TRACE FOSSILS.

Trypanites is formally defined as a cylindrical, unbranched boring in a hard substrate (such as a rock or shell) with a length up to 50 times its width (see **illustration**). Most *Trypanites* are only a few millimeters long, but some are known to be up to 12 cm (4.72 in.). They usually penetrate the substrate perpendicularly to its surface and remain straight, but they are sometimes found at oblique angles, and a

Trypanites **borings in a carbonate hardground from the Upper Ordovician of northern Kentucky. The borings descend from the top of the rock surface perpendicularly to the bedding. Each boring is filled with tiny dolomite crystals, making them stand out against the dark limestone. The longest boring is 4.7 cm (1.85 in.).**

few may curve in response to irregularities in the material they bored. Almost all *Trypanites* are found in calcareous substrates, which is evidence that the producing organisms may have used at least in part some sort of acid or other dissolving chemical to excavate the substrate. *Trypanites* is often confused with the similar boring *Palaeosabella*, which has a clavate (expanded) terminus rather than a simple rounded one. Since *Trypanites* have a circular cross section, they can sometimes be confused with predatory borings (*Oichnus*) when they pass through shells which are later liberated by erosion from their lithological (rock) context.

The organisms which produced *Trypanites* have varied over geological time. Most *Trypanites* were made by marine wormlike animals, with a few known from freshwater paleoenvironments. Very rarely remains are found in the fossil borings which can be attributed to particular groups, such as polychaetes. Similar borings are made today by sabellarid polychaete and sipunculid worms, among other groups. These borings are made as dwellings (domichnia) for filter-feeding the waters above, so they have nothing to do with predation or digestion of the substrate.

Trypanites is important to paleontologists as one of the earliest examples of bioerosion (the erosion of hard substrates by biological actions). It is the first known macroboring (a boring easily visible to the naked eye) and is found in Lower Cambrian hardgrounds about 540 million years old. *Trypanites* is surprisingly rare until the Middle and Late Ordovician, when it reappears in great abundance as part of the Ordovician Bioerosion Revolution. This was a time when many other macroborings enter the fossil record and bioerosion intensity greatly increased. *Trypanites* is the most common macroboring until the Jurassic (about 200 million years ago), when it was exceeded in abundance by the bivalve boring *Gastrochaenolites* and other more complex tubular borings. *Trypanites* is also one of the first borings to host nestling organisms, such as brachiopods, inside the empty cavities, thereby expanding ecological niche space on marine hard substrates. *See* ORDOVICIAN; PALEOECOLOGY. Mark A. Wilson

Bibliography. R. G. Bromley, On some ichnotaxa in hard substrates, with a redefinition of *Trypanites* Mägdefrau, *Paläontologische Zeitschrift*, 46:93–98, 1972; N. P. James, D. R. Kobluk, and S. G. Pemberton, The oldest macroborers: Lower Cambrian of Labrador, *Science*, 197:980–983, 1977; P. D. Taylor and M. A. Wilson, Palaeoecology and evolution of marine hard substrate communities, *Earth-Sci. Rev.*, 62:1–103, 2003; M. A. Wilson and T. J. Palmer, *Hardgrounds and Hardground Faunas*, University of Wales, Aberystwyth, Institute of Earth Studies Publ. 9, 1992.

Trypanorhyncha

An order of tapeworms of the subclass Cestoda, also known as the Tetrarhynchoidea. All are parasitic in the intestine of elasmobranch fishes. They are distinguished from all other tapeworm groups by having spiny, eversible proboscides on the head. An elongated head stalk contains the proboscis apparatus made up of a proboscis sheath and a muscular bulb. The head also bears two or four shallow, weakly muscular suckers (see **illus.**). Segment anatomy resembles that of Proteocephaloidea, except that the yolk glands are scattered. A complete life history is not known for any trypanorhynchid, although larval forms have been found in the tissues of various

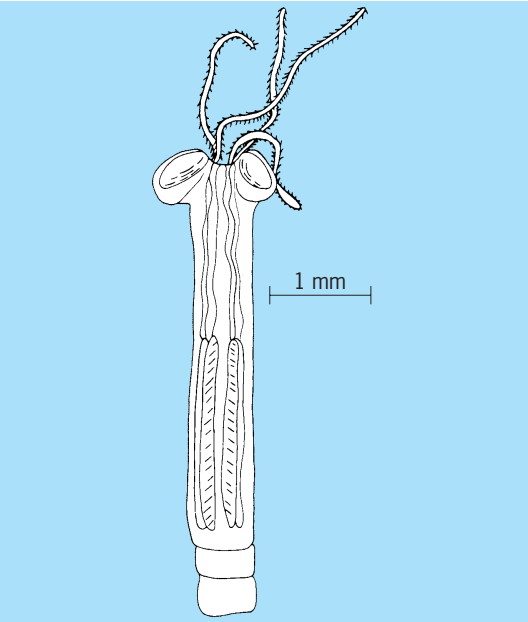

Scolex of tapeworm *Eutetrarhynchus*.

marine invertebrates and teleost fishes. *See* EUCESTODA; PROTEOCEPHALOIDEA. Clark P. Read

Trypanosomatidae

A family of Protozoa, order Kinetoplastida, containing flagellated parasites which change their morphology; that is, they exhibit polymorphism during their life cycles. The life cycles of the organisms may involve only an invertebrate host, or an invertebrate and a vertebrate host, or an invertebrate and a plant host. Several distinct morphological forms are recognized: trypanosomal, crithidial, leptomonad, and leishmanial. Differentiation into genera is dependent upon the host infected as well as the morphologic types involved. None of the stages possesses a mouth opening, and nutritive elements are absorbed through the surface of the body; that is, the organisms are saprozoic. **Figure 1** illustrates the morphologic stages and the hosts of the genera.

Morphology. All the stages possess a single nucleus and a kinetoplast composed of a rod-shaped parabasal body and a minute basal granule called the blepharoplast. In the elongated stages of the parasite, the kinetoplast may be located at the anterior end of the leptomonads, near the center in the crithidials, or at the posterior end in the trypanosomal forms. With the exception of the leishmanial stage, either an axoneme or fiberlike structure arises from the basal granule and extends immediately out of the body as a free flagellum in the leptomonad or along the edge of an undulating membrane in the crithidial and trypanosomal forms. The axoneme may terminate at the end of the undulating membrane or continue as a free flagellum. The organism moves in the direction of the protruding flagellum. The undulating membrane is a finlike structure composed of two folds of the outer pellicle of the body, with the axoneme supporting the outer edge.

Reproduction. Multiplication of the Trypanosomatidae is usually by longitudinal binary fission. For example, in the trypanosomal stage the process is as follows: The kinetoplast divides first and a second axoneme develops from the new basal granule; as the axoneme increases in length toward the anterior end, the nucleus divides; finally, when all the structures have been duplicated, the body itself splits longitudinally beginning at the anterior end (**Fig. 2**).

Taxonomy. The six generally recognized genera of the family Trypanosomatidae are *Trypanosoma*, *Leishmania*, *Leptomonas*, *Phytomonas*, *Crithidia*, and *Herpetomonas*. The first two, *Trypanosoma* and *Leishmania*, are of medical and veterinary importance. Since they are found in the blood of vertebrates, they are referred to as hemoflagellates. The other four, *Leptomonas*, *Phytomonas*, *Crithidia*, and *Herpetomonas*, occur only in invertebrates and plants.

Trypanosoma. This is the most important genus of the family Trypanosomatidae from a number of standpoints. It contains the largest number of species infecting a wide variety of hosts such as mammals, birds, fishes, amphibians, and reptiles. Although most of the species cause no damage to the hosts, there are several which produce serious diseases in humans, domesticated animals, and wild animals. The pathogenic species, prevalent in Africa, have

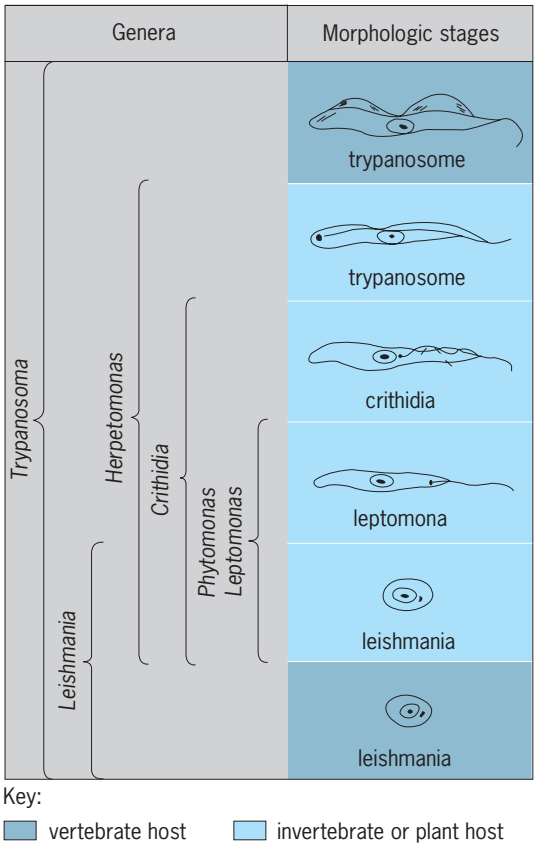

Fig. 1. Morphologic stages and the hosts of genera of Trypanosomatidae.

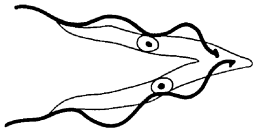

Fig. 2. A dividing trypanosome.

been responsible to a great extent for the slow development of civilization in many parts of that continent.

Developmental stages. Certain species of *Trypanosoma* possess all four developmental stages in their life cycles—trypanosomal, crithidial, leptomonad, and leishmanial forms; however, the trypanosomal stage is the most important, being found in the circulating blood of all of the vertebrate hosts. Differentiation of the species is based upon the morphology of the blood-form trypanosomes, the particular vertebrates and invertebrates serving as the hosts, the stages present in the hosts, and the course of development and location of the various stages in the hosts.

Morphology. The trypanosomal stages of the different species differ in size and shape, location of the nucleus, presence and position of the kinetoplast, and development of the undulating membrane. Trypanosomal forms range in length from less than 15 to over 80 micrometers, and may be either slender or broad. In general the nucleus is located near the center, but the kinetoplast may be at the posterior tip or some distance from it. The undulating membrane may be very prominent with many convolutions, or more or less flat and inconspicuous. There may or may not be a free flagellum at the anterior end.

Reproduction. Reproduction of the *Trypanosoma* may take place in several stages of the parasite and at different locations in the vertebrate and invertebrate hosts. Typically it is by longitudinal binary fission, but in some instances the division of the cytoplasm is delayed and multiple fission occurs.

Life history. With only rare exceptions, an invertebrate host such as a fly, bug, or leech is involved in the transmission of the *Trypanosoma* from vertebrate to vertebrate. In cyclical transmission it is necessary for the parasite to undergo developmental changes in the vector before infective, or metacyclic, trypanosomes occur. In the insect, these infective stages, depending upon the species, are located either in the salivary gland or proboscis and are transmitted by the bite, as in the case of *T. gambiense* and *T. rhodesiense*, or are located in the hindgut and are deposited on the surface of the vertebrate when the vector defecates, as in *T. cruzi*. In the latter case, the infective stage enters the body of the vertebrate host through a break in the skin, sometimes through the hole produced by the bite of the insect. Mechanical transmission can occur with the various species of *Trypanosoma* perhaps without exception, and with some it is the principal means of transmission. The trypanosomes thus obtained during a blood meal merely survive on the mouthparts of the insect and are introduced into another vertebrate when the insect feeds again within a few minutes. In *T. equiperdum* infections of horses and donkeys, contact transmission occurs during the sexual act and thus no intermediate invertebrate host is required in the life cycle.

Nonpathogenic species. Although the life cycles of the numerous species of *Trypanosoma* may differ in various respects, the one for *T. lewisi* serves to illustrate the complexity of the process for a nonpathogenic form. This parasite is worldwide in its distribution and is found in various species of rats. Under ordinary conditions it causes no apparent harm in the vertebrate host. The flea *Nosopsyllus fasciatus* transmits the organism from rat to rat. After the flea ingests the trypanosome stage in the rat's blood, a cyclical development of the parasite takes place in the intestinal tract of the invertebrate host. The long blood-form trypanosomes enter the epithelial cells lining the stomach, and reproduce by multiple fission. The modified trypanosomes which emerge from the cells migrate to the hindgut, where they attach to the lining by their anterior ends. They transform into the crithidial form and divide by binary fission. The infective, short metacyclic stages finally develop and accumulate in the rectum, from which they pass out with the feces. The rat becomes infected through ingestion of the flea's feces contaminating its body or by ingestion of the entire infected flea. For 8 or more days after reaching the bloodstream of the rat, the parasites reproduce by multiple fission in the crithidial stage and in other bizarre shapes. During this process, the individual organism divides several times without complete fission of the cytoplasm. Finally, the progeny break away from one another and develop into separate trypanosomes. At the end of this reproductive period, only typical trypanosomal forms remain. In a month or more they are destroyed by the immune response of the host.

Other interesting nonpathogenic species are *T. rotatorium* in the frog, which is transmitted by leeches; *T. granulosum* in the eel; and *T. danilewskyi* in goldfish.

Information on a number of the important pathogenic *Trypanosoma* is presented in the **table**.

Pathogenic species. Some authorities believe that the species which cause sleeping sickness, *T. gambiense* and *T. rhodesiense*, represent the wild animal species, *T. brucei*, which have become adapted to the human body; ordinarily the serum of humans is trypanocidal for *T. brucei*. In any event, the species are very similar morphologically and exhibit comparable development in the tsetse fly, *Glossina*. After the African sleeping sickness parasites enter the human body during the bite of the tsetse fly, they first multiply in the lymph and blood. Later they may invade the nervous system. Only the trypanosome stage occurs in humans. In the tsetse flies *G. palpalis* and *G. morsitans*, multiplication takes place first in the midgut while in the trypanosome stage. Later the parasites migrate to the salivary glands, where they transform into crithidial forms and multiply. Eventually metacyclic trypanosomes develop in this site. *See* DIPTERA.

Trypanosoma cruzi, the cause of Chagas' disease,

Epidemiology of *Trypanosoma* species

Species of *Trypanosoma*	Principal vertebrate hosts	Disease	Geographical distribution	Insect vectors	Mode of transmission
T. gambiense	Humans and domestic animals	Sleeping sickness	West Equatorial Africa	Tsetse flies (*Glossina*)	Cyclical; insect bite
T. rhodesiense	Probably humans and wild animals	Sleeping sickness	East Tropical Africa	Tsetse flies (*Glossina*)	Cyclical; insect bite
T. cruzi	Human, dog, armadillo, opossum, and other animals	Chagas' disease	South and Central America	Kissing bugs (*Triatoma*)	Cyclical; feces of bug
T. brucei	Domestic and wild mammals	Nagana	Tropical Africa	Tsetse flies (*Glossina*)	Cyclical insect bite
T. vivax	Domestic and wild mammals	Souma	Tropical Africa and South America	Tsetse flies (*Glossina*); stable flies (*Stomoxys*)	Cyclical (*Glossina*) and mechanical (*Stomoxys*); insect bite
T. equinum	Domestic and wild mammals	Mal de caderas	Tropical and South America	Biting flies (*Tabanus, Stomoxys*)	Mechanical; insect bite
T. evansi	Domestic and wild mammals	Surra	Asia, Australia, Madagascar	Biting flies (*Tabanus, Stomoxys*)	Mechanical; insect bite
T. hippicum	Domestic mammals, especially horses and mules	Murrina de caderas	Central America	Nonbiting flies (*Musca*)	Mechanical; by flies
T. equiperdum	Horses and donkeys	Dourine	Mediterranean countries	Usually none	Contamination; sexual act

presents a somewhat different cycle in humans and triatomid bugs. These rather large arthropods, about 1 in. (2.5 cm) in length, frequently feed around the face of the sleeping person and therefore are commonly referred to as kissing or barber bugs. In the human body the trypanosomal stages circulating in the bloodstream do not divide. All multiplication takes place in tissue cells. The parasites enter various tissues including the heart muscles, transform into leishmanial stages, and divide by binary fission until a cluster of parasites fills each cell. Before breaking out of the cells into the bloodstream, they return to the trypanosomal stage. The parasite may reenter other cells and multiply again. The trypanosome, after its ingestion by the kissing bug, multiplies first in the midgut and later in the rectum. Multiplication in the bug is primarily during the crithidial stage. Metacyclic trypanosomes develop in the rectum and pass out of the bug when it defecates. The infective stages therefore are not inoculated into the vertebrate but gain entrance into the body through contamination or breaks in the skin or through the conjunctiva of the eye. *See* HEMIPTERA.

Although *T. cruzi* is apparently present in traitomid bugs and lower animals such as the raccoon, skunk, opossum, and armadillo in various parts of the southern half of the United States, the first proved human infection in this country was not reported until 1955.

Another species, *T. rangeli*, has been recognized as a parasite of humans. In South and Central America this apparently nonpathogenic species must be differentiated from *T. cruzi* in examining the intestinal contents of triatomid bugs and the blood of humans and other vertebrates.

Leishmania. This is the second most important genus, at least from humanity's standpoint. Species of this genus occur as typical leishmanial forms in vertebrate hosts and as leptomonad stages in invertebrate hosts. Three species parasitize humans but have also been found naturally infecting dogs, cats, and perhaps other lower animals. The sand fly, *Phlebotomus*, transmits the parasite from vertebrate to vertebrate. After ingestion the leishmanial stage transforms into the leptomonad form in the gut of the fly and multiplies. The leptomonads are the infective stage for humans and are introduced by the bite of the infected fly. They enter various cells of the body such as the skin, capillaries, spleen, and liver, transform to the leishmanial stage, and multiply.

The three species of *Leishmania* are morphologically identical. Their small, oval bodies, about 5 μm in length, have a relatively large nucleus and kinetoplast, but no flagellum. The species are distinguished by their geographical distribution, the tissues they infect, and their immune reactions. They all produce serious diseases in humans which are difficult to control and treat. *Leishmania donovani* infects primarily the internal organs, causing kala azar or visceral leishmaniasis. *Leishmania tropica* is limited to the surface of the body, producing skin lesions (oriental sore) or cutaneous leishmaniasis. *Leishmania brasiliensis* is also limited to the surface of the body, but produces skin lesions (espundia, forest yaws) which frequently involve the mucous membranes of the nose, mouth, and pharynx; this is mucocutaneous leishmaniasis.

Leptomonas. In this genus, the kinetoplast is situated near the anterior end of the elongated body. The axoneme arising from the blepharoplast extends

directly out of the body as a free flagellum. Leptomonads are exclusively parasitic in invertebrates; for example, they occur in the hindgut of the common dog tick, *Ctenocephalus canis*. The nonflagellated leishmanial form is the infective stage.

Phytomonas. Morphologically similar to *Leptomonas*, phytomonads infect the latex of certain plants, for example, milkweed. Multiplication of the flagellates may cause degeneration of the infected part of the plant. Hemipterous insects which feed on latex transmit the parasite from plant to plant.

Crithidia. Species of this genus are parasitic in various arthropods. The kinetoplast is anterior to the centrally placed nucleus. There are generally a short undulating membrane and a free anterior flagellum. Depending upon the species, the organism may be found in the intestinal tract or body cavity of invertebrates including water bugs and ticks. Leptomonad and leishmanial stages develop, and encysted forms may serve as the infective stage.

Herpetomonas. All four stages, trypanosome, crithidial, leptomonas, and leishmanial, occur in the life cycle of species in this genus. The several species are exclusively parasitic in invertebrates. They infect the intestinal tracts of various species of flies. Transmission is by ingestion of encysted forms. M. M. Brooke; Horace W. Stunkard

Trypanosomiasis

A potentially fatal infection caused by parasites of the genus *Trypanosoma*.

African trypanosomiasis. The African trypanosomes, the cause of African trypanosomiasis or African sleeping sickness, are flagellated protozoan parasites. They are members of the *T. brucei* group. *Trypanosoma brucei rhodesiense* and *T. b. gambiense* cause disease in humans. *Trypanosoma brucei* is restricted to domestic and wild animals. The trypanosomes are transmitted by the tsetse fly (*Glossina*), which is restricted to the African continent. The trypanosomes are taken up in a blood meal and grow and multiply within the tsetse gut. After 2–3 weeks, depending upon environmental conditions they migrate into the salivary glands, where they become mature infective forms, and are then transmitted by the injection of infected saliva into a new host during a blood meal. The survival of the tsetse fly is dependent upon both temperature and humidity, and the fly is confined by the Sahara to the north and by the colder drier areas to the south, an area approximately the size of the United States. Approximately 50 million people live within this endemic area, and 15,000–20,000 new human cases of African trypanosomiasis are reported annually.

In humans and other mammals the trypanosomes are extracellular. During the early stages of infection, the trypanosomes are found in the blood and lymph but not in cerebrospinal fluid. There is fever, malaise, and enlarged lymph nodes. In the absence of treatment the disease becomes chronic and the trypanosomes penetrate into the cerebrospinal fluid and the brain. The symptoms are headaches, behavioral changes, and finally the characteristic sleeping stage. Without treatment the individual sleeps more and more and finally enters a comatose stage which leads to death. Treatment is more difficult if the infection is not diagnosed until the late neurological stage. The two species *T. b. rhodesiense* and *T. b. gambiense* are found in different geographical areas, have different vectors and different reservoir hosts, and produce different types of disease (see **table**).

The epidemiology of African trypanosomiasis is influenced by the environment, the abundance and type of vector, the type and availability of reservoir hosts, the type of human activities, and numbers and activities of domestic livestock. Control depends on a detailed knowledge of these factors and includes avoiding areas where there is a high density of tsetse flies; altering the environment by brush clearing; removal of reservoir hosts; use of insect traps and insecticides; and drug treatment of humans and domestic animals. By use of these protocols in an environmentally sound manner, African trypanosomiasis

Comparison of the biology of the human infective *Trypanosoma brucei gambiense* and *T. b. rhodesiense*

Characteristics	*T. b. gambiense*	*T. b. rhodesiense*
Vector	*Glossina palpalis* group, riverine tsetse	*G. morsitans* group, savana, woodland tsetse
Method of transmission	Human → Tsetse → Human ↓ Animal reservoirs	Animal reservoirs → Tsetse → Animal reservoirs ↓ Human
Resevoir hosts	Possibly kob, hartebeest, domestic pigs, dogs	Bushbuck, waterbuck, hartebeest, hyena, lion, domestic cattle, possibly warthog and giraffe
Geographical area	West and North-Central Africa	Central and East Africa
Disease	Chronic, low parasitemia, incubation period of months to years	Acute, high parasitemia; incubation period of days to weeks

can be controlled but not eradicated. *See* MEDICAL PARASITOLOGY.

Chagas' disease. *Trypanosoma cruzi*, the cause of American trypanosomiasis or Chagas' disease, is transmitted by biting insects of the Reduviidae family (subfamily Triatominae). The parasite is distributed throughout most of South and Central America, where it infects an estimated 12–20 million people and over 35 million individuals are exposed to the parasite. Over 100 different species of wild and domestic animals have also been found to be infected. Animals that live in proximity to humans such as dogs, opossums, and wood rats are important reservoir hosts. Infected animals have been detected in the southwestern United States. However, only a few infections have ever been diagnosed in individuals from the United States who have not traveled to Central or South America.

Trypanosoma cruzi is predominantly an intracellular parasite in the mammalian host. During the intracellular stage, *T. cruzi* loses its flagellum and grows predominantly in cells of the spleen, liver, lymphatic system, and cardiac, smooth, and skeletal muscle. The cells of the autonomic nervous system are also frequently invaded. The parasites grow rapidly, forming closely packed pockets of parasites called pseudocysts. Rupture of the infected cells releases these intracellular stages into the surrounding tissue spaces, where they become flagellated and invade the blood of the host. These flagellated forms are taken up in the reduviid blood meal. The trypanosomes migrate into the insect gut, where they multiply and finally are excreted.

The disease has an acute and a chronic stage. The acute stage is initiated by scratching the trypanosomes that are in insect feces into the wound produced by the insect bite. Thus, individuals are infected by contamination following the bug bite rather than by direct inoculation. The acute phase is characterized by local inflammation at the site of the bite and by fever, muscle and bone pain, headaches, and anemia. However, the symptoms are extremely varied, and a clear diagnosis is difficult to obtain. The acute symptoms disappear spontaneously within several months in 90–95% of the cases. In many individuals the infection remains silent, and death results from causes unrelated to *T. cruzi*. However, in some individuals, the disease becomes chronic, and over a period of years serious symptoms appear. One sympton is chronic myocarditis, an inflammation of the heart muscles. In some areas, Chagas' disease accounts for a majority of cardiac deaths in young adults. In other geographical regions a second condition is observed and is referred to as megaesophagus or megacolon. It is caused when *T. cruzi* invades and destroys the autonomic ganglia of the esophagus or colon. This leads to enlargement of the colon or esophagus, and in some cases death. Although there is only one known species of *T. cruzi*, different isolates appear to have differences in virulence. Possibly this explains the different symptoms observed in different geographical regions.

In villages located in endemic areas, the insect vector lives within thatched roofs, cracked walls, or trash-filled rooms, and infection usually occurs at night while the individual is asleep. In addition, individuals become infected when villagers go into forest areas in which reservoir hosts and vectors are present. Transmission also occurs by the tranfusion of blood donated by infected individuals.

There is no satisfactory chemotherapeutic agent for treatment, and a vaccine is unavailable. Therefore the methods for control of Chagas' disease are directed toward the vector. For example, houses are treated with insecticides, cracks within walls are filled, and metal roofs (replacing thatched roofs) are used to eliminate vector breeding sites. These techniques are effective in reducing the incidence of Chagas' disease. Unfortunately, they are all costly and depend upon improved socioeconomic conditions in endemic areas. *See* PARASITOLOGY.

John Richard Seed

Bibliography. J. R. Baker (ed.), *Parasitic Protozoa*, vol. 2, 2d ed., 1992; L. S. Roberts, *Foundations of Parasitology*, 5th ed., 1995.

Tsunami

A set of ocean waves caused by any large, abrupt disturbance of the sea surface. If the disturbance is close to the coastline, a local tsunami can demolish coastal communities within minutes. A very large disturbance can cause local devastation and destruction thousands of miles away. Tsunami comes from the Japanese language, meaning harbor wave.

Tsunamis rank high on the scale of natural disasters. Since 1850, they have been responsible for the loss of over 420,000 lives and billions of dollars of damage to coastal structures and habitats. Most casualties were caused by local tsunamis that occur about once per year somewhere in the world. For example, the Indian Ocean tsunami on December 26, 2004, killed about 164,000 people close to the earthquake and about 63,000 people on distant shores. Predicting when and where the next tsunami will strike is currently impossible. Once a tsunami is generated, forecasting its arrival and impact is possible through modeling and measurement technologies.

Generation. Tsunamis are most commonly generated by earthquakes in marine and coastal regions. Major tsunamis are produced by large (greater than magnitude 7), shallow-focus (less than 30 km or 19 mi deep in the Earth) earthquakes associated with the movement of oceanic and continental plates. They frequently occur in the Pacific, where dense oceanic plates slide under the lighter continental plates. When these plates fracture, they provide a vertical movement of the sea floor that allows a quick and efficient transfer of energy from the solid earth to the ocean. When a powerful (magnitude 9.3) earthquake struck the coastal region of Indonesia in 2004, the movement of the sea floor produced a tsunami in excess of 30 m (100 ft) in height along the adjacent coastline. From this source, the tsunami radiated

outward and within 2 hours had claimed 63,000 lives in Thailand, Sri Lanka, and India. *See* EARTHQUAKE; PLATE TECTONICS.

Underwater landslides associated with smaller earthquakes are also capable of generating destructive tsunamis. The tsunami that devastated the northwestern coast of Papua New Guinea on July 17, 1998, was generated by a magnitude 7 earthquake that apparently triggered a large underwater landslide. Three waves measuring more than 7 m (23 ft) high struck a 10-km (6-mi) stretch of coastline within 10 minutes of the earthquake/slump. Three coastal villages were swept completely clean by the attack, leaving nothing but sand and 2200 people dead. Other large-scale disturbances of the sea surface that can generate tsunamis are explosive volcanoes and asteroid impacts. The eruption of the volcano Krakatoa in the East Indies on August 27, 1883, produced a 30-m (100-ft) tsunami that killed over 36,000 people. In 1997, scientists discovered evidence that a 4-km-diameter (2.5-mi) asteroid landed offshore of Chile approximately 2 million years ago and produced a huge tsunami that swept over portions of South America and Antarctica. *See* ASTEROID; SEISMIC RISK; VOLCANO.

Wave propagation. Because earth movements associated with large earthquakes are thousand of square kilometers in area, any vertical movement of the sea floor immediately changes the sea surface. The resulting tsunami propagates as a set of waves whose energy is concentrated at wavelengths corresponding to the earth movements (~100 km or 60 mi), at wave heights determined by vertical displacement (~1 m or 3 ft), and at wave directions determined by the adjacent coastline geometry. Because each earthquake is unique, every tsunami has unique wavelengths, wave heights, and directionality. From a tsunami warning perspective, this makes forecasting tsunamis in real time daunting. *See* OCEAN WAVES.

Warning systems. Since 1946, the tsunami warning system in the Pacific basin has monitored earthquake activity and the passage of tsunami waves at tide gauges. However, neither seismometers nor coastal tide gauges provide data for accurately predicting the impact of a tsunami at a particular coastal location. Monitoring earthquakes gives a good estimate of the potential for tsunami generation, based on earthquake size and location, but gives no direct information about the tsunami itself. Tide gauges in harbors provide direct measurements of the tsunami, but the tsunami is significantly altered by local bathymetry and harbor shapes, which severely limits their use in forecasting tsunami impact at other locations. Partly because of these data limitations, 15 of 20 tsunami warnings issued since 1946 were considered false alarms because the tsunami that arrived was too weak to cause damage. *See* SEISMOGRAPHIC INSTRUMENTATION.

Forecasting impacts. Real-time, deep-ocean tsunami detectors (tsunameters) will provide the data necessary to make tsunami forecasts (**Fig. 1**). On November 17, 2003, the Rat Islands tsunami in Alaska provided the most comprehensive test for this forecast methodology. The magnitude-7.8 earthquake on the shelf near the Rat Islands generated a tsunami that was detected by three tsunameters located along the Aleutian Trench. It was the first tsunami detected by the newly developed tsunameter system. These real-time data combined with the model database were then used to produce the real-time model tsunami forecast (**Fig. 2**). For the first time, tsunami model predictions were obtained, during the tsunami propagation, before the waves had reached many coastlines. The initial offshore forecast was obtained immediately after preliminary earthquake parameters (location and magnitude 7.5) became available from the West Coast/Alaska Tsunami Warning Center (about 15–20 min after the earthquake). The model estimates provided expected tsunami time series at tsunameter locations. When the closest tsunameter recorded the first tsunami wave, about 80 min after the tsunami, the model predictions were compared with the deep-ocean data and the updated forecast was adjusted immediately.

These offshore model scenarios were then used

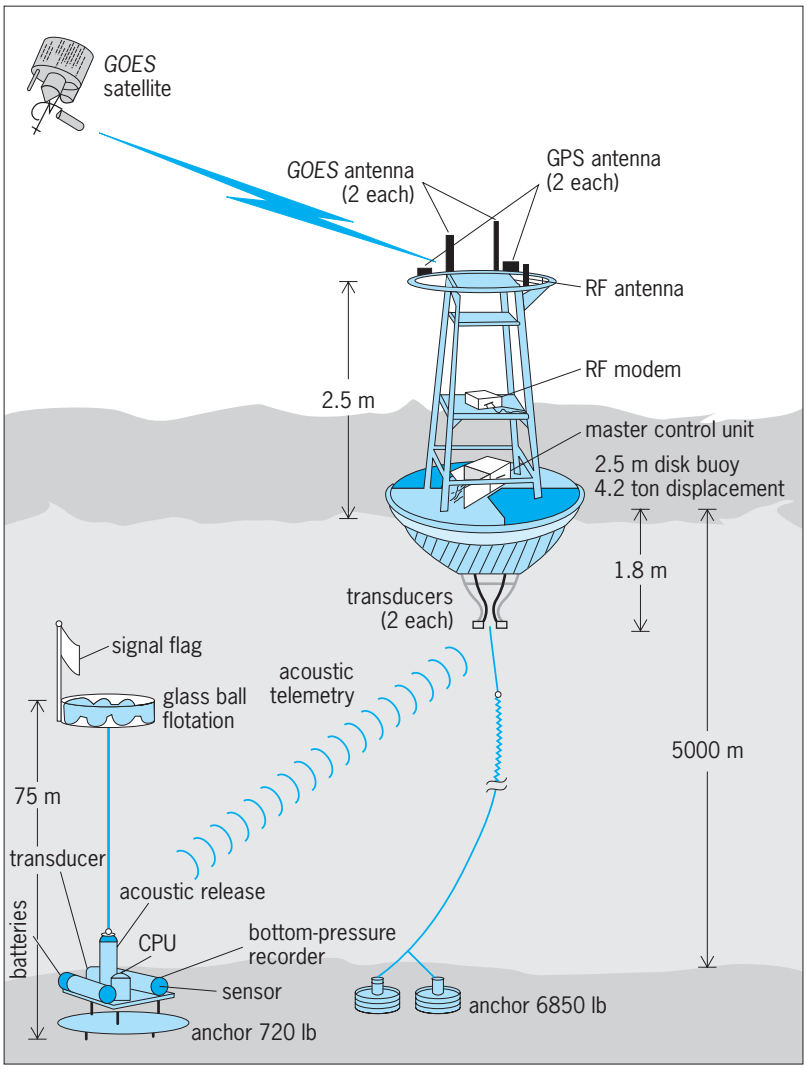

Fig. 1. Deep-ocean tsunami detection system.

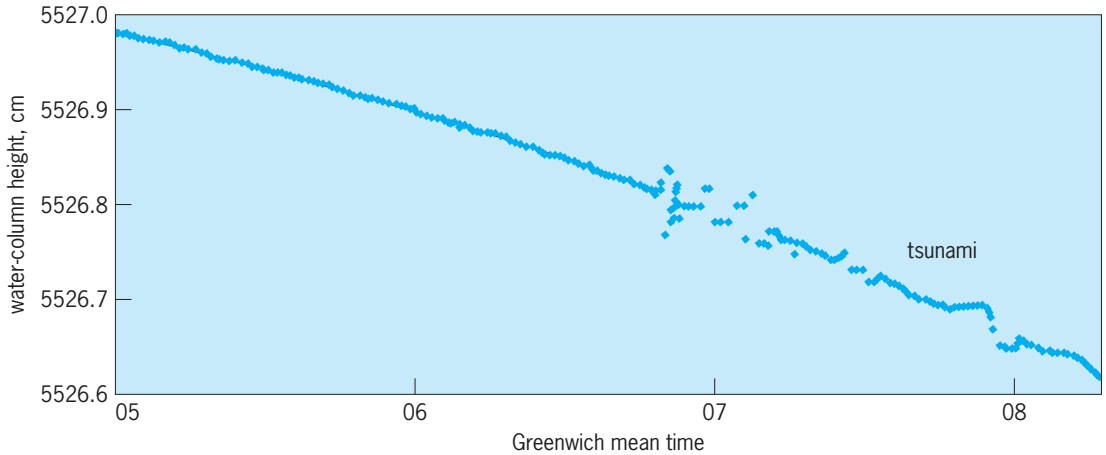

Fig. 2. Tsunami at Rat Islands, Alaska, on November 17, 2003, as measured at the tsunameter located at 50°N 171°W.

as input for the high-resolution inundation model for Hilo Bay, Hawaii. The model computed tsunami dynamics on several nested grids, with the highest spatial resolution of 30 m (100 ft) inside Hilo Bay. None of the tsunamis produced inundation at Hilo, but all of them recorded a nearly 0.5-m (1.6-ft) peak-to-trough signal at the Hilo tide gauge. Model forecast predictions for this tide gauge are compared with observed data in **Fig. 3**. The comparison demonstrates that amplitudes, arrival time, and periods of several first waves of the tsunami wave train were correctly forecasted. More tests are required to ensure that the inundation forecast will work for every tsunami likely to occur. When implemented, such a forecast will be obtained even faster, and would provide enough lead time for potential evacuation or warning cancellation for Hawaii and the west coast of the United States.

Reduction of impact. The recent development of real-time deep-ocean tsunami detectors and tsunami inundation models has given coastal communities the tools they need to reduce the impact of future tsunamis. If these tools are used in conjunction with a continuing educational program at the community level, at least 25% of the tsunami-related deaths might be averted. By contrasting the casualties from the 1993 Sea of Japan tsunami with that of the 1998 Papua New Guinea tsunami, we can conclude that these tools work. In the case of Aonae, Japan, about 15% of the population at risk died from a tsunami that struck within 10 minutes of the earthquake because the population was educated about tsunamis, evacuation plans had been developed, and a warning was issued. In the case of Warapa, Papua New Guinea, about 40% of the at risk population died from a tsunami that arrived within 15 min of the earthquake because the population was not educated, no evacuation plan was available, and no warning system existed.

Eddie N. Bernard

Bibliography. E. N. Bernard (ed.), *Developing Tsunami-Resilient Communities: The National Tsunami Hazard Mitigation Program* (NTHMP) [reprinted from *Nat. Hazards*, 35(1), 2005], Springer, The Netherlands, 2005; E. N. Bernard, The U.S. National Tsunami Hazard Mitigation Program: A successful state–federal partnership, *Nat. Hazards*, 35(1):5–24, NTHMP, 2005; F. I. González et al., The NTHMP tsunameter network, *Nat. Hazards*, 35(1):25–39, NTHMP, 2005; F. I. González et al., Progress in NTHMP hazard assessment, *Nat. Hazards*, 35(1):89–110, NTHMP, 2005; V. V. Titov et al., The global reach of the 26 December 2004 Sumatra tsunami, *Science*, 309(5743):2045–2048, 2005; V. V. Titov et al., Real-time tsunami forecasting: Challenges

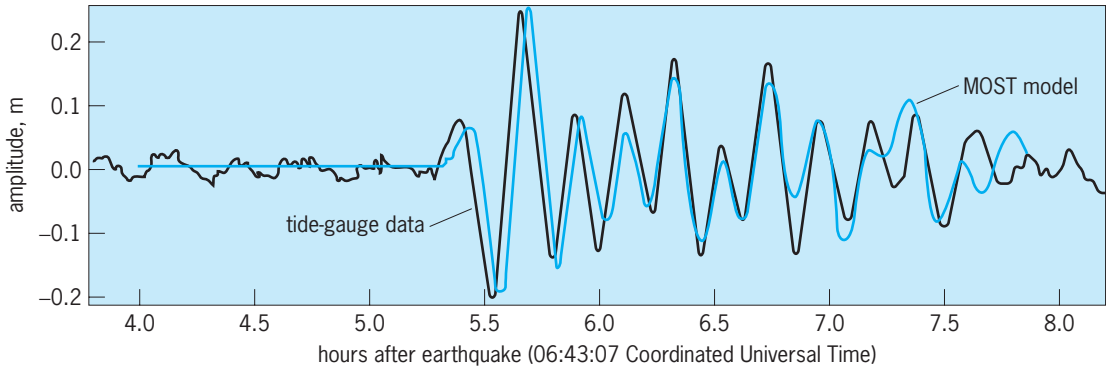

Fig. 3. Coastal forecast at Hilo, Hawaii, for 2003 Rat Islands tsunami, comparing the forecasted and measured gauge data.

and solutions, *Nat. Hazards*, 35(1):41–58, NTHMP, 2005.

Tuberculosis

An infectious disease caused by the bacillus *Mycobacterium tuberculosis*. It is primarily an infection of the lungs, but any organ system is susceptible, so its manifestations may be varied.

The tubercle bacillus was discovered by Robert Koch in 1882. Effective therapy and methods of control and prevention of tuberculosis have been developed, but the disease remains a major cause of mortality and morbidity throughout the world.

Tuberculosis is an acute health problem, particularly in the developing countries of Latin America, Africa, and Asia. The resurgence in tuberculosis between 1985 and 1993 was mainly the result of its occurrence in persons infected with human immunodeficiency virus (HIV) and with acquired immune deficiency syndrome (AIDS). In addition to individuals infected with HIV, there are other groups with a higher incidence of tuberculosis, including those having close contacts with infectious tubercular individuals; those with medical conditions such as silicosis, diabetes, end-stage renal disease, hematologic disease, and immunosuppressive therapy; intravenous drug users and alcoholics; and those in long-term care facilities, such as prisons and nursing homes. The treatment of tuberculosis has been complicated by the emergence of drug-resistant organisms, including multiple-drug-resistant tuberculosis, especially in those with HIV infection. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS).

Most new cases of clinical tuberculosis arise in individuals who have been previously infected. The eradication of tuberculosis and its prevention depend upon the detection and prophylactic treatment of infected individuals so that clinical disease does not occur, and effective treatment of those with symptomatic clinical disease so that transmission of tuberculosis to others is prevented.

Pathogenesis. *Mycobacterium tuberculosis* is transmitted by airborne droplet nuclei (1–10 micrometers in diameter) produced when an individual with active disease coughs, speaks, or sneezes. When inhaled, the droplet nuclei reach the alveoli of the lung. In susceptible individuals the organisms may then multiply and spread through lymphatics to the lymph nodes, and through the bloodstream to other sites such as the lung apices, bone marrow, kidneys, and meninges.

The development of acquired immunity in 2 to 10 weeks results in a halt to bacterial multiplication. Lesions heal and the individual remains asymptomatic. Such an individual is said to have tuberculous infection without disease, and will show a positive tuberculin test. The risk of developing active disease with clinical symptoms and positive cultures for the tubercle bacillus diminishes with time and may never occur, but is a lifelong risk.

Only close contacts such as family members or other sharing a closed environment are liable to become infected from a clinical case of tuberculous disease.

Only 5% of individuals with tuberculous infection progress to active disease. Progression occurs mainly in the first 2 years after infection; household contacts and the newly infected are thus at risk.

The classic pathologic lesion in tuberculosis is the granuloma or tubercle, the result of a cell-mediated hypersensitivity response. The tubercle consists of an aggregation of epithelioid cells including Langhans multinucleated giant cells, surrounded by a rim of monocytes, lymphocytes, and fibroblasts. Although the granuloma is not unique to tuberculosis, a distinctive feature in tuberculosis is a friable, cheeselike appearance in diseased tissue. This results from liquefaction of the epithelioid center of the tubercle.

Clinical manifestations. Many of the symptoms of tuberculosis, whether pulmonary disease or extrapulmonary disease, are nonspecific. Fatigue or tiredness, weight loss, fever, and loss of appetite may be present for months. A fever of unknown origin may be the sole indication of tuberculosis, or an individual may have an acute influenzalike illness. Erythema nodosum, a skin lesion, is occasionally associated with the disease.

The lung is the most common location for a focus of infection to flare into active disease with the acceleration of the growth of organisms. There may be complaints of cough, which can produce sputum containing mucus and pus. Though pulmonary tuberculosis is one of the major cases of blood in sputum, it is uncommon. Other chest complaints, such as shortness of breath or pain, are unusual. On examination, listening to the lungs may disclose rales or crackles and signs of pleural effusion (the escape of fluid into the lungs) or consolidation if present. In many, especially those with small infiltration, the physical examination of the chest reveals no abnormalities.

Miliary tuberculosis is a variant that results from the blood-borne dissemination of a great number of organisms resulting in the simultaneous seeding of many organ systems. The meninges, liver, bone marrow, spleen, and genitourinary system are usually involved. The term miliary refers to the lung lesions being the size of millet seeds (about 0.08 in. or 2 mm). These lung lesions are present bilaterally. Symptoms are variable and the diagnosis can be difficult.

Extrapulmonary tuberculosis is much less common than pulmonary disease. However, in individuals with AIDS, extrapulmonary tuberculosis predominates, particularly with lymph node involvement. Fluid in the lungs and lung lesions are other common manifestations of tuberculosis in AIDS. The lung is the portal of entry, and an extrapulmonary focus, seeded at the time of infection, breaks down with disease occurring.

Development of renal tuberculosis can result in symptoms of burning on urination, and blood and

white cells in the urine; or the individual may be asymptomatic. The symptoms of tuberculous meningitis are nonspecific, with acute or chronic fever, headache, irritability, and malaise.

A tuberculous pleural effusion can occur without obvious lung involvement. Fever and chest pain upon breathing are common symptoms.

Bone and joint involvement results in pain and fever at the joint site. The most common complaint is a chronic arthritis usually localized to one joint. Osteomyelitis is also usually present.

Pericardial inflammation with fluid accumulation (pericardial effusion) or constriction of the heart chambers secondary to pericardial scarring (constructive pericarditis) are two other forms of extrapulmonary disease.

Diagnosis. The principal methods of diagnosis for pulmonary tuberculosis are the tuberculin skin test, sputum smear and culture, and the chest x-ray. Culture and biopsy are important in making the diagnosis in extrapulmonary disease.

Tuberculin skin test. A positive tuberculin skin test defines tuberculous infection; it is the major and essential screening test. Some individuals with a positive test will have further studies done to rule out tuberculous disease; others will be selected for preventive drug therapy to thwart progression to disease.

Tuberculin is composed mainly of tuberculoprotein obtained from cultures of the tubercle bacillus. The Mantoux test is the standard tuberculin test. An intracutaneous injection of purified protein derivative tuberculin is performed, and the injection site is examined for reactivity (as manifested by induration) 48–72 h later. An individual with tuberculous infection develops a delayed hypersensitivity response to the tubercle bacillus, one manifestation of which is a positive reaction to the tuberculin skin test. A multiple-puncture technique is occasionally used for screening, but such a test is not as accurate as the Mantoux test, and any positive reaction has to be confirmed with a standard Mantoux test. Skin sensitivity is not affected by therapy and usually lasts for one's lifetime, though it may wane with age. *See* HYPERSENSITIVITY.

While a positive Mantoux test confirms tuberculous infection, a negative reaction does not necessarily mean tuberculosis is not present. A false negative reaction, or anergy, can occur with viral illness, AIDS, neoplastic disease, the use of immunosuppressive drug therapy, and fulminant disseminated tuberculosis, and in any condition affecting T-lymphocyte function.

Pulmonary disease. In those with symptoms suggesting clinical pulmonary disease, the chest x-ray film is essential since it reveals the extent and severity of disease. The four major radiographic abnormalities are segmental or lobar infiltration, pleural effusion, lymphadenopathy, and miliary densities. The most common pattern would be a nodular infiltrate with cavitation (see **illus.**). When an affected area heals, calcification, fibrosis, and volume loss may occur.

Activity or bacteriologic status cannot be determined from the chest film: while the film may be strongly suggestive of tuberculosis, other bacterial and fungal disease can give similar findings. Further diagnostic studies to demonstrate the organism via sputum smear and sputum culture are mandatory.

An individual with tuberculous infection without disease will have a negative chest x-ray, since the initial focus has healed and is not seen. However, there may be residual enlargement of some of the lymph nodes.

The best sputum sample for *M. tuberculosis* is a single early-morning specimen. At least three single morning specimens are obtained and sent for culture.

Extrapulmonary disease. The diagnostic approach to extrapulmonary disease depends upon the site involved, but usually specific tissue biopsy is required to identify the granulomas of tuberculosis.

The diagnosis of tuberculous pericarditis and peritonitis invariably requires biopsy. In addition to pathologic examination, all tissue specimens are cultured. The diagnosis of tuberculous meningitis is made by culture of cerebrospinal fluid. Synovial tissue (surrounding joints and connective tissue) biopsy and culture are done for tuberculosis arthritis. In genitourinary tuberculosis, urinalysis shows pus in the urine without bacteria the diagnosis depends on urine culture.

Treatment and control. The modern era of chemotherapy for tuberculosis began with the use of streptomycin in experimental tuberculosis in 1945. For the first time a drug was available that was remarkably effective against tuberculosis. Other agents followed: paraaminosalicylic acid in 1946, isoniazid and pyrazinamide in 1952, and ethambutol in 1961. By using a combination of agents for

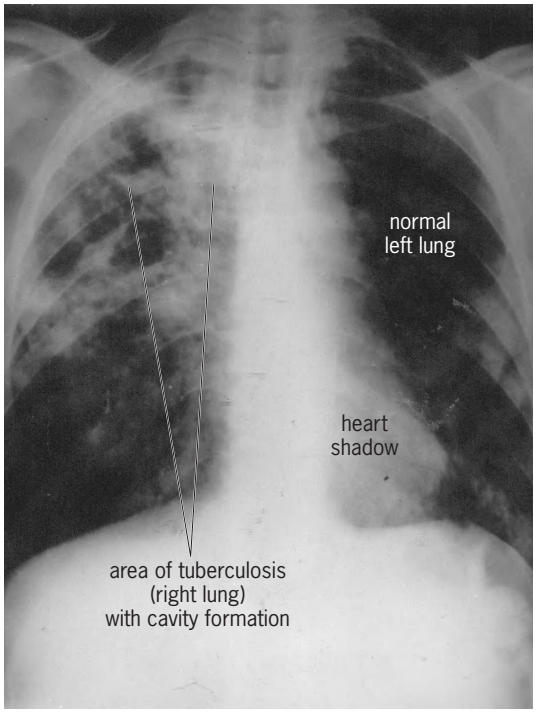

Chest radiograph: posteroanterior view showing extensive tuberculosis with cavitation in right lung.

prolonged periods, drug resistance is overcome and potentially all cases can be cured with very low relapse rates. In 1966 the most potent drug to date, rifampin, was introduced for clinical use, permitting shorter drug regimens. *See* ANTIBIOTIC; DRUG RESISTANCE.

Preventive therapy. Screening for tuberculosis relies upon the tuberculin skin test. Since mass screening of the general population is not cost-effective, screening is directed at those subpopulations who have a high risk of becoming infected or have an increased incidence of tuberculous disease. Such subpopulations might include those in correctional institutions or health care facilities. After the individual with tuberculous infection is identified and a negative chest x-ray rules out clinical tuberculous disease, the use of isoniazid therapy daily for 6 months to 1 year will prevent development of tuberculous disease.

Since isoniazid is hepatotoxic (toxic to the liver), it cannot be used in everyone with a significant tuberculin reaction, but only in those in whom the risk of progression to tuberculous disease is greater than the risk of hepatotoxicity, which can cause hepatitis. Hepatitis due to isoniazid occurs mainly in those over 35 years of age. For the individual over 35 with tuberculous infection, another risk factor must be present to justify isoniazid use.

Tuberculous disease therapy. A combination of two or more drugs is used in the initial therapy of tuberculous disease. Drug combinations are used to lessen the chance of drug-resistant organisms surviving. Valuable first-line drugs include isoniazid, rifampin, pyrazinamide, ethambutol, and streptomycin. Isoniazid and rifampin are bactericidal and especially useful. Since it is injectable, streptomycin is useful in supervised programs when patient compliance is in doubt.

Other agents with activity against *M. tuberculosis* include capreomycin, kanamycin, ethionamide, paraaminosalicylic acid, and cycloserine. These agents are considered second-line drugs, and they are used when drug resistance causes treatment failure.

The preferred treatment regimen for both pulmonary and extrapulmonary tuberculosis is a 6-month regimen of the antibiotics isoniazid, rifampin, and pyrazinamide given for 2 months, followed by isoniazid and rifampin for 4 months. Because of the problem of drug-resistant cases, ethambutol can be included in the initial regimen until the results of drug susceptibility studies are known.

Once treatment is started, improvement occurs in almost all individuals. In pulmonary tuberculosis, sputum smears are the best way to ascertain effectiveness of therapy, and smears usually become negative for acid-fast organisms in a few weeks to 3 months. In 3–5 months, cultures usually become negative. After 10–14 days of adequate therapy, the patient is considered noninfectious due to rapid reduction in the total number of organisms in the sputum. After a course of therapy has been successfully completed, the patient is considered cured.

Any treatment failure or individual relapse is usually due to drug-resistant organisms. Retreatment is guided by drug susceptibility testing. Primary drug resistance where drug-resistant organisms are present at the start of therapy once occurred in less than 3% of cases; but with disease spread from persons with drug-resistant organisms, primary resistance is becoming more common. Acquired or secondary drug resistance is resistance that develops during treatment. In any case, drug resistance is associated with individuals not taking medication correctly. *See* DRUG RESISTANCE.

The community control of tuberculosis depends on the reporting of all new suspected cases so case contacts can be evaluated and treated appropriately as indicated. Individual compliance with medication is essential. Furthermore, measures to enhance compliance, such as directly observed therapy, may be necessary. *See* MYCOBACTERIAL DISEASES.

George Lordi

Bibliography. American Thoracic Society, Control of tuberculosis in the United States, *Amer. Rev. Respir. Dis.*, 146:1623–1633, 1992; American Thoracic Society, Treatment of tuberculosis and tuberculosis infection in adults and children, *Amer. J. Respir. Crit. Care Med.*, 149:1359–1374, 1994; B. R. Bloom (ed.), *Tuberculosis: Pathogenesis, Protection and Control*, 1994; P. D. Davies (ed.), *Clinical Tuberculosis*, 1993.

Tubeworms

The name given to marine polychaete worms (particularly to many species in the family Serpulidae) which construct permanent calcareous tubes on rocks, seaweeds, dock pilings, and ship bottoms. The individual tubes with hard walls of calcite-aragonite, ranging from 0.04 to 0.4 in. (1 mm to 1 cm) in diameter and from 0.16 to 4 in. (4 mm to 10 cm) in length, are firmly cemented to any hard substrate and to each other.

Economically they are among the most important fouling organisms both on ship hulls (where they are second only to barnacles) and inside seawater cooling pipes of power stations. A moderate growth of tubeworms can add more than 12% to the fuel costs of a medium-sized freighter. Mass settlements of tubeworm species in the genera *Hydroides*, *Ficopomatus*, *Spirobranchus*, *Serpula*, and *Pomatoceros* can deposit massive layers of calcareous tubes (3.2–8 in. or 8–20 cm in thickness per year) on dock walls and other submerged marine installations. On such structures as the legs of drilling platforms, the accumulated weight can lead to major structural damage.

Tubeworms feed by filtering suspended material from the water, using a crown of ciliated pinnate tentacles. For this reason, life on a ship bottom or inside an intake pipe can provide an enhanced water flow and food supply, and thus more rapid growth. Sexual reproduction in tubeworms results in planktonic trochosphere larvae, which are the natural dispersal

stage. After a varying period of larval life, settling (which may be gregarious on suitable surfaces) is induced by a combination of chemical and physical stimuli. Initial tube secretion after settlement can be rapid (up to 0.4 in. or 1 cm per week), and a new generation of tubeworms can quickly overgrow and smother a slower-growing parental one.

About 340 valid species of serpulid tubeworms have been described. The majority are truly marine, but several species of *Ficopomatus* thrive in brackish waters of low salinity, and one species occurs in fresh waters in Karst limestone caves. The wide geographical distribution of certain abundant species owes much to human transport on the bottoms of relatively fast ships and occurred within the last 120 years. *See* ANNELIDA; POLYCHAETA.

W. D. Russell-Hunter

Tubulidentata

An order of mammals containing a single living genus *Orycteropus*, the aardvark. Aardvarks occur in suitable habitats throughout sub-Saharan Africa. This order exhibits the results of an extreme adaptation for burrowing and feeding on small food items (particularly termites and ants). *See* MAMMALIA.

Aardvarks, also known as antbears, resemble a medium-sized to large pig. The body is massive with an elongate head and a piglike snout. The tough thick skin is sparsely covered with bristly hair. The ears are large and donkeylike and can be moved independently of each other. The strong muscular tail is kangaroolike. The short thick legs possess powerful sharp claws that are used to excavate burrows and to open the nests of termites and ants which are gathered by the aardvark's long sticky tongue. Adult aardvarks lack incisor and canine teeth. The simple peglike teeth on the sides of the jaws consist of tubular dentin covered by cement ("tubule teeth"). They lack enamel and grow continuously during the animal's life. The dental formula is I 0/0 C 0/0 Pm 2/2 M 3/3 × 2 for a total of 20 teeth. Adult aardvarks have a head and body length of 1000–1580 mm (39–62 in.), a tail length of 443–710 mm (17–28 in.), and a shoulder height of 600–650 mm (23–25 in.). Most weigh 50–70 kg (110–153 lb). *See* AARDVARK.

In the past, tubulidentates were often considered closely related to ungulates. However, recent mitochondrial and nuclear gene data show a close relationship to elephant-shrews, paenungulates (hyraxes, sirenians, and proboscideans), and golden moles (Chrysochloridae). All of these ecologically divergent forms probably originated in Africa. Molecular evidence implies that they all may have arisen from a common ancestor that existed in the Cretaceous Period when Africa was isolated from other continents. Three genera of the family Orycteropodidae are known: *Leptorycteropus*, *Myorycteropus*, and *Orycteropus*. The earliest known tubulidentate (*Myorycteropus*) is from early Miocene deposits found in Kenya in East Africa. *Orycteropus gaudryi*, a species from the late Miocene epoch, is similar to *O. afer* except that the former had a greater number of cheekteeth. A relatively unspecialized form, *Leptorycteropus*, dates from the mid-Pliocene epoch. Pleistocene remains are known from France, Greece, India, and Turkey. A genus from Madagascar (*Plesiorycteropus*) may be related to them.

Donald W. Linzey

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, Freeman, 1998; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Tufa

A spongy, porous limestone formed by precipitation from evaporating spring and river waters; also known as calcareous sinter. Calcium carbonate commonly precipitates from supersaturated waters on the leaves and stems of plants growing around the springs and pools and preserves some of their plant structures (see **illus.**). Tufa tends to be fragile and friable. Tufa deposits are limited in extent and are found mainly in the youngest rocks, Pleistocene or Holocene. *See* LIMESTONE; TRAVERTINE.

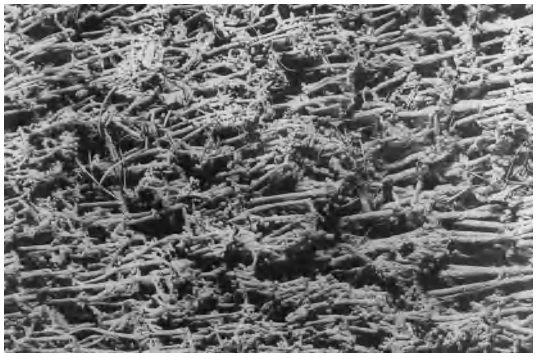

Calcareous tufa deposited on plant stems. (*Copyright © 1975 by Francis J. Pettijohn*)

Raymond Siever

Tuff

Fragmental volcanic products from explosive eruptions that are consolidated, cemented, or otherwise hardened to form solid rock. In strict scientific usage, the term "tuff" refers to consolidated volcanic ash, which by definition consists of fragments smaller than 2 mm. However, the term is also used for many pyroclastic rocks composed of fragments coarser than ash and even for pyroclastic material that has undergone limited posteruption reworking. If the thickness, temperature, and gas content of a tuff-forming pyroclastic flow are sufficiently high, the constituent fragments can become compacted and fused to form welded tuff. The term "tuff" is also used

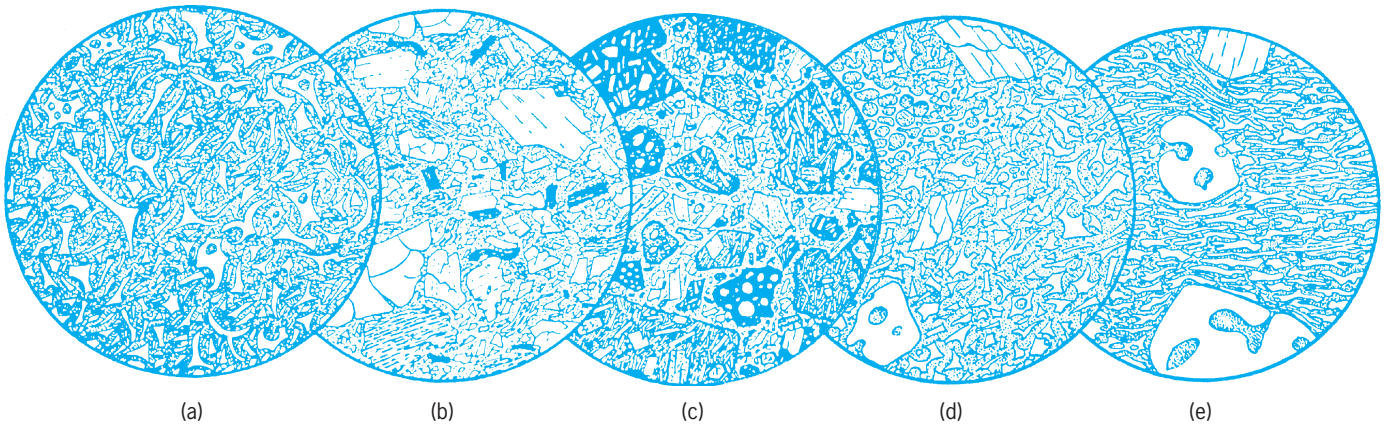

Fig. 1. Some common types of tuffs as viewed through the petrographic microscope. The field of view of the images is 2 mm for *a–c*, 3 mm for *d* and *e*. (*a*) Rhyolitic vitric tuff erupted from Mount Shasta Volcano, California, showing aggregates of glass fragments (shards) set in a matrix of very fine ash (volcanic dust). (*b*) Rhyolitic crystal tuff, Etsch Valley, Italy; broken crystals of quartz, plagioclase feldspar, and biotite set in a fine-grained matrix of glass and pumice fragments. (*c*) Andesitic lithic tuff, near Managua, Nicaragua, composed of andesitic rock fragments set in a finer-grained matrix of plagioclase and pyroxene and glass; there are no discrete large crystals. Each lithic fragment itself contains its own crystals set in a finer-grained matrix. (*d*) Tuff from the mostly unwelded top of the Bishop Tuff, an ignimbrite erupted from Long Valley Caldera, California; crystals of quartz and alkali feldspar are set in a matrix of undeformed glass shards and volcanic dust. (*e*) Welded tuff from the same locality as *d*, but the specimen was collected from the mostly welded interior part of the same ignimbrite; same constituents as in *d*, but here the glass shards are stretched, bent, and flattened by the welding (compaction) process. (*After H. Williams et al., Petrography, 2d ed., W. H. Freeman, 1982*)

in the naming of several related types of small volcanic edifices formed by hydrovolcanic eruptions, triggered by the explosive interaction of hot magma or lava with water. *See* IGNIMBRITE; PYROCLASTIC ROCKS.

As with their unconsolidated counterparts, the fragmental volcanic materials making up tuffs can vary widely in chemical and mineralogical composition, texture (crystalline or glassy), shape, and other properties. A common classification of the various types of tuff is based on the proportions of the constituent fragments: volcanic glass, crystals, and rock fragments (lithics). A tuff composed mostly of shards of glass is called vitric tuff, one containing mainly crystal is termed crystal tuff, and one in which rock fragments dominate is a lithic tuff (**Fig. 1**).

Tuffs are further distinguished by their approximate bulk chemical composition, as determined by the composition of the constituent fragments taken as a whole. Thus, depending on its bulk composition, a tuff can be more fully described as basaltic vitric tuff, andesitic crystal tuff, trachytic lithic tuff, rhyolitic vitric tuff, and the like. Vitric tuffs are generally characteristic of highly explosive and large-volume eruptions, and they can be deposited and preserved at hundreds of kilometers from their sources. Because of higher viscosity and gas content, these voluminous eruptions generally involve rhyolitic magmas; less commonly, intermediate composition magmas (andesitic, trachytic, dacitic); and, rarely, basaltic magmas. *See* IGNEOUS ROCKS; LAVA; MAGMA; VOLCANO.

Glass fragments (or shards) in tuffs are highly irregular in shape because of the violent shredding and sudden cooling (quenching) of gas-charged liquid lava when it is explosively ejected into the atmosphere. Typically, the shards are bounded by concave surfaces that represent the walls of ruptured gas bubbles (vesicles) [**Fig. 2**]. Larger glassy fragments, such as chunks of pumice not totally explosively disrupted, may contain numerous vesicles that have been preserved intact; a diagnostic feature of welded tuffs is the flattening and elongation of the vitric components. "Fresh" glass shards (that is, those unaltered by posteruption processes) are generally clear and colorless, but they also can have brownish, yellowish, or pinkish tones depending on chemical composition and trace impurities. In some samples, the glass is clouded by the suspension of tiny crystals (microlites or crystallites) of iron-titanium oxide or other dark-colored minerals. The larger crystals contained in tuffs are rarely preserved in their entirety; instead, they are broken into pieces, shattered by the force of explosive eruption. Rock fragments in tuffs are most commonly derived from solidified volcanic

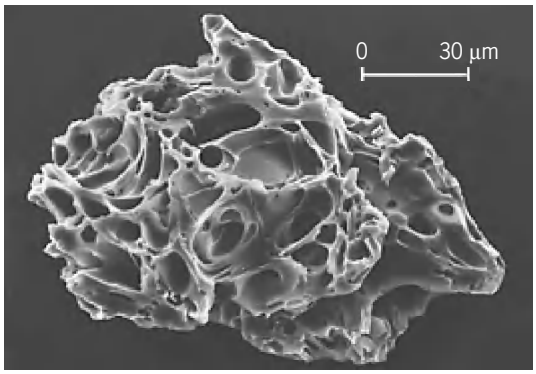

Fig. 2. Scanning electron microscope (SEM) image of a single ash particle (vitric fragment) from the May 18, 1980, eruption of Mount St. Helens, Washington. The tiny voids (vesicles) were created by expanding bubbles of volcanic gas during the rise and eruption of magma. If this fragment were further explosively shattered, the resulting tinier fragments (glass shards) would represent pieces of vesicle walls. (*Image by A. M. Sarna-Wojcicki, USGS*)

rocks erupted earlier from the same volcanic system (cognate pyroclasts), but they can also include small amounts of solid fragments unrelated to the volcano (accidental pyroclasts).

The vitric fragments of a tuff are readily altered by postdepositional processes, initially by hydration (chemical reaction with water) because of the percolation of hydrothermal fluids and ground water. For example, brown-colored glass of basaltic composition, upon hydration and weathering, is converted to yellowish or orange-colored palagonite; a pyroclastic rock composed mostly of hydrothermally altered or weathered basaltic glass is called palagonite tuff. Hydrated glassy material then begins to devitrify, that is, to crystallize to form extremely fine-grained aggregates of silica (colloidal and crystalline), feldspar, and clay minerals. Glassy fragments of rhyolitic composition are commonly altered to clay minerals of the montmorillonite (smectite) group to form a porous, light-colored rock called bentonite, widely found in volcanic regions. The postdepositional breakdown of crystals in tuffs is similar to that of the mineral constituents in most igneous rocks. The erosion and reworking of tuff by wind and water create sedimentary materials that are ultimately redeposited on land or in water, along with sediments of nonvolcanic origin, to form widespread submarine or subaerial volcaniclastic deposits. The term "tuffaceous" is used to describe sedimentary deposits which, while composed mostly of nonvolcanic fragments, contain an appreciable and recognizable component of ash-size pyroclasts. *See* CLAY MINERALS; SEDIMENTARY ROCKS.

In many volcanic regions of Italy, Mexico, and other countries, tuff is the common and preferred building stone. Many tuffs are colorful and visually attractive, soft enough to be quarried and shaped by hand and yet with sufficient structural strength to be set into walls with mortar. Dwellings and structures, including churches, were carved into thick, massive ignimbrite deposits by the inhabitants of the Cappadocia volcanic region (Anatolia, central Turkey) as early as mid-third century B.C. Robert Tilling

Bibliography. R. A. F. Cas and J. V. Wright, *Volcanic Successions: Modern and Ancient*, Chapman & Hall, London, 1988; R. V. Fisher and H.-U. Schmincke, *Pyroclastic Rocks*, Springer-Verlag, Berlin, 1984; G. Heiken, *An Atlas of Volcanic Ash: Smithsonian Contributions to the Earth Sciences*, Smithsonian Institution, Washington, DC, 1974; H. Williams, F. J. Turner, and C. M. Gilbert, *Petrography: An Introduction to the Study of Rocks in Thin Sections*, 2d ed., W. H. Freeman, San Francisco, 1982.

Tularemia

A worldwide disease caused by infection with the bacterium *Francisella tularensis*, which affects multiple animal species, including humans. Tularemia cases have been reported from North America, Europe, Russia, China, and Japan. In the winter, infection in adult humans occurs frequently from skinning infected rabbits, hares, muskrats, or beavers bare-handed; in the summer, infections occur in both adults and children from transmission of the bacteria through the bites of ticks or deer flies.

Clinical signs. Clinical signs typically occur 1–14 days after exposure. Tularemia can be difficult to differentiate from other diseases because it can have multiple clinical manifestations. Nonspecific signs frequently include fever, lethargy, anorexia, and increased pulse and respiration rates. The disease can overlap geographically with plague, and both may lead to enlarged lymph nodes (buboes). However, with tularemia, the buboes are more likely to ulcerate. If tularemic infection results from inhalation of dust from contaminated soil, hay, or grain, either pneumonia or a typhoidal syndrome can occur. Rarely, the route of entry for the bacteria is the eyes, leading to the oculoglandular type of tularemia. If organisms are ingested from soil, water, or contaminated wildlife, the oropharyngeal form can develop, characterized by abdominal pain, diarrhea, vomiting, and ulcers. *See* PLAGUE.

The mortality rate varies by species, although with treatment it is low. Ungulates are frequently infected but suffer low mortality from uncomplicated infections, except for documentation of high mortality in sheep on the range (up to 15% in untreated lambs), related to heavy tick infestations.

Tularemia is not transmitted directly from person to person. If the infected person or animal is untreated, blood remains infectious for 2 weeks and ulcerated lesions are infectious for a month. Deer flies (*Chrysops discalis*) are infective for 2 weeks, and ticks are infective throughout their lifetime (usually 2 years). Rabbit meat is infective even after being frozen for 3 years.

Ticks most frequently implicated in disease transmission include wood ticks (*Dermacentor andersoni*), dog ticks (*D. variabilis*), and Lone Star ticks (*Amblyomma americanum*). Deer flies (*C. discalis*) may also spread the organism. In Sweden, the implicated mosquito species is *Aedes cinereus*. Humans can be exposed through the bites of infected coyotes, squirrels, skunks, hogs, cats, or dogs.

Diagnosis and treatment. Confirmation of infection is usually made with a fourfold rise in specific antibody titers. Other tests, such as fluorescent antibody and culture, utilizing special media, may also be used. The liver, spleen, and lymph nodes can be enlarged, with whitish foci of necrosis.

A number of antibacterial agents are effective against *F. tularensis*, the most effective being streptomycin. Gentamicin and tobramycin may be effective; and the tetracyclines and chloramphenicol can be utilized, but relapses occur, and thus treatment must continue until the temperature has been normal for 4–5 days. Penicillin and the sulfonamides have no therapeutic effect. *See* ANTIBIOTIC. Millicent Eidson

Bibliography. A. S. Benenson (ed.), *Control of Communicable Diseases Manual*, 1995; T. Morner, The ecology of tularaemia, *Rev. Sci. Technol.*, 11:1123–1130, 1992.

Tulip tree

A tree, *Liriodendron tulipifera*, also known in forestry as yellow poplar, belonging to the magnolia family, Magnoliaceae. One of the largest and most valuable hardwoods of eastern North America, it is native from southern New England and New York westward to southern Michigan, and south to Louisiana and northern Florida. In rich, moist soil it may grow 150 ft (45 m) tall and have a diameter of 8–10 ft (2.4–3 m). *See* MAGNOLIALES.

This tree is distinguished by leaves which are squarish at the tip as if cut off, true terminal buds flattened and covered by two valvate scales, stipular scars encircling the twig, an aromatic odor resembling that of magnolia, chambered white pith, and cone-shaped fruit which is persistent in winter

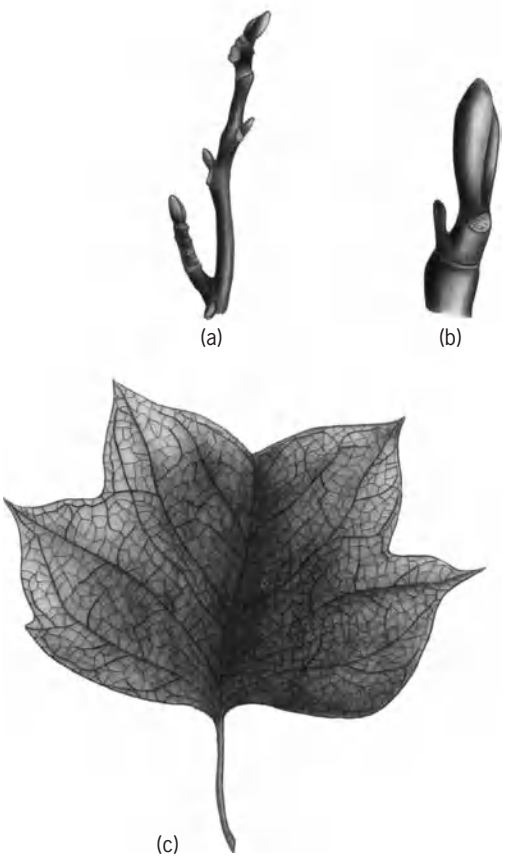

Tulip tree (*Liriodendron tulipifera*). (*a*) Twig, (*b*) terminal bud, and (*c*) leaf.

(see **illus.**). The name tulip refers to the large greenish-yellow and orange-colored flowers.

The wood of the tulip tree is light yellow to brown, hence the common name yellow poplar, which is a misnomer. It is a soft and easily worked wood, used for construction, interior finish, containers (boxes, crates, baskets), woodenware, excelsior, veneer, and sometimes for paper pulp. Because of its wide natural dimensions, the tulip tree often yields lumber as wide as 60 in. (150 cm) which is valuable for certain articles of furniture. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth P. Davis

Tumbling mill

A grinding and pulverizing machine consisting of a shell or drum rotating on a horizontal axis. The material to be reduced in size is fed into one end of the mill. The mill is also charged with grinding material such as iron balls. As the mill rotates, the material and grinding balls tumble against each other, the material being broken chiefly by attrition.

Tumbling mills are variously classified as pebble, ball, or rod depending on the grinding material, and as cylindrical, conical, or tube depending on the shell shape. *See* CRUSHING AND PULVERIZING; GRINDING MILL; PEBBLE MILL.

Ralph M. Hardgrove

Tumor

Literally, a swelling; in the past the term has been used in reference to any swelling of the body, no matter what the cause. Thus a swollen region produced by edema, congestion, or hemorrhage into a tissue has been called a tumor. However, the word is now being used almost exclusively to refer to a neoplastic mass, and the more general usage is being discarded.

Neoplasm. A neoplastic mass or neoplasm is a pathological lesion characterized by the progressive or uncontrolled proliferation of cells. The cells involved in the neoplastic growth have an intrinsic heritable abnormality such that they are not regulated properly by normal methods. The stimulus which elicits this growth is not usually known. The cellular proliferation serves no useful function and often is very detrimental. Most cells in the body can undergo neoplastic changes, and hence there are many kinds of neoplasms. All classes of vertebrates have members which have developed neoplasms. These growths also have been seen in some invertebrates and plants.

Tumors are composed of two basic components: the parenchyma, which consists of the neoplastic proliferating cells, and the stroma, which is the supporting framework of connective tissue that includes the vascular supply. The stroma is derived from normal tissue, and the amount within a tumor varies greatly. It is the parenchyma which determines the biological behavior of the neoplasms.

Benign and malignant tumors. It is common to divide tumors into benign or malignant. The decision as to which category a tumor should be assigned is usually based on information gained from gross or microscopic examination, or both. Benign neoplasms usually grow slowly, remain so localized that they cause little harm, and generally can be successfully and permanently removed. Malignant or cancerous neoplasms tend to grow rapidly, spread throughout the body, and recur if removed.

Degree of harm. Not all tumors which have been classified as benign are harmless to the host, and some can cause serious problems. Difficulties may occur as a result of mechanical pressure. As the mass of cells increases in size, it may press against another

structure. In this way a blood vessel or duct may be occluded or a vital organ compressed. Benign tumors of glandular tissue can induce illness as a consequence of overproduction of certain hormones. Clinically there may be very abnormal signs. In contrast, some malignant neoplasms are relatively harmless in that they are slow-growing and often can be successfully removed. Examples of tumors with these characteristics are certain skin cancers.

Growth patterns. Neoplasms exhibit a wide range of abnormal growth patterns. It is not always easy to decide if a given tumor should be classified as benign or malignant. Nor is it known if certain benign tumors can, on occasion, progressively change and become malignant. The early phases of many malignancies greatly resemble benign growths.

Characteristics. There are a number of characteristics which are used to differentiate benign from malignant tumors. It is important to remember that these are generalities and that it is possible to find exceptions for some of the characteristics given.

The cells of benign tumors are well differentiated. This means that the cells are very like the normal tissue in size, structure, and spatial relationship. The cells forming the tumor usually function normally. Cell proliferation usually is slow enough so that there is not a large number of immature cells. Also, because growth is relatively slow, the stroma proliferation keeps pace with that of the parenchyma, and hemorrhage and ischemic necrosis are not common. As the cellular mass increases in size, most benign tumors develop a fibrous capsule around them which separates them from the normal tissue. The cells of a benign tumor remain at the site of origin and do not spread throughout the body. Anaplasia (loss of differentiation) is not seen in benign tumors.

The cells of malignant tumors may be well differentiated, but most have some degree of anaplasia. Anaplastic cells tend to be larger than normal and are abnormal, even bizarre, in shape. The nuclei tend to be very large, and irregular, and they often stain darkly. Mitoses are seen frequently, and necrosis and hemorrhage are common. Malignant tumors may be partially but never completely encapsulated. The cells of the cancer infiltrate and destroy surrounding tissue. They have the ability to metastasize; that is, cells from the primary tumor are disseminated to other regions of the body where they are able to produce secondary tumors called metastases. In summary, the characteristics which are most important in separating malignant tumors from benign growths are their anaplasia, invasiveness, and ability to metastasize.

In most cases the formation of a neoplasm is irreversible. It results from a permanent cellular defect which is passed on to daughter cells. Tumors should undergo medical appraisal to determine what treatment, if any, is needed. N. Karle Mottet; Carol Quaife

Tumor suppressor genes. Tumor suppressor genes are a class of genes which, when mutated, predispose an individual to cancer. The mutations result in the loss of function of the particular tumor suppressor protein encoded by the gene. Although this class of genes was named for its link to human cancer, it is now clear that these genes play a critical role in the normal development, growth, and proliferation of cells and organs within the human body. The protein product of many tumor suppressor genes constrains cell growth and proliferation so that these events occur in a controlled manner. Thus, these genes appear to act in a manner antagonistic to that of oncogenes, which promote cell growth and proliferation.

Currently, the retinoblastoma (RB) p53 and p16 genes are the best-understood tumor suppressors (see **table**). Inactivating mutations in the RB gene have been observed in retinoblastomas, osteosarcomas (cancer of the bone), as well as cancers of the lung, breast, and bladder. The RB gene regulates proliferation, growth, and replication of cells. It also regulates specific transcription factors which activate genes whose protein products are involved in cell cycle progression (the cycle of growth and division of cells). Cell proliferation and division occurs in response to external cellular signals (such as growth factors). The RB gene functions by coordinating the expression of a number of genes which play a significant role in cell proliferation.

Tumor suppressor genes

Syndrome (or tumor)	Name of gene	Tumor type
Retinoblastoma	RB	Retinoblastoma, small cell lung carcinomas, osteosarcomas
Li-Fraumeni	p53	Sarcomas, breast carcinomas, brain tumors
Familial adenomatous polyposis	APC	Colon cancer
Neurofibromatosis I	NF1	Neurofibromas
Neurofibromatosis II	NF2	Schwannomas, meningiomas
Wilms' tumor	WT-1	Nephroblastoma
Von Hippel-Lindau	VHL	Renal cell carcinoma
Familial breast cancer	BRCA1	Breast, ovary
Colorectal cancer	DCC	Colon cancer
Familial melanoma	p16	Melanoma, many others
Hereditary nonpolyposis colon cancer	hMSH2	Colon cancer
	hMLH1	Colon cancer
	hPMS1,2	Colon cancer

The p16 mutations have been observed in cancers of the skin, lung, breast, brain, bone, bladder, kidney, esophagus, and pancreas. By regulating key factors involved in cell cycle progression, the p16 gene functions to modulate the activity of the RB gene. RB gene function is modulated during the cell cycle by phosphorylation (covalent addition of phosphate). The enzymes responsible for the phosphorylation of the RB gene are called cyclin-dependent kinases (kinases are enzymes which enzymatically add phosphate to their substrates; cyclins play a key role in activating cyclin-dependent kinases). Phosphorylation of the RB gene inactivates its growth suppressive function. The p16 gene specifically inhibits the cyclin-dependent kinases responsible for the phosphorylation of the RB gene. Thus, p16 functions by keeping RB in its growth-suppressive form. A delicate balance between the activity of cyclin-dependent kinases and the tumor suppressor p16 allows for controlled cell cycle progression by the RB gene. Loss of p16 prevents the RB gene from controlling cell proliferation and growth.

The tumor suppressor p53 is the most frequently mutated gene associated with the development of many different types of human cancer, including those of the breast, lung, and colon. It is also associated with the rare inherited disease, Li-Fraumeni syndrome. Affected individuals manifest an increased likelihood of breast carcinomas (invasive cancers of epithelial origin), soft tissue sarcoma (cancers of the connective tissue), brain tumors, osteosarcoma, leukemia, and adrenocortical carcinoma. Like RB and p16, p53 has a role in cell cycle regulation. In response to DNA damage, p53 can cause cells to stop replicating, thus allowing time for DNA repair. In addition, p53 functions in the cell's decision on whether to undergo programmed cell death (apoptosis), a process that plays a critical role in the normal development and functioning of many organs in the human body. Deregulated cell proliferation and escape from apoptosis appear to be two common pathways leading to tumor formation. The p53 gene functions to provide a regulated balance between these two processes. This gene is able to bind to DNA in a specific manner and activate the transcription (production of mRNA) of a number of genes in order to regulate cell proliferation. The p53 gene can induce the expression of p21, which like p16 is an inhibitor of cyclin-dependent kinases involved in cell proliferation. By inducing the expression of p21, p53 is able to cause cells to stop proliferating. There is indirect evidence that p53 may have biochemical functions apart from its ability to bind DNA. These other functions are likely to play a significant role in the ability of p53 to control apoptosis and cancer.

The normal counterparts of tumor suppressors and oncogenes collaborate with each other to ensure appropriate growth and cell proliferation within the context of the body. It has been suggested that when these genes are mutated they work together to allow tumor growth, consistent with the observation that the accumulation of multiple loss-of-function mutations in tumor suppressor genes and gain of function mutations in oncogenes within a single cell gives rise to cancer. *See* CANCER (MEDICINE); GENE; MUTATION; ONCOLOGY; TUMOR VIRUSES.

Mark Ewen

Bibliography. M. E. Ewen, The cell cycle and the retinoblastoma protein family, *Cancer Metastasis Rev.*, 13:45–66, 1994; E. R. Fearon and B. Vogelstein, A genetic model for colorectal tumorigenesis, *Cell*, 61:759–767, 1990; S. H. Friend et al., A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma, *Nature*, 323:643–646, 1986; A. J. Levine, The tumor suppressor genes, *Annu. Rev. Biochem.*, 62:623–651, 1993; R. A. Weinberg, Finding the anti-oncogene, *Sci. Amer.*, 259:44–51, 1988.

Tumor suppressor genes

A functionally heterogenous class of genes, ranging from nuclear transcription factors to cell adhesion molecules having as their common denominator that their function must be compromised during development and progression of a given tumor. The most important and most widely studied tumor suppressor is called p53. Its importance becomes evident from the fact that mutations in its gene are found in approximately 50% of all human cancers, thus constituting the most frequent alteration in a single human cancer-associated gene. The prevalence of p53 mutations in human cancer has spurred research worldwide to decipher p53 functions. It is now clear that the major function of p53 is to preserve the integrity of the genome of a cell under various conditions of cellular stress, endowing p53 with the title "guardian of the genome," coined by David Lane, one of the discoverers of p53 in 1979.

Tumorigenesis is a very complex process that is initiated by the accumulation of genetic alterations in at least two classes of genes: Activation of proto-oncogenes to oncogenes drives tumor initiation and progression by furthering uncontrolled cell growth and division, while inactivation of tumor suppressor genes leads to the elimination of important checkpoints in cellular proliferation that normally prevent uncontrolled proliferation. This implies that the products of tumor suppressor genes, the tumor suppressor proteins, exert functions that are disadvantageous to tumor development and progression. Inactivation of tumor suppressor genes usually occurs through mutations but can also occur by other means, like gene silencing by epigenetic means, for example, promoter inactivation through DNA methylation. Although both activation of proto-oncogenes and inactivation of tumor suppressor genes normally are the result of mutations as a consequence of manifold endogenous and exogenous DNA-damaging events, there is an important difference: as activation of a proto-oncogene to an oncogene leads to a "gain of function," already the mutational alteration of one of the two proto-oncogene alleles inherited from father or mother suffices to achieve an oncogenic

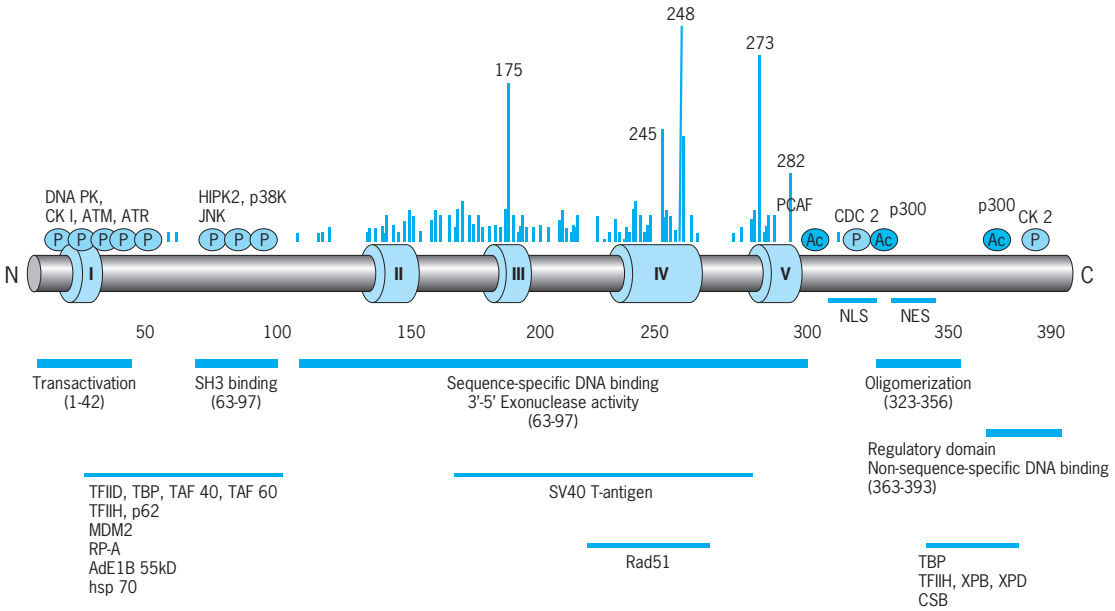

Landmarks for p53. Roman numerals represent the five regions that are conserved within p53 from all vertebrates. Known phosphorylation (P) and acetylation (Ac) sites are indicated. The vertical bars, clustered in the center of the p53 molecule, indicate amino acid residues mutated in human tumors (hot spots are identified by amino acid number). Below the molecule, horizontal bars indicate the current information concerning various domains of p53 for biological activities, p53 DNA interactions, and p53–protein complex formation; respective amino acids are given in brackets. Horizontal bars farther below mark interaction areas for various cellular and viral proteins interacting with p53. ATM indicates ATM kinase; ATR, ATR kinase; CK, casein kinase; CSB, Cockayne's syndrome B protein; DNA PK, DNA-dependent protein kinase; HIPK2, HIP kinase 2; JNK, Jun kinase; MDM2, mouse double minute protein 2; NES, nuclear export signal; NLS, main nuclear localization signal; RP-A, replication protein A; SV 40, Simian Virus 40; TAF, transcription-activating factor; TBP, TATA-Box binding protein; TF, transcription factor; XPB, xeroderma pigmentosum B protein; XPD, xeroderma pigmentosum D protein.

effect. As mutations in a tumor suppressor gene usually are characterized by a "loss of function," mutational inactivation of one allele of a tumor suppressor gene is not sufficient to eliminate the protective effect of the tumor suppressor, as the remaining nonmutated allele is still active. Inactivation of a tumor suppressor gene thus always requires that both of its two alleles are mutated.

Multifunctionality of p53. The protein p53 is truly multifunctional. It consists of several domains exhibiting a number of biochemical activities and interacting with a large variety of cellular and viral proteins (see **illustration**). The p53 core domain (amino acids 102–292) harbors at least three biochemical activities: (1) It mediates sequence-specific DNA binding to p53 consensus DNA sequences. Together with the bipartite *N*-terminal transactivation domain, this function enables p53 to act as a sequence-specific transcriptional activator. (2) The p53 core domain also recognizes complex non-B DNA structures, in both a sequence-specific and non-sequence-specific fashion. Recognition of structural features in nucleic acids also extends to the recognition of structural motifs in RNA, forming the molecular basis for p53 being a regulator of mRNA translation. (3) In addition, the p53 core domain exhibits an intrinsic 3′–5′ exonuclease activity (removes nucleotides from the ends of DNA).

As the p53 core domain is the major target for mutational inactivation of the p53 gene (see illustration), its multiple activities obviously are relevant to the tumor suppressor functions of p53. The C-terminal domain of p53 (amino acids 323–393), separated from the core domain by a flexible hinge region, harbors the p53 nuclear localization signals, an oligomerization domain, and a basic domain, which is able to bind non-sequence-specifically to DNA and RNA, and to exert an RNA and DNA reannealing activity. The basic C-terminal end also is the major regulatory domain of p53, affecting its DNA binding and p53 exonuclease activities.

All major domains of p53, but especially the p53 N- and C-termini, can interact with a large number of cellular and viral proteins. These proteins are partners of p53 in various p53 functions and regulate either positively (support) or negatively (inactivate), p53 functions. The physiologically most relevant cellular protein in this respect is MDM2, which binds to the p53 N-terminus and inhibits the transactivation function of p53, and it also targets p53 to proteasomal degradation. As the *mdm2* gene is a transcriptional target for the p53 protein, MDM2 and p53 are connected via a negative autoregulatory feedback loop.

Genomic integrity. The best-analyzed tumor suppressor function of p53 is that of a sequence-specific transcriptional activator (transactivator) of cellular genes which are involved in cell-cycle checkpoint control or apoptosis. In the absence of cellular stress, p53 is rapidly turned over and is present in normal cells only in small copy numbers of 1000–10,000 molecules. The high turnover is due to the interaction of p53 with MDM2. Various cellular stress signals, most importantly DNA damage, activate different signaling cascades, which lead to phosphorylation of p53 at N-terminal and C-terminal

phosphorylation sites. Phosphorylation of p53 first disrupts the p53–MDM2 interaction, resulting in metabolic stabilization of p53. Then p53 can act as a sequence-specific transactivator, leading to the upregulation of genes encoding cell-cycle inhibitors, like $p21^{waf1}$ or 14.3.3σ, which induce growth arrest in the G1 or G2 phase, respectively, or of repair-associated genes, like *gadd45*. The p53-induced growth arrest allows time to repair the inflicted DNA damage before it can be fixed as a mutation during the next round of DNA replication. Concerning tumor suppression, the most important p53 response is the induction of apoptosis in damaged cells, leading to their elimination. Consequently, p53 also activates a number of genes involved in apoptosis, like the *bax*, *fas-R*, *PUMA*, and several redox-related *PIG* genes. However, p53-mediated apoptosis is further complicated by the finding that induction of apoptosis by p53 involves nontranscriptional pathways in addition to transcriptional ones.

The p53 interacts with a variety of repair-associated proteins, and accumulating evidence strongly suggests that p53 is directly involved in repair processes and in the control of their fidelity. The idea of a direct involvement of p53 in repair processes is also strongly supported by the discovery that p53 exhibits an intrinsic 3′-5′ exonuclease activity. Exonucleases are required for nearly all processes of DNA metabolism, such as DNA replication, long-patch DNA repair, postreplicative mismatch repair, and DNA recombination. The p53 exonuclease activity thus strongly expands p53's possibilities as a guardian of the genome. So far, the exonuclease activity of p53 has been demonstrated to play a role in the fidelity control of homologous DNA recombination, where it removes mismatches in heteroduplex joints. Whether it also plays a role in DNA proofreading, for example, in DNA synthesized by the proofreader defective DNA polymerases α and β, is still a matter of debate. *See* CELL CYCLE (CANCER).

The p53 not only acts as a transactivator but also represses the expression of a large number of genes. In fact, in genewide microarray screens, more genes seem to be repressed than activated by p53. Although much less is known about the transrepression activity of p53 compared to its transactivation function, it is clear that transrepression of genes by p53 is even more complex than transactivation. Most likely, transrepression is not directly mediated by sequence-specific binding, but requires the interaction of p53 with other transcription factors, for example, with the Ets-1 protein, a transcription factor that activates a large number of genes that are involved in cell growth and invasion. By binding to Ets-1, p53 blocks the transcriptional activity of Ets-1, thereby exerting a tumor-suppressive function. Also in transrepression, the interaction of p53 with corepressors, like the Sin3a protein, might promote a local "closing" of the chromatin structure, thereby preventing transactivation.

Mutant p53. More than 80% of all mutations in the p53 gene are missense point mutations in the coding region of the p53 core domain, leading to the expression of a full-length p53 protein with a single amino acid exchange. Such a mutational spectrum is quite unusual for a tumor suppressor, as most tumor suppressors are inactivated by gene truncation or deletion or by promoter inactivation. The unusual mutational spectrum implies a selection advantage for tumor cells expressing such a mutant p53, and thus that mutant p53 exerts oncogenic properties. The oncogenic properties of mutant p53 are not yet understood at the molecular level. However, evidence obtained from the analyses of tumor banks strongly supports the concept that at least certain point mutations within the p53 gene not only serve to eliminate the tumor suppressor functions of p53 but enhance the aggressiveness of the respective tumors, leading to a poorer prognosis for tumor patients. In this regard, an important characteristic of at least most epithelial cancers is that mutations in the p53 gene are late events during their development, often associated with the onset of invasiveness. The p53 mutations thus are associated with tumor progression rather than with tumor initiation. The genetic instability of tumor cells expressing mutant p53 will allow the selection of more aggressive tumor cell variants and of tumor cells resisting tumor therapy. In addition, mutant p53 is suspected to actively suppress the responsiveness of tumor cells to various anticancer treatments.

Prognosis and therapy. Like no other tumor suppressor, the central role of p53 at the crossroad of cell cycle control, DNA repair, and apoptosis, as well as the fact that p53 is mutated in a large number of tumors, renders p53 a prominent target for applied tumor research. In contrast to wild-type p53, which is usually an unstable protein, mutant p53 accumulates in tumor cells. Detection of enhanced levels of p53 in tumor tissue is indicative for the expression of a mutant p53 and is already commonly used in tumor diagnosis. The value of mutant p53 as a diagnostic and prognostic marker will be enhanced by sequence analyses, allowing the classification of the aggressiveness of the respective mutant p53 as derived from the analyses of tumor banks.

The most challenging aspect of applied p53 research is the development of therapy regimens based on the p53 status of the tumor cells.

Gene therapy approaches. Wild-type p53 is able to induce apoptosis under appropriate conditions, whereas mutant p53 has lost this activity and even seems to actively block at least certain apoptotic pathways. Therefore, gene replacement, that is, introduction of wild-type p53 into tumors by various vector systems, is one of the aims of a p53-based gene therapy. In addition to the problems common to all gene therapy approaches, such as tumor-specific delivery of the therapeutic gene and its high-level expression in tumor cells only, several p53-specific problems must be resolved: the transduced wild-type p53 should not hetero-oligomerize with the endogenous mutant p53 because this would result in its inactivation; and it should specifically induce

apoptosis rather than growth arrest. Both problems can be overcome by genetically modifying the wild-type p53 to create a "super p53," which is more potent in induction of apoptosis than wild-type p53, and which contains a "foreign" oligomerization domain preventing hetero-oligomerization.

Conventional approaches. Development of low-molecular-weight compounds to be used as drugs in p53-based therapies is a promising alternative to p53-based gene therapy. One possibility is the restoration of the wild-type function of mutant p53 by drugs directly binding to the mutant p53. This has been demonstrated to work in vitro with tumor cells. However, the challenge will be to develop a drug that specifically reactivates mutant p53 without activating wild-type p53 in healthy tissues. Conversely, inactivation of the oncogenic properties of mutant p53 will possibly render otherwise resistant tumors amenable to conventional tumor therapy. A prerequisite, however, will be the identification of the oncogenic functions of mutant p53 at the molecular level. Finally, inhibition of the transactivation function of wild-type p53 in healthy tissue by a systemically applicable compound could prevent radiation-induced apoptosis in tissues surrounding the tumor, thereby allowing the use of very high doses of radiation in tumor therapy. Obviously, treatment with such compounds should not abolish p53-mediated protection of the healthy tissue from radiation-induced DNA damage, as predicted by the concept of an active involvement of wild-type p53 in repair processes in the absence of p53 transcriptional activation.

Like no other tumor-associated protein so far, p53 has stimulated both basic and clinically oriented molecular tumor research. Although current understanding of the molecular functions of p53 is still far ahead of its translation into therapies, this avenue of p53 research is picking up momentum. Given the importance of p53 in tumor treatment, any successful p53-based tumor therapy should provide a major breakthrough in the fight against cancer. *See* CANCER (MEDICINE); CELL CYCLE; GENETICS; MUTATION.

Wolfgang Deppert

Bibliography. E. Kim and W. Deppert, *Biochem. Cell Biol.*, 81(3):141–150, 2003; Reviews: p53, *Human Mutation*, vol. 21, 2003; Reviews: p53, *Oncogene*, vol. 18, no. 53, 1999.

Tumor viruses

Viruses associated with tumors can be classified in two broad categories depending on the nucleic acid in the viral genome and the type of strategy to induce malignant transformation.

RNA viruses. The ribonucleic acid (RNA) tumor viruses are retroviruses. When they infect cells, the viral RNA is copied into deoxyribonucleic acid (DNA) by reverse transcription and the DNA is inserted into the host genome, where it persists and can be inherited by subsequent generation of cells. Transformation of the infected cells can be traced to oncogenes that are carried by the viruses but are not necessary for viral replication. The viral oncogenes are closely similar to cellular genes, the proto-oncogenes, which code for components of the cellular machinery that regulates cell proliferation, differentiation, and death. Incorporation into a retrovirus may convert proto-oncogenes into oncogenes in two ways: the gene sequence may be altered or truncated so that it codes for proteins with abnormal activity; or the gene may be brought under the control of powerful viral regulators that cause its product to be made in excess or in inappropriate circumstances. Retroviruses may also exert similar oncogenic effects by insertional mutation when DNA copies of the viral RNA are integrated into the host-cell genome at a site close to or even within proto-oncogenes.

RNA tumor viruses cause leukemias, lymphomas, sarcomas, and carcinomas in fowl (avian leukosis virus, Rous sarcoma virus); rodents (Gross leukemia virus, Friend erythroleukemia virus, Harvery sarcoma virus, Moloney sarcoma virus, Bittner mammary adenocarcinoma virus); primates (simian sarcoma virus, gibbon ape leukemia virus); and other species. The human T-cell leukemia virus (HTLV) types I and II are endemic in Southeast Asian populations and cause adult T-cell leukemia and hairy-cell leukemia. *See* AVIAN LEUKOSIS; LEUKEMIA; LYMPHOMA; ROUS SARCOMA.

DNA viruses. DNA viruses replicate lytically and kill the infected cells. Transformation occurs in nonpermissive cells where the infection cannot proceed to viral replication. The transforming ability of DNA tumor viruses has been traced to several viral proteins that cooperate to stimulate cell proliferation, overriding some of the normal growth control mechanisms in the infected cell and its progeny.

Unlike retroviral oncogenes, DNA virus oncogenes are essential components of the viral genome and have no counterpart in the normal host cells. Some of these viral proteins bind to the protein products of two key tumor suppressor genes of the host cells, the retinoblastoma gene and the p53 gene, deactivating them and thereby permitting the cell to replicate its DNA and divide. Other DNA virus oncogenes interfere with the expression of cellular genes either directly or via interaction with regulatory factors. There is often a delay of several years between initial viral infection in the natural host species and the development of cancer, indicating that, in addition to virus-induced transformation, other environmental factors and genetic accidents are involved. A specific or general impairment of the host immune responses often plays an important role.

DNA tumor viruses belong to the families of papilloma, polyoma, adeno, hepadna, and herpes viruses and produce tumors of different types in frog (Luke's renal adenocarcinoma of leopard frog), chicken (Marek's disease, a herpesvirus-induced malignant lymphoma), woodchuck and ducks (hepatitis virus), rodents (Shope papilloma virus in rabbit, SV40 and adenovirus 12 in hamsters, polyoma virus

in mice, hamster, rat, guinea pigs, and rabbits), and primates (*Herpesvirus saimiri* and *H. ateles* in New World monkeys). *See* INFECTIOUS PAPILLOMATOSIS; MUTATION.

DNA tumor viruses are thought to play a role in the pathogenesis of about 15–20% of human cancers. These include African Burkitt's lymphoma, nasopharyngeal carcinoma, immunoblastic lymphomas in immunosuppressed individuals and a proportion of Hodgkin's lymphomas that are all associated with the Epstein-Barr virus of the herpes family; and liver carcinoma in Southeast Asia and tropical Africa associated with chronic hepatitis B virus infection. *See* ANIMAL VIRUS; CANCER (MEDICINE); EPSTEIN-BARR VIRUS; HODGKIN'S DISEASE; ONCOLOGY. Maria G. Masucci

Bibliography. J. M. Bishop, Viral oncogenes, *Cell*, 42:23–38, 1985; J. F. Nevins, E2F: A link between the Rb tumor suppressor protein and viral oncoproteins, *Science*, 258:424–429, 1992; R. A. Wienberg, Tumor suppressor genes, *Science*, 254:1138–1146, 1991; H. zur Hausen, Viruses in human cancers, *Science*, 254:1167–1173, 1991.

Tuna

Certain perciform (spiny-rayed) fishes in the family Scombridae. Like all other scombrids (such as mackerels, bonitos, wahoo, and sierras), tunas have a fusiform (tapering toward each end) and moderately compressed body and certain other characteristics that adapt them for sustained swimming at high speeds. The long spinous dorsal fin is depressible in a groove in the back; the pelvic fins and usually the pectoral fins are small and retractable in shallow depressions; the scales are typically small, but enlarged scales usually cover the anterior part of the body and lateral line, and form an envelopment called the corselet. The eyes protrude very little, if at all, beyond the surface of the head; the mouthparts fit snugly against the pointed head; and the gill covers fit snugly against the body. These features eliminate almost all irregularities that would cause resistance to the water. Tunas are also recognized by the finlets (independent multibranched rays, each appearing as a small fin) behind the dorsal and anal fins. The slender caudal peduncle, supported on each side by two small keels and a large median keel in between, and the lunate caudal fin are driven by powerful muscles for fast and sustained swimming. Sustained swimming depends on red muscle (comparatively thin muscle fibers containing large amounts of myoglobin and mitochondria), and the body temperature of tunas may be several degrees above water temperature. Tunas feed on a wide variety of fishes, squids, and crustaceans. *See* MUSCULAR SYSTEM; PERCIFORMES.

In American waters there are nine species of scombrids in four genera which bear the name tuna. However, there are other scombrids that are just as much tuna fishes as those that bear the name. Following is a brief description of "tunas" that occur in the western North Atlantic, the eastern North Pacific, and beyond. Practically all tunas are pelagic (that is, living in the open sea) and oceanodromous (migratory in salt water), with most species being highly migratory. Most species are very important commercial food fishes as well as game fishes.

Tribe Thunnini. The upper part of the body lacks distinct dark oblique stripes; the surface of the tongue has two longitudinal ridges; the maxilla can reach to, but does not go past, the middle of the eye; and there are 9–16 spines in the first dorsal fin.

Slender tuna (Allothunnus fallai). The body is slender, with a fork length (the length of a fish from its mouth to the fork in its tail) about five times the depth of the body; the dorsum is bluish to deep purple; and the belly is white without stripes or spots. Slender tuna are circumglobal and of minor commercial importance. The fork length is 105 cm (41 in.).

Genus Auxis. The first and second dorsal fins are widely separated, and the interpelvic process (a fleshy projection between the inner edges of the pelvic fins) is a large, single-pointed flap.

Bullet tuna (Auxis rochei). Bullet tuna are also called bullet mackerel. They inhabit the Pacific, Indian, and Atlantic oceans, including the Mediterranean Sea. The fork length is 50 cm (20 in.).

Frigate tuna (Auxis thazard). Frigate tuna are also called frigate mackerel. They inhabit the Atlantic, Indian, and Pacific oceans. The fork length is 65 cm (26 in.).

Genus Euthynnus. The body is naked beyond the corselet, and usually has black spots between the pectoral and pelvic fin bases.

Little tunny (Euthynnus alletteratus). Little tunny are reef-associated and oceanodromous. They inhabit the tropical and subtropical Atlantic, including the Mediterranean and Black seas, Caribbean Sea, and Gulf of Mexico. The total length is 122 cm (48 in.).

Kawakawa (Euthynnus affinis). The posterior portion of the back has a pattern of broken oblique stripes. They inhabit the Indo–West Pacific and eastern central Pacific and are highly migratory. The fork length is 100 cm (39 in.).

Black skipjack (Euthynnus lineatus). The black skipjack are generally iridescent blue with black dorsal markings composed of three to five horizontal stripes and variable black or dark gray spots above the pelvic fins; they occasionally have extensive longitudinal stripes of light gray on the belly. They are of minor commercial importance. They inhabit the eastern Pacific from California to Peru and the Galápagos Islands. The fork length is 84 cm (33 in.).

Skipjack tuna (Katssuwonus pelamis). These have three to five longitudinal stripes on the belly. Although highly important commercially, the skipjack tuna have been reported for ciguatera poisoning. They are highly migratory in all tropical and warm-temperate seas, and are absent from the Mediterranean and Black seas. The fork length is 108 cm (42 in.).

Genus Thunnus. There are no dark stripes or black spots on the body, which is covered with very small scales beyond the corselet.

Yellowfin tuna (Thunnus albacares). The second dorsal and anal fins are exceptionally long and falcate

(sickle-shaped) and yellow. They are found worldwide but are absent from the Mediterranean Sea. The fork length is 239 cm (94 in.).

Blackfin tuna (Thunnus atlanticus). Blackfin tuna have relatively few (19 to 25) gill rakers (bony or cartilaginous projections that point forward and inward from the gill arches), whereas other species of genus *Thunnus* have more. Their backs are metallic dark blue, their sides are silvery, and their bellies are milky white. They are found in the western Atlantic from Massachusetts to the Trinidad Islands to Rio de Janeiro, Brazil. The fork length is 108 cm (42 in.).

Bigeye tuna (Thunnus obesus). The corselet of bigeye tuna is not very distinct; the pectoral fin is moderately long; the first dorsal fin is deep yellow; the second dorsal and anal fins are light yellow; and the finlets are bright yellow edged with black. They inhabit tropical and subtropical waters of the Atlantic, Indian, and Pacific oceans, and are absent from the Mediterranean Sea. The total length is 250 cm (98 in.).

Atlantic bluefin tuna (Thunnus thynnus). These tuna have backs that are dark steel blue or nearly black; their sides and belly are silvery gray, with large silvery spots and bands; their cheeks are silvery (see **illustration**). They inhabit the western Atlantic from Canada to the Gulf of Mexico, the Caribbean Sea to Venezuela, and south to Brazil; and Norway to the Canary Islands, including the Mediterranean and Black seas. These are the largest tunas. The all-tackle angling record is a 679 kg (1494 lb) specimen caught off Nova Scotia in 1979. The total length is 458 cm (180 in.).

Pacific bluefin tuna (Thunnus orientalis). Once considered as a subspecies, the Atlantic bluefin and the Pacific bluefin are very similar, with the latter reaching a much smaller maximum size. They inhabit the Gulf of Alaska to southern California, and the Sea of Okhotska south to the northern Philippines. The fork length is 300 cm (118 in.).

Albacore (Thunnus alalunga). Albacore constitute the most distinctive species of *Thunnus*, with a long paddle-shaped pectoral fin that reaches past the second dorsal fin and often to the second finlet. They are cosmopolitan in range, inhabiting the tropical and temperate waters of all oceans, including the Mediterranean Sea. The fork length is 140 cm (55 in.).

Tribe Sardini. The upper part of the body has 5 to 10 dark oblique stripes; the surface of the tongue lacks a pair of cartilaginous longitudinal ridges; the mouth is large, with the maxilla reaching past the middle of the eye; and there are 20 to 22 spines in the first dorsal fin.

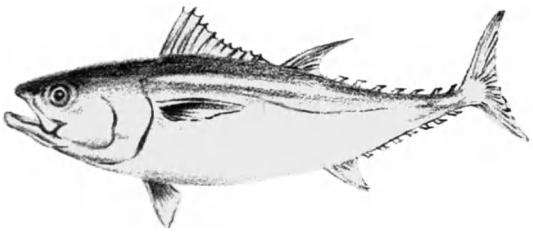

Bluefin tuna (*Thunnus thynnus*).

Atlantic bonito (Sarda sarda). The mouth of Atlantic bonito is large, with the maxilla reaching beyond the posterior margin of the eye. The body exhibits oblique dorsal stripes with a greater angle than in other species of *Sarda*. Atlantic bonito have been reported for ciguatera poisoning. They inhabit the eastern Atlantic from Norway to South Africa, including the Mediterranean and Black seas; the western Atlantic from Nova Scotia to Florida, the northern Gulf of Mexico, and Colombia to northern Argentina; and are apparently absent from most of the Caribbean Sea. The fork length is 96 cm (38 in.).

Pacific bonito (Sarda chiliensis lineolata). These exhibit oblique stripes on the back; the mouth is of moderate size. They inhabit the eastern Pacific from Alaska to Baja California. The fork length is 102 cm (40 in.).

Striped bonito (Sarda orientalis). These have virtually horizontal stripes on the back. They inhabit the Indo-Pacific, the Hawaiian Islands, and the eastern Pacific from the coast of the United States to Baja California to Peru and the Galápagos Islands. The fork length is 102 cm (40 in.).

Herbert Boschung

Bibliography. B. B. Collette, Mackerels, Family Scombridae, pp. 516–536 in B. B. Collette and G. Klein-MacPhee (eds.), *Bigelow and Schroder's Fishes of the Gulf of Maine*, 3d ed., Smithsonian Institution Press, Washington, DC, 2002; B. B. Collette, Mackerels, molecules, and morphology, pp. 149–164 in B. Séret and J.-Y. Sire (eds.), *Proceedings of the 5th Indo-Pacific Fish Conference*, Noumea, Paris, 1999; B. B. Collette, Scombridae, in W. Fisher (ed.), *FAO Species Identification Sheets for Fishery Purposes*, vol. 4: *Western Central Atlantic*, Rome, 1978; B. B. Collette and C. R. Aadland, Revision of the frigate tunas (Scombridae, *Auxis*), with descriptions of two new subspecies from the eastern Pacific, *Fish. Bull.*, 94:423–441, 1996; B. B. Collette and C. E. Nauen, *FAO Species Catalogue*, vol. 2: *Scombrids of the World*, an annotated and illustrated catalogue of tunas, mackerels, bonitos and related species known to date, FAO Fish. Synopsis, 125(2), 1983; B. B. Collette and B. R. Smith, Bluefin tuna, *Thunnus thynnus orientalis* from the Gulf of Papua, *Jap. J. Ichthyol.*, 28(2):166–168, 1981; B. W. Halstead, P. S. Auerbach, and D. R. Campbell, *A Colour Atlas of Dangerous Marine Animals*, Wolfe Medical Publications, W. S. Cowell, Ipswich, England, 1990; D. A. Olsen, D. W. Nellis, and R. S. Wood, Ciguatera in the Eastern Caribbean, *Mar. Fish. Rev.*, 46(1):13–18, 1984.

Tundra

An area supporting some vegetation beyond the northern limit of trees, between the upper limit of trees and the lower limit of perennial snow on

mountains, and on the fringes of the Antarctic continent and its neighboring islands. The term is of Lapp or Russian origin, signifying treeless plains of northern regions. Biologists, and particularly plant ecologists, sometimes use the term tundra in the sense of the vegetation of the tundra landscape. Tundra has distinctive characteristics as a kind of landscape and as a biotic community, but these are expressed with great differences according to the geographic region.

Patterns. Characteristically tundra has gentle topographic relief, and the cover consists of perennial plants a few centimeters to a meter or a little more in height. The general appearance during the growing season is that of a grassy sward in the wetter areas, a matted spongy turf on mesic sites, and a thin or sparsely tufted lawn or lichen heath on dry sites. In winter, snow mantles most of the surface with drifts shaped by topography and surface objects including plants; vegetation patterns are largely determined by protecting drifts and local areas exposed to drying and scouring effects of winter winds. By far, most tundra occurs where the mean annual temperature is below the freezing point of water, and perennial frost (permafrost) accumulates in the ground below the depth of annual thaw and to depths at least as great as 1600 ft (500 m). A substratum of permafrost, preventing downward percolation of water, and the slow decay of water-retaining humus at the soil surface serve to make the tundra surface moister during the thaw season than the precipitation on the area would suggest. Retention of water in the surface soils causes them to be subject to various disturbances during freezing and thawing, as occurs at the beginning and end of, and even during, the growing season. Where the annual thaw reaches depths of less than about 20 in. (50 cm), the soils undergo "swelling," frost heaving, frost cracking, and other processes that result in hummocks, polygonal ridges or cracks, or "soil flows" that slowly creep down slopes. As the soils are under this perennial disturbance, plant communities are unremittingly disrupted and kept actively recolonizing the same area. Thus topography, snow cover, soils, and vegetation interact to produce patterns of intricate complexity when viewed at close range.

Plant species, life-forms, and adaptations. The plants of tundra vegetation are almost exclusively perennial. A large proportion have their perennating buds less than 8 in. (20 cm) above the soil (chamaephytes in the Raunkiaer life-form system), especially among the abundant mosses and lichens. Another large group has the perennating organs at the surface of the soil (hemicryptophytes in the Raunkiaer system). Vegetative reproduction is common—by rhizomes (many of the sedges), stolons (certain grasses and the cloudberry, *Rubus chamaemorus*), or bulbils near the inflorescence (*Polygonum viviparum*, *Poa vivipara*, *Saxifraga hirculis*); thus clone formation is common in plant populations. Apomixis, the short-circuiting of the sexual reproduction process, is found frequently among flowering plants of tundra. Seed is set regularly by agamospermy, for example, in many dandelions (*Taraxacum* sp.), hawkweeds (*Hieracium* sp.), and grasses (*Calamogrostis* sp., *Poa* sp., *Festuca* sp.). The high incidence of apomixis in tundra flowering plants is coincident with high frequency of polyploidy, or multiple sets of chromosomes, in some circumstances a mechanical cause of failure of the union of gametes by the regular sexual process. Asexual reproduction and polyploidy tend to cause minor variations in plant species populations to become fixed to a greater extent than in populations at lower latitudes, and evolution tends to operate more at infraspecific levels without achieving major divergences. Adaptations are more commonly in response to physical factors of the stressful cold environment rather than to biotic factors, such as pollinators or dispersal agents, of the kinds that exert such control in the congenial warm, moist climates.

Soil conditions. Tundra soils are azonal, without distinct horizons, or weakly zonal. Soils on all but very dry and windswept sites tend to accumulate vegetable humus because low temperatures and waterlogging of soils inhibit processes of decay normally carried out by bacteria, fungi, and minute animals. Where permafrost or other impervious layers are several meters or more beneath the surface in soils with some fine-grained materials, leaching produces an Arctic Brown Soil in which there is moderately good drainage and cycling of mineral nutrients. In the greater part of tundra regions not mantled by coarse, rocky "fjell-field" materials (**Fig. 1**), the soils are more of the nature of half-bog or bog soils. These are characterized by heavy accumulations of raw or weakly decayed humus at the surface overlying a waterlogged or perennially frozen mineral horizon that is in a strongly reduced state from lack of aeration. Such boggy tundra soils are notoriously unproductive from the standpoint of cultivated plants, but they are moderately productive from the standpoint of shallowly rooted native plants. In

Fig. 1. Fjell-field tundra of the high Arctic. Sedges, mosses, and lichens form a thin and discontinuous sod. Late-persisting snowbanks are withdrawing from surfaces that are lighter in color because they lack many of the common plants, including dark-colored species of lichens. (*W. S. Benninghoff, U.S. Geological Survey*)

Fig. 2. Alpine tundra in French Alps. Altitudinal limit of trees occurs in valley behind building in middle distance. Although similar in vegetation structure to tundra of polar regions, Alpine tundras of lower latitudes are usually richer in vascular plant species than tundras of polar regions, and structure and composition of the vegetation have been modified by pasturing. (*W. S. Benninghoff, U.S. Geological Survey*)

Finland, forest plantations are being made increasingly productive on such soils by means of nutrient feeding to aerial parts. *See* SOIL.

Productivity. By reason of its occurrence where the growing season is short and where cloudiness and periods of freezing temperatures can reduce growth during the most favorable season, tundra vegetation has low annual production. Net radiation received at the Earth's surface is less than 20 kg-cal/($cm^2 \cdot$ year) [84 kilojoules/($cm^2 \cdot$ year)] for all Arctic and Antarctic tundra regions. Assuming a 2-month growing season and 2% efficiency for accumulation of green plant biomass, 1 cm^2 could accumulate biomass equivalent to 66.6 g-cal/year (279 J/year). This best value for tundra is not quite one-half the world average for wheat production and about one-eighth of high-yield wheat production. The tundra ecosystem as a whole runs on a lower energy budget than ecosystems in lower latitudes; in addition, with decomposer and reducer organisms working at lower efficiency in cold, wet soils, litter, and humus accumulate, further modifying the site in unfavorable ways. Grazing is one of the promising management techniques (**Fig. 2**) because of its assistance in speeding up the recycling of nutrients and reducing accumulation of raw humus. *See* BIOMASS; ECOLOGICAL ENERGETICS; ECOSYSTEM.

Fauna. The Arctic tundras support a considerable variety of animal life. The vertebrate herbivores consist primarily of microtine mammals (notably lemmings), hares, the grouselike ptarmigan, and caribou (or the smaller but similar reindeer of Eurasia). Microtine and hare populations undergo cyclic and wide fluctuations of numbers; these fluctuations affect the dependent populations of predators, the foxes, weasels, hawks, jaegers, and eagles. Alpine tundras generally have fewer kinds of vertebrate animals in a given area because of greater discontinuity of the habitats. Arctic and Alpine tundras have distinctive migrant bird faunas during the nesting season. Tundras of the Aleutian Islands and other oceanic islands are similar to Alpine tundras with respect to individuality of their vertebrate faunas, but the islands support more moorlike matted vegetation over peaty soils under the wetter oceanic climate. Tundras of the Antarctic continent have no vertebrate fauna strictly associated with it. Penguins and other sea birds establish breeding grounds locally on ice-free as well as fringing ice-covered areas. The only connection those birds have with the tundra ecosystem is the contribution of nutrients from the sea through their droppings. All tundras, including even those of the Antarctic, support a considerable variety of invertebrate animals, notably nematode worms, mites, and collembola on and in the soils, but some other insects as well. Soil surfaces and mosses of moist or wet tundras in the Arctic often teem with nematodes and collembola. Collembola, mites, and spiders have been found above 20,000 ft (6000 km) in the Himalayas along with certain molds, all dependent upon organic debris imported by winds from richer communities at lower altitudes. *See* COLLEMBOLA; TAIGA.

William S. Benninghoff

Bibliography. Arctic Institute of North America, *Arctic Bibliography*, 16 vols., 1953–1975; M. J. Dunbar, *Ecological Development in Polar Regions*, 1968; R. E. English, *World Regional Geography*, 1990; J. D. Ives and R. G. Barry, *Arctic and Alpine Environments*, 1974; M. C. Kellman, *Plant Geography*, 2d ed., 1980; G. A. Llano (ed.), *Antarctic Terrestrial Biology*, American Geophysical Union, Antarctic Research Series, vol. 20, 1972; J. C. F. Tedrow (ed.), *Antarctic Soils and Soil Forming Processes*, American Geophysical Union, Antarctic Research Series, vol. 8, 1966; H. E. Wright, Jr., and W. H. Osburn, *Arctic and Alpine Environments*, 1968.

Tung tree

The plant *Aleurites fordii*, a species of the spurge family (Euphorbiaceae). The tree, native to central and western China, is the source of tung oil.

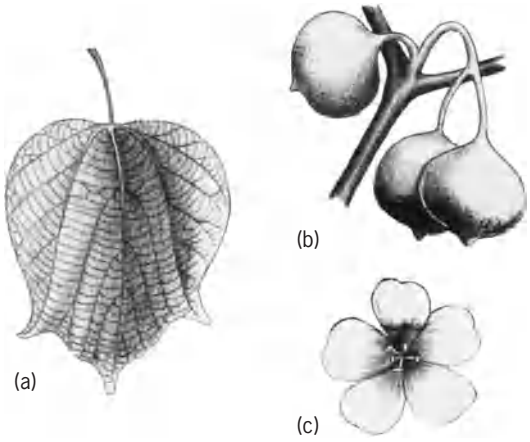

Tung tree. (*a*) Leaf. (*b*) Fruits. (*c*) Flower.

It has been grown successfully in the southern United States. The globular fruit (see **illus.**) has three to seven large, hard, rough-coated seeds containing the oil, which is expressed after the seeds have been roasted. Tung oil is used to produce a hard, quick-drying, superior varnish, which is less apt to crack than other kinds. The foliage, sap, fruit, and commercial tung meal contain a toxic saponin, which causes gastroenteritis in animals that eat it. *See* DRYING OIL; EUPHORBIALES; VARNISH.

Perry D. Strausbaugh; Earl L. Core

Tungsten

A chemical element, W, atomic number 74, and atomic weight 183.85. Naturally occurring tungsten consists of five stable isotopes having the following mass numbers and relative abundances: 180 (0.14%), 182 (26.4%), 183 (14.4%), 184 (30.6%), and 186 (28.4%). Twelve radioactive isotopes ranging from 173 to 189 also have been characterized. *See* PERIODIC TABLE.

1												13	14	15	16	17	18
1 **H**	2																2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Tungsten crystallizes in a body-centered cubic structure in which the shortest interatomic distance is 274.1 picometers at 25°C (77°F). The pure metal has a lustrous, silver-white appearance. It possesses the highest melting point, lowest vapor pressure, and the highest tensile strength at elevated temperature of all metals. Some important physical properties of tungsten are compiled in the **table**.

At room temperature tungsten is chemically resistant to water, oxygen, most acids, and aqueous alkaline solutions, but it is attacked by fluorine or a mixture of concentrated nitric and hydrofluoric acids.

Tungsten is used widely as a constituent in the alloys of other metals, since it generally enhances high-temperature strength. Several types of tool steels and some stainless steels contain tungsten. Heat-resistant alloys, also termed superalloys, are nickel-, cobalt-, or iron-based systems containing varying amounts (typically 1.5–25 wt. %) of tungsten. Wear-resistant alloys having the trade name Stellites are composed mainly of cobalt, chromium, and tungsten. *See* ALLOY; HIGH-TEMPERATURE MATERIALS.

The major use of tungsten in the United States is in the production of cutting and wear-resistant materials. Tungsten carbides (representing 60% of total tungsten consumption) are used for cutting tools, mining and drilling tools, dies, bearings, and armor-piercing projectiles.

Unalloyed tungsten (25% of tungsten consumption) in the form of wire is used as filaments in incandescent and fluorescent lamps, and as heating elements for furnaces and heaters. Because of its high electron emissivity, thorium-doped (thoriated) tungsten wire is employed for direct cathode electronic filaments. Tungsten rods find use as lamp filament supports, electrical contacts, and electrodes for arc lamps.

Tungsten compounds (5% of tungsten consumption) have a number of industrial applications. Calcium and magnesium tungstates are used as phosphors in fluorescent lights and television tubes. Sodium tungstate is employed in the fireproofing of fabrics and in the preparation of tungsten-containing dyes and pigments used in paints and printing inks. Compounds such as WO_3 and WS_2 are catalysts for various chemical processes in the petroleum industry. Both WS_2 and WSe_2 are dry, high-temperature lubricants. Other applications of tungsten compounds have been made in the glass, ceramics, and tanning industries.

Miscellaneous uses of tungsten account for the remainder (2%) of the metal consumed.

Charles Kutal

Physical properties of tungsten

Property	Value
Melting point	3410 ± 20°C (6170 ± 36°F)
Boiling point	5700 ± 200°C (10,300 ± 360°F)
Density, 27°C (81°F)	19.3 g/cm^3 (11.2 oz/in.^{3})
Specific heat, 25°C (77°F)	0.032 cal/g-°C (0.13 J/g-°C)
Heat of fusion	52.2 ± 8.7 cal/g (218 ± 36 J/g)
Vapor pressure 2027°C (3681°F)	6.4×10^{-12} atm (6.5×10^{-7} Pa)
3382°C (6120°F)	2.3×10^{-5} atm (2.3 Pa)
5470°C (9878°F)	0.53 atm (5.4×10^4 Pa)
Electrical resistivity 27°C (81°F)	5.65 microhm-cm
1027°C (1881°F)	34.1
3027°C (5481°F)	103.3
Thermal conductivity 27°C (81°F)	0.43 cal/cm-s-°C (1.8 J/cm-s-°C)
1027°C (1881°F)	0.27 (1.1)
Absorption cross section, 0.025-eV neutrons	18.5 ± 0.5 barns ($18.5 \pm 0.5 \times 10^{-24}$ cm^2)

Bibliography. F. A. Cotton and G. Wilkinson, *Advanced Inorganic Chemistry*, 6th ed., 1999; E. Lassner and W.-D. Schubert, *Tungsten: Properties, Chemistry, Technology of the Element, Alloys, and Chemical Compounds*, 1999; E. Pink and L. Bartha (eds.), *The Metallurgy of Doped Non-Sag Tungsten*, 1989.

Tunicata (Urochordata)

A subphylum of marine animals of the Chordata. They are characterized by a perforated pharynx or branchial sac used for food collection, a dorsal notochord restricted to the tail of the larva (and the adult in one class), absence of mesodermal segmentation or a recognizable coelom, and secretion of an outer covering (the test or tunic) which contains large amounts of polysaccharides related to cellulose (see **illustration**).

Three classes are usually recognized: the sessile Ascidiacea (sea squirts or ascidians); planktonic Thaliacea (salps, doliolids, and pyrosomids); and Appendicularia, minute planktonic forms with tails living inside a specialized test or house adapted for filtering and food gathering. Approximately 2000 species of Tunicata are recognizable. The group is found in all parts of the ocean.

Except for a few deep-sea carnivorous forms, tunicates feed on minute plankton and finely divided organic detritus. Food drawn into the pharynx in water currents created by ciliary or muscular activity is filtered on a mucous sheet secreted in the floor of the pharynx and rolled into a cord on the dorsal side. The food cord is passed into the esophagus and from there to the rest of the alimentary canal. Digestion is extracellular; feces are discharged into the outgoing current of water. There are no excretory organs, and most nitrogenous wastes are removed in soluble form. Concentrations of urates

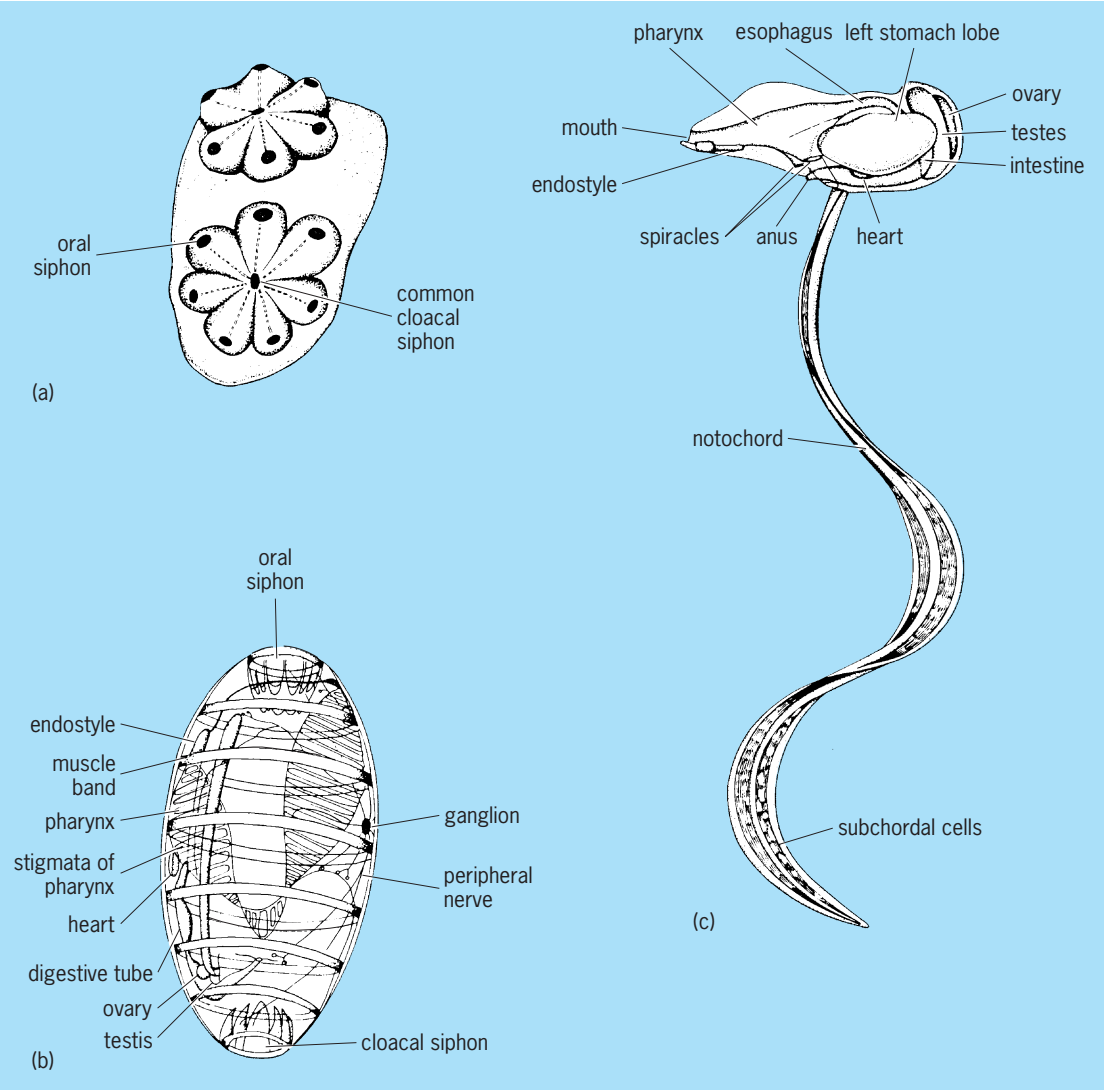

Representative Tunicata. (*a*) *Botryllus*, a compound ascidian; (*b*) *Doliolum*, a planktonic thaliacean (*after P. A. Meglitsch, Invertebrate Zoology, Oxford University Press, 1967*). (c) *Oikopleura*, a planktonic larvacean, removed from its house (*after A. Alldredge, Appendicularians, Sci. Amer., 235:94–102, 1976*).

may also accumulate throughout life in closed vesicles. The heart is tubular and reverses the direction of its beat at intervals, alternately driving blood in opposite directions through the body organs and through the pharyngeal wall. Heavy metals, usually vanadium or iron, accumulate in specialized cells in the blood; these are probably concerned with test formation or chemical defense, and are not respiratory carriers. The nervous system is simple and consists of a solid dorsal ganglion and a few peripheral nerves. The ganglion is closely associated with a neural gland discharging into the entrance of the pharynx.

Most tunicates are hermaphroditic and may be oviparous or viviparous. The larva is a minute (about 0.04 in. or 1 mm) tailed tadpole with a notochord, dorsal nerve cord, and sense organs. Larvae are free swimming for a short period and, except in Appendicularia, lose the tail and notochord at metamorphosis into an adult. Ascidians may be solitary or form colonies; Thaliacea form aggregates or colonies by asexual budding and may alternate between solitary and aggregate forms. Appendicularia are always solitary.

Tunicates have little economic importance except as fouling organisms. A few species have pharmacological properties, and a few larger ascidians are used for food. *See* APPENDICULARIA (LARVACEA); ASCIDIACEA; CHORDATA; THALIACEA. Ivan Goodbody

Bibliography. A. Alldredge, Appendicularians, *Sci. Amer.*, 235:94–102, 1976; N. J. Berrill, *The Tunicata*, 1950; P. Brien and W. Harant, Embranchement des tuniciers, in P. P. Grassé (ed.), *Traité de Zoologie*, vol. 11, 1948; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982.

Tuning

The process of adjusting the frequency of a vibrating system to obtain a desired result. The term is applicable to a wide variety of such systems, but is most commonly used in connection with musical instruments, electronic circuits, and machinery.

Musical instruments. The frequency (pitch) of stringed instruments is determined by the length, mass, and tension of vibrating strings. In members of the violin family, for example, tuning is accomplished by adjusting the tension of each string individually. In keyboard instruments, the process of tension adjustment must achieve a tempered scale in each octave throughout the range of the keyboard. The pitch of wind instruments depends on the volume and shape of a vibrating column of air. Tuning is accomplished by adjusting the shape of the column, usually its length. *See* MUSICAL ACOUSTICS; MUSICAL INSTRUMENTS.

Electronic circuits. In electronic circuits, there are a variety of frequency-determining elements. The most widely used is a combination of an inductance L (which stores energy in a magnetic field) and a capacitance C (which stores it in an electric field). The frequency of oscillation is determined by the rate of exchange of the energy between the two fields, and is inversely proportional to $\sqrt{LC}$. Tuning is accomplished by adjusting the capacitor or the inductor until the desired frequency is reached. The desired frequency may be one that matches (resonates with) another frequency. This occurs when a piano tuner adjusts a string's tension until its vibration agrees with that of a tuning fork, and when a radio receiver is tuned to a desired station. Another purpose of tuning may be to match a frequency standard, as when setting an electronic watch to keep accurate time. The frequency-determining element in such watches, as well as in radio transmitters, digital computers, and other equipment requiring precise frequency adjustment, is a vibrating quartz crystal. The frequency of vibration of such crystals can be changed over a narrow range by adjusting a capacitor connected to it. *See* QUARTZ CLOCK; RADIO RECEIVER; RESONANCE (ALTERNATING-CURRENT CIRCUITS).

At extremely high radio frequencies (in the range above a gigahertz, or 10^9 vibrations per second), the inductance and capacitance are not discrete elements, but are distributed over the surface of the apparatus to be tuned. An example is the microwave cavity, in which electromagnetic waves are trapped and reflected. The cavity is tuned by a plunger or other movable element that changes the volume of the cavity. *See* CAVITY RESONATOR; MICROWAVE.

Another function of tuning in electronics is the elimination of undesired signals. Filters for this purpose employ inductors and capacitors, or crystals. The filter is tuned to the frequency of the undesired vibration, causing it to be absorbed elsewhere in the circuit. *See* ELECTRIC FILTER.

Automatic tuning by electrical control is accomplished by a varactor diode. This is a capacitor whose capacitance depends on the direct-current (dc) voltage applied to it. The varactor serves as a portion of the capacitance of the tuned circuit. Its capacitance is controlled by a dc voltage applied to it by an associated circuit, the voltage and its polarity depending on the extent and direction of the mismatch between the desired frequency and the actual frequency. *See* VARACTOR.

The extremely high-frequency electromagnetic waves associated with visible, infrared, and ultraviolet radiation could not be tuned in the manner of the longer radio waves until the advent of the laser. The very narrow spectrum regions of radiation produced by lasers can be tuned by various means, such as temperature control or the insertion of prisms or gratings in the laser cavity. *See* LASER.

Machinery. Tuning is a technique also applied to the vibrations of machinery. The vibrating elements are typically weights associated with springs, or their equivalent inertial mass and restoring force. Suppression of undesired vibration is the usual aim of mechanical tuning, by the mechanical equivalent of an electrical filter. *See* RESONANCE (ACOUSTICS AND MECHANICS). Donald G. Fink

Tuning fork

A steel instrument consisting of two prongs and a handle which, when struck, emits a tone of fixed pitch. Because of their simple mechanical structure, purity of tone, and constant frequency, tuning forks are widely used as standards of frequency in musical acoustics. In its electrically driven form, a tuning fork serves to control electric circuits by producing frequency standards of high accuracy and stability.

A tuning fork is essentially a transverse vibrator. The amplitude of longitudinal vibration at the end of the stem is small compared with the amplitude of the transverse vibrations at the ends of the prongs

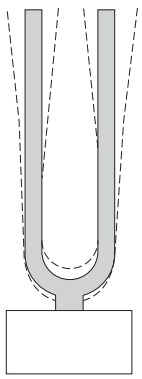

Tuning fork vibrating at its fundamental frequency.

(see **illus.**). Thus, when the stem of the fork is pressed on a sounding board or a resonance box, the vibrations persist for a considerable time since the small-amplitude vibrations transfer energy to the sounding board at a low rate.

Tuning forks are constructed to cover the entire range of audible frequencies from 20 to 20,000 Hz. The frequency of a fork varies approximately as the inverse square of the length and directly as the thickness of the prongs. The sound output of the fundamental frequency of a fork may be reinforced by attaching the stem to an air column type of resonance box having the same fundamental frequency. *See* MUSICAL ACOUSTICS; VIBRATION. Lawrence E. Kinsler

Tunnel

An underground space of substantial length, usually having a tubular shape. Tunnels can be either constructed or natural and are used as passageways, storage areas, carriageways, and utility ducts. They may also be used for mining, water supply, sewerage, flood prevention, and civil defense.

Construction. Tunnels are constructed in numerous ways. Shallow tunnels are usually constructed by burying sections of tunnel structures in trenches dug from the surface. This is a preferred method of tunneling as long as space is available and the operation will not cause disturbance to surface activities. Otherwise, tunnels can be constructed by boring underground. Short tunnels are usually bored manually or by using light machines (such as a roadheader or backhoe). If the ground is too hard to bore, a drill-and-blast method is frequently used. For long tunnels, it is more economical and much faster to use tunneling boring machines which work on the full face (complete diameter of the opening) at the same time. In uniform massive rock formations without fissures or joints, tunnels can be bored without any temporary supports to hold up the tunnel crowns. However, temporary supports are usually required because of the presence of destabilizing fissures and joints in the rock mass (**Fig. 1**). A layer of shotcrete serves as the primary lining to protect the newly exposed surface and to support the tunnel crowns as well. The shotcrete is frequently reinforced by steel meshes and, if necessary, braced by steel lattices. *See* DRILLING, GEOTECHNICAL.

In soft ground, it has become popular to use shield machines for boring and reinforced concrete segments for lining. The largest shield machine ever, 14 m (46 ft) in diameter, was used in constructing the Tokyo Trans-Bay Highway (Tokyo Wan Aqua-Line) between Kawasaki City and Kisarazu City of Japan (completed in 1997). There are various types of shield machines available to serve different purposes. The multiface shield machine (**Fig. 2**) was first used in constructing the Osaka Business Park subway station in Japan (completed in 1995). The machine is 17 m (56 ft) in width and 7.5 m (25 ft) in height and has three cutters which operate independently. Theoretically, the two side cutters can be detached from the center cutter upon the completion of station excavation, leaving the center cutter to continue boring toward the next station. However, this technique was not applied in this case, and the machine was used for the Osaka Business Park station only.

Immersed tunnels. For tunnels to be constructed across bodies of water, an alternative to boring is to lay tunnel boxes directly on the prepared seabed. These boxes, made of either steel or reinforced concrete, are prepared in dry docks and sealed at their ends by the use of bulkheads. They float as the docks are flooded, and are towed to the site by tugboats. The boxes are then flooded to allow them to sink to the seabed after they are properly positioned. Immersed tunnels are usually buried in shallow trenches dug for this purpose and covered by ballast so they will not be affected by the movement of the water. The joints between tunnel sections are made watertight by using rubber gaskets, and water is pumped out of the tunnel to make it ready for service. Among the numerous immersed tunnels the Øresund Link, completed in 1999 [3.5 km (2.2 mi)], between Denmark and Sweden is second in length to the cross-bay tunnel [5.8 km (3.6 mi)] for the Bay Area Rapid Transit system of San Francisco, California. *See* CONCRETE; STEEL.

Microtunnels. Small tunnels, such as sewer lines and water mains, are usually installed by jacking steel or concrete pipes into the ground (**Fig. 3**). The soil core inside the tubes can be removed manually or by

using "moles," which are essentially small shield tunneling machines. The alignment of the pipes is continuously monitored and adjusted. With the moles guided by a computerized navigation system, it is possible to align pipes to a precision within 100 mm (4 in.) regardless of length. This technique has been used for jacking pipes as large as 2 m (7 ft) or so in diameter for distances more than 100 m (330 ft). Pipes with smaller sizes can be jacked to distances of more than 300 m (990 ft).

Special tunneling techniques. Auxiliary measures are frequently required to ensure the safety of tunnels during boring in soft ground. Compressed air was used in the past, but is it is seldom used anymore because improper decompression may cause aeroembolism (diver's disease) in workers. Instead, grouting and ground freezing are now preferred. In the Central Artery Project in Boston, Massachusetts, ground freezing is carried out to permit three tunnels to be bored under the railway tracks leading to the South Station Railway Terminal. Roughly 1600 freezing pipes are installed to depths varying from 13.7 to 16.8 m (45 to 55 ft), and the volume of frozen soil is more than 60,000 m^3 (2,100,000 ft^3). This could be the largest undertaking of this nature.

An underpass, scheduled to be completed in 2001, is being constructed beneath the Taipei International Airport by using the Endless-Self-Advancing method. To minimize ground settlements and ensure the safety of air traffic, interlocked steel pipes are first jacked into the ground to form a protective shelter. The soil core inside the shelter is excavated at the rate of 400 mm (16 in.) per lift. Concrete segments are moved one by one into the space created, by jacking the segments behind. Each time only the last segment is jacked forward, and the jacking force is taken by the frictional resistance acting on all the rest of the segments. Cables are anchored to the first and last segments so the force acting on the last segment can be transmitted to the segments in front of it. The movements of these segments resemble the movements of centipedes. In theory, there is no limit on the length of tunnels installed by this method.

Longest tunnels. The longest tunnel of any kind is the New York City/West Delaware water supply tunnel (completed in 1944). It runs for 169 km (105 mi) from the Rondout Reservoir into the Hillview Reservoir in Yonkers, New York. The longest rail tunnel is Seikan Tunnel (53.6 km; 33.3 mi) in Japan (completed in 1985). The longest road tunnel was Saint Gotthard Tunnel (16.9 km; 10.5 mi) in Switzerland (completed in 1980). In mid2001 the title was taken over by the Laerdal Tunnel (24.5 km; 15.2 mi).

The longest undersea tunnel is the Channel Tunnel [49.4 km (30.7 mi), of which 38 km (23.6 mi) is undersea] across the English Strait. It runs from Folkestone in Britain to Calais in France. There are two running tunnels plus one service tunnel in the center, 7.6 and 4.8 m (25 and 16 ft) in internal diameter, respectively. The tunnel was officially inaugurated on May 6, 1994, when the Queen of England and President Mitterrand of France became the first official passengers to pass by train between the two countries. Eurotunnel has a concession from the British and French governments to run the tunnel until 2052.

Fig. 1. Temporary supports for a four-lane road tunnel. (*RESA Engineering Corporation*)

Fig. 2. Multiface shield tunneling machine. (*Hitachi Zosen Corporation*)

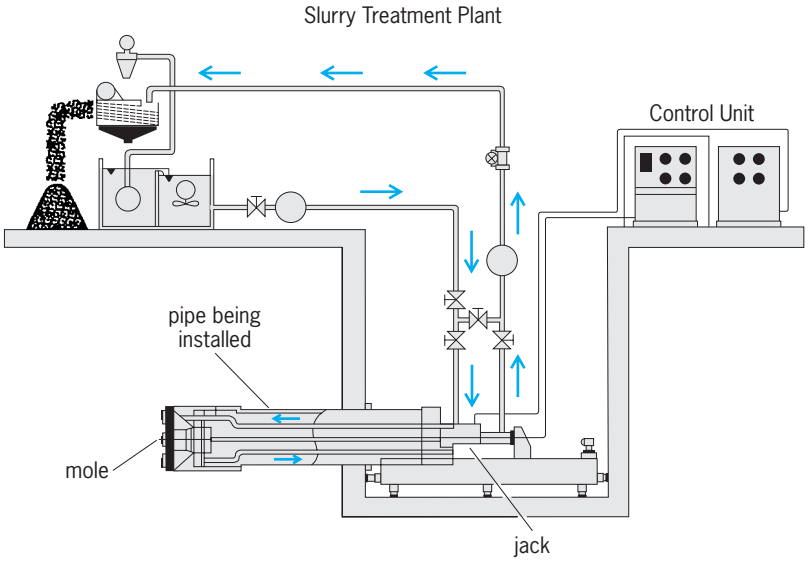

Fig. 3. Microtunneling and pipe jacking technique.

Shuttle trains, carrying up to 180 cars, and freight shuttles, carrying 28 lorries, will take 35 minutes to cross the strait. Za-Chieh Moh; Richard N. Hwang

Bibliography. C. J. Kirkland (ed.), *Eurotunnel: Engineering the Channel Tunnel*, 1995; J. O. Bickel, T. R. Kuesel, and E. H. King (eds.), *Tunnel Engineering Handbook*, 1995; Z. C. Moh et al., Underpass beneath Taipei International Airport, *Proceedings of the Conference on New Frontiers and Challenges*, Bangkok, Thailand, November 8–12, 1999; A. M. M. Wood and A. M. Wood, *Tunnelling: Management by Design*, 2000.

Tunnel diode

A two-terminal semiconductor junction device (also called the Esaki diode) which does not show rectification in the usual sense, but exhibits a negative resistance region at very low voltage in the forward-bias characteristic and a short circuit in the negative-bias direction.

The short-circuit condition exists because both the *p* and *n* regions of the device are doped with such high concentrations of the appropriate impurities that the normal barrier is rendered sufficiently thin to allow the free passage of current at zero and all negative-bias conditions. The forward-bias characteristic (**illus.** *a*) shows a maximum and a minimum in the current with a negative-resistance region between. Band-potential diagrams show the internal electronic situation existing at the current minimum (illus. *b*), the current maximum (illus. *c*), and zero bias (illus. *d*). The top of the shaded regions of the band-potential diagrams shows the level to which electrons fill the available energy levels in the valence and conduction bands of the materials forming the *pn* junction. The bottom of the conduction band is designated E_c and the top of the valence band E_v. No electrons can penetrate the forbidden energy gap between E_c and E_v except in the barrier region, where it is thin enough to allow electron transit by tunneling.

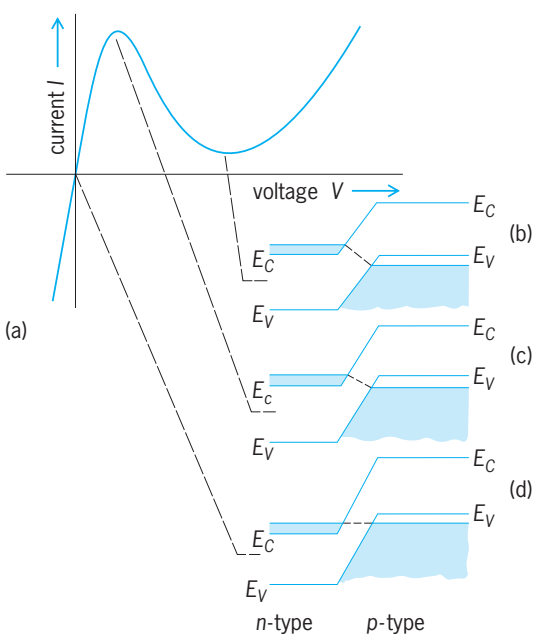

Tunnel diode characteristic. (*a*) Forward-bias voltage-current plot. (*b*) Band-potential diagram for the current minimum. (*c*) Band-potential diagram for the current maximum. (*d*) Band-potential diagram for zero bias.

The observed characteristic may be accounted for as follows. In illus. *d* the electron level is the same on both sides of the junction (zero bias). No net current flows because there is no difference in electronic energy across the junction. As forward bias is applied, tunneling current will flow since now the electrons in the *n*-type material on the left will rise to a level above those on the right (illus. *c*). As long as these electrons are still below the top of the valence band E_v on the right, current will increase. When the top of the elevated electron distribution exceeds the level of E_v, current will begin to decrease and the diode has entered the negative-resistance region. In illus. *c* the top of the electron distribution on the left is even with E_v on the right, and any further increase in bias will reduce the number of electrons available for the tunneling current. Therefore, illus. *c* corresponds to the current maximum. As bias continues to increase, the point is reached where the bottom of the conduction band E_c on the left is even with E_v on the right, and the entire electron distribution, being above E_v, is removed from the tunneling process. At this point, only normal forward-bias diffusion current flows. This current is composed of energetic electrons diffusing over the top of the barrier while remaining in the conduction band. Illustration *b* shows the point at which tunneling is no longer possible and thus corresponds to the current minimum at the end of the negative-resistance region. For further discussion of the properties of junction diodes *see* JUNCTION DIODE; SEMICONDUCTOR; ZENER DIODE. Lloyd P. Hunter

Bibliography. H. Mizuta and T. Tanoue, *The Physics and Applications of Resonant Tunnelling Diodes*, 1995; J. Singh, *Semiconductor Devices: An Introduction*, 1994; E. S. Yang, *Microelectronic Devices*, 1988.

Tunneling in solids

A quantum-mechanical process which permits electrons to penetrate from one side to the other through an extremely thin potential barrier to electron flow. The barrier would be a forbidden region if the electron were treated as a classical particle. A two-terminal electronic device in which such a barrier exists and primarily governs the transport characteristic (current-voltage curve) is called a tunnel junction. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

During the infancy of the quantum theory, L. de Broglie introduced the fundamental hypothesis that matter may be endowed with a dualistic nature—particles such as electrons, particles, and so on, may also have the characteristics of waves.

This hypothesis found expression in the definite form now known as the Schrödinger wave equation, whereby an electron or an alpha particle is represented by a solution to this equation. The nature of such solutions implies an ability to penetrate classically forbidden regions of negative kinetic energy and a probability of tunneling from one classically allowed region to another. The concept of tunneling, indeed, arises from this quantum-mechanical result. The subsequent experimental manifestations of this concept, such as high-field electron emission from cold metals, alpha decay, and so on, in the 1920s, can be regarded as one of the early triumphs of the quantum theory. *See* FIELD EMISSION; RADIOACTIVITY; SCHRÖDINGER'S WAVE EQUATION.

In the 1930s, attempts were made to understand the mechanism of electrical transport in resistive contacts between metals and rectifying metal-semiconductor contacts in terms of electron tunneling in solids. In the latter case, since a proposed theoretical model did not properly represent the actual situation, the theory predicted the wrong direction of rectification. In many cases, however, conclusive experimental evidence of tunneling was lacking, primarily because of the rudimentary stage of material science.

Tunnel diode. The invention of the transistor in 1947 spurred the progress of semiconductor technology. By the 1950s, materials technology for semiconductors such as Ge and Si was sufficiently advanced to permit the construction of well-defined semiconductor structures. The tunnel diode (also called the Esaki diode) was discovered in 1957 by L. Esaki. This discovery demonstrated the first convincing evidence of electron tunneling in solids, a phenomenon which had been clouded by questions for decades. This device is a version of the semiconductor *pn* junction diode which is made of a *p*-type semiconductor, containing mobile positive charges called holes (which correspond to the vacant electron sites), and an *n*-type semiconductor, containing mobile electrons (the electron has a negative charge). Esaki succeeded in making the densities of holes and electrons in the respective regions extremely high by doping a large amount of the appropriate impurities with an abrupt transition from one region to the other. Now, in semiconductors, the conduction band for mobile electrons is separated from the valence band for mobile holes by an energy gap, which corresponds to a forbidden region. Therefore, a narrow transition layer from *n*-type to *p*-type, 5 to 15 nm thick, consisting of the forbidden region of the energy gap, provided a tunneling barrier. Since the tunnel diode exhibits a negative incremental resistance with a rapid response, it is capable of serving as an active element for amplification, oscillation, and switching in electronic circuits at high frequencies. The discovery of the diode, however, is probably more significant from the scientific aspect because it has opened up a new field of research—tunneling in solids. *See* BAND THEORY OF SOLIDS; CIRCUIT (ELECTRONICS); HOLE STATES IN SOLIDS; JUNCTION DIODE; NEGATIVE-RESISTANCE CIRCUITS; SEMICONDUCTOR; SEMICONDUCTOR DIODE; TUNNEL DIODE.

Esaki and colleagues have explored negative resistance phenomena in semiconductors which can be observed in novel tunnel structures. One obvious question is: What would happen if two tunnel barriers are placed close together, or if a periodic barrier structure—a series of equally spaced potential barriers—is made in solids? It has been known that there is a phenomenon called the resonant transmission. Historically, resonant transmission was first demonstrated in the scattering of electrons by atoms of noble gases and is known as the Ramsauer effect. In the above-mentioned tunnel structures, it is clear that the resonant tunneling should be observed. In preparing double tunnel barriers and periodic structures with a combination of semiconductors, the resonant tunneling was experimentally demonstrated and negative resistance effects were observed. *See* SEMICONDUCTOR HETEROSTRUCTURES.

Tunnel junctions between metals. As discussed above, tunneling had been considered to be a possible electron transport mechanism between metal electrodes separated by either a narrow vacuum or a thin insulating film usually made of metal oxides. In 1960, I. Giaever demonstrated for the first time that, if one or both of the metals were in a superconducting state, the current-voltage curve in such metal tunnel junctions revealed many details of that state. At the time of Giaever's work, the first satisfactory microscopic theory of superconductivity had just been developed by J. Bardeen, L. N. Cooper, and J. R. Schrieffer (BCS theory). Giaever's technique was sensitive enough to measure the most important feature of the BCS theory—the energy gap which forms when the electrons condense into correlated, bound pairs (called Cooper pairs).

The tunneling phenomenon has been exploited in many fields. For example, small-area tunnel junctions are used for mixing and synthesis of frequencies ranging from dc to the infrared region of the spectrum. This leads to absolute frequency measurement in the infrared and provides the most accurate determination of the speed of light. *See* PARAMETRIC AMPLIFIER.

To study nonequilibrium superconducting properties, two tunnel junctions, one on top of the other sharing the middle electrode, are used. One junction seems to inject quasi-particles, while the other detects their effects on the important parameters. Tunnel junctions are also used as a spectroscopic tool to study the phonon and plasmon spectra of the metals and the vibrational spectra of complex organic molecules introduced inside the insulating barriers (tunneling spectroscopy). *See* SPECTROSCOPY; SUPERCONDUCTIVITY.

Josephson effects. Giaever's work opened the door to more detailed experimental investigations—it pioneered a new spectroscopy of high accuracy to study the superconducting state. In 1962, B. Josephson made a penetrating theoretical analysis of tunneling between two superconductors by treating the two superconductors and the coupling process

as a single system, which would be valid if the insulating oxide were sufficiently thin, say 2 nanometers. His theory predicted, in addition to the Giaever current, the existence of a supercurrent, arising from tunneling of the bound electron pairs.

This led to two startling conclusions: the dc and ac Josephson effects. The dc effect implies that a supercurrent may flow even if no voltage is applied to the junction. The ac effect implies that, at finite voltage V, there is an alternating component of the supercurrent which oscillates at a frequency of 483.6 MHz per microvolt of voltage across the junction, and is typically in the microwave range. The dc Josephson effect was soon identified among existing experimental results, while the direct observation of the ac effect eluded experimentalists for a few years. The effects are indeed quantum phenomena on a macroscopic scale. Extraordinary sensitivity of the supercurrents to applied electric and magnetic fields has led to the development of a rich variety of devices with application in wide areas of science and technology. Superconducting quantum interference devices (SQUIDs) are made of one or more Josephson junctions connected to form one or more closed superconducting loops. Owing to their unprecedented sensitivity, SQUIDs are the main building blocks of many sensitive instruments such as magnetometers, power meters, voltmeters, gradiometers, and low-temperature thermometers. These are finding wide-range application in the fields of solid-state physics, medicine, mineral exploration, oceanography, geophysics, and electronics. Josephson junction and SQUIDs are used as switches for digital applications. They are the basic elements found in the picosecond-resolution sampling oscilloscope, as well as memory and logic circuits featuring high switching speed and ultralow power dissipation, in the order of 1 microwatt. In the communication field, they are used in analog applications, such as high-frequency local oscillators, detectors, mixers, and parametric amplifiers. Furthermore, the ac Josephson effect is now used to define the volt in terms of frequency in standards laboratories, eliminating the antiquated standard cell. *See* FUNDAMENTAL CONSTANTS; JOSEPHSON EFFECT; SQUID; SUPERCONDUCTING DEVICES. Leo Esaki

Bibliography. L. L. Chang (ed.), *Resonant Tunneling in Semiconductors: Physics and Applications*, 1992; L. Esaki, Long journey into tunneling, *Science*, 183:1149–1155, 1974; I. Giaever, Electron tunneling and superconductivity, *Science*, 183:1253–1258, 1974; H. Grahn, *Semiconductor Superlattices: Growth and Electronic Properties*, 1995; B. D. Josephson, The discovery of tunneling supercurrents, *Science*, 184:527–530, 1974.

Tupelo

A tree belonging to the genus *Nyssa* of the sour gum family, Nyssaceae. The most common species is *N. sylvatica*, variously called pepperidge, black gum, or sour gum, the authorized name being black tupelo. Tupelo grows in the easternmost third of the United States; southern Ontario, Canada; and Mexico. In moist soil this tree usually ranges from 60 to 80 ft (18–24 m) in height and 2 to 3 ft (0.6–0.9 m) in diameter, but some may be 110 ft (33 m) tall and 5 ft (1.5 m) in diameter. *See* MYRTALES.

Leaf, fruit cluster, bud, and twig of black tupelo.

The tree can be identified by the comparatively small, obovate, shiny leaves (see **illus.**); by branches that develop at a wide angle from the axis; and by a chambered pith. The fruit is a small blue-black drupe, a popular food for birds. The wood is yellow to light-brown and hard to split because of the twisted grain. Tupelo wood has been used for boxes, baskets, and berry crates, and as backing on which veneers of rarer and more expensive woods are glued. It is also used for flooring, rollers in glass factories, hatters' blocks, and gunstocks. *See* FOREST AND FORESTRY; TREE. Arthur H. Graves; Kenneth P. Davis

Turbellaria

A class of the phylum Platyhelminthes commonly known as the flatworms. These animals are chiefly free-living and have simple life histories. The bodies are elongate and flat to oval or circular in cross section. Their length ranges from less than 0.04 in. (1 mm) to several centimeters, but may exceed 20 in. (50 cm) in land planaria. Large forms are often brightly colored. Smaller forms may have black, gray, or brown parenchymal pigment or may be white or transparent except for the color of ingested food and symbiotic algae. This class, which numbers some 3400 described species, is ordinarily subdivided into 12 orders: Acoela, Catenulida, Haplopharyngida, Lecithoepitheliata, Macrostomida,

Nemertodermatida, Neorhabdocoela, Polycladida, Prolecithophora, Proseriata, Temnocephalida, and Tricladida.

Economics and ecology. Turbellaria are not economically important but have proved valuable in the study of such fundamental biological problems as regeneration, metabolism, axial gradients, evolution, and adaptations to parasitism. Although widely distributed in fresh and salt water and moist soil, they are usually overlooked because of their generally small size, secretive habits, and inconspicuous color. They are seldom eaten by other animals but frequently feed on one another and may harbor commensals and parasites, chiefly Protozoa and Nematoda. Some species are themselves parasites or commensals on other aquatic invertebrates.

Morphology and physiology. Locomotion is by gliding, swimming, or muscular movements of the body wall. Respiration takes place by diffusion through the body wall since respiratory organs are lacking. The cellular or syncytial epidermis is usually covered with cilia and may contain the nematocysts of ingested coelenterates. Gland cells are of frequent occurrence, producing adhesive substances, mucus, and two types of rod-shaped secretions or rhabdoids. These are the commoner rhabdites and the longer and slenderer rhammites. Usually there is a basement membrane beneath the epidermis, inside which lies an outer circular and an inner longitudinal muscle layer, sometimes with a diagonal layer between. The muscular system also includes both parenchymal and organ muscles. Between the body wall and internal organs lies the parenchyma, a more or less compact network of mesenchyme cells. Plasma in its interstices may function as a circulatory fluid.

Digestive system. The mouth, on the midventral surface, is generally followed by the pharynx which occurs in several forms. The simple pharynx has no limiting membrane and is only slightly eversible or protrusible. The plicate pharynx lies in a deep pouch, has a limiting membrane except at its base, and is highly protrusible. The bulbous pharynx lies in a shallow pouch and is surrounded by a limiting membrane. It is eversible and occurs in two chief forms, the cask-shaped or doliiform located at the anterior end of the intestine parallel to the main body axis, and the globular or rosulate lying ventral to the intestine and perpendicular to the main body axis. A short esophagus is often present and polypharyngy may occur. The intestine may be unbranched or a two, three-, or many-branched sac with or without diverticula. It is lined with tall epithelial cells and generally has no anus. Turbellaria are carnivorous; digestion is largely intracellular.

Excretion. The protonephridia are elongated tubules, usually paired, with ciliated flame bulbs on the lateral branches and one or more external openings. They are often lacking in marine forms and are probably primarily concerned with elimination of water.

Nervous system. In primitive forms the nervous system is an epidermal network with five pairs of longitudinal nerves connected by a nerve ring. Swellings at the intersection of nerve ring and nerves represent the beginnings of a brain. With further development, a pair of cerebral ganglia is formed by fusion of these swellings and the entire system sinks into the parenchyma. Posteriorly the longitudinal nerves are reduced to one well-developed ventral pair and often a much smaller dorsal pair. Intermediate conditions between these two extremes are common. Sensory receptors are located chiefly in the head region, although tactile bristles arise from widely scattered sensory cells. Chemoreceptors consist of depressed areas of epidermis with cilia for circulating water over the sensory surface. These are located in auricles, frontal organs, and ciliated rings, pits, and furrows. Statocysts or organs of equilibrium occur chiefly in primitive marine forms. Many species have one or two pairs of photoreceptors or eyes, and in land planarians and polyclads these may become numerous.

Reproductive system and reproduction. Virtually all Turbellaria are hermaphroditic (see **illus.**). Male and female systems may be separate throughout or may have a common antrum and pore. Genital pores are usually on the midventral surface, sometimes combined with the mouth opening. The saclike gonads are compact or follicular depending upon whether they consist of a few relatively large or several to many small bodies, but some acoeles have only scattered clusters of germ cells in the parenchyma. The male system has a single pair of sperm ducts which may fuse to form a seminal duct, enlarge to form spermiducal or storage vesicles, or empty into a true seminal vesicle. The copulatory organ is usually muscular, often encloses the seminal vesicle, may be armed

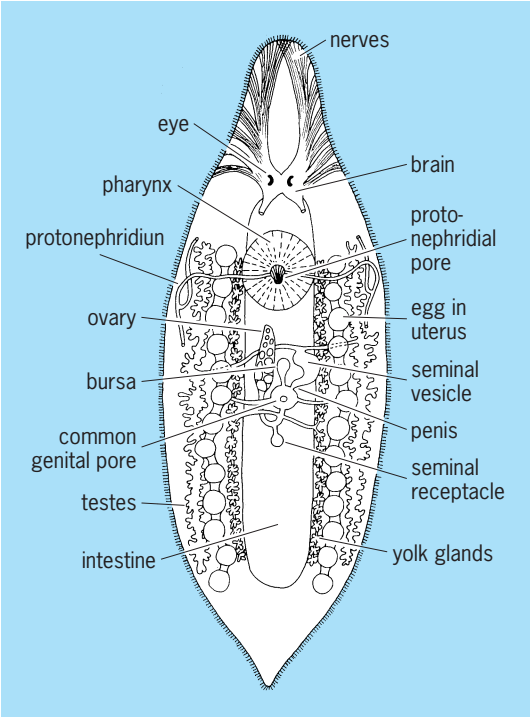

Typical hermaphroditic turbellarian, *Mesostoma ehrenbergii wardii*.

with a cuticular apparatus, and contains the ejaculatory duct. When prostate glands are present, their secretions may be stored in a prostatic vesicle but are mixed with the sperm in the seminal vesicle or in the male genital canal. Some Turbellaria produce entolecithal eggs which contain yolk, but generally the eggs are ectolecithal and yolk is derived from yolk cells. These yolk cells are probably degenerate ovocytes which may be produced throughout the gonad, in a specialized section of the gonad, or in special yolk glands which are distinct from the ovary. The female system usually has its own ducts through which its products reach the female antrum and ultimately the exterior. In the absence of ducts, fertilized eggs escape by rupture of the body wall or by way of the digestive system and the mouth. Accessory structures such as vaginae, copulatory bursae, uteri, seminal receptacles, and specialized glands may also occur. Asexual reproduction by fragmentation or by binary fission with the formation of temporary chains of zooids occurs in some rhabdocoeles and triclads. *See* ACOELA; MACROSTOMORPHA; POLYCLADIDA; TRICLADIDA. E. Ruffin Jones

Bibliography. E. Bresslau, *Turbellaria*, in W. Kükenthal (T. Krumbach, ed.), *Handbuch der Zoologie*, vol. 2, pt. 1, 1933; L. H. Hyman, *The Invertebrates*, vol. 2, 1951; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; W. D. Russell-Hunter, *A Life of Invertebrates*, 1979.

Turbidite

A bed of sediment or sedimentary rock that was deposited from a turbidity current. A turbidity current is an underwater flow produced by movement of a turbid mass of water downslope as a result of the excess weight of the turbid water as compared with the surrounding clear water. Turbidity currents are therefore a type of gravity (or density) current; such currents are kept in motion by gravity acting on relatively small differences in density between different fluids (gases or liquids) or between different parts of the same fluid mass. In some gravity currents the density difference results from a difference in temperature or salinity (for example, fresh water flowing above salt water), but in turbidity currents it results from the presence of dispersed sediment.

Turbidity currents. Many geologists would restrict the term turbidity current to flows in which the sediment dispersion is maintained by fluid turbulence (that is, in which the sediment is in true suspension). In this case, there is a need for terms to describe other types of gravity flows resulting from sediment which is in dispersion, but not as a result of the action of turbulence. The general class of such flows has been called sediment gravity flows, and there appear to be four main kinds: turbidity currents, in which sediment suspension is maintained by turbulence; subaqueous debris flows, in which the sediment grains are held in dispersion by the strength of the (generally muddy) matrix; subaqueous grain flows, in which sediment is dispersed by forces acting between the grains themselves (dispersive pressure) produced by shearing of a concentrated mass of grains; and liquefied sediment flows, in which the sediment is supported by a transient upward flow of pore fluids. Deposits of all four types of flows are generally closely associated, and it is not easy to reconstruct the exact mechanism of deposition from examination of the deposit. The term turbidite may therefore be used by some geologists in a broad sense to include deposits from all four types of sediment gravity flows.

The concept of turbidity currents as agents for the transportation and deposition of sandy (or coarser) sediment was introduced and developed mainly by P. H. Kuenen in the 1950s. Most geologists accept that a large part of the sediments both in the modern ocean basins and in the ancient stratigraphic record is composed of turbidites, even though direct observation of natural turbidity currents is extremely difficult. Evidence derived both from detailed topographic surveys of the modern continental slopes and rises, and from core samples of the surface sediments, indicates that sediment gravity flows are generated from time to time by failure of masses of sediment that have been deposited (in relatively shallow water) on the upper parts of submarine slopes. Slumping or liquefaction of sediments produces a mass of moving sediment which mixes with the surrounding water, accelerates as it moves downslope, and generally becomes a fully turbulent turbidity current. Most such currents move downslope in well-defined submarine canyons or channels, which have probably been eroded or constructed by a long series of previous flows. At the base of the slope, the turbidity currents leave the main channel to construct large submarine fans composed largely of turbidites, alternating with finer-grain sediments deposited by slow settling of mud from suspension (hemipelagic sediments). The most powerful turbidity currents may carry the finer sand and mud for many hundreds of kilometers along the sea floor, to construct the great flat submarine plains known as abyssal plains.

The most striking example of a turbidity current whose effects were actually recorded was the flow triggered by the earthquake that in 1929 struck the continental slope just south of the Grand Banks of Newfoundland. Cables crossing the North Atlantic were broken not only at the time of the earthquake near the epicenter, but also in sequence over a period of 12 h for a distance of several hundred kilometers to the south. The sequence of breaks indicates that the cables south of the epicenter were broken not by the effect of the earthquake itself, but by a large turbidity current which traveled downslope away from the epicenter at speeds of between 10 and 45 mi/h (5 and 20 m/s). Probably most turbidity currents are not of this magnitude, but it is not unusual to find individual turbidite beds that are more than a meter in thickness, and in several well-documented cases such beds have been traced laterally for distances of several tens of kilometers. *See* TURBIDITY CURRENT.

Characteristics. The term turbidite is fundamentally genetic and interpretive in nature, rather than being a descriptive term (like common rock names). Turbidites are clastic sedimentary rocks, but they may be composed of silicic grains (quartz, feldspar, rock fragments) and therefore be a type of sandstone, or they may be composed of carbonate grains and therefore be a type of limestone. A geologist's description of a rock as a turbidite is actually an expression of an opinion that the rock was deposited by a turbidity current, rather than being a description of a particular type of rock.

A combination of inference from laboratory experiments and field observations suggests that many turbidites display several, if rarely all, of a group of characteristics:

They show a sharp base and a gradational top, and there is not infrequently a progressive decline in mean grain size from the base to the top (graded bedding).

Examination of the exposed base of the bed (known as the sole) reveals a variety of structures (sole marks) that were originally produced by erosion of the mud over which the turbidity current moved. The marks cut by the head of the flow were then rapidly covered by sediment deposited by the body of the same current, and in some cases the marks have been distorted by sinking of sediment into the originally soft muds below the turbidite (to form load casts). Common types of sole marks include those scoured by the flow (scour marks), and of these the most characteristic form is scooplike, elongated parallel to the current that cut it, with the blunt end on the upcurrent side. Such structures are called flutes (or flute casts where preserved as molds on the sole of the turbidite that was deposited above them), and they are very useful for indicating the direction of flow of ancient turbidity currents. Also, presence of abundant scour marks suggests that the current which formed them was fully turbulent, and that the bed above was deposited by a turbidity current and not by some other type of (nonturbulent) sediment gravity flow. A second main type of sole marks are those that are cut by large fragments carried in the lower part of the flow (tool marks); the most common are long straight sets of grooves, parallel to the flow direction.

There is commonly a characteristic (Bouma) sequence of sedimentary structures within the bed (see **illus.**). Each characteristic structure (or group of structures) defines a division in the sequence and is identified by a letter: (a) massive (that is, no internal structures) except for some size grading; (b) plane lamination, parallel to the base of the bed (in many beds this lamination is faint and indistinct); (c) cross-lamination formed by migration of small-scale ripples, or convolute lamination formed by the distortion of ripple cross-lamination during, or soon after, its deposition; (d) an upper division of plane lamination, grading up into the next division; (e) muddy sediment (pelite). In most turbidites, only some of the Bouma divisions are represented, but the order in which the divisions occur is only very rarely reversed. Where one turbidite rests directly on another, several of the upper divisions of the first turbidite may have been eroded by the second turbidity current before the second turbidite was deposited, so that it appears at first as though there is only a single bed. Careful examination shows the erosion surface (and generally a size break between the turbidite below and the base of the one above). Such turbidites are said to be amalgamated into a single (composite) bed.

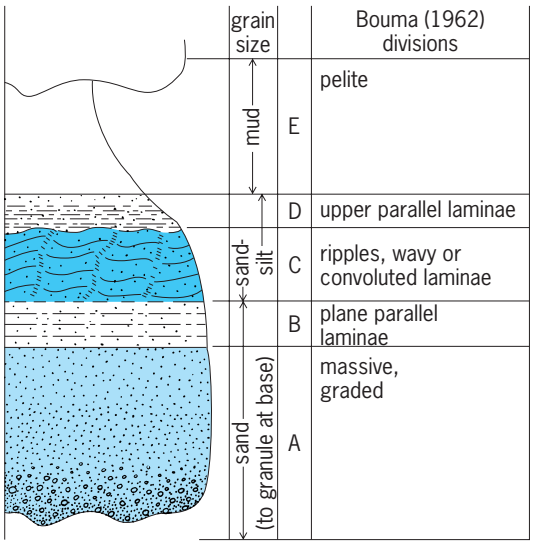

Ideal sequence of structures in a turbidite. (*After H. Blatt, G. Middleton, and R. Murray, Origin of Sedimentary Rocks, 2d ed., Prentice-Hall, 1980*)

Because many (if not all) turbidites have been very rapidly deposited, perhaps over a period of a few hours or days, they represent catastrophic events. Organisms originally living on the mud bottom may be swept away or buried. If buried, they may burrow to the surface (forming escape burrows), but thick turbidites are rarely extensively disturbed by burrowing organisms, even if the muds between them indicate that such organisms were common in the environment of deposition.

Sediment deposited rapidly is liable to liquefaction and subsequent expulsion of water to form fluid escape structures, and these are more commonly formed in turbidites than in other types of sediments.

The fossils found in turbidites are commonly broken, abraded, and displaced from shallower water, as compared with the fauna indigenous to the environment of deposition. In the case of geologically recent turbidites (Mesozoic and Cenozoic), the microfossils in the muds interbedded with turbidite beds not infrequently indicate great depths of deposition (several thousands of meters). Thick sand beds would normally be taken to indicate deposition in shallow water, so identification of the beds as turbidites is important for interpretation of the geological history of the deposits.

Identification. No single feature of a deposit is sufficient to identify it as a turbidite; all of the features

above have been described from other types of beds. A correct interpretation can be made only after analyzing all of the evidence together. Furthermore, it must be emphasized that not all turbidites are marine; there are well-documented examples of modern turbidity currents and turbidites described from lakes. Although probably most turbidites were originally deposited in water of considerable depth (hundreds to thousands of meters), it is generally difficult to be specific about estimating the depth of deposition. The most that can be said is that (in most cases) there is no sign of sedimentary structures formed by the action of waves.

Flysch. Turbidites and interbedded shales form large parts of the stratigraphic succession exposed on the continents. Many such thick turbidite-shale units were called flysch by geologists working in the era before plate tectonics (roughly before the mid-1960s). Some flysch units probably represent the deposits of ancient continental slopes and rises, but many are related to smaller basins of deposition formed during the early stages of plate subduction or continental collision, and a few may be fillings of deep-sea trenches. Many flysch formations have been strongly deformed and even metamorphosed subsequent to deposition. A few turbidite formations, however, were deposited along with suitable source beds in basins which were later raised to moderate temperatures 120–300°F (50–150°C), but which escaped strong deformation or metamorphism, and these have subsequently become prolific producers of oil or gas. Examples include several fields that produce from Cretaceous or Cenozoic sandstones in the Ventura and Great basins of California. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS; SEDIMENTARY ROCKS; SEDIMENTOLOGY.

G. V. Middleton

Bibliography. B. J. Skinner and S. C. Porter, *The Dynamic Earth: An Introduction to Physical Geology*, 5th ed., 2003; D. A. V. Stow (ed.), *Deep Water Turbidite Systems*, 1991; R. Tinterri et al. (eds.), *An Introduction to the Analysis of Ancient Turbidite Basins from an Outcrop Perspective*, 1999.

Turbidity current

A flow of water laden with sediment that moves downslope in an otherwise still body of water. The driving force of a turbidity current is obtained from the sediment, which renders the turbid water heavier than the clear water above. Turbidity currents occur in oceans, lakes, and reservoirs. They may be triggered by the direct inflow of turbid water, by wave action, by subaqueous slumps, or by anthropogenic activities such as dumping of mining tailings and dredging operations.

Turbidity currents are characterized by a well-defined front, also known as head, followed by a thinner layer known as the body of the current. They are members of a larger class of stratified flows known as gravity or density currents. A simple density current is driven by density differences produced by salinity or temperature. Salt and heat are conservative contaminants; they may be advected or diffused, but their total amount in the flow is conserved. Sediment is in general a nonconservative contaminant; it can be entrained from or deposited on the bed, thus changing the total amount of sediment in suspension. *See* DEPOSITIONAL SYSTEMS AND ENVIRONMENTS.

A turbidity current must generate enough turbulence to hold its sediment in suspension. If it is not able to do so, the sediment deposits and the current dies. The extra degree of freedom, that is, sediment entrainment and deposition, is what makes turbidity currents an interesting phenomenon. Under certain conditions, a turbidity current might erode its bed, pick up sediment, become heavier, accelerate, and pick up even more sediment, increasing its driving force in a self-reinforcing cycle akin to the formation of a snow avalanche.

Observations of turbidity currents were made in lakes and anthropogenic reservoirs long before their occurrence in the ocean became apparent. For many years, the only direct evidence of turbidity current activity in the ocean was the failure of submarine cables, such as the one associated with the Grand Banks earthquake of 1929. Following the earthquake, a large number of transoceanic telegraph cables were broken. At first, it was believed that the earthquake itself had caused the damage. However, 23 years after the earthquake an examination of the timing of the failures indicated that a turbidity current triggered by the earthquake had been responsible for the breakage.

Turbidity currents constitute a major mechanism for the transport of fluvial, littoral, and shelf sediments onto the ocean floor. These flows are considered to be responsible for the scouring of submarine and sublacustrine canyons. These canyons are often of massive proportions and rival the Grand Canyon in scale; they may be eroded directly into deltaic deposits, as in the case of the Rhone River delta in Lake Geneva, Switzerland, or into the continental shelf as in the case of Scripps and La Jolla submarine canyons in California. Only flows of substantial velocity could accomplish this excavation. The conditions under which the generation of such swift, canyon-scouring currents becomes possible have intrigued marine geologists and oceanographers for many years. Below the mouths of most canyons, turbidity currents form vast depositional fans that have many of the features of alluvial fans built by rivers and constitute major hydrocarbon reservoirs. The sedimentary deposits created by turbidity currents, known as turbidites, are a major constituent of the geological record. *See* MARINE GEOLOGY; MARINE SEDIMENTS; SUBMARINE CANYON; SUBMARINE FAN; TURBIDITE.

Marcelo H. Garcia

Bibliography. D. A. Edwards, *Turbidity Currents: Dynamics, Deposits and Reversals*, 1993; B. McCaffrey, B. Kneller, and J. Peakall, *Particulate Gravity Currents*, 2001; D. A. V. Stow (ed.), *Deep Water Turbidite Systems*, 1991.

Turbine

A machine for generating rotary mechanical power from the energy in a stream of fluid. The energy, originally in the form of head or pressure energy, is converted to velocity energy by passing through a system of stationary and moving blades in the turbine. Changes in the magnitude and direction of the fluid velocity are made to cause tangential forces on the rotating blades, producing mechanical power via the turning rotor.

The fluids most commonly used in turbines are steam, hot air or combustion products, and water. Steam raised in fossil fuel-fired boilers or nuclear reactor systems is widely used in turbines for electrical power generation, ship propulsion, and mechanical drives. The combustion gas turbine has these applications in addition to important uses in aircraft propulsion. Water turbines are used for electrical power generation. Collectively, turbines drive over 95% of the electrical generating capacity in the world. *See* ELECTRIC POWER GENERATION; GAS TURBINE; HYDRAULIC TURBINE; STEAM TURBINE; TURBINE PROPULSION; TURBOJET.

Turbines effect the conversion of fluid to mechanical energy through the principles of impulse, reaction, or a mixture of the two. For the impulse turbine, high-pressure fluid at low velocity in the boiler is expanded through the stationary nozzle to low pressure and high velocity. The blades of the turning rotor reduce the velocity of the fluid jet at constant pressure, converting kinetic energy (velocity) to mechanical energy. *See* IMPULSE TURBINE.

For the reaction turbine, the nozzles are attached to the moving rotor. The acceleration of the fluid with respect to the nozzle causes a reaction force of opposite direction to be applied to the rotor. The combination of force and velocity in the rotor produces mechanical power. *See* REACTION TURBINE.

Frederick G. Baily

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; H. P. Bloch, *A Practical Guide to Steam Turbine Technology*, 1995; M. P. Boyce, *Gas Turbine Engineering Handbook*, 3d ed., 2006; R. H. Kehlhofer, *Combined Cycle Gas and Steam Turbine Power Plants*, 2d ed., 1999; J. L. Kerrebrock, *Aircraft Engines and Gas Turbines*, 2d ed., 1992; E. Logan (ed.), *Handbook of Turbomachinery*, 2d ed., 2003.

Turbine engine subsystems

A typical aircraft gas turbine engine is composed of an assemblage of individual turbine engine subsystems and components which function together to provide thrust to propel the aircraft and power for its onboard functioning (**Fig. 1**).

Inlets. Virtually all operational aircraft gas turbine engines work with subsonic axial inlet Mach numbers at the fan or compressor inlet face (although there are experimental fans and compressors that involve supersonic flow into the compressor). It is the function of the inlet to intercept the stream tube that contains the airstream that will be ingested by the engine at any flight speed, accommodating any power level of the engine (that is, airflow requirement) and any flight attitude (such as a high yaw or pitch flight angle of attack), and to accelerate or diffuse that airstream with a minimum of pressure drop and distortion to the Mach number accepted by the fan or compressor or fan face (**Fig. 2**).

The inlet may involve a complex transition geometry where the intake is not a simple circle concentric with the engine; such a geometry is often used in turboprop engines with offset gearboxes. An additional complication is encountered in applications of engines embedded in aircraft fuselages with "cheek" intakes on either side of the aircraft. In this case, the two inlets must be joined to supply a single stable flow of air to the engine, even in the case of an aircraft in a highly yawed flight condition where one inlet may be partially blanked by separated flow from the aircraft's forebody. *See* TURBOPROP.

For low-speed aircraft such as helicopters, the inlet must accelerate the intake flow, so that the inlet is essentially a bellmouth. This generally allows the designer to provide a generous radius to the intake lip (Fig. 2*a*), which makes the inlet very tolerant of high pitch or yaw angles of attack. For aircraft with a very large range of flight speeds, ranging from takeoff to transonic cruise, the inlet must diffuse the high-speed intake flow down to the lower Mach number required by the fan or compressor. This generally requires a relatively smaller intake lip radius (Fig. 2*b*), which makes the system much more sensitive to off-design variation in airflow and to aircraft attitude.

The ultimate complication is encountered in supersonic flight, which requires a convergent/divergent diffuser to decelerate the supersonic intake flow to subsonic velocity. This may require a variable-area throat to swallow the high-loss strong shock that forms in front of the intake as the aircraft reaches supersonic flight speed. *See* SUPERSONIC DIFFUSER.

Inlet particle separators. Helicopters often land on, take off from, and hover over unprepared spaces and can churn up a dust storm that is sucked into the engine inlet. If allowed to enter the engine, this sand and dust will severely erode and damage the turbomachinery blading and will clog small air holes that supply vital cooling air to hot parts. Modern helicopter engines include inlet particle separators that pass the air through curvilinear passages in which the sand and dust particles are centrifuged and concentrated in a collection zone where they are extracted with a portion of the airflow by an ejector or a blower (**Fig. 3**). Turboprop engines may include primitive separator systems to capture ingested birds, hailstones, or other foreign objects that might damage the engine. *See* HELICOPTER.

Fans and compressors. The vast majority of fans and compressors (whose function is to increase the

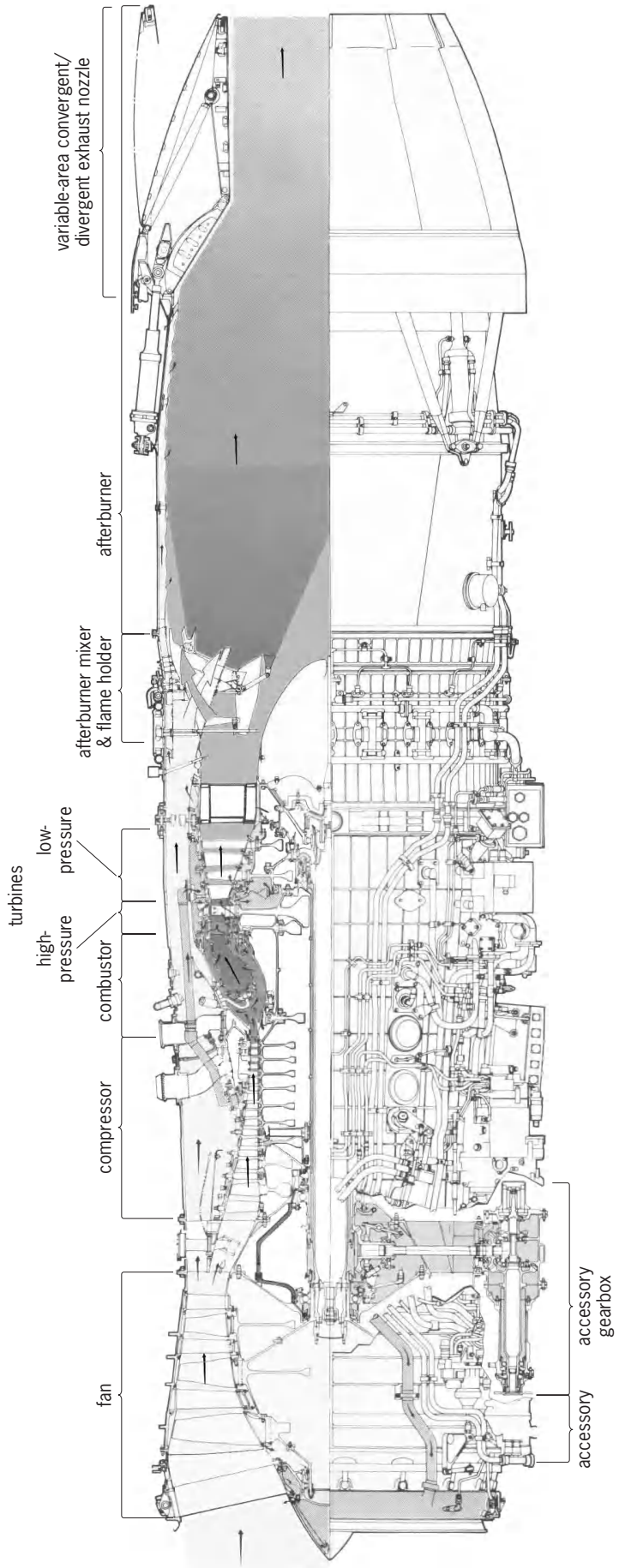

Fig. 1. Turbine engine subsystems and components of a typical aircraft gas turbine engine (afterburning low-bypass turbofan). (*Courtesy of GE Aircraft Engines*)

pressure of the incoming flow of air) of aircraft gas turbine engines are axial-flow turbo-compressors (Fig. 1). These devices consist of a series of individual stages, each comprising a rotating row (or cascade) of radially oriented airfoils (or rotor blades) followed by a stationary row of radially oriented airfoils (or stator blades). The rotating row of blades acts to increase the pressure and tangential velocity of the flow. The flow exiting the rotating row of blades is directed onto the stationary blade row which converts some of the tangential component of velocity imparted by the preceding rotor blade into a further increase in stage pressure rise.

Although the compressors of early jet engines were radial or centrifugal flow components, axial-flow compressors have been developed to achieve higher efficiency levels and lower frontal areas. They are therefore the prevailing approach in all jet engines except for the smallest sizes, where the axial-flow airfoils are not rugged enough for engine duty. In these small machines, single- and two-stage centrifugal compressors are found, as well as hybrid axi-centrifugal components consisting of one or more axial stages followed by a centrifugal stage.

Multistage compressors with pressure ratios above 5:1 or 10:1 (Fig. 1) involve extremely wide variation of flow conditions to the individual rows of airfoils in the front and the rear stages of the compressor as the speed and pressure ratios vary through the operating range. Such mismatch is aggravated by other operational circumstances such as nonuniform pressure distributions delivered by the inlet duct or rapid changes in the spool speed. Ultimately, the large angles of attack seen by individual airfoils will cause them to stall and result in unstable operation of the engine called surge. Such surge may cause flame-out or large pressure pulsations that result in physical damage to the engine. Various approaches have evolved to afford high-pressure compressors the required amounts of stall margin. A very high pressure compressor may be split into two sections—the intermediate section and the high-pressure section—each operating on a separate spool at its own speed and driven by its own turbine. An alternative or supplementary approach to design of high-pressure compressors is to provide pivots for each stator vane in forward stages of the compressor along with actuators that will vary the pitch angle of the vanes as a function of the spool speed and compressor inlet conditions to avoid the large angles of attack and the stall and surge that might ensue. *See* COMPRESSOR.

The fans of low-bypass turbofans (sometimes referred to as low-pressure compressors) are generally multistage axial-flow components, which do not differ materially from high-pressure compressors in concept or design (Fig. 1). On the other hand, medium- and high-bypass turbofans usually have a single-stage axial-flow compression stage, and may differ considerably in important respects from multistage compressors. In order to minimize the frontal area of the engines, fans are designed with a very low hub radius of the blades (compared to the blades' tip radius). They must also be designed to be much more rugged than other compressor blading in order to absorb the impact of birds, hailstones, and other foreign objects that may be drawn into the engine inlet. Fan blades in ultrahigh-bypass engines may incorporate provisions for varying their pitch in order to achieve thrust-reversing capability. *See* TURBOFAN.

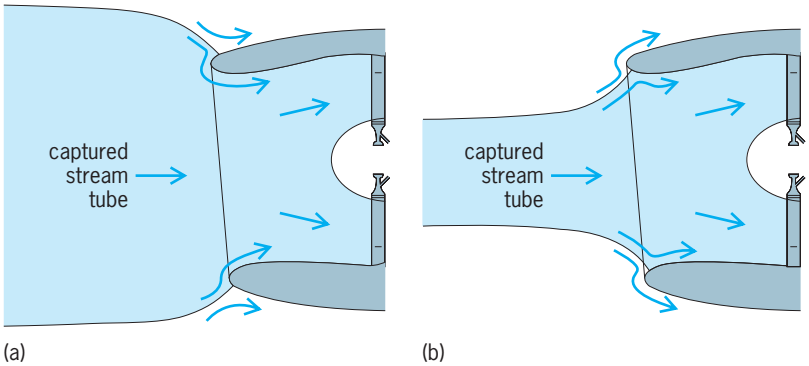

Fig. 2. Operation of a subsonic inlet at low and high flight speeds. (*a*) Flight speed lower than fan entry air velocity. (*b*) Flight speed higher than fan entry air velocity.

Combustors. The combustion chamber or combustor of the aircraft gas turbine engine (Fig. 1) provides for the burning of fuel in the airflow exiting from the compressor, and for supplying the resulting stream of high-temperature, high-pressure products of combustion to the turbine. In modern engines the combustor is usually an annular chamber, in the middle of the core between the compressor exit and the turbine entry, surrounding the center shaft of the rotor. Fuel is introduced at the upstream end through injectors, which may exploit any of several mechanisms to condition the fuel for the combustion process.

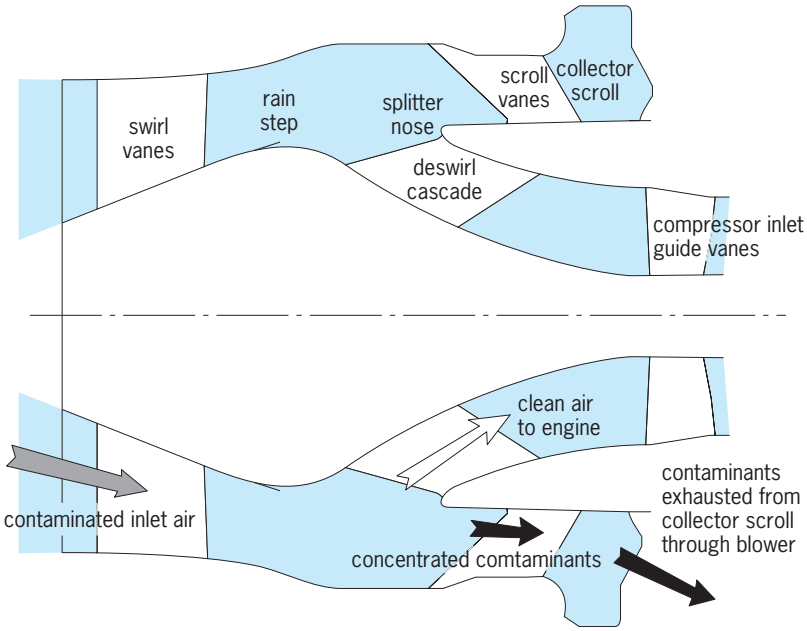

Fig. 3. Inlet particle separator. (*After M. G. Ray and J. L. Browne, T700 Engine Integral Inlet Separator, 6th European Rotorcraft and Powered Lift Forum, No. 37, University of Bristol, United Kingdom, 1980*)

With a high-pressure fuel or atomizing system, the fuel is introduced through very small orifices in a nozzle, generating a very fine spray. With an intermediate fuel pressure system, sometimes referred to as the air-blast injection system, the droplets may not be fine enough to permit efficient combustion, so various types of shear layers are introduced into the airstream in proximity to the injector to further break up the droplets. With a low-pressure fuel system and a vaporizing system, the injectors may be located in a hot region of the combustor so that the fuel is preheated and completely vaporized before entering the combustor. The combustion process is confined in a liner or shell, cooled by unburned air, which serves not only to contain the flame, but also to provide passages for air jets, lateral to the main direction of the flow, which promote and stabilize the flame. These lateral jets of unburned air also dilute its high-temperature zones to a lower and more uniform level of temperature that can be tolerated by the downstream turbine. One or more electric spark igniters are placed strategically in the chamber to light the flame in the course of starting the engine. Provision is made in the design for minimizing production of pollutants—carbon monoxide, oxides of nitrogen, hydrocarbon residues, particulates and smoke. *See* ATOMIZATION.

Turbines. Turbines in aircraft gas turbine engines extract mechanical energy from the stream of high-pressure, high-temperature airflow exiting the combustor. There may be several turbines in series in a typical engine, each driving an individual spool of the compressor ahead of it, or a fan component, or an external load like a propeller or helicopter rotor.

As shown typically in Fig. 1, each stage of turbines in modern aircraft gas turbine engines generally includes a stationary row (or cascade) of radially oriented blades or nozzle passages which accelerate and turn the incoming flow and direct it onto the following rotating row of radially oriented airfoils (or rotor blades). *See* TURBINE.

The modern gas turbine engine places enormous demands on virtually all the mechanical, aerodynamic, heat-transfer, metallurgical, and manufacturing process arts. The blades of the turbine rotate at velocities near the speed of sound, in a centrifugal acceleration field thousands of times that of gravity, in a chemically active transonic gas stream which is filled with abrasive particles and whose temperature may be well above the melting temperature of the material of which the blades are made. The blades and other parts of the turbine must also endure thousands of cycles of large changes in temperature and stress as the engines are started and shut down and as power is varied during the course of their normal operation.

There are two approaches to coping with the extremely high temperatures: by means of heat-resistant materials, and by cooling the blade with lower-temperature air. Jet engine needs have provided the incentive for a continuing progression of improved high-temperature materials, starting with steel alloys, followed by generations of cobalt and nickel alloys, and enhanced by treatments such as dispersion strengthening and by directional solidification and monocrystal casting processes. More recent developments include intermetallic alloys and even nonmetallic materials such as ceramics and composite compositions of high-temperature phases of carbon. *See* HIGH-TEMPERATURE MATERIALS.

The turbine blades are also cooled by air extracted from the compressor discharge flowpath, bypassing the combustor, and brought aboard the turbine rotor. The air is introduced into small circuitous passages cast into the turbine blades which are designed to promote conductive, convective, impingement, or transpiration cooling of the blades. Some of the spent cooling air is then conducted through small passages through the surface of the blade, and it is directed to provide film cooling over critical surface areas of the blade (**Fig. 4**). In spite of this challenging environment and duty, the refinement of all these technologies has made possible durable turbines that operate safely and reliably, and last thousands of hours

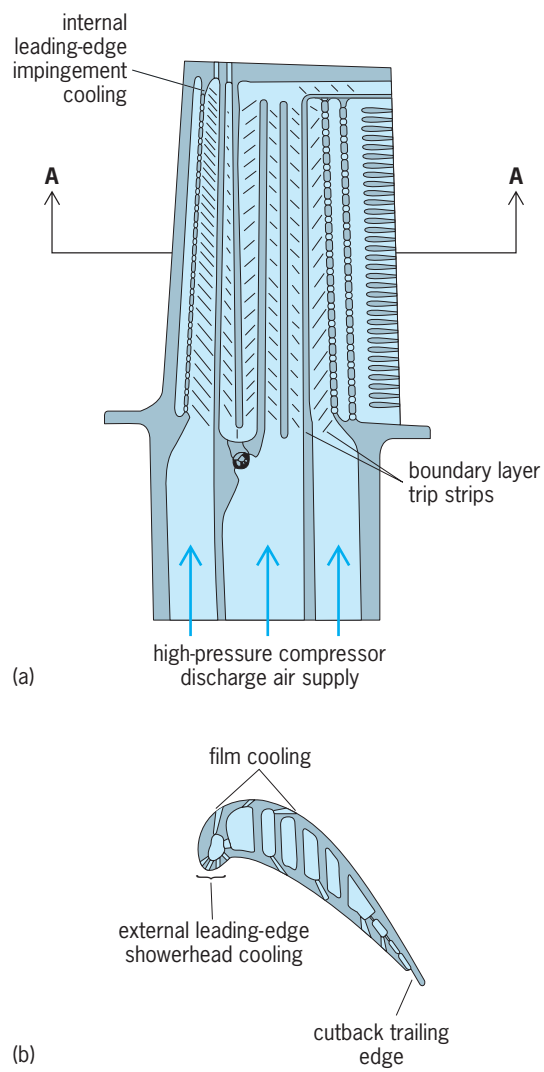

Fig. 4. Typical cooling scheme for a high-temperature turbine blade. (*a*) Plan view. (*b*) Cross-section view A–A. (*Courtesy of Pratt and Whitney*)

without overhaul or replacement, in many cases, several years "on-the-wing."

Regenerators and recuperators. A regenerator is a heat exchanger used in a gas turbine engine to transfer heat from the waste heat in the exhaust stream to the air being discharged from the compressor, which enters the combustor. The net effect is to reduce the amount of fuel required in the combustor with a net improvement in the fuel economy of the engine. Although regenerators are common in stationary and surface vehicle gas turbines, they have not yet proven useful in aircraft propulsion because the weight and bulk of current regenerators cause more deterioration in the mission performance of the aircraft than can be compensated by the possible reduction in fuel consumption.

Jet nozzles and variable exhaust nozzles. Simple jet engines and turbofan engines, which do not have an external propulsor such as a propeller, derive their thrust from the acceleration of the airflow through the turbine engine. This dictates that the energy generated in the engine as high-pressure, high-temperature airflow must be used to accelerate the airflow through an exhaust nozzle at the rear of the engine. The shape of the nozzle is very much a function of the pressure ratio across the nozzle (that is, the ratio of the pressure delivered by the engine to the nozzle, to the ambient pressure to which the exhaust is delivered). When the pressure ratio is less than the critical value (the value required to drive the stream to sonic velocity, approximately 1.9), the nozzle may have a simple convergent or conical shape. For engine operating conditions where the nozzle pressure ratio is greater than the critical value, the nozzle must be configured with an area distribution that initially converges and then diverges to accommodate the transition to supersonic conditions in the exhaust stream. In circumstances where there is a large variation of volumetric airflow from the engine in the course of its range of power level and environmental operation, the nozzle throat area may be designed to be variable to accommodate the range of operation most efficiently. *See* NOZZLE.

Afterburners. As illustrated in Fig. 1, an afterburner may be added to low-bypass turbofans and turbojets as a means of achieving thrust augmentation by burning fuel in the exhaust of the engine before it exits the engine through the exhaust nozzle. Although afterburners are relatively lightweight and are capable of augmenting thrust from 30% to 50%, they are extremely inefficient in fuel usage and are applied only in circumstances requiring short-duration bursts of thrust augmentation for extreme conditions such as takeoff, transonic acceleration in supersonic aircraft, and combat maneuvers in military aircraft.

The main components of an afterburner are a mixer (in turbofan engines) to mix the low-temperature bypass stream with the high-temperature core engine exhaust; the fuel injectors, which spray the liquid hydrocarbon fuel in the upstream end of the burner; the flame holders, bluff bodies whose wakes stabilize the location of the flame front; a shell liner, which provides containment of a film of cooler air between itself and the afterburner casing, and may also include a perforated panel whose purpose is to suppress combustion instability known as screech; and a variable exhaust nozzle, which is capable of adjusting the nozzle throat area and the exhaust area to accommodate the considerable changes in volumetric flow associated with the variation of the

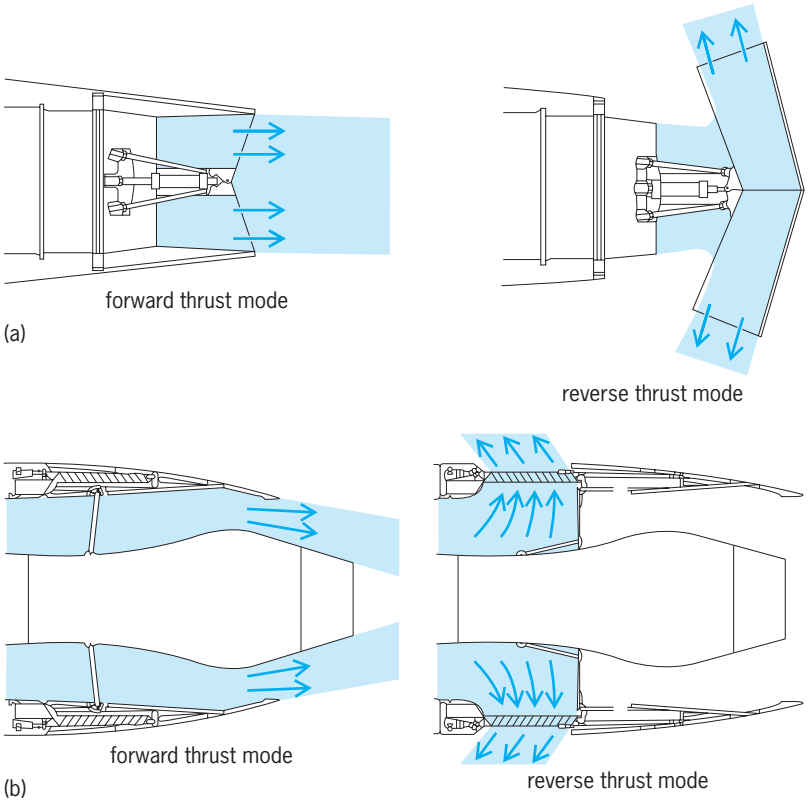

Fig. 5. Typical aircraft gas turbine thrust reversers. (*a*) Target-type reverser, typically applied to jet engines. (*b*) Blocker-door-type reverser, typically applied to turbofan engines. (*After M. J. T. Smith, Aircraft Noise, 2d ed., Cambridge University Press, Cambridge, United Kingdom, 2004*)

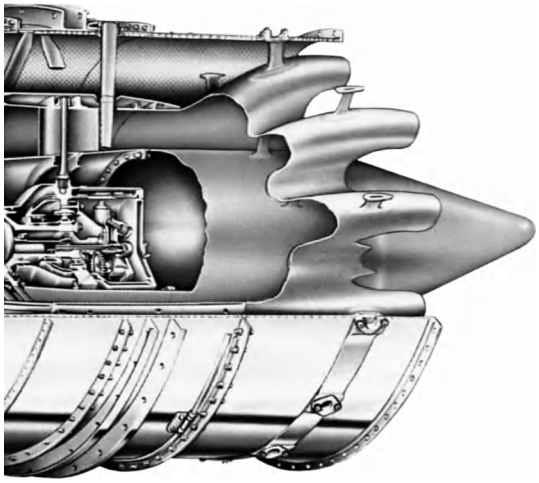

Fig. 6. Exhaust mixer in low-bypass turbofan for noise suppression and improved propulsive efficiency. (*Courtesy of Pratt & Whitney*)

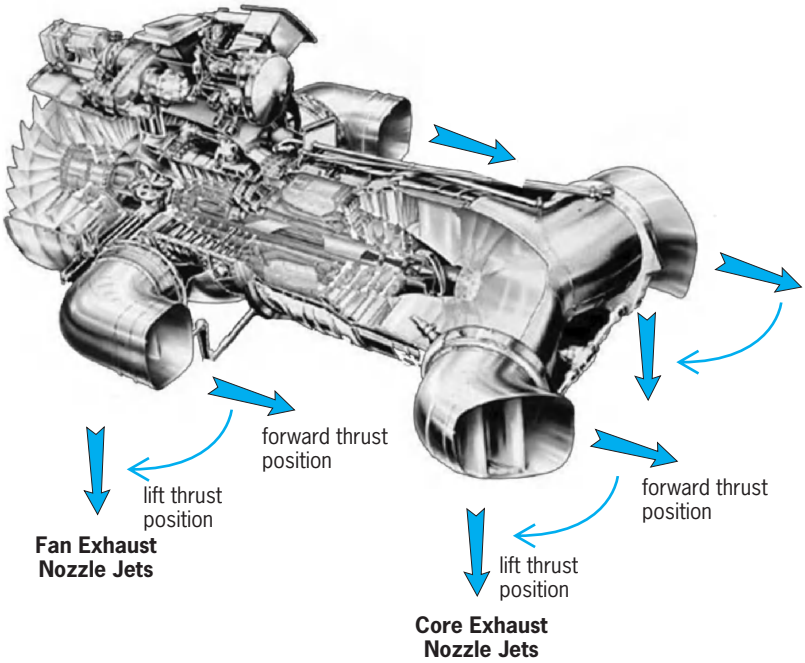

Fig. 7. Turbofan engine with swivel nozzles in both fan and core exhaust for vertical takeoff and landing. (*Courtesy of Rolls Royce*)

gas temperature rise in the afterburner (Fig. 1). *See* AFTERBURNER.

Thrust reversers. Aircraft gas turbine engines may include provision for reversing the direction of flow of the airstream handled by the propulsor in order to provide braking thrust for the aircraft, primarily during its landing roll on very short runways or runways with icy surfaces. For turboprops and ultrahigh-bypass engines, this is usually accomplished by varying the pitch angle of the propeller or of the fan blading, so that the propulsor pumps the air in the direction opposite to that required for forward flight. Simple jet engines may include a target reverser, in which a barrier is deployed in the jet stream exiting from the rear of the engine to redirect the jet in a direction with a forward component of velocity (**Fig. 5***a*). For high- and medium-bypass turbofans, this is more easily accomplished by introducing blocker doors into the fan discharge bypass stream, and simultaneously opening up alternative passages that reverse the air to a direction opposite to the flight direction (Fig. 5*b*). A retarding thrust of the order of 40% of the normal engine thrust is typically obtained by these devices.

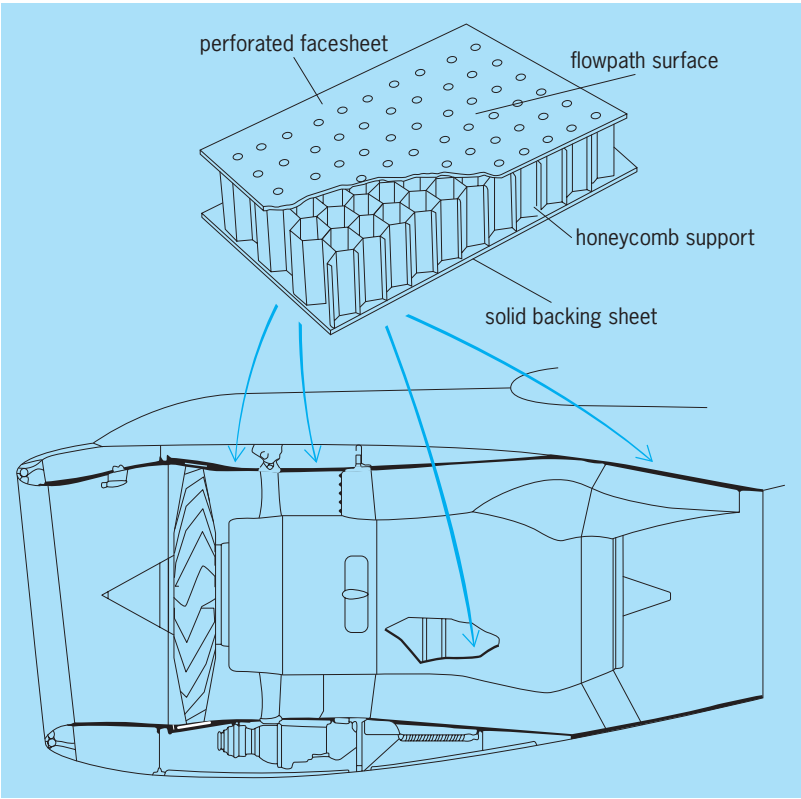

Fig. 8. Duct lined with hexagonal-cell honeycomb Helmholtz resonator cavities for turbomachinery noise suppression. (*After M. J. T. Smith, Aircraft Noise, 2d ed., Cambridge University Press, United Kingdom, 2004*)

Mixers and noise suppressors. Jet noise generally involves a broad frequency band and is dealt with by providing a highly convoluted exhaust nozzle, making a transition from the annular duct of the turbomachinery exit to a "daisy" or a multijet configuration with a extended perimeter (**Fig. 6**). The action of the nozzle's extended perimeter is to intensify the mixing and hence to shorten the length of the noise-generating shear layer between the jet and the lower-speed airflow external to the nozzle. Such suppressors were very common in earlier generations of subsonic transports that used turbojet engines. The action and the configurations are quite similar to the mixers that are used to mix the core jet with the bypass jet in mixed-flow turbofan engines to provide for noise suppression as well as the propulsive efficiency increase associated with eliminating any substantial difference between the velocities of the core and the bypass streams. They are no longer required on the very high bypass turbofans, with their much lower jet velocities, used in recent generations of subsonic transports. Supersonic transport engines use low-bypass turbofans or turbojet engines with much higher jet velocities that require such suppression devices. Mixers are also used in low-bypass afterburning turbofan engines to mix most of the bypass stream with the core exhaust to assure that the bypass stream will be burned in the afterburner (Fig. 1). *See* AERODYNAMIC SOUND; TURBOJET.

Thrust deflectors. To achieve increased agility and maneuverability, the low-bypass turbofans or turbojet engines of military combat aircraft may include provision for mechanically deflecting the jet nozzle from the direction of flight to a small angle upward or downward to provide a component of thrust in the direction of climb or descent. The nozzle may also be provided with an additional degree of freedom to deflect the exhaust to either side to provide the aircraft with increased capability to execute sharp turns. In a more extreme application, the thrust deflectors may be designed to deflect the exhaust stream to a full 90° or more to facilitate vertical takeoff and landing (**Fig. 7**). *See* VERTICAL TAKEOFF AND LANDING (VTOL).

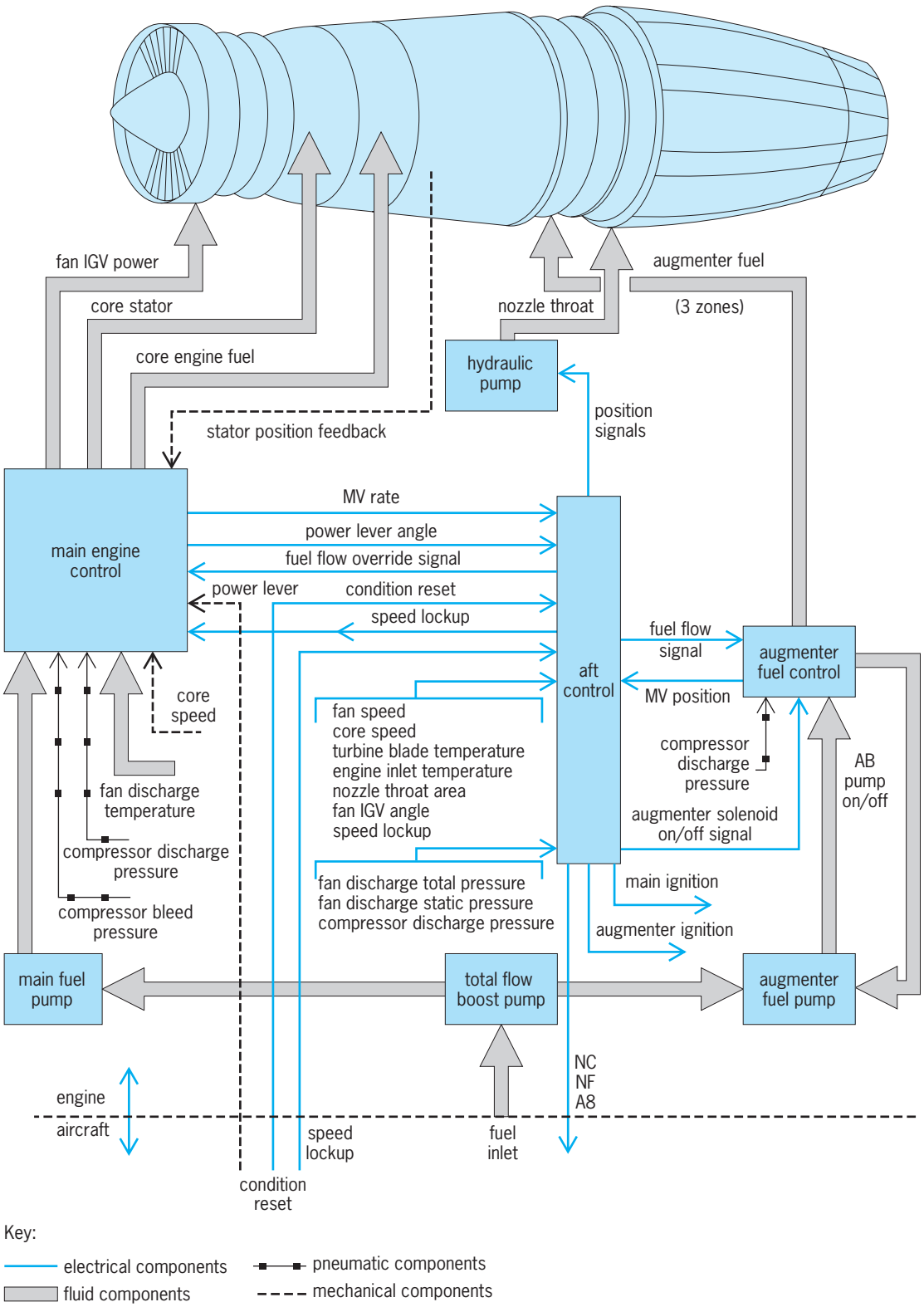

Fig. 9. Schematic of typical aircraft gas turbine engine control system. IGV = inlet guide vane; MV = motorized valve; AB = afterburner; NC = core spool speed signal; NF = fan spool speed signal; A8 = exhaust nozzle area signal. (*After T. W. Fowler, Jet Engine Propulsion Systems for Engineers, GE Aircraft Engines, Cincinnati, OH, 1989*)

Turbomachinery noise suppressors. In high-bypass-ratio turbofan engines where the exhaust jet noise is no longer dominant, turbomachinery noise becomes intrusive and must be suppressed. The generated noise is generally dominated by pure tones (blade passing frequency and its whole-number multiples). It may be suppressed by panels which include a large number of tuned Helmholtz resonating cavities that

are embedded in the walls of inlet and exhaust ducts (**Fig. 8**). *See* HELMHOLTZ RESONATOR.

Infrared suppressors and radar cross-section reduction. A key requirement of modern military aircraft engines is that they minimize their emission of infrared energy from their hot parts, and their reflection of radar microwave energy by which the aircraft might be detected and tracked by hostile missiles. Infrared suppressors are basically convoluted exhaust ducts that shield the line-of-sight of the hot turbine from any view, and provide for cooling the duct walls enough so that they themselves do not radiate detectable levels of energy.

Radar cross-section reduction also may involve convoluted ducts that shield the line-of-sight of the face of the engines' turbomachinery, which might otherwise act as a large reflector of radar microwave energy. The ducts are coated with substances that absorb microwaves so that they themselves do not contribute to the radar reflectivity of the aircraft. *See* MILITARY AIRCRAFT; RADAR-ABSORBING MATERIALS.

Accessory drives and accessories. In order to provide power to rotate an engine's accessories, a bevel gear set is generally provided at the front end of the core engine which drives a radial shaft through one of the front frame struts to an external gearbox (Fig. 1). The external gearbox includes pads for mounting and driving the accessories. Typical driven accessories might include an electric generator-starter, tachometer, fuel pump, hydraulic pump, and lubrication supply and scavenge pumps.

Control systems. A key auxiliary function that must be performed in any engine is control. The control system (**Fig. 9**) is centered on the main fuel control, which accepts a variety of signals from sensors situated around the engine that provide data indicative of the status of the engine's operation (such as measurements of key pressures and temperatures within the engine, rotative speeds of the engine's spools, actual positions of the engine's variable geometry, torque in the engine's output shaft, and rates of fuel flow to the engine's combustion systems); data indicative of the aircraft's operating conditions; and a signal from the pilot's throttle or power demand lever. On the basis of these data and of a complex program, the control energizes actuators that manipulate fuel valves and other variable geometry of the engine (such as variable-pitch stators, variable-area and vectoring exhaust nozzles, and thrust reversers) to provide stable power at the demanded level and, when requested, provide smooth and fast changes in power level. The control must also provide for starting and shutdown of the engine; must protect the engine from surge, overspeed, overtemperature, and overtorque; and in the event of any failure, must provide for residual power or for safe shutdown. Early generations of engines used hydromechanical controls, but modern systems are generally based on digital electronic devices. A starting system (a hydraulic, pneumatic, or electrical motor) is mounted on the core spool's accessory gearbox to rotate the spool up to a speed where it is pumping enough air to permit combustion and is able to generate enough energy to be self-sustaining. The fuel system includes the fuel pumps, regulating valves, flow dividers, and fuel nozzles. The ignition system includes a high voltage unit and igniter. The anti-icing system provides hot air or electrical heating to parts in the engine inlet which might otherwise become clogged with ice during flight through supercooled moisture in the atmosphere. Electrical, hydraulic, or pneumatic actuators drive the engine's variable geometry such as variable-pitch compressor stator blades and variable-area exhaust nozzles. *See* CONTROL SYSTEMS.

Fredric F. Ehrich

Bibliography. J. H. Horlock, *Axial Flow Compressors: Fluid Mechanics and Thermodynamics, Butterworths*, London, 1958, Krieger Publishing, Malabar, FL, 1973; J. H. Horlock, *Axial Flow Turbines: Fluid Mechanics and Thermodynamics, Butterworths*, London, 1966, Krieger Publishing, Malabar, FL, 1973; J. L. Kerrebrock, *Aircraft Engines and Gas Turbines*, 2d ed., 1992; M. Kroes et al., *Aircraft Powerplants*, 7th ed., 1995; J. D. Mattingly, W. H. Heiser, and D. T. Pratt, *Aircraft Engine Design*, 2d ed., AIAA, Reston, VA, 2002; G. C. Oates, *Aircraft Propulsion Systems: Technology and Design*, 1989; Rolls Royce plc, *The Jet Engine*, 5th ed., Derby, United Kingdom, 2005.

Turbine propulsion

Propulsion of a vehicle by means of a gas turbine. While gas turbines have found significant applications in stationary and mobile power plants, their light weight, compact volume, low frontal area, and long life (in comparison to reciprocating diesel or Otto-cycle internal combustion engines) make them ideal for primary propulsion of vehicles.

Since about 1940, gas turbines have come to dominate most areas of common carrier aircraft propulsion, have made significant inroads into the propulsion of surface ships, and have been incorporated into military tanks. Turbine propulsion has been a subject of development activity for application to railroad locomotives and to buses, trucks, and automobiles.

Core or gas generator. The primary power producer common to all gas turbines used for propulsion is the core or gas generator, operating on a continuous flow of air as working fluid. The air is compressed in a rotating compressor, heated at constant pressure in a combustion chamber burning a liquid hydrocarbon fuel, and expanded through a core turbine which drives the compressor. This manifestation of the Brayton thermodynamic cycle generates a continuous flow of high-pressure, high-temperature gas which is the primary source of power for a large variety of propulsion schemes. The turbine is generally run as an open cycle; that is, the airflow is ultimately exhausted to the atmosphere rather than being recycled to the inlet. *See* BRAYTON CYCLE.

For gas generators using a very high pressure, the compressor and turbine may each be divided into two subunits to make a two-spool gas generator.

A heat exchanger may be incorporated in a midstation in the compressor as an intercooler to improve the cycle efficiency. A heat exchanger may be installed as a regenerator or recuperator to transfer heat from the airflow ultimately exhausted from the engine to the airflow entering the combustor, thereby decreasing the amount of hydrocarbon fuel needed to produce a specific temperature of the gas. This also improves the cycle efficiency. *See* HEAT EXCHANGER.

Water, or a water-methanol mixture, may be injected in the airstream upstream of the combustor to augment the available power from the gas generator.

Propulsion. The residual energy available in the high-temperature, high-pressure airstream exiting from the core is used for propulsion in a variety of ways.

For traction-propelled vehicles (buses, trucks, automobiles, military tanks, and most railroad locomotives), the core feeds a power turbine which extracts the available energy from the core exhaust and provides torque to a high-speed drive shaft as motive power for the vehicle. With a free-turbine arrangement, this power turbine is a separate shaft, driving at a speed not mechanically linked to the core speed. With a fixed turbine, this power turbine is on the same shaft as the core turbine, and must drive at the same speed as the core spool. In traction vehicles the power turbine generally drives through a transmission system (that is, a gear system or hydraulic pump-motor set or electrical generator-motor set) which affords a constant- or a variable-speed reduction to provide the necessary torque-speed characteristics to the traction wheels. *See* AUTOMOTIVE TRANSMISSION.

Aircraft, ships, and high-speed land vehicles, which cannot be driven by traction, are propelled by reaction devices. Some of the ambient fluid around the vehicle (that is, the water for most ships, and the air for all other vehicles) is accelerated by some turbomachinery (a ship propeller, aircraft propeller, helicopter rotor, or a fan integrated with the core to constitute a turbofan engine). The reaction forces on this propulsion turbomachinery, induced in the process of accelerating the ambient flow, provide the propulsion thrust to the vehicle. In all these cases, motive power to the propeller or fan is provided by a power turbine extracting power from the gas generator exhaust. In the case of a jet engine, exhaust from the gas generator is accelerated through a jet nozzle, so that the reaction thrust is evolved in the gas generator rather than in an auxiliary propeller or fan. Indeed, in turboprop and turbofan engines, both forms of reaction thrust (from the stream accelerated by the propeller or fan and from the stream accelerated by the core and not fully extracted by the power turbine) are used for propulsion (**Fig. 1**). *See* JET PROPULSION; TURBOFAN; TURBOJET; TURBOPROP.

Propulsive efficiency. The efficiency of the energy generation in a turbine power plant is a function of the pressure ratio of the compression process (including the ram pressure induced in a high-speed vehicle by decelerating the ambient air to bring it aboard the vehicle), the maximum temperature of the cycle, the efficiency of compression and expansion of the turbomachinery components, and the losses associated with pressure drops in static components, leakages, and parasitic air losses.

In reaction propulsion the efficiency of the propulsion process is also a strong function of the velocity at which the propulsive exhaust jet exits from the engine. The propulsive efficiency is given approximately by Eq. (1), where a is the ratio of exhaust

$$\eta_p = \frac{2}{1 + a} \tag{1}$$

velocity (relative to the vehicle) to the vehicle's air speed. Maximum propulsive efficiency is achieved when the exhaust velocity is equal to the air speed or $a = 1$. This ideal condition can be approached, but only at the expense of making the propulsor very large and heavy, since the thrust F_n per unit airflow W_a is found to be approximately as given by Eq. (2),

$$\frac{F_n}{W_a} = V_0(a - 1) \tag{2}$$

implying that values of a near unity give vanishingly small thrust per unit airflow. Equation (2) indicates that the thrust per unit airflow is proportional to flight speed V_0, so that a relatively large-mass flow must be handled at low speed. *See* THRUST.

In the design of a turbine engine for reaction propulsion, this balance of considerations is handled by a spectrum of propulsion systems illustrated in Fig. 1. For very low flight speeds, very large propulsors are used to accelerate large amounts of air through small velocity increments, only slightly above that low flight speed, as typified by the helicopter. High-speed propulsion systems are designed to accelerate smaller amounts of air to much larger velocities, as typified by the pure jet engine. Turboprops and turbofans constitute intermediate devices in this propulsion spectrum.

Aircraft propulsion. Except for reciprocating engines for small general aviation airplanes and helicopters with less than 700 hp (520 kW), turbine engines dominate the entire spectrum of aircraft propulsion. *See* AIRCRAFT ENGINE; AIRCRAFT PROPULSION; RECIPROCATING AIRCRAFT ENGINE.

Helicopters are generally powered by one, two, or three turboshaft engines, with the output power shafts feeding a combining gearbox which drives the helicopter rotor. The engines generally have free turbines, since the rotor must be able to operate at constant rotating speed, independent of the amount of power required. The engines are fitted with special controls which automatically hold output speed and share power evenly among the engines in multiple-engine installations. *See* HELICOPTER.

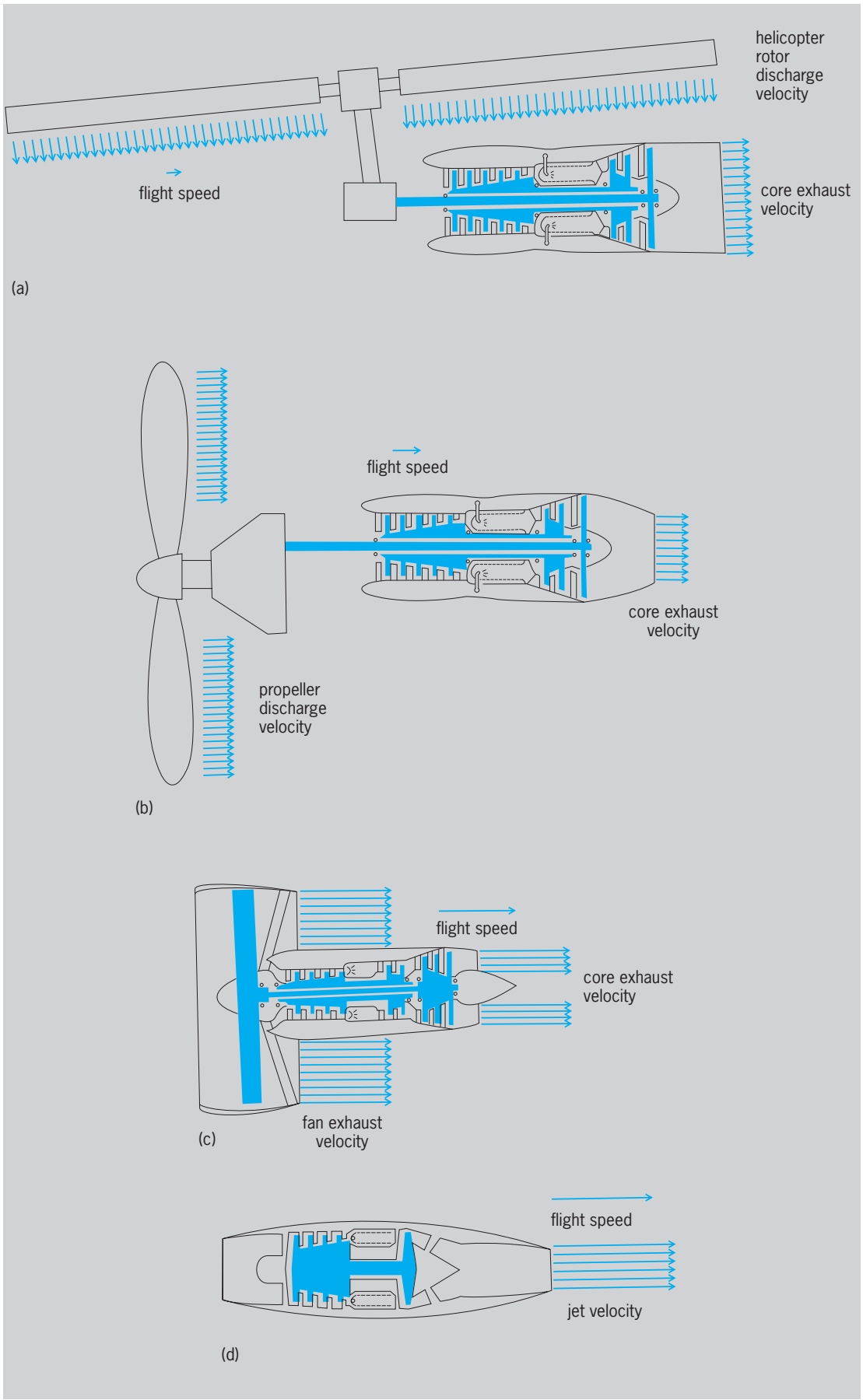

Fig. 1. Aircraft turbine propulsion systems. (*a*) Helicopter turboshaft. (*b*) Turboprop. (*c*) High-bypass turbofan. (*d*) Turbojet.

Vertical/short takeoff and landing (V/STOL) aircraft constitute a relatively new and evolving class of aircraft with an enormous variety of operational, experimental, and proposed propulsion systems. They have two categories of propulsion: conventional thrusters to propel the aircraft in forward flight, and vertical thrusters providing lift to support the aircraft in vertical and low-speed flight when the wings are not providing lift. The vertical thrust may be provided by one or a combination of schemes: deflection of the jet from the conventional thrusters; physical rotation of the nacelle or a tilt wing in which the conventional thruster is mounted; addition of a helicopter rotor to a conventional aircraft (in a convertible aircraft); addition of auxiliary vertically mounted direct-lift engines; addition of auxiliary vertically mounted fans, powered by separate engines or by shafting or compressed air drawn from the forward thrusters; and so forth. In the vertical flight mode the aircraft control and stability are dependent on differential modulated thrust from the several lifting thrusters, since the conventional aircraft control surfaces are ineffective without forward flight speed. *See* CONVERTIPLANE; SHORT TAKEOFF AND LANDING (STOL); VERTICAL TAKEOFF AND LANDING (VTOL).

Conventional takeoff and landing aircraft are generally equipped with turboprops for low subsonic flight speeds up to a Mach number of 0.7. For higher flight speeds through the transonic range, high-bypass turbofans are generally used. Hybrid engines, encompassing the advantages of the lower weight and size of the turbofan and the higher propulsive efficiency of the turboprop, are another alternative. One category is the prop-fan, where high-speed, low-diameter multibladed, counterrotating propellers are used, often in a pusher configuration appropriate to tail mounting. Although most proposed configurations use a gearbox to drive the propeller, the unique unducted fan type manages to accomplish propulsion without the use of a gearbox. Another configuration in the class of ultrahigh-bypass engines employs a shrouded fan in the conventional position before the power producer, but with a bypass ratio of 10 or more, which generally requires turboproplike features such as a gearbox drive and a variable-pitch fan to achieve thrust reversal. For propulsion at very high transonic and low supersonic flight speeds, very low-pass turbofans are generally used in which engine thrust augmentation is provided for short-duration acceleration or, in military aircraft, for combat maneuvers. Additional developments in advanced combat aircraft include variable-cycle features to accommodate extreme mixes of subsonic and supersonic duty, thrust vectoring to achieve extremes of aircraft maneuverability, and features to suppress the infrared and radar cross sections of the engine. Aircraft such as supersonic transports that require sustained flight at high supersonic speeds have been equipped with afterburning turbojet engines. *See* AFTERBURNER; MILITARY AIRCRAFT.

Ship propulsion. Turbines have also found usage in high-speed ships, where their light weight, compact envelope, fast starting, transient response characteristics, adaptability to a wide spectrum of liquid fuels, and long life are particularly suited.

Virtually all types of ships are powered by turboshaft engines which drive propellers through geared speed reducers. Many of the turbines are derived from aircraft engines: helicopter turboshafts, or turbofan engines with the fan removed, or jet engines with the addition of a power turbine. These engines are made suitable for the marine environment by rerating for the appropriate duty cycle, substitution of materials and coatings for protection from the seawater environment, addition of inlet systems to exclude ingestion of seawater, addition of appropriate exhaust stacks, provision of bases for floor mounting, addition of enclosures for acoustic isolation and thermal environmental control, and addition of starting and control systems unique to ship propulsion. *See* MARINE ENGINE.

In conventional hull- or buoyancy-borne ships, turbines are found in the high-speed end of the vehicle spectrum and, indeed, are fast coming to dominate the modern naval fleets of the world. *See* NAVAL SURFACE SHIP.

Hydrofoil ships generally have two separate power plants: small units to power the ship propellers at low forward speeds when the craft is hull-borne, and larger units to power the ship propellers or water jet propulsion units during high-speed, foil-borne operation. *See* HYDROFOIL CRAFT.

Water-borne surface-effect ships or amphibious air-cushion vehicles have two classes of turbine propulsion: prime movers to power fans which provide the air cushion to support the vehicle, and prime movers (effectively turboprops) which power air propellers to propel the vehicle. In certain installations the same turbine may power both functions. *See* AIR-CUSHION VEHICLE; SHIP POWERING, MANEUVERING, AND SEAKEEPING.

Locomotive propulsion. Over the years there has been considerable experimental and development work, and a modest amount of actual rolling stock experience, on gas-turbine-powered, traction-propelled locomotives in France, Japan, Germany, and the United States. As with ship applications, many of the gas turbines are turboshafts, adapted from shaft, fan, or jet engines developed for aeronautical use.

Various power conditioning schemes have been employed. The Sikorsky Turbotrain I, with a free turbine engine, drives directly through a gear reduction to the drive wheels, as does the Japanese National Railways KI-HA391. The Rohr/Frangeco turbine train developed for, and still in use by, Amtrak drives through a hydraulic torque converter to the drive wheels. General Electric built a series of powerful gas turbine-electric locomotives for the Union Pacific Railroad. In these units a fixed-turbine prime mover drove an electric generator through a gear reduction system, and the generated electric power drove motors at the drive wheels. One of these locomotives was modified to burn powdered coal rather than liquid petroleum.

Unique variants of turbine-powered locomotives were the very high-speed air-cushion track vehicles or aerotrains developed in the 1960s, most notably by Jean Bertin in France. Like the amphibious air-cushion vehicle, the aerotrain was supported by a pressurized air cushion, in this case developed between the vehicle and an inverted-T guideway or track. In addition to gas turbines to supply the pressurized air for the air cushion, aircraft-type turbofan or turboprop engines were used for reaction propulsion of the vehicles. Speeds on the order of 250 mi/h (400 km/h) were obtained on experimental vehicles. *See* LOCOMOTIVE.

Automotive vehicle propulsion. Beginning in the late 1940s and early 1950s, the United States auto industry ran a number of modified gas turbine engines, and in 1953 a standard passenger car was manufactured with a specially designed gas turbine engine. Over the next 15 years, five generations of gas turbine engines were installed in passenger automobiles. Although these gas-turbine-powered vehicles did not result in a marketable product capable of competing with the conventional spark-ignition internal combustion engine, they did demonstrate the gas turbine engine's smooth operation and ability to start in cold weather. *See* INTERNAL COMBUSTION ENGINE.

The two main characteristic disadvantages of the gas turbine engine for automotive operation, which must be overcome before it can displace the internal combustion engine, are its poor efficiency at part load and idle conditions, and its poor acceleration from idle condition. Another negative factor, the high production cost, relates to the use of expensive and scarce cobalt and nickel in the hot turbine sections. However, technological advances have been made in the use of ceramic components. In addition to lowering the production costs, ceramic components permit increased engine operating temperatures, around 2500°F (1370°C), which improve engine cycle efficiency. The conventional piston engine vehicle has become more expensive as a result of efforts to meet federal car mileage requirements and exhaust emission standards, causing the gas turbine engine to become more cost-competitive. *See* CERAMICS.

A free-turbine gas turbine engine, designed for 100 hp (75 kW), was developed (**Fig. 2**), which demonstrates that technology is available to achieve fuel economies equivalent to those of spark ignition

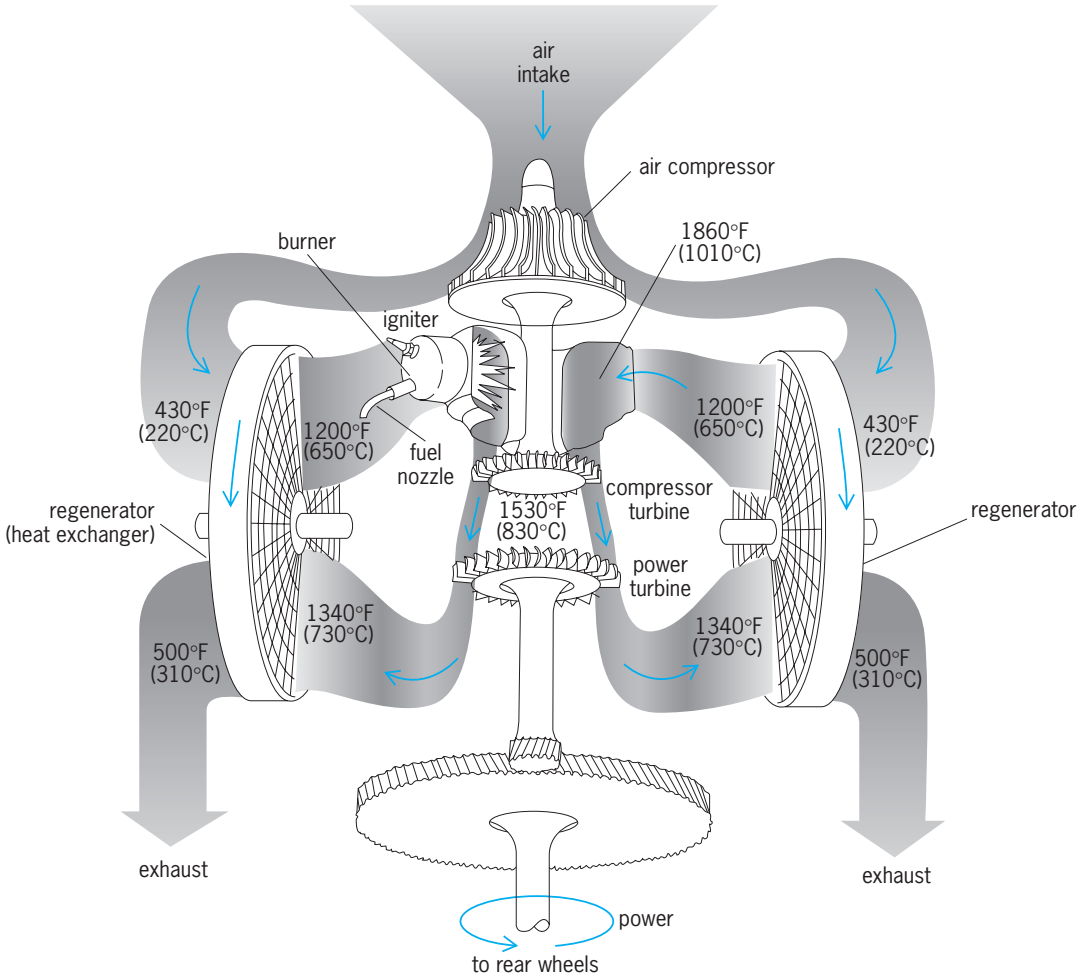

Fig. 2. Chrysler upgraded turbine engine. (*Chrysler Corp./U.S. Department of Energy*)

engines and also achieve lower exhaust emissions, easier starting, and less noise and vibration. The engine can operate on various kinds of fuel, such as coal-derived fuels, alcohols, and blends.

The leading automotive companies have also experimented with more powerful versions of automobile gas turbines for large buses and trucks. In typical long-range intercity applications, the gas turbine should retain all its intrinsic advantages and be less subject to the disability of its high idle fuel consumption. Regenerative, free-turbine engines have been developed especially to meet the requirements of heavy-duty engine applications in ground vehicles and equipment in the general range of 300–600 hp (225–450 kW). The regeneration system reduces fuel consumption and lowers exhaust temperatures.

The power transfer system is an automatic transmission without any torque converter. It offers controlled engine braking and high torque-rise performance. Gas turbine power results in less weight in the power section, the absence of a radiator, and a clean exhaust obtained with standard diesel fuels.

In 1976 the U.S. Army selected a 1500-hp (1120-kW) regenerative engine with a two-spool core and a free turbine to power its XM1 main battle tank. For cross-country duty, the gas turbine inlet is supplied with a combination inertial separator and barrier filter system to extract sand and dust from the air entering the engine. The engine runs on a wide range of fuels, including gasoline, diesel, and jet aircraft fuels. The power turbine is supplied with variable inlet guide vanes to optimize fuel consumption at all output powers. The automatic transmission provides four speeds forward and two in reverse, contains integral brakes, has continuously variable hydrostatic steering, and provides pivot steering while the transmission is in neutral. *See* ARMY ARMAMENT; GAS TURBINE; PROPULSION.

Fredric F. Ehrich

Bibliography. C. F. Foss, *Jane's Armour and Artillery*, annually; *Gas Turbine World Handbook*, annually; P. Jackson, *Jane's All the World's Aircraft*, annually; R. L. Trillo, *Jane's High-Speed Marine Craft*, annually.

Turbocharger

An air compressor or supercharger on an internal combustion piston engine that is driven by the engine exhaust gas to increase or boost the amount of fuel that can be burned in the cylinder, thereby increasing engine power and performance. On an aircraft piston engine, the turbocharger allows the engine to retain its sea-level power rating at higher altitudes despite a decrease in atmospheric pressure. *See* RECIPROCATING AIRCRAFT ENGINE; SUPERCHARGER.

Construction and operation. The turbocharger is a turbine-powered centrifugal supercharger. It consists of a radial-flow compressor and turbine mounted on a common shaft (see **illus.**). The turbine uses the energy in the exhaust gas to drive the compressor, which draws in outside air, precompresses it, and supplies it to the cylinders at a pressure above atmospheric pressure. Turbocharger speed, which may be upward of 130,000 revolutions/min, is not dependent upon engine speed, but is determined by the power balance between the turbine and the compressor.

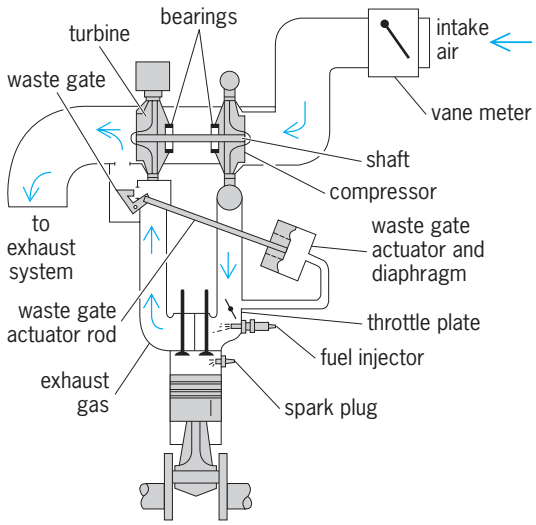

Typical turbocharger installation on a fuel-injected spark-ignition engine, using a waste gate to limit boost pressure. Charge-air cooler is not shown. (*Ford Motor Co.*)

Common turbocharger components include the rotor assembly, bearing housing, and compressor housing. The turbine housing and control devices frequently differ according to application. The compressor housing is generally fabricated of cast aluminum, and it may contain an integral bypass valve. Turbine housings are cast from nodular graphite iron and other materials, depending on exhaust temperature. The shaft bearings usually receive oil from the engine lubricating system. Engine coolant may circulate through the housing to aid in cooling. *See* ENGINE COOLING.

Some turbochargers have twin-flow housings in which the two flows do not join until just ahead of the turbine wheel inlet. This allows pulse turbocharging in which the kinetic energy of the exhaust gas is used in addition to its pressure energy. In the usual constant-pressure turbocharging, only the pressure energy of the exhaust gas is used, along with single-flow turbine housings. Such turbocharger installations are used for automotive engines.

Boost-pressure control. The faster the engine runs, the faster the turbine spins the compressor. This causes a pressure buildup that must be controlled in order to avoid engine damage. One method is to allow a portion of the exhaust gas to bypass the turbine by opening a valve or flap known as a waste gate. It can be a separate component (see illus.) or integrated into the turbine housing. As much exhaust gas is diverted as is necessary to slow the turbine down to the proper speed.

The waste gate is actuated pneumatically by control pressure that is tapped off at the pressure end of the turbocharger. Instead of a waste gate, regulation may be provided by variable turbine geometry, as in a variable-nozzle turbocharger. This allows the constant-pressure characteristics of the turbine to change continuously, more efficiently using the exhaust-gas energy.

Charge-air cooling. As the air is compressed, its temperature rises, thus reducing the efficiency of turbocharging. The use of a heat exchanger as a charge-air cooler (also known as aftercooler or intercooler) helps overcome this problem. Before entering the cylinders, the hot air from the compressor is sent through the charge-air cooler. This cools the air, thereby increasing its density. *See* HEAT EXCHANGER; INTERNAL COMBUSTION ENGINE.

Donald L. Anglin

Bibliography. *Bosch Automotive Handbook*, 1987; J. Humphries, *Automotive Supercharging and Turbocharging Systems*, 1995; Society of Automotive Engineers, *Turbocharger Performance and Application*, 1982; Society of Automotive Engineers, *Turbochargers and Turbocharged Engines*, 1979; K. Zinner, *Supercharging of Internal Combustion Engines*, 1978.

Turbodrill

A rotary tool used in drilling oil or gas wells in which the bit is rotated by a turbine motor inside the well. The principal difference between rotary and turbo drilling lies in the manner that power is applied to the rotating bit or cutting tool. In the rotary method, the bit is attached to a drill pipe, which is rotated through power supplied on the surface. In the turbodrill method, power is generated at the bottom of the hole by means of a mud-operated turbine.

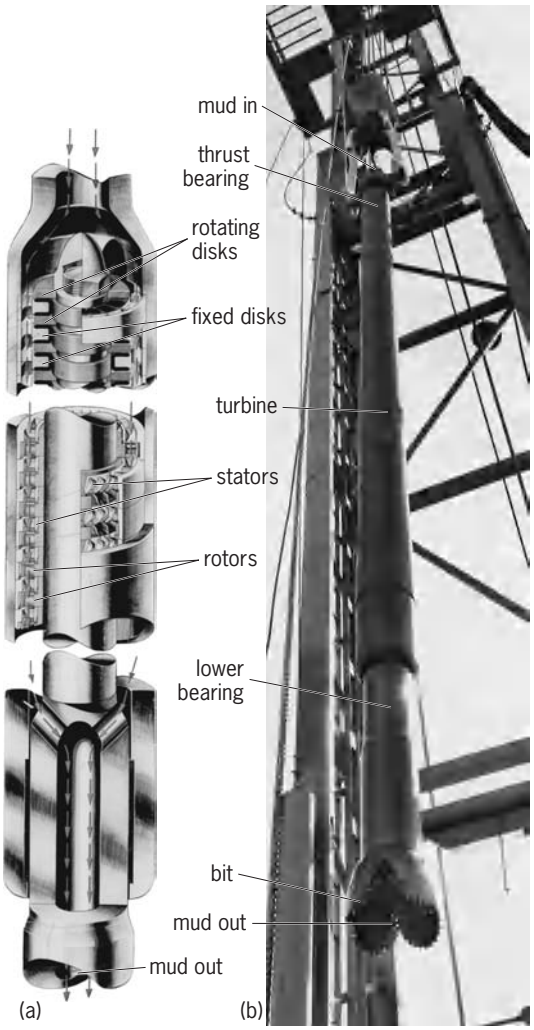

Components of a turbodrill. (*a*) Cutaway view of turbine, bearings, and bit. (*b*) Drill string suspended above hole. (*Dresser Industries*)

The turbodrill (see **illus.**) consists of four basic components: the upper, or thrust, bearing; the turbine; the lower bearing; and the bit. Most turbodrills are about 30 ft (9 m) long, with shafts about 20 ft (6 m) long. The turbodrill is attached at its top to a drill collar, or heavy length of steel pipe, that makes up the bottom end of the drill pipe extending to the surface. Once the turbodrill passes below the well head, operations on the rig floor are the same as for rotary drilling. Rotation of the drill pipe is not necessary for turbodrilling, because rotation of the bit develops through the turbine on the lower end of the drill string. It is usual practice, however, to rotate the drill pipe above the turbine slowly, at 6–8 revolutions/minute, either by means of the rotary table on the derrick floor or through torque of the turbine on the bottom. Rotation of the bit is much faster than in rotary drilling, and is usually between 500 and 1000 revolutions/minute.

In operation, mud is pumped through the drill pipe, passing through the thrust bearing and into the turbine. In the turbine, stators attached to the body of the tool divert the mud flow onto rotors attached to the shaft. This causes the shaft, which is connected to the bit, to rotate. The mud passes through a hollow part of the shaft in the lower bearing and through the bit, as in rotary drilling, to remove cuttings, cool the bit, and perform the other functions of drilling fluid. Capacity of the mud pump, which is the power source, determines rotational speed.

Two basic types of turbodrills are used in the United States. One is a standard 100-stage unit (one rotor and one stator comprise a stage); the other is a tandem turbodrill, made up of two or three standard sections. Although turbodrills have been in wide use in Russia, they are still relatively rare in the United States, where the emphasis has been on the rotary tool method of drilling. Despite faster penetration with the turbodrill, and several other advantages, widespread use of the turbodrill in the United States has been limited principally because of the faster wear of the bits, necessitating time-consuming and costly round trips to remove the drill string from the

hole and change the bit. *See* DRILLING, GEOTECHNICAL; OIL AND GAS WELL DRILLING. Ade L. Ponikvar

Turbofan

An air-breathing aircraft gas turbine engine (**Fig. 1**) with operational characteristics between those of the turbojet and the turboprop. Like the turboprop, the turbofan consists of a compressor-combustor-turbine unit, called a core or gas generator, and a power turbine. This power turbine drives a low- or medium-pressure-ratio compressor, called a fan, some or most of whose discharge bypasses the core. *See* TURBOJET; TURBOPROP.

Operating principle. The gas generator produces useful energy in the form of hot gas under pressure. Part of this energy is converted by the power turbine and the fan it drives into increased pressure of the fan airflow. This airflow is accelerated to ambient pressure through a fan jet nozzle and is thereby converted into kinetic energy. The residual core energy is converted into kinetic energy by being accelerated to ambient pressure through a separate core jet nozzle (**Fig. 2***a*). The reaction in the turbomachinery in producing both streams produces useful thrust.

Bypass ratio. Turbofans are generally characterized by their bypass ratio, the ratio of the airflow which flows around (bypasses) the core to the airflow which passes through the core. Airplanes which cruise at a flight Mach number less than 0.6 are generally propeller-driven (if they are turboprop-driven, they may be considered extremely high-bypass-ratio turbofan aircraft). In the range of Mach 0.6 to transonic flight speeds, bypass ratios of 3 to 7, where the fan imparts a very modest propulsive velocity to a very large bypass flow, are found to be most efficient and dominate aircraft applications (Fig. 1*a*). Their outstanding fuel economy also assures a low rate of exhaust emissions because of their relatively low fuel consumption. The low velocity in their exhaust jets also makes them inherently quieter than jet engines or low-bypass turbofans. For aircraft which fly supersonically or have mixed supersonic-subsonic missions, turbofans with higher exhaust jet velocities and lower bypass ratios of 0.1 to 3 are used. Such low-bypass turbofans are often equipped with afterburners and variable exhaust nozzles (Fig. 1*b*) to provide thrust augmentation (on the order of 50% greater than the thrust otherwise obtainable). Since afterburning operation involves very high specific fuel consumption, its usage is generally restricted to very short duration, typically for transonic acceleration or combat maneuver. *See* AFTERBURNER; SPECIFIC FUEL CONSUMPTION.

Where the bypass ratio is designed to be less than 0.1, the bypass air has a very minor role in affecting the propulsive efficiency of the engine. The usual

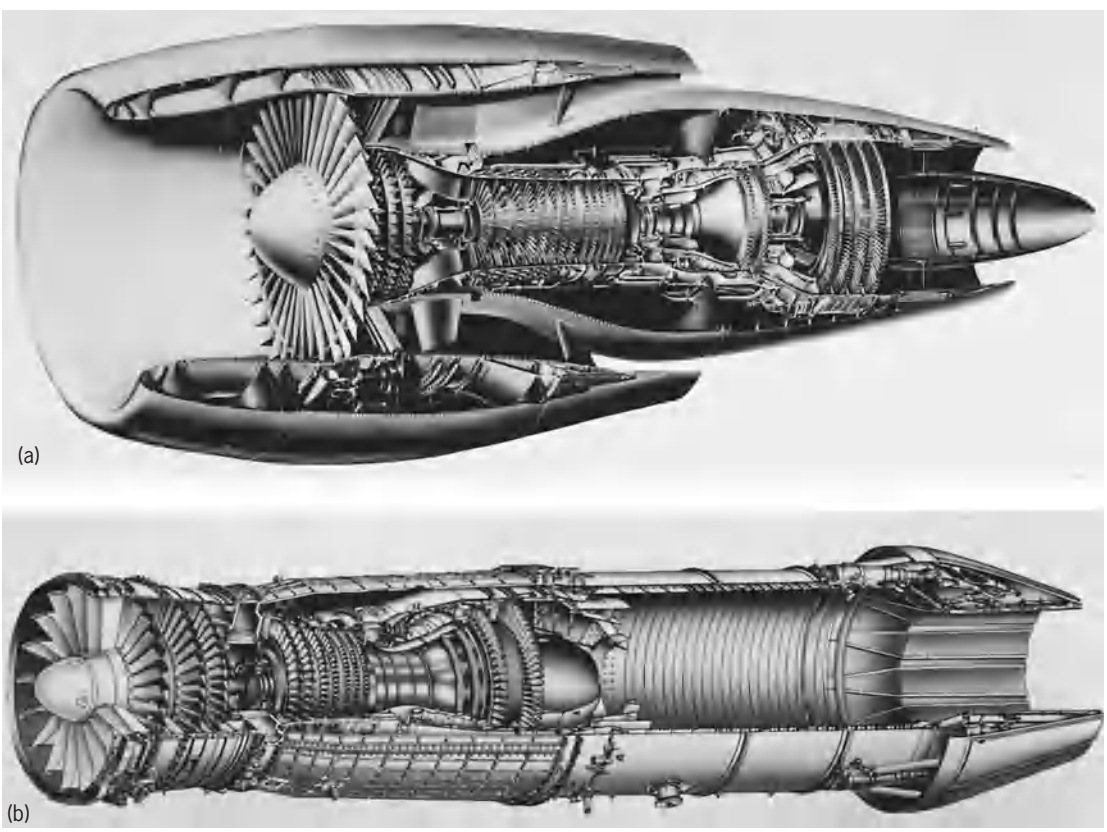

Fig. 1. Modern turbofan engines. (*a*) High-bypass engine for high subsonic flight speed. (*b*) Low-bypass afterburning engine for transonic and low-supersonic flight speeds. (*GE Aircraft Engines*)

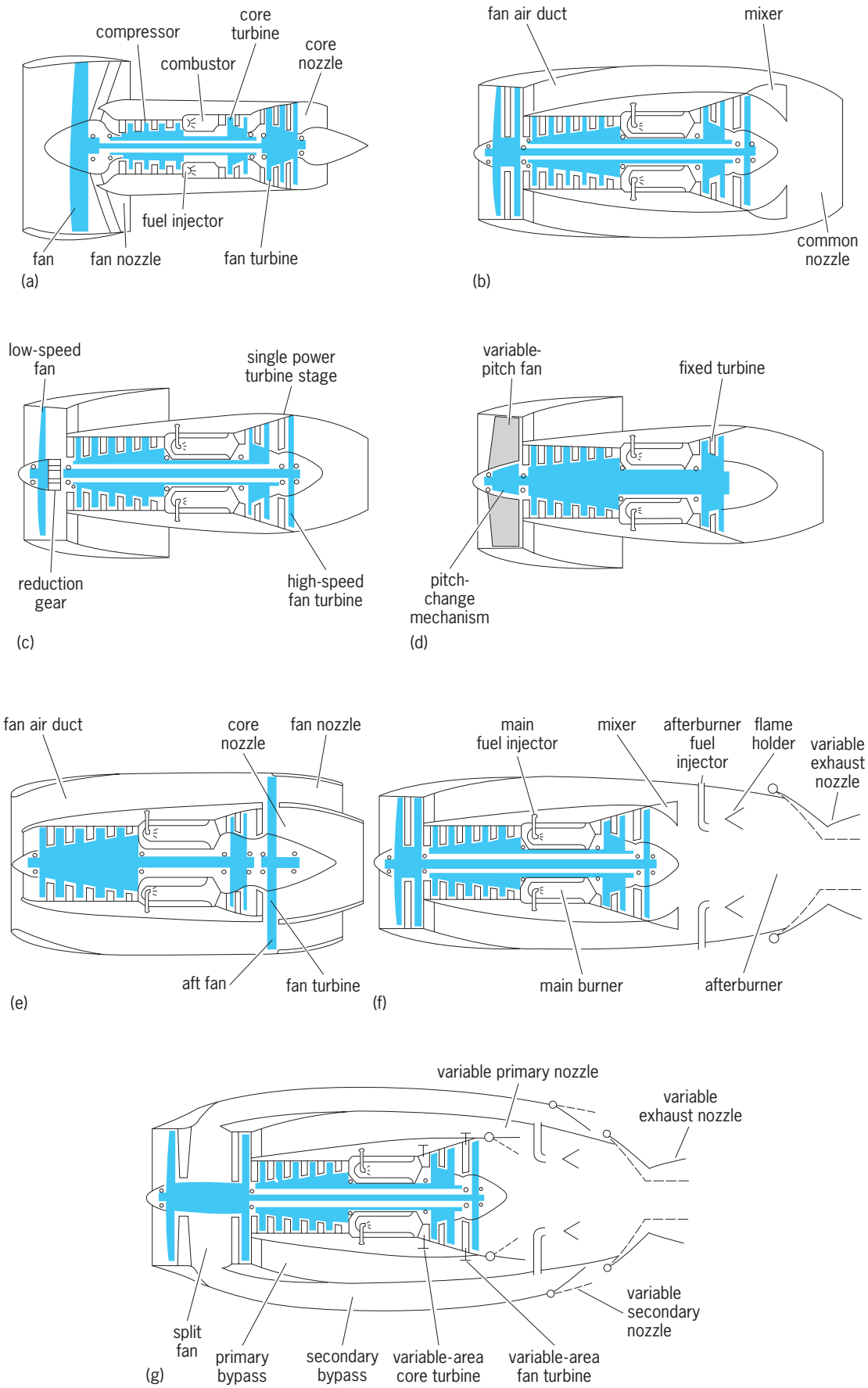

Fig. 2. Turbofan configurations. (*a*) High-bypass, separate-flow turbofan. (*b*) Mixed-flow turbofan. (*c*) Fixed-turbine, variable-pitch fan. (*d*) Ultrahigh-bypass turbofan. (*e*) Aft fan. (*f*) Afterburning turbofan. (*g*) Variable-cycle turbofan.

intent of this approach to engine design is to use the bypass air to blanket the hot end of the engine (the combustor casing, the turbine casings, and the afterburner casing) with cool air to protect the surrounding aircraft structure. In thermodynamic terms, the very low bypass turbofan engine is indistinguishable from a jet engine, where a small fraction of the compressor airflow is usually extracted to cool the structure and casings of the hot end of the engine.

Configurations. Ideally the fan and core exhaust velocities should be nearly equal. If they are not, some benefit is derived from incorporation of a mixer between the two streams and exhausting the mixed flow (Fig. 2*b*) through a common nozzle.

Variable-pitch fan blading, such as is found in propeller blading, is being applied in turbofan blading. An early application was in turbofan engines with fixed turbines (that is, where the power turbine driving the fan was part of the same rotor as the core), where the variability was used for unloading the fan during the engine starting sequence (Fig. 2*c*). *See* PROPELLER (AIRCRAFT).

A more subsequent application is in the ultrahigh-bypass turbofan, where the more conventional blocker-door thrust-reverser systems become impractical, and where variable (that is, reversible) pitch fan blading is used to provide reverse thrust (Fig. 2*d*). Another feature of such ultrahigh-bypass engines is a reduction gear, which permits the driving power turbine to rotate at a much higher speed than the large-diameter, low-speed fan. This feature allows the power turbine to be designed with fewer stages and with a lower diameter (and hence to be lighter and smaller) than would otherwise be required without a gearbox.

An early manifestation of the turbofan was the aft fan, where a row of fan blades is mounted in the rear of the engine, astride the power turbine blading (Fig. 2*e*).

Turbofans intended for service in mixed subsonic-supersonic missions are often designed with afterburners in which an additional combustor is located in the exhaust streams (Fig. 2*f*) of the turbofan to provide thrust augmentation for any thrust-limited segment of the mission. The burning may be arranged in the core stream or the bypass stream, in both, or in a mixed stream. A variable-area exhaust nozzle is generally also necessary. *See* AFTERBURNER.

Variable-cycle turbofans exploit the advantages of high-bypass operation for subsonic flight, and provide low-bypass operation for efficient supersonic cruise for mixed-mission aircraft. These involve a variety of variable-geometry components and elements in the basic turbofan: additional variable stators in the fan, variable stators in turbines, valving elements in the bypass streams, addition of a second bypass stream, more than one augmentation burner, and so forth (Fig. 2*g*).

To shorten the landing of aircraft, turbofans are often provided with reversers, devices which block the exhaust jet and partially reverse its direction, thereby reversing the thrust vector.

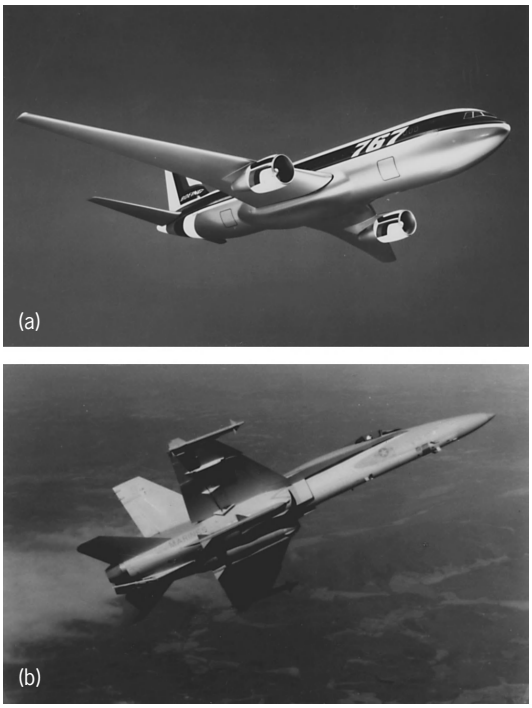

Fig. 3. Turbofan installations. (*a*) Twin-turbofan, under-the-wing installation in a modern transport aircraft (*Boeing*). (*b*) Military combat aircraft with twin low-bypass afterburning turbofan engines embedded in fuselage (*GE Aircraft Engines*).

To shorten the takeoff field length required, particularly on hot days or at high altitudes, provision can be made for water or water-methanol to be injected into the air upstream of the combustor.

Installation. Turbofan engines are most commonly installed in a pod or nacelle which provides an efficient air inlet system and a streamlined, low-drag external configuration. Such pods are commonly hung below the aircraft's wings (**Fig. 3***a*) or less commonly above the wings, or at either side of the aft end of the fuselage. A third engine can be installed in a pod at the root of the vertical tail assembly. In single- or twin-engined military aircraft, the engines may be embedded in the fuselage (Fig. 3*b*). *See* JET PROPULSION; TURBINE PROPULSION.

Fredric F. Ehrich

Bibliography. P. Hill and C. Peterson, *Mechanics and Thermodynamics of Propulsion*, 2d ed., 1992; P. Jackson (ed.), *Jane's All the World's Aircraft*, annually; M. Kroes et al., *Aircraft Powerplants,* 7th ed., 1995; Rolls-Royce plc, *The Jet Engine,* 5th ed., 2005; I. E. Treager, *Aircraft Gas Turbine Engine Technology*, 3d ed., 1995.

Turbojet

A gas turbine power plant used to propel aircraft, where the thrust is derived within the turbo- machinery in the process of accelerating the air and products of combustion out an exhaust jet nozzle (**Fig. 1**). *See* GAS TURBINE.

Operating principle. In its most elementary form (**Fig. 2***a*), the turbojet operates on the gas turbine or Brayton thermodynamic cycle. The working fluid,

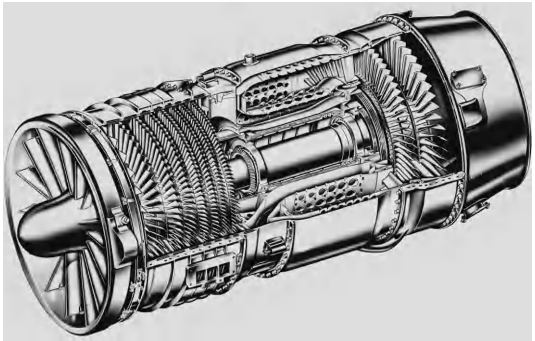

Fig. 1. Typical turbojet engine. (*GE Aircraft Engines*)

air drawn into the inlet of the engine, is first compressed in a turbo-compressor with a pressure ratio of typically 10:1 to 20:1. The high-pressure air then enters a combustion chamber, where a steady flow of a hydrocarbon fuel is introduced in either spray or vapor form and burned continuously at constant pressure. The exiting stream of hot high-pressure air, at an average temperature whose maximum value may range typically from 1800 to 2800°F (980 to 1540°C), is then expanded through a turbine, where energy is extracted to power the compressor. Because heat had been added to the air at high pressure, there is a surplus of energy left in the stream of combustion products that exits from the turbine and that can be harnessed for propulsion. *See* BRAYTON CYCLE; GAS TURBINE.

The turbojet engine derives propulsive thrust from the stream of hot high-pressure gas by ejecting the flow at high velocity through a jet nozzle. The principle of action was characterized by Isaac Newton in his second law of motion, in which he generalized the fact that the force F required to accelerate an object was simply proportional to the product of the object's mass m and the acceleration a, as in Eq. (1).

$$F = ma \tag{1}$$

In the case of the turbojet engine, rather than simply accelerate a single lump of mass, the engine acts to accelerate a continuous stream of air by raising its velocity from that at which it enters the engine, V_0, taken to be the flight speed of the aircraft, to that which leaves the exhaust nozzle, V_j. The force that the components of the engine must then exert on the airflow to achieve this acceleration is given by Eq. (2), where M is the rate of mass flow through

$$F = M(V_j - V_0) \tag{2}$$

the engine. The reaction to this force exerted by the engine's components on the airstream is then felt as thrust on the engine's mounts, which connect it to the aircraft. *See* FORCE.

There is a significant inefficiency in this process of thrust production that is manifest in the considerable amount of energy tied up in the high-temperature, high-velocity exhaust stream that leaves the engine without having been usefully harnessed. The propulsive efficiency of the thrust generation is found to be approximately given by Eq. (3), where $V_j > V_0$. The

$$\eta_p = \frac{2V_0}{V_0 + V_j} \tag{3}$$

relationship implies that high propulsive efficiency requires matching the jet velocity as closely as possible with the flight speed. In actual modern practice, therefore, simple turbojet engines are applied most often to aircraft that have very large portions of their flight profile at supersonic flight speed in order to most efficiently exploit the supersonic jet velocity and avoid prohibitively low propulsive efficiencies.

Configurations. The basic configuration of the turbojet (Fig. 2*a*) includes an inlet system, where the inlet stream of air, approaching the engine at a relative velocity equal to the flight speed of the aircraft, is decelerated to a lower velocity at which it can be drawn into the compressor. The compressor may be a multistaged axial-flow unit (Fig. 1), a centrifugal (radial-flow) compressor, or a combination axi-centrifugal component.

In modern usage, large multistage axial-flow compressors with airflows greater than 12–15 lb/s (5–7 kg/s) are typically 4–6% more efficient than

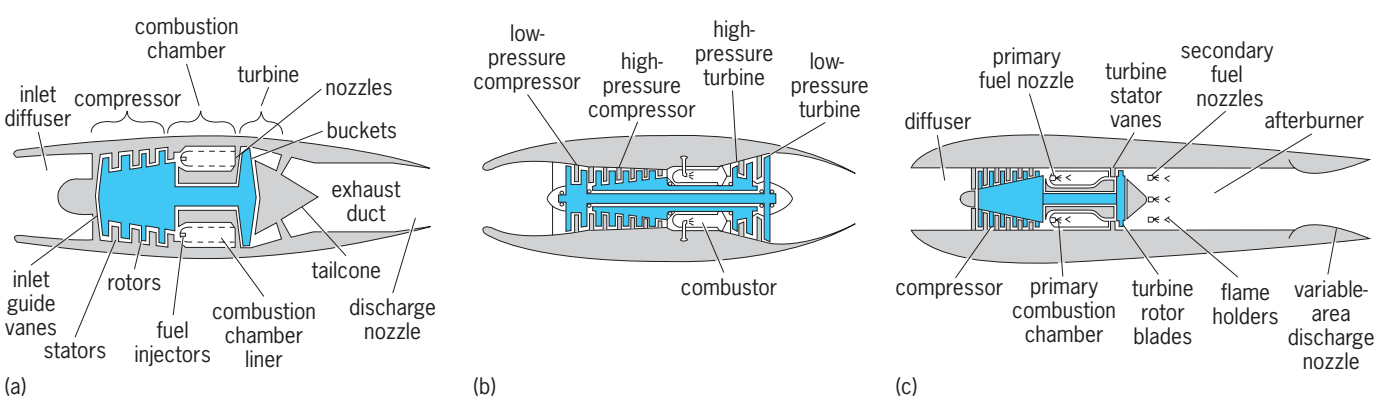

Fig. 2. Turbojet engine configurations. (*a*) Basic turbojet engine with axial-flow components. (*b*) Dual-rotor turbojet engine. (*c*) Afterburning turbojet engine.

the equivalent single-stage centrifugal compressor. Moreover, the requirement for very high tip speeds and for radially extended diffusion systems generally dictates frontal areas for centrifugal compressors that may be 30–50% greater than for the equivalent set of axial-flow stages. The large frontal area is particularly disadvantageous in fitting the engine into a low-drag nacelle for high flight speed, and in fitting a power producer within the confines of the bypass stream of a turbofan. The higher tip speeds of the centrifugal machines also dictate that they be considerably heavier than their axial-flow counterparts. For these reasons, axial-flow compressors are invariably used in modern engines except for very small units with airflows less than 12–15 lb/s (5–7 kg/s), where the small size of axial-flow blading involves disproportionate sacrifices in compressor efficiency and manufacturing complexity so that a centrifugal or axi-centrifugal unit becomes the configuration of choice.

Multistage compressors designed for high pressure ratios, above 6:1 or 7:1, have included cascades of stator vanes, each of which is pivoted so that its stagger angles may be varied at different power levels of operation to match individual-stage characteristics and avoid the phenomenon known as compressor stall or surge. In very high-pressure compressors, it may be necessary to split the compressor into two sections, each operating at its own independent rotative speed and driven by its own turbine stages (Fig. 2). *See* COMPRESSOR.

The combustors in early machines were often composed of an annular array of separate cans, each with its own fuel injector, but modern practice generally uses full annular chambers. The turbines are most often axial-flow in configuration, and have involved the application of the most advanced developments and fabrication techniques of high-temperature, high-strength alloys and, subsequently, nonmetallic materials, as well as complex internal passages to introduce cooling air, all to permit continuous operation at the very high gas temperatures that are required for high thermodynamic efficiency. The final component, the exhaust nozzle, is generally a fixed convergent unit, or, for very high values of engine pressure ratios and flight speeds, may include some divergence after the throat of the convergent section to most efficiently achieve supersonic jet velocities. *See* NOZZLE.

The thrust of a turbojet may be augmented by the addition of an afterburner (Fig. 2*c*). The afterburner also entails the use of a variable-area exhaust nozzle, since the area must be increased during afterburning to accommodate the additional heating of the exhaust air to a temperature in the range of 3400°F (1870°C), which greatly increases it volume. Although it permits the achievement of up to 50% more thrust from an engine that would otherwise be limited by the maximum temperature that can be withstood by the turbine, that thrust increment is achieved very inefficiently, that is, with very high fuel consumption. Afterburners are therefore included in engines only for short-term or emergency usage, such as in combat by military fighter aircraft. *See* AFTERBURNER.

Application. Turbojets were first used to power military combat aircraft in World War II. In the first decades after the war, turbojets were successfully applied to commercial transport aircraft, but were soon displaced by turboprop and turbofan aircraft, which were substantially more efficient at the subsonic and transonic flight speeds used, while exploiting the same principles of propulsion and enjoying much the same advantages of simplicity, light weight, and durability as the turbojet. *See* TURBOFAN; TURBOPROP.

Turbofans have also displaced turbojets in propulsion of military aircraft. Although combat aircraft are designed with supersonic flight capability where turbojets might be expected to find application, the aircraft must also spend a considerable amount of time and fuel at lower flight speeds for purposes of takeoff, climb, loiter, cruise, loiter, and approach and landing, so that low-bypass turbofans have evolved as the most efficient design arrangement. The single major current application of the turbojet has been to the supersonic commercial transport. But the next generation of supersonic transports may very well be powered by some type of turbofan, because of the need to fly substantial portions of its mission at subsonic speed and because of the difficult noise problem that a simple turbojet imposes.

Turbojets have retained a small niche in the aircraft propulsion spectrum, where their simplicity and low cost are of paramount importance, such as in short-range expendable military missiles, or where their light weight may be an overriding consideration, such as for lift jets in prospective vertical takeoff and landing aircraft. *See* AIRCRAFT PROPULSION; JET PROPULSION; TURBINE PROPULSION; VERTICAL TAKEOFF AND LANDING (VTOL).

Fredric F. Ehrich

Bibliography. P. Hill and C. Peterson, *Mechanics and Thermodynamics of Propulsion*, 2d ed., 1992; P. Jackson (ed.), *Jane's All the World's Aircraft*, annually; M. Kroes et al., *Aircraft Powerplants,* 7th ed., 1995; Rolls-Royce plc, *The Jet Engine,* 5th ed., 2005.

Turboprop

A gas turbine power plant producing shaft power to drive a propeller or propellers for aircraft propulsion. Because of its high propulsive efficiency at low flight speeds, it is the power plant of choice for short-haul and low-speed transport aircraft where the flight speeds do not exceed Mach 0.5–0.6. Developments in high-speed, highly loaded propellers have extended the range of propellers to flight speeds up to Mach 0.8–0.9, and there are prospects of these extremely efficient prop-fans assuming a much larger role in powering high-speed transport aircraft. *See* GAS TURBINE.

Power producer. As with all gas turbine engines, the basic power production in the turboprop is

accomplished in the gas generator or core of the engine, where a steady stream of air drawn into the engine inlet is compressed by a turbocompressor. The high-pressure air is next heated in a combustion chamber by burning a steady stream of hydrocarbon fuel injected in spray or vapor form. The hot, high-pressure air is then expanded in a turbine that is mounted on the same rotating shaft as the compressor and supplies the energy to drive the compressor. By virtue of the air having been heated at higher pressure, there is a surplus of energy in the turbine that may be extracted in additional turbine stages to drive

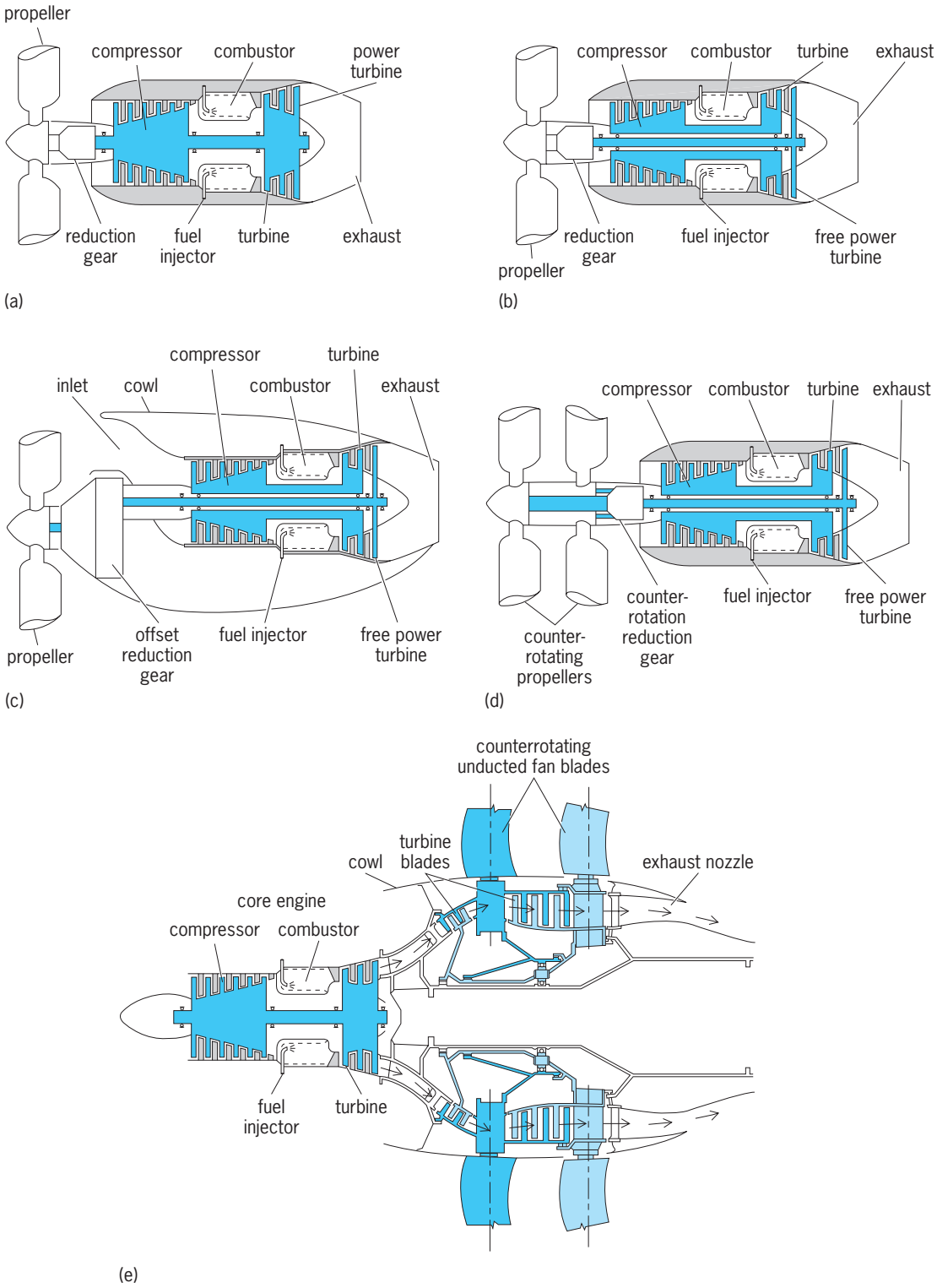

Fig. 1. Typical turboprop configurations. (*a*) Fixed-turbine turboprop. (*b*) Free-turbine turboprop. (*c*) Offset-gearbox turboprop. (*d*) Counterrotating turboprop. (*e*) Unducted fan, an ungeared, counterrotating, pusher turboprop. (*After R. C. Hawkins, Unducted fan for tomorrow's subsonic propulsion, Aerospace America, October 1984*)

a useful load, in this case a propeller or propellers.

A large variety of detailed variations are possible within the core. The compressor may be an axial-flow type, a centrifugal (that is, radial-flow) type, or a combination of stages of both types (that is, an axi-centrifugal compressor). In modern machines, the compressor may be split in two sections (a low-pressure unit followed by a high-pressure unit), each driven by its own turbine through concentric shafting, in order to achieve very high compression ratios otherwise impossible in a single spool. *See* COMPRESSOR.

Engine efficiency, measured as specific fuel consumption, is the rate of fuel flow (mass of fuel per unit time) per shaft power. A value of specific fuel consumption of 0.4 lb/(h)(hp) or 0.24 kg/(h)(kW) is typical of modern engines. The compressor pressure ratio is an important parameter in designing for a low specific fuel consumption. Pressure ratios of 12:1 to 20:1 are typical in current engines. Future engines may have pressure ratios as high as 40:1. The maximum value of the average temperature of the hot gases entering the turbine has a major effect on the specific power of the engine, the amount of horsepower generated per unit of airflow passing through the engine. The specific power is of fundamental importance since the size and weight of the engine depend strongly on the amount of airflow through the engine. A typical value of specific power in a modern engine might be 200 hp/(lb)(s) or 329 kW/(kg)(s). *See* SPECIFIC FUEL CONSUMPTION.

Configurations. The earliest and simplest turboprops had the turbine that extracted useful output power integrated (or fixed) on the same spool as the turbine that drove the compressor, forming a single dual-purpose turbine (**Fig 1***a*). In this configuration, the power producer spool drives the propeller shaft through a speed-reduction gearbox to accommodate the large-diameter, low-rotational-speed propeller.

There are several major disadvantages to this design having to do with starting, part-power efficiency, and windmilling operation of inoperative engines. These disadvantages are overcome by the free-turbine configuration (Fig. 1*b*). In this case, the load turbine is mounted on a spool separate from the gas generator and is free to drive the propeller through the propeller gearbox at a speed independent of the gas generator, and it is controlled to optimize propeller operation.

A common variant of these engines involves an offset reduction gear that calls for an offset inlet at one side of the gearbox (Fig. 1*c*).

A turboprop may also be used to drive two concentric counterrotating propellers (Fig. 1*d*). Counterrotating propellers are capable of providing significantly higher propulsive efficiency than single propellers.

Turboprops may also be designed in different configurations characterized as tractor or pusher types. In the tractor type the propeller is forward of the engine, which makes it most amenable to mounting with the propeller forward of the wing (**Fig. 2***a*).

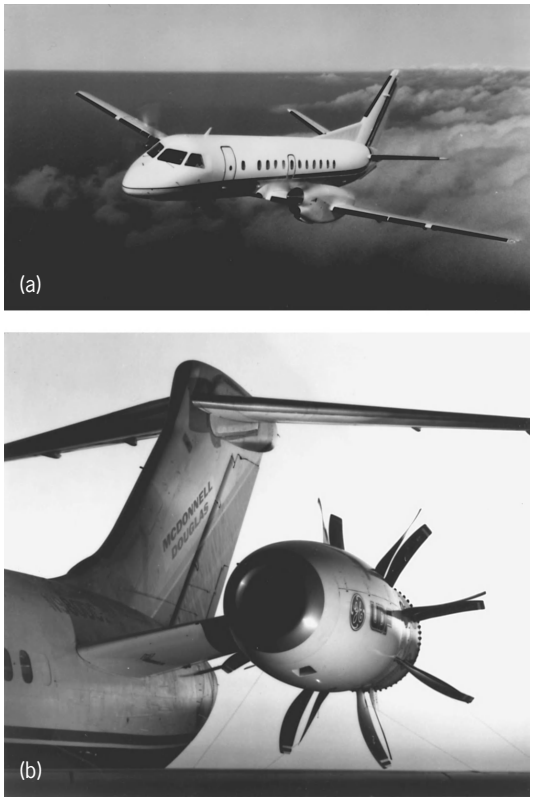

Fig. 2. Typical turboprop installations. (*a*) Conventional tractor-type single-rotation turboprop (*Saab Corp.*). (*b*) Unducted fan, a pusher-type counterrotating turboprop (*GE Aircraft Engines*).

The pusher type has the propeller mounted aft of the engine, which makes this engine amenable to being mounted aft of the wing (Fig. 2*b*).

An ingenious engine type, referred to as the unducted fan, involves the pusher configuration, combined with very efficient high-speed counterrotating propellers and a unique scheme that does away with the gearbox (Fig. 1*e* and Fig. 2*b*). *See* AIRCRAFT PROPULSION; PROPELLER (AIRCRAFT); TURBINE PROPULSION.

Fredric F. Ehrich

Bibliography. M. Kroes et al., *Aircraft Powerplants,* 7th ed., 1995; P. Jackson, (ed.), *Jane's All the World's Aircraft,* annually; Rolls-Royce plc, *The Jet Engine,* 5th ed., 2005; I. E. Treager, *Aircraft Gas Turbine Engine Technology,* 3d ed., 1995.

Turbulent flow

A fluid motion in which velocity, pressure, and other flow quantities fluctuate irregularly in time and space. **Figure 1** shows a slice of a water jet emerging from a circular orifice into a tank of still water. A small amount of fluorescent dye mixed in the jet makes it visible when suitably illuminated by laser light, and tags the water entering the tank. In this and similar realizations of the flow, there is a small region close to the orifice where the dye concentration does not vary with position, or with time at a given position. This represents a steady laminar state. Generally in laminar motion, all variations (if

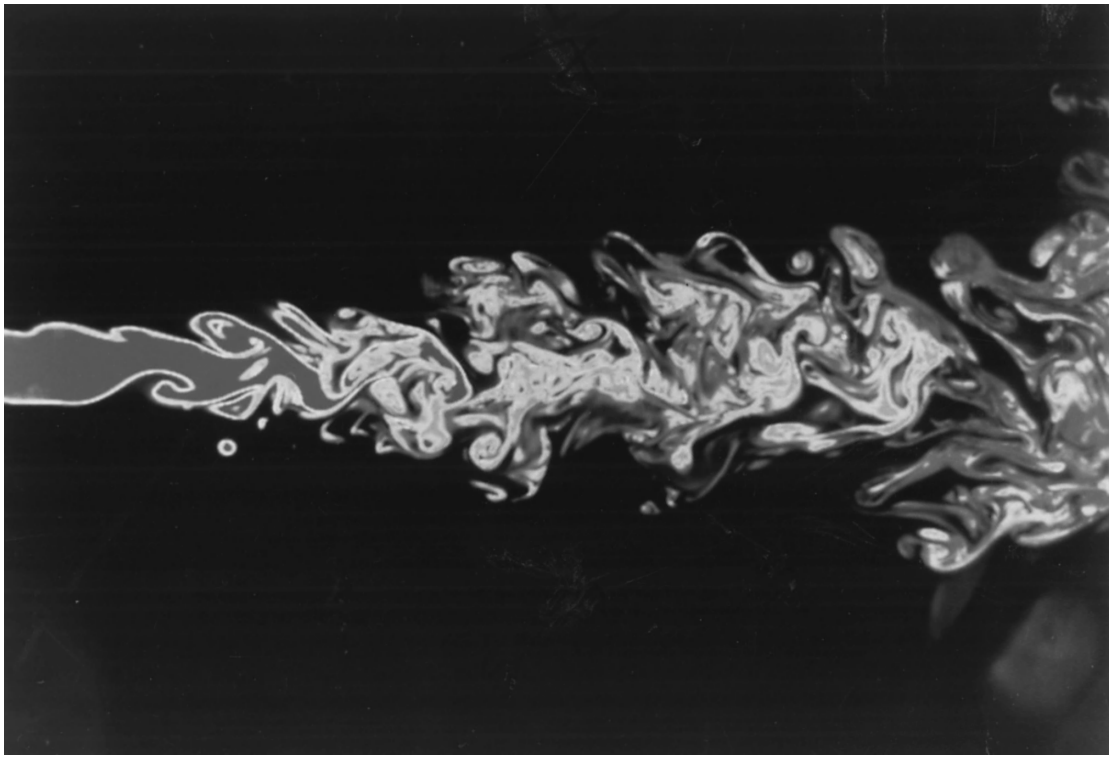

Fig. 1. Two-dimensional image of an axisymmetric water jet, obtained by the laser-induced fluorescence technique. A neodymium:yttrium-aluminum-garnet laser beam, shaped into a sheet of 250-micrometer thickness by using suitable lenses, was directed into a water tank into which the jet fluid, containing small amounts of uniformly dispersed fluorescing dye (sodium fluorescein), was emerging. The laser had a power density of 2×10^7 J/s per pulse and a pulse duration of about 10 nanoseconds. The flow is thus frozen to a good approximation. The region imaged extends from 2 diameters downstream of the orifice to about 18 diameters. The Reynolds number based on the orifice diameter, the velocity at the orifice, and the viscosity of water is about 2000. (*From R. R. Prasad and K. R. Sreenivasan, Measurement and interpretation of fractal dimension of the scalar interface in turbulent flows, Phys. Fluids A, 2:792–807, 1990*)

they occur at all) of flow quantities, such as dye concentration, fluid velocity, and pressure, are smooth and gradual in time and space. Farther downstream, the jet undergoes a transition to a new state in which the eddy patterns are complex, and flow quantities (including vorticity) fluctuate randomly in time and three-dimensional space. This is the turbulent state. *See* JET FLOW; LAMINAR FLOW.

Turbulence occurs nearly everywhere in nature [in the Earth's boundary layer (extending to more than a few hundred meters upward from the ground), the jet stream, cumulus clouds, rivers and oceans, the stellar atmosphere, interstellar gas clouds, and so forth] and in technology (in flow over airplanes, flow over turbine blades, flow of natural gas and oil in pipelines, combustion systems, and so forth). Two important characteristics of turbulence are the efficient dispersion and mixing of vorticity, heat, and contaminants. In flows over solid bodies such as airplane wings or turbine blades, or in confined flows through ducts and pipelines, turbulence is responsible for increased drag and heat transfer. Turbulence is therefore a subject of great engineering interest. On the other hand, as an example of collective interaction of many coupled degrees of freedom, it is also a subject at the forefront of classical physics. *See* DEGREE OF FREEDOM (MECHANICS); DIFFUSION; HEAT TRANSFER; PIPE FLOW; PIPELINE.

Figure 1 demonstrates the principal issues associated with turbulent flows. The first is the mechanism (or mechanisms) responsible for transition from the steady laminar state to the turbulent state even though, for both states, the governing equations (the Navier-Stokes equations) are the same, and the same smooth and symmetric boundary conditions are imposed on the flow everywhere. A second issue concerns the description of fully developed turbulence typified by the complex state far downstream of the orifice. To understand and describe the essential features of these spatial patterns, their interactions and temporal evolution, and to develop on this basis a rational theory capable of predicting flow features, is at the heart of turbulence theories. Finally, it is of technological importance to be able to alter the flow behavior to suit particular needs: Delaying transition to turbulence, or promoting it, or affecting the spread rate of the jet, or decreasing the drag of an airplane wing, or relaminarizing a turbulent flow some distance downstream of where it has become turbulent are some examples. Together, these three aspects—the origin of turbulence, the dynamics of fully developed turbulence, and the control of turbulent flows—constitute the turbulence problem. The problem assumes richer complexion when effects such as buoyancy, compressibility, electromagnetic conductivity, and chemical reactions are included. In spite of sustained efforts, turbulence

has remained unsolved. Less is known about eddy motions on the scale of centimeters and millimeters than about atomic structure on the subnanometer scale, reflecting the complexity of the turbulence problem. *See* NAVIER-STOKES EQUATION.

Origin of turbulence. A central role in determining the state of fluid motion is played by the Reynolds number. In general, a given flow undergoes a succession of instabilities with increasing Reynolds number and, at some point, turbulence appears more or less abruptly. It has long been thought that the origin of turbulence can be understood by sequentially examining the instabilities. This sequence depends on the particular flow and, in many circumstances, is sensitive to a number of details even if the gross features in a given flow are nominally fixed. The program of precisely identifying the various instabilities culminating in fully developed turbulence has not been carried out for any flow, but a careful analysis of the perturbed equations of motion has resulted in a good understanding of the first two instabilities (primary and secondary) in a variety of circumstances. *See* REYNOLDS NUMBER.

Since the onset of turbulence resembles the onset of complexity in nonlinear systems in general, the universality theories describing the onset of chaos have been thought to bear on the transition to turbulence in fluid flows. The spirit of universality is that, no matter what equations govern a low-dimensional system, its behavior in the vicinity of bifurcations depends on certain generic features in phase space. This issue is an active area of research, and the experience so far has been that the onset of chaos in special types of flows under special circumstances follows these theories, at least to a very good approximation, but the relation between chaos (or temporal stochasticity) and fluid turbulence (which possesses temporal as well as spatial randomness, and large-scale order underlying the latter) remains unclear. *See* CHAOS.

Fully developed turbulence. Some of the principal difficulties in fully developed turbulence are the following: (1) The equations of motion are nonlinear, possess no general solutions, and permit few statements of general validity to be made; there is no small parameter in the problem on the basis of which approximate solutions can be deduced rationally. (2) There is no well-understood working model of turbulence that replicates its essential properties. (3) Turbulent velocity fluctuations at small scales are strongly nongaussian, this being an essential feature. (4) The number of degrees of freedom is very large. *See* DISTRIBUTION (PROBABILITY).

An estimate of the number of degrees of freedom is given by the quantity $(L/\eta)^3$, where L is the characteristic size of the large eddy in the flow (or an upper bound for the eddy scale either excited by inherent instability or forced by an outside agency), and η is the smallest scale below which all eddy motions are damped by viscosity. This number increases with the flow Reynolds number according to its 9/4 power. Three-quarters of the way downstream from the orifice in Fig. 1, the ratio L/η is of the order of 100. Although such flows can now be computed directly, the prospect at high Reynolds numbers remains discouraging—for the atmosphere, L is of the order of a few kilometers whereas η is of the order of a millimeter—even though computational capabilities have continued to increase rapidly and parallel processing has been much considered as a tool for expanding the scope of computation. *See* CONCURRENT PROCESSING; SUPERCOMPUTER.

Quite often in engineering, the detailed motion is not of interest, but only the long-time averages or means, such as the mean velocity in a boundary layer, the mean drag of an airplane or pressure loss in a pipeline, or the mean spread rate of a jet. It is therefore desirable to rewrite the Navier-Stokes equations for the mean motion. The basis for doing this is the Reynolds decomposition, which splits the overall motion into the time mean and fluctuations about the mean. These macroscopic fluctuations transport mass, momentum, and matter (in fact, by orders of magnitude more efficiently than molecular motion), and their overall effect is thus perceived to be in the form of additional transport or stress. This physical effect manifests itself as an additional stress (called the Reynolds stress) when the Navier-Stokes equations are rewritten for the mean motion (the Reynolds equations). The problem then is one of prescribing the Reynolds stress, which contains the unknown fluctuations in quadratic form. A property of turbulence is that the Reynolds stress terms are comparable to the other terms in the Reynolds equation, even when fluctuations are a small part of the overall motion. An equation for the Reynolds stress itself can be obtained by suitably manipulating the Navier-Stokes equations, but this contains third-order terms involving fluctuations, and an equation for third-order terms involves fourth-order quantities, and so forth. Thus, at any stage of the process, which can be continued indefinitely, there are more unknowns than equations; that is, the system of equations is not closed. This is the closure problem in turbulence. The Navier-Stokes equations are themselves closed, but the presence of nonlinearity and the process of averaging result in nonclosure.

Given this situation, much of the progress in the field has been due to (1) exploratory experiments and numerical simulations of the Navier-Stokes equations at low Reynolds numbers; and (2) plausible hypotheses in conjunction with dimensional reasoning, scaling arguments, and their experimental verification.

Experiments, for long the central tool of research in turbulence, are limited to measuring a small number of parameters at a few positions in high-Reynolds-number flows. Low-Reynolds-number flows (at least some of their features) can be quantitatively mapped in three dimensions by using lasers and advanced optical techniques; this Reynolds number range is also the one for which numerical simulations are currently possible. From a combination of such studies, it has been learned, among other things, that the magnitude of the dissipation rate of turbulent kinetic

energy is independent of viscosity (even though viscosity is essential for dissipation); that the boundary between the turbulent and nonturbulent regions in high-Reynolds-number free shear flows such as jets is sharp and fractallike; that the dissipation of energy is highly intermittent in space; that some events that appear to be dynamically significant are also intermittent and perhaps quasicyclic; and that, when the flow scales are suitably coarse-grained, some degree of spatial order on scales of order L is visible even at very high Reynolds numbers, especially if the flow development in time is observed. The true significance of each of these features in accomplishing transport is still under active research. *See* FRACTALS; VISCOSITY.

The intermittency in space of the turbulent energy dissipation is shown at moderate (**Fig. 2***a*) and high (Fig. 2*b*) Reynolds numbers. The signal becomes less space-filling or more intermittent as the Reynolds number increases. In particular, the big spikes in Fig. 2*b* are many times larger than the corresponding ones in Fig. 2*a*. This intermittency, representing the fact that there is a limit to the mixing at small scales, is believed to be an important feature of turbulence. It is not entirely clear how this feature arises dynamically, but it can be modeled well by a simple multiplicative process.

A classic and celebrated hypothesis is the concept of local isotropy, which assumes that small scales of motion are isotropic irrespective of the gross orientation of the mean flow, and thus possess some universality. A second notion is the matchability between behaviors of highly disparate scale ranges so that a functional form for average quantities of interest can be determined for the intermediate scale range. For example, in the turbulent boundary layer over a flat wall, this type of argument leads to a logarithmic variation of mean velocity with height for heights large compared to the viscous scale and small compared to the overall thickness of the boundary layer. Similarly, intermediate scales that are large compared to η but small compared to L (the so-called inertial range) are expected to possess self-similarity, leading to power-law variations for the spectral densities of energy, dissipation, variance of concentration fluctuations, and so forth. These predictions have received experimental support, and, in fact, they seem to be realized under conditions where they are not necessarily expected to be valid, based on first principles. This raises the possibility that the basic theoretical arguments of this type (which, incidentally, do not make much use of the Navier-Stokes equations) have a much wider range of validity. It should be emphasized that these arguments predict an absolute number in the inertial range, but not elsewhere. *See* BOUNDARY-LAYER FLOW.

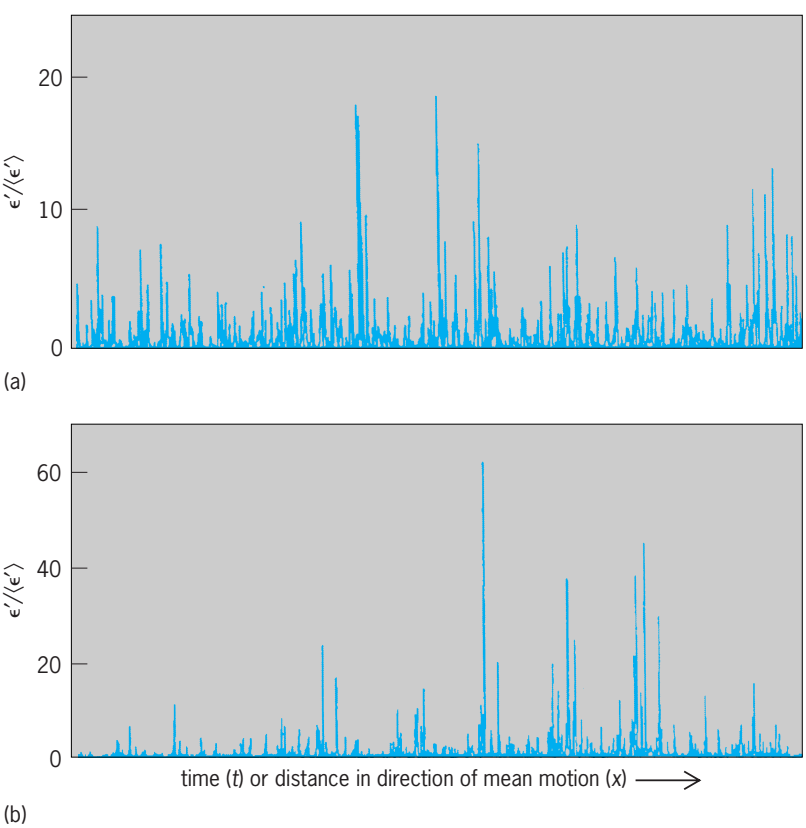

Fig. 2. Typical signals of ϵ', a component of the turbulent energy dissipation, normalized by its mean value $\langle\epsilon'\rangle$. (*a*) Signal obtained in a laboratory turbulent boundary layer at a moderate Reynolds number (defined suitably). (*b*) Signal obtained in the atmospheric surface layer at a high Reynolds number. (*After C. Meneveau and K. R. Sreenivasan, Simple multifractal cascade model for fully developed turbulence, Phys. Rev. Lett., 59:1424–1427, 1987*)

Certain specific questions of engineering interest can be answered quickly by modeling the Reynolds stress in a variety of ways and closing the Reynolds equations. The earliest model, based on analogies with molecular motion, postulated that the physical motion of eddies over well-defined distances accomplishes transport. The resulting eddy viscosity—essentially the product of a characteristic velocity scale and a length scale of turbulent motion—is the analog of molecular viscosity. Even though there are circumstances where the eddy-viscosity approach works roughly, the concept has many drawbacks and is not very useful as a general idea; in any case, the eddy viscosity varies from flow to flow and from point to point in a given flow. In the next level of models, separate equations are written for the length and velocity scales making up eddy viscosity, but there are several unknown coefficients that have to be determined empirically. There are more complex models, all of which resort to empiricism of dubious quality at some level. In spite of this drawback, they are quite useful once the limits of their validity have been established. At present, they represent a practical way of computing high-Reynolds-number flows of technological interest. Another fruitful approach is the large-eddy simulation, which models the small-scale motion but simulates large eddies on the computer.

Control of turbulent flows. Unlike several other issues in turbulence, questions concerning flow control can be posed in specific terms. However, because of this specificity, a broad-brush approach to the control problem encompassing all circumstances is unlikely to succeed. Some typical objectives are the

reduction of drag of an object such as an airplane wing, the suppression of combustion instabilities, and the suppression of vortex shedding behind bluff bodies. A surge of interest in flow control is due in part to the discovery that some turbulent flows possess a certain degree of spatial coherence at the large scale. An example of successful control, based on an unrelated idea, is the reduction of the skin friction on a flat plate by making small longitudinal grooves, the so-called riblets, on the plate surface, imitating shark skin.

Prospects. Progress in the turbulence problem depends on the capability to make accurate measurements in high-Reynolds-number flows, the increase in computer power, the invention of new tools for handling large streams of stochastic data, and a judicious combination of all of them. Unfortunately, simply computing or making measurements in a highly nonlinear system such as turbulent flow does not always add to understanding. Although several new analytical tools, graphical display capabilities, and data-compression and data-handling techniques are being explored, it is difficult to predict what true progress is likely to occur through the 1990s. It is clear, however, that turbulence will spur important activity in a number of disciplines at the forefront of science and technology; conversely, it will benefit from them. In the long run, perhaps, all these tools can enhance the qualitative understanding of turbulence; to obtain quantitative data in a specific context, reliance may always have to be placed on experiment, as well as modeling and computation that use this qualitative knowledge in a sensible way. This would resemble to some extent the situation in quantum chemistry. *See* FLUID-FLOW PRINCIPLES. K. R. Sreenivasan

Bibliography. H. Bai-Lin (ed.), *Chaos II*, 1990; D. M. Bushnell and C. B. McGinley, Turbulence control in wall flows, *Annu. Rev. Fluid Mech.*, 21:1–20, 1989; R. J. Garde, *Turbulent Flow*, 1994; J. L. Lumley, *Whither Turbulence? Or, Turbulence at Crossroads*, 1990; H. L. Swinney and J. P. Gollub (eds.), *Hydrodynamic Instabilities and the Transition to Turbulence*, 2d ed., 1985; H. Tennekes and J. L. Lumley, *A First Course in Turbulence*, 1972; A. A. Townsend, *The Structure of Turbulent Shear Flows*, 2d ed., 1976.

Turmeric

A dye or a spice obtained from the plant *Curcuma longa*, which belongs to the ginger family (Zingiberaceae). It is a stout perennial with short stem, tufted leaves, and short, thick rhizomes which contain the colorful condiment. As a natural dye, turmeric is orange-red or reddish brown, but it changes color in the presence of acids or bases. As a spice, turmeric has a decidedly musky odor and a pungent, bitter taste. It is an important item in curry and is used to flavor and color butter, cheese, pickles, and other food. *See* SPICE AND FLAVORING; ZINGIBERALES.

Perry D. Strausbaugh; Earl L. Core

Turn and bank indicator

A combination instrument which provides an aircraft pilot with two distinct pieces of information: the aircraft's rate of turn about the vertical axis, and the relationship between this rate and the aircraft's angle of bank. It is also known as the needle and ball indicator or the turn and slip indicator.

The turn needle is operated by a gyroscope and indicates the rate at which the aircraft is turning about the vertical axis in degrees per second. Semirigid mounting of the gyro permits it to rotate freely about the lateral and longitudinal axes while restricting motion about the vertical axis (see **illus.**). In a turn, gyroscopic precession causes the rotor to tilt in the direction opposite the turn with a magnitude proportional to the turn rate. A mechanical linkage converts this precession to reversed movement of a turn needle, thus indicating proper turn direction. A spring attached between the gyro assembly and the instrument case holds the gyro upright when precession force is not present and allows calibration for a given turn rate, while a damping mechanism is included to prevent excessive oscillation of the turn needle. Because of the criticality of this instrument under instrument flight conditions, it is normally powered by a separate energy source from the other flight instruments. *See* GYROSCOPE.

The bank or slip indicator is a simple inclinometer consisting of a curved glass tube containing fluid and a black ball bearing which is free to move in the fluid. The fluid provides dampening for the ball movements, and a small projection at one end of the tube contains an air bubble to compensate for fluid expansion during temperature changes. The tube is curved upward at the ends so the ball seeks the lowest point in wings-level flight. The indicator is actually a balance indication, showing the relationship between the rate of turn and the angle of bank of the aircraft. During a turn the ball experiences two forces in the horizontal plane: the component of the lift force in the horizontal plane pulling to the inside of the turn, which is created by the bank angle or the tilt of the lift force; and the centrifugal force

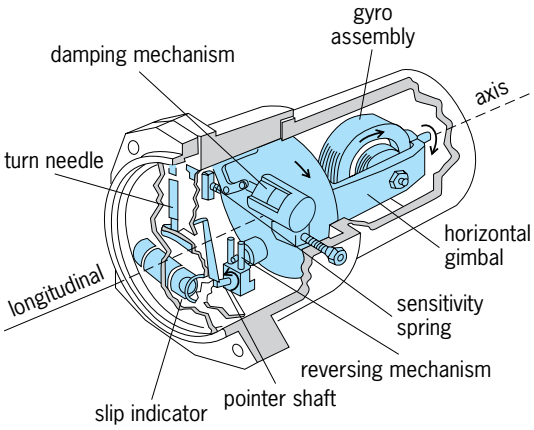

Typical turn and bank indicator. (*U.S. Air Force*)

created by the turn rate producing a force to the outside of the turn. It is the sum of these two forces that the ball displays. *See* AIRCRAFT INSTRUMENTATION; ROTATIONAL MOTION.

Grady W. Wilson

Turnbuckle

A device for tightening a rod or wire rope. Its parts are a sleeve with a screwed connection at one end and a swivel at the other or, more commonly, a sleeve with screwed connections of opposite hands (left and right) at each end so that by turning the sleeve, the connected parts will be drawn together, taking up slack and producing necessary tension (see **illus.**). Types of ends available are hook, eye, and clevis. The turnbuckle can be connected at any convenient place in the rod or rope, and several may be used in series if required. *See* ROPE; SCREW THREADS.

Paul H. Black

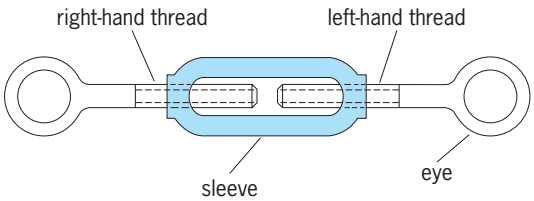

Turnbuckle with eyes.

Turning (woodworking)

The shaping of wood by rotating it in a lathe and cutting it with a chisel. The lathe consists essentially of a bed on which are mounted a headstock, a tailstock, and a tool rest (see **illus.**). The headstock is rotated by a motor and holds one end of the wood to be turned. The tailstock holds the other end of the wood, allowing it to rotate freely. The tool rest provides a fixed guide along which the operator can handle the chisels if the turning is by hand, or along which the tool is driven if the turning is mechanized.

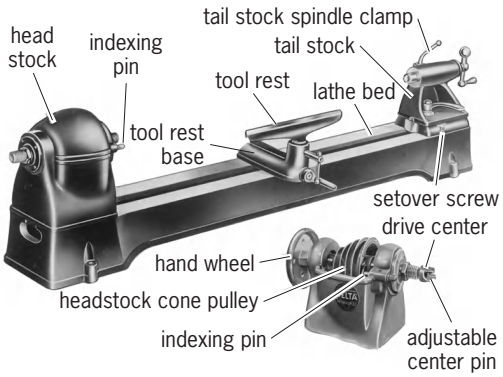

Wood-turning lathe and detail of headstock. (*Delta*)

Variously shaped chisels are used, all with longer handles than the woodworking chisels intended to be driven by mallet, and thus providing a firmer grip. A gouge is used for roughing cuts, for example, in turning the work as it comes from the sawmill to a nearly cylindrical shape. A skew chisel with a straight cutting edge is used in finishing. A parting chisel with a tapered shank is used in separating the finished work from the stock. *See* WOODWORKING.

Alan H. Tuttle

Bibliography. S. Hogbin, *Wood Turning*, 1980.

Turnip

The plant *Brassica rapa*, or *B. campestris* var. *rapa*, a cool-season, hardy crucifer of Asiatic origin belonging to the order Capparales and grown for its enlarged root and its foliage, which are eaten cooked as a vegetable (see **illus.**). The plant is an annual when planted early, a biennial if seeded late in the summer. Propagation is by seed. Popular white-fleshed varieties (cultivars) grown for their roots are Purple Top Globe and White Milan; Yellow Globe and Golden Ball are common yellow-fleshed varieties. Shogoin is a popular variety grown principally in the southern United States for turnip greens. Turnip harvesting begins when the roots are 2–3 in. (5–7.5 cm) in diameter, usually 40–70 days after planting. Principal areas of production in the United States are in the South. *See* CAPPARALES; RUTABAGA.

Turnip, of the Brassicaceae family.

H. John Carew

Turnip and rutabaga can be damaged by several diseases throughout their growth period. Seedling diseases caused by soil-inhabiting fungi may kill or debilitate young plants soon after emergence. Proper planting-site selection and preparation, along with seed spacing to avoid crowding, can minimize losses. Leaf-spotting diseases are caused by fungi that produce wind- or rain-disseminated spores. Alternaria leaf spot, anthracnose, and blackleg are important foliar diseases and may be controlled by protective fungicides. Turnip leaves grown for food are

particularly susceptible to leafspot damage. Blackleg may also damage roots before harvest and during storage.

Bacterial blackrot is often a destructive disease. Symptoms, damage, and control measures are similar to those for blackrot of cabbage.

Clubroot is a serious disease of worldwide distribution. The causal organism, which persists for several years in the soil, infects young roots; severely deformed, unattractive turnips and rutabagas are produced as a result. Infested soil should be avoided, or disease-resistant cultivars should be planted. Deformed, mottled leaves or stunted plants may be caused by several insect-transmitted viruses.

Turnips and rutabagas require adequate boron to prevent unsightly, internal breakdown of root and stem tissues. *See* CABBAGE; PLANT PATHOLOGY.

J. O. Strandberg

Bibliography. G. R. Dixon, *Vegetable Crop Diseases: American Edition*, 1981; A. F. Sherf and A. A. McNab, *Vegetable Diseases and Their Control*, 2d ed., 1986; J. C. Walker, *Diseases of Vegetable Crops*, 1952; J. C. Walker, R. H. Larson, and A. L. Taylor, *Diseases of Cabbage and Related Plants*, USDA Agr. Handb. 144, 1958.

Turquoise

A mineral of composition $CuAl_6(PO_4)_4(OH)_8 \cdot 5H_2O$ in which considerable ferrous ion (Fe^{2+}) may substitute for copper. Ferric ion (Fe^{3+}) may also substitute for part or all of the aluminum (Al), forming a complete chemical series from turquoise to chalcosiderite [$CuFe_6(PO_4)_4(OH)_8 \cdot 5H_2O$]. Turquoise with a strong sky-blue or bluish-green to apple green color is easily recognized, and such material is commonly used as a gem. Some variscite, of composition $AlPO_4 \cdot 2H_2O$ with minor chemical substitutions of Fe^{3+} and or chromium ion (Cr^{3+}) for aluminum and with a soft, clear green color, may be marketed as green turquoise.

Crystals of turquoise are rare; they are triclinic in symmetry, with space group $P\bar{1}$. Most turquoise is massive, dense, and cryptocrystalline to fine-granular. It commonly occurs as veinlets or crusts and in stalactitic or concretionary shapes. It has a hardness on the Mohs scale of about 5 to 6 and a vitreous to waxy luster. The distinctive light blue coloration of much turquoise is the result of the presence of cuprous ion (Cu^{2+}); limited substitution of the copper by Fe^{2+} produces greenish colors. *See* CRYSTAL STRUCTURE.

Occurrence. Turquoise is a secondary mineral, generally formed in arid regions by the interaction of surface waters with high-alumina igneous or sedimentary rocks. It occurs most commonly as small veins and stringers traversing more or less decomposed volcanic rocks. Since the times of antiquity, turquoise of very fine quality has been produced from a deposit in Persia (now Iran) northwest of the village of Madeň, near Nishapur. It occurs also in Siberia, Turkistan, China, the Sinai Peninsula, Germany, and France.

The southwestern United States has been a major source of turquoise, especially the states of Nevada, Arizona, New Mexico, and Colorado. Extensive deposits in the Los Cerillos Mountains, near Santa Fe, New Mexico, were mined very early by Native Americans and were a major early source of gem turquoise. However, much of the gem-quality turquoise has been depleted in the Southwest. Some mining districts that still supply gem material are the Sleeping Beauty district between Globe and Miami, Arizona; Morenci, Arizona; Crow Springs, Stormey Mountain, and Blue Moon district, Nevada; Menassa, Colorado; and Hachita and Oro Grande, New Mexico.

Imitations. Because of much demand for turquoise at relatively low prices and because of the scarcity of gem-quality turquoise, which is generally high priced, many materials that are not completely natural or are turquoise-colored imitations have appeared on the market. These varieties are known as stabilized, oil-treated, treated, reconstituted, and imitation turquoise.

Stabilized turquoise is a poor-quality natural turquoise (often referred to as chalk in the trade) that has been chemically impregnated and hardened with organic resins to improve the color and the hardness of the final product. Such material becomes very workable, permanently hardened, and stable, and can be very attractive. Major sources of such low-grade turquoise are found in Mexico and China.

Oil-treated turquoise is a material in which the color of the natural turquoise has been enhanced by impregnation with oil, paraffin, or oil-based polishes. Such treatment generally results in only temporary improvement of color, with subsequent fading.

Treated turquoise is a natural or stabilized turquoise that has been altered (dyed) to produce a change in the coloration of the natural material.

Reconstituted turquoise is fabricated by bonding dust, turquoise particles, or nuggets together with plastic resins.

Imitation turquoise results from a process in which a natural compound is treated or a synthetic compound is manufactured to closely approximate turquoise's appearance. Included are turquoise-colored plastics, glass enamel, and dyed chalcedony.

Stabilized, oil-treated, and reconstituted turquoise can be distinguished by quantitative chemical microanalysis for carbon (C), hydrogen (H), and nitrogen (N), the elements that were added to the original material in the various modification processes. Treated turquoise can generally be recognized by a well-trained eye. Physical testing usually suffices to distinguish the materials comprising imitation turquoise. *See* GEM; PHOSPHATE MINERALS.

Cornelis Klein

Bibliography. C. S. Hurlbut, Jr., and R. C. Kammerling, *Gemology*, 1991; C. Klein, and C. S. Hurlbut, Jr., *Manual of Mineralogy*, 21st ed., 1993; C. Palache, H. Berman, and C. Frondel (eds.), *System of Mineralogy*, vol. 2, 7th ed., 1951.

Twilight

The period between sunset and darkness in the evening and darkness and sunrise in the morning. The following statements apply to evening twilight; the reverse would apply to morning twilight.

The characteristic light is caused by atmospheric scattering, which transmits sunlight to the observer for some time after the Sun has set. It depends geometrically on latitude, longitude, and elevation of the observer, and on the time of the year. Physically it depends also on local conditions, particularly the weather. *See* METEOROLOGICAL OPTICS.

Three degrees of twilight are conventionally distinguished. Civil twilight ends when the center of the Sun is 6° below the horizon; if the sky is clear, it is usually practicable to carry on ordinary outdoor occupations without artificial light during civil twilight. Nautical twilight ends when the depression of the Sun is 12°; at this time both the horizon and the brighter stars are visible. Astronomical twilight ends when the depression of the Sun is 18°; at this time no trace of illumination by the Sun appears in the sky. As thus defined, the times of ending of the three sorts of twilight can be precisely calculated.

Since the angle at which the Sun sets depends on an observer's latitude, twilights are relatively short near the Equator, where the Sun sets perpendicularly, and can last for days near or above the Arctic and Antarctic circles. *See* ANTARCTIC CIRCLE; ARCTIC CIRCLE. Gerald M. Clemence; Jay M. Pasachoff

Bibliography. A. B. Meinel and M. P. Meinel, *Sunsets, Twilights, and Evening Skies*, 1991; G. V. Rozenberg, *Twilight*, 1966; V. J. Schaefer and J. A. Day, *A Field Guide to the Atmosphere*, 1999.

Twinkling stars

A phenomenon by which light from the stars, as it passes through fluctuations in the Earth's atmosphere, is rapidly modulated and redirected to make the starlight appear to flicker. Although it is familiar to those who have looked with the unaided eye at the night sky, the twinkling phenomenon affects all wavelengths that manage to penetrate the Earth's atmosphere, from the visible to the radio wavelengths. At visible wavelengths, atmospheric fluctuations are caused predominantly by temperature irregularities along the line of sight. Minor contributions are made by irregularities in atmospheric density and in water vapor content. All such irregularities introduce slight changes in the index of refraction of air, and these changes affect light waves in two ways: they modulate the intensity of the light, and they deflect the light waves in one direction and then another. An analogous phenomenon is often observed when light grazes across the surface of a hot highway: light is bent and distorted by pockets of hot air rising over the pavement's surface. At radio wavelengths, electron density irregularities in the ionosphere modulate and redirect radio waves. *See* REFRACTION OF WAVES.

Unaided-eye observations. High-altitude winds and local air currents carry atmospheric irregularities across the line of sight. Because the entrance pupil of the human eye is much smaller than the characteristic size of the atmospheric irregularities, any twinkling that is visible to the unaided eye is caused primarily by the modulation of the light intensity and not by the deflection of the light waves. Modulation is best understood as an interference phenomenon acting on adjacent light waves. Like random waves in water, sometimes the waves add constructively, causing the star to brighten, and other times the waves add destructively, causing the star to dim. *See* INTERFERENCE OF WAVES.

While stars twinkle, planets generally do not. This difference occurs because stars are essentially infinitesimally small points of light. All light from a star travels along the same path through the atmosphere, and all the light from a star is modulated simultaneously. Planets have a definite angular extension. Light from one side of the planet traverses one path through the atmosphere, while light from the other side of the planet traverses another path through the atmosphere. Along each path the modulation of light is different, so that when all the light is added together by the human eye the total modulation or twinkling is averaged out. *See* PLANET; STAR.

Telescope observations. The twinkling phenomenon is of utmost interest to astronomers who view the skies from ground-based telescopes. While modulation variations are present, it is the deflection of light that causes the most serious problems. The entrance pupil of an optical telescope is often much larger than the characteristic size of atmospheric irregularities. The composite star image produced by a large telescope is a blurry circle that results when the randomly deflected light waves are added together in an extended time exposure. To diminish atmospheric effects, telescopes are built on high mountains, and are placed at least 30–45 m (100–150 ft) above the ground. The best observatories are those where the twinkling phenomenon is minimized. Examples include the Mauna Kea Observatory in Hawaii and observatories in the foothills of the Andes Mountains in northern Chile. *See* OBSERVATORY, ASTRONOMICAL.

To completely remove the twinkling effects of the atmosphere, there are two alternatives. The first is to place a telescope in orbit above the atmosphere, as with the Hubble Space Telescope. The second alternative is to monitor the random deflections of the atmosphere and, within the telescope, to bend the deflected light back onto its original path. This optical technique is given the name adaptive optics. *See* ADAPTIVE OPTICS.

Radio sources. Small radio sources have been found to vary rapidly in brightness because of the flow of density fluctuations in the solar wind across the path between the radio source and the Earth. The method of interplanetary scintillations has proved a useful means of probing the interplanetary medium near the Sun and also of establishing the very small size of some of the radio sources. Radio sources

whose angular diameters are large do not show interplanetary scintillation at all, just as planets do not share in the stellar twinkling. Pulsars show erratic changes of amplitude in their radio emission, caused by radio scintillation both near the neutron star and in interstellar space. *See* PULSAR; RADIO ASTRONOMY; SOLAR WIND. Laird A. Thompson

Bibliography. J. M. Beckers, Adaptive optics for astronomy: principles, performance, and applications, *Annu. Rev. Astron. Astrophys.*, 31:13-62, 1993; H. G. Booker and W. E. Gordon, A theory of radio scattering in the troposphere, *Proc. Inst. Radio Eng.*, 38:401–412, 1950; B. J. Rickett, Interstellar scattering and scintillation of radio waves, *Annu. Rev. Astron. Astrophys.*, 15:479–504, 1977; N. J. Woolf, High resolution imaging from the ground, *Annu. Rev. Astron. Astrophys.*, 19:367–398, 1981; A. T. Young, Seeing: Its cause and cure, *Astrophys. J.*, 189:587–604, 1974.

Twinning (crystallography)

A process in which two or more crystals, or parts of crystals, assume orientations such that one may be brought to coincidence with the other by reflection across a plane or by rotation about an axis. Crystal twins represent a particularly symmetric kind of grain boundary; however, the energy of the twin boundary is much lower than that of the general grain boundary because some of the atoms in the twin interface are in the correct positions relative to each other.

In the general grain boundary, all the neighbors of the atoms of the interface are in distorted positions. The usual definition of a twin relationship between two crystals states that there exists a set of parallel equivalent crystal planes of atoms which is common to both twins, but that rows of atoms are discontinuous across the interface. Quite commonly, twins are mirror images of each other, as in the **illustration**. Also, it is common for a twin in a crystal to leave the nearest neighbors of the atoms in the interface unchanged in orientation, but to place the atoms in the second neighbor shell in altered positions. This feature is also true of the twin in the illustration. It is sometimes possible to create a twin in a crystal by putting an external stress on the crystal; in other cases, twins are found "grown in." *See* CRYSTAL GROWTH. Robb M. Thomson

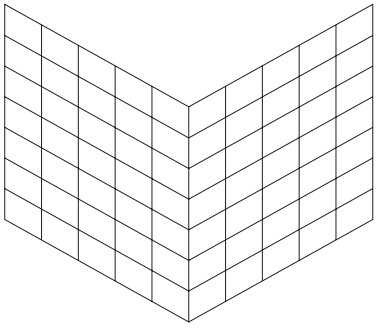

Example of a twinned crystal. One atom plane is common to each half of the crystal, but the other lines of atoms suffer a discontinuity at the twin boundary.

Bibliography. A. Holden and P. S. Morrison, *Crystals and Crystal Growing*, 1982; R. Raj and S. L. Sass (eds.), Interface Science and Engineering. '87, *Journale de Physique*, vol. 49, colloque C5, 1989; J. P. Van der Eerden, *Fundamentals of Crystal Growth*, 1993.

Twins (human)

Two babies born to a mother at one birth. Knowledge about the biological bases of twinning, as well as sophistication in techniques for data collection, research design, and analysis, have increased substantially in recent years. Consequently, twin research has been incorporated into a growing number of behavioral science and medical science research programs.

Biology of twinning. There are two types of twins, monozygotic and dizygotic. Members of a twin pair are called co-twins. Controversy surrounding the definition of a twin arose with the advent of reproductive technologies enabling the simultaneous fertilization of eggs, with separate implantation. The 1996 cloning of Dolly the lamb in Scotland directed attention toward the promises and pitfalls of possible human cloning. The unique "twinlike" relationships that would result between parents and cloned children (who would be genetically identical to their parents) also challenge current conceptions of twinship. Monozygotic twins are clones (genetically identical individuals derived from a single fertilized egg), but parents and cloned children would not be twins for several reasons, such as their differing prenatal and postnatal environments. *See* REPRODUCTIVE TECHNOLOGY.

Monozygotic twins. The division of a single fertilized egg (or zygote) between 1 and 14 days postconception results in monozygotic twins. They share virtually all their genes and, with very rare exception due to unusual embryological events, are of the same sex.

A common assumption is that because monozygotic co-twins have a shared heredity, their behavioral or physical differences are fully explained by environmental factors. These environmental influences may occur during the prenatal, perinatal, or postnatal periods. However, monozygotic twins are never exactly alike in any measured trait, and may even differ for genetic reasons. For example, the random "shutting down" of one X chromosome in every female cell soon after conception (a process called lyonization) can cause monozygotic female twins to differ in X-linked traits, such as color blindness. *See* COLOR VISION; SEX-LINKED INHERITANCE.

Sometimes chromosomes fail to separate after fertilization, causing some cells to contain the normal chromosome number (46) and others to contain an abnormal number. This process, called mosaicism, results in monozygotic co-twins who differ in chromosomal constitution. These unusual cellular processes explain the presence of monozygotic pairs in

which one co-twin is normal while the other shows a genetic anomaly reflecting a mixture of normal and abnormal cells. A rare case of monozygotic triplets including two healthy males and a co-triplet with Turner's syndrome (loss of one X chromosome) has been reported. *See* MOSAICISM.

There are several other intriguing variations of monozygotic twinning. Splitting of the zygote after day 7 or 8 may lead to mirror-image reversal in certain traits, such as handedness or direction of hair whorl. The timing of zygotic division has also been associated with placentation. Monozygotic twins resulting from earlier zygotic division have separate placentae and fetal membranes (chorion and amnion), while monozygotic twins resulting from later zygotic division share some or all of these structures. Compared with two-chorion monozygotic twins, monozygotic twins sharing a chorion are more likely to be born prematurely, to differ in birth weight, and to die early (in extreme cases). Associations between mirror-image reversal and placentation are expected, but relationships among these events do not appear to be straightforward. Should the zygote divide after 14 days, the twins may fail to separate completely. This process, known as conjoined twinning, occurs in approximately 1 monozygotic twin birth out of 200. The many varieties of conjoined twins differ as to the nature and extent of their shared anatomy. Approximately 70% of such twins are female. There do not appear to be any predisposing factors to conjoined twinning.

Dizygotic twins. Dizygotic twins result when two different eggs undergo fertilization by two different spermatozoa, not necessarily at the same time. Dizygotic twins share, on average, 50% of their genes, by descent, so that the genetic relationship between dizygotic co-twins is exactly the same as that of ordinary brothers or sisters. Theoretically, dizygotic twins may share between 0 and 100% of their genetic makeup, but most are close to the 50% average. Some dizygotic co-twins share higher or lower proportions of genes for certain traits, so they may be more or less similar in those traits. Dizygotic twins may be of the same or opposite sex, outcomes that occur with approximately equal frequency.

There are some unusual variations of dizygotic twinning. There is the possibility of polar body twinning, whereby divisions of the ovum prior to fertilization by separate spermatozoa could result in twins whose genetic relatedness falls between that of monozygotic and dizygotic twins, or between dizygotic twins and unrelated individuals. Blood chimerism, another variation, refers to the presence of more than one distinct red blood cell population, derived from two zygotes, and has been explained by connections between two placentae. In humans, chimerism can occur in dizygotic twins. New techniques estimate that chimerism occurs in 8% of dizygotic twins and 21% of dizygotic triplets. Superfecundation is the conception of dizygotic twins following separate fertilizations, usually within several days, in which case each co-twin could have a different father. Superfecundation may cause significant developmental discrepancies between co-twins due to their differing paternal heredity. Superfetation, which refers to multiple conceptions occurring several weeks or even one month apart, may be evidenced by delivery of full-term infants separated by weeks or months and by the birth or abortion of twin infants displaying differential developmental status. *See* FERTILIZATION; OOGENESIS.

Epidemiology. According to conventional twinning rates, monozygotic twins represent approximately one-third of twins born in Caucasian populations and occur at a rate of 3–4 per 1000 births. The biological events responsible for monozygotic twinning are not well understood. It is generally agreed that monozygotic twinning occurs randomly and not as a genetically transmitted tendency. Some recent evidence from Sweden suggests an increased tendency for mothers who are monozygotic twins to bear same-sex twins themselves; further work will be needed to resolve this question.

Dizygotic twinning represents approximately two-thirds of twins born in Caucasian populations. The dizygotic twinning rate is lowest among Asian populations (2 per 1000 births), intermediate among Caucasian populations (8 per 1000 births), and highest among African populations (50 per 1000 births in parts of Nigeria). The natural twinning rate increases with maternal age, up to between 35 and 39 years, and then declines. A possible causal mechanism is the increased ovarian activity that continues until a woman reaches her late thirties. Elevated levels of follicle-stimulating hormone (FSH) may increase the probability of dizygotic twinning among some women. Dizygotic twinning has also been linked to increased parity, or the number of children to which a woman has previously given birth. However, parity is associated with older maternal age, which is more closely linked to dizygotic twinning. Mothers of dizygotic twins are also significantly taller and heavier, on average, than mothers of monozygotic twins and singletons. *See* OVARY.

Dizygotic twinning appears to be genetically influenced, although the pattern of transmission within families is unclear. One study found that parents of both monozygotic and dizygotic twins are twice as often left-handed as their own same-sex singleton siblings, but the specific mechanism linking twinning and left-handedness is unknown. Recent work in New Zealand uncovered a gene in sheep associated with dizygotic twins and triplets that could assist the understanding of the genetics of human multiple birth.

A higher frequency of dizygotic twinning among lower social classes than among higher social classes has been reported. This relationship may simply reflect the larger size of the lower-class families. A relationship may exist between dizygotic twinning and dietary factors, but it could be quite complex: A reduction in the dizygotic twinning rate has been associated with reduced nutritional supply in some European countries during World War II, although there have been mixed trends regarding nutrition and overall twinning rates. Some researchers have

implicated the yam (a food with estrogenlike substances) in the extremely high dizygotic twinning rate among the Yoruba of western Nigeria, although this is not definitive. Higher and more constant food supplies may explain the increased dizygotic twinning rate in southwest Finland's archipelago of Aland and Aboland, relative to the mainland.

Increased coital rate, following extended periods of sexual abstinence, may also increase the dizygotic twinning rate. The rate of dizygotic twinning is elevated in extramarital conceptions, during the first year of marriage, and in couples immediately after soldiers return from war. It may be that, in some women, coitus stimulates ovulation, and frequent ovulation leads to superfecundation. Increased dizygotic twinning has also been associated with an increased number of marriages.

The United States' twinning rate rose 65% between 1980 and 2002 and 38% between 1990 and 2002. It went up 3% between 2001 and 2002, to 31 twin births per 1,000 births. These increases were mainly due to advances in fertility treatments (for example, in vitro fertilization and ovulation induction), but were also due to delays in child-bearing. The increase mostly involved dizygotic twinning, in which multiple ovulation and maternal age are key factors. In vitro fertilization has, to a smaller extent, increased monozygotic twinning rates, possibly due to laboratory manipulation of the early embryo. However, in 2002 the birth rate of triplets and other higher order multiples dropped slightly to 184 per 100,000 births. This represented the third such decrease in the previous four years, following a 400% increase between 1980 and 1998.

The birth rate of preterm babies, defined as infants born after less than 37 weeks' gestation, increased 12.1% in 2002, and increased by 15% since 1990. This change partly reflects the increase in the rate of multiple births, which usually occur early.

Determination of zygosity. The accurate assignment of twin pairs as monozygotic or dizygotic is crucial for avoiding misleading estimates of genetic influence upon traits of interest. Co-twin comparison of selected DNA markers (six to eight) is the most accurate (greater than 99%) scientific method for classifying twins as monozygotic or dizygotic. Except for monozygotic twins, it would be extremely rare for family members to show matching DNA patterns. Cells for analysis can be obtained easily by gently scratching the inner cheek to obtain a buccal smear. Several laboratories are equipped to receive samples by mail and to provide twin-typing results in approximately 2 weeks. Blood-typing is another objective method for classifying twin pairs. Blood-group differences identify dizygotic co-twins with complete certainty. However, an absence of blood group differences can indicate only the probability that the twins are monozygotic, because in rare cases dizygotic co-twins may share all measured blood groups because of shared parentage. Analyses of anthropometric characteristics, such as fingerprint ridgecount (dermatoglyphic analysis), ponderal index, and cephalic index, improve the accuracy of twin classification when used in conjunction with blood-typing. *See* BLOOD GROUPS; DEOXYRIBONUCLEIC ACID (DNA); FINGERPRINT.

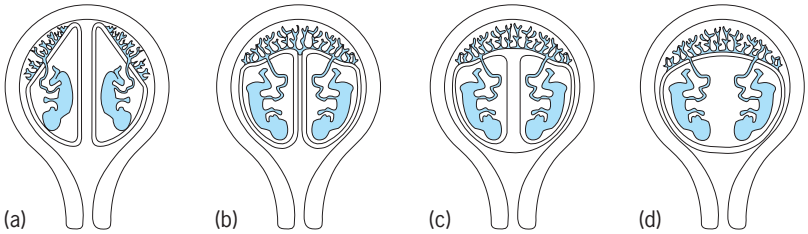

Fig. 1. Arrangements of placentae and placental membranes for monozygotic and dizygotic twin pairs in the uterus. (*a*) Monozygotic or dizygotic twins with separate amnions, chorions, and placentae. (*b*) Monozygotic or dizygotic twins with separate amnions and chorions and fused placentae. (*c*) Monozygotic twins with separate amnions and single chorion and placenta. (*d*) Monozygotic twins with single amnion, chorion, and placenta. Analysis of these arrangements can aid in diagnosing zygosity. (*After C. E. L. Potter, Fundamentals of Human Reproduction, McGraw-Hill, 1948*)

Other methods for diagnosing zygosity include physical-resemblance questionnaires and examination of placentae and placental membranes (**Fig. 1**). The more accurate of these methods is the physical resemblance questionnaire, which shows 93–99% agreement with the results from blood typing. Twins' and parents' judgments of twin type are often questionable.

Twin studies. Francis Galton reasoned late in the nineteenth century that comparative analyses of the behavioral and physical resemblance of the two types of twins would provide useful insights into the relative contributions of heredity and environment. (Although the biology of twinning was not established during Galton's time, he correctly recognized that there were two types of twins.) In particular, he suggested that greater resemblance within twin pairs derived from a single ovum (monozygotic twins), relative to twin pairs derived from separate ova (dizygotic twins), would demonstrate a contribution from genetic factors. This monozygotic–dizygotic twin comparison, called the classic twin method, assumes that differences between one-egg twins are generally due to environmental effects, given their shared heredity, and that differences between two-egg twins are due to genetic and environmental differences between them. The twin method additionally assumes that trait-relevant environmental influences are equal for both types of twins (equal environments assumption). Galton obtained detailed analyses of the life histories of twins, some of whom displayed early physical and behavioral similarities, and some an absence of early resemblance. He concluded that nature made a major contribution to many domains of human development.

Challenges to the classic twin method have been raised, despite its successful implementation in numerous behavioral science and medical science fields. Social biases refer to biases in treatment of monozygotic and dizygotic twins; for example, some critics of twin studies assert that there might be more similar treatment of monozygotic than dizygotic twins that may amplify behavioral similarity between the former and reinforce behavioral

dissimilarity between the latter. Empirical testing has, however, demonstrated that similarity of treatment is unrelated to similarity of behavior among monozygotic twins. Primary biases refer to unique prenatal influences, such as shared intrauterine circulation, which tend to reduce resemblance between monozygotic co-twins. These effects do not, however, appear to seriously affect results from twin data based upon large samples. Recruitment biases refer to differences in representation among the types of twins who volunteer for research: approximately two-thirds of adult same-sex twin volunteers are female, and two-thirds are monozygotic. Twin studies carried out in Norway and the United States (Minnesota), however, were unable to demonstrate differences in resemblance for intelligence quotient (IQ), personality traits and interests, or demographic variables between solicited pairs in which both co-twins completed questionnaires by mail, as compared with pairs in which one twin participated. Twins who volunteer for research without solicitation, however, may not be representative of the general twin population. Another concern is that some twin studies are based on small samples, limiting the power of their conclusions.

Intraclass correlation. The intraclass correlation (r_i) is the preferred measure of resemblance between co-twins. It is based on a ratio of the variation between groups to the total variation (group refers to a twin pair) and is given by the equation below, where

$$r_i = \frac{S^2b - S^2w}{S^2b + (n-1)S^2w}$$

S^2b = variance between families, S^2w = variance within families, and n = average number of individuals per family.

Heritability expresses the proportion of population variation associated with genetic differences in a given characteristic. It is a function of the particular population under study, the time of measurement, and the measuring instrument, and is therefore not a constant figure but one that reflects changes in these factors. Subtracting the dizygotic intraclass correlation from the monozygotic intraclass correlation estimates half the genetic effect because dizygotic twins share half as many genes, on average, as do monozygotic twins. Doubling the resulting value yields an estimate of the genetic effect. This procedure works only if the genetic effects are additive; that is, the effects of the genes are constant, regardless of the genes with which they are paired. However, sometimes genes behave interactively or nonadditively—that is, the effects of the genes change depending upon the genes with which they are paired. This may increase the monozygotic twin intraclass correlation and heritability estimates because monozygotic twins always share gene combinations while dizygotic twins do not.

Methods for analyzing twin data have moved beyond analyses of correlations to more complex model-fitting approaches. Researchers develop a model specifying genetic and environmental relatedness and apply it to data from twins and other relatives. They can then estimate genetic and environmental contributions to the trait, based on how well the model fits the data.

Twin and adoption studies. Twins reared apart offer informative tests of genetic and environmental influences on behavior. If monozygotic twins are separated early in infancy and raised in separate homes selected at random, their degree of similarity (intraclass correlation) provides a direct estimate of genetic effects. The study of dizygotic twins reared apart enables additional tests of the possibility of genetic interactions. Another approach is to study the adopted siblings of twins that are reared apart, a procedure that enables comparison between individuals who share environments but not genes; this is the reverse of studying separated monozygotic twins except that each twin and his/her adopted sibling are not the same age. This can be overcome by studying "virtual twins," genetically unrelated children of the same age raised together from infancy. They may consist of two adopted children or one adopted child and one near-in-age biological child in the family. They share their environments but not their genes so they truly are the reverse arrangement of monozygotic twins reared apart. One laboratory in the United States (at California State University, Fullerton) is studying these "twinlike" sibships to determine the extent to which common environments affect behavioral similarities.

Comparison of resemblance between monozygotic and dizygotic twins reared together and those reared apart also provides information on the influence of common rearing on behavioral and physical similarity. The level of personality resemblance is similar for monozygotic twins reared apart and reared together (approximately 50% of the variance is associated with genetic factors), demonstrating that common environmental influences do not contribute importantly to personality similarity. Resemblance in the body mass index (which is the weight in kilograms divided by the height, in squared meters) is only slightly lower for monozygotic twins reared apart than for monozygotic twins reared together. These analyses suggest that both genetic factors and environmental factors unique to the individual are important influences on these characteristics.

Twin gender studies. There has been recent interest in gender-related behaviors of opposite-sex twins. This is due to findings in the nonhuman literature showing that female mice and gerbils demonstrate masculine behaviors if situated next to a male fetus in utero. It is suggested that prenatal exposure to male hormones may be responsible for these results. Human twin research has produced mixed findings in this regard. It is also important to consider that female co-twins' behaviors could reflect the psychological and social effects of being raised with a same-age male sibling. However, the finding that females with male co-twins show the same frequency of otoacoustic emissions ("hums" discharged in the inner ear to raise the volume of weak sounds) as their brothers is difficult to explain without reference to opposite-sex

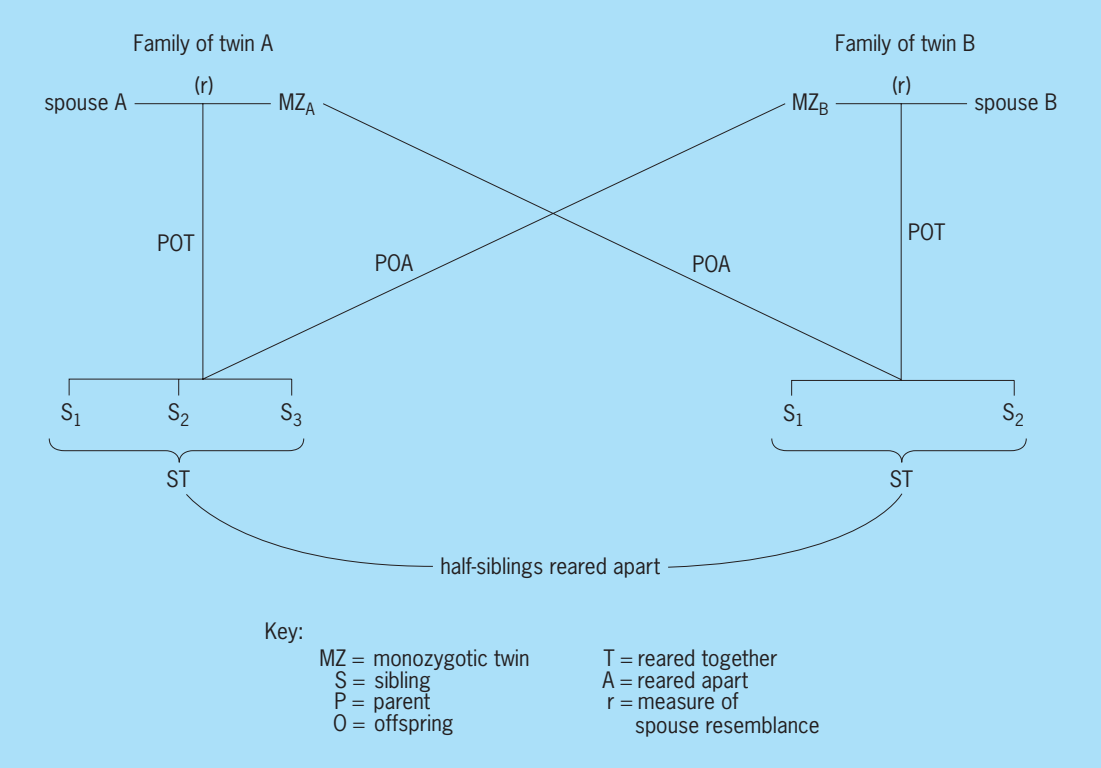

Fig. 2. Twin-family design. (*Adapted from S. Scarr-Salapatek, in K. W. Schaie, ed., Developmental Human Behavior Genetics*)

twins' prenatal biological circumstances.

Twin-family studies. The offspring of monozygotic twins share unique relationships with one another. The children of monozygotic twins are genetically equivalent to half-siblings, because they each have a parent that is genetically identical (either a twin mother or a twin father). Similarly, monozygotic twins share the same genetic relationship with the co-twin's children (nieces and nephews) as they do with their own children. This family constellation thus enables comparisons between co-twins, marital partners, twins and children, twins with nieces and nephews, full siblings, and half-siblings. Studies of mental abilities and physical characteristics that use the twin-family method are available. Studying social relationships among members of these unique families is also potentially informative with respect to genetic and social influences (**Fig. 2**).

Twins as couples. This approach to twin research compares monozygotic and dizygotic twins' social relationships with reference to their joint behaviors, such as cooperation and competition. The focus of interest is the behavior of the pair (for example, their level of cooperation or competition), not their separate behaviors as in classic twin studies. Research using this design has generally revealed greater cooperation between monozygotic twins than dizygotic twins in experimental games and tasks. These tendencies mirror the increased social closeness that monozygotic twins typically express toward one another and the increased grief they experience following the loss of their co-twin, relative to dizygotic twins. There is, however, overlap among twin types—some dizygotic twins are especially close to one another, while some monozygotic twins are less so.

QTL studies. Quantitative traits, such as height, intelligence, and certain aspects of personality, are influenced by many genes working together. People vary continuously on such traits, ranging from short to tall; average to highly intelligent; and outgoing to shy. Quantitative trait loci (QTLs) are the many genes that affect these traits, although they each differ in their effect. Identifying genes with the greatest influence on traits of interest can tell us about individual differences in the development of that trait. Now that the Human Genome Project has mapped the full complement of 20,000 to 25,000 genes, behavioral scientists and molecular geneticists have increased their efforts at finding QTLs for mental retardation, aggression, and psychopathology. It has how been shown that individuals with the apolipoprotein e4 gene are at increased risk for late-onset Alzheimer's disease. Twin studies will play important roles in the quest to find QTLs. The most valuable information may come from studying monozygotic twins who differ in their expression of disease and behavior. In the case of twins differing in schizophrenia, it may be that one twin was exposed to specific environmental factors that activate relevant genes or QTLs, while the other twin was not. Furthermore, even monozygotic twins are not genetically identical, as indicated above. Unraveling the link between these genetic differences and behavior will be an informative, albeit challenging task. *See* HUMAN GENOME PROJECT.

Twin registries. Twin studies depend on large samples in order to draw valid conclusions about behavior and development. Statistically significant findings are those occurring by chance fewer than one time in twenty. Statistical significance is associated with sample size and the size of the correlation. A .25 correlation would be statistically significant with 45 twin pairs, while a .50 correlation would be statistically significant with 12 twin pairs. Large twin samples are also required for estimating heritability, especially if the heritability of a given trait is low in the population. However, a large number of subjects does not ensure valid findings because the quality of the information may be difficult to determine. It has also been shown that heritability may be calculated more accurately and efficiently using twins reared apart rather than twins reared together. For example, 400–500 monozygotic and 400–500 dizygotic twin pairs reared together allow heritability to be estimated with the same degree of confidence as 50 monozygotic twin pairs reared apart. This is because heritability is estimated directly from the reared apart twins and indirectly from the reared together twins.

Modern twin research has come a long way in recruiting twins since Francis Galton did his famous study of nature and nurture in 1875. Galton gathered twins by sending notices to twins and relatives of twins. Respondents were asked to provide the names and addresses of other twins that they knew, thus enlarging his potential participant pool. Galton received 80 replies, although he did not indicate what percentage of his initial inquiries this number represented. Thirty-five twin pairs, presumably monozygotic, showed "close similarity" based on written responses to his questions. Galton clearly showed that twin studies could illuminate genetic and environmental influences on human behavior. However, findings from small samples gathered unsystematically urge cautious interpretation of the results. Fortunately, there is now a growing number of population-based twin registries.

The oldest twin registry in the world is the Danish Twin Registry, established in 1954. It now includes over 65,000 twin pairs, comprising 127 birth cohorts. A recent review summarized characteristics for 16 national twin registries, located in the United States, Australia, Belgium, Canada, China, Denmark, Finland, Germany, Italy, Japan, Korea, the Netherlands, Norway, Sri Lanka, Sweden, and the United Kingdom. There are also statewide twin rosters such as the Minnesota Twin Registry and the Wisconsin Twin Panel, and commercial registries such as the United Kingdom's *Gemini Genomics*, dedicated to finding disease-related genes. Investigators interested in specific conditions can try to identify specific subjects within these sources. There are, however, registries that target twins with particular conditions such as HIV exposure and chronic fatigue syndrome. Some registries include specific subgroups of twins, such as World War II veterans, Vietnam veterans, and elderly citizens. Aside from offering researchers access to large and diverse twin samples, the registries facilitate international collaboration. Studies of rare medical conditions can proceed by pooling cases across sites. Attempts at cross-cultural replication of findings can also be conducted.

Value of twin studies. Studies of monozygotic and dizygotic twins provide valuable information on genetic and environmental influences on human behavioral and physical development. Studies of twins, therefore, offer many opportunities for examining novel hypotheses and generating new explanations of observations. For example, twin studies have been used to assess predictions generated by human evolutionary biology and theories of economics. New model-fitting approaches to data analysis enable testing for the presence of specific genetic and environmental effects. Refinement of twin samples and the study of more well-defined characteristics is another important trend. In recent years, twin research has begun to address the origins of less frequently studied behaviors such as humor, love styles, athleticism, and religiosity. Continued extension and elaboration of the twin design, replication of research, and continued testing of assumptions will help researchers to improve twin methods and respond to criticism. *See* BEHAVIOR GENETICS; HUMAN GENETICS.

Nancy L. Segal

Bibliography. M. G. Bulmer, *The Biology of Twinning in Man*, Clarendon Press, Oxford, 1970; A. Busjahn (ed.), Special issue on Twin Registries across the globe: What's out there in 2002, *Twin Res.*, no. 5, 2002; F. Galton, The history of twins as a criterion of the relative powers of nature and nurture, *J. Anthropol. Inst.*, 6:391–406, 1875; International Human Genome Sequencing Consortium, Finishing the euchromatic sequence of the human genome, *Nature*, 43:931–945, 2004; K. K. Kirk and N. G. Martin (eds.), Special issue on religion, values and health, *Twin Res.*, no. 5, 1999; D. T. Lykken, S. Geisser, and A. Tellegen, Heritability estimates from twin studies: The efficiency of the MZA design (unpublished manuscript); G. A. Machin et al., Correlations of placental vascular anatomy and clinical outcomes in 69 monochorionic twin pregnancies, *Amer. J. Med. Gen.*, 61:229–236, 1996; P. McGuffin, B. Riley, and R. Plomin, Genomics and behavior: Toward behavioral genomics, *Science*, 291:1232–1249, 2001; National Center for Health Statistics, News Release, Dec. 17, 2003; N. L. Segal, *Entwined Lives: Twins and What They Tell Us About Human Behavior*, Plume, New York, 2000; N. L. Segal, Human cloning: A twin-research perspective. *Hastings Law J.*, 53:1073–1084, 2002; N. L. Segal, Virtual twins: New findings on within-family environmental influences on intelligence, *J. Educ. Psych.*, 92:442–448, 2000.

Two-phase flow

The simultaneous flow of two phases or two immiscible liquids within common boundaries. Two-phase flow has a wide range of engineering applications

such as in the power generation, chemical, and oil industries. Flows of this type are important for the design of steam generators (steam–water flow), internal combustion engines, jet engines, condensers, cooling towers, extraction and distillation processes, refrigeration systems, and pipelines for transport of gas and oil mixtures.

The most important characteristic of two-phase flow is the existence of interfaces, which separate the phases, and the associated discontinuities in the properties across the phase interfaces. Because of the deformable nature of gas-liquid and liquid-liquid interfaces, a considerable number of interface configurations are possible. Consequently, the various heat and mass transfers that occur between a two-phase mixture and a surrounding surface, as well as between the two phases, depend strongly on the two-phase flow regimes. Multiphase flow, when the flow under consideration contains more than two separate phases, is a natural extension of these principles. *See* FLUID MECHANICS; INTERFACE OF PHASES.

From a fundamental point of view, two-phase flow may be classified according to the phases involved as (1) gas-solid mixture, (2) gas-liquid mixture, (3) liquid-solid mixture, and (4) two-immiscible-liquids mixture. *See* GAS; LIQUID; PHASE EQUILIBRIUM.

Flow regimes in vertical pipes. For up-flow in vertical pipes, several different flow regimes can be observed for gas-liquid (or two-immiscible-liquids) mixtures.

Bubbly flow. Bubbly flow is the most widely known flow pattern and is recognizable from everyday experience. Here the gas phase is distributed in discrete bubbles within a liquid continuum.

Slug flow. As the bubble concentration becomes very high in a bubbly flow, coalescence may produce much larger bubbles. In this type of flow, some of the gas bubbles have nearly the same cross section as that of the channel, and their length may be many times the diameter. The gas bubbles or plugs have spherical caps and flat tails and move along like bullet-shaped bubbles (**Fig. 1**). The bubble is separated from the pipe wall by a descending liquid film.

Churn flow. If the velocity of a two-phase mixture flowing as a slug flow is increased, the flow structure will eventually become unstable and the gas plugs will begin to break up, leading to an unstable regime in which the gas flows in a chaotic manner through the liquid (Fig. 1). The liquid is mainly displaced to the channel walls and flows in oscillatory motion, upward and downward.

Annular flow. In this regime, the gas flows mainly in a central core, with the liquid flow almost completely confined to a liquid layer or film on the channel walls. This presents a more or less continuous interface to a stream consisting mainly of gas flowing in the center of the channel.

Drop flow. The drop flow regime (also called mist flow or liquid dispersed flow) is one in which there is complete dispersion of the liquid in a continuum gas flow. This is usually found within flow in heated channels which, incidentally, has been found to contain all the previous mentioned flow patterns developed in the order described here.

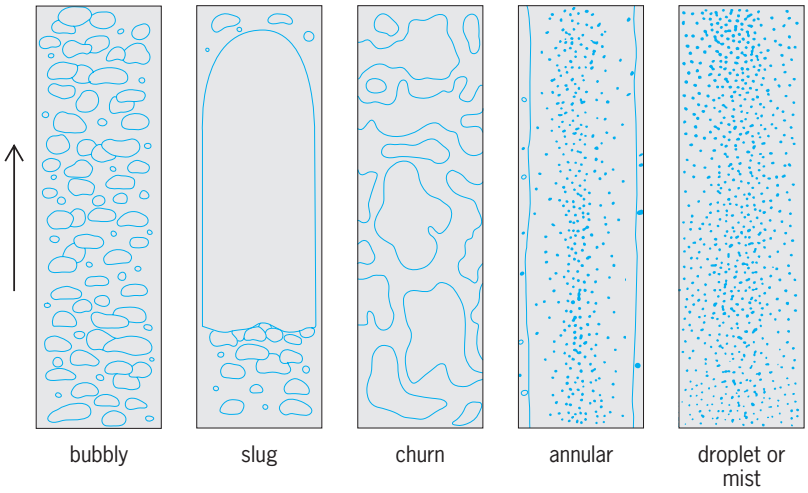

Fig. 1. Flow patterns of the gas-liquid flow in a vertical pipe.

Flow regimes in horizontal pipes. For flow in horizontal (and inclined) pipes, the flow patterns are more complex because gravity causes an asymmetric distribution of the phases (**Fig. 2**).

Bubbly flow. This is defined as in vertical flow, but there is a tendency for the bubbles to flow in the upper part of the channel (Fig. 2). At higher flow rates, the whole cross section may contain bubbles, but there will be a higher concentration toward the upper part of the duct.

Plug flow. The characteristic bullet-shaped bubbles occur as in vertical slug flow. However, in this case the plugs do not occupy as large a proportion of the cross section, and the liquid layer, separating the gas bubble from the wall, tends to be thicker at the bottom of the channel than at the top. In addition, the nose of the bubble is asymmetric (Fig. 2).

Stratified flow. At low velocities, the two fluids may flow in complete separation, with a flat interface. The liquid flows at the bottom of the channel and the gas at the top.

Wavy flow. As the gas velocity is increased in stratified flow, surface waves begin to build up on the liquid layer. This region is called wavy flow.

Slug flow. As the gas velocity is further increased in the wavy flow region, the waves grow in size, eventually reaching the top of the channel. These are propagated as slugs at high velocity and wet the whole of the channel surface, leaving a liquid film covering the surface in between the bridging waves.

Annular flow. At very high gas flow rates, the slugs become pierced, creating a gas core, and the flow becomes essentially annular. As in vertical flows, there may be some droplet entrainment and a dispersion of bubbles may exist in the liquid film. The film is very much thicker at the bottom of the channel than at the top.

Droplet flow. This is much the same as for vertical droplet flow, except the highest concentration of

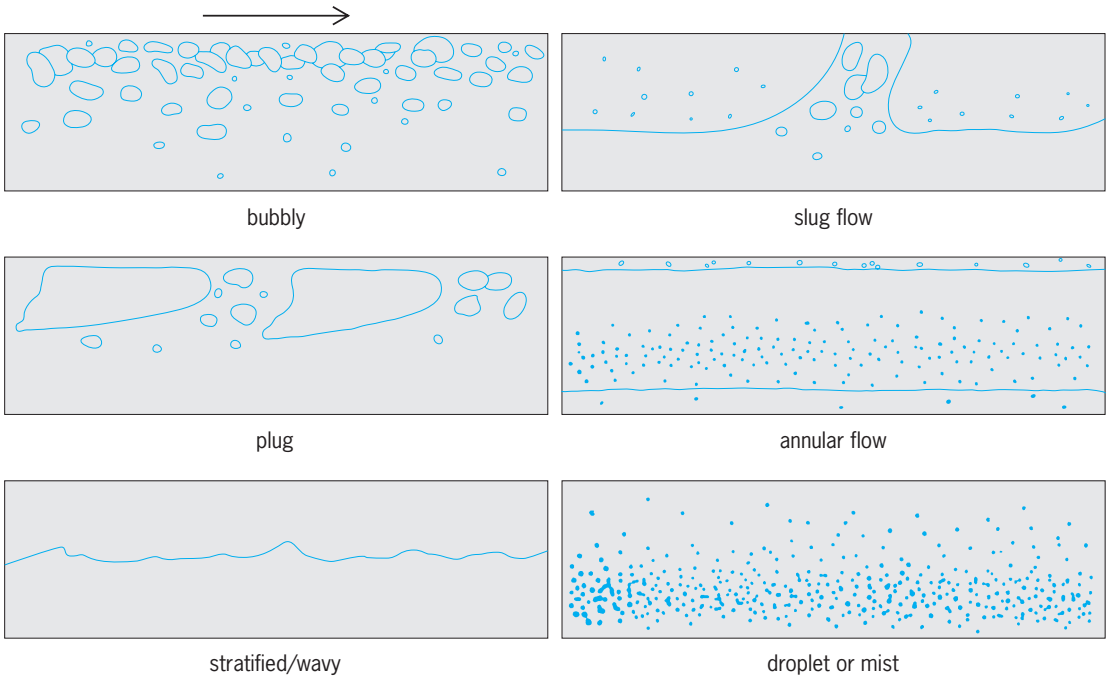

Fig. 2. Flow patterns of gas-liquid flow in a horizontal pipe.

droplets is below the pipe centerline.

Two-phase flow measurements. The primary condition for all two-phase flow is specified by the volumetric quality, which is defined as the ratio of the dispersed-phase flow rate to the total flow rate. Other required fundamental measurements include the average volumetric quality (obtained using methods such as quick-closing valves, and neutron, x-ray, and gamma-ray techniques) and local properties such as the void fraction or volume fraction, continuous- and dispersed-phase velocities, turbulent intensity, and bubble (drop) size. Dual optical and dual resistivity probes have been used for void fraction, bubble velocity, and bubble diameter measurements. The local void fraction (gas-liquid) or volume fraction (liquid-liquid) is defined (using a time-history record from a local probe) as the ratio of the time of the dispersed phase to the total measuring time. Continuous-phase velocity and turbulent intensity data have been obtained with hot-film anemometry. *See* FLOW MEASUREMENT.

Two-phase flow applications. Industrial applications of two-phase flow include systems that convert between phases, and systems that separate or mix phases without converting them (adiabatic systems).

Phase-change systems. Many of the practical cycles used to convert heat to work use a working fluid. In two or more of the components of these cycles, heat is either added to or removed from the working fluid, which may be accompanied by a phase-change process. Examples of these applications include steam generators, evaporators, and condensers, air conditioning and refrigeration systems, and steam power plants.

The majority of electricity-generating plants are variations of steam power plants. The function of the boiler in a fossil-fuel power plant is to supply the energy needed to vaporize the water, which passes through the tubes in the boiler. In nuclear power plants, the energy is released by a controlled nuclear reaction taking place in a separate reactor building. This energy is initially transferred to either pressurized water or to a liquid metal reactor coolant, with a subsequent energy transfer to the water circuit in secondary heat exchangers. Solar power plants have receivers for concentrating and collecting solar radiation in order to vaporize the working fluid. Regardless of the energy source, the vapor produced in the boiler expands, thus causing the rotation of the turbine shaft, which is connected to an electrical generator. The lower-pressure vapor leaving the turbine passes through a heat exchanger called a condenser. There it condenses on the outside of tubes carrying cooling water. The cooling water is pumped to a cooling tower, where the energy taken up in the condenser is released into the atmosphere by evaporation. *See* SOLAR ENERGY; STEAM ELECTRIC GENERATOR.

Adiabatic systems. In these types of systems, the process of phase mixing or separation occurs without heat transfer or phase change. Examples of these systems are airlift pumps, pipeline transport of gas-oil mixtures, and gas-pulverization of solid particles. *See* ADIABATIC PROCESS.

There are many different designs for phase separation that use the principles of gravity and centrifugal force. **Figure 3** shows a separating hydrocyclone in which the effect of centrifugal forces is used for solid-fluid separation or fluid-fluid separation. Hydrocyclones do not have any rotating parts, and the necessary vortex action is produced by pumping the

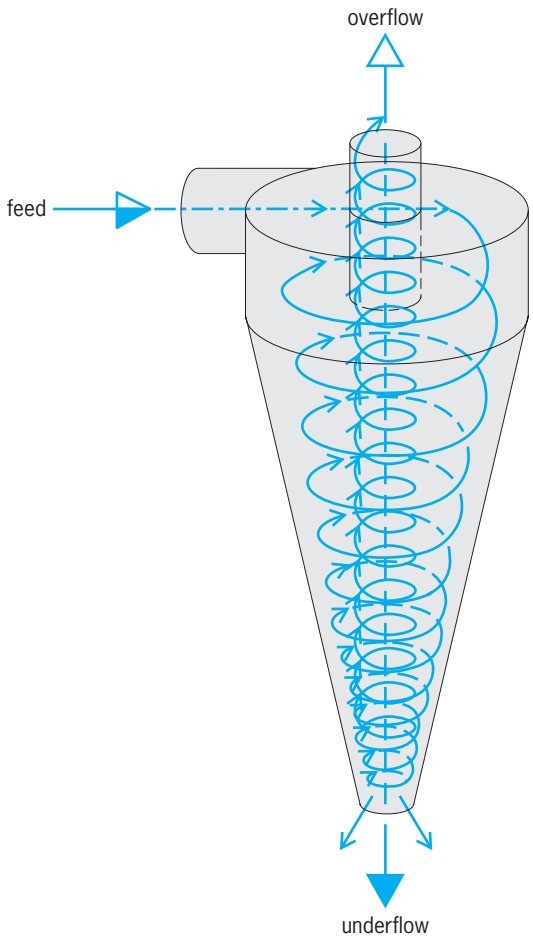

Fig. 3. Hydrocyclone phase-separation system.

fluid tangentially into a cylindrical/conical body. The end-plate, which covers the top of the cylindrical section, contains the liquid overflow pipe, which usually protrudes some distance into the cyclone body. The underflow, which carries the higher-density fluid or solids, leaves through the opening in the apex of the cone. H. H. Bruun; F. Hamad; B. K. Pierscionek

Bibliography. H. H. Bruun, *Hot-wire Anemometry: Principle and Signal Analysis*, Oxford University Press, 1995; M. J. Moran and H. N. Shapiro, *Fundamental of Engineering Thermodynamics*, John Wiley, New York, 1995; M. C. Roco, *Particulate Two-Phase Flow*, Butterworth-Heinemann, Boston, 1993; L. R. Weatherly, *Engineering Processes for Bioseparations*, Butterworth-Heinemann, Boston, 1994.

Tylenchida

An order of nematodes in which the labial region is variable and may be distinctly set off or smoothly rounded and well developed; the hexaradiate symmetry is most often retained or discernible. The amphids are porelike, oval, or slitlike, or clefts located on the lips. The internal circlet of six sensilla may be lacking. The external circle of 10 sensilla is often evident; however, these may be reduced to a visible 4 or some may be doubled. The hollow stylet is the product of the cheilostome (conus, "guiding apparatus," and framework) and the esophastome (shaft and knobs). Throughout the order and its suborders the stylet may be present or absent and may or may not be adorned with knobs. The variable esophagus is most often divisible into the corpus, isthmus, and glandular posterior bulb. The corpus is divisible into the procorpus and metacorpus. The metacorpus is generally valved but may not occur in some females and males, and the absence is characteristic of some taxa. The orifice of the dorsal esophageal gland opens either into the anterior procorpus or just anterior to the metacorporal valve. The excretory system is asymmetrical, and there is but one longitudinal collecting tubule. Females have one or two genital branches; when only one branch is present, it is anteriorly directed. Except for sex-reversed males there is only one genital branch. Males may have one (=phasmid) or more caudal papillae. The spicules are always paired and variable in shape; they may or may not be accompanied by a gubernaculum.

The order comprises five superfamilies: Tylenchoidea, Criconematoidea, Aphelenchoidea, Aphelenchoidoidea, and Sphaerularoidea.

Tylenchoidea. The superfamily Tylenchoidea comprises parasites of fungi, plants, and some insects. The most destructive and widespread endoparasitic plant nematodes are found in the families Pratylenchidae and Heteroderidae. The "meadow" or "lesion" nematode *Pratylenchus* is an obligate migratory endoparasite whose feeding results in cell death. In order to maintain themselves these nematodes must continuously seek out living cells in all directions, forming an ever-growing lesion. Nematode lesions can eventually girdle and kill roots.

Among heteroderids, species of interest include the cyst nematodes and root-knot nematodes. These two are very similar in biology but differ slightly in their life cycles and root response. They are extremely important to agriculture throughout the world and affect more than 2000 plant species.

Criconematoidea. This is a superfamily of ectoparasitic nematodes of plants ranging from migratory to highly specialized sessile forms that emulate endoparasites. Sedentary root parasitism is accomplished in several ways. In some species the spear or feeding stylet is nearly one-half the length of the body. This feature allows the nematode to stay in one place and feed on several cells by probing deeply and widely. Other species induce terminal root galls that are attractive to other individuals of the same species. The most sophisticated sessile forms embed the anterior portion of the body into the root; their feeding induces the production of giant nutritive plant cells. Two important species are *Criconemoides xenoplax* and *Tylenchulus semipenetrans*. The former, through excessive root damage, predisposes peach trees to bacterial canker, an aerial disease caused by *Pseudomonas syringae. Tylenchulus semipenetrans* is commonly known as citrus nematode and is the most important citrus parasite.

Aphelenchoidoidea and Aphelenchoidea. Aphelenchoidoidea includes parasitic forms, while Aphelenchoidea comprises mycetophagous soil inhabitants. As plant parasites, the Aphelenchoidoidea most often attack the aboveground parts; a few utilize insects as phoretic hosts to transport them from plant to plant. There are several important species of plant parasites that attack the aboveground parts, such as buds and leaves. Some important crops include strawberries, lilies, rice, chrysanthemums, and currants. An extremely important species is *Bursaphelenchus xylophilus*, on Japanese red and black pine. Seedlings can be killed by this nematode in 40 days. The nematode is transported from tree to tree by longhorn beetles in the family Cerambycidae. A single beetle, on the average, transports some 15,000 infective larvae. Another important disease is red ring of coconut, caused by *Rhadinaphelenchus cocophilus*. In this instance, the nematode is transported from tree to tree in the hemocoel of the weevil *Rhynchophorus palmarum*. Nematodes are deposited on palms when the weevil attempts to oviposit. *See* NEMATA; PLANT PATHOLOGY.

Sphaerularoidea. This is a superfamily of parasitic nematodes. Adult females are hemocoel parasites of insects and mites; a few taxa contain both plant and insect parasites. In general, nematodes belonging to this group have three distinct phases in their life cycles: two free-living and one parasitic. In the free-living phase the female gonad is single, anteriorly directed, and with few developing oocytes and a prominent uterus filled with sperm. In the parasitic phase either the body becomes grossly enlarged and degenerates to a reproductive sac, or the uterus prolapses and gonadal development takes place outside the body. The males are always free-living and not infective. The most interesting taxon is *Sphaerularia bombi*, which prolapses the uterus. When totally prolapsed, the gonad becomes the parasite and the original female a useless appendage. It is not unusual for the prolapsed gonad to attain a volume 15,000 times that of the original female. Armand R. Maggenti

Tyndall effect

Visible scattering of light along the path of a beam of light as it passes through a system containing discontinuities. The luminous path of the beam of light is called a Tyndall cone. An example is shown in the **illustration**. In colloidal systems the brilliance of the Tyndall cone is directly dependent on the magnitude of the difference in refractive index between the particle and the medium. In aqueous gold sols, where the difference in refractive index is high, strong Tyndall cones are observed.

For systems of particles with diameters less than one-twentieth the wavelength of light, the light scattered from a polychromatic beam is predominantly blue in color and is polarized to a degree which depends on the angle between the observer and the incident beam. The blue color of tobacco smoke is an example of Tyndall blue. As particles are increased in size, the blue color of scattered light disappears and the scattered radiation appears white. If this scattered light is received through a nicol prism which is oriented to extinguish the vertically polarized scattered light, the blue color appears again in increased brilliance. This is called residual blue, and its intensity varies as the inverse eighth power of the wavelength. *See* COLLOID; SCATTERING OF ELECTROMAGNETIC RADIATION. Quentin Van Winkle

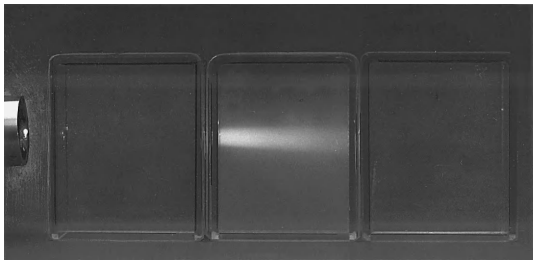

Luminous light path known as the Tyndall cone or Tyndall effect. (*Courtesy of H. Steeves and R. G. Babcock*)

Type (printing)

The intelligible images organized into readable text of various styles and sizes. Type used in printing on paper or in video display is divided into four categories: foundry, machine-cast, photocomposed, and digitized type. In the first two the face of the letter is raised on one end of a piece of metal. It is from that surface, when inked, that the impression of type was made from its invention until the 1970s. In photocomposition, the type is reproduced photographically. Digitized methods assemble dots into typographic letters, lines, or pages.

Classification. Foundry type, also known as hand type, is cast as single characters in much the same way that Johannes Gutenberg, the inventor of movable type, produced them in Mainz, Germany, about 1440.

Machine-cast metal type is produced by Linotype, Intertype, Ludlow, and Monotype machines. Very few of these machines are still in operation. All but the last cast type in lines, or slugs. The Monotype—in reality two devices, a keyboard and a caster—produces individual types that are then manually set in lines of desired lengths. The Linotype, the earliest of the mechanical composing machines, was introduced in 1886; it was invented by Ottmar Mergenthaler.

The output of the photographic type machine is the image of type on film or on photosensitized paper in negative or positive form.

The term "cold type" is applied to text matter produced on a typewriter or laser printer and to words or lines made up of individual printed characters assembled or pasted together for photographic reproduction. Digital type uses a number of printout technologies, imaging small dots into lines (called rasters) and positioning them on pages from patterns of zeros and ones called bitmaps. Because this method is electronic, more sizes and variations (for

Venetian		ABC
Old Style	French	ABC
	Dutch-English	ABC
Transitional		ABC
Modern		ABC
Contemporary	sans serif	ABC
	square serif	ABC
Scripts		ABC
Black Letter		ABC
Decorative Letters		ABC

Fig. 1. Eight type classifications.

example, condensing or expanding or other distortion) are possible.

Of the more than 50,000 type styles created since Gutenberg's time, about 6000 are in everyday use throughout the world. The most widely used method for classifying them is the serif-evolution system, based on the different shapes of the terminals or endings of letters. This provides eight classifications: Venetian, Old Style (Dutch-English and French), Transitional, Modern, Contemporary (sans serif and square serif), Scripts, Black Letter, and Decorative Letters (**Fig. 1**).

Type measurements. Type is set in sizes from 4 point to 144 point. The sizes of type in most common use are 6, 7, 8, 9, 10, 11, 12, 14, 16, 18, 24, 30, 36, 42, 48, 60, and 72 point (**Fig. 2**). Modern sizes can range from 2 point to 500 point in increments up to three decimals (that is, 9.363 point), depending on the software.

The American point system, adopted in 1886, made the unit of type measurement a point, 0.01383 in. (0.35128 mm) or nearly 1/72 in., and all type sizes are multiples of this unit. This system replaced the sixteenth-century practice of giving all type sizes names such as pearl, agate, nonpareil, brevier, long primer, and pica. Some names remained and have been assigned other functions. Agate, $5^1/_2$-point type or 14 lines of type to an inch, has come to be used for measuring newspaper advertising space. Publications quote small space rates by the agate line. Nonpareil is used to designate 6 points of space in and around type. Pica is commonly used to denote a unit of space measuring 12 points. It is applied as a dimensional unit to the length of type lines and to the width and depth of page. With the advent of desktop publishing, many software programs have standardized on 72 points equaling exactly 1 inch. With these programs, the point now equals 0.01389 in., compared with the traditional point, 0.01383 in. Some of these programs allow the user to select which point system to use.

In continental Europe the unit of type measurement is the didot point, which equals 0.0148 in.; in the United States the point is 0.0138 in. In the didot system, the horizontal unit of measurement is the cicero, which equals 12 didot points, or 0.178 in. (4.51 mm). With the increasing internationalization of design measurement systems due to the use of computers, the didot point now is rarely used.

The standard height of metal type in the United States is 0.918 in. (2.3 cm), a dimension called typehigh and one observed by photoengravers and electrotypers who make plates to be combined with cast type. Letterpress printing presses are adjusted to fit this standard height.

The hot composition type methods used metal composed of an alloy of lead, tin, and antimony. Large sizes, 72–144 point or more, were sometimes made of wood. Type size is the distance measured in points from above the ascending or capital letters to below the descending letters. In the days of metal setting, the point size of type was dependent on the point size of the body that the metal letter was cast onto. Thus, there is often a variance in the actual letter size of different fonts of the same point size, steming from the design and casting of letters in metal. Thus, the exact measurement from the top to the baseline of one 8-point A may not be the same as another 8-point A. The x-height is the height of the lowercase x, which is the typical height of the "readable" parts of the lowercase letters. It is the best determinant of type size.

Fonts. A complete complement of letters of one size and style from A to Z, together with the arabic numerals, punctuation, footnote reference marks, and symbols, is called a font (**Fig. 3**). Most fonts also include the ampersand and currency symbols and the ligatures, ff, fe, ffi, fl, ffl, æ, and œ. A letter from a different font that is included by accident in type composition is known as a wrong font. All typographic images, from letters to symbols to oriental ideographs, are called glyphs.

Type families. A family of type may be likened to the shades of a color in that it includes variations of a given type face or design. Weights are varied, from light to medium, bold, and extra bold; letters are condensed and expanded, as well as outlined, inlined, and shadowed (**Fig. 4**).

Space inserted between the lines of type to open them up vertically is called leading. Space inserted between letters of a word to open it up horizontally is known as letter spacing. Modifying space between two letters is called kerning. Modifying space between a series of letters is called tracking.

Special designs. Ornaments and rules (lines) can also be cast on machines or typeset digitally. These include individual designs, known as printer's flowers, fleurons, or florets; braces and brackets used to

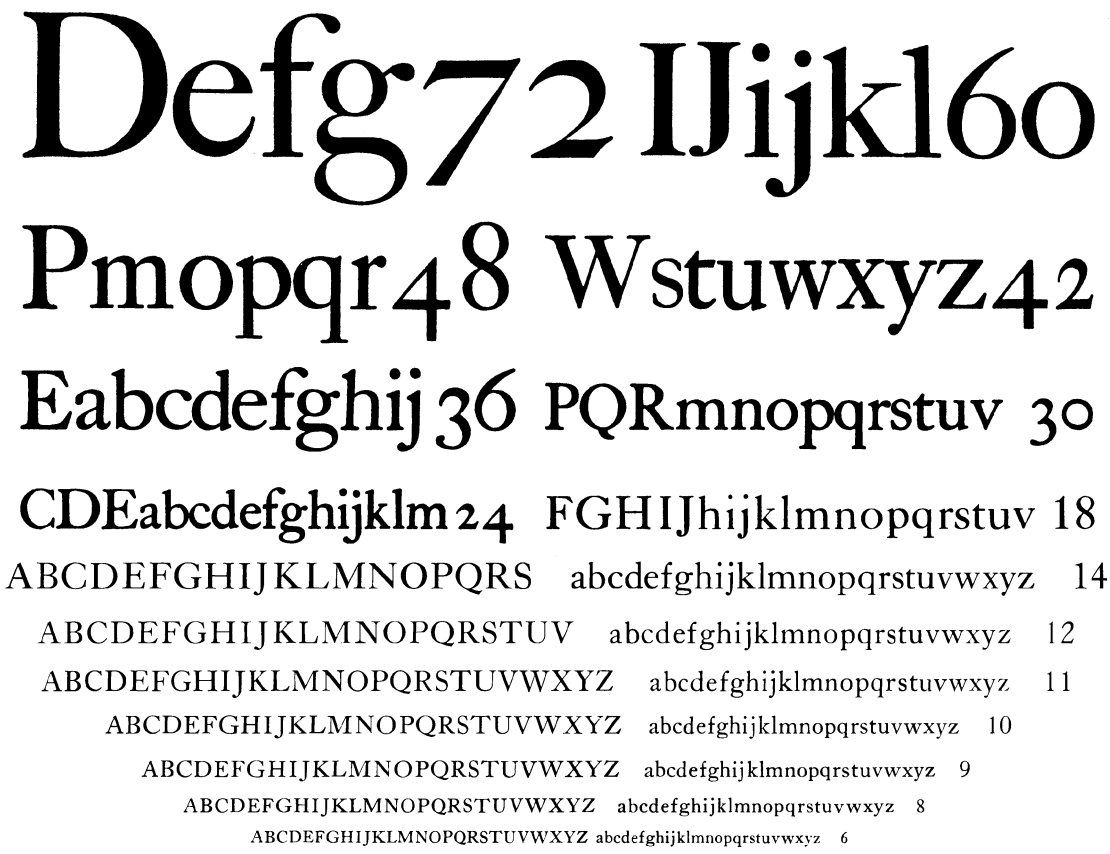

Fig. 2. Sizes of type in most common use.

enclose or connect lines of type; chemical, mathematical, astronomical, and medical signs; and signs of the zodiac. Special symbols are called pi symbols or dingbats.

The arabic numerals introduced into Europe in the twelfth century are available in many fonts as old style and modern (or lining). They are part of the fonts of typefaces derived from designs of the period 1400–1800.

Each age since the Italian Renaissance has produced memorable type designers. Among them are Nicolas Jenson, Aldus Manutius (first to use italics and small caps), Claude Garamond, William Caslon, John Baskerville, Giambattista Bodoni, Frederic W. Goudy, W. A. Dwiggins, Rudolf Koch, Paul Renner, Jan van Krimpen, Stanley Morison, Eric Gill, Bruce Rogers, Matthew Carter, Hermann Zapf, and Erik Spiekermann.

Those who stress type in their design of books, publications, and advertising are known variously as typographers, typographic designers, or type directors. The size of type you are now reading is 9 point and is set $16^1/_2$ picas wide. *See* PRINTING.

Type for desktop publishing. Most typesetting today is performed using desktop publishing (DTP), which refers to the creation of documents through the use of a personal computer (small enough to fit on a desktop), off-the-shelf software that uses the PostScript® page description language, and laser-based output devices (consisting of a raster image processor and a marking engine) that convert the digital information into images on paper and other surfaces, such

Fig. 3. A font showing roman capitals, roman lowercase, small capitals, italic capitals, and italic style.

Fig. 4. A family of type.

as printing plates for offset lithography.

The type used for desktop publishing is also digital. Most of the fonts used today are Type 1, TrueType®, or OpenType® technology.

Type 1 fonts. The font technology known as Type 1 was developed by Adobe® as part of the PostScript language, and is the most widely used and best-supported method for imaging text on high-resolution output devices such as imagesetters and platesetters. Type 1 fonts consist of two components: the screen font and the printer font. Type 1 printer fonts (also called outline fonts) contain detailed information about the character's shape, expressed as mathematical formulas called vectors. This outline information consists of the location of the anchor points (which designate the corners and other important spots on the outline of the letters) and the shape of the lines drawn between the anchor points. These data are needed by the raster image processor (RIP), where it is used to create a high-quality representation of the typeface on the output device. The printer font's outline can be sent to the RIP as part of a PostScript file, or it can be downloaded (copied through the network) as needed from the desktop publishing workstation, or it can be "resident" in the RIP's font cache (a special partition of memory used to store font outlines).

The other component of a Type 1 font is called the screen font because it was developed to provide an on-screen representation of the font's shape at the desktop publishing workstation. Since all information shown on a computer monitor must somehow be converted to pixels (picture elements) in order to be displayed, the screen font consists of a pixel-by-pixel representation of each character at the low resolution used by the monitor (72 dots per inch).

In the early days of desktop publishing, screen fonts were supplied in only a limited number of common sizes (typically 6, 8, 10, 12, 14, 16, and 24 points). Any attempt to use the font at a size that did not have a matching screen font resulted in a poor representation of the font on the screen. Since each letter was captured as a digital image, any manipulation of text outlines (such as changes in scaling or kerning) resulted in jagged type. This common scenario came to be known as "bitmapped" text, since the poor on-screen representation made it obvious that the screen font was a bitmap made of pixels instead of a smooth vector outline. This problem was largely eliminated in 1989 by the introduction of a software program called Adobe Type Manager® (ATM) that performed on-the-fly rasterization of the printer font's vector equations into bitmaps for on-screen display.

TrueType fonts. TrueType is a more recent development in digital font technology. Unlike the Type 1 fonts, TrueType fonts have only one component. The same file provides information for the computer display and the output device. This is possible because capabilities very similar to ATM (the ability to rasterize outline fonts for the monitor on-the-fly) are built into both the Macintosh and the Microsoft Windows® operating system. TrueType also has improved facilities for controlling the "hinting" of its characters when displayed at low resolution. This is very important for improving the readability of small text on low-resolution devices such as computer monitors (72 dpi) and office laser printers (300–600 dpi).

Another technical distinction can be seen in the formulas used to describe font outlines: TrueType fonts tend to have more anchor points than similar Type 1 fonts due to the method used for describing curved lines within a PostScript font (known as Bézier). TrueType fonts use less complex (but easier to render) quadratic calculations; this reduced complexity requires more points to express elaborate curves. Conversion from quadratic to Bézier (or vice versa) is not lossless, so the transformation of fonts from one technology to another may lead to very slight changes in the shape of characters.

OpenType fonts. The OpenType format was jointly developed by Microsoft and Adobe Systems and is based on both TrueType and Type 1 font formats. It is a cross-platform font format that can be used on both Macintosh and Windows systems and has an expanded character set based on the international Unicode encoding standard. As a result, OpenType supports more languages, has greater typographic capabilities, and simplifies font management requirements. In addition, OpenType is compatible and can be used in the same document with Type 1 and TrueType fonts.

Eugene M. Ettenberg; Frank J. Romano; Thomas Destree; Roger Walton

Bibliography. M. Annenberg and S. O. Saxe, *Type Foundaries of America and Their Catalogs*, 1994; A. Faiola, *Typography Primer*, 2000; R. Goldberg and F. Romano, *Digital Typography*, 2000; A. Haley, *Hot Type Designers Make Cool Fonts*, 1998; R. Mclean, *Typographers on Type: An Illustrated Anthology from William Morris to the Present Day*, 1995; A. W. White, *Type in Use: Effective Typography for Electronic Publishing*, 1999.

Type method

The nomenclatural method for providing a fixed reference point for a taxon (plural, taxa), a group of organisms. When Linnaeus, the eighteenth-century

Swedish naturalist, ushered in the modern period of systematic biology through the publication of his classical work, *Systema Naturae*, he established the basis of binomial nomenclature. By means of this method the systematist who recognizes new, or supposedly new, taxa makes these known by means of a technical description and a scientific name. If the new taxon is a species, the material (specimens or parts of specimens) which the author had before him when he described it, together with such additional material as he may have gathered later, was considered typical, and such specimens were known as types. The types represented the author's notion of the new species and served as the basis of comparison. As knowledge increased, it was realized that each taxonomic name must be tied to a single specimen—the "type"—which is a nomenclatural, not a biological, type. Having a fixed reference for each name, later workers could correct taxonomic errors and modify the limits of species taxa without subsequent confusion in scientific names.

Type designations. For the first 50 or 100 years after Linnaeus, taxonomists did not make formal type designations. In fact, James E. Smith replaced some of Linnaeus's original specimens with better ones when they became available. Later taxonomists generally designated one specimen to serve as the type (or holotype) for newly described species. For existing species with competing syntypes or cotypes, one was selected as the type, usually called the lectotype. As the number of collections and type repositories increased and became less accessible to the taxonomist, as types became lost or destroyed, and as complications arose in trying to apply the single type concept retroactively, more refined type terminology and methods of type designation were developed. Although terms are available for more than 100 kinds of "type" specimens, these fall in two major classes: (1) nomenclatural types, the name-bearing specimens (onomatophores), which serve as reference points for nomenclatural purposes, and (2) taxonomic types which serve as standards of description and reference points for authors' concepts. The first category has been sanctioned by the International Commission on Zoological Nomenclature, the second category by taxonomic practice.

Among the more commonly used terms to designate these two kinds of types, nomenclatural and taxonomic, are the following:

1. *Nomenclatural types*

a. Holotype: the single specimen designated or indicated as "the type" by the original author at the time of publication of the original description or the only specimen known at the time of the original description.

b. Syntype (also called cotype): every specimen on which an author bases an original description when no single specimen has been designated as the holotype.

c. Lectotype: one of a series of syntypes which is selected subsequent to the original description and thenceforth serves as the definitive type of the species. To be effective, such selection must be made known through publication.

d. Neotype: a specimen selected as type subsequent to the original description when the primary types are definitely known to be destroyed. Here again selections must be made known through publication.

2. *Taxonomic types*

a. Paratype: a specimen other than the holotype which is before the author at the time of original description and which is designated as such or is clearly indicated as being one of the specimens upon which the original description was based.

b. Allotype: a paratype of the opposite sex to the holotype which is designated or indicated as such.

c. Topotype: a specimen not of the original type series collected at the type locality.

d. Plesiotype: a specimen or specimens on which subsequent descriptions or figures are based.

e. Metatype: a specimen compared with the type by the author of the species and determined by the author as conspecific with it.

f. Homotype: a specimen compared with the type by another than the author of a species and determined by the former to be conspecific with it.

Stabilization of names. Just as types are required as nomenclatural reference points for the application of scientific names to species, so they are also necessary for the stabilization of names for higher taxa. This is especially true at the level of the genus, not only because there are more generic names than those of any other higher category but also because generic names provide roots for the names of several taxa immediately above them, as tribe, family, and so forth. The type for a name of a higher taxon is a lower taxon; for example, the type of a genus is a species. As with species names, the necessity for refining early generic concepts and the retroactive application of the generic type concept to the works of previous authors have resulted in the development of a rather precise series of rules and practices governing their selection. These are specified in detail in the International Rules of Zoological Nomenclature and are concerned with two principal situations: (1) cases in which the generic type is to be accepted automatically based upon the method or form utilized by the author in the original publication in which the new name was proposed (such as type by original designation, type by monotypy, type by tautonomy), and (2) cases in which a subsequent author has a certain amount of freedom in making a selection from among two or more eligible names for type species (type by subsequent designation). As with other aspects of zoological nomenclature, priority of designation, if properly carried out, is binding upon later authors, except as the International Commission of Zoological Nomenclature, in the interest of conserving current usage, may set aside the rules in a given case. The Commission is also called upon to arbitrate difficult cases of generic type interpretation, as when the author misidentifies the species designated as the type of a genus and as a result, the generic

diagnosis applies to a species quite different from the one named in his publication. In such cases, the decision is usually influenced by whether the strict application of the author's concept or of a concept based upon the actual species name will have the most adverse nomenclatural effect. *See* ZOOLOGICAL NOMENCLATURE.

Walter J. Bock

Bibliography. R. E. Blackwelder, *Taxonomy: A Text and Reference Book*, 1967; D. Hegberg, *Systematics*, 1977; International Trust for Zoology Nomenclature, *International Code of Zoological Nomenclature*, 1964; E. Mayr, *Principles of Systematic Zoology*, 2d ed., 1991; G. G. Simpson, *Principles of Animal Taxonomy*, 1990.

Typewriter

A machine that produces printed copy, character by character, as it is operated. Its essential parts are a keyboard, a set of raised characters or a thermal print head, an inked ribbon, a platen for holding paper, and a mechanism for advancing the position at which successive characters are printed. The QWERTY keyboard (named for the sequence of letters of the top row of the alphabet worked by the left hand) was designed in the 1870s. It contains a complete alphabet, along with numbers and the symbols commonly used in various languages and technical disciplines. The manual typewriter was introduced in 1874, followed by the electrically powered typewriter in 1934. By the late 1970s, electronic typewriters offered memory capability, additional automatic functions, and greater convenience. Further advances in electronic technology led to additional capabilities, including plug-in memory and function diskettes and cartridges, visual displays, nonimpact printing, and communications adapters. Although many typewriters are still in use, computers and word processing software largely have supplanted them. *See* WORD PROCESSING.

Edward W. Gore, Jr.; Robert A. Rahenkamp

McGRAW-HILL

ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY

19 U - ZYG

www.MHEST.com

U

Ulcer

A lesion on the surface of the skin or a mucous membrane characterized by a superficial loss of tissue. Ulcers are most common on the skin of the lower extremities and in the gastrointestinal tract, although they may be encountered at almost any site. The diverse causes of ulcers range from circulatory disturbances or bacterial infections to complex, multifactorial disorders. The superficial tissue sloughs, leaving a crater that extends into the underlying soft tissue, which then becomes inflamed and is subject to further injury by the original offender or secondary infection.

Peptic ulcer. Peptic ulcer is the most common ulcer of the gastrointestinal tract and refers to breaks in the mucosa of the stomach or the proximal duodenum that are produced by the action of gastric secretions. Despite a plethora of clinical and experimental studies, it is still unknown why peptic ulcers develop. However, with rare exceptions, a person who does not secrete hydrochloric acid will not develop a peptic ulcer. Ulcers of the stomach tend to develop as a result of superficial inflammation of the stomach. These individuals tend to have normal or decreased amounts of hydrochloric acid. By contrast, most individuals with peptic ulcers of the proximal duodenum secrete excessive amounts of acid. Importantly, a bacterium, *Helicobacter pylori*, has been isolated from the stomach of most people with peptic ulcers and is thought to play a causative role.

The principal complications of peptic ulcers are hemorrhage and perforation, both of which may be fatal. Peptic ulcers do not ordinarily lead to cancer, but cancers of the stomach may become ulcerated and resemble peptic ulcers. Although stress has been anecdotally related to peptic ulcers for at least a century, serious doubt has been cast upon this concept. *See* HELICOBACTER; INFLAMMATION.

Ulcerative colitis. Ulcerative colitis is a disease of the large intestine characterized by chronic diarrhea and rectal bleeding. The disorder is common in the Western world, occurring principally in young adults. Its cause is not known, but there is some evidence for a familial predisposition to the disease. Psychosomatic explanations, which were popular at one time, are no longer thought to be credible. Most individuals have intermittent attacks, with periods of partial or complete remission. Those individuals with long-standing, extensive ulcerative colitis have a higher risk of colon cancer than the general population. Hemorrhage from the injured colon and perforation of the colonic wall are serious complications. Individuals with ulcerative colitis also tend to develop diseases of other organs, including the liver, joints, and skin. *See* CANCER (MEDICINE); HEMORRHAGE.

Other types. Other ulcers of the gastrointestinal tract are caused by infectious agents. Bacterial and viral infections produce ulcers of the oral cavity. Diseases such as typhoid, tuberculosis, and bacillary dysentery and parasitic infestation with ameba lead to ulcers of the small and large intestines.

Narrowing of the arteries to the legs caused by atherosclerosis, particularly in persons with diabetes mellitus, often causes ulcers of the lower extremities. Ulcers overlying severe varicose veins of the legs can also be troublesome. Such ulcers are difficult to heal unless the circulatory problems are brought under control. *See* ARTERIOSCLEROSIS; BACILLARY DYSENTERY; DIABETES; TUBERCULOSIS. Emanuel Rubin

Bibliography. M. A. Peppercorn, *Contemporary Diagnosis and Management of Ulcerative Colitis and Proctitis*, 1995; S. Szabo and C. J. Pfeiffer, *Ulcer Disease: New Aspects of Pathogenesis and Pharmacology*, 1989.

Ultimobranchial bodies

Small, enigmatic structures which originate as terminal outpocketings from each side of the embryonic pharynx. They occur only in vertebrates, where they are almost universal but difficult to homologize.

They probably represent an expression of continued growth activity caudally, associated with pouch- or gill-forming potentialities of foregut entoderm. They are usually bilateral in mammals. The last (and sometimes next to last) "true" pharyngeal pouch and ultimobranchial primordium often unite to form a combined entity known as a caudal pharyngeal complex. During development in humans, the ultimobranchial bodies may be intimately related to the third, as well as the fourth, pharyngeal pouch. Here, as in most mammals, this complex would be the fourth or last.

Location and structure. Originally the bodies were regarded as vestigial lateral, or accessory, thyroids since, because of the mechanics of growth in mammals, they were found to join with lateral lobes of the thyroid. These "bodies" are now interpreted as relatively indifferent tissue, but capable of modification by various factors. When intimately incorporated within a growing and differentiating mammalian thyroid, they were frequently described as indistinguishably transformed (by embryonic induction) into thyroidlike tissue. They appeared to function especially in this way only during periods of thyroid activity such as would occur under the influence of the controlling hormone thyrotropin, but they also have added significance. *See* EMBRYONIC INDUCTION.

In lower vertebrates, ultimobranchial bodies do not join the thyroid but remain isolated in connective tissue between the thyroid gland and heart. Because of their position in fishes and amphibians, these structures were referred to as suprapericardial bodies.

Morphologically, ultimobranchial tissue can be highly variable, even within individuals of a single species. In lower forms, it is often vesicular, and evidence of secretory activity does occur, especially in elasmobranchs, reptiles, and birds. In mammals, it may reflect the secretory features more typical of lower animals. However, within old, atrophic or relatively inactive thyroids, the tissue is often cystic and nonsecretory. It can become mucus-secreting, and this may be reminiscent of its entodermal origin. Among mammals, depending upon the species, it has at various times been thought either to degenerate without trace, to transform by induction into functional thyroid follicles, or occasionally to persist after birth as variable, and multiple, epithelial cysts which sometimes stimulate a leukocytic reaction. Now established is the fact that, at least in sheep and rats, ultimobranchial tissue can occur almost universally as potentially dynamic, cystic metaplasias in postnatal thyroids, which may transform, under certain conditions, into active, centrifugally proliferating, stratified squamous epithelium, resulting in cellular aggregates (neoplasias).

Furthermore, because of the proximity of these ultimobranchial bodies to more cephalic "true" pouches (fifth, fourth, or third in ascending vertebrate phyla), which normally give rise to parathyroid and thymus tissue, and because foregut entoderm may possess reciprocal potentialities, this tissue on occasion may also carry induced, if not intrinsic, attributes for formation of accessory thymus or parathyroid tissue. *See* PARATHYROID GLAND; THYMUS GLAND.

Function and fate. Since in many animals the ultimobranchial bodies have been shown to be inconsistent and to possess few specific, permanent structural attributes, they have been regarded by many as not only variable and erratic in development but also relatively unstable after birth. Despite this unusual character for endocrine tissue, it has been discovered that the ultimobranchial bodies can produce a hormone distinctly different from that of other thyroid hormones. This hormone is a polypeptide and has been labeled calcitonin. Collectively, these cells constitute an endocrine system that plays a role as a delicate and perhaps subtle calcium-regulating hormone. In response to hypercalcemia (high calcium levels of blood), these ultimobranchial cells (or "C" cells) secrete large quantities of this hormone, which rapidly lowers plasma calcium by inhibiting bone resorption, thus preventing excessive osteolysis and calcium deposition in tissues. Although these bodies have been demonstrated in many lower vertebrates, ultimobranchial cells in the shark, chicken, dog, and sheep are especially rich in calcitonin. *See* HORMONE.

However, because of the morphologic instability of this "gland" within the mammalian thyroid (variable cystic changes), including that of humans, a functional instability, or lability, may be implied. It can be influenced by certain intrinsic and extrinsic factors. Not only is it subject to marked structural alterations experimentally, including susceptibility for tumorigenesis, by hormonal and dietary manipulations, such as vitamin A deficiency or carcinogens, but it is seemingly labile, normally, at least in sheep and rats. Possibly no other mammalian endocrine organ exhibits such marked structural variability and therefore, perhaps, subtle systemic dependence. It is vulnerable to change and may reflect changes in blood calcium.

John H. Van Dyke

Ultracentrifuge

A centrifuge of high or low speed which provides convection-free conditions and which is used for quantitative measurement of sedimentation velocity or sedimentation equilibrium or for the separation of solutes in liquid solutions.

The ultracentrifuge is used (1) to measure molecular weights of solutes and to provide data on molecular weight distributions in polydisperse systems; (2) to determine the frictional coefficients, and thereby the sizes and shapes, of solutes; and (3) to characterize and separate macromolecules on the basis of their buoyant densities in density gradients.

The ultracentrifuge is most widely used to study high polymers, particularly proteins, nucleic acids, viruses, and other macromolecules of biological origin. However, it is also used to study solution

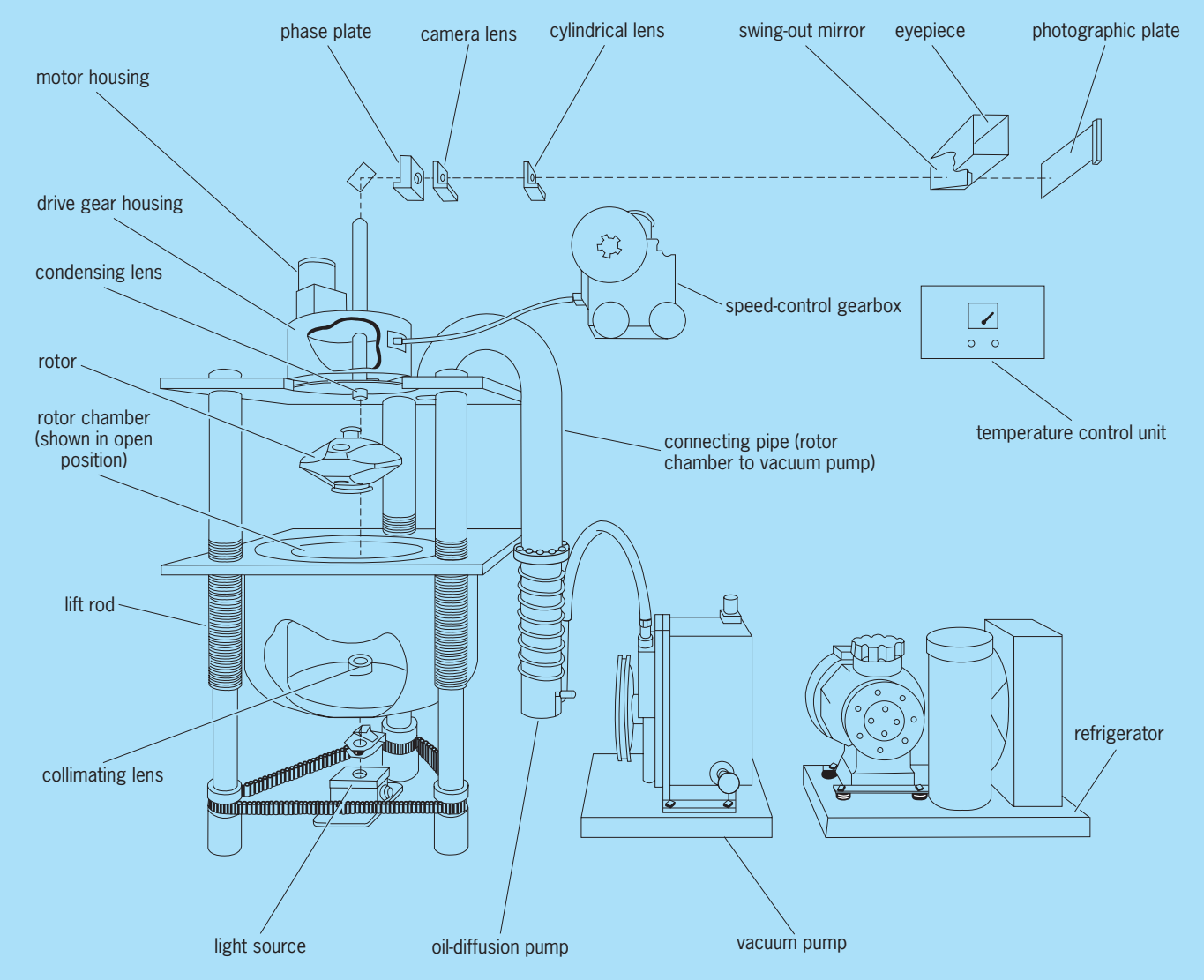

Fig. 1. Diagram of an electrically driven analytical ultracentrifuge. (*Beckman Instruments*)

properties of small solutes. In applications to macromolecules, one distinguishes between the analytical ultracentrifuge, which is used for accurate determination of sedimentation velocity or equilibrium, and the preparative ultracentrifuge, which is used to separate solutes on the basis of their sedimentation velocities or buoyant densities. *See* CENTRIFUGATION.

The application of a centrifugal field to a solution causes a net motion of the solute. If the solution is denser than the solvent, the motion will be away from the axis of rotation. The nonuniform concentration distribution produced in this way leads to an opposing diffusion flux tending to reestablish uniformity. In sedimentation-velocity experiments, sedimentation prevails over diffusion, and the solute sediments with finite velocity toward the bottom of the cell, although the concentration profile may be markedly influenced by diffusion. In sedimentation-equilibrium experiments, centrifugal and diffusive forces balance out, and an equilibrium concentration distribution results which may be analyzed by thermodynamic methods.

Instrument design. Although the first ultracentrifuges were oil-turbine- or air-driven, all modern instruments are electrically driven. The components of an electrically driven analytical ultracentrifuge are shown in **Fig. 1**. The rotor, cell, drive mechanism, and temperature control are designed to achieve high centrifugal accelerations and conditions of convectionless sedimentation. Special rotors have been designed for analytical and preparative purposes. An analytical rotor is shown in **Fig. 2**, and a preparative rotor in **Fig. 3**. Analytical cells, which contain the solution under study, are sector-shaped to allow unimpeded radial flow. The most commonly used analytical ultracentrifuge can be run at speeds between about 2000 and 60,000 revolutions/min. This fastest speed generates centrifugal forces of nearly 260,000 *g*. *See* CENTRIFUGAL FORCE.

Three types of optical systems are used in analytical ultracentrifuges to measure concentration

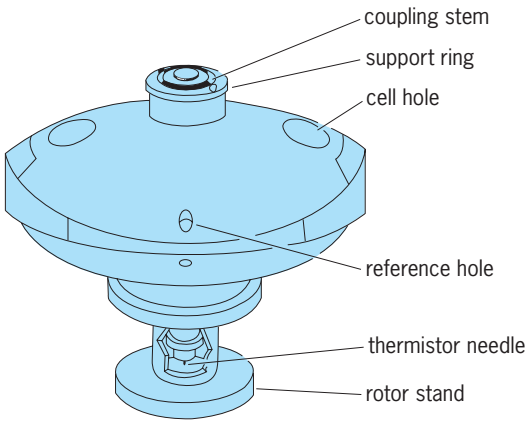

Fig. 2. Analytical rotor. (*Beckman Instruments*)

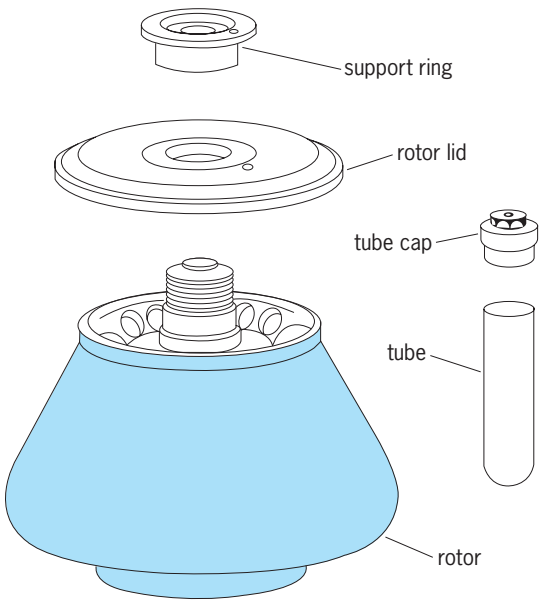

Fig. 3. Fixed-angle preparative rotor and tubes. (*Beckman Instruments*)

distributions while the instrument is in operation. The schlieren system measures refractive index gradients or, equivalently, concentration gradients dc/dr versus distance r from the axis of rotation. The Rayleigh system measures concentration c versus r by creation of interference fringes. The absorption optical system measures c versus r by visible or ultraviolet absorption by chromophoric solutes. Concentration distributions can also be determined in the preparative ultracentrifuge by optical absorption or by radioactivity of labeled solutes after sedimentation has ceased. *See* SCHLIEREN PHOTOGRAPHY.

Molecular weight determination. Four methods of molecular weight determination are available with the ultracentrifuge.

1. *Sedimentation equilibrium*. If equilibrium has been reached between sedimentation and diffusion, the relation between the solute concentration profile and molecular weight M is Eq. (1), where C_1 and

$$M = \frac{2RT \ \ln (C_2/C_1)}{(1-\overline{v}\rho)\omega^2(r_2^2 - r_1^2)} \tag{1}$$

C_2 are concentrations of solute at points r_1 and r_2 in the cell, $\overline{v}$ is the solute partial specific volume, ρ is the solution density, ω is the angular velocity, R is the gas constant, and T is the absolute temperature. For very large macromolecules, equilibrium can be attained only at low speeds, and then may require a day or more of continuous running to be established.

2. *Archibald approach-to-equilibrium method.* The condition for sedimentation equilibrium, that no net flux of solute occurs, is satisfied at the ends of the cell column throughout the approach to equilibrium. Therefore an equilibrium result for molecular weight may be obtained from the concentration distribution at either end of the solution column in much shorter times than required for complete equilibrium to be reached. The appropriate relation is Eq. (2).

$$M = \frac{RTrc}{(1-\overline{v}\rho)\omega^2(dc/dr)} \tag{2}$$

3. *Sedimentation equilibrium in a density gradient.* Binary solutions of low-molecular-weight components of different densities, such as H_2O and CsCl, will exhibit a variation in density with distance from the axis of rotation at sedimentation equilibrium. When a macromolecular species is present in such a density gradient, it will take up a position in the cell where its buoyant density equals the density of the medium. The molecular weight of the macromolecule is related to the density gradient $d\rho/dr$, the standard deviation σ of the width of the band of macromolecules, and the position of the band center r, by Eq. (3).

$$M = \frac{RT}{\overline{v}\omega^2(d\rho/dr)\sigma^2 r} \tag{3}$$

4. *Sedimentation velocity.* This method utilizes centrifugal accelerations which are great enough to produce conveniently measurable sedimentation rates for solutes. A sedimentation coefficient S is obtained which is defined by Eq. (4), where dr/dt is

$$S = \frac{dr/dt}{\omega^2 r} \tag{4}$$

the sedimentation velocity of the solute. If the diffusion coefficient D of the solute is known, the molecular weight is obtained from relationship (5). The

$$M = \frac{RTS}{D(1-\overline{v}\rho)} \tag{5}$$

measured sedimentation coefficient alone may be used to determine the molecular weight of a polymer if an S versus M curve has been calibrated with homologous polymers of known molecular weight. The calibration curve is summarized in the form of Eq. (6), where K and $a$are molecular-weight-

$$S = KM^a \tag{6}$$

independent parameters characteristic of a given polymer-solvent system.

Sedimentation velocity experiments are often run in density gradients, which stabilize the solution against convective disturbances.

If the solute is polydisperse, the molecular weights determined by the methods listed above will represent various averages over the molecular weight distribution. Most common of these are the number-average molecular weight M_n, which is the first moment of the distribution; the weight-average molecular weight M_w, which is the ratio of the second moment to the first; and the z-average molecular weight M_z, which is the ratio of the third moment to the second. By allowing the determination of these averages in a single experiment, sedimentation equilibrium is a valuable tool for characterizing molecular weight distributions in polymers. *See* MOLECULAR WEIGHT.

Frictional coefficient. Ultracentrifugation is used to characterize the sizes and shapes of solutes by determination of their translational frictional coefficient f, which is the proportionality constant between the force acting on a solute molecule and its velocity. The relation between sedimentation coefficient and frictional coefficient is given by Eq. (7), where N is the Avogadro number.

$$S = \frac{M(1 - \overline{v}\rho)}{Nf} \tag{7}$$

For a sphere of radius R, f is given by Stokes' law, Eq. (8), where η is the solvent viscosity. The frictional coefficient has also been calculated for prolate and oblate ellipsoids of revolution, rigid rods, flexible polymer chains, and chains of intermediate stiffness.

$$f = 6\pi\eta R \tag{8}$$

Buoyant density. Sedimentation equilibrium in a density gradient enables a solute to be characterized on the basis of its buoyant density, which is essentially the reciprocal of its solvated partial specific volume. This technique has been especially useful in the study of nucleic acids, where it has enabled distinctions between molecules to be made on the basis of base composition, deoxyribonucleic acid or ribonucleic acid, degree of isotopic substitution, double- or single-strandedness, and other important features. *See* DEOXYRIBONUCLEIC ACID (DNA).

Victor A. Bloomfield

Bibliography. W. Borchard, H. G. Killian, and G. Lagaly, *Progress in Analytical Ultracentrifugation*, 1992; F. Kremer and G. Lagaly (eds.), *Ultracentrifugation*, 1994; P. Sheeler, *Centrifugation in Biology and Medical Science*, 1981, reprint 1990; T. Svedberg and K. O. Pedersen, *The Ultracentrifuge*, 1940.

Ultrafast molecular processes

Various types of physical and molecular changes occurring on time scales of 10^{-14} to 10^{-9} s that are studied in the fields of photophysics, photochemistry, and photobiology. The time scale ranges from near the femtosecond regime (10^{-15} s) on the fast side, embodies the entire picosecond regime (10^{-12} s), and borders the nanosecond regime (10^{-9} s) on the slow side.

A typical experiment is initiated with an ultrashort pulse of energy—radiation (light) or particles (electrons). Rapid changes in the system under study are brought about by the absorption of these ultrashort energy pulses. Excited molecules, for example, may change their structure internally, may gain or lose protons or electrons, may react chemically, may exchange their energy with nearby molecules, or may simply lose memory of the initial dipole direction in which they were excited (dephasing). These changes can be measured by ultrafast detection methods. Such methods are usually based on some type of linear or nonlinear spectroscopic monitoring of the system, or on optical delay lines that employ the speed of light itself to create a yardstick of time (a distance of 3 mm equals a time of 10 ps). These studies are important because elementary motion, such as rotation, vibrational exchange, chemical bond breaking, and charge transfer, takes place on these ultrafast time scales. The relationship between such motions and overall physical, chemical, or biological changes can thus be observed directly. *See* CHEMICAL DYNAMICS; OPTICAL PULSES.

Ultrashort pulses. Just as a very fast shutter speed is necessary to obtain a sharp photograph of a fast-moving object, energy pulses for the study of ultrafast molecular motions must be extremely narrow in time.

Mode locking. The basic technique for producing ultrashort light pulses is mode-locking a laser. One way to do this is to put a nonlinear absorption medium in the laser cavity, which functions somewhat like a shutter. In the time domain, the developing light pulse, bouncing back and forth between the laser cavity reflectors, experiences an intensity-dependent loss each time it passes through the nonlinear absorption medium. High-intensity light penetrates the medium, and its intensity is allowed to build up in the laser cavity; weak-intensity light is blocked. This process shaves off the low-amplitude edges of a pulse, thus shortening it. One of the laser cavity reflectors has less than 100% reflectance at the laser wavelength, so part of the pulse energy is coupled out of the cavity each time the pulse reaches that reflector. In this way, a train of pulses is emitted from the mode-locked laser, with each pulse separated from the next by the round-trip time for a pulse traveling at the speed of light in the laser cavity, about 13 ns for a 2-m (6.6-ft) laser cavity length. *See* LASER; NONLINEAR OPTICS.

Dye lasers. The most convenient lasers for ultrafast molecular experiments are the mode-locked continuous-wave (CW) dye lasers. A jet stream of dye solution shaped like a thin plate is the laser gain medium. Because of the broad spectral characteristics of various dye solutions, these lasers can provide tunable wavelengths over an entire range, from the near-ultraviolet through the visible to the

near-infrared. Dye lasers are mainly of three types. In the first type, the gain jet is synchronously excited (pumped) by a mode-locked argon-ion laser or a frequency-doubled solid-state laser, such as neodymium/yttrium aluminum garnet (Nd:YAG). The pump pulse widths are usually around 70–100 ps, while the dye laser output pulses are 1–10 ps wide.

The second type of dye laser is a hybrid of the first type and a saturable absorber dye jet in the laser cavity, which reduces the laser pulse width. Subpicosecond pulses with discrete tunability can be generated with this type of dye laser.

The third type is the colliding-pulse mode-locking dye laser. This laser generates the shortest pulses, about 40 fs. The configuration is a ring cavity pumped by a continuous-wave ion laser. An absorber jet is placed at a quarter cavity length from the gain jet. Two counterpropagating pulses are adjusted to meet in the pulse-shaving saturable absorber. This introduces a fourfold greater light intensity in the absorber jet, which enhances the effectiveness of the saturable absorber.

Ultrashort pulse compression. To obtain still narrower widths, pulses can be sent through a length of single-mode optical fiber that induces a monotonic frequency sweep (chirp). The red spectral components are on the leading edge of the chirped pulse, and the blue components are on the trailing edge. The pulse is then rephased by an element such as a grating pair or a prism pair, bringing the leading and trailing edges back into coincidence. Laser pulses with widths of 6 fs, or only three or four optical cycles, have been obtained in this way. It is noteworthy that such pulses, because of the Heisenberg uncertainty principle, have a broad spectral distribution (about 3700 cm^{-1}), a property that is useful in ultrafast probe experiments. Also, ultrashort pulses with an arbitrary shape can be achieved by placing amplitude and phase filters between the rephased pair. *See* SPREAD SPECTRUM COMMUNICATION; UNCERTAINTY PRINCIPLE.

Amplification. In general, the ultrashort light pulses produced by continuous-wave mode-locking dye lasers are not intense enough for carrying out most time-resolved experiments. The peak power per pulse is only about 100 mW for 100-fs pulses. Femtosecond amplifiers have been developed that operate with a repetition rate near 100 kHz (high repetition rate) to 10 Hz (low repetition rate). One type of 10-kHz amplifier is a 1.2-mm-thick dye jet, pumped by a copper vapor laser. The pulse width of the copper vapor laser is about 15 ns and is electronically synchronized with the dye laser pulses. The dye laser pulses are set to multipass, usually six times, through the amplifier jet, the energy of an amplified single pulse reaching several microjoules ($1\ \mu J \div 100$ fs = 10 MW). The amplifiers with a lower repetition rate usually consist of a three-stage flow dye cell, which is pumped by a frequency-doubled Nd:YAG regenerative amplifier. The pump pulses are about 100 ps wide with an energy of 50–100 millijoules per pulse. The energy of the final amplified dye laser pulse can reach more than 1 mJ, while preserving the 100-fs pulse width (peak power more than 10 GW).

Supercontinuum. By passing intense ultrashort dye laser pulses into a cell containing water, or almost any kind of transparent solution, a supercontinuum (white light) output is created that has the same pulse width as that of the dye laser. The supercontinuum is produced by a combination of nonlinear optical and molecular effects. The spectral distribution of the supercontinuum can be caused to span a wavelength range from 190 to 1600 nm. An ultrashort light pulse with such broad spectral coverage is very useful as the probe pulse for time-resolved pump/probe spectroscopy.

Ultrashort electron pulses. Electron pulses of short duration can be produced from microwave linear accelerators. Single micropulses of electrons are produced by electronically gating the electron gun, and they are further bunched by synchronously applied microwave fields. Pulses containing 10^{10}–10^{11} electrons with a time width as short as 5 ps can be produced in this way. The approximately 20-MeV energy per electron in the electron pulse is over a million times greater than the energy per photon in a laser pulse. However, the high energy is not necessarily an advantage when studying molecular events, which typically require no more than a few electronvolts for their initiation. The excess energy may cause a number of events to take place simultaneously, thus complicating interpretation. *See* PARTICLE ACCELERATOR.

Detection techniques. In addition to the extremely narrow energy pulses in the time domain, special ultrafast detection techniques are required in order to preserve the time resolution of an experiment.

Streak camera detection. A streak camera operates in some ways like an oscilloscope. A slit image of the light from the event is focused onto photocathode material deposited on a transparent substrate that is housed in an evacuated image tube. The resulting electron packet is accelerated and sent through a linearly falling electric field (voltage ramp) that can be set to decay on nanosecond-to-subpicosecond time scales. The resulting electron packet is thereby deflected onto a moderately persistent phosphor screen. The degree of deflection depends on the instantaneous value of the field—electrons in the packet that arrive early encounter a larger field and are deflected more than electrons arriving later. The resulting phosphor streak therefore contains the time-intensity information of the experiment.

Pump/probe detection. In pump/probe experiments, the time resolution is limited only by the pulse width of the laser. A pump pulse is absorbed, creating a nonequilibrium concentration of intermediates. The decay of intermediates can be detected by transient absorption spectroscopy using a weak probe pulse optically delayed in time after the arrival of the strong pump pulse.

The development of the pump/supercontinuum probe technique has drawn considerable attention. In this type of experiment, an ultrashort pulse

of white light is used as the probe. The major advantages of the supercontinuum are that time-dependent spectral changes can be studied over a broad range of wavelengths and that a more complete view of the transient absorption spectrum can be obtained for analysis.

Time-correlated single-photon counting. The concepts behind this type of detection are based on nuclear counting technology. The probability distribution for emission of a single photon after an excitation event yields the time-intensity distribution of all the emitted photons. In a typical experiment, an ultrashort laser pulse is split in two by a beam-splitting reflector. One pulse impinges on a fast photodiode to generate the start signal for a time-to-amplitude converter. The other pulse, sometimes frequency-doubled, excites the sample. The input into a microchannel plate photomultiplier is adjusted to detect less than a single fluorescence photon per laser pulse. The detected signal is amplified and constitutes a stop signal. The output of the time-to-amplitude converter thus provides a voltage that is proportional to the time delay between the start and stop pulses. A multichannel analyzer is used to build up a histogram of these time delays by using many pulses from a laser with a high repetition rate. The methods of time-correlated single-photon counting are limited by photomultiplier response and detector electronics, providing time resolution no better than about 20 ps. However, the high detection sensitivity is often an advantage. There are still many interesting experiments on the 20–1000-ps time scale both in chemistry and biology that can be carried out with this detection technique.

Correlation methods. Two light signals from the same ultrafast source (autocorrelation) or from two different sources (cross-correlation), when combined in a nonlinear medium such as a crystal of potassium dihydrogen phosphate, produce a two-photon effect. Since no electronic instrumentation is involved, correlation methods can provide time resolution into the femtosecond regime. A widely used autocorrelator design is the interferometer type. If one pulse is delayed with respect to another, the intensity of the two-photon signal will be small, but when the pulses enter the nonlinear medium more nearly simultaneously, the signal increases. An asymmetric correlator has also been developed by the addition of a linear dispersive medium in one of the interferometer arms. In this way, the autocorrelator can be transformed into a powerful diagnostic instrument capable of providing a complete description, in amplitude and shape, of the ultrashort laser pulse.

The simplest asymmetric correlator uses propagation through glass. The pulse is broadened in a known manner by the linear dispersion of the glass. The broadened pulse is then cross-correlated with the original ultrashort pulse. The intensity correlation yields information on the shape of the ultrashort pulse, while its phase can be extracted from interferometric cross-correlation. The reconstruction leads to a complete determination of the complex electric-field amplitude of the pulse broadened by the glass. Taking the Fourier transform of this pulse and multiplying it by the transfer function of the glass leads to the Fourier transform of the original pulse. *See* INTERFEROMETRY.

Light-gating upconversion. A related correlation method that has been widely used is known as fluorescence upconversion. This is the best technique for resolving fluorescence on ultrafast time scales. The laser pulse is split in two. One pulse is then frequency-doubled and used to excite a sample, causing it to generate the fluorescence. The second beam is optically delayed relative to the pump pulse, and is sent into a nonlinear crystal together with the fluorescence to produce a sum-mixing frequency signal. Only that portion of the fluorescence coincident with the gating pulse is detected. The cross-correlation signal of gating pulse and fluorescence is detected, recorded, and analyzed as a function of time delay between pump and gating pulses to obtain time-intensity information about the experiment.

Biological applications. An interesting experimental application of ultrafast spectroscopy is in the resolution of the early ultrafast processes in the complex chain of events responsible for animal vision. Though various species of animals have widely different outward characteristics, the physicochemical processes through which the various organisms respond to light are remarkably similar. The overall response of the eye to light may take as long as a few hundredths of a second, yet it has been found that events on femtosecond and picosecond time scales dominate the early response following the absorption of light in the retina. *See* VISION.

A closely related photobiological process, which has received considerable attention by spectroscopists studying the femtosecond regime, is light-driven transmembrane proton pumping in purple bacteria (*Halobacterium halobium*). Each link in the complex chain of molecular processes, such as an ultrafast photoionization event, is "fingerprinted" by a somewhat different absorption spectrum, making it possible to sort out these events by pump/probe methods of ultrafast laser spectroscopy. The necessity for use of fast primary events in nature concerns overriding unwanted competing chemical and energy loss processes. The faster the wanted process, the less is the likelihood that unwanted, energy-wasting processes will take place. *See* LASER PHOTOBIOLOGY; LASER SPECTROSCOPY; SPECTROSCOPY.

G. Wilse Robinson; Ningyi Luo

Bibliography. T. Kobayashi et al. (eds.), *Ultrafast Phenomena XIV: Proceedings of the 14th International Conference, Niigata, Japan, July 25-30, 2004*, 2005; M. M. Martin and J. T. Hynes (eds.), *Femtochemistry and Femtobiology: Ultrafast Events in Molecular Science*, 2004; R. A. Mathies et al., Direct observation of the femtosecond excited-state cis-trans isomerization in bacteriorhodopsin, *Science*, 240:777–779, 1988; A. Mozumder (ed.), Early events in radiation chemistry (special issue), *Rad. Phys. Chem.*, vol. 31, no. 1, 1989; I. Tanaka, I. N. Molin, and Mustafa El-Sayed (eds.), *Ultrafast Processes in*

Chemistry and Photobiology, Chemistry in the 21st Century Monograph, 1995; A. H. Zewail, *Femtochemistry: Ultrafast Dynamics of the Chemical Bond*, World Scientific Series in the 20th Century Chemistry, vol. 3, 1994.

Ultrafiltration

A filtration process in which particles of colloidal size are retained by a filter medium while solvent plus accompanying low-molecular-weight solutes are allowed to pass through. Ultrafilters are used (1) to separate colloid from suspending medium, (2) to separate particles of one size from particles of another size, and (3) to determine the distribution of particle sizes in colloidal systems by the use of filters of graded pore size.

Ultrafilter membranes have been prepared from various types of gel-forming substances. Unglazed porcelain has been impregnated with gels such as gelatin or silicic acid. Filter paper has been impregnated with glacial acetic acid collodions of varying strengths to produce filters of graded porosity.

Another type of ultrafilter membrane is made of thin plastic sheet with millions of tiny pores evenly distributed over it. Flow rates of liquids and gases through these membranes are very high because the pore volume is 80% of the total membrane volume and the pores proceed through the filter in a direct path. Nominal pore diameters range from 10 nanometers to 5 micrometers. *See* COLLOID; FILTRATION.

Quentin Van Winkle

Ultralight aircraft

A lightweight, single-seat aircraft with low flight speed and power, used for sport or recreation. Ultralights evolved from hang gliders, and the first models were crude, home-built modifications of the Icarus, a rigid-wing, biplane, tailless hang glider popular in the mid-1970s. Newer ultralights are designed strictly as motorized aircraft and cannot be flown as hang gliders. Ultralights are sold as kits, requiring from 50 to 300 h for construction. *See* GLIDER.

Design. There have been tremendous variety of ultralight airframe configurations and control systems (see **illus.**). Airframe types include powered weight-shift hang gliders, flying wings, canard designs, antique biplane replicas, and traditional, monoplane structures with conventional tail designs. Control systems can be either weight-shift systems, two-axis controls, or conventional three-axis controls. In weight-shift designs pilots must shift their weight, using a movable seat, to change the attitude. Two-axis designs use controls to the elevator and rudder only; there are no ailerons. Modern designs use three-axis control systems; weight-shift and two-axis control systems are no longer in production. Three-axis controls resemble those of a standard airplane and include either ailerons or spoilerons (spoiler-type systems used to make turns). Some more sophisticated ultralights are equipped with wing flaps, used to steepen approach profiles and to land at very slow airspeeds. *See* AIRFRAME; FLIGHT CONTROLS.

Powered parachutes are also widely used. In this design, the pilot is suspended beneath a square parachute canopy. The parachute's engine and propellers are situated directly behind the pilot. Counterrotating propellers are employed in order to cancel out any uncommanded turning due to torque effects. Steering is controlled by manipulation of parachute risers, as in a conventional sport parachute. *See* PARACHUTE.

Typically, ultralight airframes are made of aircraft-grade aluminum tubes covered with Dacron sailcloth. Areas of stress concentration are reinforced with double-sleeving or solid aluminum components. Most ultralights are cable-braced and use aircraft-grade stainless steel cables with reinforced terminals. Wingspans average 30 ft (9 m), and glide ratios range from 7:1 to 10:1, depending primarily on gross weight and wing aspect ratio. *See* AIRFOIL; ASPECT RATIO; WING.

Ultralight engines are lightweight, two-stroke power plants with full-power values in the 28–35-hp (21–26-kW) range. They operate on a mixture of gasoline and oil, and most transmit power to the propeller via a reduction or belt drive, a simple transmission that enables the propeller to rotate at a lower, more efficient rate than the engine shaft. *See* PROPELLER (AIRCRAFT); RECIPROCATING AIRCRAFT ENGINE.

Regulation. Prior to September 1982, ultralights were unregulated in the United States. At that time the Federal Aviation Administration (FAA) enacted a regulation that set performance and operational parameters for ultralights. The limitations include a maximum empty weight of 254 lb (115 kg), a maximum level flight speed of 63 mi/h (55 knots or 28 m/s), a stall speed no higher than 27 mi/h (24 knots or 12 m/s), a maximum 5-gallon (19-liter) fuel capacity, and single-seat occupancy. Operational limitations include provisions prohibiting flight at night and over congested areas and controlled airspace. Ultralight pilots must also observe all visual flight rules (VFR). Additionally, ultralights may be used only for sport or recreational purposes and cannot receive a foreign or United States airworthiness certificate.

Since the FAA does not officially recognize ultralights as aircraft but instead treats them as vehicles, it does not require that the pilot have a pilot certificate or the vehicle have an airworthiness certificate. Instead, the FAA has approved a voluntary program of the United States Ultralight Association (USUA) that registers ultralight pilots and vehicles. The USUA also endorses instructors to teach in two-seat, ultralight-type vehicles.

In the interest of safer training, the FAA has authorized the use of heavier, more powerful two-seat, ultralight-type vehicles for use in training prospective ultralight pilots. In order to meet training needs, these vehicles may have empty weights up to 496 lb (225 kg), stall speeds of 40 mi/h (35 knots or 18 m/s), maximum cruise speeds of 80 mi/h (70 knots or

Ultralight aircraft. (*a*) Pterodactyl Ascender, a canard-type aircraft with 35-hp (26-kW) engine (*Freedom Fliers, Inc.*). (*b*) FP-303, with all-wood, latticed construction (*Fisher Flying Products*). (*c*) "Hummer" aircraft, with three-axis controls and V-tail (*Maxair*). (*d*) Quicksilver GT, which uses three-axis controls and 28-hp (21-kW) engine (*Eipper Aircraft*).

36 m/s), and maximum fuel capacities of 10 gallons (38 liters).

If an aircraft does not conform to all criteria established for ultralights, it must be registered as an experimental-category aircraft with the FAA, and the pilot must possess at least a student pilot certificate.

The National Transportation Safety Board investigated a number of fatal ultralight accidents and noted a high rate of in-flight airframe failure and loss of control, compared to accident rates for FAA-certificated lightplanes. Government officials undertook an investigation of the level of regulation of ultralights. Consequently, ultralight flying activity took a dramatic downturn in the mid-1980s. By the 1990s, the combination of safer designs and training in two-seaters brought an increase in flying activity, and a reduction in the numbers of total accidents. The number of fatal accidents has remained at a rate of about 35 per year. *See* AIRPLANE.

Thomas A. Horne

Bibliography. A. Y. Berger and N. Burr, *Berger-Burr's Ultralight and Microlight Aircraft of the World*, 2d ed., 1986; C. Hughs, *The Ultralight Pilot's Flight Training Manual*, 1995; B. Millspaugh, *Ultralight Airman's Manual*, 1987.

Ultrasonics

The science of sound waves having frequencies above the audible range, that is, above about 20,000 Hz. Original workers in this field adopted the term supersonics. However, this name was also used in the study of airflow for velocities faster than the speed of sound. The present convention is to use the term ultrasonics as defined above. Since there is no marked distinction between the propagation and the uses of sound waves above and below 20,000 Hz, the division is artificial. In this article the emphasis is on instrumentation, engineering applications, analytical uses, and medical applications. *See* SOUND.

Ultrasonic Generators and Detectors

The earliest instruments for producing ultrasonic waves in air were the Galton whistle and the Hartmann generator. These devices produce sound waves by blowing a jet of high-pressure air from a narrow slit against a sharp metal edge. The Hartmann generator raises the velocity of the jet above that of the sound waves and in effect generates standing shock waves. *See* SHOCK WAVE.

Piezoelectricity and magnetostriction. Ultrasonic transducers have two functions: transmission and

reception. There may be separate transducers for each function or a single transducer for both functions. The usual types of generators and detectorsv for air, liquids, and solids are piezoelectric and magnetostrictive transducers. Quartz and lithium niobate ($LiNbO_3$) crystals are used to produce longitudinal and transverse waves; thin-film zinc oxide (ZnO) transducers can generate longitudinal waves at frequencies up to 96 GHz. Another class of materials used to generate ultrasonic signals is the piezoelectric ceramics. In contrast to the naturally occurring piezoelectric crystals, these ceramics have a polycrystalline structure. The most commonly produced piezoelectric ceramics are lead zirconate titanate (PZT), barium titanate ($BaTiO_3$), lead titanate ($PbTiO_3$), and lead metaniobate ($PbNb_2O_6$). Composite transducers are transducers in which the radiating or receiving element is a diced piezoelectric plate with filler between the elements. They are called "composite" to account for the two disparate elements, the piezoelectric diced into rods and the compliant adhesive filler. *See* MAGNETOSTRICTION; PIEZOELECTRICITY.

Polymer transducers. Polyvinylidene fluoride (PVDF) is the most popular piezoelectric polymer material for shock sensors and for microphones and other sound pickup devices. It acts efficiently from low frequencies up to the megahertz range. The transducer is formed as a thin film and is polarized by applying a high voltage while the film is mechanically stretched. *See* MICROPHONE.

Pulse systems. These have been used to measure properties of liquids and solids. A short burst of ultrasonic waves from the transducer is sent into the medium and is reflected. By timing the received pulse from the transducer with respect to the transmitted pulse, or by a phasing technique, accurate velocity measurements can be made. Such techniques have been used widely in measuring the elastic constants of small specimens. The attenuation also can be measured by the rate at which pulses decrease with distance transmitted, but careful consideration must be given to spreading loss and to the losses in the seals connecting the transducers to the specimens.

Solid dielectric transducer. This type of transducer is made by stretching a sheet of dielectric film (typically Myla®) that is metallized on one side across a metal backplate. The conducting surface and the metal backplate form a capacitor. The separation between the surface and the plate varies because ultrasonic waves that are incident upon the surface compress gases trapped between the film and the backplate. The change in separation results in a flow of charge. These devices are used as both projectors and receivers in the frequency range 40–1000 kHz. The same transducer is used as both the projector and the detector in a pulse-echo system that is frequently used to measure range for autofocusing cameras. *See* ELECTRET TRANSDUCER.

Optoacoustics. Modulated light absorbed by a fluid or solid results in local heating and generation of an acoustic signal. Pulsed lasers can generate a very short ultrasonic signal with significant energy in the frequency range of 1 MHz to 10 GHz. Such pulses are useful in studying properties of materials and locating defects. *See* PHOTOACOUSTIC SPECTROSCOPY.

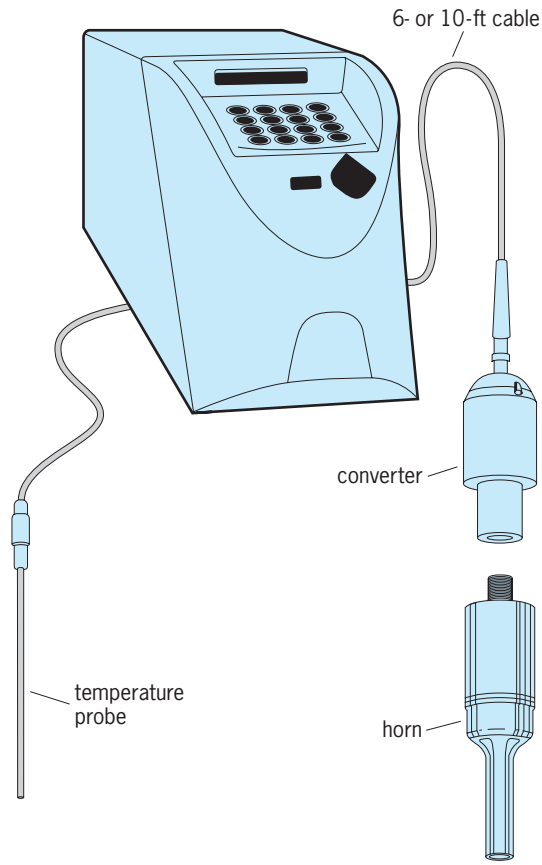

Fig. 1. Sonicator, used to produce high-intensity ultrasound. (*Misonix Inc.*)

High-power devices. High-power ultrasound (typically 600 W) can be obtained with sonicators, consisting of a converter, horn, and tip (**Fig. 1**). The converter transforms electrical energy to mechanical energy at a frequency of 20 kHz. Oscillation of piezoelectric transducers is transmitted and focused by a titanium horn that radiates energy into the liquid being treated. Horn and tip sizes are determined by the volume to be processed and the intensity desired. As the tip diameter increases, intensity or amplitude decreases.

Shear waves in liquids. A number of shear-wave transducers, most of them employing torsional or shear-wave generators of quartz, have been used to measure the shear viscosity and shear stiffness of liquids. These devices measure the acoustic resistance R and the acoustic reactance X of a plane shear wave sent into the liquid. *See* ACOUSTIC IMPEDANCE.

Measurements of this sort have shown that moderately viscous liquids have elastic as well as viscous properties, and they have been widely used in studying the motions possible in polymers and lubricating oils. A typical measurement of a polymer consisting of chlorinated biphenyls (aroclor series), shows that

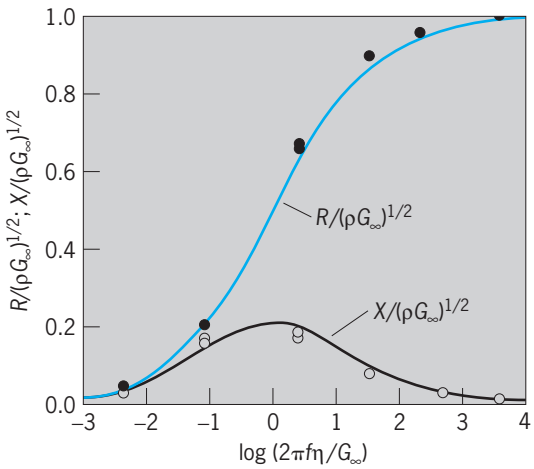

Fig. 2. Graph of $R/(\rho G_\infty)^{1/2}$ and $X/(\rho G_\infty)^{1/2}$ versus the logarithm of $2\pi f\eta/G_\infty$ for a polymer consisting of chlorinated biphenyls (aroclor series), where R is the acoustic resistance, X is the acoustic reactance, ρ is the density, G_∞ is the stiffness at infinite frequency, f is the frequency, and η is the viscosity of the liquid. Curves give theoretical values, and data points give measured values.

at low frequencies R and X are nearly equal, consistent with viscous-wave propagation (**Fig. 2**). For very high frequencies the reactance term tends to zero while the resistance term approaches a value given by Eq. (1), where ρ is the density and G_∞ the stiffness

$$R = (\rho G_\infty)^{1/2} \tag{1}$$

of the liquid at an infinite frequency. The frequency at which the reactance is a maximum (relaxation frequency) is given by Eq. (2), where η is the viscosity

$$f = \frac{G_\infty}{2\pi\eta} \tag{2}$$

of the liquid. For example, water with a stiffness of a porous solid of 10^8 newtons/m^2 (10^9 dynes/cm^2) and a viscosity of 10^{-3} pascal·s (0.01 poise) would have a relaxation frequency of 10^{10} Hz, which is above any present measurements. The viscoelastic properties of lubricating oils have been shown to contribute to the load-carrying capacity of spur gears operating at high speeds. *See* RHEOLOGY; VISCOSITY.

Engineering Applications

The engineering applications of ultrasonics can be divided into those dealing with low-amplitude sound waves and those dealing with high-amplitude (usually called macrosonics) waves.

Low-amplitude applications. Low-amplitude applications are in sonar (an underwater-detection apparatus), in the measurement of the elastic constants of gases, liquids, and solids by a determination of the velocity of propagation of sound waves, in the measurement of acoustic emission, and in a number of ultrasonic devices such as delay lines, mechanical filters, inspectoscopes, thickness gages, and surface-acoustic-wave devices. All these applications depend on the modifications that boundaries and imperfections in the materials cause in wave propagation properties. The attenuation and scattering of the sound in the media are important factors in determining the frequencies used and the sizes of the pieces that can be utilized or investigated. *See* SONAR.

Ultrasonic inspectoscopes. These transmit sound waves into a metal casting or other solid piece and determine the presence of flaws by reflections or by an interruption of the sound-wave transmission through the piece. Frequencies ranging from 500 kHz to 15 MHz are used. Such devices are among the best means for determining defects in metals, glasses, and ceramics, and they have also been applied in the inspection of automobile tires. *See* NONDESTRUCTIVE EVALUATION.

Ultrasonic thickness gages. These have been used in measuring the thickness of pieces when one side is not accessible, such as in boilers. Pulsing systems and continuous-frequency systems based on the resonance principle are used.

Surface-acoustic-wave devices. One application of ultrasonics involves the use of Rayleigh surface waves which were first studied in the case of earthquake waves on the surface of the Earth. In this mode the motion is along the surface with a small penetration into the interior. Such devices may be used in obtaining high-frequency wave filters (**Fig. 3**). Surface waves are generated by equally spaced interdigital electrode transducers. These generate a band of frequencies whose midfrequency is determined when the half Rayleigh wavelength equals the spacing of the transducer electrodes. Since the transducer sends out waves in both directions, the wave toward the back is absorbed by a resistive termination.

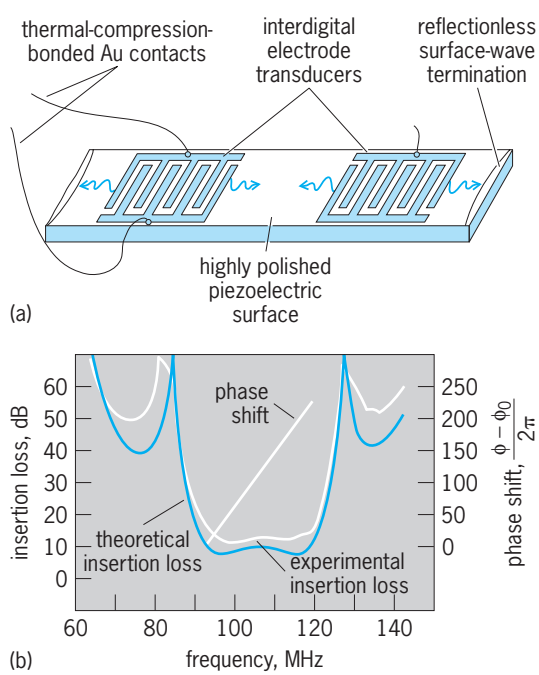

Fig. 3. Surface-wave transducers used as a bandpass filter. (*a*) Configuration. (*b*) Characteristics. Insertion loss is a form of attenuation. Phase shift is nearly linear. (*After E. A. Kraut, ed., Acoustic Surface Wave and Acoustooptic Devices, Optosonic Press, 1971*)

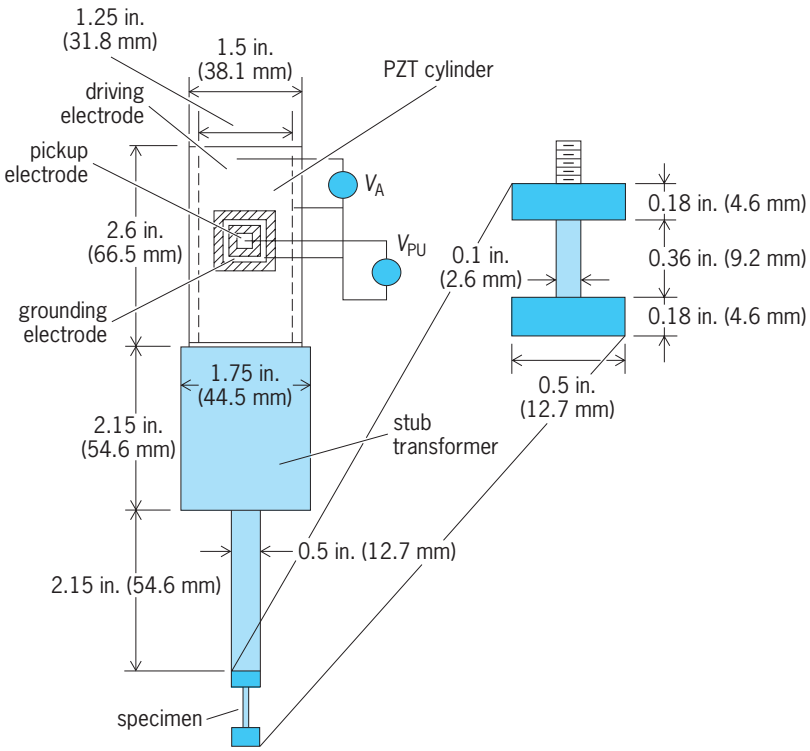

Fig. 4. Equipment for fatigue testing at ultrasonic frequencies, consisting of ultrasonic generator and specimen. V_A = applied voltage; V_{PU} = pickup voltage.

Most of the surface-wave devices use lithium niobate ($LiNbO_3$), which has a high electromechanical coupling and a low acoustic attenuation. For low-temperature coefficients, lithium tantalate ($LiTaO_3$) and S-cut quartz are sometimes used.

Surface-acoustic-wave devices are used for oscillators, nondestructive evaluation, and various processing devices. *See* SURFACE-ACOUSTIC-WAVE DEVICES.

Acoustic emission. An effect related to internal friction, motion of dislocations, and fatigue in materials is the noise in the specimens produced by strain. This is called acoustic emission. The first research on this phenomenon was done for rocks. This work started in the 1930s and was undertaken to test mine areas that were near danger regions for slides. Since the attenuation of sound waves in rocks is quite large at high frequencies, most of the measured frequencies were in the 150–10,000-Hz range. Later work has attempted to relate acoustic emission to earthquake properties. Acoustic emission and microfracturing in rocks may be directly related to the inelastic part of the stress-strain behavior, that is, to the internal friction. This suggests that the acoustic emission in rocks is connected with dislocation motion as has been established for metals.

High-amplitude applications. High-amplitude acoustic waves (macrosonic) have been used in a variety of applications involving gases, liquids, and solids. Some common applications are mentioned below.

Effects due to cavitation. A liquid subjected to high-amplitude acoustic waves can rupture, resulting in the formation of gas- and vapor-filled bubbles. When such a cavity collapses, extremely high pressures and temperatures are produced. The process, called cavitation, is the origin of a number of mechanical, chemical, and biological effects.

Cavitation plays an integral role in a wide range of processes such as ultrasonic cleaning and machining, catalysis of chemical reactions, disruption of cells, erosion of solids, degassing of liquids, emulsification of immiscible liquids, and dispersion of solids in liquids. Cavitation can also result in weak emission of light, called sonoluminescence, as discussed below. *See* CAVITATION; SONOCHEMISTRY.

Fatigue testing. Fatigue testing at ultrasonic frequencies has had considerable application because of the rapidity with which it can be accomplished. A typical device for testing (**Fig. 4**) consists of a PZT cylinder, vibrating at a frequency of the order of 20 kHz, attached to a mechanical transformer in the form of two quarter-wave stubs. The device is made from an alloy of titanium which has a low internal friction up to higher strains than any other materials. The stub transformer gives a transformation ratio proportional to the square of the ratio of the large diameter to the small diameter, in this case 12.25. The specimen, which consists of a thin shaft with two end weights, tuned to the same frequency as the transducer and the transformer, produces another factor of 5. Altogether, a linear strain of 6×10^{-3} can be produced if the material can withstand it.

By using such devices, it has been shown that the results obtained at high frequencies are closely similar to the results obtained at low frequencies, cycle for cycle. Constant total strain and constant displacement tests are possible. Positive mean load, elevated temperature, and corrosion ultrasonic fatigue tests can be performed routinely. The feasibility and advantages of performing crack propagation tests at ultrasonic frequencies have been demonstrated.

Other high-amplitude effects. Ultrasound is used widely in the cleaning of metal parts, such as in watches. The large acoustic forces actually break off particles and contaminants from metal surfaces. Ultrasound has been investigated for washing textiles.

One of the principal applications of ultrasonics to gases is particle agglomeration. This depends upon the fact that light particles can follow the rapid motion of the sound waves, whereas heavy ones cannot. Hence, light particles will strike and stick to heavy ones, reducing the number of small particles in the gas. The heavy particles eventually will fall to a collecting plate or can be drawn there by means of an electric field. This technique has been used in industry to collect fumes, dust, sulfuric acid mist, carbon black, and other substances.

Another industrial use of ultrasonics has been to produce alloys, such as lead-aluminum and lead-tin-zinc, that could not be produced by conventional metallurgical techniques. Shaking by ultrasonic means causes lead, tin, and zinc to mix.

Analytical Uses

In addition to their engineering applications, high-frequency sound waves have been used to determine the specific types of motions that can occur

in gaseous, liquid, and solid mediums. Both the velocity and attenuation of a sound wave are functions of the sound frequency. By studying the changes in these properties with changes of frequency, temperature, and pressure, indications of the motions taking place can be obtained.

Sound attenuation in fluids. In monatomic gases and monatomic liquids such as mercury, the sound attenuation can be explained as absorption due to viscosity and heat conduction. For such fluids, the attenuation A satisfies Eq. (3), where f is the frequency,

$$A = \frac{2\pi^2 f^2}{\rho v^3}\left[\frac{4}{3}\eta + \frac{(\nu - 1)K}{C_p}\right] \tag{3}$$

ρ the density, v the sound velocity, η the coefficent of viscosity, ν the ratio of specific heats, K the thermal conductivity, C_p the specific heat at constant pressure, and A the attenuation in nepers per meter, if the other quantities are in SI units.

Polyatomic liquids show additional attenuation due to relaxations of two types. Thermal relaxations, which have been demonstrated for gases and nonassociated liquids, that is, liquids that contain nonpolar molecules, occur by an interchange of energy between the longitudinal sound wave and the rotational and internal modes of motion of the gas or liquid molecules.

Structural relaxations occur for associated liquids, for polymer liquids, and also for solids. These relaxations take place when one part of the molecule moves from one position to another under the combined effect of the thermal- and sound-wave energy. A definite structure, such as that which occurs in associated liquids and polymer liquids, is required. *See* SOUND ABSORPTION.

Effects in solids. For solids, a variety of effects cause attenuation and velocity dispersion. Probably the simplest of these are thermal effects.

Thermal effects. When a solid body is compressed by an acoustic wave, the compressed part becomes hotter and the expanded part cooler. Thermal energy is transmitted from the hot part to the cool part. Since this energy comes from the acoustic wave, a loss or attenuation of the wave results which is proportional to the square of its frequency. For bars in flexural vibration, the thermal path is quite short, and the effect produced is large. Below a frequency f_0, determined by Eq. (4*a*), such a source produces an internal friction $1/Q$ given by Eq. (4*b*). Here K is the thermal conductivity, C_p the specific heat at constant pressure, f the frequency of the sound wave, ρ the density, W the width of the bar in centimeters, and Y_0^σ and Y_0^θ the adiabatic and isothermal values of Young's modulus, respectively. The velocity increases as a function of frequency (**Fig. 5**), while a corresponding internal friction occurs. If velocity and sound attenuation are graphed against frequency (Fig. 5), the shape of the resulting curves is the same in any medium with a single relaxation, although the horizontal and vertical scales vary from one medium to another.

$$f_0 = \frac{\pi K}{2C_p \rho W^2} \tag{4a}$$

$$\frac{1}{Q} = \frac{Y_0^\sigma - Y_0^\theta}{Y_0^\theta}\left(\frac{ff_0}{f^2 + f_0^2}\right) \tag{4b}$$

The thermal path l for a longitudinal wave becomes smaller as the frequency of vibration increases. It is given by Eq. (5) where v is the velocity

$$l = \frac{v}{2f} \tag{5}$$

of propagation. Inserting this expression in Eq. (4*a*) with $l = W$, the relaxation frequency for a longitudinal wave is then given by Eq. (6). Above this frequency, the material is isothermal, whereas below f_0 it is adiabatic. This frequency is above 10^{10} Hz for most materials. The attenuation for a longitudinal wave for this thermoelastic effect is given by Eq. (7), where γ is the Grüneisen constant, B the bulk modulus, α the thermal expansion coefficient, C the specific heat per unit volume, ω is $2\pi f$, ρ the density of the medium, V_l the longitudinal velocity, T the absolute temperature in kelvins, and A the attenuation in nepers per meter, if the other quantities are in SI units. This source of attenuation is quite large for metals but provides only about 4% of the thermal attenuation for insulators.

$$f_0 = \frac{C_p \rho v^2}{2\pi K} \tag{6}$$

$$A = \frac{\omega^2}{2\rho V_l^3}\left(\frac{\gamma^2 KT}{V_l^2}\right) \qquad \gamma = \frac{3B\alpha}{C} \tag{7}$$

The main thermal attenuation is provided by the Akheiser effect. This loss is determined by the thermal conductivity and the nonlinear third-order elastic moduli.

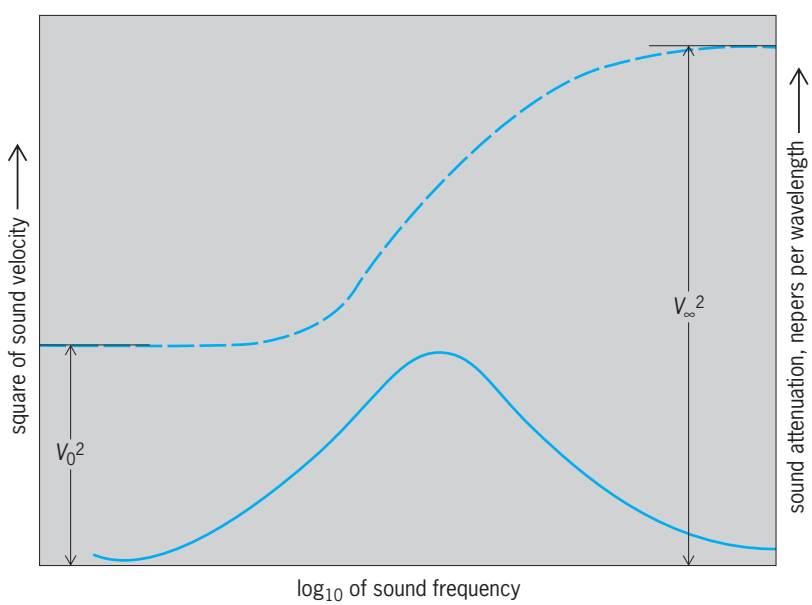

Fig. 5. Velocity dispersion (broken line) and corresponding attenuation per wavelength peak (solid line) for a medium with a single relaxation. The velocity increases as a function of frequency. One neper equals 8.7 dB.

Other relaxations. A number of relaxation phenomena are associated with the motion of impurity atoms, grain boundaries, domain boundaries, and other motions occurring in a solid. Interstitial atoms, such as nitrogen and carbon in iron, can cause an appreciable acoustic loss. These impurity atoms have preferred positions between the iron atoms in the crystal lattice. When a sound wave stretches the lattice in one direction and compresses it in a direction perpendicular to the first, the interstitial atoms, actuated by thermal energy, tend to go to the most open regions. When a compression due to the sound wave occurs, the reverse motion takes place. Since it requires a thermal activation energy H to move the impeding atoms aside, the frequency of jumping f follows Eq. (8), where f_0 is the frequency of vibration

$$f = f_0 e^{-H/(RT)} \tag{8}$$

of a nitrogen atom due to thermal motion ($\cong 10^{13}$ Hz), R the energy necessary to increase the temperature of 1 mole of atoms (6.022×10^{23} atoms) by 1°C, and T the temperature in kelvins. Since H is about 16,400 cal/mole (68,600 joules/mole) for nitrogen and R is 2 cal (8.31 joules), the relaxation frequency for this process is about 1 Hz at room temperature.

Other relaxations involving substitutional atoms have been observed at higher temperatures, since the substitutional atoms in this case have higher activation energies. Relaxations involving the rotation of grains in polycrystalline samples have been observed at high temperatures and low frequencies.

Much faster relaxations occur in magnetic processes involving the motion of domain walls in magnetic materials. A demagnetized specimen is made up of a number of domains within which the direction of magnetism is the same. Domains with directions of magnetism at right angles to or at 180° from the original direction are separated by regions called Bloch walls, in which the direction of magnetism changes from one domain to the other by small steps in the orientation of magnetism. A compressive stress in the same direction as the magnetic flux—for a positive magnetostrictive material—causes the domain to shrink, whereas it causes domains directed at 90° to expand. Hence, the domain wall moves as the stress changes from compressive to extensional. For a discussion of Bloch walls. *See* FERROMAGNETISM.

The domain walls can be held up by dislocations and other imperfections in the magnetic material, and a definite magnetic field or stress is required before the domain wall moves at all. On the reverse cycle the domain wall lags behind the applied stress (magnetic or elastic). The effect produces a hysteresis loop in the material and an acoustic loss called the microhysteresis effect. As the direction of magnetism changes, eddy currents are generated. These limit the velocity with which a domain wall can move and produce an acoustic loss for alternating stresses called the microeddy current effect. For a given size domain, there is some frequency for which the velocity is only half as large as that for low frequencies for the same applied magnetic field. The loss at this frequency is a maximum, and hence this frequency is a relaxation frequency. It can be shown that this frequency is determined by Eq. (9), where R is the

$$f_0 = \frac{R}{96\chi_0 l^2} \tag{9}$$

electrical resistivity of the material, χ_0 the initial magnetic susceptibility for a demagnetized material, and l the thickness of a domain. For nickel, for example, this frequency is of the order of 10^5 Hz. *See* DOMAIN (ELECTRICITY AND MAGNETISM); EDDY CURRENT; HYSTERESIS.

Many other relaxations occur, depending on the nature of the solid-state motion that can take place in the material. Ultrasonic measurements carried out over wide frequency and temperature ranges are powerful tools for investigating such motions.

Low-temperature data. Ultrasonic waves have provided significant information on processes that occur at temperatures near absolute zero. In liquids the most important results have been obtained for liquid helium, while for solids results have been obtained with metals at low temperatures which reveal a considerable amount of information about the mechanism of superconductivity.

Liquid helium. When helium is liquefied at its boiling point (4.2 K or −452°F) and cooled further, at the so-called lambda (λ) point (2.2 K or −455.7°F) there is a transformation of normal helium into superfluid helium. At the λ point there is an ambiguity in the sound velocity and also a high attenuation. Superfluid helium has a zero viscosity and a high thermal conductivity. The former leads to a small acoustic attenuation for normal sound, while the latter leads to the capability of transmitting thermal waves, the so-called second sound. Second sound can be initiated and detected by thermal means, and it has been found that the velocity is zero at 2.2 K (−455.7°F), rises to a maximum of 65 ft/s (20 m/s) at 1.7 K (−456.6°F), and decreases thereafter at lower temperatures. The velocity of normal sound varies from 750 ft/s (230 m/s) near absolute zero to 600 ft/s (180 m/s) near 2.2 K (−455.7°F). *See* LIQUID HELIUM; SECOND SOUND.

Attenuation at low temperatures. At very low temperatures the ultrasonic attenuation of pure normally conducting metals becomes high. For example, above 10 K (−442°F) the ultrasonic attenuation of pure tin (**Fig. 6**) is still relatively small and increases as the square of the frequency; whereas at 4 K (−452°F), at which temperature tin is still in the normal state, the attenuation is high and increases in proportion to the frequency. It has been shown that the added attenuation in the normal state is due to the transfer of momentum and energy from the acoustic wave to the free electrons in the metal. If the acoustic wavelength is greater than the electronic mean free path, this transfer determines an effective viscosity, and the attenuation increases in proportion to the square of the frequency. When the mean free path becomes longer than the acoustic wavelength, as it does at low temperatures, the energy communicated to the

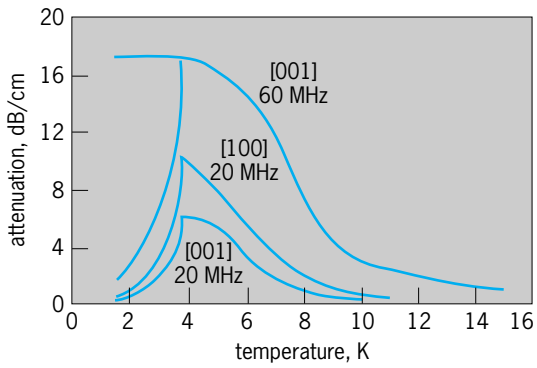

Fig. 6. Longitudinal sound-wave attenuation measurements for a single crystal of tin along the (001) axis and along the (100) axis. °F = (K × 1.8) − 459.67; 1 dB/cm = 2.5 dB/in.

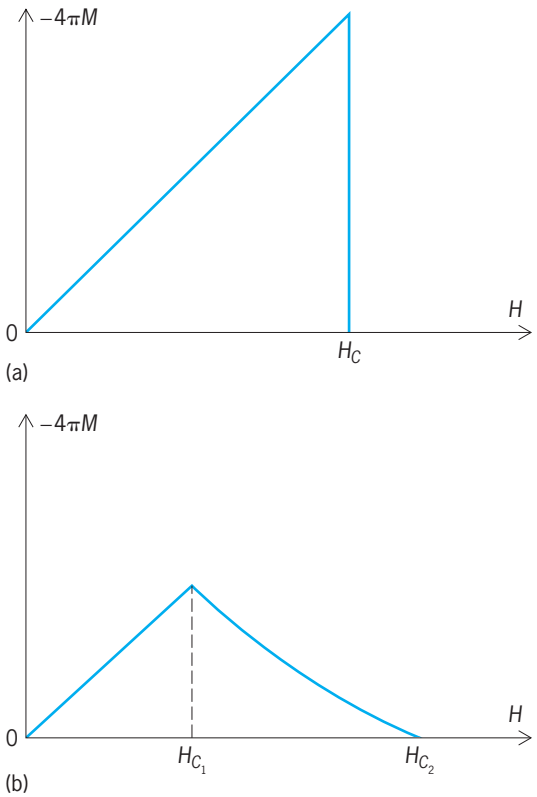

Fig. 7. Magnetization curves of long cylinders of type I and type II superconductors. The applied field *H* is directed along the axis of the cylinder.

electrons is not returned to the acoustic wave and a high attenuation results. The attenuation is proportional to the number of times the crystal lattice vibrates and hence to the frequency. *See* FREE-ELECTRON THEORY OF METALS.

As the temperature drops below the temperature at which tin becomes superconductive (3.71 K or −453.0°F), this source of attenuation drops rapidly to zero. The form of the temperature dependence (Fig. 6) has been used to confirm the Bardeen-Cooper-Schrieffer energy-gap theory of superconductivity. However, at lower frequencies, that is, from 10 to 100 MHz, losses due to dislocations can occur. These are different for the normal and superconducting states, and this difference has to be taken account of in order to determine the form of the energy-gap relation. For frequencies above 100 MHz, the attenuation due to dislocations is small compared to the electron-phonon loss, and direct measurements give the shape of the energy-gap curves. *See* SUPERCONDUCTIVITY.

Acoustic measurements are also useful for type II or high-field superconductors (HFS). For these types of superconductors, which are used for superconducting magnets, there are two critical fields, rather than the single field of type I superconductors (**Fig. 7**). In type I the magnetic flux is completely excluded from the interior of the material below H_c. For type II superconductors the magnetic flux is completely excluded from the interior only below H_{c1}. Between H_{c1} and H_{c2} the magnetic flux consists of flux vortices in the form of filaments directed along H, embedded in a superconducting material. When a dc electric current flows in a direction normal to H, each vortex experiences a force normal to its length, which causes it to move. The vortices are pinned by defects, and a finite current density is required before the vortices move. An alternating current or alternating stress causes motions of the pinned vortices which lag behind the applied forces. The result is an acoustic attenuation (**Fig. 8**). A sharp dip in the attenuation occurs near the superconducting field H_{c2}. Above H_{c2} the material is in the normal state and the attenuation rises rapidly with the field. *See* LOW-TEMPERATURE ACOUSTICS.

Resonant ultrasound spectroscopy. Resonant ultrasound spectroscopy (RUS) is an experimental technique for obtaining a complete set of elastic constants. It is based on the measurement of the free mechanical resonances of a sample of well-defined

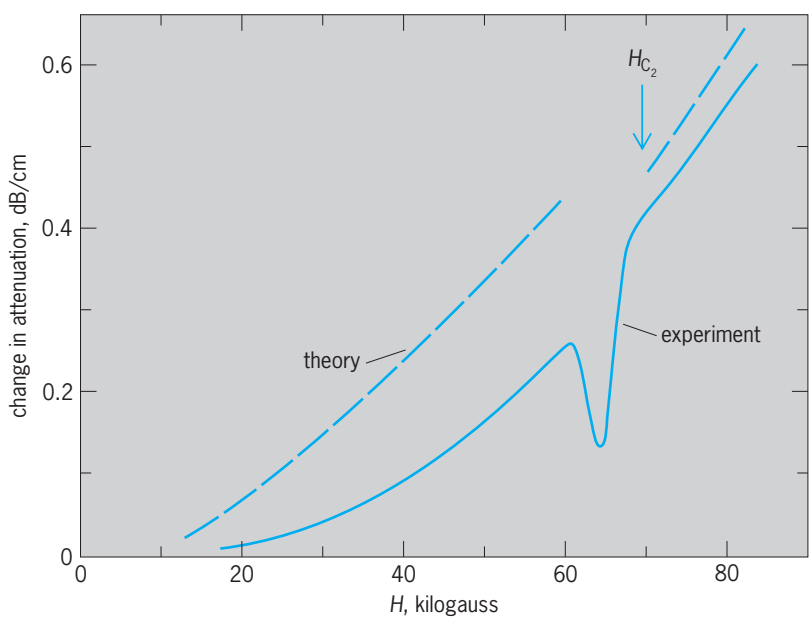

Fig. 8. Magnetic field variation of the attenuation of 9.1-MHz shear waves parallel to *H* in annealed Nb–25% Zr with temperature at 4.2 K (−452°F). 1 kilogauss = 0.1 tesla; 1 dB/cm = 2.5 dB/in.

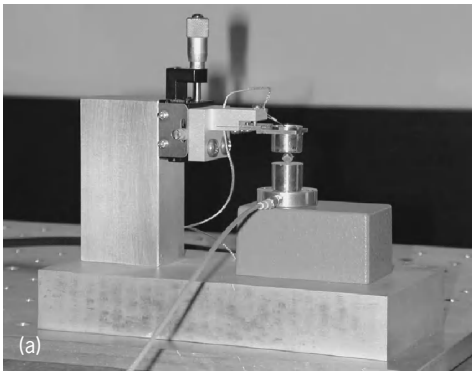

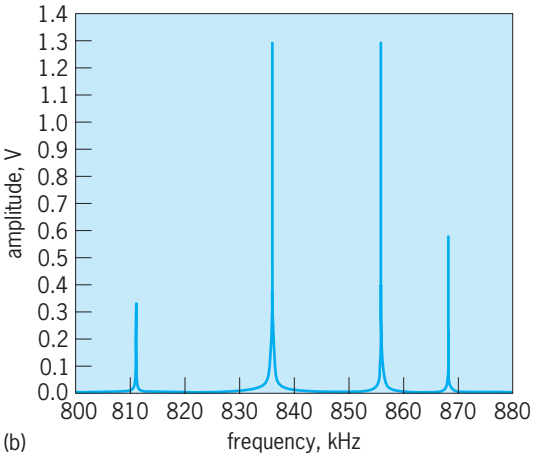

Fig. 9. Resonant ultrasound spectroscopy (RUS). (*a*) Transducers with mounted sample. (*Dynamic Resonance Systems Inc.*) (*b*) Part of a typical response of a scan.

shape, usually a sphere or rectangular parallelepiped. Given a sample's density and elastic constants, a spectrum of resonant frequencies can be calculated for free boundary conditions. It is the job of RUS to measure these resonant frequencies and fit the calculated resonances to the experimentally measured ones by adjusting the starting values for the elastic tensor.

In an RUS experiment, a specimen, typically 1–3 mm (0.04–0.12 in.) on a side, is held between two transducers with a minimal amount of pressure (**Fig. 9***a*). One transducer is used to excite the sample while the other measures its response. A few of the sharp resonances that are typically observed are shown in Fig. 9*b*. RUS does not require bonding of different transducers on the sample, as do conventional techniques, and all moduli are determined simultaneously. RUS is also sensitive to the symmetry of the object, so that certain symmetry-breaking effects (such as asphericity) are easily detected.

Sonoluminescence. Sonoluminescence is a phenomenon in which pulses of light are emitted by collapsing cavitation bubbles. There are two classes of sonoluminescence: multiple-bubble and single-bubble. Multiple-bubble sonoluminescence is associated with cavitation processes that involve many bubbles in relatively gassy liquids. Multiple-bubble cavitation is random in nature. As a result, the location and timing of multiple-bubble sonoluminescence events are largely unpredictable. Such conditions often arise in sonochemistry and industrial ultrasonic applications.

In single-bubble sonoluminescence, sonoluminescent pulses are emitted from a single, isolated bubble, trapped at the pressure antinode of a high-amplitude acoustic standing wave set up in a fluid-filled acoustic resonator. The host fluid is degassed. The bubble is driven into large-amplitude, nonlinear, radial oscillations by the levitation field. One sonoluminescent pulse is emitted in each acoustic cycle. Typical frequencies are in the range of 20–60 kHz. Single-bubble sonoluminescence can be generated with very simple apparatus and is easily visible to the unaided eye. The duration of the sonoluminescent pulse is estimated to be less than 50 picoseconds. There has been no direct measurement of the pulse duration. The emission can be extremely stable and repetitive, lasting for hours. It can also be quasiperiodic and chaotic, depending on the driving conditions. The spectrum is broad-band and increases in intensity in the ultraviolet. Slight cooling of the host liquid significantly increases the intensity. The intensity of the luminescence is sensitive to changes in the composition of gases within the bubble of less than 1%. Individual single-bubble sonoluminescence pulses are isotropic and unpolarized. Single-bubble sonoluminescence has been observed only in water and water solutions.

There is no completely satisfactory explanation for this phenomenon. The most plausible explanation is that during the highly spherically symmetric collapse of the isolated bubble, a shock wave is launched toward the interior of the bubble. When the shock wave reaches the center of the bubble, the temperature and pressure of the gas behind the shock front are both very high, leading to conditions under which light is generated. The short duration of the emission is accounted for by the fact that the conditions necessary for emission are quickly quenched as the shock wave diverges after focusing at the bubble's center.

Medical Applications

Application of ultrasonics in medicine can be generally classified as diagnostic and therapeutic. The more common of these at present is the diagnostic use of ultrasound, specifically ultrasonic imaging.

Ultrasonic imaging. The use of medical ultrasound procedures and equipment surpasses that of any other medical imaging modality. Three common imaging modes have evolved using ultrasound technology in medical imaging.

B-mode. Ultrasonic pulses are partially reflected from interfaces where a change occurs in the specific acoustic impedance. For pulse-echo imaging, an ultrasonic transducer, operating by the piezoelectric effect, emits an acoustic pulse and converts the returning echoes to voltage pulses whose amplitudes are proportional to the echo strength. The amplitude of voltage pulses can be used to modulate the intensity of the oscilloscope trace to create a B-mode

scan line. A two-dimensional B-mode image can be produced by combining a number of B-mode scan lines from different directions or positions into an image. The spatial relationship between the scan lines is preserved by recording the direction of the ultrasound beam and the range to targets, based on an assumed constant sound speed. B-mode images with more than 100 scan lines are the basis for many imaging systems used today, and video displays of digital data have replaced the simple oscilloscope traces used in early devices. The display can be updated at over 100 images per second.

Arrays of elements are used to focus and steer ultrasonic beams rather than fixed lenses and mechanical motion of an ultrasound transducer. Array systems include both linear arrays, where groups of elements are combined in a sequential fashion to scan an ultrasonic beam in a plane, and phased arrays where phasing of elements is used to steer the beam in the imaging plane. Phasing of elements is also used to focus the beam at multiple locations during repeated transmits or to continuously focus receive signals by phase coherent summation. Electronic focusing has been predominantly one-dimensional, allowing focusing within the imaging plane and orthogonal to the propagation direction. However, some systems have electronic control of the out-of-plane beam characteristics.

M-mode. Motion-mode or M-mode images are produced by combining B-mode scan lines based on their temporal relationship instead of the spatial relationship used in B-mode images. A sequence of B-mode scan lines is acquired with the same beam position and orientation, and is displayed vertically on a video screen with the most recent scan line on the right-hand side. As the next acoustic pulse is recorded, previous measurements are shifted to the left on the screen. The effect is a scrolling record of the tissue position as a function of time. Strong scatterers, such as organ boundaries, will be seen as bright spots on individual B-mode scan lines and horizontal streaks on the M-mode image if the boundary is stationary. Motion of these boundaries toward the transducer causes the streak to move upward on the display, and movement away causes the streak to move downward. This allows for easy measurement of movement. The M-mode has been used extensively to monitor heart-wall motion.

Doppler. The frequency shift caused by the reflection of sound from moving objects (acoustic Doppler effect) is used to image fluid flow in many application areas, including medicine. Blood flow in vessels can

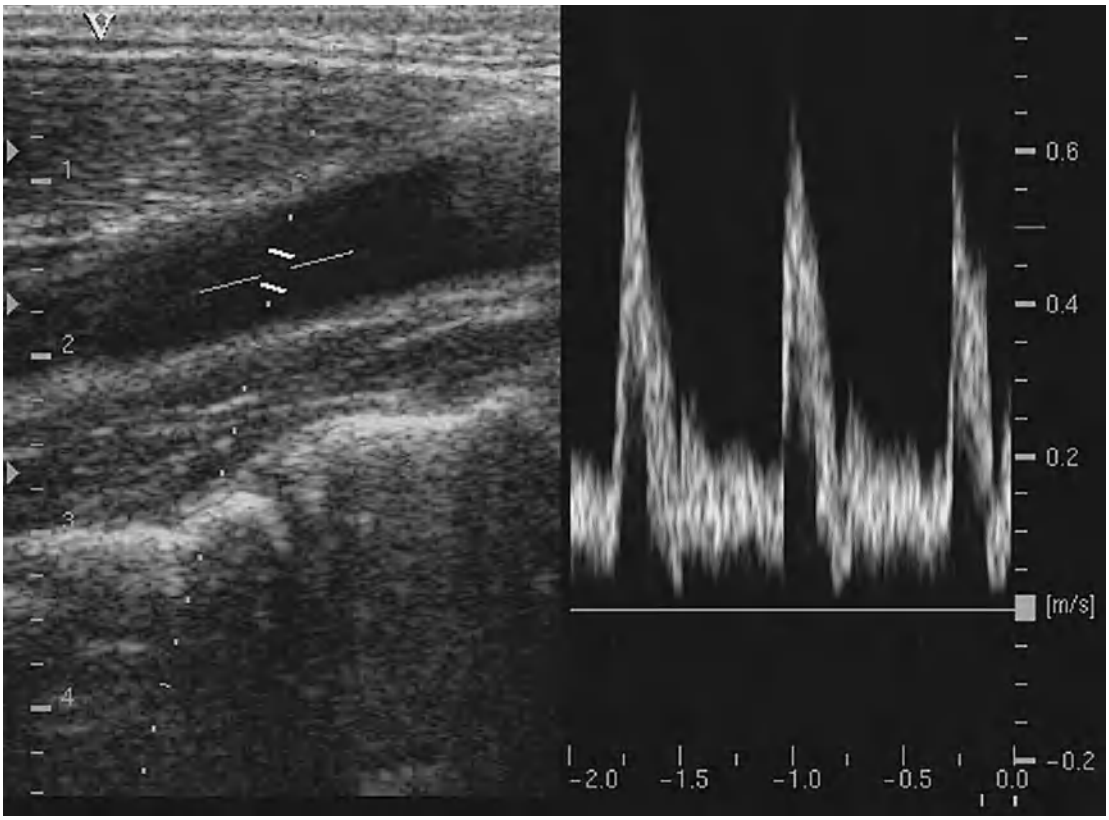

(a) (b)

Fig. 10. Duplex display from a diagnostic ultrasound scanner. (*a*) B-mode image of the human carotid artery. The depth in centimeters is shown along the left side. The three triangles at left mark the position of the electronic focal zones for transmit. The graphic in the center of the image indicates the location where the Doppler measurement is being made. (*b*) Spectral Doppler display. The abscissa (horizontal axis) is time in seconds, and the ordinate (vertical axis) is flow velocity from the angle-corrected Doppler measurement. The gray scale in the display is the relative power in the Doppler signal for the velocity and time specified by the location on the display.

be imaged within the body in real time without the use of radioactive tracers. Commercially available devices can measure pulsatile flow velocities and can overlay frequency-shift (velocity) information, coded in color, on the traditional gray-scale B-scan image of the surrounding anatomy. Power-mode Doppler, in which the integrated Doppler spectral power is estimated and encoded in color, has been added to many ultrasound systems.

Duplex and triplex modes. The imaging information from a number of these modes can be combined in display formats termed duplex or triplex modes, in which ultrasound systems can continuously update the display at slow rates (10 images per second depending on image quality). **Figure 10** is an example of a duplex display, with both a B-mode image of the human carotid artery and the corresponding Doppler spectrum of the blood flow.

Nonlinear (harmonic) imaging. Nonlinear propagation is also being used in ultrasonic imaging. As an ultrasonic wave passes through tissue, higher harmonics will be produced in the portion of the field at high amplitude. Because of focusing, the higher amplitudes are confined to a narrower beam, and scattering from objects outside this beam is reduced. Therefore, emphasizing these nonlinear signals can produce better-resolution images. Artifacts introduced in the image from undesired reflections near the transducer, such as from ribs, are also reduced because the nonlinear components are produced only after a sufficient propagation distance. *See* BIOMEDICAL ULTRASONICS; MEDICAL IMAGING; MEDICAL ULTRASONIC TOMOGRAPHY; NONLINEAR ACOUSTICS.

Ultrasound contrast agents. Ultrasound contrast agents may provide additional means for overcoming signal-to-noise problems in smaller vessels. Most of these agents are stabilized gas bubbles, some of which are surrounded by shells of albumin, polysaccharides, or lipids. Some contain perfluorocarbon gases that have low solubilities and slow diffusion rates in aqueous media, which serve to slow their dissolution. Agents are used for imaging in the heart, brain, and other anatomic sites. Nonlinear imaging techniques are also being applied to these agents due to their strong nonlinear scattering that can produce substantially greater amplitudes in harmonics than tissue and even subharmonic emissions. In addition, higher-amplitude ultrasonic pulses can eliminate the signal from some contrast agents, and monitoring the rate at which signal returns provides a measure of blood flow.

Ultrasound bioeffects. The safety of medical diagnostic ultrasound has been studied almost since its initial use. In general, the approach has been to limit the acoustic output of ultrasonic devices to levels which have not produced observable bioeffects; that is, all diagnostic ultrasound devices should be equivalent in safety and effectiveness to pre-1976 ultrasound devices. Knowledge of potential thresholds for bioeffects has improved, and two indices have been identified to provide feedback to the user on potential bioeffects. The thermal index specifies the capability of raising the temperature of tissue. The mechanical index indicates the potential for the ultrasonic field to generate acoustic cavitation, which can produce thermal, chemical, and mechanical effects. These indices provide information for determining risk versus benefit when a physician determines that important information can be obtained at acoustic levels with bioeffects potential.

Ultrasonic therapy. Ultrasonic fields of sufficient amplitude can generate bioeffects in tissues. Although diagnostic ultrasound systems try to limit the potential for these effects, therapeutic levels of ultrasound have been used in medicine for a number of applications. Conventional therapeutic ultrasound is a commonly available technique used in physical therapy. High-frequency acoustic fields (typically 1 MHz) are applied through the skin to the affected area in either a continuous wave or long pulses.

Extracorporeal shock-wave lithotripsy (ESWL) disintegrates kidney stones with a high-amplitude acoustic pulse passing through the skin of the patient. The procedure eliminates the need for extensive surgery. Bioeffects are limited to the location of the stone by using highly focused fields which are targeted on the stone by imaging techniques such as ultrasound or fluoroscopy. The stone disruption process is not completely understood, but suggested mechanisms include shear stress and cavitation, which cause initial fracturing.

Ultrasound hyperthermia is used to treat a variety of conditions including cancer by heating the tissue to produce cell death. However, the technique is experimental in most applications. The hyperthermia systems typically use continuous-wave acoustic fields or acoustic pulses of considerably longer duration than those used in extracorporeal shock-wave lithotripsy.

Ultrasound ablative therapy uses higher-amplitude pulses to generate both mechanical and thermal disruption of the tissue.

Henry E. Bass; J. Brian Fowlkes; Veerle M. Keppens

Bibliography. C. E. Brennen, *Cavitation and Bubble Dynamics*, 1995; F. A. Duck, A. C. Baker, and H. C. Staritt, *Ultrasound in Medicine*, 1998; D. Ensminger, *Ultrasonics: The Low- and High-Intensity Applications*, 2d ed., 1988; D. H. Evans and W. N. McDicken, *Doppler Ultrasound: Physics, Instrumentation, and Signal Processing*, 2000; F. W. Kremkau, *Diagnostic Ultrasound: Principles and Instruments*, 1998; A. Migliori and J. Sarrao, *Resonant Ultrasound Spectroscopy*, 1997; E. P. Papadakis (ed.), *Ultrasonic Instruments and Devices*, 1999; K. S. Suslick, *Ultrasound: Its Chemical, Physical, and Biological Effects*, 1988.

Ultraviolet astronomy

Astronomical observations carried out in the region of the electromagnetic spectrum with wavelengths from approximately 10 to 350 nanometers. The ultraviolet spectrum is divided into the extreme-ultraviolet (EUV; 10–90 nm), far-ultraviolet (FUV;

90–200 nm), and near-ultraviolet (near-UV; 200–350 nm). Ultraviolet radiation from astronomical sources contains important diagnostic information about the composition and physical conditions of these objects. This information includes atomic absorption and emission lines of all the most abundant elements in many states of ionization. The hydrogen molecule (H_2), the most abundant molecule in the universe, has its absorption and emission lines in the far-ultraviolet. Thus, ultraviolet observations make it possible to probe a very wide range of physical conditions of matter in the universe, from the very cold gas in dense interstellar regions with temperatures of perhaps 30 K (−406°F) to the hot gas found in supernova remnants and in the coronas of stars and galaxies with temperatures approaching 10^6 K. *See* ASTRONOMICAL SPECTROSCOPY; ULTRAVIOLET RADIATION.

Observations. Ultraviolet radiation with wavelengths less than 310 nm is strongly absorbed by molecules in the atmosphere of the Earth. Therefore, ultraviolet observations must be carried out by using instrumentation situated above the atmosphere. Ultraviolet astronomy began with instrumentation at high altitudes aboard sounding rockets for brief glimpses of the Sun and stars. The first major ultraviolet satellite observatories to be placed in space were the United States *Orbiting Astronomical Observatories* (*OAO*s). *OAO 2* operated from 1968 to 1972 and provided the first full survey of the many kinds of ultraviolet sources in the sky, while *OAO 3* (*Copernicus*) operated from 1972 to 1980 and obtained high-resolution spectra of bright ultraviolet-emitting stars in order to probe the composition and physical state of intervening interstellar gas and to study the stellar winds of hot stars. Also, a number of smaller satellites, including the European *TD 1* and the Dutch *ANS*, provided very important survey measurements on the ultraviolet brightnesses of astronomical sources. *See* ROCKET ASTRONOMY.

With the launch of the *International Ultraviolet Explorer* (*IUE*) into a geosynchronous orbit in 1978, the full potential of ultraviolet astronomy to probe a wide range of scientific problems became a reality. The *IUE* satellite was a collaborative project of the U.S. National Aeronautics and Space Administration (NASA), the European Space Agency, and the United Kingdom Science and Engineering Research Council. It consisted of a reflecting telescope of modest size (18 in. or 45 cm in diameter) followed by several spectrographs with ultraviolet-sensitive detectors that produce ultraviolet spectra over the wavelength region from 120 to 320 nm. Between 1978 and 1996 the *IUE* obtained approximately 104,000 ultraviolet spectra of a wide range of astronomical objects, including comets and planets, cool and hot stars, exploding stars, external galaxies, and quasars.

In 1992 the extreme-ultraviolet window to the universe was opened with the launch of NASA's *Extreme Ultraviolet Explorer* (*EUVE*) satellite. The *EUVE* contains telescopes designed to produce images of the extreme-ultraviolet sky and spectra of bright extreme-ultraviolet sources in the wavelength range from approximately 10 to 90 nm. Because of strong absorption by neutral hydrogen in the interstellar gas, it was expected that the *EUVE* would not be able to see very many galactic sources. However, the irregular distribution of that gas has permitted the *EUVE* to probe to substantial distances in some directions. Most of the sources of radiation detected with the *EUVE* over $8^1/_2$ years of operation were stars with hot active outer atmospheres (or coronae) and hot white dwarf stars. *See* WHITE DWARF STAR.

Over the period 1985–1998, a number of missions were conducted or launched from the space shuttle, including the ASTRO missions containing the Hopkins Ultraviolet Telescope, the Ultraviolet Imaging Telescope, and the Wisconsin Ultraviolet Photo-Polarimeter Experiment. In addition, ultraviolet spectroscopy at medium and very high resolution has been accomplished with the Orbiting and Retrievable Far and Extreme Ultraviolet Spectrometer and the Interstellar Medium Absorption Profile Spectrograph. The scientific data for these missions were acquired over the 7–10-day period that the shuttle orbited the Earth.

Although it had an inauspicious beginning, the Hubble Space Telescope has been the centerpiece of both ultraviolet and visible astronomy. The telescope was launched into low Earth orbit in 1990, and shortly thereafter it was discovered that its 94-in.-diameter (2.4-m) primary mirror was ground incorrectly and the first images produced by the observatory were blurry. The problem was corrected in 1993 during a space shuttle refurbishment and repair mission. The complement of instruments aboard the Hubble Space Telescope has included imaging cameras operating at ultraviolet, visual, and infrared wavelengths, and spectrographs operating at ultraviolet and visual wavelengths. The two original Hubble spectrographs, the Faint Object Spectrograph and the Goddard High Resolution Spectrograph, were replaced with the Space Telescope Imaging Spectrograph in 1997. A very sensitive spectrograph designed to observe faint ultraviolet sources, the Cosmic Origins Spectrograph, was scheduled to be deployed on the Hubble Space Telescope in 2004 with the space shuttle *Columbia*. However, the *Columbia* atmospheric reentry disaster in 2003 led to the NASA decision in 2004 to cancel all future Hubble Space Telescope Shuttle servicing missions. It appears unlikely the Cosmic Origins Spectrograph will ever be deployed on the Hubble Space Telescope. With failing gyroscopes and batteries, the Hubble Space Telescope is not expected to be operational beyond 2008 or 2009 if there are no servicing missions. *See* HUBBLE SPACE TELESCOPE; SPECTROGRAPH; TELESCOPE.

The repaired Hubble Space Telescope has fulfilled its original expectations and provided spectacular high-angular-resolution pictures (better than 0.1 arcsecond) at visible and ultraviolet wavelengths, as well as low- and high-spectral-resolution ultraviolet spectra over the wavelength range from 120 to 320 nm. The ultraviolet measurements have been

used to pursue an extremely broad range of astronomical studies, from conditions in the atmospheres of the planets to measurements relevant to the creation of the elements in the first few minutes of the big bang. *See* BIG BANG THEORY.

The *Far-Ultraviolet Spectrograph Explorer* (*FUSE*) satellite was launched into an 800-km (500-mi) circular orbit in June 1999. *FUSE* is designed to explore the universe in the 90–120-nm region of the spectrum at high spectral resolution. The far-ultraviolet region of the spectrum is rich is diagnostic information about cold and hot gas in the interstellar and intergalactic medium and in the environments of stellar and extragalactic sources. The primary goals of the mission include measurements of the deuterium-to-hydgrogen abundance ratio in the Milky Way Galaxy and the intergalactic medium, and a study of the properties of hot plasmas and cold molecular hydrogen in the Galaxy. The deuterium abundance studies are providing insights about the production of deuterium in the big bang and the subsequent destruction of the element as it is cycled through stars.

Galaxy Evolution Explorer (*GALEX*), launched in 2003, is a modest-sized ultraviolet imaging and spectroscopic survey mission that will probe star formation over 80% of the age of the universe. The ultraviolet emissions detected by the far-ultraviolet detectors aboard *GALEX* are ideally suited for revealing the presence of hot young stars in external galaxies.

The *Cosmic Hot Interstellar Plasma Spectometer* (*CHIPS*), a small satellite launched in early 2003, is designed to carry out an all-sky study of the diffuse extreme-ultraviolet background radiation at wavelengths from 9 to 26 nm. *CHIPS* observations will be used to determine the temperature, ionization conditions, and cooling mechanisms of the hot gas believed to fill the region of space in the immediate vicinity of the Sun. The early results from this mission suggested the conditions in the local hot bubble are very different than expected.

Discoveries. The important discoveries of ultraviolet astronomy span all areas of modern astronomy and astrophysics. Some of the notable discoveries in the area of solar system astronomy include new information on the upper atmospheres of the planets, including planetary aurorae and the discovery of the enormous hydrogen halos surrounding comets. In studies of the interstellar medium, ultraviolet astronomy has provided fundamental information about the molecular hydrogen content of cold interstellar clouds along with the discovery of the hot phase of the interstellar medium, which is created by the supernova explosions of stars. In stellar astronomy, ultraviolet measurements led to important insights about the processes of mass loss through stellar winds and have permitted comprehensive studies of the conditions in the outer chromospheric and coronal layers of cool stars. The *IUE*, Hubble Space Telescope, and *FUSE* observatories have contributed to the understanding of the nature of the hot gaseous corona surrounding the Milky Way Galaxy. Ultraviolet observations of exotic astronomical objects, including exploding stars, active galactic nuclei, and quasars, have provided new insights about the physical processes affecting the behavior of matter in extreme environments. The spectrographs aboard the Hubble Space Telescope have revealed the existence of large numbers of hydrogen clouds in the intergalactic medium. These intergalactic clouds may contain much more normal (baryonic) matter than exists in the known luminous galaxies and stars. The measures of the abundance of deuterium in the Milky Way Galaxy and beyond have provided important constraints on the conditions in the evolving universe when it was only several minutes old. *See* COMET; COSMOLOGY; GALAXY, EXTERNAL; INTERSTELLAR MATTER; MILKY WAY GALAXY; PLANETARY PHYSICS; QUASAR; SATELLITE (ASTRONOMY); STAR; SUPERNOVA.

Blair D. Savage

Bibliography. J. N. Bahcall and L. Spitzer, Jr., The space telescope, *Sci. Amer.*, 247(1):40–51, July 1982; S. Bowyer, Extreme ultraviolet astronomy, *Sci. Amer.*, 271(2):32–39, August 1994; J. C. Brandt et al., The Goddard High Resolution Spectrograph: Instrument, goals, and science results, *Pub. Astron. Soc. Pacific*, 106:890–908, 1994; B. D. Savage, Ultraviolet-optical space astronomy: Past, present and future, in J. A. Morse, J. M. Shull, and A. L. Kinney (eds.), Ultraviolet-Optical Space Astronomy Beyond HST, *Astron. Soc. Pacific Conf. Ser.*, 164:3–14, 1999; B. D. Savage and K. R. Sembach, Interstellar abundances from absorption line observations with the Hubble Space Telescope, *Annu. Rev. Astron. Astrophys.*, 34:279–329, 1996; L. A. Shore, IUE: Nine years of astronomy, *Astronomy*, 15(4):14–22, April 1987.

Ultraviolet lamp

A mercury-vapor lamp designed to produce ultraviolet radiation. Also, some fluorescent lamps and mercury-vapor lamps that produce light are used for ultraviolet effects. *See* FLUORESCENT LAMP; MERCURY-VAPOR LAMP; ULTRAVIOLET RADIATION.

Near-ultraviolet lamps. Ultraviolet energy in the wavelength region from 320 to 400 nanometers is known as near ultraviolet, or black light. Fluorescent and mercury lamps can be filtered so that visible energy is absorbed and emission is primarily in the black-light spectrum. The ultraviolet energy emitted is used to excite fluorescent pigments in paints, dyes, or natural materials to produce dramatic effects in advertising, decoration, and the theater; in industrial inspection, fluorescent effects are often used to detect flaws in machined parts and other products, as well as invisible laundry marks.

Middle-ultraviolet lamps. Middle ultraviolet spans the wavelength band from 280 to 320 nm. Mercury-vapor lamps are sometimes designed with pressures that produce maximum radiation in this region, using special glass bulbs that freely transmit this energy.

One such lamp type is the sunlamp. The **illustration** shows the reflector sunlamp, with a self-contained filament ballast and starting mechanism.

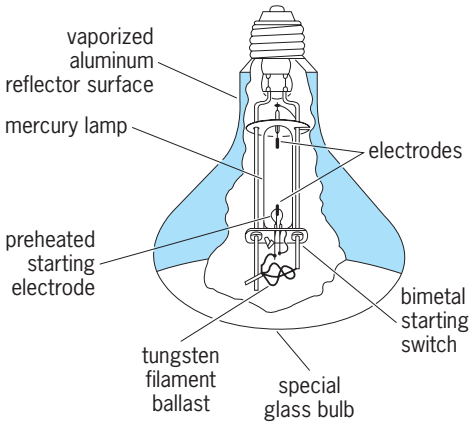

Cutaway view of ultraviolet lamp.

The reflector sunlamp combines the middle ultraviolet, which reddens the skin, with infrared energy and light from the filament to produce a suntanning effect with the sensations of warmth and brightness normally associated with sunshine. *See* SUNLAMP.

Other lamps designed for middle-ultraviolet radiation are known as photochemical lamps. They are used for a variety of tasks, including mold destruction, inspection of sheet metal for pinholes, and black-and-white printing of engineering drawings.

Far-ultraviolet lamps. Some radiation in the 220–280-nm wavelength band has the capacity to destroy certain kinds of bacteria. Mercury lamps designed to produce energy in this region (the 253.7-nm mercury line) are electrically identical with fluorescent lamps; they differ from fluorescent lamps in the absence of a phosphor coating and in the use of glass tubes that transmit far ultraviolet. Germicidal lamps are sometimes used to reduce airborne bacteria and to kill certain organisms on or near perishable products in storage or on certain products in the pharmaceutical industry. *See* ULTRAVIOLET RADIATION (BIOLOGY).

Alfred Makulec

Ultraviolet radiation

Electromagnetic radiation in the wavelength range 4–400 nanometers. The ultraviolet region begins at the short wavelength (violet) limit of visibility and extends to the wavelength of long x-rays. It is loosely divided into the near (400–300 nm), far (300– 200 nm), and extreme (below 200 nm) ultraviolet regions (see **illus.**). In the extreme ultraviolet, strong absorption of the radiation by air requires the use of evacuated apparatus; hence this region is called the vacuum ultraviolet. Important phenomena associated with ultraviolet radiation include biological effects and applications, the generation of fluorescence, and chemical analysis through characteristic absorption or fluorescence.

Biological effects of ultraviolet radiation include erythema or sunburn, pigmentation or tanning, and germicidal action. The wavelength regions responsible for these effects are indicated in the figure. Important biological uses of ultraviolet radiation include therapy, production of vitamin D, prevention and cure of rickets, and disinfection of air, water, and other substances. *See* ULTRAVIOLET RADIATION (BIOLOGY).

Fluorescence and phosphorescence are phenomena often generated as a result of the absorption of ultraviolet radiation. These phenomena are utilized in fluorescent lamps, in fluorescent dyes and pigments, in ultraviolet photography, and in phosphors. The effectiveness of ultraviolet radiation in generating fluorescence is shown in the illustration. *See* FLUORESCENCE; FLUORESCENT LAMP; PHOSPHORESCENCE; PHOTOGRAPHY.

Chemical analysis may be based on characteristic absorption of ultraviolet radiation. Alternatively, the fluorescence arising from absorption in the ultraviolet region may itself be analyzed or observed. *See* FLUORESCENCE MICROSCOPE; SPECTROSCOPY.

Sources of ultraviolet radiation include the Sun (although much solar ultraviolet radiation is absorbed in the atmosphere); arcs of elements such as carbon, hydrogen, and mercury; and incandescent bodies. The wavelengths produced by some sources of ultraviolet radiation are indicated in the illustration. *See* ULTRAVIOLET LAMP.

Artificial sources of ultraviolet light are often used to simulate the effects of solar ultraviolet radiation in the study of the deterioration of materials on exposure to sunlight. Trace amounts of chemicals which

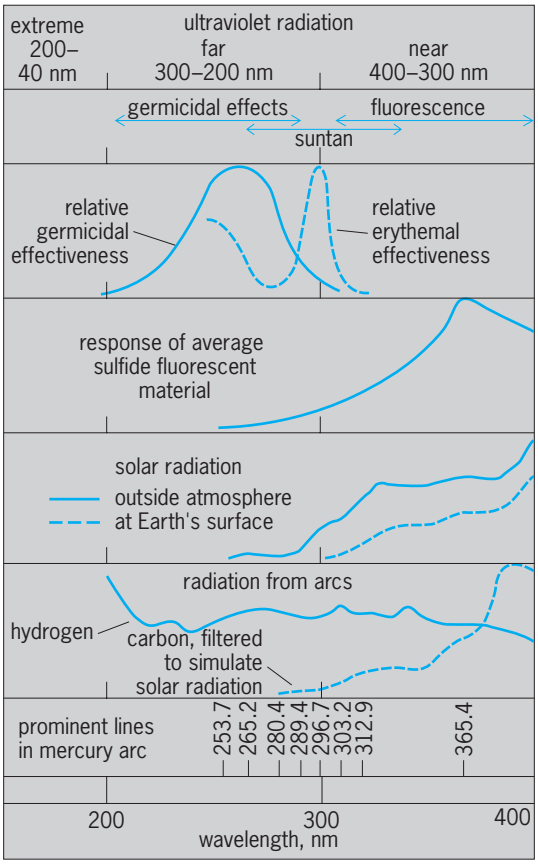

Phenomena associated with ultraviolet radiation.

strongly absorb ultraviolet radiation may effectively stabilize materials against such degradation. *See* INHIBITOR (CHEMISTRY).

Detectors of ultraviolet radiation include biological and chemical systems (the skin, the eye of an infant, or eye without a lens, and photographic materials are sensitive to this radiation), but more useful are physical detectors such as phototubes, photovoltaic and photoconductive cells, and radiometric devices.

Fred W. Billmeyer

Bibliography. R. C. Denney and R. Sinclair, *Visible and Ultraviolet Spectroscopy*, 1988; T. Frost, *Ultraviolet Spectrometry*, 1992; W. Harm, *Biological Effects of Ultraviolet Radiation*, 1980; R. E. Huffman (ed.), *Ultraviolet Technology*, 3 vols., 1986, 1988, 1989; H. Moseley, *Nonionising Radiation: Microwaves, Ultraviolet Radiation and Lasers*, 1988.

Ultraviolet radiation (biology)

The ultraviolet portion of the spectrum includes all radiations from 15 to 390 nanometers (others extend this range to 4–400 nm). Radiations shorter than 200 nm are absorbed by most substances, even by air; therefore, they are technically difficult to use in biological experimentation. Radiations between 200 and 300 nm are selectively absorbed by organic matter, and produce the best-known effects of ultraviolet radiations in organisms. Radiations between 300 and 390 nm are relatively little absorbed and are less active on organisms. Ultraviolet radiations, in contrast to x-rays, do not penetrate far into larger organisms; therefore, the effects they produce are surface effects, such as sunburn and development of D vitamins from precursors present in skin or fur. The effects of ultraviolet radiations on life have, therefore, been assayed chiefly with unicellular organisms such as bacteria, yeast, and protozoa, although suspensions of cells of higher organisms, for example, eggs and blood corpuscles, have been useful as well.

Ultraviolet radiation in sunlight at the surface of the Earth is restricted to the span from about 287 to 390 nm, although shorter wavelengths are present beyond our atmosphere, as shown by measurements with rockets. Consequently, artificial sources of the radiations are generally used in experimentation. *See* ULTRAVIOLET RADIATION.

Photobiological effects. Only the ultraviolet radiations which are absorbed can produce photobiological action. All life activities are shown to be affected by ultraviolet radiations, the effect depending upon the dosage. Small dosages activate unfertilized eggs of marine animals, reduce the rate of cell division, decrease the synthesis of nucleic acid, especially in the nucleus, reduce the motility of cilia and of contractile vacuoles, and sensitize cells to heat. Large dosages increase the permeability of cells to various substances, inhibit most synthetic processes, produce mutations, stop division of cells, decrease the rate of respiration, and may even disrupt cells. The effect of ultraviolet radiations upon cells is invariably deleterious.

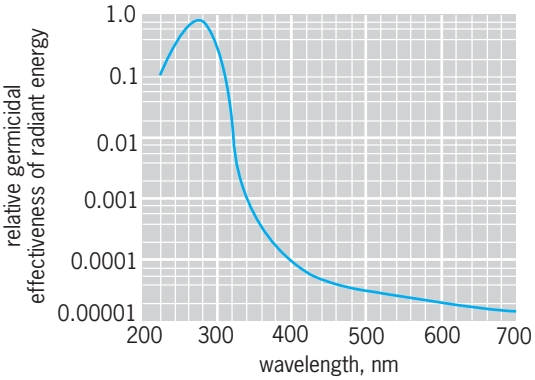

Relative bactericidal action of near-ultraviolet and visible regions for *Escherichia coli* on agar. (*After A. Hollaender, ed., Radiation Biology, vol. 2, McGraw-Hill, 1955*)

Despite their damaging effects, ultraviolet radiations are used as tools in biological research because they stop certain cell activities selectively without introduction of extraneous chemicals. They have been found especially useful in the production of mutations in microorganisms. *See* MUTATION.

Action spectra. Some wavelengths of ultraviolet radiations are more effective than others. More bacteria are killed by a given dosage of ultraviolet radiation at 260 nm than by the same dosage of radiation at 300 nm. When the bactericidal effectiveness of each of a series of wavelengths is plotted against the wavelength, the resulting curve is an action spectrum for the bactericidal effect (see **illus.**). Action spectra have been determined for many other effects of ultraviolet radiations. Each action spectrum is postulated to represent the absorption spectrum of the substance in the cell that is responsible for the particular effect. For the bactericidal effect, production of mutations, and retardation of cell division, the action spectrum suggests absorption by nucleoproteins of nucleic acid. For sensitization to heat and inhibition of ciliary movement, the action spectrum suggests absorption by ordinary proteins. For permeability, another action spectrum exists. *See* PHYSIOLOGICAL ACTION SPECTRA.

Mechanism. The effect of a given dosage of ultraviolet radiation upon protozoan cells is greater when the radiation is flashed than when continuous, that is, if a period of radiation is followed by a period of darkness. This indicates that a thermal reaction follows the primary photochemical reaction, a suggestion which is substantiated by the fact that increasing the temperature within the viable range accentuates the effect of flashing.

Action upon deoxyribonucleic acid (DNA), present in large amounts in chromosomes, consists primarily of the formation of pyrimidine dimers, chiefly between adjacent thymine residues on a strand. They interfere with replication of the DNA.

Photoreversal. The action of ultraviolet radiation on cells can be reversed to a considerable degree by simultaneous or subsequent exposure of the irradiated cells to short wavelength visible, violet and blue, or long wavelength ultraviolet light. This process

has been called photoreversal or photoreactivation. Thus, nucleic acid synthesis, inhibited by ultraviolet radiations, is resumed after exposure to visible light. At the same time, cell division, previously inhibited or retarded, is resumed. It appears that those effects of ultraviolet radiation having a nuclear site are most readily photoreversed.

Photoreversal consists of breaking of the thymine dimers into monomers, so reconstituting DNA in its original form. For this purpose a photoreactivating enzyme, which attaches itself to the dimers, is required as well as light.

Photoreversal is never complete; therefore, photoreactivated cells act as if they had been given a smaller dosage of ultraviolet radiations. It is evident that to make most effective use of ultraviolet radiations as a tool in experimental work, cells must be protected from visible light.

Dark repair. In addition to photoreactivation, it is now known that many cells kept from replication of DNA (in nonnutrient solutions) undergo dark repair. Pieces of DNA-containing dimers are excised (excision enzyme) and replaced on the basis of the information present in the other strand of DNA. Another enzyme (ligase) fastens the piece in place at the open end. Ultraviolet-resistant strains have very effective dark repair systems, while UV-sensitive ones do not.

Effects of ultraviolet on the skin. Erythema is the reddening of the skin following exposure to ultraviolet radiation of wavelength shorter than 320 nm, wavelength 296.7 nm being most effective. These radiations injure cells in the outer layer of the skin, or epidermis, liberating substances which diffuse to the inner layer of the skin, or dermis, causing enlargement of the small blood vessels. A minimal erythemal dose just induces reddening of the skin observed 10 h after exposure. A dose several times the minimal gives a sunburn, killing some cells in the epidermis after which serum and white blood cells accumulate, causing a blister. After the dried blister peels, the epidermis is temporarily thickened and pigment develops in the lower layers of the epidermis, both of these factors serving to protect against subsequent exposure to ultraviolet.

Both thickening of the epidermis and tanning may occur without blistering. Since the pigment in light-skinned races develops chiefly below the sensitive cells in the epidermis, it is not as effective as in dark-skinned races where the pigment is scattered throughout the epidermis. Consequently, the minimal erythemal dose is much higher for the dark- than for the light-skinned races.

Pigmentation or tanning also appears when the skin of young individuals is subjected to massive doses of ultraviolet radiations longer than 320 nm. Presumably this occurs by oxidation of precursors of the pigment, melanin, already present in the epidermis. Since such radiation is strong in sunlight, a skin may tan even in absence of short radiations.

Excessive exposure to ultraviolet radiation has been found to lead to cancer in mice, and it is claimed by some to cause cutaneous cancer in humans.

Clinical use. Ultraviolet radiations were once used extensively in the treatment of rickets, many skin diseases, tuberculosis other than pulmonary, especially skin tuberculosis (lupus vulgaris), and of many other diseases. The enthusiasm for sun bathing is, in part, a relic of the former importance of ultraviolet radiation as a clinical tool. Vitamin preparations, synthetic drugs, and antibiotics have either displaced ultraviolet radiations in such therapy or are used in conjunction with the radiations.

Ultraviolet radiations alone are still employed to treat rickets in individuals sensitive to vitamin D preparations. In conjunction with chemicals, they are used in treating skin diseases, for example, psoriasis, pityriasis rosea, and sometimes acne, as well as for the rare cases of sensitivity to visible light. They are also often used to sterilize air in hospitals. In some European laboratories, they are still used as adjuncts to drugs for treating lupus vulgaris and some other forms of tuberculosis. Ultraviolet radiations, however, are probably more important in research than in clinical practice. *See* RADIATION BIOLOGY; TUBERCULOSIS; VITAMIN D.

Arthur C. Giese

Bibliography. J. D. Longstreth (ed.), *Ultraviolet Radiation and Melanoma with a Special Focus on Assessing the Risks of Stratospheric Ozone Depletion*, 1987; H. Moseley, *Nonionizing Radiation: Microwaves, Ultraviolet and Laser Radiation*, 1988; J. A. Parrish (ed.), *The Effect of Ultraviolet Radiation on the Immune System*, 1983; M. Tevini (ed.), *UV-B Radiation: Effects on Humans, Animals, Plants, Microorganisms, and Materials*, 1993; World Health Organization Staff, *Solar and Ultraviolet Radiation*, 1992.

Ultrawideband (UWB) systems

An electronic system which has either an instantaneous bandwidth of at least 500 MHz or a fractional bandwidth (ratio of the instantaneous bandwidth measured at the −10 dB points to the center frequency) of at least 20%, whichever is greater. Although considered a recent breakthrough in broadband wireless technology, UWB has now experienced over 40 years of technology development.

Development of technology. Ultrawideband technology stems from work in time-domain electromagnetics begun in 1962 by Gerald F. Ross to describe fully the transient behavior of microwave networks through their characteristic impulse response. The conventional means of characterizing a linear, time-invariant system is by a swept, frequency response (that is, amplitude and phase measurements versus frequency), a very tedious endeavor for broadband systems. Ross, recognizing that a linear system could be fully characterized by its response to an impulse excitation (that is, the result of a single measurement), developed real-time hardware techniques to implement such a measurement. Mathematically, knowing the response to an impulse, the response to any arbitrary excitation could be readily determined by the convolution integral. However,

two fundamental problems existed: how to generate a close approximation to a mathematically ideal impulse (an infinitely large, infinitesimally short-duration waveform), and how to actually measure the extremely short duration responses that would be expected from such excitations.

In 1957, Leo Esaki invented the tunnel diode, the first known practical application of quantum physics. This device, with its extremely wide bandwidth (tens to hundreds of gigahertz), not only permitted subnanosecond pulse generation essential for impulse excitation but also could be used as a sensitive thresholding device for the subsequent detection of extremely short duration waveforms. In 1962, time-domain sampling oscilloscopes based on the tunnel diode were introduced for high-speed triggering and detection, first enabling the capture and display of ultrawideband waveforms. Impulse measurement techniques were subsequently applied to the analysis of wideband radiating antenna elements, in which the impulse response is the radiated electromagnetic field. In doing so, it quickly became apparent that short-pulse radar and even communications systems could be developed using the same set of tools. However, it was not until the introduction of a low-cost short-pulse receiver in 1972 to replace the expensive time-domain sampling oscilloscope that system developments in ultrawideband radar and communications accelerated rapidly. *See* ANTENNA (ELECTROMAGNETISM); TUNNEL DIODE; TUNNELING IN SOLIDS.

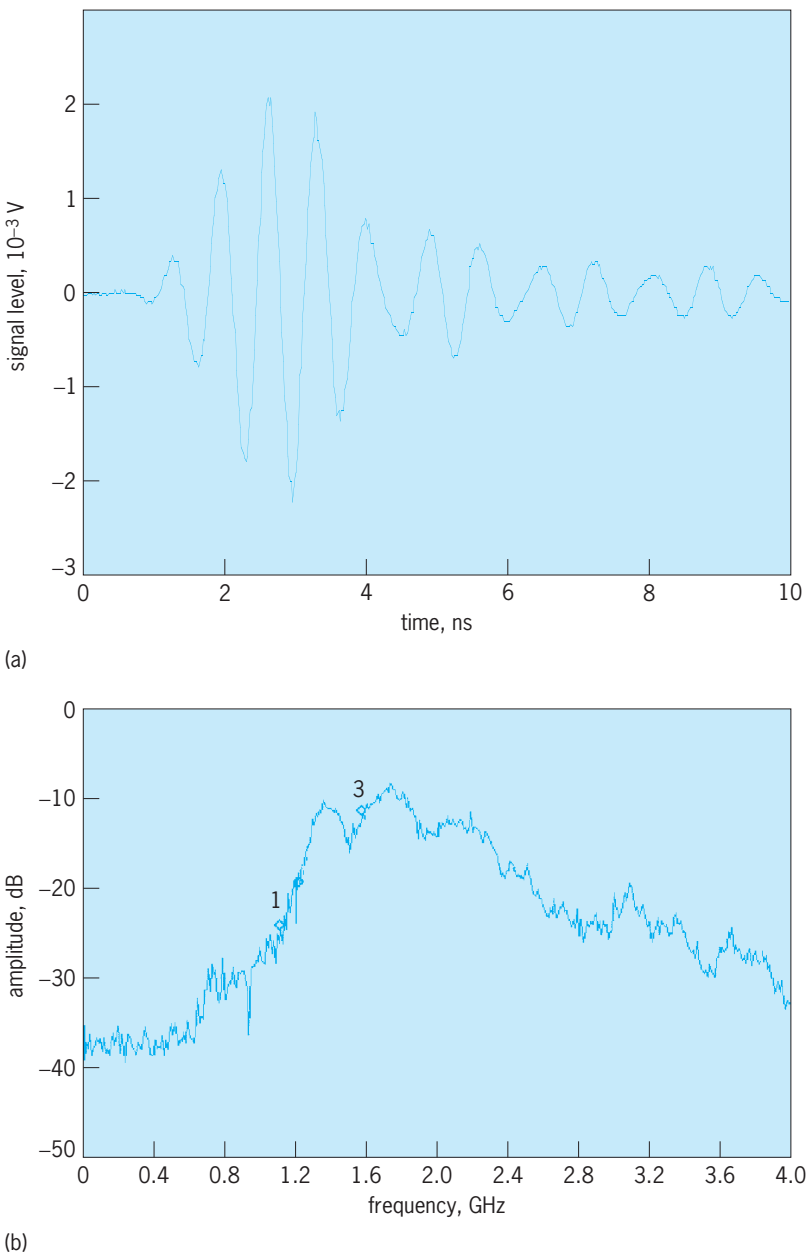

Fig. 1. Typical ultrawideband pulses in the time and frequency domains. (*a*) Typical ultrawideband transmit pulse in free space, with just a few cycles of radio-frequency energy. (*b*) Typical UWB power spectrum from impulse-excited wideband antenna. Center frequency $f_0 = 1.78$ GHz, and bandwidth at -10 dB relative to peak amplitude is 1.12 GHz, giving a fractional bandwidth of 63%.

Early techniques for ultrawideband signal generation used the fast rise (or fall) times of a baseband-generated pulse to impulse- or shock-excite a wideband antenna, which in turn would generate an electromagnetic burst consisting of only a few radio-frequency cycles of energy (**Fig. 1**). Tunnel, step-recovery, and avalanche diodes were used to produce this excitation signal, while modern techniques use the fast rise time of a semiconductor switch or gate. By varying the dimensions of the antenna, the frequency and bandwidth characteristics of the resulting ultrawideband pulse could be adjusted. Modern techniques for ultrawideband pulse generation use time-domain filtering techniques for spectral shaping.

Until the late 1980s, ultrawideband technology was alternatively referred to as baseband, carrier-free, or impulse. The term "ultrawideband" was not applied until approximately 1989, when the theory, techniques, and many hardware approaches of ultrawideband had experienced nearly 30 years of development. Somewhat paralleling the development of spread-spectrum systems after World War II, much of the early (pre-1994) development of ultrawideband systems (particularly for communications applications) was classified. The fact that these extremely short duration pulses required unique techniques for detection made them of particular interest for low-probability-of-detection applications for the government and military.

Significance of technology. The early interest in short-pulse techniques was primarily concentrated in short-range radar applications. From first principles, the shorter a radar's pulse duration, the finer the resultant range resolution. Thus, with pulse durations measured in hundreds of picoseconds, resolutions of only a few inches (several centimeters) or less were now possible. Such resolutions were orders of magnitude better than what could be achieved at that time, and significant research in the use of short-pulse techniques for radar applications continued.

From the time-scaling property of the Fourier transform relationship between time and frequency domains, the shorter the transmitted pulse, the wider the instantaneous bandwidth. Thus, since short-pulse waveforms could be inexpensively produced

directly at baseband (that is, without modulating a carrier frequency as in conventional radar), applications of ultrawideband to ground-penetrating radar and through-wall imaging followed quickly. Here, low-frequency wideband excitation translated into significantly enhanced material penetration capability. Ground-penetrating radar is used to locate buried objects, including humans in disasters such as building or mine collapses. Through-wall imaging is used for law enforcement and intelligence activities. *See* FOURIER SERIES AND TRANSFORMS; GROUND-PROBING RADAR; RADAR.

The early interest in ultrawideband for communications stemmed from the fact that short-pulse waveforms are extremely difficult to detect. Since the pulse bandwidth is spread over many hundreds of megahertz to gigahertz, a conventional (that is, narrowband) receiver will receive only a small slice of this spectral energy. The total received energy is directly proportional to the ratio of the intercept receiver to ultrawideband spread bandwidths.

These short-pulse waveforms have other unique properties as well. For example, one limiting factor affecting the performance of both mobile and indoor communications systems, particularly with narrowband systems, is the deleterious effect of multipath signal cancellation. In general, a received signal is the composite sum of the direct (or line-of-sight) signal from transmitter to receiver and a number of secondary reflections of the signal off objects between the transmitter and receiver. These "bounced" or multipath returns will arrive at the receiver later than the direct signal and, because of this time misalignment and possible phase inversion due to the reflection, will cause varying amounts of signal cancellation or distortion of the direct path. However, if the duration of the pulse is short enough, it is possible to distinguish in time between the direct and reflected multipath returns without causing any signal cancellation.

This capability is illustrated by the geometry of **Fig. 2**, in which two units, A and B, are attempting to communicate, but the signal also bounces off a wall located 3 m (10 ft) away. With the speed of light at roughly 0.3 m (1 ft) per nanosecond, a signal reflected from the wall will arrive approximately 10 ns (3-m or 10-ft path differential) later than the direct path. If the signal's pulse duration is less than this 10-ns differential, the direct and reflected path can be fully resolved without mutual interference. Thus, short-pulse techniques become of particular interest for high-reliability communications in severe multipath channels. However, if the delay spread of the channel (that is, the time over which such multipath effects are dominant) exceeds the bit duration (reciprocal bit rate) of the ultrawideband communications system, intersymbol interference (distortion) can occur and ultrawideband techniques become less effective.

Finally, because the time of arrival of a short-duration pulse can be very accurately measured using high-speed detection circuitry, such as enabled by the Esaki diode, accurate time-of-flight or distance measurements can he accomplished, opening up the use of the technology for precision localization applications.

Applications. Since the U.S. Federal Communications Commission (FCC) first established rules for the unlicensed use of ultrawideband technology in 2002, only a handful of ultrawideband designs have progressed beyond the laboratory stage.

Much of the short-pulse ultrawideband work either remains classified or has been used for applications having primarily government or military interest. Recent applications include an ultrawideband airborne wireless intercommunications system (AWICS) which uses the low-probability-of-intercept and multipath mitigation features of the ultrawideband waveform to provide a wireless communications capability within an aircraft or helicopter fuselage; ultrawideband wireless transceivers for the relay of voice, data and compressed video for covert network communications; ultrawideband perimeter intrusion radars for detecting unauthorized personnel; ultrawideband through-wall sensing radars for police and fire rescue; ultrawideband collision and obstacle avoidance radars for unmanned aerial vehicle (UAV) guidance and control; and several others.

On the commercial side, much of the lag in getting ultrawideband products to market has been because of the need for consensus on industry standards. The Institute of Electrical and Electronic Engineers (IEEE) is sponsoring two industry-wide standards for ultrawideband: IEEE 802.15.3a dealing with high-speed ultrawideband systems for wireless personal area networks (WPAN), and IEEE 802.15.4a dealing with low-speed ultrawideband systems for application to longer-range wireless networking and localization (for example, tracking of ultrawideband tags).

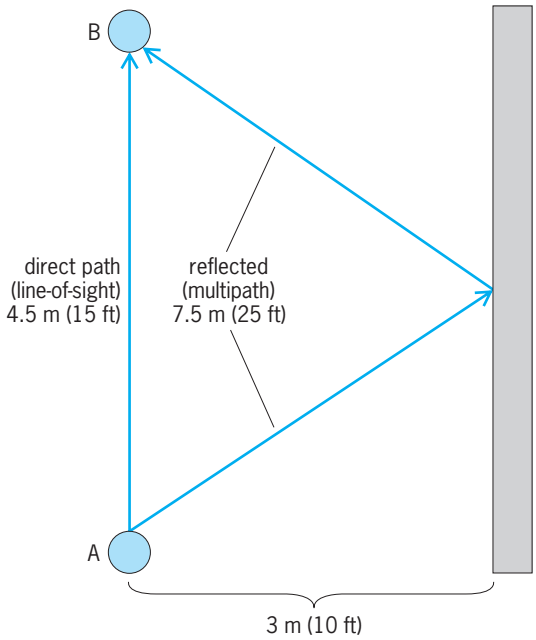

Fig. 2. Multipath geometry.

The high-speed WPAN standards group is split into two major camps, with one favoring the use of direct-sequence spread spectrum (DSSS) techniques and the other favoring the use of multiband orthogonal frequency division multiplex (M-OFDM). In both, the modulations chosen represent ultrawideband variants of more conventional waveforms. That is, instead of using short-pulse techniques to generate a large instantaneous bandwidth, these ultrawideband variants rely upon the underlying high-speed data stream to produce the requisite bandwidth. For example, DSSS techniques are used in IEEE 802.11b, cellular wireless, and various wireless home products (such as wireless telephones), while OFDM techniques are commonly found in asymmetric digital subscriber line (ASDL) services, IEEE 802.11a/g (WiFi) systems, digital audio broadcast (DAB), and digital terrestrial television broadcast in Europe and Japan. It appears that instead of a single standard for WPAN, nonstandard consumer products based on both of these technologies will emerge. *See* SPREAD SPECTRUM COMMUNICATION.

The low-speed ultrawideband localization standards group IEEE 802.15.4a in April 2005 completed the first step of soliciting proposals from industry, with the goal of reaching a consensus in the future. However, it appears that the vast majority in industry favor a physical layer (PHY) for ultrawideband localization, based on the use of short-pulse techniques. Unlike WPAN, several companies already have ultrawideband localization products on the market. According to measured data from these systems, the use of short-pulse techniques, together with precise time-of-flight measurements, enables 2- and 3D localization to within 15 cm (6 in.), even in the presence of severe multipath as is typically found in hospitals and industrial plants.

FCC report and order. As with any radio-frequency (RF) technology, both commercial and government use of ultrawideband requires appropriate frequency allocations. On February 14, 2002, the FCC approved a First Report and Order (R&O) to permit ultrawideband technology under Part 15 of its regulations. The R&O was precedent-setting in that, for the first time, the FCC had permitted intentional emissions to fall into previously restricted areas of the spectrum. The FCC waved restrictions on ultrawideband operation within certain restricted bands, but imposed additional constraints on the use of the technology for imaging, radar, and communications applications. These constraints included limitations on the field of operation and requirements for interagency coordination for the use of imaging and through-wall sensors, additional peak power constraints for all forms of ultrawideband operation, and operational restrictions on peer-to-peer indoor and handheld outdoor ultrawideband communications devices. Under this First R&O, ultrawideband communications devices are restricted to intentional operation between 3.1 and 14.6 GHz; through-wall imaging and surveillance systems are restricted to between 1.99 and 10.6 GHz (for use only by law enforcement, fire and rescue, and other designated organizations); and automotive radars are restricted to frequencies above 24.075 GHz. Ground-penetrating radar and through-wall imaging sensors were also permitted below 960 MHz.

On December 15, 2004, the FCC released a Second R&O on ultrawideband technology which opened up a broad range of new applications for both ultrawideband and wideband systems. Recognizing that modern ultrawideband devices can be spectrally confined to nonrestricted frequency bands, the FCC created two new Part 15 sections: 15.250 and 15.252.

Section 15.250 permits the operation of wideband systems (that is, systems now defined with a minimum bandwidth of at least 50 MHz) within the band 5925–7250 MHz. Except for operation on an aircraft or satellite or within a toy, wideband or ultrawideband systems operating under 15.250 have no further restrictions on their use as was previously mandated under the First R&O. Thus, indoor/outdoor, vehicular and shipboard, and other uses are now permitted at power levels equivalent to those specified in the First R&O.

The new Section 15.252 opens up two new frequency bands for the operation of wideband vehicular radar systems: 16.2–17.7 GHz and 23.12–29.0 GHz. The technology used in these frequency bands is envisioned for field-disturbance sensors providing automotive back-up assistance.

In January 2005, the United Kingdom's Office of Communications (Ofcom) released a consultation document that proposed rules for the use of ultrawideband technology within the U.K. Comments have been received from over 60 international government and corporate entities, and it is expected that the U.K. will incorporate rules for the unlicensed use of ultrawideband. However, it is anticipated that the U.K. will coordinate its rulemaking with the final recommendations from the International Telecommunication Union (ITU) Radiocommunication Bureau's Task Group 1/8 when it completes its study on the compatibility between ultrawideband devices and radiocommunication services. An ITU decision, representing the 189 member states, will have global ramifications for the introduction of ultrawideband technology.

Robert J. Fontana

Bibliography. C. L. Bennett and G. Ross, Time-domain electromagnetics and its applications, *Proc. IEEE*, 66(3):299–318, 1978; J. Foerster et al., Ultrawideband technology for short- or medium-range wireless communications, *Intel Technol. J.*, 2d Quarter, 2001; D. G. Leeper, Wireless data blaster, *Sci. Amer.*, May 2002; J. Taylor (ed.), *Introduction to Ultra-Wideband Radar Systems*, CRC Press, Boca Raton, FL, 1995; *2002 IEEE Conference on Ultra Wideband Systems and Technologies*, Institute of Electrical and Electronic Engineers, May 21–23, 2002; *2003 IEEE Conference on Ultra Wideband Systems and Technologies*, Institute of Electrical and Electronic Engineers, Nov. 16–19, 2003; *Ultra Wideband Systems 2004: Joint Conference on Ultra Wideband Systems and Technologies and International Workshop on UWB Systems*, Institute of Electrical and Electronics Engineers, May 18–21, 2004;

Ultra-Wideband, Short-Pulse Electromagnetics 1, 2, 3, and *4*, Plenum Press, New York, 1993, 1994, 1997, 1999.

Umklapp process

A concept in the theory of transport properties of solids which has to do with the interaction of three or more waves, such as lattice waves or electron waves, in a solid. In a continuum, such interactions occur only among waves described by wave vectors $\mathbf{k}_1$, $\mathbf{k}_2$, and so on, such that the interference condition, given by Eq. (1), is satisfied. The sign of $\mathbf{k}$

$$\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0 \tag{1}$$

depends on whether the wave absorbs or emits energy. Since $\hbar\mathbf{k}$ is the momentum of a quantum (or particle) described by the wave, Eq. (1) corresponds to conservation of momentum. In a crystal lattice further interactions occur, satisfying Eq. (2), where $\mathbf{b}$ is

$$\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = \mathbf{b} \tag{2}$$

any integral combination of the three inverse lattice vectors $\mathbf{b}_i$, defined by $\mathbf{a} \cdot \mathbf{b}_j = 2\pi\delta_{ij}$, the $\mathbf{a}$'s being the periodicity vectors. The processes described by Eq. (2) are Umklapp processes or flip-over processes, so called because the total momentum of the initial particles or quanta is reversed. *See* CRYSTAL STRUCTURE.

Examples of Umklapp processes are the following: (1) interactions of three lattice waves due to anharmonic lattice forces; of these, only processes described by Eq. (2) produce intrinsic thermal resistance in nonmetals; the exponential variation of the thermal resistance observed in dielectric crystals at low temperatures confirms the concept of Umklapp processes; (2) scattering of electrons by lattice waves, causing electrical and thermal resistance in metals; it has become clear that the observed properties cannot be accounted for in terms of processes described only by Eq. (1), but Umklapp processes must also be considered; (3) Bragg reflection, which can be regarded as an Umklapp process involving only two waves. *See* CONDUCTION (HEAT); ELECTRICAL CONDUCTIVITY OF METALS; THERMAL CONDUCTION IN SOLIDS; X-RAY DIFFRACTION. Paul G. Klemens

Bibliography. C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996; J. M. Ziman, *Principles of the Theory of Solids*, 2d ed., 1979.

Uncertainty principle

A fundamental principle of quantum mechanics, which asserts that it is not possible to know both the position and momentum of an object with arbitrary accuracy. This contrasts with classical physics, where the position and momentum of an object can both be known exactly. In quantum mechanics, this is no longer possible, even in principle. More precisely, the indeterminacy or uncertainty principle, derived by W. Heisenberg, asserts that the product of Δx and Δp—measures of indeterminacy of a coordinate and of momentum along that coordinate—must satisfy inequality (1). The Planck constant,

$$\Delta x \times \Delta p \gtrsim \hbar = \frac{h}{2\pi} \tag{1}$$

$h \simeq 6.6310^{-34}$ joule-second, is very small, which makes inequality (1) unimportant for the measurements that are carried out in everyday life. Nevertheless, the consequences of the inequality are critically important for the interactions between the elementary constituents of matter, and are reflected in many of the properties of matter that are ordinarily taken for granted.

Indeterminacy and atomic sizes. For example, the density of solids and liquids is set to a large degree by the uncertainty principle, because the sizes of atoms are determined with decisive help of inequality (1). In the ground state of the hydrogen atom, an electron is bound to a proton with an energy E_B equal to the sum of the potential energy (which is $-e^2/a$, because of the Coulomb attraction of two particles of charge e and $-e$, separated by the distance a) and the kinetic energy, as given in Eq. (2), where p and m_e

$$E_B = \frac{-e^2}{a} + \frac{p^2}{2m_e} \tag{2}$$

are the momentum of the electron with respect to the nucleus and its mass. In the absence of quantum indeterminacy, both p and a could be simultaneously arbitrarily small, and E_B could be arbitrarily negative. However, in quantum theory inequality (1) implies that the smallest momentum for an electron localized to within the distance a of a stationary proton is no less than $p \simeq \hbar/a$. Substituting this relation in Eq. (2) and minimizing E_B with respect to a leads to the conclusion that the radius of the hydrogen atom in the ground state is given by Eq. (3). This is

$$a_0 = \frac{h^2}{4\pi^2 m_e e^2} \simeq 0.529 \times 10^{-10} \text{ meter} \tag{3}$$

known as a radius of the first Bohr orbit of the hydrogen atom. The corresponding binding energy $E_B = -e^2/2a_0 \simeq -13.6$ eV, in agreement with much more elaborate calculations and with the measurements. *See* ATOMIC STRUCTURE AND SPECTRA.

When the atoms are tightly packed (separated by approximately $2a_0$), the typical density of the resulting material is equal to the proton mass divided by the volume of a cube of size $2a_0$, approximately 1.4 g/cm^3. While this calculation is a rough order-of-magnitude estimate, as it ignores complexities of atomic structure and the presence of neutrons in the nuclei (which would approximately double the density), and does not take into account the intricacies of atomic packing, it yields an essentially correct guess about the typical densities of solids and liquids. *See* QUANTUM THEORY OF MATTER.

Indeterminacy and the wave function. The wave-particle duality suggests a revealing perspective on quantum indeterminacy. States of quantum objects

can be represented by a wave function $\psi(x)$, which evolves in accord with the linear Schrödinger equation. Wave functions of definite wavelength λ (or wave number $k = 2\pi/\lambda$) are given by Eq. (4), and

$$\psi(x) \sim \exp(-ikx) \tag{4}$$

describe a free particle with a definite momentum $p = \hbar k = h/\lambda$. Almost any function of x (and, in particular, any wave function) can be expressed as a sum of such special plane-wave (definite-momentum) solutions, as in Eq. (5*a*). Here $\tilde{\psi}(k)$ supplies the complex weights of the different plane waves, and is known as a Fourier transform of $\psi(x)$. It can be obtained through Eq. (5*b*). The Fourier transform is

$$\psi(x) = (2\pi)^{-1/2} \int dk \tilde{\psi}(k) \exp(-ikx) \tag{5a}$$

$$\tilde{\psi}(k) = (2\pi)^{-1/2} \int dx \psi(x) \exp(ikx) \tag{5b}$$

an alternative, complete description of the same abstract state of the quantum, but in terms of the momentum rather than position: Both $\psi(x)$ and $\tilde{\psi}(k)$ contain exactly the same information. Thus, it is not possible to alter either of them [by, say, confining $\psi(x)$ to a smaller interval Δx] without affecting its transform. *See* SCHRÖDINGER'S WAVE EQUATION; WAVE MECHANICS.

It is a mathematical fact derived from the properties of Fourier transforms that a wave function $\psi(x)$ which is peaked within $\Delta(x)$ in position corresponds to $\tilde{\psi}(k)$ with a range of wave numbers, Δk, equal to or greater than approximately $1/\Delta x$ (**Fig. 1**). This leads to an equivalent statement of the uncertainty principle, Eq. (6). It turns out that the absolute minimum of this product is achieved for gaussian wave functions. These are given by Eq. (7*a*), where σ is a parameter, and their corresponding Fourier transforms by Eq. (7*b*). The ranges Δx and Δk can be

$$\Delta x \times \Delta k \gtrsim 1 \tag{6}$$

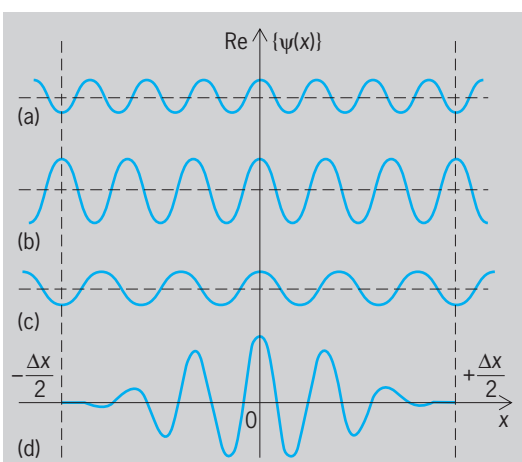

Fig. 1. Real parts of three wave functions with definite wave numbers (*a*) $k_0 + (\Delta k/2)$, (*b*) k_0, and (*c*) $k_0 - (\Delta k/2)$; and (*d*) their sum. The three wave functions add up constructively at $x = 0$, but interfere destructively at $+\Delta x/2$ and $-\Delta x/2$, resulting in a localized wave packet with a half-width inversely proportional to Δk. (*After C. Cohen-Tannoudji, B. Diu, and F. Laloë, Quantum Mechanics, Wiley, Hermann, 1977*)

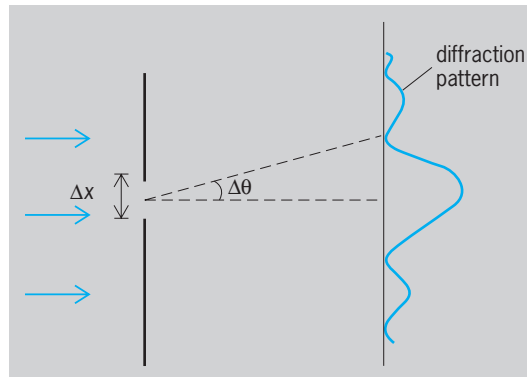

Fig. 2. Indeterminacy and a single-slit diffraction pattern. The size of the slit, Δx, is a measure of the uncertainty in position of the quantum. As the slit becomes narrower, the momentum in the direction parallel to the screen with the slit becomes more uncertain, resulting in a larger deflection angle, $\Delta\theta$, and a broader diffraction pattern. (*After R. P. Feynman, R. B. Leighton, and M. Sands, Feynman Lectures on Physics, vol. 1, Addison–Wesley, 1964, reprint 1989*)

$$\psi(x) \sim \exp(-x^2/2\sigma^2) \tag{7a}$$

$$\tilde{\psi}(k) \sim \exp(-k^2\sigma^2/2) \tag{7b}$$

defined as second moments of the probability distributions corresponding to $\psi(k)$ and $\tilde{\psi}(k)$. For the two gaussians in Eqs. (7), this yields $\Delta x = \sigma/\sqrt{2}$ and $\Delta k = 1/\sqrt{2}\sigma$. Hence, when the uncertainties are defined as dispersions, the minimum of the product of position and momentum is given by Eq. (8). The gaussian states which attain this minimum of uncertainty are sometimes called coherent states, and can be realized as a ground state of a harmonic oscillator. *See* COHERENCE; DISTRIBUTION (PROBABILITY); FOURIER SERIES AND TRANSFORMS; HARMONIC OSCILLATOR.

$$\Delta x \Delta p = \frac{\hbar}{2} \tag{8}$$

Indeterminacy and predictability. In classical physics, simultaneous knowledge of position and momentum can be used to predict the future trajectory of a particle. Quantum indeterminacy and the limitations it imposes force such classical notions of causality to be abandoned. This can be illustrated in a single-slit diffraction pattern (**Fig. 2**). Quanta arrive one by one from the left, each starting with a well-defined and identical momentum. They can pass through a screen with an opening (size Δx), so that to the right of the slit the position of each quantum in x is known with the accuracy determined by the size of the opening. However, as the wave packet is now restricted in x, its momentum along the x axis must be uncertain. As a result, when the slit becomes narrower, the diffraction pattern forming on a plate to the right of the screen (which can be thought of as a photographic plate) will become broader. Quanta will arrive one by one, blackening the photographic plate one grain at a time, but the pattern of many such

quanta will have a distinctly wavelike character. This further illustrates the wave-particle duality, characteristic of quantum theory. It also shows that prediction in quantum mechanics is statistical in character. In the example at hand, it is expressed by the probability density given by Eq. (9), which is evaluated at

$$p(x) = |\psi(x)|^2 \tag{9}$$

the point of detection, along the photographic plate.

Other uncertainty relations. Another well-known example of indeterminacy involves energy and time, as given by inequality (10). Physically, its origins are

$$\Delta E \times \Delta t \gtrsim \hbar \tag{10}$$

somewhat different from those of inequality (1), but mathematically this inequality again derives from the properties of Fourier transforms. Inequality (10) relates, for example, lifetimes of unstable states with the widths of their lines. *See* LINEWIDTH.

In quantum physics, relations similar to inequalities (1) and (10) hold for pairs of many other quantities. They demonstrate that the acquisition of the information about a quantum object cannot be usually achieved without altering its state. Much of the strangeness of quantum physics can be traced to this impossibility of separating the information about the state from the state itself. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS.

Wojciech Hubert Zurek

Bibliography. C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics*, 1977; R. P. Feynman, R. B. Leighton, and M. Sands, *Feynman Lectures on Physics*, 1964, reprint 1989; J. A. Wheeler and W. H. Zurek, *Quantum Theory and Measurement*, 1983.

Unconformity

In the stratigraphic sequence of the Earth's crust, a surface of erosion that cuts the underlying rocks and is overlain by sedimentary strata. The unconformity represents an interval of geologic time, called the hiatus, during which no deposition occurred and erosion removed preexisting rock. The result is a gap, in some cases encompassing millions of years, in the stratigraphic record.

Types. There are four kinds of unconformable relations (see **illus.**):

1. Nonconformity—underlying rocks are not stratified, such as massive crystalline rocks formed deep in the Earth.

2. Angular unconformity—underlying rocks are stratified but were deformed before being eroded, resulting in angular discordance; this was the first type to be recognized; the term unconformity was originally used to describe the geometric relationship between the underlying and overlying bedding planes.

3. Disconformity—underlying strata are undeformed and parallel to overlying strata, but separated by an evident erosion surface.

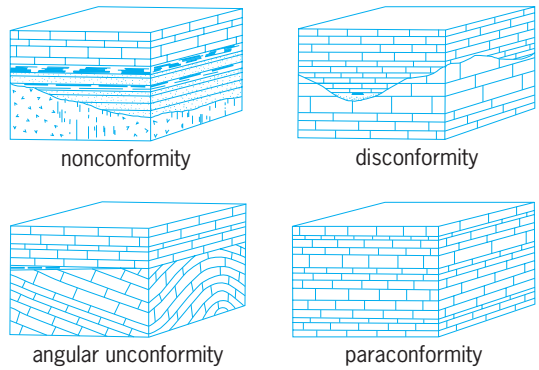

Four types of unconformity. (*After C.O. Dunbar and J. Rodgers, Principles of Stratigraphy, John Wiley and Sons, 1957*)

4. Paraconformity—strata are parallel and the erosion surface is subtle, or even indistinguishable from a simple bedding plane.

Importance in geologic record. Nonconformities imply deep or long-continued erosion. Angular unconformities are formed during episodes of mountain building; strata are deformed and then uplifted into the zone of erosion. Disconformities can be formed either by broad uplift (without deformation) or by a drop in sea level. Sea-level falls produce extensive disconformities that serve to define sequences of intervening strata. Several large-scale falls have generated continent-wide erosion surfaces; more frequent smaller-scale falls have affected only the continental margins. Marginal sequences can be recognized in the subsurface by using geophysical seismic profiling. Paraconformities are produced in the same fashion as disconformities, but their recognition is based on gaps in the biostratigraphic record; missing fossil zones indicate the breaks in deposition. On the deep ocean floor, paraconformities are formed both by physical erosion and by chemical dissolution of the sediment. The breaks correspond to climatic and tectonic changes that control the formation and circulation of oceanic bottom water. *See* STRATIGRAPHY.

Charles W. Byers

Underground mining

The extraction of ore from beneath the surface of the ground. Underground mining is also applied to deposits of industrial (nonmetallic) minerals and rocks, and underground or deep methods are used in coal mining. Some ores and industrial minerals can be recovered from beneath the ground surface by solution mining or in-place leaching using boreholes. *See* COAL MINING; SOLUTION MINING.

Underground mining involves a larger capital investment and higher production cost per ton of ore than open pit mining. It is done where mineral deposits are situated beyond the economic depth of open pit mining; it is generally applied to steeply dipping or thin deposits and to disseminated or massive deposits for which the cost of removing the overburden and the maintaining of a slope angle in

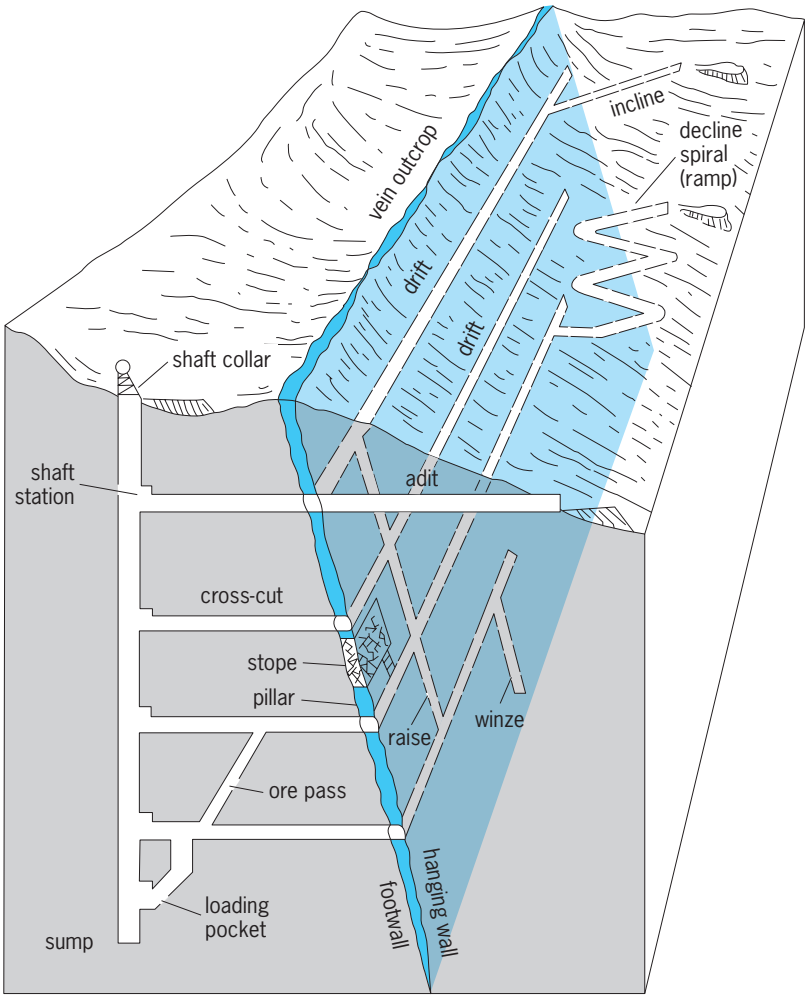

Fig. 1. Underground mining entries and workings.

adjacent waste rock would be prohibitive. In some situations, the shallower portion of a large orebody will be mined by open pit methods, and the deeper portion will be mined by underground methods. *See* OPEN PIT MINING.

Underground mine entries are by shaft, adit, incline, or spiral ramp (**Fig. 1**). Development workings, passageways for gaining access to the orebody from stations on individual mine levels, are called drifts if they follow the trend of the mineralization, and cross-cuts if they are driven across the mineralization. Workings on successive mine levels are connected by raises, passageways that are driven upward. Winzes are passageways that are sunk downward, generally from a lowermost mine level.

In a fully developed mine with a network of levels, sublevels, and raises for access, haulage, pumping, and ventilation, the ore is mined from excavations referred to as stopes. Pillars of unmined material are left between stopes and other workings for temporary or permanent natural support. In large-scale mining methods and in methods where an orebody and its overlying waste rock are allowed to break and cave under their own weight, the ore is extracted in large collective units called blocks, panels, or slices. *See* MINING.

Exploration

Exploration and development constitute the preproduction stage of underground mining. Exploration refers to the delineation of a newly discovered mineral deposit or an extension of a known deposit and to its evaluation as a prospect. During exploration, the deposit is investigated in sufficient detail to estimate its tonnage and grade, its metallurgical recovery characteristics, and its suitability for mining by various methods.

Information on the size, shape, and attitude of a deposit and information for estimating the tonnage and grade of the ore is taken from drill holes and underground exploration workings. Diamond core drilling provides intact samples of ore and rock for assaying and for detailed geologic and geotechnical study; percussion drilling provides chips of material for the recognition of ore and waste boundaries and for additional sampling. Underground exploration workings are used for bulk and detailed sampling, rock mechanics testing, and the siting of machinery for underground drilling. *See* DRILLING AND BORING, GEOTECHNICAL.

The tonnage and the grade of the material available in a mineral deposit are interrelated. The cutoff grade is the weakest mineralization that can be mined at a profit. Ore reserves are calculated in respect to the amount of ore in place at potential cutoff grades, the tonnages and average grades in identified blocks of ore, and the ultimate tonnage and grade of ore that should be available under projected conditions of recovery and wall rock dilution in mining. The suitability of a deposit for mining is determined in testing and evaluation work related to the physical and chemical nature of the ore, hydrologic conditions, and the needs for ground control. *See* ROCK MECHANICS.

Mine Development

Where high topographic relief allows for an acceptable tonnage of ore above a horizontal entry site, an adit or blind tunnel is driven as a cross-cut to the deposit or as a drift following the deposit from a portal at a favorable location for the surface plant, drainage facilities, and waste disposal. In situations where the deposit lies below or at a great distance from any portal site for an adit, entry must be made from a shaft collar or from an incline or decline portal. A large mine will commonly have a main multipurpose entry and several more shafts or adits to accommodate personnel, supplies, ventilation, communication, and additional production.

Adits. Access by adit generally provides for relatively low-cost underground mining. The broken ore from above the adit level can be brought to the portal in trains, conveyor belts, and rubber-tired trucks without the need for hoisting, and the workings can be drained without pumping. The driving of an adit is generally less expensive per unit distance of advance than the sinking of a shaft or the driving of an inclined access. In areas of low topographic relief and in the mining of deep orebodies, the sinking of a

shaft will often be a more economical approach than the driving and maintaining of a considerably longer incline or adit from a remote part of the site.

Shafts. Production shafts are generally located in stable ground on the footwall side of a dipping deposit rather than in the deposit itself or in the hanging-wall side, where protective pillars would be needed to maintain stability as mining progresses. A shaft may be inclined to follow the dip of the deposit and avoid increasingly longer cross-cuts to the ore at greater depth, but vertical shafts are more common because of their lower construction and maintenance cost per unit of depth and their better efficiency for hoisting ore. Shafts are sunk as rectangular or circular openings 15–30 ft (5–9 m) in diameter; they are equipped with a headframe and hoisting system and are lined with timber, steel forms, or concrete for ground support. Smaller shafts 5–15 ft (1.5–5 m) in diameter, generally for escapeways and ventilation, may be bored by mechanical drilling machines.

Inclines. Inclines equipped with hoists, declines for access by rubber-tired equipment, and gently inclined spiral ramps for diesel-powered truck haulage allow for direct access to relatively deep mine levels without having to transfer the ore and materials to hoisting systems.

Development workings. Development workings in the deposit consist of mine levels and sublevels, with drifts in the ore zone or in the more stable rock on the footwall side of the ore zone. Level workings serve as passageways for miners and low-profile equipment and as haulageways. In broken or unstable ground, passageways and haulageways are supported by timber sets and steel beams or arches; further stabilization is given by rock bolts, sometimes in combination with cable bolting and wire mesh, and the walls may be lined with concrete or spray-on shotcrete.

The raises that connect levels and sublevels provide for the removal of broken ore (chutes and ore passes), for access by miners, and for ventilation and supply routes.

In conventional mining and in the most common development procedures, headings are advanced in a cyclic sequence of drilling, blasting, mucking (removal of broken rock), and installing ground support. In continuous mining, the cycle is replaced by rapid excavation, a single operation in which headings are advanced by powerful tunnel boring and road header machines with teeth that break rock from the face. In situations where the uniformity and texture of the rock and ore permit development by continuous mining, the walls of the resulting passageways are smoother and more stable than would be provided by conventional cyclic operations involving blasting. *See* TUNNEL.

The continuous mining procedure of raise boring is well established. Shaft boring is used in the sinking of small-diameter ventilation shafts and escapeways. The driving of mine level development headings by cutting and boring machinery is more common in coal, potash, and salt deposits and in relatively soft sandstones and shales than in hard ore and rock.

Hydraulic breakers provide successively smaller rock sizes at development headings, and the broken rock is collected at the face by mechanical loading machinery and transferred to the mine haulage system by mobile conveyors or rubber-tired load-haul-dump machines. Haulage beyond the transfer point has been done by electric-powered locomotives with trains of cars but now is increasingly done by rubber-tired electric- or diesel-powered shuttle cars or trucks and by conveyor belt systems. In shaft mines, the broken rock is collected in underground storage pockets and loaded into skips for hoisting to the surface.

The entire sequence in mine development—the advance of headings, breaking of rock, loading, haulage, and hoisting—is increasingly automated. Teleoperated and autonomous machines have become central to every stage in mining, and new mines are developed with the use of geographic information systems (GIS) technology to accommodate the extensive communication systems and mining methods that relate to operations by remote control. *See* GEOGRAPHIC INFORMATION SYSTEMS.

Mining Methods

A fundamental condition in the choice of mining method is the strength of the ore and wall rock. Strong ore and rock permit relatively low-cost methods with naturally supported openings or with a minimum of artificial support. Weaker ore and wall rock necessitate more costly methods requiring widespread temporary or permanent artificial support such as rock bolting. Large deposits with weak ore and weak walls that collapse readily and provide suitably broken material for extraction may be mined by low-cost caving methods. Few mineral deposits are so uniform that a single method can be used without modification in all parts of the mine. Mining to an increasing depth with higher stress conditions and mining from a thicker portion of an orebody into thinner or less uniform portions will especially call for changes in method.

Naturally supported openings. The stopes remain open, essentially by their own strength, during ore extraction. Stability may be maintained to some extent by timbers, rock bolts, and accumulations of broken ore. The workings may collapse with time or may eventually need to be filled with waste material to protect workings in adjacent areas. Backfilling involves the placement of a paste of cemented waste rock or mill tailings. The methods range from gophering, an unsystematic small-scale practice, to carefully planned and executed systems using limits determined by rock mechanics investigations.

Open stoping. This is used in steeply dipping and thin orebodies with relatively strong ore and wall rock. In overhand methods the ore is stoped upward from a sill pillar by miners working on a staging composed of stulls (round timbers) and lagging (planks). With the drilling and blasting of successive small blocks of ore from the back (roof), the broken ore falls onto lower stagings and to the bottom of the stope; it is

collected on the haulage level through draw points or chutes. In underhand stoping the ore is mined downward in a series of benches, and the broken ore is scraped or hauled into a raise or ore pass for collection on a lower mine level. The width of an open stope is limited by the strength of the ore and its capability to stand unsupported. Occasional pillars, generally of waste or low-grade zones in a vein, are left for support; timber stulls may be wedged between the stope walls for stability as well as for access, and rock bolts may also be used to maintain wall stability.

Sublevel stoping. Also referred to as longhole or blasthole stoping, sublevel stoping is practiced in steeply dipping and somewhat wider orebodies with strong ore and strong walls (**Fig. 2**). Sublevel drifts and raises or slots are driven at the ends of a large block of ore so that a series of thinner horizontal slices can be provided. Miners in the sublevels drill patterns of radial holes (ring or fan drilling) or quarrylike parallel holes (slashing). Beginning at the open face of the initial slot, the ore is blasted in successive increments, and the broken ore falls directly to the bottom of the stope. A crown pillar is generally left unmined at the top of the stope to support the next major level.

Vertical crater retreat. This is a method of sublevel stoping in which large-diameter blastholes are drilled in a parallel pattern between major levels, and the ore is broken from the bottom of the stope in a sequence of localized blasts. All of the drilling, loading, and blasting are done by miners and teleoperated machinery in the upper level, so there is no need for access to the ore from below as the stope progresses upward.

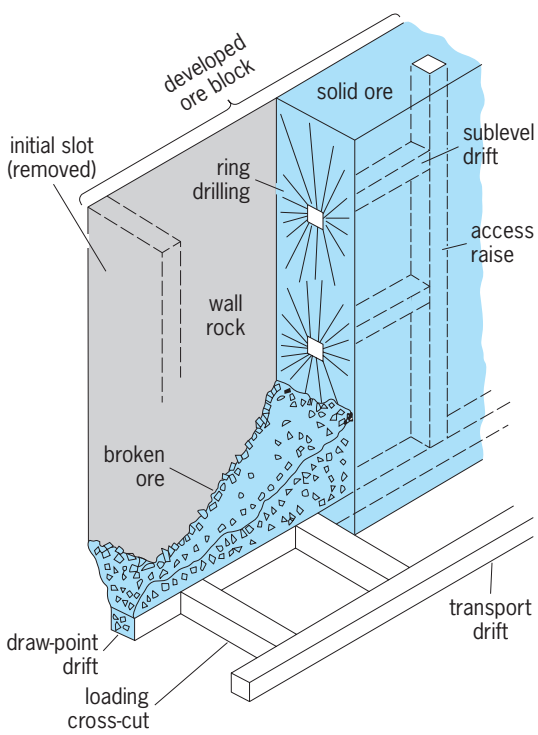

Fig. 2. Sublevel stoping, with ring drilling.

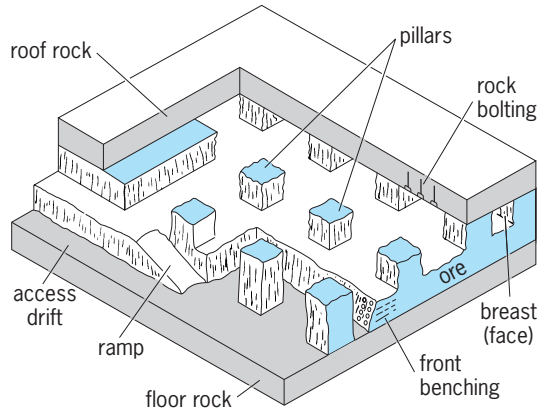

Fig. 3. Room-and-pillar mining; two-stage benching operation.

Room-and-pillar mining. This is also referred to as stope-and-pillar mining when done in a less regular pattern. Room-and-pillar mining is done in coal seams and in flat-lying or gently dipping ore and industrial mineral deposits (**Fig. 3**). It is a low-cost method of underground mining because fast-moving rubber-tired equipment can operate freely, especially in large rooms and haulageways. Thin-bedded deposits are generally mined in a single stage (pass) by conventional or continuous mining; thicker deposits are mined in a two-stage benching operation. In deposits of considerable thickness, an underground quarrying operation follows the first-stage opening of a development level for sufficient access by open-pit-type blasthole drills. Room-and-pillar mining is generally limited to depths on the order of 3000 ft (914 m) in hard-rock mines and to lesser depths in coal mines because of rock bursts and similar manifestations of high-stress concentration on the pillars. Extraction in mining generally amounts to about two-thirds of the ore in a bedded deposit, with the remaining ore being left in pillars; in places where pillars can be "robbed" and the roof allowed to settle, extraction can be increased to 90% or more. *See* ROCK BURST.

Shrinkage stoping. This is an overhand method in which broken ore accumulates in the stope, affording temporary support for the walls and a working platform for miners (**Fig. 4**). Shrinkage stoping is most applicable to steeply dipping veins with strong ore that will stand across a span and with relatively strong wall rock that would slough into the stope in places if left completely unsupported. When ore is broken, it has an expansion or swell factor; this necessitates a periodic drawing (shrinking) of some of the broken ore from the draw points and chutes to allow for continued access to the top of the stope. When all of the ore has been broken except for that left in pillars to protect the adjacent raises and mine levels, the entire content (the magazine) of the stope is drawn. The empty stope may be left open or filled with waste rock, and the pillars may eventually be mined.

Artificially supported openings. In these methods, workings are kept open during mining by using

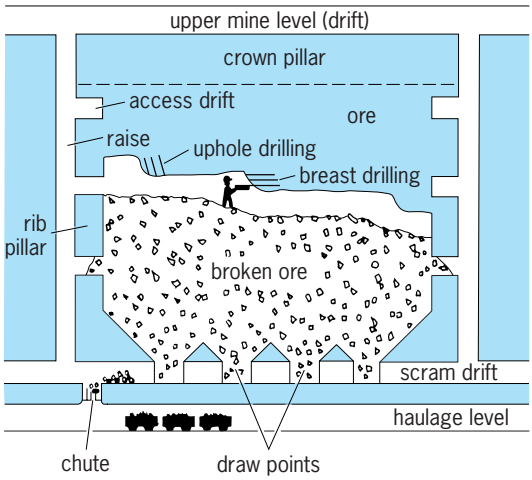

Fig. 4. Shrinkage stoping, longitudinal section.

waste material, timber, and hydraulic props. After the ore is extracted, the workings are filled to maintain stability or are allowed to cave.

Cut-and-fill stoping. This method, also referred to as drift-and-fill, is used in steeply dipping orebodies in which the ore has sufficient strength to be self-supporting but the walls are too weak to stand entirely without support (**Fig. 5**). Most cut-and-fill stoping is done overhand, with the drilling and blasting phase similar to that in shrinkage stoping; the broken ore, however, is removed from each new cut or slice along the back, and the floor of the stope is built up of waste material such as sand or mill tailings brought in by pipeline as a water slurry. The smooth and compacted or cemented fill material provides an especially suitable floor for rubber-tired machinery. Variations in cut-and-fill mining include the ramp-in-stope system, in which load-haul-dump equipment can move rapidly in and out of the stope on an inclined surface of fill material, and the less-mechanized system of resuing in narrow veins. In resuing, ore and waste material are broken separately and the waste material is left to accumulate as fill. One additional system, undercut-and-fill, is applied to bodies of weaker ore. It provides a solid artificial back of reinforced and cemented fill for the mining of successively underlying slices of ore.

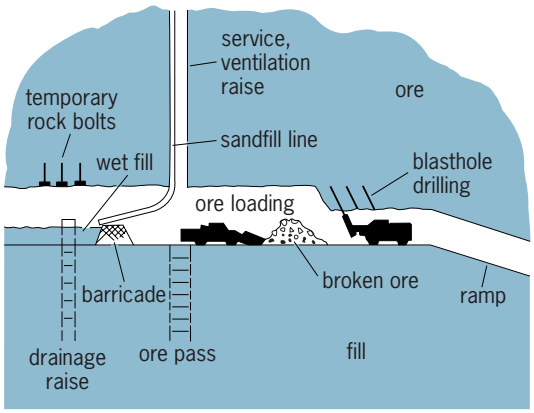

Fig. 5. Cut-and-fill stoping with sand slurry and ramp.

Square set stoping. This is a labor-intensive and high-cost method that has been classically used in situations where the ore is too weak to stand across a wide or long back and the walls are not strong enough to support themselves. A square set, a skeletal box of keyed timbers, is filled and wedged into the available space as each small block of ore is removed by drilling and blasting. Mining continues by overhand or underhand stoping, and the stope becomes a network of interlocked square sets. The sets in the mined portion of the stope are filled with mill tailings or waste rock and pillars are left between mined-out stopes for additional wall support while the remainder of the deposit is being mined. Because of its high cost, square setting is no longer in use; it has been superseded in many mines by cut-and-fill, top slicing, and sublevel caving methods.

Longwall mining. This method is applicable to uniform and extensive but relatively thin deposits. Primarily a highly mechanized and increasingly automated coal mining method at depths where rock pressures are too high for safe room-and-pillar mining, it has also been used in potash deposits and to some extent in bedded iron, copper, and uranium orebodies. In the South African deep gold mines, a form of longwall mining is used in the thin-bedded ore zones.

In longwall mining, practically all of the coal or ore is recovered except for that left in safety pillars to protect surface structures.

The basic practice is to maintain a temporary opening in a uniform line along a working face and then to allow the roof to cave onto the floor or waste fill (gob) behind the active area. In a typical mechanized longwall coal operation, the roof support units are canopies with hydraulic-powered adjustable legs or chocks that are moved ahead as the coal is shaved into slices by shearing and plowing machinery with integrated conveyor systems. In the mining of South African gold reef deposits, longwall-type mining is done by drilling and blasting; the active area is kept open by hydraulic props and timber-concrete packs, and the mined-out areas are filled to some extent by waste rock or cemented mill tailings.

Longwall mining systems allow for a high abutment pressure to build up in solid ore or coal in advance of the face, a low-pressure zone to exist in the working area just behind the face, and a normal lithostatic pressure to build up again in the mined-out and caved or gob-filled area as the face is moved ahead.

Top slice mining. Seldom used today, this method has been applied to wide and steeply dipping deposits with weak ore and weak walls. It has been of use in recovering pillars that have been left between filled stopes. It is a relatively expensive and labor-intensive method with a requirement for abundant timber, but it permits nearly total extraction of the ore. Top slicing is ultimately a caving method of mining, but the ore must first be drilled and blasted, and temporary support is needed between the taking of each successive downward slice or horizontal cut of ore. Working begins in drifts and cross-cuts on a mining floor

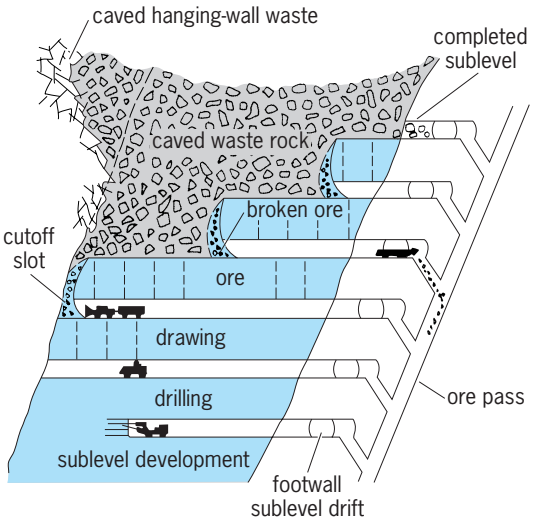

Fig. 6. Sublevel caving, with stages of development and mining.

at the top of a raise; after the driving of a series of adjacent cross-cuts so that a slice of sufficient width has been taken, a mat of timber and scrap lumber is laid down on the floor and the supporting timbers are blasted to cave the overlying rock. A new slice is mined laterally from drifts and cross-cuts under the mat, with the mat supported by timber props (stulls). A mat is again laid down, supports are blasted, and subsequent slices are mined beneath the subsiding accumulation of timber mats and waste rock.

Caving methods. These methods are used in large orebodies with relatively weak ore and with weak wallrock that will collapse as the ore is removed. Geologic conditions must permit subsidence, and the ore must be sufficiently jointed or fractured to form fragments small enough to be handled in drawpoints and raises. Ore recovery in mining is generally quite high, but a certain amount of dilution from waste rock must be accepted.

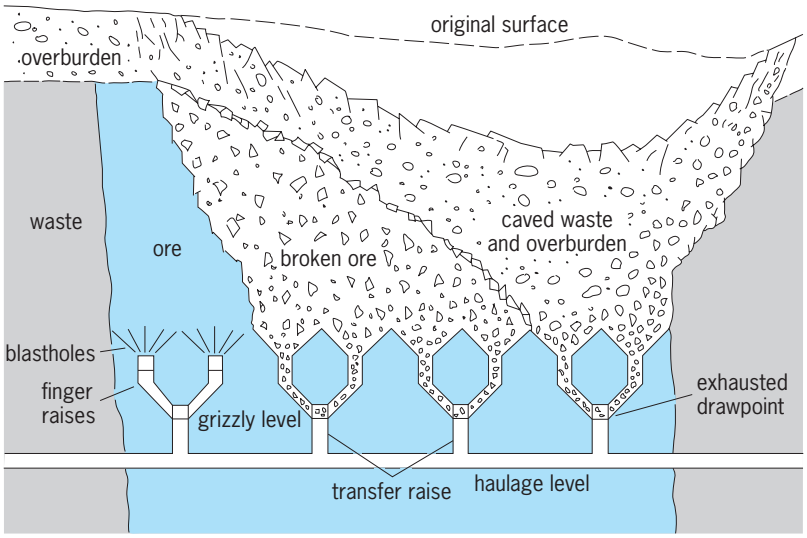

Fig. 7. Block caving, with principal haulage level, driving raises to production (grizzly) levels, and raises to workings.

Sublevel caving. This type is most suited to large and steeply dipping orebodies with weak walls and with ore that has enough stability to maintain sublevels (**Fig. 6**). It is similar to sublevel open stoping, but in this method the walls and the back are allowed to collapse. The ore is mined in downward increments that are drilled, blasted, and drawn from levels below the ore. Access drifts are driven on the footwall side of the orebody, sublevel cross-cuts are driven in ore, and fans of blastholes are drilled at intervals in the cross-cuts. A steplike succession of slices is mined in retreat from the hanging wall, with the wall rock collapsing and following the extraction of the ore. As each fan of holes is blasted, the broken ore caves into the sublevel, where it is loaded and transported to the ore pass. Broken waste rock fills the void as the ore is drawn. When an excess of waste rock begins to dilute the broken ore, the drawing is stopped and the next fan of holes is blasted.

Block caving. This is applied to large and relatively uniform bodies in which both ore and waste will cave readily (**Fig. 7**). Production on the order of 50,000–75,000 tons (45,000–68,000 metric tons) per day can be achieved at a very low mining cost, but the capital cost of a block-caving mine is high. A mine is prepared for block-caving operations by establishing a principal haulage level, driving raises to production levels (slusher or grizzly levels), and driving a larger number of raises to workings on an undercut level beneath the orebody or block to be mined. Caving is initiated by drilling and blasting a slice of ore above the undercut level and, if necessary, by excavating narrow stopes at the boundaries of the block. With the drawing of the initially broken ore, the block begins to cave under its own weight. With further drawing, the entire column of ore and overburden rock continues to subside and break upward for as much as 4000 ft (1220 m) to the surface, where a depression forms. The ore, broken and crushed in caving, flows through cone-shaped draw holes and finger raises. The finger raises are carefully monitored at draw points on the grizzly level so that the caving action is kept uniform and salient channels of subsiding waste rock are not allowed to form prematurely. Broken ore collected from finger raises reaches the haulage level through transfer raises. *See* EXPLOSIVE; MINING; PROSPECTING. William C. Peters

Bibliography. R. E. Gertsch and R. L. Bullock (eds.), *Techniques in Underground Mining*, SME, 1998; H. L. Hartman and J. M. Mutmansky, *Introductory Mining Engineering*, 2002; H. L. Hartman (ed.), *SME Mining Engineering Handbook*, 2 vol., 1992; B. Stack, *Handbook of Mining and Tunnelling Equipment*, 1982; K. S. Stout, *Mining Methods and Equipment*, 1980.

Underwater demolition

The controlled use of explosives to achieve specific underwater work requiring cutting, fragmenting, perforating, or pounding. The use of underwater explosives was pioneered by the military before and

during World War II for the removal of beach obstructions prior to amphibious landings, mine clearance, channel cutting, and the demolition of wrecked vessels.

With increasing commercial work in the oceans resulting from increased trade, complex marine salvage operations, oil and gas exploitation, and other industrial development, underwater demolition techniques have been widely applied and expanded. The technology has been developed to allow widespread and safe use of explosives as a tool in constructive underwater work.

The type of explosive material, its characteristics, and the shape and size of the charge all determine the effectiveness of explosives as a tool. Explosives are particularly effective under water because the water itself acts to "tamp" the explosion and increase its effect over that of a similar charge on the surface.

Military applications. The most common use of underwater demolition in the military remains the removal of obstacles to amphibious assault. Divers place charges on the obstacles. These charges are designed to cut or break up the obstacles.

When there is too little time or it is too dangerous to move mines or unexploded ordnance, divers place charges on or near the mine. When these charges are set off, the mine explodes and is destroyed.

Military and commercial organizations also use underwater demolition for clearing wrecked ships from harbors or waterways which must be used. In this work, advantage is taken of the ability of properly designed charges to cut or pound. Cutting charges may be used to cut away masts or to cut the ship's plates and internal structure. Pieces of the ship may be lifted out or the remainder may be pounded down with explosives to provide sufficient clearance for ships to pass over.

Another frequent task for underwater demolition that has both military and commercial use is widening or deepening a channel for the passage of shipping. Where channels must be made larger, explosive charges placed in regular patterns may be used to break up hard clay, sand, or rock bottoms so the bottom material can be removed by dredging.

Commercial applications. In addition to the wreck removal and channel widening and deepening uses described above, a wide range of commercial applications has been developed for underwater demolition. In the offshore oil and gas industries, charges may be placed and fired well below the ocean floor to open fissures in the rock, cut off steel pipe, open trenches, and remove old structures.

A large number of new explosives suited to particular tasks and the techniques to employ them have been developed. Special charge shapes and sizes allow very precise work to be done with explosives. While charges for military purposes are usually placed by divers, explosives used commercially can also be placed by crewed submersibles or remote-operated vehicles. *See* DIVING; EXPLOSIVE; SHIP SALVAGE.

William I. Milwee

Bibliography. E. I. du Pont de Nemours and Co., Inc., *Blaster's Handbook*, 1969; R. Larn and R. Whistler, *Commercial Diving Manual*, 1984; *Military Explosives*, 1991; U.S. Navy, Bureau of Ordnance, *Use of Explosives in Underwater Salvage*, 1956.

Underwater navigation

The process of directing the movements of submersible vehicles, and divers, from one point to another. The development of improved submersible vehicles, coupled with advances in saturated diving, has resulted in new requirements for underwater navigation. Various methods which have proved successful include acoustic transponder systems, dead reckoning, surface-referenced navigation from a support ship, terrain map matching, homing, and various combinations of these. The choice of the navigation system depends on such factors as precision required, area to be covered, availability of surface vessels, sea state under which they are expected to operate, and the duration of the mission. Redundant systems are routinely provided to maintain the safety of crewed vehicles. *See* DIVING; SUBMARINE; UNDERWATER VEHICLE.

Most mission requirements must be met by a combination of navigation equipments in an integrated system (**Fig. 1**). Therefore the acoustic equipments described are compatible to a large extent, and interference between sonars is minimized. Further, dead-reckoning computations make provision for periodic updating of position from other navigational equipment.

Acoustic transponders. The most accurate undersea navigation systems use acoustic transponders to precisely determine position with respect to points on the sea floor. A minimum of two transponders are normally dropped from a surface or submerged vehicle to establish a bottom-tethered array. The transponders can be interrogated by surface vessels, submersible vehicles, or divers. When the navigation solution is computed on the surface vessel, the submerged vehicle also carries a transponder, and the position of the submerged vehicle is computed relative to the bottom-tethered array. The location of these transponders can be determined by an accurately positioned surface vehicle that transmits a series of interrogation pulses from different locations. The location of the submersible is determined by a series of round-trip time measurements obtained by interrogating transponders at known locations.

The small size and low power consumption of microminiaturized computers readily allow on-board, and real-time, position determination by a submerged vehicle interrogating the transponder array. This method is widely used with both crewed and crewless submersibles as well as towed vehicles. [Crewless underwater vehicles are customarily referred to as unmanned underwater vehicles (UUVs), and are also referred to as autonomous underwater vehicles (AUVs).] Since the interrogating sonar is well below the surface layers where multipath reflections from the air-water interface do not distort

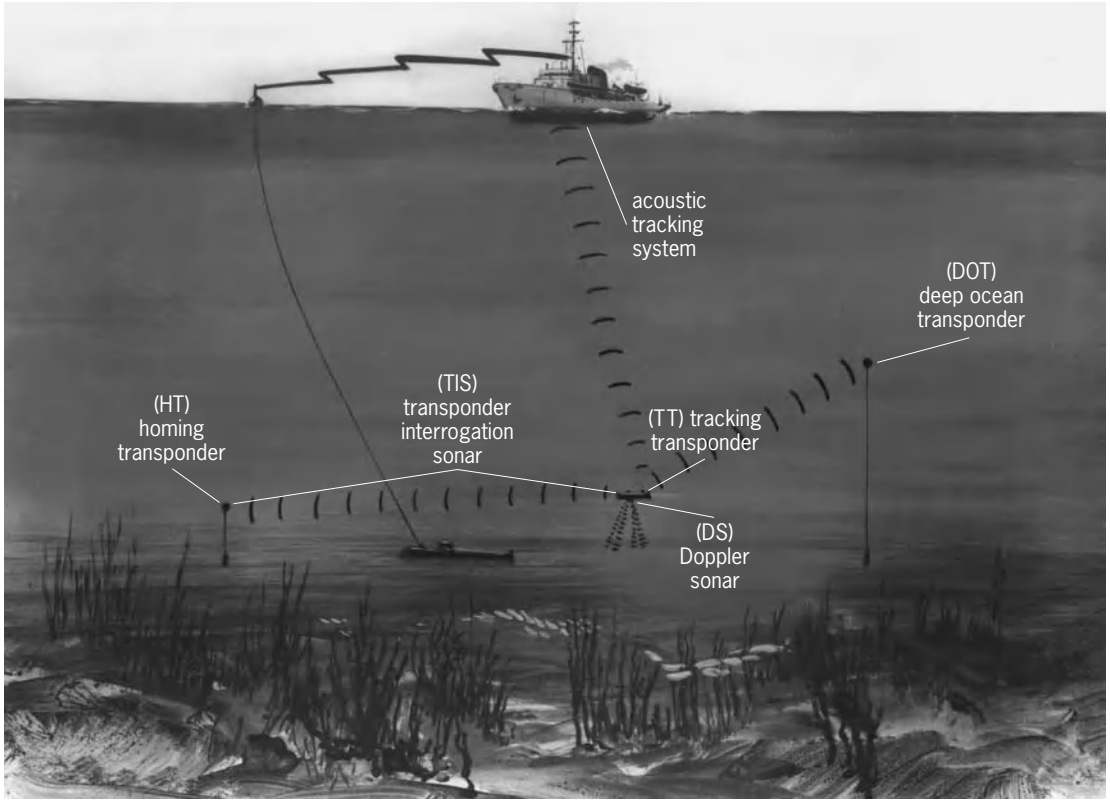

Fig. 1. Acoustic navigation systems, which employ various equipment.

received sonar signals, it provides a high-accuracy solution. The time delay between interrogation and received return (plus a fixed transponder turnaround delay) is a measure of the range of each transponder from the interrogating source. Individual transponders are identified by selected frequencies or digital codes. The differentiation, whether by frequency or code, may be applied to either the interrogate pulse or to the transponder reply. If each transponder has a unique response to interrogation, then all transponders within range can be interrogated with a common signal.

A typical deep-ocean navigation transponder (**Fig. 2**) is of cylindrical configuration and contains the electronics, battery, and transducer. A suitable float and anchor are also provided. Because of the bottom shadow found in deep water (discussed below), it is essential to place the transducer well above the sea floor. The height chosen for the deep-ocean transponder (DOT) is a compromise between range and uncertainty of position due to the watch circle, which is defined as the horizontal range over which a tethered object can vary in response to ocean currents. The frequencies employed for both reception and reply are commonly in the range of 7 to 20 kHz.

Fig. 2. Deep-ocean navigation transponder. This beacon is about 55 in. (140 cm) in length and 5 in. (12.5 cm) in diameter. (*Sona Tech, Inc.*)

The key element of any transponder design is the signal recognition circuit. This circuit must be able to recognize short interrogation pulses of varying amplitude in the presence of widely changing noise levels and to reject a large percentage of false signals due to noise bursts. The most successful circuit consists of a modified differential amplifier and a phase detector to compare the total energy received over a wide bandwidth to that in a relatively narrow band centered on the interrogation frequency. This approach rejects loud noise pulses but responds to weak interrogations. After recognition of a signal, the receiver is gated off for a short period of time to preclude multiple responses, for example, those which might occur from signals reflected off the bottom or the surface of the ocean. *See* DIFFERENTIAL AMPLIFIER; UNDERWATER SOUND.

A triangular array of transponders (**Fig. 3**) is an efficient deployment pattern since the addition of a single transponder can double the area in which the vehicle is able to search or navigate. This doubling is intuitively clear from the following argument. Three transponders define a triangle.

Adding a fourth transponder to form a parallelogram doubles the coverage area. This may be a prime consideration if the submersible must carry and deploy the transponders. The transponder spacing is determined from the estimated maximum range of the transponders. The maximum range is a complex function of vehicle altitude, transponder tether length, and the operating-area depth-temperature profile.

In deep water, acoustic rays are refracted upward by thermal gradients that change the index of refraction for propagation of sound waves, thereby creating a bottom shadow zone. Each two-way time delay measurement from the UUV to a transponder locates the UUV on a sphere that is centered at the transponder location and whose radius is the one-way distance to the transponder. The intersection of two of these spherical loci places the UUV on a circle. The addition of a vertical measurement produces two potential intersections with this circle. This two-point ambiguity limits the range at which a vehicle near the bottom can hear a transponder far more severely than attenuation of the signal by the absorption or spreading loss. Maximum practical range of transponders, such as the deep-ocean transponder is less than 7–10 nautical miles (13–18 km) near the surface and reduces to less than 4 nmi (7 km) when both the transponders and the interrogating vehicle are 300 ft (90 m) above the bottom. The compromise usually chosen is to space transponders 1–3 nmi (2–5.5 km) apart and to place more transponders where the area of operation exceeds twice the transponder spacing.

A two-transponder array is useful if additional data, such as the vertical displacement, are known. Both a three-transponder array and a two-transponder array plus vertical displacement result in a two-point ambiguity. The ambiguity can usually be resolved by a rough knowledge of the vehicle's position.

The transponder array survey-calibration error is the largest contributor to the navigation error. A transponder array is calibrated from either a surface ship or a submersible by making a pass over each transponder (to determine depth) and a pass over the center of each baseline (to determine the intertransponder spacing). Data collected from a submersible have the advantage of being taken below surface layers. The collected data are batch-processed with least-squares fit algorithms to compute the transponder positions. A number of transponders incorporate self-calibrating features. Upon receipt of a calibrate command, the transponders operate in what is commonly referred to as a sing-around mode in which they interrogate one another, and in effect generate (or collect) the data which allow determination of intertransponder spacing. The transponders may additionally be instrumented and programmed to measure pressure (to determine transponder depth) and temperature (to determine sound velocity within the array).

It is possible to solve for position within a transponder net with a hand calculator using a simplified and linearized set of range equations. However, the exact equations are nonlinear and are most expeditiously solved with a digital computer. A computer solution can dispense with linearization and simplifying assumptions, and will readily compute ray-bending corrections and harmonic mean sound velocity, and account for Earth curvature if necessary. A primary advantage of the digital computer is in the implementation of data filters to reject inconsistent transponder returns and navigation solutions. Transponder navigation systems have demonstrated excellent repeatability of data and the ability to return to the same spot on the sea floor to an accuracy of 16 ft (5 m) or less.

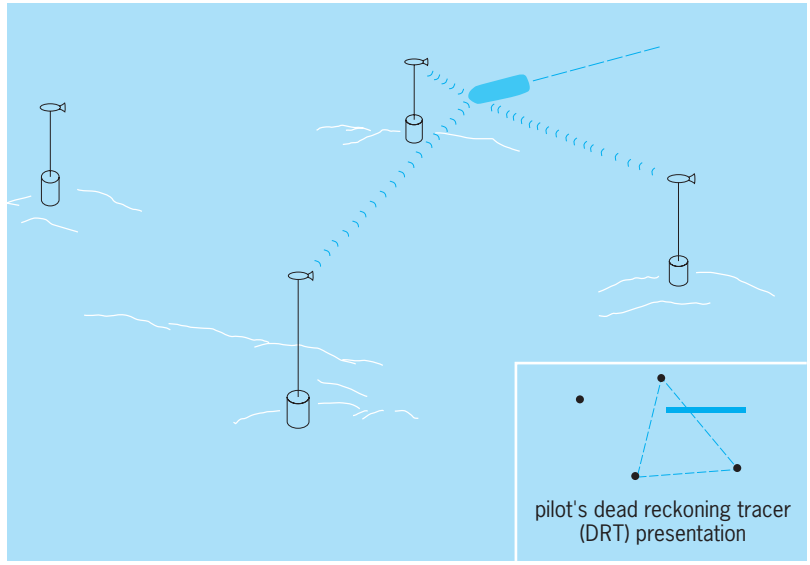

Fig. 3. Submersible navigating from underwater acoustic transponders.

Dead reckoning. An undersea vehicle needs sensors to show distance traveled and direction of travel to mechanize dead-reckoning estimation of position. The most promising sensor for operation near the sea floor is a Doppler sonar. Ground speed, fore-aft and athwartship, can be determined from the Doppler shift in frequency of signals returned from the sea floor. Pulse-type and continuous-wave Doppler sonars are available. The pulse type, which makes use of a frequency tracker circuit to lock on the received pulse, appears to offer the best accuracy, and it has the ability to operate to 600 ft (180 m) above the sea floor. Range gating to reduce noise reception is another feature possible with pulse-type Dopplers. Certain Doppler sonars operate on returns from reverberations from the water mass. This technique extends the usefulness of Doppler sonars to those situations in which the vehicle is out of range of the sea floor. Accuracy is of course much less, since the velocity measurement is referenced to a moving water mass. *See* SONAR.

Distance traveled is calculated from integration of the x and y sonar velocities. Accuracies of 0.1–0.2% of distance traveled can be obtained at altitudes ranging from about 3 to 600 ft (1 to 180 m) above the bottom. Much greater range is obtained in deep-water Doppler sonars. The technology of Doppler velocity

logs (DVLs) has been significantly improved since the mid 1990s, in part because of the increased interest in and military utility of UUVs.

Correlation velocity sonar can provide good velocity data at much greater bottom ranges. This sonar transmits pulsed energy in a broad beamwidth vertically to the bottom. Reflected echoes are received by an array of identical hydrophones, each having a broad beamwidth. The C-shaped receiving array provides fore-aft and athwartships components of velocity which are estimated independently of the velocity of sound. A correlation sonar, operating at altitudes of up to 16,000 ft (5000 m) over the sea bottom, can measure velocity components to accuracies which, when integrated, give a positional accuracy of 0.4% of distance traveled.

Since the Doppler and correlation sonars both measure in ship's body axes, a north-referenced heading measurement is required for navigation. This is usually provided by a small gyrocompass or magnetic compass. Accuracy for the dead-reckoning system is about 0.5% of distance traveled. *See* GYROCOMPASS.

In addition to Doppler sonar, some specialized submersibles carry inertial dead reckoning in the form of a small, stabilized platform. The accuracy of any dead reckoning system can be improved by receiving additional position information that is more accurate than the errors that have accumulated since the last position update. This opens up opportunities to use sonar-based underwater terrain map matching or occasional Global Positioning Satellite System (GPS) fixes collected by floating an antenna to the surface as sources of these independent position updates. *See* DEAD RECKONING; INERTIAL GUIDANCE SYSTEM.

Homing and piloting sensors. Most submersible vehicles carry additional navigation sensors. The horizontal obstacle sonar (HOS) which is a constant-transmission frequency-modulated sonar (CTFM), is normally used to detect objects ahead of the submersible (**Fig. 4**). This sonar also has a transponder channel for determining bearing and range to specially designed transponders. An altitude-depth sonar provides vertical navigation by furnishing depth and altitude off the bottom (Fig. 4). Finally, a vertical obstacle sonar (VOS) determines heights of objects in the path of the vehicle (Fig. 4). This sonar is a constant-transmission frequency-modulated type with a higher transmission frequency than the horizontal obstacle sonar. Both horizontal and vertical obstacle-avoidance sonars are useful in under-ice navigation. *See* HOMING; PILOTING.

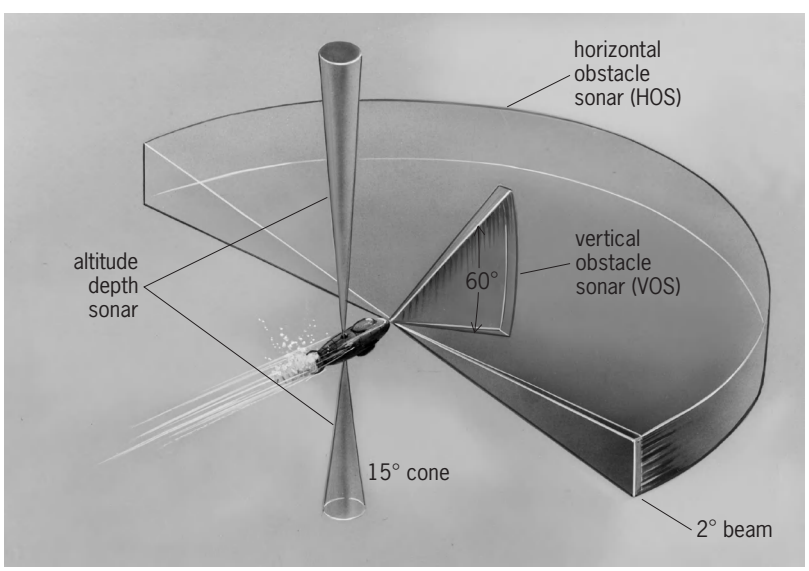

Fig. 4. Configurations of three types of pilotage sonars.

Surface reference. An acoustic tracking system allows the monitoring and vectoring of the position of a submersible from a surface vessel, where space and weight are not at a premium. The Global Positioning Satellite System or other radio techniques can be used to determine the ship's own position in geographic coordinates. Surface tracking systems are of two types: ultrashort baseline, employing orthogonal line hydrophone arrays; and short baseline, employing four separate hydrophones. *See* MARINE NAVIGATION.

When monitoring a vehicle carrying a transponder (Fig. 1), both range and direction (bearing and depression angle) can be determined. Range is computed from the time interval between interrogation and received reply, while bearing and depression angles are determined from phase differences between the received signals at different elements of the hydrophone array (first system), or from time-of-arrival differences at the separate hydrophones (second system).

The ultrashort-baseline system employs a single transducer, mounting two orthogonal arrays in an assembly less than 10 in. (25 cm) in diameter. Typical systems have achieved accuracies of 1% of slant range, within a horizontal offset of four times the water depth.

The short-baseline system consists of an array of four hydrophones mounted on a ship in two orthogonal baselines. The length of each baseline is the longest possible consistent with the ship's geometry. The hydrophones are mounted in wells and frequently can be extended below the ship's hull to reduce multipaths and ship's noise. Similar mounting is employed for the single transducer of the ultrashort-baseline system.

Since it employs time-of-arrival measurements rather than phase comparisons, the short-baseline system is less accurate than the ultrashort-baseline system, but it is usable at greater depths and over wider areas. Sometimes both systems are combined in a single installation.

The surface ship has one significant disadvantage: its acoustic equipment is located in the surface layers and is most subject to ship's noise and nonhomogeneities in the propagation medium. Therefore surface-ship tracking usually augments rather than replaces those systems carried by submersible vehicles.

Submarine navigation. Submarines must operate over wide areas of the ocean and often under highly secure and covert conditions; therefore, navigation

techniques which depend upon sonic or electromagnetic emissions, or which restrict movements of the vehicle to either near-bottom or near-surface regions, are judged to be too restrictive. The ability of inertial navigation systems to operate without frequent recourse to external position updates makes them prime candidates for submarine navigation. The most stringent navigation requirements for navigation accuracy are associated with missile-firing submarines. Contemporary sea-launched ballistic missiles (SLBMs) navigate using self-contained sensors (such as inertial sensors and star trackers). Inasmuch as they mechanize a dead-reckoning navigation solution their impact errors are sensitive to position and velocity errors at the time of launch. Improvements of other aspects of missile guidance systems have driven requirements to improve the ship's navigation system. The additional need for covert and secure long-duration navigation has resulted in the evolution of a navigation configuration that uses a pair of electrostatic gyroscope navigators (ESGNs) as the primary elements of an integrated navigation subsystem. The very stable, low-drift characteristics of the electrostatic gyroscope result in an inertial navigation system with a low and highly predictable error growth. Weapons system accuracy is further enhanced by velocity derived from both gyroscopes and secure correlation velocity sonar techniques to give a direct measure of the submarine ground speed. Vertical deflection maps, which are used for vertical-axis tilt compensation of the gyroscopes, are generated from combined satellite and oceanographic surveys of the Earth's gravity field. The relatively rare position resets are selected from either Global Positioning System satellite or bathymetric sonar data. The passive electromagnetic-log, which measures vehicle velocity relative to the water and is therefore subject to ocean current errors, is used primarily to damp the gyroscope computational (Schuler) oscillations. The various data are processed in a central navigation computer which uses statistical estimation algorithms. *See* ESTIMATION THEORY.

A class of inertial systems that was originally developed for commercial aircraft has found application in undersea vehicle navigation. The inertial components in these systems, both gyroscopes and accelerometers, are fixed to the vehicle frame in a strapped-down configuration. The mechanical gimbals and torque motors, which had traditionally isolated both gyroscopes and accelerometers from vehicle base motions, have been replaced with mathematical transformations implemented in a digital computer. Additionally, mechanical spinning-wheel gyroscopes have been replaced with the counter-rotating light beams of ring-laser gyroscopes. These solid-state configurations make extensive use of microprocessor and microminiature technology. The low power consumption, small volume, light weight, and low cost of ring-laser-gyroscope strapped-down systems makes them attractive for crewless undersea

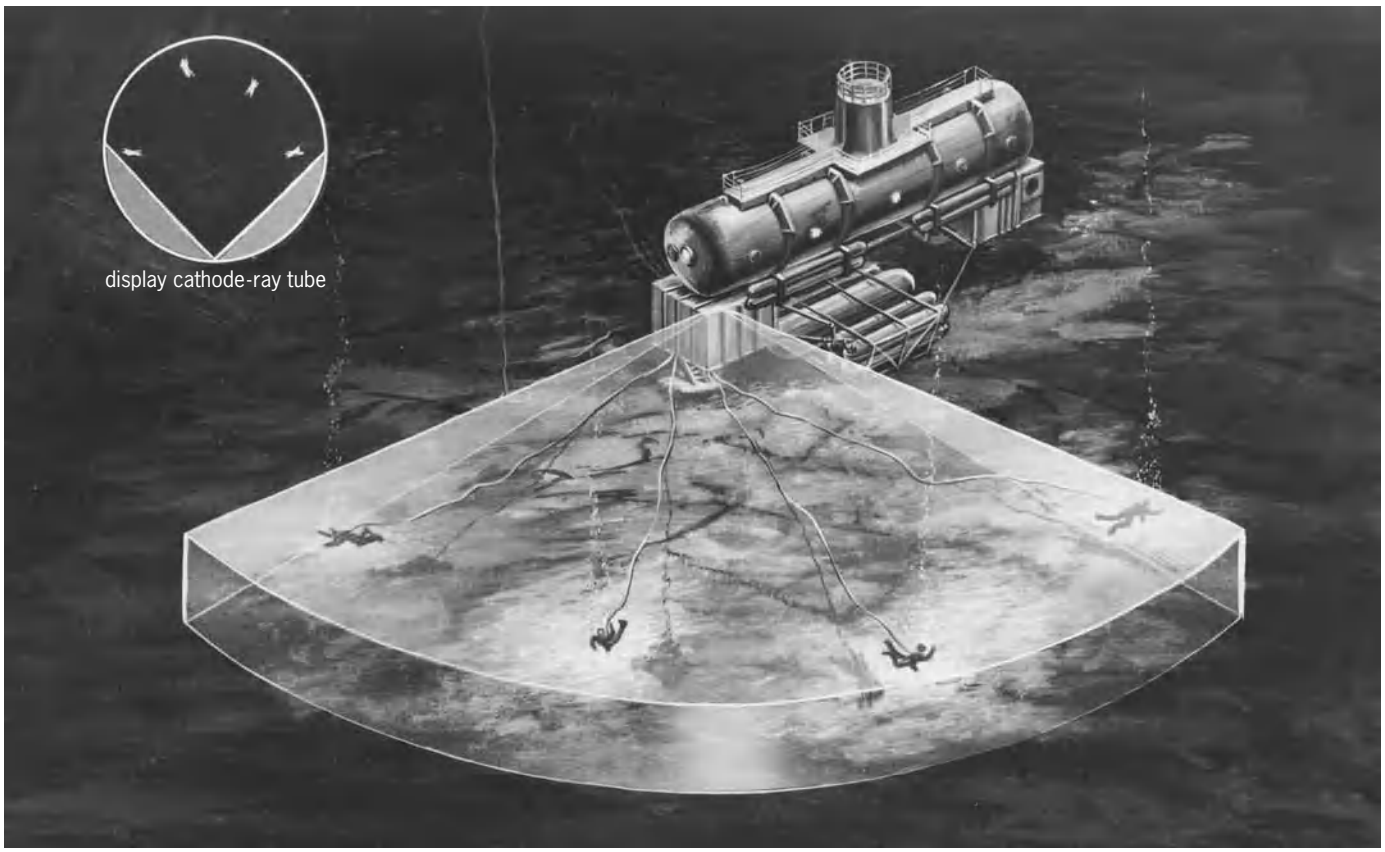

Fig. 5. Sonar for tracking of divers from an underwater habitat. Display cathode-ray tube shows positions of divers to diving officer.

vehicles, which have limited on-board battery-supplied power. Such vehicles do not have the long-duration requirements of nuclear submarine missions, and can therefore use less precise inertial navigators with external aids such as Doppler or correlation velocity sonar. *See* ACCELEROMETER; GYROSCOPE.

The increase in speed, capability, and capacity of microprocessors and computer memory, combined with corresponding decreases in size, power consumption, and cost has permitted UUVs to be given capabilities ordinarily reserved for larger crewed submarines. System reliability is approaching absolute, with increased use of fault-tolerant processor configurations and redundant sensors. The more sophisticated navigation system configurations make use of such computationally intensive techniques as bottom terrain matching, visual image processing, and object recognition. *See* FAULT-TOLERANT SYSTEMS; IMAGE PROCESSING; MICROPROCESSOR; SEMICONDUCTOR MEMORIES.

Diver navigation and tracking. Navigation has always been a difficult problem for divers. With the advent of saturation diving, undersea habitats, and transfer capsules to allow divers to work anywhere on the continental shelves, precise navigation is even more essential. A diver cannot surface from great depths without many hours or days of decompression. Therefore, before humans can be turned loose without restraining lines, there must be positive methods of diver navigation and tracking.

Some of the same sonars developed for submersible pilotage offer an answer for a diver. A horizontally swept constant-transmission frequency-modulated sonar similar to that previously described (**Fig. 5**) is well matched to the excursions of saturated divers, who are limited in allowable vertical movement. By communication with the habitat, divers who stray too far or lose their orientation can be guided back to safety.

Transponder systems similar to those described above have been employed to guide divers and uncrewed submersibles to and from work sites. The transponders, intended for operation at short ranges (up to 1000 ft or 300 m), transmit and receive in the range 100–400 kHz at low power levels. They are consequently much smaller and lighter in weight than typical deep-ocean navigation transponders (Fig. 2).

Charles J. Hrbek; Joseph A. Cestone; Emery St. George, Jr.; Richard L. Greenspan

Bibliography. W. E. Bradley, Evolution of integrated navigation for deep ocean surveying, *Oceans '90,* pp. 451–457, IEEE, 1990; J. A. Cestone et al., Latest highlights in acoustic underwater navigation, *Navigation,* 24(1):7–39, Spring 1977; D. E. DiMasa and W. K. Stewart Jr., Terrain-relative navigation for autonomous underwater vehicles, *Oceans '97*, pp. 541–546, IEEE, 1997; R. L. Greenspan, Inertial navigation technology from 1970–1995, *Navigation,* 42(1): 165–185, 1995; S. K. Hole, B. Woodward, and W. Forsythe, Design constraints and error analysis of the temporal correlation, *IEEE J. Oceanic Eng.*, 17(3): 269–279, 1992; E. Levinson and R. Majure, Accuracy enhancement techniques applied to the marine ring laser inertial navigator (MARLIN), *Navigation,* 34(1):64–86, Spring 1987; J. S. Stambaugh and R. B. Thibault, Navigation requirements for autonomous underwater vehicles, *Navigation,* 39(1):79–92, Spring 1992; J. A. Strickrott and S. Negahdaripour, On the development of an active vision system for 3-D scene reconstruction and analysis from underwater images, *Oceans '97*, pp. 626–633, IEEE, 1997; M. Uliana, F. Andreucci, and B. Papalia, The navigation system of an autonomous underwater vehicle for Antarctic exploration, *Oceans '97*, pp. 403-408, IEEE, 1997.

Underwater photography

The techniques involved in using photographic equipment underwater. Underwater photography is used to document subjects in fresh-water lakes and quarries, in temperate seas, beneath the polar ice pack, and in the deep sea, but by far the greatest percentage of underwater photography is done within sport-diving limits in the tropical oceans. The color and variety of marine life in the Caribbean Sea, Red Sea, or Indo-Pacific Ocean are diverse and compelling, creating an ever-changing portfolio of subjects to photograph.

Scientists use underwater photography to document issues, such as the health of the coral reefs, while photojournalists use underwater photography as a medium to report on the attributes of dive sites. However, the largest growth in underwater photography has come from recreational scuba divers.

Underwater photographers are faced with specific technical challenges. Water is 600 times denser than air and is predominantly blue in color. Depth affects light and creates physiological considerations for the photographer. As a result, underwater photography requires an understanding of certain principles of light beneath the sea.

Principles. As in all photography, consideration of the variables of light transmission is crucial to underwater photography. When sunlight strikes the surface of the sea, its quality and quantity change in several ways. *See* LIGHT; PHOTOGRAPHY.

Refraction. As light travels from air to a denser medium, such as water, the light rays are bent (refracted); one result is magnification of underwater objects by one-third as compared to viewing them in air. The magnification effect must be considered when estimating distances underwater, which is critical for both focus and exposure.

When using a camera equipped with a flat port (outer lens), the refractive index of the lens matches the perspective of the photographer when viewing the scene through the flat port of a facemask. Therefore, a subject that looks to be 3 ft (0.9 m) away to the photographer is actually 3 apparent feet from the lens. Apparent distances may be imprecise, and depend on a subjective guess at how distant a subject may be, but few subjects tend to hold still long enough for the distance to be measured. In this

example of 3 apparent feet, the actual measured distance would be 4 ft (1.2 m). Given the spontaneous nature of underwater photography, the popular convention has always been to work in apparent distances, and lenses are made specifically for underwater use and calibrated in apparent feet. This makes accurate distance estimation a valuable underwater photographic skill.

Using a dome port (correction lens) on the camera nullifies the effects of refraction and magnification and restores the true underwater distance. A dome port also creates a "virtual image," which typically exists a distance of about twice the diameter of the dome from the film plane. An 8-in. (20-cm) dome port will have a virtual image about 16 in. (40 cm) away, and the lens must be able to focus on this virtual image. With some lenses, a diopter (closeup lens) must be added to assure close-focus capability, for with a dome port, if the lens cannot focus on the virtual image, it will not be able to focus at all. *See* REFRACTION OF WAVES.

Absorption. Light is absorbed when it propagates through water. Variables affecting the level of light penetration include the time of day (affects the angle at which the sunlight strikes the surface of the water); cloud cover; clarity of the water; depth (light is increasingly absorbed with increasing depth); and surface conditions (if the sea is choppy, more light will be reflected off the surface and less light transmitted to the underwater scene).

Selective color filtration. Depth affects not only the quantity of light but also the quality of light. Once light passes from air to water, different wavelengths of its spectrum are absorbed as a function of the color of the water and depth. Even in the clearest tropical sea, water serves as a powerful cyan (blue-green) filter. Natural full-spectrum photographs can be taken only with available light in very shallow depths. In ideal daylight conditions and clear ocean water, photographic film fails to record red at about 15 ft (4.5 m) in depth. Orange disappears at 30 ft (9 m), yellow at 60 ft (18 m), green at 80 ft (24 m), and at greater depth only blue and black are recorded on film. To restore color, underwater photographers must use artificial light, typically from a submersible strobe (**Fig. 1**). *See* SEAWATER.

Water also absorbs and changes the quality of the light discharged by the strobe. Light is absorbed not only vertically (as a function of depth) but also horizontally (as a function of strobe-to-subject distance). For effective strobe photography, the photographer should be no more than 10 ft (3 m) from the subject, and ideally within 4 ft (1.2 m). Getting closer maximizes the color and resolution of the underwater photo, and thereby minimizes the number of lenses needed for underwater work. Telephoto lenses, for example, perform poorly underwater because their minimum focus distance puts too much water between the photographer and subject. *See* STROBOSCOPIC PHOTOGRAPHY.

Camera systems. The water column between photographer and subject degrades both the resolution of the image and the transmission of artificial light (necessary to restore color). Therefore, the most effective underwater photos are taken as close as possible to the subject, thereby creating the need for a variety of optical tools to capture subjects of various sizes within this narrow distance limitation. An above-water (topside) photographer might choose to shoot a full-figure photo of a person by using a 50-mm lens from 12 ft (3.6 m) away or by using a 24-mm wide-angle lens from 5 ft (1.5 m) away; the underwater photographer has fewer options. A subject 12 ft away will not reflect enough strobe light for an effective photo, so the photographer must use a wide-angle lens and work closer (**Fig. 2**). In this case, perhaps the best choice would be a 20-mm lens from 3 ft (0.9 m) away.

Fig. 1. Dolphin and underwater photographer—13-mm lens. The extreme wide angle of the 13-mm lens (170° diagonal coverage) was used here to capture a shot of an Atlantic bottlenose dolphin and diver. By using an extreme wide angle, it is possible to cover large subjects from a close distance, thereby maximizing color and resolution. (*Photograph by Stephen Frink*)

Since camera lenses cannot be interchanged underwater without flooding the camera, the photographer generally restricts the photographs to those appropriate to the lens. Alternatively, a photographer may carry several camera systems underwater, each dedicated to a specific focal-length lens. Increasingly, housing manufacturers are creating port options for using zoom lenses underwater. Two popular focal lengths are the 17- to 35-mm range (for wide-angle and reef scenics) and the 70- to 180-mm macro range (for fish portraits). Primary lenses still have a wider acceptance among photographers, but the quality of the zooms and the ergonomics of the new

Fig. 2. Over/under shot, Solomon Islands—16-mm lens. The over/under shot (one-half above the water and one-half below) is best done with a housed camera and dome port. An extreme wide-angle lens is typically used, although sometimes a split diopter is also used to account for the difference in refractive index and exposure values between the air and water. (*Photograph by Stephen Frink*)

generation of housings and ports make variable-focal-length zoom lenses a viable option.

There are two types of underwater cameras—amphibious and housed. Amphibious cameras may be used either underwater or topside, although some lenses are for underwater use only (known as water contact lenses) [**Fig. 3**]. A housed camera is a conventional above-water camera that has been protected from the damaging effects of seawater by a waterproof enclosure. The amphibious camera is protected by a series of O-rings, primarily located at the lens mount, film loading door, shutter release, and other places where controls are necessary. The O-rings make the system not only resistant to leaks but also impervious to dust or inclement weather when used above water.

The housed system is typically larger and bulkier than the amphibious system, but the trend in modern housings is to create a small, compact housing with optimal ergonomic placement of controls. A much wider range of lenses may be accommodated in underwater-housed systems, and custom ports are available specifically for the lenses. O-rings are also used in housings to prevent water intrusion. Some housings have mechanical levers that operate the important camera controls such as shutter release, film advance (although most housed cameras are equipped with motor drives), aperture, focus, and zoom control, although some housings now incorporate electronic controls for some functions. Clearly, the most popular camera for use with underwater housings is the single-lens reflex 35-mm, equipped

Fig. 3. Underwater photographer on a shipwreck—15-mm lens. Shipwrecks provide excellent subjects for wide-angle photography, and including a diver as an element of composition can add color to an otherwise monochromatic subject, as well as a sense of scale. (*Photograph by Stephen Frink*)

Fig. 4. Blue-cheeked butterflyfish, Red Sea—28-mm lens. Colorful tropical marine life is a fascinating subject for the underwater photographer. The challenge is to capture the composition in a natural, nonintrusive fashion. (*Photograph by Stephen Frink*)

Fig. 5. Lionfish, Red Sea—60-mm macro lens. When capturing closeup photographs of subjects that might be frightened by using framers (closeup kit or extension tubes), the housed single-lens reflex camera is the perfect tool. Not only can the photographer achieve precise composition by viewing through the lens, but also modern autofocus technology makes it easy to get a sharp picture. (*Photograph by Stephen Frink*)

with either a macro focusing lens (either a 60-mm or 105-mm macro) or a wide-angle lens (**Fig. 4**). Autofocus is one of the new technologies available in topside cameras that have contributed greatly to the ease of underwater photography.

Lenses. The standard lens for the amphibious camera is either a 35-mm or 28-mm, and at the optimal 3-ft (0.9-m) focus distance these lenses are typically used to photograph coral-reef marine life or head-and-shoulder views of divers. Conventional telephoto lenses are of little use underwater, but macro telephoto lenses (say, 105-mm or 200-mm) permit the photographer to capture small and typically skittish subjects from a greater distance, thereby not alarming them (**Fig. 5**).

For closeup documentation of small subjects, most amphibious systems provide an auxiliary diopter (supplementary lens) and focusing wand that works in conjunction with the normal lens. The diopter is placed in front of the lens, and the focus wand plus framer (an open frame that attaches to and extends from the lens, within which the subject is kept to ensure proper focus) shows what the lens sees. The advantage to this type of system is that it can be added and removed while underwater, so that if a larger fish swims by, the photographer has the option to remove the closeup kit and shoot. Another option (for typically greater magnification) is the use of an extension tube. This machined aluminum tube is inserted between the lens and the camera body (O-ring sealed), thereby changing the focal length of the lens to a magnification ratio of one-third life size (1:3) to twice life size (2:1), depending on which combination of tube and framer is used. A housed camera may use continually focusing macro lenses (closeup) to capture subjects with the standard range, but also with magnification ratios of 1:1.

In order to photograph large underwater subjects and panoramas, a wide-angle lens is necessary (**Fig. 6**). For both amphibious and housed systems, that means an angle of coverage ranging from 94° (equivalent to a 20-mm lens topside) to 180° (equivalent to a 16-mm full frame fisheye lens). The advantage to these optics is that the photographer can be within 3 ft (0.9 m) of a massive subject, yet still gain the color from the artificial light and enhance the resolution by minimizing the water column.

Fig. 6. Caribbean reef shark and diver—16-mm lens. The extraordinary depth of field of the full-frame fisheye lens is apparent here, as both the shark and the diver are in sharp focus. A little perspective distortion is evident: because the shark is nearer to the lens relative to the diver, it seems even larger and more impressive. (*Photograph by Stephen Frink*)

Modern amphibious cameras are constructed of materials that resist corrosion in salt water, such as anodized aluminum or marine-grade plastics for the body, while the best optics are made of multicoated glass designed for water contact. The interface between the glass and seawater are part of the optical formula with many amphibious lenses, particularly the wide-angle lenses, and they do not work well above water as a result. Housings are typically constructed of either anodized aluminum or polycarbonate, both of which are strong and corrosion-resistant.

One advantage of the housed system is that an existing camera system may be modified for underwater use, as opposed to buying a camera and lens system specifically dedicated to underwater photography. Since most cameras used in housings are single-lens reflex and autofocus, the level of sophistication in housed systems is quite high. Some of the existing viewfinder-type amphibious cameras are a bit more simplistic, and require the photographer to compose through a relatively imprecise viewfinder. Housings typically have some means of enlarging the single-lens reflex view for viewing through a facemask.

The primary advantages of the amphibious camera are its compact size (as compared with the housing) and ease of operation. The range of accessory lenses available for amphibious cameras tends to be extensive, covering most underwater imaging opportunities. Once familiar with the amphibious system, the photographer can usually react quickly and accurately to capture the underwater scene. There are certain small subjects that may be too skittish to capture with framers, but other than that, most of the subjects that a photographer could capture with a housing could also be captured with the amphibious camera. There are trade-offs between the ease of operation and small size of the amphibious system, as contrasted with the compositional accuracy of the single-lens reflex in a housing; yet successful photographs are commonly taken with both systems.

Stephen Frink

Deep-sea underwater photography. Photography at depths beyond the limits of sport diving—approximately 150 ft (35 m)—requires the design and use of special camera and lighting equipment. Watertight cases are required for both camera and light source, and they must be able to withstand the pressure generated by the sea. For each 33 ft (10 m) of depth, approximately one additional atmosphere ($\sim 10^2$ kilopascals) of pressure is exerted. At the greatest ocean depths, about 40,000 ft (12,000 m), a case must be able to withstand 17,600 lb/in.^{2} (1200 kg/cm^2). The windows for the lens and electrical seals must also be designed for such pressure to prevent water intrusion. Examples of such deep-sea work are the bathypteroid fish photographed at 4000 ft (1200 m) in the Gulf of Mexico (**Fig. 7**) and a view of the Romanche Trench at a depth of 25,000 ft (7625 m; **Fig. 8**).

Cylindrical cases are most common since they are stronger than rectangular ones. The cameras are typically adapted to accept the standard 100-ft (30-m) motion picture spool of 35-mm film that allows 800 individual 35-mm frames to be exposed. Thin-base cronar film permits even more exposures with cameras that can accept 400 ft (122 m) of this film type. This length allows up to 3200 photographs to be taken before changing film, an important consideration since much time is consumed in raising and lowering the camera gear from the sea.

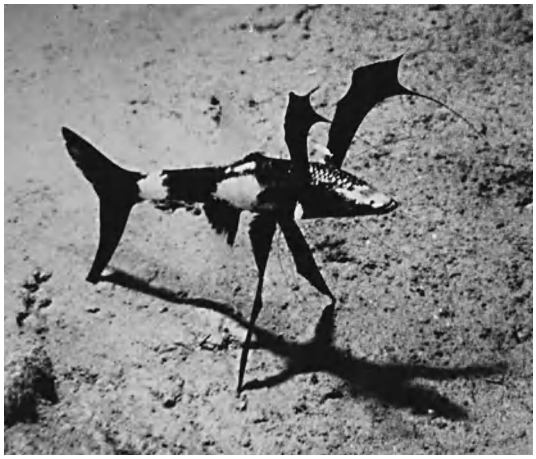

Fig. 7. Bathypteroid fish photographed from the Westinghouse diving vehicle *Deepstar* at 4000 ft (1200 m) depth in the Gulf of Mexico. Tungsten lighting was used. (*Photograph by R. Church*)

Auxiliary lighting is required, since daylight is absorbed in both intensity and hue. The electronic flash utilizing a xenon tube is almost universally used for deep-sea photographs since it is efficient and compatible in color temperature with daylight film. Corrected optics for the camera may be used behind the lens port to compensate for the different refractive index of air and water.

The camera must be positioned and triggered to render the desired photograph, and the great depths preclude a free-swimming human operator. Operation is often from a cable via sonar sensing equipment

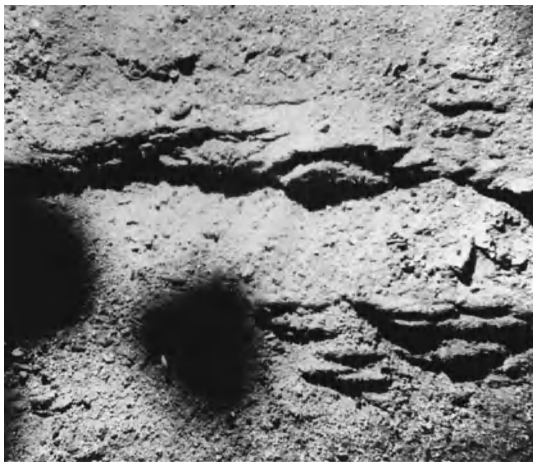

Fig. 8. Photograph taken (August 1956) in Romanche Trench, 0°10′S, 18°21′W, at depth of 25,000 ft (7625 m). A nylon cable was used from the deck of the *Calypso*. (*National Geographic Magazine, March 1958*)

or from deep-diving underwater vehicles. Bottom-sensing switches can operate deep-sea cameras for photographing the sea floor, and remotely operated vehicles (ROVs) can incorporate both video and still cameras. The ROV is tethered to either the surface or the submersible vehicle by an umbilical cable, and it is moved by its own propellers. The operator can direct the movement of the ROV, view the scene through the video camera, and record appropriate still images when the ROV is properly positioned. *See* SONAR; UNDERWATER VEHICLES.

Many deep-sea cameras have an internal data chamber that is photographed simultaneously with the deep-sea exposure. Generally, photographic digital data of 16 digits can record such information as the second, minute, and hour that the photograph was taken, as well as the compass orientation of the photograph and camera-to-subject distance. The camera-to-subject distance is fed into the camera from a sonar sensor mounted on the camera. When the camera-to-subject distance and focal length of the lens are known, it is possible to calculate the field of view for each photograph. Thus, the size of the objects in the photograph may be measured. When the data and the image on the film are inseparable because they are rendered with a simultaneous exposure, the chances of error are greatly reduced.

When an observer descends to great depths in a diving vehicle, the camera can assist in documentation by recording what is seen. Furthermore, the visual data will assist in accurate description of the observed phenomena. Elapsed-time photography with a motion picture camera in the sea is important in studying sedimentation deposits caused by tides, currents, and storms. Similarly, the observation of biological activity taken with the elapsed-time camera and then speeded up for viewing may reveal processes that cannot ordinarily be observed. *See* DIVING; PHOTOGRAPHY; UNDERWATER TELEVISION.

Harold E. Edgerton; Stephen Frink

Bibliography. J. Church, *Jim Church's Essential Guide to Composition*, 1998; J. Church, *Jim Church's Essential Guide to Nikonos Systems*, 1998; M. Edge and I. Turner (eds.), *The Underwater Photographer*, 1999; S. Frink, *Wonders of the Reef: Diving with a Camera*, 1996; A. Kohler and D. Kohler, *The Underwater Photography Handbook*, 1999; J. Liburdi and C. Sherman, *The New Guide to Sea & Sea*, 1998; M. Webster and G. Bradford (illustrator), *The Art & Technique of Underwater Photography*, 1999; N. Wu, *How To Photograph Underwater* (*How To Photograph*, vol. 2), 1994.

Underwater sound

The production, propagation, reflection, scattering, and reception of sound in seawater. The sea covers approximately 75% of the Earth's surface. In terms of exploration, visible observation of the sea is limited due to the high attenuation of light, and radar has very poor penetrability into salt water. Because of the extraordinary properties that sound has in the sea, and because of some of the inherent characteristics of the sea, acoustics is the principal means by which the sea has been explored. *See* OCEAN.

Absorption. Sound has a remarkably low loss of energy in seawater, and it is that property above all others that allows it to be used in research and application. Absorption is the loss of energy due to internal causes, such as viscosity. Over the frequency range from about 100 Hz (cycles per second) to 100 kHz, absorption is dominated by the reactions of two molecules, magnesium sulfate ($MgSO_4$) and boric acid [$B(OH)_3$]. These molecules are normally in equilibrium with their ionic constituents. The pressure variation caused by an acoustic wave changes the ionic balance and, during the passage of the pressure-varying acoustic field, it cannot return to the same equilibrium, and energy is given up. This is called chemical relaxation. At about 65 kHz magnesium sulfate dominates absorption, and boric acid is important near 1 kHz. Absorption has been measured in the laboratory and at sea (**Fig. 1**). The absorption coefficient α is the exponential loss in intensity of an acoustic signal of a given frequency per meter of its path. The ratio of intensities after traversing a distance r is given by Eq. (1). Often absorption

$$I/I_0 = e^{-\alpha r} \tag{1}$$

is defined as the power to the base 10, that is, $I/I_0 = 10^{-\alpha r}$. Due to the very low absorption coefficients

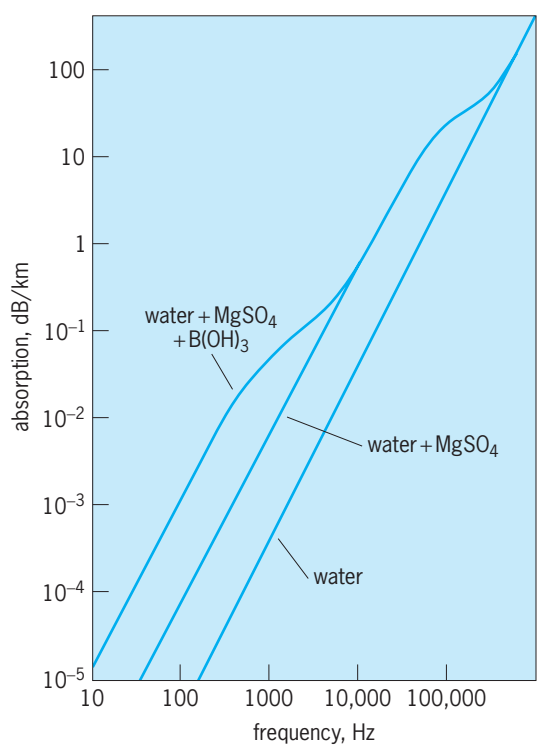

Fig. 1. Sound absorption coefficient as a function of frequency for seawater. For this case, the salinity is 35‰, temperature 4°C (39°F), pressure 100 kPa (approximately atmospheric pressure), and pH 8. Note the effect of the presence of $MgSO_4$ and $B(OH)_3$. (*After F. H. Fisher and V. P. Simmons, Sound absorption in sea water, J. Acous. Soc. Amer., 62:558–564, 1977*)

in seawater, the convention is to change its units to km^{-1}. *See* SEAWATER; SOUND; SOUND ABSORPTION.

Both magnesium sulfate and boric acid are present in very small amounts in seawater. If their concentrations were a few orders of magnitude larger, sound measurement would be as ineffective as it is with radar or optics.

Sound speed. The speed of sound in seawater and its dependence on the parameters of the sea, such as temperature, salinity, and density, have an enormous effect on acoustics in the sea. There have been several empirical equations published in the literature, differing by small amounts, but for the purposes here, the differences are unimportant. To the lowest order in temperature, salinity, and pressure is Eq. (2),

$$C = 1449.2 + 4.6T + 1.34(S - 35) + 0.016Z \quad (2)$$

where the speed of sound in seawater (C) is in meters per second, temperature (T) is in degrees Celsius, salinity (S) is in parts per thousand (‰), and depth (Z) is in meters.

Ocean environment. Generally the environmental parameter that dominates acoustic processes in oceans is the temperature, because it varies both spatially and temporally. The dynamics of the sea are very complex, and one driven by solar heating, winds, atmospheric forces, bathymetry, major ocean currents, turbulence, and the Earth's rotation. It is impossible to include all of these effects in the equations of acoustics. But by limiting the theory to the most salient features, and through reasonable approximations, some solutions can be obtained that are in agreement with experimental results. When agreement is not satisfactory, either the acoustic theory or the ocean model must be modified.

Solar heating of the upper ocean has one of the most important effects on sound propagation. As the temperature of the upper ocean increases, so does the sound speed. Winds mix the upper layer, giving rise to a layer of water of approximately constant temperature, below which is a region called the thermocline. Below that, most seawater reaches a constant temperature. All these layers depend on the season and the geographical location, and there is considerable local variation, depending on winds, cloud cover, atmospheric stability, and so on. Shallow water is even more variable due to tides, fresh-water mixing, and interactions with the sea floor. Major ocean currents, such as the Gulf Stream and Kuroshio, have major effects on acoustics. The cold and warm eddies that are spun off from these currents are present in abundance and significantly affect acoustic propagation. *See* GULF STREAM; KUROSHIO; OCEANOGRAPHY.

Units. The science of underwater sound is the study of pressure waves in the sea over the frequency range from a few hertz to a few megahertz. The International System (SI) units are the pascal (Pa) for pressure (equal to one newton per square meter) and the watt per square meter (W/m^2) for sound intensity (the flow of energy through a unit area normal to the direction of wave propagation). In acoustics, it is more convenient to refer to pressures, which are usually much smaller than a pascal, and the consequent intensities with a different reference, the decibel. Intensity in decibels (dB) is ten times the logarithm to the base ten of the measured intensity divided by a reference intensity. The unit of reference intensity is the intensity of a plane wave that has a root-mean-square sound pressure of one micropascal (1 μPa). In the past, several other reference units have been used. *See* DECIBEL; SOUND INTENSITY; SOUND PRESSURE.

Wave propagation. The mathematical equation that sound obeys is known as the wave equation. Its derivation is based on the mathematical statements of Newton's second law for fluids (the Navier-Stokes equation), the equation of continuity (which essentially states that when a fluid is compressed, its mass is conserved), and a law of compression, relating a change of volume to a change in pressure. By the mathematical manipulation of these three equations, and the assumption that only very small physical changes in the fluid are taking place, it is possible to obtain a single differential equation that connects the acoustic pressure changes in time to those in space by a single quantity, the square of the sound speed (c), which is usually a slowly varying function of both space and time. This known as the wave equation [Eq. (3)], where p is the acoustic pressure; and

$$\nabla^2 p = c^{-2} \frac{\partial^2 p}{\partial t^2} \quad (3)$$

the differential operator ∇^2, known in mathematics as the laplacian, is, in many underwater acoustics applications, given in either spherical or cylindrical coordinates. *See* FLUID-FLOW PRINCIPLES; LAPLACIAN; NAVIER-STOKES EQUATION; WAVE EQUATION.

Knowing the sound speed as a function of space and time allows for the investigation of the spatial and temporal properties of sound, at least in principle. The mathematics used to find solutions to the wave equation are the same as those that are used in other fields of physics, such as optics, radar, and seismics. *See* WAVE MOTION.

In addition to knowing the speed of sound, it is necessary to know the location and nature of the sources of sound, the location and features of the sea surface, the depth to the sea floor, and, in many applications, the physical structure of the sea floor. It is not possible to know the sound speed throughout the water column or know the boundaries exactly. Thus the solutions to the wave equation are never exact representations of nature, but estimates, with an accuracy that depends on both the quality of the knowledge of the environment and the degree to which the mathematical or numerical solutions to the wave equation represent the actual physical situation.

There are a variety of analytical and numerical solutions to the wave equation that are useful. An important parameter in selecting the most appropriate solution is the ratio of acoustic wavelength to some of the characteristic dimensions of the environment.

The wavelength λ of a sound wave of frequency f is Eq. (4). Since the speed of sound is approximately

$$\lambda = c/f \tag{4}$$

1500 m/s (5000 ft/s), typical wavelengths for 100, 1000, 10,000 and 100,000 Hz are 15, 1.5, 0.15, and 0.015 m (50, 5, 0.5, and 0.05 ft), respectively.

If the wavelength is very small compared to the dimensions of oceanographic interest (such as depth, distance, sea floor, or sea surface roughness), good approximate solutions are those of geometric optics, known as ray theory. Rays that originate at the source follow paths that are determined by Snell's law. According to Snell's law, if a ray is launched from a source where the sound speed is c_0 at an angle θ_0 from the horizontal axis, the horizontal angle θ of the ray will, at any point along its path, obey Eq. (5),

$$\cos\theta/c = \cos\theta_0/c_0 = \text{a constant} \tag{5}$$

where c is the speed of sound at a point along the ray path. This statement allows ray paths to be determined for each angle θ_0 by either analytical or numerical computation. A simple example is shown in **Fig. 2**. As will be seen later, this limit allows for good insight into the spreading and the distribution of acoustic energy. *See* GEOMETRICAL OPTICS; REFRACTION OF WAVES.

The bending of rays is called refraction. The case shown in Fig. 2 is for a linear sound speed profile. It can be shown for this case that rays will follow circular paths. There will be a limiting ray, as shown, beyond which no ray will exist. The region beyond that ray is known as the shadow zone, and is well known to limit the detectability of objects by sonars. Other, more elaborate solutions to the wave equation are required when geometric distances become comparable to wavelength, such as a sound wave near a boundary or object, or when multiple waves are present. *See* SONAR.

The interaction of sound waves with the boundaries of the sea is essential for understanding underwater sound. By examining a highly idealized model, it is possible to gain a valuable understanding of some important properties of the interaction of acoustic fields with the sea surface and sea floor. It is useful to consider the case of a plane sound wave with a frequency f. The second-order derivative with respect to time on the right-hand side of the wave equation (3) yields for that term $c^{-2}\partial^2 p/\partial t^2 = -4\pi^2 f^2 c^{-2} p$. Substituting this into Eq. (3) gives Eq. (6), which is

$$(\nabla^2 + k^2)p = 0 \tag{6}$$

known as the Helmholtz equation. The quantity k is called the wave number, defined to be $k = 2\pi f/c$. If, as shown in **Fig. 3**, the acoustic wave is directed toward a boundary at an angle θ from the horizontal axis, x, the wave can be represented by Eq. (7),

$$p_i = A\sin(kx\cos\theta - ky\sin\theta - 2\pi f t) \tag{7}$$

where A is its acoustic pressure amplitude. There will also be a reflected wave, given by Eq. (8), whose

$$p_r = B\sin(kx\cos\theta + ky\sin\theta - 2\pi f t) \tag{8}$$

acoustic pressure amplitude is B. If the medium on the other side of the boundary is also a fluid, but with a sound speed of c_0 and a wave number k_0, a good "guess" at the solution for a wave traveling away from the boundary, as shown in Fig. 3, is given by Eq. (9), where D is the acoustic pressure amplitude.

$$p_t = D\sin(k_0 x\cos\theta_0 - k_0 y\sin\theta_0 - 2\pi f t) \tag{9}$$

At the boundary, $y = 0$, there are two conditions that have to be satisfied. The pressure on both sides has to be equal, and the vertical component of the fluid displacements has to be continuous, for all values of x. It is therefore required that $k\cos\theta = k_0\cos\theta_0$, or, remembering the definitions of k and k_0, $\cos\theta/c = \cos\theta_0/c_0$, which is identical with the Snell's law of Eq. (3). Returning to the conditions that have to be met at $y = 0$, from the continuity of pressure,

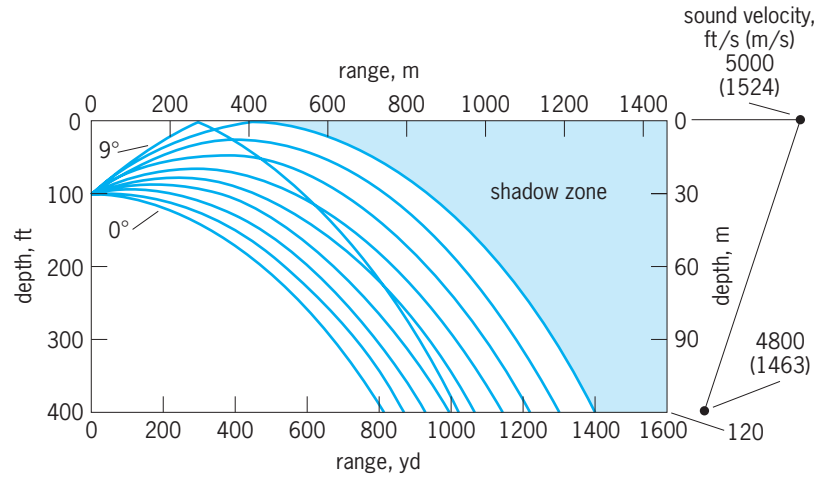

Fig. 2. Ray paths for a source near the surface of an ocean with a linear negative sound speed gradient, shown at the right. For this case, all rays follow the arcs of circles with different radii for each angle of emission. (*After R. J. Urick, Principles of Underwater Sound, 3d ed., McGraw-Hill, 1983, reprint, Peninsula Publishing, 1996*)

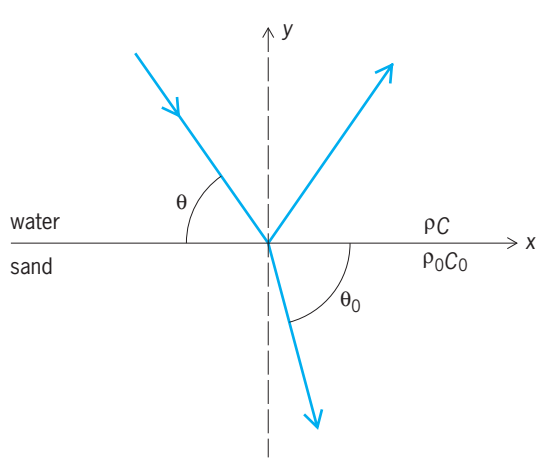

Fig. 3. Idealized model of reflection and transmission at a flat boundary of two fluids.

Eq. (10) must be satisfied.

$$A + B = D \tag{10}$$

A fluid's vertical velocity, u_y, can be shown to obey Eq. (11), where $\partial p/\partial y$ is the local vertical pressure

$$\frac{\rho \partial u_y}{\partial t} = \frac{\partial p}{\partial y} \tag{11}$$

gradient, and the density is ρ. In the upper fluid, at the boundary the total pressure is $p = p_i + p_r$. Taking the time derivative of u_y in both fluids to be equal at $y = 0$ leads to Eq. (12). Dividing Eq. (12) by Eq. (10) leads to Eq. (13) for B, where $Z = \rho_0 c_0/\rho c$.

$$\rho^{-1} k \sin\theta\{-A + B\} = -\rho_0^{-1} k_0 \sin\theta_0 D \tag{12}$$

$$B = \frac{A(Z\sin\theta - \sin\theta_0)}{Z\sin\theta + \sin\theta_0} \tag{13}$$

In the case of reflection from the air-sea surface, typical approximate values for the densities and sound speed for air and water, respectively, are 1.25 kg/m^3 (0.078 lb/ft^3) and 300 m/s (1000 ft/s), and 1000 kg/m^3 (62 lb/ft^3) and 1500 m/s (5000 ft/s). Therefore, Z is approximately 2.5×10^{-4}. It can also be shown from Snell's law that $\sin\theta_0$ will, for this case, always be a number close to 1. So the amplitude of the reflected wave will be almost exactly that of the incoming wave, only having the opposite sign. The sum of the amplitudes at the boundary of the incoming wave, p_i, and the reflected wave, p_r, is very close to zero. This boundary is called a pressure release boundary and is approximated as $p = 0$. Since $A \approx B$, it is also evident from Eq. (10) that very little energy is transmitted to the air.

A typical sandy sea floor has a sound speed approximately 10% higher and a density about twice that of seawater. From Snell's law it can be seen that at the angle $\theta = \arccos(c/c_0) \cong 25^\circ$ the direction of the refracted wave in the sea floor is parallel to the boundary. At that angle, it can be seen that $B = A$. That is, all of the acoustic energy is reflected. For angles $\theta < 25^\circ$, values of $\sin\theta_0$ become purely imaginary and B can be shown to be equal to A times a term that only represents a phase change. The term B/A is known as the reflection coefficient, and the angle $\theta_c = \arccos(c/c_0)$ is called the critical angle. This concept of total reflection is essential to the understanding of propagation in shallow water. It can be shown that the reflection is a minimum at normal incidence, implying that this is also the angle of maximum penetration into the sea floor and therefore important in the use of acoustics in the study of the internal structure of the seabed.

Although these two idealized examples do not represent the actual conditions of the sea, they do contain some of the physics that is observed. The sea surface is rarely flat, but it is always considered to be a pressure release surface. The sea floor is never a perfect fluid, and more realistic model extensions include absorption, viscosity, shear, and layering. The theories follow the same rules, namely, finding solutions in each medium and matching the boundary conditions of continuous pressure and displacement. *See* REFLECTION OF SOUND.

Deep-water propagation. During the middle years of the twentieth century, it was discovered that low-frequency sound, under some conditions, could propagate over very long distances with unusually small losses in intensity. Indeed, explosive signals were heard halfway around the Earth. The concept can be understood through the application of ray theory in deep water. Since the deep oceans are several thousand meters deep and major oceanographic features have dimensions many times larger than acoustic wavelengths, the use of geometric optics, or ray theory, is appropriate.

In many parts of the world, warm water overlays the deeper cold water. Sound speed near the surface can be dominated by the temperature, which is highest at or near the surface, and then decreases with depth until it becomes constant, in most oceans somewhere near 4°C (39°F). From Eq. (2) it can be seen that the effect of temperature is to decrease sound speed with depth, and the pressure term ($0.016Z$) causes an increase with depth. The sound speed profile as a function of depth (due to the temperature and depth) will have a minimum that is usually found to be at about 1000 m (3300 ft). An idealized picture is shown in **Fig. 4**.

For an impulsive omnidirectional sound source (typical of an explosive) that is located in depth at the minimum sound speed (Fig. 4*a*), rays are emitted in all directions and each will follow a trajectory

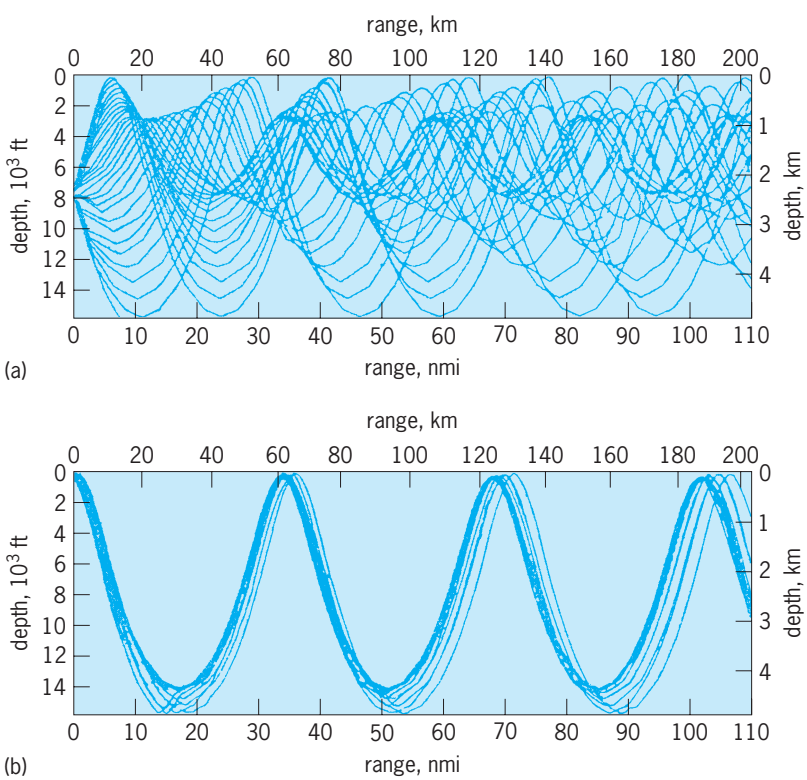

Fig. 4. Ray trace diagrams for deep-water propagation (*a*) for a source near the sound channel axis and (*b*) for a source near the surface—note the large region of shadow zones.

dictated by Snell's law. Those that are directed upward will gradually bend downward as they approach water with higher speeds. Those with steep enough trajectories will strike the sea surface. There will be a limiting ray that reaches the surface at zero degrees. Its angle of emission from the source, using Snell's law, is given by Eq. (14).

$$\theta_L = \arccos\,(c_{\min}/c_0) \tag{14}$$

All rays with smaller angles of emission will not reach the surface, and will bend toward the axis of minimum speed, intersecting the axis with the equal but opposite sign to that at which they were emitted. If the sound speed at the surface is less than that at the bottom, these rays will not reach the bottom as shown in Fig. 4. The same is true for the down-going rays. There will be a limiting ray with an angle again dictated by Snell's law, and rays that do not interact with either the surface or the sea floor and propagate outward from the source cylindrically rather than spherically.

The path that each ray will follow can be determined using Snell's law and the sound speed profile. (There are many computational methods available for this.) For a given range, it is possible to determine the family of rays that can be emitted from the source and pass through that point. It is also possible, knowing the ray paths, to compute the arrival time of each. It can be shown that the rays that arrive first are those that make the largest excursions from the channel axis, and the last to arrive is the ray that travels along the channel axis, which will be the most intense. The signal that is received on the channel axis from a distant explosive source begins with weak arrivals from the rays making the widest excursions and ends with a dramatic buildup that suddenly cuts off with the arrival of the ray that travels along the channel axis. The arrival of this last signal can give an accurate determination of the range from the source to the receiver, and is the basis for the SOFAR concept that was used for the location of downed aircraft in earlier times. *See* SOFAR.

For a sound source located near the surface, the rays that propagate without interaction with the surface or sea floor are more limited and leave large regions where sound does not penetrate (Fig. 4*b*). These shadow zones are vast, and it is only after large distances that the rays arrive near the surface. This sinuous behavior of the sound energy continues for very long distances with cyclic intervals of approximately 60–70 km (37–44 mi). The areas where there is focusing near the surface are called convergent zones. These are important for the detection of distant signals.

Of interest to the oceanographers and acousticians is the application of long-range, deep-water propagation to the study of ocean structure, such as warm and cold water eddies, the meander of ocean currents, internal waves, and global warming. For a fixed geometry of sources and receivers there will be specific ray paths, each migrating through the upper and lower depths. The arrival time of each ray is determined by the path and the sound speed along it. Experiments have shown remarkable temporal stability in many of the paths By observing the angular direction and arrival times, the paths of the arrivals can be identified. The measurement of small changes in the arrival times can be used to interpret oceanographic change. It is a form of tomography. *See* COMPUTERIZED TOMOGRAPHY.

There are many other mathematical solutions to the wave equation that are used in deep-water application. In order to obtain tractable solutions, many of these require assumptions about the homogeneity of the ocean. Of particular merit is the application of the parabolic approximation to long-range propagation for the case where there is a horizontal variation in oceanographic conditions. In most ocean applications, it is a valid to assume that the horizontal change in sound speed is sufficiently gradual to allow the range-dependent part of the Helmholtz equation to be approximated by a parabolic equation. The method is highly amenable to modern computation, and has been applied to many problems in acoustics, including shallow-water and atmospheric acoustics. *See* DIFFERENTIAL EQUATION.

Shallow-water propagation. Shallow water can be defined as the regions of the oceans associated with the continental shelves. Shallow depths are of the order of a few hundred meters or less. Unlike deep water, the sea floor plays a major role in the acoustic properties in shallow water. In shallow water, the application of higher frequencies is more suitable for both the detection of objects and for the study of some of the smaller-scale oceanographic processes. For lower frequencies, propagation over longer ranges requires the consideration of the many acoustic paths contributed by the continued interaction with the boundaries. Sound that reflects off the sea surface loses little energy. The reflections off the sea floor when the grazing angle is greater than critical are also nearly perfectly reflecting. Thus the sound field at longer ranges comprises many arrivals.

The principal method of solving the wave equation in shallow water is that of normal modes. The approach is to make the assumption that sound propagation in the ocean has no azimuthal dependence. Then the wave equation, or the Helmholtz equation, depends only on range and depth, and is separable into an equation of depth and one of range. The sound speed is assumed to have no range dependence, but it can depend on depth.

The range-dependent equation is the Bessel equation, whose solutions are well known. The vertical, or depth-dependent equation, is then solved. It will depend on the sound speed profile, the water depth, and the boundary conditions at the sea surface (the pressure release condition) and the sea floor. This leads to a set of solutions, called modes. Then, by representing the source, often as a point source at a depth z_0, the solution is obtained for any point in the water column by a sum of modes. Range dependence is difficult to accommodate to this method, but the normal-mode method is a very powerful

tool for understanding the complex way that signals travel through shallow water. Signals become distorted from their original form due to the changes in horizontal speed that result from the complex interaction of sound with the boundaries. The frequency dependence of sound speed is called dispersion. *See* BESSEL FUNCTIONS.

Ambient noise. A consequence of the remarkable transmission of sound is that unwanted sounds are transmitted just as efficiently. One of the ultimate limitations to the use of underwater sound is the ability to detect a signal above the noise. In the ocean, there are four distinct categories of ambient sound: biological, oceanographic physical processes, seismic, and anthropogenic. Seismic is in the frequency range called infrasound, and will not be covered. *See* INFRASOUND.

Biological sources. Biological sound is caused by both fish and marine mammals for the purpose of either communication or predation. The frequency range is broad. The sounds are sometimes very intense and geographically variable. At the low frequencies (those that would be called audible), the sounds of whales and other mammals dominate. Many of the species are migratory, giving rise to very high sound levels at certain times of the year at specific locations. There are currently intense studies of these creatures with the objective of assessing their numbers, their habits, and the purposes of their communications. At the higher frequencies, about 1–20 kHz, snapping shrimp dominate the sound in shallow water, especially around piers and other structures. From an engineering viewpoint, it is difficult to conceive of methods, other than avoidance, to process against these noises due to their high variability in space and time.

Physical processes. The natural physical causes of ambient noise are the sounds of breaking waves, surf, and rain noise.

There have been only a few studies of the physical mechanisms that create the sound. In the few measurements that have been made on individual breakers, there appears to be a wide variation in the frequency content. A plausible hypothesis is that the collapsing of the bubble cloud produced by the breaking wave generates a band of acoustic frequencies. This, however, has not been confirmed. Because winds are the general cause of breaking waves, this type of noise is referred to as wind-generated. It dominates the noise spectrum above about 500 Hz, but is a contribution down to at least 100 Hz. Wind-generated noise increases with wind speed, with the difference in level being about 30 dB between a very calm sea (sea state 1) and a violent sea (sea state 6). *See* OCEAN WAVES.

Anthropogenic sound. Shipping dominates the ambient sound in most oceans in the frequencies of about 10–150 Hz. The level clearly depends on the amount and type of shipping and the propagation conditions. The radiated noise from each class of ship is different, with the loudest being supertankers. Another anthropogenic sound, from oil prospecting, is seasonal and geographically limited, but the airguns and sparkers used have very high source levels.

Figure 5 shows an averaged compilation of typical sound levels that are due to anthropogenic and natural physical sounds. *See* ACOUSTIC NOISE.

Scattering and reverberation. The other source of unwanted sound is reverberation. Sound that is transmitted inevitably finds something to scatter from in the water column, at the sea surface, or at the sea floor. The scatter is usually in all directions, and some of it will return to the system that processes the return signals. Sources of scattering in the water column are fish, particulates, and physical inhomogeneities. The sea surface is, under normal sea conditions, agitated by winds and has the characteristic roughness associated with the prevailing

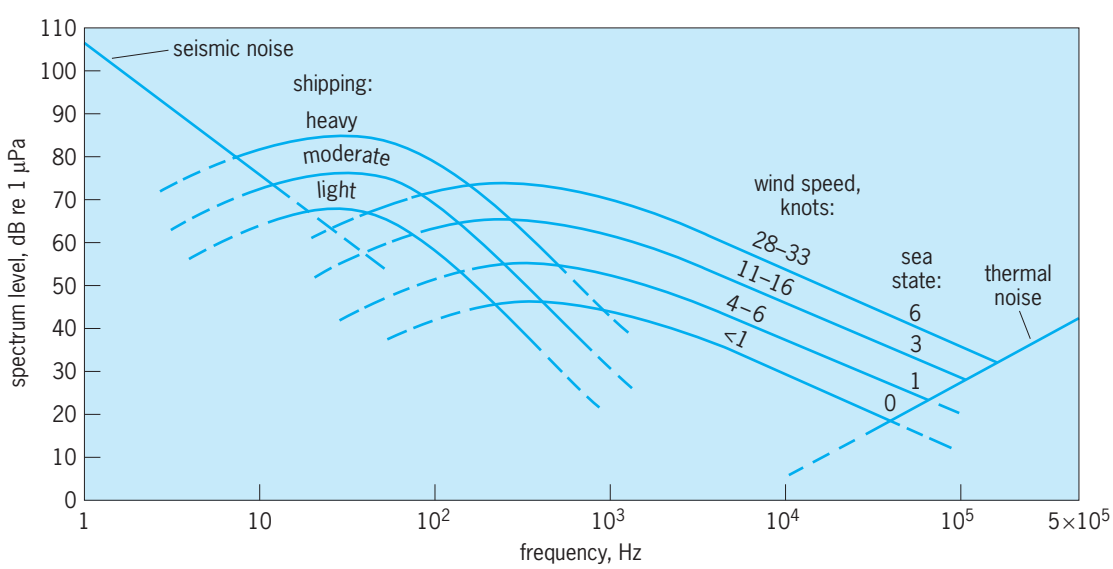

Fig. 5. Average values of ambient noise for both shipping and wind-generated sound, and representative values of seismic and thermal noise. 1 knot = 0.5 m/s. (*After R. J. Urick, Principles of Underwater Sound, 3d ed., McGraw-Hill, 1983, reprint, Peninsula Publishing, 1996*)

atmospheric conditions. Rough surfaces scatter sound with scattering strengths that depend on the roughness, the acoustic frequency (or wavelength), and the direction of the signal. The scattering is highly time-dependent, and needs to be studied with an appropriate statistical approach. The sea floor has inherent roughness and is usually inhomogeneous, both properties causing scatter. Although scatter degrades the performance of sonars, the characteristics of the return can be determined to enable its cancellation through signal processing or array design. Scattering can also be used to study the sea surface, the sea floor, fish types and distribution, and inhomogeneities in the water column. *See* SCATTERING LAYER.

There are still fundamental questions about the scattering of sound from the sea surface. At higher wind speeds, the classic theory of scattering from the sea surface underestimates the scattering strength by several decibels, enough to question either its validity or applicability. There is evidence that the layer of bubbles caused by the breaking of waves plays a significant role. Conclusive experimental proof of this hypothesis has not been obtained. The sea floor needs more study to determine the causes of scattering. That, however, may not yield a universal model because of the widely varying nature of the sea floor.

Experiments and ships. Both oceanography and ocean acoustics (underwater sound) have gained enormously from the developments in electronics, miniaturization, materials, and digital processing. It is now possible to place measuring instruments in the sea: on a buoy (either free-floating or tethered), on an anchored platform anywhere in the water column, or on underwater vehicles. The miniaturization of processing and data storage allows for the collection of large databases in real or near-real time. Communications now allow the dissemination of data to many locations at any distance. *See* INSTRUMENTED BUOYS; UNDERWATER VEHICLES.

Underwater acoustic instrumentation, whether for basic science studies, the detection of objects in the sea, or the exploitation of nature's resources, can be divided into active and passive types. Active measuring systems send and receive signals. Passive systems only receive signals. The active systems currently use everything from explosives and airguns, which create high-energy short pulses at characteristically low frequencies, to small hand-held systems with ultrahigh frequencies.

An acoustic signal is usually created by the dilatation of either magnetostrictive or piezoelectric materials. The design for their use is dictated by the desired energy, frequency, and directivity (beam width). The most common materials used for the reception of signals in both active and passive systems are piezoelectric. The receivers are commonly called hydrophones. Significant gains can be achieved in signal level by geometrically arranging sets of hydrophones or transmitting elements with the proper spacing. These are called arrays. Modern electronics and digital processing allow for many hydrophones to be placed in an array that can be towed from a ship, deployed horizontally, suspended horizontally from a stationary ship or platform, or suspended from an anchor. **Figure 6** depicts some of these concepts. Sources can be fixed or towed and, in some cases, mounted on platforms such as ship's sonars and depth finders. *See* ACOUSTIC SIGNAL PROCESSING; HYDROPHONE; MAGNETOSTRICTION; PARAMETRIC ARRAYS; PIEZOELECTRICITY.

The common requirement of all oceangoing research is a ship to carry equipment, scientists, and engineers to the site of an experiment, to deploy the instruments, maintain station, and to run on

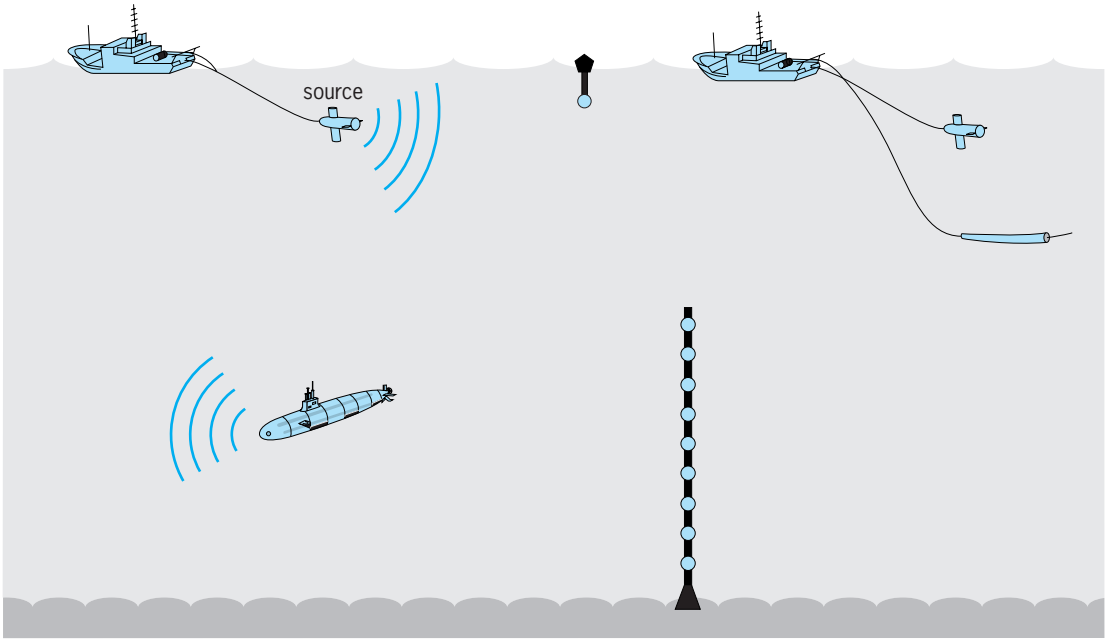

Fig. 6. Different sources and receivers used in deep-water and shallow-water experiments.

specifically determined routes. The planning and logistics for ocean experimentation are a vital part of the experience. The ships that are used for deep-sea work must be capable of withstanding storms and robust enough to be functional for the completion of experiments under difficult conditions. Nearshore operations usually do not have the same requirements. Often overlooked is the cost and tempo of working at sea. While ocean instrumentation has improved and its costs have dropped, the price of ships and the cost of operation have increased significantly. Although modern oceanographic vessels are much more capable than those of the past, they still sail at speeds of only a few knots. Seagoing is still a labor-intensive endeavor that cannot be overlooked in the planning of experiments. Sometimes aircraft are used for the deployment of sources and sensors. They are excellent within the limitations that are inherent in work with aircraft such as those on weight and volume. *See* OCEANOGRAPHIC VESSELS.

Ralph R. Goodman

Bibliography. L. Brekhovskikh and Yu. Lysanov, *Fundamentals of Ocean Acoustics*, 1982; C. S. Clay and H. Medwin, *Acoustical Oceanography*, 1977; C. S. Clay and H. Medwin, *Fundamentals of Acoustical Oceanography*, 1988; Y. Desaubes et al. (eds.), *Oceanographic And Geophysical Tomography*, 1990; G. Frisk, *Ocean and Seabed Acoustics*, 1994; B. R. Kerman (ed.), *Sea Surface Sound*, 1988; National Defense Research Committee, *Physics of Sound in the Sea*, 1945, reprint 1969 and 1989; I. S. Tolstoy and C. S. Clay, *Ocean Acoustics*, 1966, reprint 1987; R. J. Urick, *Ambient Noise in the Sea*, 1986; R. J. Urick, *Principles of Underwater Sound*, 3d. ed., 1983, reprint 1996; J. L. Worzel, M. Ewing, and C. L. Pekeris, *Propagation of Sound in the Ocean*, 1948, reprint 2000.

Underwater television

Any type of electronic camera that is located underwater in order to collect and display images. It must be packaged in a waterproof housing. The underwater camera may be packaged with its own recording device, or it can be attached to a television that is located on a ship, in a laboratory, or at a remote site. In the latter case, the images by the camera are called real-time images, as they are available at the same time that the camera is recording them. An underwater television may be used for sport, ocean exploration, industrial applications, or military purposes. Common imaged subjects are animals, coral reefs, underwater shipwrecks, and underwater structures such as piers, bridges, and offshore oil platforms.

During the daytime and at minimal depths, underwater television can be used to view objects illuminated with natural sunlight. For nighttime viewing or at very deep depths, artificial lights must be used. Light in the ocean is reduced in intensity quite severely: even in the clearest natural waters, a beam of blue or green light will be reduced in intensity by approximately 67% every 230 ft (70 m). Light that is propagating through an aqueous medium such as seawater or lake water will also be selectively reduced in intensity, based on wavelength. In most situations, the extreme reds and blues will be most severely attenuated, with the region of highest clarity being the yellow to blue-green wavelengths. *See* ELECTROMAGNETIC RADIATION.

Modern advances in television cameras have opened up a host of applications for underwater viewing, such as of standard-video, charge-injection-device, silicon-intensified-target, or charge-coupled-device cameras. These cameras can be adapted for high frame rate, low light level, or underwater color imaging. Computer modeling of underwater images can be used to predict the performance of an underwater imaging system in terms of range of viewing and quality of images as a function of the clarity of the water. The most spectacular use of an underwater television system was the discovery of the lost luxury liner *Titanic*. In order to find the *Titanic*, which was located at a depth of approximately 16,000 ft (4800 m), an underwater television system that consisted of a video camera equipped with lights was used. *See* CHARGE-COUPLED DEVICES; TELEVISION; TELEVISION CAMERA; UNDERWATER PHOTOGRAPHY.

Jules S. Jaffe

Bibliography. R. D. Ballard, The discovery of the *Titanic*, 1987; J. S. Jaffe, Computer modeling and the design of optimal underwater imaging systems, *IEEE J. Oceanic Eng.*, 15(2):101–111, 1990; R. C. Smith and K. S. Baker, Optical properties of the clearest natural waters (200–800 nm), *Appl. Optics*, 20:177–184, 1981.

Underwater vehicles

Submersible work platforms that can be operated either remotely or by an onboard crew. Unlike military submarines, which are essentially submerging ships, submersibles are not self-contained. All submersibles require some sort of surface-support vessel.

History

Perhaps the earliest underwater vehicles were the seventeenth-century ambient-pressure diving bells used mostly for recovering items from shipwrecks, in which the workers were in a dry environment but exposed to full depth pressure. Development of practical,crewed underwater vehicles began before World War I, with 1-atmosphere surface-tethered diving bells and armored diving suits. Both were intended to perform salvage tasks. Some of the bells could go to depths of several hundred feet and were equipped with external lights and a variety of tools.

Armored suits were one-person submersibles in the general shape of a diving suit. Depth capabilities to 700 ft (210 m) were possible since the operator/diver was not exposed to sea pressure. Bottom time at the work site was greatly increased over conventional diver capability. Problems with the flexibility and watertight integrity of the suits' articulating

joints limited use of these devices prior to World War II.

In 1930, the United States scientist William Beebe and engineer Otis Barton developed the *Bathysphere* ("deep sphere"), a thick-walled steel ball lowered into the sea on a steel cable. The interior provided sufficient room for two crew members and their equipment. Two quartz glass windows permitted outside viewing and photography. The cabin had a third window opening but that window was broken during installation and was covered with a steel plate. The *Bathysphere* made 32 dives from 1930 to 1934 at Bermuda, ending with a dive to 3028 ft (910 m). Beebe's classic book, *Half Mile Down*, tells the story of this pioneering deep-sea research program.

In 1950, an improved bathysphere, the *Benthoscope*, was developed by Barton. It was successfully tested off California to depths of 4500 ft (1350 m), but made few working dives. Japan developed the *Kuroshio* for oceanographic research in the early 1950s. Retired in the 1960s, it was capable of diving to 650 ft (200 m). This was the world's last operational bathysphere, although the diving bells of today are not much different in concept.

It was not until 1948 that the first practical untethered underwater vehicle was tested: the bathyscaph ("deep ship") *FNRS-2*, invented by Swiss physicist Auguste Piccard. It was an underwater free balloon, with buoyancy provided by lighter-than-water aviation gasoline contained in a large steel float (balloon). A thick-walled spherical cabin for two crew members was suspended beneath the float. The float moved vertically by adding weight (seawater) or losing weight by dropping steel ballast pellets. Small electric motors fitted with propellers provided limited horizontal and vertical movements.

From 1950 to 1978, the French Navy developed and operated bathyscaphs *FNRS-3* (1953) and *Archimede* (1963). In 1953, Piccard built his last bathyscaph, *Trieste*, which was sold to the U.S. Navy in 1958. In 1960, a Navy expedition took *Trieste* to a depth of 35,800 ft (10,700 m) at Challenger Deep near the island of Guam in the western Pacific, the deepest place in the world ocean. After the original *Trieste's* retirement in 1963, the Navy built two improved bathyscaphs called *Trieste II*. Both had a maximum depth capability of 20,000 ft (6000 m).

Archimede was retired in 1978. It was the last crewed submersible in the world capable of diving to the deepest place in the sea. Four years later, *Trieste II* was retired; it was the last operating bathyscaph in the world. This type of deep submersible vehicle had served 35 years; much of this time, it was the only way to work in the deep ocean.

These retirements were not the end of deep-diving submersibles. New construction materials and technologies have provided the means for building submersibles that are much smaller, less costly to operate, and safer than the bathyscaphs.

During the 1960s and early 1970s, divers and crewed submersibles dominated the underwater scene. However, new types of vehicles were under development in research laboratories during this time. A major development was the remotely operated vehicle (ROV), a tethered vehicle controlled from the surface of the sea. The autonomous underwater vehicle (AUV), essentially an underwater robot, was also under active development but its maturity would take much longer than the ROVs.

Almost all early submersible designs, technologies, and operational techniques were developed by navies, the U.S. Navy in particular. Military missions and related science needs were the primary forces driving research on underwater operations during this period. By the mid-1970s, civil applications were stimulated by the steep oil price increases worldwide. This led to increased offshore oil and gas exploration and production activity. Opening new resource areas, such as the North Sea, accelerated development of better tools and techniques for conventional deep-sea divers. Beginning in 2004, this pattern of technological demand began repeating itself as oil prices soared again. Exploration and production now go to depths as great as 2 mi (3 km), where only submersibles can do the undersea work. *See* DIVING; OIL AND GAS, OFFSHORE.

Types

Underwater vehicles can be grouped into three general categories: crewed deep submersible vehicles (DSVs), tethered remotely operated vehicles (ROVs), and untethered autonomous underwater vehicles (AUVs). In recent years, the term "human occupied vehicles" (HOV) has sometimes been substituted for DSV.

There are also hybrid vehicles, which combine two or three categories on board a single platform. Within each category of submersible, there are specially adapted vehicles for specific work tasks. These can be purpose-built or modifications of standard submersibles.

Crewed deep submersibles. Since the late 1950s, over 200 crewed submersibles have been put into service worldwide. By 2005, about 75 were operational; the majority support tourist operations, with marine science applications being the second greatest. There are five types of these vehicles: 1-atm untethered vehicles; 1-atm tethered vehicles, including observation/work bells; atmospheric diving suits; and diver lockout vehicles. While they differ mainly in configuration, source of power, and number of crew members, all carry a crew at 1-atm (10^2-kilopascal) pressure within a dry, pressure hull. The fifth type of crewed submersible is the wet submersible, in which the crew is exposed to full depth pressure.

The purpose of the DSV is to put the trained mind and eye to work inside the ocean, often referred to as in situ work. The earliest submersibles had very small viewing ports fitted into thick-walled steel hulls. In the mid-1960s, experimental work began on use of massive plastics (acrylics) as pressure hull materials. Today, submersibles with depth capabilities to 3300 ft (1000 m) are being manufactured with pressure hulls made entirely of acrylic. Essentially the hull is now one huge window.

Fig. 1. Two-person deep submersible vehicle *Deep Rover* with all-acrylic pressure hull. (*Deep Ocean Engineering, Inc.*)

1-atmosphere untethered vehicles. These are submersibles that carry 1–64 occupants at 1-atm pressure (**Fig. 1**). The submersible is battery-powered (though fuel cells have been tested), free-swimming, and capable of limited horizontal travel across the seabed.

The four deepest-diving crewed submersibles are operated by France (1), Russia (2), and Japan (1). Three can dive to 20,000 ft (6000 m), and Japan's *Shinkai 6500* can go to 21,323 ft (6500 m). Though the deepest place in the ocean is about 36,000 ft (11,000 m), there is no crewed vehicle at present that can reach this depth. However, a submersible designed to operate at 20,000 ft can reach 98% of the sea floor worldwide. Therefore, designing for two-thirds the maximum depth of the ocean is a reasonable engineering and cost goal.

In 2005, two programs began constructing new crewed submersibles for marine science work. At the Woods Hole Oceanographic Institution, a 21,323-ft (6500-m) vehicle is being designed to replace the 13,000-ft (4000-m) *Alvin*, built in 1964, and in the Peoples Republic of China a 22,460-ft (7,000-m) crewed submersible is being built.

Observation/work bells and 1-atm tethered vehicles. These carry a crew of two or three at 1-atm pressure. Power is supplied by batteries, or from the surface through an umbilical cable. The vehicles (**Fig. 2**) always operate with a lift cable from the surface and are designed for panoramic viewing and a high degree of maneuverability and station-keeping, and are fitted with manipulators (external mechanical arms and hands) for work tasks. Many of these tethered vehicles are capable of limited maneuvering in three dimensions. They are primarily designed for performing manipulative work around underwater structures mostly in support of offshore oil and gas development. Maximum depth is 2300 ft (700 m) [**Fig. 3**].

Atmospheric diving suits (ADS). These are designed for a one-person crew at atmospheric pressure. In operation, they are similar to the 1-atm tethered vehicles. In this case, the vehicle is anthropomorphically configured (**Fig. 4**) and power for maneuvering is manual or electrical via a surface umbilical cable. The suit offers a high degree of manipulative dexterity and maneuverability. Atmospheric diving suits derive propulsive power from thrusters or manually by the operator's walking. The maximum depth of operation is 2000 ft (600 m).

Diver lockout submersibles. These can be either part of a free-swimming vehicle or a tethered bell system. They carry both an operating crew and a dive team. One section of the vehicle's pressure hull is at 1-atm pressure, while the other section can be pressurized to ambient pressure to permit diver egress. Power may be derived from batteries, advanced air-independent energy systems (such as fuel cells), or a surface-connected umbilical cable. Diver-breathing gases and heating may be carried aboard or supplied through an umbilical cable. A maximum operating depth (for diver lockout) is about 820 ft (250 m). Only a few diver lockout vehicles still exist and they are rarely used for commercial work. Navies use free-swimming ones for special warfare missions.

The more common way of delivering commercial divers to work sites is to use a cable-lowered diving

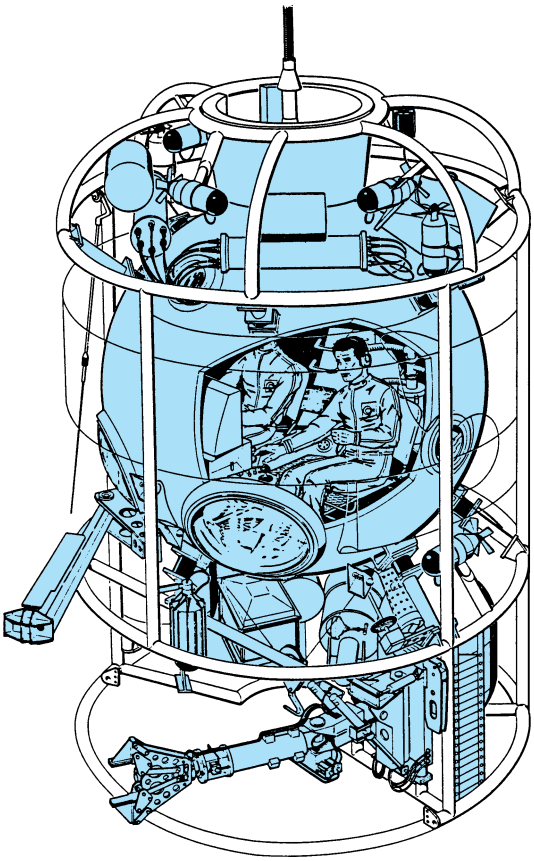

Fig. 2. Observation/work bell. (*Comex Industries*)

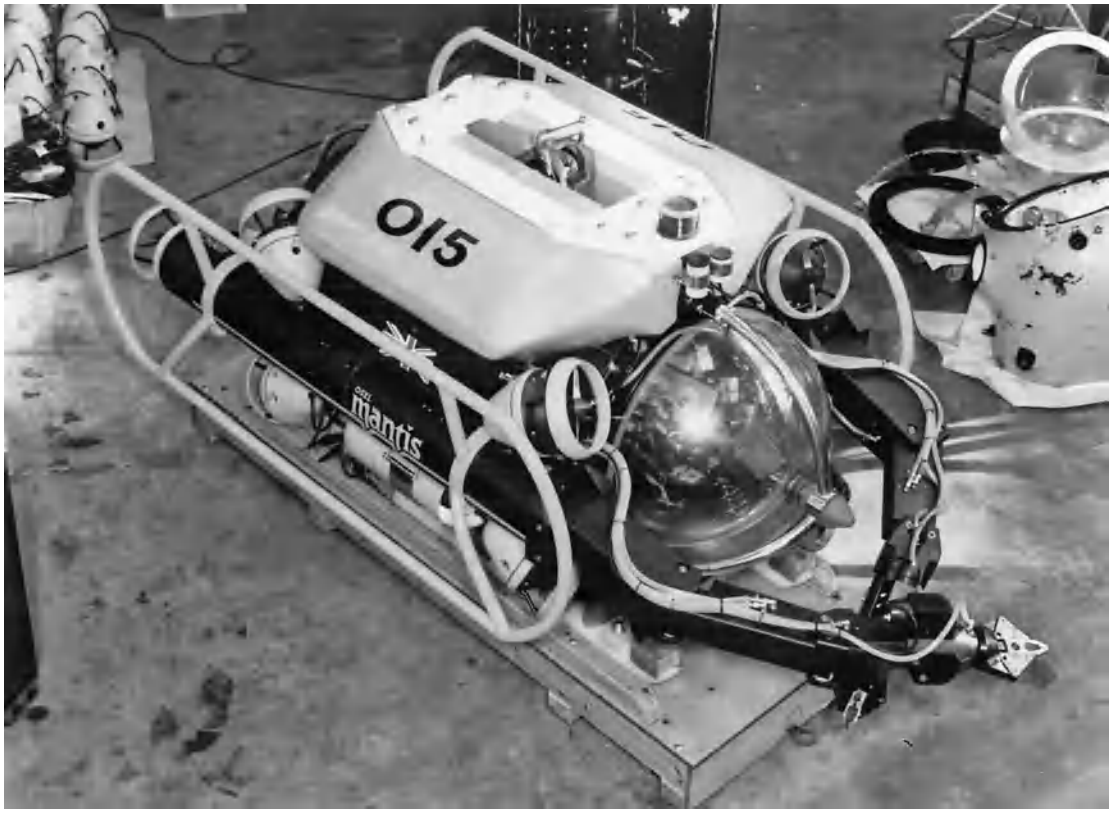

Fig. 3. One-atmosphere tethered vehicle *Mantis* (*OSEL Group, Ltd.*)

Fig. 4. HARDSUIT atmospheric diving suit. (*Oceanworks International*)

bell called a personnel transfer capsule (PTC). The divers are delivered and recovered from the dive site in a dry bell pressurized to equal the water pressure at the dive depth. A hatch in the bottom provides a means for the divers to leave and reenter the bell. On returning to the surface, the bell is mated to a deck decompression chamber (DDC) on aboard the support vessel, where the divers can decompress over periods ranging from hours to days.

Wet submersibles. There are two types of wet submersible vehicles: simple diver propulsion units in which a diver holds onto and steers through the water, and swimmer delivery vehicles (SDVs) in which the diver either rides inside (**Fig. 5**) or sits astride. In all cases, the personnel are exposed to full ambient pressure and water temperature. These vehicles are battery-powered and designed for shallow depths, not much more than 100 ft (30 m). The breathing air is carried in on-board tanks or on the diver's back. A few swimmer delivery vehicles are designed so that the interior of the cabin does not completely flood. The crew sits in a trapped air bubble, keeping the upper halves of their bodies in pressurized air. Facemasks or regulators are not necessary while passengers remain in the vehicle. While most of these vehicles have been developed for recreation, navy special warfare personnel also use similar vehicles.

Remotely operated vehicles. The first ROVs were developed in the late 1950s for naval use. By the mid-1970s, they were being widely adapted for the civil

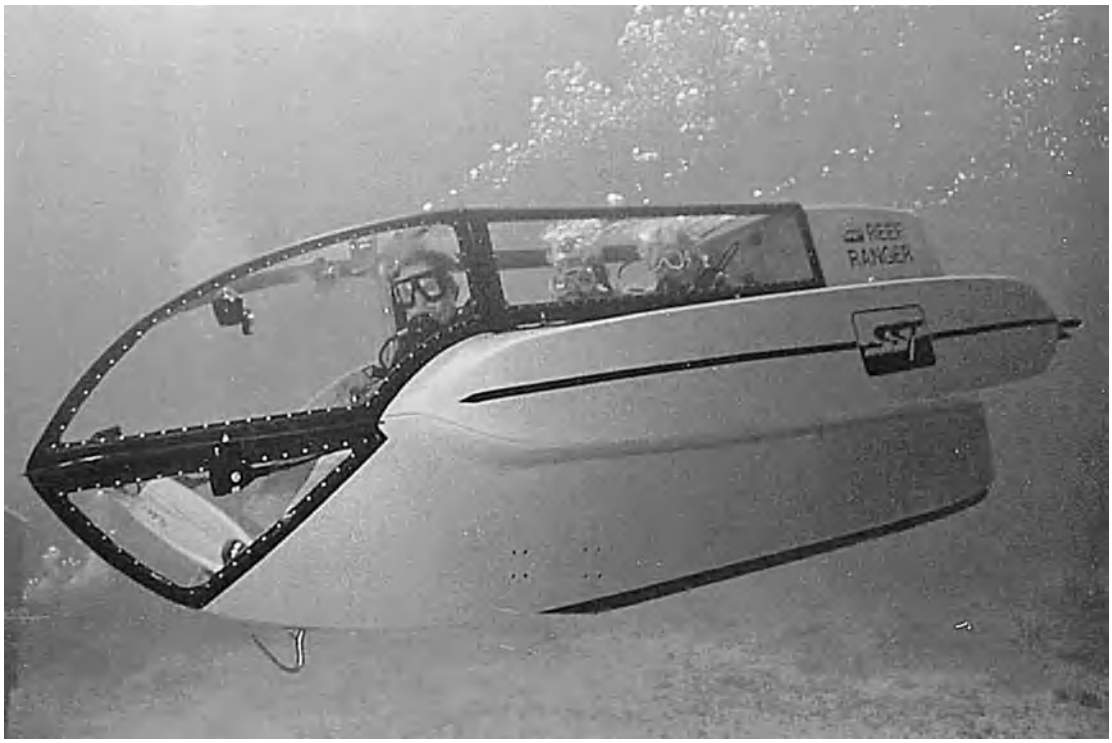

Fig. 5. Reef Ranger wet submersible. (*Submersible Systems Technology*)

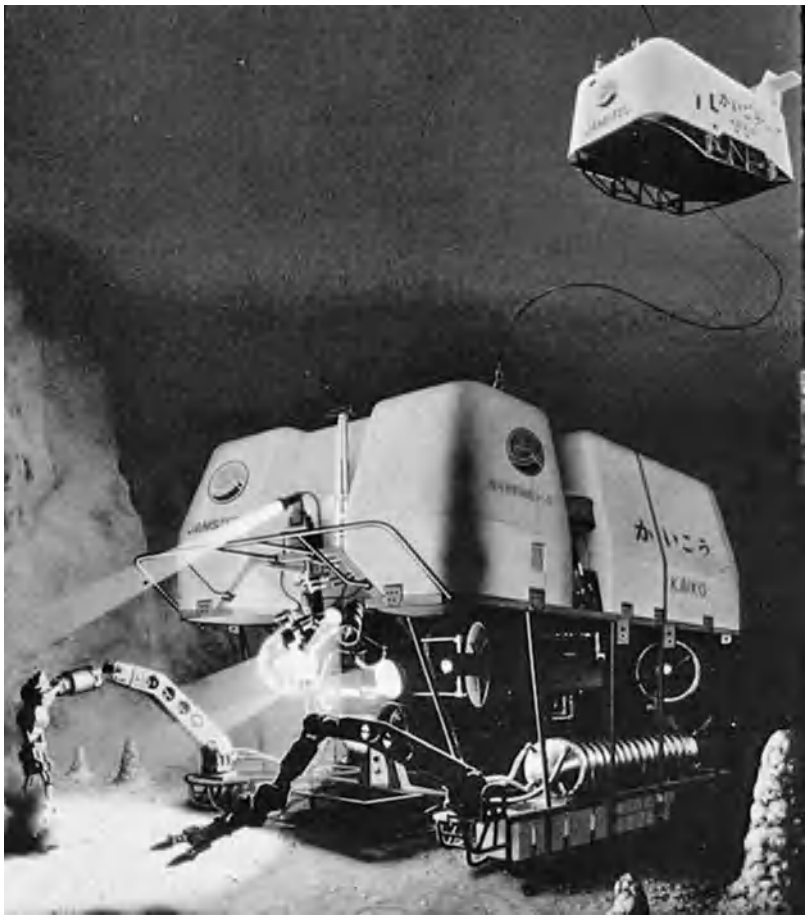

Fig. 6. Japan's 36,000-ft (11,000-m) *Kaiko*, a large working ROV. (*Jamstec & Mitsui Engineering and Shipbuilding*)

sector. In the offshore oil and gas industry, the ROV has substantially replaced both crewed submersibles and divers.

The rapid acceptance of these submersibles is due to their relatively low cost and the fact that they do not put human life at risk when undertaking hazardous missions. However, their most important attribute is that they are less complex. By virtue of their surface-connecting umbilical cable, they can operate almost indefinitely since there is no human inside requiring life support and no batteries to be recharged. There are four types of ROV: tethered free-swimming, towed, bottom-reliant, and structurally reliant.

Tethered free-swimming vehicles. These are the dominant remotely operated vehicles, and nearly 5500 have been constructed since the early 1970s. This type of ROV can be divided into two use categories: inspection vehicles (essentially swimming cameras) and work vehicles (equipped with a variety of tools and sampling devices). Costs of ROVs range from about $15,000 to several million dollars. *See* UNDERWATER PHOTOGRAPHY.

The vehicle itself is only part of an operating system, consisting of a control/display console, a power supply, a cable winch with a cable/vehicle handling apparatus, an umbilical cable (which transmits power and control commands down to the vehicle and data and television signals back to the surface), and the vehicle handling cage or "garage" (which is optional). The ROV is always equipped with a television camera, lights, and thrusters that provide three-dimensional maneuverability. In addition, there can be depth sensors, a wide array of manipulative and acoustic devices, and special

instrumentation to perform a variety of work tasks. For complex or deep-diving operations, the submersible is often equipped with an acoustic tracking system and a high-resolution sonar. This permits the surface support vessel to maintain close control of the vehicle relative to other targets on the sea floor. These vehicles can range from the size of a personal computer to that of a compact automobile. Larger ROVs are configured as basic work platforms that can use a wide variety of optional equipment according to the mission. *See* SONAR; UNDERWATER SOUND; UNDERWATER TELEVISION.

In the late 1980s, a few remotely operated vehicles with depth capabilities to 20,000 ft (6000 m) were put into service for deep ocean inspection and object recovery missions. In March 1995, the $50 million Japanese ROV *Kaiko* dove to 35,798 ft (about 11,000 m) into the Challenger Deep area off the Marianas Islands, equaling the 1960 record of the U.S. Navy's crewed Bathyscaph *Trieste*. Subsequently *Kaiko* returned to this site twice, demonstrating there are no depth limits to ROV capabilities. *Kaiko* was lost on a 6000 m dive in 2004; however, the Japanese government plans to replace it (**Fig 6**).

The vast majority of these vehicles have been operated at depths of less than 2000 ft (600 m). However, offshore oil and gas is being produced at depths greater than 3000 ft (1000 m), well beyond diver depths. Exploratory drilling occurs at 10,000 ft (3000 m) depth and it will not be long before production begins at these depths. Here, remotely operated vehicles are not an option but a necessity and several have been developed for these applications. Long-penetration vehicles are being developed to inspect the insides of several-mile-long aqueducts, ocean outfalls, and other large pipelines. For example, in 1998, a special ROV, *Phoenix*, made a record-setting 6-mi-long, (10,000-m) tunnel inspection in Finland.

Towed vehicles. Often called "sleds," these are usually equipped with instruments and imaging systems for long-range surveying, search, and reconnaissance (**Fig. 7**). As they are towed by a surface ship, their depth is controlled by adjusting the length of tow cable. Some sleds have fins to stabilize their underwater flight. There are two varieties within this class of vehicle. One type is towed above the bottom and is instrumented for scientific research or broad-area reconnaissance. The maximum depth of this type of vehicle is 20,000 ft (6000 m), though there are only a few in the world that can operate this deep. The second type is designed to be towed while in contact with the bottom to lay and bury pipelines or cables. These are basically specialized sea-floor plows.

Bottom-reliant vehicles. These are generally large, massive vehicles designed for a single-purpose work such as pipeline and cable burial, pipeline inspection, bottom excavation, and trench backfilling. They are controlled by an operator team on board a dedicated support ship, and maneuver on the bottom using

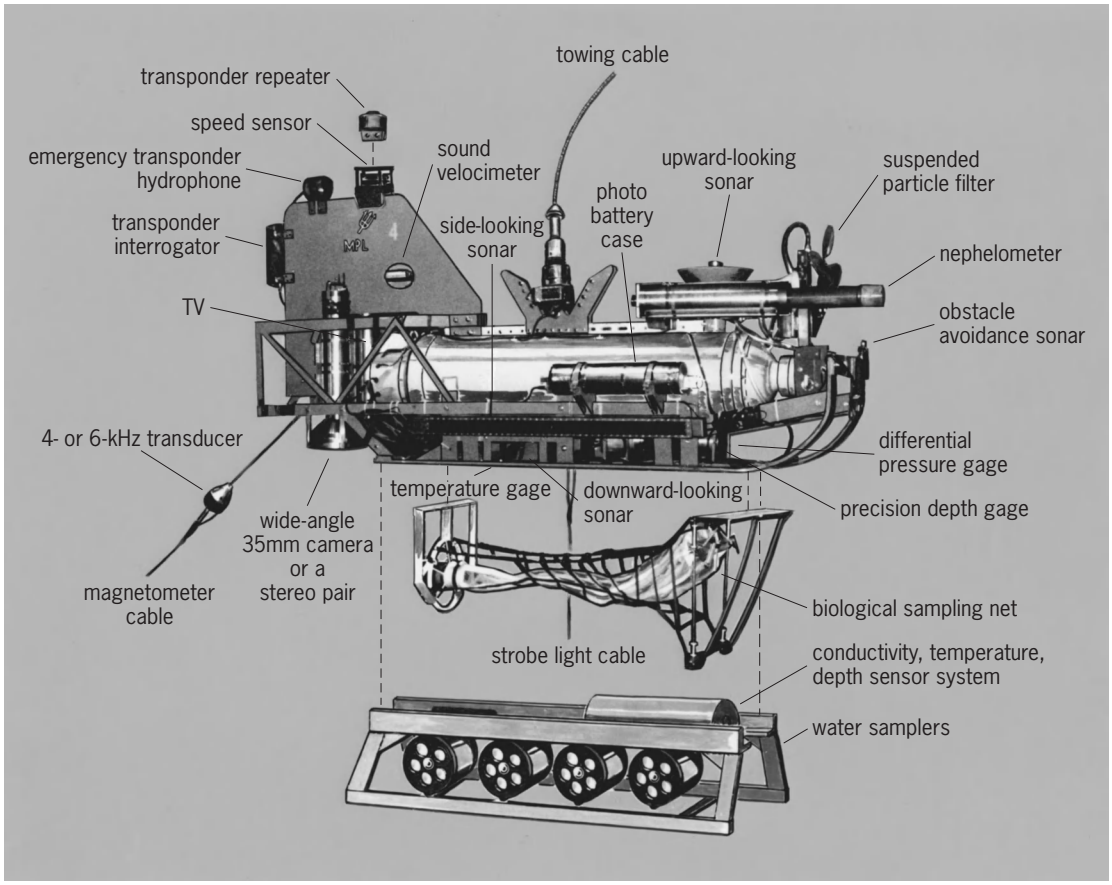

Fig. 7. Components of a towed vehicle. (*Marine Physical Laboratory, Scripps Institute of Oceanography*)

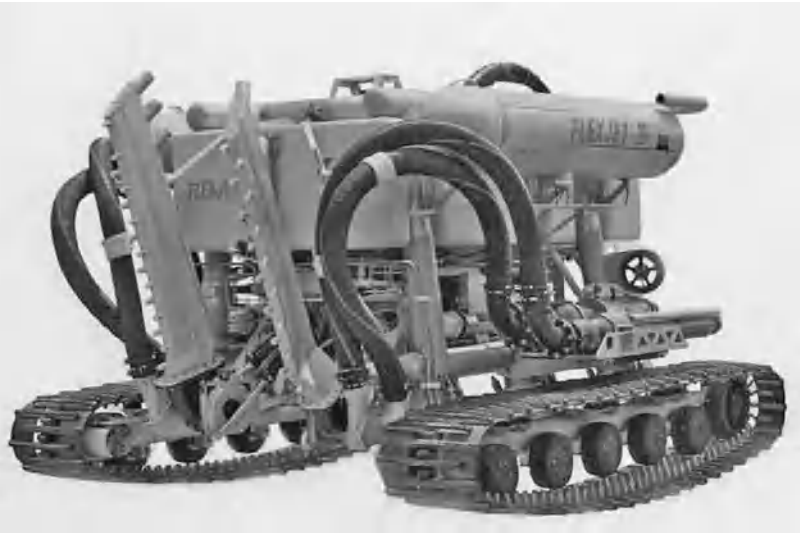

Fig. 8. Bottom-reliant vehicle. (*Sonsub International*)

wheels, caterpillar-like tracks, or Archimedes screws (**Fig. 8**).

Structurally reliant vehicles. These vehicles obtain power and control from the surface, but are designed to operate in contact with a fixed structure such as an oil platform. They can run on wheels or be propelled by push-pull rams as they crawl along a pipeline or other structure (**Fig. 9**). These vehicles are purpose-designed and conduct such tasks as ship hull and structure cleaning, inspection and maintenance of subsea production systems, and in-water repair tasks.

Autonomous underwater vehicles. These are crewless and untethered submersibles which operate independent of direct human control. Practically, AUVs are mobile instrumentation platforms with propulsion, powered by on-board batteries or fuel cells. A preprogrammed, on-board computer controls their operations, although a few experimental vehicles have been controlled by acoustic commands from the surface. The data they collect is stored and often processed on-board and, after surfacing, transmitted to other stations via satellite or radio link for further processing. Autonomous underwater vehicles' depth capability is from a few hundred feet to over 20,000 ft (6,000 m).

Fig. 9. Structurally reliant vehicle used for servicing seafloor oil well structures. (*Kvaener Brug A/S*)

First developed by the military in the early 1970s, numerous operational AUVs have entered the civil sector in recent years. Military applications include such tasks as minefield location, mapping, and neutralization; minefield installation; submarine decoys; military oceanography; and covert intelligence collection. Several navies now have operational AUV programs which actively use these vehicles in fleet operations.

Civilian tasks include site monitoring, basic oceanographic data gathering, under-ice mapping, offshore structure and pipeline inspection, and bottom mapping. Autonomous submersibles are particularly useful where human presence is not required and for long-duration operations.

Autonomous underwater vehicles span a wide range of sizes and capabilities, related to their intended missions. Large, transport-class platforms, about 33 ft (10 m) in length and 11 tons (10 metric tons), have been designed for missions requiring long endurance, high speed, large payloads, or high-power sensors. Smaller, network-class platforms, about 3.3 ft (1 m) in length and weighing 220 lb (100 kg), address missions requiring portability, multiple platforms, adaptive spatial sampling, and sustained presence in a specific region. Historically, most vehicles have been propeller-driven, but several new, variable-buoyancy, underwater gliders are emerging in the network class.

Over 100 autonomous underwater vehicles have been designed and built worldwide since 1963 by government laboratories, industry, and academia. Through the mid-1990s, most AUVs in all classes were research prototypes. By 2000, an increasing number of operational military and commercial vehicles were available, particularly in the network class. Developments in microprocessors, memory, batteries, and fuel cells are leading to higher performance in smaller platforms. Likewise, fiber optics and micro-electro-mechanical systems are spawning a new generation of compact, high-precision, low-power sensors ideally suited for payloads on autonomous underwater vehicles. The Global Positioning System (GPS) and emerging global communication services are enabling timely access to data and real-time control. *See* MICRO-ELECTRO-MECHANICAL SYSTEMS (MEMS); SATELLITE NAVIGATION SYSTEMS.

A major obstacle to more extensive use of AUVs is the limit on the amount of on-board power they can carry, thus limiting mission duration. Most submersibles of this type are powered by batteries. True long-duration missions will require an alternative to battery systems. Research is ongoing to find other power systems, such as fuel cells and solar cells which can be mounted on the vehicle. *See* BATTERY; SOLAR CELL.

Transport class. In the transport class, the *Mobile Undersea Systems Test Laboratory* is the world's

largest autonomous underwater vehicle: 33 ft (10 m) long, 4 ft (1.2 m) in diameter, weighing 10 tons (9.2 metric tons), rated to 1970 ft (600 m) depth, with an 87-mi (140-km) range and 1.1-ton (1-metric-ton) payload capacity. Other transport-class vehicles rated to full ocean depth include the French *Epaulard*, the Russian *MT-88*, the Chinese *CR-01*, and the *Theseus*, developed for the Canadian Department of National Defense. *Theseus* is rated to 3280 ft (1000 m) depth with an operating range of 466 mi (750 km) and an endurance of 100 h. In 1996, *Theseus* completed two missions under the Arctic icecap of 218 mi (350 km) and 50 h in duration.

Intermediate class. Intermediate between transport class and network class are vehicles such as the Marine Utility Vehicle System (MARIUS). At 14.8 ft (4.5 m) long, 2 × 3.6 ft (0.6 × 1.1 m) in rectangular cross section, and rated to 1970 ft (600 m) depth, MARIUS has an operating range of 31 mi (50 km) and is used for seabed inspections and environmental surveys in coastal waters. The Norwegian-built *Hugin 3000* AUVs have been quite successful commercially in doing seafloor surveys to depths as great as 9000 ft (3000 m) and for missions as long as 48 h.

Network class. Typical high-performance, propeller-driven vehicles in the network class include the Autonomous Benthic Explorer (ABE) and the Remote Environmental Monitoring Units (REMUS). Developed by the Woods Hole Oceanographic Institution, ABE is 992 lb (450 kg) in air, rated to 3.7 mi (6000 m) depth, and has an operating range of 62 mi (100 km). In 1997, ABE mapped magnetic anomalies associated with deep-ocean spreading zones in the North Pacific. REMUS, a highly portable, shallow-water vehicle is 5 ft (1.5 m) long, 8 in. (20 cm) in diameter, 88 lb (40 kg) weight in air, with up to a 50-mi (80-km) range (**Fig. 10**). REMUS has mapped water properties and performed mine reconnaissance to the edge of the surf zone. The U.S. Navy has used one of these submersibles for mine countermeasure work in the Arabian Gulf.

Network-class, buoyancy-driven vehicles are small, near-neutrally-buoyant platforms that move vertically and horizontally through the water because of small changes in buoyancy. One such vehicle, the Virtual Mooring Glider (VIRMOG), is fitted with battery-operated buoyancy controllers, and can operate with an estimated 40-day, 497-mi (800-km) range independent of temperature gradient. VIRMOG is designed for both deep- and shallow-ocean sampling, and uses changes in buoyancy for propulsion (underwater glider). Batteries are used as the energy source for the variable-buoyancy engine. One mission envisioned for VIRMOG is as a holding station at a fixed location, thus virtually emulating a mooring.

Hybrid vehicles. Hybrid vehicles are those that combine crewed vehicles, remotely operated vehicles, and divers. For example, the hybrid *DUPLUS II* can operate either as a tethered free-swimming ROV or as a 1-atm tethered crewed vehicle. This evolved to provide capability for remotely conducting those tasks for which human skills are not needed, and then to put the human at the place where those skills are required. Other hybrid examples include ROVs that can be controlled remotely from the surface or at the work site by a diver performing maintenance and repair tasks, bottom-crawling ROVs controlled by a diver to anchor a pipeline or cut a trench, and small "scout" ROVs that can "fly off" from a larger ROV. The Woods Hole Oceanographic Institution's *Argo-Jason* is an example of the fly-off system. It combines a towed vehicle (*Argo*) with a tethered free-swimming vehicle (*Jason*). *Argo* conducts large-area search or reconnaissance, while *Jason* (carried within and deployed from *Argo*) conducts detailed inspection and manipulation of objects of interest encountered during the towing phase.

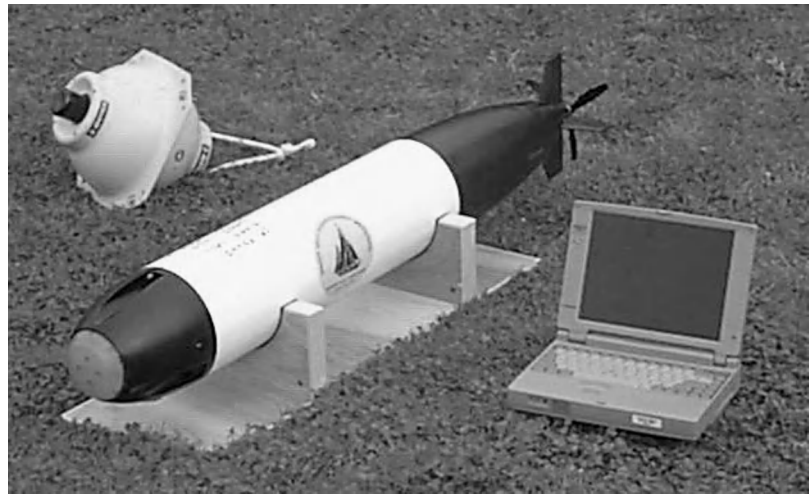

Fig. 10. REMUS is rated for continental shelf depths (355 ft; 200 m) and is compact and lightweight (40 kg; 88 lb). It is depicted with its laptop computer for downloading missions and uploading data, and one of the acoustic transponders used for navigation. (*Woods Hole Oceanographic Institution*)

Applications

Almost every conceivable underwater work task from military missions to recreational activities and from offshore oil and gas to environmental conservation has used submersibles. While there will always be a need for divers, as working depths increase the need for submersibles will continue to increase. The question of whether crewed submersibles will be replaced by ROVs and ROVs by AUVs is not valid. It is best to consider these types of vehicles, and all their variants, as a "family" of tools. The in situ worker chooses the best ones for a given job. A sample of work areas are listed below.

Offshore oil and gas activities. At one time, there was a clear distinction between the tasks that remotely operated vehicles performed in support of offshore oil and gas operations and those performed by divers and crewed submersibles. Over time, the capabilities of ROVs have increased to the point that almost every task once done by divers and crewed submersibles can now be done with ROVs, often better, and always at lower cost. As a result, crewed submersibles are rarely used in this business.

Observation, and video or photographic documentation. This is of the strongest areas of ROV uses, although in a

few cases some crewed vehicles are still used. The latter tend to be the 1-atm tethered and atmospheric-diving-suit (ADS) type submersibles. In shallow depths (<115 ft or 35 m) the diver is preferred. Observational tasks include determining the geometry and position of pipelines and cables following installation, determining the condition of a pipeline's concrete coating after installation, obtaining accurate positions of pipeline tie-in locations, leak detection monitoring, inspecting a wellhead for structural integrity, investigating accidents, and accumulating information needed to develop salvage or retrieval plans. Also included are bottom surveys such as pipeline and cable route selection, verification surveys, hardware site installation surveys, and debris identification and location mapping.

Structural nondestructive testing. This technique is used to locate and identify bent, broken, or missing structural members, identify and map debris location on structures, assess and remove marine growth, locate suspended pipeline and cable sections, perform structural-member thickness measurements, detect and measure cracks and flaws, and measure the effectiveness of the structural cathodic protection system.

Support of diver operations. Remotely operated vehicles can provide diver support and monitoring, including initial diving-gear checkout for leaks, continuous diver monitoring of job performance, safety, assistance in locating the diver's position, topside understanding of diving conditions, evaluation of the dive site in terms of diver safety prior to dive, provision of an additional and mobile light source, monitoring and inspection of the diver's work, and television or photographic documentation of the diver's work.

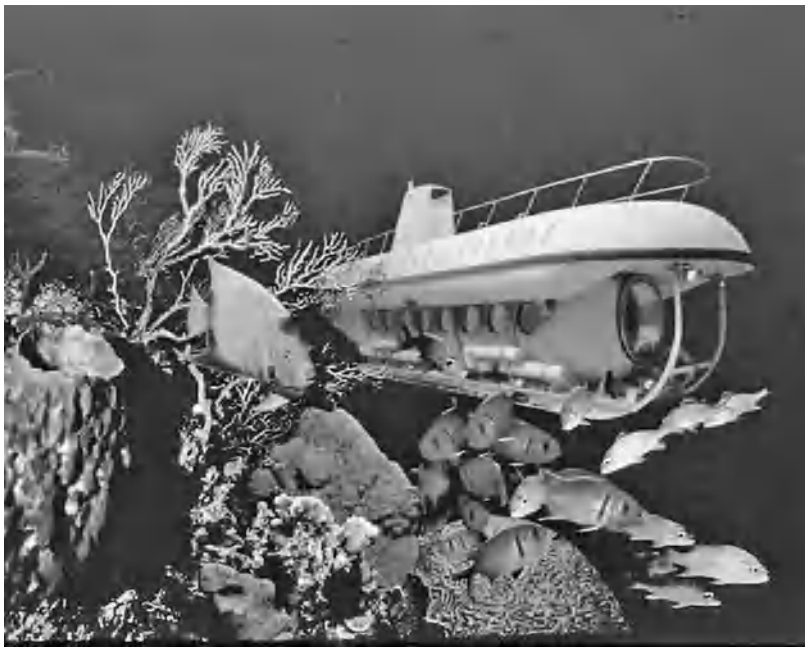

Fig. 11. Atlantis 48-passenger tourist submarine. (*Atlantis Submarines International*)

Object search, location, and retrieval assistance. Underwater vehicles are used to locate lost or abandoned objects, to determine the object's position relative to the surface salvage or retrieval vessel, and to provide assistance in attaching lift lines or cables to objects for retrieval. The majority of underwater vehicles cannot lift items weighing more than 110–220 lb (50–100 kg). Often separate lift systems (underwater elevators) are employed to work together with the submersible to bring heavy loads to the surface.

Oversight of underwater activities. Underwater vehicle are used to monitor underwater construction from surface platforms (such as ships, barges, and docks). These tasks include grouting operations, piling installation, alignment and orientation measurements during structural installation, observation of pipeline pull-in procedures, pipeline weighting and backfilling operations, and touchdown operations (for example, pipeline installations or spudding-in of drills).

Drilling support assistance. This is supplied by locating and retrieving objects, debris, and equipment; removing sediment to permit observation of components; connecting and disconnecting shackles and hydraulic/electric lines; collecting grout and bottom samples; operating valves; observing the drillstring operation and blowout preventer installation; and placing and recovering acoustic marker beacons.

Support of routine deepwater production operations. With hydrocarbon production occurring in increasingly deep waters that are well beyond diver depths, the routine management, inspection, and maintenance of seafloor structures with their associated pipelines is being done with ROVs. As the AUV technologies and techniques mature, eventually they will be used for tasks where a tethered vehicle is not feasible. In all cases, the sea-floor structures will be configured to be vehicle friendly for the mechanical arms (manipulators) and electronic eyes on the submersibles.

Marine leisure. This sector is the single largest user of crewed submersibles. Built in the early 1960s, the *Auguste Piccard* was the first tourist submarine. This 64-passenger submersible was used at the 1964 Swiss National Fair in Geneva. Crewed by pilot, copilot, and stewardess, it carried over 34,000 paying passengers during the 14-month-long fair.

In 1985, a Canadian company, Atlantis Submarines, built the 28-passenger *Atlantis I*, which was placed into service at Grand Cayman Island. This began a business sector that has employed some 60 crewed submersibles since that time. While a few of the early tourist submarines were conversions of existing DSVs, 51 were built for the business sector. The largest is the 66-passenger *Atlantis XIV* operating off Waikiki Beach, Hawaii. Thirteen *Atlantis* submarines are now in operation at Grand Cayman, Barbados; St. Thomas, U.S. Virgin Islands; Aruba; Cozumel; Hawaii; and Guam (**Fig. 11**). Tourist submersibles have operated in 34 countries throughout

the world. By 2005, they had carried nearly 12 million passengers with no serious incidents. A typical dive lasts about 1 h and goes to depths of 100–150 ft (30–46 m). There are also two small converted commercial-type submersibles at Grand Cayman that take two passengers to depths of 1000 ft (300 m).

The large passenger-carrying submersible business is a mature market sector. The best sites are occupied and few new submarines will be added to the fleet in the future. There is a new development direction, the resort-sized submersible for locations that cannot support a 40–50-passenger vehicle. These DSVs have a 2–20-passenger capacity and are better suited for destination resorts and smaller marine leisure locations. In 2005, an American company launched the *Alicia*, a 1000-ft (300-m) tourist submersible that can carry six passengers.

In addition to passenger-carrying tourist submarines, there are a number of personal or recreational submersible designs now being offered to the leisure market. Most of these are wet submersibles, although in 2005 two 1-atm DSVs were delivered to private owners.

Military. The most important applications are intelligence gathering, search, identification, location, retrieval, and neutralization of ordnances. These include explosive ordnance (mines, torpedoes, and bombs), downed aircraft, sunken vessels, and other objects of high national interest. Other military applications include hardware site and cable route survey, submarine rescue, wreck marking and destruction, topographic surveying, radiation measurements, water sampling, and water turbulence measurements and tests.

For several years, special warfare organizations, such as the U.S. Navy SEALS, have used small submersibles as swimmer delivery vehicles. In general, they are carried on submarines and the submersibles are launched and recovered while the submarine is submerged. In the larger navies, "wet" ambient-pressure vehicles are being replaced by 1-atm dry submersibles. However, many of the smaller navies in the world will continue to use the wet submersibles for covert operations.

Up through the 1970s, military research and development was the primary force behind the development of all types of submersibles. Today, almost all underwater submersible development is done in industry and academic research facilities. The military either contracts with them for services or buys this equipment from the private sector.

Scientific research. The crewed vehicle, particularly the 1-atm untethered variety, continues to be the workhorse of the scientific community. Twenty-four are in service worldwide for support of marine research. In this role, the submersible has become a very sophisticated platform from which the scientist can directly observe and measure a variety of environmental aspects, which are best understood by direct visual observation. The best-known is the Navy-owned *Alvin*, which has been operated by the Woods Hole Oceanographic Institution (WHOI) since 1964. Diving as deep as 13,000 ft (4000 m), *Alvin* has made many significant discoveries in the deep ocean. Since the early 1980s, a family of 20,000-ft (6000-m) DSVs have been used for research work. Examples of this class are the two Russian Mir submersibles in operation since 1987 (**Fig. 12**).

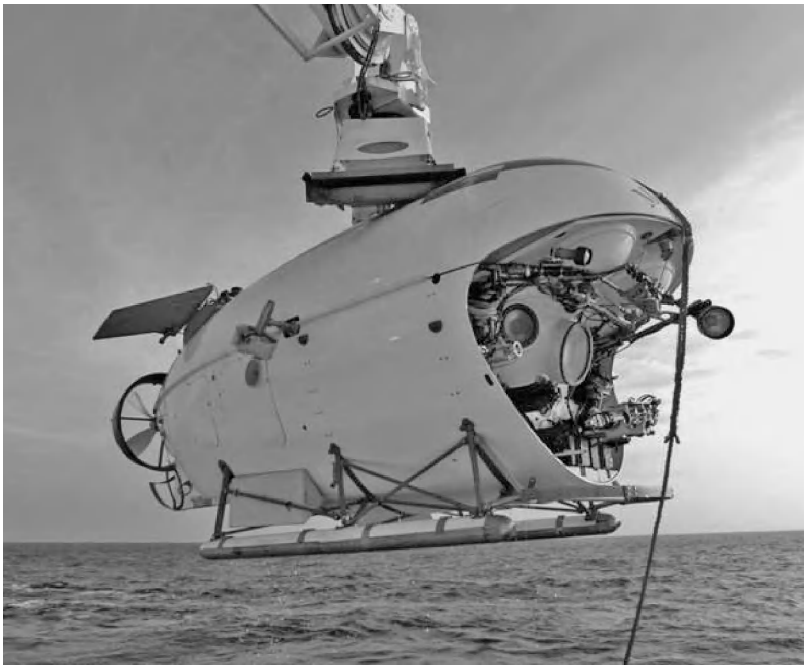

Fig. 12. One of the two Russian Mir 20,000-ft (6,000-m) manned submersibles. (*P.P. Shirshov Institute of Oceanology*)

In recent years, several scientific institutions have started to use remotely operated vehicles for research platforms. They offer an affordable submersible platform with excellent flexibility for a wide range of mission configurations. The rapid progress in development of microcomputers has made it possible to put much more capability into these small vehicles. *See* MICROPROCESSOR.

Towed vehicles have been used for research tasks since the 1960s. They are used to conduct deep-water geophysical and geological surveying and reconnaissance, midwater biological and water sampling, and collection of sea-floor sediments.

Some scientific organizations have combined the use of remotely operated and crewed vehicles to economize operations and at-sea time. For example, studies of hydrothermal vents by the crewed submersible *Alvin* were preceded by observing many square miles of sea floor with the towed vehicle *Angus* (which actually discovered the vents). *Alvin*, with its human observers, was subsequently deployed to conduct closeup observations and sampling, its primary role. Later, a small remotely operated vehicle, *Jason Jr.*, was used to swim off from *Alvin* to investigate places on the wreck that were too narrow or hazardous for the crewed vehicle to enter. This practice of using a towed vehicle

to perform a relatively high-speed, large-area reconnaissance and a crewed vehicle (sometimes with an on-board ROV) for final significant observations saves many days, perhaps months, of effort, since the crewed submersible is not an effective high-speed search and survey platform. *See* HYDROTHERMAL VENT.

The AUV has become a powerful new platform for oceanographic research where high-quality measurements are required but human presence is not. Eventually it should be possible for AUVs to make unattended transects of oceans, with the vehicle moving between a set depth and the surface while taking a variety of oceanographic measurements. Navigation updating and data transfer can be done by a radio link between the AUV and a satellite in space. As noted earlier, the major technical problem is getting enough power on board to support long-duration missions. In the 1990s, a Russian group proposed an AUV equipped with solar cells that can recharge batteries when the vehicle is on the surface transmitting data to a satellite. In 2003, a U.S. company put this type of AUV on the market.

In addition to operational AUVs, the future will bring crewed long-duration systems. Only one example exists at present, the U.S. Navy's *NR-1*. The small nuclear-powered *NR-1* can operate with mission times of up to a month.

U.S. Navy nuclear submarines have been used to do scientific research under the ice in the Arctic, the least-known ocean. However, the last of the fully ice-capable submarines was retired in 2000. Military submarines are not ideal for the long-duration scientific missions, while *NR-1* has relatively short endurance. The optimum platform would be a built-for-the-purpose nuclear submarine. In the early 1990s, a Canadian company and their Soviet Union partner proposed the development of *Ocean Shuttle*, a long-duration, multimission, crewed submarine analogous to NASA's space shuttle system. It was not built due to the end of the Cold War and a change in Canada's government. Such a submarine platform will be expensive to build and operate, but it offers the only means to do certain kinds of critical research in the world ocean. This type of underwater platform will eventually be developed because there are ocean work tasks that can only be done in this way.

Other uses. As the work capabilities of underwater vehicles have been proved, additional new opportunities have opened up. This is especially true in the uses of tethered, free-swimming ROVs. They now have a wide variety of civil uses such as fisheries resources management; inspection and maintenance of ships; law enforcement and public safety; marine pollution detection, remediation, and monitoring; treasure hunting; and underwater archeology. Many of these uses take advantage of vigorous price competition at the low end of the inspection vehicle market, where an ROV can be obtained for about $15,000. The evolution of the AUV will also follow a similar path to provide affordable systems to a wide community. While these economies are being realized, the top-end submersibles will increase in complexity, capability, and price. *See* SUBMARINE.

Don Walsh

Bibliography. American Bureau of Shipping, *Rules for Building and Classing Underwater Vehicles, Systems and Hyperbaric Facilities*, ABS, 2002; R. D. Ballard, *The Eternal Darkness: A Personal History of Deep-Sea Exploration*, Princeton University Press, 2000; W. J. Broad, *The Universe Below*, Simon & Schuster, New York, 1997; W. Forman, *The History of American Deep Submersible Operations*, Best Publishing, 1999; G. Griffiths (ed.), *Technology and Applications of Autonomous Underwater Vehicles*, CRC Press, 2002; *Jane's Underwater Technology 2005-06*, Janes Publishing, New York, 2005; Marine Board, National Research Council, *Underwater Vehicles and the National Needs*, National Academy Press, Washington, DC, 1996; G. Roberts et al. (eds.), *Guidance and Control of Underwater Vehicles 2003*, Elsevier, 2003; Underwater News & Technology, *Remote Operated Vehicles of the World 1998/99 Edition*, Technology Systems Corp., Stuart, FL, 1999.

Uninterruptible power supply

A source of high-quality alternating power to sensitive electrical equipment. An ideal uninterruptible power supply (UPS) provides a critical load with power when the main supply suffers from an over- or undervoltage condition causing interruption of the supply. It thus ensures power without break for the critical load. The UPS should be reliable and highly efficient, requiring low maintenance, low cost, and light weight. Uninterruptible power supplies were originally designed to furnish reliable and high-quality continuous power to large computer systems. They are now applied to a wide variety of critical equipment, such as medical facilities, life-support systems, financial transaction data storage and computer systems, telecommunications, industrial processing, and on-line management systems. UPS ratings range from a few hundred volt-ampere (VA) up to several thousand kilovolt-ampere (kVA). Three types of UPS systems are available: static, rotary, and hybrid static/rotary. *See* CIRCUIT (ELECTRICITY); ELECTRIC POWER TRANSMISSION.

Static UPS. Static UPS are available in ratings from 100 VA to 1 MVA, powering equipment ranging from personal computers and telecommunication systems, to medium power medical systems, to high power main systems. They are the most commonly used UPS systems, since they have high efficiency, high reliability, and low distortion. The disadvantages of static UPS are their poor performance with nonlinear and unbalanced loads, and high costs of achieving very high reliability. The principal configurations of static UPS systems are line-preferred, inverter-preferred, and line-interactive.

Line-preferred UPS. Line-preferred UPS are also referred to as standby or off-line UPS. This type of UPS is extensively employed to support different electrical loads. A line-preferred UPS is shown in **Fig. 1**. The system contains an ac/dc converter, a battery bank, a dc/ac converter, and static and transfer switches. There are three operating modes: normal mode, recovery mode, and backup mode.

In the normal mode of operation, power available on the ac line is normal, and the transfer and the static switches transfer power from the ac line to the load. The dc/ac converter and the load are in parallel. The dc/ac converter can be employed as an active filter by adding a simple control unit to the system, in the normal mode of operation. Hence, line-preferred UPS decreases the line current harmonics and improves the load power factor. Furthermore, the ac/dc converter can be controlled to reduce the harmonic content of the current on the dc side.

During recovery mode, the ac line directly feeds the load as well as the ac/dc converter to charge the battery. This mode can be activated with proper control strategy whenever needed. The ac/dc converter is not required to feed the load, because it is used only as a charger. Consequently, it allows low power rating, which makes it possible for the line-preferred UPS systems to be cheaper than the other static UPS topologies.

In backup mode, the ac line is not available, and the transfer switches disconnect the line from the ac/dc converter to prevent accidents. The dc/ac converter is turned on at the same time, and the battery unit supplies critical load via the converter until the main is within the preset tolerance again. The most important issue is the transfer time from normal mode to backup mode of operation. It is usually about $^1/_4$ line cycle, which is enough for most of the applications such as personal computers. *See* ALTERNATING CURRENT; CONVERTER; DIRECT CURRENT.

Inverter-preferred UPS. The inverter-preferred configuration, also called a double-conversion UPS or on-line UPS, is shown in **Fig. 2**. The system consists of an ac/dc converter (rectifier/charger), a battery unit, a dc/ac converter (inverter), and static (bypass) and transfer switches. The inverter-preferred UPS system has four operating modes: normal, recovery, stored energy, and bypass.

In the normal mode the transfer switch is on and the bypass switch is in the off position. The transfer switch continuously transfers power to the load from the dc/ac converter through the ac/dc converter. In this mode of operation, in contrast with the line-preferred UPS topology, the dc/ac and ac/dc converters are always on. Because double conversion occurs in this configuration, line conditioning is very good. The ac/dc converter is employed as a rectifier and charger for the load and battery. Hence, the ac/dc converter must be rated at 100% of the power demand of the load and battery unit. The dc/ac converter also must be at 100% of power load demand, since it is always on. Whenever the ac line is out of the tolerable limits, to avoid any undesirable power event between the ac line and the critical load, the transfer switch is off. This mode is

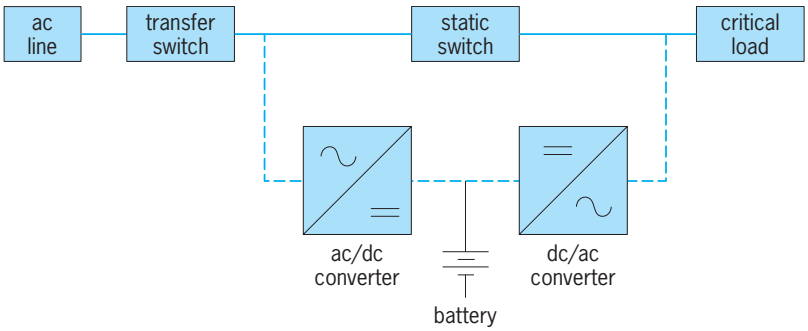

Fig. 1. Block diagram of a line-preferred UPS (off-line technology).

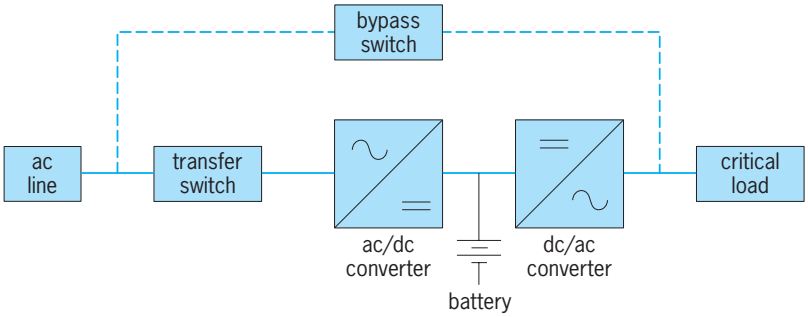

Fig. 2. Block diagram of an on-line UPS with bypass switch.

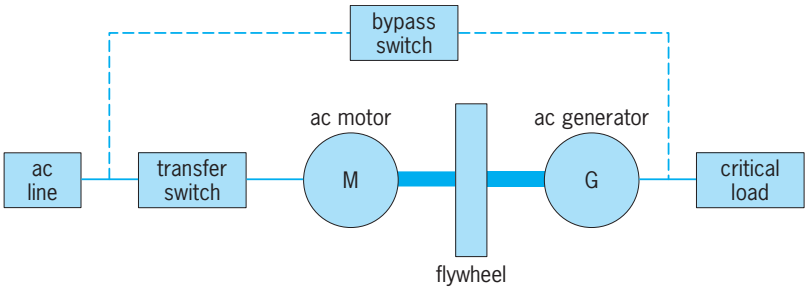

Fig. 3. Block diagram of the flywheel rotary UPS topology.

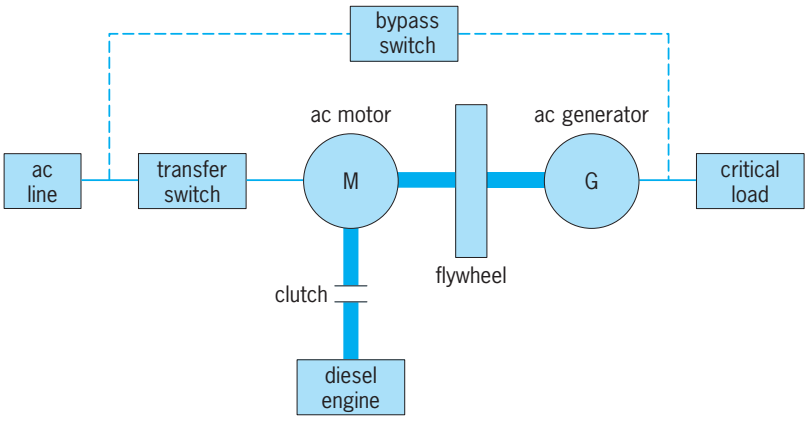

Fig. 4. Block diagram of diesel rotary UPS topology.

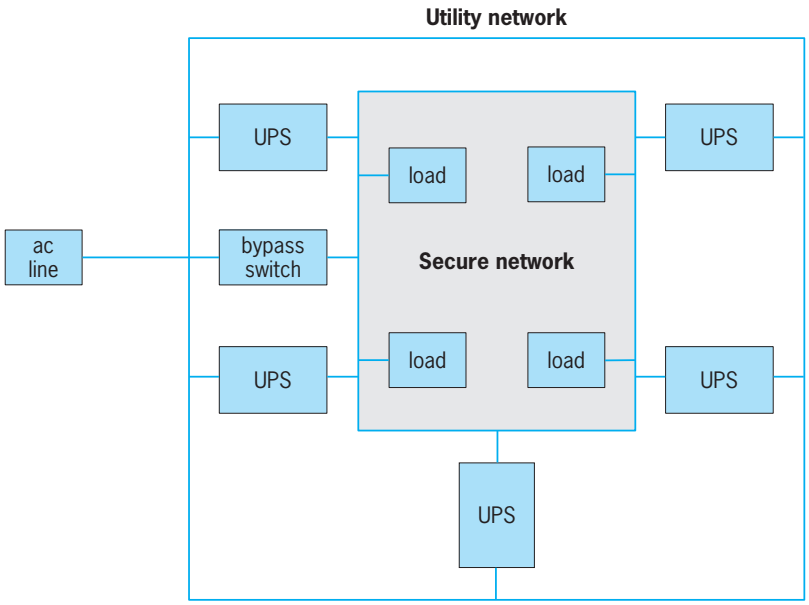

Fig. 5. Block diagram of the online distributed configuration.

known as backup mode. Since the dc/ac converter is always on, it continues to deliver power from the battery to the critical load without interruption. The UPS system returns to normal mode when the ac input voltage is within the preset tolerance again.

The main advantage of this configuration is that there is no transition time between normal and backup modes. Moreover, the inverter-preferred UPS system is more reliable than the line-preferred UPS topology; and, since the ac/dc converter is always on, the output frequency can be widely adjusted. Additionally, very wide tolerance to the input voltage differences and very good regulation of the output voltage can be realized by double-conversion topology.

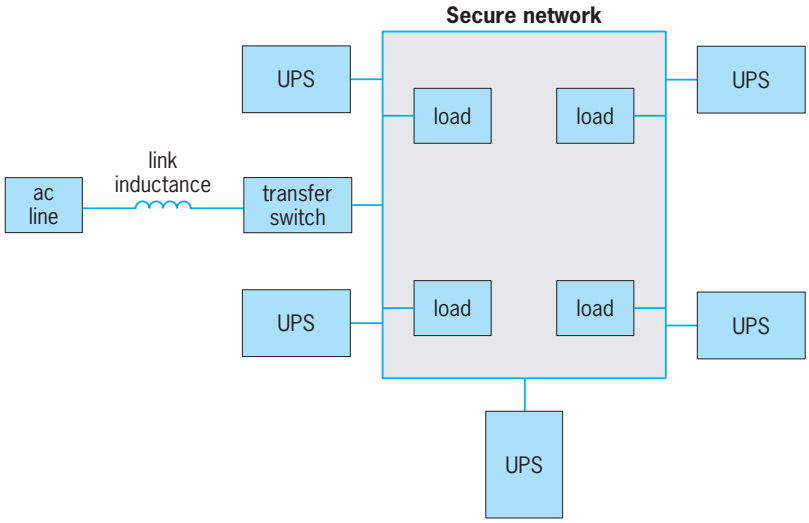

Fig. 6. Block diagram of line-interactive distributed application.

Line-interactive UPS. The Line-interactive UPS is also known as the grid-interactive UPS system. There are three types: single-conversion, in-line, and double-conversion topologies. They offer advantages such high efficiency and improved performance. For example, the single-conversion line-interactive UPS does not produce additional harmonics at the input side, a major advantage. High efficiency, low cost, and applicability in high power rating are the main advantages of the delta-conversion UPS configuration compared to the on-line UPS.

Rotary UPS. In rotary UPS systems a motor/generator set is used for energy conversion. There are two types of rotary UPS systems: flywheel rotary and diesel rotary. As shown in **Fig. 3**, a flywheel rotary UPS contains an ac motor, an ac generator, a flywheel unit, a static switch, and an optional bypass switch. The flywheel unit provides a mechanical coupling between the ac motor and ac generator, and is the energy storage component in this topology, rather than a battery as in the static UPS system. It provides dynamic energy for the generator and feeds the critical load through the generator when the ac line is out of preset tolerance. Restricted frequency range, restricted speed range, limitation on energy extraction and huge flywheel dimensions, even for a few seconds backup for the load, are the main disadvantages of this configuration.

The diesel rotary UPS is depicted in **Fig. 4**. The motor/generator set is connected to a diesel engine by means of a clutch. In normal mode, the ac line provides the power for the motor; the generator merely distributes the energy to the load. In backup mode, the ac motor is disconnected from the power line by the control unit, and the diesel engine begins to operate. It does not power the load at once, since it needs a certain amount of start-up time. The rotational energy stored in the flywheel unit provides interim power for the load until the diesel engine reaches its rated speed. After that, the clutch is engaged and the system operates as a diesel/generator set to supply the critical load. The flywheel unit does present disadvantages due to its diameter and weight. Nevertheless, the rotary UPS systems are definitely more reliable than static UPS systems. *See* ALTERNATING-CURRENT GENERATOR; CLUTCH; DIESEL ENGINE; FLYWHEEL; MOTOR.

Hybrid Static/Rotary UPS. Hybrid static/rotary UPS systems combine the features of both the static and rotary UPS systems. The main advantages of this configuration are better power quality with nonlinear loads, high reliability, frequency stability, better isolation, and low cost of maintenance are.

Applications. UPS systems provide uninterruptible and reliable power for the load including power conditioning in the system. There are two approaches to deploying UPS systems: distributed and centralized.

Distributed approach. Many parallel UPS units placed in an interconnected secure network feed the critical loads in a distributed system. Power circulates between the UPS units and the critical load

elastically. On-line and line-interactive approaches are two distributed schemes commonly used.

In the on-line distributed configuration, the secure and utility networks are connected to each other with the parallel UPS units, as shown in **Fig. 5**. There is isolation between the secure and utility networks which is provided by dc links of the UPS units. When the utility network is live, the critical loads are supplied by the ac line via the UPS units. The reactive current required by the secure network load is also provided by the UPS units. The utility network current harmonics are isolated from the secure network as well. Hence, the UPS units provide power conditioning and isolation for the system.

When the ac line is out of preset tolerance, the secure network is fed by the UPS batteries without any disruption. If one of the UPS units fails, in order to avoid power trouble, this unit is disconnected from the network instantly. Since power sharing is available between the parallel units, the other UPS units maintain power for the network, instead of failed UPS unit.

In the line-interactive distributed configuration, there is only one network instead of the utility and the secure networks, as shown in **Fig. 6**. Therefore, the system has a simpler structure compared to the online distributed topology, and the system cost is less, because the UPS units only work when the utility is in failure. Moreover, there is only one inverter in each UPS units. This characterization provides bidirectional power flow for the system. As mentioned before, there is no rectifier stage in the line-interactive distributed configuration; therefore the system is more reliable and more efficient than the online distributed approach.

When the utility is normal, the secure network is supplied by the ac mains. Simultaneously, the batteries are charged by the inverters. The reactive power demanded the secure network loads is also provided by the inverters. This configuration also achieves power line conditioning. Thanks to link inductance, the ac line absorbs all load changes with a delaying process (approximately 0.5 ms). The batteries feed the secure network loads in the backup mode of operation. The system has to provide seamless transition from the normal mode to the backup mode. To achieve this, the power angle is rapidly changed by the inverter from a negative value to a positive value. The inverter has a high switching frequency (10 kHz).

There are load sharing and system monitoring problems in this topology. Therefore, the system needs very fast digital controllers, which means high cost, a disadvantage of distributed UPS systems.

Centralized approach. Only one large UPS unit is employed in the centralized configuration, as shown in **Fig. 7**. The large UPS provides continuous operation for the whole system. This configuration is more desirable for industrial and utility applications, due to low maintenance requirements. To overcome the redundancy problem and to expand load capacity, additional costs are incurred in the centralized approach. Moreover, service and maintenance specialists are needed, also contributing to increase costs. However, such specially trained groups decrease the risk and increases the reliability of the system.

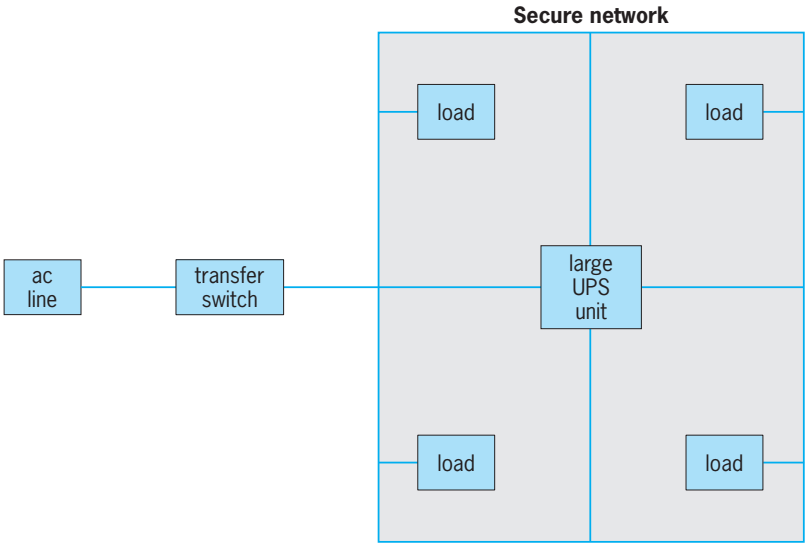

Fig. 7. Block diagram of the centralized approach.

Ayse E. Amac; Ali Emadi

Bibliography. S. B. Bekiarov and A. Emadi, Uninterruptible power supplies: Classification, operation, dynamics, and control, in *Proceedings of the 17th Annual IEEE Applied Power Electronics Conference and Exposition*, Dallas, TX, March 2002, pp. 597–604; M. C. Chandorkar, Control of distributed UPS systems, in *Proceedings of the 25th Annual IEEE Power Electronics Specialists Conference*, Taipei, Taiwan, June 1994, pp. 197–204; S. A. O. da Silva et al., A three phase line-interactive UPS system implementation with series-parallel active power line conditioning capabilities, *IEEE Trans. Ind. Applicat.*, 38(6):1581–1590, 2002; W. W. Hung and G. W. A. McDowell, Hybrid UPS for standby power systems, *Power Eng. J.*, 4(6):281–291, 1990; A. A. McLennan, Static UPS technologies, *IEE Colloquium on Uninterruptible Power Supplies*, London, February 1994, pp. 2/1–2/5; S. R. Philpott, Large scale UPS installations, in *Proceedings of the IEEE International Conference on Electrical Installation Engineering in Europe*, London, June 1993, pp. 64–68; J.-C. Wu and H.-L. Jou, A new UPS scheme provides harmonic suppression and input power factor correction, *IEEE Trans. Ind. Elec.*, 42(6):629–635, 1995.

Unit operations

A structure of logic used for synthesizing and analyzing processing schemes in the chemical and allied industries, in which the basic underlying concept is that all processing schemes can be composed from and decomposed into a series of individual, or unit, steps. If a step involves a chemical change, it is called a unit process; if physical change, a unit operation.

These unit operations are utilized for certain definite functions wherever they are employed, and their use cuts across widely different processing applications, including industries as diverse as manufacture of chemicals, fuels, pharmaceuticals, pulp and paper, processed foods, and primary metals. The unit operations approach thereby serves as a very powerful form of morphological analysis, which systematizes process design, and greatly reduces both the number of concepts that must be taught and the number of possibilities that should be considered in synthesizing a particular process.

The selection and classification of unit operations is somewhat arbitrary. Filtration is generally thought of as a single unit operation, but in reality there are a number of quite different types of filters. Distillation, solvent extraction, and adsorption are usually regarded as different unit operations, but all can serve to fulfill the same function of separating a liquid mixture into two or more products of different composition. Distillation accomplishes this through differences in volatility, solvent extraction utilizes different solubilities into an immiscible solvent, and adsorption is based on different affinities for a solid surface. All fall within the general category of separation processes, which is itself a focused area of investigation. Different separation techniques are useful or preferred for different mixtures and purposes. The invention and selection of separation techniques is an important dimension of process synthesis, along with selection among, and assemblage of, other unit operations.

Most of the unit operations are based mechanistically upon the fundamental transport processes of mass transfer, heat transfer, and fluid flow (momentum transfer). Unit operations based on fluid mechanics include fluid transport (such as pumping), mixing/agitation, filtration, clarification, thickening or sedimentation, classification, and centrifugation. Operations based on heat transfer include heat exchange, condensation, evaporation, furnaces or kilns, drying, cooling towers, and freezing or thawing. Operations that are based on mass transfer include distillation, solvent extraction, leaching, absorption or desorption, adsorption, ion exchange, humidification or dehumidification, gaseous diffusion, crystallization, and thermal diffusion. Operations that are based on mechanical principles include screening, solids handling, size reduction, flotation, magnetic separation, and electrostatic precipitation. The study of transport phenomena provides a unifying and powerful basis for an understanding of the different unit operations. *See* CHEMICAL ENGINEERING; TRANSPORT PROCESSES; UNIT PROCESSES.

For further information concerning operations based on fluid mechanics *see* CENTRIFUGATION; CLARIFICATION; FILTRATION; MIXING; PUMP; PUMPING MACHINERY; SEDIMENTATION (INDUSTRY); THICKENING. For operations based on heat transfer *see* COOLING TOWER; DRYING; EVAPORATION; HEAT EXCHANGER; KILN. For operations based on mass transfer *see* ABSORPTION; ADSORPTION; CRYSTALLIZATION; DEHUMIDIFIER; DIFFUSION; DISTILLATION; HUMIDIFICATION; ION EXCHANGE; LEACHING; SOLVENT EXTRACTION. For operations based on mechanical principle *see* ELECTROSTATIC PRECIPITATOR; FLOTATION; MAGNETIC SEPARATION METHODS; MECHANICAL SEPARATION TECHNIQUES; SOLIDS PUMP.

C. Judson King

Bibliography. W. F. Furter (ed.), *A Century of Chemical Engineering*, 1982; A. Grandison and M. Lewis, *Separation Processes*, 1995; D. Hagerty, E. Gearhard, and C. Plank, *Chemical Engineering*, 1989; C. J. King, *Separation Processes*, 2d ed., 1980; W. L. McCabe and J. C. Smith, *Unit Operations of Chemical Engineering*, 5th ed., 1993; R. Reynolds, *Unit Operations and Processes in Environmental Engineering*, 2d ed., 1996.

Unit processes

Processes that involve making chemical changes to materials, as a result of chemical reaction taking place. For instance, in the combustion of coal, the entering and leaving materials differ from each other chemically: coal and air enter, and flue gases and residues leave the combustion chamber. Combustion is therefore a unit process. Unit processes are also referred to as chemical conversions.

Together with unit operations (physical conversions), unit processes (chemical conversions) form the basic building blocks of a chemical manufacturing process. For instance, reactants are mixed together (unit operation) and allowed to react (unit process), and the product stream is separated into its constituents which then may or may not be purified further (various unit operations). Most chemical processes therefore consist of a combination of various unit operations and unit processes.

The basic tools of the chemical engineer for the design, study, or improvement of a unit process are the mass balance, the energy balance, kinetic rate of reaction, and position of equilibrium (the last is included only if the reaction does not go to completion). Together these are popularly called the four horsemen of chemical engineering, and the study or design is not complete unless all four have been addressed.

A list of unit processes can be neither rigid nor permanent. New ones are added as they are developed. A partial list follows (see separate articles on the items with asterisks):

Acetylation*
Acidolysis
Acidulation
Acylation*
Addition reaction*
Alcoholysis*
Alkali fusion reaction*
Alkylation*
Amination*
Ammoniation
Ammonolysis
Aromatization*
Calcination
Carboxylation
Causticization
Chelation*
Combustion*
Condensation reaction*
Consolidation reaction
Corrosion
Cyanoethylation*
Cyclization

Dealkylation
Decomposition
Dehydration reaction
Dehydrogenation*
Diazotization* and coupling
Double decomposition
Electrolysis*
Epoxidation*
Esterification
Etherification
Fermentation*
Fischer-Tropsch process*
Friedel-Crafts reaction*
Grignard reaction*
Haloform reaction*
Halogenation*
  bromination
  chlorination
  fluorination
  iodination
Halogenolysis
  brominolysis
  chlorinloysis
  fluorinolysis
  iodinolysis*
Hydration*
Hydroboration*
Hydrodealkylation
Hydroformylation*
Hydrogenation*
Hydrogenolysis
Hydrolysis*
Hydrorefining
  hydrodesulfurization
  hydrofining
Hydroxylation reaction*
Interesterification
Ion exchange*
Isomerization*
Methanation
  methanization
Neutralization
Nitration
Nitrogen fixation*
Oxidation process*
  also bleaching*
  detonation
  explosion and explosive*
Oxoformylation
Ozonolysis*
Photolysis*
Photosynthesis*
Polymerization*
  addition polymerization
  condensation
  polymerization
  dimerization
  photopolymerization
  trimerization
Polyoxyethylation
Pyrolysis*
  carbonization
  coking
  cracking
  destructive distillation
  dry distillation
Racemization*
Reduction
Reforming processes*
  hydroforming
  platforming
Saccharification
Saponification
Silicate formation
Substitution reaction*
Sulfonation*
Transesterification*
Vulcanization

See CHEMICAL ENGINEERING; UNIT OPERATIONS.

W. F. Furter

Unitary symmetry

A type of symmetry law, an important example of which is flavor symmetry, an approximate internal symmetry law obeyed by the strong interactions of elementary particles. According to the successful theory of strong interactions, quantum chromodynamics, flavor symmetry is the consequence of the fact that the strong force (the so-called glue force, mediated by the SU_3^{color} gauge field) is the same between any two quarks, the constituents of hadrons. It follows that if all the kinds (flavors) of quarks had the same mass, strong interactions would have the symmetry SU_N^{flavor}, where N is the number of quark flavors. As a result, if nonstrong (electromagnetic and weak) interactions were neglected, then hadrons would occur as degenerate (all having the same mass) multiplets (irreducible representations) of the group SU_N. *See* COLOR (QUANTUM MECHANICS); FLAVOR; HADRON; QUANTUM CHROMODYNAMICS; QUARKS.

Spin independence. An example of unitary symmetry is the spin independence of electric forces on slowly moving electrons. The electron has spin $^1/_2$ and thus two spin states, referred to as spin-up and spin-down. Denoting the wave functions of these two states by $|u\rangle$ and $|d\rangle$, the spin-independent physical properties (energy eigenvalues, charge density, and so on) of a system of electrons such as those in an atom, in which the forces on electrons are nearly spin independent, are unchanged by the replacements shown in Eqs. (1), where α and β are

$$\begin{aligned} |u\rangle &\to \alpha|u\rangle + \beta|d\rangle \\ |d\rangle &\to -\beta^*|u\rangle + \alpha^*|d\rangle \\ |\alpha|^2 + |\beta|^2 &= 1 \end{aligned} \tag{1}$$

complex numbers. These transformations form a group known as SU_2. They correspond to rotations of space; α and β can be expressed in terms of the three numbers which describe the rotation. It is easily seen that many-electron states decompose under the transformation; for example, using Eqs. (1), the two-electron state $(|u, d\rangle - |d, u\rangle)/\sqrt{2}$ is unchanged, and the three remaining two-electron states, $|u, u\rangle$, $|d, d\rangle$, and $(|u, d\rangle + |d, u\rangle)/\sqrt{2}$, transform to linear combinations of themselves. This is the decomposition into singlet and triplet spin states, that is, into total spin $S = 0$ and 1, respectively; the nonmixing between them is equivalent to the invariance of S to rotation. *See* ANGULAR MOMENTUM; ELECTRON SPIN; SPIN (QUANTUM MECHANICS).

SU_2 flavor symmetry (charge independence). The first flavor symmetry encountered in particle physics was the near equivalence of the proton and neutron in nuclear physics. This is called charge independence, *i*- spin (isotopic spin) invariance, or (in terms of the quark model) flavor SU_2 symmetry. The nucleons, proton and neutron, both have spin $^1/_2$ and have nearly the same mass (differing by only 0.14%), and the nuclear force between any two nucleons—whether two protons, a proton and a neutron, or two neutrons—is nearly the same. This is just like the approximate spin independence of the force between two electrons and is describable in the same way, as an SU_2 symmetry, Eqs. (1). Hence one speaks of the proton and neutron as being the two *i*-spin states of a nucleon, a particle with *i*-spin $^1/_2$. Conventionally, the proton is the *i*-spin up state and the neutron is *i*-spin down. For real spin in empty space, spin-up and spin-down are defined with respect to an arbitrary direction [if the direction is changed, the spin-up and spin-down states change by the transformation Eqs. (1)], but for *i*-spin electromagnetism permanently singles out a direction in "*i*-spin space," the nucleon's *i*-spin up state (proton) having charge e and its *i*-spin down state (neutron) having charge zero. A real-spin analog of this would be for a spin-$^1/_2$

particle carrying a magnetic moment to be always in a constant magnetic field.

The states of nuclei (bound systems of nucleons) can be approximately described as having a definite total *i*-spin (I), but since the electrostatic energy of a nucleus rises with the square of the number of its protons (Z) this description rapidly becomes worse for higher-Z nuclei.

The *i*-spin symmetry extends to all the strongly interacting particles, both baryons and mesons, all of which have *i*-spin as one of their properties. An unsuccessful attempt was made to explain this universality by supposing, for example, that the pion (pi meson), which has *i*-spin 1, was a nucleon-antinucleon bound state. But the universality is now explained in the quark model of hadrons by the fact that the two least massive quarks, the up (u) and down (d) quarks, are similar (analogous to the similarity of proton and neutron). The masses of the u and d quarks are estimated at roughly 3 MeV and 6 MeV, respectively (the neutron is heavier than the proton because the d is heavier than the u). Despite this mass ratio of 2 the u and d really do behave similarly in a hadron, because they have there a zero-point energy of the order of 200 or 300 MeV, and so the effect of their masses on their dynamics is very small.

The *i*-spin of any hadron is thus determined by its u and d quarks. For instance, the lightest (*s*-wave) $J^P = {}^3/_2{}^+$ baryons contain three quarks in a symmetric flavor state (this is required to make the overall state of the three fermions symmetric under exchange), and therefore they have *i*-spin ${}^3/_2$, 1, ${}^1/_2$, 0 according to whether 3, 2, 1, 0 of their quarks are u or d. Similarly, exchange symmetry implies that the lightest $J^P = {}^1/_2{}^+$ baryons have *i*-spin ${}^1/_2$, 0 or 1, or ${}^1/_2$ according to whether 3, 2, or 1 of their quarks are u or d. *See* I-SPIN.

SU_3 flavor symmetry. The third lightest quark is the strange quark (s), with mass about 0.15 GeV. It is named in reference to the "strange particles" (hyperons and K mesons), which were found to be produced in high-energy collisions almost as copiously as pions but did not decay rapidly into nucleons and pions despite being massive enough to do so. In terms of the quark model, one of the quarks (or antiquarks) of a strange particle is an s quark (or anti-s quark). (There are also doubly and triply strange baryons, in which two or all three of the quarks are s quarks.) In the decay of a strange particle into nonstrange particles (or of a doubly strange particle into a singly strange particle, and so on) an s quark must change into a u or d; this can happen only (as with any flavor change of a quark) by a weak interaction, not a strong interaction, and so the decay is slow. This is similar to ordinary beta decay of a neutron, which involves the weak-interaction change of a d quark into a u. *See* WEAK NUCLEAR INTERACTIONS.

If the s quark were as light as a u or d quark, then these three kinds would be equivalent (as far as the strong interactions were concerned), and any flavor-independent physical properties (energy eigenvalue, for instance) of a state of quarks would be unchanged by replacements similar to the SU_2 transformation shown in Eqs. (1) but with a third state, $|s\rangle$, involved in addition to $|u\rangle$ and $|d\rangle$, and with the complex coefficients specified by eight (rather than three) real numbers. These symmetry transformations form a group known as SU_3 [SU_N denotes the group of $N \times N$ unimodular (determinant = 1) unitary matrices]. The triplet of states $|u\rangle$, $|d\rangle$, and $|s\rangle$, equivalent under the SU_3 symmetry, is analogous to the doublet $|u\rangle$ and $|d\rangle$, equivalent under the SU_2 (*i*-spin) symmetry. [It is not to be confused with an *i*-spin triplet of SU_2. The group of rotations of space is essentially the same group as the group SU_2, which is the reason for the exact analogy described above between a real-spin doublet and a flavor SU_2 doublet (both denoted by $|u\rangle$ and $|d\rangle$), and for the use of the term *i*-spin. SU_3 is different.] The baryons described at the end of the previous section provide examples of SU_3 multiplets if the quarks in the baryon are u, d, or s, namely the octet and the decuplet (**Table 1**). *See* STRANGE PARTICLES.

TABLE 1. Lightest (*s*-wave) baryons made of *u*, *d*, or *s* quarks

Quark content*	$J^P = {}^1/_2{}^+$ *I*-spin	$J^P = {}^1/_2{}^+$ Names†	$J^P = {}^3/_2{}^+$ *I*-spin	$J^P = {}^3/_2{}^+$ Names
nnn	${}^1/_2$	$N^{0,+}$	${}^3/_2$	$\Delta^{-,0,+,++}$
nns	0; 1	Λ^0; $\Sigma^{-,0,+}$	1	$\Sigma^{-,0,+}$
nss	${}^1/_2$	$\Xi^{-,0}$	${}^1/_2$	$\Xi^{-,0}$
sss			0	Ω^-
		[$SU_3(8)$]		[$SU_3(10)$]

*n denotes a nucleon-type quark, that is, u or d.
†N^0, N^+ are neutron, proton.

Because the s quark is significantly heavier than the u and d quarks, the SU_3 (u, d, s) symmetry is only approximate; the masses of the hadrons belonging to an SU_3^{flavor} multiplet differ from one another by the order of 100 or 200 MeV. Sometimes, as in the lightest $J^p = 1^-$ meson and ${}^3/_2{}^+$ baryon multiplets, the masses depend nearly linearly on the number of s quarks contained, as one would naively expect. In other cases, such as the lightest ${}^1/_2{}^+$ baryon multiplet, the mass of the s also has an indirect effect on the mass of the baryon through the strength of the spin-spin color interaction (the color analog of the hyperfine interaction, that is, the interaction of magnetic moments) because the color-magnetic moment of the heavier s is smaller than that of the u or d, just as for ordinary magnetic moments; this is also the explanation of the difference of the Λ and Σ masses.

The SU_3^{flavor} symmetry implies relations between strong coupling constants (for example, $\pm\sqrt{3}g_{KN\Lambda} \pm g_{KN\Sigma} = 2g_{\pi NN}$) and between cross sections; those which are testable experimentally seem to agree. There are also many relations among the matrix elements of electromagnetic and weak current operators. For instance, the magnetic moments of the ${}^1/_2{}^+$ baryon octet are predicted to have the relations given by Eqs. (2), according to which all moments of the octet members follow from the

TABLE 2. The lightest (s-wave) baryons made of *u*, *d*, *s*, or *c* quarks

Quark content*	$J^P = 1/2^+$ SU$_3$ multiplet†	$J^P = 1/2^+$ Names	$J^P = 3/2^+$ SU$_3$ multiplet†	$J^P = 3/2^+$ Names
lll	(8)	N, Λ, Σ, Ξ	(10)	$\Delta, \Sigma, \Xi, \Omega$
llc	$(\bar{3})$; (6)	$\Lambda_c, \Xi_c; \Sigma_c, \Xi'_c, \Omega_c$	(6)	$\Sigma_c, \Xi_c, \Omega_c$
lcc	(3)	Ξ_{cc}, Ω_{cc}	(3)	Ξ_{cc}, Ω_{cc}
ccc			(1)	Ω_{ccc}
	[SU$_4$(20′)]		[SU$_4$(20)]	

*l denotes a light quark, that is, *u*, *d*, or *s*.
†The SU$_3$ multiplets (1), (3), $(\bar{3})$, (6), (8), and (10) are also called singlet, triplet, antitriplet, sextet, octet, and decuplet, respectively.

proton and neutron moments.

$$\begin{aligned} \mu(\Sigma^+) &= \mu(p) \\ 2\mu(\Lambda) = \mu(\Xi^0) &= -2\mu(\Sigma^0) = \mu(n) \\ \mu(\Sigma^-) = \mu(\Xi^-) &= -[\mu(p) + \mu(n)] \end{aligned} \tag{2}$$

These relations are found to be qualitatively correct, although the effect of the smaller magnetic moment of the *s* quark is clearly seen: For example, twice the ratio of the moments of the Λ and the neutron is 0.64; according to the naïve quark model, this is the ratio of the moments of the *s* and *d* quarks. *See* BARYON; MESON; WEAK NUCLEAR INTERACTIONS.

Just as the SU$_2$ (*u, d* or *i*-spin) symmetry applies to all hadrons, so does the (more approximate) SU$_3$ (*u, d, s*) symmetry. For instance, the lightest (*s*-wave) baryons whose quarks are *u*, *d*, *s*, or *c* (the next heavier quark) are described in **Table 2**. [In the naming scheme of these baryons, the capital letters specify the *i*-spin as in the SU$_3$(8) and SU$_3$(10) multiplets of Table 1. The *c* subscripts specify how many *s* quarks have been replaced by *c* quarks; each such replacement raises charge by *e* (the charge difference of *c* and *s*). For example, whereas the members of a Ξ multiplet are $\Xi^{-,0}$, those of a Ξ_{cc} are $\Xi_{cc}{}^{+,++}$. The two $1/2^+$ baryons with identical quantum numbers, Ξ_c and Ξ'_c, have different spin-flavor structure and differ in mass (by 0.11 GeV) because the color-magnetic moments of the *n*, *s*, and *c* quarks differ from one another.] Of these lightest $1/2^+$ and $3/2^+$ baryons, only those with more than one *c* quark have not yet been seen.

SU$_{4,5,6}$ flavor symmetry. The lightest of the three so-called heavy quarks, the charmed quark (*c*), has a mass of about 1.25 GeV. This large mass badly breaks the SU$_4$ flavor symmetry that would hold if *u, d, s*, and *c* all had the same mass. Although the lightest $1/2^+$ and $3/2^+$ baryons listed in Table 2 can each be described as comprising a 20-member multiplet of SU$_4$, SU$_4$ symmetry is not a useful starting point to relate properties of the multiplet members. The remaining heavy quarks, the bottom (*b*) and top (*t*) quarks, are heavier still with masses of about 4.5 and 175 GeV, respectively, and so SU$_5$ and SU$_6$ flavor symmetries are even more pointless to consider. In fact the top quark is so massive that it decays (weakly) before it has a chance to be a constituent of any well-defined hadron, its mean life being less than 10^{-24} s. *See* CHARM; J/PSI PARTICLE; UPSILON PARTICLES.

Chiral U$_N^L$ U$_N^R$ symmetry. If quarks were massless, their chirality would be a good quantum number, conserved by the glue interaction, and the flavor symmetry group would be larger, namely $U_N^L \times U_N^R$, where U_N^L is the flavor symmetry group of the left-handed quarks and U_N^R correspondingly for the right-handed quarks. Not only is this symmetry broken by the quark masses, but there is evidence (partly from numerical lattice quantum chromodynamics calculations) that it is spontaneously broken. In the approximate model where the masses of the *u* and *d* quarks are taken to be zero and the heavier quarks are ignored, the consequences of the spontaneous symmetry breaking of the $U_2^L \times U_2^R$ symmetry down to U$_2$ were worked out, in the σ model and current algebra, even before the quark model was developed, with results such as the Goldberger-Treiman relation between the strong and weak coupling constants of the charged π meson and the Adler-Weisberger relation between the Gamow-Teller and Fermi weak coupling constants of the nucleon.

Other unitary symmetries. There are unitary symmetries that are not purely internal in the strict sense, namely those in which it is supposed that the interaction between the fundamental particles is spin-independent (as for nonrelativistic electrons interacting with electrostatic forces). This was suggested by E. Wigner as an approximate symmetry in light nuclei; there the fundamental particle is a quartet, with states $|p\uparrow\rangle, |p\downarrow\rangle, |n\uparrow\rangle, |n\downarrow\rangle$(where $|p\uparrow\rangle$ means a proton with spin up, and so on), and the symmetry group is therefore SU$_4$. Similarly, SU$_6$ symmetry, resulting from combining flavor SU$_3$ with spin independence, is sometimes used in the quark model. It is known that such unitary symmetries can never be exact in a relativistic theory, but that fact does not prevent them from being approximately true for low-lying states. *See* ELEMENTARY PARTICLE.

Charles J. Goebel

Bibliography. M. Gell-Mann and Y. Ne'eman (eds.), *The Eightfold Way*, rev. ed., Perseus Publishing, 2000; G. Kane, *Modern Elementary Particle Physics*, Westview Press, 1993; D. B. Lichtenberg, *Unitary Symmetry and Elementary Particles*, 2d ed., Academic Press, 1978; U. Mosel, *Fields, Symmetries, and Quarks*, 2d ed., Springer-Verlag, 1999; L. H. Ryder, *Elementary Particles and Symmetries*, rev. ed., Gordon and Breach, 1986; N. P. Samios, M. Goldberg, and B. T. Meadows, Hadrons and SU(3): A critical review, *Rev. Mod. Phys.*, 46:49–81, 1974.

Units of measurement

Values, quantities, or magnitudes in terms of which other such are expressed. Units are grouped into systems, suitable for use in the measurement of physical quantities and in the convenient statement of laws relating physical quantities. A quantity is a measurable attribute of phenomena or matter.

A given physical quantity A, such as length, time, or energy, is the product of a numerical value or measure $\{A\}$ and a unit $[A]$. Thus Eq. (1) holds.

$$A = \{A\}[A] \tag{1}$$

The unit $[A]$ can be chosen arbitrarily, but it is desirable to define units in such a way that they are derived from a few base units by equations without numerical factors other than unity, and that the equations between numerical values of quantities have exactly the same form as the equations between the quantities. For example, the kinetic energy E of a body is given in terms of its mass M and speed V by Eq. (2), where $E = \{E\}\ [E]$, $M = \{M\}\ [M]$, $V =$

$$E = {}^1/_2 MV^2 \tag{2}$$

$\{V\}\ [V]$, and $^1/_2$ is called a definitional factor and is dimensionless. If the units of E, M, and V are defined in such a way that Eq. (3) holds, then the equation between the numerical values is Eq. (4). A system

$$[E] = [M][V]^2 \tag{3}$$

$$\{E\} = {}^1/_2\{M\}\{V\}^2 \tag{4}$$

of units defined in this way is called a coherent system. It is constructed by defining the units of a few base quantities independently; these are called base units. The units of all other quantities are defined by equations similar to Eq. (3) with no numerical factors other than unity, and are called derived units. It is a matter of choice which quantities are to be considered as the base quantities and also how many. It is desirable to have as few base quantities as possible, provided no serious inconveniences are encountered. *See* DIMENSIONAL ANALYSIS.

In 1960 the General Conference on Weights and Measures (CGPM) gave official status to a single practical system, the International System of Units, abbreviated SI in all languages. The system is a modernized version of the metric system. The SI, as subsequently extended, includes seven base units, two supplementary units, and nineteen derived units with special names. These derived units, and others without special names, are derived from the base and supplementary units in a coherent manner. A set of prefixes is used to form decimal multiples and submultiples of the SI units. *See* METRIC SYSTEM.

The CGPM and its subsidiary body, the International Committee for Weights and Measures (CIPM), have accepted certain units which are not part of the SI but which are widely used or are useful in specialized fields, for use with the SI or for temporary use in those fields. For definitions of the seven SI base units, discussion of the physical standards for realizing these units, and tables of the SI base, supplementary, and derived units and units accepted for use with the SI *see* PHYSICAL MEASUREMENT.

This article discusses how the various derived units are defined in terms of the base units, both in the SI and in other coherent systems, and how units in different systems, as well as arbitrary units which are not part of any system, are related. Various units which are not accepted for use with the SI, but whose use is nevertheless common, are discussed. However, the use of all units which are not SI units or specifically accepted for use with the SI has been discouraged, and whenever possible, units outside the SI should be replaced by SI units or by their multiples and submultiples formed by attaching SI prefixes.

Geometrical Units

Units of plane angle and solid angle are purely geometrical. The SI units of plane and solid angle are regarded as dimensionless derived units.

Plane angle units. The radian (rad), the SI unit of plane angle, is the plane angle between two radii of a circle which cut off on the circumference an arc equal in length to the radius. Since the circumference of a circle is 2π times the radius, the complete angle about a point is 2π rad. *See* RADIAN MEASURE.

The degree and its decimal submultiples can be used with the SI when the radian is not a convenient unit. By definition, 2π rad $= 360^\circ$. The minute $[1' = (1/60)^\circ]$ and the second $[1'' = (1/60)']$ can also be used.

Steradian. The steradian (sr), the SI unit of solid angle, is the solid angle which, having its vertex at the center of a sphere, cuts off an area on the surface of the sphere equal to that of a square with sides of length equal to the radius of the sphere. Since the total area of a sphere is 4π times the square of its radius, the complete solid angle about a point is 4π sr.

Mechanical Units

In mechanics, it is convenient to have three base quantities, and two of these are generally chosen to be length and time. Systems of mechanical units may be classified as absolute systems, in which the third base quantity is mass, and gravitational systems, in which the third base quantity is force.

Absolute systems. Two absolute systems of metric units are commonly employed, each named for its base units of length, mass, and time: the mks (meter-kilogram-second) absolute system, and the cgs (centimeter-gram-second) absolute system. The mks absolute system is the mechanical portion of the SI. Once the length, mass, and time units are selected, Newton's second law in the form $F = ma$ is used to define the unit of force F in terms of mass (m) units and acceleration (a) units, which involve length and time units. The unit of force derived in this way through Newton's second law is called the newton in the mks system and the dyne in the cgs system. Units of work W are defined by the relation

TABLE 1. Absolute systems of units

Quantity	Defining relation	Units: mks (SI)	cgs	British
Length (L)	Fundamental	m	cm	ft
Mass (m)	Fundamental	kg	g	lb
Time (t)	Fundamental	s	s	s
Velocity (v)	$v = L/t$	m/s	cm/s	ft/s
Acceleration (a)	$a = v/t$	m/s^2	cm/s^2	ft/s^2
Force (F)	$F = ma$	$1\ N = 1\ kg \cdot m/s^2$	$1\ dyne = 1\ g \cdot cm/s^2$	$1\ pdl = 1\ lb \cdot ft/s^2$
Work (W)	$W = FL$	$1\ J = 1\ N \cdot m$	$1\ erg = 1\ dyne \cdot cm$	ft · pdl
Power (P)	$P = W/t$	1 W = 1 J/s	erg/s	ft · pdl/s
Momentum (p)	$p = mv$	kg · m/s	g · cm/s	lb · ft/s

$W = FL$, where L represents a displacement in the direction of a force F. Units of power P in turn are defined by the relation $P = W/t$ and are therefore the ratio of work units to units of time t. A coherent absolute system of British units is based on the foot, the pound (1 lb $\cong$ 0.4536 kg), and the second; and the derived unit of force is called the poundal (pdl), a unit which is not often used. Various quantities expressed in the three absolute systems are listed in **Table 1**. *See* KINETICS (CLASSICAL MECHANICS); NEWTON'S LAWS OF MOTION.

Gravitational systems. Gravitational systems, in which the base quantities are length, force, and time, have been frequently employed by engineers, and are therefore sometimes called technical systems.

In the British gravitational system the standard force, called the pound force (lbf), is the gravitational force exerted on a 1-lb mass at a location on the Earth's surface where the acceleration of a freely falling body due to gravity has the standard value $g_n = 9.80665\ m/s^2 \cong 32.174\ ft/s^2$. More simply stated, 1 lbf is the force which imparts an acceleration of magnitude g_n to a 1-lb mass. A coherent system is obtained by using Newton's second law, $F = ma$, to define a mass unit called the slug (1 lbf = 1 slug · ft/s^2). The work unit in this system is the foot-pound force (ft · lbf). *See* FREE FALL.

In the mks gravitational system the unit of acceleration is the meter per second squared (m/s^2), and the unit of force is taken to be the Earth's gravitational force on a 1-kg mass at a location where this force imparts an acceleration of magnitude g_n. Unfortunately, this unit of force has also been called a kilogram, but to distinguish it from the unit of mass it should be called kilogram force (kgf). There is no generally accepted name for the mks gravitational unit of mass, although the name metric slug has been used. The cgs gravitational system is constructed in a similar manner. Various quantities expressed in the three gravitational systems are listed in **Table 2**.

Noncoherent gravitational systems in which force is expressed in pounds force (lbf) or a comparable kilogram force (kgf) or gram force (gf), and mass in pounds, kilograms, or grams, can be set up by writing Newton's second law in the form $g_nF = ma$.

Length units. The meter (m) is the SI base unit of length. The use of special names for decimal submultiples of the meter should be avoided, and units formed by attaching appropriate SI prefixes to the meter should be used instead. Thus the micron (μ), which was defined as 10^{-6} m, should be replaced by the micrometer (μm), which has the same value; and the millimicron (mμ), which was defined as 10^{-9} m, should be replaced by nanometer (nm). The fermi, which was defined as 10^{-15} m, and was used to measure nuclear distances, should be replaced by the femtometer (fm), which has the same value.

The angstrom (Å) is equal to 10^{-10} m. Although it has been accepted for temporary use with the SI, it is preferable to replace this unit with the nanometer, using the relation 1 Å = 0.1 nm.

The nautical mile (nmi), equal to 1852 m, has been accepted for temporary use with the SI in navigation. (This unit approximates the length of 1 minute of arc of a great circle on the surface of the Earth.) *See* NAVIGATION.

The foot (ft) is, as discussed above, the unit of length in the British systems of units, and it is also

TABLE 2. Gravitational systems of units

Quantity	Defining relation	Units: mks	cgs	British
Length (L)	Fundamental	m	cm	ft
Force (F)	Fundamental	kilogram force (kgf)	gram force (gf)	pound force (lbf)
Time (t)	Fundamental	s	s	s
Mass (m)	$m = F/a$	$kgf \cdot s^2/m$	$gf \cdot s^2/cm$	$1\ slug = 1\ lbf \cdot s^2/ft$
Velocity (v)	$v = L/t$	m/s	cm/s	ft/s
Acceleration (a)	$a = v/t$	m/s^2	cm/s^2	ft/s^2
Work (W)	$W = FL$	m · kgf	cm · gf	foot-pound force (ft · lbf)
Power (P)	$P = W/t$	m · kgf/s	cm · gf/s	ft · lbf/s
Momentum (p)	$p = mv$	kgf · s	gf · s	slug · ft/s

in customary use in the United States. Since 1959 the foot has been defined as exactly 0.3048 m. The yard (yd) is defined as exactly 3 ft or 0.9144 m. *See* LENGTH; MEASURE.

Units used to measure x-ray wavelengths. Relative measurements of x-ray wavelengths can be made to a higher accuracy than absolute measurements. That is, the ratio of two x-ray wavelengths can be determined with a higher accuracy than the ratio of either of them to the meter. The same situation holds for dimensions of crystal lattices, which are derived from x-ray wavelengths by x-ray diffraction experiments. For this reason, x-ray wavelengths and dimensions of crystal lattices have been expressed in units that are defined in terms of a standard x-ray wavelength or crystal lattice dimension. *See* X-RAY CRYSTALLOGRAPHY; X-RAYS.

Before 1965, most x-ray wavelengths were expressed in terms of the X-unit, which is approximately 10^{-13} m. The grating constant of calcite was defined to be exactly 3029.04 X-units. Subsequent absolute measurements of x-ray wavelengths with ruled gratings indicated that the X-unit exceeds 10^{-13} m by about 2 parts per thousand. Furthermore, in practice, workers in the field began using definitions of the X-unit based on various x-ray wavelengths instead of the calcite grating definition, and subsequent precise measurements indicated that these wavelength standards differed from each other and from the calcite grating standard by as much as 20 parts per million.

The X-unit has been superseded by the A* unit, introduced by J. A. Bearden in 1965, which is based on the tungsten $K\alpha_1$ line as a standard. The peak of this line is defined as exactly 0.2090100 A*. X-ray wavelength tables have been published in terms of this unit. At the time the A* unit was defined, it was thought to equal 10^{-10} m (the angstrom unit, Å) to within 5 parts per million, but the A* unit is now believed to be 20 ± 5 parts per million larger than 10^{-10} m.

Units used in astronomy. Special units whose values are obtained experimentally are used in astronomy. For their definitions *see* ASTRONOMICAL UNIT; LIGHT-YEAR; PARSEC.

The astronomical unit and parsec are accepted for use with the SI. The parsec rather than the light-year is used in technical literature.

Area units. The square meter (m^2), the SI unit of area, is the area of a square with sides of length 1 m. Other area units are defined by forming squares of various length units in the same manner. The hectare (ha) is equal to 1 square hectometer (1 hm^2) or equivalently to 10^4 m^2. Its use with the SI is permitted for expressing land or water areas. *See* AREA.

Cross sections, which measure the probability of interaction between an atomic nucleus, atom, or molecule and an incident particle, have the dimensions of area, and the appropriate SI unit for expressing them is therefore the square meter. The barn (b), a unit of cross section equal to 10^{-28} m^2, has been accepted for temporary use with the SI. Typical nuclear reactions have cross sections ranging from millibarns to several thousand barns. A related quantity, which is connected with the probability that a reaction will emit radiations in a particular direction, is the differential cross section, which has the dimensions of barns per steradian. *See* NUCLEAR REACTION.

Units of volume. The cubic meter (m^3), the SI unit of volume, is the volume of a cube with sides of length 1 m. Other units of volume are defined by forming cubes of various length units in the same manner. The liter (symbol L in the United States) is equal to 1 cubic decimeter (1 dm^3), or equivalently to 10^{-3} m^3. It has been accepted for use with the SI for measuring volumes of liquids and gases.

Time units. The second (s) is the SI base unit of time. However, other units of time in customary use, such as the minute (1 min = 60 s), hour (1 h = 60 min), and day (1 d = 24 h), are acceptable for use with the SI. *See* TIME.

Frequency units. The hertz (Hz), the SI unit of frequency, is equal to 1 cycle per second. A periodic oscillation has a frequency of n hertz if it goes through n cycles in 1 s. Other units of frequency are defined by forming reciprocals of time units in the same manner. *See* FREQUENCY (WAVE MOTION).

Speed and velocity units. The meter per second (m/s), the SI unit of speed or velocity, is the magnitude of the constant velocity at which a body traverses 1 m in 1 s. Other speed and velocity units are defined by dividing a unit of length by a unit of time in the same manner. *See* SPEED; VELOCITY.

The knot (kn) is equal to 1 nautical mile per hour (1 nmi/h); it has been accepted for temporary use with the SI.

Acceleration units. The meter per second squared (m/s^2), the SI unit of acceleration, is the acceleration of a body whose velocity changes by 1 m/s in 1 s. Other units of acceleration are defined by dividing a unit of velocity by a unit of time in the same manner. *See* ACCELERATION.

The gal or galileo (symbol Gal) is equal to 1 cm/s^2, or equivalently to 10^{-2} m/s^2. This unit and its decimal submultiple the milligal (1 mGal = 10^{-3} Gal = 10^{-5} m/s^2) are employed in geodesy and geophysics to express the acceleration of gravity, and have been accepted for temporary use with the SI. *See* EARTH, GRAVITY FIELD OF; GEODESY.

Mass units. The kilogram (kg), the SI base unit of mass, is the only SI unit whose name, for historical reasons, contains a prefix. Names of decimal multiples and submultiples of the kilogram are formed by attaching prefixes to the word gram (g). Since 1 kg = 10^3 g, 1 g = 10^{-3} kg. The metric ton (t), which is equal to 10^3 kg or 1 megagram (Mg), is permitted in commercial usage of the SI.

The pound (lb), is, as discussed above, the unit of mass in the British absolute system, and is also in customary use in the United States. In 1959 the pound was defined to be exactly 0.45359237 kg.

The slug is, as discussed above, the unit of mass in the British gravitational system. By definition 1 pound force (lbf) acting on a body of mass 1 slug produces an acceleration of 1 foot per second

squared (1 ft/s^2). The slug is equal to approximately 32.174 lb or 14.594 kg. *See* MASS.

Force units. The newton (N), the SI unit of force, is the force which imparts an acceleration of 1 meter per second squared (1 m/s^2) to a body having a mass of 1 kg.

The dyne, the cgs absolute unit of force, is the force which imparts an acceleration of 1 centimeter per second squared (1 cm/s^2) to a body having a mass of 1 g. Since 1 cm/s^2 = 10^{-2} m/s^2, and 1 g = 10^{-3} kg, it follows that 1 dyne = 10^{-5} N.

The unit of force in the British absolute system is, as discussed above, the poundal (pdl), the force which imparts an acceleration of 1 foot per second squared (1 ft/s^2) when applied to a body of mass 1 lb. One poundal is approximately 0.13825 N.

As discussed above, the units of force in the mks gravitational, cgs gravitational, and British gravitational systems are the forces which impart an acceleration equal to the standard acceleration of gravity, g_n = 9.80665 m/s^2 $\cong$ 32.174 ft/s^2, when applied to bodies having masses of 1 kg, 1 g, and 1 lb, respectively. These units are named the kilogram force (kgf), gram force (gf), and pound force (lbf), respectively. Unfortunately, these units have also been called simply the kilogram, gram, and pound, giving rise to confusion with the mass units of the same name; in proper usage the suffix "force" should always be employed. The pound has also been called the pound mass (lbm) to further distinguish mass and force units. One pound force is approximately 4.4482 N. *See* FORCE.

Pressure and stress units. The pascal (Pa), the SI unit of pressure and stress, is the pressure or stress of 1 newton per square meter (N/m^2). This is a rather small unit for most practical purposes; for example, atmospheric pressure is approximately 10^5 Pa. Thus, most pressures are most readily expressed in decimal multiples of the pascal formed by attaching the appropriate SI prefix.

The dyne per square centimeter (dyne/cm^2), the cgs absolute unit of pressure, has sometimes been called the barye, but this name is uncommon. Other units of pressure can also be formed by dividing various units of force by various units of area, such as the pound force per square inch (lbf/in.^{2}, frequently abbreviated psi).

Pressure has been frequently expressed in terms of the bar and its decimal submultiples, where 1 bar = 10^6 dynes/cm^2 = 10^5 Pa. A reference level of 1 microbar (1 μbar = 10^{-6} bar = 1 dyne/cm^2 = 0.1 Pa) is commonly used in the calibration of microphones, hydrophones, and loudspeakers. The millibar (1 mbar = 10^{-3} bar = 10^2 dynes/cm^2 = 10^2 Pa = 0.1 kPa) is commonly used in meteorology. The temporary use of the millibar with the SI has been allowed in order to permit meteorologists to communicate easily within their profession, but the kilopascal should be used in presenting meteorological data to the public. *See* SOUND PRESSURE.

Pressures are also frequently expressed in terms of the height of a column of either mercury or water which the pressure will support. This practice is very convenient in conjunction with the use of barometers or other instruments in which pressure is determined from such column heights. However, the pressure which supports a liquid column of given height depends on the acceleration due to gravity and on the density of the liquid (and, in turn, on its temperature), and these must therefore be specified for accurate work.

Two other units which have been frequently used for measuring pressure are the standard atmosphere and the torr. The standard atmosphere (atm) is exactly 101,325 Pa, which is approximately the average value of atmospheric pressure at sea level. The torr is exactly 1/760 atmosphere, or approximately 133.322 Pa. To within 1 part per million, it is equal to the pressure of a column of mercury of height 1 millimeter (1 mmHg) at a temperature of 0°C when the acceleration due to gravity has the standard value g_n = 9.80665 m/s^2. *See* ATMOSPHERE; PRESSURE; PRESSURE MEASUREMENT.

Energy and work units. The joule (J), the SI unit of energy or work, is the work done by a force of magnitude 1 newton when the point at which the force is applied is displaced 1 m in the direction of the force. Thus, joule is a short name for newton-meter (N · m) of energy or work. *See* ENERGY; WORK.

Units of energy or work in other systems are defined by forming the product of a unit of force and a unit of length in precisely the same manner as in the definition of the joule. Thus, the erg, the cgs absolute unit of energy or work, is the product of 1 dyne and 1 cm. Erg is a short name for dyne-centimeter of energy or work. Since 1 dyne = 10^{-5} N and 1 cm = 10^{-2} m, it follows that 1 erg = 10^{-7} J.

The foot-poundal (ft · pdl), the British absolute unit of energy or work, is the product of 1 poundal and 1 foot. The foot-pound, or, more properly, the foot-pound force (ft · lbf), the British gravitational unit of energy or work, is the product of 1 lbf and 1 ft. Foot-poundal and foot-pound are also names of units of torque, discussed below.

Power-time products. Sometimes energy is measured in units which are products of a unit of power and a unit of time. Since 1 watt (W) of power equals 1 joule per second (1 J/s), as discussed below, the joule is equivalent to 1 watt-second (1 W · s). In electrical power applicatons, energy is frequently measured in kilowatt-hours (kWh), where 1 kWh = (10^3 W) · (3600 s) = 3.6 × 10^6 J. *See* ELECTRICAL ENERGY MEASUREMENT.

Calorie and British thermal unit. The calorie and the British thermal unit were originally defined as the quantities of heat required to raise the temperature of a specified mass of water by a specified temperature. Usually these units have been used in connection with energy as heat, but they can also be used when referring to energy as work or energy in any other form. *See* CALORIMETRY; HEAT.

The calorie was originally defined as the quantity of heat required to raise the temperature of 1 g of air-free water 1°C under a constant pressure of 1 atm. However, the magnitude of the calorie, so defined, depends on the place on the Celsius temperature

scale at which the measurement is made. The 15°C calorie is based on the temperature interval from 14.5 to 15.5°C, but other temperature intervals have been used. At the Fifth International Conference on the Properties of Steam in London in 1956, the International (Steam) Table calorie was defined as exactly 4.1868 J; this is the type of calorie most frequently used in mechanical engineering. The thermochemical calorie, which has been used in thermochemistry in preference to the other types of calorie, is exactly 4.184 J.

Since the calorie is a relatively small unit, the kilocalorie (kcal), also called the large calorie and the kilogram-calorie, is used to designate 10^3 calories. The energy intake of the human body has been commonly expressed as the number of large calories which the eaten food will liberate as it passes through the body. Unfortunately, confusion arises because it is usually not indicated that kilocalories rather than calories are being used.

The British thermal unit (Btu) was originally defined as the quantity of heat required to raise the temperature of 1 lb of air-free water 1°F under a constant pressure of 1 atm. Again, the temperature interval must be specified; the 60° Btu, based on the interval from 59.5 to 60.5°F, has been frequently used.

The International Table Btu and the thermochemical Btu are both defined to be of such magnitude that the values of specific heat capacity of any substance are equal in size whether expressed in Btu per pound per degree Fahrenheit [Btu/(lb · °F)] or in calories per gram per degree Celsius [cal/(g · °C)], when the corresponding type of calorie is used. From this it follows that each type of Btu is equal to approximately 251.996 times the corresponding type of calorie. Then the International Table Btu is approximately 1055.056 J, and the thermochemical Btu is approximately 1054.350 J.

Electronvolt. The electronvolt (eV), whose value is experimentally determined, is frequently used to express the energies of atomic systems. It has been accepted for use with the SI. *See* ELECTRONVOLT.

Power units. The watt (W), the SI unit of power, is the power which gives rise to the production of energy at the rate of 1 joule per second (1 J/s). Other units of power can be defined by forming the ratio of a unit of energy to a unit of time in the same manner. *See* POWER.

The horsepower (hp) is equal to exactly 550 ft · lb/s, or approximately 745.700 W. It frequently has been employed to express the power generated by engines and machinery.

Torque units. The newton-meter (N · m), the SI unit is the magnitude of the torque produced by a force of 1 newton acting at a perpendicular distance of 1 m from a specified axis of rotation. The joule should never be used as a synonym for this unit: although the two units are both products of 1 newton and 1 m, the orientation of force and length is quite different in the two cases.

Units of torque in other systems are defined by forming the product of a unit of force and a unit of length in precisely the same manner as in the definition of the newton-meter. Thus the dyne-centimeter (dyne · cm), the cgs absolute unit of torque, is the product of 1 dyne and 1 cm; the erg should never be used as a synonym for this unit.

The foot-poundal (ft · pdl), the British absolute unit of torque, is the product of 1 poundal and 1 foot. The foot-pound (ft · lbf), the British gravitational unit of torque, is the product of 1 lbf and 1 ft. These units are sometimes called the poundal-foot (pdl · ft) and pound-foot (lbf · ft) to distinguish them from the units of energy or work.

The SI unit of torque does work of 1 N · m on a body which rotates through 1 rad in the direction of the torque. Some standards of metric practice therefore suggest that the SI unit of torque can be designated newton-meter per radian (N · m/rad) in order to further distinguish it from the unit of energy or work. The names of other torque units should then be similarly modified. *See* ROTATIONAL MOTION; TORQUE.

Viscosity units. For units of dynamic viscosity and kinematic viscosity *see* VISCOSITY.

Permeability units. The darcy is a commonly used unit of permeability to fluid flow. The permeability of a rock, brick, or other porous substance is 1 darcy if 1 cm^3 of a fluid of 1-centipoise viscosity will flow through a section 1 cm thick and of 1−cm^2 cross section in 1 s at a pressure difference of 1 atm. In core analysis and other measurements of petroleum-bearing rock, the unit used is the millidarcy, which is 10^{-3} darcy. *See* PETROLEUM RESERVOIR ENGINEERING; SOIL MECHANICS.

The SI unit of permeability is the permeability of a porous substance such that 1 m^3 of a fluid of viscosity 1 Pa · s will flow through a section 1 m thick and of 1-m^2 cross section in 1 s at a pressure difference of 1 Pa. This unit has no special name, but has the dimensions of 1 m^2. The SI unit of permeability is exactly 1.01325×10^{12} darcy, and 1 darcy $\equiv 0.987 \times 10^{-12}$ m^2.

Electrical Units

For a general discussion of electrical units, including the SI or mks system, three cgs systems [electrostatic system of units (esu), electromagnetic system of units (emu), and gaussian system], and definitions of the SI units ampere (A), volt (V), ohm (Ω), coulomb (C), farad (F), henry (H), and weber (Wb) *see* ELECTRICAL UNITS AND STANDARDS

This section discusses some additional SI units and some units in the cgs electromagnetic system which are frequently encountered in scientific literature in spite of the fact that their use has been discouraged.

Siemens. The siemens (S), the SI unit of electrical conductance, is the electrical conductance of a conductor in which a current of 1 ampere is produced by an electric potential difference of 1 volt. The conductance G is defined by the equation $I = GV$, where I is the current in amperes, V is the potential difference in volts, and G the conductance in siemens. The conductance of an electrical conductor in siemens is the reciprocal of its resistance in ohms. *See* CONDUCTANCE; ELECTRICAL RESISTANCE.

The siemens was formerly called the mho (℧) to illustrate the fact that the unit is the reciprocal of the ohm.

Abampere. The abampere (abA), the cgs electromagnetic unit of current, is that current which, if maintained in two straight, parallel conductors of infinite length, of negligible circular cross section, and placed 1 cm apart in vacuum, would produce between these conductors a force equal to 2 dynes per centimeter of length. *See* ELECTRIC CURRENT.

The abampere is equal to exactly 10 A. That a current of 10 A satisfies the above definition can be seen from the following argument: If two straight, parallel conductors of infinite length, separated by distance r, carry currents I_1 and I_2, then the force F on segment of one of the conductors of length l is given by Eq. (5). The constant μ_0 is the permeability of vacuum, $4\pi \times 10^{-7}$ newton per ampere squared. Thus, when the current in each wire is 1 A and the distance between the conductors is $r = 1$ m, the force between them is 2×10^{-7} newton per meter of length, as in the definition of the ampere. When $I_1 = I_2 = 10$ A, and $l = r = 1$ cm $= 10^{-2}$ m, the force between the conductors, given by Eq. (6), is the force specified in the definition of the abampere. *See* AMPÈRE'S LAW; BIOT-SAVART LAW.

$$F = \frac{\mu_0 I_1 I_2 l}{2\pi r} \tag{5}$$

$$F = \frac{(4\pi \times 10^{-7})\,(10)\,(10)\,(10^{-2})}{2\pi(10^{-2})}\ \text{newton}$$
$$= 2 \times 10^{-5}\text{newton} = 2\ \text{dynes} \tag{6}$$

Abvolt. The abvolt (abV), the cgs electromagnetic unit of electrical potential difference and electromotive force, is the difference of electrical potential between two points of a conductor carrying a constant current of 1 abA, when the power dissipated between these points is equal to 1 erg per second. Then 1 abV $= 10^{-8}$ V, as can be seen from Eq. (7).

$$1\ \text{abV} = \frac{1\ \text{erg/s}}{1\ \text{abA}} = \frac{10^{-7}\ \text{J/s}}{10\ \text{A}}$$
$$= 10^{-8}\ \frac{\text{W}}{\text{A}} = 10^{-8}\ \text{V} \tag{7}$$

The abampere and abvolt are not frequently used, but they figure in the definitions of the maxwell, gauss, oersted, and gilbert, given below.

Maxwell. The maxwell (Mx), the cgs electromagnetic unit of magnetic flux, is the magnetic flux which, linking a circuit of one turn, produces in it an electromotive force of 1 abV as it is reduced to zero in 1 s. Then 1 maxwell $= 10^{-8}$ weber, as can be seen from Eq. (8).

$$1\ \text{Mx} = 1\ \text{abV}\cdot\text{s} = 10^{-8}\ \text{V}\cdot\text{s} = 10^{-8}\ \text{Wb} \tag{8}$$

Units of magnetic flux density. The tesla (T), the SI unit of magnetic flux density (also called magnetic induction), is a magnetic flux density of 1 weber per square meter (1 Wb/m^2). The gauss (Gs), the cgs electromagnetic unit of magnetic flux density, is a magnetic flux density of 1 maxwell per square centimeter (1 Mx/cm^2). Then 1 gauss $= 10^{-4}$ tesla, as can be seen from Eq. (9).

$$1\text{Gs} = \frac{1\ \text{Mx}}{(1\ \text{cm})^2} = \frac{10^{-8}\ \text{Wb}}{(10^{-2}\ \text{m})^2}$$
$$= 10^{-4}\ \frac{\text{Wb}}{\text{m}^2} = 10^{-4}\ \text{T} \tag{9}$$

Alternative equivalent definitions for the tesla and gauss follow from Eq. (10) for the magnetic induction B, where F is the force exerted on a current element of length l and current l making angle θ with the magnetic induction vector. This equation is valid whether quantities are expressed in SI or in cgs electromagnetic units. Thus, an alternative definition for the tesla is the constant magnetic induction which exerts a force of 1 newton on a straight wire of length 1 m perpendicular to the magnetic induction vector and carrying a current of 1 ampere; that is, 1 T = 1 N/(A · m). The gauss can be defined, in precisely the same manner, as 1 dyne per abampere-centimeter [1 Gs = 1 dyne/(abA · cm)].

$$B = \frac{F}{Il\sin\theta} \tag{10}$$

Units of magnetic field strength. The SI unit of magnetic field strength is 1 ampere per meter (1 A/m), which is the magnetic field strength at a distance of 1 m from a straight conductor of infinite length and negligible circular cross section which carries a current of 2π A. (Other geometric configurations can be used to define this unit; the choice is arbitrary.) This definition is based on the definition, in the SI, of the magnetic field strength H_{SI}. At a distance r from a long straight conductor carrying current I, H_{SI} is given by Eq. (11). The left-hand side of this equation is the line integral of H_{SI} around a circular path, all of whose points are at distance r from the conductor. Substituting $r = 1$ m and $I = 2\pi$ A gives $H_{\text{SI}} = 1$ A/m.

$$2\pi r H_{\text{SI}} = I \tag{11}$$

The SI unit has sometimes been called the ampere-turn per meter because Eq. (11) can be modified to Eq. (12), where i is the current in each turn of a coil, part of which forms the long straight conductor in question, and n is the number of turns in the coil. Substituting $r = 1$ m, $i = 2\pi$ A, and $n = 1$ turn in Eq. (12) gives $H_{\text{SI}} = 1$ A · turn/m. However, the name ampere per meter has been adopted by the CGPM.

$$2\pi r H_{\text{SI}} = ni \tag{12}$$

The oersted (Oe), the cgs electromagnetic unit of magnetic field strength, is the magnetic field strength at a distance of 1 cm from a straight conductor of infinite length and negligible circular cross section which carries a current of 0.5 abA. Again, this definition is based on the definition of magnetic field strength. In the cgs electromagnetic system the magnetic field strength H_{cmu} at a distance r from a long straight conductor carrying current I is given by

Eq. (13), which differs from Eq. (11) by the factor

$$2\pi r H_{\text{emu}} = 4\pi I \tag{13}$$

4π on the right-hand side. Substituting $r = 1$ cm and $I = 0.5$ abA in this equation gives $H_{\text{emu}} = 1$ abA/cm. Thus, 1 Oe = 1 abA/cm in the cgs electromagnetic system, but the method by which the quotient is formed is different from that in the SI unit, being based on Eq. (13) rather than Eq. (11).

When the magnetic field strength in cgs electromagnetic units is 1 Oe, the magnetic field strength in SI units is found by substituting in Eq. (11) $I = 0.5$ abA and $r = 1$ cm, which gives Eq. (14). Thus

$$H_{\text{SI}} = \frac{0.5 \text{ abA}}{2\pi(1 \text{ cm})} = \frac{0.5(10 \text{ A})}{2\pi(10^{-2} \text{ m})} = \frac{10^3}{4\pi}\frac{\text{A}}{\text{m}} \tag{14}$$

1 Oe corresponds to $(10^3/4\pi)$ A/m $\cong$ 79.577 A/m.

Units of magnetic potential and mmf. The ampere serves as the SI unit of magnetic potential difference and magnetomotive force (mmf), as well as the unit of current. In the SI system, the magnetomotive force around a closed path equals the current passing through a surface enclosed by the path. Thus, 1 A is the magnetomotive force around a closed path when a current of 1 A passes through an enclosed surface. The SI unit has also been called the ampere-turn, because the magnetomotive force around a path that loops around a current coil equals the product of the number of turns in the coil and the current in each turn; however, the name ampere has been adopted by the CGPM.

The gilbert (Gb), the cgs electromagnetic unit of magnetic potential difference and magnetomotive force, is the magnetomotive force around a closed path enclosing a surface through which flows a current of $[1/(4\pi)]$ abA. This definition is based on the fact that in the cgs electromagnetic system the magnetomotive force around a closed path is 4π times the current passing through an enclosed surface. When this current is $[1/(4\pi)]$ abA, the magnetomotive force is therefore 1 abA in cgs electromagnetic units, but the unit is given the special name gilbert. When the magnetomotive force in cgs electromagnetic units is 1 gilbert, the magnetomotive force in SI units is $[1/(4\pi)]$ abA = $[10/(4\pi)]$ A. Thus, 1 Gb corresponds to $[10/(4\pi)]$ A $\cong$ 0.79577 A. *See* MAGNETOMOTIVE FORCE.

Photometric Units

Photometric units involve a new base quantity, luminous intensity. For the definition of the candela (cd), the SI unit of luminous intensity, *see* PHOTOMETRY; PHYSICAL MEASUREMENT.

For a general discussion of photometric units, including units of illuminance (illumination) and luminance, and in particular the SI units lux (lx) and candela per square meter (cd/m^2), *see* ILLUMINATION; LUMINANCE.

This section gives an explicit definition of the lumen and discusses units of luminous energy.

Lumen. The lumen (lm), the SI unit of luminous flux, is the luminous flux emitted within a unit solid angle (1 steradian) by a point source having a uniform intensity of 1 candela. It follows, therefore, that a light source having an intensity of 1 candela in every direction will be emitting a total luminous flux of 4π lumens.

The lumen is also equal to the luminous flux received on a unit surface, all points of which are at a unit distance from a point source having a uniform intensity of 1 candela.

The output of light sources is given in lumens. *See* LUMINOUS FLUX.

Luminous energy units. The lumen-second (lm · s), the SI unit of luminous energy (also called quantity of light), is the luminous energy radiated or received over a period of 1 s by a luminous flux of 1 lumen. This unit is also called the talbot. Other units of luminous energy include the lumen-hour (1 lm · h = 3600 lm · s) and the million-lumen-hour. *See* LUMINOUS ENERGY.

Radiation Units

Certain quantities and units are used particularly in the area of ionizing radiation. The special units curie, roentgen, rad, and rem, which were previously adopted for use in this area, are not coherent with the SI, but their temporary use with the SI has been approved while the transition to SI units takes place. The CGPM, acting on proposals of the International Commission on Radiation Units and Measurements (ICRU) and the International Commission on Radiological Protection (ICRP), has adopted the special names becquerel, gray, and sievert for the SI derived units of activity, absorbed dose, and dose equivalent.

Activity units. The becquerel (Bq), the SI unit of activity (radioactive disintegration rate), is the activity of a radionuclide decaying at the rate of one spontaneous nuclear transition per second. Thus 1 Bq = 1 s^{-1}.

The curie (Ci), the special unit of activity, is equal to 3.7×10^{10} Bq. (This unit was originally chosen to approximate the activity of 1 g of radium-226.) *See* RADIOACTIVITY.

Exposure units. The SI unit of exposure to ionizing radiation, 1 coulomb per kilogram (1 C/kg), is the amount of electromagnetic radiation (x-radiation or gamma radiation) which in 1 kg of pure dry air produces ion pairs carrying 1 coulomb of charge of either sign. (The ionization arising from the absorption of bremsstrahlung emitted by electrons is not to be included in measuring the charge). *See* BREMSSTRAHLUNG.

The roentgen (R), the special unit of exposure, is equal to 2.58×10^{-4} C/kg. (This is equivalent to 1 cgs electrostatic unit of charge in 1.293 mg of air, which is the mass of 1 cm^3 of air at temperature 0°C and pressure 1 atm.)

Absorbed dose and kerma units. The gray (Gy), the SI unit of absorbed dose, is the absorbed dose when the energy per unit mass imparted to matter by ionizing radiation is 1 joule per kilogram (1 J/kg).

The rad (rd), the special unit of absorbed dose, is equal to 10^{-2} Gy. In air, the number of rad per

roentgen is 0.877, and this is also approximately true in soft tissue.

Kerma (an acronym for kinetic energy released in matter) is related to absorbed dose and is measured in the same units. The kerma is the sum of the initial kinetic energies of all the charged ionizing particles liberated by uncharged ionizing particles in an element of matter divided by the mass of that element. Except for a correction due to energy lost to bremsstrahlung, the exposure is the ionization equivalent of the air kerma.

Dose equivalent units. Different types of radiation cause slightly different effects in biological tissue. For this reason, a weighted absorbed dose called the dose equivalent is used in comparing the effects of radiation on living systems. The dose equivalent is the product of the absorbed dose and various dimensionless modifying factors stipulated by the ICRP. The chief such factor, the quality factor, depends on the linear energy transfer, and this, in turn, depends on the kind of incident radiation and its energy. For electromagnetic radiation and electrons the quality factor equals 1; for heavier charged particles it is greater than 1. *See* LINEAR ENERGY TRANSFER (BIOLOGY).

The sievert (Sv), the SI unit of dose equivalent, is the dose equivalent when the absorbed dose of ionizing radiation multiplied by the stipulated dimensionless factors is 1 joule per kilogram (1 J/kg).

The rem, the special unit of dose equivalent, is equal to 10^{-2} Sv. *See* ENVIRONMENTAL RADIOACTIVITY; RADIATION BIOLOGY; RADIATION INJURY (BIOLOGY).

Unit of Catalytic Activity

In 1999, the CGPM adopted the katal (kat) as the SI unit of catalytic activity, equal to 1 mole per second (1 mol/s). In its resolution adopting the unit, the CGPM cited the importance for human health and safety of facilitating the use of SI units in the fields of medicine and biochemistry, and noted that a non-SI unit of catalytic activity called "unit," symbol U, equal to 1 micromole per minute (1μmol/min), which is not coherent with SI, has been in widespread use in these fields.

Other Units

Logarithmic measures may be used with the SI. *See* BEL; DECIBEL; NEPER; PH; VOLUME UNIT (VU). For units of loudness (sone, phon) *see* LOUDNESS. For units of sound absorption by surfaces *see* ARCHITECTURAL ACOUSTICS. For units of acoustic frequency and pitch *see* MUSICAL ACOUSTICS; PITCH. For quantities and units pertaining to sound transmission *see* SOUND. For temperature scales *see* TEMPERATURE. For units measuring quantities in chemistry *see* ATOMIC MASS UNIT; ATOMIC WEIGHT; CONCENTRATION SCALES; ELECTROCHEMICAL EQUIVALENT; GRAM-MOLECULAR WEIGHT; MOLECULAR WEIGHT; RELATIVE ATOMIC MASS; RELATIVE MOLECULAR MASS. For quantities and units pertaining to nuclear reactors *see* REACTOR PHYSICS. For units of information content *see* BIT; INFORMATION THEORY.

Jonathan F. Weil

Bibliography. Institute of Electrical and Electronics Engineers, *Standard for Use of the International System of Units (SI): The Modern Metric System*, IEEE/ASTM Std. SI 10-2002, 2002; International Commission on Radiation Units and Measurements, *Fundamental Quantities and Units for Ionizing Radiation*, ICRU Rep. 60, 1998; International Commission on Radiation Units and Measurements, *Quantities and Units in Radiation Protection Dosimetry*, ICRU Rep. 51, 1993; International Organization for Standardization, *Quantities and Units*, 1993; B. N. Taylor, *Guide for the Use of the International System of Units (SI)*, NIST Spec. Pub. 811, 1995; B. N. Taylor (ed.), *The International System of Units (SI)*, NIST Spec. Pub. 330, revised 2001; U. S. Metric Association, *Guide to the Use of the Metric System*, 15th ed., 2002.

Universal joint

A linkage that transmits rotation between two shafts whose axes are coplanar but not coinciding. The universal joint is used in almost every class of machinery: machine tools, instruments, control devices, and, most familiarly, automobiles.

Hooke's joint. A simple universal joint, known in English-speaking countries as Hooke's joint and in continental Europe as a Cardan joint, is shown in **Fig. 1***a*. It consists of two yokes *A* and *B* (Fig. 1*b*) attached to their respective shafts and connected by means of spider *S*. Angle ϕ between the shafts may have any value up to approximately 35°, if angular velocity is moderate when angle ϕ is large. Although shaft *B* must make one revolution for each revolution of shaft *A*, the instantaneous angular displacement of shaft *B* is the same as that of shaft *A* only at the end of each 90° of shaft rotation. Thus, only at four positions during each revolution is angular velocity ω_B of shaft *B* the same as angular velocity ω_A of shaft *A*. Three curves for designated values of ϕ are plotted in **Fig. 2**. These curves show deviation of ω_B from a constant ω_A as shaft *A* is turned through 180°. *See* FOUR-BAR LINKAGE.

Double Hooke's joint. The variation in angular displacement and angular velocity between driving and driven shafts, which is objectionable in many mechanisms, can be eliminated by using two Hooke's joints, with an intermediate shaft (Fig. 1*c*). This arrangement is conventional for an automobile drive shaft. The axes of the driving and driven shafts need not intersect; however, it is necessary that the axes y and y' of the two yokes attached to intermediate shaft *b* lie respectively in planes containing the axes of adjoining shafts (*a*, y, and *b* in same plane; *b*, y', and *c* in same plane), and that angle β between the driving and intermediate shafts equal angle β' between the intermediate and driven shafts.

Bendix-Weiss joint. Two intersecting, thin, bent shafts (**Fig. 3***a*) with plane *Z* through the point of contact bisecting the angle between the shafts and perpendicular to the plane containing the axes of the shafts will maintain the constant angular velocity.

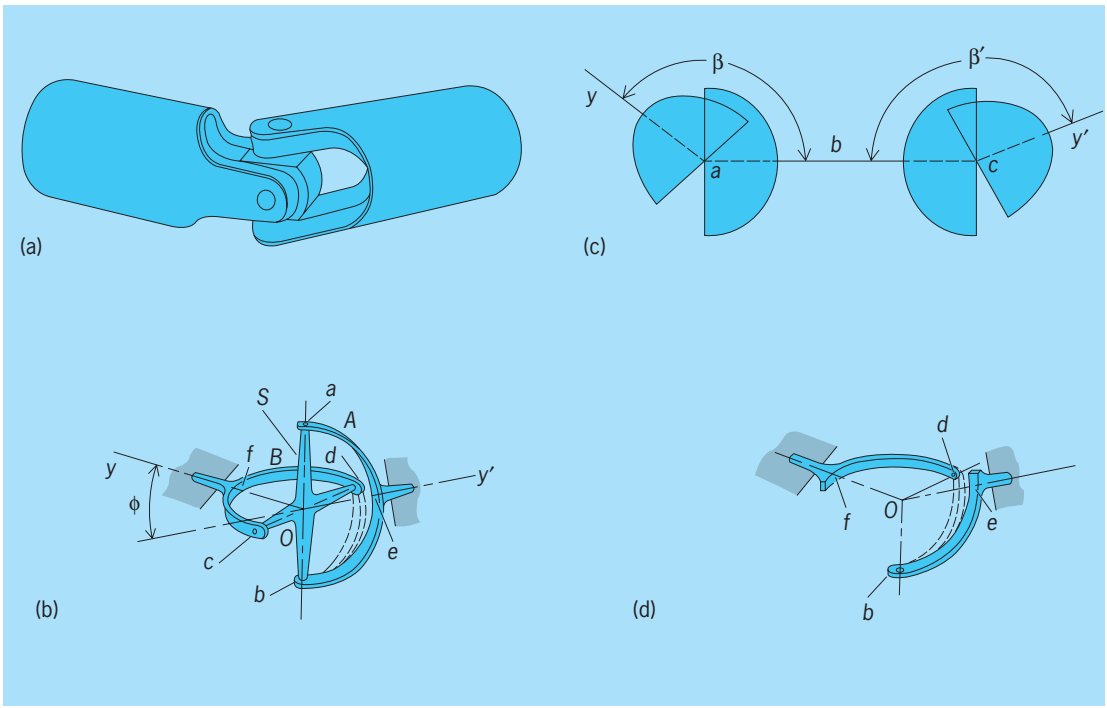

Fig. 1. Universal joints. (*a*) Simple. (*b*) Yoke and spider. (*c*) Double. (*d*) Four-bar conic linkage equivalent of yoke and spider. (*After C. W. Ham, E. J. Crane, and W. L. Rogers, Mechanics of Machinery, McGraw-Hill, 1958*)

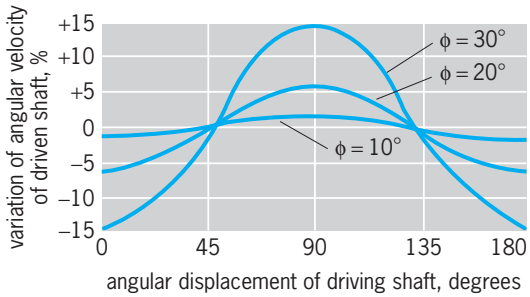

Fig. 2. Variation in the angular velocity of the driven shaft of a simple universal joint for the constant angular velocity of the driving shaft.

One practical application of this principle, in which constant angular velocity is transmitted through a single universal joint, is the joint developed by Carl Weiss and illustrated in Fig. 3*b*. Four large balls are transmitting elements, while a center ball acts as a spacer. The transmitting balls must lie in plane *Z*, as explained above. By means of milled grooves in the yoke attached to each shaft, the balls are maintained at all times in plane *Z*. This joint (**Fig. 4**) is used for a front-wheel automotive drive because it transmits unvarying angular displacement even when steering the vehicle requires a varying angle between driving and driven shafts.

Spherical four-bar linkage. The universal joint is a spatial, spherical four-bar linkage. If only half of each fork is considered as *eb* of *A* and *fd* of *B* (Fig. 1*b*) and these are assumed to be connected by spherical link *db* equal to the fixed distance between the two adjacent points of spider *S*, a four-bar conic linkage is produced (Fig. 1*d*) in which the axes of all the turning pairs intersect in *O*. With this arrangement the fork could be omitted, and there would be a kinematic equivalent of the original mechanism.

Driver and follower make complete revolutions in the same time, but the velocity ratio is not constant

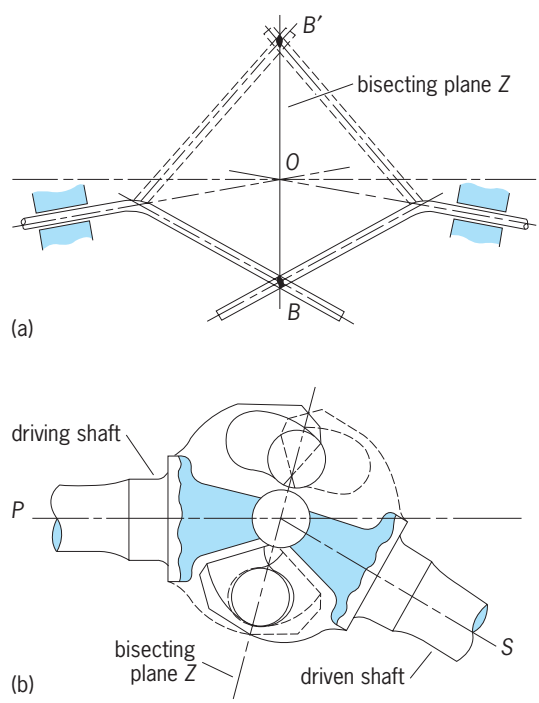

Fig. 3. Interacting shafts in sliding contact transmit constant angular velocity. (*a*) Basic configuration (*after R. T. Hinkle, Kinematics of Machines, Prentice-Hall, 1952*). (*b*) Cross section of ball-bearing adaptation.

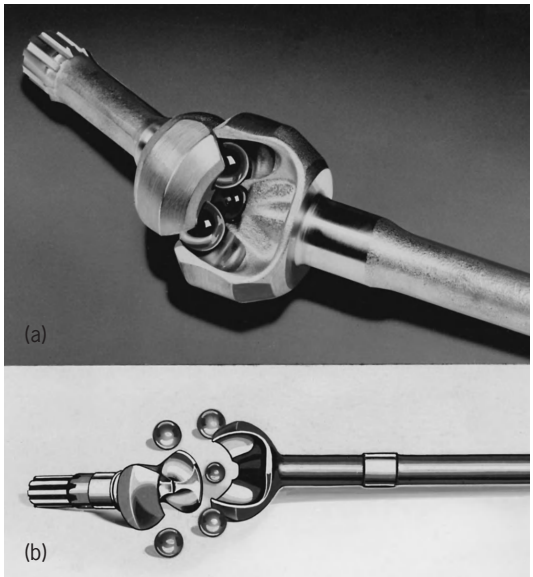

Fig. 4. Constant-velocity universal joint sometimes called a Rzeppa joint. (*a*) Partially separated cutaway. (*b*) Disassembled joint showing the arrangement; the right-hand member has been cut away. (*Bendix Aviation Corp.*)

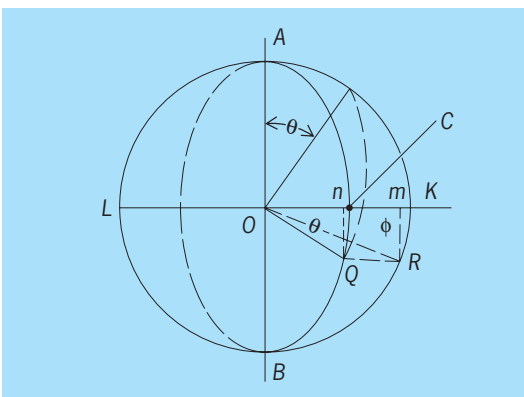

Fig. 5. Geometry of motion for spherical joint of Fig. 1*d*.

throughout the revolution as can be shown by analysis. In **Fig. 5**, COQ is the projection of the real angle described by the follower. The real component of the motion Qn of the follower is in a direction parallel to AB, and line AB is the intersection of the planes of the driver's and the follower's path. The true angle ϕ described by the follower (while the driver describes the angle θ) can be found by revolving OQ about AB as an axis into the plane of circle $AKBL$. Then OR is the true length of OQ, and ROK, designated ϕ, equals the true angle that is projected as COQ, designated θ. With this notation $\tan \phi = Rm/Om$ and $\tan \phi = On/On$. But $Qn = Rm$; hence $\tan \theta/\tan \phi = Om/On = OK/OC = 1/\cos \beta$. Douglas P. Adams

Bibliography. A. S. Hall, Jr., *Kinematics and Linkage Design*, 1986; H. H. Mabie and C. F. Reinholtz, *Mechanisms and Dynamics of Machinery*, 4th ed., 1987; G. H. Martin, *Kinematics and Dynamics of Machines*, 2d ed., 1982.

Universal motor

A series motor built to operate on either alternating current (ac) or direct current (dc). It is normally designed for capacities less than 1 hp (0.75 kW). It is usually operated at high speed, 3500 revolutions/min loaded and 8000 to 10,000 revolutions/min unloaded. For lower speeds, reduction gears are often employed, as in the case of electric hand drills or food mixers. As in all series motors, the rotor speed increases as the load decreases, and the no-load speed is limited only by friction and windage. To obtain more constant speed with variations in load, a centrifugal governor may be used to switch in or out a small resistor in series with the armature as in **Fig. 1**.

If an alternating current is applied to any dc series motor, the motor would still rotate. Since the current is reversed simultaneously in the armature and the field, the torque would pulsate but would not reverse direction. However, a universal motor designed to operate on ac should have certain modifications: laminated cores to avoid excessive eddy currents, fewer turns in the field coils than in a dc motor, and more poles and usually more commutator segments. *See* CORE LOSS; DIRECT-CURRENT MOTOR.

The series ac motor is an alternating-current commutator motor which has great flexibility of performance. It can be operated over a wide range of speeds and is readily controllable. The series ac commutator motor is in many respects similar to the dc series motor and the universal motor.

The ac series motor, like the dc series motor and the repulsion motor, consists fundamentally of these windings or their equivalent: (1) rotating armature winding, (2) stationary field winding, and (3) compensating winding (**Fig. 2**).

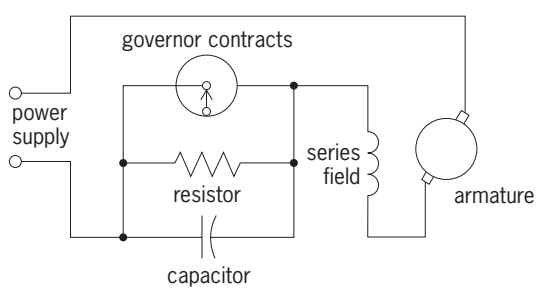

Fig. 1. Universal motor diagram.

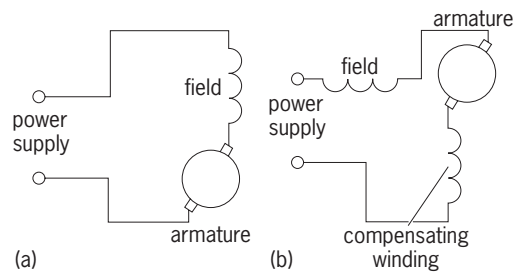

Fig. 2. Series motor diagram. (*a*) Without compensating winding. (*b*) With compensating winding.

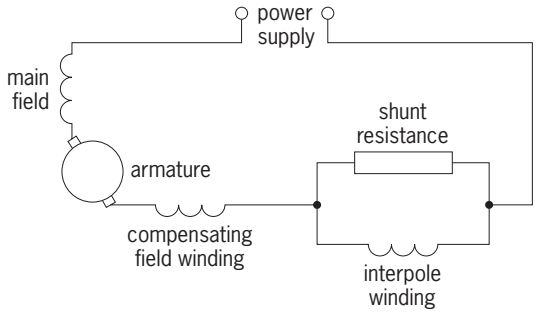

Fig. 3. Single-phase traction motor circuit diagram. (*General Electric Co.*)

A major problem in larger (up to 1000 hp or 750 kW) ac series motors is in commutation. Because of the transformer action between the field and armature coils, voltage is produced in the armature coils which are short-circuited by the brushes as the commutator bars pass under them. The coils which are short-circuited act like a short-circuited secondary of a static transformer. The resulting large currents are interrupted as the bars pass the brushes, causing bad sparking. In addition, these induced currents reduce the magnetic flux of the field and reduce the torque of the motor. Interpoles shunted with noninductive resistance are required on ac series motors as in **Fig. 3**. *See* COMMUTATION.

The single-phase commutator motor usually has a large number of turns in the armature winding, more commutator segments, and a small number of turns in the field winding, as compared with the dc motor, which is designed for relatively strong field and weaker armature. The ac motor usually has more poles and operates at a lower voltage than its dc counterpart. Irving L. Kosow

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; A. E. Fitzgerald, C. Kingsley, and S. D. Umans, *Electric Machinery*, 5th ed., 1990; G. McPherson and R. D. Laramore, *An Introduction to Electrical Machines and Transformers*, 2d ed., 1990; S. A. Nasar, *Electric Machines and Power Systems*, 1995.

Universe

The universe comprises everything in existence, including all matter and energy, and the enormous volume which contains them. The observable universe currently spans about 2.6×10^{23} km (1.6×10^{23} mi), and contains approximately 2.4×10^{52} kg (5.2×10^{52} lb) of matter, yielding an average density equivalent to a few atoms per cubic meter. Most of the universe, then, is empty space; the matter is distributed thinly throughout, forming objects and structures at a variety of different sizes. The study of this matter and energy, and its distribution, composition, and origin, is what constitutes the sciences of astronomy and cosmology.

Constituents

This article will start the cosmic survey with the more familiar physical objects—atoms, stars, planets, and galaxies—following a sequence of increasing size, building up to the largest structures in the universe. A number of lesser-known and less tangible entities, such as energy, exotic particles, and dark matter, will complete the survey. With this synopsis of the constituents of the universe in hand, this article will address the questions of where they all came from and what their ultimate fate will be.

Ordinary units of measure, such as miles, kilometers, pounds, and kilograms, are usually not convenient for dealing with the enormous sizes and masses encountered in astronomy, so astronomers have devised other units. The standard length scale in astronomy is the parsec (pc), equal to 3.1×10^{13} km (1.9×10^{13} mi). For larger distances, kiloparsecs (kpc = 1000 pc) and megaparsecs (Mpc = 1,000,000 pc) are used. The light-year, about one-third of a parsec, often appears in popular astronomy texts, but is less common at the professional level. The mass of the Sun, 2.0×10^{30} kg (4.4×10^{30} lb), is used as a standard mass unit. Temperatures are measured using the Kelvin scale. A kelvin (K) is the same as a Celsius degree, but the Kelvin scale starts from absolute zero (−273°C) instead of the freezing point of water. To recast the figures quoted above, then, the observable universe has a diameter of roughly 8400 Mpc and contains 4.4×10^{22} solar masses of material.

Baryonic matter. Most of the matter encountered in everyday life is in the form of atoms. An atom consists of a positively charged nucleus of protons and neutrons, surrounded by clouds or shells of negatively charged electrons. The protons and neutrons are responsible for most of the mass of the atom. Since both protons and neutrons belong to a class of subatomic particles known as baryons, this ordinary form of atomic matter is called baryonic matter by astronomers. The number of protons in an atom's nucleus determines many of its properties, that is, which of the roughly 100 chemical elements that atom represents. *See* BARYON; ELEMENT (CHEMISTRY); ELEMENTARY PARTICLE; ELEMENTS, GEOCHEMICAL DISTRIBUTION OF; PERIODIC TABLE.

A large fraction of the visible matter elsewhere in the universe—planets, stars, nebulae, galaxies—is also baryonic in nature, but the relative proportions of the chemical elements are very different from here on Earth. Hydrogen is by far the most abundant element in the universe, representing 75% of the total baryonic mass. Helium is also plentiful, at about 23% of the total mass. All the heavier elements, collectively referred to as metals (regardless of their solidity or conductivity), make up the remaining 2%, with carbon, nitrogen, oxygen, neon, and iron as the primary constituents. The relative chemical proportions observed on the surface of the Sun (the solar abundances) are used as a reference point for most other abundance measurements (see **table**). The exact proportions present in any given

Solar abundances of the most common elements*

Element	Symbol	Atomic number	Mass fraction	Number fraction
Hydrogen	H	1	0.73	0.92
Helium	He	2	0.25	0.078
Oxygen	O	8	0.0077	0.00061
Carbon	C	6	0.0029	0.00030
Iron	Fe	26	0.0016	0.000037
Neon	Ne	10	0.0012	0.000077
Nitrogen	N	7	0.00095	0.000084
Silicon	Si	14	0.00069	0.000030
Magnesium	Mg	12	0.00047	0.000024
Sulfur	S	16	0.00038	0.000015
Argon	Ar	18	0.00018	0.0000058
Nickel	Ni	28	0.000086	0.0000018
Calcium	Ca	20	0.000058	0.0000018
Aluminum	Al	13	0.000044	0.0000023
Sodium	Na	11	0.000030	0.0000016
Chromium	Cr	24	0.000027	0.00000065
Chlorine	Cl	17	0.000012	0.00000037
Manganese	Mn	25	0.000010	0.00000023
Phosphorus	P	15	0.0000075	0.00000030
Cobalt	Co	27	0.0000058	0.00000012

*The proportion of each element is listed both by mass (fraction of the total mass) and by number (fraction of the total number of atoms).
SOURCE: Data after C. W. Allen, *Astrophysical Quantities*, 3d ed., Althone Press, London, 1991.

astronomical object will depend on its age, environment, and formation history. *See* ELEMENTS, COSMIC ABUNDANCE OF.

Stars and stellar evolution. The most numerous components of the nearby universe, as seen in the night sky, are the stars. These pinpricks of light are actually objects much like the Sun, but they appear faint due to their extreme distance from the Earth. Stars are enormous balls of hot gas (primarily hydrogen and helium), held together by their own gravitation. They are powered by nuclear reactions deep in their interiors, where temperatures and pressures are high enough to fuse hydrogen atoms together into helium, releasing energy in the process. (This process is sometimes referred to as hydrogen burning, although it is quite different from ordinary chemical burning.) A star's total mass, which may be from 0.08 to 100 times that of the Sun, determines nearly all of its other characteristics, including temperature, luminosity (intrinsic brightness), and how long it shines before running out of fuel. The Sun is an average star, with a surface temperature of about 5800 K, and a projected lifespan of 10^{10} years. More massive stars have higher core temperatures and pressures, so the fusion reactions take place much faster, and these stars shine brighter and hotter (thus appearing white or blue in color), expending their fuel reserves in only millions of years. Cool low-mass stars are more frugal with their available core hydrogen, and are expected to shine dimly but steadily, with a dull red color, for hundreds of billions of years. When plotted on a graph of temperature versus luminosity, this trend appears as a diagonal band known as the main sequence (**Fig. 1**). *See* CARBON-NITROGEN-OXYGEN CYCLES; MASS-LUMINOSITY RELATION; NUCLEAR FUSION; PROTON-PROTON CHAIN.

This main-sequence relationship holds only for stars which are burning hydrogen to helium in their cores. Once a star has used up all of its core hydrogen, it must change its internal structure to utilize new fuel sources. First, the hydrogen-burning zone expands outward from the core, and the enhanced power production causes the star to increase significantly in both luminosity and physical size, becoming a red giant or supergiant star. Further increases in core temperature and pressure permit the heavier elements to fuse and release energy, but these fuel sources are less efficient and relatively short-lived. A low-mass star (less than a few solar masses) will burn helium for a time, but as the helium fuel is exhausted and the internal furnace wanes, the outer layers of the star are ejected into space, and the core will gradually shrink into a tiny, dense ember, a white dwarf, glowing only from its residual heat.

Higher-mass stars will burn helium, then carbon, and then a succession of even heavier elements, each for a progressively shorter time, until fusable material of any sort abruptly runs out, and the star collapses catastrophically. Much of the interior mass of the star is compacted into an ultradense core; the outer layers rebound off this core and explode into space, forming a type II supernova, which shines for several weeks at a billion times the luminosity of the Sun. Many of the heavier chemical elements are formed in massive stars and during the supernova explosion itself, and are then spread throughout nearby space to be incorporated into future generations of stars. The stellar core is usually left behind as a neutron star, a small, rapidly rotating body consisting almost entirely of neutrons. Powerful magnetic fields on the surface of a neutron star can produce radio waves, which appear to blink on and off

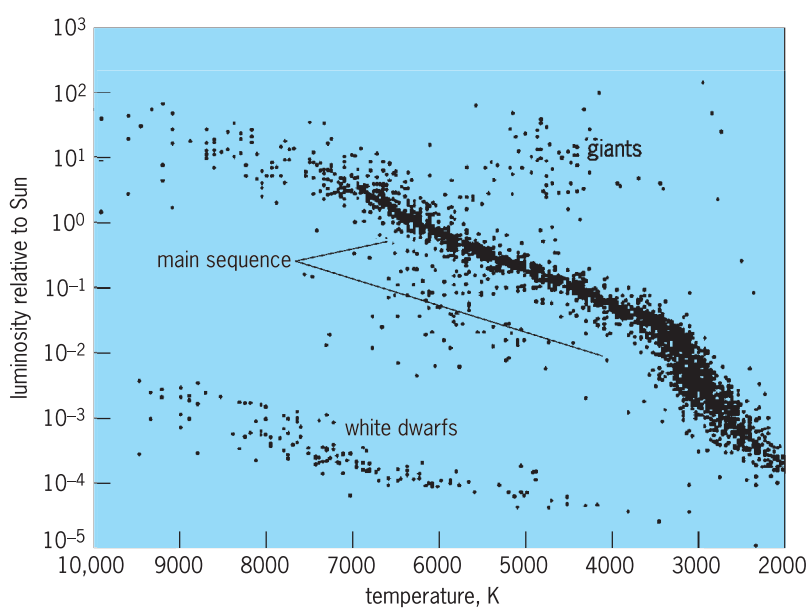

Fig. 1. Hertzsprung-Russell diagram, plotting absolute luminosity against temperature. Each dot represents a star in the local solar neighborhood. The diagonal band is the main sequence, where "normal" core hydrogen burning is taking place. Other regions of the diagram correspond to later stages of stellar evolution. (*Data from W. Gliese and H. Jahreiss, Preliminary Version of the Third Catalogue of Nearby Stars, Astron. Rechen-Institut, Heidelberg, 1991*)

as the star spins. A neutron star of this sort is called a pulsar. In the most extreme cases, the stellar core is compressed so far that it collapses to an infinitesimal point, forming a black hole, and the resulting release of energy is so great that the explosion can be seen from across the universe, appearing as a gamma-ray burst (GRB), a brief but exceptionally powerful pulse of high-energy radiation. The dense end states of stars—white dwarfs, neutron stars, and black holes—are collectively termed compact objects. *See* BLACK HOLE; GAMMA-RAY BURSTS; NEUTRON STAR; NUCLEOSYNTHESIS; PULSAR; STAR; STELLAR EVOLUTION; WHITE DWARF STAR.

Solar system. The Sun is accompanied by a number of smaller objects of various sizes and compositions. The Sun's gravitational domain extends out to almost half a parsec, but its planetary system lies much closer, within about 5×10^9 km (3×10^9 mi) of the center. The innermost planets—Mercury, Venus, Earth, and Mars—are small rocky bodies with solid surfaces and thin atmospheres. Slightly beyond the orbit of Mars is a belt containing several hundred thousand rocky asteroids, fragments of a planet that never quite formed. Farther from the Sun are the gas giants—Jupiter, Saturn, Uranus, and Neptune—with thick, turbulent, gaseous atmospheres surrounding small liquid or solid metallic cores. Pluto, usually considered the most distant of the "official" planets, is solid, but is composed predominantly of ice rather than rock, and is considered by some astronomers to be the largest and closest of a significant population of small icy bodies that orbit beyond Neptune in the Kuiper Belt. Over 750 of these Kuiper Belt Objects are currently known. Even farther outwards lies the Oort Cloud, a spherical shell containing yet more balls of ice, some of which occasionally plummet toward the inner solar system, get heated by the energy of the Sun, and appear as comets, as do some of the Kuiper Belt Objects. *See* ASTEROID; COMET; KUIPER BELT; PLANET; SOLAR SYSTEM.

Extrasolar planets. In recent years, it has become evident that the Sun is not the only star with a planetary system. Disks of particle matter have been seen around recently formed stars, and are thought to be analogous to the initial solar nebula out of which the Sun and the rest of the Earth's solar system formed. Moreover, over 100 nearby stars, most of them similar in type to the Sun, exhibit subtle periodic shifts in their motion, which are due to the gravitational influence of orbiting planetary companions. Most of these other solar systems are configured rather differently from the Sun's, with large planets orbiting close to their star, often even closer than Mercury is to the Sun. Other solar systems with smaller planets or large planets in slower, more distant orbits, may also exist, but measurement techniques must improve before such systems can be detected.

Interstellar material. The "empty" space between the stars actually contains significant amounts of matter—some of it distributed continuously, some of it concentrated in enormous dark clouds—collectively known as the interstellar medium. An atomic hydrogen component is distributed fairly smoothly, although with some variation in temperature and density. Also present are tiny solid dust particles, typically with diameters of 0.00025 cm (0.0001 in.) and smaller, and composed of graphite, silicon oxide, and complex carbon-chain and carbon-ring molecules. *See* INTERSTELLAR MATTER.

Also scattered throughout space are denser clouds of molecular hydrogen (H_2), carbon monoxide (CO), and more complex molecules, sprinkled with the heavier elements. Some of these clouds, or dark nebulae, are many parsecs in size, and contain enough raw material to form a million stars like the Sun. In fact, these clouds often are the site of star formation, when external gravitational disturbances or shock waves trigger the collapse of portions of the clouds. The densest clumps of mass within the cloud, having a stronger gravitational pull, will accumulate more of the surrounding material, increasing the pressure and temperature in the core of the clump to the point where hydrogen fusion can start. Star-forming regions, viewed with instruments such as the Hubble Space Telescope, are some of the most beautiful phenomena in the nearby universe, as new stars emerge from the residual gas surrounding their birthplaces. *See* MOLECULAR CLOUDS; PROTOSTAR.

Sometimes these clouds form self-gravitating spheres of gas which do not have the mass of approximately 0.08 solar mass that is required to support sustained internal fusion, and thus will not shine like stars, but which still glow dimly from the residual heat of their collapse. These "failed stars" or "superplanets" are called brown dwarfs, because they are very dim and red compared to real stars. A large number of brown dwarfs have been discovered since 1995, although not in sufficient numbers that they can account for the dark matter which appears to pervade the universe. *See* BROWN DWARF.

Interstellar space is also populated by high-velocity protons and electrons, which have been accelerated by supernova explosions and other mechanisms to nearly the speed of light. These energetic particles are called cosmic rays. A cosmic-ray proton, despite being just a single subatomic particle, can possess the equivalent momentum of a fast baseball. Most cosmic rays originate within the Milky Way, but some may have reached Earth from other galaxies. *See* COSMIC RAYS.

Galaxies. Stars and interstellar matter are not distributed uniformly throughout the universe, but cluster together in vast units known as galaxies, each containing from 10^7 to 10^{12} stars. Galaxies are categorized into three types—spirals, ellipticals, and irregulars—with numerous subclasses based on size and structure.

The Milky Way, the galaxy containing the Earth, is a typical spiral galaxy—a flattened disk of some 10^{11} stars (**Fig. 2***a*). An oblate central bulge, consisting mostly of older red stars, occupies the central region of the galaxy. The bulge is surrounded by a thin disk of younger stars. Within the disk, a pattern of spiral arms is defined by bright blue stars, recently formed from the gas clouds which also inhabit the disk. Both

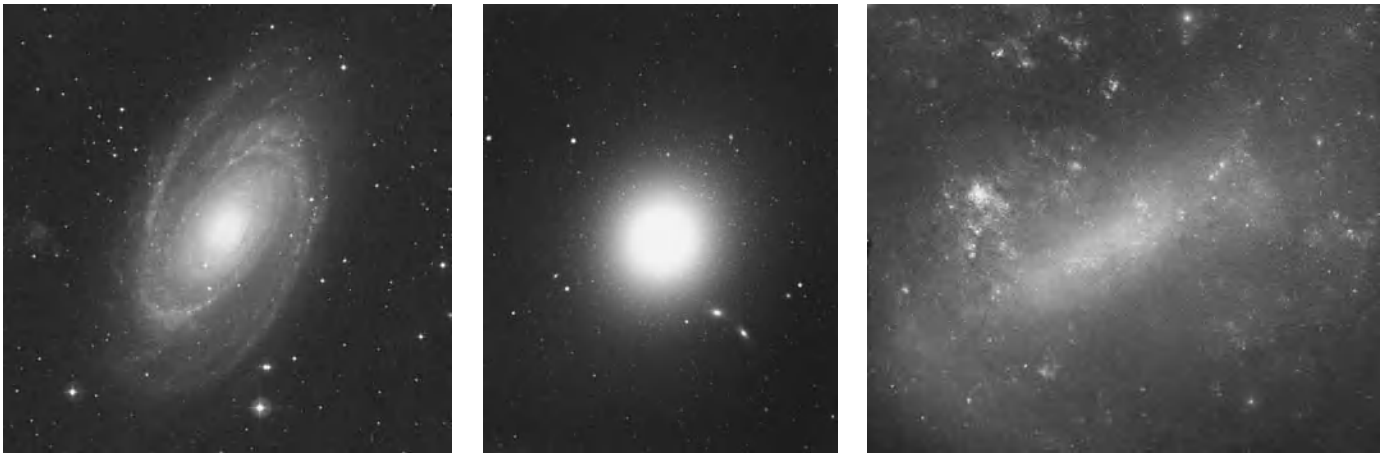

(a) (b) (c)

Fig. 2. Galaxies. (*a*) Spiral galaxy M81 (NGC 3031). The flat disk, spiral arms, and central bulge are visible. The spherical halo that surrounds both is too sparsely populated to be evident (*Digital Sky Survey*). (*b*) Elliptical galaxy M87 (NGC 4486). The central galaxy appears quite spherical. Several hundred small globular clusters (each with thousands of stars) surround M87 (© *Anglo-Australian Observatory; photograph by David Malin*). (*c*) Large Magellanic Cloud, an irregular galaxy near the Milky Way. Bright regions containing hot young stars are found throughout it (© *Anglo-Australian Observatory; photograph by David Malin*).

the bulge and the disk are embedded in the halo, a sparsely populated sphere of very old stars, some of which date back to the formation of the galaxy. The Milky Way, like many other spirals, is accompanied by an assortment of dwarf galaxies and globular clusters (smaller collections of only thousands or millions of stars), which orbit the main galaxy. Dwarf galaxies which approach too closely to the galactic center are ripped apart by the galaxy's gravitational field, and their component stars are absorbed into the halo, suggesting that other such mergers in the past may explain the current size of the Milky Way. The visible portion of the disk extends about 30 kpc from center to edge, although the orbital motions of the stars and gas strongly imply a significant component of nonluminous dark matter both within the visible galaxy and beyond. Other spiral galaxies also exhibit this characteristic, implying that dark matter is a common constituent of galaxies. *See* MILKY WAY GALAXY; STAR CLUSTERS.

Elliptical galaxies are simpler systems than the spirals, consisting mostly of a singe population of cool low-mass stars, arrayed in a spherical or oblate configuration (Fig. 2*b*). Ellipticals have already converted all of their interstellar matter into stars, so that only the long-lived red stars now remain; young blue stars, like those found in spiral arms, have long since run out of fuel and expired. Ellipticals span a much wider range in size than the spirals—dwarf ellipticals, such as those orbiting nearby spiral galaxies, may contain only a few million stars, while the giant ellipticals that inhabit the centers of major clusters have upward of 10^{12} stars. As in spirals, the stars in ellipticals orbit the center of those galaxies at speeds which require significant additional dark mass, beyond that evident in the form of stars, to keep the galaxy from flying apart. Ellipticals are more likely than spirals to be found within galaxy clusters, which suggests that collisions or interactions between galaxies play some role in the formation of the elliptical type.

The irregular type includes all galaxies not readily recognizable as spirals or ellipticals (Fig. 2*c*). Irregular galaxies are similar to spirals in that they often exhibit active star formation, but they do not share the well-ordered structure of spirals. They may be the result of violent galaxy collisions, or perhaps are smaller galaxies that were prevented from developing structure due to gravitational disruption from a nearby larger galaxy. The Milky Way's two largest companions, the Large and Small Magellanic Clouds, are irregulars. *See* GALAXY, EXTERNAL; MAGELLANIC CLOUDS.

Quasars. Supermassive black holes, similar to the stellar-sized black holes which form during the supernova explosion of a massive star but containing millions or billions of solar masses instead of only a few, are thought to exist at the centers of many galaxies, including the Milky Way. Black holes themselves cannot, by definition, emit any light or other radiation, but they can serve as highly efficient engines for converting mass into energy. As stellar or interstellar matter is drawn into the black hole, it forms an accretion disk around the central point and is heated by friction to temperatures of millions of kelvins, so that it glows brightly with a luminosity equivalent to 10^{12} Suns. At the far reaches of the universe, these galaxy cores appear as faint starlike points of light, and were thus labeled quasistellar objects or quasars when first discovered. Nearby galaxies such as the Milky Way do not appear as quasars, because the flow of fuel to the black hole has slowed to a trickle, and the central engine shines only dimly. Farther away, however, the ultraluminous galaxy cores are seen as they were many billions of years ago, when the central engines were still being supplied with copious fuel. The most distant quasars known today are located more than 90% of the way to the edge of the observable universe. *See* QUASAR.

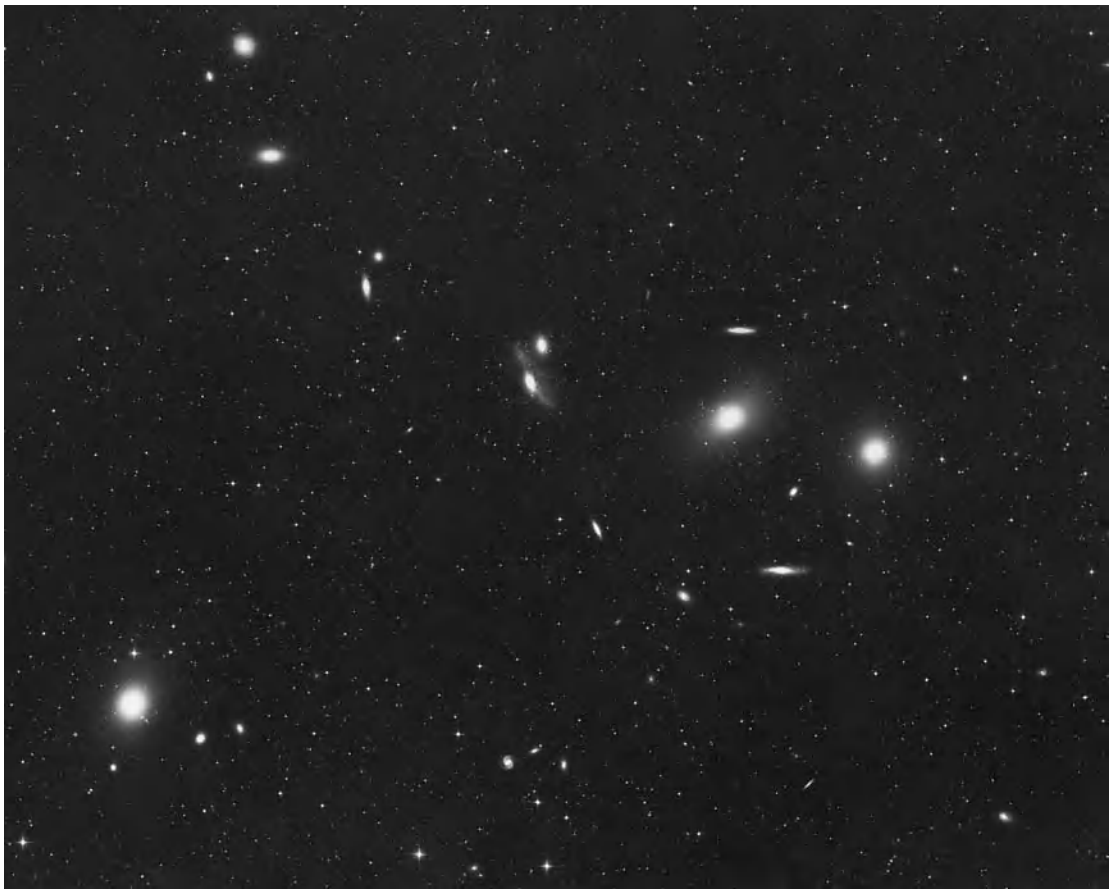

Fig. 3. Heart of the Virgo Cluster, the nearest moderately rich cluster of galaxies, including the giant elliptical galaxies M84, M86, and M87. (© *Anglo-Australian Observatory; photograph by David Malin*)

Groups, clusters, and large-scale structure. Galaxies themselves are usually bound together in groups (up to about 50 galaxies) and clusters (50 to thousands of galaxies) spanning regions 2–10 Mpc in diameter (**Fig. 3**). Like individual galaxies, groups and clusters are gravitationally self-contained systems, only on a larger scale. Clusters must contain more mass than is immediately evident—the total mass of the visible galaxies in a typical cluster is only a small fraction of the amount needed to keep the fast-moving galaxies from flying out of the cluster. There must therefore be some other mass component in addition to the galaxies, contributing its gravitational potential to keep the cluster together. Some of this "missing mass" can be attributed to hot diffuse gas, the intracluster medium, which is visible to x-ray and ultraviolet telescopes and fills the space between cluster galaxies, but the nature of the rest of the dark matter is yet to be explained. *See* LOCAL GROUP; VIRGO CLUSTER.

Clusters accumulate into yet larger entities called superclusters. Typical superclusters are on the order of 100 Mpc in size, and often take elongated or filamentary shapes. Recent redshift surveys, astronomical programs to measure the velocities of thousands or millions of distant galaxies and thus estimate their distance, have provided glimpses of these largest known levels of structure in the universe (**Fig. 4**). Groups and clusters seem to be concentrated in thin sheets, surrounding enormous voids with very few galaxies. Superclusters sit at the edges and vertices where surfaces intersect. The arrangement has been likened to a froth of soap bubbles, and may provide vital clues to the mechanisms by which structure formed out of the initially homogeneous early universe.

Antimatter. All subatomic particles have oppositely charged antiparticle counterparts. When a particle and its corresponding antiparticle collide, they annihilate, converting all their mass into energy. Both matter and antimatter are expected to have been formed in the early universe, but clearly not in precisely equal amounts, since the observable universe is predominantly matter. If slightly more matter were formed than antimatter, as predicted by current theory, then later matter-antimatter annihilation would leave behind only the "extra" matter, as observed today. Trace amounts of antimatter are still found in some astronomical environments, where new particle-antiparticle pairs are created out of free energy, in the opposite reaction from annihilation. Collisions between cosmic rays or high-energy photons can supply such energy, creating electrons and their antiparticles, positrons. *See* ANTIMATTER; ELECTRON-POSITRON PAIR PRODUCTION.

Nonbaryonic particles. Although most observable matter is baryonic in nature, several kinds of nonbaryonic matter also exist in the universe. Some

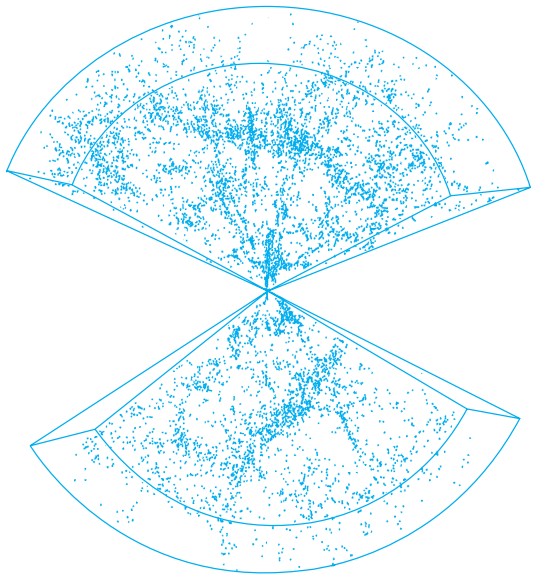

Fig. 4. Large-scale distribution of galaxies in two opposing regions of the sky, as measured by extensive redshift surveys. Each dot represents a galaxy. (*From L. N. Da Costa et al., A complete southern sky redshift survey, Astrophys. J., 424:L1–L4, 1994*)

nonbaryonic particles have been experimentally observed, and play a crucial role in explaining many processes, from nuclear fusion to supernova explosions. Other nonbaryonic particle types have been proposed on the basis of theory but have not yet been detected directly. According to current astronomical measurements, nonbaryonic matter accounts for approximately 23% of the total density of the universe, so it is important to determine the identity and characteristics of this component of the universe.

Electrons, strictly speaking, are not baryons, but they are considered as an integral part of ordinary atomic matter, and thus are usually not categorized with the more exotic nonbaryonic species described below. The net electric charge of the universe is assumed to be zero, so there are the same number of electrons as protons, but since the proton mass is almost 2000 times that of the electron, electrons make only a small contribution to the total cosmic mass. *See* ELECTRON.

A large number of other nonbaryonic particles, including mesons, muons, and taus, have been discovered in particle physics experiments. Many of these particles are quite massive (by subatomic standards), but they also have very short lifetimes—they rapidly decay into less massive particles. As such, they are considered unimportant for cosmology. *See* PARTICLE PHYSICS.

Neutrinos are electrically neutral, very low mass particles that are generated in nuclear reactions. They were once thought to be completely massless entities, but theoretical and experimental evidence now suggests that they have a tiny but nonzero mass, approximately 2×10^{-34} g (4×10^{-37} lb). Neutrinos do not interact strongly with baryons; they pass right through most ordinary matter, and even stars and planets are virtually transparent to them. This characteristic makes them useful tools for studying otherwise inaccessible environments, like the fusion core of the Sun or the centers of supernovae, but this also makes neutrinos difficult to detect and measure, as they usually pass through any sort of detection device without leaving a trace. Projects to detect neutrinos from the Sun and other astronomical sources have met with some success in recent years, offering an entirely new means of studying the universe. *See* NEUTRINO; NEUTRINO ASTRONOMY; SOLAR NEUTRINOS.

According to new theories of particle physics, a variety of other nonbaryonic particle species might also exist. Some of these hypothetical particles could be much heavier and slower than neutrinos, but like neutrinos, would only interact with baryonic matter via the gravitational force. The most likely candidate to fill this role is the so-called weakly interacting massive particle (WIMP). Lightweight axions are another possibility, along with magnetic monopoles, WIMPzillas, and neutralinos. Experiments attempting to directly detect WIMPs and other particles are currently under way, but have not yet proven the existence of any of these exotic entities. *See* MAGNETIC MONOPOLES; WEAKLY INTERACTING MASSIVE PARTICLE (WIMP).

Dark matter. A large fraction of the mass of the universe is in some form which cannot be seen, but which is evident from its gravitational effect on the motions of bright objects, such as stars and galaxies. This dark matter, or "missing mass," is present on many different scales, from galaxies to the universe as a whole. Several possibilities for the nature of the dark matter have been suggested. Numerous small dim baryonic objects, such as faint red stars, brown dwarfs, or black holes, located in the halos of galaxies, might account for some of the galactic-scale dark matter. However, searches for these MACHOs (massive compact halo objects) have turned up far fewer confirmed detections than necessary to explain all the missing mass. Furthermore, several independent lines of argument indicate that baryons can comprise only about 4% of the total mass-energy density of the universe, while the dark matter has to make up approximately 23% of the cosmic total in order to properly explain the distribution and motion of galaxies. Dark matter must therefore consist primarily of nonbaryonic particles. Neutrinos were once considered as likely candidates, but recent observations of the cosmic microwave background argue against this possibility, because neutrinos travel at nearly the speed of light and would spread out too quickly to permit the observed lumpiness of matter in the early universe. Slower-moving ("cold") particles such as WIMPs are therefore favored by current theories. Active theoretical and experimental efforts are under way to determine the characteristics of such particles, and the effect they would have on the distribution and dynamics of the luminous matter, but direct detection of these particle species will be very challenging. The identification of dark matter remains one of the key unresolved issues of modern astrophysics. *See* DARK MATTER.

Energy. Energy is a physical entity as real as matter, but somewhat less tangible, which makes it more difficult to categorize easily. Like matter, energy comes in many different forms, which can be readily transformed from one to another. Energy can also be converted to and from matter. Various forms of energy which play a significant role in the universe—gravitational, electromagnetic, kinetic, nuclear, and dark energy—are described below. *See* ENERGY.

Gravity is the predominant physical force in most astronomical contexts, and is thus responsible for the existence of most of the universe's structure, including stars, planets, galaxies, clusters, and superclusters. Gravitational potential also drives many of the key energy-generating mechanisms in the cosmos, such as nuclear fusion in stellar cores and accretion disks in quasars. *See* GRAVITATION.

Energy is readily transmitted throughout the universe via electromagnetic radiation, which may take the form of visible light, radio waves, x-rays, gamma rays, or any other part of the electromagnetic spectrum. Electromagnetic energy may be described either as waves—intertwined electric and magnetic fields rippling through empty space—or as particles called photons, which travel in straight lines like small solid objects. The universe now contains about 400 photons per cubic centimeter (6500 per cubic inch), so photons outnumber atoms by a large margin. *See* ELECTROMAGNETIC RADIATION; LIGHT; PHOTON.

Matter may contain energy (in addition to the rest-mass energy defined by relativity) by virtue of its motion. This kinetic energy may be due to large-scale bulk motion, such as a planet hurtling through space, or it may be in the form of thermal energy—the microscopic motion of an object's constituent atoms, vibrating or moving relative to each other. The higher the temperature, the faster the average motion, and the more energy is present. *See* KINETIC THEORY OF MATTER.

Large reserves of energy are stored in the force which holds atomic nuclei together. There are actually two such forces, the strong nuclear force and the weak nuclear force, which play different roles at the subatomic level. Although considerably weaker than the strong force, the weak nuclear force plays an important role in radioactive decay and the formation of isotopes of the heavy elements. The strong force maintains the nuclear structure upon which matter is based, but some of this energy can be released in other forms during nuclear reactions. When two hydrogen atoms are fused together to form helium, for instance, the hydrogen nuclei have a greater total binding energy than the resulting helium nucleus, and the difference is released as free energy (in the form of photons and neutrinos). Nuclear fusion is the process which powers all luminous stars in the universe. Energy release can also take place when a large heavy nucleus splits into two lighter nuclei with less net energy. This is the mechanism behind nuclear fission, as found in nuclear power plants and uranium/plutonium bombs. But because of the relative scarcity of heavy fissionable elements in the universe, fission does not play any appreciable role in astronomical environments. *See* NUCLEAR FISSION; NUCLEAR REACTION; RADIOACTIVITY; STRONG NUCLEAR INTERACTIONS; WEAK NUCLEAR INTERACTIONS.

Research suggests that, on cosmological scales, gravitational attraction is countered by a previously undetected form of energy, known as dark energy or the cosmological constant, which causes space to expand. A force of this sort had originally been suggested by Einstein's equations of general relativity, but the strength of this force was assumed to be zero, because no excess expansion rate had been observed in the motions of galaxies. Since 1998, however, observations of distant supernovae and the cosmic background radiation have led astronomers to the conclusion that the expansion of the universe is indeed accelerating, and that dark energy is, in fact, the dominant component of the universe. According to these data, the dark energy comprises 73% of the total density of mass and energy in the universe, "outweighing" the dark matter and baryonic matter by a factor of 3. The actual nature of the dark energy is still a mystery. *See* COSMOLOGICAL CONSTANT; DARK ENERGY; RELATIVITY.

The traditional four basic forces of nature—gravity, electromagnetism, and the strong and weak nuclear forces—may in fact all be different facets of a single fundamental force. Particle physics experiments show that under conditions of extremely high energy the weak nuclear force and electromagnetism merge into a single "electroweak" force. Theoretical efforts are being made to devise a grand unified theory under which the strong force is also incorporated, and a still more general force that also includes gravity. Such a unified force may have existed during the moments following the big bang, and then fragmented into separate forces as the universe expanded and cooled. With the discovery of the long-range "antigravity" effect of dark energy, the task of combining all of the known physical forces into a single coherent framework has become even more challenging. *See* ELECTROWEAK INTERACTION; FUNDAMENTAL INTERACTIONS; GRAND UNIFICATION THEORIES; STANDARD MODEL; SUPERSYMMETRY.

Origin, Evolution, and Fate

The universe is a dynamically evolving system; there are compelling reasons to believe that it has not always been in its current state. Most of the evidence for this assertion is indirect, as humans have not observed the cosmos for long enough to actually see things change on large scales. But by closely studying the current distribution and motion of matter and energy, and collecting the "fossil light" from distant objects, scientists have constructed a consistent picture of the creation of the universe, the big bang theory, which explains the observations fairly well. Many of the specifics are still unresolved, however, and are currently the subject of research and debate.

Big bang. The observable universe originated 13.7 billion years ago in a fiery cataclysm termed the big bang. This was not an explosion of compressed

matter and energy into a previously empty space. Instead, space itself, as well as everything contained therein, sprang from a single point of near-infinite density and temperature, and grew to the volume observed today. As it expanded, it cooled, allowing familiar forms of matter to condense from the high-energy "soup" of energy and subatomic particles that constituted the very early universe. The framework for the present-day distribution and dynamics of matter in the universe was also established at these very early times, and thus remnants of the extreme conditions that existed at the beginning can still be seen. Much of the current understanding of the big bang and the early universe hinges on two key observations: velocities of distant galaxies and the cosmic background radiation. *See* BIG BANG THEORY.

Cosmological redshifts. The relative velocity of an astronomical object, toward or away from an observer, can be determined very precisely by measuring the color distribution of the light from the object. Distinct spectral features, characteristic of the different chemical elements that are present, appear at well-established locations in the spectrum, but the entire spectrum is shifted toward shorter wavelengths (bluewards) for an object approaching the observer, or toward longer wavelengths (redwards) for a receding object. The amount of this Doppler shift is closely related to the actual relative speed of the object. *See* DOPPLER EFFECT.

If galaxies were moving randomly in the universe, observers would expect to see just as many blueshifted galaxies as redshifted galaxies, roughly half moving toward Earth and half moving away. But measurements of galaxy motions made in the early twentieth century showed a very different distribution: With the exception of a few nearby galaxies, all galaxies exhibited redshifts of varying amounts, implying that they were all moving away from Earth. Edwin Hubble extended this work by determining the distances to these galaxies, and found that recession velocity was directly proportional to distance—that the more distant galaxies were moving away faster (**Fig. 5**). The constant of proportionality relating speed and distance now bears Hubble's name, and an accurate determination of this Hubble constant (H_0) has been one of the central pursuits of modern astronomy for the past several decades. Recent studies converged to a value of H_0 = 72 kilometers per second per megaparsec.

The observation that nearly all galaxies are moving away from Earth's location in space might suggest that Earth is at a special location in the universe, at or near the center of an expanding distribution of matter. It is highly unlikely, however, that Earth would happen to be located at such a special position. Fortunately, these observations can be explained in another fashion. If space itself is expanding uniformly in all directions, and the galaxies are being carried along in this general expansion, then any point in the universe would see all other points moving away. The greater the distance between two points, the more space exists between them, and the faster this distance increases. Hubble's law is therefore a consequence of the uniform expansion of space, which causes more distant galaxies to exhibit larger redshifts because they are receding from the observer faster.

If space is expanding uniformly in all directions, then in the past the universe was smaller. All the matter and energy that is observed in the current-day cosmos would be present in this smaller volume, so the density (average mass per volume) and temperature (related to the average energy per volume) of the universe would be higher than they are today. At even earlier times, the volume would be yet smaller, and the density and temperature even higher, and at some point in the past all matter and energy may have existed in a single point of infinite temperature and density. *See* HUBBLE CONSTANT; REDSHIFT.

Cosmic background radiation. If this picture is correct, and the current universe was spawned from just such a primordial fireball, then the sky should be filled with a residual afterglow from the era when the matter in the universe was hot and emitting strongly, in much the same way that the surfaces of stars shine today. This primordial radiation will have cooled dramatically, though, since the space containing it has expanded since the radiation was emitted. From its initial characteristic temperature of about 3000 K, this cosmic background radiation, or cosmic microwave background, is expected to have cooled to about 3 K above absolute zero, so that instead of existing mostly as visible light, the energy distribution will have shifted far to the red, into the microwave radio region of the spectrum.

This microwave background was first detected by A. Penzias and R. Wilson in 1965 as persistent low-intensity static in microwave communications equipment, and was identified as being of cosmic origin. Subsequent ground-based measurements of the energy distribution of the cosmic background radiation suggested a characteristic temperature of approximately 3 K, but it was not until the *Cosmic Background Explorer* (*COBE*) satellite mission

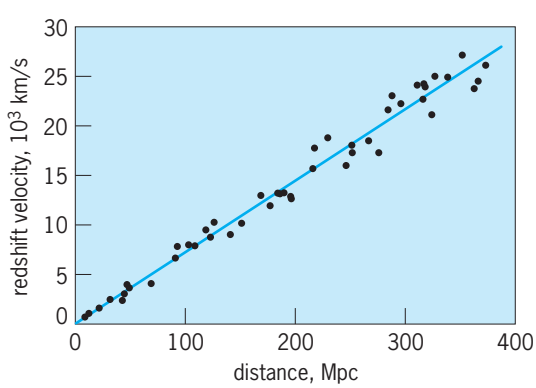

Fig. 5. Hubble diagram showing linear relationship between galaxies' distances and recession velocities. Each point represents one galaxy. The slope of the best-fit line gives the Hubble constant H_0. (*Data from A. G. Reiss et al., Obsevational evidence from supernovae for an accelerating universe and a cosmological constant, Astron. J., 116:1009–1038, 1998; and S. Perlmutter et al., Measurements of Ω and Λ from 42 high-redshift supernovae, Astrophys. J., 517:565–587, 1999*)

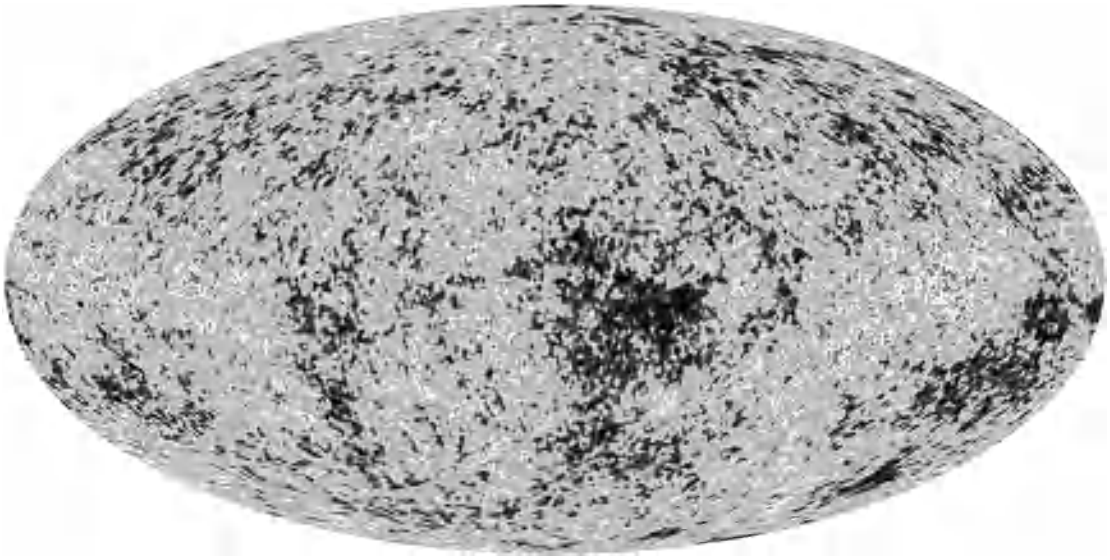

Fig. 6. Full-sky projection map of temperature variations in the cosmic background radiation, as measured by the *Wilkinson Microwave Anisotropy Probe* (*WMAP*). (*NASA/WMAP Science team*)

in 1990 that the spectrum was measured simultaneously over a wide range of wavelengths, firmly establishing the cosmic background radiation temperature at 2.726 K.

Since the cosmic background radiation is an intrinsic characteristic of the universe, observers expect to see it reaching the Earth uniformly from all directions in space, and the *COBE* measurements confirmed that this is indeed the case, once the Doppler effect from the Earth's motion through space is subtracted. Subsequent observations of the cosmic background radiation by highly sensitive microwave telescopes, including the *Wilkinson Microwave Anisotropy Probe* (*WMAP*), have revealed faint ripples in the temperature distribution of the background radiation. These patterns in the cosmic background radiation are a result of the very first clumps of matter that accreted in the early universe, as regions with slightly different densities emitted slightly hotter or cooler radiation. By measuring the amplitude and angular size of these temperature variations, researchers have been able to deduce the age, total density, and early history of the universe with unprecedented precision. *See* COSMIC BACKGROUND RADIATION; WILKINSON MICROWAVE ANISOTROPY PROBE.

Evolution of the universe. The best models for the beginning of the universe start at 10^{-43} s after the big bang itself. The conditions before this point were so extreme that the current understanding of physics is insufficient to say anything meaningful (although new theories are being developed to address this limitation). At 10^{-42} s, the universe was 10^{45} times smaller than it is today (approximately 6×10^{-18} cm, or 2×10^{-18} in.), and had a mean temperature of 10^{30} K. At such temperatures, matter as presently known cannot exist, because the energies are so enormous that even the protons and neutrons themselves are torn apart into separate sub-subatomic particles called quarks. The universe was a featureless mixture of subatomic particles and high-energy photons. *See* QUARKS.

As the universe expanded and cooled, however, more particles were permitted to form, once they could exist without being immediately destroyed. At 10^{-34} s, when the temperature had dropped to 3×10^{26} K, heavier exotic particles like magnetic monopoles could have emerged. Around this time, the rapidly enlarging structure of space may have undergone an era of even faster expansion, driven by the energy of space itself. This era of hyperfast inflation helps to explain features of the present-day observable universe, such as the uniformity of the cosmic background radiation over the entire sky, and the way in which the average mass density of the universe is high enough for structures such as stars and galaxies to form, but not so large that it would immediately recollapse upon itself. *See* INFLATIONARY UNIVERSE COSMOLOGY.

When the temperature had dropped further, to about 3×10^{12} K, protons and neutrons condensed out of the quark mixture. It was still much too hot for electrons to join them and form atoms, but more complex atomic nuclei—deuterium (proton plus a neutron), helium-3 (2 protons plus a neutron), helium-4 (2 protons plus 2 neutrons), and lithium (3 protons plus 4 neutrons)—were created, in a fashion similar to stellar core fusion. This era of big bang nucleosynthesis established the initial composition of the universe, about three-quarters hydrogen, one-quarter helium, and a smattering of lithium, from which all subsequent stellar and supernova nucleosynthesis has proceeded.

Temperatures were still too high for electrons to join these nuclei to make complete atoms, until a time about 370,000 years after the big bang. Up to this point, the universe was relatively opaque, since unattached electrons are very good at absorbing light and other electromagnetic radiation. Once the temperature fell below 3000 K, however, protons

and electrons could combine into neutral hydrogen atoms, and the universe suddenly became transparent to most wavelengths. The cosmic background radiation formed at this point, and has had little interaction with matter since, which is why observations with microwave telescopes such as *COBE* and *WMAP* reveal such detailed information about this epoch.

The formation of structure—the development of clumpiness in the universe, which appears today as stars, galaxies, and clusters—started slightly prior to matter-radiation decoupling. These very first clumps of mass left their faint imprint on the temperature distribution of the cosmic background radiation, and then continued to accumulate and grow after becoming independent from the radiation field. The exact sequence of structure formation—whether many smaller objects such as stars formed first and then accumulated to form galaxies and clusters, or whether matter clumped on galaxy- or cluster-sized scales and then subdivided into smaller bodies—is still not well understood, because of the scarcity of luminous objects and the difficulty of detecting them at such extreme distances. Sophisticated computer simulations, combined with indirect evidence from observations of the cosmic background radiation and the most distant quasars, suggest that massive stars (100 times the mass of the Sun, or greater) formed relatively early and then exploded as supernovae, pumping energy and newly formed heavier elements into the surrounding clouds of hydrogen and helium, and thus setting the stage for the formation of galaxies and stars observed today. Observations from future generations of space telescopes will help resolve many of the uncertainties regarding this era of cosmic history.

Eventually, within 10^9 years of the big bang, larger structures such as galaxies did take shape, and within them subsequent generations of stars coalesced from interstellar clouds of hydrogen. The Sun (along with its system of planets) appeared about 5×10^9 years ago. In elliptical galaxies, nearly all the gas has been turned into stars by now, while spiral and irregular galaxies continue to spawn new stars at a steady rate. Large galaxies gradually absorb and consume any dwarf galaxy companions that approach too closely, and occasionally two large galaxies will collide and merge. Meanwhile, the overall expansion of the universe continues.

Ultimate fate of the universe. After another 5×10^9 years, the Sun will run out of hydrogen fuel in its core, and after swelling up into a red giant, will collapse into a white dwarf, gradually fading in brightness as it cools. Most of the other stars that are seen in the night sky will meet a similar fate, some of them within a few million years, others lasting hundreds of billions of years before their fuel reserves run out. New generations of stars will be formed to take their place, but over time the interstellar hydrogen reserves of spiral galaxies such as the Milky Way will be exhausted, and the formation of new stars will cease, leaving only the long-lived dim red stars.

On larger scales, the universe will probably continue to expand indefinitely. It was once thought that if the average cosmic density of matter were high enough, the mutual gravitational attraction would be sufficiently strong to gradually slow the expansion to a standstill, and then cause the universe to contract, eventually collapsing in a "Big Crunch" (or "Gnab Gib"). According to the current tally of baryonic and dark matter, however, there is not enough gravitational mass in the universe to overcome the current rate of expansion. With the added repulsive force of the dark energy, the expansion rate will increase, instead of decrease. Objects such as stars, galaxies, and clusters, which are kept together by their own self-gravity, will remain largely intact, but the enormous voids separating clusters and superclusters from each other will grow ever larger. Eventually, over time scales of trillions of years, the universe will grow empty and dim, illuminated by only a few last stellar remnants, cooling slowly toward absolute zero.

Conclusion. Over the past several centuries, astronomers have assembled a fairly comprehensive and consistent view of the universe. Many of the major constituents are now known, and modern astronomical theories offer plausible explanations of how they were formed, what they are made of, and how they interact with each other. There are, however, a number of key unanswered questions, such as the nature of dark matter and dark energy, which will require many further advances in technology, observation, and theory to fully resolve. *See* COSMOLOGY.

Bradford B. Behr

Bibliography. B. W. Carroll and D. A. Ostlie, *An Introduction to Modern Astrophysics*, Addison-Wesley, 1996; E. W. Kolb and M. S. Turner, *The Early Universe*, Addison-Wesley, 1990.

Unloader

A power device for removing bulk materials from railway freight cars or highway trucks. Especially in the case of railway cars, the car structure may aid the unloading, as in the hopper-bottom dumper; here the unloader may be a vibrator which improves flow of the bulk material into a storage bin or to a conveyor. Thus an unloader is the transitional device between interplant transportation means and intraplant materials-handling facilities. *See* MATERIALS-HANDLING EQUIPMENT; TRANSPORTATION ENGINEERING.

Basically, design considerations for an unloader are similar to those for a device for removing bulk from a storage bin or warehouse with low headroom and limited access. A special consideration is the possible need for an elevating platform to bring the car or truck floor to the level of the unloading dock and to maintain the alignment as material is removed and the vehicular springs expand. Unloading may then be accomplished by portable belt, drag-chain, flight, or similar conveyor (**illus.** *a*). *See* BULK-HANDLING MACHINES.

A fully mechanized unloader includes a section of track to which the car is clamped and then

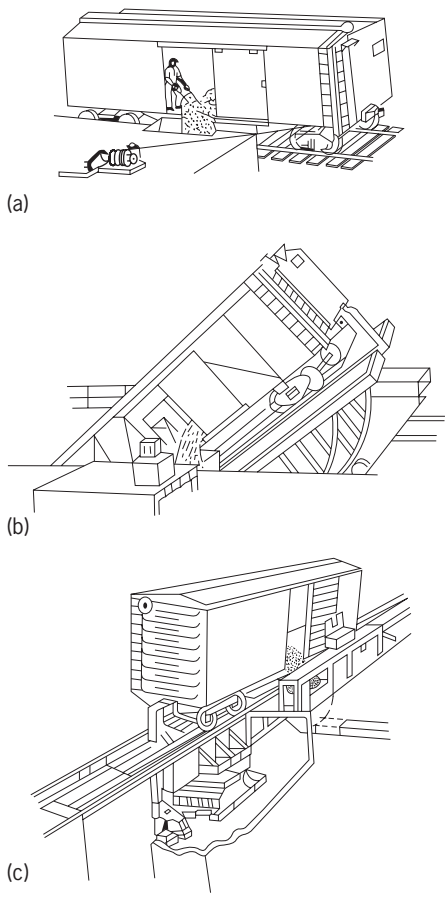

Unloader. (*a*) Power scoop. (*b*) Rocker. (*c*) Tilter. (*Link-Belt Co.*)

rocked or tilted (illus. *b* and *c*). In this way unloading rates of up to ten boxcars per hour are achieved. For less specialized unloading, dock equipment such as power scoops may be used, or yard handling equipment such as boom unloaders may be adapted.

Frank H. Rockett

Upper-atmosphere dynamics

The motion of the atmosphere above 50 km (30 mi). The predominant dynamical phenomena of the upper atmosphere are quite different from those encountered in the lower atmosphere. Among those encountered in the lower atmosphere are cyclones, anticyclones, tropical hurricanes, thunderstorms and shower clouds, tornadoes, and dust devils. Even the largest of these phenomena do not penetrate far into the upper atmosphere. Above an altitude of about 50 km (30 mi), the predominant dynamical phenomena are internal gravity waves, tides, sound waves (including infrasonic), turbulence, and large-scale circulation. These are described below.

Internal gravity waves. Except under meteorological conditions characterized by convection, the atmosphere is stable against small vertical displacements of small air parcels; this results from buoyancy forces that tend to restore displaced air parcels to their original levels. An air parcel therefore tends to oscillate around its undisturbed position at a frequency known as the Brunt-Vaisala frequency; the square of this frequency is given by Eq. (1), where

$$\omega_B^2 = -g\left(\frac{\rho_0'}{\rho_0} + \frac{g}{C^2}\right) \tag{1}$$

ρ_0 is the undisturbed density, ρ'_0 is the vertical gradient of the density, g is the acceleration of gravity, and $C = (\gamma p_0/\rho_0)^{1/2}$ is the speed of sound; γ is the ratio of the specific heats, and p_0 is the undisturbed pressure. Typical periods for the Brunt-Vaisala oscillation are in the vicinity of 5 min.

If pressure waves are generated in the atmosphere with frequencies much greater than ω_B, they propagate as sound waves. For frequencies much less than ω_B, the waves propagate as internal gravity waves; in this case, the restoring forces for the wave motion are provided primarily by buoyancy (that is, gravity) rather than by compression. These waves have peculiar properties. The wave motion is nearly horizontal, but the inclination from the horizontal is sufficient so that gravity provides the restoring force for the displaced air parcels. The phase velocity of the waves is downward, but the group velocity is upward. Such waves are presumably excited near the Earth's surface, probably in wind systems, and the energy flow is upward, reaching well into the ionosphere. Owing to the decrease in density with altitude, the amplitudes of the waves must increase with altitude in order to preserve continuity of energy flow; above 60 km (36 mi), the amplitude is so large that it dominates the observed wind profiles. An example is given in **Fig. 1**, in which a wavelike structure with a vertical wavelength of about 20 km (12 mi) can be seen. In the lower atmosphere, the gravity waves are presumably also present, but with amplitudes so small that they cannot be detected.

Above 100 km (60 mi), dissipative mechanisms come into play for internal gravity waves; these include viscosity, thermal conductivity, and eddy transport of heat and momentum. The dissipation of the waves, judged from their observed amplitudes below 100 km (60 mi), provides an energy source which is apparently large enough to be of some significance in the heat budget of the upper atmosphere.

Internal gravity waves should not be confused with surface waves, which also occur within the atmosphere. The surface waves occur at discontinuity surfaces, or in regions where the gradient is strong enough to guide the propagation of the wave in a direction perpendicular to the gradient. Such waves are frequently visible on cloud sheets.

Tides. Tides are internal gravity waves of particular frequencies. The term tidal usually implies that the exciting force is gravitational attraction by the Moon or Sun. However, it is conventional in the case of atmospheric tides to include also those waves that are excited by solar heating. One is therefore concerned with three separate excitation functions—lunar gravitation, solar gravitation, and solar heating.

Examination of barometric pressures in tropical latitudes normally discloses two maxima and two

minima per day, the maxima occurring near 1000–2200 h local time; at higher latitudes the variations are smaller and the noise due to passing weather systems is much larger, and statistical analyses are needed to find the tidal oscillations. The gravitational tidal force of the Moon is 2.2 times larger than that of the Sun, and its period is 12 h 26 min. The fact that only the 12-h oscillation is readily seen on a barograph record therefore indicates that the oscillation is of thermal origin, or that the oscillation is so sharply resonant that the 12-h tidal force of the Sun excites a much stronger oscillation than the larger 12-h 26-min tidal force of the Moon. The possibility of a thermally excited oscillation of period 12 h arises because the solar heating curve has harmonics, that is, it is not a simple sinusoidally oscillating function. Although the possibility of a resonant excitation of the 12-h oscillation by solar gravitational force has been considered and the resonant period of the atmosphere is about 12 h, the resonance is now thought to be much too broad to discriminate between the solar and lunar gravitational tidal periods. Both are amplified several times by the resonance, but the thermal excitation must be responsible for the major observed effect.

Although the lunar tide of period 12 h 26 min is not discernible on a barograph trace, it can be found from a statistical analysis, and it is about 15 times smaller than the 12-h oscillation.

Although tidal oscillations are small at the Earth's surface, they are relatively large at high atmospheric levels. Only a few days' observations of the winds near the altitude of 100 km (60 mi) is required to disclose the diurnal pattern of behavior in which the largest components have 24- and 12-h periods; the amplitudes of these tidal winds are about 100 times greater than those at the Earth's surface. The tidal wind patterns in the upper atmosphere also disclose their presence in another way, by generating electrical currents in the ionosphere through a dynamo action. These in turn give rise to diurnal variations in the geomagnetic field that can be observed at the Earth's surface. The lunar tidal oscillations also increase with altitude, but they are smaller than the solar tidal oscillations (about 2 m/s or 4 knots as compared to 20 m/s or 40 knots), so that they have not been recognized in the limited amount of wind data available. However, they are recognizable in terms of the geomagnetic variations that they produce, although a statistical analysis is needed to find them. *See* ATMOSPHERIC TIDES; GEOMAGNETIC VARIATIONS; IONOSPHERE.

Sound waves and infrasonics. Sound waves generated in the lower atmosphere may propagate upward; to maintain continuity of energy flow, the waves might be expected to grow in relative amplitude as they move into the more rarefied upper atmosphere. However, higher temperatures in the upper atmosphere refract most of the energy back toward the Earth's surface, giving rise to the phenomenon known as anomalous propagation. This involves the redirection of upward-moving sound waves back to the surface beyond the point where the source (usually a large explosion) can be heard by waves propagating along the surface. Thus there is normally a zone of silence, beyond which the source can again be heard because of anomalous propagation.

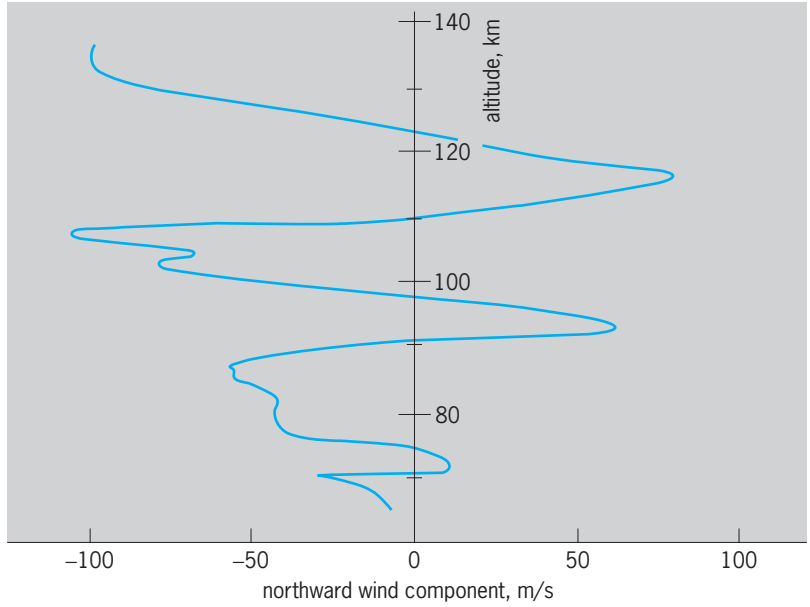

Fig. 1. Northward wind component over Wallops Island, Virginia, on June 6, 1962. 1 km = 0.6 mi; 1 m/s = 2 knots.

Sound waves or infrasonic waves (frequencies too low to be audible) generated by a thunderstorms have been considered possible sources of energy input in the upper atmosphere as the waves become shock waves on entering more rarefied regions of the atmosphere. The refraction effect mentioned above makes this unlikely, but the possibility has remained under consideration. The energy spectrum of sound waves generated by thunderstorms has been measured. A broad peak near 200 Hz was found, with no energy increase in the infrasonic region.

Infrasonic waves with periods from 20 to 80 s have been observed occasionally with detectors at the Earth's surface in connection with auroral activity, the average rate of observed occurrence being about once a day. The occurrence seems to be associated with rapidly moving auroral forms where the rate of movement is supersonic. Although this rapid motion does not represent a supersonic motion of atmospheric gas, but rather only a movement of the energetic particle source causing the aurora, nevertheless, a shock wave should be produced by the supersonic motion of the generating mechanism, the aurora. The observed direction of arrival agrees with this concept of an auroral source, the preferred direction of propagation being the direction of motion of the auroral forms. The observed amplitudes range up to 10 dynes/cm^2 (1 pascal). *See* SOUND.

Turbulence. There is clear visual evidence of turbulence in the upper atmosphere; this evidence is obtained by examination of vapor trails released from rockets or of long persisting meteor trails. The trails spread above about 110 km (66 mi) by an amount

that corresponds to molecular diffusion, but at lower altitudes the spreading is far in excess of that which can be attributed to this mechanism, and turbulence is the only known means of explaining the additional spreading. Further, in the region below 110 km (66 mi), the trails adopt a turbulent appearance, whereas at higher altitudes, where molecular diffusion predominates, the trails are smooth, as should be expected.

The source of the turbulence is not clear. The atmosphere is thermodynamically stable against vertical displacements throughout the region above the troposphere, and work has to be done against buoyancy forces in order to produce and maintain turbulence. The only apparent source of energy is internal gravity waves, either tidal or of random period. One concept is that wind shear associated with such waves may become great enough to overcome the stability and produce turbulence; the relevant criterion for this is that the Richardson number be smaller than some arbitrary value. The Richardson number essentially expresses the ratio of the stability (the work that must be done against buoyancy forces if turbulence is to be maintained) to the destabilizing forces involved in wind shear. It appears probable that the criterion is satisfied occasionally, but the trails do not give the impression of being turbulent only in the regions of strong shear. Another concept is that gravity waves may at times be of sufficiently large amplitude so that potentially colder air is brought over potentially warmer air; that is, the temperature distribution may be distorted in such a way that the adiabatic lapse rate is exceeded in some regions so that thermal convection can proceed. Again, the trails do not have the appearance of being convectively unstable in narrow regions as the concept requires.

The presence of turbulence and its associated eddy transport cannot be doubted, even though there are difficulties in understanding the energy source and the physical generation mechanism. In addition to the visual evidence provided by vapor and meteor trails, the effect of the eddy transport on atmospheric structure is such that much can be deduced about the average rate of eddy transport from a study of atmospheric structure. On this basis it is recognized that the average eddy diffusion coefficient varies from about 10^4 cm^2/s just above the tropopause up to about 10^7 cm^2/s near 110 km (66 mi), above which it is exceeded by the molecular diffusion coefficient (above 110 km, the eddy coefficient probably decreases because of the severe damping of eddies by viscosity and heat conduction). *See* DIFFUSION.

The use of an eddy diffusion coefficient permits an evaluation of the transport properties of eddies without specifying their physical properties. The generally accepted concept is that the larger eddies break up into progressively smaller eddies, which finally dissipate their energy against viscous forces. The largest eddies are responsible for most of the transport. Examination of vapor trails suggests that the largest eddy size is of the order of 100 m (330 ft); with eddies of this size, eddy velocities near 10 m/s (20 knots) would be required to produce the observed mixing rates, and the lifetimes of the largest eddies would be a few tens of seconds. *See* METEOROLOGICAL ROCKET.

Large-scale circulation. There are prevailing patterns of atmospheric circulation in the upper atmosphere, but they are very different from those that occur in the lower atmosphere, which are associated with weather systems and have complicated structures resulting from growth of instabilities. The upper atmospheric large-scale wind systems are mainly diurnal in nature and global in scale.

The main heat source that is responsible for the upper atmospheric circulation is ultraviolet radiation from the Sun, radiation that is mainly absorbed at altitudes between 100 and 200 km (60 and 120 mi). The atmosphere is not a good infrared radiator in this altitude region, so the temperature rises rapidly with altitude, providing a temperature gradient of such a magnitude that molecular conduction transfers the absorbed heat downward to altitudes below 100 km (60 mi) where the atmosphere does have the capability of radiating the energy back to space. **Figure 2** shows some typical temperature distributions through the atmosphere. Above about 300 km (180 mi), the temperature becomes roughly constant with altitude because very little energy is absorbed there (the gas is exceedingly rarefied) and the thermal conductivity is good enough under these circumstances to virtually eliminate vertical temperature gradients. *See* SOLAR RADIATION.

The region of rising temperature above 80 km (48 mi) is known as the thermosphere. The exosphere is that region of the atmosphere that is so rarefied that for many purposes collisions between molecules can be neglected; it is roughly the region above 500 km (300 mi).

The maximum temperature in the thermosphere and exosphere on a diurnal basis occurs in late afternoon and near the subsolar latitude. The diurnal maximum is about 30% greater than the minimum, which occurs in the early morning hours and at about the same latitude as the subsolar point but in the opposite hemisphere. On average, the temperature of

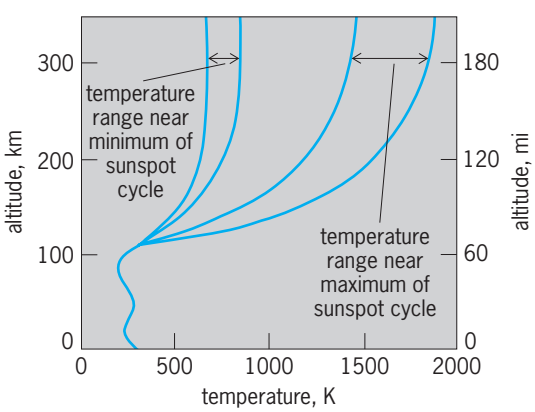

Fig. 2. Some typical temperature distributions through the atmosphere. °F = (K × 1.8) − 459.67.

the exosphere over the summer pole is about 300 K (540°F) higher than over the winter pole.

Where the atmosphere is warm, it is less dense, and the pressure fall with altitude is less rapid than where the atmosphere is cold. Thus the warmest region of the upper atmosphere is also the region of highest pressure. The resultant horizontal pressure gradients tend to make the atmosphere flow away from the warm region toward the cold. While Coriolis force comes into play to modify this tendency, in the upper atmosphere there are other effects that reduce the importance of the Coriolis force. These effects are viscosity and ion drag. Viscosity acts to reduce vertical gradients in the wind field. Ion drag occurs because the ions in the ionosphere are not free to drift across magnetic field lines; when the neutral atmosphere moves, collisions between neutral atmospheric particles and ions provide a drag force on the neutral atmospheric motion. These forces cause the global-scale wind patterns in the thermosphere to be relatively simple, mainly blowing from the warmest region toward the coldest. Wind patterns have been calculated taking all these forces (including Coriolis) into account, and the calculations are in reasonable agreement with the observations, although observations are difficult to make and are rather sparse. *See* CORIOLIS ACCELERATION.

The fact that the coldest region of the thermosphere occurs at mid or low latitude rather than over the winter polar region where sunlight is entirely lacking results from downward motion and compressional heating that occurs over the winter polar region—about 10^{-1} cm/s at 50 km (30 mi), 10 cm/s at 100 km (60 mi), and 10^2 cm/s above 150 km (90 mi). This compensates in considerable degree for the lack of solar heating there. A similar effect occurs in the vicinity of the temperature minimum; without it the minimum temperature would be much lower than it is.

An important, though sporadic, heat input in the upper atmosphere is auroral zone heating. At times of geomagnetic disturbances, important heating occurs there as a result of the inflow of energetic charged particles along the geomagnetic field lines and as a result of electric currents that cause Joule heating as they flow through the lower ionosphere. In especially strong events, the auroral zone heating completely disrupts the normal pattern of upper atmospheric circulation, and the flow in all but the lower thermosphere is from the poles to the Equator. On such occasions the auroral heating is a more important heat input to the upper atmosphere than is the absorption of solar ultraviolet radiation. More common are events where the normal circulation is only partially disrupted and both circulations are recognizable—one away from the normal diurnal temperature maximum and the other away from the poles. *See* AURORA; GEOMAGNETIC VARIATIONS; THERMOELECTRICITY.

At high latitudes, the ionosphere moves in response to electric fields imposed as a consequence of interactions between the Earth's magnetic field and the solar wind. The imposed electric field causes the ionosphere to drift in a generally antisunward direction over the polar caps (regions at higher latitudes than the auroral zone, or magnetic latitudes greater than about 68°), with a return circulation (that is, generally sunward in direction) just outside the polar caps. Although the ions are much less numerous than the neutral particles and they transfer their momentum to the neutral particles when they collide, they quickly regain their drift velocity through action of the electric field, and each ion can therefore transfer momentum many times to the neutral particles, making up for their lesser numbers by repeated collisions, each transferring momentum. Because of the persistence of ionospheric drift caused by electric fields, the neutral particles gradually pick up a drift velocity similar to that of the ions. Thus a velocity pattern similar to the polar-region ionospheric drift pattern becomes superposed on the large-scale circulation of the neutral thermosphere. Over the polar caps, the two motions are largely in the same direction, but the ionospheric circulation imposes a reverse circulation on the neutral atmosphere in a relatively narrow region just outside the polar caps. *See* ATMOSPHERIC GENERAL CIRCULATION; SOLAR WIND; WIND.

Composition. Ultraviolet radiation of suitable wavelengths can photodissociate atmospheric molecules—something of great importance in the upper atmosphere. It is even important in the stratosphere, where ozone is formed as a result of absorption by molecular oxygen of ultraviolet radiation, the important wavelengths being below 242 nanometers. Dozens of chemical reactions are actually involved in ozone chemistry, but in highly simplified terms the essential mechanism is shown by reactions (2), where M is a third body that is

$$\begin{aligned} &O_2 + h\nu \rightarrow O + O \\ &O + O_2 + M \rightarrow O_3 + M \\ &O + O_3 \rightarrow 2O_2 \end{aligned} \tag{2}$$

required for the reaction to proceed, even though it is not changed chemically by the reaction. An equilibrium between formation and destruction of ozone is established. This basic set of reactions is important in the altitude range 15–70 km (9–42 mi); it indicates in a highly simplified way how the ozone layer is formed. *See* STRATOSPHERIC OZONE; STRATOSPHERE.

At higher altitudes where ultraviolet radiation is present that cannot penetrate as deeply as the ozone layer, the importance of oxygen photodissociation is greater. As atomic oxygen becomes more plentiful and molecular oxygen less plentiful, the main loss mechanism for atomic oxygen is reaction (3).

$$O + O + M \rightarrow O_2 + M \tag{3}$$

At altitudes above 100 km (60 mi), recombination becomes very slow because of low concentrations and the resultant smaller probability of collisions. As a result, the loss of atomic oxygen occurs mainly by diffusion downward to the region below 100 km (60 mi) where recombination proceeds more rapidly.

Above 100 km (60 mi), the distribution of atmospheric gases becomes increasingly controlled by diffusion of the various constituents among themselves; the end result of the diffusion process in the gravitational field is the diffusive-equilibrium distribution of atmospheric constituents in which each constituent is distributed as if the others were not present. Atomic oxygen, being much lighter than molecular oxygen and nitrogen, falls off with altitude much more slowly than the molecular species. Above about 200 km (120 mi), atomic oxygen predominates over molecular oxygen and nitrogen. A similar situation with atomic nitrogen does not occur, because it photodissociates much more slowly; it requires much shorter-wave-length ultraviolet radiation (<80 nm), which is much less intense in solar radiation.

The photodissociation of oxygen above 100 km (60 mi), the diffusion downward of the resultant atomic oxygen to altitudes below 100 km (60 mi), and its recombination there into molecular oxygen provide a downward transport of chemical energy. Energy from ultraviolet radiation is converted partly into chemical form (by production of atomic oxygen) and partly into heat at the altitude where it is absorbed (the excess energy beyond that required for photodissociation being released as kinetic energy). The chemical energy is released below 100 km (60 mi) when recombination into molecular oxygen occurs.

The large-scale circulation of the thermosphere produces a rather surprising effect on atmospheric composition. In the lower part of the altitude region where atmospheric constituents are in diffusive equilibrium, horizontal winds are more effective in transporting lighter constituents than heavier ones. Because of the lesser rate of decrease in concentration with altitude for the lighter constituents, an effectively thicker layer is transported by the wind system. As a consequence, the large-scale circulation of the thermosphere transports relatively more atomic oxygen than molecular oxygen and nitrogen from the summer to the winter hemisphere and toward the winter polar region. The effect is so pronounced that it results in atomic oxygen enrichment of the winter thermosphere, in spite of the fact that the rate of photoproduction is very low in that hemisphere. This atomic oxygen enrichment of the winter thermosphere has an important effect on ionospheric chemistry, and the winter ionosphere is more highly ionized as a result of the enrichment. The enrichment effect is even stronger is the case of helium, because it is lighter than atomic oxygen. The enrichment in the case of helium is referred to as the winter helium bulge. *See* AERONOMY.

Planetary upper atmospheres. Data from other planets, most notably Venus and Mars, have considerably expanded understanding of upper atmospheric dynamics. Venus has a surface atmospheric pressure almost a hundred times greater than Earth, whereas Mars has a surface pressure about one-hundredth that of Earth. Surprisingly, the upper atmospheres of these two planets are much more similar to one another than to the upper atmosphere of Earth. This similarity is due to the absence of intrinsic magnetic fields large enough to shield the atmospheres from the solar wind (Mars may have an intrinsic field just on the borderline of providing some shielding effect), and to the very similar atmospheric compositions on the two planets (about 96% carbon dioxide in both cases).

The most detailed data that have been obtained are for Venus, mostly from the *Pioneer Venus* and *Venera 11* and *12* missions. These show that the solar wind confines the ionosphere on the sunward side of the planet, the interplanetary magnetic field providing the mechanism for the interaction. Since the solar wind is quite variable in intensity, the altitude at which the interface occurs is also variable, ranging from 200 to 1000 km (120 to 600 mi) but frequently in the vicinity of 400 km (240 mi). The daytime ionosphere tends to escape from the confinement region by flowing at high velocity (several kilometers per second) across the entire terminator from the day-side hemisphere to the night-side hemisphere, providing an important source for the maintenance of the night-time ionosphere.

The thermosphere of Venus is very cold compared to that of Earth, about 250 to 300 K (−10 to 80°F) on the day side and 100 K (−280°F) on the night side. The neutral thermospheric circulation is undoubtedly a flow from the day side to the night side, as both the horizontal neutral pressure gradient and momentum transfer from the rapidly flowing ionosphere act in this sense.

Measurements of the vertical distributions of atmospheric constituents can be interpreted in terms of the average rate of turbulent mixing in the lower thermosphere, and the result is a surprisingly large rate, much larger than on Earth. The rate is so high that there is some reason to doubt the correctness of the interpretation.

The upper atmosphere of Mars is very similar to that of Venus. The great difference in surface pressure on the two planets does not have any significance with regard to the upper atmosphere, and other parameters are rather similar. It is plausible to assume that internal gravity-wave activity on Mars might be much greater than on Venus because surface winds are greater; this is a direct result of the low density of the Martian atmosphere at the surface and its consequent low heat capacity. Winds blowing over mountains are likely sources of internal gravity waves, and the surface winds on Mars are known to be strong, at least at times. On Venus, on the other hand, surface winds are weak, as the high-density atmosphere can transfer large amounts of heat with only light winds. *See* MARS; VENUS.

Francis S. Johnson

Bibliography. A. Brekke, *Physics of the Upper Polar Atmosphere*, 1995; S. Chapman and R. S. Lindzen, *Atmospheric Tides, Thermal and Gravitational*, 1970; T. E. Cravens and A. F. Nagy, Aeronomy of the inner planets, *Rev. Geophys. Space Phys.*, 21:263–273, 1983; W. Dieminger, C. K. Hartmann, and R. Leitinger, *The Upper Atmosphere: Data Analysis*

and Interpretation, 1994; M. Ghil and S. Childress, *Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory, and Climate Dynamics*, 1987; C. Riege, *Fundamentals of Atmospheric Dynamics and Thermodynamics*, 1990; R. G. Roble, R. E. Dickinson, and E. C. Ridley, Global circulation and temperature structure of the thermosphere with high-latitude plasma convection, *J. Geophys. Res.*, 87:1599–1614, 1982.

Upsilon particles

A family of elementary particles whose first three members were discovered in 1977. The upsilon mesons, Υ, are the heaviest known vector mesons, with masses greater than 10 times that of the proton. They are bound states of a heavy quark and its antiquark. The quarks which bind to form the upsilons carry a new quantum number called beauty or bottomness, and they are called b-quarks or b. The mass of the b-quark is around 5 GeV. The anti-b-quark or $\overline{b}$ carries antibeauty, and therefore the upsilons carry no beauty and are often called hidden beauty states. Direct proof of the existence of the b-quark was obtained by observing the existence of B mesons which consist of a b-quark bound to a lighter quark. Twelve $b\overline{b}$-mesons have also been observed thus far. *See* QUARKS.

Heavy quarks. In addition to electric charge, all quarks carry a new nonscalar "charge," called color, which results in forces much stronger than the electric Coulomb force. There is, however, a major difference between the electric and the color forces. While the Coulomb force becomes weaker at large distances, color forces are believed to become stronger at large distances, resulting in the confinement and thus unobservability of free quarks. The theory of the color forces is known as quantum chromodynamics (QCD). *See* COLOR (QUANTUM MECHANICS); QUANTUM CHROMODYNAMICS.

A bound system of a heavy quark and its antiquark ($q\overline{q}$ or quarkonium) has very small dimensions. As a result, the color force becomes relatively weak and the quarks move with low velocity. The quark-antiquark system can therefore be described by a Schrödinger equation if the potential between the quarks is known. In a pure coulombic potential, a $q\overline{q}$ system becomes very similar to an hydrogen atom. A better analogy, however, is the case of positronium, an atomlike system consisting of an electron and a positron bound together by their mutual electric attraction. Just as for the case of positronium, which can decay into photons, the quanta of the electromagnetic field, so can the quarkonium state annihilate into gluons, the quanta of the field coupled to the color charge. According to older ideas, heavy quarkonium states should annihilate in times of the order of 10^{-24} s, while quantum chromodynamics suggests much longer lifetimes. *See* ATOMIC STRUCTURE AND SPECTRA; GLUONS; NONRELATIVISTIC QUANTUM THEORY; POSITRONIUM; QUANTUM MECHANICS.

In 1974 the first heavy vector meson was discovered. The mass of this new particle, called J/ψ, is around a 3 GeV and its width of the order of 60 keV, corresponding to a lifetime of about 10^{-19} s. The discovery of the J/ψ (a $c\overline{c}$ bound system, where c stands for the charm quark of mass around 1.5 GeV), was a triumph for quantum chromodynamics. *See* CHARM; J/PSI PARTICLE.

The upsilon mesons provide even better testing grounds for quantum chromodynamics. Because they contain the much heavier b quarks, the color force is slightly weaker than for the J/ψ case. Thus the upsilon annihilation rate was expected to be slower or, equivalently, its width to be narrower. This is indeed so: the width of the upsilon is only 40 keV while its mass is 9.4 GeV.

Potential models. Quantum chromodynamics has not yet been able to derive the exact form of the potential. It can, however, suggest its form. In particular, the strength of the potential is supposed to increase linearly at large distances and to obey a $1/r$ law at short distances, like the Coulomb potential. On this basis, the properties of the ψ and Υ families have been thoroughly calculated. The amount of experimental information is now so large that in fact the process can be reversed, and from the data the form of the potential has been extracted to great accuracy.

Family of states. Just as for the case of the hydrogen atom or of positronium, the $b\overline{b}$ pair can bind in many different ways, properly classified by the quantum numbers of the various states. The bound pair can have orbital angular momentum of 0, 1, 2, and so forth, the corresponding states being referred to as S-wave, P-wave, D-wave, and so forth. Each quark, in addition, has a spin angular momentum of $^1/_2$. The two spins can combine to give a total spin of 0 or 1. Spin-0 states, as in all quantum systems, are called singlet states, and spin-1 states, triplet states. Orbital and spin angular momentum combine to give the total angular momentum of the state. Bound states are usually classified by the principal quantum number n, the total angular momentum J, and the total spin S (see **table** and **illus.**).

S-wave states. While for a pure coulombic potential there exists an infinite number of bound states, the quarkonium potential requires that there should

Masses of the known members of the upsilon family

Meson name	Mass, GeV/c^2*
Υ	9.46037 ± 0.00021
Υ'	10.02330 ± 0.00031
Υ''	10.3553 ± 0.0005
Υ'''	10.5800 ± 0.0035
$\Upsilon(5S)$	10.865 ± 0.008
$\Upsilon(6S)$	11.019 ± 0.008
$\chi_{b,J=2}$	9.9132 ± 0.0006
$\chi_{b,J=1}$	9.8919 ± 0.0007
$\chi_{b,J=0}$	9.8598 ± 0.0013
$\chi'_{b,J=2}$	10.2685 ± 0.0004
$\chi'_{b,J=1}$	10.2552 ± 0.0005
$\chi'_{b,J=0}$	10.2321 ± 0.0006

*c = speed of light.

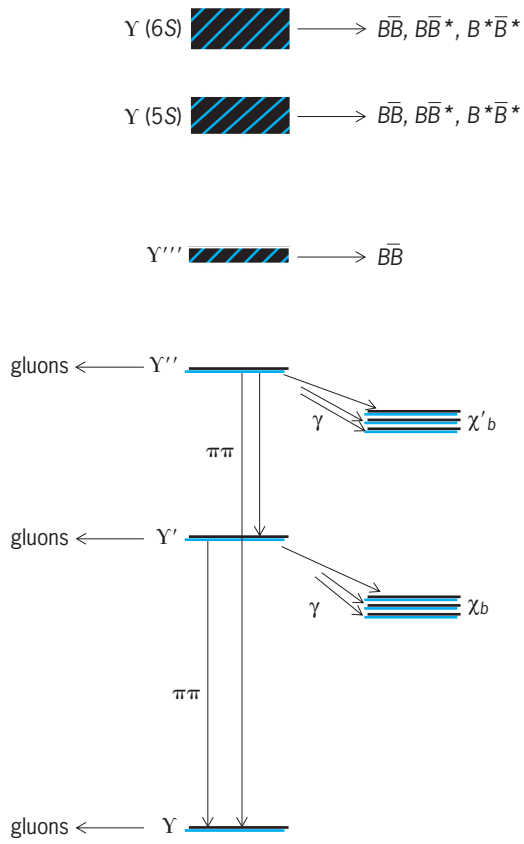

Known $b\bar{b}$ levels and their intertransitions.

be only three *S*-wave, spin-triplet, bound $b\bar{b}$ states. These are, in fact, the first three upsilon mesons discovered in 1977. The Υ mass is now known to about 20 parts per million, an accuracy close to that for the electron mass.

The fourth upsilon, Υ′′′, was discovered in 1980. In contrast to the first three upsilons, which have widths ranging from 40 to 20 keV, the width of the Υ′′′ is about 20 MeV, almost a thousand times greater, indicating that it is slightly heavier than two *B* mesons into which it decays. Direct evidence for the decay $\Upsilon''' \rightarrow B + \overline{B}$ was obtained by the CESR experiments, by observing the beta decay of the *B* mesons.

The fifth and sixth radial excitations of the triplet *S*-wave $b\bar{b}$ system were observed in 1984. These mesons are usually called Υ(5*S*) and Υ(6*S*).

P-wave states. Six spin-triplet, *P*-wave bound $b\bar{b}$ states, called χ_b, are expected to exist. These states cannot be produced in e^+e^- annihilations because they have positive parity and charge conjugation. They can, however, be produced in a two-step process given by expressions (1) and (2). In this way

$$e^+e^- \rightarrow \Upsilon \tag{1}$$

$$\Upsilon \rightarrow \chi_b + \gamma \tag{2}$$

the lowest spin-triplet *P*-wave state can be reached, after emission of a photon, from the Υ′, and the second *P*-wave state can be reached via decays of the Υ″. Because the triplet *P*-wave states have total spin of 1 and also orbital angular momentum 1, the total angular momentum of these states can be, according to the rules of quantum mechanics, 0 or 1 or 2. The spin-orbit interaction shifts the mass of states with different total angular momentum, resulting in six different states: $\chi_{b,J=2}$, $\chi_{b,1}$, $\chi_{b,0}$, $\chi'_{b,2}$, $\chi'_{b,1}$, and $\chi'_{b,0}$. The three χ'_b were discovered in 1982 by observing the decays $\Upsilon'' \rightarrow \chi + \gamma$, $\chi \rightarrow \Upsilon' + \gamma$, and $\chi \rightarrow \Upsilon + \gamma$, and the three χ_b were discovered in the same manner in 1983. Later the χ_b states were also observed by other experiments.

In addition to electromagnetic transitions between Υ's and χ's, transitions between Υ's, with emission of two pions, have also been observed. *See* ELEMENTARY PARTICLE; MESON. Paolo Franzini

Bibliography. P. Franzini and J. Lee-Franzini, Upsilon resonances, *Annu. Rev. Nucl. Part. Phys.*, 33:1–29, 1983; G. Kane, *Modern Elementary Particle Physics*: *The Fundamental Particles and Forces*, 1993; L. M. Lederman, The upsilon particle, *Sci. Amer.*, 239(4):72–80, October 1978; Review of particle properties, *Phys. Rev. D*, 50:1173–1825, 1994.

Upwelling

The phenomenon or process involving the ascending motion of water in the ocean. Vertical motions are an integral part of ocean circulation, but they are a thousand to a million times smaller than the horizontal currents. Vertical motions are inhibited by the density stratification of the ocean because with increasing depth, as the temperature decreases, the density increases, and energy must be expended to displace water vertically. The ocean is also stratified in other properties; for example, nutrient concentration generally increases with depth. Thus even weak vertical flow may cause a significant effect by advecting nutrients to a new level. *See* GEOPHYSICAL FLUID DYNAMICS.

There are two important upwelling processes. One is the slow upwelling of cold abyssal water, occurring over large areas of the world ocean, to compensate for the formation and sinking of this deep water in limited polar regions. The other is the upwelling of subsurface water into the euphotic (sunlit) zone to compensate for a horizontal divergence of the flow in the surface layer, usually caused by winds.

The upwelling of abyssal water is very slow (meters per year) and is inferred from numerical models, the distribution of properties, and the requirement that the circulation in the deep ocean balance the sinking of cooled surface water in polar regions. The residence time of the abyssal water, that is, the time from formation and sinking in polar regions to upwelling after flowing at depth through the world ocean, is hundreds of years. It had been thought that upwelling of abyssal water resulted from the downward mixing of heat across the thermocline over vast areas of the low- and mid-latitude oceans, buoying up the old abyssal waters so that they could be replaced

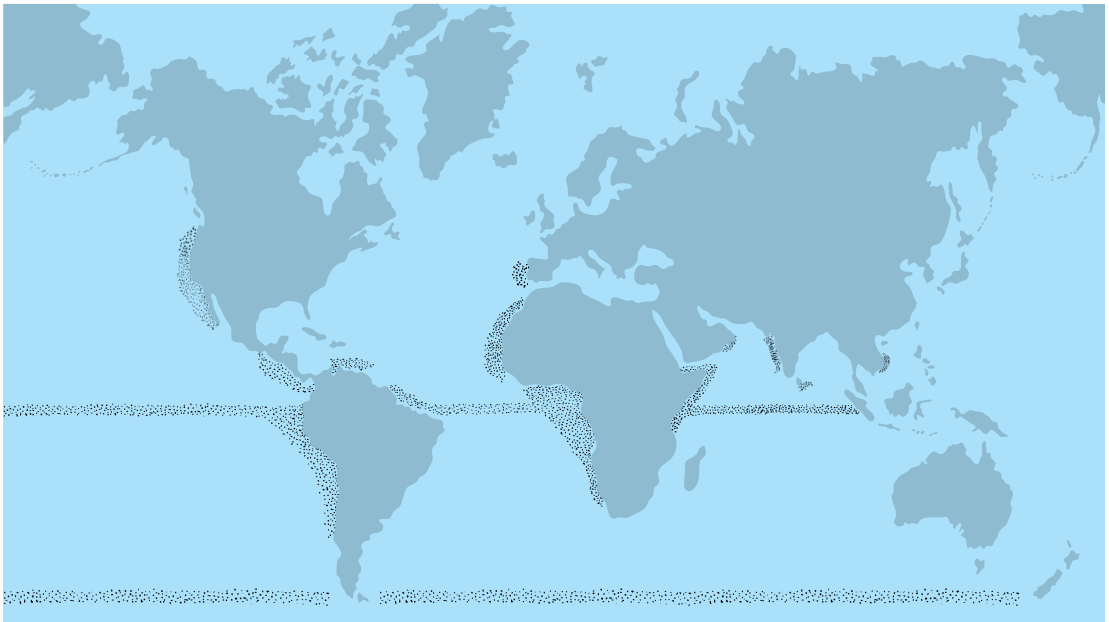

Fig. 1. Regions of major upwelling are stippled.

by new deep water flowing from the polar regions. It is now apparent that mixing in the deep ocean is insufficient to account for the amount of abyssal upwelling, and that surface forcing by the winds may be responsible.

Wind-induced divergence in the surface currents drives the more rapid (meters per day) upwelling of water from a hundred meters' depth into the euphotic zone. This upwelling can have an immediate effect on the ecology of the ocean and the world's economy. If the mechanism causing the divergence continues, cold and nutrient-rich subsurface water will continue to upwell to the surface layer and spread out horizontally. Eventually, sinking or downwelling will occur to balance this upwelling, but this may occur far from the area of upwelling. Regions of persistent upwelling are characterized by relatively cool surface water that is rich in nutrients compared with typical oceanic surface water. The upwelling flux of nutrients increases the region's fertility. Unusually high biological productivity is associated with these upwelling regions, which are often obvious in satellite images of sea surface temperature or chlorophyll distributions. *See* OCEAN CIRCULATION.

Occurrence. Upwelling may occur anywhere in the ocean where a mechanism exists to cause a divergence in the surface layer flow. It occurs consistently in both the equatorial zone and the Antarctic circumpolar ocean but is probably most conspicuous along the western coasts of the continents (**Fig. 1**). At low and middle latitudes in both hemispheres, the surface waters within a hundred kilometers of the eastern boundaries of the ocean (the west coasts) are cooler (by several degrees Celsius) and more productive biologically (by a factor of 10 or more) than water farther seaward at the same latitudes as a result of upwelling near the coast.

Wind-induced upwelling was explained qualitatively by V. W. Ekman (1905), who studied the influence of the wind on the surface currents. He showed that the effects of the Earth's rotation (the Coriolis force) and frictional forces (due to turbulence) resulted in the net movement of the surface layer being directed 90° to the right of the wind in the Northern Hemisphere and 90° to the left in the Southern Hemisphere. This wind-induced transport, called Ekman transport, is confined to a layer only several tens of meters deep. Quantitatively, the Ekman transport can be calculated from the wind stress (the frictional drag force exerted by the wind on the sea surface) divided by the Coriolis parameter (twice the Earth's rotation rate multiplied by the sine of the latitude). *See* CORIOLIS ACCELERATION; WIND; WIND STRESS.

Because of semipermanent midocean atmosphere high-pressure systems, the prevailing winds are generally equatorward along the western coasts of the continents. Thus the Ekman transport is directed away from the coasts, and this one-sided divergence necessitates an upwelling of subsurface water into the surface layer near the coast (**Fig. 2**). The upwelled water comes from depths of 50 to a few hundred meters and reaches the surface layer in a relatively narrow band extending only a few tens of kilometers from the coast. The intensity of coastal upwelling must balance the Ekman transport directed offshore since the latter must be replaced by upwelling near the coast. The upwelling index used by the U.S. National Marine Fisheries Service is exactly this. Typical Ekman transport values off Oregon and California during summer are about 1000 kg/s per meter of coastline. If this water upwells through a band extending about 10 km from the coast, the vertical velocity would be only about 0.0001 m/s (about 10 m per day). This compares with typical horizontal currents of 0.1–1 m/s (about 10–100 km per day).

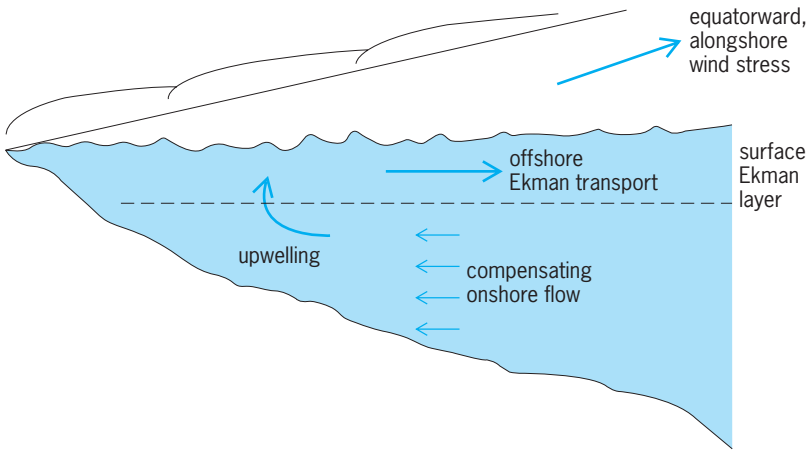

Fig. 2. Schematic diagram of coastal upwelling.

Upwelling at the Equator also results from the wind-induced divergence of the surface layer, analogous to that in coastal upwelling. The trade winds in the equatorial regions blow westward, and thus the surface water is transported away from the Equator on either side of it because the Ekman transport is to the right of the wind in the Northern Hemisphere and to the left in the Southern Hemisphere. The divergence of the Ekman transport near the Equator results in the upwelling of cool nutrient-rich water.

There is also strong upwelling in the Arabian Sea (northwestern Indian Ocean) off the east coasts of Somalia, Saudi Arabia, and Oman during May to September when the monsoon winds blow from the southwest, nearly parallel to the coast, resulting in offshore Ekman transport. The monsoon wind increases to maximum offshore, causing a further divergence in the surface flow, augmenting the coastal upwelling. The sea surface temperature at the Arabian coast may drop below 20°C, while temperatures 1000 km seaward are nearly 30°C. During the winter the winds blow in the opposite direction and no upwelling occurs. *See* MONSOON METEOROLOGY.

Very near the Antarctic continent the winds blow westward, but farther to the north the winds blow eastward. Between these winds, there is a region of diverging flow in the surface layer, which causes nutrient-rich water to upwell. The eastward winds blow unimpeded around the Antarctic continent and, with the Ekman transport directed to the left of the wind in the Southern Hemisphere, the Ekman transport diverges and causes upwelling. Large blooms of plankton occur, providing the food for krill that are eaten by whales, which is why the ocean around Antarctica was once a major whaling ground. The water, although dense, is weakly stratified, and thus water may be upwelled from great depth using relatively little energy. This wind-induced upwelling may help drive the abyssal ocean's overturning circulation. *See* ANTARCTIC OCEAN.

Effects. As a result of coastal upwelling, the water near the coast is denser, and this sets up an alongshore (equatorward) geostrophic current off the west coasts. The boundary between upwelled water and the offshore oceanic surface water often has a region of sharp temperature and density contrast (a front). The alongshore geostrophic current in this region may be very strong (a coastal jet) because of the sharp density gradient. The trajectory of an upwelling parcel of water is relatively flat; the parcel moves horizontally with the currents along the coast as it gradually rises toward the surface. The upwelled water spreads offshore and equatorward, and the area affected by coastal upwelling may eventually extend 100 km or more offshore. The flow pattern is not simple, and the boundary between upwelled and oceanic water may meander, with filaments of cold water being apparent in satellite images of the sea surface temperature along the west coasts of the continents. Along the west coasts, low stratus clouds or fog is common due to the cooling of air near the coast as it comes into contact with the cool upwelled water. The cool ocean temperatures along the California coast contrast considerably with the warmer (by as much as 10°C) temperatures along the eastern United States seaboard at the same latitudes. The early explorers to the west coasts of Africa and South America were surprised by the anomalously cool surface waters near the coast, and the Conquistadores reportedly made good use of the upwelling by hanging flagons of wine in the sea to cool while in Peruvian ports.

Biologically, coastal upwelling regions are economically important. Their primary productivity is high because nutrients are supplied to the euphotic zone by the upwelling process. This provides a base for the short efficient food chains that characterize coastal upwelling regions. It is estimated that half of the fish caught commercially are from coastal upwelling regions, which constitute only 0.1% of the ocean's area. The coast of Peru is legendary for the large population of guano birds and the anchoveta fishery on which the birds feed. Because of the relative narrowness of the coastal upwelling regions, high fish concentrations can be efficiently harvested, making the species vulnerable to overfishing. The short food chain and low species diversity make the coastal upwelling ecosystem especially vulnerable to natural environmental variations. Evidence for climate variations can be found in the marine sediments, containing the undissolved remains of high productivity. *See* BIOLOGICAL PRODUCTIVITY; MARINE ECOLOGY; MARINE FISHERIES; MARINE SEDIMENTS; SEAWATER FERTILITY.

Variability. Upwelling varies with changing winds, with latitude, and, like the weather, with the changing seasons. In some regions, there is a seasonal reversal of the winds as atmospheric pressure systems shift. An example is the west coast of North American north of about 40° latitude (northern California), where the winds from late fall until early spring are predominantly from the southwest, resulting in onshore Ekman transport and the suppression of upwelling. Off southern California, Mexico, and Peru, the upwelling continues throughout the year, but the intensity varies with latitude and season. However,

off Peru, coastal upwelling is affected by an occasional catastrophic variation—El Niño.

El Niño originally referred to the occasional appearance of excessively warm water along the coast of Peru and was once thought to be due to weakened coastal upwelling. It is now known that El Niño is a larger-scale phenomenon originating in the western equatorial Pacific that, among other effects, causes the warm and nutrient-depleted surface layer to deepen greatly off Ecuador and Peru. Local coastal upwelling continues under undiminished coastal winds during El Niño, but cool nutrient-rich water is out of reach of the upwelling process and only warm nutrient-depleted water flows into the euphotic zone, collapsing its usually high productivity. At the same time, the winds in the equatorial Pacific decrease and equatorial upwelling ceases or diminishes greatly, and the equatorial surface waters warm. The resulting change in ocean-atmosphere interaction in the equatorial region affects the weather worldwide. *See* EL NIÑO; MARITIME METEOROLOGY; SEAWATER; TROPICAL METEOROLOGY. Robert L. Smith

Bibliography. A. E. Gargett, Physical process and the maintenance of nutrient-rich euphotic zones, *Limnol. Oceanog.*, 36:1527–1545, 1991; S. J. Lentz, The surface boundary layer in coastal upwelling regions, *J. Phys. Oceanog.*, 22:1517–1539, 1992; W. J. Schmitz, Jr., On the interbasin-scale thermohaline circulation, *Rev. Geophys.*, 33:151–173, 1995; C. P. Summerhayes et al. (eds.), *Upwelling in the Ocean: Modern Processes and Ancient Records*, Wiley, 1995; J. Robert Toggweiler, The ocean's overturning circulation, *Phys. Today*, pp. 45–50, November 1994.

Uranium

A chemical element, symbol U, atomic number 92, atomic weight 238.03. The melting point is 1132°C (2070°F) and the boiling point is 3818°C (6904°F). Uranium is one of the actinide series. *See* ACTINIDE ELEMENTS; PERIODIC TABLE.

Uranium in nature is a mixture of three isotopes: ^{234}U, ^{235}U, and ^{238}U. Uranium is believed to be concentrated largely in the Earth's crust, where the average concentration is 4 parts\break per million (ppm).

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

The total uranium content of the Earth's crust to a depth of 15 mi (25 km) is calculated to be 2.2×10^{17} lb (10^{17} kg); the oceans may contain 2.2×10^{13} lb (10^{13} kg) of uranium. Several hundred uranium-containing minerals have been identified, but only a few are of commercial interest. *See* RADIOACTIVE MINERALS; URANINITE.

Because of the great importance of the fissile isotope ^{235}U, rather sophisticated industrial methods for its separation from the natural isotope mixture have been devised. The gaseous diffusion process, which in the United States is operated in three large plants (at Oak Ridge, Tennessee; Paducah, Kentucky; and Portsmouth, Ohio) has been the established industrial process. Other processes applied to the separation of uranium include the centrifuge process, in which gaseous uranium hexafluoride is separated in centrifuge cascades, the liquid thermal diffusion process, the separation nozzle, and laser excitation. *See* ISOTOPE (STABLE) SEPARATION.

Uranium is a very dense, strongly electropositive, reactive metal; it is ductile and malleable, but a poor conductor of electricity. Many uranium alloys are of great interest in nuclear technology because the pure metal is chemically active and anisotropic and has poor mechanical properties. However, cylindrical rods of pure uranium coated with silicon and canned in aluminum tubes (slugs) are used in production reactors. Uranium alloys can also be useful in diluting enriched uranium for reactors and in providing liquid fuels. Uranium depleted of the fissile isotope ^{235}U has been used in shielded containers for storage and transport of radioactive materials. *See* NUCLEAR FUELS; NUCLEAR REACTOR.

Uranium reacts with nearly all nonmetallic elements and their binary compounds. Uranium dissolves in hydrochloric acid and nitric acid, but nonoxidizing acids, such as sulfuric, phosphoric, or hydrofluoric acid, react very slowly. Uranium metal is inert to alkalies, but addition of peroxide causes formation of water-soluble peruranates. *See* URANIUM METALLURGY.

Uranium reacts reversibly with hydrogen to form UH_3 at 250°C (482°F). Correspondingly, the hydrogen isotopes form uranium deuteride, UD_3, and uranium tritide, UT_3. The uranium-oxygen system is extremely complicated. Uranium monoxide, UO, is a gaseous species which is not stable below 1800°C (3270°F). In the range UO_2 to UO_3, a large number of phases exist. The uranium halides constitute an important group of compounds. Uranium tetrafluoride is an intermediate in the preparation of the metal and the hexafluoride. Uranium hexafluoride, which is the most volatile uranium compound, is used in the isotope separation of ^{235}U and ^{238}U. The halides react with oxygen at elevated temperatures to form uranyl compounds and ultimately U_3O_8. Fritz Weigel

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; J. J. Katz et al. (eds.), *The Chemistry of the Actinide and Transactinide Elements*, 3d ed., 2006; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004.

Uranium metallurgy

The processing treatments for the production of uranium concentrates and the recovery of pure uranium compounds, as well as the conversion chemistry to produce uranium metal and the processes employed for preparing uranium alloys.

Ore processing. The procedures to recover uranium from its ores are numerous, because of the great variety in the nature of uranium minerals and associated materials and the wide range of concentration in the naturally occurring ores. Recovery of uranium requires chemical processing; however, preliminary treatment of the ore may involve a roasting operation, a physical or chemical concentration step, or a combination of these.

Roasting can bring about significant chemical change on its own, but moreover it can improve the filtering and settling characteristics of an ore for subsequent processing. Only a limited number of deposits are sufficiently rich in uranium-bearing minerals and have appropriate physical characteristics to make a preliminary physical concentration of the uranium a feasible operation. Chemical concentration is most common as a preliminary treatment.

In general, one of two leaching treatments—acid leaching and carbonate leaching—is used as the initial step in chemical concentration. The choice depends on the nature of the ore, which largely determines the efficiency and the cost of the process employed.

Acid leaching is carried out universally with sulfuric acid. The finely ground ore is treated with the dilute acid, and usually an oxidant is added to make sure that all the uranium goes into solution as the uranyl ion. When an ore is carbonate-leached, the action of sodium carbonate on the uranium forms water-soluble $[UO_2(CO_3)_3]^{4-}$ ion. This solution is stabilized by adding bicarbonate and by keeping the uranium in the oxidized state. From either the acid or carbonate leaching process, the uranium can be recovered as ammonium uranate or sodium uranate (yellow cake) by precipitation with ammonia or with sodium hydroxide. *See* LEACHING.

Recovery of pure uranium. The concentrate, whether obtained by chemical or physical means, is treated chemically to give a uranyl nitrate solution that can be further purified by solvent extraction. The impurities remain in the aqueous phase, while the uranium is extracted into the organic phase. Formerly, ether or hexone was used as an extractant; presently, tributyl phosphate (TBP) is favored. From the organic phase, the uranium is stripped by means of water, and the high-purity uranium can be recovered as nitrate crystals or by precipitation from the solution. The nitrate serves as the starting material for other compounds, such as the oxides. In large-scale processing, the nitrate is decomposed thermally to give uranium trioxide, UO_3, which is subsequently reduced with hydrogen to form uranium dioxide, UO_2. Uranium tetrafluoride, UF_4 (green salt), is prepared by treating UO_2 with hydrogen fluoride (HF) gas. *See* SOLVENT EXTRACTION.

Metal reduction. Uranium metal can be obtained from its halides by fused-salt electrolysis or by reduction with more reactive metals. The reaction of UO_2 with calcium yields metal of fair quality.

The electrolysis of UF_4 in a fused-salt bath of sodium chloride and calcium chloride has been used to prepare tons of high-purity metal. The metal is deposited as small granules on a molybdenum cathode. In the calcium reduction of UO_2, the ingredients of the charge are mixed and placed in an inert atmosphere and are heated to about 1830°F (1000°C) to form finely divided uranium and calcium oxide. Calcium hydride can be employed instead of calcium in this process. The metal from either electrolysis or the calcium-UO_2 process is leached, pressed, and sintered or melted to give solid uranium. *See* ELECTROLYSIS.

The largest tonnages of good-quality uranium have been produced by metallothermic reduction of finely divided UF_4 with calcium or magnesium in steel bombs lined with fused dolomitic oxide (Ames process). The charge, consisting of an intimate mixture of UF_4 with the reductant metal in granular form together with a suitable booster (usually calcium plus iodine), is placed into the reduction bomb, the lid is bolted down, and the bomb is heated to ignition temperature. In the case of calcium, the reaction heat is such that the charge can be ignited at room temperature, and the products are sufficiently molten to allow separation of metal and slag in the reduction vessel. For magnesium, the whole bomb has to be preheated in a heat-soaking pit. The heat generated in the reduction reaction is sufficient to melt both metal and slag, and the molten metal collects by gravity as a pool under the molten slag, where it

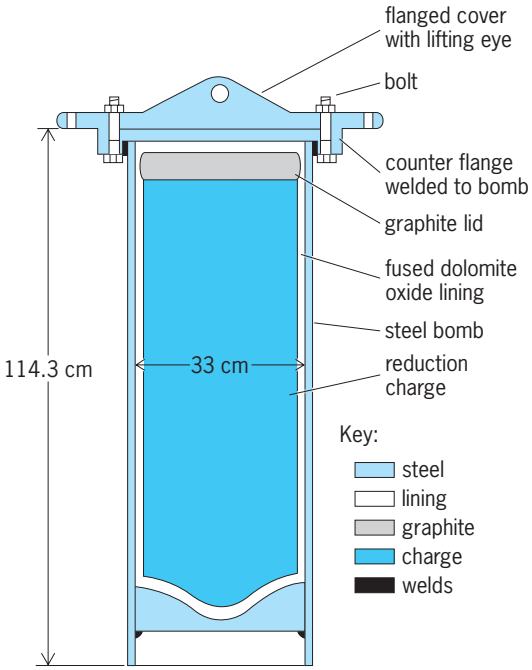

Fig. 1. Standard Derby bomb for reduction of UF_4 with magnesium by the Ames process. Capacity is 317.9 lb (144.2 kg) of uranium. 1 cm = 0.4 in.

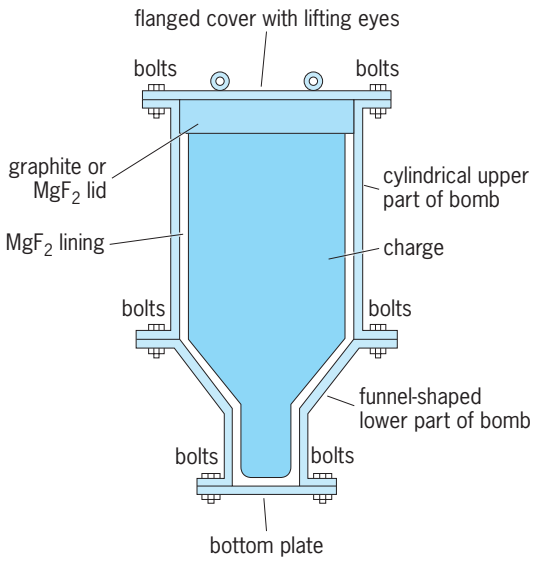

Fig. 2. Dingot bomb for large charges.

remains during solidification. The shape of the metal ingot (also referred to as a biscuit) depends on the shape of the reduction bomb. A flat, pancake-shaped ingot, which has to be remelted for reshaping, is obtained in the standard bomb (**Fig. 1**), while reduction in a Dingot bomb (**Fig. 2**) yields a so-called direct-reduction ingot (dingot), which is shaped in such a way that it may undergo further metallurgical treatment without remelting. The Dingot process is particularly suited for large-scale reductions (110 lb or 50 kg to 1.7 tons or 1.5 metric tons). The metal obtained in the metallothermic reduction process is quite pure. Metal obtained by reduction in a standard bomb has to be recast by vacuum melting in a graphite crucible and casting into graphite molds. The metal can be fabricated by conventional means with due consideration for its chemical reactivity and allotropic transformations. The low-ductility modification of uranium is usually avoided in fabrication processes. In the metallurgy of ^{233}U and ^{235}U, strict weight and configuration limitations must be imposed to avoid criticality.

Alloys. Uranium alloys are prepared by fusing the components together. All procedures following conventional metallurgical techniques, however, may have to be carried out inside inert-gas glove boxes because many alloys are attacked by oxygen or moisture. *See* NUCLEAR FUELS; NUCLEAR FUELS REPROCESSING; URANIUM. F. Weigel

Bibliography. R. C. Merritt, *Extractive Metallurgy of Uranium*, 1971, reprint 1979; Organization for Economic Cooperation and Development, *Uranium Resources: Production and Demand*, 1988.

Uranus

The first planet to be discovered with the telescope and the seventh in the order of distance from the Sun. It was found accidentally by W. Herschel in England on March 13, 1781. Herschel's home-made telescope was good enough to show that this object was not starlike; Uranus appeared as a fuzzy patch of light, not a point. At first he thought it was a comet, but subsequent calculations of the orbit demonstrated that Uranus was actually a planet, about twice as far from the Sun as Saturn and therefore in an orbit that agreed almost exactly with the prediction of the Titius-Bode relation for planetary distances. *See* PLANET.

The planet and its orbit. The main orbital elements are the semimajor axis (mean distance to the Sun) of 1.78×10^9 mi (2.87×10^9 km); the eccentricity of 0.047; the sidereal period of 83.75 years; the orbital velocity of 4.25 mi/s (6.8 km/s); and the inclination of orbital plane to ecliptic of 0.8°.

These characteristics indicate a normal, well-behaved orbital motion. Therefore it is surprising to discover that the obliquity (inclination of rotational axis to orbit plane) of Uranus is 98°, exceeded only by the obliquities of Pluto and Venus. This means that the axis is almost in the plane of the orbit, and thus the seasons on Uranus are very unusual. During summer in one hemisphere, the pole points almost directly toward the Sun while the other hemisphere is in total darkness. Forty-two years later, the situation is reversed. Thus, "day" and "night" for an observer at the north or south pole of Uranus each last more than 40 years. Recent years of solstice (pole toward Sun) are 1944 and 1985, and recent and future years of equinox (equator toward Sun) are 1966 and 2007. The south pole was almost directly facing the Sun when the *Voyager 2* spacecraft sped past the planet in January 1986. The actual period of rotation of Uranus is 17.24 h. *See* PLUTO; SPACE PROBE; VENUS.

The maximum apparent equatorial diameter of the disk of this distant planet is about 4.″1 (the Moon has an apparent diameter of 31′). The corresponding linear equatorial diameter is 31,763 mi (51,118 km). The mass is 14.5 times the mass of the Earth, a value that can be well determined from the observed motions of the satellites. The corresponding mean density is 1.27 g/cm^3, which is greater than that of Saturn even though Uranus is smaller than Saturn. This means that Uranus is richer than Saturn (or Jupiter) in elements heavier than hydrogen and helium but not nearly so rich in these elements as Earth (mean density 5.5). The same is true for Neptune; thus these two planets constitute a distinct subgroup in the outer solar system when compared with Jupiter and Saturn. *See* JUPITER; NEPTUNE; SATURN.

The apparent visual magnitude of Uranus at mean opposition, that is, when closest to Earth, is +5.5. Therefore, this planet is just visible to the unaided eye in a dark sky when its position among the stars is known. The corresponding value of the albedo, or reflectivity, is 0.5, the same as Saturn and Jupiter, and slightly higher than Neptune (0.4). The temperature that Uranus would assume in simple equilibrium with incident solar radiation at this distance is about 55 K (−360°F). This is essentially identical with the value found by direct measurement. There is thus no evidence for the existence of an internal energy

source as observed for Jupiter, Saturn, and Neptune. The temperature measured by *Voyager 2* was the same at both the sunlit and dark poles. Thermal radiation has been detected from the lower atmosphere of Uranus, indicating that the temperature increases with depth but much more slowly than in the case of the other major planets. At a wavelength of 21 cm, a temperature of 250 ± 60 K (9 ± 90°F) has been measured. *See* ALBEDO.

Through the telescope, Uranus appears as a small, slightly elliptical blue-green disk. This appearance is confirmed in the nearly featureless pictures of the planet obtained by *Voyager 2*. Uranus owes its characteristic aquamarine color to the relatively high proportion of methane in its atmosphere. This gas absorbs the orange and red wavelengths from incident sunlight, scattering back blue-green light to an observer. The methane abundance is 20 to 30 times the amount corresponding to a solar distribution of the elements, in agreement with the proportion of heavy elements in the planet deduced from its average density. In striking contrast, the proportion of helium to hydrogen, the two most abundant gases in the planet's atmosphere, is essentially equal to that observed in the Sun.

Like Jupiter and Saturn, Uranus exhibits zonal winds that are parallel to the equator, despite the unusual orientation of the planet's rotational axis. These winds appear to increase in speed with increasing latitude, unlike the winds on Jupiter and Saturn, which are strongest at the equator. They move in the same direction as the rotation of Uranus, producing a circulation pattern like that on Neptune. The causes of this circulation and the way it maintains the two poles at the same temperature have not yet been determined.

Uranus has a magnetic field equivalent to that of a bar magnet offset from the planet's rotational axis by one-third of the planetary radius and tilted by 60°. (On Earth the field is inclined only 11° to the rotational axis and offset by 0.07 of the radius.) This means that the poles of the magnetic field on Uranus occur at subtropical latitudes instead of near 80°, as on Earth. This field orientation is similar to that of Neptune and must reflect a similar (but not yet understood) mechanism for field generation on these two planets. The strength of the magnetic field on Uranus is five times that of Earth, but varies considerably over the planet. This field traps belts of electrons and protons (like Earth's Van Allen belts) in a magnetosphere whose unusual configuration is determined by the orientations of the planet's magnetic and rotational axes. Electrons and protons from these Van Allen belts bombard the upper atmosphere of Uranus and give rise to an aurora, as they do on Earth, but this aurora accurs in the subtropics instead of at high latitudes. *See* AURORA; MAGNETOSPHERE; PLANETARY PHYSICS; VAN ALLEN RADIATION.

Satellites. Before the *Voyager 2* encounter, only five satellites of Uranus had been discovered. They form a remarkably regular system with low orbital eccentricities and inclinations close to the plane of the planet's equator. The *Voyager* cameras found 11 additional small satellites closer to the planet, including two that may serve as gravitational "shepherds" for the outermost ring (see **table**). Two additional inner satellites were found from Earth in 2003. Like Jupiter, Saturn, and Neptune, Uranus has more distant, irregular satellites (moons that move in orbits with high inclinations or eccentricities). Two of these unusual moons were discovered with the Palomar 200-in. (5-m) telescope in 1997, three more were discovered

Satellites of Uranus*

Satellite	Mean distance from Uranus 10^3 mi	Mean distance from Uranus 10^3 km	Sidereal period, days	Radius mi	Radius km	Albedo (reflectivity)	Density, g/cm^3‡
Cordelia†	30.9	49.8	0.34	8	13		
Ophelia†	33.4	53.8	0.38	9	15		
Bianca	36.8	59.2	0.44	14	22		
Cressida	38.4	61.8	0.46	20	32		
Desdemona	38.9	62.6	0.47	19	30		
Juliet	40.0	64.4	0.49	26	42		
Portia	41.1	66.1	0.51	34	55		
Rosalind	43.5	70.0	0.56	19	30		
Belinda	46.8	75.3	0.62	21	34		
Puck	53.4	86.0	0.76	48	78	0.07	
Miranda	80.8	130	1.41	147	236	0.3	1.2
Ariel	119	191	2.52	360	580	0.4	1.7
Umbriel	165	266	4.14	364	585	0.2	1.4
Titania	271	436	8.71	490	789	0.3	1.7
Oberon	362	583	13.5	473	761	0.2	1.6
Caliban§	4474	7200	579	19	30	0.07¶	
Stephano§	4971	8000	677	6	10	0.07¶	
Trinculo§	5284	8504	759	3	5	0.07¶	
Sycorax§	7581	12,200	1283	37	60	0.07¶	
Prospero§	10,066	16,200	1962	9	15	0.07¶	
Setebos§	10,229	17,500	2209	9	15	0.07¶	

*Only satellites with well-determined orbits as of March 1, 2004, are listed.
†Shepherding satellites for ring.
‡Eccentricity of 0 is a circle, 1.00 is a parabola.
§Outer, retrograde satellites. Radii are derived by assuming albedos of 0.07. Densities are unknown.
¶Assumed.

in 1999, and four more since then. Only the first six to be found have reliable orbits, but the orbits of the others are being improved and the search for additional satellites continues.

The spacecraft was able to obtain pictures of the surfaces of all five of the larger satellites and resolved features on Puck as well. For reasons that are not yet understood, there is increasing evidence of internal geological activity on the larger satellites with decreasing distance from the planet. Miranda, the smallest but also the closest, revealed the most spectacular diversity of surface features (**Fig. 1**). Impact craters are present on all of these objects as well.

With densities of approximately 1.5 g/cm^3, these satellites are presumably composed of a mixture of abundant ices (dominated by water ice) and some rocky material. The generally low reflectivities and the deposits of dark material observed in some locations on the satellites' surfaces are ascribed to the presence of carbon-rich compounds. These compounds are more neutral in reflectivity than the dark material found on Saturn's satellite Iapetus, however, suggesting that more than one type of organic material is present in the outer solar system.

Rings. A system of ten rings can be observed around Uranus. They are designated by a mixture of Greek letters and Arabic numerals (**Fig. 2**). Since the rotational axis of Uranus is almost in the plane of the planet's orbit around the Sun, the rings are nearly perpendicular to this plane, which made it possible to discover nine of them when they passed in front of a star.

The *Voyager* cameras added a tenth ring, lambda, between delta and epsilon. All of these rings are remarkably narrow. The widest is epsilon, whose mean width of only 36 mi (58 km) is still six to seven times larger than that of the next widest rings. Like the small satellites, the rings are as dark as coal, with reflectivities of only 3–5%. (The Earth's Moon has a reflectivity of 11%.) The spaces between these narrow rings are often thinly populated with dust. Measurements of the diminution of starlight by the rings using a photometer on the spacecraft as well as measurements of the effect of the rings on radio transmission from the spacecraft itself suggest an absence of small particles within the epsilon ring. The origin and means of confinement of these rings are topics of ongoing research.

Cosmogony. It appears that all four giant planets have approximately equal masses of heavy elements in their interiors, equivalent to 10–20 Earth masses. The great difference between Jupiter and Saturn on the one hand, and Uranus and Neptune on the other, can then be understood in terms of differences in their atmospheres. The outer two planets acquired less gas from the original solar nebula when they formed, perhaps because the nebula was simply thinner at this large distance from its center. Since the nebular gases were predominantly hydrogen and helium, Uranus and Neptune began with a smaller endowment of these two light elements and thus ended up with mean densities larger than Jupiter and Saturn. The excess methane remarked on above was produced by the core, which was unable to trap the two lightest gases. The hydrogen and helium are therefore present in a solarlike ratio to one another.

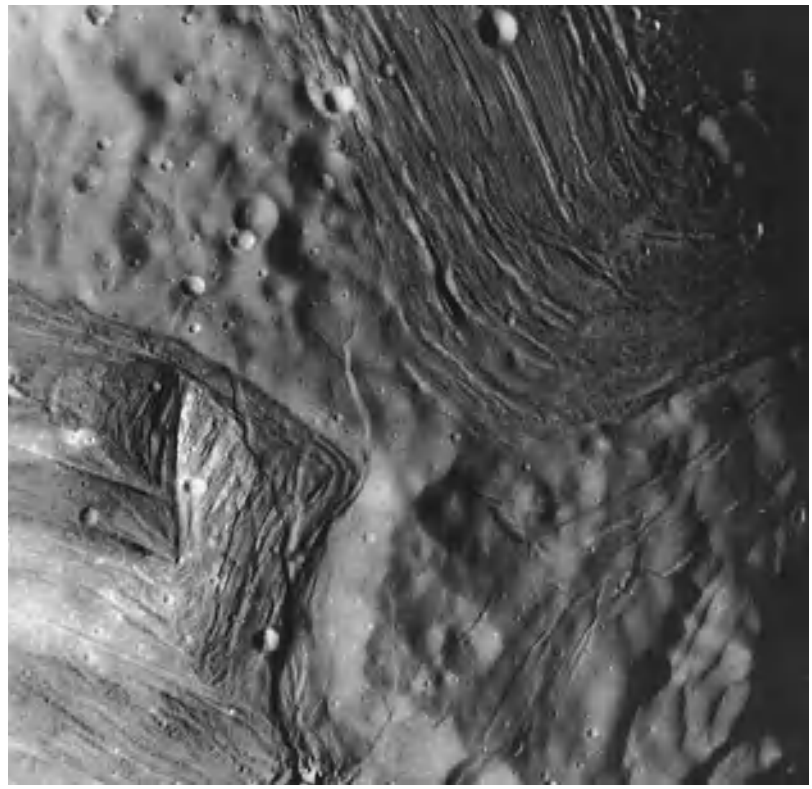

Fig. 1. Miranda viewed from *Voyager 2*. Features as small as 2000 ft (600 m) across are visible. The two regions of grooved terrain are remarkable, showing evidence of large-scale geological activity on this tiny satellite.

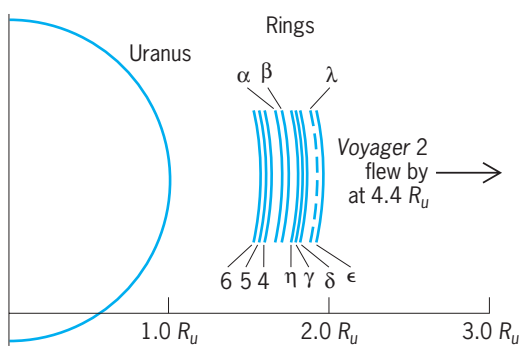

Fig. 2. Diagram of Uranus ring system, showing one hemisphere of Uranus at the left and segments of the 10 narrow rings. The rings all lie between 1.5 and 2 times the radius of Uranus (R_u) from the planet's center.

The large obliquity of Uranus and its lack of an internal heat source remain unsolved problems. The obliquity is commonly attributed to the impact by a large body during the late phases of core formation, but a detailed theory has not yet been developed.

Tobias C. Owen

Bibliography. J. K. Beatty, C. C. Petersen, and A. Chaikin (eds.), *The New Solar System*, 4th ed., Cambridge University Press, 1999; R. Greenberg and A. Brahic (eds.), *Planetary Rings*, University of Arizona Press, 1984; G. Hunt, *Uranus and the Outer Planets*, Cambridge University Press, 1982; D. Morrison

and T. Owen, *The Planetary System*, 3d ed., Addison-Wesley, 2003; Reports of *Voyager 2* encounter with Uranus, *Science*, 233:39–109, 1986; D. A. Rothery, *Satellites of the Outer Planets: Worlds in Their Own Right*, 2d ed., Oxford University Press, 1999; F. W. Taylor, *The Cambridge Photographic Guide to the Planets*, Cambridge University Press, 2001.

Urban climatology

The branch of climatology concerned with urban areas. These locales produce significant changes in the surface of the Earth and the quality of the air. In turn, surface climate in the vicinity of urban sites is altered. The era of urbanization on a worldwide scale has been accompanied by unintentional, measurable changes in city climate.

Controlling factors. The process of urbanization changes the physical surroundings and induces alterations in the energy, moisture, and motion regime near the surface. Most of these alterations may be traced to causal factors such as air pollution; anthropogenic heat; surface waterproofing; thermal properties of the surface materials; and morphology of the surface and its specific three-dimensional geometry—building spacing, height, orientation, vegetative layering, and the overall dimensions and geography of these elements. Other factors that must be considered are relief, nearness to water bodies, size of the city, population density, and land-use distributions (**Table 1**).

TABLE 1. Climatic alterations produced by cities

Element	Compared to rural environs
Contaminants	
Condensation nuclei	10 times more
Particulates	10 times more
Gaseous admixtures	5–25 times more
Radiation	
Total on horizontal surface	0–20% less
Ultraviolet: winter	30% less
summer	5% less
Sunshine duration	5–15% less
Cloudiness	
Clouds	5–10% more
Fog: winter	100% more
summer	30% more
Precipitation	
Amounts	5–15% more
Days with <5 mm (0.2 in.)	10% more
Snowfall: inner city	5–10% less
lee of city	10% more
Thunderstorms	10–15% more
Temperature	
Annual mean	0.5–3.0°C (0.9–5.4°F) more
Winter minima (average)	1–2°C (1.8–3.6°F) more
Summer maxima	1–3°C (1.8–5.4°F) more
Heating degree days	10% less
Relative humidity	
Annual mean	6% less
Winter	2% less
Summer	8% less
Wind speed	
Annual mean	20–30% less
Extreme gusts	10–20% less
Calm	5–20% more

*After H. E. Landsberg, *The Urban Climate*, Academic Press, 1981.

Urban heat island. In general, cities are warmer than their surroundings, a fact documented over a century ago. They are islands or spots on the broader, more rural surrounding land. Thus, cities produce a heat island effect on the spatial distribution of temperatures (**Fig. 1**). The timing of a maximum heat island is followed by a lag shortly after sundown (Fig. 1*c*–*e*), as urban surfaces, which absorbed and stored daytime heat, retain heat and affect the overlying air. Meantime, rural areas cool at a rapid rate (Fig. 1*d*).

There are a number of energy processes that are altered to create warming, and various features lead to those alterations (**Table 2**). City size, the morphology of the city, land-use configuration, and the geographic setting (such as relief, elevation, and regional climate) dictate the intensity of the heat island, its geographic extent, its orientation, and its persistence through time. Individual causes for heat island formation are related to city geometry, air pollution, surface materials, and anthropogenic heat emission. There are two atmospheric layers in an urban environment. Besides the planetary boundary layer (PBL) outside and extending well above the city, the first of these layers is the urban boundary layer. This layer is due to the spatially integrated heat and moisture exchanges between the city and its overlying air (**Fig. 2**). The surface of the city corresponds to the level of the urban canopy layer. Fluxes across this plane comprise those from individual units of the second layer, the urban canopy layer, such as roofs, canyon tops, trees, lawns, and roads, integrated over larger land-use divisions (for example, suburbs).

Energy and moisture balance. An understanding of changes of climate in the city is achieved by investigating the energy and moisture regimes in the urban and rural locales.

The urban energy balance is the sum of the net all-wave radiation (inputs and outputs of solar and long-wave radiation) and the energy released through anthropogenic activities. This sum is calculated by adding the latent heat flux (evapotranspiration), sensible (turbulent) heat flux, storage heat flux (into surface, buildings, and so forth), and the net heat advection (horizontal heat transfer). The latent heat flux is the product of the latent heat of vaporization and the mass equivalent of the latent heat flux. These values apply to the top of a volume that extends to sufficient depth that vertical heat and water exchange are negligible. The fluxes are determined at the top of the volume (for example, in watts per square meter).

The urban water balance is the sum of the precipitation, the piped water supply, and the water released through anthropogenic activities. This sum is calculated by adding the parameters of runoff, the measured change in storage for the period of interest, the moisture advection, and the mass equivalent of the evapotranspiration.

The moisture exchanges of cities differ in several ways from those occurring in rural settings. Precipitation can increase in, or downwind of, a

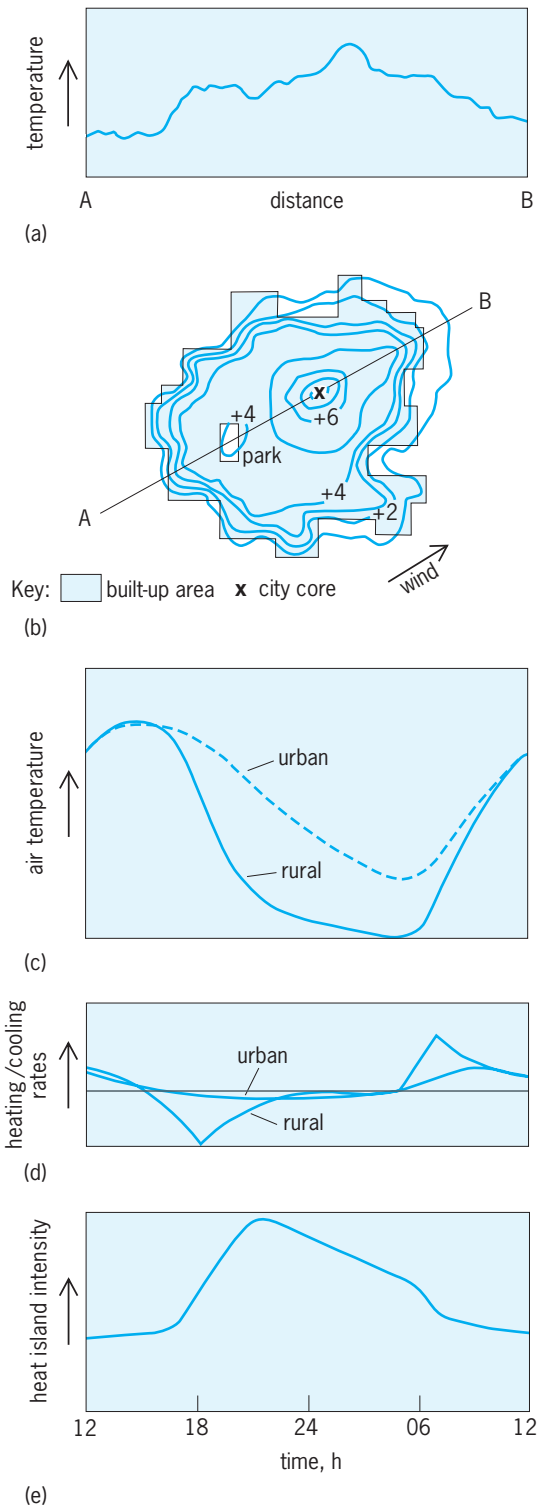

Fig. 1. Example of heat island magnitude, spatial arrangement, and timing during the day under "ideal" (calm, clear) weather. (*a*) The cross-sectional changes in temperature across the heat island spatial pattern in relation to (*b*) the plan outline of a city, (*c*) air temperature in an urban and rural locale during 24 h, (*d*) air heating and cooling rates per hour, and (*e*) resulting heat island intensity. Vertical scale units are around 2°C for *a*, *c*, and *d* and 2°C/h for *e*. Values at the contours in *b* represent temperatures in °C associated with built-up area in comparison to rural areas. °F = (°C × 1.8) + 32. (*After T. R. Oke, The energetic basis of the urban heat island, Quart. J. Roy. Meteorol. Soc., 108:1–24, 1982*)

large urban area (Table 1), and significant increases occur in the warm season, which is associated with more local convective processes. The city environment is aerodynamically rougher, contains more pollution condensation nuclei, promotes more thermal and mechanical turbulence, has an effect on the formation of convective clouds, and enhances precipitation. Development of a clear characterization of the role of urban places in affecting precipitation is complicated by the topographic settings of cities.

Since the piped water supply and the water released through anthropogenic activities are extra sources of moisture in a city, there is a net gain of moisture by the process of urbanization. Water loss through mass exchange of water (evapotranspiration) is considered to be less in a city because of a "waterproofing effect." This effect is brought about by replacement of moisture surfaces with more impervious ones with the result that the ratios of latent heat flux to net all-wave radiation have been found to exceed expectations. For example, in a temperate city, evapotranspiration has been shown to constitute 38% of the losses of the annual external water balance and 81% of the losses of the summer water balance. Assuming moisture advection to be negligible, runoff would increase in cities. Many studies have indicated increased runoff in urban environments as well as altered hydrographic characteristics of urban streams. Changes in land use have been shown to have four effects on the hydrology of an urban area: (1) peak flow is increased; (2) total runoff is increased; (3) water quality is lowered; and (4) hydrologic amenities such as the appearance of river channel and esthetic impressions are lowered. Discharge regimes are a function of the percentage of impervious surface area in cities and percentage of area served by storm sewers. *See* HYDROGRAPHY; HYDROLOGY.

Since for a given period evapotranspiration and change in water storage are typically lower, dewpoint is usually lower in the daytime in urban areas. At night, a humidity island may result in the city because of extra transfer of evapotranspiration, and water released through anthropogenic activities. The largest unknown in analyses of water balance for cities is a basic understanding of transfers of evapotranspiration from complex suburban and urban landscapes. The measured spatial variability of evapotranspiration is dependent on the complex source areas of moisture transfer to fixed observed points across the urban and suburban environment.

Air movement in cities. Urban areas affect the speeds and directions of winds on both a micro- and mesoscale. The aerodynamic roughness of cities is larger than that of surrounding rural environments, thus inducing more frictional drag. The heat island effect influences the pressure field and vertical stability of the air over the city. The city influence on wind is also a function of the gradient air motion above the city. Many studies have shown that as general winds in the region increase, the intensity of the heat island diminishes. However, the built-up environment of the city induces mechanical turbulence and may

TABLE 2. Suggested causes of the urban heat island

Energy processes that are altered to create warming	Features of urbanization underlying energy balance changes
Urban canopy layer	
Increased absorption of short-wave radiation	Canyon geometry—increased surface area and multiple reflection
Increased long-wave radiation from the sky	Air pollution—greater absorption and reemission
Decreased long-wave radiation loss	Canyon geometry—reduction of sky view factor
Anthropogenic heat source	Building and traffic heat losses
Increased sensible heat storage	Construction materials—increased thermal admittance
Decreased evapotranspiration	Construction materials—increased "waterproofing"
Decreased total turbulent heat transport	Canyon geometry—reduction of wind speed
Urban boundary layer	
Increased absorption of short-wave radiation	Air pollution—increased aerosol absorption
Anthropogenic heat source	Chimney and stack heat losses
Increased sensible heat input-entrainment from below	Canopy heat island—increased heat flux from canopy layer and roofs
Increased sensible heat input-entrainment from above	Heat island, roughness—increased turbulent entrainment

*After T. R. Oke, The energetic basis of the urban heat island, *Quart. J. Roy. Meteorol. Soc.*, 108:1–24, 1982.

speed up winds across the area by some 30%. For conditions of light regional airflow, development of an intense heat island and its effects on vertical stability and pressure field variations will tend to promote quite variable patterns of air motion in the urban area. In rural areas, air is more stable, and little motion occurs. In cities, the thermal influences induce increases in turbulence and wind speed.

Horizontal advection of air, and the heat content of that air, further complicates a complete understanding of heat fluxes from sectors of urban environments. This advective component of motion and heat transfer may play a role in the enhanced dispersion potential of pollution in urban regions. The large spatial variability of heat fluxes across urban areas has been postulated to affect plume spread. *See* ATMOSPHERIC GENERAL CIRCULATION; WIND.

Heat island influences affect the mixing height of the atmosphere—a measure of vertical pollutant mixing above the urban canopy layer. Basically, a domelike lifting of inversions occurs over the city center, with subsidence of the inversion on the perimeter of the urban region. Aircraft overflights and remote detectors measure atmospheric mixing and inversion conditions with devices such as lidar and Doppler radar. *See* DOPPLER RADAR; LIDAR.

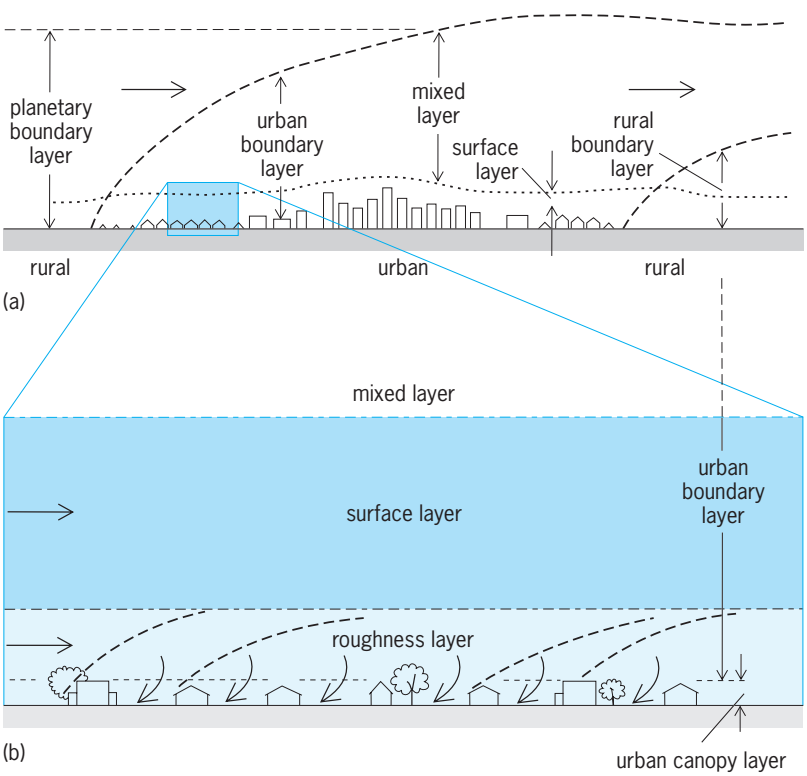

Fig. 2. Diagrams showing exchange of heat and moisture between a city and its overlying air. (*a*) Idealized arrangement of boundary-layer structures over a city; horizontal arrows indicate general direction of prevailing wind. (*b*) A blowup of a section of the rural layer in *a*, showing the mixed, surface, and roughness layers, including eddies of air shown by curved arrows, and individual mixed layers associated with land conditions, indicated by broken curved lines. (*After T. R. Oke, The urban energy balance, Prog. Phys. Geog., 12:471–508, 1988*)

Satellite climatology. Satellite platforms are used to study the distribution of heat islands and the finer instantaneous spatial details of the heat and moisture at the surface. It is possible that uncritical acceptance of remotely sensed data can lead to erroneous conclusions. In the search for in-depth physical understanding of the causes of heat islands, many questions arise. For example, what is the nature of the true surface "seen" by space platform sensors? What is the relation between temperatures interpreted by the satellite and the three-dimensional surface temperatures across the urban environment? Satellite estimates of radiant surface temperatures provide a "snapshot" of the entire city region at a resolution not made possible previously—a large advantage in assessing the spatial dimensions of a heat island. However, the resultant heat island interpretation is specific to the satellite observation system, such as the instantaneous time of the satellite pass over an area (not often at the time of maximum heat island formation) and the view angle of the resultant temperature pattern of surfaces seen only by the satellite radiometer. *See* CLIMATE MODIFICATION; CLIMATOLOGY; SATELLITE METEOROLOGY.

Anthony J. Brazel

Bibliography. J. K. S. Ching, J. F. Clarke, and J. M. Godowitch, Modulation of heat flux by different scales of advection in an urban environment, *Boundary-Layer Meteorol.*, 25:171–191, 1983; C. S. B. Grimmond and T. R. Oke, An evapotranspiration-interception model for urban areas, *Water Resources Res.*, 27:1739–1755, 1991; H. E. Landsberg, *The Urban Climate*, 1981; T. R. Oke,

The energetic basis of the urban heat island, *Quart. J. Roy. Meteorol. Soc.*, 108:1–24, 1982; T. R. Oke, The urban energy balance, *Prog. Phys. Geog.*, 12:471–508, 1988; M. Roth and T. R. Oke, Satellite-derived urban heat islands from three coastal cities and the utilization of such data in urban climatology, *Int. J. Remote Sens.*, 10:1699–1720, 1989.

Urea

A colorless crystalline compound, formula CH_4N_2O, melting point 132.7°C (270.9°F). Urea is also known as carbamide and carbonyl diamide, and has numerous trade names as well. It is highly soluble in water and is odorless in its purest state, although most samples of even high purity have an ammonia odor. The diamide of carbonic acid, urea has the structure below.

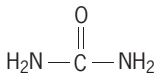

Urea occurs in nature as the major nitrogen-containing end product of protein metabolism by mammals, which excrete urea in the urine. The adult human body discharges almost 50 g (1.8 oz) of urea daily. Urea was first isolated in 1773 by G. F. Rouelle. By preparing urea from potassium cyanate (KCNO) and ammonium sulfate (NH_4SO_4) in 1828, F. Wöhler achieved a milestone, the first synthesis of an organic molecule from inorganic starting materials, and thus heralded the modern science of organic chemistry. *See* NITROGEN; PROTEIN METABOLISM.

Preparation. Prior to World War II, urea was prepared commercially by the hydrolysis (reaction with water) of calcium cyanamid ($CaCN_2$). The development during the 1940s of large-scale high-pressure reaction vessels made of nickel-chromium alloy allowed the Kolbe process to become the principal means for producing urea. This process, first developed in the 1870s, involves the reaction of ammonia (NH_3) and carbon dioxide (CO_2) at 100 atm (10 megapascals) and 150–200°C (302–392°F) in the presence of a catalyst such as an acid or a base. This reaction forms ammonium carbamate (NH_4COONH_4), which upon further heating undergoes dehydration (loss of water) to yield urea. The overall process is shown in reaction (1). This is a very important industrial chemical process, but there are numerous other methods for producing urea.

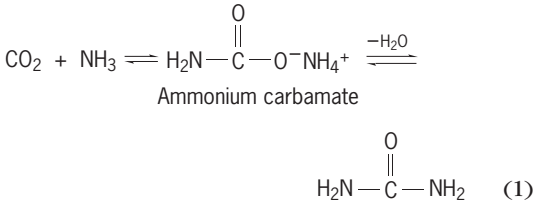

(1)

Uses. Because of its high nitrogen content (46.65% by weight), urea is a popular fertilizer. About three-fourths of the urea produced commercially is used for this purpose. After application to soil, usually as a solution in water, urea gradually undergoes hydrolysis to ammonia (or ammonium ion) and carbonate (or carbon dioxide). Another major use of urea is as an ingredient for the production of urea-formaldehyde resins, extremely effective adhesives used for laminating plywood and in manufacturing particle board, and the basis for such plastics as malamine. *See* FERTILIZER; UREA-FORMALDEHYDE RESINS.

Other uses of urea include its utilization in medicine as a diuretic. In the past, it was used to reduce intracranial and intraocular pressure, and as a topical antiseptic. It is still used for these purposes, to some extent, in veterinary medicine and animal husbandry, where it also finds application as a protein feed supplement for cattle and sheep. Urea has been used to brown baked goods such as pretzels. It is a stabilizer for nitrocellulose explosives because of its ability to neutralize the nitric acid that is formed from, and accelerates, the decomposition of the nitrocellulose. Urea was once used for flameproofing fabrics. Mixed with barium hydroxide, urea is applied to limestone monuments to slow erosion by acid rain and acidic pollutants.

Chemistry. Because of its ability to bond to itself in an orderly fashion by hydrogen bonding, urea reversibly forms inclusion complexes with straight-chained hydrocarbons having seven or more carbon atoms. These complexes consist of the hydrocarbon encased in a cylindrical wrapping of urea molecules. They are isolable as waxy solids, and can be used for separating straight-chained from branched hydrocarbons in a mixture. *See* CLATHRATE COMPOUNDS; HYDROGEN BOND.

Urea has numerous applications in synthetic organic chemistry because its nitrogen and oxygen atoms are nucleophilic and its carbon atom is electrophilic. One important example of a nucleophilic reaction of urea is its reaction with acid chlorides or esters to form *N*-acyl derivatives known as ureids. Examples of ureids are acetylurea ($CH_3CO—NHCONH_2$) and the diureid diacetylurea ($CH_3CO—NHCONH—COCH_3$). The six-membered cyclic diureids formed from urea and various substituted malonic acid esters [reaction (2)] are known as barbiturates. The barbiturates are potent, addictive

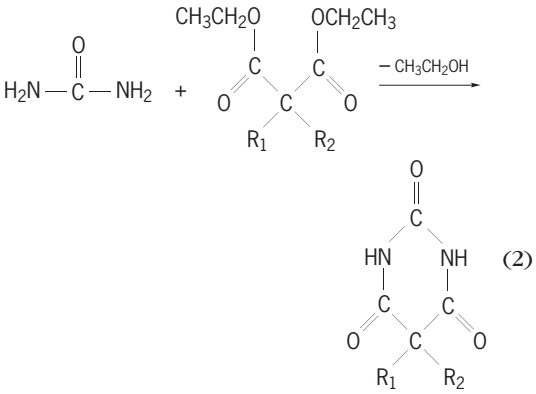

(2)

$R_1 = R_2 = H$: barbituric acid
$R_1 = R_2 = CH_3CH_2$: veronal
$R_1 = CH_3CH_2, R_2 = C_6H_5$: phenobarbital

hypnotics (sleep inducers) used in medicine. *See* BARBITURATES; ESTER.

An example of an electrophilic reaction of urea is its reaction with ethyl alcohol, under pressure, to form urethane (ethyl urethane, ethyl carbamate), as shown in reaction (3). Urethane is also synthesized

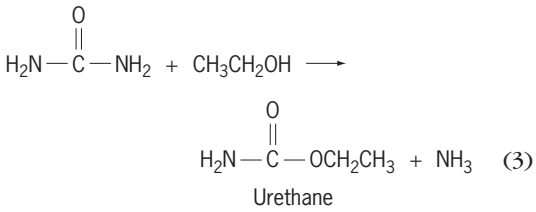

by heating urea nitrate, the salt formed from the reaction between urea and nitric acid, with ethyl alcohol. Urethane was once used as a mild hypnotic, as a topical bactericide, and in the treatment of leukemia. Its toxicity precludes its continued use for these purposes. Urethane is currently used primarily as an organic solvent and as a reagent for organic synthesis. *See* ORGANIC CHEMISTRY; ORGANIC SYNTHESIS.

Robert D. Walkup

Bibliography. L. F. Fieser and M. Fieser, *Reagents for Organic Synthesis*, vol. 1, 1967; F. A. Lowenheim and M. K. Moran, *Faith, Keyes and Clark's Industrial Chemicals*, 4th ed., 1975; B. Meyer, *Urea-Formaldehyde Resins*, 1979; T. W. Solomons, *Organic Chemistry*, 7th ed., 1999.

Urea-formaldehyde resins

The condensation products obtained by the reaction of urea or melamine with formaldehyde. Resinous condensation products of formaldehyde with other nitrogen-containing compounds, for example, aniline and amides, also belong to this group of resins but have gained only limited utility. Resins derived from the condensation of formaldehyde with urea, melamine, aniline, and *p*-toluene sulfonamide are discussed below.

Urea- and melamine-formaldehyde resins. Because the amino resins possess an excellent combination of physical properties and can be easily fabricated in a variety of colors, they are used as adhesives, laminating resins, molding compounds, paper and textile finishes, and surface coating. Methods of utilizing the resins are generally similar to methods employed for several other condensation resins, such as the phenolic or epoxy resins. First, intermediate condensation resins are prepared; then, after compounding, the combinations are cured to yield a finished product.

The intermediate condensation products are prepared by the reaction of urea or melamine formaldehyde under neutral or mildly alkaline conditions to form mono-, di-, or polymethylol derivatives, with reactions (1)–(3).

If a soluble resin is desired for surface coatings, the dimethylol derivatives are treated with an alcohol in an acidic medium to form ethers, reaction (4).

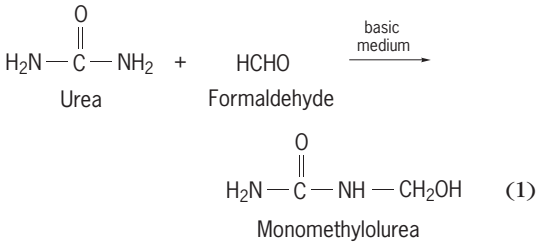

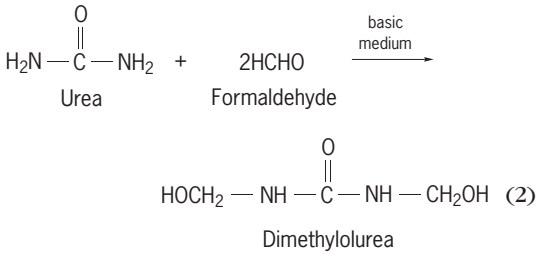

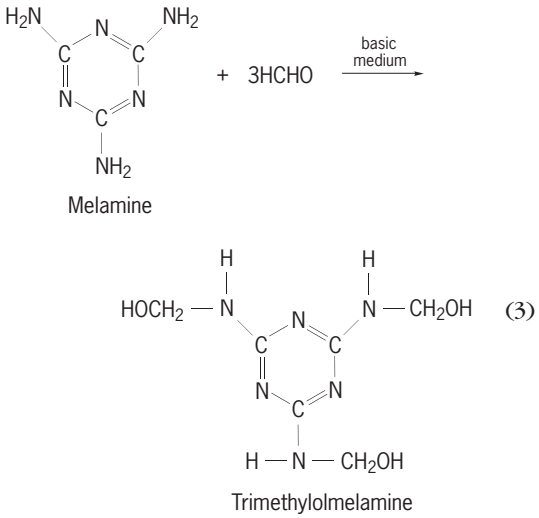

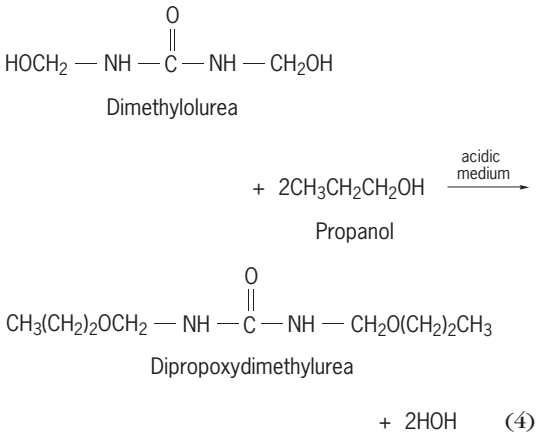

For other purposes, the low-molecular-weight methylol derivatives can then be compounded with fillers, catalysts, plasticizers, and pigments, and cured, usually by the application of heat and pressure. During the curing operation, the hydroxyl portion of a methylol group condenses with a hydrogen atom to yield water. Such a condensation can be

represented by reaction (5) for the cross-linking of methylol urea.

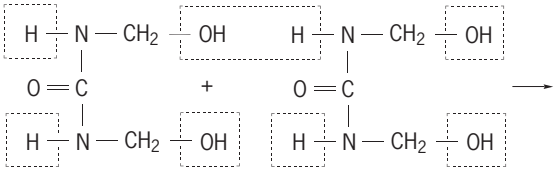

n molecules *n* molecules

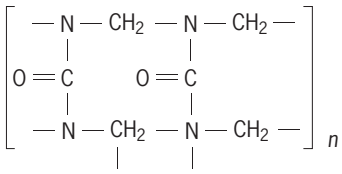

Cross-linked polymer

Other reactions are also involved in curing. For example, the methylol groups can react with hydroxyl groups of alcohols (added as plasticizers), clays, or cellulose (added as fillers). Methylol groups can also react with each other to form ether linkages.

For immediate curing, a free acid can be used as a cross-linking catalyst. However, for preparation of a stable molding powder that can be stored, a salt or ester that will liberate its corresponding acid at the high temperature used for molding usually serves as the catalyst.

As may be seen from reactions (1)–(5), the completion of the curing reaction depends on the elimination of as much water as possible. Although most of the water is lost as vapor, a hygroscopic filler, such as cellulose, asbestos, or certain clays, is added to absorb the last traces of moisture. In addition, the filler can also react with some of the methylol groups of the resin intermediates.

Although the fabrication and applications of both the urea- and melamine-type resins are generally similar, there are several important differences in properties. The melamine-type resins are more resistant to water, marring, and heat than the urea-type resins, but are somewhat more expensive.

Curing in a mold results in translucent or opaque products, depending on the nature of the fillers and pigments used. The ease with which color can be introduced into amino resins has made it possible to make very attractive articles, such as dinnerware, buttons, appliance cases, handles, and knobs. In applications where resistance to heat and scratching is important, for example, in dinnerware, the melamine resins are preferred. *See* PLASTICS PROCESSING.

Another important application of both resins is in adhesives, especially for lamination of furniture and plywood. Depending on the resin and catalyst used, adhesives that can be cured under a variety of conditions may be formulated. Melamine resins are especially valuable as adhesives for laminates of paper or fabrics, including glass cloth. *See* ADHESIVE.

Foams were developed for use in packaging and as thermal and acoustical insulation.

Soluble urea-formaldehyde and melamine-formaldehyde resins may be used to impregnate cloth or to treat paper. Curing of the resins in the cloth results in improved qualities, such as resistance to creasing or, in the case of melamine resins, in resistance to shrinkage in woods. Curing of the resins in paper results in an improvement in the strength of the paper when wet.

Although unmodified resins are common in the applications cited, modified resins or blends with other resins are often useful in some cases. In combination with alkyd resins, the ether derivatives, especially the ethers of the melamine resins, are components in the formulation of baking enamels that are hard and resistant to water and detergents. Resins modified by the use of an alkyl-substituted urea or melamine are more flexible than the unmodified resin and are thus useful in coating compositions. Blends of urea- or melamine-type resins with resins derived from resorcinol and formaldehyde are sometimes employed in adhesive compositions and as binders for the sawdust or wood chips used in the manufacture of particulate boards.

Another application of blends is the use of urea-type resin blends as impregnants for fabric used in the manufacture of minimum-care or wash-and-wear cotton goods. *See* TEXTILE CHEMISTRY.

Aniline- and sulfonamide-formaldehyde resins. In addition to urea and melamine, other compounds containing $—NH_2$ groups can condense with formaldehyde to form methylol derivatives which are capable of further reaction. Although the corresponding resins have received limited attention, two examples are discussed here, aniline and *p*-toluenesulfonamide.

In a neutral medium, the reaction of aniline with formaldehyde yields a methylol derivative that exists as a cyclic trimer. When the trimer is heated, further condensation results in a resin with the probable monomer unit

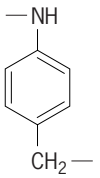

Some cross-linking through the phenyl rings may also occur.

Because of their resistance to the absorption of water, the resins have been used in electrical applications, such as insulation or panels, where their natural brown color is not objectionable.

Similarly, resins can be prepared by curing methylol derivatives resulting from the condensation of *p*-toluenesulfonamide with formaldehyde. The probable monomer unit is

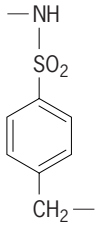

The resulting polymers are less colored than the aniline-type resins and have been employed in surface coatings. *See* FORMALDEHYDE; POLYMERIZATION; UREA. John A. Manson

Bibliography. J. R. Fried, *Polymer Science and Technology*, 2d ed., 2003; G. Odian, *Principles of Polymerization*, 4th ed., 2004; M. P. Stevens, *Polymer Chemistry: An Introduction*, 3d ed., 1998.

Uredinomycetes (rust)

An order of fungi known as plant rusts that belong to the division (phylum) Basidiomycota. In nature, all 7000 species are obligate parasites of many vascular plant species. They cause diseases known as rust on numerous cultivated crops such as asparagus, onions, beans, snapdragons, and sunflowers. Ferns, fern allies, pines and their relatives, and many other trees are also attacked. Each rust species infects one or just a few closely related plant host species.

The body of a rust fungus consists of numerous microscopic, threadlike, branching hyphae that grow inside the tissues and between the cells of the host plant. Specialized feeding structures, haustoria, grow from these hyphae into the host cells of the plant and provide nourishment for the hyphae. Hyphal cells are binucleate. Depending upon the rust species and environmental conditions, up to six different kinds of spore-producing structures, the sori, may be produced by one rust species. These sori, which are usually no more than 0.5 mm in width, may be powdery, waxy, or crustlike, and whitish, yellow, orange, brown, or blackish. An infection may produce many sori, and each sorus may produce thousands of spores. The spores are mostly wind-borne. The urediniospores of *Puccinia graminis* (the rust species that causes severe epidemics on wheat, some other cereals, and other grasses) may be disseminated for several thousand kilometers.

Rust teliospores that are produced in sori called telia are essential elements for classification. More than 125 different forms of telia have been discovered. These variations in the telia are the basis for the classification of the genera of rusts. Teliospores consist of one or more specialized probasidial cells. During the development of each of these cells, nuclei fuse and meiosis occurs. This meiosis results in a septate metabasidium from which four haploid meiospores (basidiospores) are formed. Basidiospores are forcibly ejected from the tips of their tiny stalks and are wind-disseminated. After landing on a susceptible host, under proper conditions basidiospores produce new infections. In cold regions, a rust may survive the winter as dormant, thick-walled, dark-colored teliospores. In regions of long, hot, dry summers, similar teliospores may survive the summer. In many rust species, teliospores mature and produce basidiospores with no or little dormancy, especially in the tropics. Basidiospore infections often give rise to sexual reproductive structures: the spermogonia that produce spermatia, the aecial initials (female structures), and multiple zygotes (aeciospores). Asexual spores (urediniospores) are formed by most rust species and are the stage of rusts most often found.

Although most rust species require only one host species to complete their life cycles (autoecious rusts), some of the best-known rusts require two taxonomically unrelated hosts (heteroecious rusts). For example, the rust *Cronartium ribicola* produces spermogonia and aecia on large galls on stems of pine trees, and uredinia and telia on leaves of currants and gooseberries. *Puccinia graminis* produces spermogonia and aecia on leaves of barberries, and uredinia and telia on wheat and many other grasses.

Rust fungi occur on all continents except Antarctica. They infect host plants in more than 250 vascular plant families. Rust species are especially abundant in the grass, legume, potato, coffee, and sunflower families. Under natural conditions, most rust species are in balance with their hosts and do little harm. Under agricultural ecosystems, many rusts are feared pathogens of important crops, including cereals, many beans and peas, peaches and plums, pine and eucalyptus trees, and coffee. Rust diseases are controlled most effectively by breeding resistant host varieties. This often requires the search for resistance in closely related wild relatives of crop plants followed by cross breeding to introduce wild-type resistance genes into domesticated varieties. Fungicides are used for some rusts. In the case of some heteroecious rusts, eradication of one of the hosts (the noneconomic one) has been successful. Some rust species are used to aid in biological control of weeds. *See* BASIDIOMYCOTA; EUMYCOTA; FUNGI; PLANT PATHOLOGY. Joe F. Hennen

Bibliography. G. B. Cummins and Y. Hiratsuka, *Illustrated Genera of Rust Fungi*, rev. ed., 1983; L. J. Littlefield, *Biology of the Plant Rusts: An Introduction*, 1981; K. J. Scott and A. K. Chakravorty (eds.), *The Rust Fungi*, 1982.

Uric acid

An end product of purine metabolism in humans and higher primates. Uric acid is excreted as such in the urine, and it is very poorly soluble in aqueous solutions, causing it to crystallize when concentrations of the compound in the urine are abnormally high. This leads to crystalluria (excretion of crystals in urine), hematuria (blood in the urine), infection, or urinary tract stones. Other mammals do not experience these problems, as they express an enzyme, uricase, which catalyzes the conversion of uric acid to allantoin, which is highly soluble. A nonprimate exception is the Dalmatian dog, in which high rates

of excretion of uric acid are a consequence of defective renal tubular reabsorption; thus, the uric acid is excreted before it can be oxidized. In birds and reptiles, uric acid is an excretory end product of the metabolism of proteins.

Biosynthesis and metabolism. Purines are nitrogenous bases comprising a nine-membered ring formed by the fusion of a six-membered pyridine ring and a five-membered imidazole ring (see below). Examples of purine bases include adenine and guanine, the building blocks of ribonucleic acid (RNA) and deoxyribonucleic acid (DNA); hypoxanthine; and xanthine, the immediate catabolic precursor of uric acid. As they are found in DNA, purine nucleotides are composed of the purine base, a ribose molecule, and one to three phosphate moieties. Purine nucleotides such as inosinic acid and hypoxanthine ribose phosphate (inosine 5′-monophosphate, or IMP) are synthesized de novo from small molecules, such as glutamine, bicarbonate, formate, glycine, and aspartic acid. They may also be recovered from the free bases by salvage pathways (that is, the reclaiming of purines resulting from the breakdown of cellular DNA and RNA), catalyzed by enzymes such as hypoxanthine phosphoribosyltransferase (HPRT) and adenine phosphoribosyltransferase. The enzyme which catalyzes the conversion of xanthine to uric acid, xanthine oxidase, also catalyzes the formation of xanthine from hypoxanthine. *See* NUCLEIC ACID; NUCLEOTIDE; PURINE.

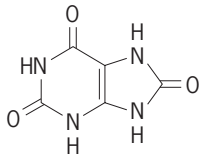

Uric acid

Hyperuricemia. Uric acid may be found in large amounts in blood (hyperuricemia) due to rapid nucleic acid turnover (for example, in leukemia), and especially due to the acute lysis of tumor cells (for example, in successful chemotherapy). In this situation, uric acid crystallization in the kidney can cause kidney failure. Uric acid concentration may also rise in the blood because of diminished renal excretion, the situation in most patients with gout; in renal disease; under conditions such as lactic acidemia or ketoacidosis, in which organic acids such as lactic acid compete with uric acid for tubular reabsorption; and in response to certain diuretics. Hyperuricemia also results from overproduction of purines through the de novo pathway, as occurs in Lesch-Nyhan disease, which is due to a genetically determined deficiency in the HPRT enzyme. Other symptoms of Lesch-Nyhan disease include mental retardation, neurologic abnormalities, and bizarre and self-injurious behavior. Chronic hyperuricemia of any cause may lead to gout, with its painful, disfiguring arthritis and potentially fatal renal failure.

Hyperuricemia that results from diminished renal excretion, but not from renal failure, can be treated with drugs that cause increased excretion of uric acid. However, these drugs are contraindicated in hyperuricemia resulting from overproduction of purines or from nucleic acid breakdown. Treatment is instead by pharmacologic inhibition of xanthine oxidase. *See* GOUT; KIDNEY DISORDERS; METABOLIC DISORDERS; METABOLISM; PROTEIN METABOLISM.

William Nyhan

Bibliography. H. Jinnah, in C. R. Scriver et al. (eds.), *The Metabolic and Molecular Bases of Inherited Disease*, p. 2573, McGraw-Hill, New York, 2001; W. L. Nyhan and P. Ozand, *Atlas of Metabolic Diseases*, p. 376, Chapman & Hall-Arnold, London, 1998.

Uridine diphosphoglucose (UDPG)

A compound in which α-glucopyranose is esterified, at carbon atom 1, with the terminal phosphate group of uridine-5′-pyrophosphate (that is, uridine diphosphate, UDP). On very mild acid hydrolysis, glucose and UDP are liberated. Uridine diphosphoglucose occurs in animal, plant, and microbial cells and is synthesized enzymatically from uridine triphosphate (UTP) and α-glucose-1-phosphate by UDP-glucose pyrophosphorylase, as shown in reaction (1).

$$\text{Glucose-1-phosphate} + \text{UTP} \rightarrow \text{UDP-glucose} + \text{pyrophosphate} \quad (1)$$

This compound functions as a key in the transformation of glucose to other sugars. It can, for example, be epimerized to UDP-galactose, oxidized to UDP-glucuronic acid (which in turn can be decarboxylated to UDP-xylose and UDP-arabinose), or reduced to UDP-rhamnose. All of these compounds are then glycosyl donors utilized in the synthesis of simple compounds (such as morphine glucuronide, sucrose, or lactose) or complex polysaccharides (such as glycogen and cellulose). The general reaction is (2).

$$\text{UDP-X} + \text{ROH} \rightarrow \text{R-X} + \text{UDP} \quad (2)$$

The interconversion of UDP-glucose and UDP-galactose is also important in the utilization of free galactose by cells. Galactose is first phosphorylated to form galactose-1-phosphate. It then reacts with UDP-glucose to yield UDP-galactose and glucose-1-phosphate, and the UDP-galactose is finally epimerized to UDP-glucose. Since the UDP-glucose is regenerated in the last step, it is a cofactor in the sequence and is not used up. The overall reaction is (3).

$$\text{Galactose} + \text{ATP} \rightarrow \text{glucose-1-phosphate} + \text{ADP} \quad (3)$$

UDP-glucose was the earliest discovered of a class of nucleoside diphosphosugar compounds. Other transformations and syntheses occur with glucose attached to other nucleotides. Also, several sugars other than glucose may be initially attached to nucleotides.

For example, cytidine diphosphoglucose (CDP) and thymidine diphosphoglucose (TDP) can be

transformed to CDP-3,6-dideoxyhexoses or TDP-L-rhamnose, which then function as glycosyl donors. Adenosine diphosphoglucose and guanosine diphosphoglucose are primary glycosyl donors; the former is the intermediate in starch biosynthesis in plants, playing a role analogous to that of UDP-glucose in glucogen synthesis in animals.

UDP-acetylglucosamine is formed from acetylglucosamine-1 phosphate and UTP, and guanosine diphosphomannose is formed from mannose-1 phosphate and GTP. The former can be epimerized to UDP-acetylgalactosamine or transformed to complex peptide derivatives of UDP-acetylmuramic acid (compounds involved in the synthesis of bacterial cell walls and the mechanism of action of penicillin). GDP-mannose can also undergo a number of transformations (for example, to GDP-mannuronic acid) or can function directly as a glycosyl donor.

In general, then, UDP-glucose is a prominent member of a family of compounds composed of sugars activated as derivatives of nucleoside diphosphates. In these active forms they can be used directly as glycosyl donors, or they may be transformed to other more complex sugars before their utilization for the biosynthesis of both simple saccharides and complex polysaccharides. *See* CARBOHYDRATE METABOLISM; NUCLEIC ACID. Jack L. Strominger

Bibliography. S. J. Baum, *Introduction to Organic and Biological Chemistry*, 1993; D. M. Greenberg (ed.), *Metabolic Pathways*, vol. 1, 3d ed., 1967; L. Stryer, *Biochemistry*, 4th ed., 1995.

Urinalysis

Laboratory examination of urine. Urine is a filtrate of the blood and is produced in the kidneys. It is a reflection of the metabolic activity of the body; conditions that affect the normal homeostatic mechanisms are often revealed by a careful analysis of the composition of the urine. *See* KIDNEY; KIDNEY DISORDERS.

Modern routine urinalysis can be divided into two basic procedures: macroscopic and chemical examination, and microscopic analysis.

Macroscopic and chemical examination. This has two separate segments, with most of the chemical examination carried out by using dipstick testing.

Macroscopic examination. This segment includes noting the color and clarity of the urine. Normal urine is pale yellow or straw colored and is usually transparent or clear; abnormal urine may vary greatly in color and may show varying degrees of cloudiness. In addition, any abnormal odor of the urine is noted. Certain drugs and infectious agents make products with distinct odors; for example, diabetic urine often has a "fruity" odor. *See* DIABETES.

The specific gravity of urine, that is, the ratio of the weight of a volume of urine to that of the same volume of water, is measured routinely. Since urine contains the products of renal metabolism, such as salt and nitrogenous wastes, it is heavier than water. Specific gravity is an indicator of the kidney's ability to concentrate or dilute the urine and thus of renal tubular function. When the kidneys are functioning poorly, the specific gravity may be fixed at 1.010 (normal range is 1.003–1.032) because the renal tubules are unable to adequately concentrate or dilute the urine.

TABLE 1. Values for a normal urine sample as measured by dipstick testing

Test or constituent	Presence or concentration
Volume	1000–1500 ml/day
Appearance	Clear; straw-colored
pH	Range 4.5–8; usually acid; diet-dependent
Protein	2–8 mg/dl per urine sample [less than 150 (mg/dl)/day]
Homoglobin (blood)	Not normally detectable
Glucose (sugar)	0–20 mg/dl
Ketone bodies	0–2 mg/dl [less than 20 (mg/dl)/day]
Bilirubin	0.02 mg/dl
Urobilinogen	0.5–4 mg/day
Nitrite (bacterial metabolic product)	Not normally detectable
Leukocyte esterase (product of neutrophils)	Not normally detectable
Specific gravity	1.003–1.032

Chemical testing. Most routine chemical urinalysis is now carried out by dipstick testing, which involves the use of plastic strips, or dipsticks, bearing pads embedded with chemical reactants and color indicators. When the strips are dipped into urine, the chemicals in the pads react with it and change color. The reaction of each pad represents a separate chemical test for a specific product in the urine. The detection of chemical abnormalities in the urine often plays a vital role in diagnosis and therapy since it provides a simple and noninvasive means of assessing renal and general metabolic function (**Table 1**). Dipstick testing includes the following categories:

1. Protein. Protein, predominantly albumin, is normally present in urine but only in small amounts (usually less than 150 mg per day). Increased protein excretion (proteinuria) is sometimes associated with heart or kidney failure or infectious diseases.

2. Glucose. Reducing sugars, of which glucose is the most important, are generally not found in normal urine. Glucose in the urine (glycosuria) is usually associated with diabetes mellitus.

3. Ketone bodies. Ketone bodies, in the form of acetone, acetoacetic acid, and beta-hydroxybutyric acid, are present in normal urine in very small quantities. Elevation of these substances in the blood (ketosis) or in the urine (ketonuria) is usually related to abnormalities in carbohydrate metabolism, for example, diabetes mellitus.

4. Blood. Hemoglobin, a component of red blood cells, is normally found in minute quantities in the urine. Excessive hemoglobin in the urine (hemoglobinuria) is a pathologic finding, since it means that blood cells are leaking out of the bloodstream.

5. Bile. Small amounts of bilirubin normally can be found in urine. Larger amounts occur in jaundice

that is produced by biliary tract obstruction. *See* PROTEIN.

Microscopic examination. The urine normally contains a wide variety of formed elements that can be identified by using a light microscope. Together these elements form the urinary sediment.

Cells. Cells in the urine originate in the bloodstream or in the epithelium lining the urinary tract. The main types of epithelial cells are renal tubular cells, transitional (urothelial) cells, and squamous cells. All three types occur in relatively small numbers in the normal sediment. In disease states that affect the nephron, increased numbers of renal tubular cells may be present in the urine. Transitional cells may show an increase after urethral catheterization or prostate surgery, as well as in the presence of bladder infections or bladder cancer. Increased numbers of squamous cells in the urine may be a urinary contaminant from the vaginal epithelium.

Blood cells occur normally in the urine in small numbers, and consist mainly of polymorphonuclear neutrophils (a type of granular leukocyte) and red blood cells. A high level of polymorphonuclear neutrophils may indicate an infection of the urinary tract. Increased red blood cells may indicate a "silent" tumor of the kidney or bladder or, when accompanied by casts, may indicate a disease specifically affecting the renal glomeruli.

Casts. These proteinaceous products of the kidney are of major importance when present in increased numbers or seen in abnormal forms, because they usually indicate intrinsic renal disease. They are formed only in the kidney—in the lumen of the distal nephron—and consist mainly of Tamm-Horsfall protein, which is unique to the kidney. Casts are cylindrical and are named on the basis of their microscopic appearance and the cells they contain (**Table 2**).

TABLE 2. Types and characteristics of normal and abnormal urinary casts

Type of cast	Characteristics
Associated with normal urine	
Hyaline	Smooth surface; low refractive index; hard to see without staining
Granular	Fine granules scattered throughout
Associated with abnormal urine	
Granular	Coarse granules scattered throughout
Waxy	High refractive index; easy to see; surface looks waxy; notched sides; ends squared off (look broken off)
Fatty	Cast surface covered with high refractile lipid globules of varying size
Pigment	Basically a hyaline cast that is colored because of absorption of an abnormal pigment such as hemoglobin or melanin
Red blood cell	Surface covered with red blood cells
White blood cell	Surface covered with polymorphonuclear neutrophils
Renal tubular epithelial	Surface covered with renal tubular epithelial cells
Bacterial/fungal	Surface admixed with polymorphonuclear neutrophils and bacteria or fungi

Microorganisms. Normally, urine is sterile, and the urinary sediment should not contain microorganisms. However, the terminal third of the urethra and the vagina can be inhabited by bacteria and yeast forms, and these may be found in the sediment in small numbers in freshly voided urine. In patients with serious urinary tract infections, microorganisms are usually present in the urine in considerably greater numbers.

Parasites are not normally found in the urine. They occur rarely in the United States but are present more frequently in other parts of the world. *Schistosoma hematobium* is the only human parasite indigenous to the urinary tract; other parasites can appear in the urine as contaminants from the vagina or the perirectal area. Lice and mites from an individual can fall into urine during voiding.

Other elements. Mucus is frequently found in urine sediment and has no known pathologic significance. A wide variety of crystals appear in the urine; their presence may be normal or may indicate an abnormal state. In addition, crystalline forms of certain medications, such as sulfa drugs, are often seen in urine sediment. Crystal identification is performed with the aid of polarized microscopy, because many crystals in the urine are birefringent. *See* POLARIZED LIGHT MICROSCOPE.

Additional tests. Other nonroutine tests may be performed on urine. Renal function can be tested, and renal clearance can be measured. Also, tests can be carried out for the presence of poisons, drugs, and abnormal metabolites, and for evidence of pregnancy. *See* CLINICAL PATHOLOGY. Meryl H. Haber

Bibliography. J. A. Freeman and M. F. Beeler, *Laboratory Medicine*: *Urinalysis and Medical Microscopy,* 2d ed., 1983; M. H. Haber, *The Urinary Sediment: A Textbook Atlas*, 1981; T. A. Stamey and R. W. Kindrachuk, *Urinary Sediment and Urinalysis: A Practical Guide for the Health Science Professional,* 1985; S. K. Strasinger, *Urinalysis and Body Fluids*, 4th ed., 2001.

Urinary bladder

A distensible, muscular sac in most vertebrates which serves as a reservoir for urine. Snakes, crocodilians, birds (with the exception of the ostrich), most lizards, and a few fish lack a urinary bladder. In these organisms, urine empties directly into the cloaca. The development of the urinary system is intimately associated with the development of the reproductive system. Three general types of urinary bladder are recognized among the vertebrates: tubal, cloacal, and allantoic. *See* URINE.

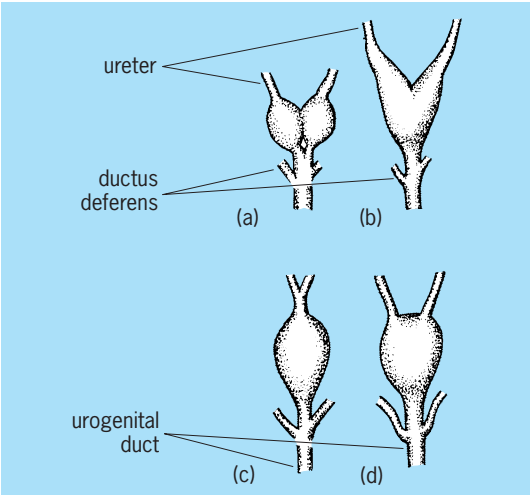

Types of tubal bladders. (*a*) Duplex. (*b*) Bilobed. (*c*) Simplex, with ureters united. (*d*) Simplex, with separate ureters. (*After H. E. Walter and L. P. Sayles, Biology of the Vertebrates, 3d ed., Macmillan, 1949*)

Tubal bladders. Most fish possess tubal bladders, that is, enlargements of the mesonephric ducts (see **illus.**). A single bladder, simplex type, results when two excretory ducts unite to form a single expanded structure. In the cod (*Gadus*) two separate bladders, duplex type, occur near the terminus of each excretory duct. These structures unite posteriorly to form a single passageway. A third type of tubal bladder is the bilobed bladder which occurs in gars as the result of fusion of the left and right urinary ducts.

Cloacal bladders. The cloacal bladder is found in monotremes, amphibians, and some dipnoans. There is no direct connection between the excretory ducts and this type of bladder. The bladder is an outpouching or diverticulum of the cloacal wall. The cloacal opening is closed by a sphincter muscle and the urine which seeps into the cloaca from the excretory ducts is forced into the bilobed bladder.

Allantoic bladders. This type of bladder is derived from the ventral wall of the cloaca and possibly the allantoic diverticulum. The role of the allantois in the formation of this type of bladder, which is found in most mammals, the turtles, and those lizards which have a bladder, is questioned by some embryologists. *See* ALLANTOIS.

The mammalian bladder is lined with a special epithelium composed of transitional cells. The muscular layer is composed of vertical, horizontal, and oblique fibers. The bladder drains through the urethra, the opening being controlled by a sphincter. Innervation is by the hypogastric sympathetic plexus and partly by parasympathetic fibers from the second and third sacral nerves. Stimulation of the parasympathetic causes the bladder muscle to contract and relaxes the internal sphincter. Micturition is a reflex act which is initiated voluntarily except in children. *See* PARASYMPATHETIC NERVOUS SYSTEM; SYMPATHETIC NERVOUS SYSTEM; URINARY SYSTEM. Charles B. Curtin

Urinary system

The system consisting of the kidneys, urinary ducts, and bladder. These structures will be discussed in this article in terms of their comparative anatomy, embryology, and physiology.

Comparative Anatomy

Similarities are not particularly evident among the many and varied types of excretory organs found among vertebrates. The variations that are encountered are undoubtedly related to problems with which vertebrates have had to cope in adapting to different environmental conditions.

Archinephros. It is now generally believed that the primitive vertebrate ancestor possessed an excretory organ referred to as an archinephros or holonephros (**Fig. 1**). This probably consisted of a pair of dorsally located ducts extending the length of the body cavity. Each duct was joined by a series of segmentally

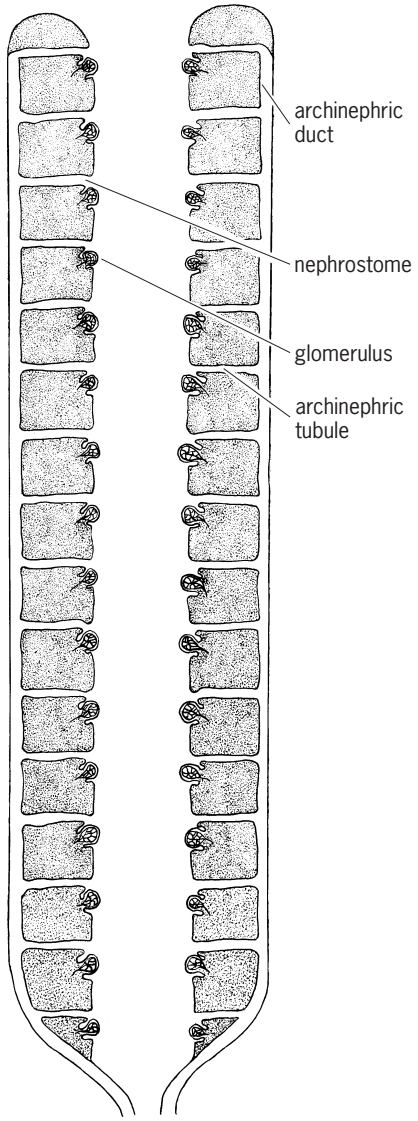

Fig. 1. Hypothetical structure of archinephros. (*After C. K. Weichert, Elements of Chordate Anatomy, 3d ed., McGraw-Hill, 1967*)

arranged tubules, one pair to each segment. The other end of each tubule opened into the body cavity by a ciliated, funnel-shaped aperture. Close to each opening was a small knot of arterial blood vessels called an external glomerulus. From this type of kidney with its archinephric duct the various kidneys of forms living today may originally have been derived. The larval form of the hagfish and the larvae of certain amphibians, the caecilians, are present-day vertebrates possessing kidneys of this type.

Anamniote kidneys. The anterior portion of the archinephric kidney persists only in the adult stage of the hagfish and of certain teleost fishes in which it is called the head kidney, or pronephros. It appears in the embryos of most vertebrates as a transitory structure that usually degenerates soon after it has formed. The remainder of the kidney posterior to the pronephros is known as the opisthonephros.

Pronephros. The importance of the pronephros lies mainly in the part it plays during development in forming the archinephric duct which persists even though the pronephros appears only as a transient structure. In some larval forms the pronephros may be important in getting rid of wastes at a time when the opisthonephros is being formed. Even in the hagfish, the head kidney has become modified from the primitive condition. Here the openings of the tubules connect with the pericardial cavity, and the fluid drained by the tubules passes into a nearby vein instead of entering the archinephric duct. Pronephros and opisthonephros become completely separated by degeneration of the portion between them.

Opisthonephros. Because the pronephros is usually a transitory structure, the opisthonephros is the more important of the two. It serves as the adult kidney in lampreys, most fishes, and amphibians. The opisthonephros differs from the pronephros in several respects. A main distinction is that the segmental arrangement of the kidney tubules is lost and many tubules may lie within the confines of a single segment. Furthermore the connection of the opisthonephric tubules with the body cavity is usually lost, and renal corpuscles with internal glomeruli are typically present. These are small knots of arterial vessels, each surrounded by a double-walled cup called Bowman's capsule; the two together form a renal corpuscle (**Fig. 2**). The internal, or visceral, layer of Bowman's capsule is actually very complexly arranged, folded, and interdigitated to follow the configurations of the individual glomerular capillaries which exhibit extensive anastomoses. The cells, called podocytes, are not squamous in type and do not furnish a continuous covering of the capillaries.

In some forms the anterior tubules of the opisthonephros lie in the same segments as posterior pronephric tubules. This indicates the transitional nature of the two. A typical opisthonephric tubule consists of a narrow neck adjacent to the renal corpuscle, followed in turn by secretory and collecting portions. The collecting portion joins the archinephric duct. The ends of several collecting tubules may unite to form a common duct which either opens into the archinephric duct or establishes an independent connection with the cloaca. Several such accessory ducts may be present.

Cyclostomes. The opisthonephros of the adult hagfish differs basically from the original archinephros only in the loss of peritoneal connections of the posterior tubules. In the adult lamprey the opisthonephros on each side consists of a long, strap-shaped body without peritoneal connections. The kidneys lie on either side of the middorsal line from which each is suspended by a mesenterylike membrane. The archinephric duct lies along the free edge of the kidney. The ducts of the two sides unite posteriorly to open into a urogenital sinus which leads to the outside through an aperture at the tip of a small urogenital papilla. Two slitlike genital pores connect the urogenital sinus with the coelom. The condition is similar in both sexes. Eggs or spermatozoa leave the body cavity by way of the genital pores, the urogenital sinus, and the urogenital aperture. Only here are the reproductive and urinary systems associated.

Fishes. There is much variation in shape of the opisthonephric kidneys of fishes, but they are fundamentally similar in structure. In some they extend the length of the coelom; in others they are short and may show various degrees of fusion. Peritoneal funnels rarely occur. Some marine teleosts lack glomeruli and thus possess aglomerular kidneys. In elasmobranchs the anterior ends of the kidneys of the male have been appropriated by the reproductive system. Modified kidney tubules, called efferent ductules, connect each testis with the archinephric duct which lies on the ventral surface of the kidney and serves as a ductus deferens for sperm transport. Accessory urinary ducts are usually present. In teleost fishes there is no connection between the testes and the opisthonephric kidneys. The posterior ends of the archinephric ducts of female fishes

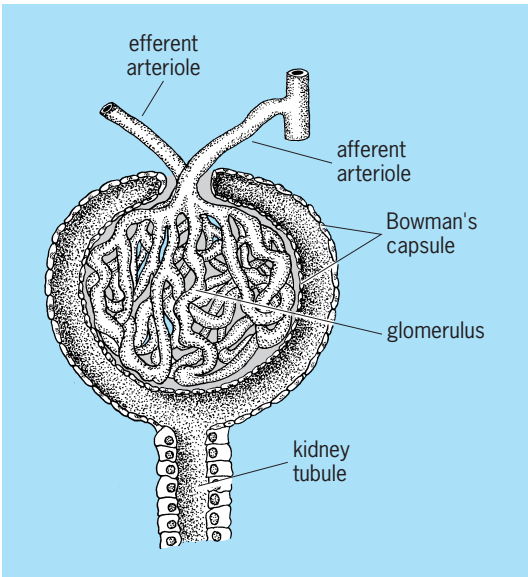

Fig. 2. Renal corpuscle. (***After C. K. Weichert, Anatomy of the Chordates, 4th ed., McGraw-Hill, 1970***)

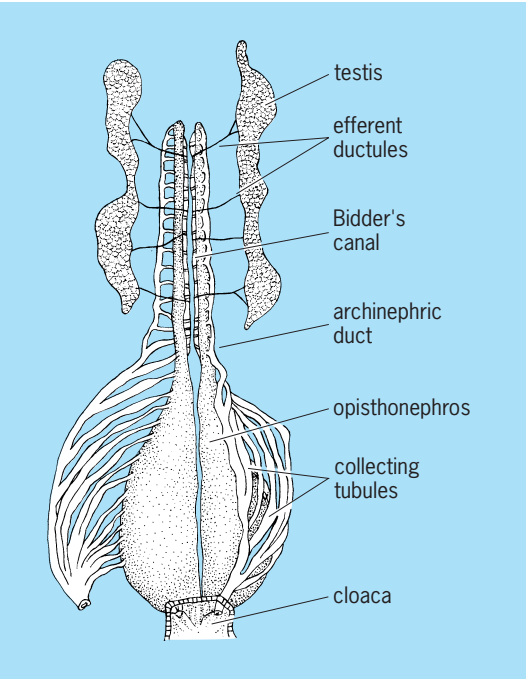

Fig. 3. Urogenital organs of male salamander, ventral view. The collecting ducts on the left are shown detached from the cloaca and spread out for clarity. (*After C. K. Weichert, Elements of Chordate Anatomy, 3d ed., McGraw-Hill, 1967*)

enter a common urinary sinus inside a small urinary papilla. The latter enters the cloaca in elasmobranchs and dipnoans, but in most other fishes it opens directly to the outside, a cloaca being absent.

Amphibians. In common with other amphibians, adult caecilians possess an opisthonephros. The kidneys of the tailed amphibians are much like those of elasmobranch fishes, the anterior ends in males being concerned with genital rather than urinary functions (**Fig. 3**). The archinephric duct courses along the lateral edge of the kidney a short distance outside the kidney proper. Numerous collecting ducts or tubules leave the kidney at intervals to join the duct. The two ducts in both sexes open separately into the cloaca. In frogs and toads the opisthonephric kidneys lie toward the posterior part of the abdominal cavity. A yellowish adrenal gland is located along the ventral surface of each. The kidneys of females are not related to the reproductive system but in males an intimate connection exists. Modified tubules, or efferent ductules, from the testes connect with kidney tubules which lead to the archinephric duct. This duct is located within the kidney along its lateral margin. Peritoneal funnels are present on the ventral sides of the kidneys in some frogs but they connect with veins rather than the kidney tubules. A thin-walled urinary bladder opens into the amphibian cloaca. It has no connection with the archinephric ducts.

Amniote kidneys. In reptiles, birds, and mammals three types of kidneys are usually recognized: the pronephros, mesonephros, and metanephros. These appear in succession during embryonic development, but only the metanephros persists in the adult. Mesonephros and metanephros actually represent different levels of the opisthonephros of lower forms, the metanephros being equivalent to the posterior portion.

The anteriorly located pronephros appears during very early development, but it soon degenerates and the more posterior mesonephros then develops. The duct of the pronephros, however, persists to become the duct of the mesonephros. This is actually the same as the archinephric duct but is usually referred to as the Wolffian duct.

The mesonephros persists for a time and then degenerates. In the meantime the metanephros has begun to develop from the region posterior to the mesonephros. A few mesonephric tubules and the Wolffian duct persist to contribute to the reproductive system of the male or to remain as vestigial structures.

Mesonephros. The tubules of the mesonephros develop in the same manner as opisthonephric tubules. Some of the anterior tubules may even form peritoneal connections. In some forms the mesonephric kidneys become voluminous structures; in others they amount to very little. In reptiles, spiny anteaters, and marsupials the mesonephros may even persist for a time after birth.

Metanephros. A metanephric kidney develops on each side posterior to the mesonephros. It is composed of essentially the same parts as the mesonephros and contains renal corpuscles, secretory tubules, and collecting tubules. No peritoneal connections are present. Each kidney has a twofold origin. An outgrowth from the posterior end of the Wolffian duct grows forward into the tissue posterior to that from which the mesonephros was derived and which is called the metanephric blastema. The outgrowth is destined to form the ureter and the collecting portion of the kidney. It branches and rebranches many times to form ultimately large numbers of fine collecting tubules. At least in mammals at the point where the outgrowth undergoes its primary divisions an expanded region forms the pelvis of the kidney.

From the blastema adjacent to the collecting tubules arise secretory tubules. Each tubule grows and becomes S-shaped. One end establishes connection with a collecting tubule; the other expands and becomes invaded by a vascular glomerular tuft so that a typical renal corpuscle is formed. Each tubule differentiates into several regions (**Fig. 4**).

Reptiles. The metanephric kidneys of reptiles lie in the posterior part of the abdominal cavity, usually in the pelvic region. They are small, compact, and often markedly lobulated. The posterior portion on each side is somewhat narrower. In some lizards the hind parts may even fuse. The degree of symmetry varies, being most divergent in snakes and limbless lizards which have notably long, narrow, lobulated kidneys in correlation with the shape of the body. One kidney may be entirely behind the other.

Snakes and crocodilians lack a urinary bladder, but most lizards and turtles have well-developed bladders which open into the cloaca. Except in turtles

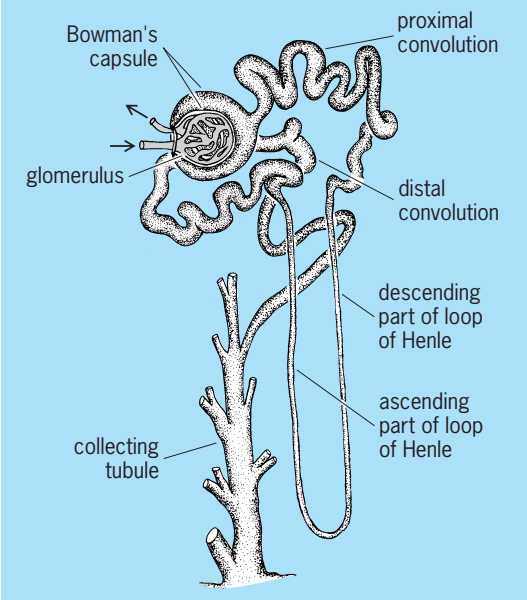

Fig. 4. Mammalian metanephric tubule, showing the renal corpuscle and secretory and collecting portions. (*After C. K. Weichert, Elements of Chordate Anatomy, 3d ed., McGraw-Hill, 1967*)

the ureters open independently into the cloaca. In turtles they connect to the bladder. Some turtles possess a pair of accessory urinary bladders which open into the cloaca. They may be used as accessory organs of respiration. In females they are reported to be filled with water which is used to soften the soil when a nest is being prepared.

Birds. The kidneys of birds are situated in the pelvic region of the body cavity; their posterior ends are usually joined. They are lobulated structures with short ureters which open independently into the cloaca.

Except for the ostrich, birds lack urinary bladders. Urinary wastes, chiefly in the form of semisolid uric acid, are eliminated through the cloaca along with the feces.

Mammals. The rather typical mammalian metanephric kidney (**Fig. 5**) is a compact, bean-shaped organ attached to the dorsal body wall outside the peritoneum. The ureter leaves the medial side at a depression, the hilum. At this point a renal vein also leaves the kidney and a renal artery and nerves enter it. The kidney is surrounded by a capsule of connective tissue under which lies the cortex. The renal corpuscles and the greater part of the secretory tubules lie entirely in the cortex. The portion of the kidney surrounded by the cortex is the medulla. It is partly composed of large areas, the renal pyramids. The outer borders of the pyramids are divided into smaller units called lobules. The collecting tubules lie within the pyramids but may extend well up into the cortex. The inner portion of each pyramid, in the form of a blunt papilla, projects into an outpocketing of the pelvis known as a minor calyx. Several minor calyces join together to enter major calyces which in turn open into the renal pelvis. The pelvis leads to the ureter which empties into the bladder, except in monotremes in which it enters the urethra. Urine, which is stored temporarily in the bladder, passes to the outside through the urethra. In males the urethra opens at the tip of the penis; in females the condition varies, for in some, as the rat and mouse, the urethra opens independently to the outside, passing through the clitoris. It usually, however, enters a vestibule which is the terminal part of the genital tract. The kidneys of mammals are markedly lobulated in the embryo, and in many forms this condition is retained throughout life. *See* KIDNEY. Charles K. Weichert

Comparative Embryology

The kidney, or nephros, of vertebrates is made up of many individual structural and functional units known as nephrons. The nephrons are derived from that portion of the embryonic mesoderm designated as the intermediate mesoderm. Ideally, this material becomes segmented, with each segment being termed a nephrotome (**Fig. 6**). Any given nephrotome contains a coelomic chamber, the nephrocoele, which opens to the adjacent body cavity via a peritoneal funnel.

The conversion of nephrotome to nephron involves the following events. There arises from the dorsolateral wall of the nephrotome a tubular outgrowth, the principal tubule, which communicates with the nephrocoele via a nephrostome; the medial wall of the nephrotome thins, flattens, and bulges inwardly coincidently with its invasion by arterial capillaries derived from the nearby dorsal aorta. The skein of capillaries makes up a glomerulus, and the wall of the nephrotome investing the glomerulus is known as a renal, or Bowman's, capsule (**Fig. 7**).

The original vertebrate nephros presumably consisted of similar nephrons throughout the length of the organ with each opening independently to the

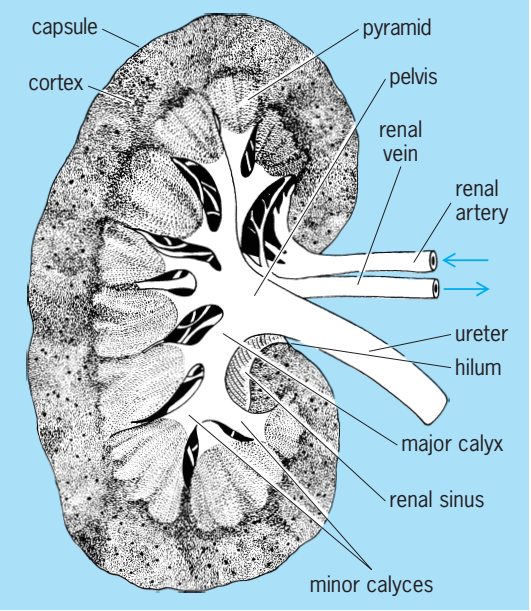

Fig. 5. Sagittal section of metanephric kidney of a human (semidiagrammatic). (*After C. K. Weichert, Elements of Chordate Anatomy, 3d ed., McGraw-Hill, 1967*)

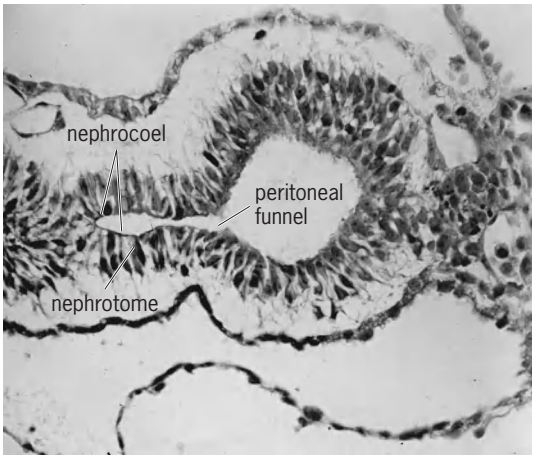

Fig. 6. Nephrotome of human embryo.

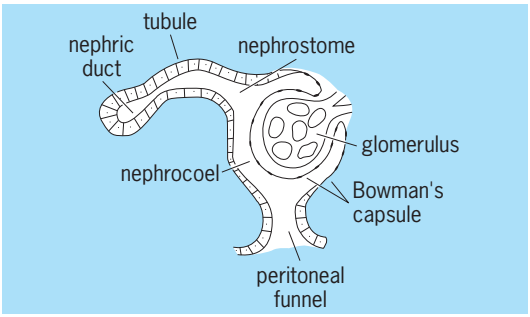

Fig. 7. Basic design of a nephron.

exterior. However, this arrangement is only hypothetical, for in all known vertebrates the nephrons empty into a common drainage, or nephric, duct which passes back alongside the nephros to terminate in the cloaca, the chamber which also receives the digestive tract. The situation is also complicated in present-day vertebrates by four variables. First, in the embryos of higher vertebrates, typical hollow nephrotomes are seldom formed; instead, nephrons differentiate without segmental arrangement within a continuous cord, the nephrogenic cord, of intermediate mesoderm. Second, as the nephros develops embryonically, its entire length does not appear at one time; instead, the nephrons appear in sequence from front to rear and the first-formed anterior ones tend to disappear before the posterior ones arise. Third, the nephrons become progressively more complex from anterior to posterior. Fourth, the manner of establishment of the nephric duct is not consistent in all vertebrates. The results of these four variables follow.

Nephros of fishes and amphibians. Embryonic development of the nephros is inaugurated within the most anterior reaches of the mesomere. This intermediate mesoderm becomes segmented into nephrotomes from each of which a nephron forms in the general manner already described. The number of these first formed nephrons varies with the species, but is always relatively small, usually three to five. These nephrons, because of their anterior position and because they are the first to appear, form the head kidney, or pronephros. Accordingly, the nephrons themselves are termed pronephric tubules, or pronephrons. As the first and most anterior pronephric tubule arises, it first extends dorsolaterally and then turns backward to join the one forming immediately behind. This one in turn joins the one behind it, and so on, thus producing a common drainage duct designated the pronephric duct. The pronephric duct, once initiated, extends itself backward along the still undifferentiated intermediate mesoderm until it joins the cloaca (**Fig. 8***a*).

With the exception of the hagfishes and some bony fishes in which it persists throughout life, the pronephros has a temporary existence. In the sharks, for instance, it is present only during early embryonic stages and has little or no functional significance, but in those forms having an active free-living larval stage, such as amphibians and lampreys, it persists in the larva as a functional organ and the individual pronephrons possess ciliated peritoneal funnels opening to the adjacent coelom. This provides for the uptake of certain materials directly from the coelom as well as from the bloodstream. Direct demonstrations of the functional capacities of pronephric tubules have come from a variety of tests. For example, the pronephrons of the larval lamprey will take up quantities of coloidal carbon which may be injected into the coelom; and through their ability to accumulate the dye phenol red, the pronephrons of frog larvae and certain bony fishes have also revealed their functional capabilities.

Whatever the length of its existence, the pronephros is supplemented and succeeded by a second generation of nephrons derived from the remainder of the intermediate mesoderm. Although

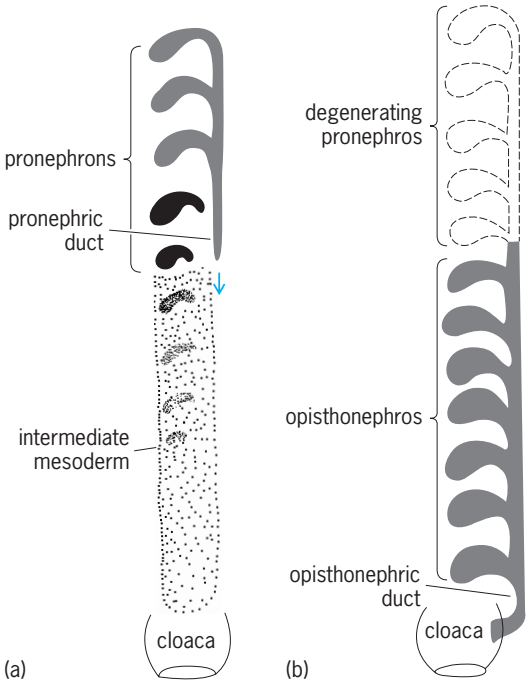

Fig. 8. Diagram of (*a*) developing pronephros and (*b*) opisthonephros.

they arise in basically the same manner and exhibit the same fundamental structure, these later-formed nephrons tend to be longer and more complex in their makeup, and ordinarily lack peritoneal funnels. Unlike their pronephric forerunners these nephrons fail to establish their own drainage duct; they join the already existing pronephric duct. Eventually, as noted, the earlier-formed pronephros disappears, leaving this later generation of nephrons to constitute the final, definitive kidney. This organ, distinctive to fishes and amphibians, is the opisthonephros, or back kidney, and the former pronephric duct which its nephrons have taken over is the opisthonephric duct (Fig. 8*b*).

Nephros of reptiles, birds, and mammals. The embryonic initiation of the nephros of higher vertebrates is customarily described as involving the establishment of a pronephros and pronephric duct as in fishes and amphibians. Although this may be true for reptiles and birds, in mammals pronephric tubules rarely if ever appear. In humans, for example, that level of intermediate mesoderm equivalent to the pronephros never gets beyond the point of conversion to a few rudimentary nephrotomes. Because definitive pronephric tubules are never provided, the pronephric duct must arise in some fashion other than by junction of the ends of pronephrons. The original nephric duct is initiated as a solid rod which splits off the dorsolateral side of the nephrogenic cord (**Fig. 9**). Once established in this fashion, the solid duct then frees itself from the parent material and as a tapering rod extends itself backward by independent terminal growth, ultimately contacting and fusing with the wall of the cloaca. The solid rod gradually hollows out and within a few days after its original establishment becomes tubular throughout.

Wolffian duct. As the nephric duct, or Wolffian duct, is being formed the first of two generations of nephrons appears. The first generation consists of a series of nephrons derived from the middle level of the intermediate mesoderm. The details of their manner of development and final form vary from one category of animals to another, but in general they conform to the pattern described above. The human embryo may be used as a specific illustration. Briefly, the nephrogenic cord paralleling the growing Wolffian duct provides an increasing number of serially arranged spherical bodies known as nephric vesicles (**Fig. 10**). These are solid at first, but become hollow, and as they do so each vesicle sends a principal tubule dorsolaterally to join the Wolffian duct. The vesicle proper will provide the capsule surrounding the later-developing glomerulus. Elongation and twisting of the principal tubule and the acquisition of a capillary network complete the nephron (**Fig. 11**).

Wolffian body. These nephrons are known as mesonephric tubules, or mesonephrons, and collectively they make up the mesonephros, or Wolffian body. This kidney is well developed in reptiles and birds, but in mammals exhibits considerable variation. In the rat embryo, for example, it is quite rudimentary and only a dozen or so abortive nephrons arise. At the other extreme is the pig embryo whose mesonephros is large and bulky and involves several hundred long and convoluted nephrons. The human mesonephros lies between these two extremes.

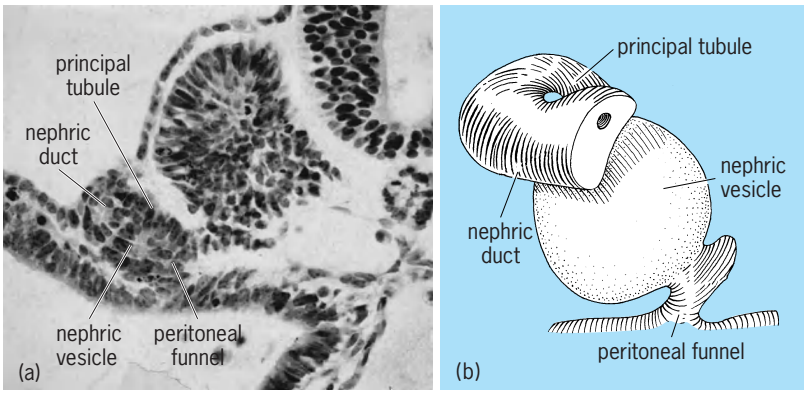

Fig. 10. Early human mesonephron. (*a*) Section. (*b*) Reconstruction.

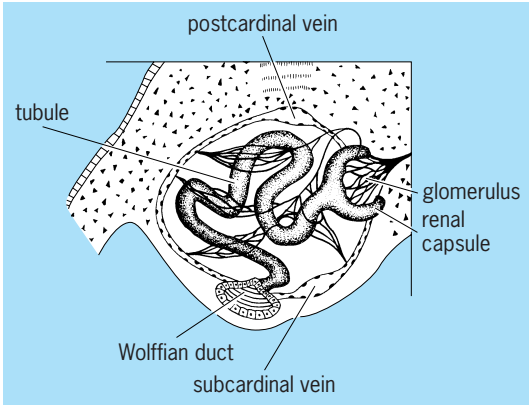

Fig. 11. Fully differentiated human mesonephron. (*After T. W. Torrey, Morphogenesis of the Vertebrates, 2d ed., John Wiley and Sons, 1968*)

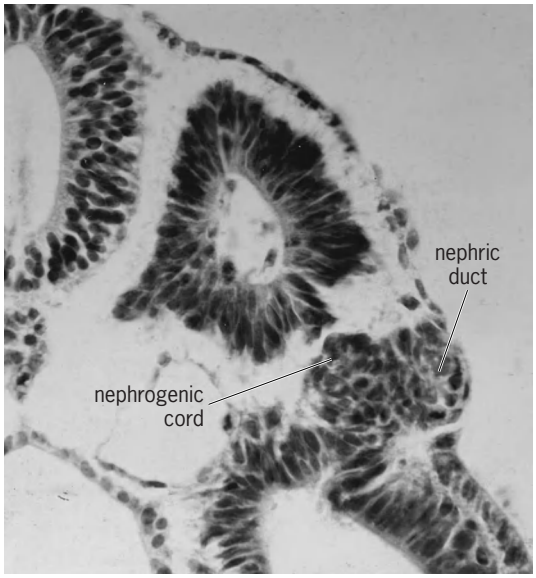

Fig. 9. Origin of nephric duct in human embryo.

Embryonic function. The variable status of the mesonephros, especially in mammals, raises the interesting question of the functional role it plays in the economy of the embryo. It is reasonable, of course, to infer function from exhibited structure. More convincing, however, are the direct experimental demonstrations of functional capacities of the mesonephrons. To this end, certain techniques employed to assess kidney functions in adult forms have been applied with profit.

The chick embryo has been a common subject for such tests. Solutions of indigo red and trypan blue injected into the vascular system ultimately appear in the mesonephrons. Another approach has been that of cultivating a fragment of mesonephros in a suitable culture medium to which an indicator such as phenol red has been added. The proximal portions of the tubules pick up the indicator and transport it to their lumina; the distal portions resorb water. Still another method has been the direct identification of nitrogenous wastes deposited in the embryonic bladder, the allantois. *See* ALLANTOIS.

Similar experiments on mammals are complicated by the intrauterine location of the embryo and its association with a placenta. Nevertheless, it has been possible either to inject suitable indicators directly into the embryonic body or to introduce them secondarily by transmission via the placenta from the maternal bloodstream. Such tests on embryos of the rabbit, cat, pig, and pouch-young opossum have provided positive demonstrations of the functioning of both the glomerular filter and the tubule proper.

Metanephros. As the mesonephros is attaining its maximum development, a second generation of nephrons is inaugurated. The history of this group is considerably more complicated than the former and runs briefly as follows. A tubular outgrowth from the Wolffian duct, the ureteric diverticulum, appears close to the entrance of the Wolffian duct into the cloaca and pushes itself anteriorly into the undifferentiated intermediate mesoderm behind the mesonephros. The distal end of this diverticulum enlarges, and the intermediate mesoderm coincidently begins to condense around the enlargement. **Figure 12***a* shows that the proximal segment of the original diverticulum is the ureter, or metanephric duct; the expanded distal end of the diverticulum is the primitive renal pelvis; the condensed mesoderm around the pelvis is the metanephric blastema.

Subsequent events pertain primarily to the pelvis and blastema. The former subdivides first to form the two future major calyces (Fig. 12*b*). These divide and subdivide until several generations of branches are produced. The earlier generations come to represent the minor calyces; the later generations become the collecting tubules (Fig. 12*c* and *d*). As the primitive pelvis carries on this program of subdivision, the blastemal tissue subdivides into a corresponding number of masses. Thus there results a nodule of blastema in association with the end of each prospective collecting tubule (**Fig. 13**). Each such nodule or sphere is the forerunner of a metanephric or uriniferous tubule. The solid sphere first becomes a vesicle which, by elongation, is transformed into a tortuous tubule (Fig. 13). The thinner-walled, blind end of the tubule becomes the capsule surrounding the glomerulus forming concomitantly. Coincidentally, the tip of the prospective collecting duct grows out to meet the end of the tubule and the two unite (Fig. 13). The uriniferous tubules, collecting tubules, and calyces are the definitive kidney or metanephros of the late embryo and adult. In terms of gross anatomy, the convoluted portions of the uriniferous tubules, with their glomerular capsules, collectively form the cortex of the kidney. These tubules lead into the collecting system and calyces making up the medulla of the kidney. All drainage ultimately converges upon the renal pelvis, which in turn leads to the ureter.

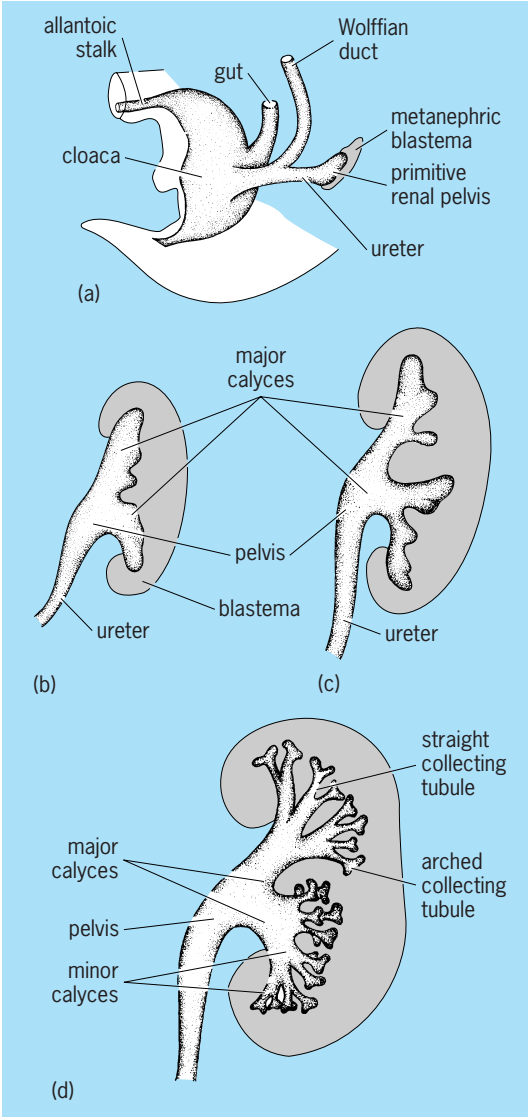

Fig. 12. Stages *a-d* in development of metanephric pelvic and blastema. (*After T. W. Torrey, Morphogenesis of the Vertebrates, 2d ed., John Wiley and Sons, 1968*)

Developmental interdependence. Experimental analyses have revealed important developmental interdependencies within the mesonephric and metanephric

systems. The original nephric duct grows back from the level of its origin to join the cloaca. In the normal course of events it is joined along the way by mesonephrons and thus becomes the Wolffian duct. If the backward extension of the duct alongside the prospective mesonephros-forming area is prevented, little or no development of the mesonephros will occur. This can be accomplished readily in the chick embryo by producing a minute wound or inserting some block in the pathway of the duct. This is interpreted to mean that the duct serves as an inductor of the mesonephric tubules; that is, differentiation of the mesonephrons depends upon some kind of stimulus provided by the nephric duct. Another consequence of such a blockage of the nephric duct is the elimination of the ureter which normally grows from it. This in turn leads to a failure of the blastema to produce metanephric tubules. The ureter is the inductor of these tubules. Development of the nephros is thus revealed as a series of steps, each dependent upon the one before, with the original nephric duct playing the starting role.

During the latter part of embryonic life the mesonephros and metanephros function simultaneously. Gradually, however, the mesonephros regresses and the metanephros assumes full responsibility for excretion. In females, the mesonephros disappears almost entirely; in males, parts of it are incorporated in the reproductive system. As a consequence of straightening and elongation of the fetal body, the kidneys come to be displaced relatively far forward in the body. *See* REPRODUCTIVE SYSTEM.

Urinary bladder. At or near the posterior ends of the nephric ducts there frequently is a reservoir for urine. This is the urinary bladder. Actually there are two basic varieties of bladders in vertebrates. One is found in fishes in which the reservoir is no more than an enlargement of the posterior end of each urinary duct. Frequently the urinary ducts are conjoined and a small bladder is formed by expansion of the common duct. The far more common type of bladder is that exhibited by tetrapods. This is a sac which originates embryonically as an outgrowth from the ventral side of the cloaca. Present in all embryonic life, it is exhibited differentially in adults. All amphibians retain the bladder, but it is lacking in snakes, crocodilians, and a few lizards; birds, also, with one or two exceptions, lack a bladder. It is present in all mammals. Because much of the developmental history of the tetrapod bladder is linked to the history of the cloaca, it will be considered in this conjunction.

Amphibians. The basic pattern is exemplified by the Amphibia. All retain the cloaca as adults and the urinary bladder arises as a diverticulum from the floor of the cloaca (**Fig. 14**). Commonly it is bilobed. There is no direct connection between the excretory ducts and the bladder. Instead, urine first passes into the cloaca and thence into the bladder. Urine is expelled by the intermittent opening of the cloacal orifice and the coincidental contractions of the muscular wall of the bladder.

Reptiles and birds. In reptiles and birds the cloaca is partly subdivided so that the intestine and urogenital ducts open into separate compartments which then join in a common outlet. In the embryo a pouch develops from the floor of the urogenital portion of the cloaca and expands to form a prominent sac known as the allantois. Ultimately the allantois enlarges to extend beyond the confines of the embryo and to fuse broadly with the outermost membrane, the chorion, surrounding the embryo (**Fig. 15**). The combined chorioallantoic membrane ultimately makes broad contact with the membrane lining the inner surface of the shell. This fusion and spreading of the allantois serves to bring an extensive blood supply adjacent to the shell and thus sets up a medium for respiratory exchange. The allantois, therefore, serves not only as a reservoir for urine but also plays a major role in embryonic respiration. *See* FETAL MEMBRANE.

In birds the entire allantois is discarded at hatching and no adult urinary bladder is retained. This is also true for most reptiles although some, notably the turtles and certain lizards, retain the base of the allantois as an adult bladder, and the remainder atrophies.

Mammals. Among mammals, only a few primitive forms retain a cloaca as adults. In all the others it is modified in such a way as to be eliminated.

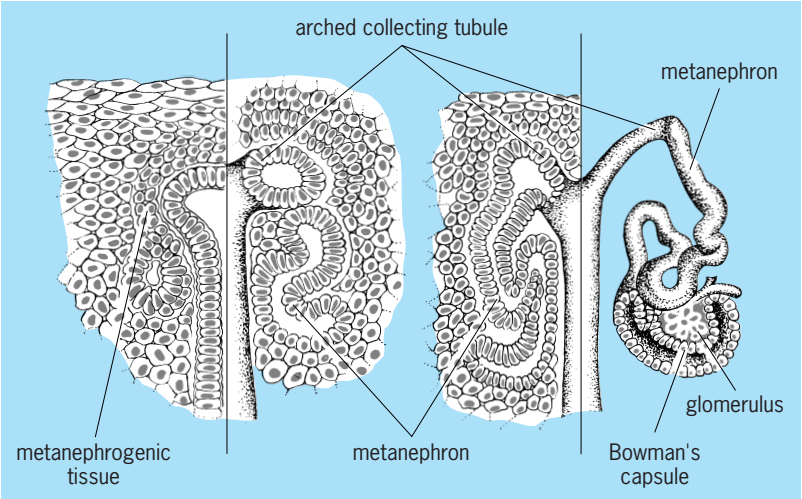

Fig. 13. Differentiating metanephron. (*After T. W. Torrey, Morphogenesis of the Vertebrates, 2d ed., John Wiley and Sons, 1968*)

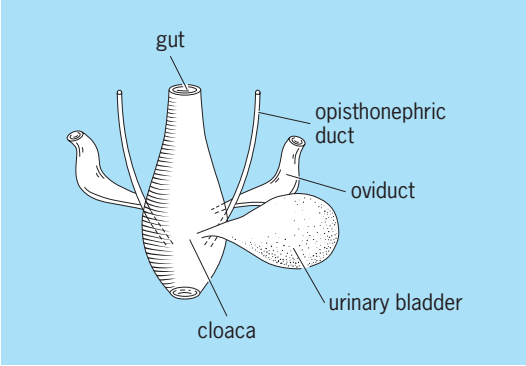

Fig. 14. Cloaca and urinary bladder of a female amphibian. (*After T. W. Torrey, Morphogenesis of the Vertebrates, 2d ed., John Wiley and Sons, 1968*)

Concomitantly, the openings of the excretory ducts are shifted and the urinary bladder is established. The pig embryo may serve to illustrate the usual events.

Almost as soon as the cloaca is established, an allantois arises from its floor. As in reptiles and birds, the allantoic sac, complete with blood vessels, expands greatly and fuses with the chorion. This association enables it to serve not only as a urinary bladder, but also to provide a major component of the placenta, so important in fetal economy. In the meantime the cloaca itself is modified. The cloaca initially ends blindly, its ventroposterior floor making contact with the embryonic skin to form the cloacal membrane (**Fig. 16***a*). A division of the cloaca into two parts, a dorsal rectum and ventral urogenital sinus, then follows. This division is effected by the urorectal fold, a crescentic fold which works backward from the angle where the allantois, excretory ducts, and gut meet until it meets the cloacal membrane (Fig. 16*b* and *c*). Not only is the cloaca itself subdivided, but the cloacal membrane is likewise divided into an anal and urethral membrane. Both these membranes eventually rupture so that the two parts of the cloaca open independently to the exterior.

The next step is the enlargement of that portion of the urogenital sinus receiving the neck of the allantois. The enlargement may involve a part of the allantoic stalk itself. This enlargement is the beginning of the definitive urinary bladder. As the bladder expands, the terminal ends of the Wolffian ducts which open to the urogenital sinus at this level are absorbed into the bladder wall. In consequence, the ureters, which stemmed initially from the Wolffian ducts, come to open directly to the growing bladder while the Wolffian ducts open somewhat behind into that part of the sinus which remains narrower and gives rise to the urethra (Fig. 16*c*). The ultimate fate of the Wolffian ducts and the final form of the urethra differ between the two sexes. The allantois per se is discarded at birth. Theodore W. Torrey

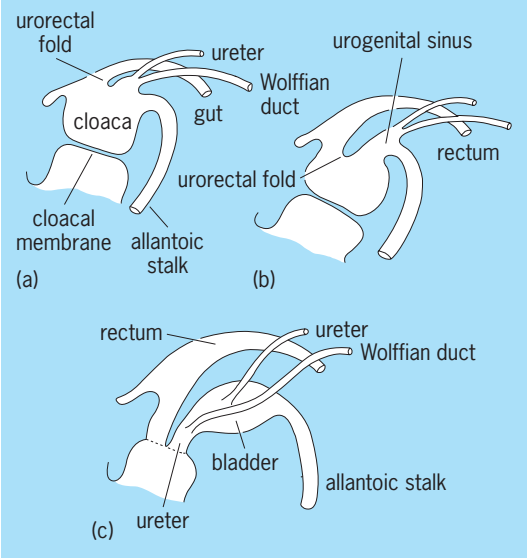

Fig. 16. Development of cloaca and urinary bladder. (*a*) Formation of cloacal membrane. (*b*) Urorectal fold formation. (*c*) Urethra formation.

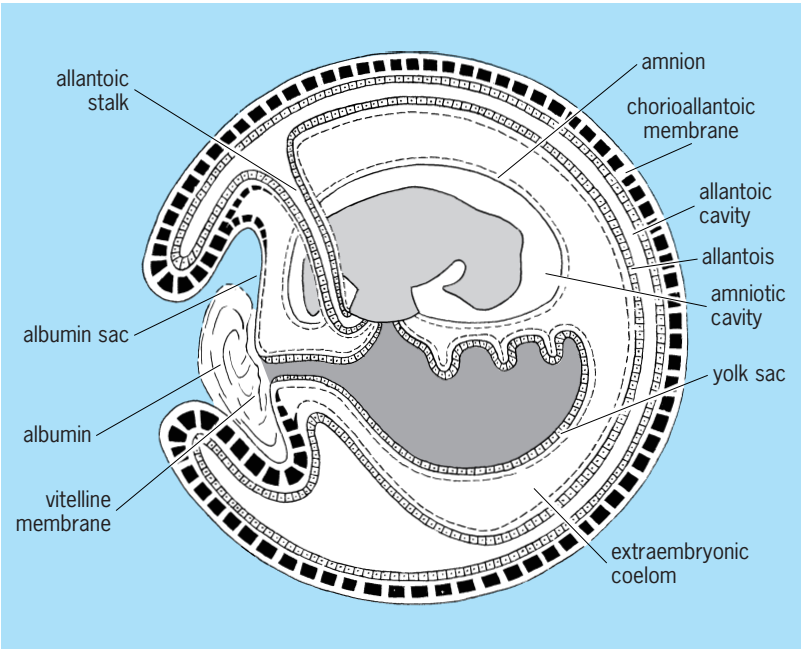

Fig. 15. Fully matured extraembryonic membranes of the chick. (*After T. W. Torrey, Morphogenesis of the Vertebrates, 2d ed., John Wiley and Sons, 1968*)

Physiology

Urine is produced by individual renal nephron units which are fundamentally similar from fish to mammals (Fig. 4); however, the basic structural and functional pattern of these nephrons varies among representatives of the vertebrate classes in accordance with changing environmental demands. Kidneys serve the general function of maintaining the chemical and physical constancy of blood and other body fluids. The most striking modifications are associated particularly with the relative amounts of water made available to the animal. Alterations in degrees of glomerular development, in the structural complexity of renal tubules, and in the architectural disposition of the various nephrons in relation to one another within the kidneys may all represent adaptations made either to conserve or eliminate water. The urinary systems of fishes, amphibians, reptiles, birds, and mammals will be compared to see how they respond to demands imposed by various fresh-water, marine, amphibious, terrestrial, and desert environments that influence their development and differentiation.

General principles. Certain basic principles underlie excretory processes in all animals, from protozoa to humans. These are regulation of volume, electrolyte balance, movement of water across cell membranes, and the elimination of nitrogenous substances.

Regulation of volume. Paleontological evidence strongly suggests that early vertebrates evolved in fresh water, or at least that progenitors of modern fishes had a long history of dwelling in fresh water. Kidneys of living vertebrates bear the imprint of

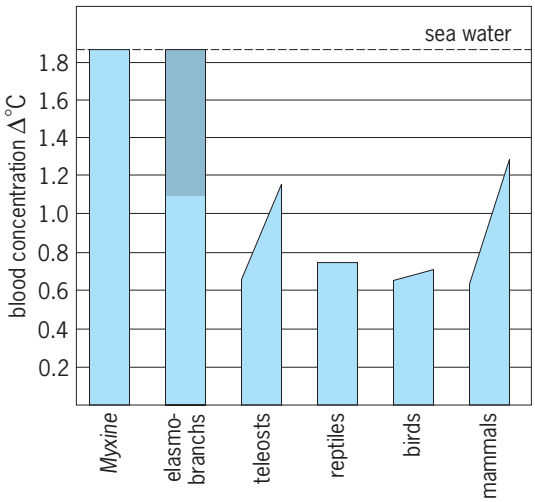

Fig. 17. Solute concentrations in blood of various marine vertebrates. Only *Myxine* and the elasmobranchs have blood isosmotic with seawater. Note the significant contributions of organic solutes, urea, and trimethylamine oxide (shaded) to the total osmotic concentration of elasmobranch blood. (*After A. P. McLockwood, Animal Body Fluids and Their Regulation, Harvard University Press, 1964*)

this early evolutionary history. Except for the primitive marine cyclostome *Myxine,* all modern vertebrates, whether marine, fresh-water, or terrestrial, have concentrations of salt in their blood only one-third or one-half that of seawater (**Fig. 17**). The early development of the glomerulus can be viewed as a device responding to the need for regulating the volume of body fluids. Hence, in a hypotonic fresh-water environment the osmotic influx of water through gills and other permeable body surfaces would be kept in balance by a simple autoregulatory system whereby a rising volume of blood results in increased hydrostatic pressure which in turn elevates the rate of glomerular filtration. Similar devices are found in fresh-water invertebrates where water may be pumped out either as the result of work done by the heart, contractile vacuoles, or cilia found in such specialized "kidneys" as flame bulbs, solenocytes, or nephridia that extract excess water from the body cavity rather than from the circulatory system. Hence, these structures which maintain a constant water content for the invertebrate animal by balancing osmotic influx with hydrostatic output have the same basic parameters as those in vertebrates that regulate the formation of lymph across the endothelial walls of capillaries.

None of the marine invertebrate phyla seem to have had a fresh-water ancestry. Members of the phylum Echinodermata, generally regarded as the closest relatives of the protovertebrate ancestors of fishes, have body fluids that are isosmotic with seawater, and they have no trace of any structure that resembles a kidney or other kind of excretory organ. *See* OSMOREGULATORY MECHANISMS.

Electrolyte balance. A system that regulates volume by producing an ultrafiltrate of blood plasma must conserve inorganic ions and other essential plasma constituents. The salt-conserving operation appears to be a primary function of the renal tubules which encapsulate the glomerulus. As the filtrate passes along their length toward the exterior, inorganic electrolytes are extracted from them through highly specific active cellular resorptive processes which restore plasma constituents to the circulatory system. The ability to form a urine hypotonic to blood is related to the length of the tubule. Animals, such as fresh-water fishes and amphibians, that produce a urine almost as dilute as distilled water have very long tubules with specially developed distal segments that are impermeable to water while they extract and restore to the blood almost every trace of salt from tubular urine before it enters the bladder. Proximal segments can also actively resorb salt, but their cells are freely permeable to water so that proximal tubular fluid is always isosmotic with respect to blood. Urine of marine fishes is always approximately isotonic with blood, and their renal tubules are short, with the specialized distal segment absent. Proximal tubule cells of all vertebrates also actively resorb glucose, amino acids, and other organic compounds from filtrate. Additionally they actively secrete foreign substances from blood into proximal tubular fluid.

Movement of water. Concentration gradients of water are attained across cells of renal tubules by water following the active movement of salt or other solute. Where water is free to follow the active resorption of sodium and covering anions, as in the proximal tubule, an isosmotic condition prevails. Where water is not free to follow salt, as in the distal segment in the absence of antidiuretic hormone, a hypotonic tubular fluid results. A special case is examined below of hypertonic urine formed when water is free to diffuse into a solution of high salt concentration surrounding mammalian collecting ducts when the cells are not correspondingly permeable to solute.

Nitrogenous end products. Of the major categories of organic foodstuffs, end products of carbohydrate and lipid metabolism are easily eliminated mainly in the form of carbon dioxide and water. Proteins, however, are more difficult to eliminate because the primary derivative of their metabolism, ammonia, is a relatively toxic compound. For animals living in an aquatic environment ammonia can be eliminated rapidly by simple diffusion through the gills. It is formed by relatively simple biochemical processes of deamination and transamination, which do not require the expenditure of free energy. However, when ammonia is not free to diffuse into an effectively limitless aquatic environment, its toxicity presents a problem, particularly to embryos of terrestrial forms that develop wholly within tightly encapsulated eggshells or cases. For these forms the detoxication of ammonia is an indispensable requirement for survival. During evolution of the vertebrates two energy-dependent biosynthetic pathways arose which incorporated potentially toxic ammonia into urea and uric acid molecules, respectively. Both of these compounds are relatively harmless, even in high concentrations, but the former needs a relatively large amount of water to ensure its elimination,

and uric acid requires a specific energy-demanding tubular secretory process to ensure its efficient excretion. *See* EXCRETION.

Methodology. The comparative physiological approach can be justified solely for information about the nature of the human's world, but it often contributes also to the solution of practical problems. For example, studies on the peculiar aglomerular kidney of the American anglerfish or goosefish (*Lophius*) led directly to the development of clinical renal clearance tests; micropuncture techniques that took advantage of the unusually large nephrons of the mudpuppy (*Necturus*) proved the hypothesis of glomerular filtration and identified specific sites along the renal tubule as loci of discrete secretory and resorptive processes; and observations on the structural and functional modifications in kidneys of the desert rat and other mammalian forms of arid regions showed how the countercurrent multiplier system operates in the renal medulla to concentrate urine. These examples show where an exploitation of experimental variables provided by nature led directly to new methods or information of general significance in renal physiology and pathology.

Rate of glomerular filtration. Following the disclosure that certain carbohydrates could not be excreted by aglomerular kidneys, a search was then undertaken to find some freely filterable nonmetabolizable carbohydrate which was also not resorbed by the renal tubules of glomerular kidneys. This led eventually to employment of the polysaccharide inulin in an application of the Fick principle to measure the rate of glomerular filtration. Knowing the total number of milligrams of inulin appearing in bladder urine per minute, one could determine the number of milliliters of filtrate needed to supply this amount by dividing the total number of milligrams excreted per minute by the number of milligrams of inulin in 1 ml of glomerular filtrate. This technique is applied to any intact animal or human patient by quantitatively collecting bladder urine, usually with the aid of a urethral catheter, and simultaneously sampling the inulin content of arterial blood. The latter is used in place of glomerular filtrate, which is virtually impossible to obtain, because the inulin contents of arterial blood and filtrate are identical. Only red blood cells and proteins are retained during the process of ultrafiltration. Hence the glomerular filtration rate (GFR) is determined by Eq. (1), where U_{in} is the amount

$$\mathrm{GFR} = \frac{U_{in}V}{P_{in}} \tag{1}$$

of inulin in 1 ml of urine, V the number of milliters of urine formed per minute, and P_{in} the amount of inulin in 1 ml of arterial plasma (filtrate).

Tubular resorption and secretion. Knowing the volume of filtrate formed per minute, one can measure the amount of any freely filterable plasma constituent being filtered as GFR $\times$ P_x, where P_x is the amount of plasma glucose, urea, amino acid, or other constituent in 1 ml of arterial blood (glomerular filtrate). The amount of this plasma constituent either resorbed or secreted by the tubules is then determined as the difference between the amounts filtered and that finally appearing in urine. For glucose, which is resorbed subsequent to filtration, this would be GFR $\times$ $P_G - U_G V$. For penicillinic acid, which is profusely secreted into urine by the renal tubule cells from postglomerular blood that supplied the tubules, it would be $U_{Pa}V -$ GFR $\times$ P_{Pa}.

For most organic compounds actively secreted or resorbed by the renal tubules subsequent to being filtered, there is a maximal limit imposed on the tubular transport rate (T_m). This maximal rate of transfer may be dependent upon the saturation of carrier capacity or upon a limitation in the amount of free energy available for moving these compounds across barriers against their own concentration gradients. Analogous compounds that are secreted or resorbed simultaneously may compete with one another for common carrier; raising the concentration of one will depress the transport rate of the other.

Renal plasma flow (RPF). The Fick principle is used to measure the amount of blood or plasma flowing through the kidneys per minute. For example, if the total milligrams of inulin excreted per minute ($U_{in}V$) is determined and divided by the number of milligrams removed from 1 ml of plasma as it traverses the kidney (arterial-venous difference or art. P_{in} − ven. P_{in}), then one knows the number of milliliters of plasma that was delivered to the kidneys per minute in order to account for the total number of milligrams of inulin excreted per minute.

However, use of inulin for this purpose calls for separate sampling of blood from the renal artery and the renal vein. Arterial blood from any part of the body can be used for concentrations of plasma constituents in renal arterial blood, but samples for venous blood must be from the renal vein, and this poses a serious problem in studies on intact animals or patients. Fortunately, certain foreign substances are secreted so profusely by the tubules that when they are infused into the circulation slowly, only low concentrations accumulate in arterial blood, and they are completely extracted while traversing the renal circulation. Inulin, excreted solely by glomerular filtration, is only 20% extracted in most mammals, but some compounds such as penicillin and *para*-aminohippuric acid (PAH) are almost completely removed by tubular secretion subsequent to glomerular filtration. So with low art. P_{PAH}, the value for ven. P_{PAH} may be assumed to be zero in the Fick equation, and the RPF is determined from Eq. (2). Thus the

$$\mathrm{RPF} = \frac{U_{PAH}V}{\text{art. } P_{PAH}} \tag{2}$$

total number of milligrams of PAH extracted from the circulation by the kidneys per minute, divided by the number of milligrams of PAH extracted from each milliliter of plasma, gives the total number of milliliters of plasma that came to the kidneys per minute in order to account for the number of milligrams of PAH extracted (excreted) per minute. The small amount of nonextracted PAH in venous blood

probably represents that fraction of blood which supplies nontubular tissue. Hence, RPF measurements made by this method are sometimes referred to as "effective" RPF, to represent that which is selectively supplied to active tubular tissue; other extraglomerular blood goes to inert fat, connective tissue, and the renal capsule.

Filtration fraction. When both GFR and RPF are known, one can obtain another hemodynamic parameter. Alterations in the fraction of plasma filtered, FF = GFR/RPF, can give insight into whether vasomotor arteriolar changes occur mainly afferent to the glomeruli or as alterations in the diameter of efferent arterioles leading away from the glomeruli. If constriction occurred solely in the afferent glomerular arterioles, one would expect both GFR and RPF to fall correspondingly, and the FF would remain constant. If, however, resistance increased exclusively in the efferent arteriole, RPF would fall as before but the effective glomerular filtration pressure would be expected to rise, thereby increasing the fraction filtered.

In most mammals an average of about one-fifth of the plasma delivered to glomeruli is filtered. Lower vertebrates have much lower filtration fractions because of the separate renal portal (venous) circulation that bypasses the glomeruli and goes directly to the tubules. Hence, the high-pressure arterial blood supply that exclusively sustains glomerular filtration in all vertebrates constitutes a relatively small fraction of the total amount of blood delivered to the kidney in inframammalian forms.

Micropuncture. With the development of exquisitely delicate techniques that enable one to deal quantitatively with nanoliter (10^{-9} liter) fluid samples, physiologists turn to direct micrurgical techniques to study specific secretory and resorptive processes in assignable loci of renal tubule and collecting ducts. Minute samples of urine can be obtained either while urine is flowing freely within the tubular lumen, while tubules are being perfused with various solutions, or while flow is stopped by injections of oil at selective sites in "split-drop" experiments. Simultaneously, blood samples can be obtained from adjacent regions, and methods are available to measure electrical potential differences and flow of current across tubule cells or across separate cell membranes on the vascular and luminal sides of tubular epithelium.

Stop-flow technique. It is also possible to localize specific secretory and resorptive processes in intact animals by clamping the ureters to allow a stationary column of tubular urine to be held in contact with the tubular epithelium for several minutes. This permits cells in the various segments to operate on the urine in the column, thereby exaggerating processes that modify the composition of tubular urine in specific regions. The clamp is then released, and while urine spurts out under pressure, several dozen serial samples are collected and subjected to chemical analyses. The first samples come from the pelvis, collecting ducts, and distal segments of the tubule; later samples are from segments more proximal to

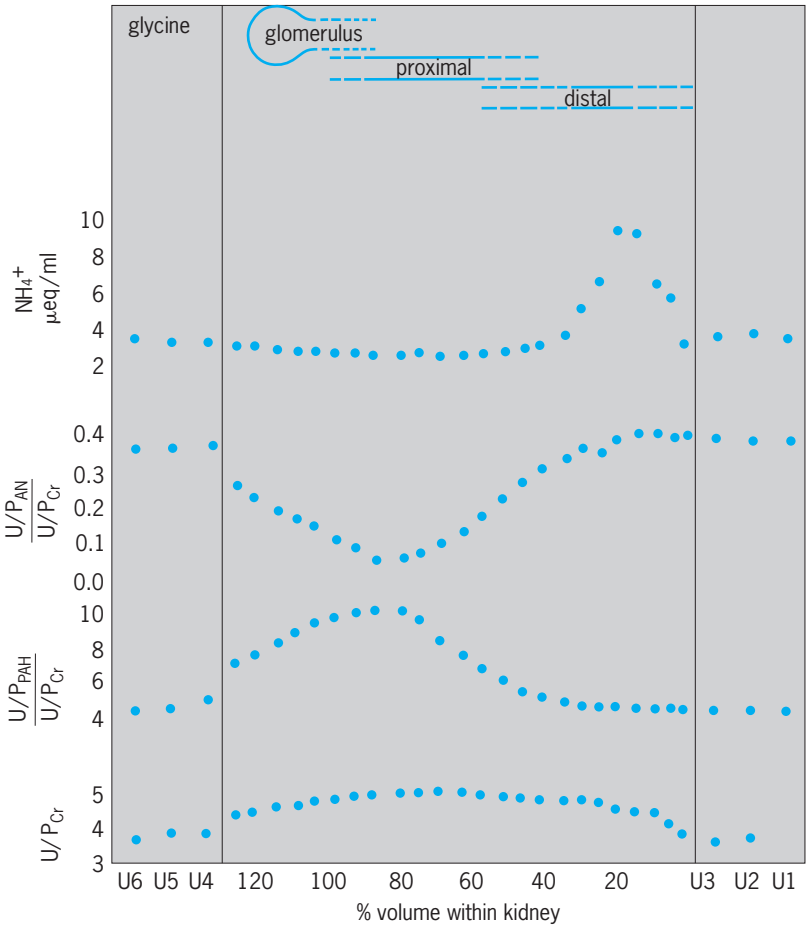

Fig. 18. Stop-flow analysis of the sites of ammonia secretion, aminonitrogen resorption, and PAH secretion. (*After R. F. Pitts, Physiology of the Kidney and Body Fluids, Year Book Medical Publishers, 1963*)

the glomerulus; and finally filtrate and urine formed after release of the clamp are sampled. Some substance excreted solely by glomerular filtration is used as a marker, and its concentration reflects reabsorption of water by the tubule segments. Any urinary component whose concentration rises with reference to the marker is assumed to be secreted in that area, and when its relative concentration falls, it is assumed to be resorbed in the segment corresponding to that particular serially collected sample. **Figure 18** shows a stop-flow analysis in which the site of reabsorption of the amino acid glycine AN is identified as being in the same proximal region in which PAH is secreted. Ammonia (NH_4^+) is secreted distally. In this case creatinine was used as a marker, and the urine/plasma (U/P) concentration ratios of PAH and NH_4^+ were factored by the corresponding U/P ratio of creatinine to correct for reabsorption of water which occurred at varying rates along the nephron during the stoppage of flow.

Culture methods. Isolated renal tubules and thin slices of kidney in suitably oxygenated, balanced salt solutions containing substances capable of being secreted continue to carry on their cellular transport activities for several hours. In the simplest preparations a colored organic compound such as phenol

red can be directly observed under the microscope as it is progressively concentrated to a very high degree within the lumen of renal tubules. Such studies have thrown considerable light on carrier-mediated transport competition and the nature of energy dependence. More complicated techniques involve use of chemical analyses and techniques for microperfusing isolated tubules with synthetic solutions of known composition. Fish kidneys, especially of the marine flounder, are particularly useful for tissue culture because they contain little cementing substance and the individual tubules can be easily teased apart. *See* TISSUE CULTURE.

Patterns of excretion. Nephrons of primitive protovertebrates probably were supplied exclusively with venous blood, and the open-ended tubules had nephrostomes that drained fluid directly from the body cavity. With the assumption of a freshwater habitus and the need for osmoregulation, it is thought that fishes acquired a glomerulus that at first was perhaps only loosely associated with the coelomostome, as in the embryonic pronephros of present-day vertebrates. Eventually a dual blood supply became possible when the glomerulus supplied with arterial blood became tightly encapsulated within the end of the tubule, but the open coelomostome persisted in some lower vertebrates. With the evolution of a homeostatically regulated internal environment and the guarantee of constant high-pressure filtration at the glomerulus, the renal portal system disappeared in mammals leaving the renal tubule solely with an arterial postglomerular bloody supply (**Fig. 19**).

The functional significance of the venous renal portal supply in a cold-blooded form such as the frog, for example, probably derives from its ability to maintain tubular secretory activity under environmental conditions that would tend to reduce arterial blood pressure to the point where glomerular filtration ceases. Glomerular activity generally is more labile in cold-blooded vertebrates than in mammals. Rates of filtration rapidly fall in fishes and amphibians when they are stressed either by evaporation or osmoregulatory water deprivation, or by cold. Lower vertebrates also rely more on tubular secretory processes for sustenance when glomerular filtration is diminished or ceases entirely and when the low-pressure venous supply of the renal portal system may be the sole source of blood for the kidney.

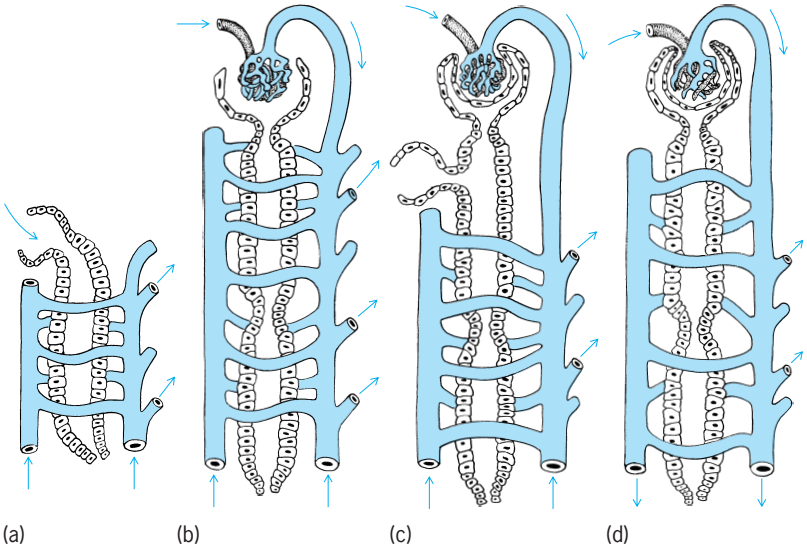

Fig. 19. Probable stages in the evolution of the vertebrate nephron; arrows indicate direction of flow. (*a*) Hypothetical protovertebrate tubule drained the body cavity through an open coelomostome with only a venous blood supply. (*b*) In earliest vertebrates the coelomostome was brought into loose association with knots of arterial blood vessels capable of filtering fluid under hydrostatic pressure. (*c*) Later evolution sealed Bowman's capsule in the glomerulus. (*d*) In mammals the renal portal venous supply and coelomostomes disappeared. (*After H. W. Smith, From Fish to Philosopher, Little, Brown, 1959*)

Teleost fishes. The urinary system of fishes plays only a secondary role in the excretion of nitrogenous waste products. As much as 90% of the nitrogen that comes from the deamination of amino acids is usually excreted at the gills in the form of ammonia. Fish in fresh water form large quantities of glomerular filtrate, and this is desalinated by their characteristically long renal tubules which eventually produce copious volumes of essentially salt-free urine. Fresh-water fishes excrete urine equivalent to 30% of their body weight per day, and the dilute urine may have freezing point depressions as low as Δ 0.07°F (Δ 0.04°C). Anadromous forms such as the salmon and steelhead trout show markedly reduced urine flow and glomerular filtration rates when in their seawater habitat.

Marine teleosts need to conserve free water in maintaining solute concentrations in blood that are only about half that of seawater (average is approximately Δ 1.62°F (Δ 0.9°C); seawater freezes at 28.54°F (−1.86°C). Production of urine is scanty; glomeruli are reduced in size and totally absent in some 20 species of marine teleosts. After intake of seawater and following intestinal absorption, blood is desalinated to maintain water balance by the gills, which remove the monovalent sodium and chloride ions, and the kidneys, which extract divalent magnesium, calcium, and sulfate ions. These ion transfers are carried out by specific energy-demanding active transport processes, which operate in opposite directions in similar branchial and renal systems of fresh-water fishes. Distal tubules are absent in marine fishes, and the urine of both the glomerular and aglomerular forms is roughly isosmotic with that of blood.

Cyclostomes. Primitive hagfishes such as *Myxine* resemble marine invertebrates and are unique among vertebrates in that their blood is isosmotic with seawater, due almost entirely to the accumulation of inorganic ions (Fig. 17). Hence no osmoregulatory mechanisms are required. Fresh-water and brackish lampreys such as *Lampetra*, on the other hand, produce copious flows of dilute urine, and their renal function resembles that of fresh-water and diadromous teleosts.

Cartilaginous fishes. Sharks, skates, rays, and chimeroids living in the sea prevent water loss to the marine environment in a way that is totally different from that of marine teleosts. They have relatively

large glomeruli and they do not drink seawater; instead the renal tubules have specialized processes capable of extracting urea and trimethylamine oxide from glomerular filtrate, and the active resorption of these organic compounds raises the osmotic concentration of their blood and body fluids to equality with seawater (Fig. 17).

Lungfishes. Urinary systems of fresh-water dipnoans function similarly to those of fresh-water teleosts. However, during estivation all urinary functions stop, and potentially toxic ammonia is then converted by the liver to urea which may accumulate in body fluids to extremely high levels without causing damage. When floods occur after periods of drought sometimes lasting several years, water is rapidly taken up osmotically, and the resulting increase in blood volume and pressure restores glomerular filtration; then a copious production of urine rapidly removes urea and other waste products accumulated during estivation. It is interesting that of the three surviving lungfishes, the Australian lungfish (*Neoceratodus forsteri*), which cannot estivate or survive out of water, is the only one incapable of synthesizing urea via the so-called ornithine cycle that accounts for the production of urea in the African and South American lungfishes, both of which can estivate.

Amphibians. The purely aquatic urodeles, such as *Necturus*, and the tadpole stages of anuran frogs and toads have patterns of urinary excretion similar to that of fresh-water teleosts. However, during metamorphosis when anurans change from gill to lung breathing, nitrogen is excreted principally in the form of urea rather than ammonia. Urea is actively secreted by the tubules of the more aquatic frogs, but it does not appear in terrestrial toads or in the unusual crab-eating frog (*Rana cancrivora*) of Southeast Asia that lives in marine mangrove swamps. Such retention of urea could serve the same osmoregulatory function that it does in marine cartilaginous fishes, that is, to help toads in an arid environment to absorb water quickly through the skin during brief showers, and to provide a more favorable free-water gradient for the marine toad in its hyperosmotic environment. Renal accommodations to semiterrestrial life in amphibians, which involve glomerular recruitment, facultative tubular reabsorption of water, and alterations in the synthesis and secretion of urea are undoubtedly implicated also in such amphibious fishes as the tree-climbing perch (*Anabas*), the mudskippers, many gobies, and such reptiles as the aquatic turtles, alligators, and many water snakes which may alternately live in and out of an aqueous environment.

Terrestrial reptiles and birds. These egg-laying forms have adapted to "uricotelism," which effectively provides for detoxification of ammonia and also for extremely efficient conservation of water. In addition to having relatively low rates of glomerular filtration, uric acid can be actively secreted by the tubules to form a urine practically saturated with urate. Actual precipitation of solid urate occurs in the allantoic bladder of embryos and in the cloaca of adults, and masses of white crystals are found as residues in hatched eggshells and feces, respectively.

Mammals. The placental form of embryonic development in mammals guaranteed an aquatic environment for the fetus which permitted retention of the "ureotelic" habitus. The renal portal system disappears in mammals, whose tubules are supplied solely with blood that has previously traversed the arterial glomerular capillaries (Fig. 19). The unique functional feature of the mammalian kidney is its ability to concentrate urine. Human urine can have four times the osmotic concentration of plasma (1200 milliosmoles/liter), and some desert rats that survive on a diet of seeds without drinking any water have urine/plasma concentration ratios as high as 17 (6.34 Δ osmoles/liter, more than six times that of seawater). More aquatic forms such as the beaver have correspondingly poor concentrating ability.

The concentration operation depends on the existence of a decreasing gradient of solute concentration that extends from the tips of the papillae in the inner medulla of the kidney outward toward the cortex. The high concentration of medullary solute is achieved by a double hairpin countercurrent multiplier system which is powered by the active removal of salt from urine while it traverses the ascending limb of Henle's loop (**Fig. 20**). The salt is redelivered to the tip of the medulla after it has diffused back into the descending limb of Henle's loop. In this way a hypertonic condition is established in fluid surrounding the terminations of the collecting ducts. Urine is concentrated by an entirely passive process as water

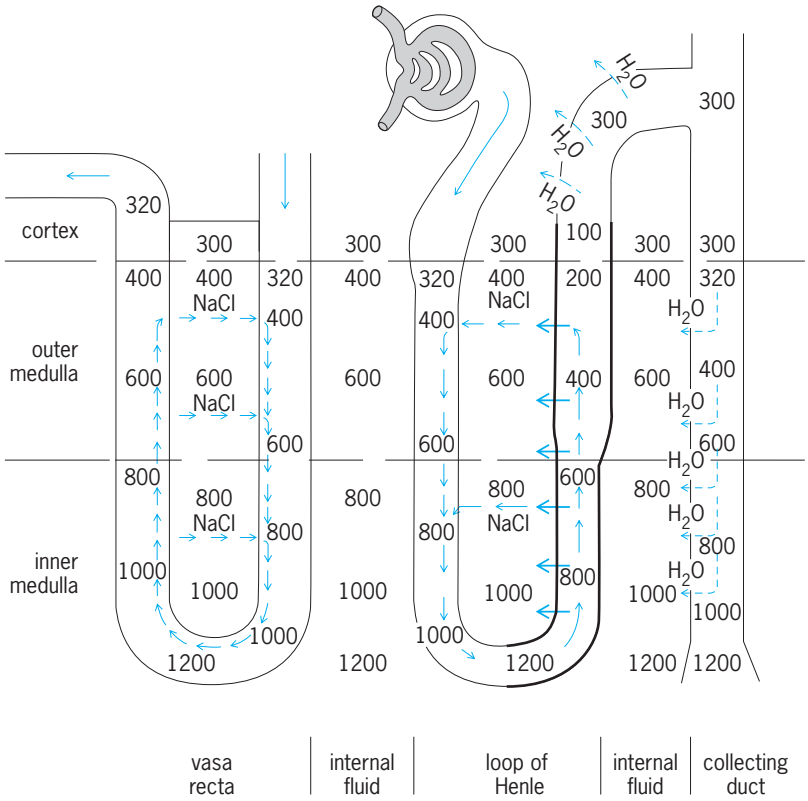

Fig. 20. The tubular countercurrent mechanism for concentrating the urine, and the exchange system in the straight hairpin blood vessels of the medulla, which minimizes loss of solute from medullary interstitial fluid. Arrows indicate direction of movement of solute. (*After A. C. Guyton, Textbook of Medical Physiology, Saunders, 1966*)

leaves the lumen of collecting ducts to come into equilibrium with the hypertonic fluid surrounding its terminations.

Urine more dilute than plasma is formed when hypotonic urine delivered by the ascending limb to the cortex traverses the lengths of distal tubules and collecting ducts and it is not free to come into equilibrium with the progressively more salty interstitial fluids of the cortex, outer medulla, and inner medulla, respectively.

Another countercurrent exchange system operating in the straight blood vessels (vasa recta) that parallel the loops of Henle prevents washout of the hypertonic interstitial fluid of the inner medulla where the blood is in osmotic equilibrium with the surrounding interstitial fluid. Because of the hairpin arrangement of the vasa recta, salt and other solutes are trapped and kept from being carried away from the medulla (left side of Fig. 20). As solute diffusing from the ascending capillary limb comes into equilibrium with more dilute interstitial fluid nearer the cortex, it enters the descending vessel to be redelivered to the inner medulla. This entirely passive process involving exchange of solute is the same as that operating in countercurrent heat exchangers.

Removal of most of the salt from glomerular filtrate takes place in the proximal convoluted tubules of the renal cortex. Sodium is actively resorbed from tubular fluid, and chloride then passively follows down an electrical gradient. Removal of solute effects the simultaneous isosmotic entry of water into the cortical interstitial fluid. This solution, containing about 300 milliosmoles of solute, is then rapidly restored to the general circulation by blood perfusing the cortical capillaries. In humans, about 80% of the glomerular filtrate is withdrawn by the time tubular fluid enters the descending limbs of Henle's loops. During water diuresis the distal tubules and collecting ducts are relatively impermeable to water, and the hypotonic urine then leaving the loops of Henle is made even more dilute by continued active extrusion of ions, finally resulting in the elimination of large volumes of urine containing little solute. In contrast, during conditions demanding conservation of water antidiuretic hormone (ADH) is released by the posterior pituitary gland, and the distal tubules and collecting ducts are then rendered freely permeable to water. Urine in the presence of ADH is isotonic by the time it reaches the middle of the distal segment, and its volume has been drastically reduced by the continued active extrusion of sodium and the associated passive diffusion of water. The small amount of isosmotic urine entering the collecting duct is progressively concentrated as it gives up water to the hypertonic medullary interstitial fluid and finally achieves a maximal concentration of solute equivalent to that of fluid surrounding the collecting ducts at the papillary tips.

Regulation mechanisms. Many years ago it was shown that injection of crude extracts of pituitary glands into frogs and toads caused rapid weight gain by accumulation of water. Now a whole family of octapeptides is known to be secreted by neurons of the hypothalamus of the brain and then transported by axons to the pituitary gland where they are stored in the neurohypophyseal lobe. Two of these, arginine vasotocin and vasopressin, oppose diuresis in terrestrial and semiterrestrial vertebrates, and they are frequently referred to as ADH. ADH increases water uptake by opening pores in the skin and urinary bladders of anurans and diminishes urinary output both by increasing the extraction of water from distal tubular urine and by reducing the rate of glomerular filtration. Secretion of ADH is regulated by osmoreceptors in the hypothalamus. Among teleosts secretions of a neurosecretory system in the posterior region of the spinal cord are passed along into a neurohemal organ, the urohypophysis, via a structural and functional arrangement similar to the hypothalamic-neurohypophyseal system of the terrestrial forms.

Renal regulation of electrolytes is accomplished by steroid hormones produced by the adrenal cortex or by homologous tissues in the lower vertebrates. In the absence of cortical steroids kidneys cannot adequately retain sodium and chloride; serum potassium levels are elevated; and there is a marked reduction in blood volume with death inevitably ensuing. Aldosterone is the most potent cortical steroid in mammals. Its release is regulated by changes in the electrolyte level of blood flowing through the brain and kidney. It also promotes sodium retention and potassium loss in sweat and gastrointestinal glands. *See* STEROID.

In the frog, however, aldosterone induces the renal tubule to retain potassium and lose sodium. Its effect on the kidney is just the opposite of that in mammals; nevertheless, the frog recovers sodium even more efficiently from bladder urine, and the net result in conserving salt is the same.

Renin is a proteolytic enzyme probably produced by specialized secretory cells in the afferent glomerular arteriole. It reacts with a plasma component to produce a nonprotein substance called angiotensin II which is capable of raising arterial blood pressure. There are several lines of evidence that suggest that the kidney's renin-angiotensin system is a prime regulator of the secretion of aldosterone by the adrenal cortex. Apparently some kind of feedback regulation mutually controls these two important secretions that regulate blood pressure, renin by affecting peripheral resistance in the arterioles, and angiotensin II by controlling retention of salt, hence blood volume and pressure. A negative feedback mechanism could be triggered by hemorrhage or some other factor, causing a temporary decrease in renal arterial blood pressure and renal blood flow. The resultant hypotension could activate the renin-angiotensin system causing the release of aldosterone, which in turn would induce the renal tubules to retain sodium and to increase blood volume, and consequently to raise blood pressure. Such a negative feedback path would then be the afferent signal needed to inactivate the renin-angiotensin system. This is speculative, but the system is of interest to both clinicians and comparative physiologists.

There is also a feedback regulatory system within the kidney itself that apparently coordinates tubular and glomerular function. The juxtaglomerular apparatus which is found in all classes of vertebrates consists of specialized tubular epithelium at the end of Henle's loop (macula densa) and of renin-containing cells in the afferent arterioles. These two separate types of specialized cells come into intimate contact with one another, and the two apposing parts always belong to the same nephron. There is considerable evidence to support the hypothesis that an increased concentration of salt in tubular urine at the macular densa site activates a sodium-sensitive feedback system which operates through the juxtaglomerular apparatus to reduce the rate of glomerular filtration and thereby adjust the tubule's load of filtered salt to the sodium reabsorptive capacity of this particular nephron. *See* ENDOCRINE MECHANISMS; URINE.

Roy P. Forster

Bibliography. R. P. Forster, Comparative vertebrate physiology and renal concepts, in *Handbook of Physiology*, sect. 8: *Renal Physiology*, pp. 161–184, American Physiological Society, 1973; D. J. Griffiths, *Urodynamics*: *The Mechanics and Hydrodynamics of the Urinary System*, 1980; A. C. Guyton, *Textbook of Medical Physiology*, 10th ed., 2000; R. Hill and G. Wyse, *Animal Physiology,* 2d ed., 1989; W. S. Hoar, *General and Comparative Physiology*, 3d ed., 1983; C. L. Prosser, *Comparative Animal Physiology*, 4th ed., 1991; T. W. Torrey, *Morphogenesis of the Vertebrates*, 4th ed., 1979.

Urinary tract disorders

Diseases affecting the outflow tract of the urinary system from the renal pelvis to the distal portion of the urethra. Disorders of the kidney itself are considered in a separate article. *See* KIDNEY DISORDERS.

Structures and Associated Disorders

Urinary tract disorders can occur in the renal pelvis, ureters, bladder, or urethra.

Renal pelvis. The renal pelvis, the structure that collects and carries urine from the kidney to the ureter, is subject to a number of disorders.

Anomalies. Anomalies of the renal pelvis are common. The pelvis may be subdivided to form a duplicated system, which may join a single ureter or empty into separate ureters. By itself such an anomaly is harmless, but if there is associated obstruction or vesicoureteral reflux (abnormal flow of urine from the bladder back into the kidney), there may be symptoms of pain, associated infection, and possibly destruction of renal tissue.

Calculi. Calculi typically arise in the renal papillae usually below the epithelial layer. The cause is usually due to a metabolic defect, which results in excessive excretion of oxalate, urate, calcium, and/or cystine. The most common stones are calcium-containing, such as calcium oxalate and calcium phosphate. Precipitating factors include dehydration, dietary excess of certain foods, and infection. Also, in the presence of a bacterial infection or foreign material, the salts, which are often in a supersaturated condition in the urine, may precipitate out. Thereafter, the calculi enlarge by accretion of the same or other salts.

After reaching a certain size, calculi break through the epithelial layer and migrate into the renal pelvis and float free. They may pass down the ureter. Some stones, particularly magnesium ammonium phosphate calculi (struvite stones), are associated with bacterial infection and may become very large and take the shape of the renal pelvis and infundibulae. These are, therefore, called staghorn calculi. In addition to causing discomfort, they can also cause obstruction and infection of the renal parenchyma, which can lead to loss of kidney function. Patients with staghorn calculi must be treated to remove the stones.

Other obstructions. Fibrous bands and aberrant vessels associated with the kidneys are usually blamed for obstruction when no other intrinsic cause can be found for ureteropelvic obstruction. In some cases, there may be intermittent obstruction from "sagging" of the kidney in the upright posture. Tumors, either intrinsic or extrinsic, and periureteral fibrosis (fibrosis around ureter) are also capable of blocking the flow of urine from the renal pelvis.

In complete acute obstruction from any cause, the backpressure in the collecting system hinders urine secretion from the kidney. This process is reversible if the obstruction is relieved. In chronic partial obstruction, the most common type, the backup of urine in the renal collecting system leads to dilation of the pelvis, known as hydronephrosis. Hydronephrosis is classified as mild, moderate, or severe and reflects the degree of dilation of the renal pelvis due to the obstruction.

Obstruction causes urinary stasis and this contributes to developing infection (pyelonephritis; if pus develops the condition is known as pyonephrosis). The resulting distension and (often) infection may destroy some or all of the renal parenchyma, with deleterious effects on renal function. The kidney may ultimately become an abnormally enlarged sac with no useful function.

Infection. Acute infection of the kidney, known as pyelonephritis, produces fever and flank pain. If there are no underlying causes (such as stone, obstruction, or tumor), the disease can be treated with antibiotics and has no serious aftereffects.

Tumors. Tumors of the pelvis are rare; when they do occur they are usually transitional-cell carcinomas with varying degrees of differentiation. They can cause obstruction and hematuria. They are often associated with similar tumors elsewhere in the urinary system. It is not known whether this represents superficial implantation of tumor or new tumor growth due to persistence of inciting factors.

Ureters. These long thin tubes that carry urine from the renal pelvis to the bladder are subject to the same diseases as the renal pelvis.

Anomalies. Duplication of the ureters can occur at any point, from a dual pelvis and fusion to form one ureter, to complete duplication and separate

implantation into the bladder. The anomaly is generally harmless unless obstruction or reflux is present, which can contribute to developing infections. Sometimes an anomalous ureter inserts into the vagina and causes dribbling of urine, for which surgical correction is required due to incontinence.

Calculi. One of the most common causes of excruciating abdominal and flank pain, often radiating to the groin or testicle, is the passage of small calculi. They originate from the renal pelvis and travel toward the bladder. If the stone becomes lodged along the course from the kidney to the urethra, intervention is required to remove the stone. If the stone causes obstruction, the ureter and renal pelvis proximal to the stone becomes dilated (hydroureteronephrosis). Again, obstruction of the urinary system can lead to renal damage and loss of renal function.

Other obstructions. Occasionally, the ureter is obstructed by external pressure from retroperitoneal fibrosis (development of a fibrous mass in the back of the abdomen behind the abdominal lining), tumor, or aortic aneurysm. In addition, any chronic obstruction of the lower urinary tract (bladder or urethra) will lead to progressive dilation of the ureters and renal pelvis. The effect is the same no matter what the cause; obstruction can lead to permanent renal damage if not relieved.

Vesicoureteral reflux. If the normal valvelike mechanism that prevents urine in the bladder from going back toward the kidney is defective, vesicoureteral reflux results. Reflux can occur during bladder filling or emptying. When associated with an infection, vesicoureteral reflux can lead to renal injury.

Tumors. These are usually transitional-cell carcinomas and do not differ appreciably from those seen in the renal pelvis and bladder.

Bladder. The urinary bladder is subject to anomalies, obstruction, inflammation, calculi, fistulae, and tumors. As the receptacle for urine from the kidneys, it is variable in shape and form depending upon the amount of urine it contains.

Anomalies. Bladder exstrophy is an uncommon congenital anomaly due to failure of both the lower abdominal wall and anterior bladder wall to close during development. The bladder appears directly through the abdominal wall defect. Occasionally a fetal structure, such as the urachus, fails to close by the time of birth and drains urine through the umbilicus. Abnormal congenital valves of the posterior urethra, known as posterior urethral valve, interfere with bladder emptying. This condition occurs only in males and is significant because it can lead to chronic renal obstruction and subsequent renal failure.

Calculi. Bladder calculi usually develop as a result of infection or obstruction. They can be of varying size and are often composed of phosphates and urates. After removal or spontaneous passage of stones, new stones tend to form unless the predisposing factors are corrected.

Obstruction. Interference with complete emptying of the bladder may be secondary to various factors. In males it is often due to prostatic enlargement, and in females it is typically due to a cystocele, which is a bulging of the bladder into the vaginal vault. Other causes of obstruction include calculi; tumors; constriction secondary to surgical instrumentation or gonococcal infection; or congenital valves. A neurogenic bladder (which occurs when the normal balance of bladder wall contraction and sphincter relaxation is altered and the bladder cannot completely empty) usually follows spinal cord injury or disease and can lead to inadequate bladder emptying.

Obstruction from any cause leads to changes in the bladder wall. The bladder wall can weaken in areas and bulge outward, causing the formation of a diverticulum. Other effects of obstruction include increased susceptibility to infection, stone formation, and backpressure on the kidneys and ureters.

Inflammation. Cystitis or inflammation of the urinary bladder is a common affliction and is characterized by dysuria (painful urination), frequency, and urgency. The organisms responsible are primarily bacterial, including *Escherichia coli*, pseudomonas, and other gram-negative organisms, although yeast and fungi can cause urinary tract infections. Tubercle bacilli are rarely involved. Bacteria can gain access to the bladder from the blood or, more commonly, ascend from the urethra. The lesions include granular or even cystic thickening of the epithelium and redness and swelling of the wall. A peculiar plaquelike lesion called malakoplakia, which is usually associated with the bladder but can be found elsewhere in the body, is known to be a reaction to *E. coli*.

Fistulas. Abnormal connections (fistulas) between the bladder and the female urethra or vagina are not unusual. These can be due to pelvic or transvaginal surgery, childbirth, or radiation. Enterovesical fistulas (abnormal connection between the bowel and bladder) can also be secondary to pelvic surgery, radiation therapy, or inflammatory bowel diseases. Fistulas can result in incontinence and chronic urinary tract infection.

Tumors. The bladder is frequently the site of tumors, which are almost always of transitional-cell origin. Although benign lesions can exist in the bladder, bladder cancer needs to be ruled out. Bladder cancer can remain superficial or it can penetrate the bladder muscle, invade neighboring structures, obstruct the ureters or urethra, and metastasize to distant organs. Malignant tumors are more frequent in cigarette smokers and in people who work near aniline dyes. These observations indicate that carcinogens excreted in the urine play a role in initiating bladder cancer.

Female urethra. This structure is not often diseased. In older women a painful raspberrylike growth of inflammatory tissue may appear at the external urethral opening and is known as a caruncle. The urethra can also be a site of transitional-cell and squamous cell malignancy.

Roy N. Barnett; Katherine S. Rhee; Martin I. Resnick

Male urethra. The male urethra is approximately 20 cm in length and is responsible for draining urine from the bladder and acting as a conduit for the delivery of semen. The first 3 cm is surrounded by the prostate gland and is termed the prostatic urethra.

The prostate gland may develop cancer, and options are available to treat this condition based on stage of the tumor or extent of disease at time of detection.

The prostatic urethra is prone to obstruction. The most common cause is a noncancerous growth of the prostate called benign prostatic hyperplasia. This growth is a natural process associated with aging and can occlude the prostatic urethra, causing progressive difficulty in urinating. Individuals may complain of straining with urination, weak stream, incomplete bladder emptying, nocturia (excessive nighttime urination), frequency, and/or double voiding. Similar symptoms may be secondary to urethral strictures which can be secondary to trauma, surgical instrumentation, or gonococcal infections. Symptoms of benign prostatic hyperplasia can be treated with medications to shrink the prostate or relax the bladder neck. If symptoms are severe or there is an inability to void, a portion of the prostate needs to be removed surgically. Less-invasive procedures such as microwave therapy, lasers, and transurethral needle ablation can be used as well. If the prostate is too large for an endoscopic procedure, it can be removed by open surgery.

Signs and Symptoms

Signs and symptoms of urinary tract disorders may be generalized or localized. General disturbances include fever and pain. Malaise can accompany fever, but it may also be due to the body's response to the toxins normally eliminated by the kidney if there is loss of kidney function. Local signs and symptoms depend upon the location of the disease. When the renal pelvis or ureters are obstructed or infected, back or abdominal pain overlying the affected organ often arises. The pain can be severe and radiate down to the groin or testicle. Other urinary tract disorders, such as bladder cancer, can present with asymptomatic hematuria (presence of blood in the urine).

Diagnosis

Techniques used to diagnose urinary tract disorders include clinical laboratory analysis of blood and urine, radiological procedures, and endoscopy.

The most common clinical laboratory procedure is urinalysis, which involves microscopic examination of the urine to identify red and white blood cells not normally present. If infection is suspected, cultures for bacteria can be done. Serum studies are useful in assessing renal function and identifying metabolic disorders. *See* URINALYSIS.

Intravenous contrast is used to help delineate and examine the urinary tract with x-rays. Stones can be diagnosed with a plain abdominal x-ray or computed tomography without contrast. In certain situations, radioisotopic dyes can be injected to determine the function of each kidney. The urinary tract can also be examined with ultrasound. *See* MEDICAL IMAGING.

Endoscopy requires the use of special instruments to directly visualize the urethra, bladder, ureters, and renal pelvis. Radiopaque dyes can also be injected into the urinary tract by this approach. When abnormal tissue is visualized, biopsies can be obtained and sent for pathologic examination.

Treatment

Treatment depends upon the type of urinary tract disorder that is present. For infections, a urine culture and sensitivity assessment should be obtained and a course of oral or intravenous antibiotics administered and adjusted appropriately according to the organism identified.

Obstructive lesions, such as benign prostatic hyperplasia, can be treated surgically or with medication to relieve lower urinary tract symptoms. Surgical intervention includes endoscopic or open procedures to remove excess prostate tissue. Treatment with medication involves the use of α-blockers and 5-α reductase inhibitors.

Stones are treated in a variety of ways including administration of medication and/or surgical management. Patients can choose to wait for the stones to pass spontaneously or can elect for surgical intervention, including electrohydraulic shock wave lithotripsy (ESWL®), ureteroscopy with lithotripsy or basket retrieval, percutaneous surgery, or open surgery.

The treatment of tumors, for example in the bladder, requires resection (surgical excision) to determine the pathological diagnosis and also the depth of invasion. If a bladder tumor is superficial, tumor resection can be followed with intravesical immunotherapy or chemotherapy to prevent recurrent tumors. More invasive disease should be treated with radical cystectomy, which involves removal of part of the urinary bladder, as well as the prostate (in men) and the uterus, fallopian tubes, ovaries, anterior vaginal wall and the urethra (in women), along with bilateral pelvic lymph node dissection. Combined radiation and chemotherapy is also an option in patients with invasive disease. Tumors of the upper tract may require nephroureterectomy (surgical removal of the kidney, ureter, and top part of the bladder) or distal ureterectomy (removal of the distal portion of the ureter) with ureteral implant, depending upon the tumor location and grade. *See* CANCER (MEDICINE); REPRODUCTIVE DISORDERS; TUMOR; URINARY SYSTEM. Katherine S. Rhee; Martin I. Resnick

Bibliography. M. I. Resnick et al., *Critical Decisions in Urology*, 3d ed., 2004; P. C. Walsh et al., *Campbell's Urology*, 8th ed., 2002.

Urine

An aqueous solution of organic and inorganic substances, mostly waste products of metabolism, which collects in the kidneys and is removed from the body primarily through the bladder and urinary tract. The kidneys maintain the internal milieu of the body by excreting these waste products and adjusting the loss of water and electrolytes to keep the body fluids relatively constant in amount and composition.

About 1.5 quarts (1500 ml) of blood courses through the kidneys each minute. As it passes through the glomeruli, the pressure of the blood, opposed by the osmotic pull of the plasma proteins, tends to drive water and salts through the wall of the glomerular tuft into the glomerular space at a rate of 7.2 in.^{3} (120 ml) per minute. This dilute glomerular filtrate courses through the kidney tubules where it is concentrated by active reabsorption of water, glucose, and essential salts. Approximately 119 parts of 120 of the total volume is reabsorbed, leaving only 1 part to be excreted as urine which, in the adult, under normal circumstances, amounts to approximately 1.2–1.5 quarts (1200–1500 ml) per day. *See* OSMOREGULATORY MECHANISMS.

The urine normally is clear and has a specific gravity of 1.017–1.020, depending upon the amount of fluid ingested, perspiration, and diet. The increase in specific gravity above that of water is due to the presence of dissolved solids, about 60% of which are organic substances such as urea, uric acid, creatinine, and ammonia; and 40% of which are inorganic substances such as sodium, chloride, calcium, potassium, phosphates, and sulfates. Its reaction is usually acid (pH 6) but this too varies with the diet. It usually has a faint yellow color due to a urochrome pigment, but the color varies depending upon the degree of concentration, and the ingestion of certain foods (for example, rhubarb) or cathartics. It usually has a characteristic aromatic odor, the cause of which is not known. *See* URIC ACID.

The chemical and physical composition of the urine frequently reflects pathological changes not only in the genitourinary tract but disease processes in other parts of the body. *See* CLINICAL PATHOLOGY; KIDNEY; URINALYSIS; URINARY SYSTEM.

Reuben Straus

Urodela

One of the orders of the class Amphibia, also known as the Caudata. The members of this order are the tailed amphibians, or salamanders, and are distinguished superficially from the frogs and toads (order Anura) by the possession of a tail, and from the caecilians (order Apoda) by the possession of limbs.

Morphology. In appearance, salamanders resemble lizards in that members of both groups normally have a relatively elongate body, four limbs, and a tail. The similarity is, however, only superficial. The most obvious external difference, the moist glandular skin of the amphibian and dry scaly skin of the reptile, is underlain by the numerous characters that distinguish reptiles from amphibians.

The vast majority of salamanders are well under a foot in length. The largest is the giant salamander of Japan and China, which may attain a length of 5 ft (1.5 m). A close relative, the hellbender, which lives in streams in the eastern United States, grows to over 2 ft (0.6 m) in length.

Four well-developed limbs are typically present in salamanders, but the legs of the mud eel (*Amphiuma*) of the southeastern United States are very tiny appendages and of no use in locomotion. The sirens (*Siren* and *Pseudobranchus*), also aquatic salamanders of the same region, have undergone even further degeneration of the limbs and retain only tiny forelimbs. Salamanders with normal limbs usually have four toes on the front feet and five on the rear, though a few have only four on all feet. A feature of the Urodela not shared with the Anura is the ability to regenerate limbs. If a leg is bitten off or otherwise lost, a salamander can grow a whole new limb.

Physiology. The moist and highly vascularized skin of a salamander serves as an organ of respiration, and in some forms shares with the buccal region virtually the whole burden. In one entire family of salamanders, the Plethodontidae, lungs are not present. Even in those forms with lungs and gills, dermal respiration plays a very important part.

The tympanum is absent in salamanders and the middle ear is degenerate. Hence these animals are deaf to airborne sounds. Undoubtedly correlated with this is the reduction of voice. In contrast to the highly vocal frogs, salamanders are mute or produce only slight squeaking or, rarely, barking sounds when annoyed. However, salamanders can detect sounds carried through ground or water.

Ecology. Salamanders live in a variety of aquatic or moist habitats. Many forms are wholly aquatic, living in streams, rivers, lakes, and ponds. Others live on land most of the year but must return to water to breed. The most advanced species live out their lives in moist places on land or, in the case of some tropical species, in trees, and never go in the water.

Reproduction. Producing the young alive rather than by laying eggs is very rare among salamanders. The eggs of salamanders may be deposited in water or in moist places on land. Presumably, aquatic breeding is the more primitive. Fertilization is external in most of the Cryptobranchoidea, internal by means of spermatophores in the remaining suborders. The spermatophore is a gelatinous capsule of sperm deposited by the male salamander in the course of an often elaborate courtship. The female salamander picks up the capsule with her cloacal lips and retains the sperm in her body until the eggs are fertilized and laid. The young salamander that hatches from the egg may have the adult morphology in the case of advanced forms, but usually it emerges as a larva. The larva of a salamander resembles the adult closely, in contrast to the larva of a frog, but differs in possessing gills and in being sexually immature, as well as in histological and other internal characters. There are salamanders in which metamorphosis from larval to adult morphology does not take place, and sexual maturity is attained in the larval condition. This condition is known as neoteny. In some forms it is a permanent, unchangeable condition, while in others the larva can be made to undergo metamorphosis

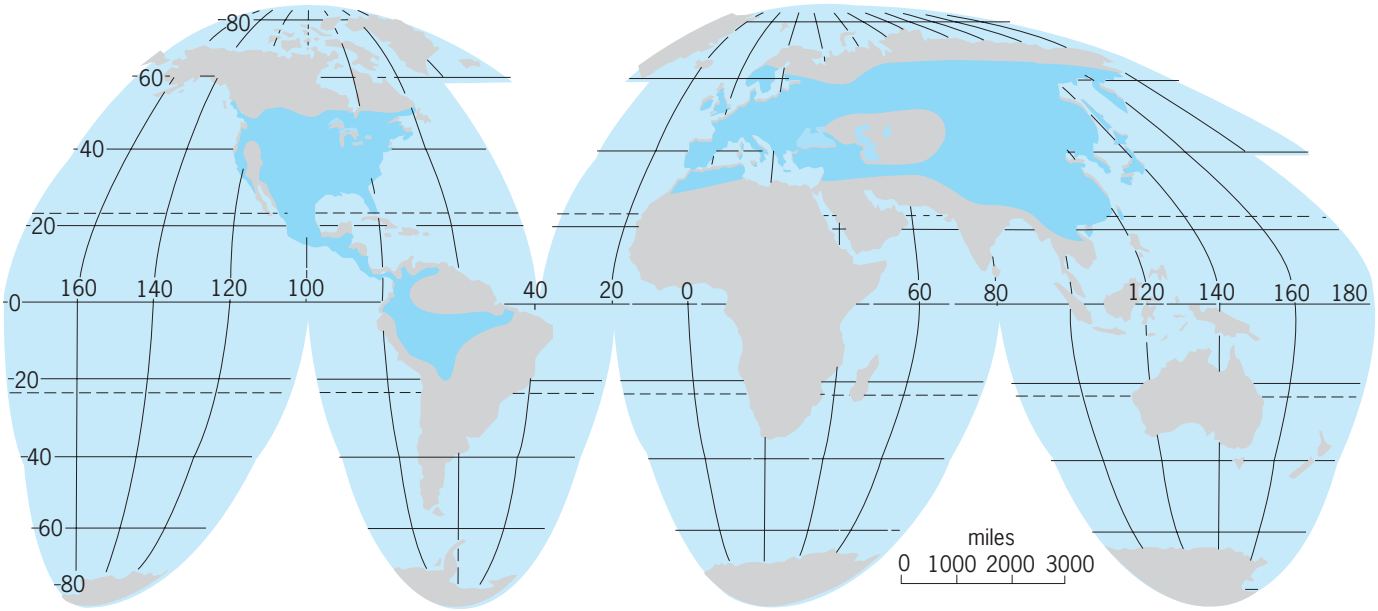

World distribution of salamanders (Urodela), indicated by colored areas of land. 1 mi = 1.6 km.

by treatment with thyroid hormone. *See* NEOTENY; THYROID GLAND.

Nutrition. Both as larva and as adults, salamanders are almost wholly carnivorous. Few species take anything but invertebrates in their diet, although larger forms may eat small vertebrates such as frogs and fish. Some are known to eat both plant material and invertebrates.

Distribution. The greatest concentration of families and genera of salamanders is in the eastern United States, and these animals are one of the few major groups of vertebrate animals that is distinctly nontropical. No salamander occurs below the Equator in the Old World, and only 18 species of the evolutionarily advanced family Plethodontidae are found as far south as South America in the New World (see **illus.**).

Classification. The Urodela are readily arranged in seven families, but authorities differ as to the relationships among these families and their classification into suborders. Three to five suborders may be recognized, and one of these is thought by some to merit ordinal rank, equivalent to the Urodela. About 300 species of salamanders are known to be living today, and these are distributed among about 54 genera.

Cryptobranchoidea. The Asiatic salamanders (family Hynobiidae) and the giant salamanders (family Cryptobranchidae) are sometimes grouped as the suborder Cryptobranchoidea. Five genera are recognized in the Hynobiidae, the most important of which is *Hynobius* with 18 species found over much of Asia from Japan to the Ural Mountains. One species even crosses the Arctic Circle, the farthest north reached by any salamander. There are only two living genera, each with a single species, in the Cryptobranchidae: the hellbenders (*Cryptobranchus*) of the eastern United States and the giant salamanders (*Megalobatrachus*) of Japan and China. *Andrias*, a fossil genus of Europe and North America, is apparently identical with *Megalobatrachus*, and if this is the case, the Asiatic giant salamander should be called *Andrias*.

The members of this primitive suborder differ from other salamanders in having external fertilization, though *Ranadon* evidently is an exception. Aquatic larvae are present in all forms, and some hynobiid larvae are peculiar in having sharp, recurved claws. Hynobiids undergo normal metamorphosis, but cryptobranchids retain many larval characteristics and are partially neotenic.

The hellbender and its giant Asiatic relative are aquatic species that live in rivers and streams. In the Hynobiidae there are both terrestrial and aquatic species.

Salamandroidea. This is the largest suborder of the Urodela with some 260 species in 5 families and 48 genera. All living salamanders belong to this group except the members of the suborder Cryptobranchoidea, which have external fertilization and lack the fusion of the angular and prearticular bones of the lower jaw seen in higher forms, and the suborder Meantes, neotenic forms lacking hindlimbs.

Ambystomatidae. The family Ambystomatidae, sometimes considered as a separate suborder, Ambystomoidea, is composed of 4 genera and 35 species limited in distribution to North America. The genus *Ambystoma*, with about 30 species, is the largest and most widespread of the family, and ranges from southern Alaska to Mexico and from the Atlantic to the Pacific. The tiger salamander (*Ambystoma tigrinum*) is found in a variety of subspecific forms from southern Canada to Mexico and over most of the United States. It lives in both arid and humid regions and is the only salamander in much of the region of the Great Plains and the Rocky Moun-

tains. Heavily forested regions of western North America are the home of the Pacific giant salamander (*Dicamptodon*). This largest of terrestrial salamanders may attain a length of 1 ft (0.3 m). All ambystomatids have aquatic larvae, but adults of most species are terrestrial, returning to the water only to breed. However, neoteny is of frequent occurrence in the family. The axolotl (*A. mexicanum*) of Mexico, which is neotenic in the natural state but will transform in captivity, is an ambystomatid.

Salamandridae. This family of North America, Europe, and Asia includes 40 species in 16 genera. The majority of forms are Eurasian, with only 2 genera and 6 species found in North America: *Taricha* on the Pacific Coast and *Notophthalmus* in the eastern United States and northeastern Mexico. These genera are endemic to North America, although an earlier, simplified classification placed several American and Eurasian species in a single genus, *Triturus*. The newts, as these forms are called, are often strikingly colored animals with brilliant red or orange underparts and darker brown or green upper surfaces. All newts have an aquatic larval state, whereas the adults are variously aquatic or terrestrial, according to species. A complicated life history is shown by the red-spotted newt of eastern North America, in which there is a land-dwelling stage (the red eft) commonly but not always interposed between the aquatic larval and the aquatic adult stages. Although most salamanders are nocturnal, the red eft and adults of *Taricha* and *Notophthalmus* are often found in daylight. These animals seem to be almost immune to attack by predators, possibly because of noxious skin secretions.

Two salamanders of the genus *Salamandra* are two European species with breeding habits that are peculiar for salamanders. *Salamandra atra*, a high-mountain form, retains the eggs and larvae within the body until they emerge as fully transformed young. *Salamandra salamandra*, a species of lower elevations, gives birth to larvae. In both species highly developed gills enable the larvae to absorb oxygen from the maternal oviduct.

Amphiumidae. Three large, eellike salamanders of the southeastern United States are the only members of the family Amphiumidae. These congo eels (*Amphiuma*) are partly neotenic aquatic animals with very tiny limbs of no use in locomotion.

Plethodontidae. This large and diverse family of about 180 species and 24 genera is characterized by the absence of lungs and the presence of a fine groove from nostril to upper lip. Except for two species of *Hydromantes* in Europe, the family is wholly American. Three species of this genus are found in California, but the greatest concentration of genera and species is in the eastern United States, with a secondary center in Mexico and Central America. The species found in South America are the world's southernmost salamanders.

The lungs are reduced or absent in some salamanders of other families, but this feature is consistent throughout the Plethodontidae. It is thought that the absence of lungs serves a hydrostatic function in stream-dwelling salamanders, and for this reason plethodontids are thought to have evolved from ancestors with such habits. However, the present-day members of this family are an ecologically diversified group. A majority of the species are terrestrial and lay eggs that develop directly into salamanders with adult morphology, skipping the aquatic larval stage. Other forms have aquatic larvae and terrestrial or aquatic adults. Many tropical species are arboreal, living in air plants, bromeliads, that retain moisture in dry seasons. At the other extreme are neotenic, subterranean species that come to light only when a cave is explored or a deep well dug.

Proteidae. The family Proteidae includes four species of *Necturus*, so-called mud puppies, all found in eastern North America, and the single species of *Proteus* in Europe. All are neotenic, aquatic animals. *Proteus* lives only in subterranean waters, whereas *Necturus* lives in lakes and streams. These two genera have sometimes been grouped in a separate suborder, Proteida, but it has also been suggested that they are not so closely related as they superficially appear to be.

Meantes. This group forms the smallest suborder of the Urodela, and these salamanders are sufficiently peculiar that they sometimes are regarded as constituting a separate order, Trachystomata. There are two genera, *Pseudobranchus* with a single species and *Siren* with two species. Commonly known as mud eels, these salamanders are found in areas of low elevation from Texas to Florida, Maryland, and Illinois. The mud eels are neotenic, retaining such larval characters as lidless eyes and external gills. The body form is eellike, with only anterior limbs being present. *Siren* reaches a length of 36 in. (91 cm), while *Pseudobranchus* is less than 9 in. (23 cm) long. Members of both genera are wholly aquatic, and are found beneath rocks in streams, and burrowing in the muck and vegetation of bogs and swamps. *See* AMPHIBIA; ANURA; APODA.

Richard G. Zweifel

Urogenital system

The combined structures composing the urinary and genital, or reproductive, organs of vertebrates. The terms urinogenital or genitourinary system are equally applicable and just as frequently used. During embryonic development, the urinary and reproductive systems are closely interrelated, as their ducts arise in associated mesodermal regions. Common passages are associated with both these systems in various vertebrates, such as the single orifice for emission of urine and sperm.

Urinary system. The principal organ of the urinary system is the kidney with its associated structures. Three morphologically distinct kidneys occur among the vertebrates. These are the pronephros, mesonephros, and metanephros. The pronephros is drained by the pronephric duct, the mesonephros by the Wolffian duct, and the metanephros by the

ureter. The principal function of the urinary system is the removal of nitrogenous waste products. *See* KIDNEY; URINARY SYSTEM.

Reproductive system. While the same structures of the urinary system occur in both males and females, specific structures such as the ovaries and testes are found in the normal female and male, respectively. The primary function of the reproductive system is the perpetuation of the species. The main reproductive structures of the female vertebrate consist of the ovary, oviduct, uterus, vagina, and external orifice. Many modifications occur among these structures in the various classes. Essentially, the male system consists of the testes and their ducts, seminal vesicles, prostate gland, Cowper's gland, and penis. Modifications of these structures also occur among the various vertebrate species. *See* REPRODUCTIVE SYSTEM.

Charles B. Curtin

Urokinase

An enzyme that is a plasminogen activator. Urokinase cleaves the plasma protein plasminogen, forming the active enzyme plasmin, which subsequently degrades fibrin. Thus urokinase is an essential component in the fibrinolytic clot-dissolving system in the human body. Other plasminogen activators include tissue plasminogen activator and streptokinase. *See* PLASMIN.

Urokinase is found in human urine and in much lower concentrations in human plasma. It is secreted by cells as a single-chain glycoprotein with a molecular weight of 54,000 and three structural domains: a human-growth-factor-like region, a kringle region or domain (a triple-loop protein structure containing three disulfide bonds), and a serine protease part. Unlike the tissue-type plasminogen activator produced by endothelial cells, urokinase has only a single kringle domain. The hydrolysis of urokinase by plasmin results in the formation of a two-chain molecule that possesses significantly increased activity as a plasminogen activator. *See* PROTEIN.

In the body, urokinase is produced by kidney cells, and its presence in the urine promotes the dissolution of any blood clots in the urine-collecting system of the kidneys or the bladder. Urokinase is also produced by a variety of tumor cells, and it is thought to be involved in the formation of tumor metastases. Human monocytes and fibroblasts have specific receptors for urokinase; the bound urokinase is functionally active. Urokinase can be inhibited by plasma inhibitor proteins. The presence of cell receptors and inhibitors suggests that the regulation of urokinase function is complex.

The use of urokinase to dissolve intravascular blood clots has necessitated the development of a method of production on a large scale. Initially, purified urokinase in the two-chain form was obtained by separation from human urine. However, the need for 400 gallons (1.5 kiloliters) of fresh urine to purify a single clinical dose of urokinase made that approach impractical. Subsequently, large-scale commercial production of urokinase was developed by using human embryonal kidney-cell culture. This was a major early achievement in biotechnology.

Urokinase is employed in clinical medicine in the treatment of venous blood clots (thrombophlebitis and pulmonary emboli), acute myocardial infarction, and arterial blood clots in the legs and arms. It is also used to prevent the accumulation of blood clots in intravenous catheters used to administer long-term chemotherapy. Like other plasminogen activators, urokinase is most effective in removing freshly formed blood clots. Unlike two-chain urokinase and streptokinase, tissue-type plasminogen activator binds to fibrin with high affinity; this suggests that tissue plasminogen activator is a safer and more effective thrombolytic agent. However, in clinical trials all three plasminogen activators appear to be similarly effective in removing intravascular blood clots. Consequently, the high cost of urokinase and tissue plasminogen activator, compared to streptokinase, has become the deciding factor in selecting a fibrinolytic agent. *See* BLOOD; ENZYME.

Philip C. Comp

Bibliography. F. Blasi, J.-D. Vassalli, and K. Dano, Urokinase-type plasminogen activator: Proenzyme, receptor, and inhibitors, *J. Cell. Biol.*, 104:801–804, 1987; P. M. Mannucci and A. D'Angelo (eds.), *Urokinase: Basic and Clinical Aspects*, 1982; W. J. Williams et al. (eds.), *Hematology*, 4th ed., 1990.

Uropygi

An order of arachnids, the tailed whip scorpions, comprising about 70 species from tropical and warm temperate Asia and the Americas. Most are dark reddish-brown, of medium to giant size, 0.72–2.6 in. (18–65 mm), the largest one being *Mastigoproctus giganteus* of the southern United States and Mexico. The elongate, flattened body bears in front a pair of greatly thickened, raptorial pedipalps set with many sharp spines and used to hold and crush insect prey. The first pair of legs is elongated and modified into feelers. The abdomen terminates in a slender, many jointed, whiplike flagellum. The uropygids are harmless, nocturnal creatures without poison glands that live in dark places and burrow into the soil. When disturbed, they expel a volatile liquid, with the strong odor of acetic acid, from a gland at the base of the tail. This accounts for the name "vinegaroon" given by many Americans to these much-feared animals. *See* ARACHNIDA.

Willis J. Gertsch

Uropygial gland

A relatively large, compact, bilobed, secretory organ located at the base of the tail (uropygium) of most birds having a keeled sternum. It is known also as

the preen, oil, or scent gland. This is the only true skin gland possessed by this class of vertebrates. The gland develops from a pair of invaginations of the dorsal ectoderm on each side of the free coccygeal vertebrae, giving rise to numerous secretory tubules surrounded by a sheath of connective tissue of mesodermal origin. Experiments with duck embryos have shown that formation of the gland is dependent upon a specific action of the subjacent mesoderm on the overlying ectoderm. Implanted mesoderm from other body regions failed to induce a gland-forming response in the ectoderm. *See* EMBRYONIC INDUCTION.

Wide variations in the size, shape, and structure of this gland occur among different species: it is large in most aquatic birds; rudimentary and nonfunctional in most goatsuckers, whippoorwills, nighthawks, and some pigeons; absent in bustards, many pigeons and parrots, and in all ratite birds (such as the ostrich, cassowary, and emu), although present in the embryos of all those named. *See* RATITES.

The glandular secretion, predominantly oily and sometimes of offensive odor (musk-duck, hoopoe, petrel), is discharged through an orifice at the tip of a nipplelike protuberance often encircled by short, bristly feathers. The act of preening induces a flow of secretion from the nipple which is transferred by the beak to the body plumage. In water fowl there is some evidence that this oily secretion assists in maintaining the water-repellent quality of the feathers, either directly or by preserving their physical structure. *See* FEATHER.

The occurrence and nonoccurrence of this gland in closely related species make its use difficult to interpret. Consequently, its function has aroused much controversy for over a century. No indispensable function has yet been agreed upon. It probably has different uses in different birds. *See* AVES; GLAND; SCENT GLAND.

Mary E. Rawles

Ursa Major

The Great Bear, a circumpolar constellation for observers at mid-northern latitudes or higher latitudes (see **illustration**). Part of this constellation is the prominent asterism (star pattern) known as the Big

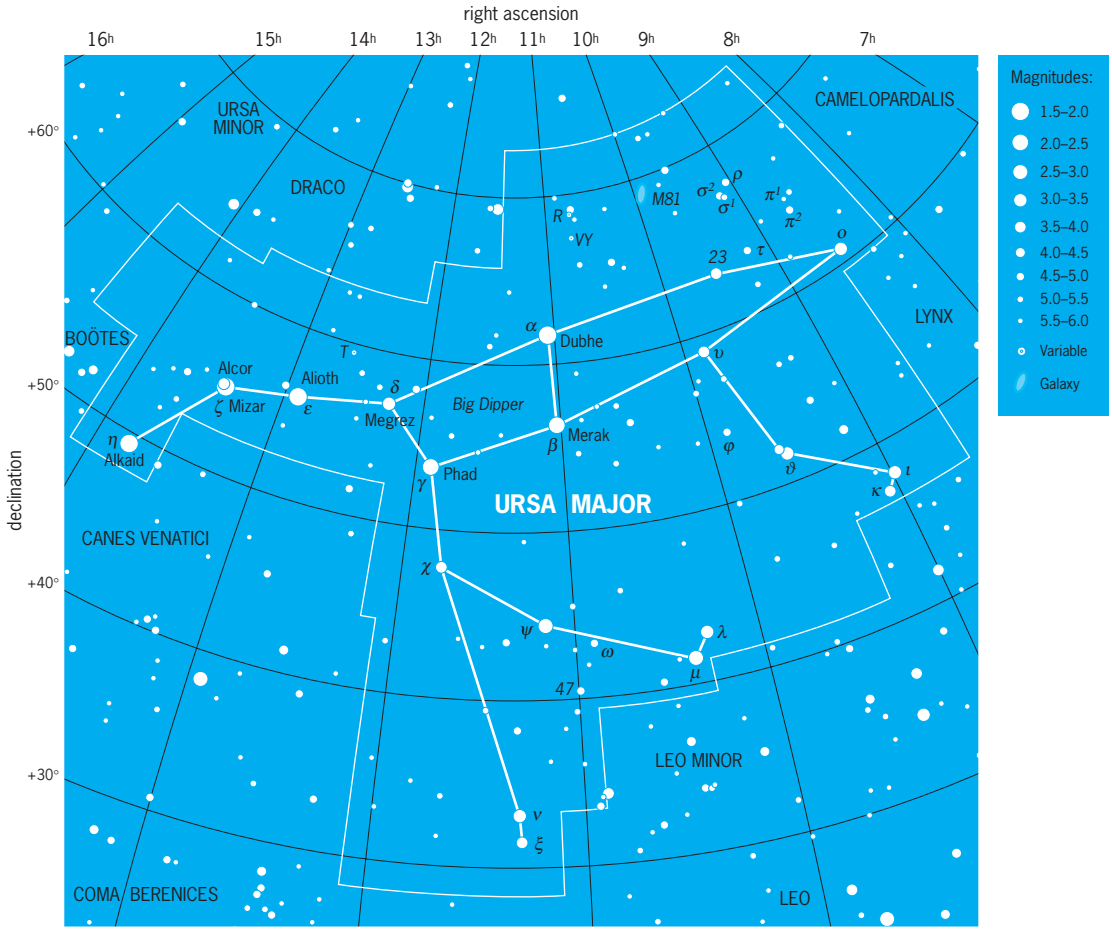

Modern boundaries of the constellation Ursa Major, the Great Bear. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

Dipper, whose bowl of four stars is linked to a tail of three more, perhaps representing a long bear's tail pulled out when the bear was swung by it to throw it into the heavens. In England, the Big Dipper is more usually known as Charles's Wain (Wagon) or the Plough.

The stars of the Dipper were labeled by Johan Bayer (1601) in alphabetical order with Greek letters, contrary to his usual practice of assigning Greek letters in order of brightness. The two stars at the head of the dipper, alpha and beta Ursae Majoris, point toward Polaris, a star that is now within 1° of the north celestial pole. Thus these pointers are very useful for formal and informal navigation. *See* POLARIS.

In the handle of the Big Dipper, the middle star Alcor has a fainter companion, Mizar, that is separately visible to the unaided eye. On inspection with a telescope, each is a double star. *See* MIZAR.

In Greek mythology, the bear is Callisto, a woman with whom Zeus had a son, perhaps changed in form by Zeus to protect her from the jealousy of Hera, Zeus's wife. In other versions of the myth, Callisto was changed into a bear in revenge by Callisto's companion Artemis or by Hera.

M81 is a prominent spiral galaxy in the northern part of Ursa Major, thus making it visible year-round from mid-northern latitudes. *See* GALAXY, EXTERNAL; MESSIER CATALOG.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Ursa Minor

The Little Bear, a circumpolar constellation from northern latitudes (see **illustration**). Part of this constellation is the prominent asterism (star pattern) known as the Little Dipper, whose stars are fainter than the corresponding stars of the Big Dipper. At the end of the handle of the Little Dipper is Polaris, at magnitude 2.0 not a particularly bright star—only the 49th brightest in the sky—but one that happens in the present epoch to be close to (almost 1° away from) the north celestial pole, the extension of the Earth's axis of rotation into the sky. Thus the Little Dipper and Polaris are useful for formal and informal navigation. *See* POLARIS.

In Greek mythology, Arcas, the son of Zeus and Callisto, was about to shoot a bear, not knowing that it was his mother. To protect her, Zeus changed Arcas into a bear and hurled both bears off by their tails to the heavens. Hera, annoyed at their being protected, then convinced Poseidon not to allow the bears to bathe in the sea, which resulted in their being circumpolar constellations, which never drop to the horizon. *See* URSA MAJOR.

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

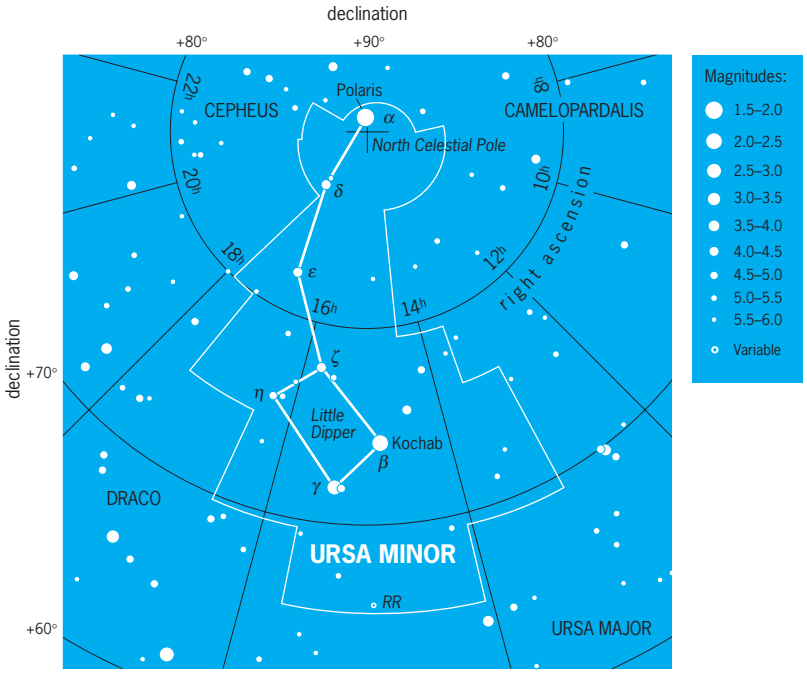

Modern boundaries of the constellation Ursa Minor, the Little Bear. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

Ustilaginomycetes (smut)

A class of the subdivision Basidiomycotina. These microscopic plant-parasitic fungi are commonly known as smut fungi. Many species may have a shorter or longer saprophytic life cycle, capable of yeastlike reproduction. Several nonparasitic yeasts have been identified as closely related to the smuts. About 1500 true Ustilaginales species are known.

The body (mycelium) of the smut fungi consists of hyphae that are thin, transparent, branched, septate, and binucleate. Usually parasitic, the fungus grows in the host tissues and gives little, if any, evidence of its presence before spore formation sets in. The smut spores (teliospores or ustilospores) are the organs of dispersion and resistance, and are thick walled, pigmented, and variously ornamented. They are formed in great number in the sori, which consist of host tissues, spore masses, and sometimes modified fungal cells or tissues.

The smut fungi develop variable sori in different organs (such as the roots, stems, leaves, inflorescences, flowers, anthers, and seeds) filled by powdery or agglutinated, usually colored (black, brown, violet, or yellow) spore masses. The spores may develop singly, in pairs, or in aggregates of spore balls. These balls may be composed entirely of fertile

spores, or of a combination of spores, sterile cells, or hyphae. Spore germination results in a basidium (promycelium or ustidium) which gives rise to hyaline basidiospores. *See* BASIDIOMYCOTA; EUMYCOTA; FUNGI.

Until recently, the Ustilaginales have been divided into two families: Ustilaginaceae (with septate basidia producing lateral and apical basidiospores) and Tilletiaceae (with unseptate basidia producing apical basidiospores). In the new classification, based also on ultrastructural and molecular biological characters, the true smut fungi are divided into 2 classes, 8 orders, 17 families, and 62 genera.

Many smut fungi produce severe losses of cereals and other cultivated plants. Some important smut diseases are dwarf bunt and the two stinking or common bunts of wheat, barley, and other Gramineae (caused by *Tilletia contraversa, T. caries*, and *T. foetida*), karnal or partial bunt of wheat (caused by *Tilletia/Neovossia indica*), partial bunt of rice (caused by *Tilletia/Neovossia horrida*), corn smut (caused by *Ustilago maydis*), loose smut of wheat, barley, and oat (caused by *U. tritici, U. nuda,* and *U. avenae*), covered smut of barley (caused by *U. hordei*), sugarcane smut (caused by *U. scitaminea*), and potato smut (caused by *Thecaphora solani*). Kálmán Ványky

Bibliography. International Mycological Institute, *C.M.I. Description of Pathogenic Fungi and Bacteria*, 1964– ; K. Ványky, *Illustrated Genera of Smut Fungi*, 1987.

Uterine disorders

Neoplastic, functional, congenital, and inflammatory diseases of the uterus and cervix. Uterine and cervical disorders are common gynecological complaints, which can be categorized in terms of increased uterine size, abnormal bleeding, and infection.

Size. The most common cause of increased uterine size is the presence of fibroids (leiomyoma). These benign smooth-muscle tumors are the most common tumors in women of reproductive age, occurring in approximately one-third of women. Fibroids vary in size from less than an inch to the equivalent of a term pregnancy. Problems arise when the fibroid presses on nearby organs such as the urinary tract. Fibroids may also cause infertility or pregnancy loss. Rarely, fibroids become malignant. With any increase in uterine size, pregnancy must be excluded and medical attention sought. Large fibroids must be surgically removed (myomectomy), and the uterus may have to be removed as well (hysterectomy). Small fibroids may be treated medically by turning off ovulation, which lowers estrogen levels and shrinks the tumors. Fibroids usually become smaller after menopause.

Abnormal bleeding. Abnormal uterine bleeding, especially in postmenopausal women, must always be evaluated. In girls and young women it is usually caused by irregular ovulation or hormonal imbalances. In women of reproductive age, problems of pregnancy must always be excluded such as miscarriage or ectopic pregnancy. Other common causes include polyps of the uterus or cervix. These benign, fleshy tumors should be removed to exclude cancer, but are usually of no consequence. Occasionally, the uterine lining (endometrium) invades the muscle of the uterus, making the uterus less able to compress the surrounding arteries during menstruation. This condition, known as adenomyosis, typically causes heavy, painful menses. *See* MENSTRUATION.

The most serious reason to seek medical attention for abnormal uterine bleeding is to exclude the possibility of cancer. Uterine and cervical cancers are the most common cancers of the female genital system. Uterine cancers usually occur in women who are obese, are hypertensive, and have diabetes mellitus; cervical cancer is most common among women who have engaged in sexual intercourse at an early age with multiple partners, especially uncircumsized ones. The diagnosis of uterine cancer is made by an endometrial biopsy or dilatation and curettage. The cornerstone of diagnosing cervical cancer is the Pap smear, which examines the cells of the cervical canal. Noninvasive or minimally invasive cancers may be treated by surgery alone. More advanced cancer requires either radiation or chemotherapy as well.

Infection. Uterine infections are not extremely common but are associated with trauma, childbirth, or miscarriage. The inflammation may be associated with a foul discharge, pain, or bleeding. Cervical infections and inflammations (cervicitis) are common, and symptoms include a foul discharge and staining or bleeding. Most inflammations are nonspecific but may become chronic. Specific agents of cervicitis include staphylococci, streptococci, gonococci, syphilis, chancroid, tuberculosis, chlamydia, and mycoplasma; appropriate antibiotics should be chosen to treat such infections. *See* REPRODUCTIVE SYSTEM DISORDERS; STAPHYLOCOCCUS; STREPTOCOCCUS; SYPHILIS; TUBERCULOSIS; UTERUS. Machelle M. Seibel

Bibliography. H. Melchoir, D. Walwiener, and G. Schlag (eds.), *Gynecology and Obstetrics*, 1994.

Uterus

The hollow, muscular womb, being an enlarged portion of the oviduct in the adult female. An adult human uterus, before pregnancy, measures approximately 3 × 2 × 1 in. (8 × 5 × 3 cm) in size and has the shape of an inverted, flattened pear. The paired Fallopian tubes enter the uterus at its upper corners; the lower, narrowed portion, which is called the cervix, projects into the vagina (see **illus.**). Normally the uterus is tilted slightly forward and lies behind the urinary bladder.

The lining, or mucosa, responds to hormonal stimulation, growing in thickness with a tremendous increase in blood vessels during the first part of the menstrual cycle. If fertilization does not occur, the thickened vascular lining is sloughed off, producing the menstrual flow at the end of the cycle, and a new

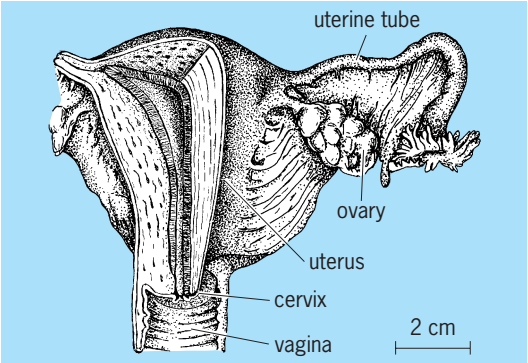

Human uterus and associated structures. (*After L. B. Arey, Developmental Anatomy, 7th ed., Saunders, 1965*)

menstrual cycle begins with growth of the mucosa. When pregnancy occurs, the mucosa continues to thicken and forms an intimate connection with the implanted and enlarging placenta. However, the maternal and fetal blood vessels remain distinct; food, oxygen, and wastes flow between these two circulatory systems, although maternal and fetal bloods do not mix. *See* MENSTRUATION; PREGNANCY.

The uterus develops from two embryonic tubes, which may remain distinct in the adult, forming two uterine chambers as in the rabbit; or may fuse to various degrees, forming a branched two-chambered uterus with a single lower portion, as in the cow; or may unite to form a single chamber, as in humans. *See* REPRODUCTIVE SYSTEM.

Walter Bock

Vaccination

Active immunization against a variety of microorganisms or their components, with the ultimate goal of protecting the host against subsequent challenge by the naturally occurring infectious agent.

The terms "vaccine" and "vaccination" were originally used only in connection with Edward Jenner's method for preventing smallpox, introduced in 1796. In 1881 Louis Pasteur proposed that these terms should be used to describe any prophylactic immunization. Vaccination now refers to active immunization against a variety of bacteria, viruses, and parasites (for example, malaria and trypanosomes). [**Tables 1** and **2**].

During the last 200 years, vaccination has controlled nine major diseases: smallpox, diphtheria, tetanus, yellow fever, pertussis, poliomyelitis, measles, mumps, and rubella. In the United States and other developed countries, where programs have been expanded to immunize virtually all children under 2 years of age, most vaccine-preventable diseases have been reduced by 95% or more from prevaccine highs (**Table 3**). In the case of smallpox, the dream of eradication has been fulfilled. Vaccinations against influenza, hepatitis B, pneumococci, and hemophilus influenza type b have made major headway against these infections. *See* SMALLPOX.

Implicit within Jenner's method of vaccinating against smallpox was the recognition of immunologic cross-reactivity together with the notion that protection can be obtained through active immunization with a different, but related, live virus. It was not until the 1880s that the next immunizing agents, vaccines against rabies and anthrax, were introduced by Pasteur. Two facts of his experiments on rabies vaccines are particularly noteworthy.

First, Pasteur found that serial passage of the rabies agent in rabbits resulted in a weakening of its virulence in dogs. For the next 100 years, Pasteur's empirical approach for attenuating the virulence of a live virus by repeated passages in cells of species different from the natural host remained the principal empirical method for developing attenuated-live-virus vaccines. During multiple passages in an animal or in tissue culture cells, mutations accumulate as the virus adapts to its new environment. These mutations adversely affect virus reproduction in the natural host, resulting in lessened virulence. Only as the molecular basis for virulence has begun to be elucidated by modern biologists has it become possible to deliberately remove the genes

TABLE 1. Some bacterial vaccines licensed in the United States

Vaccine	Year first described	Vaccine type
Typhoid	1896	Inactivated
Cholera	1896	Inactivated
Plague	1897	Inactivated
Pertussis	1926	Acellular toxoid
Bacillus Calmette-Guérin (tuberculosis)	1927	Live attenuated
Diphtheria	1888	Toxoid (protein)
Tetanus	1888	Toxoid (protein)
Pneumococcus Meningococcus Hemophilus influenza b	Post World War II	Capsular polysaccharide

TABLE 2. Some vaccines used in prevention of viral diseases of humans

Disease	Condition of virus	Route of administration	Comment
Poliomyelitis (Sabin)	Live attenuated	Oral	Recommended immunization for general human population
Poliomyelitis (Salk)	Inactivated	Subcutaneous	
Measles	Live attenuated	Subcutaneous	
Rubella*	Live attenuated	Subcutaneous	
Mumps	Live attenuated	Subcutaneous	
Mumps (Finland)	Inactivated	Subcutaneous	
Hepatitis A†	Inactivated	Intramuscular	Immunization recommended only during epidemic, travel, military service, exposure, or for selected high-risk groups
Smallpox	Vaccinia virus, an apothogenic virus related to variola	Intradermal	
Yellow fever	Live attenuated	Subcutaneous	
Influenza	Live attenuated	Oropharyngeal	
Influenza	Inactivated	Subcutaneous	
Rabies	Inactivated	Subcutaneous	
Adenovirus infection	Live attenuated	Oral by enteric coated capsules	
Japanese B encephalitis	Inactivated	Subcutaneous	
Eastern and western equine encephalitis	Inactivated	Subcutaneous	
Chickenpox‡	Live attenuated	Subcutaneous	
Hepatitis B (Heptavax-B)§	Subunit (HBsAg)	Subcutaneous	
Hepatitis B (Institut Pasteur)¶	Subunit (HBsAg) including Pre-S peptide	Subcutaneous	
Hepatitis B	Subunit (recombinant DNA from yeast)	Subcutaneous	
Cytomegalovirus disease	Live attenuated	Subcutaneous	Experimental vaccine being tested on pediatric nurses and renal transplant candidates

*Not to be administered to a postpubertal woman unless she is not pregnant and understands that it is imperative she not become so for at least 3 months after vaccination.
†Available for children over 2 years of age and for adults. Recommended for travelers to regions of high endemic disease, children living in countries with high prevalence, health-care workers, users of illicit intravenous drugs, food handlers, toddlers in day care and their supervisors, recipients of liver transplants, and persons older than 30 years with chronic liver disease.
‡Used for leukemic children and other high-risk patients.
§High-risk groups in the United States are health-care workers, selected patients on hemodialysis, children with hemophilia, residents and staff of institutes for the mentally handicapped, homosexually active males, prostitutes, users of illicit drugs, prisoners, certain military personnel, and infants and young children immigrating from high epidemic areas.
¶Used in Taiwan in 1985 to immunize all newborn babies who were at risk for infection from their mothers.

promoting virulence so as to produce attenuated viruses.

Second, Pasteur demonstrated that rabies virus retained immunogenicity even after its infectivity was inactivated by formalin and other chemicals, thereby providing the paradigm for one class of noninfectious virus vaccine, the "killed"-virus vaccine.

Attenuated-live and inactivated vaccines are the two broad classifications for vaccines. Anti-idiotype antibody vaccines and deoxyribonucleic acid (DNA) vaccines represent innovations in inactivated vaccines. Recombinant-hybrid viruses are novel members of the live-virus vaccine class recently produced by genetic engineering.

Attenuated-live vaccines. Because attenuated-live-virus vaccines reproduce in the recipient, they provoke both a broader and more intense range of antibodies and T-lymphocyte-associated immune responses than noninfectious vaccines. Live-virus vaccines have been administered subdermally (vaccinia), subcutaneously (measles), intramuscularly (pseudorabies virus), intranasally (infectious bovine

TABLE 3. Past maximum and current morbidity from vaccine preventable diseases

Disease	Number of cases: Maximum (year)	Number of cases: In 1998
Diphtheria	206,939 (1921)	1
Measles	894,134 (1941)	100
Mumps	152,209 (1968)	666
Pertussis	265,269 (1934)	7405
Polio (paralytic)	21,269 (1952)	1
Rubella	57,686 (1969)	364
Congenital rubella syndrome	20,000 (1964-65)	7
Tetanus	1,560 (1923)*	41

*Deaths in death registration states only in 1923.

rhinotracheitis), orally (trivalent Sabin poliovirus), or by oropharyngeal aerosols (influenza). Combinations of vaccines have also been used. When combined live measles, rubella, and mumps vaccines have been given by injection, the antibody response to each component has been comparable with the antibody response to the individual vaccines given separately.

Many organisms produce their pathophysiologic effects through acute infections localized to the gastrointestinal, pulmonary, nasopharyngeal, and genitourinary surfaces. These areas are bathed in mucus that contains highly important secretory immunoglobulin A (IgA). Subcutaneous, intravenous, and intramuscular immunization regimens, though effective in inducing IgG-type antibodies, are almost universally ineffective at the induction of secretory IgA antibodies. In contrast, oral administration is a feasible alternative to stimulate the common mucosal immune system, as with the live-attenuated polio vaccine. Live-virus vaccines administered through a natural route of infection often induce local immunity, which is a decided advantage. However, in the past, attenuated-live virus vaccines have been associated with several problems, such as reversion to virulence, natural spread to contacts, contaminating viruses, lability, and viral interference. *See* IMMUNOGLOBULIN; VIRULENCE; VIRUS CLASSIFICATION; VIRUS INTERFERENCE.

Noninfectious vaccines. Noninfectious vaccines include inactivated killed vaccines, subunit vaccines, synthetic peptide and biosynthetic polypeptide vaccines, oral transgenic plant vaccines, anti-idiotype antibody vaccines, DNA vaccines, and polysaccharide-protein conjugate vaccines. With most noninfectious vaccines a suitable formulation is essential to provide the optimal antigen delivery for maximal stimulation of protective immune responses. Development of new adjuvant and vector systems is pivotal to produce practical molecular vaccines.

Killed vaccines. The inactivated forms of the toxins produced by the tetanus and diphtheria bacteria, called toxoids, were among the first vaccines developed. For human use, toxoid vaccines must be highly purified to remove extraneous materials. Killed-virus vaccines consist of partially purified virus particles whose infectivity has been destroyed by treatment with chemicals or with radiation. These virus particles retain the ability to elicit an immune response and to protect from clinical disease. Though generally safer and more stable than attenuated-live-virus vaccines, killed-virus vaccines require multiple doses of high concentrations of antigen sometimes administered with adjuvant, that is, a substance that enhances the potency of the antigen. The induction of cellular immune responses may be poor and the protection induced by killed-virus vaccines may be of short duration, so that booster vaccinations are needed. Vaccines produced by killing viruses have been known to contain surviving, infectious particles that have caused outbreaks of disease (Cutter polio vaccine, foot-and-mouth-disease vaccine). Although incapable of producing direct cellular destruction, noninfectious vaccines may elicit abnormal hypersensitivity reactions.

Through the use of the oral, attenuated-live (Sabin) vaccine and the inactivated (Salk) vaccine, paralytic polio has been virtually eliminated from the United States. The last naturally occurring case of paralytic polio in the United States was in 1979. In 1997 and 1998, there were only six reported cases. All were associated with receipt of the trivalent oral, attenuated-live vaccine. An all-inactivated polio vaccine schedule has been recommended for routine childhood vaccination beginning January 1, 2000. *See* CELLULAR IMMUNOLOGY; DIPHTHERIA; POLIOMYELITIS; TETANUS.

In the 1970s, concerns developed that whole-cell pertussis vaccines caused permanent brain damage in some children. This led to a sharp decline in pertussis vaccination, waning immunity, and outbreaks of disease in several countries. Even though the general risk of serious postvaccination neurologic damage was not confirmed, a strong feeling persisted that safer and more efficacious pertussis vaccines were needed. Acellular pertussis vaccines consisting of pertussis toxoid (and in some cases, highly purified filamentous hemagglutinin and pertactin proteins) were recommended in the United Kingdom, Sweden, and Japan, and were approved in 1996 in the United States. *See* WHOOPING COUGH.

Polysaccharide vaccines. Polysaccharide vaccines consist of the polysaccharide coats, or capsules, of encapsulated bacteria. Although polysaccharides are recognized mainly by T-cell-independent mechanisms, they do not effectively produce high-level, high-affinity antibodies, nor do they induce the T-cell memory required for booster responses. However, the immunogenicity of polysaccharides can be significantly increased by covalently linking them to carrier proteins. This approach is thought to work through the recruitment of T-cell help. For example, Hemophilus influenza b vaccines developed with this conjugate approach have been very successful in inducing high, boostable, protective antibody levels in infants and have dramatically reduced Hemophilus influenza b–associated disease in countries where they are in general use. *See* INFLUENZA.

Subunit vaccines. Subunit vaccines consist of immunogenic viral proteins stripped free from whole virus particles, then purified from other irrelevant components, thereby reducing the risk of adverse reactions and residual infectious virus. Subunit vaccines require the addition of adjuvants or their formation into forms such as lipid vesicles (liposomes, virosomes) or micelles. The inability to propagate hepatitis B virus in tissue culture provided an incentive for the production of one of the first subunit vaccines, Heptavax-B, which was obtained by purifying and formalin-inactivating the surface antigen from the plasma of chronic carriers of hepatitis B. Experimental subunit vaccines have also been produced from purified glycoproteins derived from

the virus membranes of enveloped virus particles, such as parainfluenza-3, measles, rabies, influenza, respiratory syncytial virus (RSV), feline leukemia virus, bovine herpes virus-1, and pseudorabies. *See* LIPOSOMES.

Immune stimulating complexes (ISCOMs) are very immunogenic and represent an interesting delivery system for glycoprotein subunit vaccines. These complexes are really advanced liposomes that look like small balls and are made from an adjuvant.

Synthetic peptides. Chemically synthesized peptides are perceived as one alternative to conventional vaccines to elicit virus-neutralizing antibodies, or to "prime" the host for neutralizing antibody responses. A conventional approach to preparing antipeptide antibodies is conjugation of a peptide to a known protein or synthetic polymer carrier. Methods designed to avoid the use of carrier by polymerizing or cyclizing the peptides have also been reported, as has an approach known as the multiple-antigen peptide system, in which a small peptidyl core matrix is covalently bound to radially branching synthetic peptides as dendritic arms.

The peptide sequences chosen for synthesis in developing new vaccines correspond to antigenic determinants (epitopes) on the surface of virus particles that have been identified through chemical, electron-microscopic, and crystallographic analysis. Application of recombinant DNA technology has also allowed nucleotide sequences of viral genomes to be determined, from which primary amino acid sequences, secondary structure, and potential B-cell antigenic and T-cell (amphipathic) sequences may be deduced.

Carrier-free polypeptide vaccines are anticipated to be especially safe because they are chemically well defined, do not contain nucleic acids or extraneous proteins, and are unlikely to have any of the pathogens that might be present in serum or tissue extracts. Polypeptide vaccines are also extremely stable, withstanding thermal extremes and ambient conditions that might inactivate other types of vaccines under field conditions. This is important where target populations in underdeveloped countries may lack access to refrigeration equipment.

Synthetic peptide vaccines may be feasible for agents for which protection is based mainly on humoral (antibody-mediated) immunity and suitable conserved antigenic sites can be found. Synthetic peptide vaccines are more realistic for DNA than for ribonucleic acid (RNA) viruses, since DNA viruses mutate much less frequently than RNA viruses. Thus, antibodies against a single antigenic site of an RNA virus may allow too many escape mutants and, in contrast to DNA viruses, may require antibodies against different antigenic sites. Successful synthetic peptide vaccines have been produced for the simple, DNA-containing canine parvovirus and mink enteritis virus, which readily induce full protection of all vaccinated animals after a severe challenge with virulent viruses. *See* CANINE PARVOVIRUS INFECTION; FELINE PANLEUKOPENIA; PEPTIDE.

Biosynthetic polypeptide vaccines. The nucleic acid sequences encoding microbial antigens and chemically synthesized peptides can be cloned in bacterial and yeast plasmids or in viral vectors and expressed efficiently in prokaryotic and eukaryotic cells, either as genetically purified (cloned) proteins or as parts of fusion proteins. The immunogens made in this way may be secreted into the medium or may be expressed on the surface or in the cytoplasm of cultured cells. The antigens can then be purified from the culture medium or from cell lysates. The biosynthetic production of antigens is particularly useful when the microbial agent is difficult to grow in tissue culture, such as human hepatitis B virus, or dangerous, such as rabies or HIV.

Anti-idiotype antibody vaccines. The binding site of an antibody possesses a highly distinctive three-dimensional structure (the idiotype) that can itself act as an antigenic determinant to stimulate the formation of antibodies. The anti-idiotype antibodies, therefore, mimic the form of the surface antigen initially employed to induce the formation of a specific idiotypic antibody.

To prepare such vaccines, a monoclonal antibody is first produced specifically against a microbial antigen. This monoclonal antibody is injected into an animal to induce a second antibody specific to the monoclonal antibody. If the second antibody binds to the idiotype of the initial monoclonal antibody, then the antigen-recognition portion of the second antibody should be structurally analogous to the original microbial neutralization antigen. Injection of the second antibody as a vaccine will then induce antibodies against the original microbial agent. Experimental anti-idiotype antibody vaccines have been made against a number of diseases including hepatitis B virus, rabies virus, HIV, herpesvirus, poliovirus, and several bacterial and parasitic pathogens. *See* ANTIBODY.

DNA vaccines. DNA (gene) vaccines, or the direct injection of DNA plasmids that express antigens of interest, represent another approach to subunit vaccines. In animal models, DNA vaccines stimulate cell-mediated as well as antibody-based immune responses, unlike most conventional inactivated vaccines. The DNA vaccines persistently express genes encoded by the plasmid without integrating into chromosomal DNA. Gene expression from the direct intramuscular injection of DNA has been detected in mice, rats, rabbits, chicks, dogs, fish, cattle, and nonhuman primates, illustrating the diversity of species in which direct DNA vaccination without a delivery vehicle is possible. DNA vaccines have a number of advantageous properties. They can protect against viral infections in which the antibody response alone is not protective, or where there is pronounced antigenic diversity among the target strains. Protection against lethal challenges has been achieved through parenteral or mucosal immunizations, or by the use of gene guns that deliver tiny amounts of DNA-coated gold beads. The same or similar DNA plasmids can be used as vectors for subsequent immunizations, because no immune response is elicited

to the vector. The development of new DNA vaccines may be pertinent to several diseases, including acquired immune deficiency syndrome (AIDS) and hepatitis B, and will undoubtedly receive increasing attention. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS).

Adjuvants. Adjuvants are chemicals that significantly enhance speed, vigor, and persistence of strong antigens and the potency of weak antigens. Adjuvant formulations consist of adjuvants in suitable delivery vehicles, such as mineral or vegetable oil emulsions, squalene, liposomes, nonionic block polymer surfactants, and biodegradable polymer microspheres. Biodegradable microspheres can be especially useful as vehicles because of the ease of vaccine delivery by the oral route and the safety of the delivery vehicle (tissue compatibility and no secondary reactions). The most common microspheres range from a few micrometers to 200 μm in diameter.

The most common adjuvants used in immunological research are Freund's incomplete adjuvant, which consists of an aqueous antigen solution emulsified with an equal volume of paraffin oil/mannide monooleate (Arlacel A), and Freund's complete adjuvant (muramyl dipeptide delivered in oil emulsions), which also contains killed mycobacteria. The complete adjuvant consistently stimulates high and long-lasting antibody responses, which can be attributed to the slow release of antigen from the site of injection. Neither form of adjuvant is acceptable for human use because both induce pain, abscess formation, granulomas, synovial joint lesions, and other adverse reactions. The only adjuvants authorized in the United States for human use are alum salts, which are both ineffective with some antigens and capable of producing adverse reactions, such as granulomas at the injection site. The emergence of subunit vaccines created by recombinant DNA technology and of synthetic peptide vaccines has intensified the need for safer and more effective adjuvants.

Cytokines and immunomodulation. The induction of an immune response in the presence of an appropriate antigen initiates reciprocal interactions between antigen presenting cells, such as dendritic and Langerhans cells (and possibly macrophage) and T lymphocytes bearing receptors for processed forms of the antigens. The development and propagation of the immune response also involves diverse subtypes of T cells and B lymphocytes, natural killer and lymphokine activated killer cells, neutrophils, and eosinophils. The communications between these immune and inflammatory cells are mediated in large part by small proteins (about 20 kilodaltons) named cytokines. The cytokines include the interleukins (IL-1 to IL-18), gamma interferon, granulocyte-macrophage colony stimulating factor (GM-CSF), tumor necrosis factor, and transforming growth factor beta. The cytokines have various actions, for example, IL-1 (T- and B-cell maturation), gamma interferon [helper T cell (Th)1 upregulation, enhanced major histocompatibility complex expression], IL-2 (Th1 upregulation), IL-4 (Th2 upregulation), IL-10 (suppression of inflammatory cytokines, enhancement of B-cell proliferation), GM-CSF (co-migratory signals for dendritic cells), and IL-12 (stimulation of Th1 differentiation). Cytokines are species-specific and expensive, and there are concerns regarding their stability, toxicity, and potential autoimmunity. These issues have limited their use as vaccine components. *See* CYTOKINE; IMMUNOLOGY.

Immunomodulation refers to the ability of many adjuvants to modify the cytokine network and, hence, the immune response. Adjuvants can target macrophages of the Peyer's patches (lymph nodules of the intestine) to induce mucosal immune responses, or allow antigens to associate directly with Class I major histocompatibility complex (MHC) molecules on the cell surface to elicit Class I–restricted CD8+ cytotoxic T-lymphocyte responses. Particulate antigens tend to target macrophages, and macrophages can be recruited to the site of antigen deposition through the action of cytokines. Freund's complete adjuvant and lipopolysaccharide-based adjuvants drive the responses of Th1 subtype CD4+ helper cells, while alum salt, cholera toxin, and pertussis adjuvants drive the responses of Th2 subtype CD4+ helper cells. The Th1 responses are associated with interferon gamma, IL-2, and IL-12 secretions and T cells involved with strong delayed-type hypersensitivity reactions. Th1 cells also secrete tumor necrosis factor, thereby enhancing B-cell responses that elicit IgG2a, the antibody subclass most efficient in binding the serum complement proteins that enhance antigen-antibody reactions. The Th2 cells secrete the IL-4, IL-5, and IL-10 that activate the production of high levels of IgG1, IgA, and IgE by B cells. Selection of the appropriate immunoregulatory adjuvant not only leads to an enhanced immune response but also determines the isotype of IgG, which other immunoglobulins are made, and how much CD4+-directed cell-mediated immunity is generated. *See* IMMUNOLOGIC CYTOTOXICITY.

Recombinant viral vectors and chimeric viruses. Another option is the use of recombinant viruses for the delivery of vaccine antigens. Although foreign antigens may be expressed in these recombinant viruses and used as inactivated subunit vaccines, live-attenuated-viral-based vectors as vaccines have decided advantages. This is because a single vaccination without adjuvants can suffice for immunization by the live-virus vaccine. Replication of the viral-based vector within the host amplifies the amount of foreign antigen being expressed, thereby often increasing both cell-mediated and humoral immune responses to the foreign antigen. The processing and posttranslational modification of the foreign antigen within infected eukaryotic cells more closely resemble those occurring during natural infection with the pathogen from which the foreign antigen was derived. In addition, presentation of the antigen may be optimized.

Attenuated bacteria, such as *Salmonella*, as well as many different virus families have been used as vectors, including adenoviruses, herpesviruses,

papovaviruses, retroviruses, and picornaviruses. However, the most extensively developed and exploited vector has been vaccinia virus, a member of the Poxvirus family. Vaccinia is a very large virus; about 5% of its DNA either is redundant or encodes genes that are not absolutely essential for virus replication in cultured cells. Thus, the capacity of vaccinia recombinants to accommodate foreign DNA is very large. Experimental vaccinia-based-vector vaccines include those with foreign DNA inserts derived from hepatitis B virus, influenza, malaria, Epstein-Barr virus, genital herpesvirus, respiratory syncytial virus, Lassa fever virus, human T-cell leukemia virus, and AIDS virus. One of the most interesting of the vaccinia-based viral vectors is the recombinant that expresses the rabies virus glycoprotein gene. The vaccinia/rabies recombinant was incorporated in bait as an oral vaccine and is presently being used for the immunization of wild life. Although vaccinia is a comparatively safe vaccine virus, immunosuppressed people can become systemically infected and die with generalized vaccinia infection. Therefore, it is difficult to imagine vaccinia being used in humans in view of the global AIDS epidemic. An additional problem arises from the induction of vaccinia antibodies during the initial usage of the vaccinia vector so that it becomes difficult to use the vectorial vaccine twice.

Recombinant DNA techniques have also made it possible to construct poliovirus type 1/type 3 antigenic hybrid viruses in cell cultures to elicit both type 1 and type 3 neutralizing antibodies. Three poliovirus serotypes are known, and each serotype displays three neutralizing antigenic sites. A type 1/type 3 chimeric virus constructed from clones of Sabin strain derivatives would be useful in the primary vaccination of children to reduce vaccine-associated cases of poliomyelitis. *See* ANIMAL VIRUS; POLIOMYELITIS.

Diarrheal disease vaccines. Diarrheal diseases are caused by the double-stranded, segmented RNA enteroviruses, by the single-stranded RNA Norwalk-type caliciviruses, and by enterotoxigenic bacteria, such as *Vibrio cholerae, Escherichia coli*, and *Shigella dysenteriae*. These pathogens cause about 2.5 million deaths each year in children under 5 years of age, especially in developing countries where food and water are frequently contaminated. Inactivated vaccines and purified toxoids have traditionally been used to prevent cholera (Table 1) and other diarrheal diseases. Recently, however, genetic engineering techniques have enabled scientists to develop new oral vaccines, such as those obtained by genetically engineering the deletion of virulence factors from *Shigella*, *Salmonella*, and *V. cholerae*. *See* CHOLERA; DIARRHEA; ENTEROVIRUS; ESCHERICHIA; GENETIC ENGINEERING; RIBONUCLEIC ACID (RNA).

Rotavirus vaccines. Rotaviruses are the most common cause of severe diarrhea and dehyration in infants. In the developing countries, rotavirus diarrhea causes about 700,000 deaths per year. In the United States, about one-third of pediatric hospital admissions are due to rotaviruses, with about 20 deaths per year. Rotaviruses elaborate an enterotoxin (nonstructural protein NSP4) that triggers a signal transduction pathway that alters epithelial cell permeability and chloride secretion. Natural protective immunity against severe rotavirus disease is built up during the first 2–3 years of life. Studies with live attenuated oral rotavirus vaccines have also shown that the majority of severe episodes of rotavirus diarrhea are preventable by oral immunization. An oral tetravalent rotavirus vaccine was licensed in the United States in August 1998. This vaccine consists of a live-attenuated Rhesus rotavirus serotype 3 combined with human-rhesus reassortants expressing serotypes 1, 2, and 4. The federal government recommended that every Infant in the United States get the protective vaccine. Over 1 million children swallowed three doses of the vaccine, at 2, 4, and 6 months of age. However, in October 1999 the recommendation for routine vaccination of United States infants was withdrawn, and the manufacturer voluntarily withdrew the vaccine from the market, after it was found that 56 children developed a painful and potentially fatal bowel obstruction within 1 week of receiving the vaccine. Several other candidate rotavirus vaccines, including bovine-human reassortants, are currently being investigated. *See* INFANT DIARRHEA.

Oral transgenic plant vaccines. These vaccines exemplify the most recent innovative approach to vaccine development. Ultimately, the vaccine designers hope to use common, uncooked, edible plants, such as bananas and tomatoes, as transgenic plant delivery systems to express the antigens of pathogenic organisms in the gut, so as to induce secretory IgA, the predominant form of immunoglobulin found in mucosal secretions. Presently, however, only model studies have been completed using transgenic potatoes and tobacco plants. Transgenic plants have many potential advantages as vaccines. They are easy to grow at minimal cost, even in the most underdeveloped countries; and unlike bacteria and animal cells, they do not require specialist media and equipment or stringent purification protocols. Transgenic plants can be conveniently self-fertilized to produce stable, true-breeding lines propagated by conventional horticultural techniques and stored or distributed as seeds. Neither sterile syringes nor nurses would be needed for vaccination. *See* BREEDING (PLANT); SOMATIC CELL GENETICS.

Human and animal diarrheas were initially targeted to demonstrate the feasibility of oral transgenic plant vaccines. Transgenic potato and tobacco plants were engineered to express the heat-labile subunit of *E. coli* enterotoxin LT-B (LT stands for heat-labile toxin) and the related subunit of the *V. cholerae* enterotoxin (CT)-B, both of which have affinity for the cell surface receptor, monosialosyl ganglioside GM1. Transgenic plants expressing the capsid protein of the Norwalk calicivirus were also constructed. To optimize the plant cells' expression of the bacterial and viral proteins, a constitutive plant promoter from cauliflower mosaic virus was used to drive transcription, and a plant-specific termination sequence derived from

a soybean vegetative storage protein gene was inserted into the constructs. Fecal IgA secretory antibodies and serum IgG antibodies, were induced by mice fed (via a stomach tube) transgenic plants. Mice immunized with potato LT-B were also partially protected after exposure to intact LT enterotoxin. Furthermore, in preliminary human trials, volunteers who ingested the transgenic potatoes expressing the LT-B subunit developed serum and secretory neutralizing antibodies with minimal adverse side reactions. In other studies, a protective antibody response to foot-and-mouth disease virus was induced in mice following oral or parenteral immunization with alfalfa transgenic plants expressing the foot-and-mouth disease virus structural protein VP1; and neutralizing antibodies were induced by mice immunized with leaf extracts from transgenic plants expressing the glycoprotein S polypeptides of swine-transmissible gastroenteritis coronavirus. Likewise, rabies virus glycoprotein G has been expressed in transgenic tomatoes.

Cancer vaccines. Materials commonly termed cancer vaccines are usually immunotherapeutic preparations targeted to increase the immune response to mutated cellular proteins. They are intended to inhibit the growth of previously diagnosed cancers rather than to prevent the appearance of new, spontaneous malignant tumors. However, in cases in which infectious agents are strongly associated with cancer induction, vaccines capable of preventing cancers have been made available. *See* CANCER (MEDICINE); TUMOR VIRUSES.

Viral hepatitis comprises common infectious diseases of humans. There are more than 350 million carriers of the hepatitis type B virus (HBV) in the world, with about 1 million in the United States. It has been estimated that HBV causes about 1 million deaths worldwide, with at least 5000 deaths in the United States, each year. Carriers of HBV are at risk of developing hepatocellular carcinomas many years after the initial HBV infection. Hepatocellular carcinoma is one of the most common cancers of the world (ranking within the top 10). In some parts of Asia, it may be the first or second most common cancer. A nationwide HBV vaccination program was implemented in Taiwan in 1984. Since the initiation of this universal vaccination program, the incidence of hepatocellular carcinoma in Taiwanese children has halved, and the mortality rate for hepatocellular carcinoma has also decreased. Vaccination programs in Korea and Gambia have also decreased the incidence of hepatocellular carcinoma. *See* HEPATITIS.

Marek's disease virus is a cell-associated herpesvirus of poultry and causes a naturally occurring contagious malignant lymphoma in chickens. The disease has been effectively controlled for more than a decade using a live vaccine based on a non-pathogenic serologically related herpesvirus of turkeys. Cancers caused by other known infectious agents (such as, Epstein-Barr virus, Hepatitis C virus) may also ultimately be preventable by vaccination. *See* AVIAN LEUKOSIS; EPSTEIN-BARR VIRUS. Saul Kit

Bibliography. C. J. Arntzen, Pharmaceutical foodstuffs: Oral immunization with transgenic plants, *Nature Med. Vaccine Suppl.*, 4:502–503, 1998; D. Baxby, *Jenner's Smallpox Vaccine: The Riddle of Vaccinia Virus and Its Origin*, 1981; B. S. Blumberg, Hepatitis B virus, the vaccine, and the control of primary cancer of the liver, *Proc. Nat. Acad. Sci. USA*, 94:7121–7125, 1997; J. C. Cox and A. R. Coulter, Adjuvants: A classification and review of their modes of action, *Vaccine*, 15:248–256, 1997; B. Duff and P. Duff, Hepatitis A vaccine: Ready for prime time, *Obst. Gynecol.*, 91:468–471, 1998; J. L. Melnick, Virus vaccines: Principles and prospects, *Bull. WHO*, 67:105–112, 1989; S. A. Plotkin and E. A. Mortimer, Jr. (eds.), *Vaccines*, 1988; C. O. Tachet et al., Immunogenicity in humans of a recombinant bacterial antigen delivered in a transgenic potato, *Nat. Med.*, 4:607–609, 1998; *Vaccine*, periodical published by Butterworth, London, six times a year.

Vacuole

An intracellular compartment, bounded by a single membrane bilayer, which functions as a primary site of protein and metabolite degradation and recycling in animals, but serves additional complex functions in fungi and plants (**Fig. 1**). Scientists who study vacuoles also define them as the terminal product of the secretory pathway. The secretory pathway functions to transport protein and metabolite-containing membrane vesicles from sites of synthesis or uptake to the vacuole. *See* CELL MEMBRANES; CELL METABOLISM.

Animal lysosomes. In animals, a lytic vacuole known as the lysosome typically functions to process macromolecules. Lysosomes contain an acidic environment and store a large set of specialized hydrolytic enzymes that serve in the breakdown of nucleic acids, sugars, lipids, and proteins. Such macromolecules can be targeted to the lysosome from sites of synthesis. For example, proteins that assemble incorrectly in the endoplasmic reticulum (ER) can be degraded in the lysosome and their

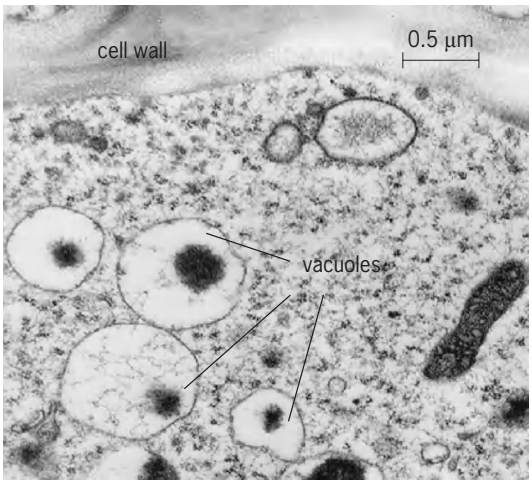

Fig. 1. Electron micrograph of a barley root tip cell, showing multiple vacuoles within the cytoplasm.

constituent amino acids recycled. Proteins that can serve as nutrients are also targeted to the lysosome from the cell surface as a result of endocytosis. An important process for the recycling of cytoplasm in eukaryotic cells is autophagy, in which molecules or organelles are encapsulated in membrane vesicles that fuse with the lysosome. *See* ENDOCYTOSIS; ENDOPLASMIC RETICULUM; LYSOSOME.

In the mammalian immune system, macrophages and neutrophils take up particles and pathogens in the process of phagocytosis, during which the pathogen is eventually digested in the lysosome. A number of diseases in humans can be caused when intracellular pathogens evade destruction in the lysosome. *Mycobacteria* species, which are responsible for severe illnesses such as tuberculosis and leprosy, undergo phagocytosis by cells of the immune system but are able to interfere with the process of transport to the lysosome for digestion. Thus, these bacteria are able to circumvent the immune system and reside as intracellular parasites. *See* PHAGOCYTOSIS.

Fungi. In fungi such as the yeast *Saccharomyces cerevisiae*, vacuoles can serve additional functions not found in animals. Besides a lytic function, they serve in the storage of ions such as sodium, calcium, potassium, and chloride, as well as amino acids for protein synthesis. Ion storage and transport across the vacuolar membrane is particularly important for maintaining the optimal osmotic balance within the cytoplasm for cellular processes. In yeast, it has been demonstrated that vacuoles can also function in the destruction and recycling of cellular organelles, such as peroxisomes, which help protect the cell from toxic oxygen-containing molecules such as peroxides. The process of peroxisome digestion by vacuoles is known as pexophagy. In fungi, many mutants are known to have defective vacuoles or no vacuoles, indicating that vacuoles are not essential under all conditions.

Plants. The most complex vacuoles are found in plants. In contrast to the vacuoles found in yeast and animals, plants contain multiple types of vacuoles with distinct functions essential for survival. Some vacuoles contain hydrolytic enzymes and store ions similar to those found in lysosomes, whereas others serve a role in storing pigments such as anthocyanins which impart color to flowers to attract insect or bird pollinators. Specialized ER-derived vacuoles in plant seeds, known as protein bodies, function in the storage of proteins called prolamines that are common in the endosperm of cereals such as rice and maize. Upon germination, the proteins are degraded and used as a source of amino acids and nitrogen for the growing plant. This is an effective survival strategy, as is the storage in seeds of toxins such as lectins and other macromolecules, such as protease inhibitors, that interfere with predation by herbivores. Such compounds are known as allelochemicals. Other toxins, such as alkaloids, are stored in vacuoles in parts of the plant, such as the leaves, which are subject to frequent herbivory. Scientists have learned that plants produce and store in their vacuoles a vast array of unique chemicals which may, in addition to their natural functions, have medicinal value. *See* PLANT CELL.

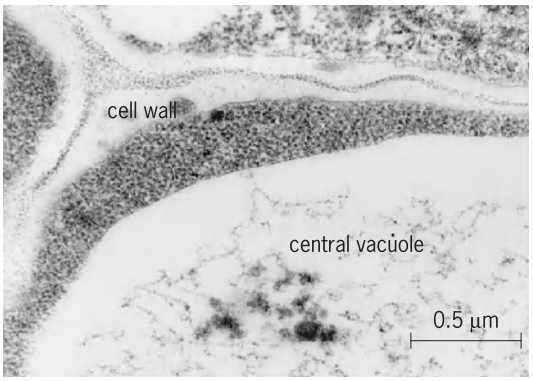

Fig. 2. Electron micrograph of a vacuole from an *Arabidopsis* root tip cell. Such large central vacuoles can accumulate osmotic pressure sufficient to provide turgor for plant cell growth.

Another unique function that vacuoles serve in plants is in cell growth. As a consequence of the accumulation of ions, metabolites, and water, plant vacuoles are under considerable internal osmotic pressure (**Fig. 2**). The vacuolar membrane in plants, known as the tonoplast, as well as the cell itself would burst under this pressure if not for the rigid wall that surrounds the cells. The resulting turgor pressure provides mechanical stability to plant stems. Loss of osmotic pressure in the vacuole due to a lack of water results in plant wilting. The osmotic pressure of the vacuoles also provides the driving force that allows plants to grow by enlarging their cell volume. Enzymes reduce the rigidity of the cell wall, which permits cell expansion under the force of turgor. This is a fundamental process in plants and explains why vacuoles can occupy as much as 95% of the volume of some cells. In fact, the turgor is sufficient to drive rapid movements. In the sensitive plant *Mimosa pudica*, rapid uptake and release of water in vacuoles of specialized cells known as pulvini are responsible for the folding of leaves when touched. *See* CELL WALLS (PLANT).

Transport to the vacuole. Vacuoles are part of a complex secretory pathway that transports protein-containing membrane vesicles from the ER to the plasma membrane or vacuole. In recent years, scientists have made considerable efforts to understand the process by which proteins are targeted to vacuoles. Proteins destined for vacuoles are synthesized at the ER, appropriately processed, then packaged into membrane transport vesicles. The vesicles fuse with the Golgi apparatus, where the proteins are repackaged into transport vesicles that are targeted to the vacuole via another membrane compartment known as the prevacuolar compartment. Each step in this process from ER to Golgi to vacuole requires the formation of new transport vesicles and the subsequent fusion of those vesicles with the next compartment. *See* GOLGI APPARATUS.

The mechanism of vesicle transport is explained by the SNARE hypothesis: Specialized proteins on

transport vesicles known as v-SNARES interact with compatible proteins known as t-SNAREs on target membranes. This interaction leads to membrane fusion. Vesicle formation requires additional specialized proteins. Although many of the basic mechanisms of protein transport are conserved between animals, fungi, and plants, some aspects are quite different. Proteins destined for the vacuole possess signals specifying their transport. In animals the sugar mannose phosphate is frequently used as a signal, whereas in fungi the targeting information resides within the sequence of amino acids specifying the protein. The signal is located at the amino-terminal end of the protein.

Plants have a diverse array of targeting signals, specified by the order of their amino acids. In fact, three different types of signals are found within plant proteins, which are located at the amino terminus, at the carboxyl terminus, or within an internal region of the protein. Other types of macromolecules appear to be transported to the vacuole via specific transport proteins located in the vacuolar membrane. For example, water is moved in and out of the vacuole via transport proteins known as aquaporins. Glenn R. Hicks; Natasha Raikhel

Bibliography. N. J. Bryant and T. H. Stevens, Vacuole biogenesis in *Saccharomyces cerevisiae:* Protein transport pathways to the yeast vacuole, *Microbiol. Mol. Biol. Rev.*, 62(1):230–247, 1998; A. Haas, Reprogramming the phagocytic pathway: Intracellular pathogens and their vacuoles, *Mol. Mem. Biol.*, 15:103–121, 1998; D. J. Klionsky and Y. Ohsumi, Vacuolar import of proteins and organelles from the cytoplasm, *Annu. Rev. Cell Dev. Biol.*, 15:1–32, 1999; S. Kornfeld, The biogenesis of lysosomes, *Annu. Rev. Cell Biol.*, 5:483–525, 1989; F. Marty, Plant vacuoles, *Plant Cell*, 11:587–599, 1999; A. A. Sanderfoot and N. V. Raikhel, The specificity of vesicle trafficking: Coat proteins and SNAREs, *Plant Cell*, 11:629–641, 1999; M. Wink, The plant vacuole: A multifunctional compartment, *J. Exp. Bot.*, 44(suppl.):231–246, 1993.

Vacuum cleaner

An electrically powered mechanical appliance for the dry removal of dust and loose dirt from rugs, fabrics, and surfaces. Portable vacuum cleaners are widely used in domestic and industrial cleaning of surfaces that cannot be wiped or brushed, such as carpets, upholstery, tapestry, or highly contoured surfaces. Small hand vacuum cleaners are used for cleaning the inside of passenger cars. Large industrial tank cleaners are used in cleaning shops. Power saws may be equipped with vacuum cleaners which collect sawdust as it is made. Schools and offices have systems of centrally located dust collectors and fans for creating vacuum; pipes lead to room outlets at which cleaning appliances are attached.

In operation, atmospheric pressure forces air which carries with it loose dirt and dust into the cleaning tool. The dirt-laden air then passes through a bag which retains the dirt. The air is exhausted through an electric fan and an additional filter. The fan provides the low internal air pressure, commonly called the vacuum.

Design of this machine centers on convenience of handling, features including a flexible air hose and tools for cleaning special surfaces and crevices, with considerable variety in location or motion of various assemblies. A critical feature is the filter action of the bag, which may be of cloth (in which case it can be emptied) or of paper (in which case it is disposable). Because at the exhaust the fan provides air at a slight pressure, the appliance may include attachments for spraying.

The great advantage of the vacuum cleaner is its retention of dirt in the filter bag. However, the force of the draft is limited by the use of atmospheric pressure. For this reason, local industrial cleaning of work in process may be by compressed air. *See* COMPRESSOR; VACUUM PUMP. Frank H. Rockett

Vacuum fusion

A technique of analytical chemistry for determining the oxygen, hydrogen, and sometimes nitrogen content of metals. The method can be applied to a wide variety of metals, the alkali and alkaline earth metals being exceptions. The range of the method extends from 1% down to a few parts per million for oxygen and nitrogen and down to fractional parts per million for hydrogen.

The metal sample is either fused or dissolved in a bath, or flux, of a second metal in a heated graphite crucible supported inside an evacuated glass or quartz vessel (see **illus.**). Oxygen is released from the metal as carbon monoxide by reaction of oxides or dissolved oxygen with carbon from the graphite crucible at high temperature. Metal nitrides dissociate, although not always quantitatively, to form elemental nitrogen. Hydrogen is evolved as elemental hydrogen.

The mixture of carbon monoxide (CO), nitrogen (N_2), and hydrogen (H_2) is analyzed to determine individual component concentrations by one of several techniques, including micro-Orsat, mass-spectrometric, and vacuum-manometric procedures. The last is in most general use. Gas quantities are determined by measurement of the pressure of the gas after confinement into a small calibrated volume. The product of the pressure and volume is

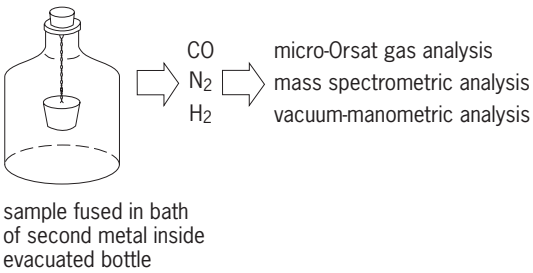

Vacuum-fusion methods of analysis.

proportional to the total number of moles of gas present, regardless of the species of gas molecules. Measurement is first made of the total quantity of CO, N_2, and H_2 collected. These gases are then oxidized to carbon dioxide (CO_2), water (H_2O), and N_2 by passing over hot copper oxide. The CO_2 and H_2O can be removed from the gas mixture successively by selective freezing or by use of chemical absorbents. The decrease in total moles of gas accompanying removal of each species permits determination of the number of moles of CO_2, N_2, and H_2O.

Furnace temperatures and bath conditions must be selected for each metal or alloy to ensure quantitative recovery of oxygen. The required furnace temperatures range from 1650°C (3000°F) for iron, nickel, and low-alloy steels to as high as 1900°C (3500°F) for metals such as titanium, zirconium, tantalum, and thorium. Iron, nickel, and low-alloy steels are usually fused without addition of other metals. High-melting metals require use of a second metal to lower their melting points. The second metal is used as a previously prepared bath or, in some instances, is added simultaneously with the sample metal. Iron, nickel, platinum, and tin are the most generally used fluxing metals.

Nitrogen is not always quantitatively evolved, and values obtained by vacuum-fusion methods for nitrogen are not generally reliable. Hydrogen is always evolved quantitatively upon fusion of the metal in a vacuum. When hydrogen alone is of interest, simpler and more rapid techniques can be used, involving vacuum extraction from metal heated to temperatures well below the melting point.

Vacuum-fusion analysis of gases and metals has found its widest industrial application, from a process-control viewpoint, in titanium, tantalum, niobium, zirconium, and molybdenum. Determination of hydrogen in steel is a routine procedure in some applications. As a research tool, the vacuum-fusion techniques have been applied to practically all of these materials except the alkali and alkaline-earth metals.

Another method, inert gas fusion, was developed for determination of gases in metals. The techniques used are similar to vacuum fusion but substitute inert gases for the vacuum environment. This method is sometimes referred to as argon-fusion. The methods of detection of the gases involved during the fusion are modified somewhat. *See* GAS AND ATMOSPHERE ANALYSIS; VACUUM METALLURGY. Frank C. Benner

Bibliography. A. Choudhury, *Vacuum Metallurgy*, 1990; J. D. Winefordner and M. M. Bursey (eds.), *Treatise on Analytical Chemistry*, vol. 11, pt. 1: *Theory and Practice*, 2d ed., 1989.

Vacuum measurement

The determination of a gas pressure that is less in magnitude than the pressure of the atmosphere. This low pressure can be expressed in terms of the height in millimeters of a column of mercury which the given pressure (vacuum) will support, referenced to zero pressure. The height of the column of mercury which the pressure will support may also be expressed in micrometers. The unit most commonly used is the torr, equal to 1 mm (0.03937 in.) of mercury (mmHg). Less common units of measurement are fractions of an atmosphere and direct measure of force per unit area. The unit of pressure in the International System (SI) is the pascal (Pa), equal to 1 newton per square meter (1 torr = 133.322 Pa).

Atmospheric pressure is sometimes used as a reference, particularly for relatively crude measurements. The pressure of the standard atmosphere is 29.92 in. or 760 mm of mercury (101,325 Pa or 14.696 lbf/in.^{2}); therefore, using atmospheric pressure as a reference, an absolute pressure of 10 in. of mercury would be expressed as 19.9 in. of mercury vacuum.

In the laboratory, measurement of vacuum is important because the vacuum level has a significant effect on most physical, chemical, and biological processes.

In industry, vacuum level is commonly measured and controlled to maintain uniformity of product and as a guide to safe operation. Vacuum measurement is used, for example, in the manufacturing of television sets (evacuation of picture tubes); in metallurgy (treatment of metals attacked by common gases); and in the pharmaceutical industry (distilling heat-sensitive compounds). *See* PROCESS CONTROL.

Pressures above 1 torr can be easily measured by familiar pressure gages, such as liquid-column gages, diaphragm-pressure gages, bellows gages, and bourdon-spring gages. At pressures below 1 torr, mechanical effects such as hysteresis, ambient errors, and vibration make these gages impractical. *See* MANOMETER; PRESSURE MEASUREMENT.

Pressures below 1 torr are best measured by gages which infer the pressure from the measurement of some other property of the gas, such as thermal conductivity or ionization. The thermocouple gage, in combination with a hot- or cold-cathode gage (ionization type), is the most widely used method of vacuum measurement today. *See* IONIZATION GAGE.

Other gages used to measure vacuum in the range of 1 torr or below are the McLeod gage, the Pirani gage, and the Knudsen gage. The McLeod gage is used as an absolute standard of vacuum measurement in the 10–10^{-4} torr (10^3–10^{-2} Pa) range. *See* MCLEOD GAGE; PIRANI GAGE.

The Knudsen gage is used to measure very low pressures. It Measure pressure in terms of the net rate of transfer of momentum (force) by molecules between two surfaces maintained at different temperatures (cold and hot plates) and separated by a distance smaller than the mean free path of the gas molecules. Richard Comeau

Bibliography. T. A. Delchar, *Vacuum Physics and Technology*, 1993; D. Hucknall, *Vacuum Technology and Applications*, 1993.

Vacuum metallurgy

The making, shaping, and treating of metals, alloys, and intermetallic and refractory metal compounds in a gaseous environment where the composition and partial pressures of the various components of the gas phase are carefully controlled. In many instances, this environment is a vacuum ranging from subatmospheric to ultrahigh vacuum (less than 760 torr or 101 kilopascals to 10^{-12} torr or 10^{-10} pascal). In other cases, reactive gases are deliberately added to the environment to produce the desired reactions, such as in reactive evaporation and sputtering processes and chemical vapor deposition. The processes in vacuum metallurgy involve liquid/solid, vapor/solid, and vapor/liquid/solid transitions. In addition, they include testing of metals in controlled environments. The **table** lists the processes, pressure ranges, and gas environments.

The early applications of vacuum metallurgy were involved with the more conventional aspects, such as extractive metallurgy, melting, and heat treatment. Since the 1970s, two important developments have occurred. First, other nonvacuum specialty melting processes, such as electroslag remelting (ESR) and plasma melting, have been developed. They have some unique aspects (for example, grain size refinement) which make them strong competitors with vacuum melting processes on the one hand and complementary melting processes on the other. For many applications, the final ingot may have been sequentially processed by a number of vacuum and nonvacuum melting processes. Second, in many high-technology applications, a single material, such as a metal or alloy, cannot meet the multiple requirements placed on it for a particular application. For example, the blades and vanes for the hot stage of a gas turbine engine have to be a composite material, with high-temperature nickel or cobalt base alloy for strength at elevated temperatures and a coating of an alloy such as nickel (or cobalt)-chromium-aluminum-yttrium to resist the severe corrosion environment. Thus a large and rapidly growing aspect of vacuum metallurgy consists of vapor deposition techniques.

Advantages. There are three basic reasons for vacuum processing of metals: elimination of contamination from the processing environment, reduction of the level of impurities in the product, and deposition with a minimum of impurities. Contamination from the processing environment includes the container for the metal and the gas phase surrounding the metal. For example, melting of titanium in a refractory crucible in air would severely contaminate the melt both from the material of the crucible and from the atmospheric oxygen, making it useless. Therefore titanium is processed in water-cooled copper crucibles (which do not react with the metal) and in vacuum. In the vacuum process, impurities, particularly oxygen, nitrogen, hydrogen, and carbon, are released from the molten metal and pumped away; and metals, alloys, and compounds are deposited with a minimum of entrained impurities.

There are numerous and varied application areas for vacuum metallurgy; some of these are discussed below.

Various metallurgical processes and their operating pressure regimes*

Processes	Approximate pressure regime, torr (Pa)	Gas environment†
Extraction and refining of metals	Several atm to 10^{-3} torr (0.133)	V
Degassing of liquid steel	400 to 1 (53,200 to 133)	V
Vacuum induction melting	100 to 10^{-6} (13,300 to 1.33×10^{-4})	V
Vacuum arc melting	1 to 10^{-3} (133 to 0.133)	V, I
Electron-beam melting	10^{-3} to 10^{-8} (0.133 to 1.33×10^{-6})	V
Levitation melting	760 to 10^{-5} (101,080 to 1.33×10^{-3})	V, I
Melting on water-cooled hearths	760 to 10^{-5} (101,080 to 1.33×10^{-3})	V, I
Zone refining	760 to 10^{-8} (101,080 to 1.33×10^{-6})	V, I
Surface treatment (carburizing, nitriding)	10^{-1} to 10^{-3} (13.3 to 0.133)	R
Heat treatment:		
High vacuum	1 to 10^{-6} (133 to 1.33×10^{-4})	V
Ultrahigh vacuum	10^{-8} to 10^{-10} (1.33×10^{-6} to 1.33×10^{-8})	V
Evaporation		
For purification and bulk deposits	10^{-2} to 10^{-7} (1.33 to 1.33×10^{-5})	V
For thin-film deposition	10^{-1} to 10^{-10} (13.3 to 1.33×10^{-8})	V, R
Sputtering to form thin-film or bulk deposits	10^{-2} to 10^{-5} (1.33 to 1.33×10^{-3})	I, R
Chemical vapor deposition	760 to 10^{-2} (101,080 to 1.33)	V, R
Electron-beam welding (except where beam is brought out to atmosphere)	$<10^{-4}$ ($<1.33 \times 10^{-2}$)	V
Brazing	10^{-2} to 10^{-5} (1.33 to 1.33×10^{-3})	V
Diffusion bonding	10^{-2} to 10^{-5} (1.33 to 1.33×10^{-3})	V
Evaluation of physical and mechanical properties	Several atm to 10^{-9} torr (1.33×10^{-7})	V
Study of nature of surfaces and reactions occurring on clean surface	$<10^{-8}$ ($<1.33 \times 10^{-6}$)	V
Mechanical working in vacuum	10^{-1} to 10^{-4} (13.3 to 1.33×10^{-2})	V
Sintering of powders in vacuum	10^{-1} to 10^{-5} (13.3 to 1.33×10^{-3})	V
Electron-beam machining	$<10^{-4}$ ($<1.33 \times 10^{-2}$)	V

*After R. F. Bunshah, *Techniques of Metals Research*, vol. 1, John Wiley and Sons, 1968.
†V = residual gases in a vacuum; I = inert gas added; R = reactive gas added.

Extractive metallurgy. Included in extractive metallurgy are the beneficiation of ores, reduction of ores or compounds to crude metals, and subsequent refining of the high-purity metals. Examples are the reduction of dolomite with ferro-silicon, refining of rare-earth metals by distillation, and preparation of pure metals by halide dissociation on a heated wire.

Melting processes. The melting processes have several functions. Among them are the refining of crude metals, the consolidation of various components into alloys, and the production of an ingot of desired shape.

The metal may be heated and melted by several heating methods which distinguish the various processes, for example, induction melting, consumable-electrode arc melting (see **illus.**), and electron-beam melting. In addition, there is the levitation melting technique where no crucible is used, the metal being supported by an electromagnetic force.

Vacuum melting and allied processes are (1) titanium alloy scrap recycling to gravimetrically separate high-density inclusions such as broken tungsten carbide tool bits and other heavy impurities; (2) removal of low-density inclusions and production of a titanium alloy ingot from scrap turnings, scrap solids, virgin sponge, aluminum, and master alloys; (3) development of a single merit process, vacuum-electroslag, to replace the two melt process for high-temperature superalloys; it uses a thin slag layer of calcium oxide/aluminum oxide on the molten pool, which reduces nitrogen and hydrogen pick-up and oxidation of the metal from the slag; melt cleanliness is improved since the slag acts as a filter; and (4) nitrogen-alloyed steels as a substitute for the more expensive nickel-alloyed steels.

Casting of shaped products. The molten metal is cast into intricate shapes by using a mold of refractory compounds usually made by the lost-wax process. A good example of this process is the production of components for aircraft jet engines by using a nickel-base high-temperature alloy. *See* METAL CASTING.

Vacuum degassing of molten steel. This is the largest single application of vacuum metallurgy. In this process, steel melted in conventional electric arc

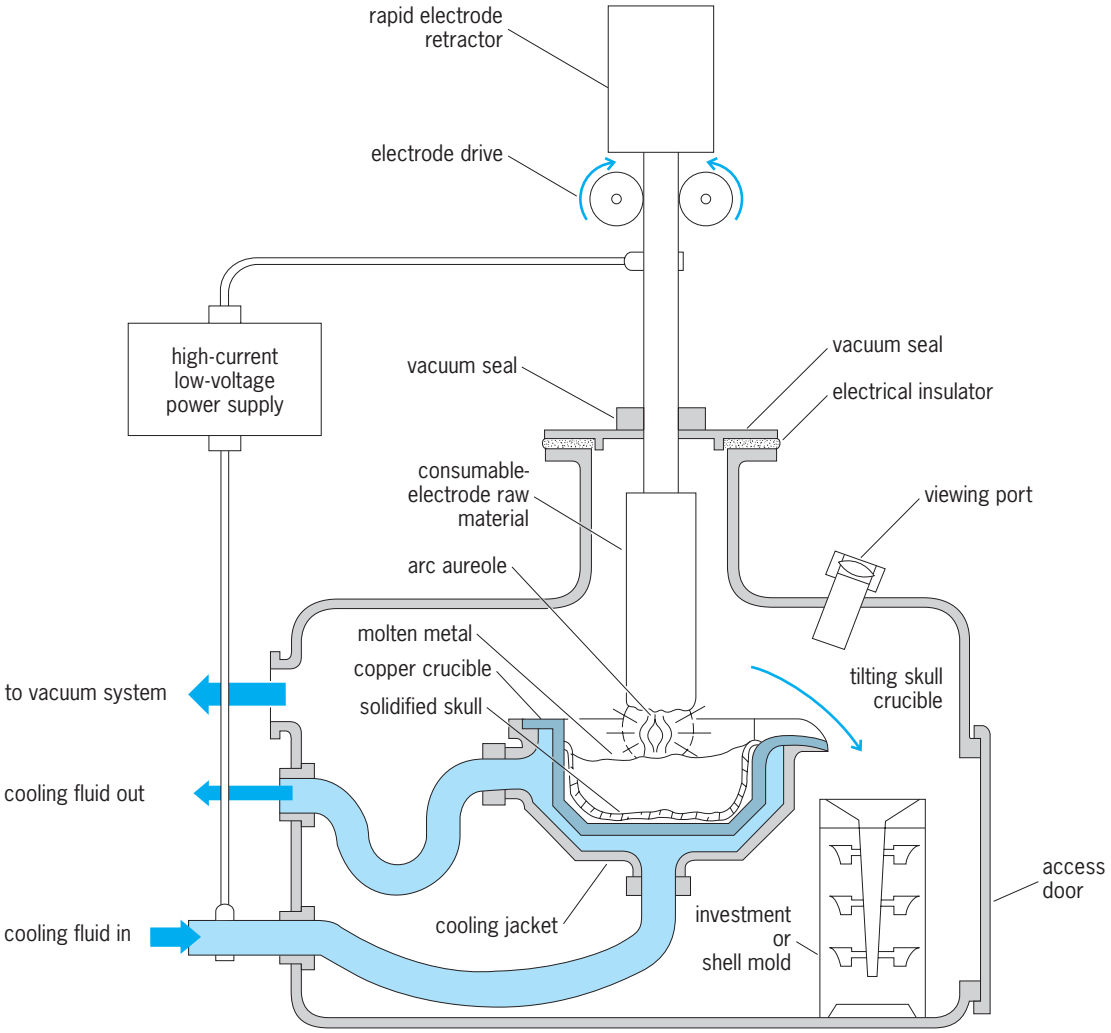

Vacuum metallurgy. Skull-crucible, consumable-electrode vacuum arc melting and casting furnace. (*After R. F. Bunshah, Techniques of Metals Research, vol. 1, John Wiley and Sons, 1968*)

furnaces is treated in vacuum before being cast into ingots. There are several processes for vacuum degassing. During the process the gases in the molten metal are released and pumped away. *See* STEEL; STEEL MANUFACTURE.

Heat treatment. Heat treatment of refractory and reactive metals (titanium, columbium, tungsten, and so forth) is carried out in vacuum to promote desired reactions in the solid metal while preventing recontamination from the atmosphere; an example is age-hardening of titanium alloys to get maximum strength.

Surface treatment. Modification of the surface properties of a material is carried out by diffusing various species such as carbon, nitrogen, boron, or aluminum from a solid, liquid, or gaseous environment. The effects of impurities in the gas phase will affect the surface properties, and hence the need for a controlled environment. In some instances ionization of the reactive gas permits the process to occur more rapidly and at lower temperatures, for example, ion nitriding and ion carburizing.

Joining. Joining processes include welding, brazing, and diffusion bonding. They are carried out in vacuum where the metals being joined would be contaminated by joining in air. A development of particular importance is electron-beam welding. *See* ARC WELDING; WELDING AND CUTTING OF MATERIALS.

Vapor deposition. This type of process includes the physical vapor deposition—evaporation, ion plating, and sputtering—and the chemical vapor deposition. The products of these processes are coatings and free-standing shapes such as sheet, foil, and tubing of thickness ranging from 20 nanometers to 25 mm. These processes have widespread application in diverse applications in the metallurgical, chemical, engineering, and electronic industries. A large variety of products are made possible by using high-vacuum coating processes.

Decorative. These include automotive trim (interior and exterior), toys, cosmetic packaging, pens and pencils, Christmas decorations, food and drink labels, costume jewelry, home hardware, eyeglass frames, packaging and wrapping materials, and watch cases.

Optically functional. These include laser optics (reflective and transmitting), architectural glazing, home mirrors, automotive rearview mirrors, eyeglass lenses, projector reflectors, camera lenses and filters, instrument optics, auto headlight reflectors, television camera optical elements, and meter faces.

Electrically functional. Included in this group are semiconductor devices, integrated circuits, capacitors, resistors, magnetic tape, disk memories, superconductors, electrostatic shielding, switch contacts, and solar cells.

Mechanically functional. These products are aircraft engine parts, aircraft landing gear, solid film lubricants, and tool bit hard coatings.

Chemically functional. In this group are corrosion-resistant fasteners, gas turbine engine blades and vanes, battery strips, and marine equipment.

Ultrafine powders. Ultrafine powders of metals and ceramics, 1–10 nm in diameter, are produced by converting the condensed metal phase to a vapor by evaporation or sputtering processes at inert or inert-plus-reactive gas pressures at about 100 millitorr so as to produce gas-phase condensation into ultrafine particles, which are then trapped on liquid-nitrogen-cooled surfaces. If a ceramic compound (oxide, carbide, nitride) powder is desired, the gas phase contains the appropriate reactive base (oxygen, hydrocarbon, or nitrogen) to form the compound.

Space processing. The space environment combines microgravity conditions and high vacuum. Results of experimental work in various space missions have demonstrated improvements in properties of metals, semiconductors, and glasses over similar material processed on Earth in a 1-*g* environment. The space environment also presents a more functional possibility of studying physical and chemical properties of materials. In a particular configuration known as the molecular shield, an extreme vacuum environment (less than 10^{-13} torr) with a very high pumping throughput is available in space. Such an environment is not possible on Earth. When combined with the micro-*g* environment of space, the potential for producing entirely unique materials is large. *See* SPACE PROCESSING.

Other processes. Other metallurgical operations carried out in vacuum are the testing of metals, electron-beam machining, study of surfaces, powder metallurgy, and mechanical working. *See* METAL COATINGS; POWDER METALLURGY.

The products of vacuum processing include the entire gamut of materials, steels, reactive metals (titanium, zirconium, and so forth), refractory metals (tungsten, molybdenum, and others), their alloys, intermetallic compounds, and refractory compounds. Vacuum metallurgy occupies a key role in the design of space processing programs. *See* PYROMETALLURGY; PYROMETALLURGY, NONFERROUS.

Rointan F. Bunshah

Bibliography. American Vacuum Society, *Transactions of the Vacuum Metallurgy Conference*, 1967, also other annual editions; J. A. Belk, *Vacuum Techniques in Metallurgy*, 1963; R. F. Bunshah, *Deposition Technologies for Films and Coatings: Developments and Applications,* 2d ed., 1994; R. F. Bunshah, History and current status of vacuum metallurgy, *J. Vac. Sci. Technol.*, A12(4):936–945, 1994; R. F. Bunshah (ed.), *Vacuum Metallurgy*, 1958; *J. Vac. Sci. Technol.,* 1964 to date.

Vacuum pump

A device that reduces the pressure of a gas (usually air) in a container. When gas in a closed container is lowered from atmospheric pressure, the operation constitutes an increase in vacuum in this container. The unit used for vacuum is a millimeter (mm) column of mercury (Hg). Another term for the mmHg unit is the torr. These terms are used interchangeably. The torr is named in honor of E. Torricelli, a

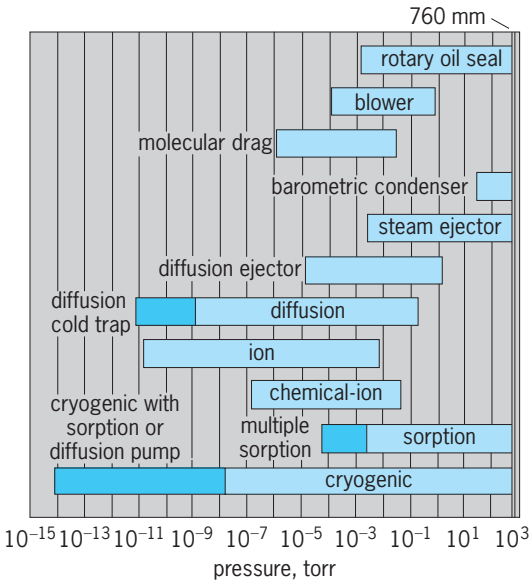

Fig. 1. Comparison of the degree of vacuum attained by important types of vacuum pumps.

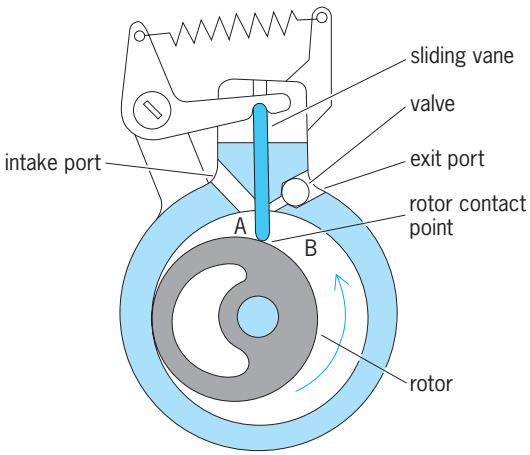

Fig. 2. Chief components of rotary oil-seal pump.

pioneer in vacuum technology. The various vacuum regions are classified as follows:

Low vacuum, 760 (atmospheric pressure) to 25 mmHg
Medium vacuum, 25 to 10^{-3} mmHg
High vacuum, 10^{-3} to 10^{-6} mmHg
Very high vacuum, 10^{-6} to 10^{-9} mmHg
Ultrahigh vacuum, 10^{-10} mmHg and beyond

A vacuum as high as 10^{-15} mmHg or one-billionth of a billionth of an atmosphere has been reached. *See* PRESSURE.

Vacuum pumps are evaluated for the degree of vacuum they can attain and for how much gas they can pump in a unit of time. **Figure 1** compares the degree of vacuum that the more important types of vacuum pumps can attain. The determination of the quantity of gas a vacuum pump handles is discussed later. At low vacuum, this quantity is usually expressed as pounds of air per hour. Any other gas is converted, via molecular weight relationship, to air equivalent for this purpose. For high vacuums, pounds of air per hour is expressed as torr liters (or mmHg liters) per second. In practice, where high vacuum is required, two or more different types of pumps are used in series. For example, a rotary oil-seal pump or steam ejector can be used with a diffusion pump to have continuity in pumping from atmospheric pressure to the high or very high vacuum range.

Mechanical pumps. One of the early vacuum pumps was similar in design to the reciprocating steam engine. A power-driven piston working in a cylinder sucked the gas out of any attached container. A more recent development is the rotary oil-seal pump (**Fig. 2**). Gas is sucked into chamber A through the opening intake port by the rotor. A sliding vane partitions chamber A from chamber B. The compressed gas that has been moved from position A to position B is pushed out of the exit port through the valve, which prevents the gas from flowing back. The valve and the rotor contact point are oil-sealed. Since each revolution sweeps out a fixed volume, it is called a constant-displacement pump.

A rotary blower pump operates by the propelling action of one or more rapidly rotating lobelike vanes. It does not use oil as a sealing medium.

The molecular drag pump is one that operates at very high speeds, as much as 16,000 revolutions/min. Pumping is accomplished by imparting a high momentum to the gas molecules by the impingement of the rapidly rotating body.

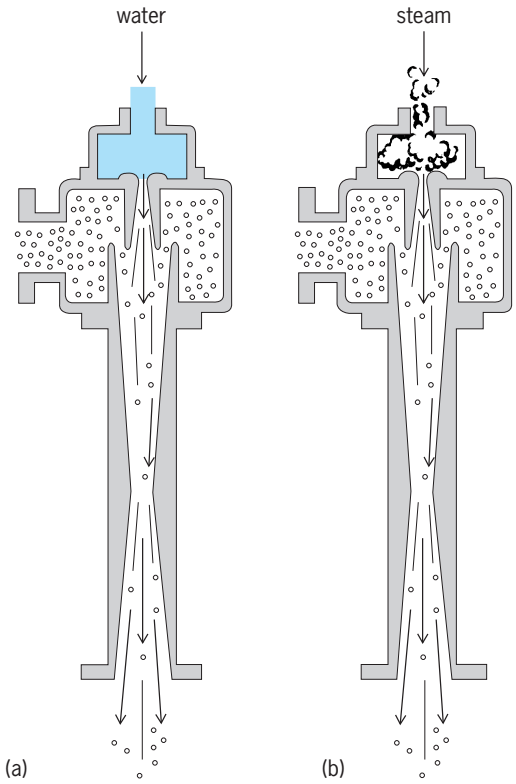

Fig. 3. Diagram of the water aspirator pump. Utilization of (*a*) water and (*b*) steam.

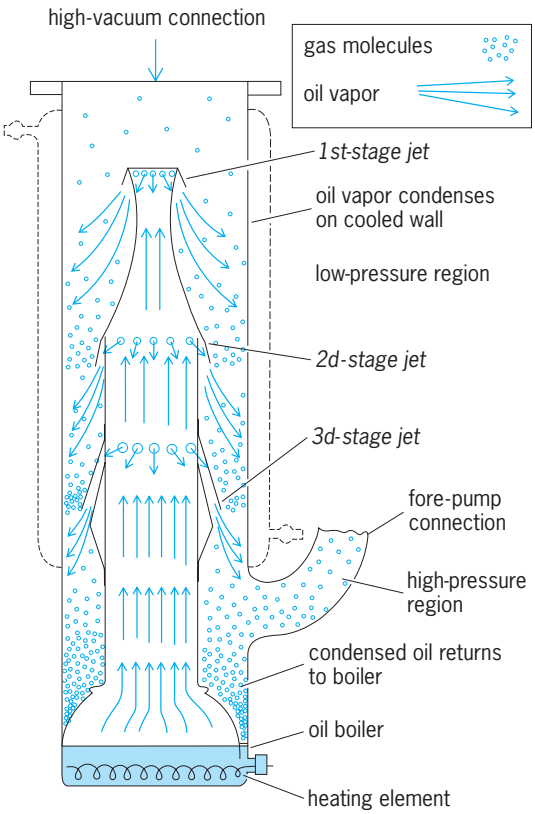

Fig. 4. Main operating features of diffusion pump.

Ejector pumps. The earliest form of this type is the water aspirator (**Fig. 3**). When water is forced under pressure through the jet nozzle, it will force the gas in the inlet chamber to go through the diffuser, thus lowering the pressure in the inlet chamber. When high-pressure steam is used instead of water, it is called a steam ejector. Steam ejectors often are combined in multiple series with barometric condensers to reach lower pressures. A barometric condenser operates by injecting water into the upper end of a vertical 35-ft (11-m) pipe whose exit is immersed in a catch basin called a hot well. The gas trapped in the inlet chamber is carried out with the water. Like the water aspirator, it can produce a vacuum (usually between 25 and 40 mmHg) equal to the vapor pressure of the water. This type of pumping is widely used in the chemical and petroleum industries.

Diffusion pumps permit high pumping speed at very high vacuum. The vapor-jet principle of the ejector pumps has been modified by using oil or mercury vapor diffusing through a jet, then condensing the vapor on the cooled wall of the pump chamber. **Figure 4** illustrates the diffusion pump operation. **Figure 5** shows a small, 1-in.-diameter (2.5-cm) diffusion pump with pumping speeds as low as 0.26 gallon/s (1 liter/s) compared for size with one of the largest pumps, 48 in. (120 cm) in diameter having a pumping speed of 26,000 gallons/s (100,000 liters/s). Sometimes to meet operating requirements one jet is built on the ejector principle. Then it is called a diffusion-ejector pump.

Specialty pumps. A number of pumps have been developed which meet special pumping requirements. In all cases it is not the principles that are new, but the way they are used. The ion pump operates electronically. Electrons that are generated by a high voltage applied to an anode and a cathode are spiraled into a long orbit by a high-intensity magnetic field (**Fig. 6**). These electrons colliding with gas molecules ionize the molecules, imparting a positive charge to them. These are attracted to, and are collected on, the cathode. Thus a pumping action takes place. Ion pumps are used for evacuating nuclear accelerators, x-ray tubes, and other specialized equipment. In a chemical-ion pump, a metal (usually titanium) is evaporated by electrical power, and the metal vapor reacts with chemically active molecules (O_2 and N_2). The reaction products condense on the walls of the pump. The inert gases that remain are removed by devices similar to that of the ion pump. Sorption pumping is the removal of gases by adsorbing and absorbing them on a granular sorbent material such as a molecular sieve held in a metal container. When this sorbent-filled container

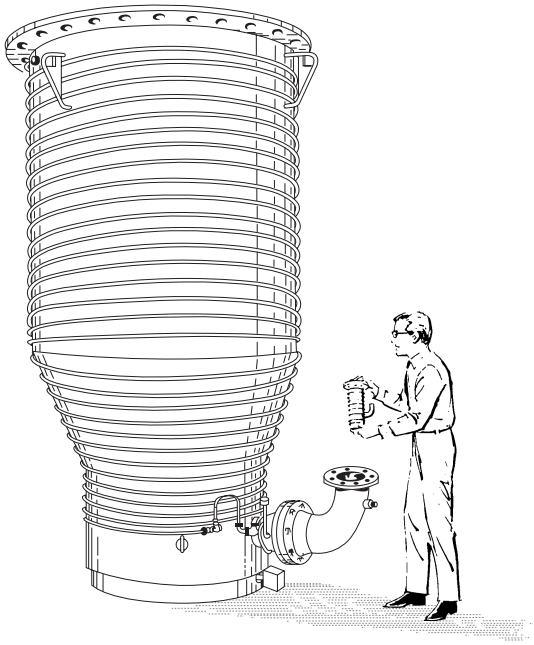

Fig. 5. Size comparison of two diffusion pumps.

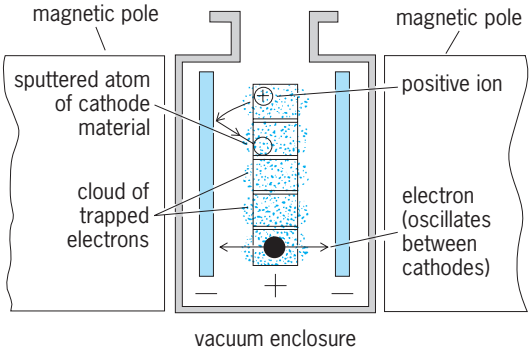

Fig. 6. Principal features of an ion pump.

is immersed in liquid nitrogen, the gas is "sorbed." This type of pump produces a quick, clean vacuum. Cryogenic pumping is accomplished by condensing gases on surfaces that are at extremely low temperatures. This is usually attained by the use of liquid or gaseous hydrogen at around −423°F (−253°C) or liquid helium at −452°F (−269°C). The equipment is very complex in design. However, it attains very high pumping speeds for low-density wind tunnels and space simulator chambers.

Vacuum pumping equations. There are two relations of practical use in vacuum technology. The first is $P_1V_1 = P_2V_2$ (at constant temperature). That is, when pressure P_1 is changed to a new pressure P_2, the volume of the gas at P_1 which is V_1 is changed to a new volume V_2.

The second is $Q = SP$. Here Q is the gas throughput (expressed as torr liters/second at a specified temperature), S is the speed of pumping (in this case liters/second), and P is the pressure (in torr) at the point where S is to be measured. The measurement of S requires a vacuum gage to obtain P and one of several methods for measuring Q. One way to determine S is to measure the amount of air at atmospheric pressure that is admitted into the vacuum chamber through a controlled, calibrated leak. This measurement gives Q, and S is then obtained from the second equation. A more precise method, especially useful for very high and ultrahigh vacuum, is the measurement of a pressure drop across an orifice of known gas conductance. The gas conductance is a mathematically determined quantity obtained from the geometry of the test portion of the vacuum chamber. *See* PUMP.

Edward S. Barnitz

Bibliography. L. G. Carpenter (ed.), *Vacuum Technology: An Introduction*, 2d ed., 1983; A. Roth, *Vacuum Technology*, 3d ed., 1990, reprint 1998.

Vacuum tube

An electron device in which use is made of the electrostatically or magnetically controlled flow of electrons in an evacuated space, or of the phenomena accompanying the electron emission, acceleration, and collection associated with the flow. In common usage, completely evacuated tubes are called hard tubes; tubes containing a small amount of deliberately added gas or vapor are called soft tubes, but they are usually included in the family of vacuum tubes.

Vacuum tubes are used as active and passive electrical circuit elements. In suitable circuitry they rectify, detect, gate, amplify, and oscillate. They can also scan and display images; generate heat, light, and radiation; ionize and accelerate atoms and molecules; and measure high vacuum.

Principle of operation. Most vacuum tubes employ the Edison effect, in which a unidirectional flow of current takes place through a vacuum by thermionic emission of electrons from a hot surface (the cathode). The electrons are collected by a nearby positively charged electrode (the anode or plate). Two-electrode Edison-effect tubes were first employed as radio-frequency detectors, but the enormous potential of the device was not realized until it was shown that the current could be controlled with large amplification ratios and great sensitivity by small electrical charges on a third, meshlike electrode (the grid) interposed between cathode and anode. From the resulting three-element tube (the triode) came all the characteristic inventions of the modern electronic age. *See* ELECTRON EMISSION; ELECTRON MOTION IN VACUUM; THERMIONIC EMISSION.

Later refinements included additional grids by which certain problems inherent in triodes were overcome, and combination tubes were developed in which several functions were combined within one vacuum envelope. Elaborations of the basic idea include camera and picture tubes and magnetrons, klystrons, and traveling-wave tubes for microwave radar. *See* KLYSTRON; MAGNETRON; MICROWAVE TUBE; PICTURE TUBE; RADAR; TELEVISION CAMERA TUBE; TRAVELING-WAVE TUBE.

Use. Originally, vacuum tubes were enclosed in cylindrical glass structures (hence the name tube). Through the walls or base of the tube were sealed the metal electrodes which emitted, controlled, or collected the electrons traversing the empty space within. Up through the 1960s, vast numbers of small tubes were used in radio and television broadcast receivers. They also found their way into industrial instruments and controls, and into the early electronic computers, but their size, fragility, inordinate power consumption, and short lifetimes made them unpopular for these more demanding purposes. Since that time, such vacuum tubes have been almost entirely replaced by the much smaller, cheaper, and longer-lived solid-state devices. Only a few standard receiving-tube types still remain in production, and are used where their electrical characteristics are particularly apt, or where environmental or circuit conditions would destroy solid-state devices. A significant exception is the cathode-ray tube (CRT), which is still made in great numbers for consumer applications. Though bulky, complex, and fragile, it is highly versatile, and so remains in favor as a display device for television, computers, and instruments of all kinds. *See* CATHODE-RAY TUBE; ELECTRONIC DISPLAY.

Other vacuum tubes in widespread use are high-power transmitting and microwave tubes; x-ray and accelerator tubes; mass spectrometer, electron microscope, night-vision, and surface-analytical tubes; and electron beam guns (partial tubes) for evaporators and welders. Since there is no prospect that any of these types will be replaced by solid-state devices, research and development on them continues, and new types occasionally appear. Such specialized tubes range in size from ultraminiature, flat devices sandwiched between glass or sapphire plates for flat displays or high-temperature uses, to giant accelerator chambers used for scientific research. In power-handling ability, they range from microwatts to megawatts. In the larger sizes especially, they may be welded, brazed, or bolted together from

ceramic-metal subassemblies. When fitted with auxiliary vacuum pumps and water-cooling systems, and integrated closely with other circuit elements for use in space, undersea, or military environments, they may be scarcely recognizable as tubes. *See* ELECTRON MICROSCOPE; IMAGE TUBE (ASTRONOMY); ION SOURCES; LIGHT AMPLIFIER; MASS SPECTROSCOPE; PARTICLE ACCELERATOR; PHOTOMULTIPLIER; PHOTOTUBE; SURFACE PHYSICS; WELDING AND CUTTING OF MATERIALS; X-RAY TUBE. Nicholas Reinhardt

Bibliography. J. W. Gewartowski and H. A. Watson, *Introduction to Electron Tubes*, 1969; A. S. Gilmour, Jr., *Microwave Tubes*, 1986; S. Okamura (ed.), *The History of Electron Tubes*, 1994; F. Rosebury, *Handbook of Electron Tube and Vacuum Techniques*, 1965, reprint 1993.

Vaginal disorders

Inflammations, trauma, tumors, congenital defects, and functional disorders of the vagina, the muscular membranous canal from the vulvar opening to the uterine cervix.

Vaginitis. The most common vaginal disorder is vaginitis, an inflammation of varying degree, characterized by itching (pruritus), typically accompanied by a discharge. Breast-feeding infants sometimes display a mild vaginal discharge due to hormones in breast milk. Young children with a vaginal discharge must be examined to exclude the presence of a foreign object or an infectious condition such as gonorrheal vulvovaginitis.

Among women of reproductive age, the three most common types of vaginitis are caused by the protozoon *Trichomonas*, the bacterium *Gardnerella*, or the monilial fungus. Postmenopausal women with little or no estrogen are most likely affected by senile vaginitis due to irritation of the thinned, atrophic mucosa. The diagnosis can usually be made by microscopic inspection of the discharge, and appropriate medication prescribed. Women with new or multiple sexual partners should also have vaginal cultures taken for chlamydia and gonorrhea. Recurrent monilia infections could be a clue to diabetes or human immunodeficiency virus (HIV) infection. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); GONORRHEA.

Other disorders. Trauma of the vagina is common. Lacerations suggest either forced intercourse or the forceful use of foreign objects. Accidental injury as a result of falling on penetrating objects may also occur. Irritation, sensitivity, or even ulceration can occur from ingredients in douches, suppositories, or lubricants, or even from a forgotten tampon.

The most important primary lesion of the vagina is carcinoma, but its incidence is low. Cancer of the cervix is much more common, and in advanced cases the cancer may extend into the vagina. Hyperpigmented areas in the vagina and labia must be monitored carefully to exclude melanoma. Benign lesions such as papillomas, fibromas, and hemangiomas are relatively uncommon and of minor consequence. *See* CANCER (MEDICINE).

Congenital malformations of the vagina are uncommon. They include a double vagina, a transverse septum, and an imperforate hymen. These conditions typically require surgical correction. Congenital absence of the vagina is important as menstrual blood cannot escape and reflexes into the pelvic cavity, a condition leading to the formation of endometriosis. Therefore, a vaginal canal must be established at the time of menarche. Congenital cysts are common but rarely significant.

The final category of vaginal disorder is vaginismus, which is characterized by spasmodic, painful, and involuntary contraction of the lower vaginal muscles. Although typically psychological in origin, it can be due to physical causes. Rarely, vaginismus is interpreted as infertility because of the inability to have intercourse. Sexual counseling is often therapeutic, and infertility can be treated by using artificial insemination. *See* REPRODUCTIVE SYSTEM DISORDERS. Machelle M. Seibel

Valence

The combining power of a chemical element for other elements as measured by the number of bonds to other atoms that one atom of the given element forms upon chemical combination; also known as valency. Valence theory concerns all the physical and chemical properties of molecules that especially depend on molecular electronic structure.

Thus, in water, H_2O or

the valence of each hydrogen atom is 1; the valence of oxygen, 2. In methane, CH_4 or

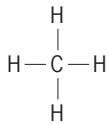

the valence of hydrogen again is 1; of carbon 4. In sodium chloride, NaCl, and carbon tetrachloride, CCl_4, the valence of chlorine is 1, and CH_2 the valence of carbon is 2.

Much more is known about a water molecule than that it contains two hydrogen atoms and one oxygen atom. Each OH distance is 9.57 nanometers and the HOH bond angle is $104°27'$. The oxygen and hydrogen ends of the molecule are negatively and positively charged, giving it a dipole moment 1.84×10^{-18} electrostatic unit (esu). The molecule absorbs infrared light strongly but is transparent to visible light. Scientists have provided quantitative understanding of these properties and many more in terms of the fundamental theory of quantum mechanics. *See* BOND ANGLE AND DISTANCE; QUANTUM CHEMISTRY.

Combining power of an element. By the 1920s the most important facts about atoms had been established experimentally. A neutral atom of atomic number Z comprises a massive nucleus of charge $+Ze$ and Z very light electrons, each of charge $-e$, where $e = 4.80 \times 10^{-10}$ esu; most of the space within the atom is empty. Atomic nuclei are immutable through ordinary chemical changes; when one molecule of hydrogen (H_2) combines with one molecule of chlorine (Cl_2) to give two molecules of hydrogen chloride (HCl), the four nuclei (two hydrogen nuclei, or protons, of charge $+1e$ and two chlorine nuclei of charge $+17e$) are unchanged. It is redistribution of electrons between atoms which constitutes chemical combination. This is what valences of atoms control, and this is what a theory of valence must explain.

Atomic structure. Understanding of molecule formation requires an understanding of the electronic structure of atoms. According to N. Bohr, electrons in an atom move in orbits much like the orbits of planets about a sun, held to the nucleus by electrical attractions for it, prevented from falling into it by centrifugal forces. A special quantum effect is operative at the atomic level, however, which possesses no analogy in the motions of planets; not all orbits are possible for an electron, but only those for which the angular momentum of the electron as it moves about the nucleus is an integer multiple of $h/2\pi$, where $h = 6.63 \times 10^{-34}$ J is Planck's constant, and for which the energy is similarly quantized. Furthermore, not more than two electrons can move in one orbit at once. *See* ATOMIC STRUCTURE AND SPECTRA; ELECTRON CONFIGURATION.

When the consequences of these ideas are worked out, there actually emerges the periodic classification of the elements. To cover just part of the periodic table, occupation of orbits by electrons in the lighter atoms are shown in the **table**, where the symbol $2p$ stands for three distinct orbits of the same energy and shape but differently oriented in space. The lowest energy orbit is $1s$, forming the K shell. Next in energy are $2s$ and $2p$, making up the L shell. The $3s$ state is still higher, in the M shell. The chemically inert gases helium (He) and neon (Ne) are characterized by closed shells of 2, and $2 + 8 = 10$ electrons, respectively. The next inert gas is argon (Ar), with a closed shell of $2 + 8 + 8 = 18$ electrons, followed by krypton (Kr) with $2 + 8 + 18 + 8 = 36$ electrons, and the others. *See* INERT GASES; PERIODIC TABLE.

Electron configuration of some atoms

			Orbit					
			K shell	L shell		M shell		
Atom		Z	$1s$	$2s$	$2p$	$3s$	$3p$	$3d$
Hydrogen	(H)	1	1	0	0	0	0	0
Helium	(He)	2	2	0	0	0	0	0
Lithium	(Li)	3	2	1	0	0	0	0
Beryllium	(Be)	4	2	2	0	0	0	0
Boron	(B)	5	2	2	1	0	0	0
Carbon	(C)	6	2	2	2	0	0	0
Nitrogen	(N)	7	2	2	3	0	0	0
Oxygen	(O)	8	2	2	4	0	0	0
Fluorine	(F)	9	2	2	5	0	0	0
Neon	(Ne)	10	2	2	6	0	0	0
Sodium	(Na)	11	2	2	6	1	0	0
Magnesium	(Mg)	12	2	2	6	2	0	0
Aluminum	(Al)	13	2	2	6	2	1	0

Rule of eight. Many of the simple facts of valence follow from the postulate that atoms combine in such a way as to seek closed-shell or inert-gas structures (rule of eight) by the transfer of electrons between them or the sharing of a pair of electrons between them. Many molecular structures may be obtained by inspection by using these rules. The electrons in the K shell are not involved in the bonding for atoms after He, nor are the electrons in the K and L shells for atoms following Ne.

Hydrogen has a valence of 1, because one more electron will give a hydrogen atom an inert-gas structure. Carbon can form four bonds, because four more electrons give it the neon electronic structure.

Bond types. The bond between two atoms is covalent if one electron in the bonding electron pair comes from each atom, as in H:H or the CH bonds in CH_4. It is coordinate covalent if both electrons come from one atom, as the boron-nitrogen bond in the compound $F_3B{-}NH_3$. If there is complete transfer of electrons from one atom to another, the bond is electrovalent or ionic, as in sodium fluoride (NaF). Bonds intermediate in type are possible; the bond in hydrogen fluoride (HF) is between covalent and ionic. An ionic bond X^+Y^- will be more stable the less the ionization potential of X and the greater the affinity of Y for electrons, that is, when X is a metallic element from the lower left corner of the periodic table and Y is a nonmetallic element from the upper right corner. Bond type can be inferred from both chemical and physical evidence. *See* CHEMICAL BONDING; ELECTRONEGATIVITY.

Bonds involving one or three electrons are known, but they are rare; H_2^+ and HeH are examples. Multiple bonds between atoms are common and important; examples are the carbon-carbon bond and the carbon-oxygen bonds in ethylene and carbon dioxide. For discussion of a bond of special importance in biology *see* HYDROGEN BOND.

Valence electrons are the electrons of an atom that can participate in chemical bonding, for example, for H and He the $1s$ electrons, for Li through Ne the $2s$ and $2p$ electrons, and for Na the $3s$ electrons.

Oxidation-reduction. As generally used, the word valence is ambiguous. Before a value can be assigned to the valence of an atom in a molecule, the electronic structure of the molecule must be known exactly, and this structure must be describable simply in terms of simple bonds. In practice, neither of these conditions is ever precisely fulfilled. A term not so ambiguous is oxidation number or valence number. Oxidation numbers are useful for the balancing of oxidation-reduction equations, but they are not related simply to ordinary valences. Thus the valence of carbon in CH_4, $CHCl_3$, and CCl_4 is 4; oxidation

numbers of carbon in these three substances are −4, +2, and +4. *See* OXIDATION-REDUCTION.

Quantum theory of valence. The above simple theory of valence is inadequate in at least three ways. First, it fails to account for many experimental facts, such as why the six C—C bonds in the molecule benzene, C_6H_6, are physically and chemically equivalent, what the electronic structures of the boron hydrides are, why the H—H bond is much stronger than the C—C bond, why CO_2 is a linear molecule but H_2O nonlinear, and what principles govern the rates of chemical combination. Second, the explanations that are offered are not physically satisfying. The stability conferred upon a molecule by the sharing of a pair of electrons by two atoms is established, but what is the real origin of this stability? And third, the theory is not detailed or quantitative enough to allow correlation and prediction of the many different properties of molecules. Dozens of properties of molecules can be measured, many to a high degree of accuracy. The ultimate theory should account for all of these quantitatively. *See* MOLECULAR STRUCTURE AND SPECTRA.

The quantum theory of valence does not possess these faults. It is based on the precise laws of physics for the atomic domain that were formulated in the 1920s by E. Schrödinger and others, the discipline known as quantum mechanics. The primitive quantum ideas of M. Planck and N. Bohr require modification to take care of the experimental fact that electrons and other particles at times act like waves. Like waves, they interfere when they are on top of one another in a manner that can be precisely calculated. According to nineteenth-century physics, an electron moving about a proton would collapse onto it. In the Bohr theory this collapse is prevented by a special quantum hypothesis; in the new mechanics it is prevented by elementary energy considerations. It would be favored by the attractive potential energy of the particle pair, but it turns out to be catastrophic for their kinetic energy. Instead of collapse a compromise is reached; the electron, or wave, is smudged out over a region surrounding the nucleus which defines the atomic size. *See* QUANTUM MECHANICS; UNCERTAINTY PRINCIPLE.

The pattern of the periodic table comes out as before. The orbits of Bohr are replaced by other entities, orbitals, which represent not the paths of the electrons but the amplitudes of the electron waves at different points in space. Furthermore, electrons are treated as if they were spinning, but only in two possible ways. The rule that generates the periodic table then is that in an atom no two electrons can occupy the same atomic orbital with the same spin. *See* EXCLUSION PRINCIPLE.

In a chemical bond, again there is interplay of kinetic and potential energies. An electron pair will tend to be shared by two atoms instead of being located on one of them if that situation is energetically favorable. The region between nuclei is more favorable for the potential energy of electron-nuclear attraction than other regions the same distance from just one nucleus. Moving in this region is not as favorable for the kinetic energy as moving on individual atoms, but the potential energy predominates when a bond is formed. The normal covalent bond may be described as two electrons occupying one molecular orbital, rather than two distinct atomic orbitals, with opposite spins because the exclusion principle is still operative. *See* MOLECULAR ORBITAL THEORY.

When a detailed examination is made of these effects with the modified theory, the stabilities of actual molecules and other of their properties are quantitatively accounted for. In particular, if two atoms approach which have low-energy atomic orbitals which overlap each other in space, and if two electrons are available, the conditions are favorable for forming, with evolution of heat, a chemical bond. It follows that the valence of an atom is given by the number of unpaired electrons it possesses, an old basic rule of valence.

The greater the overlap between two atomic orbitals, the stronger the bond that can be formed with them (criterion of maximum overlapping). This condition may be regarded as determining the shapes of molecules. Two or more orbitals of comparable energy, as 2*s* and 2*p* orbitals, can be combined (hybridized) to give orbitals concentrated along certain directions in space, and these are the orbitals that participate in directed bond formation. In the carbon atom, for instance, the four electrons in the 2*s* and 2*p* subshells are potential valence electrons. The two 2*s* electrons are paired, however, so that to make four bonds possible one of these must be promoted to a vacant 2*p* orbital. Four bonds than are possible, in various directions. Four equivalent bonds can be formed, tetrahedrally directed, as in CH_4. Three bonds in a plane and one other less strong one can be formed, as in $H_2C{=}CH_2$. In this manner L. Pauling and others have accounted for a multitude of phenomena in stereochemistry. *See* STEREOCHEMISTRY.

The peculiar bonding in benzene and other aromatic molecules has been explained, together with its consequences for chemical reactivity. The principles governing reaction rates have been formulated and applied. *See* BENZENE; DELOCALIZATION; RESONANCE (MOLECULAR STRUCTURE).

Research in valence theory has led to general and complete understanding of most aspects of molecular electronic structure, and has contributed toward acceptance of the language of modern physics as a proper language for chemistry. Considerable research in this field continues, however, because substances with new types of bonds are being synthesized constantly, and new physical methods for studying molecules are constantly revealing more intimate details of molecular structure which demand explanation. *See* CHELATION; CHEMICAL DYNAMICS; CONJUGATION AND HYPERCONJUGATION; COORDINATION CHEMISTRY; LIGAND FIELD THEORY; ORGANIC CHEMISTRY; ORGANOMETALLIC COMPOUND; STRUCTURAL CHEMISTRY. Robert G. Parr

Bibliography. H. Gray, *Electrons and Chemical Bonding*, 1964; D. J. Klein and N. Trinajstic (eds.), *Valence Bond Theory and Chemical Structure*, 1990; R. McWeeney (ed.), *Coulson's Valence*, 3d ed.,

1980; M. F. O'Dwyer, J. E. Kent, and R. D. Brown, *Valency*, 2d ed., 1989; L. Pauling, *The Nature of the Chemical Bond and the Structure of Molecules and Crystals*, 3d ed., 1960; L. Pauling and P. Pauling, *Chemistry*, 1995; A. N. Stranges, *Electrons and Valence: Development of the Theory 1900–1925*, 1982.

Valence band

The highest electronic energy band in a semiconductor or insulator which can be filled with electrons. The electrons in the valence band correspond to the valence electrons of the constituent atoms. In a semiconductor or insulator, at sufficiently low temperatures, the valence band is completely filled, and the conduction band is empty of electrons. Some of the high energy levels in the valence band may become vacant as a result of thermal excitation of electrons to higher energy bands or as a result of the presence of impurities. When some electrons are missing, the remaining ones may be redistributed among the energy levels within the valence band under an applied electric field, giving rise to an electric current. The net effect of the valence band is then equivalent to that of a few particles which are equal in number and similar in motion to the missing electrons but each of which carries a positive electronic charge. These "particles" are referred to as holes. *See* BAND THEORY OF SOLIDS; CONDUCTION BAND; ELECTRIC INSULATOR; HOLE STATES IN SOLIDS; SEMICONDUCTOR.

H. Y. Fan

Bibliography. C. Kittel, *Introduction to Solid State Physics*, 7th ed., 1996.

Value engineering

A thinking system (also called value management or value analysis) used to develop decision criteria when it is important to secure as much as possible of what is wanted from each unit of the resource used. The resource may be money, time, material, labor, space, energy, and so on. The system is unique in that it effectively uses both knowledge and creativity, and provides step-by-step techniques for maximizing the benefits from both. It promotes development of alternatives suitable for the future as well as the present. This is accomplished by identifying and studying each function that is wanted by the customer or user, then applying knowledge and creativity to achieve the desired function. Resources are converted into costs to achieve direct, meaningful comparisons. By using the methods of value engineering, 15 to 40% reduction in the required resources often results.

Application. Value engineering has applications in five broad areas: in design, purchase, and manufacture of products; in administrative groups, private or public, where the task is to achieve accomplishment through people; in all areas of social service work, such as hospitals, insurance services, or colleges; in architectural design and construction; and in development as well as research.

Approach–function study. The specific functions which the customer or user wants will be of two types: use function, which does something for the user; and esthetic function, which is selected because it pleases the user. Value engineering uses a planned approach for intensive and effective utilization of every applicable technique. It requires the development of sufficient skill in the application of enough techniques to bring into clear view a liberal number of value alternatives. Each function is systematically understood, identified, clarified, and named.

Value. Value is defined as the proper function for the least resources. Best value, then, is the attainment of the full function desired for the lowest cost achievable. The value of a function becomes a vital measure, being the lowest cost of securing the function. After arriving at this measure, high effort is made to achieve a function near its value.

The system is used to improve value in either or both of two situations: (1) The product or service as used or as planned may provide 100% of the functions the user wants, but lower costs may be needed. The system then holds those functions but achieves them at lower cost. (2) The product or service may have deficiencies, that is, it does not perform the desired functions or lacks quality, and so also lacks good value. The system aims at correcting those deficiencies, providing the functions wanted, while at same time holding the use of resources (costs) at a minimum.

Evaluation of function by comparison. Once the functions have been identified, clarified, understood, and specified, the following question must be answered: What is the lowest cost which, under the present conditions, would provide the described function? The answer is developed by comparisons to the past. These values are established by other valid comparisons, such as: How might an important portion of the function be accomplished, and what would that cost? How would that function be accomplished in a different industry or a different country, under very different conditions, and what would it cost? These values have to be compared with larger, smaller, and similar items or services and their costs. If there is no comparison, there can be no evaluation.

Often this task of creatively evaluating the function of itself brings a good answer to the problem. For example, the Navy was building 1000 landing crafts. One function was to "contain 200 gallons of gasoline." A noncombat life of 8 years was desired. The best quotation for each landing craft was $520. The function was evaluated by comparisons. How else could 200 gal be contained and what would each cost? Four 50-gal drums would have cost $25. A standard 250-gal oil tank, often used with oil burners in homes, cost $30. The $30 figure is selected as a base. Additional costs were considered for some

connections, piping, and perhaps coatings that would be needed, so the value of the function was selected to be $50. As a result of the evaluation the Navy elected to use four drums, separated into two groups of two, at a unit cost of $80. Thus, the cost for the job was $80,000 instead of $520,000.

Minimizing normal human negatives. The system, with its intense emphasis on functions, deep searches for knowledge, and constant effective creativity, built into step-by-step techniques, offsets many human traits which retard or prevent beneficial change. Some of the human negatives which act as retardants are the following: (1) Thoughts tend to follow habit patterns. (2) A decision to change, if proved wrong, may bring embarrassment. (3) Making a change may bring personal risk to the decision maker. (4) Decisions based upon sound general criteria often do not fit the specific instance. (5) The required good decision may be contrary to what is "normally" done. (6) Subjective coloring of attitudes of important people in the area makes good decisions difficult without sound objective data. (7) Decisions vital to profitability are often made by people not accountable for profits. (8) Obscure cause-and-effect relationships in some matters that create costs allow decisions to be made that will injure value. (9) While feelings are strictly personal, they may influence and often control decisions affecting value. (10) Most environments are hazardous to anything new, or to a change. Sound, objective alternatives developed by the value engineering problem solving system do much to overcome these human retardants.

Job plan. All minds must be "tuned" to work on exactly the same problem at exactly the same time. A job plan is thus developed in order to present the problems that will be faced and to establish the functions that should be accomplished. The job plan requires the following five steps: (1) gathering extensive information about the problem area; (2) analyzing the information for meaning and sense of direction; (3) doing the essential creative work; (4) judging the results; (5) creating a development planning program. The first four steps all require a different mental activity, and each has to be thoroughly completed before proceeding onto the next step.

Information step. In this step, all the facts and pertinent information (such as costs, quantities, and specifications) have to be obtained because only through complete understanding of the situation can valuable assessments be made. Assumptions have to be sorted out and reviewed to determine if they can be replaced or supported by facts. Long-standing assumptions have to be especially checked for validity.

Analysis step. This step involves the development of "function" thinking. From the information collected in the first step, functions are developed to answer such questions as: What are the problems involved, and which must be solved first? Are the solutions to these problems reasonable? What goals should be aimed for, and what steps should be taken in order to achieve them? Is any more information needed? Have any assumptions been overlooked, and are all assumptions already noted still valid? Have the best approaches to these problems been developed and, if so, what savings or benefits will result? Should better solutions be sought? What problems, if presented with better solutions, would produce even more beneficial results for the project?

Creativity step. After acquiring the relevant information and reaching an understanding of the problems through analysis, one must apply creative thinking, involving free use of the imagination. Two aspects of the mind are important here: the diverse knowledge bits that it has acquired over the years, and the ability to join these bits in various combinations constituting temporary "mental pictures" that may be unique. Thus a number of fresh alternative solutions to consider emerge in solving the overall problem or individual problems.

Judgment step. This step involves creatively studying the ideas presented in the previous step. No ideas should be thrown out. Rather, they should be developed and improved. Ideas involving monetary value should especially be studied closely and objectively in order to seek out their limitations and to try to lessen, overcome, or eliminate any negative aspects. It may be necessary to send certain ideas back to step one and to run them through the entire process again. Some ideas may have so many advantages that the lessening of their drawbacks may become the principal point of concern. And those ideas which appear to be completely thought through and seem capable of providing the greatest yield are sent on to step five.

Development planning step. In this step the best specialists and vendors are selected for consultation, and an investigation program is established to provide the most recent information on and the most current capabilities of any of the approaches that show potential.

VECP (value engineering change proposals). When this thinking system is used on government or military supply or construction situations, the same or better quality is usually secured for millions of dollars less in cost. To make it profitable for the manufacturer or contractor to hire value engineers and make this contribution to the government, a percentage of the saving, often 40 to 50%, for a time is paid on VECPs. Hundred of millions of tax dollars are being saved each year because of VECPs. In addition, resources are being conserved, and millions of dollars are being saved by owners, because of the work of the architectural firms and construction companies that have trained and qualified value engineers.

Training and qualification. Skilled consultants and some universities teach the techniques of value engineering. A professional society, the Society of American Value Engineers, with chapters in many cities, sets standards, gives examinations, and awards the citation of "Certified Value Specialist" (CVS) to qualified people.

World status. Extensive application of value engineering is growing in the United States, Japan, Germany, Sweden, France, Canada, and England. Important application is growing in Norway, Italy, Spain, Korea, Taiwan, South Africa, India, and other places. Professional societies exist in the United States, Japan, Scandinavia, France, and South Africa. *See* ACTIVITY-BASED COSTING; INDUSTRIAL ENGINEERING; METHODS ENGINEERING; OPERATIONS RESEARCH; OPTIMIZATION; PROCESS ENGINEERING; PRODUCTION ENGINEERING; PRODUCTION PLANNING.

Lawrence D. Miles

Bibliography. E. Castillo, *Extreme Value Theory in Engineering,* 1988; W. H. Copperman, *A Guide to the Contractual Aspects of Value Engineering,* 1986; R. M. Curtice, *Strategic Value Analysis: A Modern Approach to Systems and Data Planning,* 1987; M. C. Macedo, Jr., P. V. Dobrow, and J. J. O'Rourke, *Value Management for Construction*, 1978; A. Mudge, *Value Engineering: A Systematic Approach,* 1989; J. J. O'Brien, *Value Design and Construction*, 1976; D. E. Parker, *Management Application of Value Engineering: For Business and Government*, 1994; *Value Engineering and Management Digest*, monthly.

Valve

A flow-control device. This article deals with valves for fluids, liquids, and gases. Valves are used to regulate the flow of fluids in piping systems and machinery. In machinery the flow phenomenon is frequently of a pulsating or intermittent character and the valve, with its associated gear, contributes a timing feature. For electrical valves *see* ELECTRON TUBE

Pipe valves. The valves commonly used in piping systems are gate valves (**Fig. 1**), usually operated closed or wide open and seldom used for throttling; globe valves (**Fig. 2**), frequently fitted with a renewable disk and adaptable to throttling operations; check valves (**Fig. 3**), for automatically limiting flow in a piping system to a single direction; and plug cocks (**Fig. 4**), for operation in the open or closed position by turning the plug through 90° and with a shearing action to clear foreign matter from the seat.

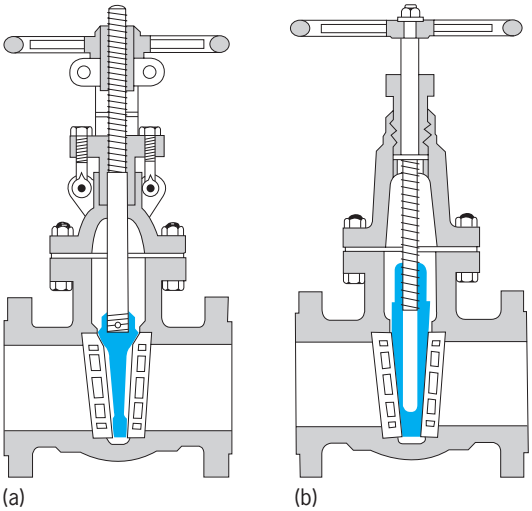

Fig. 1. Gate valves with disk gates shown in color. (*a*) Rising threaded stem shows when valve is open. (*b*) Nonrising stem valve requires less overhead.

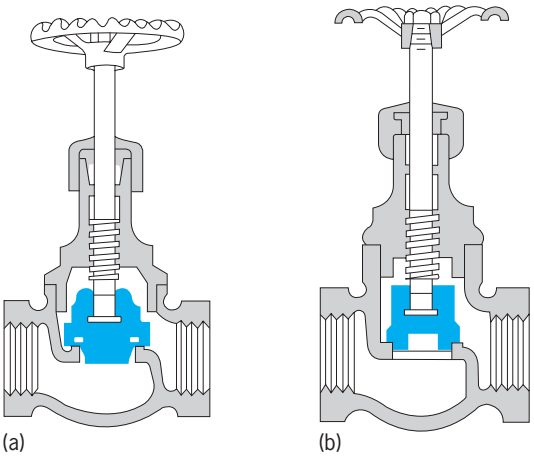

Fig. 2. Globe valves with (*a*) gasket in disk and (*b*) ground metal-faced disk.

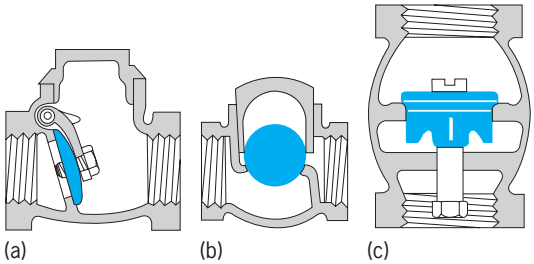

Fig. 3. Various types of straightway check valves. (*a*) Swing. (*b*) Ball. (*c*) Vertical.

Valves may have various structural features such as outside stem and yoke; packless construction; angle, as opposed to straightway flow; power instead of manual operation; and combined nonreturn and stop-valve arrangements. Valves are made in a wide assortment of materials, and a wide variety of trim, with brass or bronze for general service; cast iron for low steam pressures and temperatures (less than 250 lb/in.^{2} or 1.7 megapascals) and for hydraulic pressures below 800 lb/in.^{2} (5.5 MPa); steel and alloy steels for the highest operating pressures and temperatures (such as 5000 lb/in.^{2} or 34 MPa, 1200°F or 650°C steam); and selected metals for chemical and process applications. Most valves are manufactured and available as hardware and comply with the requirements of the ASTM, ANSI, and ASME as to material and dimensional standards. They are variously offered as flanged, screwed, welded, sweated, or compression-fitted for connection to pipe, machinery, and fittings.

Safety and relief valves are automatic protective devices for the relief of excess pressure. They are usually rigorously specified under the legal

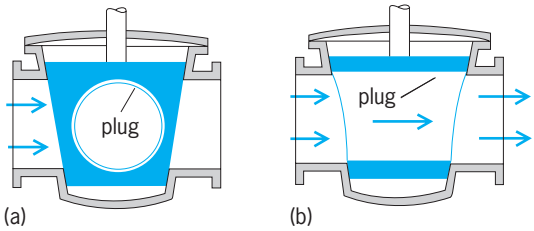

Fig. 4. Plug valve in (*a*) closed and (*b*) open positions.

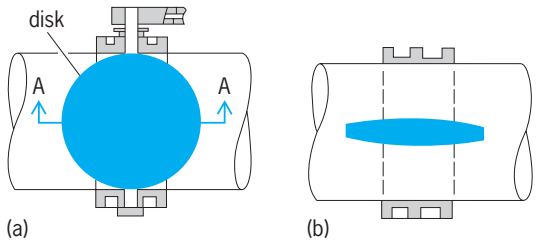

Fig. 5. Butterfly-type valve for penstock is typically 25 ft (7.6 m) in diameter. (*a*) Elevation. (*b*) Section A-A.

regulations of public authorities and insurance underwriters. They must open automatically when the pressure exceeds a predetermined value; they must allow the pressure to drop a predetermined amount before closing to avoid chattering, instability, and damage to the valve and the valve seat; they must have adjustment features for both the relieving and blowdown pressures; and they must be tamperproof after setting by responsible licensed operators. *See* SAFETY VALVE.

Hydraulic turbine valves. For hydraulic turbines and hydroelectric systems, valves and gates control water flow for (1) regulation of power output at sustained efficiency and with minimum wastage of water, and (2) safety under the inertial flow conditions of large masses of water. Valve sizes are usually large [for example, 6 ft (1.8 m) in diameter] so that power operation is necessary. Carefully streamlined construction to minimize fluid dynamic losses must be accompanied by ample provision to withstand shock and damaging effects of hydraulic inertia. Gate, butterfly (**Fig. 5**), telescoping, and needle constructions (**Fig. 6**) are variously employed. Wicket or cylinder gates regulate the flow of water to a reaction turbine at the speed ring while a governor-operated needle valve regulates flow to a Pelton impulse unit. *See* HYDRAULIC TURBINE; WATER HAMMER.

Steam-engine valves. To control the kinematics of the cycle, steam-engine valves range from simple D-slide and piston valves to multiported types. Slide valves control admission and release of steam to and from a double-acting cylinder by a single moving valve mechanism giving the necessary lap, lead, and angle of advance to accomplish the predetermined values of cutoff and compression. Multiported valves such as plug, Corliss, or poppet valves provide four valves for a double-acting cylinder. Each valve serves a single purpose of admission or exhaust for the head or crank end. The uniflow construction uses a poppet valve for admission with a row of exhaust ports alternately covered and uncovered by the engine piston. Plug and slide valves are limited to low pressures and temperatures (200 lb/in.^{2} or 1.4 MPa and 100°F or 55°C of superheat); poppet valves will operate under the maximum pressures and temperatures of steam-engine practice without warping. Many types of reversing gear have been perfected which use the same slide valve or piston valve for both forward and backward rotation of an engine, as in railroad and marine service. *See* STEAM ENGINE.

Internal combustion engine valves. Poppet valves are used almost exclusively in internal combustion reciprocating engines because of the demands for tightness with high operating pressures and temperatures. The valves (**Fig. 7**) are generally 2 in. (5 cm) in diameter or smaller on high-speed automotive-type engines, are cam-operated and spring-loaded, with their lift a small fraction of an inch, and with a gas velocity of 200–300 ft/s (60–90 m/s). They are cooled by transfer of heat to the engine jacket mostly through the valve stem. In heavy-duty units, the stem may have a partial mercury or sodium filling which aids, by inertia, in conducting heat to the

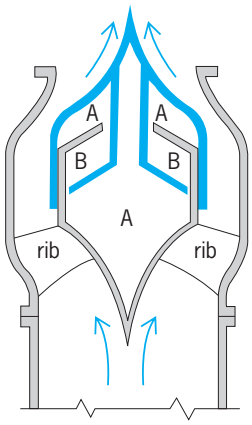

Fig. 6. Large needle valve is actuated hydraulically by pressure in chambers A to close and by pressure in annular chambers B to open.

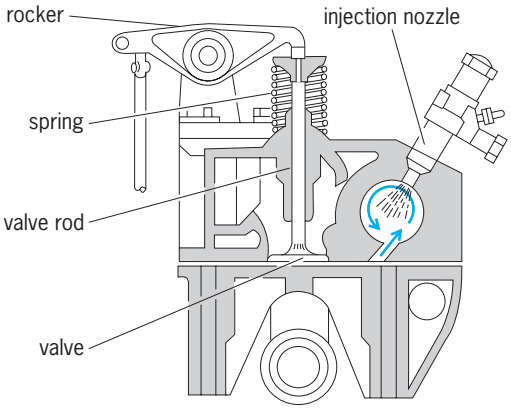

Fig. 7. Poppet valve for internal combustion engine. (*After E. A. Avallone and T. Baumeister III, eds., Marks' Standard Handbook for Mechanical Engineers, 8th ed., McGraw-Hill, 1978*)

valve stem and guides. Exhaust valves are subject to the effects of extreme temperature and must accordingly be most carefully designed and constructed of alloy materials. Two-cycle engines utilize ports, alternately covered and uncovered by the main piston, for inlet or exhaust. *See* CAM MECHANISM; INTERNAL COMBUSTION ENGINE; VALVE TRAIN.

Compressor valves. In compressors, valves are usually automatic, operating by pressure difference (below 5 lb/in.^{2} or 34 kilopascals) on the two sides of a movable, spring-loaded member and without any mechanical linkage to the moving parts of the compressor mechanism (**Fig. 8**). The practical objectives are timing, tightness, and life. Pressure drop should be a minimum. Moving parts must be light, but spring loadings should be heavy enough to avoid chatter and to give certainty to the valve action. Airspeeds are in the order of 5000 ft/min (1500 m/min) through the valves. Valves for inlet and exhaust service are frequently made interchangeable in the interest of reduced investment. On high-pressure (above 1000 lb/in.^{2} or 7 MPa) service, automatic poppet valves are usually substituted for the more common plate valves to give greater strength and tightness under loading. *See* COMPRESSOR.

Pump valves. Like those for compressors, pump valves are usually of the automatic type operating by pressure difference. The service conditions, however, are very dissimilar because of the noncompressibility of liquids; the presence of entrained solids like grit, fibers, and sludge; the corrosive potential of chemicals like acids and alkalies; the high viscosity of many liquids; the vapor pressure at the pump-operating temperature; and inertia effects accompanying discontinuity of the liquid column. Fluid speeds are low (200–300 ft/min or 60–90 m/min with cold water and 100 ft/min or 30 m/min with viscous liquids). A multiplicity of small valves (about 4 in. or 10 cm in diameter) arranged in docks with a positive suction is a common construction. Various constructions are used (**Fig. 9**). Theodore Baumeister

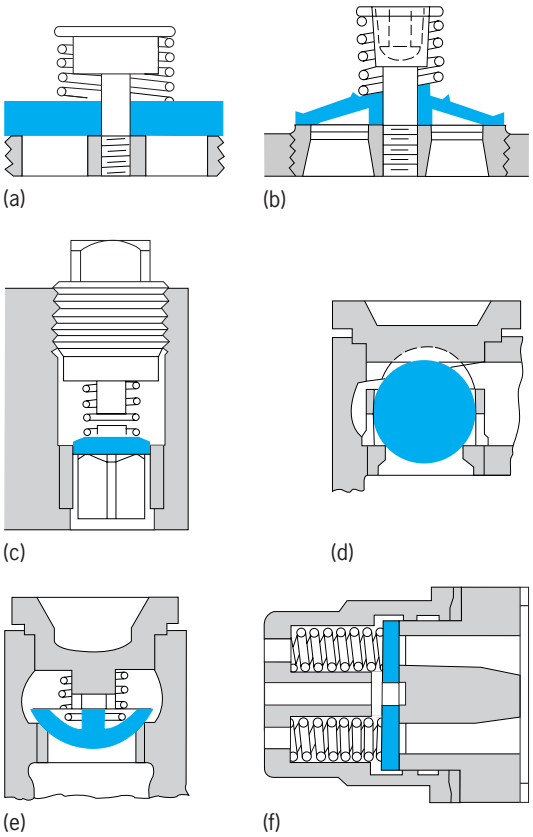

Fig. 9. Pump valves. (*a*, *b*) Disk valves are for moderate pressure. (*c*) Conical-faced wing valves are for high pressure (above 1000 lb/in.^{2} or 7 MPa). (*d*, *e*) Ball and rounded valves are for viscous liquids. (*f*) Double-ported valves are for high-speed reciprocating pump.

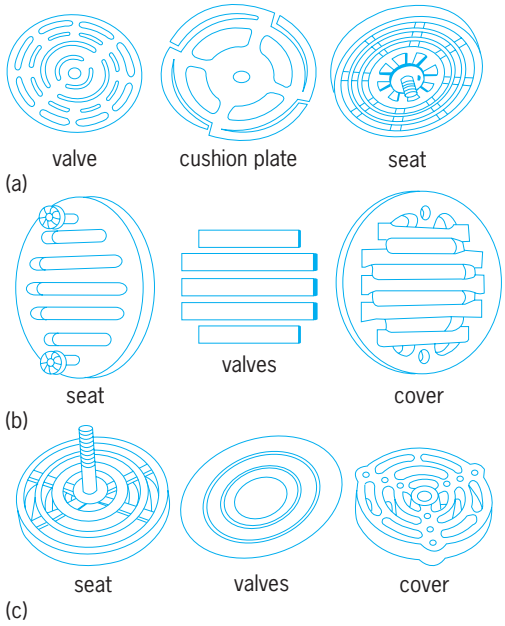

Fig. 8. Air and gas compressor valves. (*a*) Ingersoll-Rand ring plate. (*b*) Worthington feather valve. (*c*) Chicago pneumatic simplate valve. (*After E. A. Avallone and T. Baumeister III, eds., Marks' Standard Handbook for Mechanical Engineers, 8th ed., McGraw-Hill, 1978*)

Bibliography. E. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; W. Ulanski, *Valve and Actuator Technology*, 1991.

Valve train

The valves and valve-operating mechanism by which an internal combustion engine takes air or fuel-air mixture into the cylinders and discharges combustion products to the exhaust. *See* VALVE.

Mechanically, an internal combustion engine is a reciprocating pump, able to draw in a certain amount of air per minute. Since the fuel takes up little space but needs air with which to combine, the power output of an engine is limited by its air-pumping capacity. The flow through the engine should be restricted as little as possible. This is the first requirement for valves. The second is that the valves close off the cylinder firmly during the compression and power strokes. *See* INTERNAL COMBUSTION ENGINE.

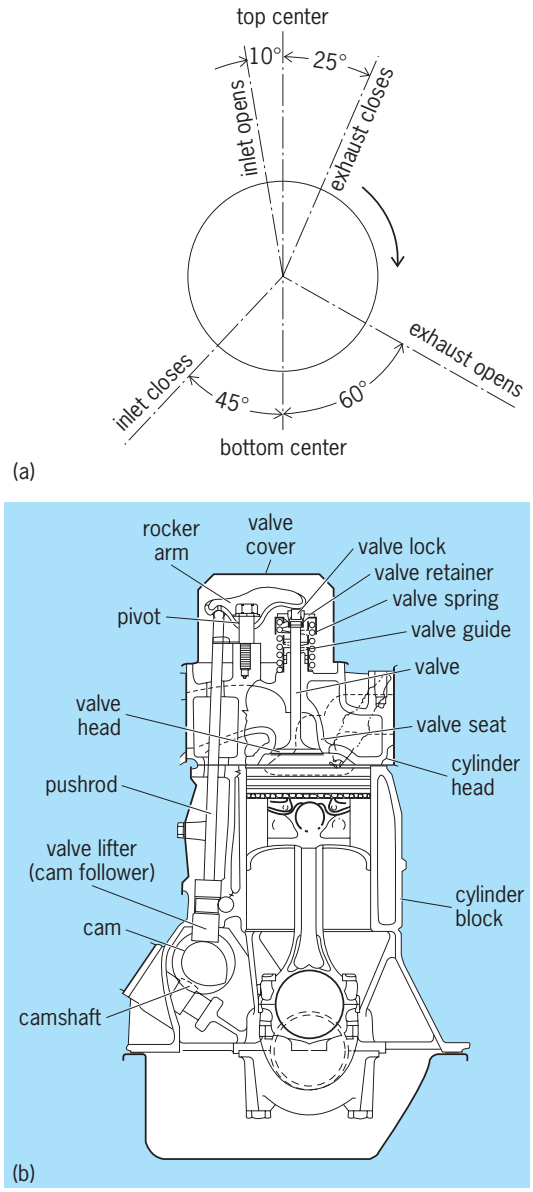

Fig. 1. Valve for a four-cycle engine. (*a*) Valve-timing diagram. (*b*) Valve train with the camshaft in the cylinder block and the valves in the cylinder head. (*Pontiac Division, General Motors Corp.*)

Valve action. In most four-stroke engines the valves are the inward-opening poppet type, with the valve head ground to fit a conical seat in the cylinder block or cylinder head (**Fig. 1**). The valve head is held concentric with its seat by a cylindrical stem running in a valve guide. The valve is held closed by a compressed helical spring. The valve is opened wide by lifting it from its seat a distance equal to approximately 25% of the valve diameter. At this lift the cylindrical opening between valve and seat will be about as large as the cross section of the flow passage to the valve seat.

The valve is streamlined and as large as possible to give maximum flow, yet of low inertia so that it follows the prescribed motion at high engine speed. Valves are usually made of a heat- and scale-resistant alloy that will keep its strength and shape at high temperature. The exhaust valve is sometimes made hollow and partially filled with metallic sodium. In operation the sodium melts and shakes back and forth, transferring heat from the valve head to the stem and valve guide. The valve-seat area is hard-faced, with inserts of a tough heat-resistant alloy used in aluminum castings and high-performance engines.

Cam action. Engine valves are usually opened by cams that rotate as part of a camshaft, which may be located in the cylinder block (Fig. 1) or cylinder head (**Fig. 2**). Riding on each cam is a cam follower or valve lifter, which may have a flat or slightly convex surface or a roller, in contact with the cam. The valve is opened by force applied to the end of the valve stem. A valve rotator may be used to rotate the valve slightly as it opens.

In many small four-stroke cycle engines, which have the camshaft and valves in the cylinder block, the force is applied through a valve lifter to the valve stem. In automotive and diesel engines with the camshaft in the cylinder block and valves in the cylinder head, the force may be applied through a mechanical linkage actuated by the cam follower (Fig. 1). The linkage consists of cam follower, pushrod, and rocker arm. The pushrod is a light rod or tube with ball ends, which carries the motion of the cam follower to the rocker arm. The rocker arm is a lever pivoted near its center so that, as the pushrod raises one end, the other end depresses the valve stem, opening the valve. *See* CAM MECHANISM.

In engines with the camshaft and valves in the cylinder head, the cam may operate the valve directly through a cup-type cam follower, or bucket tappet (Fig. 2). Other overhead-camshaft engines use a finger-type cam follower or rocker arm between the cam and the valve stem.

To ensure tight closing of the valve even after the valve stem lengthens from thermal expansion, the

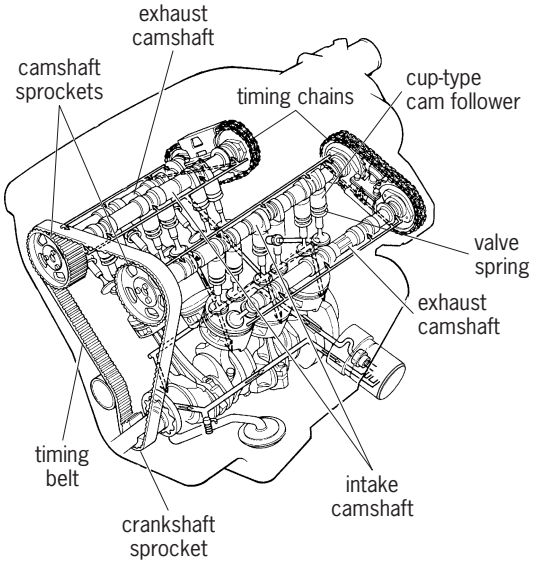

Fig. 2. Valve train of a V-6 four-cycle engine with the valves and two camshafts located in each cylinder head. The cams act directly on cup-type cam followers to open the valves. (*Ford Motor Co.*)

valve train is adjusted to provide some clearance when the follower is on the low part of the cam. The cam shape includes a ramp that reduces shock by starting the lift at about 2 ft (0.6 m) per second even though the clearance varies from time to time. To open the valve quickly, an acceleration of about 400 times gravity (for automotive-size engines) is used. Excessive acceleration deflects the valve linkage, giving false motion to the valve, causing it to close at high velocity before the closing ramp is reached. The cam surface between the opening and closing acceleration sections includes the point of maximum lift and maximum deceleration. If this part of the cam is sinusoidal and the valve spring is properly designed, the spring forces remain proportional to the forces required by the cam for the deceleration of the valve. The highest operating speed for a given lift can be obtained without the follower leaving the cam. The time that the valve is held at maximum lift is known as the dwell. *See* HYDRAULIC VALVE LIFTER.

Timing. At high piston speeds the inertia of the charge in the inlet passages causes the gas to continue to flow into the cylinder long after the piston has started up on the compression stroke. To trap the maximum amount of fresh charge in the cylinder, the closure of the inlet valve may be delayed up to about 60° after bottom center. Late inlet closing increases engine torque at high speeds but reduces torque at low speeds. To reduce the work required at high speeds to expel the products of combustion during the exhaust stroke, the exhaust valve must open considerably before bottom center. The time during which both valves are open is known as valve overlap. Donald L. Anglin

Bibliography. *Bosch Automotive Handbook*, 2d ed., 1987; W. H. Crouse and D. L. Anglin, *Automotive Engines*, 8th ed., 1994; Society of Automotive Engineers, *SAE Handbook,* 4 vols., annually.

Vampyromorpha

An order of coleoid cephalopods that contains only one species, *Vampyroteuthis infernalis* (see **illustration**). It is characterized by a flat, broad, leaflike, chitinous internal shell (gladius); eight arms around the mouth connected by a deep web; no tentacles, but two small sensory filaments that retract into pockets between the bases of the first and second arms; fingerlike cirri and a single row of suckers with chitinous rings along the arms; one pair of paddle-shaped fins on the body in adults (two pairs at the juvenile stage); very dark maroon, mostly black pigmentation; and photophores (light-producing organs) on the body, head, and arms. Vampyromorphs superficially resemble cirrate octopods, but they bear features that show distinct teuthoid (squid) relationships (for example, similar gladius, photophores, and radula), and in many characteristics are intermediate between teuthoids and octopods. Fossil records indicate the group is a descendant of Late Triassic or Jurassic forms. The sensory filaments between the first and second pair of arms are unique among all cephalopods and are thought to be tactile organs. *See* TEUTHOIDEA.

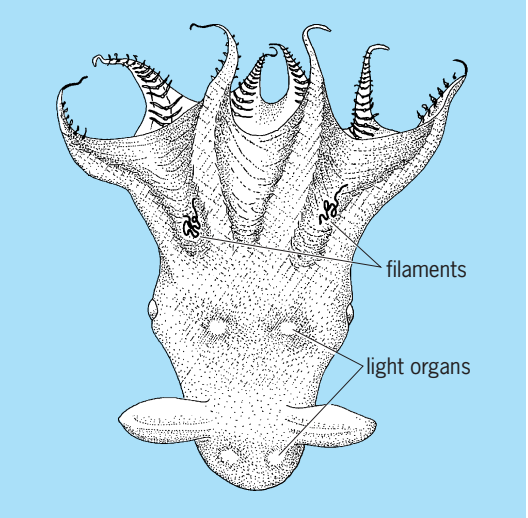

Vampyroteuthis infernalis **showing sensory filaments and photophores (light organs).**

The sexes are separate, and the male has no hectocotylus (modified arm for transferring sperm). Sperm are stored by females in seminal receptacles anterior to the eyes. The eggs are relatively large (0.12–0.16 in. or 3–4 mm in diameter), and they are laid individually in the water. The young usually occur deeper than 2900 ft (900 m), then gradually move shallower as they grow. Adults reach a total length of about 16 in. (40 cm). *See* CEPHALOPODA; COLEOIDEA; OCTOPODA.

Vampyroteuthis infernalis is a gelatinous bathypelagic (very deep, midwater living) species that inhabits worldwide tropical, subtropical, and temperate waters. It occurs mostly at depths of 1600–3800 ft (500–1200 m), occasionally to 4800 ft (150 m). Clyde F. E. Roper

Bibliography. P. R. Boyle, *Cephalopod Life Cycles*, vol. 1, 1983, vol. 2, 1987; K. N. Nesis, *Cephalopods of the World*, transl. from Russian, 1987; G. Pickford, *Vampyroteuthis infernalis* Chun, an archaic dibranchiate cephalopod, pt. 1, Dana Rep. 29, pp. 1–45, 1946, pt. 2, Dana Rep. 32, pp. 1–132, 1949; C. F. E. Roper and R. E. Young, Vertical distribution of pelagic cephalopods, *Smithson. Contrib. Zool.*, no. 209, pp. 1–51, 1975; K. M. Wilbur and M. R. Clarke (eds.), *The Mollusca*, vol. 12: *Paleontology and Neontology of Cephalopods*, 1988; R. E. Young, Homology of retractile filaments of vampire squid, *Science*, 156(3782):1633–1634, 1967.

Vanadium

A chemical element, V, with atomic number 23. Natural deposits contain two isotopes, ^{50}V (0.24%), which is weakly radioactive, and ^{51}V (99.76%). Commercially important as an oxidation catalyst, vanadium also is used in the production of alloy steel and ceramics and as a colorizing agent. Studies have demonstrated the biological occurrence of vanadium,

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**
actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

especially in marine species; in mammals, vanadium has a pronounced effect on heart muscle contraction and renal function. *See* PERIODIC TABLE; TRANSITION ELEMENTS.

Very pure vanadium is difficult to prepare because the metal is highly reactive at temperatures above the melting point of its oxide (663°C or 1225°F) from which it is produced. Vanadium is a bright white metal that is soft and ductile. It has a melting point of 1890°C (3434°F), a boiling point of 3380°C (6116°F), and a density of 6.11 g/cm^3 (3.53 oz/in.^{3}) at 18.7°C (65.7°F). The thermal and electrical conductivity of vanadium is superior to that of titanium.

At room temperature, the metal is resistant to corrosion by oxygen, salt water, alkalies, and nonoxidizing acids, the exception being hydrogen fluoride (HF). Vanadium cannot withstand the oxidizing conditions presented by nitric acid or aqua regia. At elevated temperatures it will combine with most nonmetals to form oxides, nitrides, carbides, arsenides, and other such compounds.

The physical properties of vanadium are very sensitive to interstitial impurities. The strength varies from 30,000 lb/in.^{2} (200 megapascals) in the purest form to 80,000 lb/in.^{2} (550 MPa) in the commercial grade. The melting point is markedly altered by small impurities; vanadium containing 10% carbon has a melting point of 2700°C (4892°F).

Vanadium has a low fission neutron cross section. This property combined with the metal's excellent retention of strength at elevated temperatures has made its use in atomic energy applications attractive.

Carbon and alloy steels consume more than half the vanadium produced in the United States. Many plate, structural, bar, and pipe steels contain vanadium to enhance strength and toughness. The basis for the unique properties of these carbon alloy steels is the formation of vanadium carbide. These carbides are extremely hard and wear-resistant; they do not coalesce readily, but maintain a state of fine dispersion. Many large steel forgings contain vanadium in the range 0.05–0.15%; here vanadium acts as a grain refiner, and also improves the mechanical properties of the forgings. Tool steels are another large class of vanadium-containing steels; vanadium ensures the retention of hardness and cutting ability at the elevated temperatures generated by the rapid cutting of metals.

The production of ferrovanadium, an iron alloy, is very important since the primary commercial use of vanadium is in steel. Ferrovanadium is produced by aluminum or silicon reduction of V_2O_5 in the presence of iron in an electric arc furnace. The commonly practiced aluminum reduction is exothermic, so that little additional heat from the arc is required. Silicon processing requires a two-stage reduction to achieve efficient operation. *See* FERROALLOY; STEEL MANUFACTURE.

Vanadium compounds, especially V_2O_5 and NH_4VO_3, are excellent oxidation catalysts in the chemical industry. Processes that employ such catalysts include the manufacture of polyamides, such as nylon; sulfuric acid production by the contact process; phthalic and maleic anhydride syntheses; and various oxidations of organic compounds such as the conversion of anthracene to anthraquinone, ethanol to acetaldehyde, and sugar to oxalic acid. Vanadium pentoxide is used as a mordant in dyeing and printing fabrics and in producing aniline black for the dye industry. Vanadium compounds are used in the ceramics industry for glazes and enamels. A wide range of colors can be obtained with combinations of vanadium oxide, zirconia, silica, lead, tin, zinc, cadmium, and selenium. *See* CERAMICS; DYEING; MORDANT.

Vanadium has long been recognized as an essential element in biological systems; however, the role of the metal often is obscure. Tunicates accumulate vanadium to levels 1 million times greater than the surrounding seawater. This vanadium was once thought to act as an oxygen carrier but now is believed to be an oxidation catalyst that repairs damage to the polymeric, protective tunic of these animals. The first vanadium-dependent enzyme, vanadium bromoperoxidase, was isolated from brown, red, and green marine algae (for example, *Ascophyllum nodosum*); this enzyme catalyzes the bromination of a variety of organic molecules by using hydrogen peroxide and bromide. This activity may be the source of many important brominated compounds that potentially may be used as antifungal and antineoplastic agents. *See* ENZYME.

A variety of physiological effects in mammalian systems have been reported, the most significant being in cardiovascular and renal function. Vanadate causes constriction of veins in the kidney and can alter the retention and excretion of sodium and chloride ions. Cardiac effects of vanadium are species-specific, with observed increases (rabbit and rat) and decreases (guinea pig and cat) in heart muscle contractility. Because vanadate is a potent inhibitor of Na,K-ATPase in the laboratory, it has been suggested that this is the site of the metal's action. However, the physiology of vanadate is probably more complicated, since vanadium behaves as a hormone mimic by elevating intracellular calcium ion (Ca^{2+}) levels in a process that is poorly understood. *See* BIOINORGANIC CHEMISTRY.

Vincent L. Pecovaro

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and*

Physics, 85th ed., CRC Press, 2004; A. S. Tracey and D. C. Crans (eds.), *Vanadium Compounds: Chemistry, Biochemistry, and Therapeutic Applications*, 1998.

Van Allen radiation

The high-energy, charged particles that are trapped into orbits by the geomagnetic field, forming radiation belts that surround the Earth. The belts consist primarily of electrons and protons and extend from a few hundred kilometers above the Earth to a distance of about *8* R_e (R_e = radius of Earth = 6371 km = 3959 mi). James Van Allen and coworkers discovered them in 1958 using radiation detectors carried on satellites *Explorer 1* and *3*, and they are often referred to as the Van Allen belts.

Motion of charged particles. A charged particle under the influence of the geomagnetic field follows a trajectory that can be conveniently described as a superposition of three separate motions. The first motion, produced by the magnetic force acting at right angles to both the particle velocity and the magnetic field, is a rapid spiral about magnetic field lines. As the spiraling particle moves along the field line toward either the North Pole or South Pole, the increase in magnetic field strength causes the particle to be reflected so that it bounces between the Earth's two hemispheres. Superimposed on the spiral and bounce motions is a slow east-west drift; electrons drift eastward and protons or heavier ions drift westward. Thus, individual trapped particles move completely around the Earth in a complicated pattern, their motion being constrained to lie on magnetic shells. **Figure 1** illustrates the motion of a trapped proton during several north-south reflections. Subsequent bounces will carry the proton completely around the Earth. Trapped electrons behave in a similar manner except that the direction of drift is eastward. In the absence of perturbing forces such as additional electric and magnetic fields or collisions with other particles, this motion will continue indefinitely.

Because the Earth's magnetic field is not axially symmetric about the Earth, the radiation belts are also asymmetric. For example, the trapped particles dip closer to the Earth as they drift over the Atlantic Ocean. At altitudes above a few R_e, this effect becomes small. Near the outer boundaries of the belts, at 7–10 R_e, an additional warping of the geomagnetic field results from the pressure of the solar wind.

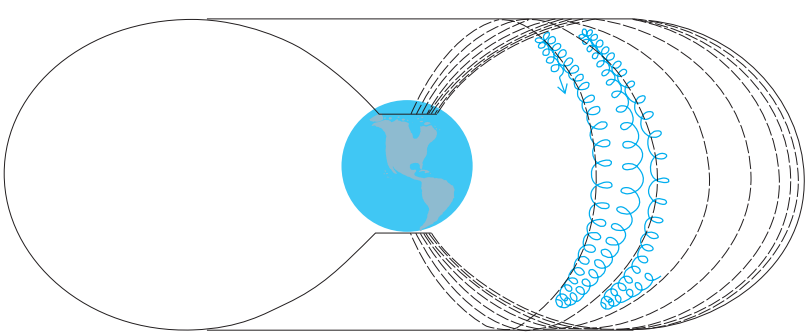

Fig. 1. Motion of a proton trapped in the geomagnetic field.

Trapped particle populations. The distribution of charged particles in the Earth's radiation belts is illustrated schematically in **Fig. 2**, where electron fluxes are depicted in the cross section on the left and proton fluxes are shown on the right. The darker shading indicates more intense fluxes of trapped particles. The spatial structure of trapped radiation shows two maxima: an inner radiation belt centered at about 1.5 R_e and an outer belt centered at 4–5 R_e. In the inner radiation belt, the most penetrating particles are protons with energies extending to several hundred megaelectronvolts (MeV). The intensity of the high-energy protons reaches 10^5 cm^{-2} s^{-1} for energies above 15 MeV. However, the flux of high-energy protons decreases rapidly with increasing distance from the Earth and becomes insignificant beyond 4 R_e. Low-energy protons with energies up to a few MeV occur throughout the stable trapping region which extends to approximately 8 R_e. Electrons are also found throughout the trapping region, with local maxima occurring in the inner and outer belts. The electron energies extend to several MeV, and in the outer radiation belt electrons are the most penetrating component.

Experiments have shown that helium nuclei (alpha particles) with energies of a few MeV are also trapped in the geomagnetic field, the maximum flux of a few hundred particles per square centimeter per second (energy >2 MeV) occurring at about 3 R_e. The flux of alpha particles is therefore only about 2×10^{-4} that of the protons. Large fluxes of trapped oxygen ions with energies up to about 20 kiloelectronvolts (keV) have also been discovered. These oxygen-ions, which are most pronounced after magnetic storms, are important in that they imply that some of the trapped ions come from the Earth's atmosphere.

The **table** gives sample values of the omnidirectional, integral fluxes of trapped electrons and protons as a function of altitude (in R_e) for positions in the equatorial plane. These values are averages, and time variations can be substantial, particularly for electrons in the outer belt. Also given in the table are values of the characteristic energy E_0, for which the flux decreases by a factor of 2.7. The table shows the systematic decrease in E_0 with increasing altitude for protons. A similar trend occurs for electrons except within the region between the inner and outer belts where the electron flux exhibits a minimum and the energy spectrum is richer in low-energy electrons.

Earth's ring current. The trapped particles drift slowly around the Earth. This motion, caused by inhomogeneities in the Earth's magnetic field, causes electrons to drift eastward and protons or heavier ions to drift westward. The resulting current, called the ring current, acts to decrease the strength of the Earth's (surface) northward magnetic field at low latitudes. During times of magnetic storms, it can store significant amounts of energy. To a first approximation, the ring current particles and radiation belt

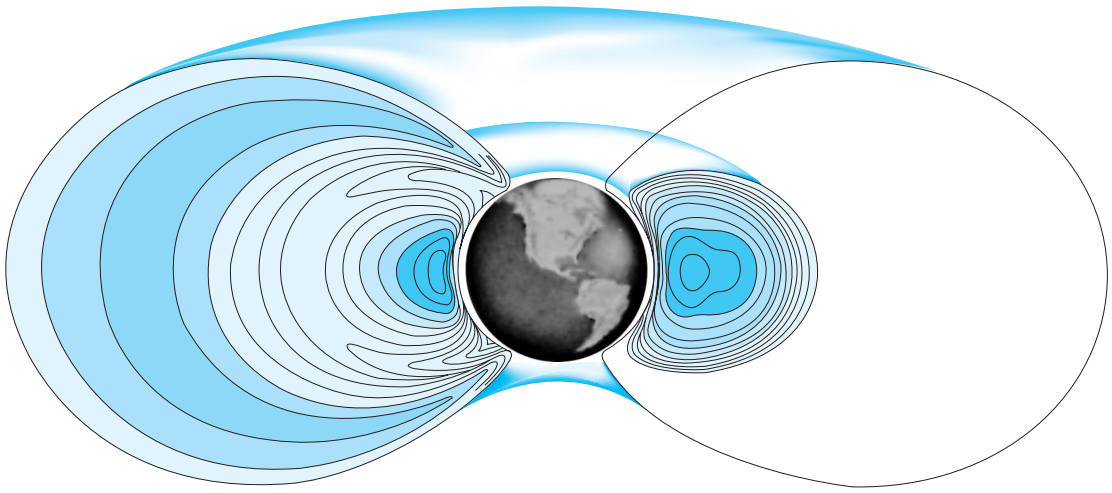

Fig. 2. Structure of Earth's radiation belts. Protons ($E > 15$ MeV) are shown on the right, and electrons ($E > 0.5$ MeV) are shown on the left. Darker shading indicates more intense fluxes of trapped particles.

particles are the same. There are, however, some important distinctions. Radiation belt particles typically refer to those particles that can penetrate into dense materials, such as instruments on board spacecraft or even astronauts. Electrons, for example, are important sources of penetrating radiation, yet contribute little to the ring current. Conversely, ring current particles carry the bulk of the total current density. *See* CURRENT DENSITY.

Time variations. The intensity, energy spectrum, and spatial distribution of particles within the radiation belts vary with time. The most dramatic variations are associated with magnetic storms, although smaller changes accompany most magnetic fluctuations. The changes are most pronounced for particles in the outer belt where the magnetic variations are the largest. During a magnetic storm, the electron flux in the outer belt may increase by an order of magnitude or more, although the response differs greatly among individual storms. Following a magnetic storm, the electron fluxes at lower altitudes usually increase as electrons, which were accelerated into the outer belt, diffuse closer to the Earth.

Figure 3 illustrates the time variations in the flux of electrons above 0.5 MeV on the Equator at 4 R_e. The upper part of the figure gives the K_p index (a measure of the magnetic activity). The higher the values of K_p, the larger are the irregular fluctuations in the geomagnetic field. It is apparent that the electron flux increases noticeably following magnetic fluctuations and decays slowly while the Earth's field is steady. The fluxes of low-energy protons in the outer belt also increase during magnetic storms, although the changes in proton fluxes are generally smaller than for electrons.

In the inner belt, the high-energy proton flux is relatively constant with time such that a very strong magnetic storm is required to alter the intensity appreciably. The most pronounced effect observed during magnetic storms is a reduction in the high-energy trapped proton flux, the reduction being greater with increasing distance from the Earth.

At the outer edges of the radiation belt, large time variations are common. A systematic diurnal variation exists, caused by the warping of the overall geomagnetic field by the pressure of the solar wind (a stream of ionized plasma emitted by the Sun). The resulting asymmetry in the field produces a diurnal variation in the trapped radiation observed at a given point above the Earth. Variations in solar-wind pressure alter the degree of warping of the magnetic field and produce variations in the flux of charged particles. *See* GEOMAGNETIC VARIATIONS; MAGNETOSPHERE; SOLAR WIND.

Omnidirectional fluxes and characteristic energies of trapped particles on the Equator

Altitude, R_e	Low-energy protons ($E > 0.4$ MeV)		High-energy protons ($E > 15$ MeV)		Electrons ($E > 0.5$ MeV)	
	Flux, $cm^{-2}\ s^{-1}$	E_0, MeV	Flux, $cm^{-2}\ s^{-1}$	E_0, MeV	Flux, $cm^{-2}\ s^{-1}$	E_0, MeV
1.25	4×10^4	2.2	1×10^4	20	4×10^7	0.9
1.5	1×10^6	2.2	1×10^5	13	6×10^7	0.4
2.0	1.6×10^7	2.1	7×10^4	8	3×10^6	0.13
2.5	1×10^8	0.6	8×10^3	8	2×10^5	0.33
3.0	2×10^8	0.4	2×10	7	6.7×10^4	0.42
3.5	1×10^8	0.34			3×10^5	0.5
4.0	3×10^7	0.3			1×10^6	0.4
5.0	4×10^6	0.13			2.5×10^6	0.35
6.0	9×10^5	0.11			9×10^5	0.25

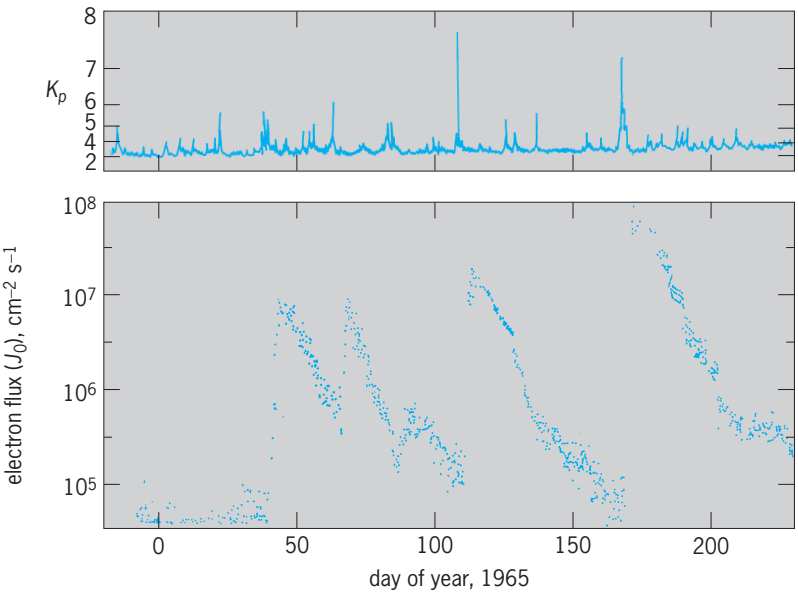

Fig. 3. Time variations in the magnetic activity index K_p and the associated changes in the flux of trapped electrons ($E > 0.5$ MeV) at an altitude of 4 R_e.

Artificial radiation belts. On at least nine occasions, high-energy electrons were injected into the geomagnetic field and trapped as a result of high-altitude nuclear detonations. The most famous of these events was the Starfish (hydrogen bomb) test on July 9, 1962, which produced an intense radiation belt lasting for many years. Following this nuclear explosion, several satellites ceased to transmit data because of loss of electrical power resulting from solar-cell damage. Shorter-lived belts were induced by three nuclear tests conducted by the Soviet Union in the fall of 1962. *See* NUCLEAR EXPLOSION.

Particle accelerators that are carried on rockets and satellites can also inject trapped electrons and ions. This type of experiment is designed to study the characteristics of the geomagnetic field and the plasma instabilities, which affect a beam of charged particles.

Origin and loss of trapped particles. It is believed that most of the very high energy protons (>50 MeV) in the inner belt result from the spontaneous decay of high-energy neutrons, produced by collisions of cosmic rays with atmospheric atoms. However, the vast majority of the radiation belt populations are ions and electrons that originate from either the atmosphere or the solar wind and are accelerated by processes only partly understood.

Theories of the sources of radiation-belt particles call on fluctuating electric and magnetic fields to accelerate the trapped particles. The geomagnetic field is continually agitated by variations in solar-wind pressure, and the changing magnetic field can accelerate charged particles by means of the induced electric fields. In addition to accelerating charged particles, the fluctuating fields also cause the particles to migrate across magnetic shells, making it possible for a particle that was initially injected at the outer boundary of the geomagnetic field to be transported to the inner belt. Detailed calculations of the acceleration and diffusion processes resulting from magnetic variations are not yet possible since there is insufficient knowledge about electric and magnetic fields in space. As knowledge improves, it will be possible to provide more exacting tests of the theories.

The physical nature of the processes by which trapped particles are ultimately lost from the belts is also poorly understood. The average lifetime for electrons varies from about a year in the center of the inner radiation belt to a few days in the outer belt. Within about 1000 km (600 mi) from the Earth, the particles are removed by collisions with neutral atoms and molecules of the Earth's atmosphere. The cumulative effect of these collisions is that the particles eventually lose so much energy that they become part of the ionosphere. At higher altitudes, the loss mechanism is probably the interaction of the trapped particles with electromagnetic waves. Waves of many frequencies and propagating modes exist throughout the magnetosphere, and their electric and magnetic fields can disrupt the trapping motion of the particles, eventually allowing them to reach the lower atmosphere. Theoretical estimates of the loss rates produced by these effects are in fair agreement with experimental values, but the severe approximations, which are required in the calculations, make the results inconclusive. Experiments have shown that electrons are also removed from the radiation belts by electromagnetic waves from high-power radio transmitters on the ground and by lightning discharges; however, the importance of these processes in controlling the radiation belts remains controversial.

Relation to other phenomena. Trapped radiation plays an important role in magnetic storms. One of the most important features of a magnetic storm is the main phase, during which the electric currents resulting from the motion of the trapped electrons and ions cause the horizontal component of the Earth's field measured at the Equator to decrease.

Electrons and protons, which are guided down magnetic field lines to the atmosphere at hight latitudes, also excite the polar aurora. It is believed that some of the same processes that accelerate auroral particles also accelerate radiation-belt particles. *See* AURORA.

The radiation belts are just one feature of the space plasma environment. For example, plasma from the Sun is continually impinging on the Earth's magnetic field, resulting in a "cavity" known as the magnetosphere. In addition, ionospheric electrons and ions form a plasma. Some of the electrons and ions making up this plasma can be accelerated and become trapped in the geomagnetic field, forming the radiation belts.

Particles are continually leaking out of the radiation belts into the atmosphere. This process produces ionization and contributes to the electron and ion population in the ionosphere. It is likely that some of the energy necessary to produce the airglow is derived from the charged particles striking

the atmosphere. The precipitating particles also initiate chemical reactions in the upper atmosphere that can, in turn, alter the concentrations of some of the minor species. For example, intense precipitation of electrons or protons can reduce the ozone content of the upper atmosphere, although it is not clear whether the effect is important on a global scale. *See* AIRGLOW; IONOSPHERE.

Radiation belts of other planets. Because the conditions that lead to the formation of Earth's radiation belts are so general, it is believed that any planet or moon that has a large enough magnetic field will also have radiation belts. Spacecraft missions to Mercury, Venus, and Mars have shown that the magnetic moments of these planets are too small to support long-term trapping of energetic particles. However, Jupiter, Saturn, Uranus, and Neptune have strong magnetic fields and very large, intense radiation belts analogous to those of the Earth. The belts for these planets have many characteristics that differ from those of the Earth because these planets are farther from the Sun, have high rotation rates, and have moons and rings immersed in their radiation belts. The rings of Saturn are efficient absorbers of energetic particles, so the radiation belts have gaps in the region of the rings. Similarly, the numerous moons of Saturn and Jupiter absorb particles, leading to minima in the radiation belt profiles at the orbital distances of the moons. On the other hand, some moons, especially Io at Jupiter, have atmospheres from which atoms escape, are ionized, and become part of the radiation belts. Hence, the radiation belts of some outer planets contain heavier ions such as oxygen, sulfur, and sodium provided by moons. Other sources of particles, such as protons from decaying neutrons, protons and helium ions from the solar wind, and ions from the planets' atmospheres, apparently exist in much the same fashion as with Earth. Comparison of the radiation belts at the different planets is giving valuable insight into the importance of the various plasma processes that form this trapped population of particles. *See* PLASMA (PHYSICS).

Space weather and radiation belts. The term "space weather" describes the conditions in space that affect Earth and its technological systems. It is a consequence of the behavior of the Sun, and the interaction of the solar wind with the Earth's magnetic field. Space weather can have many adverse effects, including damage to critical spacecraft instrumentation, degradation in the performance of navigation and communications systems, disruption of power grids, and exposure of astronauts and even aircraft passengers to hazardous radiation levels. In particular, the fluxes of electrons and protons trapped in the radiation belts can injure both personnel and equipment on board spacecraft if the vehicle is exposed to these energetic particles for a sufficiently long time. For example, on January 11, 1997, the communications satellite *Telstar 401*, which was orbiting 22,000 mi above Earth, was lost during a magnetic storm caused by a coronal mass ejection hitting the Earth's magnetosphere. Because of the spatial structure of the belts, the degree of damage will be strongly dependent on the position in space and hence on the orbit of the vehicle. In the inner belt region, high-energy protons can pass through several centimeters of aluminum structure and injure components in the interior of the spacecraft. In most other regions of the radiation belts, the trapped particles are less penetrating, and damage is confined to exposed equipment such as solar cells. Pete Riley; Martin Walt

Bibliography. G. Beynon et al. (eds.), *The Magnetosphere, the High-Latitude Ionosphere, and Their Interactions*, 1989; Y. Kamide and W. Baumjohann, *Magnetosphere-Ionosphere Coupling*, 1993; M. G. Kivelson and C. T. Russell (eds.), *Introduction to Space Physics*, 1995; G. K. Parks, *Physics of Space Plasmas*, 1991; J. G. Roederer, *Progress in Solar-Terrestrial Physics*, 1983; M. Schulz, *Earth's radiation belts, Rev. Geophys. Space Phys.*, 20:613–621, 1982; M. Schulz and L. J. Lanzerotti, *Particle Diffusion in the Radiation Belts*, 1974; J. A. Van Allen, *Origins of Magnetospheric Physics*, 1983.

Van der Waals equation

An equation of state of gases and liquids proposed by J. D. van der Waals in 1873 which takes into account the nonzero size of molecules and the attractive forces between them. He expressed the pressure p as a function of the absolute temperature T and the molar volume $V_m = V/n$, where n is the number of moles of gas molecules in a volume V:

$$p = \frac{RT}{V_m - b} - \frac{a}{V_m^2}$$

Here $R = 8.3145\ \mathrm{J\ K^{-1}\ mol^{-1}}$ is the universal gas constant, and a and b are parameters that depend on the nature of the gas. Parameter a is a measure of the strength of the attractive forces between the molecules, and b is approximately equal to four times the volume of the molecules in one mole, if those molecules can be represented as elastic spheres. The equation has no rigorous theoretical basis for real molecular systems, but is important because it was the first to take reasonable account of molecular attractions and repulsions, and to emphasize the fact that the intermolecular forces acted in the same way in both gases and liquids. It is accurate enough to account for the fact that all gases have a critical temperature T_c above which they cannot be condensed to a liquid. The expression that follows from this equation is $T_c = 8a/27\,Rb$. *See* CRITICAL PHENOMENA; GAS; INTERMOLECULAR FORCES; LIQUID.

In a gas mixture, the parameters a and b are taken to be quadratic functions of the mole fractions of the components since they are supposed to arise from the interaction of the molecules in pairs. The resulting equation for a binary mixture accounts in a qualitative but surprisingly complete way for the many kinds of gas-gas, gas-liquid, and liquid-liquid phase

equilibria that have been observed in mixtures. *See* PHASE EQUILIBRIUM.

The equation is too simple to represent quantitatively the behavior of real gases, and so the parameters a and b cannot be determined uniquely; their values depend on the ranges of density and temperature used in their determination. For this reason, the equation now has little practical value, but it remains important for its historical interest and for the concepts that led to its derivation. *See* THERMODYNAMIC PRINCIPLES. J. S. Rowlinson

Bibliography. A. Ya. Kipnis, B. E. Yavelov, and J. S. Rowlinson, *Van der Waals and Molecular Science*, Oxford University Press, 1996.

Vanilla

A choice flavoring obtained from a climbing orchid, *Vanilla fragrans*, a native of tropical American forests. The vanilla plant belongs to the orchid family and is indigenous to southeastern Mexico, where it was used by the Aztecs to flavor their cocoa. In 1510 *vagnuila* first appeared in Spain. Its fruits are pods called vanilla beans (see **illus.**). These are picked at the proper time before they have fully matured.

Vanillin (4-hydroxy-3-methoxybenzaldehyde) is the principal component of vanilla, although other components contribute to the distinctive flavor of the extract compared to synthetic vanilla. When they are harvested, the beans contain no free vanillin; it develops during the curing period from glucosides that break down during the fermentation and sweating of the beans. The sweating process consists of alternately drying the beans in sunlight and bunching them so that they heat and ferment. Sweating boxes are used in Mexico, whereas the shorter Madagascar method starts out by wilting green pods in hot water and uses blankets on which the beans can first be spread out and later rolled up for the enzymatic reactions and fermentation to take place. Further curing and dehydration occur in a warehouse. Periods of 4 weeks to 4 months may be required to develop the proper flavor and reduce the moisture content of the beans sufficiently to prevent molding. Beans can be artificially dried in ovens, but frequently an inferior-quality product results. *See* FERMENTATION; FOOD ENGINEERING.

Pods and flowers of *Vanilla fragrans*. (*USDA*)

After curing, the pods are sorted into grades based on quality. The best cured beans are 8–10 in. (20–25 cm) long, with drawn-out ends and curved bases. They are soapy or waxy to the touch, dark brown, and coated with fine crystals of vanillin, termed frost. Vanillin constitutes 1.2–3.5% of the bean, but other compounds contribute also to the aroma. In addition to the flavoring materials, vanilla beans contain fat, wax, sugar, gum, resin, and tannin. Vanilla is used in cookery, confectionery, and beverages. Vanilla extract, most used, is prepared by extracting the crushed beans with alcohol. A synthetic vanillin is made from eugenol occurring in clove oil, but the natural product is preferred. Several plants have been used as substitutes for true vanilla but these are of little value. *See* ORCHIDALES; SPICE AND FLAVORING. Perry D. Strausbaugh; Earl L. Core

The principal types of commercially used vanilla beans are the Mexican, Bourbon [Bourbon comes mainly from Madagascar, but was named after the island of Bourbon (now Island of Réunion) in the Indian Ocean, where the French started the cultivation of vanilla], South American, Javan, and Tahitian.

Vanilla extract is prepared from vanilla beans with or without one or more of the following added: sugar, dextrose, glycerol. Vanilla extract contains the soluble matters from not less than 3.3 oz of vanilla beans in 1 qt (10 g/100 ml). To be legally called vanilla extract, 1 U.S. gal (3.785 liters) of vanilla extract must contain the soluble matter from not less than 13.35 oz (378.5 g) of vanilla beans. The finished flavoring should contain at least 35% alcohol by volume to keep the solubles in solution.

In the alcoholic extraction of the vanilla flavor, the color of the extract is influenced by the quality of the beans, the strength of the alcoholic menstruum, the duration of the extraction, and the presence of glycerin, which is added to retard evaporation and to retain the flavor of the extract. Best results are obtained with three consecutive extractions at room temperature, each requiring a minimum of 5 days. The first should have a maximum alcohol content of 65%; the second, 35%; the third, 15%. To improve aroma, extracts are aged, using stainless steel or glass containers.

A standard vanilla extract is equivalent in flavoring strength, though not in quality, to a 0.7% vanillin solution. The vanillin content of pure extracts range from 0.04 to 0.12 oz/qt (0.11 to 0.35 g/100 ml), with the average at about 0.06 oz/qt (0.19 g/100 ml). Ash content, soluble ash, lead number, total acidity, and acidity other than vanillin are among the conventional indices used to detect adulteration.

The approximate average maximum levels of usage (in ppm) for vanilla extract in foods are as follows:

Beverages	200	Baked goods	1900
Ice cream, ices	3000	Syrups	54
Candies	4000	Icings	4800
Toppings	2700		

Gideon E. Livingston; Myron Solberg

Bibliography. H. B. Heath and G. A. Reineccius, *Flavor Chemistry and Technology*, 1986; R. Teranshi, R. G. Buttery, and F. Shahadi (eds.), *Flavor Chemistry: Trends and Developments*, 1989.

Vapor condenser

A heat-transfer device that reduces a thermodynamic fluid from its vapor phase to its liquid phase. The vapor condenser extracts the latent heat of vaporization from the vapor, as a higher-temperature heat source, by absorption in a heat-receiving fluid of lower temperature. The vapor to be condensed may be wet, saturated, or superheated. The heat receiver is usually water but may be a fluid such as air, a process liquid, or a gas. When the condensing of vapor is primarily used to add heat to the heat-receiving fluid, the condensing device is called a heater and is not within the normal classification of a condenser. *See* HEAT TRANSFER.

Classification by use. Condensers may be divided into two major classes according to use: those used as part of a processing system, and those used for serving engines or turbines in a steam power plant cycle.

Process condensers. In a processing system condensers selectively recover liquid from a mixed vapor, recover pure liquid from the vapor of an impure liquid, recover noncondensables from a mixture of gas and vapor, or extract heat from heat-pump, refrigeration, and cryogenic cycles. The vapor condensed may be steam or the vapor from any nonaqueous liquid. Condensers used for any purpose except as vapor condensers in the power plant cycle, even though they provide the heat sink for a process cycle, are classified as process condensers.

One of the more important uses of process condensers is that of condensate recovery. Condensers are frequently used with fractionating and distillation columns in the production of hydrocarbon liquids and in processes for liquefying gases. Similarly they are used in the production of distilled water as distiller condensers in single- and multiple-effect evaporating systems. The multistage flash evaporator with integral vapor condensers has been established as one of the more economical means for the production of potable water from seawater. This apparatus contains a multiplicity of condensers in series, each condenser producing a successively lower pressure in which flashed vapor from seawater brine is produced, condensed, and collected. Both distillate and the flashing brine streams reduce in temperature as they flow through the apparatus (**Fig. 1**).

Power cycle condensers. Used as part of the power plant cycle, condensers serve as the heat sink in the cycle and reduce the back pressure on the turbine or engine so that a maximum of heat energy becomes available as useful work. Steam (water) is the most common substance used for the power generation cycle. Other thermodynamic fluids can be used but seldom are. Another important function of a condenser in the steam power plant cycle is the recovery of condensate as the major source of pure boiler feedwater for the boilers. *See* BOILER FEEDWATER; THERMODYNAMIC CYCLE.

Classification of operation. Condensers may be further classified according to mode of operation as surface condensers or as contact condensers.

Surface condensers. In surface condensers the condensing vapor and the cooling fluid remain separated from each other by a dividing wall or walls, which form the heat-transfer surface. This heat-transfer surface is most often in the form of tubes but may also be plates or partitions of various geometries. Condensate, cooling fluid, and noncondensable gases are usually removed separately, although in some designs the condensate and noncondensable gases are withdrawn as a mixture. *See* SURFACE CONDENSER; VACUUM PUMP.

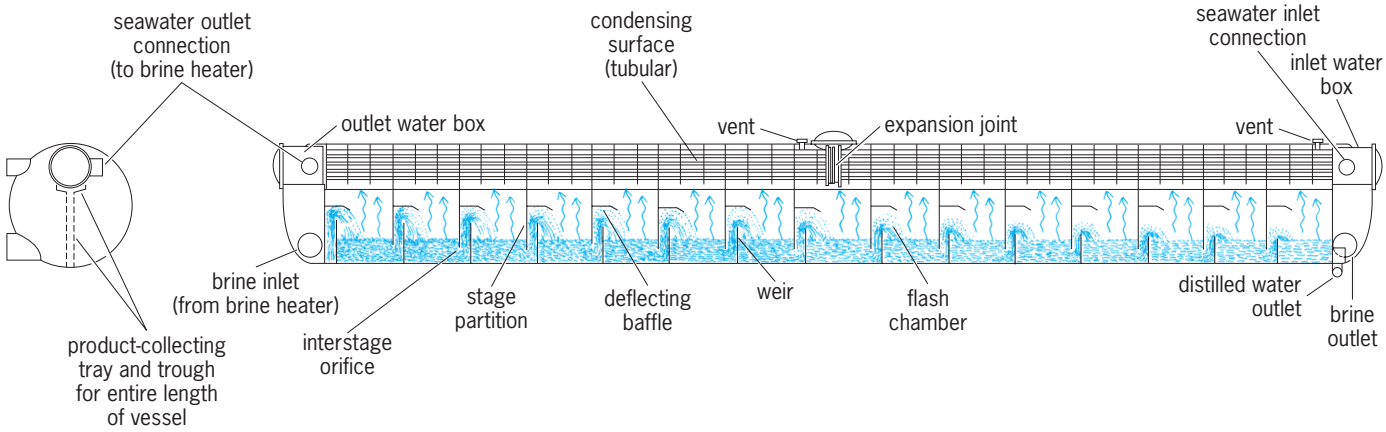

Fig. 1. Simplified views of multistage flash evaporator for production of potable water from seawater.

Contact condensers. In the contact condenser the vapor and cooling liquid come into direct contact with each other and are mixed in the condensing process. The condensed vapor and cooling liquid combine and are withdrawn from the apparatus together. The noncondensable gases are usually withdrawn separately, although in some contact condensers they are entrained in the mixed condensate and cooling liquid and removed through a common outlet. *See* CONTACT CONDENSER.

Removal of noncondensables. Condensers are required, almost without exception, to condense impure vapors, that is, vapors containing air or other noncondensable gases. Because most condensers operate at subatmospheric pressures, air leaking into the apparatus or system becomes a common cause for vapor contamination. Many process condensers are used to condense vapors whose noncondensable impurities are relatively independent of air that may leak into a vacuum system. Because the vapor supplied to the condenser is continuously reduced to liquid, the noncondensable gases in the gas-vapor mixture collect and concentrate. Their accumulation seriously affects heat transfer, and means must be provided to direct them to a suitable outlet. Most surface and contact condensers are arranged with a separate zone of heat-transfer surface within the condenser and located at the outlet end of the vapor flow path for efficient removal of the noncondensable gases through dehumidification. A vapor flow path of the displacement type, free from zones of stagnation and short circuiting, is essential for achieving maximum condensing heat-transfer rates (**Figs. 2** and **3**).

Separate external vapor condensers arranged in series with the vapor flow path of the main condenser are used when the ratio of noncondensable gases to condensing vapor is high, or when the vapor content of the noncondensable gases must be reduced to low values.

Removal of noncondensables from condensers operating at subatmospheric pressure requires vacuum pumps. The noncondensables are removed from condensers operating above atmospheric pressure by venting to atmosphere or to aftercondensers to reduce further the moisture content of the removed gases before discharging them to atmosphere or to a recovery process.

Heat-receiving fluid. Water is the most commonly used liquid for absorbing heat from condensing vapors. Liquid hydrocarbons and other chemical compounds in liquid state are used primarily as heat receivers for condensers, used as a part of a petroleum-refining or chemical-manufacturing process. Because of natural evaporation, surface waters are normally at a lower temperature than ambient air when air temperatures exceed 32°F (0°C), and thus provide the lowest-temperature cooling medium readily available. High specific heat, ease in pumping, and rapid heat transfer characterize water as an excellent medium for use as a heat receiver in a condensing system. The tendency of water to corrode metal surfaces is not usually a serious disadvantage. Thin-walled heat-transfer surfaces of the less costly corrosion-resistant alloys are generally satisfactory. Ferritic materials used for the water-containing parts of condensers can be made suitable for the corrosion environment by increasing thickness as a corrosion allowance or by the use of corrosion-resistant linings or by cathodic protection. *See* CORROSION.

Air is the gas most commonly used for absorbing heat from a condensing process. However, its low heat capacity, low density, and relatively low heat transfer preclude it from being an ideal fluid for this purpose. Its availability and the fact that air does not readily corrode or foul heat-transfer surfaces offer some advantages in its use. Shortages of water for cooling or for use as a heat sink and the danger of thermal pollution of surface water have markedly

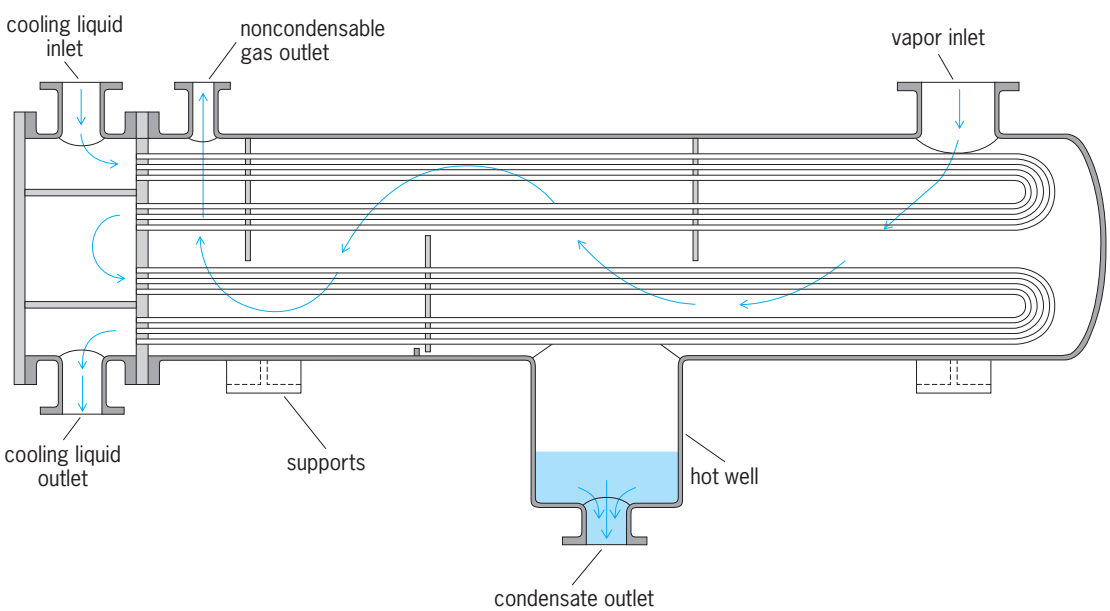

Fig. 2. Process condenser with one baffled-shell pass of condensing vapor and four tube passes of cooling liquid.

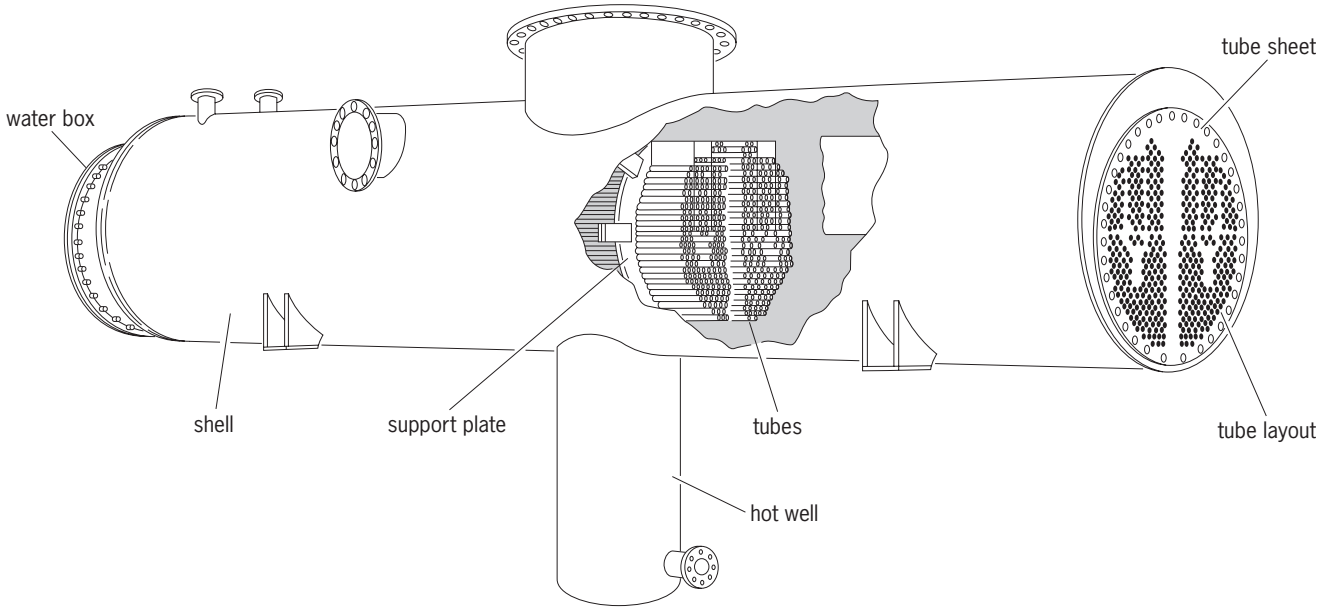

Fig. 3. Typical sections of small steam surface condenser.

increased the use of air. Application of extended-surface tubes to process condensers, to process coolers, and to steam condensers serving turbines is reasonably economical and practical where water supplies are critical.

The cooling of heat-transfer surfaces by evaporation involves the use of both water and air. The heat-transfer surfaces are continuously wetted by water. Air, blown over the wetted surfaces, absorbs the water vapor as it is released from the evaporating cooling water, thereby removing the heat released in condensers.

Condensate cooling. In those process systems where recovery of condensate is of primary importance, it is also desirable to cool the condensate below the saturation temperature of its condensing vapor. Some condensate cooling results as condensate falls from tube to tube in the process of being collected and accumulated. If additional cooling is needed, condensate cooling sections in the condenser can be provided by flooding a selected amount of surface within the condenser with condensate. Flooding is done either by installing baffles in the condensate drain zone to flood the tubes in these sections or by effecting a similar result with a loop in the condensate drain piping. Baffles are most effective in process condensers of the horizontal type, and loops are more adaptable to process condensers of the vertical type.

Condensate reheating. Modern steam surface condensers for turbines are provided with means for condensate reheating (**Fig. 4**). Condensate falling from tube to tube becomes cooled below the temperature corresponding to condenser operating pressure; unless provision is made for reheating this condensate before it is removed from the condenser, a measurable amount of heat energy is lost from the cycle, and cycle efficiency is reduced. In addition, subcooled condensate absorbs air from the condensing vapor and, if allowed to remain in the boiler feedwater, may cause serious corrosion of the feed system and boilers.

The condensate is reheated by a portion of the incoming steam. Baffles direct this steam to the condenser hot well in such a manner that its velocity

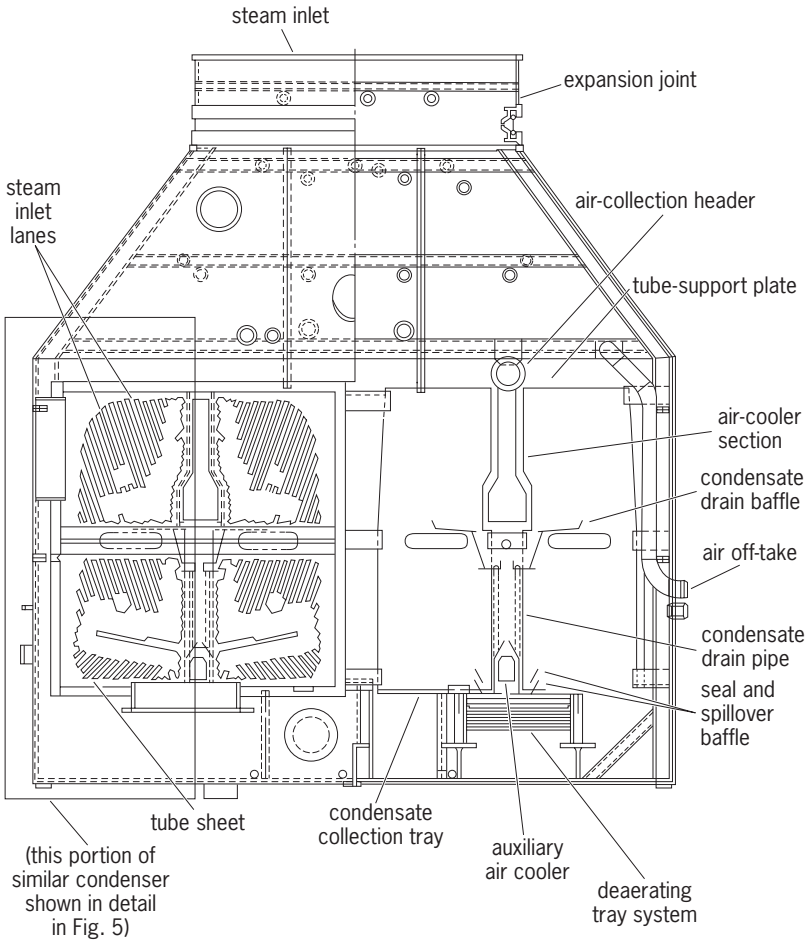

Fig. 4. Cutaway end view of steam surface condenser. (*Worthington Corp.*)

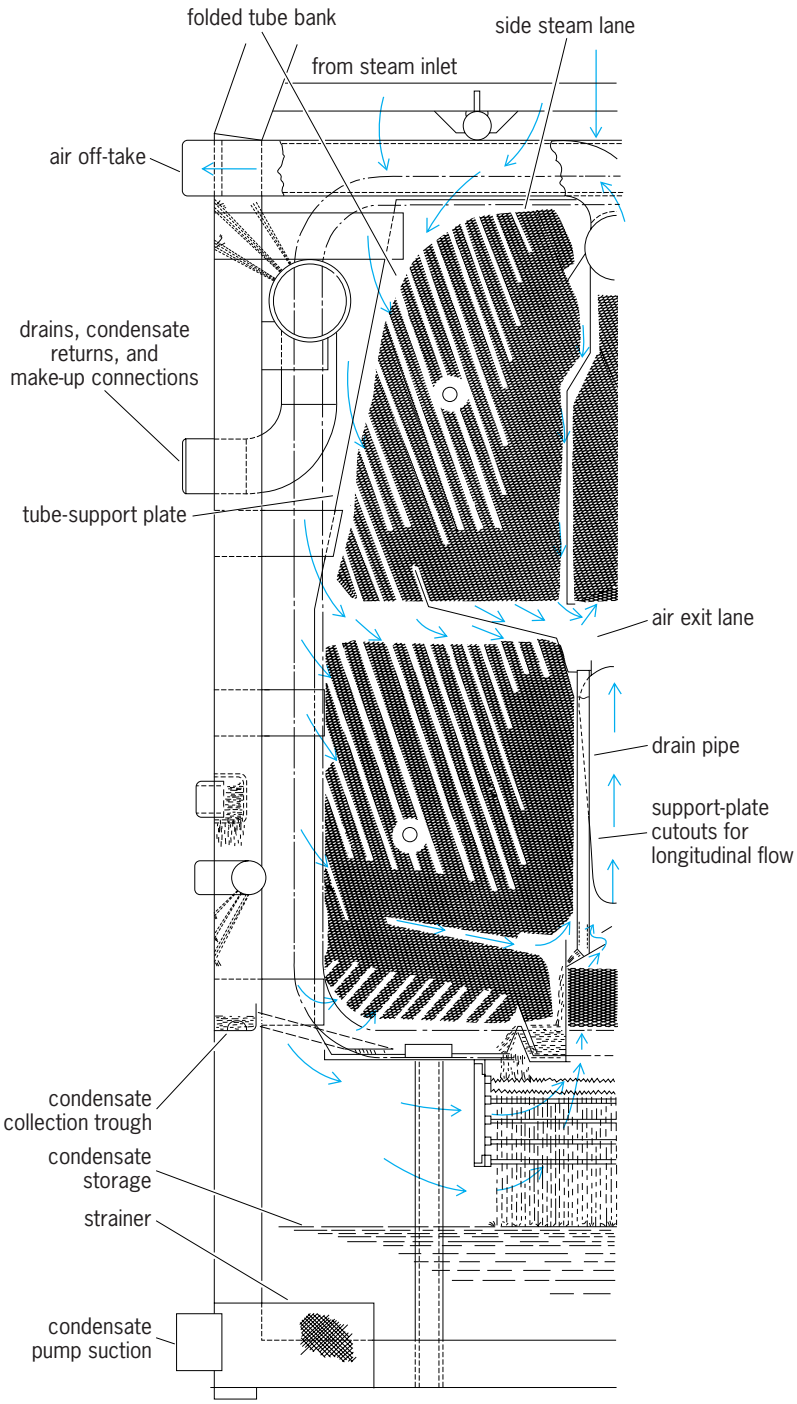

Fig. 5. Detail of a portion of a condenser that is similar to the one in Fig. 4. (*Worthington Corp.*)

energy is converted to pressure (**Fig. 5**). As a result, the local static pressure and corresponding temperature become equal to, or greater than, the static pressure and corresponding temperature at the condenser stream inlet. Condensate falling from the tube bundle to the hot well is reheated, the degree of reheat depending on the height of fall and the flow rate per unit area. If vertical height is limited, effective reheating and the associated deaeration can be achieved by redirecting the condensate falling from the tubes over collecting baffles and a series of trays or plates. This higher zone of pressure and temperature is vented along with the noncondensables removed from the condensate into the condenser tube bank and finally to the air cooler joining the mainstream of noncondensables being cooled and expelled. Under favorable conditions the condensate may reach temperatures as much as 5°F (3°C) above the temperature corresponding to static pressure at the condenser steam inlet, and the dissolved oxygen content of the deaerated condensate may be consistently less than 0.01 cm^3 per liter.

Condenser capacity. The quantity of cooling fluid required to condense a given quantity of vapor may be obtained by applying the general heat-balance equation. For a condensing vapor which may be superheated, saturated, or wet, with its condensate subcooled or reheated and with the noncondensable gases cooled, the relation is given by Eq. (1). Here

$$W_v(h_v - h_c) + W_n(h_{na} - h_{nb}) = W_f(h_{fb} - h_{fa}) \quad (1)$$

W_v = quantity of vapor; h_v = enthalpy of entering vapor; h_c = enthalpy of leaving condensate; W_n = quantity of noncondensable gases; h_{na} = enthalpy of entering noncondensables; h_{nb} = enthalpy of leaving noncondensables; W_f = quantity of cooling fluid; h_{fa} = enthalpy of entering cooling fluid; and h_{fb} = enthalpy of leaving cooling fluid. (The standard engineering practice is to express quantities in lb/h and enthalpies in Btu/lb.)

The temperature of the leaving cooling fluid is usually selected to be 5–10°F (3–6°C) less than the condensing temperature of the saturated vapor. Ordinarily, low temperature differences are associated with contact condensers; larger temperature differences are characteristic of economic design for surface condensers. Small temperature differences require enormous amounts of condensing surface and are seldom a practical design criterion.

For both contact and surface condensers the conditions for determining the hourly heat transferred can be expressed by Eq. (2) for undirectional steady-

$$Q = UA\ \Delta t_m \quad (2)$$

state heat flow. Here Q = hourly heat, Btu/h; U = overall heat-transfer coefficient, Btu/(h)(ft^2)(°F); A = area of heat-transfer surface, ft^2; and Δt_m = logarithmic mean-temperature difference, °F. Δt_m is based on the assumption of constant specific heat and substantially constant condensing temperature with the cooling-fluid temperature increasing from inlet to outlet in its flow path. U and A are generally applicable to surface condensers only.

For direct-contact condensers where surface is not readily determined, it is usual to consider heat transferred per unit volume of condensing space rather than per unit area. In this case Eq. (3) describes the

$$Q = KV\ \Delta t_m \quad (3)$$

heat flow, where K = volumetric heat-transfer coefficient, Btu/(h)(ft^3)(°F); V = volume of direct contact

condensing space, ft^3; and the other symbols are as previously defined.

The magnitude of K depends on the physical properties of the condensing vapor and heat-receiving fluid, the type of direct-contact surface (spray, tray, or packing), and the operating temperature. It is ordinarily determined experimentally for each type of contact condenser.

The magnitude of U, for use in Eq. (2), can be obtained from the summation of Eq. (4) for the individ-

$$U = \frac{1}{r_v + r_l + r_f + r_w} \tag{4}$$

ual resistances in the heat-transfer system, resistance being expressed in (h)(ft^2)(°F)/Btu. In Eq. (4), $1/r$ = conductance across any boundary, Btu/(h)(ft^2)(°F); r_v = condensing boundary resistance; r_l = cooling-fluid boundary resistance; r_f = fouling, dirt, or scale resistance; and r_w = separating wall, resistance.

Conductance of the condensing boundary $1/r_v$ can be determined analytically for pure vapors. For steam on the outside of horizontal tubes, $1/r_v$ may be 2000–2500 Btu/(h)(ft^2)(°F) [11,400–14,200 J/(s)(m^2)(°C)] for well-designed multitube condensers. Conductance of the cooling-fluid boundary $1/r_l$, for fluids flowing inside of tubes, may be determined from the relation of the Nusselt number to the product of the Reynolds and Prandtl numbers with the use of empirical constants such as those suggested by W. H. McAdams. Excellent agreement between computed values and test values results from the use of these equations. For water inside of tubes $1/r_l$ may be in the order of 1000–1400 Btu/(h)(ft^2)(°F) [5700–8000 J/(s)(m^2)(°C)] for the normal range of water velocities.

Values for conductance of the condensing boundary and the cooling-fluid boundary for some common thermodynamic fluids are listed in the **table** for cooling-fluid velocities in the range of 6–7 ft/s (1.8–2.1 m/s), condensing temperatures in the range of 100°F (38°C), and condensing rates of approximately 10 lb/(ft^2)(h) [1.4×10^{-2} kg/(m^2)(s)]. *See* CONDUCTION (HEAT); CONVECTION (HEAT).

Conductance for fouling or dirt films $1/r_f$ depends on the characteristics of the cooling fluid and the condensing vapor in relation to the accumulation (of dirt and corrosion products on the separating wall or tube surfaces) in service. In addition, the oxide film on new clear tubes or other types of metal heat-transfer surfaces contributes to overall fouling resistance r_f. For water-cooled shell and tube steam condensers a conductance of 2000 Btu/(h)(ft^2)(°F) [11,400 J/(s)(m^2)(°C)] is reasonable for copper-base alloy tubes after mechanical cleaning. Conductance of the separating wall $1/r_w$ may be computed from the thickness of the wall and the thermal conductivity of the material from which it is made. The overall heat-transfer coefficient varies with the physical properties, flow rates of the fluids, and geometry of the condenser and condensing surfaces. For steam condensers U may be in the order of 350–800 Btu/(h)(ft^2)(°F) [2000–4500 J/(s)(m^2)(°C)]; with few exceptions, heat transfer is lower with condensing vapors other than steam.

Conductance values for some thermodynamic fluids*

Fluid	Condensing vapor conductance	Cooling fluid conductance
Isopropyl alcohol	400	360
Benzene	600	520
Water	2500	1400
Ammonia	3200	2300

*Btu/(h)(ft^2)(°F) = 5.8 J/(s)(m^2)(°C).

Condenser components. The components of a representative contact condenser are shown in **Fig. 6**. The cooling-liquid distribution system shown consists of baffles and impingement surfaces to distribute the liquid uniformly and to allow it to cascade in counterflow relationship with the condensing vapor. In the illustrated condenser the cooling section for noncondensable gases coincides with the coolant distribution section where vapor is condensed and the noncondensable gases are dehumidified and concentrated. Other constructions for distributing the coolant include rings or slats made of metal, plastic, or ceramic. Most process condensers of the contact type are of counterflow design. In jet condensers noncondensables are entrained in the cooling liquid and thereby removed. The tail pipe provides the barometric leg and discharges into a sealing well, thus eliminating the need for an extraction or tail pump. Low-level contact condensers, those without barometric legs, usually require pumps to extract the cooling water.

Components of typical surface condensers are shown in Figs. 2, 3, 4, 5, and **Fig. 7**. The condensing surface consists of tubes of 0.5–1.5-in. (13–38-mm) outside diameter made of copper base alloys, less frequently of aluminum, nickel alloys, chromium steel, chromium-nickel steels, and titanium. The cooling surface for noncondensables is ordinarily the same material as the main condensing surface, although special corrosion-resistant alloys may be used when the noncondensables are especially corrosive. Usual practice is to condense the vapors on the outside of tubes (Figs. 2, 3, 4, and 5), but when the vapors are especially corrosive they may be condensed on the inside of tubes (Fig. 7). Tubes are usually plain, but for condensing vapors from low-conductance fluids, finned tubes may be used. For this purpose low fins, approximately equal in height to basic tube wall thickness, are used on the condensing side spaced so that the condensate formed does not bridge the fins and act like insulation. Gas- or air-cooled condensers use high fins on the cooling side with the extended surface generally at least 10 times the equivalent plain tube area. Bimetal tubes are most frequently used in process condensers when the corrosion environment is severe and different on one side of the tube than on the other.

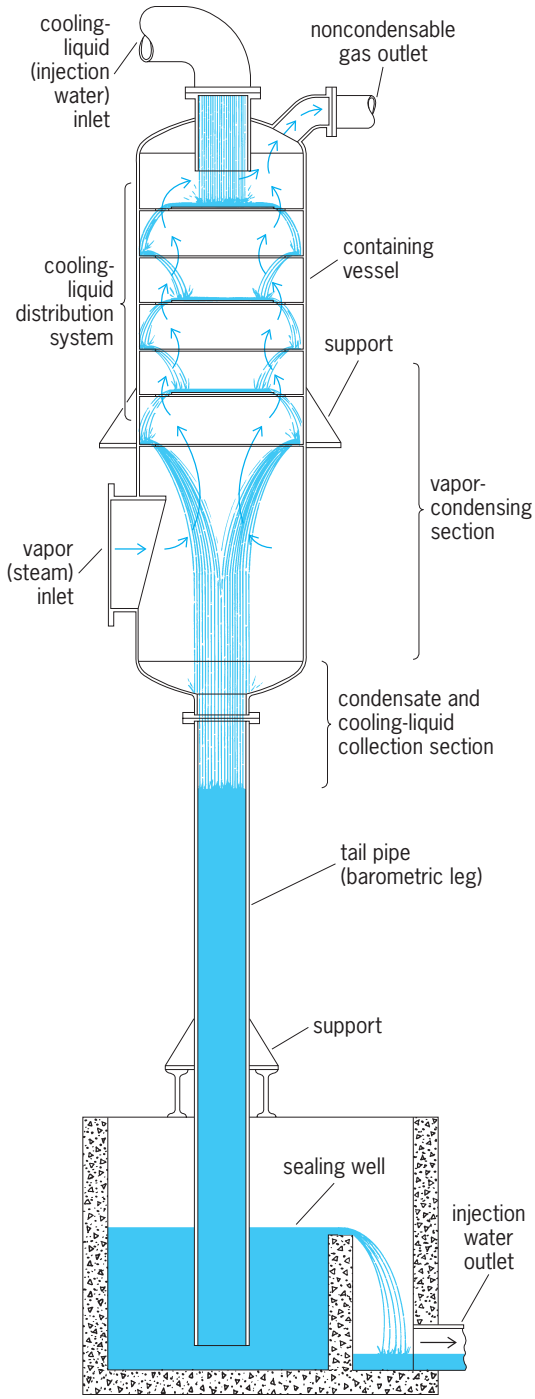

Fig. 6. Counterflow barometric-type contact condenser.

Tube vibration of magnitudes sufficient to cause tube failure is controlled by the spacing of tube support plates or baffles. Tube vibration seldom results from transmitted mechanical energy but rather from vapor velocity, causing tubes to deflect and vibrate at their natural frequency. With low-density vapors at high velocity, generally in the sonic range or greater, tube vibration is caused by the lift and drag effects of the vapor flowing around the tubes. Severe vibration is usually independent of the Kármán vortex street with low-density vapors, but may be associated with both the lift and drag effects of velocity and the Kármán vortex street at high-vapor densities even at relatively low velocities. *See* VORTEX.

The tube sheets (into which tubes are welded, expanded, or packed) are generally of Muntz metal, naval brass, copper-silicon alloy, or other corrosion-resistant alloys. Carbon-steel tube sheets may be used with carbon or chromium-nickel steel tubes. In nuclear power plants steam condenser tube sheets may be double, that is, with an air gap between to detect leakage and to prevent cooling water from leaking into the steam space. The inner tube sheets are usually made of carbon steel and the outer sheets are made of copper alloy.

The cooling fluid system for surface condensers, when liquid is used, usually consists of chambers attached to the tube sheets and arranged with inlet and outlet connections for the circulation of the cooling medium. When this fluid is liquid, the chambers are designed to distribute the liquid over the face of the tube sheets and into the tubes with little or no cavitation.

Condensers may be designed with one or more liquid passes, dependent on the thermal design conditions, the quantity of cooling liquid available or

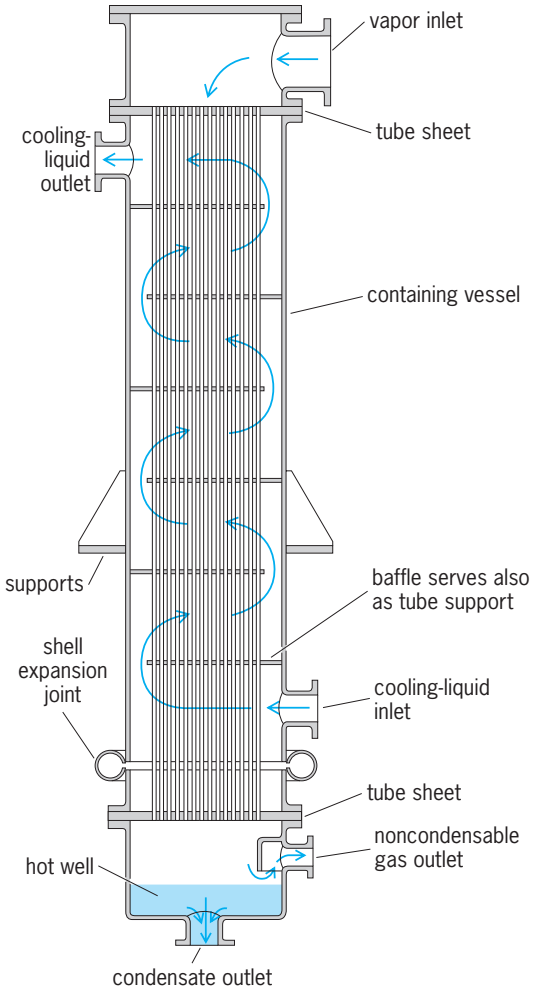

Fig. 7. Vertical-process surface condenser with condensing vapor inside the tubes and the cooling liquid outside the tubes, within the shell.

desired, and the space conditions for installation. Steam surface condensers in large steam power plants are of single-pass design except where the cooling-water supply is limited, in which case two-pass condensers are used. Large installations (those that require cooling towers, spray, or evaporation ponds as a means for controlling cooling-water temperature) ordinarily use two-pass condensers. Condensers with three or more passes seldom prove economical for serving engines or turbines. Process condensers are frequently designed with more than two cooling liquid passes and seldom with less.

Auxiliary equipment. Operation of condensers requires pumps (1) for injecting or circulating the cooling fluid, (2) for removing condensate or mixed condensate and injection water, and (3) for removing noncondensables.

Centrifugal pumps are usually used for injecting or circulating cooling water. Conventional volute pumps are used for circulating cooling water for small surface condensers and as injection water pumps for contact condensers. Mixed-flow volute pumps and axial-flow pumps of vertical design are best suited to large surface condensers.

Centrifugal pumps of the horizontal volute type, equipped with pressure-sealed stuffing boxes, are well suited for removing condensate. Vertical, multistage condensate pumps are more suited to large steam power plant installations, where they effect significant installation economies. Positive displacement pumps are used for withdrawing condensate from small condensers. Tail pumps of the conventional double-suction centrifugal type are used for pumping mixed condensate and injection water from low-level jet (contact) condensers. *See* PUMPING MACHINERY.

Pumps for removing noncondensable gases are classified as displacement and ejector types. The displacement machine is built either as a reciprocating vacuum pump, similar to a piston-type air compressor, or as a rotary machine, similar to a gear-pump or sliding-vane rotary compressor. The displacement-type vacuum pump is widely used on condensers of all sizes. It is not suitable for use at high vacuum with high noncondensable gas loads because it must be disproportionately large for such conditions. It is suitable for use with extremely large surface condensers serving turbines, because of the low noncondensable gas loads characteristic of these installations.

Steam jet ejectors are widely used as vacuum pumps. Having no moving parts, they require little or no operating attention and are simple to install. They have excellent capacity characteristics at high vacuum and are especially suited to applications where the noncondensable gas load is high. When used with steam condensers serving turbines, steam jet ejectors are equipped with surface inter- and aftercondensers. The exhaust steam is condensed and returned to the feed system, and the heat from the exhaust is also recovered in the feedwater. Thus they become highly efficient machines for removing noncondensable gases from condensers used in the generation of power. *See* CENTRIFUGAL PUMP; COOLING TOWER; DISPLACEMENT PUMP. Joseph F. Sebald

Bibliography. A. J. Chapman, *Heat Transfer*, 4th ed., 1984; J. P. Holman, *Heat Transfer*, 9th ed., 2001; L. Iev and I. Iev, *Heat Transfer Flow Dynamics*, 1990; J. R. Simonson, *Engineering Heat Transfer*, 2d ed., 1989.

Vapor cycle

A thermodynamic cycle, operating as a heat engine or a heat pump, during which the working substance is in, or passes through, the vapor state. A vapor is a substance at or near its condensation point. It may be wet, dry, or slightly superheated. One hundred percent dryness is an exactly definable condition which is only transiently encountered in practice. Vapor behavior deviates so widely from the ideal gas laws that calculation requires the use of tables and graphs that give the experimentally determined properties of the fluid.

Power and refrigeration plants. A steam power plant operates on a vapor cycle where steam is generated by boiling water at high pressure, expanding it in a prime mover, exhausting it to a condenser, where it is reduced to the liquid state at low pressure, and then returning the water by a pump to the boiler (**Fig. 1**).

In the customary vapor-compression refrigeration plant, the process is essentially reversed with the refrigerant evaporating at low temperature and pressure, being compressed to high pressure, condensed at elevated temperature, and returned as liquid refrigerant through an expansion valve to the evaporating coil (**Fig. 2**).

The Carnot cycle, between any two temperatures, gives the limit for the efficiency of the conversion of heat into work (**Fig. 3**). This efficiency is independent of the properties of the working fluid. Although the thermal efficiency is independent of the

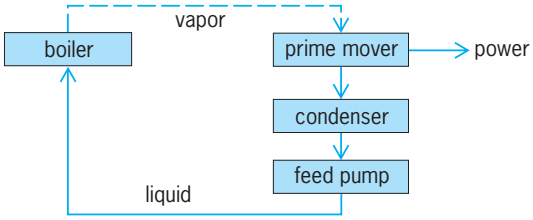

Fig. 1. Rudimentary steam power plant flow diagram.

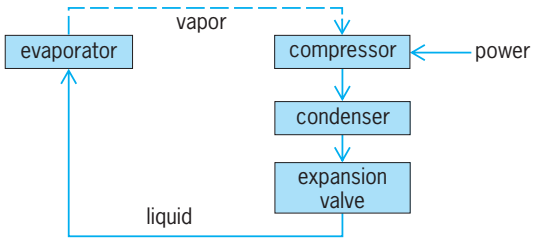

Fig. 2. Rudimentary vapor-compression refrigeration plant flow diagram.

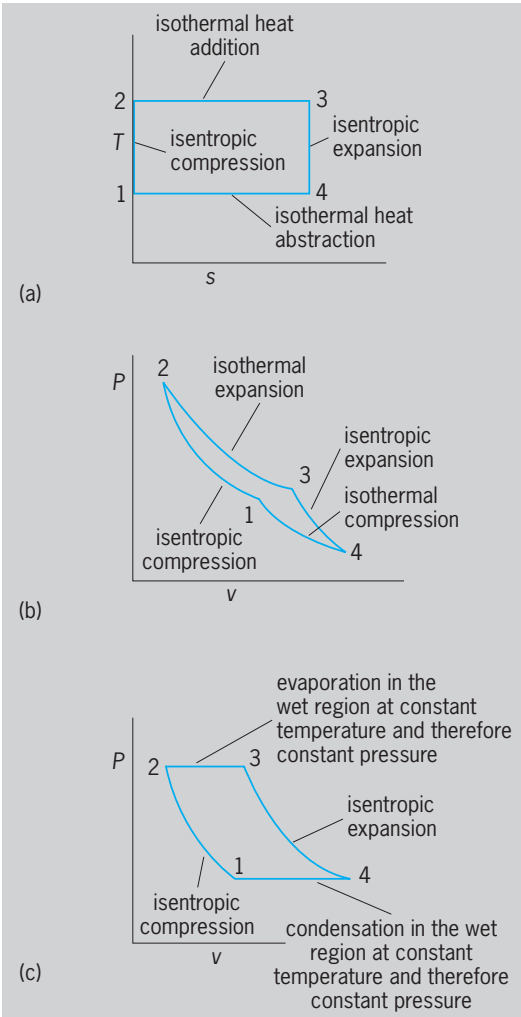

Fig. 3. Carnot cycle. (*a*) Temperature-entropy. (*b*) Pressure-volume for fixed gas. (*c*) Pressure-volume for vapor.

properties of the fluid, the mean effective pressure (and consequent physical dimensions of the engine) will be vitally influenced by choice of fluid (compare Fig. 3*b* with *c*). The Carnot cycle is not realistic for the evaluation of steam power plant performance because the cycle precludes the use of superheat and calls for the isentropic compression of vapor. It is useful, however, for specifying the limiting efficiency that a real cycle might approach. It is so used in judging performance of vapor-cycle heat engines and heat pumps. *See* CARNOT CYCLE.

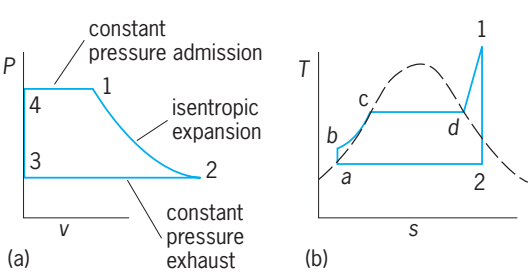

Fig. 4. Rankine cycle for heat-engine plant. (*a*) Pressure-volume diagram of prime mover. (*b*) Temperature-entropy diagram for the power plant.

The Rankine cycle is more realistic in describing the ideal performance of steam power plants and vapor-compression refrigeration systems.

Vapor steam plant. In the case of the steam power plant (Fig. 1), the Rankine cycle (**Fig. 4**) has two constant pressure phases joined by a reversible adiabatic (isentropic) phase 1–2. From the properties of the fluid, the work of the prime mover, ΔW_{PM}, is most conveniently evaluated as in Eq. (1), where h

$$\Delta W_{PM} = h_1 - h_2 \tag{1}$$

is the enthalpy, Btu/lb. The feed pump uses some of this work, ΔW_{FP}, to return the water from the condenser to the boiler so that the net output of the cycle is $\Delta W_{PM} - \Delta W_{FP}$. This net output can be related to the heat that must be added to produce steam by consideration of the *T-s* diagram (Fig. 4*b*). The area under line *bcd*–1 is the heat supplied in the boiler (heat source); phase 1–2 is the isentropic expansion in the prime mover; the area under line 2–*a* is the heat rejected to the condenser (heat sink); phase *a*–*b* is the isentropic compression of the liquid in the feed pump. Thus, the thermal efficiency is given by Eq. (2).

$$\frac{\text{Work done}}{\text{Heat added}} = \frac{\Delta W_{pm} - \Delta W_{FP}}{h_1 - h_a - \Delta W_{FP}} \tag{2}$$

There are many variables which influence the performance of the Rankine cycle. For steam the thermal efficiency is a function of pressure, temperature, and vacuum (**Fig. 5**). High pressure, high superheat, and high vacuum lead to high efficiency. *See* RANKINE CYCLE.

Vapor refrigeration plant. The Rankine cycle can be used to evaluate the performance of the vapor-compression system of refrigeration (Fig. 2). A counterclockwise path is followed on the *P-v* and *T-s* cycle diagrams (**Fig. 6**). The refrigerant enters the compressor as low-temperature, low-pressure vapor (4–1); isentropic compression follows (1–2), and then high-pressure delivery (2–3). The work to drive the compressor is $h_2 - h_1$. The machine cooling coefficient of performance *cp*, which is essentially the reciprocal of thermal efficiency, is given by Eq. (3); for a warming machine, it is given by Eq. (4).

$$cp = \frac{\text{refrigeration}}{\text{work done}} = \frac{h_1 - h_d}{h_2 - h_1} \tag{3}$$

$$cp = \frac{\text{heat delivered}}{\text{work done}} = \frac{h_2 - h_c}{h_2 - h_1} \tag{4}$$

The difference $h_1 - h_d$ is heat removed in the refrigerating coils, the area under phase *d*–1 on the *T-s* diagram. The difference $h_2 - h_c$ is the heat delivered to the condensing coils, the area under phase (2*abc*). Because the flow is throttled through the expansion valve, the enthalpy is constant, $h_c = h_d$. Some ideal performance values of a vapor-compression refrigeration system are plotted in **Fig. 7**. For further details on heat-pump vapor cycles *see* HEAT PUMP.

The remainder of this article is concerned with

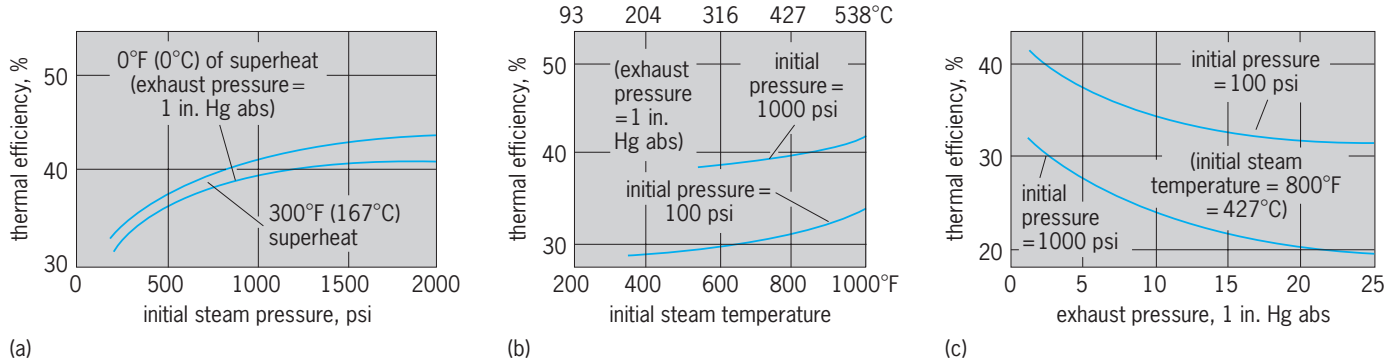

Fig. 5. Thermal efficiency of ideal Rankine steam cycle. (*a*) Steam pressure. (*b*) Steam temperature. (*c*) Vacuum. 1 psi = 6.9 kPa; 1 in. Hg abs = 3.4 kPa.

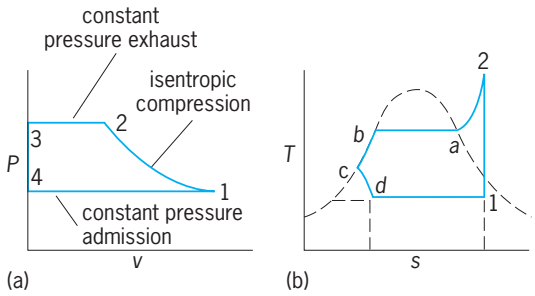

Fig. 6. Rankine cycle for heat-pump plant. (*a*) Pressure-volume diagram of compressor. (*b*) Temperature-entropy diagram for a vapor-compression refrigeration plant.

the vapor cycle as it is applied to power-generation purposes.

Regenerative heat cycle. The regenerative cycle is a modification of the simple Rankine cycle. Feed water is heated by extracted steam (**Fig. 8**). As a result, less heat must be added in the boiler to evaporate a pound of steam and in turn to deliver a kilowatt hour of work output from the associated steam engine.

A cycle with a single stage of regenerative heating can be viewed as two Rankine cycles superimposed on one another (Fig. 8*b*). In one the exhaust pressure and consequent temperature are substantially higher than in the other (point *f* versus point 2). The heat of condensation represented by the area under *c*–*f* can be used to raise the feed temperature from T_b to T_c. The area under phase *b*-*c* is smaller than under phase *c*–*f* so a fraction of a pound of steam is needed to raise 1 lb of water to the common temperature level T_c.

The principle of regeneration can be extended to multiple-stage heating with different final feed temperatures and in the limit reaching the boiler saturation temperature T_c with an infinite number of heating stages. Some consequences of the process are reflected in the data of **Fig. 9**. The gain in thermal efficiency is the consequence of reducing the quantity of heat rejected to sink 2–*a* in Fig. 8. The weight of steam flow to the prime mover for the production of a kilowatt-hour is larger than with the simple Rankine cycle (Fig. 4). But the heat required to make a pound of steam is so much less that there is an overall thermodynamic gain per kilowatt-hour. The weight flow of steam, from the prime mover to the condenser, is less than with the simple Rankine cycle. It is this reduction in heat rejected to the thermodynamic sink that raises the overall thermal efficiency of the regenerative above the nonregenerative cycle. Modern steampower practice uses the steam turbine for up to ten stages of regenerative heating.

Reheat cycle. The resuperheat or reheat cycle is another improvement in vapor cycles favored in current central station practice. Steam expanding

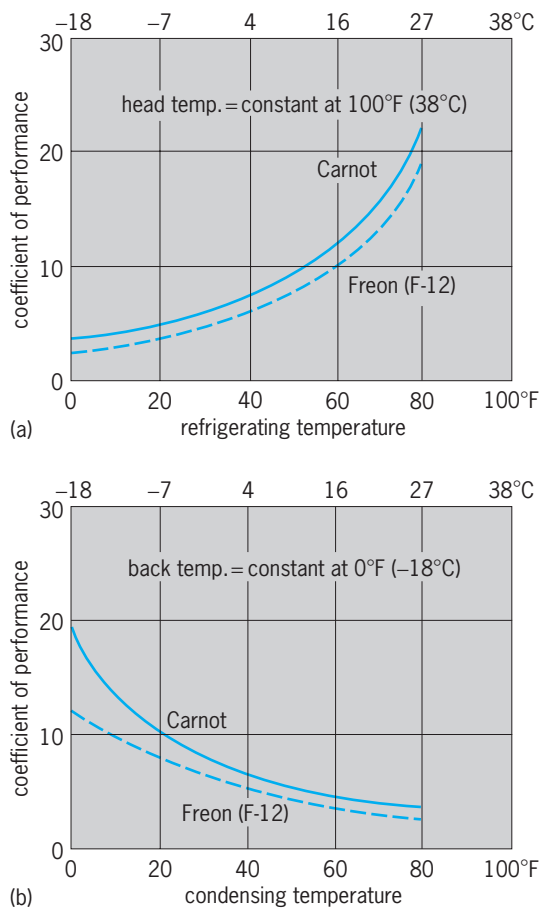

Fig. 7. Coefficient of performance of ideal Rankine and Carnot cycles as influenced by (*a*) refrigerating temperature and (*b*) condensing temperature.

isentropically (Fig. 4*b*) grows wetter with consequent increased erosion of machinery parts and loss in mechanism efficiency: Superheat tends to correct these weaknesses but metallurgical limitations fix the maximum allowable steam temperature. Reheating the steam after a partial expansion (**Fig. 10***a*) gives a practical correction. The reheating can be carried out at various pressure and temperature levels and in multiple stages. Thermal efficiency is improved over the simple Rankine cycle (Fig. 10*b*). Current practice uses a single reheat stage with temperatures approximately equal to primary steam temperature. Supercritical pressure plants favor two reheating stages.

Binary vapor cycle. Comparison of the Rankine (Fig. 4*b*), the reheat (Fig. 10*a*), and Carnot (Fig. 3*a*) cycles shows that there are considerable thermodynamic losses in the first two by failure to approach the rectangular *T-s* configuration of the Carnot cycle. The binary vapor cycle uses two fluids with totally different vapor pressures, such as mercury and water (**Fig. 11**). If a Rankine cycle using mercury is superimposed on that using steam, it is possible to operate the mercury condenser as the steam boiler, transferring the heat from the one fluid to the other. Because of the differences in the latent heats and specific heats of the two fluids, several pounds of mercury must be circulated to make 1 lb of steam. However, the combined

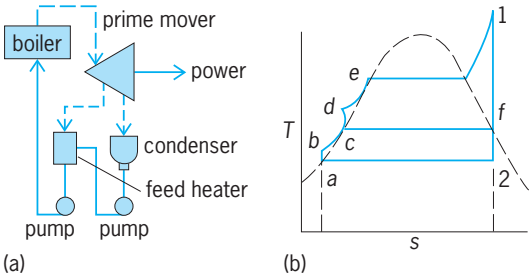

Fig. 8. Regenerative cycle. (*a*) Rudimentary steam power plant flow diagram with single-stage feed heating. (*b*) Temperature-entropy diagram.

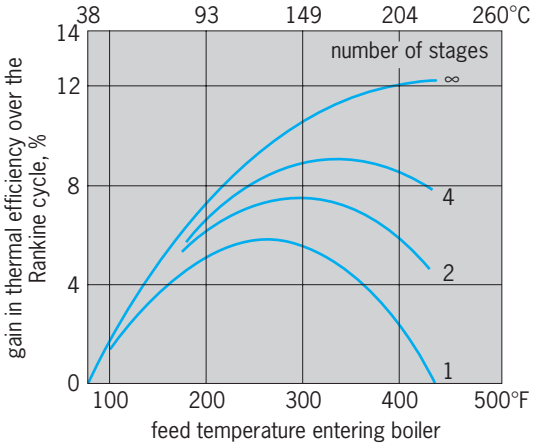

Fig. 9. Gain in thermal efficiency by use of regenerative instead of Rankine cycle; steam conditions, 400 psi (2.76 MPa) and 700°F (371°C); exhaust pressure, 1 in. Hg abs (3.4 kPa).

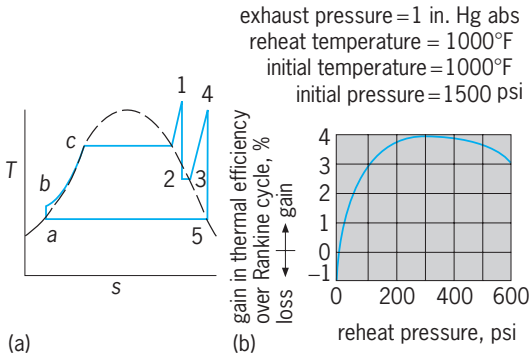

Fig. 10. Resuperheat or reheat cycle. (*a*) Temperature-entropy diagram. (*b*) Gain in thermal efficiency as function of reheat pressure; primary pressure, primary temperature, and reheat temperature constant at 1500 psi (10.3 MPa), 1000°F (538°C), and 1000°F (538°C), respectively. 1 in. Hg abs = 3.4 kPa; 1 psi = 6.9 kPa.

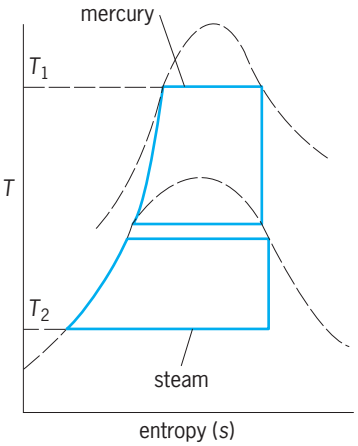

Fig. 11. Binary vapor cycle using mercury and water, temperature-entropy diagram.

T-s diagram (Fig. 11) approximates the rectangular specification with appreciable gain in thermal efficiency. For many years the binary vapor cycle was thought to have practical as well as theoretical advantages, but its economic difficulties have led to its eclipse by, and abandonment in favor of, the high-pressure, regenerative, reheat, steam cycles. *See* THERMODYNAMIC CYCLE. Theodore Baumeister

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; V. M. Faires, *Thermodynamics*, 6th ed., 1978; W. L. Haberman and J. E. John, *Engineering Thermodynamics*, 2d ed., 1989; K. A. Rolle, *Thermodynamics and Heat Power*, 4th ed., 1994; J. R. Simonson, *Engineering Heat Transfer*, 2d ed., 1989.

Vapor lamp

A source of radiant energy excited by a supply of electricity which creates a current of ionized gas between electrodes in an enclosure that contains the arc while permitting transmission of the radiant energy. Gaseous-discharge lamps or vapor lamps are given various names relating to the element

Approximate efficiencies of popular lamp types

Lamp	Nominal efficiency, lumens per watt
Mercury vapor	50
Fluorescent	70
Metal-halide	90
High-pressure sodium	110
Incandescent	20

responsible for the majority of the radiation (mercury, sodium metal-halide, xenon), to the physical attribute of the lamp (short-arc, high-pressure), or, in the case of fluorescent lamps, to the way a phosphor on the bulb wall fluoresces as a result of the lamp's low-pressure mercury-vapor excitation. *See* ARC DISCHARGE; ELECTRICAL CONDUCTION IN GASES.

Gaseous-discharge lamps are broadly and increasingly used throughout the world because the conversion of electric energy to radiant energy in a gaseous discharge provides radiation in narrow bands within the range of visible light in which the rods and cones of the eye are most sensitive. These light sources have high efficiency in conversion of electricity to light (see **table**). The popularity of high-pressure sodium lamps that emit yellow-white light for roadway lighting is due to high efficiency. This lamp's principal radiation is near the peak sensitivity of the eye, while radiation at other wavelengths is produced through a spectrum-line broadening due to high pressure.

The gaseous-discharge lamp requires a high voltage to start, but as the electric current increases, the resistance decreases, and some means must be available to limit the current and avoid lamp failure. Some gaseous-discharge lamps use a tungsten filament within the lamp envelope, which gets hot and limits the current. More commonly, a magnetic structure called a ballast (transformer) is used external to the lamp and limits the lamp current. *See* ILLUMINATION; TRANSFORMER.

For discussions of common types of vapor lamps *see* FLUORESCENT LAMP; MERCURY-VAPOR LAMP; METAL HALIDE LAMP; NEON GLOW LAMP; SODIUM-VAPOR LAMP; SUNLAMP; ULTRAVIOLET LAMP.

T. F. Neubecker

Vapor lock

Interruption of fuel flow to an engine due to blockage of passages in the fuel system by fuel vapor.

To promote easy starting, all gasolines contain volatile constituents that under some conditions, such as high ambient temperature, tend to produce more vapor than the fuel-system vents can handle. The action of an engine-mounted fuel pump, in decreasing the pressure at its inlet, tends to vaporize the fuel. If the vapor forms faster than the pump can draw it from the fuel line, the flow of fuel to a carburetor is effectively stopped and the engine stalls. Vapor lock is much less likely on a fuel-injected engine with an electric pump in the fuel tank. However, an engine with port fuel injection may experience vapor lock if the injector or fuel overheats, or if the Reid vapor pressure of the fuel is too high. *See* CARBURETOR; INTERNAL COMBUSTION ENGINE; VAPOR PRESSURE.

The tendency to form vapor in a given fuel system has been correlated with the Reid vapor pressure and ASTM distillation curve of the gasoline. With a given fuel, vapor formation can be minimized by reducing heat in the fuel system, increasing fuel-system pressure, and eliminating sudden changes in cross section or direction of fuel lines. *See* AIRCRAFT FUEL; FUEL SYSTEM; GASOLINE.

Donald L. Anglin

Bibliography. W. H. Crouse and D. L. Anglin, *Automotive Engines*, 8th ed., 1994.

Vapor pressure

The saturation pressures exerted by vapors which are in equilibrium with their liquid or solid forms. One of the most important physical properties of a liquid, the vapor pressure, enters into many thermodynamic calculations and underlies several methods for the determination of the molecular weights of substances dissolved in liquids. For a discussion of the vapor pressure relationships of solids *see* SUBLIMATION; MOLECULAR WEIGHT; SOLUTION.

If a liquid is introduced into an evacuated vessel at a given temperature, some of the liquid will vaporize, and the pressure of the vapor will attain a maximum value which is termed the vapor pressure of the liquid at that temperature. Although the quantity of liquid remaining does not diminish thereafter, the process of evaporation does not cease. A dynamic equilibrium is established, in which molecules escape from the liquid phase and return from the vapor phase at equal rates. *See* EVAPORATION.

It is important to make a distinction between the vapor pressure of a liquid, as described above, and the pressure of a vapor. The vapor pressure of a pure liquid is a unique and characteristic property of the liquid and depends only upon the temperature. A gas or vapor may, on the other hand, exert any pressure within reason, depending upon the volume to which it is confined, provided it is not in contact with its liquid phase.

Liquid-vapor equilibrium. The relationship between the vapor pressure of a liquid and the temperature is indicated by a phase diagram. The **illustration** shows the phase diagram for water, the line *OC* being the vapor pressure line for liquid water. Line *AO* is the vapor pressure line (sublimation pressure curve) for ice, and line *BO* is the liquid-solid equilibrium line. Point O is called the triple point and is the unique pressure and temperature at which a pure solid, its liquid, and its vapor can coexist in equilibrium under the pressure of the vapor alone. The triple point of water is not the familiar melting point (0°C or 32°F) but rather 0.01°C (32.018°F). The distinction is that the total pressure at the triple point is 4.58 mm (611 pascals), the common vapor pressure of solid and liquid water, whereas the

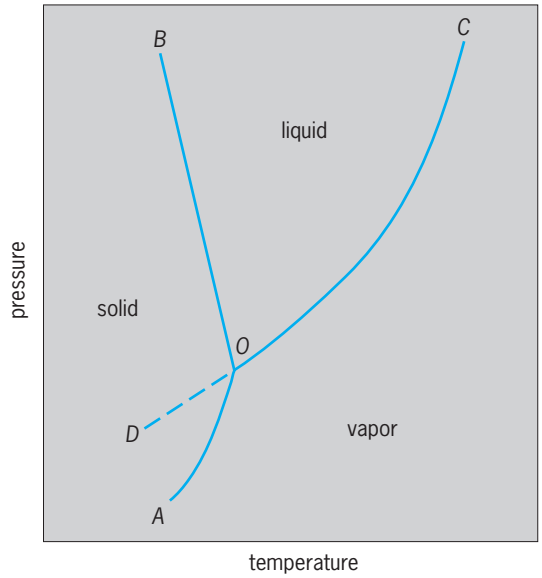

Phase diagram of the water system.

total pressure at the melting point is ordinarily 1 atm (100 kilopascals). Point *C* is called the critical point and is the point above which there is no distinction between the liquid and gaseous phases. It is not possible to liquefy a gas at temperatures above the critical temperature, regardless of the applied pressure. Necessarily, the surface tension and the latent heat of vaporization become zero at the critical point. For ordinary liquids, the vapor pressure at the critical point is usually about 50 atm (5000 kPa). There is no evidence for a critical point on the solid-liquid equilibrium line, and *B* is meant to indicate a direction, rather than a point; other phases may appear above *B*, of course, and alter the direction of the line. It is possible to undercool a liquid below its triple point if crystallization nuclei are absent. The vapor pressure of the undercooled liquid is given by the broken line *DO* and lies above the equilibrium vapor pressure of the solid. The undercooled liquid is therefore metastable, because the system tends to assume its lowest vapor pressure at equilibrium. *See* CRITICAL PHENOMENA; TRIPLE POINT.

Quantitative relations. For most liquids the relationship between the vapor pressure and temperature can be expressed by an equation having the form of (1), where a, b, and c are constants. The simpler equation (2), where m and n are constants, is often

$$\log P = a + b \log T + \frac{c}{T} \tag{1}$$

$$\log P = \frac{m}{T} + n \tag{2}$$

adequate. The change in the vapor pressure of a liquid with temperature may be expressed by the Clausius-Clapeyron equation (3), where L_v is the

$$\frac{dP}{dT} = \frac{L_v}{T\Delta V} \tag{3}$$

molar latent heat of vaporization and ΔV is the difference in molar volumes of the vapor and liquid at the temperature T. Because the molar volume of the liquid is negligible by comparison with that of the vapor except near the critical point, the Clausius-Clapeyron equation may be written as (4), where V

$$\frac{dP}{dT} = \frac{L_v}{TV} \tag{4}$$

is the molar volume of the gas. If the gas follows the ideal gas law ($PV = RT$), the equation may be further written as (5), where R is the gas constant. For

$$\frac{1}{P}\frac{dP}{dT} = \frac{L_v}{RT^2} \tag{5}$$

moderate ranges of temperature, the latent heat of vaporization is constant, and the equation may be rearranged and integrated to give the useful form of (6).

$$\log_{10} \frac{P_2}{P_1} = \frac{L_v}{2.3R}\left(\frac{1}{T_1} - \frac{1}{T_2}\right) \tag{6}$$

From the latent heat of vaporization and the vapor pressure at one temperature, the vapor pressures at other temperatures may thus be calculated.

The vapor pressure of a liquid is lowered when a substance is dissolved in it. At low solute concentrations in nonideal solutions, or at all solute concentrations in ideal solutions, the partial vapor pressure of the liquid in a solution is proportional to the mole fraction of the solvent, Eq. (7), where X_1 is the sol-

$$P_1 = X_1 P_1^0 \tag{7}$$

vent mole fraction and P_1^0 is the vapor pressure of the pure solvent. Raoult's law, as this important relationship is known, may also be expressed in the form of Eq. (8), where X_2 is the mole fraction of the solute.

$$\frac{P_1^0 - P_1}{P_1^0} = X_2 \tag{8}$$

The lowering of the vapor pressure of the solvent is proportional to the mole fraction of the solute. This relation provides the basis for the determination of the molecular weights of dissolved substances. *See* CONCENTRATION SCALES.

Vapor pressure measurement. Because the vapor pressures of liquids range widely, a number of methods have been devised for their measurement. The static method consists of introducing an excess of the liquid into an evacuated system at a given temperature, and measuring the vapor pressure with an attached mercury manometer. It is useful for vapor pressures ranging from a few millimeters of mercury up to several atmospheres. The dynamic method consists of introducing the liquid into a system in which the pressure may be varied, and noting the temperature at which the liquid boils at a given pressure. It is useful for moderately high vapor pressures. The transpiration method consists of bubbling a known volume of an inert gas at a definite pressure through the liquid. The quantity of liquid transported into the carrier gas is determined, and the vapor pressure of the liquid is calculated from

relation (9), where P_C is the carrier gas pressure

$$P_L = P_C \frac{n_L}{n_L + n_C} \tag{9}$$

and n_L and n_C are the numbers of moles of liquid vaporized and carrier gas, respectively. For the measurement of very low vapor pressures, the Knudsen effusion method may be used. This method is based upon the measurement of the mass of vapor which escapes through a very small hole in a vessel containing a liquid in equilibrium with its vapor. If the diameter of the hole is small compared with the mean free path of the gas molecules, the latter suffer no collisions in their passage through the hole. Equation (10) relates the mass of gas which passes

$$m = P\sqrt{\frac{M}{2\pi RT}} \tag{10}$$

per square centimeter per second through the hole and the vapor pressure, where m is the mass of gas per square centimeter per second and M is the molecular weight. The Knudsen method has been used with radioactive isotopes or with mass spectrometers to determine very low vapor pressures. *See* PHASE EQUILIBRIUM. Norman H. Nachtrieb

Bibliography. P. W. Atkins, *Physical Chemistry*, 6th ed., 1997; G. M. Barrow, *Physical Chemistry*, 6th ed., 1996.

Varactor

A solid-state device which has a capacitance that varies with the voltage applied across it. The name varactor is a contraction of the words variable and reactor. Typically the device consists of a reverse-biased *pn* junction that has been doped to maximize the change in capacitive reactance for a given change in the applied bias voltage. The device has two primary applications: frequency-tuning of radio-frequency circuits including frequency-modulation (FM) transmitters and solid-state receivers, and nonlinear frequency conversion in parametric oscillators and amplifiers. *See* CAPACITANCE; FREQUENCY MODULATOR; PARAMETRIC AMPLIFIER; REACTANCE.

A *pn* junction in reverse bias has two adjacent microscopic space-charge or depletion regions which function like the plates of a capacitor (see **illus.**). These depletion regions get larger as the applied reverse-bias voltage is increased; however, the increase in the width of the depletion region is not linear with bias voltage, but instead is sublinear, the exact nature of the relationship depending upon the doping profile in the *pn* junction. For example, in a *pn* junction with constant doping density, the depletion region width varies as the square root of the applied reverse-bias voltage. Because the capacitance of the device is proportional to the width of the depletion region, the nonlinear relationship between bias voltage and depletion width results in a nonlinear voltage-capacitance relationship as well. The *pn* junction doping profile is adjusted by the device designer to obtain the desired capacitive nonlinearity.

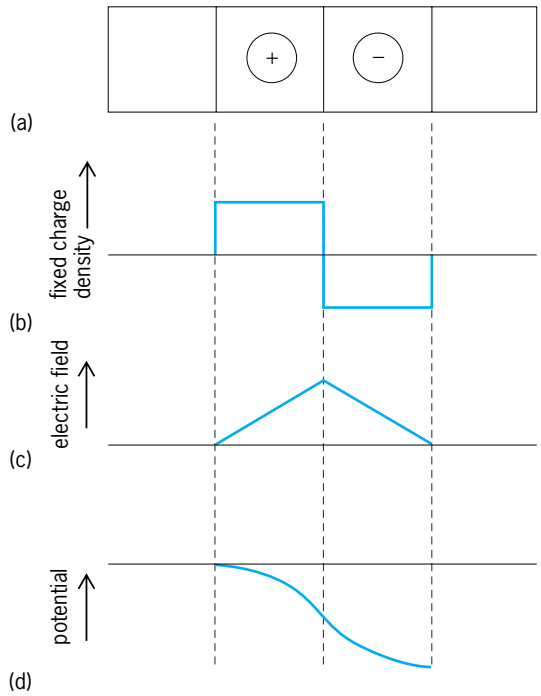

Depletion region of a *pn* junction used in a varactor diode. (*a*) Fixed charge. (*b*) Density of fixed charge stored capacitively in the depletion layer. (*c*) Built-in electric field which results from formation of the depletion region. (*d*) Electric potential.

The frequency response of the varactor is governed by the relationship between the series linear resistance of the diode and its nonlinear capacitance. The highest frequency for which the device will function properly is that at which the capacitive reactance (the reciprocal of the product of nonlinear capacitance and frequency) is equal to the series resistance of the device. Thus, designing a varactor for maximum frequency response involves choosing a doping density high enough for a small series resistance but low enough so that the capacitance of the device is small.

Varactors, as well as other solid-state devices, possess the advantage that they are compact and robust, permitting their use in hostile environments, as well as improving the reliability of the circuits in which they are employed. *See* JUNCTION DIODE; MICROWAVE SOLID-STATE DEVICES; SEMICONDUCTOR; SEMICONDUCTOR DIODE. David R. Andersen

Bibliography. D. Christiansen (ed.), *Electronics Engineers' Handbook*, 4th ed., 1996; D. H. Navon, *Semiconductor Microdevices and Materials*, 1986; L. S. Nergaard and M. Glicksman (eds.), *Microwave Solid State Engineering*, 1964; E. S. Yang, *Microelectronic Devices*, 1988.

Variable (mathematics)

A term that became part of mathematical language during the development of analytic geometry (no equivalent word is found in the writings of Archimedes and other Greek mathematicians), and which is frequently not defined in modern works,

apparently in the belief that the meaning ordinarily attached to the word is sufficiently precise. When a definition is given, it is usually to the effect that a symbol x is a variable if it may denote any member of a set S of objects. A variable is discrete or continuous according as its range (the set S) is discrete (for example, a subset of the natural numbers) or continuous (for example, all real numbers between two real numbers), respectively. Such a working definition usually suffices for the needs of mathematics. *See* ANALYTIC GEOMETRY; NUMBER THEORY; PARAMETRIC EQUATION.

Leonard Blumenthal

Variable star

A star with a detectable change in brightness that is often accompanied by other physical changes. Every star varies in brightness sometime during its life, usually in the early stages (while it is forming) or in the late stages (close to its death). Therefore, variability provides important clues about the evolution and nature of stars. Depending upon the type of star, variability in brightness can provide information about its size, radius, mass, temperature, luminosity, internal and external structure, composition, and distance from the Earth.

Hundreds of thousands of stars are known to vary in brightness. Some variable stars are very bright and can be observed with the unaided eye, while others are observable only with telescopes. The variation from maximum to minimum brightness (the range) can be anywhere from thousandths of a magnitude to many full magnitudes, and the period can be anywhere from a fraction of a second to several years. The change in brightness may be caused by intrinsic physical changes in the star or stellar system such as pulsations or eruptions; extrinsic changes, such as the eclipse of the star by one or more stars; or rotation of a star with bright or dark spots on its surface. *See* MAGNITUDE (ASTRONOMY).

Observation. Most variable stars need to be monitored in a systematic way over years and decades in order to determine their period, long-term behavior, and unusual and brightness variation changes. The simplest method of determining the brightness of variable stars is visually through a telescope or binoculars. Groups of observers worldwide, such as the American Association of Variable Star Observers (AAVSO), carry out systematic visual brightness determinations for thousands of variable stars. Group members, mostly amateur astronomers, observe with their own binoculars and telescopes. They use special star maps (finding charts) that identify the variable star and stars of constant and known brightness (comparison stars) used in estimating the brightness of the variable. Observational data provided by these dedicated observers have been used extensively by professional astronomers to analyze the behavior of the changes in brightness of variable stars, to schedule and correlate observations made with large telescopes and special instruments (both on the Earth and aboard satellites orbiting the Earth), and to make theoretical variable-star models with the aid of powerful computers in order to understand better the physical structure of these stars.

Other techniques for systematically monitoring variable stars, used by both professionals and increasing numbers of amateurs, include photographic and photoelectric photometry and charge-coupled devices (CCDs). Photographic plates provide a lasting record of the brightness of stars. Photoelectric photometry uses photomultipliers to amplify and measure the intensity of the light from a star very precisely. It is particularly useful for observing stars that have a small range of variation, from thousandths to tenths of a magnitude. CCDs are semiconductor imaging devices consisting of a two-dimensional array of photodetectors and a readout amplifier. They convert optical images into electrical signals, and are particularly useful in providing images of faint objects not seen through the eyepiece of a telescope. *See* ASTRONOMICAL IMAGING; CHARGE-COUPLED DEVICES; PHOTOMETRY; PHOTOMULTIPLIER.

Certification and nomenclature. Once a star is discovered to be changing in brightness, and its position in the sky and the kind (type) and the amount (range) of its variability are determined, it is ready to be certified and named as a variable star. The naming of a variable star is based on a system developed in the mid-1800s by F. W. A. Argelander. The first variable star discovered in a constellation is given the letter R, followed by the genitive form of the Latin name of the constellation. For example, R Leonis is the variable star R in the constellation of Leo. When other variables are discovered in that constellation, they receive the letters S, T, and so on to Z. Subsequent variables are named RR, RS to RZ, then SS to SZ, and so on to ZZ, after which follow AA, AB to AZ, and so on to QZ. (The letter J is never used in this system.) The next variable after QZ, the 334th variable in a constellation, is named V335. (For example, the variable V335 Sagittarii is the 335th variable in Sagittarius.) Stars that have been assigned Greek letters, such as o (omicron) Ceti, or small roman letters, such as g Herculis, prior to the start of this system continue to keep those names.

Until World War II, the certification and naming of a variable was done in Germany. The International Astronomical Union continued the certification and naming until 1952. This task has been carried out by the compilers of the *General Catalogue of Variable Stars* in Russia ever since. The fourth edition of this catalog and name lists that supplement it comprise 38,622 known variables designated through 2003. An additional 14,810 stars listed in the *New Catalogue of Suspected Variables* (1982) are in need of confirmation of their variability.

Not all constellations have the same number of variables, because the distribution of variables is not uniform in the Milky Way Galaxy. While the concentration is very high in the disk, particularly toward the center of the Galaxy (as in the case of the constellation Sagittarius, with 4374 variable stars listed),

the corona of the Milky Way Galaxy has the lowest number (as in the case of the constellation Caelum, which has only nine variables listed). *See* MILKY WAY GALAXY.

With the advent of CCDs, thousands of new faint variable stars are being discovered, and additional thousands of stars which vary only by millimagnitudes are being discovered as well with sensitive photoelectric photometers aboard space satellites. Yet, there may still be thousands of variable stars undiscovered, too faint, with too small variability, or too distant to be detected with present instrumention. There are several ongoing surveys of the sky, such as the All-Sky Automated Survey (ASAS), that are finding new variables every day. A popular activity for amateur astronomers is data mining the new surveys to search for unique objects for further study. *See* DATA MINING.

Classification. The light curve and the spectrum of a variable star form the basis for determining its classification. A light curve (**Fig. 1**) is the graph of the observed brightness (usually in magnitudes) of the star plotted against time (usually in Julian days). The important parameters are the shape of the light curve, the brightness range, and the regularity and the period of the brightness variation. *See* LIGHT CURVES.

The spectrum of a star helps to measure its temperature, luminosity, chemical composition, age (population), and mass. The spectra of variable stars normally range from the O or B (very hot/blue), to the M, C, and R (cool/red) spectral classes. Brown dwarfs (L and T spectral classes) also exhibit variability. Some lines in the spectra help to identify the nature of the variable. For example, metallic oxides and carbon lines, accompanied by bright lines of hydrogen, are usually seen in cool, red, long-period variables. *See* ASTRONOMICAL SPECTROSCOPY; BROWN DWARF; SPECTRAL TYPE.

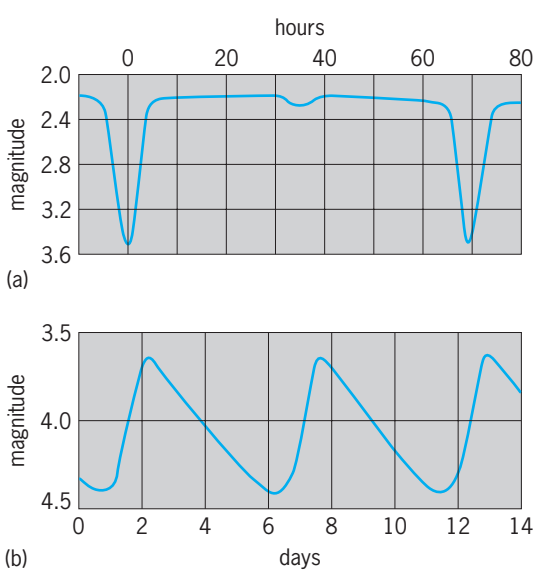

Fig. 1. Comparison of eclipsing variable stars and intrinsic variable stars. (*a*) An eclipsing variable star (Algol), with a period of 2.87 days. (*b*) A typical intrinsic variable star (Delta Cephei), with a period of 5.37 days.

Variable stars can be divided into two major types: extrinsic and intrinsic variables (Fig. 1).

Extrinsic variables. The extrinsic variables are those stars in which the variability in brightness occurs because of the occultation of one star by another object, or by the rotation of a spotted star.

In the eclipsing binary type of star, the change in the brightness of the system is the result of a geometric phenomenon. The primary eclipse occurs when the brighter star is eclipsed by the fainter, and the secondary eclipse occurs when the fainter star is eclipsed by the brighter. Information on the size, shape, and the distance to the stars from the Earth may be obtained from the eclipse phenomenon of these stars. The prototype of eclipsing binaries is Algol (β Persei), a bright star, observable with an unaided eye, that changes its brightness from 2.1 to 3.4 magnitudes as it is eclipsed every 69 hours (Fig. 1*a*). *See* BINARY STAR; ECLIPSING VARIABLE STARS.

A second type of eclipsing object is the extrasolar planetary transit. Here, a planet transits across the disk of its parent star, causing a dip in brightness similar to an eclipsing binary star. The dip is usually very small, as planets are much smaller than the star they orbit, but is detectable by careful ground-based photometry. *See* EXTRASOLAR PLANETS; TRANSIT (ASTRONOMY).

In a rotating variable, the variability is caused by rapid rotation of a star, often a member of a binary system, that has dark or bright spots on its surface, similar to sunspots. *See* BINARY STAR; ECLIPSING VARIABLE STARS.

Intrinsic variables. Intrinsic variables are those stars in which the variability of brightness occurs because of physical change in or on the star itself. These stars are divided into two classes: pulsating and eruptive (see **table**).

Pulsating variables. Periodic pulsation (contraction and expansion) of the star and its outer layer results in the variation in brightness, as well as variations in the star's temperature, spectrum, and radius, for pulsating variables, Cepheids are rare, highly luminous (supergiant), yellow pulsating stars. They vary with periods from 1 to 100 days, and have a range of variation from 0.1 to 2 magnitudes. The prototype of the class is δ Cephei, a star observable with an unaided eye which varies in brightness from magnitude 3.7 to 4.4 every 5.4 days (Fig. 1*b*). Polaris, the North Star, is also a cepheid variable with a smaller range of variation from magnitude 2.5 to 2.6 every 4 days. *See* POLARIS.

Cepheids show an important correlation between their period of variation and their relative brightness, in that those stars with longer periods are also brighter. This period-luminosity relationship was first noticed in 1912 by H. Leavitt while studying the cepheid variables in the Small Magellanic Cloud (**Fig. 2**). H. Shapley extended this relationship to cepheids in the Galaxy, and converted the relationship to period and intrinsic luminosity of these stars. In 1950, W. Baade determined the zero point of this

Representative types of intrinsic variables

Type	Range of period, days	Amplitude, magnitude	Spectra and luminosity	Example	Number known
Pulsating					
RR Lyrae	<1	<2	A to F (blue), giants	RR Lyrae	7382
Cepheids (types I and II)	1–70	0.1–2	F to G (yellow), supergiants	δ Cephei	1028
Long-period					
Mira	80–1000	2.5–6+	M (red), giants	o Ceti (Mira)	7390
Semiregular	30–1000	1–2	M (red), giants and supergiants	Z Ursae Majoris	5479
RV Tauri	30–150	Up to 3	G to K (yellow and red), supergiants	RV Tauri	134
Irregular	Irregular	Up to several magnitudes	All types		5884
Others					707
Eruptive					
Supernovae	?	15 or more		CM Tauri (Crab Nebula)	7
Novae	Centuries?	7–16	O to A, subdwarfs	GK Persei	388
Recurrent novae	20 years?	7–9		RS Ophiuchi	8
Dwarf novae	10–600	2–6	A to F, subdwarfs	SS Cygni, Z Camelopardalis	420
Flare	?	Up to 6	M, dwarfs	UV Ceti	1780
Nebular	Rapid and erratic	Up to a few magnitudes	B to M, main sequence and subgiants	T Tauri	1698
R Coronae Borealis		1–9	F to K, supergiants	R Coronae Borealis	42
Others					380

relationship, with his work on these stars in the Andromeda Galaxy. *See* ANDROMEDA GALAXY; MAGELLANIC CLOUDS.

Thus, by observing the apparent brightness change and determining the period of this variation, use of the period-luminosity relationship leads to the absolute magnitude of the star. Once the absolute magnitude of a star is known, the distance to it can be determined from the equation below,

$$M = m + 5 - 5 \log D$$

where M is the absolute magnitude, m is the apparent magnitude, and D is the distance in parsecs. Due to this important correlation, cepheids have been used to measure distances to nearby galaxies and in the determination of the distance scale of the universe.

A similar group of variable stars, W Virginis stars, are found in globular star clusters and the corona of the Milky Way Galaxy. These stars are bluer (and thus hotter) and older (population II) than the cepheids. They have periods from 10 to 30 days, and obey a similar period–luminosity relationship. They are sometimes called type II cepheids to distinguish them from the classical type I cepheids. *See* CEPHEIDS.

Pulsating stars. RR Lyrae are the second most common type of variable in the Galaxy. They have periods of less than 1 day, and have a small range of variation, from 0.5 to 1.5 magnitudes. They are particularly numerous in globular clusters, and thus are sometimes referred to as cluster variables. Their spectral classes range from A to F. The prototype of the class is RR Lyrae, which within 13.5 h varies from magnitude 7.0 to 8.1 and back. All RR Lyrae variables have the same intrinsic luminosity, of magnitude 0.5. Therefore these variables do not follow the period-luminosity relationship. Thus, if an RR Lyrae star can be identified in a star cluster and its apparent magnitude determined, the above equation and the star's known absolute magnitude can be used to obtain the distance to the cluster.

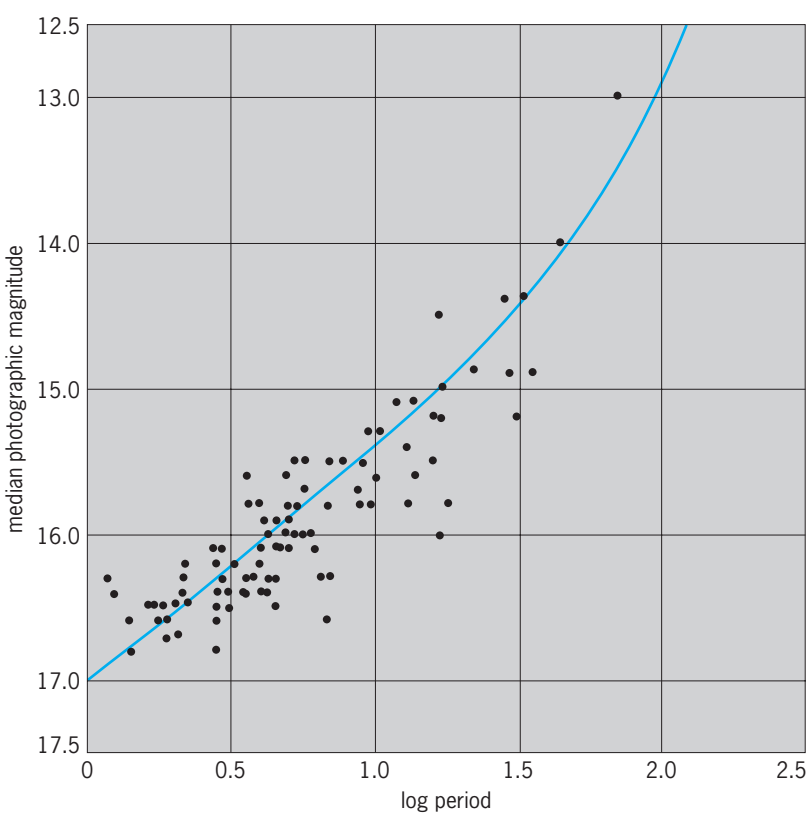

Fig. 2. Leavitt's period-luminosity curve for cepheid variables in Small Magellanic Cloud.

Long-period pulsating variable stars are the most abundant of all variables in the Milky Way Galaxy. They are red, cool, giant or supergiant stars with spectral class M or R, S, or C carbon types. Long-period variables are old stars which have evolved from the main sequence and are in the late stages of their evolution. Their size (diameter) is several hundred times that of the Sun; their mass, however, is about the same as the Sun. Their mass is concentrated in the core, which is surrounded by an extended gaseous atmosphere. *See* STELLAR EVOLUTION.

The light variability of long-period variables is due to the pulsation (contraction and expansion) of the star and its radiating surface. The prototype of long-period variables is o Ceti (Mira), which has given its name to those long-period variables that have a range of variation of 2.5 magnitudes and more. Mira variables have periods ranging from 100 to 1000 days. Although the change in brightness is periodic, the brightness of individual cycles may not be the same, some cycles being much brighter or fainter than others. Generally the rise to maximum brightness is faster than the decline to minimum brightness. *See* MIRA.

The spectra of long-period variables are abundant with lines of metallic oxides such as titanium oxide (TiO) and bright emission lines of hydrogen, seen particularly near maximum brightness. These bright lines are thought to be generated by propagating shock waves in the atmosphere of these stars.

Red giant and supergiant stars with less regularity of variation, shorter periods, and smaller ranges of variation, less than 2.5 magnitudes, are called semiregular variables. Other red variable stars that do not exhibit any regularity in their brightness change are called irregular stars.

Rare, very luminous, yellow supergiant stars that generally show alternating shallow and deep fadings are called RV Tauri stars. These pulsating variables vary by 2 to 3 magnitudes within 30 to 150 days.

Eruptive variables. Eruptive variables are those stars that have one or more eruptions—the ejection of matter into space—in their lifetime. These variables are sometimes referred to as cataclysmic variables because of the cataclysmic explosions they undergo. Cataclysmic variables are very close binary systems made up of an evolved, hot, dense white dwarf and a less evolved cool star which transfers mass onto the white dwarf accretion disk. *See* WHITE DWARF STAR.

There are many types of eruptive variables, ranging from supernovae and novae, in which spectacular explosions take place that brighten the system by many magnitudes, to nebular variables, which are young stars in the early stages of stellar formation, to R Coronae Borealis variables, which have sudden unpredictable decreases in their brightness. Only a few of the eruptive variable types will be discussed; a more complete listing of types is given in the table.

The most spectacular type of stellar explosion is a supernova, wherein a star's luminosity suddenly increases to tens of thousands of times the original brightness and the star outshines the total brightness of its galaxy. The gigantic explosion may be due to the gravitational collapse of a very hot and massive star that has exhausted the energy available from nuclear reactions; the collapse then creates enormous energy to blast outward the layers surrounding the stellar core. Alternatively, the explosion may occur in a close binary system in which one of the components is a massive white dwarf, an Earth-sized, very compact, old star, that explodes when it receives too much material from the other component. In the Milky Way Galaxy the most recent supernova observed was seen in 1604 in the constellation Serpens. However, Supernova 1987A, which exploded on February 23, 1987, in the Large Magellanic Cloud, a nearby companion galaxy to the Milky Way Galaxy, was visible to the unaided eye and was studied intensely. *See* SUPERNOVA.

The best known of the eruptive variables are stars that brighten by 7 to 16 magnitudes (less than 0.0010 of that of a supernova) within about a day. They stay at maximum brightness for a few days or weeks and then slowly fade. The word "nova," meaning new, was used for these stars. Actually a nova is not a new star, but an already existing star that due to the eruption has become very luminous and thus visible. *See* NOVA.

Recurrent novae are stars that about every 20 years or more have eruptions during which the system brightens by 7 to 9 magnitudes.

Dwarf novae have smaller-scale eruptions, in which the star brightens by 2 to 6 magnitudes within a day, stays bright 1 to 2 weeks, and then fades to the original brightness. The quasiperiodic eruptions may occur from ten to several hundred days apart. Dwarf novae are divided into two categories: U Geminorum and Z Camelopardalis stars. Z Camelopardalis stars are differentiated from U Geminorum variables by their periods of nonvariability midway between maximum and minimum brightness. These standstills can last days, months, or sometimes years.

R Coronae Borealis stars, instead of brightening through eruptions, irregularly decrease in brightness every 2 to 3 years by 1 to 9 magnitudes. These bright, highly luminous supergiant stars are rich in carbon and poor in hydrogen in composition. The decrease in brightness is caused by the veiling of the star by thick carbon clouds expelled to the star's atmosphere. *See* CATACLYSMIC VARIABLE; STAR.

Janet Akyüz Mattei; Arne Henden

Bibliography. G. North, *Observing Variable Stars, Novae and Supernovae*, 2004; C. Hellier, *Cataclysmic Variable Stars*, Springer, 2001; C. Hoffmeister, G. Richter, and W. Wenzel, *Variable Stars*, Springer, 1985; D. H. Levy, *Observing Variable Stars*, 2d ed., Cambridge University Press, 1998; J. R. Percy (ed.), *The Study of Variable Stars Using Small Telescopes*, Cambridge University Press, 1986; H. A. Smith, *RR Lyrae Stars*, Cambridge University Press, 1995; C. Sterken and C. Jaschek (eds.), *Light Curves of Variable Stars*, 2d ed., Cambridge University Press, 2005; B. Warner, *Cataclysmic Variable Stars*, Cambridge University Press, 1995.

Variational methods (physics)

Methods based on the principle that, among all possible configurations or histories of a physical system, the system realizes the one that minimizes some specified quantity. Variational methods are used in physics both for theory construction and for calculational purposes.

General principles. The earliest use of a variational principle for physics is Fermat's principle in optics, which states that when a light ray traverses a medium with nonuniform index of refraction its path is such as to minimize its travel time. The integral given in Eq. (1) expresses the time that the light takes to

$$T = \int_{\Gamma} \frac{ds\ n(s)}{c} \tag{1}$$

travel from point A to point B along a path Γ, where c is the speed of light in vacuum and n is the given index of refraction. An application of the calculus of variations to this integral makes it possible to determine the particular path Γ_0 for which T is a minimum. This problem is mathematically identical to the variational principle that determines a geodesic, the path of shortest distance, in a given geometry. In that form, the same principle determines the world lines of all objects in the general theory of relativity. *See* CALCULUS OF VARIATIONS; RELATIVITY; RIEMANNIAN GEOMETRY.

Similarly, in mechanics, Hamilton's principle for the action is defined for any system of point particles by Eq. (2). Here $L = K - V$ is the lagrangian

$$I = \int_{t_1}^{t_2} dt\ L \tag{2}$$

of the system, where K and V are its kinetic and potential energies, respectively, expressed in terms of the generalized coordinates and velocities of the particles. The integral extends over an arbitrarily prescribed path Γ in configuration space as a functions of the time t. Hamilton's principle asserts that the trajectories of all the particles are determined by the requirement that Γ be such that, for given initial and final times t_1 and t_2, the action I is a minimum; for this reason it is also called the principle of least action. If the calculus of variations is applied to implement this principle, the corresponding Euler-Lagrange equations are obtained—partial differential equations for the lagrangian, which express necessary and sufficient conditions for the trajectories to be such that the action is a minimum. These equations are the lagrangian equations of motion, in other words, Newton's equations of motion in lagrangian form. *See* ACTION; HAMILTON'S PRINCIPLE; LAGRANGE'S EQUATIONS; LAGRANGIAN FUNCTION; LEAST-ACTION PRINCIPLE; NEWTON'S LAWS OF MOTION.

The principle of least action has been generalized to infinitely many degrees of freedom, that is, fields. A Lagrange density function $\mathscr{L}$ is then defined, which is a function of the fields and their time derivatives at any given point in space and time. The action is given by Eq. (3), in which the x-integral extends over all

$$I = \int_{t_1}^{t_2} dt \int dx\ \mathscr{L} \tag{3}$$

three-dimensional space. The Euler-Lagrange equations of the principle of least action in this case are the field equations.

It is possible, in this way, to generate, for example, the Maxwell equations from the Lagrange density given in Eq. (4), where **E** and **H** are the electric

$$\mathscr{L} = \frac{1}{8\pi}(\mathbf{E}^2 - \mathbf{H}^2) + \frac{1}{c}\ \mathbf{j} \cdot \mathbf{A} - \rho\varphi \tag{4}$$

and magnetic fields respectively, **j** is the current density, **A** is the vector potential, ρ is the charge density, and φ is the scalar potential. *See* MAXWELL'S EQUATIONS; POTENTIALS.

Theory construction. For any field theory, only the Lagrange density needs to be given; the field equations are then derivable as the corresponding Euler-Lagrange equations. A similar technique makes it possible to derive the Schrödinger equation and the Dirac equation in quantum mechanics from specific Lagrange density functions. *See* NONRELATIVISTIC QUANTUM THEORY; QUANTUM MECHANICS; RELATIVISTIC QUANTUM THEORY.

This method has great procedural advantages. For example, it facilitates a check of whether the theory satisfies certain invariance principles (such as relativistic invariance or rotational invariance) by simply ascertaining whether the Lagrange density satisfies them. The corresponding conservation laws can also be derived directly from the lagrangian. In all modern quantum-field theories, it is customary to postulate the specific form of the theory by simply specifying the Lagrange density function; the variational principle then automatically leads to the field equations and implies all the conservation laws. *See* CONSERVATION LAWS (PHYSICS); QUANTUM FIELD THEORY; SYMMETRY LAWS (PHYSICS).

Calculational uses. The variational method plays an important role in quantum-mechanical calculations. For the computation of needed quantities in terms of functions that result from the solution of differential equations, it is always of great advantage to use formulas that have the special form required to make them stationary with respect to small variations of the input functions in the vicinity of the unknown, exact solutions.

For instance, the energy of a quantum-mechanical system may be expressed in terms of the expectation value of its hamiltonian, H, given by Eq. (5), with the

$$E = \frac{(\psi, H\psi)}{(\psi, \psi)} \tag{5}$$

wave function ψ set equal to the eigenfunction ψ_E of H at that energy. The expression in Eq. (5) has the property that it is stationary with respect to small variations of the function ψ in the neighborhood of ψ_E: Its first variation from $\psi = \psi_E$ vanishes. This implies that if, in the calculation, instead of the exactly correct eigenfunction ψ_E an approximation whose

norm differs from that of the correct eigenfunction by a small amount ϵ is used, then the calculated value of E will differ from the exact value by the much smaller amount ϵ^2. For the ground state of the system, that is, the state of lowest energy, this stationary value is, in fact, an absolute minimum. As a result, the number E calculated by Eq. (5) with any arbitrary function ψ is always an upper bound for the ground-state energy, and if ψ is a reasonable approximation, that bound will be close to the correct value. To find a similarly good approximation and bound for excited states, it is necessary to choose trial functions that are orthogonal to the ground state.

In practice, almost all computations of energy levels of atomic or nuclear systems are performed by using this variational method. A trial function depending on a number of parameters is adopted, and the parameters are chosen so that the calculated value of E is stationary; that is, its first derivatives with respect to the parameters all vanish. The value of E so obtained is the best approximation possible with the given set of trial functions. (This procedure is called the Ritz variational method.)

Analogous variational methods also exist for the calculation of other quantum-mechanical quantities, as well as for classical quantities that are eigenvalues of differential equations arising in dynamics. For example, the vibrational frequencies of mechanical systems can be calculated by such a method, as was done first by Lord Rayleigh. In quantum mechanics, there are several expressions for the scattering phase shift of a particle in a central potential that are stationary with respect to small variations of the scattering wave function near its exact form. R. Feynman used the variational method in the context of his path integrals to calculate the lowest energy of an electron in a polarizable crystal—a polaron. *See* FEYNMAN INTEGRAL; MINIMAL PRINCIPLES; POLARON.

Roger G. Newton

Bibliography. P. Blanchard, *Variational Methods in Mathematical Physics: A Unified Approach*, 1992; R. P. Feynman, Slow electrons in a polar crystal, *Phys. Rev.*, 97:660–665, 1955; P. Griffiths, *Exterior Differential Systems and the Calculus of Variations*, 1983; J. T. Oden and J. N. Reddy, *Variational Methods in Theoretical Mechanics*, 1983.

Varistor

Any two-terminal solid-state device in which the electric current I increases considerably faster than the voltage V. This nonlinear effect may occur over all, or only part, of the current-voltage characteristic. It is generally specified as $I \propto V^n$, where n is a number ranging from 3 to 35 depending on the type of varistor. The main use of varistors is to protect electrical and electronic equipment against high-voltage surges by shunting them to ground. *See* ELECTRIC PROTECTIVE DEVICES.

One type of varistor comprises a sintered compact of silicon carbide particles with electrical terminals at each end. It has symmetrical characteristics (the same for either polarity of voltage) with n ranging from 3 to 7. These devices are capable of application to very high power levels, for example, lightning arresters. *See* LIGHTNING AND SURGE PROTECTION.

Another symmetrical device, the metal-oxide varistor, is made of a ceramiclike material comprising zinc oxide grains and a complex amorphous intergranular material. It has a high resistance (about 10^9 ohms) at low voltage due to the high resistance of the intergranular phase, which becomes nonlinearly conducting in its control range (100–1000 V) with $n > 25$.

Semiconductor rectifiers, of either the *pn*-junction or Schottky barrier (hot carrier) types, are commonly utilized for varistors. A single rectifier has a nonsymmetrical characteristic which makes it useful as a low-voltage varistor when biased in the low-resistance (forward) polarity, and as a high-voltage varistor when biased in the high-resistance (reverse) polarity. Symmetrical rectifier varistors are made by utilizing two rectifiers connected with opposing polarity, in parallel (**Fig. 1***a*) for low-voltage operation and in series (Fig. 1*b*) for high-voltage use. For the high-voltage semiconductor varistor, n is approximately 35 in its control range, which can be designed to be anywhere from a few volts to several hundred. *See* SEMICONDUCTOR RECTIFIER.

Silicon rectifiers are fabricated mostly by solid-state diffusion techniques similar to those utilized for transistors and integrated circuits. A thin plate of *n*-type silicon is exposed to a gaseous atmosphere containing a dilute *p*-type impurity at a high temperature (about 2200°F or 1200°C). Some atoms of the *p*-type impurity diffuse into the silicon surface by displacing silicon atoms, forming a *pn* junction within about 0.001 in. (0.025 mm) of the surface.

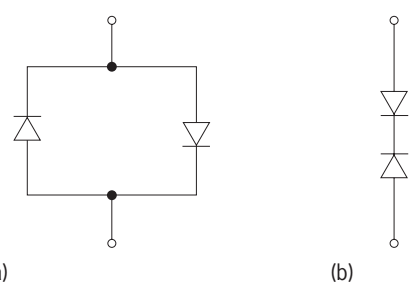

Fig. 1. Symmetrical rectifier varistors. (*a*) Low voltage. (*b*) High voltage.

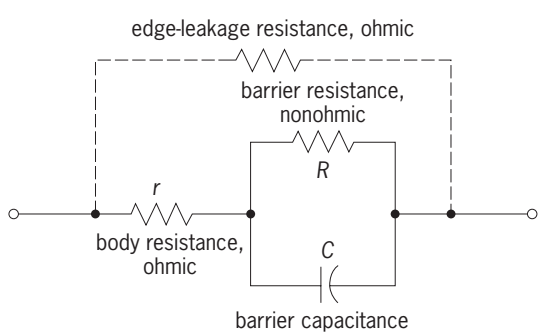

Fig. 2. Equivalent circuit of a single-rectifier varistor.

Metal electrodes are applied to *p*- and *n*-type regions by plating or evaporation. *See* INTEGRATED CIRCUITS; JUNCTION DIODE; TRANSISTOR.

The equivalent circuit of a *pn*-junction varistor is shown in **Fig. 2**. The capacitance is high, limiting high-frequency performance. I. A. Lesk

Bibliography. E. Bleuler and R. O. Haxby (eds.), *Methods of Experimental Physics*, vol. 2, pt. A, 1975; J. D. Harnden, Jr., et al., Metal-oxide varistor: A new way to suppress transients, *Electronics*, 45(21):91–95, October 9, 1972; L. M. Levinson (ed.), *Ceramic Transactions*, vol. 3: *Advances in Varistor Technology*, 1989; J. R. O'Connor and J. Smiltens (eds.), *Silicon Carbide*, 1960.

Varnish

A transparent surface coating which is applied as a liquid and then changes to a hard solid. Varnishes are solutions of resinous materials in a solvent, and dry by the evaporation of the solvent or by a chemical reaction, either with oxygen from the air or by some other means, including absorption of atmospheric moisture.

Spirit varnishes are those in which the evaporation of solvent is the only drying process; the solvent is usually alcohol, although the term is used for similar coatings made with other solvents. Shellac varnish, made by dissolving shellac in alcohol, is the most common of this type. Oleoresinous varnishes are made by treating a drying oil with a resin, usually with heat, and dissolving the reaction product in a solvent, usually a petroleum fraction; drying results from the evaporation of the solvent, followed by polymerization of the drying oil portion, a reaction which is accelerated by metallic driers added to the varnish. For a discussion of the mechanism of this drying action *see* DRIER (PAINT); SHELLAC.

Varnish coatings on wood are used to protect against abrasion, staining, and weather and to reduce the penetration of water and other materials without obscuring the grain or changing the color materially. Varnishes are used on masonry to reduce the penetration of moisture and the damage from freezing. Paper is coated with varnish to resist moisture and keep printing from being damaged. *See* PRINTING.

Shellac varnishes are used where fast-drying, slightly colored coatings of low cost are needed and where resistance to water, solvents, and weather is not required. Oleoresinous varnishes have greater resistance, but their color is usually darker. Spar varnishes are oleoresinous varnishes, often made from tung oil and a phenolic resin, which have good weather resistance; they were originally used for marine finishes.

Polyurethane varnishes exhibit superior abrasion resistance compared with other types. High gloss and gloss retention are other attractive features.

Varnish stains contain a dye dissolved in the solvent; they are used to change the color of a wood without obscuring the grain. *See* DYE; SOLVENT; SURFACE COATING. C. R. Martinson; C. N. Sisler

Varnish tree

The plant *Toxicodendron vernicifluum* (previously known as *Rhus vernicifera*), also called lacquer tree, a member of the sumac family (Anacardiaceae). It is a native of China, but has long been cultivated in Japan. When the bark is cut, it exudes a milky juice which darkens and thickens on exposure. This is the lacquer long used in China and Japan. When properly applied, the thin transparent film becomes a varnish of extreme hardness. Lacquer is a remarkably protective coating as it is not altered by acids, alkalies, alcohol, or heat up to 160°F (71°C). Nut galls, iron in solution, and gold or other metals are mixed with the lacquer before drying to make the various kinds of lacquers. The process of lacquering is technical and tedious, sometimes requiring 300–400 coats and several years to complete the finish of one item. *See* LACQUER; SAPINDALES. Perry D. Strausbaugh; Earl L. Core

Varve

Any of a variety of distinct sediment laminations or beds deposited within the span of a single year. They are formed commonly in saline or fresh-water lakes, but examples from marine environments are known as well. Usually, varves occur in repetitive series and thus comprise vertical sequences of annual cyclic deposits. Varves range in thickness from less than a millimeter (0.04 in.) to over a meter (3 ft), but typically are a few millimeters or centimeters thick.

The classic varves are found in glacial lake sediments formed during the Pleistocene ice ages. These glacial varves occur typically as couplets of light-colored silt or sand and dark clay. The relatively coarse silt-sand layers are formed during the warm summer months when meltwater inflows and sediment yields to the lake are large. During winter when meltwater inflow is greatly reduced or stopped, the fine-grained clay settles slowly to the lake bottom to deposit the fine, dark winter layer of the varve. Similar varved sediments are found commonly in modern glacier-fed lakes that undergo large seasonal variations in inflow (see **illus.**). *See* GLACIATED TERRAIN; PLEISTOCENE.

Varves can be matched sometimes between sample localities and used as tools for correlation as well as for chronological reconstructions. Varve chronologies have been erected in late Quaternary deposits in numerous areas of northern Europe and North America formerly occupied by continental ice sheets. Such varve correlations and chronologies have been used to reconstruct the recession history of the last ice sheet as well as to establish an absolute time scale back to 10,000 years. In addition to their use in dating sedimentary deposits, varves have been used to investigate sedimentation rates, cyclic deposition, climate variations, glacial histories, and as standards of comparison for other dating techniques. Varves have been identified occasionally in ancient sedimentary rocks. *See* QUATERNARY.

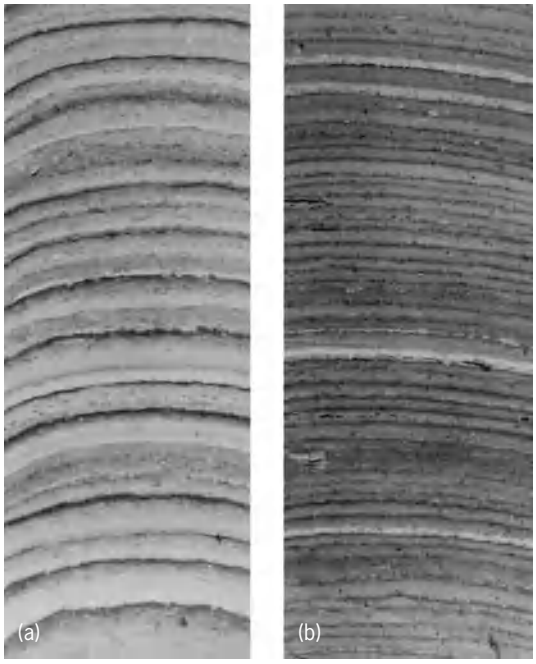

Varved sediment in two lake-bottom cores from modern glacier-fed lake (Hector lake, Alberta, Canada). Varves collected closer to the river inlet (left) are thicker and represent higher annual sedimentation rates. The length of each sequence shown is 5 cm (2 in.).

Chemically precipitated varves, typically composed of calcium carbonate, have been recognized in numerous modern lakes and ancient lacustrine deposits. These arise from a variety of mechanisms, including seasonal fluctuations in temperature, salinity, and pH; variable water motions; and biological productivity cycles, especially seasonal diatom blooms which affect pH and, consequently, calcium carbonate solubility. Such varves commonly form in lakes that are chemically stratified (meromictic). *See* GLACIAL EPOCH; MARINE SEDIMENTS; MEROMICTIC LAKE; SEDIMENTARY ROCKS; SEDIMENTOLOGY; SHALE.

Norman D. Smith

Bibliography. R. Y. Anderson et al., Meromictic lakes and varved lake sediments in North America, *USGS Bull.*, 1607:1–19, 1985; I. Banerjee, Sedimentology of Pleistocene glacial varves in Ontario, Canada, *Geol. Surv. Can. Bull.*, 226A:1–44, 1974; E. M. Leonard, Varve studies at Hector Lake, Alberta, Canada, and the relationship between glacial activity and sedimentation, *Quaternary Res.*, 25:199–214, 1986; S. D. Ludlam, Fayetteville Green Lake, New York, III: The laminated sediments, *Limnol. Oceanog.*, 14:848–857, 1969; N. D. Smith, Sedimentation processes and patterns in a glacier-fed lake with low sediment input, *Can. J. Earth Sci.*, 15:741–756, 1978.

Vascular disorders

Those disorders that involve the arteries, veins, and lymphatics.

Arteries. The arteries distribute blood to the tissues. Those arteries with proper names, such as the aorta or the renal artery, are referred to as large and medium-sized arteries. They branch into vessels of smaller size, finally terminating in the smallest arteries, the arterioles. These microscopic vessels then connect with the capillaries, where the exchange of oxygen and carbon dioxide takes place. Diseases of the large and medium-sized arteries are the major cause of morbidity and mortality in the Western world.

Arteriosclerotic vascular disease. This disorder, also known as arteriosclerosis, affects large and medium arteries and is particularly common in the arteries supplying the heart, those supplying the brain, and those supplying the lower extremities. Progressive narrowing and finally total occlusion of these arteries lead to the development of angina pectoris, myocardial infarction, stroke, and vascular insufficiency in the limbs. For unknown reasons, the process rarely involves the arteries of the arm. In the legs, it leads to pain in the muscles with walking (intermittent claudication) and, in far advanced cases, gangrene. *See* GANGRENE; INFARCTION.

Although the cause of the disease is not known, certain factors appear to predispose to its development. These include cigarette smoking, hypertension, elevated serum cholesterol, and diabetes mellitus. There can also be a genetic predisposition to the disorder. There is increasing evidence that control of the above risk factors can prevent the onset of the disease or slow its progression. *See* DIABETES; HYPERTENSION.

The disorder appears to develop in stages, with early lesions sometimes seen even in young adults. The primary lesion is the plaque, which is an accumulation of fat, smooth muscle, and collagen that develops at branch points and bifurcations in the arterial system. As the plaque gets larger, it begins to impinge on the lumen of the artery, leading to a reduction in blood flow. In later stages, the plaque undergoes degeneration, with loss of the surface covering, the appearance of calcium, necrosis, and intraplaque hemorrhage. Portions of the plaque may break off, producing emboli that may block vital areas of the tissues supplied by the artery. In its final stages, clot formation (thrombosis) at the site of the plaque can occur, which completely shuts off the blood supply to an area of tissue. The extent to which the tissue survives then depends on the capability of accessory vessels that bypass the area of occlusion to take over the function of the occluded artery. *See* EMBOLISM; THROMBOSIS.

At some sites the arteriosclerosis may be accompanied by an aneurysm, a ballooning of the weakened vessel wall. Aneurysms most commonly occur in the abdominal aorta, where rupture is not uncommon. When they develop in the femoral and popliteal arteries of the lower limb, rupture rarely occurs but thrombosis is common. The forming thrombus within the aneurysm can break away and occlude the arteries of the lower limb and foot. *See* ANEURYSM.

Patients with diabetes mellitus are prone to develop arteriosclerosis in the coronary arteries, the

carotid artery in the neck, and the arteries supplying the lower limbs. For diabetics, the prevalence of arteriosclerosis is at least 20 times as high for the leg arteries and 8 times as high for the carotid artery compared with the general population. This high rate of occurrence in arteries supplying the limbs is the main reason for the very high rate of limb loss in persons with diabetes mellitus.

Although there is evidence that arteriosclerosis may be prevented or slowed by control of those risk factors associated with its occurrence, narrowing and occlusion of the large and medium arteries continues to be the major health problem in the Western world. To treat arteriosclerotic narrowings of vessels, surgery may be done to remove or bypass the obstruction, or a balloon may be placed into the artery and guided to the site of involvement, where it is inflated in order to compress the plaque. This procedure is known as balloon angioplasty. The goal is to restore normal blood flow to the involved tissues. *See* ARTERIOSCLEROSIS.

Inflammation. Much less common are diseases that lead to arteritis, an inflammation of the large or medium arteries, with subsequent occlusion or rupture of the artery. The causes include bacterial infection, syphilis, allergic disorders, and hypersensitivity states.

Lodging of bacteria in the wall of an artery can lead to development of a mycotic aneurysm. The bacteria originate from a remote site in the body where they gain access to the circulation. It is not known why they localize within the artery wall, but when such localization occurs, the artery can become weakened, leading to development of an aneurysm that may rupture or be the site of further dissemination of the bacteria throughout the body. Mycotic aneurysms can develop, for example, in patients with an infection involving one of the valves of the heart (bacterial endocarditis).

Before penicillin was introduced, syphilis was a common cause of infection of the arterial wall. It led to destruction of the elastin and smooth muscle of large arteries such as the aorta. The wall of the aorta was ultimately weakened to the point where an aneurysm would occur. This disease is rare in the United States today presumably because of the effectiveness of penicillin in killing the causative organism. *See* SYPHILIS.

The hypersensitivity disorders that can lead to arteritis include periarteritis nodosa, lupus erythematosus, rheumatic arteritis, temporal arteritis, and giant-cell arteritis. When these affect the arterial wall, thrombosis of the involved artery is the most common outcome. The thrombosis occurs gradually and leads to symptoms reflecting the chronic reduction in blood flow to the organ. The clinical manifestations depend on the artery involved and the tissues supplied. *See* CONNECTIVE TISSUE DISEASE; HYPERSENSITIVITY.

Small arteries. Several generalized disorders can involve the small arteries. These include scleroderma, periarteritis nodosa, lupus erythematosus, rheumatoid arthritis, and dermatomyositis. The involvement can produce areas of cell death (necrosis) and ulceration of the skin, particularly of the limbs. This appears to occur because of the occlusion of the involved small arteries, with consequent loss of blood supply. Such involvement often is preceded by a type of hypersensitivity to cold, a condition known as Raynaud's syndrome. Such hypersensitivity can also occur when the small arteries and arterioles go into spasm when the hands or feet are exposed to cold. When this occurs, the digits, particularly the fingers, turn white, then blue, and finally red. This has been referred to as the triphasic color response. When it is not associated with an underlying disease such as scleroderma, it poses no risk. However, if it is associated with an underlying disorder, the small arteries of the hands and feet may become occluded, and digit ulcers or gangrene can develop.

Veins. The most common disorder of the venous circulation is varicose veins. It develops because of the loss of function of the valves in the superficial veins of the limbs. The two most commonly involved veins are the greater and lesser saphenous. In this condition, the veins are enlarged and often tortuous. If the deep veins are not involved, the varicose veins are termed primary. Varicose veins tend to cluster in families and are more common in women than men. The condition may be worsened by pregnancy due to both hormonal effects and an increase in the venous pressure in the iliac veins due to the enlarging uterus. Primary varicose veins may be cosmetically unpleasing, but they rarely cause serious symptoms in the involved limbs.

A more serious problem develops when the varicose veins are associated with destruction of valves in the deep veins, secondary to deep-vein thrombosis. Blood then is forced to flow in a retrograde direction and through the perforating veins to join the superficial veins. As a result, the superficial veins become enlarged and tortuous. Because of the combined involvement of the deep and superficial veins, leg swelling and pain are common. As time passes, brownish discoloration of the skin of the lower leg can occur. In addition, ulceration is common and can be difficult to treat. When varicose veins develop because of problems within the deep venous system, they are called secondary varicose veins.

When patients undergo major surgery or are confined to bed because of a serious medical illness, thrombosis of the veins of the leg can occur. This is commonly referred to as phlebitis but this term is not correct since inflammation is not a prominent part of the process. It is more appropriate to refer to this disorder as deep-vein thrombosis. The thrombosis occurs most commonly in the deep veins of the calf but can progress and extend proximally to involve the popliteal, femoral, and iliac veins. When this occurs, portions of the thrombi can dislodge, leading to the development of emboli in the lung. Emboli that arise from veins as large as the femoral or iliac may be quite large and can prove fatal.

Lymphatics. The lymphatics are responsible primarily for draining interstitial fluid from the tissues

in which they are found. This fluid, which is rich in protein, must be removed, or swelling (edema) will occur. When this drainage function is disturbed by obstruction of the lymphatics by tumors, parasites, or surgical excision, serious swelling can occur. This condition is called lymphedema. The pattern of swelling differs from that seen with obstruction of the venous system. For example, it will involve not only the lower leg but the foot and toes as well. In lymphedema, it is not uncommon for infection with a betahemolytic streptococcus to occur. When this happens, the patient develops shaking chills, fever, and pain in the limb. Prompt treatment with an antibiotic such as penicillin is necessary to avoid further destruction of the lymphatics and worsening of the edema. *See* EDEMA; LYMPHATIC SYSTEM.

Obstruction of the lymphatics can also be congenital or can appear unexpectedly at the time of the onset of the menses. The mechanisms are unknown, but the effects are the same as when a definite cause is found. Regardless of its cause, lymphedema is a permanent condition. Its control requires the use of compression stockings, which limit the amount of fluid that can accumulate in the legs during the day when the individual is upright.

Arteriovenous communications. Proper function of the cardiovascular system occurs when blood follows its normal course via the arteries to the capillaries and veins and then back to the right heart. If there are anomalous communications between the arteries and veins proximal to the capillaries, blood is shunted away from the capillary bed. Such shunting can be congenital or acquired.

The congenital form can occur at almost any site in the vascular system. When these communications are present, they may appear as prominent veins that are much larger and more numerous than normal. These are commonly referred to as hemangiomas. The extent to which they cause problems depends upon their location, their size, and numbers of feeding communications that are present. When they occur in the brain, the thin-walled and enlarged veins can rupture leading to intracranial hemorrhage, which may be fatal as the intracranial pressure rises often to very high levels. When they occur in the limbs, lungs, or abdomen, the problems are most often related to the amount of blood that is shunted and thus bypasses the capillary network. If large volumes of blood are shunted, there may be an extra load on the heart, leading to cardiac hypertrophy and ultimately to heart failure and death. Correction of these conditions is difficult or impossible if the malformations are extensive.

If a foreign body such as a bullet or knife passes between a major artery and vein, an artificial communication can develop that results in the shunting of a large volume of blood. When the communication is between very large vessels such as the abdominal aorta and inferior vena cava, the amount of blood shunted may be so large that heart failure occurs. These abnormal communications require surgical correction to restore normal blood flow. *See* CARDIOVASCULAR SYSTEM. D. Eugene Strandness, Jr.

Bibliography. W. R. Hazzard, R. Andres, and E. L. Bierman (eds.), *Principles of Geriatric Medicine*, 4th ed., 1998; E. Braunwald et al. (eds.), *Harrison's Principles of Internal Medicine*, 15th ed., 2001.

Vector (mathematics)

A mathematical element used to describe quantities having magnitude and direction. It can be represented by a directed line segment. Vectors are widely used in mathematical physics. *See* CALCULUS OF VECTORS. Marvin Yelles

Vector methods (physics)

Methods that make use of the behavior of physical quantities under coordinate transformations.

Definitions. From the point of view of physics, the most appropriate definition of a vector in three-dimensional space is a quantity that has three components which transform under rotations of the coordinate system like the coordinates of a point in space. What characterizes rotations is that the distance from the origin, $\sqrt{x_1^2+x_2^2+x_3^2}$ of all points $\mathbf{x}$ with cartesian coordinates x_i, $i = 1, 2, 3$, remains unchanged. Specifically, if the rotation takes the x_i to new coordinates x_i', given by Eqs. (1) [in which

$$x_i' = \sum_{j=1}^{3} a_{ij}x_j \qquad i = 1, 2, 3$$
$$\sum_{k=1}^{3} a_{ik}a_{jk} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases} \qquad \det\{a_{ij}\} = 1 \tag{1}$$

$\det\{a_{ij}\}$ is the determinant of the matrix $\{a_{ij}\}$], then the three quantities V_i, $i = 1, 2, 3$, form the components of a vector $\mathbf{V}$, if in the new coordinate system the transformed coordinates are given by Eq. (2).

$$V_i' = \sum_{j=1}^{3} a_{ij}V_j \qquad i = 1, 2, 3 \tag{2}$$

If the coordinate transformation of Eqs. (1) is such that $|a_{ij}| = 1$ for $i = j$ and $a_{ij} = 0$ for $i \neq j$, but such that $\det\{a_{ij}\} = -1$ rather than $+1$ as in Eqs. (1), then it describes a reflection, in which a right-handed coordinate system is replaced by a left-handed one. [Transformations that combine Eqs. (1) with a reflection are called improper rotations, whereas Eqs. (1) is called a proper rotation.] If, for such a transformation, Eq. (2) also holds, then $\mathbf{V}$ is called a polar vector, whereas if the components of $\mathbf{V}$ do not change sign, it is called an axial vector or pseudovector. The vector $\mathbf{V}$ can also be looked upon as a quantity with a direction, with the magnitude $\sqrt{V_1^2+V_2^2+V_3^2}$ pointing from the origin of the coordinate system to the point in space with the cartesian coordinates (V_1, V_2, V_3).

A quantity that remains invariant under a rotation of the coordinate system is called a scalar. The dot product of two vectors, given by Eq. (3), is a scalar,

and the vector product, given by Eq. (4), where ϵ_{ijk}

$$\mathbf{V} \cdot \mathbf{W} = \sum_{i=1}^{3} V_i W_i \tag{3}$$

$$(\mathbf{V} \times \mathbf{W})_i = \sum_{j,k=1}^{3} \epsilon_{ijk} V_j W_k \tag{4}$$

is antisymmetric under exchange of any two of its indices, and $\epsilon_{123} = 1$, is an axial vector if both **V** and **W** are polar. *See* CALCULUS OF VECTORS.

The importance of vectors and scalars in physics derives from the assumed isotropy of the universe, which implies that all general physical laws should have the same form in any two coordinate systems that differ only by a rotation. It is therefore useful to classify physical quantities according to their transformation properties under coordinate rotations. Examples of scalars include the mass of an object, its electric charge, its volume, its surface area, the energy of a system, and its temperature. Other quantities have a direction and thus are vectors, such as the force exerted on a body, its velocity, its acceleration, its angular momentum, and the electric and magnetic fields. Since the sum of two scalars is a scalar, and the sum of two vectors is a vector, it is important in the formulation of physical laws not to mix quantities that have different transformation properties under coordinate rotations. The sum of a vector and a scalar has no simple transformation properties; a law that equated a vector to a scalar would have different forms in different coordinate systems and would thus not be acceptable.

Examples. Newton's second law, Eq. (5), in which

$$\mathbf{F} = m\mathbf{a} \tag{5}$$

F is the force on a particle, m is its mass, and **a** is its acceleration, has the same form in any two coordinate systems that are rotated with respect to one another. This is because both **F** and **a** are vectors, and m is a scalar; therefore $\mathbf{F} - m\mathbf{a}$ is a vector, and if it vanishes in one coordinate system (that is, all three of its components vanish), then it also vanishes in any rotated system. *See* NEWTON'S LAWS OF MOTION.

If the dot product of Newton's second law is taken with the velocity **v**, the result, givenby Eq. (6),

$$\mathbf{v} \cdot \mathbf{F} = m\mathbf{v} \cdot \mathbf{a} \tag{6}$$

equates scalar quantities. The left-hand side of this equation is the rate at which work is being done, and the right-hand side is the time-derivative of $^1/_2 m\mathbf{v}^2$, which is the kinetic energy. Therefore, this equation expresses the conservation of energy in this case: the rate at which work is done on the particle equals the rate at which its kinetic energy increases. *See* CONSERVATION OF ENERGY; DIFFERENTIATION; ENERGY.

Similarly, if the vector product of Newton's second law with the position vector **x** of the particle is taken, the result equates two polar vectors. The left-hand side is the torque applied (about the origin of the coordinate system), and the right-hand side is the rate at which the angular momentum (with respect to the origin) increases. Thus the conservation of angular momentum for the particle is obtained. *See* ANGULAR MOMENTUM; RIGID-BODY DYNAMICS; TORQUE.

Maxwell's equations for the electric and magnetic fields, **E** and **B**, respectively, in vacuum, are Eqs. (7),

$$\begin{aligned} \nabla \times \mathbf{B} &= \frac{4\pi}{c}\mathbf{j} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} & \nabla \times \mathbf{E} &= -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \cdot \mathbf{E} &= 4\pi\rho & \nabla \cdot \mathbf{B} &= 0 \end{aligned} \tag{7}$$

in which **j** is the current density, ρ is the charge density, and c is the speed of light. These are vector equations, and hence are form-invariant with respect to coordinate rotations. Since the curl is a vector and the divergence is a scalar, the first two are vector equations and the second two are scalar. However, since ρ is a scalar, the third equation implies that **E** is a polar vector, and the second implies that **B** has to be an axial vector: its components do not change sign under a coordinate reflection. *See* MAXWELL'S EQUATIONS.

Generalizations. There are other physical quantities that transform like tensors under coordinate rotations. Indeed, axial vectors are really antisymmetric tensors of rank 2. Since an antisymmetric 3×3 matrix has exactly three independent entries, such a tensor in three dimensions has three components that transform under proper rotations like those of a vector. The stress tensor is an example of a physical quantity that transforms like a symmetric tensor of rank 2. *See* TENSOR ANALYSIS.

When measurements of an event in one laboratory are compared with those of the same event as performed in another laboratory moving uniformly with respect to the first, the Lorentz transformation of the special theory of relativity transforms simultaneously four quantities, the three spatial coordinates and the time. It is therefore useful to combine these four into a four-dimensional space called Minkowski space. The requirement of the special theory of relativity, that all physical laws should have the same form in any two laboratories in uniform motion with respect to one another, can consequently be most simply implemented by classifying physical quantities according to their transformation properties under Lorentz transformations, and demanding that laws should never equate quantities which transform differently. If that demand is met, the law will automatically be form-invariant with respect to such transformations. *See* LORENTZ TRANSFORMATIONS; RELATIVITY.

For that reason, the concept of a vector has been generalized to four dimensions: a four-vector is a set of four quantities that transform under Lorentz transformations just like the spatial coordinates (x_1, x_2, and x_3) and the time (t). There is, however, one important difference between the Minkowski space in which the relativistic four-vectors live and the three-dimensional coordinate space: the square of the distance that remains invariant under Lorentz transformations is $x_1^2 + x_2^2 + x_3^2 - c^2t^2$, which is required so that the speed of light c be the same for all observers. That makes the Minkowski space

noneuclidean. Similarly, the concept of a tensor is generalized to Minkowski space. *See* NONEUCLIDEAN GEOMETRY.

An example of a physical four-vector is the four-momentum of a particle, the first three components of which are its ordinary momentum (**p**), and the fourth, its energy (E) divided by c. Its invariant length is given by Eq. (8), which expresses the relativistic

$$\sqrt{E^2/c^2} - p \cdot p = mc \tag{8}$$

relation between the momentum of a particle and its energy. The form invariance of the Maxwell equations is most easily exhibited by combining the electric and magnetic fields into an antisymmetric field tensor of rank 2. The Maxwell equations then simply state that the divergence of this tensor is equal to the four-dimensional current density (which combines the three-dimensional current and the charge density), and that the divergence of its dual vanishes. *See* RELATIVISTIC ELECTRODYNAMICS.

In their most modern guise, vector methods make use of the coordinate-free concepts of exterior forms and the exterior differential calculus.

Roger G. Newton

Bibliography. R. L. Bryant, *Exterior Differential Systems*, 1991; D. A. Danielson, *Vectors and Tensors in Engineering and Physics*, 2d ed., 1997; F. A. Hinchey, *Vectors and Tensors for Engineers and Scientists*, 1976.

Vega

The fifth brightest of all stars and the third brightest in the northern sky. It will be the north polar star in about 12,000 years. In moving through the Milky Way Galaxy, the Sun is generally heading toward the position now occupied by Vega. At a distance of 7.8 parsecs (25.3 light-years, or 2.4×10^{14} km, or 1.49×10^{14} mi), Vega, or α Lyrae, is the prototypical star of spectral class A0V, indicating that it has an effective surface temperature of 9600 K (16,800°F) and derives its energy from the thermonuclear burning of hydrogen in a stable core region. Stars of this class have main-sequence lifetimes of about 5×10^8 years, 20 times shorter than the Sun. Vega's spectrum shows strong features due to hydrogen in the star's outer layers. The sharpness of these features implies that Vega is pointing its rotation axis nearly at the Sun. In comparison with the Sun, Vega is approximately 2.9 times larger in diameter, 2.5 times more massive, and 60 times more luminous. *See* PRECESSION OF EQUINOXES; SPECTRAL TYPE.

In most respects Vega is a typical A-type star made noteworthy by its proximity to the Sun. Two additional aspects of Vega are notable. First, unlike most stars of its type, Vega has no stellar companions gravitationally bound to it to form a binary star system. Second, Vega emits far more radiation at infrared wavelengths than would be expected. This radiation originates from a shell or disk of particles with a temperature of 100 K (−280°F) surrounding Vega out to a distance of 1.3×10^{10} km (8×10^9 mi), twice the radius of the solar system. The shell particles are at least 1 mm (0.4 in.) in size and are likely to be left over from Vega's formation. Near its edge, this dusty disk exhibits two prominent clumps of material opposite each other that may be the result of a gravitational resonance with a planet orbiting around Vega about midway through the disk. No such planet has yet been discovered, however. *See* INFRARED ASTRONOMY; STAR.

Harold A. McAlister

Bibliography. A. Fraknoi, D. Morrison, and S. Wolff, *Voyages Through the Universe*, 3d ed., Brooks Cole, 2004; I. Heinrichsen, H. J. Walker, and U. Klaas, Infrared mapping of the dust disk around Vega, *Monthly Notices Roy. Astron. Soc.*, 293:L78–L82, 1998; J. B. Kaler, *The Hundred Greatest Stars*, Springer, 2002.

Vegetable ivory

The seed of the tagua palm (*Phytelephas macrocarpa*) of tropical America. Each drupelike fruit contains six to nine bony seeds. The extremely hard endosperm of the seed is used as a substitute for ivory. Vegetable ivory can be carved and tooled to make buttons, chess pieces, knobs, inlays, and various ornamental articles. The hard, white seeds (coquilla nuts) of the Brazilian palm (*Attalea funifera*) are often used as a substitute for vegetable ivory in making the same or similar articles. *See* ARECALES.

Perry D. Strausbaugh; Earl L. Core

Vegetation and ecosystem mapping

The graphic portrayal of spatial distributions of vegetation, ecosystems, or their characteristics. Vegetation is one of the most conspicuous and characteristic features of the landscape and has long been a convenient way to distinguish different regions. Because vegetation provides the basic framework of terrestrial ecosystems, maps of ecosystems and biomes have been mainly vegetation maps. Resources generally must be inventoried and mapped before they can be well managed. Thus, as pressure on the Earth's natural resources grows and as natural ecosystems are increasingly disturbed, degraded, and in some cases replaced completely, the mapping of vegetation and ecosystems, at all scales and by various methods, has become more common and important. *See* BIOME; ECOSYSTEM.

Basic approaches and considerations. Probably since classical times, and certainly since the rapid development of mapmaking during the Renaissance, mappers have located vegetation in relation to natural or cultural patterns. Three approaches arose for mapping general vegetation patterns, (1) based on vegetation structure or gross physiognomy, (2) based on correlated environmental patterns, and (3) based on important floristic taxa. The environmental approach provides the least

information about the actual vegetation but succeeds in covering regions where the vegetation is poorly understood. While most earlier and many later vegetation maps involve combinations of the three approaches, most modern classification systems use a combination of physiognomic and floristic characters. *See* PLANT GEOGRAPHY.

The first world vegetation maps were produced in Europe near the end of the nineteenth century and were inconsistent since many regions of the world were not well known. As knowledge of obscure regions increased, better maps appeared for the whole world and for individual continents, countries, and regions. Perhaps the first world vegetation-based map free of environmental surrogates was that by A. W. Küchler in 1949. Excellent regional vegetation maps include those of the United States by Küchler (1964), South America by Hueck and Seibert (1972), the Mediterranean and adjacent regions by UNESCO (1969), Western Australia by Beard, and France and India by the French vegetation mapping groups at Toulouse, Grenoble, and Pondicherry (India). In the United States, excellent vegetation maps have also appeared for many states, with particularly detailed versions for California, Minnesota, and Wisconsin.

Mapping has expanded since the 1950s to involve other aspects of vegetation and ecosystems as well as new methodologies for map production. Functional processes such as primary production, decomposition rates, and climatic correlates (such as evapotranspiration) have been estimated for enough sites so that world maps can be generated. Structural aspects of ecosystems, such as total standing biomass or potential litter accumulations, are also being estimated and mapped. Quantitative maps of these processes or accumulations can be analyzed geographically to provide first estimates of important aspects of world biogeochemical budgets and resource potentials.

Computer-produced maps, using Geographic Information Systems (GIS), often coupled directly with predictive models, remote-sensing capabilities, and other techniques, have also revolutionized vegetation and ecosystem mapping. This gives scientists a powerful tool for modeling and predicting the outcome from global climate change, in that feedback from the world's vegetation can be accounted for. Before computer technology exploded in the early 1980s, the spatial scale and related resolution or grain of vegetation and ecosystem mapping was limited by the static nature of hard-copy maps. The advent of GIS technology enabled the analysis of digital maps at any spatial scale, and the only limitation was the resolution at which the data were originally mapped. In addition, GIS software is used for sophisticated spatial analyses on maps, and this was virtually impossible before. *See* CLIMATE MODELING; GEOGRAPHIC INFORMATION SYSTEMS.

Spatial scale and base maps. The concept of spatial scale is relevant to many fields, including geography, cartography, and ecology. Spatial scale is often confused with the term "resolution" or "grain" because static maps have a set scale and resolution (the smallest object that can be seen on the map). With a GIS-based map, one can zoom in and out electronically, adjusting the scale and the resolution simultaneously. The only limitation is the resolution or grain of the original electronic map. Consider, for example, an electronic map of Washington State in which the original resolution of the map is very high, that is, features as small as 10 m (33 ft) across were measured. If the whole state is in view (that is, small-scale representation), the small-sized features are not visible. However, if one zooms in on a city (that is, large-scale representation), the outlines of individual buildings will be seen.

In terms of nomenclature, a 1:24,000 scale map is large-scale, and a 1:1,000,000 scale map is small-scale. This is often reversed, which is incorrect.

The U.S. Geological Survey large-scale maps (at 1:24,000) may be too coarse for working with plant associations and diverse cultural features. They are widely available, however, and can be enlarged either photostatically or using GIS software for field work. U.S. Army Map Service sheets at 1:250,000 are available in a consistent style over wide areas. If maps of these two types have not become too outdated because of land clearing, green forest overlay printing on these maps will provide a first-order distinction between forest/woodland and other kinds of vegetation. Such maps at least can serve to show altitudinal zonation, and may further suggest relations to geologic substrate and soil-type associations if reliable soil maps can be found. *See* CARTOGRAPHY; GEOLOGY; GEOMORPHOLOGY; MAP DESIGN; MAP REPRODUCTION.

The mapping of large areas at coarser scales, even with the benefit of the remote-sensing techniques mentioned below, is of varying accuracy for regions with different economic and landscape conditions. A common denominator of mapping at a widely available and acceptable scale convention of 1:1,000,000 has been advocated.

The following hierarchy, the U.S. National Vegetation Classification (USNVC) system, was developed jointly by over a dozen federal agencies, The Nature Conservancy (TNC), and the Ecological Society of America. It is used for classifying terrestrial ecological communities, based primarily on their vegetation.

System (for example, Terrestrial)

Physiognomic levels:
- Formation Class
  - Formation Subclass
    - Formation Group
      - Formation Subgroup
        - Formation

Floristic levels:
- Alliance
  - Association

The classification system is internationally recognized, and it builds upon the UNESCO (1973) classification by incorporating the classification scheme

Summary of the classification hierarchy by the Association for Biodiversity Information

Level	Primary basis for classification	Level divisions and examples
Class	Type, height, and relative percentage of cover of the dominant, uppermost vegetation	Seven classes: Forest, Woodland, Shrubland, Dwarf-shrubland, Herbaceous, Nonvascular, and Sparse Vegetation
Subclass	For Forest, Woodland, Shrubland, and Dwarf Shrubland classes: leaf character	Three subclasses in each: evergreen, deciduous, and mixed evergreen-deciduous (no mixed evergreen-deciduous, dwarf-shrubland subclasses have been defined)
	For Herbaceous class: persistence and growth form	Four subclasses: perennial grasslands, perennial forb vegetation, annual grass and forb vegetation, and hydromorphic vegetation
	For Nonvascular class: relative dominance of nonvascular vegetation type	Three subclasses: lichens, mosses, algae
	For Sparse Vegetation class: particle sizes of the substrate features	Three subclasses: consolidated rock; boulder, gravel, cobble, or talus; and unconsolidated material (soil, sand, or ash)
Group	Varies by class: leaf characteristics, broad climatic types, presence and character of woody strata, major topographic position types or landforms	About 60 groups; example: Temperate or Subpolar Needle-leaved Evergreen Forest
Subgroup	Relative human impact	Two subgroups: Natural/Semi-natural or Cultural
Formation	Additional structural and environmental factors, including hydrology	Many; example: Saturated Temperate or Subpolar Needle-leaved Evergreen Forest
Alliance	Dominant/diagnostic species, usually of the uppermost or dominant stratum	Many; example: *Picea mariana* Saturated Forest Alliance
Association	Additional dominant/diagnostic species from any strata	Many; example: *Picea mariana/Alnus incana/ Sphagnum* Forest

of R. S. Driscoll et al. (1984), and by increasing detail with the addition of alliance and association floristic levels. It is essentially identical to the National Vegetation Classification Standard (NVCS, FGDC 1997). This system also includes global rank definitions of conservation status ranking, making it an extremely powerful tool for identifying and prioritizing communities for conservation. However, worldwide maps have not yet been produced from this classification system. The NVCS classification hierarchy is summarized in the **table.**

For large countries, wall maps with good color printing can show many details by using patterns and symbols at smaller scales. Again, GIS-based maps are not necessarily limited this way. For the United States, Küchler's 1:3,168,000 map of 116 kinds of "Potential Natural Vegetation" has considerable detail related to patterns of rivers, which are shown on the map, and mountains, which are not shown but are readily recognized by the distribution of montane conifer or meadow types. For the former Soviet Union, an eight-sheet wall map at 1:4,000,000 shows even finer detail for the actual vegetation, with 109 colored mapping units, many of them subdivided by letters and supplemented by symbols for significant species. Coarser scales were required (varying with continent and type of map) in the Soviet Union's "Physical-Geographic Atlas of the World," but there was remarkably little loss of detail.

Colors are generally required for good maps of vegetation types, but process and structural quantities can usually be shown best by black-and-white shading. Computer-printed GIS maps are easily edited, so a series of maps can be rapidly produced that represent enormous amounts of information. In addition, sophisticated spatial analyses can be done using GIS software that are not possible using conventional static maps.

Topographic base maps may be improved to show such features as slope and aspect (direction), as well as altitude, in more detail. Simple, pale contour lines provide this information without distracting from the markings superimposed to represent vegetation or ecosystems. Topographic shading makes relief and drainage stand out more sharply but may interfere with the biological patterns.

Data gathering. Two basic approaches to detailed local vegetation analysis are (1) classification, which is usually based on intensively sampled plots, called relevés or quadrats, and (2) ordination, which usually involves some form of gradient or continuum analysis. Ordination is a method for plotting and

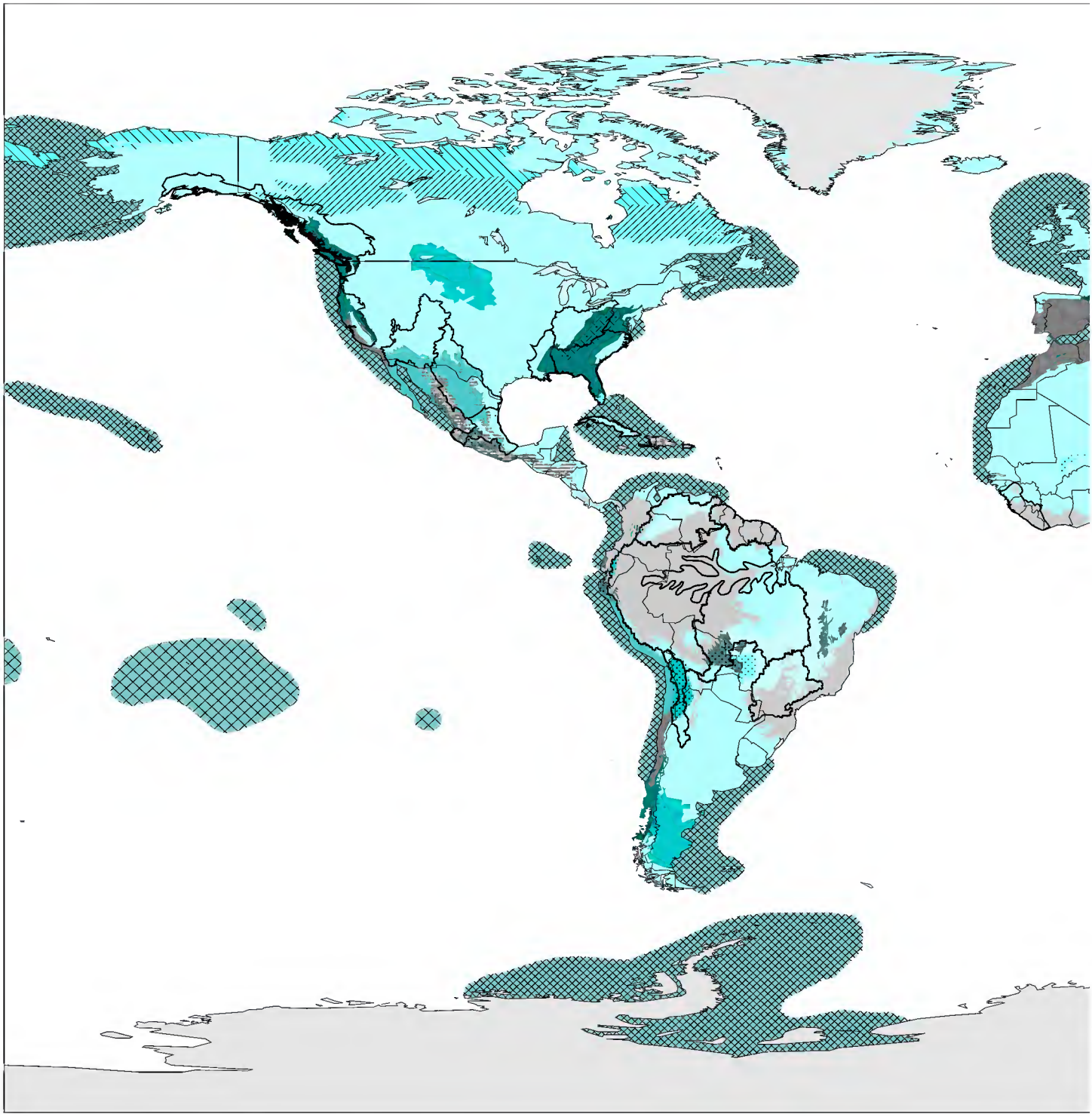

Terrestrial major habitat types.

inferentially ordering sampled vegetation plots in a usually hypothetical environmental space based on the relative similarity or dissimilarity of plots to each other. Both approaches require detailed field sampling and may yield data and patterns which are not easily mapped. More generalized mapping that focuses on vegetation types and other ecological patterns over larger areas requires more extensive data gathering, which can be obtained from remote-sensed imagery. Remote-sensed imagery is any imagery that is collected from a remote location, such as an aircraft or a satellite. There are two types of remote-sensed imagery: aerial photography and digital imagery. Aerial photography uses the familiar glass lens and photosensitive film technology. It is usually deployed from aircraft, but satellite deployment is also used. Digital imagery may or may not require a lens, but the photosensitive film has

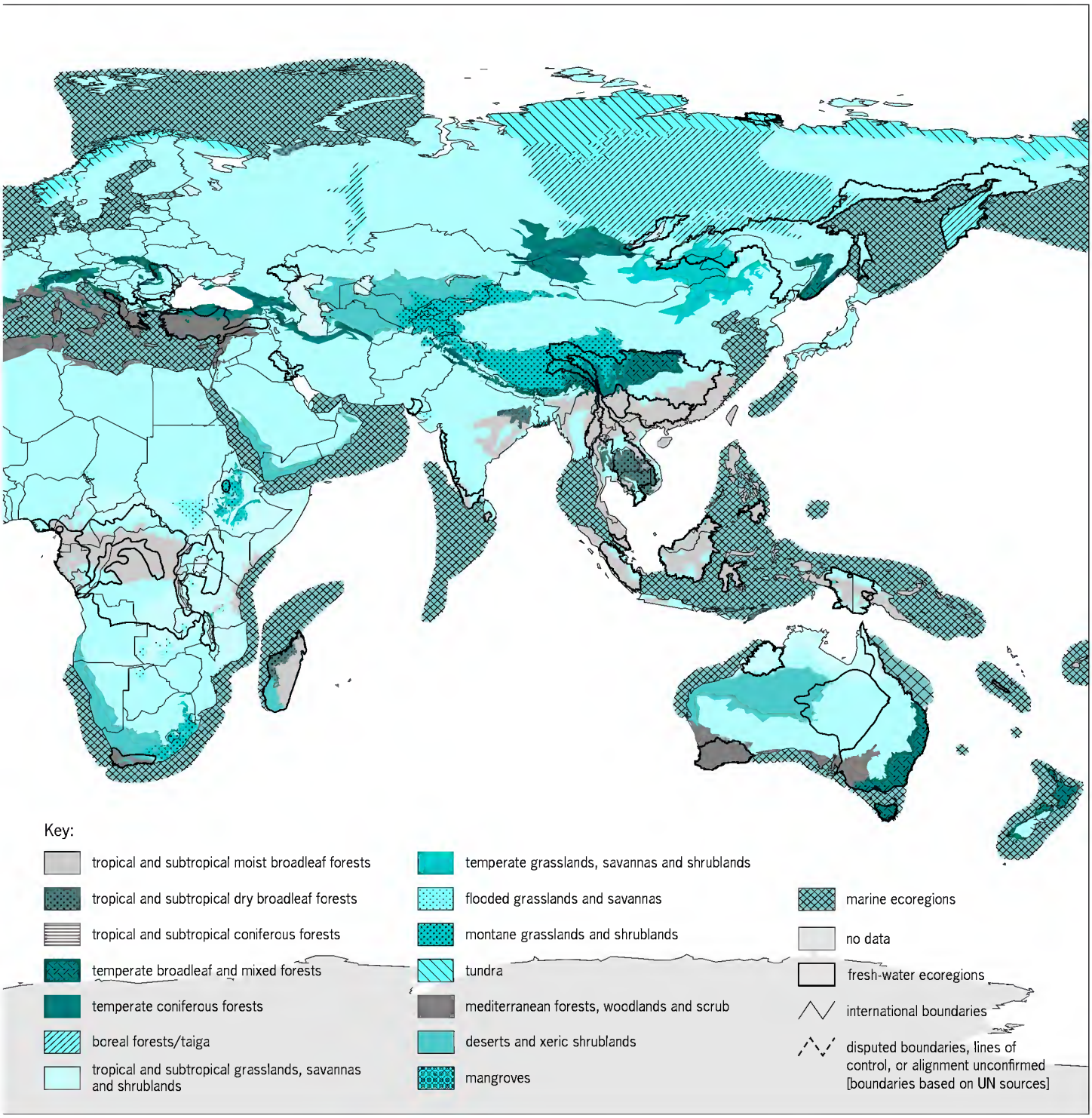

been replaced by electronic sensors that send their information to computers, where the data are stored and analyzed. Digital imagery has typically been deployed from satellites, but with improvements in technology, aircraft-deployed sensors are becoming more common. *See* AERIAL PHOTOGRAPH.

Aerial photography. Aerial photography has been used since the 1920s to provide direct images of forests and other visibly distinct vegetation cover, with or without extensive ground checking to provide species recognition and other calibration data. This has been especially useful in countries with large, sparsely populated areas. Large areas are stereoscopically covered from high altitudes for surveys, road and harvest plans, and fire control operations. In addition, very fine photographs have been made of statistically selected plots, permitting measurement of shadow and tree height and crown form to provide

a basis for estimating tree volume and biomass. *See* PHOTOGRAMMETRY.

Soil surveys are also made from aerial photographs, such as to estimate soil moisture or to measure areas for which conservation payments are made to individual farmers. But even when land is freshly plowed, there are limits to what can be inferred from the air or from observation of surface soil horizons. Local experience and "ground-truthing" are necessary, but soil variations may still obscure the apparent surface patterns. Ground-truthing involves checking (on the ground) certain elements of what was seen by aircraft or satellite.

The experienced mapper relies on the "lay of the land" (perhaps perceptibly slight differences in concavity or convexity) and the quality and vigor of vegetation to help fill in the incomplete observations of substrate conditions. The balance of intensive and extensive work and of "calibration" checks and routine extension of the mapped area may depend greatly on available time, money, and competing duties. Administrative considerations tend to favor statistical control and objective measures of unexplained variance within the somewhat arbitrarily grouped mapping units. Yet for the understanding of vegetation, soil, or other terrain variables, which correlate with the whole ecosystem's condition, nonroutine detectivelike insights are needed, as well as the orderly use of what has been guessed, observed, and confirmed. *See* GLACIATED TERRAIN; SOIL CONSERVATION.

Digital imagery. Digital imagery technologies have revolutionized the way in which vegetation and ecosystem mapping data are collected and analyzed. With the use of sophisticated electronic sensors and computers, workers can extract enormous amounts of information about the Earth's vegetation and ecosystems that was not possible using only aerial photograph analyses. Applications include agricultural crop type mapping and monitoring; forestry clear-cut and burn mapping and species delineation; and land cover and land use mapping of changes in agriculture, urban and rural areas, and biomass. In addition, workers can now measure ecosystem changes that are occurring throughout the world at spatial scales that were never practical with previous technology.

There are two general digital imagery systems: active and passive. Active systems use lasers (for example, laser induced distance and ranging, LIDAR) or radar (for example, synthetic aperture radar, SAR) to send out light or electromagnetic radiation and measure the reflected signal. Passive systems (such as *Landsat Thematic Mapper*, *TM*) measure combinations of wavelengths of electromagnetic radiation reflected differentially from the Earth. *See* LIDAR; RADAR; REMOTE SENSING.

There are three general types of passive digital imagery: panchromatic, multispectral, and hyperspectral. Panchromatic imagery shows vegetation in black and white and has limited use. It sees electromagnetic radiation that is visible to the human eye (for example, *India Remote Satellite*, *IRS*). Multispectral imagery (for example, *Landsat TM*) usually has a dozen or less detection bands, suitable for differentiating gross classes of vegetation. This type of imagery measures the visible spectrum of light (that is, red, green, and blue light) but also adds invisible spectra such as infrared (thermal) and ultraviolet. Hyperspectral imagery, with over 200 bands, can discriminate between individual plant species, or it can identify diseased areas. Most of the spectra measured are not visible to the human eye, so sophisticated computers are used to digest the vast amounts of information. In the future, there will be ultraspectral imagery, containing thousands of bands, increasing the ability to identify as yet unseen properties of the Earth's vegetation. *See* INFRARED RADIATION; ULTRAVIOLET RADIATION.

There are two ways of characterizing the resolution of digital imagery: spectral and spatial. Spectral resolution refers to how many different wavelengths of electromagnetic radiation a sensor can detect. Some hyperspectral sensors (for example, Compact Airborne Spectrographic Imager, CASI) can detect over 250 bands. Spatial resolution refers to the smallest object that can be detected by the sensor. The spatial resolution is usually expressed as a gridcell size. For example, each gridcell of the *Landsat TM* satellite is 30 m (100 ft) on a side. Therefore, anything smaller than 30 m on the ground will not be seen very clearly by the satellite. Some of the newer satellites have sensors that can detect ground features as small as 1 m (39 in.).

Active imagery technologies can gather three-dimensional information about vegetation that is impossible to perceive with passive-type remote sensing. For example, the height and even volume of the tree canopy can be measured. However, the spectral resolution is not as good; that is, species classification is by gross morphology or texture, as opposed to reflected light. Using passive remote sensing, the complex nature of how light of varying wavelengths is reflected from the Earth is measured, so a greater resolution of species types is possible.

Some of the most powerful digital imagery technologies combine active and passive sensors. For example, LIDAR, SAR, and *Landsat TM* sensors can be integrated, giving a unique three-dimensional representation of vegetation species types, morphology, and community structure. In addition, with the LIDAR technology, detailed information about the terrain and elevation is incorporated, facilitating vegetation-elevation correlation analyses.

Map examples. Local detail must be generalized greatly when one attempts an overview of the whole Earth system (see **illus.**).

As a first step, one might focus on environmental characteristics. L. Holdridge's system of life zones, for example, though without seasonal patterns, has been the basis for many maps of potential vegetation, especially in the tropical and subtropical Americas. The resulting data set has been used for predicting the influence of climate change on the world's vegetation patterns.

In addition to vegetation types, various maps have shown estimated world patterns of annual primary production, based both on generalization from measurements and on prediction from climatic relationships. Omission of animal consumption of vegetation may make the estimates low but probably will not change the striking geographic pattern shown. Plant respiration, detrital decomposition, and other basic processes have also been mapped by prediction from climatic relationships.

Structural phenomena, which have more complex environmental relations and trade-offs, have become much easier to model using remote-sensing technology, coupled with direct measurements. The map of the terrestrial storage of carbon by E. Matthews et al. (2000) is based not just on above- and below-ground live vegetation but also on carbon stored in soils. Maps of this nature become powerful tools when they are used to measure changes (natural or from human influence) over time that occur in the Earth's ecosystems. For example, the United Nations Environment Program and the World Conservation Monitoring Center (UNEP-WCMC) created a composite map of the global distribution of original and remaining forests, showing that nearly 50% of the Earth's forests have been destroyed. Details of patterns of ecosystem rates and parameters, on the scale of the world vegetation zones and on the more detailed scale showing the accelerating changes due to humans, continues to be a major challenge for future generations of ecologists.

Numerous international agencies are involved with long-term research on the Earth's ecosystems. Much of their information comes from remote-sensed data for measuring changes in the Earth's vegetation and ecosystems. In addition, further advances in technology have enabled scientists to monitor climate change and predict the impacts of global climate change on the Earth's ecosystems.

Agencies that are responsible for studying global environmental change and its impact on ecosystems include Global Change and Terrestrial Ecosystem (GCTE), Global Terrestrial Observing System (GTOS), and Global Climate Observing System (GCOS). *See* CLIMATE MODELING; GLOBAL CLIMATE CHANGE.

The significance and application of vegetation and ecosystem mapping has changed significantly since the time of Küchler over 50 years ago. The stakes are high, with worldwide deforestation, global climate change, accelerated species extinction rates, and declining biodiversity. Scientists need vast amounts of information about the planet, and they need it quickly in order to save the Earth's various ecosystems. *See* BIODIVERSITY; ECOLOGY; EXTINCTION (BIOLOGY); TERRESTRIAL ECOSYSTEM

. Blake E. Feist; Elgene O. Box

Bibliography. M. P. Anderson et al., *International Classification of Ecological Communities: Terrestrial Vegetation of the United States*, vol. II: *The National Vegetation Classification System: List of Types*, The Nature Conservancy, Arlington, VA, 1998; R. S. Driscoll et al., An Ecological Land Classification Framework for the United States, *U.S. Forest Serv. Misc. Publ.*, no. 1439, 1984; Federal Geographic Data Committee, *Vegetation Classification Standard*, FGDC-STD-005, 1997; D. H. Grossman et al., *International Classification of Ecological Communities: Terrestrial Vegetation of the United States*, vol. I: *The National Vegetation Classification System: Development, Status, and Applications*. The Nature Conservancy, Arlington, VA, 1998; R. Leemans, *Global Data Sets Collected and Compiled by the Biosphere Project*, Working Paper, IIASA-Laxenburg, Austria, 1990; E. Matthews et al., *Pilot Analysis of Global Ecosystems: Forest Ecosystems*, World Resources Institute, 2000; UNESCO, International Classification and Mapping of Vegetation, Series 6 of *Ecology and Conservation*, 1973.

Velocimeter

An instrument that measures the velocity of a flowing liquid or gas, usually understood to be the average velocity of fluid within a relatively small measuring volume. Velocimeters are distinct from flow meters, which measure spatially averaged velocity, or flow rate, across the cross section of a duct or channel. Depending on the nature of the flow and the objective of the measurement, a velocimeter may be used to measure a time-dependent flow velocity or a time-averaged value. Another characteristic of a velocimeter is its spatial resolution, namely the size of its measuring volume. Some velocimeters measure flow velocity in both magnitude and direction, whereas others only resolve a velocity component in a particular direction or plane. The term anemometer is sometimes used as synonymous to velocimeter, although, strictly speaking, the former term should apply only to instruments that measure air velocity. *See* FLOW MEASUREMENT.

Velocimetry is an essential activity in fluid mechanics research laboratories and in flow-related industrial and environmental applications. The operation of different velocimeters is based on a variety of principles:

1. Pressure tubes, including Pitot-static tubes, relate flow velocity to measurable pressure differences in the flow.

2. Thermal anemometers, such as the hot-wire and the hot-film anemometers, relate flow velocity to the rate of convective heat transfer from a heated sensor. Pulsed-wire anemometers track the convection of heat pulses injected into the flow by small sources.

3. Laser Doppler velocimeters relate flow velocity to the frequency shift of laser light scattered by fine particles transported by the flow. The same principle, but applied to sound waves, is utilized by ultrasonic Doppler velocimeters.

4. Particle tracking measures flow velocity by optically monitoring the displacement of particles transported by the flow. A refined and powerful extension of this method is particle image velocimetry.

5. Popular instruments used for monitoring wind speed, such as the cup and propeller anemometers,

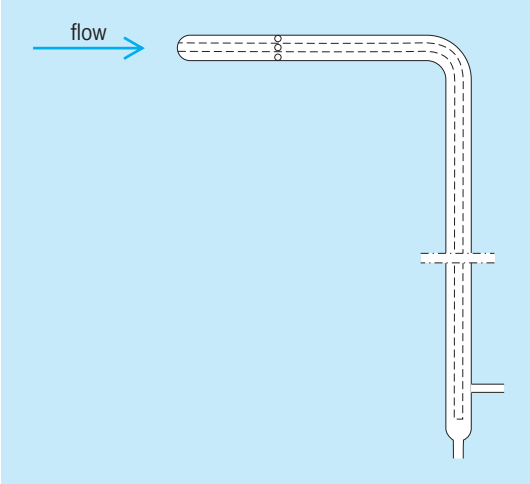

Fig. 1. Pitot-static tube.

measure flow velocity via a force or torque applied on a mechanical element.

Pressure tubes and multihole probes. Pressure tubes are thin hollow tubes inserted into a stream to measure its velocity V, static pressure P, or total pressure P_o. In idealized flow, these three parameters are related to each other through Bernoulli's equation (1), where ρ is the fluid density and the combi-

$$P_o = P + \tfrac{1}{2}\rho V^2 \qquad (1)$$

nation $\tfrac{1}{2}\rho V^2$ is called the dynamic pressure. A Pitot tube, also called total or impact tube, is a simple tube with an open end facing the flow and the other end connected through flexible tubing to a manometer or pressure transducer. A Pitot tube would ideally read the average total pressure across its face. A static tube is also hollow, but, instead, has a sealed, rounded nose facing the flow and a number of small holes on its side. It reads approximately the static pressure in the free flow. A coaxial combination of these tubes, called the Pitot-static tube (**Fig. 1**), reads directly the dynamic pressure as $P_o - P$, from which the flow velocity can be easily calculated with the use of Eq. (1). In this device, the inner tube faces the flow and reads the total pressure, whereas the outer tube is sealed in the nose, has circumferential holes, and reads the static pressure. *See* BERNOULLI'S THEOREM; MANOMETER; PITOT TUBE; PRESSURE MEASUREMENT; PRESSURE TRANSDUCER.

Pressure tubes are simple, inexpensive, and easy to operate as they do not require calibration and their output can be monitored by simple instruments, like a home-made manometer. On the other hand, they also have considerable limitations. Like any other intrusive method, they may distort the flow and its properties, particularly the static pressure. They require alignment within a few degrees with the flow direction, so they cannot be used in flows that keep changing direction or are highly turbulent, such as would be the case with the wind. Their temporal response is slow, typically requiring one minute or more to settle after they are first connected or following a sudden change in flow velocity. Bernoulli's equation is based on the assumptions of steady, incompressible (that is, with constant density) and inviscid (that is, frictionless) flow, so pressure-static tubes would be subjected to errors when used in flows that deviate significantly from these conditions. Examples of inappropriate use include measurements very close to a wall; measurements in fluids with very high viscosity, such as thick oils; measurements in very low speed flows; and when the flow speed is not very small compared to the speed of sound. For air at room temperature, this means that simple pressure tubes are usable only in the range between about 0.3 and 100 m/s (0.7 and 220 mi/h). Compressible flow relationships must be used at higher speeds. *See* COMPRESSIBLE FLOW; VISCOSITY.

Different combinations of pressure-measuring devices can be used to measure flow direction (yawmeters) or both flow direction and velocity magnitude. An example is the five-hole probe (**Fig. 2**), which consists of five thin tubes fastened together. The central tube has a square nose facing the flow and measures the total pressure, while the four other tubes have slanted noses and measure pressures that are intermediate between the total and static values, depending on the orientation of the flow velocity. Such devices require precise manufacturing, elaborate calibration over a range of angles, and a sophisticated algorithm to recover the velocity magnitude and direction from the simultaneous readings of the five pressures. *See* YAW INDICATOR.

To overcome the slow response of pressure tubes, researchers have proposed the mounting of miniature pressure transducers within the tubes. Such devices, called fast-response pressure probes, can provide measurements in transient and turbulent flows, but their outputs are also subjected to additional errors due to electronic interference and vibrations.

Thermal anemometers. In thermal anemometry, the sensor is a fine, short metallic element, inserted in the flow and supplied with an electric circuit, such that the generated Joule heat increases its temperature above that of the surrounding flow. For metallic elements, the resistance also rises above its value for the unheated sensor. If the flow velocity rises, then more heat is removed from the sensor, so that its temperature tends to decrease. In the usual thermal anemometry configuration, called constant temperature anemometry (CTA), the sensor comprises one side of a Wheatstone bridge, with the other three sides consisting of resistors. A high-gain difference amplifier senses any imbalance in the bridge as a result of a change in the sensor resistance, and provides

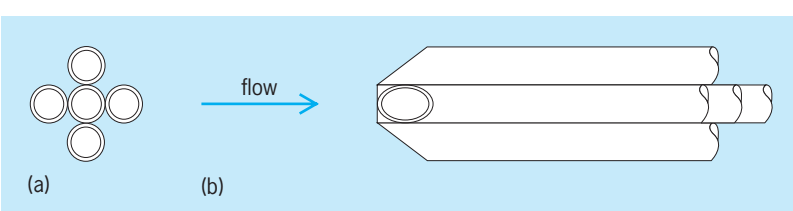

Fig. 2. Five-hole probe. (*a*) Front view. (*b*) Side view.

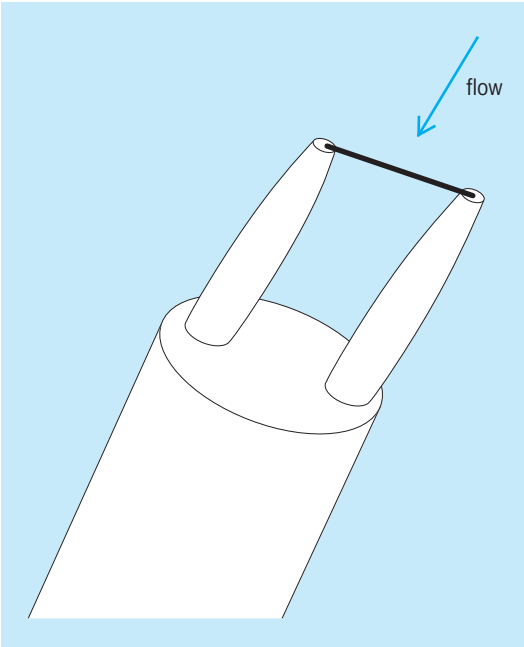

Fig. 3. Hot-wire probe.

additional feedback current that tends to increase the amount of heat generated by the sensor. This tends to increase its temperature, which, in turn, increases its resistance, thus restoring the bridge balance in such a way that the sensor resistance is always kept equal to a preset value. The voltage across the sensor is related to the flow velocity and temperature. With proper calibration, and if the flow temperature is maintained equal to that during calibration or is measured by other means, the flow velocity may be determined from the measured voltage. *See* RESISTANCE MEASUREMENT; WHEATSTONE BRIDGE.

The hot-wire anemometer (HWA) has a cylindrical sensor made of platinum-plated tungsten or platinum alloys. A single-sensor hot-wire probe has an insulated body containing two parallel metallic needles (prongs), at the free tips of which the sensor is mounted and through which current is flowing (**Fig. 3**). A cylindrical sensor is mainly sensitive to the velocity component normal to its axis, and so it cannot resolve any other velocity components. Two- and multisensor hot-wire probes (**Fig. 4**), in which the sensors are inclined with respect to the average flow direction, can, with proper angular calibration, measure two velocity components (cross-wire anemometers) or even the full velocity vector. Hot wires can be used only in gases and nonconducting liquids. Due to their fragility (a typical sensor has a diameter of 5 μm and a length of about 1 mm) and high sensitivity, they require an environment free of dust, droplets, and other particles that may impact on them and break them or alter their heat transfer characteristics. Hot films have sensors that are thin (about 1 μm thick) films of nickel, deposited on a ceramic substrate and coated by a quartz insulation, so that they are insulated from the flowing fluid and can be used in tap water and other conducting liquids. Because of their sturdier construction, they are less sensitive than hot wires to flow impurities and aerodynamic loading.

An algebraic equation for the response of hot wires and hot films may be formulated by considering that Joule heating is the energy input and convective heat transfer is nearly all of the energy output, while the sensor temperature is maintained constant so that there is no change of energy stored in the sensor. This equation connects the sensor voltage E, the flow velocity V, the flow temperature T_f, and and the sensor temperature T_s. Such an equation would also contain geometrical, physical, and electrical properties of the sensor, but such properties cannot be known precisely and also may change as the sensor ages in operation. For this reason, instead of an exact equation, thermal sensor response is described by the semiempirical relationship (King's law) given in Eq. (2), with the coefficients A, B, and n determined

$$\frac{E^2}{T_s - T_f} = A + BV^n \tag{2}$$

by calibration in a flow of known velocity and temperature, such as a small calibration jet. The optimal value of n is near 0.5, but the values of A and B depend on the properties of each individual sensor and the settings of the electrical circuit.

A great advantage of thermal anemometers is their excellent spatial and temporal resolution, making them suitable for statistical measurements in turbulent and unsteady flows. In addition to their fragility and the need for calibration, disadvantages include sensitivity to flow direction, inability to resolve flow orientation in reversing flows, and sensitivity to flow temperature and composition, besides that to flow velocity.

A great variety of thermal anemometer sensors and driving circuits are available commercially at considerable cost. A skilled technician can put together both simple hot-wire probes and electronic circuitry at a fraction of this cost, based on available designs. The recording and processing of thermal anemometer signals is best done with the use of a digital data acquisition system connected to a computer.

Laser Doppler velocimeter. The laser Doppler velocimeter (LDV), alternately referred to as a laser Doppler anemometer (LDA), utilizes the Doppler phenomenon, namely the fact that the frequency of waves scattered by a moving object is different from the frequency of incident waves, with the difference (Doppler frequency) being proportional to

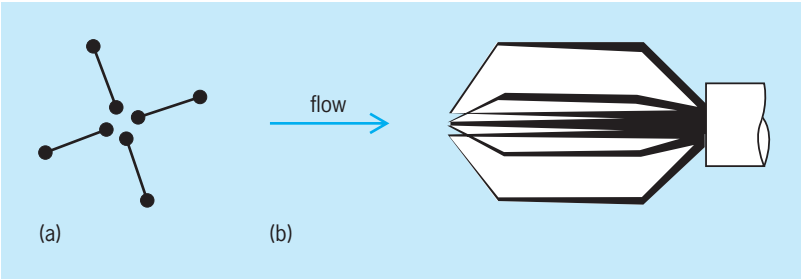

Fig. 4. Four-wire probe. (*a*) Front view. (*b*) Side view.

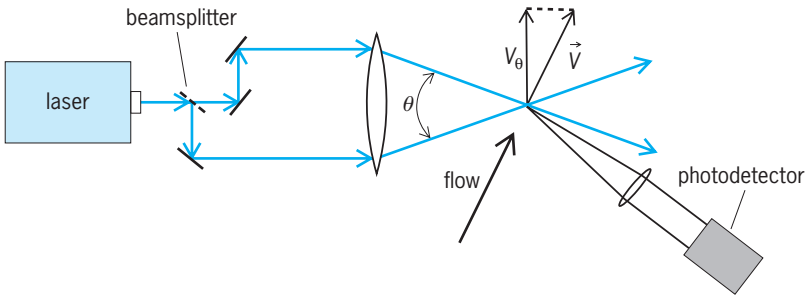

Fig. 5. Laser-Doppler velocimeter.

the component of the velocity of the object along the bisector of the angle between the directions of propagation of the incident and scattered waves. For practical reasons, an optical Doppler system must use a laser beam, which consists of high-intensity, collimated, coherent (that is, with all light waves in phase), monochromatic (that is, with a single frequency, or a small number of distinct frequencies) light. Because the Doppler frequency is far smaller than can be resolved by any available photosensitive device (photodetector, commonly a photomultiplier tube or a photodiode), the light frequency detection system is arranged in a way that it detects the Doppler frequency directly and not as the difference of two slightly unequal frequencies. The most popular configuration is the dual-beam system (**Fig. 5**). In this arrangement, the laser beam is split by a beamsplitter (for example, a partially transparent mirror or prism) and, with the use of additional optical components, forms two parallel beams. These beams are made to intersect in a small measuring volume. A receiving lens collects the light from the measuring volume and projects it on the photodetector, which provides an electric signal that is proportional to the intensity of the received light. *See* DOPPLER EFFECT; LASER.

Because homogeneous fluids, such as clean air and water, would not provide a signal, light-scattering particles are required to cross the measuring volume. Tap water and room air may contain sufficient impurities to act as scatterers, but most commonly it is preferable, or even necessary, to seed the flow, that is, to introduce a small amount of solid particles or liquid droplets in an even distribution. Compared to natural impurities, flow seeding has the advantage of allowing control of the size and shape of the particles, thus reducing errors that would be caused by particles of different masses moving at different speeds relative to the fluid. Clearly, it is important that the motion of foreign particles matches the motion of the surrounding fluid, which makes it necessary to use very small particles, typically of dimensions comparable to 1 μm.

In a dual-beam system, the velocity component V_θ normal to the bisector of the angle θ between the intersecting beams is determined from the frequency f_D of the collected light to be given by Eq. (3),

$$V_\theta = \frac{\lambda}{2\sin(\theta/2)} f_D \tag{3}$$

where λ is the wavelength of the laser light. Thus, it is only necessary to measure f_D and no calibration is required. Among the various methods for determining the frequency of the photodetector signal, the one most widely used currently is burst analysis. In this method, it is ensured that the particle concentration is sufficiently small for no more than a single particle to reside in the measuring volume at any time instant and a dedicated device computes the peak of frequency spectrum of the light emitted by a single particle (burst) by performing a fast Fourier transform (FFT). *See* FOURIER SERIES AND TRANSFORMS.

Compared to hot-wire anemometry, laser Doppler velocimetry has the advantage of being nonintrusive and usable in media containing impurities. As it measures true local velocity, it is insensitive to fluid temperature and composition. As with hot wires, the laser Doppler velocimeter system described above cannot discriminate flow direction, but this capability can be easily realized through a method called frequency shifting. In terms of frequency response and spatial resolution, laser Doppler velocimetry is inferior to hot-wire anemometry.

To measure more than one velocity component, it is necessary to use light of different frequencies. A popular approach is to separate the light of an argon-ion laser into two pairs of green and blue beams on two perpendicular planes; both pairs are made to intersect within the same measuring volume, the two colors are separated by optical filters in the receiving system, and each single-color component is processed by a separate photodetector and burst analyzer.

Laser Doppler velocimeter systems are available commercially from several suppliers and are too complex for the common user to put together from generic components. Their cost is quite high, but less-expensive educational systems are also available.

Particle image velocimeter. Particle tracking is the direct method of determining flow velocity by timing the displacement of markers that move with the flow. In the simplest case of readily visible, slowly moving, isolated particles, it is sufficient to simply employ a ruler and a stopwatch, but for the method to be applied to complex flows with high resolution, it is usually necessary to capture consecutive particle positions in rapid succession and to be able to separate and identify the images of closely spaced particles. This need has led to the development of digital particle image velocimetry (DPIV or PIV). In a basic particle image velocimetry system (**Fig. 6*a***), a laser beam is converted to a sheet of light with a small thickness by passing through a glass rod (or special lenses for improved quality). The flow is seeded with small particles (typically 1 μm or less in diameter) and two images of a small portion of the light sheet are recorded, illuminated by two light pulses. Although light pulses can be produced by using a shutter or a rotating disk to interrupt the beam of a continuous-wave laser, such as an argon-ion laser, the most powerful method is to use a twin-cavity Nd:YAG (neodymium yttrium aluminium garnet) laser, which

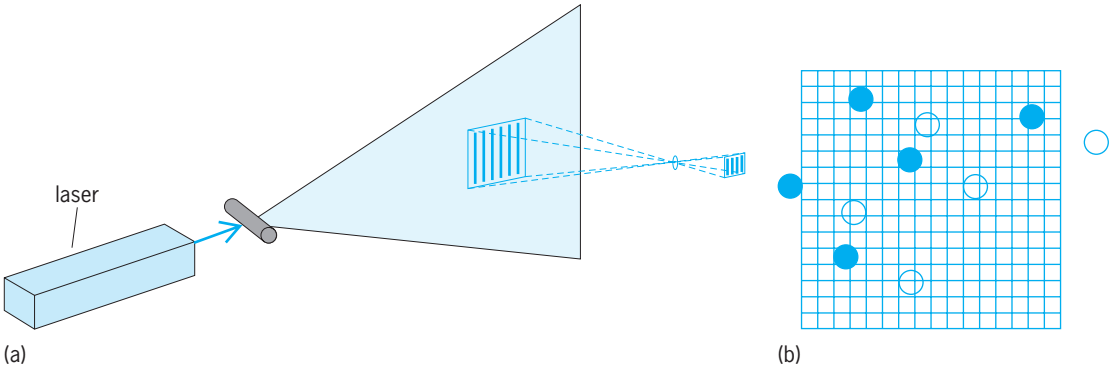

Fig. 6. Particle image velocimeter. (*a*) Apparatus. (*b*) Superposition of two images with pairs of images of particles that are within the measuring volume.

produces high-energy (100–400 mJ) light pulses separated by a time difference δt between 100 ps and 10 ns. Digital photography has advanced sufficiently to allow the independent recording of the two images on two different digital frames. When the particle concentration is sufficiently large, the superposition of the two frames contains a large number of pairs of images of the different particles that are within the measuring volume (Fig. 6*b*). These images can be sorted statistically with the use of sophisticated algorithms employing the cross-correlation technique. Then, each particle's velocity projection on the light sheet plane can be computed by dividing the particle displacement by δt.

Whereas hot-wire anemometry and laser-Doppler velocimetry measure an average flow velocity within the measuring volume, thus being point-measuring methods, particle image velocimetry provides the velocity distribution within an area. This is achieved by dividing the recorded image into smaller elements, called interrogation areas and each containing several particle images, and then determining the velocity in the center of each area by averaging the velocities of the corresponding particles. This makes it possible to draw instantaneous streamlines (that is, lines tangential to the velocity vector) in the flow and compute additional properties, such as the vorticity component normal to the image plane. For this reason, particle image velocimetry has gained great popularity. As laser, camera, and software technology keeps improving, it is expected that the power of particle image velocimetry will further increase. A variation of the basic method is stereoscopic particle image velocimetry, which uses two cameras at different angles to resolve the full three-dimensional velocity vector. Another improvement is time-resolving particle image velocimetry, which provides pairs of particle images at a succession of relatively closely spaced times, thus improving the temporal resolution of the basic method, which is much lower than those of hot-wire anemometry and laser Doppler velocimetry. Early particle image velocimetry systems were put together by individual research teams, but a variety of commercial systems are currently available. As with many other electronic devices, particle image velocimetry systems have improved in quality, while also becoming more affordable.

Ultrasonic velocimeters. The operation of the ultrasonic Doppler velocimeter is based on the same physical principle as laser Doppler velocimetry, but utilizes ultrasound rather than light. The sound is produced by a transmitter, which also alternates as a receiver, collecting scattered sound as it returns toward it. The average velocity of particles scattering sound along the path of the beam is determined from the difference in the frequencies of the transmitted and received signals. By measuring the time delay between a transmitted and a received sound pulse, it also becomes possible to determine the location of the particle, from which it is possible to obtain not only the average velocity, but also an instantaneous velocity profile along the beam path. A disadvantage of this method is that it cannot measure directly the axial velocity component in pipe and duct flows. A great advantage, compared to laser Doppler velocimetry and particle image velocimetry, is that it does not require optical accessibility, and therefore it is suitable for opaque fluids.

Another kind of ultrasonic velocimeter, used primarily in atmospheric research, has a pair of transmitter-receiver transducers facing each other. It measures the average flow velocity along the sound path from the difference in the speed of sound propagating in the two directions along this path, considering that sound travels faster if it propagates with the flow than against it. *See* SOUND; WIND MEASUREMENT.

Cup and propeller anemometers. The cup anemometer (**Fig. 7**) is a familiar device, frequently installed on towers near airports and in meteorological stations for measuring wind speed. It consists of three or four hollow hemispherical cups mounted at some radial distance from a rotating shaft and such that all the concave sides face in the same circumferential direction. When the instrument is exposed to a flowing stream, it will always rotate in the same sense, because the drag on the concave side of each cup is higher than the drag on its convex side. With its shaft vertical, the cup anemometer would rotate at an angular velocity ω which is proportional to the magnitude V_b of horizontal wind speed, but it is

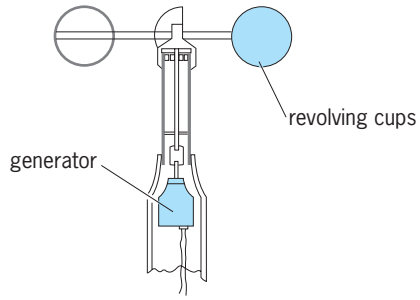

Fig. 7. Cutaway diagram of a rotating cup anemometer. (*After G. K. McMillan and D. M. Considine, eds., Process/Industrial Instruments and Controls Handbook, 5th ed., McGraw-Hill, 1999*)

incapable of resolving wind direction or the vertical wind component. For this reason, it is often accompanied by a rotating vane, which aligns with the wind direction. To measure the vertical wind magnitude, a second cup anemometer can be used, this time with its shaft horizontal. The speed of rotation is monitored by an electric generator coupled to the shaft. All that is required is the proportionality coefficient between ω and V_b, which is obtained by calibration. Although the measuring volume of cup anemometers is relatively large (typically 200 mm or 8 in. across), it is acceptable for most atmospheric applications. In gusty and strongly turbulent winds, cup anemometers tend to overspeed, that is, to read wind speeds higher than the actual one. Moreover, their accuracy deteriorates at speeds lower than 0.25 m/s (0.5 mi/h).

A related device is the propeller anemometer, which actually operates like a wind turbine. It consists of a small axial propeller mounted on a rotating shaft, whose speed of rotation is proportional to the flow velocity parallel to its axis. A commonly used device is the Gill anemometer, which has four helicoidal blades. Like cup anemometers, propeller anemometers require calibration to determine the proportionality coefficient. They also tend to overspeed in turbulent winds, but much less than cup anemometers. If not aligned with the flow direction, they tend to read values lower than the actual wind velocity projection on their axis. Three propeller anemometers mounted on mutually orthogonal shafts can be used to measure wind in both magnitude and direction. Stavros Tavoularis

Bibliography. H. H. Bruun, *Hot-Wire Anemometry*, Oxford University Press, 1995; E. O. Doebelin, *Measurement Systems: Application and Design*, 5th ed., McGraw-Hill, 2004; R. J. Goldstein (ed.), *Fluid Mechanics Measurements*, 2d ed., Taylor & Francis, 1996; F. Mayinger and O. Feldman (eds.), *Optical Measurements: Techniques and Applications*, 2d ed., Springer, 2001; M. Raffel, C. E. Willert, and J. Kompenhans, *Particle Image Velocimetry: A Practical Guide*, Springer, 1998; S. Tavoularis, *Measurement in Fluid Mechanics*, Cambridge University Press, 2005; J. G. Webster (ed.), *Mechanical Variables Measurement: Solid, Fluid, and Thermal*, CRC Press, 2000.

Velocity

The time rate of change of position of a body in a particular direction. Linear velocity is velocity along a straight line, and its magnitude is commonly measured in such units as meters per second (m/s), feet per second (ft/s), and miles per hour (mi/h). Since both a magnitude and a direction are implied in a measurement of velocity, velocity is a directed or vector quantity, and to specify a given velocity completely, the direction must always be given. The magnitude only is called the speed. *See* SPEED.

Linear velocity. A body need not move in a straight line path to possess linear velocity. The instantaneous velocity of any point of a body undergoing circular motion is a vector quantity, such as v_1 or v_2 in **Fig. 1**. When a body is constrained to move along a curved path (**Fig. 2**), it possesses at any point an

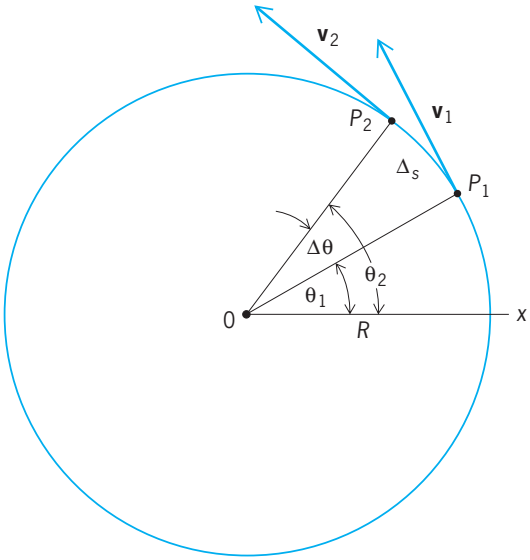

Fig. 1. Illustration of angular displacement, angular speed, and tangential velocity.

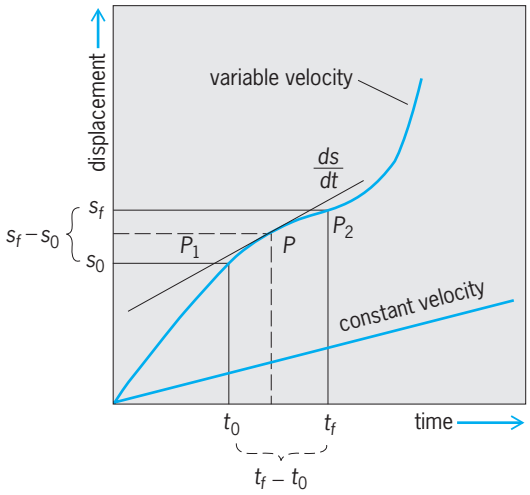

Fig. 2. Average velocity from P_1 to P_2 is $(s_f - s_0)/(t_f - t_0)$. The instantaneous velocity at point P is the limit of the ratio representing the average velocity as the interval approaches zero.

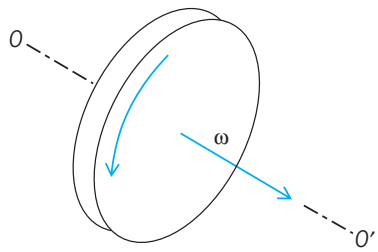

Fig. 3. Angular velocity shown as an axial vector.

instantaneous linear velocity in the direction of the tangent to the curve at that point. The average value of the linear velocity is defined as the ratio of the displacement to the elapsed time interval during which the displacement took place. The displacement of a body from an initial position s_0 to a final position s_f after time t is equal to $s_f - s_0$. The corresponding time interval is $t_f - t_0$. The magnitude of the average velocity is then given by Eq. (1), where Δs is

$$\bar{v} = \frac{\text{displacement}}{\text{elapsed time}} = \frac{s_f - s_0}{t_f - t_0} = \frac{\Delta s}{\Delta t} \qquad (1)$$

displacement and Δt is the corresponding elapsed time.

The magnitude of the instantaneous velocity v of a body is the limiting value of the foregoing ratio as the interval approaches zero. In the notation of calculus, Eq. (2), ds/dt is the instantaneous time rate of change

$$v = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = \frac{ds}{dt} \qquad (2)$$

of displacement (Fig. 2).

The velocity of a body, like its position, can only be specified relative to a particular frame of reference. Consequently, all velocities are relative. *See* RELATIVE MOTION.

Angular velocity. The representation of angular velocity ω as a vector is shown in **Fig. 3**. The vector is taken along the axis of spin. Its length is proportional to the angular speed and its direction is that in which a right-hand screw would move. If a body rotates simultaneously about two or more rectangular axes, the resultant angular velocity is the vector sum of the individual angular velocities. Thus, if a body rotates about an x axis with angular velocity ω_x, and simultaneously about a y axis with an angular velocity ω_y, the resultant angular velocity ω is the vector sum given by Eq. (3).

$$\omega = \omega_x + \omega_y \qquad (3)$$

It should be emphasized that whereas angular velocities are commutative in addition, that is, they may be added in any order, angular displacements are not commutative. *See* ROTATIONAL MOTION.

Angular displacement. Figure 1 represents a body rotating with circular motion about an axis through O perpendicular to the figure. Line OP_1 is the position of some radius in the body at a time t_1, with θ_1 being the angular displacement from a reference line. Line OP_2 is the position of the same radius at a later time t_2, with the angular displacement θ_2. Angular displacement may be measured in degrees, radians, or revolutions.

Angular speed. From Fig. 1, it is seen that the body has rotated through the angle $\Delta\theta = \theta_2 - \theta_1$ in the time $\Delta t = t_2 - t_1$. The average angular speed $\bar{\omega}$ is defined by $\bar{\omega} = \Delta\theta/\Delta t$, and the instantaneous angular speed $\omega = d\theta/dt$. Although it is customary in most scientific work to express angular speed in radians per second, it is common in engineering practice to use the units of revolutions per minute (rpm) or revolutions per second (rps).

Tangential velocity. When a particle rotates in a circular path of radius R through an angular distance $\Delta\theta$ in a time Δt, as in Fig. 1, it traverses a linear distance Δs. The average linear speed $\bar{v}$ is given by Eq. (4), since $\Delta s = R\Delta\theta$. Similarly, the instantaneous

$$\bar{v} = \frac{\Delta s}{\Delta t} = \frac{R\,\Delta\theta}{\Delta t} = \bar{\omega}R \qquad (4)$$

speed v is given by $v = \omega R$. The direction of this instantaneous speed is tangential to the circular path at the point in question. Any vector **v** drawn in this direction represents the tangential velocity.

Combined velocities. A body may have combined linear and angular motions, as is the case when the wheel of a moving automobile rolls along the ground with an angular velocity about its axle which moves with a linear velocity parallel to the pavement. In this case, a point on the rim of the tire describes a curved path called a cycloid. If a circular body rolls on the surface of a sphere, a point on the periphery of the rotating body describes a curve called an epicycloid. *See* CYCLOID; EPICYCLOID.

Rogers D. Rusk

Bibliography. J. D. Cutnell and K. W. Johnson, *Physics*, 4th ed., 1997; H. Goldstein, *Classical Mechanics*, 3d ed., 2001; D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, 6th ed., 2001; C. Kittel, W. D. Knight, and M. A. Ruderman, *Mechanics*, vol. 1, 2d ed., 1973.

Velocity analysis

A technique for the determination of the velocities of the parts of a machine or mechanism. Both graphical and analytical analyses of plane mechanisms will be discussed in this article, but of the several methods of each type available, only one of each type will be described. The graphical method will be discussed first, since the visualization which is an inherent part of the graphical analysis generally gives a better physical feel for the problem than most purely analytical methods. Analytical methods, however, are necessary for computer analyses. *See* VELOCITY.

Purposes. In a high-speed machine it is important that the inertia forces be determined. This requires an acceleration analysis of the machine, and the first step in an acceleration analysis usually is a velocity analysis. In the analysis of some machines, the velocity of a particular point in the machine may itself be the important thing to be determined in the analysis—for example, the cutting speed and return

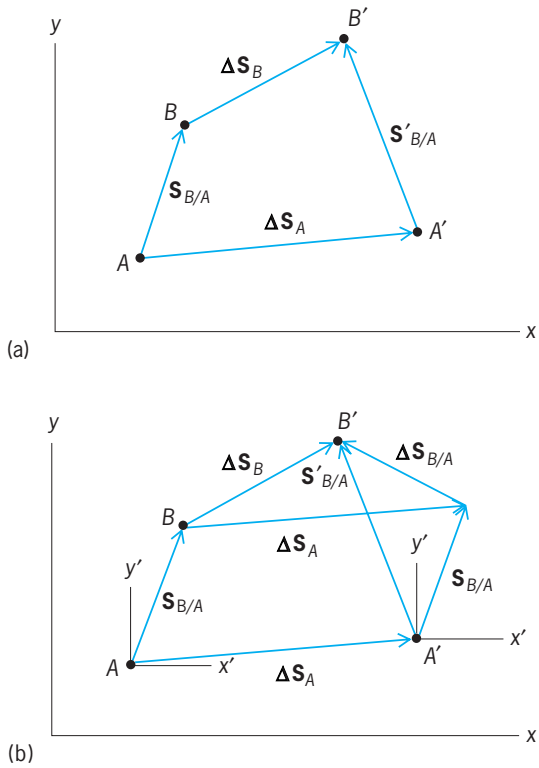

Fig. 1. Relative displacement. (*a*) Position vectors locating *B* with respect to *A* at beginning and end of a time increment. (*b*) Vectors for derivation of relative displacement equation.

speed of the cutting tool in a shaper, or the shuttle velocity in a textile machine.

Relative displacement. The method of relative velocities is widely used in velocity analysis, and the relative velocity equation is fundamental for this method of analysis. It is derived by first considering relative displacements. Referring to **Fig. 1***a*, let A and B be any two points moving in the stationary x-y coordinate system shown. Point B is located with respect to point A by the position vector $\mathbf{S}_{B/A}$. Suppose that during a time interval Δt, point B has the vector displacement $\Delta\mathbf{S}_B$ and point A the displacement $\Delta\mathbf{S}_A$. At the end of the time interval Δt, B is located with respect to A by the vector $\mathbf{S}'_{B/A}$. The vector change in $\mathbf{S}_{B/A}$ during the time interval Δt is shown in Fig. 1*b* as $\Delta\mathbf{S}_{B/A}$. From Fig. 1*b* it is evident that Eq. (1) holds, which is the relative displacement

$$\Delta\mathbf{S}_B = \Delta\mathbf{S}_A + \Delta\mathbf{S}_{B/A} \tag{1}$$

equation. It is further noted from Fig. 1*b* that $\Delta\mathbf{S}_{B/A}$ is the displacement of B as measured in the nonrotating x'-y' coordinate system which is attached to and moves with A.

Relative velocity. If the time interval Δt approaches zero, the displacements shown in Fig. 1*b* become infinitesimal and the velocity vectors representing the velocities of the points are proportional to the infinitesimal displacement vectors as shown in **Fig. 2**. From Fig. 2 the relative velocity equation (2) can

$$\mathbf{v}_B = \mathbf{v}_A + \mathbf{v}_{B/A} \tag{2}$$

be written. Equation (2) may also be derived mathematically by dividing Eq. (1) through by Δt and taking the limit as Δt approaches zero, as shown by Eq. (3) or (4). It must be clearly understood that $\mathbf{v}_{B/A}$

$$\lim_{\Delta t\to 0}\frac{\Delta\mathbf{S}_B}{\Delta t} = \lim_{\Delta t\to 0}\frac{\Delta\mathbf{S}_A}{\Delta t} + \lim_{\Delta t\to 0}\frac{\Delta\mathbf{S}_{B/A}}{\Delta t} \tag{3}$$

$$\mathbf{v}_B = \mathbf{v}_A + \mathbf{v}_{B/A} \tag{4}$$

in Eq. (2) represents the velocity of B as measured in a nonrotating coordinate system attached to A. It is usually called the velocity of B relative to A, although there has been some objection to this terminology.

Two points on a rigid body. Let body 2 of **Fig. 3** represent any mechanism link moving with general plane motion. Its angular velocity is ω (the vector representation of ω is a vector out of the paper). Equation (2) holds for any two points, and therefore holds for points A and B of the special case being considered. The relative velocity $\mathbf{v}_{B/A}$ is the velocity of B measured in the nonrotating x'-y' coordinate system attached to A. In this coordinate system B moves on a circular path with its center at A, and therefore its velocity as measured in this coordinate system has a magnitude equal to $r_{AB}\omega$ and a direction perpendicular to line AB. *See* RIGID-BODY DYNAMICS.

Coincident points on two links. Consider the mechanism of **Fig. 4***a* in which the pin at the end of link 2 is constrained to move in the slot of 3. Point P_2 is the center of the pin, and point P_3 is the point on link 3 coincident with P_2. There is no such actual physical point, but an extension built onto link 3 and overlapping

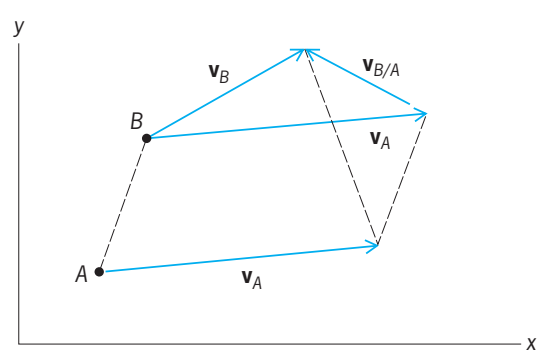

Fig. 2. Vectors for derivation of the relative velocity equation.

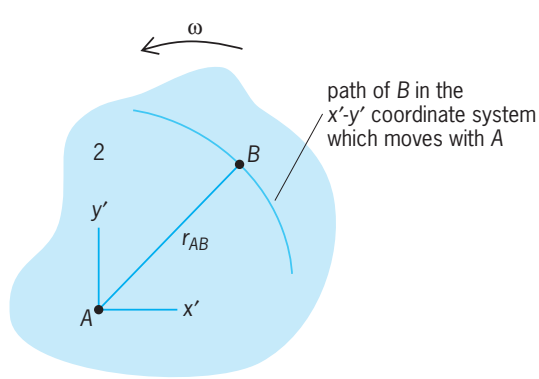

Fig. 3. Rigid body 2 moving with general plane motion.

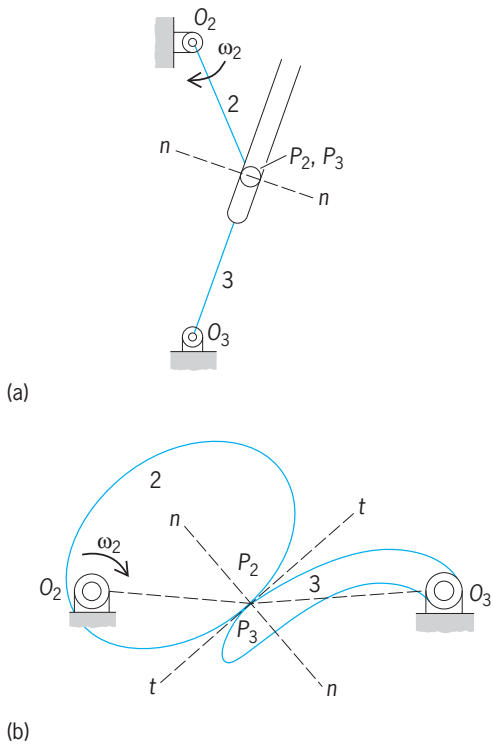

Fig. 4. Mechanisms *a* and *b* for which the relative velocity equation is applied to coincident points on different links.

the pin can be imagined. Equation (2) as applied to the coincident points P_2 and P_3 is Eq. (5). Assume

$$\mathbf{v}_{P_3} = \mathbf{v}_{P_2} + \mathbf{v}_{P_3/P_2} \tag{5}$$

the magnitude of $\mathbf{v}_{P2}$ is known. The directions of $\mathbf{v}_{P2}$ and $\mathbf{v}_{P3}$ are known. To solve a single-vector equation, there can be no more than two unknowns, so that the direction of $\mathbf{v}_{P3}/P_2$ must also be known. There can be no relative motion of points P_2 and P_3 in the direction of line *n-n* normal to the slot, so that the only relative motion possible is along the slot. Therefore the direction of $\mathbf{v}_{P3}/P_2$ is along the slot. Similarly for the direct-contact mechanism of Fig. 4*b*, the relative velocity $\mathbf{v}_{P3}/P_2$ must be along the common tangent of the contacting surfaces, line *t-t′*. *See* LINKAGE (MECHANISM).

Application example. The use of the relative velocity equation will be illustrated in the velocity analysis of a slider-crank mechanism. Slider-crank mechanisms are used in a wide variety of machines. Perhaps the most familiar example is the crank, connecting rod, and piston mechanism of the internal combustion engine. A skeleton drawing of such a mechanism is shown in **Fig. 5***a*. The moving links are numbered 2, 3, and 4, with link 1 being the fixed frame. Key points are identified with letters. It is assumed that the angular velocity of crank 2 is known. The magnitude of the velocity of point A is given by $v_A = r_{O_2A}\omega_2$. A vector representing this velocity is shown in Fig. 5*a* originating at point A, and also is laid off to scale from an origin O_v in Fig. 5*b*. The relative velocity equation, $\mathbf{v}_B = \mathbf{v}_A + \mathbf{v}_{B/A}$, is next applied to points A and B. Point B is constrained to move along a horizontal straight line so that the direction of $\mathbf{v}_B$ is known. Vector $\mathbf{v}_A$ is completely known; $\mathbf{v}_{B/A}$ is perpendicular to line AB, since A and B are two points on the same rigid body (line 3). A velocity polygon (triangle) can be drawn as in Fig. 5*b* to solve for the unknown quantities of the relative velocity equation, the magnitudes of $\mathbf{v}_B$ and $\mathbf{v}_{B/A}$. These magnitudes can be scaled directly from the polygon. Usually it would be necessary to make an analysis of this type for a number of positions of the mechanism in its motion cycle. Other mechanisms can be analyzed graphically for velocities in a similar manner. *See* SLIDER-CRANK MECHANISM.

Use of vector mathematics. In the graphical method of velocity analysis just discussed, the geometry of the mechanism is known in each phase of its motion cycle from the drawing of the mechanism in each position. In an analytical velocity analysis, the geometry is usually determined analytically; that is, a position analysis of the mechanism is performed by using trigonometry, vector mathematics, or some other analytical method as the first step in the velocity analysis.

Position analysis. One analytical method for the position analysis of mechanisms, developed by M. A. Chace, makes use of vector mathematics. It consists essentially of solving vector triangles containing two unknowns. For example, in **Fig. 6** an offset slider-crank mechanism is shown with appropriate vectors placed on the mechanism. Assuming the position of crank 2 is known as the input to the mechanism, vectors **P** and **Q** are known and may be added together to obtain vector **C**. Vectors **C**, **r**, and **s** form a vector triangle containing two unknowns, the direction of vector **r** and the magnitude of vector **s**. That is, unit vector $\hat{r}$ and s are the unknowns (carets are used in this article to designate unit vectors). The plane vector equation $\mathbf{C} + \mathbf{s} + \mathbf{r} = 0$ (or $\mathbf{C} + s\hat{s} + r\hat{r} = 0$) must be solved for the unknowns. Chace

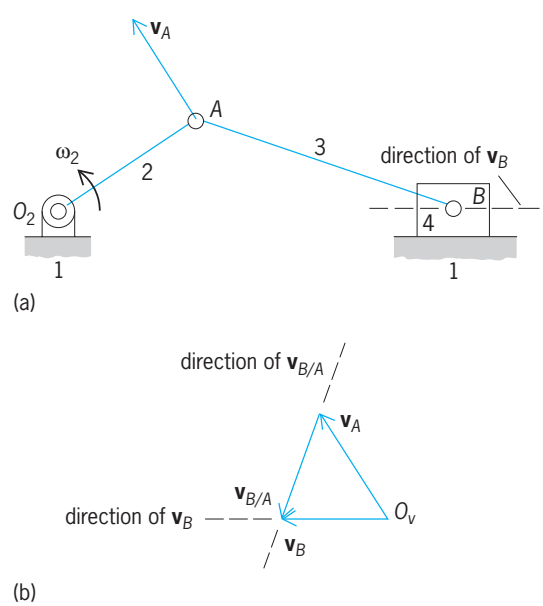

Fig. 5. Velocity analysis. (*a*) Slider-crank mechanism to be analyzed. (*b*) Velocity polygon.

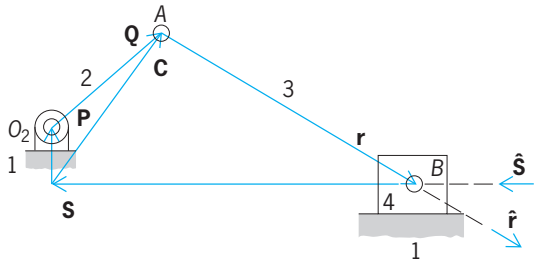

Fig. 6. Offset slider-crank mechanism showing vector triangle to be solved for position analysis of the mechanism.

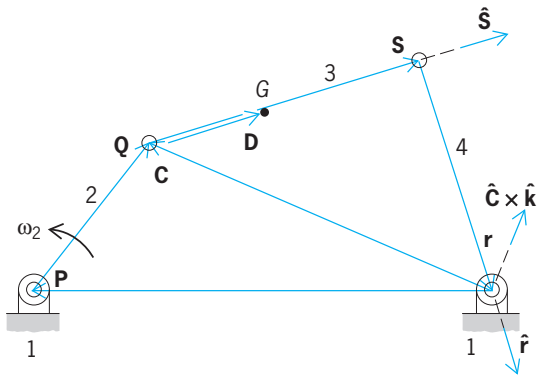

Fig. 7. Four-bar linkage showing vector triangle to be solved for the position analysis of the mechanism.

developed solutions for the plane vector equation for various cases using vector mathematics. Vector **C** is assumed always to be the known vector, and case 1 is the situation where two magnitudes, s and r, are unknown. Chace's equations for the vectors **s** and **r** for this case are Eqs. (6) and (7), where $\hat{k}$ is the unit

$$\mathbf{s} = \frac{\mathbf{C} \cdot (\hat{\mathbf{r}} \times \hat{\mathbf{k}})}{\hat{\mathbf{r}} \times (\hat{\mathbf{s}} \times \hat{\mathbf{k}})} \hat{\mathbf{s}} \tag{6}$$

$$\mathbf{r} = \frac{\mathbf{C} \cdot (\hat{\mathbf{s}} \times \hat{\mathbf{k}})}{\hat{\mathbf{s}} \cdot (\hat{\mathbf{r}} \times \hat{\mathbf{k}})} \hat{\mathbf{r}} \tag{7}$$

vector in the z direction or out of the paper.

Case 2 is the situation illustrated in Fig. 6 where there is one unknown magnitude, s, and one unknown direction, $\hat{\mathbf{r}}$, in the vector triangle. Chace's equations for the vectors **r** and **s** for this case are Eqs. (8) and (9). The lower signs on the radicals

$$\mathbf{r} = -[\mathbf{C} \cdot (\hat{\mathbf{s}} \times \hat{\mathbf{k}})](\hat{\mathbf{s}} \times \hat{\mathbf{k}}) \pm \sqrt{r^2 - [\mathbf{C} \cdot (\hat{\mathbf{s}} \times \hat{\mathbf{k}})]^2}\,\hat{\mathbf{s}} \tag{8}$$

$$\mathbf{s} = \left(-\mathbf{C} \cdot \hat{\mathbf{s}} \mp \sqrt{r^2 - [\mathbf{C} \cdot (\hat{\mathbf{s}} \times \hat{\mathbf{k}})]^2}\right) \hat{\mathbf{s}} \tag{9}$$

would be chosen for the analysis of this mechanism because **r** has a negative component in the $\hat{s}$ direction for the complete mechanism cycle.

Case 3 is the situation illustrated in **Fig. 7**. Vector **P** connecting the fixed pivots is a known vector and can be expressed in the form $\mathbf{P} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$. The position of the input crank 2 is assumed to be known, so that vector **Q** is known and can be expressed in the same form as vector **P**. Vectors **P** and **Q** can be added together to get vector **C**. Vectors **C**, **s**, and **r** form a vector triangle in which the unknowns are unit vectors $\hat{\mathbf{s}}$ and $\hat{\mathbf{r}}$. Chace's solutions for vectors **r** and **s** for this case are Eqs. (10) and (11). The upper

$$\mathbf{r} = \mp\sqrt{r^2 - \left(\frac{r^2 - s^2 + C^2}{2C}\right)^2}\,(\hat{\mathbf{C}} \times \hat{\mathbf{k}}) - \left(\frac{r^2 - s^2 + C^2}{2C}\right)\hat{\mathbf{C}} \tag{10}$$

$$\mathbf{s} = \pm\sqrt{r^2 - \left(\frac{r^2 - s^2 + C^2}{2C}\right)^2}\,(\hat{\mathbf{C}} \times \hat{\mathbf{k}}) + \left(\frac{r^2 - s^2 + C^2}{2C} - C\right)\hat{\mathbf{C}} \tag{11}$$

signs on the radicals would be selected for the analysis of the mechanism of Fig. 7, because **s** has a positive component in the $\hat{\mathbf{C}} \times \hat{\mathbf{k}}$ direction and **r** has a negative component in that direction. *See* FOUR-BAR LINKAGE.

Time derivative of a vector. The time derivative of a vector originating from a fixed origin such as the one shown in **Fig. 8** is the velocity of the point at the tip of the vector; that is, Eq. (12) holds.

$$\mathbf{v}_P = \lim_{\Delta t \to 0} \frac{\Delta \mathbf{r}}{\Delta t} = \frac{d\mathbf{r}}{dt} = \dot{\mathbf{r}} \tag{12}$$

Time derivative of a fixed-length vector. If a vector **r** in a rotating body is of fixed length and originates on the axis of rotation of the body as shown in **Fig. 9**, its time derivative is $\omega \times \mathbf{r}$, since the magnitude of $\omega \times \mathbf{r}$ is $\omega r \sin\phi$, and it is evident from Fig. 9 that the velocity of point P (the time derivative of **r**) is $\omega r \sin\phi$ in the direction of $\omega \times \mathbf{r}$. It may also be shown that the time derivative of any fixed-length vector **r** in a body is equal to $\omega \times \mathbf{r}$ even though the vector **r** does not originate on the axis of rotation of the body. In this case, however, the derivative of the vector is not the velocity of the point at the tip of the vector, but the velocity of the point at the tip of the vector relative to the point at the origin or tail of the vector.

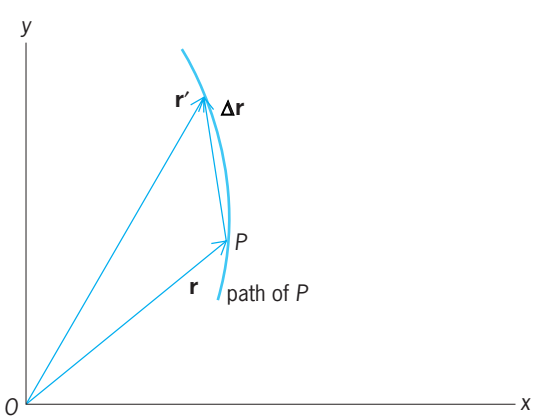

Fig. 8. Changing vector originating from a fixed origin.

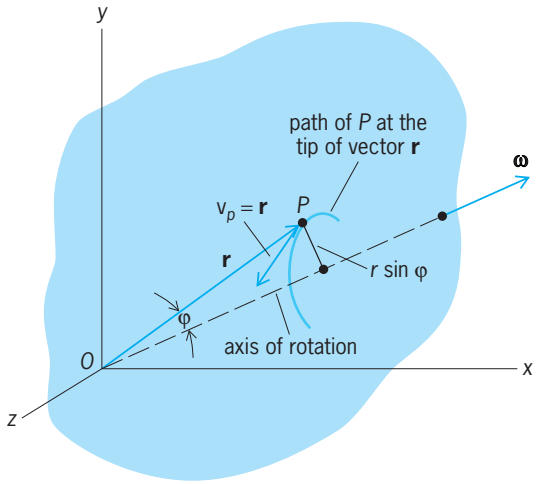

Fig. 9. Fixed-length vector in a body which originates on the axis of rotation of the body; $\mathbf{v}_p = \mathbf{r} = \omega \times \mathbf{r}$.

Four-bar mechanism. For the four-bar mechanism shown in Fig. 7, the vector loop equation (13) may

$$\mathbf{P} + \mathbf{Q} + \mathbf{s} + \mathbf{r} = 0 \qquad (13)$$

be written. Differentiating Eq. (13) with respect to time yields Eq. (14), since vectors **Q**, **s**, and **r** are all fixed-length vectors. In a plane motion mechanism the directions of ω_3 and ω_4 are known (direction $\hat{\mathbf{k}}$). The unknowns ω_3 and ω_4 may be determined by taking the dot product of each vector of Eq. (14) with

$$\omega_2 \times Q + \omega_3 \times s + \omega_4 \times r = 0 \qquad (14)$$

$\hat{\mathbf{r}}$ and $\hat{\mathbf{s}}$, respectively, reducing Eq. (14) in each case to a scalar equation with a single unknown. The solutions thus determined for ω_3 and ω_4 are Eqs. (15) and (16). With these angular velocities known, the

$$\omega_3 = \frac{-\omega_2(\hat{\mathbf{k}} \times \mathbf{Q}) \cdot \hat{\mathbf{r}}}{s(\hat{\mathbf{k}} \times \hat{\mathbf{s}}) \cdot \hat{\mathbf{r}}} \qquad (15)$$

$$\omega_4 = \frac{-\omega_2(\hat{\mathbf{k}} \times \mathbf{Q}) \cdot \hat{\mathbf{s}}}{r(\hat{\mathbf{k}} \times \hat{\mathbf{r}}) \cdot \hat{\mathbf{s}}} \qquad (16)$$

velocity of any point on the mechanism can be found. For example, the velocity of point G is given by Eq. (17).

$$\mathbf{v}_G = \omega_2 \times Q + \omega_3 \times D \qquad (17)$$

Computer subprograms for Chace solutions to the plane vector equation and for all the various vector operations have been written, thus making the position and velocity analyses (as well as the acceleration analysis) of mechanisms by computer using Chace's method quite convenient. Computer analyses are particularly useful when complex mechanisms must be analyzed in many positions of their cycle of motion. *See* STRAIGHT-LINE MECHANISM; VECTOR (MATHEMATICS). James C. Wolford

Bibliography. S. Doughty, *Mechanics of Machines*, 1988; B. Paul, *Kinematics and Dynamics of Planar Machinery*, 1979; J. Shigley and C. R. Mischke, *Machine Design Fundamentals*, 1989; J. Shigley and J. Uicker, *Theory of Machines and Mechanisms*, 2d ed., 1994.

Veneer

A thin sheet of wood of uniform thickness produced by peeling, slicing, or sawing. Depending on the manner of production and the portion of wood from which a veneer is made, the grain may be flat, vertical, or biased. Most veneer is rotary-cut from a bolt of wood, called a flitch, centered in the chucks of a lathe, which may have a capacity for logs from 2 to 16 ft (0.6 to 4.8 m) long, about 8 ft (2.4 m) being a common capacity. A nose bar bears against the flitch parallel to the center line of the lathe and a knife, also extending nearly the length of the lathe, peels off the veneer. Knives near the ends of the flitch cut the edges of the veneer.

The peeling blade advances to maintain the thickness of the veneer at the initial setting, which may be as thin as $^1/_{64}$ in. (0.4 mm), more typically in the range from $^1/_{32}$ to $^1/_{20}$ in. (0.75 to 1.25 mm), or as thick as $^3/_8$ in. (9.5 mm). The outer surface of the veneer as it emerges from under the nose bar is smooth; it is called the tight side. However, the inner surface tends to form lathe checks as the blade curls it away from the log; hence it is called the loose side.

The veneer is clipped to approximate size, dried sufficiently to prevent attack from molds and fungi (moisture typically 15% or less by weight), and bundled until needed. Each bundle contains the sheets cut from a single flitch, stacked in order as cut, and labeled as to total content. The user can then match veneers for figured panels.

Veneers cut from selected hardwoods, burls, crotches, and stumps are used for facings on furniture and the interior decoration of buildings. The art of veneering was practiced in Egypt at least as early as the fifteenth century B.C. Until the twentieth century most veneer was produced by sawing; for the hardest woods, such as ebony and oak, veneer is still sawed. During the eighteenth century veneering (in which a single relatively large veneer covers an area) and marquetry (in which numerous small pieces of wood and other materials form a pattern) reached a peak of craftsmanship in the decoration of fine furniture. During the early years of the industrial revolution, ornate inlays and highly figured veneers applied on poorly fabricated furniture created an impression that veneering was a cheap substitute for solid construction. With improved methods of drying and gluing, veneering regained its position in furniture manufacture and has been extended to wall paneling and other decorative purposes. Other uses of veneer include plate separators in storage batteries, boxes for fruit and vegetables, drums for cheeses, and crates, hampers, and baskets for transportation and storage. Spanish cedar, sliced to $^1/_{100}$ in. (0.25 mm), is used to wrap fine cigars. Most veneer is used in plywood panels. *See* PLYWOOD; WOOD PRODUCTS. Frank H. Rockett

Ventilation

The supplying of air motion in a space by circulation or by moving air through the space. Ventilation may be produced by any combination of natural or mechanical supply and exhaust. Such systems may include partial treatment such as heating, humidity control, filtering, or purification, and, in some cases, evaporative cooling. More complete treatment of the air is generally called air conditioning. *See* AIR CONDITIONING.

Natural ventilation. Natural ventilation may be provided by wind force, convection, or a combination of the two. Although largely supplanted by mechanical ventilation and air conditioning, natural ventilation still is widely used in homes, schools, and commercial and industrial buildings. It is effective and economical in areas of prevailing winds and for high industrial buildings that have hot equipment which provides the motivating convective force.

Wind-force ventilation may be provided by direct force, such as wind blowing in one side of a building and out the other side or through a monitor. This is commonly known as cross ventilation. Because of the friction and velocity losses incurred through building openings, only about 25–60% of the wind velocity is available for ventilation, depending upon whether the wind direction is perpendicular or oblique to the building openings. The resulting negative pressures on the downwind side of the building (if open to the outside) and the sizes and locations of the exhaust openings all materially affect the flow of air.

Airflow around an object creates a negative pressure on the downwind side. This principle is used to advantage in the design of a large number of building ventilators and monitors to provide basic ventilation, to assist existing convective ventilation, or to prevent backdraft down into a building. **Figure 1** illustrates the round, rotating head, and continuous roof-type ventilators. Because of the many complex forces involved, airflow capacities cannot be calculated but must be determined by testing.

The force for convection ventilation is created by the difference in weight of two air columns at different temperatures; the heavy, cool column attempts to displace the hot, light column. The pressure p exerted by a fluid column varies as the height h and density w of the fluid; that is, $p = hw$, expressed in consistent units. The basic equation for such flow is $V = \sqrt{2gh}$, again in consistent units for velocity V, height h, and gravity g. From these relations and because temperature has a direct relationship to density, the equation below may be used to estimate con-

$$Q = 9.4A\sqrt{h(t_i - t_o)}$$

vective flow (**Fig. 2**). Here Q = air passing through an opening by convective force, ft^3/min; A = free (net) area of the opening, ft^2 (the inlet areas are assumed to equal the exhaust openings); h = height from inlets to outlets, ft; t_i = average temperature of indoor air column in height h, °F; t_o = temperature of outdoor air, °F; and 9.4 = constant of proportionality, including value of 65% for effectiveness of openings.

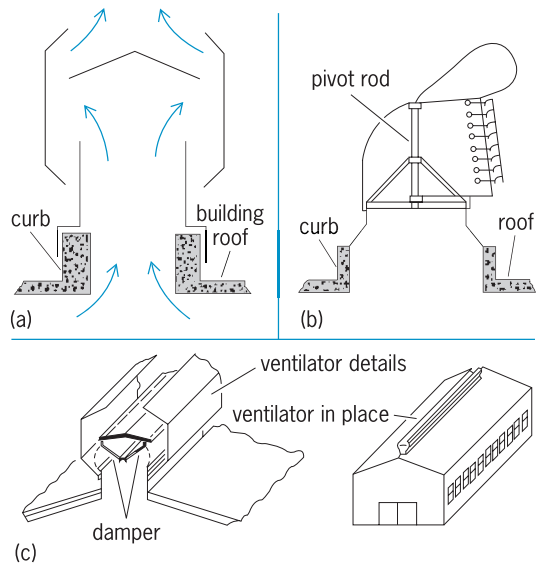

Fig. 1. Roof exhausts for natural ventilation. (*a*) Cross section of round ventilator and (*b*) of rotating-head ventilator. (*c*) Continuous roof ventilator.

This equation indicates that convective velocities tend to be low. Consequently, high buildings having large exhaust (and inlet) openings and high internal heat loads are the most practical application for convection ventilation. The best flow would be obtained by a well-shaped hole in the roof, but, because weather protection must also be provided, it is customary to exhaust through roof monitor or baffle-type exhaust ventilators as in Figs. 1 and 2 for small volumes and **Fig. 3** for large volumes.

Mechanical ventilation. Mechanical supply ventilation may be of the central type consisting of a central fan system with distributing ducts serving a large space or a number of spaces, or of the unitary type (Fig. 3) with little or no ductwork, serving a single space or a portion of a large space. Both types are employed for schools and for commercial and industrial applications. *See* DUCTED FAN.

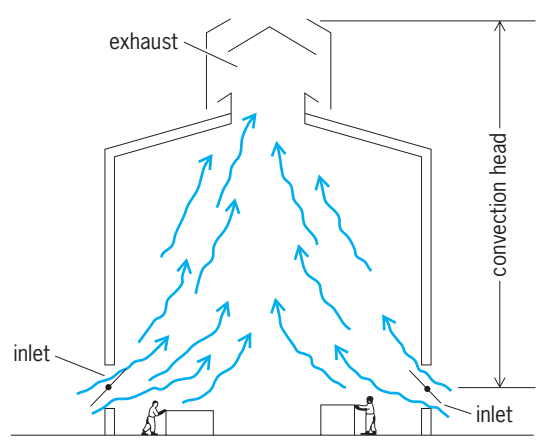

Fig. 2. Building cross section illustrating motivating convection head (height) for natural ventilation.

Central system assemblies may be custom-built, factory-prefabricated for job site assembly, or factory-assembled. Ventilating units are factory-assembled with rare exceptions. The assemblies consist of a fan and usually include air filters and air heaters of the finned-tube type. The fans may be of the propeller, axial flow, or centrifugal type. Central systems require the last two fan types because of the higher static pressures usually encountered in the distributing ducts. *See* FAN.

The mechanically powered roof ventilator is a practical source of low-cost ventilation. For summer (nontempered) supply, such units consist of fans and weather hoods, or dampers (**Fig. 4**). For all-year ventilation, heaters or other equipment may be added.

Outside air connections are generally provided for all systems. Outside air is needed in controlled quantities to remove odors and to replace air exhausted from the various building spaces and equipment. The inlets are located to minimize the intake of fumes, dust, organic materials, and pollens. Because it is never possible to find completely clean air, ventilation air filters are usually provided in the system casings in all except a few industrial or summer relief applications. This prevents clogging and poor heat transfer for the air heaters and helps to reduce the pollen and dust in the occupied areas served by the systems.

Air distribution. Duct systems to distribute and disperse the air are required for all but the small unitary systems to avoid short-circuiting and to provide adequate ventilation to all parts of the space served. Such distribution permits desirable air movement in the space without undesirable draft and temperature stratification. Ventilation systems generally serve as conveyors for adding or removing heat and humidity. The distribution system must be adequate for this purpose also. The amount of air circulated is important because too small a volume requires uncomfortably high temperatures for heating or results in high building temperature for heat removal ventilation. Similar problems occur with humidity because of the limited amounts of moisture which can be conveyed without condensation problems. If a system is oversized, it becomes unnecessarily expensive. In addition, it may be difficult to discharge the air in the space within acceptable velocity limits.

Outlet grilles and diffusers of the rectangular, square, and round type are provided to further control and distribute the supply of air within the selected throw (blow length) for each outlet. Any air column acts as a pump and will entrain many times the primary volume of air. The outlets are designed and selected to obtain maximum entrainment (and mixing) because this greatly reduces air motion and temperature stratification, making possible greater comfort for the room occupants.

Exhaust ventilation systems. Exhaust ventilation is required to remove odors, fumes, dust, and heat from an enclosed occupied space. Such exhaust may be of the natural variety previously described or may be mechanical by means of roof or wall exhaust fans or mechanical exhaust systems. The mechanical systems may have minimal ductwork or none at all, or may be provided with extensive ductwork which is used to collect localized hot air, gases, fumes, or dust from process operations. Where it is possible to do so, the process operations are enclosed or hooded to provide maximum collection efficiency with the minimum requirement of exhaust air.

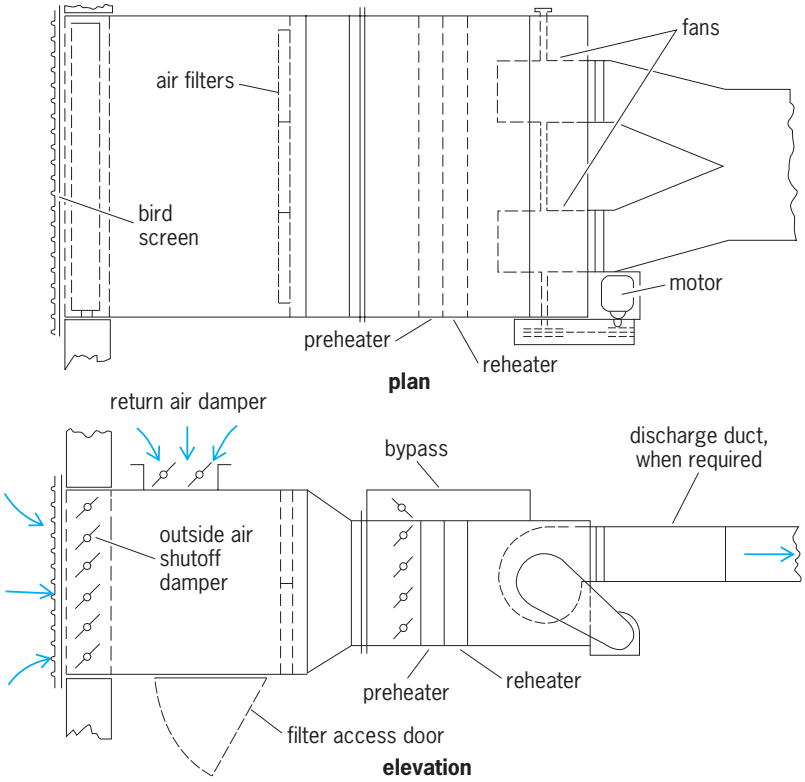

Fig. 3. Factory-assembled ventilating unit.

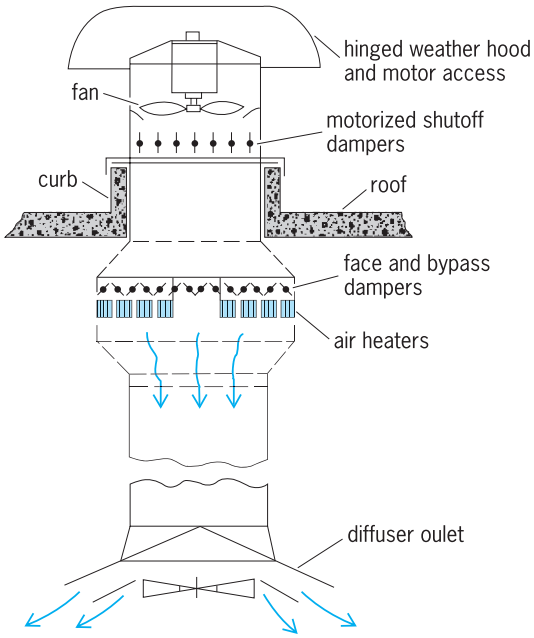

Fig. 4. Mechanically powered roof air supply unit. For summer ventilation, the portion shown in solid lines is used. Heater and dampers are for winter use. Diffuser outlet improves air distribution into ventilated space.

Because of the possibilities of recirculated or external air pollution, it is customary to remove dust and fumes where practical or where required by means of ordinary ventilation filters, more efficient washers or centrifugal collectors, or chemical scrubbers.

Where dust is conveyed in the exhaust ducts, the velocities must be adequate to lift and move the dust particles. Velocities of 3000–6000 ft/min (914–1828 m/min) are required for this purpose. Lower velocities are tenable for fume removal, but corrosion protection must be provided by selection of duct and equipment materials. The round duct is used in most dust and fume systems because of its lower friction and because of its better dust-handling characteristics.

Fans for ordinary ventilation exhaust or heat removal may be similar to supply fans. Fume and dust exhaust fans must be more rugged and are frequently of the radial-blade (paddle-wheel) type for this reason. Axial flow and conventional centrifugal fans are also used in these applications. John H. Clarke

Venus

The second planet in distance from the Sun. Venus has been called "Earth's twin" because it is similar to Earth in gross characteristics such as mass, radius, and density (see **table**). In other ways Venus is apparently different. Its atmospheric mass is almost a hundred times that of the Earth; its atmosphere is mostly carbon dioxide instead of nitrogen and oxygen; an extensive cloud layer of concentrated sulfuric acid is present; its surface temperature is an unbearable 867°F (464°C); and it rotates with a period of 243 days, and from east to west, clockwise as we look down from the north, in the opposite sense of most other planets. Some of these differences are due more to alternate evolutionary paths of the two planets than to totally different initial conditions.

Appearance. To the naked eye, Venus is the brightest starlike object in the sky. It is usually visible during the night either soon after sunset or close to sunrise. It can sometimes be seen during the daytime. As observed through a pair of binoculars or a telescope, Venus exhibits a crescentlike or gibbous appearance. Similar to the Moon, Venus ranges over a full set of phases from a "new moon" to a "full moon." This phase variation is caused by a changing fraction of the Sun-illuminated hemisphere facing toward the Earth. At new moon, only the dark nighttime side is seen, while at full moon all of the daytime hemisphere is seen. Venus undergoes a complete set of phase changes over its synodic period of 584 days. *See* PHASE (ASTRONOMY).

Clouds. The light seen coming from Venus is entirely due to sunlight that is reflected from a dense cloud layer whose top is located about 45 mi (70 km) above the surface and whose bottom lies within 30 mi (50 km) of the surface. In contrast to the Earth's approximately 50% cloud cover, the clouds of Venus are present over the entire planet. In yellow or red light they present a uniform appearance. However, the clouds show a banded and spotted pattern when viewed in ultraviolet light (**Fig. 1**). These ultraviolet markings provide information about atmospheric motions.

The clouds of Venus consist of a large number of tiny particles, about 1 micrometer in size, that are made of a water solution of concentrated sulfuric acid. Such a composition may at first seem very surprising when compared with the water clouds of the Earth's lower atmosphere. However, sulfuric acid particles are the dominant type of particles in the Earth's upper atmosphere, although the amount there is much less than the amount present in Venus's atmosphere. In the case of both planets, sulfuric acid is produced primarily from sulfur-containing gases that combine with water vapor and oxygen-containing gases. Compositional measurements made from the United States *Pioneer Venus* Sounder probe, which descended through Venus's atmosphere on December 9, 1978, show that sulfur dioxide is the principal sulfur-containing gas in Venus's atmosphere. Sulfur dioxide is also the major gas species injected into the Earth's stratosphere by volcanic explosions. Such injections cause a large, but temporary, increase in the amount of sulfuric acid there.

Atmospheric composition. By far, the chief gas species of Venus's atmosphere is carbon dioxide, which makes up 96% of the atmospheric molecules, while nitrogen accounts for almost all the remainder. Trace amounts of sulfur dioxide (150 parts per million), water vapor (20 ppm), carbon monoxide (17 ppm), argon (70 ppm), helium (12 ppm), neon (7 ppm), hydrogen chloride (0.4 ppm), and hydrogen fluoride (0.005 ppm) are present in the lower atmosphere, with the concentration of the first two of these declining dramatically near the cloud tops due to the formation of new sulfuric acid there. Chemical transformations also occur in the deeper portions of the atmosphere, aided by the high temperatures there. For example, at altitudes below about 25 mi (40 km), carbon monoxide is gradually converted into carbonyl sulfide, a gas containing carbon, oxygen, and sulfur atoms.

Carbon dioxide. In contrast to the dominance of carbon dioxide in Venus's atmosphere. The Earth's atmosphere consists mostly of nitrogen and oxygen, with carbon dioxide being present at a level of only 340 ppm. In part, this difference may stem more from temperature differences than from intrinsic differences. Over the lifetime of the Earth, an amount of carbon dioxide comparable to that in Venus's atmosphere was vented out of the Earth's hot interior. The outgassed carbon dioxide remained in the atmosphere for only a short time. Almost all of it participated with rain in dissolving land rocks. Rivers carried the dissolved rock and carbon dioxide into the oceans, where they subsequently precipitated to form carbonate rocks, such as limestone. Venus's surface is much too hot for oceans of water to be present, and hence its atmosphere has been able to retain essentially all of the carbon dioxide vented from its interior. *See* ATMOSPHERIC CHEMISTRY.

Characteristics of Venus*

Characteristics	Values
Mass	0.815 Earth's mass
Equatorial radius	0.95 Earth's equatorial radius or 3760.4 mi (6051.8 km)
Mean density	0.95 Earth's value or 5.24 g/cm^3
Range of topography	9 mi (15 km) [Earth is 12 mi (20 km)]
Gravity	0.907 Earth's gravity
Mean distance from Sun	0.723 Earth's distance
Perihelion	66,780,000 mi (107,480,000 km), 0.718 Earth's distance
Aphelion	67,690,000 mi (108,940,000 km), 0.728 Earth's distance
Orbital period	0.615 of an Earth year or 224.7 Earth days
Orbital eccentricity	0.4 Earth's value or 0.0068
Orbital inclination to Earth's orbital plane	3.39°
Sidereal rotational period	243.686 Earth days with respect to the stars
Length of day	116.75 Earth days with respect to the Sun
Atmospheric surface pressure	92 bars (1334 lb/in.^{2}), 90 times Earth's atmospheric pressure
Average surface temperature	867°F (464°C) [Earth is 59°F (15°C)]
Surface wind speed	0.7–2.2 mi/h (0.3–1.0 m/s)
Surface atmospheric composition (by volume)	96.5% carbon dioxide (CO_2), 3.5% nitrogen (N_2); minor constituents (ppm): sulfur dioxide (SO_2): 150; argon (Ar): 70; water (H_2O): 20; carbon monoxide (CO): 17; helium (He): 12; neon (Ne): 7

*Some of these numbers depend on assumptions as to when and where they were observed and some variation will be seen in the literature.

Rare gases. There are two varieties of rare gases found in planetary atmospheres: those that were derived from the gas cloud (solar nebula) from which the planets formed (primitive rare gases), and those that were produced from the radioactive decay of certain elements, such as potassium, in the interior of the planets. A fundamental finding about the composition of Venus's atmosphere is the detection of much more primitive argon and neon than in Earth's atmosphere. Furthermore, Mars's atmosphere has an even smaller amount of these rare gases than does the Earth's. These differences in the abundances of rare gases among the atmospheres of Venus, Earth, and Mars may, in part, be due to sizable differences in the rates at which their earliest atmospheres were lost. Such losses may have been caused by the blowoff of portions of these atmospheres by massive, high-velocity stray bodies that were particularly abundant in the first several hundred million years of these planets' lifetimes. The blowoff was powered by hot rock vapor generated when these bodies collided with the planets' surfaces. Mars was especially vulnerable to atmospheric blowoff because of its low mass and hence its low surface gravity. Conceivably, the Earth lost a good fraction of its earliest atmosphere when a body with a mass comparable to that of Mars hit it. Such a giant impact is considered by many as the most likely way by which the Earth's Moon formed. Since Venus and Earth have comparable amounts of carbon dioxide and nitrogen in their atmospheres and the rocks of their interiors, these atmospheric components may have been added at a later stage, perhaps from volatile-rich bodies that came from the outer solar system. *See* EARTH; MARS; MOON; SOLAR SYSTEM.

Water vapor. The amount of water vapor in Venus's atmosphere is much less (about 100,000 times) than the amount of water in the Earth's oceans. Since it is unlikely that Venus was initially endowed with so much less water than the Earth, Venus probably lost almost all of its original water over its lifetime. The loss of water from a planet may be determined primarily by the amount that is in the upper atmosphere, the stratosphere, where solar ultraviolet radiation decomposes water vapor molecules into hydrogen and oxygen. The light gas hydrogen can eventually escape the planet's gravity and be lost to space, while the leftover oxygen can combine with other gases, such as carbon monoxide, or with iron at the planet's surface. Because Venus is closer to the Sun than the Earth, its lower atmosphere was hotter,

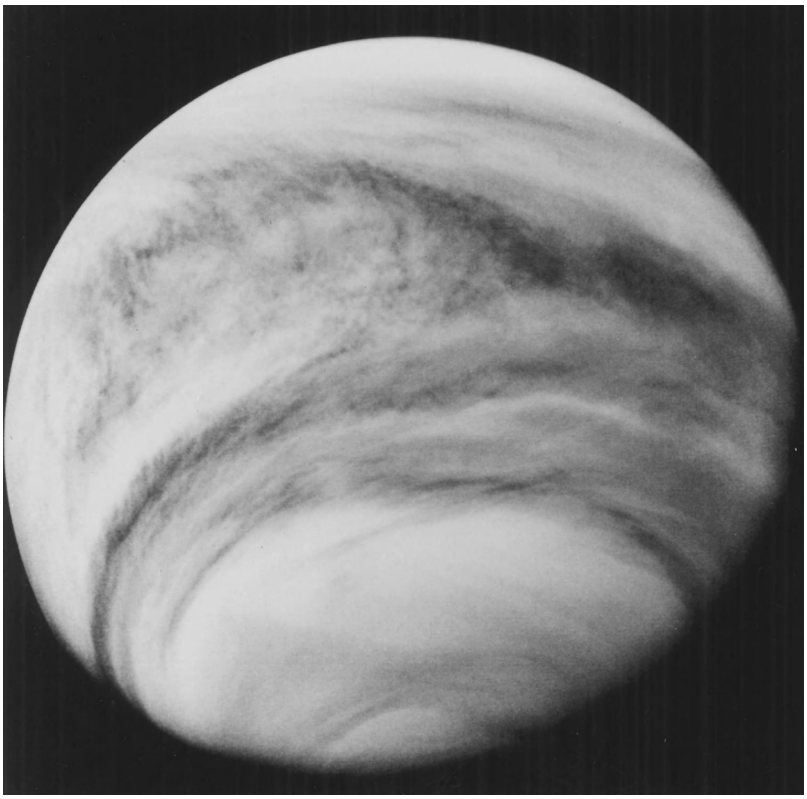

Fig. 1. Clouds of Venus photographed in ultraviolet light from the *Pioneer Venus* spacecraft. The south pole is near the bottom of the image.

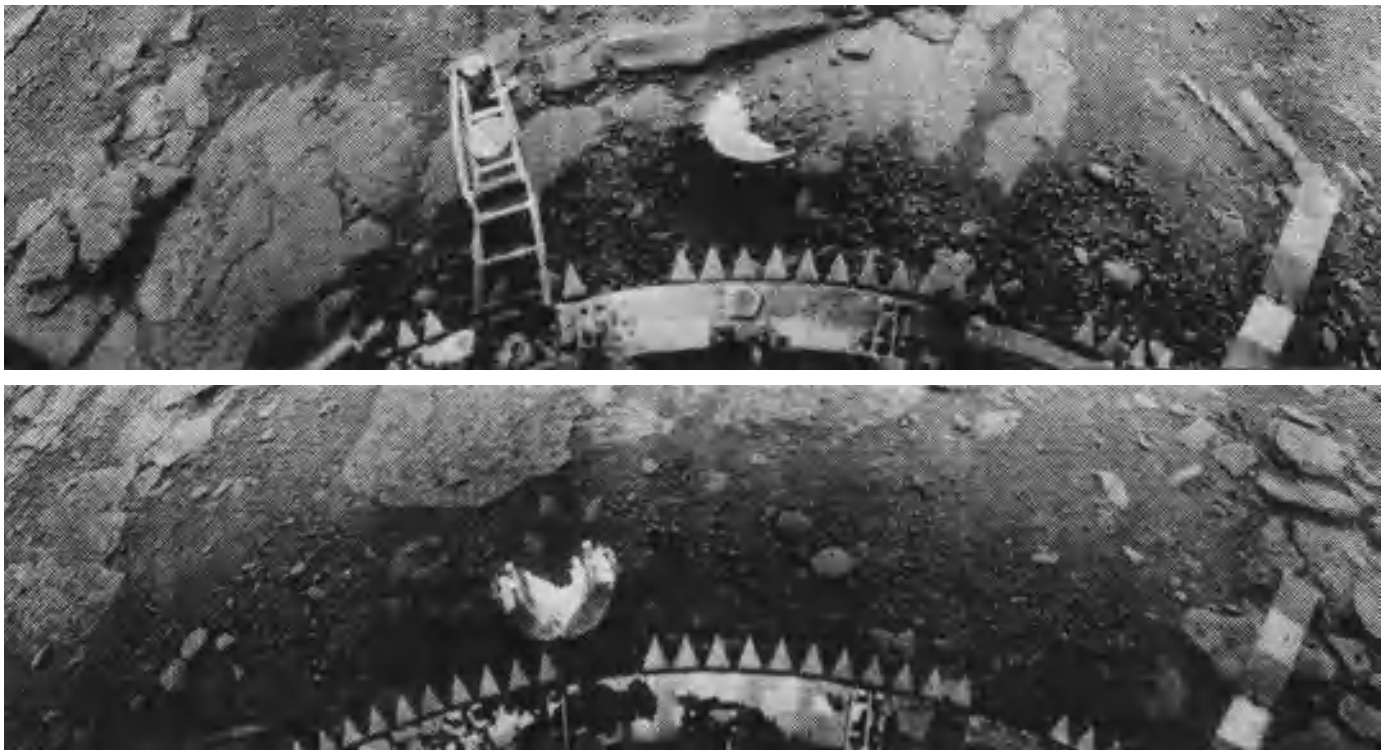

Fig. 2. Panoramic view, in two frames, of the surface of Venus obtained by the Soviet *Venera 13* spacecraft shortly after it landed. Parts of the spacecraft and apparatus extending from it can be seen at the bottoms of the frames.

water vapor was more abundant at early times there, and hence much more of the water vapor was able to penetrate into its stratosphere than was the case for the Earth. Consequently, Venus could have lost much more water over its lifetime than did the Earth. Some confirmation of this viewpoint has been given by the finding of one of the. *Pioneer Venus* experiments that there is about 100 times as much heavy water, or water containing deuterium, on Venus as on Earth. Such an enrichment of deuterium is expected since light hydrogen can more easily escape to space than heavy hydrogen.

Temperature. By detecting long-wavelength heat radiation produced at the surface, radio telescopes first showed that the surface temperature was 850°F (730 K), that is, over 600°F (350 K) higher than the boiling point of water. This result was confirmed by direct temperature measurements of the atmosphere made from Soviet and United States spacecraft that descended through Venus's atmosphere. The atmospheric temperature has a relatively cool value of −10°F (250 K), about 45°F (25 K) below the freezing point of water, near the top of the cloud layer, which is at a pressure of about 1/20 that at the Earth's surface. The temperature gradually increases with decreasing altitude until it reaches about 850°F (730 K) at the surface, where the pressure is 90 times that at the Earth's surface. Just as on Earth, surface temperatures on Venus decrease at higher elevations. At the top of Maxwell Montes, about 7.5 mi (12 km) higher than the average elevation, the temperature could be only about 650°F (620 K).

The high value of Venus's surface temperature is not due to its being closer to the Sun than the Earth. Beacuse its cloud layer reflects to space about 75% of the incident sunlight, Venus actually absorbs less solar energy than does the Earth. Rather, the high temperature is the result of a very efficient greenhouse effect that allows a small but significant fraction of the incident sunlight to penetrate to the surface (about 2.5% according to Soviet and United States spacecraft measurements), but prevents all except a negligible fraction of the heat generated by the surface from escaping directly to space. The thermal energy produced by the surface and hot lower atmosphere, which occurs at infrared wavelengths, is very effectively absorbed by the carbon dioxide, water vapor, sulfur dioxide, and sulfuric acid particles of the atmosphere. However, these materials are poor absorbers at visible wavelengths, where most of the solar energy lies. The greenhouse effect raises Venus's surface temperature by almost 900°F (500 K), but causes only a modest 60°F (35 K) rise in the surface temperature of the Earth. This difference is due to the Earth's atmosphere being partially transparent at some infrared wavelengths, thus permitting some surface heat to escape to space. *See* GREENHOUSE EFFECT.

Meteorology. The atmosphere near the cloud tops is moving with a jet-stream-like velocity of about 225 mi/h (100 m/s) from east to west in the direction of Venus's rotation. Although the solid surface rotates around in 243 days (clockwise as one looks down from the north, opposite the rotation of Earth), the clouds are observed to move around the planet

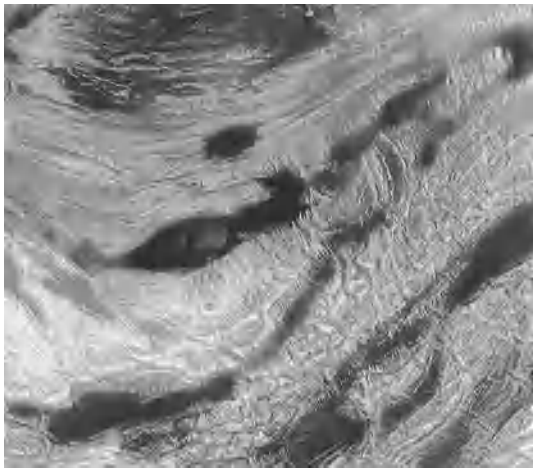

Fig. 3. *Magellan* image of the central and northern parts of the highland region Ovda Regio (north at top). Ridges running approximately east-west (horizontally) have been interpreted as resulting from crustal shortening (compression). Ridges are cut by through-going fractures and graben (fault-bounded valleys), indicating that an episode of northeast-southwest extension followed the compressional event. The youngest event was flooding of low-lying areas by smooth (radar-dark) lava flows. (*Jet Propulsion Laboratory*; *NASA*)

in the same direction with a period of about 4 days. This motion is called superrotation. Only Saturn's moon Titan shares this phenomenon with Venus. There is a small component of cloud motion of about 10 mi/h (5 m/s) from the equator toward the pole. In contrast to the situation for the Earth where only a small portion of the atmosphere moves at jet-stream speeds, the entire cloud top region on Venus moves at these large speeds. The wind speed for the most part gradually, but in a few places sharply, decreases with declining altitude and achieves values of a modest few meters per second within 6 mi (10 km) of the surface. Winds at the surface are also a few meters per second. However, wind streaks are seen on the surface, in the *Magellan* radar images, in the vicinity of the more recent impact craters, indicating that the formation of these craters produced temporary stormlike atmospheric winds near the surface. *See* SATURN.

Like the winds on the Earth, the winds on Venus are produced ultimately by differences in the amount of solar energy absorbed by different areas of the planet, such as areas at different latitudes. The large wind speeds near the cloud tops may be the result of Venus's atmosphere being very deep. The transport of momentum by the mean circulation and the eddies from the deep dense portions near the surface to the much less dense regions near the cloud tops translates sluggish motions into fast ones. At midlatitudes and in the polar regions of the Earth, the east-west wind speeds are determined primarily by a balance between variations of pressure with latitude and the Coriolis force due to the rotation of the Earth's surface, which is shared by its atmosphere. But Venus's surface rotates too slowly for the Coriolis force to be important in its atmosphere. Rather, centrifugal forces (analogous to those on a merry-go-round) owing to the rotating motions of the winds themselves around the planet may provide the balance to the pressure gradient forces. *See* ATMOSPHERIC GENERAL CIRCULATION.

Because Venus's atmosphere is massive, atmospheric motions are very effective in reducing the horizontal variations of temperature in its lower atmosphere. In the region of the clouds and at lower altitudes, temperature variations occur mostly in the north-south direction, with the equator being only a few degrees warmer than the poles close to the surface and some tens of degrees warmer within the cloud region. Above the clouds, the atmosphere is actually somewhat warmer at the poles than near the equator because of the effects of heat transported by atmospheric motions. As Venus's axis of rotation is almost exactly perpendicular to its orbital plane, its climate has little seasonal variability.

Interior. The similar mean densities of Venus and the Earth imply that Venus is made of rocks similar to those that make up the Earth. However, because Venus formed closer to the Sun, in perhaps a somewhat warmer environment, it may have initially contained a smaller amount of sulfur and water-bearing compounds. Such an environmental difference would probably not have significantly affected Venus's content of the long-lived radioactive elements uranium, potassium, and thorium. Over the lifetime of the Earth, and presumably Venus, the decay of these elements may have generated enough heat to cause these planets to become chemically

Fig. 4. *Magellan* image of a portion of Leda Planitia. The smooth (radar-dark) plains were formed by volcanic lava flows that covered the region. These flows embayed the older, radar-bright, highly fractured or chaotic highlands, sometimes called tessera, that rise out of the plains in the upper-left or northwest part of the image, as well as the circular ring structure in the lower left, which is probably an impact crater. Subsequent volcanism produced radar-bright flows in the upper right. (*Jet Propulsion Laboratory; NASA*)

differentiated, as free iron melted and sank toward their centers. Also, both planets probably formed "hot" because of gravitational energy released in bringing chunks of rock and planetesimals together to form them. In this case, they may have undergone substantial differentiation in their early histories, with an accompanying release of much of their atmospheric gases. Thus, Venus's interior may be qualitatively similar to that of the Earth in having a central iron core, a middle mantle made of rocks rich in silicon, oxygen, iron, and magnesium, and a thin outer crust containing rocks enriched in silicon in comparison with the rocks of the mantle. However, in contrast to the situation for the Earth, Venus's core may now be either entirely solid or entirely liquid, which could account for the absence of a detectable magnetic field. *See* EARTH INTERIOR.

Rotation. In contrast to the Earth and almost all other planets, Venus rotates in the opposite direction with respect to its orbital motion about the Sun. It rotates so slowly that there are only two sunrises and sunsets per Venus year. Tides raised in the body of the planet by the Sun may have greatly reduced the rate of rotation from an intially large value, similar to the Earth's, to its present low value. Because Venus's atmosphere is so massive, tides raised by the Sun in the atmosphere may have also been important, with its current rate of rotation being determined chiefly by a balance between the oppositely directed torques arising from these two types of solar tides. Venus currently presents the same face to the Earth at the times of closest approach. This suggests that the tidal forces that were exerted by the Earth also played a role and helped lock Venus in its present rotational state.

Life. The current high surface temperature and acid constitution of the clouds preclude the existence of living organisms like those that inhabit the Earth. But Venus may have had a hot ocean of water in its early history when the Sun put out less energy than it does now. If so, it is unclear whether life could have arisen then. Any early life would have been des- troyed when Venus lost its oceans and achieved its current high surface temperature.

James B. Pollack; R. Stephen Saunders

Spacecraft. Venus has been more intensely explored by spacecraft than any other planet. Some 31 (9 of them unsuccessful) United States and Soviet spacecraft have been sent to Venus. The Soviet Union *Venera* and *Vega* series of spacecraft and the United States *Mariner* and *Pioneer* series provided data on the composition of the surface and atmosphere, and the dynamics of the upper atmosphere. *Venera* landers transmitted back the first images from the surface of another planet (**Fig. 2**). The surface of Venus was seen in different places to be rocky with little granular material. The United States *Magellan* by the end of 1992 had mapped nearly 98% of the surface. Venus is the last of the inner planets to be mapped in detail but is the most important planet for understanding how the Earth evolved as a life-sustaining habitat. *See* SPACE PROBE.

Surface. *Magellan*'s mapping mission revealed a unique global volcanic and tectonic style on Venus. Broad volcanic plains make up about 85% of the surface of Venus. The rest is tectonically deformed, higher-standing terrains with complex systems of folds and faults. Regional tectonism is evident in the widespread compressional and extensional deformation of much of the surface material (**Fig. 3**). Venus apparently has had a dynamic mantle that has driven crustal warping, which may be ongoing. However, while various regions of the planet show evidence of motion, no evidence of Earth-like plate tectonics has been found. *See* PLATE TECTONICS.

Long, narrow troughs are seen in many areas where the crust has ruptured; these linear rift zones are associated with extensive broad, domical rises and shield-volcano complexes. Examples are the Sif-Gula region, Beta Regio, and Atla Regio. Large areas of the planet are covered by lava that flowed from volcanic vents (**Fig. 4**). Volcanism on Venus occurs globally, unlike Earth where volcanism is mostly restricted to linear zones that define plate boundaries.

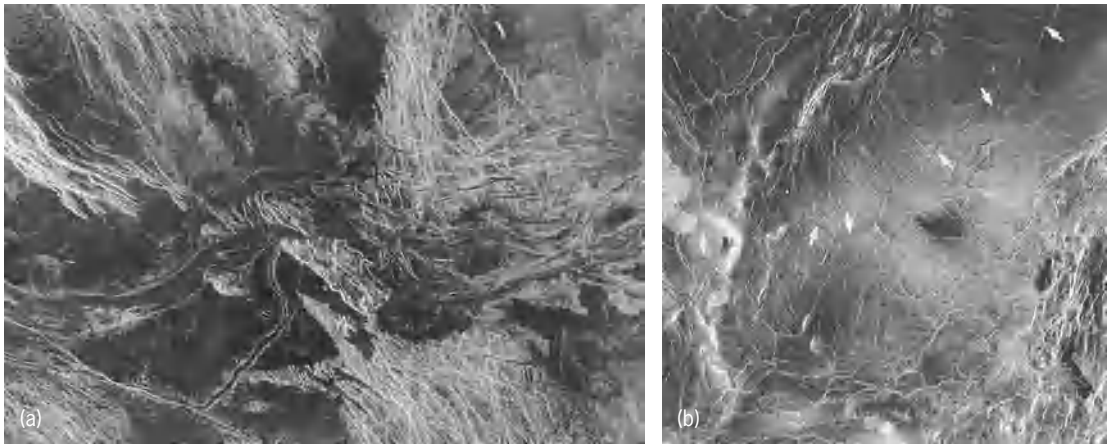

Fig. 5. ***Magellan*** **images of channels on Venus. (*a*) Part of a channel in the Lada Terra region, 750 mi (1200 km) long and 12 mi (20 km) wide, displaying numerous streamlined structures that attest to the very fluid lavas responsible for carving the channel. (*b*) A 360-mi (600-km) segment of the Hildr Channel (indicated by arrows), whose length of 4200 mi (6800 km) makes it the longest known channel in the solar system. (*Jet Propulsion Laboratory; NASA*)**

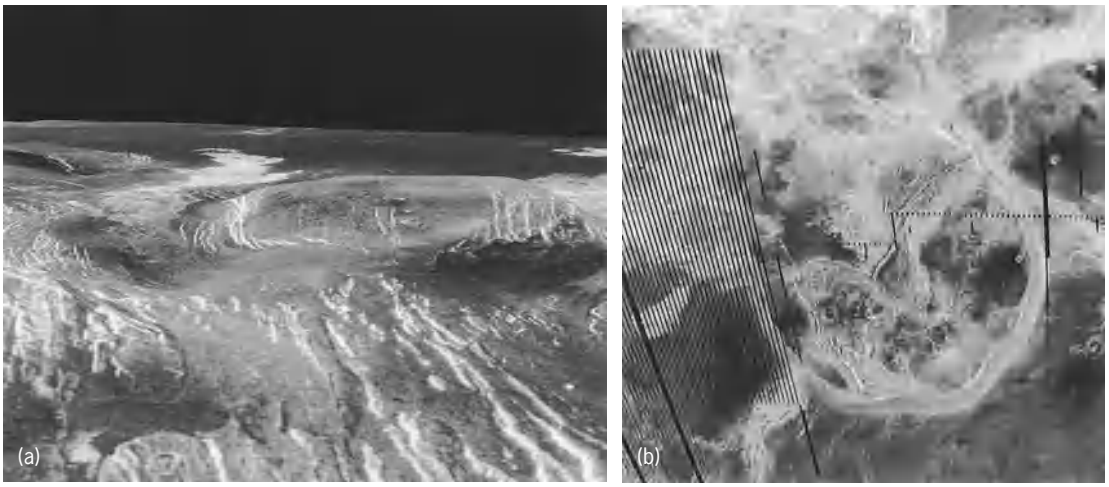

Fig. 6. ***Magellan*** **images of coronae on Venus. (*a*) Computer-simulated view of the corona Idem-Kuva, 60 mi (97 km) in diameter. Lava flows extend for hundreds of miles across the fractured plains in the background. (*b*) Artemis Chasma, whose diameter of 1300 mi (2100 km) makes it the largest corona identified. The interior contains complex systems of fractures, numerous flows, and small volcanoes, and at least two impact craters. The margin forms a steep trough with raised rims. (*Jet Propulsion Laboratory; NASA*)**

The planet also has some unexplained surface features, including long channels meandering across the plains (**Fig. 5***a*). One channel is 4200 mi (6800 km) long (Fig. 5*b*), which is longer than the Nile river. Its origin remains unknown. Many of the channels are clearly formed by lava, but even under the high temperature (850°F or 730 K) of the surface, most known volcanic lava compositions should solidify before they could flow far in open channels. A more exotic lava such as sulfur or a carbonate may have formed some of the longer channels. Another unique landform is the scat- tered flat, circular volcanoes that resemble giant pancakes. These range in diameter from about 9 to 50 mi (15 to 80 km) and are a few hundred meters thick.

Coronae are classified as volcanic-tectonic structures (**Fig. 6**). These features are oval to circular and range in diameter from 60 to 1300 mi (100 to 2100 km). Most are around 190 mi (300 km) in diameter. They have low relief, up to 1–2 mi (2–3 km) high. The margins are generally irregular concentric ridges and troughs. The interiors may be quite irregular, including various tectonic and volcanic edifices. Many coronae have associated volcanic flows. A mantle plume model has been postulated for the formation of coronae, in which the rising lower-density plume causes doming, with contemporaneous or subsequent relaxation to form the bounding ridges and troughs.

The higher elevations of Venus, Maxwell Mons for example, have distinctly higher-than-average radar reflectivity and low emissivity. One explanation for this behavior is that, at slightly lower temperature and pressure in the more elevated regions, a metallic mineral is stable and distributed throughout the surface material to create higher reflectivity. Possible minerals could be pyrrhotite, pyrite, or magnetite.

Impact craters. Impact craters are much less abundant on Venus than on the Moon or Mars. Approximately 950 impact craters have been identified in *Magellan* images (**Fig. 7**). Most appear to be unmodified by erosion. Based on the abundance of impact craters, the average age of the surface appears to be about 300 to 500 million years old. This age would imply a low rate of surface modification by erosional, or even by volcanic and tectonic, processes, and suggests that parts of the planet have undergone relatively recent volcanic or tectonic activity. On a global scale, impact craters are randomly distributed, allowing the determination of the mean surface age of less than 500 million years, but the small number of total crates provides little age information at a local scale.

Asteroids and comets that collide with Venus should have typical velocities of about 12 mi/s

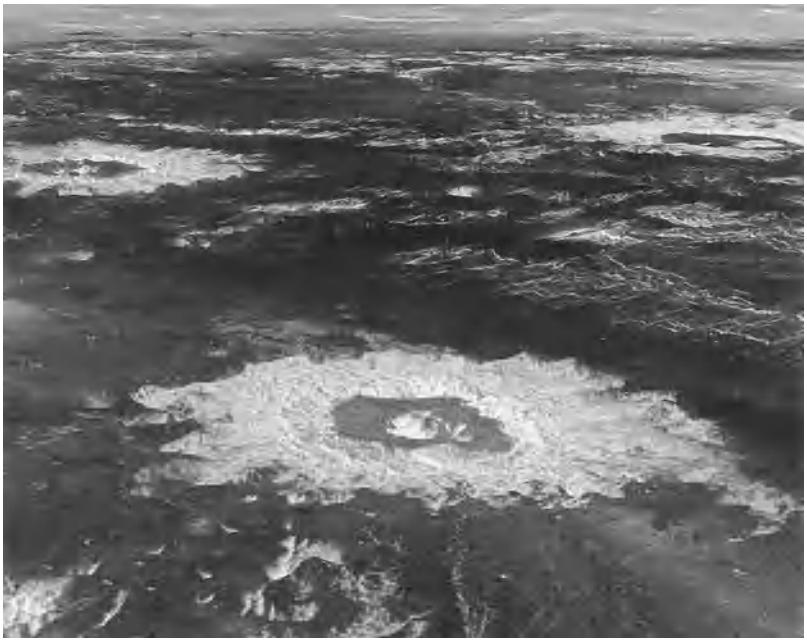

Fig. 7. Computer-simulated perspective view, based on data from *Magellan* spacecraft, showing three impact craters in the northwestern portion of Lavinia Planitia with diameters ranging from 23 to 39 mi (37 to 63 km). (*Jet Propulsion Laboratory; NASA*)

(20 km/s). Venus's thick atmosphere plays a significant role in meteorite impacts. Because the atmosphere is so dense, only craters larger than 2 mi (3 km) in diameter can form, except in crater clusters where large projectiles apparently broke up before impact, peppering an area with smaller fragments. Ordinarily, a projectile that would produce a crater smaller than about 2 mi (3 km) in diameter would vaporize or break up in its transit to the surface.

Wind features. It was of major scientific interest to discover wind activity on Venus. *Magellan*'s radar saw in many areas abundant bright and dark wind streaks near topographic barriers such as small volcanic ridges. Wind streaks have been found to be most frequent near large impact craters. The craters may have provided the material that is moved by the wind. Also, the impact process itself may have produced some streaks. Some of the larger impact craters have large parabolic features hundreds of miles long. Many of the wind streaks are in the vicinity of these parabolas. The parabolas may be bright or dark, and all are open to the west. This orientation is consistent with some interaction with the east-to-west winds tens of miles above the surface of Venus, which reach speeds of 125 mi/h (200 km/h). Particulate crater ejecta may have been distributed downwind. Alternatively, the turbulent atmospheric disturbance created by the impact may have been carried to the west by the general atmospheric circulation, creating deposition or erosion patterns in the lee of the event. Further study of the orientations of some wind streaks may provide a better understanding of Venus's global wind patterns.

History of volcanism. Currently active volcanism has not been detected on Venus. One theory postulates cessation of a violent episode of global volcanism 300 to 500 million years ago. Subsequently there would have been less volcanic activity, forming the large shield volcanoes such as Maat Mons that are scattered over the planet. It is estimated that the activity represented by the younger volcanoes is less than 5% of the older plains flows. An alternative hypothesis holds that there is an equilibrium between the formation of impact craters and the volcanic outpourings that erase them, so that the surface always appears to be about 500 million years old. *See* PLANET.

R. Stephen Saunders

Bibliography. J. K. Beatty, C. C. Petersen, and A. Chaikin (eds.), *The New Solar System*, 4th ed., 1999; S. W. Bougher, D. M. Hunten, and R. J. Phillips (eds.), *Venus II: Geology, Geophysics, Atmosphere, and Solar Wind Environment*, University of Arizona Press, 1998; P. Cattermole, *Venus: A New Geology*, 1994; H. S. Cooper, Jr., *The Evening Star: Venus Observed*, 1994; D. J. Eicher, *Magellan* scores at Venus, *Astronomy*, 19(1):34–42, January 1991; D. H. Grinspoon, *Venus Revealed*, 1997; *Magellan* reports, *J. Geophys. Res.*, 92:13,059-13,690, Aug. 25, 1992, and 92:15,921-16,382, Oct. 25, 1992; *Magellan* reports, *Science*, 252:247–312, Apr. 12, 1991; *Pioneer Venus* results, *Science*, 203:743–808, Feb. 23, 1979, and 205:41–121, July 6, 1979.

Venus' flytrap

Dionaea muscipula, an insectivorous plant of North and South Carolina (see **illus.**). The two halves of a leaf blade can move as if they were hinged along the midrib and, swinging upward and inward, the two surfaces come together. Any insect alighting on a leaf triggers this sensitive motor mechanism, and is caught between the closing halves of the leaf blade. In this trap, the insect is slowly digested by enzymes secreted by cells in the leaf. *See* INSECTIVOROUS PLANTS; NEPENTHALES; SECRETORY STRUCTURES (PLANT).

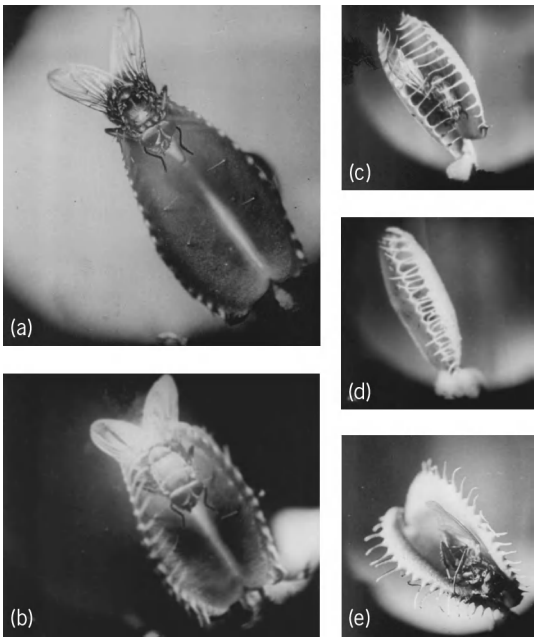

Stages *a–e* in capture and digestion of fly by leaf of Venus' flytrap. (*General Biological Supply House*)

Perry D. Strausbaugh; Earl L. Core

Vermiculite

A group of minerals common in some soils and clays and belonging to the family of minerals called layer silicates. Species within the vermiculite group are denoted as either dioctahedral vermiculite or trioctahedral vermiculite, with two or three octahedral cation sites occupied per formula unit, respectively.

Structure. Trioctahedral vermiculite is a 2:1 layer silicate with a fundamental unit similar to that of mica. An octahedral sheet forms the basis of the layer. This sheet is composed of octahedra (a polygon with six corners and eight faces) with magnesium (Mg) or similar-sized cation [iron (Fe), aluminum (Al)] as the central cation and six surrounding oxygens (O) and hydroxyls (OH). To form the 2:1 layer, the octahedral sheet is sandwiched between two opposing tetrahedral sheets. Tetrahedra consist of T cations [T = silicon (Si) or aluminum] surrounded by four oxygen atoms. *See* MICA; PHLOGOPITE.

Unlike micas, however, vermiculites have an additional octahedral sheet between the 2:1 layers. This sheet is partially occupied by magnesium [or lesser amounts of sodium (Na) or calcium (Ca)] and water (H_2O) molecules, but both occupy definite sites. The sheet has not been determined precisely. The stacking of layers in vermiculite, although potentially very complex, always occurs as a two-layer form; semirandom stacking sequences are the norm. The structure resembles chlorite also, but the interlayer octahedral sheet in vermiculite is not as fully occupied as in chlorites. *See* CHLORITE; HYDROGEN BOND.

Properties. Perfect basal cleavage develops by the layerlike structure. Density varies but is near 2.4 g/cm^3; hardness on the Mohs scale is near 1.5; and luster is pearly. Thin sheets deform easily and may be yellow to brown. Rapid heating causes some vermiculites to undergo a "popcorn" effect, where interlayer water explosively discharges to cause exfoliation, whereas other vermiculites exfoliate more slowly into wormlike threads. *See* HARDNESS SCALES.

Occurrence. Trioctahedral vermiculites form as an alteration product of phlogopite, biotite, or chlorite, whereas dioctahedral vermiculites are generally thought to form from illite. Trioctahedral vermiculites are common in well-drained soils in humid regions wherever the parent material contains trioctahedral micas. Interstratification of vermiculite and mica is common. *See* BIOTITE; ILLITE; MUSCOVITE.

Uses. Heat-treated and expanded vermiculite is used as an insulator in construction. Mixed with plaster and cement, vermiculite is used to make lightweight versions of these materials. Vermiculite is also useful as an absorbent for some environmentally hazardous liquids. *See* CLAY, COMMERCIAL; CLAY MINERALS; SILICATE MINERALS. Stephen Guggenheim

Bibliography. S. W. Bailey (ed.), *Hydrous Phyllosilicates* (*Exclusive of the Micas*), 1988.

Vernalization

The induction in plants of the competence or ripeness to flower by the influence of cold, that is, temperatures below the optimal temperature for growth. Vernalization thus concerns the first of the three phases of flower formation in plants. In the second stage, for which a certain photoperiod frequently is required, flowers are initiated. In the third stage flowers are unfolded.

Phenomena of vernalization were reported sporadically in the nineteenth century; however, intensive investigations in this field started with G. Gassner in 1910. The cold requirement for the induction of ripeness to flower is known for a large number of perennial, winter annual, and late-flowering summer annual plants. Examples are winter cereals and other grasses (strains of *Lolium, Bromus, Alopecurus, Phleum, Cynosurus*, and *Dactylis*), beets and cabbages from the genera *Beta* and *Brassica*, and species of *Lupinus, Vicia, Oenothera, Streptocarpus, Daucus, Carum, Apium, Petroselinum, Echium, Campanula, Cichorium, Scorzonera*, and *Helianthus*. These plants are prevented from floral initiation before winter by the cold requirement. Most physiological work on vernalization has been done with winter rye, henbane (*Hyoscyamus niger*), *Chrysanthemum morifolium*, and *Arabidopsis thaliana*.

Normally, the shoot apex of either a plant or an embryo must be exposed to cold for vernalization to occur. Root cuttings of *Lunaria biennis* and of *Cichorium intybus* and leaf cuttings of *Lunaria*, which produce shoot meristems in their adventitious roots, may be vernalized successfully, too, in certain cases even before the formation of shoot meristems. In unifoliate *Streptocarpus* species, vernalization takes place in the macrocotyledon.

Many plants can be vernalized in the seed stage. However, some biennial and perennial plants like *Hyoscyamus niger, Dianthus barbatus*, and some strains of *Brassica oleracea* do not respond to cold treatment during a juvenile phase, which is usually 1–3 months. In certain races of winter wheat, lettuce, and *Arabidopsis* the vernalizibility diminishes during germination, or the first days or weeks thereafter, and then increases again (that is, the vernalizing effect of a given cold treatment is relatively great in seeds, smaller in young seedlings, and greater again in older plants). Vernalization of seeds is only possible after imbibition; between 30 g of water for caryopses of some cereals and 100 g for certain papilionaceous seeds must be added to 100 g of seeds. Caryopses of cereals may even be vernalized prior to maturity, when still upon the mother plant.

Cold duration and temperature. The minimum duration of a cold treatment that will barely accelerate flower formation is 4 days in Petkus winter rye, 5 days in a Japanese radish variety, and 7 days in biennial *Hyoscyamus niger*. The optimal duration of vernalization varies. It is less than 3 weeks for the radish and those late summer cereals requiring a small amount of cold and more than 100 days for strains of *Hyoscyamus niger, Lactuca sativa*, and *Lupinus albus*.

In some instances, such as winter rye, *Hyoscyamus*, and *Arabidopsis*, the longer a vernalization treatment lasts, the lower is the optimal temperature of vernalization. It then approaches temperatures just above zero. The upper limit of vernalizing temperatures is known exactly only in some plants that have high cold requirements, such as a variety of *Allium cepa* (59–68°F or 15–20°C), sugarbeet (50–52°F or 10–11°C), *Hyoscyamus* (63°F or 17°C), and kohlrabi (about 50°F or 10°C). Temperatures below zero (down to 18°F or −8°C) have also proved effective in some cases, such as caryopses of winter cereals and *Arabidopsis* seeds, but to a rather small degree.

Most cold-requiring plants are also frost-resistant; their frost resistance, however, usually diminishes during vernalization for reasons which are not completely known. *See* COLD HARDINESS (PLANT).

Effect of photoperiodism. The vernalization effect is measured by time in days to anthesis, to ear emergence (in grasses), and to unfolding of the first flower and by the number of leaves on the main shoot axis

below the insertion of the first flower. The effect is influenced by photoperiod, light intensity, or both before, during, and after vernalization. The vernalization effect is increased when plants are exposed prior to cold treatment to low light intensity (*Arabidopsis thaliana*), to short-day conditions (*Daucus carota* and grasses like *Poa pratensis, Bromus inermis*, and *Dactylis glomerata*), or to long-day conditions (short-day plant *Chrysanthemum morifolium* var. Sunbeam). Illumination during vernalization has opposite effects in different species. In many long-day plants a vernalization effect is manifested only when vernalization is followed by long-day exposure (*Daucus carota*, many winter cereals and other grasses, and strains of *Arabidopsis* requiring much cold); in others, the greatest relative vernalization effect occurs under short-day conditions (*Beta vulgaris, Agrostemma githago*, and *Arabidopsis* strains requiring less cold). *See* PHOTOPERIODISM.

Devernalization. The vernalization effect can be annulled (devernalization) immediately after cold treatment by high temperatures (several days at 77–104°F or 25–40°C) if the preceding cold exposure did not last too long (winter rye and *Dianthus barbatus*), if it took place in darkness (*Arabidopsis*), and if the plants were not too old at the beginning of the cold treatment (*Arabidopsis*). Short days or low light intensity may have devernalizing effects, too (sugarbeet, *Oenothera biennis*, and perennial plants like certain chrysanthemums). Devernalized plants can be revernalized.

Antivernalization. Flowering may be delayed and the effect of subsequent vernalization may be reduced by a heat treatment, for example, 3–5 days at 86–104°F (30–40°C) for lettuce, *Arabidopsis*, and winter rye. This antivernalization is applied to onions in order to prevent bolting.

Genetic basis. Up to six genes have been found responsible for the cold requirement of vernalizable plants. In the case of three of these genes—*Hb, Ii*, and *Kk*, which are found in wheat, barley, rye, and *Arabidopsis*—the cold requirement is caused by the recessive alleles; in the case of the other genes—*Ss, Tt*, and *Uu*, found in wheat, barley, *Hyoscyamus*, and *Arabidopsis*—it is caused by the dominant alleles. *I* and *K* (found in barley and wheat) are epistatic to *S*, and *ss* is epistatic to *ii* and *kk*; *T* (found in *Arabidopsis*) is epistatic to *H* and *hh*. A vernalization effect is never transmitted to sexual descendants. *See* GENE ACTION.

Biochemical basis. These differences in the genetic basis of vernalization suggest that vernalization has different biochemical backgrounds in different species. Thus the results gained by application of physiological chemical methods do not form a clear picture. In winter rye vernalization occurs only in the presence of oxygen. In some cold-requiring plants, like *Hyoscyamus* and *Arabidopsis* (but not in many Campanulaceae and Gramineae), cold can be replaced by gibberellins, which may overcome an NH_3-induced inhibition of vernalization in winter wheat; in others, like sugarbeet and *Oenothera*, gibberellins can replace cold only partially. Gibberellins may have certain side effects (for example, on internode elongation and ramification), which vernalization does not have, and they are almost never effective on the seed stage. In winter wheat (but not in winter rye), vernalization can perhaps be specifically inhibited by maleic hydrazide, which promotes flower formation in cold-requiring celery, and by chloramphenicol, which increases the effect of incomplete vernalization in winter rye. In the radish vernalization can be replaced to a certain degree by ribonucleic acid (RNA), and in *Streptocarpus* and *Oenothera* vernalization can be inhibited by thiouracil and fluorouracil, respectively, which replace uracil in RNA synthesis. However, no correlation between RNA, deoxyribonucleic acid (DNA), and vernalization has been found in other cases such as winter rye and *Hyoscyamus*. Studies of the role of auxins in vernalization also have given varying results.

Vernalization may further be partially or entirely replaced in certain plants by vitamin E (winter rye, *Cichorium intybus*), kinetin (*Cichorium intybus* and *Geum urbanum*), decapitation (*Chrysanthemum morifolium* and *Geum urbanum*), heat (*Scrophularia alata*), high light intensity (winter cereals and other winter annual grasses and *Geum urbanum*), and by others. Although the vernalization effect is transmissible by grafting in *Hyoscyamus*, it has not been possible to extract from vernalized plants flower-inducing principles other than gibberellinlike substances which, for several reasons, do not seem to be identical with the end product of vernalization. *See* AUXIN; GIBBERELLIN; NUCLEIC ACID; PLANT GROWTH.

Klaus Napp-Zinn

Bibliography. J. W. Einset, *Plant Growth and Development*, 1988; D. E. Fosket, *Plant Growth and Development: A Molecular Approach*, 1994; P. S. Nobel, *Plant Physiology*, 1991; W. Ruhland (ed.), *Encyclopedia of Plant Physiology*, vols. 15 and 16, 1965; P. F. Wareing and I. D. Phillips, *Growth and Differentiation in Plants*, 3d ed., 1981.

Vernier

A short auxiliary scale placed along the main instrument scale to permit accurate fractional reading of the least main division of the main scale, invented by Pierre Vernier about 1630. The auxiliary, or vernier, scale is graduated in one or both directions from the fiducial (index) mark in numbered divisions which are fractionally shorter (in a direct vernier) or longer (in a retrograde vernier) than those on the main scale. The position of the fiducial mark (the zero mark of the vernier scale) between divisions on the main scale is indicated by the number of the graduation on the vernier scale which lines up exactly with a graduation on the main scale.

Each vernier division of the direct-reading vernier in the **illustration** is equal to 0.9 of the main scale divisions. The fiducial mark is located between 3 and 4 on the main scale, and since graduation 6 lines up with the main scale the exact reading is 3.6 ($3.6 + 6 \times 0.9 = 9.0$).

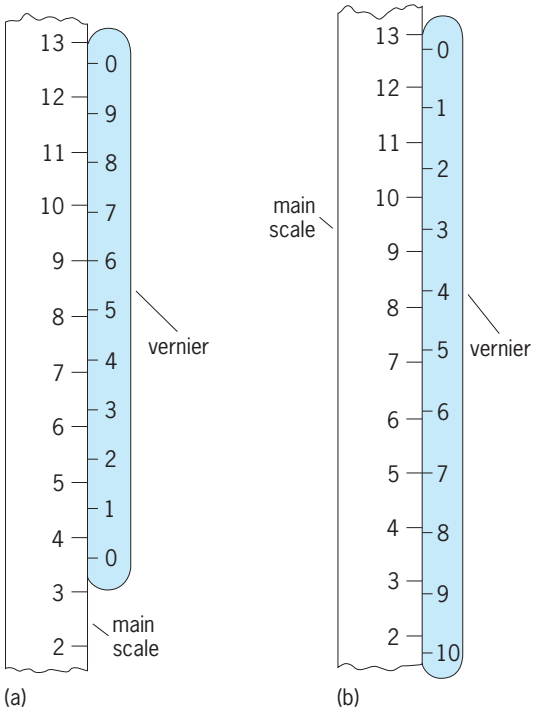

Two types of vernier scales. (*a*) Direct (reading 3.6). (*b*) Retrograde (reading 12.7).

Such a vernier scale gives a direct reading to the nearest tenth division and permits estimation, if none of the graduations on the vernier line up exactly, to 0.03 or 0.05 divisions.

Vernier scales are used in both linear and circular measurements, as on calipers, screw micrometers, and protractors on telescope mounts and surveying instruments. The divisions are usually tenths or hundredths of units, but may be quarters, twelfths, or other fractions for circular measurements. If the main scale is graduated in half degrees, the vernier divisions may be $^{11}/_{12}$ or $^{29}/_{30}$ of the main scale divisions, to permit reading to 5 or 2 minutes of arc.

The term vernier is often loosely applied to any fine adjustment, for example, to the one of two or more controls, such as valves or rheostats of different range, that causes the least control action per unit motion or other input. William A. Wildhack

Vertebra

The basic unit of the vertebral column. Collectively, the vertebrae surround and protect the spinal cord and provide some type of axial support for the body. The stresses that the vertebral column must meet change somewhat from one end of the animal to the other, and also differ greatly between aquatic and terrestrial vertebrates because the problems of support and locomotion in these media are quite different. Accordingly, vertebral structure varies widely; yet all vertebrae have many features in common.

Basic structure. A vertebra from the thoracic region of a mammal illustrates the basic morphology well (**Fig. 1**). The ventral portion consists of a disc-shaped mass of bone known as the body or centrum. An arch of bone, the neural arch, extends dorsally from the centrum and encompasses a space, the vertebral canal, in which the spinal cord lies. The bases, or roots, of the arch are narrower than other parts so that clefts, the intervertebral foramina, lie between the arch bases of successive vertebrae. Spinal nerves pass through these foramina.

Certain muscles and ligaments attach onto the spinous process, which extends dorsally from the top of the arch, and onto a pair of transverse processes, which extend laterally from the arch. One pair of articular processes, or zygapophyses, extend forward from the neural arch, and another pair extends posteriorly. Articular processes of successive vertebrae overlap and help to hold the vertebrae together. Additional articulating or reinforcing processes are present in certain other terrestrial vertebrates. The centra of adjacent vertebrae are also joined together by invertebral discs of fibrocartilage. Numerous ligaments interlace the vertebrae.

Ribs articulate onto the thoracic vertebrae. Typically, each mammalian rib has two articular surfaces—a terminal head and a tubercle situated a short distance distal to the head. The tubercle articulates with a facet located on the end of the transverse process; the head usually articulates intervertebrally onto the intervertebral disc and adjacent parts of the bordering centra.

Development. The vertebrae develop from mesenchymal cells that migrate from the sclerotomal portions of the somites and spread around the embryonic neural tube and notochord (**Fig. 2**). The first sign of vertebrae differentiation in this mesenchymal sheath is the development of small blocks of cartilage known as the dorsal arcualia. Since the spinal ganglia and nerves have previously begun to differentiate in a segmental position directly adjacent to the muscle segments, the dorsal arcualia, which are the

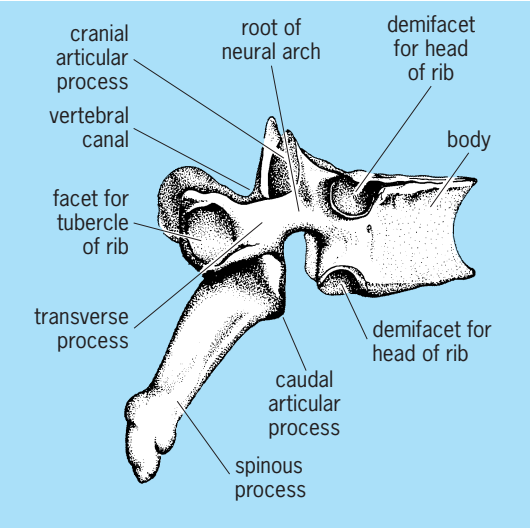

Fig. 1. Lateral view of a human thoracic vertebra. Anterior is toward the right. (*After C. M. Goss, ed., Gray's Anatomy of the Human Body, 28th ed., Lea and Febiger, 1964*)

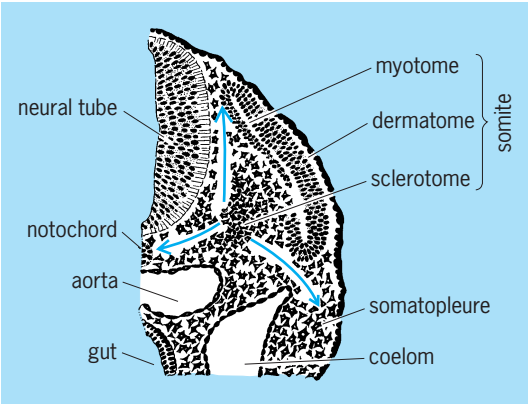

Fig. 2. Diagrammatic cross section through one side of the back of a 4-week-old human embryo. Migration of mesenchymal cells from the sclerotome is shown by arrows. (*After L. B. Arey, Developmental Anatomy, 6th ed., Saunders, 1954*)

rudiments of the neural arches, assume an intersegmental position. The rest of each vertebra, in turn, differentiates intersegmentally, where it can be acted on by muscles of adjacent segments. Most vertebrates have only one pair of dorsal arcualia per body segment, but sharks and other cartilaginous fishes have two pair. A pair of ventral arcualia differentiate adjacent to the ventral part of the notochord and develop either into ribs or, in the tail region of many vertebrates, into a ventral hemal arch that surrounds the caudal artery and vein.

The centra of the vertebrae develop around the notochord, or within its sheath, primarily from mesenchyme derived directly from the sclerotomes. In some fishes, however, bases of the arcualia contribute to the centra. Centra formation is very incomplete in cyclostomes and lungfishes, and a large notochord persists into adult life; but in most vertebrates the notochord is replaced, or nearly so, by the developing centra. A thin strand of notochord runs through the centra in cartilaginous fishes and many salamanders, and traces of the notochord may lie between successive centra. There is a trace within the intervertebral disc of mammals.

The arches and centra remain cartilaginous throughout life in cartilaginous fishes. Parts of the vertebrae may remain cartilaginous in other groups, but by the adult stage of most vertebrates the cartilage is replaced by bone.

Fishes. Since shortening of the body of fishes is resisted by the incompressible, yet flexible, vertebral column, contraction of the longitudinal fibers in the muscle segments, first on one side of the body and then on the other, causes the body to bend from side to side. The vertebral column of fishes is an important part of the locomotor system. Its role in physical support of the body is rather negligible because flesh is only slightly denser than water and needs little support in an aqueous medium. The centra of fishes are strong discs or biconcave cylinders with traces of the notochord persisting in the concavity between them. They resist compression forces well, yet permit the body to bend from side to side. Neural arches lack articular process because there are few vertical forces to be resisted.

Since the entire vertebral column of fishes is performing essentially the same function, there is little regional differentiation. Trunk vertebrae bearing ribs extend from the trunk to the tail. The first vertebra is slightly modified for articulating with the skull, but head and trunk do not move independently and there is no neck. Ribs disappear in the tail region, and the ventral arcualia form an arch of bone, the hemal arch, that encases the caudal artery and vein.

Amphibians. Terrestrial vertebrates receive little support from the surrounding air, and most propel themselves by thrusts of the limbs against the ground. The vertebral column becomes essentially a beam that suspends the body weight and transfers this to the ground by way of the limbs. The vertebral column must resist considerable vertical forces. The centra are thoroughly ossified, and the neural arches are securely united by articular processes.

There is only a little more regional differentiation in the vertebral column in ancestral amphibians than in fishes. A single vertebra is specialized for articulating with the head; no well-defined neck region is present. A series of dorsal, or trunk, vertebrae extend to the pelvic girdle. Here a single sacral vertebra and pair of ribs are specialized for articulating with the pelvic girdle. Anteriorly in terrestrial vertebrates weight is transferred from the vertebral column and ribs to the pectoral girdle by a muscular sling. Caudal vertebrae follow the sacral vertebra. Ribs articulate with all vertebrae from the head into the base of the tail. Longest near the front of the series, they are progressively shorter posteriorly.

Mammals. Regional differentiation of the vertebral column continues in reptiles, and by the mammalian stage of evolution different parts of the column are clearly specialized to subserve certain functions in addition to their general supportive role.

Cervical vertebrae. The head moves independently of the trunk, and a distinct neck region is present. With few exceptions all mammals from a shrew to a giraffe have seven cervical vertebrae. In whales and porpoises, which have readapted to an aquatic mode of life by assuming a fishlike body form, an independent motion of head and trunk would be undesirable. Seven cervical vertebrae are present, but they are much shortened and fused together. Small embryonic ribs have fused on the sides of the cervical vertebrae and contribute to the transverse process. Another diagnostic feature of most cervical vertebrae is the perforation of the transverse process by a transverse foramen that carries the vertebral artery and vein (**Fig. 3**). The first two cervical vertebrae, the atlas and axis respectively, are so modified as to permit the head to move in nearly any direction.

Thoracic vertebrae. Well-developed ribs, which play an important role in respiratory movements, articulate with the anterior trunk or thoracic vertebrae of mammals. The more anterior ribs are quite short and stout and also play an important role in weight transfer from the vertebral column to the pectoral girdle. A strong muscular sling, composed primarily of the

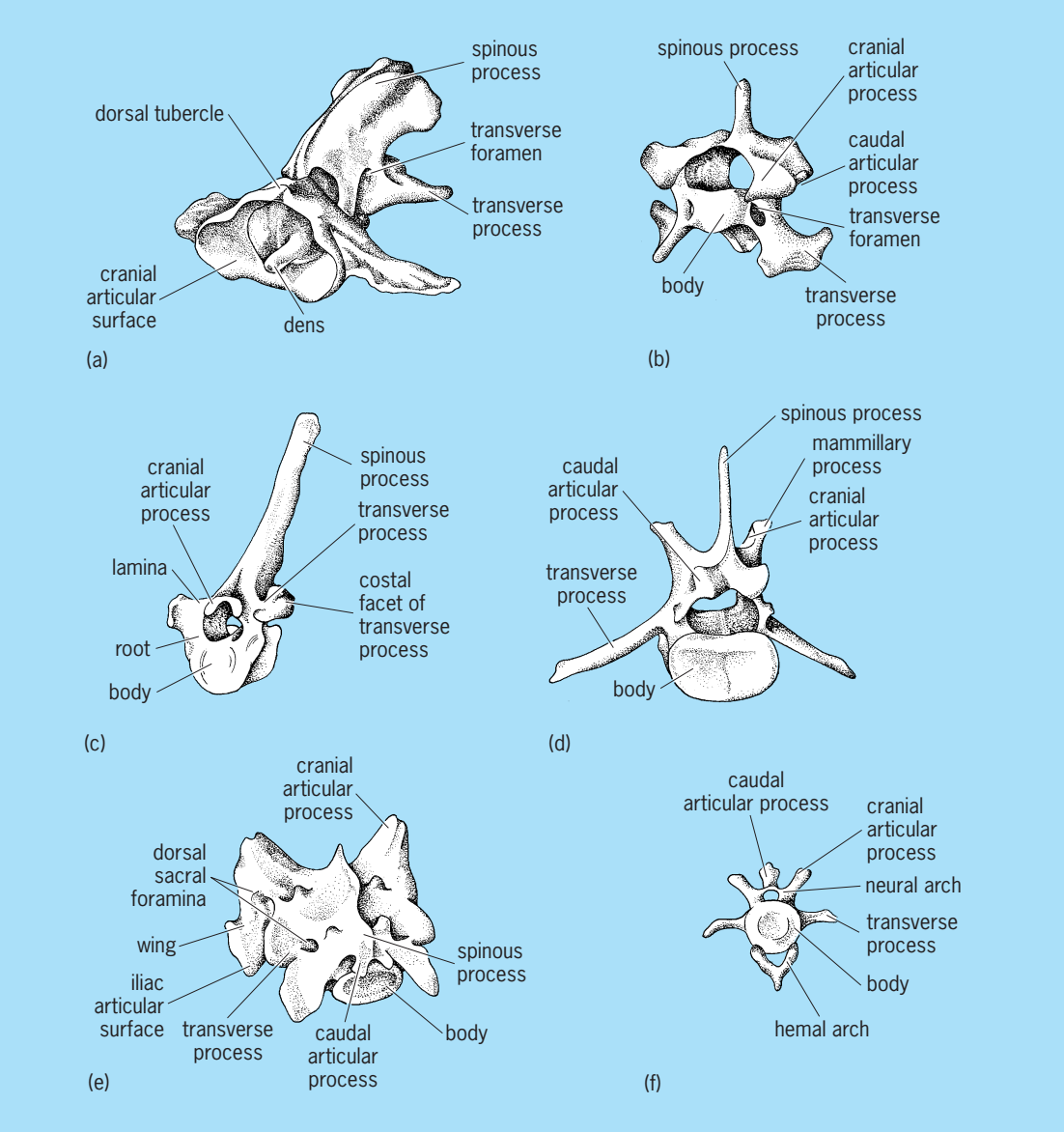

Fig. 3. Representative vertebrae of a dog. (*a*) Atlas and axis articulated, cranial lateral aspect. (*b*) Fifth cervical, cranial lateral aspect. (*c*) Sixth thoracic, cranial lateral aspect. (*d*) Fifth lumbar, caudal lateral aspect. (*e*) Sacrum caudal, lateral aspect. (*f*) Fourth caudal, cranial aspect. (*After M. E. Miller, G. C. Christensen, and H. E. Evans, Anatomy of the Dog, Saunders, 1964*)

serratus ventralis, extends from the ribs to the dorsal border of the scapula. Mammalian thoracic vertebrae, and only the thoracic vertebrae, bear distinct articular surfaces for ribs. The number of thoracic vertebrae varies between species but is on the order of 11 (bat) to 18 or 20 (horse). Humans have 12.

Lumbar vertebrae. Lumbar vertebrae occupy the posterior part of the mammal trunk region. They are characterized by relatively large transverse processes to which certain powerful back muscles attach. The transverse processes include embryonic rib rudiments that have fused onto the vertebrae. The number of lumbar vertebrae vary between species but is on the order of 5 (bat) to 8 (whale). Humans have 5.

Sacral vertebrae. Correlated with the greater efficiency of terrestrial locomotion and the need for strong support for the powerful hindlegs, the number of sacral vertebrae increased during evolution from the single one of amphibians. Reptiles usually have two, and most mammals have three which are fused together, along with their embryonic rib rudiments, into a solid complex of bone called the sacrum. In species in which there are heavy stresses in this region, the number of sacral vertebrae is increased. Humans, which are bipeds, have 5, and certain of the powerful hoofed mammals, for example, the horse, also have 5. Birds, whose hindlegs act as shock absorbers upon landing, have between 10 and 23 vertebrae fused together in their synsacrum.

Caudal vertebrae. The tail and caudal musculature no longer play an important role in the locomotion of most mammals (the Cetacea being a conspicuous exception), and the tail is greatly reduced in size. The spinal cord of mammals ends within the lumbar region, and only a few spinal nerves continue through the vertebrae canal into the tail. Caudal vertebrae are

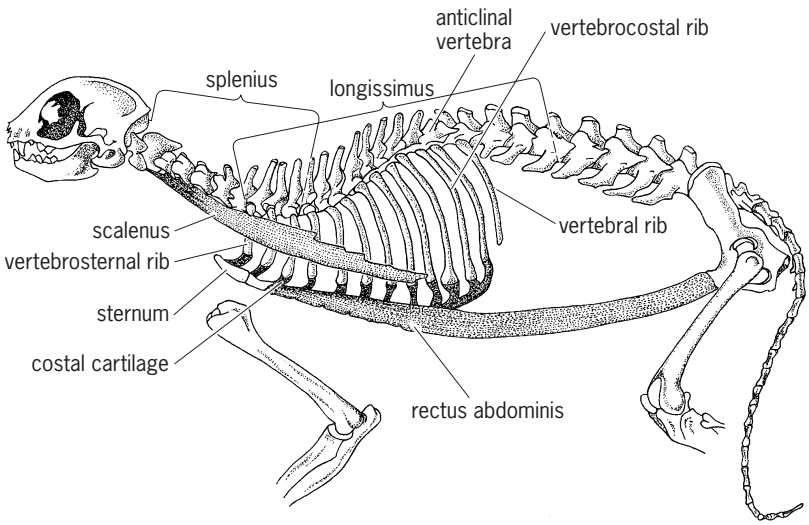

Fig. 4. Lateral view of cat skeleton to illustrate Slijper's views on biomechanics of trunk skeleton. The brackets indicate lines of action of certain muscles upon the neural spines. (*After W. F. Walker, Vertebrate Dissection, 4th ed., Saunders, 1970*)

small and become progressively incomplete as one moves distally along the tail until only centra are left. The more proximal caudal vertebrae often bear on their ventral surfaces small, V-shaped hemal arches or chevron bones. Tail length, and hence the number of caudal vertebrae, vary widely. Some oppossums have as many as 35, and humans, in which the tail is absent as an external structure, have only 3 to 5 caudal vertebrae. These form an internal coccyx to which certain anal muscles attach.

Biomechanics. The vertebral column of mammals is a complex structure that forms the primary girder of the skeleton, subserves the functions discussed above, and by its extensions and flection takes part in the locomotion of certain fast-moving quadrupeds. In order to analyze its biomechanics, some authors have compared it to a cantilever girder, but others have taken a broader view and have compared the entire trunk skeleton and its muscles to an archer's bow (**Fig. 4**). The centra form the main supporting "bow." It is under tension from dorsal muscles and would tend to straighten out if it were not for the "bowstring" (pelvic girdle, rectus abdominis, and certain ribs) counteracting this force. This hypothesis views the vertebral column as a dynamic girder capable both of providing support and of participating in the movements of the body.

The spinous processes are interpreted as muscle lever arms acting on the centra. The direction of their inclination tends to be perpendicular to the major muscle forces acting upon them. The increase in height of the spinous processes at the base of the neck and front of the thorax increases the length of these lever arms and the mechanical advantage of the muscles acting on them. Many of these muscles act to support the relatively heavy head on the end of the vertebral beam. *See* BIOMECHANICS; SKELETAL SYSTEM.

Warren F. Walker, Jr.

Bibliography. R. L. Carroll, *Vertebrate Paleontology and Evolution*, 1987; A. S. Romer and S. T. Parsons, *The Vertebrate Body*, 6th ed., 1986.

Vertebrata

The largest natural group (clade, monophyletic group) of chordate animals. Like all chordates, vertebrates have a notochord, but it is largely replaced by vertebrae (initially neural arches that surround the nerve chord) during development. Together with the living hagfishes (Hyperotreti; Myxiniiformes) and many fossil groups, vertebrates are members of the Craniata. One major theme of craniate evolution is the increasing elaboration of the brain, its cranial nerves, and many sensory organs, such as the organs of smell (olfactory), hearing (inner ear) and feel (sensory organs of the lateral line system). A unique innovation is the evolution of neural crest cells, cells that begin development along the neural crest of the embryo and which then migrate to many parts of the body, giving rise to or contributing to the formation of such structures as the cranium and the branchial arches (a remnant of which remains in humans as parts of the larynx). Vertebrates have a long fossil history. Earliest fossils of jawless vertebrates are known from the late Cambrian (490 million years before present, mybp). Jawed vertebrate fossils are known from the Ordovician (450 mybp) and this group began to predominate in the Devonian (about 360 mybp). *See* CHORDATA; NERVOUS SYSTEM (VERTEBRATE).

Formal vertebrate classifications of the past emphasized overall distinctiveness and grade of organization. Modern (evolutionary phylogenetic) classifications are based on common ancestry and genealogy. Thus, traditional classifications have been highly modified and classifications in many textbooks are out of date. Older classifications recognized a group, Agnatha, which contained lampreys and the more primitive hagfishes. Modern classifications recognize the evidence that lampreys, although jawless, are nevertheless more closely related to jawed vertebrates than to hagfishes. Older classifications placed the lungfishes and coelacanth with the bony fishes. Modern classifications place them with the legged vertebrates, because they share a common ancestor with legged vertebrates not shared with bony fishes. *See* ANIMAL SYSTEMATICS; SARCOPTERYGII.

Lampreys and fossil jawless fishes. The most basal living members of Vertebrata are the lampreys (Hyperoatia; about 35 species). Lampreys are jawless fishes whose eel-like body resembles hagfishes, but whose predatory, round, and toothed mouth is unique among vertebrates. Lampreys and many fossil jawless vertebrates that flourished in the Paleozoic have a number of unique evolutionary innovations. These include neural arches (vertebrae) that enclose the dorsal nerve chord, muscles that work to move the fin rays, vertebrate heart with atrium and ventricle closely situated, nervous regulation of the heart, and a number of important sensory innovations including the evolution of extrinsic eye muscles, the elaboration of the inner ear to two vertical semicircular canals (later to be three in jawed vertebrates), and the development of true neuromasts (sensory

hair organs) in the sensory canal system. Of course, some of the soft anatomy cannot be observed in fossils.

Jawed vertebrates. Major evolutionary innovations characterized the jawed vertebrates, Gnathostomata ("jawed mouth"). In addition to jaws and characters mentioned above, gnathostomes have paired fins or the homologue of paired fins, legs. They also have a mineralized exoskeleton that surrounds certain cartilages of the head and fins and consists of true bone (as in human leg bones) or calcified cartilage (as in many sharks). As is usual in evolution, such characters evolve stepwise; there are jawless relatives of gnathostomes with paired fins, and more primitive relatives that have mineralized bone but which lack either paired fins or jaws. One of the remarkable innovations of the gnathostome ancestor was the evolution of the adaptive immune system, the basis of our ability to fight off infection and to develop vaccines to fight many infectious diseases.

Placodermi and Chondrichthyes. There are two major living clades of gnathostomes and one major extinct clade. The extinct clade is Placodermi, a group of Paleozoic fishes characterized by heavy body armor and which includes some of the largest predatory fishes of Paleozoic waters. The smaller of the living clades is Chondrichthyes. It includes the chimeras (30 species) and sharks and rays (about 830 species). Chondrichthyes has a rich fossil history; the earliest remains are known from the Late Ordovician, and early sharks were common predators in both Paleozoic and Mesozoic seas. They are characterized by several evolutionary innovations including internal fertilization via the unique male clasper organs, and the histology of the calcified cartilage that surrounds the cartilaginous skeleton (micrometric prismatic cartilage). The lack of an obvious bony skeleton is a secondary innovation of sharks and is not primitive. In fact, Chondrichthyes fishes have bone, which form the base of each scale. *See* CHONDRICHTHYES; PLACODERMI.

Osteichthyes. The third major clade of vertebrates is Osteichthyes, a group that in traditional classification contained only the bony fishes, but which now includes the Actinopterygii (bony fishes, about 25,000 species) and the Sarcopterygii (about 24,000 species). Sarcopterygii includes the legged vertebrates (amphibians, reptiles, mammals, and birds), the living lungfishes, coelacanths, and their diverse fossil relatives commonly known as crossopterygians. *See* OSTEICHTHYES.

Current classification. A current classification of living vertebrates is shown below. Note that the ranks of this classification serve only to organize and subordinate the groups; ranks do not imply distinctiveness or uniqueness of adaptations. Many of the traditional classes, such as class Mammalia, would have lower rank in modern classifications such as this one.

Subphylum Chordata
- Infraphylum Craniata
  - Superclass Hyperotreti (hagfishes)
    - Vertebrata
      - Class Hyperoatia (lampreys)
        - Gnathostomata
          - Infraclass Chondrichthyes (chimeras, sharks, rays)
            - Osteichthyes
              - Subclass Actinopterygii (bony fishes, sturgeons, trouts, basses, tunas, etc.)
              - Sarcopterygii (lungfishes, coelacanths, amphibians, reptiles, birds, mammals)

E. O. Wiley

Vertebrate brain (evolution)

A highly complex organ consisting of sensory and motor systems that constitutes part of the nervous system. Virtually all of the brain systems that are found in mammals occur in birds, reptiles, and amphibians, as well as in fishes and sharks. These systems have become more complex and sophisticated as the adaptive requirements of the animals changed. Occasionally, new sensory systems arose, most likely as specializations of existing systems; however, some of these disappeared as animals left the aquatic world. Some sensory systems arose and declined several times in different lineages. In spite of the many changes in brain structure, some of them quite dramatic, the evolution of the nervous system, from the earliest vertebrates through to those of today, has been relatively conservative. *See* NERVOUS SYSTEM (INVERTEBRATE); NERVOUS SYSTEM (VERTEBRATE).

Central nervous system. The nervous system consists of two main divisions: the central nervous system, which is made up of the brain and the spinal cord, and the peripheral nervous system. The peripheral nervous system consists of the nerves that bring information from the outside world via the eyes, ears, nose, skin, and other sensory systems, and the nerves that carry information from the body's interior to the spinal cord and brain. These nerves also convey commands from the brain and spinal cord to the external muscles that move the skeleton, as well as to various internal organs and glands. *See* CENTRAL NERVOUS SYSTEM.

The head end of any bilaterally symmetrical animal is the most important end. Most of the sensory systems are located in or on the head because that is the ideal placement for systems that monitor the environment for danger, something good to eat, a potential mate, a safe place to hide, and anything else important to the animal's survival.

The brain consists of a variety of systems, some of which are sensory and deal with the acquisition of information from the internal and external environments. Other systems are motor and are involved with the movement of the skeletal muscles; the muscles of the internal organs, such as the heart and the digestive and respiratory systems; and the secretions of certain glands, such as the salivary and tear glands. The bulk of the brain, however, is composed of systems that are neither sensory nor motor, although

they have strong connections with both. Instead, these systems are integrative and organize, coordinate, and direct the activities of the sensory and motor systems. These integrative systems regulate such processes as sleep and wakefulness, attention, the coordination of various muscle groups, emotion, social behavior, learning, memory, thinking, planning, and other aspects of mental life. Social behavior itself is highly complex and includes such interactions between individuals as courtship, mating, parental care, and the organization and structure of groups of individuals. *See* BRAIN; MOTOR SYSTEMS.

Development. The nervous system develops in the embryo as a hollow tube. The remnants of this hollow tube in the adult are known as the brain ventricles and are filled with the cerebrospinal fluid. Two types of overall brain organization are found among vertebrates based on the relationship of the neurons to these ventricles. In brains with the laminar type of organization, the neurons have not migrated very far from the layer immediately surrounding the hollow ventricular core of the brain, which is the zone in which neurons are born (**Fig. 1**). This type of organization is typical of amphibian brains and is also found in the brains of some sharks, fishes (especially those in which their skeletons are partly or mostly made of cartilage rather than bone), and lampreys (jawless fishes). In contrast, in the brains of those vertebrates with the elaborated type of organization, the neurons have migrated from the zone around the ventricles to occupy nearly all of the interior of the brain. This type of brain organization is typical of reptiles, mammals, birds, fishes (with fully bony skeletons), as well as skates, rays, some sharks, and hagfishes (jawless fishes) [**Fig. 2**]. *See* NEURON.

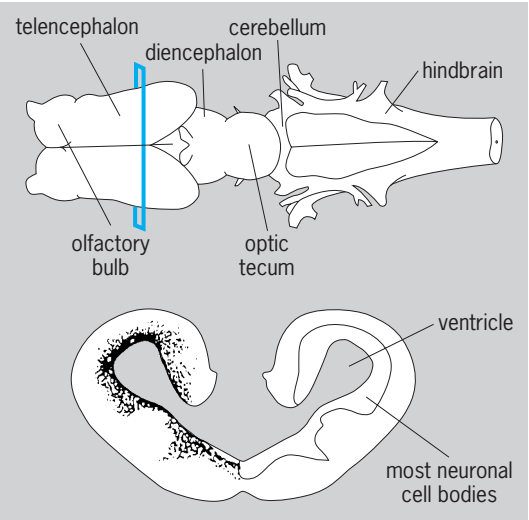

Fig. 1. Forebrain of a salamander, with a laminar brain organization. The color plane provides the cross section below. Most of the neurons remain congregated around the surface of the brain ventricle near the zone where they were "born." The cerebral hemisphere also shows invagination (rolling inward toward the midline during development). A membrane surrounding the brain (not shown) keeps the cerebrospinal fluid of the ventricles contained.

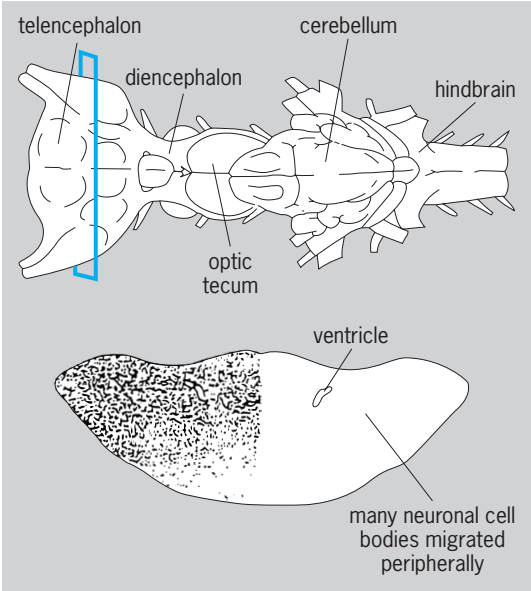

Fig. 2. Forebrain of a shark, with an elaborated brain organization. Most of the neurons have migrated away from the ventricular surface and fill virtually the entire volume of the cerebral hemisphere.

An interesting feature of the brains of bony fishes results from a developmental difference from other vertebrates. Before the embryonic nervous system rolls up to form a hollow tube, it is a relatively flat plate. In other vertebrates, the part of the plate that will form the brain rolls inward as does the rest of the nervous system to form the hollow tube, a process called invagination. In bony fishes, however, the region destined to become the brain rolls outward and downward, a process called eversion or turning outward (**Fig. 3**).

Brain size. In general, brain size or weight varies in proportion to the size of the body. In some species, however, brain size or weight is greater than would be expected for that body weight. Examples of bigger-than-expected brains are found in humans, chimpanzees, and porpoises. Although an elephant has a much larger brain than a human, it is no greater than would be expected for an animal of that body weight. Other mammals with brain weights typical for their body size are wolves, rats, and gorillas. In addition, birds have brains that are comparable to those of mammals of equivalent body weight. Some birds, such as crows, have brain weights in excess of what would be expected for their body weight. Moreover, crows have brain weights that would be expected of a small primate of equivalent body weight. In contrast to birds and mammals, amphibians, reptiles, and bony fishes have relatively small brains for their body weights, as do jawless fishes (**Fig. 4**).

Subdivisions of brain. The brain is divided into a hindbrain, a midbrain, and a forebrain.

Hindbrain. The hindbrain is a region that contains the nerve endings that receive information from the outside world and from the body interior; these are known as sensory cranial nerves. The neuron groups

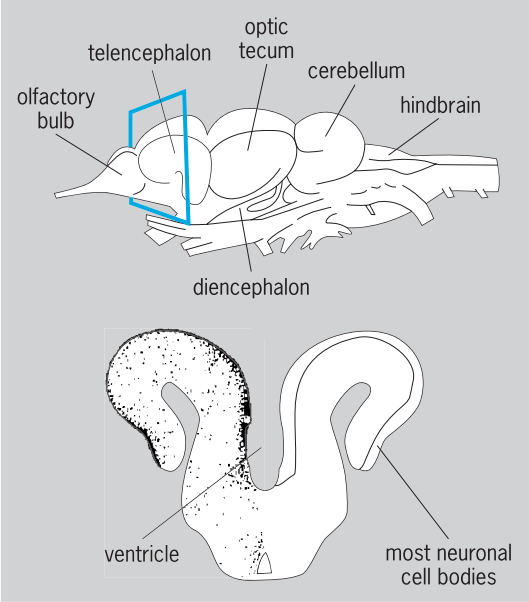

Fig. 3. Forebrain of a fish that shows eversion of the cerebral hemisphere (rolling away from the midline during development). The color plane provides the cross section below. This fish also has a laminar brain. The eversion of the cerebrum and the congregation of neurons near the ventricular surface result in the "inside-out" organization of such brains. A membrane surrounding the brain (not shown) keeps the cerebrospinal fluid of the ventricles contained.

upon which they terminate are known as sensory cranial nerves. Also found in the hindbrain are the motor nerves that control the internal and skeletal muscles and glands, which are called motor cranial nerves; the neuron groups from which they originate are known as motor cranial nerve nuclei.

Some fishes have an extraordinary sense of taste, and this is reflected in the development of their hindbrains. Such fishes are the carps and the catfishes, which are both bottom feeders that live in murky waters. In addition to having many taste buds in their mouths and throats, these fishes have thousands of taste buds studding their entire body surface. In effect, their bodies are like giant tongues that taste the water as the fish swims through it. The so-called whiskers of the catfishes, known as barbels, also are covered with large numbers of taste buds. A catfish will drag its barbels across the bottom to taste for morsels of food that might be present. Three cranial nerves are involved in the detection of taste stimuli: the facial nerve, the glossopharyngeal nerve, and the vagus nerve. The facial nerve, which innervates the skin taste buds, terminates in a large lobe on the surface of the hindbrain known as the facial lobe. The vagus and glossopharyngeal nerves, which innervate taste buds of the mouth and throat, terminate in an enormous lobe just to the side of the facial lobe, called the vagal lobe. Only carps and catfishes have these taste lobes on the hindbrain.

Many animals possess senses that humans do not possess. One such sense is the lateral line sense which derives from receptors located in the lateral line organ and can easily be seen on most bony fishes as a thin, horizontal line running the length of the body from behind the gill opening to the tail. Other lateral line organs can be found on the head and jaws. These lateral line systems contain mechanoreceptors that respond to low-frequency pressure waves that might be produced by other fishes nearby or the bow wave of a fast-swimming predator about to strike. The cranial nerves that innervate the lateral line organs terminate in a special lateral line region of the hindbrain. The lateral line sense is an effective way for animals to find out something about their world without the use of vision. Lateral line systems and a special region of the hindbrain dedicated to lateral line sense are found in fishes and sharks, jawless fishes, and bony fishes of various sorts. They are found in amphibians only in the tadpole phase of their lives and are totally lacking in birds, reptiles, and mammals. *See* LATERAL LINE SYSTEM.

Electroreception is another way of dealing with a murky environment. Scientists have described two types of electroreception: active and passive. The receptors for this sense are also located in the lateral line canals and sometimes on the skin as well. Animals with passive electroreception, such as sharks and rays, platypuses, and echidna, can detect the presence of the very weak electric fields that are generated around a living body, which they then follow to capture their prey. Animals with active

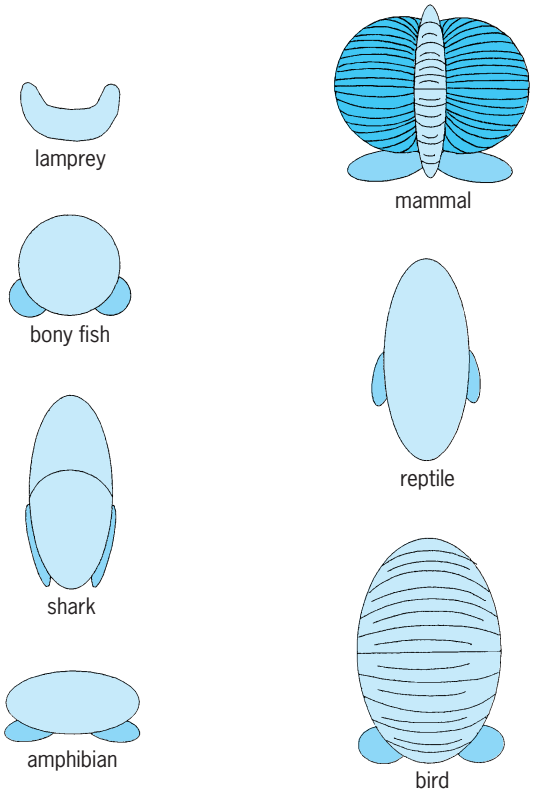

Fig. 4. Some examples of the varieties of cerebellar size and shape among vertebrates. (*Adapted from A. B. Butler and W. Hodos, Comparative Vertebrate Neuroanatomy: Adaptation and Evolution, 1966*)

electroreception generate stronger electric fields around themselves using specialized organs called electric organs. By detecting changes in the electric fields around themselves, they can derive a picture of their world including the objects and creatures in it. Electrosensory cranial nerves terminate in a region of the hindbrain known as the electrosensory area, which is close to the lateral line area. A second group of active electrosensory fishes are capable of generating electric fields that are so powerful they can stun a prey or an enemy. Among these are the electric eel, the electric catfish, and an electric shark (the torpedo). These animals also use their low-level electric fields to detect objects and creatures in the environment. *See* ELECTRIC ORGAN (BIOLOGY).

Not only did the hindbrain change in response to the great alterations in sensory evolution, but it also underwent major transformations; for example, motor-neuron groups involved in swallowing, chewing, and salivating evolved as a consequence of the transition to land and the loss of the water column to carry food from the opening of the mouth into the throat.

In addition to the sensory and motor cranial nerve nuclei, the hindbrain contains two important coordinating or integrating systems: the cerebellum and the reticular formation. The functions of the cerebellum are varied; they include the integration of a sense of balance with aspects of movement and motor learning and motor memory, as well as playing an important role in electrosensory reception. Although the internal organization of the cerebellum's neurons has remained rather conservative in evolution, the size and overall form of the cerebellum vary widely. One view is that the size of the cerebellum depends on the extent to which an animal moves through three dimensions of space rather than remaining more or less on one plane. Thus, bottom-dwelling fishes and surface-dwelling mammals would have relatively small cerebella, whereas fishes that change depth extensively and mammals that climb trees and descend to the ground, as well as flying birds and mammals, would have larger cerebella. There are, however, numerous exceptions to this rule.

The reticular formation coordinates the functions of various muscle groups that must cooperate with each other. For example, the actions of the jaws and tongue must be coordinated so that an animal does not eat its own tongue while eating its meal. The reticular formation also coordinates the motor-neuron groups that control the air column that enters and leaves the mouth and throat, which produces the various vocalizations of land animals, including speech. The reticular formation also is involved in sleep, wakefulness, and attention. In general, the reticular formation has been very conservative in its evolution, although it is best developed in reptiles, birds, and mammals.

Midbrain. The midbrain contains the motor cranial nerves that move the eyes. It also is sometimes regarded as the "map room" of the brain because it contains neuron groups that are organized to form maps of visual space, auditory space, and the body. These maps are coordinate with each other so that a sudden unexpected sound will cause the head and eyes to move to the precise region of space from which the sound originated. In those animals that make extensive use of sound localization to find their way in the world, such as owls and bats, the map areas of the midbrain are very highly developed. In addition, certain snakes, such as rattlesnakes and boa constrictors, have infrared detectors on the snout or under the eyes. These detectors can sense the minute heat from a small animal's body at a distance of 1 m (3 ft) or more. The midbrains of these animals also have infrared maps that are in register with the auditory, visual, and body maps to permit the animal to correlate all the necessary information to make a successful strike on prey in virtually total darkness.

Forebrain. The forebrain is a very complex region that consists of the thalamus, the hypothalamus, the epithalamus, and the cerebrum or telencephalon. In addition, the forebrain contains the limbic system, which has components in all regions of the forebrain as well as continuing into the midbrain. The thalamus processes and regulates a large quantity of the information that enters and emanates from the forebrain. As the cerebrum increases in size and complexity in land animals, especially in mammals and birds, the thalamus increases accordingly. The hypothalamus regulates autonomic functions, such as respiration, heart rate, blood pressure, and temperature, as well as behaviors such as feeding, drinking, courtship and reproduction, parental, territoriality, and emotional, which it controls in conjunction with the limbic system. The hypothalamus also regulates the endocrine system. The size and complexity of the hypothalamus, relative to the rest of the brain, is greatest in fishes and sharks; it declines considerably in proportion to the rest of the brain in land animals. The epithalamus contains the pineal gland, which is involved in various biological rhythms that depend on daylight, including seasonal changes such the beginning and end of the breeding season and annual migrations. In some animals, such as certain reptiles, the pineal takes on the form of an eye, located on the top of the head and known as the parietal eye. This eye has a lens and a primitive retina that capture light and transmit information, such as the amount of daylight, to the hypothalamus. The epithalamus, like the hypothalamus, is relatively smaller in the brains of land animals.

The greatest evolutionary expansion of the forebrain is seen in the cerebrum. The cerebrum consists of an outer layer, the pallium, and a series of deep structures, known as the subpallium. The subpallium is composed of the corpus striatum, the amygdala, and the hippocampus. The outer layer of the cerebrum in mammals is known as the cerebral cortex. Considerable debate surrounds the evolutionary relationship between the cerebral cortex of mammals and the pallium and subpallium of nonmammals. Most specialists, however, seem to agree that the mammalian cortex arose from the pallium and

certain regions of the subpallium. The cerebrum is relatively small in animals with laminar brains and larger in those with complex brains. Scientists have only begun to catalog the many complicated behavioral functions of the cerebrum. Among them appear to be memory, thinking and reasoning, and planning. With the advent of life on the land, the cerebrum underwent an extreme degree of elaboration in reptiles and birds and especially in mammals. *See* ENDOCRINE SYSTEM (VERTEBRATE). William Hodos

Bibliography. M. J. Benton, *Vertebrate Paleobiology*, 1990; A. B. Butler and W. Hodos, *Comparative Vertebrate Neuroanatomy: Adaption and Evolution*, 1996; R. Nieuwenhuys, H. J. ten Donkelaar, and C. Nicholson, *The Central Nervous System of Vertebrates*, 1997.

Vertical takeoff and landing (VTOL)

A flight technique in which an aircraft rises directly into the air and settles vertically onto the ground. Such aircraft do not need runways but can operate from a small pad or, in some cases, from an unprepared site. The helicopter was the first aircraft that could hover and take off and land vertically, and is now the most widely used VTOL concept. *See* HELICOPTER.

The helicopter is ideally suited for hovering flight, but in cruise its rotor must move essentially edgewise through the air, causing vibration, high drag, and large power losses. The aerodynamic efficiency of the helicopter in cruising flight is only about one-quarter of that of a good conventional airplane. The success of the helicopter in spite of these deficiencies started a wide-ranging study of aircraft concepts that could take off like a helicopter and cruise like an airplane. The term VTOL is usually used to designate the aircraft other than the helicopter that can take off and land vertically. The term V/STOL indicates an aircraft that can take off vertically when necessary, but can also use a short running takeoff, when space is available, to lift a greater load. *See* SHORT TAKEOFF AND LANDING (STOL).

Many of the VTOL concepts of the early 1950s were carried to flight test by using small, relatively inexpensive configurations to explore flying characteristics in vertical takeoff and landing and in the transition to and from conventional flight. These programs determined the design techniques needed to cope with the unique aerodynamic, stability and control, and structural response problems of VTOL operation. By the 1960s the technology had progressed to the point that more sophisticated versions of some concepts could be considered for operational test and evaluation.

VTOL concepts may be considered under two categories, advanced rotorcraft and fixed wing. The first category includes concepts featuring rotors for vertical lift similar to the helicopter, and the second generally has a fixed wing with deflected jets.

Advanced rotorcraft. This category is subdivided into two concepts: aircraft that operate like a helicopter, and those that operate like a helicopter in hover flight but tilt the rotors to act as propellers in cruise.

Fig. 1. Bell/Boeing V-22 Osprey.

Tiltrotor. The tiltrotor concept began with the Bell XV-3 in 1958. This concept is closest to the conventional helicopter and relies heavily on helicopter technology. The rotor disks are horizontal in VTOL operation and are tilted 90° to act as propellers in cruising flight. Such aircraft can have cruise efficiencies at least twice those of the helicopter, making them especially useful for helicopter missions where greater range, speed, and time on station are desired. In 1977 the first Bell XV-15 was flown in a successful test program that further matured tiltrotor technology and led to the Bell/Boeing V-22 Osprey (**Fig. 1**), the first operational tiltrotor aircraft.

The V-22, which first flew in 1989 and is about three times the size of the XV-15, has many possible military applications, including troop carrying in Marine assaults, Army medical evacuation, Navy antisubmarine warfare, Air Force special operation forces, and Coast Guard search and rescue. A civil version of this concept could help relieve airport ground and air congestion by flying from vertiports, areas that could be made more convenient, more strategically located, and smaller than airports, so that airports could be bypassed entirely for shorter-range flights.

X-wing. Another concept explored extensively during the late 1980s was the X-wing. For vertical flight, the four rotating helicopter-type blades featured blowing along the trailing and leading edges to increase hover performance. During cruise, the rotor stops and the aircraft flies in a fixed-wing mode, with two forward and two aft swept wings. The concept went through ground-based testing, and a flight research vehicle was built. Budget constraints curtailed the program before the X-wing configuration could be flight-tested.

Other concepts. Numerous advanced rotorcraft concepts are continuously being investigated in research programs (**Fig. 2**). These include designs that take off and land as a rotor system and fix the blades to act as wings for cruise flight. Other concepts feature tiltrotor technology for vertical flight, but fold the blades and employ other propulsion modes for cruise.

Fixed wing. Development of the vectored-thrust concept began in the late 1950s. The concept features an engine with four rotating nozzles that deflect

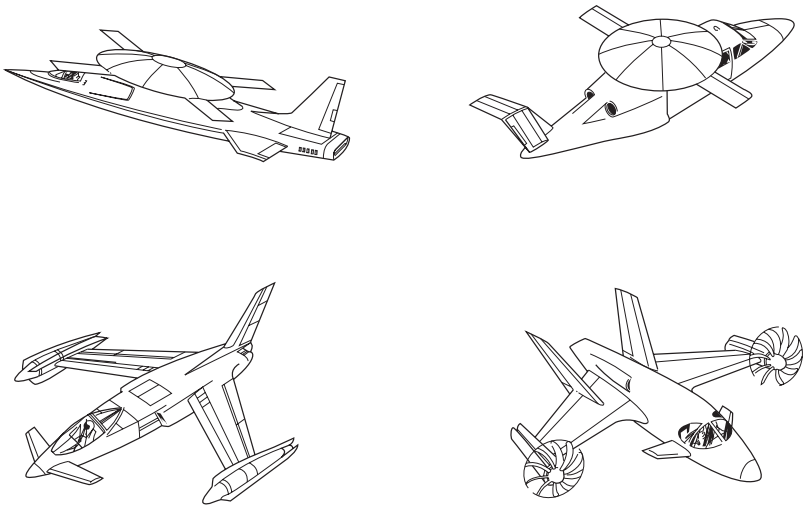

Fig. 2. Examples of advanced rotorcraft concepts.

the thrust from horizontal for conventional flight to vertical for VTOL operation. This activity led to the very successful British Harrier aircraft, which is considered to be a V/STOL aircraft because its normal mode of operation is to use a short takeoff run, when space is available, to greatly increase its payload and range capability.

The Soviet Union developed the world's second operational VTOL aircraft, the Yak 36. The aircraft features a lift-plus-lift cruise propulsion concept whereby the cruise engine has two fully vectorable nozzles for vertical and cruise flight and a pair of forward-located lift engines for lift and balance in vertical flight.

The success of the British Harrier led to the development of the Harrier II. This aircraft (**Fig. 3**) has numerous changes that provide approximately twice the range and payload of the original Harrier. The Harrier II (designated AV-8B by the U.S. Marine Corps) operates from both land and sea bases. In an amphibious operation, these aircraft fly attack and close-support missions from assault ships in support of troop-carrying helicopters. Once ashore, they are staged forward to austere sites close to the battlefront, so that they can react quickly in support of the ground troops.

Fig. 3. McDonnell Douglas/British Aerospace Harrier II aircraft.

The initial Harrier II was powered by an engine with a thrust of 21,450 lb (95,410 newtons), but in 1989 a new engine with an additional 3000 lb (13,344 N) of thrust became an option.

With the success of the Harrier, design studies and technology development programs were directed at expanding the flight envelope of this class of aircraft to include supersonic capability. The term STOVL (short takeoff vertical landing) is used to define this supersonic fighter-attack aircraft since the large benefits in payload and range of a short takeoff will be factored into the basic design. *See* CONVERTIPLANE.

W. P. Nelms

Bibliography. J. Albers, *NASA Rotorcraft Technology for the 21st Century*, AIAA Pap. 89-2066, 1989; J. W. Fozard, *The British Aerospace Harrier: Case Study in Aircraft Design*, AIAA Professional Study Series, 1978; *Proceedings of the International Powered Lift Conference*, Santa Clara, California, December 7–10, 1987; *Proceedings of the National Technical Specialists Meeting on Tactical V-STOL Aircraft*, New Bern, North Carolina, September 19–21, 1989; *Special Course on V/STOL Aerodynamics*, AGARD Rep. 710, 1984.

Vessel traffic service

A marine traffic management system, operated by a competent authority, for improving the safety and efficiency of vessel traffic and protecting the environment within designated service areas. A vessel traffic service (VTS) has the ability and the authority to interact with marine traffic and respond to developing traffic situations. It improves order and predictability on the waterway throughout its service area by using surveillance, communications, and standard operating procedures to collect, evaluate, and share information to support professional mariners' navigation decision making. Opportunities for human error are reduced through the provision of traffic information, traffic organization, and navigation-assistance services. Captains, mates, and marine pilots use the delivered vessel traffic services to direct and control the maneuvering of ships, tugs with tows, and passenger ferries.

Systems. A modern VTS system uses a variety of sensors to collect information and assemble a "traffic image" of all vessel movements, events, and conditions in its area of responsibility. These sensors may include surface radar, closed-circuit television cameras, wind speed and visibility monitors, tide and current gauges, and simple visual overlooks. The newest tool available to vessel traffic service is the automatic identification system (AIS). This consists of a shipborne device that automatically reports each ship's position, identification, destination, and other information needed for the shore-based authority to conduct traffic management. All information is gathered from these sensors and sent to the vessel traffic center using microwave or high-capacity communications lines. There, it is presented to the vessel traffic

service operators on a sophisticated, composite display so they may monitor the complete traffic situation and relay relevant information to the participating mariners. Information is typically sent to the mariner using dedicated voice radio channels. However, more vessel traffic services are coming to rely on data transfer using digitized messages instead of ship-to-shore voice communications. Information of general interest to all vessels may be sent at regular broadcast intervals. Information or advice directed at a specific ship is sent when the ship requests information or when the vessel traffic service feels the ship needs it. *See* CLOSED-CIRCUIT TELEVISION; MICROWAVE; PILOTING; RADAR.

Operations. Functions and capabilities vary among the world's more than 300 VTS systems, but their basic operational function is to provide three services: information, navigational assistance, and traffic organization.

Information service. Information service provides ships with essential information in time to contribute to the mariner's on-board navigational decision making. This may include the position, identification, or intentions of other ships, the state of the tide or current, the availability of space in an anchorage, the presence of a concentration of fishing or recreational vessels, or the vertical clearance under a bridge. It is typically information that the vessel traffic service has gathered and that is otherwise unavailable to ships.

Navigational assistance service. Navigational assistance service is provided to an individual ship when requested, or when deemed necessary to prevent an imminent disaster. The vessel traffic service provides appropriate advice to assist with on-board navigational decision making and monitors the navigation of the vessel, but never gives specific helm or engine orders. Instead, navigation assistance provides information on a vessel's ability to adhere to its intended course or route, information on its position relative to known locations or coordinates (longitude/latitude), or the distance and direction to landmarks or other objects. Navigation assistance may include advice relating to tracks or routes to be followed, or in an extreme case directions on action to take to prevent collisions and groundings. *See* MARINE NAVIGATION; NAVIGATION; SHIP ROUTING.

Traffic organization service. Traffic organization service prevents the development of dangerous maritime situations and provides for the safe and efficient navigation of traffic within the vessel traffic service area by managing vessel movements. It concerns the forward planning of movements and water-space management, and is particularly relevant in times of congestion or when the movement of special transports may affect the flow of other traffic. Monitoring traffic, enforcing adherence to governing rules and regulations, and communicating conditions and expectations to affected parties are essential elements of traffic organization. Traffic organization may include establishing and operating a system of traffic clearances in relation to the priority of movements, allocating space, and mandatory reporting of movements, as well as establishing routes to be followed, speed limits to be observed, or other appropriate measures considered necessary by the VTS authority. It may also include managing one-way traffic routes, isolating dangerous transits, scheduling movements, or prioritizing movements. Traffic organization is used to prevent vessels from meeting or overtaking each other or to ensure that vessel interactions occur at favorable locations. Where the vessel traffic service is authorized to issue traffic organization-related instructions to vessels, the instructions given are results oriented, leaving the details of execution to the vessel. The successful operation of a traffic organization regime requires established rules, good communications, positive identification of all affected vessels, and precision positioning.

A vessel traffic service also provides support to allied services including customs and law enforcement, firefighting and emergency responders, and commercial maritime interests in the port. This support is usually in the form of information on the position, identity, or intentions of vessels of interest.

Operators. In the United States, the Coast Guard has been named in law as the competent authority to operate vessel traffic services. As such, the vessel traffic service is an extension of the Coast Guard's authority to regulate marine traffic and direct the movement of vessels. This arrangement is consistent with other VTS operations throughout the world.

Exercising this considerable authority requires a dedicated, highly trained staff of professional VTS operators. Most VTS operators come to the profession after having served at sea for many years. They must first undergo extensive training in laws and regulations, local geography and operating practices, traffic monitoring, effective communications, and equipment use. Training often includes many hours of practice using simulated traffic situations. Once certified to perform the duties of a vessel traffic service operator, they are assigned a sector and a watch shift. Most vessel traffic services operate 24 h a day throughout the year.

Users. Vessel traffic services are aimed at assisting the professional mariner. This limits participation vessels of a certain size or of a certain employment. In most areas around the world, participation in a vessel traffic service is mandatory for commercial vessels greater than 270 metric tons (300 tons) or greater than 20 m (65 ft) in length. Passenger ferries are usually required to participate. Inland or harbor craft, such as tugs, may be excluded. Small fishing vessels and recreational craft rarely take part in the vessel traffic service. In some cases, these types of craft may be prohibited from participating to reduce the communications load on the system. In other areas, these craft may be allowed to participate as needed.

Routine. The process of participating in a vessel traffic service begins when a vessel checks in to inform the vessel traffic service of its position, identity, and intentions. The medium for communicating this information is usually VHF-FM voice radio. In its initial report to the vessel traffic service, a vessel will provide the vessel traffic service with a sailing plan that will serve as the vessel traffic service's

primary source of information to monitor the participating vessel and upon which it will base its traffic organization decisions. *See* RADIO BROADCASTING.

During the course of the vessel's transit through the area, the vessel traffic service will monitor its progress and adherence to the sailing plan. The vessel traffic service may occasionally provide information to the vessel relevant to opposing traffic, channel (fairway) conditions, weather, aids to navigation, or anything else that might affect the transit or influence the sailing plan. The vessel may wish to contact the vessel traffic service during its transit for updated information or to report that it must deviate from the agreed-upon sailing plan. Upon completing its transit, a vessel might make a final call to the vessel traffic service reporting "all secure" or to inform the vessel traffic service of its plans. J. Michael Sollosi

Bibliography. International Association of Marine Aids to Navigation and Lighthouse Authorities (IALA), *Vessel Traffic Services Manual*, 1998; International Maritime Organization Assembly Resolution A.857(20), *Guidelines on Vessel Traffic Services*, 1997; National Research Council, *Vessel Navigation and Traffic for Safe and Efficient Ports and Waterways: Interim Report*, National Academy Press, Washington, DC, 1996; U.S. Coast Guard Office of Vessel Traffic Management National Research Council, *Minding the Helm: Marine Navigation and Piloting*, 1994.

Vestibular system

The vestibular system subserves the bodily functions of balance and equilibrium. It accomplishes this by assessing head and body movement and position in space, generating a neural code representing this information, and distributing this code to appropriate sites located throughout the central nervous system. The centrifugal flow of information begins at sensory hair cells located within the peripheral vestibular labyrinth. These hair cells synapse chemically with primary vestibular afferent nerve fibers, causing them to fire with a frequency code of action potentials that include the parameters of head motion and position. These vestibular afferents, in turn, enter the brain and terminate within the vestibular nuclei and cerebellum. Information carried by the firing patterns of these afferents is combined within these central structures with incoming sensory information from the visual, somatosensory, cognitive, and visceral systems to compute a central representation of head and body position in space. This representation is called the gravito-inortial vector and is an important quantity that the central nervous system employs to achieve balance and equilibrium. *See* POSTURAL EQUILIBRIUM; BRAIN; NERVOUS SYSTEM (VERTEBRATE).

Vestibular function is largely reflex and unconscious in nature. We are thus often unaware of vestibular action. For example, while we are able to remain upright in complete darkness partially because of vestibular postural reflexes, and can be aware of our posture, the mechanisms responsible for this are unconscious or unrealized. Vestibular reflexes are also analogous to spinal reflexes, like the knee jerk reflex, because, there are only a few neurons or brain cells interposed in the pathway between sensory input and motor output. These vestibular reflexes originate within the vestibular labyrinth, which is subsequently connected with the eyes (vestibulo-ocular reflexes), the neck (vestibulo-colic reflexes), and the body (vestibulo-spinal reflexes). The vestibular system can influence movement, somatosensory and visual sensation, digestion, equilibrium, and state of mind. Vestibular activity modifies the firing of intrinsic cerebellar neurons, and influences spinal and brainstem motor and interneurons and superior collicular and cerebral cortical neurons, where, for example, the orientations of visual receptive fields are modified by changes in head position. *See* REFLEX.

The head can be envisioned as a sphere free to pivot upon the neck joints, which are analogous to ball and-socket joints. Thus, the head can move in three dimensions, and the appropriate labyrinthine sensors report signals that are measures of the angular velocity of the head in these three dimensions—pitch, roll, and yaw. Vertebrates also experience linear motion, for example, walking along a flat surface or climbing up and down, as well as the influence of gravity. The two components of linear acceleration—the linear acceleration vector that the brain needs to keep informed about—are the impulsive and gravitational components. Linear motion can be vertical or horizontal, and impulsive (during the initiation of linear motion) or static (gravitational). The labyrinth is thus divided into receptor areas that primarily detect angular and linear head motion.

Morphology of the vestibular labyrinth. The vestibular labyrinth is housed within the petrous portion of the temporal bone of the skull along with the cochlea, the organ of hearing (**Fig. 1**). [Anatomists named this bone the temporal bone, meaning the time bone, because this is where gray hair first appears, indicating the passage of time. Petrous refers to rock, and the petrous portion of the temporal bone is among the most dense and hard bone in the body, offering protection to the delicate cochlea and labyrinthine organs.] The receptor element or primary motion sensor within the labyrinth is the hair cell (**Fig. 2**). Hair cells respond to bending of their apical sensory hairs by changing the electrical potential across their cell membranes. These changes are called receptor potentials, and the apical surface of the hair cell thus functions as a mechanical-to-electrical transducer. Because hair cells are small with little surface area, they are isopotential throughout, and the apical receptor potentials are also sensed by the base of the cell. The cell base is a secretory organ that regulates the liberation of neurotransmitter in response to changes in the membrane potential. This neurotransmitter, probably glutamate, diffuses across the synaptic cleft to complex with chemical receptors on the peripheral or dendritic portions of fibers of the VIIIth cranial (vestibulocochlear) nerve (**Fig. 3**). The action of these chemical receptors, in turn, results in a depolarization of the membrane

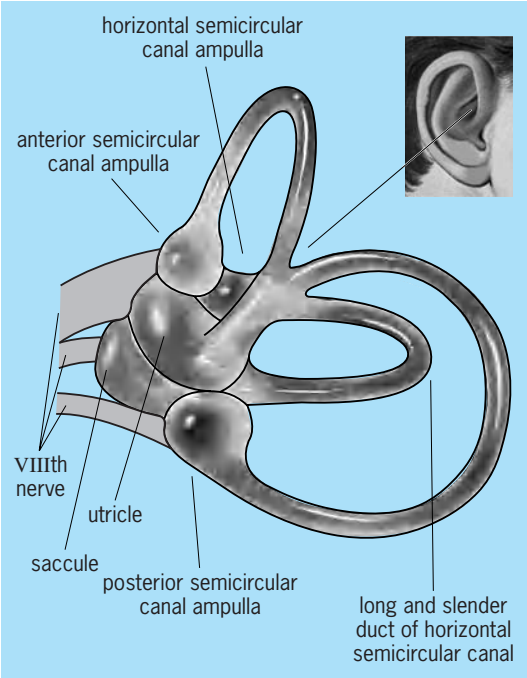

Fig. 1. The vestibular labyrinth is located within the inner ear. It communicates with the brain via the VIIIth nerve. Each of the three semicircular canals has an ampulla and a long and slender duct. The utricle primarily senses motion in an earth-parallel plane, while the saccule primarily senses motion and gravity in an earth-perpendicular plane.

of the VIIIth nerve fibers, causing the initiation of action potentials that are conducted toward the central nervous system. The frequency of this action potential train encodes the parameters of angular and linear motion. *See* BIOPOTENTIALS AND IONIC CURRENTS; EAR (VERTEBRATE); SYNAPTIC TRANSMISSION.

Hair cells are the common sensory element in both the angular and linear labyrinthine sensors as well as within the cochlea. The particular frequency of energy that hair cells sense within these diverse end organs arises because of the accessory structures surrounding the hair cells. Thus, angular motion is sensed by the semicircular canals, linear motion by the otolith organs, and sound energy by the cochlea.

Semicircular canals. The angular accelerometers, the semicircular canals, bilaterally three in number, are named for their orientation within the skull. There are two horizontal or lateral canals, two superior or anterior canals, and two inferior or posterior semicircular canals. These three canals on each side of the head are oriented in planes orthogonal to each other (Fig. 1). The three planes of space (x, y, and z or yaw, pitch, and roll) are represented by the three planes of the semicircular canals. The hair cells of each semicircular canal lie in a single sheet atop a saddle-shaped structure called the crista of the canal. The crista resides within a widened, egg-shaped portion of the canal called the ampulla (Fig. 2). Hair cells are divided into type I and II with reference to their innervation and cytoarchitectural features. The type I cell is flask-shaped with a constricted neck and a flared apical lip. The type II cell is more cylindrical. The apical surface of each cell faces a fluid with a high potassium concentration called endolymph and is covered with sensory hairs, or cilia, separated into stereocilia and a kinocilium. In the semicircular canals, a gelatinous cupula sits atop the hair cells, completely filling the lumen of the ampulla of the canal and closing it off to fluid flow. The general shape of the cupula conforms to the receptor area from which it arises. The cupula forms a diaphragm across the ampulla. It is most easily displaced at its center, where it is thinner. The hair cell kinocilium and the taller stereocilia project into the cupula and are displaced when the cupula moves. Because the cupula conforms to the shape of the crista, is attached to the ampullary wall on three sides, and is thinner in its central portions, the cilia in its central regions are subject to more displacement than those in the periphery for a given acceleration. A graded response to cupular displacement thus results. The apex of the cupula may act as a relief valve for

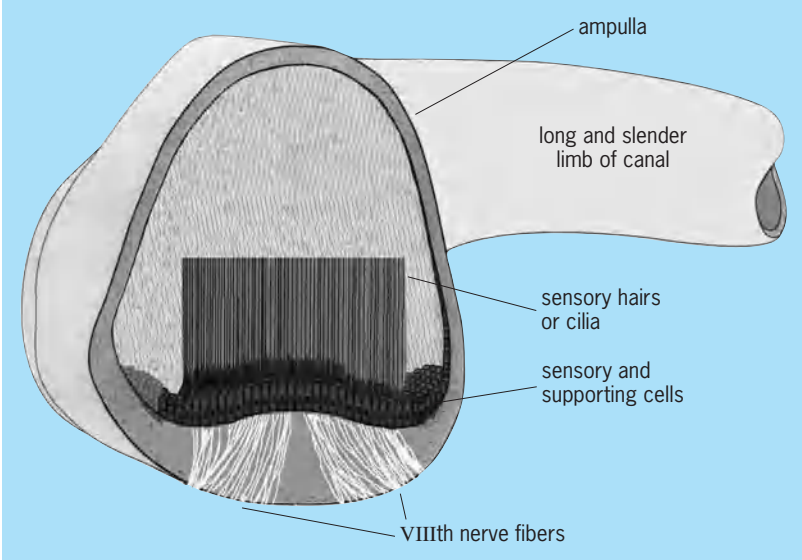

Fig. 2. Ampulla of a semicircular canal. This cross section is perpendicular to the long axis of the horizontal canal. The canal hair cell stereocilia are considerably longer than their macular counterparts, averaging about 0.1 mm in length.

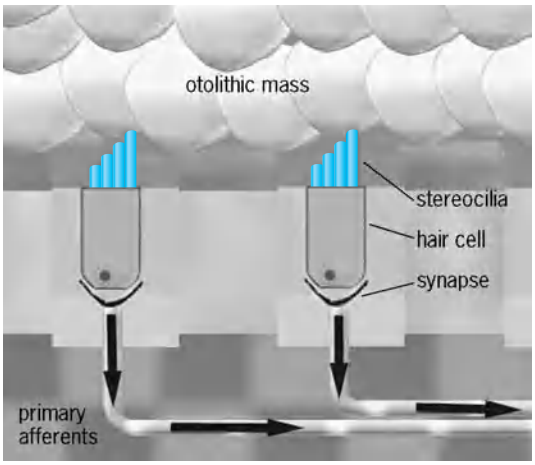

Fig. 3. Otolithic macula at rest. The arrows within the primary afferents or VIIIth nerve fibers indicate spontaneous activity in these fibers in the absence of motion of the otolithic mass relative to the hair cell stereocilia. (The otolithic membrane is not illustrated for clarity.)

extreme angular accelerational forces such as those caused by head trauma. The apical cupular surface may be displaced and ultimately detached only to reseal later. This mechanism may serve to prevent damage to receptor structures.

Otolith organs. The linear accelerometers, the otolithic organs (utricle and saccule), are bilaterally two in number. They are fluid-filled sacs containing a curved sheet of hair cells called the macula. The hair cell apical sensory hairs protrude through perforations in a membrane called the otolithic membrane to contact a mass of calcium carbonate crystals called the otolith (or stone) that is supported by the otolithic membrane (Fig. 3). The orientation of the utricular macula is roughly horizontal, and that of the saccule, roughly vertical.

Type I hair cells lie predominantly in the center of the crista and are concentrated in the striola of the macula. (The striola is a line running roughly down the center of the macula.) Type II cells are located peripherally in the crista and throughout the macula of the saccule and utricle. Type I hair cells are unique in that the afferent dendritic process of the bipolar ganglion neuron located in Scarpa's ganglion expands to form a calyx or chalice that completely surrounds the basolateral surface of the cell. Type II cells are innervated by finer fibers, and one afferent may contact up to 30–40 single hair cells. There are also the dimorphic afferents that contact both type I cells via a calyx ending and type II cells via a bouton ending. Considering the ultrastructure of the contacts between the hair cells and primary afferents, it is clear that the hair cell is the presynaptic element and the afferent the postsynaptic element within this chemical synaptic formation. Within the hair cells located at the basolateral surface opposite the afferent contact, there is almost always a typical rod or ribbon synaptic formation. In addition to the afferent synaptic processes, there are central nervous system efferent contacts arising from the efferent vestibular nuclei within the medulla and pons. These contacts terminate upon the calyx of the type I neuron and directly upon the type II hair cell and often upon its primary afferent.

The apical surfaces of vestibular hair cells are endowed with numerous cilia arranged in a characteristic fashion. The longest, the kinocilium, is eccentrically located immediately adjacent to rows of closely spaced stereocilia. The stereocilia become progressively shorter as they extend nearly to the opposite side of the cell. The hair cell is thus morphologically polarized, with the long kinocilium and tallest stereocilia at one end and the shortest stereocilia at the other end. The directional orientation of the hair cell is thus readily apparent and can be determined by noting the location of the kinocilium. The mapping of directional sensitivity has been determined for the crista and macula of the labyrinth. The least complex arrangement is found in the three semicircular canals, where the hair cells are aligned parallel to the long axes of their respective canals with the kinocilia all on the same side of the cells. In the two vertical canals, the anterior and posterior, the kinocilia are on the side opposite to the utricle, while in the horizontal canal the polarity is reversed with the kinocilia on the utricular side of the canal. In each otolithic macula the cell orientation is divided into directional hemifields separated by the striola, a line of opposed hair cells across the macular epithelium. The hair cells are oppositely oriented to each other in each hemifield.

Physiology of the vestibular labyrinth. Electrophysiological recordings have been taken from various labyrinthine elements. Elegant studies have been performed by placing microelectrodes into living hair cells, subjecting these cells to movement of their cilia, and noting changes in the membrane potential. The resting potentials of hair cells are usually between −50 and −60 mV. Hair cells respond to the movement of their cilia with transmembrane potential changes (receptor potentials) that are graded in amplitude, ranging 5–20 mV. These cells respond to static hair bundle deflections and to sinusoidal deflections with a frequency range of 0 to 150 Hz. Within this range, the output of the hair cell transduction apparatus depends on the hair bundle displacement only and not on the velocity of bundle motion (**Figs. 4** and **5**). The depolarizing phase of the receptor potential occurs upon bending of the hair bundle toward the kinocilium, and the hyperpolarizing response with bundle motion away from the kinocilium or toward the shortest stereocilium. In the semicircular canals the tallest cilia are inserted into the cupula. Estimation of the magnitude of cupular motion in the human is estimated not to exceed 3 micrometers at its extreme. Thus the sensitivity of these cells to bundle motion is extraordinary. The slope of the input/output curve can reach 20 mV per micrometer of displacement. (It has been estimated that the movement of the basilar membrane of the human cochlea is approximately equal to 1 micrometer at auditory threshold.) The depolarizing portion of the receptor potential is consistently accompanied by a decrease in input impedance of the cell.

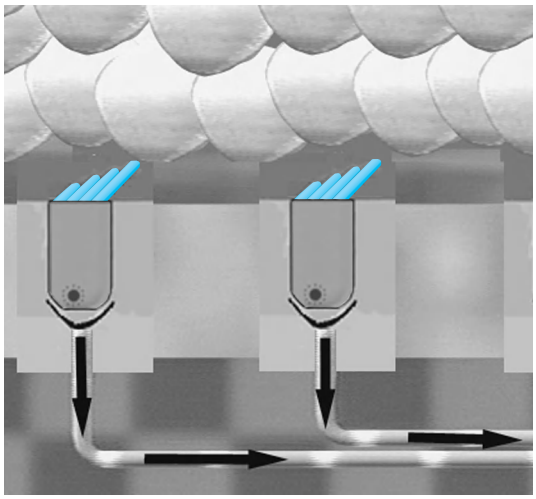

Fig. 4. Otolithic macula being stimulated by linear acceleration or gravity in its on-direction. The stereocilia are deviated toward the longest cilium (the on-direction for the hair cells). The large arrows within the primary afferents or VIIIth nerve fibers indicate increased activity in the presence of motion of the otolithic mass relative to the hair cell stereocilia.

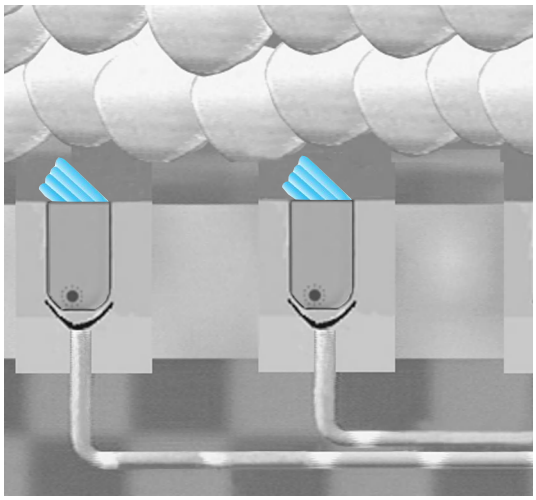

Fig. 5. Otolithic macula being stimulated by linear acceleration or gravity in its off-direction. The stereocilia are deviated away the longest cilium (the off-direction for the hair cells). The absence of arrows within the primary afferents or VIIIth nerve fibers indicates decreased or absent activity in the presence of off-direction motion of the otolithic mass relative to the hair cell stereocilia.

This suggests that the receptor potential of these vertebrate cells results from alteration of the membrane conductance to an ion whose reversal potential lies positive to the resting potential. The observation of the blocking effects of magnesium ion upon receptor potentials is consistent with a role for calcium in the transduction process. Potassium also certainly plays an important role in transduction. Thus it has been unequivocally demonstrated that hair cells can produce receptor potentials in response to the bending of their apical hairs.

The hair cell is presynaptic to the primary afferent fiber. Recording from this fiber should show modulation of any ongoing spike discharges depending upon the polarity of the receptor potential. This may be seen from single unit responses recorded within the VIIIth nerve.

Responses of the VIIIth cranial nerve. Let us consider vestibular afferents supplying the right and left horizontal semicircular canals. If an animal is placed upon a turntable and the head adjusted so that the position of the horizontal canals is parallel to the plane of rotation, rotation should modulate the receptor potentials and the horizontal canal primary afferent discharge rate. We begin rotation with a brief rapid buildup of angular acceleration, continue rotation at a constant velocity for some time, and then decelerate to a head-still position. Before stimulation begins, both horizontal canal nerves are firing with an equal resting discharge rate of around 100 impulses per second (I/S). Before considering the neural changes that might occur with stimulation, it is necessary to consider the mechanical events within the semicircular canals that result in the production of these changes. At the onset of angular acceleration the membranous labyrinth moves with the head as it is tethered to the skull. Movement of the head induces a differential pressure across the cupular membrane. This pressure causes endolymph flow relative to the canal, much as the pressure in a garden hose causes water flow out of the hose. The endolymph resists flow or movement because of its inertia, viscous drag between the fluid and canal wall, and the elastic restoring forces of the cupula and hair cell cilia. Then, during the initial angular acceleration there is a brief period of endolymph flow due to the motion of the canal and the relative stationariness of the fluid within the canal. This results in cupular deflection, say, in the utriculopedal direction in the right horizontal canal, leading to an increase in the firing rate of nerve fibers supplying the right side. Conversely, directly opposite actions take place on the left canal, effectively decreasing or even completely shutting down the firing rate of fibers supplying the left side. Moments later during the constant-velocity phase of rotation, as inertia is overcome the endolymph begins to catch up with surrounding structures, and there is no longer a relative endolymph flow acting against the cupula. Nevertheless, the displacement of the cupula continues for a brief period due to its inherent sluggishness. After a time, when all transient disturbances subside and angular velocity remains constant, the rate of firing of the nerve fibers returns to its prestimulus value. When the head is abruptly decelerated, this force acts upon the bilateral horizontal canals, causing reactions that are precisely opposite to those which occurred upon acceleration. Note that during deceleration inertia works to sustain motion in the endolymph while movement in other structures abruptly stops.

In summary, endolymph flow in a utriculopedal direction deforms the cupula, which then carries with it the cilia, causing a depolarizing receptor potential to be generated at the apex of the hair cell. This depolarization causes the basal or secretory portion of the hair cell to secrete more neurotransmitter, which increases the firing rate of the vestibular primary afferent. Reversing the direction of the endolymph flow hyperpolarizes hair cells, decreases ongoing transmitter release, and decreases the afferent firing rate to a value below the resting discharge rate. These responses are mirrored in the opposite semicircular canal. Thus acceleration and deceleration are signaled bidirectionally via a modulation of ongoing spontaneous neural activity. These actions are well correlated with the morphological polarity of hair cells in the lateral canals, where it is now known that bending of the cilia toward the kinocilium results in hair cell depolarization, whereas hyperpolarization results from bending the cilia away from the kinocilium or toward the shortest stereocilium. Just as the bilateral horizontal canals work in tandem, the right anterior and left posterior semicircular canals are roughly coplanar and work in tandem. The opposite is true for the right posterior and left anterior canals.

In the macula of the utricle or saccule, the adequate stimulus is impulsive linear acceleration or a change in the effective gravitational force caused by tilting of the head. As far as the hair cells are concerned, this is directly analogous to what happens within the crista of the semicircular canals except that the driving force, bending of the sensory hairs, is now caused by relative motion of the otolithic stone.

If the head is tilted in either pitch or roll, the otolith will tend to move relative to the head because of the effect of gravity upon its weight. This motion will impart a force on the hair bundles, causing them to bend and initiating the transduction cascade that reports this head tilt to the brain. If one experiences an impulse of linear acceleration such as beginning to walk, the inertia of the otolith will cause it to lag behind the head, thus bending the sensory hairs. When linear acceleration is complete, as in the constant-velocity phase of walking, the stone will return to its original position because of the elastic restoring forces of the otolithic membrane. Just as in the semicircular canals, these responses are bilaterally paired and signal appropriate bidirectional increases and decreases in impulsive and gravitational components of linear acceleration.

Unusual combinations of angular and linear motion can cause motion sickness or sea sickness. Terrestrial motion sickness is analogous to the space adaptation syndrome that afflicts astronauts during their sojourn into microgravity. In the case of space travel into microgravity, the weight of the otolith becomes small, nearly zero, while the mass of the otolith remains unchanged. Thus there are no responses to tilt while responses to impulsive linear accelerations persist. The semicircular canals function normally, as they have no gravitational responses. This unusual combination of impulses, that is, normal canal and impulsive linear inputs with no gravitational inputs traveling to the brain, might provoke a feeling of nausea and imbalance and confound the orientation and magnitude of the gravito-inertial vector, resulting in space adaptation syndrome or motion sickness. However, recall that there are additional sensory inputs that contribute to this vector. Astronauts report that being strapped into their couches may ameliorate their symptoms, just as sailors report that staring out at the horizon may ameliorate sea sickness.

Overview. The crista ampularis of the vestibular semicircular canals contains two types of receptor neurons or hair cells: Type I cells are flask-shaped and centrally located, while type II cells are more cylindrical and predominate in the periphery of the crista. Within the macula of the saccule and utricle, type I cells predominate in the striola where they are morphologically oppositely polarized, while type II cells are found throughout the macula. Both type I and II cells are presynaptic to axons of the vestibular portion of the VIIIth cranial nerve. Mammalian vestibular nerve fibers range in diameter from 1 to 10 micrometers and are usually grouped into two classes, thick and thin fibers. Thick fibers form a nerve calyx terminal expansion that surrounds the cell body of type I hair cells. From one to five hair cells contact these thick fibers. By contrast, the finer afferent fibers are innervated by many type II hair cells. The thick vestibular afferents discharge phasically, while the thin fibers discharge more tonically and generally have a lower sensitivity to natural stimulation. The phasic nature of the type I system is presumably due to afferent innervation by only a few hair cells. Their higher sensitivity may be ascribed to the character of the diaphragmatic displacement of the central portions of the cupula containing the cilia of the type I cells. The central portions of the cupula are thinner than the peripheral portions and undergo a greater displacement for a given head movement than the peripheral portions. The kinocilia of the type II cells located in the thicker, less mobile peripheral portions of the cupula probably endow the type II fine-fiber system with its low sensitivity to adequate stimulation.

Central projections of vestibular primary afferents. The primary afferents innervated by hair cells are the peripheral processes of bipolar neurons having cell bodies located in Scarpa's ganglion within the internal auditory meatus. The central processes of these cells contact neurons in the brainstem of the central nervous system. These centrally directed processes pass between the inferior cerebellar peduncle and the spinal trigeminal tract and bifurcate into rostrally and caudally directed fascicles which are then distributed within the vestibular nuclei and cerebellum. Thick and thin vestibular fibers terminate upon specific portions of the vestibular nuclei. A portion of the primary vestibular fibers project directly to the cerebellum via the juxtarestiform body and terminate ipsilaterally in the lingula, uvula, and nodulus. There are no direct projections to the cerebellar flocculus or paraflocculus. Second-order vestibular fibers arising from cell bodies within the vestibular nuclei also project to the cerebellum, including the flocculus and ventral paraflocculus, but do so bilaterally.

Vestibular nuclei. This complex is defined as the brainstem region where primary afferents from the labyrinth terminate. The nuclear complex is composed of four main nuclei: the superior, medial, lateral, and descending nuclei. The axonal projections of vestibular nuclear neurons travel to all parts of the neuraxis, including the brainstem, cerebellum, spinal cord, and cerebrum. Ascending fibers originating within the superior, medial, and ventrolateral nuclei travel within a fiber tract called the medial longitudinal fasciculus to terminate within the extraocular motor nuclei to subserve the vestibulo-ocular reflex. These fibers also have axon collaterals in other midbrain, pontine, and medullary nuclei that provide efference copy of the intended eye movement. The medial and ventrolateral nuclei also send fibers to the rostral extraocular motor nuclei that bifurcate, sending a branch via the descending medial longitudinal fasiculus to terminate upon cervical motoneurons and interneurons. These bifurcating fibers help to coordinate the vestibulo-ocular and vestibulo-colic reflexes to stabilize eye and head in space. Many of these vestibulo-ocular, vestibulo-spinal, and vestibulo-ocular-spinal neurons receive terminals of floccular and ventral parafloccular Purkinje neurons. In general, vestibular nucleus neurons maintain a high degree of spontaneous discharge. The motor outputs caused by the axons of vestibular neurons are thus under the modulatory control of the cerebellum. *See* MOTOR SYSTEMS.

The dorsal portion of the lateral vestibular nucleus is called Deiters' nucleus and is the origin of the ipsilateral descending vestibulo-spinal tract that terminates upon ipsilateral extensor motoneurons at several levels of the spinal cord. Dorsal Deiters' neurons are the recipient of heavy input from Purkinje cells residing within the cerebellar vermal anterior lobe. This lateral vestibulo-spinal tract and nucleus subserves the vestibulo-spinal reflexes that help maintain postural stability. The excitatory drive upon extensor motoneurons provided by the lateral vestibulo-spinal tract is counterbalanced by a similar input to flexor motoneurons that originates from the red nucleus. The descending nucleus and the caudal portions of the medial vestibular nucleus have a network of back-and-forth connections with the cerebellum. These pathways undoubtedly play a role in maintaining posture by modulating the impulse traffic that enters into motor pathways.

Efferent vestibular system. In all vertebrates, including the earliest evolved such as lamprey or myxine (hagfish), there is an efferent system that originates from cell bodies within the central nervous system and terminates upon labyrinthine hair cells and primary afferents. In primates these efferent vestibular cell bodies reside within the efferent vestibular nuclei adjacent to the abducens nuclei at the pontomedullary junction. Electrical stimulation of this efferent system causes excitation in vestibular primary afferents but inhibits hair cells. The efferent vestibular system is presently a subject of intense study but undoubtedly is in place to enhance vestibular function. It is interesting that evolution felt it necessary to modify incoming vestibular information before it could enter the central nervous system.

Stephen M. Highstein

Bibliography. P. Blazquez et al., Input of the anterior and posterior semicircular canals via interneurons carrying head velocity information to the dorsal Y group of the vestibular nuclei, *J. Neurophysiol.*, 83:2891–2904, 2000; R. Boyle et al., Neural readaptation to Earth's gravity following return from space, *J. Neurophysiol.*, 86:2118–2122, 2001; S. M. Highstein, R. D. Rabbitt, and R. Boyle, Determinants of semicircular canal response dynamics in the toadfish, *Opsanus tau*, *J. Neurophysiol.*, 75:575–596, 1996; R. A. McCrea, A. Strassman, and S. M. Highstein, Anatomical and physiological characteristics of vestibular neurons mediating the vertical vestibulo-ocular reflex of the squirrel monkey, *J. Comp. Neurol.*, 264:571–594, 1987; R. B. Silver et al., Examination of the cupula and stereocilia of the horizontal semicircular canal in the toadfish, *Opsanus tau*, *J. Comp. Neurol.*, 402:48–61, 1998.

Vestimentifera

A group of benthic marine annelid worms, discovered in 1969, that is restricted to habitats rich in sulfide (for example, hydrothermal vents and sulfide seeps). As adults they lack a mouth, gut, and anus, and are nourished by internal symbionts. All vestimentiferans live in tubes of varying hardness and rigidity (see **illustration**). Tube material is secreted by internal glands and is a mixture of chitin and protein. *See* HYDROTHERMAL VENT.

Classification. In the past, vestimentiferans had been considered to be members of the phylum Pogonophora as well as of the phylum Annelida, and it even had been suggested that they are a phylum in their own right. This was largely resulted from emphasis on their extraordinary anatomy rather than focusing on their marked similarities with tube-dwelling annelids, such as sabellid and serpulid tubeworms. They are now considered to be a part of the annelid family Siboglinidae along with other members of the former phylum Pogonophora and the bone-eating worms *Osedax*. *See* ANNELIDA.

Taxonomy. The previously erected taxonomy within Vestimentifera has been superseded following the placement of the group inside Siboglinidae. Currently there are 10 genera containing a total of 15 species: *Alaysia*, *Arcovestia*, *Escarpia*, *Ridgeia*, *Riftia*, *Lamellibrachia*, *Oasisia*, *Paraescarpia*, *Seepiophila*, and *Tevnia*. Further study is needed to establish whether other groupings apart from genera are needed.

Morphology. Vestimentiferans have four regions along their length (see illustration). The anteriormost, the obturacular region, has a mass of vascularized filaments supported by a paired structure, the obturaculum. Normally these branchial filaments protrude from the opening of the tube and act as a gill, allowing for exchange of dissolved substances between the worm's body and seawater. There may be as many as 230,000 filaments in this branchial plume.

The second region, the vestimentum, is muscular and serves to maintain the plume of branchial filaments in the open seawater by pressure on the inner tube surface at its opening. The vestimentum is the site of many glands that contribute material for lengthening the tube and thickening it near the opening. Laterally and dorsally, a pair of elongated flaps overlap one another and form a chamber that communicates with the seawater. The genital pores, both female and male, open into this chamber, and eggs and sperm pass through it.

The third region, the trunk, is a single segment and constitutes about 75% of the total length of the worm. It has a pair of large longitudinal blood vessels; the dorsal of these returns deoxygenated blood and carbon dioxide to the anterior branchial plume and, with its muscular wall, functions as a heart. The ventral vessel brings oxygenated blood from the plume to the body of the worm. The bulk of the trunk is made up of the trophosome, an organ containing masses of sulfide-oxidizing bacteria. Gonads are also present in the trunk. Much of the body wall of the trunk is lined internally by longitudinal muscles that allow the worm to withdraw into its tube for protection from predators. Glands secreting tube material add to the thickness of much of the tube posterior to the opening.

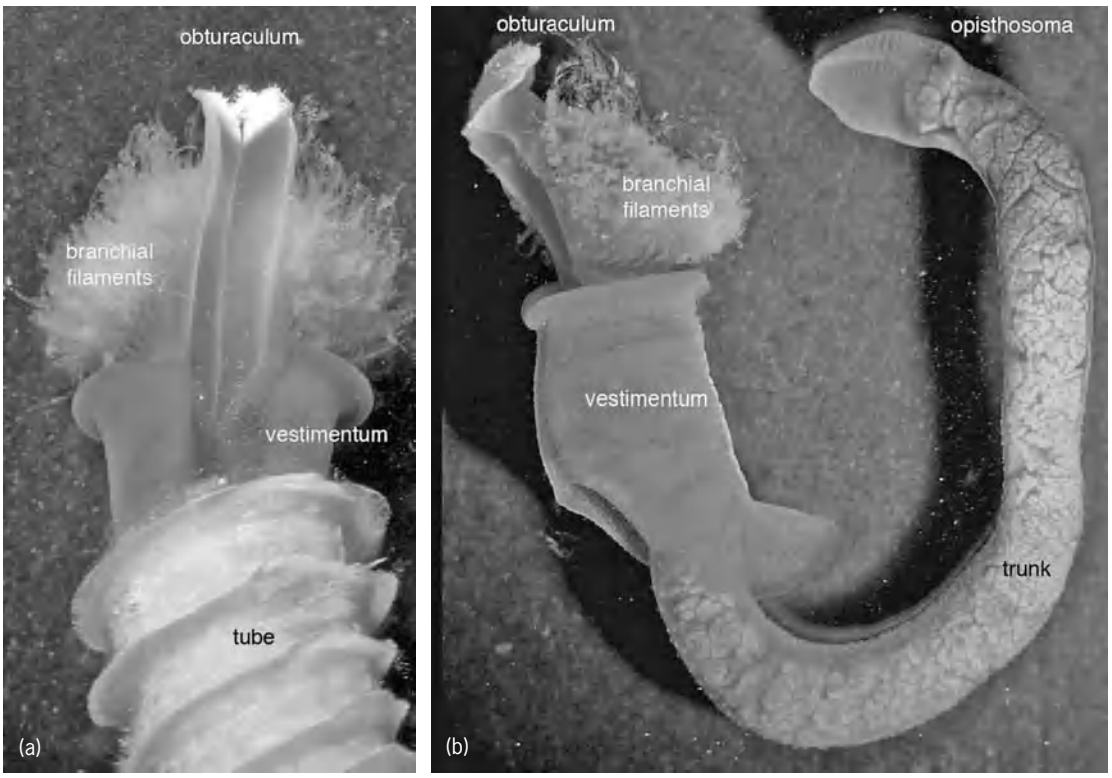

Tevnia jerichoana. **(*a*) In tube. (*b*) Removed from tube, showing four body regions.**

The fourth and most posterior region, the opisthosoma, is made up of many segments, each segment in the anterior portion bearing a row of small hooks that can be set into the inner surface of the tube. This provides an anchor against which the body of the worm retracts when the longitudinal muscles of the trunk contract during withdrawal into the tube.

Nutrition. The blood of vestimentiferans is red owing to hemoglobin, similar to that of other annelids. Oxygen and sulfide are transported by the hemoglobin, and are bound so tightly that they cannot interact as they would in normal circumstances. Thus, the symbiotic sulfide-oxidizing bacteria are able to carry out their chemoautotrophic mode of nutrition utilizing the sulfide and oxygen. The oxygen also supports the metabolism of the worm. There is no mouth, gut, or anus in postlarval worms, and their nutrition relies solely on the symbiotic bacteria contained in the bacteriocytes (specialized cells) of the trophosome. Certain cells of the worm harvest bacteria by breaking down bacteriocytes, and the food is passed on for the growth and metabolism of the worm. *See* BACTERIAL PHYSIOLOGY AND METABOLISM; HEMOGLOBIN.

Nervous system. The brain is present in the anteroventral portion of the vestimentum. The ventral nerve cord extends from the brain, separates around the ventral ciliated field, reunites, and continues posteriorly to the posterior tip of the opisthosome. Giant axons are associated with the ventral nerve cords, allowing the organism to enhance its withdrawal response into the tube. Nerve tracts pass anteriorly from the brain to each of the branchial filaments; they appear to coordinate the contractions of a series of ring muscles around the blood vessels of the branchial filaments and help move blood through the filaments back to the central region of the blood system. Other nerve tracts pass to the brain from certain sensory filaments associated with the branchial filaments.

Digestive, excretory, and reproductive systems. An excretory organ comprising a pair of nephridia, as in many other tubicolous annelids, is situated ventral and anterior to the brain. It consists of a mass of ciliated tubules that appear to unite and ultimately open into the paired excretory ducts. The ducts may further fuse to a single opening to the exterior.

A transitory digestive system has been demonstrated in one vestimentiferan species and is likely to be found across the group. Late larval stages exhibit an anterior, ventral–medial process. This snout-like structure has an opening (mouth) surrounded by cilia and leads to a continuous (intestinal) cavity proceeding to a terminal anus. When the larvae reach 400 μm long, the snout and the anus are no longer present. Only the former gut remains; however, it is isolated and contains bacteria. It is believed that when the formerly planktonic larvae settle in an area that contains high levels of sulfide, they feed on benthic bacteria, among which are sulfide-oxidizing bacteria. The sulfide-oxidizing bacteria appear to be phagocytized by cells of the midgut epithelium, where they grow and multiply. When the mouth and anus close, the midgut, with its associated bacteria, increases in size and complexity and becomes the trophosome.

Reproductive system. There are separate sexes in Vestimentifera. Mature eggs range 75–110 μm in

diameter and are the result of oogenesis in paired anterior ovaries. They pass posteriorly, then anteriorly, through paired oviducts to emerge through paired gonopores on the dorsal surface of the vestimentum.

Spermatogenesis is initiated in the testicular epithelium of tubules that open into paired sperm ducts. As many as 1500 spermatozoa may be associated with a single cytophore in the anterior portion of the sperm ducts. Near the gonopore, spermatozoa are free of the cytophore and are arranged in bundles. Spermatozoa are shed in bundles (spermatozeugmata) through paired gonopores on the dorsal surface of the vestimentum.

Larval development. Fertilization of eggs is internal in the oviducts. Sperm either transferred directly from a male or gathered from the seawater are stored in a special region of each oviduct known as a spermatheca. The eggs are fertilized as they pass through the spermathecae and are then spawned into the water and develop into trochophore larvae. Branchial filaments appear shortly thereafter. Opithosomal chaetae develop in the posterior segments. The vestimentum is formed when the larval length is about 400 μm. The obturaculum is 1 mm long. Internal bacteria are consistently present at larval lengths greater than 285 μm. *See* HYDROTHERMAL VENT. Meredith L. Jones; Greg Rouse

Bibliography. C. R. Fisher, Chemoautotrophic and methanotrophic symbioses in marine invertebrates, *Rev. Aquatic Sci.*, 2:399–436, 1990; A. Hilario, C. M. Young, and P. A. Tyler, Sperm storage, internal fertilization, and embryonic dispersal in vent and seep tubeworms (Polychaeta: Siboglinidae: Vestimentifera), *Biol. Bull.*, 208:20–28, 2005; M. L. Jones, The giant tube worms, *Oceanus*, 27:47–54, 1984; G. W. Rouse, A cladistic analysis of Siboglinidae Caullery, 1914 (Polychaeta, Annelida): Formerly the phyla Pogonophora and Vestimentifera, *Zool. J. Linn. Soc.*, 132:55–80, 2001; A. Schulze and K. M. Halanych, Siboglinid evolution shaped by habitat preference and sulfide tolerance, *Hydrobiologia*, 496:199–205, 2003; M. Webb, *Lamellibrachia barbami*, gen. nov. sp. nov. (Pogonophora), from the northeast Pacific, *Bull. Mar. Sci.*, 19:18–47, 1969.

Vesuvianite

A sorosilicate mineral of complex composition crystallizing in the tetragonal system; also known by the name idocrase. Crystals, frequently well formed, are usually prismatic with pyramidal terminations (see **illus.**). It commonly occurs in columnar aggregates but may be granular or massive. The luster is vitreous to resinous; the color is usually green or brown but may be yellow, blue, or red. Hardness is $6^1/_2$ on Mohs scale; specific gravity is 3.35–3.45.

The composition of vesuvianite is expressed by the formula $Ca_{10}Al_4(Mg,Fe)_2Si_9O_{34}(OH)_4$. Magnesium and ferrous iron are present in varying amounts, and boron or fluorine is found in some varieties. Beryllium has been reported in small amounts.

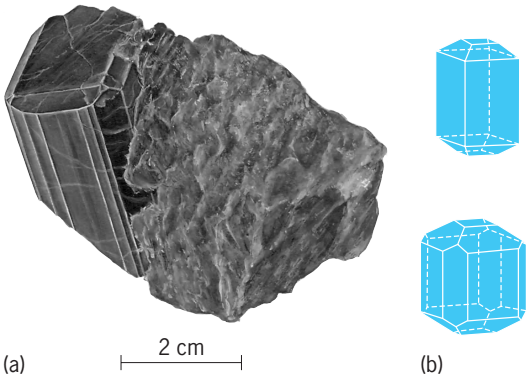

Vesuvianite. (*a*) Crystal, Christiansand, Norway (*specimen from Department of Geology, Bryn Mawr College*). (*b*) Crystal habits (*after C. Klein and C.S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., 1993*)

Vesuvianite is found characteristically in crystalline limestones resulting from contact metamorphism. It is there associated with other contact minerals such as garnet, diopside, wollastonite, and tourmaline. Noted localities are Zermatt, Switzerland; Christiansand, Norway; River Vilui, Siberia; and Chiapas, Mexico. In the United States it is found in Sanford, Maine; Franklin, New Jersey; Amity, New York; and at many contact metamorphic deposits in western states. A compact green variety resembling jade is found in California and is called californite. *See* SILICATE MINERALS. Cornelius S. Hurlbut, Jr.

Vetch

Any of a group of plants which are mostly annual and perennial legumes with weak viny stems often terminating in tendrils. There are about 150 species in the temperate zones of four continents. The leaves are compound with many leaflets. Vetches are used mainly for green manure, cover crops, hay, and pasture and silage. The seeds are used as concentrate in animal feeds, and some vetches are used as a vegetable for human consumption. Cool temperatures promote best development. Seed is sown in the fall in the Southern United States and in early spring in the Northern states. In general, are vetches more tolerant of acidic soil conditions than are most legume crops, but they have a relatively high requirement for phosphorus. Inoculation of seed with nodule-forming bacteria for symbiotic nitrogen fixation is necessary if vetch has not been previously grown in the field. *See* LEAF; LEGUME.

Traditionally vetch species have been grown in mixes with cereals for use as fodder. Although it was formerly thought that vetch species were self-fertile, recently it has been determined that some of the species are not self-fertile and actually benefit from bees or other pollinating insects for seed production. Several of the species are considered good plants for bee honey production. Identification of vetches is difficult until pods and seeds develop. *See* POLLINATION.

United States species. By the 1960s, some 35 vetch species and subspecies had been found in the United States. Many of these species have been introduced from Europe, the Far East, and Asia, where they had been used for several centuries for forage and as vegetables. Of these 35 species, 16 are native to the United States.

Vicia villosa. The vetch complex *Vicia villosa* is the most widely grown. The subspecies of this complex (*V. villosa* ssp. *villosa*), commonly known as hairy vetch, is the most winter-hardy and is mostly adapted to the eastern and southern United States, where it is grown as a winter annual or biannual. A second subspecies of this group is woollypod or smooth vetch (*V. villosa* ssp. *varia*). It is less winter-hardy than the hairy vetch type and is adapted to the Pacific Coast states, where it grows as a winter annual, and the southern United States. The first variety of woollypod vetch, Lana, was developed in 1961 in California. Recently new varieties, such as Cepello and Haymaker Plus, have been developed in Australia and are imported for use in California. Seed lots of hairy vetch that are sold often have mixtures of hairy and woollypod or smooth vetch. *Vicia villosa* ssp. *pseudocracca* occurs as a naturalized type in the northeastern United States.

Vicia sativa. The vetch complex *V. sativa* is the second most important species in use in the United States. In its native habitat of Europe and the Mediterranean regions, this vetch is known to occur as five subspecies. Common vetch (*V. sativa* ssp. *sativa*) is used throughout the United States. However, due to its it low winter hardiness, it is grown as a winter annual in the southern and western states and as a spring-summer annual in the northern states. Narrow-leaf or blackpod vetch (*V. sativa* ssp. *nigra*) occurs mostly as a weed in waste places in the United States. It spreads by volunteering and makes an excellent pasture and covercop. Underground vetch (*V. sativa* ssp. *amphicarpa*) produces pods above and below the ground and is being studied for revegetation on marginal lands of the Far East.

Vicia faba. Commonly known as the faba bean, horsebean, or broadbean, *V. faba* produces coarse, upright plants with large leaves and pods. It is one of only several vetches that are important sources of food for human beings. Some varieties are grown as vegetables. Frequently dried, faba beans have been a part of cuisines around the Mediterranean basin, and the Chinese have eaten them for more than 5000 years. It is also used in concentrate feeds for livestock in the Mediterranean basin.

This species can be divided into two subforms based on seed size. The horsebean is the large-seeded form (*V. faba* var. *faba*), and the tick or bell bean (*V. faba* var. *minuta*) is the small-seeded form, which is often used as a green manure-covercrop in the warmer regions of the United States during the winter and spring. In the colder regions, it is used during the spring and summer. In the Mediterranean basin the small-seeded form is used as a stock feed as well as a vegetable.

Fig. 1. Purple vetch in flower. (*USDA*)

Other species. Purple vetch (*V. bengalensis*) possesses poor winter hardiness; however, it is useful in the California rice-growing area, where it is used as a soil-improving crop in rotation with the rice (**Fig. 1**). Hungarian vetch (*V. pannonica*), a more winter-hardy type from Central Europe, is grown in the Pacific Northwest for forage, green manure, and seed.

One-flowered vetch (*V. articulata*) has fine leaves and stems but, lacking winter hardiness, is confined to warm regions. It can be grown under the same conditions as common vetch. Plants are smooth or nearly so, and the light-lavender-colored flowers are borne singly. Its flat seeds distinguish it from its nearest look-alike, bard or barn vetch (*V. monantha*), which is very restricted in adaptation. Bitter vetch (*V. ervilia*) is an erect vetch without tendrils. Although it is mainly found in southern Europe, it has been grown in the Pacific Northwest for use as stock feed in small quantities, for it can be toxic to livestock in larger quantities.

Bigflower or large yellow vetch (*V. grandiflora*), an introduced annual that has naturalized in the southern states, the western seaboard states, and the Great Lakes states, has promise as winter cover and reseeds freely. Woodford, a cultivar, was released in Kentucky for use on less-developed (rough) pastures.

Tiny vetch (*V. hirsuta*) is an introduced annual that has naturalized along the southern states bordering the Mississippi River, Texas, the eastern seaboard states, the Great Lakes states, and the Pacific Coast states.

Narbonne or French vetch (*V. narbonensis*), a possible wild relative of *V. fava*, is being studied in Australia, Europe, and the Middle East for use as a food, feed, and fodder crop. It has been cultivated as a forage crop in the eastern United States and has locally established as a weed in Maryland and the District of Columbia.

Weeds. Bird or purple-white tufted vetch (*V. cracca*) is a winter-hardy perennial, and its subspecies, cow or bramble vetch (*V. cracca* ssp. *tenuifolia*), is considered to be an invasive weed in the eastern United States and parts of Canada. This species complex was introduced for erosion control on cutover and burnt-over areas. Four-seeded, sparrow, or lentil vetch (*V. tetrasperma*), also considered an invasive weed, is an introduced annual that has naturalized along and east of the Mississippi River and in the Pacific states.

Crown vetch. This plant, *Coronilla varia*, is a long-lived, winter-hardy perennial legume, but it is not a true vetch. It spreads by seeds and rhizomes to form a dense, weed-free, erosion-resisting ground cover. Its greatest use is for erosion and weed control on unmowed slopes of highways, industrial developments, and military installations. Its forage value is being studied, Crown vetch has wide adaptation to variations in climate and soils. It is highly regarded for its resistance to fire. Rose, pink, and white blossoms enhance its attractiveness throughout most of the growing season (**Fig. 2**). It thrives best on well-limed and well-drained soils, requires no mowing, and rarely needs fertilization. Plantings are made with crowns (roots) or with seed that require special bacterial inoculation.

Crown vetch is slow to germinate and become established. Diseases and insects are of minor importance. Neither spittlebug nor weevil attacks crown vetch. Grazing animals do not bloat on it. Seed originally was produced only in Pennsylvania, starting in 1946; now it is grown in several states, including Ohio, New York, Nebraska, Iowa, and West Virginia. The first variety named and released was Penngift; later the Emerald (Iowa) and the Chemung (New York) varieties were named and released.

Some research has shown the palatability of crown vetch to be low except while growth is young and tender. In other grazing trials, animals were slow to accept it, but after a few days their performance on crown vetch was comparable to that on other common grass-legume pastures. Chemical analysis of crown vetch hay has shown that its crude protein and crude fiber content is similar to that of other legume hays. Digestible dry matter of crown vetch hay is below that of other grass-legume hays harvested at the same stage of maturity. Crown vetch hay is often difficult to wilt and cure. Its forage value is being studied. *See* COVER CROPS; SOIL CONSERVATION.

Fig. 2. Crown vetch. (*a*) In flower. (*b*) In seed. Each segment of a pod contains a seed.

Seed. Vetch seed is produced commercially along the Pacific Coast and in the southern Great Plains. and the South. Although most of the demand for vetch seed is supplied by production in the United States, some seed (of woollypod vetch) is imported from Australia, and seed of broad and tick beans is imported from Canada. At one time, much hairy vetch seed was imported from Europe. Vetch is grown alone and with small grains which give mechanical support to the weak stems of the vetch. Seed retains vitality for 5 years or longer. *See* SEED.

Diseases and insect pests. Diseases of vetch include anthracnose, false anthracnose gray mold, downy mildew, black stem, stem rot, root rot, rust, ovularia leaf and stem spot, several mosaic viruses, and septoria scald. Root-knot nematodes attack all varieties. Control is aided by using uninfected seed of resistant varieties on clean land. *See* NEMATA; PLANT PATHOLOGY.

Vetch is attacked by many of the insect pests of alfalfa, clover, and other forage legumes, including the pea aphid, cutworm, corn earworm, fall armyworm, vetch bruchid, grasshopper, lygus bug, and leafhopper.

Various hybrids of common vetch have been developed to be resistant to the bruchid insect and certain nematodes. *See* APHID; ENTOMOLOGY, ECONOMIC; HEMIPTERA; INSECTA.

Estimating nitrogen contribution. With the current interest in using sustainable agriculture methods to produce food in the United States, the vetches have taken on a renewed importance for those farmers who want to use legumes for crop rotation and green manure. Some of the vetch types like woollypod and hairy vetch show the ability to fix high amounts of nitrogen through their root symbiosis with nitrogen-fixing bacteria. Since these vetches and other types are being used to enrich the soil to help other crops, it is useful to be able to estimate the above nitrogen production of these vetches. In the 1950s, a method was developed by researchers at the University of Davis that is still in use today. The method involves sampling the aboveground fresh weight of the vetch crop, from which it is possible to estimate the amount of nitrogen in the sample. *See* NITROGEN FIXATION; AGRICULTURAL SOIL AND CROP PRACTICES.

Fred V. Grau; Walter Graves

Bibliography. J. A. Duke, *Handbook of Legumes of World Economic Importance*, Plenum, New York, 1981; R. R. Henson and H. A. Schoth, Vetch Culture and Uses, *USDA Farmers' Bull.*, no. 1740, 1955; F. J. Herman, Vetches in the United States, Native Naturalized and Cultivated. *USDA Agr. Handb.*, no. 168, 1960; J. H. Wiersema, J. H. Kirkbride, Jr., and C. R.

Gunn, Legume (Fabaceae) Nomenclature in the USDA Germplasm System, *USDA ARS Tech. Bull.*, no. 1757, 1990; W. A. Williams and W. L. Graves, Nitrogen contribution of annual legumes, in *Agron. Prog. Rep.*, no. 266, University of California Agricultural Experiment Station and Cooperative Extension, 1999.

Vibration

The term used to describe a continuing periodic change in the magnitude of a displacement with respect to a specified central reference. The periodic motion may range from the simple to-and-fro oscillations of a pendulum, through the more complicated vibrations of a steel plate when struck with a hammer, to the extremely complicated vibrations of large structures such as an automobile on a rough road. Vibrations are also experienced by atoms, molecules, and nuclei.

A mechanical system must possess the properties of mass and stiffness or their equivalents in order to be capable of self-supported free vibration. Stiffness implies that an alteration in the normal configuration of the system will result in a restoring force tending to return it to this configuration. Mass or inertia implies that the velocity imparted to the system in being restored to its normal configuration will cause it to overshoot this configuration. It is in consequence of the interplay of mass and stiffness that periodic vibrations in mechanical systems are possible.

Mechanical vibration. This is the term used to describe the continuing periodic motion of a solid body at any frequency. When the rate of vibration of the solid body ranges between 20 and 20,000 Hz, it may also be referred to as an acoustic vibration, for if these vibrations are transmitted to a human ear they will produce the sensation of sound. The vibration of such a solid body in contact with a fluid medium such as air or water induces the molecules of the medium to vibrate in a similar fashion and thereby transmit energy in the form of an acoustic wave. Finally, when such an acoustic wave impinges on a material body, it forces the latter into a similar acoustic vibration. In the case of the human ear it produces the sensation of sound. Thus, vibrations in solid and fluid bodies are essential to production, transmission, and reception of sound. *See* MECHANICAL VIBRATION.

One degree of freedom. Systems with one degree of freedom are those for which one space coordinate alone is sufficient to specify the system's displacement from its normal configuration. It is the simplest yet the most fundamental type of vibration system. An idealized example known as a simple oscillator consists of a point mass m fastened to one end of a massless spring and constrained to move back and forth in a line about its undisturbed position (**Fig. 1**). Although no actual acoustic vibrator is identical with this idealized example, the actual behavior of many vibrating systems when vibrating at low frequencies is similar and may be specified by giving values of a single space coordinate. They include loudspeaker cones, telephone diaphragms, microphone diaphragms, and drum membranes. *See* DEGREE OF FREEDOM (MECHANICS).

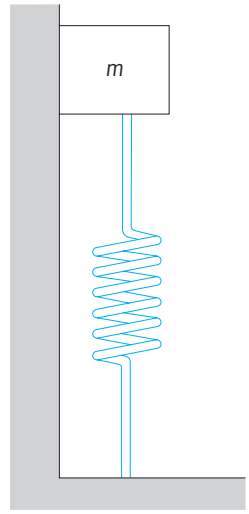

Fig. 1. Simple oscillator.

When the restoring force of the spring of a simple oscillator on its mass m is directly proportional to the displacement x of the latter from its normal position, the system vibrates in a sinusoidal manner called simple harmonic motion. This motion is identical with the projection of uniform circular motion on a diameter of a circle and is represented by the equation $x = A \sin 2\pi ft$. The frequency of vibration f is given by Eq. (1), where s is the constant of proportionality between force and stretch or compression of the movable end of the spring. The constant A represents the amplitude of the vibration, that is, the maximum displacement of mass on either side of its rest position. The magnitude of A is determined by the manner in which the motion is initially started. Note that the frequency of vibration of a simple oscillator is independent of the amplitude of its vibration. *See* HARMONIC MOTION.

$$f = \left(\frac{1}{2\pi}\right)\sqrt{s/m} \tag{1}$$

The variation in velocity v of the point mass of a simple oscillator is given by Eq. (2) and acceleration a is given by Eq. (3). Note that for a given displacement amplitude A, the velocity amplitude $2\pi fA$ is directly proportional to the frequency, and the acceleration amplitude $4\pi^2 f^2 A$ is directly proportional to the square of the frequency. Consequently, large velocity and acceleration amplitudes of a simple oscillator are more readily obtained at high than at low frequencies.

$$v = \frac{dx}{dt} = 2\pi fA \cos 2\pi ft \tag{2}$$

$$a = \frac{d^2x}{dt^2} = -4\pi^2 f^2 A \sin 2\pi ft \tag{3}$$

Energy considerations. The total mechanical energy of a simple oscillator equals the sum of the kinetic energy associated with the moving mass and the

potential energy stored in the distorted spring. When no frictional forces are present, these two types of energy are continuously being converted from one to the other and back again as the vibration cycle progresses. Their sum remains constant at a value given by Eq. (4).

$$E = \frac{sA^2}{2} = 2m(\pi f A)^2 \qquad (4)$$

When frictional or other dissipative forces are present, the simple oscillator gradually changes its mechanical energy of vibration into heat energy with an attendant reduction in its amplitude of vibration. No one simple law is capable of describing the variation of frictional forces as a material body vibrates. However, in many important cases the frictional force is opposite to the direction of motion and proportional to the velocity of the vibrating mass. For this type of damping force, the amplitude of vibration of the simple oscillator decreases exponentially with time in accordance with the equation $A = A_0 e^{-\alpha t}$, where α is a constant directly proportional to the frictional force and inversely proportional to the mass of the oscillator (**Fig. 2**). The natural frequency of free oscillation of a damped oscillator is slightly less than that of the same oscillator system without damping. The amount by which the frequency is lowered increases with increased damping. *See* DAMPING.

External driving force. A simple oscillator, or some equivalent system, is often maintained in a condition of steady vibration by the application of a periodic external driving force. After a sufficient time has elapsed, any natural free-oscillation frequency of the oscillator dies out and it vibrates solely at the frequency of the impressed driving force. When the applied driving force is sinusoidal, as represented, for example, by the function $F(\cos 2\pi ft)$, the forced vibration ultimately reaches a steady-state amplitude of $A = F/(2\pi f Z)$, where Eq. (5) applies. Z is known

$$Z = \left[R^2 + \left(2\pi f m - \frac{s}{2\pi f}\right)^2\right]^{1/2} \qquad (5)$$

as the mechanical impedance of the oscillator. From this equation it is apparent that the driven oscillator will have a maximum amplitude when the forcing frequency is near $[1/(2\pi)]\sqrt{s/m}$, the free oscillation

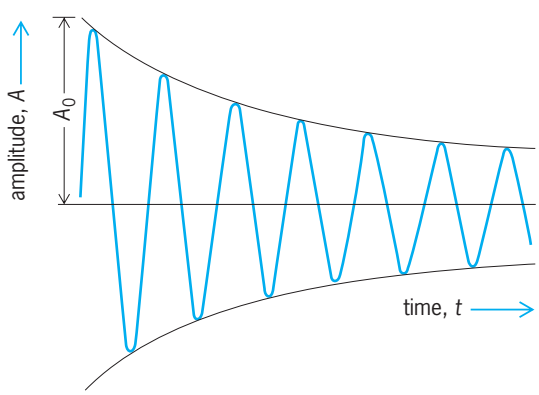

Fig. 2. Damped vibration of simple oscillator.

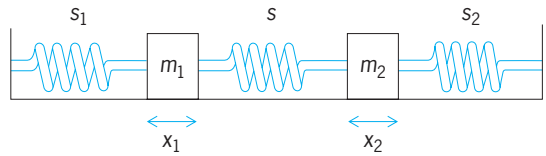

Fig. 3. Simple oscillator with two degrees of freedom.

frequency of the oscillator. *See* FORCED OSCILLATION; OSCILLATOR.

Nonlinear systems. When the amplitude of vibration of any real oscillator system becomes large, the elastic restoring force no longer is proportional to the displacement x. The term s relating force and stretch in the spring is not independent of x and may either increase or decrease as x increases. In either case, the motion no longer is sinusoidal and the frequency is not independent of the displacement amplitude A. The frequency becomes higher when s increases with increasing x and lower for a decreasing s as x increases. The latter characterizes the large-amplitude oscillations of a pendulum bob. *See* PENDULUM.

Two degrees of freedom. When two simple vibrating systems are interconnected by a flexible connection, the combined system has two degrees of freedom. Consider the simple oscillator of mass m_1 and spring s_1 connected to a second oscillator of mass m_2 and spring s_2 by means of the spring s (**Fig. 3**). The motion of m_1 is completely described by the displacement x_1 and that of m_2 by x_2. The system is said to have two degrees of freedom since two independent space coordinates are required to specify the motion.

A system having two degrees of freedom has two normal modes of vibration of respective frequencies f' and f''. Both of these frequencies differ from the respective natural frequencies f_1 and f_2 of the individual uncoupled oscillators. The larger the force constant s of the coupling spring, the larger becomes the frequency difference $(f' - f'')$ relative to $(f_1 - f_2)$; that is, f' increases relative to f_1, and f'' decreases relative to f_2.

Transverse vibrations are defined as those which occur when the vibrations of the medium in question are perpendicular to the direction of propagation of the exciting wave; longitudinal vibrations occur when the vibrations of the medium are lengthwise, along the direction of propagation of the wave. An example of a vibrating system with two degrees of freedom is given by the transverse vibrations of two masses m_1 and m_2 fastened at intermediate points on a tightly stretched wire rigidly supported at each end. Here, the higher frequency mode f' corresponds to a method of vibration in which the individual motions of the two masses are oppositely directed at all times. The lower frequency mode f'' corresponds to one in which the two masses move together in phase, that is, in the same transverse direction at all times.

Several degrees of freedom. A vibrating system is said to have several degrees of freedom if many space

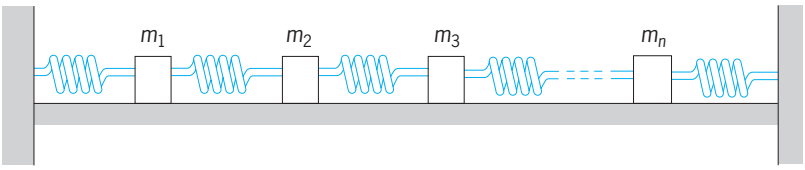

Fig. 4. Vibrating system with several degrees of freedom.

coordinates are required to describe its motion. One example is n masses $m_1, m_2, \ldots, m_n$ constrained to move in a line and interconnected by $(n - 1)$ coupling springs with additional terminal springs leading from m_1 and m_n to rigid supports (**Fig. 4**). This system has n normal modes of vibration, each of a distinct frequency.

Vibrations of elastic bodies. The primary sources of energy for producing sound waves in fluids are vibrations of elastic bodies at frequencies ranging from 20 to 20,000 Hz. In turn, the ultimate detection of sound waves by such devices as the human ear or a microphone requires an elastic body being forced into vibration by the impinging sound waves.

Vibrations of elastic bodies are not usually a simple harmonic motion of just one frequency but instead are of a complex nature, having many natural frequencies and modes of vibration. This complexity arises from numerous factors. Not all the parts of a solid body move together in phase with each other. For instance, a loudspeaker diaphragm has its mass spread over a considerable surface area whose individual sectors may vibrate with different phases and amplitudes with respect to adjacent sectors. Consequently, it often is necessary to give the displacement at each point on the diaphragm as a function of time in order to describe the motion adequately. Most solid bodies are capable of displacements in any direction in space, so that three independent space coordinates are required to specify their vibration. The vibration of a solid body is influenced by interaction with its surrounding medium. The natural frequencies and modes of vibration of a solid body depend upon its shape and dimensions.

It is sometimes possible to confine the mode of vibration to one type having a limited number of natural frequencies. The vibrations of a stretched wire are predominately transverse, although it also is possible to excite weak longitudinal vibrations along the wire. As the diameter of a wire increases relative to its length, the importance of the longitudinal mode of vibration increases until in a thin rod both transverse and longitudinal modes are readily excited. As the diameter of a rod is further increased relative to its length, it ultimately may be regarded as a thin plate capable of vibrating in a thickness mode along its axis or in transverse modes along its surface. In each of these examples it is possible to choose a method of excitation which will encourage one mode of vibration and discourage others.

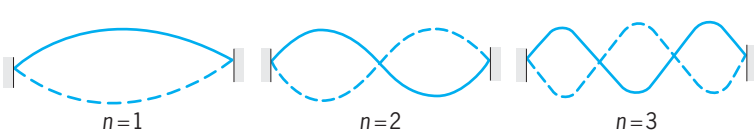

Fig. 5. First, second, and third normal modes of a vibrating string.

In any consideration of the vibration of solid bodies, characteristics of major acoustical significance include (1) the natural free vibration frequencies of the body, (2) the segmental vibration pattern for each mode of vibration, and (3) the efficiency with which the vibrations are coupled to the surrounding medium.

Vibrations of strings. The transverse vibrations of thin strings result from tension forces in the string tending to restore any displaced portion of the string to its equilibrium position. The natural frequencies f of a thin string length l rigidly supported at its ends are given by the equation $f = nc/(2l)$, where n may have any integral value 1, 2, 3, . . . and c is the velocity with which transverse waves are propagated along the string. In turn, $c = \sqrt{T/m}$, where T is the tension to which the string is stretched and m is the mass per unit length of string. The fundamental, or lowest frequency mode of vibration, of a string is $c/(2l)$. In addition, the string is capable of vibrating at harmonic overtone frequencies of 2×, 3×, 4×, . . . this frequency (**Fig. 5**). The simple integral number relationship between the various frequencies of vibration of a string leads to a pleasant harmonic tonal structure which accounts for the use of vibrating strings as the primary source of sound for such musical instruments as the violin, piano, and harp. *See* MUSICAL INSTRUMENTS.

The relative amplitudes of the various harmonic modes of vibration of a string depend upon the particular manner in which the string is initially excited. For instance, if a string is plucked at its center, the even integral (2, 4, 6, . . .) harmonic frequencies will be weak compared to the odd integral frequencies. As a string freely vibrates, nodal positions of no displacement are spaced at intervals of l/n along its length. For a string plucked at its center, the even harmonic frequencies are weak since all have a nodal position at the midpoint of the string.

Vibrations of membranes. The transverse vibrations of stretched membranes result from tension forces in the membrane tending to restore any displaced portion of the membrane to its equilibrium position. The theory of vibrations of stretched membranes of either square or rectangular shape is primarily of mathematical interest. However, the vibrations of a circular membrane have certain physical characteristics which find practical applications. These applications include parchment membranes used for drumheads and the stretched thin steel or aluminum diaphragms of condenser microphones.

When a circular membrane is rigidly clamped along its outer radius, it will vibrate at a fundamental frequency given by Eq. (6), where T is the tension to

$$f = \frac{0.384}{a}\sqrt{\frac{Tt}{\rho}} \qquad (6)$$

which the membrane is stretched, t its thickness, ρ its surface density, and a its radius. When vibrating

in this mode, the entire surface of the membrane moves in phase with a maximum amplitude at its center. The other possible modes of vibration are not integral multiples of f, but have frequency ratios of 2.295, 3.60, 4.90, and so forth, to the fundamental. When vibrating in these latter modes, one or more nodal circles are present in which the membrane is oppositely displaced on adjacent sides of the circles. The average displacement of the membrane vibrating in this manner is small, because of cancellations between the oppositely phased circular sectors. This condition leads to low acoustical efficiency and limited practical significance for either the production or reception of sound. For example, the response of a condenser microphone at first gradually falls off and then becomes very irregular at frequencies well above the fundamental of its diaphragm.

Vibrations of plates and diaphragms. The transverse vibrations of thin plates or diaphragms result from elastic restoring forces produced when their surfaces are deformed. The natural modes of vibration of such plates are involved functions of boundary shape and dimensions, whether the boundaries are free to vibrate or are rigidly clamped, and the elastic properties of its material. When a thin circular plate is rigidly clamped at its rim, its fundamental frequency is given by Eq. (7), where t is the thickness of the plate, a is its radius, ρ is its density, Y is Young's modulus, and σ is Poisson's ratio for the material of the plate. When vibrating in this mode, the entire surface of the plate moves in phase with a maximum amplitude at its center. The overtone modes of vibration have frequency ratios of 3.88, 8.70, and so on, to the fundamental. When the plate is vibrating in these latter modes, one or more nodal circles are present in which the plate is oppositely displaced on adjacent sides of the circles. The average displacement of the plate's surface under these conditions is small because of cancellations between the oppositely phased circular sectors. As in the case of circular membranes, this condition leads to low acoustical efficiency.

$$f = 0.47\,\frac{t}{a^2}\sqrt{\frac{Y}{\rho(1-\sigma^2)}} \tag{7}$$

The theory of transverse vibration of thin plates finds application in the design of such devices as telephone receiver diaphragms, horn-type loudspeaker diaphragms, underwater sound projectors used in sonar systems, and in understanding the vibrations of wall panels, floors, automobile body panels, and hull plating of ships.

The fundamental mode of vibration of a circular plate free at its rim has a frequency some 13% lower than when rigidly clamped. The relatively involved problem of transverse vibrations of a thin flat plate becomes still further complicated when the plate has a simple curved surface, and a theoretical solution is impossible when the plate is curved into the shape of a bell.

Vibrations of rods. Rod vibrations primarily consist of two simple types: (1) longitudinal vibrations along the long axis of the rod, and (2) transverse vibrations at right angles to this axis.

Longitudinal vibrations. When any plane cross section in a long thin rod is displaced longitudinally relative to adjacent planes, elastic restoring forces caused by either compression or tension in the rod tend to restore the plane to its normal position. These forces result in the propagation of longitudinal waves along the axis of the rod. The interference of two such waves traveling in opposite directions sets up a pattern of standing waves in the rod having certain discrete frequencies. The magnitude of these natural frequencies depends upon the length and material of the rod and upon the particular constraints existing at the two ends of the rod.

If the rod is either free at both ends or rigidly clamped at both ends, its fundamental frequency is given by $f = c/2l$, where l is its length, and c is the velocity with which longitudinal waves are propagated in the rod. When vibrating in its fundamental longitudinal mode, the length of the rod is one-half the wavelength of the waves being propagated along the axis of the rod. Overtone frequencies are given by integral multiples ($n = 2, 3, 4, \ldots$) of the fundamental frequency. In the case of these latter modes of vibration, nodal positions having no longitudinal displacements are spaced at intervals of $l/(2n)$ along the length of the rod.

When the rod is rigidly clamped at one end and free at the other, its fundamental frequency is given by $f = c/(4l)$. Its overtone frequencies are the odd integral ($n = 3, 5, 7, \ldots$) multiples of this frequency. Numerous types of constraints may be placed on the ends of the rod, for example, stiffness of elastic springs, inertia of masses, and so forth. When a rod free at one end is mass-loaded with a mass m at the opposite end, its fundamental frequency gradually decreases from $c/(2l)$ to $c/(4l)$ as the magnitude of the mass increases from zero to infinity.

Whenever a given rod is vibrating longitudinally in one of its natural modes, it may be supported or clamped at a nodal position without interfering with this particular mode of vibration. Since only a few modes of vibration, or in some cases only one, will have a nodal position at a given location, a judicious choice of support position may reduce or eliminate unwanted modes of vibration.

An important practical application for longitudinal vibrators is their use in sonar transducers, where both the magnetostrictive vibrations of nickel tubes and the piezoelectric vibrations of crystals are utilized.

Transverse vibrations. A long rod is capable of vibrating transversely as well as longitudinally, and it is often difficult to produce one motion without exciting the other. When a circular rod is rigidly clamped at one end and free at the other (**Fig. 6**) its fundamental frequency is $f = 0.28ac/l^2$ where c is the velocity of longitudinal waves in the material of the rod, a is its radius, and l its length. The overtone modes of vibration have frequency ratios of 6.27, 17.5, 34.4, and so on, to that of the fundamental.

If the rod were to have a rectangular instead of

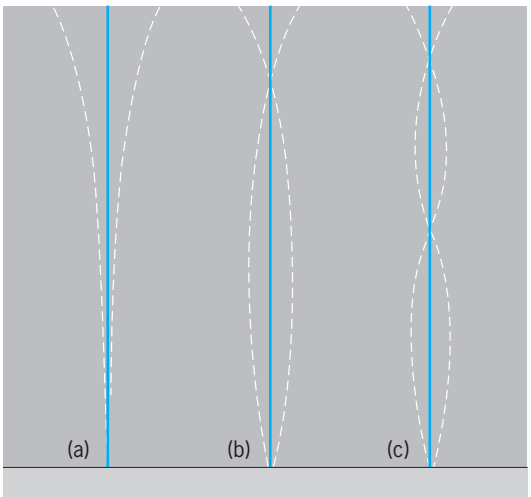

Fig. 6. Transverse vibrations of a long circular rod. (*a*) Fundamental mode. (*b*) First-overtone mode. (*c*) Second-overtone mode.

a circular cross section, its fundamental frequency would be $f = 0.16tc/l^2$, where t is the lateral thickness of the rod in the direction of transverse vibration. Thin rods or reeds of rectangular cross section clamped at one end have numerous applications, including use as a vibrating mouthpiece in the woodwind family of musical instruments and use as resonance vibrators in vibration frequency indicators. The latter type of instrument contains a large number of thin reeds, each having a different resonance frequency. When the base of the indicator is placed in contact with a vibrating object, the free end of that reed having the same natural frequency will be set into vigorous vibration, and its visual motion can be used to indicate on an appropriate scale the numerical value of the vibration frequency.

Another important type of transverse vibration is that of a rod free at both ends. The fundamental frequency of such a rod of rectangular cross section is $f = 0.33tc/l^2$, and there are overtones in ratio to this frequency of 2.75, 5.4, 8.9, and so on. The bars of a xylophone and similar musical instruments vibrate in this manner.

Vibration measuring equipment. Vibrations of solid bodies may result in the generation and transmission of unwanted noise. To determine the source of such noise and to devise methods for its elimination, it frequently becomes necessary to measure these vibrations.

When the vibration of a point in a material body is to be measured, a device must be provided with a sensing element which indicates either the displacements, velocities, or accelerations of the point. Such a device is usually either some mechanical system which indicates the characteristics of the vibration by means of a mechanical pointer, or a pickup capable of converting mechanical energy to electrical or some other form of energy. In conjunction with associated equipment, these pickups may be used to measure solely vibration amplitudes or to give a detailed picture of the entire vibration pattern. A vibrometer, or vibration meter, is a device indicating solely amplitude of vibration, while a vibrograph provides a complete oscillographic record of the vibration.

Vibration meter. A typical electrical vibration meter consists of an electromechanical pickup, adjustable attenuator, amplifier, integrating network, and a direct reading meter capable of being calibrated to read displacement, velocity, or acceleration amplitudes. Connections are also provided for oscillographic presentation, for a pair of headphones for listening to the vibration being measured, for connection to a vibration analyzer, or for connection to an electronic frequency counter.

Mechanical and mechanooptical vibrometers and vibrographs employ gear trains or mechanical or optical lever arms to magnify the vibratory motions before they are indicated or recorded. These devices are usually held by hand, with a probe contacting the vibrating surface (**Fig. 7**). Mechanical vibrometers are normally used at frequencies up to 500 Hz.

Mechanical vibrographs are instruments containing a moving paper or film on which a scribing device records the amplitude of the motion being measured. In one type, shown in **Fig. 8**, the case and recording paper themselves move in proportion to the motion being measured and, for sufficiently high frequencies, a mass supported on a soft spring with an attached scribing stylus serves as a stationary reference. In a second type, the case and recording paper are rigidly fixed, and the vibrations are transmitted

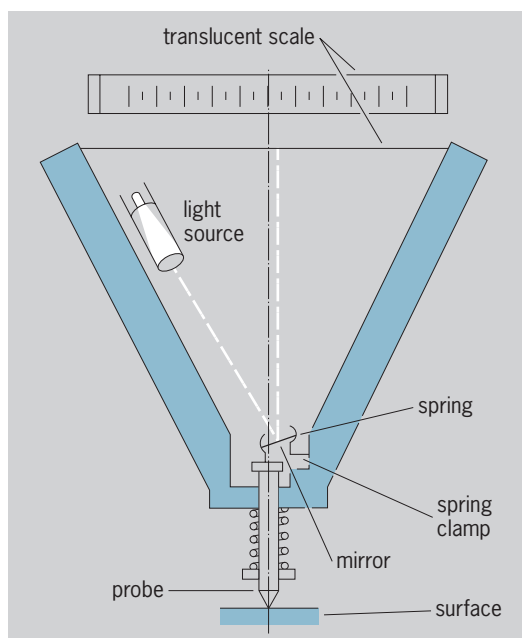

Fig. 7. Mechanooptical vibrometer. The motion given to the probe by the vibrating surface is used to rock a mirror and thereby actuate an optical lever arm. A light beam reflected from the mirror and focused onto the scale provides an indication of the vibration amplitude. (*General Electric Co.*)

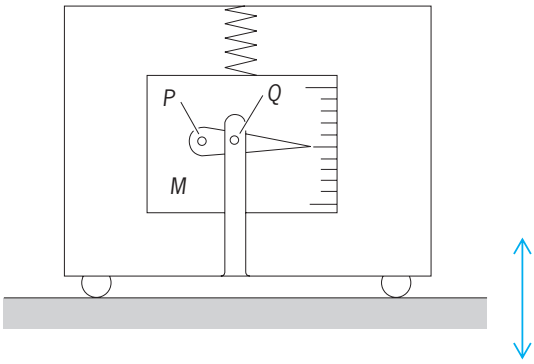

Fig. 8. Mechanical vibrograph. The scribing stylus pivots on point *P*, which is attached to the mass *M*. The case is connected to the stylus at point *Q*. (*After C. M. Harris, ed., Handbook of Noise Control, McGraw-Hill, 1957*)

to the stylus by means of a probe in contact with the vibrating body.

Resonant vibrators. The simple measurement of frequency of vibrating mechanical systems is often made with an instrument containing a series of resonant mechanical vibrators. These sensing elements are a series of cantilever-mounted reeds weighted at their free ends; their natural transverse frequencies are selected to cover a frequency spectrum of interest. Such instruments usually have a relatively small frequency range. By use of a series of these instruments, each covering a different frequency range, an entire range from about 20 to 500 Hz may be covered.

Vibration pickups. These are electromechanical transducers capable of converting mechanical vibrations into electrical voltages. Depending upon their sensing element and output characteristics, such pickups are referred to as accelerometers, velocity pickups, or displacement pickups. *See* ACCELEROMETER; SOUND; VIBRATION DAMPING; VIBRATION ISOLATION; VIBRATION PICKUP; WAVE MOTION.

Lawrence E. Kinsler

Bibliography. J. P. Den Hartog, *Mechanical Vibrations*, 4th ed., 1956, reprint 1985; A. D. Dimarogonas, *Vibrations for Engineers*, 2d ed., 1995; M. Geradin and D. Rixen, *Mechanical Vibrations*, 2d ed., 1997; S. G. Kelly, *Fundamentals of Mechanical Vibrations*, 2d ed., 2000; B. Pippard, *The Physics of Vibration*, 3d ed., 1989; W. Thomson, *Theory of Vibration with Applications*, 5th ed., 1997.

Vibration damping

The processes and techniques used for converting the mechanical vibrational energy of solids into heat energy. While vibration damping is helpful under conditions of resonance, it may be detrimental in many instances to a system at frequencies above the resonant point. This is due to the fact that the relative motion between the base of the vibration isolator and the mounted body tends to become smaller as the isolator becomes more efficient at the higher frequencies. With damping present, the force transmitted by the elastic element is unable to overcome the damping force; this leads to a resulting increase in transmissibility. *See* DAMPING; VIBRATION; VIBRATION ISOLATION.

All metal springs which include structural members such as brackets and shelves have some damping. However, such damping is insufficient for vibration isolators and must be augmented by special damping devices.

Viscous damping. Several different types of damping devices have been developed and used successfully. Probably the most familiar is that used on automobiles, which, although known as a shock absorber, is in reality a damper, and functions as a limiter to the spring system of spring constant k. The system is shown in **Fig. 1**. A piston p is attached to the body m and is arranged to move vertically through the liquid in a cylinder c which is secured to the support s. As the piston moves, the force required to cause the liquid to flow from one side of the piston to the other is approximately proportional to the velocity of the piston in the cylinder. This type of damping is known as viscous damping. The damping force is controlled by the viscosity of the liquid and by the size of the orifice in the piston. There are several disadvantages to this type of damping; for example, it is unidirectional; it is affected by temperature changes; and because the liquid is passed from one side of the piston to the other side through an opening, it is time-conscious. The opening, whether it is an orifice or the clearance between piston and cylinder, can pass only so much liquid in a given length of time. If the body to which the piston is attached is caused to displace faster than the liquid can transfer, a bottoming effect occurs. This effect is experienced by riders in automobiles when a hole in the street is hit at too fast a rate; the springs may appear to bottom out, but actually it is the shock absorber or damper. *See* SHOCK ABSORBER.

Some of the disadvantages of viscous damping may be overcome by using air instead of liquid as the damping medium. Air, being compressible, will add to the effective spring force with large displacements. If the air is housed within a flexible bellows, damping will be attainable horizontally as well as vertically. Such a system is illustrated in **Fig. 2**.

This type of damping has proved very effective

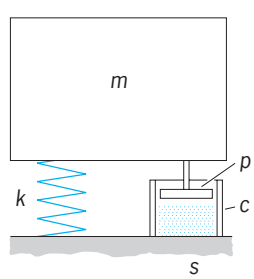

Fig. 1. Automobile shock absorber.

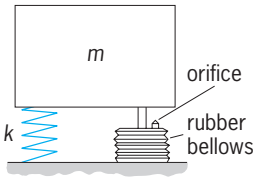

Fig. 2. System employing viscous damping with air.

in vibration isolators. The primary disadvantage occurs under conditions of high and low temperature with the change in elasticity of the rubber bellows.

Friction damping. Damping forces may be generated by causing one dry member to slide on another. This is known as dry friction or coulomb damping. A damper of this type is shown in **Fig. 3**. A pin *p* inserted in cylinder *c* and attached to body *m* is arranged to slide between two vertical spring members which are attached to the support *s*. The force exerted by the damper in opposition to the motion of the body *m* is the product of the normal force and the coefficient of friction between the pin and vertical leaf springs. The damping force is usually constant; however, if the pin is tapered, a variable force may be obtained. Friction damping is used in several commercially available isolators because it provides a simple means to control the damping forces. If necessary, independent damping systems may be provided within the same unit for vertical and horizontal motions. Some frictional dampers are effective in both vertical and horizontal directions. One such system (**Fig. 4**) comprises a load-carrying concave-convex spring made of a metal screen consisting of two parts attached by an eyelet. A damping-coil spring encircles the two attached springs. Any motion vertically or horizontally will cause the load springs to deflect vertically or tilt horizontally. Any change in position of the load springs results in the damping spring's being forced out over the surface of the concave springs; this creates a frictional force. Since the surface contact of the damping spring increases with displacement, the damping force is approximately linear.

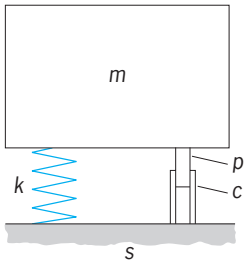

Fig. 3. System employing Coulomb damping.

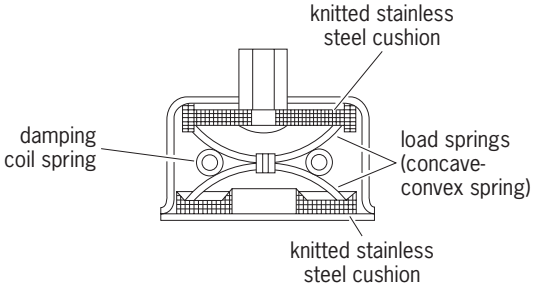

Fig. 4. Frictional damper.

Inherent damping. There are many applications where vibration dampers of an external type such as those discussed cannot be used because of space limitations, economic considerations, or the fact that the system needs very little damping. These applications make use of the inherent damping, or internal hysteresis, of such materials as rubber, felt, and cork. Vibration isolation with inherent damping is most commonly used in applications with constant motor speeds, such as air compressors, generators, and grinders.

Magnetic damping. This type of damping is attainable as a result of the electric current induced in a conductor moving through a magnetic field. The damping force can be made proportional to the velocity of the conductor moving through the field. Magnetic damping has not been used successfully in vibration isolators because its effectiveness is limited to a single direction. *See* ELECTROMAGNETIC INDUCTION.

Numerical values. Considerable emphasis is placed upon the numerical value for the damping force. Such a value is needed to predict the behavior of a damped system; however, difficulty is often encountered in such an analysis. The difficulty is that viscous damping, although susceptible to mathematical analysis and the establishment of a numerical value for the damping force, is seldom encountered in pure form in actual practice. The types of damping actually used are not well adapted to mathematical analysis. The damping value may be determined by the logarithmic decrement method, but care should be exercised when these values are used in equations of pure viscous damping. This is especially true of the two most common forms of damping—friction and inherent damping. With both these types the effect of damping will vary with the amplitude of vibration. Since the damping effect is different, the natural frequency of the system will change with amplitude of vibration. For double amplitudes of vibration, such as 0.060 in. (0.15 cm), the magnification at resonance may be one value and the resonance frequency may be so many cycles per second, while with a double amplitude of 0.020 in. (0.05 cm), the magnification may be different and the resonance period may be higher. This difference must be recognized when applying formulas derived from the analysis of pure viscous damping. *See* SPRING (MACHINES); VIBRATION MACHINE.

K. W. Johnson

Bibliography. L. L. Beranek and I. L. Ver, *Noise and Vibration Control Engineering: Principles and Applications*, 1992; J. P. Den Hartog (ed.), *Mechanical Vibrations*, 4th ed., 1956, reprint 1985; C. M. Harris, *Shock and Vibration Handbook*, 4th ed., 1995;

A. D. Nashif, D. I. Jones, and J. P. Henderson, *Vibration Damping*, 1985; S. S. Rao, *Mechanical Vibrations*, 2d ed., 1990.

Vibration isolation

The isolation, in structures, of those vibrations or motions that are classified as mechanical vibration. Vibration isolation involves the control of the supporting structure, the placement and arrangement of isolators, and control of the internal construction of the equipment to be protected.

The simplest kind of mechanical vibration has the waveform of sinusoidal motion (**Fig. 1**). Vibrations in structures, although generally more complex in waveform, exist wherever movement takes place. Such movement may be caused, for example, by the engine in an automobile, by engines or wind buffeting in aircraft, or by a punch press in a building. Delicate electronic equipment and precision instruments must normally be isolated from these motions if accurate measurements are to be obtained. *See* MECHANICAL VIBRATION; VIBRATION.

Vibration, in most cases, may be effectively isolated by placing a resilient medium, or vibration isolator, between the source of vibration and its surrounding area to reduce the magnitude of the force transmitted from a structure to its support or, alternatively, to reduce the magnitude of motion transmitted from a vibrating support to the structure. Isolating vibration at its source is commonly termed active or source isolation; isolating an instrument from its surroundings is known as passive isolation. In either case the vibration isolator is designed according to the same principles. Vibration isolators are available commercially, with published data covering their characteristics. These data include minimum and maximum rated load, natural frequency at rated load, transmissibility, damping characteristics, ultimate strength, sway space limits, and so on.

Natural frequency; resonance. The prime concern in the field of vibration isolation is the proper use of isolators under various load configurations with respect to their loading, the desired natural frequency, the position and location of the isolators, and the relationship to the structural response of the equipment to which they are attached. For the vibration isolator to be effective, the natural frequency should be approximately 0.4 times the frequency of the interfering source. The natural frequency is the frequency at which a freely vibrating mass system oscillates once it has been deflected. In the case where vibrations occur over a wide frequency range, such as in aircraft (5–2000+ Hz), the natural frequency of the isolator is established with respect to the cruising speed. This means that when the lower frequencies are traversed, such as during aircraft takeoff, the mounted equipment will pass through a condition known as resonance. Resonance is said to exist when the natural frequency of a spring (in this case an isolator) coincides with the frequency of the excitation forces. Resonance causes magnification of the input vibration and may be harmful to the equipment if not controlled within reasonable limits. To control the vibration magnification at resonance, the resilient element within the isolator must be damped. With suitable damping, the magnification factor may be held to 3 or less. *See* DAMPING; RESONANCE (ACOUSTICS AND MECHANICS); VIBRATION DAMPING.

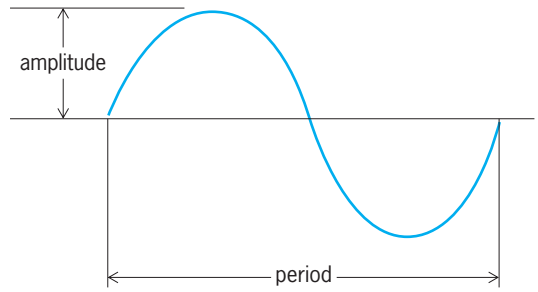

Fig. 1. Sinusoidal motion.

The vibration isolator should be considered as only one part of the isolating system, the other parts being the supporting structure that lies below the isolator, and the internal structure of the equipment that is above the isolator. When isolators are selected for use where the period of resonance is critical, it should not be forgotten that the flexibility of the supporting structure and the flexibility of the isolators are in series, so that the resultant resonant frequency of the loaded system will therefore be inversely proportional to the square root of the sum of these two flexibilities. The additional flexibility of the structure will lower the natural frequency of the system and will also result in increased displacement during resonance, caused by the presence of the undamped structure. The flexibility of the supporting structure, including the structural linkages leading to such structures, should not exceed 25% of the flexibility of the isolator. To neglect the resiliency of the support structures in providing the desired natural frequency of the system, as so many textbooks do with the assumption of rigid support, is impractical.

There are several methods for determining the load at each isolator location; one is shown in **Fig. 2**, where the loads are as given in the following

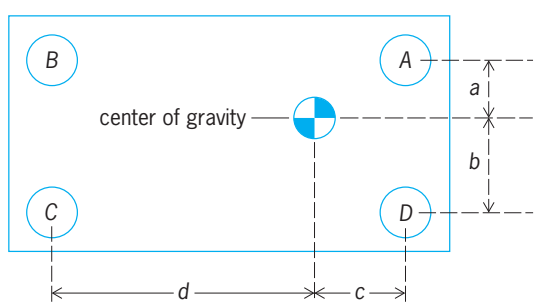

Fig. 2. Determination of isolator locations.

four equations (W is the total weight).

$$A = W\left(\frac{b}{a+b}\right)\left(\frac{d}{c+d}\right)$$
$$B = W\left(\frac{b}{a+b}\right)\left(\frac{c}{c+d}\right)$$
$$C = W\left(\frac{a}{a+b}\right)\left(\frac{c}{c+d}\right)$$
$$D = W\left(\frac{a}{a+b}\right)\left(\frac{d}{c+d}\right)$$

The next step in vibration isolation is the positioning and arrangement of the isolators with regard to the geometry of the equipment.

Location of isolators. The vibration isolators may be positioned and arranged in many different ways, all variations of three basic types, each of which requires a definite amount of space: (1) isolators attached underneath equipment, known as an underneath mounting system; (2) isolators located in the plane of the center of gravity of the equipment, known as a center-of-gravity system; (3) mountings arranged four on each side in the plane of the radius of gyration, known as a double side-mounted system or radius-of-gyration system. *See* CENTER OF GRAVITY; RADIUS OF GYRATION.

Textbooks generally treat problems relating to the mounting of equipment with resilient elements as masses in unlimited space. Under these undefined conditions any of the three systems listed may be used; however, when a space limitation is imposed, each system has a definite limitation and must be used accordingly.

Underneath mounting system. For an underneath mounting system, the most efficient for vibration isolation is one with a low natural frequency in both the vertical and horizontal axes. The most stable system is one with a high natural frequency with respect to the disturbing frequency, but in this case the vibration characteristics will suffer. Stability may be achieved with low-frequency isolators by maintaining adequate spacing of the isolators with respect to the height of the center of gravity from the mounting plane. Tests show that the isolator spacing should not be less than twice the height of the center of gravity from the mounting plane. This condition is illustrated in **Fig. 3** . The use of an underneath mounting system for equipments that exceed the condition shown in Fig. 3 should be augmented by stabilizers to provide the needed stability.

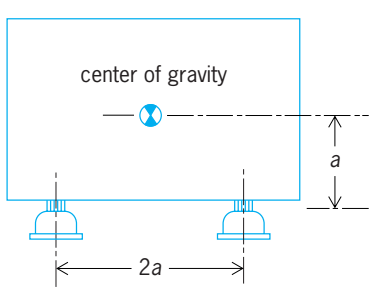

Fig. 3. Underneath mounting system.

Center-of-gravity system. Locating the isolators in the plane of the center of gravity has generally been considered the ideal system because of its ability to decouple the rotational modes of vibration. In actual use, however, such results are seldom achieved because of space limitations. The primary conditions are (1) that the isolators be located in a plane passing through the center of gravity, (2) that the distance between isolators be twice the radius of gyration of the body, and (3) that the horizontal-to-vertical stiffness of the isolators be equal. The first and last of these conditions are easy to satisfy. The second condition is not always possible where a space limitation exists. The limit of the center-of-gravity system, especially in aircraft, is generally arrived at when the height-to-width ratio of the equipment exceeds 2.8.

Radius-of-gyration system. The third system that may be used when the limit of the center-of-gravity system has been reached is a double side-mounted system or radius-of-gyration system as illustrated in **Fig. 4**. Two sets of isolators are arranged on each side. For optimum results the isolators should be located in the plane of the radius of gyration. However, since it is often difficult to determine the exact radius, acceptable results will be obtained in most instances by assuming the body to be of uniform density. Satisfactory results have been obtained with bodies having height-to-width ratios up to 5. The limitation of this system will be reached when structural rigidity of the body is such that excessive bending occurs between the upper and lower isolator locations.

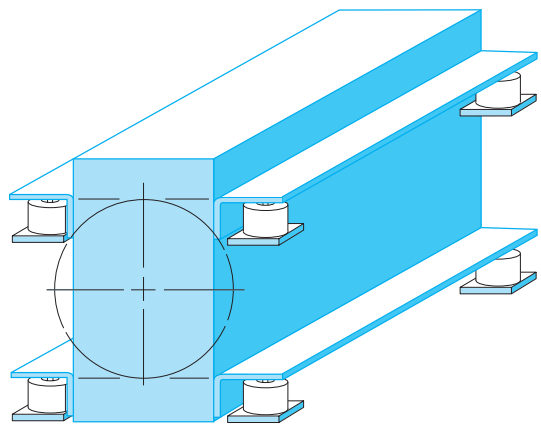

Fig. 4. Radius-of-gyration system.

The next consideration in vibration isolation is the structural rigidity of the body to be isolated. This step is of importance in that the use of incorrect structures, particularly supporting brackets for component parts, can render the other two steps useless. Supporting brackets act as springs under a vibratory condition and become resonant at their natural frequency. Should resonance of the brackets coincide with the isolator resonance, damage may occur. The resonance point for the brackets and internal component should occur when the isolators are approaching their maximum efficiency so that the

input vibration to the internal structure is at a low level. This level is normally reached at four times the natural frequency of the isolator.

The isolation of vibration may be accomplished in most instances by the use of commercial isolators. In some cases, such as in heavy industrial machinery where the supporting structure is concrete, less consideration need be given to the flexibility of the supporting structure. Also, where the body to be isolated is stationary, as in the case of compressors, the height-to-width ratio with respect to the isolator spacing is of little importance. *See* VIBRATION MACHINE.

K. W. Johnson

Vibration machine

A device for subjecting a system to controlled and reproducible mechanical vibration. Vibration machines, commonly called shake tables, are widely used in vibration measurement and analysis. There are three types of vibration machines in general use. These are the mechanical direct-drive type, the mechanical reaction type, and the electrodynamic type. Other types, such as hydraulic excitation devices, resonant systems, piezoelectric vibration generators used for instrument calibration, and machines for testing packages, have only limited or specialized applications.

Direct-drive machine. The mechanical direct-drive vibration machine (sometimes referred to as the brute-force type) is a machine in which the vibration table is forced by a positive linkage to undergo a displacement. The linkage is driven by a direct attachment to eccentrics or crankshafts. The degree of eccentricity or crank radius determines the amplitude of vibration of the table. The motion of the vibration table is approximately simple harmonic. The frequency is varied over a range of 5–60 Hz by a variable-drive unit. Automatic cycling is normally used to cycle the table to maximum frequency and return to the starting frequency. The period of cycling is usually 1–3 min. *See* HARMONIC MOTION.

The vibration table is usually designed, by the use of several cam linkages, to move in both vertical and horizontal directions. A machine of this type is pictured in **Fig. 1**.

Reaction-type machine. A vibration machine in which the forces exciting the vibrations are generated by rotating or reciprocating unbalanced masses is called a reaction-type machine. The system consists of a table which is suspended from a mounting frame attached to the floor. The unbalanced masses which are attached to a shaft are part of the floating table assembly. A separate variable-speed unit, which is attached to the mass-shaft assembly of the table by a flexible coupling, drives the system.

The amplitude of vibration is controlled by the degree of off-center setting of the rotating masses. Frequency is controlled by a rheostat in series with a dc motor. The frequency may be continuously varied over a range of 5–60 Hz. The primary advantage of this type of machine is the constant displacement provided by the rotating masses. Also, since the table is suspended, there is essentially no vibration at the floor attachment points. A reaction-type machine is pictured in **Fig. 2**.

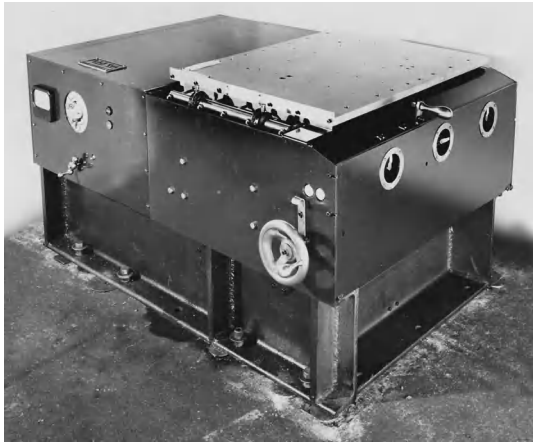

Fig. 1. Direct-drive vibration machine. (*L. A. B. Corp.*)

Electrodynamic machine. The electrodynamic vibration machine derives its vibratory force from the action of a fixed magnetic field upon a coil of wire contained in it and excited by a suitable alternating current. The armature consists of those parts that are integral with the coil and the associated elements that move in the magnetic field. This type works on the same principle as that of an electrodynamic loudspeaker. The force output is proportional to the magnitude of the current. Automatic displacement or acceleration control may be had over a frequency range of 5–2000+ Hz.

The electrodynamic vibration machine provides motion of a sinusoidal waveform and a form commonly termed random vibration.

Fig. 2. Reaction-type vibration machine. (*L. A. B. Corp.*)

Fig. 3. Complete vibration test system showing the electrodynamic shaker in the left foreground. In the rear, from left to right, are the power amplifier, control console, and instrumentation cabinets. (*Calidyne Co.*)

The primary advantage of this type of vibration machine is the wide frequency range it offers. **Figure 3** pictures a machine complete with the control console. *See* MECHANICAL VIBRATION; VIBRATION.

K. W. Johnson

Vibration pickup

An electromechanical transducer capable of converting mechanical vibrations into electrical voltages. Depending upon their sensing element and output characteristics, such pickups are referred to as accelerometers, velocity pickups, or displacement pickups.

Accelerometer. This device consists essentially of a mass which is seismically supported with respect to a surrounding case by means of a spring and guided to prevent motions other than those along the seismic direction of support. In the operating frequency range of the accelerometer, which is below its resonant frequency, the mass experiences essentially the same acceleration as does the case of the accelerometer. The mass exerts a force on the spring's support which is directly proportional to the acceleration being measured. This, in turn, is converted into an electrical voltage by means of stresses produced in a piezoelectric crystal. *See* ACCELEROMETER.

Velocity pickup. This device generates a voltage proportional to the relative velocity between two principal elements of the pickup, the two elements usually being a coil of wire and a source of magnetic field (see **illus.**). In a probe type of velocity pickup, the relative motion is between a coil, which is attached to one end of a probe having its other end in contact with the structure whose vibration is being measured, and a stationary magnetic field produced inside the hand-held case. The magnetic field is essentially stationary in space so that the voltage output of the moving coil is directly proportional to the velocity of the vibrating body.

Displacement pickup. This is a device that generates an output voltage which is directly proportional to the relative displacement between two elements of the instrument. These pickups are similar in construction and behavior to velocity pickups. The only essential difference is the use of a frequency-weighting network, required to make them direct-reading. Instrumentation problems in vibration measurements are generally reduced if measurements of small displacements at high frequencies are made

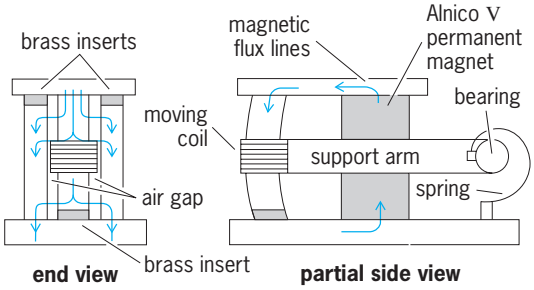

Velocity pickup. The coil swings on one end of an arm which is supported by bearings at the opposite end. The case follows the motion of the structure to which it is attached. (*MB Manufacturing Co.*)

with accelerometers or velocity pickups, since for a given displacement the acceleration varies as the square of the frequency and the velocity directly as the frequency. *See* VIBRATION. Lawrence E. Kinsler

Vibrotaction

The response of tactile nerve endings to varying forces on the skin and to oscillatory motion of the skin. Grasping, holding, and tactile exploration of an object are part of everyday experience. The performance of all these activities is dependent on the dynamic response of specialized nerve endings located in the skin.

Knowledge of the neural and psychophysical processes involved in vibrotaction is necessary to develop effective tactile communication systems, such as vibrotactile pagers, and multipin vibrotactile stimulators used to produce images on an extended skin surface. Potential users range from business people to the visually and hearing impaired: potential applications range from activities involving visual and aural sensory saturation (such as pilots in high stress situations) to those requiring sensory stimulation (for example, virtual reality).

The success of vibrotactile devices and the successful performance of manual tasks involving grasping and holding small or delicate objects ultimately depend on maintaining adequate tactile function.

Touch receptors. Small electrical impulses, or action potentials, generated by a single nerve ending can be recorded by inserting a microelectrode into the arm of an alert human subject and positioning the electrode tip in a nerve fiber. Studies of the action potentials produced when the skin is locally depressed at a fingertip have identified activity in four networks of distinctive nerve endings lying within 1–2 mm of the skin surface. These nerve endings, which transform skin motion into neural signals, consist of physically and functionally different mechanoreceptors.

Three types of mechanoreceptor in the fingertip have been implicated in vibrotaction. One of these is a population of slowly adapting mechanoreceptors (SA type I, or SAI); their action potentials persist for some time after a transient skin indentation. The two others are populations of rapidly adapting receptors (FAI and FAII types). Type I receptors respond only to skin indentation within a few millimeters of the nerve ending, whereas type II receptors also respond to stimuli at a greater distance from the nerve ending. A fourth mechanoreceptor type appears to respond primarily to skin stretch (denoted SAII).

Mechanoreceptor populations. Exploration of the surface of an object by the fingertips produces spatially and temporally varying tactile stimuli.

For example, to record the responses of single human mechanoreceptors to the Braille representation of letters, a sequence of embossed dots is scanned repeatedly over a (stationary) fingertip in one direction. Between each scan, the sequence of dots is shifted in a direction normal to the scan. The action potentials generated by the skin-stretch receptors and FAII receptors do not resolve the Braille characters. In contrast, the characters are clearly resolved by the SAI receptors, and are partially resolved by the FAI receptors. Thus, surface topography of this form and scale can be detected peripherally by individual nerve endings. The SAI receptor population also establishes the limit of resolution of gross surface features, such as edges and cracks.

When scanned by the fingertip, small-scale surface features or textures (for example, sandpaper or silk) generally require more complex neural representation in order to be detected. In the detection of surface texture, the slowly adapting and the rapidly adapting receptor populations are implicated, either together or separately, depending on the nature of the surface and details of the task. For example, the texture of finely etched surfaces, consisting of dots 1 μm in relief, 50 μm in diameter, and separated by 100 μm, can be detected when stroked across the fingertip. In this case, the texture appears to be detected by neural activity in the FAII receptor population.

Successful lifting and gripping of small or delicate objects between the thumb and fingertips (for example, an eggshell) are also dependent on vibrotaction. This is because precision grip is controlled indirectly by neural activity in the slowly adapting and rapidly

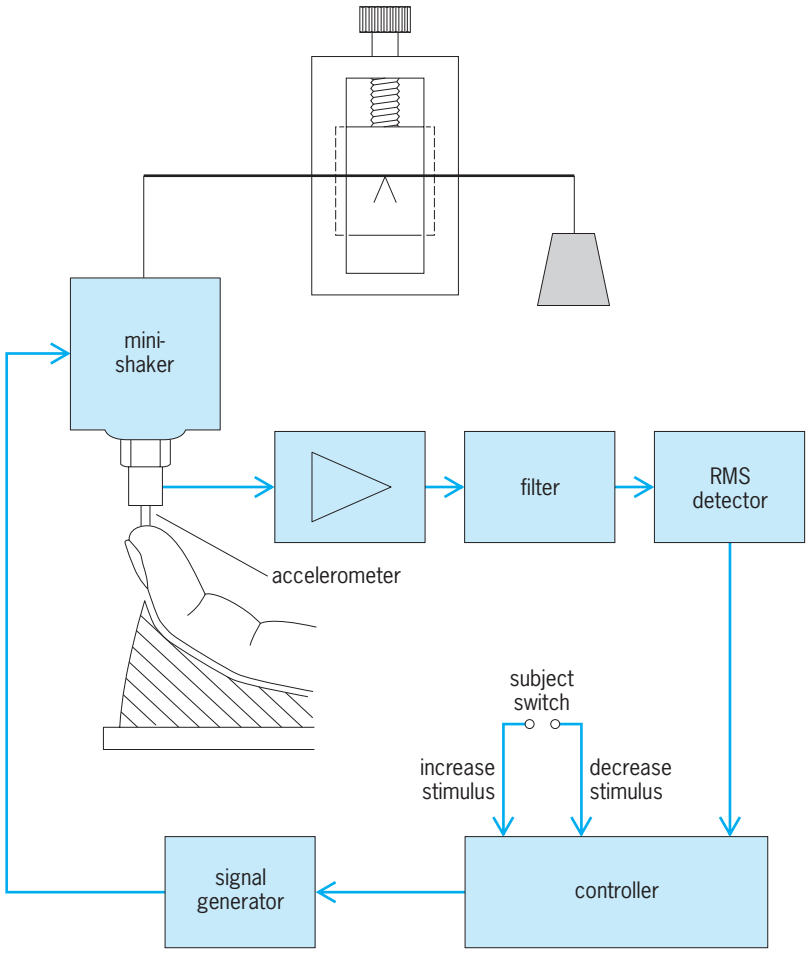

Fig. 1. Block diagram of apparatus for determining vibrotactile perception thresholds at the fingertip.

adapting mechanoreceptor populations, which triggers compensatory muscle action in response to the detection of microscopic slippage of the object.

Thresholds. The existence of four, separate neural channels at stimulation levels close to the threshold of vibrotactile perception has been confirmed by psychophysical experiments. Three of these neural channels, corresponding to the slowly adapting type 1 and both types of rapidly adapting mechanoreceptors, can determine independently the vibrotactile threshold in specific frequency ranges with appropriate stimulation. Therefore it is possible to express changes in mechanoreceptor acuity in terms of changes in vibrotactile perception thresholds at selected frequencies.

The determination of the acuity of the three mechanoreceptor populations requires apparatus such as that shown in **Fig. 1**. The subject is seated comfortably with forearm and hand supported horizontally by a contoured arm rest. A small-diameter, flat-ended, cylindrical probe is lowered onto the fingertip and positioned between the center of the whorl and the nail. The stimulator is mounted on a beam balance (the fulcrum of which may be raised or lowered) so as to maintain the extremely small contact force required between the skin and stimulating probe (typically 0.15 N or 0.5 ounce·force).

Sinusoidal bursts of vibration followed by quiescent periods are then applied to the fingertip, with amplitude initially increasing between successive bursts. Detection of the tone burst by the subject is signaled by pressing a switch, whereupon the stimulus amplitude is recorded by a computer. The vibration intensity is then decreased between bursts until the stimulus can no longer be felt, whereupon the stimulus amplitude is again recorded and the amplitude progression again reversed.

Tactograms. The results of vibrotactile threshold measurements may be conveniently summarized by a tactogram (**Fig. 2**). This graphical representation of the data is equivalent to the audiogram obtained from audiometric measurement of hearing acuity. A tactogram shows the change in threshold sensitivity from the mean value for healthy hands as a function of stimulation frequency. Thus, it can measure the differences in acuity between fingers and between individuals. A tactogram can also measure the losses in acuity resulting from injury, disease, or aging. *See* AUDIOMETRY.

One type of abnormal tactogram is characterized by similar reductions in threshold sensitivity at all frequencies and is exemplified by the data plotted as circles in Fig. 2. This similar loss in acuity occurring in each network of mechanoreceptors and associated nerve fibers is compatible with changes occurring within the whole nerve. A second type of abnormal tactogram involves elevated thresholds at frequencies mediated by only one receptor type (for example, the data plotted as squares in Fig. 2). A mechanoreceptor-specific mechanism is unlikely to occur within a nerve trunk; therefore this frequency-selective loss pattern is suggestive of neural dysfunction occurring at, or close to, the nerve endings. This pattern of abnormal threshold elevation has been observed only in the hands of manual workers who operate vibrating power tools. *See* CUTANEOUS SENSATION; MECHANORECEPTORS. Anthony J. Brammer

Bibliography. S. Bolanowski et al., Four channels mediate the mechanical aspects of touch, *J. Acous. Soc. Amer.*, 84:1680–1694, 1988; A. J. Brammer and J. E. Piercy, Rationale for measuring vibrotactile perception at the fingertips as proposed for standardization in ISO 13091-1, *Arbetslivsrapport*, 4:125–132, 2000; K. O. Johnson and S. S. Hsiao, Neural mechanisms of tactual form and texture perception, *Annu. Rev. Neurosci.*, 15:227–250, 1992; M. A. Srinivasan, J. M. Whitehouse, and R. H. LaMotte, Tactile detection of slip: Surface microgeometry and peripheral neural codes, *J. Neurophysiol.*, 63:1323–1332, 1990.

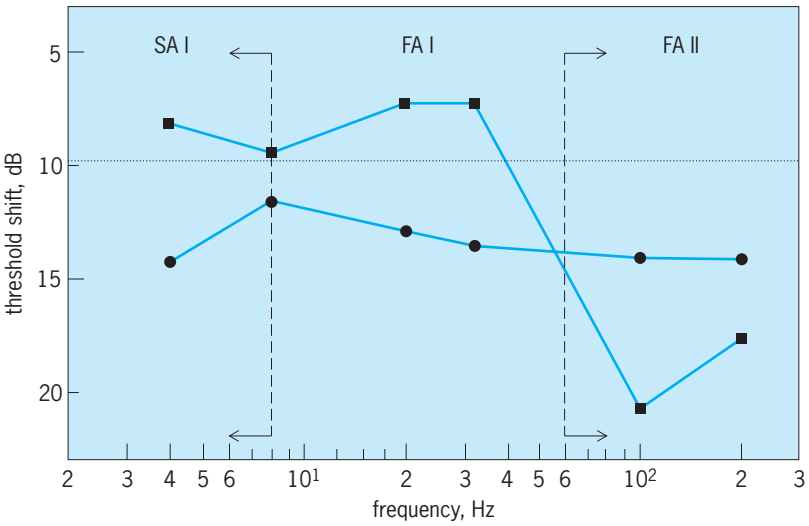

Fig. 2. Tactogram showing the approximate maximum threshold recorded at the fingertips of healthy persons (horizontal line), and examples of two types of threshold elevation (reduction in acuity) at selected frequencies (circles and squares).

Vicuna

A rare animal whose fiber makes the world's most costly and most exquisite cloth, surpassing all others in fineness and beauty. It is found in an almost inaccessible area of the Andes Mountains, at altitudes between 16,000 and 19,000 ft (4876 and 5791 m). The vicuna, one of the wildest of animals, is less than 3 ft (0.9 m) high and weighs 75–100 lb (33.7–45 kg).

A single animal yields only approximately $^1/_4$ lb (0.11 kg) of hair; thus 40 animals are required to provide enough hair for the average coat. To preserve the species, the Peruvian and Bolivian governments have placed the vicuna under their protection. Attempts to domesticate this animal have not been successful but are still being made in Peru. The fiber of the vicuna is the softest and most delicate of the known animal fibers; yet it is strong for its weight, is resilient, and has a marked degree of elasticity and surface cohesion. It is used in costly suitings and overcoat fabrics. *See* ALPACA; CAMEL'S HAIR; CASHMERE; LLAMA; MOHAIR; NATURAL FIBER; WOOL.

M. David Potter

Video amplifier

A low-pass amplifier having a bandwidth in the range from 2 to 100 MHz. Typical applications are in television receivers, cathode-ray-tube computer terminals, and pulse amplifiers. The function of a video amplifier is to amplify a signal containing high-frequency components without introducing distortion.

In a single stage of an *RC* (resistance-capacitance) coupled amplifier the high-frequency half-power limit is determined essentially by the load resistance, the internal transistor capacitances, and the shunt capacitance in the circuit. To extend the bandwidth of an *RC*-coupled amplifier, it is necessary to overcome the effects of these capacitances. In the past, this was often done by adding an inductance *L*, as shown in the **illustration**, employing shunt peaking. The principal design requirement is to choose a value for the inductance that will extend the frequency response without introducing an undesirable hump in the gain characteristic for frequencies near the upper half-power frequency.

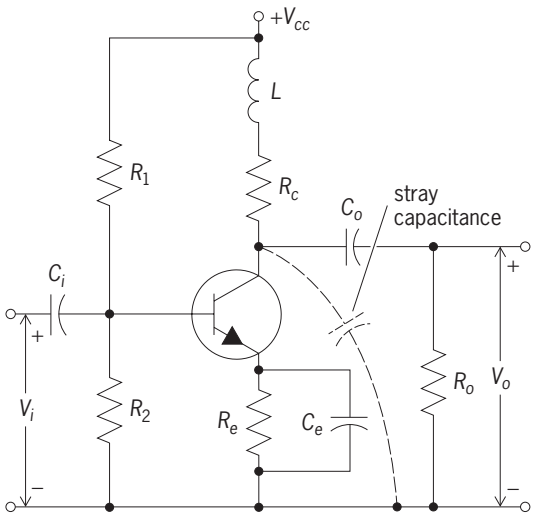

Shunt-compensated amplifier, in which effect of capacitances is overcome by adding inductance *L*.

Modern video amplifiers use specially designed integrated circuits. With one chip and an external resistor to control the voltage gain, it is possible to make a video amplifier with a bandwidth between 50 and 100 MHz having voltage gains ranging from 20 to 500. The use of integrated video amplifiers minimizes cost and space. These amplifiers also eliminate the need for the shunt-peaking inductor shown in the illustration. *See* AMPLIFIER; INTEGRATED CIRCUITS.

Harold F. Klock

Bibliography. D. Christiansen (ed.), *Electronics Engineers' Handbook*, 4th ed., 1996; J. Millman and A. Grabel, *Microelectronics*, 2d ed., 1987; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Video disk

A medium used to record, distribute, and play video information. The video disk exists in three major forms: the Digital Video Disk (DVD; sometimes also called the Digital Versatile Disk); the Laser Video disk (sometimes called the Laser Vision disk); and the Video CD. There is an audio version of the DVD, called the Digital Audio Disk, that is based on the same digital principles as the DVD. *See* COMPACT DISK.

The advantages of DVD and Laser Video disks over videocassette tapes include very high video quality (because the signal-to-noise ratio is very high); use of a full television frame on playback; multiple audio channels, digital and analog; archival quality, since there is no disk deterioration over time; the possibility of interactivity, whereby cues, branches, and so forth can be programmed in; a freeze-frame option with no loss of quality; the ability to interface and interact with computers using digital cues; the availability of rapid access to any point in the recording; the possibility of auto-stop, whereby a frame is preselected for freezing; frame-by-frame viewing, providing slow motion with no loss of quality; and programmability with certain systems. The Video CD does not possess these advantages, but is a lower-cost option with limited image quality and minimal interactivity.

Common elements. All three types of video disks have basic elements in common, notably a solid-state laser a fraction of a square millimeter in size that emits a few milliwatts of coherent light focused to a micrometer-sized spot (see **illus.**). This laser spot becomes essentially an optical stylus that performs various functions such as tracking, focusing, and reading the total integrated light. *See* LASER.

Tracking and focusing. The laser spot follows a track on the rotating disk that is either the data itself or a groove. It does this by sensing the symmetry of the laser light reflected back to a split segmented detector; if the spot is not symmetric, the position of the head is jogged until it is back to a perfectly registered spot on the detector. The jogging movement is created by a small motor, often a voice coil motor, that is driven by an electronic servo system that "closes the loop" in an iterative manner. The laser spot and the disk are kept in focus through a similar servo-based controller that uses the image of the laser on the detector.

Total integrated light. By adding up the total integrated light on the detector, it is possible to read the total light returned from the disk. While the laser light travels over the disk, differences in the intensity of the light, affected by the disk features that represent the data, are detected. These disk features can be pits in the surface (using the phase contrast of light coming off the disk) or the absorbed light in a dye or other material. This change in intensity is detected and the pulse train decoded by the system to produce a series of data bits that are then reconstructed to form a video image. This principle is used by all the above video disk applications. *See* OPTICAL RECORDING.

TABLE 1. Major changes to CD technology that created DVD

Parameter changed	Result
Code format changed from EFM to EFM+	Increased density × 1.1
Track pitch changed from 1.6 to 0.74 micrometers	Increased density × 2
More efficient error correction (ECC)	Increased density × 1.3
Smaller minimum feature (400 nm)	Increased density × 2

DVD. This is a video version of the original compact disk (CD) and has many similarities. As for CDs, the data that are encoded on the disk, and later decoded, are digital rather than analog. This means that the discrete disk pits are decoded to give a sequence of 1's and 0's, which then can be manipulated using digital techniques. The DVD uses the same conceptual method for tracking and focusing and the same constant linear velocity (CLV) rotation method as CDs. The disk dimensions are the same, with 120-mm (4.72-in.) and 80-mm (3.15-in.) outer-diameter versions. However, the DVD disk itself is only 0.6 mm thick (versus 1.2 mm); two are then laminated together to get 1.2 mm as in CDs. The inner hole diameters of DVDs and CDs are identical at 15 mm (0.6 in.).

Significant changes were made in DVD disks relative to CDs, including increased density and data rate. Through those changes and by the use of MPEG (Motion Picture Enhanced Graphics) compression technology, it was possible to create 4.7-GB disks that can contain several hours of, for example, an NTSC (National Television System Committee) quality movie. The development of data compression technology was critical to the success of DVD; it allows a single 4.7-GB disk to contain up to two versions of a typical full-length movie as well as supplemental information such as parallel sound tracks. This has been pivotal to its commercial success. Newer versions of the DVD disk have higher capacities and also are recordable (**Table 1**). *See* DATA COMMUNICATIONS; DATA COMPRESSION.

Eight to Fourteen Modulation (EFM) code, the way that a collection of 1's and 0's is translated to bytes, contains rules that must be followed to ensure that overlap of disk pits does not occur. The change to EFM+ for DVDs produced a more efficient translation method. The error correction (ECC) method used in DVD is called the 2Dimensional Product Code and is able to reduce errors on the disk from one error in 1000 bits to virtually none. The changes in track pitch and minimum feature size (the smallest feature that can be read) resulted from changing the wavelength of the laser from 780 nm to 650 nm (the shorter the wavelength, the smaller the spot) and increasing the numerical aperture of the lens system (that is, the power of the lens) that focuses the laser image; the DVD lens now has a numerical aperture of 0.6.

There are several capacity versions of the DVD disk: a single-layer, single-sided disk at 4.7 GB; a two-sided disk with one layer each side (9.4 GB); a single-sided disk with two layers per side (8.5 GB); and a two-sided disk with two layers each side (17 GB). These disks are all 1.2 mm in thickness but are made with laminates that can be read through from a single side by moving the focus point to the deeper level (two layers), or have to be turned over (two sides). Different types of DVD may also have different capabilities (**Table 2**). The major difference between random access memory (RAM) and read/write (RW) formats is the use of zones in the RAM disk that allow the constant linear velocity operating mode to have a greater degree of rapid random access, closer to a constant angular velocity (CAV) device.

DVD players/recorders are available as stand-alone devices, incorporated into computers, or as part of video entertainment devices or combination devices. Key aspects of DVD players are shown in **Table 3**.

Laser Video disk. Older than DVD technology and now largely supplanted by it, the Laser Video disk

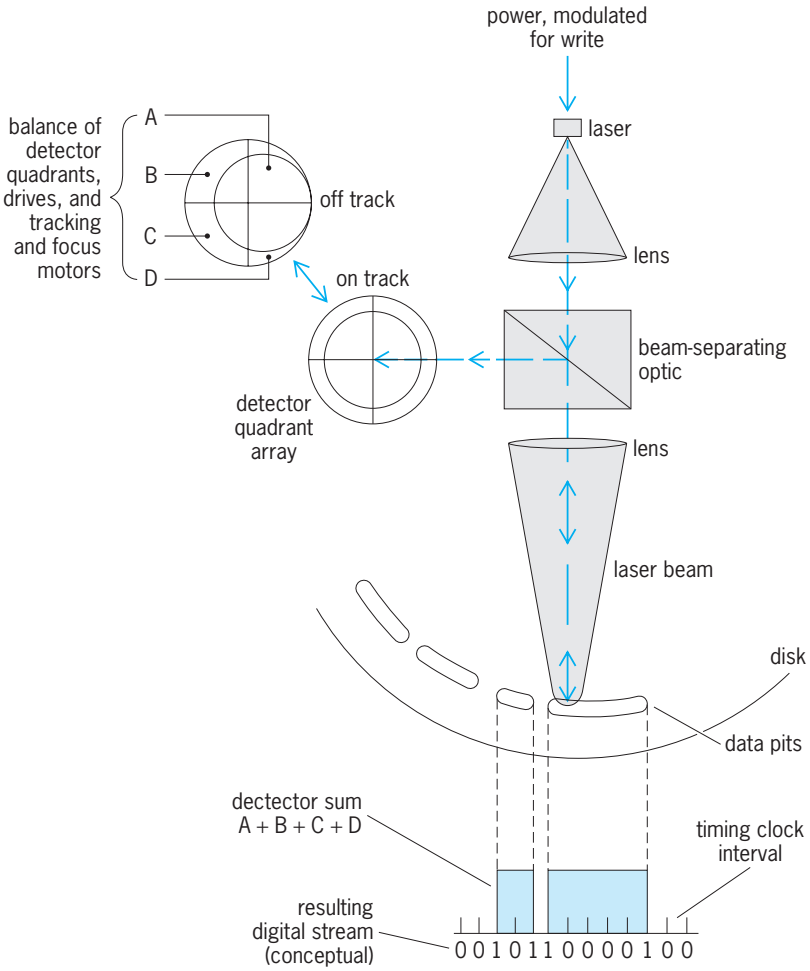

DVD operational principles.

TABLE 2. Types of DVD disks and their uses

Type	Detail	Use
DVD ROM	Read only	Used for DVD playback; not recordable
DVD RAM	Read/write	Recordable and rewritable disk
DVD R	Write once	Recordable once but not re-recordable
DVD RW+/−	Read/write	Different versions of DVD RAM

TABLE 3. Specifications of DVD players

Feature	Specification
Video component	Digital MPEG-2 (30:1 compression from 8-bit 4:2:0)
NTSC resolution	720 × 480 (or PAL resolution 720 × 576)
Audio uncompressed	16/20/24/-bit 48/96 KHz LPCM
Audio compressed	Dolby digital, MPEG1, or MPEG 1 layer 2, audio S/N 96–144 dB
Data rate	Maximum 9.8 Mbps
Rotation speed	570–1600 rpm
Channel data rate	26.16 Mbps

uses many of the same principles. The major difference is that the information is recorded on to the disk and played back in an analog form—that is, it is based on a continuously variable signal without discrete digital bits. The result is less density and less potential for data manipulation, as well as higher cost. There are advantages, however, such as in video quality.

There are three levels of Laser Video disk systems that are commercially available. Level I, the basic level, plays video information in a straightforward linear (CAV or CLV) mode. Movies are typically made at level I. Level I also includes limited interactivity, frame selection, stop and freeze, search options, and chapter stops. Level II includes a microprocessor in the player and has built-in cues in the disk. This allows preprogramming of the video disk with automatic responses and option branches. Level III requires interfacing to a host personal computer. This allows the usual functions of the personal computer to be brought into direct interactivity with the video. Hence the level III system incorporates some adaptive elements of learning into interactivity. *See* COMPUTER; DIGITAL COMPUTER; MICROCOMPUTER.

These functions have resulted in many applications for laser video disks: entertainment; video games with user interaction; instructional devices with branched, cued interactivity; point-of-sale terminals with self-directed pathways; learning systems for self-paced instruction; military training systems; simulation; medical teaching systems; multilanguage training, using dual audio channels; archival storage of pictorial, video, and digital images; and video editing of movies or television. *See* MULTIMEDIA TECHNOLOGY; SIMULATION; VIDEO GAMES.

Video CD. Using the same disk technology used in CD-ROM-XA and combining it with MPEG1 compression, it is possible to play a highly compressed version of a movie. This basically uses 650 MB to record a full movie, and as a result much quality is lost. However, this technology has been very successful in environments where cost is paramount. It is often used for illegal copies of copyright video content. It is not to be confused with an older format called CD-V, which was a hybrid of laser video and CD technology and has largely disappeared. David H. Davies

Bibliography. M. Ely and P. Delancy, *DVD Production with Video*, Butterworth-Heinemann, 2000; A. C. Luther, *Video Recording Technology*, Artech House, 1999; L. Purcell, *CD-R/DVD: Digital Recording to Optical Media*, McGraw-Hill, 2000; J. Ross, *DVD Player Fundamentals*, Prompt Publications, 2000; K. A. Schouhamer Immimk, *Codes for Mass Data Storage Systems*, Shannon Foundation, 1999; J. Taylor, *DVD Demystified*, 2d ed., McGraw-Hill, 2000.

Video games

Entertainment systems in which a computer is used to drive a video display and interact with players using a variety of input devices. Video games can be divided into arcade systems, home computers, and game consoles. The distinction between a home computer and a game console is that a computer can be used for a variety of other applications such as word processing and Internet access while a game console is specifically designed for entertainment purposes.

Arcade systems. Arcade systems are typically built for a particular game or set of related games. Everything from the choice of controls to the design on the sides and face of the unit are geared toward the game itself. The display is typically larger than that on a home computer, providing a greater sense of immersion. In some cases, specialized optics are used to further increase the sense of breadth and depth. In some racing simulations, for example, multiple screens are arranged side by side in an arc around the player in order to give an even greater feeling of being "in" the game.

Input devices on arcade units are more specialized than those on home computers or game consoles. While early arcade systems primarily used

buttons, knobs, and joysticks, the more recent ones have moved far beyond those simple devices. Driving simulations often use full race-car cockpits, and skiing and water-skiing simulators have added a pair of skis that the player can stand on in order to control his or her movements in the game. Light guns which allow the player to "shoot" the screen have proven popular with certain types of action games.

The goal of most of these high-end arcade systems is to provide an experience that cannot be duplicated on a home computer or game console, and without such innovations arcades would rapidly lose business to the growing home market. The chief disadvantage of arcade systems is that they are custom-designed on a per-game basis, which means high development costs and early obsolescence. They are also expensive to distribute, as entire arcade units must be shipped from the manufacturer. There have been at least two initiatives aimed at standardizing arcade hardware so that the units can be reused for different games, but so far there has been relatively little success in that area.

More recently, there has been considerable innovation at the high end of the arcade market. Often described as LBE (location based entertainment), these arcade systems are expensive, custom-designed, and clustered together as part of a larger entertainment complex. These facilities feature a combination of conventional arcade games and specialized high-end systems involving virtual reality gear, large rear-projection screens, motion platforms, and sophisticated input devices. *See* SIMULATION; VIRTUAL REALITY.

Game consoles. A game console typically does not include its own display, but is hooked up to a television set. This provides only limited resolution but dramatically lowers the cost of the consoles themselves. This is critical since a primary selling point of game consoles over home computers is their lower cost. The input devices on game consoles are limited to simple multipurpose controllers that are included with the console. Most consoles come with two controllers in order to support two-player competitive games.

Consoles are much less expensive than the average home computer system and often have specialized hardware for fast graphics and high-quality sound. The games are distributed either as cartridges ("carts") or compact disks (CDs) that contain the game logic programmed into read-only memory (ROM). CD-ROMs have many advantages over carts, including much greater storage capacity and lower production costs. This results in a lower cost of entry to the market, which in turn provides a wider selection of games to the consumer. However, the business model used by many companies that make console games is built on selling the console itself at cost or even taking a small loss on it, and making a significant profit on each game sold. This becomes more difficult if the distribution medium for the game is CD-ROM, since CD-ROM games may be produced by other companies. *See* COMPACT DISK.

Console systems have gone through at least four distinct generations. The first systems were very simple, often involving rudimentary logic that was custom-designed for a particular game. These quickly gave way to systems built around 8-bit microprocessors, such as the original Sega and Nintendo systems. These were superseded by the 16-bit systems which provided greatly increased memory and graphics performance. After some unsuccessful attempts at 32-bit systems, the industry moved to full 64-bit systems, with custom chipsets for sound and three-dimensional (3D) graphics. These systems often rival high-end home computers in terms of graphics capabilities. However, the high cost of market entry for console game developers (relative to the home computer game sector) has limited their selection of games. This, along with the ubiquity of high-performance personal computers in the home, threatens the survival of the console market. *See* BIT; COMPUTER GRAPHICS; COMPUTER STORAGE TECHNOLOGY; MICROPROCESSOR.

Home computers. Home computer systems have grown steadily more powerful to the point where they are more than capable of serving as game machines. Since they use an actual computer monitor rather than relying on a television set, the resolution and overall image quality is much higher than that of game consoles and often rivals or exceeds that of the arcade systems. A variety of controllers are available for the personal computer, ranging from simple analog joysticks to sophisticated input devices with full-force feedback. *See* COMPUTER PERIPHERAL DEVICES.

Production, packaging, and distribution costs for computer games are similar to those for game consoles that use CD-ROM. However, widespread Internet access allows game companies to make sample or "demo" versions of their games available for download. This helps build the market for the games, and may prove to be an effective distribution mechanism for the full retail versions as well. This will avoid the problem of limited shelf-space in computer game stores and will drastically reduce production and distribution costs. *See* INTERNET.

Home computer systems are not without disadvantages. They are typically 10 or 20 times as expensive as an average game console, which limits their market. On the other hand, most families now own a home computer, and many own more than one. Since the computer has its own monitor, the family TV is not monopolized the way it is with a game console.

The other disadvantage that home computers have when compared to arcade systems or game consoles is that there tends to be considerable diversity in terms of hardware capabilities and configuration. Developers must contend with different screen resolutions, color depths, video cards, display drivers, sound cards, input devices, and so on. Even basic information about the end user such as processor speed or the amount of installed memory is unknown during game development. This makes the testing and debugging of computer games much more challenging than game consoles or arcade units.

Types of games. Regardless of the platform (home computer, game console, arcade system), there are a number of basic genres of game. Some of these genres are more common on one type of platform than another, mostly due to technical limitations.

Action games are popular on all platforms. Sometimes called "twitch" games, they tend to be fast-paced and require good reflexes and hand-eye coordination on the part of the player. The storyline for such games is usually very basic—shoot everything that moves. The player is sometimes trying to achieve a goal, such as escaping from a dungeon or rescuing hostages, but the game play is essentially the same regardless.

Adventure games are mostly popular on home computers, though some game consoles do support them. Unlike action games, adventure games combine an element of exploration with some puzzle-solving such as finding keys, unlocking doors, or gathering information. Since these games involve a gradual accumulation of objects in an inventory, as well as the solving of various puzzles, it is essential that there be some way of saving the game state information across playing sessions. On a home computer, this is trivial. Some game consoles also provide some storage capability, although it is usually quite limited. Arcade systems do not have data storage capabilities, so they are generally not suitable for this particular genre. A related genre is role-playing, which is similar to an adventure game but with greater emphasis on playing a particular class of character (such as a magician or a thief) and working to improve that character's skills and abilities.

Simulation games usually involve operating some sort of vehicle, such as an aircraft or a high-performance race car. Flight simulators and vehicle combat simulators (such as tanks or giant walking robots) are in this category as well. Sports simulators are a slightly different category, since they are focused around an individual or team sport such as golf, hockey, football, or soccer. Since this genre shares many characteristics with action games, simulators can also be found on all popular platforms.

Strategy games are based on planning and anticipating future events. There are two subcategories: war games and system simulators. War games are often based on reenacting historic battles of land, sea, or air. The player takes on the role of a general or commander and plans a strategy that hopefully will win the war. System simulators, on the other hand, work by simulating a system such as a city, an ant colony, or an entire planet and allowing the player to make choices that affect the grown and development of that system. Anything from urban planning to maintaining a global ecosystem can be part of the challenge. Since the strategy games often take time to plan and develop, they tend to be found mostly on home computers and are not seen in arcade systems.

Simple puzzle games are the final genre, and they can be found on all platforms. However, their popularity has declined relative to the other gaming genres.

Future directions. A number of new developments will affect arcade systems, game consoles, and home computers. Most important is the steady improvement in the quality and performance of 3D graphics accelerators. Since these chips are designed specifically to handle the final stages of the graphics "pipeline," they can be optimized for that task in order to obtain maximum performance. They also relieve the main central processing unit (CPU) of the burden of doing the low-level graphics rendering, which leaves more computing cycles available for other aspects of the game.

The next area where innovation can be anticipated is the development of 3D audio systems. Just as 3D graphics accelerators take over the work of rendering images on the screen, 3D audio chipsets handle the generation and spatialization of sound. The increasing sophistication of audio hardware and software will allow sound to be positioned in three dimensions around the player as they travel through the game.

The third area that will see rapid growth is support for multiplayer gaming. Home computers already support this feature, using a standard Internet connection to provide head-to-head or team play. Game consoles will likely also gain that capability, and even some arcade units have the ability to link to other systems over custom networks.

Another important development will be the growing use of force feedback, which provides sensations such as vibration, impact, and resistance to movement. Some very high-end arcade systems have even added motion platforms, which tip and tilt the players as they travel through the game.

As the home computer market grows, the other two sectors will continue trying to find ways to compete effectively. New generations of the popular game consoles come out every couple of years, offering graphics and sound improvements that overshadow all but the highest-end home computer systems. Arcade units are attempting to differentiate themselves from both home computers and game consoles by providing ever more sophisticated input devices. This competition between platforms is likely to produce some amazingly sophisticated and compelling games. *See* HUMAN-COMPUTER INTERACTION.

Bernie Roehl

Bibliography. *Computer Gaming World Magazine*; *Game Developer Magazine*; *Next Generation Magazine*.

Video microscopy

The use of a high-quality video camera or other fast camera (such as a charge-coupled device) attached to a research-quality light microscope for the purpose of real-time or high-speed imaging of samples on a microscope stage. These images are recorded at regular intervals (often at "video rate" of 30 images per second), and the time-lapse sequence can be played back in the form of a movie. The term "video microscopy" originally referred to microscope imaging

using true video (30 frames per second) but now generally refers to rapid time-lapse imaging techniques. Video microscopy is used frequently to image small structures that move rapidly within cells as well as movement of whole cells. This motion can be quantitated, and in the case of fluorescence microscopy, changes in fluorescent intensity (reflecting the local chemical environment of the fluorescent molecule or the number of fluorescent molecules) can be quantitated as well. *See* CAMERA; MICROSCOPE.

Conventional video systems. Video microscopy is not very useful for stained biological material or specimens that have high contrast. It is most useful for specimens that have extremely weak contrast or are invisible by conventional microscopy. In practice, the video signal from the camera is digitized and enhanced at video frequency by an image processor. Contrast is enhanced first electronically by changing the offset of the camera, and then digitally by choosing the peak gray levels (usually about 30 levels) of the biological image and expanding them over 256 levels of gray. The latter operation is performed by assigning each gray level in the original image to a new value through the use of an output look-up table. *See* OPTICAL MICROSCOPE.

Image processors are able to perform operations in real time. The image processor is also used to improve image quality by averaging successive frames to remove electronic noise and by subtracting a stored image of the background to remove fixed-pattern mottle (such as dust spots) on the lenses. Subtraction of the background image at video frequency also serves to correct minor unevenness of illumination. The resulting image exhibits high contrast and additional detail that is invisible when viewed by the eye through the microscope. Microscopes equipped with Nomarski differential interference contrast optics offer the best visualization of submicroscopic objects by video microscopy. This type of microscopy permits visualization of biological objects as small as 20 nanometers and solid objects as small as 5 nm (for example, colloidal gold particles). Biological objects detected by this method are an order of magnitude smaller than the limit of resolution of light microscopes. *See* IMAGE PROCESSING; INTERFERENCE MICROSCOPE.

Charge-coupled-device systems. Charge-coupled-device (CCD) cameras use a light-sensitive silicon chip as the detector. Often the chip is cooled to very low temperatures (-20 to -40°C) to reduce noise. CCD chips have the advantage that the signal can be easily digitized and stored on a computer, and this digital signal is a linear measure of the amount of light recorded by the chip (and therefore of the number of fluorescent molecules). They are often better than conventional video cameras at recording images of fluorescent objects (since fluorescence images are often dim, and CCD cameras allow varying the length of exposure to collect more signal). Conventional video cameras are still preferred if transmitted light is used. *See* CHARGE-COUPLED DEVICES.

Motility assays. Video microscopy is used routinely to examine the structure and function of the cellular cytoskeleton and to detect the movement of submicroscopic organelles in living cells. The filaments of the cytoskeleton function as tracks on which protein "molecular motors" transport material to different locations in the cell. Microtubules (one class of cytoskeletal filaments) and membrane-enclosed vesicles, both smaller than the limit of resolution, can be detected by video microscopy in living cells and in reconstituted cell-free systems. *See* CYTOSKELETON.

The video microscope has also been used to examine the movement of objects by molecular motors during motility assays. The most common type of motility assay is the gliding filament assay. First, the motor protein of interest is adsorbed to a glass coverslip. Then the cytoskeletal filaments on which these motors move are added to the coverslip, and video microscopy is used to view the movement of filaments over the glass surface. The filaments appear to glide over the glass surface. This assay has become a standard procedure for studying the microtubule-dependent motors, kinesin and dynein. *See* CELL MOTILITY.

Another method is to adsorb the filaments to the glass surface first, then add organelles or artificial beads coated with motors, and view movement of the organelles and beads along the filaments. The latter method has been used successfully to study both microtubule-dependent motors and the actin filament-dependent motors known as myosins. Actin filaments (another class of cytoskeletal filaments) are too small to be detected by video microscopy, but vesicles moving on them can be seen.

Fluorescence ratio imaging. Rapid advances in low-light-level cameras have made it possible to expand video microscopy to study activities in living cells by using fluorescent dyes. Living cells can be stained with an ion-sensitive fluorescent dye, and changes in the fluorescence excitation or emission can be used to monitor cellular activity. The quantitation of the optical signal from living cells is best achieved by using fluorescence ratio imaging. This technique is used to measure ion concentrations such as hydrogen and calcium, and requires that the fluorescent probe be sensitive to changes in the concentration of the ion of interest at one excitation or emission wavelength but insensitive at another. The ratio of the fluorescence intensity at the two different wavelengths normalizes for optical pathlength, local probe concentration, and loss of signal due to photobleaching (loss of fluorescence due to exposure to light). The fluorescence ratio-imaging microscope requires a standard fluorescence microscope, a low-light-level detector (often a CCD camera), and an image processor. An image processor is required for image acquisition and basic arithmetic operations such as addition and subtraction; basic point operations such as intensity remapping and thresholding; basic analysis operations such as histogram and statistical calculations; and basic display operations such as image scroll, zoom, and pseudocolor display. *See* FLUORESCENCE; FLUORESCENCE MICROSCOPE.

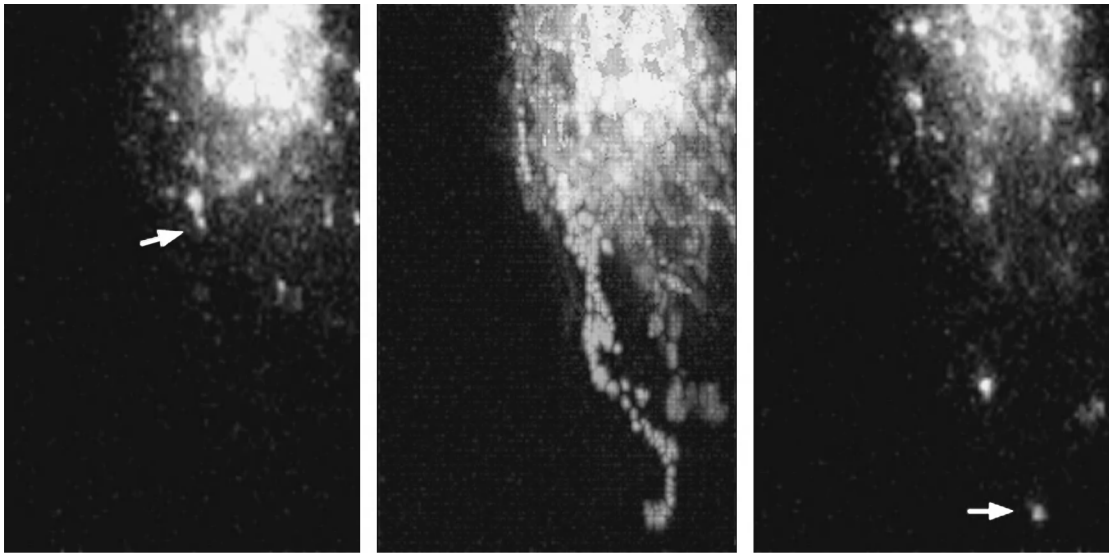

Use of video microscopy to track an intracellular object (at arrow) labeled with a GFP-tagged protein as it moves away from the Golgi apparatus (large area of bright fluorescence). The beginning frame (left) and ending frame (right) of the sequence are shown. The center frame is a superposition of the entire sequence showing the object's path as it moves through the cell.

Fluorescence recovery after photobleaching. Video microscopy in combination with fluorescent probes can also be used to measure dynamic changes in the structural components in cells or the rate of diffusion of macromolecules in cells. Since most fluorescent probes bleach when illuminated for some period at the excitation wavelength, this property can be exploited to measure dynamic changes in cells. This technique is referred to as fluorescence recovery after photobleaching, and requires that the area of the cell around the region of interest be masked to protect it from the light while the region of interest is illuminated. Fluorescent molecules from the unbleached region diffuse to the bleached region, and the rate of recovery of fluorescence can be determined. This technique has been used to demonstrate that cytoskeletal filaments exchange subunits with a monomer pool and that macromolecules diffuse slowly through the porous matrix of the cytoplasm. The technique has also been used to determine how fast the membrane proteins diffuse in membranes such as the plasma membrane. Importantly, it is easy to distinguish which molecules are fixed in place in the cell (by attachment to cytoskeleton or other structures) and which are free to diffuse. *See* CELL ORGANIZATION.

Green fluorescent protein. Until recently, attaching fluorescent probes to proteins required either chemical attachment of the fluorescent probe to the protein of interest followed by introduction of the protein to the cell, or the addition of fluorescently labeled antibodies to the protein of interest in situ. These methods work well if the protein of interest is one that is normally taken up by the cell or if it is exposed on the cell surface. However, the overwhelming majority of cellular proteins could be fluorescently labeled only with great difficulty if they were to be used in a living cell.

The green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* has been used as a fluorescent tag. GFP is fluorescent without the attachment of another molecule. Well-established recombinant DNA techniques make it possible to make fusion proteins in which GFP is attached to any cellular protein. Upon introduction of fusion protein DNA into cells in culture, the protein tagged with GFP is synthesized by the cell just like the normal cellular version of the protein. Since this protein is fluorescent, its movements inside the cell can be easily followed by video microscopy. The **illustration** shows an example of a structure visualized with a GFP-tagged protein moving within a cell. It is important to verify that the GFP tag does not interfere with the normal function of the protein. GFP variants (and even a red fluorescent protein) have been developed that allow for tagging and following up of three proteins simultaneously in the same cell. *See* BIOLUMINESCENCE; GENETIC ENGINEERING.

George M. Langford; John F. Presley

Bibliography. R. D. Allen, The microtubule as an intracellular engine, *Sci. Amer.*, 255(2):42–49, 1987; R. D. Allen and N. S. Ailen, Video-enhanced microscopy with a computer frame memory, *J. Microsc.*, 129:3–17, 1983; S. Inoue and K. Spring, *Video Microscopy: The Fundamentals*, 1997; S. Scales, R. Pepperkok, and T. E. Kreis, Visualization of ER-to-Golgi transport in living cells reveals a sequential mode of action for COPII and COPI, *Cell*, 90:1137–1148, 1997; D. L. Taylor and Y.-L. Wang (eds.), *Methods in Cell Biology*, vol. 30, 1989.

Videotelephony

A means of simultaneous, two-way communication comprising both audio and video elements. Participants in a video telephone call can both see and hear each other in real time. Videotelephony is a subset of teleconferencing, broadly defined as the various

ways and means by which people communicate with one another over some distance. Initially conceived as an extension to the telephone, videotelephony is now possible using computers with network connections. In addition to general personal use, there are specific professional applications, such as criminal justice, health care delivery, and surveillance that can greatly benefit from videotelephony. *See* TELECONFERENCING.

History. Although basic research on the technology of videotelephony dates back to 1925, the first public demonstration of the concept was by American Telephone and Telegraph Corporation (AT&T) at the 1964 New York World's Fair. The device was called Picturephone, and the high cost of the analog circuits required to support it made it very expensive and thus unsuitable for the market place. The 1970s brought the first attempts at digitization of transmissions. The video telephones comprised four parts: a standard touch-tone telephone, a small screen with a camera and loudspeaker, a control pad with user controls and a microphone, and a service unit (coder-decoder, or CODEC) that converted the analog signals to digital signals for transmission and vice versa for reception and display. This system required a wideband infrastructure (telephone circuits) with a capacity of 6.3 megabits/s. *See* ANALOG-TO-DIGITAL CONVERTER; DIGITAL-TO-ANALOG CONVERTER.

During the 1980s companies attempted to market videoconferencing for both public and private use, for example, setting up nationwide networks of videoconference rooms.

Analog video transmission. The primary limiting factor in videotelephony development has been the requirement for switched broadband circuits. The public telephone infrastructure was developed for narrow-band voice communication only. Television network broadcasts and television cable distribution systems are wideband in nature, but they are not switched between users and are only one-way. Transmission requirements can be reduced considerably by utilizing smaller screens with less spatial resolution, by reducing image refresh rates, and by delivering black-and-white images only. However, the requirements are still far beyond the capacity of ordinary telephone lines. Due to these limitations, true analog videotelephony has never really developed. Therefore, considerable attention has been devoted to digitizing and compressing video signals to make their transmission more economical. *See* CABLE TELEVISION SYSTEM; TELEPHONE SERVICE; TELEVISION NETWORKS.

Digital video transmission. Pure digitization of video signals is normally done using pulse code modulation (PCM), which requires sampling and encoding to produce a transmission rate of 45 megabits/s. The video telephones from the 1970s used the differential pulse-code modulation (DPCM) compression technique, which reduced the required transmission rate to 6.3 megabits/s. This technique transmits only the differences between successive samples. Aother compression technique popularized in the 1980s is the discrete cosine transform (DCT), which takes advantage of the fact that several adjacent points within an image essentially are the same. A third, widely used compression algorithm, called motion compensation, is based on the fact that moving objects keep the same form. Thus an image of an object must be transmitted only once, and its subsequent movement requires transmission of only its new position and orientation. *See* DATA COMPRESSION; PULSE MODULATION.

During the 1980s the International Telecommunications Union (ITU) agreed to use a standard for video coding and decoding called H.261. This standard uses a combination of the three compression algorithms and specifies standard image formats. Using the standard, good-quality video images may be communicated at transmission rates as low as 64 kilobits/s. A slightly different version of this standard is also common, called H.263. Several audio standards are in use: G.723, G.711, G.722, and G.728. Two general, higher-level videotelephony/videoconferencing standards developed by the ITU are commonly adhered to: H.320 for ISDN videotelephony and H.323 for IP videotelephony.

Infrastructure. Considerable improvements in telephone, computer, and television infrastructures have been made over the years. Telephone connections over copper cable have been replaced by optical fiber, the capacity of connections to the home has been increased and standardized, and more efficient technologies to use transmission capacity are available. Cable television connections have become interactive. Computer/data communication infrastructures, such as the Internet, World Wide Web, and high-speed, wide-band links such as digital subscriber line (DSL) and cable modems have become more widespread. These developments have converged to make video communication possible at affordable rates. *See* INTERNET; MODEM; OPTICAL COMMUNICATIONS; WORLD WIDE WEB.

The basis of communication over the Internet is the Internet Protocol (IP), which allows for data in packets to be routed to different addresses using the Transmission Control Protocol (TCP). The data in the packets may be voice, video, data, graphics, and so forth. Thus, multimedia exchanges may take place between two or more parties. The exchanges can be over the public Internet or over private corporate intranets. Some companies have started replacing their traditional telephone private branch exchanges (PBXs) with IP-based PBXs, which would theoretically allow everyone in the company to have a video or multimedia telephone—totally separate from a computer. This could provide, for example, unified messaging—the ability to store voice mail as e-mail, answer telephone calls with e-mails and vice versa. *See* ELECTRONIC MAIL; PRIVATE BRANCH EXCHANGE.

Video telephone hardware. Small residential video telephones, computer-based desktop video telephones, and small videoconferencing setups have been introduced to fulfill diverse needs. One such

commercially available residential videophone is about as big as a typical office desk telephone with a small flip-up screen that has an eyeball camera above it. Although it will work with several standards, this phone is primarily designed for use over Integrated Services Digital Network (ISDN) lines in which a residence gets three circuits; one circuit is used for control and the other two for voice and video. *See* INTEGRATED SERVICES DIGITAL NETWORK (ISDN).

An example of a computer-based desktop videophone consists of a PCI (Peripheral Component Interconnect) video/audio CODEC board to add to a personal computer, a composite color camera, audio peripherals, and visual collaboration software.

Video telephone software. Videotelephony software has been developed and made widely available that permits real-time collaboration and conferencing, including multipoint and point-to-point conferencing. Multipoint means, for example, that three people in three different locations could have a video telephone conference call in which each could see and hear the others. In addition to the basic audio and video capabilities, such software provides several other features such as a whiteboard, background file transfer, program sharing, and remote desktop sharing. John Bleiweis

Bibliography. E. M. Dickson and R. Bowers, *The Video Telephone*, 1974; S. J. Emmott and D. Travis (eds.), *Information Superhighways: Multimedia Users and Futures*, 1995; K. Finn, A. J. Sellen, and S. B. Wilbur, *Video-Mediated Communication* (*Computers, Cognition and Work*), April 1997; A. M. Noll, Anatomy of a failure: Picturephone revisited, *Telecomm. Policy*, 16:307–316, 1992; R. Schaphorst, *Videoconferencing and Videotelephony: technology and Standards*, Artech House Telecommunications Library, January 1997.

Vinegar

A food condiment containing mainly acetic acid that is produced by the bacterial oxidative fermentation of various ethanolic solutions. Vinegar, like wine and certain milk products, is one of the oldest fermented foods used by humans. Babylonian records indicate that vinegars prepared from wines and beer were widely used as early as 5000 B.C. *See* ACETIC ACID; WINE.

Types. The names for the different vinegars are based on the substrates from which they are made. These include the juices from different fruits, starchy vegetables, cereals, and distilled ethanol (ethyl alcohol). In the United States, white distilled vinegar makes up over 80% of the annual production. *See* ETHYL ALCOHOL.

Much of the white distilled vinegar is made from denatured synthetic ethanol. Nitrogen compounds, minerals, and other nutrients must be added to the alcohol medium to support the growth and metabolism of the acetic acid bacteria. Acetic acid or ethyl acetate are commonly used as the ethanol denaturants.

Most other vinegars are made from ethanolic solutions generated by a yeast fermentation. Fruit juices usually contain sufficient concentrations of hexose sugars and other yeast nutrients to support yeast growth and fermentation. Cider vinegar is produced from yeast-fermented apple juice, wine vinegar from fermented grape juice. Other juices such as pineapple, orange, and pear, as well as sugarcane syrup and molasses, can serve as fermentation substrates.

Vinegar is also produced from starchy vegetables such as potatoes and from cereals such as barley, corn, and rice. Yeasts do not ferment starch to ethanol, and therefore the starch first must be broken down (saccharified) to simple sugars, mainly glucose. Malt vinegar is made from an infusion of barley malt and other cereals in which the enzymes in the malt have converted the starch to fermentable sugars. *See* ENZYME.

Balsamic vinegar from Modena, Italy, is a costly product that is highly prized by gourmets. The juice of Trebbiano grapes is concentrated about 30% by simmering, then fermented by *Zygosaccharomyces* yeasts that tolerate high concentrations of sugar. A portion of the ferment is transferred annually to a succession of wooden casks, where the acetic acid fermentation and other enzymatic changes occur. The battery of casks may be 200 years or older; thus, the harvested vinegar is of great age. About 2 liters (2 qt) of vinegar are produced from 100 kg (220 lb) of the concentrated grape must.

Vinegar sold in the United States must contain at least 4 g of acetic acid per 100 ml of solution

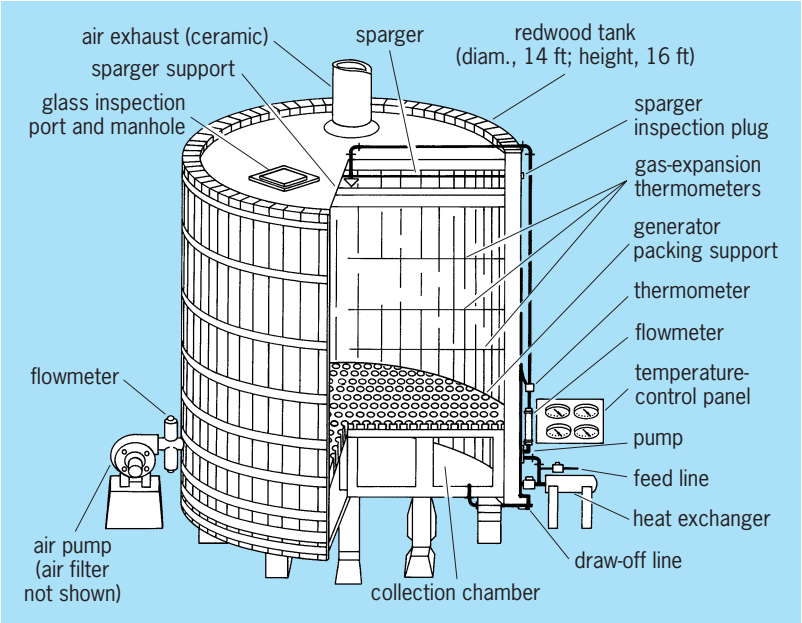

Fig. 1. Trickling, quick vinegar generator. The diameter of the redwood tank is 14 ft (4 m), and the height is 16 ft (5 m). (*After L. A. Underkofler and R. J. Hickey, eds., Industrial Fermentations, 2 vols., Chemical Publishing, 1954*)

(40-grain vinegar). Other federal specifications describe permitted color, odor, presence of trace metals, and alcohol content. Nondistilled vinegars possess distinctive colors and flavors that reflect the properties of the original substrate.

Acetic acid bacteria. Bacteria belonging to the genus *Acetobacter* are primarily responsible for the vinegar fermentation. Four species are recognized; they are distinguished on the basis of nutrient requirements, production of brown pigments, and tolerance to ethanol. All acetic acid bacteria require oxygen for growth and metabolism.

Gluconobacter is a closely related genus that also oxidizes ethanol to acetic acid. The genus is differentiated from *Acetobacter* by the fact that it cannot oxidize acetic acid to carbon dioxide and water. *Acetobacter* species carry out the latter oxidation, often after ethanol has been depleted. *See* INDUSTRIAL MICROBIOLOGY.

Production methods. A variety of methods are used for the production of vinegar. The simplest fermentations are slow and inefficient but are relatively foolproof and utilize inexpensive equipment. Submerged fermenters produce 120–150-grain vinegar in minimal time, but they are expensive and system failures can occur.

Orleans process. This is one of the older commercial methods. The fermentation is carried out in a large cask in which holes have been drilled to permit the introduction of air. Screens are used to prevent the entrance of fruit flies. The cask has a spigot for the withdrawal of finished vinegar, and the bung (stopper) contains a tube so that fresh wine or other substrate can be added without disturbing the film of vinegar bacteria. The fermentation is initiated by mixing wine with about 20% fresh vinegar. Although the process is slow and the yields of acetic acid may be under 80%, vinegar of high quality can be produced.

Trickling, quick process. In this procedure, the alcoholic substrate is sprayed on the surface of a chamber packed with beechwood shavings, corn cobs, or other carrier materials that support a slime composed of acetic acid bacteria (**Fig. 1**). Air is sparged through the system. Because considerable heat is generated, cooling of the chamber or recirculated vinegar stock is usually required. The vinegar stock is recirculated two or three times until the desired concentration of acetic acid is attained. Generators of this type may be operated continuously for several years or more.

Submerged fermentation. In submerged fermentation (**Fig. 2**), air is sucked into the bottom of a stainless steel tank and then broken into very small bubbles by the action of a high-speed rotor. Advantages of this system are that vinegar is produced at a high rate, about 10-fold faster than the trickling filter, and the efficiency may be as high as 94%. One disadvantage is that loss of aeration for even 1 min may result in death of the culture. Because of this problem, many vinegar producers have auxiliary electric generators as backups in the event of power failure.

Final processing steps are filtration to clarify the vinegar and pasteurization at 140–158°F (60–70°C) to arrest further bacterial growth and enzyme activity. When concentrated vinegar is desired, for example, 200-grain, this can be accomplished by freezing out some of the water from a 120-grain product. *See* FERMENTATION.

Don F. Splittstoesser

Bibliography. H. A. Conner and R. J. Allgeier, Vinegar: Its history and development, *Adv. Appl. Microbiol.*, 20:81–133, 1976; H. J. Peppler and D. Perlman (eds.), *Microbial Technology*, 2d ed., vol. 2, 1979; M. Plessi, A. Monzani, and D. Coppini, Determination of the monosaccharide and alcohol content of balsamic and other vinegars by enzymatic methods, *Agr. Biol. Chem.*, 52:25–30, 1988; H. J. Rehm and G. Reed (eds.), *Biotechnology*, vol. 3, 2d ed., 1994.

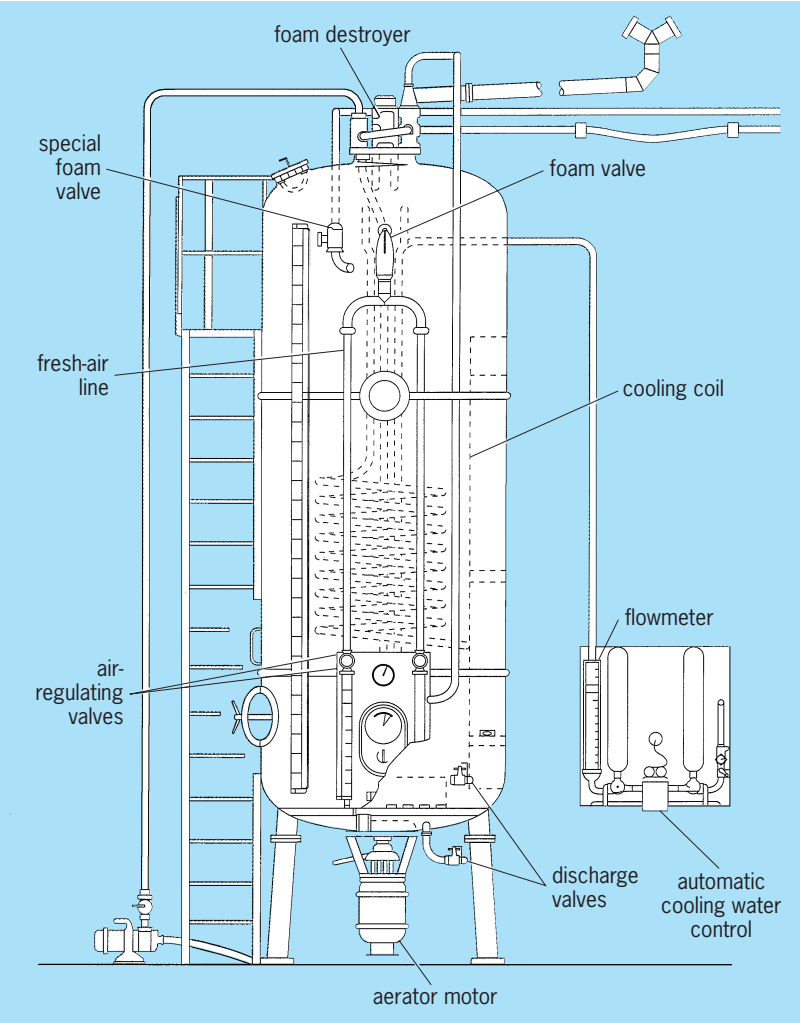

Fig. 2. Frings acetator for submerged fermentation. (*After G. B. Nickol, Microbial Technology, 2d ed., vol. 2, Academic Press, 1979*)

Violales

An order of flowering plants, division Magnoliophyta (Angiospermae), in the subclass Dilleniidae of the class Magnoliopsida (dicotyledons). The order

consists of 24 families and more than 5000 species, including most of the families that have been referred to the Englerian order Parietales. The largest families are the Flacourtiaceae (about 800 species), Violaceae (about 800 species), Cucurbitaceae (about 900 species), Begoniaceae (about 1000 species), and Passifloraceae (about 650 species). The Cucurbitaceae have often been considered to constitute a distinct order, Cucurbitales, but there is a developing consensus that their relationship is here.

The most characteristic feature of this morphologically heterogeneous order is the unilocular, compound ovary with mostly parietal placentation. When the stamens are numerous they are initiated in centrifugal sequence. The ovules have two integuments and a nucellus that is several cells thick. The more primitive members of the order, such as some of the Flacourtiaceae, are trees with stipulate, alternate leaves; perfect, polypetalous flowers with numerous centrifugal stamens; a compound pistil with free styles and parietal placentation; and seeds with a well-developed endosperm. Tendencies toward unisexuality, reduction in the number of stamens, fusion of filaments, development of a corona, reduction in the number of carpels, fusion of styles, and loss of endosperm from the seed can all be seen in the Flacourtiaceae. These are also some of the prominent characteristics used in combination to define many of the other families of the order. *See* FLOWER; LEAF; SEED.

The Violales differ from the related order Capparales (which also has parietal placentation) in usually lacking myrosin cells, in having a much higher proportion of woody species, in often having perigynous to epigynous flowers, in seldom having compound leaves, and in very often having three carpels (a rare number in the Capparales) and only seldom two (the most common number in the Capparales). *See* CAPPARALES.

Cucumbers and melons (species of *Cucumis*, in the Cucurbitaceae), pumpkins and squashes (species of *Cucurbita*), begonias, and violets (*Viola*; see **illus.**) are familiar members of the Violales. *See* CANTALOUPE; CUCUMBER; DILLENIIDAE; HONEY DEW MELON; MAGNOLIOPHYTA; MAGNOLIOPSIDA; MUSKMELON; PERSIAN MELON; PLANT TAXONOMY; PUMPKIN; SQUASH; WATERMELON.

Arthur Cronquist; T. M. Barkley

Downy yellow violet (*Viola pubescens*). (*Photograph by Elsie M. Rodgers, National Audubon Society*)

Viral inclusion bodies

Abnormal structures formed in the cytoplasm or the nucleus of the host cell (or both) during the multiplication of some viruses. These structures vary from virus to virus and may represent aggregations of mature virus particles but more often are areas where viral synthesis takes place. These areas of viral synthesis have an altered appearance when treated with typical diagnostic stains.

The presence of inclusion bodies is often important in diagnosis. Brain tissue taken from suspected cases of rabies is examined for the presence of cytoplasmic inclusions known as Negri bodies. In papillomavirus infections, inclusion bodies are found in the nuclei of the epidermal cells in which typical virus particles can be seen by electron microscopy. In human herpes simplex infection, intranuclear inclusions are found that are surrounded by a clear halo. These inclusions are sites of replication of the virus, or "virus factories." *See* HERPES; RABIES; VIRUS.

Marcia M. Pierce

Bibliography. L. Collier and J. Oxford, *Human Virology*, 2d ed., Oxford University Press, Oxford, U.K., 2000; D. M. Knight and P. M. Howley (eds.), *Fundamental Virology*, 4th ed., Lippincott, Williams & Wilkins, Philadelphia, 2001; L. M. Prescott, J. P. Harley, and D. A. Klein, *Microbiology*, 6th ed., McGraw-Hill, New York, 2005; S. Schlesinger and M. Schlesinger, Viruses, in J. Lederberg (ed.), *Encyclopedia of microbiology*, 2d ed., vol. 4, pp. 796–810, Academic Press, San Diego, 2000; B. A. A. Voyles, *The Biology of Viruses*, 2d ed., McGraw-Hill, Chicago, 2002.

Virgo

The Virgin, perhaps Astraea, in Greek mythology the goddess of justice, whose scales (the constellation Libra) are adjacent. Also, she could be Ceres, goddess of agriculture, or her daughter Proserpina who, because she ate from a forbidden pomegranate, must spend half her year with Pluto, her husband, in Hades, leaving winter behind for the rest of us. This southern zodiacal constellation is visible in the summer evening sky from midnorthern latitudes. It contains the nearest cluster of galaxies to us; the Milky Way Galaxy is in an outer part of the cluster. *See* LIBRA; VIRGO CLUSTER; ZODIAC.

Spica, the brightest star in Virgo, at first magnitude is the 16th brightest star in the sky. It is a B-star, a hot star on the main sequence of the Hertzsprung-Russell diagram, and glows blue-white. It is slightly variable in brightness. Since Spica means an ear of wheat or corn, it would match the identification of Virgo with Ceres or Proserpina. *See* SPICA.

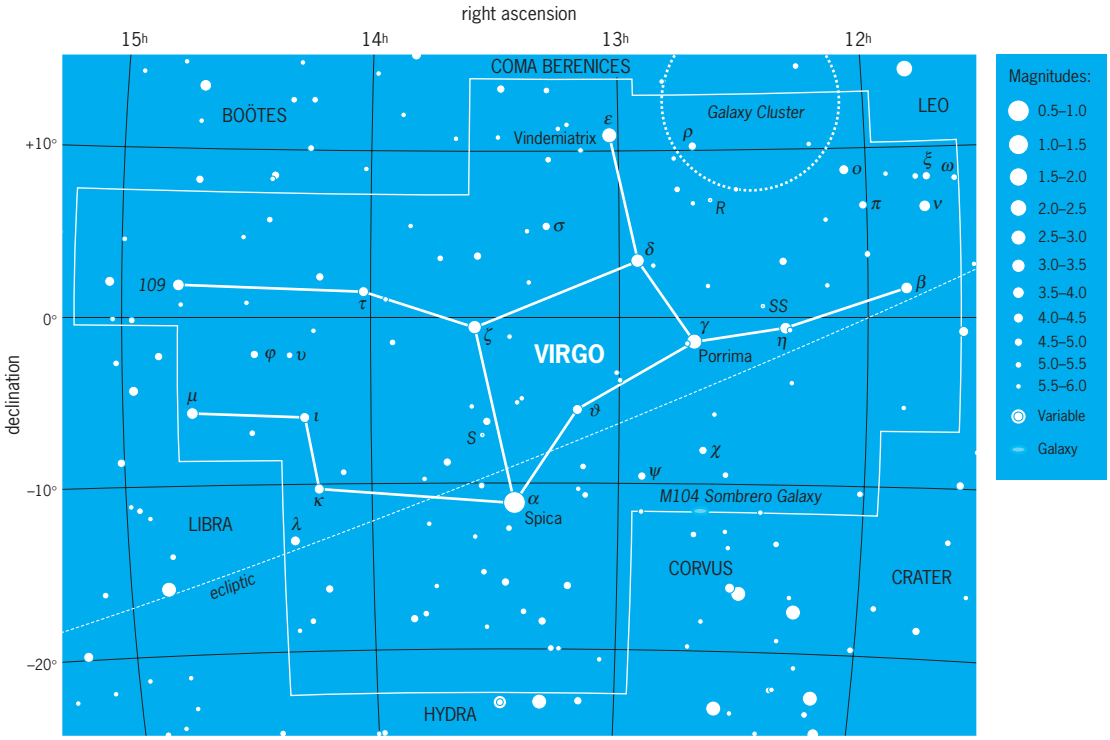

Modern boundaries of the constellation Virgo, the Virgin. The celestial equator is 0° of declination, which corresponds to celestial latitude. Right ascension corresponds to celestial longitude, with each hour of right ascension representing 15° of arc. Apparent brightness of stars is shown with dot sizes to illustrate the magnitude scale, where the brightest stars in the sky are 0th magnitude or brighter and the faintest stars that can be seen with the unaided eye at a dark site are 6th magnitude. (*Wil Tirion*)

The modern boundaries of the 88 constellations, including this one, were defined by the International Astronomical Union in 1928. *See* CONSTELLATION.

Jay M. Pasachoff

Virgo Cluster

The nearest large cluster of galaxies and one of the best studied. It dominates the distribution of bright galaxies in the sky, forming a conspicuous clump in the distribution of easily visible elliptical and spiral galaxies. The clump lies mostly in the constellation of Virgo but also extends into neighboring constellations, especially Coma Berenices. The well-known spirals M58, M61, and M100, for example, are members of the Virgo Cluster. *See* CONSTELLATION; UNIVERSE; VIRGO.

Structure. The first exhaustive exploration of the Virgo Cluster was the Du Pont Telescope survey, carried out in the 1980s by A. Sandage and B. Binggeli. Their survey showed that the cluster has the appearance of two superimposed clusters (**Fig. 1**). The two concentrations are approximately 5° apart; one is centered near the very luminous elliptical galaxy M87 and the other near M49, another giant elliptical galaxy. They lie along a nearly north–south line. The spiral galaxies are much less concentrated toward these central peaks than are either the giant ellipticals or the dwarf ellipticals. Furthermore, the spirals are more common, relative to other types, in the southern concentration than in the northern. It is also found that the dwarf elliptical galaxies are less concentrated toward the center of the entire complex than are any of the other types, possibly the result of mass segregation, with the more massive

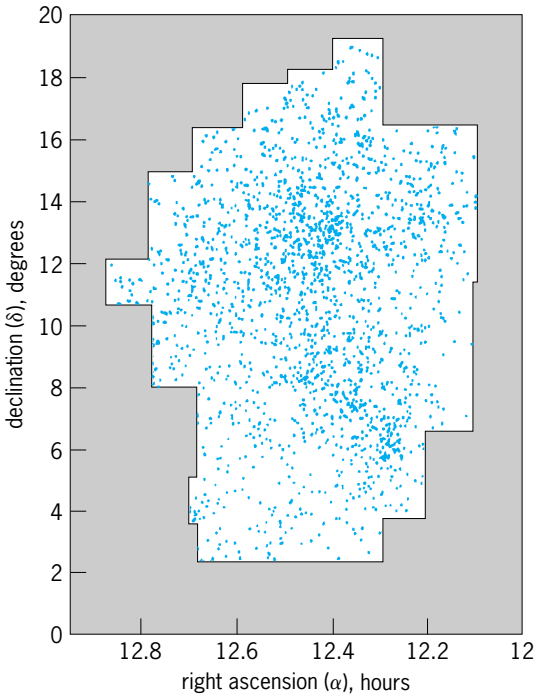

Fig. 1. Distribution of members of the Virgo Cluster on the sky, as determined in the Du Pont Telescope survey. (*After O.-G. Richter and B. Binggeli, eds., The Virgo Cluster, European Southern Observatory, 1985*)

galaxies, under the effects of gravity and encounters, having fallen toward the center of mass more than the low-mass ones.

Populations. The Virgo Cluster is an example of a mixed-population cluster, with both early-type (elliptical and S0) and late-type (spiral and irregular) galaxies present. For members brighter than apparent blue magnitudes of approximately 15.7, about 75% are early types; most of these are dwarf galaxies of low intrinsic luminosity. The remainder are mostly spiral and dwarf irregulars. The most common type of spiral is Sc, of which there are 32 members out of the whole sample of 1343 probable members brighter than this limit (**Fig. 2**).

The total population of all types and of all intrinsic luminosities is not known, but it can be estimated from the surveys and by extrapolating to the faint end of the galaxy luminosity function (the relationship between number and luminosity of galaxies). For galaxies brighter than an absolute magnitude of about −10, a total population of 1900 galaxies is calculated. This excludes the several outlying groups and clusters, such as the Local Group (which includes 30 galaxies). *See* LOCAL GROUP.

Galaxy interactions. As the nearest large cluster, the Virgo Cluster provides a good testing ground for theories of how the dense cluster environment affects the properties of galaxies. The first peculiarity of the Virgo galaxies to be noticed was the rather limited extent of the disks of the spirals. Especially evident in their distribution of neutral hydrogen, as measured at radio wavelengths, the spiral galaxies clearly have been stripped of their outer gas by encounters with other galaxies in the cluster. A related phenomenon is the presence of so-called anemic galaxies, spirals that have unusually low surface brightness and inconspicuous spiral arms. There are several such galaxies in the Virgo Cluster, the result of direct collisions or other tidal disturbances that have removed gas from the disks, inhibiting star formation there and leaving unusually pale galaxies in place of once-robust ones.

Recent studies with large optical telescopes (for instance, with the Keck Telescopes, the Very Large Telescope, the Gemini North Telescope, and other giant telescopes), together with infrared, x-ray, and radio investigations, have shown evidence of violent gravitational interactions among members. For instance, the galaxy NGC 4438, noted as peculiar years ago from Palomar Observatory photos, has been found to have a long tidal tail of debris (stars and gas), ejected by a near-collision with a neighbor. A large cloud of gas lies to one side, produced by ram-pressure stripping, apparently in an earlier collision.

A massive cloud of very hot gas that surrounds the Virgo Cluster is detected by x-ray and radio telescopes (**Fig. 3**). Apparently it is the result of the heating of both intergalactic and intragalactic gas, some of it ejected by collisions of galaxies.

Surveys from space. Surveys of the Virgo Cluster galaxies have used various large ground-based and space telescopes to examine the members in detail. Notable results of some of these surveys have shown dramatic evidence of the complicated evolution that occurs in a cluster environment. There is a large number of dwarf elliptical galaxies, many with bright nuclei, which are found to be remnants of low-mass star-forming galaxies that were stripped of their gas in the recent past. Massive galaxies like M87 (**Fig. 4**) are found to have captured intergalactic gas, which has fallen toward its nucleus, where it is swallowed by a massive black hole. Just before it disappears, the gas is violently heated and some is ejected as a jet that can be detected at virtually all wavelengths, from x-rays to radio waves. Surveys by the *Hubble Space Telescope*, the *Spitzer Infrared Telescope* (**Fig. 5**), and the *Chandra X-ray Observatory* have brought new understanding of galaxy evolution in the cluster, where interactions and collisions leave their

Fig. 2. Central area of the Virgo Cluster, showing some of its brightest galaxy members. (*Canada-France-Hawaii Telescope image, courtesy of C. Cuillandre*)

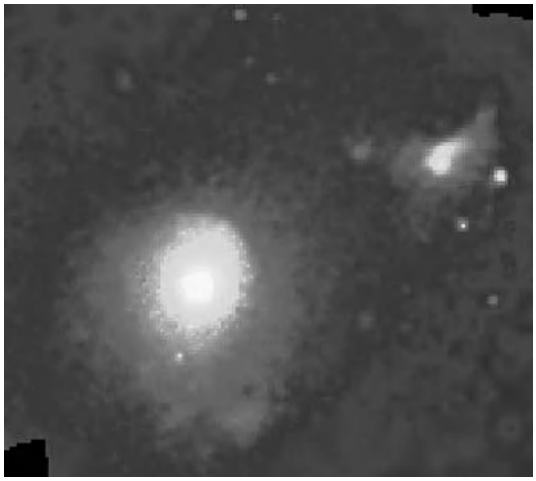

Fig. 3. X-ray image of the entire Virgo Cluster. The main blob is centered on the giant active galaxy M87 (see Fig. 4). (*ROSAT image, courtesy of NASA*)

Fig. 4. Giant elliptical galaxy M87. It is one of the brightest galaxies in the Virgo Cluster and is notable for its extremely massive central black hole. (*CFHT image, courtesy of C. Coullandre*)

indelible marks. *See* CHANDRA X-RAY OBSERVATORY; HUBBLE SPACE TELESCOPE; SPITZER SPACE TELESCOPE.

Dynamics. The dynamics of the cluster is more complex than was once thought. It is not a simple, single group, acting under gravity in a symmetrical way. The velocities in the core are complicated. The elliptical galaxies form a relatively regular, concentrated subcluster that may constitute a dynamically smoothed system, in which gravitational interactions have erased any irregularity of shape. However, the spirals have a wider distribution in space and an accompanying larger spread in velocities; they seem to be still in the process of falling into the cluster core.

The velocities can give a value for the entire cluster mass and for the ratio of mass to light, a useful ratio for gauging the amount of dark matter in a cluster. For the Virgo Cluster, the mass-to-light ratio is found to be in the range of 500 in solar units, depending on how well smoothed the velocities have become, which is still uncertain. This value indicates that the cluster contains mostly unseen matter, as the average mass-to-light ratio for a population of normal stars is about 1. The total mass of the cluster is calculated to be 4×10^{14} times the mass of the Sun. *See* COSMOLOGY.

Distance. Determining the distance to the Virgo Cluster has been one of the more important tasks of twentieth-century astronomy. The Virgo Cluster forms an important stepping-stone for the cosmic distance scale, as it provides a connection between local distance criteria and the use of the Hubble law of gauging distances. The best measurements of the distance to the Virgo Cluster are based on the study of its cepheid variables with the *Hubble Space Telescope*. Other evidence comes from study of its globular star clusters, its planetary nebulae, its novae, and its supernovae. These give a mean distance of 16 megaparsecs (5×10^7 light-years). When the infall velocity of the Milky Way Galaxy toward the center of the Virgo Cluster is taken into account, this distance indicates that the local cosmic value of the Hubble constant, which relates expansion velocity to distance, is about 70 kilometers per second

Fig. 5. Sombrero Galaxy, in an image obtained with the *Spitzer Infrared Orbiting Telescope*. The spheroidal galaxy of stars hosts a thin, spiral-shaped ring of dust, visible in the infrared because of the light emission from the warm dust grains. (*NASA Spitzer image*)

per megaparsec. *See* CEPHEIDS; GALAXY, EXTERNAL; HUBBLE CONSTANT. Paul Hodge

Bibliography. A. Diaferio and K. van der Huch (eds.), *Outskirts of Galaxy Clusters*, Cambridge University Press, 2004; O.-G. Richter and B. Binggeli (eds.), *The Virgo Cluster*, European Southern Observatory, 1985; W. C. Saslaw, *The Distribution of Galaxies: Galaxy Clustering in Cosmology*, Cambridge University Press, 2000; W. Waller and P. Hodge, *Galaxies and the Cosmic Frontier*, Harvard University Press, 2003.

Virial equation

An equation of state of gases that has additional terms beyond that for an ideal gas which account for the interactions between the molecules. The pressure p can be expressed in terms of the molar volume $V_m = V/n$ (where n is the number of moles of gas molecules in a volume V), the absolute temperature T, and the universal gas constant $R = 8.3145$ J K^{-1} mol^{-1} or, in more commonly used practical units, 0.082058 L atm K^{-1} mol^{-1} [Eq. (1)]. In the equation

$$\frac{V_m}{RT} = 1 + \frac{B_2(T)}{V_m} + \frac{B_3(T)}{V_m^2} + \frac{B_4(T)}{V_m^3} + \cdots \qquad (1)$$

the virial coefficients $B_n(T)$ are functions only of the temperature and depend on the nature of the gas. In an ideal gas, in which all interactions between the molecules can be neglected because V_m is sufficiently large, only the first term, unity, survives on the right-hand side. *See* GAS.

The equation is important because there are rigorous relations between the coefficients B_2, B_3, and so on, as well as the interactions of the molecules in pairs, triplets, and so forth. It provides a valuable route to a knowledge of the intermolecular forces. Thus if the intermolecular energy of a pair of molecules at a separation r is $u(r)$, then the second virial coefficient can be expressed as Eq. (2). Here

$$B_2(T) = 2\pi N_A \int_0^\infty (1 - e^{-u(r)/kT}) r^2\, dr \qquad (2)$$

$N_A = 6.0221 \times 10^{23}$ mol^{-1} is Avogadro's constant, and $k = R/N_A$. In a gas mixture, B_n is a polynomial of order n in the mole fractions of the components. *See* INTERMOLECULAR FORCES.

The virial equation is useful in practice because it represents the pressure accurately at low and moderate gas densities, for example, up to about 4 mol L^{-1} for nitrogen at room temperature, which corresponds to a pressure of about 100 atm (10 MPa). It is not useful at very high densities, where the series may diverge, and is inapplicable to liquids. It can be rearranged to give the ratio pV_m/RT as an expansion in powers of the pressure instead of the density, which is equally useful empirically, but the coefficients of the pressure expansion are not usually called virial coefficients, and lack any simple relation to the intermolecular forces or, in a mixture, to the composition of the gas. *See* VAN DER WAALS EQUATION. J. S. Rowlinson

Bibliography. K. E. Bett, J. S. Rowlinson, and G. Saville, *Thermodynamics for Chemical Engineers*, MIT Press, Cambridge, MA, 1975; E. A. Mason and T. H. Spurling, *The Virial Equation of State*, Pergamon, Oxford, 1969.

Virial theorem

A theorem in classical mechanics which relates the kinetic energy of a system to the virial of Clausius, as defined below. The theorem can be generalized to quantum mechanics and has widespread application. It connects the average kinetic and potential energies for systems in which the potential is a power of the radius. Since the theorem involves integral quantities such as the total kinetic energy, rather than the kinetic energies of the individual particles that may be involved, it gives valuable information on the behavior of complex systems. For example, in statistical mechanics it is intimately connected to the equipartition theorem; in astrophysics it may be used to connect the internal temperature, mass, and radius of a star and to discuss stellar stability. The virial theorem makes possible a very easy derivation of the somewhat counterintuitive result that as a star radiates energy and contracts it heats up, rather than cooling down. *See* STAR; STATISTICAL MECHANICS; STELLAR EVOLUTION.

Derivation. The derivation of the virial theorem begins with the sum, over all the particles of the system, of the dot product of the position ($\mathbf{r}$) and momentum ($\mathbf{p}$). This quantity, labeled G, is given in Eq. (1),

$$G \equiv \sum_i \mathbf{p}_i \cdot \mathbf{r}_i \qquad (1)$$

where the index i runs over all the particles in the system. (The quantity G is half of the time derivative of the spherical moment of inertia.) The time derivative of the right-hand side of Eq. (1) involves two terms. Taking the time derivatives of the momenta, and noticing that these time derivatives are the forces, leads to the last term in Eq. (2). Taking the

$$\frac{dG}{dt} = 2T + \sum_i \mathbf{F}_i \cdot \mathbf{r}_i \qquad (2)$$

time derivatives of the positions to get the velocities, and noticing that the momenta times the velocities just gives twice the kinetic energy, T, leads to the other term on the right-hand side of Eq. (2). Thus, Eq. (2) is obtained. *See* CALCULUS OF VECTORS; ENERGY; FORCE; KINETICS (CLASSICAL MECHANICS); MOMENT OF INERTIA.

To complete the derivation, this result must be averaged over time. Since the left-hand side of Eq. (2) is the time derivative of G, its average is the difference between the values of G at the beginning and end of the interval divided by the interval. If the system is periodic, it can be averaged over a period. Then the difference of the G's vanishes, and so does

the average. If the system is confined, so that G is bounded, by which is meant that the velocities and position vectors remain finite, then the system can be averaged over a long time. The average of the left-hand side then approaches 0 as the time interval approaches infinity. In either case, the virial theorem (3) is obtained, in which the angle brackets indicate

$$2\langle T\rangle = -\sum_i \langle \boldsymbol{F}_i \cdot \boldsymbol{r}_i\rangle \tag{3}$$

averages. In Eq. (3), the right-hand side is twice the virial of Clausius.

Generalizations. Forces proportional to the velocity, such as those due to a uniform external magnetic field, have no effect on the validity of the virial theorem. Indeed, the force exerted on a particle by a magnetic field is the cross product of the field and the particle's velocity; and the sum, over all the particles in the system, of the dot product of this force and the particle's position is a total derivative and averages to zero if G is either periodic or bounded, as has been assumed. In quantum mechanics, the commutator with the hamiltonian is used in place of the time derivative to obtain the same result. A relativistic derivation is somewhat more difficult. It can be done by integrating the divergence of the energy-momentum tensor dotted into a coordinate. For electromagnetic or gravitational forces, it is found that the total energy, E, is given by Eq. (4), where m_i and

$$E = \sum_i m_i c^2 \left(\frac{1 - v_i^2}{c^2}\right)^{1/2} \tag{4}$$

v_i are the mass and velocity of the ith particle and c is the velocity of light. The correct nonrelativistic result is easily obtained from this equation. *See* NONRELATIVISTIC QUANTUM THEORY; RELATIVITY.

Elementary results. In the common case that the forces are derivable from a power-law potential, V, proportional to r^k, where k is a constant, the virial is just $-k/2$ times the potential energy. Thus, in this case the virial theorem simply states that the kinetic energy is $k/2$ times the potential energy. For a system connected by Hooke's-law springs, $k = 2$, and the average kinetic and potential energies are equal. For $k = -1$, that is, for gravitational or Coulomb forces, the potential energy is minus twice the kinetic energy. *See* COULOMB'S LAW; GRAVITATION; HARMONIC MOTION.

Stars. Stars have only small electric fields because the electrical forces keep the electrons and the positive ions close together. Hence the dominant source of the potential energy is gravity. Now, in the derivation of the virial theorem, the kinetic energy is just half the sum of products of the masses and velocity squares; that is, T involves the translational kinetic energy, but not the internal kinetic energy of the molecules. Thus, it is simply related to the temperature. Other degrees of freedom, rotation and vibration of molecules, for instance, make additional contributions to the internal energy. The total internal energy may be written as in Eq. (5), recalling the usual thermodynamic definition

$$U = \frac{2T}{3(\gamma - 1)} \tag{5}$$

of γ as the ratio of the specific heats at constant pressure and constant volume. The difference between the specific heats is the gas constant, R, so that γ gets smaller as the specific heats get larger. For a nonrelativistic monatomic gas or plasma, $\gamma = 5/3$. For a relativistic gas or for photons, $\gamma = 4/3$. Endothermic reactions such as the ionization of atoms or the breakup of nuclei, both of which occur as the temperature increases, reduce the value of γ. *See* KINETIC THEORY OF MATTER; SPECIFIC HEAT.

As the star loses energy, for instance by radiation, the total energy, E, decreases. But since the total energy is related to T by Eq. (6), T may increase.

$$E = V + U = \left[-2 + \frac{2}{3(\gamma - 1)}\right]T \tag{6}$$

For γ greater than 4/3, the gravitational potential energy, V, becomes more negative; the internal energy, U, becomes more positive; and the temperature increases.

Thermal disintegration of iron nuclei sets in at several billion kelvins and drives γ below 4/3. When γ falls below 4/3, since T is certainly positive, it appears that the total energy must also be positive. This would mean that energy had to be supplied to assemble the star, which is unreasonable, since nothing does work on interstellar material to assemble it into stars. There is a fundamental error in assuming that the star is nearly static, that is to say, the time derivative of G may be neglected in Eq. (2) (or, equivalently, the second time derivative of the spherical moment of inertia may be neglected). While that is an excellent approximate for ordinary stars whose configuration evolves very slowly, it must be concluded that stars with less than 4/3 are unstable, and lose pressure support and collapse precipitously.

In real stars, γ approaches 4/3 only when the electrons become relativistic, so that the connection between kinetic energy and velocity is different from what has been used here. A more rigorous analysis, however, leads to the same conclusion. When γ falls below 4/3 in the stellar core, the core collapses and a supernova results. *See* GRAVITATIONAL COLLAPSE; SUPERNOVA.

Albert G. Petschek

Bibliography. D. D. Clayton, *Principles of Stellar Evolution and Nucleosynthesis*, 1983; G. W. Collins, II, *The Virial Theorem in Stellar Astrophysics*, 1978; H. Goldstein, *Classical Mechanics*, 3d ed., 2001; R. K. Pathria, *Statistical Mechanics*, 2d ed., 1996.

Viroids

Small infectious nucleic acid particles containing independently replicating single-stranded circular ribonucleic acids (RNAs) of 246–399 nucleotides, which are known to infect certain monocot and dicot plants. Viroids are the smallest known disease agents, with an estimated molecular weight (sum

of the atomic weights of a molecule's constituent atoms) of as little as $1.0–1.3 \times 10^4$, in marked contrast to the conventional plant virus genomes, which have molecular weights of approximately 2×10^6. Conventional viruses are made up of nucleic acid encapsulated in protein (capsid), whereas viroids are uniquely characterized by the absence of a capsid. In spite of their small size, viroid RNAs can replicate and produce characteristic disease syndromes when introduced into cells. Although the viroids thus far identified are associated with plants, based on the unusual properties of the infectious agents they may also be found to affect other forms of life, including humans. It is possible that in instances where a viral etiology of disease has been assumed but where no causal virus has been identified, viroid infectious agents are involved.

Classification. Up to 27 different viroid species have been described from widely separated geographical locations. Viroids are classified in two families. The first family, Pospiviroidae, contains 24 viroids which have a central conserved nucleotide region which contain the same nucleotide sequences. Some examples include potato spindle tuber viroid (PSTV), citrus exocortis viroid (CEV), chrysanthemum stunt viroid (CSV), and cucumber pale fruit viroid (CPFV). The second family, Avsunviroidae, contains three viroids, which do not have a central conserved region: avocado sunblotch viroid (ASV), peach latent mosaic viroid (PLMV), and chrysanthemum chlorotic mottle viroid (ChCMV).

Diseases and hosts. Diseases caused by viroids are agriculturally relevant, affecting economically important plants and an assortment of herbaceous and woody plants, such as potato, tomato, cucumber, hop, citrus, grapevine, coconut palm, fruit trees (apple, pear, plum, peach, avocado), and ornamentals (chrysanthemum and coleus).

Viroid infections in some plant species produce profound disease symptoms ranging from stunting and leaf epinasty (downward bending) to plant death, whereas infections in other species produce few detectable symptoms compared with uninoculated control plants. Viroids generally have a restricted host range, although several viroids can infect the same hosts and cause similar symptoms. For example, PSTV, CEV, CSV, and CPFV infect Rutgers tomato (*Lycopersicon esculentum*), and all except CSV cause stunting of growth, epinasty, and bunchy top. Infection with PSTV, CEV, and CSV produces similar symptoms in purple passion plant (*Gynura aurantiaca*).

Good controls are not available for diseases caused by these small infectious agents other than indexing procedures to provide viroid-free propagules. Indexing programs generally rely on the use of bioassays that are laborious and require incubation periods of several months. Sensitive biochemical assays [such as polyacrylamide gel electrophoresis, dot-blot hybridization with labeled complementary deoxyribonucleic acid (cDNA) or complementary RNA (cRNA) probes, and reverse transcription polymerase chain reaction] are faster and more reliable to use than conventional bioassays and have been developed to detect infectious viroid RNAs in infected host tissue.

Cross-protection. Although viroids have been shown to differ significantly in their nucleotide sequence (CCR of Pospirviroidae is the only portion of the viroid molecule which has conserved nuclear sequences), the phenomenon of cross-protection is known to occur in some viroid double infections. Cross-protection is defined here as the interference in symptom expression of a challenge-inoculated viroid or virus in a previously infected plant. Cross-protection has been shown to exist in viroid species of both Pospiviroidae and Avsunviroidae families and is known to be due to subtle variations in the structure of their RNA. As a result of these phenomena, plants infected with a latent or nonsymptomatic strain become protected against the challenge inoculation by a severe strain. The practical applications of cross-protection have been to indicate relationships among plant viruses and to control tobacco mosaic virus in commercial tomato plantings and tristeza of citrus in Brazil.

Cross-protection has been shown between strains of PSTV and also between strains of ChCMV. Moreover, cross-protection has been shown between strains of the closely related viroids PSTV, CEV and CSV. However, a lack of cross-protection has been shown for ChCMV strains against PSTV, CEV and CSV strains, which separates the former from the latter three viroid species (two different viroid families).

Interference with other pathogens. Interference has also been found between viroids and bacteria in double infections in chrysanthemums. The amount of pith maceration of chrysanthemum cuttings, cultivar Bonnie Jean, by *Erwinia chrysanthemi* 159 is significantly reduced by prior CSV infection. The reduction is apparent in CSV-infected plants approximately 30 days after inoculation with the viroid. This period of time coincides with CSV symptom expression and represents a period of viroid multiplication. It appears that a threshold of viroid molecules within infected tissue is necessary for three observed events: symptom development, viroid detection by polyacrylamide gel electrophoresis assay, and decreased bacterial pith maceration. This reaction differs from the cross-protection phenomenon, which involves closely related organisms (that is, strains of a single virus and viroid). Furthermore, reduced infection has been found with *Phytophthora infestans* (which causes potato late blight) in potato plants infected with PSTV. Viruses, however, have been reported to reduce or increase the severity of fungal infections. *See* PLANT VIRUSES AND VIROIDS; VIRUS.

R. K. Horst

Bibliography. T. O. Diener (ed.), *The Viroids*, 1987; T. O. Diener, *Virus Genes*, 11:119–131, 1996; R. Flores et al. (eds.), *Advances in Virus Research*, vol. 55, 2001; T. C. Hall and J. W. Davies (eds.), *Nucleic Acids in Plants*, vol. 2, 1979; J. S. Semancik et al. (eds.), *Viroids and Viroid-like Pathogens*, 1988.

Virtual acoustics

The simulation of the complex acoustic field experienced by a listener within an environment. The technology is also known as three-dimensional sound and auralization. Going beyond the simple left-right volume adjustment of normal stereo techniques, the goal is to process sounds so that they appear to come from particular locations in three-dimensional space. Although loudspeaker systems have been developed, much of the work in the field focuses on using headphones for playback and is the outgrowth of earlier analog techniques. For example, in binaural recording, the sound of an orchestra playing classical music is recorded through small microphones in the two imitation ear canals of an anthropomorphic artificial or dummy head placed in the audience of a concert hall. When the recorded piece is played back over headphones, the listener passively experiences the illusion of hearing the violins on the left and the cellos on the right, along with all the associated echoes, resonances, and ambience of the original environment. Techniques use digital signal processing to synthesize the acoustical properties that people use to localize a sound source in space. Thus, they provide the flexibility of a kind of digital dummy head, allowing a more active experience in which a listener can both design and move around or interact with a simulated acoustic environment in real time. *See* BINAURAL SOUND SYSTEM.

Psychoacoustical cues. The success of virtual acoustics is critically dependent on whether the acoustical cues used by humans to locate sounds have been adequately synthesized. Much of the present understanding of sound localization is based on Lord Rayleigh's duplex theory of 1907, which describes the role of two primary cues: interaural time differences for low-frequency signals, and interaural intensity differences for high-frequency signals. However, the duplex theory neglects the important interaction between incoming sound waves and the outer ears, or pinnae: spectral shaping by the pinnae is highly direction dependent; the absence of pinna cues degrades localization accuracy; and pinna cues are at least partially responsible for externalization, or the outside-the-head sensation. *See* EAR (VERTEBRATE).

There may be many cumulative effects on the sound as it makes its way to the eardrum, but all of these effects can be expressed as a single filtering operation much like the effects of a graphic equalizer in a stereo system. The exact nature of this filter can be measured by an experiment in which an impulse (a single, very short sound pulse or click) is produced by a loudspeaker at a particular location. The acoustic shaping by the two ears is then measured by recording the outputs of small probe microphones placed inside the ear canals of an individual or an artificial head (**Fig. 1**). If the measurement of the two ears occurs simultaneously, the responses, when taken together as a pair of filters, include an estimate of the interaural differences as well. Thus, this technique makes it possible to measure all of the relevant spatial cues together for a given source location, for a given listener, and in a given room or environment (**Fig. 2**). *See* ELECTRIC FILTER; EQUALIZER.

The spatial filters constructed from these ear-dependent characteristics are examples of finite impulse response filters, and the characteristics themselves are often referred to as head-related transfer functions (HRTFs). Filtering in the frequency domain is a point-by-point multiplication of the input signal with the left and right head-related transfer

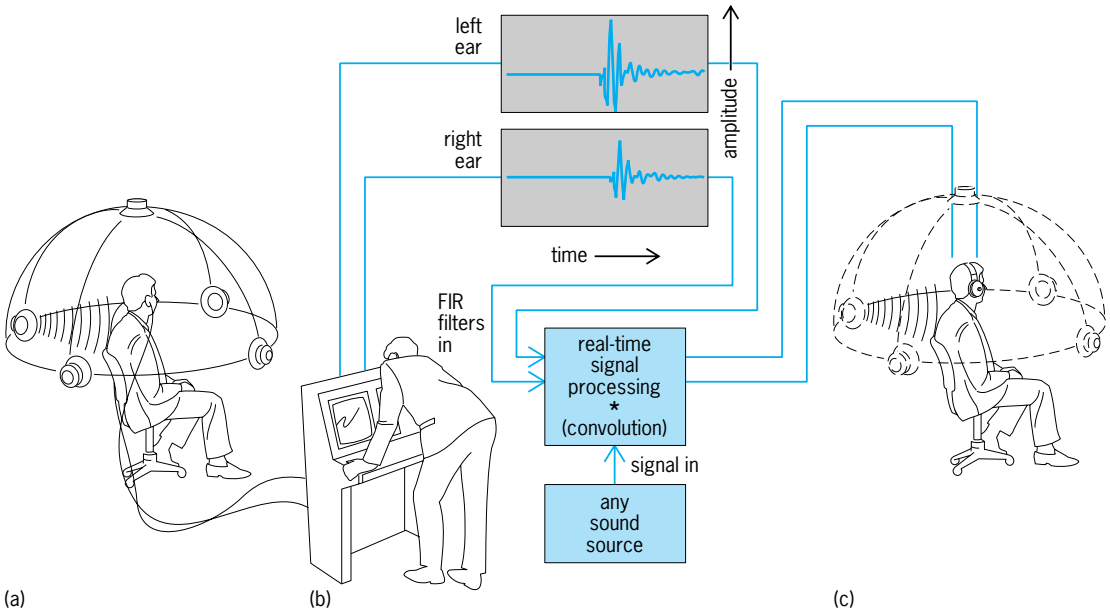

Fig. 1. Technique for synthesizing virtual acoustic sources with measurements of the head-related transfer function. (*a*) Pinnae (outer-ear) responses are measured with probe microphones. Placement of loudspeakers is illustrative only. (*b*) Pinnae transforms are digitized as finite-impulse-response (FIR) filters. An example of a pair of finite impulse responses measured for a source location at 90° to the left and 0° elevation (at ear level) is shown for the left and right ears. (*c*) Synthesized cues are played over headphones. (*After E. M. Wenzel, Localization in virtual acoustic displays, Presence: Teleoper. Virt. Environ., 1(1):80–107, 1992*)

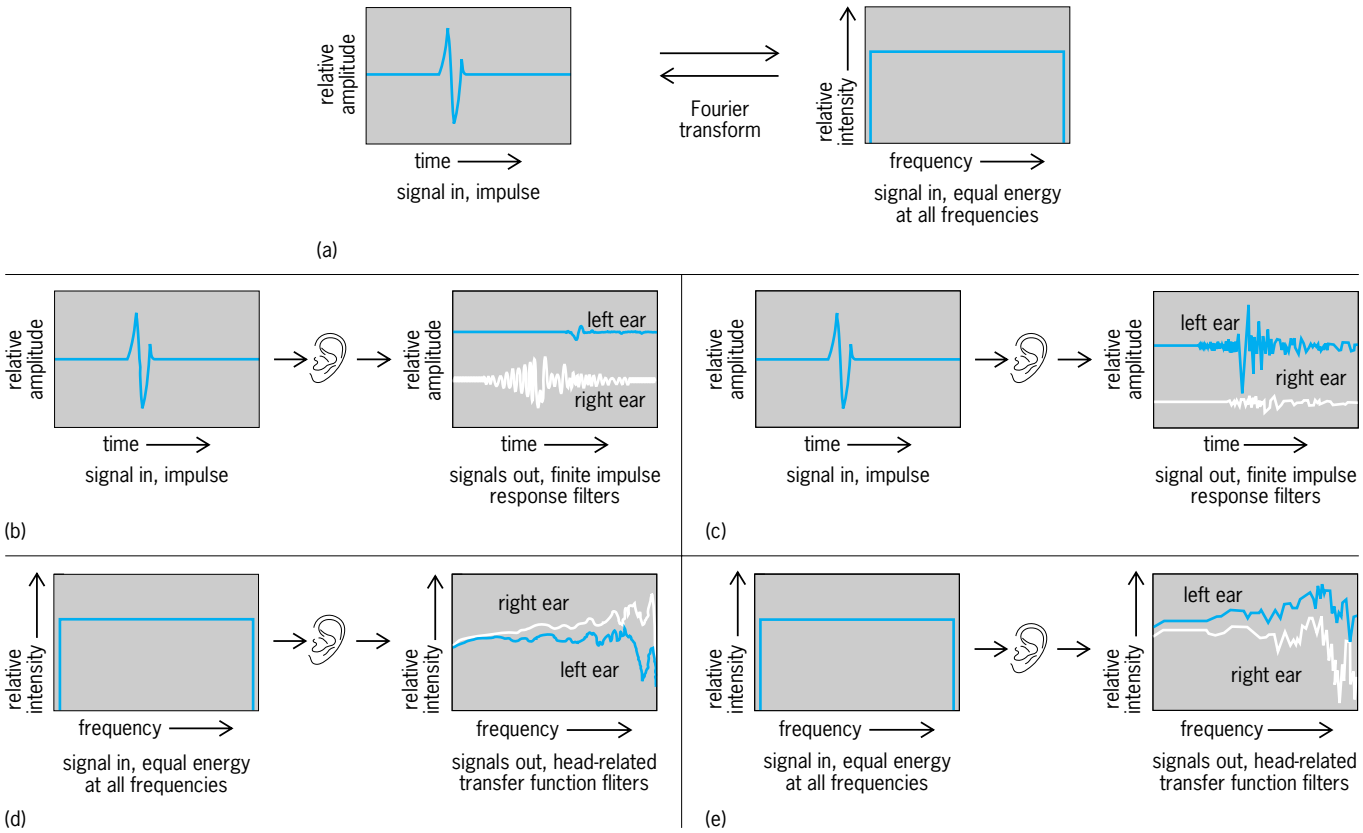

Fig. 2. Effects of spectral shaping by the pinnae (outer-ear structures). (*a*) Equivalent representations, via the Fourier transform, of a broad-band acoustic signal before interaction with the outer-ear structures, in the time domain (where the acoustic signal is an impulse) and the frequency domain (where only intensity is shown here). (*b*) Effect of the pinnae on an impulse delivered from a loudspeaker directly to the right (+90° azimuth, 0° elevation), as measured in the left-ear and right-ear canals of the individual. (*c*) Effect of the pinnae on an impulse delivered from a loudspeaker 60° to the left (−60° azimuth, 0° elevation). (*d*) Frequency-domain representation of the effect on the signal from 90° azimuth, 0° elevation. (*e*) Frequency-domain representation of the effect on the signal from −60° azimuth, 0° elevation. (*After E. M. Wenzel, Localization in virtual acoustic displays, Presence: Teleoper. Virt. Environ., 1(1):80–107, 1992*)

functions, while filtering in the time domain by the finite impulse response filters occurs via a more complex multiply-and-add operation known as convolution. *See* CONTROL SYSTEMS; FOURIER SERIES AND TRANSFORMS; INTEGRAL TRANSFORM.

By using these spatial filters, it should be possible to impose spatial characteristics on a signal so that the signal apparently emanates from the originally measured location. Of course, the localizability of the sound will also depend on other factors such as its original spectral content. Narrowband sounds (such as sine waves) are generally hard to localize, while broadband, impulsive sounds are the easiest to locate. Other acoustic features, such as the ratio of direct to reflected energy in a reverberant field, can also provide a cue to distance (closer sources correspond to larger ratios) as well as enhance the sensation of externalization. *See* REVERBERATION.

Psychophysical validation. The only conclusive test of the simulation technique is an operational one in which the localization of real and synthesized sources is directly compared. Two kinds of errors are usually observed when subjects are asked to judge the position of a stationary sound source in the free field: a relatively small error in resolution on the order of about 5–20°, and front-back reversals, in which a source in the front (or rear) is perceived by the listener as if it were in the rear (or front). Occasionally, reversals in elevation are also observed.

In studies that compare localization judgments of stationary broadband sources in the free field with judgments of virtual (headphone-presented) sources synthesized from the subjects' own head-related transfer functions, localization accuracy is generally similar for the free-field and headphone stimuli, although source elevation is less well defined and reversal rates increase from about 5% for real sources to about 10% for virtual sources. Other experiments suggest that most listeners can obtain useful directional information from nonpersonalized head-related transfer functions, particularly for the left-right dimension of azimuth, so long as these functions come from a "good localizer," who accurately localizes both real and virtual sources. Again, localization is comparable for free-field and headphone stimuli, although, for some subjects, source elevation is completely disrupted and reversal rates increase substantially. Interestingly, some people show little ability to localize source elevation for both real and virtual sounds, as if they have poorly designed pinna structures which have little impact on the spectra of incoming sounds. Listening "through" these

people's transfer functions results in a loss of elevation perception for any listener.

Real-time implementation. Real-time systems began to appear only when powerful digital signal processing chips were available. In general, these systems are intended for headphone presentation and use time-domain convolution to achieve real-time performance. Usually, one or more sources are simulated for an anechoic environment by using spatial filters chosen according to the output of a head-tracking device in order to compensate for the listener's position and orientation. Smooth motion trajectories can be simulated by interpolation between impulse responses originally measured at intervals of about 10° or more, and a simple distance cue may be provided via real-time scaling of amplitude. While such systems simulate only the direct paths from each virtual source to the listener, it is possible to freely manipulate both source and listener position in a highly dynamic, interactive display. *See* INTEGRATED CIRCUITS.

The same degree of interactivity is desirable in a more complex simulation since perceptual research suggests that errors like front-back reversals and failures of externalization can be mitigated by providing the acoustic cues available in reverberant environments. Of particular interest is the development of the image model and related ray-tracing techniques for simulating room characteristics. Extensions of these techniques implement the directional characteristics of source reflections in real time by convolution with filters based on head-related transfer functions. For example, in one implementation of the image model the walls, floor, and ceiling in an environment are simulated by calculating the mirror image of a sound source behind each surface to account for the specular reflection of the source signal. The filtering effect of surfaces such as wood or drapery can also be modeled with a separate filter whose output is delayed by the time required for the sound to propagate from each image source being represented. Finally, the delayed signals are processed with spatial filters appropriate to the direction of each reflection to create the components of the binaural output. The superposition of these components with the direct path to the source is then presented over headphones. Such dynamic modeling requires enormous computational resources for real-time implementation in a truly interactive (head-tracked) display. *See* ARCHITECTURAL ACOUSTICS; HEARING (HUMAN); PSYCHOACOUSTICS; REFLECTION OF SOUND; SOUND; TELECONFERENCING; VIRTUAL REALITY.

Elizabeth M. Wenzel

Bibliography. J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*, 1983; F. L. Wightman and D. J. Kistler, Headphone simulation of free-field listening, I: Stimulus synthesis, *J. Acous. Soc. Amer.*, 85(2):858–867, 1989; F. L. Wightman and D. J. Kistler, Resolution of front-back ambiguity in spatial hearing by listener and source movement, *J. Acous. Soc. Amer.*, 105(5):2841–2583, 1999.

Virtual manufacturing

The modeling of manufacturing systems using audiovisual or other sensory features to simulate or design alternatives for an actual manufacturing environment, or the prototyping and manufacture of a proposed product using computers. The motivation for virtual manufacturing is to enhance people's ability to predict potential problems and inefficiencies in product functionality and manufacturability before real manufacturing occurs.

The term "virtual manufacturing" became prominent in the early 1990s, in part because of the U.S. Department of Defense Virtual Manufacturing Initiative. The concept has since gained international acceptance and has broadened in scope. For the first half of the 1990s, pioneering work in this field was done by a handful of major organizations, mainly in the aerospace, earthmoving equipment, and automobile industries, plus a few specialized academic research groups. Factors accelerating the worldwide market interest in virtual manufacturing include price-performance improvements in the hardware and software technologies required, and increased awareness of its huge potential. *See* MANUFACTURING ENGINEERING; MODEL THEORY.

Framework. The concepts underlying virtual manufacturing include virtual reality, high-speed networking and software interfaces, agile manufacturing, and rapid prototyping.

Virtual reality. Virtual reality is broadly defined as the ability to create and interact in cyberspace, that is, a simulated space that represents an environment very similar to the actual environment. The subset of virtual reality that is used in virtual manufacturing is commonly known as virtual environment. Virtual environment systems differ from previously developed computer-centered systems in the extent to which real-time interaction is simulated. The perceived visual space is three-dimensional rather than two-dimensional, the human-machine interface is multimodal, and the user is immersed in the computer-generated environment; the screen separating the user and the computer becomes invisible to the user.

The virtual environment for virtual manufacturing is simulated through immersion in computer graphics coupled with an acoustic interface, domain-independent interacting devices such as wands, and domain-specific devices such as steering and brakes for cars or earthmovers or instrument clusters for airplanes. There are two main techniques for creating immersion in current virtual reality systems: head-mounted displays and stereoscopic projectors. A head-mounted display is a tracking device incorporating liquid-crystal displays or miniature cathode-ray tubes which is mounted on the head of the user to provide information on head movements for updating the visual images. An example is the Binocular Omni-Oriented Monitor (BOOM) developed commercially. Stereoscopic projectors provide a three-dimensional sensation using screen

projections of two distinct views of an object separated by a small angle based on viewing the object with the left eye only and with the right eye only. The users normally wear liquid-crystal-display shutter glasses. An example is the Cave Automatic Virtual Environment (CAVE) developed by the University of Illinois at Chicago. *See* VIRTUAL REALITY.

High-speed networking and software interfaces. The issues here are concerned with computer-aided-design (CAD) model portability among systems, trade-offs of high-detail models versus real-time interaction and display, rapid prototyping, collaborative design using virtual reality over distance, use of the Web for small- or medium-business virtual manufacturing, use of qualitative information (illumination, sound levels, ease of supervision, handicap accessibility) to design manufacturing systems, use of intelligent and autonomous agents in virtual environments, and the validity of virtual reality versus reality (quantitative testing of virtual versus real assemblies or equipment). With regard to hardware, the two extremes (in terms of cost and performance) are personal computers using Virtual Reality Modeling Language (VRML) and the Internet, and workstations using the CAVE library and asynchronous transfer mode (ATM) networks with data transfer speeds up to 622 Mbps (see **illus.**).

Demonstration of Virtual Reality Modeling Language (VRML) Browser interface with CAVE™ for 3D Immersive Analysis of Manufacturing Shopfloor. The CAVE™ virtual reality display features three room-sized rear-projected walls and a top-projected floor, each driven by workstations. (CAVE is a trademark of the University of Illinois.)

Agile manufacturing. A term that is sometimes mentioned in the context of virtual manufacturing is agile manufacturing. Agile manufacturing integrates an organization's people and technologies through innovative management and organization, knowledgeable and empowered people, and flexible and intelligent technologies. Virtual manufacturing provides a model for making rapid changes in products and processes based on customer requirements, and an agile manufacturing system attempts to implement it.

Rapid prototyping. An area in which virtual manufacturing has made an impact is in rapid prototyping, or layered manufacturing. In processes such as stereolithography, selective laser sintering, and fused deposition modeling, a CAD drawing of a part is processed to create a layered file of the part. The part is built one layer at a time, precisely depositing layer upon layer of material.

Virtual collaborative environments. Global virtual manufacturing extends the definition of virtual manufacturing to include the use of the Internet and intranets (global communications networks) for virtual component sourcing, and the use of virtual collaborative design and testing environments by multiple organizations or sites. Companies committed to global virtual manufacturing may shorten the time to market for new products, cut the cost of prototyping and preproduction engineering, enable many more variations to be tried out before committing to manufacture, and increase the range and effectiveness of quality assurance testing. Prototypes can be virtually assembled, tested, and inspected as part of production planning and operative training procedures. They can be demonstrated, market-tested, used to brief and train sales and customer staff, transmitted instantly from site to site via communications links, and modified and recycled rapidly in response to feedback.

Manufacturers and their worldwide subcontractors and main suppliers can establish agile manufacturing teams that will work together on the design, virtual prototyping, and simulated assembly of a particular product, at the same time establishing confidence in the virtual supply chain. Using the most advanced virtual reality systems, geographically remote members of the team can meet in the same virtual design environment to discuss and implement changes to virtual prototypes.

An example of a recent development in virtual collaborative environments is the projection of gestures and movements of multiple remote designers through voice-activated avatars to help explain the designers' intentions to others in real-time using high-speed ATM networks.

Monitoring and control for complex manufacturing. For monitoring and control of complex manufacturing systems, four dimensions can be conceived to express complexity: space, time, process, and network. Space concerns the physical location, layout, and flow issues critical in all manufacturing operations. Time involves the facility life cycle and operational dynamics beginning with the concurrent engineering of the production process and testing facilities during product design, extending through the maturation and decline of the first-generation products, and cycling through the same process for future generation products. Process relates to the coherent integration of engineering, management, and manufacturing processes. It permits the examination of the important yet intricate interplay of relationships between classically isolated functions,

such as the relationships between production planning and purchasing, production control and marketing, quality and maintenance, and design and manufacturing. Process decisions range from long-range operational planning to short-term machine- and device-level planning and control. The integration between various levels of aggregation is essential. Network concerns organization and infrastructure integration. While the third dimension, process, focuses on the actions, this dimension concentrates on the actors and their needs and responsibilities. Clearly including personnel, the set of actors also includes all devices, equipment, and workstations; all organizational units, be they cells, teams, departments, or factories; and all external interactors, such as customers, vendors, subcontractors, and partners. Issues such as contrasting hierarchically controlled networks with hierarchical, autonomous agent networks must be addressed. *See* CONCURRENT PROCESSING.

Virtual manufacturing techniques address issues such as designing products that can be evaluated and tested for structural properties, ergonomic functionality, and reliability without having to build actual scale models; designing products for esthetic value meeting individual customer preference; ensuring facility and equipment compliance with various federally mandated standards; facilitating remote operation and control of equipment (telemanufacturing and telerobotics); developing processes to ensure manufacturability without having to manufacture the product (for example, avoiding destructive testing); developing production plans and schedules and simulating their correctness; and educating employees on advanced manufacturing techniques with emphasis on safety.

Applications. Virtual manufacturing has been used in the automotive industry to design the interior of cars and perform virtual crash testing. The heavy-equipment industry has built electronic prototypes of earthmoving equipment instead of physical prototypes. Both industries are actively exploring virtual collaborative environments using multiple CAVEs to coordinate designs with their suppliers and showrooms. The pharmaceutical industry is exploring use of this technology for drug manufacturing to rapidly conform to the Federal Drug Administration (FDA) requirements. The electronics industry is comparing the impact of employees trained with head-mounted devices in virtual reality environments with those trained in the real environment. The telecommunications industry is using haptic interfaces, which allow users to feel their interfaces and interactions, to design their phones. Pat Banerjee

Bibliography. P. Banerjee, L. Freitag, and S. Mehrotra (eds.), *Virtual Reality in Manufacturing Research and Education: Proceedings of the Symposium Sponsored by the National Science Foundation*, 1996; A. Johnson et al., CAVERN: The CAVE Research Network, *Proceedings of the 1st International Symposium on Multimedia Virtual Laboratory*, Tokyo, March 25, 1998; Special Issue on Virtual Manufacturing, *IIE Trans.*, 30(7):581–644, 1998.

Virtual reality

A form of human-computer interaction in which a real or imaginary environment is simulated and users interact with and manipulate that world. Users travel within the simulated world by moving toward where they want to be, and interact with things in the world by grasping and manipulating simulated objects. In the most successful virtual environments, users feel that they are truly present in the simulated world and that their experience in the virtual world matches what they would experience in the environment being simulated. This sensation is referred to as engagement, immersion, or presence, and it is this quality that distinguishes virtual reality from other forms of human-computer interaction. *See* HUMAN-COMPUTER INTERACTION.

When a user interacts with a virtual environment, the computer-generated graphics display must be updated with each turn of the head or movement of the hand. The virtual environment must be able to generate and display realistic-looking views of the simulated world quickly enough that the interaction feels responsive and natural. Initially, use of the technology was limited to those applications where the cost of high-performance computer graphics hardware could be justified, such as in flight simulators. The decreasing cost of such hardware has made it possible to experiment with applying the technology in many other areas. *See* COMPUTER GRAPHICS.

Hardware. Virtual reality relies on a variety of specialized input and output devices to achieve a sense of natural interaction in a simulated environment. The goal is to go beyond traditional computer interfaces and mimic human experience with the real world. For example, in traditional computer-generated graphics, only one view of a scene is projected, rendered, and drawn on the display screen. But when a person views the real world, the left and right eyes see different views and the perceptual system uses these differences to judge the positions of objects in the scene. Also, human interaction with the world is quite rich; people move around, reach for things, pick them up, and manipulate objects directly. All the senses are involved. But in traditional computer applications, user interaction modes are limited to typing on a keyboard, moving a mouse around on a flat surface, or perhaps using a hand-held joystick. A primary goal of virtual reality is to develop strategies for creating a natural interaction environment that will contribute to a sense of presence in the simulated world. *See* STEREOSCOPY.

Tracker. The most important of the input devices used in a virtual environment, a tracker is capable of reporting its location in space and its orientation. Tracking devices can be optical, magnetic, or acoustic. Optical systems use cameras to record user activity. The images are interpreted to detect body position or hand gestures. Camera-based environments allow the user to direct system activity in a very natural and unencumbered manner, but there are a number of difficult problems with this approach. For

example, it is difficult to interpret images of hand activity when users move their hands quickly or when fingers overlap. *See* COMPUTER VISION.

Acoustic and magnetic tracking systems use a stationary element and small movable elements. The movable element is usually worn on the user's head or hand and is connected by cable to the computer system. In acoustic systems, the movable tracker element emits a high-frequency sound, which is detected by a set of stationary microphones. The strength of the signal recorded by each microphone will depend on the distance between the sound source and a particular microphone, and the variation in signal strength indicates the location of the sound. Acoustic systems are accurate to within a fraction of an inch (a few millimeters). In magnetic tracking systems, the stationary element emits a pulsed magnetic field. The movable sensor attached to the user reports its position and orientation relative to the source. Magnetic systems are accurate to 0.1 in. (2.5 mm) in position and 0.1° in rotation. Optical and acoustic systems rely on a clear line-of-sight path between the signal source and the receptor. Magnetic tracking systems do not have this requirement, but magnetic tracking can be susceptible to interference from metallic objects in the environment. *See* MICROPHONE.

A tracker is sometimes combined with a traditional computer input device, such as a mouse or a joystick. The user holds the device and can use the regular controls, such as mouse buttons, to signify when different actions should occur. Everyday objects, including treadmills and stationary bicycles, have also been adapted to serve as input devices to virtual reality systems. *See* COMPUTER PERIPHERAL DEVICES.

Data glove. An attempt to provide a truly natural input device, the data glove is outfitted with sensors that can read the angle of each of the finger joints in the hand. Wearing such a glove, users can interact with the virtual world through hand gestures, such as pointing or making a fist. Gages that measure strain or fiber optics, where the amount of transmitted light signifies the angle at a finger joint, are used in gloves to determine finger positions. *See* FIBER-OPTIC SENSOR; STRAIN GAGE.

Stereoscopic displays. The real-world visual experience is approximated in virtual environments by using stereoscopic displays. Two views of the simulated world are generated, one for each eye, and a stereoscopic display device is used to show the correct view to each eye. Two types of stereo viewing devices are the head-mounted display and the binocular omni-orientation monitor (BOOM). Each device has two display screens, and the appropriate image is shown to the left and right eye.

Stereo glasses, lightweight and inexpensive, provide another alternative. High-quality stereo glasses usually use liquid-crystal shutters that allow light to pass through only one lens at a time. Computer-generated views for the left and right eyes are alternated very quickly on a computer monitor, or perhaps on a large projection screen. A transmitter connected to the computer synchronizes the glasses to the display so that the appropriate lens is active at each instant. Switching between views is fast enough that the user sees the overall scene in stereo and is unaware of the alternating views. *See* LIQUID CRYSTALS.

The different display devices have advantages and disadvantages. The BOOM user does not need to wear any special equipment, which allows the user to look away from the display at any time. It also makes it easy for multiple users to take turns at the BOOM. However, the ability to step out of the virtual world at any moment can decrease the feeling of presence or immersion. Head-mounted displays, however, are helmetlike devices that are strapped on and require some effort to remove. Lightweight stereo glasses are easy to wear and to remove. Compared to the other devices, glasses are inexpensive enough that it is reasonable to provide a pair for everyone in a group.

Applications. Virtual reality can be applied in a variety of ways, including entertainment, education and training, manufacturing and design, and exploring scientific data and processes. In scientific and engineering research, virtual environments are used to visually explore whatever physical world phenomenon is under study. For example, a researcher can study how air travels near the surface of a wing. Virtual environments have also been used to explore whether different molecules will bond, with a force-feedback device used to convey whether molecules are attracted or repulsed. These virtual environments simulate physical phenomena occurring in the everyday world, but virtual environments are also used to explore worlds that can only be imagined, such as four-dimensional geometries, or microworlds that operate according to their own set of physical laws.

Training personnel for work in dangerous environments or with expensive equipment is best done through simulation. Airplane pilots, for example, train in flight simulators. Virtual reality can enable medical personnel to practice new surgical procedures on simulated individuals. As a form of entertainment, virtual reality is a highly engaging way to experience imaginary worlds and to play games. *See* AIRCRAFT TESTING; VIDEO GAMES.

Virtual reality also provides a way to experiment with prototype designs for new products. For example, the traditional design process for a new car involves building physical prototypes to evaluate the design. Virtual reality provides a way to simulate the new design and experiment with it, avoiding the cost of building the physical prototype. *See* COMPUTER-AIDED DESIGN AND MANUFACTURING.

M. Pauline Baker

Bibliography. W. Barfield and T. Furness (eds.), *Virtual Environments and Advanced Interface Design*, 1995; R. Earnshaw, M. Gigante, and J. Jones (eds.), *Virtual Reality Systems*, 1993; H. Rheingold, *Virtual Reality*, 1991; Special issue on virtual reality, *IEEE Comput. Graphics Appl.*, vol. 14, no. 1, January 1994; A. Wexelblat, *Virtual Reality Applications and Explorations*, 1993.

Virtual work principle

The principle stating that the total virtual work done by all the forces acting on a system in static equilibrium is zero for a set of infinitesimal virtual displacements from equilibrium. The infinitesimal displacements are called virtual because they need not be obtained by a displacement that actually occurs in the system. The virtual work is the work done by the virtual displacements, which can be arbitrary, provided they are consistent with the constraints of the system. *See* CONSTRAINT.

The principle of virtual work is equivalent to the conditions for static equilibrium of a rigid body expressed in terms of the total forces and torques. That is, the principle of virtual work can be derived from these conditions, and conversely. *See* EQUILIBRIUM OF FORCES; STATICS.

One advantage of the principle of virtual work is that it can serve as a basis for all of statics. In the solution of problems the principle of virtual work is often useful for eliminating the need for consideration of the forces of constraint, since these forces often are perpendicular to the virtual displacements and consequently do no work.

The following problem provides an application of the principle of virtual work (see **illus.**). A uniform plank of weight W rests on a smooth floor and leans against a smooth wall. The angle between the plank and the floor is ϕ. A weightless, inextensible string connects the lower end of the plank to the wall. The value of the tension T in the string is desired. Let $2b$ be the length of the plank. Then, as can be seen from the sketch, Eqs. (1) hold. When ϕ increases by

$$y = b \sin \phi \qquad x = 2b \cos \phi \tag{1}$$

$d\phi$, the resulting changes dx and dy in x and y, respectively, are given by Eqs. (2). According to the

$$dy = b \cos \phi \, d\phi \qquad dx = -2b \sin \phi \, d\phi \tag{2}$$

principle of virtual work, the total work done in the virtual displacement produced by $d\phi$ must be zero. The forces of reaction N_1 and N_2 are normal to the wall, since the wall is smooth. Consequently they do no work. The total work, therefore, is the work done by the forces W and T moving through their respective virtual displacements dy and dx. This work is given by Eq. (3). Solving for the tension T gives Eq. (4). The same result is obtained from the equilibrium conditions expressed in terms of forces and torques. When the principle of virtual work is used, the forces of reaction need not be considered.

$$0 = -2bT \sin \phi \, d\phi + Wb \cos \phi \, d\phi \tag{3}$$

$$T = (W/2) \cot \phi \tag{4}$$

Paul W. Schmidt

Bibliography. N. G. Chetave, *Theoretical Mechanics*, 1990; H. Goldstein, *Classical Mechanics*, 3d ed., 2001.

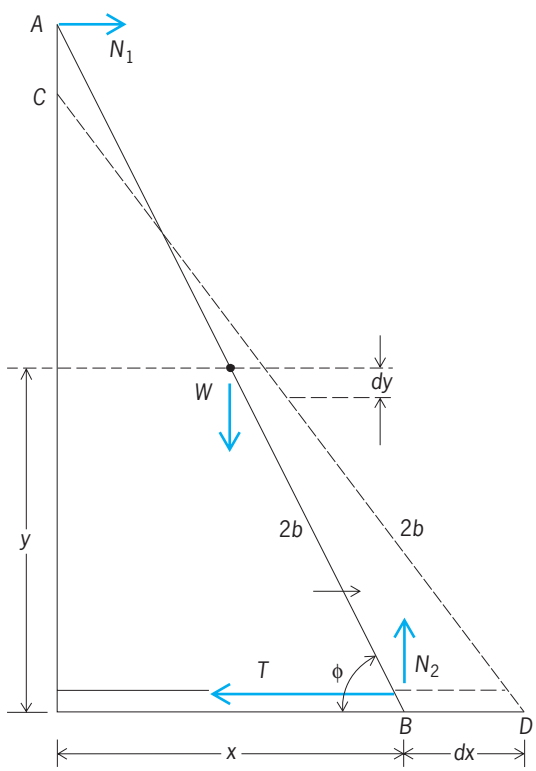

Illustration of principle of virtual work. (*After R. A. Becker, Introduction to Theoretical Mechanics, McGraw-Hill, 1954*)

Virulence

The ability of a microorganism to cause disease. Virulence and pathogenicity are often used interchangeably, but virulence may also be used to indicate the degree of pathogenicity. Infectious diseases are caused by bacteria, fungi, protozoa, viruses, and larger parasites. Scientific understanding of the underlying mechanisms of virulence has increased rapidly due to the application of the techniques of biochemistry, genetics, molecular biology, and immunology. Bacterial virulence is better understood than that of other infectious agents.

Virulence is often multifactorial, involving a complex interplay between the parasite and the host. Various host factors, including age, sex, nutritional status, genetic constitution, and the status of the immune system, affect the outcome of the parasite-host interaction. Hosts with depressed immune systems, such as transplant and cancer patients, are susceptible to microorganisms not normally pathogenic in healthy hosts. Such microorganisms are referred to as opportunistic pathogens. The attribute of virulence is present in only a small portion of the total population of microorganisms, most of which are harmless or even beneficial to humans and other animals. *See* OPPORTUNISTIC INFECTIONS.

The spread of an infectious disease usually involves the adherence of the invading pathogen to a body surface. Next, the pathogen multiplies in host tissues, resisting or evading various nonspecific host defense systems. Actual disease symptoms are from damage to host tissues caused either directly

or indirectly by the microorganism's components or products.

Most genetic information in bacteria is carried in the chromosome. However, genetic information is also carried on plasmids, which are independently replicating structures much smaller than the chromosome. Plasmids may provide bacteria with additional virulence-related capabilities (such as pilus formation, iron transport systems, toxin production, and antibiotic resistance). In some bacteria, several virulence determinants are regulated by a single genetic locus. *See* PLASMID.

Microbial cell surfaces. The microbial cell surface plays a major role in virulence. It mediates adherence and resistance to host defense systems, and contains components responsible for important biological properties, including toxic ones. In both gram-positive and gram-negative bacteria a cell wall overlays the cell membrane, and in some strains a third layer known as the capsule may be present (see **illus.**) The capsule is important in allowing the microorganisms to resist the host's phagocytic defense systems. Peptidoglycan is a polymer of amino acids and amino sugars that is common to both types of bacteria, and is responsible for the rigidity of cell walls. In gram-negative bacteria a second, outer membrane is present, of which lipopolysaccharide is a characteristic component. Both peptidoglycan and lipopolysaccharide have toxic properties. Rod-like protein structures known as pili or fimbriae are present on the surface of some gram-negative bacteria that mediate adherence to host tissues. *See* BACTERIA.

Adherence. Most infections start at a mucous membrane, where pili are important virulence factors. Piliated strains of *Neisseria gonorrhoeae* are more virulent than nonpiliated strains, since pili increase attachment of the organism to human buccal and epithelial cells. In piglets a vaccine made by using pili provided effective protection against enteropathogenic *Escherichia coli* strains. In this case, antibodies to pili were responsible for resistance to infection. The interaction between host-cell receptors and microbial surface adhesins or ligands plays a marked role in the preference of certain microorganism for particular tissues, a phenomenon known as organ or tissue tropism. In a variety of bacterial pathogens, adhesion is followed by invasion of and multiplication in host cells. *See* CELLULAR ADHESION.

Nutrition. The nutritional environment provided by host tissues is an important factor in the multiplication of pathogenic microorganisms in living organisms, but is a poorly understood aspect of the host-parasite interaction. However, something is known concerning the role of iron availability in humans, and the bacterial response to it. The amount of free iron in the body is very low because various proteins in body fluids bind iron avidly. Iron is an essential bacterial nutrient, but pathogenic bacteria are nonetheless able to proliferate in this iron-poor environment. Under these conditions, bacteria often secrete low-molecular-weight molecules, siderophores, which bind iron with great affinity. Receptor proteins for

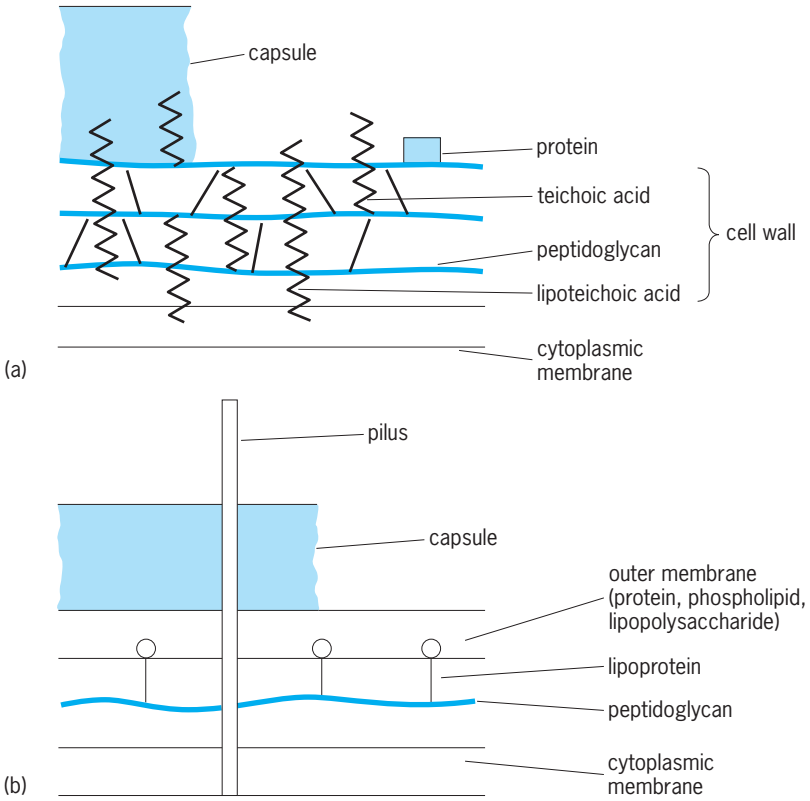

Cell surface of (*a*) gram-positive and (*b*) gram-negative bacteria.

the siderophore-iron complexes are induced in the outer membrane of gram-negative bacteria. This iron uptake system has been extensively studied in *E. coli*, and is present in a variety of other pathogenic bacteria. Commonly, the genetic information for iron uptake systems is carried on plasmids. *See* SIDEROPHORES.

Nonspecific host defenses. Two interrelated and important host defense systems are a cascade of serum proteins known as the complement system, and phagocytic cells. Host complement proteins are activated by microbial components such as lipopolysaccharide and peptidoglycan. Complement proteins promote inflammation and chemotaxis of phagocytes toward the invading organism, and render microorganisms susceptible to attachment and ingestion by phagocytes. Some complement proteins cause direct killing of gram-negative bacteria.

A variety of bacterial cell-surface components have been linked to resistance to complement-mediated killing, thereby increasing bacterial virulence. In *E. coli* these include outer membrane proteins, capsular polysaccharide, and complete lipopolysaccharide. The nature of the lipopolysaccharide is important in the resistance to complement-mediated killing of *Salmonella* species and *Shigella sonei*, while capsular polysaccharide is important in the resistance of type b *Haemophilus influenzae* and meningococci.

One of the earliest recognized virulence factors is the capsule, which allows microorganisms to resist phagocytosis. The complement component C3b and immunoglobulin G are opsonins (that is,

serum factors promoting the attachment and ingestion of microorganisms by phagocytic cells) when deposited on the microbial cell surface. The C3b and antibody are recognized by receptors in the plasma membrane of the phagocyte. In encapsulated bacteria, however, immunoglobulin G and C3b are deposited on the cell wall of the organism and are thus covered over by the capsule. This precludes their interaction with receptors on the phagocytes, and the microorganism avoids recognition and ingestion. Such a mechanism operates in some strains of *Staphylococcus aureus*, *Streptococcus pneumoniae*, and the yeast *Cryptococcus neoformans*, and may apply to other encapsulated bacteria. Antibodies directed against capsular antigens are required for efficient phagocytosis of encapsulated bacteria.

Certain bacteria, although efficiently phagocytized, are able to resist the intracellular killing action of phagocytes. In some of these cases, resistance to killing has been associated with cell-surface components (for example, a peptidoglycolipid in *Mycobacterium* species).

Specific host defenses. Some microorganisms can alter the surface that they present to the host. This antigenic variation is best understood in trypanosomes, which are protozoa. The trypanosome surface coat is composed of a protein known as the variant-surface glycoprotein to which the host develops immunity. The host defense system fails to eliminate the parasites completely because a few trypanosomes replace their coat with a different glycoprotein, requiring a new immunological response, by a process known as antigenic variation. The disease thus shows a cyclical nature, and antigenically different forms are isolated from patients in relapse. *See* CELLULAR IMMUNOLOGY; PHAGOCYTOSIS; RELAPSING FEVER; TRYPANOSOMATIDAE.

Host damage. Bacterial substances causing host damage can be divided into endotoxins, exotoxins, and tissue-degrading enzymes. Lipopolysaccharide (endotoxin) has a variety of pharmacologic and pathological effects, most of which seem to reside in the part of the molecule known as lipid A. Damaging effects of lipopolysaccharide include coagulative responses, necrotic lesions, effects on circulating leukocytes and platelets, changes in blood pressure, and fever-producing activity. Peptidoglycan shares many of the pathogenic properties of lipopolysaccharide, but is generally less potent.

Both peptidoglycan and lipopolysaccharide activate the complement system. Too much activation may damage the host. A variety of microbial cell-surface antigens can interact with antibodies and complement to form circulating immune complexes that may damage host tissues.

Exotoxins are proteins secreted by growing bacteria. Compared to endotoxins, they are less stable to heat, more specific in their actions, and more potent. Exotoxins may be divided into two major groups, membrane-damaging (phospholipases, hemolysins, lysins) or intracellular-acting. Phospholipases are produced by species of *Clostridium* and hemolysins by *Staphylococcus aureus* strains. *Staphylococcus aureus* also produces leukocidin, which creates pores in the cell membranes of polymorphonuclear leukocytes and macrophages. Examples of intracellular-acting toxins are *Corynebacterium diphtheriae* toxin, which interferes with protein synthesis and is lethal for many animal species; *Vibrio cholera* toxin, an enterotoxin causing massive diarrhea and dehydration; and *Clostridium tetanus* and *C. botulinum* toxins, which are potent neurotoxins. *See* BOTULISM; CHOLERA; DIPHTHERIA; TETANUS.

Certain generalizations can be made about these toxins. The genes specifying their production are often encoded by a prophage (a virus integrated into the bacterial chromosome) or by a plasmid. These toxins typically show a dichain construction: the alpha chain possesses the biological activity, while the beta chain mediates binding to receptors on susceptible host cells. These host receptors may actually be hormone receptors also used by the toxins. Toxins are often produced in inactive proenzyme forms that must be cleaved enzymatically to release active alpha fragments. These fragments may inhibit protein synthesis (such as in diphtheria toxin and *Pseudomonas aeruginosa* exotoxin A) or may elevate cellular cyclic adenosine monophosphate levels, altering normal ion flux (such as in cholera toxin).

A variety of bacteria secrete enzymes such as lipases, proteases, coagulases, hyaluronidases, and nucleases. These may play a role in virulence by degrading host tissues and perhaps helping nourish the parasite. In most cases, it has not been possible to make definitive statements about the role of these substances as virulence factors. *See* TOXIN.

From an evolutionary standpoint, the host-parasite relationship can be regarded as an uneasy truce. If the parasite acquires excessive virulence and kills its host, it seriously diminishes its chances of finding a new host. Furthermore, much of the apparent virulence of pathogens is actually a reflection of the reaction of the host's defense systems. *See* MEDICAL BACTERIOLOGY.

Brian Wilkinson

Bibliography. G. H. Cassell (ed.), Microbial surfaces: Determinants of virulence and host responsiveness, *Rev. Infect. Dis.* 10(suppl. 2):S273–S456, 1988; B. B. Finlay and S. Falkow, Common themes in microbial pathogenicity, *Microbiol. Rev.*, 53:210–230, 1989; M. Schaechter, G. Medoff, and D. Schlessinger, *Mechanisms of Microbial Disease*, 3d ed., 1998.

Virus

Any of the elementary agents that possess some of the properties of living systems, such as having a genome and being able to adapt to changing environments. However, viruses are not functionally active outside their host cells. Viruses share three characteristics: (1) their simple, acellular organization consisting of a nucleic acid genome surrounded by a protective protein shell, which may itself be enclosed within an envelope that includes a membrane; (2) the presence of either DNA or RNA, but

not both; and (3) their inability to reproduce independent of host cells. In essence, viruses are nucleic acid molecules, that is, genomes that can enter cells, replicate in them, and encode proteins capable of forming protective shells around them.

One debate since the discovery of the chemical nature of viruses is whether or not these agents are living microorganisms. Lacking any metabolism, they can be considered as lifeless as isolated chromosomes. They are, however, capable of reproducing their own kind abundantly and, in that sense, possess one of the major attributes of life. Some scientists believe that this is sufficient reason to term them as living. Others, however, believe that these agents should not be referred to as living, as they lack many of the basic abilities of life. It is preferable to refer to them as functionally active or inactive rather than living or dead.

The origin and evolution of viruses is a subject of much speculation among virologists. Two hypotheses have been proposed to account for their development. The first is that some of the more complex viruses, such as the poxviruses and herpesviruses, arose from small cells (probably prokaryotic) that parasitized larger, more complex cells. These parasitic cells then became simpler and more dependent on their hosts in a process known as retrograde evolution. There are several problems with this hypothesis. For example, viruses are radically different from prokaryotes; intermediate forms between the prokaryote and the complex viruses should still be in evidence but have not been found. The second hypothesis is that viruses are cellular nucleic acids that have become partially independent of the cell. Evidence that suggests this hypothesis may be true includes the observation that the nucleic acids of retroviruses contain sequences quite similar to those of normal cells.

Viruses are recognized as significant causes of disease in animals and plants. Many of the most important diseases that afflict humankind, including poliomyelitis, hepatitis, influenza, the common cold, measles, mumps, chickenpox, herpes, rubella, hemorrhagic fevers, encephalitis, and the acquired immunodeficiency syndrome (AIDS), are caused by viruses. AIDS, first diagnosed in 1981 and caused by the human immunodeficiency virus (HIV), represents an unprecedented epidemic with enormous implications for world health because case mortality rates for HIV-infected individuals exceed 90%. Current estimates of the numbers of infected human beings (intravenous drug abusers, male and female sex partners of HIV-positive individuals, and children of infected women) approach 45 million, and their numbers are still increasing rapidly in many parts of the world. So far, efforts to prevent or control HIV infection have been largely unsuccessful. Viruses also cause diseases in livestock (foot-and-mouth disease, avian influenza, hog cholera, distemper) and plants (tobacco mosaic disease, tomato ring spot, rice dwarf disease) that are of great economic importance. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); PLANT VIRUSES AND VIROIDS.

Viruses are also the simplest model systems for studying basic problems in biology. Their genomes are often no more than one-millionth the size of, for example, the human genome; yet the principles that govern the behavior of viral genes are the same as those that control the behavior of human genes. Viruses thus afford unrivaled opportunities for studying mechanisms that control the replication and expression of genetic material. *See* HUMAN GENOME PROJECT.

Morphology and size. Virus particles range in size from about 10 to 400 nanometers in diameter. The smallest viruses are about the size of a ribosome, while the largest, most complex viruses (such as vaccinia) are visible in the light microscope. However, most viruses must be visualized using the electron microscope. Although viruses differ widely in shape and size (see **illustration**), they are constructed according to certain common principles. Basically, viruses consist of nucleic acid and protein arranged in a structure known as the nucleocapsid. The nucleic acid is the genome containing the information necessary for virus multiplication and survival; the protein is arranged around the genome in the form of a layer or shell that is termed the capsid. Some viruses consist only of a naked nucleocapsid, while in others the nucleocapsid is surrounded by a membrane on the outside of which "spikes" composed of glycoproteins may be attached; this is termed the envelope. The complete virus particle is known as the virion, a term that denotes both intactness of structure and the property of infectiousness.

There are four general morphological types of capsids and virion structure. Icosahedral capsids are regular polyhedrons with 20 equilateral triangular faces and 12 vertices, appearing as spherical shapes when viewed at low power on the electron microscope. Helical capsids are shaped like hollow cylinders containing an extended viral nucleic acid. Enveloped viruses contain either helical or icosahedral capsids and have a roughly spherical shape overall. Complex viruses are neither purely icosahedral or helical and possess tails or other structures not found in simpler viruses.

Viral genomes. Viral genomes are astonishingly diverse. Some consist of deoxyribonucleic acid (DNA), others of ribonucleic acid (RNA). Some are double-stranded, others single-stranded. Some are linear, others circular. Some contain positive-sense RNA, meaning the genome can be directly read and translated into proteins, while others consist of negative-sense RNA and must be converted to a positive strand in order to be translated. Some consist of one molecule, others of several (up to 12). Their size also varies within wide limits: they range from 3000 to 280,000 base pairs if double-stranded, and from 5000 to 27,000 nucleotides if single-stranded. *See* DEOXYRIBONUCLEIC ACID (DNA); RIBONUCLEIC ACID (RNA); VIRUS CLASSIFICATION.

Viral proteins. Viral genomes encode three types of genetic information. First, they encode the structural proteins of virus particles. Second, most viral genomes can be neither replicated nor decoded (that

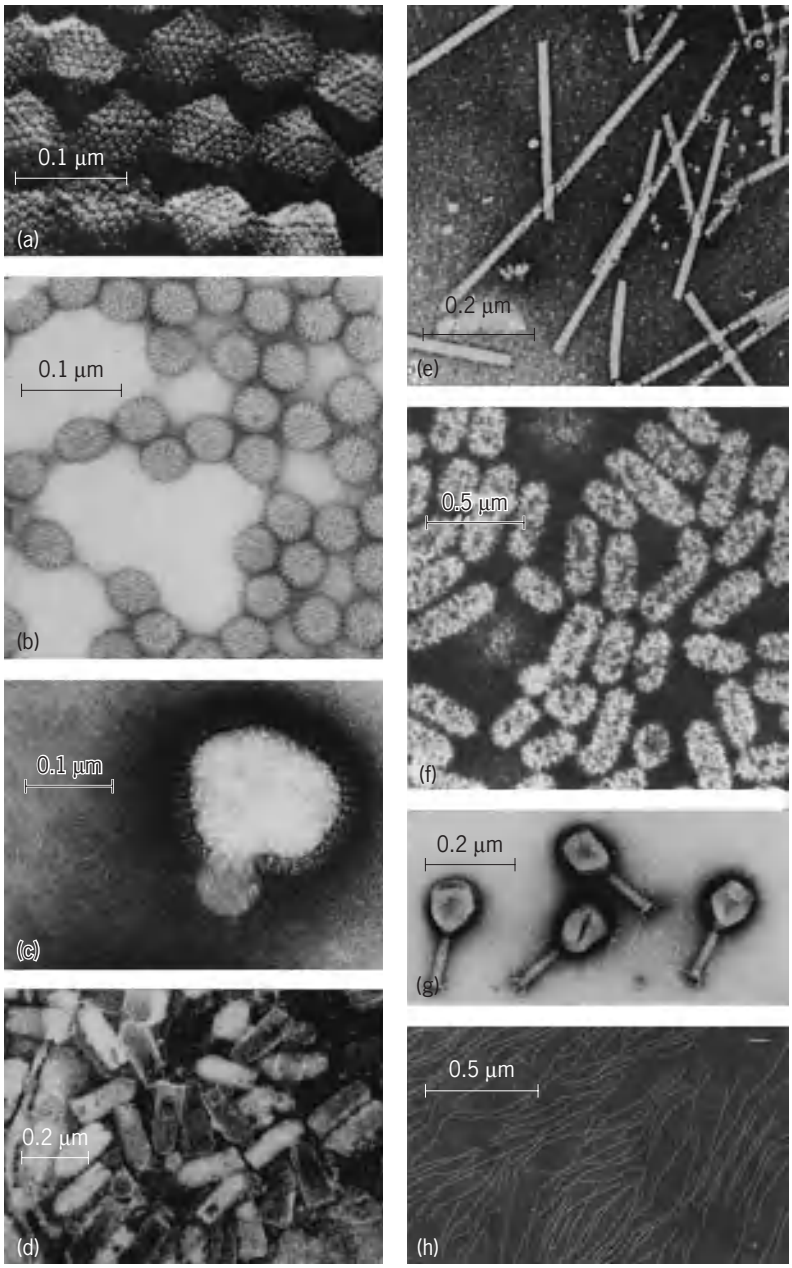

Electron micrographs of highly purified preparations of some viruses. (*a*) Adenovirus. (*b*) Rotavirus. (*c*) Influenza virus (*courtesy of George Laser*). (*d*) Vesicular stomatitis virus. (*e*) Tobacco mosaic virus. (*f*) Alfalfa mosaic virus. (*g*) T4 bacteriophage. (*h*) M13 bacteriophage.

is, transcribed into messenger RNA) by host-cell enzymes. Most viruses, therefore, encode enzymes capable of transcribing their genomes into messenger RNA molecules that are then translated by host-cell ribosomes, as well as nucleic acid polymerases capable of replicating their genomes. For example, the RNA genome of retroviruses encodes the enzyme reverse transcriptase, which is carried within the capsid when the virus infects a host cell. This enzyme allows these viruses to copy their genome and insert it into the host-cell chromosome, where it remains for long periods of time before it becomes activated by an outside stimulus. Many viruses also encode nonstructural proteins with catalytic and other functions necessary for virus particle maturation and morphogenesis. Third, many viruses encode proteins that interact with components of host-cell defense mechanisms against invading infectious agents. The more successful these proteins are in neutralizing these defenses, the more virulent viruses are and the more severe the resulting disease. Larger viruses encode specific proteins with these functions; smaller viruses do not, and for them these functions are exercised primarily by the structural proteins that they encode.

Virus–cell interaction. The two most commonly observed virus–cell interactions are the lytic interaction, which results in virus multiplication and lysis of the host cell, and the transforming interaction, which results in the integration of the viral genome

into the host genome and the permanent transformation or alteration of the host cell with respect to morphology, growth habit, and the manner in which it interacts with other cells.

Lytic interaction. The lytic virus–cell interaction is best thought of in terms of a cycle, the one-step growth cycle. This cycle involves the sequential, precisely regulated expression of the information encoded in the viral genome. The manner in which viral genomes express the information encoded in them is characteristic of each virus family.

The key features of the one-step growth cycle, which lasts, depending on the virus, from 6 to 36 hours, are as follows. The parental virus particle adsorbs (attaches) to specific receptors located on the host-cell surface, and is internalized by a process akin to phagocytosis. The viral genome is then either completely or partially released and expresses the information that it encodes by being transcribed into messenger RNA molecules (this is not necessary if the genome is positive-sense RNA), which are then translated. At the same time, the viral genome replicates either in the nucleus or in the cytoplasm, depending on the virus. Generally, DNA viruses replicate their genome in the nucleus, whereas RNA viruses carry out this process in the cytoplasm. When a sufficient amount of capsid proteins has accumulated, morphogenesis proceeds and progeny virus particles (up to 10^6 per cell for small viruses) are formed. Throughout this period, degradative and necrotic changes are elicited that result in the disintegration (lysis) of the cell. Naked viruses are released when the cells lyse; enveloped viruses are released when their nucleocapsids bud through the outer cell membrane, thereby acquiring their envelope.

When viruses infect multicellular organisms, the one-step growth cycle is repeated many times until, for one reason or another, further multiplication is arrested, or the host dies. The symptoms of virus infection vary widely, from asymptomatic infections detectable only by the formation of antibodies, to progressively more severe disease culminating in death. For every virus there are variant strains that differ in the severity of their effects on cells and host organisms: the more severe the effects, the more virulent the virus strain is said to be.

Transforming interaction. Certain viruses interact with cells not only by means of the lytic interaction but also by means of an interaction in which virus multiplication is repressed and the host cell is not destroyed. In this type of interaction, either the viral genome is integrated into the genome of the host cell or it replicates as a plasmid. The cells transformed by these viruses do not die and are capable of multiplying. Transformed bacterial cells are said to be lysogenic; they grow like normal cells, but when the integrated viral genome that they harbor is activated, which happens with a frequency of about 1 in 10^5, a viral growth cycle ensues and the cells lyse. Transformed animal and plant cells are also capable of multiplying; they often grow into tumors, and the viruses that cause such transformation are known as tumor viruses.

Tumor viruses are extremely important in studies on the mechanisms of cancer induction (oncogenesis) for two reasons. First, infection with tumor viruses generally transforms normal cells into potential tumor cells very efficiently; with all other tumorigenic agents, the frequency of transformation is much lower. As a result, studies at the biochemical and molecular level are possible on virus-induced transformation but not on transformation caused by other agents. Second, elucidation of the mechanisms by which tumor viruses transform cells has identified the causes of carcinogenesis in molecular terms.

Viruses can cause cancer in several ways. A virus may bring oncogenes into a cell and insert them into its genome. For example, the retrovirus Rous sarcoma virus carries a gene that codes for tyrosine kinase. This enzyme is responsible for the phosphorylation of many cellular proteins which are regulated by the addition of a phosphate group. By bringing in a new source of phosphorylation, the regulation of these proteins is altered and cell growth and behavior change as a result, potentially leading to cancer.

Other oncogenic viruses carry one or more very effective promoters, which can then insert next to a cellular oncogene. This promoter then stimulates transcription of the oncogene, leading to cancer. For example, some chicken retroviruses cause lymphomas when their genome is integrated next to a cellular oncogene. This mechanism is known as insertional activation or proviral insertion.

A final mechanism involves the inactivation of proteins that negatively regulate cell division. The normal regulation of cell division involves the extremely sensitive interplay of proteins that either promote or inhibit it, depending on the signals received. Certain viruses, primarily DNA-containing viruses, encode proteins that bind or inactivate proteins that inhibit cell division. These viral proteins can therefore result in the activation of cell division, promoting formation of cancer cells as well as stimulating virus reproduction. *See* CANCER (MEDICINE); ONCOLOGY; RETROVIRUS; TUMOR VIRUSES.

Antiviral chemotherapy. Because viruses enter host cells and make use of host-cell enzymes and constituents to reproduce, development of drugs to treat viral infections seemed a remote possibility for many years. A drug that would block virus reproduction would likely be toxic to the host. However, inhibitors of virus-specific enzymes and life-cycle stages have now been discovered. Most antiviral drugs in use today disrupt either viral nucleic acid synthesis or specific stages in the virus life cycle. Viral nucleic acid synthesis is almost always carried out by virus-encoded enzymes that do not exist in uninfected cells and are therefore excellent targets for antiviral chemotherapy. Stages in the virus life cycle that have been elucidated through research are also used as targets in a number of antiviral treatments; these include enzymes required for attachment to host cells and enzymes responsible for virus maturation and release.

Numerous chemical compounds have been described that inhibit the multiplication of viruses.

Only a few, however, inhibit virus multiplication efficiently in the body without undesirable side effects. The most successful of these compounds are analogs of ribo- and deoxyribonucleosides, such as iododeoxyuridine, trifluorothymidine, and adenosine arabinoside (vidarabine, ara-A) for the topical treatment of herpes simplex keratitis and encephalitis; acyclovir and gancyclovir for the treatment of herpesvirus and cytomegalovirus infections, respectively; zidovudine (AZT), lamivudine (3TC), didanosine (ddI), zalcitabine (ddC), and stavudine (d4T) for the treatment of HIV infections (AIDS); and ribavirin (Virazole®) for the treatment of respiratory syncytial virus and Lassa fever virus infections. Another viral function that has been targeted is the cleavage of polyproteins, precursors of structural proteins, to their functional components by virus-encoded proteases; examples include the HIV protease inhibitors (saquinvir, indinavir, and ritonavir) used for treatment in AIDS patients. *See* CHEMOTHERAPY; CYTOMEGALOVIRUS INFECTION; HERPES; INFLUENZA; RESPIRATORY SYNCYTIAL VIRUS; VIRUS CHEMOPROPHYLAXIS.

Other promising new approaches to antiviral chemotherapy include the use of agents that interfere with virus adsorption (primarily against HIV); agents that inhibit the enzyme neuraminidase, which is essential for the entry of influenza virus into cells and the subsequent release of its progeny from cells; agents that bind in hydrophobic pockets of viral capsids, thereby stabilizing them and interfering with uncoating (primarily against picornaviruses such as rhinoviruses); the targeted introduction of toxins into infected cells (against viruses for which the receptors are known, such as HIV, rhinoviruses, and Epstein-Barr virus); and the introduction into cells of specific antisense RNA sequences.

Antiviral agents on which much interest is focused are the interferons. Interferons are cytokines or lymphokines that regulate cellular genes concerned with cell division and the functioning of the immune system. Their formation is strongly induced by virus infection; they provide the first line of defense against viral infections until antibodies begin to form. Interferons interfere with the multiplication of viruses by preventing the translation of early viral messenger RNAs. As a result, viral capsid proteins cannot be formed and no viral progeny results. Interferon is clinically useful in the treatment of hepatitis B and C infection.

By far the most effective means of preventing viral diseases is through mobilization of the immune system by vaccines. There are two types of antiviral vaccines: inactivated and attenuated active. Most of the antiviral vaccines currently in use are of the latter kind. The principle of antiviral vaccines is that inactivated virulent or active attenuated virus particles cause the formation of antibodies that neutralize a virulent virus when it invades the body. *See* ANIMAL VIRUS; VACCINATION; VIRUS, DEFECTIVE.

W. K. Joklik; Marcia M. Pierce

Bibliography. L. Collier and J. Oxford, *Human Virology*, 2d ed., Oxford University Press, 2000; D. M. Knight and P. M. Howley (eds.), *Fundamental Virology*, 4th ed., Lippincott, Williams & Wilkins, Philadelphia, 2001; L. M. Prescott, J. P. Harley, and D. A. Klein, *Microbiology*, 6th ed., McGraw-Hill, New York, 2005; V. van Regenmortel et al. (eds.), International Committee on Taxonomy of Viruses, *Virus Taxonomy: Classification and Nomenclature of Viruses: Seventh Report of the International Committee on Taxonomy of Viruses*, 2000; S. Schlesinger and M. Schlesinger, Viruses, in J. Lederberg (editor-in-chief), *Encyclopedia of Microbiology*, 2d ed., vol. 4, pp. 796–810, Academic Press, San Diego, 2000; B. A. A. Voyles, *The Biology of Viruses*, 2d ed., McGraw-Hill, Chicago, 2002.

Virus, defective

A virus that by mutation has lost the ability to be replicated in the host cell without the aid of a helper virus. The virus particles (virions) contain all the viral structural components; they can attach, penetrate, and release their nucleic acid ribonucleic acid (RNA), deoxyribonucleic acid (DNA)] within the host cell. However, since the mutation has destroyed an essential function, new virions will not be made unless the cell was simultaneously infected with the helper virus, which can provide the missing function. Only then will the cell produce a mixed population of new helper and defective viruses. Occasionally, when their nucleic acids become integrated in the DNA of the host cell, defective viruses persist in nature by propagation from mother cell to daughter cell.

Deletion mutants. The most important group of defective viruses are deletion mutants. They are derived from their homologous nondefective (wild-type) virus through errors in the nucleic acid replication, which result in the deletion of a fragment in the newly synthesized molecules. The deletion is a consequence of an excision or of a failure in the synthesis of the fragment. The defective nucleic acid must be capable of self-replication, at least in the presence of the wild-type virus, and must combine with other viral components to form a particle in order to exit the cell. Any other types of deleted nucleic acids, although probably transiently formed, are dead-end products, undetectable among the viral progeny.

The defective RNA tumor viruses are deletion mutants. Mammalian and most avian sarcoma viruses require a nondefective leukemia virus as a helper virus. Usually the specificity for a certain type of host cell exhibited by the defective virion depends on the helper virus, indicating that one of the virion surface proteins has been furnished by the helper virus gene. These proteins are involved in interactions with cellular surface receptors, and thus determine whether a cell can serve as a host for viral infection. There are also defective leukemia viruses sometimes called acute leukemia viruses. In all cases of defective RNA tumor viruses, the deleted RNA fragment was replaced by a new segment, which carries genetic

information required for the induction of tumors or neoplasia in the host. This gene, the oncogene, is thought to have arisen by a modification of a normal cellular gene. *See* ONCOLOGY; TUMOR VIRUSES.

Defective interfering particles. Certain deletion mutants, which interfere with the replication of the helper virus, are called defective interfering particles. Cells infected with a mixture of a standard virus and its defective interfering particles produce much less of the standard virus than the same type of cells infected with the standard virus in the absence of defective interfering particles. Defective interfering particles were discovered by P. Von Magnus in 1947 in infections with influenza virus. He called them incomplete or immature particles, and their ability to attenuate the replication of influenza virus he termed autointerference. Since then, defective interfering particles have been found in most of the major groups of viruses (see **table**).

The widespread occurrence of defective interfering particles has prompted the suggestion that they must perform an important role in the evolution of viral infection. The self-limiting nature of the infection imposed by the generation of defective interfering particles must be advantageous to the virus. Since viruses require living host cells for self-replication, an unlimited spread of infection resulting in the death of all available host cells is clearly not expedient to the survival of the virus. The plausibility of this hypothesis is borne out by the observation that in Australia, during attempts to control the rabbit population with myxomavirus, a rapid drift from the original virulent virus strain to less virulent strains has taken place.

Occurrence of defective interfering particles in RNA viruses*

Virus group	Member name
Negative strand	
Rhabdo	Vesicular stomatitis
	Rabies, bovine ephemeral fever, others
Paramyxo	Sendai
	Newcastle disease
	Measles
	Mumps
Orthomyxo	Influenza
	Fowl plague
Arena	Lymphocytic choriomeningitis, Amapair, Parana, Pichinde
	Tacaribe
Bunya	Bunyamwera
	LaCrosse
	Turlock
Positive strand	
Picorna	Poliovirus
	Mengovirus
Calici	Feline calcivirus
Toga	Sindbis, Semliki Forest, West Nile, Japanese encephalitis
Corona	Mouse hepatitis
Double strand	
Reo	Reovirus
Orbi	Bluetongue virus of sheep
Fungal	Yeast-killer, others
Other	Infectious pancreatic necrosis virus

*Defective interfering particles were also observed in DNA viruses of the Papova and Herpes groups. Only one type of adenovirus (adeno-12) seems to generate these particles.

SOURCE: J. Perrault, *Curr. Topics Microbiol. Immunol.*, vol. 93, 1981.

Morphology. Defective interfering particles derived from "spherical" (icosahedral) viruses, when viewed with the electron microscope, are indistinguishable from the standard virion. Since the nucleic acid of these virions is buried in the interior of the spherical structure, the deleted nucleic acid of the defective interfering particle has very little effect on the overall particle dimensions. On the other hand, in elongated (rod or bullet-shaped) virions, structural proteins are wound around the backbone of the ribbonlike nucleic acid molecule, whose size determines the long dimension of the particle. Consequently, in these cases defective interfering particles are shorter than their homologous virions, the foreshortening of the particle being proportionate to the size of the deletion. This is illustrated in the electron micrographs of vesicular stomatitis virus and some of its particles shown in the **illustration**. A nucleic acid with a 50% deletion is contained in a defective interfering particle (illus. *a*), half the length of the homologous virions. Other, more extensive deletions are found in particles correspondingly shorter.

Genome structure. A defective interfering particle's requirements for genome replication and incorporation into a virionlike particle result in the selection of certain nucleic acid structures. In the case of DNA viruses, the selection leads to deleted DNAs, with tandem duplications of a segment which contains the origin of replication. One advantage of these structures is their more efficient replication through multiple initiations, and another may reside in a selection of a minimum DNA size, in order to form a particle. The minimum DNA size is sometimes achieved by acquisition of host-cell DNA segments. The mechanism by which these structures are formed is not understood, although many hypotheses have been proposed.

Based on the nucleic acid structure, RNA-containing defective interfering particles can be broadly classified into two groups. In one group are particles in which the deletion of the linear RNA molecule took place at one end. In the other group the deletion was internal with conserved ends. Some viruses, such as vesicular stomatitis virus, generate both types of defective interfering particles. The terminally deleted RNAs have yet another structural peculiarity: a sequence of approximately 50 terminal nucleotides (the building blocks of nucleic acids) at the two ends are inverted copies of each other (inverted terminal repeats). Such structures are not found in the homologous viral genome, and are thought to be intimately connected with the ability of these defective interfering particles to interfere with the replication of the virus. They also furnish important clues in studies of the molecular mechanisms by which these particles are generated.

Interference. It is generally thought that the reason defective interfering particles interfere with the production of their homologous viruses is that the

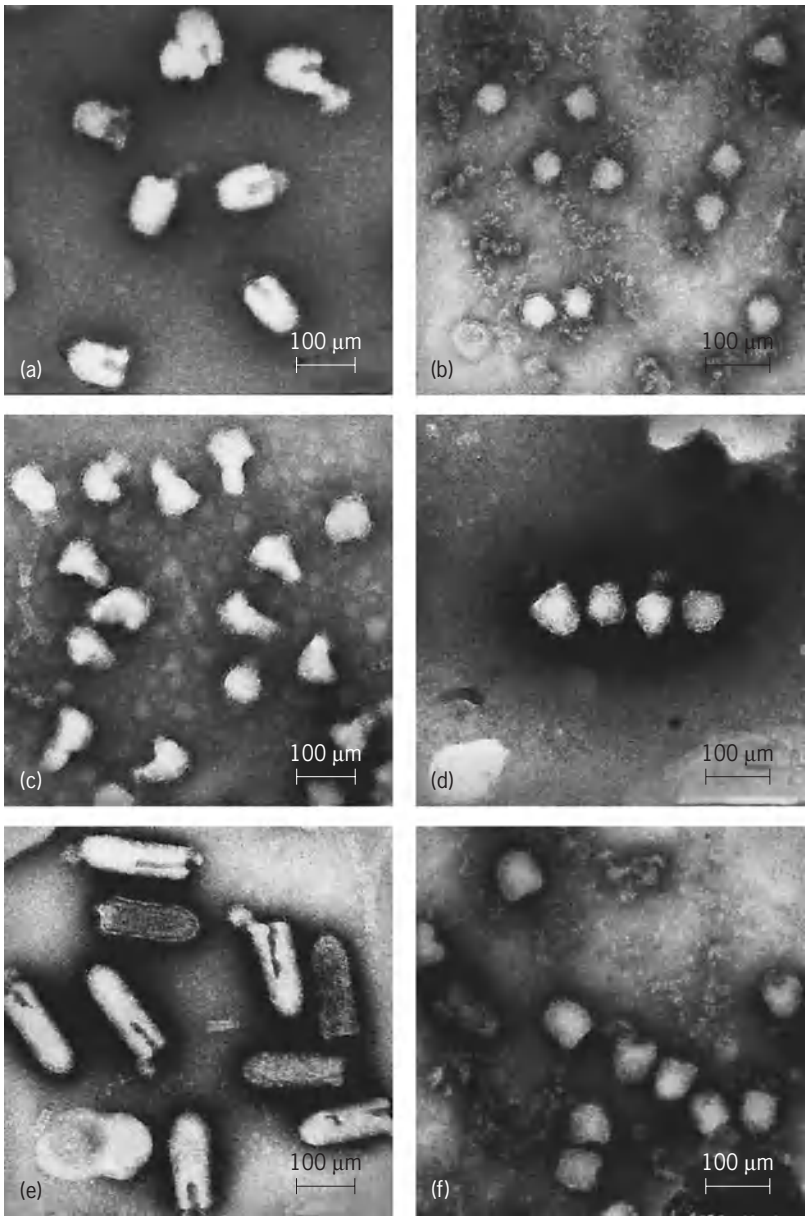

Electron micrographs of virion and defective interfering particles of vesicular stomatitis virus. (*a*) Defective interfering particle with a 50% deleted RNA. (*b*) Defective interfering particle with an 85% deleted RNA. (*c*, *d*, *f*) Defective interfering particles from various isolates, all with a 70% deleted RNA. (*e*) Vesicular stomatitis virion. (*From M. E. Reichmann, C. R. Pringle, and E. A. C. Follett, J. Virol., 8:156, 1971*)

nucleic acids of defective interfering particles have a replicative advantage over those of the homologous viruses and are capable of competing successfully for the synthetic enzymes. If at some stage of the infectious cycle these enzymes are in limited supply, they might be expected to be engaged predominantly in the synthesis of the defective-interfering-particle nucleic acid. In that case, insufficient enzyme will be available for viral nucleic acid replication. In the synthesis of defective-interfering-particle DNA with reiterated origins of replication, multiple initiation processes might tie up most of the available enzymes. Alternatively, in the case of terminally deleted defective-interfering-particle RNA, the unique inverted repeats may represent a stronger binding site for the synthetic enzymes than the corresponding but different structures of the virion RNA.

Defective interfering particles with RNAs deleted internally seem to interfere by a different mechanism. In repeated rounds of replication, a selection of defective interfering particles with progressively shorter RNA molecules has been reported. This observation suggested that the rate of RNA synthesis might be the major factor in interference. The shorter defective-interfering-particle RNA molecules, which take less time to be synthesized than the longer virion RNA, would accumulate if the rate of synthesis were the major time-determining factor. In multiple rounds of infections, the shortest defective-interfering-particle RNAs would eventually outnumber both virion RNAs and the longer defective-interfering-particle RNAs, and utilize all the available synthetic enzymes. *See* VIRUS INTERFERENCE.

Specificity of helper and interference functions. Defective interfering particles can interfere only with the cellular synthesis of the virus strain from which they were originally derived, or with virus strains closely related to it. The more distant the relationship, the weaker the interference. The extent of interference can therefore serve as a measure of virus strain relatedness, in addition to other classification criteria such as morphology and serology. Similarly, the ability to serve as a helper to a defective virus is highly specific. In some cases of defective viruses (unlike the defective-interfering-particle group), the relationship between the helper virus and defective virus is not obvious. The defective adeno-associated virus and the satellite tobacco necrosis virus are examples in which the helper-defective virus relationship is also specific. Only adenoviruses and tobacco necrosis viruses can serve as helpers, respectively.

Persistent infections. Defective viruses have been used in studies of persistent, or low-level, infections. Normally cytolytic viruses kill their host cells both in cultures and in intact organisms. Occasionally, however, some cells survive, multiply, and simultaneously produce low levels of virus. In whole organisms, persistent infections can result in periodic episodes of the acute viral disease, especially if the immune response to the virus becomes temporarily compromised. Persistent infections are also involved in degenerative brain diseases and can lead to complications, sometimes years after a bout with an acute virus disease, such as measles. Laboratory models of carrier cultures infected with several viruses have been developed. In some of these the aid of defective interfering particles in the attenuation of the virulence was found to be required for both the establishment and maintenance of the culture. Although defective interfering particles were isolated also in a few cases from naturally occurring persistent infections, it is not clear whether they play a major role in the evolution of these diseases. In many cases the diminished virulence observed in persistent infections is the property of the infectious virion itself, rather than being due to interference by defective interference particles. Frequently virions isolated from persistent infections are temperature-sensitive mutants.

Temperature-sensitive mutants. A mutation might lead to a change of a single amino acid of a protein

encoded in the viral genome. Such mutations usually precipitate a change in the three-dimensional structure of the pertinent protein, accompanied by a loss of biological activity. In some cases, however, the altered protein can maintain its original conformation at slightly lower than the normal body temperatures (for example, 31°C or 87.8°F instead of 37°C or 98.6°F). Virus mutants of this kind, called temperature-sensitive (ts) mutants, do not grow efficiently at normal body temperatures, but they grow like wild-type virus at the lower temperatures. These mutants belong to the group of conditional lethal mutants, which have played an important role in genetic studies. For research purposes they can be grown to good yields under the permissive conditions (such as low temperatures in the case of temperature-sensitive mutants), while other types of mutants often cannot be obtained in adequate quantities.

The reduced growth and virulence of temperature-sensitive mutants at normal body temperatures might serve as a basis for their selection in persistent infections. In any case, temperature-sensitive mutants have been isolated from such infections in animals or carrier cultures of foot-and-mouth disease, sendai, vesicular stomatitis, herpes, influenza, coxsackie, and some alpha viruses. It is not known whether defective interfering particles, temperature-sensitive mutants, or both constitute a major path to persistent infections and virus latency. *See* ANIMAL VIRUS; MUTATION; VIRULENCE; VIRUS INFECTION, LATENT, PERSISTENT, SLOW; VIRUS INTERFERENCE. M. E. Reichmann

Bibliography. H. Fraenkel-Conrat and R. R. Wagner (eds.), *Comprehensive Virology*, vol. 10, 1977, vol. 16, 1980, vol. 18, 1983; J. A. Levy, H. Fraenkel-Conrat, and R. A. Owens, *Virology*, 3d ed., 1995; M. E. Reichmann and W. M. Schnitzlein, Defective interfering particles of rhabdoviruses, *Curr. Topics Microbiol. Immunol.*, 86:123–168, 1979.

Virus chemoprophylaxis

The treatment of viral disease by chemical drugs has been approached experimentally through the inhibition of intracellular metabolic processes that otherwise lead to the synthesis of viral constituents. Compounds have also been tested for their capacity to inhibit adsorption, penetration, or release of infectious virus from the host cell. In order to be considered for chemoprophylaxis, a compound must have a greater specificity for inhibiting virus-directed reactions than for normal host cell reactions. Although some success has been achieved, there are no inhibitors available that can be used in the routine treatment of viral infections in humans.

Very few inhibitors have been found which have a specific effect on virus penetration or release from the host cell. Amantadine, a symmetrical amine, inhibits specifically certain members of the myxovirus group through blocking viral penetration into the host cell. When administered prophylactically, that is, in advance of exposure to the virus, amantadine is reported to have a significant protective effect in experimental animals and humans against certain influenza A strains but not against influenza B or other viruses. This protection consists mainly of a modification of the disease to a milder form.

Guanidine and 2-(α-hydroxybenzyl)-benzimidazole (HBB) specifically inhibit the replication of many picornaviruses in the test tube, for example, poliovirus, by restricting the formation of virus-specific enzymes, which indirectly results in an inhibition of the synthesis of viral ribonucleic acid (RNA) and coat proteins. In experimentally infected animals, there is no protective effect by either inhibitor. This is probably due to rapid production of drug-resistant mutants.

Methisazone (isatin-β-thiosemicarbazone) is an inhibitor of many members of the poxvirus and adenovirus groups. Treatment of infected human cells with this inhibitor results in the formation of immature, noninfectious virus particles. A derivative of this compound has been shown to be an effective prophylactic for smallpox in humans if given within 24–48 h after exposure.

Many purine and pyrimidine analogs inhibit both ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) synthesis. By the addition of ribose or deoxyribose to make the corresponding riboside or deoxyriboside, the activity of the analog can be directed preferentially toward the inhibition of RNA or DNA. The significance of this structural specificity of the sugar component for the inhibition of virus synthesis is evident in the finding that the riboside of 5-fluorouracil is a much more effective inhibitor of the growth of tobacco mosaic virus than its deoxyriboside. Ribosides of halogenated benzimidazoles are more selective inhibitors of influenza virus or poliovirus multiplication than the free benzimidazoles or their deoxyriboside. *See* NUCLEIC ACID; PYRIMIDINE.

Another difference between the two types of nucleic acid is in their base composition, since uracil is present in RNA and thymine in DNA. Pyrimidine analogs may therefore exert an action according to their structural similarity to either of these two bases (see **illus.**). With halogenated pyrimidines, the size of the halogen substituent determines the character of the compound. The van der Waals radii of fluorine and bromine resemble closely those of hydrogen and the methyl group, respectively: The size and shape of 5-bromouracil is very similar to that of thymine, and that of 5-fluorouracil is similar to that of uracil.

5-Bromouracil, especially in the form of its deoxyriboside, is an effective inhibitor of the synthesis of DNA bacteriophage but is without effect on the synthesis of RNA tobacco mosaic virus. 5-Fluorouracil, on the other hand, inhibits the growth of the RNA virus, and its action can be reversed by uridine but not by thymidine.

5-Fluoro-2′-deoxyuridine (FUDR) inhibits DNA synthesis by inhibiting the enzymatic synthesis of thymidylic acid; 5-bromo-2′-deoxyuridine (BUDR) and 5-iodo-2′-deoxyuridine (IUDR) are incorporated into DNA, resulting in the production of a faulty nucleic acid which does not function normally. All three of these halogenated deoxyuridines inhibit replication of members of the major DNA virus groups: papova-, adeno-, herpes-, and poxviruses.

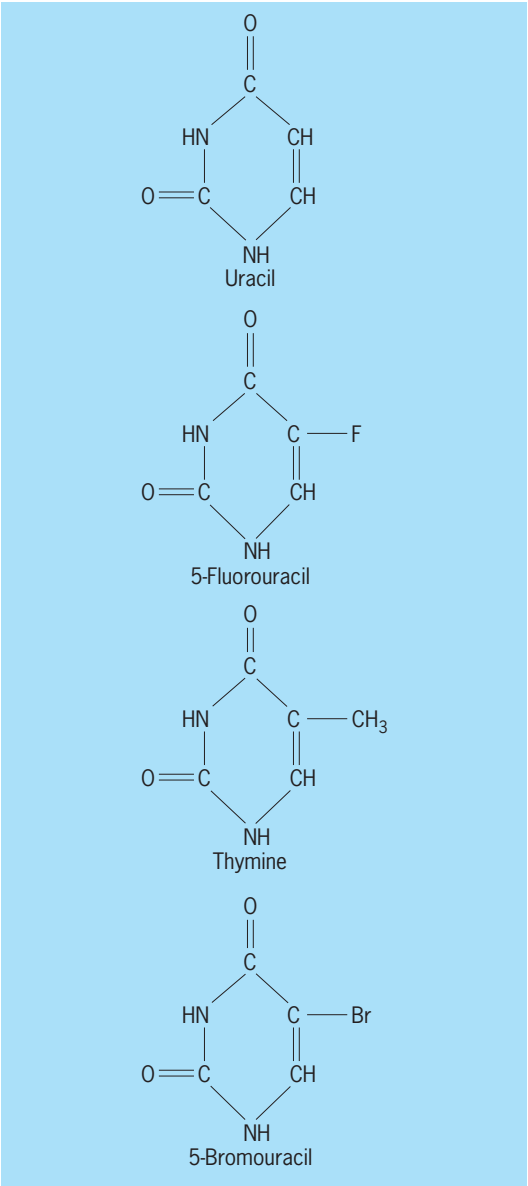

Pyrimidines uracil and thymine and their analogs.

Topical administration of IUDR is being used in humans in the treatment of corneal lesions due to herpes simplex virus. Systemic treatment of viral infections is not practicable because of the toxicity of the drugs. However, in cases of herpesvirus encephalitis, massive near-lethal doses (0.006 oz/lb or 400 mg/kg) of IUDR have been reported to lead to complete recovery.

Another pyrimidine analog, 1-β-*d*-arabinofuranosylcytosine hydrochloride (cytosine arabinoside), inhibits DNA synthesis, and thereby the replication of DNA viruses. As with IUDR, cytosine arabinoside inhibits cellular DNA synthesis and viral DNA synthesis about equally and therefore exhibits little viral specificity.

Dactinomycin (actinomycin D) inhibits DNA-dependent RNA synthesis and thus inhibits the multiplication of DNA viruses, but not most RNA viruses. However, dactinomycin inhibits the multiplication of some RNA-containing viruses (some myxoviruses and Rous sarcoma virus). The exact mechanism of this dactinomycin inhibition is not clear. Rifampicin (Rifampin) is an antibiotic whose action depends upon its preferential inhibition of bacterial RNA polymerase over animal-cell RNA polymerase. Since poxviruses carry their own RNA polymerase for synthesizing viral messenger RNA (needed for transcription of the viral genome), rifampicin was tried against this virus in culture and found to be very effective in inhibiting its replication. Although the exact mechanism of action of this compound is not known, it is thought to block assembly of DNA and proteins into the mature virus particle.

Although not practical as chemotherapeutic agents, protein inhibitors have been useful in the study of viral replication. For example, puromycin, cycloheximide, and ρ-fluorophenylalanine all inhibit synthesis of both viral and cell proteins. *See* BACTERIOPHAGE; CHEMOTHERAPY; VIRUS.

Joseph L. Melnick

Virus classification

There is no evidence that viruses possess a common ancestor or are in any way phylogenetically related. Nevertheless, classification along the lines of the Linnean system into families, genera, and species has been partially successful. In the early 1970s the International Committee on Taxonomy of Viruses (ICTV) was established. This article is a summary of the committee's most recent findings. Based on the organisms they infect, the first broad division is into vertebrate; algae, fungi, yeast, and protozoan; invertebrate; plant; and bacterial viruses. (However, viral families may fall into more than one of these classes.) Within these classes, other criteria for subdivision are used. Among these are general morphology: envelope or the lack of it; nature of the genome (deoxyribonucleic acid, DNA, or ribonucleic acid, RNA); structure of the genome [single-stranded (ss) or double-stranded (ds), linear or circular, fragmented or nonfragmented]; mechanisms of gene expression and virus replication (positive- or negative-strand RNA); serological relationship; host and tissue susceptibility; pathology (symptoms, type of disease). Information about many of these parameters has been available only since the 1950s, and was acquired after the introduction of tissue culture laboratory techniques.

Families and Genera of Viruses Infecting Vertebrates

The families of vertebrate viruses and the shapes and sizes of their characteristic members are presented in **Fig. 1**. Families are sometimes subdivided into subfamilies; the suffix -virinae may then be used. The subgroups of a family or subfamily are equivalent to the genera of the Linnean classification. *See* ANIMAL VIRUS.

DNA viruses. The vertebrate DNA viruses are divided into 10 families: Asfarviridae, Poxviridae, Iridoviridae, Hepadnaviridae, Herpesviridae, Polyomaviridae, Papillomaviridae, Adenoviridae, Circoviridae, and Parvoviridae.

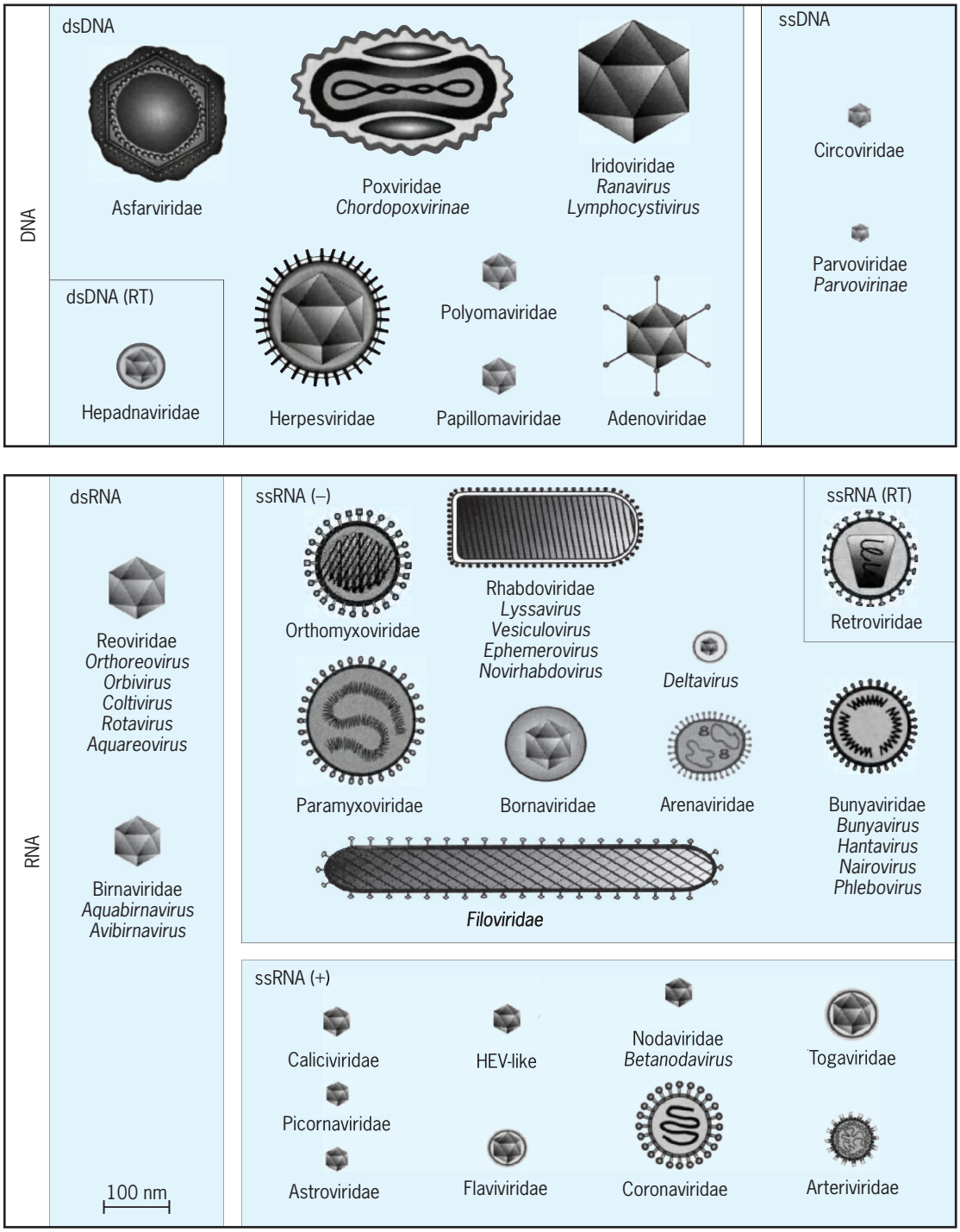

Fig. 1. Families and genera of viruses infecting vertebrates. Individual frames separate taxa of viruses containing double-stranded and single-stranded genomes. Large boxes separate taxa of viruses having DNA or RNA genomes. All diagrams have been drawn approximately to the same scale to provide an indication of the relative sizes of the viruses. (*From M. H. V. van Regenmortel et al., eds., Virus Taxonomy: 7th Report of the International Committee on Taxonomy of Viruses, p. 30, copyright 2000, with permission from Elsevier*)

Asfarviridae. The properties of this family are based on the sole genus *Asfivirus* (African swine fever virus). Virions consist of a nucleoprotein core structure (70–100 nm in diameter), surrounded by internal lipid layers and an icosahedral capsid (170–190 nm in diameter) and a lipid-containing envelope. The genome is a single molecule of linear dsDNA (170–190 kbp). African swine fever virus infects domestic and wild swine and can be transmitted by ticks, direct contact, fomites (disease-carrying objects), or mechanically by biting flies.

Poxviridae. The properties of this family are based on vaccinia virus of the *Orthopox* genus. The brick-shaped, (200 × 200 × 250)-nanometer particle contains a dumbbell-shaped nucleocapsid with a genome of dsDNA (170–250 kbp). At least 30 structural proteins have been reported. In spite of the lipid-containing envelope, most of these viruses are ether-resistant (some members of the poxviridae are, however, ether-sensitive). Generally,

members of the Poxviridae are sensitive to common detergents, formaldehyde, oxidizing agents, and temperatures greater than 40°C (104°F). The large genome encodes a great number of enzymes which lead to a more autonomous virus replication than that of any other family. Unlike any other DNA virus, the poxvirus replicates in the cell cytoplasm, independent of functions found in the cell nucleus. Poxviridae are divided into two subfamilies: Chordopoxvirinae (which infects vertebrates) and Entomopoxvirinae (which infects invertebrates; Fig. 3). Genera of Chordopoxvirinae include *Orthopoxvirus* (human smallpox), *Leporipoxvirus* (rabbit myxoma), *Avipoxvirus* (chickenpox), *Capripoxvirus* (sheeppox), *Suipoxvirus* (swinepox), *Mulluscipoxvirus* (molluscum contagiosum), *Yatapoxvirus* (Yaba monkey tumor virus), and *Parapoxvirus* (pseudocowpox). *See* SMALLPOX VIRUS.

Iridoviridae. These are mostly insect viruses except for the genus *Ranavirus* (frog viruses) and the genus *Lymphocystivirus* (fish viruses). The icosahedral virion is 120–200 nm in diameter but may be up to 350 nm, with a dsDNA genome (between 140 and 303 kbp). Several structural proteins and also lipids (not in a lipid bilayer envelope) are contained in the capsid. Viruses released from cells may have a plasma-derived outer envelope. The viruses are sensitive to treatment with ether.

Hepadnaviridae. Hepadnaviruses are spherical, sometimes pleomorphic (variably formed) viruses. The virion particles are 40–48 nm in diameter after negative staining. An outer envelope surrounds an icosahedral nucleocapsid core containing the viral genome, the viral polymerase, and associated cellular proteins. The genome consists of a single molecule of circular DNA (3.0–3.3 kbp) that is partially single stranded. The family includes two genera: *Orthohepadnavirus* (hepatitis B virus) and *Avihepadnavirus* (duck hepatitis B virus). *Hepadnavirus* infection induces overproduction of surface proteins, which are secreted as lipoprotein particles together with virus to the blood. Empty viruslike particles consisting of excess virus envelope material are present in much greater numbers than complete virions in most individuals. *See* HEPATITIS B VIRUS.

Herpesviridae. In this family the particle is 120–200 nm in diameter. Its double-stranded linear DNA (125 to 240 kbp in size) is packed in a liquid crystalline array which fills the entire volume of the icosahedral capsid (100–125 nm). This package is surrounded by a lipid bilayer envelope and several glycoproteins. The space between the icosahedral capsid and the envelope, the tegument, is filled with at least 15 distinct polypeptides. Viral infectivity is destroyed by lipid solvents and detergents. The subfamilies of this group are Alphaherpesvirinae (herpes simplex virus group), Betaherpesvirinae (cytomegalovirus group), and Gammaherpesvirinae (lymphoproliferative virus group). Epstein-Barr (EB) virus is the cause of Burkitt's lymphoma in the malaria belt and nasopharyngeal carcinoma in the Peoples Republic of China. It is also the causative agent of infectious mononucleosis. *See* CYTOMEGALOVIRUS INFECTION; EPSTEIN-BARR VIRUS; HERPES SIMPLEX VIRUS.

Polyomaviridae. This group is characterized by a nonenveloped icosahedron (72 capsomers), 40 nm in diameter, which contains five to seven unique structural proteins. Double-stranded circular DNA (5 kbp) is complexed inside the nucleocapsid to histone proteins of host cell origin. The sole genus in this family is *Polyomavirus*. Polyomaviruses include *JC polyomavirus*, which causes the rare disease known as progressive multifocal leukoencephalopathy in humans, and Simian virus 40, which causes a similar disease in rhesus monkeys.

Papillomaviridae. This family consists of nonenveloped icosahedral virions 55 nm in diameter. The capsid is composed of 72 capsomers in skewed arrangement. The genome consists of a single molecule of circular dsDNA (6800–8400 bp). The sole genus in this family is *Papillomavirus*. Species include human papillomaviruses and bovine papillomaviruses. Human papillomavirus spreads by close contact with warts. Specific human cancers, including cervical and penile cancers, have been linked to specific types of Papillomaviruses. *See* PAPILLOMAVIRUS.

Adenoviridae. This group is characterized by a nonenveloped icosahedron of 252 capsomers, 70–90 nm in diameter, which contains at least 12 unique structural proteins. The 12 vortex capsomers (pentons) exhibit one (in the mammalian group) or two (in the avian group) filamentous projections (fibers) involved in attachment to the host cell. The genome, a double-stranded linear DNA (26–45 kbp), has been sequenced for human adenovirus. The family comprises two genera, *Mastadenovirus* (mammalian, including human) and *Aviadenovirus* (avian), each containing a large number of members. *See* ADENOVIRIDAE.

Circoviridae. This family, containing the sole genus *Circovirus*, is characterized by nonenveloped icosahedral virions. The particles range in size 12–27 nm in diameter. Virions contain circular single-stranded DNA (2.0 kbp). Species in the genus include porcine circovirus and chicken anemia virus.

Parvoviridae. These virions are nonenveloped, 18–26 nm in diameter with icosahedral symmetry with 32 capsomers. Parvoviridae are the smallest viruses, containing three or four structural proteins and a ssDNA genome of 4–6 kbp. This family includes two subfamilies, Parvovirinae and Densovirinae. The subfamily includes the genera *Parvovirus*, *Erythrovirus*, and *Dependovirus*. *Dependovirus* comprises defective viruses that require a helper virus (usually an adenovirus) for their replication. The other genera, *Parvovirus* and *Erythrovirus*, are autonomous. The subfamily Densovirinae and the genus *Dependovirus* produce, without preference, particles which contain either the positive or the negative DNA strand. Parvoviruses contain only the negative strand. Densoviruses infect arthropods only and will be discussed in a subsequent section.

Single-stranded, positive-sense RNA viruses. RNA vertebrate viruses may be either single-stranded or double-stranded. The single-stranded ones are

further subdivided into positive-strand and negative-strand RNA viruses, depending on whether the RNA contains the messenger RNA (mRNA) nucleotide sequence or its complement, respectively. The RNA genes may be located on one or several RNA molecules (nonfragmented or fragmented genomes, respectively).

The single-stranded, positive-sense RNA vertebrate viruses contain eight families: Caliciviridae, Picornaviridae, Astroviridae, Flaviviridae, Nodaviridae, Coronaviridae, Togaviridae, and Arteriviridae. In addition, a genus of HEV-like viruses is found in this group but has not yet been assigned to a family.

Caliciviridae. These icosahedral particles, 27–40 nm in diameter, exhibit 32 visible cup-shaped depressions and contain 90 dimers of the major structural protein. Although their single-stranded RNA (7.4–8.3 kb) serves as mRNA, a smaller mRNA fragment has also been detected. Caliciviruses infect a broad range of animals that includes cattle, swine, and humans. Genera in this family include *Lagovirus*, *Vesivirus*, Norwalk-like viruses, and Sapporo-like viruses. The prototype of this virus group is the vesicular exanthema of swine virus (VESV).

Picornaviridae. This family contains the genera *Enterovirus* (human polio), *Cardiovirus* (mengo), *Rhinovirus* (common cold), *Hepatovirus* (hepatitis A virus), *Parechovirus* (human parechovirus), and *Aphtovirus* (foot-and-mouth disease). The nonenveloped viral capsids are icosahedral, 22–30 nm in diameter, composed of 60 identical units, consisting of three surface proteins each. Viruses contain one single-stranded RNA molecule of 7–8.5 kb in size. A small protein, VPg, is covalently bound at its 5′ end instead of the cap structure, and is polyadenylated at the 3′ end. Viruses are insensitive to ether, chloroform, or nonionic detergents. *See* ENTEROVIRUS; PICORNAVIRIDAE; RHINOVIRUS.

Astroviridae. Virions are 28–30 nm in diameter, spherical, and nonenveloped. A five- or six-pointed star is discernible on the surface of about 10% of particles. This family contains a sole genus, *Astrovirus*. A single molecule of infectious, positive-sense ssRNA (6.8–7.9 kb in size) makes up the viral genome. A poly(A) tract is located at the 3′ end of the genome. *Astroviruses* have been detected in stool samples from humans, cats, cattle, and several other animals. Astroviruses cause acute, nonbacterial gastroenteritis in children and immunocompromised adults.

HEV-like viruses. These virions are icosahedral (27–34 nm) and nonenveloped. The genome is a positive-sense ssRNA molecule of approximately 7.2 kb, with a 3′-poly(A) tail. Hepatitis E virus is associated with outbreaks and sporadic cases of fecally transmitted acute hepatitis in tropical and subtropical countries. *See* HEPATITIS E VIRUS.

Flaviviridae. These virions are 40–60 nm in diameter, spherical in shape, and contain a lipid envelope. A sole capsid protein makes up the capsid, while two or three virus-encoded membrane proteins are found in the envelope. The genome is a positive-sense ssRNA molecule, 9.6–12.3 kb in size. Genera in this group include *Flavivirus* (yellow fever virus), *Pestivirus* (bovine viral diarrhea virus), and *Hepacivirus* (hepatitis C virus). *See* YELLOW FEVER VIRUS; HEPATITIS C VIRUS.

Nodaviridae. These virions are nonenveloped, roughly spherical in shape, 29–32 nm in diameter, with icosahedral symmetry. Two ssRNA molecules (mol wt 1.1 and 0.48×10^6, respectively) serve as mRNAs and are both required for infectivity. Genera include *Alphanodavirus* (Nodamura virus) and *Betanodavirus* (striped jack nervous necrosis virus). All members of the Alphanodaviruses are insect viruses, although serological evidence suggests that Nodamura virus also naturally infects pigs. The Betanodaviruses were isolated from juvenile marine fish, and are associated with significant problems in commercial fish hatcheries.

Coronaviridae. Genera in this family include *Coronavirus* (severe acute respiratory syndrome or SARS) and *Torovirus* (equine torovirus). The spherical, enveloped particle is 120–160 nm in diameter, and exhibits large (20 nm) spikes projecting from the surface. The nucleocapsid is helical and is surrounded by an internal core shell of approximately 65 nm. One ssRNA molecule ranges in size from 27.6 to 31 kb. Coronaviruses are known to infect a large number of animals as well as humans. *See* SARS.

Togaviridae. This group contains the genera *Alphavirus* (Western equine encephalitis virus) and *Rubivirus* (rubella virus). The icosahedral nucleocapsid (40 nm in diameter) contains the capped and polyadenylated RNA (9–11.8 kb). The envelope contains two viral glycoprotein spikes. Many of these viruses are transmitted by insects and are the cause of human diseases, including several types of encephalitis. Rubella virus causes German measles, a benign childhood disease that causes severe fetal defects in pregnant women exposed to the virus. The viruses are spherical, enveloped particles 70 nm in diameter. *See* RUBELLA.

Arteriviridae. The sole genus in this family is *Arterivirus* (equine arteritis virus). These virions are spherical with a diameter of 45–60 nm. The nucleocapsid is surrounded by a lipid envelope. The genome consists of a single copy of a linear ssRNA that ranges in length from 12.7 to 15.7 kb. The host range of arteriviruses is limited, with horses, donkeys, swine, and monkeys infected by their respective species.

Single-stranded, negative-sense RNA viruses. The nucleocapsid of negative-sense RNA animal viruses contains an RNA-dependent RNA polymerase required for the transcription of the negative strand into the positive mRNAs. Virion RNA is neither capped nor polyadenylated. The group is divided into seven families: Orthomyxoviridae, Paramyxoviridae, Filoviridae, Rhabdoviridae, Bornaviridae, Arenaviridae, and Bunyaviridae. In addition, a genus known as *Deltavirus* found in this group is not assigned to a specific family.

Orthomyxoviridae. This group contains the genera *Influenzavirus A* (human influenza type A), *B*, and *C*, as well as the genus *Thogotovirus*. These are enveloped, spherical, pleomorphic particles 80–120 nm in diameter, with helical nucleocapsids. Their fragmented genome consists of 6–8 RNA

molecules, with genome sizes ranging from 10 to 14.6 kb. Surface projections for *Influenza A virus* are formed by the two major glycoproteins HA (hemagglutinin) and NA (neuraminidase). Serological classification by the World Health Organization (WHO) is based on these two neutralization antigens. *See* INFLUENZA VIRUS.

Paramyxoviridae. This group contains the subfamilies Paramyxovirinae and Pneumovirinae. The Paramyxovirinae include the genera *Respirovirus* (Sendai virus), *Rubulavirus* (mumps), and *Morbillivirus* (measles). The Pneumovirinae includes the genera *Pneumovirus* (respiratory syncytial virus) and *Metapneumovirus* (turkey rhinotracheitis virus). The enveloped, spherical particle, 150 nm in diameter, contains a single-stranded, nonfragmented RNA of approximately 15,000 nucleotides. *See* PARAMYXOVIRUS; RESPIRATORY SYNCYTIAL VIRUS; SENDAI VIRUS.

Filoviridae. This group includes two genera, the Marburg-like viruses and the Ebola-like viruses. Virions are pleomorphic but have mainly filamentous shape, sometimes with extensive branching. Length varies greatly between particles but the diameter is uniform (about 80 nm). Virions have a lipid membrane envelope derived from the host cell plasma membrane, with surface projections of viral glycoprotein of about 10 nm in length. The genome is a nonsegmented, negative-stranded, linear RNA molecule of about 19 kb in size. The natural reservoir of filoviruses is still unknown. Human disease is fatal in 25% of Marburg virus infections and in 50–90% of Ebola virus cases. *See* EBOLA VIRUS; MARBURG VIRUS; FILOVIRIDAE.

Rhabdoviridae. This group contains the genera *Vesiculovirus* (vesicular stomatitis Indiana virus), *Lyssavirus* (rabies), *Ephemerovirus* (bovine ephemeral fever virus), *Novirhabdovirus* (infectious hematopoietic necrosis virus), *Cytorhabdovirus* (lettuce necrotic yellows virus), and *Nucleorhabdovirus* (potato yellow dwarf virus). The particles are enveloped, bullet-shaped, and measure about 100–430 nm long by 45–100 nm in diameter. The outer surface of virions is covered with projections 5–10 nm long and about 3 nm in diameter. The nucleocapsid is 30–70 nm in diameter and exhibits helical symmetry. The genome consists of a single molecule of linear, negative-sense ssRNA about 11–15 kb in size. They contain five types of structural proteins: L, G, N, P, and M. Spikes are formed by G glycoprotein, and the M (matrix) protein occupies the space between the nucleocapsid and the envelope. The RNA is complexed with the N protein to form the helical nucleocapsid. Virus infectivity is rapidly inactivated at 56°C (133°F) or through exposure to lipid solvents.

Bornaviridae. The sole genus in this family is *Bornavirus* (Borna disease virus). The virions have a spherical morphology with a diameter of 90 ± 10 nm containing an internal core of 50–60 nm and a limiting outer membrane envelope that appears to be covered with projections approximately 7 nm long. The genome consists of a single molecule of a linear, nonsegmented negative-stranded RNA of 8.9 kb in size. Horses and sheep have been regarded as the main natural hosts of Borna disease virus, and the virus can cause a fatal neurological disease in these animals.

Arenaviridae. The sole genus in this family is *Arenavirus*, typified by the lymphocytic choriomeningitis virus. Its enveloped, spherical, pleomorphic particle is 50–300 nm in diameter. In section it exhibits granular appearance, due to ribosomes of host cell origin. A dense lipid envelope is covered by club-shaped projections 8–10 nm in length. The genome consists of two ssRNA molecules, L and S, of about 7.5 kb and 3.5 kb in size. Virus infectivity is inactivated by treatment with organic solvents or by exposure to ultraviolet (UV) and gamma irradiation. The reservoir hosts for this family are almost all rodents.

Bunyaviridae. This large family has been subdivided into five genera: *Bunyavirus* (bunyamwera), *Hantavirus* (Hantaan virus), *Phlebovirus* (Rift Valley fever virus), *Nairovirus* (Dugbe virus), and *Tospovirus* (tomato spotted wilt virus; Fig. 4). The virions are generally spherical or pleomorphic, 80–120 nm in diameter, and display surface glycoprotein projections of 5–10 nm embedded in a lipid bilayered envelope approximately 5 nm thick. The envelopes are usually derived from Golgi membranes in the host cell or sometimes cellular surface membranes. The genome contains three molecules of ssRNA, designated L (large), M (medium), and S (small), which total 11–19 kb. Viruses in the genera *Bunyavirus*, *Nairovirus*, and *Phlebovirus* are capable of alternately replicating in vertebrates and arthropods. Aerosol transmission occurs in hantaviral infections, while plant-to-plant transmission by thrips (tiny winged insects in the genus *Thysanoptera*) is necessary for Tospoviruses.

Deltavirus. These virions are roughly spherical with an average diameter of 36–43 nm. An outer envelope contains lipid and three envelope proteins obtained from the coinfecting helper *hepadnavirus*. An inner nucleocapsid of 19 nm comprises the RNA genome of hepatitis delta virus (HDV) and approximately 70 copies of the only HDV-encoded protein, known as delta antigen. The genome consists of a single molecule of circular negative-sense ssRNA about 1.7 kb in length. HDV requires the presence of a helper *hepadnavirus* to provide envelope proteins. It is therefore considered a subviral satellite virus. Natural HDV infection is found only in humans with hepatitis B virus as helper virus.

Double-stranded RNA viruses. The dsRNA vertebrate viruses contain two groups, the Reoviridae and the Birnaviridae.

Reoviridae. The *Reoviridae* include the nine genera *Orthoreovirus* (mammalian orthoreovirus), *Orbivirus* (bluetongue virus), *Rotavirus* (Rotavirus A), *Coltivirus* (Colorado tick fever virus), *Aquareovirus* (Aquareovirus A), *Cypovirus* (Cypovirus 1; see Fig. 3), *Fijivirus* (Fiji disease virus), *Phytoreovirus* (wound tumor virus in plants and insects), and *Oryzavirus* (rice ragged stunt virus). (See Fig. 4 for *Fjivirus*, *Phytoreovirus*, and *Oryzavirus*.) The

nonenveloped virions within this family have icosahedral symmetry but may appear spherical in shape. They each have a capsid, which is made up of concentric protein layers organized in one, two, or three distinct capsid shells. The overall diameter of the capsid is 60–80 nm. The nine genera may be divided into two groups: one group contains those viruses in which virus particles have relatively large spikes situated at the 12 vertices of the icosahedron; the second group includes those genera with relatively smooth or almost spherical virus particles lacking surface projections. Virions from some genera can leave infected cells by budding, acquiring a membrane envelope derived from the cellular membranes; however, this envelope appears to be transient in most cases. The virions contain a genome of 10, 11, or 12 linear segments of dsRNA. The properties of the *Reoviruses* vary according to the genus. *Rotavirus*, in particular, has become recognized as a frequent cause of human diarrhea. *See* ROTAVIRUS.

DNA and RNA reverse transcribing viruses. Vertebrate viruses in this category include the families Hepadnaviridae, discussed under the DNA virus category previously, and the Retroviridae.

The Retroviridae contains the genera *Alpharetrovirus* (avian leucosis virus), *Betaretrovirus* (mouse mammary tumor virus), *Gammaretrovirus* (murine leukemia virus), *Deltaretrovirus* (bovine leukemia virus), *Epsilonretrovirus* (Walleye dermal sarcoma virus), *Lentivirus* (human immunodeficiency virus or HIV), and *Spumavirus* (chimpanzee foamy virus). These virions are spherical, enveloped, and 80–100 nm in diameter. Surface projections made of glycoprotein are about 8 nm in length. The internal core contains the viral nucleocapsid, which is apparently spherical. The viral genome consists of a dimer of linear, positive-sense, ssRNA, each monomer 7–11 kb in length. There are two envelope proteins encoded by viral genes. Other proteins include a protease and a reverse transcriptase, which is essential to the process of producing a DNA copy of the RNA genome. This copy is then integrated into the host cell chromosome in a successful infection. Retroviruses have been implicated in human as well as animal leukemias, lymphomas, sarcomas, and immunodeficiencies, such as acquired immune deficiency syndrome (AIDS). *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS).

Families of Viruses Infecting Algae, Fungi, Yeast, and Protozoa

The shapes and sizes of characteristic members of the families of viruses that infect algae, fungi, yeast, and protozoa are presented in **Fig. 2**.

DNA viruses. Viruses in this group include one family, the Phycodnaviridae, and a genus known as *Rhizidiovirus* that is not assigned to a specific family.

Phycodnaviridae. Three genera include *Chlorovirus* (*Paramecium bursaria Chlorella virus 1*), *Prasinovirus* (*Micromonas pusilla virus SP1*), *Prymnesiovirus* (*Chrysochromulina brevifilum virus CbV-PW1*), and *Phaeovirus* (*Ectocarpus siliculosus virus 1*). These virions are nonenveloped, polyhedral particles with a shell surrounding an electron-dense core. They are 130–190 nm in diameter, with flexible, hairlike appendages extending from some of the vertices. The large genome consists of a dsDNA ranging from 160 to 380 kbp in size. The phycodnaviruses are ubiquitous in fresh water or seawater, and infect fresh-water and marine algae worldwide.

Rhizidiovirus. These virions are isometric, 60 nm in diameter. The genome consists of a single molecule of dsDNA. *Rhizidiomyces virus* infects the fungus *Rhizidiomyces*. The virus appears to be transmitted in a latent form in the zoospores of the fungus. No other viruses have been identified in this genus.

Double-stranded RNA viruses. RNA viruses infecting algae, fungi, yeast, and protozoa include dsRNA viruses, ssRNA viruses, and ssRNA viruses that use reverse transcriptase. The double-stranded RNA viruses include three families: Totiviridae, Partitiviridae, and Hypoviridae.

Totiviridae. Genera in this family include *Totivirus* (*Saccharomyces cerevisiae virus L-A*), *Giardiavirus* (*Giardia lamblia virus*), and *Leishmaniavirus* (*Leishmania RNA virus 1-1*). These virions have isometric symmetry and are 30–40 nm in diameter, nonenveloped, and lack surface projections. The genome consists of a single molecule of linear dsRNA, 4.6–7.0 kbp in size. These viruses are associated with latent infections of their fungal or protozoal hosts.

Partitiviridae. Genera in this family include *Partitivirus* (*Gaeumannomyces graminis virus*), *Chrysovirus* (*Penicillium chrysogenum virus*), *Alphacryptovirus* (White clover cryptic virus 1), and *Betacryptovirus* (White clover cryptic virus 2) [see Fig. 4 for *Alphacryptovirus* and *Betacryptovirus*]. These virions are isometric, nonenveloped, 30–40 nm in diameter, and contain two unrelated linear dsRNA segments (1.4–3.0 kbp in size). These viruses are associated with latent infections of their fungal and plant hosts.

Hypoviridae. The sole genus in this family is *Hypovirus* (*Cryphonectria hypovirus 1*). No true virions are associated with members of this family. Pleomorphic vesicles 50–80 nm in diameter, lacking any detectable viral structural proteins but containing dsRNA (9–13 kbp in size), are the only virus-associated particles that can be isolated from infected fungal tissue. Members infect the chestnut blight fungus, *Cryphonectria parasitica*.

Single-stranded RNA viruses. This group contains two families: Narnaviridae and Barnaviridae.

Narnaviridae. Two genera are found in this family: *Narnavirus* (*Saccharomyces cerevisiae narnavirus 20S RNA*) and *Mitovirus* (*Cryphonectria parasitica mitovirus 1-NB631*). No true virions are found associated with members of the genus *Narnavirus*. The *Narnavirus* genomes form ribonucleoprotein complexes containing linear single-stranded RNA of 2.5-kb and RNA-dependent RNA polymerases. These viruses infect the fungus *Saccharomyces cerevisiae*. The genus *Mitovirus* contains no reported virions. A single segment of dsRNA of 2.7 kb can be isolated from mitochondria of infected isolates. The virus

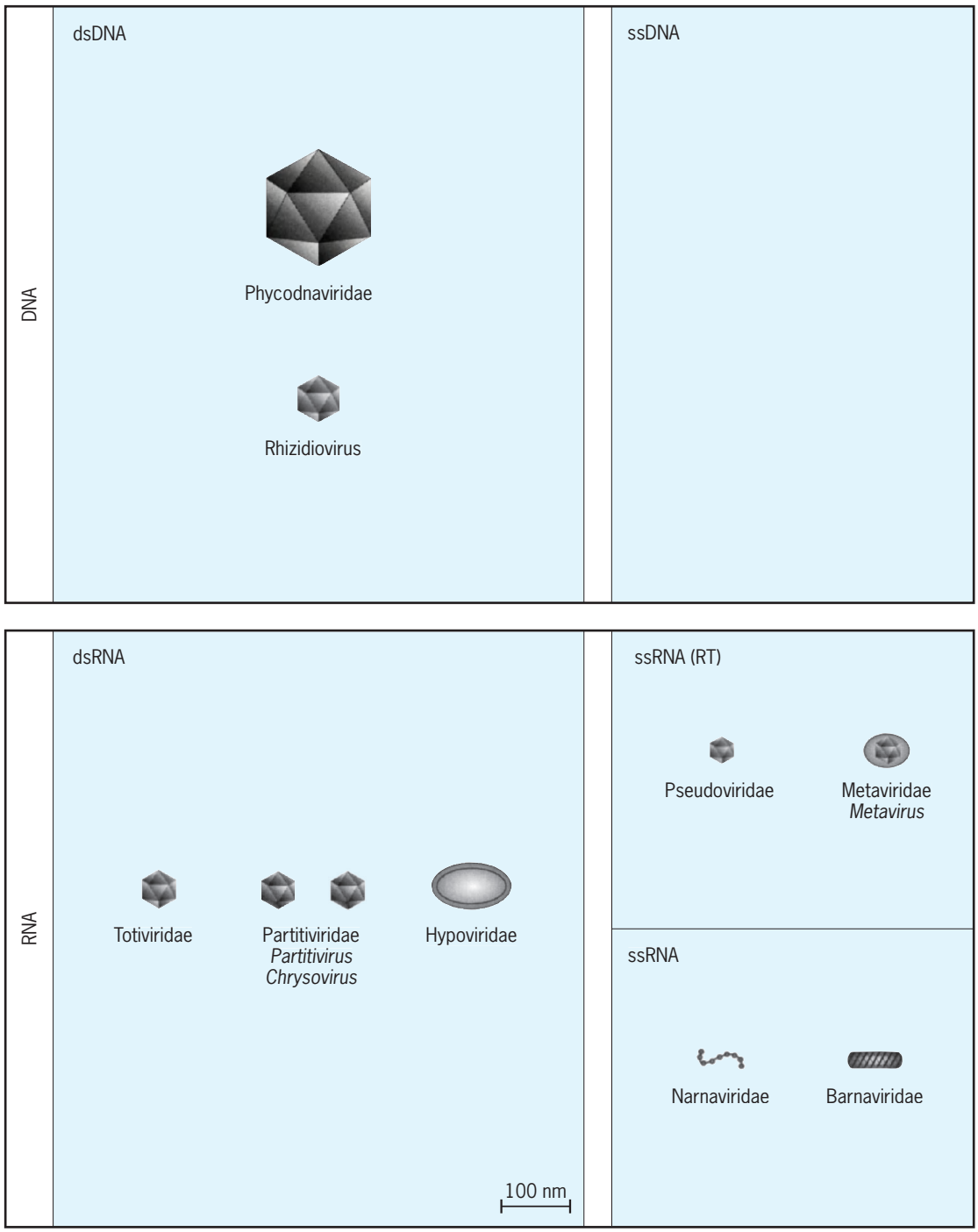

Fig. 2. Families and genera of viruses infecting algae, fungi, yeast, and protozoa. Individual frames separate taxa of viruses containing double-stranded and single-stranded genomes. Large boxes separate taxa of viruses having DNA or RNA genomes. All diagrams have been drawn approximately to the same scale to provide an indication of the relative sizes of the viruses. (*From M. H. V. van Regenmortel et al., eds., Virus Taxonomy: 7th Report of the International Committee on Taxonomy of Viruses, eds., p. 31, copyright 2000, with permission from Elsevier*)

infects members of the chestnut blight fungus, *Cryphonectria parasitica*.

Barnaviridae. The sole genus of this family is *Barnavirus* (mushroom bacilliform virus). Virions are nonenveloped, rod-shaped, and lack surface projections. Virions range 18–20 nm in width and 48–53 nm in length. The genome consists of a single linear molecule of a positive-sense ssRNA, 4.0 kb in size. The virus infects the common cultivated botton mushroom.

Single-stranded RNA viruses using reverse transcriptase. Viruses in this category include two families. *Pseudoviridae* and *Mataviridae*.

Pseudoviridae. Two genera are found in the family Pseudoviridae: *Pseudovirus* (*Saccharomyces cerevisiae Ty1 virus*) and *Hemivirus* (*Drosophila melanogaster copia virus*). The morphology of virus particles of this family is poorly known and variable. Both viruses make similar-looking particles, but those of *Saccharomyces cerevisiae Ty1*

virus are cytoplasmic whereas those of *Drosophila melanogaster copia virus* are nuclear. The major virion RNA molecule consists of a single strand of RNA of approximately 5–6 kb. A viral reverse transcriptase is used to copy the RNA and make a dsDNA "provirus," which is integrated into the host genome.

Metaviridae. Two genera are found in the family Metaviridae: *Metavirus* (*Saccharomyces cerevisiae Ty3 virus*) and *Errantivirus* (*Drosophila melanogaster gypsy virus*; **Fig. 3**). Morphology of particles is poorly characterized. Members include species which produce primarily or exclusively intracellular particles. Extracellular particles are enveloped with ovoid cores. Membranes are found in some members, apparently derived from the host cell. These viruses contain a reverse transcriptase as well as the positive strand RNAs.

Families and Genera of Viruses Infecting Invertebrates

The families of invertebrate viruses and the shapes and sizes of their characteristic members are presented in Fig. 3.

Double-stranded DNA viruses. These contain five families: Poxviridae, Baculoviridae, Iridoviridae, Polydnaviridae, and Ascoviridae.

Poxviridae. These viruses were previously described under vertebrate DNA viruses. However, the members of the subfamily Entomopoxvirinae strictly infect insects. Genera in this subfamily include *Entomopoxvirus A*, *B*, and *C*.

Baculoviridae. These are insect viruses. Two genera include *Nucleopolyhedrovirus* (*Autographa californica multiple nucleopolyhedrovirus*) and *Granulovirus* (*Cydia pomonella granulovirus*). The virions contain enveloped rod-shaped nucleocapsids. Nucleocapsid dimensions range 30–60 nm in diameter and 250–300 nm in length. The circular supercoiled dsDNA ranges 80–180 kbp in size.

Iridoviridae. These viruses were previously described under vertebrate DNA viruses. However, two genera of this family, *Iridovirus* and *Chloriridovirus*, strictly infect insects.

Polydnaviridae. These virions consist of nucleocapsids of uniform size (approximately 85 nm by 330 nm). A two-unit membrane envelope surrounds the ellipsoid nucleocapsid. Genomes consist of multiple dsDNA molecules of variable size, ranging from 2.0 to more than 31 kbp. The genera in this family are *Ichnovirus* (*Campoletis sonorensis ichnovirus*) and *Bracovirus* (*Cotesia melanoscela bracovirus*). These viruses have only been isolated from wasps; *Polydnaviruses* are unique among viruses in having obligate associations with parasitic wasps. The viral genome is inserted as a provirus into the host genome, from which new genome molecules are replicated.

Ascoviridae. These virions are bacilliform in shape, measuring about 130 nm in diameter by 200–400 nm in length. An envelope surrounds the inner particle. The sole genus in this family is *Ascovirus* (*Spodopter frugiperda ascovirus 1a*). These viruses cause disease in insects only, and are found primarily in lepidopterous larvae.

Single-stranded DNA viruses. This group includes only two families, the Parvoviridae and Circonviridae, which were previously described under vertebrate DNA viruses. The subfamily Densovirinae includes three genera responsible for infecting invertebrates: *Densovirus*, *Iteravirus*, and *Brevidensovirus*.

Double-stranded RNA viruses. These include two families: Reoviridae and Birnaviridae.

Reoviridae. These viruses were previously described under vertebrate RNA viruses. Only one genus in this family infects invertebrates: *Cypovirus* (*Cypovirus 1*). *Cypoviruses* have only been isolated from arthropods.

Birnaviridae. These viruses were previously described under vertebrate RNA viruses. Only one genus in this family infects invertebrates: *Entomobirnavirus* (*Drosophila X virus*).

Single-stranded, negative-sense RNA viruses. These include two families: Rhabdoviridae and Bunyaviridae.

Rhabdoviridae. These viruses were previously described under vertebrate RNA viruses. A large number of rhabdoviruses have not been assigned to an existing genus and there are a number of ungrouped viruses. An example is *Sigma virus*, which infects *Drosophila* species, conferring CO_2 sensitivity to infected insects.

Bunyaviridae. These viruses were previously described under vertebrate RNA viruses. They are capable of alternately replicating in vertebrates and arthropods, causing disease in the vertebrates while having little or no effect on the arthropods (ticks, mosquitoes, and other arthropod disease vectors).

Single-stranded, positive-sense RNA viruses. These include five families: Togaviridae, Flaviviridae, Picornaviridae, Tetraviridae, and Nodaviridae. They also contains the unassigned genus referred to as cricket paralysis like viruses.

Togaviridae. These viruses were previously described under vertebrate RNA viruses. The viruses in the genus *Alphavirus* are transmitted biologically between vertebrates by mosquitoes and other blood-feeding arthropods. Typically, infection in the arthropods becomes permanent.

Flaviviridae. These viruses were previously described under vertebrate RNA viruses. The viruses in the genus *Flavivirus* are usually transmitted biologically between vertebrates by mosquitoes and other blood-feeding arthropods.

Picornaviridae. These viruses were previously described under vertebrate RNA viruses. Transmission by arthropod vectors is not known, although *Encephalomyocarditis virus* has been isolated from mosquitoes and ticks.

Tetraviridae. Two genera are found in this family: *Betatetravirus* and *Omegatetravirus*. These virions are nonenveloped, roughly spherical, about 40 nm in diameter, and contain a single, positive-sense ssRNA segment of about 6.5 kb. All virus species are isolated from *Lepidoptera* species (moths and butterflies).

Nodaviridae. These viruses were previously described under vertebrate RNA viruses. All species of

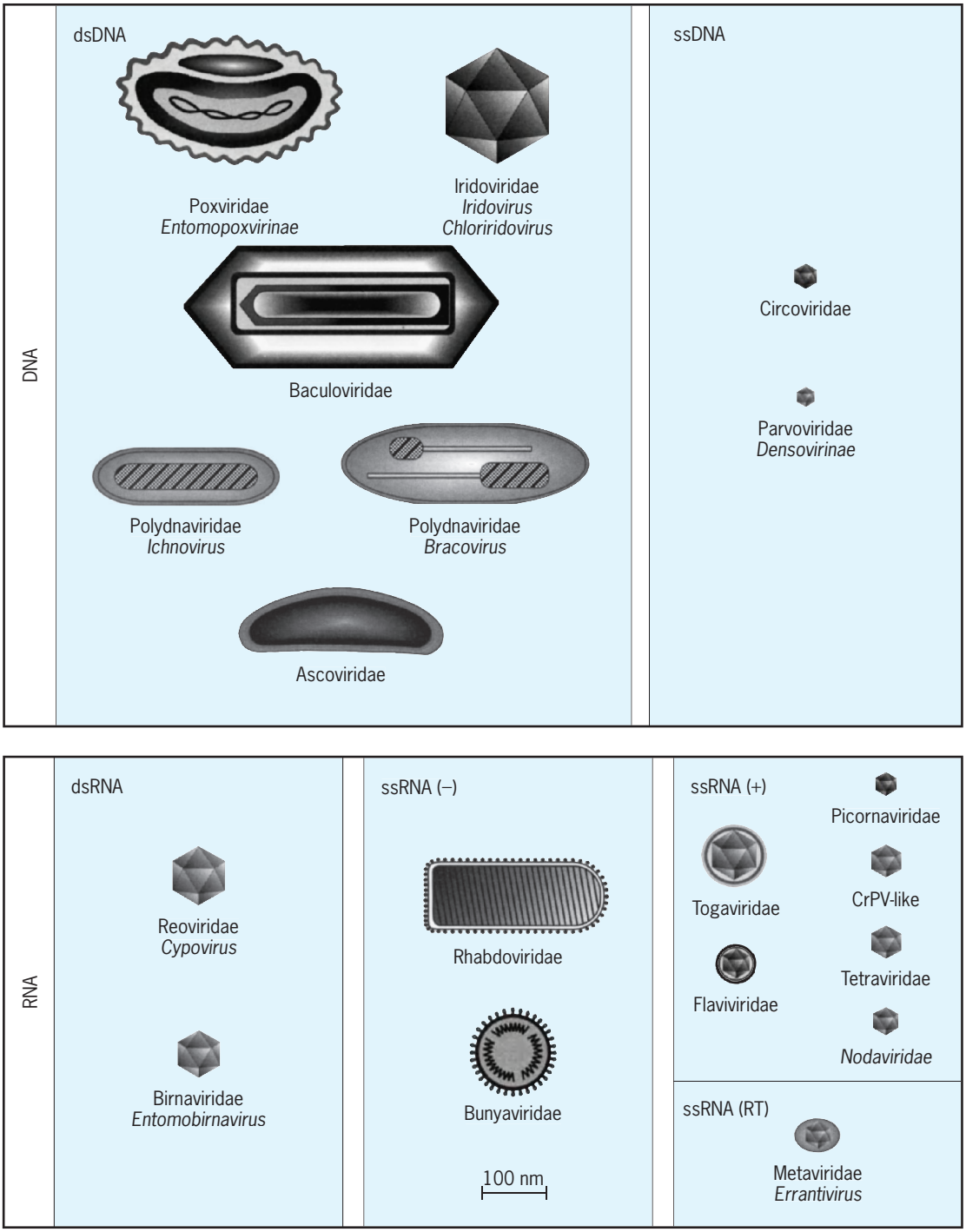

Fig. 3. Families and genera of viruses infecting invertebrates. Individual frames separate taxa of viruses containing double-stranded and single-stranded genomes. Large boxes separate taxa of viruses having DNA or RNA genomes. All diagrams have been drawn approximately to the same scale to provide an indication of the relative sizes of the viruses. (*From M. H. V. van Regenmortel et al., eds., Virus Taxonomy: 7th Report of the International Committee on Taxonomy of Viruses, p. 32, copyright 2000, with permission from Elsevier*)

the *Alphanodaviruses* were isolated in nature from insects.

Cricket paralysis-like viruses. This genus is not yet assigned to a family. These virions are roughly spherical with a diameter of 30 nm and no envelope. The virions exhibit icosahedral symmetry. The genome consists of a single molecule of linear, positive-sense ssRNA of 9–10 kb in size. All members of this genus have been isolated from invertebrate species.

Single-stranded RNA viruses using reverse transcriptase. This group contains only the family Metaviridae. These viruses were previously described under RNA viruses infecting algae, fungi, yeast, and protozoa. Members of the genus *Errantivirus* are known to infect species of *Drosophila*.

Families and Genera of Viruses Infecting Plants

The families of plant viruses and the shapes and sizes of their characteristic members are presented

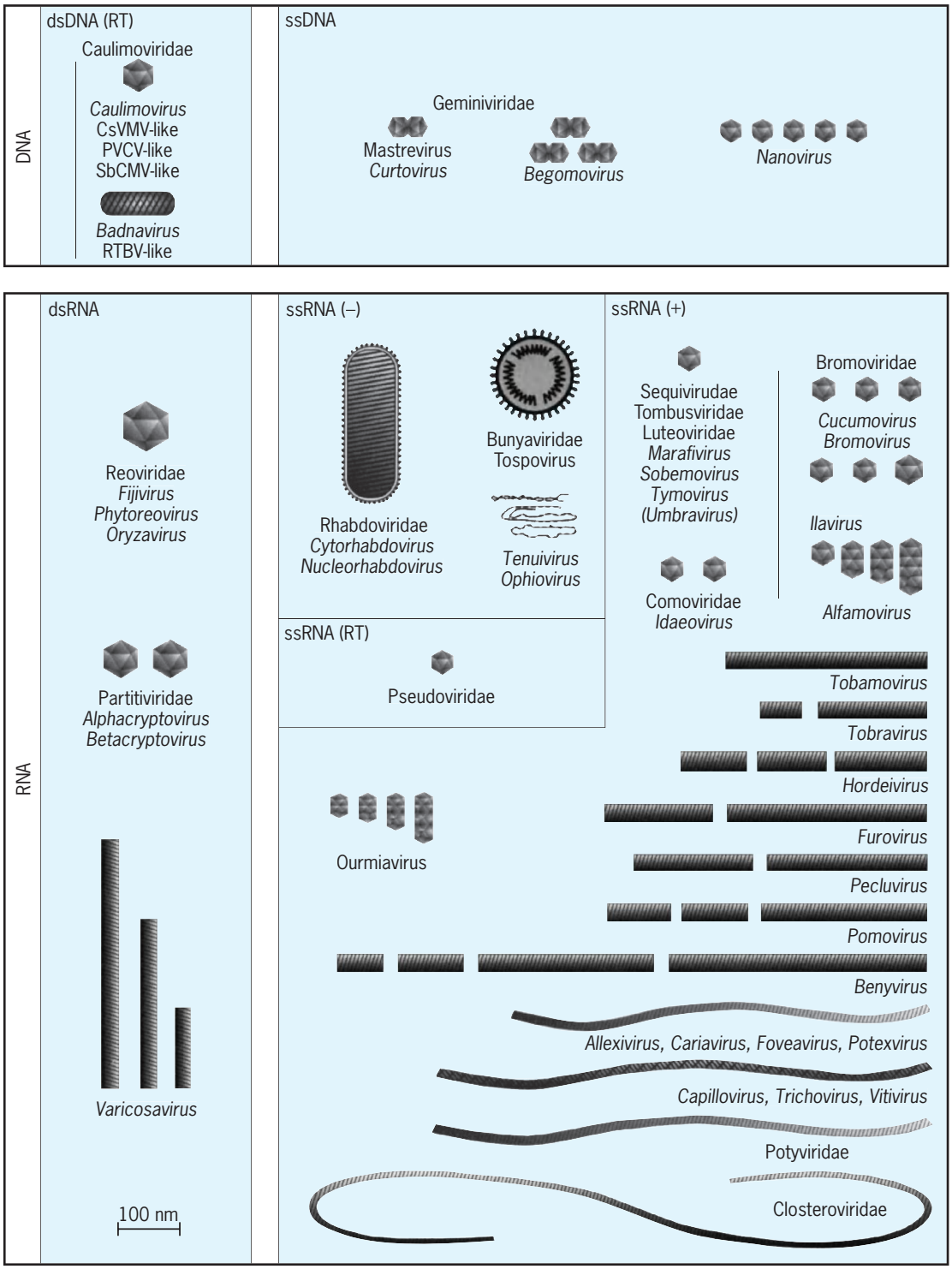

Fig. 4. Families and genera of viruses infecting plants. Individual frames separate taxa of viruses containing double-stranded and single-stranded genomes. Large boxes separate taxa of viruses having DNA or RNA genomes. All diagrams have been drawn approximately to the same scale to provide an indication of the relative sizes of the viruses. (*From M. H. V. van Regenmortel et al., eds., Virus Taxonomy: 7th Report of the International Committee on Taxonomy of Viruses, p. 33, copyright 2000, with permission from Elsevier*)

in **Fig. 4**. A number of plant viruses are placed into groups, rather than families. *See* PLANT VIRUSES AND VIROIDS.

Double-stranded DNA viruses. These include only the family Caulimoviridae. These virions are either bacilliform or isometric depending on the genus. There is no envelope. The genome consists of a single molecule of circular dsDNA of 7.2–8.1 kbp. Virions contain a reverse transcriptase. Genera in this family include *Caulimovirus*, Petunia vein clearing-like viruses, Soybean chlorotic mottle-like viruses, Cassava vein mosaic-like viruses, *Badnavirus*, and Rice tungro *bacilliform*-like viruses. The host ranges of most virus species are narrow. Many virus species are spread by vegetative propagation.

Single-stranded DNA viruses. These include one family, the Geminiviridae, and one unassigned genus called *Nanovirus*.

Geminiviridae. These virions apparently consist of two incomplete icosahedra with a total of 22 pentameric capsomers. The genome consists of a single molecule of circular ssDNA, 2.5–3.0 kb in size. Genera include *Mastrevirus*, *Curtovirus*, and *Begomovirus*. *Geminiviridae* are transmitted in nature by arthropods, in most cases by a single species.

Nanovirus. This genus is not yet assigned to a family. These virions are 17–20 nm in diameter, icosahedral, and lack envelopes. The genome is composed of several species of circular ssDNA approximately 1 kb in size. These viruses infect plants, including a range of leguminous species.

Double-stranded RNA viruses. These include two families, Reoviridae and Partitiviridae and the unassigned genus *Varicosavirus*.

Reoviridae. These viruses were previously described under vertebrate RNA viruses. Three genera in this family infect plants: *Fijivirus*, *Phytoreovirus*, and *Oryzavirus*. Viruses in these genera are typically transmitted between plants by arthropod vectors such as leafhoppers.

Partitiviridae. These viruses were previously described under RNA viruses infecting algae, fungi, yeast, and protozoa. Two genera are known to infect plants: *Alphacryptovirus* and *Betacryptovirus*. There are no known natural vectors for these viruses. The plant cryptoviruses are transmitted by ovule and by pollen to the seed embryo.

Varicosavirus. This genus is not yet assigned to a family. Virions are rod-shaped, about 320–360 nm in length, with a diameter of about 18 nm. The genome consists of dsRNA of two sizes, 7 kbp and 6.5 kbp. Viruses are transmitted by fungi of the genus *Olpidium* and are therefore soil- or waterborne. The type species for this genus is Lettuce big-vein virus.

Single-stranded negative-sense RNA viruses. These include two families, Rhabdoviridae and Bunyaviridae, plus two unassigned genera.

Rhabdoviridae. These viruses were previously described under vertebrate RNA viruses. Two genera in this family infect plants: *Cytorhabdovirus* and *Nucleorhabdovirus*. The viruses are primarily distinguished on the basis of the sites of virus maturation (cytoplasm: *Cytorhabdovirus*; nucleus: *Nucleorhabdovirus*).

Bunyaviridae. These viruses were previously described under vertebrate RNA viruses. One genus in this family is known to infect plants: *Tospovirus* (tomato spotted wilt virus). These viruses are known to be transmitted by thrips.

Tenuivirus. This genus is not yet assigned to a family. Virions have a thin filamentous shape, consisting of nucleocapsids 3–10 nm in diameter, with lengths proportional to the size of their RNA. The ssRNA genome consists of four or more segments, ranging in size from 11.9 to 9 kb. Plant hosts of tenuiviruses are all in the family Graminae.

Ophiovirus. This genus is not yet assigned to a family. Virions are naked filamentous nucleocapsids about 3 nm in diameter, with the shortest length about 760 nm. The genome consists of ssRNA of 11–12 kb in size. The natural hosts are dicotyledonous for two species, and tulip for the third species.

Single-stranded positive-sense RNA viruses. These include many families and unassigned genera, as follows.

Sequiviridae. Particles are isometric, about 30 nm in diameter. The genome consists of one molecule of infective, positive-sense RNA, 9–12 kb in size. Two genera are found in this family: *Sequivirus* (parsnip yellow fleck virus) and *Waikavirus* (rice tungro spherical virus). Transmission of these viruses is by aphids or leafhoppers; transmission is thought be dependent on a self-encoded helper protein.

Comoviridae. Virions are nonenveloped, 28–30 nm in diameter, and exhibit icosahedral symmetry. Genomes consist of two molecules of linear positive-sense ssRNA. Genera include *Comovirus* (cowpea mosaic virus), *Fabavirus* (broad bean wilt virus 1), and *Nepovirus* (tobacco ringspot virus). All members of the family have biological vectors: comoviruses are transmitted by beetles, fabaviruses by aphids, and many nepoviruses are transmitted by nematodes.

Potyviridae. Virions are flexible filaments (helical symmetry) with no envelope and are 11–15 nm in diameter. A single molecule of positive-sense ssRNA of 8.5–10 kb in size makes up the genome. Genera in this family include *Potyvirus* (potato virus Y), *Ipomovirus* (sweet potato mild mottle virus), *Macluravirus* (Maclura mosaic virus), *Rymovirus* (ryegrass mosaic virus), *Tritimovirus* (wheat streak mosaic virus), and *Bymovirus* (barley yellow mosaic virus). All members of the family *Potyviridae* form cytoplasmic cylindrical inclusion bodies during infection. Some members have a narrow host range, while most members infect an intermediate number of plants. Transmission to most hosts is accomplished by mechanical inoculation.

Sobemovirus. This genus is not yet assigned to a family. Virions are about 30 nm in diameter and exhibit icosahedral symmetry. The genome consists of a single molecule of positive-sense ssRNA, approximately 4.1–4.5 kb in size. The representative virus for this genus is southern bean mosaic virus. Sobemoviruses infect both monocotyledonous and dicotyledonous plants, but the natural host range of each species is narrow.

Marafivirus. This genus is not yet assigned to a family. These virions exhibit icosahedral symmetry, are 28–32 nm in diameter, do not have an envelope, and contain one molecule of linear, positive-sense ssRNA, about 6.5 kb in size. The representative virus for this genus is *Maize rayado fino virus*.

Luteoviridae. Virions are 25 to 30 nm in diameter, exhibit icosahedral symmetry, and have no envelope. The genome consists of a single molecule of linear, positive-sense ssRNA of about 5800 nts. Three genera are found in this family: *Luteovirus* (barley yellow dwarf virus), *Polerovirus* (potato leafroll virus), and *Enamovirus* (pea enation mosaic virus-1). Transmission is by specific aphid vectors.

Umbravirus. This genus is not yet assigned to a family. Umbraviruses do not form conventional virus

particles. Evidence suggests that the infective RNA is protected in lipid-containing structures. The representative virus is carrot mottle virus.

Tombusviridae. This is a very large family of plant viruses, with eight genera identified: *Aureusvirus*, *Avenavirus*, *Carmovirus*, *Dianthovirus*, *Machlomovirus*, *Necrovirus*, *Panicovirus*, and *Tombusvirus*. Virions exhibit an icosahedral symmetry with virions from the genera *Aureusvirus*, *Avenavirus*, *Carmovirus*, *Dianthovirus*, *Tombusvirus* having a rounded outline, a diameter of about 32–35 nm, and a granular surface due to the presence of 90 projections known as P domains. Virions in the genera *Machlomovirus*, *Necrovirus*, and *Panicovirus* lack the protruding domain and appear smooth. The genome consists of a single molecule of positive-sense, linear ssRNA ranging in size from 3.7 to 4.7 kb. The exceptions are the *Dianthovirus* virions, which contain two genomic RNAs. A characteristic virus representative is the *Tomato bushy stunt virus* (genus: *Tombusvirus*).

Tobamovirus. This genus is not yet assigned to a family. The representative for this genus is the tobacco mosaic virus, which is characterized by particles with a rigid helical rod, 18 nm in diameter and 300–310 nm in length. The genome consists of a single molecule of positive-sense linear ssRNA, 6.3–6.6 kb in size.

Tobravirus. This genus is not yet assigned to a family. Tobacco rattle virus is the best-known member of this group, which is characterized by two types of nonenveloped, rigid helical rods 21–23 nm in diameter either 180–215 nm or 46–115 nm in length, depending on the isolate. The genome consists of two molecules of linear positive-sense ssRNA; the first is 6.8 kb and the second ranges in size from 1.8 kb to about 4.5 kb.

Hordeivirus. This genus is not yet assigned to a family. Virions in this group are nonenveloped, elongated, and rigid; they are helically symmetric and about 20 nm in diameter and 110–150 nm in length. The genome contains three molecules of positive-sense ssRNA, each approximately 3–4 kb in size. The representative virus for this genus is the barley stripe mosaic virus.

Furovirus. This genus is not yet assigned to a family. Virions in this group are nonenveloped, hollow rods; they are helically symmetric and about 20 nm in diameter and 140–160 nm or 260–300 nm in length. The genome contains two molecules of linear, positive-sense ssRNA, with the first molecule about 6–7 kb and the second molecule about 3.5–3.6 kb in size. The representative virus for this genus is the soil-borne wheat mosaic virus.

Pomovirus. This genus is not yet assigned to a family. Virions in this group are nonenveloped, rod-shaped particles; they are helically symmetric and about 20 nm in diameter, with predominant lengths of 65–80, 150–160, and 290–310 nm. The genome contains three molecules of linear, positive-sense ssRNA, with the first molecule about 6 kb, the second about 3–3.5 kb, and the third 2.5–3 kb in size. The representative virus for this genus is the potato mop-top virus.

Pecluvirus. This genus is not yet assigned to a family. Virions in this group are nonenveloped, rod-shaped particles; they are helically symmetric and about 21 nm in diameter, with two predominant lengths of 190 and 245 nm. The genome contains two molecules of linear, positive-sense ssRNA, with the first molecule about 5900 nts and the second molecule about 4500 nts. The representative virus for this genus is the peanut clump virus.

Benyvirus. This genus is not yet assigned to a family. Virions in this group are nonenveloped, rod-shaped particles; they are helically symmetric and about 20 nm in diameter, with predominant lengths of about 85, 100, 265, and 390 nm. The genome contains four molecules of linear, positive-sense ssRNA, ranging in size from 1.3 kb to 6.7 kb. The representative virus for this genus is the beet necrotic yellow vein virus.

Bromoviridae. Virions in this group are either spherical with icosahedral symmetry (diameter 26–35 nm) or bacilliform (diameters of 18–26 nm and lengths from 30 to 85 nm). Genomes consist of three linear, positive-sense ssRNAs (total genome length is approximately 8 kb). Genera in this group include the *Alfamovirus* (alfalfa mosaic virus), *Bromovirus* (brome mosaic virus), *Cucumovirus* (cucumber mosaic virus), *Ilarvirus* (tobacco streak virus), and *Oleavirus* (olive latent virus 2). The natural host range of the viruses ranges from very narrow (genus *Bromovirus*) to extremely broad (genus *Cucumovirus*). Transmission predominantly occurs by insects, usually as a mechanical vector and not a persistent infection of the insect. Several viruses in this family are responsible for major disease epidemics in crop plants.

Ourmiavirus. This genus is not yet assigned to a family. Virions in this group constitute a series of particles with conical ends and cylindrical bodies 18 nm in diameter. There is no envelope. The genome contains three segments of linear, positive-sense ssRNA. The representative virus for this genus is the *Ourmia melon virus*.

Idaeovirus. This genus is not yet assigned to a family. The virions are isometric, about 33 nm in diameter, and not enveloped. The genome contains three segments of linear, positive-sense ssRNA of about 5.5 kb, 2.2 kb, and 1 kb in size. The representative virus for this genus is the raspberry bushy dwarf virus.

Closteroviridae. Virions in this group are very flexuous filaments with helical symmetry, about 12 nm in diameter, with the length varying according to the genus and/or the individual species. The genome consists of a single molecule of linear, positive-sense ssRNA ranging in size from 15.3 to 19.3 kb. Genera include *Closterovirus* (beet yellows virus) and *Crinivirus* (lettuce infectious yellows virus). The host ranges of individual virus species are restricted. Natural vectors include aphids and whiteflies.

Capillovirus. This genus is not yet assigned to a family. The virions are nonenveloped flexuous filaments with helical symmetry, about 12 nm in diameter and 640–700 nm in length. The genome contains linear, positive-sense ssRNA of 6.5–7.4 kb in size. The

representative virus for this genus is the apple stem grooving virus.

Trichovirus. This genus is not yet assigned to a family. These virions are nonenveloped flexuous filaments with helical symmetry, about 12 nm in diameter and 640–760 nm in length. The genome contains a single linear, positive-sense ssRNA of 7.5 kb in size. The representative virus for this genus is the apple chlorotic leaf spot virus.

Vitivirus. This genus is not yet assigned to a family. Virions are nonenveloped flexuous filaments with helical symmetry, about 12 nm in diameter and 725–825 nm in length. The genome contains a single linear, positive-sense ssRNA of 7.5 kb in size. The representative virus for this genus is the grapevine virus A virus.

Tymovirus. This genus is not yet assigned to a family. Virions are nonenveloped, icosahedral particles with a diameter of about 30 nm. The genome contains a single linear, positive-sense ssRNA of 6.3 kb in size. The representative virus for this genus is the turnip yellow mosaic virus.

Carlavirus. This genus is not yet assigned to a family. These virions are nonenveloped flexuous filaments with helical symmetry, 12–15 nm in diameter and 610–700 nm in length. The genome contains a single linear, positive-sense ssRNA of 7.4–7.7 kb in size. The representative virus for this genus is the carnation latent virus.

Potexvirus. This genus is not yet assigned to a family. These virions are nonenveloped flexuous filaments with helical symmetry, 13 nm in diameter and 470–580 nm in length. The genome contains a single linear, positive-sense ssRNA of 5.9–7.0 kb in size. The representative virus for this genus is the potato virus X.

Allexivirus. This genus is not yet assigned to a family. Virions in this group are nonenveloped highly flexible filaments, 12 nm in diameter and 800 nm in length. The genome contains a single linear, positive-sense ssRNA of 9.0 kb in size. The representative virus for this genus is the shallot virus X.

Foveavirus. This genus is not yet assigned to a family. These virions are nonenveloped flexuous filaments with helical symmetry, 12 nm in diameter and 800 nm in length. The genome contains a single linear, positive-sense ssRNA of 8.4–9.3 kb in size. The representative virus for this genus is the apple stem pitting virus.

Single-stranded RNA viruses using reverse transcriptase. This group contains only one family, the Pseudoviridae. These viruses were previously described under RNA viruses infecting algae, fungi, yeast, and protozoa. Members of this family are commonly referred to as LTR retrotransposons, and as such form an intrinsic and significant part of the genome of many plants.

Families and Genera of Viruses Infecting Bacteria

The families of bacteria, *Mycoplasma*, and Archeae viruses and the shapes and sizes of their characteristic members are presented in **Fig. 5**. Bacterial viruses are also known as bacteriophages or phages. They may be tailed or nontailed. *See* BACTERIOPHAGE.

Double-stranded DNA viruses. These include nine families and one unassigned genus.

Myoviridae. These virions have contractile tails with a central core surrounded by a helical contractile sheath. The head is elongated (or icosahedral), 110 by 80 nm, and the tail measures 113 by 16 nm. Particles contain at least 21 structural proteins, two or three inside the head. Many enzymatic activities are present. The tail has a collar, base plate, six fibers, six spikes, and a central tube. Fibers initiate attachment events. Genomes range in size 34–160 kbp. Genera include T4-like viruses, P1-like viruses, P2-like viruses, Mu-like viruses, SPO1-like viruses, and ϕ–H-like viruses (which infect Archeae). These viruses infect bacteria, with the exception of the ϕ–H-like viruses.

Siphoviridae. These virions have long, noncontractile, thin tails, 65–570 nm long and 7–10 nm in diameter, which are often flexible. Tails are built of stacked disks of six subunits. Genera in this family include λ-like viruses, T1-like viruses, T5-like viruses, L5-like viruses, c2-like viruses, and ψ M1-like viruses. Heads are primarily icosahedral, and genomes consist of dsDNA of varying sizes. These viruses infect bacterial species including *Enterobacteria*, *Mycobacterium*, *Lactococcus*, and the archaea species *Methanobacterium*.

Podoviridae. These virions have short, noncontractile tails about 20 × 8 nm in dimension. Genera include the T7-like viruses, P22-like viruses, and the ϕ29-like viruses. Heads are assembled first and tail subunits are added to the completed head. Phage heads are icosahedral. The genomes contain dsDNA of varying sizes. These viruses infect a wide number of bacterial species, including *Enterobacteria*, *Pseudomonas*, *Vibrio*, and *Caulobacter*.

Tectiviridae. Virus particles in this family are nonenveloped and icosahedral, 63 nm in diameter, and have 20-nm-long spikes at the 12 vertices. There are two shells, the outer rigid and 3 nm thick, while the inner is 6 nm thick, containing a lipoprotein vesicle. There are 15 virus-specific proteins associated with the lipoprotein vesicle. The genome consists of linear dsDNA of 14.7–15.7 kbp in size. During attachment to pili of gram-negative bacteria, following the ejection of nucleic acid, a transient tail of 60 nm length develops. The well- known *Enterobacteria phage PRDI* is in this group.

Corticoviridae. The virus particles are icosahedral and 60 nm in diameter, with brushlike spikes at the 12 vertices. The particle contains circular, dsDNA about 9 kbp in size. The virion also contains the enzyme transcriptase. A representative virus is *Alteromonas phage PM2.*

Plasmaviridae. This family of enveloped DNA phages includes the *Acholeplasma phage L2*. The particles are rounded, pleomorphic (50–125 nm in diameter), and show a small densely stained center. The double-stranded, circular DNA of the genome is 1.2 kbp in size. Virus is released by budding from *Acholeplasma* host cells, without killing the host.

Lipothrixviridae. These virions are rigid rods with a helical core, 410 nm long and 38 nm in diameter, with protrusions arising asymmetrically from both

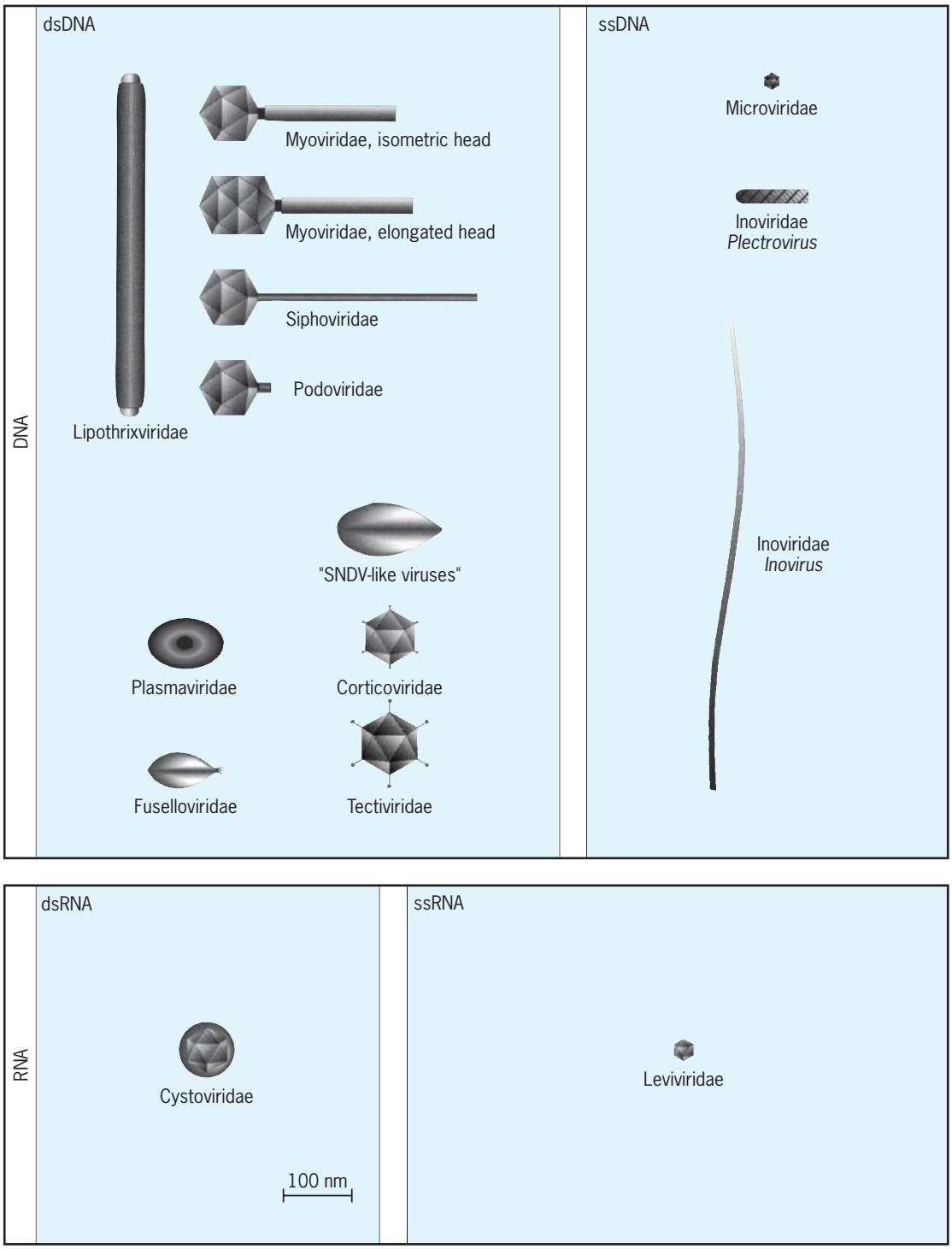

Fig. 5. Families and genera of viruses infecting bacteria. Individual frames separate taxa of viruses containing double-stranded and single-stranded genomes. Large boxes separate taxa of viruses having DNA or RNA genomes. All diagrams have been drawn approximately to the same scale to provide an indication of the relative sizes of the viruses. (*From M. H. V. van Regenmortel et al., eds., Virus Taxonomy: 7th Report of the International Committee on Taxonomy of Viruses, p. 34, copyright 2000, with permission from Elsevier*)

ends. The genome consists of one molecule of linear dsDNA of 15.9 kbp. A representative virus is *Thermoproteus virus 1*.

Rudiviridae. These virions are nonenveloped stiff rods, 23 nm in diameter by 500–780 nm in length. The genome is linear dsDNA, 33–36 kbp in size. A representative virus is *Sulfolobus virus SIRV-1*.

Fuselloviridae. These virions are lemon-shaped (60 by 100 nm), slightly flexible in appearance with short tail fibers attached to one pole. The genome contains circular, positively supercoiled dsDNA of 15.4 kbp in size. A representative virus is *Sulfolobus virus 1*.

Sulfolobus SNDV-like viruses. This genus is not yet assigned to a family. These virions are droplet-shaped, 80 by 180 nm. The genome is covalently closed circular dsDNA about 20 kbp in size. A representative virus is Sulfolobus virus SNDV.

Single-stranded DNA viruses. These include two families, Inoviridae and Microviridae.

Inoviridae. This family is characterized by nonenveloped rod-shaped phages which contain a single-stranded circular DNA genome within a cylindrical protein shell. Lengths vary from 700 nm to 2000 nm, although the diameter remains constant at about 7 nm. The genome consists of one molecule of infectious, circular positive-sense ssDNA. Genera in this family include *Inovirus* (*Enterobacteria phage M13*) and *Plectrovirus* (*Acholeplasma phage MV-L51*). Members of this family infect *Enterobacteria*, *Pseudomonas*, *Vibrio*, *Xanthomonas*, and *Acheloplasma*.

Microviridae. Virions exhibit icosahedral symmetry with projections at each of the 12 vertices. No envelope is present. Diameter of particles range from 26 to 32 nm. The genome for this family consists of one molecule of circular, positive-sense ssDNA, ranging in size from 4.4 to 6.0 kbp. Genera include *Microvirus*, *Spiromicrovirus*, *Bdellomicrovirus*, and *Chlamydiamicrovirus*. A representative virus for this family is *Enterobacteria phage ϕ X 174*. Microviridae infect *Enterobacteria*, *Bdellovibrio*, *Spiroplasma melliferum*, and *Chlamydia psittacii*.

Double-stranded RNA viruses. These include only the family Cystoviridae. The virions are enveloped, spiked bacteriophages about 85 nm in diameter. The envelope surrounds an icosahedral nucleocapsid (58 nm in diameter). The genome consists of three linear dsRNA segments of 6.4, 4.0, and 2.9 kbp in size. A representative virus is *Pseudomonas phage ϕ6*, which attaches to the pili of *Pseudomonas*.

Single-stranded RNA viruses. These include only the family Leviviridae. Virions are spherical and exhibit icosahedral symmetry (diameter of 26 nm). No envelope is present. The genome consists of one molecule of positive-sense ssRNA ranging in size from 3466 to 4276 nts. Two genera are found in this family: *Levivirus* (*Enterobacteria phage MS2*) and *Allolevivirus* (*Enterobacteria phage Qβ*). These viruses attach to pili of male *Enterobacteria*, *Caulobacter*, and *Pseudomonas*.

Marcia M. Pierce; M. E. Reichmann

Bibliography. D. M. Knight and P. M. Howley, (eds.), *Fundamental Virology*, 4th ed., Lippincott, Williams & Wilkins, Phildelphia, 2001; L. M. Prescott, J. P. Harley, and D. A. Klein, *Microbiology*, 6th ed., McGraw-Hill, New York, 2005; S. Schlesinger and M. Schlesinger, Viruses, in J. Lederberg (ed.), *Encyclopedia of Microbiology*, 2d ed., vol. 4, pp. 796–810, Academic Press, San Diego, 2000; M. H. V. Van Regenmortel et al. (eds.), *Virus Taxonomy: Seventh Report of the International Committee on Taxonomy of Viruses*, Academic Press, San Diego, 2000; B. A. A. Voyles, *The Biology of Viruses*, 2d ed., McGraw-Hill, Chicago, 2002.

Virus infection, latent, persistent, slow

Initially inapparent (covert) viral infections in which an equilibrium between the virus and the host has been established. Pathology may occur periodically, or chronically at a low level. If manifested in degenerative diseases of the central nervous system, these infections may result in death. The basis of these infections is an uncharacteristically low level of viral replication which may persist throughout the normal lifetime of the host, and which may follow the recovery from a more severe bout with the virus. The acute illness may have elicited production of antibodies, and an equilibrium between virus neutralization and antibody generation is established which may produce life-long immunity.

Persistent, slow, and latent infections are not easily differentiated. Persistent infections as described above are usually accompanied by detectable formation of antibodies without host pathology. (However, in the case of measles the persistent infection may change into a slow infection culminating in the degenerative brain disease termed subacute sclerosing panencephalitis, or SSPE.) Slow infections, while similar, eventually develop pathology. In latent infections, demonstrable presence of virus and symptoms of disease may be absent for a long time, with periodic outbreaks precipitated usually at a time when the immune system has been compromised in some way. This may occur through a variety of external influences, both physical and emotional.

Mechanisms. Several mechanisms are thought to account for the generation of latent, persistent, slow infections. Attenuation of the virulence of a virus can take place not only by immunological neutralization but also by spontaneous mutation of the virus itself. The most important group of mutants pertinent to these phenomena are the temperature-sensitive mutants, which replicate optimally at temperatures below the body temperature of the host, and very inefficiently at the elevated temperatures. The presence of temperature-sensitive mutants has been demonstrated in persistent infections of foot-and-mouth disease, coxsackie, sindbis, influenza, and equine encephalitis viruses. Another reason for diminished virulence in persistent infections may be due to the generation of defective interfering particles. Defective interfering particles are deletion mutants which are unable to replicate on their own. When they enter cells which are infected with the corresponding nonmutated virus they greatly reduce the number of infectious virus particles produced within these cells and thereby limit the extent of cell damage. The presence of defective interfering particles in persistent infections has been observed in the case of sendai, sindbis, lymphocytic choriomeningitis, and Newcastle disease viruses. Partial attenuation of viruses leading to persistent infections can also be brought about by host interferon production. *See* VIRULENCE; VIRUS, DEFECTIVE; VIRUS INTERFERENCE.

The integration of viral deoxyribonucleic acid (DNA) into the cellular genome in the case of DNA viruses, or reverse transcribed ribonucleic acid (RNA) retroviruses, often results in incomplete viral expression reminiscent of lysogeny in bacterial viruses. The afflicted cells may escape destruction by the virus itself as well as detection and destruction by the immune surveillance system. The visna

Slow virus infections*

Classification	Virus	Disease [host]
HUMANS		
RNA viruses		
Paramyxovirus	Measles variant	Subacute sclerosing panencephalitis (SSPE)
Rhabdovirus	Rabies	Rabies
Retrovirus	HTLV	T-cell leukemia, AIDS?
Picornavirus	Hepatitis A	Hepatitis
DNA viruses		
Papovavirus	JC, SV40-like	Progressive multifocal leukoencephalopathy (PML)
Unclassified	Hepatitis B	Hepatitis
ANIMALS		
RNA viruses		
Retrovirus	Visna	Meningoencephalitis [sheep]
	Maedi	Pneumonitis [sheep]
	Progressive pneumonia (PPV)	Pneumonitis [sheep]
	Equine infectious anemia	Hemolytic anemia, arteritis (equine infectious arteritis, or EIA) [horse]
	Gardner agent	Lower motor neuron [feral mouse]
	Xenotrophic C type	Hemolytic anemia, systemic lupus erythematosuslike [NZB mouse]
Arenavirus	Lymphocytic choriomeningitis (LCM)	Meningitis glomerulonephritis [mouse]
Paramyxovirus	Canine distemper	Encephalitis [dog]
Togavirus	Lactate dehydrogenase (LDV)	Elevated lactate dehydrogenase; mild nephritis [mouse]
Picornavirus	Theiler's agent	Demyelination [mouse]
DNA viruses		
Parvovirus	Aleutian disease (ADV)	Arteritis, anemia, nephritis [mink]
Papovavirus	SV40	Progressive multifocal leukoencephalopathy [monkey]
UNCONVENTIONAL AGENTS		
		Creutzfeldt-Jakob disease (CJD) [human]
		Kuru [human]
		Scrapie [sheep, goat]
		Transmissible mink encephalopathy

*After A. I. Braude, C. E. Davis, and J. Frerer (eds.), *Medical Microbiology and Infectious Diseases,* W. B. Saunders, 1981.

virus and possibly all the other retroviruses listed in the **table** belong to this group. In addition, most of the animal cancer-inducing retroviruses (not listed in the table) fall into this category, since even the nondefective retroviruses infect persistently and do not kill the host cell. *See* TUMOR VIRUSES.

Modified immunological tolerance can sometimes lead to the establishment of persistent infections. An example is lymphocytic choriomeningitis virus in mice, previously mentioned in connection with defective interfering particles. The virus can sometimes be contracted by the fetus while still in the womb, in which case a form of immunological tolerance develops with only a low level of antibodies generated at a later age. The circulating virus-antibody complexes which are at the root of the pathology of this disease are very slow to develop.

Sequestration of viruses in cells of either the peripheral or the central nervous systems probably plays an important role in the mechanism of latency. These primarily nondividing cells do not present a good environment for virus replication but can retain viruses for very long periods of time. Herpes virus, although not detected immediately, was shown to emerge when fragments of trigeminal ganglia were cultured in the laboratory. In intact organisms the virus is probably transported through the axoplasm in and out of these cells.

Slow infections. The term slow infection was introduced by Sigurdsson in 1954, while studying visna, maedi, and scrapie in Icelandic sheep. His criteria for a slow infection were as follows: a long preclinical period, extending for months or years from the time of first exposure; an extended period of a clinical course after appearance of symptoms; and confinement of pathology to a single organ or tissue system. Many of the slow infections in humans and other animals are represented in the table. A distinction is made between diseases in which a viral etiology was established and those in which the presence of virus particles cannot be demonstrated (unconventional agents). It now believed that the scrapie agent is a protein (mol wt 27,000–30,000) for which the name prion has been suggested. *See* PRION DISEASE; SCRAPIE.

Kuru, a neurologic affliction of people of the Fore and neighboring linguistic groups in New Guinea, was linked with mourning rites that included cannibalism, during which the population became infected by the Kuru agent, presumably through breaks in the skin or consumption of infected human tissue. A successful transmission of this disease to experimental animals prompted the investigation of other similar diseases, specifically the Creutzfeldt-Jakob disease, which has symptoms similar to Kuru. Eventually this rare but widespread

presenile dementia was successfully transmitted to chimpanzees. Since the discontinuation of the cannibalistic mourning rites among the Fore people, no new cases of Kuru have been observed in the subsequent generation of children.

Pathology. The tissue system most frequently affected in slow infections is the central nervous system. This is true of infections with papovaviruses (JC and other SV40-like viruses) and the unconventional agents Kuru, Creutzfeldt-Jakob disease, scrapie, and transmissible mink encephalopathy. The unconventional agents cause a degeneration that gives the tissue a spongy appearance when viewed under the light microscope; hence the name spongiform encephalopathies. The apparent sponginess is brought about by vacuolation of the neurons and the neuropil.

The papovaviruses cause demyelination through the destruction of the myelin sheath–furnishing cells, the oligodendrocytes. These symptoms, referred to as progressive multifocal leukoencephalopathy, have frequently occurred as a complication in persons with chronic lymphocytic leukemia, with Hodgkin's disease, or following treatment with immunosuppressants. Viruses, similar to JC (a human papovavirus), isolated from these cases are serologically distinct from other known papovaviruses, although they strongly cross-react with SV40.

Visna virus and subacute sclerosing panencephalitis (caused by an altered measles virus) are examples of inflammatory encephalitides. Lymphocytes and plasma cells infiltrate the central nervous system and form foci. The tissue becomes more firm; hence the name sclerosing.

The involvement of the central nervous system in slow infections manifests itself by disorientation, motor dysfunction, paralysis, and eventually death.

Latent infections. Herpes viruses are the subject of the most intensive studies of latency. The detection of infectious virus only after prolonged cultivation of explanted neural tissue implies the presence of the virus genome in a noninfectious state during latency. However, it can be argued that the infectious virus may be present in undetectably small quantities and the prolonged cultivation is necessary to increase its concentration to a detectable threshold. Despite the application of modern DNA technology to resolve this problem, the mechanism of establishment, maintenance, and reactivation of latent infections still remains obscure. *See* HERPES.

Human T-cell lymphotropic virus (HTLV) family. HTLV-I was isolated from several cases of cutaneous T-cell lymphomas in the United States. This virus is either identical or very closely related to the human adult T-cell leukemia virus (ATL), which is endemic in southwest Japan. A much rarer variant of the HTLV-I, the HTLV-II, has also been identified. HTLV-I (ATL) and HTLV-II seem to be the only human retroviruses positively identified in human malignancies. A third variant of this group, the HTLV-III or lymphadenopathy associated virus (LAV), was isolated from individuals suffering from acquired immune deficiency syndrome (AIDS). Antibodies against the last named virus have been found in 90–100% of AIDS patients, which indicates that they may have been infected with this agent. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS); ANIMAL VIRUS; VIRULENCE; VIRUS.

M. E. Reichmann

Bibliography. D. C. Bolton, M. P. McKinley, and S. B. Prusiner, Identification of a protein that purifies with the scrapie prion, *Science*, 218:1309–1310, 1982; H. Fraenkel-Conrat and R. R. Wagner (eds.), *Comprehensive Virology*, vol. 16, 1980; H. Hanfusa, M. E. Pullman, and A. Pinter (eds.), *Retrovirus and Disease*, 1989; J. A. Levy, H. Fraenkel-Conrat, and R. A. Owens., *Virology*, 3d ed., 1995; A. C. Minson, The state of the herpes genome, *Nature*, 304:477, 1983; R. Weiss, Retroviruses linked with AIDS, *Nature*, 300:12–13, 1984.

Virus interference

Inhibition of the replication of a virus by a previous infection with another virus. The two viruses may be unrelated, related, or identical. In some cases, virus interference may take place even if the first virus was inactivated. The term mutual exclusion has been applied to this phenomenon in bacterial viruses.

Several mechanisms of interference can be distinguished: (1) Inactivation of cell receptors by one virus may prevent subsequent adsorption and penetration by another virus. (2) The first virus may inhibit or modify cellular enzymes or proteins required for replication of the superinfecting virus. (3) The first virus may generate destructive enzymes or induce the cell to synthesize protective substances which prevent superinfection. (4) The first virus may generate defective interfering particles or mutants which may inhibit the replication of the infecting virus by competing with it for a protein (or enzyme) available in limited quantities; this type of viral interference has been called autointerference, and depends on a greater replicative efficiency of the defective interfering particles or mutants, compared to the infecting virus.

Examples of the first of these mechanisms are furnished by paramyxoviruses (parainfluenza), which contain in their envelope a neuraminidase enzyme. The enzyme hydrolyzes neuraminic (sialic) acid and its derivatives in the cell receptors utilized by the infecting virus or its progeny. Since closely related viruses of this group utilize the same receptors, superinfection cannot take place once the receptors have been destroyed.

Numerous examples of the second and third mechanisms are found among animal and bacterial viruses. The viral functions involved in modification of cellular proteins are usually also connected with the inhibition of normal cellular pathways. For example, poliovirus modifies the structure of cellular

ribosomes and renders them ineffective in cellular protein synthesis, while their participation in poliovirus protein synthesis is unimpaired. Superinfection with vesicular stomatitis virus (VSV) is abortive, because VSV protein synthesis is also inhibited. Similarly, ϕX 174 bacteriophage inhibits cellular deoxyribonucleic acid (DNA) synthesis as well as DNA synthesis of a superinfecting phage, and T4 phage generates enzymes which degrade DNA of the cell or of a superinfecting virus. Many animal viruses induce the synthesis of a group of host cell glycoproteins known as interferons. These interact with membranes of adjacent cells and confer on them resistance to viral infection. A variety of induced enzymatic activities results from this interaction between interferons and cells.

The generation of defective particles that interfere with the replication of the incoming virus has been reported for all major groups of animal viruses. This phenomenon may be important in persistent infection. *See* VIRUS, DEFECTIVE; VIRUS INFECTION, LATENT, PERSISTENT, SLOW.

No medical use of interference in prevention of viral infection has been made. The localized effect of the protecting virus and the requirement for its constant replication at relatively high levels reduces its efficacy as a protective agent. It is also difficult to find a protecting virus which is completely harmless to the organism.

Viral interference has been used in classifying relationships between paramyxoviruses, based on the use of identical receptors. It was also applied as a diagnostic test for rubella virus. Cells infected with this virus in tissue culture are protected from superinfection by other viruses (such as polio, measles, and influenza). The presence of rubella in the cells can be confirmed by the absence of cytopathic effects and cell death following superinfection by any of these viruses. *See* ANIMAL VIRUS; VIRULENCE; VIRUS.

M. E. Reichmann

Bibliography. H. Fraenkel-Conrat and R. R. Wagner (eds.), *Comprehensive Virology*, vol. 10, 1977; J. A. Levy, H. Fraenkel-Conrat, and R. A. Owens, *Virology*, 3d ed., 1995; S. Pestka, The purification and manufacture of human interferons, *Sci. Amer.*, 249:36–43, August 1983; M. E. Reichmann and W. M. Schnitzlein, Defective interfering particles of rhabdoviruses, *Curr. Topics Microbiol. Immunol.*, 86:123–168, 1979.

Viscosity

The material property that measures a fluid's resistance to flowing. For example, water flows from a tilted jar more quickly and easily than honey does. Honey is more viscous than water, so although gravity creates nearly the same stresses in honey and water, the more viscous fluid flows more slowly.

Shear viscosity. The viscosity can be measured in an experiment where the fluid of interest is sheared between two flat plates which are parallel to one another (**Fig. 1**). This is known as planar Couette flow. The shear flow created between the plates has the linear velocity profile given by Eq. (1), where v_x is

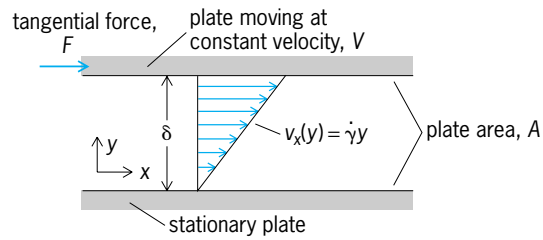

Fig. 1. Planar Couette flow. v_x = fluid velocity at distance y above the stationary plate, $\dot{\gamma}$ = velocity gradient or shear rate, δ = distance between plates.

$$v_x = \dot{\gamma}y \tag{1}$$

the velocity parallel to the plates at a perpendicular distance y above the stationary plate. The coefficient $\dot{\gamma}$, called the velocity gradient or shear rate, is given by V/δ, where δ is the distance between the plates. The shear stress is the ratio of the tangential force F needed to maintain the moving plate at a constant velocity V to the plate area A. It is expected that the shear stress increases with increasing shear rate but that the ratio of these two quantities depends only on the fluid between the plates. This ratio defines the shear viscosity, η, as in Eq. (2). The shear

$$\eta \equiv \frac{\text{shear stress}}{\text{shear rate}} = \frac{F/A}{V/\delta} \tag{2}$$

viscosity may depend on temperature, pressure, and shear rate.

Newton's law of viscosity. Isaac Newton is credited with first suggesting a model for the viscous property of fluids in 1687. Newton proposed that the resistance to flow caused by viscosity is proportional to the velocity at which the parts of the fluid are being separated from one another because of the flow. G. G. Stokes gave the first mathematical formulation of Newton's law of viscosity in 1845. Although Newton's law of viscosity is an empirical idealization, many fluids, such as low-molecular-weight liquids and dilute gases, are well characterized by it over a large range of conditions. However, many other fluids, such as polymer solution and melts, blood, ink, liquid crystals, and colloidal suspensions, are not described well by Newton's law. Such fluids are referred to as non-newtonian. *See* NON-NEWTONIAN FLUID; NON-NEWTONIAN FLUID FLOW.

For planar Couette flow (Fig. 1), the one-dimensional form of Newton's law of viscosity is given mathematically by Eq. (3), where τ_{yx} is the

$$\tau_{yx} = \mu \frac{dv_x}{dy} = \mu\dot{\gamma} \tag{3}$$

shear stress, and μ, a function of temperature and pressure, is the coefficient of viscosity or simply the viscosity. Therefore, by comparing Eqs. (2) and (3) the shear viscosity is equal to the coefficient of viscosity (that is, $\eta = \mu$) for a newtonian fluid.

TABLE 1. Viscosity conversions

Unit	poise	cp	Pa · s	$lb_m/(ft \cdot s)$	$lb_f \cdot s/ft^2$
1 poise* =	1	100	0.1	6.72×10^{-2}	2.089×10^{-3}
1 centipoise =	0.01	1	0.001	6.72×10^{-4}	2.089×10^{-5}
1 pascal-second† =	10	1000	1	0.672	2.089×10^{-2}
1 $lb_m/(ft \cdot s)$ =	14.88	1488	1.488	1	3.108×10^{-2}
1 $lb_f \cdot s/ft^2$ =	478.8	4.788×10^4	47.88	32.17	1

*1 poise = 1 dyne · s/cm^2 = 1 g/(cm · s).
†1 Pa · s = 1 kg/(m · s).

Because of this relation the shear viscosity is also often referred to as the viscosity. However, it should be clear that the two quantities are not equivalent; μ is a newtonian-model parameter, which varies only with temperature and pressure, while η is a more general material property which may vary nonlinearly with shear rate. *See* FLUID FLOW; NEWTONIAN FLUID.

Units and data. From Eqs. (2) and (3), the units of viscosity are given by force per area per inverse time. If in planar Couette flow, for example, 1 dyne of tangential force is applied for every 1 cm^2 area of plate to create a velocity gradient of 1 s^{-1}, then the fluid between the plates has a viscosity of 1 poise (=1 dyne · s/cm^2). Several viscosity units are in common use (**Table 1**). Comparison of the viscosities of different fluids (**Table 2**) demonstrates some general trends. For example, the viscosity of gases is generally much less than that of liquids. Whereas gases tend to become more viscous as temperature is increased, the opposite is true of liquids. Other data also show that increasing pressure tends to increase the viscosity of dense gases, but pressure has only a small effect on the viscosity of dilute gases and liquids. *See* DIMENSIONS (MECHANICS); UNITS OF MEASUREMENT.

Measurements. For newtonian fluids, there are several simple experiments from which the viscosity can be obtained. For example, when a solid sphere (for example, a ball bearing) of diameter D and density ρ_s falls through a fluid of viscosity μ and density ρ, the sphere eventually reaches a constant velocity, that is, the terminal velocity v_t. The flow around the sphere is well known if the Reynolds number of this flow, $\rho v_t D/\mu$, is roughly less than 0.1. Under these conditions the viscosity is given by Eq. (4), where g is

$$\mu = \frac{D^2(\rho_s - \rho)g}{18v_t} \tag{4}$$

the acceleration due to gravity. *See* CREEPING FLOW; REYNOLDS NUMBER.

For the flow of a newtonian fluid of density ρ and viscosity μ through straight conduit of circular cross section and radius R, such as a pipe or a capillary tube, the volume flow rate, Q, is proportional to the pressure drop per length, $\Delta p/L$. The viscosity is then given by the Hagen-Poiseuille law, Eq. (5). This relation assumes that the flow is laminar, which is valid when the Reynolds number for this flow, $2\rho Q/(\pi\mu R)$, is less than about 2100. *See* LAMINAR FLOW; PIPE FLOW.

$$\mu = \frac{\pi(\Delta p/L)R^4}{8Q} \tag{5}$$

The two methods of measuring the viscosity described above require assuming newtonian fluid behavior. The simple analysis used to derive Eqs. (4) and (5) fails for non-newtonian fluids. For this reason, devices have been developed to measure fluid properties in flows that closely approximate planar Couette flow (Fig. 1). One such flow can be created by shearing a fluid between a flat, circular plate and a cone (**Fig. 2**). The fluid is held in the gap between the cone and plate by the surface tension of the fluid. If

TABLE 2. Coefficients of viscosity of selected gases and liquids

Fluid	Temperature, °C (°F)	μ, centipoise
Air (gas)	−104 (−155)	0.01130
	0 (32)	0.01708
	74 (165)	0.02102
Hydrogen (gas)	0 (32)	0.00835
Ethane (gas)	0 (32)	0.00848
Water (liquid)	5 (41)	1.519
	20 (68)	1.002
	95 (203)	0.2975
Mercury (liquid)	0 (32)	1.685
Ethyl alcohol (liquid)	0 (32)	1.773
Olive oil (liquid)	10 (50)	138.0
Glycerin	0 (32)	12,110

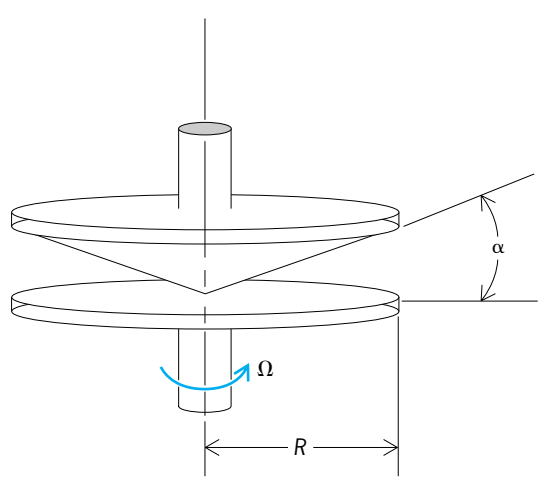

Fig. 2. Cone-and-plate geometry. α = cone angle, Ω = angular velocity of plate, R = radius of plate. (*After A. A. Collyer and P. W. Clegg, Rheological Measurement, Elsevier Applied Science, 1988*)

the cone angle α is small (roughly, less than 5°), then rotating either the plate (Fig. 2) or the cone at an angular velocity Ω generates a nearly constant shear rate of $\dot{\gamma} = \Omega/\alpha$ everywhere in the gap. The shear viscosity is proportional to the torque, $\mathscr{T}$, required to maintain steady rotation and is given by Eq. (6),

$$\eta(\dot{\gamma}) = \frac{3\mathscr{T}\alpha}{2\pi R^3 \Omega} \tag{6}$$

where R is the radius of the plate. By varying the angular velocity and measuring the applied torque, the cone-and-plate viscometer can be used to measure the shear viscosity over a large range of shear rates. Some sources of error in this measurement can arise from viscous heating, which can increase the temperature of the fluid; fluid inertia, which may cause the shear rate to vary in the gap; and nonideal geometry, which can create instabilities and nonideal flow.

Molecular origins. In classical fluid mechanics, the fluid is treated as a continuous medium without molecular structure. This approximation makes it necessary to use the concept of stress to account for the transport of momentum in the fluid caused by molecular interactions. It has been the goal of much research to derive expressions for the stress from molecular models of the fluid. One example is the kinetic theory of dilute gases.

Planar Couette flow (Fig. 1) can be considered for a dilute gas between two plates. The molecules of this gas can be modeled as nonattracting, hard spheres with diameter d and mass m. The motion of the spheres is chaotic with many collisions occurring between spheres themselves and between the spheres and the plates. The number density of spheres per volume n is small such that the mean free path is large (that is, the average distance traveled between collisions). From the kinetic theory of gases, the mean free path is

$$\lambda = 1/(\sqrt{2}\pi d^2 n)$$

and the average velocity of the gas molecules relative to the fluid velocity v is

$$\bar{c} = \sqrt{8k_B T/\pi m}$$

where k_B is Bolzmann's constant and T is the absolute temperature. The molecules exchange momentum during collisions; the slower molecules gain momentum, while the faster molecules lose momentum. For the flow described by Fig. 1, there is a flux of x momentum in the y direction which is the shear stress τ_{yx}. This model leads to a viscosity given by Eq. (7).

$$\mu = \frac{1}{3} nm\bar{c}\lambda = \frac{2}{3\pi^{3/2}} \frac{\sqrt{mk_B T}}{d^2} \tag{7}$$

See KINETIC THEORY OF MATTER; MEAN FREE PATH.

The simple hard-sphere model correctly predicts the independence of the viscosity on pressure and shear rate for very dilute gases. It also predicts that the viscosity increases as $\sqrt{T}$, which is qualitatively correct, although the viscosity of real gases depends more strongly on temperature. More refined models that employ potential energy functions lead to better quantitative agreement with data. *See* GAS; INTERMOLECULAR FORCES.

Although advanced kinetic theories of gases have been quite successful, a molecular theory of viscosity for liquids is necessarily more difficult. Whereas dilute gas molecules interact primarily in pairs as they collide, molecules in the liquid phase are in continuous interaction with many neighboring molecules. The concepts of average velocity and mean free path have little meaning for liquids. It is clear, however, that increasing temperature increases the mobility of molecules, thus allowing neighboring molecules to more easily overcome energy barriers and slip past one another. Such arguments lead to an exponential relation for the viscosity of the form given by Eq. (8),

$$\mu = A \exp\langle B/k_B T\rangle \tag{8}$$

where the constants A and B are characteristic of the liquid and must be determined experimentally. Although this is an approximate model, it does correctly give the decrease in viscosity with temperature that has been experimentally observed. *See* LIQUID.

Complex fluid behavior. Many non-newtonian fluids not only exhibit a viscosity which depends on shear rate (pseudoplastic or dilantant) but also exhibit elastic properties. These viscoelastic fluids require a large number of strain-rate-dependent material properties in addition to the shear viscosity to characterize them. The situation can become more complex when the material properties are time dependent (thixotropic or rheopectic). Fluids that are nonhomogeneous or nonisotropic require even more sophisticated analysis. The field of rheology attempts to deal with these complexities. *See* RHEOLOGY.

Lewis E. Wedgewood

Bibliography. R. B. Bird, W. E. Stewart, and E. N. Lightfoot, *Transport Phenomena*, 2d ed., Wiley, 2002; A. A. Collyer and D. W. Clegg (eds.), *Rheological Measurement*, 2d ed., Springer, 1998; D. R. Lide (ed.), *CRC Handbook Chemistry and Physics*, 86th ed., CRC Press, 2005; F. A. Morrison, *Understanding Rheology*, Oxford University Press, 2001; R. J. Silbey, R. A. Alberty, and M. G. Bawendi, *Physical Chemistry*, 4th ed., Wiley, 2004.

Vision

The sense of sight, which perceives the form, color, size, movement, and distance of objects. Of all the senses, vision provides the most detailed and extensive information about the environment. Conversely, blindness is recognized as more disabling than deafness or any other sensory handicap. In the higher animals, especially the birds and primates, the eyes and the visual areas of the central nervous system have developed a size and complexity far beyond the other sensory systems.

Visual stimuli. These are typically rays of light entering the eyes and forming images on the retina at

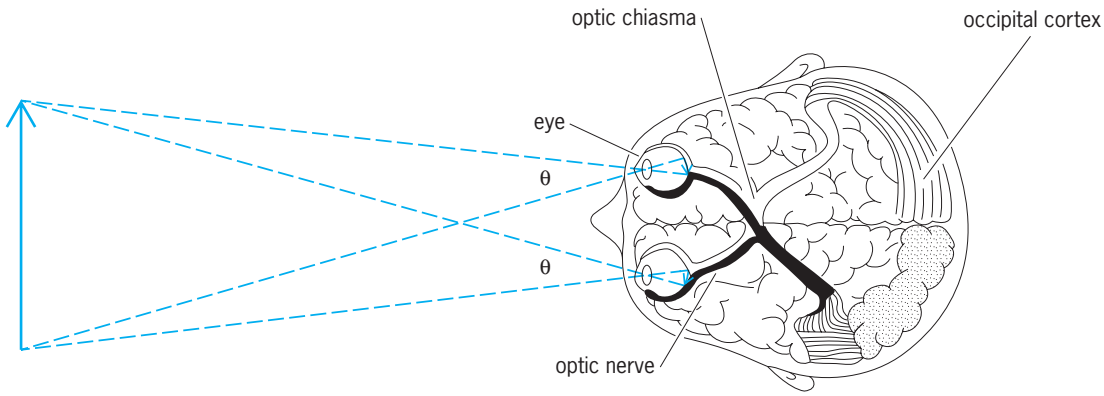

Fig. 1. Diagram showing the eyes and visual projection system. The visual angle θ is measured in degrees.

the back of the eyeball (**Fig. 1**). The intensity and wavelength characteristics of the light vary according to the light source and the object from which they are reflected. Human vision is most sensitive for light comprising the visible spectrum in the range 380–720 nanometers in wavelength. Sunlight and common sources of artificial light contain substantially all wavelengths in this range, but each source has a characteristic spectral energy distribution. In general, light stimuli can be measured by physical means with respect to their energy, dominant wavelength, and spectral purity. These three physical aspects of the light are closely related to the perceived brightness, hue, and saturation, respectively. *See* COLOR; LIGHT.

Atypical (sometimes called inadequate) stimuli for vision include momentary pressure on the eyeball, electric current through the eyes or head, a sudden blow on the back of the head, or disturbances of the central nervous system, caused by drugs, fatigue, or disease. Any of these may yield visual experiences not aroused by light. They are of interest because they show that the essentially visual character of the sensory experience is determined by the region stimulated (eyes, visual tracts) rather than by the nature of the stimulus. Indeed any observant person can detect swirling clouds or spots of "light" in total darkness or while looking at a homogeneous field such as a bright blue sky. These phenomena illustrate the spontaneous activity that is characteristic of the nervous system in general. They show that the visual system is continuously active, which in turn means that the effect of a stimulus is to modify existing activity and not merely to initiate new activity.

Anatomical basis for vision. The anatomical structures involved in vision include the eyes, optic nerves and tracts, optic thalamus, primary visual cortex, and higher visual areas of the brain. The eyes are motor organs as well as sensory; that is, each eye can turn directly toward an object to inspect it. The two eyes are coordinated in their inspection of objects, and they are able to converge for near objects and diverge for far ones. Each eye can also regulate the shape of its crystalline lens to focus the rays from the object and to form a sharp image on the retina. Furthermore, the eyes can regulate the amount of light reaching the sensitive cells on the retina by contracting and expanding the pupil of the iris. These motor responses of the eyes are examples of involuntary action that is controlled by various reflex pathways within the brain. *See* EYE (VERTEBRATE).

The process of seeing begins when light passes through the eye and is absorbed by the photoreceptors of the retina. These cells are activated by the light in such a way that electrical potentials are generated. These potentials are probably responsible for many features of the electroretinogram, an electrical response wave that can be detected by means of electrodes attached to the outside of the eye. Some of these potentials are the signs of electrochemical activity that serves to generate nerve responses in various successive neural cells (bipolars, ganglion cells, and others) in the vicinity of excitation. Finally, impulses emerge from the eye in the form of repetitive discharges in the fibers of the optic nerve. It must be emphasized, however, that the optic nerve impulses do not mirror exactly the excitation of the photoreceptors by light. Complex interactions within the retina serve to enhance certain responses and to suppress others. Furthermore it is a fact that each eye contains more than a hundred times as many photoreceptors as optic nerve fibers. Thus it would appear that much of the integrative action of the visual system has already occurred within the retina before the brain has had a chance to act.

The optic nerves from the two eyes traverse the optic chiasma. Figure 1 shows that the fibers from the inner (nasal) half of each retina cross over to the opposite side, while those from the outer (temporal) half do not cross over but remain on the same side. The effect of this arrangement is that the right visual field, which stimulates the left half of each retina, activates the left half of the thalamus and visual cortex. Conversely the left visual field affects the right half of the brain. This situation is therefore similar to that of other sensory and motor projection systems in which the left side of the body is represented by the right side of the brain and vice versa.

The visual cortex includes a projection area in the occipital lobe of each hemisphere. Here there appears to be a point-for-point correspondence between the retina of each eye and the cortex. Thus

the cortex contains a "map" or projection area, each point of which represents a point on the retina and therefore a point in visual space as seen by each eye. But the map is much too simple a model for cortical function. Vision tells much more than the location at which an object is seen. Visual tests show that other important features of an object such as its color, motion, orientation, and shape are simultaneously perceived. In monkey experiments, neurophysiologists have identified many types of cortical cells. each one responding selectively to these critical features. The two retinal maps are merged to form the cortical projection area. This merger allows the separate images from the two eyes to interact with each other in stereoscopic vision, binocular color mixture, and other phenomena. In addition to the projection areas on the right and left halves of the cortex, there are visual association areas and other brain regions that are involved in vision. Complex visual acts, such as form recognition, movement perception, and reading, are believed to depend on widespread cortical activity beyond that of the projection areas. *See* BRAIN.

Scotopic and Photopic Vision

Night animals, such as the cat or the owl, have eyes that are specialized for seeing with a minimum of light. This type of vision is called scotopic. Day animals such as the horned toad, ground squirrel, or pigeon have predominantly photopic vision. They require much more light for seeing, but their daytime vision is specialized for quick and accurate perception of fine details of color, form, and texture, and location of objects. Color vision, when it is present, is also a property of the photopic system. Human vision is duplex; humans are in the fortunate position of having both photopic and scotopic vision. Some of the chief characteristics of human scotopic and photopic vision are enumerated in the **table**.

Scotopic vision. This occurs when the rod receptors of the eye are stimulated by light. The outer limbs of the rods contain a photosensitive substance known as visual purple or rhodopsin. This substance is bleached away by the action of strong light so that the scotopic system is virtually blind in the daytime. Weak light causes little bleaching but generates neural inhibitory signals that lower the overall sensitivity of the eye. In darkness, however, the rhodopsin is regenerated by restorative reactions based on the transport of vitamin A to the retina by the blood. One experiences a temporary blindness upon walking indoors on a bright day, especially into a dark room or dimly lighted theater. As the eyes become accustomed to the dim light the scotopic system gradually begins to function. This process is known as dark adaptation. Complete dark adaptation is a slow process during which the rhodopsin is restored in the rod receptors of the retina. Faulty dark adaptation or night blindness is found in persons who lack rod receptors or have a dietary deficiency in vitamin A. These rare persons are unable to find their way about at night without the aid of strong artificial illumination. *See* VITAMIN A.

Dark adaptation is measured by an adaptometer, a device for presenting test flashes of light after various periods of time spent in the dark. The intensity of flash is varied to determine the momentary threshold for vision as dark adaptation proceeds. A 10,000-fold increase in sensitivity (that is, a reduction of 10,000 to 1 in the threshold intensity of flash) is often found to occur during a half-hour period of dark adaptation. By this time some of the rod receptors are so sensitive that only one elementary quantum (photon) of light is necessary to trigger each rod into action. A person can detect the presence of a flash of light that simultaneously affects only a few of the millions of rod receptors. Thus the scotopic sensitivity of the human eye approaches the ideal case of a receiver that is capable of responding to a single quantum of energy.

The variation of the scotopic threshold with wavelength of light is shown in the rod curve of **Fig. 2**. In spite of the variations in wavelength, the subject does not see any color when the intensity of the light is low enough to fall in the rod portion of the diagram. This scotopic vision is colorless or achromatic, in agreement with the saying that in the night all cats appear gray.

Normal photopic vision. Normal photopic vision has the characteristics enumerated in the table. Emphasis is placed on the fovea centralis, a small region at the very center of the retina of each eye.

Foveal vision. This is achieved by looking directly at objects in the daytime. In **Fig. 3** the image of a small object at F falls within a region almost exclusively populated by cone receptors. These are so closely packed together in the central fovea that their density is about 130,000 per square millimeter. Furthermore, each of the cones in the fovea is provided with

Characteristics of human vision

Characteristic	Scotopic vision	Photopic vision
Photochemical substance	Rhodopsin	Cone pigments
Receptor cells	Rods	Cones
Speed of adaptation	Slow (30 min or more)	Rapid (8 min or less)
Color discrimination	No	Yes
Region of retina	Periphery	Center
Spatial summation	Much	Little
Visual acuity	Low	High
Number of receptors per eye	120,000,000	7,000,000
Cortical representation	Small	Large
Spectral sensitivity peak	505 nm	555 nm

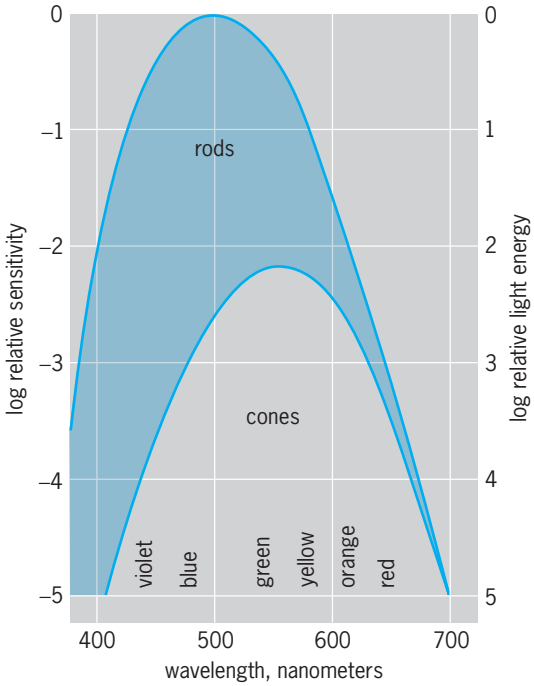

Fig. 2. Spectral sensitivity curves for human vision. The rod curve shows that scotopic vision, based on rhodopsin, is most sensitive to light of about 505 nm. The cone curve shows that photopic vision is generally less sensitive than scotopic vision, except for light at the red end of the spectrum.

a series of specialized nerve cells that process the incoming pattern of stimulation and convey it to the cortical projection area. In this way the cortex is supplied with superbly detailed information about any pattern of light that falls within the fovea centralis.

Peripheral vision. This is vision that takes place outside the fovea centralis (Fig. 3). As an example, look directly at a single letter at the center of a printed page. This letter, and a few letters immediately adjacent to it, appear clear and black because they are seen with foveal vision. The rest of the page is a blur in which the lines of print are seen as gray streaks. This is an example of peripheral vision. Vision actually extends out to more than 90° from center, so that one can detect moving objects approaching from either side. This extreme peripheral vision is comparable to night vision in that it is devoid of sharpness and color.

There is a simple anatomical explanation for the clarity of foveal vision as compared with peripheral vision. The cone receptors become less and less numerous in the retinal zones that are more and more remote from the fovea. In the extreme periphery there are scarcely any cones, and even the rod receptors are more sparsely distributed. Furthermore, the plentiful neural connections from the foveal cones are replaced in the periphery by network connections in which hundreds of receptors may activate a single optic nerve fiber. This mass action is favorable for the detection of large or dim stimuli in the periphery or at night, but it is unfavorable for visual acuity or color vision, both of which require the brain to differentiate between signals arriving from closely adjacent cone receptors.

Visual acuity. Visual acuity is defined as the ability to see fine details of an object. In Fig. 1 the small arrow at the back of each eye shows the image of the test object that is focused on the retina. Standard visual acuity is defined as the ability to see an object so small that the angle θ subtended at the eye is only 1 minute of arc, or 1/60 of a degree. At 20 ft (0.6 m) the size of such a test object is therefore only about 0.07 in., or 1.75 mm. The image of the object on the fovea (neglecting diffraction and optical aberrations that deteriorate the image) has a length of only 0.005 mm. Small as this is, it is twice the diameter of the smallest foveal cone receptor. One therefore comes to the conclusion that normal visual acuity approaches the limit imposed by diffraction and by the optical aberrations in the eye.

The specific forms of test object used for determining visual acuity yield different results. The angle θ in Fig. 1 can be made so small that it represents the size of test object that can barely be seen by the normal eye. This angular size can be infinitely small in the case of a bright point seen against a dark background. The stars at night provide a good example of this, since they are so distant that their angular subtense at the eye is practically zero. A long dark line can be detected against a bright field (for example a flagpole seen against the sky) if it subtends 1 second of arc (1/60 of a minute) at the eye. On the other hand, letters of the alphabet, as used in optometric wall charts, need to be composed of black lines 1 minute thick in order to be recognizable to the normal eye. A similar value holds for the lines of a grating (with parallel black and white lines of equal width) when viewed under good illumination.

The apparent discrepancy between acuity for single points or lines on the one hand, and for more complex forms on the other hand, can be explained by the effects of optical diffraction. This is a phenomenon resulting from the wave nature of light. It means that no optical image can ever be completely sharp and clear. The retinal image of a star is not a point but a blurred circle of light. The diameter of this blurred image is never less than 1 minute of arc,

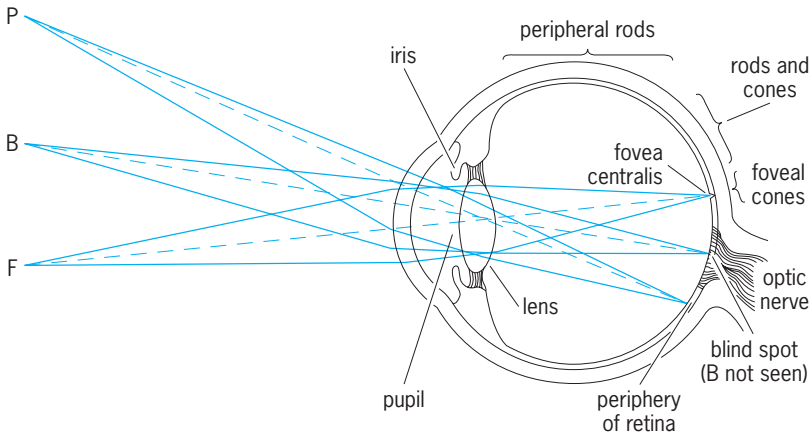

Fig. 3. Foveal and peripheral vision. Looking directly at F, the eye focuses the light from F on the central fovea, the region of clearest vision. P is seen poorly, and B not at all.

no matter how small an angle the star itself may subtend. Thus the star is seen, provided that the blurred image is noticeably brighter than the surrounding field. In the case of a grating (black and white stripe pattern) the image of each line is blurred also. If the lines are too close (less than 1 minute) together, the blurred image of one overlaps the blurred image of the next and the separate lines cannot be resolved by the eye. One comes to the conclusion, then, that the chief factors limiting the visual acuity of the fovea are (1) optical diffraction, (2) the ability to discriminate relative brightnesses within the images blurred by diffraction, and (3) the compactness of the pattern of cone receptors.

Space and Time Perception

Spatial and temporal effects are clearly apparent in the sense of sight. These two effects enable the individual to be oriented with regard to space and time in the surrounding world, especially in the perception of motion and distance. *See* PERCEPTION.

Space perception. Elementary forms of space perception are vernier and stereoscopic discrimination. Here, the eye is required to judge the relative position of one object in relation to another (**Fig. 4**). The left eye, for example, sees the lower line as displaced slightly to the right of the upper. This is known as vernier discrimination. The eye is able to distinguish fantastically small displacements of this kind, a few seconds of arc under favorable conditions. If the right eye is presented with similar lines that are oppositely displaced, then the images for the two eyes appear fused into one and the subject sees the lower line as nearer than the upper. This is the principle of the stereoscope. Again it is true that displacements of a few seconds of arc are clearly seen, this time as changes in distance. The distance judgment is made not at the level of the retina but at the cortex where the spatial patterns from the separate eyes are fused together. The fineness of vernier and stereoscopic discrimination transcends that of the retinal mosaic and suggests that some averaging mechanism must be operating in space or time or both. Furthermore, experiments with random dot stereograms demonstrate that the averaging mechanism can produce an impression of depth without visible contours in the retinal images of the two eyes.

The spatial aspects of the visual field are also of interest. As has been previously indicated, good acuity is restricted to a narrowly defined region populated by densely packed cones at the center of the visual field. A somewhat larger central region, in which the cones are somewhat more sparsely distributed, is capable of good color vision. Farther out, however, vision mainly is mediated by the rod receptors; color and form vision become extremely poor. In these peripheral regions area and intensity are reciprocally related for all small sizes of stimulus field. A stimulus patch of unit area, for example, looks the same as a patch of twice the same area and half the luminance. This high degree of areal summation is achieved by the convergence of hundreds of rod receptors upon each single optic nerve fiber. It is the basis for the ability of the dark-adapted eye to detect large objects even on a dark night.

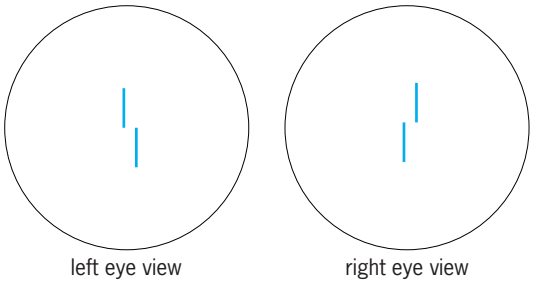

Fig. 4. Vernier and stereoscopic discriminations of space.

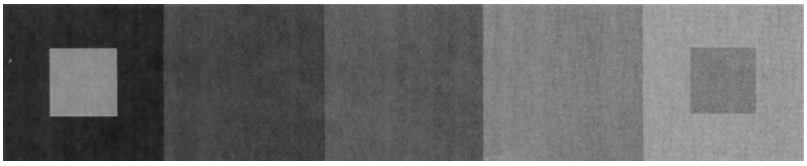

Fig. 5. Simultaneous contrast. The two small squares are physically equal, and each of the five strips is uniform. Spatial inhibitory effects cause the right square to look darker than the left square and cause the apparent darkening of the portion of each strip that is adjacent to a lighter strip.

In daytime vision, spatial inhibition, rather than summation, is most noticeable. The phenomenon of simultaneous contrast is present at a border between fields of different color or luminance. Thus a small square of gray paper appears darker on a light background than on a dark one; on a yellow background it appears bluish. The effect may originate in neural mechanisms of inhibition, such that stimulating a given region of a retina depresses the activity of regions immediately adjacent to it. This has the obvious effect of heightening contours and making forms more noticeable against their background (**Fig. 5**).

In photopic vision out of doors the eyes become light-adapted to an extent determined by such things as the time of day or the cloud cover. The luminance range of objects seen on a given day is typically no more than about 100 to 1 (two logarithmic units). Over this dynamic range the eyes appear to be capable of good spatial and temporal resolution, without the necessity for slowly adapting to the various levels of luminance.

Time perception. The temporal characteristics of vision are revealed by studying the responses of the eye to various temporal patterns of stimulation. When a light is first turned on, there is a vigorous burst of nerve impulses that travel from the eye to the brain. Continued illumination results in fewer and fewer impulses as the eye adapts itself to the given level of illumination. Turning the light off elicits another strong neural response. Afterimages are often seen at this time. A positive afterimage, resembling the original stimulus, is sometimes seen during the first fraction of a second after the light goes off. This is usually followed by a longer-lasting negative afterimage in which the color of the original object appears to be reversed. A clear afterimage may be produced by staring fixedly at the stimulus for at least

a half minute, then turning away and "projecting" the afterimage against a white or gray screen. The retinal size of the afterimage remains constant, so that its size in inches on a screen is proportional to the distance of the screen from the observer (Emmert's law). The afterimage is thought to arise chiefly from the fact that the affected region of the retina has been bleached, or otherwise changed photochemically, in such a way that its responses are different from those of the regions not stimulated. To some extent, however, visual aftereffects are due to central, rather than retinal, processes. *See* COLOR VISION.

The strength of a visual stimulus depends upon its duration as well as its intensity. Below a certain critical duration, the product of duration and intensity is found to be constant for threshold stimulation. A flash of light lasting only a few milliseconds may stimulate the eye quite strongly, providing its luminance is sufficiently high. A light of twice of the original duration will be as detectable as the first if it is given half the original luminance. However, even though two lights of the same energy may be equally detectable, a person may still be able to distinguish between them.

Voluntary eye movements enable the eyes to roam over the surface of an object of inspection. In reading, for example, the eyes typically make four to seven fixational pauses along each line of print, with short jerky motions between pauses. An individual's vision typically takes place during the pauses, so that one's awareness of the whole object is the result of integrating these separate impressions over time.

A flickering light is one that is going on and off (or undergoing lesser changes in intensity) as a function of time. At a sufficiently high flash rate (called the critical frequency of fusion, cff) the eye fails to detect the flicker, and the light pulses seem to fuse to form a steady light that cannot be distinguished from a continuous light that has the same total energy per unit of time. As the flash rate is reduced below the cff, flicker becomes noticeable, and at very low rates the light may appear more conspicuous than flashes occurring at higher frequency. The cff is often used clinically to indicate a person's visual function as influenced by drugs, fatigue, or disease. *See* PSYCHOLOGY.

Contrast sensitivity functions. Another approach to the study of vision is based on the concept of modulation, a term that has long been associated with periodic temporal changes in radio signals or in sound. But vision is particularly concerned, not with temporal variation, but with the spatial variations of light that characterize a pattern. The approach is to analyze pattern vision in terms of the contrast, or light modulation, that is measured within it. Furthermore, just as the methods of Fourier analysis have already yielded a good understanding of information transfer in radio transmission, the same methods are being applied with great success to pattern vision.

In Fourier terms, the simplest form of modulation is sinusoidal. Accordingly, an elementary form of visual pattern is the one in **Fig. 6**, where spatial frequency is 4 cycles per degree of visual angle (Fig. 1), average light intensity is I_{av}, and contrast (C) is $(I_{max} - I_{min})/(I_{max} + I_{min})$.

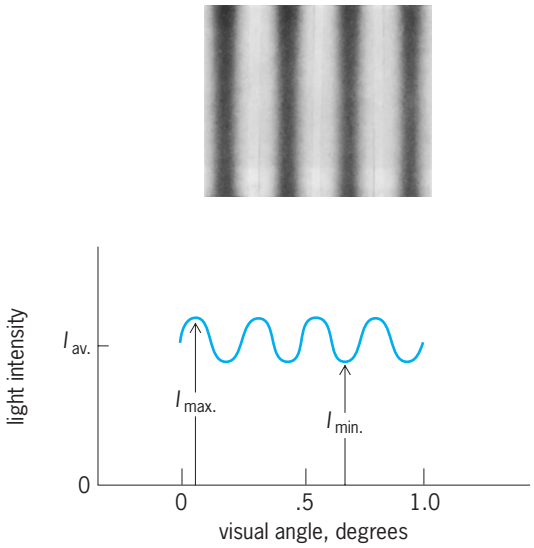

Fig. 6. Example of sinusoidal grating.

Sinusoidal gratings can conveniently be produced and controlled electronically on the face of a laboratory oscilloscope. To measure human contrast sensitivity, the contrast (C) of the grating is gradually reduced to threshold (C_0), that is, to a point where the grating is barely distinguishable from a uniform field of the same average intensity. Contrast sensitivity is defined as the reciprocal of this threshold value of contrast.

Several important facts about vision emerge from the sinusoidal grating experiments. (1) the entire range of spatial sensitivity can be explored in this way. A typical contrast sensitivity function for human vision (**Fig. 7**) peaks at about 3 to 4 cy/degree and becomes severely attenuated below 0.1 cy/deg and above 60 cy/deg. (2) prolonged viewing of a sinusoidal grating at high contrast causes a loss of contrast sensitivity for a grating of the same spatial frequency, but not for gratings of considerably higher or lower spatial frequency. This means that there

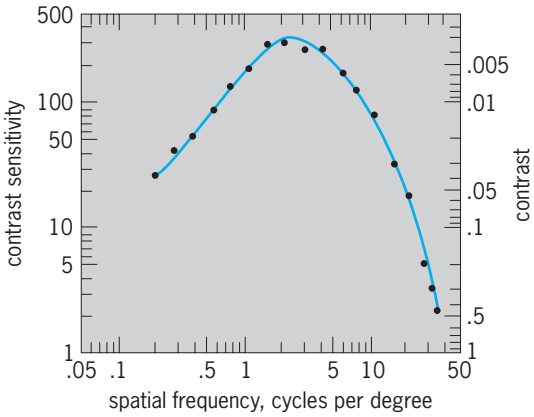

Fig. 7. Contrast sensitivity curve for a human subject. (*After F. W. Campbell and L. Maffei, Contrast and spatial frequency, Sci. Amer., 231:106–113, 1974*)

are more or less separate "channels" in the visual system for responding to spatial frequency, just as there are separate channels for color, motion, and orientation. (3) neurophysiologists have already identified single cells in the visual systems of cats and monkeys that are tuned for spatial frequency. Relatively coarse tuning is found in the cat, particularly with cells at the level of the retina or the thalamus. But in the monkey, especially in the visual cortex, some cells are found that respond only within narrow limits to particular grating frequencies. (4) the Fourier analysis of visual function is highly practical and efficient. It yields a quantitative comparison between human and animal capabilities in pattern vision. Similarly, it can be used to describe the course and severity of visual defect in patients with eye disease. Optical and lighting engineers are now using it to evaluate the conditions most favorable for vision throughout the entire frequency domain. And finally, any given pattern or scene can now be described in terms of its Fourier spatial frequency components, just as music or speech has long been analyzed in terms of its component temporal frequencies.

Lorrin A. Riggs; Charles E. Sternheim

Bibliography. H. Davson, *The Physiology of the Eye*, 4th ed., 1990; R. Held, H. W. Leibowitz, and H. L. Teuber (eds.), *Perception*, vol. 8 of *Handbook of Sensory Physiology*, 1978; D. H. Hubel, *Eye, Brain and Vision*, 1995; D. Jameson and L. M. Hurvich (eds.), *Visual Psychophysics*, pt. 4 of vol. 7, *Handbook of Sensory Physiology*, 1972; R. Jung (ed.), *Central Processing of Visual Information*, pt. 3 of vol. 7, *Handbook of Sensory Physiology*, 1973; R. Sekuler and R. Blake, *Perception*, McGraw-Hill, Boston, 2002; B. A. Wandell, *Foundations of Vision*, 1995.

Visual debugging

Visualization of computer program state and program execution to facilitate understanding and, if necessary, alteration of the program. Debuggers are universal tools for understanding what is going on when a program is executed. Using a debugger, one can execute the program in a specific environment, stop the program under specific conditions, and examine or alter the content of the program variables or pointers. Traditional command-line oriented debuggers allowed only a simple textual representation of the program variables (program state). In the Unix GNU debugger (GDB), for instance, one would enter "print *variable*" to see a textual representation of *variable* on the console:

```
(gdb) print*tree
      *tree = {value = 7, name
            = 0x8049e88"Ada
```

This textual representation did not change even when modern debuggers came with a graphical user interface. Although variable names became accessible by means of menus, the variable values were still presented as text, including structural information, such as pointers and references. Likewise, the program execution is available only as a series of isolated program stops. (Pointers are variables that contain the "addresses" of other variables.)

This article presents the most important techniques of visual debugging, each exemplified by a specific tool.

Visualizing data structures. The GNU Data Display Debugger (DDD) is a graphical front-end to a command-line debugger such as GDB, providing menus and other graphical interfaces that eventually translate into debugger commands. As a unique feature, DDD allows the visualization of data structures as graphs. The concept is simple: Double-clicking on a variable shows its value as an isolated graph node. By double-clicking on a pointer, the dereferenced value or the variable pointed is to shown as another graph node, with an edge relating pointer and dereferenced value. By subsequent double-clicking on pointers, the programmer can unfold the entire data structure (**Fig. 1**).

If a pointer points to a value that is already displayed (for example, in a circular list), no new node is created; instead, the edge is drawn to the existing value. Using this alias recognition, the programmer can quickly identify data structures that are referenced by multiple pointers.

In principle, DDD can render arbitrary data structures by means of nodes and edges. However, the programmer must choose what to unfold, as the screen size quickly limits the number of variables displayed. Nonetheless, DDD is one of the most popular debugging tools under Unix and Linux.

Animating algorithms. Lens is an experimental graphical debugging front-end that includes an algorithm animation engine. The basic idea is to associate the code with algorithm animation directives. For instance, to visualize a tree, one would draw a square for a tree node containing the value and draw edges to the left and right children. These animation directives are then automatically executed during program execution.

The drawing and animation directives of Lens allow arbitrary visual rendering of data structures, providing an effective abstraction from the details of the implementation. Also, automatic unfolding of large data structures becomes feasible, which is not possible in DDD. Finally, programmers can also specify how transitions between program states should be rendered, thus allowing animated presentations.

The disadvantage, of course, is that for each data structure to be examined (and for each abstraction level), a specific visualization must be prepared, which requires more work than the one-size-fits-all graph visualization in DDD. Thus, Lens is more a presentation than an investigation tool; its presentations, though, have been used successfully in several classrooms to teach algorithms.

Summarizing executions. Jinsight is a tool for visualizing the dynamic behavior of Java programs. In contrast to Lens and DDD, Jinsight attempts to summarize entire program runs in its visualizations: First,

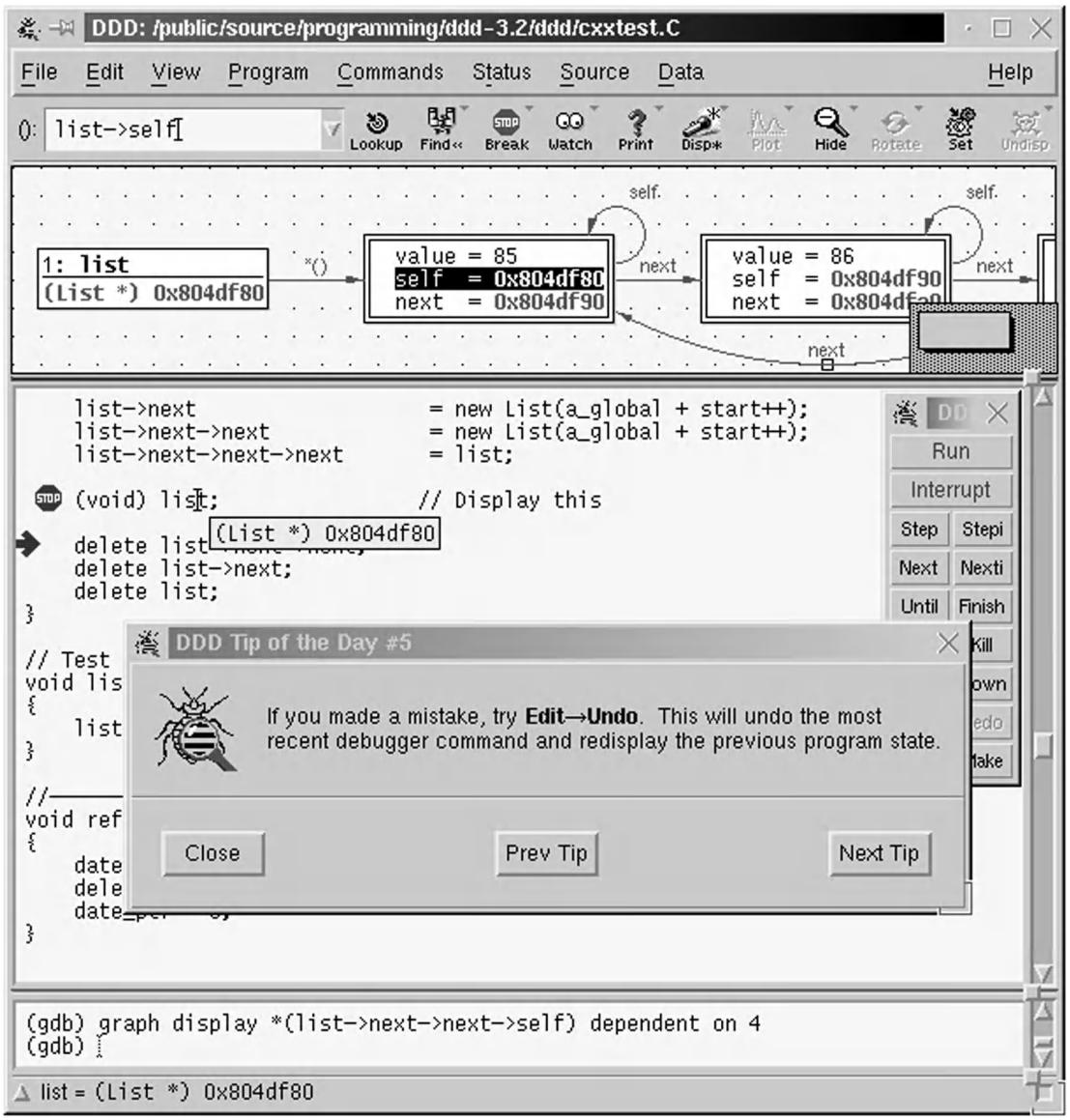

Fig. 1. Visualizing a list in DDD.

the program is executed using an extended Java virtual machine that tracks object events such as creation, interaction, or destruction during execution. At the end of the execution, this information is stored in a file which can then be examined by Jinsight.

Jinsight provides several views of the program execution. Object histograms show every single creation and destruction of arbitrary objects and highlight memory leaks. Call graphs show the sequence of invoked methods for several threads. (A thread is a portion of a program that can operate independently.) In all these views, Jinsight can identify common patterns and summarize these as well. For instance, in a call graph display, Jinsight can identify repeating sequences of function calls and summarize them to "this sequence of function calls is repeated *n* times." In practice, this effectively helps to isolate important events during program execution. Likewise, sets of objects with common properties can be summarized to a single visual representation, such as "each object of class c_1 references two objects of class c_2" (**Fig. 2**).

The drawback of Jinsight is that it allows only postmortem investigation; that is, the run can be neither explored nor influenced during execution. Also, Jinsight provides no means for exploring or rendering program states at one specific execution time. However, Jinsight is currently being extended to access program state while the program is still running, and will eventually be usable as an interactive debugger. Its innovative pattern recognition mechanisms make Jinsight the most advanced visual debugging tool available.

Conclusion. Compared to traditional debuggers, the techniques of visual debugging allow quicker exploration and understanding of what is going on in a program. Major challenges for the future include (1) finding appropriate techniques that summarize program executions and program states; (2) creating tools that can be used for both investigation and

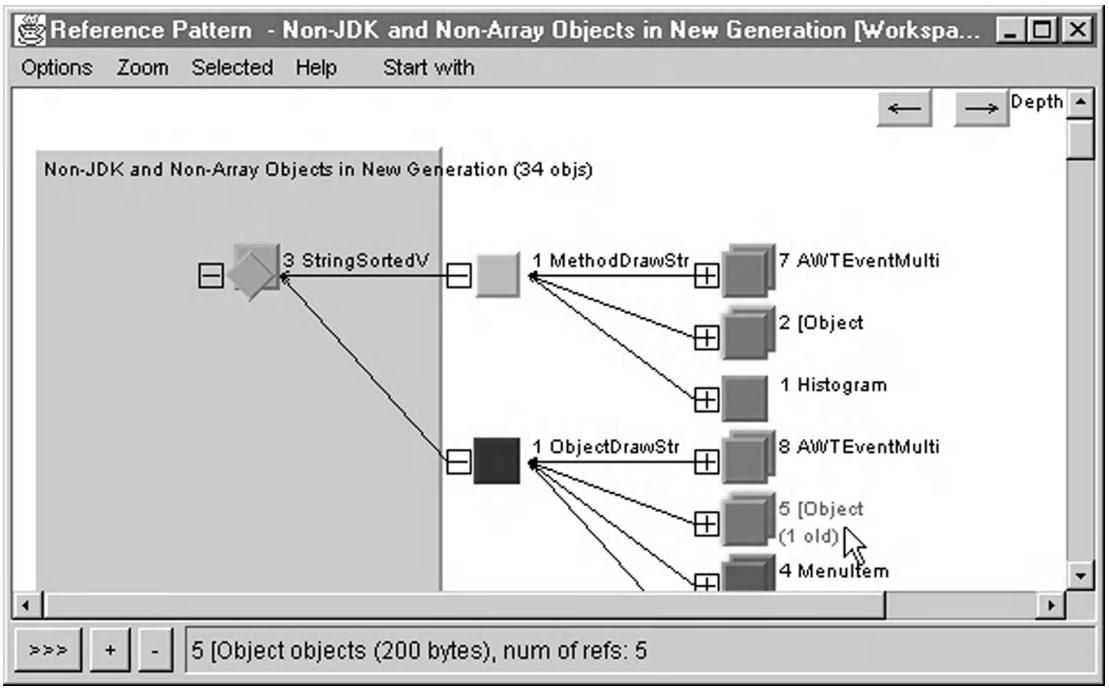

Fig. 2. Summarizing references in Jinsight.

presentation; (3) creating tools that allow exploring specific states as well as sets of states. *See* COMPUTER PROGRAMMING; PROGRAMMING LANGUAGES; SOFTWARE; SOFTWARE ENGINEERING. Andreas Zeller

Bibliography. W. de Pauw et al., Visualizing the behavior of object-oriented systems, *OOPSLA'93 Conference Proceedings* (Washington, DC), pp. 326–337, September 1993; S. Mukherjea and J. T. Stasko, Toward visual debugging: Integrating algorithm animation capabilities within a source level debugger, *ACM Trans. Comput. Hum. Interac.*, 1(3):215–244, September 1994; A. Zeller, Visual debugging with DDD, *Dr. Dobbs J.*, no. 322, pp. 21–28, March 2001.

Visual impairment

Abnormal visual acuity. The term is used to describe visual acuity substantially less than normal. The World Health Organization defines visual impairment as acuity less than 20/60 (normal being 20/20). Most states in the United States, however, will not issue an unrestricted driver's license to persons having corrected vision less than 20/40; such persons generally require special aids to read fine print and to perform other visually demanding tasks. The legal definition of blindness in the United States is visual acuity of 20/200 or worse (or severely restricted peripheral vision). The World Health Organization defines blindness as visual acuity worse than 20/400.

Visual impairment and blindness increase substantially with age. Results of a survey of an urban United States population indicated that nearly 1% of the population is visually impaired at age 40–49; this percentage increases to 7–8% by age 80. The major causes of blindness differ substantially by race (see **illus.**). Cataracts, which involve opacification of the normally clear lens, and diabetic retinopathy, which is an accumulation of fluid or the growth of abnormal blood vessels in the retina (most commonly in insulin-dependent diabetics), are important causes of blindness. A large proportion of people having blindness caused by cataracts can be treated by surgery. A growing proportion of blindness caused by diabetic retinopathy can be prevented by laser surgery.

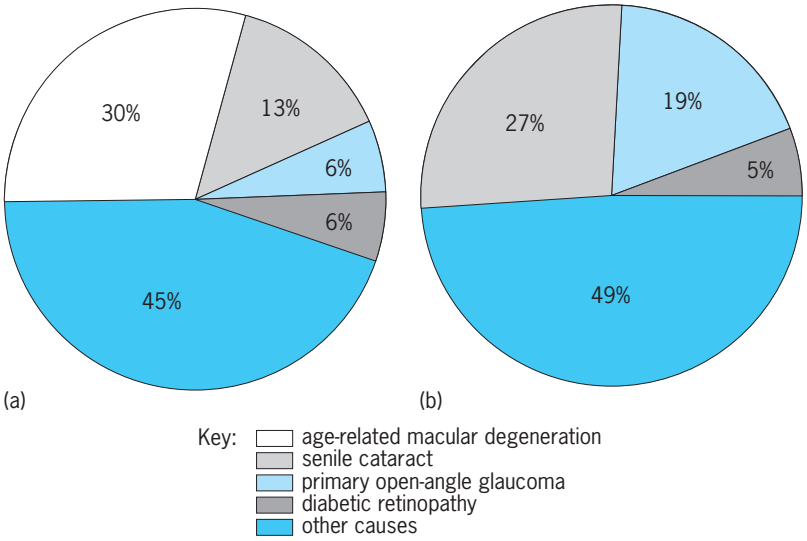

Major causes of blindness among Americans. (*a*) Whites. (*b*) Blacks. Glaucoma is more important among blacks; age-related macular degeneration is more important among whites. These preliminary data are from the Baltimore Eye Survey. Other causes of blindness include congenital cataract, corneal opacity, other types of glaucoma, hereditary retinopathy, macular scar, optic retinopathy, and trauma.

Other important causes of blindness are glaucoma and macular degeneration. Early detection of glaucoma requires routine, careful screening and examination, since many patients remain asymptomatic until much of the optic nerve is destroyed. Treatment may require drugs or surgery, the latter procedure creating an alternative pathway by which fluid can leave the eye, thereby lowering intraocular pressure to a safer level. Macular degeneration involves atrophy of that portion of the retina responsible for fine (reading) vision. A small proportion (probably less than 5%) of individuals with this condition also develop abnormal blood vessels that are amenable to laser surgery. For others, however, little can be done to prevent the onset or progression of age-related macular degeneration. However, while fine vision is often lost, peripheral vision is commonly retained. People with macular degeneration rarely suffer the total blindness of advanced glaucoma or diabetic retinopathy. *See* DIABETES; GLAUCOMA.

Visual problems are far more common in developing countries, where there are limited resources for dealing with problems that are otherwise readily treated, such as cataract, or prevented, such as trachoma or xerophthalmia ("blinding malnutrition" caused by vitamin A deficiency). In addition, in developing countries in the tropics, blinding parasitic diseases are common, particularly river blindness or onchocerciasis in Africa. Fortunately the outlook for onchocerciasis is improving with a successful, internationally organized campaign to control the black fly that spreads the disease and the development of a safer, more effective drug for treating the worm infestation that destroys vision. Trachoma, a chronic conjunctivitis that is endemic in many developing countries, is caused by the bacterium *Chlamydia trachomatis*. The infection can be treated with antibiotics or sulfonamides and prevented by proper hygiene and sanitation. *See* CATARACT; EYE DISORDERS; GLAUCOMA; VISION. Alfred Sommer

Bibliography. H. Goldstein, *The Demography of Blindness Throughout the World*, Res. Ser. 26, American Foundation for the Blind, 1980; J. M. Last (ed.), *Maxcy-Rosenau Public Health and Preventive Medicine*, 13th ed., 1992; A. Sommer, Attack blindness!, *Amer. J. Ophthalmol.*, 102:387–389, 1986; J. M. Tielsch et al., Blindness and visual impairment in an American urban population: The Baltimore Eye Survey, *Arch. Ophthalmol.*, 108:286–290, 1990.

Vitamin

An organic compound required in very small amounts for the normal functioning of the body and obtained mainly from foods. Vitamins are present in food in minute quantities compared to the other utilizable components of the diet, namely, proteins, fats, carbohydrates, and minerals. The average human adult eats about 600 g (21 oz) of food per day on a dry-weight basis, of which less than a gram is vitamins.

The discovery of the vitamins has been the result of primarily two lines of investigation: the study of nutritional disease in people and the feeding of purified diets of known composition to experimental animals. In this way vitamin deficiency diseases, known as avitaminoses, have been described. With the advent of vitamin fortification of flour, cereals, and other foods, specific vitamin deficiency diseases such as scurvy, rickets, pellagra, and beriberi have become rare in most developed countries. However, marginal deficiencies of many of the vitamins, which are difficult to detect and largely ignored, are still relatively common, and there is increasing evidence that vitamin status is related to the risk of developing chronic disease.

Synthetic and natural vitamins usually have the same biological value. Different vitamins, which are often not related to each other chemically or functionally, are conventionally divided into a fat-soluble group (vitamins A, D, E, and K) and a water-soluble group [vitamin C (ascorbic acid) and the various B vitamins]. Some animals require inositol and choline, whereas for most species these are not vitamins. The vitamins, particularly the water-soluble ones, occur almost universally throughout the animal and plant kingdoms (see **table**). *See* CHOLINE; INOSITOL.

Variations in the dietary requirements for vitamins between different animal species are usually related to the varying synthesizing abilities of the animal's tissues or gastrointestinal flora; for example, some species can synthesize vitamin C, others cannot. In addition, vitamin absorption may be hampered by certain events within the intestinal tract. For example, uptake of the fat-soluble vitamins is generally impaired by diseases of the biliary tract, due in part to the lack of adequate emulsification that results from lack of bile salts. Thiaminases in some foods may destroy thiamin within the intestinal lumen. Biotin may be rendered unabsorbable in the presence of avidin, a protein (commonly found in raw egg whites) that binds biotin.

The B vitamins function as coenzymes that catalyze many of the anabolic and catabolic reactions of living organisms necessary for the production of energy; the synthesis of tissue components, hormones, and chemical regulators; and the detoxification and degradation of waste products and toxins. However, vitamin C and the fat-soluble vitamins do not function as coenzymes. Vitamins C and E and β-carotene (a precursor of vitamin A) act as antioxidants, helping prevent tissue injury from free-radical reactions. In addition, vitamin C functions as a cofactor in hydroxylation reactions. Vitamin D has hormonelike activity in calcium metabolism; vitamin A plays a critical role in night vision, growth, and maintaining normal differentiation of epithelial tissue; and vitamin K has a unique posttranscriptional role in formation of active blood-clotting factors. *See* ANTIOXIDANT; COENZYME.

Fat-Soluble Vitamins

The fat-soluble group of vitamins comprises vitamins A, D, E, and K.

Vitamin sources and deficiencies

Recommended name	Alternative name	Best sources	Deficiency	Daily RDA for adults*
Fat-soluble				
A, β-carotene and other carotenoids	Retinol and provitamin A (for carotene)	Fish liver oils; liver; dairy products; yellow, orange, and green plants; carrots; sweet potatoes	Poor growth and night vision; blindness	1 mg (5000 IU) (as retinol)
D	D_2 or ergocalciferol; D_3 or cholecalciferol	Fish oils, fortified dairy products	Rickets, osteomalacia	5 µg (200 IU)†
E	Tocopherol	Grains and vegetable oils	Neuropathy	10 mg (15 IU)
K	K_1 or menaquinone; K_2 or phylloquinone	Green vegetables	Bleeding	80 µg
Water-soluble				
Vitamin C	Ascorbic acid	Citrus fruits, fresh vegetables, potatoes	Scurvy	60 mg‡
Thiamine	B_1	Pork, liver, whole grains	Beriberi	1.5 mg
Riboflavin	B_2	Milk, egg white, liver, leafy vegetables	Cheilosis, glossitis	1.7 mg
Niacin	Nicotinic acid and nicotinamide (or niacinamide)	Yeast, wheat germ, meats	Pellagra	19 mg§
B_6	Pyridoxine	Whole grains, yeast, egg yolk, liver	Skin disorders; convulsion in infants	2.0 mg
Biotin	—	Liver, kidney, yeast	Skin disorders	30–100 µg
Folic acid	Folate	Liver, deep-green leafy vegetables	Macrocytic anemias	0.2 mg
B_{12}	Cobalamin	Liver, meats	Pernicious anemia	2.0 µg
Pantothenic acid	—	Liver, kidney, green vegetables, egg yolk		4–7 mg

*Recommended dietary allowance (RDA; 1989) for 25–50-year-old male.
† No requirement with adequate exposure to sunlight.
‡ 100 mg for smokers.
§ 60 mg of tryptophan can replace 1 mg of niacin.

Vitamin A. This vitamin occurs in two principal forms in nature: retinol, which is found only in animal sources, and certain carotenoids (provitamins), which are found only in plant sources. Carotenoids are the compounds that give many fruits and vegetables their yellow and orange colorings. The most abundant and best known of the carotenoids is *β*-carotene, a precursor of vitamin A. It is a provitamin, because its vitamin A activity occurs upon conversion to retinol within the body. Most of this conversion takes place in the intestine. Foods rich in *β*-carotene include carrots, yellow vegetables, and dark-green leafy vegetables (for example, spinach and broccoli), pumpkins, apricots, and melons. Preformed vitamin A or retinol is found in liver, egg yolk, fish, whole milk, butter, and cheese. Both retinol and *β*-carotene are soluble in fat but not in water, and are destroyed readily by oxidation, which causes significant losses during storage and cooking. *See* CAROTENOID.

Requirements for vitamin A are expressed as retinol equivalents (RE) in order to standardize the many different sources and types of vitamin A. Many nutrition labels still express vitamin A in International Units (IU), an earlier method of measuring vitamin A. The equivalent of 1000 micrograms of retinol (1000 RE) is 5000 IU, with a typical diet containing one-half of the vitamin A preformed and one-half from carotenoids.

Vitamin A is essential for vision, for adequate growth, and for tissue differentiation. One of the primary signs of vitamin A deficiency in animals is loss of appetite, accompanied by growth retardation. The process whereby cells and tissues become "programmed" to carry out their special functions is known as differentiation. Vitamin A is necessary for normal differentiation of epithelial cells. In vitamin A deficiency, cells lose their ability to differentiate properly, and because abnormal differentiation is part of the cancer process, scientists have been investigating whether vitamin A status might be involved in some way in the prevention of cancer. *See* CANCER (MEDICINE); VISION.

One of the earliest symptoms of vitamin A deficiency is night blindness, or an impaired ability to see in dim light. Severe deficiency produces partial or total blindness (xerophthalmia). Vitamin A deficiency is by far most widespread and most serious among young children, especially in developing countries. It is the leading cause of child blindness and, in combination with other factors such as protein-calorie malnutrition or infection, is associated with high rates of child mortality. Therapeutic doses of vitamin A are distributed in targeted areas of the world to prevent xerophthalmia and to treat people in whom the early stages of blindness have already occurred. Administration of vitamin A in developing countries has resulted in a remarkable reduction in mortality of preschool-age children.

Vitamin A as retinol is well tolerated in human

adults up to levels of 50,000 IU per day; however, this is far in excess of generally recognized needs. Pregnant women should limit their intake to 10,000 IU per day, since in animals excess vitamin A is known to cause birth defects. β-Carotene, on the other hand, is the safest form of vitamin A, because it is converted by the body in a regulated way. It is poorly absorbed from the gastrointestinal tract, and its conversion to retinol becomes progressively less efficient as vitamin A status improves. Thus, high intakes of β-carotene cannot lead to hypervitaminosis A; they can, however, produce a reversible orange-yellow coloration of the skin (hypercarotenemia), which is harmless. *See* VITAMIN A.

Vitamin D. Also known as calciferol, vitamin D is essential for the proper formation of the skeleton and for mineral homeostasis. Exposure of the skin to ultraviolet light catalyzes the synthesis of vitamin D_3 (cholecalciferol) from 7-dehydrocholesterol. The other major form of the vitamin, D_2 (ergocalciferol), is the product of the ultraviolet light–induced conversion of ergosterol in plants. Vitamin D is converted to 1,25-dihydroxyvitamin D in the liver and kidney. This biochemically active form of the vitamin promotes absorption of calcium from the gut and helps maintain blood calcium levels by increasing release of calcium from the bone. *See* CALCIUM METABOLISM.

In the United States, the major source of dietary vitamin D is fortified foods such as milk. The vitamin is stable in foods; storage and cooking do not appear to affect its activity. Even moderate amounts of exposure of the skin to sunlight can provide adequate synthesis of the vitamin, although dark-skinned persons and the elderly have more limited capacity to synthesize the vitamin.

In children, vitamin deficiency results in deformation of the skeleton (rickets); in adults, it leads to a weakening of the bone, resulting from demineralization (osteomalacia). These diseases are uncommon because of food fortification with vitamin D. However, rickets does occur in infants who are breast-fed without supplemental vitamin D or exposure to sunlight, and rickets and osteomalacia can both occur in patients with kidney failure since these people cannot convert vitamin D to 1,25-dihydroxyvitamin D. The elderly are also at higher risk of vitamin deficiency as a result of a more limited ability to synthesize vitamin D in the skin when exposed to the sun, and a less efficient conversion to the active form, than the young. Some evidence suggests that poor vitamin D intake is associated with an increased risk of osteoporosis (weakening of the bones in the elderly). *See* OSTEOPOROSIS.

Vitamin D is reasonably well tolerated in adults. Intakes below 50,000 IU/day are generally not toxic; however, vitamin D is potentially toxic, especially in young children, in whom intake should not exceed 2000 IU/day. There is generally little reason, however, to exceed 400 IU/day. *See* VITAMIN D.

Vitamin E. There are at least eight naturally occurring forms of vitamin E, known as tocopherols and tocotrienols. α-Tocopherol is the most common form and also has the highest biological activity. The richest sources of vitamin E are seed oils and the margarines and shortenings made from these oils. Nuts, seeds, whole grains, leafy green vegetables, eggs, and milk are also good sources.

Vitamin E functions primarily as an antioxidant protecting the body tissues from damaging reactions (peroxidation) that arise from many normal metabolic processes and exogenous toxic agents. Thus it protects biological membranes such as those found in the nerves, muscles, and cardiovascular system; helps to prolong the life of red blood cells; and helps the body use vitamin A optimally.

Dietary deficiency is rare in humans. However, a progressive neuromuscular disease is observed in children and adults who are severely vitamin E–deficient as a result of a variety of syndromes involving malabsorption of fat. Symptoms include loss of coordination and balance and, in severe cases, loss of the ability to walk. In adults, less acute deficiencies have resulted in shortened life for red blood cells and an accumulation of so-called age pigment in certain tissues. In premature infants, a commonly occurring vitamin E deficiency is associated with a hemolytic anemia, bleeding in the brain, and damage to the retina of the eye.

The requirement for vitamin E is related to the amount of polyunsaturated fat consumed in the diet. The higher the amount of this type of fat, the more vitamin E is required; however, sources of polyunsaturated fat are usually good sources of vitamin E (for example, seed and vegetable oils). *See* FAT AND OIL (FOOD).

Vitamin E has the ability to block the formation of nitrosamines, which are chemicals that can cause cancer. Whether vitamin E helps prevent certain cancers is still under study. Epidemiological studies do suggest that the risk of cataracts is reduced with higher intake of vitamin E, and an immune-enhancing effect of vitamin E has been observed in the elderly. *See* VITAMIN E.

Vitamin K. The two naturally occurring forms of this vitamin are K_1 (menaquinone), which is synthesized by bacteria, and K_2 (phylloquinone), which is synthesized by plants. They are fat-soluble and unstable to light. Vitamin K_3 (menadione) is a synthetic form. About half of the requirements for vitamin K comes from dietary sources, predominantly green leafy vegetables, and half is synthesized in the intestinal tract by bacteria.

Vitamin K is essential for the formation of prothrombin and at least five other proteins involved in the regulation of blood clotting. The vitamin plays an essential role in the carboxylation of glutamic acid to γ-carboxyl glutamyl residues in prothrombin and other clotting proteins. Vitamin K plays a similar role in the biosynthesis of other proteins found in the plasma, bone, and kidney. However, defective coagulation, and the resultant bleeding, is the only major sign of vitamin K deficiency.

Vitamin K deficiency has been reported in some newborn infants, because the intestinal bacteria do not begin to supply large enough amounts of the

vitamin until about one week after birth. For this reason, vitamin K is given immediately after birth to prevent bleeding. In addition, infant formulas are fortified with vitamin K. Deficiency of vitamin K in adults is extremely rare. It is usually caused by prolonged treatment with antibiotics that inhibit the growth of intestinal flora, or due to anticoagulant therapy. Vitamin K antagonists are used as anticoagulants and as rodenticides. *See* BLOOD; VITAMIN K.

Water-Soluble Vitamins

The water-soluble vitamins are vitamin C and the B vitamins, which include thiamine (vitamin B_1), riboflavin (vitamin B_2), vitamin B_6, niacin, folic acid, vitamin B_{12}, biotin, and pantothenic acid.

Vitamin C. This vitamin, also known as ascorbic acid, is a white crystalline compound that is highly soluble in water and a strong reducing agent. The stability of ascorbic acid decreases with increases in temperature and pH. Destruction by oxidation is a serious problem in that a considerable quantity of the vitamin C content of foods is lost during processing, storage, and preparation. Vitamin C is found mainly in citrus fruits and green vegetables (broccoli, brussels sprouts, spinach, kale, cabbage, cauliflower), but there are other sources such as cantaloupes, mangoes, strawberries, tomatoes, sweet peppers, and potatoes.

Of all the animals studied, only the guinea pig, the red-vented bulbul bird, certain fruit-eating bats, most fish, and the primates, including humans, require a dietary source of vitamin C. The other species studied are capable of synthesizing the vitamin in the liver or kidneys.

Vitamin C is required for the hydroxylation of the amino acids proline and lysine. This is a critical step for the production of collagen, the intercellular "cement" substance that gives structure to muscles, tissues, bones, and cartilage. The vitamin also plays an important role in the synthesis of a number of neurotransmitters, which are chemicals involved in nerve conduction, and in the synthesis of carnitine, a compound involved in the oxidation of fatty acids for energy. In addition, vitamin C enhances the absorption of iron from the diet, blocks the formation of potentially carcinogenic nitrosamines in foods and in the stomach, increases the resistance to infectious diseases, and is necessary for collagen formation and optimum wound healing. Finally, the vitamin can function as an antioxidant, preventing peroxidation in foods and in the body. *See* COLLAGEN.

Deprivation of vitamin C for a sufficient period of time results in scurvy, a potentially fatal disease characterized by weakening of collagenous structures, which results in widespread hemorrhaging of the capillaries. Bleeding gums and loosening of the teeth are usually the earliest objective signs. Hemorrhages under the skin cause extreme tenderness of extremities and pain during movement. If vitamin C is not administered, gangrene and death may ensue.

People who smoke have a higher requirement for vitamin C. Epidemiological studies indicate that consumption of foods high in vitamin C is associated with a lower risk of gastrointestinal cancers, higher levels of high-density-lipoprotein (HDL) cholesterol (cholesterol associated with decreased heart disease), and lower blood pressure. Whether vitamin C itself is responsible for these beneficial effects remains to be proven. Vitamin C has also been shown to reduce the duration and severity of symptoms of the common cold. The vitamin is used in foods and for industrial purposes as a reducing agent and antioxidant. *See* ASCORBIC ACID.

Thiamine. Thiamine, also known as vitamin B_1, is a water-soluble vitamin found in many foods; pork, liver, and whole grains are particularly rich sources. It is unstable to heat, alkali, oxygen, and radiation. Because thiamine is soluble in water, about 25% of the thiamine in food is lost during the normal cooking process. A number of foods such as coffee, tea, raw fish, betel nuts, and some cereals contain antagonists to vitamin B_1. Thiamine is also lost during certain food manufacturing processes, such as the polishing of rice and the refining of flour; therefore, these foods are generally enriched with thiamine.

Thiamine is essential for carbohydrate metabolism. In the form of thiamine pyrophosphate, it is a coenzyme required for the oxidative decarboxylation of α-keto acids and for the activity of transketolase in the pentose phosphate pathway, both key steps in the breakdown of glucose to energy. Thiamine also plays a role in the conduction of nerve impulses and in aerobic metabolism.

The two principal thiamine deficiency diseases are beriberi and the Wernicke-Korsakoff syndrome. Marginal deficiencies of thiamine may produce vague symptoms such as anxiety, hysteria, nausea, depression, and loss of appetite prior to the onset of clinical symptoms of beriberi. Beriberi was once endemic in areas where polished rice made up a large part of the diet, especially Southeast Asia. Many countries now fortify rice and other cereal grains to replace the nutrients lost in processing, which has resulted in a decrease in the incidence of beriberi. Today it is the Wernicke-Korsakoff syndrome, generally associated with alcoholism, that is more likely to be encountered. The deficiency is caused by a combination of factors, including inadequate intake of the vitamin (as when alcohol replaces food), decreased absorption, and increased thiamine requirements. Symptoms range from mild confusion and depression to psychosis and coma. If treatment is delayed, memory may be impaired permanently.

Because thiamine has a high turnover rate and is not appreciably stored in the body (approximately 1 mg/day is used up in tissues), a continual supply is required. The requirement for thiamine is related to energy intake, and increases as large amounts of carbohydrates are consumed. Similarly, periods of increased metabolism such as fever raise the thiamine requirement. *See* CARBOHYDRATE METABOLISM; THIAMINE.

Riboflavin. Also known as vitamin B_2, riboflavin is a water-soluble yellow-orange fluorescent pigment that is widely distributed in nature, occurring mostly in milk, egg white, liver, and leafy vegetables. It is

stable to acid and oxidation, but is rapidly destroyed by alkali at elevated temperatures and by light.

Riboflavin deficiency results in poor growth and other pathologic changes in the skin, eyes, liver, and nerves. Riboflavin deficiency in humans is usually associated with a type of cracking at the corners of the mouth known as cheilosis; inflammation of the tongue, which appears red and glistening (glossitis); corneal vascularization accompanied by itching; and a scaly, greasy dermatitis about the corners of the nose, eyes, and ears. Because riboflavin is essential for the functioning of vitamin B_6 and niacin, and B vitamins are generally found together in the same foods, it is often difficult to isolate riboflavin deficiency from those of other B vitamin deficiencies.

The vitamin functions biochemically in two coenzymes: flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). Enzymes containing these coenzymes are known as flavoproteins. In general, these enzymes participate in oxidation-reduction reactions by accepting hydrogen ions (protons) and electrons from one substrate and transferring them to another.

Riboflavin requirements appear to be related to caloric requirements and muscular activity. *See* RIBOFLAVIN.

Vitamin B_6. This vitamin exists as three chemically related and water-soluble forms found in food: pyridoxine (pyridoxol), pyridoxal, and pyridoxamine. All three forms have equal activity for animals. They are all stable in acid but are rapidly destroyed by light in alkaline or neutral solutions. The best sources of vitamin B_6 are fish, chicken, organ meats, eggs, whole grains, beans, and nuts. Considerable losses occur in freezing of fruits and vegetables, processing of luncheon meats, and milling of cereals.

The three forms of vitamin B_6 are converted in the liver, red blood cells, and other tissues to pyridoxal phosphate and pyridoxine phosphate, which serve as coenzymes in transamination and decarboxylation reactions and other metabolic transformations of amino acids and in the metabolism of fats and nucleic acids. The vitamin is essential for the conversion of tryptophan to niacin.

Vitamin B_6 deficiency rarely occurs alone; it is commonly seen in people who are deficient in several B-complex vitamins. Clinical signs of deficiency include epileptiform convulsions, dermatitis, and anemia. Deficiency in adults is associated with depressed immune response and mental depression; in infants it leads to a variety of neurological symptoms as well as abdominal distress.

Approximately 40 drugs (for example, oral contraceptives, isonicotinic acid hydrazid, and penicillamine) are known to interfere with its metabolism. The need for pyridoxine increases proportionately with protein intake. Inadequate intake of vitamin B_6 is common, particularly in women, the elderly, alcoholics, and diabetics. Prolonged use of 500 mg or more per day of vitamin B_6 can cause ataxia and a severe sensory neuropathy. *See* VITAMIN B_6.

Niacin. Niacin, also known as nicotinic acid, is a white water-soluble powder stable to heat, acid, and alkali. It is found in biochemically active combinations as the amide, niacinamide. All living cells that have been studied have enzymic systems involving niacin. Many animals, including humans, are capable of synthesizing niacin in varying degrees from the amino acid tryptophan. Niacin is widely distributed in foods: yeasts, wheat germ, and meats, particularly organ meats, are rich sources of the vitamin. Some foods such as milk are relatively poor sources of niacin, but contain generous quantities of tryptophan.

Niacin-deficiency disease is known as pellagra and is particularly prevalent among people whose diet is largely corn. Pellagra is actually a multiple-deficiency disease; some cases respond to niacin alone while others respond only to mixtures of niacin and other B vitamins.

The pellagra-preventive potency of a diet is related not only to its niacin and tryptophan content but also to the actual availability of its niacin. There is evidence that some of the niacin content of foods cannot be released by digestive enzymes. The existence of an antiniacin material in corn has been suggested. Very high levels of niacin are used to lower blood cholesterol, but this type of treatment must be monitored carefully for side effects, such as flushing and liver damage. *See* NIACIN.

Folic acid. Folate and folacin are terms for compounds that have nutritional properties and chemical structures similar to those of folic acid (pteroylglutamic acid; PGA). Folates function metabolically as coenzymes that transport single carbon fragments from one compound to another in amino acid metabolism and nucleic acid synthesis. Deficiency of the vitamin leads to impaired cell division and to alterations of protein synthesis—effects most noticeable in rapidly growing tissues.

Folate is widely distributed in foods; liver, yeast, leafy vegetables, legumes, and some fruits are especially rich sources. As much as 50% of food folate may be destroyed during household preparation, food processing, and storage. Most dietary folates exist in the polyglutamate form (that is, with glutamic acids attached), which is converted in the wall of the small intestine to the monoglutamate form before it is absorbed into the bloodstream. The bioavailability of folate in a typical diet in the United States is about one-half that of crystalline folic acid, which is efficiently absorbed.

Deficiency is frequently characterized by megaloblastic anemia, a disease marked by the presence of abnormal red blood cells in the bone marrow, and macrocytic anemia, a disease marked by oversized red blood cells. Vitamin B_{12} deficiency causes a similar type of anemia that is responsive to folic acid. Therefore, administration of folate could mask a deficiency of vitamin B_{12}, which could be manifested later as neurological damage. Poor folic acid status is associated with increased birth defects (neural tube defects and cleft palate) in the fetus and the newborn. Use of a multivitamin supplement containing folic acid just before and during the first weeks of pregnancy has been shown to reduce the risk of

such defects substantially. Abnormal tissue development (dysplasia), characteristic of some precancerous conditions, can also be treated by folate supplementation. Poor folic acid status is prevalent in the newborn, adolescent females, pregnant women, and the elderly. *See* ANEMIA; FOLIC ACID.

Vitamin B_{12}. The terms vitamin B_{12} and cobalamin refer to all members of a group of large cobalt-containing corrinoids that can be converted to methylcobalamin or 5′-deoxyadenosylcobalamin, the two cobalamin coenzymes active in human metabolism. Cyanocobalamin is the commercially available form of vitamin B_{12} used in vitamin supplements and pharmaceuticals. This form is soluble in water and stable to heat. When given either orally or parenterally, it is converted by the removal of cyanide to the forms that are metabolically active in humans. In plasma and tissue, the predominant forms are methylcobalamin, adenosylcobalamin, and hydroxocobalamin.

Bacteria, fungi, and algae can synthesize vitamin B_{12}, but yeasts, higher plants, and animals cannot. In the human diet, vitamin B_{12} is supplied primarily by animal products, where it has accumulated from bacterial synthesis. Plant foods are essentially devoid of vitamin B_{12}.

Vitamin B_{12} deficiency results in macrocytic, megaloblastic anemia, in neurological symptoms due to demyelination of the nerves, and in other less specific symptoms (such as sore tongue or weakness). Neuropsychiatric manifestations of vitamin B_{12} deficiency are seen in the absence of anemia, particularly in the elderly. It is rare for a diet to be deficient in vitamin B_{12}; most of the vitamin B_{12} deficiency seen in the United States is due to inadequate absorption. Deficiency can be produced by strict vegetarian diets devoid of meat, eggs, and dairy products.

The intestinal absorption of vitamin B_{12} is mediated by a highly specific binding glycoprotein (Castle's intrinsic factor), which is secreted in the stomach. In pernicious anemia, vitamin B_{12} is not absorbed from the diet or reabsorbed from the bile due to lack of intrinsic factor activity. Vitamin B_{12} is generally injected into individuals with pernicious anemia or those who have undergone gastrectomy to prevent anemia and neuropathy. *See* VITAMIN B_{12}.

Biotin. This is a sulfur-containing B-complex vitamin that is very widespread in nature. Biotin is only sparingly soluble in water; it is stable in boiling water solutions. It can be destroyed by oxidizing agents, acids, and alkalies, and under some conditions by oxidation in the presence of rancid fats. Biotin functions as a coenzyme in carbon dioxide fixation and decarboxylation enzymes; it plays an important role in metabolism of carbohydrates, fats, and proteins.

Raw egg white contains a protein (avidin) that binds biotin, rendering it unavailable. Thus biotin deficiency can be induced by the ingestion of large amounts of avidin. This condition is characterized by anorexia, nausea, vomiting, mental depression, hair loss accompanied by a dry scaly dermatitis, and an increase in serum cholesterol and bile pigments.

Dietary deficiency of biotin is very rarely encountered except in infants. A nutritional deficiency of biotin that has been observed in infants under 6 months of age is characterized by skin disorders (seborrheic dermatitis). *See* BIOTIN.

Pantothenic acid. A member of the B-vitamin group, pantothenic acid is a light-yellow, viscous oil that is readily soluble in water; it is usually obtained as the calcium salt. Pantothenic acid is widely distributed; among its best sources are liver, kidneys, fresh green vegetables, and egg yolks. Loss of the vitamin during cooking is minimal, as it is present in stable conjugated form in food.

Pantothenic acid plays its primary physiological roles as a component of the coenzyme A molecule and of the acyl carrier protein of fatty acid synthetase. Reactions with these substances are important in the release of energy from carbohydrates; in gluconeogenesis; in the synthesis and degradation of fatty acids; and in the synthesis of such vital compounds as sterols and steroid hormones, porphyrins, and acetylcholine.

Evidence of dietary deficiency in humans has not been recognized clinically; however, experimentally produced deficiency has resulted in headache, fatigue, impaired motor coordination, paresthesia (prickling or tingling of the skin), muscle cramps, and gastrointestinal disturbances. A syndrome known as burning feet, which was observed among prisoners of war and malnourished subjects in the Far East, is believed by some authorities to represent a specific deficiency in humans, since it responds to pantothenic acid but not the other B vitamins. Pantothenic acid deficiency is difficult to identify in humans, because it is usually accompanied by general malnutrition and deficiency of the other B vitamins.

A recommended dietary allowance (RDA) has not been set, but an intake of 4–7 mg for an adult is considered safe and adequate. *See* NUTRITION; PANTOTHENIC ACID.

Lawrence J. Machlin

Bibliography. A. Bendick and C. E. Butterworth (eds.), *Micronutrients in Health and in Disease Prevention*, 1991; W. Friedrich, *Vitamins*, 1988; L. J. Machlin, *Handbook of Vitamins*, 2d ed., 1990; National Research Council, *Recommended Dietary Allowances* 10th ed., 1989; M. E. Shils and V. R. Young, *Modern Nutrition in Health and Disease*, 9th ed., 1999.

Vitamin A

A pale-yellow alcohol, soluble in fat but not in water. In pure form, it is readily destroyed by oxidation and light, which may cause losses during storage.

Source. Vitamin A is found in all animal tissues, although it is particularly concentrated in the liver. There are two different dietary sources for the vitamin: animal sources which contain vitamin A itself, mostly in the form of retinyl esters, and plant sources which contain carotenoids that are converted to

vitamin A in animal tissues such as the absorptive cells in the intestine. The most vitamin A–enriched animal food source is fish liver oil. Plant carotenoids are found in green and yellow fruits and vegetables such as carrots, apricots, asparagus, broccoli, and green leafy vegetables. Some fresh-water fish contain vitamin A_2, which differs slightly from vitamin A in structure. *See* CAROTENOID.

Bioassay. The vitamin A activity of carotenoids varies with their chemical structure, with β-carotene being the most potent. One international unit (IU) of vitamin A has been set at 0.3 microgram of vitamin A or 0.6 μg of β-carotene. This is somewhat confusing since the efficiency of conversion of β-carotene to vitamin A becomes greater in a deficiency state. Although biological assays are sometimes used, vitamin A and carotene are usually determined by spectrophotometric techniques. *See* MASS SPECTROMETRY.

Physiological activities. In vitamin A deficiency, the epithelial tissues of many organs are affected. Growth failure occurs, and young animals can suffer from neurological symptoms resulting from pressures on the central nervous system. Changes occur in the skin, mouth, respiratory tract, urogenital tract, and some glands. Vitamin A deficiency is also strongly associated with depressed immune function and higher morbidity and mortality due to infectious diseases such as diarrhea, measles, and respiratory infections. A severe manifestation of vitamin A deficiency is night blindness and inflammation of the eyes (xerophthalmia), followed by irreversible blindness. The symptoms seen in vitamin A deficiency reflect the multiple roles of this compound in animals. These roles are fulfilled by two compounds that are synthesized from vitamin A in the body: vitamin A aldehyde (retinaldehyde), which is critical for vision, and vitamin A acid (retinoic acid), which controls many physiological functions in both the embryo and the adult. The similar chemical structures of vitamin A, retinaldehyde, and retinoic acid are shown at right.

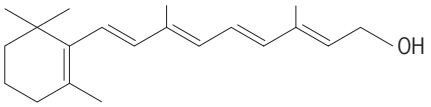

Vitamin A: all-*trans*-retinol

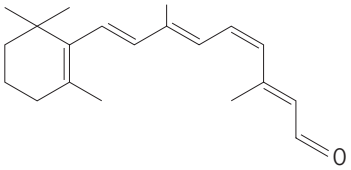

11-*cis*-Retinaldehyde

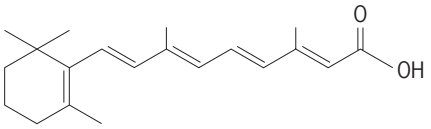

All-*trans*-retinoic acid

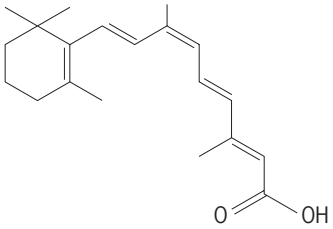

9-*cis*-Retinoic acid

Retinaldehyde. Retinaldehyde has a critical role in sight. This compound binds to a protein termed opsin to generate the visual pigment rhodopsin in the retina. Rhodopsin is the "visual antenna" that responds to light by sending signals to the brain. Following absorption of light, rhodopsin releases the bound vitamin A compound. Resynthesis of rhodopsin by providing fresh molecules of retinaldehyde is therefore essential for normal vision. *See* VISION.

Retinoic acid. Retinoic acid is the active vitamin A metabolite that is critical for many processes, such as cell growth and differentiation, in the embryo and in the adult. Retinoic acid exerts its effects by modulating the expression of many genes. This activity is due to the ability of the compound to control the activities of proteins known as retinoid nuclear receptors. When activated by retinoic acid, these proteins bind to deoxyribonucleic acid (DNA) and regulate gene transcription. Retinoic acid is currently used in treatment of various pathologies such as skin disorders and certain types of cancer. *See* DEOXYRIBONUCLEIC ACID (DNA).

Requirements. Studies of many mammalian species suggested that approximately 20 IU (6 μg) of vitamin A per kilogram of body weight will support growth and prevent symptoms of deficiency. The current intake recommendations of vitamin A in the United States is 3 mg/day and about 1 mg/day in the European Union. Vitamin A deficiency is a serious public health problem in many countries but is essentially nonexistent in western Europe and the United States. The excessive intake of vitamins has resulted in cases of vitamin A toxicity in people receiving 25 to several hundred times the requirement over prolonged periods. Discontinuance of vitamin A supplementation in these individuals is usually followed by a dramatic disappearance of symptoms of toxicity in a few days. *See* VITAMIN.

Stanley N. Gershoff; Noa Noy

Bibliography. D. Mangelsdorf, K. Umesono, and R. M. Evans, The retinoid receptors, in M. B. Sporn et al. (eds.), *The Retinoids: Biology, Chemistry, and Medicine*, Raven Press, New York, 1994; N. Noy, Vitamin A and carotenoids, in M. H. Stipanuk (ed.), *Biochemical and Physiological Bases of Human Nutrition*, W. B. Saunders, Philadelphia, 1999; S. A. Ross et al., Retinoids in embryonal development, *Physiol. Rev.*, 80:1021–54, 2000; R. D. Semba., Vitamin A, immunity and infection, *Clin. Infect. Dis.*, 19:489–499, 1994.

Vitamin B_6

A vitamin which exists as three chemically related and water-soluble forms found in food, as shown here. All three forms have equal activity for animals

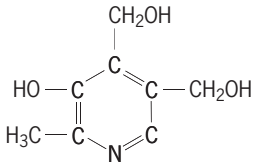

Pyridoxine (pyridoxol)

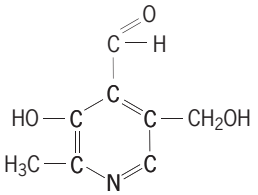

Pyridoxal

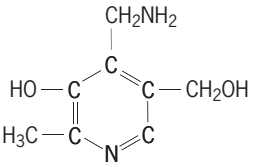

Pyridoxamine

and yeast, but pyridoxal and pyridoxamine have several thousand times the activity of pyridoxine for some bacteria. Vitamin B_6 is stable in acid, but is rapidly destroyed by light in alkaline or neutral solutions.

Bioassay. Chemical assays for vitamin B_6 are laborious; microbiological assays are commonly used. They often give results which appear much too high when compared to animal-feeding tests. This is because vitamin B_6 is conjugated with other substances in foods, and the acid hydrolysis procedures used in preparing samples for microbiological assay release quantities of the vitamin not released by normal digestive processes. *See* BIOASSAY.

Dietary requirements. Most studies of vitamin B_6 deficiency have been done on animals, since a deficiency of this vitamin in humans is very rare. Vitamin B_6 deficiency is accompanied by poor growth, dermatitis, microcytic anemia, epileptiform convulsions, and kidney and adrenal lesions. There is evidence that some women in the third trimester of pregnancy may have a special requirement for vitamin B_6 in that its administration often relieves the nausea of pregnancy. Some types of human dermatitis respond to local application of this vitamin. Experimental studies have suggested that vitamin B_6–deficient animals are deficient in available insulin and growth hormone. Marked abnormalities in lipid, carbohydrate, and protein metabolism are caused by this deficiency.

Vitamin B_6–deficient animals are oxaluric and may form kidney stones. Atherosclerosis and impaired antibody formation have been reported. Vitamin B_6–responsible anemias have been reported in humans, and it has been suggested that supplemental vitamin B_6 may protect against dental caries, particularly during pregnancy.

It is difficult to set requirements for vitamin B_6, since no single set of assay conditions or criteria has received universal acceptance. Based on animal experiments, a number of dietary factors probably affect the vitamin B_6 requirement. High-protein diets increase the requirement. High-carbohydrate diets probably increase intestinal synthesis of vitamin B_6, while unsaturated fatty acids decrease the requirement. Adults probably require about 1.5–2 mg per day, and a dietary intake of 0.4 mg per day would probably be satisfactory for most infants.

Biochemistry. Vitamin B_6 functions as a coenzyme in the form of pyridoxal 5-phosphate. The value of pyridoxine and pyridoxamine lies in the ability of tissues to convert them into pyridoxal. Vitamin B_6–containing enzymes are included in aminotransferase reactions (transamination), important mechanisms for the synthesis of amino acids by the tissues from α-keto acids. An example is shown below.

Glutamic acid + pyruvic acid $\rightleftharpoons$

α-ketoglutaric acid + alanine

Vitamin B_6 is also involved in amino acid decarboxylation reactions. The transformations of histidine to histamine and of aspartic acid to β-alanine are examples of these. This vitamin has a special function in the metabolism of tryptophan and is necessary for the conversion of tryptophan to niacin. In most vitamin B_6–deficient animals, xanthurenic acid, a metabolite of tryptophan, is excreted in abnormal quantities, and this has been used as the basis for tests of the adequacy of vitamin B_6 nutrition. This vitamin also has a special function in the metabolism of the sulfur-containing amino acids. *See* COENZYME; TRANSAMINATION; VITAMIN. Stanley N. Gershoff

Bibliography. M. H. Briggs, *Vitamins in Human Biology and Medicine*, 1981; S. S. Hendler, *Vitamin and Mineral Encyclopedia,* 1990; G. B. Tryfiates, *Vitamin B_6: Metabolism and Its Role in Growth*, 1980.

Vitamin B_{12}

A group of closely related polypyrrole compounds containing trivalent cobalt (see **illus.**); often called cobalamin. The vitamin is a dark-red crystalline compound; in aqueous solution and at room temperature it is most stable at pH 4–7.

Source and assay. In general, vitamin B_{12} is synthesized by microorganisms, not by plants, and is found in animal tissues as a result of intestinal synthesis or ingestion. It is now thought that the animal-protein factor, which was studied widely, is at least in part vitamin B_{12}. Vitamin B_{12} is usually determined microbiologically, although rat and chick growth assays are also available. *See* BIOASSAY.

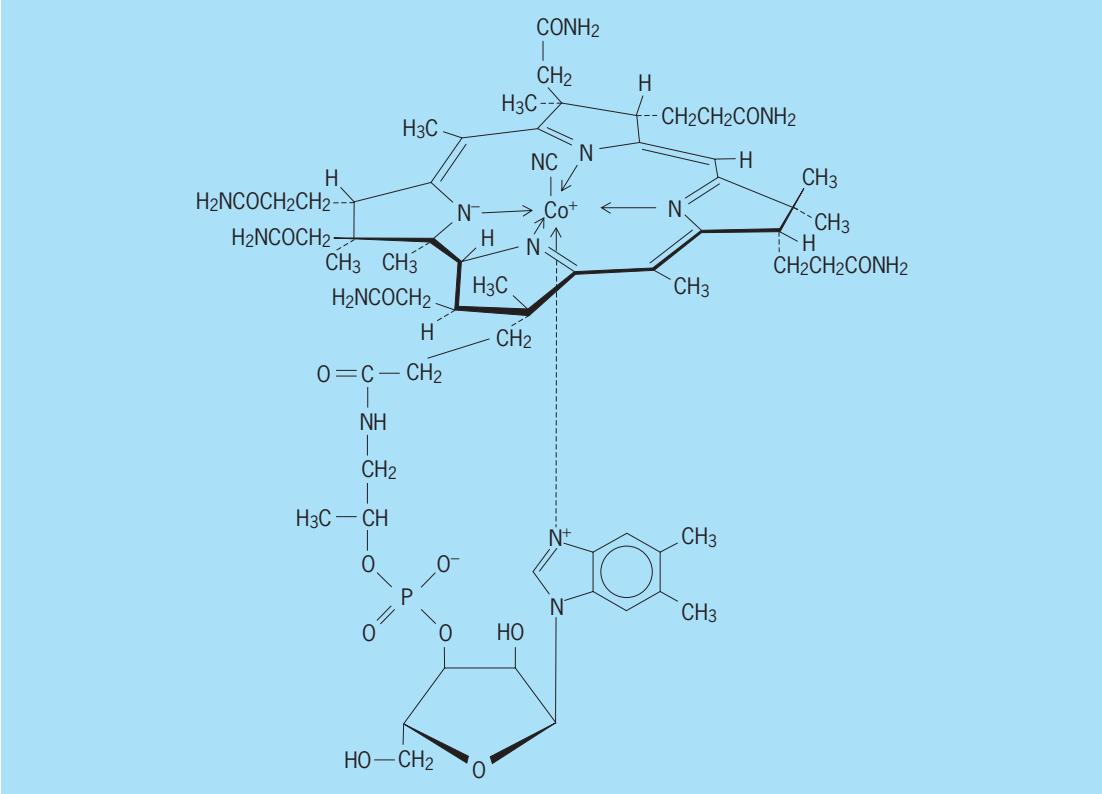

Structural formula for vitamin B_{12}.

Physiology and requirements. Vitamin B_{12} deficiency in animals is characterized primarily by anemia and neuropathy. In humans, this deficiency is called pernicious anemia. People suffering from this disease lack a factor secreted in normal gastric juice which, by affecting absorption directly and by protecting vitamin B_{12} from intestinal destruction, enables the vitamin to be absorbed. This factor, called intrinsic factor, appears to be one or more mucoproteins. Its action in facilitating the absorption of vitamin B_{12} is calcium-dependent. Vitamin B_{12} deficiency in humans may be caused by a genetically controlled lack of intrinsic factor, gastrectomy, malabsorption disease, or vegetarianism. Body stores of the vitamin are such that it takes approximately 5 years for a deficiency state to appear clinically following gastrectomy.

Requirements for vitamin B_{12} are increased by reproduction or hyperthyroidism. Of the known vitamins, B_{12} is the most active biologically. A daily injection of 1 microgram of vitamin B_{12} will prevent the recurrence of symptoms in people with pernicious anemia. For normal people a diet containing 3–5 μg per day (providing 1–1.5 μg absorbed) will satisfy vitamin B_{12} requirements. Some of the vitamin B_{12} requirement is probably met by intestinal synthesis.

Biochemistry. Several coenzymes of vitamin B_{12} exist. They are included in isomerization reactions (methylmalonyl–CoA mutase, glutamate mutase), conversion by dioldehydrases of 1,2-diols to deoxyaldehydes, methylation reactions (formation of methionine and thymidine 5′-phosphate), and conversion of ribose to deoxyribose. *See* VITAMIN.

Stanley N. Gershoff

Bibliography. S. S. Hendler, *Vitamin and Mineral Encyclopedia*, 1990.

Vitamin D

Either of two fat-soluble sterol-like compounds, ergocalciferol (vitamin D_2) and activated cholecalciferol (vitamin D_3). Vitamin D_2 is formed from the irradiation of ergosterol, a plant sterol. However, vitamin D_3 is normally manufactured in the skin, where ultraviolet light activates the compound 7-dehydrocholesterol. Vitamins D_3 and D_2 are about equal in activity in all mammals except New World monkeys and birds, in which vitamin D_2 is approximately one-tenth as active as vitamin D_3. *See* VITAMIN.

Assay. In crystalline form, vitamin D compounds are determined by their ultraviolet absorption maximum by means of spectrophotometers. However, at concentrations normally found in foods and some vitamin pills, this method is not sufficiently sensitive. Generally, high-performance liquid chromatography is used to determine vitamin D levels in high-potency pills and other concentrates and food. Adequacy of vitamin D production in skin is assessed by the measurement of its major circulating metabolite, 25-hydroxyvitamin D. This is done by

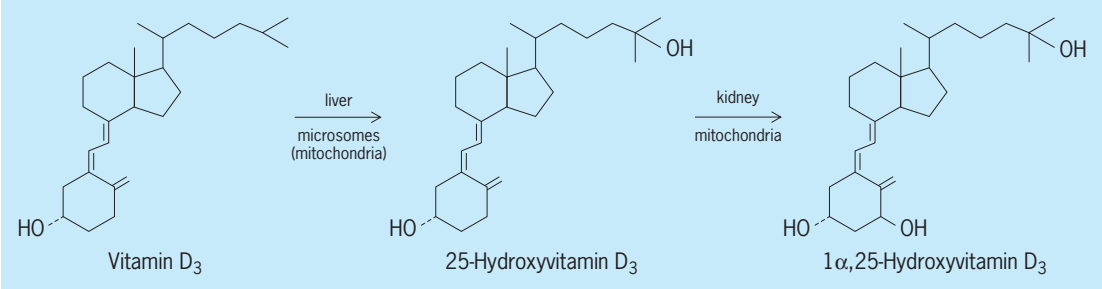

Conversion of vitamin D to its active form in the body. Vitamin D_3 can be obtained from the diet, or it can be produced in skin by the ultraviolet irradiation of 7-dehydrocholesterol. 1α,25-Dihydroxyvitamin D_3 is believed to be the only active form of vitamin D in the body.

immunoassay, radioreceptor assay, or high-performance liquid chromatography. The active form of vitamin D, 1, 25-dihydroxyvitamin D_3, can be determined only by micromethods, such as radioreceptor assay, radioimmunoassay, or reporter gene cellular assay. Normal plasma levels of vitamin D_3 are 1–2 μg/mL of 25-hydroxyvitamin D, 15–20 μg/mL and of 1,25-diphydroxyvitamin D_3, 35–40 picograms/mL. *See* BIOASSAY; IMMUNOASSAY; LIQUID CHROMATOGRAPHY; RADIOIMMUNOASSAY.

Biochemistry. Vitamin D as acquired from the diet or produced in the skin is biologically inactive. It must be metabolized by the liver to produce 25-hydroxyvitamin D_3. However, this compound is also biologically inactive under physiological circumstances and must be activated by the kidney to produce the final vitamin D hormone, 1,25-dihydroxyvitamin D_3 (see **illus.**). This hormonal form of vitamin D plays an essential role in stimulating intestinal absorption of calcium and phosphorus, in the mobilization of calcium from bone, and in renal reabsorption of calcium. This hormone acts very much like other steroid hormones in that it binds to a nuclear receptor that stimulates the cell to produce proteins that carry out the functions of vitamin D. The molecular details of the mechanisms that these proteins utilize to increase absorption of calcium and phosphorus, mobilize calcium from bone, and activate renal reabsorption of calcium remain under investigation.

The function of vitamin D has been expanded beyond regulating plasma calcium and phosphorus levels, and hence healing the diseases of rickets and osteomalacia. It is now known that the vitamin D hormone controls parathyroid gland growth and production of the parathyroid hormone. It is an immunomodulator and is being developed for use in autoimmune diseases. Vitamin D hormone also appears to play a role in the regulation of insulin production or secretion. It is also used to treat the skin disorder psoriasis. These new sites of action of vitamin D are under intense investigation. *See* HORMONE.

Sources. Vitamin D is largely absent from the food supply. It is found in large amounts in fish liver oils; cod liver oil has long been known to be an important source of vitamin D. Except for unusual circumstances, it is not found in the plant world. It is in low concentrations in milk that is not fortified and in eggs because it is stored in the egg yolk to support growth and development of the chicken embryo. Fortified foods are the major dietary source of vitamin D, but the major overall source is the production of vitamin D in skin by exposure to sunlight or ultraviolet irradiation. In winter months at temperate latitudes, insufficient amounts of vitamin D are produced in skin, and unless it is replaced by a dietary source, danger of insufficiency exists.

Deficiency. A deficiency of vitamin D in growing animals results in the disease rickets, which is marked by a failure to mineralize the skeleton because of insufficient amounts of calcium and phosphorus in the plasma due to insufficient absorption from the gut and insufficient reabsorption in the kidney. A similar disease, osteomalacia, occurs in adult animals. Osteomalacia is characterized by a failure to calcify newly forming organic matrix of bone. By far the most serious disorder of vitamin D deficiency is the low-blood calcium levels which result in convulsions known as hypocalcemic tetany. Under circumstances of low-blood calcium levels, nerves continuously excite muscle, giving rise to continuous muscle contractions or convulsions. Moderate deficiency of vitamin D may contribute to osteoporosis, especially in the elderly. *See* BONE; OSTEOPOROSIS.

Toxicity. When doses of vitamin D exceed 5000–10,000 IU per day, danger of intoxication is present. Intoxication results from high plasma calcium concentrations that lead to calcification of kidney, heart, aorta, and subcutaneous tissues. Large doses of vitamin D can be lethal. Doses above 800 IU per day should never be taken unless prescribed by a physician.

Dietary requirements. The recommended daily requirement for vitamin D_3 is 10 micrograms or 400 international units (IU). Higher requirements are reported for the elderly and for rapidly growing adolescents: 20 μg or 800 IU per day. It is possible that the average requirement is lower than 10 μg per day. The exact absolute requirement has never been determined.

Therapeutic applications. The major applications of 1,25-dihydroxyvitamin D_3 and its analogs are in treatment of the bone disease associated with kidney

failure. It is also used in many countries to treat the bone-thinning disease, osteoporosis. For example, in Japan, it and its analog, 1α-hydroxyvitamin D_3, are the major treatments for osteoporosis. It is also used for the treatment of some genetic vitamin D–resistance diseases and for patients who lack parathyroid glands. More recent applications are for certain autoimmune diseases and for the treatment of the skin disorder psoriasis. Hector F. Deluca

Bibliography. D. Feldman, F. H. Glorieux, and J. W. Pike (eds.), *Vitamin D*, Academic Press, San Diego, 1997.

Vitamin E

A group of compounds, α, β, γ, and δ tocopherols, that have a chromanol ring and phytyl side chain and are widely distributed in nature, especially in edible vegetable oils (wheat germ, sunflower, cottonseed, safflower, canola, soybean, and corn oil). Unprocessed grains, nuts, and vegetables are other sources. When used as supplements, there are two vitamin E products available: natural source RRR α-tocopherol and synthetic all-rac-α-tocopherol. The

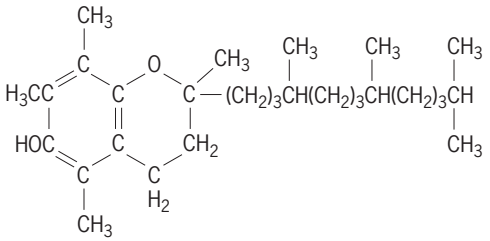

α-Tocopherol

latter is a mixture of eight different stereoisomers, of which only one is RRR α-tocopherol.

Physiological activities. The terms "vitamin E" and "α-tocopherol" are frequently used interchangeably in human nutrition, but it is imperative to distinguish between supplements of RRR α-tocopherol and of synthetic α-tocopherol because their biological activity is different. Since the major function of vitamin E is to serve as a chain-breaking antioxidant, protecting cell membranes against free-radical damage, the most potent form of the vitamin should be used as a supplement. Although gastrointestinal absorption of all forms of vitamin E is equivalent, the subsequent physiological steps are sharply in favor of the RRR form. This action is mediated by a cellular liver transfer protein that is specific for the RRR form of α-tocopherol. It maintains the plasma level by selectively choosing the RRR form and recycling it into plasma lipoproteins for distribution of the vitamin to every tissue and organ in the body. The α-tocopherol transfer protein is so efficient in preserving circulating α-tocopherol that vitamin E deficiency is a rarity among humans except in diseases of malabsorption, lipoprotein abnormalities, or mutations in the gene responsible for the synthesis of the α-tocopherol transfer protein.

Deficiency. When deficiencies of the vitamin occur in humans as a consequence of acquired malabsorption of genetic abnormalities of lipoproteins or of the transfer protein, the major symptoms that develop are in the nervous system. Ataxia (lack of muscular coordination) and other neurologic symptoms result in severe incoordination and subsequent musculoskeletal changes. Fortunately, adequate supplementation with huge doses of vitamin E, when used early enough, can prevent the neurologic syndrome from developing, or can at least stabilize the neuropathology that has developed and prevent further deterioration.

Dietary requirements. The recommended daily allowance for vitamin E is 15 milligrams, which is present in the usual Western diet. Reports of studies with supplemental vitamin E are confusing because they do not specify whether natural or synthetic α-tocopherol is employed and the dose used is frequently in International Units (IU). (Since IU were originally derived from biological testing in rodents, a weight basis in milligrams is more precise.) Although it has been traditionally accepted that the biologic effective relation between natural and synthetic vitamin E is 1.36:1 (and 1 mg RRR α-tocopherol equals 1.49 IU), the true ratio is more likely to be 2:1.

Therapeutic applications. Some of the pathology of a number of degenerative brain diseases have been ascribed to oxidative stress in the central nervous system. These include Alzheimer's disease. Down syndrome, and perhaps even Parkinson's disease and Friedreich's ataxia. The possibility of favorably affecting the development and course of these diseases has made supplemental vitamin E therapy an attractive choice. Studies of treatment programs of 150 mg of the natural-source vitamin E show higher stable plasma levels without any signs of toxicity. *See* ALZHEIMER'S DISEASE; DOWN SYNDROME; PARKINSON'S DISEASE; VITAMIN. Herbert J. Kayden

Bibliography. *Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids*, Food and Nutrition Board, Institute of Medicine, National Academy Press, Washington DC, 2000; H. J. Kayden and M. G. Traber, Absorption, lipoprotein transport, and regulation of plasma concentrations of vitamin E in humans, *J. Lipid Res.*, 34:343, 1993.

Vitamin K

A group of compounds derived from 2-methyl-1,4-naphthoquinone that prevent bleeding in mammals and birds. Vitamin K_1 (phylloquinone) is produced by green plants; a related form, vitamin K_2 (menaquinone), is produced by intestinal bacteria. Chemically synthesized forms include vitamin K_1, K_2, and menadione (vitamin K_3). All of these compounds are fat-soluble liquids at room temperature that become biologically inactive when exposed to light or alkali. *See* VITAMIN.

Physiological activities. Vitamin K_1 (phylloquinone) is a photosynthesis cofactor in plants, and

various forms of vitamin K_2 (menaquinones) participate in energy transfer reactions in bacteria (**illus.** *a*). Mammals and many other animals need vitamin K hydroquinone (the active form) as a cofactor for the specific synthesis of the amino acid γ-carboxyglutamate (Gla) in certain proteins, which enable the proteins to bind calcium and phospholipids with high specificity. The synthesis of one Gla residue generates one vitamin K 2,3-epoxide molecule, which may be reduced again to its hydroquinone form for the next reaction (see illus.). Gla is an essential part of coagulation factors and of other proteins that regulate blood clotting (hence the disruption of blood clotting and internal bleeding associated with vitamin K deficiency). Two Gla-containing proteins (osteocalcin and matrix Gla protein) control mineralization of bone and arterial walls; others, such as gas6, influence cell growth. Gla proteins have also been found in corals and in the toxins of certain marine snails; however, their exact role in these organisms has not been defined. *See* AMINO ACIDS; BLOOD; COENZYME; PROTEIN.

Analysis. Vitamin K in biological samples can be measured with great specificity by highperformance liquid chromatography (HPLC) or gas chromatography–mass spectrometry (GC-MS), usually after extraction and prepurification. Because the concentrations in most biological samples are low (1–500 picomoles per liter in human tissue), detection methods have to be very sensitive. Vitamin K sufficiency can be assessed with immunologic and other assays that detect the lack of Gla residues in specific vitamin K–dependent proteins such as prothrombin and osteocalcin. *See* CHROMATOGRAPHY; IMMUNOASSAY.

Requirements and deficiency. Humans depend on continuous vitamin K supplies, since storage is minimal. Good dietary sources of vitamin K_1 are green vegetables and fruits; certain fermented Asian foods, especially natto, have a high vitamin K_2 content. There is no indication that the consumption of large amounts of vitamin K, especially in the form of natural foods, has unfavorable effects on health.

Intestinal bacteria produce vitamin K_2, and under most circumstances enough is absorbed to prevent bleeding. However, spontaneous bleeding occurs if both dietary intake and production by intestinal bacteria are persistently low. In many countries, newborn infants routinely receive one-time injections of vitamin K to reduce their risk of cerebral hemorrhage. Other health risks related to inadequate vitamin K intake may include accelerated loss of bone minerals and hardening of arteries.

Antagonists. Various natural and synthetic 4-hydroxycoumarins are potent vitamin K antagonists that inhibit regeneration of the active hydroquinone form. Vitamin K antagonists are commonly used in medicine to prevent thrombosis related to operations, damaged heart valves, or abnormal heart rhythm. The antagonists have to be individually dosed, and people using this type of medication must keep their vitamin K intake constant. Antagonists are also used as a slow-acting rodent poison. The effects of accidental or intentional poisoning can be blunted by administration of vitamin K. *See* ENZYME INHIBITION; THROMBOSIS.

Martin Kohlmeier

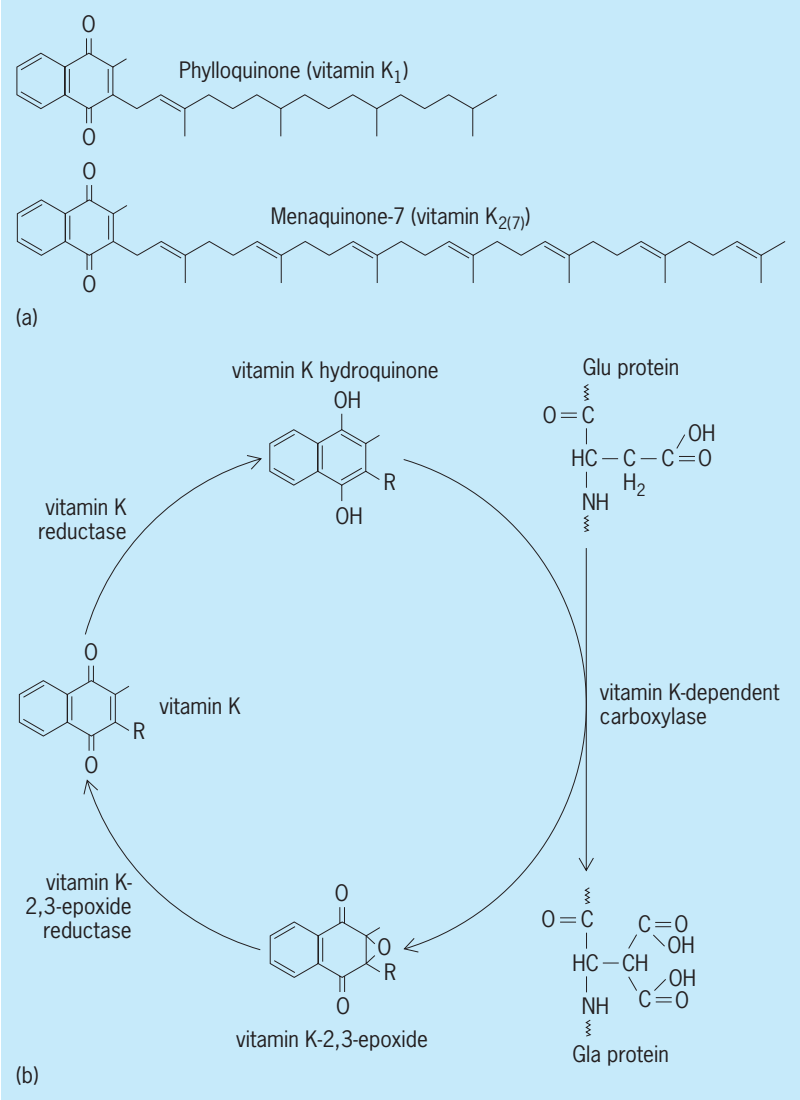

Vitamin K. (*a*) Natural forms. (*b*) Regeneration of the hydroquinone for γ-carboxyglutamate (Gla) synthesis; Glu = glutamic acid. Vitamin K functions by carboxylating certain glutamate residues in a few proteins. With each carboxylation reaction, a vitamin K molecule becomes oxidized. A microsomal enzyme system in liver, bone, and other tissues regenerates the active form, vitamin K hydroquinone. Phylloquinone and the menaquinones from bacteria work equally well in these reactions. They differ only in length and degree of saturation of their side chains; the redox-active ring system is the same.

Bibliography. J. J. Lipsky, Nutritional sources of vitamin K, *Mayo Clinic Proc.*, 69:462–466, 1994; M. J. Shearer, Vitamin K, *Lancet*, 345:229–234, 1995; M. J. Shearer, A. Bach, and M. Kohlmeier, Chemistry, nutritional sources, tissue distribution and metabolism of vitamin K with special reference to bone health, *J Nutr.*, 126:1181S–1186S, 1996.

Vivianite

A hydrated ferrous phosphate mineral, with the formula $Fe_3(PO_4)_2 \cdot 8H_2O$, for which the vivianite group is named. Other important members of this

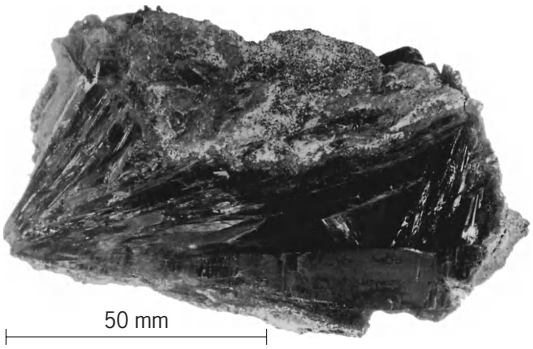

Vivianite from Wannon River, Victoria, Australia. (*American Museum of Natural History specimen*)

group are annabergite and erythrite.

Usually ferric iron is present as the result of oxidation, although this does not change the crystal structure appreciably. It crystallizes in the monoclinic system, with crystals generally prismatic (see **illus.**). Vivianite also occurs in earthy form and as globular and encrusting masses of fibrous structure. Crystals are colorless and transparent when fresh. Oxidation changes the color progressively to pale blue, greenish blue, dark blue, or bluish black.

Vivianite is widespread as a secondary mineral in gossans of metallic ore deposits, in weathered Mn-Fe phosphatic pegmatites, and in alluvial and sedimentary deposits associated with bone and other organic remains. As odontolite in fossil bone or teeth, it is often mistaken for turquois.

Erythrite and annabergite are hydrated arsenates of cobalt and nickel, $(Co,Ni)_3(AsO_4)_2{\cdot}8H_2O$. Cobalt and nickel substitute mutually to form a complete series between the end members $Co_3(AsO_4)_2{\cdot}8H_2O$ and $Ni_3(AsO_4)_2{\cdot}8H_2O$. Erythrite includes the half of the series with Co > Ni, while annabergite includes the half with Ni > Co. Both minerals crystallize in the monoclinic system. Crystals, although rare and of poor quality, are prismatic to acicular, and transparent to translucent. Erythrite commonly occurs as radial groups, globular shapes, and earth masses. Annabergite develops as finely crystalline crusts and as earthy masses. Erythrite is crimson red and peach red. Increasing the Ni content changes the color progressively to pale pink, pale green, and apple green in annabergite.

These secondary minerals develop during oxidation of cobalt and nickel arsenides. Erythrite is the more common; both occur in Europe and North America. Wayne R. Lowell

Vocal cords

The pair of elastic, fibered bands inside the human larynx. The cords are covered with a mucous membrane and pass horizontally backward from the thyroid cartilage (Adam's apple) to insert on the smaller, paired arytenoid cartilages at the back of the larynx. The vocal cords act as sphincters for air regulation and may be vibrated to produce sounds. Separation, approximation, and alteration of tension are produced by action of laryngeal muscles acting on the pivoting arytenoids. Innervation is through branches of the vagus nerve. Vibration of the cords produces fundamental sounds and overtones. These can be modified by the strength of the air current, the size and shape of the glottis (the opening between the cords), and tension in the cords. *See* LARYNX.

Among mammals only humans produce an elaborate articulate speech, although the basic structures for sound production are similar in all and rudiments of speech exist in many primates and other mammals. Sound in birds is produced in the syrinx, located at the ventral end of the trachea; many species are excellent mimics and some, such as the parrots, can achieve a high degree of speechlike vocalization. *See* ANIMAL COMMUNICATION; SPEECH. Walter Bock

Voice over IP

A technology that transports voice using data packets instead of circuit switching, which the traditional public switched telephone network (PSTN) uses. Voice over IP (VoIP), using packet technology, allows for more efficient transport of voice while providing the quality of service and reliability of PSTN.

Comparison of VoIP and PSTN. The present public switched telephone network transfers voice by converting speech received from the telephone into 64-kbps (kilobits per second) digital data and transporting it in a timeslot (limited time interval) that is periodically inserted in a higher-capacity signal, a procedure referred to as time-division multiplexing (TDM). Two timeslots, one for each direction, are allocated for each phone call. These timeslots are set up by the signaling function of the public switched telephone network and kept in use for the duration of the call. The timeslots are switched in and out of use based on the calls taking place at any point in time, and are referred to as being circuit-switched. *See* TELEPHONE SERVICE.

Although the public switched telephone network can carry nonvoice data, it does so with a modem. The modem fits into the analog bandwidth of a standard telephone, allowing it to fit into the same two 64-kbps timeslots required by voice. Again these timeslots are reserved for use during the length of the call and required even when an end user does not have data to send or receive, a procedure which is not very efficient for data service, which is typically asymmetrical and bursty.

By contrast, voice over IP uses IP packets to carry speech over a data network. Only those packets that contain speech need to be transported, thereby allowing voice over IP to improve bandwidth efficiency by transporting packets only in the direction of the call participant who is listening. To further improve efficiency, voice over IP uses speech-compression algorithms to reduce speech from 64 to 2.4–8 kbps. This helps to offset the additional bytes of information (overhead) required by the packet headers, which contain the information needed to route, check, and reassemble the packets

(**Fig. 1**). These packets are transmitted every 10 milliseconds, resulting in a bandwidth of 48 kbps during speech. *See* DATA COMMUNICATIONS; PACKET SWITCHING.

Unlike the public switched telephone network using time-division multiplexing, the IP network does not allocate specific timeslots for a particular user, although newer technology does make it possible to guarantee bandwidth over the network. This allows the IP network to take advantage of silent periods of a normal call by not sending any packets, further increasing its efficiency.

While the public switched telephone network was developed for voice and backfitted for data, IP was developed for data and is being backfitted for voice. The ubiquity of IP allows the convergence of new services dependent on voice and data. Voice over IP is being expanded to support applications such as videophone, video conferencing, and whiteboard conferencing (teleconferencing involving the use of whiteboards, which are areas on display screens on which multiple users can write or draw). *See* TELECONFERENCING; VIDEOTELEPHONY.

VoIP-PSTN interoperation. Voice over IP can take place over any portion of the transmission path of a particular call, even over the public switched telephone network. In its original instantiation, voice over IP used the public switched telephone network by compressing speech, placing it into IP packets, and transporting it from one computer to another using a modem signal. Today service providers are using packet technology in addition to the public switched telephone network for transmitting voice, using gateways to provide the interface between circuit-switched and packet-switched technology.

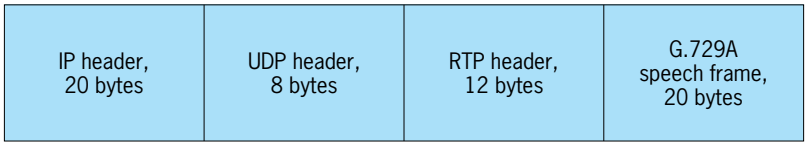

Fig. 1. Internet Protocol (IP) packet with voice over IP speech frame (encoded data) and headers. UDP = user datagram protocol. RTP = real-time protocol. G.729A is an audio data compression algorithm for voice.

The most critical development has been in technology that allows the public switched telephone network to communicate with the IP network, transforming an IP call to a public switched telephone network call, and vice versa. **Figure 2** is an example of a network that combines voice over IP and the public switched telephone network. In this example, a standard analog phone user, caller A, calls an IP phone user, caller B. Once the number is dialed, the switch processing the number uses the Signaling System 7 (SS7) protocol to send this request to the signaling gateway, which converts the SS7 messages into a Session Initiation Protocol (SIP) message. The call proceeds to caller B, who answers the call on either an SIP telephone or a personal computer running SIP client software. Once the media gateway controller detects that caller B picked up the telephone, it enables the packet voice gateway. Caller B's telephony equipment (either a voice over IP phone or a personal computer) negotiates capabilities between itself and the media gateway in the form of security and speech coding algorithms. At the completion of these negotiations, the connection between the media gateway and caller B is completed and voice is carried between callers A and B. The packet voice gateway or media gateway converts time-division multiplexing used on the public

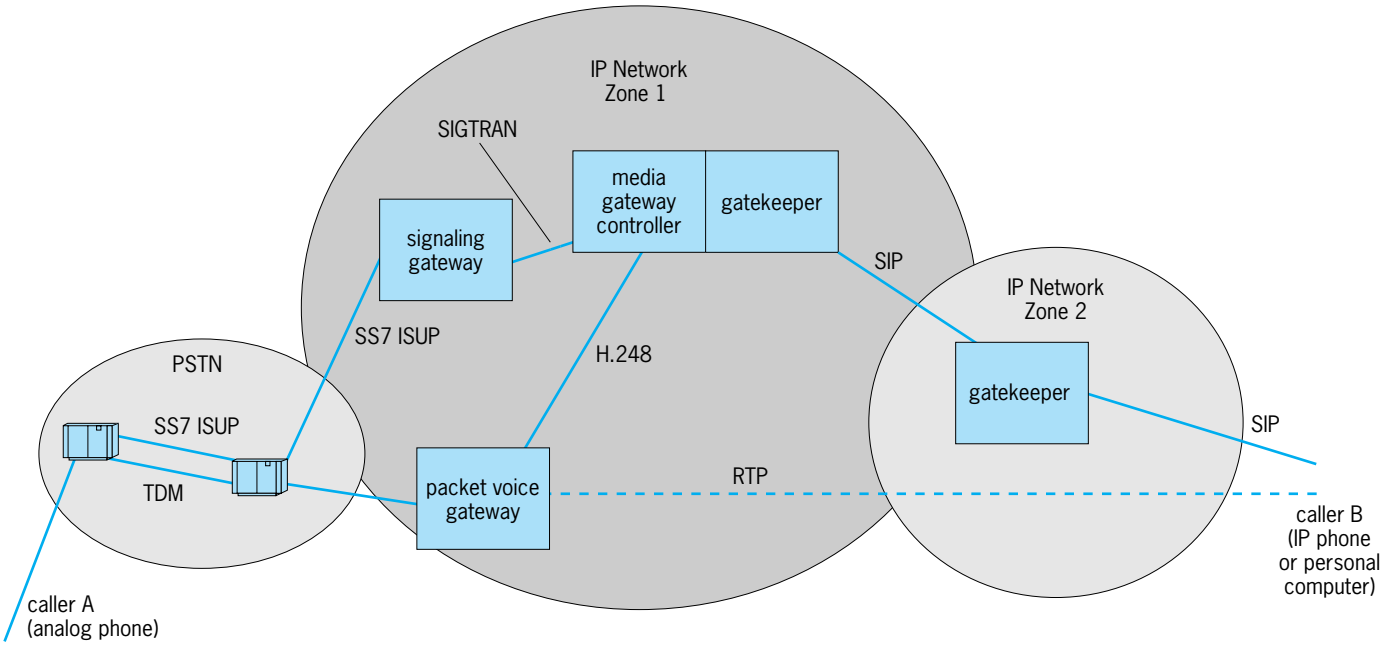

Fig. 2. Example of network operation between voice over IP and the public switched telephone network (PSTN). SS7 = Signaling System 7. ISUP = ISDN User Part. TDM = time-division multiplexing. SIGTRAN = Signaling Transport. RTP = real-time protocol. SIP = Session Initiation Protocol.

switched telephone network to and from voice being carried in real-time protocol (RTP) packets in the Internet network. *See* INTERNET.

This example happens to use SIP for the voice over Internet Protocol within the Internet network, but either the Media Gateway Control Protocol (MGCP) or H.323 might also be used. The Media Gateway Control Protocol attempts to look very similar to the public switched telephone network by adding services and control information to a central network, allowing the endpoints to be dumb. Both H.323 and SIP use a distributed architecture with intelligent endpoints and call control elements such as signaling gateways and gatekeepers, which help make sure that a call is routed to the appropriate destination. The distributed model provides for increased flexibility of applications, while the centralized approach provides an easier management of call control and provisioning (programming or instructing equipment to operate in a desired way).

Other common examples for voice over IP network configurations are end-to-end, business or enterprise services, and trunking. End-to-end, which was the initial application for voice over IP, is typically associated with low-quality calls, as the calls often go over the public switched telephone network with the end users being bandwidth-constrained. Business or enterprise applications are associated with advanced services, and their operation over high-speed local-area networks provides plenty of bandwidth. Trunking is the most recent application for voice over IP and allows service providers to transport efficiently both voice and data over their networks.

Security. Most people think of the public switched telephone network as being secure (which it is, relatively speaking), since the network resources are dedicated for the duration of a call. Packet networks or the Internet are much less secure. They are more often compared to party lines, where a piece of equipment is told to listen only for packets intended for it, but can operate in a promiscuous mode where all packets on the network can be received. This is undesirable for end users, but even more problematic is the ease with which hackers have been able to attack IP networks.

Internet Protocol networks have been broken into worldwide, and in numerous cases caused equipment to be taken off line or rendered temporarily useless. In order to build reliability into a voice over IP network, service providers must be able to increase its security level to the level of the public switched telephone network. Numerous tools and protocols have been developed for helping to secure IP networks such as firewalls, intrusion detection and prevention elements, virus scanners, and virtual private network (VPN) appliances. IP Security (IPSec), Secure Sockets Layer (SSL), and voice over IP–specific protocols are being developed, but new advances in voice over Internet Protocols and applications constantly pose new challenges to IP networks that must repeatedly upgrade their security measures to prevent network outages.

Quality of service. The public switched telephone network based on time-division multiplexing has fixed delays, while IP networks have variable delays based on the path that the data take through the network and the degree of congestion of the network. These issues can disrupt the voice by causing packets to arrive out of order or be dropped. Voice over IP software is designed to make adjustments for out-of-order packets, and speech algorithms help to adjust for lost packets, but they cannot overcome these impairments by themselves.

When delays become too great, the normal flow of a conversation is disrupted. For example, when one party finishes talking there will be an inordinate amount of time between the end of speech and a response from the far end. To address this issue, protocols [such as Multiprotocol Label Switching (MPLS) and differentiated services] are being implemented to allow for the transport of delay-sensitive data to improve performance. These protocols use information contained within packet headers to help prioritize packets so that those carrying speech (or other delay-sensitive traffic) can be routed prior to other traffic. In the case of MPLS, more overhead is added to the packet header.

Companies are also starting to increase the available bandwidth for Internet traffic and limiting the amount of traffic that can go across a particular link to improve real-time performance. When few people are trying to access or use any link, the delay is small enough not to be noticeable; but when more data go across the network, delays increase to the point where call quality is unacceptable. Home users are improving performance by using cable modems and digital subscriber lines (DSL) to increase the available bandwidth for voice over IP.

Cost savings. The first reason for using voice over IP was the cost savings that resulted from allowing home users to bypass the charges associated with long-distance calls. Voice over IP allows service providers to bypass access fees charged for using the final mile to the home since it is presently charged as data. Ultimately, these fees may change and voice over IP may become as regulated as the present public switched telephone network, thereby reducing much of today's cost advantage. Daniel Heer

Bibliography. K. Camp, *IP Telephone Demystified*, McGraw-Hill, 2002; D. Collins, *Carrier Grade Voice over IP*, 2d ed., McGraw-Hill, 2002; D. Minoli and E. Minoli, *Delivering Voice over IP Networks*, 2d ed., Wiley, 2002.

Voice response

The generation of synthetic speech signals in order to convey information to listeners, usually based upon a verbal or textual request by the users. This speech synthesis typically employs a computer program and requires access to storage of portions of speech previously spoken by humans. The naturalness of the synthetic voice depends on several factors, including the vocabulary of words to

pronounce, the amount of stored speech, and the complexity of the synthesis programs. The most basic voice response simply plays back appropriate short verbal responses, which are only copies of human speech signals stored using digital sampling technology. The most universal systems, on the other hand, are capable of transforming any given text into comprehensible speech for a given language. These latter systems so far exist for only 20 or so of the world's major languages, and are flawed in producing speech that, while usually intelligible, sounds unnatural.

Voice response is also known as text-to-speech synthesis (TTS) because the task usually has as input a textual message (to be spoken by the machine). The text could be in tabular form (for example, reading aloud a set of numbers), or, more typically, formatted as normal sentences. Speech synthesizers are much more flexible and universal than their speech-recognition counterparts, for which human talkers must significantly constrain their verbal input to the machines in order to achieve accurate recognition. In TTS, a computer database usually determines the text to be synthetically spoken, following an automatic analysis of each user request. The user may pose the request in response to a menu of inquiries (for example, by an automated telephone dialogue, by pushing a sequence of handset keys, or by a series of brief verbal responses). Thus, the term "voice response" is used to describe the synthetic speech as an output to a user inquiry. The value of such a synthetic voice is the capability of efficiently receiving information from a computer without needing a computer screen or printer. Given the prevalence of telephones, as well as the difficulty of reading small computer screens on many portable computer devices, voice response is a convenient way to get data. *See* SPEECH RECOGNITION.

Capabilities. Voice response systems have advanced significantly in recent decades. Mechanical synthesizers approximating the actual airflow in the human vocal tract have existed, at least in primitive form, for centuries. However, the advent of practical computers in the 1960s spurred the development of much better synthetic speech. Prior to the introduction of special-purpose digital signal processing (DSP) chips in the late 1970s, synthetic speech was limited to large computers. Nowadays, voice response is becoming widespread on inexpensive devices.

While TTS products vary significantly in quality and cost, most produce generally intelligible speech, but all suffer some failings in achieving naturalness. They often sound as if they have a significant (nonhuman, machine) accent. A machine that can talk with proficiency equal to that of native speakers (that is, synthetic speech which listeners cannot distinguish from human speech) is still far from realization. Natural speech output is feasible if the desired vocabulary is very limited, and the system merely concatenates lengthy stored speech units (that is, outputs a sequence of previously spoken words or phrases, stored as coded speech). Thus, complete TTS systems which accept any input text in a chosen language (including new words as well as typographical errors) must be distinguished from more primitive voice-response systems of very limited vocabulary.

Design trade-offs. Major advances in commercial synthesizers are due to improvements in computer technology and synthesis methodology. The design of TTS trades off conflicting demands: maximum speech quality, but minimum memory space, algorithmic complexity, and computational speed. Inferior sound quality is usually due to inadequate modeling of three aspects of human speech production: coarticulation, intonation, and vocal-tract excitation.

Minimum-complexity, large-memory systems. The simplest approach to voice response is to digitally sample natural speech and output the samples later as needed. While compact disks store audio signals using 16-bit sampling at 44,100 samples per second, such high bit rates are rarely used for TTS, due to cost. A more common Nyquist sampling rate is 10,000 samples per second, which still preserves sound frequencies up to almost 5 kHz, allowing quite natural speech. High-frequency energy in fricative sounds is severely attenuated (but less so than on telephone lines), but this usually has little impact on intelligibility. Straightforward sampling [that is, linear pulse-code modulation (PCM)] requires 12 bits per sample, which requires memory at 120 kbits/s. Such high data rates are prohibitive except for applications with very small vocabularies. Even in cases with more limited bandwidth (for example, 8000 samples per second in telephone applications) and more advanced coding schemes [ranging from simple logarithmic coding at 64 kbits/s to code-excited linear prediction (CELP) at 4–10 kbits/s], the straightforward playback approach is unacceptable for general TTS. Despite rapidly decreasing costs for computer memory, it will remain impossible to store all the necessary speech signals except for applications with very restricted vocabulary needs. *See* COMPACT DISK; DATA COMPRESSION; INFORMATION THEORY; PULSE MODULATION.

Generally, voice response requires the conversion of any input text into a speech waveform via algorithms which transform previously coded speech. Such speech synthesis can be characterized by the size and type of speech units concatenated to yield the output, as well as by the method used to code, store, and synthesize. If entire phrases or sentences serve as units, output quality can be very high, but a well-chosen speaker must pronounce each phrase (during a system development stage) with timing and intonation appropriate for all sentences in which it could later be needed. Thus, if a phrase (of one or more words) could occur in several syntactic contexts, its pronunciation should be recorded with appropriate variation. Merely concatenating words which were originally spoken separately often leads to lowered output quality. The duration, spectrum, and pitch of stored units must be adjusted during concatenation since such unit features vary in sentential context (especially for smaller speech units, which need more frequent concatenations).

Minimum-storage, complex systems. A voice response system which minimizes memory needs generates synthetic speech from sequences of brief basic sounds and has great flexibility. (Like the other extreme noted above, such cases have practical applications, but most TTS lies somewhere in between.) Since most languages have only 30–40 phonemes (distinct linguistic sounds), storing units of such size and number is trivial. However, the spectral features of these short concatenated sounds (lasting 50–200 ms) must be adjusted at their frequent boundaries to avoid severely discontinuous speech. Normal pronunciation of each phoneme in an utterance depends heavily on its phonetic context (for example, on neighboring phonemes, intonation, and speaking rate). The adjustment process and the need to calculate an appropriate intonation for each context lead to complicated synthesizers with correspondingly less natural output speech.

Commercial synthesizers are primarily based on concatenation of word or phoneme units, although syllables, demisyllables, and diphones have been tried. (Diphones are phoneme-sized units, divided in the middle of each phoneme, thus preserving in each diphone the transition between adjacent phonemes.) The need to smooth spectral parameters at the boundaries between units decreases as concatenation units increase in size, because of the fewer boundaries in the synthesized speech. Such smoothing is much simpler when the joined units have similar spectra at the boundaries (as is the case with diphones). However, coarticulation (the spectral effects of adjacent sounds on each other) often extends over several phonemes. For example, in "strew," a speaker's lips will round in anticipation of the vowel during the preceding three consonants, /s/, /t/, and /r/, with corresponding acoustic effects which short diphone units cannot account for.

Storage of varied speech units. Current synthesizers usually compromise between the extremes of minimizing storage and complexity. One approach is to store thousands of speech units of varying size, which can be automatically extracted from natural speech. In contrast to automatic speech recognition, where segmentation of speech into pertinent units is very difficult, TTS training exploits prior knowledge of the text (the training speaker reads a furnished text).

The modifications needed to concatenate speech units for synthesis are similar to those of other speech applications such as time expansion and compression, which involve the slowing down or acceleration of stored utterances. Such modifications can be as simple as cut-and-splice, where individual pitch periods are duplicated or deleted periodically. (A pitch period is the time interval between successive vocal-cord or vocal-fold closures, 4–15 ms.) For more natural-sounding time-scale and pitch-scale modifications, more elaborate methods using discrete Fourier transforms or sine-wave models can be employed. *See* FOURIER SERIES AND TRANSFORMS.

Memory size. English has over 300,000 words, although only 50,000 can be considered common (and most people use only about 5000). These words can be generated more efficiently using about 12,000 morphemes, the basic meaningful elements that make up words. For example, the word "antidisestablishmentarianism" consists of the root morpheme "establish" plus two prefix and four suffix morphemes. Approximately 80,000 vectors of spectral parameters (each representing the acoustical output of a vocal tract shape) would be needed for a morpheme memory, whereas a word memory could exceed 1,000,000 entries. If the speech unit is smaller than a syllable, little memory savings are gained by limiting vocabulary; 2000 diphones can generate many thousands of words, but a much smaller vocabulary of 1000 words would still require 1000 diphones.

For large-memory TTS, a serious issue is speaker fatigue, since each speaker whose voice is simulated must utter all speech units as uniformly as possible. Speakers usually find this very arduous for more than a few thousand short phrases. Furthermore, the longer the time span of recording, the more likely is nonuniformity to be present, and the resulting uneven units lead to rough-sounding speech.

Synthesis method. Most voice-response systems are terminal-analog synthesizers, meaning that they model the speech output after filtering by the vocal tract, without explicitly accounting for articulator movements. (Alternative articulatory synthesizers have not been practical so far, due to the difficulty of obtaining accurate three-dimensional vocal-tract representations and of modeling the system with a small set of parameters.) Until recently, waveform methods for synthesis, were used only with small vocabularies (for example, a few minutes of speech). For greater flexibility, parametric methods are necessary and have dominated the field. Such synthesis, however, has often been limited in speech quality owing to inadequate modeling of natural speech production. A standard speech coding method called linear predictive coding (LPC) has often been used to store speech units efficiently and play them back. This suffices for many simple applications (for example, telephone directory assistance). However, if a standard voice and small vocabulary are insufficient, the synthesizer manufacturer must process speech spoken by specific users to establish a custom vocabulary. Such as hoc processing can yield very compact representations, but is expensive.

Unrestricted-text (TTS) systems. Synthesizers that accept general text as input need a linguistic processor to convert the text into phonetic symbols in order to access the appropriate stored speech units. One task is to convert letters into phonemes. This may be as simple as a table look-up: a computer dictionary with an entry for each word in the chosen language, noting its pronunciation (including syllable stress), syntactic category, and possibly some semantic information. Many systems also have language-dependent rules, which examine the context of each letter in a word to determine how it is pronounced; for example, the letter [p] in English is pronounced /p/, except before the letter [h] (for example, in "telephone"; however, it has normal pronunciation in "cupholder"). English needs hundreds of such rules.

TTS often employs these rules as a back-up procedure to handle new words, foreign words, and typographical mistakes (that is, cases not in the dictionary). *See* PHONETICS.

Languages in which spelling follows phonetics more closely (for example, Spanish) can be modeled with very few rules since each letter has normally only one pronunciation. (The same holds for character-based languages such as Chinese, but those alphabets are much larger than 26 letters.) Letter-to-phoneme rules developed manually for many languages are capable of high precision, especially when combined with a dictionary to handle exceptional cases. Errors in phonetic transcription with such advanced systems are almost always due to proper nouns (such as names) or foreign words. Such words may obey rules different from those of the modeled language and are often capitalized (or italicized) in normal text (and can thus be automatically identified for special processing).

Letter-to-phoneme rule sets produced via neural networks have been much less successful than those developed manually. Neural nets have found greater success in speech recognition, where the number of possible utterances and acoustic variation is enormous, compared to the number of words in a language. *See* NEURAL NETWORK.

Some systems employ a word decomposition algorithm, which attempts to strip prefixes and suffixes from each word in an input text. Since there are only a few dozen such affixes in most languages and since they can affect pronunciation (for example, the third vowel in "algebra" versus "algebraic"), such a decomposition procedure can reduce memory requirements (at the expense of extra computation). On the other hand, syntactic (and less often, semantic) information is easily stored in a dictionary. Combined with a natural language parser to determine linguistic structures in the input text, this allows artificial specification of intonation by rule, that is, locating pitch and duration effects to simulate natural intonation. Many synthesizers still forego parsers as too complex or unreliable, and use simplistic intonation rules. Poor handling of intonation is a major reason why much TTS sounds unnatural.

Formant synthesis. Traditionally, TTS has employed a cascade or parallel structure of digital filters, each simulating one resonance (formant) in the vocal tract being simulated. The filters' excitation is a regularly spaced sequence of impulses (for periodic voiced speech) and pseudo-random noise (for unvoiced speech). The cascade structure approximates speech spectra well for vowels and allows simple control with one amplitude parameter. The lowest four formant frequencies and three corresponding bandwidths vary as a function of time. (Variation in higher frequencies has little perceptual effect.) *See* DIGITAL FILTER.

Linear predictive coding (LPC) synthesis. TTS has often avoided the need to manually develope coarticulation rules (for formant synthesis), instead using the automatic method of LPC. The filter for LPC has a simpler structure than that for formant synthesis becuase all spectral properties of speech (except for intensity and periodicity, which present equal challenges to most TTS methods) are included in the LPC spectral parameters (coefficients). A lattice filter is often used because the multiplier coefficients (modeling boundaries between adjacent cylindrical sections of a vocal tract) can be linearly interpolated between successive frames (for smoother speech) without yielding an unstable filter. (A frame is the length of time over which speech analysis is performed, 10–30 ms.) Formant synthesis is more flexible than LPC in allowing simple transformations to simulate different voices. Modifying formant frequencies can alter speaker-related aspects of voice much more easily than with LPC. *See* ELECTRIC FILTER.

Waveform concatenation. A more recent TTS alternative is waveform synthesis, which can yield very high quality speech at the cost of increased memory and decreased flexibility. In the PSOLA (pitch-synchronous overlap-and-add) method, for example, brief speech waveform units are concatenated, effectively fading out one unit while fading in the next, typically with synchronized pitch periods from adjacent units. As with diphone synthesis methods, PSOLA is not readily modified to simulate voices other than that of the training speaker; for many TTS applications, one voice suffices. Early PSOLA methods suffered from a need to manually segment pitch periods in the stored speech units, problems in smoothing at unit boundaries, and a requirement for large memory (for example, 80 kbits/s). Modifications have lessened some drawbacks, while retaining freedom from intensive computation at synthesis time. For example, the MBROLA (multiband resynthesis) method resynthesizes (once) the entire diphone database at a constant average pitch period, which allows simple linear time interpolation at synthesis time and smoother spectral matching.

Synthesis of intonation. The problem of determining an appropriate intonation for each input text continues to confound TTS. In simple voice response, the stored units are large (for example, phrases), and pitch and intensity are usually stored explicitly with the spectral parameters or implicitly in the signals of waveform synthesizers. However, when smaller units are concatenated, the synthetic speech sounds unnatural unless the intonation is adjusted for context. Intonation varies significantly among languages. Although automatic statistical methods show some promise, intonation analysis has mostly been manual.

Three prosodic parameters contribute to intonation: pitch, duration, and intensity. Many intensity variations are phoneme-dependent, and stressed syllables are more intense. Pitch (due to the vibration rate of the vocal cords) is the most problematic of the intonation parameters, owing to its significant variation, both at a frame-by-frame level and globally across the sentence. In English, lexically stressed syllables of emphasized words cause pitch changes to cue the boundaries of syntactic phrases. Pitch usually follows a basic falling pattern, with superimposed excursions above and below a declination line.

Compared to pitch, duration is more tightly linked to phonemics; for example, vowels are longer than consonants, stressed syllables longer than unstressed ones, and consonants shorter in clusters. Vowels are longer when prior to voiced consonants than unvoiced ones. When speaking rate varies, vowels tend to expand or compress more than consonants do.

A major difficulty in specifying natural intonation is the lack of reliable markers in most input text to indicate intonational boundaries. Sentence-final and clause-final punctuation (. ? ! : ;) are reliable places for pauses. However, many sentences feature long sequences of words with, at most, only commas (which do not correspond well to intonation). In most European languages, a break often occurs after a (high-information) "content" word followed by a "function" word. Highlighting the final word in such sequences (with durational lengthening and a pitch rise) is often appropriate.

Different languages. Simple voice-response systems work equally well for all languages since they just play back previously stored speech units. For general TTS, however, major synthesizer components are highly language-dependent. The front end of TTS systems, dealing with letter-to-phoneme rules, the relationship between text and intonation, and different sets of phonemes, is language-dependent. The back end, representing simulation of the vocal track via digital filters, is relatively invariant across languages. Even languages with sounds (for example, clicks) other than the usual pulmonic egressives require only simple modifications.

Practical speech synthesis. Commercial synthesizers are widely available for about 10 languages. They often combine software, memory, and processing chips, and range from expensive systems providing close-to-natural speech to inexpensive personal computer programs. General digital signal processing chips are widely used for TTS. Current microprocessors can easily handle the speeds for synthesis, and indeed synthesizers exist entirely in software. Memory requirements can still be a concern, especially for some of the newer waveform concatenation systems. *See* MICROPROCESSOR.

Prospects. Speech synthesis is increasingly popular, as the cost and size of computer memory decreases. Limited-vocabulary voice response yields high quality and suffices for many applications. Increasingly, inexpensive memory has been exploited by using large inventories of speech units to overcome some coarticulation and intonation problems. These trends follow those of automatic speech recognition methods, whose stochastic methods involve very simple network models, but require massive amounts of training and memory to accommodate the large amount of variability in the way that speakers talk. While TTS does not need to model such speaker variability, it must handle well the large amount of variability (even within one speaker's voice) across the many different phonetic contexts met in normal speech. Eventually, increased understanding of how humans produce and perceive speech will yield more efficient TTS, and some combination of stochastic and knowledge-based methods will yield synthetic speech quite similar to that of humans. *See* SPEECH.

Douglas O'Shaughnessy

Bibliography. J. Allen, Overview of text-to-speech systems, in S. Furui and M. Sondhi (eds.), *Advances in Speech Signal Processing*, pp. 741–790, Marcel Dekker, New York, 1992; T. Dutoit, *An Introduction to Text-to-Speech Synthesis*, Kluwer, 1997; D. Klatt, Review of text-to-speech conversion for English, *J. Acous. Soc. Amer.*, 82:737–793, 1987.

Volatilization

The process of converting a chemical substance from a liquid or solid state to a gaseous or vapor state. Other terms used to describe the same process are vaporization, distillation, and sublimation. A substance can often be separated from another by volatilization and can then be recovered by condensation of the vapor. The substance can be made to volatilize more rapidly either by heating to increase its vapor pressure or by removal of the vapor using a stream of inert gas or a vacuum pump. Heating procedures include the volatilization of water, of mercury, or of arsenic trichloride to separate these substances from interfering elements. Chemical reactions are sometimes utilized to produce volatile products as in the release of carbon dioxide from carbonates, of ammonia in the Kjeldahl method for the determination of nitrogen, and of sulfur dioxide in the determination of sulfur in steel. Volatilization methods are generally characterized by great simplicity and ease of operation, except when high temperatures or highly corrosion-resistant materials are needed. *See* CHEMICAL SEPARATION TECHNIQUES; DISTILLATION; SUBLIMATION; VAPOR PRESSURE.

Louis Gordon; Royce W. Murray

Volcanic glass

A natural glass formed by rapid cooling of magma. Magmas typically comprise crystals and bubbles of gas within a silicate liquid. On slow cooling, the liquid portion of the magma usually crystallizes, but if cooling is sufficiently rapid, it may convert to glass—an amorphous, metastable solid that lacks the long-range microscopic order characteristic of crystalline solids. *See* LAVA.

Formation. Because it is in general easier to form glasses from viscous liquids than from highly fluid liquids, the conditions required for the formation of glasses in nature vary with the chemical composition of the liquids. For example, silica-rich, rhyolitic magmas frequently quench to glass during explosive eruptions and make up the bulk of the solid material in many pyroclastic deposits (usually as shards, pumice lumps, and other fragments); but they also can erupt quiescently to form massive glassy rocks (known as obsidian, the most common source of volcanic glass on land) even in the slowly cooled interiors of flows tens of meters thick. In contrast, more

basic, basaltic glasses (sometimes known as tachylite) are less common and rarely form in more than small quantities unless rapidly cooled in a volcanic eruption. Peles hair is an example of basaltic glass formed in this way. Because magmas cool much more rapidly when erupted under water, basaltic glasses are commonly retrieved from the quenched margins of submarine basalt flows. Small quantities of glass are frequently found in the interstices between crystals in largely crystalline rocks or as small (on the order of a few to a few hundred micrometers in diameter) inclusions trapped during growth of crystals. The low probability of nucleation in such small volumes may be a factor in the vitrification of even inviscid liquids in these inclusions. *See* OBSIDIAN; RHYOLITE.

Significance of composition. The chemical compositions of magmatic liquids are key to understanding the origin of diversity in igneous rocks. Crystalline and partially crystalline rocks are frequently formed from crystal + liquid mixtures into which crystals have been concentrated by, for example, gravitational settling or rising, or from which liquid has been preferentially extracted. It is not always possible to recognize or correct for the effects of such processes on the compositions of crystalline rocks and thereby to reconstruct actual liquid compositions. There has long been considerable interest in volcanic glasses among geologists because, provided that they are not subsequently chemically altered, these glasses can provide unequivocal information on the compositions of liquids involved in igneous processes.

Igneous petrogenesis. Detailed study of unaltered volcanic glasses has led to several important conclusions about igneous petrogenesis that were unavailable or insecure on the basis of crystalline rock studies. Examples of important results from the study of glasses from the submarine basalts of the mid-oceanic ridge environment include (1) determination of the variations in liquid chemical composition within individual provinces and at different parts of the worldwide ridge system, providing the primary constraints on the processes leading to the formation of the Earth's most abundant magma type; (2) measurement of the oxidation state of primitive magmas, and by inference, of the suboceanic mantle; and (3) measurement of the concentrations and isotopic characteristics of volatile components (for example, carbon dioxide, water, sulfur, and chlorine) in relatively undegassed magmas, providing crucial inputs into understanding of the sources and concentrations of these components in the mantle. The study of fresh obsidians and of glass inclusions in crystals in rhyolites has led to similar insights into the evolution of silicic magmas, such as the determination of the preeruptive volatile contents of rhyolitic liquids (for example, typically 4–6 wt % water, dissolved both as molecules of water and hydroxyl groups); the depths of the chambers from which such magmas erupt and the temperatures in such chambers; and the mechanisms by which relatively dry obsidian flows form (for example, by collapse of foams formed during degassing of water-rich rhyolitic magmas as they rise toward the surface). *See* IGNEOUS ROCKS; MID-OCEANIC RIDGE.

Duration. Volcanic glass is highly susceptible to alteration when it is exposed to water, even in low-temperature environments such as on or close to the Earth's surface. Basaltic glasses rapidly alter to a poorly characterized, partially crystalline material known as palagonite. Typically obsidians initially remain largely glassy upon hydration, but their compositions usually change drastically during the process, principally by exchange of oxygen, alkalies, and silica with the aqueous phase. Thicknesses of palagonite on basalt and of hydrated rinds on obsidian can, under favorable circumstances, be used to determine the dates of volcanic eruptions. Perlite, with a characteristic texture of spherical cracks, and pitchstone, defined on the basis of its luster, are the result of extensive hydration of obsidian. Hydrated glass tends to devitric (that is, crystallize) given sufficient time and even minor heating. Consequently, terrestrial glasses are rare in pre-Tertiary rocks. However, glasses that are free of water and have not been reheated have survived for several billions of years in lunar rocks and meteorites. *See* MAGMA; PITCHSTONE.

Edward M. Stolper

Volcano

A mountain or hill, generally steep-sided, formed by accumulation of magma (molten rock with associated gas and crystals) erupted through openings or volcanic vents in the Earth's crust; the term volcano also refers to the vent itself. During the evolution of a long-lived volcano, a permanent shift in the locus of principal vent activity can produce a satellitic volcanic accumulation as large as or larger than the parent volcano, in effect forming a new volcano on the flanks of the old.

Planetary exploration has revealed dramatic evidence of volcanoes and their products on the Earth's Moon, Mars, Mercury, Venus, and the moons of Jupiter (**Fig. 1**), Neptune, and Uranus on a scale much more vast than on Earth. For example, Olympus Mons, a gigantic shield volcano on Mars about 600 km (375 mi) in diameter, is larger across than the length of the Hawaiian Islands. However, only the products and landforms of terrestrial volcanic activity are described here. *See* VOLCANOLOGY; MARS; MERCURY (PLANET); MOON; NEPTUNE; URANUS; VENUS.

Volcanic vents. Volcanic vents, channelways for magma to ascend toward the surface, can be grouped into two general types: fissure and central (pipelike). Magma consolidating below the surface in fissures or pipes forms a variety of igneous bodies (**Fig. 2**), but magma breaking the surface produces fissure or pipe eruptions (**Figs. 3** and **4**). Fissures, most of them less than 10 ft (3 m) wide, may form in the summit region of a volcano, on its flanks, or near its base; central vents tend to be restricted to the summit area of a volcano. For some volcanoes or volcanic regions,

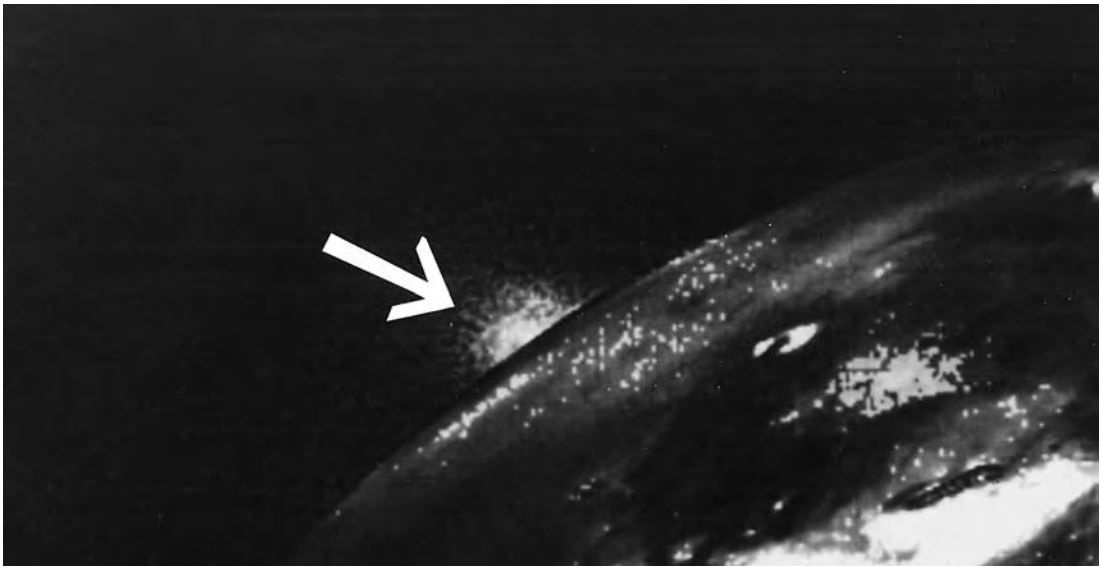

Fig. 1. Picture taken by *Voyager 2* in July 1979 showing volcanic plumes (arrow) rising approximately 60–110 mi (95–175 km) above the surface of Io, a moon of Jupiter. [In 1997 and 1999 the *Galileo* spacecraft also captured images of Io's eruptive activity.] (*NASA*)

swarms of fissure vents are clustered in swaths called rift zones.

Volcanic products. Magma erupted onto the Earth's surface is called lava. If the lava is chilled and solidifies quickly, it forms volcanic glass; slower rates of chilling result in greater crystallization before complete solidification. Lava may accrete near the vent to form various minor structures or may pour out in streams called lava flows, which may travel many tens of miles from the vents. During more violent eruption, lava torn into fragments and hurled into the air is called pyroclastic (fire-broken materials). The general term "tephra" is applied to poorly consolidated pyroclastic debris regardless of fragment or particle size. *See* CRYSTALLIZATION; LAVA; MAGMA; VOLCANIC GLASS.

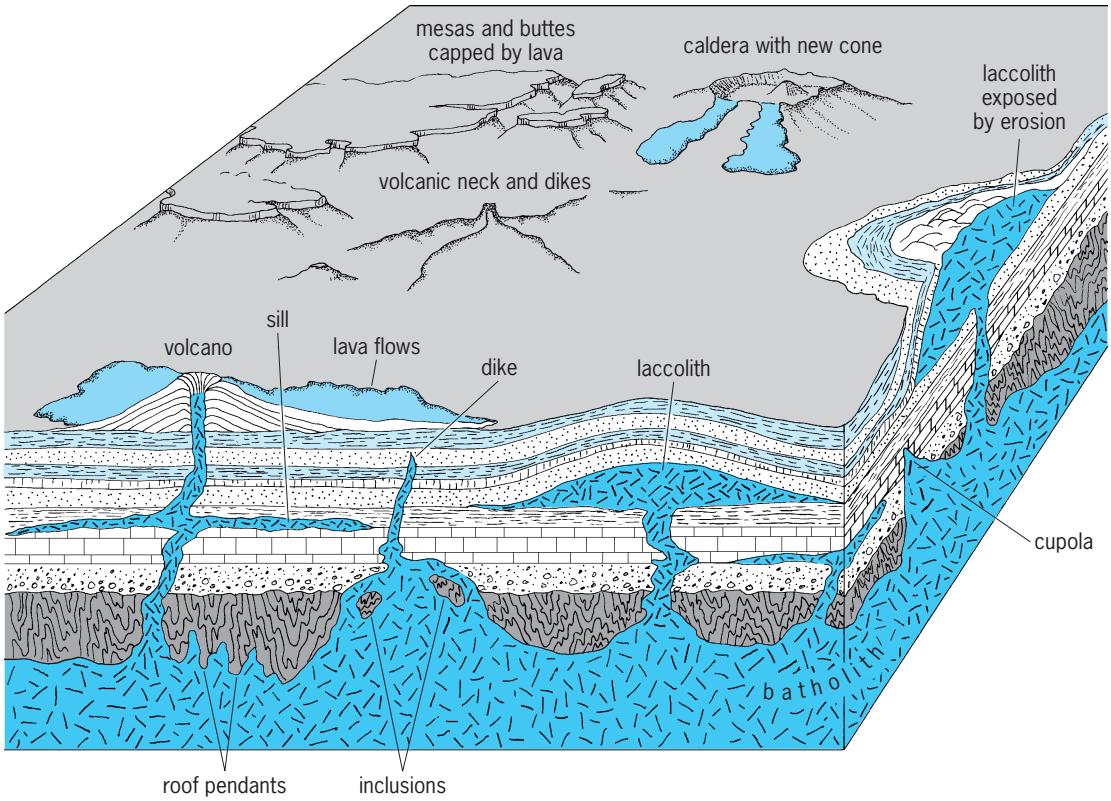

Fig. 2. Cross-sectional diagram illustrating igneous bodies in relation to geological structures and surface forms. (*After W. H. Emmons et al., eds., Geology: Principles and Processes, 5th ed., McGraw-Hill, 1960*)

The character of a volcanic eruption depends largely on the viscosity of the liquid lava; viscosity, for the purposes of this discussion, can be considered as a stickiness coefficient of the magma or lava. In general, mafic (basaltic and similar) lavas are less viscous (more fluid) than more silicic (dacitic and similar) lavas. Eruptions of mafic lavas commonly are nonexplosive or only weakly explosive and produce mostly lava flows; eruptions of silicic lavas are typically explosive and yield predominantly pyroclastic material. Some produce only pyroclastic material.

Lava flows. Lava flows are classified in terms of their surface characteristics. Flows showing smooth or hummocky, gently undulating surfaces and crusts locally wrinkled into ropelike forms are called pahoehoe. Flows showing very rough irregular surfaces covered by jagged spinose fragments resembling furnace clinker are called aa (**Fig. 5**). The terms pahoehoe and aa are of Hawaiian origin. Flows in which the fragments that constitute the upper part of the flow are fairly smooth-sided polygons are called block lava. Fluid basaltic and related mafic lavas characteristically form pahoehoe or aa flows, or flows intermediate in character between these two end members. In contrast, more viscous lavas such as andesites more commonly form block lava flows. All lava flows that have poured out on land contain various amounts of open cavities (vesicles), which mark the sites of gas bubbles formed as the rising magma reached regions of increasingly lower pressure before breaching the surface to erupt. Over geologic time, vesicles can be filled by minerals precipitated from circulating mineralizing fluids or ground water. Where pahoehoe flows enter bodies of water or wet ground, they may form heaps of irregular ellipsoids, in cross section somewhat resembling sacks of grain or pillows. Basaltic pillow lavas form an overwhelming bulk of the ocean floor. *See* ANDESITE.

Pyroclastic materials. Magma, at depth and under great pressure, contains gas in solution, but as it rises into regions of lower pressure near the surface of the Earth, the gas starts to exsolve and escape from the liquid. The gas generally escapes readily from fluid lavas, with little or no explosion; in more viscous liquids, however, the gas may acquire considerable pressure before it escapes and bursts forth in strong explosions. Thus, the emergence at the surface of highly gas-charged, viscous silicic lava may be attended by a sudden frothing as the contained gas rapidly exsolves and vesiculates. The sudden expansion of the gas may tear the froth into countless tiny shreds, each of which chills virtually instantaneously to form fragments of volcanic glass. Continued ebullition of gas results in a mass of small solid or quasisolid fragments, each surrounded by an envelope of still-expanding gas that pushes against all adjacent expanding envelopes. The net effect of this process is the isolation of the solid fragments from contact with one another, and the entire mass obtains an expansive quality that is enhanced further by the expansion of heated air trapped in the hot, moving mass. The result is a very mobile suspension of incandescent solid fragments in gas which may flow at great speed down slopes and spread out to great distances from the erupting vents, forming extensive deposits having nearly flat surfaces. When these ash flows come to rest, they commonly are still so hot that the fragments of glass stick together or even merge in the center of the deposit to form a layer of solid black obsidian. The resulting deposits are known as welded tuffs or ignimbrites; the degree of welding of ash flows is largely dependent on the

Fig. 3. Aerial view of 165-ft-high (50-m) lava fountains spurting from three en-echelon fissure vents on the northeast rift zone of Mauna Loa volcano, Hawaii, feeding a massive lava flow during its July 1975 eruption. Mauna Loa erupted again during March–April 1984, producing two 16-mi-long (25-km) flows that reached within 3–6 mi (5–10 km) of Hilo, the largest city on the island of Hawaii. (*R. T. Holcomb, USGS*)

Fig. 4. Parícutin volcano, Mexico, February 20, 1944, showing typical eruptive activity from a pipelike vent to form a cinder cone. (*Courtesy of T. Nichols*)

Fig. 5. Active pahoehoe lava flow (left) lapping against and overriding previously erupted aa lava (right) during the 1972–1974 eruption of Kilauea volcano, Hawaii. (*R. I. Tilling, USGS*)

temperature and thickness of the deposit. Rapidly moving incandescent ash flows have been called glowing avalanches or nuée ardentes; they can be highly destructive to life and property. For example, in 1902 a devastating nuée ardente produced during a violent eruption of Mont Pelée (island of Martinique in the Lesser Antilles) virtually destroyed the entire city of St. Pierre, killing some 30,000 persons. More recently, nuées ardentes generated during the March-April 1982 eruption of El Chichón volcano (State of Chiapas, southeastern Mexico) wiped out all settlements within 5 mi (8 km) of the volcano, resulting in the loss of more than 2000 people (**Fig. 6**). *See* OBSIDIAN; TUFF.

Pyroclastic materials are classified by the nature of the shreds and fragments: essential—the erupting molten lava; accessory—solidified lava of previous eruptions of the same volcano; and accidental—solid material of still older, not necessarily volcanic, rocks from the crust beneath the volcano. The classification of pyroclastic materials also considers size, shape, and consistency of constituent fragments and matrix. Blobs or drops of material still liquid enough to assume rounded or aerodynamically drawn-out forms during flight are known as lapilli (if 0.16–1.2 in. or 4–32 mm in average diameter) and bombs (if greater than 1.2 in. or 32 mm). Depending on their final shapes when they strike the ground, bombs are variously called cow-dung bombs, spindle or fusiform bombs, or ribbon bombs.

Fig. 6. Ruins of church (circled), the only visible remains of the village of Francisco León after being swept by nuées ardentes from the 1982 eruption of El Chichón volcano, southeastern Mexico. (*R. I. Tilling, USGS*)

Irregular fragments of frothy lava of bomb or lapilli size are called scoria or cinder; if the fragments are sufficiently plastic to flatten or splash as they hit, they are called spatter. The still-molten fragments of spatter often adhere to each other to form welded spatter, or agglutinate. Angular fragments larger than 1.2 in. (32 mm) either solid or too viscous to assume rounded forms during flight are known as blocks; their accumulation forms a volcanic breccia. Ejecta smaller than 0.16 in. (4 mm) are called ash, and those smaller than 0.01 in. (0.25 mm) are called dust. Indurated (hardened) volcanic ash or dust is called tuff.

The term "tephra" had been applied liberally to all pyroclastic materials, but this term should be reserved, as originally defined, for those pyroclastic deposits of air-fall origin regardless of the size of the ejected material. The individual pyroclastic fragments—the building blocks of pyroclastic rocks—are called pyroclasts. *See* PYROCLASTIC ROCKS.

Volcanic gases. In general, water vapor is the most abundant constituent in volcanic gases; the water is mostly of meteoric (atmospheric) origin, but in some volcanoes can have a significant magmatic or juvenile component. Excluding water vapor, the most abundant gases are the various species of carbon, sulfur, hydrogen, chlorine, and fluorine (such as CO_2, CO, SO_2, SO_3, H_2, H_2S, Cl_2, F_2, HCl). As discussed above, the related processes of rapid exsolution, expansion, and release of gas from the magma as it ascends near the Earth's crust provide the energy for driving and sustaining volcanic eruptions. Volcanic gases, as well as the erupted solid volcanic materials, can pose a hazard to humans. In August 1986, at Lake Nyos, a volcanic lake occupying the crater of a geologically young volcano in Cameroon (western Africa), the sudden overturn of deep lake waters caused the massive release of carbon dioxide (CO_2) gas of magmatic origin, asphyxiating more than 1700 people in low-lying areas.

Volcanic aerosols. Violent volcanic explosions may throw dust and aerosols high into the stratosphere, where it may drift across the surface of the globe for many thousands of miles. Fine ash and dust from eruptions of Icelandic volcanoes have fallen in the streets of Moscow. Studies have shown that the particles in the eruption cloud are mostly angular bits of lava, many of them glassy; but some glass spheroids also are present, as well as liquid droplets of hydrous solutions of sulfuric acid and various sulfates and chlorides. Most of the solid particles in the volcanic cloud settle out within a few days, and nearly

all settle out within a few weeks, but the gaseous aerosols (principally sulfuric acid droplets) may remain suspended in the stratosphere for several years. Such stratospheric clouds of volcanic aerosols, if sufficiently voluminous and long-lived, can have an impact on global climate. *See* ACID RAIN; AEROSOL; AIR POLLUTION.

Volcanic mudflows. Mudflows are common on steep-side volcanoes where poorly indurated or nonwelded pyroclastic material is abundant. They may form by eruptions involving the water of a crater lake; after breaching its confining walls, the water sweeps down the mountainside, incorporating and mixing with loose volcanic debris to form a slurry. Such slurries can be quite dense and commonly have a consistency similar to wet concrete. Mudflows may also form by hot or cold volcanic debris avalanches descending into streams or onto snow or ice. Probably by far the most common cause, however, is simply heavy rain saturating a thick cover of loose unstable pyroclastic material on the steep slope of the volcano, transforming the material into a mobile, water-saturated "mud," which can rush downslope at a speed as great as 50–55 mi (80–90 km) per hour. Such a dense, fast-moving mass can be highly destructive, sweeping up everything loose in its path. Volcanic mudflows can be hot or cold; they are sometimes called lahars, a term from Indonesia, where they have taken a heavy toll in property and human lives (**Fig. 7**). Fast-moving mudflows triggered by a small eruption on November 13, 1985, at Nevado del Ruiz, Colombia, killed about 25,000 people. The Ruiz catastrophe was the worst volcanic disaster in the twentieth century since the devastation associated with the eruption of Mont Pelée, Martinique in 1902.

Somewhat related to mudflows are the great floods of water, known in Iceland as Jökulhlaup, that result from rapid melting of ice by volcanic eruption beneath a glacier. A subglacial eruption in October 1996 at Grímsvötn Volcano beneath the Vatnajökull ice sheet, Iceland, produced abundant hyaloclastite. In November 1996 the catastrophic release of impounded melt water from the eruption produced a destructive glacial outburst flood that affected more than 270 mi^2 (750 km^2) and destroyed or severely damaged several bridges. *See* GLACIOLOGY.

Volcanic landforms. Much of the Earth's solid surface, on land and below the sea, has been shaped by volcanic activity. Landscape features of volcanic origin may be either positive (constructional) forms, the result of accumulation of volcanic materials, or negative forms, the result of the lack of accumulation or collapse.

Major edifices. The gross form of a volcano is largely determined by the viscosity and mode of eruption of the volcanic products, even though other factors may be operative locally on a smaller scale.

Not all volcanoes show a graceful, symmetrical cone shape, such as that exemplified by Mount Fuji, Japan, or Mayon volcano, Philippines (**Fig. 8**). In reality, most volcanoes, especially those near tectonic plate boundaries, are more irregular, though of grossly conical shape. Such volcanoes, called

Fig. 7. Houses in the Cibanjaran River buried by lahars (volcanic mudflows) generated during the 1982–1983 eruption of Galunggung volcano, Province of West Java, Indonesia. (*R. T. Holcomb, USGS*)

Fig. 8. Mayon, a symmetrical composite volcano near Legaspi, Luzon Island, Philippines. In September-October 1984, its eruption caused the evacuation of tens of thousands of residents from 36 villages. (*G. A. MacDonald*, *USGS*)

stratovolcanoes or composite volcanoes, typically erupt explosively and are composed dominantly of andesitic, relatively viscous and short lava flows, interlayered with beds of ash and cinder that thin away from the principal vents. Volcanoes constructed primarily of fluid basaltic lava flows, which may spread great distances from the vents, typically are gentle-sloped, broadly upward convex structures that resemble a Germanic warrior's shield in form. Such shield volcanoes, classic examples of which are Mauna Loa and Mauna Kea volcanoes, Hawaii (**Fig. 9**), tend to form in oceanic intraplate regions and are associated with hot-spot volcanism. The shape and size of a volcano can vary widely between the simple forms of composite and shield volcanoes, depending on magma viscosity, eruptive style (explosive versus nonexplosive), migration of vent locations, duration and complexity of eruptive history, and posteruption modifications (**Fig. 10***a*). Large shield volcanoes are much larger than the largest composite volcanoes (Fig. 10*b*).

Fig. 9. Mauna Kea volcano, 13,800 ft (4208 m) high, viewed from sea level at Hilo Bay, Hawaii. Mauna Kea shows the classic, gently sloping, broadly convex form of a shield volcano, whose surface has been peppered with numerous smaller, steeper-sided cinder cones. (*R. I. Tilling*, *USGS*)

Some of the largest volcanic edifices are not shaped like the composite or shield volcanoes. In certain regions of the world, voluminous extrusions of very fluid basaltic lava from dispersed fissure swarms have built broad, nearly flat-topped accumulations, some covering hundreds of thousands of square miles with volumes of several tens of thousands of cubic miles. These voluminous outpourings of lava are known as flood basalts or plateau basalts. Two examples in the United States are the Columbia River Plateau (**Fig. 11**) and the Snake River Plain volcanic fields; similar large features exist in India (Deccan Plateau) and South America (Paraná Basin). *See* BASALT.

The only volcanic products other than basalt that are sufficiently fluid and voluminous to form extensive volcanic plateaus or plains are ash flows. Some of the larger examples of these ash-flow or ignimbrite plains, such as the central part of North Island of New Zealand, the Jemez Mountains region of New Mexico, and the Yellowstone Plateau of Montana-Wyoming, involve repeated eruptions of ash flows covering many thousands of square miles.

Minor structures. In general, minor volcanic structures originate or develop on or near the major edifice during a single eruptive event or a short span of activity, whereas the major edifices represent the result of repeated activity spanning many thousands or even millions of years.

Lava too viscous to flow readily may accrete in or near the vent to form a steep-sided heap known as a lava or volcanic dome; such domes commonly develop following vigorous explosive activity at composite volcanoes (**Figs. 12** and **13**). Slender spires that thrust through apertures in such a dome are termed spines. The famous spine of Mont Pelée (Martinique, Lesser Antilles), formed during the eruption of 1902, reached a height of over 1280 ft (390 m), but like most such spines, was very short-lived. Other near-vent forms that may develop on the volcano include volcanic shield, cinder cone, spatter cone (pipe vent), spatter rampart (fissure vent), ash cone, tuff cone, and tuff ring (**Fig. 14**).

Major edifices are commonly modified by negative or depressional landforms. Small bowl-shaped depressions that are formed by explosion, or by failure of pyroclastic ejecta to accumulate directly above a vent, are known as craters. Most of them are found at the summit or on the flanks of volcanic cones, but some are well away from any cones.

Larger depressions at the summit of volcanoes are formed by collapse of the summit region as the support beneath it is removed by the rapid withdrawal of magma, usually by surface eruption but sometimes by subsurface migration of magma within the volcano. A depression formed by collapse is called a pit or collapse crater, or if larger than about 1.2 mi (2 km) in diameter, a caldera. Perhaps the best-known examples of calderas in the United States are Crater Lake, Oregon; Kilauea caldera, Hawaii; Yellowstone National Park (Wyoming); and Valles caldera, Jemez Mountains, New Mexico. Larger, though less regular and obvious, features of similar origin are known as

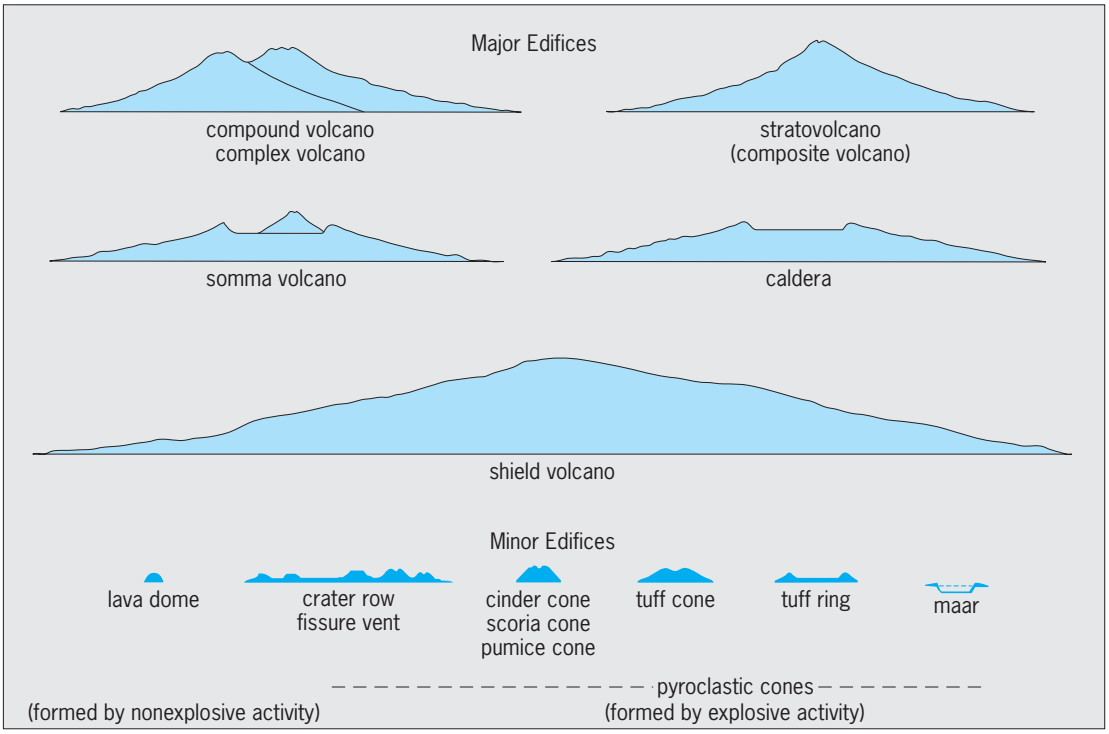

(a)

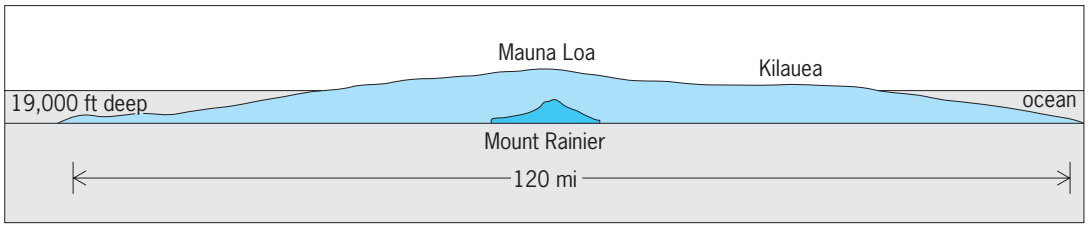

(b)

Fig. 10. Profiles of volcanoes. (*a*) Some common types of volcanoes; the relative sizes shown are only approximate and dimensions vary greatly within each group. For the major edifices, the vertical exaggeration is about 2:1; for the minor edifices, about 4:1 (*after T. Simkin and L. Siebert, Volcanoes of the World, 2d ed., Geoscience Press, 1994*). (*b*) Hawaiian shield volcanoes (Mauna Loa and Kilauea) compared with Mount Rainier (Washington), one of the larger composite volcanoes of the Cascade Range, drawn at the same scale (no vertical exaggeration) (*from R. I. Tilling et al., Eruptions of Hawaiian Volcanoes: Past, Present, and Future, USGS, 1987*).

volcanic-tectonic depressions. Like many calderas, their formation commonly, if not always, is associated with the explosive eruption of great volumes of ash flows. *See* IGNEOUS ROCKS; PETROLOGY.

Evolving landforms. Excluding postformation, gradual changes caused by weathering and erosion, volcanic landforms commonly undergo rapid changes during the course of an eruption, depending on its duration, type of eruptive activity, and so on. Such short-term, abrupt changes are dramatically illustrated by the changes in the summit configuration of Mount St. Helens (Washington) since 1980. Figure 13*a* is a photograph taken on May 17, 1980, from the north one day before the catastrophic eruption, showing the symmetrical form of a typical composite volcano (compare with Fig. 8). Figure 13*b* shows the same view, photographed after the catastrophic eruption of May 18, which removed the upper 1300 ft (400 m) of the summit and formed a 1.1 × 2.1 mi (1.7 × 3.3 km) amphitheater-shaped crater. Continued growth of the volcanic dome inside the crater (Fig. 12) could ultimately restore Mount St. Helens to its former shape. An intermediate stage in such a future restoration might be represented by Bezymianny volcano (Kamchatka, Russia), photographed in 1982 (Fig. 13*c*). The large crater, which formed in 1956 by a similar explosive eruption, is largely filled by lava of a growing volcanic dome.

Submarine volcanism. More than 80% of the Earth's crust is of volcanic origin, and about three-quarters of all the Earth's active volcanism takes place not on land but deep beneath the oceans. Such deep submarine volcanism occurs along the spreading ridges that zigzag for thousands of miles across the ocean floor, and it is exposed above sea level only in Iceland (**Fig. 15**). Because of the logistical difficulties in making direct observations posed by the great ocean depths, no deep submarine volcanic activity has

Fig. 11. A high cliff near Little Goose Dam, Washington, exposing a series of thick plateau basalts making up part of the Columbia River Plateau in the northwestern region of the United States. A car (circled) on the highway at the base of the cliff gives scale. (*D. A. Swanson, USGS*)

been actually observed during eruption by scientists. However, evidence that deep-sea eruptions are happening is clearly indicated by (1) eruption-induced earthquake activity recorded by seismic and acoustic monitoring networks; (2) the presence of high-temperature deep-ocean floor hydrothermal vents called smokers; (3) episodic short-lived but widespread hydrothermal discharges, measured and mapped as thermal and geochemical anomalies in the ocean water above the spreading ridges; and (4) the detection of new lava flows in certain segments of the oceanic ridge system (for example, Juan de Fuca and Gorda ridges off the coast of Oregon and Washington) where there had been none observed previously, as shown by repeated mapping of changes in the bathymetry and imaging of surface features. *See* HYDROTHERMAL VENT; MID-OCEANIC RIDGE.

Fig. 12. Ash and gas plume rises from the volcanic dome growing inside the large, amphitheaterlike crater formed by the May 1980 eruption of Mount St. Helens. In this picture (taken May 1983), the dome measured about 2800 ft (850 m) long, 2600 ft (800 m) wide, and 750 ft (230 m) high; the dome continued to grow until October 1986. (*L. Topinka, USGS*)

In contrast, shallow submarine volcanism has been directly observed and, in some cases, well studied. Volcanic eruptions in shallow water are very similar in character to those on land but, on average, are probably somewhat more explosive, owing to heating of water and resultant violent generation of supercritical steam. The glassy ash of cones formed in this way commonly is altered, probably by ordinary weathering processes, to brownish palagonite tuff. Such cones, like Diamond Head in Honolulu, usually have broader, flatter profiles than those characteristic of cinder cones.

Much of the ocean basin appears to be floored by basaltic lava, which, judging from its apparent density, is much less vesicular than the lavas of the basaltic cones that rise above it to form most of the oceanic islands. The pressure of the overlying water at great depths in the oceans may prevent the exsolution and explosive escape of gas from erupting lava and greatly reduce the vesiculation of the lava itself. Recent studies of submarine volcanism via deep submersible research vessels and dredged samples show that the volcanic landforms and flow features of basaltic lavas observed on land may be present on the ocean floor in addition to the ubiquitous pillow lavas. *See* OCEANIC ISLANDS.

Although pyroclastic materials in the ordinary sense probably seldom, if ever, form in deep water, lava flows may shatter on contact with the water, forming masses of glassy sand-size fragments resembling ash. Such material, called hyaloclastite, may form in water of any depth and is commonly associated with pillow lavas. Great volumes of hyaloclastite

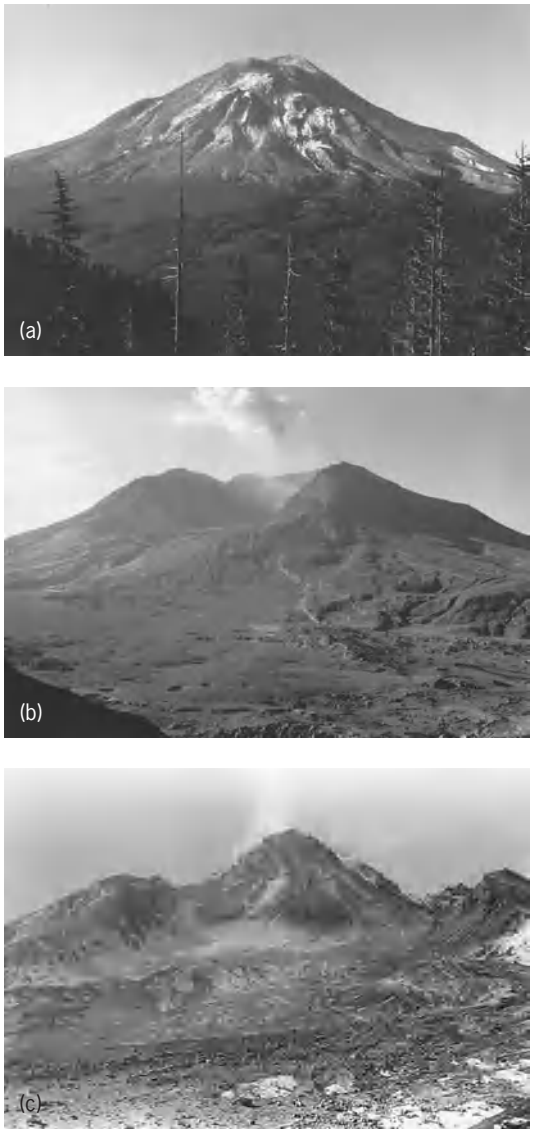

Fig. 13. Changes in the summit of Mount St. Helens. (*a*, *b*) Actual changes since 1980 (*photographs by Harry Glicken*, *USGS*). (*c*) A growing volcanic dome rises above the crater rim of Bezymianny volcano (Russia, 1982); this view may represent a possible future configuration of Mount St. Helens (*Institute of Volcanology*, *Petropavlovsk*, *Kamchatka*).

were formed in melt water by eruptions beneath glaciers in Iceland. Like ordinary basaltic ash, hyaloclastite readily alters to palagonite.

Fumaroles and hot springs. Vents at which volcanic gases issue without lava or after the eruption are known as fumaroles. They are found on active volcanoes during and between eruptions and on dormant volcanoes, persisting long after the volcano itself has become inactive. Fumarolic gases include water vapor, sulfur gases, hydrochloric and hydrofluoric acids, carbon dioxide and monoxide, and others in less abundance. They transport and deposit at the surface small amounts of many common metals. Temperatures of the escaping gases may reach 930–1470°F (500°–800°C), and the halogen gases and metals generally are found in the high-temperature fumaroles. Lower-temperature fumaroles, in which sulfur gases predominate along with steam, are called solfataras; still cooler ones liberating predominantly carbon gases are called mofettes. Researchers in submarine studies of the East Pacific Rise (20°N, off the mouth of the Gulf of California) and of the Gorda Ridge (~43°N, off the coast of Oregon) have photographed undersea fumaroles and sampled metallic sulfides they were emitting.

Fumaroles grade into hot springs and geysers. The water of most, if not all, hot springs is predominantly of meteoric origin, and is not water liberated from magma. Some hot springs appear to result simply from water circulating to warm regions at great depths in the Earth's crust, but in many the heat is of volcanic origin and the water may contain volcanic gases. Indeed, the heat may be derived wholly from rising hot volcanic gases. Some naturally heated water and steam are being developed commercially as geothermal energy to provide electricity or beneficial heat for space heating. *See* GEOTHERMAL POWER; GEYSER.

Distribution of volcanoes. Over 500 active volcanoes are known on the Earth, mostly along or near the boundaries of the dozen or so lithospheric plates that compose the Earth's solid surface (Fig. 15). These rigid plates, which range in thickness from 30 to 90 mi (50 to 150 km) and consist of both crustal and upper mantle material, form the lithosphere and move relative to one another above a hotter, more plastic zone in the mantle called the asthenosphere. *See* ASTHENOSPHERE; LITHOSPHERE; PLATE TECTONICS.

Lithospheric plates show three distinct types of boundaries: divergent or spreading margins—adjacent plates are pulling apart; convergent margins (subduction zones)—plates are moving toward each other and one is being destroyed; and transform margins—one plate is sliding horizontally past another. All these types of plate motion are well demonstrated in the Circum-Pacific region, in which many active volcanoes form the so-called Ring of Fire (Fig. 15).

Along spreading boundaries, new basaltic magma formed by partial melting of mantle material moves into the tensional fissures, solidifying as dikes or feeding lava flows on the ocean floor, accreting new material to the lithosphere. *See* RIFT VALLEY.

The Earth has remained more or less constant in size for the past few hundreds of millions of years; thus the accretion of the lithosphere by volcanism at the spreading boundaries must be compensated for by the destruction of lithosphere elsewhere. The compensatory consumption of lithosphere is assumed to take place along the convergent plate boundaries, where one plate plunges beneath (or is subducted by) another plate. Where the plate moves downward in a subduction zone, the lithosphere is depressed, forming deep troughs such as the Japan and Mariana trenches. As the sinking lithospheric slab enters the hot underlying mantle, it is partially melted, yielding magma that rises through the edge of the overlying plate to produce volcanic activity at

Fig. 14. Aerial view of a string of cinder cones ("crater row" example in Fig. 10*a*) along the fissure vent of the Lakagígar (Laki), Iceland, eruption of 1783, which resulted in the most voluminous outpouring of lava in recorded history. (*R. W. Decker, USGS*)

Fig. 15. Generalized diagram showing the major lithospheric plates and some of the active volcanoes of the Earth. Some scientists subdivide the Indo-Australian plate into the Indian and Australian plates.

the surface. The explosive volcanoes of the Circum-Pacific Ring of Fire, with their predominantly andesitic viscous magmas, have been generated by subduction processes. *See* PLATE TECTONICS.

Some volcanoes, however, are not associated with plate boundaries, and many of these so-called intraplate volcanoes form roughly linear chains in the interior parts of the oceanic plates, for example, the Hawaiian-Emperor, Austral, Society, and Line archipelagoes in the Pacific Basin. Intraplate volcanism also has resulted in voluminous outpourings of fluid lava to form extensive plateau basalts, or of more viscous and siliceous pyroclastic products to form ash flow plains.

Geologic data show that the Hawaiian-Emperor chain becomes progressively younger from northwest to southeast. This relationship, in conjunction with the parallelism in trend of the linear volcanic chain and plate motion, provides the best evidence for the so-called hot-spot or melting-spot model to account for the origin and development of the Hawaiian-Emperor and other linear volcanic chains. According to this model, as the northwest-moving Pacific plate passes over a hot, magma-generating spot in the mantle, the magma formed by partial melting rises through the oceanic lithosphere to form a volcano. This volcano is then gradually and steadily carried northwestward away from the hot spot, until it becomes isolated from the magma-generating source and dies, and a new volcano forms behind it over the hot spot. This process continues to form the entire linear volcanic chain. Although most scientists accept the geometric and dynamic aspects of the hot-spot model for the origin of linear intraplate volcanic chains, fundamental questions regarding the origin, persistence, location, or possible migration of hot spots remain. *See* EARTH; HOT SPOTS (GEOLOGY); MARINE GEOLOGY; PACIFIC ISLANDS. Robert I. Tilling

Bibliography. M. H. Carr, *Volcanic Processes in the Solar System*, 1987; R. W. Decker and B. Decker, *Volcanoes*, 3d ed., 1998; R. W. Decker, T. L. Wright, and P. H. Stauffer (eds.), *Volcanism in Hawaii*, 1987; P. Francis, *Volcanoes: A Planetary Perspective*, 1993; B. L. Foxworthy and M. Hill, *Volcanic Eruptions of 1980 at Mount St. Helens: The First 100 Days*, 1982; S. L. Harris, *Fire Mountains of the West: The Cascades and Mono Lake Valcanoes*, 1988; W. J. Kious and R. I. Tilling, *This Dynamic Planet: The Story of Plate Tectonics*, 1996; P. W. Lipman and D. R. Mullineaux (eds.), *The 1980 Eruptions of Mount St. Helens, Washington*, 1981; G. A. Macdonald, *Volcanoes*, 1972; G. A. Macdonald, A. T. Abbott, and F. L. Peterson, *Volcanoes in the Sea: The Geology of Hawaii*, 2d ed., 1983; L. McClelland et al., *Global Volcanism, 1975–1985*, 1989; H. Sigurdsson et al. (eds.), *Encyclopedia of Volcanoes*, 2000; T. Simkin and L. Siebert, *Volcanoes of the World: A Regional Directory, Gazetteer, and Chronology of Volcanism During the Last 10,000 Years*, rev. ed., 1994; T. Simkin and R. S. Fiske, *Krakatau 1983: The Volcanic Eruption and Its Effects*, 1983; R. I. Tilling, *Eruptions of Mount St. Helens: Past, Present, and Future*, rev. ed., 1990; R. I. Tilling, *Volcanoes*, 1982; R. I. Tilling, C. C. Heliker, and T. L. Wright, *Eruptions of Hawaiian Volcanoes: Past, Present, and Future*, 1987; H. Williams and A. R. McBirney, *Volcanology*, 1979.

Volcanology

The scientific study of volcanic phenomena, especially the processes, products, and hazards associated with active or potentially active volcanoes. It focuses on eruptive activity that has occurred within the past 10,000 years of the Earth's history, particularly eruptions during recorded history. Strictly speaking, it emphasizes the surface eruption of magmas and related gases, and the structures, deposits, and other effects produced thereby. Broadly speaking, however, volcanology includes all studies germane to the generation, storage, and transport of magma, because the surface eruption of magma represents the culmination of diverse physicochemical processes at depth. This article considers the activity of erupting volcanoes and the nature of erupting lavas. For a discussion of the distribution of volcanoes and the surface structures and deposits produced by them *See* PLATE TECTONICS; VOLCANO.

Volcanoes and humans. From the dawn of civilization, volcanic eruptions have intruded into human affairs, producing death and destruction, bewilderment, fear, superstition, and, ultimately, scientific curiosity. Deities or supernatural events, directly or indirectly linked to volcanoes and eruptions, figure prominently in the legends and myths of civilizations that developed in or near regions of active volcanism. During the last 500 years, at least 200,000 people have lost their lives as a result of volcanic eruptions (**Table 1**).

Three eruptions in the 1980s appreciably increased public awareness of volcanic activity and of volcanology. The eruption of Mount St. Helens (Washington) on May 18, 1980 (**Fig. 1**), caused the worst volcanic disaster in the history of the United States, resulting in the loss of 57 lives. Yet, this eruption was much less destructive than other historic eruptions (Table 1). In March-April 1982, El Chichón, an obscure and largely forgotten volcano in southeastern Mexico, produced three major explosive bursts, which obliterated all settlements within a radius of about 5 mi (8 km) and perhaps caused more than 2000 deaths. Not only was this eruption the most destructive in Mexico's history, but some atmospheric scientists have claimed that the massive injection of sulfate aerosols into the stratosphere by El Chichón may have affected global climate, perhaps lowering the average temperature in the Northern Hemisphere by 0.18 or 0.36°F (0.1 or 0.2°C).

On November 13, 1985, a very small-volume eruption (0.007 mi^3 or 0.029 km^3) occurred at the summit crater of 17,680-ft-high (5389-m), glacier-capped Volcán Nevado del Ruiz, Colombia, the northernmost active volcano in the Andes. Despite the small amount of material erupted, the hot ejecta mixed

TABLE 1. Some historical volcanic eruptions

Volcano	Year	Estimated casualties	Principal causes of death
Merapi (Indonesia)	1006	>1,000	Explosions
Kelut (Indonesia)	1586	10,000	Lahars (mudflows)
Vesuvius (Italy)	1631	18,000	Lava flows, mudflows
Etna (Italy)	1669	10,000	Lava flows, explosions
Merapi (Indonesia)	1672	>300	Nuées ardentes, lahars
Awu (Indonesia)	1711	3,200	Lahars
Papandayan (Indonesia)	1772	2,957	Explosions
Laki (Iceland)	1783	10,000	Lava flows, volcanic gas, starvation*
Asama (Japan)	1783	1,151	Lava flows, lahars
Unzen (Japan)	1792	15,000	Lahars, tsunami†
Mayon (Philippines)	1814	1,200	Nuées ardentes, lava flows
Tambora (Indonesia)	1815	92,000	Starvation*
Galunggung (Indonesia)	1822	4,000	Lahars
Awu (Indonesia)	1856	2,800	Lahars
Krakatau (Indonesia)	1883	36,000	Tsunami†
Awu (Indonesia)	1892	1,500	Nuées ardentes, lahars
Mont Pelée, Martinique (West Indies)	1902	36,000	Nuées ardentes
Soufrière, St. Vincent (West Indies)	1902	1,565	Nuées ardentes
Taal (Philippines)	1911	1,332	Explosions
Kelut (Indonesia)	1919	5,000	Lahars
Lamington (Papua New Guinea)	1951	3,000	Nuées ardentes, explosions
Merapi (Indonesia)	1951	1,300	Lahars
Agung (Indonesia)	1963	3,800	Nuées ardentes, lahars
Taal (Philippines)	1965	350	Explosions
Mount St. Helens (United States)	1980	57	Lateral blast, mudflows
El Chichón (Mexico)	1982	>2,000	Explosions, nuées ardentes
Nevado del Ruiz (Colombia)	1985	>25,000	Mudflows
Unzen (Japan)	1991	41	Nuées ardentes
Pinatubo (Philippines)	1991	>300	Nuées ardentes, mudflows, ash fall (roof collapse)
Merapi (Indonesia)	1994	>41	Nuées ardentes from dome collapse
Soufriere Hills, Montserrat (West Indies)	1997	19	Nuées ardentes

*Deaths directly attributable to the destruction or reduction of food crops, livestock, agricultural lands, pasturage, and other disruptions of food chain.
†A Japanese word now commonly used worldwide for earthquake- or eruption-triggered seismic sea waves.

with melted snow and ice to generate highly mobile mudflows that swept down the steep drainages flanking the volcano. These mudflows killed more than 25,000 people downvalley, resulting in the second-worst volcanic disaster in the twentieth century (the worst is the 1902 Mont Pelée eruption at Martinique).

In 1991, two eruptions captured worldwide attention. On June 3, pyroclastic flows (nuées ardentes) triggered by the collapse of a new lava dome at the summit of Unzen Volcano, Kyushu, Japan, killed 41 people, including three volcanologists filming the volcanic activity. Intermittent weaker activity persisted at Unzen, and periodic collapses of the still-growing lava dome continued to pose a volcanic hazard to the city of Shimabara downslope from the volcano. During the course of the Unzen eruption, at least 12,000 people evacuated their homes temporarily. A much larger explosive eruption occurred at Mount Pinatubo, Luzon, Philippines, on June 15–16, 1991, after a dormancy of about 600 years. This eruption caused more than 300 fatalities and widespread destruction of structures, civil works, and cropland, and forced the evacuation of nearly 80,000 people, including 17,000 U.S. military personnel and their dependents stationed at Clark Air Base. Most of the fatalities attributed to the June 15 climactic eruption were caused by the collapse of roofs, laden with ash wetted by heavy rains of typhoon Yunya, which struck the island of Luzon at the same time. Since the 1991 eruption, destructive mudflows triggered by heavy rainfall during the monsoon seasons have caused additional fatalities and considerable property damage; such posteruption mudflows will pose a continuing volcano hazards into the twenty-first century, when the debris-choked valleys draining Mount Pinatubo are expected to reestablish the preeruption stream gradients. The Pinatubo eruption ranks as the second largest eruption in the world in the twentieth century, after that of Novarupta (Katmai), Alaska, in 1912. Moreover, Pinatubo injected into the stratosphere at least twice the volume of aerosols as did El Chichón in 1982, and the resultant stratospheric volcanic cloud affected the global climate until the mid-1990s. *See* CLIMATE HISTORY.

In the Caribbean region, Mont Pelée had been responsible for the world's worst volcanic disaster, in 1902 (Table 1). In mid-1995, Soufriere Hills, a volcano on the island of Montserrat (British West Indies), which had been dormant for more than three centuries, began to erupt. The most intense activity occurred in 1996–1997, mostly involving nuées ardentes triggered by a series of collapses of actively growing, unstable lava domes; sporadic weaker

activity has continued. This eruption is not large and has produced few fatalities, but it has caused tremendous socioeconomic and political impact. All of the island's means of livelihood and infrastructure have been lost, and the present population (3500–4000) is only about a third of that before the eruption. As of 2000, the British and Montserrat governments were still undecided about the long-range plans for the island's rehabilitation and possible return of the evacuated population.

On average, about 50 to 60 volcanoes worldwide are active each year. About half of these constitute continuing activity that began the previous year, and the remainder are new eruptions. Analysis of historic records indicates that eruptions comparable in size to that of Mount St. Helens or El Chichón tend to occur about once or twice per decade, and larger eruptions such as Pinatubo about once per one or two centuries. On a global basis, eruptions the size of that at Nevado del Ruiz in November 1985 are orders of magnitude more frequent.

Scientific inquiry. It was not until the nineteenth century that serious scientific inquiry into volcanic phenomena became part of the rapidly developing science of geology. Even though the building of a small observatory was completed in 1847 on the flank of Mount Vesuvius (Italy), modern volcanology perhaps began with the founding of well-instrumented observations at Asama Volcano (Japan) in 1911 and at Kilauea Volcano (Hawaii) in 1912. The Hawaiian Volcano Observatory, located on Kilauea's caldera rim, began to conduct systematic and continuous monitoring of seismic activity preceding, accompanying, and following eruptions, as well as other geological, geophysical, and geochemical observations and investigations. Operated by the U.S. Geological Survey (USGS), the Hawaiian Volcano Observatory pioneered and refined most of the commonly used volcano-monitoring techniques that are employed by other observatories studying active volcanoes elsewhere, principally in Iceland, Indonesia, Italy, Japan, New Zealand, Lesser Antilles (Caribbean), Philippines, and Kamchatka (Russia). In response to the Mount St. Helens eruption in 1980, the David A. Johnston Cascades Volcano Observatory was established by the U.S. Geological Survey. A sister observatory to the Hawaiian observatory, it monitors the eruptions of Mount St. Helens and serves as the center for the study of the other potentially active volcanoes of the Cascade Range (in California, Oregon, and Washington).

In March 1988, the USGS, in a cooperative program with the State of Alaska and the university of Alaska, established the Alaska Volcano Observatory, with facilities and staff in both Anchorage and Fairbanks.

In 1999 the USGS formally designated a long-term program of volcano-monitoring studies at Long Valley Caldera (east-central California) as the Long Valley Observatory. The last volcanic activity in the Long Valley region was about 200 years ago. Since May 1980 the caldera has exhibited measurable volcanic unrest, as seen by greatly increased seismicity activity and an accumulated ground uplift of nearly 2 ft (0.7 m). *See* CALDERA.

Fig. 1. Climactic eruption of Mount St. Helens on May 18, 1980, about 5 h after the beginning of activity. The plume of ash and gases reached an altitude of about 15 mi (24 km). (*Photograph by R. M. Krimmel, USGS*)

Nature of magmas. The eruptive characteristics, products, and resulting landforms of a volcano are determined predominantly by the composition and physical properties of the magmas involved in the volcanic processes (**Table 2**). Formed by partial melting of existing solid rock in the Earth's lower crust or upper mantle, the discrete blebs of magma consist of liquid rock (silicate melt) and dissolved gases. Driven by buoyancy, the magma blebs, which are lighter than the surrounding rock, coalesce as they rise toward the surface to form larger masses. *See* IGNEOUS ROCKS; LITHOSPHERE.

Concentration of volatiles. During its ascent, the magma enters zones of lower temperature and pressure and begins to crystallize, producing crystals suspended in the liquid—physically analogous to the formation of ice crystals when water begins to freeze. Other solid fragments may also be incorporated from

TABLE 2. Generalized relationships between magma composition, relative viscosity, and common eruptive characteristics

Magma composition	Relative viscosity	Common eruptive characteristics
Basaltic	Fluidal	Lava fountains, flows, and pools
Andesitic	Less fluidal	Lava flows, explosive ejecta, ashfalls, and pyroclastic flows
Dacitic-rhyolitic	Viscous	Explosive ejecta, ashfalls, pyroclastic flows, and lava domes

the walls and roof of the conduit through which the magma is rising. As crystallization progresses, volatiles and the more soluble silicate components are concentrated in the remaining liquid. *See* PHENOCRYST; XENOLITH.

At some point during magma ascent, decreasing confining pressure and increasing concentration of volatiles in the residual liquid initiate the separation of gas from the liquid. From that point on to its eruption, the magma consists of three phases: liquid, solid, and gas. Volcanic gases generally are predominantly water; other gases include various compounds of carbon, sulfur, hydrogen, chlorine, and fluorine. All volcanic gases also contain minor amounts of nitrogen, argon, and other inert gases, largely the result of atmospheric contamination at or near the surface.

In laboratory experiments, at a temperature of 2000°F (1100°C) and pressure of 5 kilobars (500 megapascals), a melt of rhyolitic composition can contain in solution about 10% (by weight) of water, a basaltic melt about 8%. At lower pressure, the solubility of water in any magma decreases correspondingly. With continued ascent, water and other volatiles in excess of their solubilities in the magma will exsolve, vesiculate, and ultimately increase "gas pressure" of the magma to provide the driving force for eruptions. This process may be compared with the uncorking of a bottle of champagne, especially if it has been shaken; the gas separates from the wine and forms bubbles, which in turn expand violently (explode) when the cork is removed suddenly.

The actual proportion of gas to lava liberated during eruptions cannot be directly determined; the amount of gas can be lower or higher than the values from laboratory experiments depending on the actual crystallization and degassing histories of the magmas. For many eruptions, volatiles measured in the lava constitute less than 1% (by weight) of the lava erupted during the same interval. In initially unsaturated magmas, high gas pressures may be developed by supersaturation of volatiles as their residual liquid phase becomes concentrated during crystallization. In recent decades, a combination of refined laboratory methods (to analyze melt inclusions and phenocrysts) and modern remote-sensing techniques (to measure volcanic gases in atmosphere) have been used to obtain data for well-studied eruptions. For example, data obtained at Mount St. Helens (1980), Redoubt Volcano, Alaska (1989), and Mount Pinatubo (1991) indicate that many magma systems were gas-saturated at the time of eruption. Some magmas, however, undergo considerable preeruption degassing during shallow subsurface storage and transport. *See* CRYSTALLIZATION; MAGMA.

Part of the gas liberated at volcanoes probably comes from the same deep-seated source as the silicate portion of the magma, but some may be of shallower origin. Part of the steam may be the result of near-surface oxidation of deep-seated hydrogen or interaction of hot magma or rock with ground water or geothermal fluids in the proximity of the reservoir-conduit system. Some of the oxidation of the sulfur gases must have taken place close to the surface. At some volcanoes, such as Vesuvius, the carbon gases may derive in part from reaction of the magma with limestone at shallow depth. Ammonia and hydrocarbon components present in some gases probably are derived from the organic constituents of sedimentary rocks near the surface.

In some eruptions, such as the 1924 eruption of Kilauea, Hawaii, temperatures are low and the gas is wholly or very largely steam. Fresh magmatic material may be entirely absent. In these phreatic (steam-blast) eruptions, the steam is simply heated ground water from the rocks adjacent to the magma reservoir and volcanic conduit. In other eruptions, such as that of Parícutin, Mexico, in 1943, the large volume of steam given off simultaneously with lava and smaller amounts of magmatic gas far exceeds the theoretical saturation limit of the magma, indicating that volatilized ground water was involved.

Physical properties. Temperatures of erupting magmas have been measured in lava flows and lakes, pyroclastic deposits, and volcanic vents by means of infrared sensors, optical pyrometers, and thermocouples. Reasonably good and consistent measurements have been obtained for basaltic magmas erupted from Kilauea and Mauna Loa volcanoes, Hawaii, and a few other volcanoes. Measured temperatures typically range between 2100 and 2200°F (1150 and 1200°C), and many measurements in cooling Hawaiian lava lakes indicate that the basalt becomes completely solid at about 1800°F (980°C). Perhaps the most reliable temperature determinations of Hawaiian lavas are obtained from experimentally calibrated geothermometers, involving the precise chemical analysis of the abundance of calcium or magnesium in the glass matrix (the quenched liquid phase). At Nyamlagira volcano, central Africa, in vents and in flows close to vents, temperatures ranged from about 1900 to 2000°F (1040 to 1095°C). Temperatures during the 1950–1951 eruption of Oshima, Japan, were in the same range. Locally, temperatures as high as 2550°F (1400°C) have been reported but not well documented; these anomalously high temperatures have been ascribed to the burning of volcanic gases in vents. *See* GEOLOGIC THERMOMETRY.

Temperature measurements on more silicic lavas are few and much less accurate because of the greater violence of the eruptions and the necessity of working at considerable distances as a safety precaution. In general, however, they suggest lower temperatures of eruption than those for mafic lavas. For example, for andesitic and more silicic lavas, available temperature estimates have ranged from about 1800 to 1330°F (980 to 720°C). Thermocouple measurements, made only 20 h after eruption of dacitic pyroclastic flows of Mount St. Helens (August 1980), yielded a temperature range of 1337–1540°F (725–838°C).

A few field measurements of the viscosity of flowing basic lavas have been made by means of penetrometers (instruments that measure the rate of

TABLE 3. General relationships between types of eruptions and some eruptive characteristics*

Volcanic explosivity index (VEI)	0†	1	2	3	4	5	6	7	8
Description	Nonexplosive	Small	Moderate	Mod-Large	Large	Very large			
Volume of ejecta, m^3‡	$<10^4$	10^4–10^6	10^6–10^7	10^7–10^8	10^8–10^9	10^9–10^{10}	10^{10}–10^{11}	10^{11}–10^{12}	$>10^{12}$
Ash plume height, km§	< 0.1	0.1–1	1–5	3–15	10–25	>25			
Qualitative description	Gentle, effusive		Explosive		Severe, violent, terrific	Cataclysmic, paroxysmal, colossal			
Eruption type	Hawaiian	Strombolian	Vulcanian		Plinian	Ultraplinian			
Eruptions (total in file)	699	845	3477	869	278	84	39	4	0

*Modified from T. Simkin and L. Siebert, *Volcanoes of the World: A Regional Directory, Gazetteer, and Chronology of Volcanism During the Last 10,000 Years,* 2d ed., 1994.
†Nonexplosive "Hawaiian" eruptions are assigned a VEI of zero regardless of the volume of lava.
‡1 m^3 = 35.3 ft^3.
§For VEIs 0–2, given as kilometers above crater; for VEIs 3–8, given as kilometers above sea level. 1 km = 0.6 mi.

penetration into liquid of a slender rod under a given strength of thrust) and by the shearing resistance to the turning of a vane immersed in the liquid. Viscosities also have been calculated from observed rates of flow in channels of known dimensions and slope or have been estimated from the chemical composition of the lava (by extrapolation of laboratory data on viscosities of simple molten silicate compounds). The best direct determinations of viscosity are vane-shear measurements in a lava lake at Kilauea, which yielded a best estimate of about 1900 poises (190 Pa · s) at a temperature of 2100°F (1150°C). Calculations based on rate of flow at both Kilauea and Mauna Loa, Hawaii, gave viscosities of 3000–4000 poises (300–400 Pa · s) for lava close to the vents, increasing at greater distances from the vents as the lava cools and stiffens to immobility. At Hekla, Iceland, a somewhat more silicic lava in the vent had a viscosity of about 10,000 poises (1000 Pa · s); and at Oshima, Japan, the lowest viscosities in two streams near the vent during the 1951 eruption were 5600 and 18,000 poises (560 and 1800 Pa · s), respectively. In general, empirical observations of flow behavior of more silicic lavas indicate that they are more viscous than basaltic lavas, but direct field measurements of their viscosities have not yet been made. *See* LAVA; MAGMA; PYROCLASTIC ROCKS; VISCOSITY.

Types of volcanic eruptions. The character of a volcanic eruption is determined largely by the viscosity of the liquid phase of the erupting magma and the abundance and condition of the gas it contains. Viscosity is in turn affected by such factors as the chemical composition and temperature of the liquid, the load of suspended solid crystals and xenoliths, the abundance of gas, and the degree of vesiculation. In very fluid lavas, small gas bubbles form gradually, and generally are able to rise through the liquid, coalescing to some extent to form larger bubbles, and escape freely at the surface with only minor disturbance. In more viscous lavas, the escape of gas is less free and produces minor explosions as the bubbles burst their way out of the liquid. In still more viscous lavas, at times there appears to be a tendency for the essentially simultaneous formation of large numbers of small bubbles throughout a large volume of liquid. The subsequent violent expansion of these bubbles during eruption shreds the frothy liquid into tiny fragments, generating explosive showers of volcanic ash and dust, accompanied by some larger blocks (volcanic "bombs"); or it may produce an outpouring of a fluidized slurry of gas, semisolid bits of magma froth, and entrained blocks to form high-velocity pyroclastic flows, surges, and glowing avalanches (nuées ardentes). Also, rising gases may accumulate beneath a solid or highly viscous plug, clogging the vent until it acquires enough pressure to cause rupture and attendant explosion that hurls out fragments of the disrupted plug.

Types of eruptions customarily are designated by the name of a volcano or volcanic area that is characterized by that sort of activity (**Table 3**), even though all volcanoes show different modes of eruptive activity on occasion and even at different times during a single eruption.

Eruptions of the most fluid lava, in which relatively small amounts of gas escape freely with little explosion, are designated Hawaiian eruptions. Most of the lava is extruded as successive, thin flows that travel many miles from their vents. Lava clots or spatter thrown into the air in fountains (**Fig. 2**) may remain fluid enough to flatten out on striking the ground, and commonly to weld themselves to form cones of spatter and cinder. An occasional feature of Hawaiian activity is the lava lake, a pool of liquid lava with convectional circulation that occupies a preexisting shallow depression or pit crater. Recent data, however, indicate that eruptions of Hawaiian volcanoes—the namesake for the "Hawaiian" (dominantly effusive) type of eruptive activity—may not be as nonexplosive as suggested by observations of the historical eruptions at the Kilauea and Mauna Loa volcanoes. Geological and dating studies of prehistoric volcanic ash deposits demonstrate that the frequency of explosive eruptions from Kilauea is comparable to those for many of the composite volcanoes of the Cascade Range of the Pacific Northwest.

Strombolian eruptions are somewhat more explosive eruptions of lava, with greater viscosity, and produce a larger proportion of pyroclastic material. Many of the volcanic bombs and lapilli assume rounded or drawn-out forms during flight, but commonly are sufficiently solid to retain these shapes on impact.

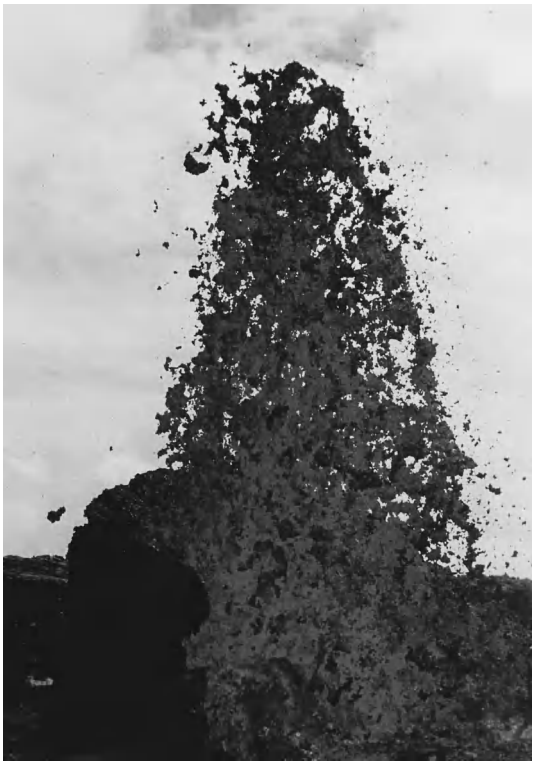

Fig. 2. Typical of "Hawaiian eruptions," an 80-ft-high (25-m) fountain of fluid basaltic lava plays during an eruption of Kilauea volcano, Hawaii, in 1973. Compare with the plinian eruption of Mount St. Helens in Fig. 1, involving much more viscous dacitic lava. (*Photograph by R. I. Tilling, USGS*)

Generally still more explosive are the vulcanian type of eruptions. Angular blocks of viscous or solid lava are hurled out, commonly accompanied by voluminous clouds of ash but with little or no lava flow.

Peléean eruptions are characterized by the heaping up of viscous lava over and around the vent to form a steep-sided hill or volcanic dome. Explosions, or collapses of portions of the dome, may result in glowing avalanches (nuées ardentes).

Plinian eruptions are paroxysmal eruptions of great violence—named after Pliny the Elder, who was killed in A.D. 79 while observing the eruption of Vesuvius—and are characterized by voluminous explosive ejections of pumice and by ash flows. The copious expulsion of viscous siliceous magma commonly is accompanied by collapse of the summit of the volcano, forming a caldera, or by collapse of the broader region, forming a volcano-tectonic depression. The term ultraplinian has been used occasionally by some volcanologists to describe especially vigorous plinian activity.

In contrast to the foregoing magmatic eruptions, some low-temperature but vigorous ultravulcanian explosions throw out fragments of preexisting volcanic or nonvolcanic rocks, accompanied by little or no new magmatic material. Certain explosion pipes and pits, known as diatremes and maars, have been produced by ultravulcanian explosions.

Volatilization of meteoric or ground water when it comes in contact with hot solid rocks in the vicinity of the volcano's magma reservoir causes the usually mild, but occasionally violent, disturbances known as phreatic explosions. No fresh magmatic material is erupted during phreatic (steam-blast) activity, which may precede major magmatic eruptions. The term "phreatomagmatic" describes a highly variable type of volcanic activity that results from the complex interaction between fresh magma/lava and subsurface or surface water (ground water, hydrothermal water, meteoric water, seawater, lake water).

The differences between the different types of eruptions are gradational but ultimately are dependent on the variation in explosivity and "size" of the eruption. Some volcanologists have proposed the volcanic explosivity index (VEI) to attempt to standardize the assignment of the size of an explosive eruption, using volume of eruptive products, duration, height of ash plume, and other criteria. Table 3 shows the general relationships and the necessarily arbitrary distinctions between eruption type, explosivity, and eruptive volume for nearly 6300 eruptions during the Holocene (the past 10,000 years of the Earth's history), for which such information is known or can be reasonably estimated. Of these eruptions, about 11% are nonexplosive Hawaiian-type eruptions (always assigned a VEI of 0 regardless of eruptive volume). Explosive eruptions considered to be small to moderate in size (VEI 1 to 2) constitute about 69% of all eruptions and are gradationally classified as Hawaiian, strombolian, or vulcanian. Only 127 plinian or ultraplinian eruptions rate VEIs of 5 or greater (very large). The May 18, 1980, eruption of Mount St. Helens rated a VEI of 5, but just barely. During the past 10,000 years, only four eruptions rated VEIs of 7, including the 1815 eruption of Tambora (Indonesia), the largest known eruption in the world in recorded history. Several of the large pre-Holocene caldera-forming eruptions would qualify for VEI ratings of 8. The VEI scale, like the Richter magnitude scale for earthquakes, is open-ended; that is, there is no maximum rating. However, to date, no eruptions of VEI 9 size (eruptive volume of 10^4 km^3) have been recognized, and magma reservoirs of such volume are virtually unknown in the geologic record. *See* HOLOCENE.

Monitoring active volcanoes. A major component of the science of volcanology is the systematic and, preferably, continuous monitoring of active and potentially active volcanoes. Scientific observations and measurements—of the visible and invisible changes in a volcano and its surroundings—between eruptions are as important, perhaps even more crucial, than during eruptions. Measurable phenomena important in volcano monitoring include earthquakes; ground movements; variations in gas compositions; and deviations in local gravity, electrical, and magnetic fields. These phenomena reflect pressure and stresses induced by subsurface magma movements and or pressurization of the hydrothermal envelope surrounding the magma reservoir.

Seismicity and ground deformation. The monitoring of volcanic seismicity and ground deformations before, during, and following eruptions has provided

the most useful and reliable information. From the many decades of study of the active Hawaiian volcanoes and volcanoes elsewhere, it has been determined that a volcano generally undergoes measurable ground deformation when magma is fed into its near-surface reservoir-conduit system.

Vertical and horizontal ground displacements and slope changes can easily be detected and measured precisely by existing geodetic techniques. Slope changes can be measured with a precision of a microradian or less by various electronic-mechanical "tilt-meters," or with somewhat less precision by leveling of short-sided arrays of benchmarks. Vertical displacements of benchmarks on a volcano, relative to a reference benchmark (or tide gage) unaffected by the volcano, can be determined to a few parts per million by leveling surveys. Horizontal distance changes between benchmarks can be measured with similar precision by various electronic distance measurement instruments.

Advances in geodetic applications of satellite positioning and other forms of space geodesy, especially the Global Positioning System (GPS), suggest that conventional ground-deformation monitoring techniques will be supplanted by satellite-based monitoring systems, if the acquisition and maintenance costs can be decreased substantially. Toward the end of the twentieth century, significant progress was made in the development and testing of near-real-time GPS monitoring networks at selected volcanic systems [for example, Augustine Volcano (Alaska), Long Valley Caldera (California)]. Another promising satellite-based geodetic technique is the InSAR (interferometric synthetic aperture radar) method, which is capable of detecting ground movements over extensive areas (that is, over the entire volcano). This technique involves the interferometric analyses of one or more pairs of SAR images acquired at time intervals; coherent ground movement is revealed by ringlike or other regular anomalies reflecting differences (interferences) in topography between the time-separated pair of satellite images being compared. The InSAR technique, while not yet a routine volcano-monitoring tool, has been shown to be successful at Yellowstone (Wyoming) and Long Valley calderas and some other volcanoes [such as, Mount Etna (Italy) and Okmok (Alaska)]. *See* GEODESY; SATELLITE NAVIGATION SYSTEMS; SEISMOGRAPHIC INSTRUMENTATION.

The ground deformation is related to, and accompanied by, intense earthquake activity reflecting the subsurface ruptures of the confining rocks of the expanding volcanic reservoir in response to the increased pressure exerted by infilling magma. Modern volcano observations employ well-designed seismic networks to monitor volcanic seismicity continuously in order to track subsurface movement of magma between and during eruptions. For well-studied volcanoes, experience has shown that premonitory seismicity usually provides the earliest signals of impending activity. Great advances have been made in volcano seismology since the early 1980s; in particular, recent studies demonstrate that the occurrence of, and variations in, long-period (low-frequency) events and volcanic tremor provide valuable insights in inferring the movement of magma and/or hydrothermal fluids before and during eruptions.

The number of earthquakes and the magnitude of ground deformation gradually increase as the magma reservoir swells or inflates until some critical strength threshold is exceeded and major, rapid migration of magma ensues to feed a surface eruption or a subsurface intrusion (magma drains from the reservoir and is injected into another part of the volcanic edifice without breaching the surface). With the onset of eruptive or intrusive activity, pressure on the "volcanic plumbing system" is relieved and the reservoir abruptly shrinks or deflates, causing flattening of slope (tilt), reduction in vertical or horizontal distances between surface points, and decrease in earthquake frequency. One or more of these inflation-deflation cycles may take place during an eruption, depending on its duration (**Figs. 3** and **4**). *See* EARTHQUAKE; SEISMOLOGY.

Gas emission and other indicators. As magma ascends, the emission rate or composition of gases exsolving from it and those generated by interaction of ground water and hot rock or magma may change with time. The systematic measurement at volcanic vents and fumaroles of such variations, though still largely experimental, shows much promise as another volcano monitoring tool. With modern methods (such as gas chromatography and mass spectroscopy), the composition of gas can be determined routinely soon after its collection. A field gas chromatograph has been developed and has been tested at several active volcanoes, including Etna (Italy), Kilauea (Hawaii), and Merapi (Indonesia). At Kilauea, volcanic gases have been monitored at 25 sites by sampling twice a week and analyzing chromatographically more than 10 different gas species. *See* GAS CHROMATOGRAPHY.

Some success has been obtained at Etna, Mount St. Helens, Galunggung (Indonesia), Mount Pinatubo (Luzon, Philippines) and other volcanoes in measuring the fluctuation in the emission rate of sulfur dioxide by means of a correlation spectrometer (COSPEC) in both ground-based and airborne modes. The emission rate of carbon dioxide (CO_2) was monitored remotely, using a modified infrared spectrophometer during some of the 1980–1981 eruptions of Mount St. Helens. Another recently developed remote gas-monitoring technique uses Fourier transform infrared (FTIR) spectroscopy for measurement of sulfur dioxide (SO_2), the ratio of sulfur dioxide to hydrochloric acid (SO_2/HCl), and silicon tetrafluoride (SiF_4) in volcanic plumes; this technique has been applied at the Kilauea (Hawaii), Etna and Vulcano (Italy), and Asama and Unzen (Japan) volcanoes. The continuous monitoring of emission at selected sites, together with periodic regional surveys, has demonstrated that the rate of carbon dioxide emission in the Long Valley Caldera region greatly increased in 1989. In areas of highest emission, the high CO_2 concentrations in the soil are killing trees by denying their roots oxygen and by interfering with

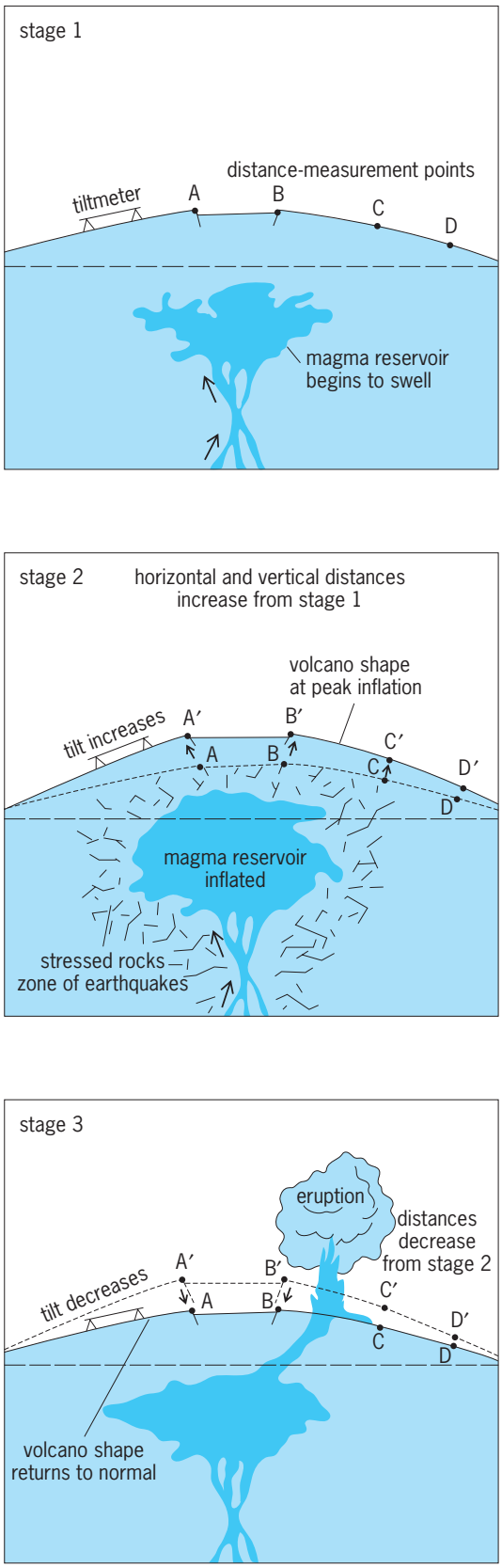

Fig. 3. Schematic diagrams showing three commonly observed stages in the course of an inflation-deflation cycle during a typical Hawaiian eruption. At stage 1, inflation begins; it peaks at stage 2. Stage 3 is the eruption-deflation stage. (*After R. I. Tilling, Monitoring Active Volcanoes, USGS, 1983*)

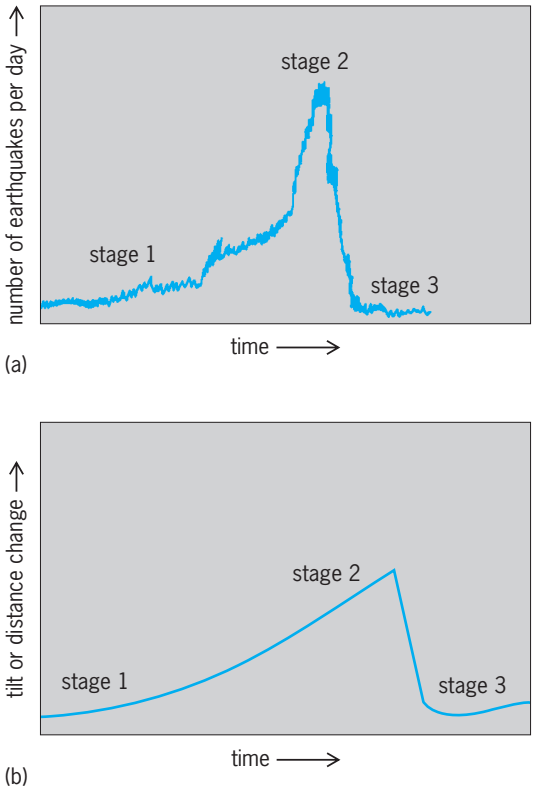

Fig. 4. Idealized graphs of (*a*) earthquake frequency and (*b*) tilt or distance changes as a function of time during the three stages of Fig. 3. (*After R. I. Tilling, Monitoring Active Volcanoes, USGS, 1983*)

nutrient uptake. Large but gradually declining volumes of CO_2 gas continue to seep from this volcanic system. The periodic determination of the absorption of certain gases (such as chlorine and sulfur) in alkaline solutions provides a time-integrated measure of their emission.

A significant advance in the monitoring of volcanic gas emission was the accidental discovery in the early 1980s that, with modification in computer algorithm and data processing, the Total Ozone Mapping Spectrometer (TOMS) instrument aboard the *Nimbus 7* satellite can be used to measure the output of sulfur dioxide during an eruption and to track the movement and, ultimately, the dissipation of the sulfur dioxide–containing stratospheric volcanic clouds. The combination of COSPEC and TOMS measurements of sulfur dioxide provides a powerful volcano-monitoring tool. *See* SPECTROSCOPY.

Periodic measurement of gas composition and emission rate, though providing important information about the volcanic system, does not, however, give the critical data on short-term or continuous fluctuations—as can seismic and ground deformation monitoring—that might augur an impending eruption. Some attempts have begun to develop continuous real-time monitoring of certain of the relatively abundant and nonreactive gases, such as hydrogen, helium, and carbon dioxide. Particularly encouraging is the continuous monitoring of hydrogen emission, utilizing electrochemical sensors and

satellite telemetry, that has been tested at Mount St. Helens, Kilauea, Mauna Loa, and Long Valley caldera (California) in the United States, and at Vulcano and other sites in Italy. Nonetheless, gas-monitoring techniques must be considered largely experimental, in large measure because of the fugitive nature of gases, their complex pathways from magmatic source to measurement site, and interactions with the hydrology of the volcano and its surroundings.

Even more experimental than gas monitoring are techniques involving measurements of changes in the thermal regime, in gravitational or geomagnetic field strength, and in various measurable "geoelectrical" parameters, such as self-potential, resistivity, and very low-frequency signals from distant sources. All of these methods are premised on the fundamental model that an inflating (or expanding) magma reservoir causes thermal anomalies and related magma-induced changes in bulk density, electrical and electromagnetic properties, piezoelectric response, and so forth, of the volcanic edifice.

Primarily because of the complicating effects of near-surface convective heat transfer, thermal monitoring of volcanoes generally has proved to be nondiagnostic, although monitoring of temperature variations in crater lakes at some volcanoes—Taal (Philippines) and Kelut (Indonesia)—has given useful information.

Although modern gravity meters can detect changes as small as 5 microGal (0.05 micrometer/s^2), interpretation of the results of gravity surveys is highly dependent on the ability to discriminate between elevation-induced and mass-difference changes. Preliminary analysis of some electrical self-potential data at Kilauea suggests that this technique may provide more advanced notice of impending eruptions than seismic and ground deformation monitoring. However, much more data and testing are required before geoelectrical monitoring techniques can be considered to be as routinely and universally applicable as seismic and ground deformation methods. *See* EARTH, GRAVITY FIELD OF; GEOELECTRICITY; GEOPHYSICAL EXPLORATION.

Volcanic hazards mitigation. Volcanoes are in effect windows into the Earth's interior; thus research in volcanology, in contributing to an improved understanding of volcanic phenomena, provides special insights into the chemical and physical processes operative at depth. However, volcanology also serves an immediate, perhaps more important, role in the mitigation of volcanic and related hydrologic hazards (mudflows, floods, and so on). Progress toward hazards mitigation can best be advanced by a combined approach. One aspect is the preparation of comprehensive volcanic hazards assessments of all active and potentially active volcanoes, including a volcanic risk map for use by government officials in regional and local land-use planning to avoid high-density development in high-risk areas. The other component involves improvement of predictive capability by upgrading volcano-monitoring methods and facilities to adequately study more of the most dangerous volcanoes. An improved capability for eruption forecasts

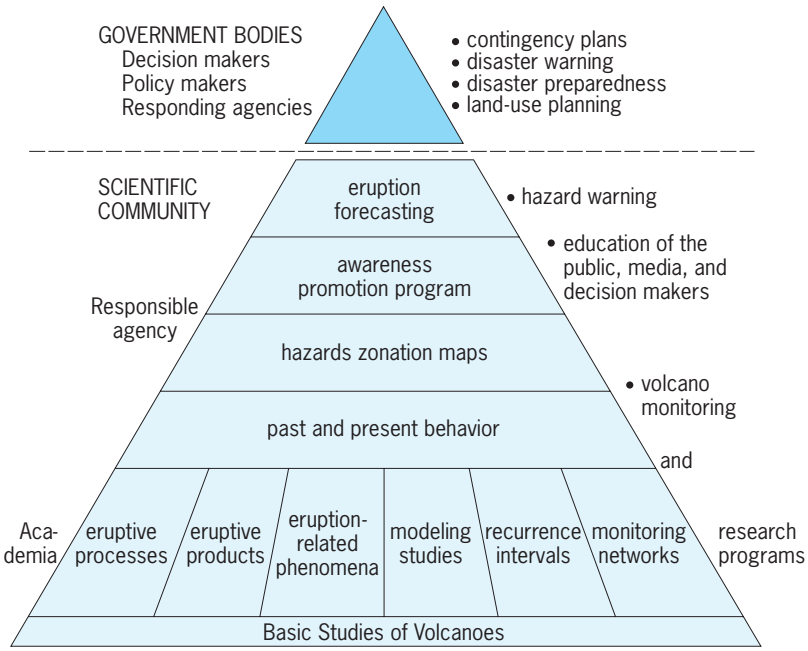

Fig. 5. Schematic diagram showing the general process of volcanic hazards assessment and development of mitigation strategies. (*After R. I. Tilling, ed., Volcanic Hazards, American Geophysical Union, 1989*)

and predictions would permit timely warnings of impending activity, and give emergency-response officials more lead time for preparation of contingency plans and orderly evacuation, if necessary.

Both these approaches must be buttressed by long-term basic field and laboratory studies to obtain the most complete understanding of the volcano's prehistoric eruptive record (recurrence interval of eruptions, nature and extent of eruptive products, and so on). Progress in mitigation of volcanic hazards must be built on a strong foundation of basic and specialized studies of volcanoes (**Fig. 5**). The separation of the apex from the rest of the triangle reflects the fact that "decision makers" must consider administrative and socioeconomic factors in addition to scientific information from volcanology. Nonetheless, volcanologists and other physical scientists—individually and collectively—must step out of their traditional academic roles and work actively with social scientists, emergency-management officials, educators, the news media, and the general public to increase public awareness of volcanoes and their potential hazards. Recent volcanic crises and disasters have shown that scientific data—no matter how much or how precise—serve no purpose in volcanic-risk management unless they are communicated effectively to, and acted upon in a timely manner by, the civil authorities.

One of the best examples of good volcanic hazards assessment and eruption forecasts is provided by the Mount St. Helens experience. As early as 1975, scientists predicted that this volcano was the one in the conterminous United States most likely to reawaken and to erupt, possibly before the end of the century. This prophecy was followed in 1978 by a detailed analysis of the types, magnitudes, and areal extents

Fig. 6. Climatic eruption of Mount Pinatubo (Luzon, Philippines) on June 15, 1991, with U.S. Clark Air Base in the foreground. The base of the eruption column measures about 16 mi (25 km) across. (*USAF photograph by Robert Lapointe*)

of potential volcanic hazards that might be expected from a future eruption of Mount St. Helens. The volcano erupted catastrophically on May 18, 1980, and the volcanic events and associated hazards largely followed the scenario outlined earlier. Because of improved volcano monitoring with the establishment of the David A. Johnston Cascades Volcano Observatory, all of the eruptions of Mount St. Helens since June 1980 have been predicted successfully.

The response by United States and Philippine volcanologists to the reawakening of Mount Pinatubo in early April 1991 (**Fig. 6**) constitutes another successful case in volcanic hazards mitigation. Seismic and volcanic gas monitoring studies prompted scientists to recommend to the U.S. and Philippine governments evacuation of the surrounding region and so more than 200,000 people were moved. About 36 h later, the cataclysmic eruption of June 15–16 took place, obliterating the region within a 5–6-mi (8–10-km) radius. Given the huge volume of the eruption (preliminary estimate is about 0.5 mi^3 or 2 km^3) and the widespread devastation, the death toll almost certainly would have been tens of thousands, rather than hundreds.

In contrast, the other two major volcanic disasters in the 1980s—El Chichón (1982) and Nevado del Ruiz (1985) [Table 1]—are sobering examples of failures in mitigation of hazards. In the case of El Chichón, the eruption came as a surprise because there were no geologic data about its eruptive history and no preeruption volcanic hazards assessment or monitoring. The Ruiz catastrophe, however, is a more tragic case because considerable geologic data existed; warning signs were recognized a year earlier and limited volcano monitoring was initiated; and a preliminary hazards-zonation map, produced more than a month before the destructive eruption, correctly pinpointed the areas of greatest hazard. Sufficient warnings were given by the scientists on site at Ruiz, but for reasons still not clear, effective evacuation and other emergency measures were not implemented by the government authorities.

In contrast, the volcanic crisis at Rabaul Caldera (Papua New Guinea) in the mid-1980s offers an excellent example of how an effective hazards-mitigation program can save lives, even though the anticipated eruption came a decade later than the original forecast. Beginning in the early 1970s, Rabaul began to exhibit signs of volcanic unrest, as seen by periodic earthquake swarms beneath the caldera and by ground uplift. During 1983 and 1984, the volcano's activity increased dramatically, with the monthly counts of earthquakes approaching 10,000 late in 1983 (**Fig. 7***a*). This sudden escalation in seismicity was accompanied by sharply increasing rates of ground deformation, as indicated by geodetic monitoring. In October 1983, after considering socioeconomic factors and the scientific information provided by the Rabaul Volcano Observatory on the status of the volcanic unrest, Rabaul government officials declared a stage-2 alert, which implied that an eruption would occur within a few months. In response to this declaration, the citizens of Rabaul Town were made aware of the need to prepare for possible evacuation, staging areas for evacuation were designated, certain roads were widened to serve as evacuation routes, and several evacuation drills were conducted.

The rate of seismicity and ground deformation continued to increase for another 6 months following the declaration of the stage-2 alert, but then the level of unrest declined rapidly (Fig. 7*a*). The expected eruption did not happen, and the officials ended the alert in November 1984. For the next 10 years, caldera activity fluctuated with relatively low levels, but slightly higher than pre-1983 rates. Then, on September 19, 1994, following only 27 hours of precursory seismic and ground-deformation activity, explosive eruptions began at Vulcan and Tavurvur, the two vents on opposite sides of the caldera that have been the sites of previous historical eruptions (Fig. 7*b*). While Rabaul Town suffered massive destruction (principally from ash accumulated and roof collapse) and over 50,000 people were displaced, fewer than 10 people were killed (several from automobile accidents)—quite remarkable considering the rapid onset of the eruption with virtually no public warning to the population. The main reason for the low fatalities of the 1994 eruption was that people quickly "self-evacuated" with the first light ash falls. The people apparently had learned their lessons well years earlier, from the heightened awareness of volcanic hazards generated by the 1983–1985 volcanic crisis as well as from memories or stories of the 1937 eruption (similarly involving eruptive outbreaks at Vulcan and Tavurvur) among the long-time residents and their families.

Volcanic ash and aviation safety. During recent decades, with the advent of high-performance jet engines, an unrecognized volcano hazard emerged: in-flight encounters between jet aircraft and volcanic ash clouds produced by powerful explosive eruptions. This hazard stems from the following: (1) volcanic ash clouds are not detectable by the aircraft's onboard radar instrumentation; and (2) if the gritty,

jagged ash particles are ingested into the aircraft's jet engines, the high operating temperatures can partially melt the ash. Severe abrasion and ash accumulation within the engine, along with adherence of melted ash to critical engine parts and openings, combine to degrade engine performance and, at worst, can cause engine flameout and power loss. Since the early 1970s, more than 60 volcanic ash–aircraft encounters have occurred, with several of the aircraft experiencing total power loss and requiring emergency landings; fortunately, to date, no fatal crashes have resulted from such encounters. However, many millions of dollars of damage to aircraft have been incurred; for example, an encounter between a Boeing-747 jetliner and the ash cloud from an eruption of Redoubt Volcano (Alaska) in December 1989 required more than $80 million dollars to replace all four engines and repair other damage. Volcanologists worldwide are now working closely with the air-traffic controllers, civil aviation organizations, and the air-carrier industry to mitigate the hazards of volcanic ash to civil aviation. A part of this effort involves the operation of nine regional Volcanic Ash Advisory Centers around the world to provide early warning of explosive eruptions and to issue advisories of potentially dangerous volcanic ash clouds produced by them.

Impact of explosive volcanism on global climate. Large explosive eruptions that eject copious amounts of volcanic aerosols into the stratosphere also can affect climate on a global basis. For example, long-lingering stratospheric volcanic clouds—for example, from the great 1883 eruption of Krakatau (between Java and Sumatra, Indonesia), the 1963 eruption of Agung volcano (Bali, Indonesia), and the 1982 eruption of El Chichón (Chiapas, Mexico)—produced spectacular sunrises and sunsets all over the Earth for many months because of the interaction of suspended aerosols and the atmosphere. Studies indicate that the sulfate aerosols in volcanic clouds form a layer of sulfuric acid droplets. This layer tends to cool the troposphere by reflecting solar radiation, and to warm the stratosphere by absorbing radiated Earth heat; in general, the combined effect is to lower the Earth's surface temperature.

The 1980 Mount St. Helens eruption apparently had minimal climatic impact, producing at most only a decrease of 0.18°F (0.1°C) in average temperature for the Northern Hemisphere. In contrast, the 1982 El Chichón and the 1991 Pinatubo eruptions lowered temperature by 0.36–0.9°F (0.2–0.5°C). Significantly, the Mount St. Helens magma contains much less sulfur [<0.01 wt % of sulfur trioxide (SO_3)] than those of El Chichón or Pinatubo, 1–4 and 0.3–0.5 wt % respectively. Perhaps the best-known example of eruption-induced climate change is that associated with the eruption of Tambora volcano (Indonesia) in 1815, the most powerful eruption in historical time. Average temperature in the Northern Hemisphere was lowered by as much as 0.9°F (0.5°C), and parts of the United States and Canada experienced unusual summer frosts and loss of crops; 1816 was known as the year without summer. On a local scale, the fallout from gas-rich eruption plumes and associated acid rains can cause damage to crops and exposed metal parts of dwellings. *See* ACID RAIN; AEROSOL; AIR POLLUTION.

Robert I. Tilling

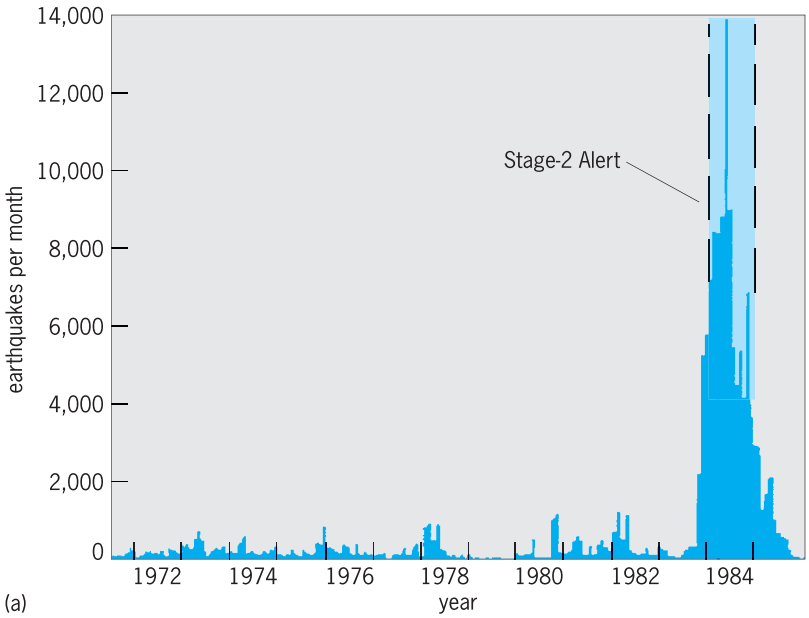

Fig. 7. Activity at Rabaul Caldera, Papua New Guinea. (*a*) Earthquakes per month from July 1971 through late 1985. The sharp increase in earthquakes in late 1983, combined with a cumulative uplift of about 2 m (6.6 ft) since 1973, prompted government officials to declare a stage-2 alert (in effect October 29, 1983, to November 22, 1984), in anticipation of a possible eruption within a few months. However, the eruption did not occur until mid-September 1994, about a decade after the alert was lifted. (*After R. I. Tilling, The role of monitoring in forecasting volcanic events, in B. McGuire et al., eds., Monitoring Active Volcanoes: Strategies, Procedures and Techniques, pp. 369–401, UCL Press Limited, London, 1995*) (*b*) Photograph, taken by the space shuttle astronauts, of the 18-km-high (11-mi) volcanic plume produced by the Rabaul eruption, about 24 hours after it began on September 19, 1994. (*NASA*)

Bibliography. R. J. Blong, *Volcanic Hazards: A Sourcebook on the Effects of Eruptions*, 1984; F. R. Boyd, Jr. (ed.), *Explosive Volcanism: Inception, Evolution, and Hazards*, 1984; T. J. Casadevall (ed.), *Volcanic Ash and Aviation Safety: Proceedings of the 1st International Symposium on Volcanic Ash and Aviation Safety*, 1994; R. W. Decker and

B. Decker, *Volcanoes*, 3d ed., 1998; R. W. Decker, T. L. Wright, and P. H. Stauffer (eds.), *Volcanism in Hawaii*, 1987; B. L. Foxworthy and M. Hill, *Volcanic Eruptions of 1980 at Mount St. Helens: The First 100 Days*, 1982; P. Francis, *Volcanoes: A Planetary Perspective*, 1993; R. W. Johnson, *Volcanic Eruptions and Climate Change*, 1993; P. W. Lipman and D. R. Mullineaux (eds.), *The 1980 Eruptions of Mount St. Helens, Washington*, 1981; G. A. Macdonald, *Volcanoes*, 1972; W. J. McGuire, C. R. J. Kilburn, and J. B. Murray (eds.), *Monitoring Active Volcanoes: Strategies, Procedures, and Techniques*, 1995; C. G. Newhall and R. S. Punongbayan (eds.), *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines*, 1996; R. Scarpa and R. I. Tilling, *Monitoring and Mitigation of Volcano Hazards*, 1996; P. D. Sheets and D. K. Grayson (eds.), *Volcanic Activity and Human Ecology*, 1979; H. Sigurdsson et al. (eds.), *Encyclopedia of Volcanoes*, 2000; T. Simkin and L. Siebert, *Volcanoes of the World: A Regional Directory, Gazetteer, and Chronology of Volcanism During the Last 10,000 Years*, rev. ed., 1994; R. I. Tilling, *Eruptions of Mount St. Helens: Past, Present, and Future*, rev. ed., 1990; R. I. Tilling (ed.), *Volcanic Hazards*, 1989; R. I. Tilling, C. C. Heliker, and T. L. Wright, *Eruptions of Hawaiian Volcanoes: Past, Present, and Future*, 1987; R. I. Tilling and P. W. Lipman, *Lessons in Reducing Volcano Risk*, 1993; H. Williams and A. R. McBirney, *Volcanology*, 1979.

Volt-ampere

The unit of total apparent power delivered to a load by an alternating current. A circuit branch with E volts across its two terminals, carrying I amperes from one to the other, is said to be receiving EI volt-amperes of apparent power, whatever may be the phase lag θ of current behind voltage. If such a load is driven through a transformer, the volt-amperes into the transformer primary is the same number as the volt-amperes into the load. Other combinations of circuit branches and their volt-amperes (for example, polyphase systems) must be handled carefully, using Eqs. (1) and (2). In Eq. (2), the volt-ampere reactive

$$\begin{aligned}\text{Real power} &= \text{average power delivered to load}\\ &= \text{apparent power} \times \text{power factor}\\ &= EI\cos\theta \quad \text{watts} \qquad (1)\end{aligned}$$

$$\begin{aligned}&\text{Reactive power}\\ &\quad = \text{apparent power} \times \text{reactive power factor}\\ &\quad = EI\sin\theta \quad \text{volt-amperes reactive} \qquad (2)\end{aligned}$$

(var) is 1 volt x 1 ampere when voltage and current are 90° out of phase. The reactive power factor is positive for inductive-reactance loads and negative for capacitive-reactance loads.

Practical computation of real power, reactive power, and apparent power is usually done with complex-number algebra using the geometrical diagram of **Fig. 1**. Clearly Eqs. (3) and (4) hold.

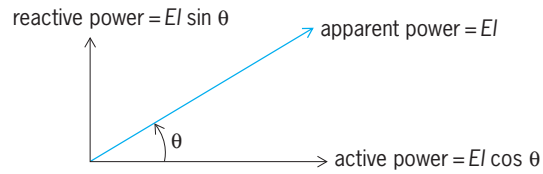

Fig. 1. Apparent, active, and reactive power.

$$\begin{aligned}&(\text{Apparent power})^2\\ &\quad = (\text{real power})^2 + (\text{reactive power})^2 \qquad (3)\\ &(\text{Volt-amperes})^2 = (\text{watts})^2 + (\text{vars})^2 \qquad (4)\end{aligned}$$

See ALTERNATING-CURRENT CIRCUIT THEORY.

Measurement. Alternating-current voltage, current, and real (or active) power can be measured by voltmeters, ammeters, and wattmeters. Reactive power, the average product of instantaneous current and out-of-phase voltage, can be measured with a varmeter (**Fig. 2**), which is in essence a wattmeter movement with provision for adjusting potential-coil current to phase shift of 90° exactly. A special meter to register volt-amperes is seldom used. *See* AMMETER; WATTMETER.

Combinations. Electrical systems of practical interest usually have many branches and possibly several driving generators. The composite voltage-and-current sums given by Eqs. (5) and (6) can be found

$$\sum_{\substack{\text{All}\\ \text{generators}}} \begin{pmatrix}\text{real}\\ \text{power in}\end{pmatrix} = \sum_{\substack{\text{All}\\ \text{branches}}} \begin{pmatrix}\text{real power}\\ \text{into branch}\end{pmatrix} \qquad (5)$$

$$\sum_{\substack{\text{All}\\ \text{generators}}} \begin{pmatrix}\text{reactive}\\ \text{power in}\end{pmatrix} = \sum_{\substack{\text{All}\\ \text{branches}}} \begin{pmatrix}\text{vars into}\\ \text{branch}\end{pmatrix} \qquad (6)$$

for such combinations. These relations are much used by power system engineers, and show that apparent power does not add arithmetically when several circuit branches are combined. The rule of addition for real power, Eq. (5), is a result of the conservation of energy. The rule of addition for reactive power, Eq. (6)—with proper subtraction of

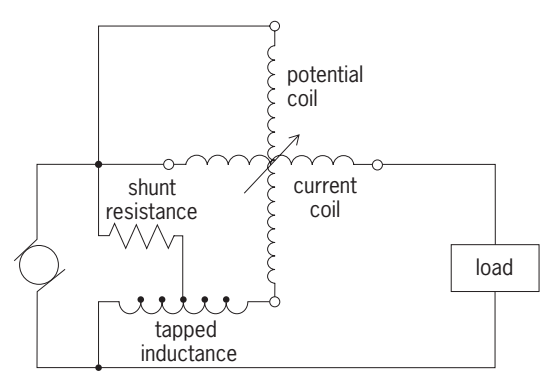

Fig. 2. Single-phase reactive power measurement which uses an inductance.

capacitive vars from inductive vars—is a mathematical statement and must be proved as such.

When the sums of volt-amperes, watts, and volt-amperes reactive are meaningful for multigenerator, multiload systems, the meanings are those obtained from Eqs. (5) and (6). A case of special interest is that of polyphase systems. Real power (watts) always exists, and the composite sums of volt-amperes and vars can be useful if the system is approximately balanced. A badly unbalanced polyphase system is not often described in terms of apparent power. *See* ELECTRIC POWER MEASUREMENT. Mark G. Foster

Bibliography. D. A. Bell, *Electric Circuits*, 6th ed., 1998; L. Chua, C. A. Desoer, and E. Kuh, *Linear and Nonlinear Circuits*, 1987; J. W. Nilsson and S. A. Riedel, *Electric Circuits*, 7th ed., 2004; L. M. Thompson, *Electrical Measurements and Calibration*, 2d ed., 1994.

Voltage amplifier

An electronic circuit whose function is to accept an input voltage and produce a magnified, accurate replica of this voltage as an output voltage. The voltage gain of the amplifier is the amplitude ratio of the output voltage to the input voltage. Often, electronic amplifiers designed to operate in different environments are categorized by criteria other than their voltage gain, even though they are voltage amplifiers in fact. Many specialized circuits are designed to provide voltage amplification. *See* AUDIO AMPLIFIER; CASCODE AMPLIFIER; VIDEO AMPLIFIER.

Voltage amplifiers are distinguished from other categories of amplifiers whose ability to amplify voltages, or lack thereof, is of secondary importance. Amplifiers in other categories usually are designed to deliver power gain (power amplifiers, including push-pull amplifiers) or to isolate one part of a circuit from another (buffers and emitter followers). Power amplifiers may or may not have voltage gain, while buffers and emitter followers generally produce power gain without a corresponding voltage gain. *See* BUFFERS (ELECTRONICS); EMITTER FOLLOWER; POWER AMPLIFIER; PUSH-PULL AMPLIFIER.

Ideal amplifier. A simplified equivalent circuit (**Fig. 1**) can be used to represent a physical voltage source an amplifier, and a resistive load. The voltage source model contains an ideal voltage source and a source resistance. The source resistance, R_S, characterizes the inability of a physical voltage source to maintain its voltage at full value when a load is connected across its terminals, while the load resistance, R_L, represents the physical fact that a load must extract energy from the source.

If the load resistance is connected directly to the source, Ohm's law dictates that the voltage across the load is always less than the ideal source voltage. In fact, the ratio of the voltage to the ideal source voltage equals the ratio of the load resistance to the sum of the load and source resistances. Thus, the load voltage is much less than the ideal source voltage if

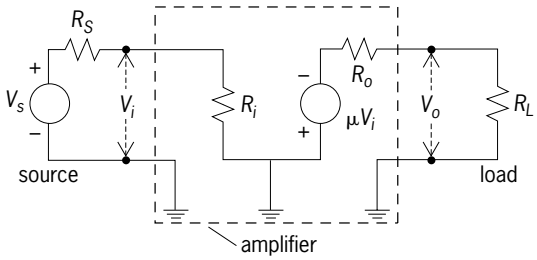

Fig. 1. Voltage amplifier circuit model (with voltage-controlled voltage source). V_s = ideal source voltage, R_S = source resistance, V_i = input voltage, R_i = input resistance, μV_i = voltage-controlled voltage source, R_o = output resistance, V_o = output voltage, R_L = load resistance.

the source resistance is much greater than the load resistance.

In the circuit model under consideration (Fig. 1), the amplifier is characterized as a black box with an input resistance, R_i; a voltage V_i across the input terminals; a voltage controlled voltage source, μV_i; and an output resistance, R_o. The input resistance of the amplifier acts as a load on the voltage source, and the output of the amplifier creates a new ideal voltage source, μV_i, connected to the load through an output resistance of R_o ohms. To correspond to the convention applied to transistor amplifier modeling, the polarity of this controlled, voltage source is assumed negative with respect to ground. Again, equations can be derived from Ohm's law relating the input voltage to the ideal source voltage, and relating the load voltage to the voltage-controlled voltage source. By combining these equations, the voltage gain of the amplifier, that is, the ratio of the load voltage to the ideal source voltage, is found to be given by Eq. (1).

$$G = -\frac{R_i}{R_i + R_s}\frac{R_L}{R_L + R_o}\mu \qquad (1)$$

If the input resistance is much greater than the source resistance and the output resistance is much less than the load resistance, the magnitude of the voltage gain of the amplifier is approximately μ, where μ typically ranges from one to several million depending on the application and on the internal design of the amplifier. In some cases, the open-circuit voltage of the source is adequately high, and it is necessary only for the amplifier to offset the effect of the load resistance on the source by interposing a unity-gain ($\mu = 1$) voltage amplifier with high input resistance (much greater than the source resistance) and low output resistance (much less than the load resistance) between the source and the load. A voltage amplifier with these characteristics is known as a buffer amplifier.

Alternative amplifier models. Other equivalent circuits are often used to represent voltage amplifiers (**Fig. 2**). For situations where the amplifier input and output resistances are relatively high, the ideal voltage-controlled voltage source of the circuit model discussed above (Fig. 1) is replaced by a voltage-controlled current source, $g_m V_i$, shunted by

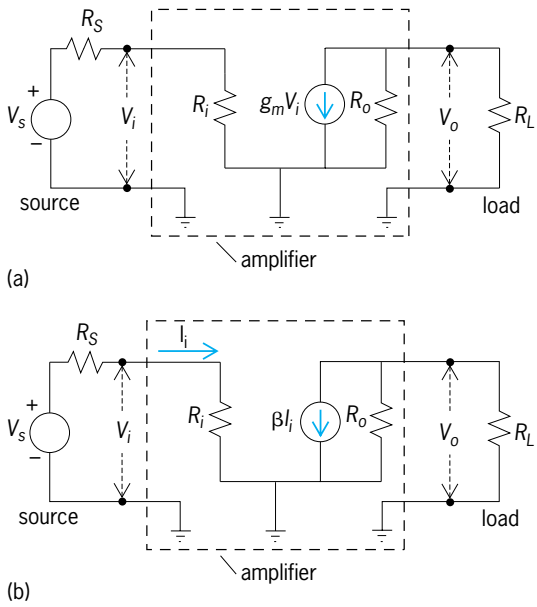

Fig. 2. Alternative voltage amplifier circuit models. (*a*) Model with voltage-controlled current source, g_mV_i. (*b*) Model with current-controlled current source, βI_i, where I_i = input current. Other symbols are as in Fig. 1.

the output resistance, R_o (Fig. 2*a*). Exact equivalence between the two circuits is obtained if g_m equals μ divided by the output resistance, and the amplifier voltage gain (for input resistance much greater than source resistance and output resistance much greater than load resistance) can be approximated as the negative of the product of g_m and the load resistance.

A third equivalent circuit model (Fig. 2*b*) for the amplifier is used primarily where the amplifier input resistance is relatively low and its output resistance is high. This circuit uses a current-controlled current source, βI_i, in its output, where I_i is the input current flowing into the amplifier. The model is equivalent to the voltage-controlled current-source model (Fig. 2*a*) if β equals the product of g_m and the input resistance. Here the amplifier gain (for source resistance much greater than input resistance and output resistance much greater than load resistance) can be approximated as the negative of beta times the load resistance divided by the source resistance.

The two current-source models (Fig. 2*a* and *b*) closely resemble the small-signal equivalent circuits used to model single-transistor voltage amplifier circuits: the voltage-controlled current-source model (Fig. 2*a*) usually is used to model a junction field-effect transistor (JFET) amplifier, while the current-controlled current-source model (Fig. 2*b*) is used for the bipolar junction transistor (BJT) amplifier (**Fig. 3**). *See* TRANSISTOR.

Transistor biasing. Neither the field-effect transistor nor the bipolar junction transistor amplifier will operate properly without proper gate (JFET) or base (BJT) bias voltages applied in series with the signal voltage. These bias circuits can be modeled as ideal voltage sources (Fig. 3). The bias and signal voltages are chosen so that the total input voltage—bias plus signal—will not cut off or saturate the amplifier for any value in the range of the input signal voltage. Cutoff will occur, for example, if the base-to-emitter voltage in a bipolar junction transistor drops below about 0.6 V. At cutoff, the transistor output current drops to zero and remains there until the base-to-emitter voltage again increases above the cutoff value. Similarly, if the bipolar junction transistor base-to-emitter voltage is too high, the transistor will saturate and its output voltage will remain steady at a voltage close to zero until the input voltage drops below the saturation level.

In addition to a bias voltage source, well-designed bipolar transistor amplifiers (Fig. 3*b*) require negative feedback at dc to protect the transistor from thermal runaway. The transistor dissipates power in operation, and this power dissipation manifests itself as an increase in the transistor temperature. The transistor collector current is a sensitive temperature-dependent function of its base-to-emitter bias voltage and will rapidly increase with increasing temperature; it takes less bias voltage to obtain the desired collector current as the temperature of the transistor increases. The increasing current further increases the transistor temperature, and this positive feedback process continues until the transistor saturates. Placing a resistor in the emitter circuit causes the base-to-emitter voltage to decrease as the collector current increases, limiting the increase in collector current with increasing temperature. With sufficient negative feedback, this variation of collector current with temperature is controlled easily: a capacitor in the emitter circuit acts as a short circuit at the amplifier operating frequencies, preventing the negative

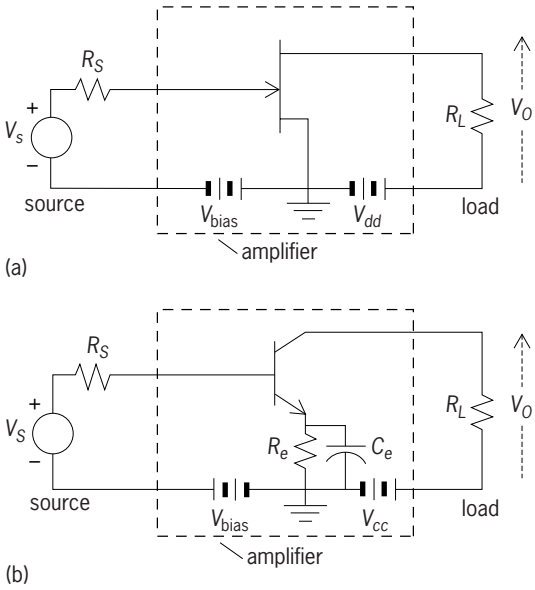

Fig. 3. Basic single-transistor voltage amplifier circuits. (*a*) An *n*-channel junction field-effect transistor common-source amplifier. V_{bias} = bias voltage, V_{dd} = drain supply voltage. (*b*) Bipolar transistor common-emitter amplifier. R_e = emitter resistor, C_e = emitter capacitor, V_{cc} = collector supply voltage. Other symbols are as in Fig. 1.

feedback resistor from reducing the amplifier gain. *See* BIAS (ELECTRONICS).

Interstage coupling. To obtain high gain, cascades of single amplifier circuits are used, usually with a coupling network, actually a simple filter, inserted between the stages of amplification. One such filter is a high-pass network formed by a coupling capacitor, the output resistances of the driving stage, and the input resistance of the driven stage. Since dc voltages are blocked by the capacitor, this ac coupling permits independently setting dc bias voltages for each amplifier stage in the cascade. The coupling network also rejects signals with ac frequency components below a cutoff. The capacitor must be sufficiently large not to attenuate any of the frequencies that are to be amplified. If dc is to be amplified, a direct-coupled amplifier is required, and the design is somewhat more complicated since dc bias voltages on each transistor now cannot be set independently. *See* DIRECT-COUPLED AMPLIFIER; ELECTRIC FILTER.

High-frequency transistor models. The simple models used above for voltage amplifiers and the closely related transistor models are valid only at relatively low operating frequencies. In higher-frequency applications, the parasitic (undesired) capacitance that exists between transistor elements, and sometimes even the small inductance of the transistor leads, must be taken into account. More accurate transistor models incorporate the most important of these parasitic effects.

In most high-frequency amplification applications, gain is desired usually over only a limited range of frequencies, and band-pass filters are used as the interstage coupling networks. These filters are usually realized as inductance-capacitance circuits chosen to resonate, or to impedance-match, the transistor to its load at the center of the frequency band of interest. The transistor parasitic capacitances, supplemented by external capacitors and inductors, are tuned out by the matching network, and parasitic attenuation effects are minimized quite effectively. *See* INTERMEDIATE-FREQUENCY AMPLIFIER; RADIO-FREQUENCY AMPLIFIER.

Balanced and unbalanced amplification. The amplifiers and their models discussed above are called single-ended amplifiers, since their input and output voltages are referred to a common reference point which by convention is called ground. These single-ended circuits, while satisfactory for most noncritical applications, have several weaknesses which degrade their performance in high-gain, weak-signal applications. Their unbalanced construction and their use of a common ground point for return currents makes them susceptible to noise pick-up. Voltage induced in the signal path when extraneous ground currents from other sources flow through the amplifier input signal ground lines is one source of noise. Another is capacitive or inductive coupling between the input signal lines and nearby current-carrying conductors. These coupling mechanisms create noise voltages at the amplifier inputs, which then are amplified along with the desired voltage. To minimize noise on sensitive signal lines, special balanced differential amplifier circuits are often used in critical amplifier applications. Differential amplifiers are designed to have equal impedances to ground for each side of the signal line and to have an output voltage proportional to the difference of the voltages from each signal line to ground. This symmetry cancels common-mode noise voltages, voltages which tend to appear on each of the signal lines as equal voltages to ground. Proper circuit design, with attention to the symmetry of the input circuit construction, can ensure that the majority of undesired noise pick-up will be common-mode noise and, hence, will be attenuated by the differential amplifier. *See* DIFFERENTIAL AMPLIFIER; INSTRUMENTATION AMPLIFIER.

Single-ended amplifiers require multiple resistor networks and interstage coupling capacitors to establish the proper bias conditions for multistage linear amplification. In contrast, differential amplifiers can be biased by using just a few resistors, and their stages also can be direct coupled, obviating the need for coupling capacitors. Since resistors and capacitors are difficult to realize on integrated circuit chips and since they take up much more valuable chip space than do transistors, balanced differential circuits are used almost exclusively in low-frequency and medium-frequency integrated-circuit amplifiers. Discrete transistor single-ended amplifier circuits dominate only at microwave frequencies where it is difficult to control parasitic effects and to achieve high-gain, stable operation with integrated circuits. *See* INTEGRATED CIRCUITS; MICROWAVE.

Differential amplifier. The differential amplifier circuit, as it is usually constructed on a chip (**Fig. 4**), uses transistors configured as current mirrors to set the proper currents through each amplifier transistor and to serve as high-resistance active loads. The pair of transistors that form the first mirror set the total

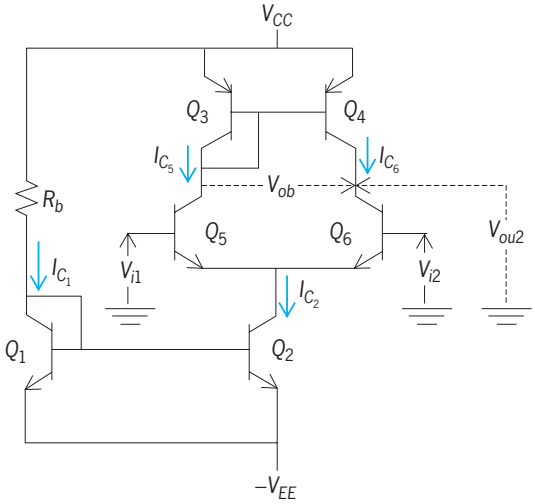

Fig. 4. Bipolar, integrated-circuit differential amplifier. Transistors Q_1, Q_2 form the first current mirror; Q_3, Q_4 form the second current mirror; Q_5, Q_6 are amplifier transistors; V_{i1} and V_{i2} are voltages applied to the input terminals; V_{ou2} is unbalanced output voltage between the collector of Q_6 and ground; V_{ob} is balanced output between collectors Q_5 and Q_6; V_{CC} and V_{EE} are supply voltages; currents $I_{C1} = I_{C2} = I_{C5} = I_{C6} = (V_{CC} + V_{EE} - 0.6)/R_B$.

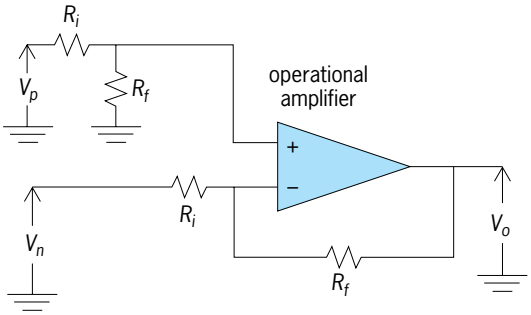

Fig. 5. Operational-amplifier-based difference amplifier, with input voltages V_p and V_n, output voltage V_o, two input resistors R_i, and two other resistors R_f.

collector current through the pair of amplifier transistors. To the signals being amplified, the second current-mirror transistor also appears as a very high resistance in the common emitter lead of the amplifier transistors. *See* CURRENT SOURCES AND MIRRORS.

The second current mirror equalizes the current through the amplifier transistors and also acts as a very high output load resistance on each amplifier transistor. The gain of this single differential stage will be very high if its output voltage is sensed by a circuit with a correspondingly high input resistance.

The differential amplifier can be used either with balanced inputs (with voltages applied to both the input terminals) or with unbalanced ones (where either of these voltages is set to zero by grounding the appropriate terminal). Similarly, the amplifier output voltage will be unbalanced with respect to ground when it is taken between the collector of either of the amplifier transistors and ground, and will be balanced when the output is taken from across the two collectors.

In most applications, needed amplification is incorporated directly into integrated circuits that perform more complex system functions. A single integrated circuit, for example, may contain all of the active (transistor) circuitry needed to build a complete frequency-modulation (FM) radio receiver. Included in the circuit functions contained on the chip are a radio-frequency amplifier, an intermediate-frequency amplifier, and a audio amplifier. The integrated circuit, a handful of discrete components, and a power supply make up the radio. *See* RADIO RECEIVER.

Operational amplifier. In cases where a voltage amplifier is required for some special purpose, operational amplifiers are often used to fill the need. The operational amplifier is an integrated circuit containing a cascade of differential amplifier stages, usually followed by a push-pull amplifier acting as a buffer. The differential voltage gain of the operational amplifier is very high, about 100,000 at low frequencies, while its input impedance is in the megohm range and its output impedance is usually under 100 ohms. The amplifier is designed to be used in a negative-feedback configuration, where the desired gain is controlled by a resistive voltage divider feeding a fraction of the output voltage to the inverting input of the operational amplifier.

A difference amplifier can be created from an operational amplifier and a few resistors (**Fig. 5**). The amplifier both amplifies and subtracts its two input voltages, V_p and V_n. In the case that the two input resistors have the same value, R_i, and the two other resistors in the circuit also have the same value, R_f, Eq. (2) holds for the output voltage, V_o.

$$V_o = \frac{R_f}{R_i}(V_p - V_n) \qquad (2)$$

With needed amplification built into many integrated circuits and with the availability of operational amplifiers for special-purpose amplification needs, there is seldom a need to design and build a voltage amplifier from discrete components. *See* AMPLIFIER; OPERATIONAL AMPLIFIER. Philip V. Lopresti

Bibliography. R. L. Geiger, P. E. Allen, and N. R. Strader, *VLSI Design Techniques for Analog and Digital Circuits*, 1990; P. Gray and R. Meyer, *Analysis and Design of Analog Integrated Circuits*, 4th ed. 2001; P. Horowitz and W. Hill, *The Art of Electronics*, 2d ed., 1989; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., 1997.

Voltage measurement

Determination of the difference in electrostatic potential between two points. The unit of voltage in the International System of Units (SI) is the volt, defined as the potential difference between two points of a conducting wire carrying a constant current of 1 ampere when the power dissipated between these two points is equal to 1 watt.

Voltage standards. These comprise Josephson junction standards, Weston standard cells, and electronic standards.

Josephson junction standards. The discovery of the Josephson effect in superconducting tunnel junctions in 1962 made possible the operation of a quantum standard of voltage. By 1972 this standard was in widespread use in national standards laboratories to supervise their Weston-cell working standards. The resolution and stability achieved by the quantum standard (1 in 10^9) were better by far than the uncertainty to which its relation to the SI unit of voltage was known (1 in 10^5). The need for some international agreement at a higher level of accuracy was resolved by defining the properties of the Josephson junction as a frequency-to-voltage converter. Although it was recognized at the time that the value chosen was somewhat arbitrary, it became possible to have international standards of adequate precision to support trade in precision electrical measuring instruments.

Subsequent work on the realization of the electrical units has enabled the relationship between the SI unit of voltage and the standard based on the Josephson effect to be established with an uncertainty believed to be slightly better than 1 in 10^7. As a result, the International Committee for Weights and Measures (CIPM; the international body concerned with measurement standards) recommended

a new value for the Josephson constant k_{J-90} given by Eq. (1),which would be used throughout the world

$$k_{J-90} = 483{,}597.9 \text{ GHz/V} \tag{1}$$

from January 1, 1990, onward. As a result the values assigned to the working standards of various laboratories were changed on that date, and all subsequent calibrations carried out by them take account of this. The new value was chosen to be consistent with the SI unit within the limits of current knowledge.

A Josephson junction consists of two superconductors, such as lead or niobium, separated by a thin insulating or poorly conducting barrier across which conduction can occur by the tunneling process. The tunnel barrier may be formed by controlled oxidation or evaporation of a material such as silicon oxide. The dimensions of the junction have to be considerably less than 0.04 in. (1 mm). When such a junction is irradiated by radio-frequency energy, usually in the frequency range 10–100 GHz, its voltage–current characteristic becomes broken into a stepped form. There is considerable hysteresis and overlapping of the steps (**Fig. 1**). The height V of each voltage step is given by Eq. (2), where f is the

$$V = \frac{h}{2e} f \tag{2}$$

irradiating frequency, h is Planck's constant, and e is the electron charge. This voltage can be measured across a single junction. Since the junction must be immersed in a helium cryostat at a temperature of a few kelvins in order to produce the superconducting effect, it is convenient to include a large part of the measuring system (**Fig. 2**) in the cryostat. The ratio between the standard cell and the Josephson step voltages can be measured in terms of the ratio of two resistors in the system (Fig. 2), a ratio which has to be established by a separate experiment. Another Josephson junction can be used as a null detector, and hence the step voltage of a few millivolts can be compared with that of a room-temperature standard cell to within less than 1 part in 10^8. *See* SQUID.

By using photolithographic techniques, it has proved possible to construct arrays of Josephson

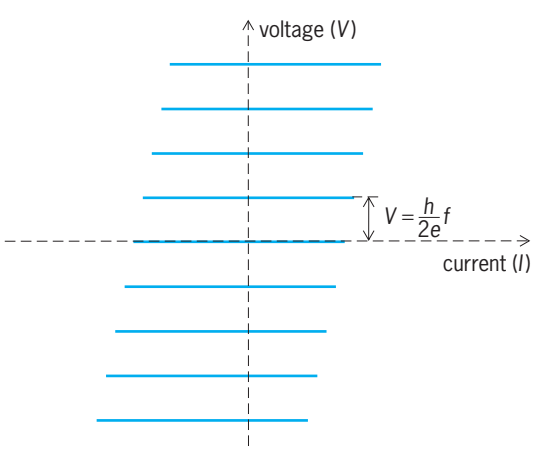

Fig. 1. Voltage–current characteristic of ideal Josephson junction.

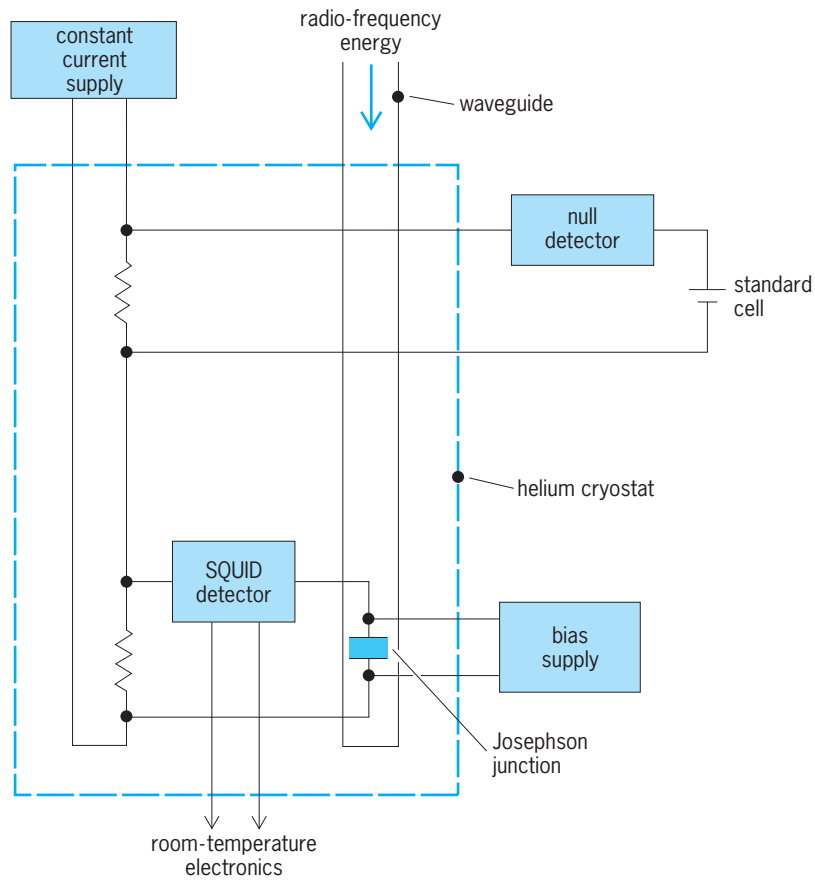

Fig. 2. Simpified Josephson junction measuring system.

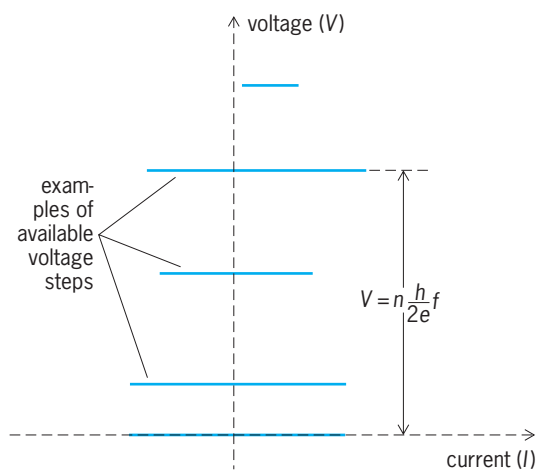

Fig. 3. Voltage–current characteristic of Josephson junction array.

junctions that consist of thousands of separate devices that can function simultaneously. In such an array each device can operate at a relatively high voltage step that extends across the zero-current axis (**Fig. 3**), and so the array may operate with a voltage of over 10 V across it. By using such a system, measurements can be made directly against other reference devices with a precision limited only by their stability and noise properties. *See* JOSEPHSON EFFECT; SUPERCONDUCTING DEVICES.

Standard cells. The saturated Weston cell, which consists of electrodes of mercury and mercury-cadmium amalgam in an electrolyte of saturated cadmium sulfate solution, has been in use since the latter part of the nineteenth century as a portable room-temperature voltage standard. *See* ELECTROMOTIVE FORCE (CELLS).

Although many designs of digital voltmeter are unsuitable for use with Weston cells as a result of excessive input currents, it is possible to choose a suitable model. By using one of these or a relatively simple potentiometer, standard-cell voltages can be compared to better than 1 part in 10^7. By comparing the voltage of each cell against the mean of a group it is possible to detect and eliminate less stable cells, and thus to maintain a voltage standard of high stability. Saturated standard cells are very sensitive to environmental conditions and to the passage of any current through them, and are therefore rather unsatisfactory as reference standards and even more so as transfer standards. Nevertheless, the stability of good cells may be better than 100 nanovolts per year. *See* POTENTIOMETER.

Electronic voltage standards. Diodes used for voltage regulation (Zener diodes) pass a current that increases rapidly when the voltage rises above some critical value, and hence tend to act as constant voltage devices. If such diodes are fed from a well-stabilized source, for example, if they are used in an operational amplifier configuration, they can provide voltage references that may have a stability as good as 1 in 10^6. The temperature coefficient of simple Zener diodes is a function of voltage and operating current. This coefficient can be greatly reduced by the use of compensated diodes, in which one or more forward-biased junctions (having a negative temperature coefficient of voltage) are connected in series with a Zener breakdown junction chosen to have a positive coefficient. A further possibility is the combination of transistor and Zener junctions to provide a compensated reference element. Reference-quality diodes can also be constructed as part of an integrated circuit that includes built-in temperature stabilization (**Fig. 4**). External operational amplifiers and other components are used to complete the voltage and temperature regulation circuits. Like Weston cells, the best diodes are obtained by selection. Since they are less sensitive to environmental conditions and overloads than standard cells, they are much more suitable for use as transfer standards and are also in almost universal use as voltage references in electronic measuring instruments. *See* INTEGRATED CIRCUITS; OPERATIONAL AMPLIFIER; ZENER DIODE.

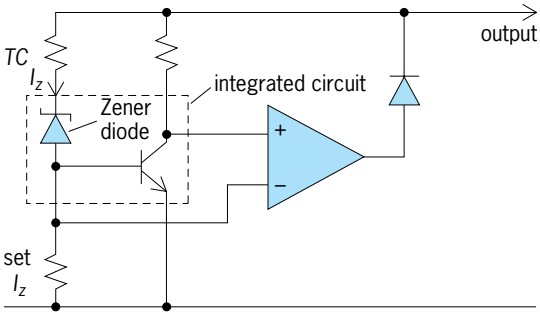

Fig. 4. Reference circuit using an integrated reference amplifier. The resistor "set I_z" allows current I_z flowing through the Zener diode to be adjusted independently of the transistor current. The resistor *TC* allows the temperature coefficient of the output voltage to be adjusted.

Realization of the unit of voltage. Reference standards of voltage are obtained through the conventional value of k_{J-90} recommended by the CIPM, the Josephson effect, and a measurement of frequency. Accurate realization of the SI unit involves exceedingly difficult measurements, for example, combining a realization of the ampere by using a current balance with the unit of resistance obtained through a calculable capacitor. For a full discussion of the unit and its relationship to other electrical units. *See* ELECTRICAL UNITS AND STANDARDS.

Direct-current voltage measurement. The chief types of instruments for measuring direct-current (constant) voltage are potentiometers, resistive voltage dividers, pointer instruments, and electronic voltmeters.

Potentiometer measurements. The most fundamental dc voltage measurements from 0 to a little over 10 V can now be made by direct comparison against Josephson systems. At a slightly lower accuracy level and in the range 0 to 2 V, precision potentiometers are used, in conjunction with very low-noise electronic amplifiers or photocoupled galvanometer detectors. Potentiometers are capable of self-calibration, since only linearity is important, and can give accurate measurements down to a few nanovolts. When electronic amplifiers are used, it may often be more convenient to measure small residual unbalance voltages, rather than to seek an exact balance. *See* AMPLIFIER; GALVANOMETER.

In all dc low-voltage measurements, care is necessary to avoid thermal emf's. All wiring must be of pure copper, which may be plated with gold but not tin or nickel, and any movement of conductors must be avoided. It is necessary to wait several seconds after any switch operation to allow thermal emf's that are generated to die away.

Resistive voltage dividers. Voltage measurements of voltages above 2 V are made by using resistive dividers. These are tapped chains of wire-wound resistors, often immersed in oil, which can be self-calibrated for linearity by using a buildup method. Instruments for use up to 1 kV, with tappings typically in a binary or binary-coded decimal series from 1 V, are known as volt ratio boxes, and normally provide uncertainties down to a few parts per million. Another configuration allows the equalization of a string of resistors, all operating at their appropriate power level, by means of an internal bridge. The use of series-parallel arrangements can provide certain easily adjusted ratios (**Fig. 5**).

Higher voltages can be measured by extending such chains, but as the voltage increases above about 15 kV, increasing attention must be paid to avoid any sharp edges or corners, which could give rise

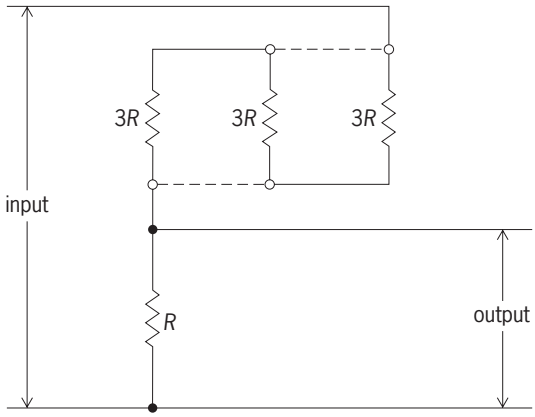

Fig. 5. Achievement of a 10:1 resistance ratio by a series-parallel connection. The single resistor *R* is adjusted to be equal in value to the three resistors of value 3*R* connected in parallel. When all the resistors are reconnected in series, a ratio of exactly 10 is produced.

to corona discharges or breakdown. High-voltage dividers for use up to 100 kV with an uncertainty of about 1 in 10^5, and to 1 MV with an uncertainty of about 1 in 10^4, have been made. *See* CORONA DISCHARGE; ELECTRICAL BREAKDOWN.

DC pointer instruments. For most of the twentieth century the principal dc indicating voltmeters have been moving-coil milliammeters, usually giving full-scale deflection with a current between 20 microamperes and 1 milliampere and provided with a suitable series resistor. Many of these will certainly continue to be used for many years, giving an uncertainty of about 1% of full-scale deflection.

Electronic voltmeters. Since about 1950 it has become increasingly common to use electronic amplifiers to drive pointer instruments in order to reduce the loading of the measured signal, reduce the possibility of mechanical damage to the meter through overload, and provide greatly increased sensitivity. It is possible to obtain instruments having a full-scale deflection of less than 1 nanovolt, and such instruments are now widely used in laboratory work as null detectors.

A further development, the replacement of the analog indicator by a digital display, produced the digital voltmeter (DVM). This type has now become the principal means used for voltage measurement at all levels of accuracy, even beyond one part in 10^7, and at all voltages up to 1 kV. Essentially, digital voltmeters consist of a power supply, which may be fed by either mains or batteries; a voltage reference, usually provided by a Zener diode; an analog-to-digital converter; and a digital display system. This design provides measurement over a basic range from zero to a few volts, or up to 20 V. Additional lower ranges may be provided by amplifiers, and higher ranges by resistive attenuators. The accuracy on the basic range is limited to that of the analog-to-digital converter. *See* ANALOG-TO-DIGITAL CONVERTER; ELECTRONIC POWER SUPPLY.

Most modern digital voltmeters use an analog-to-digital converter based on a version of the charge balance principle. In such converters the charge accumulated from the input signal during a fixed time by an integrator is balanced by a reference current of opposite polarity. This current is applied for the time necessary to reach charge balance, which is proportional to the input signal. The time is measured by counting clock pulses, suitably scaled and displayed. Microprocessors are used extensively in these instruments. *See* MICROPROCESSOR.

Alternating-current voltage measurements. The mean voltage of any number of complete cycles of an ac waveform is always zero. Any departure from zero indicates some dc content. By its nature, an ac signal is always changing. It is therefore necessary to define what is being measured in alternating voltage measurements. The value of greatest interest is the root-mean-square (rms) value, since this is a direct measure of the energy content. The root-mean-square value V_{rms} is defined in terms of the instantaneous value V by Eq. (3). The rectified mean

$$V_{\text{rms}} = \sqrt{\frac{1}{T}\int_0^T V^2 \quad \underset{\lim T\to\infty}{dt}} \tag{3}$$

value is also important since dc instruments fitted with rectifiers usually respond to the mean of the modulus of the signal, often (incorrectly) referred to as the mean. One definition of V_{mean} is given by Eq. (4).

$$V_{\text{mean}} = \frac{1}{T}\int_0^T |V| \quad \underset{\lim T\to\infty}{dt} \tag{4}$$

In any waveform there is a fixed relationship between the peak, root-mean-square, and rectified mean values. For a sinusoidal waveform, the root-mean-square and peak values are related by Eq. (5), and the root-mean-square and rectified mean by Eq. (6).

$$V_{\text{rms}} = \frac{1}{\sqrt{2}} V_{\text{peak}} \tag{5}$$

$$V_{\text{rms}} = \frac{\pi}{2\sqrt{2}} V_{\text{mean}} \tag{6}$$

It has become conventional for instrument scales to be marked with the root-mean-square value that a sine wave would have, if it gave the same deflection as the actual voltage applied. Mean sensing instruments thus have a built-in scaling factor of $1.1107198 = \pi/(2\sqrt{2})$. An alternating voltage is rarely purely sinusoidal, and departure from a sine form is usually specified in terms of harmonic content. *See* NONSINUSOIDAL WAVEFORM; WAVEFORM; WAVEFORM DETERMINATION.

Thermal transfer. Since the working standards of voltage are of the direct-current type, all ac measurements have to be referred to dc through transfer devices or conversion systems. Any instrument that responds to both ac and dc voltage can be used as a transfer device, but the transfer errors may need to be determined. The most accurate transfer device is the multijunction thermal converter (MJTC). This has a calculated ac-dc difference well below 1 part per million over its useful frequency range of about

30 Hz to 10 kHz, and the good agreement between different devices of this type supports this value. These devices are used to select and calibrate single-junction converters, which are more readily available, have only slightly inferior performance, and are useful from 10 Hz to hundreds of megahertz.

Single-junction thermal converters (SJTCs) consist of a thin wire, usually of nickel-chrome alloy, which is heated by the incoming signal to produce a rise in temperature of about 250°F (140°C). The temperature rise is detected by a single thermo- couple junction fixed in an insulating bead near the center of the wire. The whole device is enclosed in an evacuated glass envelope. The heating current is usually 5 or 10 mA. The device is calibrated by comparison with a similar device whose performance has already been established or a multijunction converter, by the sequential application of ac, dc, reversed dc, and ac signals. The pairs of results for ac and dc are averaged to eliminate the effect of drift and current reversal. More complex sequences can be used to obtain further improvements. By this means the ac-dc difference of the best devices can be certified to a level of about 5 parts per million (ppm) at low frequencies, degrading with frequency to about 15 ppm at 1 MHz.

In commercial multirange instruments, similar single-junction thermal converters are combined with switched series resistors of special design to give voltage ranges from 1 V to about 1 kV. The measurement at the highest level of accuracy of an alternating voltage is thus made by establishing a dc voltage equivalent to the root-mean-square ac voltage through the use of such an instrument. Thermal converters are also made in which other methods of temperature sensing are used, for example, the effect on the base-to-emitter voltage of a planar transistor. *See* THERMAL CONVERTERS; THERMOCOUPLE; TRANSISTOR.

AC conversion. A variety of techniques can be used to convert an ac signal into a dc equivalent automatically. All multimeters and most ac meters make use of ac-dc conversion to provide ac ranges. These are usually based on electronic circuits. Rectifiers provide the most simple example. *See* MULTIMETER.

In a commonly used system (**Fig. 6**), the signal to be measured is applied, through a relay contact, to a thermal converter. In order to improve sensitivity, a modified single-junction thermal converter may be used in which there are two or three elements in a single package, each with its own thermocouple. The output of the thermal converter is measured by a very sensitive, high-resolution analog-to-digital converter, and the digital value memorized. When a measurement is required, the relay is operated, and the thermal converter receives its input, through a different relay contact, from a dc power supply, the amplitude of which is controlled by a digital and analog feedback loop in order to bring the analog-to-digital converter output back to the memorized level. The dc signal is a converted value of the ac input and can be measured. Modern versions of this type of instrument make use of microprocessors to control the conversion process, enhance the speed of operation, and include corrections for some of the errors in the device and range-setting components.

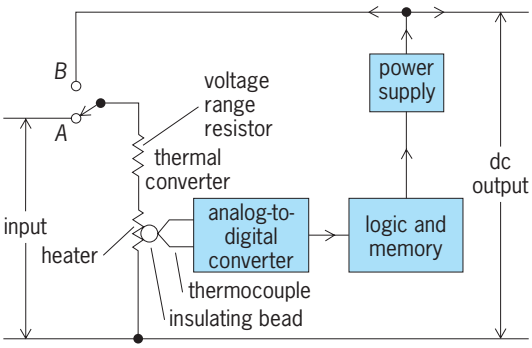

Fig. 6. Thermal ac-dc converter system. The thermal converter receives the signal to be measured through relay contact *A*, and the input from the dc power supply through relay contact *B*.

Voltage ratio devices. Two distinct forms of transformer are used as ratio standards at power and audio frequencies, namely, voltage transformers and inductive voltage dividers.

Voltage transformers have separate primary and secondary windings and operate with the secondary circuit very lightly loaded. High-permeability cores are operated at a low flux density to give good linearity. The primary/secondary voltage ratio then closely approaches the turns ratio, and phase errors are very small. Ratio and phase errors can be measured by bridge techniques using inductive or capactive dividers, and once measured are normally stable through the life of the transformer. A good reference voltage transformer may have errors less than 50 ppm. Voltage transformers are used chiefly at power frequencies and voltages up to 500 kV for calibration purposes. *See* INSTRUMENT TRANSFORMER; TRANSFORMER.

Inductive voltage dividers are basically autotransformers, usually toroidal in form, in which a tapped primary winding is wound on a core of very high permeability, usually of supermumetal. To achieve exact equality of division between windings, the winding consists of a rope of 10 (for a decade divider) similar strands of insulated wire, which are then connected end to end to give the successive sections. This construction provides an equality of division better than 1 ppm. In a two-stage divider the toroidal magnetic core is subdivided, and a second winding arranged to thread only part of it. This second winding is arranged to carry a current that generates sufficient flux in the one core to provide almost all the magnetization for the whole transformer, usually by having the same number of turns as the main winding and being connected to the same source (**Fig. 7**). The equality of division is improved to better than 1 in 10^8 for a single decade of a two-stage divider. Such dividers can be connected in a decade arrangement to give variable ratios with the same resolution, and are used as a building block in precision ac measurements. *See* AUTOTRANSFORMER; INDUCTIVE VOLTAGE DIVIDER.

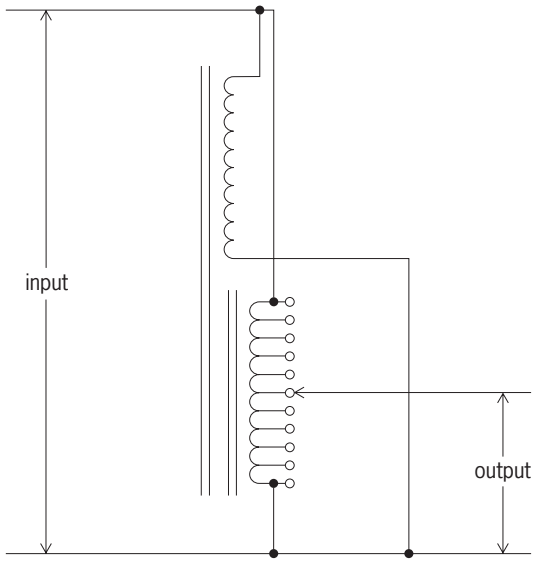

Fig. 7. Two-stage inductive voltage divider transformer.

Digital voltmeters. As in the dc case, digital voltmeters are now probably the instruments in widest use for ac voltage measurement. The simplest use diode rectification of the ac to provide a dc signal, which is then amplified and displayed as in dc instruments. This provides a signal proportional to the rectified mean. For most purposes an arithmetic adjustment is made, and the root-mean-square value of a sinusoidal voltage that would give the same signal is displayed. Several application-specific analog integrated circuits have been developed for use in instruments that are required to respond to the root-mean-square value of the ac input. More refined circuits, based on the logarithmic properties of transistors or the Gilbert analog multiplier circuit, have been developed for use in precision instruments. The best design, in which changes in the gain of the conversion circuit are automatically compensated, achieves errors less than 10 ppm at low and audio frequencies.

Sampling digital voltmeters are also used, in which the applied voltage is switched for a time very short compared with the period of the signal into a sample-and-hold circuit, of which the essential element is a small capacitor. The voltage retained can then be digitized without any need for haste. At low frequencies this approach offers high accuracy and great versatility, since the voltages can be processed or analyzed as desired. At higher frequencies, for example, in the microwave region, it also makes possible the presentation and processing of fast voltage waveforms using conventional circuits. *See* OSCILLOSCOPE.

AC pointer instruments. The analog methods mentioned above produce dc output voltages that are proportional to the ac values, and hence at lower levels of accuracy can be used to drive dc pointer instruments. However, most pointer ac instruments are either of the rectifier type, sensitive to the rectified mean value and calibrated to indicate the root-mean-square value of an equivalent sinusoidal voltage, or of the moving-iron type, which provides a square-law scale. Electrostatic voltmeters, which also produce a square-law scale, are sometimes used for voltages in the range 1–10 kV.

Frequencies above 1 MHz. Voltage measurements at radio frequencies are made by the use of rectifier instruments at frequencies up to a few hundred megahertz, single-junction converters at frequencies up to 500 MHz, or matched bolometers or calorimeters. At these higher frequencies the use of a voltage at a point must be linked to information regarding the transmission system in which it is measured, and most instruments effectively measure the power in a matched transmission line, usually of 50 ohms characteristic impedance, and deduce the voltage from it. *See* BOLOMETER; MICROWAVE MEASUREMENTS; MICROWAVE POWER MEASUREMENT; TRANSMISSION LINES.

Pulse measurements. Pulse voltage measurements are made most simply by transferring the pulse waveform to an oscilloscope, the deflection sensitivity of which can be calibrated by using low-frequency sine waves or dc. Digital sampling techniques may also be used. *See* ELECTRICAL MEASUREMENTS. R. B. D. Knight

Bibliography. C. H. Dix and A. E. Bailey, Electrical standards of measurement, *Proc. IEE*, vol. 122, no. 10R, October 1975; S. Geczy, *Basic Electrical Measurements*, 1984; F. K. Harris, *Electrical Measurements*, 1952, reprint 1975; H. W. Hellwig (ed.), Conference of Precision Electromagnetic Measurements, *IEEE Trans. Instrum. Meas.*, vol. 32, no. 1, March 1983; *International System of Units (SI)*, NBS Spec. Publ. 330, 1991; B. Noltingk (ed.), *Instrumentation Reference Book*, 2d ed., 1995; M. V. Reissland, *Electrical Measurements: Fundamentals, Concepts, Applications*, 1989; M. G. Say (ed.), *Electrical Engineer's Reference Book*, 15th ed., 1993.

Voltage-multiplier circuit

A circuit which produces a dc output voltage that is a fixed multiple of the input dc voltage or the peak voltage of an ac input waveform. This fixed multiple is approximately an integer (± 2, ± 3, ± 4, . . .). Voltage multipliers are used to produce a high dc voltage where modest load current is required. The circuit can be implemented with diodes and capacitors or with switched capacitors.

Diode-capacitor voltage multipliers. A diode passes conventional current in only one direction. In the diode-capacitor voltage doubler, this current through one of the diodes charges one of the capacitors to approximately the peak value of the input voltage. The series combination of the input source and the capacitor generates a voltage across the diode that varies from approximately 0 to -2 times the peak input voltage. The diode is reverse biased except for short intervals to replace charge drained from the capacitor to supply the load current or leakage. A second diode and capacitor form another peak detecting circuit with the voltage across the first diode as the input. This results in a negative dc output voltage equal to twice the peak of the input waveform. The polarity of the output can be reversed by simply

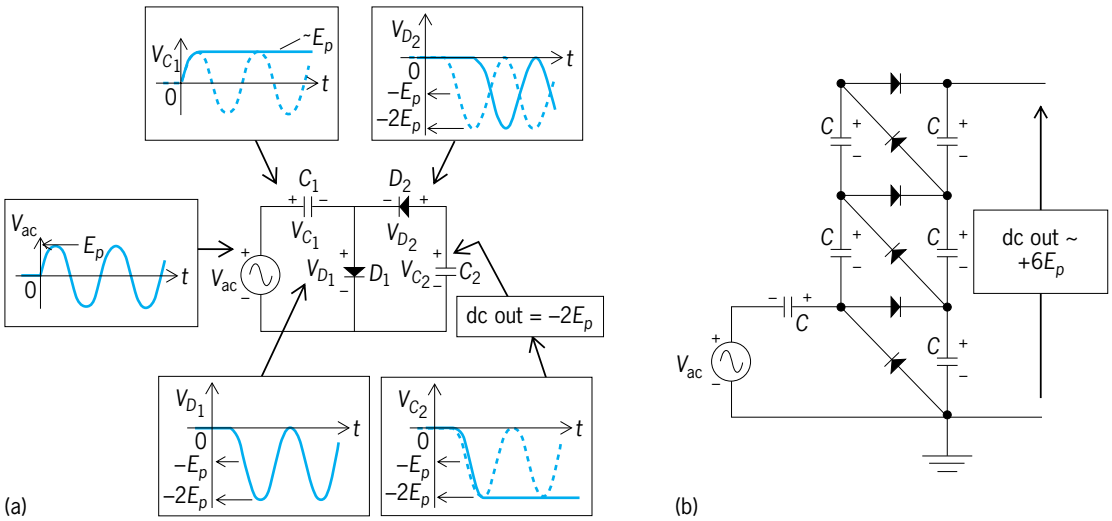

Fig. 1. Diode-capacitor voltage-multiplier circuits. (*a*) Negative voltage doubler. Diodes pass conventional current only in the directions of the arrows. Waveforms of input ac voltage, V_{ac} with peak value E_p, and of voltages at diodes D_1 and D_2 and capacitors C_1 and C_2 are also shown. (*b*) Positive, times-six multiplier with three stages, each with two diodes and two capacitors, *C*.

reversing both diodes. Another version of the voltage doubler produces two outputs, one equal to the peak input voltage and the other to its negative, using the same components. This is realized by connecting a positive and a negative peak detector across the input. *See* CAPACITOR; DIODE.

A positive times-six voltage multiplier circuit (**Fig. 1***b*) operates on the same principle as the voltage doubler circuit. Each stage, consisting of two diodes and two capacitors, adds two to the multiplication factor. The forward-bias voltage drop across each diode that results when charge is forced on the capacitors will cause the output to be less than that for an ideal. Also, any current required by the load will produce an ac ripple in the dc output. If the first capacitance is much greater than the second in the voltage-doubler circuit discussed above (Fig. 1*a*), the peak-to-peak ripple voltage can be approximated by the equation below, where I_{dc} and E_{dc} are the output

$$\text{Peak-to-peak ripple} = \frac{I_{dc}}{fC_2} = \frac{E_{dc}}{fC_2R_{load}}$$

current and voltage respectively, f is the fundamental frequency of the input ac waveform, C_2 is the second capacitance, and R_{load} is the load voltage. *See* RIPPLE VOLTAGE.

Switched-capacitor applications. A typical voltage-multiplier circuit that uses switched capacitors (**Fig. 2***a*) is a dc-to-dc converter that generates +10-V and −10-V dc sources from a single +5-V dc source. A common application for this circuit is to use a +5-V dc power supply to power integrated circuits which require positive and negative voltages. Such switched-capacitor applications are well suited to integrated circuits because the transistor switches and the associated control circuitry take up very little space and the added cost is negligible. The switches are frequently shown as mechanical switches on circuit diagrams (Fig. 2*b*) for simplicity, but they are actually transistors—typically, parallel positive and negative metal-oxide-semiconductor (PMOS and NMOS) field-effect transistors forming a transmission gate switch. The switches are alternately connected to one set of terminals and then the other. Special nonoverlapping clock signals (Fig. 2*b*) must be provided to ensure that both sets of terminals are never on or connected at the same time. In one switch

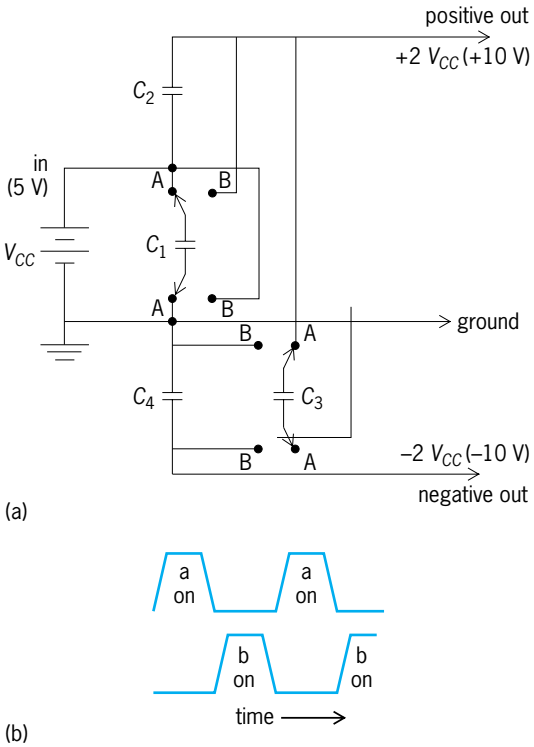

Fig. 2. Dual switched-capacitor dc-to-dc converter that generates +10-V and −10-V dc sources from a single +5-V dc source. (*a*) Simplified circuit diagram. Switches are alternately connected to terminals A and B. C_1, C_2, C_3, and C_4 are capacitors. V_{CC} is the supply voltage. (*b*) Waveforms of nonoverlapping clock signals.

position, one of the capacitors is connected to the input, and it rapidly charges to the supply voltage. In the other switch position, this capacitor is connected to a second capacitor. This is the switched-capacitor equivalent to a low-pass filter, and the second capacitor will charge up to the supply voltage with an equivalent time constant equal to C_2/fC_1, where C_1 and C_2 are the first and second capacitances, and f is the clock frequency. A voltage doubler is realized at the positive output of the circuit. To generate the negative output, a third capacitor is first connected across the positive output and then across a fourth capacitor. *See* ELECTRONIC POWER SUPPLY; INTEGRATED CIRCUITS; SWITCHED CAPACITOR; TRANSISTOR.

Applications using diode multipliers. Voltage-multiplier circuits were historically used to generate the high voltage required for cathode-ray tubes and photomultipliers, whose current requirements are very small. The voltage-doubler circuit discussed above (Fig. 1*a*) can be used as a dc-to-dc converter to generate a negative-bias voltage for an integrated circuit that has only a positive supply voltage. The ac voltage source is replaced with an internal ring or relaxation oscillator that generates a square wave, for example, 0 to 5 V dc. The voltage across the first capacitor is less than the dc supply voltage because of the forward-bias voltage drop across the diode. The ac voltage across the first diode will be a square wave ranging from approximately the negative of the supply voltage to 0. The second diode and capacitor detect the peak negative voltage, which is approximately −3.2 V dc for a +5-V dc supply. This voltage is adequate to provide the necessary negative bias required by some integrated circuits. This application of a voltage doubler circuit is called a substrate pump or a substrate bias generator. *See* CATHODE-RAY TUBE; PHOTOMULTIPLIER.

Norman G. Dillman

Bibliography. S. G. Burns and P. R. Bond, *Principles of Electronic Circuits*, 2d ed., 1996; R. C. Dorf (ed.), *The Electrical Engineering Handbook*, 2d ed., 1997; E. D. Fabricius, *Introduction to VLSI Design*, 1990; C. A. Schuler, *Electronics: Principles and Applications*, 5th ed., 1999.

Voltage regulation

The change in voltage magnitude that occurs when the load (at a specified power factor) is reduced from the rated or nominal value to zero, with no intentional manual readjustment of any voltage control, expressed in percent of nominal full-load voltage. Voltage regulation is a convenient measure of the sensitivity of a device to changes in loading. This concept is commonly used for transmission or distribution circuits as well as power supply equipment, including constant-voltage transformers, rectifiers, outdoor coupling capacitors, direct-current generators, induction frequency converters, synchronous generators, line regulation circuits, and thyristor converters. A common form of expressing voltage regulation in percent is shown in the equation below.

$$\text{Voltage regulation} = \frac{\text{no-load voltage} - \text{full-load voltage}}{\text{full-load voltage}} \times 100$$

As an example of the use of this equation, one can compute the regulation of the supply voltage to an appliance, home, office, or an entire city. This gives a measure of the need for special voltage-regulating equipment, particularly if there are voltage-sensitive components in the load. *See* GENERATOR; TRANSFORMER; VOLTAGE REGULATOR.

Paul M. Anderson

Bibliography. D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; B. R. Gungor, *Power Systems*, 1988; F. Jay (ed.), *IEEE Standard Dictionary of Electrical and Electronic Terms*, IEEE Std100-1988, 4th ed., 1988.

Voltage regulator

A device or circuit that maintains a load voltage nearly constant over a range of variations of input voltage and load current. Voltage regulators are used wherever the unregulated voltage would vary more than can be tolerated by the electrical equipment using that voltage. Alternating-current distribution feeders use regulators to keep the voltage supplied to the user within a prescribed range. Electronic equipment often has voltage regulators in dc power supplies.

Electronic Voltage Regulators

A dc power supply is an essential component in any electronic system. **Figure 1** shows a Thévenin equivalent circuit of a power supply voltage source having an open-circuit (no-load) voltage V_0 and output resistance R_0. When a load with resistance R_L is connected to the output of this supply, a load current $I_L = V_L/R_L$ flows through R_0, resulting in a drop in load voltage V_L as given by Eq. (1).

$$V_L = V_0 - I_L R_0 \tag{1}$$

See THÉVENIN'S THEOREM (ELECTRIC NETWORKS).

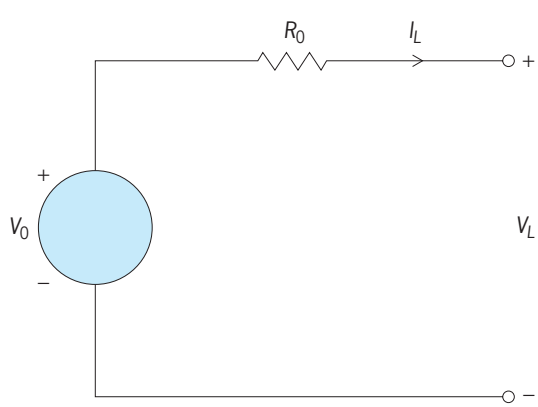

Fig. 1. Equivalent circuit of a basic dc power supply.

A basic power supply as modeled above exhibits two undersirable characteristics. The first is that the load voltage V_L decreases with increasing load current I_L. This effect can be severe in supplies with large effective output resistances. Such supplies are said to have poor load regulation. The second problem is that V_L depends directly on the source V_0. In practice, V_0 might be derived from a relatively inaccurate source such as the ac line voltage (after suitable rectification and filtering) or a battery. Fluctuations in V_0 reflect directly onto the voltage experienced by the load. In this case, the supply is said to have poor line regulation. An ideal power supply would exhibit perfect load regulation (V_L independent of I_L) and perfect line regulation (V_L independent of V_0) and would provide an output voltage of the form of Eq. (2), where V_{ref} is a well-defined reference voltage and k is a constant scaling factor. The task of an electronic voltage regulator is to provide an output voltage characteristic that closely approximates the ideal of Eq. (2), given an unregulated supply voltage as an input.

$$V_L = kV_{ref} \tag{2}$$

Feedback network. Typically, electronic voltage regulators employ a feedback network of the form shown in **Fig. 2**. Here, a high-gain amplifier compares a fraction of the load voltage V_L/k with a constant reference V_{ref}. Any difference between these two voltages is amplified and used to control a series pass device in a manner whereby this difference is minimized. For an ideal amplifier with zero offset and infinite voltage gain, the difference is reduced to zero and the ideal relationship of Eq. (2) is realized. *See* FEEDBACK CIRCUIT.

Monolithic voltage regulator chips. The wide range of applications for electronic voltage regulators has led to the development of these circuits in fully monolothic integrated circuit technology, where all or most of the required circuit components are realized on a single chip of silicon. Offering various output current and voltage ratings, and output voltages of either positive or negative polarity, several commercial regulator integrated circuits are now available to suit the requirements of most applications. The designs of these regulators have matured and have become rather sophisticated. In addition to implementation of the high-gain feedback amplifier, the series pass element, and an accurate voltage reference, all on a single silicon die, built-in protection against overload conditions (such as output short circuits and excessive operating temperature) is now standard. Novel circuit-design, processing, and packaging techniques have been developed and implemented to achieve increased accuracy, temperature stability, efficiency, reliability, and power-handling capability, while reducing package size and cost. *See* INTEGRATED CIRCUITS.

Example. In order to explain how monolithic voltage regulator chips work, a standard commercial integrated circuit is used as an example. A simplified schematic diagram showing the essential details of the regulator is given in **Fig. 3**. Although circuit design details vary considerably between commercially available regulators, depending on the design objectives, specifications, technology used, and manufacturer, the concepts implemented in the circuit of Fig. 3 are typical of the state of the art. This regulator integrated circuit was designed primarily for supplying power to logic circuits.

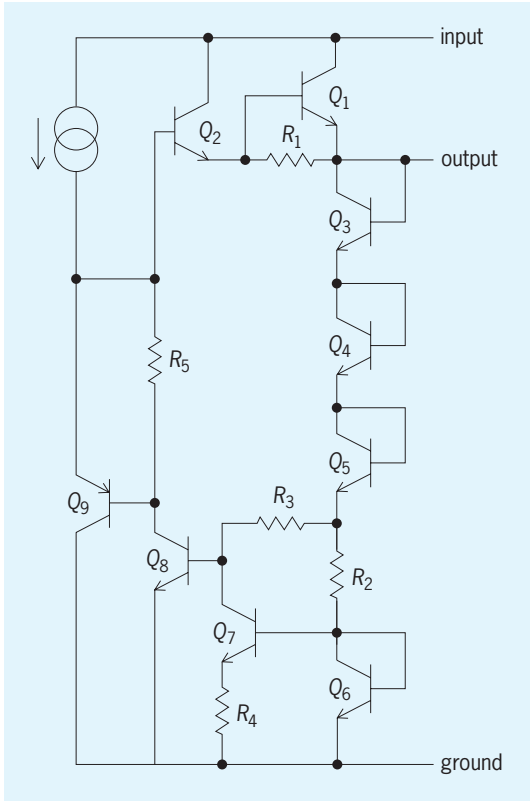

Fig. 3. Simplified circuit schematic of the LM109 voltage regulator. (*After R. J. Widlar, New developments in IC voltage regulators, IEEE J. Sol. State Circ., SC-6:2–7, February 1971*)

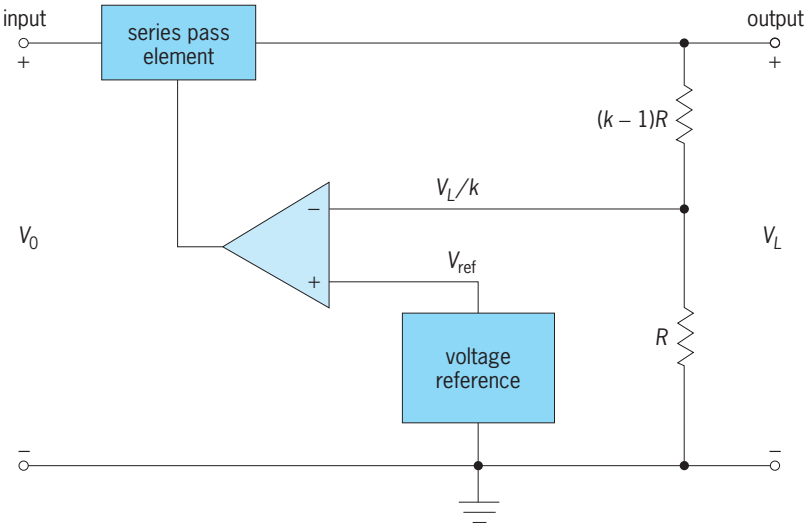

Fig. 2. Voltage regulator feedback circuit.

The circuit makes use of a basic property of bipolar transistors: the base-emitter voltage V_{BE} of a transistor has a negative dependence on temperature (that is, V_{BE} decreases with increasing temperature), while the difference ΔV_{BE} in the base-emitter voltages of two transistors biased at constant currents is directly proportional to temperature. Thus, it is possible to generate a weighted sum of V_{BE} and ΔV_{BE} of the form of Eq. (3), and to choose the weights k_1 and k_2 such

$$V_{\text{ref}} = k_1 V_{BE} + k_2 \Delta V_{BE} \tag{3}$$

that the two temperature variations cancel, resulting in a reference voltage V_{ref} that is insensitive to temperature.

Referring to the circuit shown in Fig. 3, the voltage across resistor R_4 is given by $V_{BE6} - V_{BE7} = \Delta V_{BE}$. Since transistor Q_7 and resistor R_3 conduct approximately the same current as R_4, the voltage across R_3 is $(\Delta V_{BE}/R_4)R_3$. The output voltage is simply the sum of the base-emitter voltages of transistors Q_3–Q_5, the voltage across R_3, and the base-emitter voltage of Q_8. Assuming that Q_3–Q_5 and Q_8 have the same base-emitter voltage V_{BE}, the output voltage can be written as Eq. (4), which has the same form as Eq. (3)

$$V_{\text{ref}} = \frac{R_3}{R_4} \Delta V_{BE} + 4V_{BE} \tag{4}$$

with $k_1 = (R_3/R_4)$ and $k_2 = 4$. Thus, by appropriately choosing the resistance ratio R_3/R_4, a temperature-stable reference is obtained.

Comparing this circuit with the block diagram of Fig. 2, transistors Q_8 and Q_9 perform the amplifying function shown, while Q_1 is the series pass element that supplies the load current. An additional device Q_2 is used to provide the required base-current drive for Q_1. In this circuit, the output load voltage is the samee as the generated reference voltage [that is, $k = 1$ in Eq. (2)].

The complete regulator circuit is based essentially on the simplified schematic of Fig. 3, but includes additional circuitry for improving accuracy and protecting the device under adverse operating conditions. For example, the output current is sensed on-chip and, in case of an accidental short circuit in the load, this current is limited to a safe value. However, even though the output current is limited, excessive power dissipation can cause the chip to overheat. Additional thermal protection circuitry included on-chip further limits the output current when the chip temperature reaches about 175°C (347°F).

Characteristics. This monolithic regulator provides a fixed 5-V output voltage at load currents up to 1 A and is extremely simple to use. The only external components required are two small capacitors (connected from the input and output terminals to ground) for improving high-frequency stability and transient response. The output voltage of the regulator changes by less than 50 mV for input voltages ranging from 7 to 35 V and load currents from 0 to 1 A. The temperature drift in the output voltage is less than 0.02%/°C (0.01%/°F) over the rated ambient operating temperature range (−55°C to 125°C or −67 to 257°F).

Ashok P. Nedungadi

Power-System Voltage Regulators

Voltage regulators are used in the generation, transmission, and distribution of electric power.

Power transmission and distribution. Voltage regulators are used on distribution feeders to maintain voltage constant, irrespective of changes in either load current or supply voltage. Voltage variations must be minimized for the efficient operation of industrial equipment and for the satisfactory functioning of domestic appliances, television in particular. Voltage is controlled at the system generators, but this alone is inadequate because each generator supplies many feeders of diverse impedance and load characteristics. Regulators are applied either in substations to control voltage on a bus or individual feeder or on the line to reregulate the outlying portions of the system. These regulators are variable autotransformers with the primary connected across the line. The secondary, in which an adjustable voltage is induced, is connected in series with the line to boost or buck the voltage. *See* AUTOTRANSFORMER; ELECTRIC DISTRIBUTION SYSTEMS; ELECTRIC POWER SUBSTATION.

A control and drive provide automatic operation. A voltage-regulating relay senses output voltage. When the voltage is either above or below the band of acceptable voltage maintained by the regulator, this relay causes the motor to operate and change the regulator position to raise or lower the voltage as required to bring it back within the control band. It is desirable to maintain constant voltage at the average load center out on the feeder rather than at the regulator terminals. Hence the control circuit includes a line-drop compensator with resistance and reactance elements that can be adjusted to represent line impedance. These impedances carry current proportional to circuit load, thereby simulating the voltage drop between the regulator and the load center, and modify the voltage sensed by the voltage-regulating relay. *See* RELAY.

Two principal types of feeder regulators are used: the step regulator, which provides increments or steps of voltage change, and the induction regulator, which provides continuous voltage adjustment.

Step voltage regulator. The transformation ratio of the autotransformer in this regulator is adjusted by a voltage selector switch which changes the secondary winding tap connected to the line. **Figure 4** shows the most commonly used circuit.

Switching is performed without interrupting load current by means of the two switch fingers in the selector switch. When the switch moves from the full-tap position shown in Fig. 4, one finger contacts the next tap before the other finger leaves the first tap; this constitutes a tap-bridging position. Switching reactors limit the current circulating between the bridged taps. Most regulators are designed to operate continuously in these bridging positions (as well as the full-tap positions) to provide voltages midway between the voltages of adjacent taps. Thus the voltage

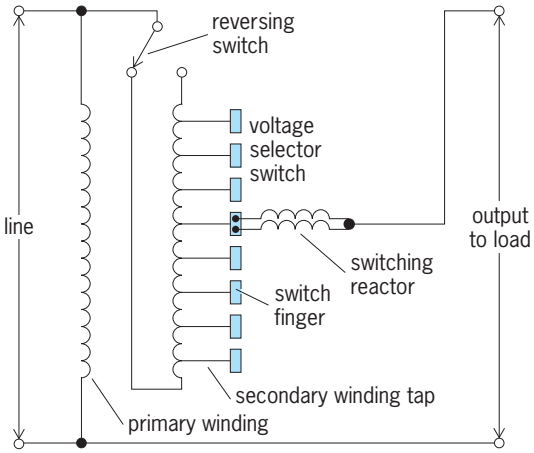

Fig. 4. Step voltage regulator circuit.

step between adjacent positions is half-tap voltage, and the number of winding taps required is half the number of operating positions in the boosting range. The automatic reversing switch changes the polarity of the secondary relative to the primary, thus providing a bucking range equal to the boosting range. Single- and three-phase designs are available with a range of ±10% in 16 or 32 steps. Other design variations may employ single-finger switching and operate continuously only in full-tap positions.

Load tap-changing transformers are often used to provide both regulation and transformation from one voltage level to another. They are similar to step voltage regulators, except that they have separate high- and low-voltage windings.

Induction voltage regulator. This is similar in structure to a wound-rotor induction motor with rotor restrained so that it moves only to adjust voltage. The primary winding on the rotor is magnetically coupled with the series secondary winding on the stator. *See* INDUCTION MOTOR; MOTOR.

The principle of a two-pole single-phase regulator is illustrated in **Fig. 5**. Secondary voltage is continuously adjusted from full buck to full boost by changing the relative angular position of these windings through 180 electrical degrees.

Single-phase regulators require an additional permanently short-circuited rotor winding that is in space quadrature to the primary winding. Without this winding, the reactance of the regulator to the line current flowing in the secondary would be excessive in the neutral region between buck and boost. Three-phase induction regulators, if built on a three-phase core, do not require a short-circuited winding. Such induction regulators inherently introduce phase shift between primary and secondary voltages and are no longer supplied for feeder regulation.

Other regulators. Other types of regulators are also used outside the United States. One construction is a transformer structure with moving coils to change coupling; another has contacts moving over the exposed conductors to provide a large number of small, discrete steps.

Line voltage may be increased by drawing leading-power-factor current through the line reactance. Static capacitors, shunt-connected in fixed or automatically switched banks, are often applied near the loads to raise voltage. The increase in voltage is not limited solely to the vicinity of the capacitor. They also help compensate for the usual system condition

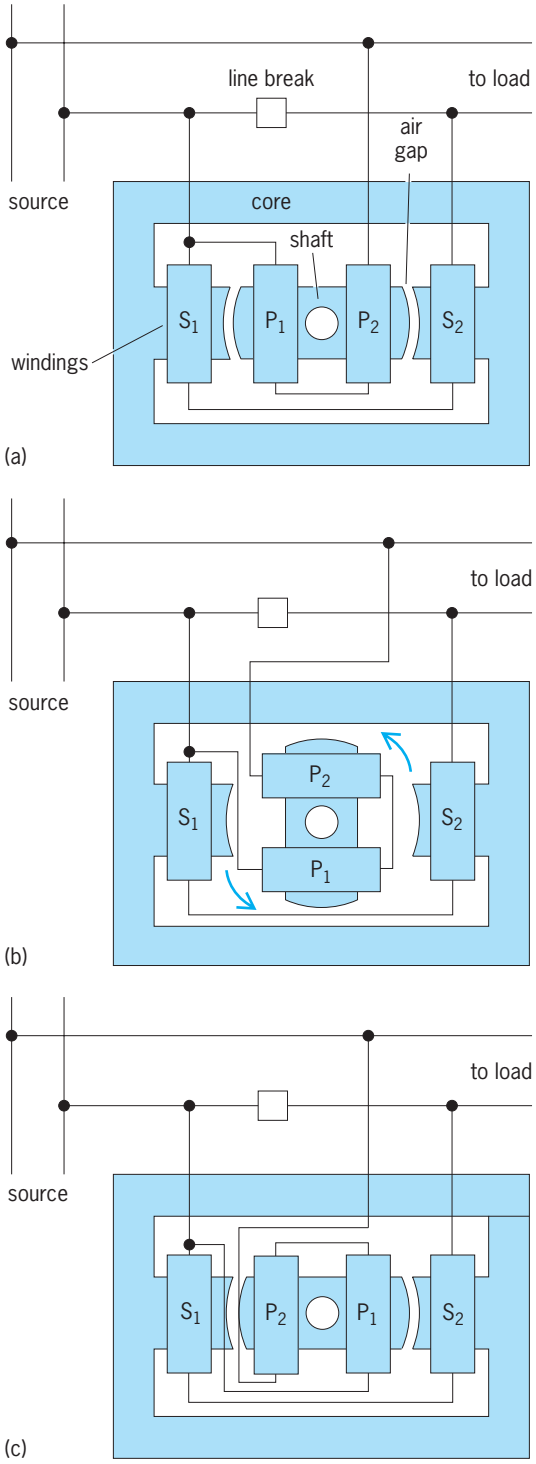

Fig. 5. Induction voltage regulator. (*a*) Full boost position. (*b*) Neutral position. (*c*) Full buck position of rotor. *P* coils are rotor primary coils, *S* coils are stationary secondary coils. (*After B. G. A. Skrotski, ed., Electric Transmission and Distribution, McGraw-Hill, 1954*)

of lagging power factor. The application of series capacitors on fluctuating loads is increasing.

Infrequently, synchronous machines called synchronous condensers, overexcited to draw leading current, are connected to lines. Their use has decreased with the availability of low-cost reliable static capacitors.

Conditions may exist where the inherent static capacity of the circuit is excessive. The line voltage may rise on cable systems or lightly loaded lines; shunt reactors may be used to neutralize the capacity of such systems. *See* STATIC VAR COMPENSATOR.

Donnell D. MacCarthy

Power generation. Voltage regulators are used on rotating machines in power generation applications to automatically control the field excitation so as to maintain a desired machine output voltage. Rotating machines, both small (down to 1 kW) and large (up to 1,000,000 kW), are the predominant means of power generation throughout the world, and voltage regulators of varying design and sophistication are employed on most of them. Even ac generators (or alternators) in automotive applications employ voltage regulators utilizing similar principles. *See* ALTERNATING-CURRENT GENERATOR; DIRECT-CURRENT GENERATOR; ELECTRIC POWER GENERATION; ELECTRIC ROTATING MACHINERY; GENERATOR.

The basic concept of a voltage regulator applied to an ac machine is shown in **Fig. 6**. The voltage regulator senses the output voltage of the ac machine (either directly in a low-voltage machine or through step-down voltage transformers in high-voltage machines) and compares this voltage feedback signal with a voltage reference. This voltage reference can be either fixed or adjustable. The difference between the reference (desired) and feedback (actual) voltages is the voltage error signal. The voltage regulator amplifies this error signal and in more sophisticated systems provides various stabilizing and limiting functions. The resulting voltage regulator output will determine the ac generator field excitation, either directly or through an exciter.

When the voltage error signal is positive, the voltage regulator will act to boost (increase) excitation to the field winding to help maintain the desired output voltage. When the error signal is negative, the voltage regulator will act to buck (decrease) excitation. In most generator applications it is not feasible to operate with fixed, uncontrolled field excitation, because the required level of field excitation to maintain rated output voltage increases as active power [measured in kilowatts (kW)] or reactive power [measured in kilovolt-amperes reactive] (kVAr)] loading is increased. Without a voltage regulator, the generator voltage will typically drop to unacceptable levels as loading is increased, and potentially will cause loss of synchronism with the power system for applications in which the generator is synchronized to a power grid. The voltage regulator will sense the tendency of the generator output voltage to drop with increasing loading, and will appropriately increase field excitation level in order to maintain the desired output voltage automatically.

Exciters. In most utility power generation applications, the ac generator is a synchronous machine, having a rotating dc field winding and a stationary three-phase armature winding. In these cases, the voltage regulator senses the output or terminal voltage of the three-phase armature winding and adjusts the excitation level to the field winding. In large synchronous machine applications, the power requirements of the field winding are significant enough to require a device called an exciter. The exciter can be either a static power amplifier comprising solid-state power conversion devices (such as thyristors or silicon controlled rectifiers) or a rotating machine, which can be either a dc machine (in older applications) or an ac machine whose output is rectified. The dc output of a static or rotating exciter is then delivered to the rotating synchronous machine field winding by means of slip rings. A brushless exciter is a type of rotating exciter that has a rotating armature and rectifier and thus provides field excitation without the need for slip rings. The combination of the voltage regulator and exciter is commonly called the

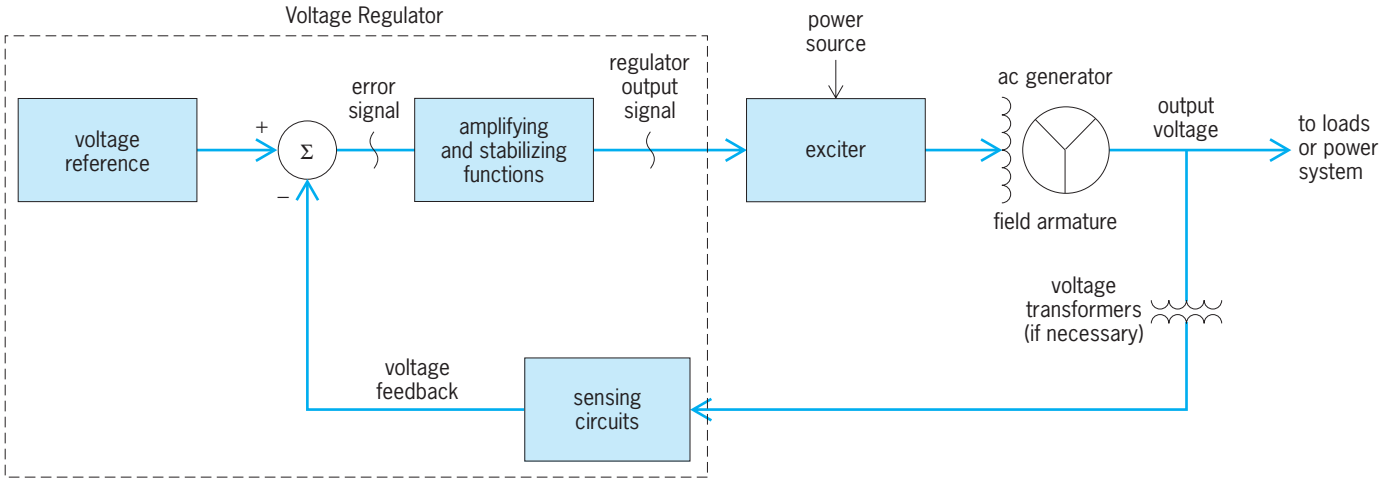

Fig. 6. Simplified block diagram of voltage regulator application.

excitation system. *See* SEMICONDUCTOR RECTIFIER; SLIP RINGS.

Regulator technology. Early voltage regulators were electromechanical devices which controlled exciter output by turning large rheostats or switching resistors with movable contacts. These regulators were slow and often required a deadband to prevent moving parts from quickly wearing out. The first solid-state continuously acting voltage regulators employed magnetic amplifiers for control. Modern analog voltage regulators utilize integrated circuits containing operational amplifiers and other electronic devices. Digital microprocessor or computer-based voltage regulators are becoming more prevalent. *See* MICROPROCESSOR; OPERATIONAL AMPLIFIER; RHEOSTAT.

Forcing capability. In addition to simply regulating steady-state generator output voltage, voltage regulators as part of an excitation system typically provide transient forcing capability during short circuits and other disturbances, allowing the exciter output to temporarily increase above that required for steady-state conditions. This capability helps the generator to more quickly recover from such faults, and thus enhances its stability or its tendency to remain in synchronism with the power system.

Limiting and portective functions. Modern voltage regulators also usually provide various limiting the protective functions which are designed to protect both the excitation system and the generator from such conditions as overexcitation (which would overload the exciter or overheat the generator field winding), underexcitation (which would jeopardize the stability of the generator connection with the power system), generator overvoltage, and excessive generator flux (as sensed by the ratio of generator output voltage versus frequency). The limiting functions provide control action to override the normal voltage regulator control signal in an attempt to provide continued operation while operating within the capability of the equipment. The protective functions back up the limiting functions and automatically remove the generator from service if the limiting functions fail to keep the operating quantities within limits. J. D. Hurley

Bibliography. P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 4th ed., Wiley, 2001; *National Semiconductor Linear Data Book*, vol. 1, 1987; R. J. Traister, *Voltage Regulator Circuit Manual*, 1989.

Volume control systems

Systems that maintain proper audio signal levels in applications such as sound recording, public address systems, and broadcasting. Two types of electronic devices, compressors and limiters, perform this operation automatically without the need for human intervention, but differ in the way that they perform.

A compressor slowly varies its gain so as to maintain its output volume level at some constant average value. When used to process the audio signal from a microphone, a compressor equalizes the loudness of speech from different talkers. It compresses the dynamic range of the signal. A limiter controls the peak levels of an audio signal. Limiters are used to prevent overmodulation of transmitters in broadcast facilities, to prevent peak clipping in public address system audio amplifiers, and to prevent overload of audio recorders. A compressor followed by a limiter can be used to dramatically enhance the loudness of an audio signal without an increase in the peak level. This is a common practice in broadcasting to increase signal range and to maximize program loudness. *See* GAIN.

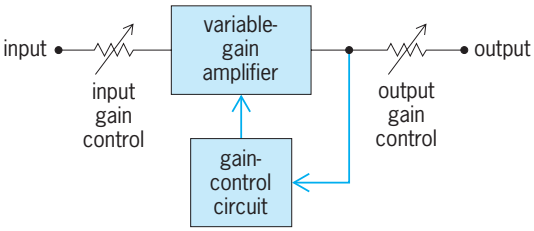

Fig. 1. Basic block diagram of a compressor and a limiter.

The input of a compressor or a limiter is applied to a variable-gain amplifier (**Fig. 1**). A gain-control circuit monitors the level of the output and generates a voltage which controls the amplifier gain. Manual controls at the input and output allow setting the levels for proper interface to external devices. *See* AUTOMATIC GAIN CONTROL (AGC).

Figure 2 illustrates the typical variation of output level versus input level for a compressor. The threshold V_t is the input voltage below which the

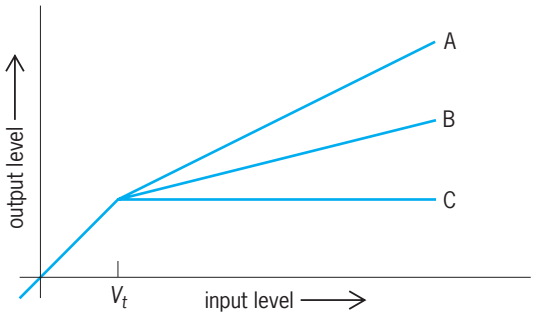

Fig. 2. Variation of compressor output level versus input level for three compression ratios.

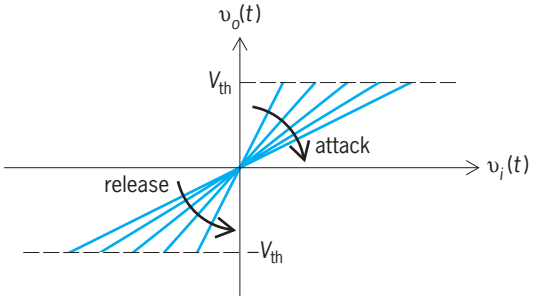

Fig. 3. Dynamic output-versus-input [$v_o(t)$-versus-$v_i(t)$] characteristic of a limiter, showing how the limiter gain decreases when its peak output voltage "tries" to exceed the threshold voltage V_{th}.

compressor gain does not vary with input. For an input greater than V_t, curves are shown for three different compression ratios. The compression ratio is the decibel change in input level required for a 1-dB change in output level. For curve A, it is 2 dB; for curve B, it is 4 dB; for curve C, it is infinite. When the input level increases, the time required for the gain to decrease is called the attack time. When the input level decreases, the time required for the gain to increase is called the release time. Typically, the attack time for a compressor is 10–100 ms and the release time is 1 to 10 s. A gated compressor is one in which the gain "freezes" when the input signal falls below threshold. This prevents low-level background sounds from being increased in amplitude.

Figure 3 illustrates the dynamic output-versus-input [$v_o(t)$-versus-$v_i(t)$] characteristics of a limiter. The gain corresponds to the slope of the $v_o(t)$-versus-$v_i(t)$ curve. For an output voltage in the range $-V_{th} \leq v_o(t) \leq V_{th}$, the gain is constant. When $v_o(t)$ "tries" to leave this range, the gain drops rapidly to limit the peak output level. This is indicated by the clockwise rotation of the slope on the graph. After the peak, the gain gradually returns to its original value. Typically, a limiter has an attack time of 1–10 ms and a release time of 10–100 ms. *See* LIMITER CIRCUIT; RADIO BROADCASTING; SOUND RECORDING; SOUND-REINFORCEMENT SYSTEM; SOUND-REPRODUCING SYSTEMS.

W. Marshall Leach, Jr.

Bibliography. G. Davis and R. Jones, *Sound Reinforcement Handbook*, 2d ed., 1990; J. M. Eargle, *Handbook of Recording Engineering*, 3d ed., 1996; D. M. Huber and R. E. Runstein, *Modern Recording Techniques*, 5th ed., 2001; D. Lyver and G. Swain son, *Basics of Video Sound*, 2d ed., 1999.

Volume unit (vu)

A unit used to measure the strength of electrical waves produced by microphones. In sound recording, broadcasting, and public address systems, microphones convert acoustical signals into electrical waves that are nonperiodic and cannot be simply described in terms of voltage, current, power, or frequency. To provide a practical means of assigning a numerical value to the strength of such waves, the concept of "volume" is used. The volume of an audio program wave is the magnitude of the wave as measured with a standard volume indicator. This is a meter calibrated to read in volume units (vu, pronounced vee-you).

A vu meter consists of an attenuator, a rectifier, and a readout device. The latter is usually an electromechanical d'Arsonval meter or a light-emitting diode (LED) array. In **Fig. 1** the block diagram of a vu meter is shown with a sine-wave input. The attenuator output is applied to the full-wave rectifier which drives the meter. The meter reading is proportional to the amplitude of the input. **Figure 2***a* illustrates a vu meter with a d'Arsonval meter readout. The level is indicated by the motion of the meter

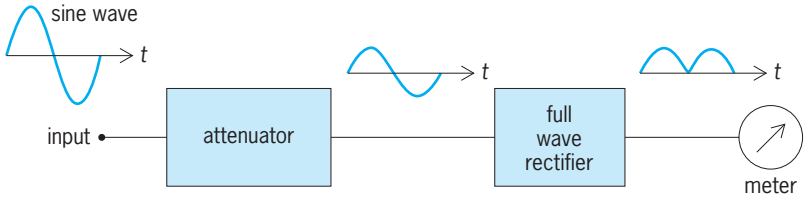

Fig. 1. Block diagram of a vu meter showing voltage waveforms for a sine-wave input.

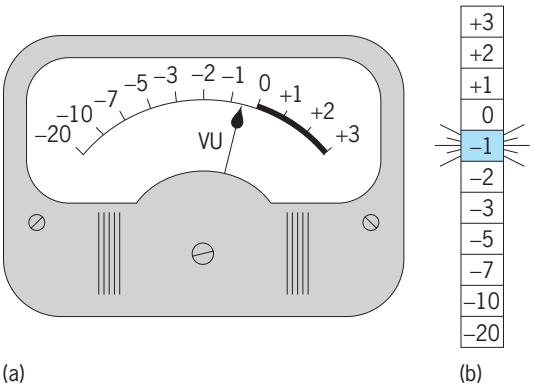

Fig. 2. Typical readouts for (*a*) a d'Arsonval vu meter and (*b*) an LED array vu meter.

needle. Figure 2*b* illustrates a vu meter with an LED readout. The level is indicated by the cell that is lit from behind by an LED. Unlike the d'Arsonval meter, the LED array must be driven by an electronic circuit. *See* LIGHT-EMITTING DIODE; RECTIFIER.

The attenuator is adjusted so that the meter reads 0 vu with a sine-wave input of 1 kHz and an amplitude that depends on the application. For example, it might be set so that 0 vu represents 1 V rms. With the sine-wave input, the meter is calibrated so that its reading corresponds to the decibel (dB) change in voltage from the reference level. *See* DECIBEL; SINE WAVE.

William M. Leach, Jr.

Bibliography. G. Davis and R. Jones, *Sound Reinforcement Handbook*, 1990; J. M. Eargle, *Handbook of Recording Engineering*, 1996; D. M. Huber and R. E. Runstein, *Modern Recording Techniques*, 1995; D. Lyver, *Basics of Video Sound*, 1995.

Volumetric efficiency

In describing an engine or gas compressor, the ratio of volume of working substance actually admitted, measured at specified temperature and pressure, to the full piston displacement volume. For a liquid-fuel engine, such as a diesel engine, volumetric efficiency is the ratio of volume of air drawn into a cylinder to the piston displacement. For a gas-fuel engine, such as a gasoline engine with carburetor, throttle body, or port injection, volumetric efficiency is based on the charge of fuel and air drawn into the cylinder. *See* COMPRESSOR; ENGINE.

Volumetric efficiency of naturally aspirated automobile and aircraft reciprocating engines may be 85–90% at rated speed. Supercharging or turbocharging increases volumetric efficiency, giving values

over 100%. At low speeds and wide-open throttle, volumetric efficiency is high. At high speeds, resistance to airflow through the carburetor or fuel-injection throttle body, the manifolds, and valve ports decreases volumetric efficiency. Two-cycle engines with crankcase compression may have a volumetric efficiency of about 60% at low speed and less than 50% at high speed. Air compressors, refrigerator compressors, and dry vacuum pumps are generally specified for a volumetric efficiency of 60–90%, dependent upon the compressor ratio, type of valve, valve gear, and machine speed. *See* INTERNAL COMBUSTION ENGINE; RECIPROCATING AIRCRAFT ENGINE.

Donald L. Anglin

Bibliography. W. H. Crouse and D. L. Anglin, *Automotive Engines*, 8th ed., 1994.

Volvocales

A large order of green algae (Chlorophyceae) comprising all forms that normally are flagellate and motile. In zoological classification, it is called Volvocida and placed in the class Phytomastigophora. The cells are solitary or united into colonies of definite structure (coenobia), often with morphological and functional differentiation among the component cells. Some taxonomists place the unicellular forms in a separate order, Chlamydomonadales. Unicells that have volvocalean cytological features but are nonflagellate and sedentary in their vegetative phase are considered here to constitute the order Tetrasporales. Alternatively, these sedentary forms may be retained in the Chlamydomonadales or Volvocales. *See* CHLOROPHYCEAE.

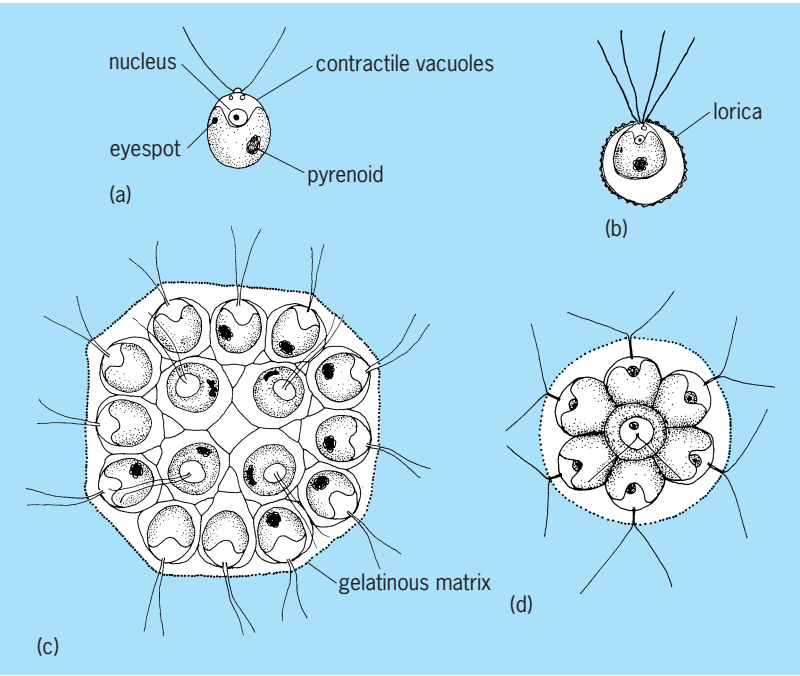

Fig. 1. Representative genera of Volvocales. (*a*) *Chlamydomonas* (Chlamydomonadaceae). (*b*) *Pedinopera* (Phacotaceae), showing granulate lorica in optical section. (*c*) *Gonium* and (*d*) *Pandorina*, both members of the Volvocaceae.

Unicellular forms. *Chlamydomona* (**Fig. 1***a*), which is presumed to be similar to ancestral Volvocales, has a very small cell body (usually 7–40 micrometers in greatest dimension) that is spherical, ellipsoid, or pyriform. The wall (or theca), which consists primarily of glycoproteins, is usually smooth, but may have blunt protuberances. The cell bears two smooth flagella of equal length at its apex, with contractile vacuoles and an eyespot at the anterior end. The chloroplast is usually cup-shaped, but it may be H-shaped, perforate, or stellate. Chloroplasts contain one or more pyrenoids. As in all Volvocales, the cells are uninucleate.

Other genera with chlamydomonad features are distinguished on the basis of flagellar number, presence or absence of photosynthetic pigments, presence or absence of pyrenoids, and cell shape. Together with *Chlamydomonas* they constitute the family Chlamydomonadaceae. Two other families of unicellular Volvocales are recognized: Dunaliellaceae, in which the unicells are naked; and Phacotaceae, in which the unicells are surrounded by a lorica—a rigid, hyaline or dark brown, wall-like structure, which is often impregnated with compounds of calcium, iron, or manganese and which may be sculptured (Fig. 1*b*).

Asexual reproduction in unicellular Volvocales is by bipartition (frequently repetitive) of a protoplast, often during a transitory nonmotile period. The progeny is released by rupture or gelatinization of the parental wall. Sexual reproduction is isogamous, anisogamous in varying degrees, or oogamous. The zygote develops a thick ornamented wall that protects it during adverse periods (as when a pool becomes dry). Meiosis occurs during germination of the zygote, so that all stages in the life history except the zygote are haploid. Many species have a temporary nonmotile gelatinous stage (palmella stage) in which the cells readily develop flagella and again become motile.

Unicellular Volvocales are ubiquitous in fresh, brackish, and coastal marine waters and in such terrestrial habitats as soil, snow, and ice. They are especially abundant in organically enriched bodies of fresh water. Nearly 50 genera with more than 800 species are recognized. Because it is common and easily cultured, *Chlamydomonas* is frequently used as an experimental organism in cytological, biochemical, and genetic studies. Certain physiological strains of *Dunaliella* are grown in mass culture as a source of carotenoids.

Colonial (coenobial) forms. Most colonial Volvocales are placed in the family Volvocaceae. In this family, coenobia are composed of *Chlamydomonas*-like biflagellate cells embedded in a firm gelatinous matrix. Coenobia have the form of a slightly curved plate or are ellipsoid to spherical. The component cells are arranged according to a predetermined pattern. A phylogenetic series, involving an increase in number of cells and the development of polarity, results in morphological and functional differentiation of the component cells. Their number is a multiple of 2, ranging from 4 (*Gonium*) to about

50,000 (*Volvox*). Each cell is surrounded by a gelatinous envelope, which may remain distinct or become confluent with those of adjacent cells. Cytoplasmic strands connect adjacent cells in *Gonium* and in some species of *Volvox*. Vegetative cells are spherical, ovoid, pyriform, or hemispherical.

Asexual reproduction is by repeated bipartition of the protoplast of an ordinary or specialized cell to form a new coenobium (autocoenobium) that is a miniature of the parent. All cell divisions are longitudinal and in a regular sequence. At the eight-celled stage, the autocoenobium has the form of a curved plate. In more advanced genera, further divisions produce a hollow sphere of very small cells, which sinks into the gelatinous matrix of the parent. There is an opening to the surface at the anterior end of the developing autocoenobium (the phialopore). The flagella are exserted into the cavity. After the last division, the phialopore expands, its lips fold back, and the hollow sphere turns inside out (everts) so that the flagella regain their normal orientation, directed outward from the coenobium. The daughter coenobium revolves slowly within the enlarged gelatinous wall of the parental cell and eventually escapes through a pore to the outside.

Sexual reproduction forms a phylogenetic series from isogamy through anisogamy to oogamy. The zygote secretes a thick wall and, after a period of dormancy, germinates meiotically to produce one to four zoospores which, like motile gametes, are biflagellate. Each zoospore divides repeatedly to form a coenobium (except in *Gonium*, where the four zoospores remain together as a coenobium, even though they are genetically heterotypic).

Gonium (Fig. 1*c*), with 4–32 cells arranged in a slightly curved quadrangular plate, is the simplest member of the family. All cells form autocoenobia simultaneously. Sexual reproduction is isogamous. In *Pandorina* (Fig. 1*d*), 16 (rarely 8 or 32) cells are arranged compactly in a hollow subspherical or ellipsoid coenobium. All eyespots are the same size, or those on one side are markedly larger, thus establishing polarity. All cells form autocoenobia simultaneously. After the last division, the bowl-shaped juvenile coenobium everts to form a sphere. Sexual reproduction is anisogamous. In *Volvulina*, the cells (usually 16) are hemispherical, the flat side facing outward, and are arranged in transverse tiers at a conspicuous distance from one another. Individual cell envelopes are visible. Eyespots in the anterior tier are larger than those in other tiers and may be lacking in the posterior half.

The coenobium of *Eudorina* comprises 8–64 (usually 32) spherical cells embedded in a homogeneous spherical or obovoid matrix that may have mammillate projections posteriorly. The cells are often in distinct transverse tiers (4–8–8–8–4). In certain species the eyespot is progressively smaller from anterior to posterior poles and may be lacking in the lowermost tier. All cells of a coenobium, or less often only certain cells, produce autocoenobia asexually in a manner typical of the family. Sexual reproduction involves an advanced type of anisogamy in which a packet of spindle-shaped male gametes resulting from repeated bipartition of a cell of a male coenobium swims as a unit to the female coenobium, where it dissociates into individual gametes. Although the female gamete is flagellate, it retains its position within its coenobium during sexual fusion.

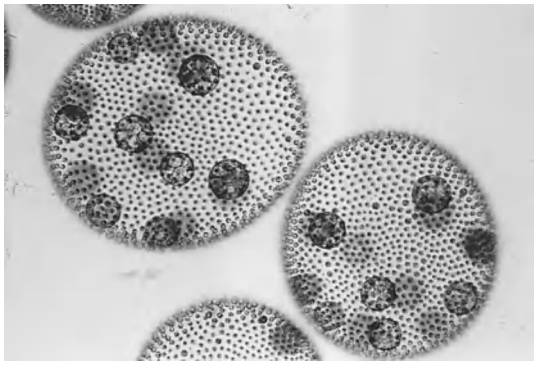

Fig. 2. *Volvox carteri*, young coenobia showing large gonidia. (*R. C. Starr, University of Texas*)

Platydorina has flattened, horseshoe-shaped coenobia, with the cells (32, 64, or 128) arranged in one layer. Depending upon the species, certain cells, varying from 4 to half the coenobium, remain small and vegetative. These may be at the anterior pole or scattered among the remaining cells, which are two or three times larger. The large cells (gonidia), which often lose their eyespot and flagella, have a massive chloroplast with several pyrenoids. All gonidia may divide simultaneously to produce autocoenobia. Sexual reproduction is similar to that in *Eudorina*, except that the female gamete may lose its flagella prior to fertilization.

Coenobia of *Volvox* (**Fig. 2**) are relatively large (0.5–1.5 mm in diameter), comprising about 500 to 50,000 cells in a single layer just within the periphery of a spherical or ellipsoid gelatinous matrix. Individual gelatinous cell envelopes are visible in most species. The cavity is filled with a watery gelatinous substance. In all embryonic coenobia, cells have cytoplasmic connections. During maturation these may either break or remain as delicate or broad strands. Eyespots decrease in size from anterior to posterior poles. During maturation of a coenobium, certain cells (2–50) in the posterior half differentiate into gonidia by enlarging, producing additional pyrenoids, and losing their eyespot and flagella. The other, smaller cells of the coenobium remain vegetative. Development of a gonidium into a daughter coenobium is typical of the order, involving eversion (**Fig. 3***a*). Sexual reproduction is oogamous. Most species are heterothallic. Homothallic species produce either unisexual or bisexual coenobia. Male gametes are produced in a packet (curved plate or compressed sphere) by enlarged cells resembling gonidia in a manner similar to the asexual development of an autocoenobium (Fig. 3*b*). Cells that function as eggs also resemble gonidia (Fig. 3*c*). The packet of male gametes dissociates as it approaches an egg.

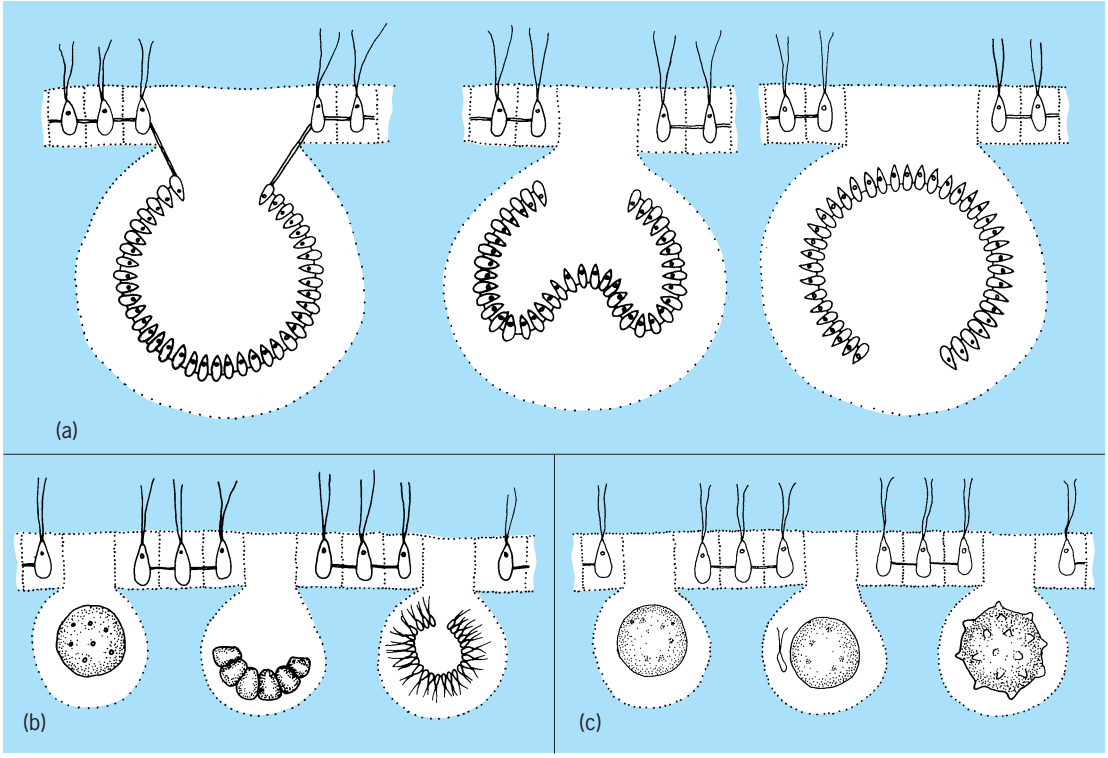

Fig. 3. Reproduction in *Volvox*. (*a*) Asexual development of an autocoenobium, showing eversion. (*b*) Development of packet of sperm from gonidium. (*c*) Fertilization of egg by sperm, resulting in spiny-walled zygote. (*After G. M. Smith, Cryptogamic Botany, 2d ed., McGraw-Hill, 1955*)

The zygote secretes a thick wall, accumulates hematochrome, and, following a period of dormancy, germinates meiotically. The result is a single zoospore, which divides repeatedly to produce a coenobium.

The Volvocaceae are abundant in organically enriched bodies of fresh water, especially temporary pools. There are 9 genera and about 48 species.

Paul C. Silva; Richard L. Moe

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae: Structure and Reproduction*, 2d ed., 1997; G. M. Smith, *The Fresh-Water Algae of the United States*, 1950.

Volvocida

An order of the class Phytamastigophorea. The protozoa, also known as the Phytomonadida, are grass-green, but a few are colorless (*Polytoma*). Individual cells may be as small as 8 micrometers. They closely relate the flagellates to the algae. They have one flagellum (*Pedimonas*), two (*Chlamydomonas*), four (*Carteria*) (**Fig. 1**), or eight (*Polyblepharides*). Plural flagella are usually equal. The group is large and about one-fourth of the approximately 100 genera form palmelloid or dendroid colonies (*Tetraspora*, *Chlorangium*), with flagella only in reproductive cells. Cell walls are of cellulose and often they are thick. Chromatophores contain the same chlorophylls as higher plants. Pyrenoids have the usual form commonly found with starch as the reserve material. *See* CILIA AND FLAGELLA.

Unicellular species predominate, but at least 16 genera form colonies. These are irregular, linear, flat, or spherical. *Volvox globator* is a hollow sphere often 1 mm in diameter. Its mucilaginous periphery contains thousands of biflagellate zooids (**Fig. 2***a*). The zooids of *V. aureus* are pear-shaped and attached by threadlike protoplasmic strands to each other; the mucilage draws away from the thick outer boundary, but several cells are incorporated into a deep wedge

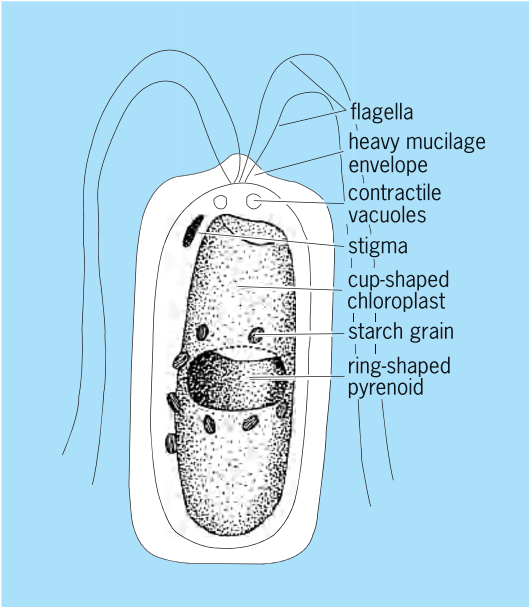

Fig. 1. *Carteria*, showing equal plural flagella.

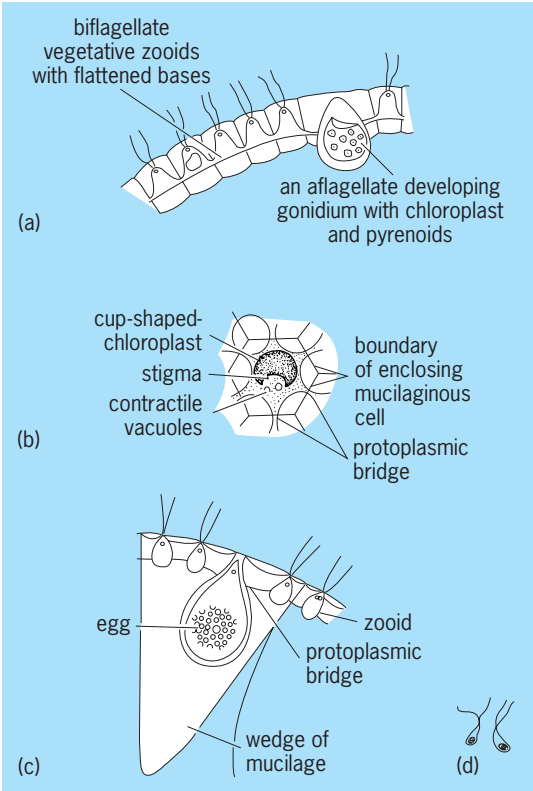

Fig. 2. Morphological features of *Volvox*. (*a*) *Volvox globator*, sectional view; (*b*) surface view of a single zooid. (*c*) *Volvox aureus*, sectional view; (*d*) sperm, enlarged about eight times by comparison with *c*.

of mucilage (Fig. 2*c*). Some zooids are gonidia that form new colonies asexually by a complicated multiple division. Isogamy, the production of like gametes which fuse, is common in Phytomonadida, but *V. globator* is monoecious, producing large eggs from certain zooids and packets of small sperm from others. Eggs, liberated into the hollow sphere, are fertilized by sperm. Thick rough walls envelop the fertilized eggs which eventually produce new colonies. Thus *V. globator* illustrates sexual reproduction comparable to that of higher animals. Most phytomonads are fresh-water inhabitants; some are terrestrial, some marine. *See* PHYTAMASTIGOPHOREA; PROTOZOA; REPRODUCTION (ANIMAL). James B. Lackey

VOR (VHF omnidirectional range)

A short-range air navigation aid, which provides azimuth aid by visual means of cockpit instruments. A VOR system provides properly equipped aircraft with bearing information relative to the VOR station and magnetic north. The VOR system is used for landing, terminal, and en route guidance. It also gives virtually static-free regular weather broadcasts, special flight instructions, and voice and code station identification. The VOR service operates in the very high frequency (VHF) band between 108 and 118 MHz, sharing alternate channels with the localizer in the instrument landing system. Typically, VOR stations are co-located with a distance measuring equipment (DME) system or a tactical navigation (TACAN) system. The combined systems are referred to as VOR/DME or VORTAC stations and provide both azimuth and distance information. *See* DISTANCE-MEASURING EQUIPMENT; INSTRUMENT LANDING SYSTEM (ILS); TACAN.

VOR/DME beacons provide air navigation services throughout the world. However, the emergence of satellite navigation systems has motivated civil aviation administrations to reconsider their continued support of these services. In the United States, the Federal Radionavigation Plan of 2001 states that "VOR/DME will continue to provide navigation services for enroute through non-precision approach phases of flight throughout the transition to satellite-based navigation." The Federal Aviation Administration (FAA) plans to reduce (but not eliminate) VOR/DME services in the U.S. National Airspace beginning in 2010 based on the anticipated decrease in the use of VOR/DME for enroute navigation and instrument approach. Any reduction will occur in phases and VOR/DME will remain in service at selected airports where it will provide a backup to satellite-based services. *See* SATELLITE NAVIGATION SYSTEMS.

There are two types of VOR equipment. One employs a four-loop antenna array and is called conventional VOR. The other, Doppler VOR, having a 50-loop or larger antenna array located around a single-loop carrier antenna, is based on the Doppler principle and is designed for installation at locations which present especially difficult siting problems due to multipath conditions. Although the two types differ in design, from the standpoint of air navigation they function in essentially the same manner and can be received by the same equipment. Both types are used to define route intersections and most domestic airways. *See* DOPPLER EFFECT; DOPPLER VOR.

Operating principles. The VOR operates on the principle that the phase difference between two signals can be employed as a means of determining azimuth if one of the signals maintains a fixed phase through 360°, so it can be used as a reference, while the other is made to vary as a direct function of azimuth. The phase difference between these two signals will then equal the azimuth of the aircraft. In practice, two demodulated 30-Hz signals are used. These are called the reference-phase and variable-phase signals.

The standard antenna system for the conventional VOR transmitter is composed of four VHF loops, mounted in a square of 48 in. (1.22 m), above a metallic counterpoise. Each diagonal pair of loops is driven from the output of a goniometer (a mechanical beam-former) and produces a figure-of-eight sideband pattern which rotates clockwise. All four loops also are driven in-phase with carrier-frequency power to radiate a circular pattern. Addition of the two patterns causes a rotating limacon [a plane curve whose equation in polar coordinates (r, A) is $r = a \cos \theta + b$], which, when detected in an

airborne receiver, produces a variable-phase signal. *See* ANTENNA (ELECTROMAGNETISM).

A 9960-Hz subcarrier amplitude modulates the carrier frequency on the circular pattern. Furthermore, this subcarrier is frequency-modulated (FM) at 30 Hz with a deviation ratio of 16 for a total deviation of 480 Hz. The aircraft receiver separates the 30-Hz amplitude-modulated (AM) variable signal and the 9960-Hz subcarrier into two separate paths for processing. The 30-Hz signal is filtered and compared to the 30-Hz frequency-modulated signal from the subcarrier. The subcarrier is limited and demodulated by a frequency discriminator circuit. It is also filtered and compared to the 30-Hz amplitude-modulated signal for the azimuth selected by the pilot. *See* AMPLITUDE MODULATION; FREQUENCY MODULATION; FREQUENCY-MODULATION DETECTOR.

Most of the VOR stations within the United States include all-solid-state equipment and give remote indications of station availability to air-traffic control personnel and remote maintenance capability to technicians. This feature ensures reliability and reduces maintenance costs since one technician is able to service several stations.

Capabilities. The VOR system shows magnetic bearing to a VOR station and deviation from a selected course, in degrees. The ground-station errors are approximately $\pm 1.4^\circ$. A VOR system allows an essentially continuous update of deviation from a selected course based on internal operations at a 30-update-per-second rate. Initialization is typically less than 1 minute after turn-on. VOR systems have line-of-sight limitations that could limit ground coverage to 30 nautical miles (55 km) or less. At altitudes above 5000 ft (1500 m) the range is approximately 100 nmi (185 km), and above 20,000 ft (6000 m) the range will approach 200 nmi (370 km). Availability depends on the location of the VOR stations and is tied into these coverage considerations.

The capacity of VOR systems is unlimited. Their reliability approaches 100%, and no ambiguity is possible for a VOR system. VOR provides system integrity by removing a signal from use within 10 seconds of an out-of-tolerance condition detected by an independent monitor. *See* ELECTRONIC NAVIGATION SYSTEMS; RHO-THETA SYSTEM.

William I. Thompson III; Robert B. Flint; Richard L. Greenspan

Bibliography. P. Enge et al., Terrestrial radio navigation technologies, *Navigation* (Journal of the Institute of Navigation), 42(1):61–108, 1995; *1999 Federal Radionavigation Plan*, pp. C25–C27, U.S. Department of Defense Rep. DOD-4650.5 and U.S. Department of Transportation Rep. DOT-VNTSC-RSPA-98-1, Washington, DC.

Vortex

In common usage, a fluid motion dominated by rotation about an isolated curved line in space, as in a tornado, a whirlpool, a hurricane, or a similar natural phenomenon. The importance of vortices is due to two characteristics: general fluid flows can be represented by a superposition of vortices; and vortices, once created, have a persistence that increases as the effects of viscosity are reduced. The aerodynamic lift forces and most other contributors to the forces and moments on aircraft and other bodies moving through fluids do not exist in the absence of vortices. *See* AERODYNAMIC FORCE; HURRICANE; TORNADO; WATERSPOUT.

Vorticity. The strength of rotation is measured by a vector called the vorticity, ω, defined as the curl of the velocity vector. A flow devoid of vorticity is known as irrotational flow or potential flow, since the velocity vector can be expressed as the gradient of a scalar function, the velocity potential. The spatial distribution of the vorticity vector provides a precise characterization of the rotation effects in fluids, and the nature of what subjectively and popularly would be called a vortex. *See* CALCULUS OF VECTORS; LAPLACE'S IRROTATIONAL MOTION.

The vorticity vector field can be constructed by measuring the instantaneous angular velocity of small masses of fluid. The vorticity vector is twice the local angular velocity vector. Starting at any arbitrary point in the fluid, a line, called a vortex line, can be drawn everywhere parallel to the vorticity vector.

Vortex tube. A bundle of vortex lines defines a tubular region of space, called a vortex tube, with a boundary surface that no vortex line crosses.

Two simple rules follow from the definitions: (1) a vortex tube must either close on itself or end on a boundary of the fluid (including extending to "infinity" if the fluid is imagined to fill all space); and (2) at every cross section of a given vortex tube, the area integral of the normal vorticity has the same value at any given instant. The area integral is, by Stokes' theorem, equal to a line integral around the periphery of the tube, namely, the line integral of the velocity component parallel to the direction of the closed curve defining the line integral. This quantity is also known as the circulation around the line, so at an instant of time a vortex tube has a unique value of the circulation applying to all cross sections (**Fig. 1**). The circulation is a measure of the speed at which fluid circulates around the path.

In many flows of interest the magnitude of the vorticity vector is small except for the interior of one tube, or a relatively few of them. Each tubular surface defines a vortex core in which the vorticity vector is large. Such situations conform most closely to the popular description of a flow as a vortex. The

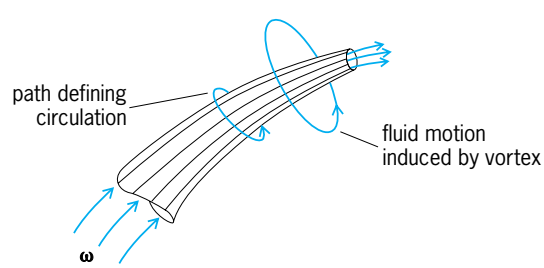

Fig. 1. Vortex tube; ω is the vorticity.

large angular velocity in the core causes a large centrifugal force on each element of fluid. To maintain mainly circular motion around the core, the centrifugal force must be balanced by a rapid decrease of pressure at the center of the vortex core. This large relative vacuum is responsible for most of the damage caused by concentrated atmospheric vortices such as tornadoes. The low pressure in the core is also responsible for the clear track of the wing-tip vortices trailing behind aircraft. These contrails are made visible by condensation of water vapor in the low-pressure cores. *See* CENTRIFUGAL FORCE.

Line vortex idealization. The core of a concentrated vortex usually has a diameter that is small compared to other dimensions associated with the vortex, such as its length or the radius of curvature of its center line. This leads to a convenient idealization, the line vortex or vortex filament, which takes the vorticity to be exactly zero outside the tube, and imagines the core area to be reduced to zero while keeping the circulation fixed. The model may be visualized by supposing that the vortex tube is straight and infinitely long, and all the motion is due to its presence. Then the motion in the irrotational region outside the core is said to be induced by the vortex, and consists of fluid moving in circular paths around the core with linear speed decreasing inversely with distance from the core center. While the fluid revolves around the vortex core, fluid particles in the irrotational region have no angular velocity around their mass centers; the situation is like a seat on a carnival ferris wheel. The resulting model has a singularity with infinite velocity at the vortex centerline—an unrealistic feature—but properly and conveniently represents the motion outside the core of the vortex being modeled.

Vortex sheet. Vortex lines confined to a layer rather than a tube describe fluid motion of a different character. This is most easily visualized when the direction of the vorticity does not vary, so all of the vortex lines are straight and parallel (**Fig. 2**). Assuming the vorticity has zero magnitude outside the layer, this vortex layer represents a flow with a different speed and direction on either side of the layer. Such a change in speed occurs at the edge of wakes produced by wind passing over an obstacle. Reducing the thickness of this layer of vorticity to zero leads to an idealization known as a vortex sheet, a surface in space across which there is a finite jump in velocity tangent to the surface (Fig. 2). Vortex sheets have a tendency to roll up because of self-induction. Aerodynamicists approximate flows in the wake of aircraft wings by vortex sheets, and the roll-up of this sheet leads to a pair of concentrated vortices trailing behind. The lift force on an aircraft depends on the formation of this trailing vortex system, and its strength is proportional to the weight of the aircraft. If the air is sufficiently humid, these vortices sometimes are visible due to condensation of water vapor in the vortex cores, which have low pressure and temperature. The trailing vortices in **Fig. 3** are made visible by injection of smoke near the wing tips. *See* WING.

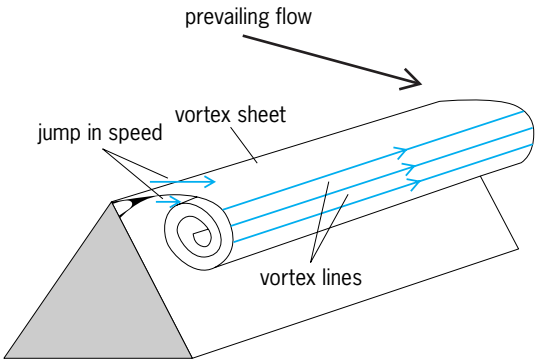

Fig. 2. Separation of flow from a sharp corner, showing a vortex sheet and its roll-up.

Fig. 3. Trailing vortex pair behind a Boeing 727 airplane (*NASA Dryden Flight Research Center file photograph ECN 3831*)

Persistence of vortex lines. In a fluid with spatially uniform density acted on by conservative external forces, the only mechanisms capable of changing the vorticity of a small element of fluid of fixed mass about its center of mass are frictional forces applied by the surrounding fluid by viscous action, and deformation of the element shape. If frictional effects are also sufficiently small to be ignored (a fluid under this approximation is called inviscid), then the circulation around any closed path that moves with the fluid is time independent. From this rule, known as Kelvin's theorem, it can be deduced that fluid particles with zero vorticity continue to have zero vorticity; and that fluid particles composing a vortex tube initially identified in a fluid continue to be a vortex tube, and the circulation for the tube is time independent. This result is usually paraphrased as "vortex lines move with the fluid." This conclusion can be extended to a special but important class of compressible motion known as barotropic flows, for

Fig. 4. Photograph of the Union City, Oklahoma, tornado of May 24, 1973. (*Courtesy of NOAA Photo Library, NOAA Central Library; OAR/ERL/National Severe Storms Laboratory (NSSL)*)

which the density depends only on the pressure. *See* FLUID-FLOW PRINCIPLES; KELVIN'S CIRCULATION THEOREM.

Vortex tubes that close on themselves roughly in the shape of a circle are not uncommon and are called vortex rings. They can readily be produced by a "vortex cannon," essentially a chamber with a circular opening opposite a drumhead. Striking the drumhead sends a puff of air that forms into a circular vortex ring outside the device. In a space without other sources of motion, a vortex ring travels surprisingly large distances: its motion is due to self-induction of the vortex lines and illustrates both induction and the notion that "vortex lines move with the fluid."

Generation of vorticity. Flows do not occur without vorticity in some region of the flow volume. Kelvin's theorem does not address the origin of vorticity. Vorticity is produced by viscosity acting at flow boundaries, by density variation, by rotation of the entire flow system, or by the action of nonconservative forces acting on the fluid. In many flows important in engineering, viscous production at boundaries is the most prevalent source of vorticity, and this often takes place only in relatively thin zones (boundary layers) adjacent to solid walls. *See* BOUNDARY-LAYER FLOW.

In natural flows that take place in the atmosphere and Earths bodies of water, vorticity is generated by viscous action too, but predominately by density variation and planetary rotation.

The last source of vorticity generation is due to action of a nonconservative force. The most prevalent nonconservative force is the electromagnetic Lorentz force. This acts in electrically conducting fluids, such as salt water, liquid metals, and ionized gases (plasmas), when electrical currents flow in the fluid in the presence of a magnetic field. This is a dominant source of vorticity in astrophysical flows, such as the sun. *See* MAGNETOHYDRODYNAMICS.

Intensification of vortices. When weak distributions of vorticity become concentrated in space we say that a vortex is formed. In flows without significant density variation or nonconservative forces, intensification occurs most commonly by one or both of two processes: vortex sheet roll-up and vortex tube stretching. Roll-up occurs because of the self-induction of vorticity elements, which causes the vortex sheet to wind increasingly tightly into tubes at the sheet edges. In this process, the vorticity in the sheet is pulled into the cores of the tubes (Fig. 2).

Vortex stretching is easiest to visualize when Kelvin's theorem applies. Then a vortex tube always encloses the same fluid material as well as the same vortex lines. If the fluid at one end of the segment of the vortex tube in Fig. 1 is moving away from the other, the tube segment stretches and its cross section decreases, a consequence of its fixed volume. The reduction of cross-sectional area squeezes the vortex lines in the bundle, increasing the vorticity in inverse proportion to the area. The speeds around the periphery of the tube increase in inverse proportion to the circumference of the tube. The increased speeds due to inward contraction of the tube is due to the same principle, conservation of angular momentum, used by a skater to increase the speed of a spin by drawing the arms inwards. *See* ANGULAR MOMENTUM.

The concentration of vorticity in a tornado and in a bathtub vortex is both due to vortex stretching. In both cases, weak background vorticity is concentrated as fluid is brought inwards. In a tornado (**Fig. 4**), this is produced by an updraft of air aloft which draws air at lower levels inwards. Draining of water produces the comparable effect in a bathtub vortex.

Sidney Leibovich

Bibliography. G. K. Batchelor, *An Introduction to Fluid Dynamics*, Cambridge University Press, 1967; H. Lamb, *Hydrodynamics*, 6th ed., 1932, reprint, Dover, 1993; H. J. Lugt, *Vortex Flow in Nature and Technology*, 1983, reprint, Krieger, 1995; P. G. Saffman, *Vortex Dynamics*, Cambridge University Press, 1992.

Vorticity

A vector proportional to the local angular velocity of a fluid element in a flowing fluid. The vorticity, $\vec{\omega}$, is a derived quantity in fluid mechanics, defined, for a flow field with velocity $\vec{u}$, by Eq. (1). As the curl of

$$\vec{\omega} = \nabla \times \vec{u} \tag{1}$$

the velocity vector, the vorticity is a vector with the dimensions of both a frequency and an angular velocity: $[\vec{\omega}] = [s^{-1}]$. As shown below, the component of vorticity along a particular axis is related to the rate of rotation of the fluid about the axis. For this

reason, flows for which $\vec{\omega} = 0$ are described as irrotational. *See* CALCULUS OF VECTORS; DIMENSIONAL ANALYSIS; VELOCITY.

Circulation. A property that is closely related to vorticity is the fluid circulation, Γ, defined, for any closed contour, C, in a fluid, by Eq. (2). In this def-

$$\Gamma = \oint_C \vec{u} \cdot d\vec{l} = \iint_S \vec{\omega} \cdot \vec{n}\, dS \qquad (2)$$

inition, $\oint_C$ indicates the conventional counterclockwise contour integral around the contour C, $\vec{l}$ is a unit vector tangent to the contour, S is an arbitrary curved surface bounded by the contour C, and $\vec{n}$ is a unit vector normal to this surface. [The equality of the two integrals in Eq. (2) may be deduced by the application of Stokes' theorem.] The circulation is thus a scalar quantity equal to the integrated component of vorticity normal to the surface around which Γ is taken. Physically it represents the average fluid spin in a given area. Circulation is important because the Kutta-Joukowski law of aerodynamics states that the lift generated by a two-dimensional airfoil is $L = \varrho U \Gamma$. In this expression, ϱ is the fluid density, U is the free-stream velocity, and Γ is the bound circulation of the airfoil, defined conventionally as the negative of the definition above. As a consequence, the lift generated by an airplane wing is proportional to the circulation around it. *See* AERODYNAMICS; AIRFOIL; STOKES' THEOREM; SUBSONIC FLIGHT.

Example: solid-body rotation. Perhaps the simplest flow with vorticity is fluid in solid-body rotation. Such a flow is ultimately achieved when a cylindrical tank of water is spun about its axis (say, the z axis) at a constant rate, Ω. For this flow the vorticity is $\omega_z = 2\Omega$, which is equal to twice the angular velocity of the fluid element about the center of the turntable. This angular velocity is also precisely the angular velocity of each infinitesimal fluid element about its own axes, because the fluid is in solid-body motion.

The solid-body rotation flow is not the only one for which the vorticity is twice the angular velocity of a fluid element. For any flow, $\vec{\omega} \cdot \vec{a}$ (the component of $\vec{\omega}$ along the direction $\vec{a}$, where $\vec{a}$ is an arbitrary unit vector) is twice the mean angular velocity of any two line segments, through the point where $\vec{\omega}$ is evaluated, that form an orthogonal coordinate system with $\vec{a}$. For the solid-body rotation flow, the circulation around a circle of radius r centered on the origin is $\Gamma = \pi r^2 \times (2\Omega)$. Consistent with the circulation-vorticity relationship explained above, the circulation calculated here is indeed the area of the circle (πr^2) times the uniform vorticity (2Ω).

Example: irrotational vortex. Consider a flow with only a tangential velocity component, of the form $v_\theta = \Gamma/2\pi r$. Any closed loop taken around the origin of such a flow has the constant circulation Γ. Any closed loop that does not include the origin has $\Gamma = 0$. In other words, this velocity field represents a flow with infinite vorticity at the origin but zero vorticity elsewhere; i.e. the vortex, called an "irrotational vortex," is irrotational everywhere except at the origin. Such a flow is physically unattainable because the action of viscosity prohibits the infinite velocity gradients of this flow. The irrotational vortex is a mathematically useful concept, however, and forms a key element of potential flow theory. *See* POTENTIAL FLOW.

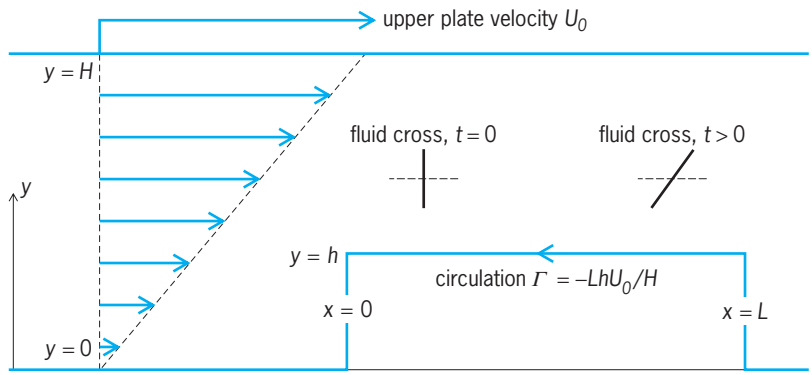

Fig. 1. Vorticity and circulation in a laminar plane Couette flow.

Example: laminar plane Couette flow. Another flow that sheds light on the concept of vorticity is the laminar Couette flow between two plates. Consider two infinite plates parallel to the x–z plane: a stationary one at $y = 0$ and a second one moving at velocity U_0 in the x direction, located at $y = H$ (**Fig. 1**). The velocity between the plates is given by $u = U_0(y/H)$, and the streamlines of the flow are thus all parallel to the x axis. The vorticity is given by $\omega_z = -U_0/H$, a constant. Surprisingly, in this parallel flow there is vorticity. The vorticity is a consequence of the fact that a cross placed in this flow would have its horizontal arm remain horizontal, whereas its vertical arm would be rotated clockwise. *See* LAMINAR FLOW.

It is instructive to consider a contour in this flow bounding the region $0 < x < L$, $0 < y < h$. The circulation about this contour is $\Gamma = -Lh(U_0/H)$. As was true for the solid-body-rotation flow, which also had constant vorticity throughout the domain, the circulation here is equal to the area of the surface (Lh) about which the circulation is measured, times the vorticity.

Vortex line and vortex tube. With a sense of what vorticity is, it is possible to contemplate what produces it and what causes it to change. It is convenient to start by defining a vortex line as a line that is everywhere tangent to the local vorticity vector (analogous to a streamline). A series of adjacent vortex lines is referred to as a vortex tube. The first Helmholtz vortex law states that at any instant in time the circulation about all loops taken around the exterior of a vortex tube is the same. Thus, vortex tubes must either form loops entirely within a fluid or terminate at some fluid boundary. *See* VORTEX.

Kelvin's theorem. Kelvin's theorem considers how the circulation Γ around a material loop in a fluid (a loop that moves with the fluid) varies in

time. Starting with the Navier-Stokes equations, Lord Kelvin showed that if (1) the flow is inviscid along the loop, (2) the fluid is subject only to potential body forces, and (3) the fluid pressure is a function of density alone, then the rate of change of Γ is 0. In other words, the circulation around a material loop is time-independent. Kelvin's theorem may also be stated slightly differently: subject to the above three constraints, vortex lines are material lines, convected with the local fluid velocity. *See* KELVIN'S CIRCULATION THEOREM; NAVIER-STOKES EQUATION.

Kelvin's theorem provides much insight into vorticity. First, consider an incompressible swirling flow in a converging section of a tube. Neglecting the three confounding effects mentioned above, as the flow converges the circulation about a material loop taken around the tube circumference remains constant. Since the flow area is diminished, this constancy of circulation implies that the mean vorticity is increased—the fluid spins more rapidly. The physical explanation for this enhancement of vorticity is that the fluid in the converging tube is stretched axially (to conserve fluid mass). This axial stretching increases the rotation rate of the fluid, for the same reason that a figure skater spins more rapidly with arms stretched overhead. The very strong winds associated with tornadoes are a result of the stretching of vortex lines in the atmosphere caused by a combination of updrafts and wind shear. *See* TORNADO.

Now, consider a cylindrical tank of air, spun about its axis and in solid-body rotation, that is gradually compressed by reducing the radius of the tank. Subject to the three constraints above, Kelvin's theorem shows that since the circulation is constant and the flow area is reduced, the vorticity is intensified. The vorticity is intensified here in order to preserve the angular momentum of the fluid. This intensification is important in engines, where, during the compression stroke, vorticity oriented parallel to the piston face (referred to as tumble) increases. *See* ANGULAR MOMENTUM; ENGINE.

Generation. Kelvin's theorem can tell us what happens when vorticity is already present in a flow, but it sheds no light on how vorticity is generated. To answer this question, it is useful to consider situations for which Kelvin's theorem is inapplicable: flow with viscosity, with nonpotential body forces, and for which the pressure is not solely a function of the density.

The action of viscosity has two effects on vorticity. One effect of viscosity is to cause the diffusion of vorticity in a fluid. The diffusion of vorticity is related to the transfer of angular momentum from one rotating object to another, by means of a frictional interface between the bodies. A classic example of this diffusion of vorticity is the Lamb-Oseen vortex. The Lamb-Oseen vortex is a two-dimensional time-dependent flow. For simplicity, one may imagine that the flow is in the x-y plane. The flow is initially an irrotational vortex of circulation Γ_o, in which the fluid travels in concentric circles at speeds inversely proportional to the distance from the center of flow. The vorticity is then zero except at the exact center of the vortex, where it is infinite. At times $t > 0$ viscosity is allowed to act. Starting with the Navier-Stokes equations; assuming that the fluid is incompressible, has constant density, and is subject only to potential body forces; and taking into account the two-dimensional axisymmetric character of the flow, it can be shown that the vorticity at times $t > 0$ and distance r from the center of flow is given by Eq. (3), where ν is the kinematic viscosity. Subjected

$$\omega_z = \frac{\Gamma_o}{4\pi\nu t}\exp\left(\frac{-r^2}{4\nu t}\right) \tag{3}$$

to viscosity, the flow develops a solid-body-rotation region, with constant vorticity, near the vortex center. The effect of vorticity is to cause the size of this region to increase, thus diffusing the initially concentrated vorticity (**Fig. 2**). The corresponding tangential velocity profiles around the vortex of Fig. 2 are shown in **Fig. 3**. *See* VISCOSITY.

A second effect of viscosity is the generation of vorticity at a wall where there is a pressure gradient at the wall. Vorticity present in the boundary layer over a flat plate is entirely vorticity that has been generated at the leading edge of the plate, where a pressure gradient exists, which has been convected downstream.

A common example of a nonpotential body force is the Coriolis force, which is present in a rotating frame of reference. This force generates vorticity in a fluid, and is a major cause of the large-scale circulation in the atmosphere and oceans. *See* CORIOLIS ACCELERATION.

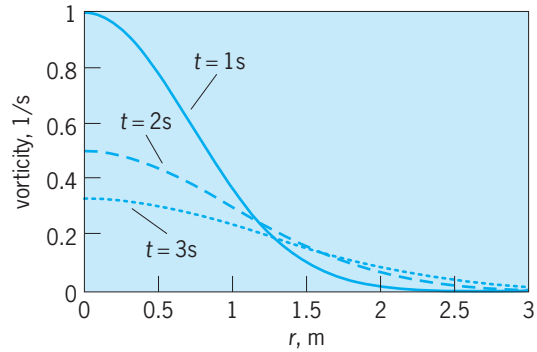

Fig. 2. Vorticity distribution in a Lamb-Oseen vortex. Here, $4\nu = 1$ and $\Gamma_o = \pi$ in SI units.

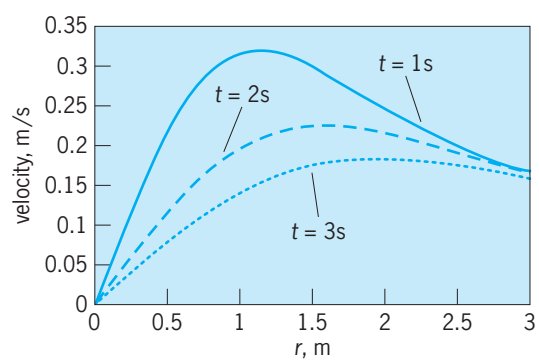

Fig. 3. Velocity distribution in a Lamb-Oseen vortex.

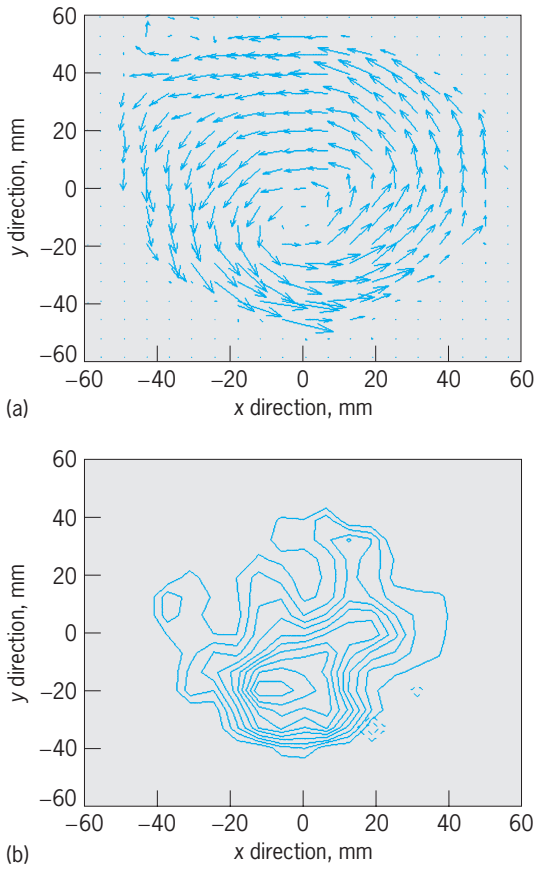

Fig. 4. Velocities and vorticity for the flow in a cardiac assist device. (*a*) Vector plot of velocities, generated by analyzing a particle image velocimetry (PIV) image taken parallel to one surface of the device. (*b*) Contour plot of the vorticity. Two small regions of negative vorticity are shown as broken contours. 1 mm = 0.04 in. (*Doug Hart and Hayden Huang, Massachusetts Institute of Technology*)

There are many flows for which the pressure may not be solely a function of the density (so-called baroclinic flows), such as the flow of gas with heat addition and the flow of water with salinity variations. Pressure gradients in such flows generate vorticity. This source of vorticity is called baroclinic torque, and is important in atmospheric flow, buoyancy-driven flow, and oceanographic flow. The physical mechanism underlying baroclinic torque is similar to that which produces a righting moment on a sailboat that is heeled over. *See* BAROCLINIC FIELD; DYNAMIC INSTABILITY; DYNAMIC METEOROLOGY.

Measurement. It is virtually impossible to measure the vorticity in a flow directly. Rather, one normally measures fluid velocities at many locations in space and then takes spatial derivatives of the velocities (since the vorticity is the curl of the velocity) to find the vorticity. Owing to the fact that derivatives of the velocity must be determined, accurate velocity information is required at a number of locations in space. Such information is commonly obtained in steady flows by means of laser Doppler velocimetry (LDV), hot-wire anemometry (HWA), or particle image velocimetry (PIV), but in unsteady flows only PIV may typically be used. **Figure 4***a* shows an example of such an unsteady flow in a cardiac assist device. Typical velocities in this device are about 50 mm/s (2 in./s). A strong vortex is present, centered slightly below and to the left of the center. Illustration *b* shows the vorticity distribution in this flow, which was computed using the data of illus. *a*. The region of highest vorticity corresponds with the center of the vortex.

Computational fluid dynamics. The governing equations of fluid mechanics may be cast in the form of differential equations for the vorticity. These are, in some (fairly rare) situations, more readily solved than the original governing equations. Since vortices form an important aspect of many fluid flows, methods that track the motion of vortices (such as large eddy simulation) have been highly successful in flow modeling. Since circulation is central to aerodynamic lift, numerical methods that treat lifting surfaces as a number of vortices (such as the vortex lattice method) can be used to model the aerodynamics of planes. *See* COMPUTATIONAL FLUID DYNAMICS; FLUID-FLOW PRINCIPLES.

Sheldon I. Green

Bibliography. S. I. Green (ed.), *Fluid Vortices*, Kluwer, Boston, 1995; A. J. Majda and A. L. Bertozzi, *Vorticity and Incompressible Flow*, Cambridge University Press, New York, 2002; P. S. Marcus, Jupiter great red spot and other vortices, *Annu. Rev. Astron. Astrophys.*, 31:523–573, 1993; P. H. Renard et al., Dynamics of flame/vortex interactions, *Prog. Energy Combust. Sci.*, 26(3):225–282, 2000; V. J. Rossow, Lift-generated vortex wakes of subsonic transport aircraft, *Prog. Aerosp. Sci.*, 35(6):507–660, 1999; T. Sarpkaya, A critical review of the intrinsic nature of vortex-induced vibrations, *J. Fluids Struct.*, 19(4):389–447, May 2004; C. H. K. Williamson and R. Govardhan, Vortex-induced vibrations, *Annu. Rev. Fluid Mech.*, 36:413–455, 2004.

Wage incentives

Plans that link employee compensation to some measure of company success. Wage incentive plans throughout the United States are varied. Some plans endeavor to link compensation to the company's performance history, others to department or unit performance, others to individual or team success. However, the majority of wage incentive plans are linked to individual achievement.

Modern wage incentive systems provide extra earnings when achievement exceeds some established bench mark. However, workers receive only a base rate, which in some installations may be reduced when achievements become less than the established bench marks. Thus, modern wage incentive plans promote the concept that workers share in both the rewards and the risks in doing business.

Experience has proven that properly designed and managed wage incentive plans that tie pay closely to performance are a way that companies can increase financial rewards and achieve gains in both productivity and quality. Plans must be developed so that workers understand that their individual efforts are directly related to the established benchmark goals. Specific incentive plans attached to individuals or small groups such as teams rather than divisions or whole plants usually are more effective.

Established benchmarks often include more than a given number of pieces produced. Such factors as product quality, plant safety, and customer satisfaction are involved in establishing the wage incentive plan. A plant that applies wage incentives throughout the organization will usually have a different plan for services than it has for direct labor, where output, quality, and safety readily are measurable.

The following criteria help assure the success of the wage incentive installation.

1. *Benchmark goals.* These should be established by careful measurement and should be attainable. Workers should understand how the benchmarks were established and realize that the workers can readily meet the targets when the members of their group have been properly trained and are giving normal effort.

2. *Meaningful goals*. Quantity goals are easily understood. However, goals based on quality must be based upon readily understood targets such as "no more than 1% defective parts per 1000" or "no more than one day of lost time due to accidents per month."

3. *Employee participation*. As the wage incentive plans are being developed, it is important to give employees an opportunity to make suggestions.

4. *Flexibility*. It is important to adjust plans as necessary in order to take into account changing methods, facilities, products, and customers.

The principal wage-incentive plans are piecework, standard hour, gain sharing, and profit sharing.

Piecework. Implied in the term piecework is the concept that all standards are expressed in money and that workers are rewarded in direct proportion to output. Under straight piecework, the day rate is not guaranteed. Since in the United States federal law requires a minimum guaranteed hourly rate for full-time employees, straight piecework is not used on a large scale. The reasons for the popularity of piecework are that it is easily understood by the worker, easily applied, and easily kept up to date as methods or product mix change. It is widely used in the fruit and farm harvest industry. When used in conjunction with the minimum wage laws, the base rate is guaranteed, and this is usually established by job evaluation. Output of the worker to the point of the guaranteed base rate usually is made a requirement of employment, and output above the guaranteed base rate is compensated in direct proportion to work produced.

Standard hour plan. The standard hour plan with a guaranteed base rate, established by job evaluation, is a form of incentive wage payment that is used frequently. The difference between the standard

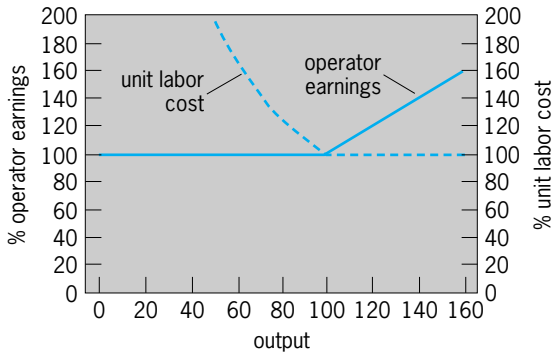

Relationship of labor cost and operator earnings to operator output.

hour plan and piecework is that under the standard hour plan, standards are expressed in time rather than money (see **illus.**). Operator earnings are based upon standard hours earned multiplied by base rate per hour. If the worker earns fewer standard hours than clock hours, the wage paid is calculated as clock hours times base rate per hour.

Certain requisites for a sound standard hour plan have been established. It has been proven that a sound standard hour wage-incentive plan will increase the rate of production with no detrimental effect of quality, lower overall unit costs, reduce supervision costs, and promote increased earnings of employees. Before installation of a wage-incentive plan based upon output is made, there should be incorporated a policy of methods standardization so that sound work measurement can be accomplished. Also, established base rates should be equitable and should provide for a sufficient spread between job classes to recognize the positions that demand more skill, effort, and responsibility.

Fair standards of performance must be developed before a valid standard hour installation can take place. With both piecework and standard hour plans, standards should be based on some form of work measurement such as time study, fundamental motion data, standard data, formula, or work sampling. The plan should give workers the opportunity to earn approximately 20–35% above their base rate if they are normally skilled, are properly trained, and execute high effort continually. The terms of the plan should be understood easily by all who are compensated by it. The plan should guarantee the basic hourly rate set by job evaluation; this basic rate should be a good living wage comparable to the prevailing wage rate of the area for each job in question. There should be a range of pay rates for each job, and these should be related to total performance—a concept that includes quality, reliability, safety, and attendance as well as output. *See* WORK MEASUREMENT.

Sound standard hour wage-incentive systems have proven to be one of the most effective motivators to increase productivity. Productivity, which is the measure of output based upon following a prescribed method and utilizing average skill and effort, is assigned a value of 100%. This value frequently is increased to 135%, and sometimes a higher percentage, under a well designed and administered standard hour incentive installation.

In those organizations where a detailed methods study has not been made and with no work measurement or wage incentives, productivity typically is about 50%. With methods analysis followed by work measurement and good supervision, the level of productivity usually rises to 90–100%. A standard hour wage-incentive installation can improve this to 125–135%. *See* PRODUCTIVITY.

Gain-sharing plans. These plans are also referred to as productivity sharing plans. Gain-sharing plans are characterized by sharing the benefits of improved productivity, cost reduction, quality improvement, customer satisfaction, and accident reduction. These plans are often used in addition to some other plan such as straight day work or a standard hour plan. Gain-sharing plans have become popular, since there are several factors other than productivity that can enhance the success of an industry or business.

Gain-sharing incentive plans can provide motivation for savings in material (both direct uses and factory supplies), reduction in rejects and scrap, improvement in product quality, accident prevention, and improved customer satisfaction. The plans tie the incentive to bench marks. Performance that achieves results better than established benchmarks is converted into monetary values that are shared by workers and the company. Benchmark values are usually based upon average historical performance for a period of time, usually one year. Often the bench mark is a production value, which is the monetary difference between sales and purchases, and consequently represents the value that all employees have contributed. Previous performance, for the past year, is considered standard. Value added can be improved by (1) savings in raw materials, purchased parts, supplies, fuel, and power; (2) reduction in rejects, rework, and scrap; (3) improvement in product overall quality; (4) reduction in worker absenteeism; (5) reduction in customer allowances, including less warranty service; (6) increase in the volume of output; (7) reduction in lost time resulting from accidents; and (8) greater productivity without increases in time inputs by hourly or salaried employees.

Under gain-sharing plans, management usually computes incentives on a monthly basis, and this is identified as a separate check from the regular paycheck (typically issued weekly). Customarily, only two-thirds of the incentive earned in a given pay period is distributed. The remaining third is placed in a reserve fund to be used any month that performance falls below standard.

Profit sharing. Profit sharing may be thought of as a form of gain sharing, since it refers to any procedure where an employer pays to all employees, in addition to the standard rates of regular pay, special current or deferred sums based not only upon individual or group performance but on the prosperity of the business as a whole. Most profit-sharing systems are either cash or deferred plans. Cash plans give

periodic distributions of money from the profits of the business, while deferred plans feature the periodic investment of portions of the profits for employees. A successful profit-sharing program depends on the profits of the company, which frequently are not under the control of the labor force. In periods of low profits or of losses, the plan may actually weaken rather than strengthen employee morale. *See* WORK STANDARDIZATION. Benjamin W. Niebel

Bibliography. C. H. Fay and R. W. Beatty, *The Compensation Source Book*, 1988; H. Gleckman, S. Atchison, and T. Smart, Bonus pay: Buzzword or bonanza, *Bus. Week*, pp. 62–64, November 14, 1994; J. Hill, Employees blunder into rewards, *U.S.A. Today*, May 22, 1990; B. W. Niebel, *Motion and Time Study*, 9th ed., 1993; G. Salvendy (ed.), *Handbook of Industrial Engineering,* 3d ed., 2001.

Wake flow

The flow downstream of a body immersed in a stream, or the flow behind a body propagating through a fluid otherwise at rest. Wakes are narrow elongated regions aligned with the flow direction and filled with large and small eddies. The wake eddies of a bridge pier immersed in a river stream, or of a ship propelled through the water, are often visible on the surface. On windy days, similar wakes form downwind of towers, smoke stacks, or other structures, but such eddies in the air are not generally visible unless some smoke or dust is entrained in them. Wakes are sustained for very large distances downstream of a body. Ship wakes retain their turbulent character for miles behind a vessel and can be detected by special satellites hours after their generation. Similarly, condensation in the wake of aircraft sometimes looks like narrow braided clouds, traversing the sky.

Turbulence in the wake of bluff bodies consists of all sizes of eddies, which interact with each other in their unruly motion. Yet, out of this chaos emerges some organization, whereby large groups of eddies form a well-ordered sequence of vortices. These vortices are rolling and moving downstream much like the waves on the surface of the ocean, and for this reason they are often referred to as rollers. The sense of rotation of these vortices alternates, and their spacing is quite regular. As a result, they can drive a structure that they encounter, or they can exert on the body that created them a force alternating in sign with the same frequency as that of the formation of the vortices. Such rollers can be experienced when following in the wake of a large truck. *See* KÁRMÁN VORTEX STREET; TURBULENT FLOW.

The flow over bodies that generate steady and significant lifing forces, that is, nonbluff bodies, forms wakes dominated by vortices that are aligned with the direction of the oncoming stream. Such vortices are often called streamers. Typical streamers are the tip vortices that are shed from the tips of airplane wings (see **illustration**). Tip vortices pose grave danger to following aircraft, especially in cases of landing. These vortices continue spinning over the landing area minutes after an aircraft has landed and taxied away, and could induce violent motions in another aircraft with catastrophic results. This phenomenon extends the time between landings and is the major cause of airport congestion. *See* WING.

Flow visualization of aircraft wing-tip vortices. (*NASA*)

The same type of streamwise vortices is encountered over other types of engineering structures with different effects. For example, streamers are shed off the corners of buildings during a violent storm, creating very large suction forces that damage the roofing material or lift the entire roof. *See* FLUID FLOW; FLUID-FLOW PRINCIPLES; VORTEX. Demetri P. Telionis

Bibliography. R. L. Panton, *Incompressible Flow*, Wiley Interscience, 3d ed., 2005; F. M. White, *Fluid Mechanics*, McGraw-Hill, 5th ed., 2003.

Wall construction

Methods for constructing walls for buildings. Walls are constructed in different forms and of various materials to serve several functions. Exterior walls protect the building interior from external environmental effects such as heat and cold, sunlight, ultraviolet radiation, rain and snow, and sound, while containing desirable interior environmental conditions. Walls are also designed to provide resistance to passage of fire for some defined period of time, such as a one-hour wall. Walls often contain doors and windows, which provide for controlled passage of environmental factors and people through the wall line.

Structure and design. Walls are designed to be strong enough to safely resist the horizontal and vertical forces imposed upon them, as defined by building codes. Such loads include wind forces, self-weight, possibly the weights of walls and floors from above, the effects of expansion and contraction as generated by temperature and humidity variations as well as by certain impacts, and the wear and tear of interior occupancy. Some walls are also designed to provide

diaphragm strength to help keep the entire building erect. Diaphragm strength refers to the ability of a wall to resist forces applied in the plane of the wall. Thus, a design to resist wind or earthquake forces may utilize walls with diaphragm strength, oriented parallel to such forces. If a wall has windows or doors, they require frame and lintel structure within the wall itself. *See* LOADS, DYNAMIC; LOADS, TRANSVERSE.

Older walls typically had to help provide support for the floors and walls above, and therefore they were classified as bearing walls. Modern engineering has led to lighter buildings that can be supported efficiently by structural frames of steel, concrete, or wood, instead of by the walls. This has led to the development of curtain walls, which are nonbearing and provide primarily for enclosure, not for structural support of the building.

Types of modern walls. Modern building walls may be designed to serve as either bearing walls or curtain walls (**Fig. 1**) or as a combination of both (**Fig. 2**) in response to the design requirements of the building as a whole. Both types may appear similar when complete, but their sequence of construction is usually different.

Bearing walls. Bearing-wall construction may be masonry, cast-in-place or precast reinforced concrete, studs and sheathing, and composite types. The design loads in bearing walls are the vertical loading from above, plus horizontal loads, both

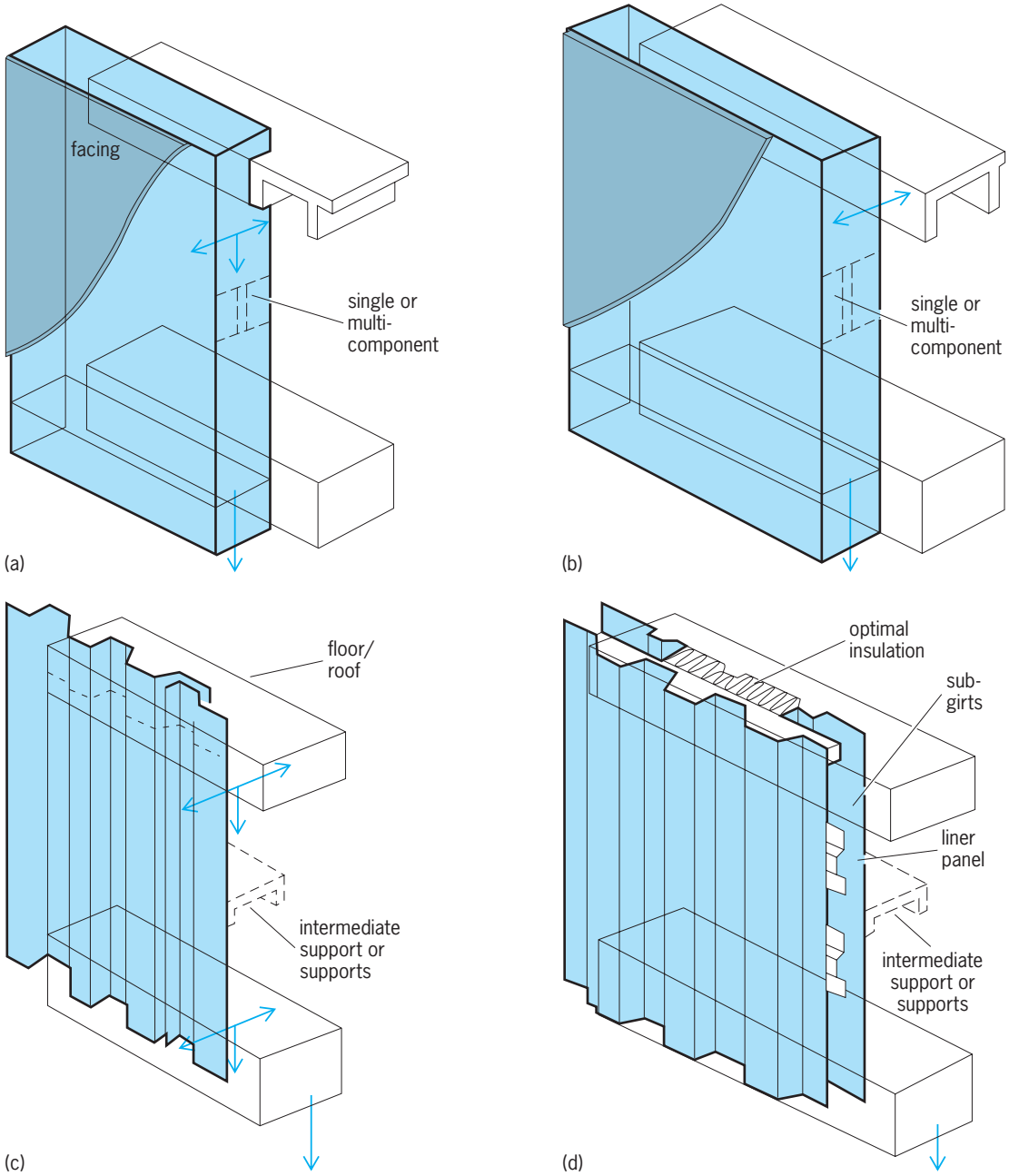

Fig. 1. Basic bearing- and curtain-wall construction. Arrows refer to force vectors, such as vertical dead loads, or horizontal wind loads. (*a*) Bearing wall. (*b*) Nonbearing (but self-supporting) wall. (*c*) Formed metal curtain wall, single thickness. (*d*) Formed metal sandwich curtain wall. (*After Sweets Catalog File: General Building and Renovation, 1993*)

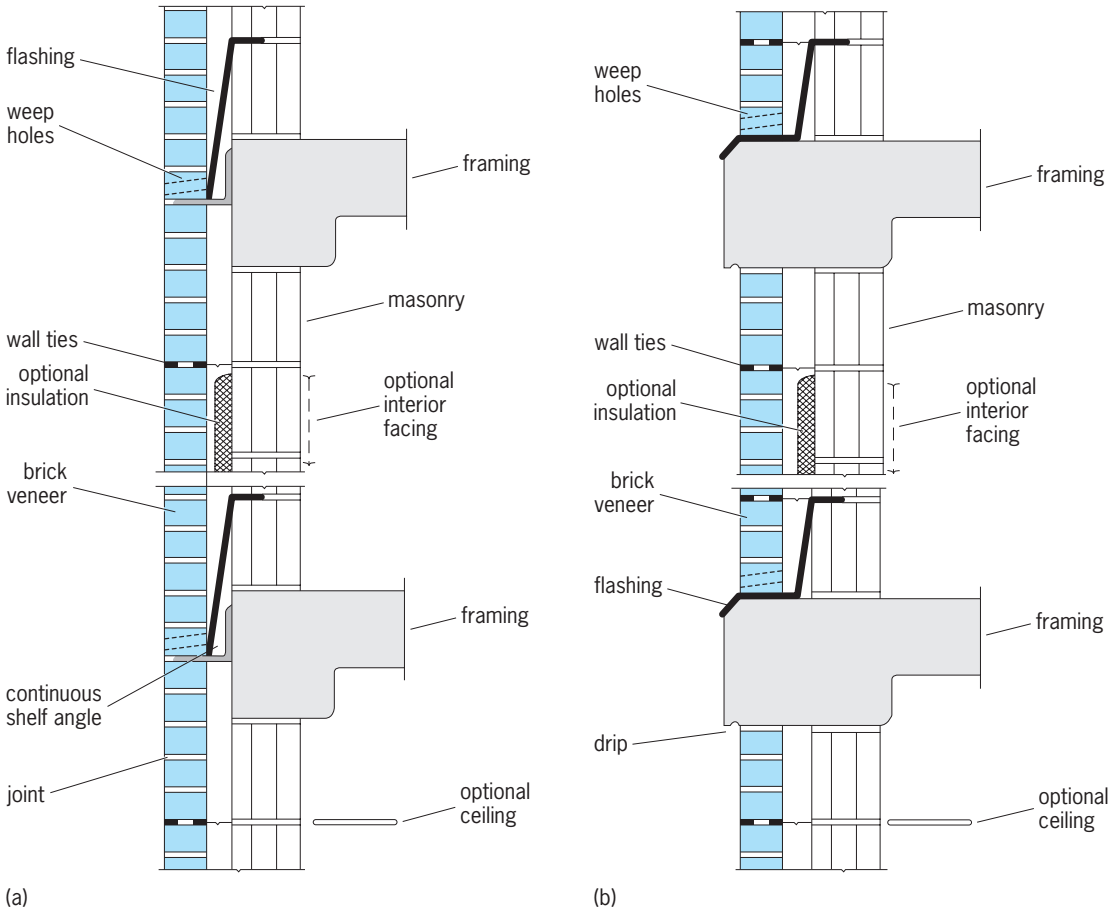

Fig. 2. Details of some masonry wall construction. (*a*) Continuous veneer. (*b*) Interrupted veneer. (*After Sweets Catalog File: General Building and Renovation, 1993*)

perpendicular and parallel to the wall plane. Bearing walls must be erected before supported building components above can be erected.

Curtain walls. Curtain-wall construction takes several forms, including lighter versions of those used for bearing walls. These walls can also comprise assemblies of corrugated metal sheets, glass panels, or ceramic-coated metal panels, each laterally supported by light subframing members. The curtain wall can be erected after the building frame is completed, since it receives vertical support by spandrel beams, or relieving angles, at the wall line (Fig. 2). The primary design loading for a curtain wall is lateral loading, perpendicular to the wall, due to wind and interior occupancy.

Masonry walls. These are a traditional, common, and durable form of wall construction used in both bearing and curtain walls. They are designed in accordance with building codes and are constructed by individual placement of bricks, blocks of stone, cinder concrete, cut stone, or combinations of these. The units are bonded together by mortar, which is composed of cement, sand, water, and sometimes lime; the mortar may be chemically enhanced by latex or epoxy additives to provide greater tensile strength. Sometimes hollow-block masonry walls are reinforced by vertical rebars, grouted into the block openings. *See* MORTAR.

Masonry walls are usually built up of at least two separate thickness, called wythes. An example is a wall composed of 4-in. (10-cm) brick and 8-in. (20-cm) concrete block. The wythes may be bonded together to form a solid wall, or they may have a space between them to form a cavity wall. The two wythes are usually structurally connected so as to act essentially as a single structural unit.

In the solid wall, the wythes are bonded to each other by mortar. Such solid walls commonly offer little insulation value, and may allow some moisture penetration from wind-driven rains unless special veneers or coatings are provided on the exterior or interior surfaces.

In the cavity wall, the wythes are typically connected to each other across the cavity by spaced steel wall ties or by welded steel trussed or parallel rod reinforcement, embedded in opposite mortar joints in each wythe. Special ties may be used to permit small movements between the two wythes.

The cavity is often used for inclusion of insulation such as fiberglass, styrofoam, or urethane foam; it also allows for drainage of any condensed or leaked moisture. Such moisture falls downward by gravity and is intercepted by waterproof metal, plastic, or asphalted fabric flashing, which spans across and downward through the cavity. This flashing causes such water to leave the wall through weep holes,

which are small holes placed in the mortar joints and spaced several feet apart (Fig. 2).

Some masonry walls, as well as other types of walls, may have a veneer of special materials such as thin brick; ceramic tile; or cementitious, elastomeric, or other water-resistant coatings; these are affixed to the exterior or interior faces for esthetic or protective reasons. *See* CONCRETE; MASONRY.

Reinforced concrete walls. These are used for both strength and esthetic purposes. Such walls may be cast in place or precast, and they may be bearing or curtain walls. Some precast concrete walls are constructed of tee-shaped or rectangular prestressed concrete beams, which are more commonly used for floor or roof deck construction. They are placed vertically, side by side, and caulked at adjacent edges. Sometimes the casting encloses insulation. In other solid types, insulation is attached on the inside face of the casting and is then protected by an inside facing material. *See* CONCRETE BEAM; REINFORCED CONCRETE.

Stud and sheathing walls. These walls are a light type of wall construction, commonly used in residential or other light construction where they usually serve as light bearing walls. They usually consist of wood sheathing nailed to wood or steel studs, usually with the dimensions 2 × 4 in. (5 × 10 cm) or 2 × 6 in. (5 × 15 cm), and spaced at 16 in. (40 cm) or 24 in. (60 cm) on center—all common building module dimensions.

The interior sides of the studs are usually covered with an attached facing material. This is often sheetrock, which is a sandwich of gypsum between cardboard facings. Insulation is usually enclosed within the stud wall cavities. A vapor barrier may also be provided; this is a coating or a membrane that minimizes humidity penetration and possible condensation within the wall.

Composite walls are essentially a more substantial form of stud walls. They are constructed of cementitious materials, such as weatherproof sheetrock or precast concrete as an exterior sheathing, and sheetrock as an interior surface finish. *See* GYPSUM; PRECAST CONCRETE.

Prefabricated walls. These are commonly used for curtain-wall construction and are frequently known as prefab walls. Prefabricated walls are usually made of corrugated steel or aluminum sheets, although they are sometimes constructed of fiber-reinforced plastic sheets, fastened to light horizontal beams (girts) spaced several feet apart (Fig. 1). The corrugations usually span vertically to deliver the lateral wind loading to the horizontal girts, which then carry such loading to the primary vertical framing structure. Prefab walls are often made of sandwich construction: outside corrugated sheets, an inside liner of flat or corrugated sheet, and an enclosed insulation are fastened together by screws to form a thin, effective sandwich wall. Sometimes the inner and outer sheets and the enclosed insulation are factory-fabricated by adhesion to form a very strong and stiff composite sandwich wall plank. These usually have tongue-and-groove vertical edges to permit sealed joints when the units are erected at the building site by being fastened to framing girts.

Panel walls. Glass, metal, or ceramic-coated metal panel walls are a common type of curtain wall used in high-rise construction. They are typically assembled as a sandwich by using glass, formed metal, or ceramic-coated metal sheets on the outside, and some form of liner, including possibly masonry, on the inside; insulation is enclosed. Sometimes such walls include masonry and insulation, about 3 ft (0.9 m) high, to form a fire-resistant wainscot, with the upper section of only glass, in one or more plys. Some modern building walls are made predominantly of glass sheets alone, fastened or adhered at the edges by elastomer adhesive/sealants.

Other types. Tilt-up walls are sometimes used for construction efficiency. Here, a wall of any of the various types is fabricated in a horizontal position at ground level, and it is then tilted up and connected at its edges to adjacent tilt-up wall sections. Interior partitions are a lighter form of wall used to separate interior areas in buildings. They are usually nonbearing, constructed as thinner versions of some of the standard wall types; and they are often designed for some resistance to fire and sound. Retaining walls are used as exterior walls of basements to resist outside soil pressure. They are usually of reinforced concrete; however, where the basement depth or exterior soil height is low, the wall may be constructed as a masonry wall. *See* BUILDINGS; RETAINING WALL; STRUCTURAL MATERIALS. Milton Alpern

Bibliography. American Society of Civil Engineers, *Building Code Requirements for Masonry Structures*, 1990; American Society of Civil Engineers, *Minimum Design Loads for Buildings and Other Structures*, 1990; D. K. Ballast, *Architect's Handbook of Construction Detailing*, 1990; C. Beall, *Masonry Design and Detailing for Architects, Engineers and Contractors*, 4th ed., 1997; H. J. Cowan, *Design of Reinforced Concrete Structures*, 2d ed., 1988; K. F. Faherty and T. G. Williamson (eds.), *Wood Engineering and Construction Handbook*, 3d ed., 1998.

Walnut

This name is applied to about a dozen species of large deciduous trees widely distributed over temperate North and South America, southeastern Europe, and central and eastern Asia. The genus (*Juglans*) is characterized by pinnately compound aromatic leaves and chambered or laminate pith (**Fig. 1**). The staminate (male) flowers are borne in unbranched catkins on the previous season's growth, and the pistillate (female) flowers are terminal on the current season's shoots. Pollination is by wind. The shells of the nuts of most species are deeply furrowed or sculptured.

The plants fall into two natural groups, one characterized by a prominent band of hairs on the upper edge of the leaf scar, pointed buds, and usually elongate nuts. The leaf scars of the other group lack the cushion of hairs, and the nuts are spherical or nearly

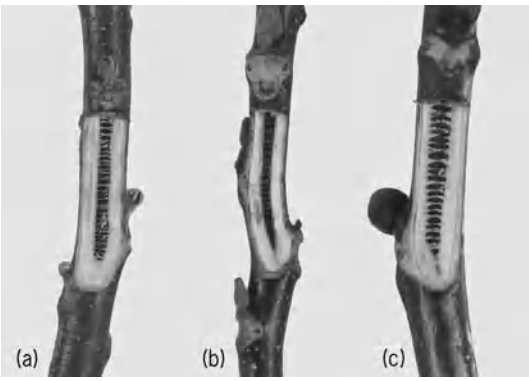

Fig. 1. Pith of (*a*) black walnut (*Juglans nigra*), (*b*) butternut (*J. cinera*), and (*c*) English walnut (*J. regia*).

Fig. 2. Branches showing buds and leaf scars of (*a*) black walnut (*Juglans nigra*), (*b*) butternut (*J. cinera*), and (*c*) English walnut (*J. regia*).

so. Of the six North American species, five are in the latter group and one, the butternut or white walnut (*J. cinera*), in the former (Fig. 1 and **Fig. 2**).

Two species, the black walnut (*J. nigra*) and the Persian or English walnut (*J. regia*), are of primary importance for their timber and nuts. The butternut finds local use in the northeastern United States. The other species are sparingly used as shade trees, as grafting stocks, and as sources of nuts.

The black walnut and other plants in the walnut family produce a toxic substance called juglone (a napthaquinone), which according to present evidence suppresses the growth of many plants when their roots come in contact with walnut roots. Tomatoes, alfalfa, rhododendron, apple, and pine are particularly affected. However, some plants, notably some of the grasses and the black raspberry, appear to be stimulated by the walnut roots. Some recent research relates laminitis in horses to their being bedded down with fresh black walnut wood chips. The chemistry of the injury is not well understood.

Black walnut. The black walnut is a large tree; it can reach a height of over 100 feet (30 m) and a diameter of 3–4 ft (0.9–1.2 m). It is native to the hardwood forests of the Central Mississippi Valley and the Appalachian region of North America.

Walnuts thrive on rich, well-drained, alluvial soils well supplied with moisture and are most abundant in the river valleys and along streams. They have, however, been planted beyond their natural range and succeed on good agricultural soils over a wide area.

The pinnately compound leaves have 15–20 leaflets, and the stamens and pistils are borne separately (monoecious) as in other members of the genus. The drupelike spherical fruit $1^1/_2$–$2^1/_2$ in. (38–64 mm) long is borne singly or in clusters of two or three. It consists of an outer indehiscent fleshy husk about $^1/_2$ in. (13 mm) thick enclosing the hard, rough-shelled nut $1^1/_4$–$1^1/_2$ in. (32–38 mm) in length. The kernels formed from the cotyledons are enclosed in membranous seed coats which turn brown if nuts remain in husks (**Fig. 3**). The dark-brown, strong, durable wood is rated as North America's best for gunstocks, fine furniture, and veneer. The annual demand for high-quality logs exceeds the supply. This is in part due to the export of walnut logs.

Nut production is mostly confined to the millions of trees in the forests and farmsteads. Formerly the nuts were gathered in large quantities, hulled, dried, and cracked for home use, or the operation was a cottage industry. With the advent of cracking machines about 1935, shelling and processing plants developed in the Central Mississippi Valley. Also, there is the beginning of a black walnut orchard industry which aims at production of both timber and nuts. Under orchard conditions, the culture and harvesting of the black walnut crop will undoubtedly be mechanized as with the Persian or English walnut.

The average kernel yield of seedling black walnuts in shell is 10–12%. Since the early 1900s the Northern Nut Growers Association has located and propagated many varieties (cultivars) with 20–30% kernel. Some of the more important of these are the Thomas, Burns, Emma K, Victoria, Stambaugh, Snyder, and Sparrow. There is a shortage of grafted trees at the present time. Although these cultivars are much better than the average wild seedlings, they have not been good enough to be commercially profitable or raised in sufficient quantity to establish a special market. They are, however, planted for home use.

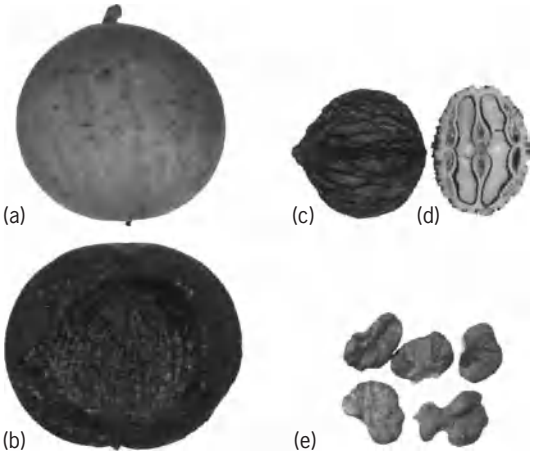

Fig. 3. Black walnut (*Juglans nigra*). (*a*) Mature fruit. (*b*) Nut with half of hull cut away. (*c*) Nut without husk. (*d*) Nut cut transversely to show shell structure. (*e*) Kernels in quarters.

Black walnut kernels were an important source of food for the American Indians and the early settlers in the United States. Now they are extensively used in ice cream, candy, and baked goods. Kernels are graded at the processing plant according to size of the pieces. In the cracking process part of the kernels are recovered as whole quarters and, with the fragments larger than $^1/_6$ in. (4 mm), are sold in plastic bags, glass jars, or metal cans. The dense hard shells of the nuts are ground and used for drilling mud in oil fields, for abrasives for polishing metal castings, and for paint fillers and other industrial uses.

English walnut. The most important of the nut-bearing trees belonging to the genus *Juglans* is the English walnut. Native to central Asia and Asia Minor, the tree was distributed widely in ancient times throughout the temperate and subtropical climatic zones of the Old World. It is also called Persian walnut, which is considered a more appropriate name. The large, round-headed, long-lived trees produce the circassian walnut lumber of commerce that is valued for gunstocks, furniture, veneer, and paneling. In France, the Balkans, and elsewhere the trees are raised for both nuts and lumber.

The leaves are pinnately compound with five to nine entire leaflets, variable in shape and size. The trees are monoecious, with the staminate flowers borne in catkins and the nutlets terminal on the current season's growth. The oval fruit, $1^1/_2$–2 in. (38–51 mm) long, in good varieties, consists of a fleshy outer hull about $^1/_4$ in. (6 mm) thick which splits irregularly at maturity, freeing the nut (**Fig. 4**). The wild-type nuts of central Asian origin are small and hard-shelled.

Through centuries of selection, large thin-shelled types have been developed, varying widely in their adaptation to different climates. Thus it is important in any commercial venture to choose varieties (cultivars) for their characteristics related to hardiness, total heat requirements, dormancy, bearing habit, and other factors. The commercial Persian walnut industry in the United States is centered in northern and central California, where the climate resembles that of the Mediterannean region where walnuts have thrived for centuries. Earlier plantings in Oregon and Washington have been reduced because of losses from winter injury. Planting has also shifted from southern California because of low yields related to failure of the trees to break dormancy, and to damage to the nuts from sunburn. The bulk of the California walnut crop is produced from orchards of older varieties such as Franquette, Hartley, Ashley, Eureka, Chico, Amigo, and Payne.

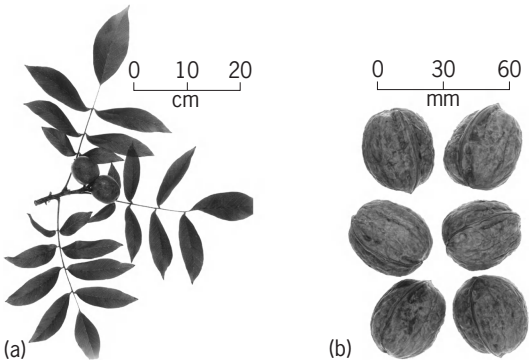

Fig. 4. English walnut (*Juglans regia*). (*a*) Twig with leaves and fruit. (*b*) Hulled nuts.

New varieties produced by breeding give greatly increased yields related to their having a high percentage of lateral buds that produce nuts. Typically in Persian walnuts pistillate flowers arise from the terminal or subterminal buds. The newer, heavy-yielding cultivars may fruit from 80% of lateral buds. Some of the cultivars with this characteristic are the Payne, Ashley, Chandler, Howard, and Sunland. Because the pistillate and staminate flowers are borne separately (monoecious) and may mature at different times (dichogamy), cultivars should be planted that will shed the wind-borne pollen at the time the pistillate flowers are receptive, either by planting cultivars whose flowers mature at the same time or by planting combinations of cultivars that shed pollen at the right time to pollinate all receptive pistils. The cultivars Amigo and Chico are used in commercial orchards particularly for their value as pollenizers. Collecting and storing pollen for artificial release at the right time is a practical procedure.

Production in the world other than in the United States and western Europe is mostly from seedling trees which receive little care. In the better orchards of California, grafted trees of high-yielding varieties are given modern highly mechanized intensive care. Propagation of the cultivars is by patch budding or whip grafting. A variety of stocks have been used. Persian walnut stock is the most compatible, but lost favor earlier because of susceptibility to root rot. California black walnut (*J. hindsii*) and hybrid (Paradox) stocks have been used, but show graft union failure from blackline disease as the trees age. The use of *J. regia* stocks is now increasing.

For good production the planting site should have good air drainage to prevent frost damage and a deep, well-drained, neutral or slightly alkaline soil. Frequent fog or rain may make disease control difficult.

Planting distance in the older orchards has been 50 ft (15 m) for mature trees. In modern orchards using early bearing cultivars which give heavy crops from lateral buds, recommended distance is 25–30 ft (7.5–9 m) with the option of removing half the trees when they begin to crowd, which will be in about 12–15 years.

Pruning is done to develop a strong framework which will support the crop. The young tree is trained with a central leader until four or five well-spaced permanent branches are developed. Then the center is opened up to form a modified leader pattern. Cultivars that bear nuts only on terminal buds should not be cut back to restrain growth. Those bearing nuts from lateral buds may be cut back without seriously affecting yield.

Good cultural practice requires weed control, cover crops, fertilization, and irrigation. An ample water supply is essential to good production. This

must be related to the highly mechanized harvesting procedures.

The more important insect pests are various husk flies, weevils, curculio, and codling moth. Troublesome diseases are walnut anthracnose, walnut blight, and various root rots. Some types of witches'-broom are related to zinc deficiency, but recent research indicates that a microplasma may be involved for which there is now no control. Appropriate pest and disease controls are essential to producing a commercial crop.

Yields of walnuts vary widely in different orchards depending on differences in site, soil, and many other factors. Fluctuations occur in overall production from year to year depending on seasonal conditions particularly related to frost, rainfall, and alternate bearing cycle. Yields of 4 tons in shell per acre (9 metric tons/hectare) have been reported, but 2–3 tons (4.5–6.75 metric tons) is more likely, and the 5-year average of all orchards is about 1 ton/acre (2.25 metric tons/hectare).

Harvesting and processing in most of the producing countries is done with local hand labor. In California, however, operations are highly mechanized. Before harvest, the ground under the trees is cleaned and leveled. Machines shake the nuts from the trees and sweep them into windrows, from which they are picked up mechanically and taken in bulk to the processing plant, where they are hulled, washed, and dried. About 25% of the crop to be marketed in the shells is sized and bleached. The others are machine-shelled, and the kernels are graded for color by electronic machines, sized, and packaged for the retail trade. The Diamond Walnut Growers, Inc., handles over 50% of the crop and determines quality and grade. The Walnut Control Board decides the amounts available for export and the proportion to be shelled.

California and Oregon produce all commercial quantities of walnuts in the United States. Other producing countries are France, Italy, and Turkey. Russia and China also produce commercial quantities of walnuts.

More than half of the English walnuts are shelled, and the kernels are used in a great variety of ways in baking and confectionery, and are also salted. The others are sold in shell to the consumer. In foreign countries many of the nuts are used locally as food.

Carpathian walnut. There is considerable interest among producers in the central and eastern United States in growing the Carpathian walnut. This strain of Persian walnuts was introduced from the Carpathian Mountains in the 1920s by Paul Crath. Prior to that time the attempts to grow Persian walnuts in the eastern United States were with varieties mostly of Mediterranean origin which failed because the trees were killed at temperatures of about 15° (−9°C). The Carpathian introductions, when dormant, withstood temperatures of −35°F (−37°C) and attracted much attention. Seeds and seedling trees were distributed widely through eastern North American from Ontario, Canada, southward into the northern parts of the Southern states. From these trees and others of outstanding hardiness, selections have been made, named, and propagated. Some of the more promising cultivars are Hansen, Colby, Broadview, Metcalfe, McKinster, and Somers. There is a concerted effort by the Northern Nut Growers Association to secure hardy varieties adapted to the central and northeastern United States from all sources. The term Carpathian walnut is now applied to these regardless of origin. Most of the trees offered by nurseries are seedlings, many of which bear nuts of poor quality, have a late bearing age, and fail to produce crops. Causes of failure may be late spring frost, lack of pollination, and pest and disease damage. There is no commercial Persian walnut industry in the eastern United States, but many trees have been planted by amateurs in a search for better varieties. There is a limited market for seed nuts; otherwise production is absorbed locally.

Butternut. The butternut, or white walnut, is the hardiest of the American species of *Juglans*, its range extending from Maine and New Brunswick, westward to Ontario, south to Arkansas and the mountains of Georgia. The trees are adapted to upland soils and in the northeastern United States are one of the few nut-bearing species. Distinguishing characters are the pointed buds and the cushion of hairs above the leaf scar (Figs. 1 and 2). The nuts, which are borne in clusters of two or three, are ellipsoid and pointed and are enclosed in a tight indehiscent husk which is covered with sticky hairs. The shell is deeply sculptured with jagged ridges.

The butternut, as compared with the black walnut, has a shallow root system and a more spreading crown. Trees are often short-lived because of defoliation by fungus disease. The wood is lighter-colored and less dense than the black walnut, and is used for interior finish and furniture. The husks of the nuts are formerly used in dying cloth greenish yellow.

Butternuts are used locally for their highly flavored oily kernels. Nuts from most of the wild trees are difficult to crack. Some progress has been made in selecting clones with good cracking quality.

Laurence H. MacDaniels

Diseases. Black walnut, English walnut, and butternut are susceptible to a number of leaf, stem, and root diseases. The seriousness of specific diseases varies greatly with tree species. Walnut anthracnose, caused by the fungus *Gnomonia leptostyla*, with the imperfect stage *Marssoninna juglandis*, is the most serious foliar disease of black walnut. The anthracnose fungus also attacks butternut, English walnut, Hinds walnut, and California walnut. Symptoms appear on the leaflets as small, circular to irregularly shaped, brown to black areas, often surrounded with a yellowish halo (**Fig. 5**). Two or more spots may later merge to form a large blighted area. Affected leaflets may turn brown and drop early. Disease losses can be reduced by the application of nitrogen fertilizers and protective sprays of the fungicide benomyl.

Two other foliar pathogens of black walnut cause diseases of lesser importance. The fungus *Microstroma juglandis* causes a white mold or a yellow leaf blotch. The fungus *Cristulariella pyramidalis*

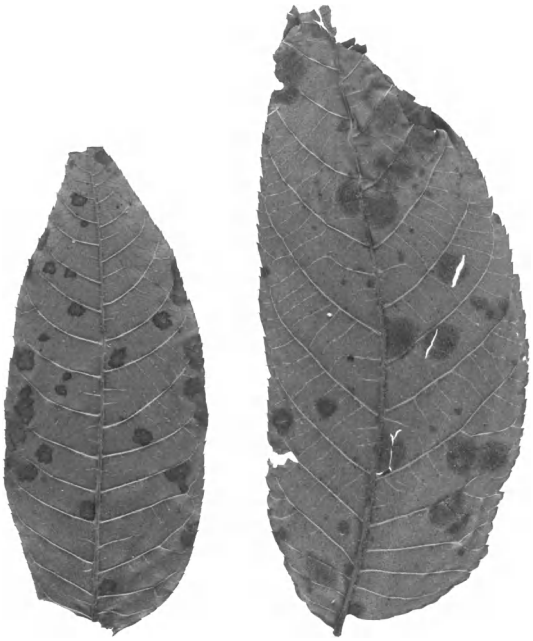

Fig. 5. Walnut leaves showing anthracnose symptoms.

has caused severe premature defoliation of trees in Illinois and Ohio.

Bacterial blight, caused by *Xanthomonas juglandis*, is especially severe on English walnut. Buds, leaves, and shoots are attacked, but the disease is most serious and destructive when quality and yield of nuts are reduced. Small reddish-brown spots occur on the leaves, and black, slightly depressed spots are observed on stems and nuts. The blight can be controlled with timely and thorough applications of copper compounds.

A canker disease caused by the fungus *Sirococcus* sp. is causing a decline of butternut throughout much of its range. The canker does not appear to pose a serious threat to black walnut.

Phythophthora-induced root and crown rots have seriously curtailed seedling production of black and English walnut. Infected seedlings occur commonly in poorly drained portions of the nursery beds. The damage continues during winter storage. Some control has been obtained with fungicide drenches and dips. Species of *Cylindrocladium* also may be serious root rot pathogens of black walnut seedlings.

Not all defoliation of walnut trees is due to pathogenic organisms. The leaves on trees growing on poor sites or suffering from drought may turn yellow and drop during August or September as a direct result of environmental stress factors. Even under normal conditions, walnut trees drop their leaves earlier in the autumn than do many other species of trees. *See* PLANT PATHOLOGY. Dan Neely

Bibliography. R. A. Jaynes (ed.), *Nut Tree Culture in North America*, Northern Nut Growers Ass., 1979; D. Neely, R. Phares, and B. Weber, Cristulariella leaf spot associated with defoliation of black walnut plantations in Illinois, *Plant Dis. Rep.*, 60:587–590, 1976; B. Thompson, *Black Walnut for Profit*, 1976; U.S. Forest Service, Walnut insects and diseases, *Workshop Proceedings, June 13–14, 1978, Carbondale, Illinois*, 1979; J. G. Woodroof, *Tree Nuts*, 1979.

Warm-air heating system

In a general sense, a heating system which circulates warm air. Under this definition both a room-heating stove and a steam blast coil circulate warm air. Strictly speaking, however, a warm-air system is one containing a direct-fired furnace surrounded by a bonnet through which air circulates to be heated (see **illus.**).

When air circulation is obtained by natural gravity action, the system is referred to as a gravity warm-air system. If positive air circulation is provided by means of a centrifugal fan (referred to in the industry as a blower), the system is referred to as a forced-air heating system.

Direct-fired furnaces are available for burning of solid, liquid, or gaseous fuels, although in recent years oil and gas fuels have been most commonly used. Furnaces have also been designed which have air circulating over electrical resistance heaters. A completely equipped furnace-blower package consists of furnace, burner, bonnet, blower, filter, and accessories. The furnace shell is usually of welded steel. The burner supplies a positively metered rate of fuel and a proportionate amount of air for combustion. A casing, or jacket, encloses the furnace and provides a passage for the air to be circulated over the heated furnace shell. The casing is insulated and contains openings to which return-air and warm-air ducts can be attached. The blower circulates air against static pressures, usually less than 1 in. (2.5 cm) water gage. The air filter removes dust particles from the circulating air. The most common type of air filter is composed of 1- to 2-in.-thick (2.5- to 5.0-cm) fibrous matting, although electrostatic precipitators are sometimes used. *See* GAS FURNACE; OIL BURNER.

Accessories to assure effective operation include automatic electrical controls for operation of burner and blower and safety control devices for protection against (1) faulty ignition of burner and (2) excessive air temperatures.

Ratings of warm-air furnaces are established from tests made in laboratories under industry-specified

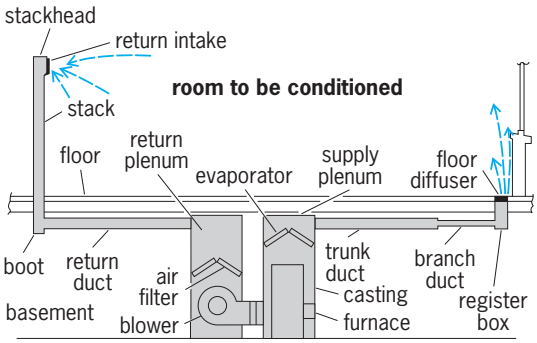

Air passage in a warm-air duct system. (*After S. Konzo, J. R. Carroll, and H. D. Bareither, Winter Air Conditioning, Industrial Press, 1958*)

conditions. The tests commonly include heat-input rate, bonnet capacity, and register delivery. Heat-input rate is the heat released inside the furnace by the combustion of fuel, in Btu/h. Bonnet capacity refers to the heat transferred to the circulating air, in Btu/h. Register delivery is the estimated heat available at the registers in the room after allowance for heat loss from the ducts has been made, in Btu/h.

The recommended method for selection of a furnace is to estimate the total heat loss from the structure under design weather conditions, including the losses through the floor and from the basement, and to choose a furnace whose bonnet capacity rating is equal to, or greater than, the total design heat loss.

The complete forced-air heating system consists of the furnace-blower package unit; the return-air intake, or grille, together with return-air ducts leading from the grille to the return-air plenum chamber at the furnace; and the supply trunk duct and branch ducts leading to the registers located in the different spaces to be heated.

The forced-air system in recent years has no longer been confined to residential installations. The extreme flexibility of the system, as well as the diversity of furnace types, has resulted in widespread use of the forced-air furnace installations in the following types of installations, both domestic and commercial: residences with basement, crawl space, or with concrete floor slab; apartment buildings with individual furnaces for each apartment; churches with several furnaces for different zones of the building; commercial buildings with summer-winter arrangements; and industrial buildings with individual furnace-duct systems in each zone. *See* COMFORT HEATING. Seichi Konzo

Bibliography. American Society of Heating, Refrigerating, and Air-Conditioning Engineers, *Equipment*, 1992, *Fundamentals*, 1994; F. C. McQuiston and J. D. Parker, *Heating, Ventilating, and Air Conditioning: Analysis and Design*, 5th ed., 2000.

Washer

A flattened, ring-shaped device used to improve the tightness of a screw fastener. Three types of washer are in common use: plain, spring-lock, and antiturn (tooth-lock) washers. Standard plain washers are used to protect a part from damage or to provide for a wider distribution of the load. Because a plain washer will not prevent a nut from turning, a locking-type washer should be used to prevent a bolt or nut from loosening under vibration (see **illus.**). For industrial applications, spring-lock washers are intended to compensate for possible loosening between assembled parts and to facilitate assembly and disassembly. Lock washers create a continuous pressure between the parts and the fastener. The antiturn-type washers may be externally serrated, internally serrated, or both. The bent teeth bite into the bearing surface to prevent the nut from turning and the fastening from loosening under vibration. To speed up assembly, a variety of permanent preassembled bolt-and-washer and nut-and-washer combinations are available. *See* BOLT; NUT (ENGINEERING); SCREW FASTENER. Warren J. Luzadder

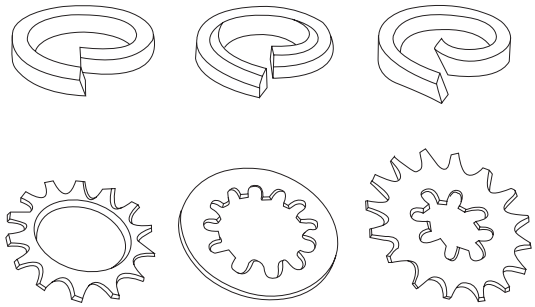

Lock washers. (*After W. J. Luzadder, Fundamentals of Engineering Drawing, 6th ed., Prentice-Hall, 1971*)

Wastewater reuse

The use of treated wastewater effluent for beneficial purposes such as irrigation, industrial processes, and municipal water supply. Several developments have prompted wastewater reuse, including shortages of freshwater, stringent requirements for wastewater effluent quality, and advancements in treatment technology.

Shortage of freshwater. Freshwater is used for municipal, industrial, and irrigation purposes. Municipal and industrial water uses are nonconsumptive, whereas irrigation is consumptive, since most of this water is lost by evaporation and transpiration and is not available for reuse. Worldwide, the total cultivated area is about 1 billion hectares, of which only about one-third is irrigated. To increase food production, more irrigation water will be required, especially in developing countries.

The Earth's surface continually receives huge amounts of freshwater as precipitation. However, most precipitation is not available for human use due to its uneven distribution. The imbalance between the freshwater supply and demand has become more serious in recent years, with the available supply diminishing due to pollution and rapidly increasing demand. The World Health Organization estimates that within the next 50 years more than 40% of the world's population will live in countries facing water stress or water scarcity.

Requirements for wastewater effluent quality. A traditional sanitary engineering system deals with the municipal water supply and wastewater disposal (**Fig. 1**). Cities usually obtain their freshwater supply from surface or ground-water sources. After use, about 80% of the freshwater supply becomes wastewater, which is collected by a sewerage system and conveyed to a treatment plant. The wastewater is treated and then discharged as effluent into receiving water such as a river, estuary, coastal water, or ground-water aquifer.

Prior to 1970, most cities in the United States and other countries provided no more than primary wastewater treatment, which is a sedimentation

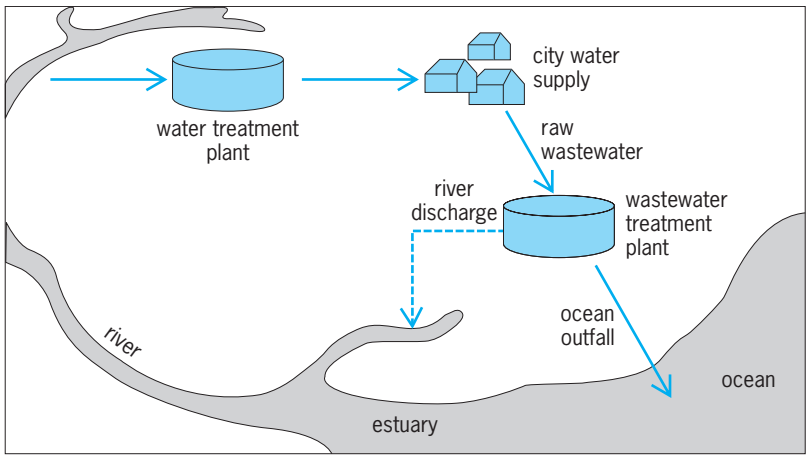

Fig. 1. Traditional water and wastewater engineering systems.

process that removes suspended solids. To restore and maintain the chemical, physical, and biological integrity of the nation's waters, Congress enacted the 1972 Federal Water Pollution Control Act Amendments (now called the Clean Water Act), which established nationwide minimum treatment requirements for all wastewater. For municipal wastewater discharge, the minimum is secondary treatment, which removes 85% of the biochemical oxygen demand (BOD) and total suspended solids. BOD is a measurement of oxygen-demanding organic wastes. Secondary treatment can be achieved by biochemical decomposition of organic waste materials, followed by further sedimentation. For industrial wastewater discharge, the minimum treatment is defined as the best practical control technology currently available. In situations in which these minimum treatment levels are not sufficient, the act requires additional treatment. *See* SEDIMENTATION (INDUSTRY); SEWAGE SOLIDS.

Water quality standards and required treatment processes. Wastewater reuse is classified as nonpotable and potable. Nonpotable uses of wastewater effluent include irrigation, industrial process or cooling water, and recreational impoundments (such as lakes or reservoirs for public use). Potable uses of high-quality wastewater effluent can be achieved by sending it back to water supply reservoirs, to potable ground-water aquifers, or directly to water treatment plants.

There are no federal regulations governing wastewater reuse in the United States. All regulations are enforced at the state level. As the U.S. Clean Water Act requires, municipal wastewater must receive at least secondary treatment; therefore, the reuse potential of wastewater depends on levels of postsecondary (or advanced/tertiary) treatment imposed. The Hawaii Department of Health has guidelines for three types of reclaimed water and the treatment and limitations of each. The highest-quality reclaimed water, designated R-1 water, is wastewater that has received secondary treatment followed by filtration and intense disinfection. R-1 water is deemed acceptable for public contact and can be used for nearly all nonpotable purposes. Next in quality is R-2 water, which is wastewater that has received secondary treatment followed by disinfection. The lowest-quality reclaimed water, designated R-3 water, is wastewater that has received secondary treatment but no disinfection. The reuse of R-2 and R-3 waters is restricted. *See* FILTRATION; SEWAGE TREATMENT.

Fecal coliform bacteria found in the intestinal tract of warm-blooded animals are used as an indicator of microbiological purity. The amount of fecal coliform in a water sample is measured in colony-forming units (CFU). Most states set a maximum of 23 CFU/100 milliliters for unrestricted irrigation use or irrigation of edible crops, sports fields, and public parks. California and Arizona set a maximum of 2.2 CFU/100 ml for higher-quality reclaimed water that can be used for recreational impoundment.

Artificial ground-water recharge of reclaimed water into potable aquifers can be accomplished by using injection wells or by surface spreading. The reclaimed water must meet drinking water standards if it is recharged into potable aquifers, and it must meet drinking water standards after percolation through the vadose zone if it is recharged by surface spreading. The vadose zone (or unsaturated zone) occurs immediately below the land surface, where soil pores (or interconnected openings) may contain either air or water.

Heavy metals and other trace toxic substances can enter municipal wastewater systems that receive industrial discharge. In these cases, special monitoring and advanced treatment of wastewater effluent before reuse are necessary. Advanced treatment methods, such as carbon adsorption, precipitation, and pressure-driven membrane processes, are used to further remove dissolved solids or salts, nutrients, and organic and inorganic chemicals. *See* ACTIVATED CARBON; PRECIPITATION (CHEMISTRY); WATER TREATMENT.

Pressure-driven membrane technology. Pressure-driven membranes are divided into four groups based on membrane pore size: microfiltration (MF), ultrafiltration (UF), nanofiltration (NF), and reverse osmosis (RO). The required pressure for each membrane type is inversely proportional to the pore size. Microfiltration is used to remove the largest particles and requires the least pressure, whereas the reverse osmosis is used to remove the smallest particles and requires the greatest pressure. Microfiltration and ultrafiltration use a sieve process, in which particles are removed because the membrane openings are smaller than the particles' size. When two solutions of salts are separated by a semipermeable membrane, osmosis allows water molecules to pass through the membrane from the solution of lower salt concentration into the solution of higher salt concentration. The reverse osmosis process, which reverses the direction of water, can be accomplished by applying to the higher-salt-concentration solution a pressure that is higher than the natural osmotic pressure. Nanofiltraton uses the sieve and reverse osmosis processes. *See* MEMBRANE SEPARATIONS; OSMOSIS; ULTRAFILTRATION.

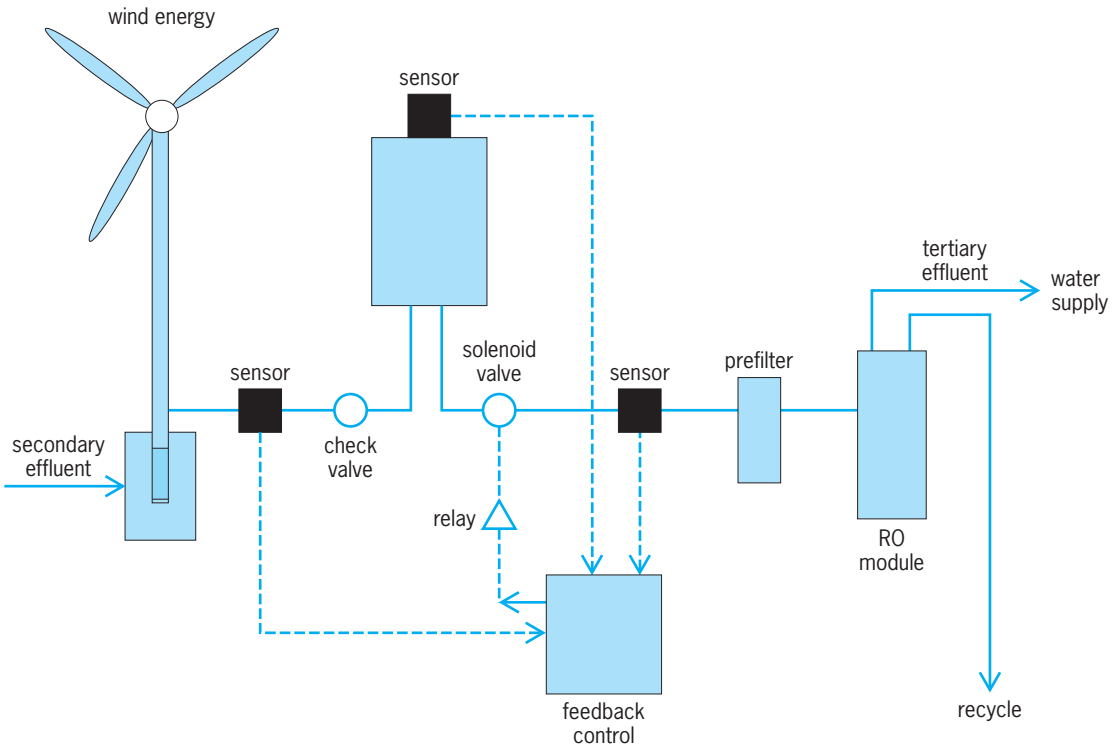

Fig. 2. Wastewater-reuse system with a wind-driven reverse osmosis process.

Reverse osmosis was developed in the 1960s with the advent of the asymmetric cellulose acetate membrane. Since then, its popularity has steadily risen because of advances made in membrane technology with the use of different materials and membrane configurations. Progress has also been made in overcoming membrane fouling and extending the life of the membrane.

Pressure-driven membrane technology is the key to expanding wastewater reuse in the future. Pressure-driven membranes can provide effective disinfection as well as total dissolved solids (TDS) removal. Pressure-driven membranes are ideal for postsecondary treatment of municipal wastewater in coastal areas. In these areas, brackish water or seawater enters the sewer infrastructure due to infiltration and inflow, causing high salinity or TDS in the wastewater flow into treatment plants. Recently, a new treatment technology called membrane bioreactor (MBR) was developed by combining micro- or ultrafiltration membranes with secondary or biological wastewater treatment processes. The city of San Diego is developing a plan to use reverse osmosis, along with a few pretreatment and posttreatment processes, to reclaim and purify its wastewater effluent before returning it to the city's water supply reservoir. Water management authorities in the Netherlands are also studying the use of membrane technology in wastewater treatment.

The high amount of energy required to create and maintain feed water pressure is a limiting factor for expanding the use of pressure-driven membranes in water and wastewater treatment systems. To address this problem, new technology is being developed that integrates the use of natural energy with pressure-driven membranes. A prototype wastewater reuse system with a wind-driven reverse osmosis process was constructed by University of Hawaii researchers on Coconut Island off the windward coast of Oahu. This system consists of a 30-ft-tall (10-m) multivaned windmill, an ultra-low-pressure RO membrane, a flow/pressure stabilizer, a prefilter, and a feedback control mechanism (**Fig. 2**). Under a moderate wind speed of 11 mi/h (5 m/s), this system can process a flow of 5000 gallons/day (19 m^3/day) and reduce the TDS content from 3100 mg/L to 50 mg/L. Another new technology is ocean-wave-driven RO desalination, which was successfully tested in the Caribbean islands by University of Delaware researchers. *See* WATER DESALINATION; WIND POWER.

As the shortage of freshwater becomes more serious and the quality of wastewater effluent improves, reuse of municipal wastewater will become one of the most attractive water resource management alternatives. Advances in health risk assessment will make treated wastewater a widely acceptable freshwater source. The field of water and wastewater engineering is undergoing rapid changes. It is expected that, through proper source control and treatment, water and wastewater will be indistinguishable. *See* WATER SUPPLY ENGINEERING.

Clark C. K. Liu

Bibliography. R. Aertgeerts and A. Angelakis (eds.), *State of the Art Report Health Risks in Aquifer Recharge Using Reclaimed Water*, World Health Organization, Regional Office for Europe, Copenhagen, 2003; H. Bouwer, Role of groundwater recharge in treatment and storage of wastewater

for reuse, *Water Sci. Technol.*, 24(9):295–302, 1991; Y. Kamiyama et al., New thin-film composite reverse osmosis membranes and spiral wound modules, *Desalination*, 51:79–92, 1984; T. Richardson and R. Trussel, Taking the plunge, *Civ. Eng.*, 67(9):42–45, 1997; D. Seckler et al., *World Water Demand and Supply, 1990 to 2025: Scenarios and Issues*, Res. Rep. 19, International Water Management Institute, Colombo, Sri Lanka, 1998; U.S. Environmental Protection Agency, *Municipal Wastewater Reuse: Selected Readings on Water Reuse*, EPA430/09-91-002, Washington, DC, 1991.

Watch

A portable instrument that measures time. The direct forerunner of the watch was the pendulum clock, invented by Christian Huygens in 1656.

Principles of operation. The creation of the pendulum clock began with the insights of Galileo in the fifteenth century. Galileo, attending services at the Cathedral of Pisa, noticed that when a chandelier was caused to swing back and forth the long or short swings were completed in the same amount of time. He knew that this was because during the long swing the increased gravitational pull made the chandelier swing faster and during the short swing the decreased gravitational pull made the chandelier swing slower. The result was that the swings—long or short—took the same amount of time. Galileo verified this by his timing apparatus: his pulse. Christian Huygens used this principal "of equal time for long or short swing" to invent the first pendulum clock in 1656. For this clock, power was supplied not by gravity but by a wound spring; whether the spring is tightly or loosely wound, the pendulum (like the chandelier) swings in equal time. *See* PENDULUM.

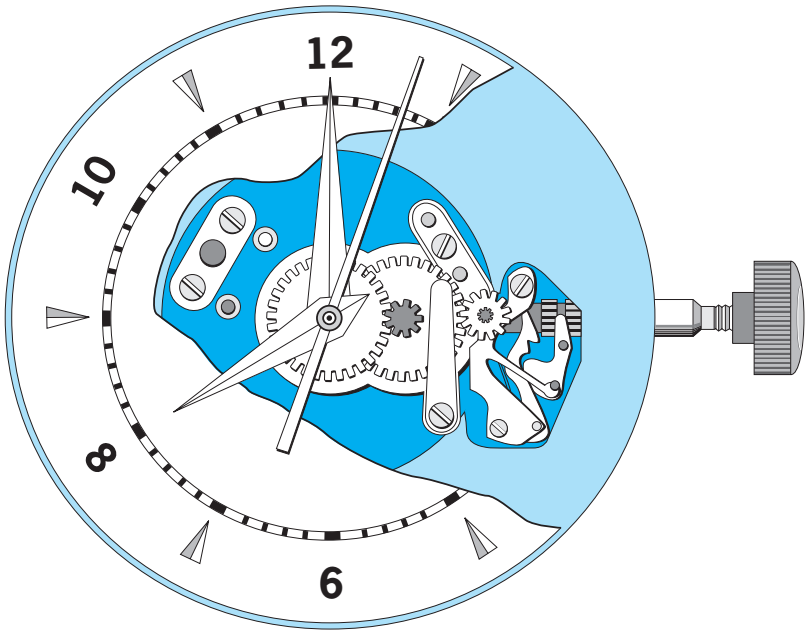

Cutaway view of classic mechanical watch.

The principles discovered by Galileo and Huygens are the precursors of all modern timepieces, which operate by using the following components: a part that swings back and forth in equal intervals called an oscillator (for example, a pendulum, balance wheel, tuning fork, or quartz crystal); a power source that energizes the swings (for example, a mainspring, battery, or capacitors); and a system that counts the swings (see **illustration**). The earliest watches consisted of a "swinger" (basically a balance wheel with an attached hairspring) as an oscillator; a mainspring as a power source; and a dial and hands to count the swings. This type of model is still manufactured and is comparatively expensive. Some are called chronometers and are precise timekeepers. But most watches today are quartz watches. *See* CLOCK; TUNING FORK.

Evolution of the quartz watch. Quartz watches are the result of the evolution of the clock mechanism. The first step was to replace the mechanical power (the mainspring) with electrical power (the battery). The next major step was to replace the oscillator (the balance wheel, with its attached hairspring) with a tuning fork and an electronic circuit; although this mechanism was eventually discontinued, the electronic circuit was retained in later watches. The next step, which was truly revolutionary, was to replace the tuning fork with a quartz crystal. Whereas a pendulum swings once per second, and a balance wheel oscillates 5 times a second, the standard quartz crystal oscillates 32,768 times a second.

Quartz crystals. The advent of the quartz crystal is the most revolutionary advance in watchmaking in the twentieth century. The properties of the quartz crystal were discovered by Pierre and Jacques Curie. They experimented with natural crystals and discovered piezoelectricity (the principle that when crystals are stressed they generate electricity). The most responsive was the quartz crystal. Since the quartz crystal produced extremely stable oscillations, it was applied to frequency control units in radio broadcasting. Then, in 1928, Bell Laboratories created the first quartz crystal clock. But before a quartz watch could be produced an integrated circuit had to be created. *See* PIEZOELECTRICITY; QUARTZ CLOCK.

Integrated circuits. The integrated circuit, or microchip, is one of the most revolutionary inventions of the twentieth century. It is used in every type of electronic device—from computers to antimissile systems. In the past, electronic circuits consisted of vacuum tubes, diodes, transistors, capacitors, and resistors, which were connected by wires and fastened to a baseboard. Depending on the number of elements involved, such a circuit could require an entire room or building. By the process of miniaturization the electronic circuit was transferred into the integrated circuit (microchip). In a watch, the microchip reduces the frequency of the quartz vibrations (32,768 times per second) to standard intervals. The microchip was patented in 1959, and in 1967 Japan and Switzerland simultaneously introduced quartz watches at the Basel Fair. *See* INTEGRATED CIRCUITS.

Stepping motors. In a quartz watch the quartz crystal requires an intermediary to reduce the oscillations to activate the setting system for dial and hands. This is accomplished with a stepping motor, which is energized by the integrated circuit (microchip). In the earliest quartz watch models "slave" oscillators (for example, balance wheels or tuning forks), activated by the microchip, were used to reduce the oscillations. However, eventually all manufactures began to use stepping motors. *See* STEPPING MOTOR.

Solid-state technology. After the quartz watch with its stepping motor was invented another historic development took place—the solid-state quartz, a watch that displayed the time with numbers (digits) rather than with hands and a dial. Whereas the hands on the dial of a stepping-motor quartz watch are moved by a set of gears attached to a setting mechanism, and an integrated circuit divides the quartz oscillations to one pulse per second (or fraction thereof), in a solid-state watch the time display is formed strictly electronically and inert matter forms the digits on the display screen. The integrated circuit controls the voltages and currents that activate the display.

With solid-state technology the revolution from a mechanical movement to a quartz movement was finally completed. No trace of the original mechanical movement remains in quartz watches, which rely solely upon electronic impulses. Freed from moving parts the full potential of the integrated circuit was realized. The integrated circuit in the solid-state watch is actually a miniature computer.

Liquid-crystal technology. An important advance in solid-state watches resulted from the discovery that liquid crystals could be rearranged and form configurations by electrical excitation. Light-emitting diodes (LEDs) were replaced by liquid crystal displays (LCDs), which are now used in all digital-reading solid-state quartz watches. *See* ELECTRONIC DISPLAY; LIGHT-EMITTING DIODE; LIQUID CRYSTALS.

Further advances. Since the quartz watch is also a miniaturized computer, it may contain many additional features. For example, some watches can import information such as appointments, birthdays, and addresses from a personal computer and display them on their screen. For the vision-impaired, watches can be made to "talk." For the outdoor sportsman, a watch exists that provides readouts of height, depth, and barometric pressure measurements. Some watches can display caloric consumption, based on exercise, if programmed with the wearer's age, weight, height, and sex; and there is a watch in which blood pressure and pulse rates appear on the screen and which flashes a warning across the screen if the wearer's blood pressure is too high. The face of a watch, whether analog (dials and hands) or digital, can in some cases be made to light up in the dark, a feature which has saved lives in emergency situations.

One notable advance in quartz watch technology started with a watch that picked up time signals from an astronomical observatory and thereby corrected the oscillations of the quartz crystal. This was accomplished by adding a miniature receiving set and antennae to the watch. This led to the development of a watch that is used as an entrance opener in over 300 resorts worldwide—it contains a computer chip which stores the code numbers of the entrances and an antenna which transmits the code to the entrance gate.

There are watches that have a camera attached. When a button is pressed the display screen becomes a viewfinder, and another button snaps the picture; the pictures can then be transferred to a computer. There are watches that receive and play music from the Internet or a compact disk; phone watches that contain a built-in microphone for outgoing calls and a built-in speaker for incoming calls; and a TV/phone watch that can receive television programs and also contains a phone (if a call comes in when the TV is on the watch automatically turns on the phone function). The advances possible in watch technology are as boundless as the capabilities of computer technology. *See* DIGITAL COMPUTER; TELEPHONE.

Benjamin Matz

Bibliography. T. Hyltin, *The Digital Electronic Watch*, 1978; B. Matz, *The History and Development of the Quartz Watch*, 1999; T. R. Reid, *The Chip: How Two Americans Invented the Microchip and Launched a Revolution*, 1984; H. Tait, *Clocks and Watches*, 1983.

Water

The chemical compound with two atoms of hydrogen and one atom of oxygen in each of its molecules. It is formed by the direct reaction (1) of hydrogen

$$2H_2 + O_2 \rightarrow 2H_2O \qquad (1)$$

with oxygen. The other compound of hydrogen and oxygen, hydrogen peroxide, readily decomposes to form water, reaction (2). Water also is formed in the

$$2H_2O_2 \rightarrow 2H_2O + O_2 \qquad (2)$$

combustion of hydrogen-containing compounds, in the pyrolysis of hydrates, and in animal metabolism. Some properties of water are given in the **table**.

Gaseous state. Water vapor consists of water molecules which move nearly independently of each other. The relative positions of the atoms in a water

Properties of water

Property	Value
Freezing point	0°C
Density of ice, 0°C	0.92 g/cm^3
Density of water, 0°C	1.00 g/cm^3
Heat of fusion	80 cal/g (335 J/g)
Boiling point	100°C
Heat of vaporization	540 cal/g (2260 J/g)
Critical temperature	347°C
Critical pressure	217 atm (22.0 MPa)
Specific electrical conductivity at 25°C	1×10^{-7}/ohm-cm
Dielectric constant, 25°C	78

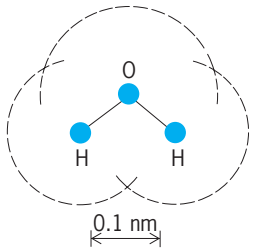

Fig. 1. Water molecule.

molecule are shown in **Fig. 1**. The dotted circles show the effective sizes of the isolated atoms. The atoms are held together in the molecule by chemical bonds which are very polar, the hydrogen end of each bond being electrically positive relative to the oxygen. When two molecules near each other are suitably oriented, the positive hydrogen of one molecule attracts the negative oxygen of the other, and while in this orientation, the repulsion of the like charges is comparatively small. The net attraction is strong enough to hold the molecules together in many circumstances and is called a hydrogen bond. *See* CHEMICAL BONDING; ELECTRONEGATIVITY; GAS; VALENCE.

When heated above 1200°C (2200°F), water vapor dissociates appreciably to form hydrogen atoms and hydroxyl free radicals, reaction (3). These prod-

$$H_2O \rightarrow H + OH \tag{3}$$

ucts recombine completely to form water when the temperature is lowered. Water vapor also undergoes most of the chemical reactions of liquid water and, at very high concentrations, even shows some of the unusual solvent properties of liquid water. Above 374°C (705°F), water vapor may be compressed to any density without liquefying, and at a density as high as 0.4 g/cm^3, it can dissolve appreciable quantities of salt. These conditions of high temperature and pressure are found in efficient steam power plants. *See* HYDROGEN BOND.

Solid state. Ordinary ice consists of water molecules joined together by hydrogen bonds in a regular arrangement, as shown in **Fig. 2**. The circles represent only the positions of the atoms, but if the sizes, as indicated in Fig. 1, are superimposed upon the figure, then it appears that there is considerable empty space between the molecules. This unusual feature is a result of the strong and directional hydrogen bonds taking precedence over all other intermolecular forces in determining the structure of the crystal. If the water molecules were rearranged to reduce the amount of empty space, their relative orientations would no longer be so well suited for hydrogen bonds. This rearrangement can be produced by compressing ice to pressures in excess of 2000 atm (14 megapascals). Altogether five different crystalline forms of solid water have been produced in this way, the form obtained depending upon the final pressure and temperature. They are all more dense than water, and all revert to ordinary ice when the pressure is reduced. *See* CRYSTAL STRUCTURE.

Liquid state. The molecules in liquid water also are held together by hydrogen bonds. When ice melts, many of the hydrogen bonds are broken, and those that remain are not numerous enough to keep the molecules in a regular arrangement. Many of the unusual properties of liquid water may be understood in terms of the hydrogen bonds which remain. As water is heated from 0°C (32°F), it contracts until 4°C (39°F) is reached and then begins the expansion which is normally associated with increasing temperature. This phenomenon and the increase in density when ice melts both result from a breaking down of the open, hydrogen-bonded structure as the temperature is raised. The viscosity of water decreases tenfold as the temperature is raised from 0 to 100°C (32 to 212°F), and this also is associated with the decrease of icelike character in the water as the hydrogen bonds are disrupted by increasing thermal agitation. Even at 100°C (212°F), the hydrogen bonds influence the properties of water strongly, for it has a high boiling point and a high heat of vaporization compared with other substances of similar molecular weight. *See* LIQUID.

The electrical conductivity of water is at least 1,000,000 times larger than that of most other nonmetallic liquids at room temperature. The current in this case is carried by ions produced by the dissociation of water according to reaction (4). This reaction

$$H_2O \rightleftharpoons H^+ + OH^- \tag{4}$$

is reversible, and equilibrium is reached rapidly, so there is a definite concentration of H^+ and OH^- ions in pure water. At 25°C (77°F), this concentration is 10^{-7} mole/liter of each species or about 10^{14} ions/ml. This concentration of ions is affected by the temperature or by the presence of solutes in the water. *See* ACID AND BASE; HYDROGEN ION.

Pure water, either solid or liquid, is blue if viewed through a thickness of more than 6 ft (2 m). The other colors often observed are due to impurities.

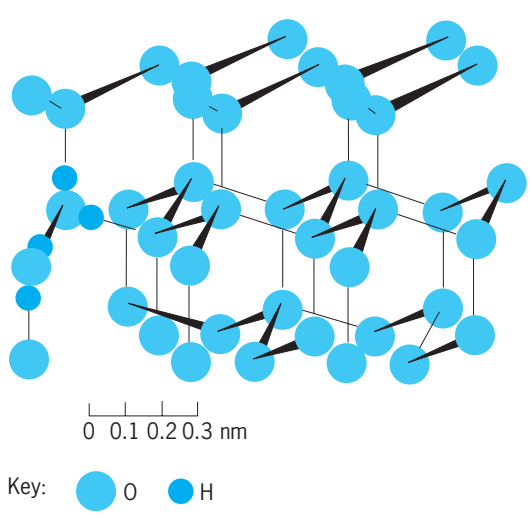

Fig. 2. Structure of ice. The hydrogen atoms are omitted for all but two water molecules.

Solutions in water. Water is an excellent solvent for many substances, but particularly for those which dissociate to form ions. Its principal scientific and industrial use as a solvent is to furnish a medium for purifying such substances and for carrying out reactions between them. *See* SOLUTION; SOLVENT.

Among the substances which dissolve in water with little or no ionization and which are very soluble are ethanol and ammonia. These are examples of molecules which are able to form hydrogen bonds with water molecules, although, except for the hydrogen of the OH group in ethanol, it is the hydrogen of the water that makes the hydrogen bond. On the other hand, substances which cannot interact strongly with water, either by ionization or by hydrogen bonding, are only sparingly soluble in it. Examples of such substances are benzene, mercury, and phosphorus. For discussions of another important class of solutions in water *see* COLLOID; SURFACTANT.

Chemical properties. Water is not a strong oxidizing agent, although it may enhance the oxidizing action of other oxidizing agents, notably oxygen. Examples of the oxidizing action of water itself are its reactions with the alkali and alkaline-earth metals, even in the cold; for instance reaction (5), and its reactions with iron and carbon at elevated temperatures, reactions (6) and (7). Reaction (7) is used commercially to

$$Ca + 2H_2O \rightarrow Ca^{2+} + 2OH^- + H_2 \qquad (5)$$

$$3Fe + 4H_2O \rightarrow Fe_3O_4 + 4H_2 \qquad (6)$$

$$C + H_2O \rightarrow CO + H_2 \qquad (7)$$

produce a gaseous fuel from solid coke. The gaseous mixture, $CO + H_2$, called water gas, is formed when steam is passed over coke heated to 600°C (1100°F).

Water is an even poorer reducing agent than oxidizing agent. One of the few substances that it reduces rapidly is fluorine, but this reaction is complicated. Chlorine is reduced only very slowly in the cold, according to reaction (8).

$$2Cl_2 + 2H_2O \rightarrow O_2 + 4H^+ + 4Cl^- \qquad (8)$$

An example of another sort of oxidation-reduction reaction in which water plays an essential role beyond that of the solvent is the disproportionation of chlorine, reaction (9), which is fast and incomplete

$$Cl_2 + H_2O \rightarrow HOCl + H^+ + Cl^- \qquad (9)$$

in neutral solution but goes to completion if base is added. *See* OXIDATION-REDUCTION.

Substances with strong acidic or basic character react with water. For example, calcium oxide, a basic oxide, reacts in a process called the slaking of lime, reaction (10). Sulfur trioxide, an acidic oxide, also reacts, reaction (11). This reaction occurs in

$$CaO + H_2O \rightarrow Ca(OH)_2 \qquad (10)$$

$$SO_3 + H_2O \rightarrow H_2SO_4 \qquad (11)$$

the contact process for the manufacture of sulfuric acid. Both of these reactions evolve enough heat to produce fires or explosions unless precautions are taken.

Another type of substance with strong acidic character is an acid chloride. Two examples and their reactions with water are boron trichloride, reaction (12), and acetyl chloride, reaction (13). These are

$$BCl_3 + 3H_2O \rightarrow H_3BO_3 + 3HCl \qquad (12)$$

$$CH_3COCl + H_2O \rightarrow CH_3CO_2H + HCl \qquad (13)$$

often termed hydrolysis reactions, as in the reaction of an ester with water, for instance ethyl acetate, reaction (14). A hydrolysis reaction of a different sort is that of calcium carbide, used in the production of acetylene, reaction (15).

$$H_3C\overset{\overset{\displaystyle O}{\|}}{C}OC_2H_5 + H_2O \longrightarrow H_3C\overset{\overset{\displaystyle O}{\|}}{C}OH + HOC_2H_5 \qquad (14)$$

$$CaC_2 + 2H_2O \rightarrow Ca(OH)_2 + C_2H_2 \qquad (15)$$

See HYDROLYSIS.

Water reacts with a variety of substances to form solid compounds in which the water molecule is intact, but in which it becomes a part of the structure of the solid. Such compounds are called hydrates, and are formed frequently with the evolution of considerable amounts of heat. Examples range from the hydrates of simple and double salts, calcium chloride hexahydrate, $CaCl_2 \cdot 6H_2O$, and ammonium aluminum alum, $NH_4Al(SO_4)_2 \cdot 12H_2O$, to the gas hydrates which are stable only at low temperatures, for example, chlorine hydrate, $Cl_2 \cdot 6H_2O$, and xenon hydrate, $Xe \cdot 6H_2O$. *See* CLATHRATE COMPOUNDS; HYDRATE.

For various aspects of water, its uses, and occurrence *see* HEAVY WATER; HYDROGEN; HYDROLOGY; IRRIGATION (AGRICULTURE); OXYGEN; PLANT-WATER RELATIONS; PRECIPITATION (METEOROLOGY); SEAWATER; TRIPLE POINT; VAPOR PRESSURE; WATER POLLUTION; WATER SOFTENING; WATER SUPPLY ENGINEERING; WATER TABLE; WATER TREATMENT; WATERPOWER.

Harold L. Friedman

Water-borne disease

Disease acquired by drinking water contaminated at its source or in the distribution system, or by direct contact with environmental and recreational waters. Water-borne disease results from infection with pathogenic microorganisms or chemical poisoning.

These pathogenic microorganisms include viruses, bacteria, protozoans, and helminths. A number of microbial pathogens transmitted by the fecal-oral route are commonly acquired from water in developing countries where sanitation is poor. Viral pathogens transmitted via fecally contaminated water include hepatitis viruses A and E. Important bacterial pathogens transmitted via fecally

contaminated water in the developing world are *Vibrio cholerae*, enterotoxigenic *Escherichia coli*, *Shigella*, and *Salmonella enterica* serotype Typhi. Water-borne protozoan pathogens in the developing world include *Giardia lamblia* and *Entamoeba histolytica*. The major water-borne helminthic infection is schistosomiasis; however, transmission is not fecal-oral. Eggs of the trematode *Schistosoma* hatch in fresh water, and the larvae invade specific snail hosts. Snails release *Schistosoma* forms called cercariae which penetrate the skin when humans contact infested waters. Eggs are excreted from infected humans in feces or urine, depending on the *Schistosoma* species. Another water-borne helmenthic infection is dracunculiasis (guinea worm infection). The nematode, *Dracunculus medinensis*, is acquired by ingestion of water containing copepods infested with larvae infectious for humans. Larvae infectious for copepods are released by infected humans into water through skin ulcers caused by the adult worm. Dracunculiasis has been the object of an effective global eradication program involving water filtration, elimination of the copepod vector, and education of populations at risk.

In developed countries, fecal contamination of drinking water supplies is less likely. However, there have been outbreaks of diseases such as shigellosis and giardiasis associated with lapses in proper water treatment, such as cross-contamination of waste-water systems and potable water supplies. Animals are therefore more likely to play a role in water-borne disease in developed countries. Bacterial pathogens acquired from animal feces such as nontyphoid *S. enterica*, *Campylobacter jejuni*, and *E. coli* serotype O157:H7 have caused outbreaks of water-borne disease in developed countries where water is not properly chlorinated. Hikers frequently acquire *G. lamblia* infections from drinking untreated lake and stream water. *Giardia lamblia* may have animal reservoirs and can persist in the environment. A recently recognized pathogen apparently resistant to standard chlorination and filtration practices is the protozoan *Cryptosporidium parvum*. This organism is found in the feces of farm animals and may enter water supplies through agricultural runoff. Very large urban outbreaks of diarrheal disease acquired through municipal water supplies have been caused by this pathogen. Filtration reduces the risk of infection but does not completely eliminate this pathogen, for which the infectious dose for humans appears to be very small. A marine dinoflagellate pathogenic for fish, *Pfiesteria piscicida*, has recently been implicated on the Atlantic coast of the United States in disease syndromes involving memory loss. However, the studies are not yet conclusive.

Chemical poisoning of drinking water supplies causes disease in both developing and developed countries. Lead, copper, and cadmium have been frequently involved. *See* CHOLERA; ESCHERICHIA; MEDICAL PARASITOLOGY; SCHISTOSOMIASIS.

Steve L. Moseley

Bibliography. A. S. Benenson, *Control of Communicable Diseases in Man*, 16th ed., American Public Health Association Publications, 1995.

Water conservation

The protection, development, and efficient management of water resources for beneficial purposes. Water's physical and chemical properties make it essential to life and civilization. It is combined with carbon dioxide by green plants in the synthesis of carbohydrates, from which all other foods are formed. It is a highly efficient medium for dissolving and transporting nutrients through the soil and throughout the bodies of plants and animals. It can also carry deadly organisms and toxic wastes. Water is a raw material that is indispensable for many domestic and industrial purposes.

Water Distribution and Cycles

Water occupies more than 71% of the Earth's surface. Over 97% of the (1.386×10^9 km^3) of water is in oceans, and a little over 2.5% is fresh water. Icecaps, glaciers, and other forms of permanent snow that cover the Antarctic, Arctic, and mountainous regions make up 77% of the fresh water; ground water is 22% of all fresh water; and the other 1% of fresh water is divided among the lakes (61%), and atmosphere and soil moisture (39%), and streams (less than 0.4%). Although streams (which include rivers) constitute the smallest percentage of the water distribution, they provide the majority of water used by people.

Hydrologic cycle. The hydrologic cycle is the continuous circulation of water from the ocean to the atmosphere to the land and back to the ocean (**Fig. 1**). Water is temporarily stored in streams, lakes, soil, ice, and underground before it becomes available for use. The ocean stores most of the water, which reenters the cycle when it evaporates due to solar energy. Water is carried through the atmosphere, becoming precipitation (in the form of rain or snow) under the right conditions.

Precipitation. Precipitation occurs when warm air masses that carry moisture are cooled, which decreases their capacity to retain moisture to the point at which water is released.

The source of the air masses and the pressure and temperature gradients aloft determine the form, intensity, and duration of precipitation. Precipitation may occur as cyclonic or low-pressure storms (mostly a winter phenomenon responsible for widespread rains), as thunderstorms of high intensity and limited area, or as mountain storms wherein warm air dumps its moisture when lifted and cooled in crossing high land barriers. *See* PRECIPITATION (METEOROLOGY).

Precipitation is characteristically irregular. Generally, the more humid an area and the nearer the ocean, the more evenly distributed is the rainfall. Whatever the annual average, rainfall in arid regions

tends to vary widely from year to year and to fall in a few heavy downpours. Large floods and active erosion result from heavy and prolonged rainfall or rapid melt of large volumes of snow. Flash floods often follow local intense thunderstorms.

Storage. As rain falls to the Earth, it is caught by vegetation. Although some water is stored on the vegetation itself, most of it runs off onto the ground. The small portion that is stored on the vegetation is called interception. Interception results in evaporation of the water back into the atmosphere. Water that does reach the ground can be absorbed by the soil, remain on the surface and run off into lakes and streams, evaporate from the ground surface back into the atmosphere, or be utilized by vegetation for photosynthesis or transpiration. *See* PLANT-WATER RELATIONS.

Water that seeps into the soil and is held in place as soil moisture by capillary forces is called infiltration. Water that moves over the ground surface and runs off into lakes and streams is called overland flow. Older soil water will eventually become displaced with infiltrating water as soil moisture increases with increasing precipitation, and will result in subsurface runoff. Subsurface runoff can move laterally through the soil or vertically into the ground-water zone, where soil pores and rock are completely filled with water. Ground water will slowly move into streams, lakes, and wetlands and provide surface water.

Recharge. The full recharge (replenishment) of water sources varies according to each water storage element. For example, the full recharge of permafrost is approximately 10,000 years, and polar ice occurs every 9700 years. Alpine glacial recharge is around 1600 years. Recharge of oceanic waters occurs about 2500 years, and ground-water recharge can average 1400 years. Recharge of lakes, bogs, and wetlands ranges between 5 and 17 years, whereas recharge of the atmosphere and stream channel networks ranges between 8 and 16 days.

The hydrologic cycle is a critical framework for identifying how human activity modifies land and water resources. Human activities alter land surface conditions, how water is stored in various parts of the hydrologic cycle, how quickly that storage can be recharged, and how water is distributed across the landscape. Understanding what aspects of the hydrologic cycle are affected by human activities can help to target conservation measures more effectively.

Underground water. Usable ground water in the United States is estimated to be 4.75×10^{10} acre-feet. Annual runoff from the land averages 1.299×10^9 acre-feet (1.602×10^{12} m^3). The volume of ground water greatly exceeds that of all fresh-water lakes and reservoirs combined. It occurs in several geologic formations (aquifers) and at various depths. Ground water under pressure is known as artesian water, and it may become available either by natural or artificial flowing wells. Ground water, if abundant, may maintain streams and springs during extended dry periods. It originates from precipitation of various ages as determined by measurements

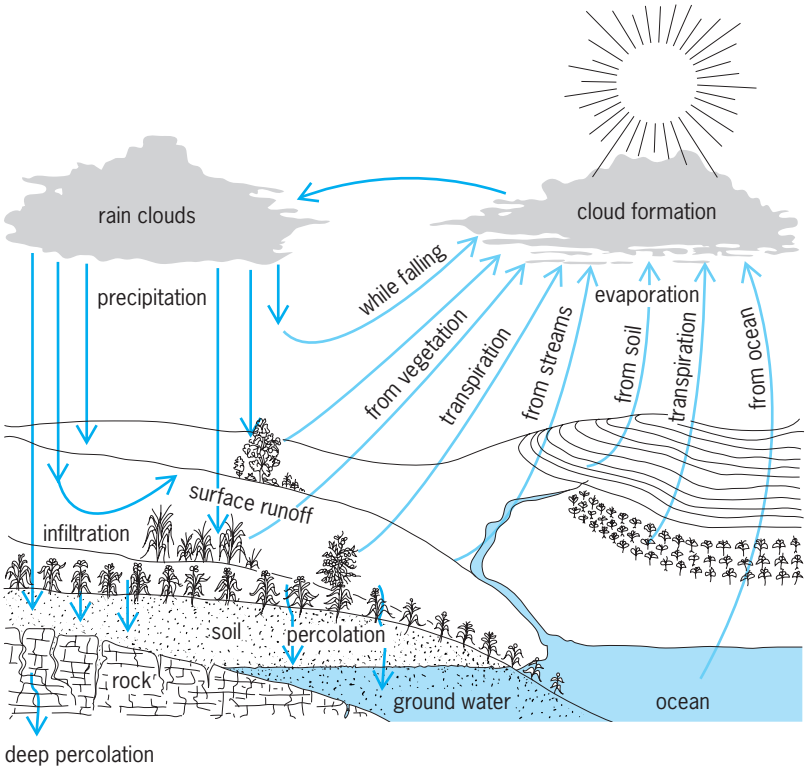

Fig. 1. Hydrologic cycle. (*Water, USDA Yearbook, 1955*)

of the decay of tritium, a radioisotope of hydrogen found in ground water. The water table is the upper level of saturated ground-water accumulation. It may appear at the surface of the Earth in marshes, swamps, lakes, or streams, or hundreds of feet down. Seeps or springs occur where the contour of the ground intercepts the water table. In seeps the water oozes out, whereas springs have distinct flow. Water tables fluctuate according to the source and extent of recharge areas, the amount and distribution of rainfall, and the rate of extraction. The yield of aquifers depends on the porosity of their materials. The yield represents that portion of water which drains out by gravity and becomes available by pumping. Shallow ground water (down to 50 ft or 15 m) is trapped by dug or driven wells, but deep sources require drilled wells. The volume of shallow wells may vary greatly in accordance with fluctuations in rainfall and degree of withdrawal. *See* GROUND-WATER HYDROLOGY.

Surface water. Streams supply most of the water needs of the United States. Lakes, ponds, swamps, and marshes, like reservoirs, represent stored streamflow. The natural lakes in the United States are calculated to contain 1.3×10^{10} acre-feet (1.6×10^{13} m^3). Swamps and other wetlands along river deltas, around the borders of interior lakes, and in coastal regions add millions more to the surface supplies. The oceans and salty or brackish sounds, bays, bayous, or estuaries represent almost unlimited potential fresh-water sources. Brackish waters are being used increasingly by industry for cooling and flushing. Reservoirs, dammed lakes, farm ponds, and other small impoundments have a combined usable

storage of 3×10^8 acre-feet (3.7×10^{11} m^3). The smaller ones furnish water for livestock, irrigation, fire protection, flash-flood protection, fish and waterfowl, and recreation. However, most artificial storage is in reservoirs of over 5000 acre-feet (6.165×10^6 m^3). Lake Mead, located in Arizona and Nevada and formed by Hoover Dam, is the largest (227 mi^2 or 588 km^2) of the 1300 reservoirs, and it contains 10% of the total stored-water capacity, or over 3.1×10^7 acre-feet (3.8×10^{10} m^3). These structures regulate streamflow to provide more dependable supplies during dry periods when natural runoff is low and demands are high. They store excess waters in wet periods, thus mitigating damaging floods. *See* DAM; SURFACE WATER.

Human Impacts

Nearly every human activity—from agriculture to transportation to daily living—relies on water resources and affects the availability and quality of those resources. Water resource development has played a role in flood control, agricultural production, industrial and energy development, fish and wildlife resource management, navigation, and a host of other activities. As a result of these impacts, natural hydrologic features have changed through time, pollution has decreased the quality of remaining water resources, and global climate change may affect the distribution of water in the future. Effective and efficient water resource management and planning will need to consider all these factors.

Patterns of water distribution and use. Water availability varies substantially between geographic regions, but it is also affected strongly by the population of the region. Asia, for example, has an extremely large total runoff but the lowest per-capita water availability (**Table 1**). In addition, nearly 40% of the world's population lives in areas that experience severe to moderate water stress. Thus the combination of water and population distribution has resulted in a large difference in per-capita water use between countries (**Fig. 2**).

Total and per-capita water availability by continent

Continent	Total water availability, km^3/year	Per-capita water availability, 1000 m^3/person/year
Europe	2,900	4.23
North and Central America	7,890	17.40
Africa	4,050	5.72
Asia	13,510	3.92
South America	12,030	38.20
Australia and Oceania	2,360	32.20

Worldwide, nearly 4000 km^3 of water is withdrawn every year from surface and ground waters. This is a sixfold increase from the levels withdrawn in 1900 (since which time population has increased four times). Agriculture accounts for the greatest proportion of water use, with about two-thirds of water withdrawals and 85% of water consumption. It also accounts for a great proportion of the increase in water use, with irrigated cropland more than doubling globally since 1960. However, in Europe and

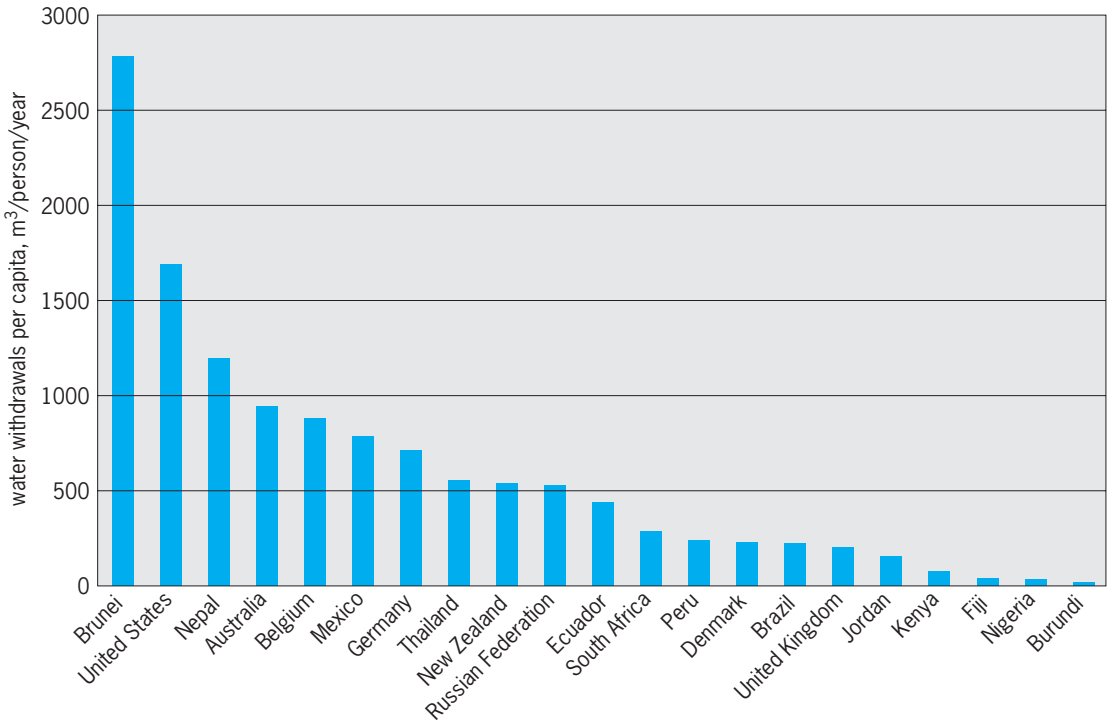

Fig. 2. Estimated per-capita water withdrawals (m^3/person) for selected countries in the year 2000. (*From P. H. Gleick, The World's Water: The Biennial Report on Freshwater Resources, 2000/2001, Pacific Institute for Studies in Development, Environment and Safety, Island Press, Washington, DC, 2001*)

North America particularly, industry consumes a large proportion of available water; industrial uses for water are anticipated to grow on other continents as well.

Changes in hydrologic features. Land development has substantially affected the distribution of water resources. It is estimated that one-half of the natural wetlands in the world have been lost in the last century. In some areas, such as California, wetland loss is estimated to be greater than 90%. The vast majority of wetlands have been lost to diking and draining associated with agricultural development, but urban and industrial changes have reduced wetlands as well. River channels have also been altered to enhance irrigation, navigation, power production, and a variety of other human activities.

Ground-water resources have been depleted in the last century, with many aquifers or artesian sources being depleted more rapidly than they can be recharged. This is called ground-water overdraft. In the United States, ground-water overdraft is a serious problem in the High Plains from Nebraska to Texas and in parts of California and Arizona. In fact, major land subsidence due to ground-water overdraft has occurred in these areas as well as in Japan, Mexico, and Taiwan.

All these changes have resulted in alterations in the natural hydrologic cycle and in the availability of water for human use.

Water pollution. Streams have traditionally served for waste disposal. Towns and cities, industries, and mines provide thousands of pollution sources. Pollution dilution requires large amounts of water. Treatment at the source is safer and less wasteful than flushing untreated or poorly treated wastes downstream. However, sufficient flows must be released to permit the streams to dilute, assimilate, and carry away the treated effluents. *See* WATER POLLUTION.

Global climate change. The availability of fresh water is also likely to be affected by global climate change. There is substantial evidence that global temperatures have risen and will continue to rise. Although the precise effects of this temperature risk on water distribution are challenging to predict, most models of climate change do anticipate increased global precipitation. It is likely that some areas, particularly those at mid to high latitudes, will become wetter, but the increased precipitation will be more seasonal than current patterns. Other areas are likely to receive less precipitation than they do currently. In addition, many models predict increases in the intensity and frequency of severe droughts and floods in at least some regions.

These changes will affect natural stream flow patterns, soil moisture, ground-water recharge, and thus the timing and intensity of human demands for fresh-water supplies. Future conservation and management efforts must consider these changes. *See* GLOBAL CLIMATE CHANGE.

Land management. Land management vitally influences the distribution and character of runoff. Inadequate vegetation or surface organic matter; compaction of farm, ranch, or forest soils by heavy vehicles; frequent crop-harvesting operations; repeated burning; or excessive trampling by livestock, deer, or elk all expose the soil to the destructive energy of rainfall or rapid snowmelt. On such lands little water enters the soil, soil particles are dislodged and quickly washed into watercourses, and gullies may form. *See* LAND-USE PLANNING; SOIL CONSERVATION.

Progress and technologies. There are a variety of measures that can be taken to reduce water consumption; many of them have already been proven effective. In the United States, for example, per-capita water usage dropped 20% from 1980 to 1995. In many cases, improvements to existing systems would contribute to water savings. In the United States, an average of 15% of the water in public supply systems (for cities with populations greater than 10,000) is unaccounted for, and presumably lost. In Jordan, it has been estimated that in the relatively recent past, 30–50% of the domestic water supply is lost due to flaws in the delivery network. In Mexico City, it has been estimated that losses from the water supply system are great enough to support a city the size of Rome.

Improvements can also be achieved by changing industrial and agricultural practices. Agricultural water consumption, the largest single use of water in the world, has an estimated overall water use efficiency of 40%. More effective use of water in agricultural systems can be achieved, for example, with more efficient delivery methods such as drip irrigation. More accurate assessment of soil and plant moisture can allow targeted delivery of water at appropriate times. In industrial settings, recycling and more efficient water use has tremendous potential to reduce water consumption. This is particularly true in the energy sector, where water is used for cooling. Overall, industrial water usage dropped by 30% in California between 1980 and 1990, with some sectors achieving even greater reductions. Japan has achieved a 25% reduction in industrial water use since the 1970s. Additional potential to reduce this usage still exists even in locations such as California where many conservation measures are already in place.

Residential water consumption can also be reduced through conservation measures. For example, prior to 1994, the typical family of four in the United States used 96 gallons (360 liters) of water each day for toilet flushing. New high-efficiency, low-flow toilets can reduce the water required to flush by 70% or more. Additional savings are possible with efficient faucet fixtures and appliances.

Political Concerns

Water conservation in the United States faces a number of institutional as well as technological challenges.

Water rights. In the United States, early rights to water followed the riparian doctrine, which grants the property owner reasonable use of surface waters flowing past his land unimpaired by upstream

landowners. The drier West, however, has favored the appropriation doctrine, which advocates the prior right of the person who first applied the water for beneficial purposes, whether or not his land adjoins the stream. Rights to ground water are generally governed by the same doctrines. Both doctrines are undergoing intensive study.

State laws generally are designed to protect riparian owners against pollution. States administer the regulatory provisions of their pollution-control laws, develop water quality standards and waste-treatment requirements, and supervise construction and maintenance standards of public service water systems. Some states can also regulate ground-water use to prevent serious overdrafts. Artesian wells may have to be capped, permits may be required for drilling new wells, or reasonable use may have to be demonstrated.

Federal responsibilities consist largely of financial support or other stimulation of state and local water management. Federal legislation permits court action on suits involving interstate streams where states fail to take corrective action following persistent failure of a community or industry to comply with minimum waste-treatment requirements. Federal legislation generally requires that benefits of water development projects equal or exceed the costs. It specifies that certain costs be allocated among local beneficiaries but that most of the expense be assumed by the federal government. In 1955, however, the Presidential Advisory Committee on Water Resources Policy recommended that cost sharing be based on benefits received, and that power, industrial, and municipal water-supply beneficiaries pay full cost. These phases of water resource development present difficult and complex questions, because many imponderables enter into the estimates of probable monetary and social benefits from given projects as well as into the cost allocation aspect.

Watershed control. This approach to planning, development, and management rests on the established interdependence of water, land, and people. Conditions on the land are often directly reflected in the behavior of streamflow and in the accumulation of ground water. The integrated approach on smaller watersheds is illustrated by projects under the Watershed Protection and Flood Prevention Act of 1954 (Public Law 566) as amended by P. L. 1018 in 1956. This act originally applied to floods on the smaller tributaries whose watersheds largely are agricultural, but more recently the application has been broadened to include mixed farm and residential areas. Damages from such frequent floods equal half the national total from all floods. Coordination of structures and land-use practices is sought to prevent erosion, promote infiltration, and retard high flows. The Natural Resources Conservation and Forest Services of the Department of Agriculture administer the program. The Natural Resources Conservation Service cooperates with other federal and state agencies and operates primarily through the more than 2000 soil conservation districts.

River basins may be large and complex watersheds. For example, the Tennessee River Basin comprises 40,000 mi^2 (103,600 km^2) in contrast to the 390-mi^2 (1010-km^2) upper limit specified in Public Law 1018. Basin projects may involve systems of multipurpose storage reservoirs, intensive programs of watershed protection, and improvement and management of farm, forest, range, and urban lands. They may call for scientific research, industrial development, health and educational programs, and financial arrangements to stimulate local initiative. The most complete development to date is the Tennessee River Basin, where well-planned cooperative activities have encompassed a wide variety of integrated land and water developments, services, and research.

Water conservation organizations. Organizations for meeting water problems take various forms. Local or intrastate drainage, irrigation, water-supply, or flood-control activities may be handled by special districts, soil conservation districts, or multipurpose state conservancy districts with powers to levy assessments. Interstate compacts have served limited functions on a regional level. To date, Congress has not given serious consideration to proposals for establishing a special federal agency with powers to review and coordinate the recommendations and activities of development services such as the Corps of Engineers, Bureau of Reclamation, Fish and Wildlife Service, Natural Resources Conservation Service, and Forest Service and to resolve conflicts among agencies and citizen groups. However, Congress has passed legislation that has some effects on water use and quality, including the Clean Water Act, the National Environmental Policy Act, the Resource Conservation and Recovery Act, the Safe Drinking Water Act, and the National Energy Policy Act.

In addition, there are a variety of nongovernmental organizations at the local, state, national, and international levels that seek to implement water conservation measures and educate residential, agricultural, and industrial users about available water conservation technologies.

International agreements. Because watersheds often span political boundaries, many efforts to conserve and manage water will require cooperation between states and countries. Many countries currently have international treaties addressing water allocation and utilization. The United States, for example, has treaties with both Mexico and Canada to deal with specific streams such as the Rio Grande and Colorado, or with boundary waters generally, as provided in the treaty with Canada. Similarly, India and Bangladesh have signed a treaty concerning the water in the Ganges River, and Israel and Jordan have an agreement governing the allocation of water from the Jordan River. In 1997, the United Nations adopted the Convention on the Law of the Non-navigational Uses of International Watercourses, which includes an obligation not to cause significant harm to other watercourse states, as well as provisions for dispute resolution. In addition, in 1996 the Global Water Partnership and the

World Water Council were formed for the purpose of addressing ongoing international water concerns. Bernard Frank; Michelle McClure; George Press

Coastal Water Conservation

Most coastal waters less than 300 ft (100 m) deep were dry land 15,000 years ago. The North Sea was a peat bog, for example, and one could walk from England to France or from Siberia to Alaska. As the glaciers retreated, these exposed continental shelves began to fill with water until they now constitute 10% of the world's ocean area. The average depth of the present continental shelf is 183 ft (60 m), with a width extending 47 mi (75 km) from shore. The salinity of coastal water ranges from 35% (100% ocean water) at the seaward edge of the continental shelf (660 ft or 200 m depth contour) to 0% (100% fresh water) within coastal estuaries and bays at the shoreward edge of the shelf. The annual temperature range of mid-latitude shelf water is 68°F (20°C) off New York and 50°F (10°C) off Oregon, with less temperature change in tropical waters. *See* CONTINENTAL MARGIN.

Resources. Because of the shallow bottom, compared with the deep ocean of 16,500 ft (5000 m) depth, organic matter is transformed to nutrients and recycled (returned to the water column) faster on the continental shelves. The growth of plants in the sunlit regions of these relatively shallow areas is thus 10 times that of the open ocean, and the rest of the shelf food web is similarly more productive. Approximately 99% of the world's fish catch is taken from these rich shelves. As a result of their accessibility and commercial value, coastal waters have been the object of extensive scientific studies.

Withdrawal of the Wisconsin glacier and buildup of the native Amerindic populations about 10,000 years ago led to simple harvesting of the living resources of the United States continental shelves. Since colonial days, however, this coastal region has been the focus of increased exploitation with little thought given to the impact of these activities. After the discovery of codfish in the New World by Cabot in 1497, a "foreign" fishing fleet was inaugurated by the French in 1502, the Portuguese in 1506, the Spanish in 1540, and the English in 1578; the first "domestic" fishery of the United States was initiated by the ill-fated Roanoke colony in 1586. The adjacent human population then grew from a few Indian settlements scattered along the coast to the present east coast megalopolis, housed in an almost continuous urban development from Norfolk, Virginia, to Portland, Maine. By 1970, continued fishing pressure of the foreign and domestic fleets had reduced the fish stocks of the northeast continental shelf to approximately 25% of their virgin biomass. *See* MARINE FISHERIES.

Pollutant impacts. At the same time, attempts at waste control in colonial days began as early as 1675 with a proclamation by the governor of New York against dispersal of refuse within the harbor, yet the New York urban effluent expanded until the percent saturation of dissolved oxygen of the harbor halved between 1910 and 1930. The amount of trace metals in the New York Bight Apex sediments now exceeds that of the outer shelf by as much as a hundredfold. Questions about the impact of extended offshore United States jurisdiction of fisheries, construction of ocean sewage outfalls, dredging, beach erosion, and emplacement of pipelines are hotly debated issues in coastal communities that depend on revenue from commercial fishing, tourism, and other forms of recreational activities. As a result of these possibly conflicting uses of the coastal zone, multidisciplinary research on this ecosystem has been intensified over the years by the U.S. Department of Energy (DOE), the National Oceanic and Atmospheric Administration (NOAA), the U.S. Geological Survey (USGS), the Environmental Protection Agency (EPA), the Bureau of Land Management (BLM), and the National Science Foundation (NSF).

People have come to realize that dilution of wastes by marine waters can no longer be considered a simple or permanent removal process within either the open ocean or nearshore waters. The increasing utilization of the continental shelf for oil drilling and transport, siting of nuclear power plants, and various types of planned and inadvertent waste disposal, as well as for food and recreation, requires careful management of human activities in this ecosystem. Nearshore waters are presently subject, of course, to both atmospheric and coastal input of pollutants in the form of heavy metals, synthetic chemicals, petroleum hydrocarbons, radionuclides, and other urban wastes.

However, overfishing is an additional human-induced stress. For example, the sardine fishery collapsed off the California coast, herring stocks are down off the east coast, and the world's largest fishery, for anchovy off Peru, has been reduced to less than 10% of its peak harvest in the late 1960s. Determination of what is the cause and which is the direct effect within a perturbation response of the food web of this highly variable continental shelf ecosystem is a difficult matter. One must be able to specify the consequences of human impact within the context of natural variability; for example, populations of sardines exhibited large fluctuations within the geological record off California before a fishery was initiated.

Furthermore, physical transport of pollutants, their modification by the coastal food web, and demonstration of transfer to humans are sequential problems of increasing complexity on the continental shelf. For example, after 30 years of discharge of mercury into the sea, the origin of the Minimata neurological disease of Japan was finally traced to human consumption of fish and shellfish containing methyl mercuric chloride. The Itai itai disease is now attributed to ingestion of food with high cadmium levels. Discharges of chlorinated hydrocarbons, such as DDT off California, polychlorinated biphenyl (PCB) in both the Hudson River, New York, and within Escambia Bay, Florida, mirex in the Gulf of Mexico, and vinyl chloride in the North Sea, have also led to inhibition of algal photosynthesis, large mortality of

shrimp, and reproductive failure of birds and fish.

Oil spills constitute an estimated annual input of 2×10^6 tons of petroleum to the continental shelves with an unresolved ecological impact; another 2×10^6 tons of petrochemicals is added each year from river and sewer runoff. Fission and neutron-activation products of coastal reactors, such as San Onofre (California), Hanford (Washington), and Wind-scale (United Kingdom), are concentrated in marine food chains with, for example, cesium-137 found in muscle tissue of fish, ruthenium-106 in seaweed, zinc-65 in oysters, and cerium-144 in phytoplankton their somatic and genetic effects on humans are presumably minimal. Finally, disposal of dissolved and floatable waste material from New York City has been implicated as a possible factor in both shellfish loss off New Jersey and the occasional closure of Long Island beaches.

Simulation models. One approach to quantitatively assess the above pollutant impacts is to construct simulation models of the coastal food web in a systems analysis of the continental shelf. Models of physical transport of pollutants have been the most successful, for example, as in studies of beach fouling by oil. Incorporation of additional biological and chemical terms in a simulation model, however, requires dosage response functions of the natural organisms to each class of pollutants, as well as a quantitative description of the "normal" food web interactions of the continental shelf.

Toxicity levels in terms of median lethal concentrations (LC50) of metals, pesticides, biofouling agents (such as chlorine), PCB, and petroleum fractions have been determined only for organisms that can be cultured in the laboratory. The actual form of the pollutant, such as methyl mercuric chloride or chloramine, and its concentration in the marine environment, however, are not always known. Furthermore, the actual contribution of a pollutant to mortality of organisms within a coastal food web is additionally confounded by the lack of understanding of natural mortality. Natural death on the continental shelf is a poorly known process. Nevertheless, there are some clear-cut examples of pollutant impacts on the coastal zone and, in these cases, management decisions are continually being made to correct these situations. *See* FOOD WEB.

Sewage. For example, raw sewage contains pathogenic bacteria that cause human diseases such as typhoid, typhus, and hepatitis. These and various gastroenteric diseases may be contracted from eating raw shellfish that live in sewage-polluted waters. Health authorities have closed more than a million acres (4000 km^2) of the best shellfish beds and put segments of the shellfish industry abruptly out of business, with economic losses running to tens of millions of dollars per year. Purification of sewage is thus absolutely necessary for healthy coastal waters. The cost increases as the degree of treatment is intensified, however. Most treatments remove only a fraction of dissolved fertilizing minerals such as nitrates and phosphates. These nutrients from sewage plants overfertilize coastal waters where the effluent is discharged, and can at times lead to oxygen depletion of bottom waters. The cost of this additional removal of nutrients must now be weighed against their potential damage to the coastal ecosystem. *See* SEWAGE TREATMENT.

Toxic materials. Insecticides also reach coastal waters via runoff from the land, often causing fish kills. Any amount above one-tenth part of insecticide to a million parts of water can be lethal to some fish for most of the following: DDT, parathion, malathion, endrin, dieldrin, toxaphene, lindane, and heptachlor. Contamination of fish eggs by DDT is fatal to a high proportion of young. Insecticides function mainly as paralytic nerve poisons with resulting lack of coordination, erratic behavior, loss of equilibrium, muscle spasms, convulsions, and finally suffocation. Federal and state legislation has all but eliminated DDT from future use in the United States. *See* INSECTICIDE; PESTICIDE.

Other chemical pollutants such as metals, acids, and gases result from industrial activities. Paper and pulp mills discharge wastes that are dangerous to aquatic life because the wastes have a high oxygen demand and deplete oxygen. Other factories discharge lead, copper, zinc, nickel, mercury, cadmium, and cobalt, which are toxic to coastal life in concentrations as low as 0.5 part per million (ppm). Cyanide, sulfide, ammonia, chlorine, fluorine, and their combined compounds are also poisonous. To prevent chemical pollution of the environment, factories are required to remove contaminants from their wastes before discharging them into coastal waters or into local sewage systems.

Oil pollution arises from various sources. Most cases of fish poisoning are from accidental spillage from tankers, storage depots, or wells. However, slow but constant leakage from refineries ruins waterways and is difficult to remedy. Oysters seem unable to breed in the vicinity of refineries. Enclosed ocean regions take longer to recover from oil spills than open coastal areas. Careless handling at plants also results in water pollution by poisonous byproducts, such as cresols and phenols that are toxic in amounts of 5–10 ppm. In past years tankers used to pump oil into the water while cleaning their tanks, but this and cleanup procedures after oil spills are being corrected by stronger federal laws.

Thermal pollution. Thermal pollution is caused by the discharge of hot water from power plants or factories and from desalination plants. Power plants are the main source of heated discharges. They are placed at the coast or on bays to secure a ready source of seawater coolant. A large power installation may pump in 10^6 gal/min (63 m^3/s) and discharge it at a temperature approximately 18°F (10°C) above that of the ambient water. Although temperatures of coastal waters range from summer highs of 95°F (35°C) in southern lagoons to winter lows of 30.3°F (−1°C) in northern estuaries, each has a typical pattern of seasonal temperature to which life there has adapted. In a shallow bay with restricted tidal flow, the rise in temperature can cause gross alterations to the natural ecology. Federal standards prohibit heating of

coastal waters by more than 0.9°F (0.5°C). *See* THERMAL ECOLOGY.

Dredging. Finally, dredging waters to fill wetlands for house lots, parking lots, or industrial sites destroys the marshes that provide sanctuary for waterfowl and for the young of estuarine fishes. As the bay bottom is torn up, the loosened sediments shift about with the current and settle in thick masses on the bottom, suffocating animals and plants. In this way, the marshes are eliminated and the adjoining bays are degraded as aquatic life zones. The northeast Atlantic states have lost 45,000 acres (182 km^2) of coastal wetlands in only 10 years, and San Francisco Bay has been nearly half obliterated by filling. Dredging to remove sand and gravel has the same disruptive effects as dredging for landfill or other purposes, whether the sand and gravel are sold for profit or used to replenish beach sand eroded away by storms. The dredging of boat channels adds to the siltation problem, and disposal of dredge spoils is being regulated in coastal areas.

Management for the future. Human populations have grown to a level where they now can have serious impacts on coastal waters. Past experience suggests that human-induced stress is most likely to lead to species replacement by undesirable forms rather than a decrease in the organic production of the ecosystem. Any societal action must now be considered in the context of what is known about the shelf ecosystem, what management decision is required, what perturbation events are likely to ensue, and what the societal costs are in using renewable and nonrenewable coastal water resources. Prediction of such perturbation events has both immediate and future value to humans in terms of management and conservation options, such as removal of shellfish before depletion of bottom oxygen, the best mode of sewage treatment, preservation of coastal species, and a decrease of toxicant levels within the coastal food web. As one moves from prediction of meteorological events to biological changes of the coastal food web, however, increasing sources of error emerge in the predictions. Since the mid-1800s, humans have introduced more sources of environmental variability to the continental shelf than this coastal ecosystem has encountered during the last 10,000 years. Nevertheless, sufficient information on continental shelf processes is emerging to suggest that specification of management options by delineation of cause and effect within a perturbation response of the coastal zone is a feasible goal.

John J. Walsh

Bibliography. P. E. Black, *Conservation of Water and Related Land Resources*, 2d ed., 1988; T. Dunne and L. B. Leopold, *Water in Environmental Planning*, W. H. Freeman, New York, 1978; P. H. Gleick, *The World's Water: The Biennial Report on Freshwater Resources, 2000/2001*, Publication of the Pacific Institute for Studies in Development, Environment and Safety, Island Press, Washington, DC, 2001; A. W. Hornslow, *Water Quality Data: Analysis and Interpretation*, 1995; N. T. Kottegoda, *Stochastic Water Resources Technology*, 1980; J. W. Moore, *Balancing the Needs of Water Use*, 1988; Soil and Water Conservation Society, *Sustainable Agricultural Systems*, 1990; F. Van der Leeden, F. L. Troise, and D. K. Todd, *The Water Encyclopedia*, 2d ed., Geraghty and Miller Groundwater Series, Lewis Publishers, Chelsea, MI, 1990.

Water desalination

The removal of dissolved minerals (including salts) from seawater or brackish water. This may occur naturally as part of the hydrologic cycle or as an engineered process. Engineered water desalination processes, producing potable water from seawater or brackish water, have become important because many regions throughout the world suffer from water shortages caused by the uneven distribution of the natural water supply and by human use. The capacity of installed desalination plants around the world at the end of 1966 was 200,000 m^3/day (53,000,000 gal/day); by 2002, it had increased to 32,400,000 m^3/day (8,560,000,000 gal/day). *See* WATER SUPPLY ENGINEERING.

Seawater, brackish water, and fresh water have different levels of salinity, which is often expressed by the total dissolved solids (TDS) concentration. Seawater has a TDS concentration of about 35,000 mg/L, and brackish water has a TDS concentration of 1000–10,000 mg/L. Water is considered fresh when its TDS concentration is below 500 mg/L, which is the secondary (voluntary) drinking water standard for the United States. Salinity is also expressed by the water's chloride concentration, which is about half of its TDS concentration. *See* SEAWATER.

Water desalination processes separate feedwater into two streams: a fresh-water stream with a TDS concentration much less than that of the feedwater, and a brine stream with a TDS concentration higher than that of the feedwater.

Thermal processes. Distillation is a process that turns seawater into vapor by boiling, and then condenses the vapor to produce fresh water. Boiling water is an energy-intensive operation, requiring about 4.2 kilojoules of energy (or latent heat) to raise the temperature of 1 kg of water by 1°C. After water reaches its boiling point, another 2257 kJ of energy (or the heat of vaporization) is required to convert it to vapor. The boiling point depends on ambient atmospheric pressure—at lower pressure, the boiling point of water is lower. Therefore, keeping water boiling can be accomplished either by providing a constant energy supply or by reducing the ambient atmospheric pressure. *See* DISTILLATION.

Multistage flash distillation (MFD). The MFD process, developed in the 1960s, is still one of the most popular desalination processes. An MFD system consists of serially arranged cylindrical tanks, called flash chambers, in which the pressure is reduced (**Fig. 1**). Seawater is discharged into the first chamber and pressure reduction lowers the water's boiling point, causing vapor to "flash off" the water surface. The water vapor is then condensed, and the resulting

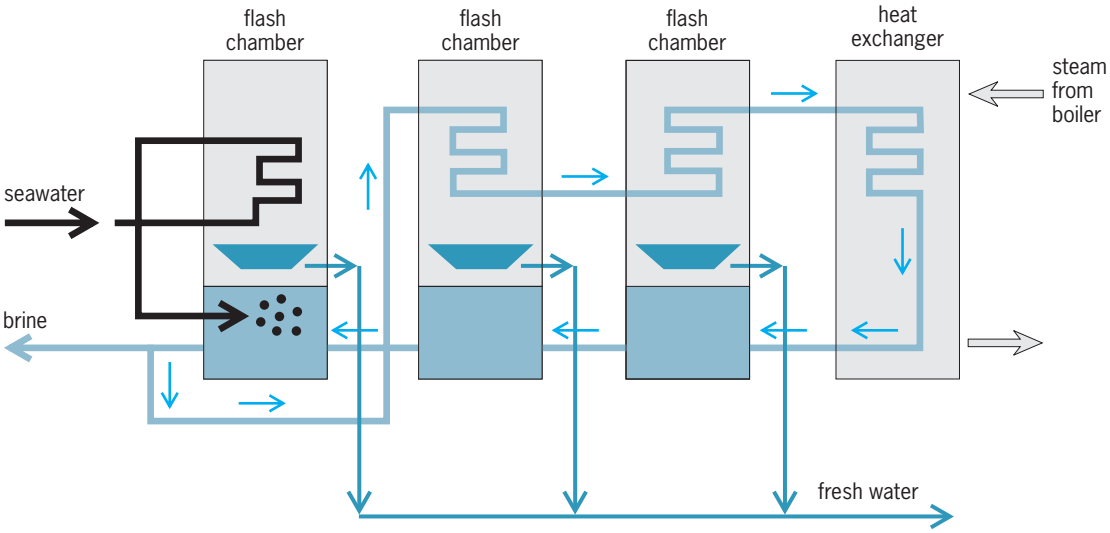

Fig. 1. Multistage flash distillation process.

fresh water is collected. The outflow of concentrated brine from one chamber becomes the inflow of the next chamber, and so on. In this process, a corresponding lower pressure must be maintained in each subsequent chamber for flashing-off to occur.

Multieffect distillation (MED). MED, an alternative distillation process, is used in places where high-temperature steam is available at reasonable cost. In an MED unit, feedwater flows into the top of a series of evaporators and falls as a thin film down the inside of long vertical tubes. The steam surrounding these tubes causes the feedwater to boil as it falls. The resulting vapor is led away, while brine collects at the bottom of the unit. Part of the feed steam condenses to a fresh-water product. The surface of long vertical tubes of all the other units is heated by the steam produced in each preceding unit. *See* EVAPORATION; EVAPORATOR; STEAM.

One major problem with both the MFD and MED processes is scale formation. Substances such as calcium sulfate in feedwater have low solubility in warmer water, and precipitate as the temperature rises, forming scales on the equipment surface. A lower operating temperature can reduce the scale problem, but not without decreasing the thermal efficiency. More research and development on scale control is needed.

Vacuum freezing. When seawater is frozen by reducing the pressure and temperature, the dissolved minerals or salts separate from the ice crystals. After freezing, the ice crystals are washed to remove the brine and melted to produce fresh water. In principle, freezing desalination processes should consume less energy than distillation processes, but no freezing process has been commercially successful. In terms of process engineering, it is more difficult to handle ice than steam.

Membrane processes. When a salt solution is separated from pure water by a semipermeable membrane, water tends to diffuse through the membrane into the salt solution. This is the well-known natural phenomenon called osmosis. *See* OSMOSIS.

Reverse osmosis. Reverse osmosis, the process that causes water in a salt solution to move through a semipermeable membrane to the freshwater side, is accomplished by applying pressure in excess of the natural osmotic pressure to the salt solution (**Fig. 2**). The operational pressure of reverse osmosis for seawater desalination is much higher than that

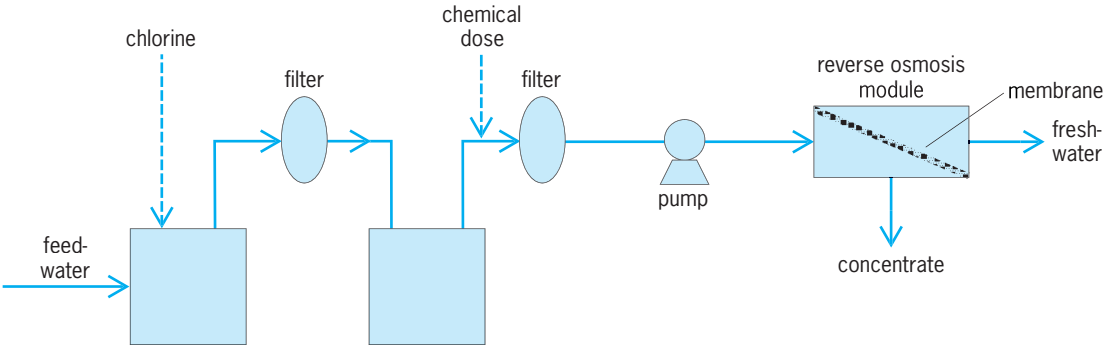

Fig. 2. Flow chart of reverse osmosis desalination.

for brackish water, as the osmotic pressure of seawater at a TDS concentration of 35,000 mg/L is about 2700 kPa (400 psi) while the osmotic pressure of brackish water at a TDS concentration of 3000 mg/L is only about 230 kPa (30 psi).

Electrodialysis. Salts dissociate into positively and negatively charged ions in water. The electrodialysis process uses semipermeable and ion-specific membranes, which allow the passage of either positively or negatively charged ions while blocking the passage of the oppositely charged ions. An electrodialysis membrane unit consists of a number of cell pairs bound together with electrodes on the outside. These cells contain an anion exchange membrane and cation exchange membrane (**Fig. 3**). Feedwater passes simultaneously in parallel paths through all of the cells, separating the product (water) and ion concentrate. *See* DIALYSIS; ION EXCHANGE.

Electrodialysis was successfully introduced in the early 1960s (about 10 years before the introduction of reverse osmosis) as a commercial means of desalting brackish water. This method is cost effective because the unit can be operated at low temperature and pressure.

Membrane fouling. A major problem in the membrane processes of water desalination is fouling, which is the scaling or plugging of membrane surfaces over time by organic and inorganic substances in the feedwater. Fouling prevention requires the pretreatment of the feedwater or the addition of antiscalants. There are three membrane-cleaning methods. In hydraulic cleaning, the flow direction is changed (backflushing) to remove fouling at the membrane surface. In mechanical cleaning, the membrane is cleared with sponge balls. In chemical cleaning, the fouled membrane is washed with chemical agents such as acid for mineral scale or alkali for organic matter. New fouling-resistant membrane materials are being developed by studying the physicochemical and biological interactions between membrane surface and foulants and antifouling agents. *See* MEMBRANE SEPARATIONS.

Processes using natural energy. Desalination processes are energy-intensive. Research continues for improving the energy efficiency of existing processes (which use conventional energy), and new technology is being developed for water desalination processes that are driven by natural energy.

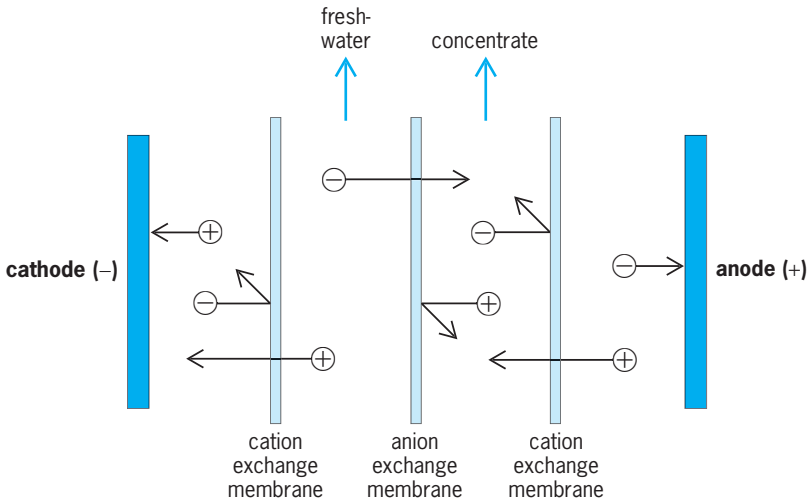

Fig. 3. Electrodialysis process.

Open-cycle OTEC system. Ocean thermal energy conversion (OTEC) uses the ocean's natural thermal gradient between surface water and water at a depth of 1000 m (3300 ft) or more to drive a power-producing cycle. In an open-cycle OTEC system, warm seawater is the working fluid. The warm seawater is flash-evaporated in a vacuum chamber to produce steam. The steam expands through a turbine that is coupled to a generator to produce electricity. The steam existing the turbine is condensed to fresh water, using cold, deep-ocean water (**Fig. 4**).

Wind-powered brackish-water desalination. At present, most interest in wind energy is focused on converting it to electric power. However, converting wind power to electric power and then converting the electric power to hydraulic power for water desalination is not energy efficient or cost effective. A prototype wind-powered reverse-osmosis system, which converts wind power directly to hydraulic power of feedwater into the reverse-osmosis module, was developed by researchers at the University of Hawaii. This prototype system was successfully tested on

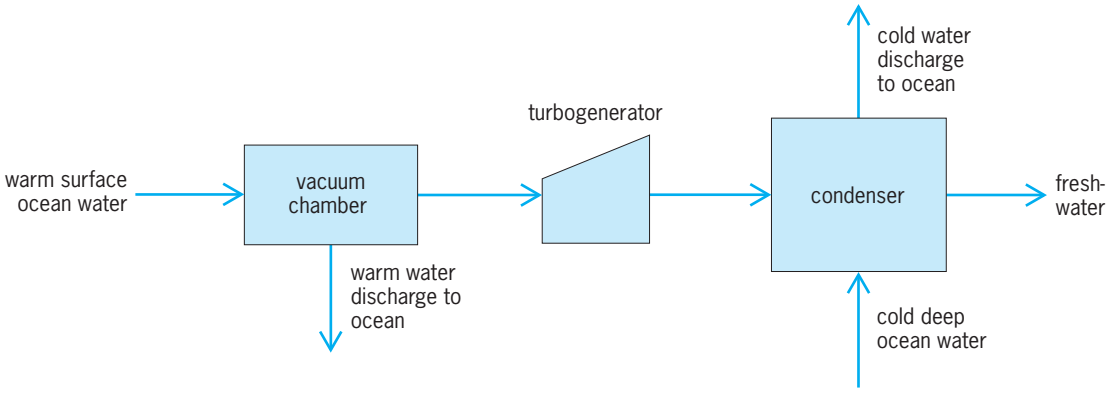

Fig. 4. Open-cycle OTEC system.

Coconut Island off the windward coast of Oahu, Hawaii. *See* WIND POWER.

Wave-driven seawater desalination. An ocean wave–powered reverse-osmosis desalination system was developed at the University of Delaware. This system, which consists of a wave pump and a reverse-osmosis module, was successfully tested at two Caribbean locations. *See* ENERGY SOURCES.

Clark C. K. Liu; Jae-Woo Park

Bibliography. D. Block and J. Valenzuela, *Thermoeconomic Optimization of OC-OTEC Electricity and Water Production Plants*, Solar Energy Research Institute, Golden, CO, SERI/STR-251-2603, 1985; R. G. Gutman, *Membrane Filtration*, Adam Hilger, Bristol, England, 1987; D. Hicks et al., DELBUOY: Ocean wave-powered seawater reverse osmosis desalination system, *Desalination*, 73:81–94, 1989; M. Mulder, *Basic Principles of Membrane Technology*, 2d ed., Kluwer Academic, Dordrecht, 1996; National Research Council, *Review of the Desalination and Water Purification Technology Roadmap*, National Academies Press, 2004; K. Wangnick, *1998 IDA Worldwide Desalination Plants Inventory*, Rep. 15, International Desalination Association, Topsfield, MA, 1998.

Water hammer

The propagation in a liquid of an acoustic wave that is caused by a rapid change in fluid velocity. Such relatively sudden changes in the liquid velocity are due to events such as the operation of pumps or valves in pipelines, the collapse of vapor bubbles within the liquid, underwater explosions, or the impact of water following the rapid expulsion of air from a vent or a partially open valve. Alternative terms such as pressure transients, pressure surge, hydraulic transients, and hydraulic shock are often employed. Although the physics and mathematical characterization of water hammer and underwater acoustics (employed in sonar) are identical, underwater sound is always associated with very small pressure changes compared to the potential of moderate to very large pressure differences associated with water hammer. *See* CAVITATION; SOUND; UNDERWATER SOUND.

A pressure change Δp is always associated with the rapid velocity change ΔV across a water hammer wave, as formulated from the basic physics of mass and momentum conservation by the Joukowsky equation, $\Delta p = -\rho a \Delta V$. Here ρ is the liquid mass density and a is the sonic velocity of the pressure wave in the fluid medium. In a pipe, this velocity depends on the ratio of the bulk modulus of the liquid to the elastic modulus of the pipe wall, and on the ratio of the inside diameter of the pipe to the wall thickness. In water in a very rigid pipe or in a tank, or even the sea, the acoustic velocity is approximately 1440 m/s (4720 ft/s), a value many times that of any liquid velocity.

For an extremely sudden change of the flow of water (assumed instantaneous) in a steel pipe, the effect of water hammer is dramatic (see **table**). However, the acoustic velocity and the pressure rise associated with water hammer are somewhat reduced by the elasticity of the pipe wall. For very elastic pipes such as polyethylene, the acoustic velocity, and the resultant pressure rise for each unit of velocity extinction, is reduced by a much greater amount. *See* ELASTICITY.

Acoustic velocity and pressure change for water in steel pipe*

Inside diameter of pipe ÷ wall thickness	Acoustic velocity (a), m/s (ft/s)	Pressure change (Δp), kPa (psi)
20	1310 (4300)	399 (57.8)
40	1210 (3970)	368 (53.4)
60	1131 (3710)	344 (49.9)
80	1064 (3490)	324 (46.9)

*For velocity change (ΔV) of 0.3048 m/s (1 ft/s).

Fortunately, liquid-handling systems in engineering are designed so that water hammer does not result from sudden closure, but is limited to more gradual flow changes initiated by valves or other devices. The dramatic pressure rise (or drop) results (see table) can be significantly reduced by reflections of the original wave from pipe-area changes, tanks, reservoirs, and so forth. Although the Joukowsky equation applies across every wavelet, the effect of complete valve closure over a period of time greater than a minimum critical time can be quite beneficial. This critical time is the time required for an acoustic wave to propagate twice the distance along the pipe from the point of wave creation to the location of the first pipe-area change. *See* HYDRODYNAMICS; PIPE FLOW.

C. Samuel Martin

Bibliography. M. H. Chaudhry, *Applied Hydraulic Transients*, 2d ed., 1987; F. J. Moody, *Introduction to Unsteady Thermofluid Mechanics*, 1990; B. B. Sharp, *Water Hammer: Practical Solutions*, 1996; E. B. Wylie and V. L. Streeter, *Fluid Transients in Systems*, 1993.

Water-jet cutting

The use of high-pressure water jets, which may contain abrasive powder, for cutting and removing materials. For example, water accelerated up to twice the speed of sound [343 m/s (1125 ft/s) at 20°C (68°F)] can penetrate and cut rock in a few seconds.

Among the methods of cutting metal and nonmetallic materials, pure and abrasive water-jet cutting techniques have a distinct advantage because of their versatility and speed (see **table**). They can cut all materials, including hard-to-machine materials such as super-alloy, Kevlar, and boron carbide. They can also easily cut aerospace materials such as graphite composite and titanium, and brittle materials such as advanced ceramics, granite, marble, and glass (**Fig. 1**). The pure water-jet technique with a relatively high flow rate (2–10 gal/min or 8–38 L/min) is

Cutting speeds of abrasive water jets

Material	Thickness in.	Thickness mm	Speed in./min	Speed mm/min
Aluminium	0.25	6.25	9–40	225–1000
	1.00	25	1–10	25–250
Steel	0.25	6.25	5–20	125–500
	1.00	25	0.5–5	12.5–750
Titanium	0.25	6.25	5–30	125–750
	1.00	25	1–7	25–175
Superalloy	0.25	6.25	8–15	200–375
	1.00	25	0.1–0.5	2.5–12.5
Kevlar	1.00	25	2–10	50–250
Graphite composite	1.00	25	3–10	75–250

widely used by the construction industry for applications such as road repair and tunnel boring. It is also used by the food industry to cut candy and chocolate bars, meats, vegetables, and fruits. Precision enhancement of the technology has led to medical applications, including orthopedic surgery for hip joint replacement. Other biomedical applications include removing clots from blood vessels and a micro water jet for corneal surgery.

The advantages of pure and abrasive water-jet cutting are (1) absence of thermal distortion and work hardening; (2) noncontact during cutting, thus eliminating tool wear and contact force; and (3) omnidirectional cutting, allowing the cutting of complex shapes and contours.

Although the use of the water-jet system is rapidly growing, the technique has some drawbacks and limitations. Water-jet technology has not yet developed fully for high-tolerance and -precision machining. The initial capital investment for the system, including the motion-control equipment and operating costs, is relatively high. And the noise level (80 adjusted decibels) is somewhat high, but the system can be specially designed to isolate the noise source. *See* MACHINABILITY OF METALS.

Cutting principle. The water-jet pump and its delivery system are designed to produce a high-velocity jet stream within a relatively short trajectory distance, since the kinetic energy of the water and abrasive particles is directly proportional to the square of the jet velocity. In abrasive jet cutting applications, the abrasives entrained in the jet stream usually attain approximately 80% of the water-droplet velocity at the nozzle tip. The jet cuts the material by a rapid erosion process, when its force exceeds the compressive strength of the material. Erosion mechanics is a highly dynamic and material-dependent phenomenon involving shear and tensile failure due to localized stress fields. Since the area eroded by the abrasive is also swept by the water stream, the heat generated during the cutting is dissipated immediately, resulting in a small rise in temperature (less than 90°F or 50°C) in the workpiece. Therefore, no thermal distortion or work hardening is associated with water-jet cutting. The cutting by rapid erosion also significantly reduces the actual force exerted on the material, enabling the water jet to cut fragile or deformable materials such as glass and honeycomb structures. Unlike traditional cutting and machining such as turning and drilling, where the cutter is fed by a continuous and constant level of energy during the entire cutting operation, the abrasive water-jet stream loses energy along its cutting path. The cutting power of the jet stream decreases from the top of the material to the bottom, leaving a tapered kerf (slit) and striation marks on the lower portion of the cut surface. This phenomenon is typical of high-energy beam-cutting applications such as laser and electron-beam cutting. *See* JET FLOW; METAL, MECHANICAL PROPERTIES OF; SHEAR.

Operation. A typical ultrahigh-pressure water-jet cutting system consists of two stages of pumping equipment, a high-pressure water-jet delivery system, a mechanical manipulator for motion control, and a discharge catcher system.

Pump. The first stage of pumping is carried out by a radial displacement pump, which can pressurize the hydraulic oil up to 3000 lb/in.^{2} (20.7 megapascals). The hydraulic oil then drives an intensifier pump (**Fig. 2**). The typical intensification ratio of this piston type of pump is 1:20, which is the area ratio of two ends of the piston. The intensifier pump can pressurize water up to 60,000 lb/in.^{2} (414 MPa). The tap water supplied to the intensifier pump passes through several stages of filters (10-0.5-micrometer range) at a boosted water-feed pressure of 80 lb/in.^{2} (0.555 MPa). The high-pressure water from the intensifier pump is channeled through an accumulator to level off pressure fluctuations created by the plunger motion of the pump. The pressurized water is then carried to a cutting station by means of stainless-steel high-pressure tubing and swivel joints or by a flexible high-pressure hose which can withstand pressures

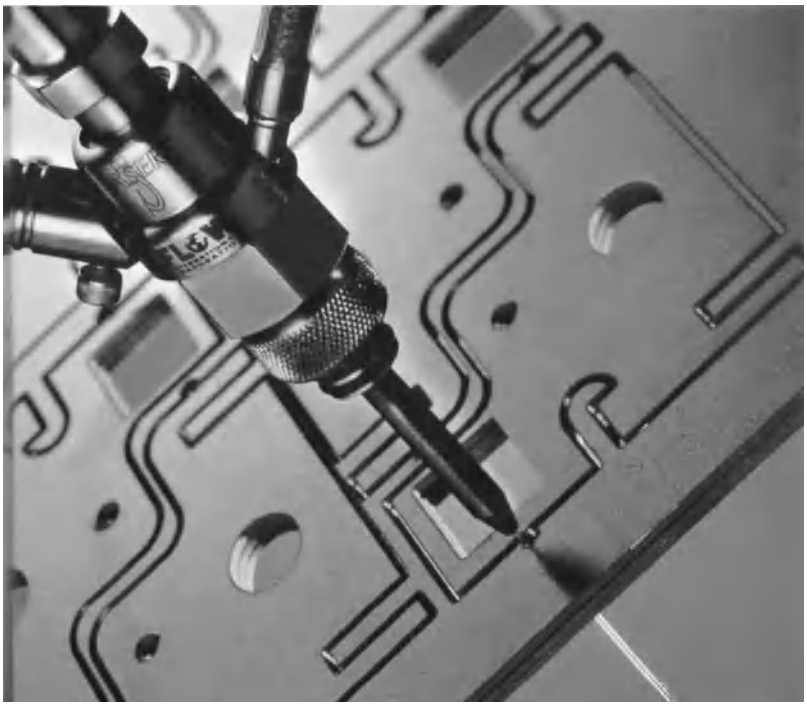

Fig. 1. Abrasive water-jet cutting of 0.5-in.-thick (12.5-mm) titanium at a pressure of 45,000 lb/in.^{2} (310 MPa). (*Flow International Corp.*)

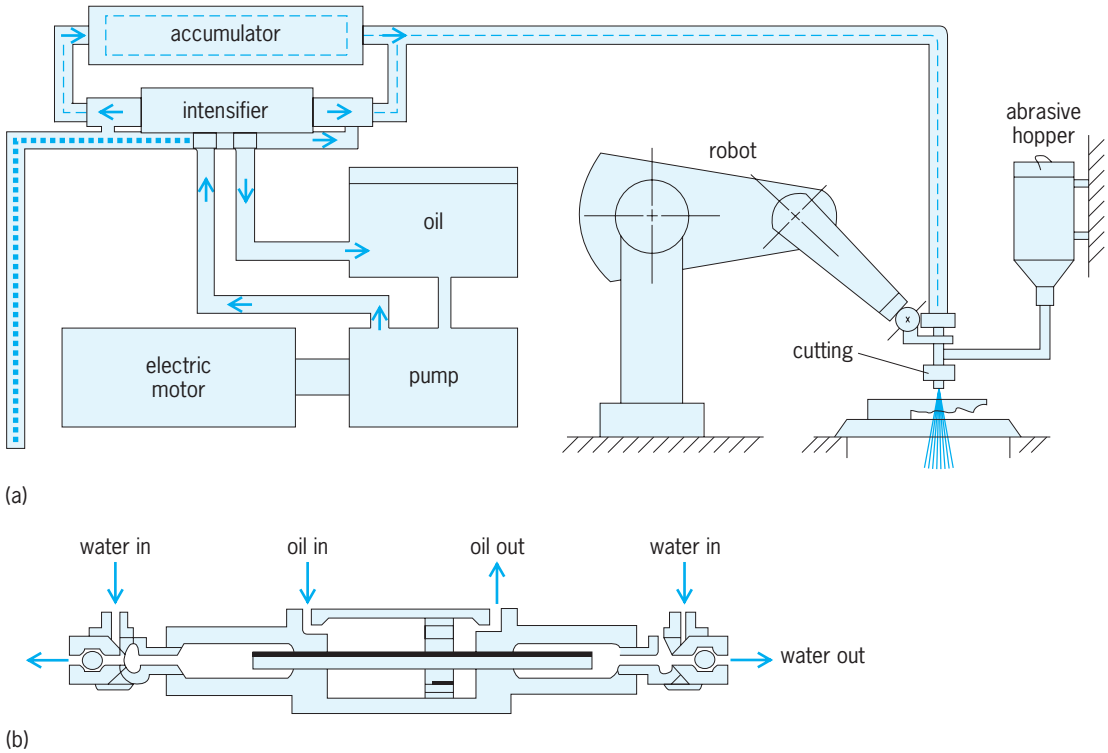

Fig. 2. Water-jet cutting system. (*a*) Water-jet pump and motion control system. (*b*) Double-acting intensifier pump.

up to 100,000 lb/in.^{2} (690 MPa). *See* DISPLACEMENT PUMP.

Jet delivery system. The high-pressure water delivered to a nozzle assembly is first converted into a high-speed jet through an orifice assembly which houses an on–off control valve and orifice. The jet created through the orifice can be manipulated by motion-control equipment such as a robot or *x-y* table to cut a desired shape of workpiece. For cutting hard and dense materials such as steel, stone, and ceramics, the water jet is mixed with abrasive particles in a mixing chamber (**Fig. 3**) before being discharged through the nozzle. The orifice is made of sapphire or diamond, and it lasts up to 200 and 2000 h, respectively. The nozzle (often called the mixing tube) is made of tungsten carbide, which has a life of 5 h under normal operating conditions. A new nozzle material, a composite carbide, can extend the life of the nozzle up to 65 h. The coherent water jet emerging from the tip of the nozzle can attain a speed of 2000 ft/s (610 m/s) at a normal operating pressure of 45,000 lb/in.^{2} (310 MPa). Recent advances in high-pressure pump technology have led to the development of several new single-stage pumps, which also can deliver up to 45,000 lb/in.^{2} but with higher flow rates.

Process control parameters. The precision and efficiency of water-jet cutting can be controlled by the process parameters, including the water pressure; abrasive type, size, and flow rate; orifice and nozzle size; standoff distance; cutting angle; traverse rate; and target material strength (Fig. 3). The velocity of the jet stream as well as the velocity of the entrained abrasive (usually 20% less than the water-droplet velocity) is primarily governed by the water pressure. The jet velocity v can be calculated from the equation below, where P is the water pressure, ρ is the water

$$V = K\sqrt{\frac{2P}{\rho}}$$

density, and K is a dimensionless system constant, of order unity, which includes the effects of water compressibility, orifice efficiency, and the abrasive and

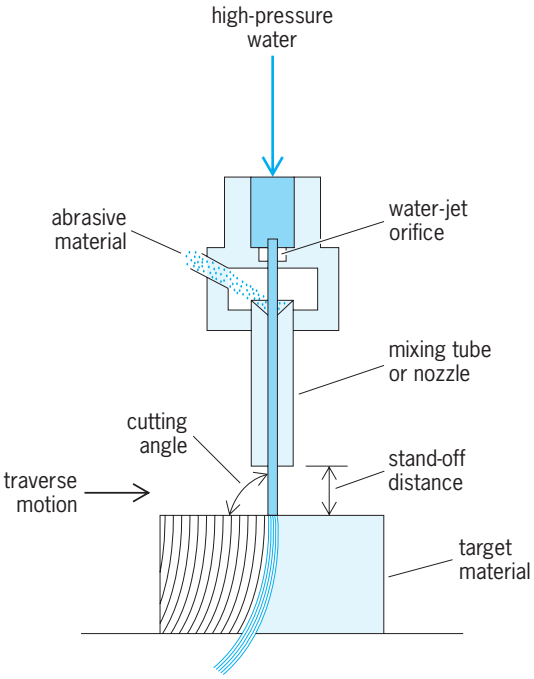

Fig. 3. Typical abrasive water-jet cutting head.

water mass flow rates. The water flow rate is controlled by the water pressure and orifice configuration. The cutting efficiency generally increases as the water flow rate increases. Since the water flow rate is limited by the pump capacity, the orifice size for typical commercially available pumps (20–100 hp, 15–75 kW) ranges between 0.004 and 0.022 in. (0.1 and 0.55 mm). The corresponding range of water flow rate is 0.5–2 gal/min (1.89–7.57 L/min). The optimized pairing in size selection for orifice and nozzle is also an important factor in increasing the efficiency of the cutting operation. The optimum pairing ratio (orifice diameter/nozzle diameter) is 0.3–0.4.

Abrasives. A mixture of water and abrasives is used for cutting hard materials. Garnet is the most commonly used abrasive in industrial cutting operations. Other abrasives such as aluminum oxide, silicon carbide, silica sand, glass bead, and steel grit are also used in special-purpose cutting and cleaning operations. The rate of workpiece erosion, and therefore the cutting efficiency, is very dependent on the material used. Aluminum oxide is an ideal abrasive for cutting brittle materials such as advanced ceramics and stones, yielding a twofold increase in cutting power as compared to garnet. However, aluminum oxide is not much more effective than garnet for cutting ductile materials. The abrasive size and flow rate also greatly affect the quality of the cut surface as well as the cutting efficiency. The range of mesh sizes commonly used for industrial applications is 50–120 (particle size 300–125 μm or 0.012–0.005 in.), and the range of abrasive flow rates is 0.5–1.5 lb/min (3.8–11.4 g/s). *See* ABRASIVE.

Standoff and traverse speed. The standoff distance of the nozzle tip from the workpiece ranges 0–0.2 in. (0–5 mm). A standoff in this range has a minimal influence on the cutting efficiency. The traverse speed of the cutting head is the most critical parameter. The depth of cutting as well as the surface finish is primarily governed by the traverse speed of the nozzle tip, which is controlled by mechanical manipulation. Therefore, the traverse speed is used as a single control parameter in most nonprecision cutting operations.

Thomas J. Kim

Bibliography. H. Louis et al., Investigation of human bones cut with high pressure water-jet, *Proceedings of 16th International Conference on Water Jetting*, October 2002; D. Miller, Micro abrasive water-jet cutting, *Proceedings of 2001 WJTA American Waterjet Conference*, August 2001; A. W. Momber (ed.), *Water Jet Applications in Construction Engineering*, 1998; A. W. Momber and R. Kovacevic, *Principles of Abrasive Water Jet Machining*, 1998; D. A. Summers, *Waterjetting Technology*, 1995; R. A. Tikhomirov et al., *High Pressure Jetting*, 1992.

Water pollution

A change in the chemical, physical, biological, and radiological quality of water that is injurious to its existing, intended, or potential uses. (for example, boating, waterskiing, swimming, the consumption of fish, and the health of aquatic organisms and ecosystems). The term "water pollution" generally refers to human-induced (anthropogenic) changes to water quality. Thus, the discharge of toxic chemicals from a pipe or the release of livestock waste into a nearby water body is considered pollution. Conversely, nutrients that originate from animals in the wild (for example, a herd of elk) or toxins that originate from natural processes (for example, red tides) are not considered pollution.

Contaminants. The contamination of ground water, rivers, lakes, wetlands, estuaries, and oceans can threaten the health of humans and aquatic life. Sources of water pollution are generally divided into two categories (**Table 1**). The first is point-source pollution, in which contaminants are discharged from a discrete location. Sewage outfalls and the 1989 Exxon *Valdez* oil spill are examples of point-source pollution. The second category is non-point-source or diffuse pollution, referring to all of the other discharges that deliver contaminants to water bodies. Acid rain and unconfined runoff from agricultural or urban areas are examples of non-point-source pollution. The principal contaminants of water are shown in (**Table 2**). These include toxic chemicals, nutrients and biodegradable organics, and bacterial and viral pathogens.

Point sources. Since the passage of the Clean Water Act in 1972, there has been considerable progress in reducing the amount of pollution originating from municipal and industrial point sources. Most wastewater treatment plants are designed to accomplish the removal of suspended solids, biodegradable organics, and pathogenic organisms. The impact of the discharge of biodegradable organics can be measured in terms of the buildup of sludge deposits and depletion of the dissolved oxygen resources of water bodies. This situation led to requirements for secondary treatment of wastewaters. Similarly, concern over the toxicity caused by the discharge of heavy metals in treated effluents led to the development of effective pretreatment programs.

Nonpoint sources. Reducing contamination from nonpoint sources has been considerably more difficult, in part because these inputs are widely distributed and highly variable. To better understand the relationship between water quality and chemical use, land use, climate geology, topography, and soils, Congress appropriated funds for the National Water-Quality Assessment (NAWQA) Program in 1991. The NAWQA Program is an ongoing investigation by the U.S. Geological Survey in river basins and aquifers around the country.

Transport and transformation processes. Contaminants discharged to water bodies are subject to a variety of transport and transformation processes and operations that can alter their composition. The physical, chemical, and biological processes that control the fate of the contaminants discharged to water bodies are numerous and varied. It is convenient to divide them into transport processes that affect all water-quality parameters in the same way and fate and transformation processes which are constituent-specific.

TABLE 1. Identification of point and non-point sources of contaminants discharged to water bodies*

Point sources	Non-point (diffuse) sources
Municipal and industrial wastewater effluents	Return flow from irrigated agriculture (specifically excluded from point source definition by Congress)
Runoff and leachate from solid waste disposal sites	Other agricultural and silvicultural runoff and infiltration from sources other than confined concentrated animal operations
Runoff and infiltrated water from concentrated animal feeding operations	Unconfined pastures of animals and runoff from range land
Runoff from industrial sites not connected to storm sewers	Urban runoff from sewered communities with a population of less than 100,000 not causing a significant water-quality problem
Storm sewer outfalls in urban centers with a population of more than 100,000	Urban runoff from unsewered areas
Combined sewer overflows	Runoff from small or scattered (less than 2 hectares or 5 acres) construction sites
Leachate from solid waste disposal sites	Septic tank surfacing in areas of failing septic tank systems and leaching of septic tank effluents
Runoff and drainage water from active mines, both surface and underground, and from oil fields	Wet and dry atmospheric deposition over a water surface (including acid rainfall)
Other sources, such as discharges from vessels, damaged storage tanks, and storage piles of chemicals	Flow from abandoned mines (surface and underground), including inactive roads,tailings, and spoil piles
Runoff from construction sites that are larger than 2 hectares (5 acres)	Activities on land that generate wastes and contaminants, such as deforestation and logging; wetland drainage and conversion; channeling of streams; building of levees, dams, causeways, and flow diversion facilities on navigable waters; construction, and development of land; interurban transportation; military training, maneuvers, and exercises; and mass outdoor recreation

*After V. Novotny and H. Olem, *Water Quality: Prevention, Identification, and Management of Diffuse Pollution*, Van Nostrand Reinhold, 1994.

After initial dilution, contaminants discharged to a water body are transported by two basic processes, advection and dispersion. Advection refers to the transport of a constituent resulting from the flow of the water in which the constituent is dissolved or suspended. Turbulent velocity fluctuations, in conjunction with concentration gradients and molecular diffusion, lead to a mass transport phenomenon called dispersion.

The principal fate and transformation processes that affect contaminants discharged to the environment are operative in most water bodies (**Table 3**). The relative importance of individual fate and transformation processes will be site-specific and will depend on the water-quality parameter under evaluation. For example, deoxygenation brought about by bacterial activity, surface reaeration, sediment oxygen demand, and photosynthesis and respiration are of major importance in assessing the oxygen resources of a stream.

Harmful effects on human health. Water pollution can threaten human health when pollutants enter the body via skin exposure or through the direct consumption of contaminated food or drinking water. For example, many states have issued fish consumption advisories following the detection of mercury in fish tissues. Other priority pollutants, including dichlorodiphenyl trichloroethane (DDT) and polychlorinated biphenyls (PCBs), persist in the natural environment and bioaccumulate in the tissues of aquatic organisms. These persistent organic pollutants are transferred up the food chain (in a process called biomagnification), and they can reach levels of concern in fish species that are eaten by humans. Finally, bacteria and viral pathogens can pose a public health risk for those who drink contaminated water or eat raw shellfish from polluted water bodies. *See* ENVIRONMENTAL TOXICOLOGY; FOOD WEB.

TABLE 2. Contaminants of concern found in point and non-point discharges to water bodies

Contaminants	Concern
Atmospheric pollutants	Acid rain leads to acid (low pH) conditions; deposition of potentially toxic constituents
Biodegradable organics	Depletion of natural oxygen resources and the development of septic conditions
Current-use pesticides	Toxicity to aquatic biota
Dissolved inorganics (such as total dissolved solids)	Inorganic constituents added by usage; reclamation and reuse applications
Heat	Growth of undesirable aquatic life; threat to sensitive life forms, upset of ecological balances
Heavy metals	Toxicity to aquatic biota; many metals are also classified as priority pollutants
Nutrients (nitrogen and phosphorus)	Growth of undesirable aquatic life; eutrophication
Pathogenic organisms	Communicable diseases
Priority organic pollutants	Suspected carcinogenicity, mutagenicity, teratogenicity, or high acute toxicity; many priority pollutants resist conventional treatment methods (known as refractory organics); long-term effects are unknown
Suspended solids (inorganic and organic)	Siltation; formation of sludge deposits and anaerobic conditions

TABLE 3. Fate and transformation processes affecting contaminants discharged to water bodies

Process	Constituents affected
Adsorption and desorption	Metals; trace organics; NH_4^+; PO_4^{-3}
Bacterial conversion, aerobic and anaerobic	Biochemical oxygen demand; nitrification; denitrification; sulfate reduction; anaerobic fermentation (in bottom sediments); conversion of priority organic pollutants
Chemical reactions (hydrolysis, ion exchange, oxidation-reduction, and so on)	Decomposition of organic compounds; specific ion exchange; element substitution
Filtration	Suspended matter; colloidal particles
Flocculation	Suspended matter; colloidal particles
Gas absorption and desorption	O_2; CO_2; CH_4; NH_3; H_2S
Natural decay	Plants; animals; protists (algae, fungi, protozoa); eubacteria (most bacteria); archaebacteria; viruses; radioactive substances
Photochemical reactions	Oxidation of inorganic and organic compounds
Photosynthesis and respiration	Algae, duckweed; submerged macrophites; NH_4^+; PO_4^{-3}; pH
Sedimentation	Suspended solids
Sediment oxygen demand	O_2, particulate biochemical oxygen demand
Surface reaeration	O_2; CO_2
Volatilization	Volatile organic compounds; NH_3; CH_4; H_2S, other gases

Harmful effects on aquatic species. Contaminants have a significant impact on aquatic ecosystems. for example, enrichment of water bodies with nutrients (principally nitrogen and phosphorus) can result in the growth of algae and other aquatic plants that shade or clog streams. If wastewater containing biodegradable organic matter is discharged into a stream with inadequate dissolved oxygen, the water downstream of the point of discharge (typically an outfall) will become anaerobic and will be turbid and dark. Settleable solids, if present, will be deposited on the streambed, and anaerobic decomposition will occur. Over the reach of stream where the dissolved-oxygen concentration is zero, a zone of putrefaction will occur with the production of hydrogen sulfide, ammonia, and other odorous gases. Because many fish species require a minimum of 4–5 mg of dissolved oxygen per liter of water, they will be unable to survive in this portion of the stream. In addition to reductions in dissolved oxygen, aquatic species are sensitive to changes in other physical habitat factors, including pH, temperature, and suspended solids.

Direct exposures to toxic chemicals is also a health concern for individual aquatic plants and animals. For example, pesticides are used to kill undesirable or nuisance organisms in many urban and agricultural areas. These chemicals are frequently transported to lakes and rivers via runoff, and they can have unintended and harmful effects on aquatic life. Obvious signs of contaminant exposure in fish from polluted environments include lesions, tumors, and skeletal deformities. Toxic chemicals have also been shown to reduce the growth, survival, reproductive output, and disease resistance of exposed organisms. These effects, while subtle, can have important consequences for the viability of aquatic populations and communities. *See* INSECTICIDE.

Effluent discharge. Wastewater discharges are most commonly controlled through effluent standards and discharge permits. In the United States, the National Pollution Discharge Elimination System (NPDES), administered by the individual states with federal Environmental Protection Agency (EPA) oversight, is used for the control of wastewater discharges. Under this system, discharge permits are issued with limits on the quantity and quality of effluents. These limits are based on a case-by-case evaluation of potential environmental impacts and, in the case of multiple dischargers, on waste load allocation studies aimed at distributing the available assimilative capacity of the water body. Discharge permits are designed as an enforcement tool, with the ultimate goal of meeting ambient water-quality standards.

Water-quality standards and criteria. Water-quality standards are sets of qualitative and quantitative criteria designed to maintain or enhance the quality of receiving waters. In the United States, these standards are promulgated by the individual states. Receiving waters are divided into several classes depending on their uses, existing or intended, with different sets of criteria designed to protect uses such as drinking water supply, bathing, boating, fresh-water and shellfish harvesting, and outdoor sports for seawater.

Toxicity studies. For toxic compounds, chemical-specific or whole-effluent toxicity studies are used to develop standards and criteria. In the chemical-specific approach, individual criteria are used for each toxic chemical detected in the wastewater. Criteria can be developed to protect aquatic life against acute and chronic effects and to safeguard humans against deleterious health effects, including cancer. The chemical-specific approach, however, does not consider the possible additive, antagonistic, or synergistic effects of multiple chemicals. The biological availability of the compound, which depends on its form in the wastewater, is also not considered in this approach.

The whole-effluent toxicity approach can be used to overcome the shortcomings of the chemical-specific approach. In the whole-effluent approach, toxicity or bioassay tests are used to determine the concentration at which the wastewater induces acute or chronic toxicity effects. In bioassay testing, selected organisms are exposed to effluent diluted in various ratios with samples of receiving water. At

various points during the test, the organisms affected by various effects, such as lower reproduction rates, reduced growth, or death, are quantified. To protect aquatic life discharge limits are established based on the results of the tests. *See* HAZARDOUS WASTE; SEWAGE TREATMENT.

Nathaniel Scholz; George Tchobanoglous

Bibliography. V. Novotny, *Water Quality: Diffuse Pollution and Watershed Management*, 2d ed., 2002; G. Tchobanoglous, F. L. Burton, and H. D. Stensel, *Wastewater Engineering: Treatment and Reuse*, 4th ed., 2002; G. Tchobanoglous and E. D. Schroeder, *Water Quality: Characteristics, Modeling, Modification*, 1985; K. M. Vigil, *Clean Water: An Introduction to Water Quality and Pollution Control*, 2003.

Water resources

The Earth's water supply and its natural distribution. Although water is a renewable resource, which is continually being replaced by precipitation, it is not evenly distributed and is scarce in many areas.

Water reservoirs. Water is stored on the Earth's surface in a number of places called reservoirs (**Table 1**).

Oceans. By far the largest reservoir is the ocean, which contains 96% of the Earth's water and occupies more than two-thirds of the Earth's surface. Ocean water, being saline, is not generally available for human consumption, although it can be used for some purposes, mainly thermoelectric power. *See* HYDROSPHERE; OCEAN.

Glaciers. Fresh water makes up only about 4% of the Earth's water. The largest freshwater reservoir is glacial ice, at 3%. Most of this ice (about 85%) occurs as continental glaciers in Antarctica and less than 10% in the Greenland ice sheet. Alpine or mountain glaciers, which occur in mountain valleys on the continents, contain a small part of the total ice. *See* HYDROLOGY; TERRESTRIAL WATER.

Ground water. The largest reservoir of available fresh water is ground water (1.05% of total water), which is stored in the pores and spaces in rocks, sand, gravel, and soil under the Earth's surface. The top plane of the ground water is referred to as the water table, below which all the spaces are filled with water. About half of the ground water occurs quite near the Earth's surface (<0.8 km) and this is an important source of water for human consumption. Although shallow ground water is continually being refilled by precipitation trickling down to the water table, the rate of recharge is very slow and often takes hundreds or thousands of years. This makes many ground-water aquifers a nonrenewable resource. The rest of the ground water, while at greater depths, does not occur much deeper than a few kilometers, where the pressure of the overlying rock becomes so great that pore space disappears. Deep ground water is harder to recover and is more likely to be saline. A smaller amount of water occurs in the soil above the water table, where both air and water fill the pore spaces; this water is referred to as soil moisture and is tightly held in the pores. *See* AQUIFER; ARTESIAN SYSTEM.

Lakes, rivers, and other reservoirs. Fresh-water lakes and rivers on the Earth's surface contain only 0.01% of the Earth's water. This water is generally available for human consumption. There is also an even smaller reservoir of water in the atmosphere (0.001%), where the water occurs as water vapor gas. The smallest reservoir of water occurs in the biosphere, within plants and animals (0.0006%).

To summarize, the main fresh-water resources available for humans on the Earth's surface are ground water and lake and river water, which together only constitute about 1.1% of the Earth's total water. *See* GROUND-WATER HYDROLOGY; SURFACE WATER.

Hydrologic cycle. Water does not permanently remain in any one reservoir on the Earth but is continually in motion through the hydrologic or water cycle (**Fig. 1**). Residence time is the length of time water spends in a particular reservoir before being removed through the water cycle. The ocean has a residence time of 36,000 years, the time it would take river flow to replace the water in the ocean. *See* ATMOSPHERIC GENERAL CIRCULATION.

The water cycle is driven by solar energy. Water is evaporated from the oceans to the atmosphere in the form of water vapor gas (a flux of 435,400 km^3/year). Part of this water vapor cools and is precipitated back to the oceans in the form of rain (398,000 km^3/year). Some water vapor (37,400 km^3/year) is transported through the atmosphere to the continents where it joins water vapor evaporated from the land (69,600 km^3/year). The vapor cools and precipitates out as rain or snow (107,000 km^3/year). Much of this precipitation runs along the Earth's surface to rivers, which flow to the oceans. A small amount of precipitation trickles down to join the ground water.

TABLE 1. World water*

Location	Surface area, km^2	Water volume, 10^3 km^3	Percent of total
World ocean	362,000,000	1,400,000	95.96
Mixed layer		50,000	
Thermocline		460,000	
Abyssal		890,000	
Glacial ice	18,000,000	43,400	2.97
Ground water		15,300	1.05
Lakes, fresh	855,000	125	0.009
Rivers		1.7	0.0001
Soil moisture		65	0.0045
Atmosphere†		15.5	0.001
Biosphere		2	0.0006
TOTAL	510,000,000	1,459,000	100

*SOURCE: E. K. Berner and R. A. Berner, *Global Environment: Water, Air and Geochemical Cycles*, Prentice Hall, 1996; National Research Council (NRC), *Global Change in the Geosphere—Biosphere*, National Academy Press, 1986; R. L. Nace, Terrestrial water, in AccessScience@McGraw-Hill, http://www.accessscience.com.DOI 10.1036/1097–8542.685800, last modified July 15, 2002 (surface area).

†As liquid water.

The flow of water in rivers to the ocean, which can be estimated reliably, is about 37,400 km^3/year. The water vapor flux from the oceans to the land is set equal to the river flux from the land to the oceans.

The total amount of water on the Earth's surface in the various reservoirs remains roughly constant over time. The general belief is that the amount of water on or near the Earth's surface has not changed greatly since 3.8 billion years ago. *See* SURFACE WATER.

Rivers. The rivers with the greatest volume of water discharged to the oceans are listed in **Table 2**. The first 13 rivers make up about 38% of the total water discharge to the oceans, which is approximately 37,400 km^3/yr. The Amazon River of Brazil and central South America alone comprises 17% of the total. There are three American rivers that are among the world's 17 largest rivers: the Mississippi, St. Lawrence, and Columbia rivers. *See* RIVER.

Size and distribution. The amount of river runoff from the continents is determined by the difference between precipitation and evaporation, or net precipitation in the river basin. Ultimately, the amount of precipitation is determined by the atmospheric circulation of water vapor. Evaporation rates depend upon temperature, and decrease with increasing distance from the equator. Most of the world's large rivers occur in two belts where precipitation exceeds evaporation, an equatorial belt from 10°N latitude to 10°S latitude and a temperate belt from 30 to 60°N and S latitude.

The two largest rivers, the Amazon and the Congo rivers, occur in the equatorial belt. The temperate belt has two major rivers, the Mississippi in the United States and the Yangtze in China. Although precipitation in the subarctic regions (60°N–70°N) is low, there is also low evaporation, resulting in net precipitation and a number of large rivers such as the Yenisei and the Lena in northern Russia and the Mackenzie in the Canadian Arctic. The 20–30°N and S belt, between the equatorial and temperate zones, has many deserts and few rivers because of descending dry air from the overall atmospheric circulation.

TABLE 2. Major world rivers by discharge to ocean

	River	Location	Discharge, km^3/yr
1	Amazon	South America	6300
2	Zaire (Congo)	Africa (Congo)	1250
3	Orinoco	South America	1100
4	Yangtze (Chiang)	Asia (China)	900
5	Bramaputra	Asia (India)	603
6	Mississippi	United States	580
7	Yenisei	Asia (Russia)	560
8	Lena	Asia (Russia)	525
9	Mekong	Asia (Vietnam)	470
10	Ganges	Asia (India)	450
11	St. Lawrence	North America	447
12	Parana	South America	429
13	Irrawaddy	Asia (Burma)	428
15	Mackenzie	North America (Canada)	306
17	Columbia	United States	251
	TOTAL all rivers		37,400

SOURCE: E. K. Berner and R. A. Berner, *Global Environment: Water, Air, and Geochemical Cycles*, Prentice Hall, Upper Saddle River, NJ, 1996.

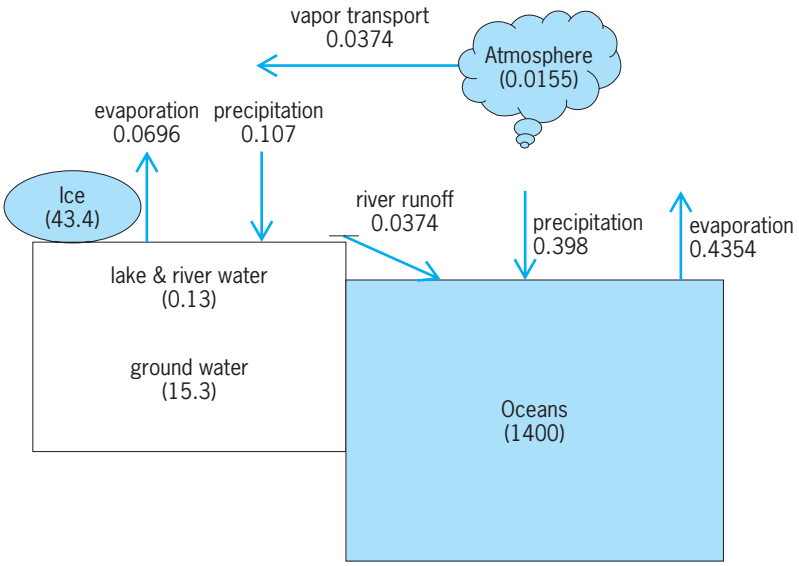

Fig. 1. Water cycle. Volumes of reservoirs in parentheses (in 10^6 km^3); fluxes in 10^6 km^3/yr. (*Data source: E. K. Berner and Robert A. Berner, Global Environment: Water, Air, and Geochemical Cycles, Prentice Hall, Upper Saddle River, NJ, 1996*)

Usage problems. The result of the uneven distribution of net precipitation and world rivers is that many areas do not have adequate water resources. For example, large parts of the flow of the Amazon and Congo rivers, which account for 18.5% of the total global river flow, are geographically inaccessible to populated areas. The remote northern rivers of North America and Eurasia (Mackenzie, Lena, and Yenisei) are also inaccessible. In total, about 20% of world river flow is geographically inaccessible to populated areas and thus not available for human use.

The western part of the United States is semiarid to arid, with precipitation generally declining as one moves west from the Mississippi River. The Colorado River, for example, flows through a generally arid region. Water is allocated by law to various states and withdrawn from the Colorado along its U.S. route until there is almost no flow by the time the river reaches Mexico. The water that remains contains high concentrations of dissolved minerals because of irrigation use and evaporation from large reservoirs behind dams, and therefore Mexico's share of the Colorado River must be desalinated. There are disputes over Colorado River water between California, Arizona, Nevada, and Mexico. The northwestern United States is an exception to the generally dry western United States, with a major river, the Columbia, fed by heavy precipitation in the coastal mountain ranges.

Rivers in other dry parts of the world, such as the Nile in Egypt, have had their flow greatly reduced due to dams and irrigation. At times, the Nile is reduced to zero flow.

Lakes. Freshwater lakes only make up 0.009% of the worlds water by volume (125,000 km^3), but they are important water resources. The world's major freshwater lakes are listed in **Table 3** in order of surface area.

TABLE 3. Major world freshwater lakes

	Lake	Location	Area, km^2	Maximum depth, m	Length, km
1	Superior	United States–Canada	82,103	406	563
2	Victoria	Africa	69,485	82	402
3	Huron	United States–Canada	59,570	229	331
4	Michigan	United States	57,757	281	494
5	Tanganyika	Africa	32,893	1470	676
6	Baikal	Russia	31,500	1620	636
7	Great Bear	Canada	31,329	446	306
8	Nyasa	Africa	28,879	695	579
9	Great Slave	Canada	28,570	614	480
10	Erie	United States–Canada	25,667	64	388
11	Winnipeg	Canada	24,390	18	428
12	Ontario	United States–Canada	19,011	244	311

SOURCE: U.S. Geological Survey.

Size and distribution. The American-Canadian Great Lakes (Superior, Michigan, Huron, Erie, and Ontario) are among the world's major lakes. They contain approximately 25,500 km^3 of water, or one-fifth of the world's fresh water, and are connected with the water finally flowing eastward from Lake Ontario into the St. Lawrence River and on to the Atlantic. The Great Lakes were formed when the last Pleistocene glacier melted about 15,000 years ago and owe their origin to glacial features: glacial erosion, deposits of sand and gravel left by the melting ice, and the general downwarping of the Earth's crust by the weight of glacial ice in north central North America. In addition to the Great Lakes, there are large numbers of small glacial lakes in Canada, Minnesota, Michigan, western New York, and New England. The Finger Lakes in upstate New York are conspicuous examples of glacial erosion. Two other large glacial lakes and many smaller lakes occur in northwestern Canada, including Great Bear Lake and Great Slave Lake, and Lake Winnipeg in central Canada. *See* LAKE.

Lake Baikal in Russian Asia is the world's largest lake by volume (23,000 km^3) because it is also the deepest at 1620 m. It alone contains about one-fifth of the Earth's fresh lake water. Lake Baikal occurs in a rift valley, a narrow downdropped valley with steep faults on either side. This type of geologic setting contributes to its great depth. A number of large, deep African lakes, such as Lake Tanganyika and Lake Nyasa, are also formed in rift valleys. Lake Tanganyika, which is almost as deep as Lake Baikal, also contains a large volume of freshwater at 20,850 km^3.

Lake Chad in central Africa used to be one of the world's largest freshwater lakes by surface area although it was very shallow (5–8 m). However, a combination of extensive use for irrigation and long-term drought have reduced its size by 90% from 25,750 km^2 in the 1960s to 2150 km^2 in 2001.

There are several very large saline lakes. In fact, the world's largest lake, the Caspian Sea (at 371,000 km^2) at the border between Asia and Europe, is saline. It was named a "sea" because of its salinity. The Aral Sea is another large saline lake, although it has been shrunk extensively by the use of its water for irrigation; its original size was 67,340 km^2 and it is now about half that large. In addition, the salty beds left behind the dried-up lake are exposed to the wind, resulting in dust storms that cause severe air pollution. The largest American saline lake, the Great Salt Lake in Utah, which is more saline than seawater, is the remnant of an originally freshwater glacier lake.

Usage problems. One major problem with freshwater lake quality in the United States has been contamination of lakes with nitrogen and phosphorus. This causes overgrowth of algae and green murky water as well as loss of oxygen in the bottom water when the algae die and decay. Phosphorus has been removed from detergents, improving lake water quality, but there remains an ongoing problem with phosphate and nitrogen from sewage effluent and lawn and agricultural runoff.

Acid lakes, particularly in the northeastern United States and in Scandinavia, result from the production of air pollution upwind from the lakes. Sulfuric acid pollution comes from coal burning power plants and nitric acid from automotive pollution and power plants. The air pollution is carried by the atmospheric circulation to susceptible lakes and then falls as acid rain.

United States water use. Freshwater use in the United States, which totaled 345 billion gallons/day in 2000, is shown in **Fig. 2**. The two largest uses, at about 40% each, are irrigation of crops and thermoelectric power generation. Public water supply accounts for 12%, followed by industry at 6%. The

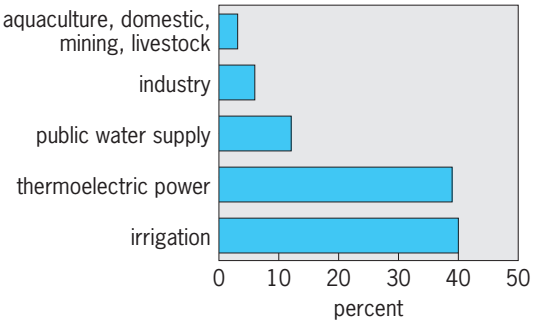

Fig. 2. Freshwater use in the United States in 2000. (*Data source: U.S. Geological Survey Cir. 1268, 2004*)

last four uses are less than 1% each: aquaculture, domestic (well) water, mining, and livestock.

The sources of freshwater are surface water (76%) and ground water at (24%). The largest use of ground water is for irrigation, which has increased to 42% of the total water use or 58 billion gallons/day. Ground water also makes up 37% of the public water supply. These two uses make up most of the total groundwater withdrawal.

The largest withdrawal of ground water occurs in three states, California, Texas and Nebraska, where ground water is used primarily for irrigation. The Ogallala Aquifer, which lies under the High Plains states and is a big ground-water source, was drawn down by more than 100 ft from 1940 to 1980 in southwest Kansas, Texas, and the Oklahoma panhandle. However, better conservation methods have since been implemented, such as more efficient irrigation to reduce water loss.

Worldwide water use. The uses of water worldwide are 70% for agriculture, 10% for domestic purposes such as drinking water, and 20% for industry (more than half of which is used for hydropower). Countries that have scarce water include a number in the belt of low precipitation (20–30° N and S latitude) such as the northern tier of Africa (Mauritania, Algeria, Morocco, Libya, Niger, and Egypt) and the Middle East (Saudi Arabia, Palestine, Syria, and Jordan). Worldwide there are 500 million people in countries with scarce water.

Water for human consumption is unsafe in many places, particularly in the developing countries. It is estimated that as much as 80% of diseases in developing countries are water-related, and 1.7 million people, often children, die from these diseases mainly in Africa and southeast Asia. Typical diseases are diarrhea, cholera, typhoid, and malaria. The main problem is that unsafe disposal of human and animal waste contaminates water for domestic use and irrigation.

More than 50% of the water used by industry (20% of the total) is used for hydropower plants. These plants provide one-fifth of the world's electricity. Hydropower is relatively clean and nonpolluting and is renewable. Dams used for hydropower generation also store water resources for agricultural irrigation, flood prevention, and domestic use. The downside of dams is their displacement of people living in the area of the reservoir. Industrial uses of water can lead to pollution of rivers and aquifers by heavy metals (such as mercury and lead) and persistent organic pollutants (such as DDT and PCB's from electrical insulation).

Agriculture uses 70% of water worldwide, primarily for irrigation. About 65% of irrigation water is "consumed" in distribution and application and by crops and not available for reuse. Irrigation can be wasteful of water and can lead to salt buildup in soils if the soil is poorly drained. Agricultural and lawn runoff often cause overfertilization of water from nitrate and phosphate, causing algal blooms and loss of oxygen in bottom water of rivers, lakes, and estuaries. There have also been problems with agricultural pesticides polluting ground and surface water.

Desalinization of salt water currently supplies only about 0.1% of fresh water. It is expensive since it requires a lot of energy. Thus, it is used primarily for drinking water in water-poor areas. *See* WATER DESALINATION.

Elizabeth K. Berner

Bibliography. E. K. Berner and R. A. Berner, *Global Environment: Water, Air and Geochemical Cycles*, Prentice Hall, Upper Saddle River, NJ, 1996; R. Clarke and J. King, *The Water Atlas*, New Press, New York, 2004; P. H. Gleick, *The World's Water 2004–2005: The Biennial Report on Freshwater Resources (World's Water)*, Island Press, Washington, DC, 2005; P. H. Gleick, Making every drop count, *Sci. Amer.*, 284(2):40–45, 2001; S. S. Hutson et al., Estimated use of water in the United States in 2000, U.S. Geological Survey Circ. 1268, 2004; R. L. Nace, Terrestrial water, in AccessScience@McGraw-Hill, http://www.accessscience.com.DOI 10.1036/1097-8542.685800, last modified July 15, 2002; S. L. Postel, G. C. Daly, and P. R. Ehrlich, Human appropriation of renewable fresh water, *Science*, 271:785–88, Feb. 9, 1996.

Water softening

The process of removing divalent cations, usually calcium or magnesium, from water. When a sample of water contains more than 120 mg of these ions per liter (0.016 oz/gal), expressed in terms of calcium carbonate ($CaCO_3$), it is generally classified as a hard water. Hard waters are frequently unsuitable for many industrial and domestic purposes because of their soap-destroying power and tendency to form scale in equipment such as boilers, pipelines, and engine jackets. Therefore it is necessary to treat the water either to remove or to alter the constituents for it to be fit for the proposed use.

The principal water-softening processes are precipitation, cation exchange, electrical methods, or combinations of these. The factors to be considered in the choice of a softening process include the raw-water quality, the end use of softened water, the cost of softening chemicals, and the ways and costs of disposing of waste streams.

Lime-soda processes. The principal reactions (1)–(7) involved in the lime-soda processes are the

$$CO_2 + Ca(OH)_2 = CaCO_3 \downarrow + H_2O \quad (1)$$

$$Ca(HCO_3)_2 + Ca(OH)_2 = 2CaCO_3 \downarrow + 2H_2O \quad (2)$$

$$Mg(HCO_3)_2 + Ca(OH)_2 = CaCO_3 \downarrow + MgCO_3 + 2H_2O \quad (3)$$

$$MgCO_3 + Ca(OH)_2 = CaCO_3 \downarrow + Mg(OH)_2 \quad (4)$$

$$2NaHCO_3 + Ca(OH)_2 = CaCO_3 \downarrow + Na_2CO_3 + 2H_2O \quad (5)$$

$$MgSO_4 + Ca(OH)_2 = Mg(OH)_2 \downarrow + CaSO_4 \quad (6)$$

$$CaSO_4 + Na_2CO_3 = CaCO_3 \downarrow + Na_2SO_4 \quad (7)$$

precipitation of bicarbonates of calcium and magnesium through the addition of lime, and precipitation of noncarbonates (sulfates, chlorides, and nitrates of magnesium and calcium) by reacting with soda ash.

In reactions (1) and (5), carbon dioxide (CO_2) and sodium bicarbonate ($NaHCO_3$) are not part of the hardness but will consume lime [$Ca(OH)_2$] and therefore must be considered in calculating the required lime amount. Reactions (2), (3), and (4) indicate the removal of carbonate hardness, while reactions (6) and (7) show the removal of magnesium and calcium noncarbonate hardness ($MgSO_4$ and $CaSO_4$), respectively. It is evident from these equations that the amounts of lime and soda ash (Na_2CO_3) required to soften a water may be calculated from the concentrations of free carbon dioxide, bicarbonate, magnesium, and noncarbonate hardnesses.

The types of equipment required for lime–soda ash softening processes generally includes rapid-mixing, flocculation, and sedimentation basins; chemical handling and feeding equipment; conventional rapid sand filters; and, sometimes, carbon dioxide–generating equipment. The conventional rapid-mixing, flocculation, and sedimentation basins are used to carry out the softening process. The chemical handling and feeding equipment includes the lime feeder, slaker, filters, and sludge-handling pumps. A typical precipitator type of water softening plant is shown in the **illustration**. The usual acceptable level of total hardness is 70–120 mg/liter (0.0093–0.016 oz/gal), expressed as $CaCO_3$. As soda ash is more costly than lime, it is advantageous to decrease the carbonate hardness to its lowest practical level. This would allow the maximum concentration of noncarbonate to remain and thus reduce softening costs. Moreover, it would require approximately 190% more soda ash than lime to remove the same amount of hardness.

Complete removal of calcium and magnesium hardnesses is not possible due to the slight solubility of calcium carbonate and magnesium hydroxide. Although, theoretically, it is feasible to reduce the hardness of water to about 25 ppm, in actual practice the range is up to 50–60 ppm or higher.

Variations in the lime-soda process. The fact that satisfactory hardness reduction may still cause deposit formation stimulated the development of methods designed to prevent scale formation by improving the stability of softened water. These methods include (1) conventional lime–soda ash softening followed by recarbonation, which makes it possible to reduce the calcium hardness to 15–20 mg/liter (0.0020–0.0027 oz/gal); (2) excess-lime treatment, a method most appropriate for waters containing 40 mg/liter (0.005 oz/gal) or more magnesium hardness (as $CaCO_3$) and usually requiring neutralization of the excess lime remaining in the solution either by recarbonation or by the use of bypassed raw water (split treatment) after treatment; (3) the hot-lime process, which takes advantage of the high reaction rates and rapid settling of the precipitates due to the decreased water viscosity at high temperatures (200–212°F or 93–100°C), with exhausted steam supplemented by live steam generally used for heating; (4) the combined lime-soda and zeolite process, a method claimed to be most suitable for water containing a high degree of temporary hardness, as water so produced contains a small amount of sodium bicarbonate.

Disposal of the large quantities of lime-softening sludge is a serious environmental problem and must be well planned in advance. The principal methods for disposal are discharge into water streams; discharge into sewers; vacuum filtration to form cake for landfill or as agricultural lime; ponding or lagooning; and recalcination of the dried sludge. The specific disposal method depends, among other things, on local regulations, effects on environment, and costs.

Ion-exchange softening. The basic principle involved in using ion-exchange processes for water

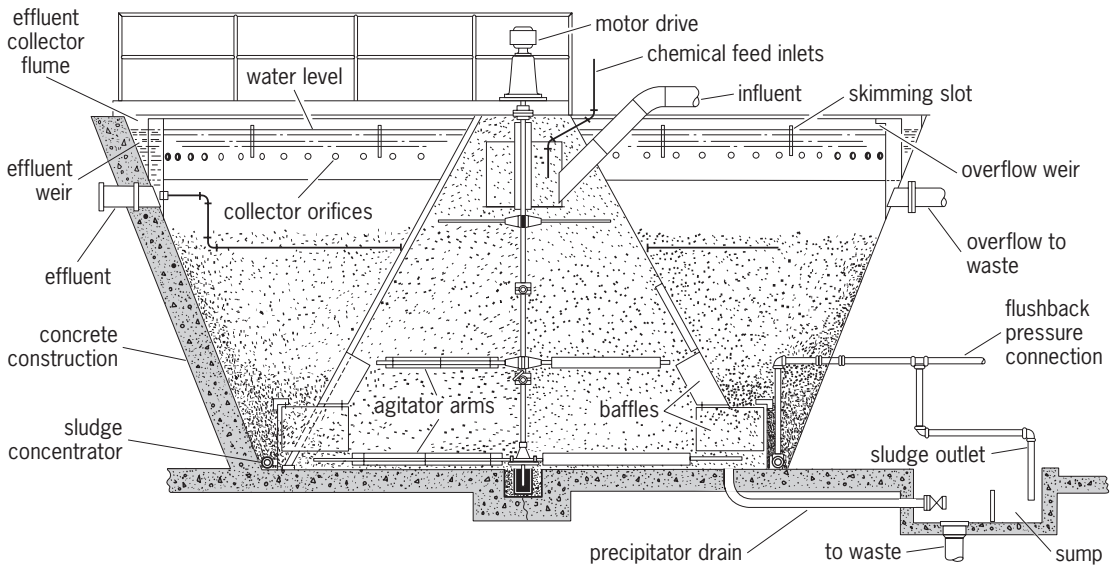

Cold lime-soda water softener, the sludge-blanket type with vertical concrete or steel precipitator.

softening is the utilization of the ability of certain insoluble substances to exchange cations with other cations dissolved in water. They are most appropriate for waters with high noncarbonate hardness or high magnesium content as the use of lime or lime-soda may be quite expensive in these cases. The exchangers used in the processes are zeolites and resins. One of the major advantages in using ion-exchange softening is that the process is continuous and the exchanger can be regenerated. Because the exchange reaction is reversible, after the cations in the exchanger are exhausted, the exchanger can be regenerated with a solution of sodium chloride. *See* ION EXCHANGE.

Zeolite process. Zeolites are crystalline alumino silicates which display cation-exchange properties. In a typical zeolite-softening plant, the hard water enters the top of a zeolite bed and filters down through the bed. Soft water is drawn off from the bottom. Regeneration is accomplished by flowing brine through the bed to remove calcium, and other noncarbonate hardnesses from the zeolite. Washing is generally done by passing water upward to remove traces of brine. The advantages of zeolite processes include production of low water hardness, application to raw water of variable hardness, operation under pressure, and low costs. *See* ZEOLITE.

Cation-exchange resins. The most commonly used cation-exchange resins are sulfonated copolymers of styrene and divinylbenzene. Raw water is generally charged downward through a resin bed which is designed to operate either with pressure or with gravity flow. Regeneration of the resin is accomplished by passing brine uniformly through the bed at a controlled rate through a distributor system. Problems associated with using cation-exchange resins are fouling of the exchanger by the presence of iron, manganese, and aluminum; loss of exchange capacity due to turbidity and bacterial slimes; and swelling caused by the presence of oxidizing agents. *See* WATER TREATMENT. Yi Hua Ma

Bibliography. American Water Works Association, Inc., *Water Quality and Treatment*, 5th ed., 1999; M. W. Harza, *Water Treatment: Principles and Design*, 2d ed., 2005; A. P. Sincero and G. A. Sincero, *Physical-Chemical Treatment of Water and Wastewter*, 2002; Nalco Chemical Co., *The Nalco Water Handbook*, 2d ed., 1988.

Water supply engineering

A branch of civil engineering concerned with the development of sources of supply, transmission, distribution, and treatment of water. The term is used most frequently for municipal water works, but applies also to water systems for industry, irrigation, water reuse, and other purposes.

Sources of Water Supply

Underground waters, rivers, lakes, and reservoirs, the primary sources of freshwater, are replenished by rainfall. Some of this water flows to the sea through surface and underground channels, some is taken up by vegetation, and some is lost by evaporation. Freshwater may also be extracted from saline or brackish waters. Water not intended for potable use, such as for irrigation, ground-water replenishment, salt-water intrusion barriers, or habitat restoration, can be supplied by reusing highly treated wastewater.

Ground water. Water obtained from subsurface sources, such as sand and gravel and porous or fractured rocks, is called ground water. Ground water flows toward points of discharge in river valleys and, in some areas, along the seacoast. The flow takes place in water-bearing strata known as aquifers. The velocity may be a few feet to several miles per year, depending upon the permeability of the aquifer and the hydraulic gradient or slope. A steep gradient or slope indicates relatively high pressure, or head, forcing the water through the aquifer. When the gradient is flat, the pressure forcing the water is small. When the velocity is extremely low, the water is likely to be highly mineralized. If there is no movement, the water is rarely fit for use. *See* AQUIFER; GROUND-WATER HYDROLOGY.

Permeability is a measure of the ease with which water flows through an aquifer. Coarse sands and gravels, and limestone with large solution passages, have high permeability. Fine sand, clay, silt, and dense rocks (unless badly fractured) have low permeability.

Water table. In an unconfined stratum, the water table is the top or surface of the ground water. It may be within a few inches of the ground surface or hundreds of feet below. Normally it follows the topography. Aquifers confined between impervious strata may carry water under pressure. If a well is sunk into such an aquifer and the pressure is sufficient, water may be forced to the surface, resulting in an artesian well. The water-table elevation and artesian pressure may vary substantially with the seasons, depending upon the amount of rainfall recharging the aquifer and the amount of water taken from the aquifer. If pumpage exceeds recharge for an extended period, the aquifer is depleted and the water supply lost. *See* WATER TABLE.

Salt-water intrusion. Normally the ground-water flow is toward the sea. This normal flow may be reversed by overpumping and lowering of the water table or artesian pressure in an aquifer. Salt water flowing into the fresh-water aquifer being pumped is called salt-water intrusion.

Springs. Springs occur at the base of sloping ground or in depressions where the surface elevation is below the water table, or below the hydraulic gradient in an artesian aquifer from which the water can escape. Artesian springs are fed through cracks in the overburden or through other natural channels extending from the aquifer under pressure to the surface. *See* SPRING (HYDROLOGY).

Wells. Wells are vertical openings, excavated or drilled, from the ground surface to a water-bearing stratum or aquifer. Pumping a well lowers the water level in it, which in turn forces water to flow from the

aquifer. Thick, permeable aquifers may yield several million gallons daily with a drawdown (lowering) of only a few feet. Thin aquifers, or impermeable aquifers, may require several times as much drawdown for the same yields, and frequently yield only small supplies.

Dug wells, several feet in diameter, are frequently used to reach shallow aquifers, particularly for small domestic and farm supplies. They furnish small quantities of water, even if the soils penetrated are relatively impervious. Large-capacity dug wells or caisson wells, in coarse sand and gravel, are used frequently for municipal supplies. Drilled wells are sometimes several thousand feet deep.

The portion of a well above the aquifer is lined with concrete, stone, or steel casing, except where the well is through rock that stands without support. The portion of the well in the aquifer is built with open-joint masonry or screens to admit the water into the well. Metal screens, made of perforated sheets or of wire wound around supporting ribs, are used most frequently. The screens are galvanized iron, bronze, or stainless steel, depending upon the corrosiveness of the water and the expected life of the well. Plastic screens are sometimes used.

The distance between wells must be sufficient to avoid harmful interference when the wells are pumped. In general, economical well spacing varies directly with the quantity of water to be pumped and inversely with the permeability and thickness of the aquifer. It may range from a few feet to a mile or more.

Infiltration galleries are shafts or passages extending horizontally through an aquifer to intercept the ground water. They are equivalent to a row of closely spaced wells and are most successful in thin aquifers along the shore of rivers, at depths of less than 75 ft (23 m). These systems are sometimes called bank filtration because the river water is filtered in a natural or engineered system before it reaches the collector. The galleries are built in open cuts or by tunneling, usually with perforated or porous liners to screen out the aquifer material and to support the overburden.

Ranney wells consist of a center caisson with horizontal, perforated pipes extending radially into the aquifer. They are particularly applicable to the development of thin aquifers at shallow depths, and sometimes fed by river water from the surface.

Specially designed pumps, of small diameter to fit inside well casings, are used in all well installations, except in flowing artesian wells or where the water level in the well is high enough for direct suction lift by a pump on the surface (about 15 ft or 5 m maximum). Well pumps are set some distance below the water level, so that they are submerged even after the drawdown is established. Well-pump settings of 100 ft (30 m) are common, and they may exceed 300 ft (90 m) where the ground-water level is low. Multiple-stage centrifugal pumps are used most generally. They are driven by motors at the surface through vertical shafts, or by waterproof motors attached directly below the pumps. Wells are sometimes pumped by air lift, that is, by injecting compressed air through a pipe to the bottom of the well. *See* ARTESIAN SYSTEMS; WELL.

Surface water. Natural sources, such as rivers and lakes, and impounding reservoirs are sources of surface water. *See* DAM; RESERVOIR; SURFACE WATER.

Water is withdrawn from rivers, lakes, and reservoirs through intakes. The simplest intakes are pipes extending from the shore into deep water, with or without a simple crib and screen over the outer end. Intakes for large municipal supplies may consist of large conduits or tunnels extending to elaborate cribs of wood or masonry containing screens, gates, and operating mechanisms. Intakes in reservoirs are frequently built as integral parts of the dam and may have multiple ports at several levels to permit selection of the best water, which may vary seasonally. The location of intakes in rivers and lakes must take into consideration water quality, depth of water, likelihood of freezing, and possible interference with navigation. Reservoir intakes are usually designed for gravity flow through the dam or its abutments.

Wastewater reuse. Because of increasing demands on freshwater sources, it is necessary to use alternative water sources. Wastewater, or treated municipal sewage, is a water resource that can be effectively used in some cases. When sewage is treated by advanced methods, such as coagulation, filtration, membranes, or oxidation, the resulting water quality is often as good as or better than the water source that would normally receive it. This improved effluent can be used for nonpotable applications such as irrigating golf courses and agriculture, replenishing ground-water supplies and salt-water intrusion barriers, or restoring habitats. Federal or state guidelines ensure the safety of this water for its intended reuse. In some areas of the world, highly treated sewage effluent is effectively used as a direct supply at drinking water treatment plants. *See* WASTEWATER REUSE.

Water desalination. In many arid coastal communities, the demand for freshwater is outstripping the economically available supplies. One source of freshwater is treated saline or brackish water from the ocean or an estuary. Water desalination is typically achieved by treating the saline water through a reverse-osmosis process to separate salt ions from the water and recover the freshwater. This process requires a lot of energy and results in a portion of the water going to waste containing concentrated ions and pollutants. *See* WATER DESALINATION.

Transmission and Distribution

The water from the source must be transmitted to the community or area to be served and distributed to the individual customers after proper treatment.

Transmission mains. The major supply conduits, or feeders, from the source to the distribution system are called mains or aqueducts.

Canals. The oldest and simplest types of aqueducts, especially for transmitting large quantities of water, are canals. Canals are used where they can be built economically to follow the hydraulic gradient or slope of the flowing water and are open

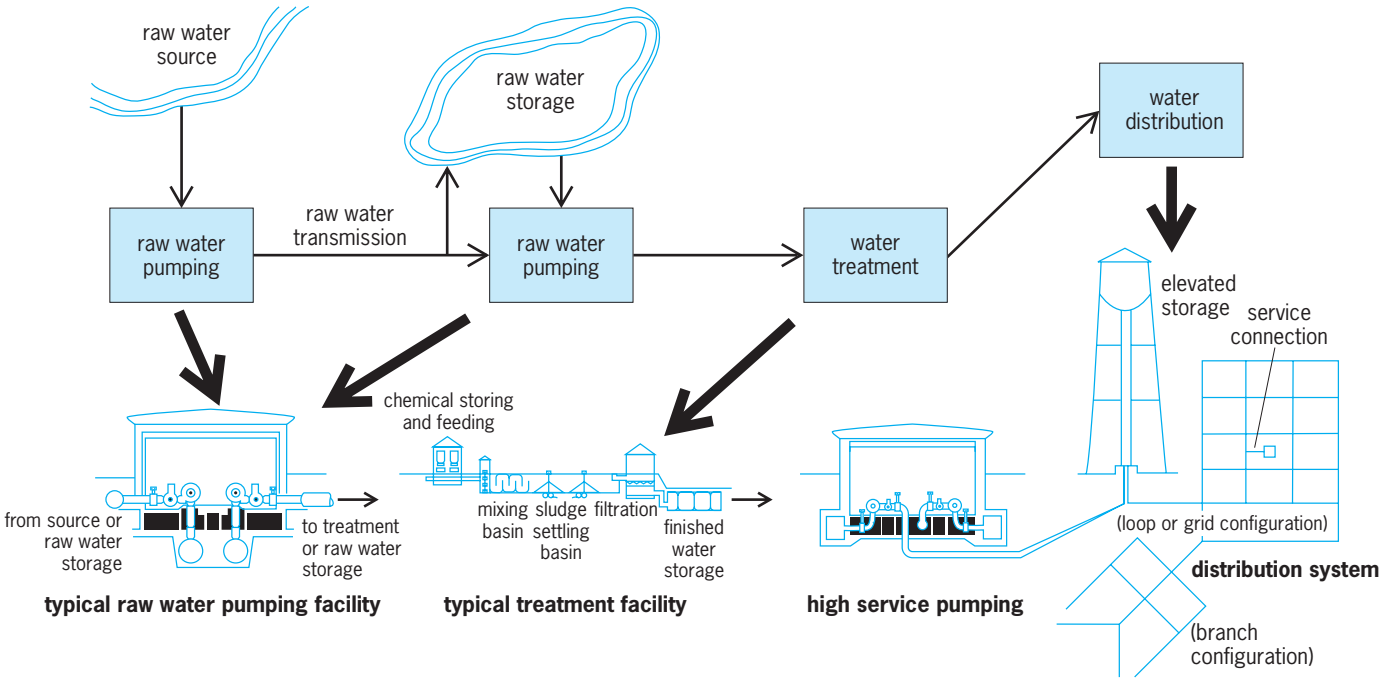

Components of a water utility. (*M. J. Cullinane, Methodologies for the Evaluation of Water Distribution Systems Reliability/Availability, Ph.D. dissertation, University of Texas at Austin, May 1989*).

to the atmosphere. If the soil is suitable, the canals are excavated with sloping sides and are not lined. Otherwise, concrete or asphalt linings are used. Gravity canals are carried across streams or other low places by wooden or steel flumes, or under the streams by pressure pipes known as inverted siphons. *See* CANAL.

Tunnels. Used to transmit water through ridges or hills, tunnels may follow the hydraulic grade line and flow by gravity or may be built below the grade line to operate under considerable pressure. Rock tunnels may be lined to prevent the overburden from collapsing, to prevent leakage, or to reduce friction losses by providing a smooth interior. *See* TUNNEL.

Pipelines. Pipelines are a common type of transmission main, especially for moderate supplies not requiring large aqueducts or canals. Pipes are made of cast iron, ductile iron, steel, reinforced concrete, cement-asbestos, or wood. Pipeline material is determined by cost, durability, ease of installation and maintenance, and resistance to corrosion. The pipeline must be large enough to deliver the required amount of water and strong enough to withstand the maximum gravity or pumping pressure. Pipelines are usually buried in the ground for protection and coolness. *See* PIPELINE.

Distribution system. Included in the distribution system are the network of smaller mains branching off from the transmission mains, the house services and meters, the fire hydrants, and the distribution storage reservoirs. The network is composed of transmission or feeder mains, usually 12 in. (30 cm) or more in diameter, and lateral mains along each street, or in some cities along alleys between the streets. The mains are installed in grids so that lateral mains can be fed from both ends where possible. Mains fed from one direction only are called dead ends; they are less reliable and do not furnish as much water for fire protection as do mains with the grid. Valves at intersections of mains permit a leaking or damaged section of pipe to be shut off with minimum interruption of water service to adjacent areas. Water distributed for potable use typically originates at a water treatment facility (see **illustration**).

House services. Called house services, the small pipes are usually of iron, copper, or plastic extending from the water main in the street to the customer's meter at the curb line or in the cellar. In many cities, each service is metered, and the customer's bill is based on the water actually used.

Fire hydrants. Fire hydrants have a vertical barrel extending to the depth of the water main, a quick-opening valve with an operating nut at the top, and connections threaded to receive a fire hose. Hydrants must be reliable, and they must drain upon closing to prevent freezing.

Distribution reservoirs. These are used to supplement the source of supply and transmission system during peak demands, and to provide water during a temporary failure of the supply system. In small waterworks, the reservoirs usually equal at least one day's water consumption. In larger systems, the reservoirs are relatively smaller but adequate to meet fire-fighting demands. Ground storage reservoirs, elevated tanks, and standpipes are used as distribution reservoirs.

Ground storage reservoirs, if on high ground, can feed the distribution system by gravity, but otherwise it is necessary to pump water from the reservoir into the distribution system. Circular steel tanks and basins built of earth embankments, concrete, or rock masonry are used. Earthen reservoirs are

usually lined to prevent leakage or entrance of dirty water. The reservoirs should be covered to protect the water from dust, rubbish, bird droppings, and the growth of algae, but many older reservoirs without covers are in use, with some in large cities.

Elevated storage reservoirs are tanks on towers, or high cylindrical standpipes resting on the ground. Storage reservoirs are built high enough so that the reservoir will maintain adequate pressure in the distribution system at all times.

Elevated tanks are usually made of steel plate, mounted on steel towers. Wood is sometimes used for industrial and temporary installations. Standpipes are made of steel plate, strong enough to withstand the pressure of the water column. The required capacity of a standpipe is greater than that of an elevated tank because only the upper portion of a standpipe is sufficiently elevated for normal use.

Distribution-system design. To ensure the proper location and size of feeder mains and laterals to meet normal and peak water demands, a distribution system must be expertly designed. As the water flows from the source of supply or distribution reservoir across a city, the water pressure in the pipes is lowered by friction. The pressures required for adequate service depend upon the height of buildings, need for fire protection, and other factors, but 40 lb/in.^{2} (275 kilopascals) is the minimum pressure needed for good service. Higher pressures for fire fighting are obtained by booster pumps on fire engines that take water from fire hydrants. In small towns, adequate hydrant flows are the controlling factor in determining water-main size. In larger communities, the peak demands for air conditioning and lawn sprinkling during the summer months control the size of water main needed. The capacity of a distribution system is usually determined by opening fire hydrants and measuring simultaneously the discharge and the pressure drop in the system. The performance of the system when delivering more or less water than during the test can be computed from the pressure drops recorded.

An important factor in the economical operation of municipal water supplies is the quantity of water lost from distribution because of leaky joints, cracked water mains, and services abandoned but not properly shut off. Unaccounted-for water, including unavoidable slippage of customers' meters, may range from 10% in extremely well- managed systems to 30–40% in poor systems. The quantities flowing in feeder mains, the friction losses, and the amount of leakage are frequently measured by means of pitometer surveys. A pitometer is a portable meter that can be inserted in a water main under pressure to measure the velocity of flow, and thus the quantity of flow. *See* FLOW MEASUREMENT.

Pumping stations. Pumps are required wherever the source of supply is not high enough in elevation to provide gravity flow and adequate pressure in the distribution system. The pumps are high or low head depending upon the topography and pressures required. Booster pumps are installed on pipelines to increase the pressure and discharge, and adjacent to ground storage tanks for pumping water into distribution systems. Pumping stations usually include two or more pumps, each of sufficient capacity to meet demands when one unit is down for repairs or maintenance. The station must also include piping and valves arranged so that a break can be isolated quickly without cutting the whole station out of service.

Centrifugal pumps have replaced steam-driven reciprocating pumps in modern practice, although many of the old units continue to give good service. The centrifugal pumps are driven by electric motors, steam turbines, or diesel engines, with gasoline engines frequently used for standby service. The centrifugal pumps that are used most commonly are designed so that the quantity of water delivered decreases as the pumping head or lift increases. *See* CENTRIFUGAL PUMP; PUMP.

Automatic control of pumping stations is provided to adjust pump operations to variations in water demand. The controls start and stop pumps of different capacity as required. In the event of mishap or failure of a unit, alarms are sounded. The controls are activated by the water level in a reservoir or tank, the pressure in a water main, or the rate of the flow through a meter. A remote control system for pumps is often used, with the signals being transmitted over telephone wires. Karl G. Linden; Richard Hazen

Water Treatment

Drinking water comes from surface and groundwater sources. Surface waters normally contain suspended matter, pathogenic organisms, and dissolved organic substances. Ground water normally contains dissolved minerals and gases. Both require treatment. Conventional water treatment processes include pretreatment, aeration, rapid mix, coagulation and flocculation, sedimentation, filtration, disinfection, and other unit processes to meet specific requirements (see illustration).

Aeration (air or oxygen into water) and air stripping (water into air) primarily are used to remove dissolved gases, such as hydrogen sulfide, which cause taste and odor, as well as to oxidize iron and manganese.

Rapid mix. Rapid (flash) mix is the unit operation where a coagulant chemical, such as alum, is applied and uniformly dispersed. Typically, the tank is equipped with a mechanical mixer and sized for a detention time of 30–60 s.

Coagulation and flocculation. The chemical process (coagulation) of neutralizing the chemical and electrostatic forces that repel particles is followed by the physical process (flocculation) of encouraging particle collisions and agglomerating particles for successful removal by gravity sedimentation. Conventional systems include parallel, rectangular basins equipped with slow-speed mixers to stimulate flocculation. Flocculation time, typically not less than 30 min, is a function of mixing intensity and coagulant dosage.

Sedimentation. Settleable particulate matter, chemical floc, and chemical precipitates are

removed by gravity settling. A typical facility has multiple basins with a working detention time of 2–4 h. Inlet and outlet arrangements, velocity through the basin, hydraulic and solids loading rates, and sludge storage and removal are important design considerations. *See* SEDIMENTATION (INDUSTRY).

Filtration. Using a porous media, the filtration process passes water while capturing colloidal and suspended solids not removed during sedimentation. Typical filter designs include slow, rapid, or pressure filters, using sand, anthracite, or other granular media singularly or in combination. Filter backwashing facilities remove trapped solids and restore the filter's effectiveness. *See* FILTRATION.

Disinfection. A primary drinking-water treatment objective is disinfection to destroy pathogens and other harmful organisms. Chlorination, using chlorine or chloramines, is the traditional disinfection technique, but chlorine dioxide is sometimes used. However, chlorine is not effective against some important pathogens. The use of chlorine in some waters produces chlorinated by-products, some of which are known or suspected carcinogens. Ozonation and ultraviolet (UV) irradiation are principal disinfection alternatives. Ozone is a strong oxidant and is effective against most pathogens. In the presence of bromide, ozonation of water can form bromate (BrO_3^-), which is a carcinogen. UV irradiation is very effective against most pathogens, including the chlorine-resistant pathogen *Cryptosporidium*. UV treatment is not known to produce any by-products of health significance.

Advanced processes. Other treatment processes have specialty applications for specific water quality challenges such as water hardness, iron and manganese that cause staining and color, contaminants that cause taste and odor, presence of natural compounds that form by-products during chlorination, and harmful organic or inorganic pollutants.

For treating high-mineral-content waters, ion exchange is most frequently used for softening, the process of removing calcium and magnesium of hard water. Ion exchange can also be used for deionization when ultrapure water is needed for industrial applications or for removing harmful inorganic pollutants such as arsenic. Reverse osmosis, electrodialysis, ultrafiltration, and other membrane processes are especially applicable for treating water that contains color and some complex organics or inorganics, as well as for desalination. *See* ION EXCHANGE; WATER SOFTENING.

Chemical precipitation is used in water softening, iron and manganese removal, and in other special cases for removing heavy metals. Enhanced coagulation can be used to remove natural organic matter and reduce the level of chlorinated organic contaminants formed during chlorination. Activated carbon (granular or powdered) is used for removing dissolved natural and synthetic organics, taste and odor compounds, and color. *See* ACTIVATED CARBON; PRECIPITATION (CHEMISTRY); WATER TREATMENT.

Oxidation processes, such as ozone, potassium permanganate, and chlorination, are used at the beginning of a water treatment plant to control taste and odor-causing compounds, reduce the microbiological load on the downstream treatment processes, and oxidize any reduced constituents. Advanced oxidation, using hydroxyl radicals produced from hydrogen peroxide in combination with either ozone or UV irradiation, may be implemented to destroy complex organic pollutants remaining in the water after filtration. Karl G. Linden; Robert A. Corbitt

Bibliography. American Society of Civil Engineers and American Water Works Association, *Water Treatment Plant Design*, 4th ed., 2004; American Water Works Association, *Water Quality and Treatment*, 5th ed., 1999; L. S. Clescerl, A. E. Greenberg, and A. D. Eaton (eds.), *Standard Methods for the Examination of Water and Wastewater*, 20th ed., 1999; R. A. Corbitt (ed.), *Standard Handbook of Environmental Engineering*, 2d ed., 1999; M. Watson Harza, *Water Treatment: Principles and Design*, 2d ed., 2005.

Water table

The upper surface of the zone of saturation in permeable rocks not confined by impermeable rocks. It may also be defined as the surface underground at which the water is at atmospheric pressure. Saturated rock may extend a little above this level, but the water in it is held up above the water table by capillarity and is under less than atmospheric pressure; therefore, it is the lower part of the capillary fringe and is not free to flow into a well by gravity. Below the water table, water is free to move under the influence of gravity. The position of the water table is shown by the level at which water stands in wells penetrating an unconfined water-bearing formation.

Where a well penetrates only impermeable material, there is no water table and the well is dry. But if the well passes through impermeable rock into water-bearing material whose hydrostatic head is higher than the level of the bottom of the impermeable rock, water will rise approximately to the level it would have assumed if the whole column of rock penetrated had been permeable. This is called artesian water, and the surface to which it rises is called the piezometric surface. *See* ARTESIAN SYSTEMS.

The water table is not a level surface but has irregularities that are commonly related to, though less pronounced than, those of the land surface. Also, it is not stationary but fluctuates with the seasons and from year to year. It generally declines during the summer months, when vegetation uses most of the water that falls as precipitation, and rises during the late winter and spring, when the demands of vegetation are low. The water table usually reaches its lowest point after the end of the growing season and its highest point just before the beginning of the growing season. Superimposed on the annual fluctuations are fluctuations of longer period which are controlled by climatic variations. The water table is also affected by withdrawals, as by

pumping from wells. *See* GROUND-WATER HYDROLOGY.

Albert N. Sayre; Ray K. Linsley

Bibliography. W. Back (ed.), *Hydrogeology*, 1989; S. N. Davis and R. J. M. DeWiest, *Hydrogeology*, 1991; L. Huisman and T. N. Olsthorn, *Groundwater Recharge*, 1983; R. C. Pyne, *Groundwater Recharge Through Wells: A Guide to Aquifer Storage Recovery*, 1995.

Water treatment

Physical and chemical processes for making water suitable for human consumption and other purposes. Drinking water must be pathogenically safe, free from toxic or harmful chemical or substances, and comparatively free of turbidity, color, and taste- or odor-producing substances. Excessive hardness and high concentration of dissolved solids are also undesirable, particularly for boiler feedwater and industrial purposes. The treatment processes commonly used are sedimentation, coagulation, filtration, disinfection, softening, and aeration. A recently developed water treatment technology uses pressure-driven membranes as a barrier to pollutants.

Sedimentation. Silt, clay, and other fine material settle if the water is allowed to stand or flow at low velocity. Sedimentation occurs naturally in reservoirs and is accomplished in treatment plants using basins or settling tanks. The detention time in a settling basin may range from an hour to several days. The water may flow horizontally through the basin, with solids settling to the bottom, or may flow vertically upward at a low velocity so that the particles settle through the rising water. Settling basins are most effective if shallow and rarely exceed 3–6 m (10–20 ft) in depth. Basic sedimentation does not remove extremely fine or colloidal material within a reasonable time; therefore, it is used principally as a preliminary process to other treatment methods or following the coagulation–flocculation process.

Coagulation and flocculation. The chemical process (coagulation) of neutralizing the chemical and electrostatic forces that repel particles is followed by the physical process (flocculation) of encouraging particle collisions and agglomerating particles for successful removal by gravity sedimentation. Thus, extremely fine particles and colloidal material are combined into larger masses by this process. These masses, called floc, are large enough to settle in basins and to be caught on the surface of filters. Waters high in organic material and iron may coagulate naturally with gentle mixing. The term "coagulation" is usually applied to chemical coagulation, in which iron or aluminum salts are added to the water to form insoluble hydroxide floc. The floc is a feathery, absorbent substance to which color-producing colloids, bacteria, fine particles, and other substances become attached and are removed from the water.

The coagulant dose is a function of the physical and chemical character of the raw water, the adequacy of settling basins and filters, and the degree of purification required. Moderately turbid water coagulates more easily than perfectly clear water, but extremely turbid water requires more coagulant. Coagulation is more effective at higher temperatures. Lime, soda ash, or caustic soda may be required in addition to the coagulant to provide sufficient alkalinity for the formation of floc, and regulation of the pH (hydrogen-ion concentration) is usually desirable. Powdered limestone, clay, bentonite, or silica are sometimes added as coagulant aids to strengthen and weight the floc, and a wide variety of polymers developed in recent years are used for the same purpose.

Filtration. Suspended solids, colloidal material, bacteria, and other organisms are filtered out by passing the water through a bed of sand or pulverized coal, or through a matrix of fibrous material supported on a perforated core. Filtration of turbid or highly colored water usually follows sedimentation or coagulation and sedimentation. Soluble materials, such as salts and metals in ionic form, are not removed by filtration. *See* FILTRATION.

Slow sand filters. Used first in England around 1850, slow sand filters consist of beds of sand 51–122 cm (20–48 in.) deep, through which the water is passed at fairly low rates of $2.4–9.4 \times 10^4$ m^3/ha ($2.5–10 \times 10^6$ gal/acre). The bed sizes range from a fraction of an acre in small plants to several acres in large plants. An underdrain system of graded gravel and perforated pipes transmits the water from the filters to the point of discharge. The sand is usually fine, ranging 0.2–0.5 mm in diameter. The top of the filter clogs from bacterial growth with use, and a thin layer of dirty sand is scraped from the filter periodically to maintain proper flow-through capacity.

Slow sand filters operate satisfactorily with reasonably clear waters but clog rapidly with turbid waters. The filters are covered in cold climates to prevent the formation of ice and to facilitate operation in the winter. In milder climates, they are often open. Slow sand filters have a high bacteriological efficiency, but few have been built since the development of water disinfection, because of the large area required, the high construction cost, and the labor needed to clean the filters and to handle the filter sand. Slow sand filters are still used in many English and European cities, but have not been built in the United States since 1950, and few remain in operation.

Rapid sand filters. These operate at rates of $1.17–2.34 \times 10^6$ m^3/(ha/day) [$1.25–2.5 \times 10^8$ gal/(acre/day)], or 25 to 50 times faster than the slow sand filters. The high rate of operation is made possible by coagulation and sedimentation ahead of filtration to remove the heaviest part of the load, the use of coarse sand, and facilities for backwashing the filter to keep the bed clean. The filter beds are small, generally ranging from 14 m^2 (150 ft^2) in small plants to 140 m^2 (1500 ft^2) in the largest plants. The filters consist of a layer of sand or, occasionally, crushed anthracite coal 46–62 cm (18–24 in.) deep, resting on graded layers of gravel above an underdrain system. The sand is coarse, 0.4–1.0 mm in diameter, depending upon the raw water quality and pretreatment, but the grain size must be fairly uniform to assure proper

backwashing. The underdrain system serves both to collect the filtered water and to distribute the wash water under the filters when they are being washed. Several types of underdrains are used, including perforated pipes, perforated false bottoms of concrete, and tile and porous plates.

Filters are backwashed at rates 5–10 times the filtering rate. The wash water passes upward through the sand and out of the filters by way of wash-water gutters and drains. Washing agitates the sand bed and releases the dirt to flow out of the filter with the wash water. The quantity of water used for washing ranges from 1 to 10% of the total output, depending upon the turbidity of the water applied to filters and the efficiency of the filter design. Combination air and water filter washes are popular in Europe but are not often used in the United States.

Municipal and large-capacity filters for industry usually are built in concrete boxes or in open tanks of wood and steel. The flow through the sand may be caused by gravity, or the water may be forced through the sand under pressure by pumping. Pressure filters can be operated at higher rates than gravity filters, because of the greater head available to force the water through the sand. However, excessive pressure may increase the effluent turbidity, and bacteria may appear in the discharge water. For this reason, and because pressure filters are difficult to inspect and keep in good order, open gravity filters are favored for public water supplies.

Diatomaceous earth filters. Swimming-pool installations and small water supplies frequently use this type of filter. The filters consist of a medium or septum supporting a layer of diatomaceous earth through which the water is passed. A filter layer is built up by the addition of diatomaceous earth to the water. When the pressure loss becomes excessive, filters must be backwashed and a fresh layer of diatomaceous earth applied. Filter rates of $1.7–4.1 \times 10^{-3}$ $m^3/(s/m^2)$ [2.5–6 gal/(min/ft^2)] are attained.

Disinfection. There are several water treatment methods designed to kill harmful microorganisms, particularly pathogenic bacteria, spores, viruses, and protozoa. The application of chlorine or chlorine compounds is the most common method; however, its use is under increasing scrutiny. Emerging disinfection methods include the use of ultraviolet light and ozone, while boiling is used as a household emergency measure.

Chlorination is practiced almost universally in public water supplies in North America. It is sometimes the sole treatment of clear, uncontaminated waters. In conventional water treatment plants, it is used along with coagulation–flocculation and filtration. Chlorination, in the form of free chlorine or chloramines, is used also to protect against contamination of water in distribution mains and reservoirs after purification.

Chlorine gas is relatively economical and easy to apply in large systems. For small works, calcium hypochlorite or sodium hypochlorite is frequently used. Regardless of which form is used, the dose varies with the water quality and degree of contamination. Clear, uncontaminated water can be disinfected with small doses, usually less than 1 part per million, whereas contaminated water may require several times as much. The amount of chlorine taken up by organic matter and minerals in water is known as the chlorine demand. For proper disinfection, the dose must exceed the demand so that free chlorine remains in the water for adequate contact time to achieve the desired disinfection. Chlorine is applied before filtration, after filtration, or at both times. *See* CHLORINE; HYPOHALOUS ACID.

Chlorination alone is not reliable for the treatment of contaminated or turbid water. A sudden increase in the chlorine demand may absorb the full dose and provide no residual chlorine for disinfection, and it cannot be assumed that the chlorine will penetrate particles of organic matter, which may shield pathogens.

Although chlorine is currently an integral part of most water treatment plants, there are a number of concerns related to the use of chlorine. Chlorine sometimes causes objectionable tastes or odors in water. This may be due to an excessive chlorine dose, but more frequently it is caused by a combination of chlorine and organic matter, such as algae, in the water. Some algae, relatively unobjectionable in the natural state, produce unbearable tastes after chlorination. In other cases, strong chlorine doses oxidize the organic matter completely and produce odor-free water. When chlorine combines with natural organic matter present in the water, a number of chlorinated by-products can be formed, many of which are known or suspected carcinogens. Some of these disinfection by-products are regulated, and water plants need to meet strict standards to adhere to public health goals. In addition to formation of by-products, chlorine is not effective against some important protozoan pathogens. For these reasons, disinfection alternatives such as ultraviolet (UV) irradiation or ozone are supplementing chlorination. Excessive chlorine may be removed by dechlorination with sulfur dioxide. However, for protection of the distribution system, it is important to maintain some residual chlorine all the way to the consumer's tap. To aid in maintaining residual chlorine, ammonia is often added to create chloramines, which are more stable over time than free chlorine. The formation of chloramines has the added advantage of minimizing formation of by-products in the distribution system.

Ozonation and UV irradiation are the principal disinfection alternatives to chlorination. Ozone is a strong oxidant produced in a gaseous form and bubbled into water, where it dissolves. Ozone is highly reactive and is consumed quickly during treatment. The mechanism of ozone attack on pathogens is oxidation of the outer membrane or of the nucleic acids. Ozone is very effective against most pathogens; however, cold temperature can negatively affect its performance for some pathogens, such as *Cryptosporidium* oocysts. In the presence of bromide, ozonation of water can form bromate (BrO_3^-), which is a carcinogen. Bromate formation can be minimized through appropriate engineering design. UV

irradiation is produced from mercury vapor lamps, similar to household fluorescent lamps. These lamps are suspended in the water that flows through the system. A UV contact time of only a few seconds is needed for water disinfection. UV light emitted from the lamps in the wavelength range of 240–280 nanometers is very effective against most pathogens, with peak effectiveness around 260 nm, which coincides with the peak absorbance of deoxyribonucleic acid (DNA). The mechanism of action for UV disinfection is the transfer of UV energy from the photon to the microorganism's DNA, damaging sites that render the pathogen unable to reproduce, and thus effectively dead. UV has become of great interest to the water treatment industry for disinfection because it is extremely effective against the chlorine-resistant pathogen *Cryptosporidium*. UV treatment is not known to produce any by-products of health significance. One drawback with both ozone and UV energy is that they do not stay in the water to protect it as it moves from the treatment plant through the distribution system. *See* OZONE; ULTRAVIOLET RADIATION.

One of the functions of environmental engineers in protecting public health through water treatment is to balance the chemical risks from using too much chlorine with the microbial risks of not using enough. This balance of risks can be mitigated through the tailored addition of other disinfection processes, such as UV irradiation and ozone, which complement the effectiveness of chlorine. A multiple-barrier disinfection approach will likely be used more frequently in the future to help minimize by-products and optimize killing of all types of waterborne pathogens.

Granular activated carbon. Fine-grain activated carbon has been used for years to reduce objectionable tastes and odors, harmful organic compounds, and heavy metals in public water supplies. Since the capacity of granular activated carbon (GAC) for this use is great, no special facilities are normally designed for taste and odor control. Instead, the carbon is put on top of the filters or is used as the filter medium during rapid filtration.

Granulated activated carbon has been proposed by the Environmental Protection Agency (EPA) to eliminate or reduce potentially carcinogenic organic substances in public water supplies. The harmful substances, such as chlorinated organic by-products, are derived from naturally occurring humic and fulvic acids combining with chlorine and bromine. A nationwide survey of chlorinated public water supplies indicated many contained regulated and unregulated chlorinated by-products. As a step toward correcting the situation, the EPA has developed federal regulations limiting the maximum concentration of two classes of chlorinated by-products, trihalomethanes and haloacetic acids, to 0.08 and 0.06 mg per liter, respectively, in the distribution system. Communities not meeting these limitations must adjust their disinfection procedures in a manner that decreases the by-products to acceptable levels. Where such adjustments are found to be insufficient, granulated activated carbon facilities or disinfection alternatives are being installed. GAC facilities have been built at only a few places, and there is little reliable operating and cost data. *See* HALOGENATED HYDROCARBON; HUMUS.

Designing the contact chamber to pass water plant effluent through a bed of granulated activated carbon is relatively simple. The real job and cost of the process are in the periodic removal, cleaning, regeneration, and returning of the granulated activated carbon to the contact chambers, and adding new material to make up for losses. Several regeneration furnaces are available, including infrared furnace, stream, and multiple hearths. All of these involve the operation and maintenance of machinery and equipment, and a competent labor force. The frequency of regeneration varies greatly, since it depends not only upon how much organics must be removed, but upon what other natural or synthetic chemicals are present in the water that interfere with organics removal, and the type of granulated activated carbon used.

With these factors in mind, the overall cost of facilities and operations can be as high as the cost of conventional water treatment. Since changes in raw water quality may affect the performance of the granulated activated carbon, some irregularity can be expected. Design, construction, and installation of equipment and monitoring facilities should not be started without reliable, expert advice and some bench or pilot testing. *See* ACTIVATED CARBON.

Water softening. Water softening is the process of removing the "hardness" caused by calcium and magnesium salts. These salts make washing difficult, waste soap, and cause unpleasant scums and stains in households and laundries. They are especially harmful in boiler feedwater because of their tendency to form scales.

Municipal water softening is common where the natural water has a hardness in excess of 150 parts per million. Two methods are used: (1) The water is treated with lime and soda ash to precipitate the calcium and magnesium as carbonate and hydroxide, after which the water is filtered. (2) The water is passed through a porous cation exchanger which has the ability of substituting sodium ions in the exchange medium for calcium and magnesium in the water. The exchange medium may be a natural sand known as zeolite or may be manufactured from organic resins. It must be recharged periodically by backwashing with brine.

For high-pressure steam boilers or some other industrial processes, almost complete deionization of water is needed, and treatment includes both cation and anion exchangers. Lime-soda plants are similar to water purification plants, with coagulation, settling, and filtration. Zeolite or cation-exchange plants are usually built of steel tanks equipped for backwashing the media with salt brine. If the water is turbid, filtration ahead of zeolite softening may be required. *See* ION EXCHANGE; WATER SOFTENING; ZEOLITE.

Aeration. Aeration is a process of exposing water to air by dividing the water into small drops, by

forcing air through the water, or by a combination of both. The first method uses jets, fountains, waterfalls, and riffles; in the second, compressed air is admitted to the bottom of a tank through perforated pipes or porous plates; in the third, drops of water are met by a stream of air produced by a fan.

Aeration is used to add oxygen to water and to remove carbon dioxide, hydrogen sulfide, and taste-producing gases or vapors. Aeration is also used in iron-removal plants to oxidize the iron ahead of the sedimentation or filtration processes. *See* WATER POLLUTION.

Karl G. Linden; Richard Hazen

Pressure-driven membranes. Membranes act as selective barriers to pollutants, where some constituents pass through while others are blocked. To move materials across a membrane, a driving force, such as pressure, is needed. Pressure-driven membranes have recently been developed and implemented as a water treatment technology for separating suspended and dissolved materials from pure water. Pressure-driven membranes are categorized as low-pressure and high-pressure membranes. Low-pressure membranes include microfiltration and ultrafiltration. High-pressure membranes include nanofiltration and reverse osmosis.

Membranes can operate under positive or negative (vacuum) pressure and are typically characterized by their molecular weight cut-off (MWCO), which is the molecular mass of the chemical for which the membrane has a rejection capability of greater than 90%. They are sometime also characterized by pore size, referring to the diameter of the micropores in a membrane surface material, or the size of the particulate that can be retained. The **illustration** presents a number of different membranes and the types of pollutants that can be removed. The material that a membrane is made of can also affect removal of a contaminant due to chemical interactions.

Membranes are configured in modules, consisting of individual membrane fibers bundled together in a cartridge. A membrane treatment system has a number of modules arranged on a rack, sized by the number of modules per flow to be treated. In addition to the modules, membrane systems need to have provisions for backwashing and chemical cleaning. The membrane fibers may treat water in a cross-flow manner where water flows tangentially to the membrane surface, or a dead-end configuration where all water passes through the membrane. When the pressure across the membrane builds up, the membranes are backwashed and returned to the starting pressure. This backwashing may occur every 20 to 60 min for low-pressure membranes. Chemical cleaning is necessary when the pressure does not return to baseline values over time.

Microfiltration. Membranes that are used mainly for separation of particles, larger microorganisms, and minor removal of dissolved organic matter or pollutants are microfiltration (MF) membranes. These can replace or augment traditional sand media filtration. Depending upon the chemical characteristics of the membrane material, some dissolved matter may be removed. The pore size of microfiltration membranes is typically around 1 micrometer, which will remove larger pathogens.

Ultrafiltration. A low-pressure membrane with a smaller pore size is the ultrafiltration (UF) membrane. This membrane is also typically used for removing suspended and colloidal particles and microorganisms, but is more efficient at removing smaller materials with a size around 0.05 μm. Thus, ultrafiltration can remove bacteria and some larger viruses, as well as some dissolved matter. Sometimes ultrafiltration is used in combination with microfiltration, which provides preremoval of larger particles, minimizing

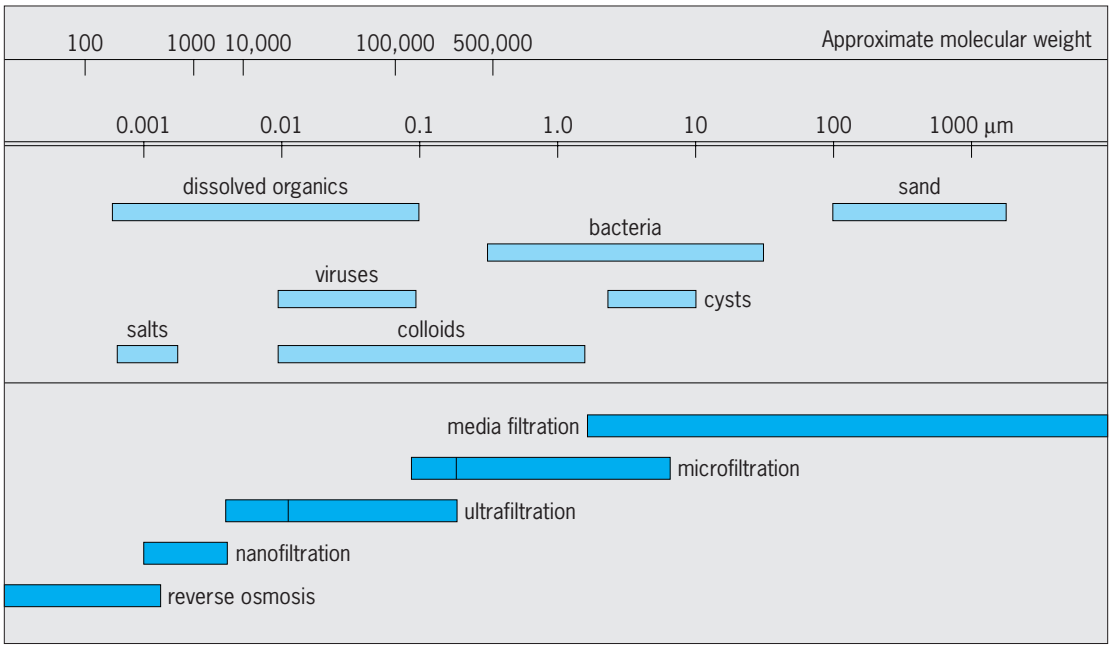

Molecular weight cut-off and pore sizes for membranes relative to various water quality concerns. (*EPA 815-C-01-001*)

clogging and frequency of backwashing and cleaning for the ultrafiltration membrane.

Nanofiltration. A high-pressure membrane that provides very good removal of many dissolved contaminants and viruses is the nanofiltration membranes. Nanofiltration (NF) removes molecules in the nanometer range and with a molecular weight greater than 1000 daltons. Thus, NF is capable of removing some elements associated with hardness (Ca and Mg), as well as bacteria and most viruses. Nanofiltration can also remove pesticides and other organic contaminants from water, as well as natural organics that can lead to the formation of chlorinated organic by-products. The charge of the particles can affect the removal efficiency, and small-molecule pollutants are typically not removed well.

Advanced treatment. Some treatment processes have specialty applications for removing or destroying contaminants that cause an undesirable taste and odor, natural compounds that may lead to formation of by-products during chlorination, and harmful organic or inorganic pollutants. Although not a threat to human health, taste- and odor-causing compounds, such as geosmin (musty odor) and methylisoborneol (MIB), derived from algae in water can lead to major esthetic problems with water and the perception that it is not fit to drink. If not removed, natural organic matter in surface waters can serve as precursor to the formation of chlorinated by-products. Organic contaminants, such as synthetically derived pesticides, pharmaceutical and personal care products, endocrine-disrupting contaminants, and industrial chemicals and solvents, are known or suspected to be harmful to human and aquatic health and are sometimes found in raw drinking water supplies. Inorganic contaminants, such as arsenic, perchlorate, nitrate, and chromium, can enter the water naturally or from a waste source. Because many of these contaminants can be difficult to remove using typical water treatment processes, advanced treatment is needed where their presence is a water-quality or health threat.

Ion exchange, for treating high-mineral-content waters, is most frequently used for softening. Ion exchange can also be used for deionization when ultrapure water is needed for industrial applications or removal of harmful inorganic pollutants such as arsenic. Reverse osmosis, electrodialysis, nanofiltration, and other membrane processes are especially applicable for treatment of water sources containing color and some complex organics or inorganics, including desalination. *See* MEMBRANE SEPARATIONS; WATER DESALINATION.

Chemical precipitation is used in water softening, iron and manganese removal, and other special cases for removal of heavy metals. Enhanced coagulation can be used to remove natural organic matter to reduce the level of chlorinated organic contaminants formed. Activated carbon (granular or powdered) applications are used for removal of dissolved natural and synthetic organics, taste and odor compounds, and color. *See* PRECIPITATION (CHEMISTRY).

Oxidation processes, such as ozone, potassium permanganate, and chlorination, are used at the beginning of a water treatment plant to control taste- and odor-causing compounds, reduce the microbiological load on the downstream treatment processes, and oxidize any reduced constituents. Ozone or advanced oxidation using hydroxyl radicals, produced from hydrogen peroxide in combination with either ozone or UV irradiation, may be implemented to destroy the numerous complex organic pollutants remaining in the water after filtration. *See* ENVIRONMENTAL ENGINEERING; WATER SUPPLY ENGINEERING.

Karl G. Linden

Bibliography. American Society of Civil Engineers and American Water Works Association, *Water Treatment Plant Design*, 4th ed., 2004; American Water Works Association, *Water Quality and Treatment*, 5th ed., 1999; L. S. Clescerl, A. E. Greenberg, and A. D. Eaton (eds.), *Standard Methods for the Examination of Water and Wastewater*, 20th ed., 1999; R. A. Corbitt (ed.), *Standard Handbook of Environmental Engineering*, 2d ed., 1999; M. Watson Harza, *Water Treatment: Principles and Design*, 2d ed., 2005.

Water-tube boiler

A steam boiler in which water circulates within tubes and heat is applied from outside the tubes. The outstanding feature of the water-tube boiler is the use of small tubes exposed to the products of combustion and connected to steam and water drums which are shielded from these high-temperature gases. Thus, possible failure of boiler parts exposed to direct heat transfer is restricted to the small-diameter tubes and, in the event of failure, the energy released is reduced and explosion hazards are minimized.

The water-tube construction facilitates greater boiler capacity by increasing the length and the number of tubes and by using higher pressure, since the

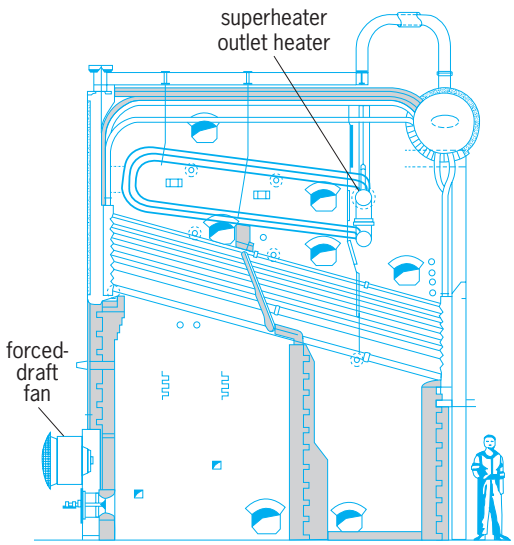

Fig. 1. Straight-tube-type boiler.

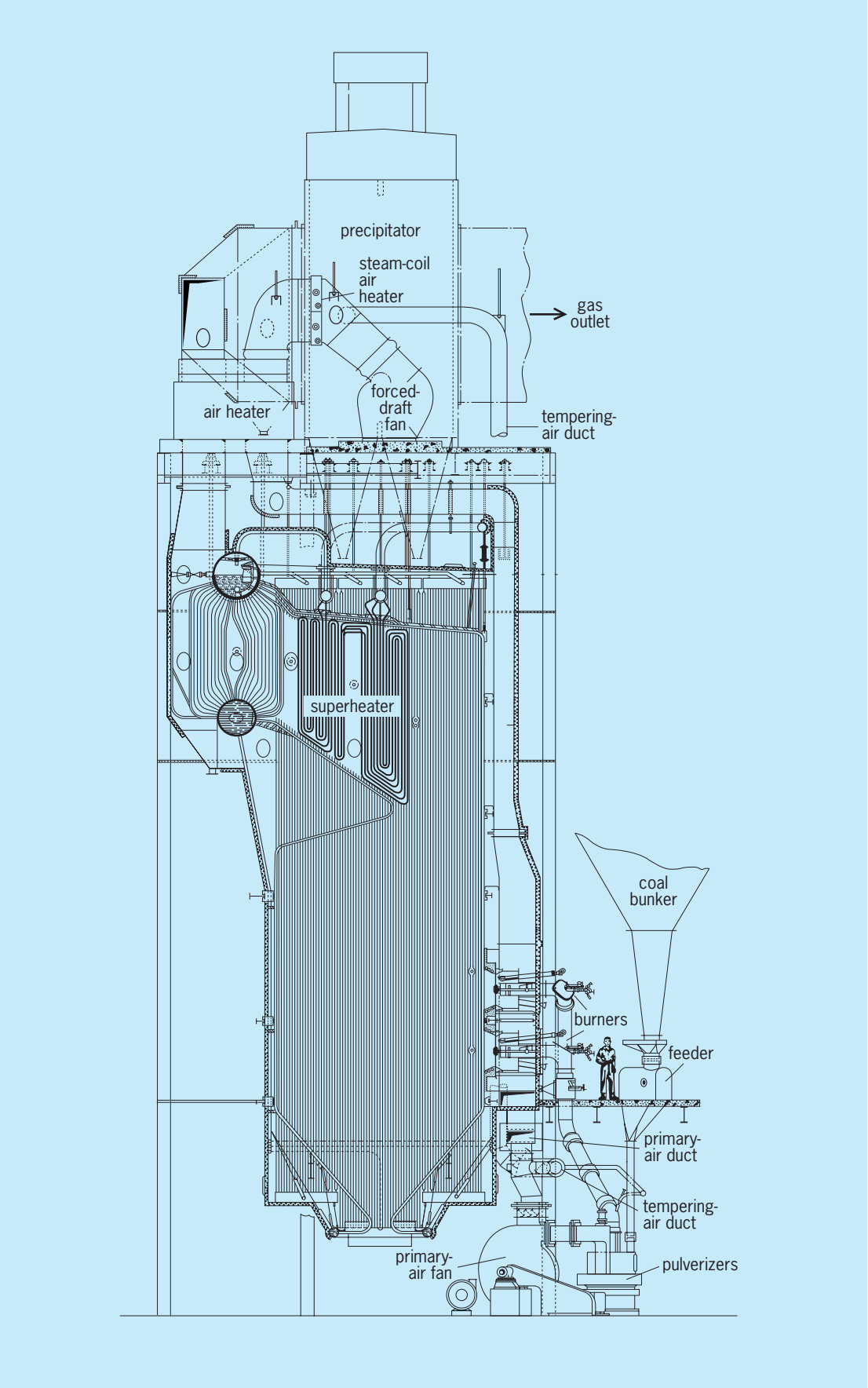

Fig. 2. Bent-tube-type boiler.

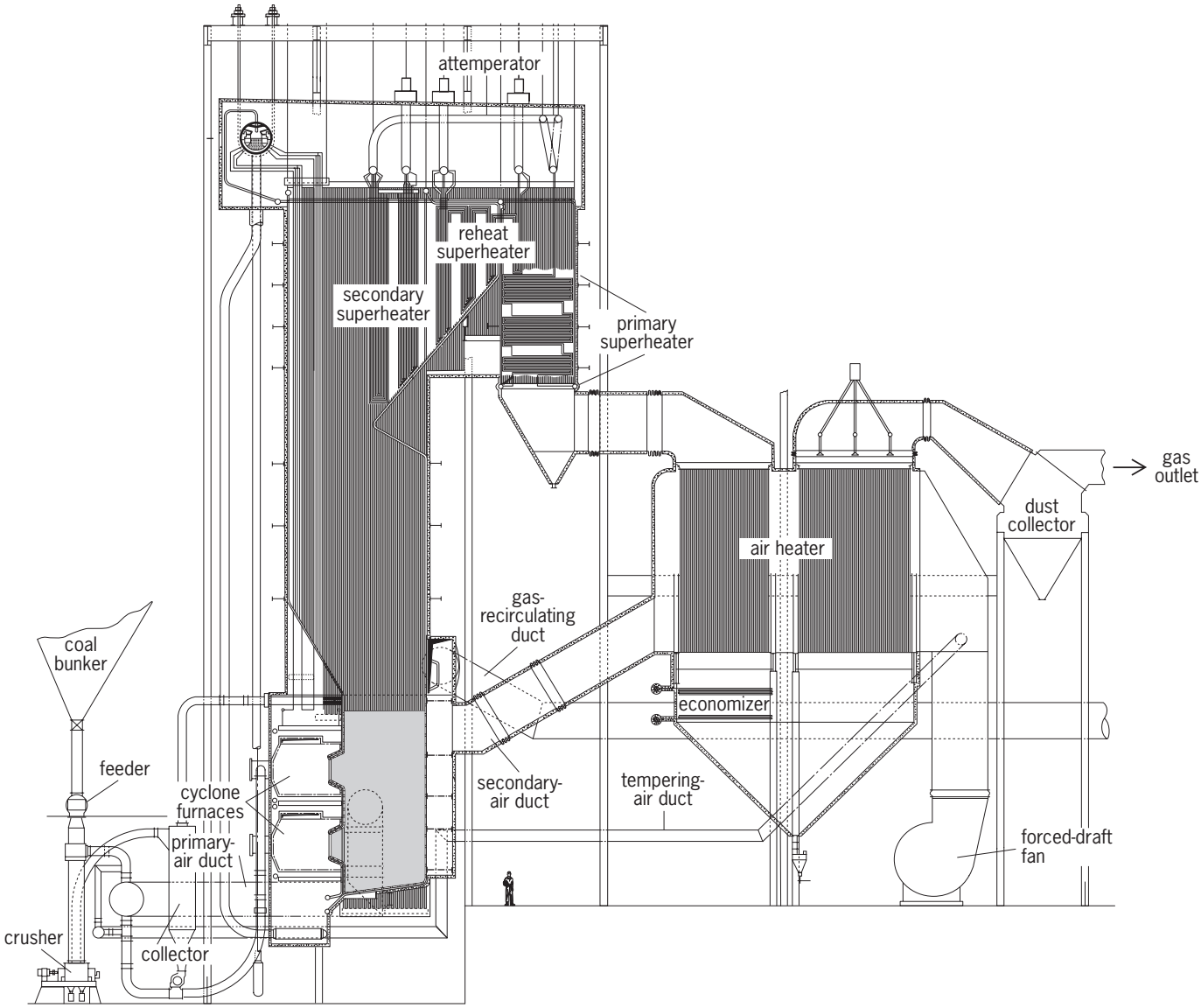

Fig. 3. Radiant-type boiler.

relatively small-diameter tubes do not require an abnormal increase in thickness as the internal pressure is increased. In addition, water-tube boilers offer great versatility in arrangement, and this permits efficient use of the furnace, superheater, and other heat-recovery components.

There are many types of water-tube boilers but, in general, they can be grouped into two categories: the straight-tube and the bent-tube types. In essence, both types consist of banks of parallel tubes which are connected to, or by, headers or drums. The early water-tube boilers were fitted with refractory furnaces. However, most modern water-tube boilers utilize a water-cooled surface in the furnace, and this surface is an integral part of the boiler's circulatory system. Further, modern water-tube boilers generally incorporate the use of superheaters, economizers, or air heaters to utilize more efficiently the heat from the fuel and to provide steam at a high potential for useful work in an engine or turbine. *See* BOILER ECONOMIZER; SUPERHEATER.

Straight-tube boiler. The straight-tube boiler (**Fig. 1**), often called the header-type boiler, has the advantage of direct accessibility for internal inspection and cleaning through handholes, located opposite each tube end, in the headers. The steam-generating sections are joined to one or more steam-and-water drums located above and parallel or transverse to the boiler tube bank. The circulation of water in the downcomer headers and of the water-steam mixture in the boiler tubes and riser headers is the result of the differential density between the water in the downcomers and the steam and water mixture in the heated boiler tubes and riser portion of the circuit. In many designs the tube bank can be baffled to increase the rate of flow of the products of combustion and, thus, improve heat transfer and the resultant absorption efficiency.

The straight-tube boiler is not applicable to high-pressure designs because of header limitations, and capacity is restricted by space requirements. Thus,

bent-tube instead of header-type boilers are used in most modern industrial installations.

Bent-tube boiler. In bent-tube boilers (**Fig. 2**), commonly referred to as drum-type boilers, the boiler tubes terminate in upper and lower steam and water drums which have few access openings. Although internal inspection is restricted, the tubes can be mechanically or chemically cleaned, and developments in water treatment and cleaning methods have overcome the early objections to the use of bent tubes. Circulation in drum-type boilers, as in header-type boilers, is due to the difference between the density of the water in the downcomer tubes and the water-steam mixtures in the riser tubes.

Because of the greater slope of tubes, or even the use of vertical tubes, in drum-type boilers, there is less possibility of steam pockets forming in the tubes and, thus, the rate of heat absorption can be increased appreciably. Consequently, drum-type boilers can be used for both low and high pressures and for capacities ranging from a few thousand to more than a million pounds of steam per hour. In addition, they can be designed to meet almost any space allotment and can accommodate heat-absorbing components such as superheaters, reheaters, economizers, and air heaters. An evolution of the bent-tube boiler is the radiant-type boiler in which the boiler tubes (steam-generating surface) form the boundary of the furnace and the containment for the superheater, reheater, and economizer surface (**Fig. 3**). Such boilers can deliver several million pounds of steam per hour at high steam pressures and temperatures. Further, by eliminating the steam drum and using a once-through flow of water and steam, they can operate at supercritical steam pressures. *See* BOILER; BOILER FEEDWATER; STEAM; STEAM-GENERATING FURNACE; STEAM-GENERATING UNIT.

George W. Kessler

Bibliography. E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; A. G. Blokh, *Heat Transfer in Steam Boiler Furnaces*, 1987; J. J. Jackson, *Steam Boiler Operation: Principles and Practices*, 2d ed., 1987.

Water tunnel (research and testing)

A hydrodynamic facility used for research, test, and evaluation, comprising a well-guided and controlled stream of water in which items for test are placed. The water tunnel is in many ways similar in appearance, arrangement, and operation to a subsonic wind tunnel. It is related to and complementary to the towing tank, in which the test item, usually a scale model of a ship or ship component, is towed through stationary water and evaluated through observation and measurement. In a water tunnel the test item is held stationary while the water is circulated around it. Many water tunnels are capable of operation with variable internal pressure to simulate the phenomenon of cavitation. *See* CAVITATION; TOWING TANK; WIND TUNNEL.

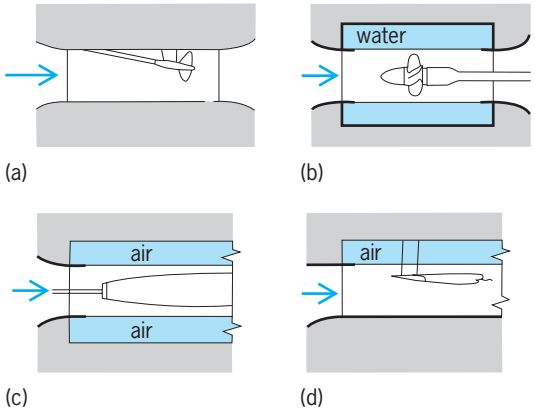

Fig. 1. Types of water-tunnel test sections with typical models. (*a*) Closed throat. (*b*) Open throat. (*c*) Free jet. (*d*) Free surface.

Classification. Water tunnels may be classified, in part, by the type of test section used. The most common section is the closed throat (**Fig. 1***a*) in which the test section flow has solid boundaries. The advantage of this arrangement is its simplicity and efficiency, but the model must be small relative to the tunnel cross section to avoid large wall effects. Small wall effects are theoretically correctable. In an open-throat test section (Fig. 1*b*) the water jet passes through a water-filled chamber of larger diameter. This minimizes wall effects, and many tunnels dedicated to propeller testing use such an arrangement. When very low test section cavitation numbers are required, or for fully cavitating flows, a free jet (Fig. 1*c*) in which the water jet passes through an air-filled chamber is useful. However, capture of the free jet and removal of excess entrained air prior to recirculation is not easily achieved over a broad range of test conditions. To study cavity flows on surface-piercing components or hydrofoils which operate near a free surface, a free-surface tunnel (Fig. 1*d*) is required in which three sides of the water flow are bounded by solid walls and the upper surface is open to air at controlled pressure levels. This arrangement is often referred to as a water channel. *See* HYDROFOIL CRAFT; OPEN CHANNEL.

Another distinction among water tunnel types is whether or not they recirculate the flow. If the tunnel is nonrecirculating, water may "blow down" from a pressurized or elevated water tank or water may be diverted from a continuous source such as a waterfall or dam on a river. The blow-down tunnel has a limited test period proportional to the size of the storage tank. All water tunnels make use of transparent test-section viewing windows.

Construction. The largest water tunnel in the world is the William B. Morgan Large Cavitation Channel, which is operated by the U.S. Navy and is shown in **Fig. 2**. Typical of most designs, it is of the recirculating, closed-throat type with a carefully designed contraction ahead of the test section. An axial-flow pump located in the lower horizontal leg impels water through the circuit and test section. The strong contraction ensures good velocity uniformity in the

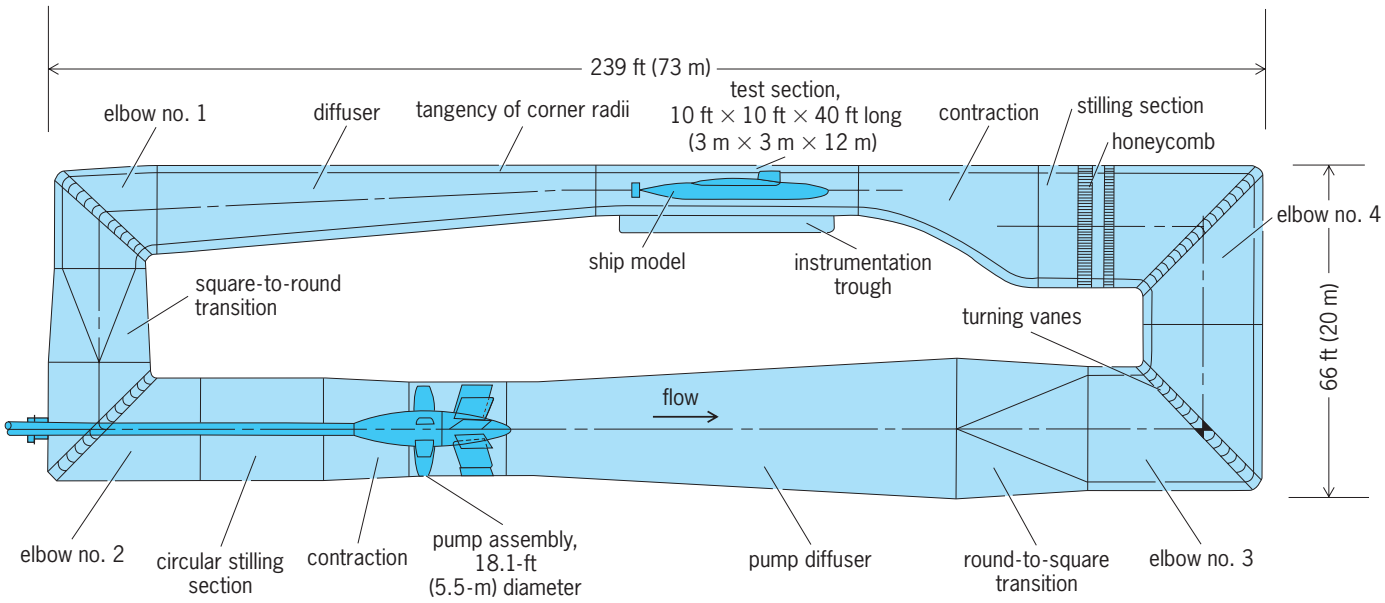

Fig. 2. Elevation of the William B. Morgan Large Cavitation Channel operated by the U.S. Naval Surface Warfare Center, Carderock Division. The water tunnel is located in Memphis, Tennessee.

test section, by the change of potential energy (pressure), which is uniform across the contraction entrance, to kinetic energy (velocity) of the test section. Turbulence is reduced by passage through a stilling section containing two honeycombs. The size of a water tunnel is generally specified in terms of the dimensions of the test section where measurements are performed. The tunnel shown in Fig. 2 has a rectangular test section of 10 ft by 10 ft by 40 ft long (3 m by 3 m by 12 m). *See* FLUID-FLOW PRINCIPLES; TURBULENT FLOW.

After passing through the text section, the velocity of the water in a recirculating tunnel is reduced to avoid cavitation and excess energy losses in other parts of the loop. Reduction of velocity takes place in the diffuser, along with an increase in pressure which helps to suppress possible cavitation of the turning vanes. Turning vanes serve to direct the flow uniformly and with minimal losses through a 90° turn at each elbow. Relatively uniform flow input to the pump is required to decrease the likelihood of undesirable vibration and pump blade cavitation.

Pressure control. To observe the cavitation characteristics of propellers and other test bodies, the water tunnel internal pressure must be varied independently of the velocity in the test section. Generally, for a model of a surface ship tested at speeds lower than full scale, it is necessary to decrease pressure to match full-scale cavitation conditions. If the model test speed is higher than full scale, the pressure must be increased. For submarines and similar applications which experience high full-scale operating pressures, it is also necessary to increase the tunnel pressure above the normal atmospheric level.

Closed- and open-throat tunnels. Most tunnels are of the closed- or open-throat type, operated with the entire test circuit filled with water. Consequently, they are well suited to investigations of completely submerged bodies or of conduits and turbomachinery such as pumps or turbines that are completely filled with water. Tunnels dedicated to turbomachinery are frequently referred to as test stands. Surface-ship hulls can be tested mounted flush to the top of a water tunnel test section, with the normally submerged portion of the hull exposed to the flow. The rigid wall or a specially constructed ground board takes the place of the free surface.

Free-surface tunnels. Often the behavior of the water surface around the ship has an important effect on the cavitation characteristics of the propeller or other appendages under test. For these cases, tests must be performed in the free-surface type of tunnel, or channel. The principal problem associated with free-surface tunnels is elimination of excess entrained air from the recaptured water downstream of the test section. For this reason, several free-surface tunnels have been designed to discharge water directly into a large reservoir after passage through the test section. A common practice for recirculating free-surface tunnels is to skim the top 10% of the flow from the test section and divert this to a reservoir, while recirculating the remainder after passage through a deaeration section.

Air content control. In closed or open types of tunnels, control of air content in the form of bubbles and dissolved gases is considered to be critical for the accurate replication of cavitation phenomena at model scale. Total gas content may be reduced by exposing the circulating water to reduced ambient pressure for an extended period of time. Dissolved air gradually comes out of solution to form bubbles which can be drawn off. Residual bubbles, or bubbles generated by the test body, can be removed during circulation in the tunnel by two processes: resorption and collection. Resorption takes place when part of the circuit is at a much lower height (higher static pressure) than the test section. The increased pressure forces free bubbles into solution. An alternative method is to slow the flow to a very low speed just upstream or downstream of the test section, and to collect free

bubbles in trays or tubes and guide them to air evacuation points. Total removal of dissolved gas and bubbles is not desired for cavitation tests. Rather, a stable bubble population which allows good correlation of cavitation inception and cavitation patterns with realistic prototype conditions is sought. Although not a feature of most existing water tunnels, aeration devices to inject desired bubble populations are being incorporated into the design of some tunnels. Because bubbles are elastic systems, they can absorb acoustic energy. For measurements of noise generated by noncavitating bodies, it is especially important to control that part of the free bubble population which is responsible for absorption of radiated noise in the frequency range of interest.

Applications. The water tunnel is applied in much the same way as a subsonic wind tunnel, the principal difference being the possibility of vaporization of the water due to local decrease in pressure (cavitation) or the existence of a free surface and attendant interface phenomena such as waves, ventilation, and spray. Water tunnels are used to investigate the dynamics, hydrodynamics, and cavitation of submerged and semisubmerged bodies such as propellers, ships, submarines, torpedoes, and hydrofoils and of turbomachinery. They are also indispensable to research on general flow phenomena in liquids. An application of great importance has been the acoustic characterization of propellers under both cavitating and noncavitating conditions. This requires low background-noise level in the frequency range of interest. Thus, turning vanes, diffuser, and impeller must be designed to strict acoustic as well as hydrodynamic criteria. *See* HYDRODYNAMICS.

The first use of water tunnels was to identify the onset of thrust breakdown of propellers due to excessive cavitation. This is still an important application, extending to other types of marine propulsors and turbomachinery as well. In a typical test of a propeller, the thrust and torque are measured as a function of rate of rotation and speed of advance through the water, for various pressures. Cavitation inception is observed, visually using stroboscopic lighting and acoustically using hydrophones. When this information is combined with that derived from towing-tank tests of the hull alone and in combination with a propeller, the thrust, torque, and power characteristics of the full-scale ship, including the operating point of best efficiency, can then be predicted. *See* HYDROPHONE; PROPELLER (MARINE CRAFT); STROBOSCOPIC PHOTOGRAPHY.

In sufficiently large water tunnels, such as the one in Fig. 2, the propeller can be tested with a completely appended model hull. This provides a more realistic operating environment. It also enables pressures induced on the hull by the propeller to be measured, and full-scale vibration levels to be predicted. The acoustic radiation of the propeller is also influenced by the presence of the hull. Noise and vibration characteristics of the hull-propeller combination are important for military surface ships and submarines because of the widespread use of acoustic detection techniques. For merchant ships, noise and vibration are important for habitability and avoidance of maintenance of vibration-sensitive components. *See* ACOUSTIC NOISE; MECHANICAL VIBRATION.

Limitations. The limitations of water tunnels in some respects are similar to those of wind tunnels. To scale dynamic effects (such as the flow induced forces) between the model and full-scale test objects, it is desirable to match the Reynolds numbers, Re, between the model and full-scale flows. The Reynolds number is defined as $\mathrm{Re} = UL\rho/\mu$, where U is a reference flow speed (typically the average upstream flow speed), L is a reference length (typically based on the size of the object being considered), and ρ and μ are the liquid density and dynamic viscosity, respectively. For flows with liquid-gas interfaces, the Froude number, Fr, which is defined as $\mathrm{Fr} = U/(gL)^{1/2}$, may also be considered. Here, g is the gravitational acceleration. In general, it is not possible to match both Re and Fr of test and full-scale flows, but is it often sufficient to conduct the model test above a threshold value of the Reynolds number while attempting to match the Froude number. *See* FROUDE NUMBER; REYNOLDS NUMBER.

Since the model scale is usually smaller in size than the full scale, the model-scale flow velocity might be increased to achieve dynamic similarity. Besides the obvious increase in power that increased velocity requires, it becomes necessary to increase pressure to maintain cavitation similitude. The cavitation number, σ, is defined as $\sigma = (P - P_v)/(1/2)\rho U^2$, where P is a reference pressure (typically the static pressure of the flow upstream of the test object), and P_v is the vapor pressure of the liquid. To achieve cavitation similarity, it is desirable to match the cavitation number between the model and full-scale flows. The need to increase the pressure at the model scale increases the required strength (shell thickness) of the water tunnel structure. This pressure constraint may be more restrictive than the power constraint, as pressure must rise with the square of the velocity. Therefore, water channel speeds are generally restricted to less than 50 knots (25 m/s) for relatively large tunnels. *See* DIMENSIONLESS GROUPS; DYNAMIC SIMILARITY; MODEL THEORY. Richard Stone Rothblum; Robert J. Etter; Steven L. Ceccio

Bibliography. C. E. Brennen, *Cavitation and Bubble Dynamics*, Oxford University Press, 1995; W. F. Brownell, *Two New Hydromechanics Research Facilities at the David Taylor Model Basin*, DTMB Rep. 1690, December 1962; R. J. Etter and M. B. Wilson, The large cavitation channel, in *Proceedings of the 23rd American Towing Tank Conference*, New Orleans, 1992; R. J. Etter et al., High Reynold's number experimentation in the U.S. Navy's William B. Morgan Large Cavitation Channel, *Meas. Sci. Technol.*, 16:1701–1709, 2005; H. D. Harper, A modernized control system for the DTNSRDC 36 inch variable pressure water tunnel, *Proceedings of the 20th American Towing Tank Conference*, Hoboken, 1983; E. V. Lewis (ed.), *Principles of Naval Architecture*, 3 vols., 2d ed., Society of Naval Architects and Marine Engineers, 1990; H. Lindgren and E. Bjärne, *Ten Years of Research in the SSPA Large Cavitation Tunnel*, SSPA Publ. 86, 1980.

Watermelon

The edible fruit of *Citrullus lanatus*, of the family Cucurbitaceae. The plant is an annual prostrate vine with multiple stems alternating from short nodes near the base of the main axis, reaching lengths of 10 to 15 ft (3 to 4.5 m). Short vine or "bush" type cultivars exist but are not commercially useful. In a typical monoecious plant, pistillate flowers occur at every seventh node, with staminate flowers at intervening nodes. Cultivated plants range from 250 to 1000 hills per acre (625 to 2500 hills per hectare) according to variety and region. Commercially, an average yield of one mature melon per plant is acceptable.

Varieties and production. The numerous open-pollinated cultivars of watermelon are highly diverse in fruit size (5–85 lb or 2.3–38.3 kg), shape (round, oval, oblong-cylindrical), rind color (very light to very dark green and often striped or mottled), flesh color (red, pink, orange, yellow, white), and seed size and color. The flesh contains 6–12% sugar, depending upon variety and condition of growth, with 8% sugar being acceptable on most markets, corresponding to juice refractometer readings of 9 or 10% total soluble solids. A few low-sugar cultivars called citrons are planted only for livestock. Such citron types grow wild in parts of Africa and are probably a significant factor in the survival of native animals in drought periods.

Seedless watermelons were introduced in 1948 in Japan. These are more difficult and costly to produce, yet they have become popular in some Asian countries. Seedless triploid watermelons are notably superior in flesh quality. Equally superior flesh quality occurs in some tetraploid varieties originated in the United States which are semiseedless and have extra firm flesh and hard rinds. Due to the extra stable flesh quality, these crops may be left on the vine for a single harvest.

The watermelon species is native to Africa, where it still occurs in the wild state in semiarid areas and is also widely cultivated by native farmers, both for human consumption and for livestock. Watermelons are grown in most tropical and subtropical and many temperate climates. They are grown locally in most of the United States, but in large volumes for interstate shipment primarily in the southern states and in California, Indiana, and Texas.

Food products. Watermelon juice of the sweet cultivars can be reduced to edible sugar and syrup. Watermelon rinds are either hard or soft, and this affects the handling performance of different cultivars. Varieties that best resist breakage in transit have firm interior flesh although the rind may be neither very thick nor very hard. Sweet pickles made from rinds have some commercial importance. Soft-rind varieties are best for that purpose.

Watermelon seeds are relished as food in some Near East countries and in China, and certain varieties are grown primarily for that purpose. There is a small industry in Iran where grilled (roasted) watermelon seeds are bagged and sold like popcorn. Salt-preserved watermelon seeds are often eaten in China.

Health value. Fresh watermelon flesh has a unique melting quality that has proved impossible to preserve in palatable form through any processing technique. The flesh consists of water (91%), fiber, and sugar, and little else of obvious nutritional value, but may have some as yet unproved health-promoting value. It is said to have diuretic properties, and both frozen concentrate and canned juice have been available for the treatment of nephritis. The seeds also are said to contain substances effective in the control of hypertension. C. F. Andrus

Diseases. Over 35 diseases of watermelon have been described, but only 10 are of economic importance.

Soil-borne diseases. Fusarium wilt, caused by the fungus *Fusarium oxysporum* f. sp. *niveum*, occurs wherever watermelon is grown. The symptoms are wilting of the foliage and a brown discoloration of the vascular tissue in the taproot. *Fusarium* spp. and other fungi also cause damping-off, a disease of seedling watermelons resulting in their collapse and death. Root-knot nematodes (*Meloidogyne* spp.) are endoparasitic wormlike organisms that cause gall-like swelling on the roots.

Foliar diseases. Gummy stem blight and anthracnose occur on leaves, stems, and fruits. The fungus causing gummy stem blight, *Mycosphaerella citrullina*, produces large leaf lesions and is common on the stem and crown, while anthracnose, caused by the fungus *Glomerella cingulata* var. *orbiculare*, is particularly severe on the fruit (see **illus.**). Three fungal diseases restricted to the leaves are downy

Anthracnose lesions on a watermelon fruit.

mildew, caused by *Pseudoperonospora cubensis*, Cercospora leaf spot, caused by *Cercospora citrullina*, and Alternaria leaf spot, caused by *Alternaria cucumerina*. All three pathogens produce brown leaf spots with either large yellow borders (downy mildew), white lesion centers (Cercospora), or concentric rings of alternating light and dark brown tissues (Alternaria).

Watermelon mosaic is an aphid-transmitted disease caused by two viruses. Watermelon mosaic virus (WMV) 1 infects only cucurbits, while WMV 2 also infects legumes. Both produce mottling, distortion, and mosaic patterns on leaves and fruits.

Control. To control fusarium wilt and root-knot nematodes, a long-term rotation of 4–5 years between crops is recommended. A seed treatment with fungicidal chemicals is recommended to control damping-off. Most foliar diseases are controlled with fungicidal sprays that are applied to the leaves. Resistant cultivars are available for the control of fusarium wilt and anthracnose. There are no known economic controls for watermelon mosaic viruses. *See* PLANT PATHOLOGY.

N. C. Schenck

Bibliography. W. C. Adlerz, Spring aphid flights and incidence of watermelon mosaic 1 and 2 in Florida, *Phytopathology*, 64:350–353, 1974; D. L. Hopkins, Fungicidal control of downy mildew and gummy stem blight of watermelon, *Proc. Flor. State Hort. Soc.*, 85:108–110, 1972; O. C. Maloy, *Plant Disease Control: Principles and Practice*, 1993; G. K. Parris, *Diseases of Watermelon*, Flor. Agr. Exp. Sta. Bull. 491, 1952; N. C. Schenck, Watermelon disease incidence in central Florida, *Plant Dis. Rep.*, 44:556–558, 1960.

Waterpower

Power developed from movement of masses of water. Such movement is of two kinds: (1) the falling of streams through the force of gravity, and (2) the rising and falling of tides through lunar (and solar) gravitation.

While that part of solar energy expended to lift water vapor against Earth gravity is a minute fraction of the total, the absolute amount of energy that is theoretically recoverable from resulting streams is an enormous but unknown quantity. Of this, but a tiny portion is actually suitable for harnessing.

Water resources. The contribution of waterpower installations to the nation's electric power supply at the beginning of World War II was about 30%. While the output from hydroelectric plants has grown, their contribution to the total electric power has dropped to about 13%, because steam-electric plants have grown at a much more rapid rate.

As of March 1980, the Energy Information Agency of the Department of Energy estimated that by the end of 1993 the total developed capacity of waterpower installations could be 99,346 MW. However, that figure does not include a number of potential sites that could at some future time be considered. In theory, the Energy Information Agency estimates an eventual maximum potential of 187,000 MW, most of the undeveloped sites being in the Pacific Northwest and Alaska. Only a fraction of that will be developed for a variety of reasons. The most attractive sites have already been utilized. Hydro plants, with their initial high cost and generally long distances from major load centers, must compete with the large, efficient, fuel-fired stations, and the burgeoning, economical, large nuclear plants. Large dam sites usually must be justified not alone on the value of the power developed, but also on the benefits from flood control, irrigation, and recreation. Problems of migrating fish, conservation, and preservation of esthetic values are also factors. On the other hand, waterpower developments add greatly to power-system flexibility in meeting peak and emergency loads. Modern excavation and tunneling techniques are lowering construction costs. The economies of lowhead sites are improved by the new, efficient, axial-flow turbines of the tubular type. *See* ENERGY SOURCES.

Silting. The capacity of hydro plants cannot be counted on for perpetuity because of gradual filling of reservoirs with sediment. This effect is serious for irrigation, flood control, and navigation. Even when a lake behind a power dam becomes filled completely with silt, electric power can be generated on the run-of-the-river flow, although output would vary with stream flow.

The rate of silting varies widely with drainage basins. Because the Columbia River carries comparatively little silt, the reservoirs at Grand Coulee and Bonneville dams should have lives of many hundreds of years. The Colorado River, on the other hand, is muddy. In the first 13.7 years after Hoover Dam went into operation in 1935, 1424×10^6 acre-feet (175,600 hectare-meters) of silt was dumped into Lake Mead. That is equivalent to a layer 1 ft deep over 2225 mi^2 (or 1 m deep over 1756 km^2). This inflow of silt has been diminished about 22% by the construction of other dams upstream, for example, the Glen Canyon Dam. It is expected that Lake Mead will have a useful life of more than 500 years.

Pumped storage. In pumped-storage hydroelectric systems, water is pumped from a stream or lake to a reservoir at a higher elevation. Pumping up to a storage reservoir is most commonly done by reversing the hydraulic turbine and generator. The generator becomes a motor driving the turbine as a pump. Power is drawn from the power system at night or on weekends when demand is low. It is not practical to shut down large, high-temperature steam stations or nuclear units for a few hours at night or even over a weekend. Because they must run anyway, the cost of pumping power is low, whereas the power generated from pumped storage at peak periods is valuable. Also, the pumped-storage system provides a means of supplying power quickly in an emergency situation, for example, during the failure of a large

steam or nuclear unit. A pumped-storage system can be changed over from pumping to generation in 2 to 5 min. *See* PUMPED STORAGE.

Tidal power. A portion of the kinetic energy of the rotation of the Earth appears as ocean tides. The mean tide of all the oceans has been calculated as 2 ft (0.6 m), and the mean power as 5.4×10^{10} hp (40 TW) or, on a yearly basis, the equivalent of 3.6×10^3 kWh (4×10^{19} joules). Unfortunately, only a minute amount of this is likely to be harnessed for use. For tidal sites to be of sufficient engineering interest, the fall would have to be at least 15 ft (4.5 m). There are few such falls, and some of these are in remote areas. The only tide-power sites that have received serious attention are on the Severn River in England, the Rance River and Mont St. Michel in northern France, the San José and Deseado rivers of Argentina, the Petitocodiac and Memram-cook estuaries in the Bay of Fundy, Canada, the Passamaquoddy River where Maine joins New Brunswick, Canada, and the Cambridge Gulf of Western Australia.

The Passamaquoddy site, which has a potential of 1800 MW (peak), is the only important tidal-power prospect in the United States. However, engineers do not consider its electrical output to be economically competitive with power that is produced by other means.

A second major handicap to tidal power is that, with a simple, single-basin installation, power is available only when there is a several-feet difference between levels in the sea and the basin. Thus, firm power is not available. Also, periods of generation occur in consonance with the tide—not necessarily when power is needed.

The only major tidal power plant in operation is the one near the mouth of the Rance River in Normandy, France. This plant operates on 40-ft (12−m) tides. It began operation in 1967. It consists of twenty-four 10-MW bulb-type turbine-generator units of novel design. The system embodies a reservoir into which sea water is pumped during off-peak hours. Turbines are then run as pumps, power being drawn from the French electrical grid. The plant produces 5×10^8 kWh (1.8×10^{15} J) annually, including a significant amount of firm power.

Tidal power is an appealing and dramatic technique, and some other large plants may be constructed. However, the total contribution of the tides to the world's energy supply will be miniscule. *See* HYDROLOGY; TIDAL POWER. Charles A. Scarlott

Application. The basic relation for power P in kilowatt from a hydrosite is $P = QH/11.1$, where Q is water flow in ft^3/s, and H is head in feet. Actual power will be less as occasioned by inefficiencies such as (1) hydraulic losses in conduit and turbines; (2) mechanical losses in bearings; and (3) electrical losses in generators, station use, and transmission. Overall efficiency is always high, usually in excess of 80% to the station bus bars.

Choice of site. The competitive position of a hydro project must be judged by the cost and reliability of the output at the point of use or market. In most hydro developments, the bulk of the investment is in structures for the collection, control, regulation, and disposal of the water. Electrical transmission frequently adds a substantial financial burden because of remoteness of the hydrosite from the market. The incremental cost for waterwheels, generators, switches, yard, transformers, and water conduit is often a smaller fraction of the total investment than is the cost for the basic structures, real estate, and transmission facilities. Long life is characteristic of hydroelectric installations, and the annual carrying charges of 6–12% on the investment are a minimum for the power field. Operating and maintenance costs are lower than for other types of generating stations.

The fundamental elements of potential power, as given in the equation above, are runoff Q and head H. Despite the apparent basic simplicities of the relation, the technical and economic development of a hydrosite is a complex problem. No two sites are alike, so that the opportunity for standardization of structures and equipment is nearly nonexistent. The head would appear to be a simple surveying problem based largely on topography. However, geologic conditions, as revealed by core drillings, can eliminate an otherwise economically desirable site. Runoff is complicated, especially when records of flow are inadequate. Hydrology is basic to an understanding of water flow and its variations. Runoff must be related to precipitation and to the disposal of precipitation. It is vitally influenced by climatic conditions, seasonal changes, temperature and humidity of the atmosphere, meteorological phenomena, character of the watershed, infiltration, seepage, evaporation, percolation, and transpiration. Hydrographic data are essential in order to show the variations of runoff over a period of many years. Reservoirs, by providing storage, reduce the extremes of flow variation, which are often as high as 100 to 1 or occasionally 1000 to 1.

Economic factors. The economic factors affecting the capacity to be installed, which must be evaluated on any project, include load requirements, runoff, head, development cost, operating cost, value of output, alternative methods of generation, flood control, navigation, rights of other industries on the stream (such as fishing and lumbering), and national defense. Some of these factors are components of multipurpose developments with their attendant problems in the proper allocation of costs to the several purposes. The prevalence of government construction, ownership, and operation, with its subsidized financial formulas which are so different from those for investor-owned projects, further complicates economic evaluation. Many people and groups are parties of interest in the harnessing of hydrosites, and stringent government regulations prevail, including those of the U.S. Corps of Engineers, the Federal Power Commission, the Bureau of Reclamation, the Geological Survey, and the Securities and Exchange Commission.

Capacity. Prime capacity is that which is continuously available. Firm capacity is much larger and is dependent upon interconnection with other power plants and the extent to which load curves permit variable-capacity operation. The incremental cost for additional turbine-generator capacity is small, so that many alternatives for economic development of a site must be considered. The alternatives include a wide variety of base load, peak load, run-of-river, and pumped-storage plants. All are concerned with fitting installed capacity, runoff, and storage to the load curve of the power system and to give minimum cost over the life of the installation. In this evaluation it is essential clearly to distinguish capacity (kW) from energy (kWh) as they are not interchangeable. In any practical evaluation of water power in this electrical era, it should be recognized that the most favorable economics will be found with an interconnected electric system where the different methods of generating power are complementary as well as competitive.

As noted above, there is an increasing tendency in many areas to allocate hydro capacity to peaking service and to foster pumped-water storage for the same objective. Pumped storage, to be practical, requires the use of two reservoirs for the storage of water—one reservoir at considerably higher elevation, say, 500 to 1000 ft (150 to 300 m). A reversible pump-turbine operates alternatively (1) to raise water from the lower to the upper reservoir during off-peak periods, and (2) to generate power during peak-load periods by letting the water flow in the opposite direction through the turbine. Proximity of favorable sites on an interconnected electrical transmission system reduces the investment burden. Under such circumstances the return of 2 kWh on-peak for 3 kWh pumping off-peak has proven to be an attractive method of economically utilizing interconnected fossil-fuel, nuclear-fuel, and hydro power plants. *See* ELECTRIC POWER GENERATION; HYDRAULIC TURBINE; POWER PLANT; NUCLEAR REACTOR.

Theodore Baumeister

Bibliography. American Society of Civil Engineers, *Civil Engineering Guidelines for Planning & Designing Hydroelectric Developments*, vol. 5: *Pumped Storage and Tidal Power*, 1989; Annual statistical report for 1980, *Elec. World*, 193(6):49–80, March 15, 1980; E. A. Avallone and T. Baumeister III (eds.), *Marks' Standard Handbook for Mechanical Engineers*, 10th ed., 1996; B. W. Clowes (ed.), *Waterpower Eighty-seven*, 3 vols., 1988; A. J. Eberhardt (ed.), *Waterpower, 1989*, 3 vols., 1989; Edison Electric Institute, *Statistical Year Book of the Electric Utility Industry*, annually; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 1999; M. G. Jog, *Hydroelectric and Pumped Storage*, 1989; D. Marier and L. Stoikaen (eds.), *Alternative Sources of Energy-Hydro*, No. 76, 1985; L. Moniton, M. LeNir, and J. Roux, *Micro Hydro-Electric Power Stations*, 1985; J. Raabe, *Hydro Power*, 1984; M. J. Roluti (ed.), *Waterpower, Nineteen-Eighty-Five*, 3 vols., 1986; *Water for Energy: Proceedings of the 3d International Symposium on Wave, Tidal, OTEC, & Small Scale Hydro Energy*, 1986.

Waterspout

An intense columnar vortex (not necessarily containing a funnel-shaped cloud) of small horizontal extent, over water. Typical visible vortex diameters are of the order of 33 ft (10 m), but a few large waterspouts may exceed 330 ft (100 m) across. In the case of Florida waterspouts, only rarely does the visible funnel extend from parent cloudbase to sea surface. Like the tornado, most of the visible funnel is condensate. Therefore, the extension of the funnel cloud downward depends upon the distribution of ambient water vapor, ambient temperature, and pressure drop due to the vortex circulation strength.

These vortices are most frequently observed during the warm season in the oceanic tropics and subtropics. In particular, the probable world maximum frequency for waterspout activity is over the Florida Keys. Another very prolific waterspout region is the entire southeast Florida Coast, although Tampa Bay has had the greatest number of damaging waterspouts. During the 1969 Lower Florida Keys Waterspout Project, nearly 500 separate waterspout events were documented for the period May to September.

All waterspouts undergo a regular life cycle composed of five discrete but overlapping stages.

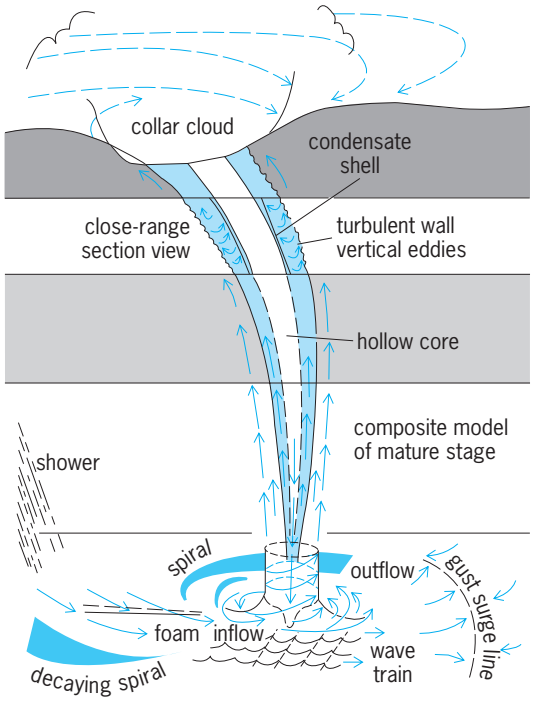

Fig. 1. Composite schematic model of a mature waterspout. For scaling reference, the maximum funnel diameters in this stage, just below the collar cloud, range from 10 to 460 ft (3 to 140 m).

Fig. 2. Decaying waterspout (in life-cycle stage 5) with the visible funnel retracting from the sea surface upward to parent cloudbase.

(1) The dark-spot stage is characterized by a prominent light-colored disk on the sea surface, surrounded by a dark patch diffuse on its outer edges—the dark spot may or may not have a small funnel above it initially, but signifies a complete vortex column extending from cloud-base to sea surface. (2) The spiral-pattern stage, the primary growth phase of the waterspout, is characterized by development of alternating dark- and light-colored bands spiraling around the dark spot on the sea surface. (3) The spray ring (incipient spray vortex) stage is characterized by a concentrated spray ring around the dark spot, with a lengthening funnel cloud above. (4) The mature waterspout stage (**Fig. 1**) is characterized by a spray vortex of maximum intensity and organization, the gradual weakening of the spiral pattern, and maximum funnel cloud length and diameter. (5) The decay stage occurs when the waterspout dissipates (often abruptly) as it is intercepted by the cool downdrafts from a nearby rain shower (**Fig. 2**).

Not every waterspout observed from its inception evolved through all other stages; however, the dark-spot stage and the mature waterspout stage have the greatest range of duration and often make up the bulk of a waterspout's lifetime (8–15 min average). In general, the larger-diameter waterspouts tend to be longer-lived and more intense; indeed, for south Florida waterspouts, about 10% or more each year equal or exceed tornadic size and intensity (maximum windspeeds just outside the calm "eye" of the spray vortex above 165 ft/s or 50 m/s). On the other hand, only about 1 out of 10 waterspouts in this region rotates anticyclonically (clockwise). Most waterspouts are spawned by cloudlines, which initially have average cloudtops of only 12,000–15,000 ft (3500–4500 m) in the Florida Keys. Waterspouts and tornadoes are qualitatively similar, differing only in certain quantitative aspects: tornadoes are usually more intense, move faster, and have longer lifetimes—especially maxi-tornadoes. Tornadoes are associated with intense, baroclinic (frontal), synoptic-scale disturbances with attendant strong vertical wind shear, while waterspouts are associated with weak, quasibarotropic disturbances (weak thermal gradients) and consequent weak vertical wind shear. *See* TORNADO; WIND.

Joseph H. Golden

Bibliography. J. H. Golden, An assessment of waterspout frequencies along the U.S. East and Gulf Coasts, *J. Appl. Meteorol.*, 16:231–236, 1977; J. H. Golden, The life-cycle of the Florida Keys waterspout, I, *J. Appl. Meteorol.*, 13:676–692, 1974; J. H. Golden, The lower Florida Keys waterspout project, May–September 1969, *Bull. Amer. Meteorol. Soc.*, 51:235–236, 1970; J. H. Golden, Scale interaction implications for the waterspout life cycle, II, *J. Appl. Meteorol.*, 13:676–692, 1974; J. H. Golden, Some statistical aspects of waterspout formation, *Weatherwise*, 26:108–117, 1973; J. H. Golden and D. Purcell, Life cycle of the Union City, Oklahoma, tornado and comparison with waterspouts, *Mon. Weath. Rev.*, 106(1):3–11, 1978.

Watt balance

An electromechanical apparatus for establishing the watt as an SI electrical unit. Prior to January 1, 1990, there were voltage units in use in the United States, France, the United Kingdom, Russia, and so on which differed from each other (and, with hindsight, from the SI volt) by up to 9 parts per million (ppm). These units were based on results from various current balances. Then results obtained from a different kind of apparatus at the National Physical Laboratory in the United Kingdom and the National Institute for Science and Technology in the United States, which were based on the simple principle described below, enabled the electrical units to be put on a sound SI basis. The accuracy in deriving the SI unit of voltage in this way was considered to be better than 0.2 ppm. Since that date these apparatuses have continued to be refined, and others are being developed in metrological laboratories, with the more ambitious objective of defining the kilogram in terms of fundamental physical constants instead of having to rely on a carefully preserved artifact, the cylinder of platinum-iridium alloy kept at the International Bureau of Weights and Measures. To do this, the accuracy achieved by the apparatus will have to be demonstrated to be of the order of 0.01 ppm. *See* CURRENT BALANCE; ELECTRICAL UNITS AND STANDARDS.

Working principle. A conductor whose total length is l is wound into a horizontal circular coil and placed in a horizontal magnetic flux B which is everywhere perpendicular to the coil (see **illustration**). If a current I flows in the conductor, the consequent vertical force can be opposed by a mass M which is subject to the earth's gravitational acceleration, g. That is, Eq. (1) is satisfied.

$$Mg = BIl \tag{1}$$

See MAGNETISM.

If, in a separate measurement, the coil is moved vertically with velocity u, a voltage V will appear

across it and will be given by Eq. (2).

$$V = Blu \tag{2}$$

See ELECTROMAGNETIC INDUCTION; FARADAY'S LAW OF INDUCTION.

Combining Eqs. (1) and (2) eliminates the product Bl, which is too difficult to measure with the required accuracy, and gives Eq. (3). This equation

$$VI = Mgu \tag{3}$$

relates the product VI to Mgu watts as measured directly in terms of the SI units of mass, length, and time.

The ambition of the present effort is to modify the SI by defining voltage and current units in terms of the Josephson effect constant, $2e/h$, and the quantum Hall effect constant, h/e^2. The effect of this would be to redefine the mass unit, the kilogram, in terms of a defined value of Planck's constant, h, the second, and the velocity of light, which defines the meter. *See* HALL EFFECT; JOSEPHSON EFFECT.

Advantages of this approach. Equation (3) can be derived in a much more general manner to bring out two points vital to attaining the required accuracy:

1. Equation (3) contains no terms relating to actual expenditure of energy. That is, the relation is concerned with virtual work only, and no real work such as friction in the bearings of any balance used to judge equality of the electromagnetic and gravitational forces, or joule heating of the coil by the current, is relevant to its exactness.

2. The geomertical shape of the coil and the spatial distribution of the magnetic flux are also irrelevant. Circular and rectangular coils and linear and radial fluxes from permanent or superconducting magnets have been employed. All that matters is that coil and flux should combine to give a vertical force, a condition which can be tested in various experimental ways.

A further advantage of this approach to linking the electrical and base SI units is that the magnitude of the various quantities can easily be designed to be of a size suitable for measurement at the highest level of accuracy possible. To illustrate this, a kilogram weight can be opposed by a few milliamperes of current which can be measured in terms of the potential difference it produces across a 100-ohm resistor. A velocity of a few millimeters a second, which is measurable as a few thousand optical fringes a second by a laser interferometer, can produce a volt of emf which can be compared directly with the output of an array of Josephson junctions. Present-day technology can measure g to 0.001 ppm accuracy by interferometric distance measurement and timing of a freely falling object. Nevertheless, attaining an overall accuracy of the order of 0.01 ppm from watt balance measurements demands unprecedented skills and meticulous attention to detail in mechanical, optical, and electrical metrology. As well as creating the satisfaction of progressing towards replacing the last SI base unit based on an artifact, undertaking watt balance measurements advances capabilities in these areas of metrology. *See* PHYSICAL MEASUREMENT.

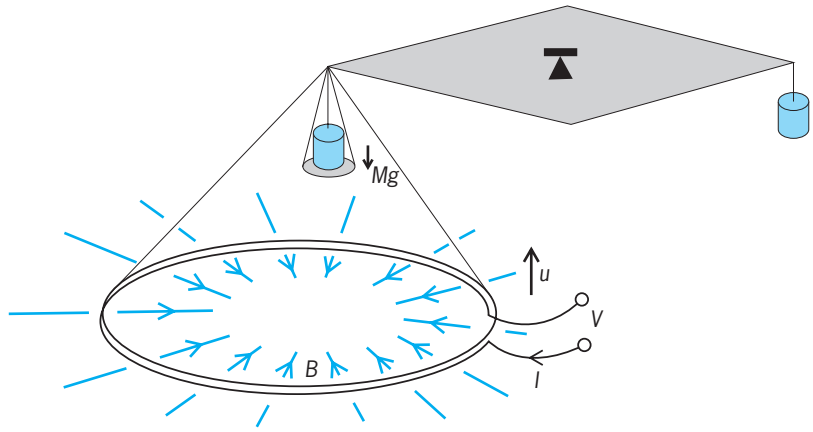

Diagram of the Watt balance, illustrating the working principle of the apparatus. Symbols are defined in text.

Bryan P. Kibble

Bibliography. A. Eichenberger et al., Tracing Planck's constant to the kilogram by electromechanical methods, *Metrologia*, 40:356–365, 2003; I. Robinson, Redefining the kilogram, *Phys. World*, 17(5):31–35, May 2004.

Watt-hour meter

An electrical energy meter, that is, an electricity meter that measures and registers the integral, with respect to time, of the power in the circuit in which it is connected. This instrument can be considered as having two parts: a transducer, which converts the power into a mechanical or electrical signal, and a counter, which integrates and displays the value of the total energy that has passed through the meter. Either or both of these parts can be based on mechanical or electronic principles.

In its wholly mechanical form the transducer is an electric motor designed so that its torque is proportional to the electric power in the circuit. The motor spindle carries a conducting disk that rotates between the poles of one or more strong permanent magnets. These provide a braking torque that is proportional to the disk rotational speed, so the motor runs at a rate that accurately represents the circuit power. The integrating register is simply connected to the motor through a gear train that gives the required movement of the dials in relation to the passage of electrical energy. *See* MOTOR.

Mechanical meters. These can measure either dc or ac energy.

DC energy measurement. Some form of commutator motor similar to those with a shunt field winding is commonly used for dc energy measurement. It is most convenient for the field to carry the circuit current, while the armature is fed with a signal from the circuit voltage. The motor is designed to have very

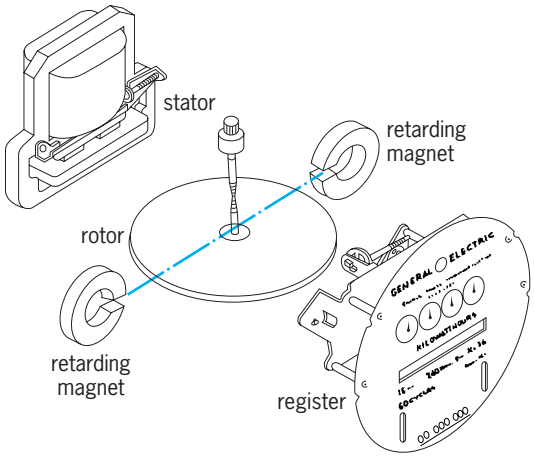

Fig. 1. Basic elements of an induction-type watt-hour meter. (*General Electric Co.*)

low friction and to start easily in order to maintain a linear torque-to-power relationship over as wide a range as possible. Some early designs were similar to the homopolar motor, in which the armature is a copper disk with the current entering through a pool of mercury. Owing to the wish to reduce mercury in the environment, because of its toxicity, these are no longer likely to be installed.

An entirely different approach to dc energy metering depends on the change in the rate of swing of a pendulum that carries a coil on its bob and swings over a second, fixed coil. The swinging coil carries a signal proportional to the voltage, while the fixed coil carries a signal depending on the load current. A duplicate system is arranged with one of the current directions reversed. The rate of one pendulum is increased by the power being measured, and the other decreased, as a result of the magnetic forces between the fixed and swinging coils. A pair of escapements is controlled by the pendulums, and the difference in their rates is obtained by using a mechanical differential. The rate of rotation of the differential output is proportional to the circuit power and is used to drive a mechanical register. The obvious error arising from mismatch of the pendulums is eliminated by regularly reversing the directions of the currents in the swinging coils and the direction of mechanical drive. These instruments are capable of 0.1% accuracy. *See* DIFFERENTIAL; PENDULUM.

AC energy measurement. The Ferraris, or induction-type, meter is used for ac energy measurement (**Fig. 1**). The stator (**Fig. 2**) carries two windings. The potential winding is fitted onto the single limb of the laminated iron structure. It is arranged to have a highly inductive impedance in order that the current flowing in it is in quadrature with the applied voltage. The current coils are wound on a pair of limbs situated on the other side of the structure. The rotor disk is mounted on a vertical axle with low-friction bearings and is arranged so that it passes through gaps in the stator and one or more permanent magnets. The two fluxes from the current and potential coils induce currents in the disk that interact with them to produce a torque proportional to the circuit power. The magnets produce a retarding torque that is proportional to the rotor speed, and is therefore proportional to the power. The disk drives a mechanical register through a worm drive on its axle. A typical rate of disk rotation is 200 revolutions per kilowatt-hour. An ordinary energy meter will easily achieve an accuracy of 2% over a wide range of loads; precision models may reach 0.1%.

Polyphase energy meters may be constructed with two or three stators driving the same disk, or individual disks mounted on the same axle. Care is required in these designs if interaction between the elements is to be avoided. Additional indicators can be fitted to indicate the maximum demand that has been made during, for example, any half-hour period since the instrument was reset. Other modifications enable the meter to record reactive energy, in kilovolt-ampere-reactive-hours, instead of real energy, in kilowatt-hours. Remote metering can be achieved by arranging for the rotation of the disk to provide electrical impulses through contact closures or optical rotation detectors. *See* VOLT-AMPERE.

Ferraris meters are in very wide use and measure the consumption of the vast majority of domestic and industrial users of electric power throughout the world. The design has been refined over many years, and its simplicity, combined with mass-production methods, has held manufacturing costs to a level at which it was hard for any other approach to compete. However, the demand on the one hand for much greater precision and on the other hand for greater complexity of measurement has led to the successful development of a variety of electronic watt-hour meters.

Electronic meters. The electronic watt transducer is a solid-state circuit that performs the multiplication of current and voltage signals, and delivers an output in the form of a pulse train at a rate proportional to power. The most widely used principle is the time-division multiplier, in which pulses are modulated in duration and amplitude by the voltage and current

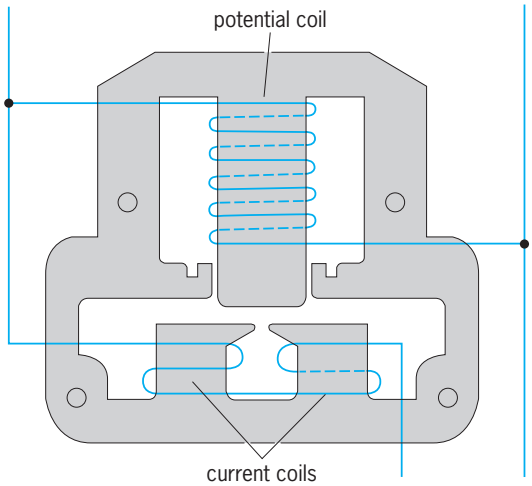

Fig. 2. Stator of induction-type watt-hour meter, showing the windings. (*General Electric Co.*)

signals, respectively. A suitably scaled version of the voltage signal is compared in a comparator with the output of a triangle or sawtooth waveform. The comparator output is a train of pulses whose durations are proportional to the quasi-instantaneous values of the voltage. In effect, this is a sampling system, and the pulse rate needs to be high compared with the signal frequency. Rates of 20–100 kHz are typical. The comparator output is then modulated to produce pulses having amplitudes proportional to the current signal. The resulting pulse train is then smoothed, giving a dc voltage signal that represents the power level. *See* COMPARATOR; PULSE MODULATION.

A further section converts this power signal into another train of pulses, but the pulse rate is now arranged to be proportional to the power. One method of achieving this is to apply the dc signal to an integrator. Each time the integrator output exceeds a threshold level, a single pulse is generated. This has its amplitude and duration very accurately defined and is applied to the integrator input with the polarity required to return the integrator toward zero. In the long term, pulses will be generated at the correct rate to be proportional to the input current regardless of the actual threshold level or integrator components.

The simplest solid-state watt-hour meter is completed by adding an electronic register to record the energy consumed. Precision electronic energy meters can give errors less than 0.005%. Electronic energy meters are not directly competitive with mechanical versions where a simple meter at the 2% level is required. However, for any application where more elaborate features are needed, the electronic model is superior and the absence of any visible moving parts reduces the opportunity for fraud. Electronic instruments are available in which six registers are provided, to record consumption at four different times of day and two levels of maximum demand. Automatic switching of heating loads can be built in, controlled by internal clocks, signals carried on the power lines, or radio signals. Prepayment and recovery of debts can be arranged through special meters that are used in conjunction with electronic keys, magnetic cards, or smart cards. Advanced designs are based on the independent measurement of voltage, current, and phase, with a microprocessor carrying out the calculation of any required parameter, including kilowatt-hours, kilovolt-ampere-reactive hours, and power factor. *See* MICROPROCESSOR.

Intelligent, or smart, meters can provide a wide variety of load and tariff-control functions, as well as remote reading of energy consumption. Different approaches are used around the world in order to meet local requirements and use available facilities. One-way or two-way communication may be operated, using radio, telephone, or the power line itself. In the United States, 900-MHz radio links have been used to enable a moving van to interrogate all the meters in a locality. In the United Kingdom, where this frequency band is used for other purposes, over 1.5 million meters are in use which select multiple tariffs and switch heating loads under the control of a radio teleswitch, operated by phase modulation of a broadcast transmission at 180 kHz. A different system uses dedicated transmissions in the 183.5–184.5-MHz band. In continental Europe, where many low-frequency systems have been in use for a long period, signals are superimposed on the power network over a low-frequency band at 9–95 kHz. Although the use of the network is very attractive, being under the control of the corporation directly concerned and saving payments to third parties, there are serious difficulties in its use owing to the way in which its transmission characteristics change with load. Broad-spectrum transmissions are employed, and reliable communication at 200 baud is achieved.

Hybrid and special meters. Electronic and mechanical techniques can be combined in a variety of ways. Signals from a mechanical transducer may be used to operate electronic registers in order to obtain the advantages of the facilities that they can provide. A mechanical impulse register may be used in conjunction with an electronic transducer, where it is considered important to maintain a record without the need for batteries or nonvolatile memory elements.

Watt-hour meters that operate at potentials of 100–250 V and currents up to 100 A are widely manufactured. At higher levels, voltage or current transformers are used to reduce the signals handled by the meter to more convenient values, frequently 110 V and 5 A. By this means it is possible to carry out energy metering at any level required, including hundreds of kilovolts and tens of kiloamperes. *See* ELECTRICAL ENERGY MEASUREMENT; ELECTRIC POWER MEASUREMENT; INSTRUMENT TRANSFORMER; TRANSFORMER. R. B. D. Knight

Bibliography. T. Brewer, *An Introduction to Electrical Measurements*, 2d ed., 2001; *Electric Meters Code for Electricity Metering*, Amer. Nat. Stand. C-12.1, 2001; R. Dettmer, Meter power, *IEE Rev.*,38(6):237–240, 1992; Edison Electric Institute, *Handbook for Electricity Metering*, 9th ed., 1999; P. S. Filipski, A TDM wattmeter with 0.5-MHz carrier frequency, *IEEE Trans. Instrum. Meas.*, IM-39:15–18, 1990; F. K. Harris, *Electrical Measurements*, 1952, reprint 1975; M. V. Reissland, *Electrical Measurements*, 1989; *Watthour Meters*, Amer. Nat. Stand. C-12.10, 1987.

Wattmeter

An instrument that measures electric power. For a complete discussion of power measurement in various types of electric circuits. *See* ELECTRIC POWER MEASUREMENT.

A variety of wattmeters are available to measure the power in ac circuits. They are generally classified by names descriptive of their operating principles. Determination of power in dc circuits is almost always done by separate measurements of voltage and current. However, some of the instruments described will also function in dc circuits, if desired.

Electrodynamic wattmeter. Probably the most useful instrument in the measurement of ac power at

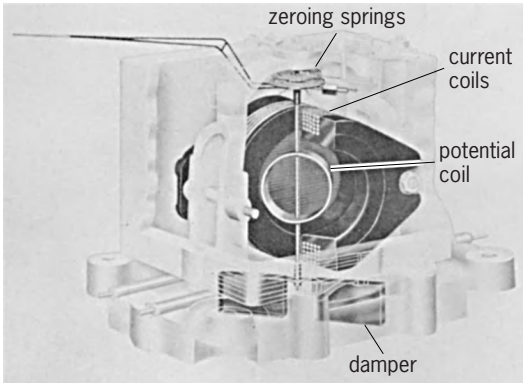

Fig. 1. Single-element electrodynamic wattmeter. (*Weston Instruments, Division of Daystrom, Inc.*)

commercial frequencies is the indicating (deflecting) electrodynamic wattmeter (**Fig. 1**). It is similar in principle to the double-coil dc ammeter or voltmeter in that it depends on the interaction of the fields of two sets of coils, one fixed and the other movable. The moving coil is suspended, or pivoted, so that it is free to rotate through a limited angle about an axis perpendicular to that of the fixed coils. As a single-phase wattmeter, the moving (potential) coil, usually constructed of fine wire, carries a current proportional to the voltage applied to the measured circuit, and the fixed (current) coils carry the load current. This arrangement of coils is due to the practical necessity of designing current coils of relatively heavy conductors to carry large values of current. The potential coil can be lighter because the operating current is limited to low values. *See* AMMETER.

If there is no iron or disturbing magnetic field due to current in neighboring conductors, then the instantaneous torque is proportional to the product of the instantaneous current in the potential coil and the instantaneous current in the current coils. Since the instantaneous current in the potential coil is proportional to the instantaneous voltage across the circuit, it follows that the instantaneous torque is proportional to the product of the instantaneous voltage and the instantaneous current in the current coils.

The moving system, however, is designed with sufficient inertia that it is unable to follow the rapid alternations of the alterating current, but it will rotate, opposed by light zeroing springs, to a position corresponding to the average torque, which is proportional to the average power being supplied.

To avoid a loss of power in the instrument due to the current flowing in the potential coil, a fine-wire coil is wound with the current coils and connected in series with the potential coil. Its effect is to cancel the magnetic effect of the potential coil current from the field of the current coils. If this compensation is correct, with the load circuit open, the instrument will read zero.

The presence of inductance in the potential circuit would normally introduce phase displacement between voltage and current, which is theoretically inadmissible. Simple tuning with capacity would introduce frequency error, but capacity in series with the moving coil and shunted by noninductive resistance is found to reduce the net reactance to a tolerably low value. *See* ALTERNATING-CURRENT CIRCUIT THEORY; CAPACITANCE; INDUCTANCE; REACTANCE.

Other disturbing effects are those of ambient temperature, eddy currents in metal structures close to the moving coil, transformer effect due to mutual induction between current and potential coils, and skin effect in which a change in distribution of current in the current coil causes frequency error. In precision instruments these effects are taken care of as much as possible by proper choice of materials and by shielding. *See* EDDY CURRENT; SKIN EFFECT (ELECTRICITY); TRANSFORMER.

This type of wattmeter is capable of great accuracy within its range and is commonly used as a laboratory standard. It has the great advantage of being usable as a transfer instrument; that is, it can be calibrated with direct current and used for ac measurements. It is used chiefly at power frequencies. For reasons of manufacturing convenience rather than on technical grounds, commercially available instruments are limited to full scale ranges of 300 V and 10 A. External resistive multipliers may be used to extend voltage range. Shunting of current coils is not recommended. The preferred method of extending the range of both current and voltage is through current and potential transformers, by which the voltages and current in the load circuit are reduced to nominal values by definite transformation ratios without introducing appreciable phase errors, or, if such errors are introduced, their values are known and may be accounted for. *See* INSTRUMENT TRANSFORMER.

Quarter-squares multiplying wattmeter. The mathematical relationship of Eq. (1) can be used to obtain

$$4XY = (X + Y)^2 - (X - Y)^2 \qquad (1)$$

the effect of an analog multiplier. If signals representing the voltage (X) and current (Y) are combined to form new signals equal to the sum and difference, and these signals are then squared and the difference of the squares is evaluated, the resulting output is proportional to the required product. Equation (1) shows that a scale factor of 4 is introduced, which gives the method its name. Several different wattmeters are based on this principle.

Thermal wattmeter. A thermal converter consists of a resistive heater in close thermal contact with one or more thermocouples. When current flows through the heater, the temperature rises. Thermocouples give an output voltage proportional to the temperature difference between their junctions, in this case proportional to the square of the current, and so make suitable transducers for the construction of thermal wattmeters. *See* ELECTRICAL RESISTANCE.

In one possible arrangement (**Fig. 2**), four devices are used in a bridge configuration. The original voltage and current signals are first converted into currents within the working range of the thermal converters. The sum of these two currents flows through two of the devices, while the difference flows through the other two devices. The

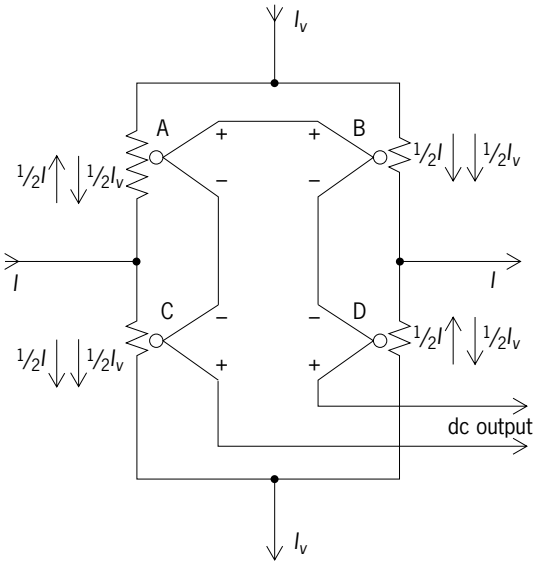

Fig. 2. Thermal wattmeter circuit. I_v is the voltage signal; I is the current signal; B and C are thermal converters with current $\frac{1}{2}(I + I_v)$; A and D are thermal converters with current $\frac{1}{2}(I - I_v)$.

required difference between the dc outputs is obtained by suitable connection of the four thermocouples. Close matching of the thermal converters is necessary in order to ensure that the subtraction accurately eliminates the signals corresponding to the squares (X^2 and Y^2) of the voltage and current inputs. Instruments of this type can have a wide bandwidth, owing to the simplicity of the thermal converter, but accuracy is limited to a few percent.

Considerable improvements in accuracy result from the use of a single multijunction thermal converter switched sequentially between the sum and a second signal obtained by combining the difference with a third, dc signal. This dc signal is derived by a feedback arrangement designed to cause the two inputs to become equal. Rearranging Eq. (1) shows that this dc signal gives the required power measurement. Wattmeters of this design can have errors as low as 10 parts per million at power frequencies. *See* THERMAL CONVERTERS; THERMOCOUPLE; THERMOELECTRICITY.

Electrostatic wattmeter. The electrostatic force between two conductors is proportional to the product of the square of the potential difference between them and the rate of change of capacitance with displacement. A differential electrostatic instrument may therefore be used to construct a quarter-squares wattmeter. In spite of the problems of matching the capacitance changes of the two elements and the small forces available, electrostatic wattmeters were used as standards for many years. Uncertainties of the order of 50 parts per million could be obtained with a useful frequency response up to 100 kHz, but the working levels of current and voltage were very limited.

Digital wattmeter. Digital wattmeters combine the advantages of electronic signal processing and a high-resolution, easily read display. Electrical readout of the measurement is also possible. A variety of electronic techniques for carrying out the necessary multiplication of the signals representing the current and voltage have been used. Usually the electronic multiplier is an analog system which gives as its output a voltage proportional to the power indication required. This voltage is then converted into digital form in one of the standard ways. Many of the multipliers were originally developed for use in analog computers. *See* ANALOG COMPUTER; ELECTRONIC DISPLAY.

Hall-effect multiplier. The Hall-effect multiplier makes use of the relationship between the transverse (Hall) voltage which appears across a suitable semiconductor device, the current flowing through it, and an applied magnetic field. The voltage is proportional to the product of the device current and the magnetic field. Multiplication can be performed in all four quadrants (that is, when the current and field are, respectively, both positive, positive and negative, negative and positive, or both negative). This system can give wide bandwidth, but accuracy is limited by the small output from the device and its poor drift characteristics. *See* HALL EFFECT.

Pulse height/width multiplier. The pulse height/width multiplier is in widespread use. In one version the width of pulses in a train having constant frequency is arranged to be controlled by one signal, generally the voltage, while their height is controlled by the other. The resulting signal is smoothed by a filter, which gives an output proportional to the required product. Some ingenuity is required to extend the method to work in four quadrants, and the bandwidth is limited by the need for the pulse frequency to be higher than the highest-frequency component of the power to be measured. At 60 Hz, errors as low as 10 parts per million can be realized. *See* ELECTRIC FILTER; PULSE MODULATION.

Electronic quarter-squares wattmeter. Electronic versions of the quarter-squares multiplier have been developed. The combination of the voltage and current signals may be carried out by using operational amplifiers. As described above, thermal instruments can achieve very high accuracy. Instruments have also been made which use electronic squaring circuits, similar to those developed for root-mean-square sensing. In such instruments, the conventional expression for power is replaced by the difference between the average values of the squares in Eq. (1) over long time periods, the evaluation of these two squares being carried out separately. The relative phase of the sum and difference signals is not significant, unlike the relative phase between the voltage and current signals, which is critical to the measurement. The technique is attractive for high-frequency applications, where minimizing phase errors is a major design problem. *See* OPERATIONAL AMPLIFIER.

A variant of the electronic quarter-squares multiplier (**Fig. 3**) makes use of the relationship in Eq. (2).

$$2XY = (X + Y)^2 - X^2 - Y^2 \qquad (2)$$

This simplifies the phase problems further, since only

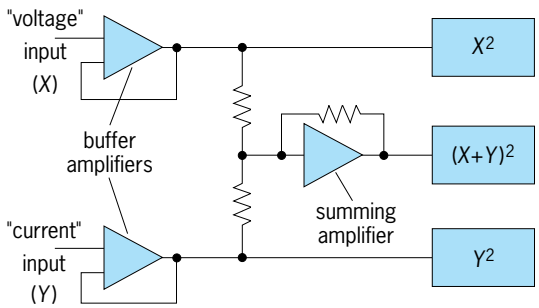

Fig. 3. Wideband electronic wattmeter.

the sum needs to be produced. Since the measurements of the three terms are independent, X^2, Y^2, and $(X + Y)^2$ can be evaluated by using conventional high-accuracy digital voltmeters, with the square being obtained arithmetically from the root-mean-square value indicated. This technique has been applied to an audio-frequency power standard which has errors of less than 50 parts per million at 3 kHz and 200 parts per million at 30 kHz.

Digital sampling wattmeter. A different approach uses sampling techniques (**Fig. 4**). Sets of coincident samples of the voltage and current signals are obtained, using sample-and-hold units. The resulting instantaneous values are converted into digital form by fast analog-to-digital converters. The pairs of measurements are multiplied by using a computer, which accumulates the products to provide a value for the power. A steady result can be obtained by accumulating samples over a time (and number of samples) which is long compared with the period of the waveform. Another strategy is to carefully control the timing of the samples so that N samples are taken in exactly M cycles, where N and M have no common factors. This has been shown to be a very powerful method, as the response of instruments using the method can extend to frequencies limited by the sample-and-hold units rather than the sampling rate. Other arrangements can include the use of random sample timing or interpolation between samples. The accuracy is limited by the performances of available sample-and-hold units and analog-to-digital converters, which are approaching the performances of thermal devices. *See* ANALOG-TO-DIGITAL CONVERTER.

Polyphase wattmeter. The instruments thus far considered are designed for single-phase power measurement. In polyphase circuits, the total power is the algebraic sum of the power in each phase. This summation is assisted by simple modifications of single-phase instruments. *See* ALTERNATING CURRENT.

For example, an electrodynamic wattmeter may contain a second coil system similar to the coils of a single-phase meter, with the second potential coil on the same shaft as the potential coil in the first system. The two systems are mounted in the same case and are designed to have matched characteristics, but care is taken that there is no magnetic interaction between them. The deflection is proportional to the sum of the torques of the two elements; thus, total power is read from the instrument scale. Electrical connections are the same as for two single-phase wattmeters. Electronic and other wattmeters in which the measurement of power appears as a voltage or digital signal may be extended to polyphase applications by suitably combining the outputs from a number of single-phase transducers.

Loading effects. When a wattmeter is used in real measurement situations, there is a problem in connecting the instrument without introducing a source of error, as the power consumed in either the voltage- or current-measuring circuit will be included in the measured power. If the voltage circuit is connected directly across the load, the current circuit will sense the current consumed by it. Conversely, if the voltage circuit is connected on the supply side of the current sensor, it will include the voltage drop in the current circuit in the measured voltage. As most wattmeters are used at a voltage close to a nominal value, the errors caused by losses in the voltage circuit are almost constant. It is therefore possible to design instruments which incorporate compensation for these losses, and the most commonly used connection applies the voltage sensing circuit directly across the load.

Range extension. The full-scale ranges of voltage and current of all types of instrument are extended by the use of electronic amplifiers, voltage attenuators, voltage transformers, or current transformers. Such scaling devices commonly bring the signals to the nominal level of 110 V, 5 A. In this way, power at extremely high voltages and currents can be measured, with all connections to the wattmeter being in light, modestly insulated cables. *See* AMPLIFIER.

Calibration. When wattmeters are being calibrated, it is not necessary to operate the test and standard instruments in a situation where real power exists. If the voltage circuits are connected in parallel and the current inputs are connected in series, both instruments are deflected by the same product of voltage and current, although only sufficient power is consumed as is required to operate the measurement circuits. This technique is known as phantom power calibration. When the wattmeter is operated in this way, loading effects do not appear. In order to ensure accurate calibration, instruments providing compensation of loading effect usually make

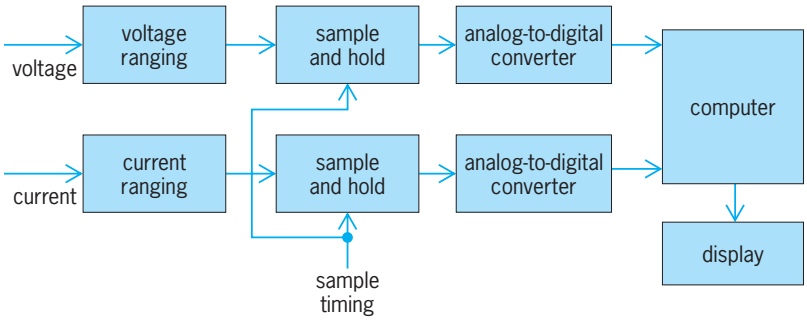

Fig. 4. Block diagram of digital sampling wattmeter.

provision for the compensation to be switched off in these circumstances. R. B. D. Knight

Bibliography. F. K. Harris, *Electrical Measurements*, 1952, reprint 1975; A. D. Helfrick and W. D. Cooper, *Modern Electronic Instrumentation and Measurement Techniques*, 1990; M. V. Reissland, *Electrical Measurements: Fundamentals, Concepts, Applications*, 1989; M. G. Say (ed.), *Electrical Engineer's Reference Book*, 15th ed., 1993; L. Schnell (ed.), *Technology of Electrical Measurements*, 1993; L. M. Thompson, *Electrical Measurements and Calibration*, 2d ed., 1994.

Wave (physics)

The general term applied to the description of a disturbance which propagates from one point in a medium to other points without giving the medium as a whole any permanent displacement.

Waves are generally described in terms of their amplitude, and how the amplitude varies with both space and time. The actual description of the wave amplitude involves a solution of the wave equation and the particular boundary conditions for the case being studied. In the cases most often considered, the wave equation is simplified to a second-order, linear, partial differential equation. This equation for a one-dimensional space coordinate x is written as Eq. (1), where ϕ is the amplitude, c is the wave velocity, and t is the time coordinate. The generalized solutions of this equation are of the form given by Eq. (2). The first term indicates a wave traveling in the positive x direction at a velocity c, and the second term a wave traveling in the negative x direction. The functions F and G are determined by the particular properties of the boundary conditions of the problem. In the one-dimensional case these are usually sine or cosine waves. The velocity of propagation c is proportional to the square root of the ratio of the elastic to the inertial constants of the medium. *See* SINE WAVE; WAVE EQUATION; WAVE MOTION.

$$\frac{\partial^2\phi}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} \tag{1}$$

$$\phi = F(x - ct) + G(x + ct) \tag{2}$$

Acoustic waves, or sound waves, are a particular kind of the general class of elastic waves. Elastic waves are propagated in media having two properties, inertia and elasticity. Elasticity of the medium is required in order to provide a force which tends to restore a displaced particle of the medium to its original position. Inertia is required to enable the displaced particle to transfer momentum to an adjoining particle. A shear wave, or rotational wave, is a wave in an elastic medium which causes an element of the medium to change its shape without a change of volume.

Electromagnetic waves (for example, light waves and radio waves) are not elastic waves and therefore can travel through a vacuum. The velocity of the wave depends on the medium through which the wave travels, but in a vacuum it is a constant, c, approximately equal to 3×10^8 m/s (9.8×10^8 ft/s). *See* ELECTROMAGNETIC WAVE; SHOCK WAVE. William J. Galloway

Wave equation

The name given to certain partial differential equations in classical and quantum physics which relate the spatial and time dependence of physical functions. In this article the classical and quantum wave equations are discussed separately, with the classical equations first for historical reasons.

Classical wave equation. In classical physics the name wave equation is given to the linear, homogeneous partial differential equations which have the form of Eq. (1). Here v is a parameter with the dimensions of velocity; $\mathbf{r}$ represents the space coordinates x, y, z; t is the time; and ∇^2 is Laplace's operator defined by Eq. (2). The function $f(\mathbf{r}, t)$ is a physical observable; that is, it can be measured and consequently must be a real function. Laplace's operator may also be written in polar coordinates $\mathbf{r}$, θ, and ϕ; if f depends only on the radial distance r, the wave equation has the simple form shown by Eq. (3).

$$\left[\nabla^2 - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\right] f(\mathbf{r}, t) = 0 \tag{1}$$

$$\nabla^2 \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{2}$$

$$\left[\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} - \frac{1}{v^2}\frac{\partial^2}{\partial t^2}\right] f(r, t) = 0 \tag{3}$$

Motion of a string. The simplest example of a wave equation in classical physics is that governing the transverse motion of a string under tension and constrained to move in a plane. In this case the wave equation is given by Eq. (4), where $y(x, t)$ is the transverse displacement of the string from its equilibrium position at time t. If the string is thought of as being made up of many separate particles, x is the label for the particle. The velocity parameter is $v = \sqrt{T/\lambda}$, with λ the linear mass density and T the tension in the string, both uniform. Equation (4) is valid when the slope of the string is small, that is, $\partial y/\partial x \ll 1$. A solution to Eq. (4) is any function of the form $f(x - vt)$, where f is arbitrary except that it must be twice differentiable. Since at $t = 0$ one would have $y(x, 0) = f(x)$, $f(x)$ specifies the initial shape of the string. Since $f(x - vt)$ differs from $f(x)$ only by a translation along the x axis of magnitude vt, it is apparent that the shape of the string moves along the x axis as a wave with shape determined by $f(x)$ and with velocity v, which is called the wave propagation velocity. A point to note is that v is not the transverse velocity $\partial y/\partial t$ of the particles in the string. In fact,

$$\frac{\partial^2 y(x,t)}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 y(x,t)}{\partial t^2} \tag{4}$$

since Eq. (5) holds, the transverse velocity for the

$$\frac{\partial y}{\partial t} = -v\frac{\partial y}{\partial x} \tag{5}$$

solution $f(x - vt)$ is always small and opposite in sign to the slope. Another solution to Eq. (4) could be $g(x + ct)$. which corresponds to the shape of the string moving in the negative x direction, that is, with velocity $-v$. The most general solution to Eq. (4) is the linear combination given by Eq. (6).

$$y(x, t) = a^{+}f(x - vt) + a^{-}g(x + vt) \tag{6}$$

The particular choice of constants $a\pm$ is determined by the shape of the string at $t = 0$ and by the transverse velocity at $t = 0$, that is, by $y(x, 0)$ and $\partial y(x, t)/\partial t|_{t=0}$.

Acoustical waves. A second type of classical physical situation in which the wave equation, Eq. (1), supplies a mathematical description of the physical reality is the propagation of pressure waves in a fluid medium. Such waves are called acoustical waves, the propagation of sound being an example. Suppose that x is the undisturbed position of a particle in the fluid and let the displacement at time t be $z(x, t)$ so that those particles that have position x when undisturbed have position $x + z(x, t)$ at time t after being disturbed. If $\partial z/\partial x \ll 1$, the acoustic approximation, then the displacement $z(x, t)$ satisfies Eq. (7), where

$$\frac{\partial^2 z(x, t)}{\partial x^2} = \frac{1}{\overline{v}^2}\frac{\partial^2 z(x, t)}{\partial t^2} \tag{7}$$

$\overline{v}^2 = dP(\rho)/d\rho$ at $\rho = \rho_0$, with ρ the density of the fluid, ρ_0 the undisturbed density, and $P(\rho)$ the pressure. The acoustic approximation must be used to derive Eq. (7) from the density equation, the equation of state for the fluid, and Newton's second law of motion. If the fluid is specified to be air, then $\overline{v}$ is the velocity of sound. Of course, for acoustical waves and vibrating strings there may be three-dimensional motion in the general case, satisfying Eq. (1) with the appropriate parameters. *See* SOUND; WAVE MOTION IN FLUIDS.

Electromagnetic waves. A third example of a classical physical situation in which Eq. (1) gives a description of the phenomena is afforded by electromagnetic waves. In a region of space in which the charge and current densities are zero, Maxwell's equations for the photon lead to the wave equations (8). Here **E** is

$$\begin{aligned}\left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]\mathbf{E}(\mathbf{r}, t) &= 0\\ \left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]\mathbf{B}(\mathbf{r}, t) &= 0\end{aligned} \tag{8}$$

the electric field strength and **B** is the magnetic flux density; they are both vectors in ordinary space. The parameter c is the speed of light in vacuum. A vector potential $\mathbf{A}(\mathbf{r}, t)$ and a scalar potential $\phi(\mathbf{r}, t)$ may be introduced by the identifications given in Eqs. (9),

$$\begin{aligned}\mathbf{B}(\mathbf{r}, t) &= \nabla \times \mathbf{A}(\mathbf{r}, t)\\ \mathbf{E}(\mathbf{r}, t) &= -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}\end{aligned} \tag{9}$$

where $\nabla \times$ is the curl operator and ∇ the gradient operator. The potentials then also satisfy the wave equation. *See* POTENTIALS.

If relativistic notation is used so that $A_x = A_1$, $A_y = A_2$, $A_z = A_3$, and $A_4 = i\phi$ with $i = \sqrt{-1}$, then Eq. (10) can be derived (all Greek indices run from

$$\left[\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right]A_\mu(\mathbf{r}, t) = 0 \tag{10}$$

1 to 4). This wave equation may be written in a form that is manifestly covariant (form invariant) with respect to the Lorentz transformations by the further identifications $x = x_1$, $y = x_2$, $z = x_3$, and $x_4 = ict$ so that the operator in Eq. (9) can be written as Eq. (11),

$$\begin{aligned}&\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\\ &\quad = \frac{\partial^2}{\partial {x_1}^2} + \frac{\partial^2}{\partial {x_2}^2} + \frac{\partial^2}{\partial {x_3}^2} + \frac{\partial^2}{\partial {x_4}^2} \equiv \frac{\partial}{\partial x_\mu}\frac{\partial}{\partial x_\mu}\end{aligned} \tag{11}$$

where the repeated index in the second equation is defined to mean that a sum is to be taken, as indicated above (Einstein summation convention). Then the manifestly covariant form of Eq. (10) is written as Eq. (12). Both $x = (\mathbf{r}, ict)$ and $A = (\mathbf{A}, i\phi)$ are Lorentz

$$\frac{\partial}{\partial x_\mu}\frac{\partial}{\partial x_\mu}A_\nu(\mathbf{r}, t) = 0 \tag{12}$$

four-vectors, and the product of two four-vectors, summed over the repeated index, is a scalar, such as

$$\frac{\partial}{\partial x_\mu}\frac{\partial}{\partial x_\mu}$$

See ELECTROMAGNETIC RADIATION; MAXWELL'S EQUATIONS; RELATIVITY; WAVE (PHYSICS); WAVE MOTION.

Quantum wave equation. The above discussion leads to the consideration of the quantum-mechanical wave equations. The nonrelativistic Schrödinger equation is an example. In this article only the relativistic quantum-mechanical wave equations are considered. *See* NONRELATIVISTIC QUANTUM THEORY.

Schrödinger-Klein-Gordon equation. The first example may be found from Eq. (12) by subtracting from the operator $(\partial/\partial x_\mu)/(\partial/\partial x_\mu)$ a constant term $(m^2c^2)/\hbar^2$, where m is the mass and $\hbar$ is Planck's constant divided by 2π. The new equation is Eq. (13) and is

$$\left[\frac{\partial}{\partial x_\mu}\frac{\partial}{\partial x_\mu} - \frac{m^2c^2}{\hbar^2}\right]\varphi(\mathbf{r}, t) = 0 \tag{13}$$

called the Schrödinger-Klein-Gordon equation. Every free elementary particle has a wave function $\varphi(\mathbf{r}, t)$ which must satisfy this equation although $\varphi(\mathbf{r}, t)$ are different types of functions for particles with different spins. If $\varphi(\mathbf{r}, t)$ has the form of a plane wave given by Eq. (14) corresponding to a particle with momen-

$$\varphi(\mathbf{r}, t) = \chi(\mathbf{P})\exp\left[\frac{i}{\hbar(\mathbf{P}\cdot\mathbf{x} - Et)}\right] \tag{14}$$

tum **P** and total energy E, then insertion of Eq. (14) into Eq. (13) yields Eq. (15). $\chi(\mathbf{P})$ in Eqs. (14) and

(15) can be any function of the momentum **P**. Since

$$\frac{1}{\hbar^2}\left[E^2 - c^2p^2 - m^2c^4\right]\chi(\mathbf{P}) = 0 \qquad (15)$$

the term in brackets in Eq. (15) is the relativistic connection between energy and momentum for a free particle, it is identically zero; therefore $\varphi(\mathbf{r}, t)$ defined by Eq. (14) satisfies Eq. (13). Historically, Eq. (15) came before Eq. (13).

Dirac equation. All the above wave equations are second-order in both space and time derivatives. A different type of relativistic quantum-mechanical wave equation is the Dirac equation, which is first-order in both space and time derivatives and has the form shown in Eq. (16). Here α is a three-component

$$\left[\frac{c\hbar}{i}\alpha \cdot \nabla + mc^2\beta\right]\psi(\mathbf{r}, t) = i\hbar\frac{\partial\psi(\mathbf{r}, t)}{\partial t} \qquad (16)$$

object whose components are each a 4×4 matrix, β is a 4×4 matrix, and the 4×1 column matrix $\psi(\mathbf{r}, t)$ is a wave function corresponding to a particle of intrinsic spin $^1/_2$.

If ψ_1, ψ_2, ψ_3, ψ_4 are the four components of $\psi(\mathbf{r}, t)$, then Eq. (16) is the matrix form of the system of Eq. (17), where $\alpha_i^{\sigma\tau}$ is the $\sigma\tau$ element of

$$\frac{c\hbar}{i}\sum_{\tau=1}^{4}\left(\alpha_1^{\sigma\tau}\frac{\partial}{\partial x_1} + \alpha_2^{\sigma\tau}\frac{\partial}{\partial x_2} + \alpha_3^{\sigma\tau}\frac{\partial}{\partial x_3}\right)\psi_\tau + \sum_{\tau=1}^{4}\beta_{\sigma\tau}mc^2\Psi_\tau = i\hbar\frac{\partial\psi_\sigma}{\partial t} \qquad (17)$$

the 4×4 matrix which forms the ith component of α in Eq. (16). Because it is necessary that each component of the solution $\psi(\mathbf{r}, t)$ of Eq. (16) also satisfy Eq. (13), α and β have certain anticommutation relations with each other which imply that the product of any of these matrices with itself is the 4×4 unit matrix. Equation (16) may be written in a manifestly covariant form by making the definitions $\gamma \equiv -i\beta\alpha$, $\gamma_4 \equiv \beta$ and by rearranging terms in Eq. (16). The result is (in index notation) Eq. (18).

$$\left[\gamma_\mu\frac{\partial}{\partial x_\mu} + \frac{mc}{\hbar}\right]\psi(\mathbf{r}, t) = 0 \qquad (18)$$

The term $\beta\alpha$ is interpreted as a three-component object with matrices $\beta\alpha_1$, $\beta\alpha_2$, and $\beta\alpha_3$ as components. The quantum-mechanical functions such as $\psi(\mathbf{r}, t)$ and $\varphi(\mathbf{r}, t)$ differ from their classical counterparts $y(\mathbf{r}, t)$ and so forth, in that the quantum functions are not directly observable quantities and are, in fact, complex. In general, only bilinear combinations of the quantum-mechanical wave functions, suitably averaged with certain operators, and over some region of space, are physically observable.

The Dirac equation is the most prominent example of the type of wave equation (relativistic quantum mechanical) that is linear in the space and time derivatives.

Rarita-Schwinger equation. A second example is the Rarita-Schwinger equation (19). This has the same

$$\left[\gamma_\mu\frac{\partial}{\partial x_\mu} + \frac{mc}{\hbar}\right]\psi_\nu(\mathbf{r}, t) = 0 \qquad (19)$$

form as the Dirac equation except that the wave function $\psi_\nu(\mathbf{r}, t)$ is now a 4×1 column matrix with each of the four components being independently a four-vector so that $\psi_\nu(\mathbf{r}, t)$ has 16 components. Equation (19) is intended to describe a free relativistic particle (and antiparticle) with intrinsic spin $^3/_2$. For this, only eight independent complex components are needed; in addition to Eq. (19), the auxiliary condition defined by Eq. (20) is required.

$$\gamma_\nu\psi_\nu(\mathbf{r}, t) = 0 \qquad (20)$$

Equations (19) and (20) together imply a third equation, Eq. (21). Exactly as in the Dirac equation,

$$\frac{\partial}{\partial x_\nu}\psi_\nu(\mathbf{r}, t) = 0 \qquad (21)$$

each component of the Rarita-Schwinger wave function satisfies Eq. (13). The necessity of auxiliary conditions, such as Eqs. (20) and (21), to eliminate redundant components in the wave function is a characteristic of the approach to relativistic wave equations which involve only first derivatives of space and time and certain matrices. These considerations for spins $^1/_2$ and $^3/_2$ may be extended to higher half-integral spins by including more four-vector indices on the wave function.

Duffin-Kemmer-Petiau equation. Another way of writing such a first-order matrix-differential wave equation is to make all the components of the wave function matrix components. The wave function $\overline{\psi}(\mathbf{r}, t)$ is then a matrix satisfying a wave equation of the form of Eq. (22), where the four β_μ are a set of singular ma-

$$\left[\beta_\mu\frac{\partial}{\partial x_\mu} + \frac{mc}{\hbar}\right]\overline{\psi}(\mathbf{r}, t) = 0 \qquad (22)$$

trices whose structure depends on the intrinsic spin of the particle to be described. For spin zero the β_μ are 5×5 matrices and for spin 1 they are 10×10. Of course, each component of $\overline{\psi}(\mathbf{r}, t)$ must also satisfy Eq. (13). Equation (22) is called the Duffin-Kemmer-Petiau equation.

Tensor equations. Free, relativistic quantum-mechanical particles with tensor equations may also be described. The best-known examples of this type are the Proca equations for spin 1 which have the form shown in Eqs. (23), and so are differential relations

$$F_{\alpha\beta} = \frac{\partial\omega_\beta}{\partial x_\alpha} - \frac{\partial\omega_\alpha}{\partial x_\beta} \qquad (23a)$$

$$\frac{\partial F_{\alpha\beta}}{\partial x_a} = \frac{m^2c^2}{\hbar^2}\omega_\beta \qquad (23b)$$

between a four-vector field $\omega_\beta(\mathbf{r}, t)$ and a second-rank tensor field $F_{\alpha\beta}(\mathbf{r}, t)$. By differentiating Eq. (23*b*) with respect to x_β, the auxiliary condition shown in Eq. (24) can be obtained; by inserting Eq. (23*a*) into

$$\frac{\partial}{\partial x_\beta}\omega_\beta(\mathbf{r}, t) = 0 \qquad (24)$$

Eq. (23*b*) and using Eq. (24), Eq. (13) can be satisfied by each of the components of ω_β. Thus it is again apparent that auxiliary conditions are needed to

eliminate redundant components from the quantum-mechanical equations of motion.

Spinor equations. The most general set of wave equations and auxiliary conditions capable of describing quantum-mechanically a free relativistic particle with mass and internal integral or half-integral spin are the Dirac-Fierz-Pauli spinor wave equations.

Equations with minimum number of components. A characteristic of all the above quantum-mechanical formulations is that the function satisfying the wave equation has more components than necessary to describe a free particle and antiparticle with the given spin. A description receiving much attention is that involving only irreducible representations of the homogeneous Lorentz group.

With this description a system with spin s is described by a wave function with $2s + 1$ components for the particle and $2s + 1$ components for the antiparticle, the minimum number for describing such a system. For the wave equation in this description, there may be only first-order time derivatives and complicated spatial derivatives, or there may be uniform dependence of the wave equation on both space and time derivatives, the order depending on the spin.

In many cases involving the $2(2s + 1)$-component description of a particle-antiparticle with intrinsic spin s, it is possible to decouple the two sets of $2s + 1$ components in the wave function by setting the mass m to zero. The wave equation then applies to a massless particle or antiparticle. For example, Eq. (16) with $m = 0$ becomes Eq. (25), in which

$$\frac{c\hbar}{i}\alpha \cdot \nabla\psi(\mathbf{r}, t) = i\hbar\frac{\partial\psi}{\partial_t}(\mathbf{r}, t) \tag{25}$$

the wave function $\psi(\mathbf{r}, t)$ now also refers to $m = 0$. With the particular choice of representation of Dirac matrices of Eq. (26), where $\boldsymbol{\sigma}$ are the Pauli spin ma-

$$\alpha = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 - & \boldsymbol{\sigma} \end{pmatrix} \tag{26}$$

trices, the upper two components of the wave function are decoupled from the lower two components. The wave equation for the latter is Eq. (27), where

$$-\frac{c\hbar}{i}\sigma \cdot \nabla\chi(\mathbf{r}, t) = i\hbar\frac{\partial}{\partial_t}\chi(\mathbf{r}, t) \tag{27}$$

$\chi(\mathbf{r}, t)$ is the lower two components of $\psi(\mathbf{r}, t)$. Equation (27) is known as the two-component neutrino equation, and the related equation for the upper two components represents a wave equation for the antineutrino, the antiparticle of a neutrino.

Wave equations like Eq. (27), with the Pauli matrices replaced by $\mathbf{s}/s$ and $\chi(\mathbf{r}, t)$ having $2s + 1$ components, are known to be appropriate for spin s massless particles when auxiliary conditions are included to restrict the wave function to two independent components as obtained in the two-component neutrino theory. This formulaton also includes the photon as a spin-1 example of a massless particle and offers an alternative description to Maxwell's equations discussed above. Such a description shows great promise for future developments in the field of relativistic quantum-mechanical wave equations. *See* QUANTUM MECHANICS; RELATIVISTIC QUANTUM THEORY.

David L. Weaver

Bibliography. J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics*, 1964; E. M. Corson, *Introduction to Tensors, Spinors and Relativistic Wave Equations*, 2d ed., 1981; C. A. Coulson and A. Jeffrey, *Waves: A Mathematical Approach to the Common Types of Wave Motion*, 2d ed., 1978; C. Itzykson and J. B. Zuber, *Quantum Field Theory*, 1980; G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers*, 2d ed., 1968; P. M. Morse and H. Feshbach, *Methods of Theoretical Physics*, 2 vols., 1953; M. D. Scadron, *Advanced Quantum Theory and Its Application through Feynman Diagrams*, 2d ed., 1991.

Wave mechanics

The modern theory of matter holding that elementary particles (such as electrons, protons, and neutrons) have wavelike properties. By 1915, experiments on the diffraction (bending) of x-rays into special directions by crystals had established that x-rays were electromagnetic waves, akin to visible and infrared light, but of much shorter wavelength than those other electromagnetic radiations. However, in 1923 A. H. Compton showed that observations on x-rays scattered by, for example, a graphite target could be quantitatively predicted via the hypothesis that the x-ray scattering from each individual target atom resulted from elastic (billiard-ball-like) collisions between the comparatively slowly moving atomic electrons and what were in effect particles (now commonly called photons) in the incident x-radiation, with each incident photon having an energy equal to the product of Planck's constant (6.6×10^{-34} joule second) and the speed of light divided by the wavelength, and momentum equal to Planck's constant divided by the wavelength. Thus, some experiments with electromagnetic radiation seemingly can be understood only by visualizing the radiation as waves, while other experiments on the same radiation seemingly require that the radiation be visualized as a stream of particles. This quite unintuitive wave-particle duality manifested by electromagnetic radiation is all the more remarkable in that the particlelike properties associated with the radiation, namely the energy and momentum of its photons, are given in terms of its wavelength, a concept that seemingly has no meaning except in a wave context. *See* COMPTON EFFECT; PHOTON; X-RAY DIFFRACTION.

In 1924 L. de Broglie postulated that the wave-particle duality which had been demonstrated for electromagnetic radiation also was a property of the elementary particles, such as electrons and protons, making up the atoms and molecules forming ordinary matter. In particular, de Broglie postulated that a particle has an associated wavelength obeying the same relation as was found to hold for photons, namely: the wavelength equals Planck's

constant divided by the particle's momentum (as customarily defined in elementary mechanics). This hypothesis was verified in 1927 in an experiment in which a beam of electrons having known momentum is diffracted by a crystal into special directions. As in the x-ray diffraction experiments discussed above, such diffraction seems understandable only on the hypothesis that the electrons are waves. Furthermore, the wavelength of the electrons in the incident beam, computed via the very same formula as was used to derive the x-ray wavelengths in the x-ray diffraction experiments, agreed precisely with the de Broglie relation. *See* ELECTRON DIFFRACTION.

Subsequent experiments have confirmed that not merely electrons but material particles in general, such as neutrons and neutral sodium atoms, manifest the wave-particle duality and obey the de Broglie relation. The de Broglie relation and the qualitative wave-particle duality concept have been incorporated into the highly successful modern theory of quantum mechanics. *See* DE BROGLIE WAVELENGTH; INTERFERENCE OF WAVES; QUANTUM MECHANICS. Edward Gerjuoy

Bibliography. E. Fermi and L. Marshall, Interference phenomena of slow neutrons, *Phys. Rev.*, 71:666–667, 1947; D. Halliday, R. Resnick, and K. S. Krane, *Physics*, vol. 2, 4th ed., 1992; D. W. Keith et al., Diffraction of atoms by a transmission grating, *Phys. Rev. Lett.*, 61:1580–1583, 1988; F. K. Richtmyer, E. H. Kennard, and J. N. Cooper, *Introduction to Modern Physics*, 6th ed., 1969.

Wave motion

The process by which a disturbance at one point in space is propagated to another point more remote from the source with no net transport of the material of the medium itself. For example, sound is a form of wave motion; wind is not. Wave motion can occur only in a medium in which energy can be stored in both kinetic and potential form. In a mechanical medium, kinetic energy results from inertia and is stored in the velocity of the molecules, while potential energy results from elasticity and is stored in the displacement of the molecules.

In a free traveling wave (as distinguished from a stationary or standing wave) one part of the medium disturbs an adjacent part, thereby imparting energy to it. This portion of the medium, in turn, disturbs another part, thereby causing a flow of energy in a given direction away from the source (see **illus.**). More technically, wave propagation is the result of kinetic energy at one point being transferred into potential energy at an adjacent point, and vice versa. The rate of travel of the disturbance, or velocity of propagation, is determined by the constants of the medium. A stationary wave is the combination of two waves of the same frequency and strength traveling in opposite directions so that no net transfer of energy away from the source takes place. A standing wave is the same but with the returning wave (toward the source) being of lesser intensity than the outwardly traveling wave so that a net transfer of energy away from the source does take place.

Wave motion can occur in a vacuum (electromagnetic waves), in gases (sound waves), in liquids (hydrodynamic waves), and in solids (vibration waves). Electromagnetic waves can also travel in gases, liquids, and solids provided that the electrical conductivity of the medium is not perfect or that the imaginary part of the dielectric constant is not infinitely great. By current usage, elastic waves propagated in gases, liquids, and solids, regardless of whether one can hear them or not, are called acoustic waves.

In this article, it is shown that wave motion can exist in any continuous medium in which potential and kinetic energy can be stored. The wave equation is derived, and solutions for electromagnetic and mechanical media are given. It is shown that longitudinal waves may exist in all mechanical media, and that transverse waves may exist in electromagnetic and solid mechanical media. In all cases, viscosity or heat losses in mechanical media, and resistance and magnetic losses in electromagnetic media are neglected. Only homogeneous and isotropic media are considered. For information related to the present discussion *see* DIFFRACTION; ELECTROMAGNETIC RADIATION; HARMONIC MOTION; HUYGENS' PRINCIPLE; INTERFERENCE OF WAVES; LIGHT; MAXWELL'S EQUATIONS; REFRACTION OF WAVES; SOUND; SUPERPOSITION PRINCIPLE; VIBRATION; WAVE EQUATION; WAVE MOTION IN LIQUIDS.

Fundamental relations. A wave is commonly referred to in terms of either its wavelength or its frequency. In any type of wave motion, these two quantities are related to a third quantity, velocity of propagation, by Eq. (1), where f = frequency, λ =

$$f\lambda = c \tag{1}$$

wavelength, and c = velocity of propagation. The period T is the reciprocal of the frequency, and the amplitude A is the maximum magnitude taken on by the variable of the wave at a given point in space. The physical significance of these quantities is presented in the illustration. It is a basic property of wave motion that the frequency of a wave remains constant under all circumstances except for a relative motion between the source of the wave and the

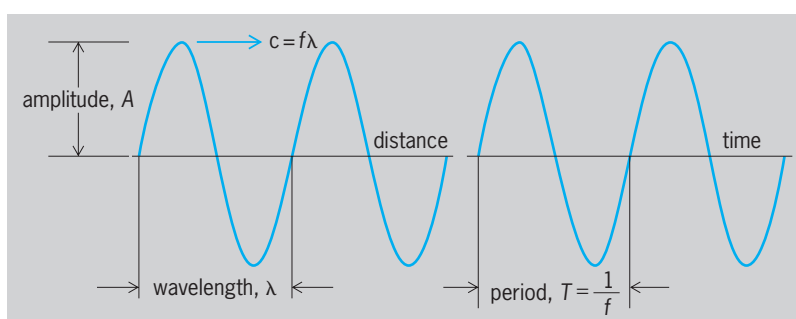

Relation between frequency, wavelength, and velocity in wave propagation.

observer. The case of a frequency shift due to relative motion, known as the Doppler effect, has many interesting applications. On the other hand, the velocity of propagation is dependent on the properties of the medium (and sometimes also on the frequency), and the wavelength will vary with the velocity in accordance with Eq. (1). *See* DOPPLER EFFECT.

Electromagnetic waves. The media in which electromagnetic waves travel possess no elasticity or inertia, but the ability to store energy in the electric and magnetic fields. The electric field corresponds in every respect to the field of an irrotational fluid motion, and the mathematical formulations of the motions of acoustic waves and electromagnetic waves are similar.

In the customary manner, in an electromagnetic medium one defines the vectorial factors **E** as the electric field strength and **H** as the magnetic field strength, with both magnitudes and directions at every point in space.

Next, it is assumed that the medium is homogeneous and isotropic (that is, that the dielectric constant κ and the permeability μ are constants and scalars), that there are no applied electromotive forces in the portion of the medium being dealt with, and that the electrical conductivity σ of the medium is zero. (Waves can be propagated in media where σ is greater than zero but less than infinity.) Then the field equations can be written as Eqs. (2)–(5), where

$$\frac{\kappa}{c}\frac{\partial \mathbf{E}}{\partial t} = \text{curl } \mathbf{H} \tag{2}$$

$$-\frac{\mu}{c}\frac{\partial \mathbf{H}}{\partial t} = \text{curl } \mathbf{E} \tag{3}$$

$$\text{div } \mathbf{H} = 0 \tag{4}$$

$$\text{div } \mathbf{E} = 0 \tag{5}$$

c will be shown to be the velocity of propagation of the electromagnetic wave in a vacuum, where $\kappa = \mu = 1$. *See* CALCULUS OF VECTORS.

Wave equation. J. C. Maxwell recognized about 1863 that these basic equations could be combined to yield an equation resembling the wave equation for mechanical wave motion. Thus he predicted the existence of electromagnetic waves which had not been suspected before. Later, electromagnetic waves proved to be identical with light waves.

The combination of Eqs. (2)–(5) yields the wave equation (6). In a vacuum, the wave equation becomes Eq. (7) and also Eq. (8), where ∇^2 is a scalar

$$\frac{\kappa\mu}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = \nabla^2 \mathbf{E} \tag{6}$$

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} = c^2 \nabla^2 \mathbf{E} \tag{7}$$

$$\frac{\partial^2 \mathbf{H}}{\partial t^2} = c^2 \nabla^2 \mathbf{H} \tag{8}$$

operator called the laplacian. *See* LAPLACIAN.

Plane wave propagation. For simplicity, particular solutions of the wave equation for the case of plane wave propagation will now be sought. A wave is called plane when a family of parallel planes can be taken in the field such that the electric and magnetic field strengths are constant in magnitude and direction at all points of any given member of the family; the planes are called wavefronts, the direction perpendicular to them the wave normal. If the axis of x is taken in the direction of the wave normal, then the wavefronts are parallel to the plane of yz.

Since, for this case, **E** and **H** are to be constant in any one wavefront, the partial derivatives with respect to y or z must vanish. Then the field equations reveal that Eq. (9) is valid. Therefore, insofar as wave

$$\frac{\partial E_x}{\partial t} = \frac{\partial E_x}{\partial x} = \frac{\partial H_x}{\partial t} = \frac{\partial H_x}{\partial x} = 0 \tag{9}$$

propagation is concerned, Eq. (10) holds.

$$E_x = H_x = 0 \tag{10}$$

The meaning of Eq. (10) is that neither **E** nor **H** can have a periodically changing component in the direction in which the wave is traveling. That is to say, the waves are not longitudinal but transverse.

Of the remaining four equations developing out of the field equations, two of them connect E_y and H_z, and the other two E_z and H_y. Dealing with the pairs, one obtains the one-dimensional wave equation (11),

$$\frac{\partial^2 (\,)}{\partial t^2} = c^2 \frac{\partial^2 (\,)}{\partial x^2} \tag{11}$$

where E_y, E_z, H_y, or H_z may be inserted within the parentheses.

The general solution of Eq. (11) can be written in the form of Eq. (12). is The first term of Eq. (12)

$$E_y = F_1(x - ct) + F_2(x + ct) \tag{12}$$

associated with a wave traveling in the positive x direction, and the second term with a wave traveling in the minus x direction. This is true because the argument of F_1 is the same whenever $x = ct$, so that as t becomes greater, the position of a wavefront moves in the positive direction of x. Since for F_2, $-x = ct$, the opposite is true.

If only the outward traveling wave is considered, Eq. (13) holds. Such a wave is transmitted without change of shape, because its position x can always

$$E_y = F_1(x - ct) \tag{13}$$

equal ct. That is, at any given elapsed time t, E_y will have the same value at $x = ct$ as it had at $x = t = 0$. Also, because $c = x/t$, c has the dimensions of a velocity and is the speed at which the wave travels.

Light is also an electromagnetic process so that one should be able to state the optical properties of a substance once its electrical constants are given. However, at the high frequencies of visible light, there is generally dispersion in the medium so that

κ and μ vary with frequency, and this variation must be taken into account if Maxwell's equations are to give a description of optical phenomena which fits measured data.

Acoustic waves in gases. In gases, the existence of inertia (resulting in kinetic energy) and elasticity (resulting in potential energy) is obvious. For example, imagine taking from the medium a small packet of gas enclosed by a weightless deformable membrane. When this packet is squeezed, it is found to have stiffness (or elasticity). Its mass equals the mass of the air molecules contained therein. Also, the deformable membrane keeps the mass inside constant.

From this description, it can be seen that it is a relatively simple matter to describe the motion of such a packet as a function of time. The mass element is labeled in any convenient manner, the most common being by its location at any convenient time. This method is called the lagrangian description of motion.

To derive the equations for wave motion in a gas, one expresses the concepts of inertia by Newton's second law, of elasticity by the perfect gas law, and of the deformable packet by the conservation of mass law. The packet is assumed to be rectangular in shape. To find the equation of motion, the supposition is made that the box is situated in a medium in which the pressure p changes in space at a space rate given by Eq. (14), where **i**, **j**, and **k** are unit vectors

$$\text{grad}\, p \equiv \mathbf{i}\frac{\partial p}{\partial x} + \mathbf{j}\frac{\partial p}{\partial y} + \mathbf{k}\frac{\partial p}{\partial z} \tag{14}$$

in the x, y, and z directions, respectively, and p is the pressure at a point.

The net force **f** acting to move the box in some direction is equal to the vector summation of the gradients in force across the three pairs of faces of the packet times the respective separations of these faces; in the positive direction, Eq. (15) holds. Here

$$\mathbf{f} = -\left[\mathbf{i}\left(\frac{\partial p}{\partial x}\Delta x\right)\Delta y\Delta z + \mathbf{j}\left(\frac{\partial p}{\partial y}\Delta y\right)\Delta x\Delta z + \mathbf{k}\left(\frac{\partial p}{\partial z}\Delta z\right)\Delta x\Delta y\right] = -V\text{grad}\, p \tag{15}$$

V is the the average volume of the packet. The positive gradient causes an acceleration of the box in the negative direction of x.

By Newton's second law, the force acting to accelerate the packet is given by Eq. (16), where **q** is the

$$\mathbf{f} = \rho' V\frac{\partial \mathbf{q}}{\partial t} \tag{16}$$

average vector velocity of the gas in the packet, ρ' the average density of the gas in the packet, and $\rho' V$ the total (constant) mass of the gas in the packet. In writing Eq. (16), it has been assumed that the packet is never displaced from its equilibrium position by a significant part of a wavelength of sound.

Combining Eqs. (15) and (16) yields Eq. (17). In

$$-\text{grad}\, p = \rho_0\frac{\partial \mathbf{q}}{\partial t} \tag{17}$$

keeping with the approximation just stated, it has been assumed that ρ' (the instantaneous density) does not appreciably deviate from the average density ρ_0. These approximations are acceptable, providing the sound pressures are below about 100 dynes/cm^2 (10 newtons/m^2).

Assuming adiabatic expansions and contractions of the gas, and that the gas is perfect, Eq. (18) is

$$\frac{dP}{P} = -\frac{\gamma dV}{V} \tag{18}$$

obtained. Here P is the total pressure in the gas and γ is the ratio of specific heat at constant pressure to specific heat at constant volume.

P and V can be defined by Eqs. (19), where p and

$$\begin{aligned} P &= P_0 + \\ V &= V_0 + \tau \end{aligned} \tag{19}$$

τ are time-varying quantities, and P_0 and V_0 are equilibrium values. If inequalities (20) hold, Eq. (21) is valid.

$$\begin{aligned} p &\ll P_0 \\ \tau &\ll V_0 \end{aligned} \tag{20}$$

$$\frac{1}{P_0}\frac{\partial p}{\partial t} = \frac{-\gamma}{V_0}\frac{\partial \tau}{\partial t} \tag{21}$$

To satisfy the law of mass conservation, one writes Eq. (22) where ξ is the average vector displacement

$$\tau = V_0\, \text{div}\, \xi \tag{22}$$

of the box.

Differentiation of Eq. (22) with respect to time, and substitution of it in Eq. (21), yields Eq. (23).

$$\frac{\partial p}{\partial t} = -\gamma P_0\, \text{div}\, \mathbf{q} \tag{23}$$

The elimination of p from Eqs. (17) and (23) yields the wave equation (24), where by definition Eq. (25) applies.

$$\frac{\partial^2 p}{\partial t^2} = c^2\nabla^2 p \tag{24}$$

$$c^2 = \frac{\gamma P_0}{p_0} \tag{25}$$

One-dimensional plane waves. An acoustic wave is called plane when a family of parallel planes can be taken in the medium such that the pressures and particle velocities are constant in magnitude and the particle velocities are constant in direction at all points of any given member of the family.

Since p and **q** are to be constant in any one wavefront, the partial derivatives with respect to y or z must vanish, and in Eq. (23), Eq. (26) applies. Since

$$\text{div}\, \mathbf{q} = \frac{\partial q_x}{\partial x} \tag{26}$$

Eq. (26) reveals that the x component of **q** is the only component of **q** remaining in the equations basic to the wave equation, the wave is longitudinal.

The one-dimensional wave equation becomes Eq. (27).

$$\frac{\partial^2 p}{\partial t^2} = c^2 \frac{\partial^2 p}{\partial x^2} \tag{27}$$

The general solution to Eq. (27) is exactly the same as that for electromagnetic waves and is given by Eq. (12). The discussion following Eq. (12) is also valid here.

One-dimensional spherical waves. In free space, it is frequently desired to express mathematically the radiation of sound from a nondirectional source. In this case the sound wave expands as it travels away from the source and the wavefront is always a spherical surface. The operator on the right side of Eq. (24) can be written in a form suitable to spherical coordinates.

Assuming equal radiation in all directions, the wave equation in one-dimensional spherical coordinates then becomes Eq. (28).

$$\frac{\partial^2 p}{\partial r^2} + \frac{2}{r}\frac{\partial p}{\partial r} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} \tag{28}$$

Differentiation shows that Eq. (28) can also be written as Eq. (29).

$$\frac{\partial^2 (pr)}{\partial t^2} = c^2 \frac{\partial^2 (pr)}{\partial r^2} \tag{29}$$

Equation (29) has the same form as that of Eq. (27). Hence, the same formal solution applies to either equation, except that the dependent variable is $p(x,t)$ in one case and $pr(r, t)$ in the other case.

The solution to Eq. (29) for the outward traveling wave only (free space) is thus given by Eq. (30).

$$p = \frac{1}{r} F_1(r - ct) \tag{30}$$

Note that, just as for the plane wave, the wave is propagated without change of shape. However, the magnitude of the sound pressure decreases inversely with distance owing to the spreading of the wave as it propagates.

Acoustic waves in liquids. Acoustic waves in liquids obey the same equations as those in gases. The velocity of propagation is greater in liquids than in gases, and viscous losses are often higher. For information on an important example of wave motion in liquids *see* UNDERWATER SOUND.

Elastic and flexural waves in solids. Several different types of acoustic waves may exist in solids, depending on the different manners in which potential energy is stored in the solid. The wave equations associated with several of these types are reviewed in the following paragraphs.

Waves on flexible stretched strings. The wave equation for a flexible stretched string is given by Eq. (31),

$$\frac{\partial^2 \xi_y}{\partial t^2} = c^2 \frac{\partial^2 \xi_y}{\partial x^2} \tag{31}$$

where ξ_y is the displacement of the string at a point x along the string in the y direction (perpendicular to the string). The speed of propagation c is equal to the square root of the ratio of the tension (in dynes) to the linear density of the string (in g/cm^2).

The solution to this equation is identical to that for Eq. (11). Because the motion of the elements of the string is perpendicular to the string, at least for small displacements, the waves are said to be transverse.

Flexural waves in bars. For the purposes of this article, it is assumed that a string has tension with negligible stiffness, while a bar has stiffness without tension. When a bar is bent, its lower half is compressed and its upper half is stretched, or vice versa. When the bending force is removed, the bar attempts to regain its equilibrium position. The restoring force is due to the moment of the forces about the neutral plane in the bar and is related to the cross-sectional dimensions and the Young's modulus of the material.

The transverse wave equation describing the motion of such a bar is Eq. (32), where ξ_y is the displacement perpendicular to the neutral plane of the bar,

$$\frac{\partial^2 \xi_y}{\partial t^2} = -\kappa^2 \frac{Y}{\rho}\frac{\partial^4 \xi_y}{\partial x^4} \tag{32}$$

Y is Young's modulus, p is the density of the bar, and κ is the radius of gyration of the cross section. Values of κ for some of the simpler cross-sectional shapes are given by Eqs. (33).

$$\kappa = \frac{a}{\sqrt{12}} \tag{33a}$$

Rectangle: Length b parallel to center line, width a perpendicular to center line, Circle: Radius a,

$$\kappa = \frac{a}{2} \tag{33b}$$

Circular ring: Outer radius a, inner radius b,

$$\kappa = 0.5(a^2 + b^2)^{1/2} \tag{33c}$$

Equation (32) differs from the usual wave equation (11) in that it has a fourth derivative with respect to x instead of a second derivative.

The function $F_1(x - ct)$ is not a solution, so that a bar satisfying Eq. (32) cannot have waves traveling along it with a velocity independent of frequency and with an unchanged shape.

If one considers excitation of the bar by one frequency f at a time, a solution to Eq. (32) may be written as Eq. (34), where $\mu = (f^2\rho/4\pi^2 Y\kappa^2)^{1/4}$,

$$\xi_y = A \cos\,[2\pi(\mu x - ft) + \phi] \tag{34}$$

$f = 2\pi\mu^2\kappa\sqrt{Y/\rho}$, A = amplitude, and ϕ = phase angle.

The velocity of propagation of the wave is equal to $(f/\mu) = (4\pi^2 Y\kappa^2/\rho)^{1/4}\sqrt{f}$. It obviously depends on the frequency f of the exciting wave. Such a velocity for a simple harmonic wave is called the phase velocity. For a complex wave with several components differing in frequency, the shape must change with distance of travel. A bar is sometimes said to be a dispersing medium for waves of bending. *See* PHASE VELOCITY.

There is also a possibility for a longitudinal wave in such a bar. In this case the bar is excited at an end

in the direction of its longitudinal axis. The wave equation is identical to Eq. (31), except that ξ_x replaces ξ_y. The motions of the particles of the bar are in a line with the longitudinal axis of the bar. Longitudinal and transverse waves may, and generally do, coexist.

Waves in plates. Wave motion in a plate is similar to that in a bar. The bending of a plate compresses the material on the inside of the bend and stretches it on the outside. But when a material is compressed, it tries to spread out in a direction perpendicular to the compressed force, so that when a plate is bent downward in one direction there is a tendency for it to curl up in a direction at right angles to the bend. The ratio of the sideways spreading to the compression is called Poisson's ratio and is designated by the letter σ.

The wave equation for the plate is given by Eq. (35), where η is the displacement of the plate

$$\Delta^4\eta + \frac{12}{(hc_L)^2}\frac{\partial^2\eta}{\partial t^2} = 0 \qquad (35)$$

perpendicular to its surface, h is the thickness of the plate, and c_L is the longitudinal plate velocity given by Eq. (36).

$$c_L = \sqrt{\frac{Y}{\rho(1-\sigma^2)}} \qquad (36)$$

From a solution to the equation it is possible to find the velocity of propagation for the transverse (bending) wave, as in Eq. (37). Just as for the bar, the

$$c_B = \sqrt{1.8hc_L}\sqrt{f} \qquad (37)$$

velocity of propagation is dependent on frequency, and the plate is said to be a dispersing medium.

Steady-state wave motion. Very frequently the source of a wave is steady and the wave produced by it is periodic, at least for a long period of time compared to the buildup and decay times of the wave. When this is true, it is convenient to specify the functions $F_1(x - ct)$ and $F_2(x + ct)$ of Eq. (12) by a summation of sinusoidal functions, as in Eq. (38), where

$$F(x - ct) = \sum_\nu A_\nu \cos\left[\omega_\nu\left(t - \frac{x}{c}\right) + \theta_\nu\right] \qquad (38)$$

$\omega_\nu = 2\pi f_\nu$; f_ν is the frequency of the νth component of the wave, and θ_ν is the phase angle of the νth component, which is determined when $x = t = 0$.

From the well-known theory of Fourier series, any function $F_1(x - ct)$, if it repeats on itself periodically, can be represented by a linear summation of sine or cosine wave functions. The component with the lowest frequency f_1 is the fundamental component or the first harmonic; each higher component, called the $(\nu - 1)$th overtone, or the νth harmonic, has a frequency equal to νf_1. In other words, ν is an integer, and the waveform $F_1(x - ct)$ is represented by a linear summation of cosine terms with frequency components f_ν that are harmonically related (f, $2f$, $3f$, $4f$, and so forth) and at $t = x = 0$ have phase angles $\theta_1, \theta_2, \theta_3$, and so forth, that may differ from zero. *See* FOURIER SERIES AND TRANSFORMS.

As a further simplification, each component of Eq. (38) is usually written as Eq. (39), where $k = \omega/c$

$$P = A\cos(\omega t - kx + \theta) \qquad (39)$$

is called the wave number. The period of this component is T and equals $1/f$. As before, θ is the phase angle in radians.

Note that A is the peak amplitude of the component wave. Generally, in acoustics, measuring devices read the root-mean-square value of a wave, so that the intensity of the wave is designated by $A_{\text{rms}} = A/\sqrt{2}$, and the strength of a wave represented by Eq. (38) is given by Eq. (40).

$$A_{\text{rms}} = \sqrt{A_{1_{\text{rms}}}^2 + A_{2_{\text{rms}}}^2 + \cdots} \qquad (40)$$

See ROOT-MEAN-SQUARE. Leo L. Beranek

Bibliography. L. L. Beranek, *Acoustics*, 1954, paper 1986; H. Georgi, *The Physics of Waves*, 1992; L. E. Kinsler et al., *Fundamentals of Acoustics*, 4th ed., 1999; P. Lorrain, *Electromagnetic Fields and Waves*, 3d ed., 1995; I. G. Main, *Vibrations and Waves in Physics*, 3d ed., 1993; H. J. Pain, *Physics of Vibrations and Waves*, 5th ed., 1999.

Wave optics

The branch of optics which treats of light (or electromagnetic radiation in general) with explicit recognition of its wave nature. The counterpart to wave optics is ray optics or geometrical optics, which does not assume any wave character but treats the propagation of light as a straight-line phenomenon except for changes of direction induced by reflection or refraction. *See* GEOMETRICAL OPTICS; OPTICS.

Any optical phenomenon which is correctly describable in terms of geometrical optics can also be correctly described in terms of wave optics. However, the many phenomena of interference, diffraction, and polarization are incontrovertible evidence of the wave nature of light, and geometrical optics often gives an incomplete or incorrect description of the behavior of light in an optical system. This is especially true if changes of refractive index occur within a space which is of the order of several wavelengths of the light. *See* DIFFRACTION; ELECTROMAGNETIC WAVE; INTERFERENCE OF WAVES; POLARIZED LIGHT. Richard C. Lord

Wave-shaping circuits

Electronic circuits used to create or modify specified time-varying electrical voltage or current waveforms using combinations of active electronic devices, such as transistors or analog or digital integrated circuits, and resistors, capacitors, and inductors. Most wave-shaping circuits are used to generate periodic waveforms. *See* INTEGRATED CIRCUITS; TRANSISTOR.

The common periodic waveforms include the square wave, the sine and rectified sine waves, the sawtooth and triangular waves, and the periodic arbitrary wave. The arbitrary wave can be made to conform to any shape during the duration of one period. This shape then is followed for each successive cycle. *See* FUNCTION GENERATOR; WAVEFORM.

A number of traditional electronic and electromechanical circuits are used to generate these waveforms. Sine-wave generators and *LC*, *RC*, and beat-frequency oscillators are used to generate sine waves; rectifiers, consisting of diode combinations interposed between sine-wave sources and resistive loads, produce rectified sine waves; multivibrators can generate square waves; electronic integrating circuits operating on square waves create triangular waves; and electronic relaxation oscillators can produce sawtooth waves. *See* ALTERNATING CURRENT; DIODE; MULTIVIBRATOR; OPERATIONAL AMPLIFIER; OSCILLATOR; RECTIFIER.

In many applications, generation of these standard waveforms is now implemented using digital circuits. Digital logic or microprocessors generate a sequence of numbers which represent the desired waveform mathematically. These numerical values then are converted to continuous-time waveforms by passing them through a digital-to-analog converter. Digital waveform generation methods have the ability to generate waveforms of arbitrary shape, a capability lacking in the traditional approaches. *See* CIRCUIT (ELECTRONICS); DIGITAL-TO-ANALOG CONVERTER; LOGIC CIRCUITS; MICROPROCESSOR.

Philip V. Lopresti

Waveform

The pictorial representation of the form or shape of a wave, obtained by plotting the amplitude of the wave with respect to time. There are an infinite number of possible waveforms (see **illus.**). One such waveform is the square wave, in which a quantity such as voltage alternately assumes two discrete values during repeating periods of time. Other waveforms of particular interest in electronics are the sine wave and rectified sine wave, the sawtooth wave and triangular wave, and the arbitrary wave—a recurrent waveform which takes on an arbitrary shape over one complete cycle; this shape is then repeated in successive cycles.

Each of these waveforms has a shape which repeats periodically in time. It is possible to characterize any of them mathematically by a Fourier series, a weighted sum of terms consisting of the basic periodic trigonometric functions: sines and cosines. A periodic waveform thus can be represented as a constant or dc term, plus a sum of harmonically related sine and cosine terms where the sine and cosine frequencies are integral multiples of the fundamental frequency. The Fourier series is given beside each waveform in the illustration as a function of time t, where E_m is the maximum value of the wave and T is the period. *See* FOURIER SERIES AND TRANSFORMS; NONSINUSOIDAL WAVEFORM; WAVE-SHAPING CIRCUITS.

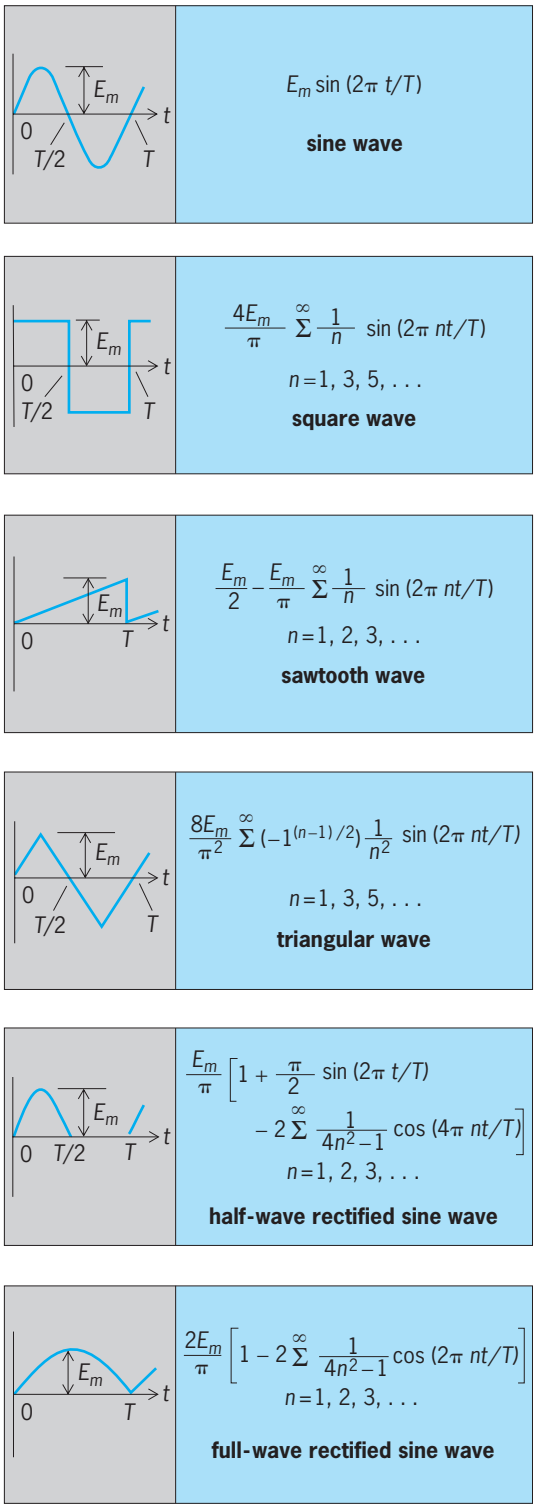

Common electrical waveforms.

Philip V. Lopresti

Waveform determination

A waveform describes the variation of a quantity with respect to time. The necessary measurements to determine a waveform are normally carried out and

presented in one of two ways: the amplitude may be presented as a function of time (time domain), or an analysis may be given of the relative amplitudes and phases of the frequency components (frequency domain). Although the simplest instruments measure and display the information in the same domain, it is possible to convert the results from the time domain to the frequency domain and vice verse by mathematical processing.

Waveforms may be divided into two classes, depending on whether the signal is repeated at regular intervals or represents a unique event. The former signal is defined as a periodic or continuous wave, the latter as an aperiodic signal or transient. *See* ELECTRIC TRANSIENT; NONSINUSOIDAL WAVEFORM; WAVEFORM.

Time domain. The oscilloscope is an example of an instrument that measures and displays directly in the time domain, by deflecting an electron beam in a vertical direction in accordance with the signal while scanning at a uniform rate in the horizontal direction. The position of the beam is revealed by a fluorescent screen. *See* GRAPHIC RECORDING INSTRUMENTS; OSCILLOSCOPE.

Frequency domain. Several methods may be used to obtain the spectral content of a waveform. In the simplest, the signal is applied to a filter that is manually tuned in turn to each frequency that is expected to be present. In order to automate the measurement, the tuning of a filter may be varied by a linear, logarithmic, or other sweep and the resulting output displayed. However, any transient signal that occurs at a frequency that is different from the filter frequency will be missed. This is an important restriction that limits the technique to continuous waveforms. Specific instruments that display the frequency components of the signal are often called spectrum analyzers or harmonic analyzers. *See* ELECTRIC FILTER; HETERODYNE PRINCIPLE; SPECTRUM ANALYZER.

For applications such as the analysis of music and speech, it is necessary to simultaneously measure the intensity variation at a number of frequencies. Instead of the swept filter, an array of separate fixed tuned filters may be used, each adjusted to respond to a slightly different frequency. The amplitude of the signal in each filter is sampled in turn and displayed, giving a histogram of the frequency components of the waveform. This approach, though simple in concept, has largely been replaced by digital techniques, which are simpler to implement and offer greater flexibility. The time variation of the signals at each frequency can be analyzed as a separate waveform; this is called spectral analysis.

Spectral analysis extends analysis of signals in the frequency domain to nonperiodic and other signals by introducing the concept of power spectral density (PSD), in which the power estimates are displayed as a continuous function of frequency instead of being limited to discrete harmonics. The data collected during a given time interval either can be used to provide the mean power spectral density over the whole period or can be subdivided to show how the power spectral density is varying with time. In the latter case, the power spectral density estimates will be less accurate. A further extension of the technique allows additional dimensions to the array of data so that the technique can be applied to, for example, image analysis. These modern techniques of waveform determination provide the crucial key to advances in fields as widely separated as the identification of radar echoes, biomedical signal analysis, speech processing, and the transmission of audio and video signals over links of very low bandwidth. *See* DATA COMPRESSION; IMAGE PROCESSING; SIGNAL PROCESSING; SPEECH RECOGNITION.

Digital techniques. Many modern instruments use techniques in which the signal to be measured is sampled and digitized. The key elements of digital oscilloscopes are a high-speed analog-to-digital converter, sufficient high-speed memory to store the results, and a display. Because the signal is continuously digitized and stored in memory, these instruments are ideal for measuring transient waveforms. As the data can be read out of digital storage at any convenient speed, low-bandwidth display circuits are adequate. *See* ANALOG-TO-DIGITAL CONVERTER.

Once the data have been collected in digital form, they can be processed in many different ways. The application of a discrete Fourier transformation (DFT) to the amplitude data enables the information to be presented in the frequency domain. In this way, harmonic distortion that was invisible on a directly displayed waveform can be made obvious. The fast Fourier transform (FFT) is a method by which the arithmetic involved is simplified and thus accelerated. Wavelet transforms are another powerful technique that can be used to analyze waveforms. These are well suited to showing localized intermittent periodicities such as natural events and transients. *See* FOURIER SERIES AND TRANSFORMS; SPEECH RECOGNITION; WAVELETS.

Typical desktop computers can carry out fast Fourier transform transformations on 1024 data points in less than 0.1 s. This allows transformations to be carried out on speech signals in real time. Some applications may need more processing power than is available with a typical desktop computer. Digital signal processor and field-programmable gate array integrated circuits can carry out fast Fourier transforms at much higher speeds. The best can carry out transformations on 1024 data points at sample rates of over 350 kHz (in 2.85 μs). The availability of such computing power makes it practical to design instruments that collect the data in the most convenient way and display the results in the form most useful to the user. *See* LOGIC CIRCUITS.

Parametric representation. Often, waveforms contain too much detailed information for easy comparison. Instead, they can be represented by simple parameters, such as the pulse duration or transition duration. Because it is important for business that everyone agrees common definitions, the Institute of Electrical and Electronic Engineers has written a standard which sets out definitions of all these parameters and how they should be measured. The transition durations (also commonly known as rise

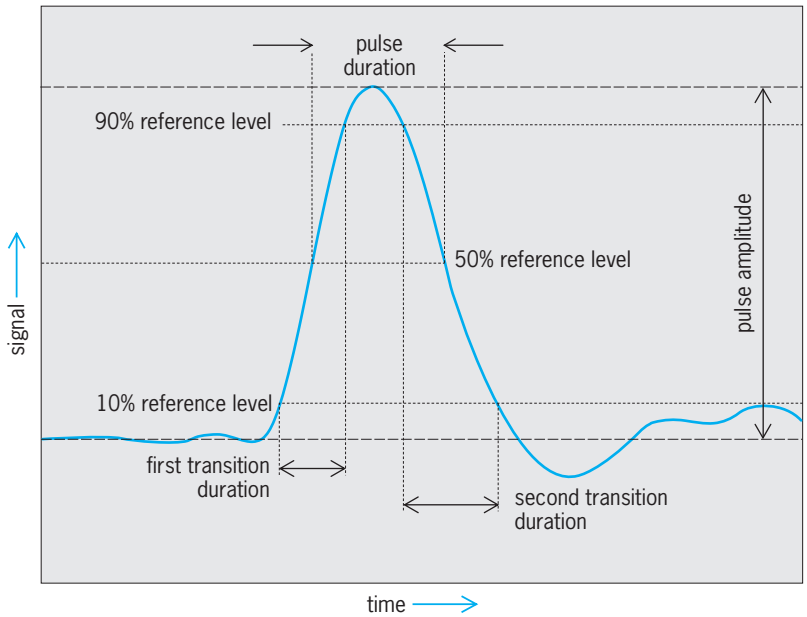

Typical pulse waveform showing the definitions of the transition duration and pulse duration parameters.

time and fall time) are measured between the 10% and the 90% pulse amplitude reference levels. The pulse duration (also commonly known as full width at half maximum or FWHM) is measured at the 50% reference level. The pulse duration and transition durations are the key parameters often used to characterize a pulse waveform (see **illustration**).

Very short pulses. All instruments used for waveform determination have a finite temporal resolution, and this is related to the instrument bandwidth. A larger-bandwidth instrument will have a shorter temporal resolution and vice versa. The temporal resolution is defined in terms of the pulse duration of the recorded waveform for an input pulse of infinitesimally short duration. How much the waveform is distorted by this temporal response depends on the difference between the response time of the instrument and the duration of the artifacts within a pulse. Equivalently, it depends on whether the spectral content of the pulse includes frequencies that are higher than the maximum to which the instrument is sensitive, its bandwidth. Very short pulses are loosely defined as those for which the available measurement instruments will significantly distort the waveform. Thus, whether or not a pulse is said to be very short depends on the available measurement instruments. Pulses with durations within one order of magnitude of the temporal response of the instrument might be considered to be distorted significantly. In such cases, the duration is specified by the time between the points of half the maximum signal, the pulse duration (see illustration).

The oscilloscope is widely used for recording electrical pulses. Sampling oscilloscopes can be used only for repetitive pulses, and transient pulses are recorded by real-time oscilloscopes and transient digitizers, which are about five times slower. In a sampling oscilloscope, a sufficiently large number of samples is taken to define the waveform. The sampling rate is not necessarily high and a different part of the waveform is measured for each trigger event.

Optical pulses can be converted to electrical pulses by devices such as a photodiode and recorded by the oscilloscope. Where a photodiode is used, the combined response of the photodiode and recording instrument must be considered. Instruments that combine a photodiode and a sampling oscilloscope together are used to measure optical communications data signals. *See* OPTICAL DETECTORS.

Very short pulses for which the waveform can be unambiguously determined by a sampling oscilloscope include electrical pulses where the durations of artifacts within the pulse are at least about 5 picoseconds. Electric fields in microwave circuits with frequencies from gigahertz to terahertz can be measured using electro-optic sampling. The electric field of a circuit changes the birefringence of an electrooptically active crystal, modifying its behavior in transmitting optical radiation. The change in birefringence is probed using a femtosecond laser, resulting in time resolutions in the subpicosecond range. *See* BIREFRINGENCE; ELECTROOPTICS; LASER.

Typically, optical pulse widths of less than 10 femtoseconds can be generated in the laboratory by commercial laser systems. To give an idea of scale, an 85-fs light pulse is about 25 μm (0.001 in.) long. These short light pulses can be used for many applications such as absolute frequency standards, to generate terahertz radiation, or to generate even shorter pulses in the extreme ultraviolet region. The durations of these pulses are estimated by autocorrelation. This technique does not give the pulse shape but only the FWHM for an estimated pulse shape. An extension of the autocorrelation technique allows the shapes of optical pulses of subpicosecond duration to be measured using the frequency-resolved optical gating (FROG) technique. The method is analogous to an oscilloscope but almost a thousand times faster. *See* FREQUENCY MEASUREMENT; OPTICAL PULSES; SUBMILLIMETER-WAVE TECHNOLOGY; ULTRAFAST MOLECULAR PROCESSES.

The streak camera is the fastest instrument used to record pulse shape directly. It is an electron imaging tube that converts an optical pulse to a distribution of intensities across a phosphor screen, recording the pattern using a photodiode array. The response of the streak camera is presently limited to about 500 ps.

Often the technology used to generate a short pulse is common to that of the measurement instrument, so that both are limited to a similar response. The resulting waveform is the convolution of the pulse shape with the instrument response. Mathematical techniques can be used to reduce this distortion, depending on the extent to which the response of the instrument is known or calculable. This process is called deconvolution.

David A Humphreys

Bibliography. R. N. Bracewell, *The Fourier Transform and Its Applications*, 3d ed., McGraw-Hill, 1999; *IEEE Standard on Transitions, Pulses and*

Related Waveforms, IEEE Std 181, 2003; S. M. Kay, *Modern Spectral Estimation: Theory and Application*, Prentice Hall, 1988, paper 1999; L. S. Marple, *Digital Spectral Analysis with Applications in C, FORTRAN, and MATLAB*, Prentice Hall, 1987; N. S. Nahman, Picosecond-domain waveform measurements, *Proc. IEEE*, 66:441–454, 1978; S. Prentiss, *The Complete Book of Oscilloscopes*, 2d ed., 1992.

Waveform generator

A term usually applied to an instrument that generates voltage waves using digital technology. These digital instruments have largely supplanted the analog instruments that traditionally served in this role. Digital instruments have a number of important advantages over analog ones: they are more flexible and can be used to generate waveforms that are beyond the capabilities of analog instruments, they produce more accurate replicas of the desired mathematical waveforms, and their output is more stable and predictable over time. *See* FUNCTION GENERATOR; SIGNAL GENERATOR; WAVE-SHAPING CIRCUTS.

The basic components of these instruments are: a digital signal memory that contains numeric samples of the desired signal waveform, a sequencing system that accesses this memory in a predetermined pattern, a high-accuracy digital-to-analog converter that transforms the contents of the accessed memory location to proportional analog voltages, a low-pass filter to smooth the output of the converter, and a digital clock—the timing device that sets the memory access rate. *See* DIGITAL-TO-ANALOG CONVERTER; ELECTRIC FILTER; SEMICONDUCTOR MEMORIES.

The two most common approaches to implementing these instruments are direct digital synthesis (DDS) and arbitrary waveform generation (AWG). Each has its own advantages and each plays an important role. Direct digital synthesis typically is used to generate high-accuracy versions of standard analog signal waveforms: sine, triangular, and square waves. Instruments that employ arbitrary waveform generation are much more flexible and can generate a repetitive signal waveform that takes on values limited only by the word size of each memory cell and by the number of memory locations in the instrument. Instruments that use direct digital synthesis can be programmed to change their output from one form to another essentially instantaneously, while devices that use arbitrary waveform generation are limited by the time it takes to replace their signal memory contents with new signal patterns.

Direct digital synthesis. Direct digital synthesis employs a fixed-rate digital-to-analog converter followed by a high-quality fixed-frequency low-pass filter. A phase increment register and a phase accumulator control the output frequency of the generator. To illustrate the concept, note that one full cycle of a sine wave is traversed as its angle covers 360° and assume that consecutive memory locations hold values of the sine function in one-degree increments: sin(0°), sin(1°), . . . , sin(359°). Sequentially scanning and converting these memory locations will take 360 "ticks" of the timing clock to generate one cycle of the output wave and the resulting frequency would be 1/360 the frequency of the clock. The phase increment register sets the memory step size and the phase accumulator increases by this amount for each clock tick. Thus, the accumulator contains a memory address representing the angle and the addressed memory cell contains the value of the sine of the angle. The direct digital synthesis memory bank is often called a look-up table since the approach is to "look up" values of the function to be generated. Since the sine of 360° is identical to the sine of 0°, and since the function repeats periodically every 360°, the accumulator is programmed to automatically return to 0° when its angle tries to pass 359°, and the conversion process repeats indefinitely.

For a 360-MHz clock and an increment size of 1°, the signal frequency will be 1 MHz. Increasing the

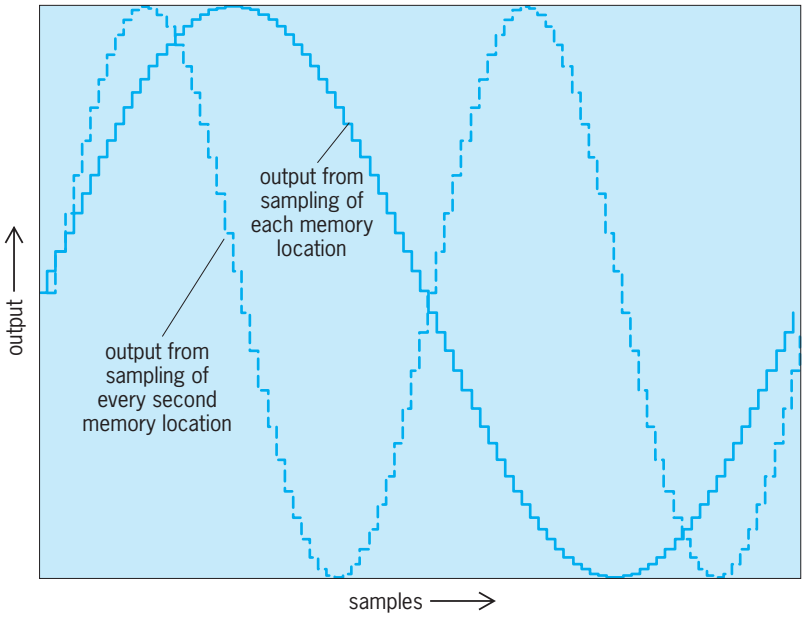

Fig. 1. Stepped-response output of direct digital synthesis (DDS) digital-to-analog converter.

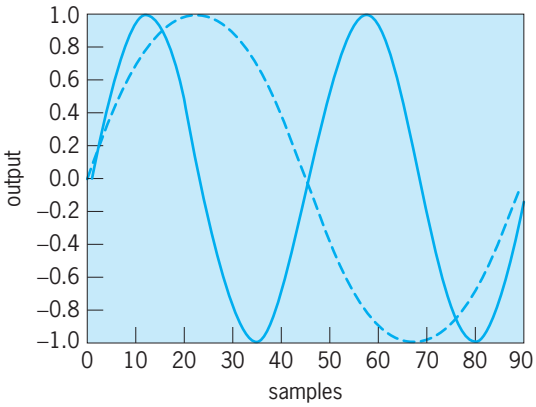

Fig. 2. Direct digital synthesis (DDS) output smoothed after low-pass filtering.

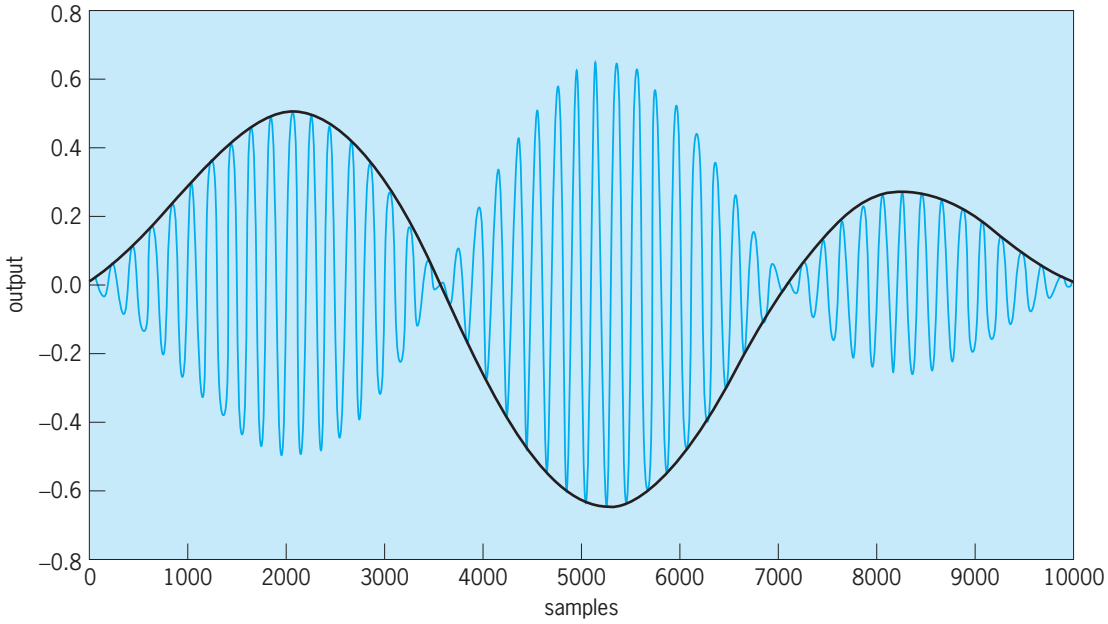

Fig. 3. Sinusoidal carrier modulated by complex pulses.

phase increment register value to 2° will access every second memory location, the sine wave will complete one cycle in 180 clock "ticks," and the signal frequency will be 1/180 the frequency of the clock, or 2 MHz in this case.

In practice, the minimum phase increment available for an instrument is much smaller than 1°. Some high-performance generators contain phase increments smaller than 1 billionth of a cycle.

The concept is illustrated in **Figs. 1** and **2**. Figure 1 shows the output of the digital-to-analog converter for a 90-sample sine wave look-up table, where each successive location contains the sine of an angle 4° greater than the preceding one. Sampling each memory location in sequence generates the first wave and sampling every second location generates the second. The steps in the figure correspond to the increments in the digital-to-analog converter output as each memory location is accessed. Figure 2 shows the smoothed, low-pass filtered output for each generated frequency. The physical output frequency depends on the timing clock frequency.

Arbitrary waveform generation. Arbitrary waveform generation is conceptually simpler than direct digital synthesis. As in direct digital synthesis, numerical values for a waveform are stored in consecutive memory locations, but the locations are always accessed sequentially with no skipping. The values inserted in each location can be arbitrary—hence the name arbitrary waveform generator. As in the direct digital synthesis case, the waveforms are repetitive, returning to the start point as the end of the signal memory is reached.

Sine waves and the rest of the waveforms that can be generated by direct digital synthesis methods also can be generated by an arbitrary waveform generator, but this is not the principal use for these instruments. Their ability to generate almost any conceivable periodic waveform is usually used to generate waveforms with much more complex shapes. These values can be chosen from preconfigured signals that come with the instrument, extracted from real-world data via analog-to-digital conversion, copied from other programs such as spreadsheets, or entered point-by-point by hand.

They can, for example, mimic the behavior of the quadrature amplitude modulation (QAM) systems that form the backbone of most digital communications links. A two-channel arbitrary waveform generator can generate QAM signals consisting of a sine waveform with amplitude and phase variations on one channel and a cosine waveform with separate amplitude and phase variations on the other, combine these waves, and distort them in ways that correspond to imperfect transmitter hardware or to poor communications links. The arbitrary waveform generator thus can generate the input signals needed to submit a link receiver to carefully controlled stress tests. *See* AMPLITUDE MODULATOR; DATA COMMUNICATIONS; MODULATION.

Figure 3 illustrates an example of how an arbitrary waveform generator might generate a complex signal, in this case the amplitude modulation of the sinusoidal carrier of a QAM transmitter by a sequence of pulses. Such a waveform cannot be generated by an analog function generator or by a waveform generator based on direct digital synthesis.

Philip V. Lopresti

Bibliography. B.-G. Goldberg, *Digital Frequency Synthesis Demystified*, LLH Technology, Eagle Rock, VA, 1999; J. H. Reed, *Software Radio: A Modern Approach to Radio Engineering*, Prentice-Hall, Upper Saddle River, NJ, 2002.

Waveguide

A device that constrains or guides the propagation of electromagnetic radiation along a path defined by the physical construction of the guide. Electromagnetic

waves may propagate in space, as radio waves, but for many purposes waves need to be guided with minimum loss from the generating point to a point of application. Several guiding systems are of importance, including two-conductor transmission lines, various forms of striplines used in microwave integrated circuits, hollow-pipe waveguides, and dielectric waveguides. Since transmission lines and striplines are discussed in other articles, the emphasis in this article is on hollow-pipe and dielectric guides. Hollow-pipe guides are used primarily in the microwave region of the spectrum, dielectric guides primarily in the optical region. *See* COAXIAL CABLE; ELECTROMAGNETIC WAVE TRANSMISSION; MICROWAVE; TRANSMISSION LINES.

Hollow-Pipe Waveguides

Hollow-pipe waveguides consist of a dielectric region, usually air, surrounded by a closed good conductor such as silver, copper, aluminum, or brass. The cross-sectional shape is usually rectangular (**Fig. 1**), but may be circular or of a variety of other shapes. Voltage and current concepts, so useful for transmission lines, are not so useful for waveguides. Distributions of electric and magnetic fields, obtained from Maxwell's equations, are needed. *See* MAXWELL'S EQUATIONS.

Solution of the field equations shows that there is an infinite number of modes for these guides, where a mode is a solution that maintains its transverse pattern but attenuates and shifts in phase as it propagates along the guide. These modal patterns may be likened to the resonant modes of a drumhead, with low-order modes having only a few transverse variations and high-order modes having many.

All modes in these guides have cutoff properties. Below a certain critical frequency, the mode will not propagate but only attenuates if excited; above the cutoff frequency, though, it may propagate with a finite phase velocity and small attenuation (zero attenuation for perfect dielectric and conductor). Higher-order modes usually have higher cutoff frequencies.

If the guiding pipe is perfectly conducting, the modes may be divided into transverse magnetic (TM) and transverse electric (TE) types. For the former, the magnetic field is confined to the transverse plane, while the electric field is so confined for the latter. These classifications remain useful for the practical metallic conductors used for such guides.

Principal mode in rectangular. The principal mode in a guide of rectangular cross section (Fig. 1), with air dielectric, constitutes the most commonly used guide and mode. The electric field of the principal mode has only a vertical component with no variations vertically, and a half-wave sinusoidal variation in the transverse (x) direction. The magnetic field has only transverse (x) and axial (z) components, forming closed loops surrounding the time-varying electric field (Fig. 1). The current flow on the interior of the metal walls is normal to the tangential magnetic field at the walls and has both axial and transverse components. The electric field is confined to the transverse plane, so the wave is a transverse electric type. Since the electric field has

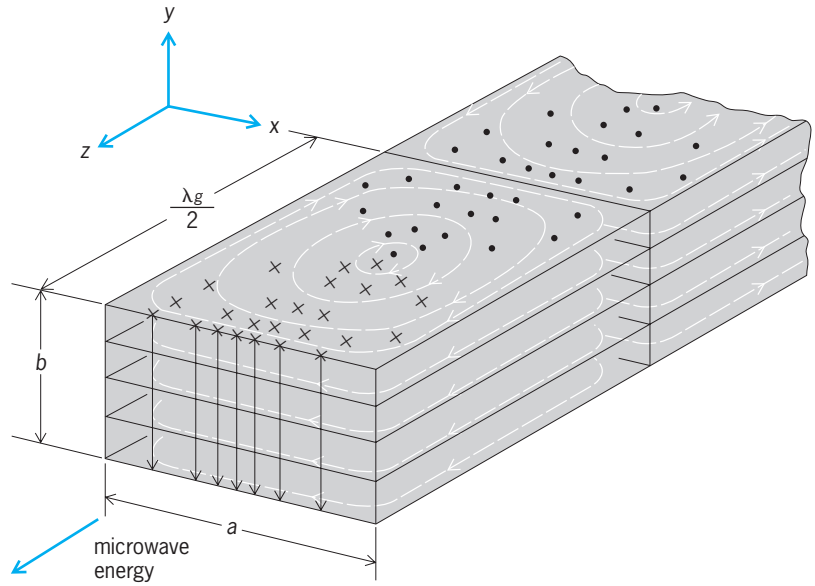

Fig. 1. Rectangular hollow-pipe waveguide with field patterns for a propagating TE_{10} mode. Solid lines indicate electric field and broken lines magnetic field; ×'s indicate energy *E* pointing down; solid circles indicate *E* pointing up. (*After MIT Radar School Staff, Principles of Radar, 1952*)

one variation in x and none in y, the designation is TE_{10}.

For a given guide of width a, the mode will propagate average power only for frequencies higher than a cutoff frequency f_c, which corresponds to a frequency for which the width is just a half-wavelength. For frequencies below f_c, the longer half-wavelength will not "fit" in the guide and, if excited, will only attenuate exponentially with z, without any average power transfer. Thus, hollow-pipe waveguides used for power transfer must have transverse dimensions that are of the order of wavelength, so they are most used in the microwave region of the spectrum.

The expressions for the fields of this TE_{10} mode, obtained from Maxwell's equations subject to the boundary conditions of the walls (taken here as perfectly conducting), are given by Eqs. (1), (2), and (3).

$$E_y = E_0 \sin\frac{\pi x}{a}\cos\omega\left(t - \frac{z}{v_p}\right) \tag{1}$$

$$H_x = -\frac{E_0}{\mu v_p}\sin\frac{\pi x}{a}\cos\omega\left(t - \frac{z}{v_p}\right) \tag{2}$$

$$H_z = -\frac{\pi E_0}{\omega\mu a}\cos\frac{\pi x}{a}\sin\omega\left(t - \frac{z}{v_p}\right) \tag{3}$$

Here a is the guide width; $\omega = 2\pi f$, where f is the frequency; μ is the permeability, equal to $4\pi \times 10^{-7}$ henry per meter for nonmagnetic dielectrics; E_0 is the electric field strength in volts per meter; and v_p is the phase velocity in the axial direction. This velocity is given by Eq. (4), where c is the velocity of light

$$v_p = \frac{c}{\sqrt{\epsilon_r}}\left[1 - \left(\frac{f_c}{f}\right)^2\right]^{-1/2} \tag{4}$$

in vacuum; ϵ_r is the relative permittivity or dielectric

constant; and f_c is the cutoff frequency, equal to $c/(2a\sqrt{\epsilon_r})$. Equation (4) shows that the phase velocity becomes imaginary, indicating no real propagation of energy, for $f < f_c$. The guide wavelength λ_g for $f > f_c$ is defined as the axial distance for which a guided mode changes phase by 2π, and is a useful concept. For the TE_{10} mode, and a nonmagnetic dielectric, it is given by Eq. (5), where λ_0 is the free-space wavelength.

$$\lambda_g = \frac{v_p}{f} = \frac{\lambda_0}{\sqrt{\epsilon_r}}\left[1 - \left(\frac{f_c}{f}\right)^2\right]^{-1/2} \tag{5}$$

See PHASE VELOCITY.

The axial power flow of the mode is obtained from the transverse field components E_y and H_x from Poynting's theorem and is found to be given by Eq. (6). This also becomes imaginary when v_p is imaginary, that is, for $f < f_c$ *See* POYNTING'S VECTOR.

$$P = \frac{abE_0^2}{4\mu v_p} \qquad \text{watts} \tag{6}$$

The maximum power transfer capability of this guide is limited by the breakdown electric field at the guide center ($x = a/2$). Taking air breakdown as 15 kV/cm, the power limit for a guide at 10 GHz would be about 260 kW.

Power or energy is said to travel at a velocity v_g, called group velocity, given for a nonmagnetic dielectric by Eq. (7). Group velocity is actually the velocity of the envelope of a modulated wave, so it could represent the envelope of a short pulse of the microwave signal. Both v_p and v_g are functions of frequency, so these guides are dispersive and the pulse spreads or changes shape as it propagates. *See* GROUP VELOCITY.

$$v_g = \frac{c}{\sqrt{\epsilon_r}}\left[1 - \left(\frac{f_c}{f}\right)^2\right]^{1/2} \tag{7}$$

Higher-order modes in rectangular guides. The waveguide of rectangular cross section has other modal solutions (**Fig. 2**). TM_{21}, for example, denotes a wave with magnetic field confined to the transverse plane, having two half-sine variations with x and one with y. The cutoff frequency for either the TM or TE mode with m variations in x and n with y, for a nonmagnetic dielectric, is given by Eq. (8). (The numbers m or n, but not both, can be zero for transverse electric waves, but both must be nonzero for transverse magnetic waves.) The cutoff frequency increases with m and n (for a given guide size), and, if $a > b$, that for the TE_{10} mode is the lowest.

$$(f_c)_{mn} = \frac{c}{\sqrt{\epsilon_r}}\left[\left(\frac{m}{2a}\right)^2 + \left(\frac{n}{2b}\right)^2\right]^{1/2} \tag{8}$$

It is usually desirable to select the guide size so that only one mode propagates, so that there is not interference among several modes propagating with different phase velocities. For example, for the microwave X-band ($f \sim 10$ GHz), a standard guide size (interior dimension) is 0.9 by 0.4 in. (2.3 by 1.0 cm). The cutoff frequency for the TE_{10} mode is then 6.56 GHz so that $f > f_c$ for this mode; for the TE_{01} mode it is 14.76 GHz, and it is 16.15 GHz for TE_{11} or TM_{11}. The higher-order modes may be excited at the beginning of the guide, and at discontinuities, but if $f < f_c$ they are localized and do not propagate. Their reactive energy may produce reflections or other circuit effects that must be taken into account in the matching process.

Although the expressions given are for guides with ideal conductors and dielectrics, good conductors and dielectrics do not change these much but do produce an attenuation of the mode as it propagates. Most guides are air-filled, and typical attenuations for the TE_{10} mode in a rectangular guide are 0.075 dB/ft (0.25 dB/m) at 10 GHz.

Circular hollow-pipe waveguides. For some applications, the symmetry of a circular-cylindrical guide is desirable. These also have an infinite number of mode types that may be divided into transverse electric and transverse magnetic classes when losses are small. Cutoff frequencies for the modes of circular

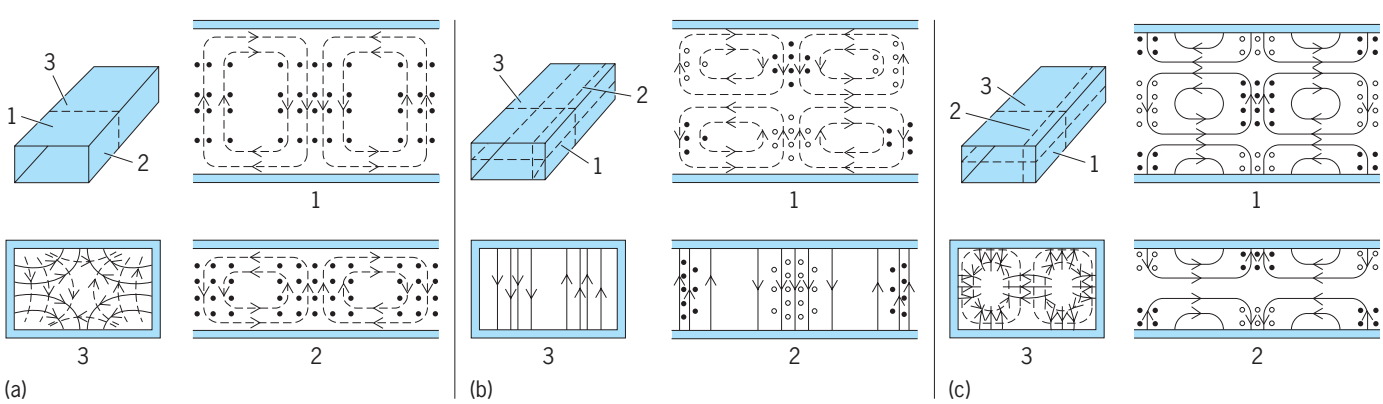

Fig. 2. Various modes of rectangular hollow-pipe guide. Picture of waveguide in each part of figure shows orientations of waveguide walls or cross sections, identified as 1, 2, and 3, on which fields are shown. Solid lines indicate electric field and broken lines magnetic field; circles indicate field perpendicular to plane of diagram, with solid circles indicating field pointing out of page and open circles indicating field pointing into page; circles accompanying solid lines indicate magnetic field, and those accompanying broken lines indicate electric field. (*a*) TE_{11} mode. (*b*) TE_{20} mode. (*c*) TM_{21} mode. (*After S. Ramo, J. R. Whinnery, and T. Van Duzer, Fields and Waves in Communication Electronics, 3d ed., John Wiley and Sons, 1994*)

guides with a nonmagnetic dielectric are given by Eqs. (9) and (10), where a is radius of the cylinder, p_{nl}

$$(f_c)_{\mathrm{TM}_{nl}} = \frac{p_{nl}c}{2\pi a\sqrt{\epsilon_r}} \tag{9}$$

$$(f_c)_{\mathrm{TE}_{nl}} = \frac{p'_{nl}c}{2\pi a\sqrt{\epsilon_r}} \tag{10}$$

represents the lth zero of the nth-order Bessel function, and p'_{nl} is the lth zero of its derivative. The first subscript in the designations TE_{nl} or TM_{nl} denotes the number of variations circumferentially, and the second subscript denotes the number of variations radially. The principal mode, or mode with longest cutoff wavelength (lowest cutoff frequency), is the TE_{11} mode, which, apart from the curvature required by the boundaries, is analogous to the principal TE_{10} mode of rectangular guides. The TE_{11} circular mode, however, has two polarizations with the same cutoff wavelength (called degenerate modes). The two may be excited in phase quadrature to produce fields with circular polarization, which may be desirable in certain rotating joints.

A second mode of special interest is the TE_{01} mode (or any TE_{0l} mode), in which the attenuation from conducting losses in the walls decreases with increasing frequency. It was consequently considered promising as a medium for long-distance, broadband communication, but the simpler optical fiber has proved the appropriate medium for these broadband links.

Other cross sections. An infinite number of other guide shapes is possible. A guide of elliptic cross section eliminates the degeneracy of modes in the circular guide, giving mode patterns somewhat between those of the circular and rectangular guides. A rectangular guide with a central ridge (called a ridge waveguide) gives a lower cutoff frequency than the rectangular guide with the same width, but also has lower power-handling capabilities because of the narrower gap in the region of highest field.

Circuit elements. A variety of circuit elements is required for exciting desired modes, filtering, coupling to passive and active elements, and other necessary networking functions. Excitation of a particular mode may be by probes, along the direction of the mode's electric field, by loops normal to magnetic field lines, or by the charge streams of an active vacuum-tube or semiconductor device placed within the guide.

A thin diaphragm introduced from the top of the guide (or the bottom, or both) may be represented as a capacitive shunt element in a transmission-line model. A diaphragm introduced from the side acts as an inductive shunt element. Circuit representations for a variety of probes, posts, irises, and other waveguide elements have also been analyzed and tabulated. Combinations of these elements may then be used to synthesize filters, reflectors, or matching elements in the guide.

Other important waveguide elements are the directional coupler, isolator, and magic-T. In the directional coupler, there is coupling to an auxiliary guide

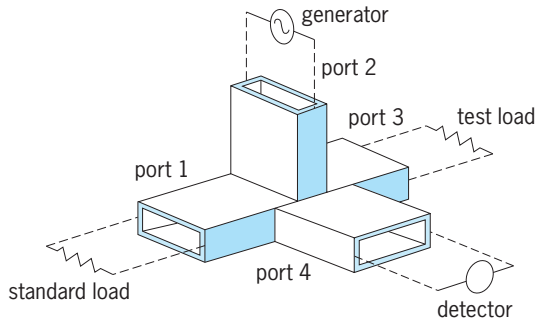

Fig. 3. Magic-T network in rectangular hollow-pipe waveguide used as a bridge. (*After S. Ramo, J. R. Whinnery, and T. Van Duzer, Fields and Waves in Communication Electronics, 3d ed., John Wiley and Sons, 1994*)

in such a way that the output of one of its ports is proportional to the wave traveling in the forward direction, and the output of the other is proportional to the reverse wave. The isolator makes use of the nonreciprocal properties of ferrites with an applied magnetic field to pass the forward-traveling wave of the guide but to eliminate the reflected wave. The magic-T (**Fig. 3**), and other forms of the microwave hybrid network, gives no output in port 4 when ports 1 and 3 are balanced, so it may be used in a bridge arrangement to compare a test load with a standard load. It is especially useful in the balanced mixers of heterodyne systems. E-plane T junctions, H-plane T junctions, and Y junctions are simple forms that are used in branching networks. *See* DIRECTIONAL COUPLER; GYRATOR.

Dielectric Waveguides

A dielectric waveguide consists of one dielectric material, called the core, surrounded by a different dielectric, called the cladding. The permittivity (dielectric constant), or refractive index, of the core is larger than that of the cladding, and under proper conditions electromagnetic energy is confined largely to the core through the phenomenon of total reflection at the boundary between the two dielectrics. *See* PERMITTIVITY; REFLECTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES.

Early dielectric guides were so lossy that they could be used only over short distances. Dielectric light pipes found surgical and laboratory use, and dielectric rods, called polyrod antennas, were used as radiators in certain World War II microwave radars. In 1969 silica fibers were developed with attenuations of 32 dB/mi (20 dB/km), low enough to be of use for optical communication applications. Since then, further improvements have reduced losses to as low as 0.3 dB/mi (0.2 dB/km). Fiber guides are now the basis for a worldwide optical communication network. *See* FIBER-OPTICS IMAGING.

Planar, rectangular, and thin-film forms of dielectric guides are also important in guiding optical energy from one device to another in optoelectronic and integrated optic devices, for example, from a semiconductor laser to an electrooptic modulator on a gallium arsenide substrate. Because of the simpler geometry, the planar forms will be used to explain

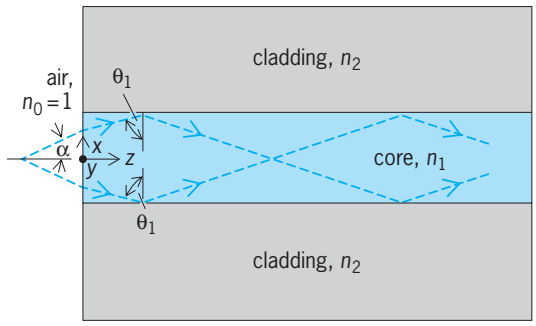

Fig. 4. Planar dielectric waveguide with ray paths that make up guided waves when angle θ_1 is greater than the critical angle for total reflection. Cladding refractive index n_2 is smaller than core refractive index n_1. Transverse direction (y direction) is perpendicular to the plane of the diagram.

the principle. *See* INTEGRATED OPTICS; LASER; OPTICAL MODULATORS.

Planar dielectric guides. The principle of dielectric guiding of electromagnetic energy can be illustrated in the symmetrical dielectric slab guide (**Fig. 4**), in which a slab of core material with refractive index n_1 is placed between two slabs of cladding with a refractive index n_2 that is smaller than n_1. Rays (normal to the wavefronts of plane waves) will be refracted in going from air to the core, and then reflected at the interface between core and cladding. If the angle of incidence θ_1 is greater than the critical angle for total reflection from this surface, all energy is reflected. The critical angle θ_c is given by Eq. (11). For angles

$$\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) \tag{11}$$

θ_1 greater than θ_c, the energy is thus trapped by the guiding system with no average power leaking out; for steeper angles ($\theta_1 < \theta_c$), there is some transmission to the outer medium on each reflection, representing a transmission loss to the propagating wave. Thus, the condition $\theta_1 > \theta_c$ is required for guiding without leakage or radiation loss.

In tracing the rays from the air at the input to the guide, there is some maximum angle α for rays to satisfy the above condition for guiding. The sine of this angle is the numerical aperture (NA) and is given by Eq. (12). Only rays within the wedge of angle $2\,\alpha_{\max}$

$$\mathrm{NA} = \sin\alpha_{\max} = \sqrt{n_1^2 - n_2^2} \tag{12}$$

satisfy the condition for total reflection. Because of the square-root relationship, even small differences between n_1 and n_2 yield reasonable values of the numerical aperture. For example, if $n_1 = 1.50$ and $n_2 = 1.49$, NA $= 0.173$ and $\alpha_{\max} \approx 10^0$.

The ray picture is clear and accurate, but the modes of this guiding system can also be found by obtaining appropriate solutions of Maxwell's equations. If variations in the transverse direction are negligible, these divide into transverse electric and transverse magnetic types, as in the hollow-pipe guides. The field does extend into the cladding dielectric (**Fig. 5**), but is evanescent (that is, it dies off exponentially with distance away from the interface). A parameter v is useful in telling which modes are guided and which radiate into the cladding. This parameter is defined by Eq. (13), where $2a$ is the height of the

$$v = \frac{2\pi a}{\lambda}\,\mathrm{NA} \tag{13}$$

slab, λ is the free-space wavelength, and the numerical aperture NA is defined by Eq. (12). The TE_0 mode is guided for all values of v; but higher-order modes, TE_m or TM_m, are guided only above some cutoff v defined by Eq. (14). To increase the v parameter of a

$$(v)_{\mathrm{cutoff}} = \frac{m\pi}{2} \tag{14}$$

mode so that it is above its cutoff value, it is possible to increase the slab height a, increase the frequency (decrease λ), or increase NA by making the difference between n_1 and n_2 greater.

Modes in the zigzag ray picture are defined as those rays for which the phase shift along the ray, including the phases of the reflections, is a multiple of 2π after a complete path of double reflection.

Rectangular and quasirectangular guides. For applications to integrated optics, in which the dielectric guide is used to direct optical energy from one component to another, it is necessary to have some lateral confinement as well as the vertical confinement provided by the layers of the slab guide (Fig. 4). The lateral confinement is provided by surrounding the guiding core with a dielectric of lower refractive index on the sides as well as on top and bottom, or processes such as ion implantation or diffusion may produce an inhomogeneous region of higher index than the surroundings to effect the confining region. *See* ION IMPLANTATION.

Optical fibers. By far the most important dielectric guide at present is the optical fiber used for optical communications. Here the round core is surrounded by a cladding of slightly lower refractive index. The combination is surrounded by a protective jacket to prevent corrosion and give added strength, but this jacket plays no role in the optical guiding. The principle of guiding is the same as that described for the slab guide. From the ray point of view, rays are totally reflected from the boundary between n_1 and n_2

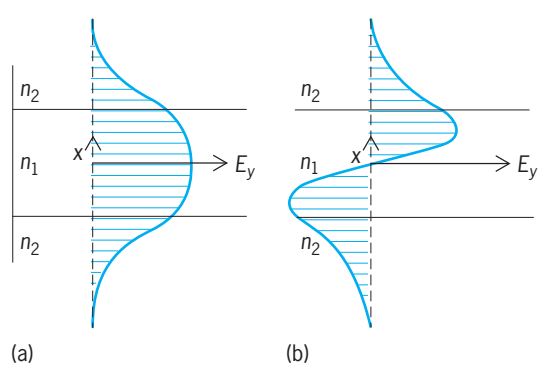

Fig. 5. Distributions of electric field E_y versus coordinate distance normal to planar dielectric guide axis for (*a*) lowest-order ($m = 0$) and (*b*) first-order ($m = 1$) transverse electric modes.

for angles flatter than the critical angle defined by Eq. (11). Numerical aperture (NA) is here also a useful concept and is given by Eq. (12).

Solutions of Maxwell's equations in circular cylindrical coordinates yield the modes of this guiding system. These divide into transverse electric and transverse magnetic classes only for modes that have no circumferential variation. Other modes have all field components and are called hybrid modes, HE_{mn} or EH_{mn}, depending respectively upon whether energy is predominantly magnetic or electric. The principal mode is the HE_{11}, which has no cutoff frequency. The cutoff frequency of other modes is determined by the v parameter of Eq. (13), with a interpreted as the radius of the core.

Losses. Fibers used in optical communications are almost universally of silica (silicon dioxide), possibly with germanium or other ions as dopants. The very low attenuations achieved came about by removal of undesired impurities. The lowest values of attenuation are about 0.3 dB/mi (0.2 dB/km), occurring for a wavelength near 1.5 micrometer. This is near the theoretical limit for Rayleigh scattering. Another local minimum of about 0.8 dB/mi (0.5 dB/km) occurs for a wavelength near 1.3 μm, which is a wavelength of lower dispersion. Attenuation increases rapidly for shorter wavelengths because of the wavelength dependence of Rayleigh scattering, and also increases at wavelengths longer than 1.6 μm because of the tail of an infrared absorption band. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; SCATTERING OF ELECTROMAGNETIC RADIATION.

Polymeric or plastic optical fibers have larger losses than silica but have found use in short-distance data-link applications. There is also work on a number of materials with losses lower than silica. For example, the fluorozirconate glasses promise very low losses in the mid-infrared region of the spectrum. *See* FIBER-OPTIC CIRCUIT; OPTICAL MATERIALS.

Radiation losses may occur even in the guided regime if there are excessive bends or corrugations in the fiber. Thus it is important to design the fiber and its cabling to avoid such perturbations.

Dispersion. All dielectric guides are dispersive in that different frequency components of the signal travel at different velocities. The variation of group velocity with frequency is the relevant quantity for calculating the distortion of the modulation envelope of an analog signal, or the pulse shape of a digital signal. This quantity is called the group-velocity dispersion. It may arise in part from the variation of refractive index with frequency (called material dispersion), and in part from the basic waveguide characteristics themselves (called waveguide dispersion). For a single-mode fiber (one with the v parameter below 2.4, so that only the HE_{11} mode is guided), material dispersion is the larger component except near a wavelength of 1.3 μm, at which point the combination becomes zero. With special dopings, or multiple claddings, this zero dispersion point can be shifted to the region of minimum attenuation, around 1.5 μm.

For the single-mode fiber, the spread in arrival time for a narrow input pulse is given by Eq. (15), where L

$$\Delta\tau = DL\,\Delta\lambda \qquad \text{picoseconds} \tag{15}$$

is the fiber length in kilometers, $\Delta\lambda$ the wavelength spread (from source, signal, or both) in nanometers, and D a dispersion quantity given in picoseconds per kilometer of length and nanometer of wavelength spread.

For fibers whose v parameter is greater than 2.4, more than one mode is guided, and these modes travel with different group velocities even at a single wavelength. This causes a spread in arrival time (called intermode distortion or intermode dispersion). Since modes near cutoff travel with velocities near that of the cladding, and modes well above cutoff near that of the core, this delay for length L is given approximately by Eq. (16), where c is the velocity of light in vacuum.

$$(\Delta\tau)_{\text{multimode}} \approx \frac{(n_1 - n_2)L}{c} \tag{16}$$

Graded-index fibers. The fibers discussed so far have a sharp discontinuity of refractive index between core and cladding and are called step-index fibers. In graded-index fibers the refractive index varies within the core, decreasing with increasing radius from a maximum at the axis (**Fig. 6**). Multimode graded-index fibers have appreciably less intermode distortion than multimode step-index fibers. Index grading can also play a role in shifting the zero dispersion wavelength of single-mode fibers.

Nonlinear effects. The refractive index of silica and other glasses depends to a slight degree on the intensity of the optical signal. Although this is a small effect, it can add up over the length of a fiber to produce both undesirable and desirable effects. Intermodulation among separate signals is the most important undesirable effect. Among the useful phenomena are pulse compression and amplification using the Raman effect. Soliton propagation, in which self-phase modulation from the nonlinear

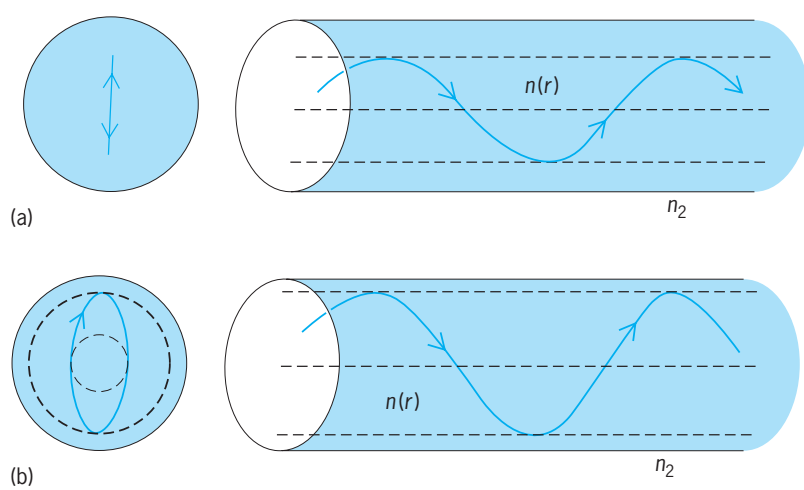

Fig. 6. Ray paths within the core of a graded-index fiber. Both end-on projection and oblique view of each path are shown. (*a*) A path through the meridian. (*b*) A skew path.

effect just compensates for group-velocity dispersion so that a short pulse propagates long distances without spreading, is especially interesting and has potential for high-data-rate communication systems. *See* NONLINEAR OPTICS; RAMAN EFFECT; SOLITON.

Fiber amplifiers. The capabilities of fiber for information transmission was greatly improved by the development of fiber amplifiers. Most important has been the silica fiber doped with erbium, providing amplification by a laser-type mechanism at a wavelength of 1.54 μm. This is the wavelength of minimum attenuation for silica fibers. Pumping of the erbium laser line is provided by semiconductor laser diodes at wavelengths of 0.98 or 1.43 μm. These amplifiers have greatly extended the transmission distance of point-to-point fiber lines and have permitted more branches in local-area uses of optical fibers. *See* OPTICAL COMMUNICATIONS; OPTICAL FIBERS.

John R. Whinnery

Bibliography. F. W. France, *Optical Fibre Lasers and Amplifiers*, 1991; H. A. Haus and J. R. Melcher, *Electromagnetic Fields and Energy*, 1989; D. Marcuse, *Theory of Dielectric Optical Waveguides*, 2d ed., 1991; N. Marcuvitz et al. (eds.), *Waveguide Handbook*, 1986; A. D. Olver, *Microwave and Optical Transmission*, 1992; S. Ramo, J. R. Whinnery, and T. VanDuzer, *Fields and Waves in Communication Electronics*, 3d ed., 1994; A. W. Snyder and J. D. Love, *Optical Waveguide Theory*, 1983.

Wavelength

The distance between two points on a wave which have the same value and the same rate of change of the value of a parameter, for example, electric intensity, characterizing the wave. The wavelength, usually designated by the Greek letter λ, is equal to the speed of propagation c of the wave divided by the frequency of vibration f; that is, $\lambda = c/f$ (see **illus.**). *See* WAVE (PHYSICS).

The wavelength for a sound of a given frequency varies greatly, depending upon the speed of propagation in the medium in which the sound is moving. For example, a sound wave having a frequency of 1000 Hz would have a wavelength of approximately 1 ft (0.3 m) in air, 412 ft (1.3 m) in water, and 17 ft (5.1 m) in steel. The wavelength of electromagnetic waves depends on the velocity of light in the material in which the waves are traveling. *See* WAVE MOTION.

William J. Galloway

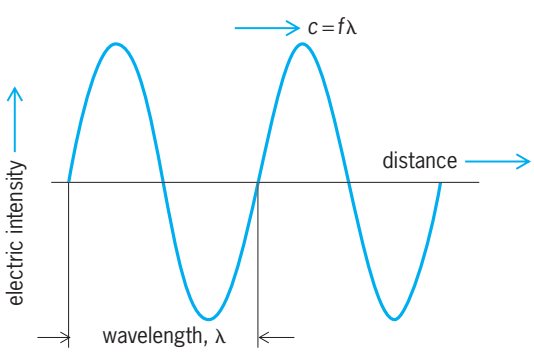

Wavelength λ and related quantities.

Wavelength measurement

Determination of the distance between successive wavefronts of equal phase of a wave. This article discusses wavelength measurement of electromagnetic waves in the radio, microwave, infrared, and optical regions. *See* ELECTROMAGNETIC RADIATION; WAVE MOTION; WAVELENGTH.

Wavelength by frequency measurement. From the relation $\lambda = c/f$ between wavelength λ, speed c, and frequency f, the wavelength of a wave motion can be calculated if the speed is known and its frequency is measured. The ease and accuracy of electronic counting and timing make frequency measurement the most precise of all physical measurements. This method of wavelength determination is thus one of the most accurate, but only if the speed (phase velocity) is known. In free space the speed of an electromagnetic wave c_0 is, through the 1983 definition of the meter, fixed at exactly 299,792,458 m/s ($\approx$186282.397 mi/s), or roughly 300,000 km/s. Unless otherwise specified, it is general practice to quote the wavelength of an electromagnetic wave as the free-space value λ_0, given by Eq. (1).

$$\lambda_0 = \frac{c_0}{f} \tag{1}$$

See FREQUENCY MEASUREMENT; PHASE VELOCITY.

Radio and microwave regions. The presence of any dielectric material (such as air) or any magnetic matter with a permeability greater then unity will cause the wave to travel at a velocity lower than its free-space value. The speed is also altered if the waves pass through an aperture, are focused by a lens or mirror, or are constrained by a waveguide or transmission line. In such cases it may be more appropriate to measure the wavelength directly. In the pioneering experiments on radio waves, it was found that standing waves existed in space whenever reflections occurred and that these provided a convenient means of measuring the wavelength. It thus became the convention to characterize waves by their wavelength rather than by their frequency, as is now more commonly the case. Specifying the frequency is preferred because, unlike the wavelength, it is independent of the speed of propagation and does not change as the wave moves from one medium to another. *See* ELECTROMAGNETIC WAVE TRANSMISSION; TRANSMISSION LINES; WAVEGUIDE.

Interferometer methods. In the microwave region the wavelengths are sufficiently short that it is convenient to measure them by using interferometer techniques directly analogous to those used with light.

In a typical interferometer used in the millimeter wavelength range (**Fig. 1**), a microwave beam is directed at a beamsplitter, which splits the beam into two parts, A and B, by partial reflection. The A beam

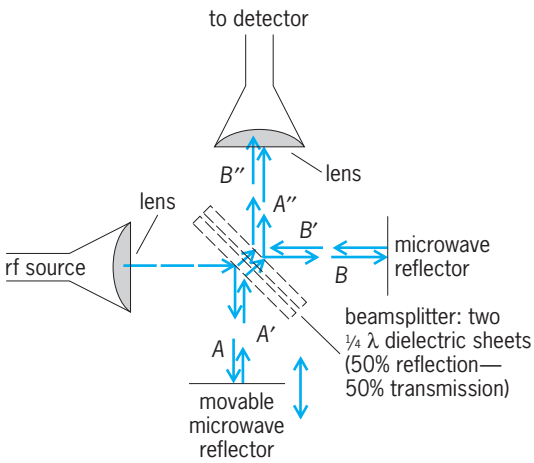

Fig. 1. Wavelength measurement by the Michelson interferometer used at millimeter wavelengths.

is reflected to a movable reflector and reflected again as A'. The beamsplitter transmits part of this as A''. The transmitted part B of the original beam is reflected by a fixed microwave reflector as B'. This is partially reflected by the beamsplitter as B''. The beams A'' and B'' combine to form standing waves, which are then detected. Movement of the movable reflector causes the position of the standing wave to move, which causes the detected signal to pass through successive points of maximum and minimum amplitude. The distance between points of successive maxima or minima is one-half wavelength. This distance may be determined from the motion of the movable reflector. *See* INTERFEROMETRY.

Standing-wave methods. A crude form of interferometer, used by early radio experimenters, does not require a beamsplitter. Instead it relies upon the standing-wave formed when a wave is reflected onto itself, as would occur, for example, with beams B and B' in the interferometer discussed above (Fig. 1). The maxima and minima of this standing wave can be observed by moving a suitable detector to various points in the field. Under ideal conditions the distance between successive minima or maxima is half the wavelength and can be measured with a measuring tape; but the detector and stray reflections can easily distort the observed pattern.

Lecher wires. In the Lecher-wire wavemeter the wave is made to travel down a transmission line. With this simple device, wavelength is measured by sliding a short circuit along the line and observing the cyclic variations of a power indicator. The distance between two successive absorption maxima or minima is half the wavelength ($\lambda/2$); thus by a simple length measurement the wavelength is measured directly.

Microwave wavemeters. These wavemeters make use of resonant coaxial-line sections or cavities as tuned elements. The two general types of microwave wavemeters are the absorption, or reaction, type and the transmission type. Wavemeters of low or medium selectivity are frequently used as coarse measuring devices to establish the general range of frequency of operation of a system before applying more refined and complex methods for accurate frequency checking. *See* CAVITY RESONATOR; COAXIAL CABLE.

The dimensions of the cavity determine the resonant frequency of a resonant-cavity microwave wavemeter (**Fig. 2**). A signal is fed in from either a coaxial line or waveguide, and energy is fed out to a suitable detector by a second coaxial line. The cavity is tuned by means of a micrometer-driven plunger, which may be calibrated in terms of wavelength. *See* MICROWAVE.

Tuned circuits. For wavelengths greater than a few meters, the dimensions of transmission lines and resonant cavities become inconveniently large. It is then more convenient to use resonant systems made from inductors and capacitors. With calibrated values of inductance L and capacitance C, it is possible to provide a scale calibrated in wavelength or frequency. For low-loss circuits the resonant frequency is $1/(2\pi\sqrt{LC})$. Absorption-type wavemeters (**Fig. 3**), often constructed with the principal inductance as a plug-in coil, are used for frequency or wavelength measurement up to frequencies of approximately 1000 MHz. *See* RESONANCE (ALTERNATING-CURRENT CIRCUITS).

Infrared and optical regions. The free-space wavelengths of monochromatic visible and infrared radiations can be derived directly from Eq. (1) if their corresponding frequencies are known. Air wavelengths can be determined by dividing the vacuum wavelength by the refractive index of air. The optical frequencies of a number of standards have been precisely determined by measurement of their frequency referenced to the cesium-133 primary frequency standard, using femtosecond combs. The most accurate standards published to date are based on transitions in single trapped ions and have an uncertainty within a factor of 3 or better of the cesium standard. These optical standards are used for the realization of the unit of length, but are also expected to become secondary representations of the second. *See* LENGTH; LIGHT; PHYSICAL MEASUREMENT; WAVELENGTH STANDARDS; FREQUENCY MEASUREMENT.

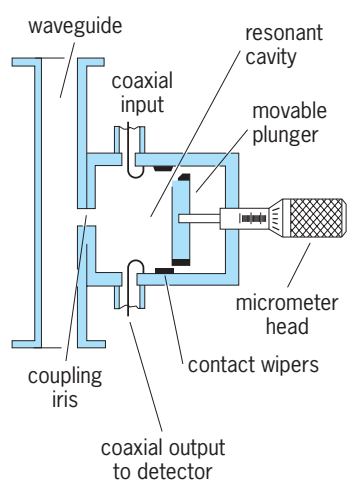

Fig. 2. Typical construction of a resonant-cavity wavemeter providing either coaxial or waveguide inputs.

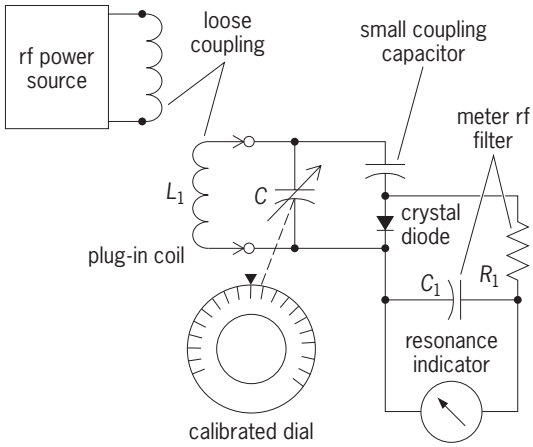

Fig. 3. Schematic diagram of inductance-capacitance type of absorption wavemeter (for frequencies between approximately 50 kHz and 1000 MHz).

Dispersion methods. Wavelength values to an accuracy of 1 part in 10^5 can be determined with a spectrometer, spectrograph, or monochromator, in which a prism or diffraction grating is used as a dispersive element. Each wavelength forms a line image of the entrance slit at a particular angle. An unknown wavelength can be determined by interpolation with the pattern formed by a lamp emitting the tabulated characteristic wavelengths of a particular element. Atomic reference data for this purpose are published, for example, on the Internet (see, for example, http://physics.nist.gov/PhysRefData/ASD/index.html). Care must be taken in using tabulated wavelength values because the values may be either for vacuum or for the refractive index of standard air. (Air wavelengths are smaller than vacuum values by about 1 part in 3000.) *See* DIFFRACTION GRATING; OPTICAL PRISM; SPECTROSCOPY.

Use of interferometers. The most precise wavelength measurements use an interferometer to compare the unknown wavelength λ_1 with a standard wavelength λ_2. Usually either the two-beam Michelson form or the multiple-beam Fabry-Perot form of interferometer is used. The general equation, applicable to both forms, is (2), in which θ is the angle incidence, t is

$$(m_1 + f_1)\lambda_1 = (m_2 + f_2)\lambda_2 = 2nt\cos\theta \qquad (2)$$

the real or virtual separation of the reflectors, and n is the refractive index of the medium between the reflectors. Thus at any arbitrary reference point in the interference pattern, Eq. (3) holds, where m_1 and

$$\lambda_1 = \frac{\lambda_2(m_2 + f_2)}{m_1 + f_1}$$

m_2 are integers, usually called the orders of interference, and f_1 and f_2 are fractions. Maximum transmission of light of wavelength λ occurs when its corresponding fractional order f is zero. Interferometer measurements to an accuracy of a few megahertz are necessary to determine the mode order number of femtosecond comb systems used to measure optical frequencies.

Several forms of interferometric "wavemeter" are commercially available for use with laser radiations. One of their principal uses is to measure the wavelength emitted by a tunable laser, for example a titanium-sapphire or diode laser, to better than 1 part in 10^6, so that it may be tuned into coincidence with a desired spectral transition. One form of wavemeter has a retroreflector that moves to and fro along a track. This reflector forms part of a two-beam interferometer so that sinusoidal intensity signals are generated. The signals for the unknown and standard radiations are counted electronically, giving totals that correspond to the order numbers m_1 and m_2. Another form of laser wavemeter, suitable for use with pulsed as well as continuous lasers, uses the wedged Fizeau interferometer with a detector array to measure the spacing and position of the interference pattern. *See* LASER; LASER SPECTROSCOPY.

Fourier transform method. When a number of wavelengths are mixed together in the input to a moving-carriage two-beam interferometer, the output signal is the summation of the many separate sine-wave signals having different periods. A Fourier analysis of this composite signal enables the separate wavelengths to be identified. This Fourier transform method is particularly useful for the measurement of complex spectra in the infrared. *See* FOURIER SERIES AND TRANSFORMS; INFRARED SPECTROSCOPY.

W. R. C. Rowley; Geoffrey P. Barwood

Bibliography. G. H. Bryant, *Principles of Microwave Measurements*, Institution of Electrical Engineers, 1993; P. Hariharan, *Optical Interferometry*, Academic Press, 2d ed., 2003; J. M. Hollas, *Modern Spectroscopy*, 4th ed., Wiley, 2004; T. S. Laverghetta, *Modern Microwave Measurements and Techniques*, rev. ed., Artech House, 1989; P. J. Mohr and B. N. Taylor, CODATA recommended values of the fundamental physical constants, *Rev. Mod. Phys.*, 77:1–107, 2005; T. J. Quinn, *Mise en pratique* of the definition of the metre (2001), *Metrologia*, 40:103–133, 2003; B. N. Taylor and W. D. Phillips, *Precision Measurements and Fundamental Constants II*, Nat. Bur. Stand. Spec. Publ. 617, 1984.

Wavelets

The elementary building blocks in a mathematical tool for analyzing functions. The functions to be analyzed can be very diverse; examples are solutions of a differential equation, and one- and two-dimensional signals. The tool itself, the wavelet transform, is the result of a synthesis of ideas from many different fields, ranging from pure mathematics to quantum physics and electrical engineering.

In many practical applications, it is desirable to extract frequency information from a signal— in particular, which frequencies are present and their respective importance. An example is the decomposition into spectral lines in spectroscopy. The tool that is generally used to achieve this is the Fourier transform. Many applications, however, concern so-called nonstationary signals, in which the makeup

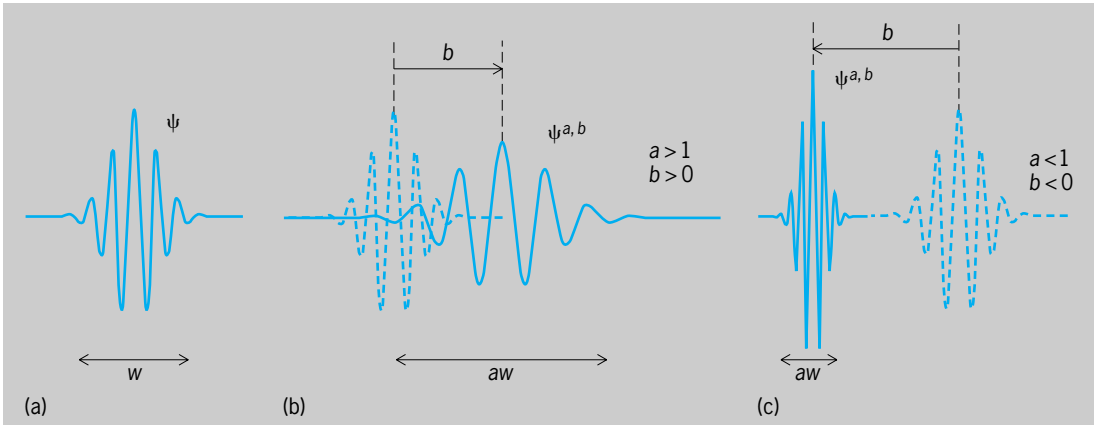

Fig. 1. Members of a wavelet family: an original and two shifted and dilated copies. (*a*) Original wavelet ψ with width *w*. (*b*) Large-scale wavelet $\psi^{a,b}$, with width *aw* that is larger than the original width *w*. (*c*) Fine-scale wavelet $\psi^{a,b}$ with narrow width.

of the different frequency components is constantly shifting. An example is music, where this shifting nature has been recognized for centuries by the standard notation, which tells a musician which note (frequency information) to play when and how long (time information). For signals of this nature, a time-frequency representation is needed. *See* FOURIER SERIES AND TRANSFORMS.

There exist many different mathematical tools leading to a time-frequency representation of a given signal $f(t)$, each with its own strengths and weaknesses. The wavelet transform is such a time-frequency analysis tool. Its particular strength lies in its ability to deal well with transient high-frequency phenomena, such as sudden peaks or discontinuities, as well as with the smoother portions of the signal. (An example is a crack in the sound from a damaged record, or the attack at the start of a music note.) The wavelet transform is less well adapted to harmonically oscillating parts in the signal, for which Fourier-type methods are more indicated.

Wavelet families. A family of wavelets is derived from one function by dilation and translation. The original function ψ is typically smooth and well concentrated; the whole family is then given by Eq. (1),

$$\psi^{a,b}(x) = |a|^{-1/2}\psi\left(\frac{x-b}{a}\right) \tag{1}$$

where a and b are real numbers and a is greater than 0. If the original wavelet ψ is centered on 0 and has width w, then $\psi^{a,b}$ is centered on b and has width aw (**Fig. 1**). The parameter b indicates the time location (if the variable x stands for time), while a gives the scale. The scale parameter is related to frequency: If the original wavelet has approximately k oscillations over its width w, corresponding (coarsely speaking) to a frequency $\omega_0 = k/w$, then $\psi^{a,b}$ will have the same number of oscillations for a width aw, corresponding to frequency $k/aw = \omega_0/a$. Large a corresponds therefore to long wavelets with low frequency, and very small a to narrow, high-frequency wavelets. In practice the range of scales can bridge several orders of magnitude.

Signal analysis. Once a family of wavelets is given, it can be used to analyze signals. Wavelets can be used to characterize or analyze a signal $f(x)$ in two different ways. The first is to correlate the signal f with all the wavelets in the family, that is, to compute the wavelet transform, given by Eq. (2). The second is

$$(Wf)(a,b) = \int f(x)\psi^{a,b}(x)\,dx \tag{2}$$

to try to write f as a superposition of wavelets, with weights which are called the wavelet coefficients. In the case where both the scale and translation parameters are allowed to vary continuously, the superposition is an integral, given by Eq. (3). If the wavelet

$$f(x) = \int_0^\infty \int_{-\infty}^\infty (wf)(a,b)\psi^{a,b}(x)\frac{da\,db}{a^2} \tag{3}$$

ψ is a real function with good localization and zero integral, then the weights in the superposition can be chosen to satisfy Eq. (4), where C_ψ is a constant

$$(wf)(a,b) = C_\psi(Wf)(a,b) \tag{4}$$

depending only on ψ (not on f). This shows that the two approaches, characterizing a function f by means of its inner products with $\psi^{a,b}$, as in Eq. (2), or decomposing it, as in Eq. (3), into a superposition of $\psi^{a,b}$, amount to the same thing.

Discrete parameters. In Eqs. (2) and (3), the scale and localization parameters a and b vary continuously; this is the continuous wavelet transform, a useful analysis tool, but very redundant. In another

Fig. 2. Two examples of wavelets which generate an orthonormal basis, and which are associated with fast algorithms.

version of the wavelet transform, a and b are restricted to a discrete lattice of values, such as $a = 2^j$, $b = 2^j k$, where j and k are integers. In this case the superposition integral in Eq. (3) is replaced by the sum in Eq. (5), where the factors on the right-hand side are given by Eqs. (6). In order for Eq. (5) to hold,

$$f(x) = \sum_{j,k} \langle f, \psi_{j,k} \rangle \psi_{j,k}(x) \tag{5}$$

$$\psi_{j,k}(x) = 2^{-j/2} \psi(2^{-j}x - k) \tag{6a}$$

$$\langle f, \psi_{j,k} \rangle = \int f(x) \psi_{j,k}(x)\, dx \tag{6b}$$

many more restrictions on ψ have to be imposed than in the continuous case. However, the wavelets $\psi_{j,k}$ are orthonormal, and there is no redundancy in the reproduction of f by the sequence $\langle f, \psi_{j,k} \rangle$ [**Fig. 2**]. *See* NONRELATIVISTIC QUANTUM THEORY.

Data compression. Smooth orthonormal wavelet bases are interesting because their smoothness makes them suitable for the decomposition of smooth functions, whereas the different scales in the wavelet family make it possible to fine-tune the decomposition locally near singularities or transients. The combination of smoothness and scaling in orthonormal wavelet bases implies that for many functions or signals f of interest, large numbers of fine-scale coefficients are negligibly small, because they are localized in smooth, uneventful parts of f, which are already well captured by the wider-scale wavelets. Because these tiny coefficients can be neglected without appreciable distortion in the reconstruction of f, data compression is possible. *See* DATA COMPRESSION.

Discrete variables. In the formulas above, x is a continuous variable. Similar formulas can be written if x varies discretely, on a lattice; the integrals then become sums. Most practical implementations use such a discrete version. In this case the orthonormal wavelet transform (and its generalizations) can be implemented via fast algorithms [with complexity of order $O(N)$ for data with N samples] that are related to subband filtering in electrical engineering.

Applications. Applications of wavelets include various forms of data compression (such as for images and fingerprints), data analysis (nuclear magnetic resonance, radar, seismograms, and sound), and numerical analysis (fast solvers for partial differential equations). *See* DIFFERENTIAL EQUATION; INTEGRAL TRANSFORM; NUMERICAL ANALYSIS. Ingrid Daubechies

Bibliography. I. Daubechies, *Ten Lectures on Wavelets*, 1992; Y. Meyer, *Ondelettes*, 1990, transl. as *Wavelets*, 1993; Y. Meyer, *Wavelets: Algorithms and Applications*, 1993.

Wavellite

A hydrated phosphate of aluminum mineral with composition $Al_3(OH)_3(PO_4)_2 \cdot 5H_2O$, in which small amounts of fluorine and iron may substitute for the hydroxyl group (OH) and aluminum (Al), respectively. Wavellite crystallizes in the orthorhombic system. The crystals are stout to long prismatic, but are rare. Wavellite commonly occurs as globular aggregates of fibrous structure (see **illus.**) and as encrusting and stalactitic masses. Wavellite crystals range in color from colorless and white to different shades of blue, green, yellow, brown, and black.

Wavellite in globular aggregates, found in Devonshire, England. (*Specimen from Department of Geology, Bryn Mawr College*)

Wavellite is a widespread secondary mineral occurring in small amounts in crevices of low-grade metamorphic aluminous rocks, in limonite, and in phosphate rock deposits. Found at many places in Europe and North America, it is also abundant in tin veins at Llallagua, Bolivia. *See* PHOSPHATE MINERALS.

Wayne R. Lowell

Wax, animal and vegetable

Any of the substances containing esters of higher fatty acids and long-chain monohydric alcohols. From a practical standpoint, this definition is inadequate because it allows liquids such as sperm whale oil and jojoba oil to be called waxes, and it fails to indicate the complexity of waxes. While waxes do contain wax esters, they are seldom if ever pure. They are usually mixtures that may contain high-molecular-weight acids, alcohols, esters, ketones, hydrocarbons, sterols, diesters, hydroxyacids, and so forth, as well as the wax esters. *See* ESTER.

Properties. Practical wax formulators use physical, rather than chemical, properties to define waxes. Waxes must: be a solid at 20°C (68°F); be crystalline; melt above 40°C (104°F) without decomposition,

Physical properties of selected waxes

Substance	Melting point, °C (°F)	Acid value*	Saponification value†	Iodine value‡	Acetyl value§
Beeswax	63 (145)	19	92	10	15
Candelilla	67 (153)	15	55	33	
Carnauba	84 (183)	5	83	10	55
Spermaceti	45 (113)	1	120	2	
Wool wax	40 (104)	20	100	30	

*mg KOH/g of wax to neutralize free acid.
†mg KOH/g of wax to saponify esters.
‡cg I_2/g of wax absorbed; indicates amount of unsaturation.
§mg KOH to neutralize acetic acid from saponifying 1 g acetylated wax; indicates amount of free hydroxyl.

have relatively low viscosity above the melting point, have consistency and solubility properties that are strongly dependent upon temperature; and be capable of being polished under slight pressure. Typical properties of several waxes are shown in the **table**.

Sources. Wax sources are abundant; therefore it is the quantity and ease of recovery that generally determine whether a wax becomes commercially valuable. Waxes have been isolated from the outer layers of bacteria, the roots, stems, leaves, fruit, and flowers of plants, the exudates of insects, the skin and hair of some animals, and the bodies of certain marine and land animals.

Important waxes. Carnauba wax is extracted from an exudate on the leaves of the carnauba palm (*Copernicia prunifera*). The only significant production is in northeastern Brazil. Leaves are cut from the tree and allowed to dry. During drying the leaf shrinks, and the wax is no longer adherent. Machines beat the leaves to free the wax, and so produce a mixture of crude wax and leaf particles which is melted and stained to isolate the crude product. Carnauba is valuable because it has a relatively high melting point. It buffs to a high luster without showing the polish lines. *See* CARNAUBA WAX.

Candelilla wax is obtained from a coating on the stem of *Euphorbia antisyphilitica*, a leafless desert shrub. It is fairly common in the Big Bend country of Texas, but the commercial product comes from the states of Chihuahua and Coahuila in Mexico. The shrubs are pulled up when they are about 1 m tall, bundled, and heated in water. The wax rises to the surface, and is skimmed off. Candelilla wax contains a tacky resin, as well as large amounts of hydrocarbon. It has strong adhesive properties, and the resin imparts antislip characteristics to formulated waxes.

Beeswax is an exudate of the honeybee. Most of the commercial beeswax comes from Africa. Hollow logs hung from trees are used by wild bees as hives. When the logs are presumed to be full, they are removed, and the wax is recovered by melting. Beeswax is an important ingredient in many cosmetic formulations. Considerable quantities are used in candles, with lesser amounts going into polishes, modeling, and pattern making. Beeswax is a good solvent for other waxes; it is a plasticizer, and promotes adhesion. *See* BEEKEEPING.

Wool wax, obtained from sheep wool, is refined to produce lanolin, which is very hydrophilic compared to other waxes. It is widely used in cosmetics and pharmaceutical creams and salves. Its property of softening dry skin without affecting natural moisture transpiration is probably due to the large amounts of sterols and alcohols present. *See* LANOLIN.

Other waxes. Many other waxes have been of commercial importance from time to time, but their availability or cost has forced them from the market. Sperm oil and spermaceti are extracted from the blubber and head cavities of the sperm whale (*Physeter macrocephalus*). Spermaceti has properties that are useful in cosmetics, but since the sperm whale is an endangered species, it cannot be hunted. Instead, the spermaceti market is satisfied by esters derived from fatty acids and synthetic fatty alcohols. Ouricouri wax is similar to carnauba. It is a coating on palm leaves, but unlike carnauba the wax must be scraped since it does not flake free from the leaf. The decrease in sales of ouricouri is probably due to high labor costs. Retamo wax is similar to candelilla. It comes from a desert shrub in northwestern Argentina. Government controls raised the price so high that it is no longer on the market. Sugarcane wax is extracted from the filtrate in refining cane sugar.

Analysis. Until the advent of gas-liquid chromatography, it was virtually impossible to analyze waxes, but the important characteristics of a wax are still based on physical properties or empirical chemical reactions. *See* FAT AND OIL.

Howard M. Hickman

Wax, petroleum

A substance produced primarily from the dewaxing of lubricating-oil fractions of petroleum. It may be of either the crystalline or microcrystalline type. The crystalline wax is produced from distillate lubricating fractions, whereas the microcrystalline wax is obtained from the residual lubricating fractions of the crude oil. The melting-point range for refined crystalline waxes is 120–150°F (48–65°C) while the petrolatum or microcrystalline waxes have melting points in the range of 150–175°F (65–79°C).

Petroleum waxes constitute approximately 90% of all wax used in industry today. The other 10% comprises vegetable and animal waxes used primarily as specialty waxes. *See* WAX, ANIMAL AND VEGETABLE.

Petroleum wax and petrolatum production in the United States has varied depending upon the economic demand.

Uses. Petroleum wax has a wide variety of uses. It is used to coat paper products, to blend with other waxes for the manufacture of candles, in the manufacture of electrical equipment and many polishes for home and industry, and as a source material for oxidized products which are being more widely used in industry. The softer waxes, such as petroleum jelly, after proper purification, are being used as medicinal products.

The petrolata have many uses, principally as components of blended waxes used for coating paper, electrical insulation, and water- and rustproofing materials, as components of crayons and printing inks, and in the manufacture of specialties. *See* PETROLATUM.

In addition to melting point, other properties are important in establishing the quality and end use of waxes. These are tensile strength, hardness, stability, odor and taste, oil content, blocking and gloss characteristics, scuff resistance, and sealing strength.

Purification of crude wax. The slack wax obtained from the primary dewaxing operation on the lubricating fraction must be purified further in most cases before marketing. The first operation in this purification is the removal of all or most of the oil associated with the wax. The deoiling operation on the wax is conducted by either the sweating operation or the solvent deoiling operation.

The old and widely used sweating process is, in reality, a fractional separation of crude wax (slack wax) obtained from a dewaxing process. The separation by melting points is accomplished in large pans or sweaters contained in a well-insulated building which can be heated uniformly at a slow rate to effect the melting-point separations. To accomplish this, molten slack wax is placed in the sweater pans and is cooled until a solid cake is formed. The sweater pans are then gradually warmed and the waxes are separated—the lowest-melting-point fractions being separated as liquid fractions or cuts. As the temperature of the sweater is increased, the higher-melting-point waxes are removed. The cuts or fractions are then blended to provide the desired range of melting point. The melting points of the fractions vary from 110 to approximately 160°F (43 to 71°C). These fractions or blends are usually treated with sulfuric acid followed by clay filtration and deodorizing for the removal of impurities. The finished wax is molded into various shapes for commercial sale and distribution to industry. The commercial grades of paraffin wax are known as crude scale and fully refined waxes, and they vary in melting-point range.

The solvent deoiling and fractionation process is a relatively new method of separating wax into its various melting-point fractions. It rapidly is replacing the sweating process because of its flexibility and wider range of application. This process is quite similar to the solvent dewaxing process. The slack wax to be deoiled or fractionated is dissolved in a suitable solvent and the mixture chilled to precipitate the desired melting-point grade of wax. The wax is separated by filtration from the solvent solution which retains the more soluble waxes and oil. Further chilling and filtration steps separate additional wax fractions. Additional treatment with acid and clay results in high-quality waxes.

The petrolatum type of wax cannot be made by the sweating processes. These waxes are made directly from residual lubricating oil from which they are separated either by the solvent dewaxing process or by the centrifuge dewaxing method. Fractional separation of petrolatum into various melting points can be accomplished by the solvent deoiling process. One of the first methods of obtaining petrolatum was through the recovery of the wax bottoms that had settled out of wax-bearing crude oils during storage. *See* DEWAXING OF PETROLEUM; PETROLEUM PROCESSING AND REFINING; PETROLEUM PRODUCTS.

Wayne E. Kuhn

Weak nuclear interactions

Fundamental interactions of nature that play a significant role in elementary-particle and nuclear physics, and are distinguished from other such interactions by special properties such as participation of all the fundamental fermions and failure to conserve parity and to respect particle–antiparticle symmetry. According to present understanding, the four fundamental forces of nature are the gravitational, the electromagnetic, the strong, and the weak. The weak force has short range (less than 10^{-17} m) and at low energies is very feeble compared to the strong and electromagnetic forces, but it can be distinguished from the latter by its special character. For example, all of matter (with the possible exception of dark matter) appears to consist of certain basic constituents, the quarks and leptons, collectively called the fundamental fermions (**Table 1**). While only quarks participate in strong interactions, and only the quarks and the charged leptons ($e^{\pm}$, $\mu^{\pm}$, $\tau^{\pm}$) participate in electromagnetic interactions, all of the fundamental fermions, including neutrinos, engage in weak interactions. Moreover, while the strong and electromagnetic interactions respect spatial inversion symme-

TABLE 1. Fundamental fermions*

Leptons Electric charge	Family 1	Family 2	Family 3
0	ν_e	ν_μ	ν_τ
$-e$	e^-	μ^-	τ^-
Quarks† Electric charge	Family 1	Family 2	Family 3
$2e/3$	u	c	t
$-e/3$	d	s	b

*Corresponding to each fermion, there is an antifermion with the same mass and opposite electric charge.
†Quarks are confined in mesons and baryons and do not appear as free particles.

TABLE 2. Classification of representative weak interactions*

Interaction	Charged	Neutral
Leptonic		
	$\mu \to e\nu\nu$	$\nu_\mu e \to \nu_\mu e$
	$\nu_e e \to \nu_e e$	$\nu_e e \to \nu_e e$
	$\tau \to l\nu\nu$	$e^+e^- \to \nu_\mu \overline{\nu}_\mu$
Semileptonic		
Meson	$\pi^+ \to \mu\nu, e\nu, \pi^0 e\nu$	
	$K^+ \to \mu\nu, e\nu, \pi^0 e\nu$	
	$K_L^0 \to \pi^\pm \mu\nu, \pi^\pm e\nu$	
Baryon	$B \to B' l\nu$	$\nu N \to \nu N, \nu N\pi, \nu NX$
	$\mu^- B \to B'\nu$	$\nu + D \to n + p + \nu$
	$\nu B \to B' l$	$eN \to eN, eX$
Nonleptonic		
Meson	$K \to \pi\pi$	
	$K \to 3\pi$	
	$D \to KK, K\pi, K2\pi, \ldots$	
	$B \to D\pi, DK, \ldots$	
Baryon	$\Lambda \to N\pi$	$NN \to NN$
	$\Sigma \to N\pi$	
	$\Xi \to N\pi$	

*N refers to a nucleon, B refers to a baryon, l refers to a charged lepton, and X refers to highly excited hadronic matter in deep inelastic collisions. Some reactions (for example, $\nu_e e \to \nu_e e$) can occur by neutral as well as charged weak interactions. The reaction $eN \to eN$ occurs by electromagnetic as well as weak interactions. The reaction $NN \to NN$ occurs by strong and electromagnetic as well as weak interactions.

try (parity) and are also particle–antiparticle (charge conjugation) symmetric, the weak interaction violates these two symmetries. *See* FUNDAMENTAL INTERACTIONS; LEPTON; PARITY (QUANTUM MECHANICS); QUARKS; SYMMETRY LAWS (PHYSICS).

Weak interactions are classified as charged or neutral, depending on whether or not a particle participating in a weak reaction suffers a change of electric charge of one electronic unit (**Table 2**). For example, the neutrino–nucleon scattering reaction (1) is

$$\nu_\mu + n \to \mu^- + p \tag{1}$$

charged because the neutrino ν_μ, which has zero charge, transforms into a negative muon μ^-, while simultaneously a neutron n (zero charge) becomes a proton p (positive charge). Actually the valence quark composition of a neutron is two d quarks and a u quark, while that of a proton is two u quarks and a d quark. Thus in reaction (1) a d quark transforms to a u quark. The weak scattering reaction (2) is neutral

$$\nu_\mu + p \to \nu_\mu + p \tag{2}$$

because the final neutrino and proton retain their initial charges. Observed charged weak interactions include nuclear beta decay and electron capture, muon capture on nuclei, decays of the μ and τ leptons and the $\pi^\pm$ and K mesons, the slow hyperon decays, decays of c- and b-quarked mesons and baryons, and the decay of the top quark. Charged neutrino–nucleon and neutrino–lepton scattering reactions, and the phenomenon of neutrino oscillations, are additional manifestations of the charged weak interaction. Neutral weak interactions, which were first observed in 1973, include but are not limited to neutral neutrino–nucleon and neutrino–lepton scattering as well as the weak electron–nucleon reaction (3), which can also occur by electromagnetic

$$e^- + N \to e^- + N \tag{3}$$

interaction. *See* BARYON; ELEMENTARY PARTICLE; HYPERON; MESON; NEUTRAL CURRENTS.

Early history. Study of weak interactions began with the discovery of radioactivity by H. Becquerel in 1896 and the recognition shortly thereafter that in one form of radioactivity the decaying nucleus emits "beta rays" (electrons). Thus nuclear beta decay was the first known weak process. In 1914 J. Chadwick observed that the electrons in beta decay of a given isotopic sample are emitted with a continuous spectrum of energies. This result and subsequent observations in the 1920s led to a crisis, because the energy available to the electron in beta decay is essentially the difference in rest energies of the initial and final nuclei, which is a fixed quantity. Thus, it appeared that the principle of energy conservation was violated in beta decay. In order to rescue that principle, W. Pauli proposed in 1930 and again in 1933 that a neutral particle of small or vanishing rest mass and half-integral spin (later called the neutrino) is emitted along with the electron in nuclear beta decay, that it shares the available fixed energy with the electron (leading to the continous electron energy spectrum), and that it escapes observation because of its feeble interactions with surrounding matter. *See* BETA PARTICLES; NEUTRINO; RADIOACTIVITY.

Fermi's theory. In 1934 E. Fermi proposed a theory of beta decay based on Pauli's neutrino hypothesis. In Fermi's theory, beta decay is described by the coupling of two charged weak "current densities" at one common space-time point. For example, in neutron beta decay, reaction (4), one of these current densi-

$$n \to p + e^- + \overline{\nu}_e \tag{4}$$

ties describes the neutron-to-proton transformation, while the other accounts for the simultaneous creation of the electron and antineutrino. Each current density behaves as a four-vector (V) under Lorentz transformations. The strength of the coupling is characterized by the "Fermi coupling constant" G_F, a quantity that must be determined by experiment and is found to be given by Eqs. (5) or in units where $\hbar = c = 1$ (where $\hbar$ is Planck's constant divided by 2π and c is the speed of light) by Eq. (6).

$$\begin{aligned} G_F &= 1.43584 \pm 0.00001 \times 10^{-49}\,\mathrm{erg \cdot cm^3} \\ &= 1.43584 \pm 0.00001 \times 10^{-62}\,\mathrm{J \cdot m^3} \end{aligned} \tag{5}$$

$$G_F = 1.16637 \pm 0.00001 \times 10^{-5}\,\mathrm{GeV^{-2}} \tag{6}$$

See RELATIVITY.

Fermi's theory gave a good phenomenological account of many aspects of nuclear beta decay, especially when generalized by G. Gamow and E. Teller in 1936. It also anticipated a number of future developments, for in the 25 years following Fermi's proposal, new weak processes were found in addition to nuclear beta decay, with the discovery of new

elementary particles and elucidation of their decay schemes. Gradually it became clear that these bear many similarities to nuclear beta decay and appear to be different manifestations of a universal "Fermi" interaction. However, it was recognized quite early (by W. Heisenberg in 1936, among others) that the Fermi theory, while undoubtedly very useful, cannot be fundamental, because it leads to a failure of unitarity (nonconservation of probability) when it is applied to high-energy processes such as neutrino–electron scattering. A related and serious shortcoming is that the Fermi theory is not renormalizable: It contains incurable divergences that occur in the calculation of higher-order corrections. *See* RENORMALIZATION.

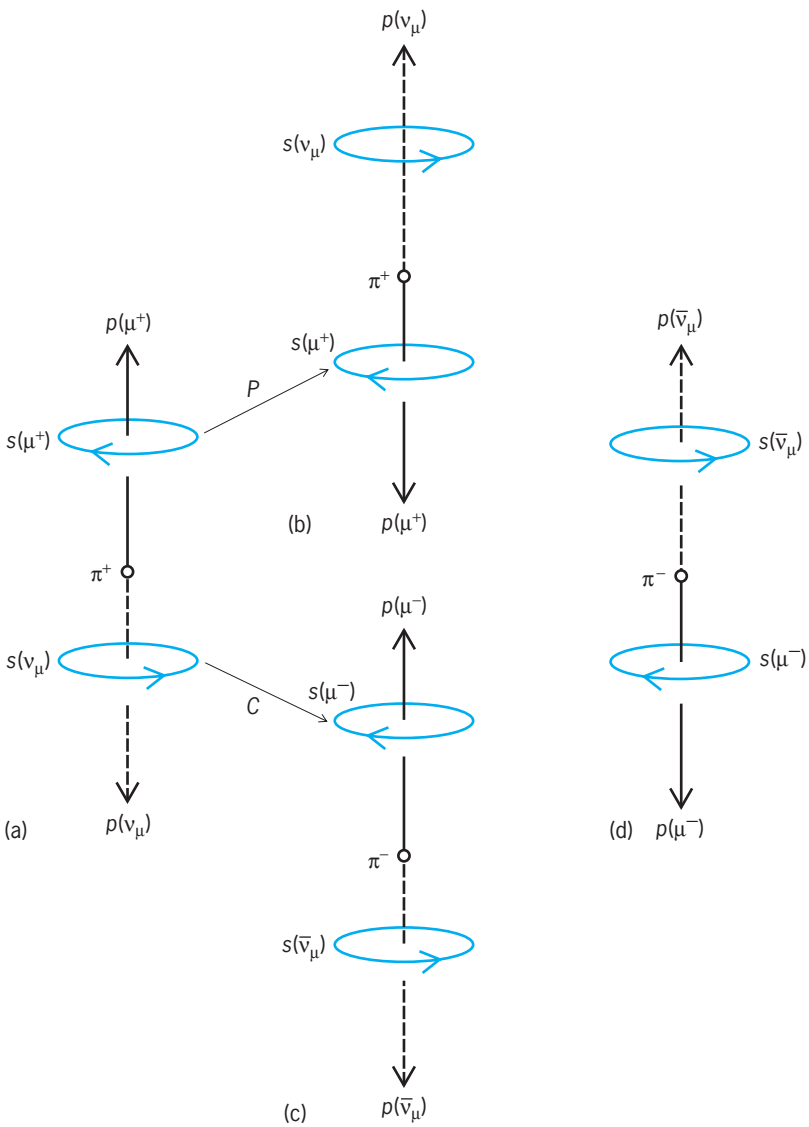

Fig. 1. Weak decay of the pion, illustrating parity (*P*) and charge conjugation (*C*) violation in weak interactions. (*a*) The decay $\pi^+ \rightarrow \mu^+ + \nu_\mu$. Experiment shows that μ^+ spin [here labeled $s(\mu^+)$] is opposed to its momentum $p(\mu^+)$. In the pion rest frame, the neutrino momentum $p(\nu_\mu)$ must be equal and opposite to $p(\mu^+)$. Also, since the pion has zero spin angular momentum, the neutrino spin $s(\nu_\mu)$ must be opposed to $s(\mu^+)$. (*b*) The result of a parity transformation on *a*. The momenta $p(\mu^+)$ and $p(\nu_\mu)$ are reversed, but the spins $s(\mu^+)$ and $s(\nu_\mu)$ remain unchanged. Here $p(\mu^+)$ and $s(\mu^+)$ are parallel, which is never observed. (*c*) The result of a charge conjugation transformation on *a*. Here the spins and momenta are left unchanged, but $\pi^+ \rightarrow \pi^-$, $\mu^+ \rightarrow \mu^-$, $\nu_\mu \rightarrow \overline{\nu}_\mu$. The arrangement shown here is never observed; experiment shows that in π^- decay the μ^- spin and momentum are always parallel. (*d*) The result of a *CP* transformation on *a*. This figure correctly represents the decay $\pi^- \rightarrow \mu^- + \overline{\nu}_\mu$

Parity nonconservation. In the 1950s the accumulation of data on the weak decays of what is now known as K^+ mesons led to a peculiarity called the tau–theta ($\tau - \theta$) puzzle. A search for the solution of this puzzle induced T. D. Lee and C. N. Yang in 1956 to propose a dramatic, important, and far-reaching hypothesis: Parity is violated in weak interactions. Their ideas were almost immediately subjected to experimental test and quickly vindicated. An example of parity violation appears in the weak decay of the charged pion: $\pi^+ \rightarrow \mu^+ + \nu_\mu$ (**Fig. 1**). In observed $\pi^+ \rightarrow \mu^+ + \nu_\mu$ decay (Fig. 1*a*), the μ^+ spin is found experimentally to be opposite to its linear momentum: helicity $h(\mu^+) = -1$. Since the π^+ spin is zero and angular momentum as well as linear momentum is conserved, the neutrino must be emitted in the opposite direction to that of the muon, and its helicity must also be $h(\nu_\mu) = -1$. A parity (*P*) transformation is equivalent to a mirror reflection in the plane containing the pion and perpendicular to the lepton linear momenta. This transformation reverses the muon and neutrino linear momenta but leaves their spins invariant. Thus under *P*, $h(\mu^+)$ and $h(\nu_\mu)$ are reversed, but this outcome (Fig. 1*b*) is never observed. Since parity invariance requires that a physical process and its mirror image occur with the same probability, one has here a "maximal" parity violation. One can also make a charge conjugation transformation on the particles of Fig. 1*a*. This changes the particles into the corresponding antiparticles but leaves the spins and linear momenta, and hence the helicities, invariant (Fig. 1*c*). However, the process of Fig. 1*c* is never observed, for experiment shows that in π^- decay, $h(\mu^-) = +1$. Thus, charge conjugation symmetry (*C*) is also violated maximally. On the other hand, a combined charge conjugation and parity (*CP*) transformation on $\pi^+ \rightarrow \mu^+ + \nu_\mu$ results in observed $\pi^- \rightarrow \mu^- + \overline{\nu}_\mu$ decay (Fig. 1*d*). *See* HELICITY (QUANTUM MECHANICS).

V-A law. The discovery of *P* and *C* violation inspired intense experimental and theoretical activity that led to the V-A law, proposed by E. Sudarshan and R. Marshak, and independently by R. Feynman and M. Gell-Mann, in 1958. In this formulation Fermi's theory was generalized in two respects: first, parity violation was accounted for by including an axial vector (A) as well as a vector (V) portion in each weak current density; and second, the Fermi-type couplings were generalized to describe not only beta decay but also muon decay, decays of charged pions, and so forth. This scheme was thus able to give an excellent account of the observed features of low-energy charged weak processes, including the newly discovered parity-violating effects, but since it was nothing more than an enhanced version of Fermi's theory, it suffered from the same defects: It is not renormalizable and leads to difficulties (breakdown of unitarity) at high energies. A radical change in point of view was thus necessary.

Emergence of the standard model. The path to a better theory starts with consideration of the possibility that a weak interaction does not occur at a single point in space–time as in Fermi-type

theories, but proceeds instead through the exchange of an intermediate boson. This idea arises naturally by analogy with the original meson theory of nuclear forces (H. Yukawa, 1935) and with quantum electrodynamics (QED), where the Coulomb force between two charged particles occurs by exchange of a photon, the spin-1 quantum of the electromagnetic field (**Fig. 2***a*). The Coulomb force is long-range (it varies as $1/r^2$), which means that the photon mass is zero. However, experiment shows that the weak force, like the nuclear force, has very short range. Therefore, if intermediate bosons (also spin-1 quanta) mediate the weak interaction, they must be massive. Also, there must be at least three such bosons: a charged W^- and its charge conjugate W^+ to transmit the charged weak interaction (Fig. 2*b*), and a neutral boson Z^0 for the neutral weak interaction (Fig. 2*c*). Thus, the problem may be stated as follows: How can one construct a renormalizable theory of weak interactions in which massive vector bosons, both charged and neutral, are exchanged between fundamental fermions?

The solution was indeed found in a theory uniting weak and electromagnetic interactions, and it was based on a combination of subtle ideas. First is the notion that the theory must be invariant under local gauge transformations, which are phase transformations of a quantum field that can vary in an arbitrary manner from one space–time point to another. Quantum electrodynamics is such a theory. Here the gauge symmetry group is U(1), the group of all 1×1 unitary matrices. The quanta of the electromagnetic field that emerge from these considerations are photons, which carry no charge. In 1954, Yang and R. Mills considered an analogous theory that employs instead the gauge group SU(2), the group of all 2×2 unitary matrices with determinant equal to unity. In this way Yang and Mills obtained a renormalizable theory similar to quantum electrodynamics but containing a triplet of spin-1 quanta with "charges" $+g$, $-g$, and zero, where g is the coupling constant in their theory. These quanta seem to be attractive candidates for the weak vector bosons. However, the Yang–Mills quanta are massless, while weak vector bosons must necessarily be massive. *See* GAUGE THEORY.

The second important idea, which overcomes this obstacle, is spontaneous symmetry breaking. This means that one has a quantum field theory possessing a certain symmetry not shared by the ground state of the system. It was shown by J. Goldstone in 1961 that when this situation occurs in a non-gauge field theory, a massless spin-0 excitation—the so-called Goldstone boson—emerges, corresponding to each degree of freedom in which the symmetry is broken. There is no experimental evidence for such bosons, so it would appear that spontaneous symmetry breaking does not occur. However, the proof of Goldstone's theorem is based on two assumptions, and in a gauge theory (for example, electrodynamics or the Yang-Mills theory) these assumptions cannot both be valid. Thus gauge theories evade the Goldstone theorem, as was first noted by P. Higgs in 1964. A most extraordinary result of these arguments is that the difficulties associated with massless

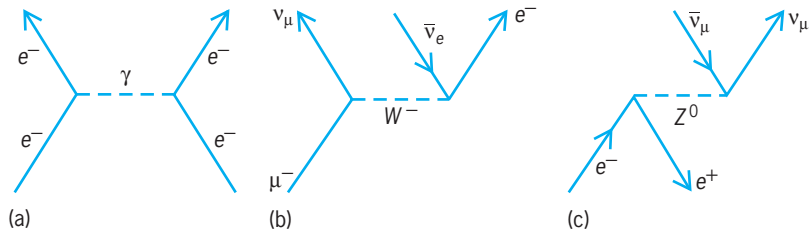

Fig. 2. Schematic diagrams of representative intermediate boson exchange processes between fermions. In these diagrams time increases upward. (*a*) Electron–electron scattering by photon exchange (electromagnetic interaction). (*b*) Muon decay ($\mu^- \rightarrow e^- + \overline{\nu}_e + \nu_\mu$) by W^- exchange (charged weak interaction). The corresponding charge conjugate process ($\mu^+ \rightarrow e^+ + \nu_e + \overline{\nu}_\mu$) by W^+ exchange occurs with the same total probability. (c) Neutrino pair production ($e^- + e^+ \rightarrow \nu_\mu + \overline{\nu}_\mu$) by Z^0 exchange (neutral weak interaction).

vector bosons and Goldstone bosons can be made to neutralize one another: By a suitable transformation, the Goldstone bosons disappear and simultaneously the vector bosons acquire mass. An inevitable consequence, however, is the appearance of at least one type of massive scalar particle, called a Higgs boson. Although Higgs bosons are required by the theory, no definite prescription can be given in this theory for their mass, which can be anywhere between about 7 GeV/c^2 and about 1000 GeV/c^2 a priori. Employing these ideas and others contributed earlier by S. Glashow, S. Weinberg (1967) and independently A. Salam (1968) constructed a successful gauge theory of the Yang-Mills type combining weak and electromagnetic interactions, and this theory was proved to be renormalizable by G. 't Hooft in 1971. In 1974, the Glashow-Weinberg-Salam theory was combined with a new gauge theory of strong interactions, called quantum chromodynamics, and the synthesis became known as the standard model. *See* ELECTROWEAK INTERACTION; HIGGS BOSON; QUANTUM CHROMODYNAMICS; STANDARD MODEL; SYMMETRY BREAKING.

The electroweak sector of the standard model contains two gauge symmetry groups, $SU(2)_L$ and U(1), with coupling constants g and g', respectively. [The subscript L refers to "left-handed fermions" (and "right-handed antifermions"), meaning that parity violation is built into the theory from the beginning.] As mentioned previously, there is a triplet of gauge fields associated with SU(2) that have "SU(2) charges" $+g$, $-g$, and 0, while there is only a single neutral field in U(1). In the standard model, the physical electromagnetic and neutral weak fields are constructed as orthogonal linear combinations of the neutral SU(2) and U(1) fields. These linear combinations are characterized by a weak mixing angle, $\theta_W = \tan^{-1}(g'/g)$, which is found to be $\theta_W = 0.49$ rad. Various relationships [Eqs. (7)–(9)] between g, θ_W, the masses m_W and m_Z,

$$\sqrt{4\pi\alpha} = g \sin\theta_W \tag{7}$$

$$m_z = \frac{m_W}{\cos\theta_W} \tag{8}$$

$$G_F = \frac{g^2}{2^{5/2} m_W^2} \tag{9}$$

Fermi's constant G_F, and the fine-structure constant

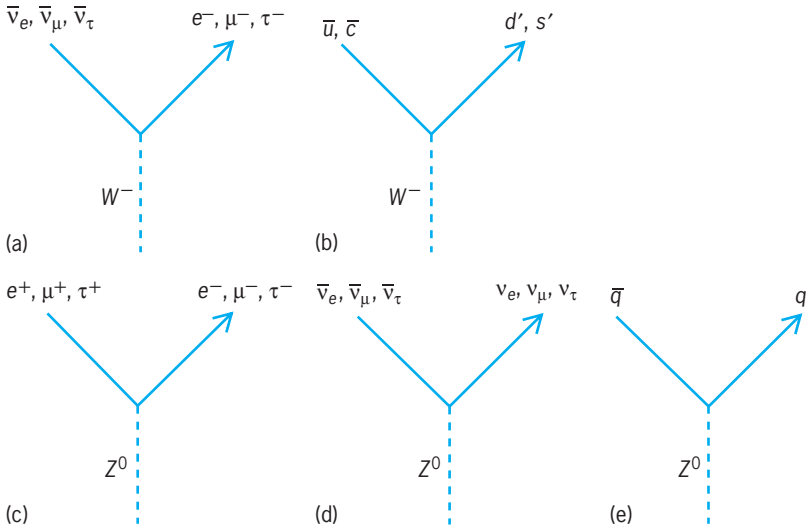

Fig. 3. Decay modes of intermediate vector bosons W^- and Z^0. (*a*) W^- decay to a charged lepton and an antineutrino. (*b*) W^- decay to a quark and an antiquark. Here, d' and s' refer to the weak eigenstates obtained from the quark mass eigenstates *d*, *s*, and *b* by the CKM matrix transformation. (The decay $W^- \rightarrow \bar{t}b'$ cannot occur because the mass of the top quark is too large. Instead the chief mode of decay of a top quark is $t \rightarrow W^+b$.) (*c*) Z^0 decay to a charged lepton–antilepton pair. (*d*) Z^0 decay to a neutrino–antineutrino pair. (*e*) Z^0 decay to a quark–antiquark pair. Here, *q* refers to a *u*, *d*, *c*, *s*, or *b* quark, while $\bar{q}$ refers to the corresponding antiquark. (The Z^0 cannot decay to a $t\bar{t}$ pair, again because the mass of the top quark is too large.)

$\alpha = 1/137.036$ (which is the fundamental parameter of quantum electrodynamics) follow from the mixing and from the requirement that the new theory conform to the known and valid results of quantum electrodynamics and the V-A law for low-energy charged weak interactions. The standard model also predicts the existence and detailed properties of neutral weak interactions, which were first observed in 1973, about 5 years after the theory's creation.

Verification of the standard model. Experiments carried out between 1973 and 1980 confirmed the correctness of the theory for low-energy neutral weak interactions (that is, in reactions where the energy transmitted by the intermediate boson Z^0 is small compared to its rest energy). These included neutrino–electron scattering, neutrino–nucleon scattering, parity violation in atoms, and scattering of polarized electrons on nucleons. The latter two cases involve interference between the weak and electromagnetic interactions. Then, in 1982–1983 at the European Center for Nuclear Research (CERN), the intermediate bosons $W^\pm$ and Z^0 were directly produced in energetic proton–antiproton ($p\bar{p}$) collisions, and detected by means of their decay products (**Fig. 3**). Subsequently, experiments were carried out at the Large Electron-Positron Collider (LEP) at CERN between 1989 and 2000, and at the Stanford Linear Collider (SLC) of the Stanford Linear Accelerator Center (SLAC), at more or less the same time, to investigate the properties of Z and W bosons in great detail. Precise measurements of the W mass were also undertaken at the Tevatron at Fermilab. In these experiments, the predictions of the standard electroweak model concerning the masses of W and Z, their couplings to leptons and quarks, and their relation to the fine-structure constant, including quantum-field-theoretic "radiative" corrections to first nonvanishing ("one-loop") order, were verified to remarkable precision. These radiative corrections include effects due to the top quark and the Higgs boson. They may thus be employed in conjunction with experimental data to determine the top quark mass m_t with small uncertainty, and to put useful lower and upper bounds on the Higgs boson mass m_H. The value of m_t obtained in this way agrees with the value from direct observations of the top quark that have been made since 1995–1996 at Fermilab. Incontrovertible observations of the Higgs boson do not yet exist, but a lower limit exists on m_H from the negative results of experimental searches so far. This limit is consistent with the limits obtained from radiative corrections (**Table 3**). It is widely expected that direct detection of the Higgs boson will be achieved by experiments at the Large Hadron Collider (LHC) at CERN, which should begin operation in 2007. *See* INTERMEDIATE VECTOR BOSON; PARTICLE ACCELERATOR.

Cabibbo-Kobayashi-Maskawa matrix. Although the fundamental quark states for strong interactions are as shown in Table 1, extensive data on charged weak decays of mesons and baryons from diverse experiments reveal the necessity of a unitary transformation of quark states *d*, *s*, and *b* to a new basis suitable for charged weak interactions. The first evidence for this was uncovered in the 1960s by N. Cabibbo, at a time when only the three "light" quarks, *u*, *d*, and *s* (and their corresponding antiquarks), were recognized. Cabibbo described the necessary linear transformation of *d* and *s* quark states by a single

TABLE 3. Some important physical quantities in weak interactions

Lepton parameters	
Electron mass	$m_e = 0.51099892 \pm 0.00000004$ MeV/c^2
Muon mass	$m_\mu = 105.658369 \pm 0.000009$ MeV/c^2
Muon mean life	$\tau_\mu = (2.19703 \pm 0.00004) \times 10^{-6}$ s
Tau lepton mass	$m_\tau = 1776.99^{+0.29}_{-0.26}$ MeV/c^2
Tau mean life	$\tau_\tau = (2.906 \pm .011) \times 10^{-13}$ s
Direct mass limits on neutrinos from conservation of energy and momentum in various decays and reactions	
$m(\nu_e) \leq 3$ eV/c^2	
$m(\nu_\mu) \leq 0.19$ MeV/c^2	
$m(\nu_\tau) \leq 18.2$ MeV/c^2	
Number of neutrino families from Z^0 decay width (LEP data): 2.994 ± 0.012	
Quark masses	
$m_u = 1.5 - 4.5$ MeV/c^2	
$m_d = 5 - 8.5$ MeV/c^2	
$m_s = 80 - 155$ MeV/c^2	
$m_c = 1.290^{+0.040}_{-0.045}$ GeV/c^2	
$m_b = 4.206 \pm 0.031$ GeV/c^2	
$m_t = 178.0 \pm 4.3$ GeV/c^2 (direct observation at Fermilab)	
$= 177.9 \pm 4.4$ GeV/c^2 (from standard model radiative corrections)	
Intermediate boson masses	
$m\gamma < 6 \times 10^{-17}$ eV/c^2	
$m_W = 80.425 \pm 0.038$ GeV/c^2	
$m_Z = 91.1876 \pm 0.0021$ GeV/c^2	
Higgs boson mass	
$m_H \geq 114.4$ GeV/c^2 (from negative results of direct searches for the Higgs boson)	
$m_H = 113^{+56}_{-40}$ GeV/c^2 (from standard model radiative corrections)	

"rotation angle." In 1973 M. Kobayashi and T. Maskawa proposed a generalization of these ideas to three generations of quarks (before direct experimental evidence for the third generation was available).

Assuming that there are three and only three quark generations, the transformation of d, s, and b states to the new basis, d', s', and b', suitable for charged weak interactions, is implemented by means of a 3×3 unitary matrix, called the Cabibbo-Kobayashi-Maskawa (CKM) matrix, as in Eq. (10). It may be

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{10}$$

shown quite generally that the various matrix elements V_{ij} can be expressed in terms of three rotation angles, θ_{12}, θ_{23}, and θ_{13} (which generalize the original Cabibbo angle), and one complex phase δ. In the standard model, it is assumed that this phase alone accounts for the phenomenon of *CP* violation. No reliable theoretical model exists to calculate the matrix elements V_{ij} from first principles, and their origin is still a mystery.

CPT invariance and CP violation. All experimental evidence is consistent with *CPT* (combined charge conjugation, parity, and time reversal) invariance of strong, electromagnetic, and weak interactions, and there are compelling theoretical reasons to believe that *CPT* invariance is a universally valid symmetry. After *P* and *C* violation of the weak interaction was discovered in 1956–1957, it appeared for some years that the combined symmetry *CP* was also valid. However, it was found in 1964 that *CP* invariance is violated in certain weak decays of neutral *K* mesons. *See* CPT THEOREM.

The neutral *K* mesons, K^0 and $\overline{K}^0$, are charge conjugates of one another (their valence quark compositions are $K^0 = d\overline{s}$, $\overline{K}^0 = \overline{d}s$) and they possess definite strangeness eigenvalues (+1 and −1, respectively). Hence, it is appropriate to think in terms of the states K^0 and $\overline{K}^0$ when considering the strong and electromagnetic interactions, which conserve strangeness. However, K^0 and $\overline{K}^0$ do not possess definite lifetimes for weak decay, nor do they possess definite masses, because the charged weak interactions do not conserve strangeness. Instead, there exist two independent linear combinations of K^0 and $\overline{K}^0$ states, called K_S^0 and K_L^0, which do have definite (and distinct) lifetimes and masses. This causes the phenomenon of K^0–$\overline{K}^0$ mixing or strangeness oscillations, first described in the 1950s by Gell-Mann and A. Pais, and independently by T. Nakano and K. Nishijima. The short-lived K_S^0 decays in only two significant modes, $\pi^+\pi^-$ and $\pi^0\pi^0$, and each of these final states has *CP* eigenvalue +1. There are many known modes of decay of K_L^0, including the fully allowed decay to $\pi^+\pi^-\pi^0$, a final state which predominantly has *CP* eigenvalue −1. Thus, if *CP* were conserved, the decays $K_L^0 \to \pi^+\pi^-$ and $K_L^0 \to \pi^0\pi^0$ would be strictly forbidden. However, J. Cronin, V. Fitch, and coworkers observed such decays with small but finite probability in 1964, which implies *CP* violation, and these results have been verified in great detail in the years since.

The data reveal that the largest *CP*-violating effect occurs in the decay amplitudes of K_L^0 to various final states ("indirect" *CP* violation) but there is also a small so-called direct *CP* violation in K^0–$\overline{K}^0$ mixing. The standard model allows for the possibility of direct as well as indirect *CP* violation in kaon decay, but theoretical prediction of direct *CP* violation within the standard model framework has large uncertainties due to the strong interaction. Hence it is not yet possible to conclude that the standard model satisfactorily accounts for *CP* violation in neutral kaon decay.

CP violation has also been observed in the $B^0 - \overline{B}^0$ system (at the BABAR installation at SLAC, and at the BELLE installation in Japan). In these experiments, e^+e^- collisions produce $b\overline{b}$ quark pairs which "hadronize" to form a $B^0 = b\overline{d}$, $\overline{B}^0 = \overline{b}d$ meson pair (or $B_s^0 = b\overline{s}$ and $\overline{B}_s^0 = \overline{b}s$ mesons) with equal and opposite momenta in the center-of-momentum frame. Because of "particle mixing" through the weak interaction, which is somewhat analogous to that in the neutral kaon system, each meson evolves continuously into its charge conjugate and back again as time elapses. However, at some point one of the mesons decays. Observation of this "tagged" decay serves to identify whether the tagged meson was B^0 or $\overline{B}^0$ when it decayed, implying that the surviving meson is of the other type. The "tagging" also starts a clock for measuring the time-dependent probability for the surviving meson to decay to a final state with definite *CP* eigenvalue (for example, $J/\psi K_L^0$). In this way a *CP*-violating asymmetry in the time-dependent decay probabilities can be measured. The resulting effects are large, and for the decay modes observed so far, they are in agreement with predictions of the standard model. However, there is reason to believe that some rare decay modes of B^0 or B_s^0 mesons, yet to be observed with precision in these experiments, may exhibit *CP*-violating effects that must be accounted for by descriptions that go beyond the standard model.

Given *CPT* invariance, *CP* violation necessarily implies *T* violation. Thus efforts have been made for many years to observe *T* violation elsewhere than in the neutral kaon or neutral bottom meson systems. One particular case is of special interest: the search for permanent spin electric dipole moments of the neutron and the electron.

These particles, among others, possess spin $^1/_2$ and have spin magnetic dipole moments. However, they cannot possess electric dipole moments unless both parity (*P*) and time reversal (*T*) invariance are violated. Of course, parity is indeed violated in the weak interaction and *CP* violation (equivalent to *T* violation) also occurs. Thus, electric dipole moments can exist through radiative corrections to the electromagnetic interaction arising from the weak interaction and the *CP*-violating mechanism. However, it can be shown that so long as *CP* violation is

described by the standard model, the electric dipole moments of the neutron and the electron would be far too small by many orders of magnitude for any practical experiment to detect. On the other hand, these electric dipole moments could be within experimental reach according to various plausible models of new physics beyond the standard model (for example, supersymmetric models of various types). Motivated by this possibility, sensitive searches for the neutron and electron electric dipole moments have been carried out, and new searches are being prepared. However, only upper limits on electric dipole moments have been obtained so far. *See* DIPOLE MOMENT; TIME REVERSAL INVARIANCE; SUPERSYMMETRY.

Weak interactions of neutrinos. According to the standard model, neutrinos and antineutrinos have zero mass, and ν_e, ν_μ, ν_τ have helicity $h = -1$ (are left-handed), whereas $\overline{\nu}_e$, $\overline{\nu}_\mu$, $\overline{\nu}_\tau$ have $h = +1$ (are right-handed). Is this description correct? Are neutrinos really massless, like photons, or do they possess small but nonzero mass? If at least one neutrino species has mass, the possibility exists that the neutrino weak interaction eigenstates ν_e, ν_μ, ν_τ are not the same as the neutrino states of definite mass (the latter are called ν_1, ν_2, ν_3). This idea of neutrino mixing is suggested by analogy with quark mixing, where, as noted earlier, the quark weak eigenstates, d', s', and b' are not the same as the mass eigenstates, d, s, and b, but are related to the latter by the CKM matrix of Eq. (10).

In fact, in four distinct experiments performed since 1998 with atmospheric neutrinos, accelerator neutrinos, and reactor neutrinos in Japan, and with solar neutrinos in Canada, compelling evidence has been found that at least two neutrino species do have nonzero mass, and that neutrino mixing actually occurs. The mixing is described by a unitary 3×3 matrix U (the "lepton mixing matrix") in Eq. (11).

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \qquad (11)$$

The matrix U, like the CKM matrix, is characterized by three "lepton" mixing angles, θ_{12}, θ_{23}, and θ_{13}, and a phase δ, although the numerical values of the leptonic quantities are not the same as the analogous CKM quantities.

Neutrino mixing causes the remarkable phenomenon of neutrino oscillations, in which a neutrino created in a definite weak eigenstate at some initial time $t = 0$ transforms into a superposition of all neutrino weak eigenstates as time elapses. For example, a solar neutrino is born in the core of the Sun as ν_e, but by the time it arrives at a detector on Earth it is in a superposition of states ν_e, ν_μ, and ν_τ. The probability to find such a neutrino at the detector in the state ν_e is therefore less than unity, and a detector that is sensitive only to ν_e will thus register fewer counts than it would have if neutrino oscillations did not occur. On the other hand, a detector that is equally sensitive to each of the neutrino weak eigenstates, ν_e, ν_μ, and ν_τ, will not suffer any such loss of counts. Experiments carried out at the Sudbury Neutrino Observatory (SNO) in Canada using both schemes of detection thus demonstrated the existence of solar neutrino oscillations and at the same time verified that the solar neutrino flux is as predicted by detailed theoretical models of the Sun.

The leptons (e^-, μ^-, τ^-) and their respective antileptons (e^+, μ^+, τ^+) have opposite electric charges and opposite magnetic moments, and are therefore different from one another. Because of this difference they are called Dirac particles. Each neutrino of a given type may also be different from its corresponding antineutrino (may be a Dirac neutrino). However, neutrinos are electrically neutral, and it is therefore theoretically possible for a neutrino of a given type to be identical to its corresponding antineutrino (in which case it is referred to as a Majorana neutrino). The distinction between these two possibilities has physical meaning only if the neutrino has nonzero mass. Thus, now that the existence of neutrinos with mass has been demonstrated, the question "Are neutrinos of the Dirac type or the Majorana type?" becomes especially significant. It can be addressed experimentally as follows. Nuclei exist for which ordinary beta decay is energetically forbidden or highly suppressed by conservation of angular momentum, but which are unstable for double beta decay, a second-order process in which two electrons are emitted. Ordinarily these two electrons are accompanied by two antineutrinos ("$\beta\beta_2$" decay). However, neutrinoless double beta decay ("$\beta\beta_0$" decay), although very difficult to observe, is also possible, but only if neutrinos are of the Majorana type. Experimental searches for $\beta\beta_0$ decay are being carried out, but they have not yet yielded definitive results.

Eugene D. Commins

Bibliography. T.-P. Cheng and L.-F. Li, *Gauge Theory of Elementary Particle Physics*, Oxford University Press, 1984, reprinted with corrections, 1988; E. D. Commins and P. H. Bucksbaum, *Weak Interactions of Leptons and Quarks*, Cambridge University Press, 1983; M. Fukugita and T. Yanagida, *Physics of Neutrinos*, Springer, Berlin, 2003; F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, Wiley, New York, 1984; B. R. Holstein, *Weak Interactions in Nuclei*, Princeton University Press, 1989; P. Langacker (ed.), *Precision Tests of the Standard Electroweak Model*, World Scientific, Singapore, 1995; L. B. Okun, *Particle Physics: The Quest for the Substance of Substance*, Harwood, London, 1985.

Weakly interacting massive particle (WIMP)

A hypothetical elementary particle that might make up most of the matter in the universe, and that is also predicted to exist in supersymmetry theory. Most matter is detected only through its gravitational effects; this "dark matter" has not been observed to emit, absorb, or reflect light of any wavelength. The total amount of dark matter appears to be approximately ten times as great as all the ordinary matter in the universe, and about one hundred times as great

as all the visible matter. The nature of the dark matter is not yet known, although many experiments are under way to try to discover it directly or indirectly. *See* COSMOLOGY; UNIVERSE.

Almost all the currently available data in elementary particle physics can be accounted for by a theory called the standard model, in which matter is made of quarks (the building blocks of protons and neutrons) and leptons (including electrons and neutrinos), while the strong, weak, and electromagnetic forces are transmitted by particles like the photon (the carrier of electromagnetic forces). However, the standard model does not predict the existence of any particle—say, χ—that could be the dark matter. Most efforts to go beyond the standard model of particle physics have been based on the idea of supersymmetry, and most versions of supersymmetry predict that there will be a stable weakly interacting massive particle (WIMP) that would be a natural candidate for the χ particles. Dark matter made of WIMPs would be "cold" dark matter (CDM), and a version of CDM theory has become the standard theory of structure formation in cosmology. Its predictions agree very well with the observed properties of the universe on large scales, but there may be disagreements on scales smaller than galaxies. *See* ELEMENTARY PARTICLE; STANDARD MODEL; SUPERSYMMETRY.

Evidence for weakly interacting dark matter. There is now abundant evidence for dark matter around galaxies and clusters of galaxies, and on larger scales in the universe. In the solar system, the more distant a planet is from the Sun, the slower its velocity around the Sun. This is as expected from standard gravity theory and mechanics, since almost all the gravitating mass in the solar system is in the Sun at its center. But gas and satellites at large distances from galaxies have orbital velocities similar to those at smaller distances from the center, which indicates that most of the mass in the galaxy must not be near the center, where most stars are, but in a roughly spherical dark matter halo that extends to perhaps ten times the optical size of the galaxy and has a mass at least ten times that of all the stars. Confirmation of the existence of such dark-matter halos has come from gravitational lensing observations, showing that light from more distant galaxies is bent by the gravity of nearer galaxies. *See* CELESTIAL MECHANICS; GALAXY, EXTERNAL; GRAVITATIONAL LENS; MILKY WAY GALAXY.

Large, nearly spherical halos are expected to form by gravitational collapse in the expanding universe. Once such a halo forms, the gas in it (mostly hydrogen and helium) can continue to collide and convert kinetic energy to radiation, sinking closer to the center of attraction. If the gas retains its angular momentum, it will spin faster as its radius decreases, and will ultimately form a disk. When much of the gas has formed into stars, such a galaxy would be a spiral, the most abundant type of large galaxy. Collisions between such disk galaxies may produce most elliptical galaxies and bulges of spiral galaxies. But if the dark matter can interact only weakly—that is, via the weak interactions, carried by the gauge particles (generalizations of the photon) W^+, W^-, and Z^0—but not strongly or electromagnetically, that would explain why χ particles cannot lose energy by radiation, but must form extended dark halos. Weakly interacting χ particles would also not be concentrated in minerals on Earth, or anywhere else in the disk of the Milky Way Galaxy for that matter, in agreement with observations. *See* INTERMEDIATE VECTOR BOSON; WEAK NUCLEAR INTERACTIONS.

There is also much evidence for dark matter in clusters of galaxies. The astronomer Fritz Zwicky pointed out in 1933 that the galaxies in one nearby cluster were moving at such high speeds that they would not be held together gravitationally unless there was much more mass than was indicated by the light from their stars. This same was subsequently found to be true of other clusters. Later, similar conclusions were reached from x-ray observations and gravitational lensing observations of clusters. *See* X-RAY ASTRONOMY.

Measurements of the typical angular sizes of structures in the cosmic background radiation—the heat radiation of the big bang—provide information about the composition of the universe, and again lead to the conclusion that most of the matter in the universe is dark matter. All of these data indicate that the total amount of dark matter is about 30–40% of the critical density required to close the universe, while the total amount of ordinary matter is only about 4–5% of the critical density, and the amount of visible matter is less than half a percent of the critical density. The density in units of critical density is usually written as Ω, a dimensionless number; the universe will expand forever if the total average density $\Omega < 1$, and will eventually collapse if $\Omega > 1$. Expressed this way, the average, or cosmological, dark-matter density is $\Omega_{dm} \approx 0.3$–0.4, the density of all ordinary matter (baryons) is $\Omega_b \approx 0.04$–0.05, while the density of visible matter is only $\Omega_{vis} \approx 0.005$. *See* BIG BANG THEORY; COSMIC BACKGROUND RADIATION.

Nature of dark matter. Since about 1980, when the evidence for dark matter became convincing, there have been various theories proposed concerning its possible composition. That dark matter is invisible suggests that the χ particles are electrically neutral. The leading candidates for the dark-matter particle are axions and supersymmetric WIMPs, since both were predicted on the basis of particle physics theories that were motivated by laboratory experiments, and it was only subsequently appreciated that they could solve the dark-matter mystery.

Supersymmetric WIMPs. Supersymmetry is the hypothesis that there is a relationship between the two known classes of particles, bosons and fermions. The bosons, such as the photon and the graviton, are responsible for the forces of nature. The fermions—including the atomic building blocks, the proton, neutron, and electron—are the matter particles. According to supersymmetry, for every kind of boson in the universe, there must also be a corresponding fermion with the same electric charge and very similar interactions with other particles. For example, the supersymmetric analog of the photon has been named the photino, and like the photon, it would be

electrically neutral. Since these hypothetical supersymmetric partner particles have not been discovered yet, if supersymmetry is right their masses must be too large for them to have been produced at current particle accelerators. Thus far, the evidence for supersymmetry is only indirect, but if the theory is right many supersymmetric partner particles should be produced at accelerators such as the Large Hadron Collider (LHC) being built in Geneva, Switzerland and scheduled to start operating in 2006. *See* PARTICLE ACCELERATOR; PHOTON; QUANTUM STATISTICS.

In most versions of supersymmetry, there is a new conserved quantum number called R-parity, and the ordinary particles all have R-parity equal to $+1$, while the superpartner particles all have R-parity equal to -1. Thus while heavier superpartner particles can decay into lighter ones (plus ordinary particles), the lightest superpartner (LSP) must be stable, since there is no lighter negative R-parity particle into which it can decay. Thus, the universe should be full of LSP particles if the temperatures in the early universe was high enough to produce supersymmetric partner particles—which is the case according to most theories of the early universe. *See* CONSERVATION LAWS (PHYSICS); INFLATIONARY UNIVERSE COSMOLOGY.

This argument was first published in 1982, with the gravitino (the superpartner of the graviton) as the LSP. In currently favored versions of supersymmetry theory, however, the LSP is a neutralino (the superpartner of some other neutral particle) such as the photino; the theory of such dark-matter particles was published in 1984. Neutralinos would be massive but have interactions like those of neutrinos and other particles that interact only via the weak interaction (and gravity); thus they would be supersymmetric WIMPs. These hypothetical particles would mostly go through whole planets without ever interacting. Nevertheless, it is feasible to search for such WIMPs with very sensitive new instruments. *See* GRAVITON.

Searching for WIMPs. Efforts to detect WIMPs directly are based on detecting their scattering from nuclei. The weak nature of the interactions of WIMPs means that the probability of their interacting is very small. But for the few that do interact, the amount of energy delivered in the collision is large enough to detect with sensitive instruments. Since the WIMPs are supposed to be gravitationally bound to the Milky Way Galaxy, their velocities v are approximately 300 km/s (180 mi/s), a thousandth the speed of light, c. It follows that the ratio of their kinetic energy ($= mv^2/2$ for these nonrelativistic particles of mass m) to their rest mass energy (mc^2) will be $^1/_2(v/c)^2 \approx 0.5 \times 10^{-6}$. If their rest mass is approximately one hundred times the mass of a proton—the sort of mass expected since the charged superpartners have not yet been produced at particle accelerators—then their kinetic energy will be about 50 kiloelectronvolts. All this kinetic energy would be transferred as recoil energy to an equal-mass nucleus struck head on. This would have two different sorts of effects in a crystal: vibrations and electronic excitations. The vibrations would quickly become random heat vibrations, and these can be detected in crystals cooled to a small fraction of a degree above absolute zero. The excited electrons could be detected as a current if the crystal is a semiconductor, or as emitted light if the crystal is transparent and has the appropriate electronic structure. It is very useful to detect both vibrations and electron excitations, since the main background of events that could be mistaken for WIMP interactions is scattering of electrons in the detector by energetic photons (Compton scattering). Since the background events would put more energy into electron excitations than into crystal lattice vibrations, this can be used to help tell whether a possible signal is coming from WIMP scattering. Many experiments of this type are being built around the world. It is desirable to put these detectors deep underground to shield them from cosmic rays. *See* COMPTON EFFECT; COSMIC RAYS; LATTICE VIBRATIONS; PARTICLE DETECTOR; SCINTILLATION COUNTER.

WIMPs can also be detected indirectly, for example by looking for particles coming from their annihilation. Supersymmetric WIMPs (neutralinos) are their own antiparticles, and their annihilation products include energetic photons (gamma rays) and neutrinos, as well as particle-antiparticle pairs, such as electrons and positrons, or protons and antiprotons. One possible signal is gamma rays from WIMP annihilation at the center of the Milky Way Galaxy, where the WIMP density is elevated because of the structure of dark-matter halos, and perhaps further elevated (or perhaps alternatively decreased—this is a controversial issue) by the presence of a black hole at the galactic center. The Gamma-ray Large Area Space Telescope (GLAST), scheduled for launch in 2006, will look for this signal with greatly increased sensitivity. Another important signal that has been searched for is neutrinos coming from WIMP annihilation in the center of the Sun or the center of the Earth, where some will settle after being slowed down by scattering off atomic nuclei in the Sun or Earth. *See* ANTIMATTER; BLACK HOLE; GAMMA-RAY ASTRONOMY; NEUTRINO ASTRONOMY.

WIMPs are also expected to be produced at accelerators such as the LHC from rapid decays of heavier supersymmetric partner particles, and this could be where they are discovered first if they are not seen before that in direct or indirect search experiments. Failure to see supersymmetric particles at LHC energies would mean that current ideas about supersymmetry are wrong.

WIMPs as cold dark matter. The CDM theory was developed to describe how galaxies and clusters of galaxies would form in a universe in which the dark matter is WIMPs or other particles that would have been "cold" in the early universe—that is, moving so sluggishly that their motion can be neglected. "Hot" dark matter (HDM) had earlier been considered. HDM refers to light particles like neutrinos, thought to have rest energies of perhaps a few electronvolts. Such particles would be moving at nearly the speed of light in the hot early universe, and this would inhibit formation of galaxies. The first

structures to form in an HDM universe would be objects with the mass of a supercluster of galaxies, and galaxies would have to form subsequently through their fragmentation. But superclusters are still forming today, while most galaxies are old. That HDM predicted the wrong sequence of cosmogony led to its downfall, while CDM is a hierarchical structure formation theory in which smaller things form earlier and merge to form larger and larger structures, up to clusters of galaxies, the largest gravitationally bound objects. *See* NEUTRINO.

The distribution and properties of galaxies predicted by the CDM theory agree beautifully with observations. Only at the centers of galaxies is there a possible discrepancy between theory and observation: theory predicts a density cusp at the center of dark-matter halos, while some observations suggest a more uniform dark-matter density there. Although such a cusp would involve only a tiny fraction of a galaxy's mass, this possible discrepancy is a serious concern for the theory that the dark matter is mostly WIMPs. Joel R. Primack

Bibliography. G. Blumenthal et al., Formation of galaxies and large-scale structure with cold dark matter, *Nature*, 311:517–525, 1984; G. Jungman, M. Kamionkowski, and K. Griest, Supersymmetric dark matter, *Phys. Rep.*, 267:195–373, 1996; B. Moore, Evidence against dissipationless dark matter from observations of galaxy halos, *Nature*, 370:629–231, 1994; J. R. Primack, D. Seckel, and B. Sadoulet, Detection of cosmic dark matter, *Annu. Rev. Nucl. Part. Sci.*, 38:751–807, 1988.

Wear

The removal of material from a solid surface as a result of sliding action. It constitutes the main reason why the artifacts of society (automobiles, washing machines, tape recorders, cameras, clothing) become useless and have to be replaced. There are a few uses of the wear phenomenon, but in the great majority of cases wear is a nuisance, and a tremendous expenditure of human and material resources is required to overcome the effects.

For reasons which are hard to explain, wear did not emerge as a technical field until the 1950s. Before then, the feeling seems to have been that wear was an inevitable accompaniment of sliding, and that nothing could be done about it. Perhaps the key event which, rather indirectly, led to people's change in attitude was the building of nuclear reactors as part of the atomic bomb program in the 1940s. These reactors produced radioisotopes of the common engineering metals (iron, copper, chromium, and so on), and these allowed the use of extremely sensitive radiotracer techniques for measuring wear. For example, it became possible to measure the very small amount of wear which occurs when one metal is pressed with a force of a kilogram against another and then slides over it.

Thus, for the first time it became possible to measure wear while it was occurring, rather than by before-and-after measurements. As result of this research, it was established by about 1950 that there were four principal forms of wear—adhesive, abrasive, corrosive, and surface fatigue—and it soon became possible to work out their mechanisms and to express the amount of wear in quantitative terms.

Adhesive wear. This is the only universal form of wear, and in many sliding systems it is also the most important. It arises from the fact that, during sliding, regions of adhesive bonding, called junctions, form between the sliding surfaces. If one of these junctions does not break along its original interface, then a chunk from one of the sliding surfaces will have been transferred to the other surface. In this way, an adhesive wear particle will have been formed. Initially adhering to the other surface, adhesive particles soon become loose and can disappear from the sliding system.

The volume of adhesive wear is governed by the equation $V = kLx/3p$, where V is the total volume of adhesive wear, k is a nondimension constant called the wear coefficient, x is the total distance of sliding, and p is the indentation hardness of the surface expressed as a stress. Typical values of k are given in **Table 1**.

A very great reduction in wear, by factors of up to a million, can be produced in metallic sliding systems by using a good lubricant. Also, there is a great advantage in making unlubricated sliding systems nonmetallic, and well-lubricated systems metallic.

Adhesive wear causes two types of failure, one being a wear-out mode which occurs after long periods of sliding because too much material has been removed, and a seizure mode which occurs in systems which generate wear particles larger than the clearance, thus producing jamming. Since many material combinations give wear particles larger than 10 micrometers, it is dangerous to reduce the clearance of any sliding system below this value. *See* FRICTION.

Abrasive wear. This is the wear produced by a hard, sharp surface sliding against a softer one and digging out a groove. The abrasive agent may be one of the surfaces (such as a file), or it may be a third component (such as sand particles in a bearing abrading material from each surface). Abrasive wear, like adhesive wear, obeys the equation given above; typical values of k are given in **Table 2**.

It will be seen that abrasive wear coefficients are large compared to adhesive ones. Thus, the introduction of abrasive particles into a sliding system can greatly increase the wear rate; automobiles, for example, have air and oil filters to catch abrasive particles before they can produce damage.

TABLE 1. Values of the wear coefficient for adhesive wear

Surface condition	Metal on metal	Nonmetal on metal, or nonmetal on nonmetal
Clean	10^{-2} to 10^{-4}	10^{-5}
Poor to fair lubricant	10^{-4} to 10^{-6}	3×10^{-6}
Good lubricant	10^{-6} to 10^{-8}	10^{-6}

TABLE 2. Values of the wear coefficient for abrasive wear

Process	*k* value
Sharp file	2×10^{-1}
Sandpapering	5×10^{-2}
Loose abrasive grains	5×10^{-3}
Polishing	5×10^{-4}

Corrosive wear. This form of wear arises when a sliding surface is in a corrosive environment, and the sliding action continuously removes the protective corrosion product, thus exposing fresh surface to further corrosive attack. No satisfactory quantitative expression of corrosive wear yet exists, but when analyzing it in terms of the above equation, *k* values are obtained ranging all the way from less than 10^{-5} for surfaces in a gently corrosive environment, to above 10^{-2} for surfaces under severe corrosive attack. *See* CORROSION.

Surface fatigue wear. This is the wear that occurs as result of the formation and growth of cracks. It is the main form of wear of rolling devices such as ball bearings, wheels on rails, and gears. During continued rolling, a crack forms at or just below the surface and gradually grows until a large particle is lifted right out of the surface.

Uses. Most manifestations of wear are highly objectionable, but the phenomenon does have a few uses. Thus, a number of systems for recording information (pencil and paper, chalk and blackboard) operate via a wear mechanism. Some methods of preparing solid surfaces (filling, sandpapering, sandblasting) also make use of wear. *See* ABRASIVE.

Wear is often used to study the operation of sliding systems. Thus, by looking at the wear of someone's shoes, it can be seen whether the wearer is walking properly. In analyzing accidents and failures, one of the main procedures is to look for signs of wear in the wrong place. Wear is specially useful in this case because it can be used in systems that are no longer operative.

Ernest Rabinowicz

Bibliography. B. Bhushan and B. Gupta, *Handbook of Tribology: Materials, Coatings, and Surface Treatments*, 1991; E. Rabinowicz, *Friction and Wear of Materials*, 2d ed., 1995; N. P. Suh, *Tribophysics*, 1986; N. P. Suh and N. Saka (eds.), *Fundamentals of Tribology*, 1980; R. B. Waterhouse (ed.), *Fretting Fatigue*, 1991.

Weasel

The common name for 14 species belonging to the family Mustelidae. These carnivores are native to every continent except Australia and Antarctica (see **table**). They are also absent from Madagascar and most oceanic islands.

Morphology. Weasels have a long slender body and short legs (see **illustration**). Most are brown or reddish-brown on the back and whitish or yellowish on the underside, although *Poecilictis* and *Poecilogale* are black with white spots and/or stripes. During winter in northern areas, the fur of some species, such as the long-tailed and short-tailed weasels, changes entirely to white except for the tip of the tail, which may remain black. The limbs possess five digits with nonretractile, curved claws. Ears are short and rounded. The dental formula is I 3/3, C 1/1, PM 3/3, M 1/2 × 2 for a total of 34 teeth. *See* DENTITION.

Long-tailed weasel (*Mustela frenata*). (*Alden M. Johnson © California Academy of Sciences*)

North American species. Three species occur in North America. The long-tailed weasel (*Mustela frenata*) has a head and body length ranging 203–260 mm (8–10 in.), a tail length of 76–152 mm (4–6 in.), and weighs 85–340 g (3–12 oz). The short-tailed weasel (*M. erminea*) has a head and body length between 170 and 325 mm (6.5 and 12.5 in.), a tail length of 42–120 mm (1.75–4.5 in.), and a weight between 42 and 365 g (1.5 and 12.5 oz). The least weasel (*M. nivalis*) is the smallest carnivore in North America. Adults are usually between 165 and 220 mm (6.5 and 8.5 in.) long including a tail

TABLE 1. Names and distribution of weasal species

Scientific and common name	Distribution
Mustela felipei (Colombian weasel)	South America
M. africana (tropical weasel)	South America
M. nivalis (least or common weasel)	North America, Europe, Asia, Africa
M. erminea (short-tailed weasel; ermine; stoat)	North America, Europe
M. frenata (long-tailed weasel)	North and South America
M. altacia (mountain weasel)	Asia
M. kathiah (yellow-bellied weasel)	Asia
M. sibirica (Siberian weasel)	Asia
M. lutreolina (Indonesian mountain weasel)	Sumatra, Java
M. nudipes (Malaysian weasel)	Malaysia, Borneo
M. strigidorsa (back-striped weasel)	Southern Asia
Lyncodon patagonicus (Patagonian weasel)	Argentina, Chile
Poecilictis libyca (North African striped weasel)	Africa
Poecilogale albinucha (African striped weasel)	Africa

25–45 mm (1–2 in.) long. They weigh between 25 and 65 g (1 and 2.3 oz). Male weasels are usually much larger than females.

Behavior and habitat. Weasels are very quick and their movements are difficult to follow. When running, the back is strongly arched at each bound and the tail is held straight out behind or at an angle. Weasels have keen senses of smell, sight, and hearing and are tireless hunters. They are voracious feeders, catching animals their own size or larger for food. They feed primarily on mice, rats, shrews, and moles, although earthworms, insects, frogs, lizards, rabbits, squirrels, snakes, and small birds may also be eaten. Weasels sever the prey's spinal column with one or two strong, swift bites at the base of the skull. The weasels' slim body and short legs enable it to hunt in the narrow openings of stone walls, under logs and rocks, and in rodent burrows. A weasel can follow a mouse to the end of the mouse's burrow and can squeeze through knotholes into chicken coops. Weasels raid farms and often kill more chickens than they need for food. As a result, many farmers dislike weasels even though they destroy farmyard pests. Although chiefly ground-dwelling animals, many weasels are excellent tree climbers and are able to raid birds' nests.

Weasels are more active at night than during the day, and they remain active in winter. They inhabit farmland as well as woodlands and swamps. Hedgerows and brushy field borders provide good habitat. Areas in the vicinity of water seem to be preferred. They make their dens in rock piles, under tree stumps, in sawmill slab piles, in abandoned rodent burrows, or occasionally in barns or other buildings. They sometimes store food in these dens and make nests from the fur, feathers, bones, and other remains of prey.

Reproduction. Most female weasels produce one annual litter of four to eight young. Ovulation is induced by copulation. Delayed implantation of the fertilized eggs occurs in many species. For example, embryos of the long-tailed weasel do not become implanted in the wall of the uterus for many months. Thus, whereas breeding normally occurs in July or August, the young are not born until April or early May of the following year. The embryos are implanted only 21 to 28 days before birth, even though the total gestation period may range from 205 to 337 days (average 279 days). Least weasels do not experience delayed implantation. They have a total gestation of approximately 35 days and may have up to 3 litters a year.

Weasels have numerous enemies including snakes, hawks, owls, foxes, cats, and humans. Like skunks, weasels possess well-developed anal scent glands that discharge a foul-smelling liquid called musk when threatened or attacked. Some long-tailed weasels have lived in the wild up to three years. Captives have lived at least five years. *See* CARNIVORA.

Donald W. Linzey

Bibliography. D. W. Linzey, *The Mammals of Virginia*, The McDonald & Woodward Publishing Company, 1998; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford Limited, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., The Johns Hopkins University Press, 1999.

Weather

The state of the atmosphere, as determined by the simultaneous occurrence of several meteorological phenomena at a geographical locality or over broad areas of the Earth. When such a collection of weather elements is part of an interrelated physical structure of the atmosphere, it is termed a weather system, and includes phenomena at all elevations above the ground. More popularly, weather refers to a certain state of the atmosphere as it affects humans' activities on the Earth's surface. In this sense, it is often taken to include such related phenomena as waves at sea and floods on land.

An orderly association of weather elements accompanying a typical weather system of the Northern Hemisphere may be illustrated by a large anticyclone, or high-pressure region. In such a "high," extending over an area of hundreds of square miles, the usually gentle winds circulate clockwise around the high-pressure center. This system often brings fair weather locally, which implies a bright sunny day with few clouds. The temperature may vary widely depending on season and time of day. However, a cyclone or low-pressure region is frequently associated with a dark cloudy sky with driving rain (or snow) and strong winds which circulate counterclockwise about a low-pressure center of the Northern Hemisphere.

A weather element is any individual physical feature of the atmosphere. At a given locality, at least seven such elements may be observed at any one time. These are clouds, precipitation, temperature, humidity, wind, pressure, and visibility. Each principal element is divided into many subtypes. For a discussion of a characteristic local combination of several elements, as they might be observed at a U.S. Weather Bureau station, *See* WEATHER MAP.

The various forms of precipitation are included by international agreement among the hydrometeors, which comprise all the visible features in the atmosphere, besides clouds, that are due to water in its various forms. For convenience in processing weather data and information, this definition is made to include some phenomena not due to water, such as dust and smoke. Some of the more common hydrometeors include rain, snow, fog, hail, dew, and frost.

Both a physical (or genetic) and a descriptive classification of clouds and hydrometeors have been devised. The World Meteorological Organization, which among many other activities coordinates the taking of weather observations among the nations of the world, recognizes at least 36 cloud types and 100 classes of hydrometeors.

Certain optical and electrical phenomena have long been observed among weather elements, including lightning, aurora, solar or lunar corona, and

halo. *See* AIR MASS; ATMOSPHERE; ATMOSPHERIC GENERAL CIRCULATION; CLOUD; FRONT; METEOROLOGY; PRECIPITATION (METEOROLOGY); STORM; WEATHER OBSERVATIONS; WIND. Philip F. Clapp

Bibliography. E. Aguado, J. E. Burt, and J. Burt, *Understanding Weather and Climate*, 3d ed., 2003; C. D. Ahrens, *Meteorology Today With Infotrac: An Introduction to Weather, Climate, and the Environment*, 6th ed., 1999; R. C. Barry and R. J. Chorley, *Atmosphere, Weather, and Climate*, 8th ed., 2003; W. R. Cotton and R. A. Pielke, *Human Impacts on Weather and Climate*, 1995.

Weather forecasting and prediction

Processes for formulating and disseminating information about future weather conditions based upon the collection and analysis of meteorological observations. Weather forecasts may be classified according to the space and time scale of the predicted phenomena. Atmospheric fluctuations with a length of less than 100 m (330 ft) and a period of less than 100 s are considered to be turbulent. Prediction of turbulence extends only to establishing its statistical properties, insofar as these are determined by the thermal and dynamic stability of the air and by the aerodynamic roughness of the underlying surface. The study of atmospheric turbulence is called micrometeorology; it is of importance for understanding the diffusion of air pollutants and other aspects of the climate near the ground. Standard meteorological observations are made with sampling techniques that filter out the influence of turbulence. Common terminology distinguishes among three classes of phenomena with a scale that is larger than the turbulent microscale: the mesoscale, synoptic scale, and planetary scale.

The mesoscale includes all moist convection phenomena, ranging from individual cloud cells up to the convective cloud complexes associated with prefrontal squall lines, tropical storms, and the intertropical convergence zone. Also included among mesoscale phenomena are the sea breeze, mountain valley circulations, and the detailed structure of frontal inversions. Because most mesoscale phenomena have time periods less than 12 h, they are little influenced by the rotation of the Earth. The prediction of mesoscale phenomena is an area of active research. Most forecasting methods depend upon empirical rules or the short-range extrapolation of current observations, particularly those provided by radar and geostationary satellites. Forecasts are usually couched in probabilistic terms to reflect the sporadic character of the phenomena. Since many mesoscale phenomena pose serious threats to life and property, it is the practice to issue advisories of potential occurrence significantly in advance. These "watch" advisories encourage the public to attain a degree of readiness appropriate to the potential hazard. Once the phenomenon is considered to be imminent, the advisory is changed to a "warning," with the expectation that the public will take immediate action to prevent the loss of life.

The next-largest scale of weather events is called the synoptic scale, because the network of meteorological stations making simultaneous, or synoptic, observations serves to define the phenomena. The migratory storm systems of the extratropics are synoptic-scale events, as are the undulating wind currents of the upper-air circulation which accompany the storms. The storms are associated with barometric minima, variously called lows, depressions, or cyclones. The sense of the wind rotation about the storm is counterclockwise in the Northern Hemisphere, but clockwise in the Southern Hemisphere. This effect, called geostrophy, is due to the rotation of the Earth and the relatively long period, 3–7 days, of the storm life cycle. Significant progress has been made in the numerical prediction of synoptic-scale phenomena.

Planetary-scale phenomena are persistent, quasi-stationary perturbations of the global circulation of the air with horizontal dimensions comparable to the radius of the Earth. These dominant features of the general circulation appear to be correlated with the major orographic features of the globe and with the latent and sensible heat sources provided by the oceans. They tend to control the paths followed by the synoptic-scale storms, and to draw upon the synoptic transients for an additional source of heat and momentum. Long-range weather forecasts must account for the slow evolution of the planetary-scale circulations. To the extent that the planetary-scale centers of action can be correctly predicted, the path and frequency of migratory storm systems can be estimated. The problem of long-range forecasting blends into the question of climate variation. *See* ATMOSPHERE; METEOROLOGICAL INSTRUMENTATION; MICROMETEOROLOGY; WEATHER OBSERVATIONS. Joseph P. Gerrity

The synoptic method of forecasting consists of the simultaneous collection of weather observations, and the plotting and analysis of these data on geographical maps. An experienced analyst, having studied several of these maps in chronological succession, can follow the movement and intensification of weather systems and forecast their positions. This forecasting technique requires the regular and frequent use of large networks of data. *See* WEATHER MAP.

The synoptic forecasts are limited because they are almost exclusively based upon surface observations. The stratosphere, a layer of air generally located above levels from 5.5 to 8 mi (9 to 13 km) and characterized by temperatures remaining constant or increasing with elevation, acts as a cap on much of the world's weather. A trackable instrument suspended beneath a rubber balloon (the radiosonde) is routinely deployed from a network of stations. The tracking of the balloon allows for wind measurements to be taken at various elevations throughout its ascent. Such a network of more than 2000 upper-air stations around the world reports its observations twice daily. Satellite measurements and aircraft

observations supplement the radiosonde network and are utilized to improve the analysis and tracking of weather systems.

The basic laws of hydrodynamics and thermodynamics can be used to define the current state of the atmosphere, and to compute its future state. This concept marked the birth of the dynamic method of forecasting, though its practical application was realized only after 1950, when sufficiently powerful computers became available to produce the necessary calculations efficiently. The first conceptual model of the life cycle of a surface cyclone showed that cyclonic storms typically formed along fronts, that is, boundaries between cold and warm air masses. The model showed that precipitation is associated with active cold and warm fronts, and that this precipitation wraps around the cyclone as it intensifies. Though subsequent work demonstrated that fronts are not crucial to cyclogenesis, the Norwegian frontal cyclone model is still used in weather map analyses. *See* CYCLONE; FRONT.

It was not until the late 1930s that routine aerological observations had a substantial impact on forecasting. By the 1950s, qualitatively reasonable numerical forecasts of the atmosphere were produced. Numerical weather prediction techniques, in addition to being applied to short-range weather prediction, are used in such research studies as air-pollutant transport and the effects of greenhouse gases on global climate change. *See* AIR POLLUTION; GREENHOUSE EFFECT; JET STREAM; UPPER-ATMOSPHERE DYNAMICS.

Though numerical forecasts continue to improve, statistical forecast techniques, once used exclusively with observational data available at the time of the forecast, are now used in conjunction with numerical output to predict the weather. Statistical methods, based upon a historical comparison of actual weather conditions with large samples of output from the same numerical model, routinely play a role in the prediction of surface temperatures and precipitation probabilities.

The recognition that small, barely detectable differences in the initial analysis of a forecast model often lead to very large errors in a 12–48-h forecast has led to experiments with ensemble forecasting. This method uses results from several numerical forecasts to produce the statistical mean and standard deviation of the forecasts. Success with ensemble forecasting suggests it can be a useful tool in enhancing prediction skill and in assessing the atmosphere's predictability. *See* DYNAMIC METEOROLOGY.

The best weather forecasts result from application of the synoptic method to the latest numerical and statistical information. The forecaster has an ever-increasing number of valuable tools with which to work. Numerical forecast models, owing to faster computers, are capable of explicitly resolving mesoscale weather systems. Geostationary satellites allow for continuous tracking of such dangerous weather systems as hurricanes. The increased routine use of Doppler radar, automated commercial aircraft observations, and use of data derived from wind and temperature profiler soundings promise to give added capability in tracking and forecasting mesoscale weather disturbances. Very high-frequency and ultrahigh-frequency Doppler radars may be used to provide detailed wind soundings. These wind profilers, if located sufficiently close to one another, will allow for the hourly tracking of mesoscale disturbances aloft. Ground- and satellite-based microwave radiometric measurements are being used to construct temperature and moisture soundings of the atmosphere. The assimilation of such data at varying times in the numerical model forecast cycle offers the promise of improved prediction and improved utilization of data and products by forecasters assisted by increasingly powerful interactive computer systems. Medium-range forecasts, ranging up to 2 weeks, may be improved from knowledge of forecast skill in relationship to the form of the planetary-scale atmospheric circulation. *See* DOPPLER RADAR; MESOMETEOROLOGY; METEOROLOGICAL RADAR; METEOROLOGICAL ROCKET; METEOROLOGICAL SATELLITES; RADAR METEOROLOGY; SATELLITE METEOROLOGY.

John R. Gyakum

Numerical Weather Prediction

Numerical weather prediction is the prediction of weather phenomena by the numerical solution of the equations governing the motion and changes of condition of the atmosphere.

The laws of motion of the atmosphere may be expressed as a set of partial differential equations relating the temporal rates of change of the meteorological variables to their instantaneous distribution in space. These equations are developed in dynamic meteorology. In principle, a prediction for a finite time interval can be obtained by summing a succession of infinitesimal time changes of the meteorological variables, each of which is determined by their distribution at a given instant of time. However, the nonlinearity of the equations and the complexity and multiplicity of the data make this process impossible in practice. Instead, it is necessary to resort to numerical approximation techniques in which successive changes in the variables are calculated for small but finite time intervals over a domain spanning part or all of the atmosphere. Even so, the amount of computation is vast, and numerical weather prediction remained only a dream until the advent of the modern computer.

The accuracy of numerical weather prediction depends on (1) an understanding of the physical laws of atmospheric behavior; (2) the ability to define through observations and analysis the state of the atmosphere at the initial time of the forecast; and (3) the accuracy with which the solutions of the continuous equations describing the rate of change of atmospheric variables are approximated by numerical means. The greatest success has been achieved in predicting the motion of the large-scale (>1000 mi or 1600 km) pressure systems in the atmosphere for relatively short periods of time (1–5 days). For such space and time scales, the poorly understood energy sources and frictional dissipative forces may

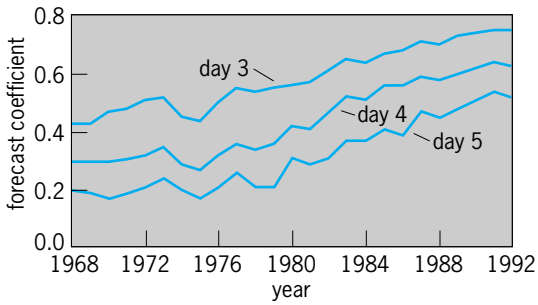

Fig. 1. Correlation coefficients between 3-day, 4-day, and 5-day forecasts of sea-level pressure made by the National Meteorological Center and observed sea-level pressure. Domain of verification is North America and adjacent ocean areas. (*Courtesy Wayman Baker, National Meteorological Center*)

be approximated by relatively simple formulations, and rather coarse horizontal resolutions (60–120 mi or 100–200 km) may be used.

Great progress has been made in improving the accuracy of numerical weather prediction models. Forecasts for 3, 4, and 5 days have steadily improved; the 5-day forecast is now as accurate as the 3-day forecast was in the early 1970s (**Fig. 1**). Because of the increasing accuracy of numerical forecasts, numerical models have become the basis for medium-range (1–10 days) forecasts made by the weather services of most countries. *See* MESOMETEOROLOGY.

Cloud and precipitation prediction. If, to the standard dynamic variables, the density of water vapor is added, it becomes possible to predict clouds and precipitation in addition to the air motion. When a parcel of air containing a fixed quantity of water vapor ascends, it expands adiabatically and cools until it becomes saturated. Continued ascent produces clouds and precipitation. The most successful predictions made by this method are obtained in regions of strong rising motion, whether induced by forced orographic ascent or by horizontal convergence in well-developed cyclones. The physics and mechanics of the convective cloud-formation process make the prediction of convective cloud and showery precipitation more difficult.

Global prediction. In 1955 the first operational numerical weather prediction model was introduced at the National Meteorological Center. This simplified barotropic model consisted of only one layer, and therefore it could model only the temporal variation of the mean vertical structure of the atmosphere. By the early 1990s, the speed of computers had increased sufficiently to permit the development of multilevel (usually about 10–20) models that could resolve the vertical variation of the wind, temperature, and moisture. These multilevel models predict the fundamental meteorological variables for large scales of motion. Global models with horizontal resolutions as fine as 125 mi (200 km) are being used by weather services in several countries.

Global numerical weather prediction models require the most powerful computers to complete a 10-day forecast in a reasonable amount of time. For example, a 10-day forecast with a 15-layer global model with horizontal resolution of 90 mi (150 km) requires approximately 10^{12} calculations. A supercomputer capable of performing 10^8 arithmetic operations per second would then require 10^4 s or 2.8 h to complete a single 10-day forecast.

Numerical models of climate. While global models were being implemented for operational weather prediction 1–10 days in advance, similar research models were being developed that could be applied for climate studies by running for much longer time periods. The extension of numerical predictions to long time intervals (many years) requires a more accurate numerical representation of the energy transfer and turbulent dissipative processes within the atmosphere and at the air-earth boundary, as well as greatly augmented computing-machine speeds and capacities.

With state-of-the-art computers, it is impossible to run global climate models with the same high resolution as numerical weather prediction models; for example, a hundred-year climate simulation with the numerical weather prediction model that requires 2.8 h for a 10-day forecast would require 10,220 h (1.2 years) of computer time. Therefore, climate models must be run at lower horizontal resolutions than numerical weather prediction models (typically 250 mi or 400 km).

Predictions of mean conditions over the large areas resolvable by climate models are feasible because it is possible to incorporate into the prediction equations estimates of the energy sources and sinks—estimates that may be inaccurate in detail but generally correct in the mean. Thus, long-term simulations of climate models with coarse horizontal resolutions have yielded simulations of mean circulations that strongly resemble those of the atmosphere. These simulations have been useful in explaining the principal features of the Earth's climate, even though it is impossible to predict the daily fluctuations of weather for extended periods. Climate models have also been used successfully to explain paleoclimatic variations, and are being applied to predict future changes in the climate induced by changes in the atmospheric composition or characteristics of the Earth's surface due to human activities. *See* CLIMATE HISTORY; CLIMATE MODIFICATION.

Limited-area models. Although the relatively coarse grids in global models are necessary for economical reasons, they are sources of two major types of forecast error. First, the truncation errors introduced when the continuous differential equations are replaced with approximations of finite resolution cause erroneous behavior of the scales of motion that are resolved by the models. Second, the neglect of scales of motion too small to be resolved by the mesh (for example, thunderstorms) may cause errors in the larger scales of motion. In an effort to simultaneously reduce both of these errors, models with considerably finer meshes have been tested. However, the price of reducing the mesh has been the necessity of covering smaller domains in order to keep the total computational effort within computer capability. Thus the main operational limited-area model run at the National Meteorological Center has a mesh length of

approximately 50 mi (80 km) on a side and covers a limited region approximately two times larger than North America. Because the side boundaries of this model lie in meteorologically active regions, the variables on the boundaries must be updated during the forecast. A typical procedure is to interpolate these required future values on the boundary from a coarse-mesh global model that is run first. Although this method is simple in concept, there are mathematical problems associated with it, including overspecification of some variables on the fine mesh. Nevertheless, limited-area models have made significant improvements in the accuracy of short-range numerical forecasts over the United States.

Even the small mesh sizes of the operational limited-area models are far too coarse to resolve the detailed structure of many important atmospheric phenomena, including hurricanes, thunderstorms, sea- and land-breeze circulations, mountain waves, and a variety of air-pollution phenomena. Considerable effort has gone into developing specialized research models with appropriate mesh sizes to study these and other small-scale systems. Thus, fully three-dimensional hurricane models with mesh sizes of 5 mi (8 km) simulate many of the features of real hurricanes. On even smaller scales, models with horizontal resolutions of a few hundred meters reproduce many of the observed features in the life cycle of thunderstorms and squall lines. It would be misleading, however, to imply that models of these phenomena differ from the large-scale models only in their resolution. In fact, physical processes that are negligible on large scales become important for some of the phenomena on smaller scales. For example, the drag of precipitation on the surrounding air is important in simulating thunderstorms, but not for modeling large scales of motion. Thus the details of precipitation processes, condensation, evaporation, freezing, and melting must be incorporated into realistic cloud models.

In another class of special models, chemical reactions between trace gases are considered. For example, in models of urban photochemical smog, predictive equations for the concentration of oxides of nitrogen, oxygen, ozone, and reactive hydrocarbons are solved. These equations contain transport and diffusion effects by the wind as well as reactions with solar radiation and other gases. Such air-chemistry models become far more complex than atmospheric models as the number of constituent gases and permitted reactions increases. *See* COMPUTER PROGRAMMING; DIGITAL COMPUTER; MODEL THEORY; SUPERCOMPUTER.

Richard A. Anthes

Modern weather analysis and forecasting procedures involve (1) the collection of global meteorological surface and upper-air observations in place (for example, ships of opportunity) and remotely sensed (for example, satellite sensors) instrumentation platforms; (2) the preparation of global surface and upper air pressure, temperature, moisture, and wind analyses at frequent time intervals based upon these observations; (3) the solution of a closed set of highly nonlinear equations governing atmospheric dynamical motions by numerical means using the most sophisticated computers to predict the future state of the global atmosphere from the initial conditions specified by the analyses in step 2; (4) the application of statistical procedures such as model output statistics (MOS) to the atmospheric simulations obtained from the dynamical models to predict a wide variety of weather elements of interest to potential users by objective means; and (5) the monitoring the intervention by experienced human forecasters at appropriate stages in this objective forecasting process to improve the quality and breadth of the forecast information communicated to a wide spectrum of users in the private and public sectors.

Data collection and analysis. Surface meteorological observations are routinely collected from a vast continental data network, with the majority of these observations obtained from the middle latitudes of both hemispheres. Commercial ships of opportunity, military vessels, and moored and drifting buoys provide similar in-place measurements from oceanic regions, although the data density is biased toward the principal global shipping lanes. Information on winds, pressure, temperature, and moisture throughout the troposphere and into the stratosphere is routinely collected from (1) balloon-borne instrumentation packages (radiosonde observations) and commercial and military aircraft which sample the free atmosphere directly; (2) ground-based remote-sensing instrumentation such as wind profilers (vertically pointing Doppler radars), the National Weather Service Doppler radar network, and lidars; and (3) special sensors deployed on board polar orbiting or geostationary satellites. The remotely sensed observations obtained from meteorological satellites have been especially helpful in providing crucial measurements of areally and vertically averaged temperature, moisture, and winds in data sparse (mostly oceanic) regions of the world. Such measurements are necessary to accommodate modern numerical weather prediction practices and to enable forecasters to continuously monitor global storm (such as hurricane) activity. *See* LIDAR; METEOROLOGICAL INSTRUMENTATION; RADAR METEOROLOGY.

At major operational weather prediction centers such as the National Center for Environmental Prediction (NCEP, formerly known as the National Meteorological Center or the European Centre for Medium Range Weather Forecasts, ECMWF), the global meteorological observations are routinely collected, quality-checked, and mapped for monitoring purposes by humans responsible for overseeing the forecast process. The basic observations collected by NCEP are also disseminated on-line by Internet Data Distribution (IDD) or in mapped form by digital facsimile (DIFAX) to prospective users in academia and the public and private sectors. At NCEP and ECMWF (and other centers) the global observational data stream is further machine-processed in order to prepare a full three-dimensional set of global surface and upper air analyses of selected meteorological fields at representative time periods. The typical horizontal and vertical resolution of the globally gridded analyses is about 60 mi (100 km) and 25–50 hectapascals (millibars), respectively. Preparation of these

analyses requires a first guess field from the previous numerical model forecast against which the updated observations are quality-checked and objectively analyzed to produce an updated global gridded set of meteorological analyses.

These updated analyses are modified as part of a numerical procedure designed to ensure that the gridded meteorological fields are dynamically consistent and suitable for direct computation in the new forecast cycle. This data assimilation, analysis, and initialization procedure, known as four-dimensional data assimilation, was at one time performed four times daily at the standard synoptic times of 0000, 0600, 1200, and 1800 UTC. The four-dimensional data assimilation is performed almost continuously, given that advanced observational technologies (such as wind profilers; automated surface observations; automated aircraft-measured temperature, moisture, and wind observations) have ensured the availability of observations at other than the standard synoptic times mentioned above. An example is the rapid update cycle (RUC) mesoscale analysis and forecast system used at NCEP to produce real-time surface and upper air analyses over the United States and vicinity every 3 h and short-range (out to 12 h) forecasts. The rapid update cycle system takes advantage of high-frequency data assimilation techniques to enable forecasters to monitor rapidly evolving mesoscale weather features.

Operational models. The forecast models used at NCEP and ECMWF and the other operational prediction centers are based upon the primitive equations that govern hydrodynamical and thermodynamical processes in the atmosphere. The closed set of equations consists of the three momentum equations (east-west, north-south, and the vertical direction), a thermodynamic equation, an equation of state, a continuity equation, and an equation describing the hydrological cycle. Physical processes such as the seasonal cycle in atmospheric radiation, solar and long-wave radiation, the diurnal heating cycle over land and water, surface heat, moisture and momentum fluxes, mixed-phase effects in clouds, latent heat release associated with stratiform and convective precipitation, and frictional effects are modeled explicitly or computed indirectly by means of parametrization techniques. A commonly used vertical coordinate in operational prediction models is the sigma coordinate, defined as the ratio of the pressure at any point in the atmosphere to the surface (station) pressure.

The NCEP operational global prediction model is known as the medium-range forecast (MRF) model. The medium-range forecast is run once daily out to 14 days from the 0000 UTC analysis and initialization cycle in support of medium-range and extended-range weather prediction activities at NCEP. The medium-range forecast run commences shortly before 0900 UTC in order to allow sufficient time for delayed 0000 UTC surface and upper air data to reach NCEP and to be included in the analysis and initialization cycle.

The aviation forecast model (AVN) is run from the 1200 UTC analysis and initialization cycle in support of NCEP's global aviation responsibilities. The aviation forecast model and medium-range forecast are identical models except that the AVN is run with a data cut off of about 3 h and only out to 72 h in order to minimize the time necessary to get the model forecast information to the field. The MRF runs with a vertical resolution of 38 levels, a horizontal resolution of about 63 mi (105 km) [triangular truncation in spectral space] for the first 8 days, and a degraded horizontal resolution of about 130 mi (210 km) for the last 6 days of the 14-day operational forecast.

The medium-range forecast is also the backbone of the ensemble forecasting effort (a series of parallel medium-range forecast runs with slightly changed initial conditions and medium-range forecast runs from different time periods) that was instituted in the early 1990s to provide forecaster guidance in support of medium-range forecasting (6–10-day) activities at NCEP. The scientific basis for ensemble prediction is that model forecasts for the medium and extended range should be considered stochastic rather than deterministic in nature in recognition of the very large forecast differences that can occur on these time scales between two model runs initialized with only very slight differences in initial conditions. Continuing advances in computer power have made it practical to begin operational ensemble numerical weather prediction. Given that each member of the ensemble (the ensemble consists of 38 different model runs) represents an equally likely model forecast outcome, the spread of the ensemble forecasts is taken as a measure of the potential skill of the forecasts. Forecasters look for clustering of the ensemble members around a particular solution as indicative of higher probability outcomes as part of a process inevitably known as "forecasting the forecasts". Ensemble forecasting techniques are also taking hold at other operational centers such as the ECMWF, and they are likely to spread to regional models such as the NCEP Eta and Regional Analysis and Forecast (RAFS) models in the nea future.

Regional forecast models such as the NCEP Eta model are utilized to help predict atmospheric circulation patterns associated with convective processes and terrain forcing. Given the increasing importance of forecasting significant mesoscale weather events (such as squall lines, flash flood occurrences, and heavy snow bands) that occur on time scales of a few hours and space scales of a few hundred kilometers, the Eta model will continue to be developed at NCEP in support of regional model guidance to local forecasters. An important challenge to mesoscale models is to make better precipitation forecasts, especially of significant convective weather events. Success in this endeavor will require increased use of satellite- and land-based data sets, especially measurements of water vapor, that are coming on line.

Forecast products and forecast skill. These are classified as longer term (greater than 2 weeks) and shorter term.

Longer term. The Climate Prediction Center (CPC) of NCEP initiated a long-lead climate outlook program for the United States in early 1995. The scientific basis for this initiative rests upon the understanding

of the coupled nature of global atmospheric and oceanic circulations on intraseasonal and interannual time scales. Particular attention has been focused on the climatic effects arising from low-frequency changes in global atmospheric wind, pressure, and moisture patterns associated with evolving tropical oceanic sea surface temperature and thermocline depth anomalies known as part of the El Niño–Southern Oscillation (ENSO) climate signal. Statistically significant temperature and moisture anomalies have been identified over portions of the conterminous United States (and in many other regions of the world) that can occur in association with an ENSO event, especially during the cooler half of the year. For example, in winter there is a tendency for the southern United States from Texas to the Carolinas, Georgia, and Florida to be wetter and cooler than normal, while much of the Pacific Northwest eastward across southern Canada and the northern United States to the Great Lakes is warmer and drier than normal during the warm phase (anomalies in sea surface temperature in the tropical Pacific from the Dateline eastward to South America are positive, while sea-level pressure anomalies are positive and negative, respectively, over the western versus central and eastern Pacific) of an ENSO event. *See* EL NIÑO; TROPICAL METEOROLOGY.

Similarly, the tropical convective activity that usually maximizes over the western Pacific oceanic warm pool (sea surface temperature about 86°F or 30°C) tends to shift eastward toward the central and, occasionally, eastern Pacific during the warm phase of an ENSO event. Under such conditions, drought and heat-wave conditions may prevail over portions of Australia and Indonesia, while conditions become much wetter in the normally arid coastal areas along western Peru and northern Chile. During the exceptionally strong ENSO event of 1982–1983, extreme climatic anomalies were observed at many tropical and midlatitude locations around the world. During the cold phase of ENSO (known as La Niña) when anomalies in the sea surface temperature in the tropical eastern Pacific are negative and the easterly trade winds are stronger than average, many of these global climatic anomalies may reverse sign, such as occurred during the 1988–1989 cold event. Increased knowledge and understanding of the global aspects of atmospheric and oceanic low-frequency phenomena have resulted from combined observational, theoretical, and numerical research investigations that in part were stimulated by the extraordinary warm ENSO event of 1982–1983. The knowledge is now being applied toward the preparation of long-lead seasonal forecasts at the Climate Prediction Center and elsewhere. *See* DROUGHT.

In the extended forecasting procedure of the Climate Prediction Center, a total of 26 probability anomaly maps are prepared (one each for temperature and precipitation for equally likely below-normal and above-normal categories) for each of 13 lead times from 0.5 to 12.5 months. The probabilistic format (begun in 1982 with the issuance of monthly outlooks of temperature and precipitation for the United States) and accompanying skill assessment of these long-lead forecasts allows sophisticated users in the public and private sectors to make better use of the outlooks for a variety of business and planning purposes. The long-lead climate outlooks are made from a combination of the following methods: (1) a statistical procedure known as canonical correlation analysis that relates prior anomalies of global sea surface temperatures, Northern Hemisphere 700-hPa (mbar) heights, and temperature and precipitation anomalies over the United States to subsequent anomalies; (2) another statistical procedure known as optimal climate normals that projects the persistence of recent temperature and precipitation anomalies to future months and seasons; and (3) forecasts of temperature and precipitation anomalies obtained from a coupled atmosphere-ocean general circulation model integrated forward in time 6 months by using selected sea surface temperature anomalies.

Shorter term. An important aspect of the forecasting process is the interpretation and dissemination of model forecasts to a wide variety of users and the general public. Specialized NCEP centers (in addition to the Climate Prediction Center) such as the National Severe Storms Forecasting Center (responsible for issuing watches and warnings of severe weather events), the Storm Prediction Center (charged with monitoring a broad range of winter and summer storms), the National Hurricane Center (responsible for monitoring the hurricane threat), and the Hydrological Meteorological Center (responsible for monitoring river basin water storage, runoff, and flood potential) are charged with providing guidance to local National Weather Service forecast offices about a variety of weather hazards. Experienced meteorologists and hydrologists at these centers are responsible for the interpretation of NCEP model and statistical forecasts so as to provide additional guidance and information on upcoming important weather events to individual offices of the National Weather Service. As an example, the Heavy Precipitation Branch of the NCEP has the responsibility of providing forecasts of significant rain and snow amount areas over the continental United States out to 48 h in advance. The experienced forecasters in the Heavy Precipitation Branch have been able to maintain their skill over the model objective precipitation forecasts since the mid-1970s, despite demonstrated model forecasting improvements, through a combination of detailed knowledge of regional and local mesometeorology, an understanding of systematic model biases and how to correct for them, and a continuing ability to select the best model on any given day in a form of human ensemble forecasting.

The skill of numerical weather prediction forecasts of temperature and winds in the free atmosphere [for example, the 500 hPa (mbar) level] continues to improve steadily. For example, since the mid-1980s the root-mean-square vector error of the 24-h 250-hPa (mbar) wind forecasts produced by the operational global NCEP model has declined from 20 mi/h to 16 mi/h (10 m s^{-1} to 7.5 m s^{-1}). Experience with medium-range-forecast/ECMWF day five predictions of mean sea-level

pressure [500 hPa (mbar) heights] over North America indicates that standardized anomaly correlation coefficient scores (a standardized anomaly is the anomaly divided by the climatological standard deviation and is used to avoid overweighting anomalies at higher latitudes relative to lower latitudes) range from 30 to 65 (experience suggests a standardized score >20 has operational forecast value) with the higher (lower) scores occurring in winter (summer). At NCEP the human forecasters have been able to improve upon the medium-range-forecast anomaly correlation coefficient scores for day five mean sea-level pressure forecasts.

Forecasts of surface weather elements such as temperature, precipitation type and amount, ceiling, and visibility are prepared objectively by a statistical-dynamical model output statistics (MOS) technique. The cornerstone of the MOS approach is the preparation of a series of regression equations for individual weather elements at specified locations, with the regression coefficients determined from predictors provided by operational NCEP models. MOS guidance is relayed to National Weather Service offices and a wide variety of external users. The MOS product has become very competitive with the human forecaster as measured by standard skill scores. Over the 25-year period ending in 1992, the mean absolute error of National Weather Service maximum and minimum temperatures issued to the general public decreased by about 1.08°F (0.6°C). To put this number in perspective, the skill of 36-h temperature forecasts issued in the late 1990s is comparable to the skill of similar 12-h forecasts made in the late 1960s. A similar "24-h gain" in skill level has been registered for routine 12-h probability-of-precipitation forecasts of the National Weather Service. Skill levels are a function of the forecast type and projection. Skill (as measured relative to a climatological control by using a standard skill measure such as the Brier score) in day-to-day temperature forecasting approaches zero by 7–8 days in advance. For probability-of-precipitation forecasts the zero skill level is typically reached 2–3 days earlier. Precipitation-amount forecasts typically show little skill beyond 3 days, while probability-of-thunderstorm forecasts may be skillful only out to 1–2 days ahead.

These varying skill levels reflect the fact that existing numerical prediction models such as the medium-range forecast have become very good at making large-scale circulation and temperature forecasts, but are less successful in making weather forecasts. An example is the prediction of precipitation amount and type given the occurrence of precipitation and convection. Each of these forecasts is progressively more difficult because of the increasing importance of mesoscale processes to the overall skill of the forecast. *See* PRECIPITATION (METEOROLOGY).

Lance F. Bosart

Nowcasting

Nowcasting is a form of very short range weather forecasting. The term nowcasting is sometimes used loosely to refer to any area-specific forecast for the period up to 12 h ahead that is based on very detailed observational data. However, nowcasting should probably be defined more restrictively as the detailed description of the current weather along with forecasts obtained by extrapolation up to about 2 h ahead. Useful extrapolation forecasts can be obtained for longer periods in many situations, but in some weather situations the accuracy of extrapolation forecasts diminishes quickly with time as a result of the development or decay of the weather systems.

Comparison with traditional forecasting. Much weather forecasting is based on the widely spaced observations of temperature, humidity, and wind obtained from a worldwide network of balloon-borne radiosondes. These data are used as input to numerical-dynamical weather prediction models in which the equations of motion, mass continuity, and thermodynamics are solved for large portions of the atmosphere. The resulting forecasts are general in nature. Although these general forecasts for one or more days ahead have improved in line with continuing developments of the numerical-dynamical models, there have not been corresponding improvements in local forecasts for the period up to 12 h ahead. The problem is that most of the mathematical models cope adequately with only the large weather systems, such as cyclones and anticyclones, referred to by meteorologists as synoptic-scale weather systems. *See* WEATHER.

Nowcasting involves the use of very detailed and frequent meteorological observations, especially remote-sensing observations, to provide a precise description of the "now" situation from which very short-range forecasts can be obtained by extrapolation. Of particular value are the patterns of cloud, temperature, and humidity which can be obtained from geostationary satellites, and the fields of rainfall and wind measured by networks of ground-based radars. These kinds of observations enable the weather forecaster to keep track of smaller-scale events such as squall lines, fronts and thunderstorm clusters, and various terrain-induced phenomena such as land-sea breezes and mountain-valley winds. Meteorologists refer to these systems as mesoscale weather systems because their scale is intermediate between the large or synoptic-scale cyclones and the very small or microscale features such as boundary-layer turbulence. *See* FRONT; METEOROLOGICAL SATELLITES; SQUALL LINE; STORM.

There is a distinction between traditional forecasting and nowcasting in terms of the lead time and quality of the forecast (**Fig. 2**). The two approaches are complementary; each has its place according to the lead time and detail required of the forecast. Conceptually, nowcasting is a simple procedure. However, vast amounts of data are involved, and it is only since the development of the necessary digital data-processing, transmission, and display facilities that it has been possible for nowcasting to be available economically. *See* DATA COMMUNICATIONS.

Information technology. A key problem in nowcasting is that of combining diverse and complex data streams, especially conventional meteorological data with remote-sensing data. It is widely held that this combination should be achieved by using digital

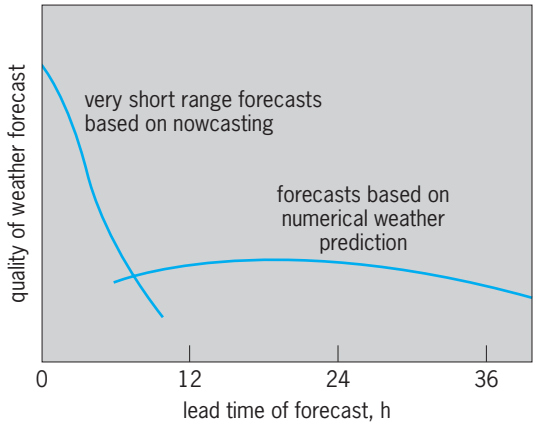

Fig. 2. The quality of weather forecasts, defined as the product of the accuracy and detail achievable, depicted schematically as a function of lead time for two different forecasting methods. Traditional forecasts based on numerical weather prediction deal with synoptic-scale events such as cyclones and anticyclones; very short range forecasts based on the nowcasting approach resolve so-called mesoscale events, such as thunderstorms, fronts, and squall lines.

data sets displayed on interactive video displays—the so-called workstation concept. By means of advanced human-computer interaction techniques, the weather forecaster will eventually be able to analyze the merged data sets by using a light pen or a finger on touch-sensitive television screens. Various automatic procedures can be implemented to help the forecaster carry out analyses and extrapolation forecasts, but the incomplete nature of the data sets is such that the forecaster will almost always be in the position of needing to fine-tune the products subjectively. The idea behind the forecasting workstation is to simplify the routine chores of basic data manipulation so that the forecaster is given the maximum opportunity to exercise judgment within the context of what is otherwise a highly automated system.

Very short-range forecast products are highly perishable: they must be disseminated promptly if they are not to lose their value. Advances in technology offer means for the rapid tailoring and dissemination of the digital forecast information and for presenting the material in convenient customer-oriented formats. In some cases, dissemination may be by direct computer-to-computer link with the user's control system. In another cases, the user will benefit from methods of visual presentation using such media as cable television or viewdata.

Nowcasting and mesoscale models. There is some evidence of a gap in forecasting capability between nowcasting and traditional synoptic-scale numerical weather prediction. This occurs for forecast lead times between about 6 and 12 h, that is, for periods when development and decay are beginning to invalidate forecasts by simple extrapolation (Fig. 2). To some extent the forecaster can identify some of the likely developments by interpreting the nowcast information in the light of local climatologies and conceptual life-cycle models of weather systems. But the best way to forecast changes is to use numerical-dynamical methods. It would be natural to assume that an immediate way would be to incorporate the detailed (but usually incomplete) nowcast data as input to numerical-dynamical models with a finer resolution than those presently in operational use; these are the so-called mesoscale numerical models. Research into the use of detailed observational data in mesoscale models has been actively pursued, but significant technical difficulties have slowed progress. Keith A. Browning

Extended-Range Prediction

Forecasts of time averages of atmospheric variables, for example, sea surface temperature, where the lead time for the prediction is more than 2 weeks, are termed long-range or extended-range climate predictions. The National Weather Service issues extended-range predictions of monthly and seasonal average temperature and precipitation near the middle of each month; these are known as climate outlooks. Monthly outlooks have a single lead time of 2 weeks. The seasonal outlooks have lead times ranging from 2 weeks to 12 $^1/_2$ months in 1-month increments. Thus, in each month a set consisting of 1 monthly and 13 seasonal long-lead-time climate outlooks is released.

Climate versus weather. When the daily average temperature (that is, the average of the high and the low temperatures for each day) is plotted as a graph for a given point, the result is a wildly oscillating curve (**Fig. 3**). Such a curve represents weather, the details of which are considered to be unpredictable beyond about 2 weeks. A curve for the 30-year average of the daily average temperature represents the climatology (Fig. 3). Finally, a 90-day running average can be applied to the curve that represents the daily average (the heaviest curve in Fig. 3); this represents the short-term seasonal climate behavior. The differences between this curve and the climatology represent the kind of temperature and, by analogy, the pressure predictions discussed in this section.

Scientific basis. Prior to the early 1980s, the techniques used to produce extended-range predictions assumed that the 700-mbar height field integrates the effects of long-lived atmospheric processes, which produce the observed 700-mbar anomalies. These signals can sometimes be isolated via linear statistical techniques, even though in general the physical causes of the anomalies are unknown. Predictions made by these techniques were slightly more accurate than climatology. During the 1980s, it was discovered that the ENSO produces, by far, the largest measurable short-term variability of middle-latitude climate of any low-frequency phenomenon known.

The name El Niño applies to the warm phase of the Southern Oscillation. On average, every 4–5 years the sea surface temperature in the central and eastern equatorial Pacific Ocean warms 1°C or more above its 30-year average. This is accompanied by weakening of the Pacific trade winds, a decrease of sea-level pressures near Tahiti in the eastern tropical Pacific, an increase of sea-level pressures in northern Australia, an increase in the depth of the oceanic thermocline in the tropical eastern Pacific Ocean (that is, anomalously warm water replaces climatologically cold water), and a shift of deep cumulus convection and rainfall from the western equatorial Pacific into

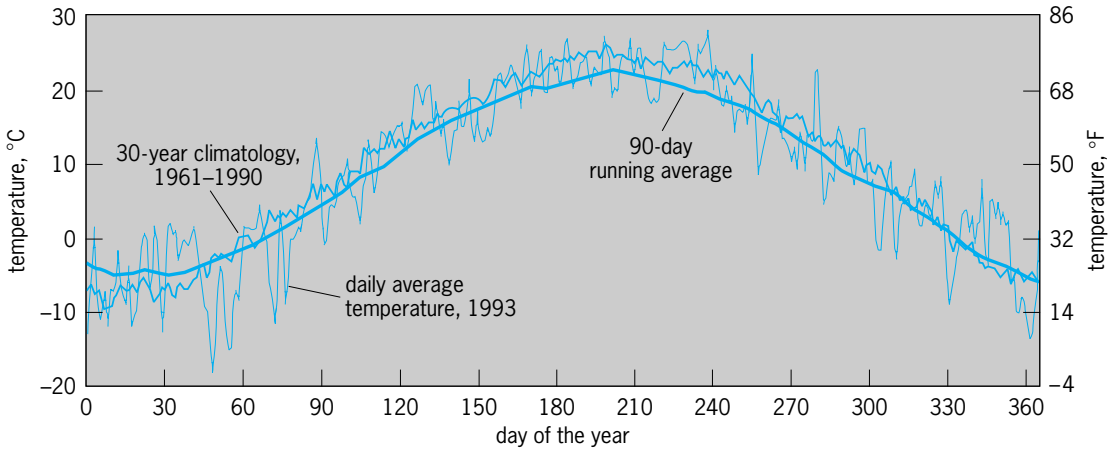

Fig. 3. Daily average temperature at Des Moines, Iowa. The thinnest curve indicates daily average temperature for January 1–December 31, 1994. The thickest curve indicates the 90-day running mean. The remaining curve indicates the 30-year average of daily average temperature, 1961–1990.

the central and eastern Pacific. The shift in deep convection alters the global atmospheric circulation, perturbing the climatological transports of heat and momentum. The warm phase often reaches this mature state by late December and normally lasts up to 18 months. This pattern is sometimes followed by a transition to the cold phase of the Southern Oscillation, although the periodicity is not particularly regular; for example, warm-phase conditions have on occasion lasted for three or more consecutive years (1939–1941, 1991–1995).

The anomalous tropospheric warming in the eastern tropical Pacific Ocean during the warm phase strengthens the Hadley circulation, including the subtropical jet. This energized subtropical jet contributes directly to a greater frequency and strength of winter and springtime storms over the southern United States, the effect sometimes including California, but it almost always brings enhanced precipitation to Texas and the Gulf states. The polar jet is also affected, often being constrained to lie north of its normal position over North America. This leads to warmer and drier conditions over sections of the northern United States. Opposite conditions tend to prevail during the cold phase of the Southern Oscillation, although the amplitude of the surface anomalies tends to be smaller than during the warm phase.

Models. The long life span and pronounced effect that ENSO has upon short-term climate variability, over the United States and elsewhere, clearly presents an opportunity to make predictions based upon the status of ENSO with a lead time of as much as a year. The statistical models traditionally used in long-range prediction have been reformulated to take advantage of this knowledge. This enhanced physical understanding has permitted, for the first time, the development and use in operational seasonal prediction of a physical model like those used for numerical weather prediction, which couples the ocean and the atmosphere. Such models should improve seasonal prediction. *See* ATMOSPHERIC GENERAL CIRCULATION; JET STREAM; TROPOSPHERE.

Physically based coupled model predictions of seasonal averages are achieved at the National Weather Service by using a global spectral medium-range forecast model modified to have improved response to sea surface temperature forcing the atmosphere in the tropics. Two 6-month runs of the model, driven by two different sets of sea surface temperature, are made. The first of the runs uses sea surface temperature persisting from a half month ago, while the second sea surface temperature fields are predicted by a three-member ensemble of coupled ocean/atmosphere model runs.

Three examples of objective empirical prediction methods are mixed persistence analog, canonical correlation analysis, and optimal climate normals. The mixed persistence analog consists of first determining, from 45 years of observations, the 10 years that are best analogs to one or more of the most recently observed seasonal means of United States surface temperature, the sea surface temperature, the average temperature of the lower 1.8 mi (3 km) of the atmosphere, and the seasonal mean 700-hPa (mbar) height-anomaly pattern. The mix of predictors used is based upon the results of forecasts made in prior years and varies with the season. A final analog is formed from a similarity-based weighted composite of these 10 individual analog years. The temperature pattern represented by this final analog is combined, via multiple linear regression, with the most recently observed seasonal mean temperature anomaly pattern (that is, persistence).

Canonical correlation analysis is a multivariate statistical technique that relates future spatial and temporal patterns of United States surface temperature or pressure to spatial and temporal patterns of observed seasonal mean tropical Pacific sea surface temperature, hemispheric 700-hPa (mbar) height, and United States surface temperature. Thus, predictor spatial patterns at different prior times can be related to predict and spatial patterns at different future times.

The optimal climate normals method is based upon the simple notion that it is possible to obtain a better estimate of the average state of the climate (seasonal mean temperature or pressure) for a future season than the standard 30-year climatology

by computing, for that season, the average of the most recent *k* years, where *k* is less than 30 years. For operational purposes the value of *k* is fixed at 10 years for temperature and 15 years for pressure. The optimal climate normals method isolates effects of trends or recent multiyear regimes in seasonal observed time series.

Accuracy. The accuracy of long-range outlooks has always been modest because the predictions must encompass a large number of possible outcomes, while the observed single event against which the outlook is verified includes the noise created by the specific synoptic disturbances that actually occur and that are unpredictable on monthly and seasonal time scales. According to some estimates of potential predictability, the noise is generally larger than the signal in middle latitudes.

An empirical verification method called cross-validation is used to evaluate most statistical models. In cross-validation the model being evaluated is formulated by using all observations except for one or more years, one of which is then predicted by the model. This is done, in turn, for all the years for which observations are available. The observations for each of the years withheld are then compared with their associated outlooks and the forecast accuracy is determined.

For coupled model forecasts, cross-validation is not used. Rather, the model must be rerun and verified for at least 10 years (cases) for a given season. These coupled model reruns, based upon observed sea surface temperatures, indicate that these forecasts of temperature have a cold-season United States average maximum correlation with observations of about 0.50 and a warm-season minimum score near zero. The average maximum score is reduced by about half when strong ENSO years are excluded. Coupled model precipitation forecasts over the United States have early winter and early spring accuracy maxima (correlations with observations of 0.30–0.40) when ENSO cases are included. Correlations between forecast and observed precipitation fields for this parameter also decline to about 0.10–0.15 when strong ENSO cases are excluded.

When forecasts of seasonal mean temperature made using the canonical correlation technique are correlated with observations, accuracy maxima in winter and summer and minima in spring and autumn are obtained. The score at lead times of 1, 4, and 7 months is nearly the same. Only at lead times of 10 and 13 months does the correlation drop to about half its value at the shorter lead times. As measured by the correlation between forecasts and observations, temperature forecasts made by the optimal climate normals technique have summer and winter maxima and minima in late winter–early spring and autumn. One interesting consequence of these temperature outlook accuracy properties is that, at certain times of the year, short-lead predictions may have lower expected accuracy than much longer lead predictions.

The accuracy of precipitation outlooks is generally lower than that of temperature outlooks. Precipitation predictions of the canonical correlation analysis and the optimal climate normals techniques have wintertime maximum correlations with observations. However, the former has low summertime accuracy, while the latter has relatively high accuracy during that season. For both methods, areas of high expected accuracy on the precipitation maps are much smaller and more spotty than those for temperature. This is due to the fact that precipitation is much more noisy (that is, variable on small space and time scales) than temperature.

Forecasts of seasonal mean temperature made by canonical correlation analysis and optimal climate normals have average correlations with observations in excess of 0.5 over large portions of the eastern United States during January through March and over the Great Basin during July through September.

Formulating an outlook. Forecasts of monthly and seasonal mean surface temperature and total precipitation, for the conterminous United States, Alaska, and Hawaii, are derived from a quasiobjective combination of canonical correlation analysis, optimal climate normals, and the coupled ocean-atmosphere model. For each method, maps consisting of standardized forecast values at stations are produced by computer. The physical size of the plotted numbers indicates the accuracy of the model. Two forecasters examine the map or maps for each lead time and variable, and resolve any conflicts among the methods. The statistical confidence in the resulting forecast will be low in regions where two or more equally reliable tools are in opposition or where the expected accuracy is low for all methods. Confidence is relatively high when tools agree and when the expected accuracy is high.

Information about accuracy and confidence is quantified through the use of probability anomalies to indicate the degree of confidence available from the tools. The climatological probabilities, with respect to which the outlook anomalies deviate, are determined well ahead of time. This is done by first dividing the observations of seasonal mean temperature and total precipitation during 1961–1990 into three equally likely categories for each of the stations at which forecasts are made. These are BELOW, NORMAL, and ABOVE for temperature, and BELOW, MODERATE, and ABOVE for precipitation. The likelihood of these categories is $^1/_3$-$^1/_3$-$^1/_3$ and is designated CL (for climatological probabilities) on the maps. Regions in which the correlation between the forecasts made by the tools and the observations, as determined by cross-validation, is less than 0.30 are given this designation. Elsewhere, nonzero probability anomalies—the numerical size of which is based upon the correlation scores of the tools, the degree of agreement among the tools, and the amplitude of the composite predicted anomalies—are assigned. Finally, the forecasters sketch the probability anomaly maps. *See* WEATHER MAP. Edward A. O'Lenic

Bibliography. C. D. Ahrens, *Meteorology Today With Infotrac: An Introduction to Weather, Climate, and the Environment*, 6th ed., 1999; K. A. Browning (ed.), *Nowcasting*, 1982; R. J. Doviak and D. S. Zrnic, *Doppler Radar & Weather Observations*, 2d ed., 1993; E. W. Friday, Jr., The modernization and associated restructuring of the National Weather

Service: An overview, *Bull. Amer. Meteorol. Soc.*, 75:43–52, 1994; J. R. Holton, *An Introduction to Dynamic Meteorology*, 3d ed., 1992; F. K. Lutgens, E. J. Tarbuck, and D. Tasa, *The Atmosphere: An Introduction to Meteorology*, 9th ed., 2003; F. Nebeker, *Calculating the Weather: Meteorology in the 20th Century*, 1995; P. Santurette and C. Georgiev, *Weather Analysis and Forecasting: Applying Satellite Water Vapor Imagery and Potential Vorticity Analysis*, 2005; S. Tracton and E. Kalnay, Operational ensemble prediction at the National Meteorological Center: Practical aspects, *Weath. Forecast.*, 8:379–398, 1993.

Weather map

A map or a series of maps that is used to depict the evolution and life cycle of atmospheric phenomena at selected times at the surface and in the free atmosphere. Weather maps are used for the analysis and display of in-place observational measurements and computer-generated analysis and forecast fields derived from weather and climate prediction models by research and operational meteorologists, government research laboratories, and commercial firms. Similar analyses derived from sophisticated computer forecast models are displayed in map form for forecast periods of 10–14 days in advance to provide guidance for human weather forecasters. *See* METEOROLOGICAL INSTRUMENTATION; WEATHER OBSERVATIONS.

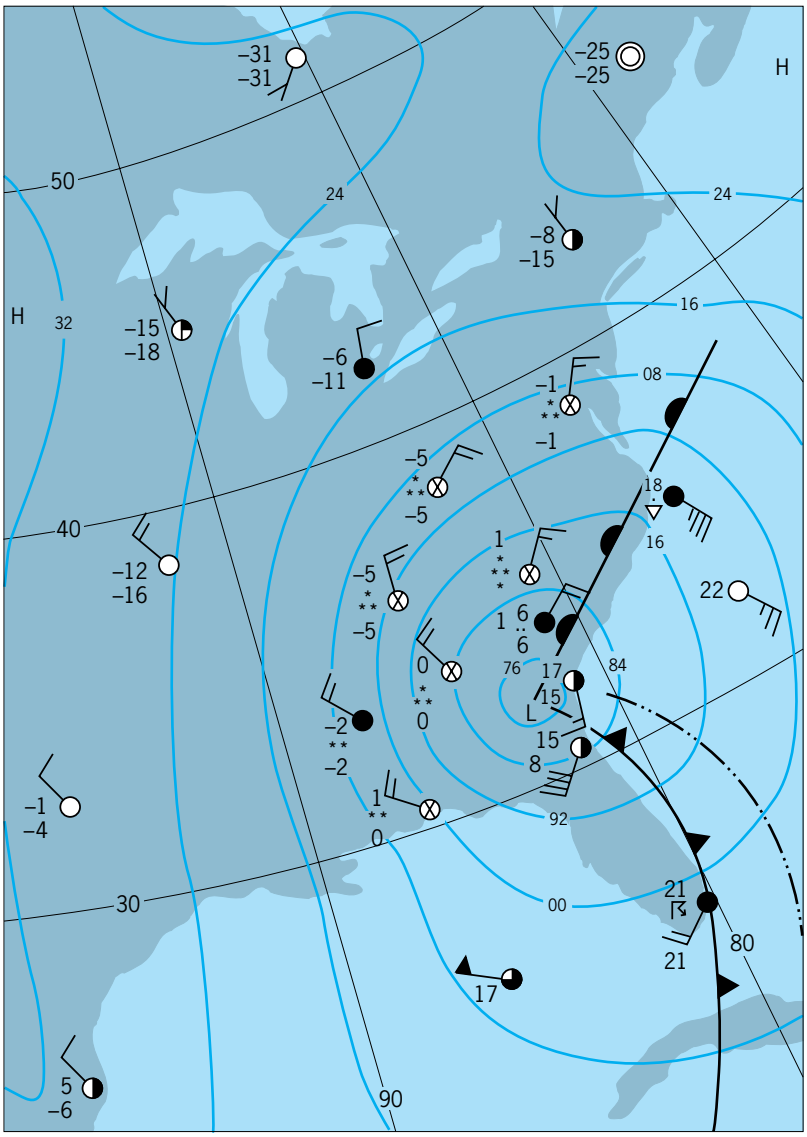

Sea-level map illustrating isobars (solid lines every 8 hectopascals or 8 millibars) for 1200 UTC 13 March 1993. Plotted numbers adjacent to each station denote temperature (top) and dewpoint temperature (bottom) in ○C. Present weather is shown by conventional symbols. Wind speed is measured in m s^{-1} (1 pennant = 25 m s^{-1}; full barb = 5 m s^{-1}; half barb = 2.5 m s^{-1}). Bold lines and dash–double dot line denote conventional fronts and squall line position.

A sea-level weather map was prepared for the "Storm of the Century" at 1200 UTC (7 a.m. EST) 13 March 1993 (see **illus.**). At this time an intense 973-hectopascal (973-millibar) surface. cyclone was centered over southeastern Georgia. Heavy snow was falling from the Florida panhandle to southern New York, while severe weather, including heavy thunderstorms with hail and tornadoes, was reported across many portions of the Florida peninsula equatorward across Cuba and into Central America. The wind pattern at 500 hPa (500 millibar) was especially favorable for a major storm along the coast of eastern North America as the surface cyclone was situated beneath a strengthening southerly flow between a deepening trough over the Gulf of Mexico and an intensifying ridge over New England and the Canadian Maritime Provinces. Winds in the jet stream near 300 hPa (300 millibar) over the eastern United States exceeded 175 knots (90 m s^{-1}) as the storm intensified rapidly.

Rapid advances in computer technology and visualization techniques, as well as the continued explosive growth of the Internet distribution of global weather observations, satellite and radar imagery, and model analysis and forecast fields, have revolutionized how weather, climate, and forecast data and information can be conveyed to both the general public and sophisticated users in the public and commercial sectors. People and organizations with access to the Internet can access weather and climate information in a variety of digital or map forms in support of a wind range of professional and personal activities. *See* CLIMATOLOGY; METEOROLOGICAL SATELLITES; METEOROLOGY; RADAR METEOROLOGY; WEATHER FORECASTING AND PREDICTION.

Lance F. Bosart

Bibliography. C. D. Ahrens, *Meteorology Today: An Introduction to Weather, Climate and the Environment*, 6th ed., 1999; J. Williams, *The USA Today Weather Almanac*, 1995.

Weather modification

Human influence on the weather and, ultimately, climate. This can be either intentional, as with cloud seeding to clear fog from airports or to increase precipitation, or unintentional, as with air pollution, which increases aerosol concentrations and reduces sunlight. Weather is considered to be the day-to-day variations of the environment—temperature, cloudiness, relative humidity, wind-speed, visibility, and

precipitation. Climate, on the other hand, reflects the average and extremes of these variables, changing on a seasonal basis. Weather change may lead to climate change, which is assessed over a period of years. *See* CLIMATE HISTORY.

Specific processes of weather modification are as follows: (1) Change of precipitation intensity and distribution result from changes in the colloidal stability of clouds. For example, seeding of supercooled water clouds with dry ice (solid carbon dioxide, CO_2) or silver iodide (AgI) leads to ice crystal growth and fall-out; layer clouds may dissipate, convective clouds may grow. (2) Radiation change results from changes of aerosol or clouds (deliberately with a smoke screen, or unintentionally with air pollution from combustion), from changes in the gaseous constituents of the atmosphere (as with carbon dioxide from fossil fuel combustion), and from changes in the ability of surfaces to reflect or scatter back sunlight (as replacing farmland by houses.) (3) Change of wind regime results from change in surface roughness and heat input, for example, replacing forests with farmland.

Ice phase and cloud seeding. Water, when present in clouds in the atmosphere as droplets about 10 micrometers in diameter, often supercools—that is, exists as a metastable liquid—to temperatures as low as −40°F (−40°C). Random motion of water molecules in the liquid leads to the formation of clusters with an ice configuration; below −40°F a cluster can quickly grow to freeze the whole droplet (homogeneous nucleation). The vapor pressure over supercooled water is greater than over ice at the same temperature (**Fig. 1**), leading to preferential growth of ice crystals. A small ice crystal introduced into such a supercooled cloud grows to become a visible snow crystal about 1 mm in diameter in a few minutes, and is sufficiently large to fall out as precipitation (**Fig. 2**). Under natural conditions, ice crystals are only rarely found in growing clouds with temperature entirely above 14°F (−10°C). Below this temperature, ice crystals are found with increasing frequency as the temperature lowers. Here insoluble impurities (minerals, silver iodide, bacteria) nucleate ice at temperatures as high as 14 to 23°F (−10 to −5°C) and with increasing frequency as the temperature lowers to −40°F.

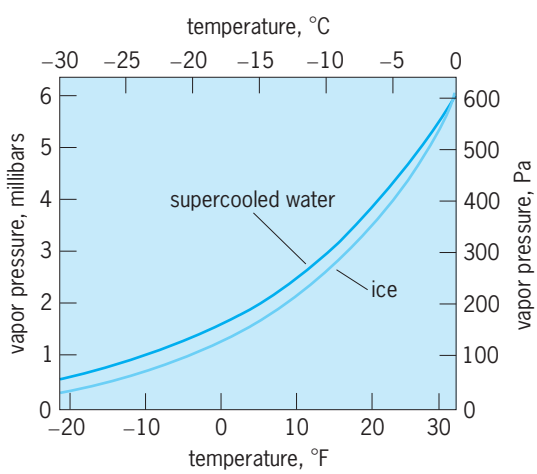

Fig. 1. Change with temperature of the vapor pressure of water vapor in equilibrium with supercooled water and ice.

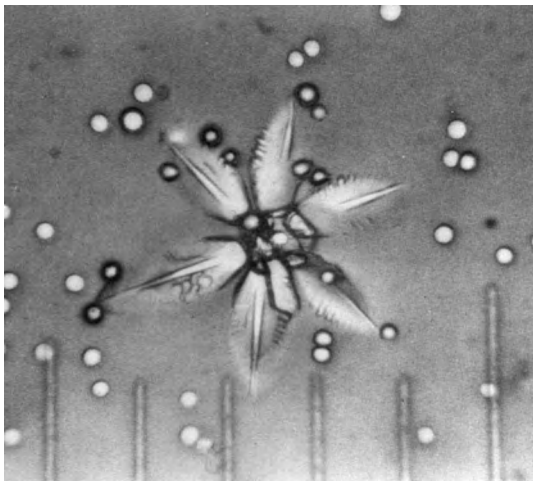

Fig. 2. Replica of cloud droplets and an ice crystal made during an aircraft penetration of a convective cloud in Montana. Each scale division represents 100 micrometers.

Aircraft measurements show a wide variation of ice crystal concentrations at a given temperature in different cloud types. Lenticular wave clouds formed near mountains sometimes are ice free at −31°F (−35°C); convective clouds over the ocean sometimes contain more than one crystal per liter of cloudy air just below 14°F (−10°C). Within this temperature range, ice crystals are nucleated on solid impurities of about 0.1 μm, having an atomic structure resembling ice (heterogeneous nucleation) carried upward from the Earth's surface—usually minerals such as kaolinite (clay) or possibly organic materials such as bacteria resulting from leaf decay.

Supercooled clouds may be induced to snow artificially by introduction of ice crystals in sufficient concentrations. This can be achieved in two ways. In the first method, cloudy air is cooled locally by dry-ice pellets at −108°F (−78°C) dropped from an aircraft. As they fall, they cool air in their wake below −40°F (−40°C), where droplets form and freeze. One kilogram of carbon dioxide produces 10^{14} crystals. A similar effect may be achieved by evaporation of liquid propane. The second method involves generating an impurity aerosol (diameter of about 0.05 μm) whose crystallographic structure is similar to ice, such as silver iodide or metaldehyde, and dispersing it into the cloud to form ice crystals [1 g AgI gives about 10^{15} (23°F or −5°C) to 10^{17} (5°F or −15°C) ice crystals].

Such an aerosol may be made by combustion of AgI-acetone solutions with a complexing additive. The aerosol generated at the ground is carried aloft under convective or turbulent conditions. Otherwise aircraft may carry continuous burners into the cloud, or drop flares containing silver compounds such as silver iodate ($AgIO_3$) which transforms to silver iodide (AgI) following combustion. Chlorine-containing compounds may be added to provide a silver iodate–silver chloride (AgI–AgCl) complex which gives a tenfold enhancement of ice nuclei.

Such flares may be dropped above cloud top, to burn as they fall through the cloud. The design of the pyrotechnic must be such that the smoke fails to coagulate prior to dilution, or the number of nuclei is dramatically reduced. Rockets have also been used to inject an explosive charge containing silver iodide into the cloud.

Historically, several projects employing these techniques have claimed an enhancement of precipitation by 10–15%. Assessment has been based on statistical evaluation of snowfall or rainfall at the ground. Since rainfall (along with most other meteorological quantities) is a highly variable quantity, and the effects produced are usually not large, such statistical tests need to be carried out over many seasons. It is necessary to evaluate whether any apparent rainfall enhancement as a departure from a mean is to be attributed to such variability or to the seeding process itself. Two techniques have been applied: use of an unseeded control area and cross-correlation of measured rainfall, and randomized seeding of one area and correlation of rainfall on seeded and nonseeded days. Each technique has problems of interpretation, since weather patterns producing precipitation change dramatically from day to day and from one year to the next. It might seem desirable to acquire for seeding trials as long a data record as practicable; however, too long a data record may lead to complications because of the possibility of changes on a climatic time scale unrelated to the changes being sought.

Physical evaluation of the seeding effectiveness is vital in removing uncertainties of an inadequate statistical base and the wide variability of a small number of precipitation situations. Aircraft penetration of seeded clouds reveals the presence of regions of supercooled water, a prerequisite, and whether or not the seeding aerosol has reached these regions to produce ice crystals in sufficient numbers (about one crystal per liter of air) to give measurable precipitation. Under ideal conditions, rainfall from the melting of snow has been measured by radar techniques and supplemented by rain gage measurements at the ground.

The most striking effects from seeding are observed in a stratiform cloud layer, or a layer of fog at the ground; dry-ice seeding from an aircraft typically produces an ice-crystal track some 0.3 mi (0.5 km) wide (**Fig. 3**). This procedure is used for clearing airports of overnight fog; it works well with cloud temperatures below about 25°F (−4°C). It has been suggested that this technique be used for clearing clouds over an urban area to allow sunlight to melt snow on the streets. By contrast, seeding cumulus clouds produces a visible result only rarely. In 1947, workers in Australia succeeded in causing a single cumulus cloud in a large number of such clouds to grow by dry-ice seeding. This was an example of dynamical seeding, resulting when release of latent heat by growth of crystals and freezing of droplets is sufficient to cause additional cloud convection. This is more likely when the cloud contains large quantities of supercooled water (2–4 $g \cdot m^{-3}$),

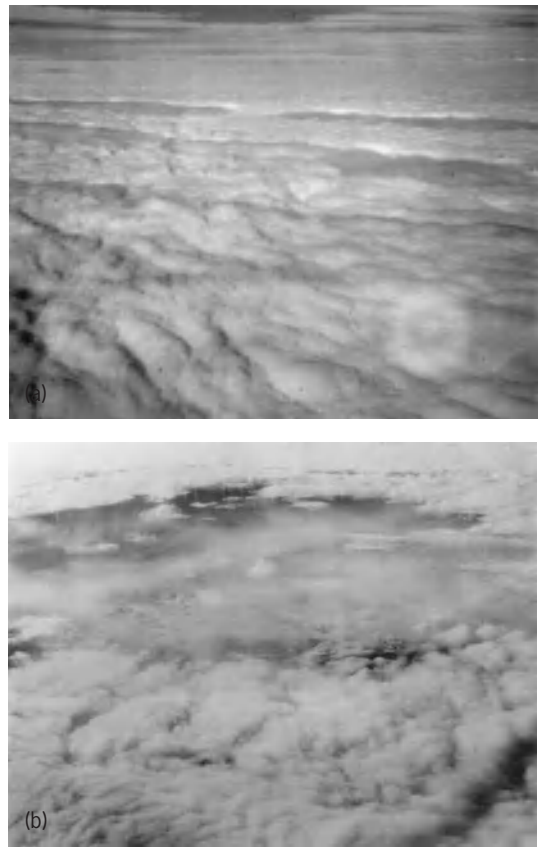

Fig. 3. Cloud dissipation. (*a*) Three lines in a stratocumulus cloud layer 15 min after seeding. (*b*) Opening in the stratocumulus layer 70 min after seeding. (*U.S. Army ECOM, Fort Monmouth, New Jersey*)

particularly in the form of supercooled raindrops. An effect is more likely when such seeding is timed to enhance the maximum natural updraft. It has been suggested that seeding clusters of clouds leads to an overall increase in rainfall, since more moist air is incorporated into the more vigorously growing region than would occur naturally.

It is evident that ice phase modification is possible only when ice crystals are not provided by natural processes—either by direct nucleation or by fallout or downward mixing of ice crystals from colder regions aloft into supercooled clouds below. A lack of crystals is often the case for ground fog or stratus cloud below a clear sky; also, for a field of developing cumuli. Developing clouds resulting from orographic lift over a mountain range offer similar opportunities implemented in many cloud seeding projects undertaken in the Sierra Nevada (California) and the Rocky Mountains. Statistical evaluation of several studies of convective cloud modification has found inconclusive results (Colorado, Florida, Switzerland), and it is evident that the possibilities of seeding for enhanced rainfall are less obvious than had first been thought.

On occasion, natural processes produce secondary ice particles from primary particles resulting from nucleation on mineral aerosol. Growth of soft hail (graupel) from supercooled cloud having an adequate spread of droplet sizes at temperatures

between 24.8° and 17.6°F (−4 and −8°C) gives such particles which rapidly grow to columns some 50–100 micrometers long (Hallett-Mossop process). Under favorable conditions, one ice particle is produced for each 50 m (164 ft) fall of a soft hail particle in the cloud, a sufficient number ot make a significant enhancement of ice concentration and ultimately precipitation. A very small number of particles—well below the detection limit in aircraft measurement—reproduce exponentially under the right conditions. A further process occurs when ice particles fall into dry or warmer air and evaporate or melt, resulting in small ice particles which may be recycled into the cloud in weak updrafts. These considerations suggest that ice cloud seeding can be effective only under a rather limited range of conditions—when natural ice production processes are not effective. Such situations occur when cool bases (<41°F; 5°C) and continental aerosol are present, where ice produced from low temperatures aloft is not recycled.

The pattern of subsequent weather modification attempts has involved searching for situations with regions having large amounts of supercooled water which persist for several hours, seeding in a controlled way, and evaluating the effect produced both by aircraft and by radar. It is being realized that seeding opportunities may be quite narrow. As a result, to become more effective in increasing precipitation may require more skill in locating seeding agents in developing clouds than is currently available. *See* CLOUD PHYSICS; FOG; PRECIPITATION (METEOROLOGY).

Hail modification. Efforts have also been made to reduce hail by similar seeding techniques. This could be accomplished by freezing every droplet in a cloud so that hailstones no longer grow, a technique that would require an unrealistic amount of seeding aerosol and would present serious distribution problems. An alternative and more realistic approach would reduce the sizes of all hailstones so that they are less than a few millimeters in diameter on falling to the ground. This would minimize crop damage. It has been suggested that hail modification can be achieved by seeding the hail-forming part of the cloud so that a given amount of cloud water is distributed over a larger number of smaller stones. Careful control of induced ice concentrations in different parts of the supercooled liquid water cloud would be required, depending on the natural hailstone concentration. This is well beyond available technology. *See* HAIL.

Hurricane modification. It has been suggested that hurricanes might be seeded, either to reduce their maximum wind strength or to change their direction of motion. However, research has revealed that large quantities of ice are often present naturally. Thus systematic large-scale seeding would not be effective. Several hurricanes were seeded during the period 1950–1975 with inconclusive results. Regions of supercooled water cloud do exist asymmetrically around the eye, associated with strong updrafts and sometimes frequent lightning. This offers the possibility of "leading" the hurricane in a different direction by selective seeding in one quadrant. Once lighting has occurred in any system, it implies the existence of an efficient ice-producing process and suggests that further seeding may be less than effective. Again, seeding opportunities may thus be very narrow. *See* HURRICANE; LIGHTNING.

Warm cloud modification. Coalescence precipitation occurs by the collision and coalescence of droplets of diameter greater than 30 μm formed by condensation with slower-falling, smaller droplets. Continuing coalescence eventually results in raindrops several millimeters in diameter. Such larger drops form on occasional large hygroscopic nuclei of sodium chloride (NaCl) or ammonium sulfate [$(NH_4)_2SO_4$] or by special mixing and evaporation mechanisms. This process produces rain not only in clouds whose temperature lies entirely above the freezing point ("warm" clouds), as is often the case in Hawaii, but sometimes also occurs in supercooled clouds. It is evident that many clouds fail to rain at all, and it has been suggested that the larger cloud droplets are not present in sufficient numbers and that seeding with larger droplets might initiate precipitation. Hygroscopic materials—such as sodium chloride—have been dispersed in airborne trails to test such hypotheses; local effects in the cloud microphysics occur, but subsequent enhancement of precipitation remains to be demonstrated. A further hypothesis is that hygroscopic seeding may influence the drop size distribution to provide conditions appropriate for secondary ice formation; while this is possible, the task of modifying the whole of the cloud in just the right place appears formidable. The inverse process—of producing many small droplets by additional nuclei from combustion such as agricultural burning—has also been suggested as leading to a rainfall decrease; this hypothesis is also unproven. Attempts to clear warm fog by high-pressure, high-volume fire-fighting sprays similar in size to 0.08-in. (2-mm) raindrops have shown some measure of success, doubling the visibility over a region of a few hundred yards.

The use of sound—as shock waves from explosives carried aloft on balloons or by ringing church bells—characterized nineteenth-century efforts in rainmaking and hail dissipation. It appears to have had only psychological value, and has been shown to be quite ineffective in causing significant changes in the cloud, either by nucleating ice crystals or by causing droplet coalescence.

Heating effects. Somewhat greater cloudiness and precipitation sometimes exist near cities, which act as "heat islands" and give enhanced convective activity and deeper clouds later in the afternoon. The effect has been found both in Chicago and St. Louis; it results from an increase in the amount of sunlight absorbed by roads and buildings and changes of surface wind flow. More dramatic effects occur over natural extended heat sources such as forest fires and also volcanoes (**Fig. 4**). Under certain conditions, setting a prairie fire will produce a thunderstorm which will extinguish the fire; this skill was practiced only in past times. *See* TERRESTRIAL RADIATION.

Fig. 4. Convective cloud forming over a forest fire.

Radiation modification. Particulates produced by forest fires and by fossil fuel combustion in industrial society pass into the air and are efficiently removed during natural precipitation processes. These particles influence the radiation balance and may be responsible for reducing sunlight and local visibility. The effect may be present on a worldwide basis; particulates have a half-life before removal of about 1 week, and can be carried considerable distances during this time. The overall result is a decrease in direct sunlight, which is somewhat compensated for by an increase in diffuse radiation scattered from other directions. The net result is a decrease in surface radiation available for photosynthesis by about 5%. The atmosphere above is heated by comparable amounts, and calculations show that this heat is lost to space more rapidly than if it were absorbed at the surface; the net result is a slight atmospheric cooling.

Average anthropogenic particle loading is around 20% of natural loading, and each is substantially less than peaks resulting from major volcanic eruptions. It has been suggested that a large effect could result from smoke produced by widespread nuclear war—the nuclear winter. This would lead to cooling the Earth's surface and heating aloft. Particles that enter the stratosphere may remain there several years and influence the radiation over this time. Similar suggestions were made for the effect of the smoke from the Kuwait oil fires following the Gulf War. This smoke was lofted only to about 3 mi (5 km), so it did not enter the stratosphere and was removed by precipitation within a few days along its trajectory; there was no detectable global effect. Contrails from high-level aircraft give a similar effect, increasing high-level cloudiness (optical thickness) in well-traveled parts of the globe. The increasing frequency of travel by jet aircraft in regions of the Northern Hemisphere could ultimately be of global importance. The effect of widespread injection of exhaust nuclei from jet aircraft at high levels (6–7 mi or 10–12 km) may be important for changing the opacity of existing cloud and for serving as particles for chemical reactions. The residence time in the upper troposphere is longer (a few weeks) and may lead to enhanced opacity of cirrus in well-traveled routes.

Greenhouse effect. In addition to particulate emissions, fossil fuel combustion increases atmospheric content of carbon dioxide, about 50% of the emissions remaining in the atmosphere, the remainder being absorbed by the oceans and contributing to enhanced photosynthesis. Carbon dioxide (CO_2) has increased by 15% since 1850, with further increases projected until fossil fuel becomes exhausted in the twenty-first century (**Fig. 5**). Other gases (for example, methane produced by rice paddies) may have a smaller (20%) but similar effect (**Fig. 6**). There is a suggestion that the recent upward tendency in the Earth's temperature may be an indication of this effect (**Fig. 7**). Calculations confirm that increased carbon dioxide would lead to an atmospheric warming, but this would be at least in part offset by the radiational cooling from articles and additional cloud formation. Such calculations are suggestive, but not sufficiently precise to predict the overall effect with confidence.

Aerosol from human activities and from episodal natural sources such as extensive forest fires and large volcanic events may have a more important worldwide effect than originally thought. Larger, more hygroscopic aerosol particles (cloud condensation nuclei) may compete with the natural background produced by ocean spray, gas reactions in the atmosphere from ocean-produced dimethylsulfide, and more universally occurring small forest fires. The result is clouds with a larger concentration of smaller droplets and of greater areal extent.

Urban pollution gives rise to this effect in low-level stratiform clouds as occur off the California coast and in the East Atlantic from industrial pollution from Europe. High-level cirrus clouds are similarly influenced by volcanic aerosols and possibly jet aircraft exhaust. Such clouds result in a higher albedo—more sunlight is scattered back to space. Such clouds form a thicker blanket and result in the emission of less radiation to space because they are colder than the underlying surface. Such changes may significantly reduce the greenhouse effect resulting from increased CO_2. *See* AIR POLLUTION; ALBEDO; GREENHOUSE EFFECT.

Ozone hole. Ozone (O_3) is normally produced at altitudes between 12 and 30 mi (20 and 50 km) in the atmosphere by ultraviolet light absorption by ordinary atmospheric oxygen. The ozone absorbs further ultraviolet radiation at these levels, providing a shield to such radiation, which has harmful biological effects at the Earth's surface.

Measurements of ozone over Antarctica in the late 1980s showed systematic wintertime decreases at 6–12 mi (10–20 km). This has been attributed to the use of manufactured fluorocarbon compounds, which are used in refrigerators, in aerosol sprays, and in the manufacture of foamed insulation. These gases are unreactive in the lower atmosphere; however, at high altitudes they become dissociated and

react with ozone. The rate of reaction is increased by adsorption on particles in the presence of nitrogen oxides (produced worldwide by soils, lightning, fertilizer use, and combustion). This leads to ozone depletion. As ozone loss extends to sunlit regions of the globe, enhanced skin cancer rates and possible damage to plants and algae may result.

The presence of aerosol at high levels in the stratosphere (15 mi or 25 km) where temperatures are low (−139°F or −95°C) leads to the presence of polar stratospheric clouds in the form of nitric acid trihydrate and, at a few degrees colder, ordinary water ice. These particles provide a surface on which chemical reactions leading to ozone destruction occur. Under normal conditions the cloud particles grow to a size of about 50–100 μm and fall out to lower levels and higher temperatures where they evaporate; the clouds lack persistence. In the presence of aerosols of volcanic origin in high concentrations (as after the Pinatubo eruption), reaction rates are much greater because of the larger surface area available; also the more numerous particles are smaller and fall out more slowly. Thus, should low temperatures occur and form clouds, the ozone destruction rate is much greater. The low temperatures are produced by dynamical (lifting) processes and, to a lesser extent, radiation loss in the winter pole; the frequency is greater in the south because the wintertime circumpolar vortex is steadier and more symmetric than in the north, where the upper-level flow is more influenced by the continents. It follows that although the ozone loss in the south is greater, the ozone loss in the north is carried toward midlatitudes because of the enhanced north/south flow in waves around the polar vortex.

The so-called depth of the ozone hole (measured in Dobson units, related to the total amount of ozone) and the area of the effect have been increasing irregularly since the hole was discovered in 1985 above Antarctica. It has subsequently been found above the Arctic. The actual effect, besides depending on the chemical composition of the stratosphere, depends on the amount and area of cooling, on the mixing in and out of the vortex, and somewhat unpredictably, on the aerosol concentration.

Ozone is also produced and reaches damaging concentrations during periods of air pollution in quiescent air as in the Los Angeles Basin in summer or in an extensive slow-moving anticyclone. In this case, ozone results from a sunlight-induced reaction requiring the presence of volatile organic compounds and nitrogen oxides. Important is the presence of a temperature inversion at a height of about 0.6 mi (1 km) to trap the products of the reactions. Urban pollution provides the organic compounds and the nitrogen oxides; over a much wider area, agriculture may also provide nitrogen oxides from fertilized soil and combustion. Such situations exist in the eastern United States and in regions of Europe, China, and Japan. *See* AEROSOL; STRATOSPHERE.

Acid rain. Sulfur is a trace component of the atmosphere, being part of the biological cycle of plants and animals. Emissions of sulfur dioxide (SO_2), particularly from fossil fuel combustion and ore refining and reduction processes, add about half as much again on a worldwide basis. Problems arise from very high concentrations of sulfur dioxide produced in industrial areas such as the northeast United States and the Ruhr Valley in Europe. These gases eventually react to form sulfate ions (SO_4^{2-}), a process enhanced by the presence of nitrogen oxides from automobile exhausts. The reactions take place in cloud

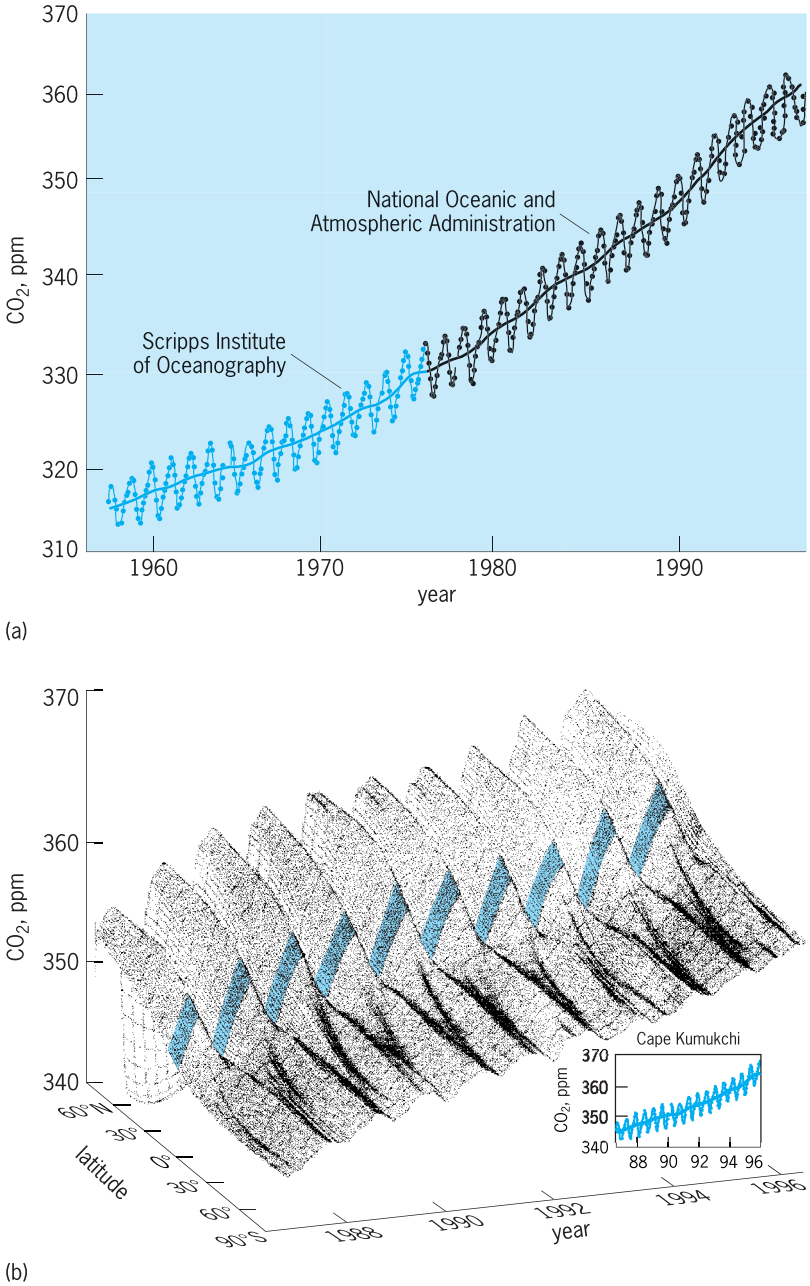

Fig. 5. Carbon dioxide concentrations. (*a*) At Mauna Loa, Hawaii. Annual fluctuations result from seasonal release and absorption of CO_2 by vegetation, primarily in the larger landmasses of the Northern Hemisphere. The steady increase is evident, as are fluctuations of the increase from year to year, probably dependent on climate effects on vegetation behavior. (*b*) Wordwide, from National Oceanic and Atmospheric Administration measurements. Note lower values in the Southern Hemisphere, showing lack of transport on an annual basis of air between the two hemispheres. The annual cycle is still present, but less in magnitude. (*After National Oceanic and Atmospheric Administration; Scripps Institute of Oceanography*)

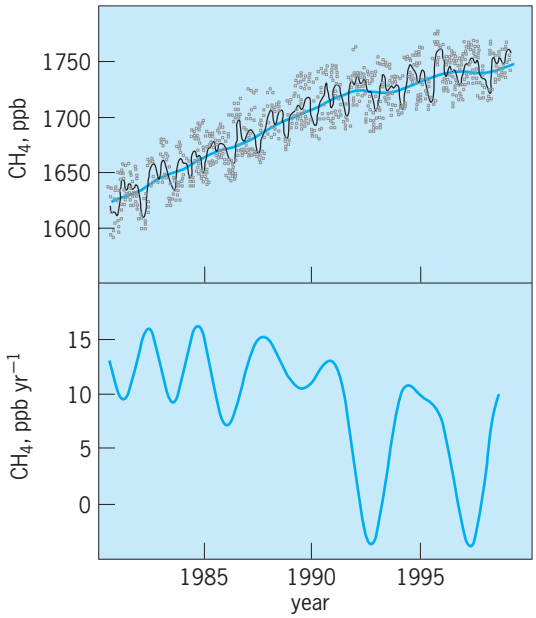

Fig. 6. Concentration of methane CH_4 gas, measured at Mauna Loa Observatory. Note an increasing tendency in the first decade of measurements, lessening in the second decade. The annual cycle is more complex than for carbon dioxide, suggesting a more complex vegetative source. (*After National Oceanic and Atmospheric Administration*)

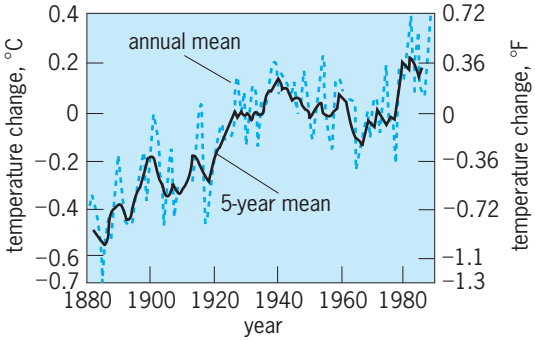

Fig. 7. Global temperature changes, 1880–1987. (*After R. A. Kerr, The weather in the wake of El Niño, Science, 242:883, 1988*)

droplets, which are removed during winter frontal precipitation and summer convective showers, to be deposited as acid precipitation often hundreds of miles downwind of the source. Rain is always slightly acidic (pH 5.7) because of the absorption of atmospheric carbon dioxide. The effect of pollution is to increase this acidity; values in the range pH 2–3 have been measured. It has been suggested that this leads to damage in spacific types of vegetation (some pines and conifers) growing in soils which lack buffering capacity, and to kills of fish in acidic spring runoff. Low-pH fogs have occasionally been observed in the Los Angeles Basin. Freezing on tress of acid fog containing nitrate (NO_3^- and sulfate ions may lead to enhanced acidity and damage at temperatures a few degrees below freezing. The actual extent of the effects is subject to disagreement because of uncertain measurement techniques of earlier years and lack of systematic observations of tree damage from natural causes over a long time period. Further uncertainty results from less frequent (and more intense) forest fires because of control strategies, leading to older stands of trees that are more susceptible to damage. *See* ACID RAIN; PH. John Hallett

Bibliography. B. Bolin et al., *The Greenhouse Effect, Climatic Change, and Ecosystems*, 1986; A. S. Dennis, *Weather Modification by Cloud Seeding*, International Geophysics Series, vol. 24, 1980; T. Conway et al., Evidence for interannual variability of the carbon cycle from the National Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network, *J. Geophys. Res.*, 99(D11):22,831–22,855, 1994; J. T. Houghton et al. (eds.), *Climate Change 1995: The Science of Climate Change*, 1995; National Research Council, *Rethinking the Ozone Problem in Urban and Regional Air Pollution*, 1991; S. F. Singer, *The Changing Global Environment*, 1975; C. Spence, *The Rainmakers: American Pluviculture to World War II*, 1980; World Meteorological Organization and United Nations Environmental Program, *Developing Policies for Responding to Climatic Change*, April 1988.

Weather observations

The measuring, recording, and transmitting of data of the variable elements of weather. In the United States the National Weather Service (NWS), a division of the National Oceanic and Atmospheric Administration (NOAA), has as one of its primary responsibilities the acquisition of meteorological information. Most observations are taken by federal employees, although private citizens and private organizations such as the utilities are also involved. The data are sent by various communication methods to National Weather Service field offices and to the National Meteorological Center near Washington, D.C.

At the Center, the raw data are fed into large computers that are programmed to plot, analyze, and process the data and also to make prognostic weather charts. The processed data and the forecast guidance are then distributed by special National Weather Service systems and conventional telecommunications to field offices, other government agencies, and private meteorologists. They in turn prepare forecasts and warnings based on both processed and raw data. *See* WEATHER MAP.

Surface observations. A wide variety of meteorological data are required to satisfy the needs of meteorologists, climatologists, and users in marine activities, forestry, agriculture, aviation, and other fields. This has led to a dual surface-observation program: the Synoptic Weather Program and the Basic Observations Program. *See* AERONAUTICAL METEOROLOGY; AGRICULTURAL METEOROLOGY; INDUSTRIAL METEOROLOGY.

The Synoptic Weather Program is designed to assist in the preparation of forecasts and to provide data for international exchange. Worldwide surface observations are taken at standard times [0000, 0600,

1200, and 1800 Universal Time Coordinated (UTC)] and sent in synoptic code.

The Basic Observations Program routinely provides meteorological data every hour. Special observations are taken at any intervening time to report significant weather events or changes. The vast majority of stations involved in these observing programs in the United States are operated by National Weather Service and Federal Aviation Administration personnel and are referred to as first-order stations. Observation sites are located primarily at airports; a few are in urban centers. At these sites, human observers report the weather elements.

Present weather. This consists of a number of hydrometeors, such as liquid or frozen precipitation, fog, thunderstorms, showers, and tornadoes, and of lithometeors, such as haze, dust, smog, dust devils, and blowing sand. The amount of cloudiness is also reported. *See* FOG; METEOROLOGICAL OPTICS; PRECIPITATION (METEOROLOGY); SMOG; THUNDERSTORM; TORNADO.

Pressure. This measurement is read from either a mercury or precision aneroid barometer located at the station. A microbarograph provides a continuous record of the pressure, from which changes in specific intervals of time are reported. Pressure changes are frequently quite helpful in short-range prediction of weather events. *See* AIR PRESSURE.

Temperature and humidity. These are measured by a hygrothermometer, located near the center of the runway complex at many airport stations. The readings are transmitted to the observation site. Readouts are either digital or dial. The temperature dial indicator is equipped with pointers to determine maximum and minimum temperature extremes; after the extremes are recorded, pointers are normally reset every 6 h. At other sites, temperature and humidity (wet-bulb temperature) are read from mercury thermometers in ventilated wooden shelters located in a grassy area. *See* HUMIDITY; HYGROMETER; TEMPERATURE MEASUREMENT.

Wind speed and direction. These measurements are telemetered into most airport stations. The equipment, consisting of an anemometer and a wind vane, is located near the center of the runway complex at participating airports; elsewhere it is placed in an unsheltered area. *See* WIND MEASUREMENT.

Visibility. This is determined from markers at known distances during daylight or from lights at known distances at night. At jet terminals, transmissometers parallel to the instrumented runway are calibrated in units of runway visibility. At airports equipped with high-intensity runway lights, runway visual range—the distance a pilot can see high-intensity runway lights along the runway while approaching for a landing—is reported. This is determined by a computer, using transmissivity, background illumination, and the intensity setting of the runway lights.

Clouds and ceilings. Various types of clouds and their heights are reported. The lowest height of opaque clouds covering half or more of the sky is known as the ceiling, and is normally measured by a ceilometer at first-order stations. The ceilometer projects a modulated light on the cloud base. At a known distance, the spot of light is electronically scanned day and night, and the cloud height is computed trigonometrically. *See* CLOUD.

Automated surface observing system. The National Weather Service, Federal Aviation Administration, and Department of Defense have been jointly engaged in the development and implementation of the Automated Surface Observing System, which is directed at providing surface observations through the use of modern sensor and computer technology. These systems measure a variety of weather elements, perform quality control, and process, display, and communicate surface information through automated techniques. These agencies have implemented more than 1000 systems in the United States to provide near-continuous (up to every minute) observations of most of the elements that have been provided manually.

Some of the information provided by the Automated Surface Observing System differs from that provided by human weather observers. For example, with this system visibility is measured at one or more visibility sensor locations rather than as the prevailing visibility (representative of a large area) observed by a human. Another difference is in the cloud observations. The human weather observer looks all around the horizon and determines how much of the sky is covered by clouds and how high the cloud bases are—a kind of area integration. The cloud sensor of the Automated Surface Observing System keeps track of clouds that move directly overhead. Its processing system tracks the clouds that have passed overhead over a 30-min period and uses that information to estimate cloud heights and amounts over a larger area—a time integration. Other information from satellites, weather radar, and lightning detection systems supplements observations made by the Automated Surface Observing System.

Other surface measurements. Several of the National Weather Service stations also measure solar radiation and the ozone content of the air, and they collect precipitation samples for chemical analysis. Automated meteorological observing stations have been established to provide basic weather observations at locations too remote for human operation, at stations staffed with part-time personnel, or where complete observation programs are not needed.

In order to better serve the interest of special users, several second-order stations have been established. For example, the U.S. Coast Guard provides limited observations along the coastline every 3 h; local airport operators provide aviation weather observations when needed for aircraft operations; and thousands of cooperative citizens record temperatures and precipitation for climatological purposes. Many of the cooperative stations also report snow density, river stage, and rates of stream discharge and evaporation. Over ocean and lake areas synoptic reports are received from moving ships.

Upper-air observations. Since the 1930s, upper-air observations have been made by the National

Weather Service with radiosondes. The radiosonde is a small, expendable instrument package that is suspended below a 6-ft-diameter (2-m) balloon filled with hydrogen or helium. As the radiosonde is carried aloft, sensors on it measure profiles of pressure, temperature, and relative humidity. These sensors are linked to a battery-powered, 300-mW radio transmitter that sends the sensor measurements to a sensitive ground receiver on a radio frequency ranging 1668.4–1700.0 MHz. By tracking the position of the radiosonde in flight with a radio direction finder or radio navigation system, such as Loran or the Global Positioning System (GPS), information on wind speed and direction aloft is also obtained.

The radiosonde flight can last in excess of 2 h, and the radiosonde can rise over 100,000 ft (30 km) and drift more than 125 mi (200 km) from the release point. The radiosonde is exposed to temperatures as cold as −121°F (−85°C) and air pressures only a few hundredths of that on the Earth's surface.

Although all the data from the flight are used, data from the surface to the 400-millibar (40-kilopascal) pressure level (about 23,000 ft or 7 km) are considered minimally acceptable for National Weather Service operations. Thus, a flight may be deemed a failure and a second radiosonde released, if the balloon bursts before reaching that pressure level or if more than 6 min of pressure or temperature data between the surface and that pressure level are missing.

When the balloon has expanded beyond its elastic limit and bursts (at a diameter of about 20 ft or 6 m), a small parachute slows the descent of the radiosonde, minimizing the danger to lives and property. Only about 15% of the approximately 75,000 radiosondes released by the National Weather Service each year are found and returned to the Service for reconditioning. These rebuilt radiosondes are used again, saving the Service the cost of a new instrument.

Worldwide, there are over 900 upper-air observation stations. Most are located in the Northern Hemisphere, and observations are usually taken at the same time each day (0000 or 1200 UTC), 365 days per year. Observations are made by the National Weather Service at 92 stations—69 in the conterminous United States, 13 in Alaska, 9 in the Pacific, and 1 in Puerto Rico. Data are exchanged between countries through international agreements.

Understanding and accurately predicting changes in the atmosphere requires adequate observations of the upper atmosphere. Radiosonde observations, plus routine aircraft reports, radar, and satellite observations, provide meteorologists with a three-dimensional picture of the atmosphere. *See* LORAN; METEOROLOGICAL INSTRUMENTATION; SATELLITE NAVIGATION SYSTEMS; WEATHER OBSERVATIONS; WIND MEASUREMENT.

Radar observations. National Weather Service, Federal Aviation Administration, and Department of Defense weather radars distributed throughout the United States are used to observe precipitation within a radius of about 250 nmi (460 km), and associated wind fields (utilizing the Doppler principle) within about 125 nmi (230 km). The primary component of this set of weather radars is known as NEXRAD (Next Generation Weather Radar), but significant weather radar observations are also provided by various Federal Aviation Administration systems. All of these radars are operated in an automated, computer-managed manner (with oversight by human personnel) to detect weather targets and provide standardized weather information products. The products provide meteorologists and other users with information on rainfall intensity, likelihood of tornadoes or severe thunderstorms, projected paths of individual storms (both ambient and within-storm wind fields), and heights of storms. This information is most helpful for short-range (up to 3 h) forecasts and warnings. The products are distributed to offices of the National Weather Service, Department of Defense, and Federal Aviation Administration and to a wide variety of other governmental and commercial users, including broadcast media. *See* DOPPLER RADAR; RADAR METEOROLOGY.

Satellite observations. Geostationary weather satellites near 22,000 mi (36,000 km) above the Earth transmit pictures depicting the cloud cover over vast expanses of the hemisphere. Using still photographs and animated images, the meteorologist can determine, among other things, areas of potentially severe weather and the motion of clouds and fog. In addition, the satellite does an outstanding job of tracking hurricanes over the ocean where few other observations are taken. Polar orbiting satellites also provide imagery for forecaster use and temperature soundings to supplement ground-based observations in numerical prediction models of the atmosphere. These data are used not only within the United States but also by many other nations of the world engaged in weather forecasting. *See* HURRICANE; METEOROLOGICAL SATELLITES.

Lightning detection systems. The technology for detecting and locating lightning strikes developed in response to a variety of needs, ranging from safety of space vehicle launches to utility company operations to weather analysis and forecasting. Ground-based lightning detection systems, which cover most of the United States and some other countries, detect the electromagnetic wave that emanates from the lightning path as the lightning strikes the ground. By the year 2000, it is anticipated that satellites will carry optical devices to detect all types of lightning over the Earth. Meanwhile, private organizations and government agencies have been cooperating in implementing lightning detection systems and in applying lightning information to both operations and research. Lightning information has proven to be operationally valuable to a wide variety of users and as a supplement to other observing systems, particularly radar and satellites. *See* LIGHTNING; LIGHTNING AND SURGE PROTECTION.

Density of observation networks. Spacing of observation points is strongly influenced by population density, with many observations near coasts, in valleys, and along main transportation routes. The required density of stations varies for each meteorological user, and there is no universally acceptable

and economically realistic plan. However, meteorologists agree on broad objectives for first-order stations. Roughly, these are surface stations not more than 90 mi (200 km) apart, climatological stations not more than 25 mi (40 km) apart, and upper-air and radar stations not more than 200 mi (320 km) apart. There is a need for a much denser real-time network to complement the first-order stations in order to improve short-range forecasts and warnings. *See* MESOMETEOROLOGY; METEOROLOGY; WEATHER FORECASTING AND PREDICTION.

Frederick S. Zbar; Ronald L. Lavoie

Weathering processes

The response of geologic materials to the environment (physical, chemical, and biological) at or near the Earth's surface. This response typically results in a reduction in size of the weathering materials; some may become as tiny as ions in solution.

Types. The agents and energies that activate weathering processes and the products resulting therefrom have been classified traditionally as physical and chemical in type. In classic physical weathering, rock materials are broken by action of mechanical forces into smaller fragments without change in chemical composition, whereas in chemical weathering the process is characterized by change in chemical composition. In practice, the two processes commonly overlap, almost inseparably. For example, diminution in particle size facilitates chemical reactivity, and an increase in volume of the products during chemical reaction may physically disintegrate the reactants.

Viewed broadly, environments of weathering and the suite of products from each may be categorized in terms of climate, such as desert, arctic, or tropical rain forest. In cold and dry climates, physical weathering predominates and produces angularity in both rock particles and surficial landforms. In warm humid climates, chemical and biochemical weathering yields rounded rock masses, and hydrated and oxidized mineral compounds which may be developed at great depths.

Agents. Within each environment, specific agents of weathering may be recognized and correlated with the types of effects they produce. Important agents of weathering are water in all surface occurrences (rain, soil and ground water, streams, and ocean); the atmosphere (H_2O, O_2, CO_2, wind); temperature (ambient and changing, especially at the freezing point of water); insolation (on large bare surfaces); ice (in soil and glaciers); gravity; plants (bacteria and macroforms); animals (micro and macro, including humans). Human modifications of otherwise geologic weathering that have increased exponentially during recent centuries include construction, tillage, lumbering, use of fire, chemically active industry (fumes, liquid, and solid effluents), and manipulation of geologic water systems.

Products. Products of physical weathering include jointed (horizontal and vertical) rock masses, disintegrated granules, frost-riven soil and surface rock, and rock and soil flows.

Products of chemical weathering include many which have been widely adapted to important economic and technologic uses. Such products include the soil, and the clays used in making ceramic structural products, whitewares, refractories, various fillers and coating of paper, portland cement, absorbents, and vanadium. These are the relatively insoluble products of weathering; characteristically they occur in clays, siltstones, and shales. Sand-size particles resulting from both physical and chemical weathering may accumulate as sandstones.

After precipitation, the relatively soluble products of chemical weathering give rise to products and rocks such as limestone, gypsum, rock salt, silica, and phosphate and potassium compounds useful as fertilizers.

Products of weathering that occur in colloidal sizes, also important qualitatively and quantitatively, are included in the preceding listings.

Processes of chemical weathering. Chemical reactions involving water and gaseous O_2 and CO_2 are probably the most important or abundant weathering processes on Earth. In sharp contrast, on the Moon, which is devoid of such an atmosphere, there is essentially no hydration, oxidation, or carbonation. Aqueous dissolution of rocks and minerals is probably the simplest or most straightforward process of chemical weathering. Solution rapidly removes rock salt (NaCl) and gypsum ($CaSO_4 \cdot 2H_2O$), but more slowly corrodes carbonate, silicate, and oxide rocks.

Hydrolysis. Water dissolves O_2 and CO_2 from the air (possibly 10 times more CO_2 from soil atmosphere), enabling it to oxidize and carbonate, as well as to hydrolyze rocks susceptible to those reactions. For example, Fe in silicate minerals is oxidized to Fe_2O_3, thereby removing Fe from the silicate structure and disrupting that network and making it more vulnerable to further breakdown. Oxidizing water reacts with metallic sulfides to produce the several sulfur-bearing acids, among them sulfuric acid, which is a powerful weathering reagent in itself. The metal constituents of the original sulfides typically become hydroxides or oxides. Fumes containing SO_2, Cl_2, or F_2 from combustion of coal, from smelters, or from industrial furnaces generally combine downwind with water vapor (humidity), rain, fog, or dew to form weathering-effective acids.

Aqueous dissolution of CO_2 produces carbonic acid, which has acidic and complexing (carbonate) properties. Dolostone (dolomite) and limestone (calcite) are quickly dissolved as Ca and Mg bicarbonates in carbonic acid, possibly producing topographic sinkholes, caves, and other karstic features, in addition to erosionally lowering the surface of those rocks. Turbulent and rapid flow of water on carbonate rocks markedly increases the rate of dissolution. The less soluble quartz, chert, clay, or iron oxides contained in dissolving limestone are left behind. Monuments and other structures composed of limestone and marble are similarly attacked.

Silicate rocks are attacked primarily by hydrolysis in a general reaction as shown below, where M refers

$$\mathrm{MAlSiAlO}_n + \mathrm{HOH} \rightarrow$$
$$\mathrm{M} + \mathrm{OH} + \mathrm{HSiO}_n + \mathrm{H(M)AlSiAlO}_n + \mathrm{Al(OH)}_3$$

to metal cations (K, Na, Ca, Mg), subscript n denotes an unspecified ratio of atoms, and the Al following Si substitutes for $\overline{\mathrm{Si}}$. Thus there are formed, by hydrolysis, soluble alkali-metal hydroxides, soluble silica (the ionic distribution depends upon pH), and relatively insoluble clay mineral (or zeolite), or less commonly, hydrated alumina. If the hydrolysis takes place at pH 9.5 or higher, both silica and alumina will be relatively soluble and mobile. They may then be separated and form bauxite ($Al_2O_3 \cdot nH_2O$). Under more acid conditions, clay minerals are formed.

Adding hydrogen ions to the hydrolyzing system increases the rate of reaction. Carbonic acid, formed when the carbon dioxide of the air and soil dissolves in water, is a source of hydrogen ions which accelerate the reaction. Organic (humic) and other acids participate in the hydrolysis. Strongly complexing organic acids may mobilize (complex) in solution Al more effectively than Si from Al-silicate minerals. Solubilization in and precipitation from organic solutions are therefore sensitive to both Eh, the oxidation potential, and pH. Another major source of hydrogen ions is their production in the ionic atmosphere about the rootlets of growing plants. During plant growth and metabolism, hydrogen ions are evolved. These are exchanged by the roots for nutrient cations (K^+, Ca^{2+}, Mg^{2+}) present in nearby clay colloids and rocks. Thus, the process of nutrition of plants is simultaneously a process of weathering of rocks. Hence, the energy which drives plant growth and is indirectly derived from the Sun likewise furnishes some of the energy for weathering of rocks.

Plant activity. Plants that are primitive in development apparently possess higher energies of cation exchange than do those that are more advanced. Lichens derive nutrient cations from fresh rock without intermediary soil. It is difficult to assess quantitatively the extent to which bacteria in the soil, and those coating interstices among mineral grains, accomplish chemical rock weathering, but some pedologists consider bacteria to be a major agent. *See* SOIL MICROBIOLOGY.

Rootlets of macroplants may sorb nutrients from adjacent soil when the mean free-bonding energy of the rootlet exceeds the crystal-bonding energy of mean free-bonding energies of clay minerals or organic substances by which they hold individual nutrient ions in polyionic systems in the soil. Hence, plant nutrition and the activity of agriculture occupy an intermediate position in the weathering sequence between fresh rock-forming minerals and intensely weathered "final" products of weathering (**Figs. 1** and **2**). Chelating organic substances extract cations from rocks, implementing rock breakdown. Partial weathering makes the rock constituents more available to plants, but extended weathering removes the nutrient materials entirely.

Results of chemical weathering. As shown by the hydrolysis reaction, the products from it may be broadly grouped into relatively soluble and relatively insoluble categories. The ultimate destination of the soluble products is the ocean, where they are concentrated in solution or removed by precipitation. Potassium released in solution by weathering, although as soluble as sodium, is more tightly sorbed by clay minerals and may be fixed in crystals of hydrous mica. Dissolved potassium is therefore less abundant than sodium in seawater. Magnesium may

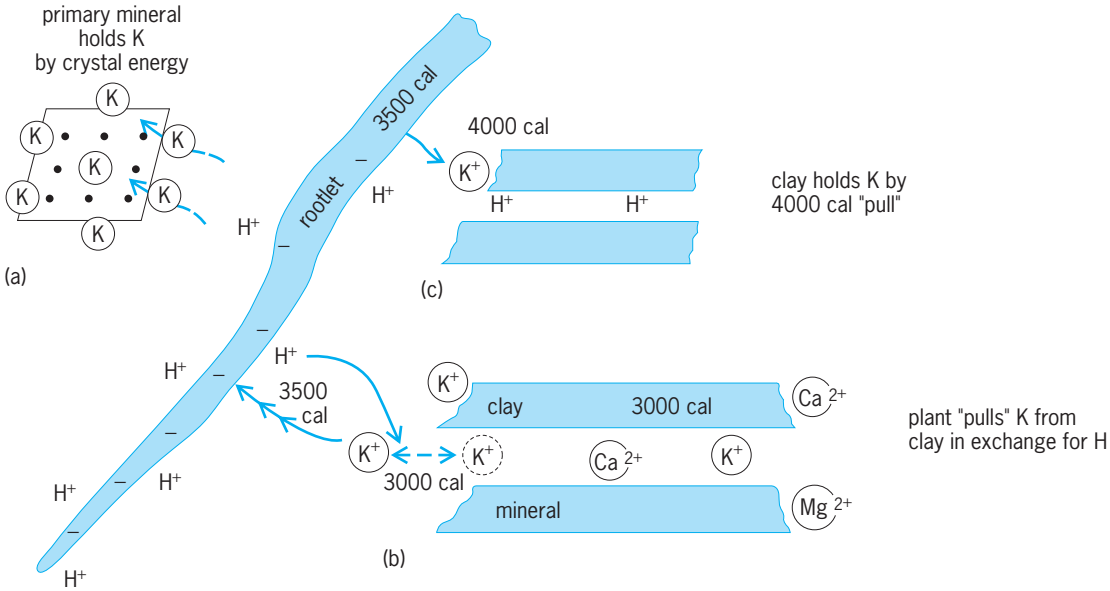

Fig. 1. Exchange-energy relationships between a rootlet and three minerals. (*a*) A potassium-bearing, primary silicate mineral. (*b*) A clay mineral well stocked with exchangeable metal cations. (*c*) A clay mineral scantily stocked with metal cations. The exchange bonding energy (calories per gram-equivalent weight) of K for H in the rootlet exceeds that in only the well-stocked clay mineral which is thus the only one of the three minerals from which nutrient ions can be taken. 1 cal = 4.18 J. (*After W. D. Keller, Mineral and chemical alluviation in a unique pedological example, J. Sediment. Petrol., 31:80–86, 1961*)

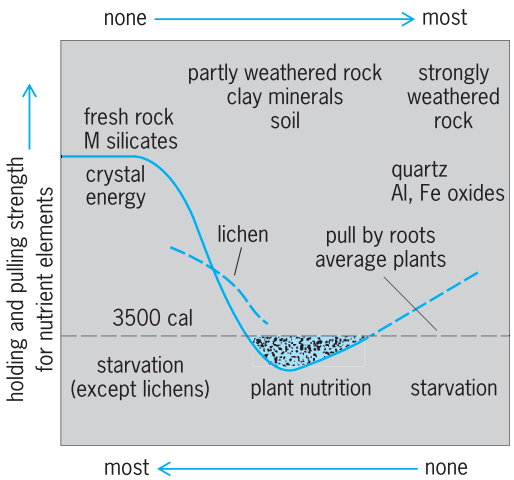

Fig. 2. Comparison of binding energy (calories per gram-equivalent weight) on nutrient metal cation (indicated by M) by rocks and soil minerals with exchange binding energy on cation by plant roots. The relation of plant nutrition to weathering and abundance of nutrients is shown. 1 cal = 4.18 J. (*After W. D. Keller, Mineral and chemical alluviation in a unique pedological example, J. Sediment. Petrol., 31:80–86, 1961*)

be incorporated in chloritic varieties of clay minerals. *See* CLAY MINERALS.

The most abundant weathering products of silicate rocks are the clay minerals. Weathering (hydrolysis) taking place in an environment such that high concentrations of calcium, magnesium, and iron (particularly ferrous) are built up tends to produce the smectite group of clays. Such a high concentration of ions occurs where evaporation exceeds precipitation, ground-water drainage is poor, or hydrolysis is rapid (as in weathering of volcanic dust). The kaolin group of clay minerals is developed where rainfall exceeds evaporation and leaching is intense. Oxidation of iron is then ordinarily high. Under conditions of very drastic leaching and continual wetting of the rocks, as in a tropical rain forest, silica and most cations dissolve, leaving hydrated oxides of alumina and ferric iron (bauxite and laterite). Rising ground–water solutions may carry Al and Fe upward and, because of evaporation or oxidation of organic complexes, leave deposits of both in the tropical subsoil. A high K^+/H^+ ratio in the aqueous-weathering system of Al-silicates yields the illite clay mineral (disordered K-mica). Weathering processes apparently reach a state of near-equilibrium with respect to kaolinite or smectite in environments such as those that prevailed where thick, valuable deposits of the clays were formed. In contrast, surface-exposed weathering of boulders and outcrops yields highly varied and changing products, quasimineral compounds, and rock wreckage.

Clay minerals, although relatively stable products of weathering in one environment, may be decomposed if subjected to more drastic leaching in another environment by processes of the removal of exchangeable cations, the more tightly fixed potassium of illite (hydrous mica) and possibly silica. Clay minerals are said to be degraded when their structures are partly destroyed. Entirely desilicated clays become bauxite or laterite. *See* BAUXITE; LATERITE.

Walter D. Keller

Bibliography. D. Atkinson, *Access to Geography: Weathering, Slopes and Landforms*, 2005; W. J. Bland and D. Rolls, *Weathering: An Introduction to the Basic Principles*, 1998; M. J. Johnsson and A. Basu (eds.), *Processes Controlling the Composition of Clastic Sediments*, 1994; A. Lerman and M. Meybeck (eds.), *Physical and Chemical Weathering in Geochemical Cycles*, 1988; R. Littke, *Deposition, Diagenesis, and Weathering of Organic Matter-Rich Sediments*, 1993.

Wedge

A piece of rigid material whose two major surfaces make an acute angle θ. It is closely related to the inclined plane and is used to multiply the applied force and to change the direction in which it acts.

For the simplest case, in which there is no friction, the forces are those shown in **Fig. 1***a*. Force F is the smaller (applied) force and Q is the larger (useful), force to be exerted. In the absence of friction, forces must act normal to the surfaces; thus, the actual force on the inclined surface is not Q but a larger force F_n (Fig. 1*b*). Summing up forces in the vertical and horizontal directions gives Eqs. (1); combining the

$$\begin{aligned} Q - F_n \cos\theta &= 0 \\ F_n \sin\theta - F &= 0 \end{aligned} \tag{1}$$

expressions thus given for F and Q, and solving for

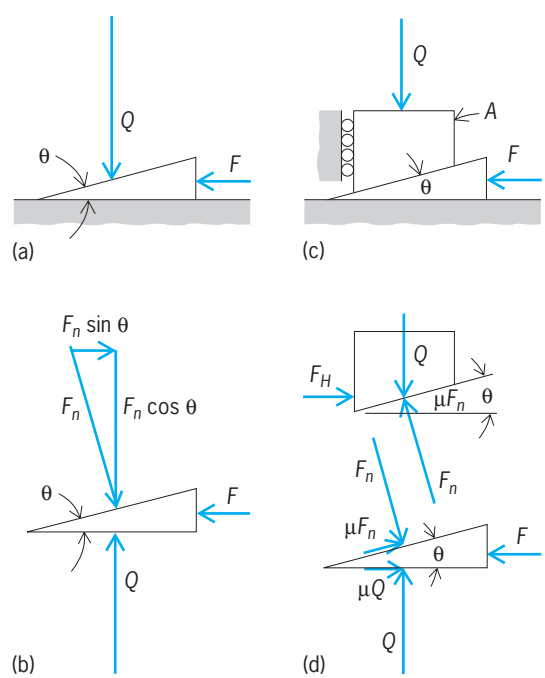

Fig. 1. Forces acting on a wedge : (*a*,*b*) without friction and (*c*,*d*) with friction.

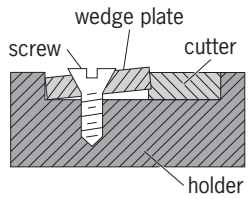

Fig. 2. Wedge plate used as clamp to retain cutter in holder.

F gives Eq. (2). Equation (3) gives the mechanical

$$F = Q \tan\theta \tag{2}$$

$$\text{M.A.} = \frac{Q}{F} = \frac{1}{\tan\theta} \tag{3}$$

advantage (M.A.), the ratio of the useful force to the applied force, as derived from Eq. (2).

In reality, friction is always present, and the analysis becomes considerably more complicated because friction forces will act to oppose relative motion (sliding) between the contacting surfaces. The simplest case is shown in Fig. 1*c*. Here the coefficient of friction μ between the wedge and adjacent members is constant, and the massless block A, upon which Q acts, is constrained without friction to move vertically. The analysis of the forces acting on the block and the wedge, shown in Fig. 1*d*, yields Eq. (4), and the mechanical advantage in this case is given by Eq. (5).

$$F = Q\frac{(1-\mu^2)\tan\theta + 2\mu}{1-\mu\tan\theta} \tag{4}$$

$$\text{M.A.} = \frac{1-\mu\tan\theta}{(1-\mu^2)\ \tan\theta + 2\mu} \tag{5}$$

In Eq. (4) it can be seen that the required force increases rapidly as μ increases, becoming infinite when $\mu \tan\theta = 1$, or $\mu = 1/\tan\theta$. At this point the mechanical advantage goes to zero.

When the block A is moving downward, the friction forces, μF_n and μQ, reverse their directions, and force F acts to restrain the motion. For this case, the force is given by Eq. (6).

$$F = Q\frac{(1-\mu^2)\tan\theta - 2\mu}{1+\mu\tan\theta} \tag{6}$$

In Eq. (6) it can be seen that the numerator decreases as the coefficient of friction μ increases, and the sign of F can become negative for a sufficiently large value of μ. When F becomes negative, the wedge must be pulled out to permit block A to move downward. Such a system is called self-locking. No simple equation exists for calculating the value of μ required for self-locking, but for practical purposes the system will be self-locking when $\mu \geq (\tan\theta)/2$.

Some applications of the wedge are in splitting wood, in raising the platform of low-lift platform trucks, in cone clutches, in V-belt drives, and in fastening parts together (**Fig. 2**). *See* INCLINED PLANE; MECHANICAL ADVANTAGE; SIMPLE MACHINE.

Richard M. Phelan

Weeds

Unwanted plants or plants whose negative values outweigh the positive values in a given situation. Weeds cost growers millions of dollars each year in reduced yield and quality of agricultural products. Especially in tropical areas, irrigation systems have become unusable because of clogging with weeds. Weeds can harbor deleterious disease organisms and insects that harm crops cost livestock producers millions of dollars. In addition, weeds can cause allergic reactions and serious skin problems (poison ivy), break up pavement, slow or stop water flow in municipal water supplies, interfere with power lines, cause fire hazards around buildings and along railroad tracks, and produce poisonous plant parts. *See* ALLERGY.

Characteristics. The most serious weeds are those that succeed in invading new areas and surviving at the expense of other plants by monopolizing light, nutrients, and water, or by releasing chemicals detrimental to the growth of surrounding vegetation (allelopathy). In the plant kingdom, dozens of species have been shown to release allelopathic chemicals from roots, leaves, and stems. The chemicals can be of many different types, including alkaloids, phenolics, terpenoids, and nitriles. Prolific seed production, seed dormancy, rapid seed germination, vigorous seedling growth, and a rapid, spreading growth habit are all beneficial to a weed for invasion and survival. *See* ALLELOPATHY.

Classification. Weeds can be classified as summer annuals, which germinate in the spring, set seed, and die in the fall (crabgrass); winter annuals, which germinate in the fall, set seed, and die in the spring (common chickweed); biennials, which germinate one year, overwinter, set seed, and die the following summer (wild carrot); simple perennials, which live for several years but spread only by seed (dandelion); and creeping perennials, which live for several years and can spread both by seed and by underground roots or rhizomes (field bindweed).

Control methods. Hand pulling, fire, flooding, and tillage are useful for controlling weeds in many situations. Insects and pathogens have also been introduced to control certain weed species. Techniques such as herbicides, computerization of spray and tillage technology, remote sensing for weed mapping and identification, and laser treatment have also been explored.

Various chemicals have always been used to suppress vegetation, but the discovery of the herbicidal properties of 2,4-D during World War II provided the stimulus for a search for new herbicides, resulting in a steady flow of new chemical tools. There are approximately 150–180 chemicals in hundreds of different formulations sold commercially as herbicides throughout the world. Some are mixed into the soil before the crop is planted, some are applied to the soil surface, and some are applied to the foliage of emerged weeds. Often, weeds can be killed without injury to the crop, because the crop fails to absorb the herbicide through a protective cuticle

on the leaves or a deep root system, degrades the herbicide before it can exert a phytotoxic effect, or contains a different form of a crucial enzyme that the herbicide blocks in the weed.

Herbicides can be applied safely and have resulted in large improvements in the availability and quality of food. They have increased the feasibility of no-till agriculture, leading to significant reductions in soil erosion. They commonly kill weeds by disrupting a physiological process that is not present in animals. Examples are the triazines, ureas, and uracils that block photosynthesis, or glyphosate and the sulfonylureas that block synthesis of essential amino acids. Consequently, many herbicides are very low in toxicity to animals. Some, however, are moderately high in toxicity and must be used carefully. *See* HERBICIDE; PHOTOSYNTHESIS.

Herbicide resistance. Although highly efficient herbicides have ben introduced, there has been a tendency to use the same herbicides continuously over several seasons because they performed more satisfactorily than the previous herbicide. However, rare individual weeds that are genetically resistant to the herbicide have flourished and reproduced, leading to populations of weeds resistant to that herbicide. For example, several dozen species have been identified as resistant to triazine herbicides, and resistant populations have developed in response to sulfonylureas, dinitrobenzeneamines, aryloxyphenoxypropionic acids, and others. Arnold P. Appleby

Biological control. Biological control involves the use of natural enemies (parasites, pathogens, and predators) to control pest populations. In the case of weed pests, the primary natural enemy groups utilized are arthropods, fungal pathogens, and vertebrates. The two major approaches include classical and inundative biological control. Biocontrol can be a highly effective and cost-efficient means of controlling weeds without the use of chemical herbicides.

Classical methods. Classical biocontrol (also termed the inoculative or importation method) is based on the principle of population regulation by natural enemies. Most naturalized weeds leave behind their natural enemies when they colonize new areas, and then can increase to significant densities. Classical biocontrol involves the importation of natural enemies, usually from the area of origin of the weed (and preferably from a part of its native range that is a good climatic match with the intended control area), and their field release as biocontrol agents. Imported biocontrol agents must be host-specific to the target weed.

Inundative methods. Inundative control is the mass production and periodic release of large numbers of biocontrol agents to achieve controlling densities. It can be used where existing populations of agents are lacking or where existing populations that are not self-sustaining at high, controlling densities can be augmented. This method is most frequently employed with fungal pathogens, where typically an agent can be applied as a mycoherbicide. One of the chief advantages of this method is that it can be integrated with conventional farming practices on cultivated croplands.

Arthropod agents. Arthropods, especially insects, are heavily utilized as imported biocontrol agents in uncultivated environments such as grasslands and aquatic systems. The most important insect orders for this are Coleoptera, Lepidoptera, Diptera, and Hemiptera. Among the many weeds successfully controlled by imported insects are prickly-pear cacti (*Opuntia* spp.) in Australia, India and South Africa; Klamath weed (*Hypericum perforatum*), which was reduced to 1% or less of its former abundance in Canada and the United States; tansy ragwort (*Senecio jacobaea*) and alligatorweed (*Alternanthera philoxeroides*) in the United States; and salvinia (*Salvinia molesta*) in Australia. *See* ARTHROPODA; COLEOPTERA; DIPTERA; HEMIPTERA; LEPIDOPTERA.

Fungal agents. Rusts (Uredinales) are the fungal group most frequently employed as imported agents in classical programs. The rust *Puccinia chondrillina*, imported from Italy, controlled the narrow-leaf form of skeletonweed (*Chondrilla juncea*) in Australia. Formulations of spores of foreign and endemic pathogens can be used as mycoherbicides in inundative applications. *See* FUNGI.

Vertebrate agents. Vertebrate animal agents typically do not possess a high degree of host-plant specificity, and their feeding has to be carefully managed to focus it on the target weeds. Goats are sometimes used to control brush in and surrounding urban areas. Fish, especially the grass carp (*Ctenopharyngodon idella*), are used to control submerged aquatic weeds. Sterile, triploid grass carp have reduced hydrilla (*Hydrilla verticillata*) biomass by more than 90% in irrigation canals in southern California. *See* AGRICULTURAL SOIL AND CROP PRACTICES. Charles E. Turner

Bibliography. F. M. Ashton and T. J. Monaco, *Weed Science: Principles and Practices*, 4th ed., 2002; K. L. S. Harley and I. W. Forno, *Biological Control of Weeds: A Handbook for Practitioners and Students*, 1992; H. M. LeBaron and J. Gressel (eds.), *Herbicide Resistance in Plants*, 1982; S. R. Radosevich and J. S. Holt, *Weed Ecology*, 2d ed., 1997; E. C. Rice, *Biological Control of Weeds and Plant Diseases; Advances in Allelopathy*, 1995; D. O. TeBeest (ed.), *Microbial Control of Weeds*, 1991.

Weight

The gravitational weight of a body is the force with which the Earth attracts the body. By extension, the term is also used for the attraction of the Sun or a planet on a nearby body. This force is proportional to the body's mass and depends on the location. Because the distance from the surface to the center of the Earth decreases at higher latitudes, and because the centrifugal force of the Earth's rotation is greatest at the Equator, the observed weight of a body is smallest at the Equator and largest at the poles. The difference is sizable, about 1 part in 300. At a given

location, the weight of a body is highest at the surface of the Earth; it diminishes with altitude and with the depth below the surface. For example, the weight of a body diminishes by about 0.1% if it is raised 2 mi (3 km) above the Earth's surface or taken 4 mi (6 km) below the surface. Weight also depends to a smaller but measurable degree on the density of the Earth's crust below the body. Weight is measured by several procedures. *See* BALANCE.

Since weight is a force, it is expressed in force units. In the United States, the commonest unit of weight is the pound, sometimes written pound force or pound weight, to distinguish it from the mass unit, pound. Pound weight is the weight of a 1-pound mass at a location where the acceleration of gravity is 32.174 ft/s^2 (9.80665 m/s^2). Where the acceleration of gravity is g, the weight of a 1-pound mass is g/32.174 pound weight.

In terms of the international kilogram the pound avoirdupois is now defined as being exactly equivalent to 0.45359237 kilogram. In relation to three smaller avoirdupois weight units, 1 pound equals 16 ounces, 256 drams, and 7000 grains. Since the relation of the pound to the kilogram is divisible by 7, 1 grain equals 0.06479891 gram, exactly.

Besides the avoirdupois pound there is the troy or apothecary pound, in which there are 5760 grains, so that this pound is equal to 576/700 avoirdupois pound. In general, the term pound is taken to mean the avoirdupois pound unless definitely stated otherwise.

The carat is a unit of weight used in evaluating precious stones, and 1 carat equals 200 milligrams. *See* MEASURE. Howard S. Bean

Bibliography. A. V. Astin, *Refinement of Values for the Yard and Pound*, Fed. Regist. Doc. 59-5442, July 1, 1959; D. Halliday, R. Resnick, and J. Walker, *Fundamentals of Physics*, 6th ed., 2001; L. V. Judson and L. E. Barbrow, *Units and Systems of Weights and Measures: Their Origin, Development and Present Status*, Nat. Inst. Stand. Tech. Lett. Circ. LC 1035, 1985; U. S. Department of Commerce, National Institute of Standards and Technology, *Factors for High-Precision Conversion*, Lett. Circ. LC 1071, 1976.

Weightlessness

A condition induced by the effective lack of resistance to gravitational force on an object or organism, sometimes known as free fall.

Explanation. Isaac Newton provided mathematical explanation for weightlessness in 1687. He was the first person to define gravitation as a force of attraction existing between all bodies, and as such one of the many forces of nature. In his classical theory of gravitation, Newton implied that each body in the universe, from a planet to a molecule, attracts every other body. This gravitational force dictates the orbit of the planets about the Sun, the ocean tides on Earth, and the adherence of humans to Earth. Gravitation is not altered by changes in chemical composition or physical state, as are electrical and magnetic forces; and furthermore, no means have been found to negate gravitational attraction. Newton proposed the law of universal gravitation, which states that two bodies of matter in the universe attract each other with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between their centers. In symbols, this is expressed as $F = Gm_1m_2/d^2$, where F is the gravitational force, m_1 and m_2 the masses of the two bodies, d the distance between them, and G a constant of proportionality called the universal constant of gravity. At the Earth's surface, $F = m_1(Gm_2/d^2) = m_1a$, where m_2 is the Earth's mass, d is the Earth's radius, and F the weight of mass m_1. The average acceleration of gravity $a = Gm_2/d^2$ is approximately 32.2 ft/s^2 or 9.8 m/s^2 at the Earth's surface, which is defined as 1 g. According to this law, even a small increase in the distance between bodies will produce a large decrease in the gravitational force, since the force decreases with the square of the distance. As a body moves from the Earth's surface to a location an infinite distance from the Earth, the gravitational force approaches zero and the body approaches weightlessness. In the true sense, a body can be weightless only when it is an infinite distance from all other objects.

Newton's second law of motion, which states that a force F acting on a body produces acceleration, explains the relationship of gravity to weight. This force acting on a body is equal to the mass m of the body times its accleration a, or $F = ma$. When this law is applied to explain weightlessness, it is evident that the weight force, or the force of gravity, produces the acceleration of gravity, which is the acceleration experienced by a freely falling body. Near the Earth's surface, freely falling bodies fall with a constant acceleration of gravity g, referred to as a standard gravity. It can then be stated that weight W is equal to the mass of the body times the acceleration of gravity, or $W = mg$. Since the weights of different bodies are proportional to their masses, in a uniform gravitational field, weight is frequently but inaccurately used to signify mass. One pound of weight is the force needed at a given location to accelerate a 1-lb mass at a rate equal to the acceleration of gravity at that location. Although weight and mass are both referred to as pounds in the English system of units, they are not the same; mass is the amount of matter present in a body irrespective of its location. Conversely, weight varies with the location of the body with regard to the gravitational field in which it is located (for example, Sun or Earth). Therefore, while the mass of a body stays constant anywhere in the universe, the weight of a body changes as the gravitational force imposed upon it is changed; when that gravitational force is reduced to zero, the body becomes weightless.

Weightlessness is also defined as a condition in which no acceleration, whether of gravity or any other force, can be detected by an object or organism within the system in question. According to Albert Einstein's principle of equivalence, there is no way to distinguish between the force of gravitational

fields and the forces due to inertial motion. When a gravitational force on a body is opposed by an equal and opposite inertial force, a weightless state is produced. This is based on the fact that the mass that determines the gravitational force of a body is the same as the mass related to the acceleration produced by an inertial force of any kind. These inertial force have no external physical origin, but are the consequences of an accelerated state of motion. Because of inertia, a moving object always tends to follow a straight line. When a person swings a bucket by the handle in a large circle, the person feels a pull on the hand, because inertial force (also called centrifugal force in this case) tends to keep the bucket moving in a straight line, while the bucket holder exerts a counterforce constraining the bucket to move along the circle. A similar situation exists in a spaceship orbiting the Earth 200 mi (320 km) above the Earth's surface, where the gravitational field is only slightly weaker than at sea level. The ship is in a state of free fall and therefore weightless. However, while it is falling or in effect being pulled toward Earth by the Earth's gravitational attractive force, the inertial force of the moving ship is directed radially outward from the Earth. Consequently, the ship follows an orbital path. *See* GRAVITATION; GRAVITY; INERTIA; NEWTON'S LAWS OF MOTION.

Realization. Weightlessness is not easily achieved. On Earth, weightlessness can be effectively created for only a few seconds by free fall using a drop tower, or for 20 or 30 s by flying a keplerian parabola in an aircraft. Rockets have been used and improved since the 1950s to achieve low-gravity conditions for scientific studies. Significant biological and materials-science research is conducted in space and on rockets that maintain microgravity conditions for 12–18 min. However, most biological and materials processes occur over a longer time frame, and require space flight to achieve effective weightlessness for periods long enough to conduct meaningful experiments. Both crewless satellites and crewed spacecraft that orbit the Earth for days and even months provide weightless conditions that can truly be used as a tool for basic and applied research (**Figs. 1** and **2**). *See* ROCKET ASTRONOMY; SCIENTIFIC AND APPLICATIONS SATELLITES.

Effects. The physical properties of the space flight environment, predominantly weightlessness, greatly impact the performance of experiments and provide the opportunity to study the effects of gravity. Efforts in fluid physics, combustion research, fundamental physics, material sciences, and biology take advantage of the weightlessness of space flight to better understand processes normally obscured by the presence of gravity.

Biological effects. Gravity is a constantly present major feature of the physical environment on the surface of the Earth. The Earth's gravity exerts a pulling force on all objects, tending to shape or deform them. Galileo recognized that as objects on the Earth's surface increased in mass the gravitational loading or weight increased, which acted to limit the size of the object. Comparing the weight-bearing musculoskeletal systems of different species of land mammals provides an example of gravity acting as a determinant of biologic function and morphology. When a mouse, human, and elephant, with respective masses of 0.7 oz (20 g), 150 lb (70 kg), and 15,000 lb (7000 kg), are compared, the skeletal fractions of their total body mass are seen to increase from 5% for the mouse to 14% for the human and 27% for the elephant; thus, the larger the animal, the larger the supporting-tissue fraction of the total body. Since gravity has persisted throughout the evolution and development of life on Earth, it seems reasonable to speculate that it is intimately involved in the form and function of most, if not all, of the organisms found on Earth. These, as well as other observations, have led to the belief that exposure of organisms to weightlessness would produce profound biological effects.

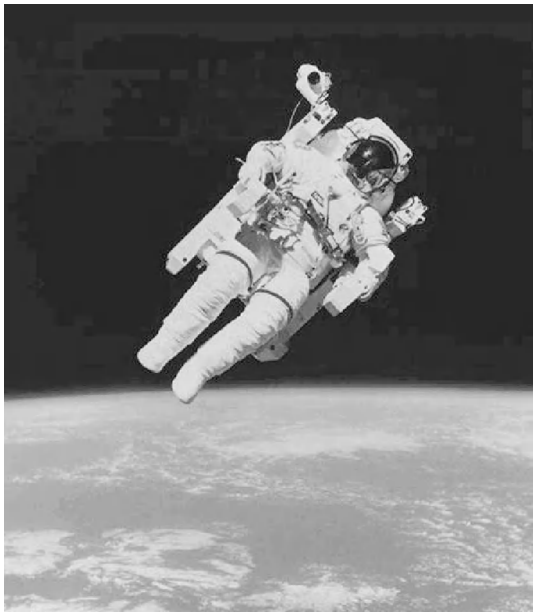

Fig. 1. Weightless astronaut Bruce McCandless during the first free (untethered) extravehicular activity outside the space shuttle *Challenger* in February 1984. Wearing the manned maneuvering unit (MMU) backpack system, McCandless traveled up to 300 ft (90 m) from the orbiter. (*NASA*)

Fig. 2. Astronaut blood flow being monitored on the middeck of the Earth-orbiting space shuttle *Atlantis* flown in October 1989. (*NASA*)

To date, biological experimentation in space and data derived from space have been limited by the number and duration of space flights. Space-flight crews have experienced weightlessness for periods up to a year, and the main body of physiological data on the effects of weightlessness on humans has been derived from observations made on these few crew members. Physiological changes occurred in most systems of the body. The mechanisms involved in most of the observed changes are, however, still unknown.

Experiments using cells have defined a pattern of change in a number of genes. These genes are specific to the structure of the supporting systems within individual cells. There appears to be a down regulation of the genes associated with the cytoskeleton, which internally supports cells. The fact that a body that is the size of a cell possess the means to sense a change in the gravitational environment suggests an amplification of physical forces via an as yet unidentified signaling pathway. Though the response to altered gravity appears to be sensed, other factors, possibly other physical forces, appear to be playing a role in the responses. In muscle tissue, organoids that have a physical load (stretch) placed on them still exhibit changes associated with exposure to the weightless environment. Thus, the changes at the cellular level are due not solely to the absence of gravity (loading) but to other physical properties of the environment. The mechanisms involved in most of the observed changes in biological systems are still unknown. These processes will be investigated on the International Space Station in a number of uniquely designed pieces of equipment such as the European Space Agency's modular culture system. *See* AEROSPACE MEDICINE; SPACE BIOLOGY.

Behavior of materials. In space, weightlessness produces flight problems. In the absence of gravity, the fuels, battery electrolytes, and other liquids do not sink to fixed levels, but coalesce in response to surface tension and the geometry of their containment. Small parts and all installation procedures must be kept scrupulously clean to keep foreign bodies of metallic dust or shavings from floating free during weightlessness and contaminating the atmosphere. Engineering problems exist in the design of feeding and waste management systems as well as laboratory equipment, plant growth chambers, animal holding facilities, and air revitalization and water reclamation systems. All of these must be designed to function in an environment where fluids float in droplets, and unattached or untethered articles drift away. *See* SPACECRAFT PROPULSION; SURFACE TENSION.

In space, sedimentation, buoyancy, and convection do not occur. This allows surface tension to play a primary role in the shape of liquids, and gas and liquid distribution and flow. Further, the role of diffusion and convection in mass transport can be distinguished in a weightless environment. *See* BUOYANCY; CONVECTION (HEAT); DIFFUSION; SEDIMENTATION (INDUSTRY).

The effect of weightlessness on materials offers a revolutionary new opportunity in materials processing. Gravity affects the mixing, separating, welding, heating, and cooling of materials on Earth. In a mixture of molecules, the light ones float to the top while the heavy ones settle to the bottom of the mixture. When light and heavy metals are heated and melted on Earth, the heavier components settle before the mixture can cool to a solid state. In the gravity environment of Earth, convective mixing occurs when heat is introduced into a material during a purifying or refining process. In electrophoresis, a process used to separate, characterize, and analyze certain biological materials, convective currents arise in the presence of gravity, and unwanted particle mixing occurs. Similarly, some characteristics of inorganic materials such as crystalline perfection, homogeneity of precipitation in multiphase systems, and purity are affected adversely by gravity when processing occurs on Earth. *See* ELECTROPHORESIS.

The unusual effect of weightlessness on the behavior of materials was demonstrated on the United States *Skylab* spacecraft in the early 1970s. It was found that in this near-weightless condition, liquid masses formed free-flying globules that could be tethered with a string; vapor grew into unlimited-size bubbles in boiling liquids; combustion products built up around burning materials and smothered the flames (**Fig. 3**); a dispersion of two immiscible liquids was more than 10^5 times more stable in space than on Earth; and reduced gravitational stresses of intramolecular bonding produced unusually large, nearly structurally perfect crystals.

Large germanium selenide crystals were grown on *Skylab*. These space-grown crystals had an almost perfect structure and quality, and one was six times larger than expected. Indium antimonide crystals that were formed in space by melting and then cooling demonstrated the shaping effects of surface tension and adhesion. In a third experiment the prediction that liquid float zones, the liquid zones

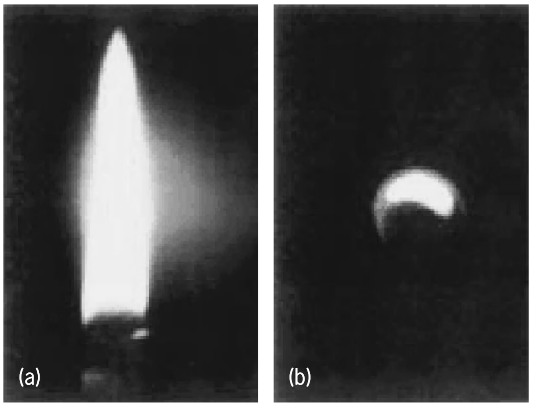

Fig. 3. Effect of weightlessness on combustion. (a) Flame on Earth is elongated. (b) Flame in space forms a circle because the lack of convection (heat rising) does not allow waste material to flow away from the flame. In an environment with no airflow, this causes the candle to burn itself out. (*NASA*)

suspended between two stationary disks, would be longer in space was confirmed. This experiment demonstrated the advantage of conducting floatzone refining, an important process in the manufacturing of high-grade electronics materials, in space. *See* CRYSTAL GROWTH; ZONE REFINING.

A series of protein crystal growth experiments were conducted on the space shuttle starting in 1983. The intent was to eliminate the density-driven convective flow and sedimentation that occurs on Earth and interferes with the formation of single large crystals. On Earth, most biological macromolecules are poorly ordered, diffract poorly, and are not suitable for the high-resolution diffraction analysis that is used to determine their three-dimensional molecular structure. Porcine elastase, interferon, and isocitrate lyase crystals grown in the space shuttle were larger, were more uniform, and yielded diffraction data of significantly higher resolution than the best crystals grown on Earth. This advance will allow scientists to identify the biological function of these macromolecules, which is determined by their structure (**Fig. 4**). *See* CRYSTALLOGRAPHY.

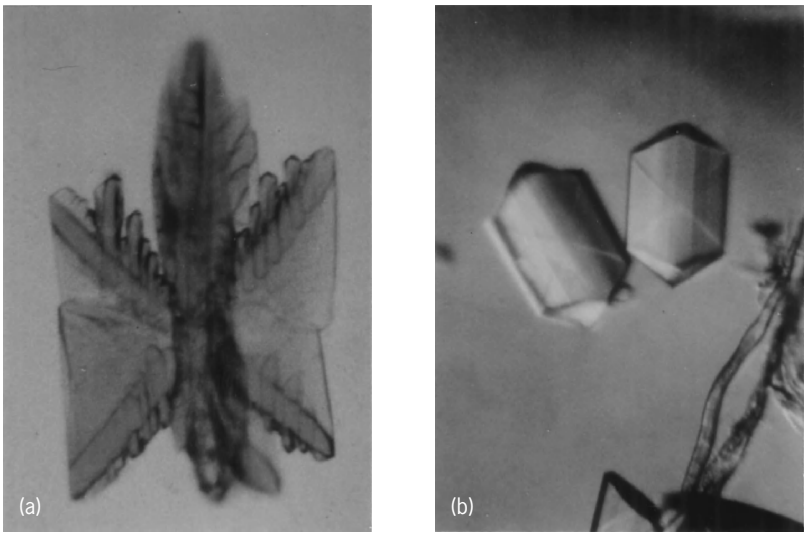

Fig. 4. Comparison of isocitrate lyase crystals (*a*) grown on Earth with typical dendritic morphology of Earth-grown crystals (*from P. C. Weber, Du Pont de Nemours & Co.*) and (*b*) grown on the space shuttle in September 1988 (*from L. J. DeLucas et al., Protein crystal growth in microgravity, Science 246:651–654, 1989*).

Russia (formerly the Soviet Union) too is actively pursuing materials science in space. A device for smelting bismuth, tin, lead, and cadmium, and a crystal growth experiment were evaluated on *Salyut 5* (1976–1977). On *Salyut 6*, in orbit from 1977 to 1982, several materials-processing furnaces were appraised and infrared-sensitive semiconductors were produced. Superconductors, eutectics, alloys, pure metals, glass, ionic crystals, and metal oxides were produced in other experiments. The first long-duration crew of *Salyut 7*, launched in April 1982, used a new computer-controlled 300-lb (135-kg) materials-processing furnace to produce several pounds of semiconductor monocrystals.

The *Mir* space station, in orbit from 1986 to 2001, has provided numerous additional opportunities to conduct materials-science and biotechnology experiments in the microgravity environment. Single crystals of semiconductor materials with improved characteristics have been grown. A tungsten-aluminum alloy composite material that cannot be produced on Earth was obtained in space; the material is lightweight and possesses good strength and wear resistance. Other experiments have studied the convergence of cells in microgravity for the purpose of obtaining very pure biologically active substances to serve as a basis for producing unique vaccines. Biotechnical experiments, such as the purification of proteins using electrophoresis, are performed in installations that are reported to be much more productive than on Earth. A technological module called *Kristall* was launched to *Mir* in 1990, providing the space station with a miniature plant for space production of ultrapure crystals and medical preparations. Analysis of these experiments is ongoing.

These efforts will be expanded using facilities on the International Space Station. A number of the international partners have developed laboratories to take advantage of, and investigate the influence of, weightlessness. These facilities are directed at the growth of purified crystals to better define the functional components of their structures, processing of materials to obtain a uniformity not achieved on Earth, and expanding basic knowledge of the effects of gravity on structures and physical reactions. An example of these efforts is the Granada Crystallization Box experiment. The experiment container, launched in August 2001, allows observations of the effect of variable levels of convection on the growth of protein crystals. This information will be used to compare different methods of growing crystals and model crystal growth in the absence of gravity. *See* SPACE FLIGHT; SPACE PROCESSING; SPACE STATION.

Thora Waters Halstead; Charles Edwin Wade

Bibliography. R. Berisio et al., Effects of microgravity on the crystal quality of a collagen-like polypeptide, *Acta Crystallog. D Biol. Crystallog.*, 56:55–61, 2000; T. G. Hammond et al., Gene expression in space, *Nat. Med.*, 5:359, 1999; Y. Kamotani, S. Ostrach, and A. Pline, Analysis of velocity data take in Surface Tension Driven Convection Experiment in microgravity, *Phys. Fluids*, 6:3601–3609, 1994; H. S. Lee, H. Merte, and F. P. Chiaramonte, The pool boiling curve in microgravity, *J. Thermophys. Heat Transfer*, 11:216–222, 1997; D. C. Walther, A. C. Fernandez-Pello, and D. L. Urban, Space shuttle based microgravity smoldering combustion experiments, *Combustion and Flame*, 1162:398–414, 1999.

Welded joint

The joining of two or more metallic components by introducing fused metal (welding rod) into a fillet between the components or by raising the temperature

of their surfaces or edges to the fusion temperature and applying pressure (flash welding).

Types. In a lap weld, the edges of a plate are lapped one over the other and the edge of one is welded to the surface of the other (**Fig. 1**).

In a butt weld, the edge of one plate is brought in line with the edge of a second plate and the joint is filled with welding metal or the two edges are resistance-heated and pressed together to fuse.

For a fillet weld, the edge of one plate is brought against the surface of another not in the same plane and welding metal is fused in the corner between the two plated, thus forming a fillet. The joint can be welded on one or both sides.

A weldment is a cast steel, forged steel, or machined steel component that is assembled to plates or structural steel shapes or other weldments by welding to form a machine part (**Fig. 2**).

A spot weld is an electrical-resistance lap weld wherein lapped surfaces are brought to high temperature by the resistance to a low-voltage, high-amperage current between two water-cooled electrodes (**Fig. 3**). This weld is used primarily for thin sheet stock.

A seam weld is similar in production to a spot weld except that the electrodes are rollers or wheels rotating at constant speed; they produce a long narrow weld. The butted edges of pipe or tubing can be welded by a modified seam weld.

Strength. Because welded joints are usually exposed to a complex stress pattern as a result of the high temperature gradients present when the weld is made, it is customary to design joints by use of arbitrary and simplified equations and generous safety factors.

The force F of direct loading, and consequently the stress S, is applied directly along or across a weld. The stress-force equation is then simply $F = SA$, in which A is the area of the plane of failure (**Fig. 4**).

For eccentric loading, the force F causes longitudinal and transverse forces of varying magnitudes along the weld. The stress is found by assuming that rotation occurs around the centroid of the welded area and that the shear stress due to the torque about the centroid is vectorially added to the direct shear caused by the force. Several points in the welded area must be checked to establish the maximum vectorial summation of the direct and eccentric stress. Maximum shear stress S_{max} is Fdr_{max}/J_{total}, where J_{total} is total polar moment of inertia of weld failure sections about the centroid, d is the distance from force F to the centroid, and r_{max} is the radius from the centroid to the farthest part of the weld. *See* STRUCTURAL CONNECTIONS; WELDING AND CUTTING OF MATERIALS.

L. Sigfred Linderoth, Jr.

Bibliography. A. D. Althouse et al., *Modern Welding*, 9th ed., 2000; T. B. Jefferson and D. T. Jefferson (eds.), *The Welding Encyclopedia*, 18th ed., 1988.

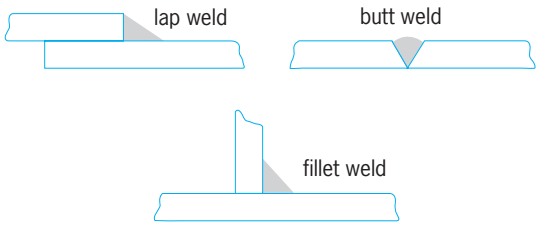

Fig. 1. Three types of welded joints.

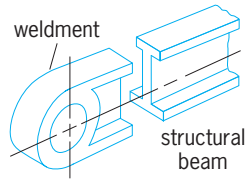

Fig. 2. Weldment is a preformed component designed to be welded to other structural or machine components.

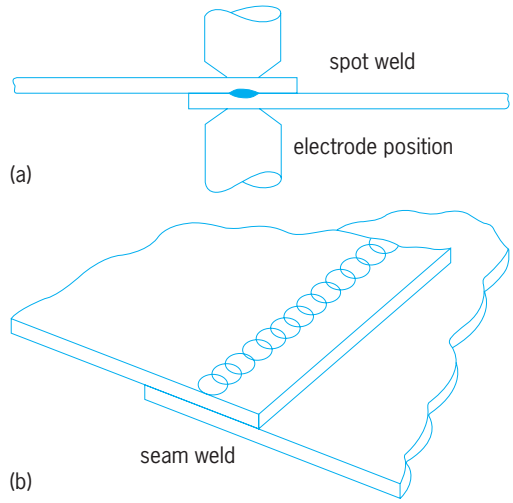

Fig. 3. Common electrical-resistance welds. (*a*) Spot weld, used mainly for thin sheet stock. (*b*) Seam weld, used to weld butted edges of pipe or tubing.

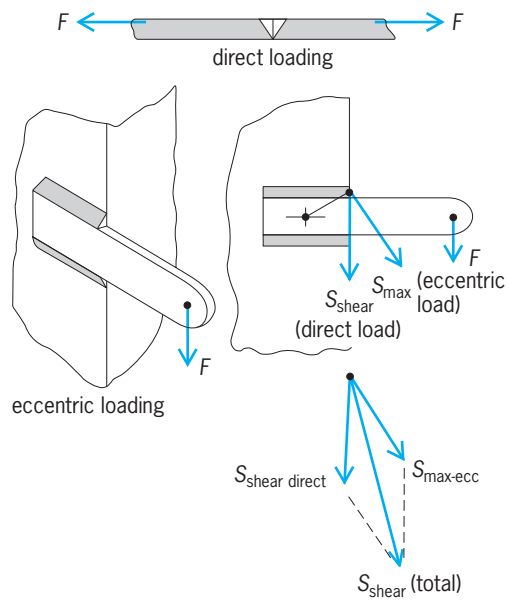

Fig. 4. Loading forces on a welded joint.

Welding and cutting of materials

Processes based on heat to join and sever metals. Welding and cutting are grouped together because, in many manufacturing operations, severing precedes welding and involves the same production personnel. Welding is one of the joining processes, others being riveting, bolting, gluing, and adhesive bonding. *See* BOLTED JOINT; WELDED JOINT.

The American Welding Society's definition of welding is "a metal-joining process wherein coalescence is produced by heating to suitable temperatures with or without the application of pressure, and with or without the use of filler metal." Brazing is defined as "a group of welding processes wherein coalescence is produced by heating to suitable temperature and by using a filler metal, having a liquidus above 800°F and below the solidus of the base metals. The filler metal is distributed between the closely fitted surfaces of the joint by capillary attraction." Soldering is similar in principle, except that the melting point of solder is below 800°F. The adhesion of solder depends not so much on alloying as on its keying into small irregularities in the surfaces to be joined. For comparison of metal joints. *See* BRAZING; JOINT (STRUCTURES); SOLDERING.

Cutting is one of the severing and material-shaping processes, some others being sawing, drilling, and planning. Thermal cutting is defined as a group of cutting processes wherein the severing or removing of metals is effected by melting or by the chemical reaction of oxygen with the metal at elevated temperatures. Welding and cutting are widely used in building ships, machinery, boilers, spacevehicles, structures, atomic reactors, aircraft, railroad cars, missiles, automobiles, buses and trailers, and pressure vessels, as well as in constructing piping and storage tanks of steel, stainless steel, aluminum, nickel, copper, lead, titanium, tantalum, and their alloys. For many products, welding is the only joining process that achieves the desired economy and properties, particularly leak-tightness. *See* TORCH.

Nearly all industrial welding involves fusion. The edges or surfaces to be welded are brought to the molten state. The liquid metal bridges the gap between the parts. After the source of welding heat has been removed, the liquid solidifies, thus joining or welding the parts together. The principal sources of heat for fusion welding are electric arc, electric resistance, flame, laser, and electron-beam. The flame in gas welding is provided by the combustion of a fuel gas.

Arc welding. The greatest volume of welding is done with arc welding processes.

Shielded metal-arc welding. This arc welding process is by far the most widely used of the various electric-arc welding processes. Like the other electric-arc welding processes, it employs the heat of the electric arc to bring the work to be welded and a consumable electrode to a molten state. The work is made part of an electric circuit known as the welding circuit. Arc welding with consumable electrodes is more widely practiced than welding with nonconsumable electrodes. A consumable electrode is melted continuously by the arc, one pole of which is the electrode, the other pole being the metal to be welded (**Fig. 1**). Generally, the consumable electrode is a wire of the same chemical composition as the work. The arc melts the electrode and some of the base metal to form a pool of weld metal, which after freezing becomes the weld. A nonconsumable electrode is tungsten or carbon, both of which have high melting points and are not consumed, except slowly by vaporization.

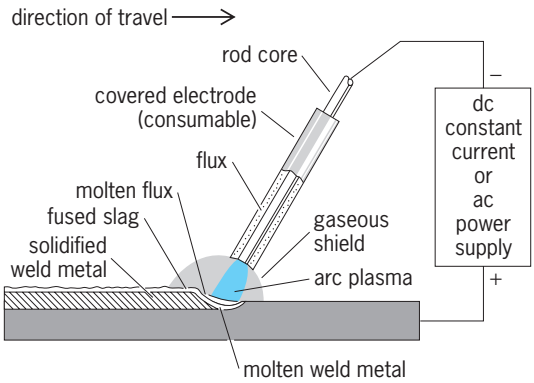

Fig. 1. Shielded metal-arc welding.

A covered electrode consists of a solid metal core; in some instances there are tubular-type cores which are covered with material serving electrical and metallurgical purposes. Electrically, the covering insulates the metal core from accidental contact with adjacent material during welding and provides an arc free from interruptions. Metallurgically, the covering may provide gas- and slag-forming ingredients to protect the weld from the air, and it may supply deoxidizers or alloying elements to produce sound welds having specified chemical composition. In addition, the covered (or coated) electrode provides scavengers to refine the grain structure of the weld metal. As the arc consumes the electrode, the operator manually feeds it into the weld.

About 6.35×10^8 lb (2.88×10^8 kg) of covered electrodes are manufactured every year with diameters ranging from $^1/_{16}$ to $^1/_4$ in. (1.6 to 6.4 mm) in stainless steel, $^5/_{64}$ to $^5/_{16}$ in. (2.0 to 7.9 mm) in mild steel, $^3/_{32}$ to $^5/_8$ in. (2.4 to 15.9 mm) in tool steel, and comparable diameters for the nonferrous alloys. The weight of the covering may be 10–50% of the weight of the covered electrode. For welding mild steel, covered electrodes are made in a number of types, which have been classified by the American Welding Society. Class E-6010 has a covering consisting principally of cellulose. The electrode can be used readily to weld flat, vertical, and overhead surfaces. The prefix E designates an electrode. In the specification for mild-steel-covered arc welding electrodes, the first two digits stand for a minimum tensile strength of the deposited metal in the as-welded condition in 1000 lb/in.^{2} (6.895 megapascals). The third digit stands for the welding position or positions in which the electrode will make a satisfactory deposit, and the last digit classes the performance

characteristics of the electrode and identifies the types of power sources on which it can be used. Class E-6012 electrodes have a thin covering high in rutile (titanium dioxide) and are adapted to welding across gaps up to $^1/_4$ in. (6.4 mm). Class E-6016 electrodes are the low-hydrogen class. The calcium carbonate in the covering evolves carbon dioxide during welding. The covering is nearly free from hydrogen, and the weld, therefore, is free from defects due to hydrogen. Class E-6024 electrodes have about 40% iron powder in the covering, which makes possible high rates of depositing weld metal. The rate of depositing steel weld metal from steel electrodes in any arc welding process ranges from 0.03 to 0.06 lb/(min)(100 A) [14 to 27 g/(min)(100 A)] of welding current.

Cored electrodes consist of a tube formed from strip and filled with slag-forming, arc-stabilizing, and alloying materials. For welding mild steel, the cored electrode is continuous. In automatic welding the arc is guided mechanically along the joint, whereas in semiautomatic welding the operator guides the arc manually by means of a torch. In some processes the arc is in air. In others a gas such as carbon dioxide protects the arc from the air.

Submerged-arc welding. This process is performed with a continuous electrode and a granular flux composed of silicates with or without deoxidizers and alloying elements. The flux is piled to a depth of $^1/_2$ to 2 in. (12.7 to 50.8 mm) along the joint. The arc melts some of the flux and is submerged in the liquid slag so produced. Although submerged-arc welding often uses current of about 40 nanometers, it can use as much as 400 nm, far above that usable in any other arc welding process. High currents enable the weld to penetrate more deeply below the surface of the work than in other arc welding processes. Welds of exceptionally good quality are produced in a wide range of metals greater than $^1/_{16}$ in. (1.6 mm) thick. Carbon, alloy, or stainless steels up to $^1/_2$ in. (12.7 mm) thick are safely welded in one pass, while thicker materials are multipass-welded. Weld-metal deposition rates (pounds of metal deposited in 1 h of operation), arc travel speeds, and weld completion rates are superior to any other process available today (**Fig. 2**).

Gas metal-arc welding. Inert-gas-shielded metal-arc welding with consumable, continuous electrodes (called GMAW for gas metal-arc welding) requires no flux and produces welds without a slag cover. The arc is in an atmosphere of argon or helium supplied at 5–100 ft^3/h (0.14–2.8 m^3/h) from the gas cup of the torch. Inert-gas shielding is particularly advantageous in welding reactive metals, such as titanium, which are susceptible to atmospheric contamination, and in welding metals, such as aluminum and stainless steel, which are susceptible to porosity (**Fig. 3**).

Carbon dioxide welding. Carbon dioxide welding with continuous electrodes is similar to gas metal-arc welding except that carbon dioxide (CO_2) is used as a low-cost shielding gas. The process is restricted to carbon and some low-alloy steels. To prevent porosity resulting from the formation of carbon monoxide by the reaction of the CO_2 with the carbon in the steel, electrodes for CO_2 welding contain deoxidizers such as manganese, silicon, aluminum, titanium, and chromium.

Nonconsumable arc welding electrodes are not deposited as part of the weld metal. The electrode is one pole of the arc, usually the electron-emitting (and therefore cooler) negative pole. The other pole is the work, which is melted. If additional metal is required to fill the joint, a filler wire is fed into the arc. Only tungsten and carbon have sufficiently high melting points to provide requisite electron emission at the high currents used in arc welding. Carbon electrodes seldom are used for welding because they vaporize more rapidly than tungsten.

Gas tungsten-arc welding. Inert-gas shielding is essential with tungsten electrodes; consequently the term "gas tungsten-arc welding" (GTAW) is used. A common electrode diameter is $^1/_8$ in. (3.2 mm) for 150 A. The process is adapted to welding thin material, $^1/_8$ in. (3.2 mm) thick and less, and to the root, or first pass, because the operator can control penetration more readily than with most other arc welding processes. The process is especially adapted for welding light-gage work requiring the utmost in quality or finish, because of the precise heat control possible and the ability of weld with or without filler metal (**Fig. 4**). Direct current (electrode negative) is used

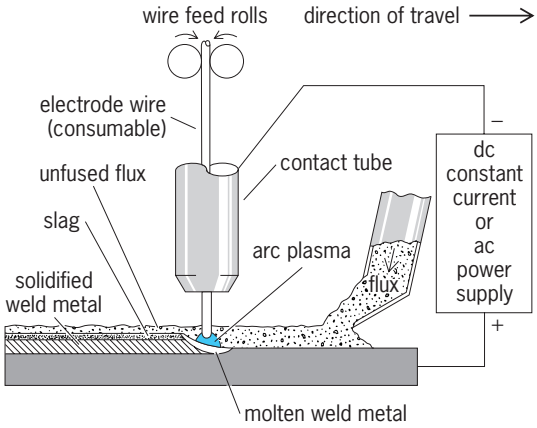

Fig. 2. Submerged-arc welding.

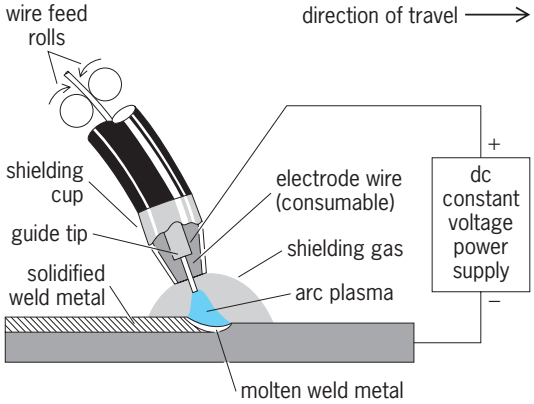

Fig. 3. Gas metal-arc welding.

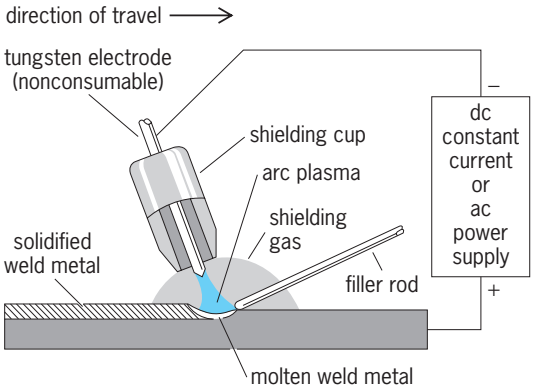

Fig. 4. Gas tungsten-arc welding.

in welding ferrous materials, copper, and nickel alloys. For aluminum and magnesium, alternating current is required because oxide film on the work is removed by reverse polarity (electrode positive). To prevent extinction of the arc as the electric potential passes through zero, a small high-frequency current (100,000 Hz) must be superimposed on the 60-Hz welding current.

Plasma-arc welding. This is a process which utilizes a plasma produced by the heat of a constricted electric-arc–gas mixture. Shielding gas may be an inert gas or a mixture of gases. Two arc modes are used to generate a plasma: transferred arc and nontransferred arc.

Plasma-arc welding resembles gas tungsten-arc welding in its use of an inert gas, but differs from it in the use of a constricting orifice. The transferred-arc mode is usually preferred. The advantages of plasma-arc welding have been observed primarily in material thicknesses greater than $^3/_{32}$ in. (2.4 mm). In such thicknesses, a significant difference is observed in the weld puddle; it is known as the keyhole effect. Surface tension causes the molten metal to flow around the keyhole to form the weld. This keyhole can be observed during the welding operation and is an indication of complete and uniform weld penetration. Another process, plasma-MIG, combines the features of plasma-arc and inert-gas metal-arc processes. The narrow arc is used for deep penetration welding of thick material or high-speed welding of thin plate. *See* ARGON.

Flux-cored arc welding. Flux-cored arc welding is somewhat like submerged arc and shielded metal-arc welding except that the flux is encased in a metal sheath instead of being laid over the wire. Although wire is automatically fed from a coil, the equipment is portable and more versatile than submerged-arc. The weld metal is shielded by the melted flux and by a gaseous medium, either externally supplied or evolved from the flux. The two principal variations in the process employ a gas-shielded electrode that requires additional shielding in the form of CO_2 around the arc and weld puddle, or a self-shielded electrode that generates its own shielding (no external shielding gas is supplied; **Fig. 5**).

The classification system for flux-cored electrodes follows the general pattern used in other filler metal classifications. In a typical designation, E70T-1, the prefix E indicates an electrode as in other electrode classification systems; the number 70 refers to the minimum as-welded tensile strength in 1000-lb/in.^{2} (6,895-MPa) units, the letter T indicates that the electrode is of tubular construction (is a flux-cored electrode), the suffix 1 places the electrode in a particular grouping based on the chemical composition of deposited weld metal, method of shielding, and adaptability of the electrode for single- or multiple-pass use.

Electroslag welding. This welding process is initiated much like the conventional submerged-arc process by starting an electric arc beneath a layer of granular welding flux. As soon as a sufficiently thick layer of hot molten slag is formed, all arc action stops and current passes from the electrode to the workpiece through the conductive slag. At this point, the process is truly electroslag welding.

Heat generated by the resistance to the current through the molten slag is sufficient to fuse the edges of the workpiece and the welding electrode. Since no arc exists, the welding action is quiet and spatter-free. The interior temperature of the bath is in the vicinity of 3500°F (1925°C). The liquid metal coming from the filler metal and the fused base metal collects in a pool beneath the slag bath and slowly solidifies to form the weld.

The wire electroslag process has been used for welding plates ranging in thickness from approximately $^1/_2$ to 20 in. (1.25 to 50 cm). The welding of greater thicknesses is technically feasible, but requires custom installations. With each electrode, the process will deposit from 25 to 45 lb (11 to 20 kg) of filler metal per hour. The diameter of the electrode wire generally used is $^1/_8$ in. (3.2 mm). Materials that have been joined by the electroslag welding process include mild and low-alloy carbon steels, stainless and high-alloy steels, and titanium.

Gas tungsten-arc spot welding. A combination of gas tungsten-arc and resistance spot welding has also been developed. Gas tungsten-arc spot welding is quite versatile and can be used to produce excellent-quality spot welds. Argon is used to shield

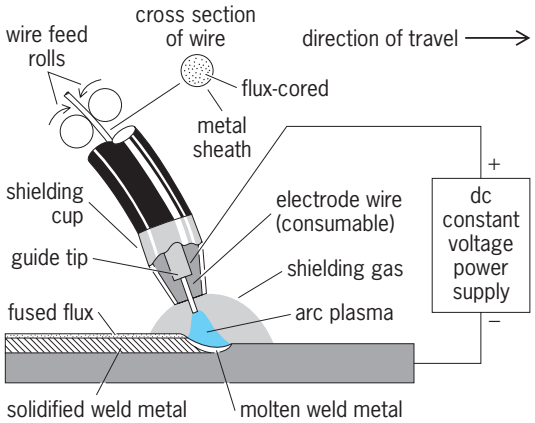

Fig. 5. Flux-cored wire welding.

the tungsten electrode and the face of the spot weld. On materials that require protection from air, the opposite side of the spot weld is shielded with either argon or helium. *See* ARC WELDING.

Resistance welding. Resistance welding processes are widely used in the manufacture of sheet metal assemblies, such as automobile bodies, aircraft, missiles, railroad cars, buses, and trailers. By definition, the required heat at the joints to be welded is generated by the resistance offered through the work parts to the relatively short-time flow of low-voltage, high-density electric current. Force is always applied before, during, and after applying current to assure a continuous electrical circuit and to forge the heated parts together.

In the most widely used processes—spot, seam, flash, and projection welding—the heat melts the surfaces to be welded. Every resistance weld involves a sequence of electrical energy and mechanical pressure. The sequence is provided by a control, which governs timing of both. For example, a control meters the 10 cycles of 60-Hz power at 10,000 A and also the electrode force at 500 lb (227 kg) required for spot welding two sheets of 0.04-in. (1-mm) steel. The welding operator merely presses the button that sets the control in operation (**Fig. 6**).

Spot welding. In spot welding, the sheets to be welded are held between electrodes of hard, high-conductivity copper alloy. They conduct heat away from the electrode-to-sheet contacts. The resistance heat at the sheet-to-sheet contact fuses the sheets together. Electrode pressure provides uniform sheet-to-sheet resistance and prevents expulsion of metal between the sheets. *See* SPOT WELDING.

Seam welding. Seam welding is a process for making overlapping spot welds by means of rotating electrode wheels. The joints can be made gas-tight and liquid-tight, as in tubing and stainless steel electric refrigerators.

Flash welding. In flash welding, the surfaces to be welded are held lightly in contact while a high current flows through the few contact points. Melting occurs instantaneously at these "bridges." Violent vaporization ensues, with expulsion of small particles of hot metal as visible flash. Flashing is continued until the surfaces are coated with a layer of molten metal. They are then squeezed together to form the weld. Flash welding is used to a considerable extent in the manufacture of automotive and aircraft products. The process is also used widely in the manufacture of household appliances, refrigerators, and farm implements. *See* FLASH WELDING.

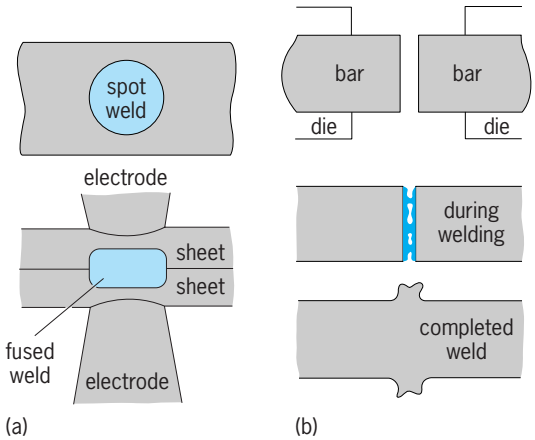

Fig. 6. Examples of resistance welding. (*a*) Spot welding (*b*) Flash welding.

Three other welding processes similar to flash welding should be mentioned: upset, upset butt and tube, and percussion welding. The greatest use of upset butt welding is in wire mills and in the welding of products made from wire. Equipment using the percussion welding process is found in several industries, mainly those in the electrical contact or component field.

Projection welding. In projection welding, a projection may be embossed on a sheet of metal. It may be produced on a solid piece of metal by machining, or it may be produced on an edge in a punch press in several ways. The height may be anywhere from a few thousandths of an inch (2.5 cm) to $^1/_8$ in. (3.2 cm) or more, depending upon the job. The purpose of projections is to localize the current and pressure at predetermined points. In this modification of the spot welding process, the concentration of the welding current is determined by the preparation of the workpieces rather than by the size and shape of the electrodes. The major portion of the heat tends to develop in the part bearing the projections during the welding operation. Cross-wire welding, which is used extensively in making wire mesh and other forms of wire products, is a form of projection welding. *See* RESISTANCE WELDING.

Gas welding. A gas flame is a less concentrated and lower-temperature source of heat than an arc. For this reason, a welding torch is often used for brazing, for welding thin material, and for applications requiring low gradients of temperature to avoid cracking, as in welding cast iron. The oxyacetylene flame, using a mixture of equal volumes of oxygen and acetylene (C_2H_2), has a higher temperature than any other commercially available fuel gas combination. High flame temperature is essential for the rapid localized heating required in welding. Several sizes and types of torch tips are available to secure flames of different intensities and dispersions. The ratio of oxygen to acetylene governs the chemical action of the flame. If the ratio is higher than unity, for example 1.1:1, the flame is oxidizing and is useful for preventing hydrogen porosity in brass. With a ratio below unity, 0.95:1, an acetylene feather appears in the flame, carbon is picked up by molten steel, and mill scale on the surface of steel is reduced. A neutral flame with a ratio of 1:1 is used for steel.

Flux is not used in welding steel but is necessary for gas-welding nonferrous metals. The use of other fuel gases, such as hydrogen, is used in the welding of metals of low melting point, such as lead, aluminum, and magnesium.

Electron beam welding. Electron beam welding is a process wherein coalescence is produced by heat

obtained from a concentrated beam composed primarily of high-velocity electrons impinging upon the surfaces to be joined. Electron beam welding equipment, which utilizes a heat source of electrons accelerated by an electric field to extremely high speeds and focused to a sharp beam by electrostatic or electromagnetic fields, is used for welding a wide range of metals in thicknesses ranging from foils to extremely thick sections. The welding equipment is manufactured for welding in high vacuum of thicknesses of 0.1 to 0.01 micrometer, in medium vacuum of 50 to 300 μm, and out of vacuum in an air or inert-gas atmosphere. The basic welding equipment is designed to operate in one of two voltage ranges. So-called low-voltage welding units operate at 15 to 60 kV, while high-voltage welders operate at 100 to 200 kV. A weld of very high purity may be obtained by this technique. Of all welding techniques, electron beam welding seems most suited for use in outer space. *See* ELECTRON MOTION IN VACUUM; VACUUM METALLURGY. Mel M. Schwartz

Laser welding. In the continuous laser beam welding process, the laser beam is focused on the workpiece surface, and at the point of focus the metal vaporizes under the beam and a vapor column extending deep into the base metal is created. The vapor column is surrounded by a liquid pool in equilibrium with the vapor (**Fig. 7**). By traversing the path to be joined with the focused beam, autogenous deep-penetration fusion welds are created (Fig. 7). Depth-to-width ratios for these welds generally exceed four to one, and are often as great as ten to one.

Low-power lasers such as pulsed ruby or pulsed yttrium-aluminum-garnet (YAG) are suitable for welding thin-gage materials such as electronic components. These lasers can produce millisecond pulses of up to 100 kW for spot welding.

Deep-penetration autogenous welds are formed primarily with continuous multikilowatt gas laser systems, the most highly developed and widely used of which is the carbon dioxide laser, operating in the far-infrared region at a wavelength of 10,600 nm. These lasers, when focused, are capable of producing power densities adequate to initiate deep penetration, a phenomenon which was previously possible only with electron beams.

The laser welding process is ideally suited to automation, since the beam constitutes a clean, remote heat source which can be readily shaped and directed by using reflective optics. A vacuum environment is not required, and a choice of shielding gasses and gas shield configurations may be adapted to specific welding situations. Unlike the somewhat similar electron beam welding process, laser welding does not generate x-rays.

The laser welding process has been shown to be suitable for a variety of metals and alloys, including ferrous materials, aluminum alloys, titanium alloys, lead, copper and copper-nickel, and most superalloys and refractory metals. Laser welds have been fabricated in these materials, which are generally ductile and porosity-free, and have average tensile strengths equivalent to or exceeding those of the base metal. High weld strengths are the result of high cooling rates, which result from the low energy input per unit length of weld which is characteristic of the laser. These high cooling rates often are advantageous in that they produce desirable structures in the welds.

Laser beam welding typically requires good joint fit-up, and is a high-speed process ideally suited to automation. As a result of the generally high cost of laser welding equipment in excess of 1-kW power level, the process is most successfully applied to high-volume applications or to critical applications requiring unique weld characteristics. *See* LASER WELDING. Edward M. Breinan

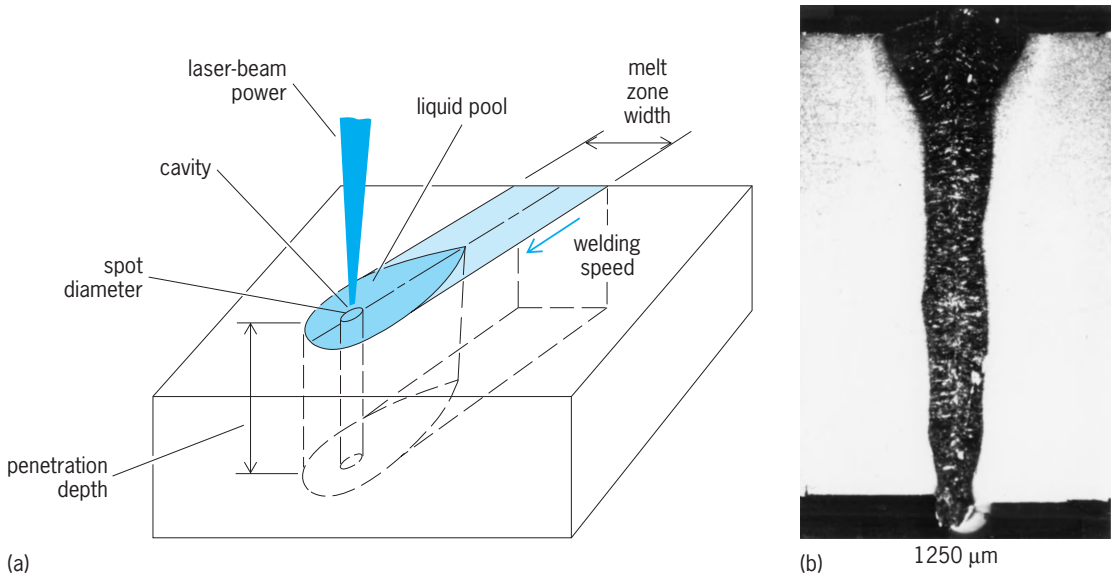

Fig. 7. Deep-penetration welding characteristics. (*a*) Schematic diagram of the deep-penetration laser welding process. The laser beam creates a vapor column (deep-penetration cavity) and deposits energy through the thickness of the workpiece. (*b*) Typical weld.

Inertia welding. Inertia welding is one type of friction welding process which utilizes the frictional heat generated at a pair of rubbing surfaces to raise the temperature to such a degree that the two parts can be forged together and form a solid weld. The process is called inertia welding because the energy required for making the weld comes from a rotating flywheel system built into the head stock of the machine. Somewhat similar to an engine lathe, the inertia welding machine also has a tail stock where one side of the workpiece is held and can be pushed forward and retracted by a hydraulic ram.

During inertia welding, one part of the workpiece held in the spindle chuck of the head stock is accelerated rapidly to a preset spindle speed. The drive power is then cut off while the nonrotating workpiece on the tail stock is pushed against the rotating part under a large thrust force. The friction at the interface brings the spindle to full stop in a matter of seconds. In the meantime, the kinetic energy stored in the spindle-flywheel system is converted into frictional heat which causes the temperature at the weld to rise quickly so that a forge weld can be formed. Because of the high pressure, the original materials at the rubbing surfaces are first softened and then squeezed out of the interface to form a flash; thus, the process brings the nascent sublayers of the material close together and results in a metallurgical bond. The weld is formed at solid state without melting; it has a finer grain structure and narrower heat-affected zone as compared with any fusion welding process, for example, electric arc welding. This often results in higher strength at the weld than at the base material.

Inertia welding is commonly used to join dissimilar materials which cannot be achieved easily or economically by any other means. Material pairs of large varieties can be welded by this process, for example, steels to aluminum alloys or copper, copper to titanium or zirconium alloys, and many others. In addition to these advantages, inertia welding machines can be readily automated for mass production. The automotive and oil drill industries are the major users of the process. Typical applications include engine exhaust valves, shock absorber parts, universal joints, axle housings, hydraulic cylinder components, and oil drill pipes. The process is also applied to weld stainless steel to plain carbon steel for outboard motor shafts, and gas turbine parts in the aircraft industry. *See* INERTIA WELDING. K. K. Wang

Other fusion welding processes. Among other fusion welding processes are high-frequency, thermit, stud, and plastic welding.

High-frequency welding may be divided into two categories: (1) high-frequency resistance welding, in which the high-frequency current is introduced into the work by direct electrical contact; and (2) high-frequency induction welding, in which the high-frequency current is induced in the work by an inductor coil, without electrical contact. The proximity effect concentrates the current in the edges to be joined, as in the manufacture of metal tubes.

In thermit welding, the joint is produced by heating with superheated liquid metal and slag resulting from a chemical reaction between a metal oxide and aluminum, with or without the application of pressure. Filler metal, when used, is obtained from the liquid metal.

There are two basic types of stud welding; arc stud welding and capacitor-discharge stud welding. Arc stud welding, still the more widely used of the two basic stud welding processes, is similar in many respects to manual shielded metal-arc welding. The heat necessary for end welding of studs is developed by passage of current through an arc from the stud (electrode) to the plate (work) to which the stud is to be welded. Capacitor-discharge stud welding derives its heat from an arc produced by a rapid discharge of stored electrical energy with pressure applied during or immediately following the electrical discharge.

There are several different ways of joining plastics to each other: hot-gas welding, friction welding, heated tool welding, and ultrasonic welding. Welding is a convenient method of joining most thermoplastic materials. The procedure is similar to that followed in metal welding when the oxyacetylene process is used. The strength of the weld, however, may vary according to the type of plastic material. The strongest joints are usually obtained from plastics with the highest degree of polymerization.

Diffusion bonding. Diffusion bonding (also known as diffusion welding) is a solid-state process in which joining is accomplished without a liquid interface (brazing) or the creation of a cast product via melting and resolidification (welding). Diffusion bonding produces solid-state coalescence between two materials under the following conditions: (1) joining occurs at a temperature below the melting point of the materials to be joined; (2) coalescence of contacting surfaces is produced with loads below those that would cause macroscopic deformation to the part; and (3) a bonding aid can be used, such as an interface foil or coating, to either facilitate bonding or prevent the creation of brittle phases between dissimilar materials.

Diffusion bonding facilitates the joining of materials to produce components with no abrupt discontinuity in the microstructure and with a minimum of deformation. The process is generally considered complete when cavities fully close at the closely fitting (faying) surfaces.

Formation of a diffusion weld occurs in three stages: initial contact of the interfaces; a time-dependent deformation of the same interfaces, further establishing intimate interfacial contact; and diffusion-controlled elimination of the original interface. The final step may occur as the result of grain growth across the interface, the solution or dispersion of interfacial contaminant, or a simple diffusion of atoms along or across the original interface (**Fig. 8**).

Although not a part of the mechanism of diffusion bonding, the use of intermediate materials is of considerable importance. Intermediate materials can be used to promote diffusion at much lower temperatures than for self welding, promote plastic flow and

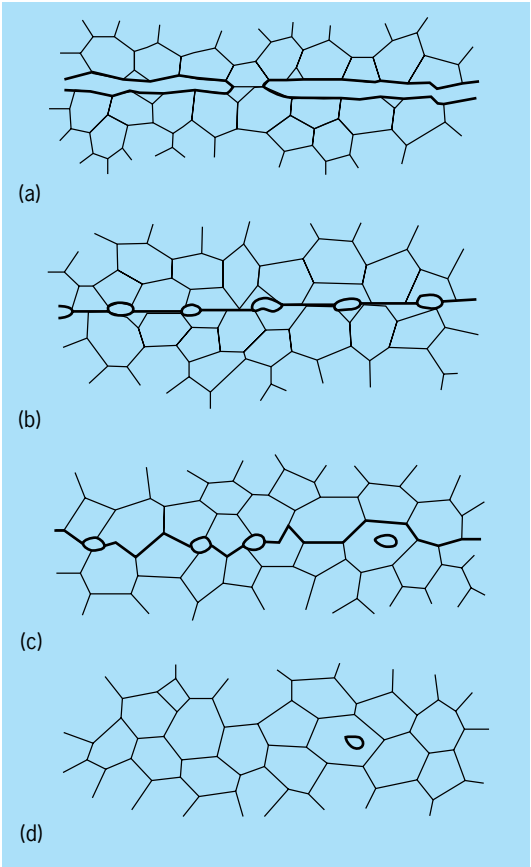

Fig. 8. Sequence of metallurgical stages in the diffusion bonding process. (*a*) Initial contact: limited to a few asperities (room temperature). (*b*) First stage: deformation of surface asperities by plastic flow and creep. (*c*) Second stage: grain boundary diffusion of atoms to the voids and grain boundary migration. (*d*) Third stage: volume diffusion of atoms to the voids.

surface conformance at lower pressures, prevent the formation of intermetallic compounds, and provide clean surfaces. The intermediate layer also permits welding in very short times. Intermediate or activating agent materials are commonly used in the form of foil, electroplates, or vapor deposits.

In practice, oxide-free conditions exist for only a limited number of materials. Accordingly, the properties of real surfaces limit and impede the extent of diffusion bonding. Titanium alloys as a group are the most notable exception. At temperatures greater than 1560°F (850°C), the titanium alloy can readily dissolve minor amounts of adsorbed gases and thin surface oxide films and diffuse them away from the bonding surfaces so that they will not impede the formation of the required metallic bonds across the bond interface.

Joint designs for producing parts that utilize intermediate diffusion aids to produce a diffusion weld are similar to those used for conventional brazing. The fundamental butt and lap joints with modifications such as scarf joints are used universally.

Materials. A number of materials can be joined by diffusion bonding; some are easier to diffusion-weld than others.

Aluminum and its alloys are difficult to join by diffusion bonding because of the tightly adhering oxide film. When the oxide is removed, the process can take place at temperatures of 849–1000°F (454–538°C) for up to 4 h. The time-temperature requirements can be considerably reduced by means of a diffusion aid such as silver or copper.

Other metals that can be joined by this process include beryllium and its alloys, copper and its alloys, heat-resistant cobalt or nickel alloys, various steels, columbium and its alloys, tantalum and its alloys, titanium and its alloys, and zirconium and its alloys.

Materials other than metals that can be joined by diffusion bonding include a number of composites, composite-metal combinations, and combinations of dissimilar metals and ceramics (such as ferrous and nonferrous materials).

Applications. Many industries have taken advantage of the benefits of the diffusion-bonding process, especially the aircraft industry, where the interest in titanium alloys is significant. Another application is in the refrigeration industry, where aluminum tubing can be substituted for copper in refrigerant line sets and the aluminum tubing subsequently joined to a brass end fitting. The metallurgical weld in this case is created without melting either the brass or the aluminum. Mel M. Schwartz

Transient liquid-phase bonding. Compared to diffusion bonding, transient liquid-phase (TLP) bonding (also know as diffusion brazing) has the advantage of being much more tolerant of the presence of surface oxide layers on the substrates. It is also suitable for a wider range of geometries than diffusion bonding. However, unlike diffusion bonding, an interlayer is always required. As a result, the bond can never have exactly the same composition as the substrates. In some cases, the bonding times required to complete isothermal solidification and subsequent homogenization of TLP-bonded joints can be very long.

Transient liquid-phase bonding relies on the use of an interlayer, which is common with the brazing process. When heated to a suitable temperature, the interlayer forms a liquid between the materials to be joined (commonly known as the substrates or workpieces). In brazing, the composition of the interlayer is selected so that the liquid is formed at a temperature below the melting point of the substrates. This allows a wide variety of materials and geometries to be joined. Unfortunately, the use of a relatively low-melting-temperature interlayer makes conventional braze joints unsuitable for use in applications that involve exposure to high temperatures. In contrast, compositional changes that occur during transient liquid-phase bonding can result in a bond with a structure and properties very close to those of the substrates as in welding.

Process. The diffusion brazing process proceeds as shown in **Fig. 9** for a joint between two substrates consisting of some metal A. The substrates are joined using an interlayer consisting of metal A plus another metal B (Fig. 9*a*). Addition of B serves to reduce the melting point of the interlayer below that

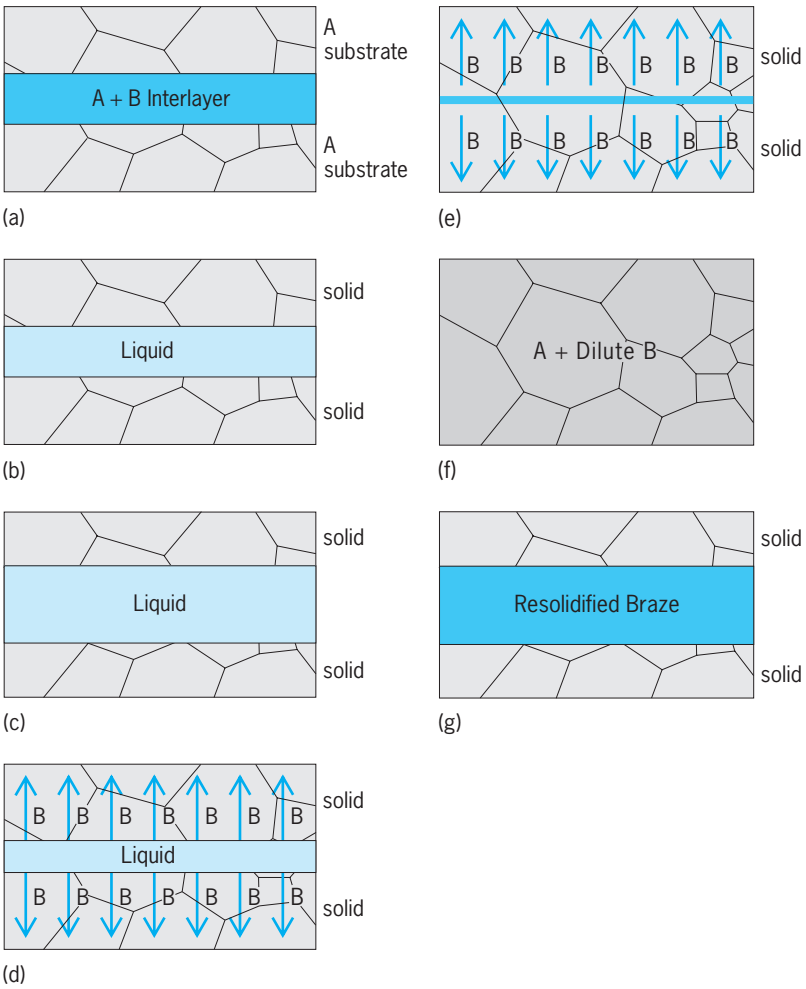

Fig. 9. Diffusing brazing process, showing the steps involved in the formation of a joint by transient liquid-phase bonding. (*a*) Initial condition. (*b*) Melting of interlayer. (*c*) Dissolution. (*d*) Isothermal Solidification. (*e*) Homogenization. (*f*) Final joint; diffusion brazing process, showing the steps involved in the formation of a joint by transient liquid phase bonding the crystals (grains) making up the substrate material extend across the bond. In comparison, (*g*) a conventional braze joint has gone through steps *a*, *b*, and *c*, and has then cooled.

of the substrates. The process temperature is chosen so that the interlayer melts but the substrates remain solid (Fig. 9*b*). After the interlayer melts, there is some tendency for the substrates to dissolve into the liquid and so the liquid initially becomes wider (Fig. 9*c*). Subsequently, atoms of metal B diffuse into the substrates (Fig. 9*d*). Because metal B reduces the melting point of the interlayer, the joint gradually resolidifies, as this metal diffuses away into the substrates (Fig. 9*d*). Normally, liquids resolidify only when cooled, but the compositional changes occurring during bonding allow resolidification of the joint while holding at a constant temperature. This is known as isothermal solidification. After isothermal solidification is completed, atoms of metal B continue to diffuse (Fig. 9*e*), and the result can be a homogenized joint that is almost identical to the substrates (Fig. 9*f*) and has very similar properties (unlike a conventional braze joint, as shown in{\uppercase{break}} Fig. 9*g*). Note that there is also a variant of the transient liquid-phase bonding process in which the interlayer starts as a pure metal and forms a liquid by reaction with the substrates.

Materials and applications. Transient liquid-phase bonding is used for both manufacturing and repairing components for high-temperature use, especially in gas turbines (for both aircraft engines and electric power generation). Mostly, the substrates involved are nickel-base alloys, and boron is often used to reduce the melting temperature of the interlayer.

William F. Gale

Explosive welding. Explosive welding is an important metal-joining technique because it can handle unique joining problems and large areas. Also known as explosive bonding, explosive welding is accomplished by a high-velocity oblique impact between two metals. The impact must have sufficient energy to cause the colliding metal surfaces to flow hydrodynamically when they intimately contact one another in order to promote solid-state bonding. Oblique impact is important because conservation of momentum allows for a reentrant jetting action that is due to hydrodynamic flow of the faying metal surfaces. The jet is ejected outward from the collision apex between the metals and produces a cleaning action by scarfing or effacing the metal surfaces. The resulting virgin metal surfaces are then compressed together under high pressure from the explosion, which promotes atomistic bonding.

The controlled application of the enormous power generated by detonating explosives permits cleaning of metal surfaces of contaminants such as films of oxide and nitride, oil, or adsorbed gases, and immediate pressing these thoroughly clean surfaces together at pressures of several million newtons per square meter. Further, the relatively small amount of energy involved with the process gives rise to minimum melt accumulation at the metal-metal interface. These are ideal conditions for effective welding.

Materials. Combinations of metals such as aluminum to steel and titanium to steel, which form brittle intermetallics when exposed to the elevated temperatures of conventional welding methods, are readily joined by explosive welding. The only metals that cannot be welded by the process are those too brittle to withstand the impact of the explosive impulse. In general, any metal or alloy possessing sufficient ductility and impact resistance such that it will not fracture as a result of the dynamic forces involved may be explosively welded. Cladding metals with elongation as low as 5% have been explosively welded. *See* CLADDING.

Process. To achieve a metallurgical bond between two similar or dissimilar pieces of metal, the atomic structure of the first metal must be brought sufficiently close to the atomic structure of the second metal that the cohesive atomic forces of each array of atoms can effectively act on the other. This normally requires an intimacy of contact on the order of an interatomic spacing. Because metal and metal-alloy surfaces in a normal atmospheric environment are always covered with films of oxides, nitrides, and adsorbed gases, the surfaces of the mating metal plates, even under very high direct pressure, will not bond metallurgically because sufficient intimacy of contact cannot be established.

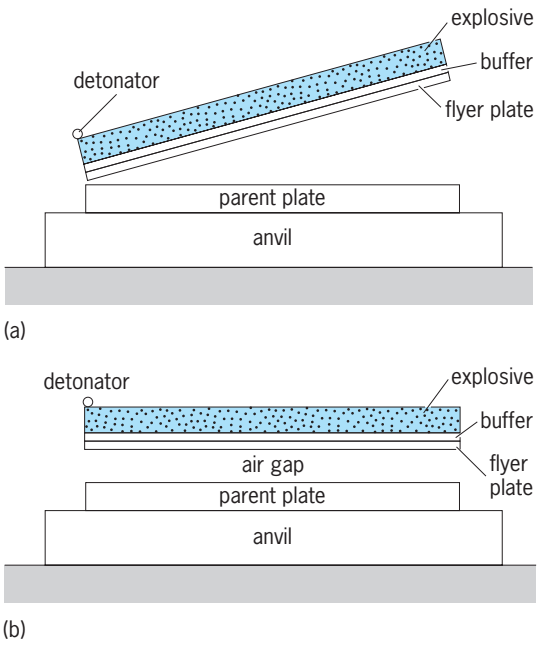

Fig. 10. Explosive welding arrangements. (*a*) Inclined. (*b*) Parallel.

The removal or effacement of surface films and the establishment of the necessary intimacy of contact must be achieved in any metallurgical bonding process. Fusion welding, brazing, and soldering are examples of joining processes that effectively disrupt surface films by dissolution or melting, and they establish the necessary intimacy of contact by wetting of the molten phase. However, processes such as friction welding and ultrasonic welding disrupt surface films by severe localized deformation of the mating surfaces while the surfaces are maintained in intimate contact by high pressure.

The process is basically simple. Typical arrangements are inclined or parallel (**Fig. 10**). The prime metal is oriented to the backer metal, with an air gap separating them. The explosive (in plastic, liquid, or granular form) is placed uniformly over the prime metal. The base of the backer rests on an anvil appropriate to the thickness of the backer and to the allowable distortion of the final product.

Applications. The widest application of explosive welding is in the area of cladding and welding of flat plate and sheet, which involves the simplest and most straightforward procedures. The major commercial application is the sheet and plate composite; the total thickness of sheet and plate that have been welded ranges from a few mils to over a foot, with cladding metal thickness ranging from 0.03 mm to more than 38.1 mm. The sizes of sheet and plate explosively welded range from small coupons to areas as large as 10 ft × 26 ft (3 m × 8 m).

Explosively welded products utilized commercially include clad tubular products with either internal or external clads. An example is the use of titanium–stainless steel transition joints in the Apollo lunar excursion module.

Welding metallurgy. The effects of welding on a product often govern its suitability for service. The weld metal, the zone close to the weld (called the heat-affected zone), and the thermal stresses and strains caused by welding may combine to yield a product different in strength, ductility, leak-tightness, and dimensional stability from that of the unwelded base metal. Weld metal frequently is compared with cast metal, with which it shares the characteristic of freezing in a columnar dendritic pattern in the direction opposite to the direction of withdrawal of heat. Unlike cast metal, weld metal usually has a free surface exposed to the air, and it is free from shrinkage cavities. Weld metal rapidly freezes compared with most castings and is comparatively fine-grained and free from chemical segregation. Porosity must be guarded against. It is a result of evolution of gas, such as nitrogen from steel, during freezing of the gas-saturated weld pool. Porosity also may result from a chemical reaction in the weld metal, such as the reaction between carbon and oxygen in liquid steel. Finally, porosity may occur as a result of inadequate inert-gas coverage during welding because of poor fixturing or poor welding conditions and procedures. Hot cracking of the weld metal during the last stages of solidification may be prevented either by fixturing to avoid strain during freezing or by adding suitable alloying elements to the weld, as in the addition of manganese to steel weld metal to combine with sulfur.

The heat-affected zone may be softer than the base metal, especially in alloys that have been cold-rolled or heat-treated for high strength. In some instances, for example the welding of annealed air-hardening steels, the heat-affected zone may be hard and brittle. Preheating and postwelding heat treatment may be required to avoid brittleness or restore strength. To achieve higher notch-impact value in steel weld metal than can be attained in a single pass, two or more passes must be deposited to take advantage of the grain-refining (hence toughening) effect of each pass on the heat-affected zone of the preceding pass. Welding heat causes distortion, which can be minimized by suitable choice of fixturing, joint design, and sequence of welding. The welded part also contains internal stresses, often up to yield strength locally in the vicinity of the weld. These may be reduced after welding by stress-relief heat treatment, which entails, in the case of steel, heating the welded structure to 1200°F (650°C) and cooling slowly. *See* HEAT TREATMENT (METALLURGY); METALLURGY.

Mel M. Schwartz

Microstructure modeling. The performance of a weld depends on its microstructure—that is, the spatial arrangement of its phases. The crystal structures or elemental compositions observed in the microstructure differ in the various regions of the welded structure, such as the base metal, the heat-affected zone (HAZ), and the fusion (melted metal) zone (FZ). During a typical welding process, microstructure evolution is affected largely by the rapid increase and decrease in temperature, which varies from region to region in the weld as shown in **Fig. 11**. These

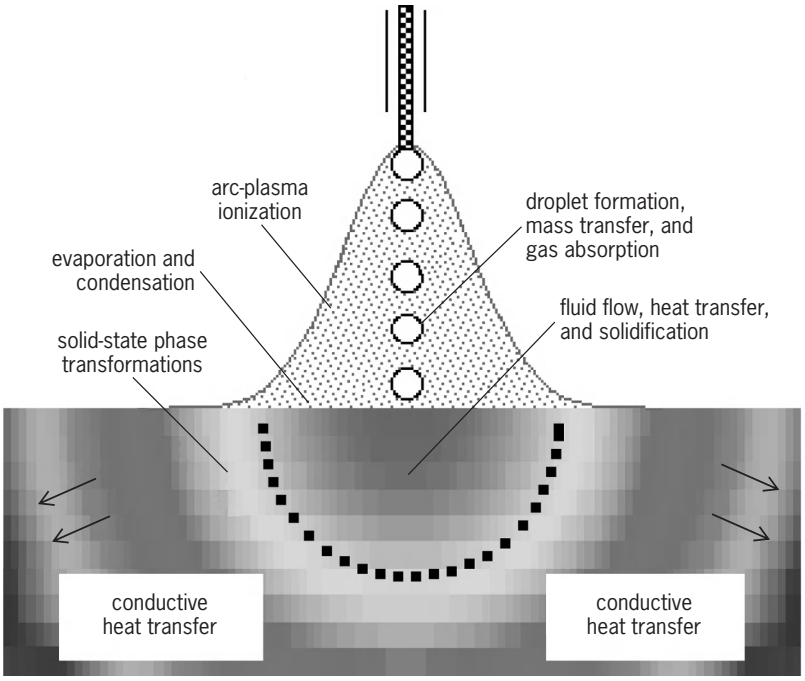

Fig. 11. Overview of physical processes in a typical arc welding process. The gray tints indicate temperature gradations.

temperature variations lead to severe microstructural gradients over a distance on the order of a few millimeters or less. *See* CRYSTAL STRUCTURE; METALLOGRAPHY.

Research is under way to describe these effects using computational models and experimental techniques to determine the best welding processes and process parameters for a given material. This is not a trivial task, considering the various physical processes that occur during welding (Fig. 11). The models are used to describe the microstructure evolution in the heat-affected zone and fusion (melted metal) zone for a given weld cooling rate and composition. Weld modeling also relies on thermal modeling efforts to relate the microstructure evolution to welding process effects. *See* MODEL THEORY.

Heat-affected zone modeling. Changes in the heat-affected zone, which reaches temperatures just below the melting point of the material, are mainly due to the response of the material to the thermal effects of solid-state phase transformations. If the material is a single-phase material, the grains in the heat-affected zone will grow as shown in **Fig. 12***a*. This may lead to a reduction in strength of the heat-affected zone. Grain growth kinetics as a function of weld thermal cycles can be described using analytical models or Monte Carlo simulations. *See* MONTE CARLO METHOD.

If the microstructure contains many phases, the changes may involve transformation from one phase to another based on composition and temperature gradients. These changes are closely linked to the thermodynamic stability of the phases and the diffusivity of the various elements in the phases.

Stainless steel that has both ferrite (body-centered cubic crystal structure) and austenite (face-centered cubic structure) phases undergoes changes as shown

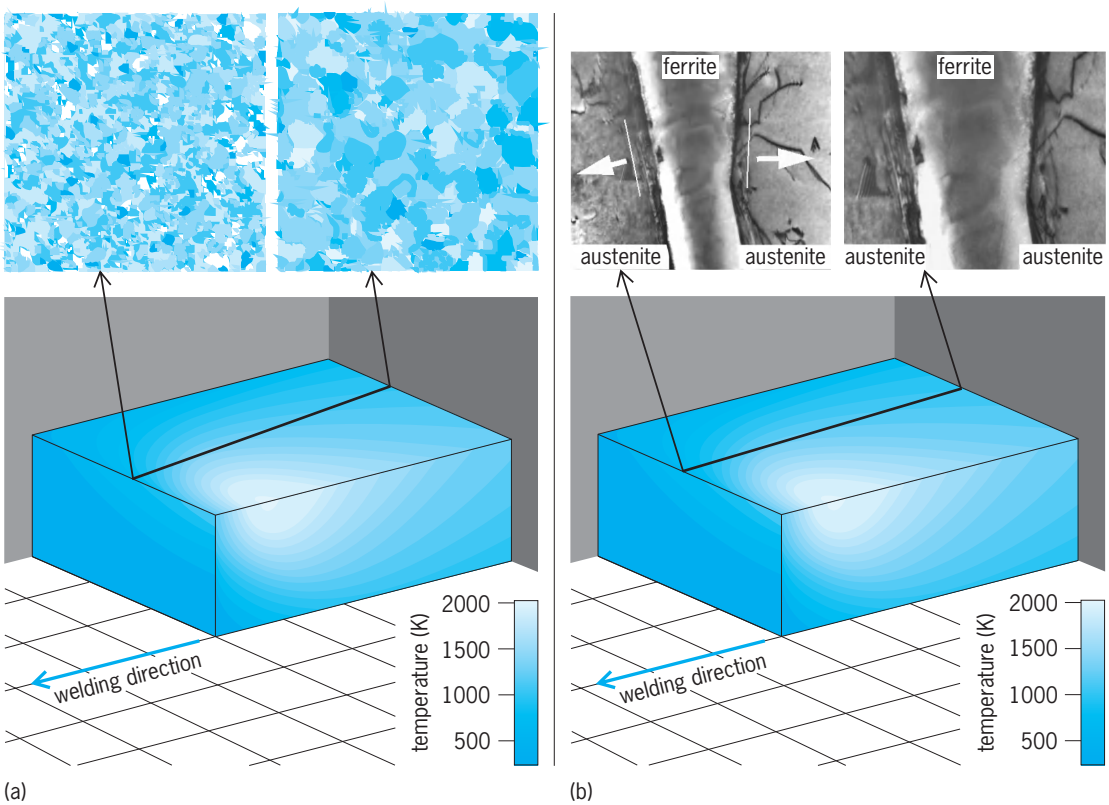

Fig. 12. Schematic representation of (*a*) grain growth in a single-phase material (colors indicate grains with different orientations) and (*b*) phase growth in a two-phase material (arrows in the multiphase materials show the direction of the boundary movement).

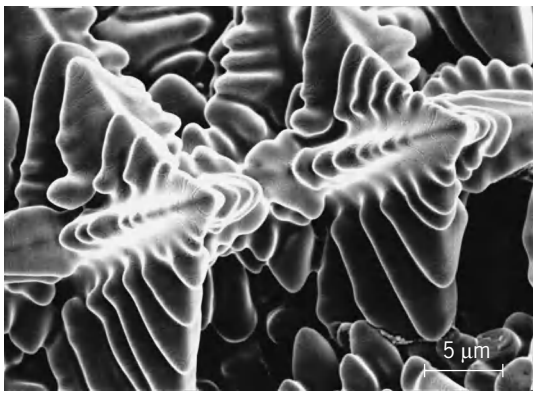

Fig. 13. Scanning electron micrograph showing a dendritic microstructure in a nickel-base superalloy weld.

in Fig. 12*b*. As the temperature increases, the stability of ferrite in the weld increases, and the ferrite grows into the austenite, leading to a larger fraction of ferrite. This growth involves diffusion of chromium, nickel, and carbon in both the phases. For the other compositions or thermal cycles of stainless steel, the ferrite may dissolve and the austenite may grow during welding. These complex changes may occur during weld cooling. The effects during heating and cooling do not cancel each other. The above changes may be beneficial or deleterious depending upon the service requirements of the welds. *See* STAINLESS STEEL.

Some of the solid-solid transformations in the heat-affected zone that occur in low-alloy (carbon) steel without diffusion (displacive transformations) can be modeled with thermodynamics and analytical expressions for the rate of the interface motion.

Fusion-zone modeling. Modeling the fusion zone is complicated by the need for a description of liquid to solid transformation (solidification), as well as the subsequent solid-state transformation. Some of the microstructural details in the fusion zone that are important are grain structure, grain morphology, partitioning of elements between liquid and solid, selection of phases in multiphase alloys, and physical defects such as cracks. During weld solidification, the solid phase grows epitaxially from the grains at the fusion line. The solidification substructure may be in the form of cells or dendrites, depending upon the thermal and compositional conditions near the solid-liquid interface (**Fig. 13**). With the growth of these dendrites, the alloying elements partition to the liquid, according to their thermodynamic properties. As the solidification continues, the partitioning of elements becomes more severe, leading possibly to nucleation of new grains in the liquid metal. At the end of solidification, cooling to room temperature results in other solid-state transformations. Modeling the fusion zone is complex, because in addition to solidification and solid-state transformation, the influence that solidification has on subsequent solid-state transformations must be considered. *See* CRACKING; EPITAXIAL STRUCTURES.

Grain structure evolution in single-phase materials has been modeled using Monte Carlo techniques, as well as some of the new modeling methods that use thermodynamic calculations to describe the solidification characteristics of multicomponent and multiphase alloys. For example, the fusion zone of stainless steel may contain either the austenite phase or the ferrite phase or a mixture of both, depending upon the composition and the cooling rate. The presence of ferrite in welds can be beneficial to avoid cracking but may be deleterious for corrosion resistance. Therefore, there is a need to develop models to describe the competition between these phases during welding. The relative thermodynamic stability between the ferrite and austenite phase can be correlated to the amount of ferrite that remains at room temperature. In terms of composition, adding chromium to stainless steel results in more ferrite formation, and adding nickel to stainless steel results in more austenite formation in welds. *See* CORROSION.

Thermodynamic and kinetic models are increasingly used to describe microstructure evolution in welds and in integrated models that describe the overall performance of the welds.

S. S. Babu; J. M. Vitek; S. A. David

Thermal cutting. Thermal cutting is used for a wide variety of work of the heavier type, such as structural work and shipbuilding, where the highest degree of dimensional accuracy is unimportant. Thermal cutting can be done manually or automatically, and in either case the torch can be brought to the work. The cut can be straight-line, curved, or beveled. The principal cutting process is oxygen cutting, which can be used for most mild steels. The cutting torch tip has a central orifice for oxygen and peripheral orifices for the preheat flame (**Fig. 14**). The flame preheats the steel to the ignition temperature, whereupon the oxygen converts the steel to iron oxide, which issues from the bottom of the cut as a molten spray. An oxygen cutting machine can cut a 1-in. (25.4-mm) steel plate at a speed of 18 in./min (46 cm/min), using an oxygen orifice of 0.0595-in. (1.51-mm) diameter and an oxygen pressure of 35 lb/in.^{2} (240 kilopascals) and consuming oxygen at the rate of 150 ft^3/h (4.25 m^3/h). Oxygen lancing is similar to cutting

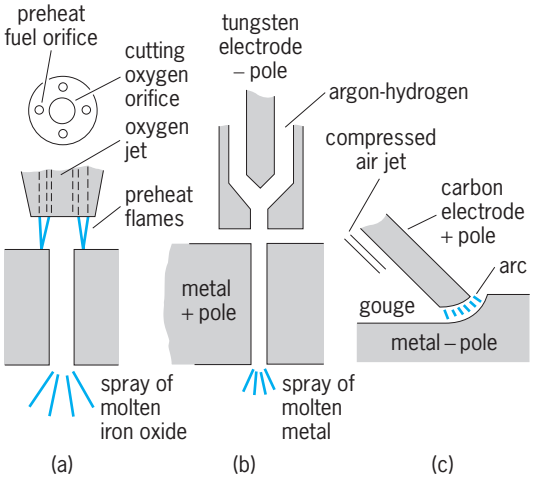

Fig. 14. Three thermal cutting processes. (*a*) Oxygen cutting. (*b*) Gas tungsten-arc cutting. (*c*) Air carbon-arc cutting.

Typical conditions for plasma arc cutting of aluminum

Thickness, in. (mm)	Speed, in./min (cm/min)	Orifice diameter, in. (mm)	Current, A	Gas flow,* ft^3/h (m^3/h)
$^1/_4$ (6.35)	300 (762)	$^1/_8$ (3.2)	300	80 (2.3) Ar–40 (1.1) H_2 130 (3.7) N_2–30 (0.85) H_2
$^1/_2$ (12.7)	200 (508)	$^1/_8$ (3.2)	300	65 (1.8) Ar–35 (0.99) H_2 140 (4.0) N_2–60 (1.7) H_2
1 (25.4)	90 (23)	$^5/_{32}$ (4.0)	400	65 (1.9) Ar–35 (0.99) H_2 140 (4.0) N_2–60 (1.7) H_2
2 (50.8)	20 (51)	$^5/_{32}$ (4.0)	400	65 (1.9) Ar–35 (0.99) H_2 140 (4.0) N_2–60 (1.7) H_2
3 (76)	15 (38)	$^3/_{16}$ (4.8)	450	130 (3.7) Ar–70 (2.0) H_2 140 (4.0) N_2–60 (1.7) H_2
4 (102)	12 (30)	$^3/_{16}$ (4.8)	450	130 (3.7) Ar–70 (2.0) H_2 140 (4.0) N_2–60 (1.7) H_2

*Either the mixture with argon-hydrogen or with nitrogen-hydrogen may be used.

except that an oxygen pipe replaces the torch for cutting steel up to 6 ft (1.8 m) thick. For highly alloyed and stainless steel cutting, it may be necessary to use a special process such as flux injection or powder cutting or some of the newer arc cutting processes described below. The selection of cutting process and type of operation (manual or mechanized) depends on the material that is being cut and the ultimate use of the product. In powder cutting, iron powder is injected into the oxygen in one process to increase the iron oxide content of the slag and thus lower its melting point.

There are seven cutting processes which utilize an arc. Some have been replaced by better and more efficient methods. These include oxygen arc cutting, air carbon-arc cutting, gas metal and gas tungsten-arc cutting, and plasma arc cutting.

In oxygen arc cutting, metals are severed by a combination of melting, oxidizing, diluting, fluxing, and mechanical ejection. For oxidation-resistant metals, the cutting mechanism is more of a melting action. Oxygen arc cutting electrodes were developed primarily for use in underwater cutting and were later applied for cutting in air. In both applications, they can cut ferrous and nonferrous metals in practically any thickness and position.

Air carbon-arc cutting is essentially a progressive melting of material which is blown away by high-velocity jets of compressed air. Base-metal melting is much less than in oxygen cutting. The process is primarily physical rather than chemical; hence it can be used on most metals.

Gas metal-arc cutting was developed soon after the commercial introduction of the gas metal-arc welding process. Conventional gas metal-arc welding equipment can be used for cutting, although some modifications have been made in the equipment designed especially for cutting operations. Gas tungsten-arc cutting with a nonconstricted arc can be used to sever nonferrous metals and stainless steel in thicknesses up to $^1/_2$ in. (12.7 mm), using standard gas tungsten-arc welding equipment. Metals cut include aluminum, magnesium, copper, silicon bronze, nickel, copper nickel, and various stainless steels. This method can be used for either manual or mechanized cutting. Good-quality cuts can be made with an argon-hydrogen gas mixture containing 65–80% argon. Nitrogen can also be used, but the quality of cut is not as good as that produced with an argon-hydrogen mixture.

Plasma arc cutting employs an extremely high-temperature, high-velocity constricted arc between the electrode in the plasma torch and the piece to be cut. Where inert gases are used, the cutting process depends upon thermal action alone. When cutting such materials as mild steel and cast iron, increased cutting speeds can be achieved by using oxygen-bearing cutting gases. The process can be used to cut any metal (see **table**). The plasma cutting process can be used in the general fabrication of all metals in many industries, such as shipyard, aircraft, chemical, nuclear, and pressure-vessel. It also finds application in the forging and casting industry. The process can be used in manual or mechanized operation and is especially adaptable to automation.

The laser is a valuable tool for drilling, cutting, and eventually milling of virtually any metal or ceramic or other inorganic solid. The mechanism by which a laser beam removes material from the surface being worked usually involves a combination of melting and evaporation, although with some materials, such as carbon and certain ceramics, the mechanism is purely one of evaporation. Any solid material can be cut with the laser beam. In addition to the pulsed ruby laser, continuous CO_2 and pulsed YAG lasers are of importance in drilling, cutting, and trimming. The potential for laser cutting of reactive metals has been greatly enhanced by a technique where oxygen is added to the continuous-laser cutting process. On especially reactive material the cutting rate and penetration are greatly improved because of the exothermic reaction and creation of additional heat

due to the oxidation of the reactive material. *See* LASER. Mel M. Schwartz

Bibliography. American Society for Metals, *Metals Handbook*, 10th ed., vol. 6: *Welding, Brazing, and Soldering*, 1993; American Welding Society, *Welding Handbook*, 8th ed., vol. 1, 1987, vol. 2, 1991, vol. 3, 1980; T. J. Blazynski, *Explosive Welding, Forming and Compaction*, 1983; H. Cary, *Modern Welding Technology*, 3d ed., 1993; H. Cerjak (ed.), *Mathematical Modeling of Weld Phenomena*, vols. 1–4, Institute of Materials, London, 1993, 1995, 1997, 1998; B. Crossland, *Explosive Welding of Metals and Its Application*, 1982; Deere and Co. Staff, *Welding*, 3d ed., 1991; D. S. Duvall, W. A. Owczarski, and D. F. Paulonis, TLP bonding: A new method for joining heat resistant alloys, *Welding J.*, 53(4):203–214, 1974; Ø. Grong, *Metallurgical Modelling of Welding*, 2d ed., Institute of Materials, London, 1997; T. B. Jefferson and D. T. Jefferson (eds.), *Welding Encyclopedia*, 18th ed., 1988; R. A. Lindberg and N. R. Braton, *Welding and Other Joining Processes*, 1985; G. R. Purdy (ed.), *Fundamentals and Applications of Ternary Diffusion*, Pergamon Press, New York, 1990; C. Stinchcomb, *Welding Technology Today: Principles and Practices*, 1989.

Well

An excavation made to extract water, oil, gas, brine, or other fluid substance from the earth. Wells are the source of about one quarter of the water supplies in the United States, all gas and oil, and most of the industrial brines and sulfur. Water wells are recorded in the earliest historic documents and probably originated during periods of drought when ancient humans attempted to reach water by digging a shallow excavation at the site of springs that had ceased to flow or in the dry channels of rivers. Later they learned to dig deeper for water where it did not issue at the land surface. The first wells were dug by hand, and some large deep ones were provided with elaborate ramps which enabled those drawing water to walk or even drive a donkey down to the water level. In the Near East for thousands of years wells (khanats) have been constructed by tunneling nearly horizontally into the outwash gravel to tap ground water which flows down the tunnel to its entrance. Horizontal wells are also used in Hawaii to skim fresh ground water from underlying salt ground water.

Most modern wells are of the drilled type. In the past the Chinese developed a crude percussion drill which enabled them to dig wells a few inches in diameter, lined with bamboo, to depths of a few thousand feet. In 1833 at Grenelle in Artois, France, a well was drilled to a depth of 1798 ft (548 m) and lined with steel casing. Its location provided the term artesian, which originally referred only to flowing wells. Drilling for brine in the United States began in West Virginia in 1808, and for oil and gas in Pennsylvania in 1859. Drilling for water in the United States began in the 1820s, but did not become common practice until the latter part of the nineteenth century. Dug wells are almost obsolete, because of the greater speed of drilling and the greater efficiency of drilled wells. *See* ARTESIAN SYSTEMS; OIL AND GAS WELL DRILLING.

Drilled wells, commonly 2–36 in. (5–90 cm) in diameter, usually are fitted with a steel tube or casing inserted in the drilled hole to the desired depth. Where the water-bearing formation is competent to stand without support, the casing is set, or finished, at the top of solid rock. Where there is danger of caving, as in sand or gravel, the casing is carried below the top of the water-bearing bed, and a perforated pipe or screen extends below the casing to the bottom of the bed. In fine-grained sand an envelope of coarse sand or fine gravel is commonly placed around the screen. In coarser-grained aquifers the screen is directly in contact with the aquifer. In either case the construction includes a considerable period of pumping, surging, or other treatment (called well development), during which the finer particles of the formation are drawn into the well and removed. This process substantially increases the initial yield of the well, in some cases by several times. Well screens may become encrusted with deposition of minerals from the ground water. Surging, acid treatment, or use of small explosive charges may be required to renew the well yield.

Most wells of large capacity are equipped with pumps of the deep-well turbine type to lift the water to the surface. However, where the water surface is less than 20 ft (6 m) below the land surface, they may be equipped with suction-type pumps. For efficiency the pump and motor should be selected to fit the expected discharge and lift. Many small-capacity farm and ranch wells are fitted with reciprocating pumps powered by windmills. Domestic wells are commonly equipped with electrically powered submersible centrifugal or jet pumps. If the well casing is bent or damaged, an air lift may be used. *See* PUMPING MACHINERY.

When a well is pumped, the pressure head at the well is lowered and a hydraulic gradient toward the well is established which causes water to flow toward the well. This lowering of head is called drawdown. One commonly used method for judging the ultimate yield of a well is to measure the yield per unit of drawdown; this is the specific capacity of the well, commonly expressed in gallons per minute per foot of drawdown as determined from pump tests after the well is completed. *See* GROUND-WATER HYDROLOGY. Albert N. Sayre; Ray K. Linsley

Bibliography. J. Balek, *Groundwater Resources Assessment*, 1989; S. N. Davis and R. J. M. DeWiest, *Hydrogeology*, 1991; D. K. Todd, *Groundwater Hydrology*, 2d ed., 1980.

Well logging

The technique of making measurements in drill holes with probes designed to measure the physical and chemical properties of rocks and their contained fluids. Over a million holes are drilled annually around

TABLE 1. Logging methods

Logs and curves	Property investigated	Primary uses
Electrical		
Self-potential	Natural currents	Defining sand/shale sequences and bed thickness; determining correlation water salinity
Normal resistivity	Resistivity	Shallow investigation of invaded zone to aid in quantifying hydrocarbon aiding saturations; correlation
Lateral resistivity	Resistivity	Moderate to deep investigation of formation; used on older logging devices
Induction	Resistivity	Moderate to deep investigation of formations in high-resistivity muds
Focused	Resistivity	Shallow to deep investigation of formations in relatively low-resistivity muds
Microresistivity	Resistivity	Very shallow investigation to detect permeable rocks and aid in measurement of porosity
Dielectric constant	Dielectric properties	Distinguishing fresh water from oil in open hole
Microresistivity imaging	Resistivity	Obtaining two-dimensional microresistivity measurements giving gray scale or color pattern of borehole wall to produce corelike images; especially useful in identifying fractures, analyzing thinly bedded formations or structural features, and recognizing special features such as conglomerate pebbles and vugs
Acoustical		
Velocity	Interval transit time of compressional waves	Determining porosity, lithology, abnormal pressures; constructing synthetic seismograms; calibrating seismic records; generation of near-borehole seismic profiles
Wave train	Attenuation of acoustic waves	Detecting fractures, gas zones; determining cement bonding, mechanical properties of rocks
Nuclear		
Gamma ray	Natural radioactivity	Determining shale content, correlation, bed thickness; locating radioactive tracers
Spectral gamma ray	Energy spectrum of natural radioactivity	Identifying radioactive elements to help determine lithology
Density	Bulk density	Determining porosity; helping to identify gas zones
Neutron	Hydrogen content	Determining porosity; helping to identify gas zones
Pulsed neutron capture	Capture cross section of elements	Distinguishing between salt water and hydrocarbon behind casing
Geochemical spectrometry	Relative abundance of certain elements	Identifying clay types and other lithologies; detecting oil behind casing in presence of fresh water
Nuclear magnetism	Hydrogen atoms in larger pore spaces	Identifying productive zones; distinguishing heavy oil from fresh water
Production		
Flow meter	Fluid flow	Determining location and rate of fluid entry
Fluid density	Density of fluid in borehole	Determining type of fluid
Temperature	Temperature	Locating fluid entry, especially gas
Radioactive tracer	Fluid	Tracing flow path of radioactive material released in borehole
Noise	Fluid entry and flow	Detecting fluid movement inside or behind pipe and locating entry
Capacitance	Type fluid in borehole	Diagnosing fluid interfaces
Casing inspection	Electromagnetic, acoustic, or properties of casing	Evaluating corrosion, damage, or perforated holes in casing
Other		
Caliper	Borehole diameter	Estimating cement volumes behind casing; supplementing log analysis
Dipmeter	Azimuth and inclination of borehole and correlation of beds in borehole	Determining dip and strike of bedding, and geometry of hole
Borehole gravity meter	Bulk density	Determining density further from borehole; calibrating surface gravity surveys

the world to probe for oil, gas, solid minerals, and fresh water, and to better understand the subsurface geology. Much information can be obtained from samples of rock brought to the surface in cores or bit cuttings, or from other clues while drilling, such as penetration rate; but the greatest amount of information comes from well logs. *See* GEOPHYSICAL EXPLORATION.

Well logs result from a probe lowered into the borehole at the end of an insulated cable. While less common and generally more expensive, certain logs may also be recorded while drilling by means of special tools in the drill string above the bit (known as measurement while drilling, or MWD). Whether by cable or by measurement while drilling, the results are recorded graphically or digitally as a function of depth. These records are known as geophysical well logs, petrophysical logs, or more commonly well logs, or simply logs.

Recording information from boreholes has become more challenging as holes deviate radically from the vertical. It is becoming routine to log holes that are deviated 90° at a selected depth and penetrate over 2000 ft (610 m) in a horizontal direction.

An enormous number of logging devices have been developed since the first experiments with borehole resistivity measurements in 1927. From a few dozen basic logging tools, several hundred different types and variations of curves can now be recorded; however, in any one hole only about a dozen or fewer curves are commonly run. An outline of logging methods and some of their main applications are shown in **Table 1**.

In addition to these logging methods, there are a number of closely related electromechanical tools that can be run on the same wire line used for logging. Although these tools perform important functions in the borehole, they are not actually well logs. Some of these tools and their primary purposes are listed in **Table 2**.

Although the most common uses of logs are for correlation of geological strata and location of

TABLE 2. Electromechanical devices run on logging cables

Tool	Purpose
Sidewall sampling	Retrieving small core samples of formations from the side of an open borehole at selected intervals by using steel cylinders fired into the borehole wall (percussion samples) or a small rotary coring bit that penetrates the formation horizontally from a tool suspended in the well bore (rotary sample)
Wire-line tester	Testing formation pressures and fluids at selected intervals from cased hole or open hole
Perforators	Perforating casing for completion with bullets or shaped charge jets
Pipe cutters	Severing drill pipe or casing with mechanical, explosive, or chemical devices
Directional survey tool	Determining azimuth and inclination of borehole at selected intervals
Wire-line packer	Setting a plug in casing or tubing
Free-point indicator	Determining where pipe is stuck
Wireline shot	Loosening pipe so that it may be backed off threaded joint above where it is stuck
Fluid sampler	Catching a sample of fluid at a selected depth in borehole
Television or photographic cameras	Inspecting casing or other borehole conditions
Impression block	Determining nature of top of "fish" or lost tools in hole

hydrocarbon zones, there are many other important subsurface parameters that need to be detected or measured. Also, different borehole and formation conditions can require different tools to measure the same basic property. In petroleum engineering, logs are used to identify potential reservoir rock; determine bed thickness; determine porosity; estimate permeability; locate hydrocarbons; estimate water salinity; quantify amount of hydrocarbons; estimate type and rate of fluid production; estimate formation pressure; identify fracture zones; measure borehole inclination and azimuth; measure hole diameter; aid in setting casing; evaluate quality of cement bonding; locate entry, rate, and type of fluid into borehole; and trace material injected into formations (for example, to detect artificial fractures or to monitor injected fluids during supplementary recovery of hydrocarbons). In geology and geophysics, they are used to correlate between wells; locate faults; determine dip and strike of beds; identify lithology; deduce environmental deposition of sediments; determine thermal and pressure gradients; create synthetic seismograms; calibrate seismic amplitude anomalies to help identify hydrocarbons from surface geophysics; calibrate seismic with velocity surveys; generate vertical seismic profiles to give a more detailed seismic picture near the borehole; and calibrate gravity surveys with borehole gravity meter. Other applications include locating fresh-water aquifers; locating solid materials; and studying soil and rock conditions for foundations of large structures.

Electrical devices. These employ instruments that generate data based on electrical measurements. *See* ELECTRICAL MEASUREMENTS; ROCK, ELECTRICAL PROPERTIES OF.

Spontaneous potential. The spontaneous potential, usually recorded along with resistivity curves, is a simple but valuable aid to help geologists correlate from one well to another and to assist them in inferring the depositional environment of the sediments. It defines permeable zones from surrounding nonpermeable shales and can be used to estimate the salinity of formation water. In the early 1930s, observations were made in boreholes of naturally occurring currents. These currents, now referred to as spontaneous potential, self-potential, or simply SP, are recorded from a simple electrode on the probe with a reference electrode usually grounded at the surface. Spontaneous potential values opposite shales are relatively constant; however, when the electrode passes a permeable bed containing water more saline than the borehole fluid, the spontaneous potential will deflect to the left or in the negative direction (**Fig. 1**). The greater the salinity contrast between formation water and borehole water-based mud, the greater the spontaneous potential deflection. It is this electrochemical property that allows estimates of formation water salinity to be made.

Borehole size and bed thickness along with the density, chemical composition, and invasion of mud into the formation will also affect the magnitude of the spontaneous potential. The presence of hydrocarbons in the permeable zone may suppress the spontaneous potential a slight amount, but there is still formation water trapped in the smaller pores which will contribute to the spontaneous potential deflection.

Clays in permeable sand will also tend to suppress the spontaneous potential. Because the amount of clays deposited with sands can give valuable clues to the environment of deposition, the spontaneous potential shape is used by geologists to help piece together the subsurface geology. An example of the spontaneous potential curve, along with several other curves, is shown in Fig. 1. Note how the shaly sands tend to suppress the spontaneous potential toward the bottom of the lower sand interval. This sand could have been deposited as part of an ancient marine beach.

Resistivity. While resistivity was the first logging measurement to be made, it is still one of the most important physical properties to record. It is easy to visualize that a sand filled with salt water will conduct electricity much easier (and therefore have a low resistivity) than a sand with most of the salt water replaced with nonconducting oil or gas. The basic measurement of resistivity, the ohm-meter, is the electrical resistance of a cube 1 m on a side; it is the reciprocal of conductivity. Like most log measurements, the resistivity log response depends not only upon water saturation but also upon secondary properties of the rock/fluid system. Quantification

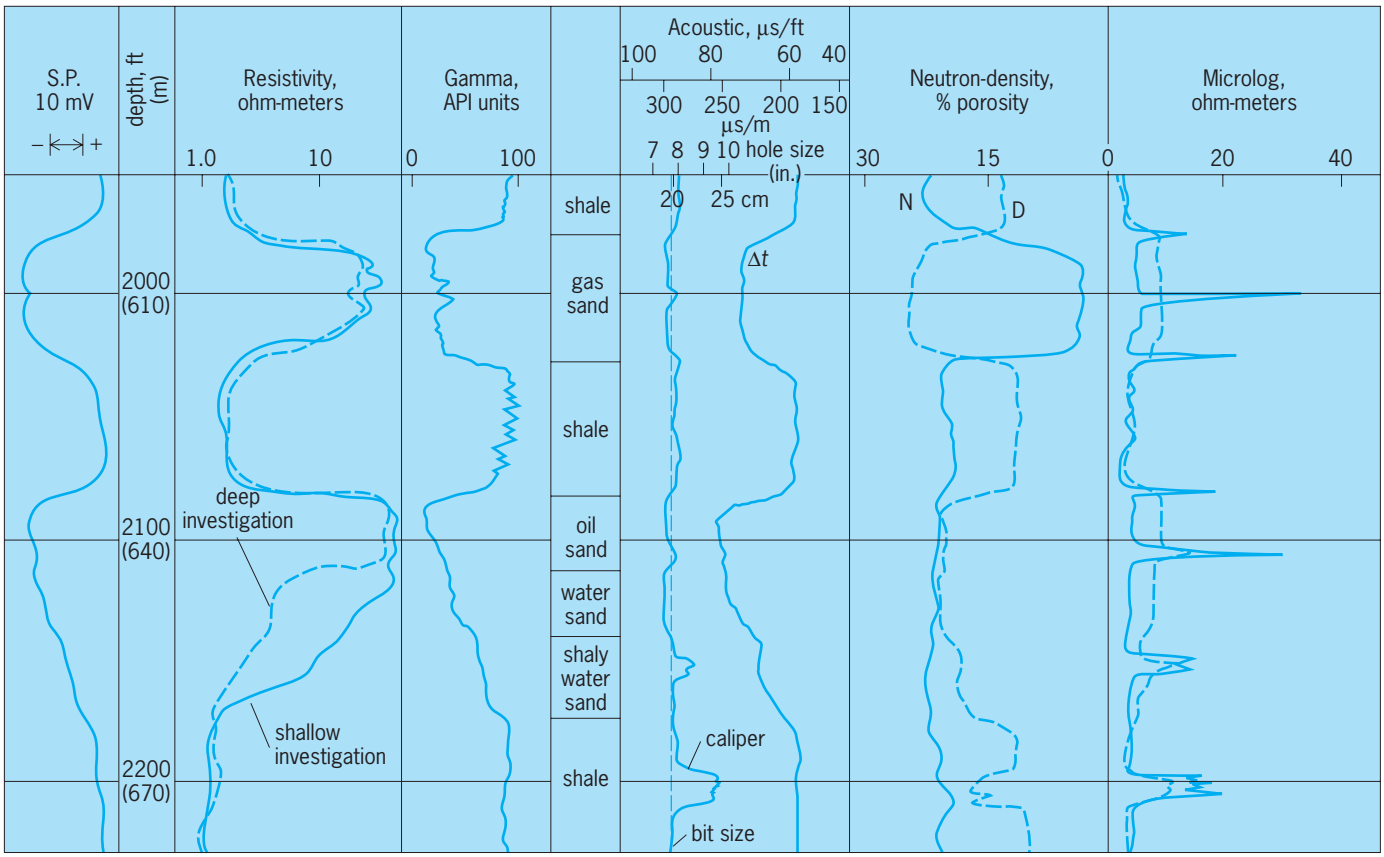

Fig. 1. Example of log response in sand/shale sequence.

of the water saturation S_w depends on knowledge of the rock pore space or porosity ϕ, cementation exponent m, saturation exponent n, and water resistivity R_w. In addition, the response characteristics of the particular tool for the surrounding physical environment, including mud and formation geometry, should be understood. In 1942 G. E. Archie published an empirical relation between water saturation, rock and fluid properties, and true resistivity R_t. This classic equation (1) is still widely used.

$$S_w = \left[\frac{R_w \phi^{-m}}{R_t}\right]^{1/n} \tag{1}$$

Where significant amounts of conductive minerals such as clay are also in the formation, allowance must be made for their influence on the total resistivity. Many improvements have been made over the years in tools to record resistivity, but the primary purpose—to determine hydrocarbon saturation ($1 - S_w$) by calculation of the true formation resistivity—remains the same.

Some resistivity devices (listed below) are designed to investigate deep into the formation away from the borehole, while others respond to the formation near the borehole. In general, the resistivity tools with shallow investigation can define thin beds better, but are more influenced by the annular zone of mud filtrate invasion. The resistivity tools with deep investigation approach true resistivity, but still may need correction because of the influence of filtrate invasion, borehole geometry, and bed thickness.

1. Normal and lateral. The short normal, long normal, and lateral curves were common on older electrical logs and had shallow, intermediate, and deep investigations, respectively. The short normal, which has 16–18 in. (41–46 cm) spacing, is still used on the induction electrical log. In spite of its shallow investigation, it has been one of the basic geological correlation tools.

2. Induction log. The induction log is probably the most commonly run logging tool today; it is designed for medium and deep investigation. As shown in **Fig. 2**, this tool consists of one or more coils that generate an alternating electromagnetic field. The induced currents which are dependent upon the conductivity of the formation are then detected by receiver coils. Multiple coils are used to focus the measurement deep into the formation; this minimizes effects of the borehole, invaded zone, and nearby zones. The induction log works best in relatively resistive borehole fluids such as fresh-water or oil-based muds.

Modern induction tools are arrayed to enhance vertical resolution and radial investigation. Beds as thin as 1 ft can be resolved while providing images of formation resistivity that reflect bedding, hydrocarbon content, and invasion profiles. The display of this information can be enhanced be color images.

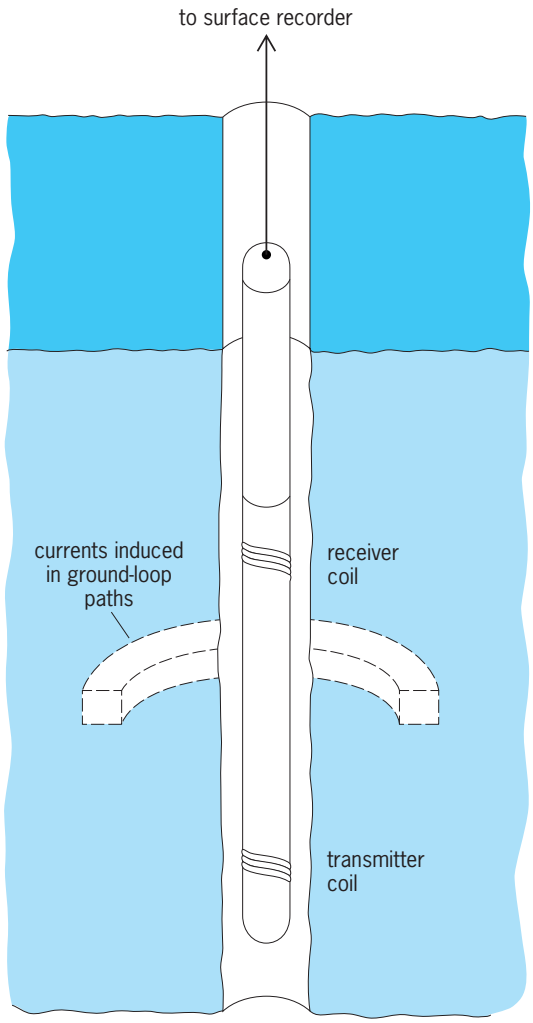

Fig. 2. Schematic of basic two-coil induction system.

3. Focused log. Where the borehole mud is more saline than the formation water, a focused log such as the Laterolog is recommended. In this tool a sheet of current from an electrode is forced into the formation by bucking currents from surrounding electrodes. This bucking current keeps the measuring current thin and allows excellent resolution of thin beds.

4. Microresistivity. There are several microresistivity tools where relatively shallow investigations of resistivity are made from pads pressed against the borehole wall. These auxiliary logs aid in the interpretation of true resistivity and porosity. One of these tools, the microlog (also known as minilog or contact log), is of particular interest. It is the oldest of all pad devices. This tool combines two resistivity measurements, one measuring about $1^1/_2$ in. (4 cm) away from the pad and the other about 4 in. (10 cm). When the pad is pressed against mud cake built up over a permeable zone, a separation appears between the two curves (Fig. 1). Shales or other nonpermeable zones do not have such separation; therefore, under favorable conditions the microlog can give excellent vertical resolution of permeable zones.

5. Microresistivity imaging. A tool to image the borehole wall electrically was introduced in 1985. This tool employs a dense array of sensors on a four-arm pad tool similar to the dipmeter. These sensors give a high-resolution electrical image of varying borehole wall conductivity that is displayed with a variable-intensity gray scale or in color. This gives an image of the borehole similar to a core photograph. The tool is especially useful in determining microheterogeneities such as fractures, bedding detail, pebbles, and vugs. This tool may give clues to productivity and depositional environment that, in absence of a core, could be missed by the other well-logging tools.

Acoustical devices. Quantitative interpretation of hydrocarbon saturation requires a knowledge of porosity. One of the most widely used porosity logs is the acoustic log, also known as a sonic log, acoustilog, or velocity log. This log is used for many other purposes, including identification of lithology, prediction of pressures, and assisting with geophysical interpretation. In addition, information from the amplitude of the acoustic wave aids in detection of fractures and gas zones, determination of mechanical properties of rocks, and analysis of cement bond behind pipe. *See* ACOUSTIC EMISSION.

The acoustical log measures the shortest time for sound waves to travel through 1 ft (or 1 m) of formation. This interval travel time [Δt; Eq (2)], measured

$$\Delta t = \frac{t_{(R_1)} + t_{(R_4)} - t_{(R_2)} - t_{(R_3)}}{4} \qquad (2)$$

in microseconds per foot (or per meter), is the inverse of velocity. A sketch of a borehole compen-

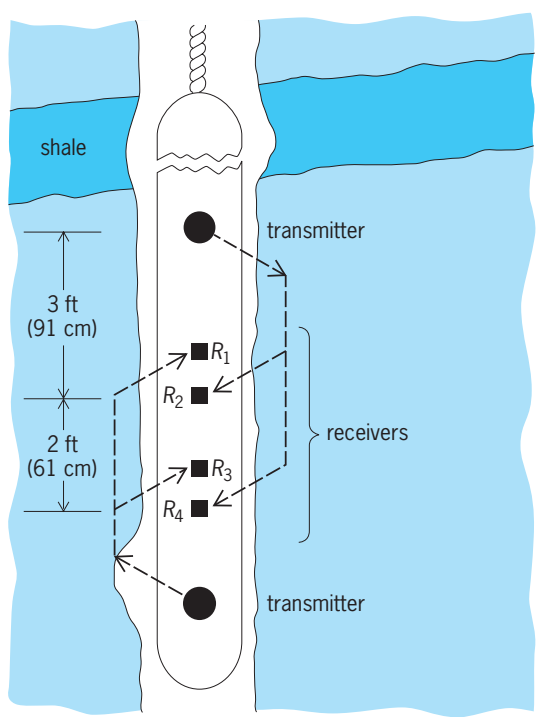

Fig. 3. Diagram of borehole compensated acoustic tool.

sated acoustic tool is shown in **Fig. 3**. The dual receivers with each transmitter are arranged so that travel time through mud and hole rugosity will not detract from measuring formation velocity. The formation travel time is dependent upon the porosity ϕ, the travel time of the fluids in the pore space (Δt_f), and the travel time of the rock matrix (Δt_{ma}). An empirical formula for clean sands known as the Wyllie time average equation relates porosity ϕ_{sv} to the sonic velocity interval travel time Δt of a water-filled rock as shown in Eq. (3). While Δt depends only

$$\phi_{sv} = \frac{\Delta t - \Delta t_{ma}}{\Delta t_f - \Delta t_{ma}} \tag{3}$$

upon the time of the first arrival of the compressional wave, other information may be extracted from the amplitude of the compressional wave and later arrivals.

In open hole logging, the amplitude of the acoustical wave train may be attenuated by natural fractures or faults. In cased hole logging, the cement bond log is an important aid in determining cement bonding in the annular space between the casing and the formation. This tool measures the amplitude of an acoustical signal imposed upon the casing. A good cement bond against the pipe will prevent it from ringing and therefore will greatly attenuate the amplitude of the compressional wave before it reaches the receiver. Supplemental information, including a display of acoustical wave trains, gives further insights into the degree of bonding.

Acoustical devices are closely linked to seismic interpretation. From the acoustical log (and density log if available) a synthetic seismic trace may be created to help correlate the seismic section with the detailed well-bore logs. By lowering multiple acoustical pickup devices or geophones into the borehole, while activating energy sources on the surface, acoustical travel time may be measured for seismic calibration. A similar arrangement is used for vertical seismic profiling, which generates a more detailed seismic section near the borehole. Vertical seismic profiling displays may also be generated from energy sources in the borehole, with geophones placed in the same borehole, at the surface, or in a nearby borehole (cross-well tomography). *See* SEISMIC EXPLORATION FOR OIL AND GAS.

Nuclear devices. These are instruments which generate data based on measurements of nuclear particles.

Gamma ray. In most cases the gamma-ray log may be considered as a shale log; that is, clays or shaly rocks will give a higher radioactive count than clean sands or carbonates. Like the spontaneous potential, this provides a good correlation curve, defines bed thicknesses, and aids interpretation of environmental deposition. It is usually a more diagnostic indicator of beds in a carbonate sequence than the spontaneous potential, and unlike spontaneous potential, can detect formations behind casing. The gamma log is relatively unaffected by borehole conditions. Often it is used to trace material injected into a formation, such as a radioactively tagged sand used as a proppant in a hydraulic fracture. Measurement is made in terms of arbitrary units, API (American Petroleum Institute) units. *See* GAMMA-RAY DETECTORS.

Spectral gamma ray. Almost all the naturally occurring radioactivity that is detected by the gamma-ray log comes from three elements: potassium, thorium, and uranium. Although it is usually assumed that high radioactive values on the gamma log come from potassium-rich shales, the spectral analysis log allows verification of the contribution from each element. Experience with this log has shown that a great number of "shales" are, in reality, uranium-rich stringers in sands, dolomites, and limestones. Some of these formations that were originally assumed to be impermeable shales have turned out to be hydrocarbon productive zones. The spectral breakdown also allows for more sophisticated geological environmental interpretation.

Density. Another important porosity tool is the density log; this tool emits a beam of gamma rays into the rock from a source such as cesium-137. These gamma rays interact with electrons in the formation through Compton scattering. The resulting lower-energy gamma rays are sensed by two detectors above the source. Although the tool responds to electron density, for most formation rocks the apparent bulk density is practically identical to the actual bulk density. For a few substances such as sylvite and rock salt (and to a lesser extent gypsum, anhydrite, and coal), small corrections are needed to arrive at true bulk density values. With knowledge of the bulk density ρ_b and reasonable assumptions of rock matrix density ρ_{ma} and fluid density ρ_f, the porosity derived from a density log (ϕ_D) can be calculated from Eq. (4).

$$\phi_D = \frac{\rho_{ma} - \rho_b}{\rho_{ma} - \rho_f} \tag{4}$$

An improved version of the density tool includes a measurement of photoelectric absorption. The magnitude of this measurement depends primarily upon lithology. However, if whole mud with barite, which has a very high photoelectric cross section, fills open natural fractures in a formation, the tool will respond accordingly. This effect, combined with high vertical resolution, can be used to identify open natural fractures.

Neutron. Another important device for determination of porosity is the neutron log. These logs respond primarily to hydrogen atoms. Therefore, in clean or shale-free formations the neutron log reflects the amount of liquid-filled porosity. Because the neutron log does not recognize gas-filled porosity, it can be compared with another porosity device such as the density to detect gas-filled zones. It can also be used in combination with other porosity logs for interpretation of lithology, including shaly sands. In neutron logging, high-energy neutrons are emitted from a source such as plutonium and beryllium. When these neutrons collide with hydrogen atoms which have a nucleus of nearly the same mass, the

greatest amount of energy is lost. When the neutrons have slowed down to thermal velocities after successive collisions, they are subject to capture by nuclei of other atoms; this causes an emission of gamma rays. Either these gamma rays or the low-energy neutrons are counted by the detector in the sands. *See* NEUTRON.

Pulsed neutron capture. Besides resistivity, another distinquishing physical characteristic is that the chlorine atom in salt water has a significantly higher capture cross section than the hydrogen or carbon atoms in oil or gas. Since resistivity measurements of the formation cannot be made in cased holes, a new tool sensitive to capture cross section was developed to distinguish hydrocarbon from salt water behind casing. The pulsed neutron capture tool employs a neutron generator that repeatedly emits pulses of high-energy neutrons. This is done by using a miniaccelerator to bombard a tritium target with deuterium atoms. After these neutrons are slowed down to the thermal state, they are captured by nuclei of the various atoms surrounding the tool. With each capture a corresponding emission of gamma rays occurs; these can be detected a short distance from the source. The total or bulk capture cross section (Σ_t) of the formation is the sum of the component cross sections of the rock matrix (Σ_{ma}) and the fluids such as water (Σ_w) and hydrocarbons (Σ_h) within the rock pores. This may be expressed as Eq. (5). Therefore, water

$$\Sigma_t = \Sigma_{ma}(1-\phi) + \Sigma_w S_w \phi + \Sigma_b (1-S_w)\phi \qquad (5)$$

saturation S_w can be estimated where there is sufficient contrast in capture cross section between the water and hydrocarbons, and reasonable estimates can be made of the other variables. This tool has proved to be a valuable aid in identifying bypassed oil stringers in cased holes. Also it is used to monitor changes in the saltwater level within a producing horizon.

Geochemical spectrometry. There are various tools and combinations of tools designed to measure the abundance of elements commonly found in reservoir rocks. In general, these tools are designed to enhance but not replace the more traditional logging tools. Tools such as the carbon/oxygen log or gamma spectrometry tool can measure elements such as carbon, oxygen, calcium, silicon, hydrogen, iron, and sulfur and capture cross sections (sigma). In addition to giving more detailed lithologic information, these tools can directly measure hydrocarbons in zones with fresh or unknown-salinity waters in open hole or through casing. Since oil is composed largely of carbon, an abundance of this element in relation to others can provide a diagnostic indication of oil in the formation. This tool is similar to the pulsed neutron capture tool in that it utilizes a pulsed 14-MeV neutron source and gamma-ray detector. However, it is optimized for production of desired prompt gamma rays from inelastic scattering of neutrons by carbon in the formation. Because changes in lithology can also affect the carbon/oxygen ratio, a ratio of calcium/silicon is recorded to help distinguish a change in carbonate content from hydrocarbon saturation.

The geochemical logging tool combines many of the features of other tools, including the gamma spectrometry tool, the natural spectral gamma-ray tool, and the compensated neutron tool (which carries a special californium neutron source). In addition to the elemental analysis associated with such tools, a measurement can be made of aluminum weight concentration. Aluminum is a key element in clay minerals, micas, and feldspars. Titanium and gadolinium are also identified with this tool. It is possible to estimate concentrations of a missing element, for example, magnesium, not normally measured by spectrometric methods. Estimates of magnesium are important to identify dolomites. The wide variety of elemental data generated by a geochemical tool is valuable for detailed evaluation of a formation.

Nuclear magnetism log. The nuclear magnetism tool provides important auxiliary information for evaluation of subsurface fluids. Basically, measurements are made by applying a strong, magnetic field to the formation for a limited time (about 2 s). After the imposed field is relaxed, measurements are made of the free precession of protons (hydrogen atom nuclei) in pore space immediately adjacent to the borehole. Two independent parameters are derived from these measurements: free-fluid index I_{Ff} and thermal relaxation time T_1. Water molecules in very fine pores such as those found in clays will not respond. The magnitude of I_{Ff} is a value which approaches total porosity in clean formations with relatively large pores. The value of T_1 is dependent upon fluid–solid interactions and pore size. This tool is of great importance because of its ability to estimate permeability (at least within an order of magnitude). By killing signals from water treated with paramagnetic ions that have been injected into an oil zone, the nuclear magnetism log can measure the amount of residual oil with great accuracy.

Production logging. A very important suite of tools has been developed for production logging. These tools, usually run after the casing has been set, are designed to locate fluid movement behind the casing and into the well bore, to detect the type of fluid, and to determine the flow rate. Production logging tools include various types of flow meters, fluid density devices, sensitive thermometers, radioactive tracer devices, noise loggers, and capacitance logs. In addition, these tools usually record some correlation curve such as gamma-ray or collar locator to tie the production curves to the formation and casing. Closely related to production tools are logging devices that magnetically, acoustically, or mechanically inspect the condition of the casing in the borehole.

Other devices. There are a number of miscellaneous devices that give additional information.

Caliper. The caliper log is an important auxiliary curve that records the diameter of the hole with depth. This is essential for estimating the amount of cement needed in the annular space between the

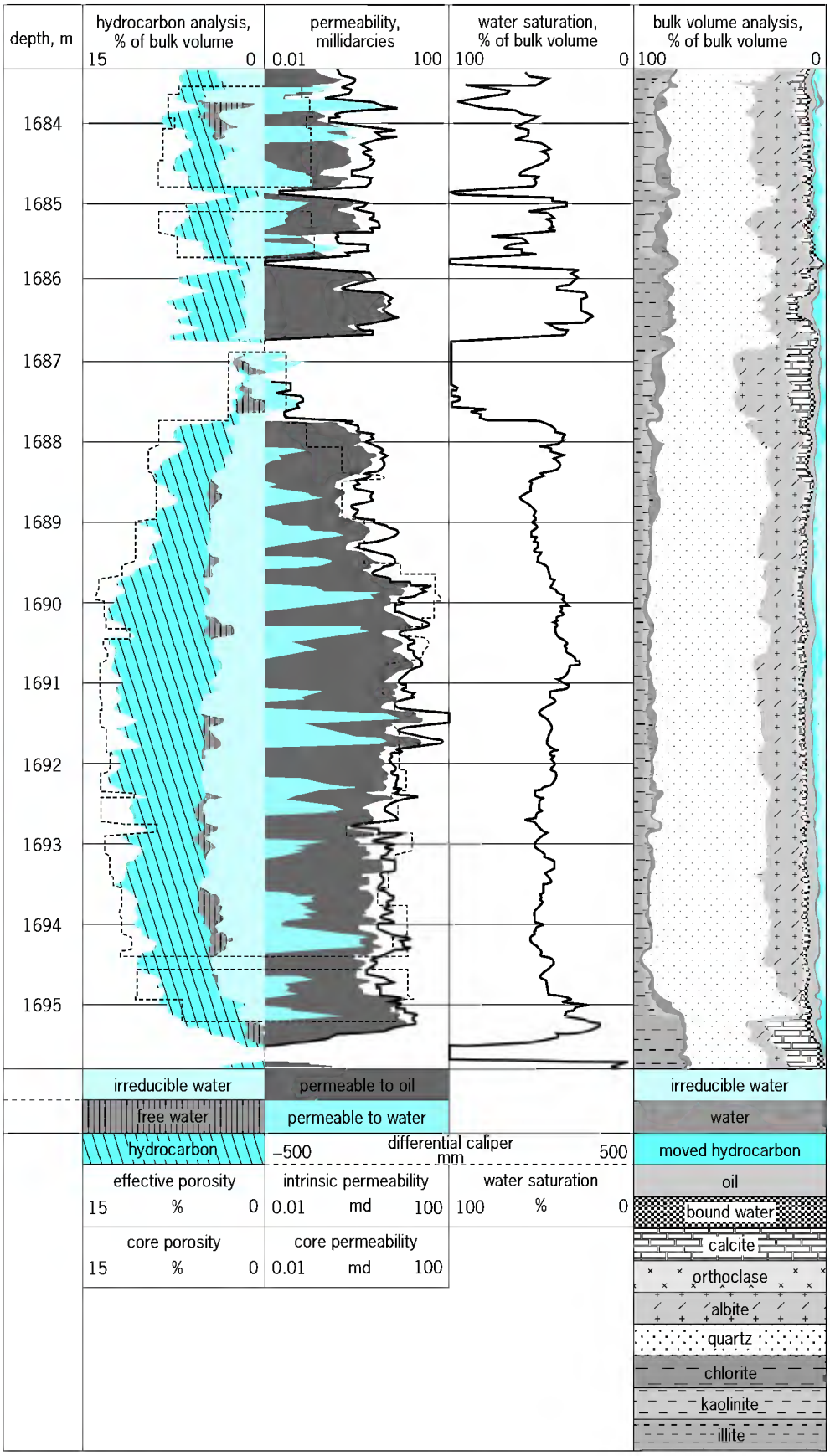

Fig. 4. Example of a computed log. In the hydrocarbon analysis column, the broken line represents core porosity. In the permeability column, the broken line represents core permeability, and the solid line represents intrinsic permeability. 1 m = 3.3 ft; 1 mm = 0.04 in. md = millidarcies.

casing and formation. It is also aids the interpretation of other curves where hole diameter or hole rugosity affects the readings. When a mud cake builds up over a permeable zone, the caliper can often detect such permeable streaks with excellent vertical resolution. The caliper is commonly run with each of the porosity tools.

Dipmeter. The dipmeter has three or four pads pressed against different sides of the hole. In the most common type of tool, these pads record microresistivity curves which can be correlated to determine depth displacement of formations from one side of the borehole to another. This information, combined with borehole size, deviation, and azimuth, is used to determine formation dip and azimuth. Such information is useful in correlating from well to well and in structural and stratigraphic interpretations. In addition, information necessary to plot the course of the borehole is provided by this tool.

Borehole gravity meter. Precise measurements of relative gravity can be made at intervals in the borehole to give estimates of formation density. These measurements are determined by large volumes of rocks that extend from tens to hundreds of feet away from the borehole. They are unaffected by borehole conditions such as mud type, hole rugosity, invaded zone, or casing and cement. Primary applications of the borehole gravity surveys include detection of significant porous zones or ore bodies away from the borehole, detection of oil and gas zones behind casing, and vertical density profiling to help interpret surface gravity mapping.

Dielectric constant. The dielectric constant log reflects the relative change in the dielectric constant of the rock fluids being surveyed. Since the relative dielectric constant of fresh water is significantly greater than oil, this tool can be used to distinguish oil from fresh water. The tools developed for open-hole logging measure phase and amplitude of a 16- and 30-MHz electromagnetic wave and at much higher frequencies such as 1.1×10^9 Hz. At these high frequencies, conductivity becomes much less important than dielectric permittivity, unlike the much lower frequencies used in the induction log. Because of the very shallow depth of investigation, this tool works best in smooth boreholes.

Computer analysis of logs. Well-log analysis has taken full advantage of the emergence of digital computing and the ability to take small, rugged, and powerful computers to the well site. Log data are recorded on magnetic disks or tapes as they are run, and measurements are provided through mathematical analysis. In addition, well-log data may be transmitted from the well site to the oil company headquarters through a high-speed digital satellite communication network. **Figure 4** illustrates an advanced petrophysical analysis with lithologic, fluid volume, and reservoir permeability calculations. Computation has been performed on a foot-by-foot basis with measurements from electric, nuclear, acoustic, and nuclear magnetism logs as well as rock cuttings from the borehole. Of particular note is the estimation of irreducible (nonmovable) water volume and log-derived permeability by interpretation of the nuclear magnetism and additional porosity logs. Rock-core measurements show good agreement with in-place log evaluation. In addition to analysis for open-hole logging, special programs are designed for cased-hole logging, solid mineral analysis, dip and structure of formations, synthetic seismic traces, and many others. In fact, the digital information from practically every curve can be incorporated into some type of analysis or interpretation program. *See* DIGITAL COMPUTER; MICROCOMPUTER.

Tools for measuring basic physical and chemical properties from a wellbore are now sophisticated and reliable. These tools combined with computer-aided modeling and interpretation programs have greatly enhanced the ability of oil and gas companies to evaluate subsurface geology, fluid content, and productivity. Richard E. Wyman

Bibliography. T. Darling, *Well Logging and Formation EvaIuation (Gulf Drilling Guides)*, 2005; J. R. Hearst et al., *Well Logging for Physical Properties: A Handbook for Geophysicists, Geologists, and Engineers*, 2d ed., 2000; J. R. Jorden and F. L. Campbell, *Well Logging II: Electric and Acoustic Logging*, 1986; S. M. Luthi, *Geological Well Logs: Their Use in Reservoir Modeling*, 2001; R. Ransom, *Practical Formation Evaluation*, 1995.

Wellpoint systems

A method of keeping an excavated area dry by intercepting the flow of ground water with pipe wells located around the excavation area. Intercepting the flow before it reaches the excavated area also improves the stability of the edge of the excavation, permitting steeper bank slopes and often eliminating the need for supporting or shoring the banks. *See* CONSTRUCTION METHODS; GROUND-WATER HYDROLOGY; WELL.

Wellpoint systems are most effective in coarse-grained soils, such as gravel or sand. They are not effective in fine soils, such as silts and clays, where the small size of the pores between grains restricts the flow of water.

The basic components of a wellpoint system are the wellpoint, the riser pipe, the header pipe or manifold, and the pump (**Fig. 1**).

The wellpoint consists of a perforated pipe about 4 ft (1.2 m) long and about 2 in. (5 cm) in diameter. It is equipped with a ball valve to regulate the flow of water, a screen to prevent the entry of sand during pumping, and a jetting tip. The steel riser pipe to which the wellpoint is attached is slightly smaller in diameter than the wellpoint. Its purpose is to bring the ground water to the surface, where it is collected by the horizontal manifold pipe or header pipe, which is 6–10 in. (10–25 cm) in diameter. Wellpoints and riser pipes are spaced 3–6 ft (0.9–1.8 m) apart depending upon the characteristics of the ground. The wellpoint and riser pipe are

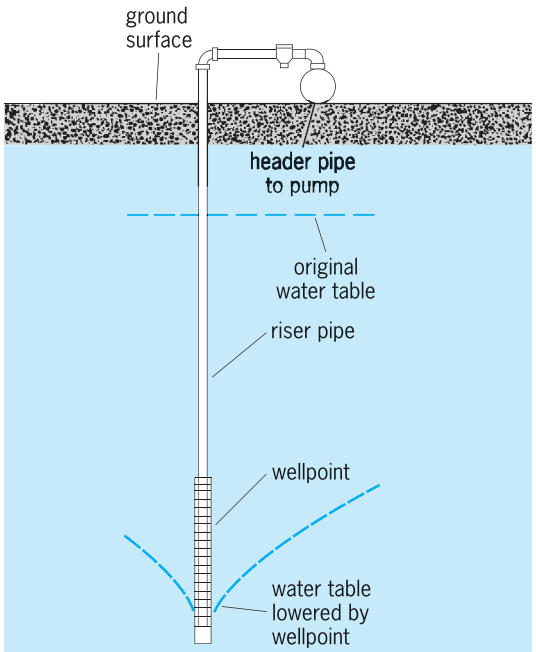

Fig. 1. Components of a wellpoint system.

Fig. 2. Layout of a three-tier wellpoint system. (*Griffin Wellpoint Co.*)

usually driven into the ground to the required depth, generally 2–5 ft (0.5–1.5 m) below the excavation, by jetting, that is, by displacing the ground below it with water under pressure.

Where the ground-water stratum is overlaid with hard clay or heavy gravel, jetting may not be able to pierce the tough material and a special boring attachment is used to carry the wellpoint down.

The pumps are located above the water table and collect the water from the header pipes for discharge away from the excavation area. The pumps are centrifugal pumps, with vacuum pumps to increase the suction lift. The suction lift in each wellpoint rarely exceeds 18–22 ft (5–7 m). Consequently, where the depth of the wellpoint below ground level exceeds this depth, it is necessary to install two or more tiers of wellpoint systems, each with its own assembly of wellpoints, header pipes, and pumps. A three-tier wellpoint system is shown in **Fig. 2**. *See* PUMP; PUMPING MACHINERY. William Hershleder

Welwitschiales

An order of the class Cycadopsida having one species, *Welwitschia mirabilis*. This plant is native to the very arid deserts of southwestern Africa. Its appearance is bizarre—it has a very short, unbranched, woody stem (sometimes to 3 ft or 1 m in diameter), which is cushion- or saucer-shaped and tapers quickly to a long taproot. There are only two leaves, and these persist throughout the life of the plant. The leaf is broadly strap-shaped, as wide as the stem, firm, and leathery, and gradually splits lengthwise between the veins. The leaf (of indefinite growth) develops from a meristem at its point of connection to the stem. The species is dioecious; the cones are borne on branched axes originating between the crown of the stem and the base of the leaf. The order is known only from the one living species and not from any fossils. *See* CYCADOPSIDA; PINOPHYTA; PLANT KINGDOM. Thomas A. Zanoni

Wentzel-Kramers-Brillouin method

A special technique for obtaining an approximation to the solutions of the one-dimensional time-independent Schrödinger equation, valid when the wavelength of the solution varies slowly with position; also known as the WKB method. It is named after G. Wentzel, H. A. Kramers, and L. Brillouin, who independently in 1962 contributed to its understanding in the quantum-mechanical application. It was however, studied earlier by J. Liouville (1837), Lord Rayleigh (1912), and H. Jeffreys (1923). It is also called the BWK method, and the JWKB method, the classical approximation, the quasiclassical approximation, and the phase integral method.

The system to be considered is a particle of mass m moving nonrelativistically in a potential $V(x)$. The Schrödinger equation for stationary states of energy E is Eq. (1), where $p^2(x)$ is defined by Eq. (2) and $\hbar$ is Planck's constant divided by 2π.

$$\frac{d^2\psi}{dx^2} + \frac{p^2(x)\psi}{\hbar^2} = 0 \qquad (1)$$

$$p^2(x) = 2m[E - V(x)] \qquad (2)$$

Two types of regions are to be considered as illustrated in **Fig. 1**: (I) The classically allowed type in which $p^2(x)$ is positive. Here $p(x)$ is defined to be the positive root, and it is the magnitude of the classical momentum. (II) The classically unallowed type in which $p^2(x)$ is negative. Here $p(x)$ is defined to be $i|p(x)|$. Typically $p^2(x)$ goes linearly through zero at the point where the region changes from classically allowed to unallowed. Such a point x_1, where $p^2(x_1) = 0$, is a turning point of the classical motion.

The solutions of differential equation (1) can be written, in a formal way, as Eq. (3). In this expression one uses all the upper signs uniformly or all the lower signs uniformly to obtain the two independent solutions. The constants in the indefinite integrals are the integration constants for the differential equation. The symbol $\sqrt{\ }$ indicates that the root of all that follows is to be taken, choosing the sign as suggested by the convention given above for $p(x)$. One verifies this formal expression of the solutions through Eqs. (4). Thus if the process indicated

$$\psi = \exp\left[\pm\frac{i}{\hbar}\int dx\sqrt{p^2(x)} \pm i\hbar\frac{d}{dx}\sqrt{p^2(x)} \pm i\hbar\frac{d}{dx}\sqrt{\cdots}\right] \quad (3)$$

$$\frac{d}{dx}\exp\left[\ \right] = \left[\pm\frac{i}{\hbar}\sqrt{p^2(x)} \pm i\hbar\frac{d}{dx}\sqrt{\cdots}\right]\exp\left[\ \right] \quad (4a)$$

$$\frac{d^2}{dx^2}\exp\left[\ \right] = \left[\pm\frac{i}{\hbar}\frac{d}{dx}\sqrt{p^2(x)\pm\cdots}\right]\exp\left[\ \right] + \left[-\frac{p^2(x)}{\hbar^2} \pm \frac{i}{\hbar}\frac{d}{dx}\sqrt{\cdots}\right] + \exp\left[\ \right] = -\frac{p^2(x)}{\hbar^2}\exp\left[\ \right] \quad (4b)$$

in Eq. (3) converges, it produces the solutions of the differential equation.

The WKB approximation to the solutions is found by evaluating the process to two orders only, Eq. (5). In a classically unallowed region the approximate solutions are usually written as Eq. (6).

$$\psi \simeq \exp\left[\pm\frac{i}{\hbar}\int dx\sqrt{p^2(x)\pm i\hbar\frac{dp}{dx}(x)}\right] \simeq \exp\left[\pm\frac{i}{\hbar}\int dx\left\{p\pm\frac{i\hbar}{2p}\frac{dp}{dx}\right\}\right] = \frac{1}{\sqrt{p}}\exp\left[\pm\frac{i}{\hbar}\int dx\,p(x)\right] \quad (5)$$

$$\psi = \frac{1}{\sqrt{|p|}}\exp\left[\mp\frac{1}{\hbar}\int dx|p(x)|\right] \quad (6)$$

Evidently the requirement for the applicability of the approximation is that $\hbar\, dp/dx$ should be small compared to p^2. In terms of the local de Broglie wavelength $\lambda = h/p$, this requirement is that $d\lambda/dx$ should be small compared to 2π. The WKB approximation therefore applies in the case of slowly varying wavelength.

Connection formulas. The approximation ordinarily applies in any region which does not include turning points, but it breaks down as $p^2(x)$ approaches zero, and in fact the approximate solution diverges as $p^{-1/2}$ at a turning point. There are connection formulas which relate the approximate solutions in a classically allowed region to the approximate solutions in a neighboring classically unallowed region. For the case illustrated in Fig. 1 the formulas are given by Eqs. (7). The integrals are always written in these

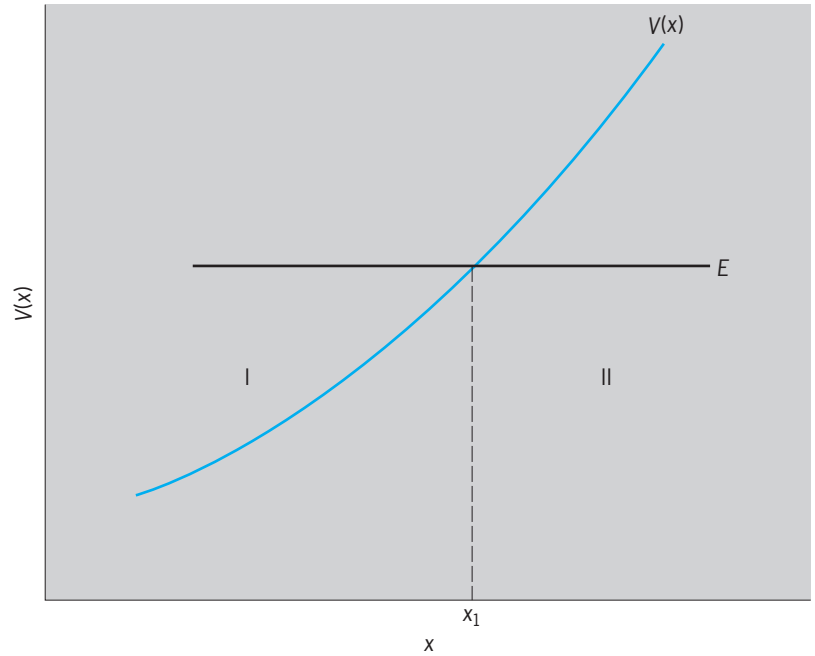

Fig. 1. Classically allowed region I to the left of a classically unallowed region II.

Fig. 2. Classically allowed region I to the right of classically unallowed region II.

$$\frac{1}{\sqrt{p}}\sin\left[\frac{1}{\hbar}\int_x^{x_1} p(\xi)\,d\xi + \frac{\pi}{4}\right] \leftrightarrow \frac{1}{2}\frac{1}{\sqrt{|p|}}\exp\left[-\frac{1}{\hbar}\int_{x_1}^{x}|p(\xi)|\,d\xi\right] \quad (7a)$$

$$\frac{1}{\sqrt{p}}\cos\left[\frac{1}{\hbar}\int_x^{x_1} p(\xi)\,d\xi + \frac{\pi}{4}\right] \leftrightarrow$$

$$\frac{1}{\sqrt{|p|}}\exp\left[\frac{1}{\hbar}\int_x^{x_1} |p(\xi)|\,d\xi\right] \quad (7b)$$

formulas so that they are increasing functions of x away from the turning point. Similar formulas, given in Eqs. (8), apply when the allowed region is to the right, as illustrated in **Fig. 2**. The proof of these formulas is too long to be given in this article.

$$\frac{1}{2}\frac{1}{\sqrt{|p|}}\exp\left[-\frac{1}{\hbar}\int_x^{x_2} |p(\xi)|\,d\xi\right] \leftrightarrow$$

$$\frac{1}{\sqrt{p}}\sin\left[\frac{1}{\hbar}\int_x^{x_2} |p(\xi)|\,d\xi + \frac{\pi}{4}\right] \quad (8a)$$

$$\frac{1}{\sqrt{|p|}}\exp\left[\frac{1}{\hbar}\int_x^{x_2} |p(\xi)|\,d\xi\right] \leftrightarrow$$

$$\frac{1}{\sqrt{p}}\cos\left[\frac{1}{\hbar}\int_{x_2}^{x} p(\xi)\,d\xi + \frac{\pi}{4}\right] \quad (8b)$$

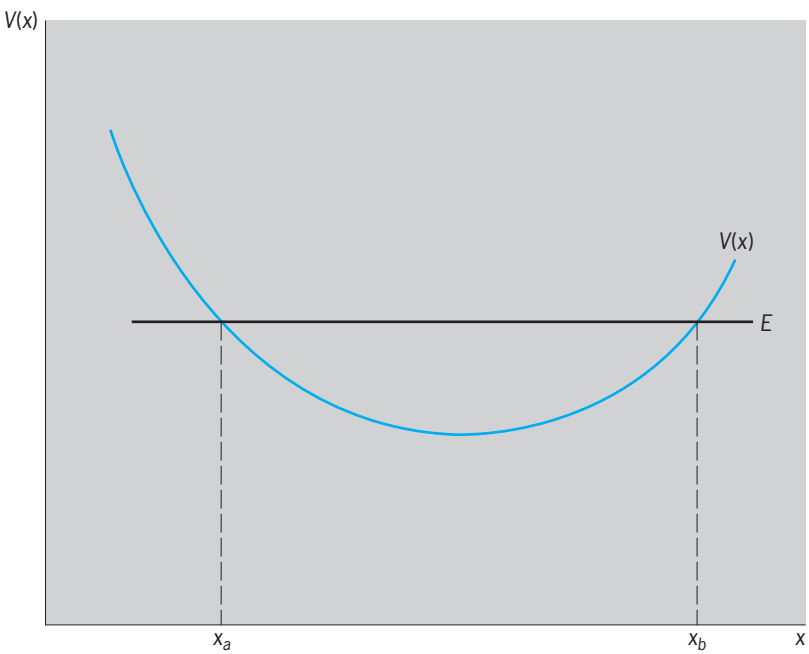

Fig. 3. Potential well with turning points at x_a and x_b.

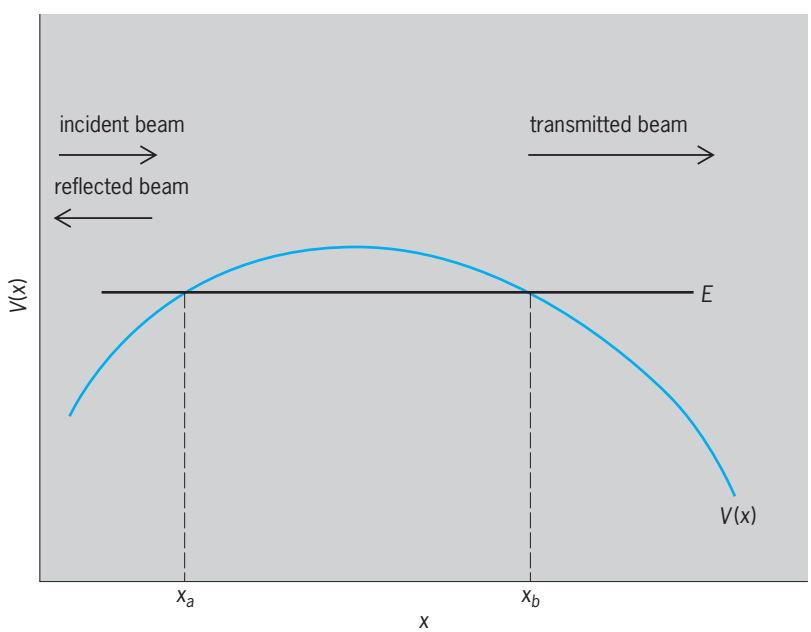

Fig. 4. Potential barrier with turning points at x_a and x_b.

Potential well. An important application of the WKB approximation is in obtaining an approximation to the allowed energy levels in a potential well. Suppose the well is as shown in **Fig. 3**. Only integrable functions, with decreasing exponential dependence, may occur in the classically unallowed regions to the left of x_a and to the right of x_b. Applying connection formulas (8*a*) and (7*a*) at the turning points x_a and x_b, one obtains two expressions for the approximate solution in the allowed region, Eq. (9). In order for these two expressions to coincide, it is necessary that Eq. (10) be satisfied, where

$$\frac{A}{\sqrt{p}}\sin\left[\frac{1}{\hbar}\int_{x_a}^{x} p(\xi)\,d\xi + \frac{\pi}{4}\right] = \frac{B}{\sqrt{p}}\sin\left[\frac{1}{\hbar}\int_x^{x_b} p(\xi)\,d\xi + \frac{\pi}{4}\right] \quad (9)$$

$$\frac{1}{\hbar}\int_{x_a}^{x_b} p(\xi)\,d\xi = \left\{n + \frac{1}{2}\right\}\pi \quad (10)$$

n indicates one of the values 0,1,2,3, ... This condition alternatively can be written as in Eq. (11). For

$$2\int_{x_a}^{x_b} \sqrt{2m\,[E - V(\xi)]}\,d\xi = \left\{n + \frac{1}{2}\right\}h \quad (11)$$

each allowed value of n there is an energy E_n that satisfies this condition; these are the allowed energy levels of the potential well, in the WKB approximation.

It is at this point that one can see the connection between modern quantum mechanics and the old quantum theory of N. Bohr and A. Sommerfeld. In the old quantum theory the allowed energy levels were found by setting the integral of the momentum, over a complete cycle, equal to an integer times Planck's constant. This condition coincides with Eq. (11), except that half-integer numbers are called for.

Barrier penetration. Another important application of the WKB approximation is in problems of barrier penetration. Consider a one-dimensional potential barrier as shown in **Fig. 4**, and suppose there is a particle beam of energy E incident on the barrier from the left. In the classical problem the particles of energy less than the peak potential energy would simply be reflected, but in the quantum-mechanical problem the particle beam is partly reflected and partly transmitted through the classically unallowed

region. The transmission coefficient D is defined as the ratio of the transmitted particle current to the incident current. By applying the connection formulas straightforwardly at the turning points and approximating for small D, one finds Eq. (12). This result

$$D = \exp\left[-\frac{2}{\hbar}\int_a^b |p(\xi)|\, d\xi\right] \qquad (12)$$

has wide applicability. It was used, for example, in the Folwer-Nordheim theory of field emission and in G. Gamow's explanation of alpha decay. *See* FIELD EMISSION; RADIOACTIVITY.

Generalization. The WKB approximation is based on the exponential function, the function that occurs in the exact solution of the free-particle Schrödinger equation. S. C. Miller and R. H. Good, Jr., developed a generalization of the WKB approximation, based on the functions that occur in the solution of an arbitrary comparison equation. The comparison equation is chosen so as to be similar to the equation under study and so as to be soluble, in the same spirit as perturbation theory. Qualitatively better approximations can be set up this way; for example, in the potential well problem, approximate solutions finite through both turning points can be found. *See* NONRELATIVISTIC QUANTUM THEORY; PERTURBATION (QUANTUM MECHANICS); QUANTUM MECHANICS. Roland H. Good, Jr.

Bibliography. V. P. Maslov, *The Complex WKB Method for Nonlinear Equations*, 1994; S. C. Miller and R. H. Good, Jr., A WKB-type approximation for the Schrödinger equation, *Phys. Rev.*, 91:174–179, 1953; E. L. Murphy and R. H. Good, Jr., WKB connection formulas, *J. Math. Phys.*, 43:251–254, 1964; D. Park, *Introduction to the Quantum Theory*, 3d ed., 1992; L. I. Schiff, *Quantum Mechanics*, 3d ed., 1968.

West Indies

An archipelago, including the Bahamas, the Greater Antilles, the Lesser Antilles, and other islands, curving 2500 mi (4000 km) from Yucatan Peninsula and southeastern Florida to northern Venezuela and enclosing the Caribbean Sea (see **illus.**). Situated between latitude 10° and 27°N and longitude 59° and 85°W, in the zone of the northeast trade winds, the West Indies have a subtropical and predominantly oceanic climate, with even warmth and steady breezes. Temperatures vary little from season to season, ranging from means of 80–85°F (27–29°C) in July to 70–78°F (21–26°C) in January at sea level. Freezing is unknown, and the hottest temperatures rarely exceed 90°F (32°C). Precipitation ranges from a low of 25–50 in. (64–127 cm) a year on low-lying islands and drier coasts up to 300 in. (7.6 m) on the highest peaks, which are almost perpetually cloud-capped. At lower elevations, rainfall is erratic from year to year and from season to season, but reaches a maximum in the summer and fall, when the northeast trades are replaced by light, variable winds. This is also the season of hurricanes, destructive tropical cyclones which sweep west and northwest across the Caribbean, sparing only the southernmost islands. The winter months are generally dry, and there is frequently a shorter dry season in July or August. *See* TROPICAL METEOROLOGY.

The West Indian flora is chiefly derived from Central and South America, but there are a number of endemic species, notably palms; many mainland plants failed to colonize the islands, which are floristically poor. The effect of isolation and small size is evident in the meager character of West Indian fauna. Animal species are limited; there are few mammals and no large ones, except for domesticated animals and, especially on Hispaniola, feral cattle, goats, pigs, and horses.

The pre-Columbian flora and fauna of the islands, originally densely wooded, were largely transformed by European conquest and colonization. Sugarcane and cotton now occupy much of the lowlands, while bananas, tobacco, cacao, citrus, coffee, Old World pasture grasses, and American, African, and Asian subsistence crops mantle the foothills and mountain slopes. Mangroves and other salt-water vegetation still fringe most of the islands; cactus, thorn bush, and dry scrub occur on protected dry leeward sites, while other low-lying areas are savannas with scattered palms and pines. Higher up, where droughts are briefer and less intense, mixed conifer and broadleaf forests exist, with rainforest above. Windswept and sodden summits support palm brake and elfin woodland. *See* COTTON; MANGROVE; SUGARCANE.

Bahama Islands. The Bahamas (4404 mi^2 or 11,406 km^2, population 244,692), together with the Turks and Caicos islands (166 mi^2 or 430 km^2, population 2287) to the southwest, make up about 3000 islands, rocks, and cays stretching more than 600 mi (960 km) southward from Grand Bahama Island, 50 mi (80 km) east of West Palm Beach, Florida. Most of the islands (Andros Island, 1600 mi^2 or 4100 km^2 is largest) lie on the Great and Little Bahama banks, slightly submerged limestone platforms separated from Florida and Cuba by the shallow Straits of Florida and Old Bahama Channel. Made of calcareous sand of marine origin and flanked on the east by a coral barrier reef, the Bahamas are flat and low, exceeding 200 ft (60 m) in height only on Cat Island (400 ft or 120 m). Precipitation is consequently scanty and unreliable, ranging from 40 in./year (102 cm/year) at Nassau, New Providence Island, to 25 in. (64 cm) at Grand Turk. What little soil exists is generally fertile, but the significant produce of land and sea is salt and sponges. Formerly a British colony, the Bahamas became independent in 1973; the Turks and Caicos islands are British.

Greater Antilles. These islands constitute the major share of the West Indies. They represent a partly submerged extension of an ancient Central American mountain chain, which stretches eastward from Hispaniola beneath 75-mi-wide (120-km) Mona Passage through Puerto Rico and the Virgin Islands to 40-mi-wide (64-km) Anegada Passage. Westward from Hispaniola the range splits in two. The southern

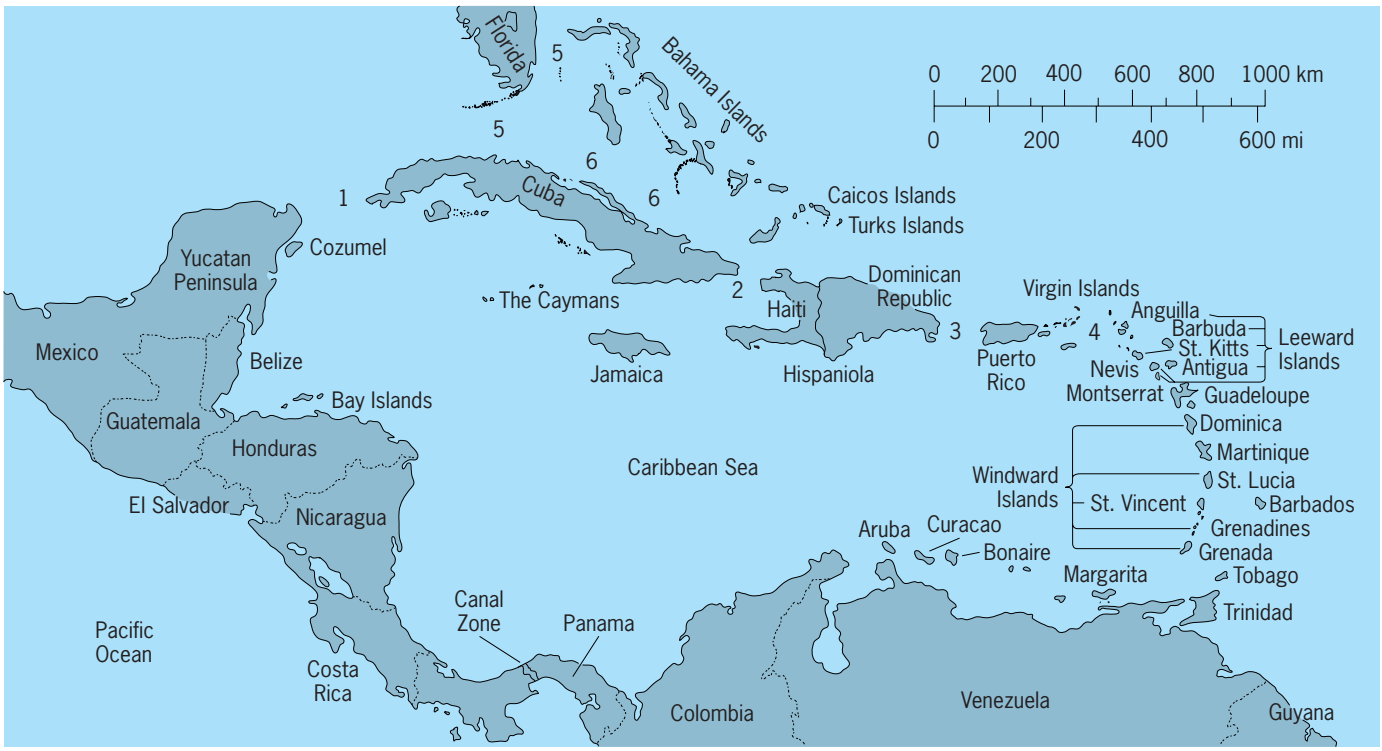

Island locations of the West Indies. Outstanding inter-island seaways are 1, Yucatan Channel; 2, Windward Passage; 3, Mona Passage; 4, Anegada Passage; 5, Straits of Florida; and 6, Old Bahama Channel. 1 mi = 1.6 km.

prong runs from the southwestern peninsula of Haiti into the sea, emerges in the Blue Mountains of Jamaica, and reappears to the west in Swan and Roatán islands off northeast Honduras. The northern prong runs from Haiti's northwest peninsula across 55-mi-wide (89-km) Windward Passage, reappearing in the Sierra Maestra of southeastern Cuba and again, further west, in the Cayman Islands (100 mi^2 or 259 km^2; population 10,419, British), low, flat limestone islands with fringing coral reefs, and reaches the shore of Central America north of the Bay of Honduras. Between these two prongs runs Cayman Trough (Bartlett Deep and Oriente Deep), which reaches 23,748 ft (7238 m) below sea level. North of the entire structure, separating Puerto Rico from the Bahama platform, is Puerto Rico Trench, 30,180 ft (9199 m) below sea level in Milkwaukee Depth. The extremes of elevation, the intense faulting and folding of most of the strata, gravity anomalies, and earthquakes all attest to the tremendous pressures to which the whole area has been subjected.

Cuba. With a length about 760 mi (1220 km) and area of 44,218 mi^2 (114,524 km^2; population 10,073,000), this is the largest of the West Indies. Only its southeastern mountain range, the Sierra Maestra, which reaches 6429 ft (1960 m) at Pico Turquino, links it structurally and physiographically with the other Greater Antilles. The rest of the long, narrow island is essentially a level or gently sloping limestone plateau resembling southern Florida, the Bahamas, and, west of the broad but shallow Yucatan Channel, the peninsula of Yucatan. The low-lying plateau is significantly broken by the Grupo de Guamuhayo (3792 ft or 1156 m at Pico San Juan) in the center of the island and by the Cordillera de los Organos at the western end, the latter a broken limestone country of karst topography, with landforms produced by chemical solution rather than by subaerial erosion. Lacking a central dividing mountain range, Cuba gets less rainfall (about 50 in. or 125 cm annually) than others of the Greater Antilles, but since there is no impediment to the trade winds, the south coast receives as much as the north. Only the western end—closest to the usual hurricane track—is much wetter. The typical vegetation is a thorn-scrub steppe of pine and palmetto interspersed with short grass. Among numerous offshore islands, the Isle of Pines (862 mi^2 or 2233 km^2), to the southwest, is largest. *See* KARST TOPOGRAPHY.

Jamaica. This island (4244 mi^2 or 10,991 km^2; population 2,214,120) is built up on an east-west trending igneous and metamorphic core, which reaches its greatest elevation in the Blue Mountains (Blue Mountain Peak, 7402 ft or 2256 m) at the eastern end of the island and is exposed at lower heights in the center. These steep and rugged mountains are subject to landslides and severe erosion where deforested. To the west and on their north and south flanks they are overlain, at about 3000 ft (900 m) above sea level, by a limestone plateau. But the limestone is so intricately dissected that little remains of the plateau surface. Karst topography is widespread and in the Cockpit Country of north-central Jamaica takes the form of sinkholes 500 ft (150 m) deep and a quarter mile (400 m) across. A by-product of weathering in some of the limestone glades is bauxite ore;

Jamaica is estimated to possess 600,000,000 commercially utilizable tons (550,000,000 metric tons). Thanks to their high alumina content, their surface occurrence, and their proximity to North American reduction plants, the Jamaican deposits are among the world's most productive. Level land is practically confined to alluvial plains, of which the largest is along the south-central coast. Formerly a British colony, Jamaica became an independent nation in 1962.

Hispaniola. With a total area of 29,530 mi^2 (76,484 km^2), the eastern two-thirds of this island is the Dominican Republic (population 5,518,430); the western third is Haiti (population 5,054,090). It is the most mountainous and structurally complex of the Greater Antilles. The Cordillera Central traverses the island from southeast to northwest at 8000–9000 ft (2400–2700 m), culminating in Pico Trujillo (about 10,200 ft or 3100 m), the highest elevation in the West Indies. There are two other major ranges. In the north, entirely within the Dominican Republic, the 4000-ft (1200-m) Cordillera Septentrionale runs eastward into the Peninsula of Samaná. In the southwest, another range forms the Tiburón Peninsula of Haiti and extends eastward into the Sierra Baharuco of the Dominican Republic; though narrower, it is almost as high as the Cordillera Central, reaching 7920 ft (2414 m) in the western Massif de la Hotte and 8793 ft (2680 m) toward the east in the Massif de la Selle. Between the southwest and central ranges runs a deep structural trench, the Cul de Sac, a continuation of the Cayman Trough between Cuba and Jamaica. The Cul de Sac contains two brackish lakes, of which the easternmost—Lake Enriquillo in the Dominican Republic—is 160 ft (49 m) below sea level.

Besides the Cul de Sac, the only extensive level areas in Hispaniola are the plain separating the northern and central ranges (Plaine du Nord in Haiti, the Cibão in the Dominican Republic), the extensive Seibo lowland in southeastern Dominican Republic, and the plain of the Artibonite River at the western end of the Cordillera Central. Above the last at about 1000 ft (300 m) lies a rolling intermontane basin, the Plaine Centrale. The largest offshore island, Ile de la Gonave (254 mi^2 or 658 km^2, population 27,000), lies in the Golfe des Gonaïves, almost encircled by the two western peninsulas; Ile de la Tortue (70 mi^2 or 181 km^2) lies off Cap Haitien to the north.

Hispaniola's climate is less oceanic than that of the other West Indies. The trade winds fail to penetrate to many sheltered locations, which warm up like continental areas in summer. Because of the heat, some of the plains are subject to drought even though they receive considerable rainfall; the Cul de Sac gets almost 50 in. or 127 cm annually, but evaporation is so intense that the vegetation is of semiarid type.

Puerto Rico. An intricately dissected east-west trending mountain core dominates the topography of this island (3423 mi^2 or 8866 km^2; population 3,187,550). The highest part rises to 4400 ft (1341 m) in Cerro de Punta. To the south is a narrow, arid coastal plain. The northern coastal plain, broader and comparatively well watered, is separated from the mountains by dissected limestone terraces, except in the northeast, where the Sierra de Luquillo (3500 ft or 1067 m), an outlier of the central range, reaches close to the sea. Little of the natural vegetation remains anywhere. A Spanish dependency until 1898, Puerto Rico, together with its off-shore islands, Mono (20 mi^2 or 52 km^2) to the west, Vieques (51 mi^2 or 132 km^2; population 10,000), and Culebra (11 mi^2 or 28 km^2; population 1000) to the east, has since been an American possession and is now a self-governing commonwealth under the United States.

Virgin Islands. These small islands with a total area of about 200 mi^2 (518 km^2) appear as eastern outliers of the Puerto Rico mountain mass but were submerged much more recently and hence display marine physical features—notable barrier reefs—superimposed on continental rocks. Anegada, at the extreme northeast of the Greater Antillean chain, is a flat coral reef raised 15 ft (4.5 m) above sea level; the other islands have steep slopes and a minimum of flat land, except for St. Croix, south of the main range. Maximum elevation, however, is only 1780 ft (543 m), at Mount Sage (Tortola); rainfall is low and erratic; vegetation consists of a scanty dry scrub woodland. The Virgin Islands are divided between the United States (St. Croix, St. Thomas, and St. John, total 133 mi^2 or 344 km^2; population 71,236) and Great Britain (Tortola, Virgin Gorda, Jost Van Dyke, and about 50 other islets of which six are inhabited, in all 59 mi^2 or 153 km^2; population 14,024).

Lesser Antilles. The numerous islands with a total area of about 2500 mi^2 or 6500 km^2 (population 1,406,000) compose an arc marking the eastern margins of the Caribbean Sea. Built up on a submarine ridge, they stretch 500 mi (800 km) from the eastern end of the Greater Antilles at Anegada Passage in the north to a 90-mi-wide (145-km) channel between Grenada and Tobago in the south. Mostly oceanic, the Lesser Antilles are festooned in two, perhaps three, separate chains: an inner ring of volcanic islands; an outer line of predominantly limestone islands to the east; and Barbados, farthest east. The largest, Guadeloupe, is a composite: the western half (Basse Terre) volcanic, the eastern (Grande Terre) limestone. From north to south, the main islands in the two chains are shown in the **table**. Formerly British colonies, Barbados became independent in 1966, Grenada in 1974, Dominica in 1978, and St. Lucia and St. Vincent (including the Grenadines) in 1979; except Montserrat, the remaining British islands are self-governing Associated States. Martinique and Guadeloupe, with the latter's small dependencies, are departments of France. The Dutch islands belong to the Netherlands Antilles.

Main volcanic arc. In this chain, islands are spaced at regular intervals, with channels between them nowhere more than 30 mi (48 km) wide. With hard basaltic cores covered by ash and other igneous material, these islands are the remnants of volcanoes in various stages of decay and subaerial erosion. At the north and south ends of the chain there has

Lesser Antilles

Islands	Maximum elevation, ft (m)	Area, mi² (km²)
Main volcanic arc		
Saba (Dutch)	2850 (869)	5 (13)
St. Eustatius (Dutch)	1980 (604)	8 (21)
St. Kitts (Br.)	3711 (1131)	65 (168)
Nevis (Br.)	3300 (1006)	36 (93)
Redonda (Br.)	975 (297)	0.3 (0.8)
Montserrat (Br.)	3002 (915)	39 (101)
Guadeloupe, west (Fr.)	4869 (1484)	327 (847)
Les Saintes (Fr.)	1036 (316)	5 (13)
Dominica	4747 (1447)	290 (751)
Martinique (Fr.)	4429 (1350)	417 (1080)
St. Lucia	3145 (959)	238 (616)
St. Vincent	4048 (1234)	133 (344)
The Grenadines	980 (299)	30 (78)
Grenada	2750 (838)	119 (308)
Outliers		
Sombrero (Br.)	40 (12)	0.2 (0.5)
Anguilla (Br.)	200 (61)	35 (91)
St. Martin (half Fr., half Dutch)	1350 (411)	33 (85)
St. Barthélemy (Fr.)	990 (302)	8 (21)
Barbuda (Br.)	200 (61)	60 (155)
Antigua (Br.)	1319 (402)	108 (280)
Guadeloupe, east (Fr.)	1300 (396)	226 (585)
La Désirade (Fr.)	930 (283)	8 (21)
Marie Galante (Fr.)	670 (204)	61 (158)
Barbados	1104 (336)	167 (433)

been no recorded post-Columbian volcanic activity, but from St. Kitts to the Grenadines the cones are still active; in 1902, St. Vincent's Soufrière and Martinique's Mount Pelée erupted, causing widespread damage and killing 30,000 Martiniquans. Smaller eruptions occurred in Guadeloupe (1956) and north of Grenada (1966 submarine). Solfataras, or emissions of hot sulfurous vapors and gases, and fumaroles, or eruptions of steam from subsurface lava, evince continued volcanic activity, and earthquake shocks are frequent. All extremely mountainous, these islands rise in some cases thousands of feet sheer out of the ocean; the Pitons of St. Lucia are the most spectacular example. There is little level land, and the dense populations (400/mi² or 150/km²) occupy all but a small part of even the steepest slopes.

Outliers. By way of contrast, the outlying islands are low-lying and flat or only moderately hilly. They consist of ancient submerged volcanoes capped by sedimentary strata, now slightly tilted. Less fertile and more arid than the inner volcanic arc, they support only a degraded scrub vegetation and are, for the most part, less densely populated than the islands to the west.

Barbados. Located at the extreme southeast of the chain of outliers, this is remnant of an ancient continent. Coral limestone escarpments ascend to an elevation of 1104 ft (336 m) in Mount Hillaby. Sedimentary strata of continental origin underlie the coral and appear at the surface in the rugged, deeply dissected Scotland District of the northeast. Elsewhere, the undulating and intensely cultivated surface is reminiscent of an English landscape.

Offshore islands. There are a number of islands near the coasts of Central and South America.

Trinidad and Tobago. With a combined area of 1981 mi² (5130 km²) and population of 1,179,100, these are detached fragments of the Caribbean Coastal Range of the Andes in northern Venezuela, from which Trinidad is separated by the Gulf of Paria. In climate and culture they resemble the Lesser Antilles, but their landforms and vegetation are more continental than insular. They gained independence from the British in 1962.

Trinidad (1864 mi² or 4828 km²) is made up of three mountain ranges with lowlands and swamps between them. The forested northern range, which reaches a height of 3085 ft (940 m), is an east-west trending continuation of the Península de Paria of Venezuela to the west, cut off by the Bocas del Dragón and a chain of small islands. The central range bisects the island from southeast to northwest; the southern range, like the central range only 1000 ft (300 m) high, forms the southern rim of Trinidad, separated from Orinoco delta by the Boca de la Sierpe. Between the central and southern ranges are deposits of asphalt, in Pitch Lake, and of oil, with the refining of imported oil the mainstay of the local economy.

Tobago (116 mi² or 300 km²) is aligned along a densely wooded ridge which rises to 1800 ft (550 m) in the northeast and slopes off to the southwest.

Southern Caribbean. Off northern South America several islands appear, like Trinidad and Tobago, as outliers of Andean ranges. Venezuelan Margarita (444 mi² or 1150 km²), a mountainous island which reaches 4800 ft (1460 m) at its western end, and flat Tortuga are remnants of the Caribbean Coastal Range. Farther west, the Netherlands Leeward Islands of Aruba, Bonaire, and Curaçao (71, 120, and 178 mi², or 184, 311, and 461 km²), together with a few small Venezuelan islands, represent an extension of the peninsula of Guajira in Colombia. The Dutch islands are worn-down remnants of crystalline rocks, from which most of an overlying limestone has been eroded. Low-lying and exposed to the desiccating sweep of the northeast trades, they are arid (mean annual rainfall 20 in. or 50 cm) and support only xerophytic, or drought-resistant, vegetation.

Western Caribbean. At the western end of the Caribbean lie San Andrés (9 mi² or 23 km²) and Providencia (12 mi² or 31 km²) east of Nicaragua but dependencies of Colombia; and the Swan Islands (2 mi² or 5 km²), northwest of Nicaragua but United States possessions. Much closer to the Central American coast are the San Blas, or Mulatas, Islands of Panama; the Corn Islands (6 mi² or 16 km²) of Nicaragua, on lease to the United States; the Bay Islands of Honduras and British Honduras; and Cozumel (189 mi² or 490 mi² including its outlier Mujeres) of Mexico, east of Yucatan. Providencia is a complex volcanic plug (1200 ft or 365 m), San Andrés and Swan are limestone similar to the Caymans, the Corn Islands are of porous lava, and the Bay Islands are outliers of the Central American range that emerges farther east in Jamaica.

David Lowenthal

West Nile Virus

An arbovirus, first identified in the West Nile area of Uganda in the early 1930s, that has been an increasing threat in North America since the late 1990s. Arboviruses (***ar***thropod-***bo***rne viruses) are carried by arthropods and transmitted to the host through the bite of the insect. West Nile virus (WNV) found its way to North America in 1999, possibly through international travel, the importation of infected birds or mosquitoes, or migration of infected birds. Severe infection with WNV can result in viral encephalitis, a dangerous and sometimes fatal inflammation of the brain. Since 1999, WNV infections in humans, birds, and mosquitoes have been reported from all states except Hawaii, Alaska, and Oregon. In 2004, there were 2539 cases of human WNV infection in 41 states in the United States. The incidence of WNV encephalitis in the United States is seasonal, with peaks corresponding to the times of the year during which adult mosquitoes are active.

Transmission and epidemiology. In order to be effectively transmitted, arboviruses must have the ability to infect vertebrates (such as birds or humans) as well as the invertebrate vector. The virus must then be able to grow in the vertebrate's bloodstream and multiply (this is known as a "viremia"), remaining in the blood long enough for infection of a second invertebrate vector. Finally, it must infect the second invertebrate's salivary gland to allow transmission to other vertebrate hosts. For WNV, this cycle occurs when the *Culex* mosquito bites an infected bird, taking up the virus and becoming infected in its turn, after which the mosquito bites another bird and continues the transmission cycle.

Humans are usually "dead-end hosts," meaning they cannot spread WNV back to the vector because they do not maintain a persistent viremia. It is now recognized, however, that the virus can be transmitted by blood transfusions and organ transplants. Despite this development, the reservoirs for this virus are birds, not humans. Reservoirs are susceptible hosts that allow the reinfection of other arthropods. Therefore, there is no transmission of WNV from human to mosquito. WNV has been detected in at least 138 species of birds. Although birds infected with WNV may become ill or die (especially crows and jays), most survive infection. Bird reservoirs that are hosts for the virus will experience an infectious viremia for 1 to 4 days after exposure. After this time, the bird will either succumb to the infection or will develop life-long immunity to WNV. For continued transmission of the virus, sufficient vectors must feed on the infectious host to ensure that some mosquitoes survive long enough to feed again on a new reservoir host.

While infection in horses, dogs, and cats has been documented, these animals are not known to develop infectious-level viremias, and are therefore likely to be dead-end hosts. WNV does not appear to cause severe illness in either dogs or cats. Cases of WNV disease in horses have been documented using detection of WNV-neutralizing antibodies. In contrast to dog and cat infection, approximately 40% of equine WNV cases result in the death of the horse.

Pathogenesis and symptoms. WNV is a member of the family of Flaviviruses, a group of enveloped viruses that contain ribonucleic acid (RNA) as their nucleic acid and include Yellow fever virus, St. Louis encephalitis virus, and Dengue fever viruses. Flaviviruses are typically associated with mild systemic disease, but can cause severe infections including encephalitis. Mosquitoes acquire the flaviviruses through feeding on the blood of a viremic vertebrate host. The virus then travels from the midgut of the mosquito to the salivary glands, where it reproduces itself in great numbers. The salivary glands release virus into the saliva, where it is transferred to the next victim during a blood meal. On biting a host, the mosquito regurgitates the virus-containing saliva into the victim's bloodstream. Upon entering the host, WNV travels through the host's plasma to tissue sites for which it has a "tropism," or preference. The initial symptoms include fever, chills, headache, and backaches caused by the resulting viremia. Most viral infections do not progress beyond this stage; however, a secondary viremia may occur, which can result in infection of organs such as the brain, liver, and the skin. This can lead to fatality in the infected host.

The febrile illness associated with WNV infection is frequently nondescript, although a rash and arthritis may manifest in some patients 3 to 14 days after exposure. Lymphadenopathy (enlargement or disease of lymph nodes) may be prominent in West Nile fever. During the acute stage, the brain membranes may become involved, although this rarely leads to disease in most individuals. Typically, fatal encephalitis occurs in the elderly and in younger people with underlying illnesses, such as diabetes, emphysema, acquired immune deficiency syndrome (AIDS), and chronic diseases of the liver, kidneys, or heart. Only one in five people infected with WNV show symptoms, and less than 1 in 100 develop serious complications from the infection.

Diagnosis. WNV infection in a patient is presumptively diagnosed based on clinical symptoms and the patient's medical history. Laboratory testing is required for a confirmed diagnosis, involving a test for antibodies that are produced very early in the infected person. These antibodies can be measured in blood or cerebrospinal fluid. This blood test is generally positive in most infected people within 8 days of onset of symptoms. The current tests for human serologic diagnosis of WNV are the Centers for Disease Control and Prevention (CDC)-defined IgM and IgG enzyme-linked immunosorbent assay (ELISA). An ELISA uses antigens or "pieces" of the WNV to detect the presence of immunoglobulins or antibodies that may be present in a patient's serum (the liquid component of blood). When these antibodies are present, a positive ELISA test will result. However, because the ELISA can cross-react with other flaviviruses (such as St. Louis encephalitis virus), it is considered a screening test only and a virus

neutralization test must be conducted to confirm that WNV is indeed the virus present in the patient.

Control and prevention. Prevention and control of arbovirus disease is best accomplished through a comprehensive mosquito management program; personal, household, and community prevention through public education; and targeted prevention programs aimed at educating high-risk groups.

Mosquito surveillance and reduction. Mosquito management programs must be established at the local level and be capable of surveillance sensitive enough to detect WNV transmission associated with increased risk of disease in humans or domestic animals. Effective mosquito control includes a sustained surveillance program to detect vector species and to identify and map their immature habitats by season. Larval mosquito surveillance is performed by sampling a wide range of aquatic habitats for the presence of pest and vector species. Adult mosquito surveillance is used to monitor species presence and relative abundance of adult mosquitoes in an area.

The detection of the *Culex* mosquito in an area indicates the possible presence of WNV as well. Traps for mosquitoes use carbon dioxide as bait (because mosquitoes are attracted to the carbon dioxide that humans exhale; thus, the traps simulate a living human). After capture, WNV can then be detected in individual mosquitoes. Source reduction of mosquitoes is the elimination of mosquito larval habitats. This is the most effective method for controlling mosquitoes, and can be as simple as proper disposal of used tires, cleaning of rain gutters, bird baths, and unused swimming pools. Sanitation is a major part of integrated vector management programs, and is central to controlling WNV infection in human and mammal populations.

Personal, community, and household prevention. Individuals can use DEET (*N*,*N*-diethyl-*m*-toluamide or *N*,*N*-diethyl-3-methylbenamide)-based repellents on skin and clothing as their first defense against WNV infection, in addition to avoiding the outdoors during prime mosquito-biting hours (dusk to dawn). Households may be protected by the elimination of mosquito-breeding sites and the repair or installation of window screens. Communities can protect against WNV disease through active reporting of dead birds and organized mosquito control measures.

Targeted prevention. Targeted prevention includes seeking out those groups that are most at-risk for severe disease with WNV, including those over age 50 and those who are immunocompromised. Individuals who work outside extensively or engage in recreation outdoors in endemic areas are also at greater risk for WNV infection and should be also targeted with disease-preventing information.

Treatment and preventive vaccine. Treatment for WNV encephalitis and fever is largely supportive: possible hospitalization, administration of intravenous fluids, respiratory support (ventilation), and prevention of secondary infections. There is no specific antiviral therapy for WNV infection.

No vaccine is available for use to prevent WNV disease in humans. One research group has successfully used recombinant deoxyribonucleic acid (DNA) vaccines to induce protective immune responses in mice and horses, but this success has yet to be transferred to humans. *See* ARBOVIRAL ENCEPHALIDES; MOSQUITO; VIRUS; ZOONOSES. Marcia M. Pierce

Bibliography. R. W. Bauman, *Microbiology: Alternate Edition with Diseases by Body System*, Benjamin Cummings, 2006; P. R. Murray et al., *Medical Microbiology*, Mosby, 2002; E. W. Nester et al., *Microbiology: A Human Perspective*, 4th ed., McGraw-Hill, 2004; L. M. Prescott, J. P. Harley, and D. A. Klein, *Microbiology*, 6th ed., McGraw-Hill, 2005.

Wetlands

Ecosystems that form transitional areas between terrestrial and aquatic components of a landscape. Typically they are shallow-water to intermittently flooded ecosystems, which results in their unique combination of hydrology, soils, and vegetation. Examples of wetlands include swamps, fresh- and salt-water marshes, bogs, fens, playas, vernal pools and ponds, floodplains, organic and mineral soil flats, and tundra. As transitional elements in the landscape, wetlands often develop at the interface between drier uplands such as forests and farmlands, and deep-water aquatic systems such as lakes, rivers, estuaries, and oceans (see **illus.**). Thus, wetland ecosystems are characterized by the presence of water that flows over, ponds on the surface of, or saturates the soil for at least some portion of the year.

Features. Regulations supporting the U.S. Federal Clean Water Act [33 CFR 328.3 (b)] define wetlands as "those areas that are inundated or saturated by surface or ground water at a frequency and duration sufficient to support, and that under normal circumstances do support, a prevalence of vegetation typically adapted for life in saturated soil conditions." A more inclusive definition of a wetland was adopted by the international community through the 1971 Ramsar Convention, which states that wetlands are "areas of marsh, fen, peat, and or water, whether natural or artificial, permanent or temporary, with water that is static or flowing, fresh, brackish or salt, including areas of marine water the depth of which at low tide does not exceed six meters."

Wetland soils can be either mineral (composed of varying percentages of sand, silt, or clay) or organic (containing 12–20% organic matter). Through their texture, structure, and landscape position, soils control the rate of water movement into and through the soil profile (the vertical succession of soil layers). Retention of water and organic carbon in the soil environment controls oxidation-reduction and other biogeochemical reactions that facilitate the functioning of wetland soils. Wetland soils are regularly wet, deprived of oxygen, or both. *See* BIOGEOCHEMISTRY; OXIDATION-REDUCTION.

Vegetated wetlands are dominated by plant species, called hydrophytes, that are adapted to live in water or under saturated soil conditions.

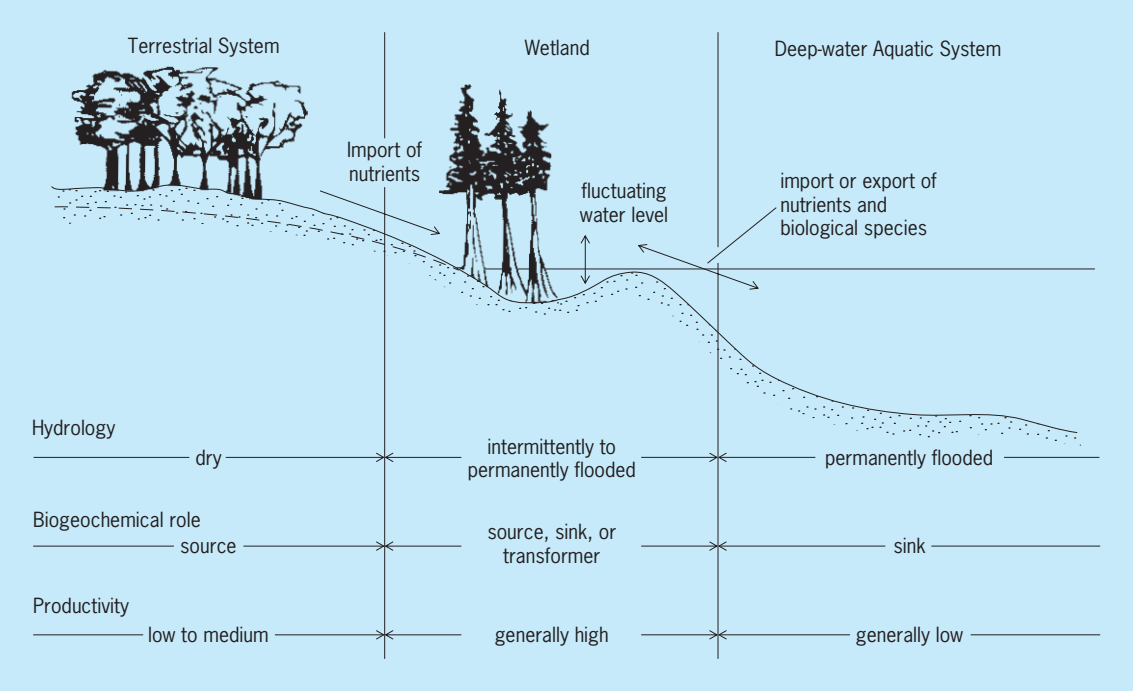

Relationships of wetlands to terrestrial systems and deep-water aquatic systems and comparison of their characteristics. (*After W. J. Mitsch and J. G. Gosselink, Wetlands, 2d ed., Van Nostrand Reinhold, 1993*)

Adaptations that allow plants to survive in a waterlogged environment include morphological features, such as pneumatophores (the "knees," or exposed roots, of the bald cypress), buttressed tree trunks, shallow root systems, floating leaves, hypertrophied lenticels, inflated plant parts (such as aerenchyma), and adventitious roots, to name a few. Physiological adaptations also allow plants to survive in a wetland environment. These adaptations include the ability of plants to transfer oxygen from the root system into the soil immediately surrounding the root (rhizosphere oxidation); the reduction or elimination of ethanol accumulation due to low concentrations of alcohol dehydrogenase (ADH); and the ability to concentrate malate (a nontoxic metabolite) instead of ethanol in the root system. *See* ROOT (BOTANY).

Wetlands differ with respect to their origin, position in the landscape, and hydrologic and biotic characteristics. For example, work in the United States has focused on the hydrology (water sources and transport processes) as well as the geomorphic position of wetlands in the landscape. This hydrogeomorphic approach, developed by M. M. Brinson and colleagues, recognizes and uses the fundamental physical properties that define wetland ecosystems to distinguish among classes of wetlands that occur in riverine, depressional, estuarine or lake fringe, mineral or organic soil flats, and slope environments. Ecosystem functions (such as subsurface water storage) in the broad categories of hydrology, biogeochemistry, plant community, and faunal support/habitat are specific to each class of wetland (see **Table 1**).

Global extent. The extent of wetlands in the world is estimated to be 2–3 × 10^6 mi^2 (5–8 × 10^6 km^2), or about 4–6% of the Earth's land surface. Wetlands occur in humid, cool regions such as bogs, fens, and tundra of Canada, Alaska, and northern Europe and Asia; along rivers and streams as riparian wetlands, seasonally flooded forests, and backswamps; in the deltas of the great rivers, including the Amazon, Mekong, Danube, Mississippi, Nile, and Rhone; along temperate, subtropical, and tropical coastlines as salt marshes, mud flats, and mangrove swamps; and even in arid regions as inland salt flats,

TABLE 1. Wetland functions typical of riverine ecosystems in North America as identified by the hydrogeomorphic approach*

Functional group	Function
Hydrology	Dynamic surface water storage
	Long-term surface water storage
	Energy dissipation
	Subsurface storage of water
	Ground-water flow or discharge
Biogeochemistry	Nutrient cycling
	Removal of elements and compounds
	Retention of particulates
	Organic carbon export
Plant community	Maintenance of native plant community
	Detrital biomass production
Faunal support/habitat	Spatial structure of habitat
	Interspersion and connectivity
	Distribution and abundance of native invertebrate fauna
	Maintenance of the distribution and abundance of native verbrate fauna

*From M. M. Brinson et al., *Guidebook for Application of Hydrogeomorphic Assessments to Riverine Wetlands*, Tech. Rep. WRP-DE-11, U.S. Army Corps of Engineers Waterways Experiment Station, Vicksburg, MS, 1995.

seasonally inundated playas, riparian systems, and vernal pools. Wetlands are found on every continent except Antarctica and in every clime from the tropics to the frozen tundra. Rice paddies, which comprise another 500,000–600,000 mi^2 (1.3–1.5 × 10^6 km^2), can be considered as a type of domesticated wetland of great value to human societies worldwide. In the contiguous United States, an estimated 160,000 mi^2 (420,000 km^2) of wetlands remain, with 490,000 mi^2 (1,270,000 km^2) in Canada and 266,000 mi^2 (688,000 km^2) in Alaska. Most Canadian and Alaskan wetlands are considered northern peatlands (containing histosols, or soils dominated by organic material). *See* BOG; MANGROVE; MUSKEG; PLAYA; SALT MARSH; TUNDRA.

Societal values. Compared to other ecosystems, wetlands are often an extremely productive part of the landscape. Wetlands support a rich variety of waterfowl and aquatic organisms, and represent one of the highest levels of species diversity and richness of any ecosystem. From a conservation viewpoint, wetlands are an extremely important habitat for rare and endangered species. For example, over 63% of reptiles, 68% of birds, and 75% of amphibians listed as threatened or endangered in the United States are associated at some point in their life cycle with wetlands.

Wetlands often serve as natural filters for human and naturally generated nutrients, organic materials, and contaminants. The ability to retain, process, or transform nutrients, organic matter, and contaminants is called assimilative capacity, and is strongly related to wetland soil texture and vegetation. The assimilative capacity of wetlands has led to many projects that use wetland ecosystems for wastewater treatment and for improving water quality. Wetlands have been constructed to process domestic sewage, coal mine drainage, industrial wastes, urban runoff, and agricultural drainage waters. Wetlands also have been shown to prevent downstream flooding and, in some cases, to prevent ground-water depletion as well as to protect shorelines from storm damage. *See* GROUND-WATER HYDROLOGY.

Management. The best wetland management practices enhance the natural processes of wetlands by maintaining conditions as close to the natural hydrology of the wetland as possible, including hydrologic connections with adjacent rivers, lakes, and estuaries. Marsh management for wildlife, particularly waterfowl, often has translated to water-level manipulation achieved by the construction of dikes, weirs, control gates, and pumps.

Losses. Wetlands can be easily affected by drainage projects and upstream hydrologic modifications. As a result, the world's wetlands are becoming a threatened landscape. Loss of wetlands worldwide currently is estimated at 50%. Although in a few regions of Scandinavia, North America, Africa, and Australia some natural wetlands still can be found in large tracts, in some countries, especially in Europe, natural wetlands are nearly extinct or, at best, heavily managed. It has been estimated that the contiguous United States has lost 53% of its wetlands over the past 200 years, with several states in the upper Midwest losing 80% or more. Wetlands in California have been reduced, according to some estimates, to less than 10% of their original distribution.

Wetland loss results primarily from habitat destruction, alteration of wetland hydrology, and landscape fragmentation (**Table 2**). Global warming may soon be added to this list, although the exact loss of coastal wetlands due to sea-level rise is not well documented. Worldwide, destruction of wetland ecosystems primarily has been through the conversion of wetlands to agricultural land. For example, over 80% of the Mississippi riverine bottomlands, and well over 90% of historic wetlands in Ohio, have been converted to crop production. Other wetland destruction has occurred through the harvesting of timber. The practice of clear cutting and the construction of logging roads are two major forestry activities that directly destroy wetlands, although degradation

TABLE 2. Human actions leading to wetland loss*

Cause of loss	Floodplains	Rivers	Lakes	Peatlands	Swamps
Drainage for agriculture, forestry, and mosquito control	C	C	P	C	C
Dredging and channelization for navigation and flood protection	A	P	A	A	A
Filling for solid waste disposal, roads, and commercial, industrial, or residential development	P	P	P	A	A
Conversion for aquaculture	C	C	P	A	A
Construction of dikes, dams, and seawalls for flood and storm control, water supply, and irrigation	C	C	C	A	A
Discharge of pesticides, herbicides, domestic and industrial waste. agricultural runoff, and sediment	C	C	C	A	A
Mining of wetlands for peat, coal, gravel, phosphate, and other materials	P	A	C	C	C
Logging and shifting cultivation	C	P	A	C	C
Ground water abstraction	P	C	A	A	A
Fire	C	P	A	C	C
Sediment diversion by dams, deep channels, and other structures	C	C	P	A	A
Hydrological alteration by canals, roads, and other structures	C	C	C	P	P
Subsidence due to extraction of ground water, oil, gas, and other minerals	C	C	A	A	A

*C = common and important cause of loss. P = present but not a major causes of loss. A = absent or exceptional.
SOURCE: From P. Harrison and F. Pearce, *AAAS A has of Population and Environment*, University of California Press, Berkeley, 2001; based on data of United Nation Environment Programme.

through other forms of timber extraction also is widespread. Peat extraction for fuel and horticulture occurs primarily in the Northern Hemisphere (Great Britain, North America, and northern Eurasia), and is responsible for the loss of millions of hectares of peatlands. Other worldwide activities that have led to the direct loss of wetlands include sand and gravel mining of riverbeds and banks and the development of mariculture (cultivation of marine species) through the conversion of salt marshes and mangrove forests into commercial impoundments.

Hydrologic modifications that destroy, alter, and degrade wetland systems include the construction of dams and water diversions, ground-water extraction, and the artificial manipulation of the amount, timing, and periodicity of water delivery. The primary impact of landscape fragmentation on wetland ecosystems is the disruption and degradation of wildlife migratory corridors, reducing the connectivity of wildlife habitats and rendering wetland habitats too small, too degraded, or otherwise irreversibly altered to support the critical life stages of plants and animals.

Protection policies. The heavy losses of wetlands in the world, coupled with the recognized values of these systems, have led to a number of policy initiatives, especially in the United States, where estimated losses have been as high as 300,000 acres (1200 km^2) per year. The federal government has no specific wetland protection law. Instead, it has relied primarily on presidential executive orders, a "no net loss" policy, and Sections 404 and 401 of the Clean Water Act of 1977 for wetland protection. The principal agencies involved in Clean Water Act jurisdiction include the Environmental Protection Agency, Army Corps of Engineers, Fish and Wildlife Service, and National Marine Fisheries Service. The efforts of these regulatory agencies with Clean Water Act jurisdiction have been augmented by a wetland protection program in agriculture ("Swampbuster" Provisions of the Food Security Act) administered by the Natural Resources Conservation Service of the Department of Agriculture. Swampbuster provisions deny federal subsidies to any farm owner who converts certain types of wetlands to farmland. Canada, Sweden, and other countries also have developed wetland protection regulations.

International cooperation. International agreements, particularly the Ramsar Convention and the North American Waterfowl Management Plan, govern the protection of wetland ecosystems worldwide. The Ramsar Convention is an international treaty for the conservation of wetland ecosystems. Named for the city in Iran where the Convention on Wetlands took place in 1971, the Ramsar Convention maintains a list of "Wetlands of International Importance." Inclusion on the list is determined by the international significance of a wetland, with an emphasis on the its ability to support migratory waterfowl. Africa and North America hold the largest percentage of Ramsar wetlands (21% and 23%, respectively), while South America supports another 16%. As of 2000, a total of 66,840,000 hectares in more than 890 sites is protected under the Ramsar Convention.

The North American Waterfowl Management Plan, agreed to in 1986 by the United States and Canada, has a goal of conserving and restoring about 9300 mi^2 (24,000 km^2) of waterfowl wetland habitat. This treaty was developed partly in response to a steep decline in waterfowl in both countries in the early 1980s.

Restoration and creation. Wetland restoration usually refers to the rehabilitation of degraded or hydrologically altered wetlands, often involving the reestablishment of vegetation. Wetland enhancement generally refers to the targeted restoration of one or a set of ecosystem functions over others, for example, the focused restoration of a breeding habitat for rare, threatened, or endangered amphibians. Wetland creation refers to the construction of wetlands where they did not exist before. This type of activity typically involves much more extensive hydrology and soils engineering. Created wetlands are also called constructed or artificial wetlands. Some regulatory agencies distinguish between constructed wetlands and created wetlands, with the former term used to designate wetlands built for wastewater or storm-water treatment, and the latter term used for wetlands developed on nonwetland sites to produce or replace a natural habitat.

Restoring, enhancing, or creating a wetland requires a comprehensive understanding of hydrology and ecology, as well as engineering skills. Failure of wetland creation and restoration usually has resulted from one or more of the following circumstances: poor understanding of the restored site's hydrology: lack of appreciation for the intrinsic dynamics of wetlands and their ecology; unrealistic expectations as to what constitutes a successfully created or restored wetland ecosystem; and not allowing enough time for the wetland ecosystem to develop.

There is a great deal of interest in wetland restoration and creation, particularly in response to the "no net loss" policy in the United States that requires replacement of wetlands when they are unavoidably filled for development. In addition, thousands of wetlands have been constructed throughout the world as part of an economical and ecological approach to improving water quality, especially for domestic wastewater, non-point-source pollution from agriculture, and coal mine drainage. Although it is difficult to estimate the extent of created and restored wetlands in the world, rates of wetland gain probably do not equal the rates of wetland loss. *See* DAM; ECOSYSTEM; ESTUARINE OCEANOGRAPHY; HYDROLOGY; RESTORATION ECOLOGY; RIVER ENGINEERING.

William J. Mitsch; Peggy L. Fiedler; Lyndon C. Lee; Scott R. Stewart

Bibliography. B. L. Bedford, D. J. Leopold, and J. P. Gibbs, Wetland ecosystems, in S. A. Levin (editor in chief), *Encyclopedia of Biodiversity*, vol. 5, Academic Press, New York, 2001; C. M. Finlayson and M. Moser (eds.), *Wetlands*, 1991; D. A. Hammer, *Creating Freshwater Wetlands*, 2d ed., Lewis Publishers, 1996; D. A. Hammer (ed.), *Constructed Wetlands*

for Wastewater Treatment, Lewis Publishers, Chelsea, MI, 1989; P. Harrison and F. Pearce, *AAAS Atlas of Population and Environment*, University of California Press, Berkeley, 2001; R. H. Kadlec and R. L. Knight, *Treatment Wetlands*, Lewis Publishers/CRC Press, Boca Raton, FL, 1995; E. Maltby, *Waterlogged Wealth*, Earthscan Publishers, London, 1986; W. J. Mitsch (ed.), *Global Wetlands: Old World and New*, Elsevier, Amsterdam, 1994; W. J. Mitsch and J. G. Gosselink, *Wetlands*, 3d ed., Wiley, New York, 2000; T. C. Winter, Hydrologic studies of wetlands in the northern prairie, in A. van del Valk (ed.), *Northern Prairie Wetlands*, Iowa State University Press, Ames, 1989.

Wheat

A food grain crop. Wheat is the most widely grown food crop in the world, and is increasing in production. It ranks first in world crop production and is the national food staple of 43 countries. At least one-third of the world's population depends on wheat as its main staple. The principal food use of wheat is as bread, either leavened or unleavened.

The United States is second to Russia in total production, but the average yield per acre in the United States is about twice that of Russia. Other major wheat-producing countries in the world are Canada, China, India, France, Argentina, and Australia. Highest per acre yields (60 bushels or 1.63 metric tons, and above) are produced in western Europe. The change to day-length-insensitive, short-straw and stiff-straw varieties which can be heavily fertilized and irrigated has resulted in yield increases, especially in food-deficit countries from about 35°N latitude to the Equator (**Fig. 1**). Wheat is an important commodity of international trade. Of the total wheat production, about one-fifth moves in world trade. The United States leads in exports. Other leading export countries include Canada, Australia, France, and Argentina.

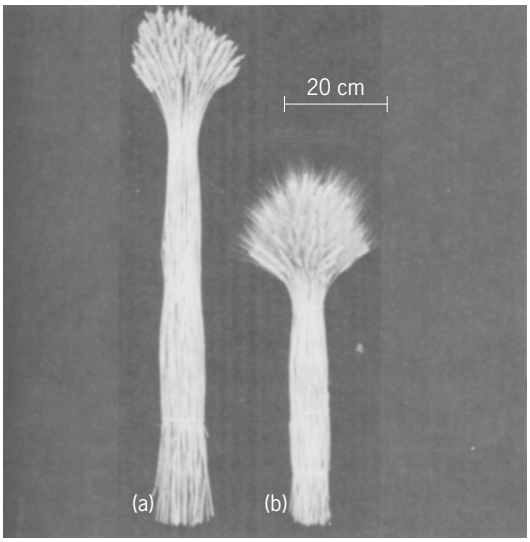

Fig. 1. Sheaves of wheat: (*a*) tall, awnless variety; (*b*) short, awned or bearded variety. Both tall and short types may have awns.

Adaptation. Wheat is best adapted to a cool dry climate, but is grown in a wide range of soils and climates. Much of the world's wheat is seeded in the fall season and, after being dormant or growing very slowly during winter, it makes rapid growth in the spring and develops grain for harvest in early summer. In the United States, about 80% of the wheat is fall-seeded. Wheat competes with other crops where the fall-seeded type makes good use of limited rainfall and where the spring-seeded type makes good use of the short frost-free growing season. Medium-textured and well-drained soils are best for wheat since it cannot withstand long periods of wet ground.

General cultural practices. Clean, weed-free seed of an adapted variety treated to control seed-borne diseases, along with good germination, increases the prospect of a satisfactory crop. Wheat is usually seeded at a depth of 1–2 in. (2.5–5 cm), or deeper when topsoil is dry. A firm seed bed, together with soil firmly packed over the seed, improves stands and reduces losses from winter kill and similar hazards. Wheat may be seeded as part of a rotation following corn, soybeans, or other small grain crops in humid regions; a year of fallow during which moisture is accumulated and weeds kept under control often precedes wheat in areas of low rainfall. Seeding time depends on soil moisture, soil temperature, presence of certain insects and disease organisms, and other factors. Seeding rate averages about 60 lb/acre (67 kg/hectare) in the United States but may be as high as 120 lb/acre (134 kg/ha) in the eastern United States and as low as 30 lb/acre (33 kg/ha) in the arid West.

Most wheat is harvested directly from the field with a combined harvester-thresher, or it may be wind-rowed first and then combined when dry. Wheat grain should contain no more than 13% moisture for safe storage, should be cool, and free from insects. Forced-air dryers are used to reduce the moisture content of grain harvested at above 13%. *See* AGRICULTURAL MACHINERY.

Market classes. Wheat for milling is classified according to hardness, color, and best use. In the United States, there are seven official market classes of which the following five are the most important: (1) hard red winter, for bread; (2) hard red spring, for bread and rolls; (3) soft red winter, for cake and pastries; (4) white, for bread, breakfast foods, and pastries; and (5) durum, for macaroni products.

Distribution of types and varieties. Variety adaptation, production practices, and crop hazards are widely different in the major areas of production in the United States. Therefore, the United States can be divided into six wheat-growing regions (**Fig. 2**), although the boundaries overlap: region I, the Northeast and Ohio Valley; region II, the Southeast; region III, the Northern Plains; region IV, the Central Plains; region V, the Pacific Northwest; and region VI, California and Arizona.

Regions I and II have the highest rainfall of the six, and wheat of soft texture with low protein

generally is grown. The crop is fall-seeded. General crop rotations are used in which corn and soybeans take prominent places. Yorkstar, Genesee, and Ionia, all white grains, are the leading varieties in region I. In region II, Arthur, a soft red winter type, is the leading variety in most of the states.

In region III, spring-sown wheat predominates. The leading varieties of durum wheat are Rolette and Leeds, while the leading hard red spring varieties are Era, Waldron, and Olaf. Some fall-sown wheat is grown in the western portion of this region, but the winters are severe and the crop often is lost. In general, the region has low rainfall; hence tillage and rotations are designed to conserve as much moisture as possible.

In region IV, little else but hard red winter wheat is grown. Subregion IVA requires a cold-hardy variety, and subregion IVB, with its hot summers, requires an early maturing variety for most dependability. In subregion IVA, leading varieties in recent years have been Winalta and Cheyenne; in subregion IVB, Scout, Centurk, Triumph, and Sturdy. This region has rather low rainfall, and tillage, including summer fallow, is especially timed to conserve moisture for the next crop.

In region V, wheat of several types is grown but soft white and hard red winter predominate. Much of region V uses summer fallow in the cropping sequence. A small acreage is irrigated. Nugaines and Hyslop have been the most important white wheat varieties in this region. Paha and Moro, club wheats with white grain, have been grown also. Wanser is the leading hard red winter variety, and Komar and Thatcher hard red spring wheats continue to be found on a small acreage.

Region VI grows day-length-insensitive common wheats with red or white grain, or amber durum. Irrigation is common for wheat in this region. The mild winters allow nonhardy varieties to be utilized from fall seeding. The leading varieties in recent years have been Anza and Inia 66.

In all of these regions, variety changes occur every year as improved forms are introduced; hence, the varieties named here represent types but may not continue to be the particular varieties actually produced in a given year. Some wheat is irrigated in regions IV, V, and VI, but the major acreage is rainfed.

Northwestern Mexico grows varieties similar to those found adapted in region VI and utilizes similar production methods. In Canada the varieties used are similar to those used in region III. Production practices in Canada are also similar to those used in region III, with heavy emphasis on dry-land farming. The leading variety of hard red spring wheat in Canada is Neepawa.

Characteristics. The wheat inflorescence is a spike bearing sessile spikelets arranged alternately on a zigzag rachis. Two, three, or more florets may develop in each spikelet and bear grains. The florets are composed of three stamens and a single ovary with its style and stigma enclosed by a lemma and palea (**Fig. 3**). The lemma may be awned or awnless.

Fig. 2. Adaptation regions for wheat in the United States. (*USDA*)

The grain may be white, red (brown), or purple, and it may be hard or soft in texture. Size of the grain or caryopsis may be large, as in durum, or very small, as in shot wheat (*Triticum sphaerococcum*). Wheats vary in plant height and in the ability to produce tillers. The stems are usually hollow.

Wheat is normally a self-pollinating plant, although out-crosses do occur. When a breeder deliberately makes crosses, the anthers must be removed before

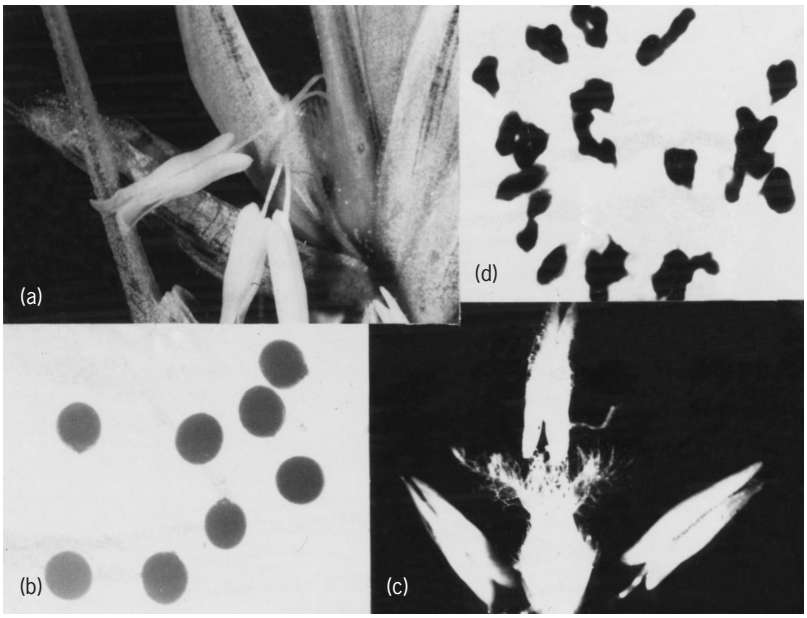

Fig. 3. Flower parts of a male fertile wheat plant: (*a*) three stamens; (*b*) pollen grains; (*c*) single ovary with feathery stigma; and (*d*) normal chromosomes.

they shed pollen, and fresh pollen must be transferred from the second parent in the cross to the stigmas while they are receptive. Certain breeding stocks of wheat have been found that are male-sterile and shed no pollen. These have been used successfully to develop hybrid wheats without resorting to hand pollination procedures.

The wheat grain is composed of the endosperm and embryo enclosed by bran layers. The endosperm portion is principally starch and is therefore used as energy food. Wheat is also an important protein source, especially for those people who use wheat as their main staple. Since wheat protein is deficient in lysine and some of the other essential amino acids that are important in human nutrition, breeding for improved amino acid balance or nutritional value is a primary objective in many breeding projects. *See* SEED.

Origin and relationships. Wheat was gathered for food in prehistoric times and has been grown as a crop for perhaps 7000 years. The exact center of origin of cultivated wheat has not been established, but the most likely area is the foothills of southwestern Iran, southeastern Turkey, and adjacent parts of Iraq and Syria with an extension to Israel. Wild *Aegilops* and *Triticum* are still found in these regions.

Botanically, wheat is a member of the grass family to which rice, barley, corn, and several other cereal grain crops also belong. The *Triticum* genus includes a wide range of wheat forms (**Fig. 4**). They are often grouped by their polyploid chromosome number series of 7, 14, and 21 pairs. In Fig. 4, *a* and *b* have 7 pairs (diploid), *c* to *l* have 14 pairs (tetraploid), and *m* to *q* have 21 pairs (hexaploid) of chromosomes.

Taxonomic studies place the goat grasses (*Aegilops*) and wheat (*Triticum*) in one genus, *Triticum*. They intercross to some extent, and it is believed that *Aegilops* species have contributed one or more groups of chromosomes to the tetraploid and hexaploid species of *Triticum*. Breeders have transferred individual genetic characteristics from goat grasses to wheat. The forms illustrated in Fig. 4 have been intercrossed, and in this way resistance to wheat rusts has been transferred from durum and emmer to common wheat. Wheat has been crossed with rye (*Secale*) and with *Agropyron* (a grass), and by the use of irradiation, segments of alien chromatin have been transferred to wheat, sometimes with beneficial results. New forms, called *Triticale*, have been derived from crossing rye and wheat followed by doubling the chromosomes in the hybrid. The most successful of these is based on a cross of rye and durum wheat in which there are 14 pairs of wheat and 7 pairs of rye chromosomes in the new species. Several variations occur.

Since there is much genetic similarity among the three genomes (sets of 7) of common wheat, the loss of a chromosome of one genome is partially compensated by similar chromosomes in the other genomes. This allows the development of 21 different genetic stocks, each deficient for one chromosome or one chromosome pair. The use of such stocks has greatly accelerated genetic studies and has made work more precise. An extension of this technique permits the substitution of a whole chromosome in one variety for the same chromosome in another variety. This has value in genetic studies and is also useful in variety improvement. *See* BREEDING (PLANT).

Fig. 4. Spikes of some varieties of wheat. (*a*) Wild and (*b*) cultivated forms of einkorn (*Triticum monococcum*); (*c*) wild emmer (*T. dicoccoides*); (*d*) emmer (*T. dicoccum*); (*e*) durum (*T. durum*); (*f*) poulard (*T. turgidum*); (*g*, *h*) durum (*T. durum*); (*i*) Polish (*T. polonicum*); (*j*) Persian (*T. carthlicum*); (*k*) emmer (*T. dicoccum*); (*l*) timopheevi (*T. timopheevi*); (*m*) spelt (*T. spelta*); (*n*) bread or common (*T. aestivum*); (*o*) club (*T. compactum*); (*p*) shot (*T. sphaerococcum*); and (*q*) spelt (*T. spelta*). (*Courtesy of H. Kihara*)

Variety development. Most countries in which wheat is grown have wheat breeding programs in which the objective is to develop more productive and more stable varieties (cultivars). Many methods are combined in these programs, but in nearly all of them specially selected parent types are crossbred followed by pure-line selection among the progeny to develop new combinations of merit. Varieties and genetic types from all over the world become candidate parents to provide the desired recombinations of good quality, winter and drought hardiness, straw strength, yield, and disease resistance. Wheats must be bred for specific milling processes and to provide quality end-use products. Many new varieties have complex pedigrees.

A number of worldwide collections of diverse wheats are available to breeders. The collection maintained by the U.S. Department of Agriculture contains about 35,000 items. Teams of breeders, geneticists, entomologists, plant pathologists, and

cereal chemists are found at major breeding centers to give comprehensiveness to the breeding of the new varieties. *See* GRAIN CROPS. Louis P. Reitz

Diseases

Estimates of annual loss of wheat due to disease range from 10 to 20%. More than 100 diseases of wheat have been described, and about 50 may be economically important. Stem rust, leaf rust, loose smut, and stinking smut are potentially destructive and have caused enormous losses in the past. Currently these diseases are better controlled. Virus diseases and the root and basal stem rots are widespread and cause significant loss.

The loss in North America from black stem rust, caused by the fungus *Puccinia graminis* f. sp. *tritici*, in earlier years has been as high as 3×10^8 bu (1.1×10^7 m^3). This fungus is heteroecious, with its sexual phase and aeciospores produced on the common barberry (*Berberis vulgaris*). The aeciospores from barberry germinate and invade wheat stems and leaves, producing urediniospores in reddish-brown pustules (**Fig. 5**) and later teliospores in black pustules. The teliospores over-winter and then germinate in the spring, producing basidiospores which initiate the life cycle on the barberry in areas where this shrub is present. Stem rust can also survive over winter as urediniospores on wheat in southern Texas, Mexico, and other countries with mild climates. Spread by airborne urediniospores into more northern wheat growing areas then occurs during the spring and summer. Eradication of the common barberry in much of North America has helped to reduce losses from stem rust by eliminating the sexual stage, preventing the development of new pathogenic races by hybridization, and also by eliminating aeciospores, one of the sources of inoculum for wheat.

Fig. 5. Black stem rust of wheat urediniospore stage.

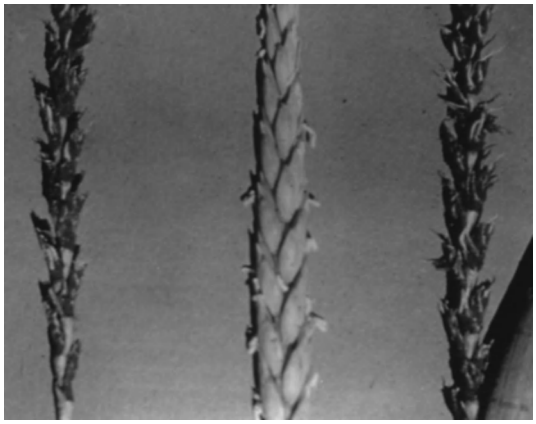

Fig. 6. Loose smut of wheat-diseased and healthy spikes.

Losses from stem rust and from the leaf rust fungus, *Puccinia recondita*, have been reduced in recent years by using cultivars of wheat with genetic resistance to the prevalent physiologic or pathologic races of both pathogens.

The common smut diseases of wheat are loose smut, caused by *Ustilago nuda*, and stinking smut, caused by *Tilletia* species. Both fungi are systemic in their host. At flowering the entire head (spike) is transformed into a mass of smut spores in the loose smut disease (**Fig. 6**). Spores develop where the grains usually form in stinking smut, but the accessory parts of the spike are not affected. The diseased heads therefore are not as conspicuous as in loose smut. An odor resembling decaying fish is associated with stinking smut. Both fungi are seed-borne, as spores externally in case of stinking smut and as mycelium within the healthy grain in the loose smut disease. Spores of stinking smut also survive in the soil in some areas. When seed with smut spores or with internal mycelium are sown, the smut fungi grow into the young seedlings, but usually there are no distinct symptoms until heading. Chemical seed treatment fungicides are effective controls of the smut diseases except where the spores survive in the soil. Resistant cultivars are also available for some areas of wheat culture. *See* FUNGISTAT AND FUNGICIDE.

Since 1950, several viral diseases have become common and destructive. The symptoms of these diseases include mosaic and mottling of foliage, stunting, rosetting, and yellowing. The virus causing wheat streak mosaic is transmitted by the mite *Aceria tulipae*. Other grains and grasses are also susceptible. The vectors of barley yellow dwarf virus are several species of aphids. Wheat spindle streak and soil-borne mosaic are caused by soil-borne viruses that are transmitted by *Polymyxa graminis*, a soil-borne fungus. Zoospores of *Polymyxa* with viral particles invade wheat roots, thereby effectively inoculating the plants. Late planting, crop rotation, and the use of resistant cultivars are some of the management practices that aid in reducing losses from these

diseases.

There are several damaging root, crown, and basal stem rot diseases in many wheat-growing areas of the world. Stunting, early kill of plants, lodging, and premature ripening are the symptoms. The causal fungi include *Fusarium* species, *Gaeumannomyces graminis*, *Pseudocercosporella herpotrichoides*, and *Bipolaris sorokiniana*. These fungi persist in the soil or crop residues from year to year. Crop rotation, late planting of winter wheat, and ensuring adequate soil fertility are management practices that reduce losses. *See* PLANT PATHOLOGY.

C. Wayne Ellett

Wheat Grain Proteins

Most of the protein of wheat grain is found in the endosperm, the inner part of the kernel that yields white flour in the milling process. The endosperm—about 75% of the wheat kernel—consists largely of starch granules surrounded by a proteinaceous matrix. The protein makes up about 8–16% of the grain, but lower or higher protein percentages are not unusual. *See* PROTEIN.

Most of the protein in the endosperm is associated with gluten, the cohesive and elastic protein mass that remains after a dough is kneaded in water or in an aqueous salt solution to wash away the starch granules. The ability to form gluten is unique to wheat (although the term "corn gluten" has been applied somewhat inappropriately in recent years to a maize protein fraction that lacks gluten-like viscoelastic properties). Gluten proteins are storage proteins, which are broken down to provide nitrogen and amino acids for the synthesis of proteins and other molecules needed by the new plant upon seed germination. In addition to gluten, endosperm proteins include 15–25% of proteins that have been classified as albumins or globulins based on their solubilities. Many of these proteins may, like gluten, serve as storage proteins, although various proteins of the albumin-globulin fraction inhibit protease, amylase, or other enzymes. These inhibitor proteins may protect the grain against insect or fungal attack. Various other proteins occur in the bran (outer layer) or germ (embryo) fractions obtained during milling, but these constitute a small part of the total protein. *See* ALBUMIN; GLOBULIN.

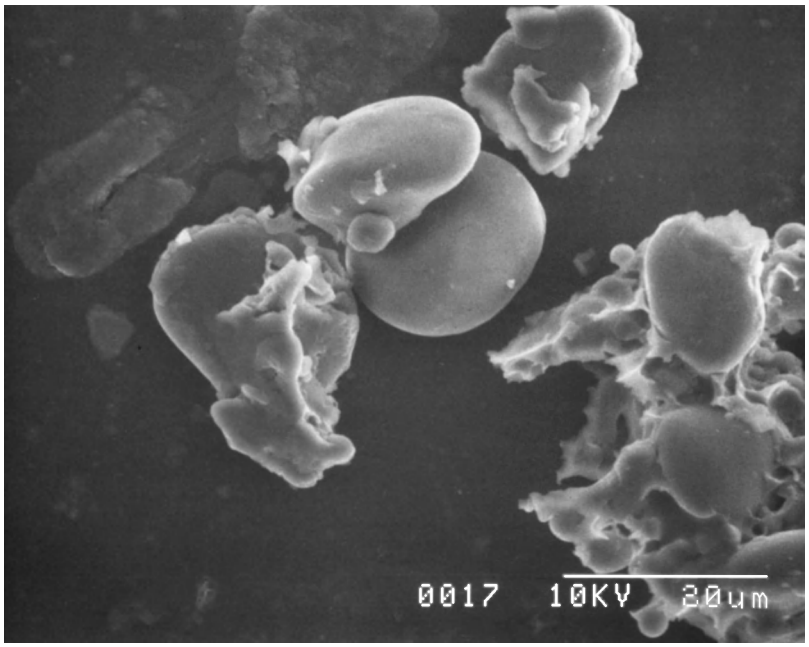

Fig. 7. Scanning electron micrograph of flour particles from a hard red winter wheat cultivar (Len). Large and small spherical objects are starch granules. The irregularly shaped material attached to the starch granules corresponds to protein deposits that surround the starch granules in the endosperm cells. (*Micrograph by Gregory M. Glenn, USDA, Albany, CA*)

During grain development when storage protein and starch synthesis occur, storage proteins are deposited in the endosperm as protein bodies. These protein bodies and any membranes surrounding them are disrupted during the desiccation phase that leads to the dry mature grain. Consequently, the protein ends up as a largely continuous matrix surrounding the starch granules. Adherence of the endosperm to the surrounding protein matrix differs in hard and soft wheats. In soft-milling cultivars, the matrix separates more readily from the granules, generally leaving them intact; however, in hard-milling cultivars, the matrix and granules adhere so strongly that starch granules are occasionally cleaved during milling. For large-scale flour milling, machinery must be adjusted appropriately for grain hardness in order to achieve good yields of white flour (endosperm). Soft wheats are commonly used for pastry or noodle making, whereas the hard wheats are more commonly used for breadmaking. **Figure 7** is a scanning electron micrograph of some flour particles from a hard red winter wheat cultivar. *See* STARCH.

Gliadin and glutenin. Traditionally, gluten proteins were divided into two main fractions on the basis of solubility (gliadin) or insolubility (glutenin) in aqueous alcohol solutions (such as 70% ethanol-water or 50% *N*-propanol–water). This is a somewhat crude fractionation, but it is now recognized that the gliadin fraction consists almost entirely of monomeric (having single polypeptide chains) proteins, and the glutenin fraction consists almost entirely of polymers, in which certain gluten proteins are linked together by disulfide bonds. The proteins of glutenin polymers are often called subunits to distinguish them from the monomeric gliadins. Thus, just as polypeptides or proteins are polymers of amino acids, glutenin is a higher-order polymer in which the monomers are proteins.

All gluten proteins have substantial amounts of proline and glutamine, 2 of the usual 20 amino acids found in proteins. Consequently, the proteins are often classified as prolamins (although in earlier work, this term was usually reserved for the gliadin proteins). Prolamins are also found in related cereal grains, including rye, barley, and corn.

Gluten can be separated by two-dimensional electrophoretic or chromatographic methods into about 40–50 components, but the actual number of components is bound to be significantly higher based on the number of gluten protein genes found in the wheat genome. Gluten proteins are about equally divided between gliadin and glutenin; however,

the actual number of components varies among cultivars. Structurally, the gliadins can be divided into α-, γ-, and ω-gliadins on the basis of their amino acid sequences. Gutenin subunits are usually divided into a high-molecular-weight group (mol wt ~70,000–100,000) and a low-molecular-weight group (mol wt ~30,000–42,000). Of all the endosperm proteins, the amounts and types of high-molecular-weight glutenin subunits show the strongest correlations with variations in mixing and baking quality among cultivars. *See* CHROMATOGRAPHY; ELECTROPHORESIS.

Breadmaking. All gluten proteins are strongly cohesive due not only to hydrogen bonding between the many glutamine residues but also to hydrophobic and ionic interactions. It is, however, the polymeric glutenin fraction that contributes elasticity to the hydrated gluten proteins, as in a wheat flour-and-water dough. Wheat cultivars vary widely in quality for breadmaking and other food uses. For breadmaking, dough should have a moderately high level of elasticity and good stability for mixing. The molecular weight distribution of the glutenin polymers is an important factor; a shift toward higher molecular weight (larger polymers on the average) leads to a stronger, more elastic, more viscous dough. Certain glutenin subunits facilitate polymer formation by having two or more cysteine residues that participate in intermolecular disulfide bond formation, whereas others having only a single cysteine residue available for intermolecular disulfide bond formation interfere with formation of larger polymers.

In addition to intrinsic genetic factors in any given wheat cultivar, there are strong environmental effects (soil conditions, weather conditions) on breadmaking qualtiy. The molecular weight distribution of the glutenin polymers or the ratio of monomeric proteins to polymeric proteins in the grain seems likely to play a role in variations that arise from environmental effects. Changes in these characteristics may result from changes in the degree of expression of specific protein components, even though the same qualitative complement is present in a given cultivar. A good balance of dough elasticity versus dough extensibility is important to the wheat processors, and this balance differs for different types of products.

Celiac disease. In some individuals, various gluten proteins and the albumin-globulin proteins of wheat endsoperm elicit allergic reactions, such as respiratory distress or skin wheals. However, wheat and its close relatives, rye and barley, are notable for their involvement in celiac disease (gluten-sensitive enteropathy). Celiac disease is also likely to involve an immune response, but one that appears to differ from allergic reactions involving primarily IgE-type immunoglobulins. In susceptible individuals, gluten proteins and their counterparts in rye and barley produce changes in the small intestinal mucosa that, in more severe cases, damage the absorptive epithelium. Because almost all nutrients—proteins, carbohydrates, calcium, iron, vitamins, and so forth—are absorbed by way of the epithelium, there can be a myriad of wide-ranging symptoms, from diarrhea and loss of weight to osteoporosis and anemia.

Although the exact mechanism that initiates processes leading to mucosal damage in celiac disease is unknown, a current hypothesis involves the binding of peptides derived from gluten to major histocompatibility complex class II proteins, which in turn present these peptides to T cells. Damage may result from cytokine release as a consequence of T-cell activation. All gliadin proteins, and probably some of the glutenin subunits, have the activating factor, which corresponds to particular combinations of amino acids in the peptide sequences and almost certainly includes glutamine and proline residues. Although gliadins have 260 or more amino acid residues in their polypeptide chains, synthetic peptides having sequences corresponding to as few as 12 amino acids of the α-type gliadin sequence have been reported to be active. However, the unifying structural principle of active peptides remains to be determined. *See* ALLERGY; GASTROINTESTINAL TRACT DISORDERS; IMMUNITY.

Donald D. Kasarda

Processing

Milling has evolved from rudimentary crushing or cracking to sophisticated separation and refining. The main purpose of milling is isolation of the starch-protein matrix, that is, separation of the endosperm from the high-fiber bran and high-lipid germ. Under optimal conditions, milling yields a high-quality, uniformly colored flour with a relatively stable shelf-life. The flours of hard wheats (11 to 13% protein) develop strong gluten complexes during mixing and are therefore suitable for making bread. Whole soft wheats (9 to 11% protein) yield flours that are used primarily for cakes, cookies, and pastries. Durum wheat is used to produce a relatively coarse flour, semolina, used for manufacture of pasta products.

Cleaning. Freshly harvested grain may contain sticks, stones, soil, chaff, and other classes of mixed grains. Wheat must be clean and somewhat moist in order to be milled to a highly-quality flour. Even small amounts of extraneous materials can produce wide variations in the acceptability and subsequent baking performance of wheat flour.

Middling streams are segregated and directed to reduction rolls, which have smooth surfaces and turn at equal speeds. They further reduce the size of relatively pure chunks of endosperm. After each reduction roll, the stream is sized. Particles that are the proper size are removed as finished flour; oversized particles are scalped from the screens and returned to a reduction stream.

The term flour extraction describes the proportion (by weight) of flour obtained from a given weight of wheat and, therefore, the degree to which endosperm has been separated or refined from the whole kernel. The higher the flour extraction, the less bran has been removed from the flour. The total of all individual flour streams constitutes straight flour, generally 72% of the wheat kernel. Individual flour streams comprising straight flour may be blended to produce the various commercial flour grades.

Developments in milling technology have enabled single-pass regrinding of flours in a pin mill to reduce

particle size to ranges from 1 to 50 micrometers. This regrinding process followed by air classification of particles by size and density enables partitioning of protein (fine) and starch (coarse) fractions. In this manner, air-classified flours may be established with specified levels of protein or starch, each with designated functional properties. *See* FOOD MANUFACTURING.

Mark A. Uebersax

Wheat Biotechnology

The application of biotechnological tools such as deoxyribonucleic acid (DNA) sequencing and plant transformation has had a major impact on the understanding of the contribution of the wheat endosperm prolamine proteins, particularly the high-molecular-weight (HMW) glutenins, to wheat protein and dough chemistry.

DNA sequencing. The isolation and DNA sequencing of the HMW-glutenin genes has enabled scientists to completely understand the primary structure of these proteins, which was extremely difficult to determine using just the proteins themselves because of the large repetitive regions in their amino acid sequences. These primary protein amino acid sequences revealed the two main features likely to be responsible for the unique characteristics of the HMW glutenins: (1) the long, central, repetitive domains theorized to contribute some of the elastic properties of doughs, and (2) the placement in the two ends of each polypeptide of most disulfide bond-forming cysteine amino acid residues necessary for glutenin polymer formation. These genes have also been used to express the proteins in microbial systems, a technique that allows easier isolation and study of individual natural and modified HMW glutenins—studies not feasible using whole wheat seeds with large numbers of chemically similar proteins. Synthesis of these wheat recombinant HMW glutenins are resulting in new understandings of these critical proteins. *See* BIOTECHNOLOGY; DEOXYRIBONUCLEIC ACID (DNA).

Plant transformation. The HMW-glutenin genes were also the first genes to be used in experiments in bioengineering wheat, in attempts to specifically alter an agronomic or economically important trait. In the first experiments, natural or modified HMW-glutenin genes were inserted into wheat cells, and whole fertile plants were regenerated (transgenic plants). When transgenic seeds of later generations were tested for dough physical properties, it was found that addition of these HMW-glutenin genes led to the synthesis of additional proteins in the wheat seed. These transformation experiments indicate that it is possible to obtain stable, high-level expression of HMW-glutenin transgenes in wheat, with dramatic changes in the physical properties of the resulting dough. However, this needs to be confirmed by large-scale field testing of these transgenic wheat lines under different environmental conditions. However, it is now that in the next few years transgenic wheat lines will make major contributions to the understanding of the function of the glutenin proteins. In addition, by further understanding the molecular basis of the glutenin contribution to dough properties, it will be possible to design new wheat varieties for targeted end uses. *See* GENETIC ENGINEERING.

One advantage of modern biotechnological techniques is that genetic traits can be moved among plants with an ease and degree never before possible. For example, it is expected that the addition of the wheat HMW glutenins to other seeds, such as barley or maize, will alter the processing characteristics of flours from those crops and potentially lead to novel products and new markets. Although such biotechnological tools are extremely powerful and will become even more important in the study of wheat quality, these tools must still be coupled with studies of protein chemistry and dough functionality to relate genetic changes to the phenotypic traits important for wheat utilization.

Olin D. Anderson

Bibliography. Y. P. Bajaj, *Wheat*, 1991; F. Barro et al., *Nat. Biotech.*, 15:1295–1299, 1997; A. E. Blechl and O. D. Anderson, Expression of a novel high-molecular-weight glutenin subunit gene in transgenic wheat, *Nat. Biotech.*, 14:875–879, 1996; L. T. Evans and W. J. Peacock (eds.), *Wheat Science: Today and Tomorrow*, 1981; D. D. Kasarda, Gluten and gliadin: Precipitating factors in coeliac disease, in M. Maki, P. Collin, and J. K. Visakorpi (eds.), *Proceedings of the 7th International Symposium on Coeliac Disease* pp. 195–212, Coeliac Disease Study Group, Institute of Medical Technology, University of Tampere, Finland; J. H. Klippart, *The Wheat Plant: Its Origin*, *Culture*, *Growth*, *Development*, *Composition*, *Varieties Together with Information on Corn Culture*, 2 vols., 1980; W. Q. Loegering, J. W. Hendrix, and L. E. Browder, *The Rust Diseases of Wheat*, USDA Agr. Handb. 334, November 1967; M. Mäki and P. Collin, Coeliac disease, *Lancet*, 349:1755–1759, 1997; S. A. Matz, *Cereal Technology*, 1970; Y. Pomeranz, *Wheat*, *Chemistry and Technology*, 2 vols., 3d ed., 1988; E. J. Pyler, *Baking Science and Technology*, 2 vols., 1988; L. P. Reitz, *Wheat in the United States*, USDA Agr. Inform. Bull, 386, 1976; P. R. Shewry, N. G. Halford, and A. S. Tatham, The high molecular weight subunits of wheat glutenin, *J. Cereal Sci.*, 15:105–120, 1992; P. R. Shewry, M. J. Miles, and A. S. Tatham, The prolamin storage proteins of wheat and related cereals, *Prog. Biophys. Mol. Biol.*, 61:37–59, 1994; N. Shani, et al., Role of the amino-and carboxy-terminal regions in the folding and oligomerization of wheat high molecular weight glutenin subunits, *Plant Physiol.*, 98:433–441, 1991; R. Tkachuk and W. Bushuk, *Gluten Proteins*, 1990; W. T. Yamazaki, *Soft Wheat: Production*, *Breeding*, *Milling*, *and Uses*, 1981.

Wheatstone bridge

An extensively used electrical network of four resistances R_{1-4} (or impedances Z_{1-4} if alternating currents are involved). This network (see **illustration**), was first described by S. H. Christie in 1833, only 7 years after Georg S. Ohm discovered the relation-

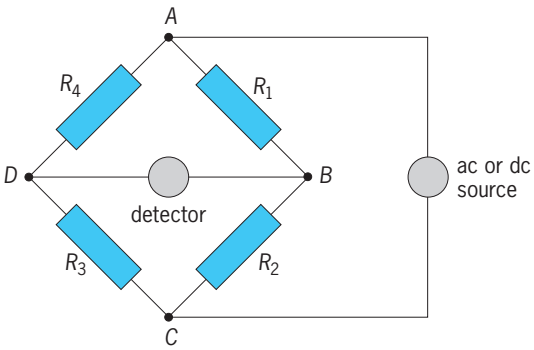

Wheatstone bridge circuit.

ship between voltage and current. Since 1843, when Charles Wheatstone called attention to Christie's work, Wheatstone's name has been associated with this network. *See* ELECTRICAL IMPEDANCE; ELECTRICAL RESISTANCE; OHM'S LAW.

The potentials at nodes B and D with respect to node C (or A), depend only on the ratios R_1/R_2 and R_4/R_3 respectively, and some or all of the values of $R_{1\text{-}4}$ can be adjusted until these potentials are equal. A detector connected between B and D will then indicate zero current or voltage and Eq. (1) will be satisfied [or Eq. (2) for alternating-current impedances].

$$R_1/R_2 = R_4/R_3 \tag{1}$$

$$Z_1/Z_2 = Z_4/Z_3 \tag{2}$$

The network possesses the following useful properties:

1. Obtaining the detector indication of zero current or voltage (which is often termed "balancing the bridge") and the consequent validity of Eq. (1) or Eq. (2) is independent of the internal resistance (or impedance) of the source or detector.

2. The same result is obtained if the positions of source and detector in the network are interchanged. This is a simple example of reciprocity in linear systems. *See* RECIPROCITY PRINCIPLE.

3. From Eq. (1) or Eq. (2), the value of one of the four resistances or impedances can be calculated if one of the others and the ratio of the remaining two are known.

4. The balanced network will eliminate unwanted electrical signals (interference and pickup) between the source and circuitry connected to B and D.

The response of the detector to a deviation from the condition of Eq. (1) or Eq. (2) depends on the values of $R_{1\text{-}4}$ or $Z_{1\text{-}4}$ as well as the internal resistances or impedances of both source and detector. This response can be calculated by network analysis. *See* NETWORK THEORY; RESISTANCE MEASUREMENT.

Bryan P. Kibble

Wheel and axle

A wheel and its axle or, more generally, two wheels with different diameters or a wheel and drum, as in a windlass, rigidly connected together so that they rotate as a unit on a common axis. The principle of operation is the same as that of the lever in that, for static equilibrium, the summation of torques about the axis of rotation equals zero. Where flexible members, such as ropes, have been firmly attached to a wheel and drum (see **illus.**) and the machine is mounted on frictionless bearings, Eq. (1) applies.

$$F_1R_1 - F_2R_2 = 0 \tag{1}$$

If F_1 is the input and F_2 the output, the mechanical advantage is given by Eq. (2).

$$MA = \frac{F_2}{F_1} = \frac{R_1}{R_2} \tag{2}$$

If the bearings are not frictionless, a friction torque T_f will oppose the rotation. For the case with F_1 the input, the friction torque will act in the same direction as the load or output torque F_2R_2, and the equilibrium equation takes the form of Eq. (3), which may be solved for individual components as in Eq. (4)

$$F_1R_1 - F_2R_2 - T_f = 0 \tag{3}$$

$$F_2 = \frac{F_1R_1 - T_f}{R_2} \tag{4}$$

See ANTIFRICTION BEARING; TORQUE.

Because all members are rigidly connected, all torques act through the same angular displacement; the efficiency η is then given by Eq. (5).

$$\eta = \frac{F_2R_2}{F_1R_1} \times 100 = \frac{F_1R_1 - T_f}{F_1R_1} \times 100$$

$$= \left(1 - \frac{T_f}{F_1R_1}\right) 100 \tag{5}$$

The main difference between the lever and the wheel and axle is that the wheel and axle permits the forces to operate through a much greater distance. In the illustration the wheel and drum could be allowed to rotate any number of revolutions if the ropes were wrapped the required number of times around each before they were attached. *See* LEVER.

The wheel and axle is seldom encountered in the exact form shown in the illustration, but the principle of the torques about the axis of rotation

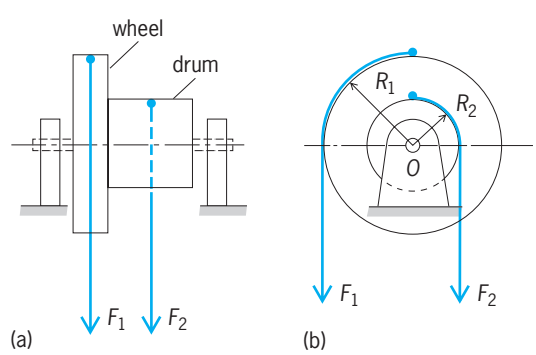

Wheel and axle. (*a*) Side view. (*b*) Front view.

equaling zero finds widespread applications in machines such as the belt drive. *See* FORCE; SIMPLE MACHINE.

Richard M. Phelan

Wheelbase

The distance in the direction of travel from front to rear wheels of a vehicle, measured between centers of ground contact under each wheel. For a vehicle with two rear axles, the rear measuring point is on the ground midway between rear axles. Tread of a vehicle is the distance perpendicular to the direction of travel between front wheels, or between rear wheels, measured from centers of ground contact.

Frank H. Rockett

White dwarf star

The smallest kind of ordinary star, about the size of the Earth. In most white dwarf stars, an amount of matter equal to 60% of the Sun's mass is compressed into this small volume. One quarter liter (about 1 cup) of white dwarf material has a mass of 600 tons. Several thousand white dwarf stars have been discovered.

Some of the other properties of white dwarf stars reach extreme values. The hottest known stars are either white dwarfs or stars that are just about to become white dwarfs, and have temperatures of a few hundred thousand kelvins. A few white dwarf stars have very strong magnetic fields, with field strengths exceeding 10^8 gauss (10^4 teslas), many times stronger than can be generated in laboratories. Some white dwarf stars rotate every few minutes, and some others rotate so slowly that no spin has been detected over several decades. The coolest white dwarf stars emit less energy per second than any other type of visible star. *See* HIGH MAGNETIC FIELDS.

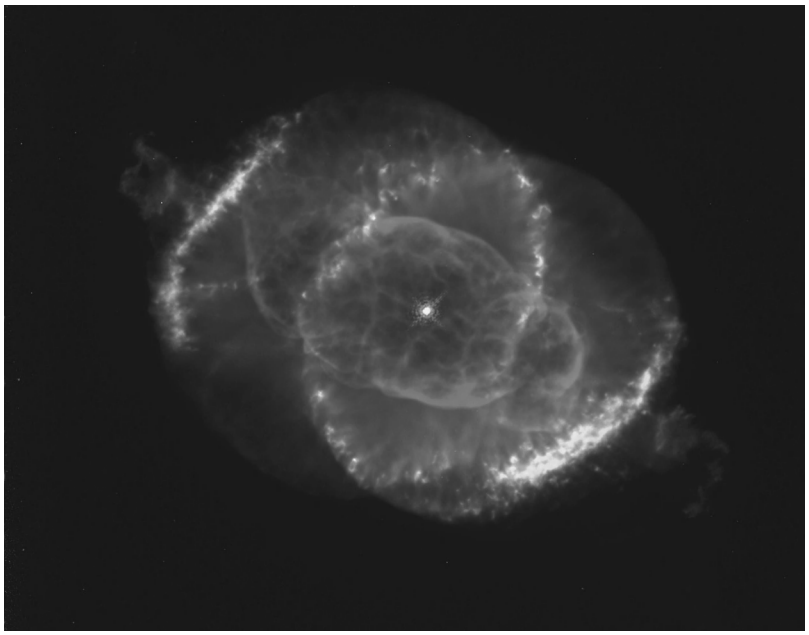

Fig. 1. Hubble Space Telescope image of the planetary nebula NGC 6543, nicknamed the Cat's Eye Nebula, surrounding a very hot, tiny blue star which is about to become a white dwarf star. (*P. Harrington, NASA*)

Astrophysical context. While these extreme stellar properties give white dwarfs some intrinsic interest, they fit into an astrophysical context that connects them to the universe at large. They are the final stages in the life cycles of low-mass stars like the Sun, the most common types of stars. At present, the Sun is on the main sequence, fusing hydrogen to form helium in its core. Five billion years from now, the hydrogen in the center of the Sun will run out, and the Sun will become a red giant star as it turns to other nuclear reactions to provide its internal heat. Stars like the Sun will become only hot enough inside that they will fuse helium nuclei to form a mixture of carbon and oxygen. They will then reach the end of the nuclear fusion road, and will no longer be able to generate their own sources of energy. At this point in their life cycle, they will be on the way to becoming white dwarf stars, the most common types of stellar remnants. *See* NUCLEOSYNTHESIS; STELLAR EVOLUTION.

Just before a star becomes a white dwarf star, it will shed its outer layers as its stellar wind strengthens at the very end of the red giant stage of its life cycle. In many and perhaps all cases, the expelled gas will become visible as a glowing gas cloud called a planetary nebula. The confusing name of these objects was given by nineteenth-century astronomers, who knew nothing of their origin but surmised only that these roundish gas clouds looked somewhat like Uranus and Neptune. **Figure 1** shows an interesting planetary nebula: the dot in the middle of the swirls of gas is a very hot, blue star which is about to become a white dwarf. *See* PLANETARY NEBULA.

These connections to the prior phases of stellar life cycles relate white dwarf stars to the rest of astrophysics. For example, the formation of white dwarf stars changes the chemical composition of the Milky Way Galaxy as a whole. When stars eject mass at the end of their lives, this ejected mass enriches the interstellar medium, the gas between the stars. This mass will eventually condense to become second- and third-generation stars and planets. The ejected mass is enriched in elements such as carbon, oxygen, and nitrogen, which are critical for life and human civilization. *See* INTERSTELLAR MATTER.

Structure. A star becomes a bona fide white dwarf when it contracts to a state where heat pressure plays a very small role in determining its internal structure. When the first white dwarfs were discovered in the 1920s, senior astronomers such as Arthur Eddington and Henry Norris Russell were astonished because it seemed impossible that stellar material could be compressed to such enormously high densities. Subrahmanyan Chandrasekhar, traveling from India to England to begin his graduate studies in astronomy, first recognized that the material in white dwarf stars was a previously unrecognized form of matter, called degenerate matter. The electron clouds in degenerate matter actually touch each other, and this touching gives white dwarf matter a resistance

to compression. This resistance to compression balances the force of gravity.

The initial reaction to these ideas of Chandrasekhar was one of disbelief. However, his theory of white dwarf stars can be tested by the use of the mass-radius relation. Since the heat in a star plays no role in its internal structure, a star of a particular mass can have only a particular radius, as long as it has the same interior chemical composition as that of all other white dwarf stars. Chandrasekhar's theoretical explanation predicts that all white dwarfs composed of a carbon/oxygen mix should fall along the same line in a graph which plots stellar mass against stellar radius. Later modifications to his theory suggest that some hot white dwarf stars should lie slightly above Chandrasekhar's line, since heat pressure plays a small role in their structure.

Tests of this theory began in the 1960s, but early data were fragmentary. The launching of the European *Hipparcos* spacecraft in 1989 provided astronomers with the distances to a large number of stars and permitted a more definitive test. The results are shown in **Fig. 2**, in which the general agreement between Chandrasekhar's theory and the observations is apparent. Several stars fall below the line for carbon, which the interior of white dwarf stars is expected to contain. Astronomers are still undecided as to whether some white dwarf stars have interiors made of iron or whether there are some observational uncertainties that need to be cleared up.

Life cycle. Superficially, the life cycle of white dwarf stars is simple. They have no nuclear energy sources, so they simply cool. The hottest star to have compacted itself to the white dwarf state is H1504+65, with a temperature of 190,000 K (350,000°F). The coolest white dwarf stars have temperatures near 3800 K (6400°F), still hotter than the coolest main-sequence stars. The internal structure of these two stars is virtually the same. Consequently, white dwarf stars are spread all over the bottom of the Hertzsprung-Russell diagram, a graphical tool which is used to classify stars. *See* HERTZSPRUNG-RUSSELL DIAGRAM.

The temperatures of the coolest white dwarf stars can be used to determine how long they have cooled. The numbers of very cool white dwarf stars drop off significantly when their temperatures reach approximately 5000 K (8500°F), about the temperature of the Sun. White dwarf stars which are this cool have been cooling for several billion years. More quantitative calculations of their lifetimes and allowance for the length of time it takes a star to become a white dwarf star can provide an estimate of the age of the stellar population of the Milky Way Galaxy. The derived age is 8–10 billion years. This age is marginally consistent with measurements of the ages of other rather old objects in the universe such as the globular star clusters, which are thought to be at least 12 billion years old.

However, the life cycle of white dwarf stars, when considered in a little more depth, is more complex. The surface layers of white dwarfs are much more varied than is the case of other types of stars. The most obvious variation is in the ratio of hydrogen to helium, the two most common elements in the universe. In virtually every other type of star, the hydrogen/helium ratio is near its cosmic value of approximately 10:1. However, white dwarf stars come in two basic flavors, hydrogen-rich and helium-rich. The hydrogen-rich stars have a surface composition which is very hydrogen-rich indeed, with no traces of helium. In some of the best-observed cases, the hydrogen/helium ratios exceed 10^7. In the helium-rich star, the surfaces are very much depleted in hydrogen. In the best-observed cases, the hydrogen/helium ratio is less than 10^{-7}. This variation in the surface hydrogen/helium ratio, over 14 orders of magnitude, is one of the most extraordinary chemical peculiarities in stellar astronomy. It seems likely that at least part of the variation comes from subtle

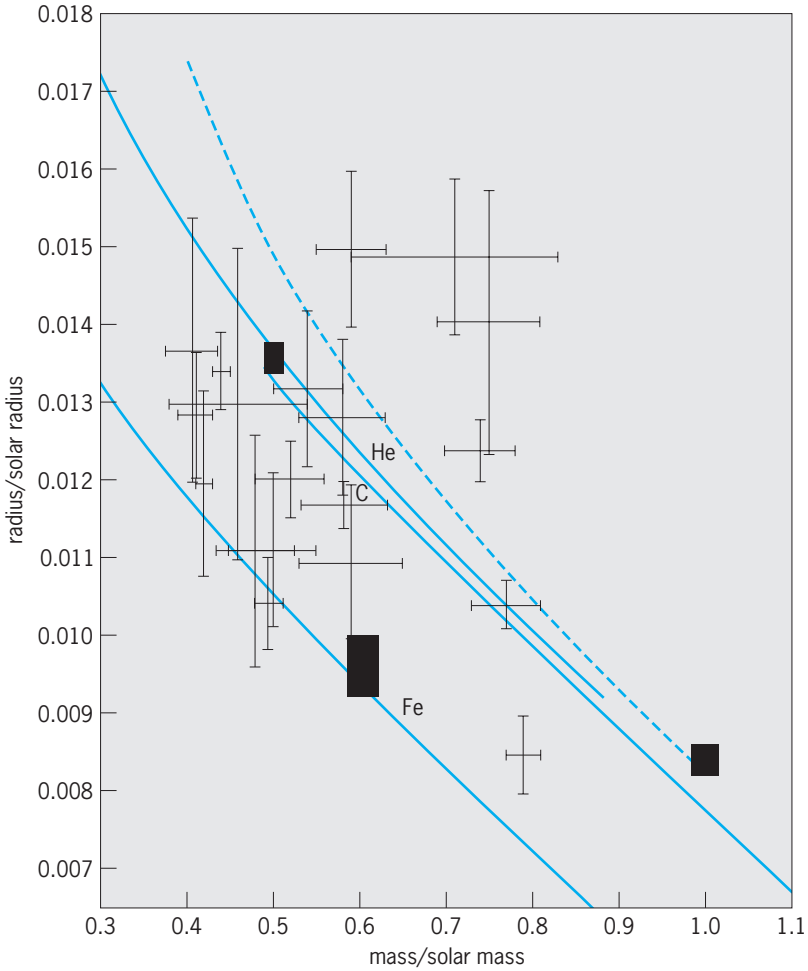

Fig. 2. Mass-radius relation for white dwarf stars. The solid lines give the relationship between mass and radius which is expected for white dwarf stars composed of helium (He), carbon (C) or a carbon/oxygen mixture, and iron (Fe). The broken line gives a revised version of the theory, which includes the hydrogen surface layer as well as the carbon interior. The data points and error bars give the position of a number of observed white dwarf stars. The mass and radius of each star are at the intersection of the vertical and horizontal lines. The three boxes are the positions of white dwarf stars in binaries, which are the measurements of highest precision and fewest assumptions. The agreement of these boxes with the theoretical lines supports Chandrasekhar's theory. (*After H. Shipman and J. Provencal, Confirming a Nobel Prize winning theory with the Hipparcos spacecraft: Accurately determining distances and diameters of white dwarf stars, in J.-E. Solheim and E. Meistas, eds., 11th European Workshop on White Dwarf Stars, pp. 15–23, Astronomical Society of the Pacific San Francisco, 1999*)

differences in the way that white dwarfs form in the late stages of red giant life cycles.

When white dwarf stars reach certain temperatures, sizes, and chemical compositions, their interiors become unstable and they vary in the amount of light that they emit. Waves of brighter and darker parts of the surface sweep around the star, producing the variations in their light from what are technically called nonradial pulsations. The periods of white dwarf stars range from 20 minutes to several hours. As a result, their pulsations are best observed by worldwide networks of telescopes, since any individual site can observe only a few pulsational cycles before the star sets. Slow changes in the temperature and diameter of one white dwarf star, PG1159–035, have been detected in the study of its pulsations. The time scale is roughly a million years, about what is expected from understanding of stellar evolution. Similar changes should be observable in other stars.

Perhaps the most extreme and least understood aspect of white dwarf stars is their high magnetic fields, which range from 1 megagauss (100 teslas) to a few hundred megagauss (a few tens of thousands of teslas). It is believed that these fields are fossil fields, the remnants of much weaker fields which exist in some stars that contain a strong, overall magnetic field. The fields in white dwarfs are discovered because they shift the wavelengths of spectral lines in the stars' surfaces. *See* STAR; ZEEMAN EFFECT. Harry L. Shipman

Bibliography. J. B. Kaler, *Stars and Their Spectra*, 1989; J.-E. Solheim and E. G. Meistas (eds.), *11th European Workshop on White Dwarfs*, 1999; V. Weidemann, Masses and evolutionary status of white dwarf stars and their progenitors, *Annu. Rev. Astron. Astrophys.*, 28:103–137, 1990.

Whooping cough

An acute infection of the tracheobronchial tree caused by *Bordetella pertussis*, a bacteria species exclusive to infected humans. The disease (also known as pertussis) follows a prolonged course beginning with a runny nose, and finally develops into violent coughing, followed by a slow period of recovery. The coughing stage can last 2–4 weeks, with a whooping sound created by an exhausted individual rapidly breathing in through a narrowed glottis after a series of wrenching coughs. The classical disease occurs in children 1–5 years of age, but in immunized populations infants are at greatest risk and adults with attenuated (and unrecognized) disease constitute a major source of transmission to others.

Bordetella pertussis is highly infectious, particularly following face-to-face contact with an individual who is coughing. The inhaled bacteria attach specifically to the hairlike cilia of cells lining the respiratory airway and remain there without invading host tissues. The disease is caused by structural components and extracellular toxins elaborated by *B. pertussis* from this superficial niche. Multiple virulence factors produced by the organism play important roles at various stages of pertussis.

For decades a vaccine produced from whole *B. pertussis* cells and combined with diphtheria and tetanus toxoids has been used throughout the world for routine childhood immunization. This vaccine has been effective in prevention of childhood pertussis, although immunity declines after 12 years. This whole-cell vaccine has significant side effects (pain, swelling, fever). Concern over vaccine morbidity has caused immunization rates to decline in some developed countries. These drops in immunization rates have often been followed by widespread outbreaks of disease, including deaths. Considerable effort has been directed toward the development of a vaccine which would minimize side effects but maintain efficacy through use of purified *B. pertussis* virulence factors. A new acellular vaccine containing filamentous hemagglutinin and pertussis toxin is available and has fewer side effects than the whole-cell vaccine. The acellular vaccine has been in use for some time in Sweden, Japan, and England. It has been approved for primary immunization in the United States. *See* DIPHTHERIA; TETANUS; VACCINATION.

Although *B. pertussis* is susceptible to many antibiotics, their use has little effect once the disease reaches the coughing stage. Erythromycin is effective in preventing spread to close contacts and in the early stage. *See* DIPHTHERIA. Kenneth J. Ryan

Bibliography. K. J. Ryan (ed.), *Sherris Medical Microbiology*, 3d ed., 1994.

Wide-area networks

Communication networks that are regional, nationwide, or worldwide in geographic area, with a minimum distance typical of that between major metropolitan areas. Smaller networks include metropolitan and local-area networks. A communication network provides common transmission, multiplexing, and switching functions that enable users to transport data between many sources and many destinations. Under ideal circumstances, the data that arrive at the destination are identical to the data that were sent. The rate of arrival of bits at any point in the network is said to be the data rate at that point and is typically measured in bits per second. These bits may come from one source or from a multiplicity of sources. The capacity of a network to transmit at a cerain data rate is known as its bandwidth. *See* LOCAL-AREA NETWORKS.

Network fundamentals. There are several fundamental attributes and concepts that facilitate the accurate transmission of data within and between digital networks. To communicate between computers, a set of rules, formats, and delivery procedures known as protocols must be established. Once these are agreed to, an Interface Control Design (ICD) is created as the basis for the communications between the sender and receiver of data.

Part of the communications protocol allows for definition of where the packets of digital data are to be routed. Each packet of data contains the unique

address of a computer or other network as its destination. The routing of the data is known as packet switching since the nodes in the network can switch the packet to various transmission paths. Networks allow alternate or redundant communications paths in the case of network transmission problems. Networks are interconnected by means of routers. These devices allow packets received from one network to be passed along to another network, while insulating each network from the internal characteristics of the other. Thus, two networks with different internal transmission speeds can be connected successfully by means of a router. *See* PACKET SWITCHING.

Another part of the communications protocol allows for including error detection and correction information in the data packets. The destination computer or network will verify the data in the packet utilizing the error control data, such as a checksum. If the data are found to be corrupted, a negative acknowledgment (NACK) will be returned to the sender, and the packet must be resent.

Protocols are also used to implement flow control. This allows the receiving computer or network to communicate back to the sender when it can or cannot receive additional data. This is important in that both computers and networks vary in terms of processing speeds and bandwidth capacity.

Advances. Since the early 1990s, the ubiquity and deployment of wide-area networks (WANs) such as terrestrial, wireless, and satellite networks have exploded around the world. It is now possible for nearly any person or entity to interface with a network. Some of these entities include personal hand-held communications devices, home and office appliances (for example computers, telephones, televisions, and security monitors), automobiles and trucks, and land- or water-based infrastructures (such as highway systems, oil rigs, and irrigation systems). *See* DIGITAL COMPUTER.

Protocols. Certain protocols have become standards for a majority of wide-area networks. Asynchronous Transfer Mode (ATM) is a protocol used in business-to-business (B2B) communications when high data rates are needed. Typically, one ATM port supports 45 megabits per second (Mbps). Frame Relay is another business-to-business protocol. The advantage of Frame Relay is that the data rate can be scaled to the individual company's needs. As an example, one company may purchase a 512 kilobits per second (kbps) frame relay circuit, whereas a company with lower requirements may purchase a 256-kbps circuit. These two protocols can also coexist in a hybrid network by utilizing devices known as protocol converters. Companies use this flexibility to place voice, video, and data simultaneously on the same circuits using a packet switching methodology. The third protocol, and the one having the most worldwide impact on both business and personal communications, is the Internet Protocol, or IP. *See* INTEGRATED SERVICES DIGITAL NETWORK (ISDN).

Transmission media and architecture. Wide-area networks may operate on a mix of transmission media for either fixed or mobile applications. Fixed applications mean that the receiver of digital data is stationary. Examples of wireline transmission media for fixed applications are fiber-optic cable, copper wire, and coaxial cable. Many advances in the deployment of wireline media have significantly increased their capacity for handling digital data. For copper, a technique known as digital subscriber line (DSL) allows for transmission in excess of 1 Mbps over regular phone lines. In fiber, a technique known as wave division multiplexing (WDM) allows the simultaneous transmission of different streams of digital data over each spectral component of the light wave. This allows bundles of fiber-optic cable to transport billions of bits (gigabits) and even trillions of bits (terabits) of data per second. Advances in software and hardware-based digital compression also contribute to the ability to transmit more information on a specific medium over a set amount of time. *See* COMMUNICATIONS CABLE; OPTICAL COMMUNICATIONS; OPTICAL FIBERS; TRANSMISSION LINES.

Wide-area networks also operate over a variety of wireless media. Wireless media can support either fixed or mobile applications. Another common distinction in wireless networks is whether it is point-to-point or point-to-multipoint. In point-to-point the originating transmission has one receiver, whereas in point-to-multipoint the originating transmission has multiple receivers. Examples of wireless media include radio-wave, microwave, cellular, and satellite. Different wireless media operate over different parts of the radio spectrum. The allocation and control of this spectrum is managed by the Federal Communications Commission (FCC). *See* COMMUNICATIONS SATELLITE; MOBILE RADIO; RADIO SPECTRUM ALLOCATIONS; RADIO-WAVE PROPAGATION.

Content. There are many different types of content transmitted over WANs. Examples of content are data, voice, video, audio, paging messages, and fax. By virtue of the ability to digitize all of this content, the major difference in transmission requirements is the bandwidth, or capacity, required to transmit digital packets of any type of content. For instance, it takes much more bandwidth to transmit digital video (1-megabit range) than it does to transmit digital audio (64-kilobit range). Certain types of content, such as video and audio, require a network and protocol that deliver the content in real time. The data packets must arrive at the destination in a particular sequence. Other types of content, such as financial data, do not have this requirement. To accommodate these differences, a specific type of WAN architecture, in terms of speed and capacity, must be designed and sized for the content application. *See* DATA COMMUNICATIONS; ELECTRICAL COMMUNICATIONS; FACSIMILE.

Digital compression. In order to utilize the bandwidth capacity of WANs more efficiently, digital compression techniques are now used in many applications. Fundamentally, digital compression reduces the amount of bits in data packets by removing repetitive strings of bits and replacing them with shorter packets that numerically describe the amount of repetitive data. An example is a television picture

with a preponderance of blue sky. Using digital compression, this part of the picture can be represented with a relatively small amount of data. At the receiver site, data are decompressed using the information in the compression packets. The original content can then be almost fully reconstructed for display or processing. Very powerful software compression techniques have been implemented in a wide variety of user applications and consumer devices, such as digital cable and satellite television. This has been made possible by tremendous advances in microprocessor computing power and chipset memory capacities, allowing the implementation of very complex compression algorithms (or software programs) to squeeze more data into existing bandwidth using economical integrated circuit chips for decompression. *See* DATA COMPRESSION.

Digital data security. With massive amounts of information being transmitted over WANs using both public and private infrastructure, data security is increasingly important. In order to secure digital data transmitted over networks, encryption techniques have been developed. Encryption of digital data involves using hardware and software to manipulate the bits in a data packet, making it unrecognizable and unusable to anyone not authorized to use the data. An authorized receiving device on the network contains the corresponding decryption hardware and software to restore the data content to its original format. Often in digital networks, the data are both compressed and encrypted in the same transmission process. Due to the value of the information on the network, it may come under attack by unauthorized parties attempting to decrypt the data. To combat this, some networks are employing various forms of "renewable security," which can be updated from time to time. An example is the use of smartcards in receiving devices, which can be replaced when an encryption algorithm is upgraded. *See* COMMUNICATIONS SCRAMBLING; COMPUTER SECURITY; CRYPTOGRAPHY.

Deployment costs. In deploying WAN infrastructure, there are cost and performance trade-offs involved in determining the appropriate topology of the network. For wireline-based networks, the wire must be either buried underground or strung on existing poles. For wireless-based networks, transmitters must be installed on towers or buildings, or communications satellites must be launched into orbit. Sometimes the business model may dictate that a hybrid of several of these topologies be used. There are several satellite orbital topologies used for different networks: geosynchronous satellites (GEO) for satellite television, medium Earth orbit (MEO) and low Earth orbit (LEO) for telephone and Internet applications. A specific deployment cost factor encountered is known as the "last mile" phenomenon. This refers to the costs to deploy the network to the farthest geographic region of desired coverage. For instance, wireline topographics that are cost-effective in urban areas often cannot be justified in some suburban or rural areas. In this case, use of a hybrid wireless transmission or satellite coverage could be used to fill in these hard-to-access areas. The cost of deploying to these areas must be justified by either the additional business benefit or expected subscriber revenue. Another instance is the type of connection between the wideband transmission medium (such as fiber optics) and the subscriber homes it passes.

Internet. The Internet, using the IP, has become by far the most ubiquitous WAN in the world. Bandwidth and computing power have progressed to the point that Internet users are able to transmit almost any type of content over the network at higher speeds and reasonably good quality. Everything from video clips, electronic mail (e-mail), telephone calls (called Voice Over IP), to digitized x-rays can be transmitted over the Internet. There are three basic variations of IP networks. First is the overall Internet itself which encompasses all personal and business users of the Internet. The second type of IP-based networks is intranets. An intranet is usually deployed within a specific organization or company. A company intranet may be used to manage human resources and financial processes and keep employees updated on company news. The third type of IP-based networks is extranets. Typically, extranets are used to link multiple organizations or companies for some common business purpose. For example, an extranet may be created between an automaker and its parts suppliers to facilitate the processes of ordering, invoicing, and inventory control between the companies. In both intranets and extranets, a technology known as a firewall is employed to prevent unauthorized access to the network. Firewalls are a combination of hardware and software that filter out any unauthorized URL (Uniform Resources Locator) from the network. The URL is the basic unique address or location for any Web site or other Internet service. *See* ELECTRONIC MAIL; INTERNET.

Richard L. Jensen

Bibliography. H.C. Berkowitz, *WAN Survival Guide: Strategies for VPNs and Multiservice Networks*, 2000; R. Cahn and R. Chan, *Wide Area Network Design: Concepts and Tools for Optimization*, 1998; S. Feit, *Wide Area High Speed Networks*, 1999; J. S. Marcus, *Designing Wide Area Networks and Internetworks: A Practical Guide*, 1999; R. W. McCarty (ed.) and M. McGregor, *Cisco WAN Quick Start*, 2000; M. McGregor, *Cisco IOS 12.0 Wide Area Networking Solutions*, 1999.

Wiedemann-Franz law

An empirical law of physics which states that the ratio of the thermal conductivity of a metal to its electrical conductivity is a constant times the absolute temperature, as given by Eq. (1). Here K_c is the ther-

$$K_c = L_0 \sigma T \tag{1}$$

mal conductivity due to the conduction electrons, σ is the electrical conductivity, T is the absolute temperature, and L_0 is known as the Lorentz number. For the case of a degenerate electron gas, the value of L_0

is given by Eq. (2), where k is the Boltzmann con-

$$L_0 = \frac{\pi^2}{3}\left(\frac{k}{e}\right)^2 \tag{2}$$

stant and e is the electronic charge. The Wiedemann-Franz law provides an important check on theories of electrical and thermal conductivity. *See* CONDUCTION (HEAT); ELECTRICAL CONDUCTIVITY OF METALS; FREE-ELECTRON THEORY OF METALS; THERMAL CONDUCTION IN SOLIDS.

Frank J. Blatt

Bibliography. H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids*, 2d ed., 1986; C. Kittel, *Introduction to Solid-State Physics.*, 7th ed., 1995; J. M. Ziman, *Principles of the Theory of Solids*, 2d ed., 1979.

Wilkinson Microwave Anisotropy Probe

A spacecraft which is precisely measuring the cosmic microwave background (CMB) radiation. The *Wilkinson Microwave Anisotropy Probe* (*WMAP*) is a space mission of the National Aeronautics and Space Administration (NASA) that has put fundamental theories of the nature of the universe to a precise test. Since August 2001, *WMAP* has continually surveyed the full sky, mapping out tiny differences in the temperature of the cosmic microwave background radiation, which is the radiant heat from the big bang. A fossil remnant of the hot big bang, the cosmic microwave background permeates the universe and is seen today with an average temperature of only 2.725 kelvins. Tiny variations about this average temperature were first discovered by NASA's *Cosmic Background Explorer* (*COBE*) mission. *WMAP* followed up on the *COBE* results by characterizing the detailed statistical nature of the temperature variations of the cosmic microwave background (called anisotropy), revealing a wealth of detail about the global properties of the universe.

History. In developing the quantitative details of the big bang theory in the late 1940s, Ralph Alpher noted that a remnant radiation should still exist, cooled by the expansion of the universe from its hot and dense earliest stages. In 1965, the cosmic microwave background was discovered accidentally by Arno Penzias and Robert Wilson. In 1992, NASA launched the *COBE* mission, which measured the precise spectrum of the cosmic microwave background and detected its temperature variations at large angular scales (greater than 7°). Recognizing that a wealth of information about the physical conditions of the early universe is encoded in the statistics of the variations at finer angular scales, scientists proposed *WMAP* to map the cosmic microwave background with higher resolution (and greater sensitivity) than *COBE*. *See* BIG BANG THEORY.

WMAP is one of a series of NASA medium-class explorer (MIDEX) satellites. The satellite was launched by a Delta 2 rocket on June 30, 2001, at 3:46 p.m. EDT (1946 UTC) from the Cape Canaveral Air Force Station in Florida. Originally called the *Microwave Anisotropy Probe* (*MAP*), the satellite was renamed in 2003 to honor the memory and accomplishments of David T. Wilkinson, a member of the science team and a pioneer in cosmic microwave background studies.

The mission was proposed to NASA in 1995, selected in April 1996, and confirmed for development in 1997. The satellite, with its single instrument, was built, tested, and launched in only 4 years.

Control of systematic measurement errors. The *WMAP* science requirements dictated that the relative cosmic microwave background temperature be measured accurately over the full sky. The overriding design requirement was to control systematic errors that would otherwise contaminate the measurements. To achieve this, *WMAP* uses differential microwave radiometers that measure temperature differences between pairs of spots on the sky.

To facilitate the separation of the cosmic microwave background from foreground signals from the Milky Way Galaxy, *WMAP* uses polarization-sensitive radiometers at five separate frequency bands centered at 23, 33, 41, 61, and 94 GHz (wavelengths of 13, 9.1, 7.3, 4.9, and 3.2 mm). There are 1, 1, 2, 2, and 4 independent pairs of feed horns per frequency, respectively, with beam sizes of 0.88, 0.66, 0.51, 0.35, and 0.22°. The radiometers are rapidly modulated with a 2.5-kHz phase switch. Amplitude calibration (that is, the determination of the temperature scale of the instrument output signal) relies on the in-flight modulation of the cosmic microwave background temperature caused by the Doppler shift arising from the satellite's motion. Calibration of the beam pattern and angular resolution relies on in-flight observations of Jupiter.

Dual back-to-back Gregorian primary reflectors with apertures of 1.4 × 1.6 m (4.6 × 5.2 ft) focus the microwave radiation from two spots on the sky roughly 140° apart and feed these signals to 10 separate differential receivers that are in an assembly directly underneath the primary optics (**Fig. 1**). Large radiators between the primary optics passively cool the sensitive amplifiers in the receiver assembly below 90 K (−290°F). The bottom half of the spacecraft provides the necessary avionics functions, such as command and data collection electronics, attitude (pointing) control and determination, power services, and a propulsion system. The entire observatory is kept in continuous shadow by a large deployed sunshield, which includes the solar panels that power *WMAP*. *See* ANTENNA (ELECTROMAGNETISM).

Orbit. *WMAP* observes the sky from an orbit about the second Sun–Earth Lagrangian point (L2), 1.5×10^6 km (0.9×10^6 mi) from Earth (**Fig. 2***a*). L2 is about four times farther from Earth than the Moon. This vantage point offers an exceptionally stable environment since the observatory can always point away from the Sun, Earth, and Moon while maintaining an unobstructed view of deep space. *See* CELESTIAL MECHANICS.

Scanning the full sky is an important part of

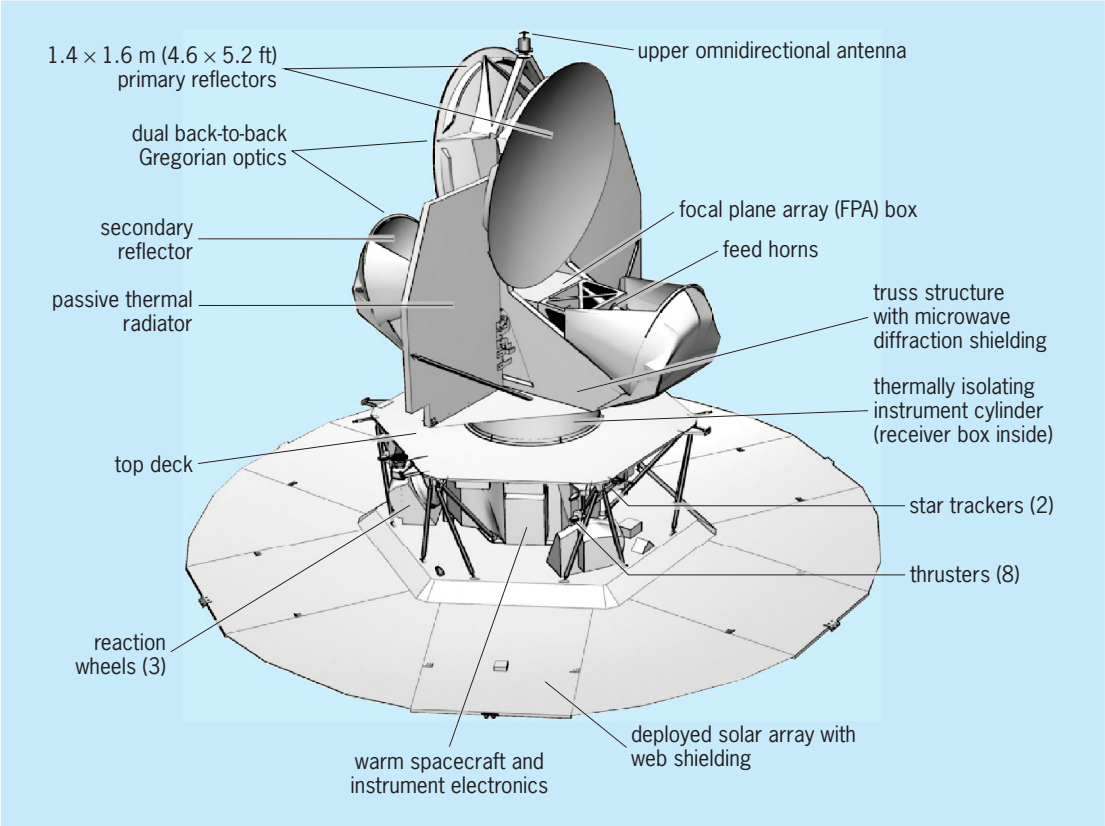

Fig. 1. ***Wilkinson Microwave Anisotropy Probe (WMAP)*** **spacecraft. (*NASA/WMAP Science Team*)**

WMAP's mission. *WMAP* scans the sky in such a way as to cover about 30% of the sky each day (Fig. 2*b*); as the L2 point follows the Earth around the Sun, *WMAP* observes the full sky every 6 months.

The attitude is controlled using two star trackers, two gyroscopes, coarse and fine sun sensors, and three reaction wheels. The satellite spins at 0.464 revolution per minute (~2 min per spin) and precesses at 0.017 revolution per minute (1 h per precession) about a 22.5° cone on the WMAP–Sun

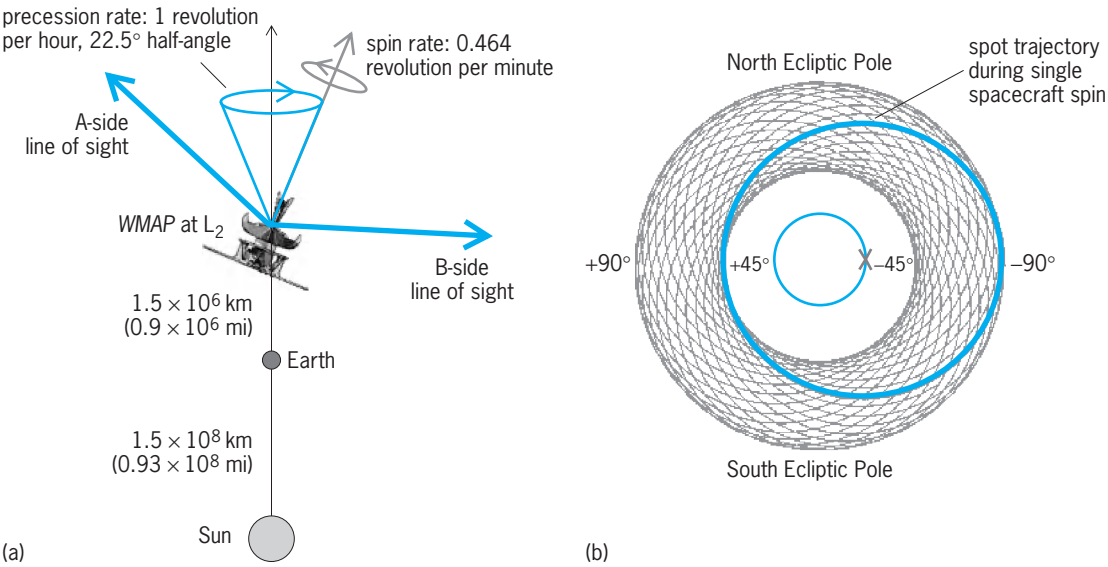

Fig. 2. ***WMAP*** **sky-scan strategy. (*a*) Location of the satellite relative to the Earth and Sun (not to scale), and the orientation of the satellite and of the lines of sight to the spots on the sky focused by the primary reflectors, which continually change as the satellite spins and precesses. (*b*) Sky pattern showing the trajectory of one of the spots focused by the primary reflectors. The ecliptic is horizontal and passes through the center of the diagram. Celestial longitude relative to the antisolar point is indicated. The precession of the spin creates the doughnut pattern that moves along the ecliptic in the course of a year. Full-sky coverage is achieved every six months. (*NASA/WMAP Science Team*)**

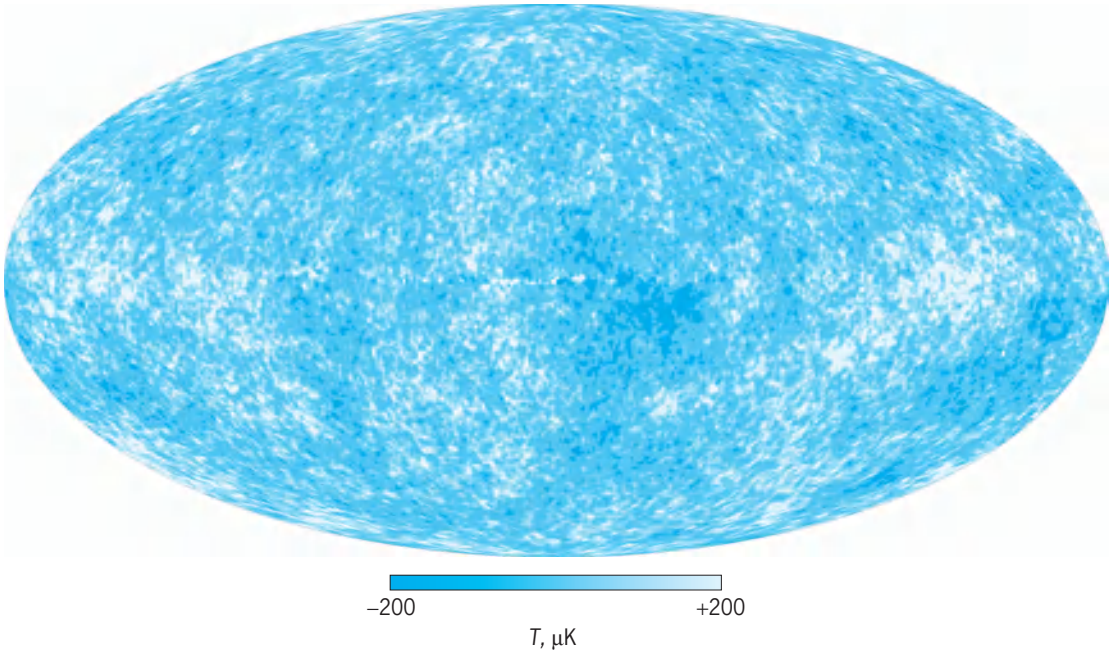

Fig. 3. ***WMAP*** **full-sky cosmic microwave anisotropy map. The microwave light captured in this picture from** ***WMAP*** **is from 379,000 years after the big bang, over 13 × 10^9 years ago. It is a "baby picture" of the universe, equivalent to taking a picture of an 80-year-old person on the day of his or her birth. Shades indicate "warmer" (light) and "cooler" (dark) spots; scale bar indicates variation (*T*) from average temperature. The oval shape is a projection of the full sky in galactic coordinates. This map is a composite from the five** ***WMAP*** **frequencies, and the effects of the Earth's proper motion are removed. (*NASA/WMAP Science Team*)**

line. Blowdown hydrazine propulsion with eight thrusters were used to achieve the L2 orbit, and are also used for station-keeping. *See* GYROSCOPE; STAR TRACKER.

Spacecraft design. The spacecraft structure is made of carbon composite and aluminum materials. The total observatory mass at launch was 840 kg (1850 lb). Communications employ two omnidirectional antennas and a fixed medium-gain antenna, which is used at 667 kilobits per second for daily downlinks to the 70-m (230-ft) antennas of the Deep Space Network. A 3.1-m^2 (33-ft^2) gallium arsenide germanium (GaAs/Ge) solar array oriented 22.5° off the full sun line, and a 23 A-h nickel-hydrogen (NiH) battery provide the required 419 W of power. There are no eclipses in the observing phase of the mission. The design lifetime of 27 months has been exceeded. *See* SPACECRAFT GROUND INSTRUMENTATION.

First results. The first results from the *WMAP* mission were reported on February 11, 2003. This release of the first year of flight data included the most detailed full-sky "baby picture" of the universe taken so far (**Fig. 3**). The *WMAP* data were compared and combined with other diverse cosmic measurements (based on galaxy clustering, supernovae, and so forth), and a new unified and precise understanding of the universe emerged.

The universe is 13.7 × 10^9 years old, with a margin of error of about 1%. The cosmic microwave signal in the *WMAP* map is from 379,000 years after the big bang. The expansion rate of the universe (the Hubble constant) value is $H_0 = 71$ (km/s)/Mpc, with a margin of error less than 5%. The contents of the universe include 4% baryons, 23% cold dark matter, and 73% dark energy. Fast-moving neutrinos do not play a major role in the evolution of structure in the universe. In fact, if they had played such a role, they would have prevented the early clumping of gas in the universe, delaying the emergence of the first stars, in conflict with the new *WMAP* data. *See* COSMOLOGY; DARK ENERGY; DARK MATTER; HUBBLE CONSTANT; NEUTRINO; UNIVERSE.

The *WMAP* data place new constraints on the dark energy. One possibility for the dark energy, Albert Einstein's cosmological constant, is fully consistent with the *WMAP* data. In this scenario the universe will expand forever. However, if the dark energy is due to some other physical effect, the conclusion about the fate of the universe could change. *See* COSMOLOGICAL CONSTANT.

The polarized data also provide new evidence for inflation, the rapid expansion of the universe a fraction of a second after its birth. Many particular inflation models are ruled out; others are supported with this new evidence. The inflationary theory predicts that density is very close to the critical density, producing a flat (Euclidean) universe. *WMAP* has determined, within the limits of instrument error, that the universe is flat. Thus, the results support the big bang and inflation theories. *See* COSMIC BACKGROUND RADIATION; INFLATIONARY UNIVERSE COSMOLOGY.

Charles L. Bennett

Bibliography. C. L. Bennett et al., First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Preliminary maps and basic results, *Astrophys. J. Suppl.*, 148:1, 2003; C. L. Bennett et al., The Microwave Anisotropy Probe (MAP) mission, *Astrophys. J.*, 583:1–23, 2003; M. Chown, *Afterglow of Creation:*

From the Fireball to the Discovery of Cosmic Ripples, University Science Books, 1996; A. H. Guth, *The Inflationary Universe: The Quest for a New Theory of Cosmic Origins*, Perseus Books Group, 1998; M. D. Lemonick, *Echo of the Big Bang*, Princeton University Press, 2003; D. N. Spergel et al., First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Determination of cosmological parameters, *Astrophys. J. Suppl.*, 148:175, 2003.

Willemite

A nesosilicate mineral, composition Zn_2SiO_4, crystallizing in the hexagonal system. It is usually massive or granular with a vitreous luster; crystals are rare. The mineral has a basal cleavage and may be variously colored, most commonly green, red, or brown. Hardness is $5^1/_2$ on Mohs scale; specific gravity is 3.9–4.2. Willemite forms a valuable ore of zinc at Franklin, New Jersey. At this famous zinc deposit willemite fluoresces a yellow-green and is found in crystalline limestone associated with franklinite, zincite, and many other rarer minerals. With the exception of the occurrence in New Jersey, willemite is a rare mineral. *See* SILICATE MINERALS.

Cornelius S. Hurlbut, Jr.

Willow

A deciduous tree and shrub of the genus *Salix*, order Salicales, common along streams and in wet places in the United States, Europe, and China. The twigs are often yellow-green and bear alternate leaves which are characteristically long, narrow, and pointed, usually with fine teeth along the margins (see **illus.**). Flowers occur in catkins; the staminate consist of one to twelve (usually two) stamens, with one or two nectar glands; the pistillate have a single ovary of two carpels, with one to four small glands. The fruit contains several silky seeds. The species are wind- or insect-pollinated and hybridize freely, which often makes identification difficult. *See* SALICALES.

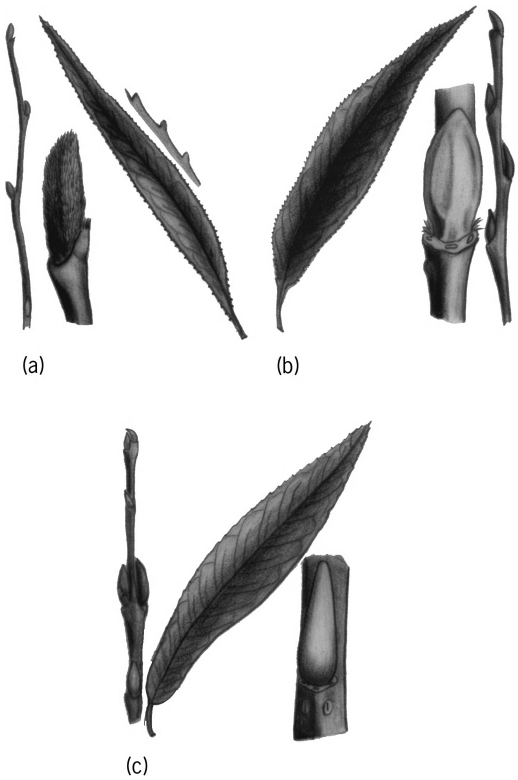

Willow diagnostic features. (*a*) Babylon weeping willow (*Salix babylonia*), twig, terminal bud, and leaf. (*b*) Crack willow (*S. fragilis*), leaf, lateral bud, and twig. (*c*) Purple-osier willow (*S. purpurea*), twig, leaf, and lateral bud.

Willow lumber is used for fuel and in making charcoal, excelsior, ball bats, boxes, crates, boats, waterwheels, and wicker furniture. However, the wood often warps badly in curing unless great care is taken, and it is usually too brittle to be used for heavy construction. The tough, pliable shoots of many species are used to make baskets; the bark of other species is used for tanning. Willows are of great value in checking soil erosion. A few species are ornamental shade trees. *See* FOREST AND FORESTRY; TREE.

James F. Ferry

Wind

The motion of air relative to the Earth's surface. The term usually refers to horizontal air motion, as distinguished from vertical motion, and to air motion averaged over a chosen period of 1–3 min. Micrometeorological circulations (air motion over periods of the order of a few seconds) and others small enough in extent to be obscured by this averaging are thereby eliminated. The choice of the 1- to 3-min interval has proven suitable for the study of the hour-to-hour and day-to-day changes in the atmospheric circulation pattern, and the larger-scale aspects of the atmospheric general circulation.

The direct effects of wind near the surface of the Earth are manifested by soil erosion, the character of vegetation, damage to structures, and the production of waves on water surfaces. At higher levels wind directly affects aircraft, missile and rocket operations, and dispersion of industrial pollutants, radioactive products of nuclear explosions, dust, volcanic debris, and other material. Directly or indirectly, wind is responsible for the production and transport of clouds and precipitation and for the transport of cold and warm air masses from one region to another. *See* ATMOSPHERIC GENERAL CIRCULATION; WIND MEASUREMENT.

Cyclonic and anticyclonic circulation. Each is a portion of the pattern of airflow within which the streamlines (which indicate the pattern of wind direction at any instant) are curved so as to indicate rotation of air about some central point of the cyclone or anticyclone. The rotation is considered cyclonic if it is in the same sense as the rotation of the surface of the Earth about the local vertical, and is considered anticyclonic if in the opposite sense. Thus, in a cyclonic circulation, the streamlines indicate counterclockwise (clockwise for anticylonic) rotation of air about a central point on the Northern Hemisphere or clockwise (counterclockwise for anticyclonic) rotation about a point on the Southern Hemisphere.

When the streamlines close completely about the central point, the pattern is denoted respectively a cyclone or an anticyclone. Since the gradient wind represents a good approximation to the actual wind, the center of a cyclone tends strongly to be a point of minimum atmospheric pressure on a horizontal surface. Thus the terms cyclone, low-pressure area, or low are often used to denote essentially the same phenomenon. In accord with the requirements of the gradient wind relationship, the center of an anticyclone tends to coincide with a point of maximum pressure on a horizontal surface, and the terms anticyclone, high-pressure area, or high are often used interchangeably.

Cyclones and anticyclones are numerous in the lower troposphere at all latitudes. At higher levels the occurrence of cyclones and anticyclones tends to be restricted to subpolar and subtropical latitudes, respectively. In middle latitudes the flow aloft is mainly westerly, but the streamlines exhibit wave-like oscillations connecting adjacent regions of anticyclonic circulation (ridges) and of cyclonic circulation (troughs).

Although the atmosphere is never in a completely undisturbed state, it is customary to refer to cyclonic and anticyclonic circulations specifically as atmospheric disturbances. Cyclones, anticyclones, ridges, and troughs are intimately associated with the production and transport of clouds and precipitation, and hence convey a connotation of disturbed meteorological conditions.

A more rigorous definition of circulation is often employed, in which the circulation C over an arbitrary area bounded by the closed curve S is given by Eq. (1), where the integration is taken completely

$$C = \oint v_t \, dS \tag{1}$$

around the boundary of the area. Here v refers to the wind at a point on the boundary, the subscript t denotes the component of this wind parallel to the boundary, and dS is a line element of the boundary. The component v_t is considered positive or negative according to whether it represents cyclonic or anticyclonic circulation along the boundary S. In this context, the circulation may be positive (cyclonic) or negative (anticyclonic) even when the streamlines within the area are straight, since the distribution of wind speed affects the value of C. *See* ATMOSPHERE; ATMOSPHERIC WAVES, UPPER SYNOPTIC; CLOUD; GEOSTROPHIC WIND; GRADIENT WIND; PRECIPITATION (METEOROLOGY); STORM.

Convergent or divergent patterns. These are said to occur in areas in which the (horizontal) wind flow and distribution of air density is such as to produce a net accumulation or depletion, respectively, of mass of air. Rigorously, the mean horizontal mass divergence D over an arbitrary area A bounded by the closed curve S is given by Eq. (2), where the integration is taken completely around the boundary of

$$D = \frac{1}{A} \oint \rho v_n \, dS \tag{2}$$

the area. Here ρ is the density of air, v refers to the wind at a point on the boundary, and the subscript n denotes the component of this wind perpendicular to the boundary, and dS is an element of the boundary. The component v_n is taken positive when it is directed outward across the boundary and negative when it is directed inward. Convergence is thus synonymous with negative divergence. If spatial variations of density are neglected, the analogous concept of velocity divergence and convergence applies.

The horizontal mass divergence or convergence is intimately related to the vertical component of motion. For example, since local temporal rates of change of air density are relatively small, there must be a net vertical export of mass from a volume in which horizontal mass convergence is taking place. Only thus can the total mass of air within the volume remain approximately constant. In particular, if the lower surface of this volume coincides with a level ground surface, upward motion must occur across the upper surface of this volume. Similarly, there must be downward motion immediately above such a region of horizontal mass divergence.

The horizontal mass divergence or convergence is closely related to the circulation. In a convergent wind pattern the circulation of the air tends to become more cyclonic; in a divergent wind pattern the circulation of the air tends to become more anticyclonic.

Regions which lie in the path of an approaching cyclone are characterized by a convergent wind pattern in the lower troposphere and by upward vertical motion throughout most of the troposphere. Since the upward motion tends to produce condensation of water vapor in the rising air current, abundant cloudiness and precipitation typically occur in this region. Conversely, the area in advance of an anticyclone is characterized by a divergent wind pattern in the lower troposphere and by downward vertical motion throughout most of the troposphere. In such a region, clouds and precipitation tend to be scarce or entirely lacking.

A convergent surface wind field is typical of fronts. As the warm and cold currents impinge at the front, the warm air tends to rise over the cold air, producing the typical frontal band of cloudiness and precipitation. *See* FRONT.

Zonal surface winds. Such patterns result from a longitudinal averaging of the surface circulation. This averaging typically reveals a zone of weak variable winds near the Equator (the doldrums) flanked by northeasterly trade winds in the Northern Hemisphere and southeasterly trade winds in the Southern Hemisphere, extending poleward in each instance to about latitude 30°. The doldrum belt, particularly at places and times at which it is so narrow that the trade winds from the two hemispheres impinge upon it quite sharply, is designated the intertropical convergence zone, or ITCZ. The resulting convergent wind field is associated with abundant cloudiness and locally heavy rainfall. A westerly average of zonal surface winds prevails poleward of the trade wind belts and dominates the middle latitudes of

both hemispheres. The westerlies are separated from the trade winds by the subtropical high-pressure belt, which occurs between latitudes 30 and 35° (the horse latitudes), and are bounded on the poleward side in each hemisphere between latitudes 55 and 60° by the subpolar trough of low pressure. Numerous cyclones and anticyclones progress eastward in the zone of prevailing westerlies, producing the abrupt day-to-day changes of wind, temperature, and weather which typify these regions. Poleward of the subpolar low-pressure troughs, polar easterlies are observed.

The position and intensity of the zonal surface wind systems vary systematically from season to season and irregularly from week to week. In general the systems are most intense and are displaced toward the Equator in a given hemisphere during winter. In this season the subtropical easterlies and prevailing westerlies attain mean speeds of about 15 knots (7.7 m/s), while the polar easterlies are somewhat weaker. In summer the systems are displaced toward the pole by 5 to 10° of latitude and weaken to about one-half their winter strength.

When the pattern of wind circulation is averaged with respect to time instead of longitude, striking differences between the Northern and Southern hemispheres are found. On the Southern Hemisphere, variations from longitude to longitude are relatively small due to the predominance of ocean, and the averaged pattern is described quite well in terms of the zonal surface wind belts. On the Northern Hemisphere there are large differences from longitude to longitude because of the ocean-continent contrast. In winter, for example, the subpolar trough is mainly manifested in two prominent low centers, the Icelandic low and the Aleutian low. The subtropical ridge line is drawn northward in effect over the continents and is seen as a powerful and extensive high-pressure area over Asia and as a relatively weak area of high pressure over North America. In summer the Aleutian and Icelandic lows are weak or entirely absent, while extensive areas of low pressure over the southern portions of Asia and western North America interrupt the subtropical high-pressure belt. *See* CLIMATOLOGY; MONSOON METEOROLOGY.

Upper air circulation. Longitudinal averaging indicates a predominance of westerly winds. These westerlies typically increase with elevation and culminate in the average jet stream, which is found in lower middle latitudes near the tropopause at elevations between 35,000 and 40,000 ft (10.7 and 12.2 km). The subtropical ridge line aloft is found equatorward of its surface counterpart and easterlies occur at upper levels over the equatorward portions of the trade wind belts. In high latitudes, weak westerlies aloft are found over the surface polar easterlies. Seasonal and irregular fluctuations of the circulation aloft are similar to those which characterize the surface winds. *See* JET STREAM.

Variation of wind with height. The rate of change of the wind vector with respect to height is called the wind shear. When the wind direction changes in a clockwise sense with increasing height, the wind is said to be veering with height. For example, if a southerly wind at the surface becomes southwesterly a kilometer above the ground, and westerly in the midtroposphere, the wind veers with height. If the geostrophic wind veers with height, then as a consequence of hydrostatic balance, there is warm advection, that is, warmer air is being blown toward colder air. Veering with height often occurs east of surface cyclones, and west of surface anticyclones in midlatitudes. When the wind direction changes in a counterclockwise sense with increasing height, the wind is said to be backing with height. For example, if a northerly wind at the surface becomes northwesterly a kilometer above the ground, and westerly in the midtroposphere, the wind backs with height. If the geostrophic wind backs with height, then there is cold advection, that is, colder air is being blown toward warmer air. Backing with height often occurs east of surface anticyclones and west of surface cyclones.

The winds may also veer with height as a result of surface friction. Typically, at the surface there is a component of wind across isobars from higher to lower pressure, while above the surface the wind tends to blow parallel to the isobars.

The variation of wind speed and direction with height has a significant impact on the type of thunderstorms which can form. Severe thunderstorms and tornadoes often form when the winds veer and increase in strength with height. *See* THUNDERSTORM; TORNADO.

The expressions veering and backing are used also to describe the change of wind direction with respect to time. For example, north of the track of a low in the Northern Hemisphere the winds back with respect to time, while south of the track of a low the winds veer with time.

Minor terrestrial winds. In this category are circulations of relatively small scale, attributable indirectly to the character of the Earth's surface. One example, the land and sea breeze, is a circulation driven by pronounced heating or cooling of a given area in comparison with little heating or cooling in a horizontally adjacent area. During the day, air rises over the strongly heated land and is replaced by a horizontal breeze from the relatively cool sea. At night, air sinks over the cool land and spreads out over the now relatively warm sea.

Another example is formed by the mountain and valley winds. These result from cooling and heating, respectively, of the mountain slopes relative to the horizontally adjacent free air above the valley floor. During the day, air flows up from the valley along the strongly heated mountain slopes, but at night, air flows down the relatively cold mountain slopes toward the valley bottom. A similar type of descending current of cooled air is often observed along the sloping surface of a glacier. This night-time air drainage, under proper topographical circumstances, can lead to the accumulation of a pool of extremely cold air in nearby valley bottoms.

Local winds. These commonly represent modifications by local topography of a circulation of large

scale. They are often capricious and violent in nature and are sometimes characterized by extremely low relative humidity. Examples are the mistral which blows down the Rhone Valley in the south of France, the bora which blows down the gorges leading to the coast of the Adriatic Sea, the foehn winds which blow down the Alpine valleys, the williwaws which are characteristic of the fiords of the Alaskan coast and the Aleutian Islands, and the chinook which is observed on the eastern slopes of the Rocky Mountains. Local names are also given in some instances to currents of somewhat larger scale which are less directly related to topography. Examples of this type of wind are the norther, which represents the rapid flow of cold air from Canada down the plains east of the Rockies and along the east coast of Mexico into Central America; the nor' easter of New England, which is part of the wind circulation about intense cyclones centered offshore along the Middle Atlantic coastal states; and the sirocco, a southerly wind current from the Sahara which is common on the coast of North Africa and sometimes crosses the Mediterranean Sea. *See* CHINOOK; SIROCCO.

Frederick Sanders; Howard Bluestein

Bibliography. J. A. Dutton, *The Ceaseless Wind: An Introduction to the Theory of Atmospheric Motion*, 1976, reprint 1986; J. R. Holton, *An Introduction to Dynamic Meteorology*, 3d ed., 1992; J. M. Wallace and P. V. Hobbs, *Atmospheric Sciences: An Introductory Survey*, 1977.

Wind measurement

The determination of three parameters: the size of an air sample, its speed, and its direction of motion. Air movement or wind is a vector that is specified by speed and direction; meteorological convention indicates wind direction is the direction from which the wind blows (for example, a southeast wind blows toward the northwest). Anemometers measure wind speed, while wind vanes indicate direction. On average, the wind blows horizontally over flat terrain; however, gusts, thermals, cloud outflows, and many other conditions have associated with them significant short-term vertical wind components. While research wind instruments typically measure both horizontal and vertical air movement, operational and personal wind sensors measure only the horizontal component.

There are many types of wind measurement instruments. In-situ devices measure characteristics of air in contact with the instrument; often they are referred to as immersion sensors because they are immersed in the fluid (air) they measure. Remote wind sensors make measurements without physical contact with the portion of the atmosphere measured. Active remote sensors emit electromagnetic (for example, light or radio waves) or sound waves into the atmosphere and measure the amount and nature of the electromagnetic or acoustic power returned from the atmosphere.

Sampling and measurement principles. Various attributes of instrument performance are used to characterize or compare different instruments. Precision, accuracy, and resolution are perhaps the most widely used attributes. Accuracy refers to the degree to which an instrument can measure the true or correct value. Instrument resolution is the smallest change in an atmospheric variable that can be measured—an instrument's resolution is the same as or better than its accuracy. Precision refers to the ability of an instrument to repeatedly measure the same atmospheric state many times. An instrument may be very precise, yet not be particularly accurate. Errors of imprecision are random—they do not occur in a preferred direction away from the true value. Instrument bias refers to measurement errors that consistently underestimate or overestimate the true value. Threshold response is the value of an atmospheric variable at which an instrument begins to respond. A mechanical anemometer may not register a response until the wind speed exceeds a certain value, for example, 0.5 m/s (1 mi/h). At this threshold speed, the force on the anemometer's cups or propeller is able to exceed the friction in the instrument's bearings. Other useful attributes of wind instruments are time constant and distance constant. Time constant is the length of time that the instrument requires to respond to an instantaneous change in atmospheric state; and it is expressed as the time required to indicate 63% of the actual change. Distance constant refers to the amount of air movement (meters) required to indicate a similar amount of change in either wind speed or wind direction.

The quality of atmospheric measurements depends not only on the performance attributes of the instruments but also on the sampling and siting characteristics. The representativeness of a measurement can be affected significantly by the location or siting of the instrument. A low-level wind measurement, for example, can represent the general wind flow pattern of the region, or it can represent the disturbance caused by nearby buildings or vegetation; both measurements are representative, but of very different conditions. Similar considerations exist for all in situ measurements and many remote measurements. Sampling characteristics refer to the manner in which the output signal from the instrument is processed. Sampling rate is the frequency at which the output signal is detected or observed. A wind sock may be sampled visually (that is, observed) at a rate of once per hour, whereas the sampling rate for the digital output of a sonic anemometer may be 20 per second. The length of time that samples are accumulated and subsequently processed is the sampling period. In sampling a phenomenon that has an inherent periodicity, the sampling rate must be greater than the frequency of the atmospheric fluctuations in order to avoid aliasing. If the sampling is aliased, high-frequency phenomena will be indicated as an unknown enhancement of lower-frequency processes. Often the sampled signal is filtered to remove certain frequencies from the processed data set.

High-frequency fluctuations may represent instrument or electronic noise and are removed with a low-pass filter. Low-frequency fluctuations, such as long-term trends, are removed with a high-pass filter. Using a combination of low- and high-pass filters constitutes a band-pass filter that passes the intermediate frequencies.

Ground-based in-situ instruments. There are three principal categories of ground-based in-place wind sensors: mechanical, pressure, and sonic. Mechanical systems are the most common. Virtually all anemometers determine wind speed by measuring the rotational speed of a propeller or a circular array of cups. Wind direction is measured in mechanical devices by the direction of a vertically oriented vane or fin mounted on a horizontal shaft. In a steady wind, the forces on the two sides of the vane are equal. As the wind changes direction, the force on the upwind or windward side of the vane initially exceeds that on the downwind or leeward side until the vane rotates into the wind and the forces again are balanced. Most wind instruments measure wind speed and direction separately, although the propeller vane measures the horizontal wind vector by mounting a propeller at the windward end of a horizontal shaft with a vane on its leeward end. Cups, propellers, and vanes have the common (although usually minor) disadvantage that their inertia causes some overresponse to changes in speed or direction.

The pressure anemometer measures the dynamic pressure exerted by the wind. A Pitot tube—an L-shaped piece of circular tubing open at the windward end—is oriented into the airstream and measures the sum of the static atmospheric pressure and the dynamic pressure of the wind. The dynamic pressure is proportional to the square of the wind speed. A manometer or differential pressure transducer measures the pressure difference between the Pitot tube and a reference. Because the pressure differential responds to the wind speed squared, pressure anemometers are very responsive and accurate at high wind speeds but poorly responsive at speeds less than 1–2 m/s (2–4 mi/h). They frequently are used in wind tunnels and on airplanes; when used for ambient measurements, they must be oriented into the wind by mounting on a vane. *See* PITOT TUBE.

The measurement principle of the sonic anemometer is the dependence of the speed of a sound wave on the wind speed. A sound wave traveling with the wind moves with the sum of the speed of sound and the wind speed, while a wave traveling into the wind has a reduced speed. The three components of the wind vector (vertical, north, and east) are measured by using three pairs of sonic transducers to measure the effective sound speed in three different directions. Sonic anemometers were traditionally used as research instruments, but they have become operational instruments because of their increased reliability and somewhat reduced costs. *See* ANEMOMETER.

Upper-air in-situ measurements. Vertical profiles (usually called soundings) of winds and temperature, humidity, and pressure from the ground surface to pressure altitudes around 25 hectopascals or millibars (25 km or 15 mi) and higher are obtained operationally with a balloon-borne radiosonde. There are nearly 100 radiosonde stations in the United States that launch radiosondes twice daily at 00Z and 12Z (Greenwich mean time or universal standard time) and more than 700 global stations that launch at least one radiosonde daily. The radiosonde consists of expendable sensors, electronics, and radio transmitter to telemeter the measurements back to a receiver at the launch station. Winds are not measured directly, but are determined from the movement of the radiosonde suspended beneath the balloon. One class of wind measurement techniques actually tracks the balloon with one of three methods: (1) an optical system using a theodolite tracks the balloon's azimuth and elevation; (2) a radio theodolite tracks the radio signal from the radiosonde transmitter; and (3) radar systems track a radar retro reflector suspended from the balloon to obtain slant range, azimuth, and elevation. The first two methods require knowledge of the balloon height, which is determined from the pressure measurement. Optical theodolite measurements require clear conditions between the ground station and the balloon. All three methods suffer form a loss of accuracy as the distance to the balloon increases and its elevation angle decreases. *See* SURVEYING INSTRUMENTS; TELEMETERING.

The second class of radiosonde wind measurement techniques uses any of two different radio navigation systems. One system uses the Loran-C radio navigation system. A transponder on the balloon receives the navigation (time-of-arrival) signals and rebroadcasts them back to the ground station, where the wind is determined by tracking the change in position of the balloon. This system requires a pressure measurement to determine the height of the balloon. Another type of navigation-based wind-finding technique uses the Global Positioning System (GPS). It can measure very accurately the actual velocity of the radiosonde, and with some GPS radiosonde receivers it is also possible to measure latitude, longitude, and altitude. Two major advantages of these devices are the high accuracy and precision of the wind measurements and the worldwide coverage of GPS. *See* AIR NAVIGATION; LORAN; SATELLITE NAVIGATION SYSTEMS.

Dropwindsondes are the airborne counterpart to conventional radiosondes and are used primarily in research studies. They are ejected from aircraft and float to the ground on a parachute, measuring winds in the same way as the various standard radio navigation wind-finding methods.

Remote wind sensors. Advances in active remote sensing techniques have led to the transition of radar wind profilers from research devices to operational instruments. Wind profilers transmit a short pulse of radio-frequency electromagnetic energy into the atmosphere, where a small fraction is scattered back from the atmosphere to the profiler's antenna. The intensity of the energy returned from the atmosphere depends on minor irregularities in the index

of refraction—a measure of the temperature, humidity, and pressure characteristics of the air. These irregularities (sometimes referred to as turbulent eddies) are transported with the wind and so act as tracers of air movement. As a result, the frequency of the returned signal is changed (the Doppler shift) according to the speed of the wind along the pointing (radial) direction of the transmitted electromagnetic pulse. Measuring the Doppler shift of the returned signal provides a measure of the radial wind speed component as a function of distance from the antenna (that is, height above the ground). The height increment in which an independent wind speed measurement can be made is equal to one-half the length of the transmitted radio-frequency pulse. By alternately pointing the transmitted beam in various directions, the three-dimensional wind vector can be resolved over successive height intervals (ranges). Some wind profilers can resolve the wind within height intervals about 60 m (200 ft). All wind profilers operate in clear air. The lower-frequency profilers (50 and 400 mhz) are insensitive to rain, while higher-frequency profilers (21 GHz) respond to both clear air and rain; in some cases it is possible to estimate both the wind speed and the rain-drop sizes.

A newer type of radar of radar wind profiler is the spaced antenna system. It measures the wind by sensing the shift in the returned radio signal across a horizontal array of antennas. Spaced antenna systems have the advantage that the wind vector is measured from the return of single, vertically pointed (transmitted) beam, thereby decreasing both the sampling period and the sampled volume of air. Beam-pointing Doppler systems, on the other hand, measure the wind over a broad cone of air formed by three or five different beams. The sampling period required to obtain a representative vertical profile is of order one minute with the spaced antenna profiler, considerably shorter than the Doppler beam-swinging systems.

Lidar is very similar to radar except that it uses a laser to emit pulses of light, which are scattered by small particles in the air. Sodar (sound detection and ranging) refers to wind profilers that use an acoustic pulse to detect the wind in much the same way as radar wind profiles. The acoustic pulses are scattered back to the receiver on the ground by temperature fluctuations in the air. The height coverage with sodar is typically less than with radar or lidar—usually no more than a few kilometers—but the height resolution can be considerably better; mini-sodars can resolve winds over height intervals of a few meters. A major disadvantage of sodar is the irritation of the noise of the acoustic pulses. Unlike the radar wind profiles that are being used as part of operational weather observing systems, sodar and lidar remain primarily research instruments. *See* LIDAR; METEOROLOGICAL INSTRUMENTATION; METEOROLOGICAL RADAR; WIND.

Walter F. Dabberdt

Bibliography. F. V. Brock (ed.), *Meteorological Measurement Systems*, 1995; F. Dobson, L. Hasse, and R. Davis, *Air-Sea Interaction*, 1980; L. J. Fritschen and L. W. Gay, *Environmental Instrumentation*, 1979; D. H. Lenschow (ed.), *Probing the Atmospheric Boundary Layer*, 1986; A. Wexler, *Measurement of Humidity in the Free Atmosphere near the Surface of the Earth: Meteorological Observations and Instrumentation*, Meteorol. Monog. 33, American Meteorological Society, 1970; World Meteorological Organization, *Guide to Meteorological Instruments and Methods of Observation*, WMO 8, 1983.

Wind power

The extraction of kinetic energy from the wind and conversion of it into a useful type of energy: thermal, mechanical, or electrical. Wind power has been used for centuries.

Early development. Windmills were used extensively in the Middle East by the eleventh century and in Europe by the thirteenth century. By the fourteenth century the Dutch used them for draining marshes and lakes. By the end of the nineteenth century, in Denmark there were about 2500 electricity-producing windmills with a generating capacity of about 30 MW, and about 4600 windmills being used for agricultural applications. Farmers and ranchers in the western United States began using small water-pumping windmills in the midnineteenth century. These wind machines were simple durable multibladed turbines with a high starting torque. An estimated 6.5 million machines were sold between 1880 and 1930, and in the 1930s the annual energy output was equivalent to 10^9 kWh per year. Beginning about 1925, small 0.2–3 kW horizontal-axis machines supplying dc electricity to operate appliances, usually at 32 V, were used in rural areas of the United States (**Fig. 1**). The creation of the Rural Electrification

Fig. 1. A small dc horizontal-axis wind machine for rural use.

Administration in 1936 soon brought electrical service to these rural areas and doomed the widespread use of the small machines.

In 1941, a 175-ft-diameter (53-m) 1250-kW-capacity wind turbine was added to the generating system of Central Vermont Public Service Corporation. The machine was located on a 2000-ft (600-m) knob near Rutland, Vermont, and operated until 1945, when one of the blade spars broke and a blade was thrown 750 ft (230 m). The machine was scrapped after the accident, and for all practical purposes the development of wind power in the United States was discontinued until the early 1970s.

Energy in the wind. It has been estimated that the total wind power in the atmosphere averages about 3.6×10^{12} kW, which is an annual energy of about 107,000 quads (1 quad = 2.931×10^{11} kWh). Obviously, only a fraction of this wind energy can be extracted, estimated to be a maximum of 4000 quads per year. The power P in the wind per unit area is given by Eq. (1), where V is wind speed in m/s

$$P = \frac{\rho V^3}{2} \qquad \mathrm{W/m^2} \tag{1}$$

and ρ is the air density in kg/m^3, which varies from about 1.225 at sea level to 75% this value at 11,000 ft (3000 m) elevation. According to what is commonly known as the Betz limit, a maximum of 59% of this power can be extracted by a wind machine. Practical machines actually extract from 5 to 45% of the available power. Because the available wind power varies with the cube of wind speed, it is very important to find areas with high average wind speeds to locate wind machines. Examples of good areas in the United States are certain mountain passes and ridges in California and Hawaii, and the flat plains in the midwestern regions of the United States. *See* WIND.

Wind-produced electricity. Most research on wind power has been concerned with producing electricity. This effort restarted in the United States in the early 1970s. At present, nearly all the electricity produced in the United States is generated by fossil fuel plants (oil, natural gas, and coal), nuclear plants, and hydroelectric plants. There is concern about fossil fuel generation because the carbon dioxide produced when fossil fuels are burned contributes to the greenhouse effect. Disposal of radioactive waste and concerns about accidents have greatly diminished the expected role of nuclear power, and there is limited hydroelectric power left available for development. The best new sources are conservation and renewable generation. *See* ELECTRIC POWER GENERATION; ENERGY SOURCES; NUCLEAR POWER.

Wind power is a renewable energy source that has virtually no environmental problems. However, wind power has limitations. Wind machines are expensive and can be located only where there is adequate wind. These high-wind areas may not be easily accessible or near existing high-voltage lines for transmitting the wind-generated energy. Another disadvantage occurs because the demand for electricity varies with time, and electricity production must follow the demand cycle. Since wind power varies randomly, it may not be available when needed. The storage of electrical energy is difficult and expensive, so that wind power must be used in parallel with some other type of generator or with nonelectrical storage. Wind power teamed with hydroelectric generators is attractive because the water can be used for energy storage, and operation with underground compressed-air storage is another option. *See* ENERGY STORAGE; PUMPED STORAGE; WATERPOWER.

At present, the best locations for wind power are sites where average wind speeds are high and where nearby conventional generation has a high incremental cost, so that when wind power is available it can produce large cost savings. This combination of factors occurs in California. As a result, by the end of 1985 there were about 13,370 wind generators operating in that state with a total capacity of approximately 1120 MW (a large nuclear plant is about 1100 MW). Most of these machines are grouped into wind farms in mountain passes.

Types of wind machines. The most common type of wind turbine for producing electricity has a horizontal axis, with two or more aerodynamic blades mounted on the horizontal shaft. The wind machine shown in Fig. 1 is a small horizontal-axis turbine. With a horizontal-axis machine, the blade tips can travel at several times the wind speed, which results in a high efficiency. The blade shape is designed by using the same aerodynamic theory as for aircraft. *See* PROPELLER (AIRCRAFT).

Several large wind machines were built by the National Aeronautics and Space Administration under the sponsorship of the Department of Energy. The Mod-0 had a two-bladed 125-ft-diameter (38-m) rotor, which rotated at 40 revolutions/min and drove a 100-kW synchronous generator. The rotor height was 100 ft (30.5 m). The Mod-1 had a two-bladed, 202-ft-diameter (61.5-m), 35-revolutions/min rotor, which drove a 2000-kW synchronous generator. The hub height was 140 ft (42.7 m). The Mod-2 (**Fig. 2**) had a two-bladed, 300-ft (91.4 m), 17.5-revolutions/min rotor driving a 2500-kW synchronous generator. The hub height was 200 ft (61 m). The Mod-0 and a later version named the Mod-0A operated satisfactorily. The Mod-1 and Mod-2 had serious problems.

Many other small horizontal-axis machines are produced. The wind farms in California are mainly composed of these machines, with sizes ranging from a few kilowatts to over 200 kW per machine. Vertical-axis machines have also been investigated. The most common vertical-axis machine is the Darrieus or egg-beater. Two or more curved blades are attached to each end of a vertical shaft. The Darrieus machine has not been as widely used as the horizontal-axis machine for producing electricity.

Estimating energy output. The energy output from a wind machine can be estimated from the graph of machine power output versus wind speed and the wind characteristics at a site. A simplified graph of

Fig. 2. The 2500-kW Mod-2 wind machine at Goodnoe Hills near Goldendale, Washington.

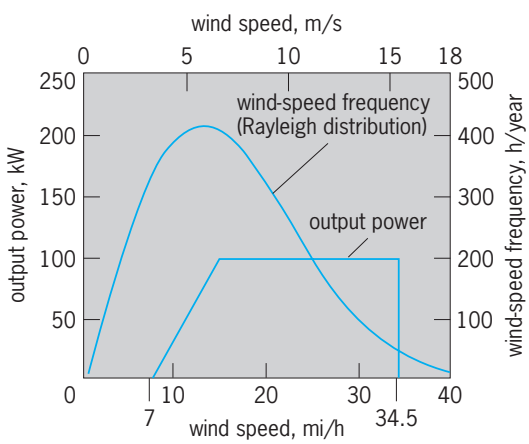

Fig. 3. Approximate Mod-0 power output curve and Rayleigh wind-speed distribution for average wind speed of 16 mi/h (7.2 m/s). The sharp cutoff in output power at a wind speed of 34.5 mi/h (15.4 m/s) is caused by the wind machine shutting down for safety at high speeds.

output power as a function of wind speed measured at a height of 30 ft (9.1 m) for the Mod-0 wind machine is shown in **Fig. 3**.

Often the wind characteristic at a site is simply specified by the average wind speed. However, wind speed V is a random variable about this average value. The Rayleigh distribution, given in Eq. (2), often

$$f(V) = \frac{2V}{c^2} e^{-(V/c)^2} \qquad (2)$$

approximates the wind-speed variation. Here, $f(V)$ is the frequency with which various wind speeds occur, and c is related to the average wind speed V_{av} by Eq. (3). The Rayleigh distribution of wind-speed

$$c = \frac{2V_{av}}{\sqrt{\pi}} \qquad (3)$$

frequencies for $V_{av} = 16$ mi/h (7.2 m/s) is shown in Fig. 3. This curve gives the hours per year the wind is in each 1-mi/h range. For example, it can be seen that the wind would be at 20 mi/h about 320 h per year. *See* PROBABILITY; STATISTICS.

The yearly wind-machine output can be determined by multiplying the output at each wind speed by the number of hours that the wind is at that speed and then summing for all wind speeds. For the example of Fig. 3, it is calculated that this wind machine would produce 590,000 kWh per year if it was available for operation 100% of the time. This means the wind machine has an average output of 67% of its maximum, which is a very high percentage. However, a wind site with an average wind speed of 16 mi/h (7.2 m/s) at 30 ft (9.1 m) elevation is also an exceptionally good wind site. Gary Thomann

Bibliography. T. Ackerman (ed.), *Wind Power in Power Systems*, 2005; T. Burton et al., *Wind Energy Handbook*, 2001; J. F. Manwell et al., *Wind Energy Explained*, 2002; D. A. Spera (ed.), *Wind Turbine Technology: Fundamental Concepts of Wind Turbine Engineering*, 1994.

Wind rose

A diagram in which statistical information concerning the direction and speed of the wind at a particular location may be conveniently summarized. In the standard wind rose a line segment is drawn in each of perhaps eight compass directions from a common origin (see **illus.**). The length of a particular

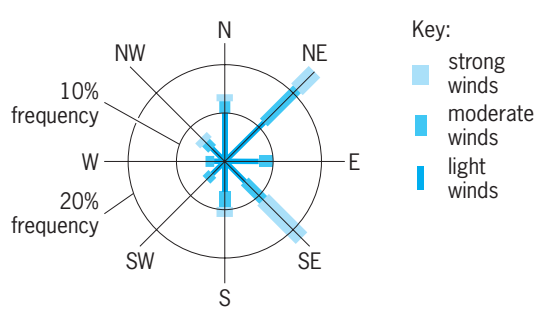

Standard wind rose.

segment is proportional to the frequency with which winds blow from that direction. Parts of a given segment are given various thicknesses, indicating frequencies of occurrence of various classes of wind speed from the given direction. *See* WIND MEASUREMENT. Frederick Sanders

Wind stress

The drag or tangential force per unit area exerted on the surface of the Earth by the adjacent layer of moving air. Erosion of ground surfaces and the production of waves on water surfaces are manifestations of wind stress. Surface wind stress determines the exchange of momentum between the Earth and the atmosphere and exerts a strong influence on the typical variation of wind through the lowest kilometer of the atmosphere. Estimated values of the surface wind stress range up to several dynes per square centimeter (0.1 pascal), depending on the nature of the surface and the character of the adjacent airflow. *See* METEOROLOGY.

Internal horizontal stresses. Significant stresses arise within the lower atmosphere because of the strong shear of the wind between the slowly moving air near the ground and the more rapidly moving air a kilometer above and because of the turbulent nature of the airflow in this region. The turbulent eddies referred to here have characteristic dimensions ranging up to a few hundreds of meters.

The effectiveness of this turbulent viscosity in the transferring of momentum from level to level may, in favorable circumstances, be a million times as great as the effectiveness of purely molecular viscosity. These turbulent viscous stresses exert an indirect effect on the surface wind stress by effecting the transfer of large amounts of momentum from higher levels down to levels adjacent to the surface of the Earth.

The torque exerted by the surface wind stress averaged over the entire Earth and over a sufficiently long period of time must be equal to zero. Otherwise the net torque acting between the Earth and the atmosphere would tend to alter the rate of rotation of the Earth about its axis and to alter the mean wind circulation of the atmosphere. In the atmosphere the angular momentum supplied, in effect, by the Earth in regions of easterly surface flow balances the angular momentum drained by the Earth in regions of westerly wind flow, so that the average torque exerted by the surface wind stress is indeed zero, or very nearly so. The average magnitude of the surface wind stress is, of course, not equal to zero. It determines the average rate at which kinetic energy of the winds is dissipated by surface friction. *See* ANGULAR MOMENTUM; ATMOSPHERE.

Wind pressure. This is the force exerted by the wind per unit area of solid surface exposed normal to the wind direction and is also known as dynamic pressure. In contrast to shearing stresses, the wind pressure arises from the difference in pressure between the windward and lee sides of the exposed surface. Wind pressure thus represents a substantial force when the wind speed is high. *See* WIND.

Frederick Sanders

Stress over sea. The drag or tangential force of the wind on the sea is expressed in units of dynes per square centimeter or micronewtons per square meter but is normally taken to represent the mean drag over an undefined area, perhaps several kilometers square, containing many waves. It is usually related to an appropriate time and space average of the wind near the sea surface (at 10 m above the mean level, for example).

Over most surfaces overland, in the absence of vertical density gradients, the drag varies as the square of the wind speed and is expressed, following aerodynamic usage, by the equation shown below, where

$$\tau = C_D \rho v^2$$

τ is the shearing stress, ρ is the air density, and v is the wind speed at a specified height, say 10 m. For a fixed (uniform) surface geometry, or roughness, C_D is a constant known as the drag coefficient. *See* AERODYNAMIC FORCE.

Over a water surface a drag coefficient is defined in the same way, but since the surface geometry varies with the wind speed, C_D is to be regarded as a variable. The drag coefficient over the sea is an important quantity in both meteorology and oceanography since it relates the wind speed to the drag, which generates ocean waves, drives the ocean currents, and sets the scale of the atmospheric turbulence that transfers water vapor and heat from the ocean to the atmosphere to provide the energy for clouds and weather systems. *See* ATMOSPHERIC GENERAL CIRCULATION; MARITIME METEOROLOGY; OCEAN CIRCULATION; OCEAN WAVES.

The drag coefficient of the sea surface depends on the wave field and on the turbulent structure of the flow in the air and the water. Present knowledge of the complicated fluid mechanics involved is not sufficient to allow theoretical calculation of it.

Methods of measurement. Many methods have been used to measure the drag of the wind on the sea, but they are of four main types. The first technique attempts to get information from the wind tilt of enclosed bodies of water. When wind blows over a lake, the surface becomes tilted upward toward the lee end. Measurements of the mean slope, which typically would be of the order of 1 cm in 10 km, can provide estimates of the wind drag. There are theoretical difficulties associated with end effects, but the major difficulties are practical: steady isothermal conditions are required, and although the surface waves can be filtered out to permit observations of the necessary accuracy, it is often hard to eliminate the longperiod oscillations of the whole body of water, which are known as seiches. *See* SEICHE.

The second method involves the measurement of the vertical of the mean wind in the lowest few meters of air over the sea. When air and water are at the same temperature, wind speed increases as the logarithm of the height, and its rate of increase provides a measure of the wind drag. This method is empirical, but it is well established; the difficulties are of engineering rather than of science. It is difficult to get instruments of the necessary precision (the wind speed has to be measured to about 1 cm/s) and even harder to expose them over water in such a way that their supports do not interfere with the wind or the waves or the currents. For this reason many

of the observations have been made in lakes and in shallow water, although some workers have developed techniques for making wind measurements accurately enough from small ships or buoys.

The third technique again involves measuring the vertical distribution of the mean wind, but in this case to greater heights. In steady conditions the drag of the wind on the sea can be inferred from the component of the wind across the isobars; therefore, for this method accurate measurements of the distribution of atmospheric pressure, as well as accurate wind observations, are necessary. This method does not work well in middle latitudes because in those areas the wind blows at only a small angle to the isobars in the lowest kilometer or so. It works better in the trade winds, where attempts to use it have met with moderate success.

The fourth technique again requires detailed wind measurements, in this case detailed in time by using an instrument of rapid response, reacting, for example, in about one-tenth of a second. Measurements are made of both the horizontal and vertical components of the wind, and their mean product provides the drag. There are again some theoretical difficulties in the interpretation of the results, but they can probably be overcome. The major difficulties are again of getting suitable instruments suitably exposed, and of the large amount of numerical analysis that has to be done to get the results. On the other hand, the method provides detailed information on the structure of the turbulent flow, and it is being increasingly used. A further attraction of this method is that estimates of evaporation and direct heat transfer can be made by measuring fluctuations of humidity and temperature as well as those of the vertical velocity component. The drag can also be inferred from observations of the fluctuations at high frequencies by deducing the energy dissipation, with the assumption that the smallscale turbulence is isotropic.

Measured values. There is substantial agreement that the drag of the wind on the sea is small relative to that of a fixed soil surface with the same geometry. It is largely independent of the fetch and so seems to depends less on the larger waves than on the short waves and ripples. Surface-active agents, which affect the shortest waves, may therefore be important.

J. R. Garratt provided evidence that C_D increases slowly with wind speed υ from about 10^{-3} at 5 m · s^{-1} to about 2×10^{-3} at 20 m · s^{-1}, possibly reaching 4×10^{-3} at 50 m · s^{-1}.

This is consistent with the notion that the effective aerodynamic roughness of the sea surface is itself determined by the local shearing stress and the acceleration of gravity. But the mechanism remains obscure. Henry Charnock

Bibliography. S. L. Hess, *Introduction to Theoretical Meteorology*, 1959, reprint 1979; D. Houghton and F. Sanders, *Weather at Sea*, 3d ed., 1998; F. K. Lutgens and E. J. Tarbuck, *The Atmosphere: An Introduction to Meteorology*, 8th ed., 2000; A. Miller and J. C. Thompson, *Elements of Meteorology*, 6th ed., 1999; P. Sachs, *Wind Forces in Engineering* 2d ed., 1978; S. D. Smith, Coefficients for sea surface wind stress, heat flux and wind profiles as a function of wind speed and temperature, *J. Geophys. Res.*, 93(C12):15467–15472, 1988.

Wind tunnel

A duct in which the effects of airflow past objects can be determined. The steady-state forces on a body held still in moving air are the same as those when the body moves through still air, given the same body shape, speed, and air properties. Scaling laws permit the use of models rather than full-scale aircraft. Models are less costly and may be modified more easily than aircraft, and conditions may be simulated that would be impossible or dangerous in flight.

Related aerodynamic research equipment includes ballistic guns, drop models, rockets, whirling arms, rocket-powered sleds, and flight tests. The wide range of problems studied requires a similarly wide range of specialized wind tunnels. *See* AIRCRAFT TESTING.

Uses and Methods

Determinations commonly made with wind tunnels are as follows:

- Drag of airplanes and missiles
- Lifting characteristics of winged vehicles or lifting bodies
- Static stability of aircraft, including missiles
- Dynamic stability derivatives of aircraft
- Torques required to deflect control surfaces
- Pressure distributions to determine air loads
- Flutter characteristics of flexible aircraft
- Distribution and rate of heat transfer to aircraft parts for cooling and structural design
- Performance of air-breathing engines and inlet airflow
- Performance of propellers
- Conditions for safe release of bombs and missiles
- Effect of wind on buildings, bridges, signs, automobiles, and other nonflying structures
- Nature of smoke flow from factories or ships
- Nature of wind flow around geographic features

Test conditions. The tunnels in which experiments are made should closely match the conditions of full-scale flight, but it is not always possible to match all the scale parameters. An attempt is always made to match scale and the effects of compressibility of the air. *See* DIMENSIONAL ANALYSIS; DIMENSIONLESS GROUPS; DYNAMIC SIMILARITY; MODEL THEORY.

If compressibility effects are neglected, scale is duplicated and flow patterns are similar when model tests and full-scale flight are at the same Reynolds number (Re). The Reynolds number is proportional to the product of speed and a characteristic length. *See* REYNOLDS NUMBER.

Compressibility effects are similar when the ratio of remote airspeed to the velocity of sound, called

Mach number (M), is duplicated. *See* MACH NUMBER.

At flight speeds below a Mach number of approximately 0.7, one need be concerned only with the Reynolds number, since the effects of the Mach number are small and may be readily calculated. It is usually sufficient to require that the Reynolds number be greater than 1,500,000. If the flight speed is above M = 0.7, the effects of compressibility are less predictable and the Mach number must be matched. If the Reynolds number exceeds 4,000,000 the scale match is usually adequate. However, research has identified problems in the transonic speed range associated with a Reynolds number above 100,000,000. No wind tunnels are available that can match that combination of Reynolds number and M. *See* TRANSONIC FLIGHT.

Special tests require the matching of other similarity parameters. Where dynamic characteristics are desired, the mass and inertia must be scaled with the forces to produce similar flight paths. If the model to be simulated is flexible, its elastic properties must be matched.

Data taken from models in wind tunnels suffer from extraneous effects due to the structure that supports the model and from the tunnel walls. The effects of support interference are minimized by design and are evaluated by careful experiments. The interference of the walls is calculated from theoretical considerations and is evaluated for each tunnel and type of model.

Measurements. Most data are acquired from wind tunnels through measurement of forces and moments, surface pressures, changes produced in the airstream by the model, local temperatures, and motions of dynamically scaled models, and by visual studies.

Force and moment measurements. A balance system separates and measures the six components of the total force. The three forces, taken parallel and perpendicular to the flight path, are drag, lift, and side force. The three moments about these axes are yawing moment, rolling moment, and pitching moment, respectively.

Surface pressure. Surface pressures are measured by connecting orifices flush with the model surface to pressure-measuring devices. Local air load, total surface load, moment about a control surface hinge line, boundary-layer characteristics, and local Mach number may be obtained from pressure data.

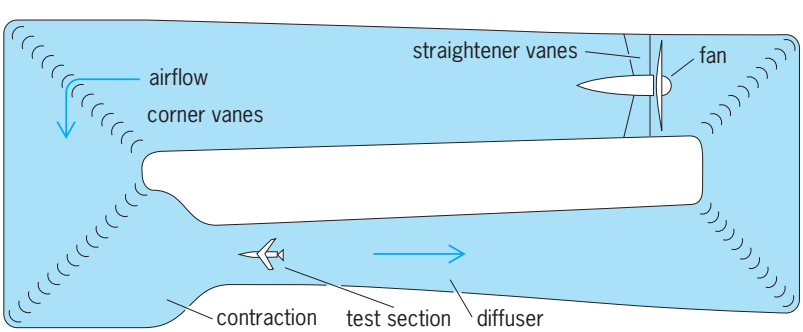

Fig. 1. Outline plan view of low-speed return-flow wind tunnel.

Reaction of model on airstream. Measurements of stream changes produced by the model may be interpreted in terms of forces and moments on the model. In two-dimensional tunnels, where the model spans the tunnel, it is possible to determine the lift and center of pressure by measuring the pressure changes on the floor and ceiling of the tunnel. The parasite drag of a wing section may be determined by measuring the total pressure of the air which has passed over the model and calculating its loss of momentum.

Surface temperature. Measurements of surface temperatures indicate the rate of heat transfer or define the amount of cooling that may be necessary.

Dynamic measurements. In elastically and dynamically scaled models used for flutter testing, measurements of amplitude and frequency of motion are made by using accelerometers and strain gages in the structure. In free-flight models, such as bomb or missile drop tests, data are frequently obtained photographically. *See* ACCELEROMETER; FLUTTER (AERONAUTICS); STRAIN GAGE.

Flow visualization. At low speeds, smoke and tufts are often used to show flow direction. A mixture of lampblack and kerosine painted on the model shows the surface streamlines. A suspension of talcum powder and a detergent in water does the same, is cleaner, and has a more pleasant odor.

At velocities near or above the speed of sound, some flow features may be made visible by optical devices. *See* INTERFEROMETRY; SCHLIEREN PHOTOGRAPHY.

Types

There are various types of wind tunnels, including low-speed tunnels; V/STOL tunnels; transonic, supersonic, and hypersonic tunnels; nonaeronautical tunnels; and hypervelocity tunnels.

Low-speed wind tunnel. This tunnel has a speed up to 300 mi/h (480 km/h). A low-speed tunnel has the essential features of most wind tunnels. These include a device to drive the air, a duct that provides smooth, steady, parallel flow of air in the test section, and instrumentation to measure the desired information. The tunnel may be open-circuit or return-flow (**Fig. 1**). The test section may be closed (with walls), or may be an open jet. When air-burning engines are to be tested, an open-circuit tunnel is appropriate. Most tunnels use a return duct, primarily to ensure smooth flow. The shape of the tunnel, continuously expanding from test section through the fan and around to the settling chamber, is chosen to reduce friction losses in the return circuit and to provide the powerful smoothing effect of a large contracting section just upstream of the test section. Except for special-purpose tunnels, most low-speed tunnels have a closed test section. The propulsion device (fan) causes an increase of pressure to overcome the friction losses in the tunnel circuit. Low-speed tunnels of convenient working-size test section (8 × 12 ft or 2.5 × 3.6 m) can test models at adequate Reynolds numbers with speeds up to 250 mi/h (400 km/h).

In a well-designed low-speed tunnel having a closed return, the kinetic energy of the air flowing

through the test section per second may be 6 to 10 times the rate at which energy is being supplied by the fan. This energy ratio (ER) may be used to estimate the power required to drive a tunnel by using the formula given below, where A_{ts} is test-section

$$\text{Kilowatts required} = \frac{13.13}{(ER)} A_{ts} \left(\frac{V}{100}\right)^3 \sigma$$

cross-section area in square meters, V is air velocity in kilometers per hour, and σ is the ratio of the air density to the density at sea level. The energy ratio may be as low as 1 or 2 for an open-circuit tunnel and is typically about 6 to 8 for most low-speed tunnels.

Lifting models experience interference in the tunnel because the walls restrict the downward flow behind the wing. The measured lift would only be achieved in free air at a larger angle of attack, and so a correction (typically less than 10%) is added to the angle of attack. The measured drag is reduced because of the angle-of-attack interference, and a correction which is a product of the lift and the angle correction is added to the measured drag. When the model has a tail, it experiences a different angle interference from that of the wing, and so a correction to the pitching moment is necessary. The drag and pitching moment corrections are of significant magnitude, but are well understood and are used confidently.

Since the model scale is usually smaller than the full-scale airplane, the Reynolds number is not matched and some scale effects appear in the data. Typically, the maximum lift of a wing is too small and the minimum drag is too large. The error can be substantial and is not subject to simple analysis. A large body of experience with these effects permits the engineer to estimate the effects of Reynolds number mismatch.

V/STOL wind tunnel. This type is a newer development of low-speed wind tunnels having a large very-low-speed section to permit testing of aircraft designed for vertical or short takeoff and landing (V/STOL) while operating in the region between vertical flight and cruising flight. In very slow flight, the flow is characterized by large downwash angles from the lifting system. The lift is usually developed by the use of power applied to a rotor, fan, jet, or jet flap. Because the rotating parts or the jet flows operate at high speeds, the flight Mach number becomes the important parameter to match. This requires that tunnel airspeeds equal flight speeds (from 20 to 100 mi/h or 32 to 160 km/h). Large test sections are required to allow the power-driven downwash to leave the model without too much distortion due to the proximity of the tunnel walls.

For economy reasons, most V/STOL tunnels are built with a standard low-speed test section (up to 300 mi/h or 480 km/h) in tandem downstream of the V/STOL test section (up to 100 mi/h or 160 km/h). This two-test-section wind tunnel provides a wide range of test conditions and has advantages in speed control at very low speeds (**Fig. 2**).

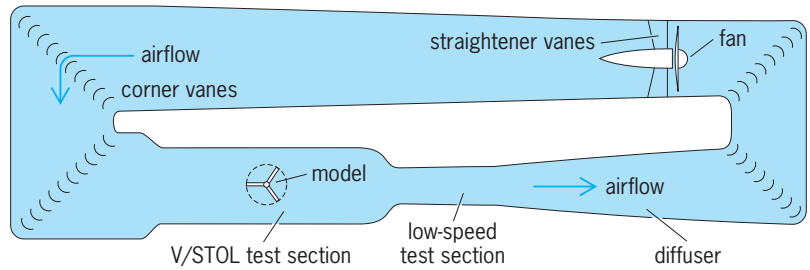

Fig. 2. Combined V/STOL and low-speed wind tunnel with two test sections.

Instrumentation and test methods are similar to those of the low-speed tunnel, but models are usually more complicated because of the power systems required. Wall interference problems are yielding to research, but corrections are still difficult. Research has continued in this area. *See* SHORT TAKEOFF AND LANDING (STOL); VERTICAL TAKEOFF AND LANDING (VTOL).

Transonic tunnel. This type is a high-speed tunnel, capable of testing at speeds near the speed of sound, at Mach numbers from 0.7 to 1.4. In compressible fluid flow in a closed duct, sonic speed can occur only where the cross section is a minimum. Therefore, sonic speed occurs where the model is located. Additional power only causes areas of supersonic flow downstream. *See* COMPRESSIBLE FLOW.

To achieve transonic flow, the test section is surrounded by a large plenum chamber and the test-section walls are vented to this chamber by means of perforations or slots parallel to the airstream. The function of the vents is as follows. Shock waves are formed in the flow near the model when sonic speed is reached. Shock waves are reflected from a solid boundary but with a change of sign (that is, expansion waves instead of pressure waves) from a free boundary. By providing sufficient free surfaces in the form of perforations or slots, reflections which would alter the flow around the model are greatly reduced or canceled. Ventilation has proved to be useful at Mach numbers from 0.9 to 1.2, where reflections are particularly troublesome. The raw data from ventilated test sections are usually used without correction except for scale effect. *See* SHOCK WAVE.

The energy ratio for a transonic tunnel may be as large as 10 at subsonic speeds, and drop to 3 at transonic speed, leading to a power requirement of up to 5000 kW/m^2 of test section.

Supersonic wind tunnel. This tunnel is capable of test speeds corresponding to Mach numbers from 1.0 to 5. For supersonic speeds, the test section must be designed for the particular Mach number desired. When air accelerates past the speed of sound, it expands more rapidly than it accelerates, so that the tunnel must be larger downstream than it is at the minimum section, where the speed is sonic. The final Mach number is uniquely determined by the ratio of the final area to the throat area (**Fig. 3**)

Continuous-flow supersonic tunnels may require 20,000–50,000 kW/m^2 of test section. Most industrial tunnels operate intermittently from energy stored in high-pressure air tanks. The air is discharged

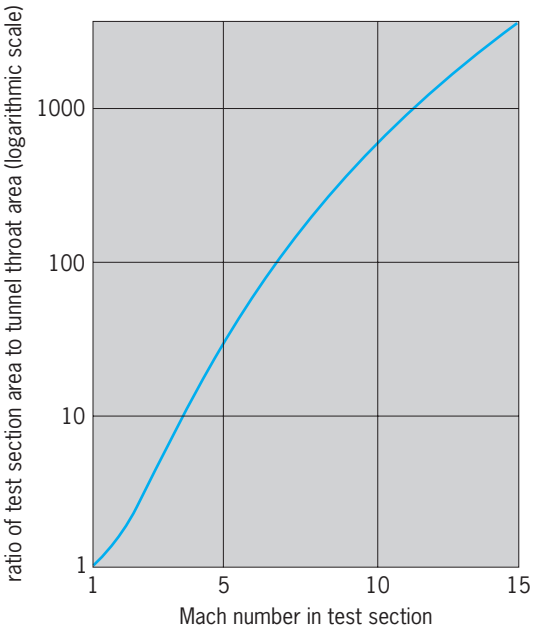

Fig. 3. Ratio of test section areas to throat area for supersonic velocity in a wind tunnel.

through a fast-acting regulator valve to the tunnel and exhausted to the atmosphere (**Fig. 4***a*).

The successful development of rapid recording instrumentation has made short run times possible.

The indraft, or vacuum-driven, tunnel allows atmospheric-pressure air to flow through the wind tunnel into a vacuum tank (Fig. 4*b*). Advantages of the indraft system include a constant (atmospheric) stagnation pressure and temperature, greater safety, and less noise and cost. Advantages of the blowdown type include higher Reynolds numbers and the ability to vary the Reynolds number over a wide range by varying the stagnation pressure.

The principal problems in the design of a supersonic tunnel are to provide sufficient pressure ratio to start and sustain the flow, and to supply adequately dry air. Drying the air to a dew point of about −40°F (−40°C) at atmospheric pressure may be required to avoid condensation shocks and uneven flow.

No wall interference effects are felt in supersonic flow, provided reflecting shock waves do not strike the model. Adjustment of results for scale effects (Reynolds number) must usually be made. *See* SUPERSONIC FLIGHT.

Hypersonic wind tunnel. This type is a supersonic tunnel operating in the range of Mach numbers from 5 to 15. When the Mach number is above 4 or 5, extremely high pressure ratios are required and the temperature drop in the test section is so large that liquefaction of the air would result unless the air were first heated (**Fig. 5**). Most hypersonic tunnels are intermittent types and use both high-pressure stagnation air and a vacuum-discharge tank to obtain the overall pressure ratio required (**Fig. 6**).

Heat is added by using electrical or gas-fired heaters or by passing the air through a bed of ceramic pebbles which have been previously heated. The nozzle may require cooling. The test section usually operates at a very low density, which means low Reynolds number and thick boundary layers.

Data from hypersonic wind tunnels are used directly but are subject to serious Reynolds number effects because of the low density of the flow. *See* HYPERSONIC FLIGHT.

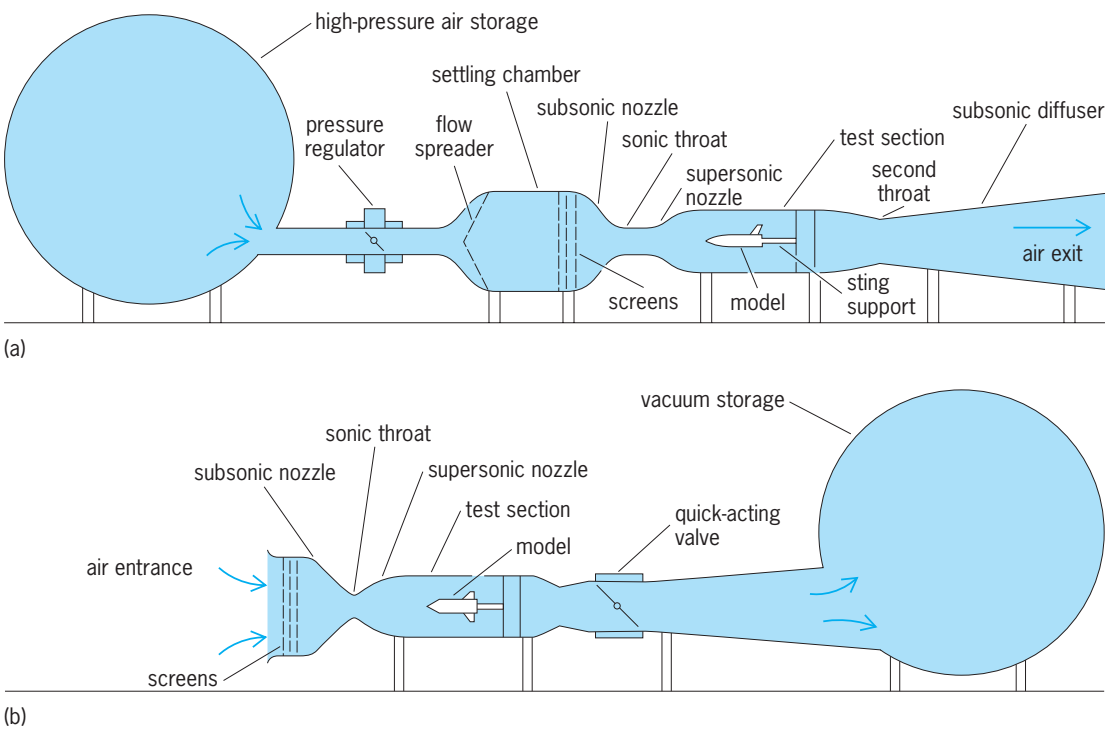

Fig. 4. Elements of (*a*) intermittent supersonic blowdown wind tunnel and (*b*) intermittent indraft supersonic wind tunnel.

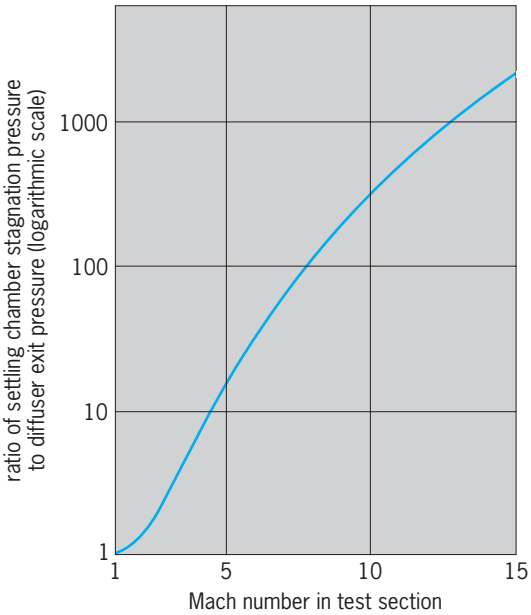

Fig. 5. Pressure ratio required to start and sustain supersonic airflow. These theoretical values assume normal shock wave at the end of the test section.

Nonaeronautical wind tunnels. Testing for the effects of wind on buildings, automobiles, bridges and other nonflying objects often requires wind tunnels of special sizes and shapes.

Meteorological wind tunnels have test sections whose length is 10 to 15 times as long as they are high (**Fig. 7**). In order to develop a thick boundary layer to match the natural wind, roughness elements such as cubes or tall spirelike structures may be located on the tunnel floor upstream of the model. The roof may be adjustable to permit matching of natural zero pressure gradient. These tunnels may be used to measure the forces on buildings, and how the flow around them affects the immediate neighborhood. Dispersal of smoke and other pollutants from factories may be studied. Snowdrift development, wind power generation, and dynamic testing of elastically scaled buildings or other structures are also done in meteorological wind tunnels.

Automobile wind tunnels are often very similar to aeronautical tunnels when force data are needed. For cooling tests, the tunnel may provide a jet of air not much larger than the car itself.

Special tunnels have been built for testing the dynamic character of suspension bridge models in which the test section was 50 times as wide as it was high.

Use of computers. Use of computers is rapidly changing the way in which wind tunnels are operated. The calculations involved in the conversion of raw data, in the form of forces and pressures, to final coefficient form are relatively simple but voluminous. More of this work, called data reduction, has been done by computers as their speed and memory capacity has increased. Advances in computer capability have led to changes from batch processing at the end of each day, to processing at the end of each run, to on-line data reduction on a point-by-point basis. Coefficients are plotted on graphic terminals while a test is in process. The test engineer can observe the progress of the test and can recall previous tests from storage for comparison. *See* COMPUTER GRAPHICS.

The computer is also used to operate the tunnel, automatically following a preprogrammed plan of test. The computer can control the tunnel speed, cause the required changes of pitch and yaw angles, and turn on the data system when specified test conditions are established.

The computer has made possible more detailed analysis of flow in wind tunnels, and has made adaptive wall tunnels possible. When the flow characteristics of a given model in free air can be calculated at a distance corresponding to the tunnel walls, the walls may be adjusted to match the free air streamlines. When this is done, the interference of the tunnel walls vanishes and no correction to the data is needed. The adjustment of the walls may be actual physical motion or virtual movement produced by injecting or extracting air through slots or holes in the walls. The techniques may be used for V/STOL or for transonic testing where corrections may be large and uncertain. Such tunnels may be called smart wind tunnels or adaptive tunnels. *See* DIGITAL COMPUTER.

Robert G. Joppa

Hypervelocity wind tunnels. Hypervelocity wind tunnels are tunnels possessing speed and temperature capabilities greater than those of hypersonic tunnels. They encompass the velocity range greater than 5000 ft/s (1500 m/s at Mach numbers 1 up to about 25) and are necessary for conducting aerodynamic research on vehicles such as ballistic-missile and satellite reentry bodies. *See* ATMOSPHERIC ENTRY.

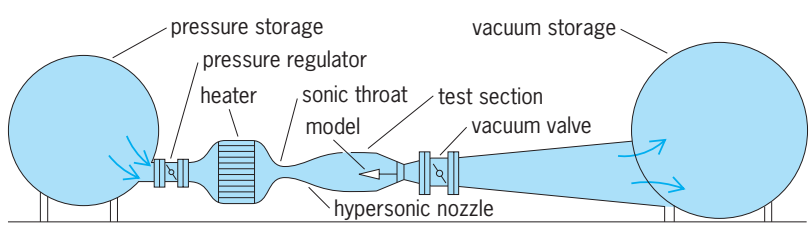

Fig. 6. Hypersonic wind tunnel with both pressure blowdown and vacuum indraft.

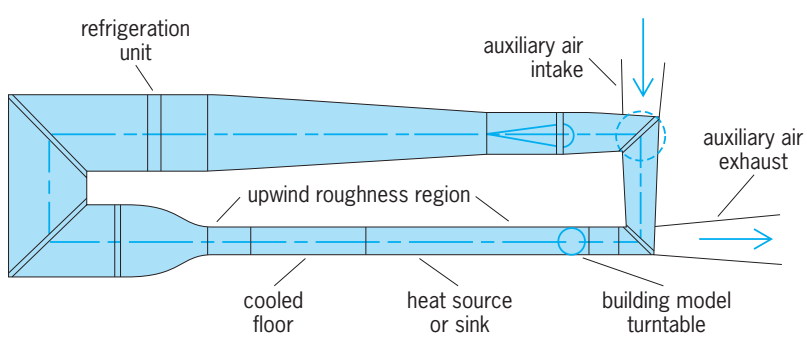

Fig. 7. Meteorological wind tunnel. (*After W. H. Rae, Jr., and A. Pope, Low-Speed Wind Tunnel Testing, 2d ed., Wiley-Interscience, 1984*)

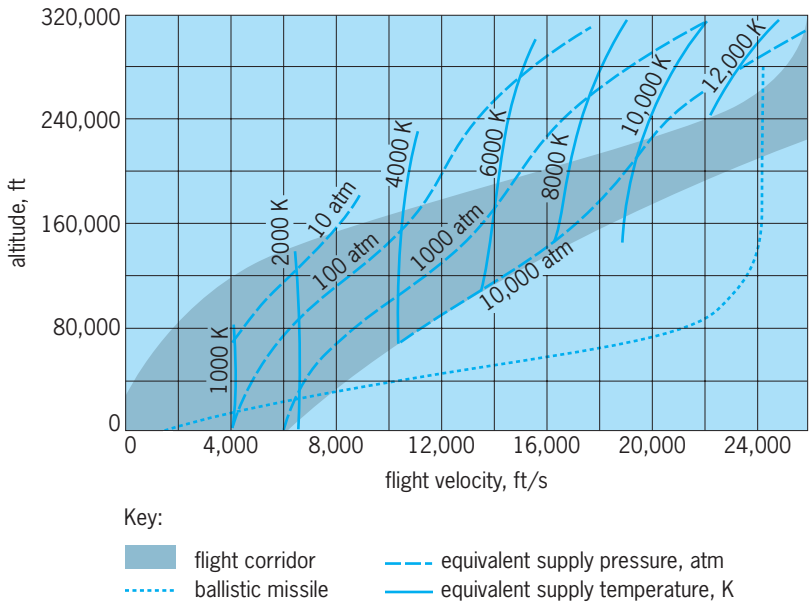

Fig. 8. Wind tunnel simulation of hypervelocity flight requires extremely high supply (stagnation) temperatures and pressures. 1 ft. = 0.3 m; 1 ft/s = 0.3 m/s; 1 atm = 10^2 kPa = 14.7 lb/in.^{2}; °F = (K × 1.8) − 460.

Wind-tunnel simulation of hypervelocity flight conditions requires extremely high supply (stagnation) temperatures and pressures, as may be seen in **Fig. 8**. These high supply temperatures are required for expanding the tunnel airflow through the nozzle to hypervelocities in the test section; similarly, high supply pressures are necessary to duplicate flight-altitude ambient pressure in the test section. Hypervelocity tunnels operate at considerably higher supply temperatures and pressures than do hypersonic tunnels.

The "flight corridor" indicated in Fig. 8 represents an approximate altitude-velocity band wherein flight of vehicles supported by aerodynamic lift and centrifugal force can be sustained. The upper boundary of the flight corridor is fixed by the minimum allowable atmospheric density; the lower boundary is fixed by the maximum allowable aerodynamic heating rate. This corridor encompasses the flight spectrum of the proposed hypersonic transport or boost-glide vehicles. A typical trajectory of a ballistic missile from its entry into the Earth's atmosphere is shown in Fig. 8. The missile is able to tolerate the higher heating rates encountered below the corridor of continuous flight because of its short flight duration.

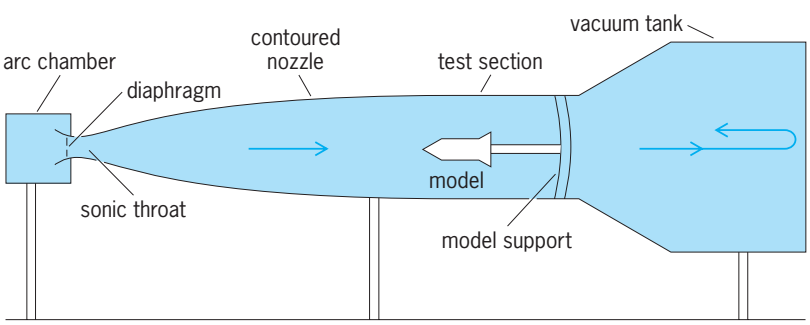

Fig. 9. Hypervelocity hotshot wind tunnel.

Complete duplication of all flight conditions in a hypervelocity tunnel is extremely difficult, if not impossible. For example, complete simulation of flight at 18,000 ft/s (5500 m/s) at 150,000-ft (45,000-m) altitude would require a tunnel supply temperature of 9000 K (16,000°F) and a supply pressure of 30,000 atm (450,000 lb/in.^{2} absolute or 3 gigapascals). Operation at such a high temperature is difficult, and operation at such a high pressure is not feasible at present. However, increasing the flight altitude to 300,000 ft (90,000 m) at the same velocity reduces the supply pressure required to a value that allows duplication to be achieved. Interpretation of test results is difficult, however, because the test gas is not in equilibrium owing to its rapid expansion from a highly ionized and dissociated state in the reservoir.

The major problem in designing a hypervelocity tunnel is the prevention of structural damage due to heating. Even though testing times may last only a fraction of a second, the high supply temperature (combined with the high pressure) results in a large heat flux, which reaches a maximum at the sonic throat. Damage at this point can result in large flow distortions, as may be inferred from the following example: A 0.025-in.-diameter (0.64-mm) sonic throat is required to establish a Mach 25 flow in a 115-cm-diameter (45-in.) test section. A change in throat diameter of only 0.002 in. (0.05 mm) would change the test-section Mach number by approximately 1.0.

The major types of currently operating hypervelocity tunnels are described below.

Hotshot tunnel. Mach numbers of 10–27 with testing times of 10–100 milliseconds can be obtained with this tunnel. The principle of operation (**Fig. 9**) is to prepressurize the arc chamber to 500–10,000 lb/in.^{2} abs (3.4–70 MPa) and evacuate the rest of the tunnel to a pressure of about 1 micrometer of mercury (approximately 10^{-6} atm or 0.1 pascal). The two chambers are separated by a plastic or a metal diaphragm. Electrical energy (capacitance or inductance storage of up to 10^8 joules) is then discharged into the arc chamber. This energy increases the air temperature and pressure, resulting in rupture of the diaphragm. The heated gas is then accelerated in the nozzle to provide hypervelocity flow. The test is usually terminated by a quick-acting dump valve which releases the stagnation chamber gas through a large port rather than requiring all the hot gases to pass through the sonic throat. This minimizes damage to both the throat and the interior of the arc chamber. The time at which this dump valve is actuated (that is, the effective run time) is influenced by a number of factors: (1) the run time necessary to obtain the desired data; (2) the time at which flow will "break down" at the model due to insufficient pressure ratio (resulting from either too large a decay in stagnation pressure or too large a buildup in vacuum tank backpressure); or (3) the maximum time the arc chamber or sonic throat can withstand the high heat-transfer rates.

Hotshot tunnels have been designed to operate at arc-chamber (supply) pressures and temperatures as high as 100,000 lb/in.^{2} abs (700 MPa) and 10,000 K (18,000°F), respectively. However, melting or oxidation of the tungsten throats (diameters of 0.020–0.40 in. or 0.5–10 mm) has, in general, limited useful arc-chamber pressures and temperatures to approximately 30,000 lb/in.^{2} abs (200 MPa) and 4000 K (7000°F). Nitrogen, rather than air, is used as a test medium in some tunnels to avoid oxidation of the tungsten throats. Other problems are the flow contamination from solid particles (throat or electrode melting) or foreign gases (pyrolysis of plastic diaphragms and insulators) and depletion of the test gas oxygen (when air is used).

Hotshot tunnels are used for heat transfer, force, dynamic stability, and pressure model tests; data are recorded on short-response-time oscillographs or on magnetic tape. Schlieren or shadowgraph photographs and high-speed motion pictures of the luminous airflow about the model provide data on flow formation, shock waves, and flow separation.

Shock tunnel. Test Mach numbers of 6 to 25 (velocities of 4000–15,000 ft/s or 1200–4500 m/s) with testing times from 0.5 to 4 ms can be obtained in this type of facility. The major sections of a shock tunnel (**Fig. 10**) consist of the driver, a constant-area tube or driven section, and a nozzle and dump tank. A high-velocity shock wave is produced by the rupture of the diaphragm separating the high-pressure driver section, containing helium, hydrogen, or a combustible mixture, from the low-pressure driven section, containing air or nitrogen. As the shock advances, it heats and accelerates the working gas behind it in the driven tube, and a region of uniform conditions is created between the shock wave and the gas interface (the contact surface separating the driver gas and the driven gas). The shock is reflected at the nozzle throat and passes back up the driven section, and the high-pressure high-temperature driven gas ruptures the throat diaphragm and accelerates to a high Mach number in the nozzle.

In a conventional shock tunnel, the reflected shock is again reflected at the gas interface and returns downstream; the run ends when the shock arrives at the nozzle throat for the second time. Run times of 4 ms are obtained by use of the "tailored interface" technique, which consists of fixing the properties of the gas on both sides of the interface so that the primary reflected shock wave produces equal pressures and velocities on each side of the interface. When this condition is met, the shock passes through the interface without reflection.

Working-gas (supply) pressures are normally in the range of 6000–20,000 lb/in.^{2} abs (40–140 MPa), while gas (supply) temperatures up to 10,000 K (17,5000°F) are achievable. Initial pressure in the vacuum dump tank is about 1–5 μm of mercury (0.1–0.7–0.7 Pa).

The main advantage of the shock tunnel over the hotshot tunnel is that the test medium is less contaminated with metallic particles, although some difficulty has been experienced with driver-gas contamination due to mixing at the interface. The main disadvantage of the shock tunnel is the shorter run time.

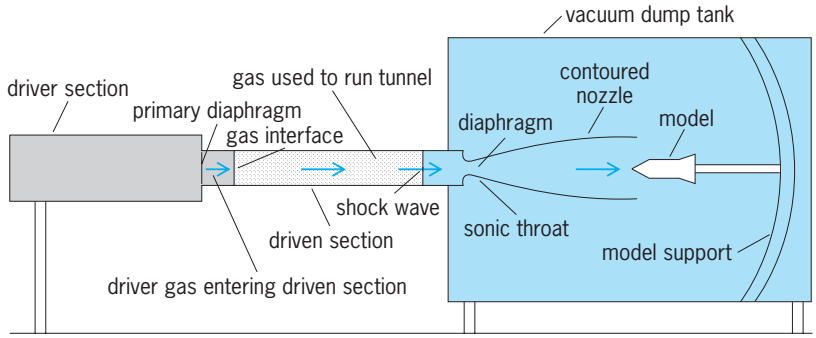

Fig. 10. Hypervelocity shock wind tunnel.

Instrumentation with response times of less than a millisecond has been developed for measurement of heat transfer, pressure distribution, and aerodynamic forces on models. A particular problem in measuring aerodynamic forces on models arises from the necessity of making the measurements while the model is still vibrating from the impact absorbed when the starting shock passed over it. This vibration produces inertial forces which must be measured with accelerometers and accounted for in the force balance data reduction. *See* SHOCK TUBE.

Graphite heater blowdown tunnel. This type of tunnel has not yet earned a nickname, so it is labeled in a very generic way. Test Mach numbers up to 14 with testing times from 0.25 to 2.5 s can be obtained with this relatively new design of hypervelocity tunnel. A Mach 18 facility has been designed but not yet built. In operation, nitrogen at room temperature is used to fill the heater (**Fig. 11**) to a pressure on the order of one-fourth of the desired test pressure. Electrical power (up to a megawatt) is then applied to a graphite heater element which heats the gas at constant volume, increasing its pressure to the desired stagnation pressure. The nitrogen is confined in the heater by a double diaphragm which is ruptured by overpressuring the interdiaphragm volume to begin the run. The decay in stagnation pressure associated with the hotshot tunnel is eliminated by introduction of controlled pressure cold gas into the bottom of the heater to push the hot test gas up, in a pistonlike fashion, out of the heater. The time over

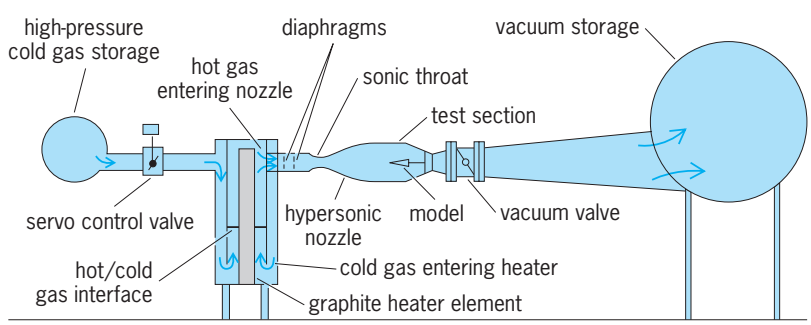

Fig. 11. Hypervelocity graphite heater blowdown tunnel.

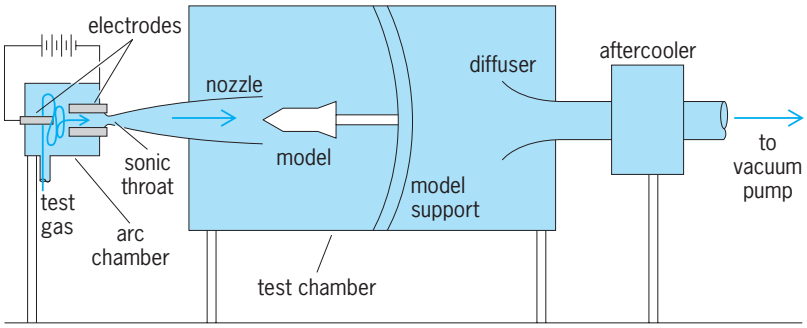

Fig. 12. Hypervelocity plasma-jet wind tunnel using dc power.

which valid data may be taken is terminated by arrival of the hot/cold gas interface at the nozzle. Control of the cold nitrogen pressure is maintained by very fast-acting electrohydraulic servo-driven control valves. Maximum supply pressures of 20,000 lb/in.^{2} abs (140 MPa) and supply temperatures of 3200°R (1800 K or 2750°F) have been obtained in this type of facility.

The main advantage of this type of tunnel over the shock tunnel is the relatively long run times. During the test time available in this facility, it is possible to make a complete high speed angle-of-attack sweep from 0 to 20 or 25°, obtaining heat transfer, pressure, or aerodynamic force data over that entire angle-of-attack range. In either a shock tunnel or a hotshot, the model is preset to given angle of attack, and data are obtained only at that condition during the run. Its main disadvantage lies in its complexity and resultant high cost.

Plasma-jet tunnel. This tunnel has the capability of developing the highest temperature (approximately 20,000 K or 35,000°F) and the longest run time (several minutes) of any hypervelocity tunnel (**Fig. 12**). It is arc-heated and utilizes either dc or ac power. The arc heater consists of a water-cooled cathode, an injection (or vortex) chamber, and a water-cooled hollow anode.

The function of the injection chamber is the generation of a strong vortex to stabilize the arc in the center of the heater. The vortex is created by the injection of high-pressure air (or other test gases) in a tangential direction at the outer edge of the vortex chamber. After injection, the test gas flows to the center of the injection chamber, through the hollow anode, and out the nozzle throat. Since the angular momentum of the test gas will be conserved as it flows to the center of the chamber, its tangential velocity will increase and its density will decrease. This effect creates a pronounced radial density gradient in the vortex chamber and hollow anode, causing the arc to remain along the center line, where the density is lowest.

A magnetic field is used to rotate the arc so that the arc-attachment points at the cathode and anode do not become overheated. Further cooling of the anode is provided by the flow of test gas, since the swirling motion of the flow tends to keep cooler gas along the anode surface. A secondary benefit of this phenomenon is an increased heater efficiency because of reduced heat transfer to the anode and nozzle throat. Overall efficiencies as high as 60–70% can be obtained with a gas-stabilized arc heater.

The hypervelocity facility is used primarily for heat-transfer tests, reentry-shape ablation studies, and magnetoaerodynamic studies. Relatively high velocities (up to 20,000 ft/s or 6000 m/s) are achievable, but few plasma tunnels have been proposed or built for obtaining Mach numbers greater than 10. Maximum operating supply pressure is about 3500 lb/in.^{2} abs (24 MPa). At high pressures the power requirements are greatly increased and arc stabilization is difficult. Plasma-jet units with power requirements of 1–60 MW are in operation or under construction; units with power requirements of several hundred megawatts have been proposed.

Disadvantages are electrode contamination and large nonuniformities in the test-gas stream, high power requirements, and operating pressure limitations. Instrumentation consists of high-speed photographic equipment, calorimeter, pitot and mass flow probes, microwave and electron beam apparatus, and a spectrograph.

Other hypervelocity wind tunnels are the gun tunnel, shock tube, expansion tube, and wave superheater tunnel. Randall C. Maydew; Donald D. McBride

Instrumentation

Specifically designed mechanical, electrical, and optical devices are used to measure effects of air flow across a model. Wind tunnel testing has always required instrumentation that is rugged, reliable, and accurate. Increased costs for energy and worker-power has brought demands for shorter run time and higher productivity. To fulfill these needs, modern instrumentation has been developed which can acquire accurate data quickly and provide information directly to the research engineer during the test. Measurements common to most tunnels are pressure, temperature, turbulence, and flow direction.

Pressure measurement. The earliest method used to measure a large number of pressures in tunnel tests was with multitube liquid manometer boards. A photograph of the board was taken at each special test configuration. Later, the height of each column was measured on the photo to provide the data.

Many techniques have since been developed to provide direct conversion of pressure to electrical signals. Individual strain gage, capacitance, or force balance transducers may be used, but when several hundred pressures are read this method is expensive and requires much set-up time. The mechanically scanned multiport sampling valve which switches, in sequence, 48 pressure tubes to a single transducer is often used. Since, for an accurate reading, time must be allowed for the pressure to settle after it is switched, the maximum reading rate is about 20 pressures per second.

An electronically scanned pressure (ESP) measuring system allows pressure sensors to be sampled at 10 kHz. The heart of the system is the electronically scanned pressure module (**Fig. 13**) which contains 32 silicon pressure transducers, an electronic

multiplexer, and a calibration valve in a volume less than 4.2 in.^{3} (69 cm^3). The modules may be mounted within a tunnel model to avoid the slow repsonse caused by long pneumatic tube lengths. Because all transducers are periodically calibrated in place against an accurate pressure standard, the system accuracy is within $\pm$ 0.15% of full scale.

Some types of testing, such as engine inlet dynamics, require measurement of pressure fluctuation up to 20 kHz. Rakes containing as many as forty 0.08-in.-diameter (2-mm) semiconductor strain gage transducers are used for this purpose. *See* PRESSURE MEASUREMENT.

Temperature measurement. Most temperature measurements in tunnel tests are made with thermocouples, thermistors, and platinum resistance temperature sensors. The airstream temperature measured is the stagnation or total temperature. The static temperature is computed from this measurement.

Heat transfer measurements may be made by constructing a thin-wall tunnel model with thermocouples closely spaced on the inside surface of the wall. Individual heat transfer gages can be built by using a thin copper disk with a fine thermocouple attached, mounted in an insulated holder. The gage is installed with the sensing disk flush with the outside model surface. A heat transfer gage used in shock tunnels is constructed as a thin-film resistance thermometer. Using a 0.1-μm-thick platinum film deposited on a Pyrex glass substrate, a response time of 10^{-7} s is obtained. *See* TEMPERATURE MEASUREMENT.

Turbulence measurement. The level of velocity fluctuations is important in tunnel testing because it influences the point on a model at which the boundary layer changes from laminar to turbulent. This point of transition affects the aerodynamic drag forces on the model. Two instruments used to measure turbulence are the thermal anemometer and the laser velocimeter. *See* BOUNDARY-LAYER FLOW; TURBULENT FLOW.

The thermal anemometer is used to obtain instantaneous velocity and flow angle measurements. The sensor contains a fine wire (**Fig. 14***a*) or a conducting film on a ceramic substrate (Fig. 14*b*) which is electrically heated to a temperature greater than the surrounding air. Since the amount of heat conducted away from the sensor depends on the local air velocity, its temperature will change accordingly. The resistance of the sensor varies with its temperature, which allows its use in a Wheatstone bridge circuit to provide a signal related to flow. *See* ANEMOMETER.

The hot wire or film by itself cannot respond to changes in fluid velocity at frequencies above 500 Hz. With electronic compensation, this response can be increased to over 1 MHz. With the constant temperature circuit, shown in **Fig. 15**, a servoamplifier is

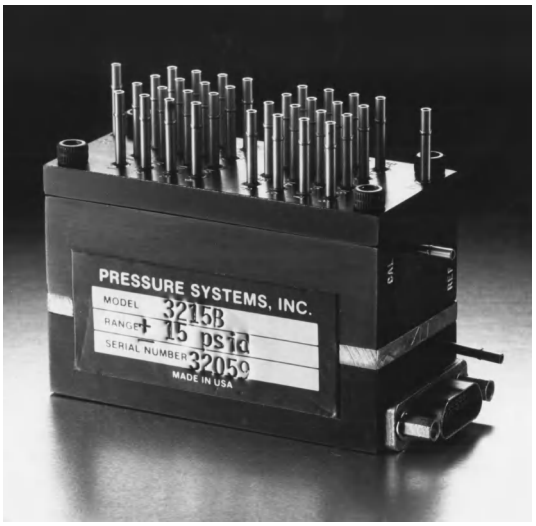

Fig. 13. A 32-transducer electronically scanned pressure module. (*Pressure Systems, Inc.*)

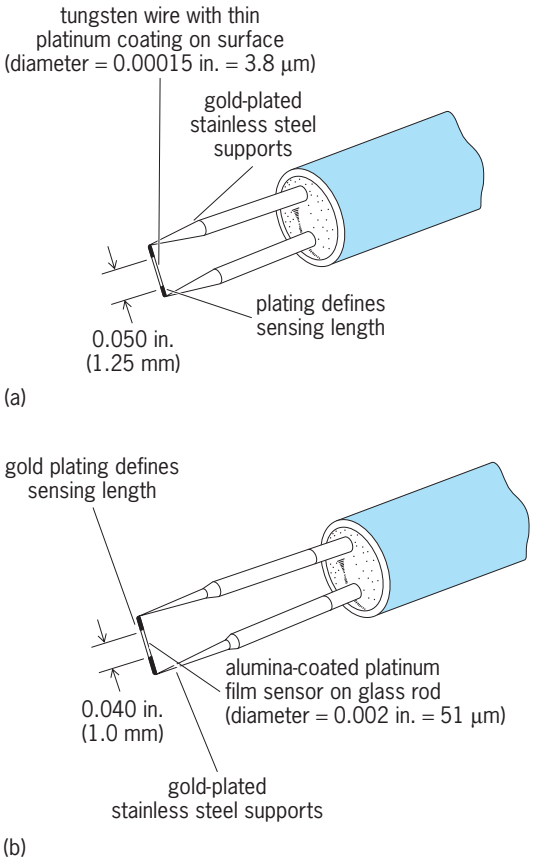

Fig. 14. Thermal anemometers with (*a*) hot-wire sensor and (*b*) cyclindrical hot-film sensor. (*TSI Inc.*)

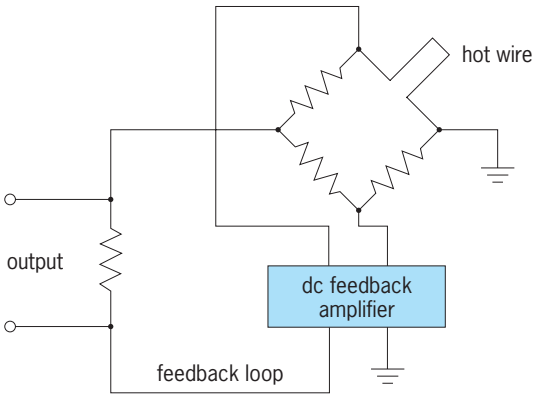

Fig. 15. Constant-temperature thermal anemometer circuit.

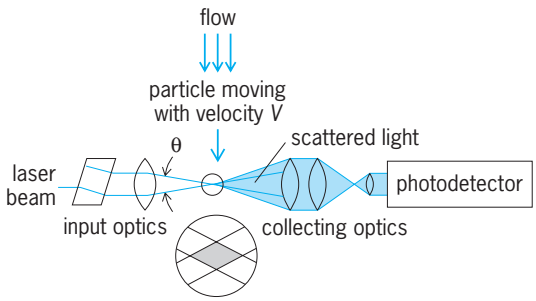

Fig. 16. Configuration of laser velocimeter.

used which senses an unbalance in the bridge due to a change in sensor resistance and feeds back a current to the bridge to restore balance. In this way the wire is held at a constant instantaneous resistance, and thus a constant temperature. Therefore, no thermal lag occurs, and the feedback current is a direct measure of the turbulence.

The laser velocimeter (LV) is an optical technique that provides remote measurement of mean velocity and turbulence in the flow field of a tunnel model. Since no mechanical probe is used, there is no disturbance to the flow. As shown in **Fig. 16**, a laser beam is optically split into two equal-intensity beams and focused to a point in the flow field. At this point, the light waves from each beam interfere constructively and destructively, forming a set of interference fringes. Extremely small aerosol particles (about 1-μm diameter) are injected upstream of the focal point. As each of these particles follows the airstream through the bright and dark fringes, the light that it scatters will vary in intensity. The scattered light is imaged onto a photodetector which produces an electrical signal proportional to the light intensity. The component of velocity normal to the optical axis in the plane of the two input laser beams can be determined from the signal frequency with the equation below, where V is the component velocity

$$V = \frac{\lambda f}{2 \sin\left(\frac{\theta}{2}\right)}$$

(m/s), f is the signal frequency (Hz), λ is the wavelength of the laser beams (m), and θ is the angle between the two laser beams. *See* LASER.

By using a two-color laser beam and additional optics, velocities in two directions can be measured simultaneously. A two-color laser velocimeter with blue and green beams is shown in **Fig. 17**.

A wide variety of flow visualization techniques are available for observing turbulent flow. Tufts attached to the model show flow patterns near the surface. China clay, lampblack, or fluorescent dyes supsended in viscous oil are used for studying flow in the boundary layer. Sublimation of fluorene or azobenzene is used to study extremely thin boundary layers. Schlieren photos allow the flow to be visualized in supersonic airstreams.

A method developed to visualize vortex patterns is called the laser sheet. A laser beam is fanned out into a thin vertical sheet and projected normal to the tunnel axis. A fine mist injected into the tunnel scatters the projected light and forms a visible screen. Any disturbance of the uniform flow field, such as that produced by vortices, appears as dark holes within the vapor screen.

Direction probes. The flow direction and the degree to which the flow is parallel to the tunnel walls is vital in the analysis of pressure and forces on a test model in a wind tunnel. Several devices are used, depending on the flow velocity (**Fig. 18**).

Yaw sphere. In subsonic flow the pressure distribution on the surface of a sphere can be used to

Fig. 17. Laser velocimeter used for turboprop flowfield analysis in wind tunnel at NASA Lewis Research Center.

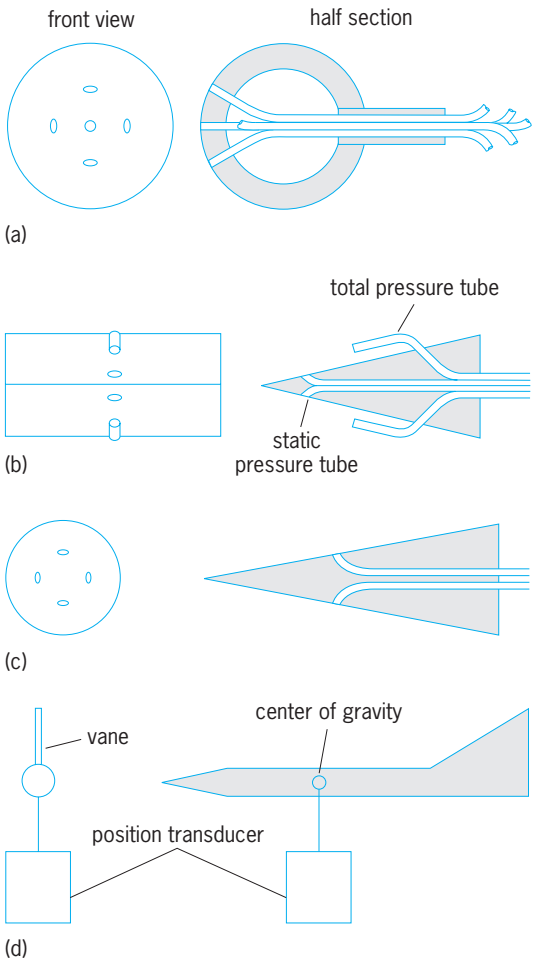

Fig. 18. Direction-measuring instruments. (*a*) Yaw sphere. (*b*) Wedge. (*c*) Cone. (*d*) Vane.

determine the flow direction of the air passing over the sphere. In a typical probe, yaw angle and pitch angle can be determined simultaneously. The ratio of the difference between opposite static pressures to the average of these pressures should be calibrated against the flow angles before the probe is used in a tunnel.

Wedge. In supersonic flow the sphere is not usable; instead, the wedge serves in its classical use as a means to determine the velocity of a supersonic stream and as an angle-of-flow meter. The wedge is a two-dimensional device, so that pitch and yaw direction measurements must be made separately. When the wedge is used only for directional measurement, the total-pressure tubes are not needed if the Mach number is known. The angle of flow can be obtained from tabulated data for wedges in supersonic flow. The wedge with static-pressure tubes can be used only to indicate subsonic flow angles, but first must be calibrated.

Cone. Similar to the wedge, the cone is usable in both supersonic and high subsonic flow fields to determine flow angles. The cone can be used to measure yaw and pitch angles simultaneously. However, like the sphere, it must be calibrated.

Vane. When space is not restricting, the mechanical vane can give direct-reading angles of yaw or pitch without detailed computations. The vane is mounted at its center of gravity on a calibrated position transducer and is read remotely. Daniel J. Shramo

Bibliography. J. D. Anderson, Jr., *Hypersonic and High Temperature Gas Dynamics*, rev. ed., 2000; J. B. Barlow, W. H. Rae, Jr., and A. Pope, *Low-Speed Wind Tunnel Testing*, 3d ed., 1999; F. Durst, A. Melling, and J. H. Whitelaw, *Principles and Practices of Laser-Doppler Anemometry*, 2d ed., 1981; N. Isyumov (ed.), *Wind Tunnel Studies of Buildings and Structures*, rev. ed., 1998; A. Pope and K. L. Goin, *High-Speed Wind Tunnel Testing*, 1965, reprint 1978.

Windings in electric machinery

Windings can be classified in two groups: armature windings and field windings. The armature winding is the main current-carrying winding in which the electromotive force (emf) or counter-emf of rotation is induced. The current in the armature winding is known as the armature current. The field winding produces the magnetic field in the machine. The current in the field winding is known as the field or exciting current. *See* ELECTRIC ROTATING MACHINERY; GENERATOR; MOTOR.

The location of the winding depends upon the type of machine. The armature windings of dc motors and generators are located on the rotor, since they must operate in conjunction with the commutator, and the field windings are mounted on stator field poles. *See* DIRECT-CURRENT GENERATOR; DIRECT-CURRENT MOTOR.

Alternating-current synchronous motors and generators are normally constructed with the armature winding on the stator and the field winding on the rotor. There is no clear distinction between the armature and field windings of ac induction motors or generators. One winding may carry the main current of the machine and also establish the magnetic field. It is customary to use the terms stator

Fig. 1. Salient-pole field winding on rotor of synchronous motor. (*Allis-Chalmers*)

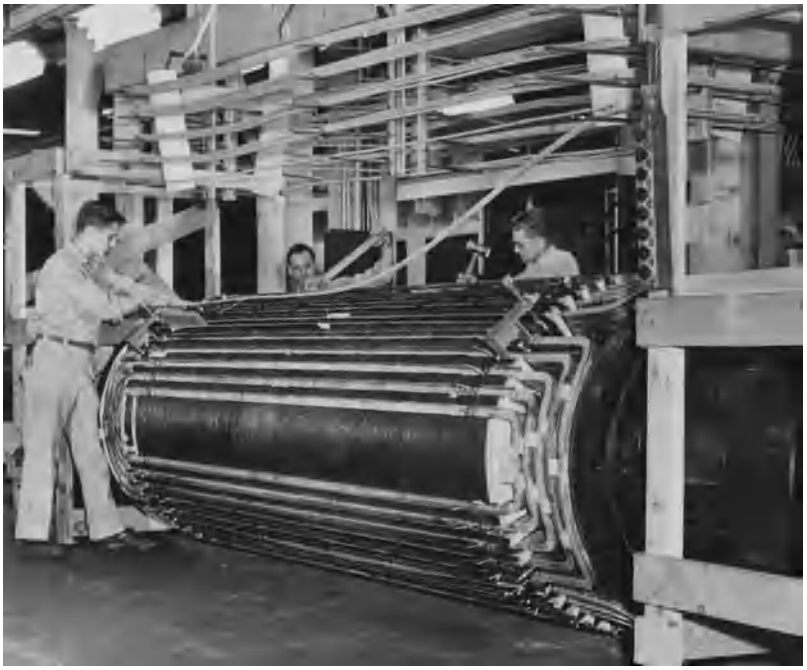

Fig. 2. Winding field on cylindrical rotor. (*National Electric Coil Co.*)

winding and rotor winding to identify induction motor windings. The word armature, when used with induction motors, applies to the winding connected to the power source (usually the stator). *See* ALTERNATING-CURRENT GENERATOR; ALTERNATING-CURRENT MOTOR; SYNCHRONOUS MOTOR.

Field windings. Field windings produce a magnetic pole fixed in space with respect to the magnetic structure on which they are mounted. If salient-pole construction is employed, the winding turns are concentrated around the pole core (**Fig. 1**). If cylindrical construction is employed, the field winding is constructed of elongated concentric loops embedded in slots cut in the surface of the field structure. In the case of dc motors and generators with compensating windings, both forms of construction may be found in one machine.

Concentrated field windings for salient-pole ac synchronous machines and for shunt fields of dc and series ac machines are form-wound of many turns of insulated copper wire. The completed coil is taped and impregnated to hold its shape and to fit around the pole core. Series field coils are normally concentrated. They are wound of larger wire, or of insulated copper strip, in order to carry the armature current without overheating.

Field coils for cylindrical construction are made of insulated rectangular copper or aluminum bars (**Fig. 2**). Individual coils are form-wound with several turns per coil. The dimensions of the coils are selected to provide a suitable space distribution of field flux. The coil is held in the slots by nonmagnetic wedges. The coil ends are held in place against mechanical forces by nonmagnetic bands.

Rotor field coils are connected to the external circuit through slip rings, on which carbon brushes rest to conduct the exciting current. *See* SLIP RINGS.

Armature windings. These windings carry ac current. They take many forms, depending upon the type and capacity of the machine and specific design requirements. Armature windings have an active portion, which lies in slots in the magnetic circuit, and end turns and end connections, which are external to the airgap. The treatment of the end connection determines the appearance and operating characteristics of the winding. Usually the end turns are formed to make diamond-shaped coils for ease in inserting and bracing the windings. A coil may have a single turn, or it may have many turns. Each turn may have a single strand, or it may have several strands of copper electrically in parallel. A coil has two coil sides, which lie in slots approximately a pole pitch apart for dc machines, whereas most ac machines have short-pitch windings in which the coil pitch is less than 180 electrical degrees. The pitch of a coil is found by taking the quotient of the slots separating the coil sides and the slots in 180 electrical degrees. Windings may be single layer, one coil side per slot, or double layer, two coil sides per slot, in which case one side of each coil will rest in the bottom half of its slot and the other coil side at the top of its slot.

Direct-current armature windings. These may be lap, or parallel, windings, or they may be wave, or series, windings. **Figure 3** shows a four-pole lap dc armature winding and a four-pole wave dc aramature winding. The two positive brushes of the lap winding would be connected together and the two negative brushes would be connected together. The lap and wave windings have the following general properties:

1. Coil ends of lap windings are fastened to adjacent commutator bars. Coil ends of wave windings are fastened to commutator bars approximately two pole pitches apart.
2. The lap winding has as many parallel paths between the line terminals as there are poles, and the conductors of a path lie under adjacent poles. The wave winding has two parallel paths between the line terminals regardless of the number of poles, and conductors of either path lie under all poles.
3. The lap winding has as many sets of brushes as there are numbers of poles, whereas wave windings need only two brush sets. Usually, however, the

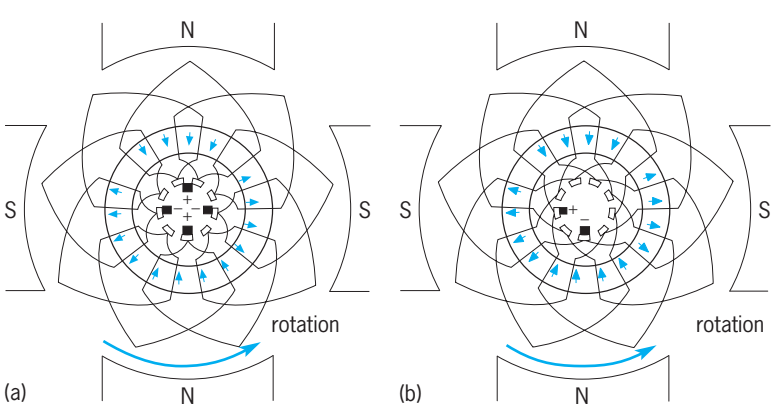

Fig. 3. Armature windings. (*a*) Simple four-pole lap winding. (*b*) Simple four-pole wave winding.

machine carries as many brush sets as there are poles to provide greater current-carrying capacity for the commutator. Brush positions are such that the coils are short-circuited only when they are in the position of minimum induced voltage.

Lap windings are adapted to high-current machines because they may have more than two parallel paths, whereas the wave windings are adapted to small-capacity machines and high-voltage machines because of the series connection of the coils.

Alternating-current armature windings. These may be single or polyphase, full or fractional pitch, Y-connected or delta-connected. Practically all except small-capacity ac machines have three-phase windings, with coils distributed around the entire armature periphery for better utilization of space and material in the machine. Winding factors are used to evaluate the performance of a winding. The generated rms voltage per phase is $E = 4.44\, k_d k_p f N_{pb} \Phi$, where k_d is the distribution factor, k_p the pitch factor, f the frequency in hertz, N_{pb} the total turns in series per phase, and Φ the flux per pole in webers per square meter.

The distribution factor k_d of a winding gives the ratio of the voltage output of a distributed winding to the voltage output of a concentrated winding. This is shown by Eq. (1), where n is number of slots over

$$k_d = \frac{\sin(n\gamma/2)}{n \sin(\gamma/2)} \tag{1}$$

which one phase is distributed and γ is electrical angle between slots. Fractional-pitch windings are employed to reduce the magnitude of harmonics, to improve wave shape, to save copper, and to permit end connections between coils to be made more easily.

The pitch factor k_p of a winding gives the ratio of the voltage output of a fractional-pitch coil to the voltage output of a full-pitch coil. This is shown by Eq. (2), where ρ is electrical angle between coil sides.

$$k_p = \cos\frac{\pi - \rho}{2} \tag{2}$$

Since the angle γ and the angle ρ in the equations for k_d and k_p change with frequency, they may be selected to eliminate or minimize certain harmonic frequencies.

Figure 4 shows the active conductors of a distributed, three-phase, two-pole, fractional-pitch, armature winding wherein a, b, and c are conductors of three phases displaced 120 electrical degrees apart. Coil a_1, $-a_1$ spans $\frac{5}{6}$ of a pole pitch, or 150 electrical degrees.

Stator and wound-rotor windings of polyphase induction motors conform to the above descriptions. Single-phase induction motor stators may be wound in the same manner, except for the number of phases; or they may be formed of several concentric loops of wire. The windings are usually form-wound from one continuous length of wire, but are not taped. This facilitates the installation of the winding in small, partly closed slots. For additional information on these and squirrel-cage windings *see* INDUCTION MOTOR

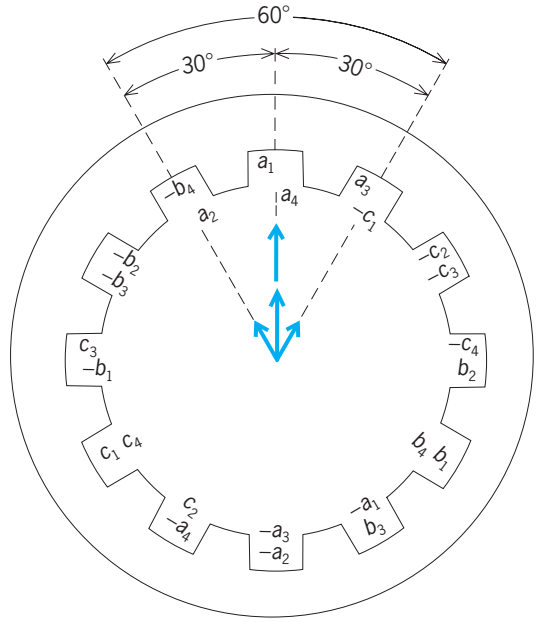

Fig. 4. Distributed, three-phase, two-pole, fractional-pitch armature winding. The voltage vector diagram is given. (*After A. E. Fitzgerald and C. Kingsley, Electric Machinery, McGraw-Hill, 1952*)

Arthur R. Eckels

Bibliography. R. T. Anderson et al., *Large Rotating Machine Windings*, 1969, reprint 1981; A. E. Fitzgerald, C. Kingsley, and S. D. Umans, *Electric Machinery*, 5th ed., 1990; M. Liwschitz-Garik and C. Gentilini, *Winding Alternating Current Machines*, 1950, reprint 1975; S. A. Nasar, *Electrical Machines and Power Systems*, 1995.

Wine

Technically, the juice of any fruit that has been subjected to an alcoholic fermentation. However, both in general usage and by legal definition the term wine is defined as a beverage produced by the alcoholic fermentation of the juice of grapes.

Grapes possess two attributes that contribute to their preeminence in wine production. The juice of fresh ripe grapes is mildly acidic (pH 3.0–3.5) and contains about 21–25% fermentable sugar (mostly glucose and fructose). These properties favor the multiplication of naturally occurring yeasts which, under anaerobic conditions, carry out an alcoholic fermentation that yields ethanol and carbon dioxide as the major end products. The final result is a beverage with sufficient ethanol concentration (usually 12.5% by volume) to render it relatively stable if it is kept in sealed containers, and sufficient in tartness and flavor to be interesting to the palate. Other fruits yield stable and flavorful fermented beverages only by considerable artifice, usually involving supplementation of their juice with fermentable sugar

and with acid. *See* ETHYL ALCOHOL; FRUCTOSE; GLUCOSE.

Grape species. Most wines are made from various cultivars of the classic European wine grape species, *Vitis vinifera*. However, in the United States east of the Mississippi, wine is also made from varieties of two native American grape species, *V. labrusca* (varieties descended from the wild "Fox Grape" still found in New England forests) and, to a lesser extent, *V. rotundifolia*. The juice of *V. labrusca* grape varieties is often too acid for production of salable wine. In order to derive a palatable wine from these grapes, their musts (that is, the juice) may be legally diluted with up to 35% by volume of water or sugar solution, prior to fermentation (without disclosure of the fact to the consumer). This process is euphemistically termed amelioration. In other countries the admixture of water or a sugar solution to musts in the production of wine is forbidden; however, musts may be legally sugared in Germany, provided the products are labeled *Verbessert*. Increasingly, *V. vinifera–V. labrusca* hybrids have been planted in many winegrowing areas of the eastern United States. Wines from the best of these crosses have little of the methyl anthranilate odor associated with wines from *V. labrusca* varieties. *See* GRAPE.

Wine types. Wines are classified on the basis of the presence or absence of carbon dioxide, on alcohol content, color, and relative sweetness.

Alcohol content. Wines whose dissolved carbon dioxide content is insufficient to produce long-lasting effervescence when served are called still wines, and are divided on the basis of their alcohol content into three general categories: table wines, fortified wines, and so-called soft or light wines. Table wines, containing 9–14% alcohol by volume, are usually used to accompany the main courses of a meal; their alcohol content derives directly from the sugar present in the original grape juice. Fortified wines have alcohol concentrations up to 20% by volume, achieved by adding brandy (spirits distilled from wine) at some stage of their production. They are sometimes known as dessert wines because such wines are often served at the end of a meal. Established as a class in the early 1980s, light wines, which contain only about 6–7% alcohol, are often prepared by removing some of the alcohol from table wines by vacuum distillation. However, some of the flavor constituents are also removed by this process, and the final product is essentially similar to a table wine which has been diluted with water. Reverse osmosis technology is used to produce other liquids legally classified as wine. In this procedure, standard wine is routed, under pressure, through a filtering chamber housing membranes permeable to substances having a nominal molecular weight of less than 500, thereby separating the standard wine into a permeate and a retentate, the latter containing reduced levels of alcohol and water and legally sold as wine.

Color. White wines (usually pale straw in color) are most often made from light-skinned grape varieties, but can also be made from dark-skinned varieties inasmuch as the juice of most wine grapes, light-skinned or dark-skinned, is unpigmented. When white wines are made from dark-skinned grapes, the juice is separated from the skins immediately after the grapes are crushed, before the anthocyanin pigments of the skins begin to dissolve. Relatively short contact (2–6 h) of colorless juice with the skins of crushed dark-skinned grapes allows extraction of sufficient color to yield final products that are various shades of pink. Technically, such products are rosé wines, but they may be labeled as white wines (for example, White Zinfandel) or even as gray (as in *vin gris*).

In red wine production the skins of crushed, dark-skinned grapes are allowed to remain in contact with the grape juice for longer periods of time during the course of the alcoholic fermentation, and the pigments of the skins become increasingly soluble as the fermentation proceeds; the maximum solubility of the pigments occurs when the alcohol concentration of the fermenting juice reaches about 6% by volume. However, long contact of skins and fermenting juice does not necessarily deepen the color of the final wine: when the cells of the grape skins die as the fermentation proceeds, they themselves become adsorbers of the very pigments that were extracted from them earlier. Thus, for maximum color to remain in the wine it may be necessary to remove the fermenting juice from the skins before the fermentation has gone to completion. Removal of the skins from a red wine fermentation before the fermentation is complete is also sometimes done for other reasons. If only a light red color is desired in the final product, the juice is separated from the skins early. This happens also when appreciable tannin extraction (which roughly parallels color extraction) must be avoided, as when wine is specifically being made for early consumption.

Relative sweetness. Purely on a compositional basis, wines are classified as dry or sweet, terms referring to the relative sugar content of the wine. Products whose initial sugar content has been completely, or almost completely, fermented away are dry wines. Sweet wines contain residual sugar and have a recognizably sweet taste.

The occurrence of residual sugar in wines may be achieved naturally or by technical means. For example, if a wine is made from partly desiccated grapes (that is, either raisined, or affected by the mold *Botrytis cinerea* and partly dehydrated), the must may contain more sugar than can be fermented to ethanol by wine yeasts before the ethanol concentration becomes inhibitory (most wine yeasts are intolerant of alcohol concentrations above 14%, but some can tolerate up to 16%). Thus wines made from such musts will, naturally, contain residual sugar. Examples are the Sauternes of France, the *Beerenauslesen* and *Trockenbeerenauslesen* of Germany, the *Vendange tardive* wines of Alsace, the *vini passiti* of Italy, and the Late Harvest wines of California.

Other wines are made to contain residual sugar by the technique of stopping the fermentation of the sugar in the musts of grapes of normal ripeness (that is, grapes neither raisined nor botrytised) by

cooling the fermenting juice and continuing the fermentation under pressure once the desired residual sugar concentration has been reached (both environmental factors acting to halt yeast activity), then filtering the partially fermented product free of yeast cells. This technique is widely used in the production of slightly sweet German white wines. Halting fermentation can also be achieved by adding brandy to a fermenting must before all the sugar has been fermented away. This is the technique used in the production of Porto, in Portugal, and of Porto-type dessert wines, red or white, elsewhere.

Finally, to impart requisite sweetness, the addition of unfermented grape juice (*muté* or *Restsüsse*) preserved with sulfur dioxide to wines that have been fermented to dryness is also sometimes employed. Unfortunately, this technique yields products tasting of grape juice and containing relatively high concentrations of sulfur dioxide, a substance that, in its free state, is not only organoleptically unpleasant but also harmful to asthmatics and other sensitive individuals.

Wine production. Winemaking operations vary widely with the type of wine being produced. In the production of table wines the grapes, harvested when they have reached the desired sugar and acid levels, are normally introduced into a crusher-stemmer, a mechanical device that crushes the berries but not the seeds, and removes the stems. Removing the stems is a standard but not universal practice. Some fine red wines of Burgundy are fermented with the stems in order that stem tannins may be extracted into the wine. The skins of the grape variety Pinot noir from which these wines are made are themselves low in tannins which, as reducing agents, help protect wines from oxidation and give red wines longevity. The stems are also not removed prior to fermentation in the production of the red wines of the Beaujolais. In this region of France, large fermenters are filled with whole grape bunches and sealed almost completely, with only a small opening being provided at the top of the vats for gas escape. The weight of the grapes at the top crushes some of the berries at the bottom, and a fermentation of the crushed berries begins, releasing carbon dioxide which, being more dense than air, forms a layer above the fermenting juice. As the carbon dioxide layer rises and respiration of the cells of the grape skins above the crushed berries is prevented, the cells die and the grape skins burst, releasing more berry juice to be fermented. Eventually all the berries are crushed in this fashion, without the intervention of mechanical crushing machines. This process, known as carbonic maceration, results in red wines particularly fruity-vinous in fragrance.

When white table wines are being fermented, the skins and seeds (a mixture known as pomace) are separated from the juice, and the juice is drawn off either to stainless steel or glass-lined fermenters or, more traditionally, to wooden barrels. The barrels are positioned on their sides and filled to only 65% of their capacity to allow space for the frothing which inevitably occurs during the fermentation. Red wines are fermented in open vats, usually made of wood, but sometimes of stainless steel. As the fermentation of a red wine proceeds, the skins of the crushed grape berries fill with carbon dioxide and rise to the surface, forming a cap layer. This cap must be "punched down" or otherwise remixed with the fermenting juice periodically, so that proper extraction of skin constituents may continue and so that the heat released during the fermentation does not raise the temperature in the cap to the point at which the yeasts are killed, causing the fermentation to "stick" or halt. As the temperature of the fermenting liquid must be controlled, large fermentation tanks are often equipped with heat transfer coils; the temperature of the fermenters is usually maintained in the range 75–86°F (24–30°C).

After the initial fermentation the wine is cellared; red wines are separated from the pomace with a press. The now-alcoholic pomace can be used for distilling purposes (in Burgundy the distilled product is called *marc* and in Italy *grappa*), or it can be used as fertilizer. Red wines are almost always cellared in wooden barrels, in which the alcoholic fermentation is completed and the wine clears itself.

Because extractives leach into the wine from the wood, the wood type and barrel size—the latter determining the wood-to-wine ratio—are important factors in the taste of the final product. American white oak (particularly *Quercus alba,* although other American *Quercus* species are also utilized) is still widely used in the United States for aging table wines and American oak has been found to be the wood of choice for aging Jerez (that is, sherry) wines in Spain. However, European oak (*Q. robur* and *Q. sessilis*) from Limousin or Nevers, France, or Slovenia, Yugoslavia, is used in Europe and increasingly in the United States for aging fine wines from the best *V. vinifera* varieties. In California, large storage casks are often of redwood (*Sequoia sempervirens*), which is essentially taste-neutral in character.

As either red or white wines clear, the yeast cells, and the precipitates that form during fermentation (largely potassium bitartrate, the commercial cream of tartar), settle as a sediment (lees) at the bottom of the storage container. Periodically, and depending on the type of wine being produced, the new wine is drawn off (racked) from this sediment, and is transferred to clean containers. Aging is then continued, and many red wines are kept in wood for 2 years before being bottled. Some white wines are not aged in wood at all, but are merely racked and stored in glass-lined tanks until finally stabilized and bottled. Before bottling, wines may be fined by the addition of substances such as egg white, sturgeon-derived isin-glass, or bentonite clay to help settle the finely dispersed particles that have been suspended in the wine. Some wines are filtered, but tight filtering of red wines is controversial. The most famous (and most expensive) red table wines of the world are not clarified in this way, but are purposely allowed to go to market containing microscopic particles that act as nuclei for slow precipitate formation. Upon long aging, such wines throw sediments in the bottle,

consisting largely of oxidized tannins and anthocyanin coprecipitates. Some wines are cold-stabilized, that is, refrigerated for a short period of time, to bring down excess potassium bitartrate, which might precipitate in the bottle when the product is chilled. In the production of ordinary wine the use of ion-exchange resins for the prevention of potassium bitartrate precipitation is now common throughout the world, although it is not permitted in France. Some inexpensive and poorly made wines are pasteurized before bottling by a brief exposure to a temperature of 140°F (60°C) in the absence of air. This process, while preventing wine spoilage by microorganisms, causes hydroxymethyl furfural formation in the wine, and deleteriously alters the wine's flavor; hence it is avoided whenever possible.

At the time of bottling, white wines, in particular, are usually dosed with a small amount of sulfur dioxide to take advantage of its antiseptic and antioxidative effects. Total (that is, free and chemically bound) sulfur dioxide of 75–150 mg/liter, and free sulfur dioxide of 10–30 mg/liter are usual in well-made white wines. Well-made red table wines usually contain total sulfur dioxide of less than 0.0033 oz/quart (100 mg/liter), most of which is present in various chemically bound forms (sulfonates, for example) by the time the wine is sold.

Fermentation. Almost all naturally fermented table wines are produced by an alcoholic fermentation carried out principally by strains of the sporogenous yeasts *Saccharomyces cerevisiae* var. *ellipsoideus* or var. *oviformis*. The actual fermentation of the grape must, however, if it were allowed to occur completely in the absence of sulfur dioxide and as the result of the activities of the naturally present microscopic flora, would start with the multiplication of asporogenous yeasts such as *Kloeckera apiculata*. These so-called wild yeasts are usually not tolerant of alcohol concentrations in excess of 4% and so would be killed as the fermentation proceeded, to be succeeded then by various *Saccharomyces*. However, since the wild yeasts produce end products other than ethanol and carbon dioxide during their metabolism of glucose, and some of these products, such as acetic acid, are not desirable in wine, their multiplication in fermenting musts is usually discouraged. This is accomplished by the addition of a small amount of sulfur dioxide (in aqueous solution; or in the form of aqueous solutions of potassium bisulfite or metabisulfite) to the must as soon as the grapes are crushed, because most wild yeasts are inhibited by even low concentrations of sulfur dioxide. Addition of sulfur dioxide at this stage is also useful in preventing browning reactions.

In Europe the grapes in long-established wine-growing areas are generally so well covered with *Saccharomyces* cells that satisfactory fermentations take place, after an initial light sulfiting of the musts. In United States practice a pure culture of an appropriate *Saccharomyces* is added after the sulfiting, usually as a water-based slurry of commercially available, dehydrated yeast cells.

Yeasts are an important part of the flora of the grape "bloom," but they are not its only members. Bacteria and molds may also be present. Two types of bacteria are specially significant to the winemaker. Certain heterofermentative lactic acid bacteria (*Leuconostoc oenos*), if allowed to multiply, bring about a malo-lactic conversion. They adventitiously convert the dicarboxylic acid, malic acid, of grape juice to the monocarboxylic acid, lactic acid, and carbon dioxide, and at the same time, by other reactions, impart certain fragrance and flavor constituents to the wine. Acetic acid bacteria (*Acetobacter* sp.), if allowed to multiply, oxidize the ethanol in the wine to acetic acid, thus turning the wine to vinegar. The growth of the lactic acid bacteria is encouraged (by the use of as little sulfur dioxide in the sulfiting as possible, and by extended contact of the wine with its lees) by winemakers who use grapes that are not completely ripe at harvest and whose wines would be too tart were it not for the diminution of acidity these bacteria are able to mediate. The growth of acetic acid bacteria is discouraged by keeping the wine away from air as much as possible and keeping the barrel surfaces that are exposed to both wine and air (such as around the bung holes), clean and antiseptic by scrubbing them with a weak solution of sulfur dioxide in water. *See* FERMENTATION; YEAST.

Oxidized wines. Although most wines are deliberately kept away from air as much as possible during cellaring, two types of wine specifically depend on oxidations for their characteristic fragrances and flavors. These are the *vins jaunes* (yellow wines) of the Jura in France, and the sherry wines of the Jerez de la Frontera region of Spain. The oxygen-wine interaction in both cases is mediated by oxidative yeasts (*Saccharomyces beticus*) that form a pellicle (the flor) on the surface of the wine, which is kept deliberately in partly filled casks open to the air. The yeasts oxidize some of the ethanol to acetaldehyde, some of which may form acetals while some also polymerizes to yield other aldehydes, helping to produce the characteristic fragrance of the products.

Although the *vin jaune* of the Jura is vintage-dated, the wines of Jerez are not, these latter reaching the public with a constancy of character that comes about as the result of the wine being of constant average age when released. This condition is achieved by the *solera* system, that is, the uniform fractional blending of younger wines with old. Most so-called sherry wines of the United States are made in an entirely different way: they obtain their characteristic caramellike bouquets and tastes by being heated at 112–150°F (44.5–65.5°C) for several months in the absence of air. Therefore, they would more accurately be known as American "Baked Wine," rather than sherry, since the heating process they undergo is known as baking. Their method of production thus resembles that of the wines of Madeira, which are heated in the same manner.

Sparkling wines. Sparkling wines have, as type species, the effervescent wine of Champagne, in France. The wines from this region are made by allowing a secondary fermentation of dry white still wine (a wine made from only certain grape varieties,

the Pinot noir, Pinot meunier, and the Chardonnay, all not fully ripe when picked) to which sugar and a flocculating yeast have been added. The mixture is put into thick-walled bottles strong enough to withstand more than the internal pressure of the 3–4 atm (300–400 kilopascals) at about 50°F (10°C) in the final wine. Subsequent to its secondary fermentation, by French law Champagne remains in contact with the yeast sediment in the bottle for at least 1 year, before that sediment is moved to the neck of the bottle, there to be frozen when the bottle is prepared for release. At that time the bottle is opened, the frozen sediment removed, and the wine dosed with a solution of sugar in brandy. Champagnes dosed with sugar liqueur are labeled, in order of increasing sweetness, *brut* (very dry), *sec* (dry), *extra-sec* (extra-dry), *demi-doux* (half-sweet), or *doux* (sweet).

Because the clarification of sparkling wine by the Champagne process is labor-intensive, a less expensive transfer process has been introduced into sparkling wine production in areas other than Champagne. In this method the secondary fermentation occurs in small bottles, but at the time the product is being prepared for market the bottles are emptied, under counterpressure, into a collecting vessel, the sediment is removed by filtration, and the clarified wine is dosed and transferred to new bottles. Such wine may be labeled "bottled fermented."

Sparkling wines are also produced by the so-called Charmat or bulk process, in which secondary fermentation takes place in sealed, large (1000–3000 gal or 3800–11,400 liters) glass-lined or stainless steel pressure tanks. The final product must be separated from the yeast soon after the secondary fermentation is completed, to prevent formation of hydrogen sulfide in quantities sufficient to foul the wine. It thus cannot have the autolyzed yeast character that constitutes part of the distinctive flavor and fragrance of Champagne.

Carbonated sparkling wines are made by charging brilliantly clear still wines with carbon dioxide at 32°F (0°C). They cannot usually be aged, however, because they often throw sediments in the bottle, which are stirred up by the escaping gas as soon as the bottle is opened, creating an unsightly product.

Wine quality. The most important factors influencing the quality of wine in a general way are the grape variety and the particular expression of it that is allowed by the environmental conditions of the growing region. While the technical skill of the winemaker is critical, production of distinctive and quality products requires distinctive and quality raw material.

The most esteemed table wine cultivars in the western world are the low-yield but distinctive noble varieties: Chardonnay and White Riesling (each yielding white wine), and Cabernet-Sauvignon and Pinot noir (each yielding red wine). In terms of quality the most suitable regions for growing these varieties are cool regions whose temperature summations lie between 1800 and 2500 degree-days (the degree-day summation is the summation of the difference between the average daily temperature, in °F, for every day of the growing season, and 50°F). Examples of such regions are the famous wine-growing areas of Bordeaux, Burgundy, the Mosel, and the Rheingau in Europe, and the southern portion of the Napa Valley and the peaks of the Santa Cruz Mountains in California. In areas of lower heat summation the grapes will not fully ripen, and in regions of higher heat summation the mature fruit lacks acidity and other needed characteristics. Other environmental factors that influence grape characteristics are vine exposure, soil drainage, and usually to a lesser extent the chemical composition of the vineyard soil.

Edward J. Wawszkiewicz

Bibliography. M. A. Amerine (ed.), *Wine Production and Technology in the United States*, 1981; M. A. Amerine and M. A. Joslyn, *Table Wines: The Technology of Their Production*, 2d ed., 1970; S. F. Anderson and D. Anderson, *Winemaking*, 1989; R. S. Jackson, *Wine Science: Principles and Applications*, 2d ed., 2000; C. S. Ough and M. A. Amerine, *Methods Analysis of Musts and Wines*, 2d ed., 1988.

Wing

The planar surface of a heavier-than-air body whose primary purpose is to produce lift during flight. A wing is composed of a series of airfoil-shaped components joined together to form one integral structure. On an airplane, the wing also provides storage space for the fuel and undercarriage, and is often used as the structural anchor point for the propulsion systems. The wing also includes important flight-control surfaces, such as the aileron, flaps, slats, and spoilers. *See* AERODYNAMIC FORCE; AIRCRAFT PROPULSION; AIRFOIL.

Wing planform definitions. Several key geometric parameters are used to define the wing planform (**Fig. 1**). The chord of a wing section (c) is defined as the straight-line distance between the leading and trailing edges. Other important parameters are the

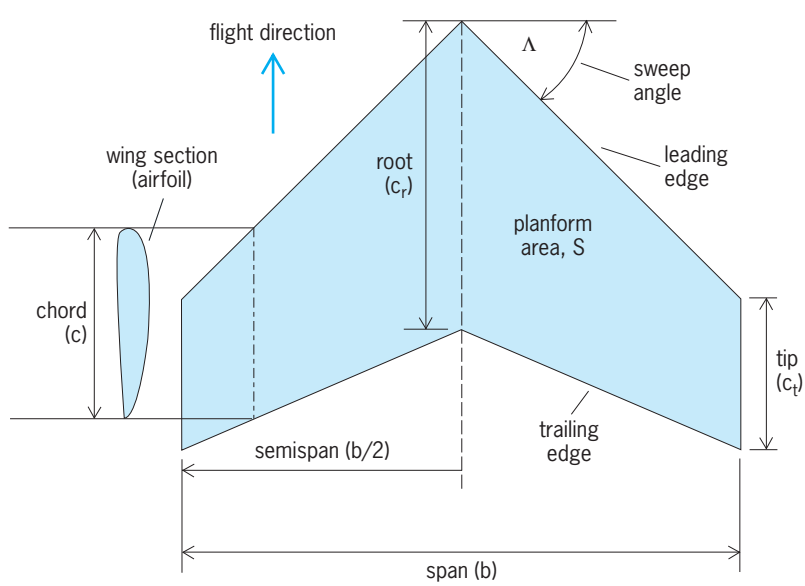

Fig. 1. Wing planform definitions.

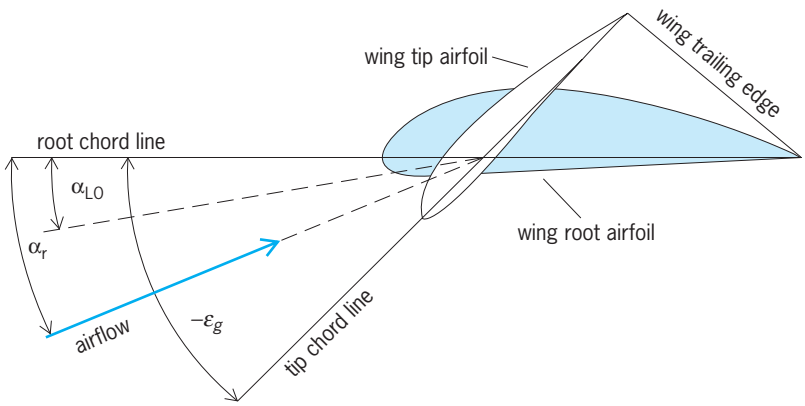

Fig. 2. Wing twist definitions.

leading-edge sweep angle (Λ), the root chord (c_r), the tip chord (c_t), the span (b), the planform area (S), the taper ratio (c_t/c_r), the mean geometric chord ($c_g = S/b$), and the aspect ratio ($A_R = b/c_g = b^2/S$).

Wing twist definitions. The geometric twist at the wing tip (ε_g) is defined as the angle between the root chord and the tip chord (**Fig. 2**). All angles are normally measured in degrees. Twist is defined as positive when the tip is rotated nose up relative to the root chord, and is referred to as wash-in. If the tip is rotated nose down the twist is negative and is called wash-out. As explained later, geometric twist is used to control the air flow over the wing and is, thus, an important wing parameter. The wing angle of attack (α_r) is defined as the angle between the root chord line and the oncoming airflow, known as the free-stream velocity (V_∞). A wing will have a particular angle of attack at which no net lift is produced, and this is known as the wing zero-lift angle (α_{L0}).

Subsonic distribution of lift. It is important to calculate the variation of lift across the wing span, from root to tip, since this is the first step in working out the shear force and bending moment at the wing root which will determine whether the wing structure will fail or not during flight. The shape of the spanwise lift distribution will be influenced by the speed of the air approaching the wing (V_∞) relative to the speed of sound in air (a) which is measured by the Mach number and defined as $M_\infty = V_\infty/a$. If $M_\infty < 1.0$, the flow is classified as subsonic, while, if $M_\infty > 1.0$, the flow is supersonic. The speed of sound decreases slowly with altitude, and has a value at sea level of 340.3 m/s (1116.45 ft/s or 661.5 knots). *See* MACH NUMBER; SOUND.

During subsonic flight, an airfoil produces lift by forcing the air to follow a path over the upper surface which has a larger amount of curvature than that over the lower surface. This flow field causes the air to have a lower static pressure on the upper surface relative to the flow underneath the wing. At each wing tip the lower surface high-pressure air is free to move onto the upper surface, and, as the wing moves forward, two corkscrew flow structures are trailed behind the airplane (**Fig. 3**). These structures are referred to as the wing-tip vortices, and on a humid day they often become visible as water condenses out in the low-pressure vortex core to form two vapor trails behind the airplane. The wing-tip vortices induce a downward flow, known as the downwash (w, measured in the same units as the free-stream velocity), not only behind the wing, but also ahead of the wing. The greater downwash at the wing tip causes the tip airfoil section to experience a much lower angle of attack than the root airfoil, which results in the lift being zero at the tips and a maximum at the wing root (Fig. 3). In the period 1912–1918, Ludwig Prandtl and his colleagues in Germany developed a mathematical model of subsonic flow, known as lifting-line theory, which allows the spanwise variation of lift to be calculated for a given wing geometry, the first step in obtaining the important wing root shear force and bending moment. *See* BERNOULLI'S THEOREM; VORTEX; WAKE FLOW.

Downwash also gives rise to a drag component known as the induced drag (or drag due to lift), which increases as the wing aspect ratio decreases. At low angles of attack, wings which have a large aspect ratio (such as those of gliders and the Lockheed U2) produce more lift for a lower induced drag than wings of low aspect ratio. Thus, high-aspect-ratio wings have a higher lift-to-drag ratio and are often referred to as being aerodynamically efficient. *See* SUBSONIC FLIGHT.

Wing sweep. As the air flows over the wing, changes in static air pressure are transmitted through the air at the speed of sound. If the airspeed anywhere over the wing approaches the speed of sound, these pressure signals cannot travel farther upstream than this point and add together to form a shock wave at this location. If the pressure signal is small, a very weak shock wave forms, which is known as a Mach wave. When the air passes through a shock wave its velocity decreases and the static pressure increases. One undesirable outcome of the formation of shock waves over the wing is a large increase in drag on the airplane. *See* AERODYNAMIC WAVE DRAG; SHOCK WAVE.

On many aircraft, the leading edge of the wing is swept back by an angle Λ degrees (Fig. 1). The effect of sweep relies on the fact that it is only the airspeed normal to the leading edge which controls the

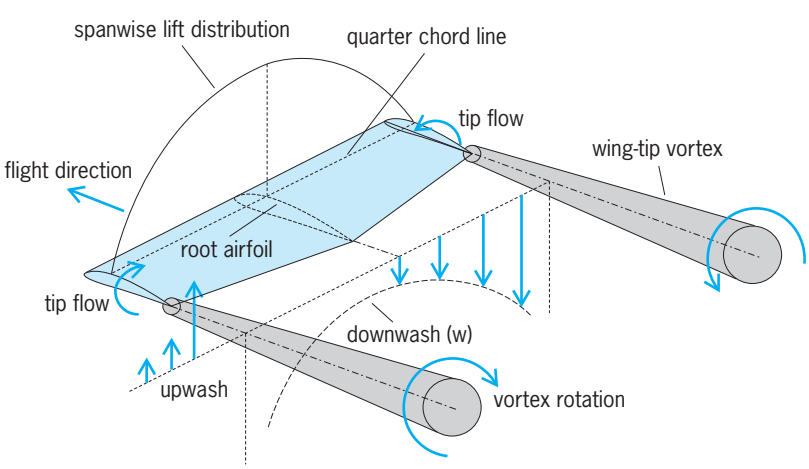

Fig. 3. Subsonic wing-tip vortices.

variation of air velocity over the upper and lower surfaces of the wing. Thus, even though the free-stream Mach number (M_∞) may be equal to unity, the airspeed normal to the leading edge ($M_\infty \cos \Lambda$) may be small enough to avoid increased drag due to shock-wave formation over the wing. Typical values of sweepback for modern transport aircraft range from 25 to 37°. *See* TRANSONIC FLIGHT.

Supersonic distribution of lift. When a wing flies at supersonic speeds ($M_\infty > 1.0$), any source of pressure disturbance can only influence the flow within a downstream conical volume called the Mach cone which extends from the source and whose curved surface is defined by Mach waves (very weak shock waves). Mach waves slope rearwards at an angle of $\sin \sigma_M = (1/M_\infty)$ to the horizontal (**Fig. 4**). One major difference between subsonic and supersonic wing flow is that in supersonic flow the downwash of the wing tip vortices is only felt within the Mach cone and not over the entire wing span or fuselage (Fig. 4). Identifying the regions which are, or are not, influenced by downwash is an important first step in calculating the spanwise distribution of lift which controls the structural loads on the wing during flight. *See* SUPERSONIC FLIGHT.

Wing stall. A wing is said to be stalled when the lift starts to decrease for an increase in angle of attack. In general, this happens when the flow ceases to follow the exact contour of the airfoil geometry, which is known as flow separation. The flow over airfoil A in **Fig. 5** has flow separation at the trailing edge. As the wing angle of attack is increased, the separation spreads over a region which is often referred to as a stall cell. Roll control of the airplane is achieved using control surfaces known as ailerons. These surfaces increase or decrease the camber of the wing section (Fig. 5) which, respectively, increases or decreases the lift. If the lift is increased on the right wing and decreased on the left, the airplane will bank (roll around the fuselage centerline) and begin to fly in a curved path. A wing is normally tapered to reduce the wing root structural loads due to weight, and rear swept to delay the formation of shock waves over the wing upper surface during cruise. However, both taper and sweep increase the likelihood of the stall cell first forming over the ailerons. By introducing negative geometric wing twist, the stall cell is forced inboard toward the wing root, essentially by allowing the wing tips to fly at a lower angle of attack than the root. This maintains smooth airflow over the ailerons and allows the pilot to maintain roll control of the airplane at all times. A similar effect to geometric twist can be achieved by carefully varying the wing airfoil section between root and tip; this is called aerodynamic twist. *See* AILERON.

Wing loading. The total lift produced by a wing is given by $L = \frac{1}{2}\rho_\infty V_\infty^2 S C_L$ where C_L is the wing lift coefficient and is dimensionless. When the air density ρ_∞ is measured in kg/m^3, V_∞ in m/s, and S in m^2, the lift has the units of newton, N. For a given value of free-stream velocity (V_∞) and wing planform area (S), the wing lift coefficient increases with angle of attack (α_r) up until the stall point; this data may be obtained from wind-tunnel tests of a scaled-down model of the wing. It is important to note that, for a given angle of attack, the wing lift is controlled by the square of the flight speed and the wing area.

Wing loading is defined as the airplane mass (W) divided by the wing planform area (S). During takeoff, where the airspeed is low, a large wing area is required to produce sufficient lift to get the airplane airborne, and this results in a low wing loading. However, drag on the wing increases with planform area. A large drag at cruise would be undesirable, especially since, at the higher airspeed, a

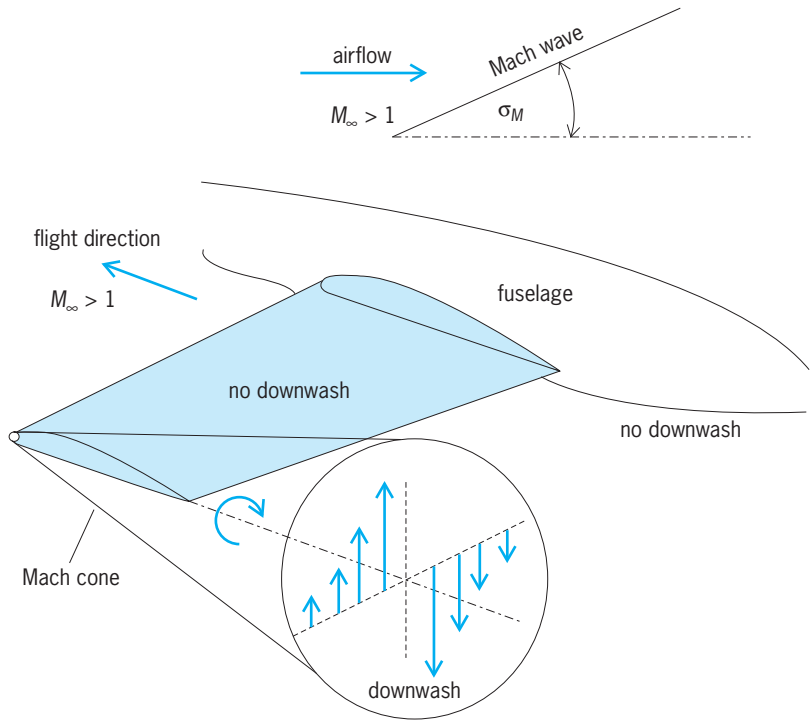

Fig. 4. Supersonic downwash. Inset shows geometry of Mach wave.

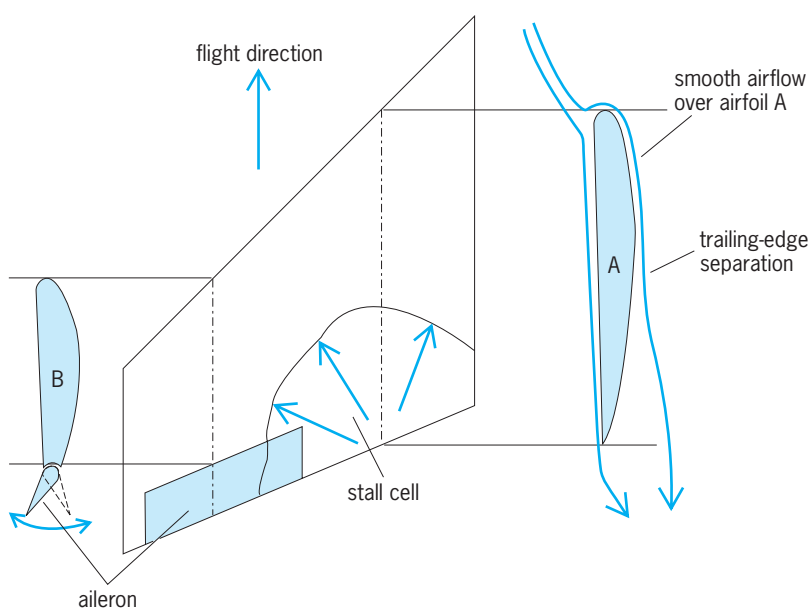

Fig. 5. Stall cell formation.

Typical wing loadings for birds and airplanes

Example	Wing loading, kg/m^2*
Swallow	1.5
Hummingbird	2.4
Golden eagle	7.1
Canada goose	20
General aviation aircraft	Values around 100
Commercial passenger aircraft	100 to 150
Combat aircraft	360 to 580
F-16 fighter aircraft	361

*1 kg/m^2 = 0.2049 lb/ft^2.

smaller wing area could produce the required lift with a higher wing loading, and aircraft with a high wing loading have a greater maximum airspeed than those with lower wing loadings. A compromise is made by increasing the wing area during takeoff and landing through the use of high-lift devices. Typical wing loadings of birds and airplanes are given in the **table**.

High-lift devices. Wing high-lift devices, specifically referred to as slats and flaps, are deployed during take-off and landing when high lift is required at slow flight speeds (**Fig. 6**). However, they are not used during cruise (Fig. 6*a*), since they greatly increase the drag on the airplane. When deployed, these control surfaces increase both the camber and chord of the cruise airfoil section (Fig. 6*b*). Increasing the camber increases the curvature of the air as it flows over the airfoil, which further decreases the upper-surface static air pressure (over the cruise configuration) and, thus, increases wing lift. The increased chord increases the wing planform area (*S*), which also increases wing lift. After touchdown, spoilers are deployed, which increase the drag on the airplane and quickly reduce the lift, thus applying the weight of the airplane on the braking wheels. Spoilers can also be used individually during high-speed flight to reduce lift on one wing only and thus roll the aircraft, and, when used together, to act as an airbrake to slow the aircraft quickly. *See* AIRPLANE; WING STRUCTURE.

Andrew J. Niven

Bibliography. I. H. Abbott and A. E. von Doenhoff, *Theory of Wing Sections*, Dover, 1949; H. Ashley and M. Landahl, *Aerodynamics of Wings and Bodies*, Dover, 1985; J. D. Anderson, Jr., *Introduction to Flight*, 4th ed., McGraw-Hill, 2000; M. Davies (ed.), *The Standard Handbook for Aeronautical and Astronautical Engineers*, McGraw-Hill, 2002; H. Schlichting and E. Truckenbrodt, *Aerodynamics of the Airplane*, McGraw-Hill, 1979; D. Stinton, *The Design of the Aeroplane*, 2d ed., Blackwell Science, 2001.

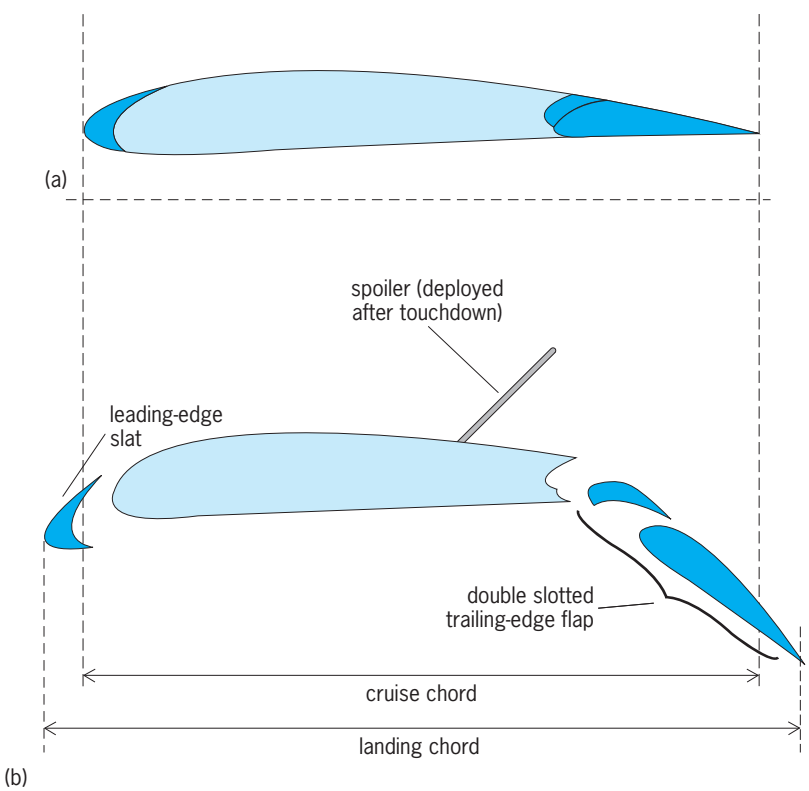

Fig. 6. High-lift devices. (*a*) Configuration during cruise. (*b*) Configuration during landing.

Wing structure

In an aircraft, the combination of outside fairing panels that provide the aerodynamic lifting surfaces and the inside supporting members that transmit the lifting force to the fuselage. The structure of a wing is an integration of the environment external to the vehicle wing, the aerodynamic shape of the wing, and the proposed use of the vehicle. The interaction of these three aspects of design leads to the selection of material, to the general structural layout, and, finally, to the detailed choice of structural shapes, material thickness, joints, and attachments. The result is a structural framework covered with a metal skin that also contributes to the load-carrying function. *See* SUPERCRITICAL WING; WING.

Structural materials. Wing structure has evolved from the early use of wood, doped canvas, and wire. The first general change was the replacement of wood with metal frameworks; wood and canvas structure is still found, however, in lightweight personal aircraft. Doped canvas was replaced by light metal skins, which served only as fairing in the beginning but which were later designed to provide a portion of the structural strength. Today, aluminum alloy outer skins are prime structural elements on all commercial transports and on the great majority of military craft.

Magnesium, steel, and titanium are also used in internal primary structure and in local skin areas. For materials used in the wing structure of research and other exotic aircraft *see* ATMOSPHERIC ENTRY.

Design of a wing structure begins logically with a derivation of loads on the wing, both flight and net, as they are affected by particular design specifications; then the design proceeds through a choice of material to the final phases, when the configurations of major structural units such as the prime wing box

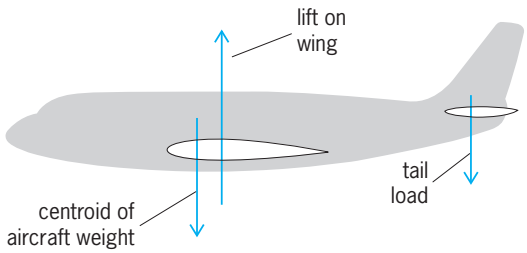

Fig. 1. Schematic diagram of major air loads.

and the leading- and trailing-edge subassemblies are determined.

Total flight load. Flow of air past the wing's airfoil shape creates a lifting force. Total lifting force L, in pounds (newtons), on an entire wing is given by the equation below, where ρ is air density in slugs per cubic foot (kilograms per cubic meter), V is velocity

$$L = \frac{\rho V^2 C_L S}{2}$$

in feet per second (meters per second), S is wing area in square feet (square meters), and C_L is lift coefficient, which is a nondimensional coefficient derived from wind tunnel tests. *See* AERODYNAMIC FORCE; AIRFOIL.

On a particular craft, the lifting force required of a wing depends on the craft's balance. For stability, the aircraft is designed to maintain the centroid of any permissible loading condition forward of the center of lift on the total wing (**Fig. 1**).

The total load on the wing in normal cruising flight is the weight of the aircraft plus the balancing tail load. The total maximum anticipated load in service is much greater than this because the aircraft will encounter atmospheric gusts and because it is required to maneuver. The designed breaking strength of the structure is normally set at 1.5 times the maximum anticipated loads, the 1.5 multiplier being a factor of safety. The resulting load is called the ultimate design load.

Maneuver loads. An aircraft must be capable of climbing, descending, turning, or, if it is a fighter, of executing rather violent acrobatics. Any such changes from a statically balanced condition produce load increments. These loads are generated when the pilot induces a sudden change in the angle of attack of the aircraft by manipulating the engine throttles and controls. An abrupt change in speed or angle of attack alters the lift coefficient or velocity and produces a change in total lift. The permissible degree of maneuverability of any specific model is based on its intended use. The **table** presents general information on ranges of maximum anticipated load for various types of craft. Down increments of maneuver load are limited to two times the weight for transports and to somewhat higher values for the others. The ultimate design load is 1.5 times the value shown. *See* FLIGHT CONTROLS.

Maximum anticipated upward flight loads

Type of craft	Required lift force on craft
Transport	2.5 × weight
Bomber	2.5–4 × weight
Fighter	6–9 × weight
Special research	Up to 12 × weight

The multipliers in the table, which are called load factors, reflect the intended usage of the craft. Experience has shown these to be adequate, but the pilot must be aware of the maneuver limitations of the craft. It is possible to exceed these loads in flight. **Figure 2** shows a simplified presentation of load factor and factor of safety applied to total external wing load. Tail load is assumed zero for simplicity of presentation. The presentation shows that the external air load equals the aircraft weight during cruise. In a 2.5-g maneuver, maximum anticipated air load is 2.5 times the aircraft weight, the 2.5 being a design factor. On this basis, ultimate design load is 1.5 × 2.5, or 3.75, times the aircraft weight, the 1.5 being the safety factor.

Gust loads. A second significant load produced on any craft's wings is an atmospheric gust. A high-velocity vertical gust produces an abrupt change in the direction of airflow over the wing. An upward gust produces an upward incremental force, and a down gust produces a downward incremental gust. The U.S. Civil Aeronautics Board (CAB) specifies the multiplying gust load factor.

For commercial aircraft, the normal gust velocity in feet per second (1 ft/s = 0.3 m/s) is specified as $30K$ at cruise speed, $15K$ at dive speed, and $40K$ for the speed recommended for flying under more turbulent conditions (less than cruise speed), where K is a coefficient dependent on wing loading, which is the ratio of aircraft weight to wing area. The recommended speed during turbulent conditions is determined from the aerodynamics of each aircraft type. The value K is about 1.20 for many typical current designs. The factors must be derived for both upward and downward gusts, which produce loads on

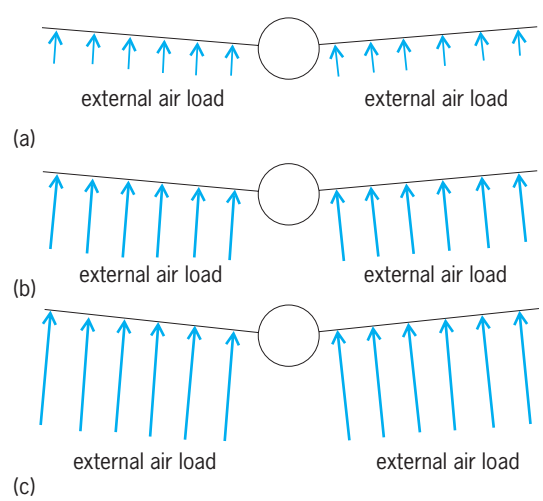

Fig. 2. Load factor and safety factor applied to total external air load. (*a*) External air load during cruise. (*b*) Maximum anticipated air load in 2. 5-*g* maneuver. (*c*) Ultimate design load with 1.5 factor of safety.

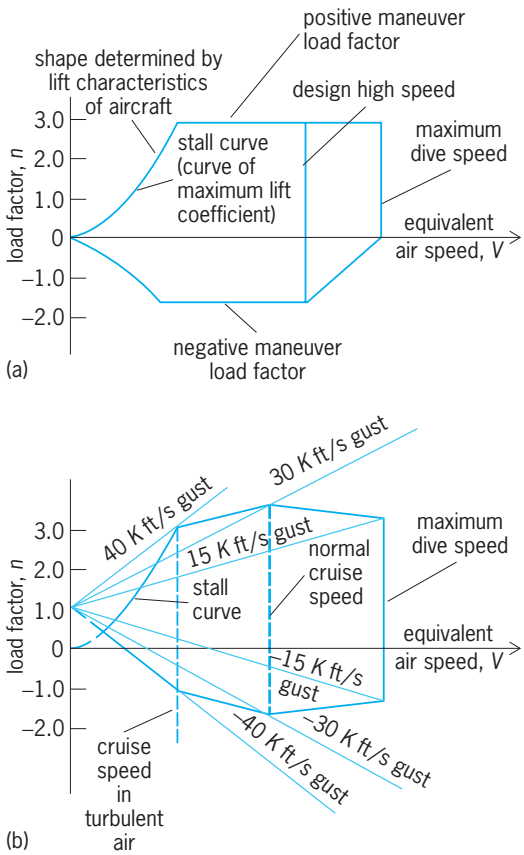

Fig. 3. *V-n* diagrams of maximum anticipated flight loads. (*a*) Maneuver envelopes. (*b*) Gust envelope. 1 ft/s = 0.3 m/s.

the wing comparable to the up and down maneuver loads. The most critical among these loads must be determined for each element of the wing structure.

V-n diagram. The design speeds and load factors for any design are summarized in a graph called the $V-n$ diagram (**Fig. 3**). The curve is derived from the specified requirements for the particular use of the vehicle. The curve applies to the craft as a whole. Balance considerations depicted in Fig. 1 permit the derivation of the total flight loads.

Landing and takeoff loads. Landing and takeoff loads are specified as part of the CAB air regulation for commercial vehicles and in connection with ground loads for military vehicles. Commercial aircraft must be capable of anticipated descent velocities of 10 ft/s (3 m/s) at design landing weight and 6 ft/s (1.8 m/s) at design takeoff weight. These descent velocities must be considered at different landing attitudes of the craft. Anticipated weight and centroids of design conditions are established by considering the vehicle usage. In addition, one-wheel landings, lateral drift, braking, turning, and pivoting are evaluated and defined in terms of loads. The dynamic effects of spinning the landing gear wheels at touchdown and the subsequent spring-back of the gear and its attaching structure are critical conditions that are also evaluated. In general, the specified loadings and state-of-the-art development of landing gear shock-absorbing devices produce actual vertical loads on each main gear approximately equal to the aircraft weight. Maximum aft loads on the main gear are of the same order of magnitude, and side loads are 60–80% of this. This load envelope at landing usually determines the design of structure in the wing to which the main gear attaches. Because of the large twisting forces induced by the drag loads applied at ground level, which must be carried up and into the wing, this envelope is also critical for significant portions of the inboard wing structure. Ten or so landing conditions affect specific portions of the wing. The ultimate reaction of these loads is, of course, the weight and rotational inertia of the fuselage. The gear is expected to withstand actual anticipated loads without permanent distortion and to withstand ultimate design loads without structural failure. *See* LANDING GEAR.

Net wing loads. The maneuvers and gusts acting on the craft also generate accelerations and inertia effects. Therefore, the net loads on the wing, which the structure must sustain, are the net sum of the previously derived loads and the counteracting inertial effects. A simple example of a flight condition is presented in **Fig. 4** for a hypothetical 100,000-lb (45,000-kg) aircraft in which the tail load is assumed to be negligible. In essence, the wing is first assumed to be isolated from the rest of the craft. Then the air loads and inertia loads are applied, and the net of these is balanced with the remainder of the craft and applied to the wing through the wing-to-fuselage connections.

From 15 to 25 different flight conditions are normally derived to satisfy possible critical conditions. These flight conditions encompass variations of

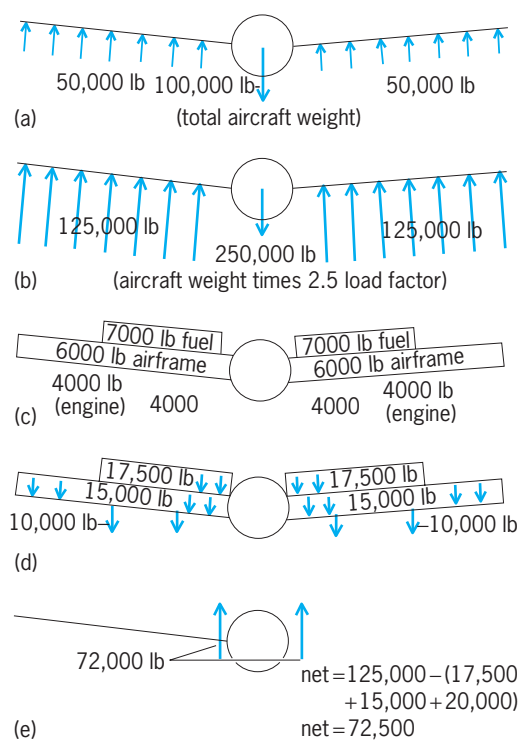

Fig. 4. Derivation of net loads. (*a*) External air load during normal cruise. (*b*) Total wing load during 2.5-*g* maneuver. (*c*) Wing dead weight at 1 *g*. (*d*) Wing weight times 2.5 load factor. (*e*) Net shear load to fuselage. 1 lb = 0.45 kg.

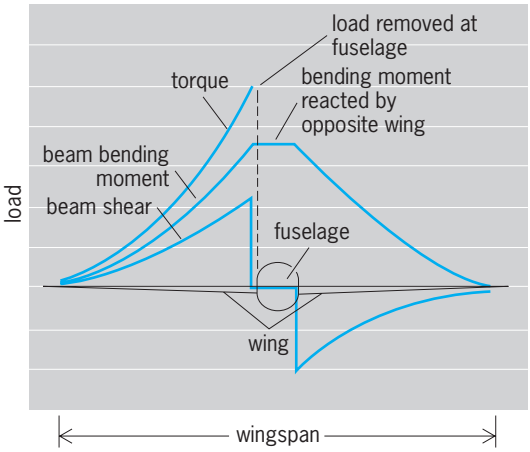

Fig. 5. Shear, bending moment, and torque curves.

aircraft gross weight and centroid location, fuel loadings, upward and downward maneuvers and gusts, special control surface and landing flap conditions, and any unusual dynamic loadings, such as dynamic interactions of unsteady airloads and structural stiffness. Dynamic loads are also affected by concentrated weights such as engines, tip tanks, special pods, and gun or rocket reactions. *See* AEROELASTICITY.

Wing structure is examined at a number of wing cross sections in order to taper the weight of the material from the tip to the inner section of the wing. This is done by summing the net loads from the tip of the wing to the section in question. This calculation yields the vertical and drag shear forces, the bending forces produced by them as they are carried in toward the fuselage, and the twisting or torsion forces produced by an eccentricity of the vertical and drag forces about the twisting center of the structure. These forces are presented as beam (vertical) and chord (drag) shear curves, beam and chord bending moments, and torque (**Fig. 5**).

Choice of material. The derivation of the net loads permits a quantitative consideration of the general structural framework. Design of the structure, however, is inseparable from the choice of material. The thickness and shape of the individual pieces also affect strength in lightweight, efficient structures because of compression stability modes.

High static strength and light weight—coupled with reasonable rigidity characteristics, good corrosion resistance, fatigue endurance, cost and formability—are prime considerations in choosing airframe material. Highly loaded, efficient airframes built since 1930 have found these qualities in aluminum. Alumium has withstood the technical competition of steel, titanium, magnesium, wood, and reinforced plastics, although complete designs have been manufactured from each of these competitive materials to try them with complete realism. At present, composite materials are being introduced into both military and civilian aircraft at an increasing rate. *See* COMPOSITE MATERIAL.

Aluminum has been somewhat supplanted by the requirements for high-supersonic and hypersonic speeds in the atmosphere (the thermal thicket) and by the demands of space technology (the atmospheric reentry problem). But even in the case of aerospace vehicles with these problems, designs incorporating protective shields and cooling of the structure have been considered to permit the use of aluminum with its favorable static strength, corrosion resistance, fatigue strength, cost, and formability, at so light a weight. *See* AEROTHERMODYNAMICS; HYPERSONIC FLIGHT.

It does appear, however, that the demands for speed interacting with the physical environment are causing aluminum to be replaced by titanium and other alloys to some extent. A material's creep characteristics and strength at high temperature are now important added parameters.

Authoritative data on strength properties of aircraft metals are available. **Figures 6–9** present in

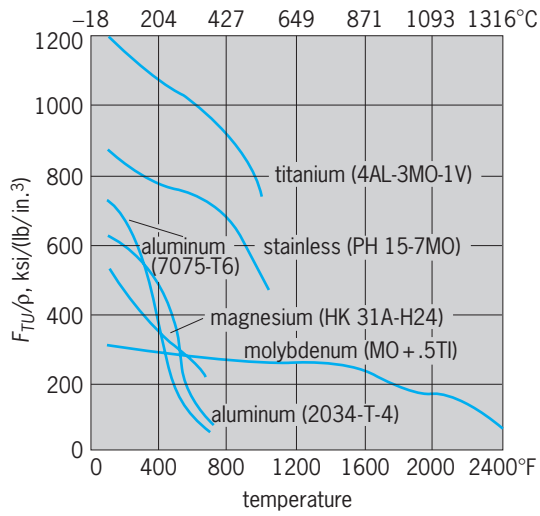

Fig. 6. Graph showing ultimate tensile strength F_{TU} divided by density ρ for various materials. 1 ksi/(lb/in.^{3}) = 0.25 MPa/(g/cm^3).

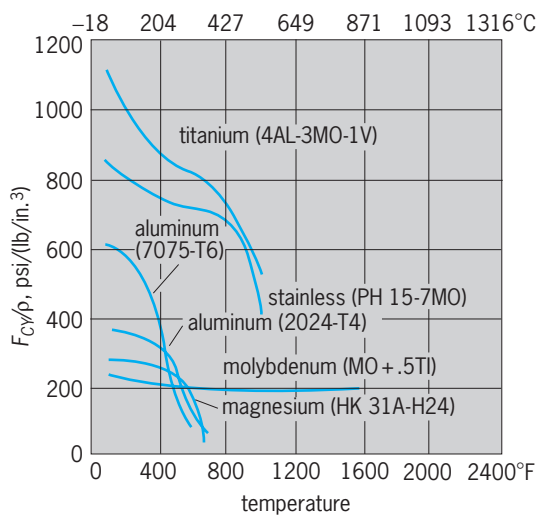

Fig. 7. Graph showing compressive yield strength F_{CY} divided by density ρ for various materials. 1 psi/(lb/in.^{3}) = 0.25 kPa/(g/cm^3).

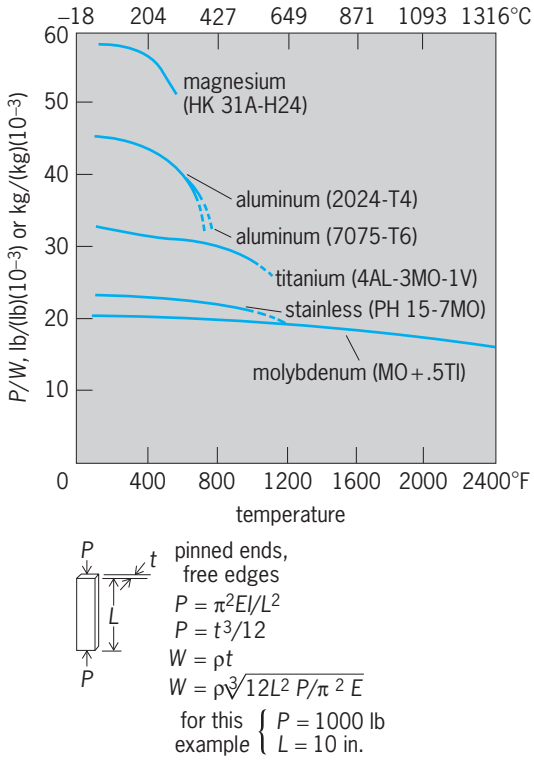

Fig. 8. Critical compression load *P* divided by weight *W* in elastic region (Euler column) for various materials. 1000 lb = 454 kg. 10 in. = 25.4 cm.

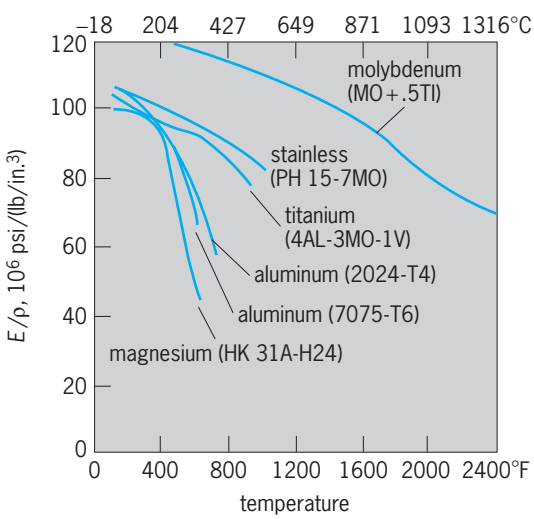

Fig. 9. Modulus of elasticity *E* divided by density ρ for various materials. 106 psi/(lb/in.^{3}) = 0.25 GPa/(g/cm^3).

graphical form comparisons of strength and stiffness with temperature. These curves indicate the superiority of aluminum at room temperature, and the disparity among materials as the temperature increases. For example, the ultimate tensile strength F_{TU} per unit weight is useful for application where tension loads predominate, such as on the lower surface of the wing (Fig. 6). The favorable tension ratios of titanium and molybdenum are negated by cost or formability, as well as by consideration of all properties shown in Figs. 7, 8, and 9. The compressive yield strength (F_{CY}) per unit weight is useful where compressive stresses and compression stability considerations are prime factors, such as on the upper surface of the wing (Figs. 7 and 8). The plot of Young's modulus of elasticity per unit weight is a useful parameter for stiffness considerations (Fig. 9). *See* STRESS AND STRAIN; YOUNG'S MODULUS.

Hypersonic flight in the atmosphere and the reentry problem emphasize characteristics such as strength at high temperature, thermal conductivity, specific heat, coefficient of expansion, melting point, creep, and oxidation temperature, as well as the properties already mentioned. Data on the thermophysical properties for many of the solid materials that may be used in the future are being measured.

Prime structural framework. An airframe wing is essentially two cantilever beams joined together. Each wing tip is the free end of the cantilever, and the centerline of the vehicle represents the plane where the two fixed ends of the cantilevers are joined (**Fig. 10**). The prime load-carrying portion of these cantilevers is a box beam made up usually of two or more vertical webs, plus a major portion of the upper and lower skins of the wing, which serve as chords of the beam. This box section also provides torsional strength and rigidity (**Fig. 11**). Normally the prime box is designed to carry all the primary structural loads; these include all beam shears and bending moments, all drag shears and bending moments, and the torsional or twisting loads. *See* CANTILEVER.

Leading and trailing edge portions of the wing, forward and aft of the prime box respectively, help to provide the airfoil shape required. These portions are designed to minimize their participation in the major load-carrying function. Where participation is forced by the detail design, the fasteners and

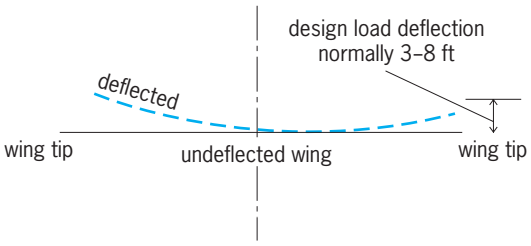

Fig. 10. Wing tips deflect upward in normal flight. 3–8 ft = 0.9–2.4 m.

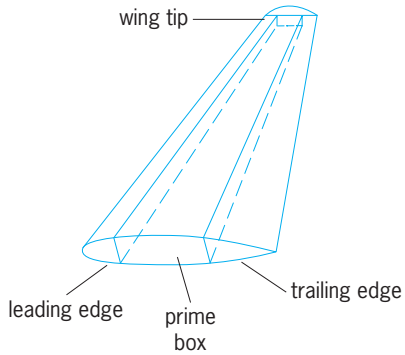

Fig. 11. Prime box of wing half-span.

materials reflect this, but normally the prime box strength is not reduced. An overlapping conservative assumption thus results.

Beam shear material. The vertical webs of box beams are the prime load path for the beam or vertical shears. These webs are comparable functionally to the webs of plate girders in highway or railroad bridges. The prime difference is the elastic buckling permitted in the aircraft web and consequently the much lighter material thickness. This buckling is clearly shown in **Fig. 12**, a photograph of a major aircraft beam under laboratory test where the shears imposed are buckling the web. When the load is removed, the buckles will disappear. This type of beam is called a partial tension field beam. *See* PLATE GIRDER.

The web material of these beams is designed by net flight loads in the outer portion of the wing. The beam and torsional or twisting loads will normally combine to produce the loads which design the webs near the front of the wing. This is usually a large angle of attack condition. A low angle of attack condition, or one involving use of the control surfaces in the trailing edge region, normally produces design shears in the rear webs. In the inner portions of the wing, especially inboard of the landing gear, the landing conditions produce shears which exceed the flight loads by significant amounts. These loadings require substantially thicker webs. Aluminum webs of 0.020 in. (0.5 mm) thickness may be anticipated in the outer portions of the wing. These increase toward the wing root, and thicknesses of 0.125–0.188 in. (3.2–4.8 mm) may occur at the inner section.

One of the beams from a twin-engine transport is shown in **Fig. 13.** A portion of the beam at the centerline of the aircraft is of truss construction to permit access for equipment and for inspection and maintenance.

Top of wing box. The top of the prime box is the compression chord of the cantilever beam. It is also the portion of the box where the greatest variations in construction are found among the products of various manufacturers. The desirability of minimum weight and the requirements of the various positive and negative bending conditions have produced numerous configurations (**Fig. 14**).

These cover structures are designed as beam columns or plate structures. They are subjected to axial load by the beam and chord bending moments, coupled with lateral loads from air pressures and, if applicable, fuel tank pressures. The final design is achieved by an iterative process. Approximate material sizes are established for a complete cross section, that is, for top of the wing box, bottom, and connecting corner members. Detail stresses are calculated by methods fully described in texts. Repetitive refinement of material sizes finally results in an efficient top cover based on minimum weight and practical construction methods. Here, too, the gages begin with a minimum of 0.010-in. (0.25-mm) aluminum in honeycomb or 0.020-in. (0.5-mm) in corrugated forms. The plate designs begin with 0.050- to 0.060-in. (1.25- to 1.5-mm) minimum and reach thicknesses of 0.750 in. (19 mm) at inboard sections on some of the largest craft. The basic material thicknesses are determined for the compression loads and are virtually always adequate for load reversals which would put the top of the box in tension. Frequently, elements of the top of the box, including splices, are tested during the development of the craft to ensure adequate strength for the final design. *See* AIRCRAFT TESTING; BEAM COLUMN.

Fig. 12. Partial tension field beam under test.

Fig. 13. Center section of main wing beam of a commercial transport.

Bottom of wing box. The bottom surface of the prime box is the tension chord of the beam. The bottom surface is most frequently a skin, tapering in thickness from a minimum at the wing tip to much heavier gages at the root. Specific gages are on the order of two-thirds to three-fourths as thick as those for the top cover of the wing box because the allowable tensions are always higher than the permissible stresses for compression stability. Fatigue strength is a significant consideration in bottom surface design. Analyses and tests are used to provide the most reliable answers available. Spanwise reinforcing members on the bottom surface bring the compression strength up to requirements for load reversals. The lower surface must normally carry 35–40% of its tension allowable as a compression design load. Framing members around access openings help carry the stresses around such structural discontinuities in the lower wing frame (**Fig. 15**).

Wing ribs. At numerous places within the wing box, bulkhead-type structures called ribs are located.

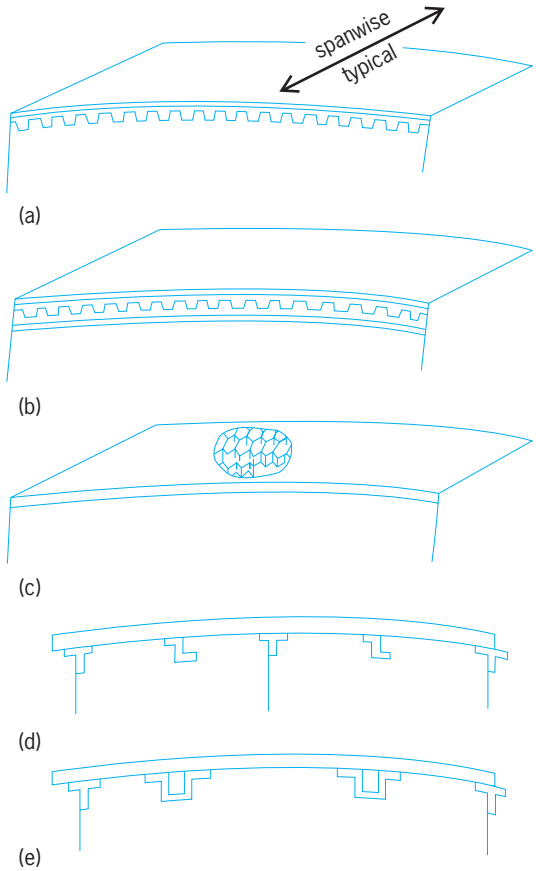

Fig. 14. Types of top cover structure. (*a*) Single skin and corrugation. (*b*) Double skin and corrugation. (*c*) Honeycomb with two skins. (*d*) Thick plate and intermediate stringer. (*e*) Plate with relatively heavy stringers.

These internal structures serve to maintain the rectangular box shape and to cut down the unsupported length of compression cover structures, to separate fuel tanks, and to distribute concentrated loads from guns, bombs, landing gear, or engines into the prime box. They are also located at any wing cross section where major load redistributions occur.

Fig. 15. Lower surface framework of outer wing of a commercial transport.

Wing weights and statistics. The weight of the aircraft divided by the wing area is called the wing load. From 1940 to 1960 wing loadings steadily increased from 35 to 100 lb/ft^2 (171 to 488 kg/m^2) on high-speed aircraft, and they have continued to increase since then.

While wing loading or intensity of loading has been increasing, the thickness of the wing has decreased from approximately 10% of chord (the distance from the leading edge to trailing edge) to 5% of chord. Thus, higher loadings on a thinner beam have increased the challenge to the designer to provide more strength in a shallower space. This has been done at the cost of increasing weights. In 1940, 4 lb/ft^2 (20 kg/m^2) represented a reasonable allowance for average wing weight. In 1960, a high-speed aircraft required about 10 lb/ft^2 (49 kg/m^2). Supersonic speeds demand specific weights well above 10 lb/ft^2.

Leading-edge structure. The leading edge, or most forward portion of the wing, serves an important aerodynamic function in establishing smooth airflow and efficient lifting power for the wing. This area sustains the highest aerodynamic pressures.

Structurally, the leading edge is an appendage, a fairing whose local loads must be supported by the prime wing box. This structure takes the form of cantilever beams or arch-type structures from the leading edge back to the front vertical beam (**Fig. 16**). These leading-edge structures are normally of 0.032- to 0.972-in. (0.8- to 1.8-mm) aluminum, with skin covers in the same thickness range.

Frequently the leading edge of an aircraft incorporates anti-icing provisions. Two prime systems have been used in the past. One type is a rubber bladder that is stretched over the leading edge and is alternately inflated slightly and deflated. This action cracks any ice formation, and the airstream sweeps it away. A second method is the hot leading edge. In this system, air that is heated by special heaters or by the engine exhaust is passed along the leading edge to melt the ice. This heated air normally exhausts at the wing tip.

Numerous aircraft have leading-edge slats. These small airfoil sections are pulled out and forward automatically by air pressures and serve to maintain smooth airflow to a higher angle of attack than the basic large airfoil would sustain. As the angle of attack is decreased, these devices are withdrawn into the basic airfoil by springs (**Fig. 17**). Slats are usually designed in short two-hinge lengths to eliminate wing deflection effects.

Trailing-edge structures. The trailing-edge structure is noteworthy for its various auxiliary devices which assist in aircraft control or in reducing aircraft landing speeds. Various forms of flaps, ailerons, and spoilers, with their hydraulic or electrical control mechanisms, fill the volume of space aft of the rear beam or spar (**Fig. 18**).

Landing flaps. Landing flaps consist of a movable airfoil-shaped structure located aft of the rear beam or spar. They extend about two-thirds of the span of the wing. Their aerodynamic function is to increase

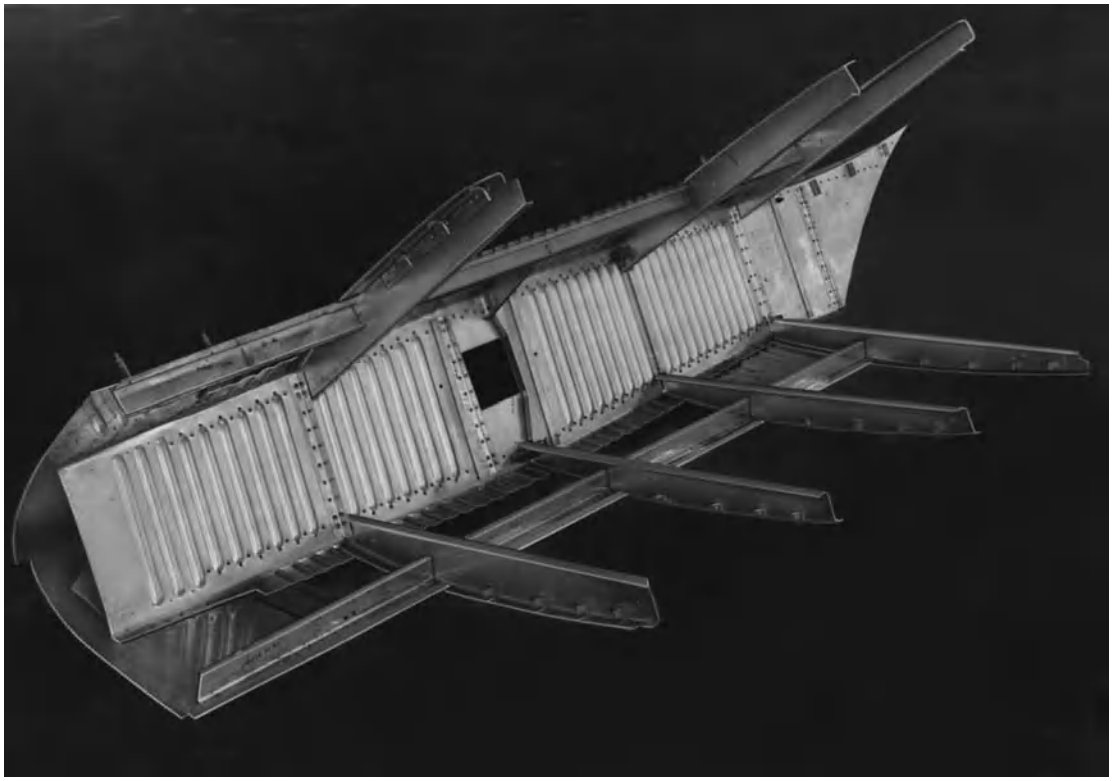

Fig. 16. Section of wing leading edge of a commercial transport.

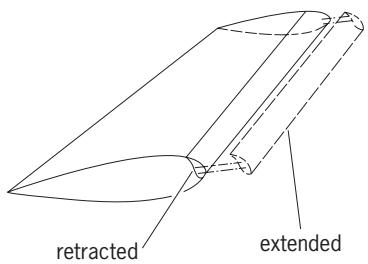

Fig. 17. Leading-edge slat is usually in short two-hinge lengths.

substantially the lift, thereby permitting lower takeoff and landing speeds.

Flaps may be hung from the wing on two, three, or four hinges. Structurally they represent a simple or continuous beam, depending on the number of hinges. Where more than two hinges exist, the wing deflection effects must be considered in addition to local air loads. That is, as the wing deflects it loads the flap in bending it to the shape of the deflected wing. Flaps normally have a single spanwise main beam. Leading-edge ribs and trailing-edge ribs are attached to this main beam, and the airfoil shape is skinned over (**Fig. 19**). Normal gages in these aluminum structures are between 0.020 and 0.040 in. (0.5 and 1.0 mm). The hinge supports are cast or forged aluminum members. The hinges may be plain or roller bearings. *See* ELEVATOR (AIRCRAFT).

Ailerons. Ailerons are located near the tips of the wings in the trailing edge. Their prime aerodynamic function is roll control of the aircraft. Structurally, the aileron is similar to the landing flap. *See* AILERON.

Spoilers. Spoilers are a special form of control surface and perform the aerodynamic function of the aileron, that is, roll control. They are located on the upper surface of the trailing edge at about midsemispan and deflect only upward by pilot control. The cross section of the spoiler is usually rectangular. It is a small box beam consisting of two spanwise beams plus ribs and covering skin.

Building the trailing edge. The various aerodynamic surfaces in the trailing edge region require 5–20 hinges whose locations must be closely controlled. The

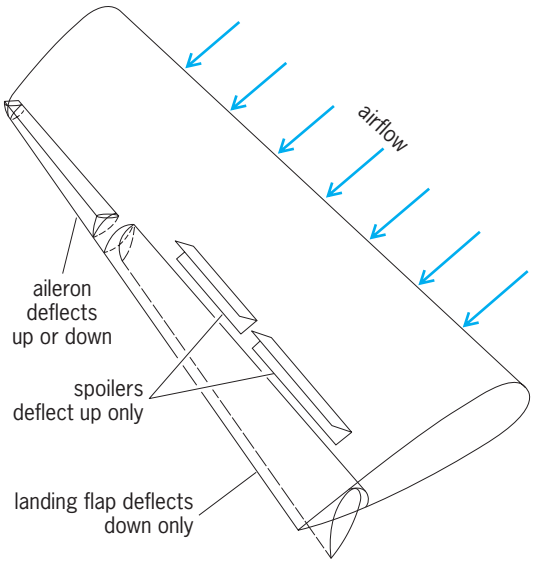

Fig. 18. Trailing-edge surfaces.

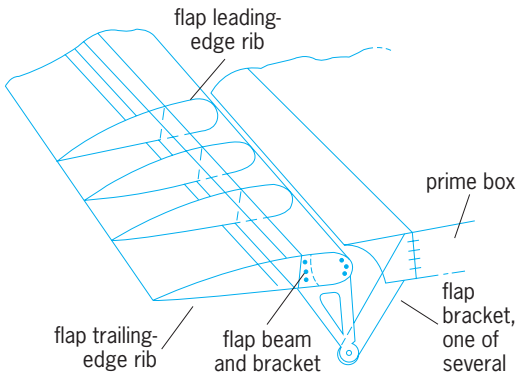

Fig. 19. Flap structure with skin removed. Although external hinges are shown, designs employing internal tracks and hinges are also used.

Fig. 20. Trailing edge with hinges for movable surfaces.

usual method is to build the fixed trailing-edge structure on the rear beam or rear spar of the wing with a master tool holding the locations of the hinge points for the movable surface (**Fig. 20**). The movable surfaces are then built from a matching hinge location tool. Thus, when the assembled surface is brought to the wing it is certain to fit without binding.

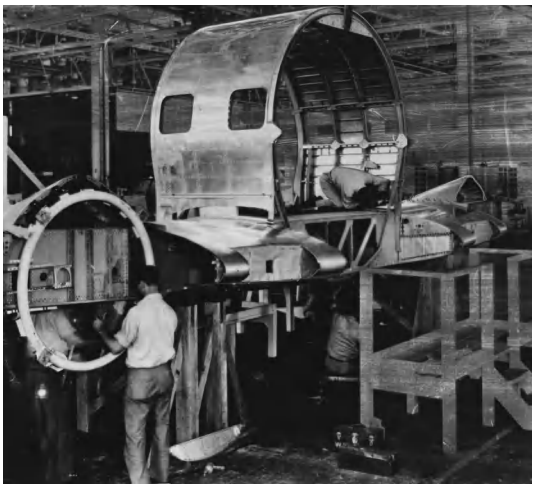

Fig. 21. Center wing-to-fuselage connection.

Wing-to-fuselage structure. The structural heart of the aircraft is the wing-to-fuselage joint (**Fig. 21**). This connection is usually the most complex in form and in analysis. Two major structural elements, the wing and the fuselage, with major loads running at right angles to one another, must be joined and analyzed for consistent deformations. The detail methods of indeterminate structural analyses are used. The calculations may require anywhere from 10 to 100 redundants. High-speed digital computers are a necessity.

The wing structures which have been described are attached to heavy aluminum ring frames in the fuselage; these rings distribute the wing loads to the fuselage skins. *See* FUSELAGE.

Assembly. The assembly of the wing can be traced by means of the illustrations. Figure 13 is representative of a main spanwise beam. To this, leading-edge structures, similar to that shown in Fig. 16, are added to produce a leading-edge and front-beam assembly. In the same manner, a beam and the fixed trailing-edge structures are joined together. Such an assembly is shown in Fig. 20. The rear beam and flap and aileron brackets are clearly indicated. These two major assemblies are then positioned in a major tool jig, and the intermediate beams, ribs, framing members, and top and bottom skins are progressively attached. Figure 15 shows such an assembly just before the attachment of the lower skins which close in the wing box. Figure 15 is an outer wing which eventually is attached as a unit to the center portion of the wing and fuselage. The center wing portion is shown in Fig. 21. The circular tubular structure at the left edge of Fig. 21 is part of the engine nacelle. *See* AIRFRAME.

Harvey J. Hoge

Bibliography. American Society for Testing and Materials, *Damage Tolerance in Aircraft Structures*, STP 486, 1971; J. Cutler, *Understanding Aircraft Structures*, 3d ed., 1999; B. K. Donaldson, *Analysis of Aircraft Structures: An Introduction*, 1993; T. H. Megson, *Aircraft Structures for Engineering Students*, 2d ed., 1990; D. J. Peery and J. J. Azar, *Aircraft Structures*, 2d ed., 1982; S. P. Timoshenko and J. Gere, *Theory of Elastic Stability*, 2d ed., 1961.

Winged bean

A plant (*Psophocarpus tetragonolobus*), also known as four-cornered bean, asparagus pea, goa bean, and manila bean, in the family Leguminosae. It is a climbing perennial that is usually grown as an annual. It has been suggested that it originated either in East Africa or Southeast Asia, but there is more evidence to support an African origin. However, Southeast Asia and the highlands of Papua New Guinea represent two foci of its domestication.

Traditionally, winged bean is grown as a backyard vegetable in Southeast Asia and a few islands of the Pacific. It is grown as a field crop in Burma and Papua New Guinea. However, between 1980 and 1990 winged bean was introduced throughout the tropical world.

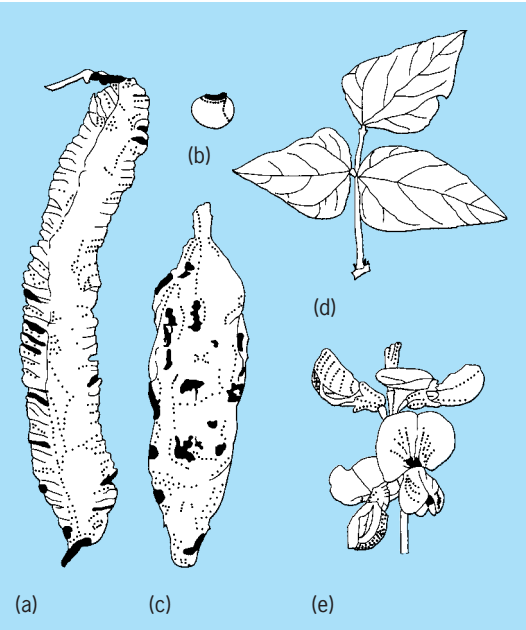

Edible parts of the winged bean: (*a*) pod, (*b*) seed, (*c*) root-tuber, (*d*) leaf, (*e*) flowers.

Nutritional value. Almost all parts of this plant are edible and are rich sources of protein. The green pods, tubers, and young leaves can be used as vegetables, and the flowers can be added to salads (see **illus.**). The dry seeds are similar to soybeans and can be used for extracting edible oil, feeding animals, and making milk and traditional Southeast Asian foods such as tempeh, tofu, and miso. Flour from the winged bean can also be used as a protein supplement in bread making.

The winged bean seed is nutritionally the most important product, containing 30–43% protein and 11–24% fat. The amino acid composition is similar to that of soybean, being low in sulfur-containing amino acids. The saturated/unsaturated fatty acid ratio is 1:3. The tocopherol content is high. The oil is refined easily and is reasonably stable. Phosphorus and zinc occur in significant quantities, and the vitamins thiamine and riboflavin are present in amounts comparable to that of other grain legumes. Tubers contain 8–10% protein on a fresh-weight basis, although the essential amino acids occur in low amounts. Immature green pods contain 1–3% protein and are rich sources of calcium, iron, and vitamin A. Leaves contain 5–7% protein on a fresh-weight basis and large amounts of vitamin A, vitamin C, and essential minerals.

Cultivation and harvesting. The winged bean is best adapted to the equatorial climate. Although it is grown at altitudes up to 6600 ft (2000 m) in Burma and Papua New Guinea, it does not tolerate frost. Short days (less than 12 h of sunlight) are necessary for both flower and tuber initiation. It can grow on a variety of soil types, but good drainage is necessary. It nodulates in association with the cowpea group of *Rhizobium* strains, which are widely distributed in the tropical soils. It responds to phosphorus fertilizers and also to potassium if phosphorus is not limiting.

In the tropics, sowing should be done at the commencement of the rainy season. Irrigation is necessary if a prolonged dry spell occurs before crop maturity. Weed control is needed in the first 4–6 weeks. Plants need to be supported by providing stakes or making a trellis; otherwise, the yield is seriously reduced. Fresh tender pods are harvested for vegetable when they are about 80% grown. Harvesting for seed is commenced as pods begin to dry out. This should be done periodically because if the pods are allowed to remain on the plant too long, pod splitting and seed shattering may occur. In the crop raised for tubers, pruning flowers and young pods increases the yield. Tubers are harvested by digging at the first sign of crop maturity.

Winged bean is a semidomesticated plant, and despite great potential it is a valuable source of food and feed mainly in subsistence farming. Plant breeding to evolve self-supporting growth habit, reduction in pod splitting and seed shattering, and synchronization of pod maturity will be necessary in order to develop winged bean varieties that can be grown broadly as an acreage crop.

Fresh harvested green pods keep poorly and should be marketed within 24 h. Tubers can be stored a little longer. The seed appears resistant to a number of storage insect pests, but seed viability declines quickly under tropical conditions. Seed intended for sowing should not be stored for longer than is necessary.

Diseases. A number of diseases and insect pests may limit winged bean yield. The most widespread and damaging disease appears to be false rust or orange gall (caused by *Synchytrium psophocarpi*). Dark leaf spot (caused by *Pseudocercospora psophocarpi*) and powdery mildew (caused by *Erysiphe cichoracearum*) are also important. Root knot nematodes (*Meloidoyne incognita*, *M. javanica*, and *M. arenaria*) cause galling of infected roots. The bean pod borer (*Maruca testulalis*) and other insect pests such as *Mylabris afzelli*, *M. pustulata*, *Heliothis armigera*, and *Icerya purchasi* have also been reported to cause damage to winged bean. *See* AGRICULTURAL SOIL AND CROP PRACTICES; BREEDING (PLANT); PLANT PATHOLOGY; ROSALES.

Tanveer N. Khan

Bibliography. T. N. Khan, *Winged Bean Production in the Tropics*, FAO Plant Prod. Protect. Pap. 38, Food and Agriculture Organization of the United Nations, 1982; R. J. Summerfield and E. H. Roberts (eds.), *Grain Legume Crops*, 1985.

Wire

A thread or slender rod of metal. Wire is usually circular in cross section and is flexible. If it is of such a diameter or composition that it is fairly stiff, it is termed rod. The wire may be of several small twisted or woven strands, but if used for lifting or in a structure, it is classed as cable. Wire may be used structurally in

tension, as in a suspension bridge, or as an electrical conductor, as in a power line. The working of metal into wire greatly increases its tensile strength. Thus, a cable of stranded small-diameter wires is stronger as well as more flexible than a corresponding solid rod. Wire may be treated or coated with various substances to protect it from corrosion or environmental influences. In addition, electrical conducting wire is usually covered with insulating material. *See* ELECTRICAL INSULATION; MAGNET WIRE. Frank H. Rockett

Wire drawing

The reduction of the diameter of a metal rod or wire by pulling it through a die. The working region of dies are typically conical (**Fig. 1**). The tensile stress

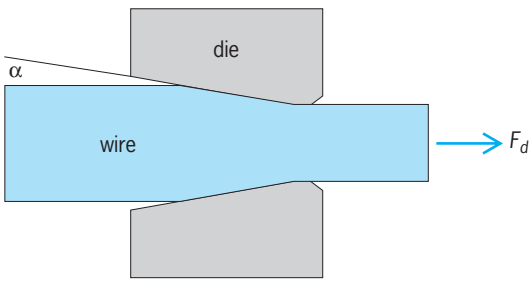

Fig. 1. Wire being drawn through a die.

on the drawn wire, that is, the drawing stress, must be less than the wire's yield strength. Otherwise the drawn section will yield and fail without pulling the undrawn wire though the die. Because of this limitation on the drawing stress, there is a maximum reduction that can be achieved in a single drawing pass. *See* STRESS AND STRAIN.

Efficiency. A simple analysis of drawing can be made with a work balance. The actual work per volume expended in drawing, w_a, can be divided into three terms: the ideal work per volume w_i, the frictional work per volume w_f, and the redundant work per volume w_r [Eq. (1)].

$$w_a = w_i + w_f + w_r \tag{1}$$

The total actual work W_a, is equal to F_dL, where F_d is the drawing force and L is the length drawn. Since the volume of the wire drawn is AL, where A is the cross-sectional area of the drawn wire, the work per volume w_a is equal to F_d/A, which is the stress on the drawn section σ_d [Eq. (2)].

$$w_a = \sigma_d \tag{2}$$

The term w_i is the work that would be required to make the same reduction in an imaginary tension test in which no necking occurs. This equals the area under the stress-strain curve. [Eq. (3)]. If work hardening is neglected, as is reasonable after a few passes, the integral $\int\sigma\ d\varepsilon$ may be approximated by $\sigma_{av}\Delta\varepsilon$, where σ_{av} is the average flow stress and $\Delta\varepsilon$ is the strain imposed during that pass.

$$w_i = \int \sigma\, d\varepsilon \tag{3}$$

The term w_f is simply the work against friction. It increases with decreasing die angle α, because the contact area between the wire and die increases with the imposed strain $\Delta\varepsilon$. For a constant coeficient of friction μ, Eq. (4) applies. If, instead of a constant friction coefficient, there is constant shear stress mk at the interface between the wire, the approximation is expressed by Eq. (5), where k is the shear strength of the wire.

$$w_f = w_i\mu \cot\alpha = \sigma_{av}\Delta\varepsilon\mu \cot\alpha \tag{4}$$

$$w_f = w_i m/(2\sin\alpha) = \sigma_{av}\Delta\varepsilon m/(2\sin\alpha) \tag{5}$$

The term w_r reflects the plastic work that is in excess of the ideal work required to produce the shape change. As a wire passes through a die, the surface layers are sheared relative to the center (**Fig. 2**). A simple upperbound model suggests that this term is independent of the strain $\Delta\varepsilon$ and depends only on flow stress and the die angle [Eq. (6)].

$$w_r = (2/3)\sigma_{av}\tan\alpha \tag{6}$$

It is often useful to define a mechanical efficiency η as in Eq. (7). Substituting Eqs. (3), (4), and (6), Eq. (8) can be written.

$$\eta = w_i/w_a \tag{7}$$

$$\begin{aligned}\eta &= \sigma_{av}\Delta\varepsilon/(\sigma_{av}\Delta\varepsilon + \sigma_{av}\Delta\varepsilon\mu\cot\alpha + (2/3)\sigma_{av}\tan\alpha)\\ &= 1/(1 + \mu\cot\alpha + (2/3)\Delta\varepsilon\tan\alpha)\end{aligned} \tag{8}$$

The variation of each of the work terms with die angle is shown schematically in **Fig. 3**. As the die angle α is increased, the redundant work increases and the frictional work decreases. The efficiency predicted by Eq. (8) is plotted in **Fig. 4** as a function of die angle for several reductions. Because Eq. (6) is an upperbound for w_r, a somewhat higher efficiency may be expected in practice. In any case, it is clear that there is an optimum die angle α for any given reduction and that α increases with increasing reduction.

The maximum reduction per pass is reached when the drawing stress σ_d equals the yield strength of the drawn wire. Except for the first few passes, work

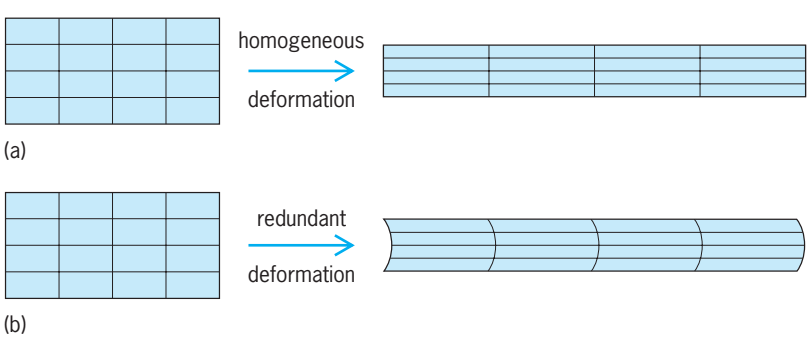

Fig. 2. Comparison of (*a*) ideal and (*b*) redundant deformation.

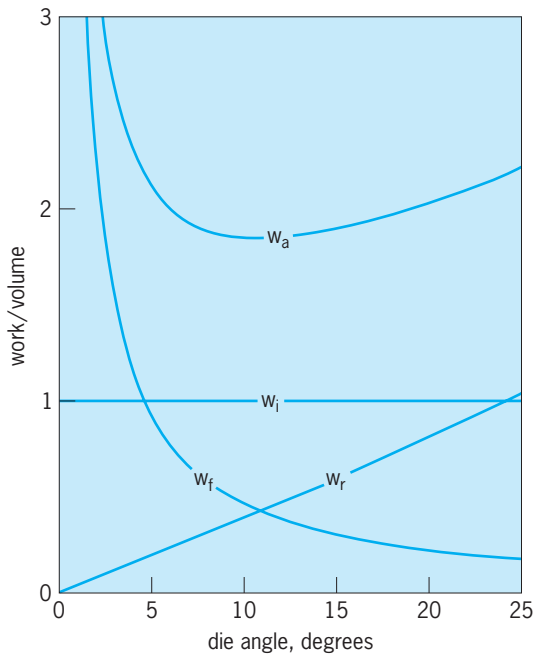

Fig. 3. Variation of the work terms with die angle according to Eqs. (1), (3), (4), and (6), with $\varepsilon = 0.3$ and $\eta = 0.08$.

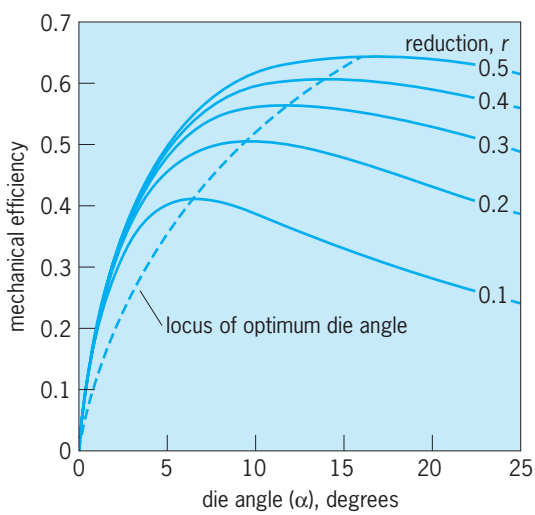

Fig. 4. Dependence of the efficiency on die angle for several reductions. A friction coefficient of $\mu = 0.08$ was assumed.

hardening can be neglected so the condition corresponds to $\sigma_d = \sigma_{av}$. Since $\sigma_d = w_a = (1/\eta)\sigma_{av}\Delta\varepsilon$, $\sigma_{av} = (1/\eta)\sigma_{av}\Delta\varepsilon$, or Eq. (9) applies. With a typical

$$\Delta\varepsilon = \eta \tag{9}$$

efficiency of 65%, $\Delta\varepsilon = 0.65$. This corresponds to an area reduction r_A as in Eq. (10), or 48%. The cor-

$$r_A = (A_o - A_f)/A_o = 1 - \exp(-\Delta\varepsilon) \tag{10}$$

responding diameter reduction $r_D = (D_o - D_f)/D_o = 1 - \exp(-\Delta\varepsilon/2)$, or 28%. In wire or rod production, multiple passes are used, without intermediate anneals, to produce fine wires. Reductions per pass are kept well below the maximum indicated by Eq. (9).

Uniformity. The homogeneity of deformation depends on the ratio Δ of the mean diameter of the wire in the deformation zone D_{av} to the contact length L between the wire and the die [Eq. (11)]. For

$$\Delta = D_{av}/L = 2\tan\alpha/(\Delta D/D_{av}) \tag{11}$$

$\Delta \leq 1$, the deformation is relatively uniform. Low die angles and high reductions per pass promote uniformity of product. However, with low reductions and high die angles, Δ may be greater than 1. Appreciable surface-to-center gradients of hardness and residual tensile stresses at the surface result when $\Delta > 1$. The inhomogeneity of deformation can be characterized by an inhomogeneity factor I.F. $= (H_s - H_c)/H_c$, where H_s and H_c are the surface and centerline hardnesses. For a material whose strain hardening is approximated by $\sigma = K\varepsilon^n$ an approximate relation is Eq. (12). **Figure 5** shows the predicted value of I.F.

$$\text{I.F.} = (1 + \tan\alpha)^n - 1 \tag{12}$$

according to Eq. (12) for drawn copper strip and wire together with experimental results. Equation (12) overpredicts I.F., but both the equation and the data indicate that the inhomogeneity increases with increasing die angle and decreasing reduction. *See* PLASTIC DEFORMATION OF METAL.

Crystallography. After large drawing reductions, wires or rods develop crystallographic textures or preferred orientations of grains. The textures are characteristic of the crystal structure of the metal. With face-centered cubic (fcc) metals, [100] and [111] directions become aligned with the wire axis. The relative amounts of the two orientations depends on the specific stacking fault energy of the metal. For aluminum and other metals with a very high stacking energy, the [111] component predominates, whereas [100] is the major component for silver. Alloys with even lower stacking fault energies have a predominant [111] component. For the latter materials, mechanical twinning is a dominant

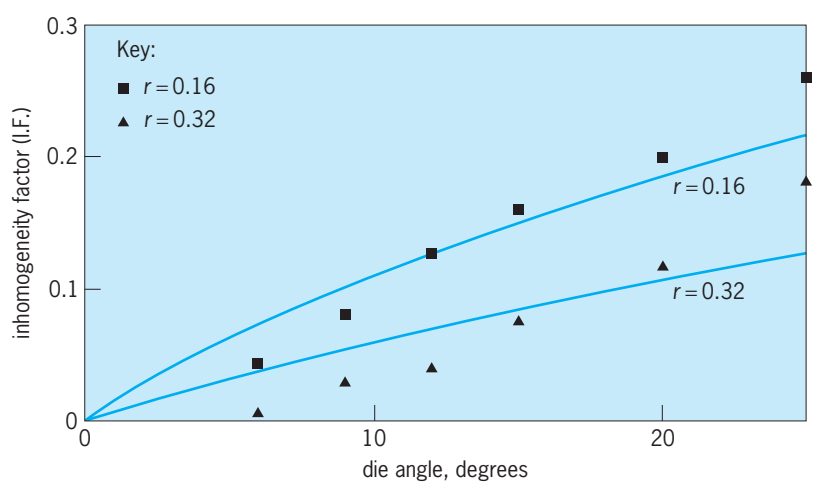

Fig. 5. Comparison of the predictions of I.F. from Eq. (12) with experiments on copper.

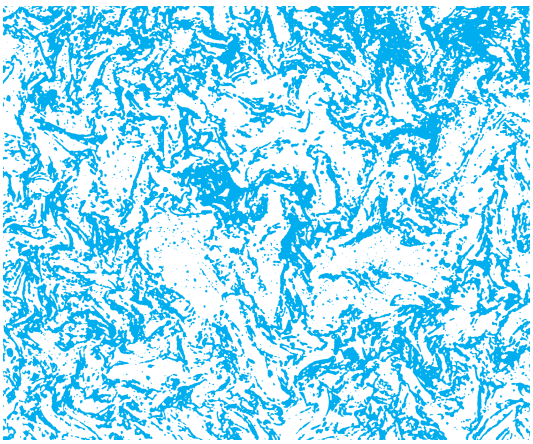

Fig. 6. Microstructure of a tungsten wire cold-drawn to an 87% reduction of area, viewed along the wire axis.

deformation mechanism. *See* ALLOY; CRYSTAL STRUCTURE; TWINNING (CRYSTALLOGRAPHY).

All body-centered cubic (bcc) metals develop a wire texture with only [110] aligned with the wire axis. Once this texture develops, the orientation of the slip systems is such that grains, as they elongate parallel to [110], thin in only the lateral [001] direction, with no strain in the lateral [110] direction. Neighboring grains must curl around each other to maintain compatibility [Eq. (4)]. This results in a characteristic microstructure (**Fig. 6**). Wire drawing of hexagonal close-packed metals forms a wire texture with the *c* axis normal to the wire axis. With this texture, grains thin in only the lateral direction normal to the *c* axis, so the microstructure formed is similar to that of bcc wires. William F. Hosford

Bibliography. W. F. Hosford, *The Mechanics of Crystals and Textured Polycrystals*, Oxford University Press, Oxford, 1993; E. M. Mielnik, *Metalworking Science and Engineering*, McGraw-Hill, New York, 1991.

Wireless fidelity (Wi-Fi)

In general, the wireless local area network (LAN) technology based on the Institute of Electrical and Electronics Engineers (IEEE) 802.11 standard. It enables computing devices to wirelessly exchange data with each other or with a wired network over a distance of up to about 300 ft (90 m), in a normal office environment, using unlicensed portions of the radio-frequency spectrum. Strictly speaking, Wi-Fi® technology refers to wireless LAN technologies that have passed interoperability tests designed by the Wi-Fi Alliance, an industry organization of vendors of 802.11 wireless LAN products. Such tests are focused on selected portions of the IEEE 802.11 standard and, sometimes, draft standards. *See* LOCAL-AREA NETWORKS.

The IEEE 802.11 standard consists of medium access control (MAC) specifications and physical layer (PHY) specifications. The MAC specifications define how a wireless LAN entity exchanges data with others using a shared wireless medium. The PHY specifications define the wireless signals that carry the exchanged data between wireless LAN entities and the wireless channels over which the wireless signals are transmitted. IEEE 802.11 was introduced in 1997 with a MAC specification and three PHY specifications. It has been amended with many extensions, including a MAC security specification 802.11i (2004) and three high-speed PHY specifications: 802.11b for 2.4 GHz (1999), 802.11a for 5 GHz (1999), and 802.11g for 2.4 GHz (2003). The IEEE 802.11 standard is still in fast evolution, as is Wi-Fi technology. *See* RADIO SPECTRUM ALLOCATIONS.

IEEE 802.11 network architecture. IEEE 802.11 defines two networking components, a station and an access point. A station has a MAC layer and a PHY layer; it is responsible for sending or receiving data to or from another station over the wireless medium, and does not relay data for any other station. An access point is a special station whose main task is to relay data for stations associated with it.

The data transmitted by the MAC layer are packed into frames. A frame consists of a header, a payload, and a tail. The header contains control information such as frame type, frame length, MAC addresses, and time window to be reserved for transmission of the frame. The payload can accommodate up to 2304 bytes of data, which may be encrypted. The tail is a 32-bit cyclical redundancy check (CRC) code that can tell a receiving MAC layer whether the received frame is corrupted.

IEEE 802.11 defines three wireless LAN architectures: independent basic service set (independent BSS), infrastructure BSS, and extended service set (ESS). An independent BSS (**Fig. 1***a*) is a wireless LAN consisting of stations only, and is usually set up on a temporary basis for a specific purpose. Every station competes for the wireless medium to send data frames to another station directly, and there is no guarantee that any two stations in the same independent BSS are in direct communication range. An infrastructure BSS (Fig. 1*b*) consists of an access point and a number of stations associated with it. The access point may be attached to a network of any kind, but most likely a wired LAN, called a distribution system (DS). Each station communicates only with the access point, and the access point relays data frames exchanged between these stations or between them and the DS. An ESS (Fig. 1*c*) consists of one or more infrastructure BSSs with all access points attached to a DS. If an ESS has two or more infrastructure BSSs, stations can utilize inter-BSS mobility by changing association from one access point to another without interrupting communications. *See* DATA COMMUNICATIONS.

IEEE 802.11 MAC. The IEEE 802.11 MAC provides three functions: fair access to the shared wireless medium with optional quality of service, reliable delivery of data frames over the wireless medium, and security protection.

Fair access. The basic medium access mechanism

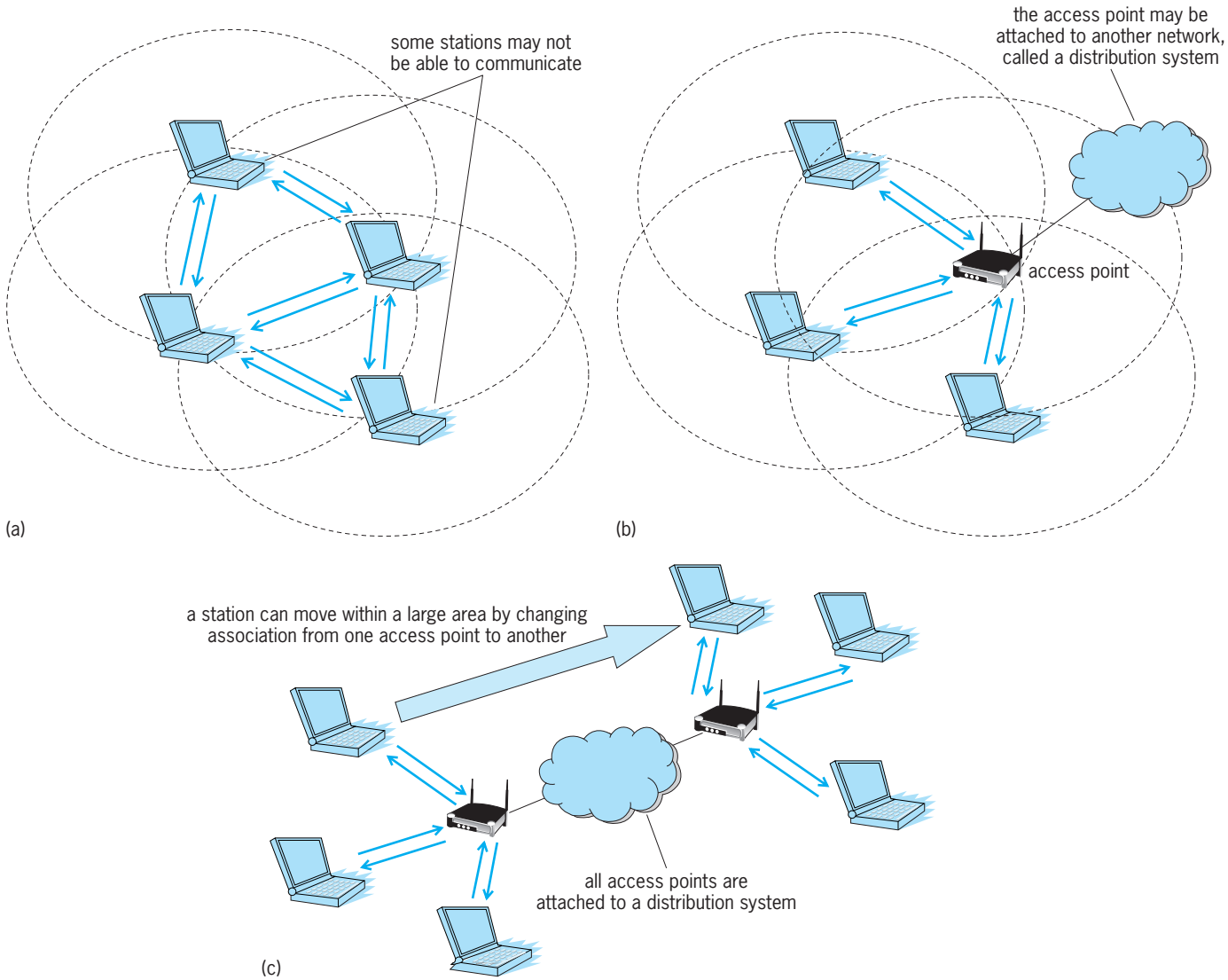

Fig. 1. Wireless LAN architectures. (*a*) Independent basic service set. (*b*) Infrastructure basic service set. (*c*) Extended service set.

employed by the IEEE 802.11 MAC is called carrier sense multiple access with collision avoidance (CSMA/CA). CSMA/CA is a wait-before-send protocol; that is, after the shared wireless medium becomes available, a station must wait a required period plus a back-off time before it can transmit a data frame, provided that the wireless medium is still available at the end of back off (**Fig. 2**). The back-off time is randomly selected by the station within a specified time period, called a contention window. CSMA/CA ensures that stations with the same contention window size and the same required waiting period have an equal probability to seize the shared wireless medium to transmit. It can reduce transmission collisions, thus achieving efficient utilization of the wireless medium, by randomizing the back-off time among station and enlarging the contention window size in case of transmission failure (which often signals congestion).

Based on CSMA/CA, the IEEE 802.11 MAC defines two medium access methods: distributed coordination function (DCF) and point coordination function (PCF). DCF assumes all stations use the same contention window range and the same required waiting period [either a DCF interframe space (DIFS) time or an extended interframe space (EIFS) time]. Since every station has an equal probability to transmit at every transmission attempt, DCF cannot support quality-of-service features such as traffic prioritization and guaranteed bandwidth.

PCF is an optional medium access method that can provide differentiated services for two classes of stations. It runs in conjunction with DCF and must be carried out by an access point. The access point maintains a list of stations that request higher quality of service. It periodically enters a contention-free period from DCF by transmitting a contention-free beacon frame, whose header announces the time window reserved for the forthcoming contention-free period. The access point transmits the beacon frame without back-off; it waits only a PCF interframe space (PIFS) time (shorter than DIFS and EIFS) after

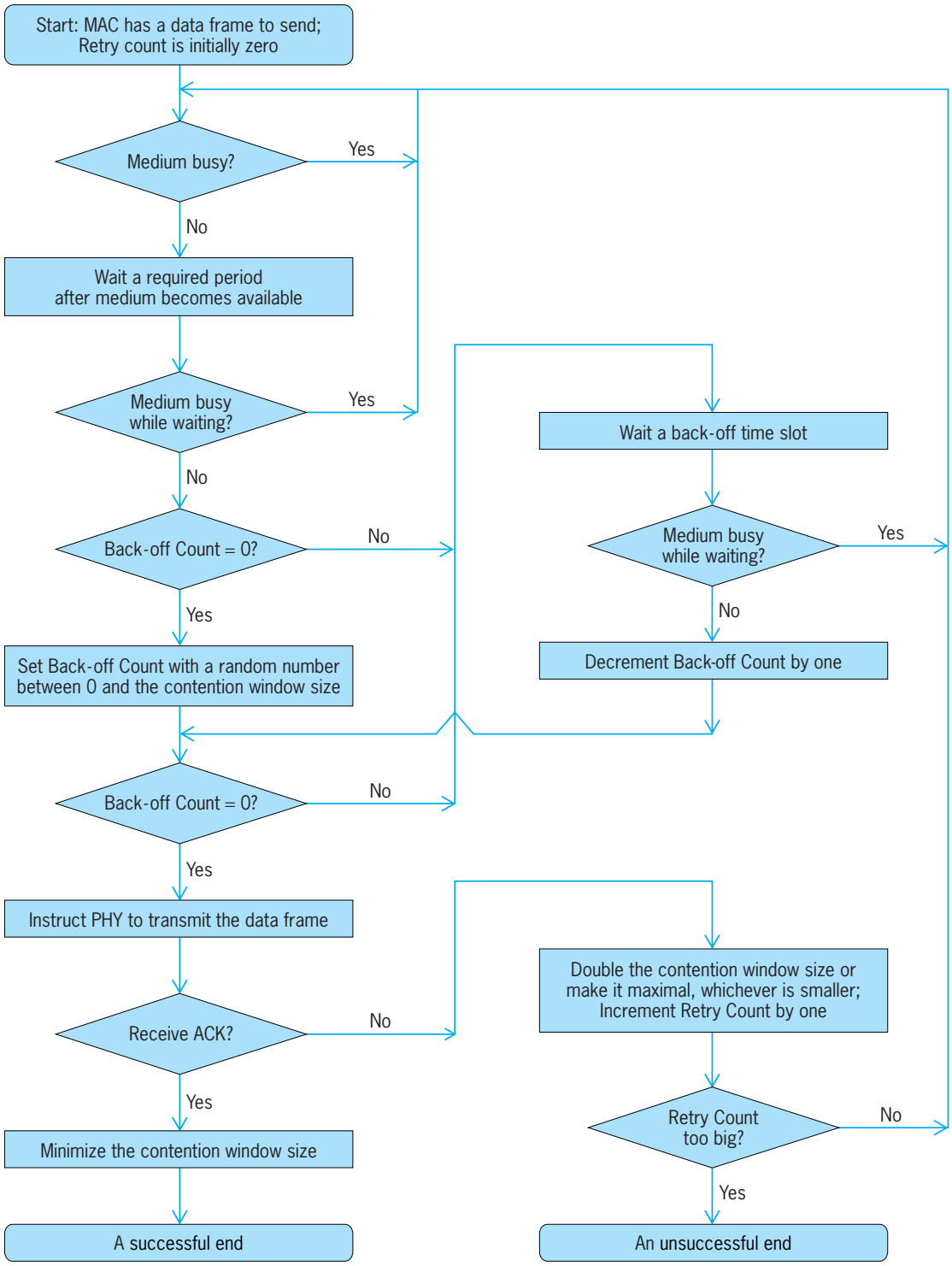

Fig. 2. Sending a data frame using carrier sense multiple access with collision avoidance (CSMA/CA).

the wireless medium becomes available. Therefore, none of the stations in the access point's radio range can compete for the wireless medium when the access point transmits the beacon frame, and they will not compete for the wireless medium during this period after receiving the beacon frame. The access point then polls stations on the list one by one and relays traffic for them with a guaranteed bandwidth and delay bound in this period.

To serve more complicated quality-of-service requirements, new medium access methods are developed in the 802.11e draft standard. In order to respond to market demand, the Wi-Fi Alliance started an interoperability test in 2004 for one of these methods, called EDCA in the 802.11e draft standard, or Wi-Fi multimedia (WMM) by the Wi-Fi Alliance. WMM provides differentiated services for four traffic classes. They are voice (highest priority),

video, best-effort data, and background data (lowest priority). WMM assigns a smaller contention window range and a shorter required waiting time to data frames with a higher priority. As a result, high-priority data frames have a greater probability to be transmitted than low-priority data frames at every transmission attempt. The bandwidth demanded by high-priority data frames can be guaranteed in a statistical sense regardless of how many low-priority data frames are waiting to be transmitted.

Reliable delivery. Since an 802.11 wireless LAN operates in unlicensed spectrum, the wireless medium could be very noisy due to the existence of other types of radio devices operating on the same band and not conforming to CSMA/CA. Thus, data frames can be frequently corrupted. In order to ensure reliable delivery of data frames over the noisy wireless medium, the IEEE 802.11 MAC employs a data frame exchange protocol for transmission of every data frame. It has a two-message form and a four-message form.

The two-message form is suitable for transmission of short data frames. It works as follows: (1) Station A sends a data frame using CSMA/CA to Station B, with the frame header announcing a reserved time window equal to the summation of the time needed to transmit the frame, a short interframe space (SIFS) time, and the time needed to transmit the expected acknowledge (ACK) frame. (2) If the data frame is correctly received, Station B sends a short ACK frame to Station A after waiting a SIFS time. (3) If Station A does not receive the ACK frame by the end of the reserved time window, it tries to resend the data frame using CSMA/CA. Note that the SIFS time is shorter than various required waiting time such as DIFS, EIFS, and PIFS time. This can prevent access points and stations that did not correctly receive the header of the data frame from interrupting the protocol.

The four-message form is designed to overcome the hidden node problem. A hidden node is a station that is outside the sending station's radio range but inside the receiving station's radio range (**Fig. 3**). Even if both the sending station and the hidden node run CSMA/CA, their data frames could still collide at the receiving station, because neither of them can detect the other's signal or time window reservation announcement in the frame header. The four-message data frame exchange protocol works as follows: (1) Station A sends a short request-to-send (RTS) frame using CSMA/CA to Station B, which tells the time window that must be reserved for the forthcoming protocol exchange. (2) Upon the reception of the RTS frame, Station B sends a short clear-to-send (CTS) frame to Station A after waiting only a SIFS time, which announces a time window reserved for the remained protocol exchanges to all stations in Station B's radio range, including hidden nodes to Station A, thus preventing them from sending during that period. (3) Upon the reception of the CTS frame, Station A sends the data frame to Station B after waiting a SIFS time. (4) If the data frame is correctly received, Station B sends a short ACK frame to Station A after waiting a SIFS time. (5) If Station A does not receive the CTS frame or the ACK frame on time, it tries to resend the RTS frame using CSMA/CA.

Security protection. Since anyone in the range of a wireless LAN can receive data frames transmitted over the wireless medium and can send data frames to stations or access points on the wireless LAN, authentication and data encryption must be implemented for access control and privacy. Initially, the IEEE 802.11 MAC defined the wired equivalent privacy (WEP) protocol to serve these needs. Unfortunately, WEP has significant security flaws. It allows authentication spoofing (falsification of a trusted identity). More seriously, an eavesdropper can break a 104-bit WEP key after collecting about 4 to 6 million encrypted data frames, which can be done in a time as short as a few hours. In order to replace WEP, the 802.11i standard has been developed. It specifies two authentication methods: preshared key (PSK) authentication (for simple wireless LAN configurations) and the IEEE 802.1x/extensible authentication protocol (EAP) authentication (for centralized access control on large wireless LANs); and two encryption methods: temporal key integrity protocol (TKIP) with message integrity check (MIC) and advanced encryption standard (AES).

TKIP uses the same stream cipher algorithm as WEP in data encryption, but corrects the security flaw in generating input to the stream cipher

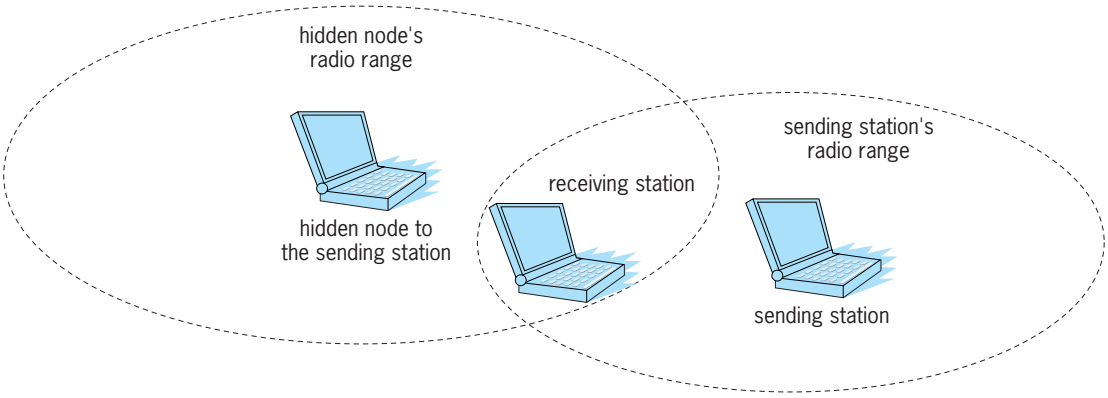

Fig. 3. Hidden node problem for a wireless network.

TABLE 1. The 2.4-GHZ frequency plan and power regulation applied to 802.11b

Channel numbers	Channel center frequencies, MHz	North America, 1000 mW (maximum output power)	ETSI (European Telecommunications Standards Institute) regulatory domain, 100 mW (effective isotropic radiation power)	Japan, 10 mW/MHz (220 mW maximum output power)
1	2412	X	X	X
2	2417	X	X	X
3	2422	X	X	X
4	2427	X	X	X
5	2432	X	X	X
6	2437	X	X	X
7	2442	X	X	X
8	2447	X	X	X
9	2452	X	X	X
10	2457	X	X	X
11	2462	X	X	X
12	2467		X	X
13	2472		X	X
14	2484			X

algorithm. It also introduces a MIC code associated with every encrypted data frame, such that the receiving station can verify authenticity of the frame. AES is a block cipher algorithm that demands significant computing power on stations, but it is considered much safer than the stream cipher algorithm used in TKIP. It combines encryption and MIC.

In the Wi-Fi Alliance's vocabulary, the wireless LAN security technology consisting of PSK authentication, the 802.1x/EAP authentication, and TKIP encryption with MIC is called Wi-Fi protected access (WPA); and that of PSK authentication, the 802.1x/EAP authentication, and the AES encryption is called WPA2. *See* COMPUTER SECURITY; CRYPTOGRAPHY.

IEEE 802.11 PHY. As of 2005 three 802.11 PHY had been put in use: 802.11b, 802.11a, and 802.11g, with 802.11b and 802.11g having a large installed base.

The 802.11b standard defines a direct sequence spread spectrum (DSSS) radio PHY for the 2.4-GHz industrial, scientific, and medical (ISM)

TABLE 2. The 5-GHz frequency plan and power regulation applied to 802.11a

		United States		CEPT (European Conference of Postal and Telecommunications Administrations) regulatory domain	
Channel numbers	Channel center frequencies, MHz	Channel allocation	Power regulation (maximum output power with up to 6 dBi antenna gain)	Channel allocation	Power regulation (effective isotropic radiation power)
36	5180	X	40 mW	X	200 mW
40	5200	X		X	
44	5220	X		X	
48	5240	X		X	
52	5260	X	200 mW	X	200 mW
56	5280	X		X	
60	5300	X		X	
64	5320	X		X	
100	5500			X	1000 mW
104	5520			X	
108	5540			X	
112	5560			X	
116	5580			X	
120	5600			X	
124	5620			X	
128	5640			X	
132	5660			X	
136	5680			X	
140	5700			X	
149	5745	X	800 mW		
153	5765	X			
157	5785	X			
161	5805	X			

band. The detailed worldwide frequency plan is shown in **Table 1**. In North America, there are 11 channels. Each channel spans 22 MHz; the center-frequency distance from that of adjacent channels is 5 MHz. Every channel partially overlaps on several adjacent channels, and thus there are effectively only three nonoverlapping channels available. The 802.11b standard specifies four data rates for the DSSS radio: 1, 2, 5.5, and 11 megabits per second (Mbps). The major interfering sources for 802.11b-based wireless LAN are microwave ovens, 2.4-GHz cordless phones, and Bluetooth® devices.

The 802.11a standard defines an orthogonal frequency division multiplexing (OFDM) radio PHY for the 5-GHz unlicensed national information infrastructure (U-NII) band. The detailed frequency plan for the United States and Europe is shown in **Table 2**. In the United States, there are 12 nonoverlapping channels available in three groups with different limits on transmitter power. Special spectrum issues such as power regulation and channel spacing for 802.11a wireless LAN operations in the 5-GHz band in Europe and Japan are specified in the 802.11h and 802.11j standards, respectively. The 802.11a standard defines eight data rates for the OFDM radio: 6, 9, 12, 18, 24, 36, 48, and 54 Mbps. Compared with 802.11b, 802.11a offers higher speed and more nonoverlapping channels but in general has a smaller operating range since 5-GHz wireless signals are attenuated faster than 2.4-GHz ones. For these reasons, 802.11a is more suitable than 802.11b for dense, high-speed installations.

The 802.11g standard defines an OFDM radio PHY that supports the same high data rates offered by 802.11a but remains compatible with the popular 802.11b. Specifically, the compatibility means that (1) every 802.11g station supports 802.11b PHY, so it can communicate with an 802.11b station using 802.11b; and (2) in an infrastructure BSS consisting of an 802.11g access point and a mixture of 802.11g and 802.11b stations, the 802.11g stations can communicate with the 802.11g access point using 802.11g for high speed, while 802.11b stations can communicate with the 802.11g access point using 802.11b, with no interference between 802.11g stations and 802.11b stations. Hui Luo

Bibliography. ANSI/IEEE Std 802.11, *Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications*, 1999; *Amendment to IEEE Std 802.11, 1999 Edition (Reaff 2003), IEEE Standard for Information technology—Telecommunications and information exchange between system—Local and metropolitan area networks—Specific requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications—Amendment 6: Medium Access Control (MAC) Security Enhancements*, 2004; IEEE Std 802.11a, *Supplement to IEEE Standard for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: High-Speed Physical Layer in the 5 GHz Band*, 1999; IEEE Std 802.11b, *Supplement to IEEE Standard for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Higher-Speed Physical Layer Extension in the 2.4 GHz Band*, 1999; IEEE Std 802.11g, *Supplement to IEEE Standard for Information technology—Telecommunications and information exchange between systems—Local and metropolitan area networks—Specific requirements—Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Further Higher Data Rate Extension in the 2.4 GHz Band*, 2003; B. O'Hara and Al Petrick, *IEEE 802.11 Handbook: A Designer's Companion,* 2d ed., IEEE Press, 2005.

Wiring

A system of electric conductors, components, and apparatus for conveying electric power from a source to the point of use. In general, electric wiring for light and power must convey energy safely and reliably with low power losses, and must deliver it to the point of use in adequate quantity at rated voltage. To accomplish this, many types of electric wiring systems and components are used.

Electric wiring systems are designed to provide a practically constant voltage to the load within the capacity limits of the system. There are a few exceptions, notably series street-lighting circuits that operate at constant current.

In the United States, the methods and materials used in the wiring of buildings are governed as to minimum requirements by the National Electrical Code, municipal ordinances, and, in a few instances, state laws. The National Electrical Code is a standard approved by the American National Standards Institute (ANSI). Most materials used in wiring systems for light and power are tested and listed by Underwriters Laboratories, Inc. (UL). *See* ELECTRICAL CODES.

The building wiring system originates at a source of electric power, conventionally the distribution lines or network of an electric utility system. Power may also be supplied from a privately owned generating plant or, for emergency supply, a standby engine-generator or battery.

The connection from the supply to the building system through the metering devices, main disconnecting means, and main overcurrent protection constitute the "service entrance." The conductors, cables or busways, are known as service conductors. The switch and fuse or circuit breaker, serving as

the disconnecting means and the main overcurrent protection, are called the service equipment. Up to six individual switches or circuit breakers may be used for the service equipment to a single building, with the exclusion of certain specifically identified or required service disconnects (such as a fire pump disconnect).

As a rule, only one set of service conductors to a building is permitted, and conductors that run in parallel are considered one set. Large industrial plants, commercial buildings, and institutions are often served from more than one source. Separate service entrances are sometimes provided for emergency lighting, fire pumps, and similar loads.

Commercial and industrial systems. Three-phase wiring systems are generally used to conform to the supply systems. Energy is transformed to the desired voltage levels by a bank of three single-phase transformers or by a single three-phase transformer. The transformers may be connected in either a delta or Y configuration. In the delta configuration the ends of the transformer windings are connected together, and line conductors are connected to these points. A three-phase, three-wire system is thus formed, from which a single-phase line can be obtained from any two conductors. In the Y configuration one end of each transformer winding is connected to a common point, and line conductors are connected to the other ends of the transformer windings. This also forms a three-phase, three-wire system. A line wire is also often connected to the common point, forming a three-phase, four-wire system, from which single-phase circuits may be obtained between the common wire and any other. In typical three-phase transformers, the windings are internally connected at the factory in either delta-delta or delta-wye configurations.

Service provided at the primary voltage of the utility distribution system, typically 13,800 or 4160 V, is termed primary service. Service provided at secondary or utilization voltage, typically 120/208 or 277/480 V, is called secondary service.

Primary service. Service at primary voltage levels is often provided for large industrial, commercial, and institutional buildings, where the higher voltage can be used to advantage for power distribution within the buildings or utility rates are less expensive when service is taken at primary voltage.

Where primary service is provided, power is distributed at primary voltage from the main switchboard through feeders to load-center substations installed at appropriate locations throughout the building (**Fig. 1**). The load-center substation consists of a high-voltage disconnect switch, a set of transformers, and a low-voltage switchboard enclosed in a heavy sheet-metal housing.

In practice, several feeder arrangements are employed. These include (1) single primary: a single primary feeder serving several substations; (2) multiprimary: individual primary feeders to each substation; (3) loop primary: two primary feeders serving several substations, interconnected to form a ring or loop; and (4) primary selective: two primary feeders serving several substations, connectable to either feeder by a selector switch.

Wiring from the low-voltage-switchboard end of the substation to the load follows conventional utilization-voltage practice.

In large industrial plants, the secondary circuits of several substations are sometimes interconnected by feeder ties through switches or circuit breakers. The switches may be closed to form a network, if the transformers are suitable for parallel operation, or operated to transfer the load from one substation to another.

Enclosed bus-bar systems, called busways, are frequently used in the wiring system of industrial plants and large buildings. Busways are made and shipped in standard lengths with a wide variety of fittings. They are connected together on the job and installed as service-entrance conductors or feeders. One type of busway is designed to receive bus plugs at intervals along its length, and thus functions both as a feeder and as a distribution board or panelboard. The plugs, which tap the feeder, may contain fuses or circuit breakers, but the tap is often carried down to a more readily accessible overcurrent protective device. *See* BUS-BAR.

Secondary service. This type of service supplies power to the building at utilization voltage. Most

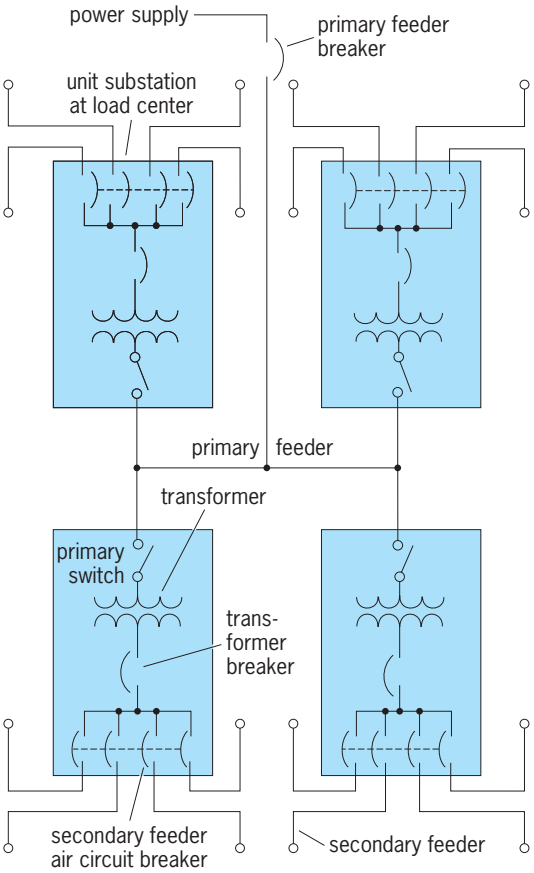

Fig. 1. Typical single primary distribution has a feeder from power supply to four load-center substations.

secondary services in the United States are 120/208 V, three-phase, four-wire, or 120/240 V, single-phase, three-wire serving both light and power. In some communities, separate light and power services are provided, typically 120/240 V, single-phase for lighting and 440 V, three-phase, delta for power.

Three-phase, four-wire services are almost universally Y configured, with the neutral tapped from the center of the Y. In some instances, a three-phase, four-wire service may be delta configured, with the neutral connected at the center tap of one phase. Such services provide typically 240 V, three-phase, three-wire for power and 120/240 V, single-phase, three-wire for lighting.

For relatively large buildings where the loads are predominantly fluorescent lighting and power (as for air conditioning), the service is often 277/480 V, three-phase, four-wire, supplying 480 V for power and 277 V, phase-to-neutral, for the lighting fixtures.

Distribution switchboards. From the service entrance, power is carried in feeders to the main switchboard, then to distribution panelboards (**Fig. 2**). Smaller feeders extend from the distribution panelboards to light and power panelboards. Branch circuits then carry power to the outlets serving the various lighting fixtures, plug receptacles, motors, or other utilization equipment. *See* BRANCH CIRCUIT.

The main distribution switchboard may also include the service equipment in its assembly. It consists of a group of switches and fuses or circuit breakers in a sheet-metal enclosure. It provides individual disconnecting and overcurrent protection for each feeder. Such equipment typically employs barriers between the service and distribution portions of the switchboard. In large buildings, additional distribution panelboards may be located at load centers.

Light and power panelboards provide individual disconnecting and overcurrent protection for the branch circuits. The circuit breakers of lighting panelboards are sometimes used as switches to operate the lighting circuits, when identified for such use.

Plug-receptacle power at 120/240 V, such as power for small appliances and business machines, in buildings provided with 277/480 V supply, is obtained from transformers. A feeder circuit from a 277/480 V distribution board or panelboard energizes the transformer primary. The secondary feeder serves a separate 120/240 V panelboard. The branch circuits serving the plug-receptacle outlets are conventional.

Wiring methods. Methods of wiring in common use for light and power circuits are as follows: (1) insulated wires and cables in metallic or nonmetallic raceways; (2) nonmetallic sheathed cables; (3) metallic armored cables; (4) busways; (5) copper-jacketed, mineral-insulated cables; (6) aluminum-sheathed cables; (7) nonmetallic sheathed and armored cables in cable support systems; and (8) open insulated wiring on solid insulators (knob and tube). *See* CONDUCTOR (ELECTRICITY).

Raceways in which insulated conductors may be installed are of several types: (1) rigid metal conduit; (2) electric metallic and nonmetallic tubing (EMT & ENT); (3) flexible metal and nonmetallic conduit; (4) liquid-tight flexible metal and nonmetallic conduit; (5) surface metal and nonmetallic raceway; (6) underfloor raceway; (7) cellular floor raceway; (8) rigid nonmetallic conduit; and (9) wireway.

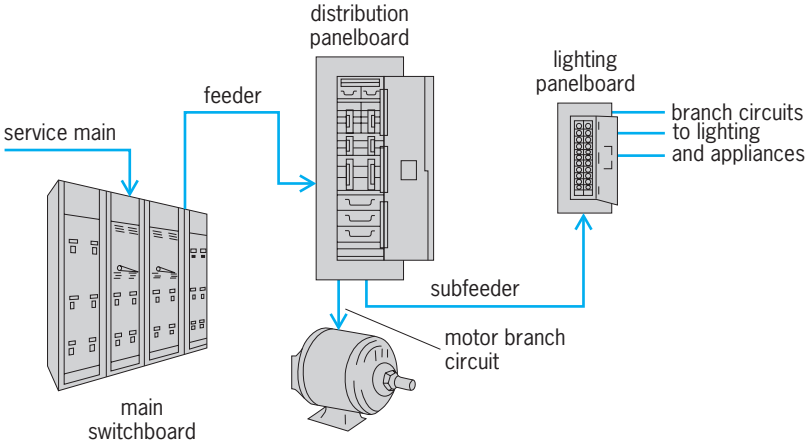

Fig. 2. Distribution wiring system has components to supply utilization voltage for several different types of load.

Insulated conductors may also be run in "cable tray," which is not considered a raceway but a cable support system.

The selection of the wiring method or methods is governed by a variety of considerations, which usually include code rules limiting the use of certain types of wiring materials; suitability for structural and environmental conditions; installation (exposed or concealed); accessibility for changes and alterations; and costs. Several methods may be employed together, for example, feeder busway risers in a multistory office building with the rest of the wiring in rigid conduit and underfloor raceways.

Circuit design. The design of a particular wiring system is developed by considering the various loads, establishing the branch-circuit and feeder requirements, and then determining the service-entrance requirements. Outlets for lighting fixtures, motors, portable appliances, and other utilization devices are indicated on the building plans, and the load requirement of each outlet is noted in watts or horsepower. Lighting fixtures and plug receptacles are then grouped on branch circuits and connections to the lighting panelboard indicated. Lighting and power panelboards are preferably located in the approximate center of the loads they serve; however, other considerations may require other locations. Panelboards in commercial and institutional buildings usually are located in corridors or electric closets. The size and number of panelboards are determined by the number of branch circuits to be served. *See* ELECTRIC POWER SYSTEMS.

Conductor sizes. The size of wires and cables used in electrical wiring systems is expressed in terms of the American Wire Gage (AWG), known also as the Brown and Sharpe (B&S) gage. Size designations run from No. 14, the smallest size commonly used in

wiring systems for light and power, to No. 4/0, the largest size in the gage. Sizes larger than No. 4/0 are designated by their cross-section areas expressed in circular mils. The largest size in practical usage is 2,000,000 circular mils (1013 mm^2). A circular mil is the area of a circle 0.001 in. (25.4 micrometers) in diameter.

Conductor capacity. The current-carrying capacity of wiring conductors is determined by the maximum insulation temperature that can be tolerated and the rate at which heat can be dissipated. All conductors offer some resistance to the flow of electric current. Consequently, heat is produced in the conductor by the flow of current I through its resistance R. The amount of heat is determined by the square of the current in amperes times the resistance in ohms (I^2R). Conductor heat is dispersed through the insulation and the surrounding raceway, and cable sheath or enclosure to the air. *See* ELECTRICAL INSULATION.

In practice, maximum current-carrying capacity of conductors is set forth in standard tables developed from laboratory tests and field experience. The National Electrical Code specifies the maximum current-carrying capacity of conductors. For any given size of conductor, the maximum capacity varies with the type of installation (in air or in raceways) and the maximum safe temperature of the insulation. Approved values are reduced for high ambient temperatures (greater than 86°F or 30°C) and for more than three current-carrying conductors in a single raceway or cable.

Feeders supplying several motors must be rated at not less than 125% of the full-load current of the largest motor plus 100% of the full-load currents of the remaining motors.

Feeders serving continuous loads that are likely to operate for 3 h or more (as office-building lighting) should not be loaded to more than 80% of rated capacity or have a capacity equal to 125% of the continuous load. The size of the feeder conductors will often be determined by the permissible voltage drop, which may require larger conductors than would be required by current-carrying capacity considerations alone.

Lighting loads. Lighting branch circuits may be loaded to 80% of circuit capacity. However, there is a reasonable probability that the lighting equipment will be added or replaced at some future time by equipment of higher output and greater load. Therefore, in modern practice, lighting branch circuits are loaded only to about 50% capacity, typically not more than 1200 W on a 20-A branch circuit.

Lighting branch circuits are usually rated at 20 A. Smaller 15-A branch circuits are used mostly in residences. Larger 30-, 40-, and 50-A branch circuits are limited to serving heavy-duty lampholders or appliances specially approved for connection to such circuits. The minimum conductor size for 20-A branch circuits is No. 12; however, on long runs larger conductors may be required to avoid excessive voltage drop. A common practice is to use No. 12 conductors between outlets and No. 10 conductors for the connection between the first outlet and the lighting panelboard.

Motor loads. These power loads are usually served by individual branch circuits, which must be rated at not less than 125% of the full-load rating of the motor.

Feeder and service-entrance design. The sum of the branch-circuit loads, including additional capacity for future load, determines the feeder load. The National Electrical Code provides a table of factors for various occupancies giving the minimum unit load per square foot and demand factors that may be applied. Demand factors are applied in installations where the loads are diversified and not likely to occur at one time. The number of feeders and their loads determine the number and size of distribution panelboard circuit elements required. The sum of the feeder loads determines the size and capacity of the service-entrance conductors and equipment.

Voltage drop. A drop in voltage along a conductor is a characteristic of all electric circuits. The voltage drop in a circuit causes the voltage at the load to be less than that applied to the circuit.

Wire or cable circuits are designed to carry a certain load but, whether the conductor be copper, aluminum, or other metal, the resistance characteristics impede the flow of current.

Since conductor resistance is proportional to conductor length, longer circuits are especially susceptible to excessive voltage drops. The amount of voltage drop E_D is determined by the current I in amperes times the resistance R in ohms ($E_D = IR$).

Good practice in circuit design dictates the following percentage values for maximum voltage drop: 5% from service entrance to any panelboard; 2% from panelboard to any outlet on branch circuit; 4% in feeders and 1% in branch circuit to motor; and 2% total in conductors to electric heating equipment.

The feeder voltage drop is calculated by formulas derived from Ohm's law and the resistance of the conductor. For three-wire, three-phase circuits (neglecting inductance), the voltage drop E_D is given in the equation $E_D = 1.732KIL/CM$, in which K is the resistance of a circular mil-foot of wire (for copper, 10.8 ohms), I is the current in amperes, L is the circuit length (source to load) in feet, and CM is the cross-section area of the conductor in circular mils.

Copper loss. This characteristic of a circuit is related to voltage drop. It is a power loss, designated as the I^2R loss, and is often expressed in percentage as the ratio of the wattage loss to the wattage delivered to the circuit.

Circuit protection. In wiring systems of high capacity (typically 1200 A and above) supplied from utility networks of large capability, overcurrent protective devices of high interrupting capacity are required. Circuit breakers of special design or current-limiting fuses are employed in such installations. In some cases, current-limiting reactors or bus-bar arrangements presenting appreciable reactance under short-circuit conditions are inserted in the service conductors. *See* CIRCUIT BREAKER; ELECTRIC

PROTECTIVE DEVICES; FUSE (ELECTRICITY).

Residential systems. For any house, the wiring system is an installed hookup of electrical conductors and components that provide for carrying electrical energy from the utility company lines to lighting fixtures, appliances, and receptacles. Every residential wiring system must conform in design and installation methods to the regulations of the National Electrical Code and any other codes that apply in a particular locality.

Every residential electrical system can be broken down into four basic categories: service entrance equipment; circuit wiring and raceways; outlet boxes and wiring devices; and fittings, connectors, and accessories.

Service entrance. As discussed above, the part of the wiring system that connects directly to the utility supply line is referred to as the service entrance. Depending upon the type of utility line serving the house, there are two basic types: overhead and underground. Service-entrance conductors are typically assemblies that consist of two or more insulated conductors combined with one bare conductor that is used as the grounded neutral, all encased in an outer nonmetallic jacket.

In the vast majority of cases, these conductors feed the service panelboard, which is an enclosure that is designed to simplify residential service-entrance installations carrying currents up to 200 A for 120/240 V, single-phase, three-wire systems. The service panelboard typically uses circuit breakers, although fuses, or a combination of fuses and circuit breakers, may be used in older installations to provide control and protection for the incoming service-entrance conductors and the branch-circuit wires, which are carried throughout the house to supply the lighting, appliances, and receptacles.

Circuit wiring and raceways. Branch circuits are run typically from the service panelboard in a residence to the various electrical outlets that must be supplied. Circuits to plug receptacle outlets and lighting fixture outlets generally are rated at 15 or 20 A and utilize two wires: one hot conductor and one neutral conductor. Circuits rated above 30 A are used only for supplying individual appliances, such as air conditioners, electric ranges, stoves, heaters, and clothes washers and dryers, and may include two hot conductors and a neutral.

A number of different types of cable are available for use in running branch circuits in the house and underground outdoors for lighting and receptacle circuits around yards, patios, and similar areas. The three most commonly used types are nonmetallic-sheathed cable (type NM), armored cable (type AC), and underground feeder cable (type UF).

Nonmetallic sheathed cable (type NM) is an assembly of two or more insulated conductors and usually a bare or an insulated ground conductor. For use in conformity with the National Electrical Code, any such cable assembly must be an approved type NM cable in sizes No. 14 to No. 2 American Wire Gage with copper conductors, and in sizes No. 12 to No. 2 with aluminum conductors. A corrosion-resistant type of nonmetallic sheathed cable (type NMC) is also available for use in damp or corrosive locations. Such conductors may be used in residential buildings of any construction type or height.

Armored cable (type AC) is a fabricated assembly of two or more insulated conductors in an interlocked, flexible metal jacket with a bare, internal, bonding conductor in intimate contact with the armor for its entire length.

Underground feeder cable is a factory-fabricated cable of one or more insulated conductors contained within an overall nonmetallic sheath that is suitable for direct burial. Typical applications include primarily lighting, as well as power loads other than swimming pools, hot tubs, fountains, and the like.

Outlet boxes and wiring devices. A wide variety of small boxes are employed at various points in the wiring system where electrical power is supplied to lighting fixtures, motors, fixed appliances, and cord-connected devices and appliances. Such boxes are available in both metallic and nonmetallic types. Metallic boxes are used in all wiring systems, while the nonmetallic types are generally used with type NM cable systems.

Wiring boxes are commonly divided into several categories according to construction and application, as follows.

Octagonal and round outlet boxes are typically used for ceiling outlets.

Oblong and square boxes are commonly referred to as gang boxes because they permit the mounting of a number of standard wiring devices in a single box. A four-gang box, for example, will accommodate four standard wiring devices mounted side by side. Such boxes usually employ one of a variety of covers.

Sectional switch boxes have removable sides to allow adding boxes together, or ganging of boxes, to accommodate installation of several wiring devices side by side at one location. These boxes do not require covers; they accommodate standard flush wiring devices with standard flush wall plates.

Utility boxes are all-purpose boxes designed for surface mounting.

Wiring devices encompass the wide range of types and sizes of convenience receptacles, wall switches, cord plugs, and connectors.

Convenience receptacles are made in many types and sizes to provide for the connection of supply cords from lighting units, motor-operated devices, and other appliances. Such devices are classified according to the number and arrangement of the contact slots and the type of mounting, which are based on their voltage and current rating. *See* ELECTRICAL CONNECTOR.

Standard receptacles are made in single and duplex (two receptacle outlets on a single yoke) types for use in a single gang box. Interchangeable-type receptacles permit different combinations of receptacles, switches, and pilot lights to be assembled as desired. Up to four devices may be installed in a single gang

box on a single strap. Combination devices are similar to the interchangeable type in that a combination of switches, receptacles, and pilot lights, if desired, can be installed as a single-device gang box. However, combination devices are not field-assembled, but are manufactured in the various configurations and, as with other standard switches and receptacle, are enclosed by molded bodies. Additionally, many so-called pilot devices have been replaced with integral light-emitting diode (LED) indicators.

Receptacle outlets installed on 15- and 20-A branch circuits must be of the grounding type to ensure compliance with the National Electrical Code. Replacements for non-grounding-type receptacles in existing residences must be nongrounding receptacles or ground-fault circuit-interrupter-type (GFCI) receptacles, or nongrounding types provided with upstream GFCI protection. Additionally, ground-fault protection must be provided for receptacles installed in the bathroom, the kitchen countertop, outdoor locations, and where they are readily accessible or not dedicated to specific appliances in attached or detached garages and basements. The required ground-fault protection may be achieved with a Class A ground-fault circuit-interrupter breakers or receptacles. *See* GROUNDING.

Wiring devices also include a range of small switches. Such switches are used in nearly every type of interior wiring system to control branch-circuit equipment (lighting units or appliances). Generally, such switches operate on a snap-action principle, with a toggle handle, push button, or rotary handle operating a contact mechanism that gives definite on and off positions.

Fittings, connectors, and accessories. In electrical work, a wide variety of small hardware devices are used to provide complete installations of conductors in conduit or cable systems. These devices include mechanical connecting and fastening items that complete the raceway system, and items that provide for the connection of current-carrying conductors.

Joseph F. McPartland; Brian J. McPartland

Bibliography. T. Croft and W. I. Summers (eds.), *American Electricians Handbook*, 14th ed., 2002; D. G. Fink and H. W. Beaty (eds.), *Standard Handbook for Electrical Engineers*, 14th ed., 2000; Institute of Electrical and Electronics Engineers, *The Recommended Practice for Electrical Power Systems in Commercial Buildings*, Std. 241, 1990; J. F. McPartland, *National Electrical Code Handbook*, 25th ed., 2005; J. F. McPartland and B. J. McPartland, *Handbook of Practical Electrical Design*, 3d ed., 1999; National Fire Protection Association, *National Electrical Code 2005*, 2005; H. P. Richter and W. C. Schwan, *Practical Electrical Wiring*, 19th ed., 2005.

Wiring diagram

A drawing illustrating electrical and mechanical relationships between parts on a component that require interconnection by electrical wiring.

A wiring diagram is distingushed from an electrical schematic in that the arrangement of the schematic bears no necessary relationship to the mechanical arrangement of the electrical elements in the component. The wiring diagram provides an accurate picture of how the wiring on the elements and between them should appear in order that the electrical wiring technician can install the wiring in the manner that will best contribute to the performance of the device.

The degree of symbolism used in a wiring diagram depends on the extent of standardization in the particular field. For example, in telephone switchboard wiring, which consists of many standardized repetitive operations, extensive symbolism is used. When the exact physical location of wiring is important, as in radio-frequency devices where electromagnetic and electrostatic coupling between wires is appreciable, the diagram can be quite pictorial.

Wiring diagrams also include such information as type of wire, color coding, methods of wire termination, and methods of wire and cable clamping. *See* SCHEMATIC DRAWING.

Robert W. Mann

Witherite

The mineral form of barium carbonate. Witherite has orthorhombic symmetry and the aragonite structure type. Crystals, often twinned, may appear hexagonal in outline (see **illus.**). It may be white or gray with yellow, brown, or green tints. Its hardness is 3.5 and its specific gravity 4.3.

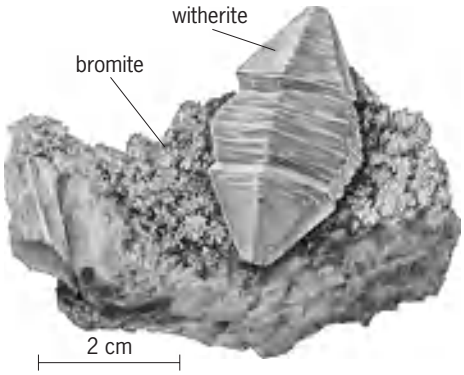

Large crystal with bromite from Fallowfield, Northumberland, England. (*Specimen from Department of Geology, Bryn Mawr College*)

Witherite may be found in veins with barite and galena. It is found in many places in Europe, and large crystals occur at Rosiclare, Illinois. *See* BARIUM; CARBONATE MINERALS.

Robert I. Harker

Wolf-Rayet star

A type of hot, luminous star that is distinguished by its extremely dense and fast wind. The spectacularly bright, discrete bands of atomic emission from these winds greatly facilitated their discovery with the

aid of a visual spectroscope by the French astronomers Charles Wolf and Georges Rayet at the Paris Observatory in 1867. *See* ASTRONOMICAL SPECTROSCOPY.

The Wolf-Rayet phenomenon includes a typical phase in the advanced evolution of a massive star (about 20–100 times the Sun's mass at birth), or sometimes (15% of cases) among lower-mass stars in the planetary nebula stage. Current evidence suggests that most massive Wolf-Rayet stars are the compact, hot cores remaining after most of the initial hydrogen-rich material in the central region of a massive star has been fused via nuclear burning to heavier elements. The surrounding unburnt, hydrogen-rich layers have been mostly removed by the efficient stellar wind. In a few rare cases, the Wolf-Rayet phenomenon is seen in the most massive, thus luminous, stars on the main sequence, even before central hydrogen-burning is completed (**Fig. 1**). *See* PLANETARY NEBULA.

The first fusion process in the core of a massive star involves the conversion of hydrogen into helium via the carbon-nitrogen-oxygen (CNO) cycle, in which helium and nitrogen are enhanced at the expense of the initially abundant hydrogen and traces of carbon and oxygen. When such fusion products are visible in the winds, a WN-type Wolf-Rayet star is seen (**Fig. 2**), whose spectrum is dominated by Doppler-broadened atomic lines of helium and nitrogen in various stages of ionization. Later, when the second fusion process occurs, helium is converted mainly into carbon and oxygen, with nitrogen being virtually destroyed. A WC-type Wolf-Rayet star is then seen with lines mainly of carbon and helium. A brief oxygen-rich phase may occur (WO) toward the end of the helium-burning phase, after which all subsequent exotic nuclear-burning phases are so rapid that they remain hidden by the much more slowly changing stellar surface. At that point, it is believed that the Wolf-Rayet star will explode as a supernova, resulting in the collapse to a black hole in most cases and an associated gamma-ray burst in very rare instances. *See* BLACK HOLE; DOPPLER EFFECT; GAMMA-RAY BURSTS; NUCLEOSYNTHESIS; SUPERNOVA.

Beneath the dense winds that often hide the stellar surface, massive Wolf-Rayet stars have surface temperatures ranging from 30,000 to 150,000 K (54,000 to 270,000°F), radii of 1 to 15 solar units, and luminosities of 10^5 to 10^6 times that of the Sun. They are losing matter at a rate that is typically 10^9 times that of the Sun's wind, at speeds ranging from 1000 to 3000 km/s (600 to 1800 mi/s), clearly exceeding the minimum speed required to escape from the star. *See* SOLAR WIND.

Although massive Wolf-Rayet stars appear to be rare (only about 250 are known so far in the Milky Way Galaxy, out of an estimated total population of 1000–2000), all massive stars likely pass through a Wolf-Rayet stage toward the end of their relatively short lives. The Wolf-Rayet phase lasts at most about 10% of the total lifetime of a massive star, the latter being some 2–6 million years. *See* STAR; STELLAR EVOLUTION.

Anthony F. J. Moffat

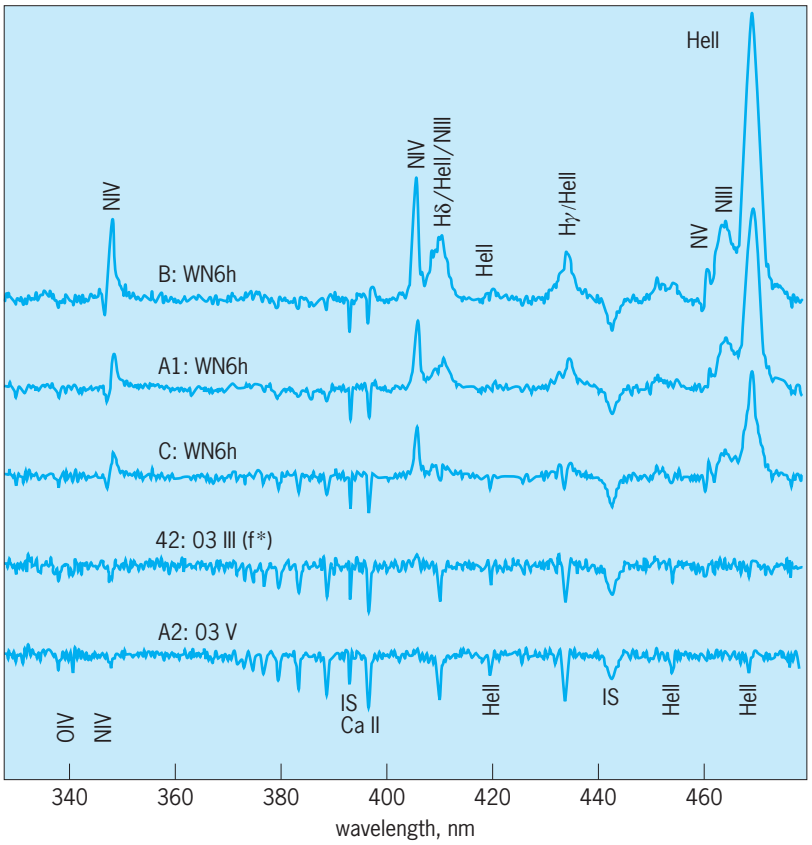

Fig. 1. Brightness-sorted montage of Hubble Space Telescope spectra for the 5 most luminous stars in the core of the extremely dense, young galactic cluster NGC 3603, showing its three hydrogen-rich, hydrogen-burning Wolf-Rayet stars, A1, B, and C. These three stars display strong Wolf-Rayet–like emission lines compared to the other normal, fainter O-type stars. (*From L. Drissen, The stellar content of NGC 3603, in K. A. van der Hucht, G. Koenigsberger, and P. R. J. Eenens, eds., Wolf-Rayet Phenomena in Massive Stars and Starburst Galaxies, International Astronomical Union Symposium 193, p. 403, Astronomical Society of the Pacific, 1999*)

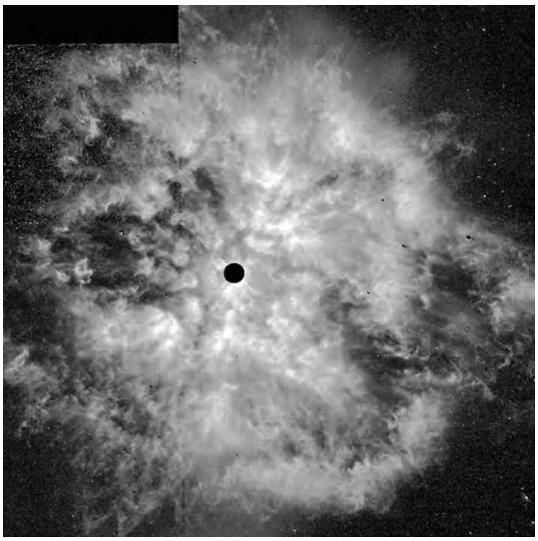

Fig. 2. Hubble Space Telescope image in the light of hydrogen-alpha of the Wolf-Rayet star WR 124 (blotted out in the center) and its surroundings. This is a particularly interesting fast-moving, hyrdogen-poor WN star, whose fast wind is catching up and strongly interacting with the interstellar medium and the slow wind of its immediate progenitor, a so-called luminous blue variable star or a red supergiant. (*From Y. Grosdidier et al., HST WFPC2/Hα imagery of the nebula M1-67: A clumpy LBV wind imprinting itself on the nebular structure?, Astrophys. J., 506:L127–L131, 1998*)

Bibliography. H. J. G. L. M. Lamers and J. P. Cassinelli, *Introduction to Stellar Winds*, Cambridge University Press, 1999; A. Maeder and P. S. Conti, Massive star populations in nearby galaxies, in G. Burbidge (ed.), *Annu. Rev. Astron. Astrophys.*, 32:227–275, 1994; K. A. van der Hucht, The VIIth Catalogue of galactic Wolf-Rayet stars, *New Astron. Rev.*, 45(3):135-232, 2001; K. A. van der Hucht, A. Herrero, and C. Esteban (eds.), *A Massive Star Odyssey: From Main Sequence to Supernova, International Astronomical Union Symposium 212*, Astronomical Society of the Pacific, 2003; K. A. van der Hucht, G. Koenigsberger, and P. R. J. Eenens (eds.), *Wolf-Rayet Phenomena in Massive Stars and Starburst Galaxies, International Astronomical Union Symposium 193*, Astronomical Society of the Pacific, 1999; J. M. Vreux et al. (eds.), *Wolf-Rayet Stars in the Framework of Stellar Evolution, 33d Liege International Astrophysical Colloquium*, 1996.

Wolframite

A mineral with chemical composition $(Fe,Mn)WO_4$, intermediate between ferberite, the iron tungstate, and huebnerite, the manganese tungstate, which form a complete solid solution series. *See* HUEBNERITE.

Wolframite occurs commonly in short, brownish-black, monoclinic, prismatic, bladed crystals (see **illus.**). It is difficult to dissolve in acids, but wolframite high in iron fuses readily to a magnetic globule.

Wolframite is probably the most important tungsten mineral. A quick and easy test for tungsten is to fuse the mineral powder with charcoal and sodium carbonate and boil the residue in hydrochloric acid with a few grains of granulated metallic tin. The presence of tungsten gives the solution a prussian-blue color. The major use of tungsten is in making ferrous (steel) alloys, nonferrous alloys, tungsten carbide, metallic tungsten, and tungsten chemicals.

Wolframite is found associated with quartz in veins in the peripheral areas of granitic bodies. It is also found in veins associated with sulfide minerals such as pyrite, chalcopyrite, arsenopyrite, and bismuthinite, together with cassiterite, molybdenite, hematite, magnetite, tourmaline, and apatite. It also occurs in placers.

China is the major producer of wolframite. Tungsten minerals of the wolframite series occur in many areas of the western United States; the major producing district is Boulder and northern Gilpin counties in Colorado. *See* TUNGSTEN.

Edward C. T. Chao

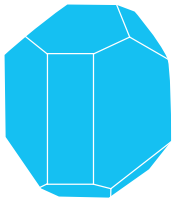

Wolframite crystal habit. (*After C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*)

Wollastonite

A mineral inosilicate with composition $CaSiO_3$. It crystallizes in the triclinic system in tabular crystals (see **illus.**). More commonly it is massive, or in cleavable to fibrous aggregates. There are two good cleavages parallel to the front and basal pinacoids yielding elongated cleavage fragments. Hardness is 5–5$^1/_2$ on Mohs scale; specific gravity is 2.85. On the cleavages the luster is pearly or silky; the color is white to gray. Wollastonite is the most common of three polymorphic forms of $CaSiO_3$, the other two being pseudowollastonite and parawollastonite. Pseudowollastonite, a high-temperature triclinic form, is very uncommon in rocks but may be a constituent of synthetic CaO-SiO_2 systems and of slags and glasses. Parawollastonite, a monoclinic form, is only rarely found in Ca-rich rocks. Wollastonite, by far the most common polymorph, occurs abundantly in impure limestones that have undergone contact metamorphism. Resulting assemblages may consist of calcite-diopside-wollastonite with variable amounts of tremolite, clinozoisite, and grossularite. Wollastonite occurs sporadically in regionally metamorphosed calcareous sediments as well. It is found in large masses in the Black Forest of Germany; Brittany, France; Chiapas, Mexico; and Willsboro, New York, where it is mined as a ceramic material. *See* SILICATE MINERALS.

Cornelius S. Hurlbut, Jr.

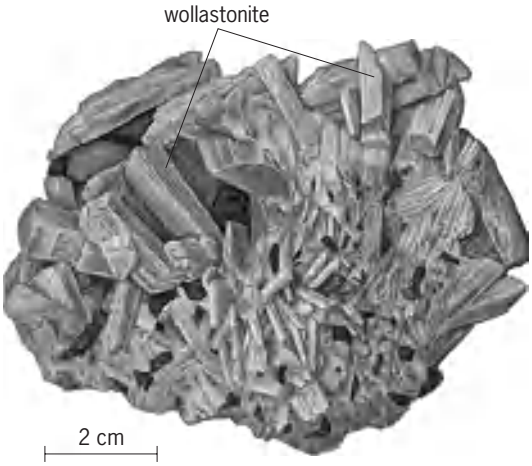

Triclinic tabular crystals embedded in limestone, Santa Fe, Chiapas, Mexico. (*Specimen from Department of Geology, Bryn Mawr College*)

Wolverine

The largest member of the family Mustelidae, to which weasels, mink, fishers, martens, ferrets, badgers, and otters belong.

Morphology. The powerfully built wolverine (*Gulo gulo*) has an overall shaggy appearance (see **illustration**). The heavy body is covered with long guard hairs that overlie a coarse, dense, kinky, woolly underfur. The general coloration is blackish brown. A light brown band extends along each side of the body from shoulder to rump before merging at the base of the tail. Most wolverines have a light silver facial mask. The large head is broad and somewhat rounded with small, wide-set eyes and short, rounded, well-furred ears. The legs, feet, and most of the relatively short, shaggy tail are dark. The legs are massive, stocky, and powerful, and the feet are large. The head and tail are carried lower than the somewhat arched back, and the lumbering plantigrade gait gives the impression of clumsiness. The palms and soles are densely haired. The claws are pale and sharply curved. A pair of anal scent glands are present which secrete a yellowish-brown fluid. The dental formula is I 3/3, C 1/1, PM 4/4, M 1/2 × 2 = 38 teeth. Adults have a head and body length of 650–1050 mm (25–41 in.), a tail length of 170–260 mm (6.5–10 in.), and weigh 7–32 kg (15–70 lb). Males are generally 25–30% heavier and 10% greater in most body measurements than females of similar age. *See* DENTITION.

Distribution and ecology. The wolverine has a circumpolar distribution, occurring in the tundra and taiga zones throughout northern Europe, Asia, and North America. It is uncommon in the United States and most of Canada and is very rare in Scandinavia. It is active all year and may be found in forests, mountains, or open plains. In mountainous areas, it moves to lower elevations during winter. Nests of grass and leaves are constructed inside the den, which may be in a cave or rock crevice, beneath a fallen tree, or in a burrow made by another animal. Wolverines are primarily terrestrial, but they can also climb trees and are excellent swimmers. These mammals have a keen sense of smell, but they lack keen senses of vision and hearing. Wolverines are solitary wanderers except during the breeding season. They occur at relatively low population densities and are territorial. They have enormous home ranges; average yearly home ranges in Montana are 422 km^2 for males and 388 km^2 for females. In south-central Alaska, the average home range is 535 km^2 for males and 105 km^2 for females with young.

Mating and reproduction. Although mating occurs from April to July, wolverines exhibit delayed implantation—that is, implantation of the fertilized eggs does not occur until November to March. Although total gestation may last from 215 to 272 days, active gestation (after implantation) is only 30 to 40 days. Births occur from January to April. Females produce litters of two to four about every 2 years. Young wolverines nurse for 8–10 weeks, leave their mother in the autumn, and reach sexual maturity in their second or third year of life. Longevity in the wild is approximately 8 to 10 years; however, it is possible for captive wolverines to live more than 17 years.

Dietary habits. Food consists of carrion (reindeer, red deer, moose, caribou), eggs of ground-nesting birds, lemmings, and berries. The tenacity and strength of the wolverine reportedly allow it to drive bears and cougars from their kills. Wolverines will attack large mammals when they are sick or trapped in deep snow; their large feet act like snowshoes, spreading the animal's weight so that it can run over soft, deep snow that slows down its heavy, hoofed prey.

Wolverine (*Gulo gulo*) in Kalispell, Montana. (*Photo by Gerald and Bliff Corsi, © 1999 California Academy of Sciences*)

Status. Humans, through direct persecution, deforestation, encroachment, and urbanization, are the only important enemy of the wolverine. It has been classified as vulnerable by the International Union for the Conservation of Nature (IUCN) and has been designated as endangered in eastern Canada. *See* BADGER; CARNIVORA; FERRET; FISHER; MARTEN; MINK; OTTER; WEASEL.

Donald W. Linzey

Bibliography. D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford Limited, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Wood anatomy

Wood is composed mostly of hollow, elongated, spindle-shaped cells that are arranged parallel to each other along the trunk of a tree. The characteristics of these fibrous cells and their arrangement affect strength properties, appearance, resistance to penetration by water and chemicals, resistance to decay, and many other properties.

Gross features of wood. Just under the bark of a tree is a thin layer of cells, not visible to the naked eye, called the cambium. Here cells divide and eventually differentiate to form bark tissue to the outside of the cambium and wood or xylem tissue to the inside. This newly formed wood (termed sapwood) contains many living cells and conducts sap upward in the tree. Eventually, the inner sapwood cells become inactive and are transformed into heartwood. This transformation is often accompanied by the formation of extractives that darken the wood, make it less porous, and sometimes provide more resistance to decay. The center of the trunk is the pith, the soft

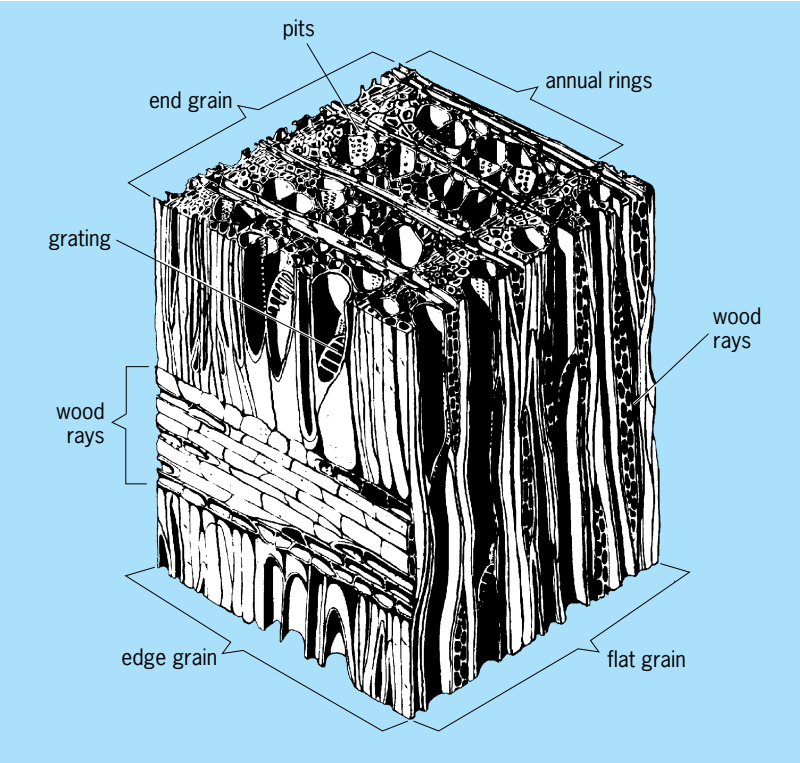

Fig. 1. Structure of a typical hardwood. (*USDA*)

tissue about which the first wood growth takes place in the newly formed twigs. *See* STEM.

The combined concentric bands of light and dark areas constitute annual growth rings. The age of a tree may be determined by counting these rings at the stump. *See* XYLEM.

In temperate climates, trees often produce distinct growth layers. These increments are called growth rings or annual rings when associated with yearly growth; many tropical trees, however, lack growth rings. These rings vary in width according to environmental conditions. Where there is visible contrast within a single growth ring, the first-formed layer is called earlywood and the remainder latewood. The earlywood cells are usually larger and the cell walls thinner than the latewood cells. With the naked eye or a hand lens, earlywood is shown to be generally lighter in color than latewood.

Because of the extreme structural variations in wood, there are many possibilities for selecting a species for a specific purpose. Some species (for example, spruce) combine light weight with relatively high stiffness and bending strength. Very heavy woods (for example, lignumvitae) are extremely hard and resistant to abrasion. A very light wood (such as balsa) has high thermal insulation value; hickory has extremely high shock resistance; mahogany has excellent dimensional stability.

Many mechanical properties of wood, such as bending strength, crushing strength, and hardness, depend upon the density of wood; the heavier woods are generally stronger. Wood density is determined largely by the relative thickness of the cell wall and the proportions of thick- and thin-walled cells present. *See* WOOD PROPERTIES.

Typical hardwood. The horizontal plane of a block of hardwood (for example, oak or maple) corresponds to a minute portion of the top surface of a stump or end surface of a log. The vertical plane corresponds to a surface cut parallel to the radius and parallel to the wood rays. The vertical plane corresponds to a surface cut at right angles to the radius and the wood rays, or tangentially within the log. In hardwoods, these three major planes along which wood may be cut are known commonly as end-grain, quarter-sawed (edge-grain) and plain-sawed (flat-grain) surfaces (**Fig. 1**).

Hardwoods have specialized structures called vessels for conducting sap upward. Vessels are a series of relatively large cells with open ends, set one above the other and continuing as open passages for long distances. In most hardwoods, the ends of the individual cells are entirely open; in others, they are separated by a grating. On the end grain, vessels appear as holes and are termed pores. The size, shape, and arrangement of pores vary considerably between species, but are relatively constant within a species.

Most smaller cells on the end grain are wood fibers which are the strength-giving elements of hardwoods. They usually have small cavities and relatively thick walls. Thin places or pits in the walls of the wood fibers and vessels allow sap to pass from one cavity to another. Wood rays are strips of short horizontal cells that extend in a radial direction. Their function is food storage and lateral conduction. Most of the rays in flat-grain surfaces are two to five cells wide, but their width and height vary in different species of hardwoods from 1 to more than 50 cells wide and from less than 100 to more than 4 in. (10 cm) in height. *See* PARENCHYMA; SECRETORY STRUCTURES (PLANT).

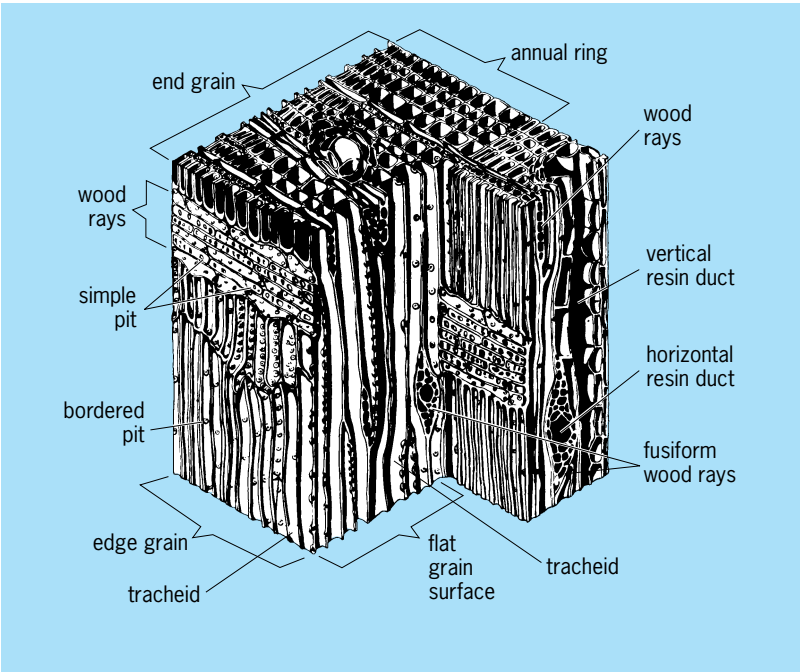

Fig. 2. Structure of a typical softwood. (*USDA*)

Typical softwood. The rectangular units that make up the end grain of softwood are sections through long vertical cells called tracheids or fibers (**Fig. 2**). Because softwoods do not contain vessel cells, the tracheids serve the dual function of transporting sap vertically and giving strength to the wood. Softwood fibers range from about 0.1 to 0.3 in. (3 to 8 mm) in length.

The wood rays store and distribute sap horizontally. Fusiform wood rays are rays with horizontal resin ducts at their centers. In the center of the end grain is a vertical resin duct. However, some softwoods, such as cedar and true fir, do not have resin ducts. The annual ring is often divided into an earlywood zone composed of thin-walled cells and a latewood zone composed of thicker-walled cells.

Sap passes from ray parenchyma cells through simple pits, unthickened portions of the cell wall, to tracheids or vice versa. Bordered pits have their margins overhung by the surrounding cell walls, but still function as passageways for sap to move from one cell to another.

Cell walls. The principal compound in mature wood cells is cellulose, a polysaccharide of repeating glucose molecules which may reach 4 μm in length. These cellulose molecules are arranged in an orderly manner into structures about 10–25 nm wide called microfibrils. This ordered arrangement in certain parts (micelles) gives the cell wall crystalline properties that can be observed in polarized light with a light microscope. The microfibrils wind together like strands in a cable to form macrofibrils that measure about 0.5 μm in width and may reach 4 μm in length. These cables are as strong as an equivalent thickness of steel.

This framework of cellulose macrofibrils is cross-linked with hemicelluloses, pectins, and lignin. Lignin, the second most abundant polymer found in plants, gives the cell wall rigidity and the substance that cements the cells together. *See* CELL WALLS (PLANT); CELLULOSE; LIGNIN; PECTIN.

Wood identification. Naked-eye field identification of unknown woods can often be made on the basis of readily visible characteristics, such as color, odor, density, or grain pattern. Observing the smoothed transverse surface with the aid of a hand lens increases the accuracy of identification, especially hardwood identification. However, for the most accurate identification, the naked eye, hand lens, and light microscope are used to examine the transverse, radial, and tangential surfaces and the various characteristics therein of the unknown wood. Wood descriptions, dichotomous keys, edge-punched cards, tables, photographs, and computer-assisted systems are also available to help in the identification process. *See* OPTICAL MICROSCOPE; PLANT ANATOMY; TREE.

Regis B. Miller

Bibliography. H. A. Core, W. A. Cote, and A. C. Day, *Wood Structure and Identification*, 2d ed., 1979; R. B. Hoadley, *Identifying Wood: Accurate Results with Simple Tools*, 1990.

Wood chemicals

Chemicals obtained from wood. The practice was carried out in the past, and continues wherever technical utility and economic conditions have combined to make it feasible. Woody plants comprise the greatest part of the organic materials produced by photosynthesis on a renewable basis, and were the precursors of the fossil coal deposits. Future shortages of the fossil hydrocarbons from which most organic chemicals are derived may result in the economic feasibility of the production of these chemicals from wood.

Wood is a mixture of three natural polymers—cellulose, hemicelluloses, and lignin—in an approximate abundance of 50:25:25. In addition to these polymeric cell wall components which make up the major portion of the wood, different species contain varying amounts and kinds of extraneous materials called extractives. Cellulose is a long-chain polymer of glucose which is embedded in an amorphous matrix of the hemicelluloses and lignin. Hemicelluloses are shorter or branched polymers of five- and six-carbon sugars other than glucose. Lignin is a three-dimensional polymer formed of phenylpropane units. Thus the nature of the chemicals derived from wood depends on the wood component involved. *See* CELLULOSE; HEMICELLULOSE.

Destructive distillation. In the past, destructive distillation of wood was employed to produce charcoal and was once an important industry, of which only vestiges remain. Volatile organic chemicals which were recovered from the distillate were acetic acid, acetone, and methanol. In addition, wood pyrolysis yielded tar oil fractions useful for medicinals, smoking meats, disinfectants, and weed killers. The resinous exudates from pine trees provided turpentine and rosin, while extracts of hardwoods and various barks yielded tannins useful in making leather. *See* ACETIC ACID; ACETONE; LEATHER AND FUR PROCESSING; METHANOL.

Modern processes. Chemicals derived from wood at present include bark products, cellulose, cellulose esters, cellulose ethers, charcoal, dimethyl sulfoxide, ethyl alcohol, fatty acids, furfural, hemicellulose extracts, kraft lignin, lignin sulfonates, pine oil, rayons, rosin, sugars, tall oil, turpentine, and vanillin.

Most of these are either direct products or byproducts of wood pulping, in which the lignin that cements the wood fibers together and stiffens them is dissolved away from the cellulose. High-purity chemical cellulose or dissolving pulp is the starting material for such polymeric cellulose derivatives as viscose rayon and cellophane (regenerated celluloses from the xanthate derivative in fiber or film form), cellulose esters such as the acetate and butyrate for fiber, film, and molding applications, and cellulose ethers such as carboxymethylcellulose, ethylcellulose, and hydroxyethylcellulose for use as gums.

The chemical fragments of the cell wall polymers which end up in solution after pulping can be isolated from the pulping liquors and used. These are principally lignin derivatives. Sulfonated lignins from

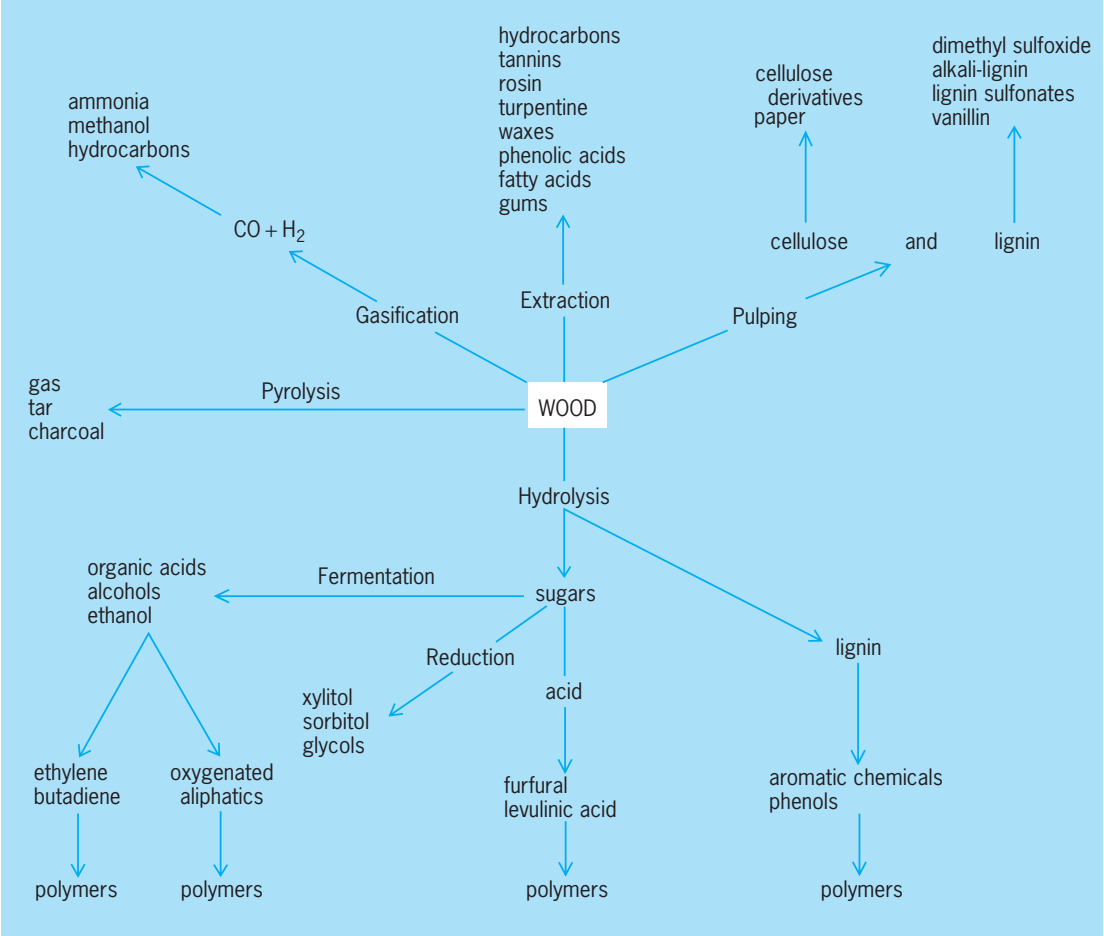

Chemical pathways for obtaining chemicals from wood.

the sulfite pulping process can be precipitated and used as tanning agents, adhesives, binders, dispersants, and so on. Mild alkaline oxidation of lignin sulfonates yields vanillin for flavoring and odorant applications. Alkali lignin from sulfate or kraft black liquor can be used as an extender for resins, for rubber reinforcement, and in emulsion stabilization. A volatile product from kraft black liquor is dimethyl sulfoxide, useful as a solvent. *See* VANILLA.

Sugars in spent sulfite liquor can be fermented by yeast to produce ethyl alcohol and food and fodder supplements. Oleoresinous wood components recovered from the draft pulping process include turpentine, pine oils, fatty acids, rosin, and tall oil. *See* ETHYL ALCOHOL; PINE TERPENE; ROSIN; TALL OIL.

Arabinogalactan, a hemicellulose gum extracted from larch, can be used in place of gum arabic. Bark extracts include phenolic acids from various conifers useful as extenders for synthetic resin adhesives and as binders, as well as waxes extracted from Douglas-fir bark which can be used for general wax applications. *See* WAX, ANIMAL AND VEGETABLE.

Wood hydrolysis converts the carbohydrate polymers in wood to simple sugars by chemical reaction with water in the presence of acid catalysts. Cellulose yields glucose, while softwood hemicellulose principally xylose. The hexoses can be readily fermented to ethanol. Acid treatment of hexoses yields hydroxymethylfurfural which is rapidly converted to levulinic acid, while pentoses yield furfural.

Potential chemicals. Considerable development effort has been devoted to the conversion of renewable biomass, of which wood is the major component, into the chemicals usually derived from petroleum. Processes for which technical feasibility has been demonstrated are shown in the **illustration**. Economic feasibility is influenced by fossil hydrocarbon cost and availability. *See* BIOMASS.

The wood cell wall polymers in their natural mixed state can be broken down into simpler compounds by nonselective drastic processes of pyrolysis and gasification in the same way that coal is converted into chemicals. In gasification, wood is heated at temperatures of about 1000°C (1830°F) to form a mixture of carbon monoxide and hydrogen as the major products with accompanying low yields of hydrocarbons as by-products. The carbon monoxide and hydrogen, as in coal gasification, may be converted into ammonia, methanol, or hydrocarbons. Pyrolysis or thermal degradation of wood in the absence of oxygen at lower temperatures converts the wood to charcoal, gas, tar, and oil. *See* COAL GASIFICATION; PYROLYSIS.

Selective processing of the wood components can

retain more carbon-carbon bonds and provide substantial yields of a wide variety of chemicals. Hydrolysis of cellulose to glucose by high-temperature dilute acids, low-temperature strong acids, or enzymes is the first step in its utilization for chemicals. Ethanol from glucose fermentation may be used as an industrial chemical, as a fuel for internal combustion engines, or as an intermediate for the production of other chemicals.

Ethanol may be dehydrated to ethylene, which is the building block for many other organic intermediates and polymers. By oxidation of ethanol to acetaldehyde, an intermediate for further production of acetic acid and anhydride, acrylonitrile, vinyl acetate, and butadiene may be obtained.

Fermentation of glucose may alternatively be directed to lactic acid, acetone, butanol, isopropanol, glycerin, and so on. Glucose hydrogenation can yield glycols as well as sorbitol. *See* FERMENTATION.

Hydrolysis of hemicelluloses yields mannose, which can be processed along with glucose, and xylose, which can be converted to furfural or xylitol. The aromatic lignin component of wood can be converted into phenols and other aromatic chemicals such as benzene under various hydrogenation conditions followed by dealkylation and dehydroxylation. *See* ORGANIC SYNTHESIS; WOOD PRODUCTS.

Irving S. Goldstein

Bibliography. I. S. Goldstein, *Organic Chemicals from Biomass*, 1981; P. Hakkila, *Utilization of Residual Forest Biomass*, 1989; E. J. Soltes, *Wood and Agricultural Residues: Research on Use for Feed*, *Fuels*, and *Chemicals*, 1983; D. L. Wise, *Organic Chemicals from Biomass*, 1983.

Wood composites

Wood-based composites and a combination of two or more elements held together by a matrix. By this definition, what we call "solid wood" is a composite. Solid wood is a three-dimensional composite composed of cellulose, hemicelluloses, and lignin polymers with smaller amounts of inorganics and extractives held together in a lignin matrix. Among the advantages of developing wood composites from solid wood are the ability to: (1) use trees or tree parts that are too small for lumber or veneer production (includes thinning, limbs, brush, and other forest biomass), (2) use underutilized wood species, (3) use mixed wood species, (4) remove defects, (5) create more uniform materials, (6) develop materials that are stronger than the original solid wood, (7) make materials of different shapes, and (8) produce both large-volume low-value and low-volume value-added materials. A beneficial by-product is the removal of biomass from overcrowded forests, especially in the western United States, to improve forest health and to reduce the fuel load that increases the threat of catastrophic fires.

Historically, wood was used only in its solid form as large timbers, or lumber. As the availability of large-diameter trees decreased and the price increased, the wood industry looked to replace large timber products and solid lumber with reconstituted wood products made using smaller diameter trees, and manufacturing, saw and pulp mill wastes.

Wood elements. Solid wood can be broken down into smaller and smaller elements, that is, veneers, flakes, strands, chips, flakes, particles, fiber, and cellulose. As the element size becomes smaller, uniformity increases and composites made from these smaller elements become more like a true material, that is, consistent, uniform, continuous, predictable, and reproducible. As the size of the composite element gets smaller, it is also possible to either remove defects (such as knots, cracks, and checks) or redistribute them to reduce their effect on product properties. Size reduction, however, requires energy and may increase water consumption. Adhesion and the development of the composite matrix is also a critical issue.

Adhesives. The basic adhesives most commonly used for wood composites are formaldehyde, urea, melamine, phenol, resorcinol, and isocyanates. Despite the apparent simplicity in terms of families of chemicals, the formulations are highly complex mixtures of chemicals and additives depending on the specific application. Wood composites used outdoors are usually produced using phenol, resorcinol, or isocyanate adhesives.

Types of composites. Wood composites can be classified in several ways: by density, element type, process, or application. Classification by specific gravity includes high-, medium-, and low-density composites (see **illustration**). Classification by element type includes strandboard, waferboard, chipboard, flakeboard, particleboard, and fiberboard. Classification by process includes wet or dry production processes. Classification by application would include products such as insulation board, underlayment composites, and structural and core composites.

Strandboard. Structural composite lumber (SCL) can be made using strands of small-diameter wood glued parallel to the length. SCL products include oriented strand lumber (OSL) and parallel strand lumber (PSL). Laminated strand lumber (LSL), oriented strand board (OSB), and OSL are produced using different lengths and sizes of strands. LSL uses strands that are about 0.3 m in length while OSB is produced from shorter strands. PSL is made from strands that are 3 mm thick, approximately 19 mm wide, and 0.6 m in length. Usually Douglas-fir, southern pines, western hemlock, and yellow poplar are used but other species are also used. The major adhesives used to produce SCL products are phenol-formaldehyde or isocyanates. All of these SCL products are used as replacements for solid wood and have a specific gravity of 0.5 to 0.8.

Plywood. The modern plywood industry began around 1910 but the furniture industry had used veneers over solid wood for several hundred years before that. Plywood is usually made using veneers placed at right angles to each other to increase dimensional stability and add strength.

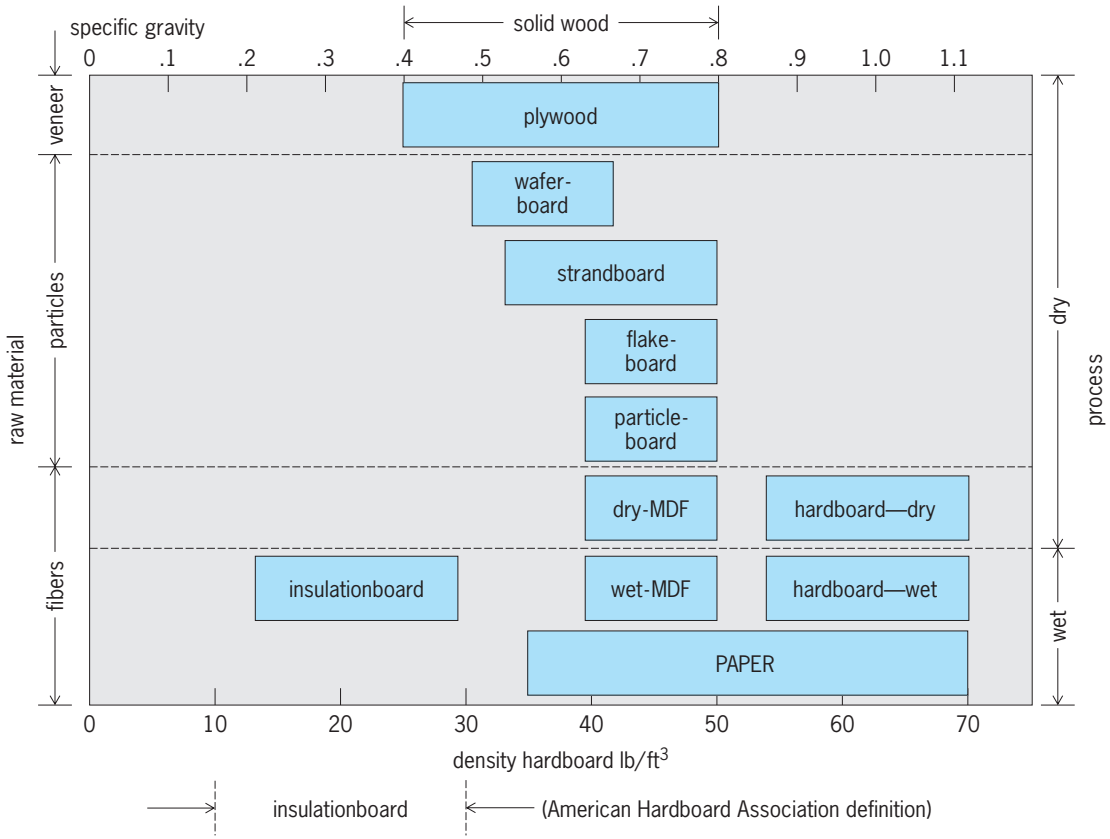

Wood composites classified by density, element type, process, and application.

Plywood is mainly used as underlayment in roofs and flooring.

Waferboard and flakeboard. The wafer board and flakeboard industries started in the 1960s. The specific gravity of wafer and flakeboard is usually between 0.6 and 0.8 and made using a waterproof adhesive such as phenol formaldehyde or an isocyanates. Flakes used in flakeboard production vary in size from 1 to 2 cm wide, 2 to 3 cm in length, and 2–5 mm in thickness. Wafers are almost as wide as they are long, while flakes are much longer than they are wide. Wafers are also thicker than flakes. These are used as the structural skin over wall and floor joists. *See* PLYWOOD.

Particleboard. The particleboard industry started in the 1940s. Particleboard is produced using particles that vary in length from about 2 mm to 1 cm. The specific gravity of particleboard is between 0.6 and 0.8 and is usually produced from softwoods such as Douglas-fir, southern pines, or other low-value wood sources. Today, particleboard is used in furniture using metal or nylon screw-together joints.

Hardboards. The hardboard industry started around 1950 and is produced mainly by a dry process with a specific gravity between 0.8 and 1.1. Dry formed hardboard is made using a fiber made by mechanical separation of the fibers from the wood. The medium-density fiberboard (MDF) industries started in the early 1960s and are usually made from fiber processed by aqueous high-temperature refining.

Three-dimensional shapes. Strand type boards can be formed into three-dimensional shapes but, in general, all of the products described above are produced in flat sheets and used in two-dimensional designs. Three-dimensional composites can also be made using mainly particles and fibers and the final shape can be formed in the hot press during manufacturing. *See* WOOD ENGINEERING DESIGN; WOOD PRODUCTS; WOOD PROPERTIES.

Roger M. Rowell

Bibliography. *Panel Trends*, 2002, Composite Panel Association. Gaithersburg, MD. 2002; *Regional Production & Market Outlook for Structural Panels and Engineered Wood Products, 2002–2007*, APA—Engineered Wood Association, Tacoma, WA, 2002; *Wood Product Demand and the Environment: Proceedings of an International Conference Sponsored by the Forest Products Research Society*, Madison, WI, 1992; *Wood Products for Engineered Structures: Issues Affecting Growth and Acceptance of Engineered Wood Products*, Forest Products Research Society, Madison, WI, 1993.

Wood degradation

Decay of the components of wood. Despite its highly integrated matrix of cellulose, hemicellulose, and lignin, which gives wood superior strength properties and a marked resistance to chemical and microbial attack, a variety of organisms and processes are capable of degrading wood. The decay process is a continuum, often involving a number of organisms over many years. Wood degrading agents are both biotic and abiotic, and include heat, strong acids or

bases, organic chemicals, mechanical wear, and sunlight (uv degradation).

Abiotic degradation. Heat degrades both cellulose and hemicellulose, reducing strength and causing the wood to darken. At temperatures above 451°F (219°C), combustion occurs.

Because wood is resistant to many chemicals, it has long been used in cooling towers. Strong acids eventually degrade the carbohydrate portion of wood, reducing its strength. Strong bases attack the lignin, leaving the wood appearance bleached and white; this effect forms the basis for the Kraft pulping of wood chips for paper production. Other chemicals, such as concentrated organics or salt solutions, can also disrupt the lignocellulosic matrix, reducing material properties of the wood. *See* PAPER.

Sunlight, primarily through the action of ultraviolet light, also degrades wood through the creation of free radicals which then degrade the wood polymers. Ultraviolet degradation extends only a few cell layers from the surface, but can become severe when water or wind continually removes weakened wood, exposing nondegraded wood beneath to ultraviolet light. In most instances, however, ultraviolet degradation causes the wood to change to a gray to brown color which is highly valued in some markets.

Mechanical wear of wood can occur in a variety of environments. For example, floating docks in marinas may shift up and down along a wood pile, continually abrading the wood and reducing the effective pile circumference. Wood railway ties are subjected to repeated shock loading as rail cars pass over. Ties in mainline track often fail from mechanical effects, long before other agents of deterioration can cause damage.

Abiotic degradation is often difficult to distinguish from biotic damage. However, the absence of biological organisms plus increases in pH or the location of the damage can provide important clues concerning the nature of the damage.

Biotic degradation. Biotic damage can occur from a variety of agents, including bacteria, fungi, insects, marine borers, and birds and animals. Birds and animals generally cause mechanical damage in isolated instances.

Biotic requirements. All biotic agents have four basic requirements: adequate temperature (32–104°F or 0–40°C) with most optima between 77–90°F (25–32°C), oxygen (or other suitable terminal election acceptor), water, and a food source. Temperature ranges for growth of decay organisms vary widely, and temperature is generally not controllable as a means for limiting the activities of wood-degrading agents.

Although oxygen is required for activity of most wood-degrading organisms, some bacteria are capable of anaerobic growth. Furthermore, many fungi are capable of growing at very low oxygen tensions or can become dormant under low oxygen conditions and then resume growth once oxygen levels improve. Many lumber mills create low oxygen conditions by submerging logs in ponds to limit biologic attack until the wood can be sawn and dried. Prolonged submersion, however, permits bacterial degradation of wood pit membranes, increasing wood permeability and causing logs to sink to the bottom of the pond.

Water is a critical element for biotic decay agents: it serves as reactant in degradative reactions, a medium for diffusion of enzymes into wood and degradative products back to the organism, and a wood swelling agent. Wood is considered most decay susceptible when the moisture content exceeds 30% by weight. At this level, called the fiber saturation point; free water is present in the wood. Thus, keeping wood dry is the primary method for protecting against decay, although some organisms have adapted to survive in wood at lower moisture levels.

Bacteria. Bacteria are primitive, single-celled organisms that generally are not major degraders of wood products, but they can damage pit membranes, thereby increasing permeability, and some are capable of cell wall degradation. Bacteria appear to be most important in submerged environments, highly fertilized agricultural soils, and other regimes where growth by more common fungal agents is limited. The role of these agents and their interactions with other wood-degrading organisms remain poorly defined.

Fungal organisms. Fungi are among the most important wood-degrading organisms because they play an important role in terrestrial carbon cycling. Wood-degrading fungi can be classified as molds, stainers, soft rotters, brown rotters, and white rotters on the basis of the attack patterns. Molds, stainers, and soft rotters are members of the ascomycetes and the deuteromycetes (Fungi Imperfecti). Molds primarily colonize the surface of freshly sawn wood, attacking readily available carbohydrates in the ray cells. These fungi cause cosmetic damage to the wood due to the production of pigmented spores (which can be brushed from the wood surface), and they can increase wood permeability, leading to uneven finishing of wood. *See* ASCOMYCOTA; DEUTEROMYCOTINA.

Stain fungi cause more serious damage since they are cellulolytic and can remove large portions of the ray cells. The hyphae of stain fungi are dark pigmented, giving the wood a bluish color. Stain fungi cause some loss in toughness and increase permeability, but have little effect on most properties. Both molds and stain fungi can be controlled by rapid drying after sawing or by application of prophylactic fungicides shortly after sawing.

Soft rot fungi are capable of significant degradation of wood carbohydrates in specialized environments, particularly where wood is wet or exposed to exogenous nutrients. Soft rot attack is often concentrated near the wood surface, and the decay quickly reduces strength. These fungi also exhibit significant preservative tolerance and are an important factor in degradation of some wood species.

Brown rot fungi are basidiomycetes and use wood carbohydrates in a manner which reduces strength rapidly at the early stages of attack. Brown rotters can remove up to 70% of the wood mass, leaving a cracked brown degraded mass of modified lignin. Brown rot fungi are most prevalent in coniferous

wood species and are associated with dramatic drops in wood strength at very early stages of decay. As a result, early detection of these fungi is an important aspect of a wood inspection program.

White rot fungi are also basidiomycetes, but these species can utilize all wood components, removing up to 97% of the wood. White rot fungi are most prevalent on hardwoods, possibly owing to a reduced lignin content in these species. The ability of white rot fungi to degrade complex ring structures has been exploited for degradation of pesticides and pretreating chips for manufacturing paper. *See* BASIDIOMYCOTA; FUNGI; PESTICIDE.

Insects. A number of insects have evolved to attack wood, including termites (Isoptera), beetles (Coleoptera), and bees and ants (Hymenoptera). Termites are social insects with highly organized worker, soldier, and reproductive castes. They are the most important wood-degrading insects in most environments, and their activity causes severe economic losses. Termite damage is generally difficult to detect because the workers do not forage outside. As a result, most damage is detected when reproductives swarm from the nest to start new colonies or when wood is so badly damaged that it fails. Most termites require wet wood, so breaking wood-soil contact and preventing wetting are preventive measures. Drenching the soil around a structure with a termiticide to develop a barrier is also commonly used. *See* COLEOPTERA; HYMENOPTERA; ISOPTERA.

A variety of beetles attack freshly fallen trees with the bark attached, but most do not attack finished products, although their damage may be found in the material. Of the wood-attacking beetles, the powder post beetles (comprising families Anobiidae, Bostrichiidae, and Lyctidae) are among the most important. Powder post beetles often cause problems in structures which are subject to infrequent inspections, and are a problem in museums. Damage is usually detected by the presence of piles of finely powdered frass below a wood member or by the presence of small exit holes on the wood surface. Powder post beetle infestations are often localized, and control is usually effected by fumigation with methyl bromide or sulfuryl fluoride.

Among the Hymenoptera which attack wood are the carpenter bees (Anthophoridae) and carpenter ants (Formicidae). Neither of these insects uses the wood as a food source. Carpenter bees excavate long galleries along the grain and provision cells along the tunnel with captured insects. Carpenter ants are social insects with a colony structure similar to the termites. They are commonly found in moist wood, and their damage can be substantial. Insecticides sprayed outside the nest and transported into the nest by workers are most effective since they have a higher probability of killing the queen.

Marine borers. In saline environments, marine borers can cause significant wood losses. While marine borers are often considered to be detrimental, these organisms are important recyclers of coarse woody debris in marine environments and their droppings provide a nutrient source to deep-ocean communities. Three groups of marine borers [shipworms, pholads, and gribbles (*Limnoria*)] cause most wood damage in these areas. Shipworms and pholads are clamlike mollusks which begin life as free-swimming larvae which then settle on the wood surface. Pholads continue to remain in their shells and gradually create pear-shaped cavities near the wood surface, from which they filter-feed in the surrounding water. Shipworms use two small shells at the top of the head to rasp a wood cavity and penetrate deeper into the wood, where they can grow to lengths of up to 5 ft (1.5 m) in 2–3 years. Shipworms are difficult to detect, and their internal damage can lead to wood failure in as little as 2 years. Both shipworms and pholads can be excluded from wood by treatment with creosote, while shipworms can also be controlled with a number of inorganic arsenicals. *See* SHIPWORM.

Limnoria are mobile crustaceans which tunnel into wood near the surface, but are not believed to use wood as a food source. The tunnels weaken the wood, which gradually erodes because of wave action. The freshly exposed wood is also attacked, creating a cycle of attack and erosion which eventually results in wood failure. Damage to wood by *Limnoria* is most often prevented by treating the wood with creosote or inorganic arsenicals. In areas where pholads and *Limnoria* attack, dual creosote and inorganic arsenical treatments are required. *See* BORING BIVALVES.

Wood protection. Protecting wood from degradation can take a number of forms. By far the simplest method is to employ designs which limit wood exposure to moisture, and the effectiveness of these methods can be seen in many centuries-old wood buildings in Asia and Europe. In some cases, however, water exclusion is not possible and alternative methods must be employed. The simplest of these methods is the use of heartwood from naturally durable species. Sapwood of any species has no natural durability, but some wood species produce heartwood which is resistant to biological attack. Decay- or insect-resistant species include redwood (*Sequoia sempervirens*), western red cedar (*Thuja plicata*), and ekki (*Lophira alata*), while marine-borer-resistant heartwoods include greenheart (*Ocotea rodiaei*) and ekki. Most marine-borer-resistant woods contain high levels of silica which discourages marine borer attack, while species resistant to terrestrial decay agents often contain toxic phenolics.

Wood can also be protected from degradation by spraying, dipping, soaking, or pressure treatment with preservatives. Protecting wood from mold and stain fungi is most often accomplished by dipping or spraying the wood with a fungicide shortly after sawing. These treatments generally provide up to 6 months of protection to the wood surface. *See* WOOD PROPERTIES.

Jeffrey J. Morrell

Bibliography. R. A. Eaton and M. D. Hale, *Wood: Decay, Pests and Protection*, 1993; A. S. Panshin and C. de Zeeuw, *Textbook of Wood Technology*, 4th ed., 1980; R. A. Zabel and J. J. Morrell, *Wood Microbiology: Decay and Its Prevention*, 1997.

Wood engineering design

The process of creating products, components, and structural systems with wood and wood-based materials. Wood engineering design applies concepts of engineering in the design of systems and products that must carry loads and perform in a safe and serviceable fashion. Common examples include structural systems such as buildings or electric power transmission structures, components such as trusses or prefabricated stressed-skin panels, and products such as furniture or pallets and containers. The design process considers the shape, size, physical and mechanical properties of the materials, type and size of the connections, and the type of system response needed to resist both stationary and moving (dynamic) loads, and function satisfactorily in the end-use environment. *See* ENGINEERING DESIGN; STRUCTURAL DESIGN.

Wood is used in both light frame or heavy timber structures. Light frame structures consist of many relatively small wood elements such as lumber covered with a sheathing material such as plywood. The lumber and sheathing are connected to act together as a system in resisting loads; an example is a residential house wood floor system where the plywood is nailed to lumber bending members or joists. In this system, no one joist is heavily loaded because the sheathing spreads the load out over many joists. Service factors such as deflection or vibration often govern the design of floor systems rather than strength. Light frame systems are often designed as diaphragms or shear walls to resist lateral forces resulting from wind or earthquake. *See* FLOOR CONSTRUCTION.

In heavy timber construction, such as bridges or industrial buildings, there is less reliance on system action and, in general, large beams or columns carry more load transmitted through decking or panel assemblies. Strength, rather than deflection, often governs the selection of member size and connections. There are many variants of wood construction using poles, wood shells, folded plates, prefabricated panels, logs, and combinations with other materials.

Solid wood. Structural dimension lumber is the most commonly used solid wood engineering material, particularly for light frame structures as well as for residential roof and floor trusses. Nonstructural grades of lumber assure satisfactory performance in furniture frames, pallets and containers, truck flooring, and a myriad or other products. The strength, stiffness, and other properties of structural lumber are determined by grading each piece. A visual stress rating is determined by looking at each piece and assigning it to a strength category based on the observed defects and growth characteristics. Machine stress rating classifies lumber by using a machine to nondestructively test each piece. This method is supplemented by additional human control for characteristics which cannot be physically tested. The stress rating process accounts for the different inherent properties of clear, defect-free wood and the influence of growth characteristics such as knots or sloping grain.

Mechanical properties. The use of specific wood types for engineering and other purposes is related to the wood's mechanical properties. Growth and manufacturing characteristics such as knots, grain deviation, and wane reduce strength of structural lumber, and increase its variability. Allowable safe design strength and stiffness for lumber depends on the lumber grade. The tension perpendicular to grain and the shear strength for wood are relatively low and dictate that wood products should not be used where these stresses may dominate. For example, short deep wood beams with a span/depth ratio less than 12 may not be efficient. Also, high loads should not be fastened to the lower side of wood beams because of cross-grain tension.

Dry wood is generally ductile and flexible with excellent energy absorption characteristics except in tension perpendicular to grain and shear where failure may be preceded by little warning. The strength of wood members depends on the duration of an applied load under some conditions. In general, wood has greater resistance to short-term loads than to long-term loads. Limits on the deflection of beams and the buckling capacity of long columns may govern the design of these elements because of moderate stiffness. Wood columns that are short or fully braced are usually economic if buckling is avoided. An example is the use of wood in pilings, piers, and utility poles.

High temperatures, above 150°F (65°C), can cause permanent strength and stiffness reductions in wood under prolonged exposure. Increasing wood moisture content can also reduce some properties and require that wood be treated to prevent biodeterioration. *See* FOUNDATIONS; WOOD ANATOMY; WOOD PROCESSING; WOOD PROPERTIES.

Engineered wood composites. Wood composites are products composed of wood elements that have been glued together to make a different, more useful or more economical product than solid sawn wood. Plywood is a common example of a wood-composite sheathing panel product where layers of veneer are glued together. Plywood is used as a sheathing material in light frame wood buildings, wood pallets, and containers to distribute the applied forces to beams of lumber or other materials. Shear strength, bending strength, and stiffness are the most important properties for these applications and may be engineered into the panel by adjusting the species and quality of veneer used in the manufacture.

Glue-laminated timber (glulam) is a similar example where layers of lumber are glued together to form larger, longer, and stronger structural elements than can be obtained with solid sawn lumber alone. Glulam beams can reach lengths over 100 ft (30 m), with cross sections exceeding 12 × 36 in. (300 × 900 mm). Curved glulam beams and arches are manufactured to maintain the intrinsic strength of wood by keeping the grain direction running along the length of the element.

Laminated veneer lumber is manufactured by

gluing together layers of dry veneer into a large panel with the grain direction of all layers oriented in the same direction. The panel is then ripped and crosscut to form a lumberlike product. The depth of the lumber is determined by sawing, and the width is determined by the number of veneer layers. Laminated veneer lumber generally has greater strength than solid sawn lumber because the defects and growth characteristics are distributed through-out the element rather than concentrated in a knot or some other characteristic. In addition, this type of lumber can be made from one or many species and cut into sizes not obtainable with solid sawn lumber.

The strength of laminated veneer lumber is determined by testing and is ensured by continuous quality control over the manufacturing process. Laminated veneer lumber is used for high-strength applications where straightness is important, such as truss chords, composite I beams, garage and window headers, scaffold planking, and furniture frames. Laminated veneer lumber can also be made with some curvature by pressing the glued veneer in a curved form before the adhesive cures.

Oriented strand board. This composite panel sheathing product is made by gluing thin flakes or compressed strands under pressure and heat. The flakes or strands are lined up and arranged in three to five layers which are oriented at right angles to one another. The center layer may have strands in a random orientation. This orientation by layers gives more uniform strength and stiffness to the panel in the same manner as alternating veneer grain direction does in plywood. Oriented strand board is used the same way as plywood and has similar strength and stiffness properties. In addition, oriented strand board manufacturers can use irregularly shaped logs and low-density species to make the flakes and strands. The properties are ensured by test and continuous quality control of the manufacturing process.

Composite I beams. Usually made of several wood-based materials, composite I beams are used as lumber substitutes for bending applications. This product looks like an I beam with the outer flanges designed to carry the bending moment through axial compression and tension. The center of the I or web is designed to resist the beam shear stresses. Laminated veneer lumber or solid-sawn lumber is used as the flange material which is glued to a web made of plywood or oriented strand board. Composite I beams are used in either light frame or heavy timber structures where a very uniform economical product is required. Special care must be taken to ensure the lateral stability and bearing capability of this product because of the relatively thin web.

Connections. A major advantage of wood and wood-based engineering materials is the ease with which they can be connected. Nails, staples, and screws are field installed in light frame structures by hand or with simple power equipment. This creates adequate connections that have high energy absorption and ductility. Individual connections, however, are generally not engineered with light frame construction. Fastening schedules, which specify number, size, and type of fastener for a specific application, have been developed by product manufacturers and building code officials.

The development of modern wood structural systems depends on efficient, economical connections systems to transfer forces. Prefabricated steel connectors installed with nails or bolts, such as the joist hanger, tie-downs, and anchors, have enabled engineered wood construction to meet increasing safety and serviceability demands in high-wind and earthquake areas. The punched-tooth metal connector plate has enabled the development of the prefabricated wood truss which is prevalent in both roof and floor systems. The metal connector plated truss is the most highly engineered component in the light frame structural system, and each truss connection is individually engineered.

Adhesive connections are efficient means of force transfer but require a high level of technology and quality control. They are commonly used in glulam where finger jointing provides a means of developing any length of lumber. The webs are glued to the flanges in composite I beams. Plywood is glued to the lumber ribs in stressed-skin panels to develop the full system action of the panel. However, these adhesive connections are made in a factory under controlled conditions. Although there are exceptions, the most common field use is with elastomeric adhesive connections to enhance the stiffness of the nail connection between sheathing and solid wood, thus improving the overall service performance of floor systems.

Bolts, shear plates, split rings, and a myriad of special connectors are used with glulam and heavy timber construction. Traditional connections such as mortise and tenon joints are found in timber frame construction. In general, heavy timber connections are more expensive and more localized than with light frame construction. Each connection is designed independently, and strength is the most important criterion. There are two general design philosophies: the connection should always be stronger than the more costly connected members whose failure could be catastrophic to the structure; and connections should be designed to control failure, avoid catastrophe, and allow for repair without the need for member replacement. Either philosophy, or both, can be used depending on the design objectives, the type of expected loads, and the opportunity. *See* WOOD PRODUCTS. Thomas E. McLain

Bibliography. J. Bodig and B. A. Jayne, *Mechanics of Wood and Wood Composites*, 1993; K. F. Faherty and T. G. Williamson, *Wood Engineering and Construction Handbook*, 3d ed., 1998; Forest Products Laboratory Staff, *Wood Engineering Handbook*, 2d ed., 1990; A. D. Freas, R. C. Moody, and L. A. Soltis (eds.), *Wood: Engineering Design Concepts*, 1983.

Wood processing

Peeling, slicing, sawing, and chemically altering hardwoods and softwoods to form finished products such as boards or veneer; particles or chips for making

paper, particle, or fiber products; and fuel. *See* PAPER; VENEER.

Most logs are converted to boards in a sawmill that consists of a large circular or band saw, a carriage that holds the log and moves past the saw, and small circular saws that remove excess bark and defects from the edges and ends of the boards. One method is to saw the log to boards with a single pass through several saw blades mounted on a single shaft (a gang saw). Sometimes, the outside of the log is converted to boards or chips until a rectangular center or cant remains. The cant is then processed to boards with a gang saw.

Drying. A high percentage of the weight of freshly cut or green wood is water. The cells of green wood contain free water in the cell cavities and bound water in the cell walls. When all the free water has been extracted and before any of the bound water has been removed, the wood is said to be at the fiber saturation point. This is a moisture content of about 28% of the ovendry weight of the wood. Above the fiber saturation point, wood does not change dimension with a change in moisture content.

As the moisture content falls below the fiber saturation point, the bound water leaves the cell walls and the wood shrinks. But it does not shrink uniformly because wood is anisotropic; that is, it does not have the same properties in all directions. In drying a cross section of a log, wood shrinks very little longitudinally (the direction that is parallel to the long axis of the log). Shrinkage around the outside of the log (tangential) is about twice that from the center of the log to the bark (radial) and can be as much as 10% of the green dimension. A dry piece of wood does not have the same dimensions and often not the same shape as when green. During the drying process, differential shrinkage can cause internal stresses in the wood. If not controlled, this can result in defects such as cracks, splits, and warp.

Below the fiber saturation point, wood takes on and gives off water molecules depending on the relative humidity of the air around it and swells and shrinks accordingly. The drying process is complicated because during the initial stages, when the outside of a board drops below the fiber saturation point and the inside remains above it, the outside shell is restrained from shrinking. The outside shell will also set and become larger than if the shell were unrestrained. If drying is too rapid during this period, surface checks and splits can occur. As drying continues, the outer shell of the board eventually restrains the inner core from shrinking. If drying in this latter stage is too rapid, internal checks called honeycombs can occur. *See* WOOD PROPERTIES.

A dry kiln is used to control the rate of moisture loss. Essentially, a dry kiln is a room in which a controlled flow of air at a specified temperature and relative humidity is passed through layers of carefully stacked lumber. A method of venting moisture-laden air outside the kiln is necessary. Kiln schedules that safely remove moisture have been developed for most commercially important species. For easily dried species (most softwoods and some low-density hardwoods), schedules are based on time at specific temperatures.

At the end of the drying process, lumber can contain drying stresses. Casehardening occurs when the center of the board is subjected to tension and the outside of the board is being compressed. For boards that are to be processed further, such as to furniture or cabinet parts, it is imperative that these stresses be relieved. The final steps in the kiln-drying process are called stress equalizing and conditioning.

The final moisture content at a given temperature and relative humidity is called the equilibrium moisture content. In the United States, the equilibrium moisture content usually is about 6% in heated buildings in the winter and about 12% in the summer. Therefore, products for interior use should be manufactured at moisture contents of about 7–8% to avoid in-service product problems. *See* WOOD PROPERTIES.

Machining. Wood is machined to bring it to a specific size and shape for fastening, gluing, or finishing. With the exception of lasers, which have a limited application at this time, all machining is based on a sharpened wedge that is used to sever wood fibers. Tools for sawing, boring holes, planing, and shaping, as well as the particles in sandpaper, use some version of the sharpened wedge.

The process of passing a single blade over the workpiece as with a hand plane is called orthogonal cutting. There are three tool angles that affect machining quality (**Fig. 1**). The cutting or rake angle is the angle from a line perpendicular to the workpiece to the face of the tool. The sharpness angle is measured from the face to the back of the tool. The clearance angle is from the back of the tool to the new surface of the workpiece. In general, the cutting angle determines the type of wood chip; the sharpness angle affects tool life; and the clearance angle must be sufficient to prevent rubbing of the new surface with the back of the blade.

There are three basic types of wood chips formed by knife machining: Large cutting angles cause splitting ahead of the blade (type I chip; **Fig. 2**). For each

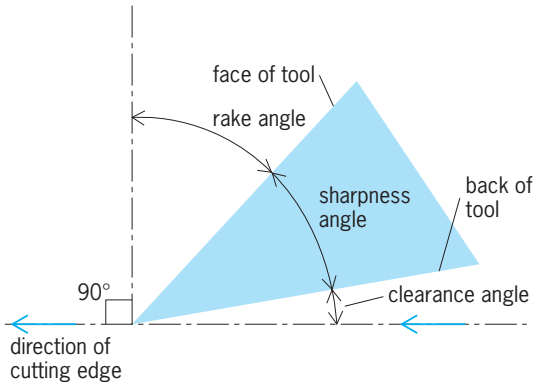

Fig. 1. The geometry of a cutting tool is described by the rake angle, measured from a line perpendicular to the direction of travel to the tool face; the sharpness angle, measured between the face and back of the tool; and the clearance angle, measured between the back of the knife and its direction of travel. (*After R. B. Hoadley, Understanding Wood, Taunton Press, 1981*)

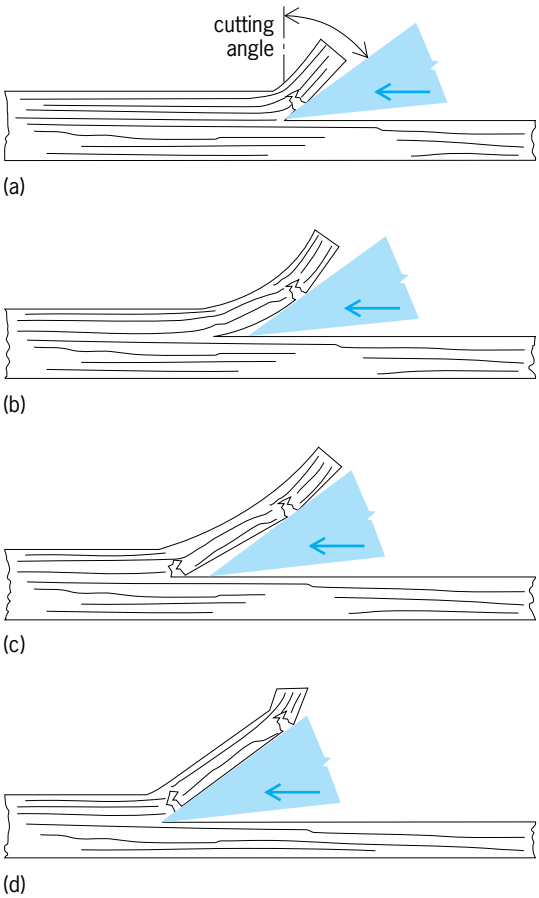

Fig. 2. Cutting action in planing wood. (*a*) Cut begins. (*b*) Chip (type I) bends, as it slides up the knife, and the wood fails ahead of the edge because of tension perpendicular to the grain. (*c*) Chip breaks. (*d*) Next segment of the cut begins. (*After R. B. Hoadley, Understanding Wood, Taunton Press, 1981*)

species and grain angle, there is a range of cutting angles that will cause a continuous chip or planer shaving and generate a smooth surface (type II chip). Small cutting angles cause the knife to plow through the wood and form a type III chip.

Small sharpness angles with large cutting and clearance angles offer the best of all possibilities. However, because relatively little tool material is present, tool dulling is relatively rapid. As the tool cutting edge rounds with wear, the cutting and clearance angles are changed at the point of fiber severance. In effect, a cutting edge that originally made a type II chip can dull so that the cutting angle becomes small and makes a type III chip.

When several blades are mounted in a rotating cutterhead and the workpiece is fed past, the process is called peripheral milling. It is difficult to manually make all cutting edges protrude equally. A process called jointing grinds a small amount from the back of the protruding cutting edge. If the protruding edge is excessive, a negative clearance angle will be present on a sharp knife and possibly cause surface defects (raised or loosened grain).

The most important factor affecting surface smoothness in knife planing is whether the knife is machining with or against the grain (**Fig. 3**). It is rare that the grain will be parallel to the surface for the entire length of a long, clear piece. In addition, the grain in and around knots often poses a problem for straight knife planers as torn or chipped grain can result. Abrasive planers that use belts of sandpaper with large grits to rough-plane are used to prevent tearing and chipping. *See* WOOD ENGINEERING DESIGN.

Veneer cutting. Veneer cutting is considered a special case of orthogonal machining in which the type II chip is the objective rather than a smooth surface on the core. Veneer is made by peeling (rotary cutting) or slicing and can be cut in thicknesses from $^1/_{100}$ to $^1/_4$ in. In rotary cutting of veneer, a log is chucked on both ends and turned while a knife advances against it. The rate of advance and the speed of log rotation determine the thickness of the veneer. For slicing, a block of wood is held against a frame that moves up and down diagonally. The slicing knife is advanced between cuts to set the thickness of the veneer.

The side of the veneer in contact with the knife can crack or check severely as it moves sharply away from the flitch or core, particularly when the veneer is thick. The loose side (with lathe or knife checks) and the tight side (without checks) can be identified by flexing the veneer. In flexing, the checks in the loose side open slightly, allowing the veneer to bend easily while the tight side remains fairly stiff. Paint or other finishes may fail if the loose side is mistakenly used as the exposed face in plywood or inlaid veneer products. To reduce or prevent lathe checks, a nose or pressure bar is set next to the knife so that the wood is compressed slightly just before it is cut.

Composite-type products. Wood is ground to fibers for hardboard, medium-density fiberboard, and paper products. It is sliced and flaked for

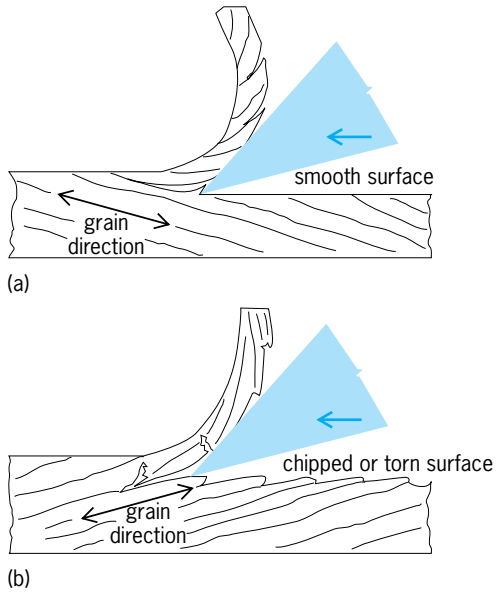

Fig. 3. Cutting direction and wood grain. (*a*) Cutting with the grain leaves a very smooth surface. (*b*) Cutting against the grain leaves a chipped surface. (*After R. B. Hoadley, Understanding Wood, Taunton Press, 1981*)

particleboard products, including wafer boards and oriented strand boards. Whether made from waste products (sawdust, planer shavings, slabs, edgings) or roundwood, the individual particles generally exhibit the anisotropy and hygroscopicity of larger pieces of wood. The negative effects of these properties are minimized to the degree that the three wood directions (longitudinal, tangential, and radial) are distributed more or less randomly. *See* WOOD PRODUCTS.

Charles J. Gatchell

Bibliography. J. G. Haygreen and J. L. Bowyers, *Forest Products and Wood Science: An Introduction*, 3d ed., 1996; R. B. Hoadley, *Understanding Wood: A Craftsman's Guide to Wood Technology*, 2d ed., 2000; U.S. Department of Agriculture, *Dry Kiln Operator's Manual*, Agr. Handb. 188 (rev.), 1991; U.S. Department of Agriculture, *Wood Handbook: Wood as an Engineering Material*, Agr. Handb. 72 (rev.), 1974.

Wood products

Products, such as veneer, plywood, laminate of products, particleboard, waferboard, pulp and paper, hardboard, and fiberboard, made from the stems and branches of coniferous (softwood) and deciduous (hardwood) tree species. The living portion of the tree is the region closest to the bark and is commonly referred to as sapwood; the dead portion of the tree is called heartwood. In many species, especially hardwoods, the heartwood changes color because of chemical changes in it. The heartwood of walnut, for example, is dark brown and the sapwood, almost white.

Wood is one of the strongest natural materials for its weight. A microscopic view reveals thousands of hollow-tubed fibers held together with a chemical called lignin. These hollow-tubed fibers give wood its tremendous strength for its light weight. These fibers, after the lignin bonding material is removed, make paper. In addition to lignin, wood is composed of other chemicals, including cellulose and hemicellulose. *See* CELLULOSE; LIGNIN; WOOD ANATOMY.

Solid wood products. This category includes lumber products, veneer, and structural plywood.

Lumber. Most small log sawmills try to maximize value and yield by automating as much of the manufacturing process as possible. The basic process of cutting lumber involves producing as many rectangular pieces of lumber as possible from a round tapered log. There are only a few sawing solutions, out of millions possible, that will yield the most lumber from any given log or larger timber.

The first stage of the process involves bucking (cross-cutting) the logs into sawmilllength logs (8–24 ft or 2.4–7 m). If the tree-length logs (which can reach a maximum length of 65–70 ft or 20–21 m) are longer than what can be transported over the roadways, the bucking is done in the forest. Otherwise, the bucking is done at the sawmill. In many cases, the logs are electronically scanned for geometric shape and, based on lumber market values, the logs are bucked to optimum length by computer. After bucking, the logs are ready for primary breakdown, or cutting on the first, or headrig, sawing machine. The goal of the headrig is to cut the log into as few pieces as possible and pass them to a machine center downstream in the flow. The log is first scanned for geometric shape, and a computer determines how best to saw it. The log is automatically positioned, and the headrig saws are automatically set. The log then is processed into a combination of side boards and a larger center cant which is cut further downstream. In some areas, a headrig may consist of chipping devices that convert the rounded outside strips of log directly into chips for paper instead of cutting the strips with a saw. After processing through the headrig, the side boards are processed through a board edger. This computerized machine removes the rounded edges (wane) from the sides of the board. The side boards are scanned by lasers and precisely edged by circular saws.

The center cants from the logs are processed through a computerized sawing machine called a rotary gang saw; the machine uses circular saws, and there are up to eight saws spaced evenly apart on a central rotating shaft (arbor). These machines are among the most accurate in the sawmill. Because the center cant makes up a very large proportion of the volume of the log, accuracy is very important.

Following edging and cant processing, the lumber is trimmed to accurate length. A computerized trimmer scans the boards for shape and surface defects and then trims them minimally. Some boards may be reprocessed through a resaw prior to trimming. The resaw uses a band saw to remove any irregular shapes from lumber. Following the trimming operation, the lumber leaves the sawmill, where it is processed through an automated sorting and stacking machine. Each piece of lumber is electronically measured and transported to a bin where boards of similar dimension are stored. When enough lumber of the same size is accumulated, the lumber is stacked for air drying, kiln drying, or planing. The planing process puts a smooth surface on the faces and edges of the board. Following planing, the lumber is graded for structural or appearance characteristics, then packaged for shipment.

Veneer and structural plywood. Structural plywood is constructed from individual sheets of veneer, often with the grain of the veneer in perpendicular directions in alternating plies. The most common construction is three-, four-, and five-ply panels. The alternating plies give superior strength and dimensional stability.

Plywood is used in many applications, including roof sheathing, siding, and floor underlayment. The veneer is rotary peeled from short logs (called blocks) which are slightly over 8 ft (2.4 m) in length. The process begins by bucking the logs into plywood block lengths (8–9 ft or 2–3 m). This process, as with sawmilling, may be computerized. Following bucking, the blocks are usually steam conditioned prior to peeling so that the wood does not split and the knots are softened.

The plywood blocks are electronically scanned and precisely positioned in the veneer lathe. The lathe peels the veneer from the log with a veneer knife. The veneer leaves the lathe and proceeds to a computer automated veneer clipper that clips out the defects (such as knots and splits). The types of veneer produced by the clipper are full, 4 × 8 ft (1.2 × 2.4 m) sheets; half, 2 × 8 ft (0.6 × 2.4 m) sheets; and strip, less than 2 ft (0.6 m) in width. The full and half sheets are automatically stacked, and the strip is hand stacked and stored to await drying.

The plywood veneer is segregated into heartwood and sapwood prior to drying. The sapwood has much more moisture in it, and requires longer drying times than heartwood. Drying times range 5–15 min. After drying, the veneer is again stacked to await panel lay-up and gluing. The individual pieces of veneer are then either manually or semiautomatically processed through a lay-up line. The lay-up line places the veneer in its proper orientation and applies adhesive to each piece. Often, the cross plies are placed on the line manually.

After the lay-up process, the panels are loaded into a hot press where the temperature is raised to polymerize the adhesive. The panels are removed from the press and proceed through a patching process where defects are routed out of the panels and a synthetic liquid patch material is applied. This patch material also polymerizes because of the heat of the panel or external heaters. Finally, the panels are trimmed up square, and sometimes sanded and grooved for siding. The panels are then graded and packaged for shipment to the customer.

Composite products. Such products include laminated products made from lumber, particleboard, waferboard, and oriented strand board.

Laminated products. Laminated products are composite products, made from lumber, parallel laminated veneer, and sometimes plywood, particleboard, or other fiber product. The most common types are the laminated beam products composed of individual pieces of lumber glued together with a phenol resorcinol–type adhesive. Laminated beams are constructed by placing high-quality straight-grained pieces of lumber on the top and bottom, where tension and compression stresses are the greatest, and lower-quality lumber in the center section, where these stresses are lower.

Another form of composite beam is constructed from individual members made from parallel laminated veneer lumber. This type of lumber has the advantage that it can be made into any length, thus creating beams to span large sections. These beams can be made entirely of parallel laminated veneer lumber, or with the top and bottom flange made from parallel laminated veneer lumber, and the center web made from plywood or flakeboard. Some of these products use solid lumber for the flange material and resemble an I beam. All structural laminated products have the advantage over lumber in that much of the natural variation due to defects is removed, and wood structural members can be made much larger than the typical 2 × 12 in. (51 × 305 mm) lumber product.

Particleboard. Nonstructural particleboard is another type of composite product that is usually made from sawdust or planer shavings. It is sometimes made from flaked roundwood. This type of particleboard is one of the most widely used forms of wood product. Often hidden from view, it is used as a substrate under hardwood veneer or plastic laminates. It is commonly used in furniture, cabinets, shelving, and paneling. Particleboard is made by drying, screening, and sorting the sawdust and planer shavings into different size classifications. This particleboard (furnish) is then placed in an adhesive blender where a urea-formaldehyde adhesive is applied. Following blending, the furnish is transported to a forming machine where the particleboard mat of fibers is created. Commonly, the forming machine creates a loose mat consisting of fine particles on the outside and coarser particles on the inside. After the mat is formed, it is transported to a multiopening hot press where it is compressed into the final panel thickness and the adhesive is cured. Following pressing, the large panels are cooled, trimmed, and cut into smaller finished sizes. They are next sanded, and graded prior to shipment to furniture and other manufacturers.

Waferboard and oriented strandboard. Waferboard and oriented strandboard are structural panels made from flakes or strands, and are usually created from very small trees. Unlike nonstructural particleboard, waferboard is designed for use in applications similar to those of plywood. Waferboard and oriented strandboard can have flakes or strands oriented in the same direction, thus giving the board greater strength in the long axis. The most common type of waferboard is made from randomly oriented flakes. The process of making waferboard is similar to that of particleboard except that the wood material used is in the form of flakes instead of particles. The adhesive is a phenol formaldehyde which is usually applied in a powdered form.

Fiber products. The most common fiber products result from pulping processes that involve the chemical modification of woodchips, sawdust, and planer shavings. Such products include pulp and paper, hardboard, and fiberboard.

Pulp and paper. Paper making begins with the pulping process. Pulp is made from wood chips created in the lumber manufacturing process, small roundwood which is chipped, and recycled paper. The fibers in the chips must be separated from each other, by mechanically grinding the fibers or chemically dissolving the lignin from them. The most common chemical processes are sulfite and sulfate (kraft). Following the pulping process, the fibers are washed to remove pulping chemicals or impurities. In some processes (for example, writing papers), the fibers are bleached.

During the beating and refining process, the fibers are flattened to create more fiber bonding sites. This process is necessary to give paper added strength. The fibers are suspended in a water slurry and processed through a Fourdrinier machine. During the wet stage, the water is removed from the pulp, and

the resulting mat of fibers is transported on a screen to the wet pressing and drying section, where the mat is passed through a series of hot rollers. Dry, finished paper emerges from the end of this section, and it is placed in rolls for further manufacture into paper products. *See* PAPER.

Hardboard. Hardboard is a medium- to high-density wood fiber product made in sheets from $^1/_{16}$ to $^1/_2$ in. (1.6 to 12.7 mm). Hardboard is used in furniture, cabinets, garage door panels, vinyl overlaid wall panels, and pegboard. It is made by either a wet or dry process. Fibers are produced from chips through a thermomechanical process which retains the lignin that is used to bond the fibers together. During the wet process, the pulp is mixed with water and placed on a screen where the water drains away and the pulp mat is retained. The mat is placed in a prepress, where the excess water is removed. The fiber mat, still on the screen, is then placed in a hot press where the panel is formed. During the dry process, air is used instead of water to form the mat prior to pressing. After the hardboard is pressed, it is tempered by using an oil soak process or high-temperature exposure.

Medium-density fiberboard. Medium-density fiberboard is used in many of the same applications where particleboard is used. It can be used in siding and is especially well suited to cabinet and door panels where edges are exposed. Unlike particleboard which has a rough edge, medium-density fiberboard has a very fine edge that can be molded very well. Medium-density fiberboard is produced in much the same way that dry-processed hardboard is produced in its early stages. The chips are thermomechanically pulped or refined prior to forming into a dry mat. Following refining, medium-density fiberboard is produced in a fashion similar to that of particleboard. The dry pulp is sprayed with adhesive (usually urea-formaldehyde or phenol-formaldehyde) and formed into a dry mat prior to pressing in a multiopening hot press. After the panels have been formed, they are cooled and cut to smaller final product sizes prior to shipment. *See* WOOD PROCESSING. Terence D. Brown

Bibliography. R. F. Baldwin, *Plywood Manufacturing Practices*, 1984; C. J. Biermann, *Essentials of Papermaking*, 1993; T. D. Brown, *Quality Control in Lumber Manufacturing*, 1982; J. G. Haygreen and J. L. Boywer, *Forest Products and Wood Science*, 3d ed., 1996; T. M. Maloney, *Modern Particleboard and Dry Process Fiberboard*, 1972, updated 1993; E. M. Williston, *Lumber Manufacturing*, 1988.

Wood properties

Physical and mechanical characteristics of wood which are controlled by specific anatomy, moisture content, and to a lesser extent, mineral and extractive content. The properties are also influenced by wood's directional nature, which results in markedly different properties in the longitudinal, tangential, and radial directions or axes (see **illus.**). Wood properties within a species vary greatly from tree to tree and within a single axis. The physical properties (other than appearance) are moisture content, shrinkage, density, permeability, and thermal and electrical properties.

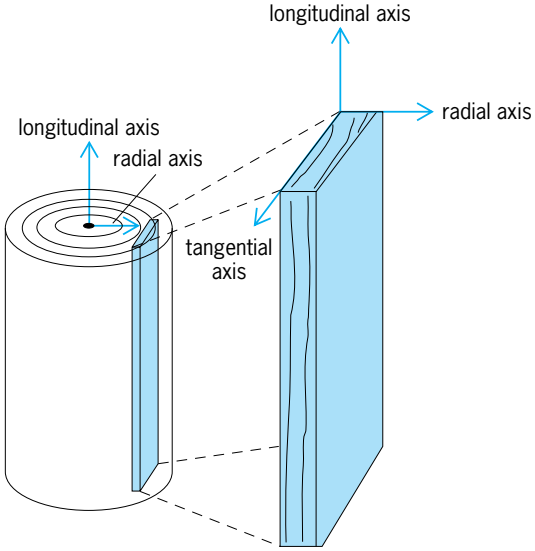

Coordinate system for a log or section of a tree (left) and a piece of lumber from the tree (right).

Moisture content. Moisture content is a major factor in the processing of wood because it influences all physical and mechanical properties, and durability and performance during use. Normal in-use moisture content of processed wood that has been dried ranges 8–13%. Water is held in wood in three ways: as water chemically bonded to hydroxyl groups in the wood substance of the cell wall, or as either free water or water vapor in the cell cavities. When the free water is removed from the cell cavities and the cell wall is saturated to its maximum, the wood is said to be at its fiber saturation point, which ranges 25–35% moisture content for most species. Wood can chemically adsorb or desorb moisture as water vapor below the fiber saturation point. Because of the capillary structure of the cavities, wood can also mechanically absorb liquid water.

Moisture content for wood is expressed on either a fractional or percentage basis. Moisture content is defined as the ratio of the mass of water contained in the wood to the mass of the same sample of dry wood. For hardwoods in the green state, the moisture content of sapwood and heartwood is approximately equal, ranging 45–150% according to species. The moisture content of softwoods in the green condition differs for heartwood and sapwood, with heartwood ranging 30–120% and sapwood 100–250% according to species. Thus, the average moisture content of undried softwoods is affected by the percentages of heartwood and sapwood in the sample.

Shrinkage. Shrinkage occurs when wood loses moisture below the fiber saturation point. Above that point, wood is dimensionally stable. The amount of the shrinkage depends on its direction relative to grain orientation and the amount of moisture lost

below the fiber saturation point. Volumetric shrinkage is about 10–19% depending on species. Wood shrinks significantly more in the radial and tangential directions than in the longitudinal direction.

Density. The density of wood is determined by the amount of cell wall substance and the volume of voids caused by the cell cavities (lumens) of the fibers. Density can vary widely across a growth or annual ring. For species with distinct growth rings, the earlywood (light bands on the cross section of wood) consists of fast-growing, relatively thin-walled cells or fibers (that is, less dense wood), and the latewood is composed of thick-walled, slow-growing cells or fibers (that is, more dense wood). The percentage of earlywood and latewood in each growth ring determines the overall density of a wood sample.

Permeability. Permeability is a measure of the flow characteristics of a liquid or gas through wood as a result of the total pressure gradient. Permeability is influenced by the anatomy of the wood cells, the direction of flow (radial, tangential, and longitudinal), and the properties of the fluid being measured. The longitudinal flow is greater than the flow in either the radial or tangential directions. The longitudinal flow is relatively unrestricted through the cavities of the cells, while the tangential and radial flows are affected by the quantity, size, and condition of pits in the cell wall. Permeability is also affected by the species, by whether the wood is sapwood or heartwood, and by the chemical and physical properties of the fluid.

Thermal properties. The primary thermal properties of wood are conductivity, specific heat, and coefficient of thermal expansion. When wood is ovendry it is a poor conductor of heat, whereas when wood has a high moisture content it is an excellent conductor. The conductivity of wood is determined by density, moisture content, and direction of conduction. Thermal conductivity in the transverse directions (radial and tangential) is approximately equal. Conductivity in the longitudinal direction is greater than in the transverse directions. The greater conductivity in the longitudinal direction is due to cell walls which are less interrupted than in the transverse directions.

For most processing operations, the dominant heating direction is transverse. Thermal conductivity is important to wood processing because heating—whether for drying, curing, pressing, or conditioning—is an integral step.

Specific heat of wood is dependent on moisture content and, to less extent, on temperature. *See* SPECIFIC HEAT.

Electrical properties. Dry wood is an excellent insulator. By measuring wood's electrical resistance, electrical moisture meters accurately determine the moisture content of wood in the 5–25% range. Two other electrical properties of interest are the dielectric constant and the dielectric power factor for alternating current. These dielectric properties are dependent on density, moisture content, frequency of current, grain orientation, and temperature. The power factor is a measure of the stored energy that is converted to heat. This power factor is also affected by frequency, moisture content, and temperature. The dielectric character of wood is used as a working principle in moisture meters to measure the moisture content of wood over its entire moisture range.

Mechanical properties. The mechanical properties of wood include elastic, strength, and vibration characteristics. These properties are dependent upon species, grain orientation, moisture content, loading rate, and size and location of natural characteristics such as knots.

Because wood is an orthotropic material, it has unique and independent mechanical properties in each of three mutually perpendicular axes—longitudinal, radial, and tangential. This orthotropic nature of wood is interrupted by naturally occurring characteristics such as knots that, depending on size and location, can decrease the stiffness and strength of the wood.

Elastic properties. Wood is both an elastic and plastic material. Elasticity manifests itself during loading and at moisture contents and temperatures that occur in most service uses of wood. Wood is primarily elastic to its proportional limit during loading at room temperature, at which point the wood becomes plastic and starts an irrecoverable flow. The elastic stiffness or modulus of elasticity of wood is dependent on grain orientation, moisture content, species, temperature, and rate of loading.

The stiffness of wood in the longitudinal (fiber) direction is utilized in the manufacture of composite products such as oriented strand board, in which the grain or fiber direction is controlled. The stiffness of wood may be predicted by using the ovendry density of the wood and a correction factor for moisture content. In industrial applications, the measured stiffness of lumber can be used to predict its strength rating nondestructively. *See* ELASTICITY.

Strength. The strength of wood, like its elastic properties, is dependent upon rate of loading, species, moisture content, orientation, temperature, size and location of natural characteristics such as knots, and specimen size. The strength of individual wood fibers in the longitudinal direction can be significantly greater than that of larger samples with their complex anatomy and many defects. As with stiffness, the excellent strength characteristics of wood in the direction of the fiber can be maximized during the manufacture of wood composites by controlling fiber alignment.

Vibration. Damping and sound velocity are two primary vibration phenomena of interest in structural applications. Damping occurs when internal friction dissipates mechanical energy as heat. The damping characteristic of wood results in quick dissipation of all energy for most frequencies of vibration. Because of its high damping capacity, wood is an excellent material in floors and other structural components, and it is excellent for resisting the forces resulting from earthquakes.

The velocity of a sound wave through wood can be used to estimate mechanical stiffness and strength:

the higher the velocity, the higher the stiffness and strength. Like other properties of wood, the velocity of sound along the three principal axes differs. The three directional velocities are proportional to the wood's stiffness values in these directions. Sound velocity in the longitudinal direction is two to four times greater than in the transverse directions. *See* WOOD ANATOMY. Jame B. Wilson

Bibliography. J. Bodig and B. A. Jayne, *Mechanics of Wood and Wood Composites*, 1982; Forest Products Laboratory Staff, *Wood Handbook: Wood as an Engineering Material*, 1987, revised 2000; J. F. Siau, *Transport Processes in Wood*, 1984; C. Skaar, *Wood-Water Relations*, 1988.

Woodward-Hoffmann rule

A concept which can predict or explain the stereochemistry of certain types of reactions in organic chemistry. It is also described as the conservation of orbital symmetry, and is named for its developers, R. B. Woodward and Roald Hoffmann. The rule applies to a limited group of reactions, called pericyclic, which are characterized by being more or less concerted (that is, one-step, without a distinct intermediate between reactants and products) and having a cyclic arrangement of the reacting atoms of the molecule in the transition state. Most pericyclic reactions fall into one of three major classes, examples of which will illustrate the use of the rule. *See* PERICYCLIC REACTION.

Electrocyclic reactions. These reactions are defined as the interconversion of a linear π system, containing n π electrons, and a cyclic molecule containing $(\eta - 2)$ π electrons which is formed by joining the ends of the linear molecule. This is exemplified by the thermal ring opening of a cyclobutene to a butadiene, reaction (1).

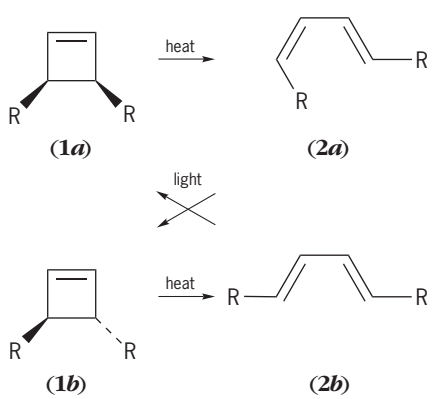

(1)

The reaction is stereospecific in that (**1*a***; R substituents cis) gives only (**2*a***) and none of (**2*b***). Conversely, cyclobutene (**1*b***) gives (**2*b***) but not (**2*a***). This mode of reaction is termed conrotatory, since both R groups rotate in the same direction (clockwise or counterclockwise when viewed edge-on) in going from reactant to product. On irradiation with ultraviolet light the butadienes recyclize to cyclobutenes. This reaction is disrotatory since (**2*a***) → (**1*b***) and (**2*b***) → (**1*a***). Similar stereospecificity is observed in the cyclization of substituted hexatrienes (**3**) and octatetraenes (**5**) to cyclohexadienes (**4**) and cyclooctatrienes (**6**), respectively [reactions (2) and (3)].

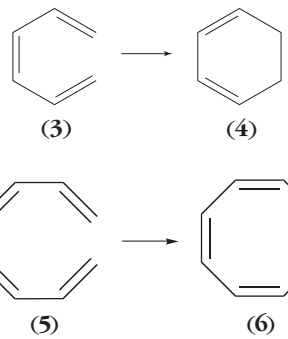

(2)

(3)

To see how these results are explained requires a knowledge of the nature, specifically the symmetry, of the molecular orbitals most involved in the reaction, such as the relevant occupied orbitals of butadiene and cyclobutene (**Figs. 1** and **2**). The signs (+ or −) in the orbitals indicate the phase of the wave function in that region of space. *See* MOLECULAR ORBITAL THEORY.

The orbitals are labeled S or A according to their symmetry with respect to rotation by 180° about an axis in the plane of the molecules (Fig. 1). This

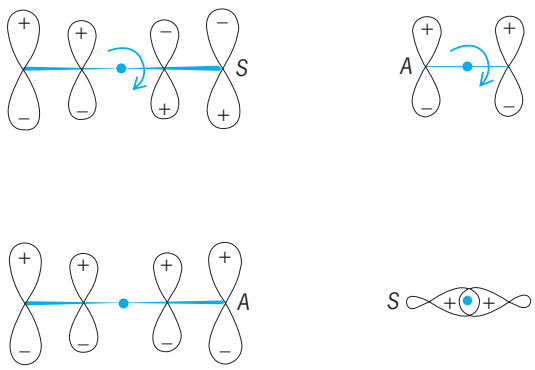

Fig. 1. Conrotatory cyclization, butadiene: axial symmetry.

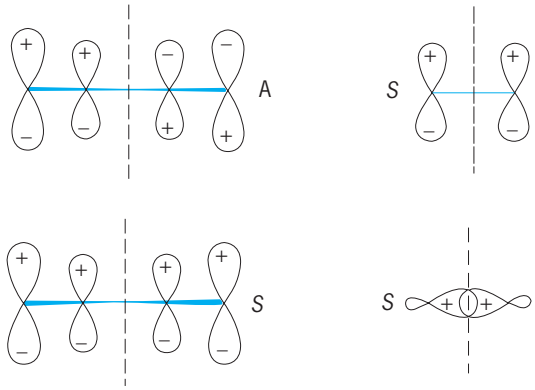

Fig. 2. Disrotatory cyclization, cyclobutene: planar symmetry.

Electrocyclic reactions of polyenes

n	Thermal reaction	Photochemical reaction
$4q$ ($q = 4, 8, \ldots$)	Conrotatory	Disrotatory
$4q + 2$ ($q = 2, 6, \ldots$)	Disrotatory	Conrotatory

element of symmetry is maintained throughout the conrotatory cyclization. The same orbitals and their symmetry with respect to reflection through the plane that bisects the molecules holds for cyclobutene (Fig. 2). The molecule retains this element of symmetry in its disrotatory cyclization. The fact that in the conrotatory mode of cyclization both butadiene and cyclobutene have the same number of symmetric occupied orbitals (one each) shows that the reaction can take place; that is, it is symmetry-allowed. In the disrotatory mode there are two occupied symmetric orbitals in cyclobutene, but only one in butadiene. Thus in the latter path, orbital symmetry is not conserved, and the reaction is said to be symmetry-forbidden.

The general rules for electrocyclic reactions of polyenes with n π electrons are summarized in the **table**, where q is an integer.

Cycloaddition reactions. The simplest example of this type of pericyclic reaction is the combination of two unsaturated molecules, ends to ends, to form a cyclic molecule with four fewer π electrons, for example, reactions (4) and (5). When these

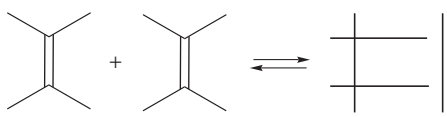

(4)

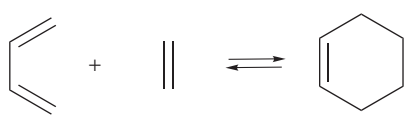

(5)

reactions are concerted, the Woodward-Hoffmann rule correctly predicts the experimentally observed stereochemistry. A generalization which can be made is that cycloadditions of two or more unsaturated molecules are most facile when there are a total of $(4q + 2)$ π electrons ($q > 0$). Cycloadditions (or the reverse reactions) involving $4q$ electrons ($q > 0$) are symmetry-allowed, but the stereochemistry of the required transition state is usually sufficiently strained that the reaction proceeds in two steps.

Sigmatropic reactions. This type of reaction is best defined by reference to the examples shown in reactions (6)–(8). Note that in each reaction there is no

(6)

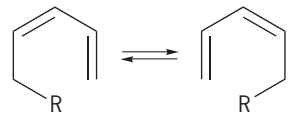

(7)

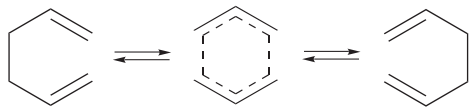

(8)

overall change in the number of σ or π bonds. The name sigmatropic is derived from the fact that there is a change in the location of the σ bond connecting the upper and lower fragments of the molecule. The individual reactions are classified in terms of the number of atoms from each fragment in the cyclic transition state. Thus the above examples are sigmatropic reactions of order [1,3], [1,5], and [3,3], respectively. All of the above and many other such reactions are known, and in each case the Woodward-Hoffmann rule predicts the experimentally observed stereochemistry.

There are several other miscellaneous pericyclic reactions to which orbital symmetry theory applies. Among these are the cheletropic reactions, for example, (9) and the Ene reaction (10). Although the

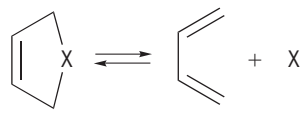

(9)

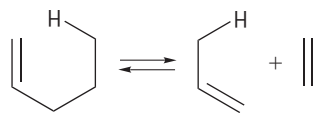

(10)

examples have shown simple hydrocarbons as reactants and products, the rules apply also to molecules containing atoms other than carbon and hydrogen. A methylene group may be replaced by an oxygen or nitrogen atom, for example, and the conclusions from orbital symmetry are unaltered.

It should be recognized that the primary explanation for the occurrence of any chemical reaction is found in either the strengths of the reactant's or the products' bonds or both. If the reaction is pericyclic, the conservation of orbital symmetry can affect the rate of the reaction. Thus if two reactions or two mechanisms for the same reaction are possible, the one which is symmetry-allowed will be much faster than the one which is symmetry-forbidden. If, for reasons unrelated to orbital symmetry, no symmetry-allowed reaction is possible for an apparently pericyclic reaction, the reaction may be observed to proceed via a forbidden path. Usually, however, it is found on careful study that the reaction avoids orbital symmetry control by proceeding in two steps via a noncyclic intermediate; that is, it is not pericyclic. *See* ORGANIC REACTION MECHANISM; STEREOCHEMISTRY.

David L. Dalrymple

Bibliography. S. N. Ege, *Organic Chemistry: Structure and Reactivity*, 4th ed., 1999; E. A. Halevi, *Orbital Symmetry and Reaction Mechanism: The OCAMS View*, 1992; R. Hoffmann and R. B. Woodward, The conservation of orbital symmetry, *Account. Chem. Res.*, 1:17, 1968; P. Laszlo, *Organic Reactions: Simplicity and Logic*, 1996; R. B. Woodward and R. Hoffmann, *The Conservation of Orbital Symmetry*, 1970.

Woodworking

The shaping and assembling of wood and wood products into finished articles such as mold patterns, furniture, window sashes and frames, and boats. The pronounced grain of wood requires modifications in the working techniques when cutting with the grain and when cutting across it. Five principal woodworking operations are sawing, planing, steam bending, gluing, and finishing. To shape round pieces, wood is worked on a lathe. *See* TURNING (WOODWORKING).

Sawing. Wood is sawed by cutting or splitting its fibers by the continuous action of a series of uniformly spaced teeth alternately staggered to move in closely parallel work planes. Action of the cutting teeth produces a path or kerf of uniform width through the workpiece from which the fibers have been severed and removed. Sawing across the grain or cell structure of the wood is called crosscutting. Cutting parallel with the grain of the piece is referred to as ripping. Saw teeth are bent alternately to the left and right to provide clearance for the blade. Some blades include straight raker teeth for cleaning fibers from the cut.

Woodcutting saws may be classed as either handsaws or power-operated saws. Either group consists of numerous types and designs.

Handsaws. Each type of handsaw is designed to accomplish one specific type of sawing operation most effectively.

Crosscut handsaws are made with about 8 teeth per inch of length. Ripsaws for cutting with the grain usually have about $5^1/_2$ teeth per inch (2 per centimeter). Finetooth saws for finishing or cabinetwork may have as many as 10–16 teeth per inch (about 4–6 per centimeter) of length.

A backsaw is a fine-tooth saw with its upper edge stiffened to ensure straight cuts. Keyhole saws with their narrow, tapered blades are used for cutout work where sharp turns are required. A compass saw has a handle with several attachable blades of varying widths, making it suitable for a variety of work. Coping saws have narrow blades usually about $^1/_8$ in. (0.3 cm) wide. The blade is held taut in a frame which is equipped with a handle. The narrow blade and high-backed frame make the saw suitable for shaping or cutout work.

Power saws. Power-operated woodworking saws are usually combined with auxiliary equipment that enables them to perform various sawing operations.

Bench or circular saws are the common woodworking type of power saw (**Fig. 1**). Depending on construction, either a power arbor or worktable may be raised, lowered, or tilted. The saw may be used for crosscutting, ripping, or resawing, and for beveled as well as tapered cuts. Molding cutters along with those designed for making rabbet, tendon, and dado joints are also used. Accessories for sanding, buffing, and polishing are available for most models.

Band saws are basically a flexible band of steel running over two vertical pulleys (**Fig. 2**). The band or blade has teeth on one side and is operated under tension. The wide distance or throat between the cutting portion of the blade and the rear blade guide and support arm adapt the band saw for cutout work or sawing on large flat pieces.

Scroll saws are used for work similar to that performed by the band saw. The continuous-band type of blade is replaced by a short, vertically reciprocating blade (**Fig. 3**).

Radial saws have their circular blade mounted above the worktable. The blade and motor are suspended from an overarm that allows travel across the workpiece: a pivot permits cuts to be taken at any angle (**Fig. 4**). Usually the saw may be raised and lowered as well as tilted at an angle. Crosscutting, ripping, mitering, and beveling may be performed; accessories and attachments permit other circular cutting tool operations such as dadoing, molding, drilling, and sanding.

Portable handsaws consist of a circular blade and electric motor plus the necessary frame, handles, baseplate, and guards. An electric cord of reasonable length permits the saw to be manually moved and positioned for the desired cut. Some models may be fastened to a special frame or table.

Portable saber saws are compact units consisting of an electric motor, a straight saw blade driven by a

Fig. 1. Bench circular saw with tilting arbor is used for parting or slotting, and can make cuts as long as working space permits. (*Delta International Machinery Corp., Pittsburgh*)

Fig. 2. Narrow bandsaw, a flexible band of steel, can make curved as well as straight cuts even in thick pieces. (*Delta International Machinery Corp., Pittsburgh*)

Fig. 3. Scroll saw, a short, vertically reciprocating blade, cuts sharp turns in thin pieces. (*Delta International Machinery Corp., Pittsburgh*)

Fig. 4. Radial saw allows wood to be clamped in position while it is being cut. (*Delta International Machinery Corp., Pittsburgh*)

reciprocating mechanism, handle, baseplate, electric cord, and other necessary parts. The lightness of the saw and its narrow blade fastened at only one end make it adaptable to many types of cutting including cutout work and shaping.

Planing. Flat or uniformly contoured surfaces of wood are roughed down, smoothed, or made level by the shaving and cutting action of a wide-edged blade or blades. Planing may be accomplished either manually or by power-operated tools.

Hand planes. Manually operated planes are classified as either bench or block types. A bench plane is used for shaving with the grain of the wood, whereas a block plane is designed for cutting across the grain. A block plane is usually small; bench planes vary in size and type. The common bench types are the smoothing, jack, fore, and jointer.

The smoothing plane, 5.5–10 in. (14–25 cm) long, is best suited for smoothing small areas. The

Fig. 5. Six-inch-long (15-cm) bed, jointer planer found in woodworking shops is used for surfacing, rabbeting, beveling, tapering, molding, and cutting round tenons. (*Delta International Machinery Corp., Pittsburgh*)

somewhat larger jack plane may be used for roughing down or leveling. Fore and jointer planes are still larger in size, with the latter being approximately 22–24 in. (55–60 cm) long. It is used to plane long surfaces.

Special hand planes are the rabbet plane, used to cut recesses for rabbet joints; the model maker's, used to remove excess wood from a curved surface; the scrub or roughing plane, which has heavy, rounded blades making it suitable for cleaning up rough boards; and the circular plane, with a flexible steel bottom that may be adjusted to fit a curved surface.

Power planes. Power-operated planes vary in size and design according to the application and type of work handled. The planer usually found in the woodworking shop is the jointer, frequently called a jointer planer (**Fig. 5**).

The jointer is designed with its cutting blades or knives fastened in a rotating cutterhead. The lengths of the knives classify the machine as to the width of board that it can surface.

Two tables, one in front of the cutterhead and one behind it, support the workpiece as it is pushed through the path of the knives. The front table is set lower than the highest point on the arc of the rotating knives by the amount to be planed from the board. The rear table must be aligned exactly with the high point of the knives so that the workpiece will not pass through the cutting path at an angle.

A tilting fence or guide is provided along one side of the tables. The opposite side of the front table has a ledge which is used to support the workpiece for rabbeting cuts.

Steam bending. Wooden members are bent or formed to a desired shape by pressure after they have been softened or plasticized by heat and moisture. If thick pieces of wood are to be bent to a permanent shape without breaking, some form of softening or plasticizing such as steaming is necessary.

When a piece of wood is bent, its outer or convex side is actually stretched in tension while its concave side is simultaneously compressed. Actually, plasticized wood can be stretched but little. It can, however, be compressed a considerable amount. When a piece of plasticized wood is successfully bent, the deformation is chiefly compression distributed almost uniformly over the curved portion. Curvature results from many minute folds, wrinkles, and slippages in the compressed area.

Steaming. Although soaking wood in water softens it somewhat, a combination of heat and moisture can produce a degree of plasticity approximately 10 times that of dry wood at normal temperatures. Wood need not be steamed to its maximum plasticity in order to be bent. Wood steamed at atmospheric pressure bends in most cases, as does wood steamed at higher pressures. Higher pressures also tend to overplasticize the wood: this results in an increased number of bending failures. Treatment of wood with boiling water has approximately the same effect as saturating it with steam at atmospheric pressure. The boiling water treatment is usually employed only when a portion of a wooden piece requires softening.

Dry wood that has a moisture content of 12% or less must have its moisture increased to approximately 15% to make it suitable for moderate bends. If more severe bends are required, additional moisture must be added to the surface areas. Wood that already has a moisture content of 20–25% needs no further saturation, even for severe bending.

The time required for steaming is directly related to the amount of moisture already present in the wood. In general, dry stock is steamed 1 h per inch (0.4 h/cm) of thickness and green stock, 0.5 h per inch (0.2 h/cm) of thickness. Steaming is usually done in a closed retort suitable for zero-pressure or low-gage-pressure steam. Steam should enter the retort or steam box through water standing in the bottom so that steam in the box will be wet or saturated.

Wood-bending methods may be classed as made without end pressure (free bends) or made with end pressure. On thick pieces, only slight curvatures are feasible by free bending. Bending with end pressure is necessary to obtain the required compression and to prevent tensile failures in moderate or severe bends. The most common method of bending with end pressure is by means of a metal strap fitted with end blocks or clamps. The strap is placed against the convex side of the piece, and bending pressure is applied at the end of the strap by some suitable means (**Fig. 6**). The end blocks must be applying end pressure simultaneously on the wood to prevent tensile stress and also to supply the necessary compression.

Other devices used for bending plasticized wood are variable-position rollers, hot-plate form presses, and special mechanical devices.

When a workpiece is removed from the bending device, it has a tendency to spring back somewhat. This is counteracted by holding the piece in position until it has dried or set. Frequently springback is compensated for by overbending.

Fig. 6. Bent boat rib is prepared for drying with end pressure applied by the strap and the rods placed against the convex side of the piece. Wood stays assist in holding the work in position. (*Forest Products Laboratory*, *Madison*, *Wisconsin*)

The selection of a species of wood to use as a moderately bent member is governed primarily by its suitability and availability. However, if a severe bend is required, the wood must be selected chiefly for its bending qualities. In general, hardwoods are of better bending quality than softwoods. Pieces should be of fairly straight grain and free of knots or other defects. White and red oak, hickory, elm, ash, beech, birch, maple, mahogany, walnut, and sweet gum are species commonly used for bending.

Adhesive bonding and gluing. Wood pieces may be fastened together by the adhesive qualities of a substance that sets or hardens into a permanent bond. Adhesives for wood are of two principal types, synthetic and natural-origin. The term glue was first applied to bonding materials of natural origin, while adhesive has been used to describe those of synthetic composition. The terms are used interchangeably, but adhesive better covers all types of materials in use. *See* ADHESIVE.

Synthetic adhesives include phenolics, ureas, melamines, polyvinyl resin emulsions, hot melts, epoxies, contacts, mastics, and various combinations of specific adhesives. A hardener or setting agent is usually required to convert synthetic adhesives from liquid to solid. These agents may be furnished separately for addition to the resin before use, or they may be already present in the resin as supplied. Since each type of synthetic adhesive has its own characteristics or requirements for use and application, such as temperature and pressure, and for service conditions, they should be studied and considered on an individual basis before use. Some types require special equipment such as heating facilities and presses, which tend to limit them to commercial use. Some thermosetting resins require temperatures of over 300°F (150°C) and pressures as high as 250 lb/in.^{2} (1.7 megapascals) depending on the wood, which tend to limit them to commercial use. *See* POLYVINYL RESINS; UREA-FORMALDEHYDE RESINS.

While adhesives of natural origin, such as animal casein, soybean, starch, and blood glues, are used to bond wood in some plants and shops, they have been replaced largely by synthetics. Animal glue is probably the natural adhesive most widely used, although casein glue is being used a great deal for structural laminating. As with the synthetic adhesives, the conditions for most suitable use and service of the natural adhesives vary, and should also be considered for use on an individual basis.

Prior to any gluing, the mating surfaces should be clean, smooth, and properly fitted or matched. Surfaces should be machined smooth and true, and be essentially free from rough machine marks, chipped or loosened grain, and surface irregularities.

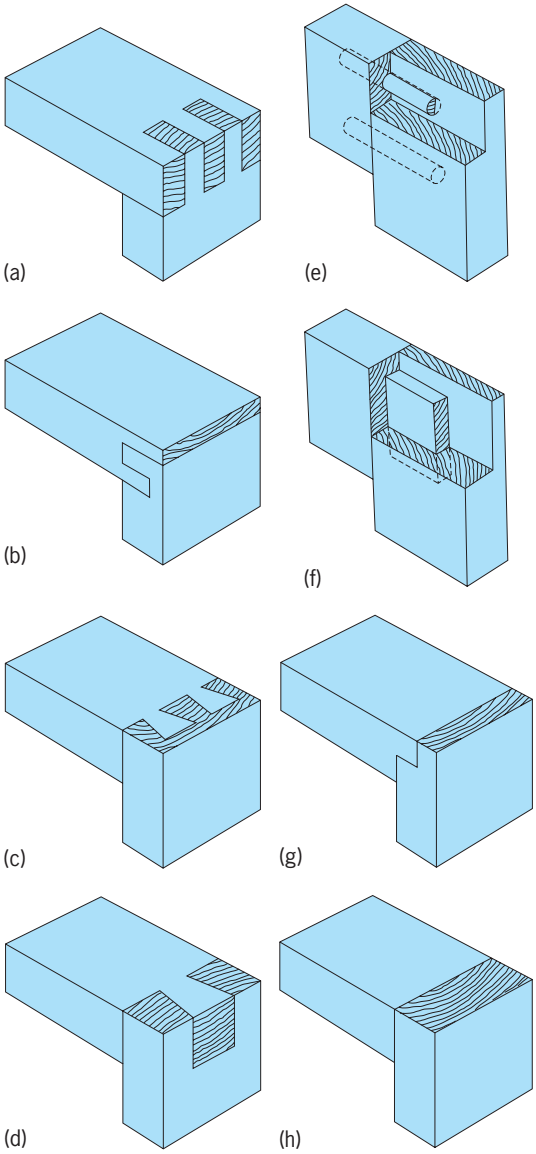

Fig. 7. Various types of corner joints: (*a*) slip or lock corner; (*b*) dado tongue and rabbet; (*c*) blind dovetail; (*d*) dovetail; (*e*) dowel; (*f*) mortise and tenon; (*g*) shouldered corner; (*h*) butt end to side grain. (*U.S. Department of Agriculture*, *Forest Products Laboratory*, *Madison*, *Wisconsin*)

Moisture content of the wood before gluing is important because it affects the quality of the bond and the performance of the glued product in service. Satisfactory adhesion to wood is obtained with most adhesives when the wood moisture content is about 6–17%, and with some glues up to 25%. Glued joints will remain most nearly free from stresses if the moisture content of the parts (when the glue sets) equals the average moisture content that the product will experience in service.

If the end grain surfaces of two pieces of wood are glued together, a butt joint is formed. However, the gluing of square-end butt joints usually does not result in a sufficiently strong and permanent joint to meet the requirements of ordinary or commercial service. Mitered joints (usually cut at a 45° angle with the grain) must also be treated essentially as butt joints for gluing purposes. Seven types of corner joints are shown in **Fig. 7**.

To obtain acceptable strength in pieces spliced together endwise, it is necessary to make a scarf, finger, or other sloped joint (**Fig. 8**). During the gluing operation, end slippage should be prevented to keep parts in proper alignment, and under sufficient and uniform pressure to get maximum joint strength. Even plain scarf joints with a low slope are not as strong as clear wood (of the same quality) in tension parallel to the grain.

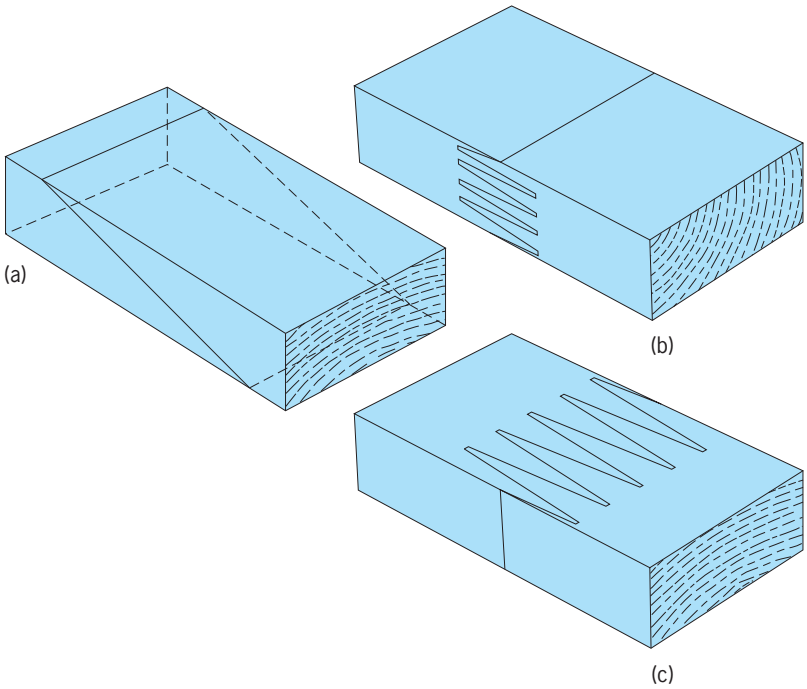

Fig. 8. End joints for splicing lumber: (*a*) scarf joint; (*b*) horizontal finger joint; (*c*) vertical finger joint. (*U. S. Department of Agriculture, Forest Products Laboratory, Madison, Wisconsin*)

Finishing. The finishing operation is the preparation and sealing or covering of a surface with a suitable substance in order to preserve it or to give it a desired appearance. The preparation and conditioning of a surface may include cleaning, sanding, use of steel wool, removing or covering nails and screws, gluing or fastening loose pieces, filling cracks and holes with putty or crack filler, shellacking, and dusting. An inconsequential item with a painted surface does not require the thorough surface preparation that a piece of fine furniture does. The quality of surface conditioning directly affects the end result.

Surface preparation. Wooden surfaces to be painted are cleaned and sanded. Sanding is done with the grain of the wood, using a sandpaper block to keep the surface even, or a power sander (**Fig. 9**). Use of steel wool instead of sandpaper in corners and on rounded surfaces is advisable. Nails and screws should be removed or sunk below the surface, and these holes, as well as any cracks, should be filled with putty or crack filler. New wood is covered with a coating of thinned orange shellac before applying paint. The undercoat tends to prevent the paint from drying unevenly and thus producing flat or glossy spots over knots or other irregularities in the wood.

If the surface is to be varnished, stained, or finished in its natural color, extra care must be taken in sanding, filling in, or smoothing holes and imperfections in the surface. The filler material should be tinted to match the wood. After the filler material has hardened, the entire surface should be sanded with coarse sandpaper. When the rough spots have been removed, sanding should be finished with a No. 1 sandpaper. On fine hardwood surfaces or on furniture from which the varnish has been removed, only fine grades of sandpaper running from No. 0 down to No. 000 should be used. Rubbing must be with the grain, and grit and dust must be removed frequently. If paint remover has been used, the surfaces should be thoroughly cleaned with turpentine and then washed with hot water and soap to remove any wax.

Finish application. Methods used to apply finish coverings to wooden surfaces vary with the substance. In any case, the surface should be wiped clear of dust just prior to applying the material. The

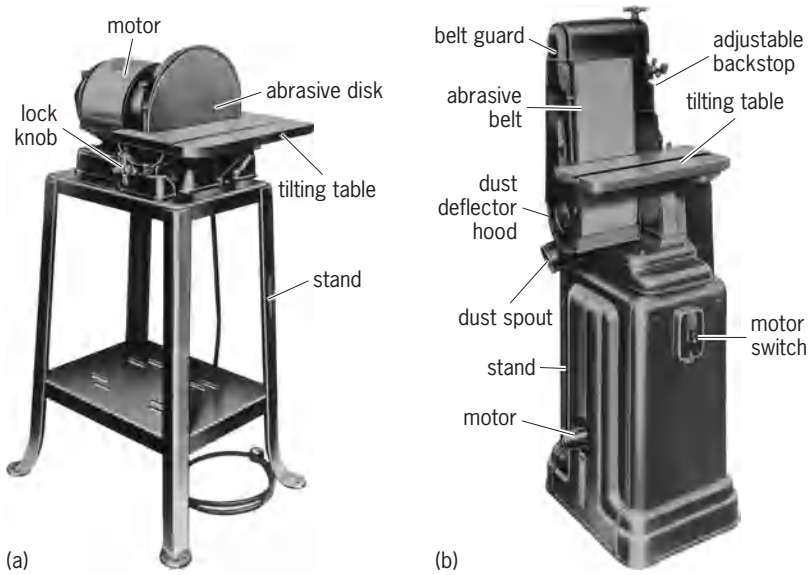

Fig. 9. Types of power sander: (*a*) disk; (*b*) belt. (*Delta International Machinery Corp., Pittsburgh*)

atmosphere and equipment used should be as dust-free as possible.

Paint may be applied by brush or spray as the situation permits. The priming coat and each additional one should be thoroughly dry before applying the next one. On woodwork or painted furniture after each coat prior to the final one, the surface should be sanded lightly to remove any dust or particles stuck to the paint. If a brush is used, the paint should be applied with long even strokes. A thin coat goes on more evenly and tends to prevent the paint from running or forming lumpy spots at the corners and edges.

Stains penetrate the pores of wood but do not fill or close them. Oil stains are the ones most commonly used except on furniture, where water stains are the rule. A few moments after a stain has been applied to a wood surface, any surplus should be wiped off with a lint-free cloth. The time that a stain remains on the wood before being wiped off determines the resulting shade. Although shellac or varnish may be applied over a stain, usually a liquid or paste filler is used to seal the pores in the wood. If varnish is used, two or three coats are generally required.

Varnishes should be applied to a surface quickly and liberally. After most of the varnish in the brush has been applied on an area, it should be spread out evenly. Strokes should be with the grain, and all marks or laps should be brushed out with a nearly dry brush. Varnish used on the first coat should be thinned with about a quart of turpentine to 4 quarts of varnish.

Each coat of varnish should be allowed to dry for several days before applying the next. Sanding should be done with No. 000 sandpaper. Thorough removal of dust is a necessity. Varnishing should be attempted only at temperatures of 65°F (18°C) or above.

Shellac, which is thinner than varnish, dries quickly and must be applied to a surface rapidly. All parts of the surface should be carefully covered with the shellac. Several hours should be allowed for drying between coats. White shellac is used for light woods with a natural finish, while orange shellac is used on all other types of wood. In general, shellac should be used with the same procedure that applies to varnish.

Linseed oil or paste wax is frequently used in finishing furniture. Rubbing may include the use of pumice and alcohol. *See* WOOD PROPERTIES. Alan H. Tuttle

Bibliography. Forest Product Laboratory, *Wood Handbook*, USDA Forest Service, Agr. Handb. 72, rev. August 1974; M. Selbo, *Adhesive Bonding of Wood*, Tech. Bull. 1512, USDA Forest Service, August 1975.

Wool

A textile fiber made of the undercoat of various animals, especially sheep; it may also be obtained from angora, goat, camel, alpaca, llama, and vicuna. Wool provides warmth and physical comfort that cotton and linen fabrics cannot give.

Sheep are generally shorn of their fleeces in the spring, but the time of shearing varies in different parts of the world. Sheep are not washed before shearing. Sometimes they are dipped into an antiseptic bath, but only when prescribed by law. Formerly, sheep were shorn by hand, but today the fleeces are usually removed in one piece by machine clippers, which shear more closely and faster than hand clippers. Wool shorn from young sheep differs in quality from that of older sheep, and wool from live sheep is different from that of dead sheep.

In the United States domestic wool reaches the mill in loosely packed bags; imported wool comes in tightly compressed bales. Each fleece contains different grades, or sorts, of wool, and the raw stock must be carefully graded and segregated according to length, diameter, and quality of fiber. Wool from different parts of the body of the lamb differs greatly. The shoulders and sides generally yield the best quality of wool, because the fibers from those parts are longer, softer, and finer.

Wool technology is that branch of animal science concerned with investigating the structure, growth characteristics, and chemical properties of wool affecting and determining commercial use. The term wool covers the fibers of sheep, angora goats, camels, alpacas, llamas, and vicunas. In this article, however, wool technology refers to sheep fiber only.

Fiber structure. Wool is epidermal in origin, a complex, organized structure growing from a follicle buried in the dermis of the skin (**Fig. 1**). Associated with the follicle are two glands, the sebaceous and the sudiferous. The sebaceous gland secretes wool grease which, when refined, is lanolin. The sudiferous gland secretes sweat, or suint. Collectively, grease and suint in the raw fleece are called yolk, a substance which is widely used in the pharmaceutical and cosmetic industries for lanolin compounds because it can be absorbed by the human skin. Both

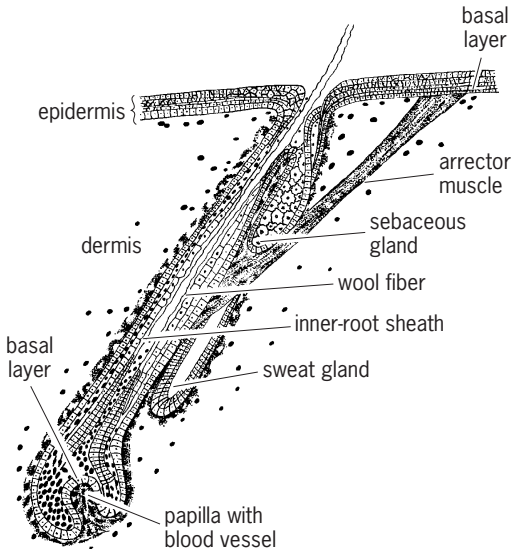

Fig. 1. Longitudinal section of a completely developed nonmedullated wool follicle, including sebaceous and sweat glands.

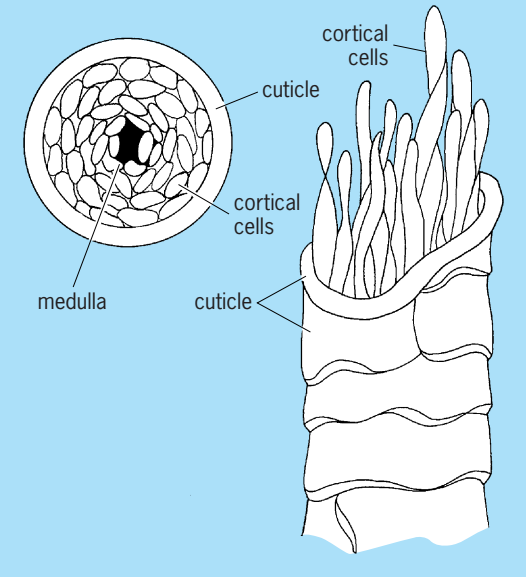

Fig. 2. Cuticle, cortex, and medulla of wool fiber.

glands open into the follicle. Their secretions function as lubricants and protectants for the wool fiber as it grows.

Growth. Blood capillaries in the papilla nourish the growing fiber, which consists of two parts, the root and the shaft. The root, or living part, is beneath the skin surface. The shaft, which protrudes from the mouth of the follicle, is dead. Physical and chemical differences between the root and shaft are listed in **Table 1**.

TABLE 1. Physical and chemical differences between root and shaft of wool fiber*

Root	Shaft
Soft and easily crushed	Tough and horny
Cells roundish	Cells elongated
Positive test for nucleic acid	Negative test for nucleic acid
Nuclei stained with hematoxylin	Nuclei unstained with hematoxylin
Cytoplasm granular in appearance	Cells distinctly fibrous
Not birefringent	Birefringent
Positive test for sulfhydryl groups	Negative test for sulfhydryl groups
No Allwoerden reaction with chlorine water	Many large Allwoerden "sacs"

*After J. M. Matthews and H. R. Mauersberger, *Textile Fibers*, 6th ed., Wiley, 1954.

A cross section of the shaft (**Fig. 2**) reveals three layers: the cuticle, the cortex, and the medulla. The last normally is absent in fine wools and infrequent in improved medium wools.

The cuticle, or epidermis, of the fiber is scalelike, and has overlapping serrated cells, the free ends of which point toward the tip of the fiber (**Fig. 3**).

The cortical layer lies beneath the cuticle. Its long spindle cells contribute tensile strength and elasticity to the wool fiber. The cortex is further divided bilaterally into a paracortex and an orthocortex. This is demonstrated by the use of a preferential dye (acid or basic), which is taken up by the orthocortex, and may be seen in a cross-sectional view.

In medium and coarse wools, the third layer, the medulla, comprises superimposed, honeycomblike cells filled with air. Medullation is a problem to the manufacturer inasmuch as fibers possessing it have lower spinning properties and are lustrous, straight, and coarse. In piece-dyed fabrics, they produce a skittery effect if they are dyed a lighter shade.

Physical properties. Wool's major physical characteristics include fiber diameter (fineness or grade), staple length, and clean wool yield. Also significant are soundness, color, luster, and content of vegetable matter. Grade refers specifically to mean fiber diameter and its variability. Fiber diameter is the most important manufacturing characteristic. Fleeces are commercially graded visually through observation and handling by persons of long experience in the industry. Degree of crimp and relative softness of the fleece are important deciding factors employed by the graders (**Fig. 4**).

Grading systems. Two systems of grading fleeces are practiced, the American and the spinning count. The seven grades of the American system are inadequate for modern manufacturing systems. The spinning-count system (**Table 2**) provides for 14 grades, the range in fiber diameter for the various spinning counts proposed by the American Society for Testing and Materials (ASTM).

Measuring devices. For meticulous laboratory research, fiber diameter is expressed in micrometers (μm) according to methods set forth by the ASTM. Devices utilized in such work include the Hardy sectioning device (**Fig. 5**) and the micronaire (**Fig. 6**).

Advantages of grading. The accurate grading of wool is important to producer and manufacturer for the following reasons: (1) it is mechanically impossible to make soft, full-handling worsted fabric from coarse wool; (2) manufacturers must use graded wools for

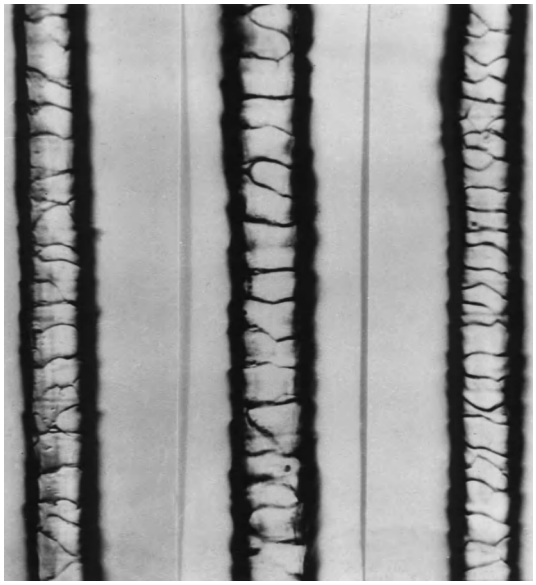

Fig. 3. Scale formation of Delaine Merino wool fibers.

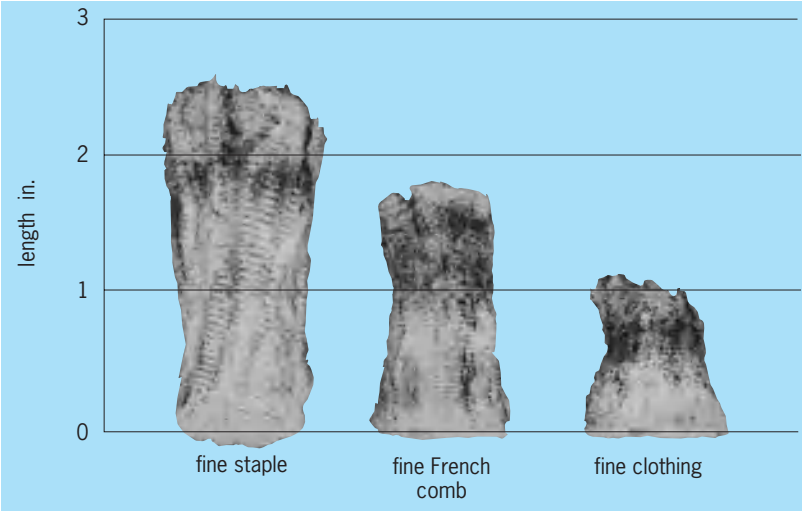

Fig. 4. Fine wool staples showing distinct crimp (wavy) pattern. The length shown for each classification is the minimum for that classification. 1 in. = 2.5 cm.

TABLE 2. Proposed specifications for grade or fineness of wool (ASTM)*

	Fineness range, μm (avg. diam.)			Fineness range, μm (avg. diam.)	
Grade	Min	Max	Grade	Min	Max
80s	17.7	19.1	54s	27.9	29.3
70s	19.2	20.5	50s	29.4	30.9
64s	20.6	22.0	48s	31.0	32.6
62s	22.1	23.4	46s	32.7	34.3
60s	23.5	24.9	44s	34.4	36.1
58s	25.0	26.4	40s	36.2	38.0
56s	26.5	27.8	36s	38.1	40.2

*Numerical terms for grade are used internationally and represent the maximum spinning capacity of wool of that fineness.

securing desired effects in finished goods; (3) textile machinery either is designed to handle certain grades exclusively, or must be adjusted for each grade; (4) graded wools gain a market advantage; (5) ranch-graded wools give the producer an index of the variability of the flock so that corrective measures can be applied in the selecting of breeding sheep; and (6) shrinkage estimates are more accurate on graded lines of wool than on ungraded lines.

Length. While wool is being graded, it is classified by length into three major categories: staple, French-combing, and clothing (**Table 3**). Usually longer wools within grade are more valuable than shorter wools because they shrink less and are less wasteful in manufacture. As fiber diameter increases, staple length increases. The length for staple classification also increases for each successively coarser grade.

Clean wool yield. Extraneous matter in unclean wool is extremely variable in quantity. Such matter consists of grease, suint, dirt, vegetable matter, and moisture. All of these except moisture are removed in the scouring (washing) process. The loss in weight may vary from 32 to 78%, and is called shrinkage; the clean scoured wool remaining is the yield.

Because wool is unmanageable after washing, the fiber is dipped in, or sprayed with, a light emulsion of olive or mineral oil to prevent it from becoming brittle and to lubricate it for the spinning operation. If the wool is to be dyed in the raw stock, it is dyed at this stage.

The percentage of extraneous matter in wool is a major factor in determining its market value. Estimates of the yield are determined either by laboratory testing of samples, or through visual appraisal by the buyer. Because of the variability in shrinkage according to length and grade, visual appraisal of clean wool yield may be inaccurate. The core test is a more reliable and consistent method.

Core test. By means of a motor-driven pressure tube, a core sample of wool is withdrawn from a bag

Fig. 5. Hardy sectioning device used in laboratory research for measuring wool fiber diameter.

Fig. 6. The micronaire is a device for measuring wool fiber diameter, which is expressed in microns (micrometers).

TABLE 3. Commercial lengths and grades of wool

American system:	Fine	1/2 Blood	3/8 Blood	1/4 Blood	Low 1/4 blood	Common	Braid
Spinning count:	80s–70s, 64s	62s–60s	58s–56	54s–50s	48s–46s	44s	40s–36s
Commercial length classes	**Staple length by grade in inches***						
Staple	2.5 and longer	3.0 and longer	3.5 and longer	4.0 and longer	4.5 and longer	5.0 and longer	5.5 and longer
Good French-combing	2.0	2.5	3.0	3.5			
Average French-combing	1.5	2.0	2.0	2.5			
Short French-combing	1.0	1.5					
Clothing and stubby	Under 1.0	Under 1.5	Under 2.0	Under 2.5	Under 4.5	Under 5.0	Under 5.5

*The length designations are based on unstretched staple length and represent a minimum length for the bulk of the staples in a sample. 1 in. = 2.5 cm.

or bale. Core-sampling patterns and the number of cores taken are prescribed according to the number of bags or bales in the lot. Each sample is then tested by standard procedures for moisture, ash, vegetable matter, and grease content, and is then adjusted to the standard condition of 14% impurities or 12.0% moisture, 0.5% ash, and 1.5% residual grease.

Chemical characteristics. Wool is primarily a protein, keratin (**Table 4**). Its molecular structure consists of long polypeptide chains in which 17 or more known amino acids are linked together by the disulfide groups of cystine. Reagents which alter the disulfide linkages change the physical characteristics of the fiber as a whole. Such reagents are oxidizers, reducers, alkalies, and light. The amphoteric nature of wool is a ready-made tool for the dyer and colorist.

Action of halogens. Treatment of the fiber with the halogens leads to absorption and chemical change. Chlorination causes wool to become yellow, harsh, and lustrous, and to lose its felting characteristics, with a corresponding rise in rate of dye absorption.

Action of heat. If briefly heated in dry air at 100–105°F (37–40°C), wool becomes harsh and loses strength and moisture. Normal moisture, softness, and strength are regained upon return to moist, cool conditions. Wool decomposes over extended periods of heat treatment.

Action of cold. At subzero temperatures, wool remains pliable and undergoes no perceptible chemical change.

Action of water and steam. Normally, wool is insoluble in water, though highly hygroscopic. If wool is boiled for 2 h, a weight loss of about 25% occurs. Conversely, the fiber diameter will swell approximately 10% with no damaging effects during brief periods of steaming (212°F or 100°C). Wool is more efficiently converted into yarn when it has been kept in a warm, moist environment; however, prolonged steaming causes loss of strength. Wool is not allowed to become absolutely dry. Usually 12–16% of the moisture

TABLE 4. Amino acid composition of wool*

Amino acid	Approximate percentage present in wool	Percentage of residue by weight	Percentage of side chain by weight
Glycine	6.5	4.49	0.09
Alanine	4.4	3.52	0.74
Serine	9.41	7.80	2.76
Proline	6.75	5.69	2.46
Valine	4.72	3.99	1.73
Threonine	6.76	5.74	2.59
Cystine	12.72†	10.83	4.89
Leucine isomers	11.3	9.75	4.92
Aspartic acid	7.27	6.28	3.22
Lysine	3.3	2.89	1.63
Glutamic acid	15.27	13.40	7.58
Methionine	0.71	0.62	0.36
Histidine	0.7	0.62	0.37
Hydroxylysine	0.21	0.19	0.11
Phenylalanine	3.75	3.34	2.07
Arginine	10.4	9.33	5.97
Tyrosine	5.8	5.23	3.43
Tryptophan	0.7	0.64	0.45
Total	110.67	94.80	45.37
Ammonia nitrogen	1.18	−0.30	−0.30
Total, corrected for ammonia nitrogen		94.50	45.07

*After J. M. Matthews.
†Based on 3.55% total sulfur and subtracting methionine sulfur.

is left in the wool to condition it for subsequent handling.

Plasticity. Under moist conditions wool becomes plastic. Its shape can be altered and its affinity for dye changed. Knowledge of this characteristic enables the manufacturer to set yarns and to produce desired color effects. Plasticity increases with rising temperatures to a point at which stretched fibers become permanently set and will not return to their normal state.

Technological advances. The inherent advantages of wool have been exploited, and its limitations as a textile fiber have been overcome by the application of technology to manufacturing processes. The use of the insecticide dieldrin as a dye renders wool mothproof for life. Permanent pleats have been imparted to garments which are shrinkproofed and can be home laundered. Each such technological advance enables wool to hold its competitive place in the field of textile manufacture and use. *See* TEXTILE.

Thomas D. Watkins

Reuse of wool fibers. There has never been a sufficient supply of new wool stocks to take care of a steadily increasing demand for wool. To meet this situation, wool fibers have had to be recovered from old clothing, rags of all kinds, and waste from wool manufacturing. This wool is variously called salvaged, reclaimed, reworked, or remanufactured, but it is best known in the textile industry as shoddy. This term is misunderstood by the average consumer, who is inclined to believe that wool fabric containing remanufactured fibers is necessarily of inferior quality.

The hardier, although less resilient, remanufactured fibers, when obtained from good original stock and combined with new wool from sheep, add durability to the soft new wool. Thus, remanufactured fibers contribute ability to withstand hard wear, although there is some sacrifice in warmth, softness of texture, and resiliency. They also make wool clothing available at lower prices.

To correct wrong impressions concerning the use of remanufactured wool, and also to protect consumers against unscrupulous practices, the United States government passed the Wool Products Labeling Act. This provides that every article of wool clothing must be labeled according to the type of wool used in its manufacture. The label must state (1) amount of wool fiber in the fabric; (2) percentage by weight of new or virgin wool fibers; (3) percentages of reprocessed or reused fibers; (4) percentage of each fiber other than wool, if such fibers constitute 5% or more of the total; (5) aggregate of other fibers; and (6) nonfibrous loading, filling, or adulterating substance.

The term wool, according to the United States government standards, must always mean new wool, not made up from any form of wool product. New wool comes directly from a fleece. It has never been previously spun, woven, felted, or worn.

Reprocessed wool. According to the government classification, reprocessed wool is that which has been reclaimed and remanufactured from unused wool materials. Such materials may be combings and scraps of wool obtained during the manufacturing processes, sample swatches, or pieces of all-wool cloth from apparel manufacturing.

Reused wool. The United States government gives the special classification of reused wool to fiber that is salvaged from all kinds of used consumers' goods.

Virgin wool. This term is now used by the textile industry to designate new wool from a lamb's fleece, but the term is too all-inclusive to serve as a criterion of quality. Although the term testifies to the fact that virgin wool does not contain remanufactured-wool fibers, it does not distinguish between the less desirable fibers of a fleece and a specially fine quality of wool. Virgin wool may also include pulled or dead wool, which may be of definitely inferior quality. One should not feel that a fabric labeled "100% new wool" is necessarily more serviceable than one containing any of the remanufactured wool fibers.

Wool of different grades may be blended or mixed together. It is not uncommon for inferior grades to be mixed with the better grades. The use of a mixture with a coarser grade of fiber is a legitimate practice if the purpose is to make a better-wearing and less expensive product, provided the label on the finished goods indicates a true description of the raw materials used. *See* ALPACA; CAMEL'S HAIR; CASHMERE; LLAMA; MOHAIR; SHEEP; VICUNA.

M. David Potter

Word processing

The use of a computer and specialized software to write, edit, format, print, and save text. In addition to these basic capabilities, the latest word processors enable users to perform a variety of advanced functions. Although the advanced features vary among the many word processing applications, most of the latest software facilitates the exchange of information between different computer applications, allows easy access to the World Wide Web for page editing and linking, and enables groups of writers to work together on a common project. The latest word processors continue to provide increased features with each new revision. *See* COMPUTER; SOFTWARE; WORLD WIDE WEB.

Writing. Writing is accomplished by using the computer's typewriterlike keyboard. Most of the characters shown on the keyboard can be typed. The characters appear on the computer screen as they are typed. A finite number of characters can be typed across the computer screen. The word processor "knows" when the user has reached this limit and automatically moves the cursor to the next line for uninterrupted typing.

The position on the computer screen where a character can be typed is marked by a blinking cursor, which is most commonly a vertical bar or underline character. The cursor can be positioned anywhere on the screen by using the mouse, or the keys marked with arrows on the keyboard. When using the arrow keys, the cursor moves in the direction

shown on the key. However, if there are no characters typed on the computer screen, the cursor will not move. *See* COMPUTER PERIPHERAL DEVICES; ELECTRONIC DISPLAY.

In addition to writing, the latest word processors provide tools to create and insert drawings anywhere in the document. Typical features allow users to draw lines, rectangles, circles, and arrowheads, and to add text. Some even enable the creation of three-dimensional effects. A variety of color options are available for lines and text, and users can also fill or shade an object, such as a circle or rectangle, with a specified color.

Editing. Editing is the process of changing an existing document. Users can, correct typographical errors, add new sentences or paragraphs, move entire blocks of text to a different location, delete portions of the document, copy text and paste it somewhere else in the document, or insert text or graphics from an entirely different document.

Most word processing programs can automatically correct many basic typographical errors, such as misspelled words, two successive capital letters in a word, and failure to capitalize the first letter of the names of days and of the first word in a sentence. Some other helpful editing tools commonly found in word processors include an automatic spelling checker, a thesaurus, and a grammar checker.

Formatting. Formatting enables users to define the appearance of the elements in a document, such as the font and type size of all headings and text, the left, right, top, and bottom margins of each page, and the space before and after sentences and paragraphs. Most word processors allow all the elements in a document to be formatted at once. This is accomplished by applying a "style." Typically, a style is selected from a library of styles provided with the word processor. There are styles for many different types of documents, including letters, resumes, and brochures. Once an overall style has been applied to a document, the style characteristics of any element within the document can be changed.

Word processors are approaching the formatting power of full-featured desktop publishing applications. The formatted page can be viewed on the computer screen exactly as it will be printed. This is referred to as "what you see is what you get" (WYSIWYG). In addition, users can add text and graphics anywhere on a page, and blocks of text and graphics can be uniquely formatted on the same page. For instance, one block of text can be in single-column format, while another block of text on the same page is in multiple-column format.

Printing. The most common printing operation is to send the document electronically from the word processor to a printer in order to produce a paper copy. The quality of the printed document is a function of the printer and, to a lesser degree, the paper, and can vary significantly. The output quality, measured in dots per inch (dpi), is directly proportional to the quality of the ink. The higher the dpi, the sharper the printed image and the more time required to print.

Other common print features include printing to a file, printing to a fax machine, and previewing pages before printing.

Printing to a file. Word processors allow a document to be printed to a file rather than to a printer. This feature is commonly used to send files directly to a print shop. Many print shops have printing equipment that can read these files directly and produce extremely high-quality printed output. *See* DATA COMMUNICATIONS.

Printing to a fax machine. Another common print feature is the ability to send an electronic document to a fax machine, which produces a much higher-quality fax than sending paper copies from one fax machine to another. *See* FACSIMILE.

Print previewing. Print previewing allows users to view the pages of a document before printing. This option displays the pages exactly as they will be printed, which allows for further editing, if required, and also saves paper.

Saving. Saving is the process of storing a copy of the electronic document as a file on a floppy or compact disk or the computer's hard disk drive. Word processors can be configured to automatically perform a save operation while working on the document, which helps prevent the loss of work.

The "Save As" feature is used to store a copy of the document under a different name or to save the file to a new location on the disk, a network node, or some other storage portable media such as a diskette or Zip disk. *See* COMPUTER STORAGE TECHNOLOGY.

Exchanging information between applications. Many business and home users purchase word processors as part of a group of applications. The other applications are typically spreadsheet, database, presentation, and personal information management applications. This package of applications is referred to as a software suite. Software suites allow for easy exchange of information. Portions of a spreadsheet, for instance, can be copied and pasted directly into a word processing document. Entire files can even be inserted. In this way, information such as text files, drawings, pictures, and spreadsheets, can be combined into a single document.

This exchange capability is made possible by a function referred to as object linking and embedding (OLE). Linking refers to the ability to access information from more than one application simultaneously; embedding refers to the ability to incorporate a copy of information from one application into another. Embedded information is referred to as an object. In order to link and embed information, both applications must support OLE.

Creating and editing Web pages. Word processing programs can be used to create and edit Web pages for the Internet. Some programs have built-in features for guiding a novice Web page designer through the entire creation process. Once created, Web pages can be edited and saved as easily as any file on a computer's hard drive. Links can be inserted into a Web page that can access other Web addresses. Also, Web pages can be viewed within the word processor, so an Internet browser is not necessary for checking

edits that have been made. In addition, existing word processing documents can easily be converted into Web pages. *See* INTERNET; WORLD WIDE WEB.

Working in groups. The latest word processors have many features for allowing groups of people to work together on the same document. For instance, multiple versions of a document can be saved to a single file for version control; access levels can be assigned so that only a select group of people can make changes to a document; edits can be marked with the date, time, and editor's name; and text colors can be assigned to differentiate editors. In addition, some word processors have editing features that include highlighting text, drawing lines through text to represent deleted text, and using red underscoring to identify changed text. Carlos Quiroga

Bibliography. A. Diller, *LaTeX Line by Line: Tips and Techniques for Document Processing*, 2d ed., John Wiley, 1999; R. Person, *Using Microsoft Word 97*, Que, 1997; R. Williams, *The PC Is Not a Typewriter: A Style Manual for Creating Professional-Level Type on Your Personal Computer*, Peachpit Press, 1992.

Work

In physics, the term work refers to the transference of energy that occurs when a force is applied to a body that is moving in such a way that the force has a component in the direction of the body's motion. Therefore work is done on a weight that is being lifted, or on a spring that is being stretched or compressed, or on a gas that is undergoing compression in a cylinder.

When the force acting on a moving body is constant in magnitude and direction, the amount of work done is defined as the product of just two factors: the component of the force in the direction of motion, and the distance moved by the point of application of the force. Thus the defining equation for work W is (1), where f and s are the magnitudes of the force

$$W = f \cos \phi \cdot s \qquad (1)$$

and displacement, respectively, and ϕ is the angle between these two vector quantities (**Fig. 1**). Because $f \cos \phi \cdot s = f \cdot s \cos \phi$, work may be defined alternatively as the product of the force and the component of the displacement in the direction of the force. In

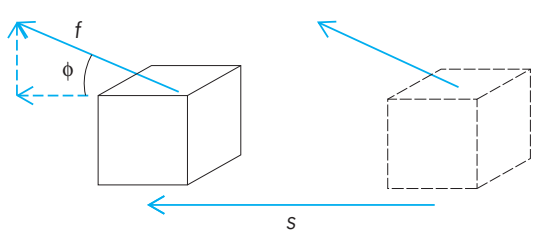

Fig. 1. Work done by constant force *f* is *fs* cos ϕ.

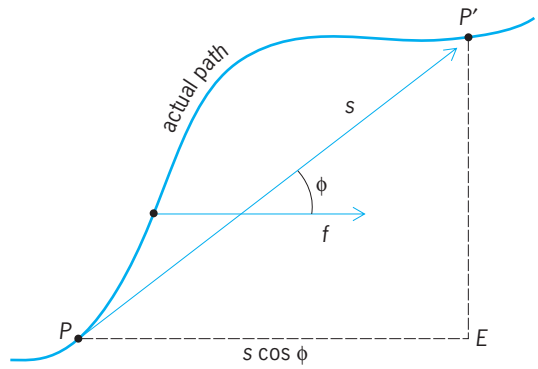

Fig. 2. Work done in traversing any path connecting points *P* and *P′* is *f* · $\overline{PE}$, assuming the force *f* to be constant in magnitude and direction.

Fig. 2 the work done by the constant force f when the application point moves along the curved path from P to P', and therefore undergoes the displacement $\overline{PP'}$, is $f \cdot \overline{PP'} \cos \phi$, or $\overline{PE}$.

Work is a scalar quantity. Consequently, to find the total work on a moving body by several different forces, the work of each may be computed separately and the ordinary algebraic sum taken.

Examples and sign conventions. Suppose that a car slowly rolls forward a distance of 10 m along a straight driveway while a man pushes on it with a constant magnitude of 200 newtons of force (200 N) and let Eq. (1) be used to compute the work W done under each of the following circumstances: (1) If the man pushes straight forward, in the direction of the car's displacement, then $\phi = 0°$, $\cos \phi = 1$, and $W = 200 \text{ N} \times 1 \times 10 \text{ m} = 2000 \text{ N} \cdot \text{m} = 2000$ joules; (2) if he pushes in a sideways direction making an angle ϕ of 60° with the displacement, then $\cos 60° = 0.50$ and $W = 1000$ joules; (3) if he pushes against the side of the car and therefore at right angles to the displacement, $\phi = 90°$, $\cos \phi = 0$, and $W = 0$; (4) if he pushes or pulls backward, in the direction opposite to the car's displacement, $\phi = 180°$, $\cos \phi = -1$, and $W = -2000$ J.

Notice that the work done is positive in sign whenever the force or any component of it is in the same direction as the displacement; one then says that work is being done *by* the agent exerting the force (in the example, the man) and *on* the moving body (the car). The work is said to be negative whenever the direction of the force or force component is opposite to that of the displacement; then work is said to be done *on* the agent (the man) and *by* the moving body (the car). From the point of view of energy, an agent doing positive work is losing energy to the body on which the work is done, and one doing negative work is gaining energy from that body.

Units of work and energy. These consist of the product of any force unit and any distance unit. Units in common use are the foot-pound, the foot-poundal, the erg, and the joule. The product of any power unit and any time unit is also a unit of work or energy. Thus the horsepower-hour (hp-h) is equivalent, in view of the definition of the

horsepower, to 550 ft-lbf/s × 3600 s, or 1,980,000 ft-lbf, or (1,980,000)(0.3048 m)(4.45 N) = 2,684,520 J. Similarly, the watt-hour is 1 J/s × 3600 s, or 3600 J; and the kilowatt-hour is 3,600,000 J.

Work of a torque. When a body which is mounted on a fixed axis is acted upon by a constant torque of magnitude τ and turns through an angle θ (radians), the work done by the torque is $\tau\theta$.

Work principle. This principle, which is a generalization from experiments on many types of machines, asserts that, during any given time, the work of the forces applied to the machine is equal to the work of the forces resisting the motion of the machine, whether these resisting forces arise from gravity, friction, molecular interactions, or inertia. When the resisting force is gravity, the work of this force is mgh, where mg is the weight of the body and h is the vertical distance through which the body's center of gravity is raised. Note that if a body is moving in a horizontal direction, h is zero and no work is done by or against the gravitational force of the Earth. If a person holds an object or carries it across level ground, she does no net work against gravity; yet she becomes fatigued because her tensed muscles continually contract and relax in minute motions, and in walking she alternately raises and lowers the object and herself.

The resisting force may be due to molecular forces, as when a coiled elastic spring is being compressed or stretched. From Hooke's law, the average resisting force in the spring is $-{}^1/_2ks$, where k is the force constant of the spring and s is the displacement of the end of the spring from its normal position; hence the work of this elastic force is $-{}^1/_2ks^2$. *See* HOOKE'S LAW.

If a machine has any part of mass m that is undergoing an acceleration of magnitude a, the resisting force $-ma$ which the part offers because of its inertia involves work that must be taken into account; the same principle applies to the resisting torque $-I\alpha$ if any rotating part of moment of inertia I undergoes an angular acceleration α.

When the resisting force arises from friction between solid surfaces, the work of the frictional force is $\mu f_n s$, where μ is the coefficient of friction for the pair of surfaces, f_n is the normal force for the pair of surfaces together, and s is the displacement of the one surface relative to the other during the time under consideration. The frictional force μf_n and the displacement s giving rise to it are always opposite in direction ($\phi = 180°$). *See* FRICTION.

The work done by any conservative force, such as a gravitational, elastic, or electrostatic force, during a displacement of a body from one point to another has the important property of being path-independent: its value depends only on the initial and final positions of the body, not upon the path traversed between these two positions. On the other hand, the work done by any nonconservative force, such as friction due to air, depends on the path followed and not alone on the initial and final positions, for the direction of such a force varies with the path,

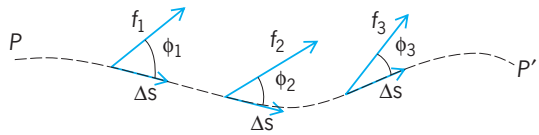

Fig. 3. Work done by a variable force.

being at every point of the path tangential to it. *See* FORCE.

Since work is a measure of energy transfer, it can be calculated from gains and losses of energy. It is useful, however, to define work in terms of forces and distances or torques and angles because these quantities are often easier to measure than energy changes, especially if energy changes are produced by nonconservative forces.

Work of a variable force. If the force varies in magnitude and direction along the path $\overline{PP'}$ of its point of application, one must first divide the whole path into parts of length Δs, each so short that the force component $f \cos \phi$ may be regarded as constant while the point of application traverses it (**Fig. 3**). Equation (1) can then be applied to each small part and the resulting increments of work added to find the total work done. Various devices are available for measuring the force component as a function of position along the path. Then a work diagram can be plotted (**Fig. 4**). The total work done between positions s_1 and s_2 is represented by the area under the resulting curve between s_1 and s_2 and can be computed by measuring this area, due allowance being made for the scale in which the diagram is drawn.

For an infinitely small displacement ds of the point of application of the force, the increment of work dW is given by Eq. (2), a differential expression that

$$dW = f \cos \phi \, ds \tag{2}$$

provides the most general definition of the concept of work. In the language of vector analysis, dW is the scalar product of the vector quantities $\mathbf{f}$ and ds; Eq. (2) then takes the form $dW = \mathbf{f} \cdot ds$. If the force is a known continuous function of the displacement, the total work done in a finite displacement from point P to point P' of the path is obtained by evaluating the

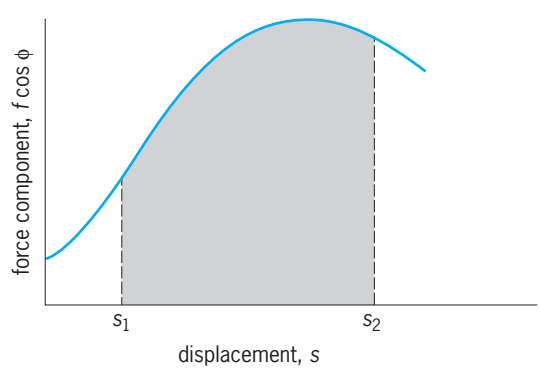

Fig. 4. Work diagram.

line integral in Eq. (3).

$$W = \int_P^{P'} f \cos\phi \, ds = \int_P^{P'} \mathbf{f} \cdot ds \qquad (3)$$

When a variable torque of magnitude τ acts on a body mounted on a fixed axis, the work done is given by

$$W = \int_{\theta_1}^{\theta_2} \tau \, d\theta$$

where $\theta_2 - \theta_1$ is the total angular displacement expressed in radians. *See* ENERGY. Leo Nedelsky

Bibliography. F. Bueche, *Principles of Physics*, 6th ed., 1994; D. Halliday and R. Resnick, *Fundamentals of Physics*, 6th ed., 2000; H. C. Ohanian, *The Principles of Physics*, 1994.

Work function (electronics)

A quantity with the dimensions of energy which determines the thermionic emission of a solid at a given temperature. The thermionic electron current density J emitted by the surface of a hot conductor at a temperature T is given by the Richardson-Dushman formula, $J = AT^2e^{-\phi/kT}$, where A is a constant, k is Boltzmann's constant ($= 1.38 \times 10^{-23}$ joule per degree Celsius) and ϕ is the work function; the last may be determined from a plot of $\log(J/T^2)$ versus $1/T$. For metals, ϕ may also be determined by measuring the photoemission as a function of the frequency of the incident electromagnetic radiation; ϕ is then equal to the minimum (threshold) frequency for which electron emission is observed times Planck's constant h ($= 6.63 \times 10^{-}$ joule second). The work function of a solid is usually expressed in electronvolts (1 eV is the energy gained by an electron as it passes through a potential difference of 1 V, and is equal to 1.60×10^{-19} J). A list of average values of work functions (in electronvolts) for metals is given in the **table**.

The work function of metals varies from one crystal plane to another and also varies slightly with temperature (approximately 10^{-4} eV/degree). For a metal, the work function has a simple interpretation. At absolute zero, the energy of the most energetic electrons in a metal is referred to as the Fermi energy; the work function of a metal is then equal to the energy required to raise an electron with the Fermi energy to the energy level corresponding to an electron at rest in vacuum. The work function of a semiconductor or an insulator has the same interpretation, but in these materials the Fermi level is in general not occupied by electrons and thus has a more abstract meaning. *See* FIELD EMISSION; PHOTOEMISSION; THERMIONIC EMISSION. Adrianus J. Dekker

Average values of work functions for metals, in electronvolts

Metal	Value	Metal	Value	Metal	Value
Al	4.20	Cs	1.93	Na	2.28
Ag	4.46	Cu	4.45	Ni	4.96
Au	4.89	Fe	4.44	Pd	4.98
Ba	2.51	K	2.22	Pt	5.36
Cd	4.10	Li	2.48	Ta	4.13
Co	4.41	Mg	3.67	W	4.54
Cr	4.60	Mo	4.24	Zn	4.29

Work function (thermodynamics)

The thermodynamic function better known as the Helmholtz energy, $A = U - TS$, where U is the internal energy, T is the thermodynamic (absolute) temperature, and S is the entropy of the system. At constant temperature, the change in work function is equal to the maximum work that can be done by a system ($\Delta A = w_{\max}$). *See* FREE ENERGY. P. W. Atkins

Work measurement

The determination of a set of parameters associated with a task.

Rationale. There are four reasons, common to most organizations whether profit seeking or not, why time, effort, and money are spent to measure the amount of time a job takes. The fifth, pay by results, is used only by a minority of organizations.

Cost accounting. If the length of time it takes to do the job is not known, the charge for the job cannot be determined. Without knowing the time/job ratio, too much or too little may be charged. As a result, the product may be overpriced and, in the long run, sales will be lost, or if too little is charged, revenue will be lost.

Evaluation of alternatives. Without knowing the time (and thus the cost) of a job, decisions regarding whether to make or buy an item, whether to mechanize or not, whether to advertise or not, could be wrong.

Acceptable day's work. Suppose a worker makes 100 widgets a day. A standard of comparison is needed to determine whether this is superior, average, or poor performance. If the standard is 75 a day, the worker is to be praised; if the standard is 200 a day, the worker is to be questioned.

Scheduling. Managers need to know the time/job ratio to make reasonable decisions for production schedules, how many people to assign to the job, how much equipment is required, and so on.

Pay by results. In a small number of organizations, pay is based on the units produced. In most organizations, pay is by the hour, week, or month and, as long as output is "reasonable," pay continues. But if pay depends upon units produced, the organization needs to know how long the task takes so a reasonable pay rate/unit can be set. *See* WAGE INCENTIVES.

Given the decision to record the amount of time per job, it is important to emphasize that first the job should be properly designed. Recording time for a poorly designed job is a mistake. The job should be improved (that is, improve productivity) before doing work measurement. *See* PRODUCTIVITY.

There are three common ways to determine time

NO.	ELEMENTS	SPEED	FEED	UPPER LINE: SUBTRACTED TIME 1	2	3	4	5	6	7	8	LOWER LINE: READING 9	10	11	12	13	14	15	MIN. TIME	AV. TIME	Std. Dev.	OCC. PER CYCLE	EFFORT RATING	NORMAL TIME
1	Get pen; TP Ready to use			3.1																3.1		.25	100	.775
2	Get letter & envelope: TP Ready to check			2.6	2.7	3.5	3.0													2.95		1	90	2.655
3	Read letter: TP Read last letter: TP Read last word *			18.2 (100)	21.5 (125)	42.1 (225)	55.0 (205)	(A)				18.2/100 = .182/word 21.5/125 = .172 42.1/225 = .187								.180/word			100	.180/word
4	Sign and aside letter and envelope: TP RL paper			3.0	2.7	3.1	3.0																	
5	Return pen: TP RL pen						3.6													3.6		.25	110	.990
	* Number of words in bracket																							
																						$\frac{4.42}{.9} = 4.91$		
																						$\frac{.180}{9} = .2$/word		

FOREIGN ELEMENTS:	TOOLS, JIGS, GAUGES, PATTERNS, ETC.:					
(A) Made 1 correction						
	OVERALL EFFORT RATING	BEGIN	END	ELAPSED	UNITS FINISHED	ACTUAL TIME PER PIECE

Elemental breakdown of a job: signing a letter. The end of the element is the termination point (TP). The times are seconds. It is assumed that four letters are considered for signature at a time, and a standard will assume no changes are needed in the letter. Elements 1, 2, 4, and 5 are constant elements. Element 3 is a variable element as the time varies with the number of words read. Allowances are assumed as 10%.

per job: stopwatch time study (sequential observations), occurrence sampling (nonsequential observations), and standard data.

Stopwatch time study. Stopwatch time study has the great advantage of flexibility—it can be used for almost any existing job. Workers have become familiar with it. It is reasonable in cost and gives reasonable accuracy. However, it does require the worker to be rated. Once the initial cost of a standard data system has been incurred, standard data may be the lowest-cost, most accurate, and most accepted technique.

There are eight steps to determine standard time:

1. *Development of a good method to be timed.*

2. *Selection of an operator to time.* In some cases, there is no choice as only one person does the job. If there is a choice, a typical or average worker should be used rather than the best. (The best is useful for productivity analysis.) An average worker is used because rating is more accurate for that worker than for an extreme worker and because other workers are more willing to accept a time standard determined on an average worker.

3. *Preparation for timing.* The job must be broken down into elements, indicating the termination points (see **illus.**). Elements permit reuses of the data, they give good internal consistency checks, they permit different ratings for different elements, and they improve methods descriptions.

4. *Selection of a timing technique.* There are four common alternatives: one watch with continuous hand movement; one watch with snapback; three watches with snapback; and an electronic watch with a hold circuit. The one watch with continuous hand movement formerly was the recommended procedure but has been made obsolete by the improved technology and lower costs of the third and fourth methods. While the first method gives accurate results, it requires excessive clerical work. The last two methods give even better accuracy and substantially reduced clerical work, at a minor incremental equipment cost for the watch. The second method has low clerical costs, but has considerable potential for inaccuracy.

5. *Determination of the number of observations.* The number can be determined by using statistical techniques from standard tables used by the organization. For example, the organization may require 200 observations for any element time less than .002 h, 175 for times from .002 to .003, and so on.

6. *Elimination of foreign elements, irregular elements, and outliers from the data.* Foreign elements are observations which are not allowed directly as part of the time standard—they may be included in

some cases indirectly as allowances. Examples of foreign elements would be speaking to a supervisor, watching someone walk by, and lighting a cigarette. An irregular element (such as oiling of the machine) occurs at infrequent intervals. The data should be used, but the number of units between elements must be determined. An outlier is an abnormally high or low value; it can be eliminated only for specific known reasons or with standard statistical tests—never merely because it is high or low.

7. *Performance rating.* Completion of the first six steps gives the recorded time. The normal time is needed—the time that a typical experienced worker would take under ideal conditions and without a break: Recorded time × rating = normal time.

8. *Allowances.* Allowances generally fall into personal, fatigue, and delay allowances. Personal allowances would be for getting a drink of water; fatigue allowances are given for physically or mentally demanding work; and delay allowances are given for uncontrollable delays. Total allowances generally range from 5 to 20% of the standard time for the job or operation. Standard time is calculated from normal time divided by (1 − allowance percent).

Occurrence sampling standards. Occurrence sampling is also called work sampling or ratio-delay sampling. If time study is a "movie," then occurrence sampling is a "series of snapshots."

The primary advantage of this approach may be that occurrence sampling standards are obtained from data gathered over a relatively long time period, so the sample is likely to be representative of the universe. That is, the time from the study is likely to be representative of the long-run performance of the worker. Another advantage is that occurrence sampling can be used when "production" is not continuous. For example, when timing phone calls from customers, calls may occur only 5 to 10 times per day and at irregular intervals. Rating generally is not used as all work is assumed to be at a 100% pace. For work done at paces different from 100%, this will cause errors in the accuracy of the standard. Thus occurrence sampling standards probably should not be used for incentive pay purposes.

To set a standard using occurrence sampling, it is necessary to do an occurrence sample plus record the units produced during the time the occurrence sample takes place. For example, consider the mechanic who does tune-ups. The individual should be observed 100 times over a 10-day period, with records of each period of idle time and of each truck or car tune-up. The output during the 10 days (say 5 trucks and 13 cars) is recorded as well as the scheduled work time (say 450 min/day).

Assume the study showed idle = 10%, truck tune-up = 36%, and car tune-up = 45%. Then working time was .9(4500) = 4050 min; a truck tune-up takes .36(4050)/5 = 292 min, and a car tune-up takes 168/.9 = 187 min.

In making the occurrence sample there are two steps: getting a sample whose size gives the desired trade-off between cost of the study and risk of an inaccurate estimate, and obtaining a sample representative of the population.

TABLE 1. Confidence levels for z levels on occurrence sampling formula*

z (number of standard deviations)	Corresponding confidence level, %
±1.0	68
±1.64	90
±1.96	95
±2.0	95.45
±3.0	99.73

*Use the table of the normal distribution (not shown here) for other values.
SOURCE: S. A. Konz, *Work Design*, published by Grid, Columbus, Ohio, 1979.

The required number of observations in the sample can be determined from the formula below,

$$A = z\sigma_p$$

where $A = sp$ = absolute accuracy desired, decimal; s = relative accuracy desired, decimal; p = mean percent occurrence, decimal; z = number of standard deviations for confidence level desired (**Table 1**); n = number of observations; and σ_p = standard deviation of a percent, $\sqrt{p(1-p)/n}$.

For example, the problem might be to determine whether or not to add an additional telephone line, as customers have been complaining the line is always busy. Management requests a study, indicating that they want a relative accuracy of ±10% and a confidence of 90%. The person doing the study must decide how many observations (samples) to take. First, from preliminary judgment, it is guessed that the lines are busy 60% of the time—that is, $p = .60$. Then $A = sp$ or .10 (.60) = .06. The number of standard deviations corresponding to a 90% confidence level is 1.64. The estimated value of p (.6) is then substituted into the formula for standard deviation of a percent. The resulting equation is $.06 = (1.64)\sqrt{(.6)(.4)/n}$. Solving for n, which is 179, and assuming 2 weeks would be needed to give a representative sample, 180 observations must be taken at the rate of 18/day for 10 days.

To make the sample representative, stratification, influence, and periodicity must be considered. Stratification means to divide the sample into strata (layers). Thus there may be 10 strata (5 days/week for 2 weeks), 20 strata (morning versus afternoon for each day of the sample), or some other division such as local calls versus long-distance. Influence means that it is not desirable that the behavior of the individual changes because of being observed. For this example, observations should be taken in such a way that neither customers nor employee change their use of the telephone because they are being sampled. Periodicity refers to existing patterns of behavior coincident with specific times of the day. The sample must not take too many or too few observations at these special times.

Assume that from the sample of 180 the lines were busy 90 times, so $p = 90/180 = .5$. Therefore,

$A = 1.64\sqrt{(.5)(.5)/180}$ so $A = 6.1\%$. Thus the conclusion is that the phones are busy 50% (±6.1%) of the time with a confidence of 90%—if the situation does not change. The stratification information indicates that the percent busy was 40% on Monday and Wednesday, 50% on Tuesday, and 60% on Thursday and Friday.

To set a time standard for phone calls, all that is needed is to record the number of calls during the time of the study. It is already known that the line was busy 50% of the time of 480 min/day × 10 days × .50 = 2400 min. If there were 512 calls during the 2 weeks, then time/call is 2400/512 = 4.7 min/call.

From this information, several actions are possible, such as installing another telephone line, estimating the cost of dealing with customers over the telephone, and estimating how much extra time is needed for the telephone on Thursday and Friday (and how much is available on other days). However, a new phone line will change the situation, so the times from the sample may not be representative of the situation after the change.

Standard data standards. Reuse of previous times (standard data) is an alternative to measuring new times for an operation. Lower cost, consistency, and ahead-of-production are the three advantages. Looking up the time to walk 160 ft (50 m) from a table of standard times rather than doing a special study to measure the time to walk 160 ft (50 m) saves the time of the person setting the standard. However, this operating cost must be balanced by the capital cost of setting up the table of standard times for walking. Thus the capital cost must be justified by many uses—that is, standard data are economically justifiable only for standard repetitive elements. A second advantage is consistency. In setting the time to walk 160 ft (50 m), a number of studies can be used, so any rating or measurement errors would tend to average out. In addition, since a table or formula is used, every analyst always gets the same answer; rating and judgment are minimized. A third advantage is that the time can be estimated prior to production. Timing requires an experienced operator, a work station with tools, and product. But times often are wanted ahead of production for determining which alternative work method to use and such.

There are three levels of detail: micro, elemental, and macro (**Table 2**). Microlevel systems have times of the smallest component ranging from about .01 to 1 s. Components usually come from a predetermined time system such as methods-time-measurement (MTM) or Work-Factor. Elemental level systems have the time of the smallest component, ranging from about 1 to 1000 s. Components come from time study or microlevel combinations. Macrolevel systems have times ranging upward from about 1000 s. Components come from elemental-level combinations, from time studies, and from occurrence sampling.

For example, assume a standard time is needed for signing a business letter (see illus.). It may be decided to break the task into five elements: (1) get a pen, (2) get a letter and envelope, (3) read the letter, (4) sign and set aside the letter and envelope, and (5) return pen. For simplicity it may be assumed that all letters are examined in batches of four, and no corrections need to be made. Element 3, read the letter, is a variable element in that it varies with the letter length. The remaining elements are constant. Element 3 time might be 5. 05N, where N = the number of words in the letter. If the total time for elements 1, 2, 4, and 5 was 150 time-measurement units (TMU), then total time for signing a letter is 150 + 5.05N. Thus a 100-word letter takes 150 + 505 = 655 TMU, and a 200-word letter takes 150 + 1010 = 1160 TMU. Allowances should be added to both times, but rating is not necessary because it is built into the MTM time. *See* METHODS ENGINEERING; PERFORMANCE RATING.

Stephan A. Konz

TABLE 2. Three levels of detail for standard time systems

Microsystem (typical component time range from .01 to 1 s; MTM nomenclature)

Element	*Code*	*Time*
Reach	R10C	12.9 TMU*
Grasp	G4B	9.1
Move	M10B	12.2
Position	P1SE	5.6
Release	RL1	2.0

Elemental system (typical component time range from 1 to 1000 s)

Element	*Time*
Get equipment	1.5 min
Polish shoes	3.5
Put equipment away	2.0

Macrosystem (typical component times vary upward from 1000 s)

Element	*Time*
Load truck	2.5 h
Drive truck 200 km	4.0
Unload truck	3.4

*27.8 TMU = 1 s; 1 s = .036 TMU.
SOURCE: S. A. Konz, *Work Design*, published by Grid, Columbus, Ohio, 1979.

Bibliography. S. A. Konz, *Work Design: Industrial Ergonomics*, 5th ed., 1999.

Work standardization

The establishment of uniformity of technical procedures, administrative procedures, working conditions, tools, equipment, workplace arrangements, operation and motion sequences, materials, quality requirements, and similar factors which affect the performance of work. It involves the concepts of design standardization applied to the performance of jobs or operations in industry or business. The task of work standardization thus follows the design process and precedes the establishment of systems of work measurement. The latter are designed to enforce work standards. *See* DESIGN STANDARDS; WORK MEASUREMENT.

Work standardization is a part of methods engineering and usually precedes the setting of time

standards. There would be little point in establishing time standards until the method by which the work is to be done has been stan- dardized.

The objectives of work standardizations are lower costs, greater productivity, improved quality of workmanship, greater safety, and quicker and better development of skills among workers. Work standardization also tests the ingenuity of industrial engineers and of production or operations managers at all levels. Sometimes, ideas leading to work standardization are derived from employee suggestions. More often, however, work standards become part of an adversarial process in which they, in effect, establish work rules and are thus involved in collective bargaining and other aspects of labor relations, just as is work measurement in general.

Work standardization often leads to simplification in that it presents opportunities for eliminating and consolidating unit operations of various kinds. Searching for a common approach also requires a careful definition of what needs to be done, which is also often a source of improvements.

Work standardization often resolves larger operations into smaller elements, and the eventual standard is a combination of such smaller ones. This procedure must not be used without caution, however. The issue of independence of elements is always present; the whole may or may not equal the sum of the parts. Subsequent work measurement, especially when done by methods of statistical analysis that focus on finding outliers, that is, truly abnormal results, must then be used to revise or reinterpret work standards.

One of the best known of the more formal techniques of work standardization is group technology. This is the careful description of a heterogeneous lot of machine or other piece parts with a view to discovering as many common features in materials and dimensions as can be identified. It is then possible to start a rather large lot of a basic part through the production process, doing the common operations on all of them. Any changes or additional operations required to produce the final different parts can then be made at a later stage. The economy is realized in being able to do the identical jobs at one time.

It may prove economical to do some of the common operations on all the parts even when they are not really required, rather than set the jobs up separately. Parts like shafts, cover plates, tooling components, or mounts for electronic circuitry are among the many applications of group technology. To implement this, a special form of "cellular" layout is required which parallels the branching characteristics of the production system. Opportunities also exist for this approach in the chemical processing industries in that basic mixes can be produced, with later additives or other unit operations providing the required product differentiation.

There has been considerable progress in computerized systems to facilitate group technology. The techniques are closely linked to computer-aided design and to formalized codes that permit the detailed description of many operations. This is a necessary prerequisite to later specification of the identities. In several of these systems, the total information is finally recorded on punched cards or other memory media, with some including a microfilm of the drawing of the part required.

Simpler and less formal forms of work standardization can occur just about everywhere, in both the private and public sector. Whenever a good idea is being put forth dealing with the more efficient performance of a detailed operation, it is worthwhile to study its possible application to a wider area. Conversely, if standard practice now exists in many parts of the business, it may be advantageous to apply it to activities hitherto left unstandardized.

However, the economics of work standardization must be carefully balanced against added costs such as new equipment or tooling, carrying greater inventories, and extra handling. The establishment of larger production batches does not necessarily produce economy of scale. Some forms of group technology, for instance, lose some of their advantages when a numerically controlled machine can quickly make changes in the parts it processes as it goes along.

There is also a possibility that work standardization may lead to repetitious and monotonous operations where previously there was a degree of flexibility and variety permitted; in such a case, work standardization becomes undesirable for a variety of operational and social reasons. Great caution must also be taken to see that administrative costs do not come to exceed any operational savings realized. *See* METHODS ENGINEERING; PERFORMANCE RATING.

John E. Ullman

World Wide Web

A part of the Internet that contains linked text, image, sound, and video documents. Before the World Wide Web (WWW), information retrieval on the Internet was text-based and required that users know basic UNIX commands. The World Wide Web has gained popularity largely because of its ease of use (point-and-click graphical interface) and multimedia capabilities, as well as its convenient access to other types of Internet services (such as e-mail, Telnet, and Usenet). *See* INTERNET.

Improvements in networking technology, the falling cost of computer hardware and networking equipment, and increased bandwidth have helped the Web to contain richer content. The Web is the fastest medium for transferring information and has universal reach (crossing geographical and time boundaries). It is also easy to access information from millions of Web sites using search engines (systems that collect and index Web pages, and store searchable lists of these pages). The Web's unified networking protocols make its use seamless, transparent, and portable. As the Web has evolved, it has incorporated complementary new technologies for developing online commerce, as well as video on demand and mobile Web (Web-enabled phone), to name a few.

Development. In 1989, Tim Berners-Lee and colleagues at the European Organization for Nuclear Research (CERN) created the Hypertext Transfer Protocol (HTTP), a communication protocol for transmitting linked documents between computers. HTTP is the basis for the World Wide Web and follows the TCP/IP for the client-server model of computing. In 1992, Berners-Lee developed a software program (client), called a Web browser, which could retrieve and display text documents under the protocol. The following year, Marc Andreessen at the National Center for Supercomputing Applications created a graphical Web browser called Mosaic. Mosaic allowed users to "browse" the Web using the familiar point-and-click functionality of the Macintosh and Windows operating systems. The first commercial Web browser was released in April 1994. *See* CLIENT-SERVER SYSTEM; DATA COMMUNICATIONS.

Linking documents. A hypertext, in contrast to a conventional text, does not have to be read sequentially. A hypertext contains linked documents, and the reader chooses a path or order to follow. A link or hyperlink is a highlighted (underlined, colored, and so forth) word, phrase, or image which, when chosen, sends the user to another document (text, image, audio, video, or animation). Hyperlinks are similar to references in a scientific paper or cross references in a dictionary. In electronic documents, these cross references can be followed by a mouse click, and the documents can be anywhere in the world. A document can contain various links to other documents on different computers anywhere on the Internet, permitting easy Web browsing (**Fig. 1**).

Individual documents are called Web pages, and a collection of related documents is called a Web site. All Web documents are assigned a unique Internet address called a Uniform Resource Locator (URL) by which they can be accessed by all Web browsers. A URL (such as http://www.hq.nasa.gov/office/procurement/index.html) identifies the communication protocol used by the site (http), its location [domain name or server (www.hq.nasa.gov)], the path to the server (office/procurement), and the type of document (html).

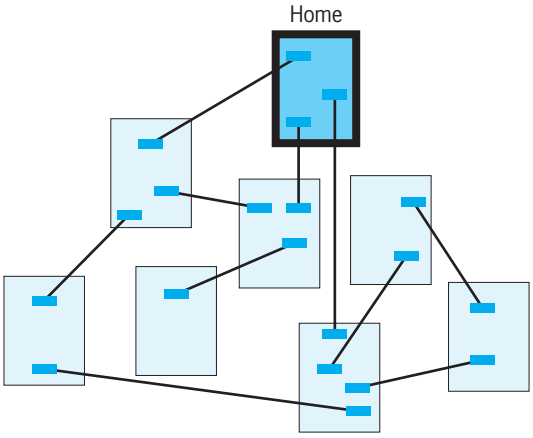

Fig. 1. In the World Wide Web, browsing pathways extend to electronic documents containing hyperlinks.

HTML. The language used to create and link documents is called Hypertext Markup Language (HTML). Markup is the process of adding information to a document that is not part of the content but identifies the structure or elements. Markup languages are not new. HTML is based on the Standard Generalized Markup Language (SGML), which became an ISO standard in 1986 and has been used in the publishing business and for technical documentation. Whereas SGML focuses on encoding the logical structure and content of a document, rather than its display formatting or even the medium in which the document will be displayed, HTML specifically encodes the format of the hypertext document for publication on the World Wide Web.

Dynamic Web pages. Though the initial format for creating a Web site was pure HTML, new and extended HTML has the ability to include programming language scripts such as common gateway interface (CGI), active server page (ASP), and Java server page (JSP), which can be used to create dynamic and interactive Web pages as opposed to just static HTML text. Dynamic Web pages allow users to create forms for transactions and data collection; perform searches on a database or on a particular Web site; create counters and track the domain names of visitors; customize Web pages to meet individual user preferences; create Web pages on the fly; and create interactive Web sites.

Extensible Markup Language (XML). XML, developed by the World Wide Web Consortium, is another derivative of SGML and is rapidly becoming the standard information protocol for all commercial software such as office tools, messaging, and distributed databases. XML is a flexible way to create common information formats and share both the format and the data on the World Wide Web, intranets, and other Web-based services. For example, automobile manufacturers might agree on a standard or common way to describe the information about a car (engine capacity, horsepower, tire specifications, mileage, transmission details, and so forth) and then describe the car information format with XML. Such a standard way of describing data would enable business users to access the manufacturers' Web pages, gather data, and then use the data for their own purpose. Any individual or group of individuals or companies that want to share information in a consistent standard way can use XML. *See* DATABASE MANAGEMENT SYSTEM.

XML is a text-generation language compared to HTML. HTML could capture only the format and the content of the text, whereas XML can capture the data related to the text. XML, like HTML, contains markup symbols (called tags) to describe the content of a Web page. For example, the XML tag <EngineCap> indicates that the data to follow is the Engine Capacity. This means that an XML file can be processed purely as data by a program, it can be stored with similar data on another computer or, just as an HTML file, it can be displayed. This is of immense use to businesses as they can now use XML for sharing data using a Web interface.

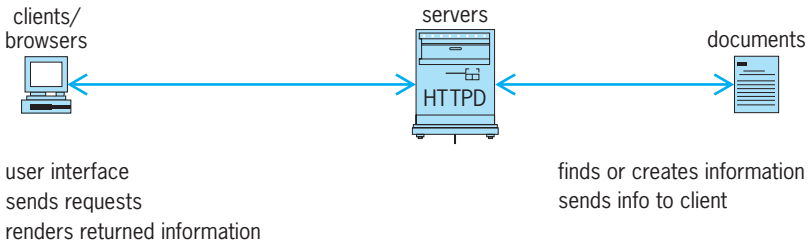

Fig. 2. Simplified view of how the Web works.

World Wide Web Consortium (W3C). The W3C is an international consortium (companies, research foundations, scientists, technologists, and universities) run by the Institut National de Recherche en Informatique et en Automatique, the Massachusetts Institute of Technology, and Keio University. The organization's purpose is to develop open (freely distributed) standards so that the Web can evolve in a single direction instead of among competing standards. The W3C is the chief standards body for HTTP and HTML. It develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential as a forum for information, commerce, and communication.

Web architecture. In the Internet client-server system, the client (computer or program) requests documents, and the server (computer or software package) sends the requested information to the client. When a user clicks on a hyperlink, the browser (client) orders a document from the server, which delivers the requested document (or an error message if it is unavailable). The browser then presents the document to the user (**Fig. 2**). Browsers exist for most popular operating systems (such as Apple Macintosh, Microsoft Windows, and Unix). *See* OPERATING SYSTEM.

The computer or program that processes document requests is called a Web server. There is another computer, called an application server (also known as an app server), that executes commands requested by the Web server to fetch data from databases. In addition, the Web server has a database (called the DB server), which is a repository of data and content. Each page of a Web site is usually stored on a server as a separate file.

Anyone can create a Web server and, for the most part, anyone can read what is online. The reason it all works is that everyone is using the same standards. However, this increases the threat of misuse of the Web by hackers or through virus attacks on Web sites. Organizations all over the world are sensitive to such harmful possibilities. Thus security has become one of the most important issues.

Security. A firewall is a set of hardware and software programs that protects the resources of a private network from users in other networks. The firewall has the ability to restrict access to the Web server by identifying the request source's Internet address. Thus it screens unauthorized access and prevents hacking, that is, inhibits outsiders from accessing private data resources. Cryptography is beginning to play a major role in Web security. Public Key Infrastructure (PKI), 128-bit encryption, and digital certificates have been developed to protect online resources. *See* COMPUTER SECURITY.

Alladi Venkatesh

Bibliography. M. Abrams, *World Wide Web: Beyond the Basics*, Prentice Hall, Englewood Cliffs, NJ, 1998; J. Nielson, *Designing Web Usability*, New Riders, Indianapolis, 2000; C. Rolland (ed.), *Information Systems in the WWW Environment*, Chapman and Hall, London, 1998; K. Sutherland, *Understanding the Internet: A Clear Guide to Internet Technologies*, Butterworth-Heinemann, Oxford, 2000.

Wrought iron

As defined by the American Society for Testing and Materials, "a ferrous material, aggregated from a solidifying mass of pasty particles of highly refined metallic iron, with which, without subsequent fusion, is incorporated a minutely and uniformly distributed quantity of slag."

This slag is a ferrous silicate resulting from the oxidizing reactions of refining, and it varies in amount from 1 to 3% in various types of final product. It is in purely mechanical association with the iron base metal, as contrasted with the alloying relationships of the metalloids present in steel.

History. Wrought iron had a dominant position in meeting the need for a forgeable, nonbrittle, ferrous metal, from before the time of historical record until the advent of the age of steel.

The early primitive methods were finally succeeded by the puddling furnace of H. Cort in 1784. Such wrought iron played a major role in the industrial revolution until the steelmaking inventions of H. Bessemer in 1856, and the open-hearth regenerative furnace of W. Siemens several years later.

Metallurgy. Whether the product is wrought iron or wrought steel, similar chemical reactions are involved in refining a molten charge of pig iron to a composition approaching pure iron. The associated metalloids are oxidized—carbon to form carbon monoxide; and silicon, manganese, and phosphorus with appropriate additions to form a liquid insoluble slag.

The vital demarcation between the puddling and steelmaking furnace was the limited temperature of the former. The puddling furnace was below the 2700°F (1482°C) needed to keep the refined iron molten. The freezing point of iron rises as refining proceeds. Practically all the 3–4% carbon in the original charge is removed, and the base metal progressively solidifies. Manual manipulation accompanied by internal reactions results in a spongy mass of virtually pure iron, impregnated with iron silicate slag. Subsequent squeezing and rolling operations eject most of the still-fluid slag; but the small percentage finally retained is elongated into the distributed threadlike structure characteristic of quality wrought iron.

In steelmaking operations, on the contrary, the high temperature maintains fluidity of both metal and

slag, with liquation and separation of the latter. The slag-free metal is poured into molds, ready for the subsequent rolling operations.

A distinguishing characteristic of wrought iron is a fragmented or irregular fracture, as contrasted with a fibrous or crystalline type in steel. Metallographic analysis shows that this results from the fiberlike slag inclusions that are finely and uniformly intermingled with the iron base metal.

The development of steelmaking processes prompted prediction of the doom of wrought iron. Puddling required hard manual labor, with low output both in unit mass and total tonnage, and accompanying high cost. However, although wrought iron once held competitive merit in certain uses, notably where corrosion- and shock-resistance were important, very little wrought iron is produced today.

Production. Efforts to improve production by replacing hand puddling with mechanical puddling were unsuccessful because of the resulting nonuniformity in different batches due to lack of control. However, the Aston process, which approached the problem on the basis of metallurgical principles, has been used successfully to make wrought iron synthetically. The process is conducted in three main steps.

1. Pig iron is refined by standard steelmaking operations (the Bessemer converter is used).

2. An iron silicate slag of desired composition is melted in a rotary furnace (iron ore and sand are the basic ingredients).

3. The metal is poured in a steady stream into a vessel containing molten slag at a temperature below the freezing point of the iron, causing a sponge ball of wrought iron to solidify.

During instantaneous and progressive solidification of the metal, its dissolved gases are liberated with sufficient force to disintegrate the metal. The large volume of liquid slag is decanted and reused, and the sponge ball with its entrained slag is compacted into a large ingot while at a welding heat.

The result is wrought iron of controlled quality. Unit masses and all mechanical equipment closely parallel those of steelmaking. Diversification of product is attained, which was not feasible with the limitations of the puddling process. *See* CAST IRON; IRON ALLOYS; IRON METALLURGY; PYROMETALLURGY; STEEL; STEEL MANUFACTURE.

James Aston; John F. Wallace

Wulfenite

A mineral consisting of lead molybdate, $PbMoO_4$. Wulfenite occurs commonly in yellow, orange, red, and grayish-white crystals. They may be tetragonal,

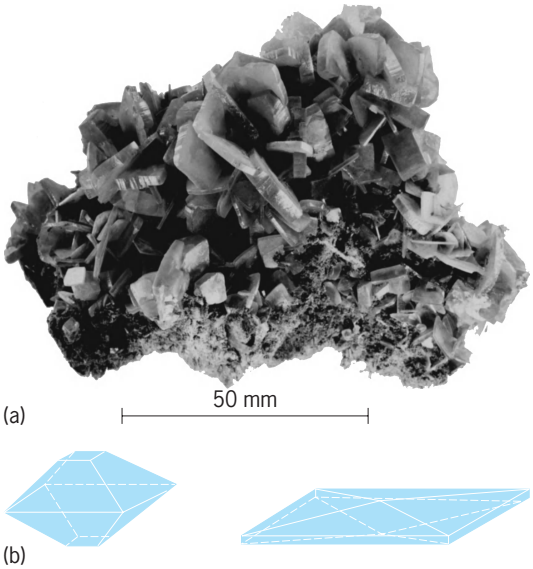

Wulfenite. (*a*) Thin tabular crystals on limonite, Organ Mountains, New Mexico (*American Museum of Natural History specimen*). (*b*) Crystal habits (*after C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*).

tabular (see **illus.**), or pyramidal, with a luster from adamantine to resinous. Wulfenite may also be massive or granular. Its fracture is uneven. Its hardness is 2.7–3 and its specific gravity 6.5–7. Its streak is white. It is easily fusible and is decomposed by hydrochloric or nitric acid with the separation of molybdic oxide.

Wulfenite occurs as a secondary mineral in the oxidized zone (of veins) of lead deposits associated with pyromorphite, cerussite, vanadinite, and other oxide zone minerals such as goethite and calcite.

Wulfenite is found in numerous localities in the western and southwestern United States. Brilliant orange tabular wulfenite crystals up to 2 in. (5 cm) in size have been found from the Red Cloud and Hamburg mines in Yuma County, Arizona. *See* MOLYBDENUM.

Edward C. T. Chao

Wurtzilite

A black, infusible carbonaceous substance occurring in Uinta County, Utah. It is insoluble in carbon disulfide, has a density of about 1.05, and consists of 79–80% carbon, 10.5–12.5% hydrogen, 4–6% sulfur, 1.8–2.2% nitrogen, and some oxygen.

Wurtzilite is derived from shale beds deposited near the close of Eocene (Green River) time. The material was introduced into the calcareous shale beds as a fluid after which it polymerized to form nodules or veins. *See* ASPHALT AND ASPHALTITE; IMPSONITE.

Irving A. Breger

X

X-ray astronomy

The study of x-ray emission from extrasolar sources, including virtually all types of astronomical objects from stars to galaxies and quasars. This space science requires balloons, rockets, or space satellites to carry experiments above most (or all) of the Earth's atmosphere, which would otherwise absorb the radiation. The x-ray region of the electromagnetic spectrum extends from wavelengths of about 10 picometers to a few tens of nanometers, with shorter wavelengths corresponding to higher-energy photons (1 nm corresponds to about 1000 eV). X-ray astronomy is traditionally divided into broad bands—soft and hard—depending on the energy of the radiation being studied. Observations in the soft band (below about 10 keV) must be carried out above the atmosphere, while hard x-ray observations can be made at high altitudes achievable by balloons. There is a limit to the observable spectrum at low energies due to absorption by neutral interstellar hydrogen gas. This low-energy cutoff depends on viewing direction, but can be below 250 eV. In the directions toward the galactic center, the cutoff can be above 2000 eV. At high energies (short wavelengths), space remains transparent through the entire x-ray spectrum. *See* ELECTROMAGNETIC RADIATION; ROCKET ASTRONOMY; X-RAYS.

X-ray observatories. In 1962, instruments aboard a rocket flight detected the first nonsolar x-ray source, later identified with a star in the constellation Scorpius (Sco X-1). The first exploratory phase in x-ray astronomy was carried out with sounding rockets and balloon flights during the late 1960s. During the early 1970s *SAS-A* (or *Uhuru*) and other satellites conducted the first full-sky surveys for x-ray emissions.

Einstein Observatory. In 1978 a major advance came to the field with the launch of the *Einstein Observatory (HEAO 2)*. This satellite introduced the use of focusing high-resolution optics to x-ray astronomy. The ensuing increase in sensitivity and the new ability to obtain images resulted in a qualitative change in the scope of the field. The images obtained with the *Einstein Observatory* during its $2^1/_2$ year operational lifetime had a major impact on many areas of study. *See* X-RAY TELESCOPE.

ROSAT. The *Roentgen Satellite (ROSAT)*, which was launched in 1990 and operated for almost 9 years, surveyed the entire sky with higher sensitivity than had been previously achieved; one result was an all-sky catalog of sources with nearly 125,000 objects, a more than twenty-fold increase in the number of known x-ray sources. *ROSAT* was also used to observe specific sources or regions of the sky for searches of extremely faint objects. The observatory used two imaging detectors. One was a high-resolution device similar to the *Einstein Observatory* High Resolution Imager (HRI), which could be moved into the focus of the x-ray telescope to obtain pictures with the best sharpness. A second detector, the Position Sensitive Proportional Counter (PSPC), provided a larger-sized image with less resolution but more information on the energy content of the x-rays from a source.

ASCA. Japan's *ASCA* (*Advanced Satellite for Cosmology and Astrophysics*, formerly *ASTRO-D*) spacecraft was launched into orbit in 1993 and operated for $7^1/_2$ years. It used lightweight modest-resolution x-ray optics that functioned at higher energies than the *Einstein* or *ROSAT* telescopes, and had advanced-technology charge-coupled-device (CCD) detectors that provided both good image quality and improved spectral resolution over previous imaging detectors. These detectors provided detailed information on the energy content of many types of x-ray sources, allowing the physical conditions at these objects to be measured. *See* CHARGE-COUPLED DEVICES.

RXTE. In late 1995 the National Aeronautics and Space Administration (NASA) launched the *Rossi X-Ray Timing Explorer (RXTE)*. This satellite measures the time variations in strong x-ray sources in order to understand the fundamental sources of

energy for these objects. Sources can be observed up to fairly high energies, which is important in resolving the different physical mechanisms responsible for the observed emission.

BeppoSAX. The Italian-Dutch *BeppoSAX* satellite, launched in 1996, can observe x-ray sources at energies ranging from 0.1 to 300 keV (although imaging is restricted to energies below 10 keV), with good energy resolution. Previously, it required the combined observations of several spacecraft to record a spectrum over this very broad energy range. In addition, two wide-field cameras can monitor the long-term variability of x-ray sources and detect transient x-ray phenomena.

The satellite is equipped with a gamma-ray burst monitor, and can quickly and accurately locate the positions of the x-ray afterglows of gamma-ray bursts that are within the field of view of the wide-field cameras. This capability makes it possible to spot the optical afterglows of some of the bursts as well, and thereby locate them even more precisely. In at least one case, observation of the optical spectrum has allowed a direct determination of the distance to the burst (about 3×10^9 light-years), strengthening the case that the bursts are generally at cosmological distances. *See* GAMMA-RAY ASTRONOMY.

Chandra. NASA's *Chandra X-ray Observatory* (*CXO*), launched in July 1999, is equipped with a 10-meter-focal-length (33-foot) telescope consisting of four pairs of mirrors arranged in concentric rings. It can focus x-rays more than 10 times as sharply as any previous x-ray telescope, with an angular resolution better than one half arcsecond, comparable to that of the best ground-based optical telescopes. The combination of high resolution with a large collecting area, sensitivity to x-rays from 0.1 to 10.0 keV, and long observation times allows *Chandra* to observe extremely faint sources.

At the telescope focus, images with the best sharpness are obtained with the High Resolution Camera (HRC), which, like similar devices on *Einstein* and *ROSAT*, is based on microchannel plates. Higher energy resolution is obtained with a second detector, the Advanced CCD Imaging Spectrometer (ACIS), with 10 charge-coupled devices similar to those used on ASCA. Transmission gratings can be swung into the path of the x-rays to achieve even better energy resolution, surpassing previous instruments. The gratings consist of meshes of fine gold wires whose spacing is precisely measured, and they deflect (diffract) the x-rays though an angle that depends on the x-ray energy. High-resolution x-ray spectroscopy enables the temperature, ionization, and chemical composition of x-ray sources to be explored. *See* DIFFRACTION GRATING.

Chandra follows an elongated orbit that takes it a third of the way to the Moon every 64 hours. This enables it to spend most of its time above the radiation belts surrounding the Earth (which can give false readings and even damage instruments) and to carry out uninterrupted long-duration observations.

XMM-Newton. In December 1999, the European Space Agency launched the *X-ray Multimirror Mission*, renamed *XMM-Newton*. The spacecraft, nearly 11 m (36 ft) long, has three x-ray telescopes, each of which has 58 concentric mirrors nested just a few millimeters apart, giving them each an effective area for collecting radiation even larger than that of *Chandra*. A charge-coupled device camera, the European Photon Imaging Camera, is located at the prime focus of each of the telescopes. For high-resolution spectroscopy, two of the telescopes are each equipped with a Reflection Grating Spectrometer. The reflection grating is a mirror with narrowly spaced grooves that reflects about half of the x-rays from the telescope to a secondary focus (with its own charge-coupled device camera), but at an angle that depends on the x-ray energy. The observatory is particularly sensitive to harder x-rays. *XMM-Newton* also has an Optical Monitor, a 30-cm (12-in.) optical-ultraviolet telescope that is aligned with the spacecraft's main telescopes and is equipped with a powerful image intensifier and a charge-coupled device camera. It allows x-ray sources to be observed simultaneously at visible and ultraviolet wavelengths. *XMM-Newton* follows a highly elliptical orbit, similar to that of *Chandra*.

Types of sources. X-ray astronomy is traditionally divided into specialties depending upon the origin of sources—galactic or extragalactic—and the energy range of observations—soft (below 10 keV) and hard (above 10 keV). Most observations have been in the soft band. Sky surveys, particularly the *ROSAT* all-sky survey, have located tens of thousands of sources at a sensitivity of about 1/100,000 of the strength of the brightest source, Sco X-1. This is comparable to a survey of the sky in visible light which extends from the brightest stars to about 12th-magnitude stars. There are over 1,000,000 stars in this range of brightness compared with the 125,000 x-ray sources detected. Some of these sources are concentrated along the galactic equator. Such a concentration corresponds to objects in the Milky Way Galaxy, particularly in the disk, which contains most of the galactic stars and the spiral arms. Other x-ray sources are uniformly spread over the sky, associated mainly with extragalactic objects, such as individual galaxies, clusters of galaxies, and quasars. *See* MILKY WAY GALAXY.

Galactic sources. Galactic x-ray sources have been identified with different types of unusual objects. With the advent of the *Einstein Observatory* and the great increase in sensitivity obtained by using focusing x-ray optics, it became possible to detect x-ray emission from stars similar to the Sun.

Normal stars. X-ray emission from the hot corona of the Sun was detected as early as the 1950s by using sounding rockets. The level of x-ray emission from the Sun is about a million times less than the emission in visible light, with occasional increases associated with solar flares. Low levels of x-ray emission have been detected from all types of normal stars. This surprising result has revived interest in models of stellar structure and evolution, which must now be reevaluated. In these sources the x-ray emission is most likely coming from the corona of the stars, with energy being transported from the stellar interior, through

the surface, and into the extended atmosphere. In all cases, the x-ray emission is a small fraction of the total luminosity of the star and is mainly interesting with regard to the problem of how energy is transported from the stellar interior to the surface. *See* SOLAR CORONA; STAR; SUN.

Supernova remnants. The Crab Nebula, which was the first nonsolar x-ray source to be identified with a specific celestial object, is a supernova remnant left over from the explosive death of a star. This particular supernova remnant is about 900 years old and contains a rapidly rotating neutron star at its center as well as a nebula consisting of hot gas and energetic particles. The neutron star is the remaining core of the star which exploded, and its rotation causes energetic electrons trapped in the magnetic fields of the nebula to radiate pulses of electromagnetic energy in radio waves, visible light, and even x-rays. About 10% of the x-ray emission from the Crab is pulsed radiation from the neutron star; the rest is extended emission associated with the nebula. *See* CRAB NEBULA.

Many other galactic supernova remnants have been detected as x-ray sources, and there have also been detections of supernova remnants in the nearest neighboring galaxies, especially in the Large Magellanic Cloud. The youngest known supernova remnant in the Milky Way Galaxy, Cas-A in the constellation Cassiopeia, is about 350 years old. The expanding shell of the explosion can easily be seen in x-ray images (see **illus.**). Analysis of the spectrum of x-ray emission shows that heavy elements, formed by nucleosynthesis within the precursor star, have been blown off during the explosion and are mixing into the interstellar medium. Observations by the *Chandra X-ray Observatory* revealed for the first time a point source near the center of Cas A which is believed to be the neutron star produced in the supernova event. Chandra observations show that the concentrations of heavy elements are arranged in a manner that echoes the ordering of the elements in the shell-burning precursor star. *See* NEUTRON STAR; NUCLEOSYNTHESIS; PULSAR; SUPERNOVA.

X-ray stars. While normal stars radiate some energy in the x-ray band, most of their luminosity is output in the visible region of the spectrum. X-ray surveys have discovered a class of objects where the majority of energy is radiated in the x-ray portion of the electromagnetic spectrum. These are known as x-ray stars. The majority of galactic x-ray stars are located in the plane of the Milky Way Galaxy, mostly within the spiral arms. This location coincides with the distribution of bright young stars in the Galaxy, and it is very likely that this class of stars goes through a stage of copious x-ray emission as the stars evolve.

Some x-ray stars have been located in the spherical halo of the Galaxy, some far from the plane, and in globular clusters. Both regions are known to contain the oldest stars in the Galaxy, and the x-ray sources there may be representative of a different class than those in the plane. Very accurate x-ray locations, followed by optical identifications and detailed studies, may help to separate x-ray stars into classes and identify different phenomena responsible for their emission. There are also transient sources which suddenly appear and reach high levels of emission rapidly, then slowly fade over periods of weeks to months. This is similar to the behavior of optical novae. *See* NOVA.

Chandra X-ray Observatory **High-Resolution Camera image of the supernova remnant Cas A. The bright point near the center of the remnant is believed to be the neutron star that is produced during the core collapse of the star that exploded. (*S. S. Murray*)**

Compact sources and binary systems. With accurate locations, it has been possible to identify x-ray stars with binary star systems. The combination of x-ray and optical data leads to the conclusion that these systems usually consist of a relatively normal visible star and one subluminous star in a gravitationally bound orbit about each other. The x-ray star is often a collapsed object, such as a white dwarf, neutron star, or even a black hole. Such compact sources form the majority of galactic emitters.

These sources are further classified by their variability, ranging from time scales as long as years to as short as milliseconds. Some sources have periodic variations resulting from the orientation of their orbits in binary star systems relative to the Earth. When an x-ray-emitting star orbits around a normal star, and is periodically blocked from view, eclipses in the x-ray emission are observed. In some instances the x-rays appear pulsed, with a period of a few seconds or less. This is similar to observations of radio pulsars and is believed to be due to the rotation of a neutron star beaming x-rays toward the Earth.

These eclipsing, x-ray pulsars provide excellent natural laboratories for studying the effects of extremely high gravitational and magnetic fields. The theoretical model for x-ray emission in these systems consists of matter being transferred from the normal star to the compact object. This process, known as accretion, usually leads to the creation of a disk of infalling material spiraling down toward the surface of the compact star. During infall the matter reaches very high temperatures as gravitational energy is converted to heat, which is then radiated as x-rays. The

details of matter transfer, heating, and radiation are active areas of astrophysical research, involving complex relationships between the very strong gravity and magnetic fields that are present in these systems. It is believed that similar processes are involved regardless of whether the compact object is a white dwarf, neutron star, or black hole. However, different types of detailed behavior are expected for each of these objects. Rotating neutron stars lead to x-ray pulsars. For white dwarfs, there appear to be fluctuations in x-ray intensity on time scales of hours to days that might be indicative of changes at the surface of the star where accreted material is collecting, and then flashing in a burst of thermonuclear energy release. In the case of black hole candidates, there are variations in x-ray intensity at extremely short time scales that indicate very small regions of x-ray emission.

Detailed studies of the orbital motions of binary star systems, using both optical and x-ray information, allow estimates to be made of the masses of the total binary system and its component stars. In this manner, it is possible to place a range on the mass of the system's unseen compact object. Theoretical limits on the maximum allowable mass for a white dwarf and neutron star can then be used to help classify the compact object. In this way, for example, a black hole has been identified as the most likely candidate for the x-ray source Cyg X-1, in the constellation Cygnus.

Another type of time behavior is represented by the so-called bursters. These sources emit at some constant level, with occasional short bursts or flares of increased brightness. The bursts last for several minutes, and often there are long quiescent periods followed by active intervals. These outbursts may well be instabilities in the x-ray emission processes associated with this type of source. *See* BINARY STAR; BLACK HOLE; CATACLYSMIC VARIABLE; STELLAR EVOLUTION; WHITE DWARF STAR.

Extragalactic sources. As with galactic x-ray astronomy, virtually all types of objects have been detected as x-ray sources as a result of the increased sensitivity available. These distant objects are radiating enormous quantities of energy in the x-ray band, in some cases more than visible light. Among the more interesting types of objects detected have been apparent normal galaxies, galaxies with active nuclei, radio galaxies, clusters of galaxies, and quasars.

Galaxies and quasars. The nearest spiral galaxy is M31, the Andromeda Galaxy. Imaging x-ray telescopes resolve it into individual stellar sources. The distribution and association with galactic features is similar to the Milky Way. There appears to be very faint emission associated with M31 as a whole, probably due to hot interstellar gas, again similar to diffuse galactic emission detected in the Milky Way. There is also a weak x-ray source located at the center of M31 where a supermassive black hole with about 3×10^7 times the mass of the Sun has been detected. The low x-ray brightness of the nucleus of M31 is an example of a weak nuclear galactic source. *See* ANDROMEDA GALAXY.

Observations of seemingly normal galaxies have led to the discovery of a new class of x-ray source which might be labeled optically dull. These objects emit unusually large fluxes of x-rays, much more than the summed emission of their stars, but appear to be otherwise normal in their optical and radio properties. The mechanisms that produce these x-rays are not well understood. In some cases, detailed studies of optical emission from such galaxies have led to the discovery of signs of nuclear activity in the form of weak emission lines.

In active nuclei galaxies (those with strong optical emission lines and nonthermal continuum spectra), x-ray emission is usually orders of magnitude in excess of that from normal galaxies. It is clearly associated with the galaxy itself and not the summed emission from stars. In most cases the emission comes from the galactic nucleus, and must be confined to a relatively small region on the basis of variability and lack of structure at the current observational limits of a few seconds of arc. Many astrophysicists believe that galactic nuclei are the sites of massive black holes, at least 10^6–10^9 times the mass of the Sun, and that the radiation from these objects is due to gravitational energy released by infalling material. Some scientists speculate that this is a common feature of all galaxies (perhaps with the massive black holes being as small as 10^6 solar masses in so-called normal galaxies), and that the broad range of properties, such as x-ray luminosity, reflects the size of the black hole and availability of infalling matter. Thus, normal galaxies have dormant nuclei where there is little or no material available, while Seyfert galaxies are quite active and quasars represent the extreme case of this mechanism.

In a unified picture of active galaxies, many of the observable properties depend on the viewing angle of the observer, as well as the size of the nuclear black hole. In this model, the black hole lies at the center of an accretion disk (shaped like a doughnut), and beyond this component other clouds of material that give rise to the optical emission lines. If the observer's line of sight is along the axis of the black hole and accretion disk, then the intense x-ray source is seen directly, whereas if the viewing angle is in the plane of the accretion disk, the intense x-ray source is not seen and the object will appear to be much weaker. X-ray observations are particularly relevant in understanding the nature of galactic nuclei, since the high-energy emission is a more direct probe of the basic energy source than optical, infrared, or radio. *See* QUASAR.

Clusters of galaxies. Clusters of galaxies are collections of hundreds to thousands of individual galaxies which form a gravitationally bound system. They are among the largest aggregates of matter in the universe (along with clusters of clusters), and can be detected at very large distances. This allows them to be studied at early epochs in the development of the universe. The space between galaxies in such clusters has been found to contain hot (approximately 10^7–10^8 K) tenuous gas which glows in x-rays. The mass of matter in this gas is equal to (or exceeds)

the mass of the visible galaxies; it contains heavy elements which were not present in the primordial mixture from which stars and galaxies are believed to have formed.

Studies of clusters can be used to help determine the order of formation in the early universe. There are two competing scenarios for the relationship between the time that objects form and their sizes. In one, the largest structures form first and then these evolve to form clusters of galaxies and galaxies. In the other, galaxies form first and then clump together to form the clusters and superclusters. X-ray observations of clusters of galaxies that have subclustering are taken to be examples of cluster formation in process, giving support to the hierarchy of small-to-large formation processes. However, optical observations of the distribution of galaxies over large regions of the universe show large-scale structures that resemble bubbles and voids, with the galaxies tending to be found on the surfaces of these bubbles. These observations are taken to be evidence for the initial formation of large-scale structures followed by smaller objects. Theoretical models tend to include aspects of both scenarios, but further observational data will be needed to clarify the ideas and test the models in more detail. *See* COSMOLOGY; GALAXY, EXTERNAL; UNIVERSE.

X-ray background. In addition to observing a bright source later identified as Sco X-1, the 1962 rocket experiment discovered a uniform background of x-ray emission around the sky. This observation was confirmed by subsequent experiments, and the background was found to contain at least two components. At low energies the background is dominated by galactic emission, while above about 1 keV the radiation comes from beyond the Milky Way Galaxy. The galactic background appears to be diffuse and varies depending upon the viewing direction. It is most likely due to hot (10^6 K) interstellar gas that was heated by a supernova explosion. Because the Milky Way is not very transparent at low x-ray energies, this hot gas is probably nearby (within a few hundred light-years).

The harder component has two possible sources. Either there is a truly diffuse source of emission, such as hot gas or very high energy particles, or the background is the superposition of many discrete sources associated with various extragalactic objects. The spectrum of the background has a smooth shape that can be well approximated by the combined spectra of several types of extragalactic sources or by that expected from a very hot gas. Energy considerations rule out high-energy particles as a source of the background, but large numbers of extragalactic objects have been found by using source counts. Thus, observed classes of individual sources account for at least 80%, and perhaps all, of the background observed. *Chandra* and *XMM-Newton* observations have confirmed this accounting. *ROSAT* and *Chandra* together resolved 90% of the background at energies up to 2 keV into distinct sources, and *Chandra* and *XMM-Newton* resolved most of the background between 2 and 10 keV. All these results limit any remaining contribution from a diffuse hot gas to be, at most, a few percent, and this limits the total mass of hot gas that can be present in the universe to an amount less than that required for gravitational closure. That is, there is not enough light-emitting or normal matter presently identified in the universe to overcome the initial expansion of the big bang and cause a latter contraction. If such matter exists, it must be some heretofore-unobservable form, called dark matter, such as very dim stars; small, isolated black holes; or exotic particles.

The source-count data from deep x-ray surveys indicate that the mix of different classes of extragalactic sources varies with their distance. The strongest and closest extragalactic sources consist mainly of clusters of galaxies and active galaxies such as Seyfert types. At fainter fluxes and thus greater distances, there are mainly active galaxies such as quasars. These observations provide evidence for rapid evolution of quasars in the sense that they were more common in the earlier universe than in the present epoch. *See* ASTROPHYSICS, HIGH-ENERGY.

Stephen S. Murray

Bibliography. B. Aschenbach., H. M. Hahn, and J. E. Trumper, *The Invisible Sky: Rosat and the Age of X-ray Astronomy*, 1998; X. Barcons and A. C. Fabian (eds.), *The X-ray Background*, 1992; W. Lewin, J. Van Paradijs, and E. Van Den Heuvel (eds.), *X-ray Binaries*, 1995; L. Scarsi et al. (eds.), *The Active X-ray Sky: Results from BeppoSAX and RXTE*, 1998; F. D. Seward and P. A. Charles, *Exploring the X-ray Universe*, 1995; W. Tucker and K. Tucker, *Revealing the Universe: The Making of the Chandra X-ray Observatory*, 2001.

X-ray crystallography

The study of crystal structures by x-ray diffraction techniques. The prediction in 1912 by the German physicist Max von Laue that crystals might be employed as natural diffraction gratings in the study of x-rays was experimentally verified in the same year by W. Friedrich and P. Knipping, who obtained diffraction patterns photographically by the so-called Laue method. Almost immediately after (1913), W. Lawrence Bragg not only successfully analyzed the structures of sodium chloride and potassium chloride by Laue photographs but also developed a simple treatment of x-ray scattering by a crystal (the Bragg law) which proved much easier to apply than the more complicated but equivalent Laue theory of diffraction. The availability of the first x-ray spectrometer, constructed by his father, William H. Bragg, as well as the substitution of monochromatic (single-wavelength) for polychromatic x-ray radiation, enabled W. L. Bragg to determine a number of simple crystal structures, including those of diamond; zincblende, ZnS; fluorspar, CaF_2; and pyrites, FeF_2.

For the next 50 years x-ray crystallography was used to determine only a limited number of crystal structures, with relatively low accuracy. This

situation has completely changed since the 1960s due to a combination of several factors: (1) the development of the so-called direct methods for the determination of both centrosymmetric and noncentrosymmetric crystal structures in a straightforward and rapid fashion; (2) the development of computer-controlled automatic x-ray diffractometers for rapidly collecting relatively precise intensity data (at low temperatures); and (3) the availability of highly sophisticated computers for conveniently handling all numerical calculations in the solution and least-squares refinement of each crystal structure. Advantages of modern single-crystal x-ray structural analysis are (1) it is the most powerful method available for the detailed characterization of any crystalline material (including macromolecules such as proteins and nucleic acids); (2) it provides highly accurate stereochemical information (including absolute configurations for optically active species) which in general is complete and unambiguous; and (3) it enables the structures of compounds containing 100–200 independent atoms to be determined very rapidly, that is, in 1–2 days. In the 1960s, 1–2 years were commonly required to work out the structures of relatively simple compounds with less than 20 independent atoms. The tremendous impact of modern x-ray crystallography as an analytical structural tool is apparent from the fact that thousands of structures of organic, inorganic, organometallic, and biological compounds (both small and large) are determined each year by x-ray diffraction. For the theoretical and experimental aspects of x-ray diffraction *see* X-RAY DIFFRACTION.

Bragg's law. Structurally, a crystal is a three-dimensional periodic arrangement in space of atoms, groups of atoms, or molecules. If the periodicity of this pattern extends throughout a given piece of material, one speaks of a single crystal. The exact structure of any given crystal is determined if the locations of all atoms making up the three-dimensional periodic pattern called the unit cell are known. The very close and periodic arrangement of the atoms in a crystal permits it to act as a diffraction grating for x-rays. W. Lawrence Bragg treated the phenomenon of the interference of x-rays with crystals as if the x-rays were being reflected by successive parallel equidistant planes of atoms in the crystal. His important equation relating the perpendicular spacing d of lattice planes in a crystal, the glancing angle θ_{hkl} of the reflected beam, and the x-ray wavelength λ is Eq. (1). This expression provides the basic condition that the difference in path length for waves reflected from successive planes must be an integral number of wavelengths $n\lambda$ in order for the waves reflected from a given set of lattice planes to be in phase with one another. Instead of referring to the nth order of a reflected beam, modern crystallographers customarily redefine Bragg's equation as Eq. (2), where d_{hkl} represents the perpendicular interplanar distance between adjacent lattice planes having the Miller indices (hkl). According to this viewpoint any nth-order diffraction maxima for waves reflected from a set of planes (hkl) with spacing d_{hkl} is equal to a first-order reflection due to a parallel set of planes (nh, nk, nl) with the perpendicular distance given by Eq. (3). For a discussion of Bragg's relation *see* X-RAY POWDER METHODS.

$$n\lambda = 2d \sin\theta_{hkl} \tag{1}$$

$$\lambda = 2d_{hkl} \sin\theta_{hkl} \tag{2}$$

$$d_{nh,nk,nl} = \frac{d_{hkl}}{n} \tag{3}$$

For a crystal with known unit-cell size and shape and with arbitrary orientation in a parallel beam of x-rays of known wavelength, several questions must be answered. First, it must be ascertained what plane, if any, is obeying Bragg's law and is reflecting the rays. Second, the direction of the reflected x-ray must be accurately measured. The Bragg law treatment does not permit these questions to be answered readily, and for this reason the concept of a reciprocal lattice model was introduced. *See* CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Reciprocal lattice. The development and application of the reciprocal lattice to x-ray crystallography have been primarily the results of work by P. P. Ewald (1913 and 1921), J. D. Bernal (1926), and M. J. Buerger (1935). In general, a reciprocal lattice consists of a three-dimensional array of points which is related to the crystal lattice (commonly called the direct lattice) in that each set of planes (hkl) in the crystal lattice is represented in reciprocal space by a point denoted by the coordinates hkl (without parentheses). **Figure 1** illustrates in two dimensions the reciprocal lattice geometrically produced in the following way from a unit crystal cell. An origin is chosen and given the coordinates 000. To every set of parallel planes of Miller index (hkl) in direct space, a reciprocal lattice vector H_{hkl} which is perpendicular to this set of planes is constructed

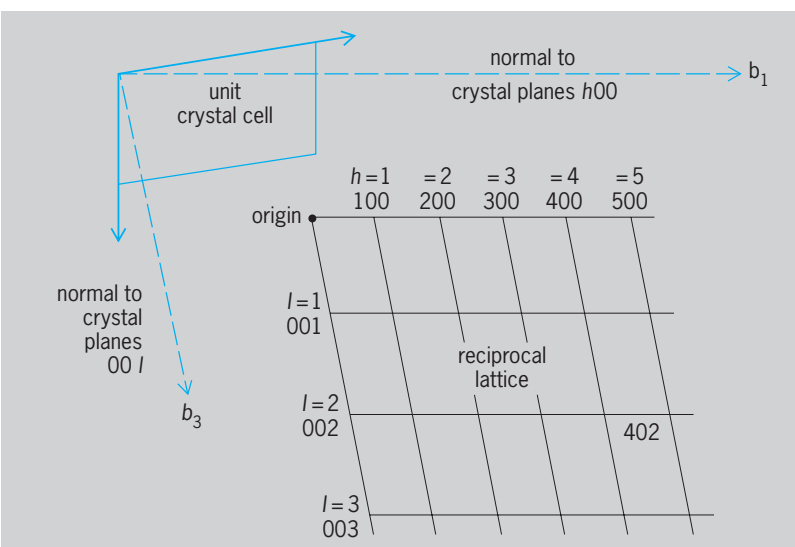

Fig. 1. Reciprocal lattice, shown in two dimensions. Each point in this lattice is reciprocal to a set of planes in the direct lattice.

from the reciprocal lattice origin at a distance inversely proportional to the interplanar spacing d_{hkl}. Thus, the set of parallel direct-lattice planes ($\bar{4}02$) with an interplanar spacing $d_{\bar{4}02}$ is represented in reciprocal space in Fig. 1 by the point $\bar{4}02$. The normal vector $H_{\bar{4}02}$, which is directed along the perpendicular of the ($\bar{4}02$) direct-lattice planes from the origin of the reciprocal lattice to the point with coordinates $\bar{4}02$, is of length $|H_{\bar{4}02}| = K/d_{\bar{4}02}$, where K is an arbitrary constant. The value of K which simply scales the reciprocal lattice is normally taken as unity or as the wavelength λ; for purposes of this discussion it is taken as 1. It can be shown that this procedure applied to a three-dimensional periodic crystal lattice will result in a three-dimensional reciprocal lattice. Each point of coordinates *hkl* in this reciprocal lattice will be reciprocal toa set of planes (*hkl*) in the direct crystal lattice.

Sphere of reflection. The geometrical interpretation of x-ray diffraction from a crystal can be best interpreted by the use of a "sphere of reflection" in reciprocal space (**Fig. 2**). A sphere of radius $1/\lambda$ (corresponding to $K = 1$) is drawn with the direction of the primary x-ray beam (that is, both the incident and transmitted rays) denoted by the unit vector s_0 assumed to travel along a diameter. The crystal C is imagined to be at the center of this sphere of reflection. The origin 000 of the reciprocal lattice is placed at the point O, where the transmitted beam emerges from the sphere of reflection. A diffracted beam will be formed if the surface of the sphere of reflection intercepts any reciprocal lattice point (other than the origin O). In Fig. 2 the reciprocal lattice point P lies on the surface of the sphere of reflection. Hence, two important results follow: (1) the set of crystal planes (*hkl*) to which this point is the reciprocal obeys Bragg's law and reflects the incident x-ray beam; (2) the direction of the diffracted rays will be from the point C at the center of the sphere to where the point P lies on the surface of the sphere. The angle of diffraction between the diffracted beam s and the primary beam s_0 is $2\theta_{hkl}$.

The reciprocal lattice, by virtue of its definition, is tied to the actual lattice insofar as orientation is concerned. As the crystal (and therefore the crystal lattice) is turned about an axis of rotation, the reciprocal lattice, in turning a similar angle about a parallel axis through its origin O, passes through the sphere of reflection (which is fixed for the primary x-ray beam that is stationary relative to the movement of the crystal). In the rotating crystal method all diffraction maxima that can possibly be recorded by any chosen x-ray wavelength λ must be represented by those points *hkl* of the reciprocal lattice that cut the sphere of reflection. These points of possible diffraction maxima all lie within a sphere of radius $2/\lambda$ called the limiting sphere. This geometrical interpretation of Bragg's law enables one to understand readily both the geometry of reflection of x-rays and the manipulation of all the single-crystal cameras and diffractometers now in use.

Crystal structure determination. A modern structural analysis by x-ray diffraction usually involves the following sequential steps: (1) selection and preliminary x-ray characterization of a suitable single crystal followed by collection of intensity data normally with a diffractometer; (2) structural determination by solution of the phase problem to give a correct trial structure; (3) least-squares refinement of the trial structure; and (4) calculation of structural parameters and presentation and display of results.

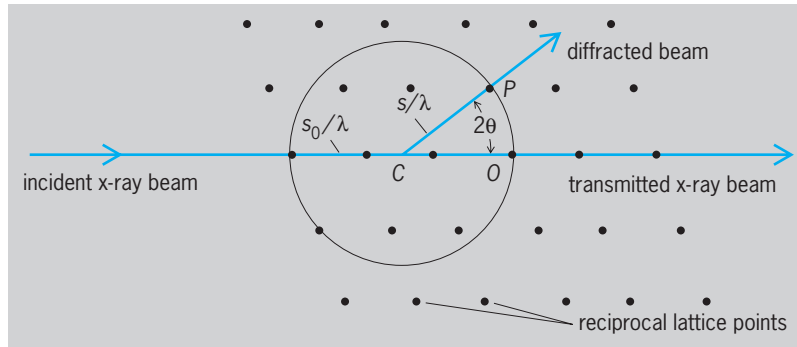

Fig. 2. Sphere of reflection in reciprocal space. Dots represent reciprocal lattice points. Diffraction occurs when any reciprocal lattice point *hkl* cuts the sphere of reflection, as at point *P*.

Collection of intensity data. Since x-rays are scattered by the electrons of the atoms, the intensity of each diffracted beam depends on the positions of the atoms in the unit cell. Any alteration in atomic coordinates would result in changes of the intensities relative to one another. Hence, the first important step in a crystal-structure determination is concerned with the collection of the intensity data by a recording technique which effectively measures the intensity at each reciprocal lattice point.

In the early days of x-ray crystallography, the classical Bragg x-ray spectrometer was used with an ionization chamber to collect data on the intensity of x-ray reflections. It was soon found possible and more convenient to use photographic recording for such data collection. Until the early 1960s the great majority of structural determinations were based on photographically recorded intensities which were usually visually estimated. Nevertheless, this process of photographic data collection was time-consuming, and normally several months were needed to record, judge, average, and scale the intensities of three-dimensional film data. The estimated level of precision of the relative intensities obtained photographically was 15–20%, which was generally sufficient for the solution of stereochemical and conformational problems but was not sufficient for a reliable determination of atomic thermal motion or bonding. Since about 1953 highly stabilized x-ray diffraction units and sensitive, reliable detectors such as scintillation counters have been developed. This modern instrumentation, coupled with the demand for making the intensity collection from crystals both more rapid and more accurate (especially that from proteins because of their instability and large number of data to be recorded), has resulted in a return to direct recording by counter techniques. The commercial availability of automatic

diffractometers which use computerized circuitry to synchronize the movement of the crystal and the detector has effectively revolutionized data collection from crystals such that upward of 1000–3000 diffracted beams can be obtained daily with errors normally less than 3%. *See* SCINTILLATION COUNTER.

Prior to a preliminary x-ray analysis, a single crystal (normally 0.05–0.5 mm in a given dimension) is optically chosen, mounted either on a glass fiber with epoxy or inside a thin-walled glass capillary, and placed on a goniometer head attached to the diffractometer. The preliminary x-ray characterization involves both optical and x-ray alignment with automatic indexing of diffracted beams and selection of a unit cell, from which lattice constants and an orientation matrix are calculated. After selection of data collection parameters, the computer controlling the diffractometer uses the orientation matrix, the *hkl* indices, and wavelength of the monochromatic x-ray beam (normally Cu$K\alpha$ or Mo$K\alpha$ radiation) to calculate angle settings for each diffracted beam to be measured. The intensity $I(hkl)$ and its estimated standard deviation $\sigma(I)$ are measured individually for each reflection with a sensitive scintillation counter detector. Data collection involves the computer-recording of the $I(hkl)$ and $\sigma(I)$ for all independent h,k,l reflections until an entire diffraction pattern has been collected. For a typical organic compound with 20 nonhydrogen atoms, the diffraction pattern may possess approximately 2000 independent reflections. Data reduction involves the correction of the raw $I(hkl)$'s for background, possible radiation-induced crystal decay, Lorentz-polarization effects, and crystal absorption effects, after which the corrected intensities are converted to observed structure factor amplitudes $|F_{hkl}|_o$. The space group symmetry of the unit cell is deduced by examination of the tabulated systematic absences of the $I(hkl)$'s.

Solution of phase problem. Although the size and shape of the unit cell determine the geometry of the diffraction maxima, the intensity of each reflection is determined by the number, character, and distribution of the atoms within the unit cell. The second stage in a structural analysis is the solution of the phase relations among the diffracted beams (that is, the phase problem) from which a correct trial structure can be obtained. The intensity of each diffracted beam is related to the square of its amplitude; that is, Eq. (4)

$$I_{hkl} = k\,|F_{hkl}|_o^2 \tag{4}$$

holds, where $|F_{hkl}|_o$ is the observed structure factor amplitude. Each diffracted beam has not only a characteristic intensity but also a characteristic phase angle α_{hkl} associated with it which expresses the degree to which the diffracted beam is in plane with the other diffracted beams.

Because the electron density is a real, positive quantity which varies continuously and periodically in a crystal, the electron scattering density $\rho(xyz)$ is derivable from the three-dimensional Fourier series as in Eq. (5), where $\rho(xyz)$ represents the electron

$$\rho(xyz) = \frac{1}{V}\sum_h\sum_k\sum_l |F_{hkl}|_o \times \cos[2\pi(hx + ky + lz) - \alpha_{hkl}] \tag{5}$$

density at any point with fractional coordinates x, y, z in the unit cell of volume V. Hence, if the characteristic amplitude $|F_{hkl}|_o$ and its phase α_{hkl} for each diffracted beam are known, the electron density can be calculated at fractional grid points in the unit cell. A properly phased three-dimensional electron-density map effectively provides peaks characteristic of direct images of atoms in the unit cell. The phase problem in x-ray crystallography arises because experimental measurements yield only the magnitudes of the structure factors $|F_{hkl}|_o$ but not the phases α_{hkl}. Hence, a structural analysis involves a search for the characteristic phases to be utilized together with the observed amplitudes in order to obtain a three-dimensional electron-density map and thereby to determine the crystal structure.

The structure factor itself is related to the scattering by the atoms in the unit cell by Eq. (6), where

$$F_{hkl} = \sum\nolimits_n f_n T_n \exp[2\pi i(hx_n + ky_n + lz_n)] \tag{6}$$

x_n, y_n, z_n are the fractional coordinates of atom n along the three crystallographic axes. The f_n's are the individual atomic scattering factors which are known for an atom at rest. At zero 2θ angle of scattering for which all the electrons of an atom scatter in phase, f_n is equal to the atomic number. The T_n's are the individual modifications which cause a decrease of the f_n's as a result of thermal motion. If an atom n is assumed to vibrate equally in all directions, its thermal behavior is designated by an isotopic temperature factor B_n (in nm^2) given by Eq. (7); a large

$$T_n = \exp(-B_n \sin^2\theta_{hkl}/\lambda^2) \tag{7}$$

B_n value (greater than 0.05 nm^2) corresponds to a large thermal displacement of an atom from its mean position and hence to a small T_n value. In general, atomic thermal motion varies with direction such that the isotropic B_n of an atom n is replaced by six atomic thermal parameters which describe an ellipsoidal electron distribution. Hence, the thermal motion of an atom is conventionally represented by a probability thermal ellipsoid (that is, an atom has a given probability, usually 50%, of being inside the thermal ellipsoid). If the positions of the atoms in the unit cell are known, the complex structure factor for each diffracted beam *hkl* can be calculated and both its magnitude and phase obtained by Eq. (8).

$$F_{hkl} = A_{hkl} + iB_{hkl} \tag{8}$$

$$A_{hkl} = \sum\nolimits_n f_n T_n \cos 2\pi(hx_n + ky_n + lz_n)$$

$$B_{hkl} = \sum\nolimits_n f_n T_n \sin 2\pi(hx_n + ky_n + lz_n)$$

$$i = \sqrt{-1}$$

Consequently, the calculated structure factor amplitude $|F_{bkl}|_c$ and the associated phase α_{bkl} can be obtained from Eq. (9).

$$|F_{bkl}|_c = \sqrt{A_{bkl}^2 + B_{bkl}^2} \tag{9}$$

$$\tan \alpha_{bkl} = \frac{B_{bkl}}{A_{bkl}}$$

For centrosymmetric crystals which have a center of symmetry located at the corner of each unit cell, the structure factor F_{bkl} simplifies to a real rather than a complex number; that is, since for each atom at x, y, z, there must exist a centrosymmetrically related atom at $-x, -y, -z$, it follows that $B_{bkl} = 0$ and $F_{bkl} = \pm A_{bkl}$. The phases of the diffracted beams then are restricted to being either completely in phase (that is, $\alpha_{\mathrm{hkl}} = 0^\circ$ corresponding to $F_{bkl} = +A_{bkl}$) or completely out of phase (that is, $\alpha_{bkl} = 180^\circ$ corresponding to $F_{bkl} = -A_{bkl}$) with one another.

Various means of deducing the phase angle of each diffracted beam have been used. Prior to the advent of direct methods, the main approach used to overcome the phase problem was based on the discovery by A. L. Patterson in 1935 that when the experimentally known quantities $|F_{bkl}|_0^2$ instead of F_{bkl} are used as coefficients in the Fourier series, the maxima in this summation then correspond not to atomic centers but rather to a map of interatomic vectors. This vector map represents a superposition of all interatomic vectors between pairs of atoms translated to the origin of the unit cell. Since the peak height for a given vector between two atoms is approximately proportional to the product of their atomic numbers, the vectors between heavy atoms (that is, those of high atomic number) usually stand out strongly against the background of heavy-light and light-light atom vectors. Consequently, the Patterson map is especially applicable to the structure determination of compounds containing a small number of relatively heavy atoms because approximate coordinates of these heavy atoms can be obtained provided the heavy-atom vectors are correctly recognized in the Patterson map. Computers have enabled the development of a number of powerful techniques in unraveling the Patterson vector map in terms of atomic coordinates; these include multiple superpositions of parts of the Patterson map and "image-seeking" with known vectors.

Direct methods of phase determination presently constitute the universal approach to solving the phase problem for both centrosymmetric and noncentrosymmetric crystal structures. Direct methods make use of probability theory to generate an adequate set of phases by consideration solely of statistical relationships among the observed structure factor amplitudes of the diffracted beams. In order to utilize direct methods, the observed structure factor amplitudes are first converted to normalized structure factor amplitudes defined by Eq. (10), where ϵ is a small integer (normally 1) that is space-group-dependent. The distribution of the $|E_{bkl}|$'s is in principle independent of the size and content of the unit cell. The application of the so-called tangent formula, first derived by J. Karle and H. Hauptman in 1956, to reflections with large $|E_{bkl}|$'s provides sets of probable phase relationships. These E-value reflections with their most probable phases are normally then utilized to calculate Fourier E-maps (that is, the $|E_{\mathrm{bkl}}|$'s are coefficients instead of the $|F_{bkl}|_o$'s) whose peak maxima are generally sufficiently resolved that they can be interpreted in terms of an approximate structure. Direct-methods computer programs such as MULTAN have been particularly successful in routinely solving the crystal structures of a wide variety of biological molecules. Although these computer programs cannot as yet be used alone to determine the structures of macromolecules such as proteins, extensive research is being carried out to apply the methodology to the structural solution of large molecules.

$$|E_{bkl}| = \left(\frac{|F_{bkl}|_o^2}{\epsilon_i \sum_n f_n^2} \right)^{1/2} \tag{10}$$

Fourier electron-density and ΔF maps. Any resulting trial structure consisting of initial coordinates for some of the atoms is often sufficient for location of the other atoms by the application of the method of successive Fourier electron-density maps. The phases calculated from the initial parameters of the presumably known atoms, together with the observed $|F_{bkl}|$, are utilized to compute a density map. New coordinates are obtained not only for the peaks corresponding to these known atoms but also for the other

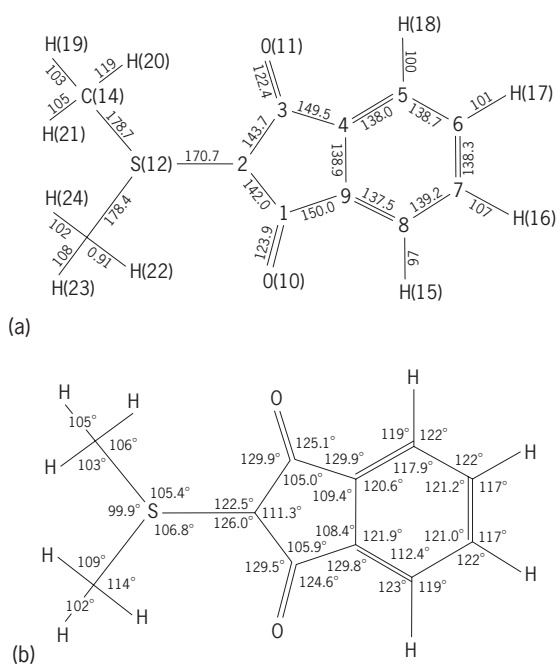

Fig. 3. Configuration of 2-dimethylsulfuranylidene-1,3-indanedione molecule which was refined by least squares to a final *R* value of 0.028 for 1170 independent diffractometry-collected *hkl* data. (*a*) Bond lengths in picometers. (*b*) Bond angles. Estimated standard deviations involving the nonhydrogen atoms are in the 0.3–0.5-pm range for distances and 0.1–0.3° range for bond angles, while those involving the hydrogen atoms are approximately 5 pm and 2.8°.

peaks which it is hoped can be interpreted from stereochemical considerations as being due to additional atoms in the structure. The Fourier process is then reiterated, with the new phases calculated from the modified coordinates of the previous set of atoms plus the coordinates of newly located atoms. If a correct distinction between the "true" and the "false" peaks is made, the electron-density function usually converges to give the entire crystal structure.

Another Fourier synthesis which is often used is a difference synthesis, $\Delta\rho(xyz) = \rho_o(xyz) - \rho_c(xyz)$, which is given by Eq. (11), where ΔF is given by Eq. (12) and α_{hkl} is the calculated phase of $|F(hkl)_c|$.

$$\Delta\rho(xyz) = \frac{1}{V}\sum_h\sum_k\sum_l \Delta F \exp(i\alpha_{hkl}) \times \exp[-2\pi i(hx + ly + lz)] \quad (11)$$

$$\Delta F = |F(hkl)|_o - |F(hkl)|_c \quad (12)$$

The particular virtue of this ΔF map is that it eliminates many of the false maxima which are associated with termination of series errors. A difference map is commonly used to locate hydrogen atoms whose positions may be obscured in a regular Fourier electron density map by the relatively heavy nonhydrogen atoms. It provides a means of subtracting out the electron density of the other atoms in the structure.

Refinement of parameters. Once the phase problem is solved and the approximately correct trial structure is known, the next step of the structural analysis involves the refinement of the positional and thermal parameters of the atoms. Normally this refinement is carried out analytically by the application of a nonlinear least-squares procedure in which a weighted quantity such as $\Sigma w[|F(hkl)|_o - |F(hkl)|_c]^2$ (where the weights w are appropriate to the experiment) is minimized with respect to the parameters. This method of refinement, which gives the best fit between the calculated and observed structure factor amplitudes, yields the most precise values for the parameters. Although there is no single reliable method for directly assessing the accuracy of a structural determination, a criterion commonly used is the unweighted reliability factor or discrepancy

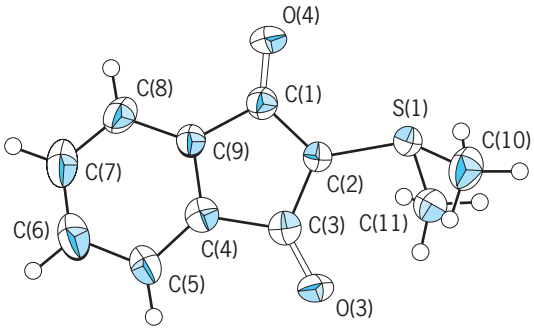

Fig. 4. Molecular configuration of the crystallographically independent $C_{11}O_2H_{10}S$ molecule. The thermal ellipsoids for the nonhydrogen atoms, which were refined with anisotropic thermal parameters, are of 50% probability.

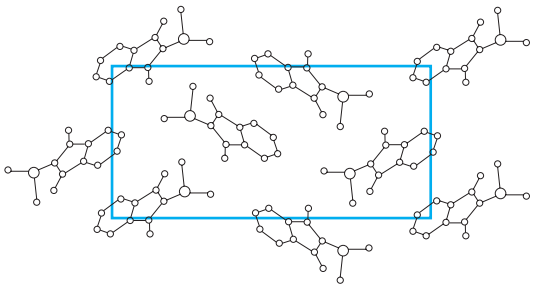

Fig. 5. Unit-cell diagram showing the arrangement of the four symmetry-related $C_{11}O_2H_{10}S$ molecules.

index R_1, defined by Eq. (13) as the summation of

$$R_1 = \frac{\sum_{hkl}\left|\,|F(hkl)|_o - |F(hkl|_c\,\right|}{\sum_{hkl}|F(hkl)|_o} \quad (13)$$

the absolute difference in the observed and calculated structure amplitudes divided by the summation of the observed amplitudes. The better the structure, including the atomic coordinates and thermal parameters, is known, the more nearly will the calculated amplitudes agree with the observed ones, and hence the lower will be the R_1 value. Discrepancy values of "finished" modern structural analyses found in the literature vary from approximately 0.12 to less than 0.02, depending upon a number of factors, such as the complexity of the structure and the number and quality of the data obtained.

Before the development of large electronic computers the determination and refinement of a crystal structure of more than about 20 atoms was not feasible. Nowadays, single-crystal analyses of uncomplicated structures may require the location of as many as 200 atoms, but there is no high correlation between the complexity of the structural determination and the number of atoms involved.

Calculations and presentations of results. The output of the least-squares refinement provides atomic parameters, namely, three x, y, z coordinates and either one isotropic or six anisotropic thermal parameters per atom, depending upon the thermal model used in the refinement process. These refined parameters and their estimated standard deviations may be used in the last step of the structural analysis to obtain the following information: (1) interatomic distances and bond angles with estimated standard deviations; (2) "best" least-squares planes of specified groups of atoms; (3) torsional angles; and (4) atomic thermal behavior. Interactive graphics programs on an appropriate computer terminal may then be utilized to generate various plots of the structure (**Figs. 3–5**).

Lawrence F. Dahl

Bibliography. T. L. Blundell and L. N. Johnson, *Protein Crystallography*, 1976; J. P. Glusker and K. N. Trueblood, *Crystal Structure Analysis: A Primer*, 2d ed., 1985; M. F. C. Ladd and R. A. Palmer, *Structure Determination by X-ray Crystallography*, 3d ed., 1993; P. Luger, *Modern X-ray Analysis on Single*

Crystals, 1980; G. H. Stout and L. H. Jensen, *X-ray Structure Determination: A Practical Guide,* 2d ed., 1989.

X-ray diffraction

The scattering of x-rays by matter with accompanying variation in intensity in different directions due to interference effects. X-ray diffraction is one of the most important tools of solid-state chemistry, since it constitutes a powerful and readily available method for determining atomic arrangements in matter. X-ray diffraction methods depend upon the fact that x-ray wavelengths of the order of 1 nanometer are readily available and that this is the order of magnitude of atomic dimensions. When an x-ray beam falls on matter, scattered x-radiation is produced by all the atoms. These scattered waves spread out spherically from all the atoms in the sample, and the interference effects of the scattered radiation from the different atoms cause the intensity of the scattered radiation to exhibit maxima and minima in various directions. *See* DIFFRACTION.

Some of the uses of x-ray diffraction are (1) differentiation between crystalline and amorphous materials; (2) determination of the structure of crystalline materials (crystal axes, size and shape of the unit cell, positions of the atoms in the unit cell); (3) determination of electron distribution within the atoms, and throughout the unit cell; (4) determination of the orientation of single crystals; (5) determination of the texture of polygrained materials; (6) identification of crystalline phases and measurement of the relative proportions; (7) measurement of limits of solid solubility, and determination of phase diagrams; (8) measurement of strain and small grain size; (9) measurement of various kinds of randomness, disorder, and imperfections in crystals; and (10) determination of radial distribution functions for amorphous solids and liquids.

For the study of crystal structure by x-ray diffraction techniques *see* X-RAY CRYSTALLOGRAPHY.

Diffraction Theory

When x-rays fall on the atoms of a substance, the scattered radiation is of two kinds: Compton modified scattering of increased wavelength which is incoherent with respect to the primary beam, and unmodified scattering coherent with the primary beam. Because of interference effects from the unmodified scattering by the different atoms of the sample, the intensity of unmodified scattering varies in different directions. A diagram of this variation in direction of intensity of unmodified scattering is called the diffraction pattern of the substance. This pattern is determined by the kinds of atoms and their arrangement in the sample; for simple structures the atomic arrangement is readily deduced from the diffraction pattern. *See* COMPTON EFFECT.

The atomic scattering factor f is defined as the ratio of the amplitude of unmodified scattering by an atom to the amplitude of scattering by a free electron, which scatters according to classical theory. In general, f is a real number which decreases with $(\sin\theta)/\lambda$, where θ is the grazing angle and λ the wavelength, from an initial value $f = Z$, where Z is the number of electrons in the atom. However, if the x-ray wavelength is close to an absorption edge of the atom, f becomes complex. For a definition of absorption edges *see* X-RAY FLUORESCENCE ANALYSIS.

If the electron density in the atom has spherical symmetry and if the x-ray wavelength is small compared to all the absorption-edge wavelengths, Eq. (1)

$$f = \int_0^\infty 4\pi r^2 \rho(r) \frac{\sin kr}{kr} dr \tag{1}$$

holds. Here $k = 4\pi(\sin\theta)/\lambda$ and $\rho(r)$ is the electron density (electrons per unit volume).

A crystalline structure is one in which a unit of structure called the unit cell repeats at regular intervals in three dimensions. The repetition in space is determined by three noncoplanar vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, called the crystal axes. The positions of the atoms in the unit cell are expressed by a set of base vectors $\mathbf{r}_n$. The position of atom n in the unit cell $q_1q_2q_3$ is given by Eq. (2).

$$\mathbf{R}_{nq} = q_1\mathbf{a}_1 + q_2\mathbf{a}_2 + q_3\mathbf{a}_3 + \mathbf{r}_n \tag{2}$$

See CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

For a crystal containing $N_1N_2N_3$ repetitions in the $\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3$ directions, the intensity of unmodified scattering is given by Eq. (3). Here I_e is the intensity, at

$$\begin{aligned} I &= I_e \sum_{nq} f_n \exp\left[\frac{2\pi i}{\lambda}(\mathbf{s}-\mathbf{s}_0)\times\mathbf{R}_{nq}\right] \\ &\quad \times \sum_{n'q'} f_{n'} \exp\left[\frac{-2\pi i}{\lambda}(\mathbf{s}-\mathbf{s}_0)\cdot\mathbf{R}_{n'q'}\right] \\ &= I_e FF^* \frac{\sin^2[(\pi/\lambda)(\mathbf{s}-\mathbf{s}_0)\cdot N_1\mathbf{a}_1]}{\sin^2[(\pi/\lambda)(\mathbf{s}-\mathbf{s}_0)\cdot\mathbf{a}_1]} \\ &\quad \times \frac{\sin^2[(\pi/\lambda)(\mathbf{s}-\mathbf{s}_0)\cdot N_2\mathbf{a}_2]}{\sin^2[(\pi/\lambda)(\mathbf{s}-\mathbf{s}_0)\cdot\mathbf{a}_2]} \\ &\quad \times \frac{\sin^2[(\pi/\lambda)(\mathbf{s}-\mathbf{s}_0)\cdot N_3\mathbf{a}_3]}{\sin^2[(\pi/\lambda)(\mathbf{s}-\mathbf{s}_0)\cdot\mathbf{a}_3]} \end{aligned} \tag{3}$$

a distance R and angle 2θ, scattered by a free electron according to classical theory, and $FF* = |F|^2$. For an unpolarized primary beam of intensity I_o, Eq. (4)

$$I_e = I_o \frac{e^4}{m^2c^4R^2}\left(\frac{1+\cos^2 2\theta}{2}\right) \tag{4}$$

holds, and $e^4/(m^2c^4) = 7.94 \times 10^{-26}$ cm^2 if R is expressed in centimeters. Here m is the mass of the electron, c is the velocity of light, and F is the structure factor, a complex quantity given by a summation over all the atoms of the unit cell as in Eq. (5), where

$$F = \sum_n f_n \exp\left[\frac{2\pi i}{\lambda}(\mathbf{s}-\mathbf{s}_0)\cdot\mathbf{r}_n\right] \tag{5}$$

$\mathbf{s}_0$ and $\mathbf{s}$ are unit vectors in the directions of the primary and diffracted beams.

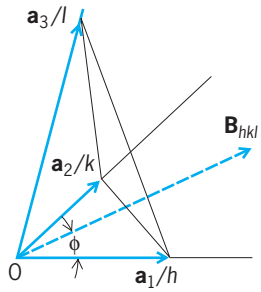

Fig. 1. Crystallographic planes with Miller indices *hkl*.

Laue equations and Bragg's law. The condition for a crystalline reflection is that the three quotients of Eq. (3) exhibit maxima, and this occurs if all three denominators vanish. Expressing the denominators in terms of three integers α, β, and γ, the three Laue equations (6) are obtained. These express the condition for a diffracted beam.

$$\begin{aligned}(\mathbf{s}-\mathbf{s}_0)\cdot\mathbf{a}_1 &= \alpha\lambda\\ (\mathbf{s}-\mathbf{s}_0)\cdot\mathbf{a}_2 &= \beta\lambda\\ (\mathbf{s}-\mathbf{s}_0)\cdot\mathbf{a}_3 &= \gamma\lambda\end{aligned} \tag{6}$$

It is convenient to introduce the concept of sets of crystallographic planes. As illustrated by **Fig. 1**, the set of planes with Miller indices *hkl* is a set of parallel equidistant planes, one of which passes through the origin, and the next nearest makes intercepts $\mathbf{a}_1/h$, $\mathbf{a}_2/k$, $\mathbf{a}_3/l$ on the three crystallographic axes.

In terms of sets of planes *hkl*, the diffraction conditions are expressed by the Bragg law, Eq. (7), where θ

$$\lambda = 2d_{hkl}\sin\theta \tag{7}$$

is the angle which the primary and diffracted beams make with the planes *hkl* and d_{hkl} is the spacing of the set. As seen from **Fig. 2**, the Bragg law is simply the condition that the path difference for rays diffracted from two successive *hkl* planes be one wavelength. In the early days of x-ray diffraction, the Bragg law was written $n\lambda = 2d\sin\theta$, and $n = 1$, 2, 3 corresponded to first-, second-, and third-order diffraction from the planes of spacing d. That notation has been largely dropped, and instead of being called second-order diffraction from planes *hkl*, it is called diffraction from the planes $2h$, $2k$, $2l$. For an extended discussion of the Bragg law *see* X-RAY POWDER METHODS.

Reciprocal lattice. The understanding and interpretation of x-ray diffraction in crystals is greatly facilitated by the concept of a reciprocal lattice. In terms of the crystal axes $\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3$, three reciprocal vectors are defined by Eqs. (8). From these definitions it follows that Eq. (9) is valid. In terms of integers *hkl*, the terminal points of the vectors in Eq. (10) generate a

$$\mathbf{b}_1 = \frac{\mathbf{a}_2\times\mathbf{a}_3}{a_1\cdot\mathbf{a}_2\times\mathbf{a}_3} \qquad \mathbf{b}_2 = \frac{\mathbf{a}_3\times\mathbf{a}_1}{a_1\cdot\mathbf{a}_2\times\mathbf{a}_3} \qquad \mathbf{b}_3 = \frac{\mathbf{a}_1\times\mathbf{a}_2}{\mathbf{a}_1\cdot\mathbf{a}_2\times\mathbf{a}_3} \tag{8}$$

$$\mathbf{a}_i\cdot\mathbf{b}_j = \begin{cases}1 & i=j\\ 0 & i\neq j\end{cases} \tag{9}$$

$$\mathbf{B}_{hkl} = h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3 \tag{10}$$

lattice of points called the reciprocal lattice. Each point in the lattice is specified by the integers *hkl*, and the vectors $\mathbf{B}_{hkl}$ represent two important properties of the sets of *hkl* planes: (1) $\mathbf{B}_{hkl}$ is perpendicular to the *hkl* planes, and (2) $|\mathbf{B}_{hkl}| = 1/d_{hkl}$. These two relations are readily proved from the geometry of Fig. 1. As seen, $\mathbf{a}_2/k - \mathbf{a}_1/h$ and $\mathbf{a}_3/l - \mathbf{a}_2/k$ are vectors lying in the *hkl* plane. From Eqs. (9) and (10), Eqs. (11)

$$\begin{aligned}\left(\frac{\mathbf{a}_2}{k}-\frac{\mathbf{a}_1}{h}\right)\cdot\mathbf{B}_{hkl} &= 0\\ \left(\frac{\mathbf{a}_3}{l}-\frac{\mathbf{a}_2}{k}\right)\cdot\mathbf{B}_{hkl} &= 0\end{aligned} \tag{11}$$

are obtained, and hence $\mathbf{B}_{hkl}$ is perpendicular to the planes *hkl*. The spacing of the planes *hkl* is given by Eq. (12).

$$d_{hkl} = \left|\frac{\mathbf{a}_1}{h}\right|\cos\phi = \frac{\mathbf{a}_1}{h}\cdot\frac{\mathbf{B}_{hkl}}{|\mathbf{B}_{hkl}|} = \frac{1}{|\mathbf{B}_{hkl}|} \tag{12}$$

Equivalence of the three Laue equations and the Bragg law can be shown as follows: Any vector $\mathbf{r}$ can be expressed by Eq. (13). Let $\mathbf{r}$ be the vector

$$\mathbf{r} = (\mathbf{r}\cdot\mathbf{a}_1)\,\mathbf{b}_1 + (\mathbf{r}\cdot\mathbf{a}_2)\,\mathbf{b}_2 + (\mathbf{r}\cdot\mathbf{a}_3)\mathbf{b}_3 \tag{13}$$

$(\mathbf{s}-\mathbf{s}_0)$ and combine it with the three Laue equations and Eq. (13) to obtain Eq. (14). The Bragg law

$$\mathbf{s}-\mathbf{s}_0 = \lambda\,(\alpha\mathbf{b}_1 + \beta\mathbf{b}_2 + \gamma\mathbf{b}_3) \tag{14}$$

can be written in vector form as Eq. (15) since the

$$\mathbf{s}-\mathbf{s}_0 = \lambda\mathbf{B}_{hkl} = \lambda(h\mathbf{b}_1 + k\mathbf{b}_2 + l\mathbf{b}_3) \tag{15}$$

usual form of the Bragg law is simply an equality in the magnitudes of the vectors: $|\mathbf{s}-\mathbf{s}_0| = 2\sin\theta$ and $|\mathbf{B}_{hkl}| = 1/k_{hkl}$. Comparison of Eqs. (14) and (15) shows that the integers α, β, γ of the three Laue equations are simply the Miller indices *hkl* of the Bragg law.

The positions of the atoms in the unit cell are represented by a set of atomic coordinates x_n, y_n, z_n

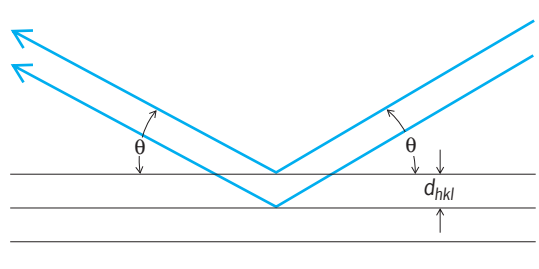

Fig. 2. Interference conditions involved in Bragg law.

such that for atom n Eq. (16) holds. For a Bragg law

$$\mathbf{r}_n = \mathbf{a}_1 x_n + \mathbf{a}_2 y_n + \mathbf{a}_3 z_n \tag{16}$$

reflection hkl, the structure factor takes the simple form of Eq. (17).

$$F_{hkl} = \sum_n f_n \exp\left[2\pi i\left(hx_n + ky_n + lz_n\right)\right] \tag{17}$$

Integrated intensity. In general, the intensity of a Bragg reflection, as expressed by Eq. (3), is not an experimentally measurable quantity. Other factors, such as the degree of mosaic structure in the crystal and the degree of parallelism of the primary beam, have a profound influence on the measured diffracted intensity for any setting of the crystal. To obtain measurements characteristic of the crystalline structure, it is necessary to adopt a more useful concept, the integrated intensity. For a small single crystal, it is postulated that the crystal is to be turned at constant angular velocity ω through the Bragg law position, and that the total diffracted energy of the reflection is to be measured. The integrated intensity E is then given by Eq. (18), where $d\alpha$ is a change in

$$E = \int\int I\frac{d\alpha}{\omega}dA \tag{18}$$

orientation of the crystal and dA is an element of area at the point of observation.

Most of the equations used in x-ray diffraction studies are derived on the assumption that the intensity of the diffracted beam is so small that any interaction with the primary beam can be neglected. These are classed as the equations for the ideally imperfect crystal. For powder samples, in which the individual crystals are extremely small, and for highly deformed single crystals, if the intensity of the diffracted beam is small, the ideally imperfect crystal is usually a good approximation. For the ideally perfect crystal, it is necessary to use a more elaborate theory which allows for the interaction of diffracted radiation with the primary beam. In general, it is the integrated intensity which is measured, and theory shows that the integrated intensity for an ideally imperfect crystal is larger than that for the ideally perfect crystal. Many of the crystalline samples used for x-ray diffraction studies are not ideally imperfect, and the measured integrated intensity is accordingly less than that predicted by the ideally imperfect crystal formulas which are used in the interpretation. The situation is usually handled by adding a correction factor called the extinction correction to the formulas for the integrated intensity from the ideally imperfect crystal.

Atomic coordinates. To have complete information about a crystalline structure, it is necessary to know all the atomic coordinates $x_n y_n z_n$ of the n atoms making up the unit cell. The atomic coordinates appear in the structure factor as given by Eq. (17), and sometimes the coordinates are obtained directly from structure factor values. Another way is to plot the electron density in the unit cell and infer the atomic positions from peaks in the electron density function. The electron density in the unit cell is given by the triple Fourier series shown in Eq. (19)

$$\rho(xyz) = \frac{1}{V}\sum_h\sum_k\sum_l F_{hkl} \times \exp\left[-2\pi i\left(\frac{hz}{a} + \frac{ky}{b} + \frac{lz}{c}\right)\right] \tag{19}$$

for which the coefficients are simply the structure factors F_{hkl}. However, from experimental measurements of either an intensity or an integrated intensity, values for $|f_{\mathrm{hkl}}|^2$ can be obtained. These yield the magnitude of F_{hkl} but not the phase. This is the most serious limitation to a straightforward determination of crystalline structures by x-ray diffraction methods. The ambiguity in the phase of F_{hkl} prevents the use of the Fourier plot of Eq. (19) as a general method for determining any crystalline structure.

Simple structures are uniquely determined by combining the x-ray intensity results with space group theory. The space group of a crystal is the repeating spatial arrangement of symmetry elements which the structure displays. Considering all the possible symmetry elements which can exist in a crystalline structure, group theory shows that there are only 230 essentially different possible combinations, and these constitute the 230 space groups. A knowledge of the macroscopic symmetry of the crystal, coupled with the systematically vanishing x-ray reflections, usually determines the space group. The limitations imposed by the space group on the possible atomic positions, coupled with the limitations imposed by the measured $|F_{hkl}|^2$, often allow a complete and unique structure determination for not too complicated structures. For highly complex structures, it is never certain that x-ray diffraction analysis can yield a complete structure determination. Additional techniques such as the isomorphous replacement by heavy atoms, the use of Patterson plots, and the determination of phase relations from inequalities are used with success on some of the complex structures. *See* GROUP THEORY.

Many structures of interest in solid-state chemistry exhibit various kinds of randomness and imperfections. The precise nature of these is sometimes of more interest than the ideal average structure. Randomness and imperfections in a structure show themselves by producing a diffuse intensity in addition to the sharp Bragg reflections. The temperature vibration of the atoms produces a diffuse intensity called temperature diffuse scattering. Quantitative measurements of this scattering lead to values for the velocity of high-frequency elastic waves and to a complete experimental determination of the spectrum of the elastic waves which constitute the thermal vibrations of the crystal. In alloys showing order-disorder changes, the short-range order parameters are obtained from quantitative measurement of the diffuse intensity which results from randomness in the atomic arrangement.

Crystalline Diffraction

The techniques employed in the study of crystalline substances are discussed below.

Laue method. The Laue pattern uses polychromatic x-rays provided by the continuous spectrum from an x-ray tube operated at 35–50 kV. The transmission Laue pattern is obtained by passing a finely collimated beam through a thin single crystal and recording the diffracted beams on a photographic film placed several centimeters beyond the crystal. For each set of planes *hkl*, θ is fixed, and the Bragg law is satisfied by selecting the proper λ from the primary beam. In a Laue pattern, the different diffracted beams have different wavelengths, and their directions are determined solely by the orientations of the *hkl* planes. Transmission Laue patterns were once used for structure determinations, but their many disadvantages have made them practically obsolete.

On the other hand, the back-reflection Laue pattern is used a great deal in the study of the orientation of crystals. The back-reflection Laue camera is shown schematically in **Fig. 3**. The polychromatic beam enters through a hole in the x-ray film and falls on a single crystal whose orientation can be set as desired by a system of goniometer circles. Diffracted beams bent through angles 2θ approaching 180° are registered on the photographic film. For cubic crystals it is very easy to read the crystal orientation from a back-reflection Laue pattern, and the patterns find considerable use in the cutting of single-crystal metal ingots.

Rotating crystal method. The original rotating crystal method was employed in the Bragg spectrometer. A sufficiently monochromatic beam, of wavelength of the order of 0.1 nanometer, is obtained by using the strong $K\alpha_1\alpha_2$ doublet with a filter which suppresses the $K\beta$ line and much of the continuous spectrum. The beam is collimated by a system of slits and then falls on the large extended face of a single crystal as shown by **Fig. 4**. Originally the diffracted beam was measured with an ionization chamber, but Geiger counters and proportional counters have largely replaced the ionization chamber. Both the crystal and the chamber turn about the spectrometer axis.

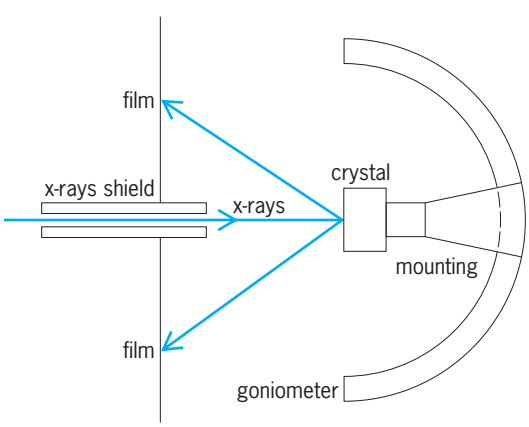

Fig. 3. Schematic of back-reflection Laue camera.

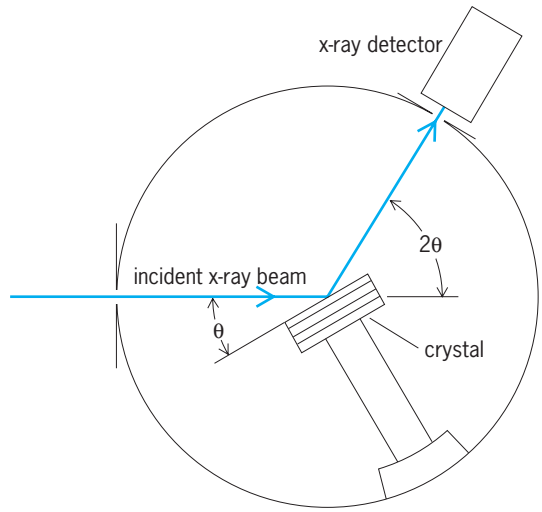

Fig. 4. Schematic of Bragg spectrometer.

The Bragg spectrometer has been used extensively in obtaining quantitative measurements of the integrated intensity from planes parallel to the face of the crystal. The chamber is set at the correct 2θ-angle with a slit so wide that all of the radiation reflected from the crystal can enter and be measured. The crystal is turned at constant angular speed ω through the Bragg law position, and the total diffracted energy E received by the ionization chamber during this process is measured. Similar readings with the chamber set on either side of the peak give a background correction. For this type of measurement, the integrated intensity E is given by Eq. (20), where P_0 is the power

$$E = \frac{P_0}{2\mu\omega}\frac{e^4}{m^2c^4}\frac{\lambda^3F^2}{v^2} \times \left(\frac{1+\cos^2 2\theta}{2\sin 2\theta}\right)\exp[-2M] \quad (20)$$

of the primary beam, μ is the linear absorption coefficient in the crystal, ω is the angular velocity of the crystal, λ is the x-ray wavelength, F is the structure factor, v is the volume of the unit cell, and $\exp[-2M]$ is the Debye factor allowing for temperature vibration. When more than one kind of atom is present, this factor must be incorporated in F^2.

Measurements of the integrated intensity E give quantitative values of F^2 directly. When a Geiger counter is used in place of an ionization chamber for this type of measurement, it is necessary to employ a narrow counterslit and traverse the counter through the reflected beam since the sensitivity of a Geiger counter is not constant over a large window opening.

The rotation camera, which is frequently used for structure determinations, is illustrated in **Fig. 5**. The monochromatic primary beam $\mathbf{s}_0$ falls on a small single crystal at O. The crystal is mounted with one of its axes (say, $\mathbf{a}_3$) vertical, and it rotates with constant velocity about the vertical axis during the exposure. The various diffracted beams are registered on a cylindrical film concentric with the axis of rotation. For a rotation about $\mathbf{a}_3$ it follows that $\mathbf{s}_0 \cdot \mathbf{a}_3 = 0$, and

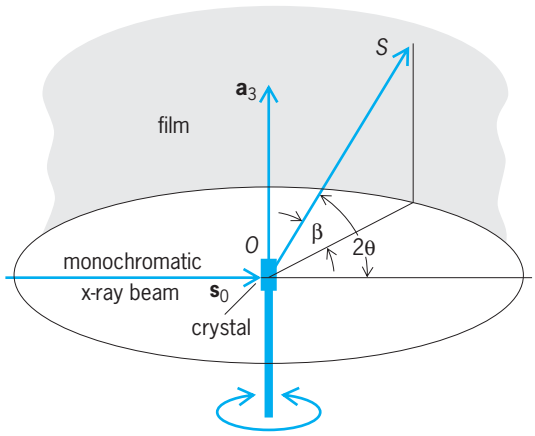

Fig. 5. Schematic of rotation camera.

the third Laue equation gives Eq. (21). The diffracted

$$\sin\beta = \frac{l\lambda}{|\mathbf{a}_3|} \tag{21}$$

beams form the elements of a set of cones, and the intersection of these cones with the cylindrical film gives a set of horizontal lines of diffraction spots. This type of pattern is called a rotation pattern, and the horizontal rows of spots are called layer lines. As seen from Eq. (21), the measured values of $\sin\beta$ give directly the length of the axis about which the crystal was rotated, and the layer line in which a spot occurs gives the l index of the reflection. Similar rotations about the other two axes give corresponding information. More elaborate variations of the rotation method, such as those of the Weissenberg and the precession cameras, involve a motion of the film in addition to the rotation of the crystal.

Powder method. The powder method involves the diffraction of a collimated monochromatic beam from a sample containing an enormous number of tiny crystals having random orientation. Since about 1950 an increasing number of powder patterns studies have been made with Geiger counter, or proportional counter, diffractometers. The apparatus is shown schematically in **Fig. 6**. X-rays diverging from a target at T fall on the sample at O, the sample being a flat-faced briquet of powder. Diffracted radiation from the sample passes through the receiving slit at s and enters the Geiger counter. During the operation the sample turns at angular velocity ω and the counter at 2ω. The distances TO and OS are made equal to satisfy approximate focusing conditions. A filter F before the receiving slit gives the effect of a sufficiently monochromatic beam. A chart recording of the amplified output of the Geiger counter gives directly a plot of intensity versus scattering angle 2θ.

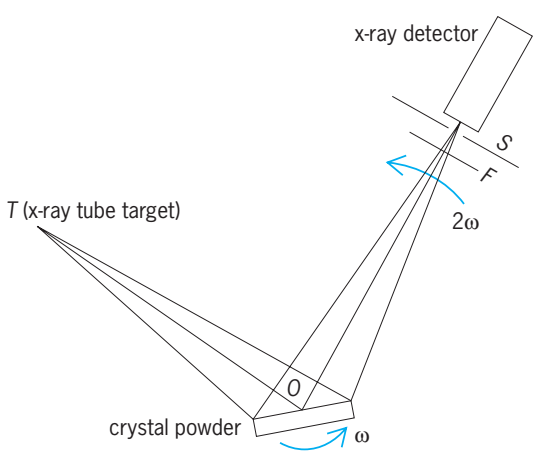

Fig. 6. Schematic representation of the Geiger counter diffractometer for powder samples.

Noncrystalline Diffraction

For a noncrystalline substance such as a glass or a liquid, a more general expression for the intensity of diffracted radiation is required. If the instantaneous position of each atom in the sample is represented by a vector $\mathbf{r}_n$, the diffracted intensity is given by Eq. (22). A particularly useful variation of Eq. (22), given as Eq. (23), is obtained by computing the average intensity $\langle I\rangle$ when the sample as a rigid array is allowed to take with equal probability all orientations in space. In Eq. (23), $k = 4\pi(\sin\theta)/\lambda$ and $r_{qn} = |\mathbf{r}_q - \mathbf{r}_n|$.

$$I = I_e \sum_q \sum_n f_q f_n \times \exp\left[\frac{2\pi i}{\lambda}(\mathbf{s}-\mathbf{s}_0)\cdot(\mathbf{r}_q-\mathbf{r}_n)\right] \tag{22}$$

$$\langle I\rangle = I_e \sum_q \sum_n f_q f_n \frac{\sin(kr_{qn})}{kr_{qn}} \tag{23}$$

The fact that there are fairly definite nearest-neighbor and second-neighbor distances in a glass or liquid means that Eq. (23) will show peaks and dips when the intensity is plotted against $(\sin\theta)/\lambda$. Peaks and dips in an x-ray diffraction pattern merely indicate the existence of preferred interatomic distances, not that the material is necessarily crystalline. X-ray patterns of noncrystalline materials are usually analyzed by a Fourier inversion of Eq. (23), which yields a radial distribution function giving the probability of finding neighboring atoms at any distance from an average atom. Bertram E. Warren

Gases. Gases and liquids are found to give rise to x-ray diffraction patterns characterized by one or more halos or interference rings which are usually somewhat diffuse. These diffraction patterns, which are similar to those for glasses and amorphous solids, are due to interference effects depending both upon the electronic distribution of each of the individual atoms or molecules and upon their relative positions in the system.

For monatomic gases the only appreciable interference effects giving rise to a distribution of scattered intensities are those produced by the electronic distribution about each nucleus. These interference effects giving rise to so-called coherent intensities are the result of the interference of the individual waves scattered by electrons in different parts of the atom. The electronic distribution of an atom is described

in terms of a characteristic atomic scattering factor which is defined as the ratio of the resultant amplitude scattered by an atom to the amplitude that a free electron would scatter under the same conditions. At zero-angle scattering the atomic scattering factor is equal to the atomic number of the atom. The coherent intensity in a given direction is proportional to the square of the atomic scattering factor. If it is assumed that the electronic distribution is spherically symmetrical, the atomic scattering factors can be readily obtained from the observed intensities. For molecular gases the interference effects depend not only on the scattering factors of the atoms but also on their relative positions in the molecule. One can observe only an average intensity scattered over a period of time during which the molecules have taken innumerable positions with respect to the incident beam. Interference effects due to the relative packing of the atoms or molecules can be neglected for dilute gases but not for dense gases.

As in the case of x-ray diffraction by crystals, light atoms such as hydrogen are difficult to detect in the presence of heavy atoms. Because of the shorter exposure times required, electron diffraction rather than x-ray diffraction has been used to study the structures of gaseous molecules. Both methods seem to be comparable in view of the accuracy of the intensity measurements and the technical difficulties involved.

Liquids. One cannot, as in the cases of dilute gases and crystalline solids, derive unambiguous, detailed descriptions of liquid structures from diffraction data. Nevertheless, diffraction studies of liquids do provide most useful information. Instead of comparing the experimental intensity distributions with theoretical distributions computed for various models, the experimental results are usually provided in the form of a radial distribution function which specifies the density of atoms or electrons as a function of the radial distance from any reference atom or electron in the system without any prior assumptions about the structure. From the radial distribution function one can obtain (1) the average interatomic distances most frequently occurring in the structure corresponding to the positions of the first, second, and possibly third nearest neighbors; (2) the distribution of distances; and (3) the average coordination number for each interatomic distance. The interpretation of these diffraction patterns given by the radial distribution function usually is not straightforward, and in general it can be said only that a certain assumed structural model and arrangement is not inconsistent with the observed diffraction data. The models considered represent only a description of the time-average environment about any given atom or molecule within the liquid.

There are great experimental difficulties in obtaining accurate intensity data. The sources of error are many for a detailed treatment the reader is referred to publications of C. Finback. A brief description of some results obtained by x-ray diffraction of liquids is given below.

Liquid elements. The radial distribution function, first used in a study of liquid mercury, has been applied to a considerable number of liquid elements mainly to compare their physical properties in the liquid and crystalline states. In most cases a lower first mean coordination number is found in the liquid state; exceptions are liquid gallium, bismuth, germanium, and lithium. The radial distribution curves give direct evidence for the existence of molecules in some liquid elements (for example, N_2, O_2 Cl_2, and P_4) and imply the existence of more complicated atomic aggregates in a few cases. Argon and helium have been extensively studied in the liquid and vapor states over wide ranges of temperature and pressure.

Liquid water and solutions. A prime example illustrating the considerable structural information made available from modern x-ray liquid diffractometry investigations is the detailed analysis of liquid water, which revealed the following significant features: (1) there are distinct structural deviations of water molecules from a uniform distribution of distances to about 0.8 nm at room temperature; (2) the first prominent maximum, corresponding to near-neighbor interactions, shifts gradually from 0.282 nm at 39.2°F (4°C) to 0.294 nm at 392°F (200°C); (3) the average coordination number in liquid water from 39.2 to 392°F (4 to 200°C) is approximately constant and slightly larger than four; and (4) the radial distribution of oxygen atoms in water at 39.2°F (4°C) is not significantly different from that in deuterium oxide at the same temperature. Comparison of calculated radial distribution functions for various proposed liquid water models (which are sufficiently defined at the molecular level) based on those derived from patterns of liquid water have shown that the only realistic model which gives agreement with data from both large- and small-angle x-ray scattering is related to a modification of the ordinary hexagonal ice structure. This solid-state structure is similar to that of the hexagonal form of silicon dioxide, tridymite, with each oxygen atom tetrahedrally surrounded by neighboring oxygen atoms to give layers of puckered six-membered rings with dodecahedral cavities large enough (radius 0.295 nm) to accommodate a water molecule.

In terms of an average configuration, the liquid water phase may be regarded as a "mixture" model comprising network water molecules forming a slightly expanded ordinary ice structure (each oxygen atom forming nearly four hydrogen bonds with neighboring oxygen atoms) and the cavity water molecules interacting with the network by less specific but by no means negligible forces. It must be emphasized that both kinds of water molecules instantaneously exist in environments which are distorted from the average, as implied by sizable root-mean-square variations in interatomic distance. *See* HYDROGEN BOND; WATER.

Radial distribution curves for concentrated $FeCl_3$ solutions indicate a large degree of local ordering of the ions with formation of $Fe^{3+}-Cl^-$ complexes. Studies on metal-metal solutions, colloidal solutions, and molecular solutions have been made. More definite results have been obtained for concentrated solutions of strongly scattering solutes in weakly scattering solvents. Examples are the proof of the existence of a polymeric species in aqueous $Bi(ClO_4)_3$,

evidence that in aqueous solution the HgX_4^{2-} anions (X = Cl, Br, and I) are tetrahedral, and definite evidence of ion-pair formation in aqueous BaI_2.

Molten salts. A molten salt is considered to be a loose and expanded imitation of the solid with the same coordination scheme and short-range order. Careful x-ray diffraction studies of a number of molten salts have indicated that melts do not possess such quasi-crystalline structures but instead have quite open structures with a wide variety of individual ion coordinations. Interpretations of radial distribution functions for several other molten salts have been made. Liquid $AlCl_3$ appears to consist mainly of Al_2Cl_6 molecules; liquid SnI_4 is composed of independent tetrahedral molecules. The results for other molten salts are not as conclusive. *See* ELECTRON DIFFRACTION; NEUTRON DIFFRACTION. Lawrence F. Dahl

Bibliography. C. S. Barrett and T. B. Massalski, *Structure of Metals*, 3d ed., 1980; M. J. Buerger, *Crystal Structure Analysis*, 1960, reprint 1980; B. D. Cullity, *Elements of X-ray Diffraction*, 3d ed., 2001; D. W. Hukins, *X-ray Diffraction by Disordered and Ordered Systems,* 1981; G. H. Stout and L. H. Jensen, *X-ray Structure Determination*, 2d ed., 1989.

X-ray fluorescence analysis

A nondestructive physical method used for chemical elemental analysis of materials in the solid or liquid state. The specimen is irradiated by photons or charged particles of sufficient energy to cause its elements to emit (fluoresce) their characteristic x-ray line spectra. The detection system allows the determination of the energies of the emitted lines and their intensities. Elements in the specimen are identified by their spectral line energies or wavelengths for qualitative analysis, and the intensities are related to their concentrations for quantitative analysis. Computers are widely used in this field, both for automated data collection and for reducing the x-ray data to weight-percent and atomic-percent chemical composition or area-related mass (of films). *See* FLUORESCENCE.

The materials to be analyzed may be solids, powders, liquids, or thin foils and films. The crystalline state normally has no effect on the analysis, nor has the state of chemical bonding, except for very light elements. All elements above atomic number 12 can be routinely analyzed in a concentration range from 0.1 to 100 wt %. Special techniques are required for the analysis of elements with lower atomic numbers (4–11) or of lower concentrations, and for trace analysis. The counting times required for analysis range from a few seconds to several minutes per element, depending upon specimen characteristics and required accuracy; but they may be much longer for trace analysis and thin films. The results are in good agreement with wet chemical and other methods of analysis. The method is generally nondestructive for most inorganic materials in that a suitably prepared specimen is not altered by the analytical process.

Basis of method. The theory of the method has its origin in the classic work by H. G. J. Moseley, who

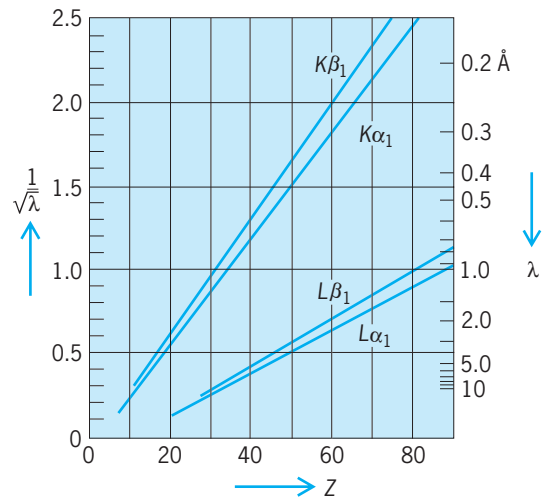

Fig. 1. Plot of Moseley's law, showing dependence of characteristic x-ray-line wavelengths λ on atomic number *Z*. 1 Å = 0.1 nm. (*After Philips Tech. Rev., vol. 17, no. 10, 1956*)

in 1913 measured x-ray wavelengths of a series of elements. He found that each element had a simple x-ray spectrum and characteristic wavelengths, and that there was a linear relationship between $1/\sqrt{\lambda}$ and Z, where λ is the x-ray wavelength and Z is the atomic number of the element emitting the x-ray. For example, a plot of Moseley's law can be used to show the K and L x-ray lines (**Fig. 1**). Aside from the discovery of the element hafnium in zirconium ores by G. von Hevesy, only a few practical uses of the relationship were reported until about 1950, when the introduction of modern x-ray equipment made it feasible to use x-rays for routine spectrochemical analysis of a large variety of materials.

An x-ray source is used to irradiate the specimen, and the emitted x-ray fluorescence radiation is analyzed with a spectrometer. The fluorescence radiation is diffracted by a crystal at different angles in order to separate the wavelengths and identify the elements, and the concentrations are determined from the relative intensities. Scintillation or gas proportional counters are generally used as detectors. This procedure is widely used and is called the wavelength dispersive method.

Around 1965, lithium-drifted silicon and germanium [Si(Li) and Ge(Li)] solid-state detectors became available for x-ray analysis. These detectors have better energy resolution, and the average pulse amplitudes are directly proportional to the energies of the x-ray quanta, which can be sorted electronically with a multichannel pulse-height analyzer. This eliminates the need for the crystal and is called the energy dispersive method. Recent developments include cryogenically cooled detectors based on superconducting tunnel junctions. They combine a far better energy resolution with the ability to detect the emission lines from very light elements.

X-ray spectra. The origin of x-ray spectra may be understood from the simple Bohr model of the atom in which the electrons are arranged in orbits within the K, L, M, . . . shells. If a particle or photon

with sufficient energy is absorbed by the atom, an electron may be ejected from one of the inner shells and is promptly replaced by an electron from one of the outer shells. This results in the emission of a characteristic x-ray spectral line whose energy is equal to the difference of the binding energies of the two orbits involved in the electron transition. The new vacancy is filled by an additional transition from the outer shells, and this is repeated until the outermost vacancy is filled by a free electron. The sum of energies of all photons emitted during the vacancy-refilling cascade is the ionization energy. The energy of the emitted line from the first transition in the cascade has a slightly lower energy than the ionization energy. For example, the ionization energy for the copper K shell is 8.98 keV, and the observed lines have energies of 8.90 keV (Cu$K\beta$) and 8.04 keV (Cu$K\alpha$); the corresponding wavelengths are 0.138, 0.139, and 0.154 nanometer. Altogether the energies of x-ray K-lines extend over three orders of magnitude from 0.111 keV (11.2 nm, Be$K\alpha$) to 114.45 keV (0.0108 nm, U$K\beta_2$).

Optical emission lines result from resonant electron transitions in the outer (valence) shells, producing complex spectra with a large number of lines. By contrast, the x-ray lines arise only from a limited number of transitions between the high-energy levels of the inner shells, so that the x-ray spectrum of an element consists of relatively few lines. They are always initiated by a primary ionization event. Lines are named after the shell where the corresponding electron transition ends ($K, L, M, \ldots$ lines). The most probable transition yielding the highest line intensity in this series is named alpha, followed by beta, gamma, and others, and the indices 1, 2, 3, . . . define a specific transition within the subseries. Depending on the number of energy sublevels in each shell, there are usually only a few important lines in the K spectrum ($K\beta$, $K\alpha_1$, $K\alpha_2$) and a dozen or more lines in the L spectrum. The M lines are rarely used in x-ray analysis.

Auger effect. Occasionally, instead of the emission of the characteristic photon in the course of an electron transition, inner atomic absorption occurs (internal conversion or the Auger effect) when the photon appears to ionize the atom in an additional shell. The existence of an intermediate fluorescent photon is, however, denied by the quantum-mechanical explanation of the Auger effect and should serve only as an aid to illustrate the energy transfer. The ejected Auger electron has a well-defined energy, namely, the energy of the internally absorbed (virtual) photon minus its ionization energy, and can be used for chemical analysis. The probability that no Auger effect occurs, that is, that the photon is actually emitted from the atom and can be used for analysis, is called fluorescence yield. It is thereby the complementary probability to the Auger effect and is higher than 50% for K-shell ionization of elements with atomic numbers above 31 (gallium). For low-atomic-number elements, the Auger effect dominates and the fluorescence is low. This is one of the main reasons for the difficulties in the analysis of very light elements, such as berillium, boron, and carbon, where the fluorescent yield is only 10^{-4} to 10^{-3}. *See* AUGER EFFECT; ELECTRON SPECTROSCOPY.

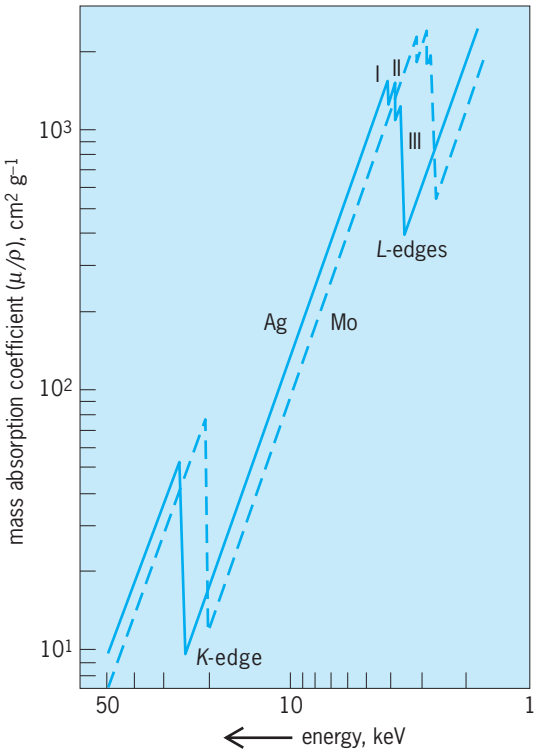

Fig. 2. Mass absorption coefficients of molybdenum (Mo) and silver (Ag) in the 1–50-keV region. Roman numerals indicate edges associated with subshells of the L shell.

X-ray absorption. The type of absorption of the photon or particle leading to the original ionization of the atom is called photoabsorption, to distinguish it from absorption by coherent scattering or Compton scattering. The probability of photoabsorption decreases gradually with increasing photon (or particle) energy, but abruptly increases by an order of magnitude when the photon energy exceeds the ionization energy of a shell. This energy is also called the absorption-edge energy (shown for the K and L edges of molybdenum and silver in **Fig. 2**). Thus the x-rays with energies just higher than the absorption-edge energy are most efficient in generating x-ray fluorescence. The efficiency decreases as the photon energy E is further increased from the edge approximately as $1/E^3$ or λ^3. Photons with smaller energies than the absorption edge have no effect in exciting fluorescence.

The absorption of x-rays is usually given as a mass absorption coefficient μ/ρ (usually expressed in cm^2 g^{-1}) and is independent of the physical state of the material. If more than one element is present, the weighted average of the coefficients of the individual elements is used. Tables of mass absorption coefficients have been compiled. The decrease of intensity of x-rays as they traverse the material is given by the linear absorption coefficient μ (usually expressed in cm^{-1}), obtained by multiplying the mass absorption coefficient by the density ρ of the material. The intensity decreases to $e^{-\mu x}$ of its original

value when the x-rays pass through a layer x centimeters thick.

Radiation sources. There are two general methods for producing x-ray spectra for fluorescence analysis excitation by photons and excitation by charged particles. The most common method is to expose the specimen to the entire spectrum emitted from a standard x-ray tube. It is sometimes modified by using a secondary target material (or monochromator) outside the x-ray tube to excite fluorescence. This has the advantage of selecting the most efficient energy close to the absorption edge of the element to be analyzed and reducing or not exciting other interfering elements, but the intensity is reduced by two or three orders of magnitude. Further alternatives are radioactive sources and synchrotron radiation.

The other method, used in electron microscopes and the electron microprobe, uses an electron beam directly on the specimen, and each element generates its own x-ray spectrum, under electron bombardment, as in an x-ray tube. *See* ELECTRON MICROSCOPE.

X-ray tubes. The radiative spectrum from an x-ray tube consists of continuous radiation (bremsstrahlung) and characteristic lines. Continuous radiation is emitted in the course of scattering (that is, deceleration) of electrons by the nuclei of the target atoms. Characteristic radiation is excited by electrons similarly to excitement by photons, and comes from the electronic shells. The primary x-ray-tube targets are usually tungsten, copper, rhodium, molybdenum, silver, and chromium. It is usually necessary to avoid the use of a tube whose target is identical to that of an element in the specimen, because the line spectrum from the target is scattered through the system, adding to the element signal. It is also desirable to select a target whose characteristic line energies lie closely above the absorption edges of the elements to be analyzed. For example, the WL lines and CuK lines are more efficient in exciting fluorescence in the transition elements chromium to copper than are the MoK lines; RhL lines are most useful to excite K lines of elements below sulfur in the periodic table. Tubes for fluorescence analysis usually have a single thin beryllium window placed at the side of the tube. *See* BREMSSTRAHLUNG.

Equipment is normally operated at x-ray-tube voltages of 20–60 kV in dc operation at up to 3 kW or more with water cooling. These voltages generate the K spectra of all the elements up to the rare earths and the L spectra of the higher-atomic-number elements. Since the detector is moved from point to point, it is essential to have a constant primary intensity and to stabilize the voltage and tube current. *See* X-RAY TUBE.

Radioactive isotopes. Radioactive isotopes that produce x-rays, such as iron-55 (MnK x-rays) and americium-241 (NpL x-rays), are used in place of an x-ray tube to excite fluorescence in some applications. These sources are much weaker than x-ray tubes and must be placed close to the specimen. They are often used in field applications where portability and size may be considerations. Alpha particles have been occasionally used. An example is the excitation source in the α-proton x-ray spectrometer (APXS) built into the Mars exploration vehicle *Sojourner* (Mars *Pathfinder* mission 1997/1998; **Fig. 3**). *See* RADIOACTIVITY.

Synchrotron radiation. Synchrotron radiation has many potential advantages. The continuous radiation is

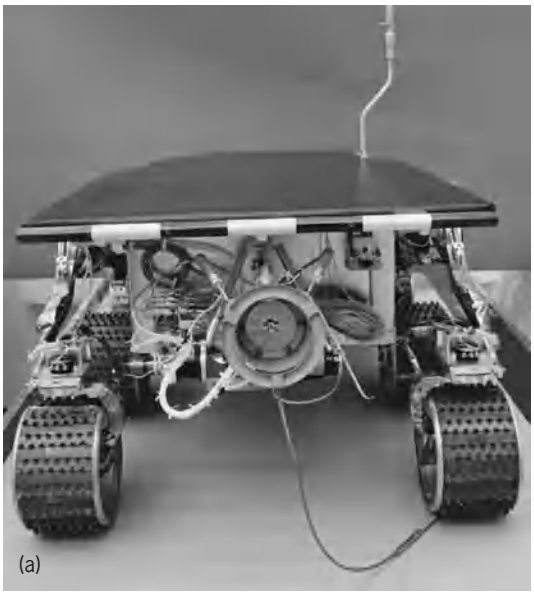

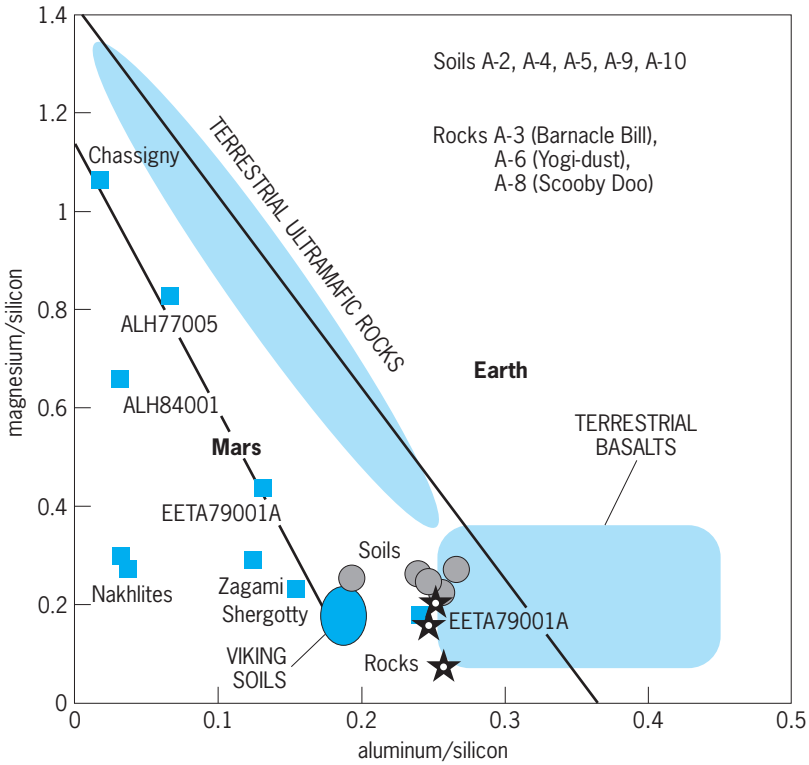

Fig. 3. Alpha-Proton X-ray Spectrometer (APXS) used on Mars *Pathfinder* mission of 1997/1998. (*a*) Mars rover *Sojourner*, rear view showing spectrometer (*copyright © 1997, Jet Propulsion Laboratory, California Institute of Technology, and the National Aeronautics and Space Administration*). (*b*) Comparison of chemical composition of rocks on Earth, of various meteorites found on Earth but presumably originating from Mars, and materials analyzed by APXS near the landing site on Mars.

several orders of magnitude more intense than that of x-ray tubes and can be used with a crystal spectrometer. In addition, a tunable crystal monochromator can be placed in the incident beam to select the optimum wavelength for fluorescing each element in the specimen. Because of its high intensity and parallelism, a very narrow beam of synchrotron radiation can be masked out in order to illuminate individual spots or grains of inhomogeneous materials. Another application is ultra-trace analysis. *See* SYNCHROTRON RADIATION.

Crystal spectrometer. A single-crystal plate is used to separate the various wavelengths emitted by the specimen. Diffraction from the crystal occurs according to Bragg's law, Eq. (1), where n is a small

$$n\lambda = 2d \sin\theta \tag{1}$$

integer giving the order of reflection, λ the wavelength, d the spacing of the particular set of lattice planes of the crystal that are properly oriented to reflect, and θ the angle between those lattice planes and the incident ray. *See* X-RAY CRYSTALLOGRAPHY.

Reflection for a particular λ and d occurs only at an angle 2θ with respect to the incident ray, and it is therefore necessary to maintain the correct angular relationship of the crystal planes at one-half the detector angle. This is done by the goniometer, which is geared to rotate the crystal at one-half the angular speed of the counter tube, and therefore both are always in the correct position to receive the various wavelengths emitted by the specimen (**Fig. 4**). For a given d, there is only one angle (for each order of reflection) at which each wavelength is reflected, the angle increasing with increasing wavelength. The identification of elements by the reflection angles for their emission lines is greatly simplified by modern computer-controlled spectrometers. The angular separation of the lines, or the dispersion, given by Eq. (2), increases with decreasing d. It is thus

$$\frac{d\theta}{d\lambda} = \frac{n}{2d\cos\theta} \tag{2}$$

easy to increase the dispersion simply by selecting a crystal with a smaller d. Reducing d also limits the maximum wavelength that can be measured since $\lambda = 2d$ at $2\theta = 180^\circ$; the maximum 2θ angle that can be reached in practice with the goniometer is about 150°.

Soller slits. The crystals are usually mosaic, and the reflection is spread over a small angular range. To increase the resolution, that is, decrease the line breadth, it is necessary to limit the angular range over which a wavelength is recorded. Parallel or Soller slits are used for this purpose (Fig. 4). These slits consist of thin (0.002-in. or 0.05-mm) equally spaced flat foils of materials such as nickel and iron, and the angular aperture is determined by the length and spacing. A typical set for fine collimation would have 0.005-in. (0.13-mm) spacings and 4-in. (100-mm) length with angular aperture 0.15° and cross section 0.28 in. (7.11 mm) square. Wider angular apertures of up to a few degrees are used with multilayer mirrors for light-element analysis. The absorption of the foils is sufficiently high to prevent rays that are inclined by more than the angular aperture to extend beyond the specimen area and enter the counter tube. Two sets of parallel slits may be used, one set between the specimen and crystal and the other between crystal and detector. This greatly increases the resolution and peak-to-background ratio, and causes a relatively small loss of peak intensity.

Diffracting crystals. Crystals commonly used in spectrometers are lithium fluoride (LiF) with reflecting plane (200) or (220), silicon (111) and (220),

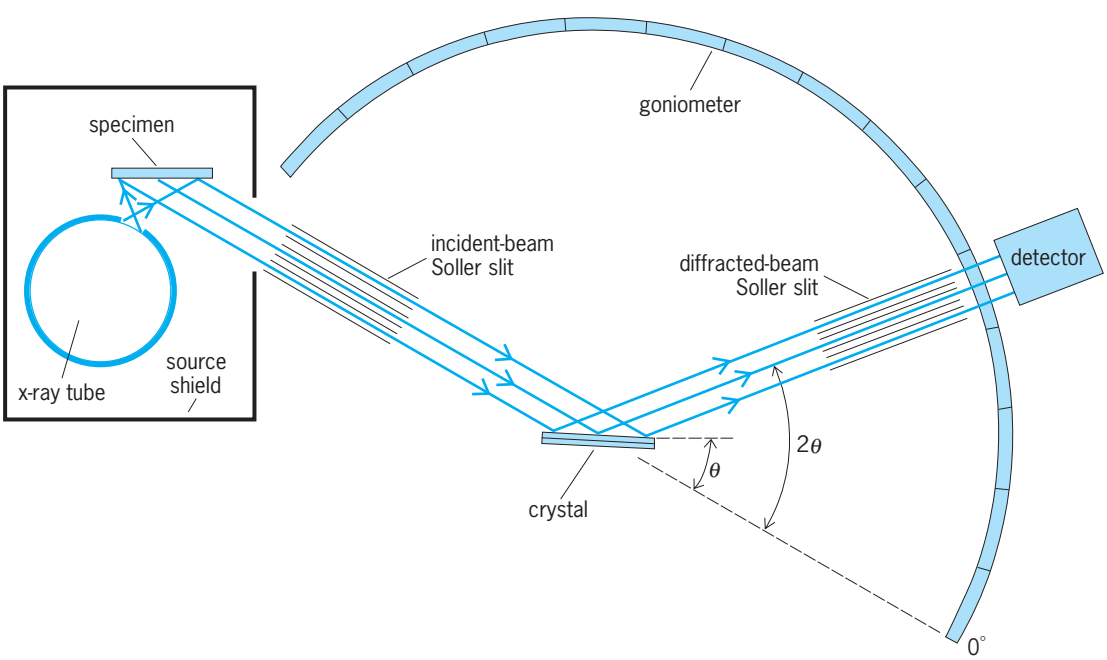

Fig. 4. X-ray fluorescence spectrograph (not to scale). Diffracted-beam Soller slit is optional.

pentaerythritol (001), acid phthalates of potassium and thallium (001), and ethylene diamine *d*-tartrate (020). It is essential that the crystal be of good quality to obtain sharp, symmetrical reflections. Unless the crystal is homogeneous, the reflection may be distorted, and portions of the reflections may occur at slightly different angles. Such effects would decrease the peak intensities of the wavelengths by varying amounts, causing errors in the analysis.

Multilayer mirrors. The longest wavelength that can be routinely analyzed with a natural crystal is around 2.4 nm (O$K\alpha$). Multilayer structures are employed as dispersive devices for lighter elements. They consist of a periodic stack of layer pairs alternating a heavy element (with high scattering power for x-rays) and light elements (serving as a spacer). The scattered partial waves from the heavy-element layers interfere constructively at certain angles in a way similar to that in crystals, but can have much longer wavelengths corresponding to the layer spacing.

Rapid analysis systems. In certain industrial applications such as the manufacture of cement, steels, and glass, and in geological exploration, large numbers of specimens containing up to a dozen or more elements must be rapidly analyzed. In some cases, the analysis must be done in a few minutes to correct the composition of a furnace that is standing by. Generally the same qualitative compositions have to be routinely analyzed, and instead of sequentially scanning over the wavelength regions, a number (up to 30) of fixed crystals and detectors are positioned around the specimen in order to allow simultaneous measurements of several elements at peak and background positions. Automated trays load the specimens into the spectrometer.

Detectors. The detectors generally used in crystal spectrometers are scintillation counters with thin beryllium windows and thallium-activated sodium iodide [NaI(Tl)] crystals for higher energies (above 4 keV), and gas flow counters with very low absorbing windows and argon/methane gas for the low-energy region (below 6 keV). A single-channel pulse-amplitude analyzer limits photon counting to a selected energy interval to improve the peak-to-background ratio and to eliminate higher-order reflections. However, no sharp energy separation is possible due to the rather limited energy resolution of these detectors. *See* GAMMA-RAY DETECTORS; PARTICLE DETECTOR; SCINTILLATION COUNTER.

Energy dispersive systems. Solid-state detectors with good energy resolution are used in conjunction with a multichannel pulse-amplitude analyzer. No crystals are required, and the detector and specimen are stationary during the measurement. The method is used with either electron-beam excitation in electron microscopes or with x-ray-tube sources. The photons of various energies are registered, and their energies are determined as soon as they enter the detector. As this occurs statistically for the various fluorescence line energies, the acquisition of the spectral data appears to be simultaneous for all lines.

Solid-state detectors. Lithium-drifted silicon [Si(Li)] detectors are generally used for the lower energies of fluorescence analysis, while lithium-drifted germanium [Ge(Li)] detectors are more often used for nuclear high-energy gamma-ray detection. The energy resolution of good Si(Li) detectors is below 130 eV (full width at one-half maximum) for Mn$K\alpha$ radiation. The lithium-drifted detectors require cooling during operation, for which liquid nitrogen is often used. *See* JUNCTION DETECTOR.

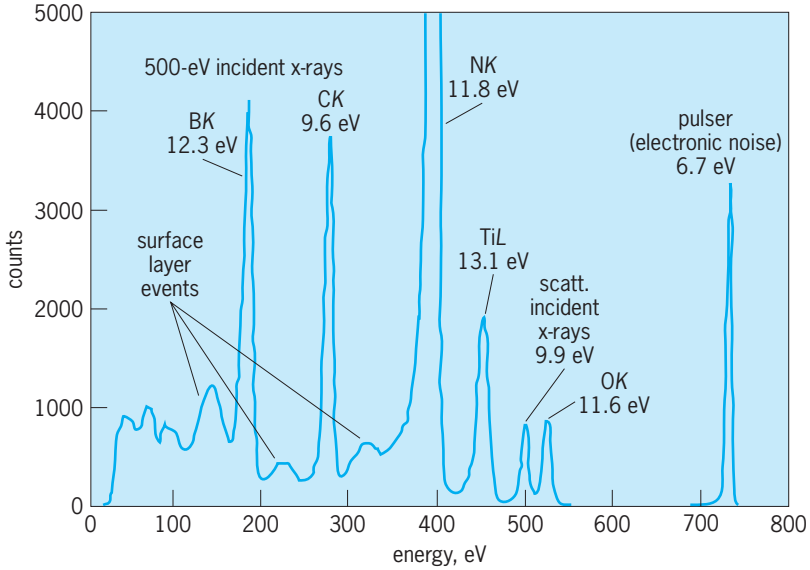

Fig. 5. Spectrum of boron nitride partially covered with titanium powder obtained with a cryogenically cooled superconducting tunnel junction detector. The energy resolution of all lines up to several hundred electronvolts is around 10–12 eV. A crystal spectrometer with a multilayer mirror would have a resolution of about 16 eV at BKα. (*After M. Frank et al., Cryogenic high-resolution x-ray spectrometers for SR-XRF and microanalysis, J. Synchrotron Rad., 5:515–517, 1998*)

The resolution of the detector is closely linked to its temperature. Some types allow operation at room temperature with degraded resolution, or with Peltier cooling stages. The most recent development are superconducting tunneling junction devices, which are operated at liquid helium temperature. Their energy resolution is comparable to wavelength dispersive spectrometers or is even much better, particularly for light elements (**Fig. 5**).

Analyzer. The output signals from the detector are fed into the analyzer, where the photon counts are stored in memory locations (1024–8192 channels are generally used) that are related to the energies of these photons. This also allows visual observation on a cathode-ray-tube screen of the accumulated spectrum and of the simultaneous counting process. Analyzers are usually provided with cursor markers to easily identify the peaks in the spectrum. Computer memories can be used for storage of the spectral counts, thus providing efficient access to computer routines for further data evaluation.

Use. Energy dispersive x-ray spectrometers are useful to accumulate spectra in short time intervals (for example, 1 min) that often allow a preliminary interpretation of the qualitative and quantitative composition of the specimen. The instruments are comparatively small, because they are designed to accept a large aperture of radiation. They require only low-power x-ray tubes that sometimes can be air-cooled.

Limitations. An important limitation of energy dispersive systems with Si(Li) detectors is the energy resolution, which is about an order of magnitude poorer in the lower energy region than that of crystal spectrometers. For example, the $K\alpha$ lines of the transition elements overlap with the $K\beta$ lines of the element preceding it in atomic number, causing severe analytical difficulties in an important region of the spectrum. The peak-to-background ratio is significantly lower than in crystal spectrometers because of the lower resolution. Another limitation is that the maximum number of photons that can be processed by the electronic circuits is limited to about 15,000–50,000 counts per second. This is the total photon count from the entire detected spectral region. Trace elements with low count rates in a matrix of high-count elements are therefore difficult to detect with sufficient statistical accuracy. Various attempts have been made to overcome this drawback by selectively exciting the elements of interest by using selective filters or secondary targets, which also greatly reduces the amount of x-ray-tube radiation that is scattered into the detector.

Microanalysis. The electron microprobe is widely used for elemental analysis of small areas. An electron beam of 1 micrometer (or smaller) is used, and the x-ray spectrum is analyzed with a focusing (curved) crystal spectrometer or with an energy dispersive solid-state detector. Usually two or three spectrometers are used to cover different spectral regions. Light elements down to beryllium, boron, and carbon can be detected. An important use of the method is in point-to-point analysis with a few cubic micrometers of spatial resolution. *X-Y* plots of any element can be made by moving the specimen to determine the elemental distribution.

Figure 6 illustrates the spectra obtained with three of the most frequently used methods of analysis. The specimen, a high-temperature alloy of the type used in aerospace and other industries, was prepared by the National Institute of Standards and Technology with stated composition in weight percent: molybdenum (Mo) 3.13, niobium (Nb) 4.98, nickel (Ni) 51.5, cobalt (Co) 0.76, iron (Fe) 19.8, chromium (Cr) 17.4, titanium (Ti) 0.85, and aluminum (Al) 0.085, total 99.27%.

Figure 6*a* shows the high-resolution spectrum obtained in about an hour with a lithium fluoride (LiF; 200) crystal spectrometer using 50-kV, 12-milliampere x-ray-tube excitation and scintillation counter. This spectrum also contains the second-order (II) and third-order (III) crystal reflections of molybdenum and niobium whose $K\beta_1$ and $K\beta_3$ components are resolved. The lower resolution of the energy dispersive method is shown in Fig. 6*b*, recorded in about 10 min using 50-kV, 2-microampere x-ray-tube excitation, Si(Li) detector, and 40 eV per channel (about 400 channels are shown). The spectral range includes the unresolved molybdenum and niobium *L* lines and titanium. Figure 6*c* is an energy dispersive spectrum excited by a 25-keV electron beam. The molybdenum and niobium spectra are weakly excited at this low voltage and are not visible on the scale used in the plot. The differences in the relative intensities of the lines in the spectra arise from differences in the conditions of excitation and detection, and they illustrate the necessity of using the proper correction factors for each method of analysis to derive the correct weight percent composition.

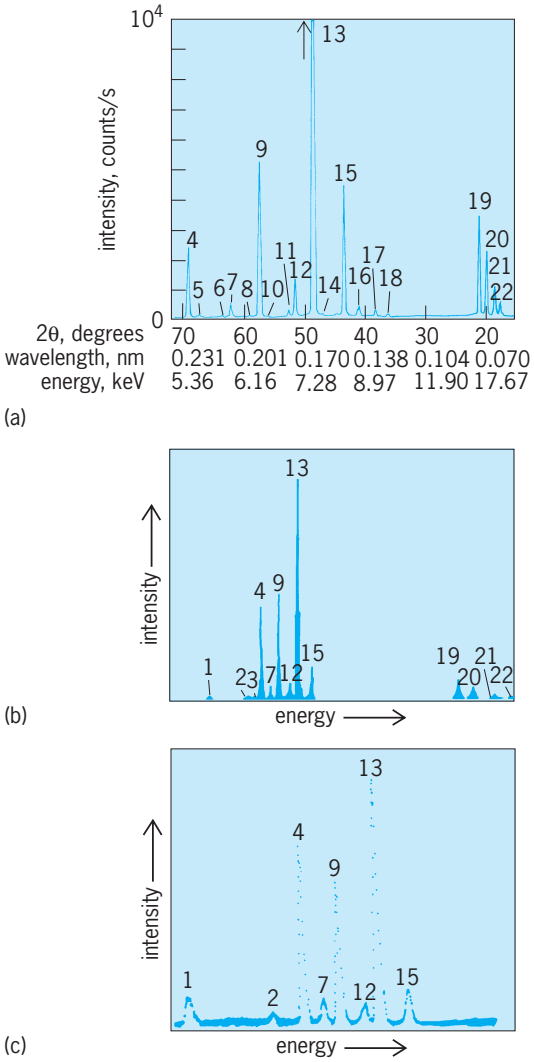

Fig. 6. Fluorescence spectra of high-temperature alloy obtained with (*a*) crystal spectrometer, (*b*) energy dispersive method with x-ray-tube excitation, and (*c*) energy dispersive method with electron-beam excitation. Spectral lines: 1, Mo + Nb$L\alpha$ + $L\beta$. 2, Ti$K\alpha$. 3, Ti$K\beta$. 4, Cr$K\alpha$. 5, Nb$K\alpha_{1,2}$III. 6, Mo$K\alpha_{1,2}$III. 7, Cr$K\beta$. 8, Nb$K\beta_{1,3}$III. 9, Fe$K\alpha$. 10, Mo$K\beta$III. 11, Co$K\alpha$. 12, Fe$K\beta$. 13, Ni$K\alpha$. 14, Co$K\beta$. 15, Ni$K\beta$. 16, Mo$K\alpha_{1,2}$II. 17, Nb$K\beta_{1,3}$II. 18, Mo$K\beta_{1,3}$. 19, Nb$K\alpha$. 20, Mo$K\alpha$. 21, Nb$K\beta_{1,3}$. 22, Mo$K\beta_{1,3}$.

Specimen preparation. The specimens may be in the form of powders, briquettes, solids, thin films, or liquids. The surface exposed to the primary x-ray beam must be flat, smooth, and representative of the sample as a whole, because usually only a thin surface layer contributes to the fluorescent beam in a highly absorbing specimen. The thickness of this layer is called information depth and may be only a micrometer or less for electron-beam excitation and 10–100 μm or more for x-rays. The degree of surface roughness, which is difficult to measure quantitatively, causes losses in intensity and results in

errors in the analysis. Consequently, solid samples are generally polished; and then, if necessary, they are lightly etched or specially cleaned to remove contaminants. This is particularly important when light elements are measured. Special care must be taken when a measured element is a constituent of such surface contamination.

Powders. Powders are processed in one of two ways. The first is to press the ground material into briquettes. The pressure should be several tons per square centimeter (1 ton/cm^2 equals approximately 15,000 lb/in.^{2} or 100 megapascals), and in most cases organic binders have to be used to improve the mechanical stability. The second way is to use fusion techniques, where the powders (mostly mineralogical or metal oxides) are dissolved at high temperatures in borax or similar chemicals, and glassy pellets are obtained after cooling. The advantage of the second method is a high homogeneity of the specimen and a reduction of interelement effects; but the intensities are reduced.

Liquids. Liquids can be analyzed by using small containers with a thin window cover. Examples are sulfur determination in oils during the refining process, lubrication oil additives, the composition of slurries, and the determination of lead, zinc, and other elements in ore processing. Low concentrations of elements in solution can be concentrated with specific ion-exchange resins and collected on filter papers for analysis. Gases containing solid particles can be filtered and the composition of the particles determined as for atmospheric aerosol filters for environmental studies. In certain industrial applications, liquids are continuously analyzed while flowing through a pipe system with a thin window in the x-ray apparatus.

Quantitative analysis. The observed fluorescent intensities must be corrected by various factors to determine the concentrations. These include the spectral distribution of the exciting radiation, absorption, fluorescence yield, and others. Two general methods have been developed to make these corrections: the fundamental parameter method and the empirical parameter method.

Fundamental parameter method. In the fundamental parameter method, a physical model of the excitation is developed and described mathematically. The method derives its name from the fact that the physical constants, like absorption coefficients and atomic transition probabilities, are also called fundamental parameters. Primary and secondary excitation are taken into account; the first is the amount of fluorescent radiation directly excited by the x-ray tube. Secondary excitation is caused by other elements in the same specimen, whose fluorescent radiation has sufficient energy to excite the characteristic radiation of the analyzed element. In practical applications, the count rate must be calibrated for each element by comparing it to the count rate from a standard of accurately predetermined composition. A standard may contain several elements or can be a pure element.

The fundamental parameter method is capable of accuracies around 1% (absolute weight percentage) for higher concentrations, and between 2 and 10% (relative) for low concentrations. The method has the advantage of allowing the use of pure-element standards. Significantly higher accuracies can be obtained with standard specimens of similar composition to the unknown.

The fundamental parameter method can also be used to determine thickness and chemical composition of thin films.

Empirical parameter method. The empirical parameter method is based upon simple mathematical approximation functions, whose coefficients (empirical parameters) are determined from the count rates and concentrations of standards. A widely used set of approximation functions is given by Eq. (3), where c_i

$$\frac{C_i}{r_i} = \frac{1 + \sum_{j \neq i}^{n} \alpha_{ij} C_j}{R_i} \qquad i = 1, \ldots, n \qquad (3)$$

is the concentration of the analyzed element i in the unknown specimen, r is the corresponding count rate, R_i is the count rate from a pure-element specimen i, C_j are the concentrations of the other elements in the unknown specimen, n is the number of elements, and α_i are the empirical parameters (also called alpha coefficients).

A minimum of $n - 1$ standard specimens, each of which contains the full set of n elements (or a correspondingly higher number, if they contain fewer elements), is required to calculate the empirical parameters, α_{ij}, before actual analysis of an unknown is possible. In practical applications, however, at least twice as many standards should be used to obtain good accuracy, thus requiring considerable effort in standard preparations. The empirical parameter method is therefore mainly used in routine applications, where large numbers of similar specimens must be analyzed. The accuracy of the method depends upon the concentration range covered by the standards; around $\pm 0.1\%$ or better can be obtained if a set of well-analyzed standards with similar compositions to the unknowns are used. If pure-element standards are not available, the pure-element counts rates, R_i in Eq. (3), can also be determined by computation from additional multielement standards.

Trace analysis. There are two distinct analytical tasks that are called trace analysis: the detection or quantification of small amounts of a material (possibly a pure element), and the determination of very low concentrations in an abundantly available sample. In both cases, the relationship between concentration and count rates is practically linear. The minimum detection limit is defined by that amount or concentration for which the peak is just statistically significant above background level B, usually $3B^{1/2}$. The background arising from scattered continuous radiation from the x-ray tube is a limiting factor in determining the peak-to-background ratio. Since intensity measurements can theoretically be made arbitrarily accurate by using long counting times, the

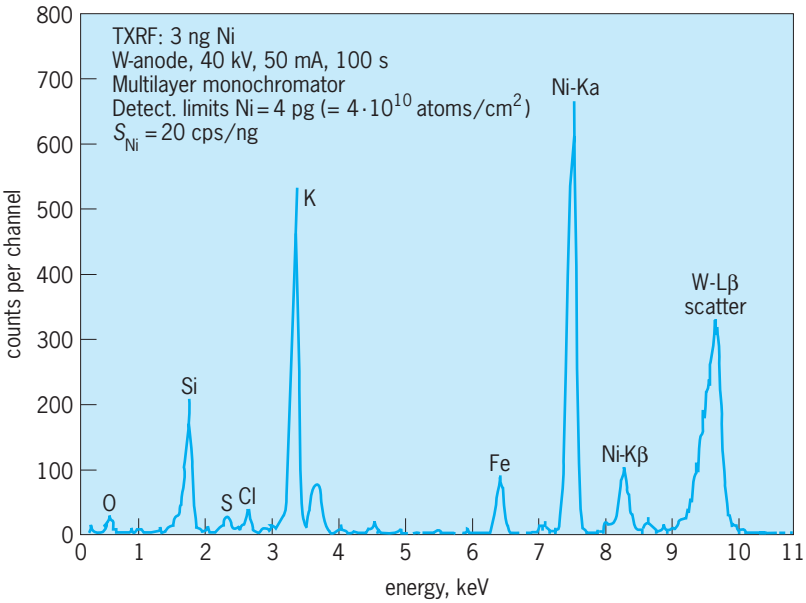

Fig. 7. Example of trace analysis by total reflection x-ray fluorescent analysis (TXRF). A droplet containing 3 ng dissolved nickel (Ni) was applied to a substrate (silicon-wafer), dried, and measured. The sensitivity *S* in this particular setup was 20 counts per second and per nanogram Ni, and the theoretical detection limit for 100 s counting time was 4 picograms, corresponding to 4×10^{10} atoms/cm^2. The elements sulfur (S), potassium (K), and iron (Fe) are contaminants of the solvent, and the silicon (Si) and oxygen lines originate mainly from a thin silicon dioxide (SiO_2) layer on top of the wafer. (*Data provided by P. Wobrauschek, Atominstitut der Österreichischen Universitäten, Vienna*).

minimum detection limits could be indefinitely low. However, in practice, the limiting factors are the background level and long-term instrument drift. Depending upon excitation conditions, matrix, and counting times, traces in the parts-per-million region may be detected with conventional instruments, and in the parts per trillion region by total reflection x-ray fluorescence.

Total reflection XRF (TRXFA). Ultra-trace analysis by x-ray fluorescence is possible by a special technique and instrumentation which is based upon background suppression by total reflection of the primary x-ray beam. The physical explanation is that the index of refraction of x-rays is very slightly smaller than 1, and a beam impinging at a flat surface at angles of a few tenths of a degree is totally reflected without noticeably penetrating the material. In practice, a substrate of a silicon-single crystal (such as a wafer) is used and a small droplet of dissolved analyte material applied and dried. The x-ray beam penetrates only the sample material, not the substrate. A Si(Li) energy dispersive detector is placed at close distance to the specimen. With conventional x-ray tubes, detection limits in the picogram range have been reported, and in the femtogram range by using synchrotron radiation. Total reflection x-ray fluorescence analysis instrumentation is commercially available (**Fig. 7**).

Thin-film analysis. As a rule of thumb, materials with a thickness exceeding a few hundred micrometers can be considered "infinitely thick" from the viewpoint of x-ray fluorescence. This limit decreases by a factor 5–20 for light elements. The intensities of thinner specimens are correspondingly lower, depending upon element, matrix, and experimental setup. In the analysis of very thin films (a few tens of nanometers), the count rates are a linear function of element concentration and of film thickness. Absorption and interelement effects must be taken into account in the analysis of thicker films and foils. This can be done with special fundamental parameter methods, but it requires adequate computing power for efficient evaluation of data.

Fundamental parameter methods allow the determination of thickness and element concentrations of thin films as well as individual layers in multilayer structures. Limitations apply to common elements of two or more layers and with respect to very light elements.

Limitations on accuracy. In both the fundamental parameter and empirical parameter methods, limitations of the accuracy are due mainly to uncertainties in the composition of the standards and variations in the specimen preparation; intensity fluctuations due to counting statistics and instrument instabilities may also contribute.

Supplemental methods. As in all analytical methods, it is sometimes necessary to supplement the chemical data from fluorescence analysis with data by other methods to properly characterize the material. The first three elements in the periodic table (hydrogen, helium, lithium) cannot be measured by x-ray fluorescence, because none of their emission lines are in the x-ray regime. The light elements beryllium through magnesium (including such important elements as carbon, oxygen, and nitrogen) can be measured, but frequently with difficulties. Often they are crucial in the characterization of a specimen, such as carbon in steels, and oxygen in rocks and oxide samples, which may require optical emission, atomic absorption, Auger and electron spectroscopy, or other analytical methods. *See* ANALYTICAL CHEMISTRY; ATOMIC SPECTROMETRY; SURFACE PHYSICS.

An important supplementary method is x-ray polycrystalline diffraction, in which the crystalline chemical phases are identified by comparing the pattern of the unknown with standard patterns. Computer methods are widely used to search the 40,000 phases currently contained in the Powder Diffraction File published by the International Center for Diffraction Data, Newtown Square, Pennsylvania. Mixtures of phases can be quantitatively determined, and there are no limitations on the chemistry of the substances. By combining the chemical data from fluorescence with the phase data from diffraction, the relation between the constituents of the sample and its properties can be established. *See* X-RAY DIFFRACTION.

Applications. X-ray fluorescence analysis is widely used for compositional control in large-scale industrial processing of metals and alloys, cements, the petroleum industry, and inorganic chemicals. Among the many other major applications are geological exploration and mineralogical analysis, soils and plants, glasses, corrosion products, the analysis of raw materials, and the measurement of plating coating thickness. It is an important method in materials characterization for research and technology,

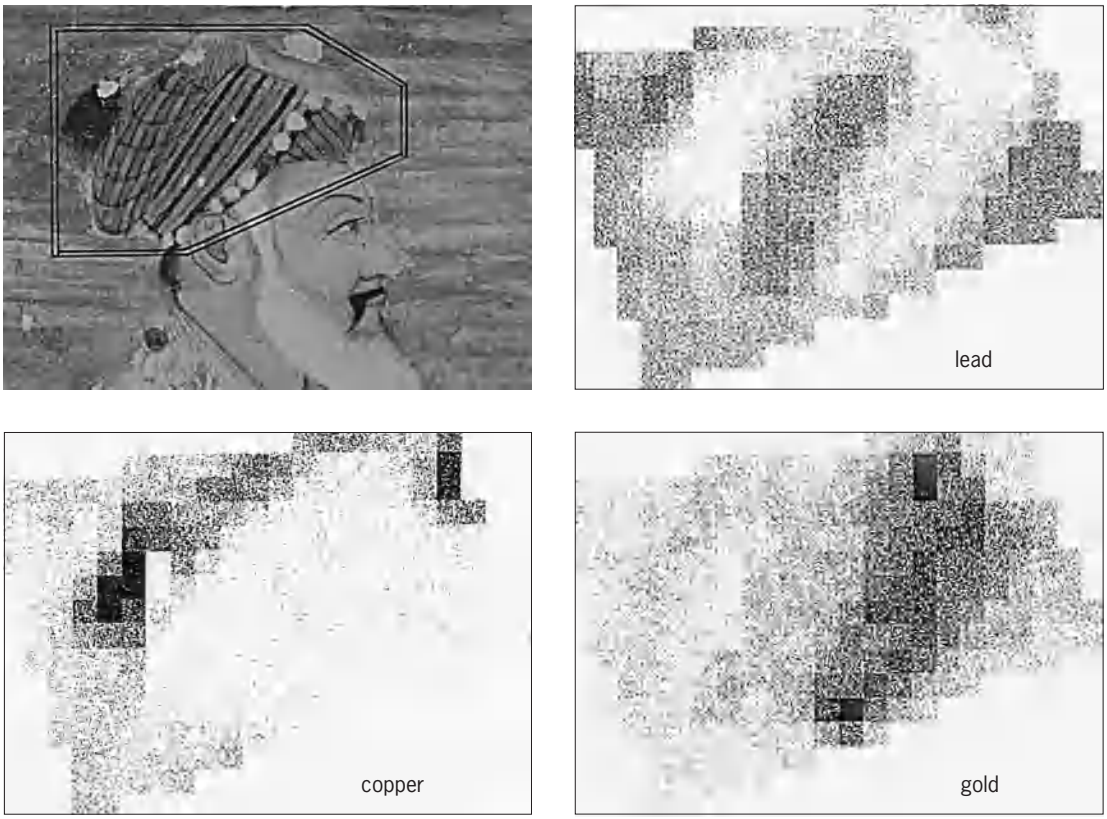

Fig. 8. Analysis of pigments in an Indian miniature, Mughal period, seventeenth century, Schloss Schönbrunn, Vienna. In the detail, the headdress (marked area) is approximately 2.1 × 1.6 in. (50 × 40 mm) and was measured in pixel steps of 0.1 in. (2.5 mm). The distribution of the elements copper, lead, and gold is shown, indicating usage of lead-white, minimum (red), azurite (blue), malachite (green), and metallic gold. (*Analysis by M. Schreiner, Akademie der Bildenden Künste, Vienna. Copyright, Österreichisches Bundesdenkmalamt, Vienna*)

providing chemical information without destroying the sample. It is the only feasible method for many complex analyses that would require extremely long times by conventional wet chemical methods on materials such as the refractory metals, high-speed cutting steels, and complex alloys.

Besides the large-scale industrial applications, the method has been used in a variety of analyses in the medical field, for environment protection and pollution control, and for many research applications. Examples are trace analysis of heavy metals in blood; analysis of airborne particles, historic coins, potteries, lead and barium in Roman skeletons, and various elements in archeological specimens; analysis of pigments to establish authenticity of a painting (**Fig. 8**); quality control of noble metals in alumina-based exhaust catalysts for cars; and analysis of ash and sulfur in coals, slags from furnace products, and surface deposits on bulk metals. The method is also widely used in forensic problems, where it is often combined with x-ray powder diffraction. Remote analysis of rocks using x-ray spectrometers carried by spacecraft and stellar landers has proven to be a valuable source of information in search of the origin of the solar system and its planets. William Parrish; Michael Mantler

Bibliography. *Advances in X-ray Analysis*, annually; E. P. Bertin, *Introduction to X-ray Spectrometric Analysis*, 1978; K. F. J. Heinrich, *Electron Beam X-ray Microanalysis*, 1981; K. F. J. Heinrich et al. (eds.), Energy Dispersive X-ray Spectrometry, *NBS Spec. Publ.*, no. 604, 1981; R. Jenkins, *X-ray Fluorescence Spectrometry*, 1988; R. Jenkins, R. W. Gould, and D. Gedke, *Quantitative X-ray Spectrometry*, 2d ed., 1995; G. R. Lachance, F. Claisse, and H. Chassin, *Quantitative X-ray Fluorescence Analysis: Theory and Application*, 1995; K. L. Williams, *Introduction to X-ray Spectrometry*, 1987.

X-ray microscope

A term used to describe a technique and an instrument or combination of instruments which utilize x-radiation for chemical analysis and for magnification of 100–1000 diameters. The resolution possible is about 0.25 micrometer. X-ray microscopy is a relatively recent development among the microscopic techniques. The contrast in the x-ray microscopic image is caused by varying x-ray attenuation in the specimen. The advantage of x-ray microscopy is that it yields quantitative chemical information, besides structural information, about objects, including those which are opaque to light. It is a reliable ultramicrochemical analytical technique by which amounts of elements and weights of samples as small as 10^{-12} to 10^{-14} g can be analyzed with an error of only a few percent. *See* X-RAY OPTICS.

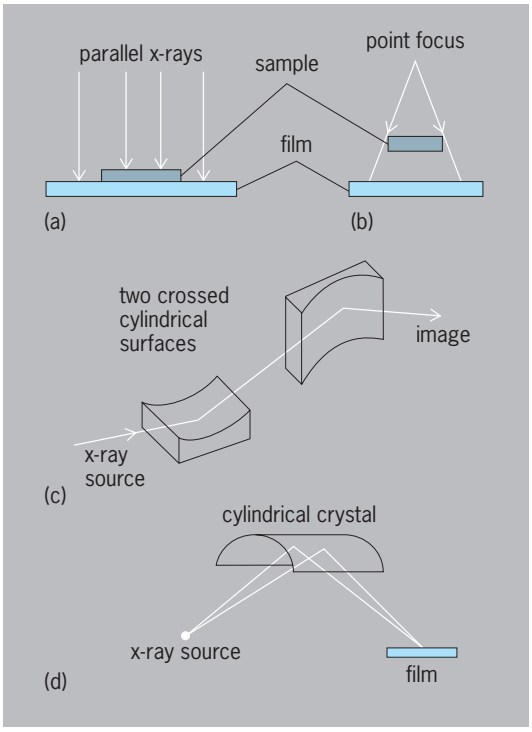

Principles for x-ray microscopy. (*a*) Contact microradiography. (*b*) Projection x-ray microscopy. (*c*) Reflection x-ray microscopy. (*d*) X-ray image spectrography.

Principles of image formation. There are four general principles of x-ray microscopy: (1) contact microradiography (**illus.** *a*), (2) projection x-ray microscopy (illus. *b*), (3) reflection x-ray microscopy (illus. *c*), and (4) x-ray image spectrography (illus. *d*). For a discussion of the first two *see* MICRORADIOGRAPHY

Contact microradiography. In contact microradiography the thin specimen is placed in close contact with an extremely fine-grained photographic emulsion which has a resolution of more than 1000 lines/mm, and radiographed with x-rays of suitable wavelength. Thus an absorption image in scale 1:1 is obtained, and this image is subsequently viewed in a light microscope. The maximal resolution is that of the optical microscope (0.25 μm), but the image has more information, which can be obtained by examining the microradiogram in the electron microscope. The optical microscope gives more information because it is possible to correlate the intensity of light in the image with the primary attenuation in the specimen. This is not at present possible in the electron microscope. The electron microscope gives only high magnification, but it is not possible to draw conclusions about the quantitative attenuation of electrons in the specimen in order to draw conclusions about the specimen composition. *See* ELECTRON MICROSCOPE.

Projection x-ray microscopy. Projection x-ray microscopy, or x-ray shadow microscopy, is based on the possibility of producing an extremely fine x-ray focal spot. This is achieved by an electronic lens system similar to that in the electron microscope. The fine focal spot is produced on a very thin metal foil which serves as a transmission target. The x-rays are generated on the target by the impact of the electrons. The sample is placed near the target, and the primary magnification depends on the ratios of the distances from focal spot to sample and sample to film. Resolution is of the same order as the size of the focal spot; the best value is about 0.1 μm in favorable objects.

Reflection x-ray microscopy. The method of reflection x-ray microscopy is based on the fact that the refractive index for x-rays in solids is a very small amount less than 1. Thus at grazing incidence (that is, incidence at very small angles), the x-rays are totally reflected, and if the reflecting surface is made cylindrical, there will be a focusing action in one dimension. By crossing two such surfaces a true image formation can be obtained, although with some astigmatism, which can be corrected by giving the surfaces a complicated optical shape. The resolution by this procedure is about 0.5–1 μm.

X-ray image spectrography. X-ray image spectrography utilizes Bragg reflections in a cylindrically bent crystal and produces slightly enlarged emission images; this technique is best classified as a micromodification of x-ray fluorescence analysis. The resolution is about 50 μm. *See* X-RAY DIFFRACTION; X-RAY FLUORESCENCE ANALYSIS.

Chemistry of specimen. The penetration of x-rays through material varies with the amount and composition of the attenuator (the material) and the wavelength of the x-rays. The specimen must be thin in order to obtain high resolution, and therefore soft (low voltage) and ultrasoft x-rays must be used. For biological specimens, x-rays of energies between 0.2 and 5 kV (about 5.0–0.2 nanometers in wavelength) are generally used. Thin metallurgical specimens can be examined with harder x-rays because of their higher x-ray absorption. In general the x-ray absorption increases with increasing wavelength of the x-rays (softer rays) and increasing atomic number of the elements composing the specimen. Thus a microscopic structure containing elements of high atomic number of sufficient concentration embedded in a matrix of elements with low atomic numbers will show up as a heavily x-ray-absorbing structure. At certain wavelengths in the x-ray-absorption spectrum, discontinuities appear: the so-called x-ray-absorption edges. The position of these edges in the spectrum is unique for each element; thus they can be utilized for identification of elements in microscopic structures in a specimen. *See* X-RAYS.

Quantitative microscopy. By measuring the variation of density in the x-ray microscopic image the x-ray attenuation can be calculated. If the x-ray microscopic image is recorded with x-rays of suitable wavelength (often monochromatic x-rays), certain chemical characteristics of the specimen can be quantitatively assessed. If a certain element is to be determined in a microstructure, two x-ray microscopic images are recorded with monochromatic x-rays with wavelengths

on either side and close to the absorption edge for the particular element. In this way elementary analysis can be performed on specimens weighing not more than 10^{-10} to 10^{-12} g with a relatively high degree of accuracy. Instead of using a photographic film to record the x-ray image, various types of detectors (Geiger-Müller tubes) are used to measure the variations of x-ray transmission in the sample. Such techniques, especially in the form of scanning, may become more useful in the future.

By proper selection of x-ray wavelength the dry weight, water content, and the content of certain other compounds can be determined in cellular structures down to about 1 μm in size. Thus weights as small as 10^{-14} g can be determined with an analytical error of only a few percent.

Applications. In biology, x-ray microscopy has been utilized for the quantitative determination of the dry weight, water content, and elementary composition of many tissues, for example, nerve cells, various types of secretory cells, the individual bands in muscle fibers, chromosomes, and parts thereof. Especially in studying mineralized tissues much new information has been gained through the use of x-ray microscopy.

The capillary circulation in the living animal can be studied by x-ray microscopy. Thorotrast and Umbradrie, contrast media, may be injected into the bloodstream and successive microradiograms recorded. The procedure is called microangiography and has been applied to the study of the finest blood vessels.

As indicated previously, soft and ultrasoft x-rays must be used for the x-ray microscopy of thin sections or smears of biologic tissue. These long-wavelength x-rays must be generated in specially designed x-ray tubes. The sample and photographic emulsion are enclosed within the high vacuum of the x-ray tube. The pictures obtained show the distribution of dry weight (mass) within the cells and tissues in addition to information on the structures in the specimen.

It is possible to determine the thickness of thick samples by using oblique incidence of the x-rays. With this technique, the thickness of nerve fibers and of some constituents of bone tissue has been determined.

By tilting the film and specimen at a certain angle, stereoscopic microradiograms have been obtained. When thick sections of tissue have been used, the three-dimensional arrangement of bone cells in bone tissue and the three-dimensional image of the capillary net in the circulatory system have been made available for study.

X-ray microscopy has been applied to biology, medical research, mineralogy, and metallurgy among other fields. Arne Engstrom

Bibliography. A. J. Morgan, *X-ray Microanalysis in Electron Microscopy for Biologists,* 1985; D. Sayre et al. (eds.), *X-ray Microscopy Two*, 1988; K. Shinohara and K. Yada (eds.), *X-ray Microscopy in Biology and Medicine*, 1990.

X-ray optics

By analogy with the science of optics, those aspects of x-ray physics in which x-rays exhibit properties similar to those of light waves. X-ray optics may also be defined as the science of manipulating x-rays with instruments analogous to those used in visible-light optics. These instruments employ optical elements such as mirrors to focus and deflect x-rays, zone plates to form images, and diffraction gratings to analyze x-rays into their spectral components. X-ray optics is important in many fields, including x-ray astronomy, biology, medical research, thermonuclear fusion, and x-ray microlithography. It is essential to the construction of instruments that manipulate and analyze x-rays from synchrotrons and particle storage rings for synchrotron radiation research. *See* GEOMETRICAL OPTICS; OPTICS; PHYSICAL OPTICS; X-RAYS.

X-ray refraction and absorption. When W. C. Roentgen discovered x-rays in 1895, he unsuccessfully attempted to reflect, refract, and focus them with mirrors, prisms, and lenses of various materials. The reason for his lack of success became evident after it was established that x-rays are electromagnetic waves of very short wavelength for which the refractive index of all materials is smaller than unity by a only a small decrement. In addition, x-rays are absorbed by materials. The refractive index can be written as a complex quantity, as in Eq. (1),

$$\tilde{n} = 1 - \delta - i\beta \tag{1}$$

where $1 - \delta$ represents the real part, n, of the refractive index and β is the absorption index. These quantities are strongly dependent on the wavelength of the x-rays (see **table**) and the material. X-rays of wavelength about 0.1 nanometer or less are called hard x-rays and are relatively penetrating, while x-rays of wavelength 1–10 nm are less penetrating and are called soft x-rays. Radiation in the wavelength range 10–50 nm, called the extreme-ultraviolet (EUV) region, is very strongly absorbed by most materials. Values of δ remain very small throughout the x-ray and extreme-ultraviolet regions with the consequence that radiation is very weakly refracted by any material. Thus lenses for x-rays would have to be very strongly curved and very thick to achieve an appreciable focusing effect. However, because the absorption index, β, is so high in comparison, such thick lenses would absorb most of the incident radiation, making such lenses impractical. *See* ABSORPTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES; ULTRAVIOLET RADIATION.

Dependence of the complex refractive index of copper on x-ray wavelength

	Refractive index parameter	
Wavelength, nm	δ	β
0.1	1.1×10^{-5}	8.7×10^{-7}
2.0	2.8×10^{-3}	6.0×10^{-3}
15.0	4.2×10^{-2}	6.9×10^{-2}

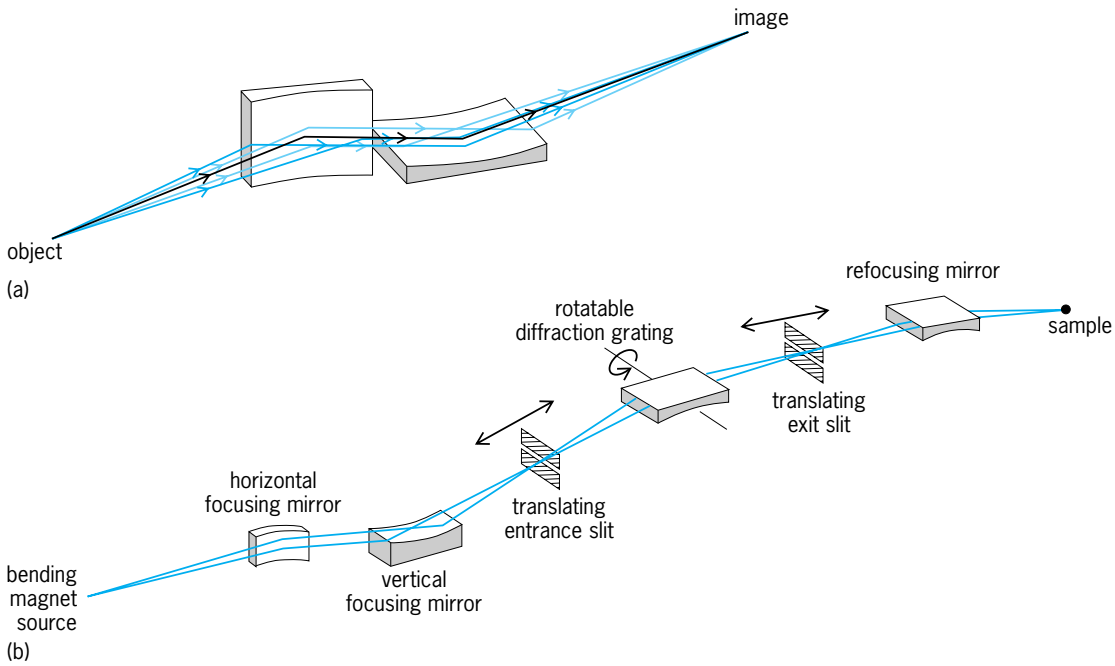

Fig. 1. Microscope principle that corrects the astigmatism associated with glancing-incidence spherical mirrors. (*a*) Refocusing of x-rays to a point source by using a combination of two spherical mirrors (*after J. H. Underwood and D. T. Attwood, The renaissance of x-ray optics, Phys. Today, 37(4):44–50, April 1984*). (*b*) Use of this principle in a synchrotron radiation beam line to collect, analyze, and refocus the x-rays emitted by the circulating electrons.

X-ray mirrors. If radiation is incident normally (that is, perpendicular) to a surface between two media of differing refractive index, the fraction of the energy that is reflected is $^1/_4(\delta^2 + \beta^2)$. This is clearly impractically small for a normal-incidence mirror for x-rays. However, useful mirrors can be constructed by using the principle of total reflection. If electromagnetic waves are incident on the boundary between one material of refractive index n_1 and another of lower refractive index n_2, there exists an angle of incidence I_c, called the critical angle, given by Eq. (2). If the angle

$$\sin I_c = \frac{n_2}{n_1} \tag{2}$$

of incidence (the angle of incident radiation with respect to the normal to the surface) is greater than this critical angle, all the wave energy is reflected back into the first medium. This phenomenon can be seen when looking upward into an aquarium tank; objects in the tank are reflected in the surface of the water, which acts as a perfect mirror. An analogous situation occurs for x-rays. Since the refractive index for all materials is slightly less than 1, x-rays incident from vacuum (or air) on a polished surface of, say, a metal encounter a lower refractive index and there exists a critical angle given by $\sin I_c = 1 - \delta$. Since δ is very small, I_c is very close to 90°. In this case the angle of incidence is customarily measured from the tangent to the surface rather than from the normal, and the angle $\theta_c = 90^\circ - I_c$ is termed the angle of glancing (or grazing) incidence. This angle is typically in the range 0.1–1.0°. *See* REFLECTION OF ELECTROMAGNETIC RADIATION.

Reflecting x-ray optics. Glancing incidence reflection of x-rays was first demonstrated in 1922. One of the first applications was the absolute measurement of the wavelength of x-rays, using a reflecting diffraction grating consisting of numerous parallel grooves of precisely known separation ruled on a reflecting surface. However, attempts to construct a reflecting system to produce x-ray images were long frustrated by the severe image defects, or aberrations, suffered by mirrors used at glancing incidence. The most severe of these is astigmatism; whereas a mirror in normal incidence forms reasonably good images of a point source, at glancing incidence these become line foci separated by a large distance. In 1948 P. Kirkpatrick and A. Baez solved this problem by crossing two spherical mirrors (**Fig. 1***a*). However, microscopy using reflection optics has never achieved wide acceptance for biological or medical research because of the difficulty of removing the remaining aberrations, and image degradation caused by scattering from mirror-surface imperfections. Such microscopes have been used to photograph the implosion of fuel pellets in laser-fusion experiments. *See* ABERRATION (OPTICS); DIFFRACTION GRATING; NUCLEAR FUSION; X-RAY MICROSCOPE.

Optics in the Kirkpatrick-Baez configuration plays an important role in constructing beam lines for research at synchrotrons and storage rings. A typical beam line consists of three to seven glancing-incidence mirrors and diffraction gratings (Fig. 1*b*). These components perform the functions of collecting the x-rays emitted by the circulating electrons, separating the x-rays into spectral components by using monochromators, and focusing the x-rays onto experimental samples. *See* SYNCHROTRON RADIATION.

Astigmatism can alternatively be corrected with a single glancing-incidence mirror having a toroidal or tubelike shape. Such a mirror must have a radius of curvature in the direction along the tube axis which is $1/\sin^2 \theta_c$ times greater than the radius in the perpendicular direction. Such mirrors are more difficult to make than the spherical or cylindrical forms used in the Kirkpatrick-Baez system. A true image-forming device requires two such reflecting elements to correct additional aberrations such as coma. Such a two-mirror device was invented by H. Wolter in 1952. The first mirror of a Wolter microscope is an ellipse of revolution, the second a confocal hyperboloid. A few microscopes of this type have been used in laser-fusion research and in synchrotron radiation microscopy. Woltier x-ray telescopes, designed to observe objects at infinity, have for the first element a paraboloid of revolution in place of an ellipsoid. They have been used extensively in x-ray astronomy aboard orbiting spacecraft. *See* X-RAY ASTRONOMY; X-RAY TELESCOPE.

Multilayer optics. Although the reflectivity of surfaces at glancing angles greater than the critical angle is very small, this reflectivity can be enhanced by depositing a stack of ultrathin films having alternately high and low values of δ on the surface (**Fig. 2**). The individual thicknesses of these films is adjusted so that the reflections from each interface add in phase at the top of the stack in exact analogy to the multilayer mirrors used for visible light. However, whereas visible multilayers require film thicknesses of hundreds of nanometers, in the x-ray region the thickness of each film must be between 1 and 100 nm. Such ultrathin films can be made by a variety of vacuum deposition methods, commonly sputtering and evaporation. The response of these artificial multilayers is strongly wavelength-selective. For incident x-rays of wavelength λ, the reflectivity of the stack has peaks at the angles θ_n given by the Bragg equation (3). Ordinarily the multilayers are used in

$$n\lambda = 2d \sin \theta_n \qquad (3)$$

the first order, that is, $n = 1$. *See* X-RAY DIFFRACTION.

As a coating for glancing-incidence optics, multilayers allow a mirror to be used at a shorter wavelength (higher x-ray energy) for a given glancing angle, increasing the projected area and thus the collection efficiency of the mirror. At wavelengths longer than 3 or 4 nm, multilayer mirrors can be used to make normal-incidence mirrors of relatively high reflecting power. For example, stacks consisting of alternating layers of molybdenum and silicon can have reflectivities as high as 65% at wavelengths of 13 nm and longer. These mirrors have been used to construct optical systems that are exact analogs of mirror optics used for visible light. For example, normal-incidence x-ray telescopes have photographed the Sun's hot outer atmosphere at wavelengths of around 18 nm. Multilayer optics at a wavelength of 13.5 nm can be used to perform x-ray microlithography by the projection method to print features of dimensions less than 100 nm.

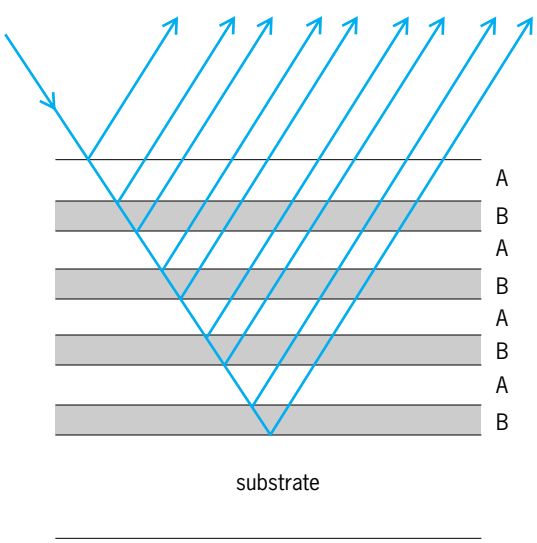

Fig. 2. Reflection of x-rays by a mirror, with alternating layers of materials A and B. The weak reflected beams from each of the interfaces combine with the correct phase relationship to form a strong reflected beam.

Crystal optics. Crystals are natural multilayer structures and thus can reflect x-rays. Many crystals can be bent elastically (mica, quartz, silicon) or plastically (lithium fluoride) to make x-ray focusing reflectors. These are used in devices such as x-ray spectrometers, electron-beam microprobes, and diffraction cameras to focus the radiation from a small source or specimen on a film or detector. Until the advent of image-forming optics based on mirrors and zone plates, the subject of x-ray diffraction by crystals was called x-ray optics. *See* X-RAY CRYSTALLOGRAPHY; X-RAY SPECTROMETRY.

Zone plates. Zone plates are diffraction devices that focus x-rays and form images. They are diffracting masks consisting of concentric circular zones of equal area, and are alternately transparent and opaque to x-rays. Whereas mirrors and lenses focus radiation by adjusting the phase at each point of the wavefront, zone plates act by blocking out those regions of the wavefront whose phase is more than a half-period different from that at the plate center. Thus a zone plate acts as a kind of x-ray lens. Zone-plate microscopy is the most promising candidate method for x-ray microscopy of biological specimens. *See* DIFFRACTION.

James H. Underwood

Bibliography. B. Aschenbach, X-ray telescopes, *Rep. Prog. Phys.*, 48:579–629, 1984; G. Margaritondo, *Introduction to Synchrotron Radiation*, 1988; E. Spiller, *Soft X-ray Optics*, 1994.

X-ray powder methods

Physical techniques used for the identification of substances, and for other types of analyses, principally for crystalline materials in the solid state. In these techniques, a monochromatic beam of x-rays is directed onto a polycrystalline (powder) specimen, producing a diffraction pattern that is recorded on

film or with a diffractometer. This x-ray pattern is a fundamental and unique property resulting from the atomic arrangement of the diffracting substance. Different substances have different atomic arrangements or crystal structures, and hence no two chemically distinct substances give identical diffraction patterns. Identification may be made by comparing the pattern of the unknown substance with patterns of known substances in a manner analogous to the identification of people by their fingerprints. The analytical information is different from that obtained by chemical or spectrographic analysis. X-ray identification of chemical compounds indicates the constituent elements and shows how they are combined.

The x-ray powder method is widely used in fundamental and applied research; for instance, it is used in the analysis of raw materials and finished products, in phase-diagram investigations, in following the course of solid-state chemical reactions, and in the study of minerals, ores, rocks, metals, chemicals, and many other types of material. The use of x-ray powder diffraction methods to determine the actual atomic arrangement, which has been important in the study of chemical bonds, crystal physics, and crystal chemistry, is described in related articles. *See* X-RAY CRYSTALLOGRAPHY; X-RAY DIFFRACTION.

Lattice geometry. The atoms in crystalline substances are arranged in a symmetrical three-dimensional pattern: some atomic arrangement is repeated by the symmetry of the crystal along straight lines throughout the crystal. The array of points and lines in **Fig. 1** outlines a lattice or framework of a typical crystal in which the third dimension is normal to the plane of the drawing. The smallest group of atoms which has the symmetry of the entire pattern is called the unit cell. There are several ways in which the unit cell is selected. In this case it was drawn parallel to the crystallographic axes a, b, and c. The traces of the various lattice planes (normal to the drawing) are indicated by heavy lines. The method used by crystallographers to identify these planes is as follows. For a given set of planes, count the number of planes crossed from one lattice point to the next along a, then repeat the procedure along b and c. The resulting numbers, called $hk\ell$, respectively, are known as the Miller indices of that set of planes, and assignment of indices to each line is called indexing the pattern.

The spacings d between the planes are related to the Miller indices and the unit cell dimensions. In crystals of the cubic system, the crystallographic axes are normal to each other and have the same length, $a = b = c$; and d is given by Eq. (1). Similar

$$d = \frac{a_o}{\sqrt{(h^2 + K^2 + \ell^2)}} \qquad (1)$$

relations exist for the other five crystal systems. The relations for the low-symmetry systems, monoclinic and triclinic, are much more complicated. Various types of charts to facilitate indexing tetragonal, orthorhombic, and hexagonal substances are available.

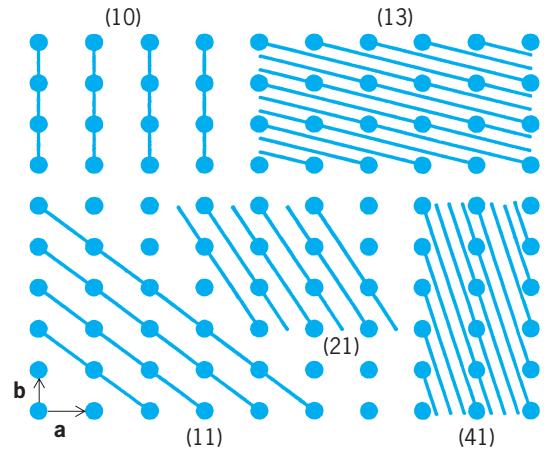

Fig. 1. Lattice of a typical crystal. The third dimension is normal to the plane of the drawing. Traces of various lattice planes are indicated by heavy lines, and their Miller indices are given.

See CRYSTAL STRUCTURE; CRYSTALLOGRAPHY.

Bragg's law. Atomic diameters have approximately the same dimensions as the wavelengths of x-rays, and therefore the crystal can act as a three-dimensional grating for x-rays in a manner analogous to the diffraction of ultraviolet or visible light by a ruled one-dimensional grating. Under appropriate conditions, the electrons around each atom scatter the incident x-ray beam in a coherent manner and in certain specific directions; the scattering from billions of atoms is in phase at the same time. *See* DIFFRACTION GRATING.

Shortly after the discovery of x-ray diffraction by M. von Laue in 1912, this complex phenomenon was formulated in a simple relation by William H. Bragg and William L. Bragg, in which the diffraction is visualized as a reflection from a large number of parallel planes. (This should not be confused with the total reflection of x-rays, which occurs at very small grazing angles from highly polished surfaces.) They showed that when two or more parallel rays of the same wavelength are incident at the same glancing angle Θ to a set of atomic planes, the path difference of the reflected rays from adjacent planes is one wavelength. This may be expressed as Eq. (2), where

$$\lambda = 2d_{hk\ell} \times \sin\Theta \qquad (2)$$

λ is the wavelength of the incident x-rays and $d_{hk\ell}$ the interplanar spacing between the $(hk\ell)$ planes of atoms. Thus the conditions for x-ray reflection are very restrictive because these is only one angle Θ at which the x-rays of a given wavelength are reflected by a particular set of atomic planes of spacing d. For a given wavelength, larger d spacings appear at smaller angles. Moreover, differentiation of the Bragg equation yields Eq. (3). This equation shows that the shift

$$\frac{\Delta\Theta}{\Delta d} = \frac{\tan\Theta}{d} \qquad (3)$$

in line position $\Delta\Theta$ due to a change of lattice spacing Δd increases as the tangent of the angle and reaches a maximum at $\Theta = 90^\circ$ (reflection angle $2\Theta = 180^\circ$),

where tan Θ is infinite. Hence the highest-angle lines in the back-reflection region of the pattern are the most sensitive to changes in the lattice spacings and consequently supply the most accurate data for the measurement of the unit cell dimensions.

Characteristics of powder patterns. Many materials are not available in the form of large single crystals, and moreover it is impractical to obtain all the x-ray reflections from single crystals for identification purposes. If the sample does not already exist in polycrystalline form, it may be pulverized. When a fine-grained powder consisting of thousands of small, randomly oriented crystallites is exposed to the x-ray beam, all the possible reflections from the various sets of atomic planes can occur simultaneously. **Figure 2** shows two examples of diffraction patterns recorded for barium titanate ($BaTiO_3$). Four basic pieces of information can be derived from the experimental pattern: (1) the 2Θ value from which the *d*-spacing can be calculated; (2) the absolute intensity, from which relative intensities can be calculated; (3) the peak width; (4) the form of the background.

For historical reasons having to do with data storage and reproduction problems, and experimental pattern is typically reduced to yield a table of *d*-values and relative intensities. These data are called the reduced pattern.

The complexity of the pattern is determined primarily by the symmetry of the substance rather than by its chemical composition. Hence, chemically complex compounds can have patterns that are nearly as simple as those from iron or copper. The number of lines also increases with unit cell size. When two or more substances are present in the sample, the pattern of each substance appears independent of the others. This makes the identifications more difficult, but mixtures containing as many as six substances have been successfully analyzed. It is also possible to make quantitative analyses of the mixtures by comparing the relative intensities of one or more principal lines of each substance. Usually some reference standards of known chemical composition are prepared to facilitate the interpretation.

One of the most important characteristics of powder patterns is that isostructural substances (that is, substances with the same crystal structure) give similar diffraction patterns. **Figure 3** shows diffraction patterns of the various phases of quartz (α-SiO_2), which have different crystal structures. Indeed, many substances may occur in two or more crystal structures; that is, they may have polymorphic forms but the same chemical composition. Such forms may be caused by slight differences in chemical preparation, different heat treatments, and other factors. *See* POLYMORPHISM (CRYSTALLOGRAPHY); QUARTZ.

When a substance is strained or plastically deformed, the x-ray lines broaden, and when the strain is removed by annealing, the lines return to their original sharpness. Figure 2*a* shows the diffractometer recording of a sample of barium titanate that has been strained by crushing, while Fig. 2*b* shows the same substance in an unstrained state. X-ray patterns can thus be used to follow the course of heat treatment or other processes used to remove strains.

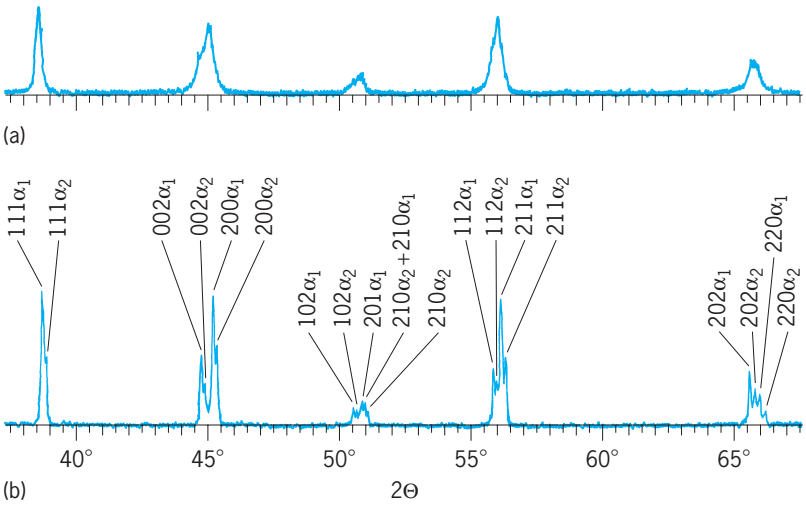

Fig. 2. Diffractograms of barium titanate ($BaTiO_3$). (*a*) Strained specimen with broadened diffraction lines. (*b*) Unstrained specimen. (*After Transactions of the Instrumentation and Measurements Conference, Stockholm, 1952*)

If the crystallites do not have a completely random orientation, the line shapes and relative intensities will change accordingly. For example, in rolling thin sheets of metal or in drawing wire, the crystallites align themselves in special ways, depending on the mechanical conditions of the process. The special or preferred orientation of the crystallites gives x-ray diffraction patterns which may be markedly different from those of the random crystallites. Similar conditions may arise in electroplating, where the plating conditions or the substrate may cause the crystallites to have a nonrandom orientation. Comparison of the random and oriented x-ray patterns shows the degree of orientation in the sample. *See* ELECTROPLATING OF METALS.

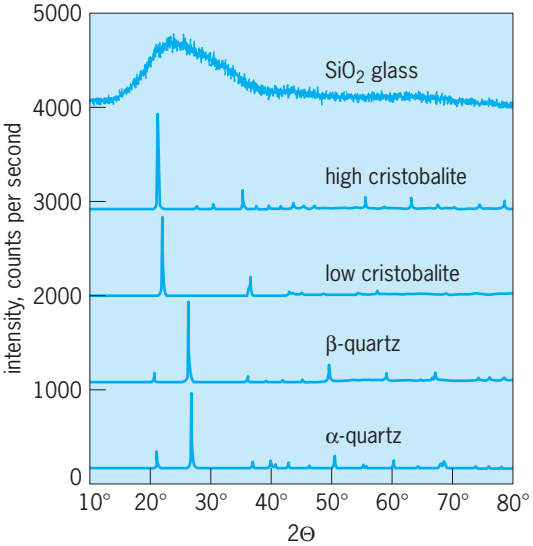

Fig. 3. Diffraction patterns of polymorphic phases of SiO_2 (quartz). (*After R. Jenkins and R. L. Snyder, Introduction to X-ray Powder Diffractometry, Wiley, 1996*)

TABLE 1. Types of x-ray diffraction databases

Name	Content	Location
Powder Diffraction File (PDF)*	Inorganic, organic, mineral, and so on	Newtown Square, PA
Cambridge Structural Database (CSD)†	Organic, organo-metallic	Cambridge, U.K.
Inorganic Crystal Structure Database (ICSD)‡	Inorganic materials	Karlsruhe, Germany
NRCC Metals Data File (CRYSTMET)	Metals and alloys	Ottawa, Canada
Protein Data Bank (PDB)	Structure of macromolecules	Brookhaven, NY
NIST Crystal Data [NBS(CD)]	Inorganic and organic unit cells	Gaithersburg, MD

*Maintained by International Centre for Diffraction Data (ICDD).
†Maintained by Cambridge Crystallographic Data Centre (CCDC).
‡Maintained by Fachinformationszentrum (FIZ).

When the crystallites are very small, for instance, about 1 micrometer, the x-ray lines broaden by an amount which increases with decreasing size. Comparisons with the line breadths of samples of larger crystallite sizes may lead to a measure of the average crystallite size in the specimen.

Amorphous materials such as glasses and liquids give patterns which consist of only a few broad lines superimposed on a continuous background. As an example, the upper diffractogram in Fig. 3 shows an amorphous pattern. Patterns of various amorphous substances closely resemble each other, and hence the method is impractical for identification. On the other hand, since the two types of patterns are so different, the method is ideally suited to distinguish between crystalline and amorphous substances and to determine the degree of crystallinity of substances between the two extremes. The study of the progress of devitrification (crystallization) of a glass and similar problems are frequently accomplished with x-rays. Certain chemical and structural properties of liquids might be ascertained from the diffraction patterns when the liquids are frozen to form crystals. *See* AMORPHOUS SOLID; GLASS; LIQUID.

There are also many smaller changes in the x-ray pattern which may reveal important information. In substitutional solid solutions, for example, atoms of different elements may substitute for one another and occupy the same relative positions as in the pure metals. The substitutions of solute atoms occur on the same lattice sites occupied by the solvent atoms, but are randomly distributed. If the atoms are of different size, the average unit cell size will change accordingly. In simple cases, it is possible to determine the chemical composition of intermediate members by measuring the unit cell dimensions because there is often a nearly linear relationship between the two. In interstitial solid solutions, atoms are added to the empty spaces in the structure and there is little, if any, change in the dimensions. *See* ALLOY STRUCTURES; SOLID SOLUTION.

Diffraction databases. There are a number of databases available for x-ray diffraction work (**Table 1**). The majority of these databases are designed and maintained for the single-crystal community rather than for the powder community. Nevertheless, much cross fertilization takes place. For example, a number of the patterns in the Powder Diffraction File (PDF) are calculated from single-crystal data of the type contained in the other databases.

While these databases have proven their usefulness in a wide range of applications, there has been little attempt to exploit a combination of them. Agreements allow mutual use of the PDF, the Cambridge Structural Database (CSD), and the Inorganic Crystal Structure Database (ICSD). Three major advantages have accrued from this cooperation. (1) By use of cross-reference "hooks" for each database entry, the user has access to experimental powder data from the PDF and structural information from the ICSD, permitting the full modeling of the experimental pattern. (2) The PDF can be supplemented using powder patterns calculated from structural data. (3) The combined efforts of the various editorial groups can only help to improve the overall content and quality of diffraction data.

In addition, the International Centre for Diffraction Data (ICDD), which maintains the PDF, archives fully digitized raw data sets (PDF-3). The availability of the trace of the original experimental data can often give useful background information for the editorial process. In addition, the success rate of the search-match process has been dramatically improved by a strategy based on searching the whole observed pattern with its background (not just the reduced pattern), and on adding candidate phases together to compose, rather than decompose, an observed multiphase pattern.

Powder Diffraction File. The Powder Diffraction File is a collection of single-phase x-ray powder diffraction patterns in the form of tables of the interplanar spacings (d) and relative intensities (I^{rel}) characteristic of the compound. New and updated diffraction patterns are continually added to the PDF, and an automated editorial system allows their detailed review. Currently 2500 such patterns are added each year, comprising 1900 inorganic patterns and 600 organic patterns. There is a continuing effort to ensure that a significant proportion of the new phases represent current needs and trends in industry and research. The master database of powder patterns is continually undergoing revision and updating, but in order to ensure that all database users have the opportunity to work with the same version, a frozen version of the master database is produced each year as the PDF-2.

The PDF-2 contains a series of individual data

33-1161 ★

SiO_2

Silicon Oxide

Quartz, syn

Rad. CuKα$_1$ **λ** 1.540598 **Filter** Mono. **d-sp** Diff.
Cut off **Int.** Diffractometer **I/I$_{cor.}$** 3.6
Ref. *Natl. Bur. Stand. (U.S.) Monogr. 25*, **18** 61 (1981)

Sys. Hexagonal **S.G.** P3$_2$21 (154)
a 4.9133(2) **b** **c** 5.4053(4) **A** **C** 1.1001
α **β** **γ** **Z** 3 **mp**
Ref. Ibid.

D$_x$ 2.65 **D$_m$** 2.66 **SS/FOM** F$_{30}$=77(.013,31)

εα **nωβ** 1.544 **εγ** 1.553 **Sign** + **2V**
Ref. Swanson, Fuyat, *Natl. Bur. Stand. (U.S.), Circ. 539*, **3** 24 (1954)

Color Colorless
Pattern taken at 25 C. Sample from the Glass Section at NBS, Gaithersburg, Maryland, USA, ground single-crystals of optical quality. Pattern reviewed by Holzer, J., McCarthy, G., North Dakota State University, Fargo, North Dakota, USA, *ICDD Grant-in-Aid* (1990). Agrees well with experimental and calculated patterns. O_2Si type. Quartz group. Also called: silica. Also called: low quartz. Silicon used as internal standard. PSC: hP9. To replace 5-490 and validated by calculated pattern. Plus 6 additional reflections to 0.9089.

dÅ	Int	hkℓ	dÅ	Int	hkℓ
4.257	22	100	1.1532	1	311
3.342	100	101	1.1405	<1	204
2.457	8	110	1.1143	<1	303
2.282	8	102	1.0813	2	312
2.237	4	111	1.0635	<1	400
2.127	6	200	1.0476	1	105
1.9792	4	201	1.0438	<1	401
1.8179	14	112	1.0347	<1	214
1.8021	<1	003	1.0150	1	223
1.6719	4	202	0.9898	1	402
1.6591	2	103	0.9873	1	313
1.6082	<1	210	0.9783	<1	304
1.5418	9	211	0.9762	1	320
1.4536	1	113	0.9636	<1	205
1.4189	<1	300			
1.3820	6	212			
1.3752	7	203			
1.3718	8	301			
1.2880	2	104			
1.2558	2	302			
1.2285	1	220			
1.1999	2	213			
1.1978	1	221			
1.1843	3	114			
1.1804	3	310			

Fig. 4. Example of an image in the PDF-2 (Powder Diffraction File).

sets (**Fig. 4**), each of which contains a list of *d/I* pairs, a chemical formula, a name, a unique identification (PDF) number, and a reference to the primary source. Supplemental data may be added, including Miller indices for all lines, unit cell and space group data, physical constants, and experimental details. Because the number of patterns is large, special ways of organizing the *d*'s and *I*'s into subfiles have been devised, such as minerals, zeolites, and pharmaceuticals. A smaller version of the database called PDF-1 is produced mainly for search programs on minicomputers with limited disk storage. As of 2001, about 160,000 patterns of the PDF-2 database require about 620 megabytes of storage, and PDF-1 about 39 megabytes.

Instrumentation. There are many types of powder diffractometer available ranging from simple laboratory instruments to versatile and complex instruments using a synchrotron source. Specialized instruments allow recording of diffraction patterns under nonambient conditions, including variable temperature, pressure, and atmosphere. Completely automated equipment for x-ray analysis is available. Most laboratory instruments consist of a high-voltage generator which provides stabilized voltage for the x-ray tube, so that the x-ray source intensity varies by less than 1%. A diffractometer goniometer is mounted on a table in front of the x-ray tube window. Electronic circuits use an x-ray detector to convert the diffracted x-ray photons to measurable voltage pulses, and to record the diffraction data.

The number of elements useful for x-ray tube targets is limited to a few, of which copper is the most commonly used. The copper K spectrum consists of a few lines superimposed on a broad continuum. These lines are the intense Cu Kα_1/Kα_2 doublet and the weaker Cu Kβ_1/β_3 doublet. In the α doublet, the Kα_1 line at 0.1541 nanometer is about twice as intense as the Kα_2 line at 0.1544 nm. The individual components of the Kβ_1/β_3 doublet are not usually resolved, and appear as a single line of average wavelength 0.1392 nm. The continuum begins at a wavelength which depends on the applied voltage, rises rapidly to a broad maximum, and then diminishes with increasing wavelength. The intensity of the continuum is mainly dependent on the x-ray tube voltage, but it also increases with atomic number of the target element. The entire spectrum is scattered and diffracted with varying efficiencies by the polycrystalline specimen. The interpretation of the diffraction pattern is simplified if the Kβ line is eliminated. This has traditionally been accomplished by inserting a β-filter in the x-ray path. The filter consists of a thin nickel foil which almost completely absorbs the Kβ radiation and transmits the Kα. More modern diffractometers employ a primary or diffracted beam monochromator, or a proportional Si(Li) detector to achieve the same result.

The two principal methods of recording diffraction patterns are by the use of film and with x-ray counter tubes. In the film method (**Fig. 5*a***), a small amount of the powder is glued to a thin glass fiber and rotated continuously in the center of a powder camera to give better randomness. The latter is essentially a light-tight cylindrical enclosure, usually about 4.5 in. (11 cm) in diameter and provided with collimators to define the x-ray beam. The strip of x-ray film is placed against the inside circumference of the cylinder concentric with the specimen axis. Two holes are punched in the film, one to admit the

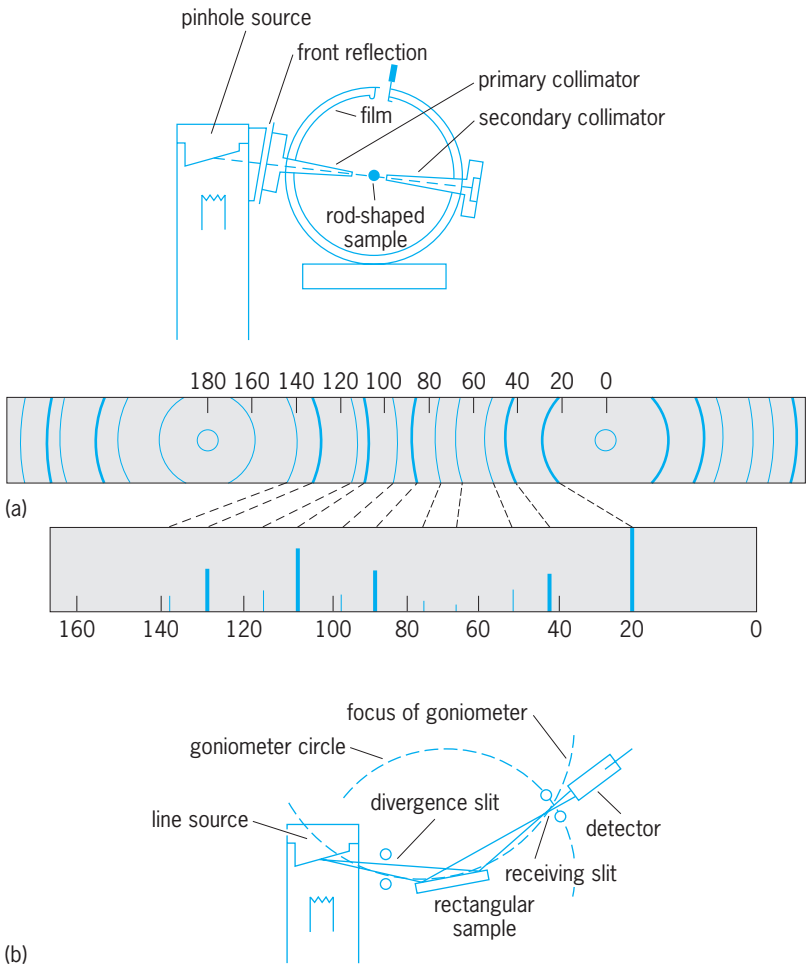

Fig. 5. Instrumentation for the measurement of x-ray powder patterns. (*a*) Debye-Scherrer camera. (*b*) Powder diffractometer.

collimator, and the other to allow the undiffracted beam to pass through. Exposures of 1–4 hours are required. Each properly oriented crystallite may produce a reflection that appears as a spot on the film. For each set of atomic planes there will be a large number of crystallites of various orientations, each producing its own spot, and all the reflected rays will form the surface of a cone, with its apex at the specimen and subtending the angle 4Θ on the film. The various sets of atomic planes thus produce a series of concentric cones, and these appear on the film as a series of arcs whose curvature depends on the reflection angle. After the film is developed and dried, it is laid flat and the linear distance is measured from the position on the film where the direct beam passed through to each of the arcs. Since, the wavelength λ and the camera diameter are known and since allowance may be made for film shrinkage in development, these linear distances may be converted into angular values, from which d may be calculated for each set of planes by substituting in the Bragg equation.

The more common diffractometer method employs a counter tube instead of film (Fig. 5*b*). The specimen is prepared with a flat surface which is exposed to the x-ray beam. The counter tube always points toward the specimen, and the reflection data are recorded line by line. The divergent primary x-ray beam converges, after reflection from the specimen, to a narrow slit in front of the counter tube. By rotating the counter tube at twice the angular speed of the specimen so that the specimen surface always makes an angle Θ with the primary beam when the counter tube is at 2Θ, sharp well-resolved lines are obtained at all reflection angles. This instrument makes possible the direct and accurate measurement of intensities and reflection angles. The peak-to-background ratios, resolution, and line shapes are far superior to those obtained by the cylindrical film camera. The operation is completely automatic.

Phase identification. Since every crystalline material gives, at least in principle, a unique x-ray diffraction pattern, study of diffraction patterns from unknown phases offers a powerful means of qualitative identification by comparing an x-ray pattern from the material to be analyzed with a file of single-phase reference patterns. Although the potential for qualitative phase identification was certainly recognized from the very early days of x-ray diffraction, the first attempts to list reference patterns were not published in detail until the late 1930s, along with means of archiving and retrieval of patterns. These methods still provide the basis of many search/match methods in use today. All databases require some type of index system to allow access to information contained within the database. For searching print databases, it is common practice to supply a combination index plus search manual for each of the main subsets of the PDF. A number of manual searching methods have been developed based on the three methods in common use: the alphabetical method, the Hanawalt method and the Fink method. The alphabetical index is designed to permit a rapid systematic search for all patterns with a specified chemical content. The Hanawalt method involves grouping the patterns in the PDF according to the d-value of the 100% intensity line. Each interval is sorted on the d-value of the second most intense line. Subsequent lines are listed in order of decreasing intensity. The Fink method indexes a pattern on its eight largest d-spacing lines, and eight separate entries are made using a cyclic permutation of d-values. An *Alphabetical Index* and a *Hanawalt Search Manual* are published annually, and a *Fink Method Search Manual* is published at irregular intervals. **Table 2** illustrates the various entry methods employed in the common indexes. The *Alphabetical Index* is a chemistry-based index using only elemental information. The *Hanawalt Index* is an intensity-driven index since it employs only the strongest lines for searching. The *Fink Index* is a d-spacing-driven index since it employs mainly the largest d-values. There have been numerous attempts to completely automate the search-match process resulting in several very successful commercial products, which permit routine external and internal standard calibration, precision alignment checks, and new levels of accuracy and precision in data collection and analysis.

Quality of experimental data. The ability to recognize a reference pattern in an unknown material

TABLE 2. Types of PDF data searching indexes

Index	Entry method	Search parameters
Alphabetic	Chemistry	Permuted elemental symbols
Hanawalt	*I*/*d*	Three strongest lines
Fink	*d*/*I*	First eight lines
EISI	Chemistry/*d*	Low/high *Z* elements; *d*-spacing
Boolean	Various	*d*-spacings, chemistry, strong lines, CODEN,* physical properties, functional groups, and so on

* Designation assigned to a periodical title by the Chemical Abstracts Service.

strongly depends on the quality of the *d*'s and *I*'s in both the reference material and the unknown sample. One of the principal problems in the identification of materials by comparison of an experimental pattern with reference patterns is the variability in the quality of the data. For Debye-Scherrer camera data, one should assume an error window of $\pm\Delta 2\Theta = 0.1^\circ$; for normal diffractometer data, one typically assumes a $\pm\Delta 2\Theta = 0.05^\circ$; and for internal standard corrected diffractometer or Guinier camera data, a $\Delta 2\Theta$ window as low as 0.01° may be assumed. Although the experimentally measured parameter is generally the 2Θ value, the search-match parameter is invariably the *d*-value. Unfortunately, the error relationship between 2Θ and *d* is nonlinear. The most useful lines for phase identification are the low-angle lines, and these are also the lines subject to the largest error in *d*. As the quality of both reference and experimental patterns improves, the problem of pattern recognition becomes easier. There has been a major effort to meet the ever-increasing demand for the higher-quality data needed because of improved instrumentation and better techniques. The introduction of the computer for data collection, treatment, and processing has improved the quality of measured *d*-spacings, leading to an ongoing need for improvement in the quality of reference patterns. For example, the modern automated powder diffractometer can produce *d*-spacing accuracies of about 1 part per 1000 for all but the larger *d*-values. This quality of data corresponds to an average angular error of 0.01–0.03° in 2Θ.

Ron Jenkins

Bibliography. D. L. Bish and J. E. Post, *Modern Powder Diffraction*, Mineralogical Society of America, 1989; B. D. Cullity, *Elements of X-ray Diffraction*, 2d ed., Addison-Wesley, 1978; R. Jenkins and R. L. Snyder, *Introduction to X-ray Powder Diffractometry*, Wiley, 1996; H. P. Klug and L. E. Alexander, *X-ray Diffraction Procedures: For Polycrystalline and Amorphous Materials*, 2d ed., 1974.

X-ray spectrometry

A rapid and economical technique for quantitative analysis of the elemental composition of specimens. It differs from x-ray diffraction, whose purpose is the identification of crystalline compounds. It differs from spectrometry in the visible region of the spectrum in that the x-ray photons have energies of thousands of electronvolts and come from tightly bound inner-shell electrons in the atoms, whereas visible photons come from the outer electrons and have energies of only a few electronvolts. *See* X-RAY DIFFRACTION.

In x-ray spectrometry the irradiation of a sample by high-energy electrons, protons, or photons ionizes some of the atoms, which then emit characteristic x-rays whose wavelength λ depends on the atomic number *Z* of the element ($\lambda \propto 1/Z^2$), and whose intensity is related to the concentration of that element. Generally speaking, the characteristic x-ray lines are independent of the physical state (solid or liquid) and of the type of compound (valence) in which an element is present, because the x-ray emission comes from inner, well-shielded electrons in the atom. **Figure 1** illustrates the removal of one of the innermost, K-shell, electrons by a high-energy photon. The photon energy must be greater than the binding energy of the electron; the difference in energy appears as the kinetic energy of the ejected electron. The K-ionized atom is unstable, and one of the L- or M-shell electrons drops into the K-shell vacancy within 10^{-14} to 10^{-6} s. As this transition occurs, a characteristic x-ray photon is emitted with an energy equal to the difference in energy between the K and the L (or M) shell, or an additional electron, called an Auger electron, is ejected from the atom. Either the x-rays or the Auger electrons may be used for analysis, but in this article the discussion is concerned exclusively with the x-rays. *See* AUGER EFFECT.

Figure 2 shows some of the allowed transitions and the naming of the lines. There is a selection rule in atomic physics which says that only certain ones of the outer electrons are allowed to fill a vacancy in an inner shell. The rule can only be stated in terms of quantum mechanics: the transition must be from one shell to another, and the second (orbital) quantum number must change by ± 1, that is, $p \rightleftarrows s$, $d \rightleftarrows p$, and so on. **Figure 3** shows the λ versus *Z* relationship for the strongest *K*- and *L*-series lines. Figure 2 shows that some of the transitions may come from the valence shell, in which case there will be slight alteration of wavelength or line shape of characteristic

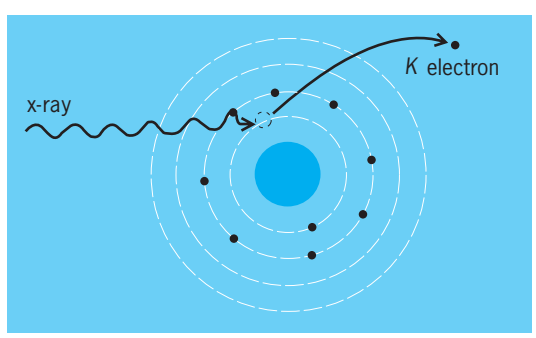

Fig. 1. Removal of a *K* electron from an atom by a primary x-ray photon. (*After L. S. Birks, X-ray Spectrochemical Analysis, 2d ed., Wiley-Interscience, 1969*)

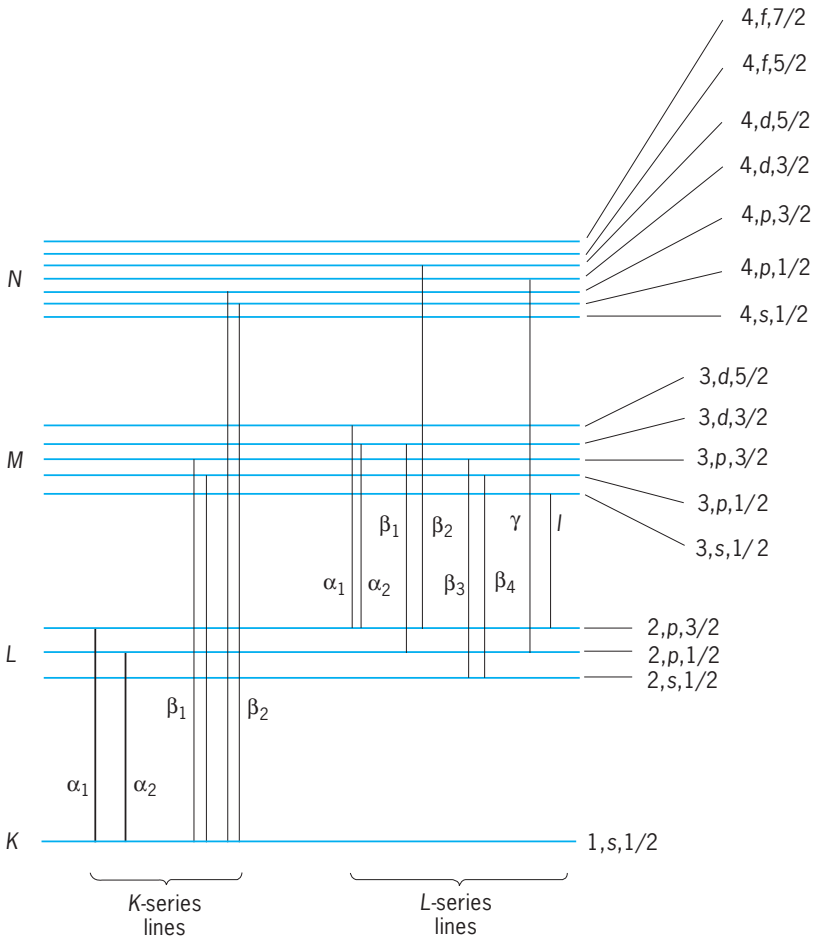

Fig. 2. Partial energy-level diagram showing the transitions leading to the *K* and *L* series lines. (*After L. S. Birks, X-ray Spectrochemical Analysis, 2d ed., Wiley-Interscience, 1969*)

lines with valence. This alteration can be measured for multivalent elements such as sulfur by using spectrometers designed for high resolution (**Fig. 4**).

Spectrum analysis. Photon generation of characteristic spectra is the most common and is called x-ray fluorescence. It is carried out with an x-ray tube as the source of primary radiation. There are two ways of analyzing the spectra: wavelength dispersion and energy dispersion.

Wavelength dispersion. This is shown in **Fig. 5***a*. The characteristic emission from the sample is usually excited by a chromium or tungsten target x-ray tube which is operated at 2–3 kW. The emitted radiation is limited to a parallel beam by the blade collimator and is diffracted, one wavelength at a time, by an analyzer crystal. Bragg's law ($n\lambda = 2d \sin \theta$) relates the diffraction angle θ to the wavelength λ for crystal planes with an interatomic-spacing distance d. The term n is the order of the diffraction, 1, 2, 3, etc., as it is in classical optics. As shown in Fig. 3, the characteristic wavelength decreases as the atomic number increases. The $2d$ spacing of the analyzing crystal must be greater than the wavelength being diffracted, but if it is too much greater the spectral lines will be crowded toward small θ. It has become the practice to use a crystal of lithium fluoride for elements of atomic number Z greater than 20, a pentaerythritol crystal for Z between 13 and 20, and a potassium acid phthalate crystal for Z between 8 and 13. X-ray fluorescence analysis is not generally suitable below about $Z = 9$, but with special instruments and techniques it can be extended down to about $Z = 5$.

The detectors used for wavelength dispersion are either gas proportional counters or scintillation counters. Both of these count individual photons, but the gas counters are most suitable for wavelengths longer than about 0.2 nanometer, while the scintillation counters are most suitable for shorter wavelengths. The amplitude of the output pulse for each photon is proportional to the energy of the x-ray photon it represents. However, the statistical variation in the amplitude for each specific photon energy means that characteristic lines from neighboring elements are not resolved by either gas proportional

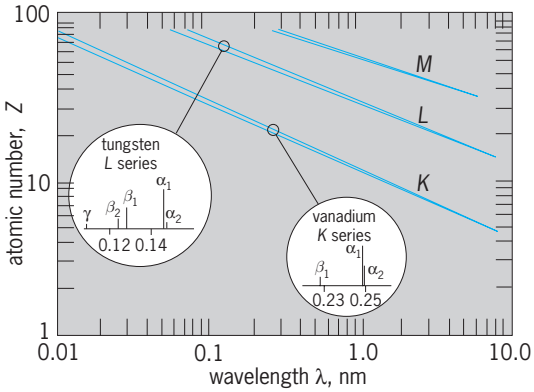

Fig. 3. Atomic number *Z* versus wavelength λ of the characteristic lines ($\lambda \propto 1/Z^2$). (*After L. S. Birks, X-ray Spectrochemical Analysis, 2d ed., Wiley-Interscience, 1969*)

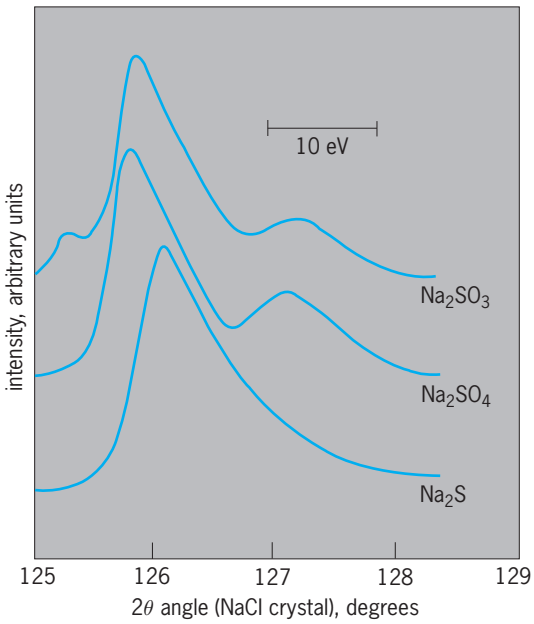

Fig. 4. Variations of the sulfur *Kβ* line with valence state of sulfur. (*After L. S. Birks and J. V. Gilfrich, X-ray fluorescence analysis of the concentration and valence state of sulfur in pollution samples, Spectrochim. Acta, 33B, no. 7:305, 1978*)

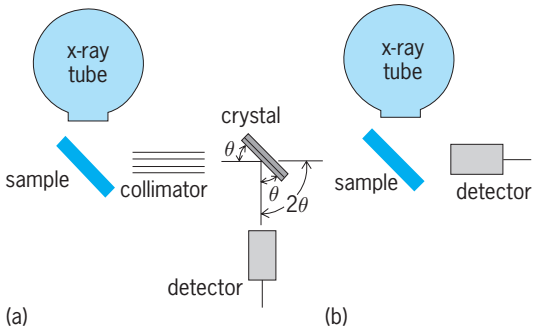

Fig. 5. Wavelength and energy dispersion methods. (*a*) Wavelength dispersion with a crystal spectrometer. (*b*) Energy dispersion with a solid-state detector.

or scintillation counters. **Figure 6** shows the resolution for the gas proportional counter alone and with a crystal spectrometer, as well as for the silicon solid-state detector used for energy dispersion analysis, which is discussed below; resolution with a scintillation counter is almost a factor-of-3 worse than a gas proportional counter. In wavelength dispersion it is the resolution of the crystal spectrometer, however, which determines the separation between neighboring wavelengths, and the crystal resolution is better than any of the detectors.

Energy dispersion. In this method all of the radiation emitted by the sample enters an energy-sensitive detector, usually a silicon solid-state detector (Fig. 5*b*). Such detectors are operated at liquid-nitrogen temperature to reduce electronic noise and allow an energy resolution of about 150 eV (Fig. 6). This resolution is adequate to distinguish the $K\alpha$ and $K\beta$ lines of a single element, but not adequate to separate the $K\beta$ line of one element from the $K\alpha$ line of the next higher atomic number element (for example, Cr$K\beta$ from Mn$K\alpha$). **Figure 7** compares the wavelength and energy-dispersion spectra from a sample containing a range of atomic number elements. Resolution is better with the crystal spectrometer for most of the spectral range of interest.

In spite of its relatively poor resolution, energy dispersion has become widely accepted because of two advantages it has over wavelength dispersion. First, all the characteristic lines are recorded simultaneously, which makes the method faster than a scanning crystal spectrometer (but not as fast as the multiple crystal-spectrometer instruments). Second, the solid angle of radiation accepted by the detector is 10–100 times greater than the solid angle accepted by the crystal; this allows the primary-source power to be reduced proportionally. With energy dispersion, it is feasible to use electron excitation at beam currents below 10^{-8} ampere in the scanning electron microscope, compared to beam currents of about 10^{-6} A required in electron probes used with crystal spectrometers. Likewise, with energy dispersion, it is feasible to use proton or alpha-particle beams from Van de Graaff or cyclotron accelerators. For proton beams, the proton energy must be 1–5 MeV, compared to electron energies of 10–50 keV for the same x-ray yield; for alpha particles, the energy should be even higher, 10–50 MeV. Compared to photon excitation, positive-ion excitation results in lower background intensity and improves the limit of detection by as much as a factor of 10. On the other hand, direct electron excitation results in much higher background intensity because it is generated in the sample rather than merely scattered by the sample; this degrades the limit of detection by a factor of 10–100.

With energy dispersion, it is even feasible to use some of the radioactive isotope sources as the primary radiation to excite the characteristic x-ray spectra in the sample. There are isotopes such as tritium which undergo beta decay and, when used in a metal matrix, produce continuum and characteristic matrix x-rays; other isotopes such as americium-241 undergo alpha decay and produce gamma rays or x-rays. The advantage of isotope sources is their compact size and the elimination of power supplies and so forth. However, advances in low-power air-cooled x-ray tubes (10–50 W) and compact power supplies have largely eliminated the need for isotope sources.

Types of samples and data interpretation. Two general classes of samples are analyzed routinely by x-ray spectrometry: thin and bulk samples.

Thin samples. Samples smaller than several milligrams per square centimeter such as air-pollution particles collected on a filter, constitute such a thin layer that the primary radiation can penetrate easily and excite emission from each element in proportion to the mass per unit area of that element. Quantitative calibration for each element is accomplished experimentally by determining the sensitivity S in photons per second per microgram per square centimeter for that element. Accuracy (or more properly, precision) depends on the number N of photons counted; the

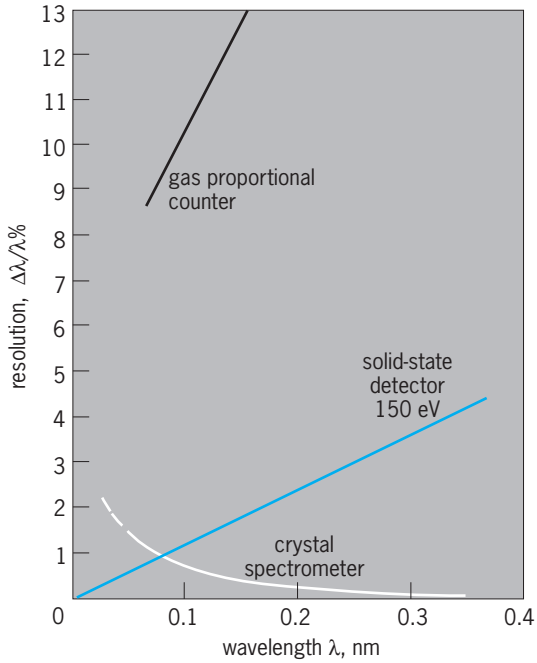

Fig. 6. Resolution of a solid-state detector and a gas proportional counter alone or in conjunction with a crystal spectrometer.

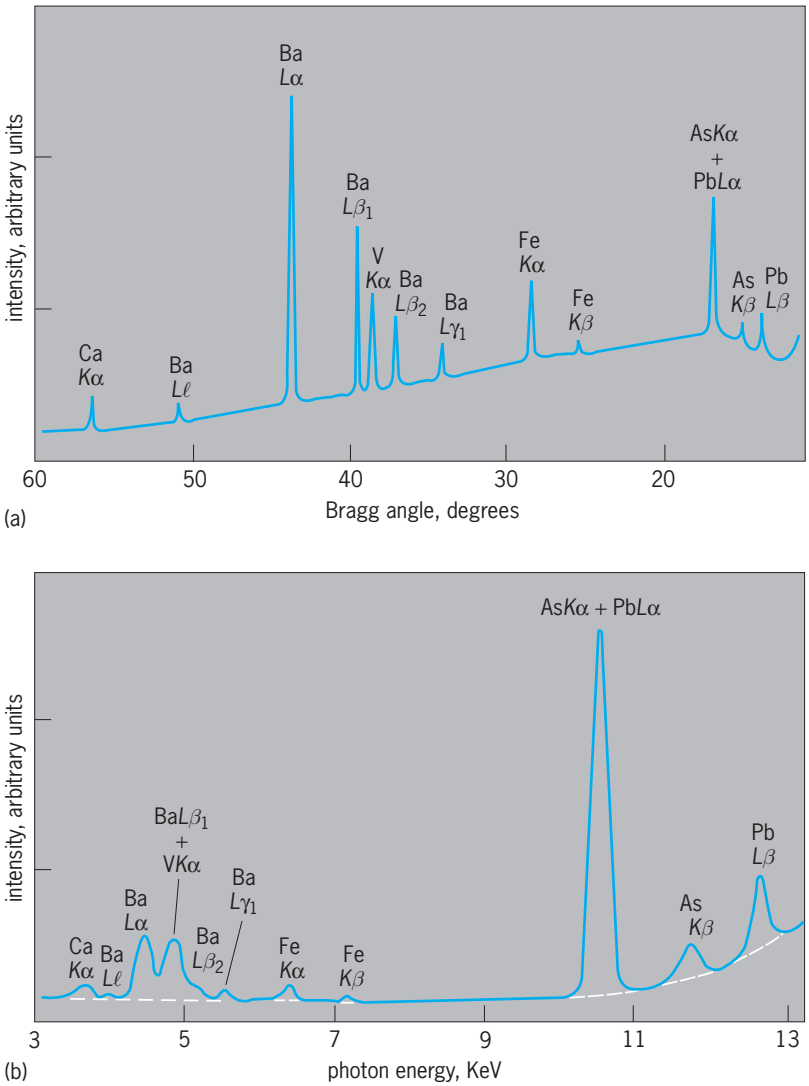

Fig. 7. X-ray spectra of the same sample as measured by (*a*) wavelength dispersion and (*b*) energy dispersion. (*After L. S. Birks, Pinpointing airborne pollutants, Environ. Sci. Tech., 12:150, 1978*)

expected standard deviation σ of a measurement is approximated by $\sigma = \sqrt{N}$. Limit of detection C_L depends not only on sensitivity but on the background intensity as well. If the measured intensity of the background is N_B and that at the line peak is N_p, the limit of detection is defined as the amount of material C_L which gives a signal above background of $3\sigma_B$, that is, $N_p - N_B = 3\sqrt{N_B}$ and $C_L = 3\sqrt{N_B}/(S \times t)$, where t is the counting interval in seconds. Thus it is important to minimize the background intensity N_B by careful instrumental design to eliminate scattering from material other than the sample, and also to minimize the mass per square centimeter of the substrate on which the sample is mounted. With photon excitation and optimized design, the limit of detection varies from about 1 nanogram/cm^2 for elements around atomic number 20 to 10–50 ng/cm^2 for the extremes of high or low Z. The limit is approximately the same for either wavelength or energy dispersion when photon excitation is used. As was stated in the previous portion of this subject area, the limit of detection can be improved by about a factor of 10 with proton excitation, but the analysis is then generally limited to energy dispersion.

Bulk samples. These consist of solids, powder, or liquids and present quite a different problem from thin samples. X-ray absorption limits the penetration of the primary radiation through the sample matrix and the depth from which characteristic radiation may emerge. In addition, the characteristic radiation from some of the elements in the specimen may excite the characteristic radiation of other elements by secondary fluorescence. Matrix absorption and secondary fluorescence depend on the sample composition and determine the x-ray intensity versus composition relationship for each element. Thus, instead of a single, linear calibration as was described for thin samples, the calibration is generally nonlinear, and a family of curves is needed for each element as the matrix composition changes.

If I_{ai} is the intensity from a pure sample of element i, and I_{bi} is the intensity from an unknown concentration of element i in a matrix of other elements, then the most useful parameter is the relative x-ray intensity $R_i = I_{bi}/I_{ai}$. To a first approximation, R_i can be expressed in terms of the concentration C_i and the absorption term μ which incorporates both the absorption of the incident primary radiation and the emerging characteristic radiation, as shown in Eq. (1), where A is a constant which depends on a

$$R_i = \frac{AC_i}{\mu} \tag{1}$$

number of instrumental parameters.

In x-ray spectrometry the analyst can measure R_i but cannot determine C_i directly, because μ depends on C_i and the concentration of each other element C_j as well. However, it is possible to write an equation containing R_i and C_i in terms of individual absorption and secondary-fluorescence effects of each element on the intensity from each other element. The expression is Eq. (2), where Σ is the sum of a

$$\frac{C_i}{R_i} = 1 + \Sigma\alpha_{ij}C_j \tag{2}$$

number of terms, one for each other element in the specimen. The coefficient α_{ij} means the effect on element i by the presence of element j; it is often referred to as an influence coefficient and may be determined experimentally by measuring mixtures of elements i and j. The terms may also be calculated from some of the fundamental properties of atoms and radiation.

There are many variations of the mathematical expressions for determining concentration from x-ray intensity; most of them require computers to evaluate. Whatever the method of data interpretation, a few generalizations may be made about the analysis of bulk specimens. The accuracy (precision) is about 1–2% of the amount of an element present for major constituents, but degrades to 5–10% of the amount present for concentrations to 10–100 ppm. The limit of detection is generally about 1 ppm for

middle-range atomic number elements, but varies from less than 0.1 ppm for metals in biological tissue to 10 ppm or more for low-*Z* elements such as carbon in a middle-*Z* matrix such as steel.

Improvements and limitations. X-ray spectrometry generally does not require any separation of elements before measuring, because the x-ray lines are easily resolved. However, preconcentration methods are sometimes useful as a means for improving the limit of detection. An example is the precipitation or ion-exchange collection of soluble elements in water. Likewise, dilution is sometimes useful to reduce matrix variability or inhomogeneity. An example is the solution of mineral samples in borax.

One limitation of x-ray spectrometry is the progressive difficulty of measurement below atomic number 11. There are several reasons for this, including the strong absorption of such long-wavelength radiation by the x-ray tube and detector windows and also the reduced intensity due to a lower number of x-ray photons emitted per atom ionized (the fluorescent yield factor). Although the elements from boron (5) to fluorine (9) may be measured with specially designed equipment, such measurement cannot be considered routine. In practice, photoelectron spectroscopy and Auger electron spectroscopy are favored for the lower-atomic-number elements, but electron methods may also require special instrumentation and techniques. *See* ELECTRON SPECTROSCOPY; SPECTROSCOPY. L. S. Birks

Bibliography. B. K. Agarwal, *X-ray Spectroscopy*, 2d ed., 1991; E. P. Bertin, *Principles and Practices of X-ray Spectrometric Analysis,* 2d ed., 1975; K. Debertin and R. G. Helmer, *Gamma- and X-ray Spectrometry with Semiconductor Detectors*, 2005; R. E. Van Grieken and A. A. Markowicz (eds.), *Handbook of X-ray Spectrometry Revised and Expanded* (*Practical Spectroscopy, V. 29*), 2d ed., 2001; A. Meisel, *X-ray Spectra and Chemical Binding,* 1989; K. Tsuji, J. Injuk, and R. Van Grieken (eds.), *X-ray Spectrometry: Recent Technological Advances*, 2004.

X-ray telescope

An instrument designed to collect and detect x-rays emitted from a source outside the Earth's atmosphere and to resolve the x-rays into an image. Absorption by the atmosphere requires that x-ray telescopes be carried to high altitudes. Balloons are used for detection systems designed for higher-energy (hard) x-ray observations, whereas rockets and satellites are required for softer x-ray detectors. *See* X-RAY ASTRONOMY; X-RAYS.

Image formation. An image-forming telescope lens for x-ray wavelengths can be based either on the phenomenon of total external reflection at a surface where the index of refraction changes (grazing-incidence telescope) or on the principles of constructive interference (multilayer telescope).

Grazing-incidence telescope. In the case of x-rays, the index of refraction in matter is slightly less than unity. By applications of Snell's law, the condition for total external reflection is that the radiation be incident at small grazing angles, less than a critical angle of about 1°, to the reflecting surface. The value of the critical angle depends on the wavelength of the radiation and the material used. As the wavelength decreases (higher energy), the grazing angle required is smaller; as the atomic number (*Z*) of the material increases, the grazing angle for a given wavelength increases. The detailed reflectivity as a function of energy is complicated by x-ray absorption edges in the material, as well as the density of the material and its surface properties. *See* REFLECTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES.

Based on these properties, x-ray mirrors have been constructed which focus an image in two dimensions. Various configurations of surfaces are possible. In the Wolter type I x-ray telescope (**Fig. 1**), which has been built and used successfully for x-ray astronomy on the *Einstein* and *ROSAT* observatories, two surfaces of revolution are used to produce a high-quality image. These mirrors are manufactured from fused quartz, which is coated with nickel after being shaped and polished. The surfaces are extremely smooth in order to focus well, and the shape of the mirror must also be within narrow tolerances. The *Einstein Observatory* mirrors were the first ever built for nonsolar x-ray imaging. They were four nested pairs ranging about 2–1 ft (0.6–0.3 m) in diameter and 4 ft (1.2 m) in length, and the total area of the polished surface was about the same as that of the 200-in. (5-m) telescope at Palomar Mountain; yet the effective area of the telescope was only about 150 in.^{2} (1000 cm^2), because of the grazing-angle geometry.

Both the *Einstein* and *ROSAT* observatories were limited by their optical designs to very low energies. For *Einstein* the telescope cut off at energies above 4.5 keV; that is, x-rays of higher energy were not reflected because their grazing angle (set by the shape of the telescope and its focal length) was too large. For *ROSAT* the high-energy cutoff occurred at 2.5 keV. Subsequent telescopes had higher cutoffs: 12 keV for ASCA, 10 keV for *BeppoSAX* and *Chandra*, and 15 keV for *XMM-Newton*.

Multilayer telescope. A second type of x-ray telescope is based on the principles of constructive

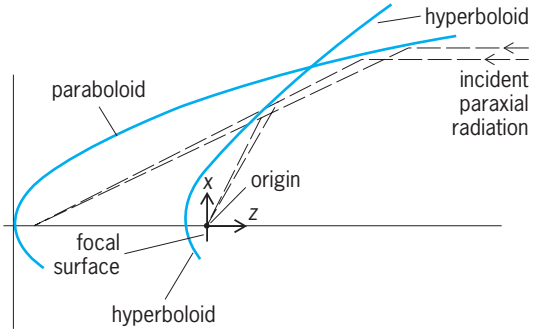

Fig. 1. Wolter type I x-ray telescope. Two surfaces of revolution (paraboloid and hyperboloid) reflect the incident x-rays to a common focus. The horizontal scale is greatly compressed. The *y* axis is perpendicular to the plane of the figure.

interference in extremely thin layers of material deposited on a mirror surface. Unlike the grazing-incidence telescope, multilayer telescopes do not require the x-rays to strike at shallow angles in order to be reflected. Instead these mirrors are similar to normal optical telescope mirrors where the incoming radiation strikes the mirror at nearly normal incidence to be reflected and focused. *See* INTERFERENCE OF WAVES.

The multilayers are coatings of specially selected materials that have crystal structures of regularly spaced atoms. They are evaporated onto a mirror surface that has been highly polished so that it is the correct shape to focus x-rays and is so smooth that the average bumpiness on the mirror surface is comparable to the wavelength of the x-rays being reflected. This is the same quality of surface that is required for grazing-incidence mirrors, but is somewhat easier to achieve when the surface being polished is nearly flat. Several layers of coatings are used in order to achieve high reflectivity, and each layer must be of a precise thickness and composition to work properly. These coatings are deposited in vacuum to achieve the desired high degree of cleanliness and purity of the multilayers if they are to work efficiently. *See* ARTIFICIALLY LAYERED STRUCTURES.

The advantage of a multilayer mirror is that all of the mirror area is used in collecting the x-ray radiation, whereas the grazing-incidence telescopes have only a small projected area of the actual mirror surface collecting radiation. There is an offsetting disadvantage, as the multilayer mirror reflects x-rays only within a very narrow range of energy while the grazing-incidence mirror reflects over a broad range of energies. The effect is similar to using a narrow-band filter with an optical telescope, and in some cases this can be very useful.

The first astronomical use of multilayer, normal-incidence x-ray telescopes was in photographing the Sun. Images of the solar surface of very high angular resolution were made in a selected wavelength interval near 6.7 nanometers. They showed that the filamentary structure of solar flares extends down to as small a size as can be resolved in the pictures, a result that has direct consequences for the theoretical work being done on understanding these flares. The Sun is an excellent object for observation with multilayer telescopes because it is so bright (being nearby) that there is enough flux for detection over even a narrow energy band.

Image detection. The telescope mirrors focus x-rays, producing an image in two dimensions in the same manner as the lenses of optical telescopes result in images of the sky. Suitable devices are required to detect and record these images, completing the functional requirements of an astronomical telescope. Various types of x-ray detectors have been developed for this purpose. These are position-sensitive devices which in effect are electronic cameras suitable for x-ray wavelengths.

Microchannel plates. The high angular resolution of the grazing-incidence telescope requires a camera that has correspondingly good spatial resolution. One type of detector uses microchannel plates (MCP; **Fig. 2**) and yields about 20 micrometers resolution for x-rays in the soft energy band about 100–10,000 eV). The microchannel plate is an array of small hollow tubes or channels (about 10–15 μm in diameter) which are processed to have high secondary electron yield from their inner walls. A single x-ray striking the surface of a channel may produce a free electron. The electric field produced by placing a high voltage across the microchannel plate accelerates this electron, which collides with the wall of the tube to produce more electrons. This results in a cascade of electrons through the channel, multiplying in number until a sufficient signal is produced to be recorded electronically, giving the location of the event. *See* IMAGE TUBE (ASTRONOMY); PHOTOMULTIPLIER.

Charge-coupled devices. Another type of detector developed for x-ray imaging applications is the charge-coupled device (CCD). This is a solid-state detector that consists of microscopic silicon picture elements (pixels) in which electronic charges produced by the passage of an x-ray photon are collected. The charge-coupled device is periodically read out by shifting the collected charges from one pixel to the next, as in a bucket brigade. The charge is detected in a sensitive amplifier, and the corresponding position is determined by keeping count of the number of charge transfers that were made. Typical charge-coupled devices have pixels that are 15–25 μm on a side, and there are up to 4096 × 4096 pixels on one such device. The charge-coupled device not only records the position of an event but can also yield information on the energy of the photon. Thus this type of detector is useful as an imaging spectrometer, providing a spectrally resolved image of an x-ray source

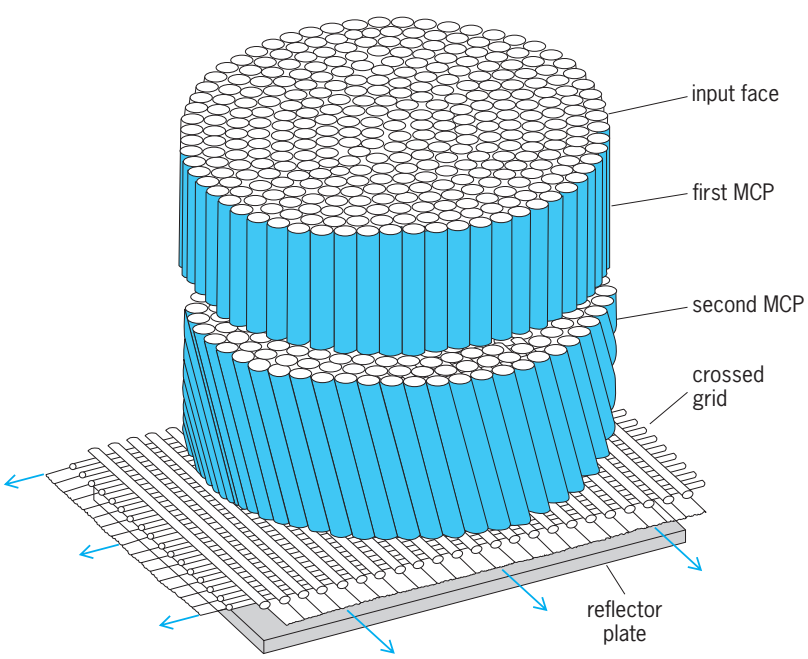

Fig. 2. Blow-up (not to scale) of the major components of a high-resolution imaging detector for an x-ray telescope. Microchannel plates (MCP) detect incident radiation and amplify the signal for position determination by the grid of crossed wires below.

which can be analyzed for the energy distribution of x-rays in different parts of the image. Such devices are needed to study supernova remnants and other extended sources in order to understand the conditions within such objects and how they vary. *See* CHARGE-COUPLED DEVICES.

Position-sensitive proportional counters. Yet another type of detector used with x-ray telescopes is a gas-filled counter in which x-rays are photoelectrically absorbed, yielding an electron which is detected by the ionization it produces in the gas. By operation of the counter in the proportional mode and use of planes of wires to localize the electrical signals, the position and amplitude of each event can be recorded. These detectors generally have lower spatial resolution (about 200 μm) than do microchannel plate or charge-coupled-device detectors, but they can be made larger in size and provide better energy resolution than microchannel plate detectors. Position-sensitive proportional counters were used in the *Einstein Observatory* and in *ROSAT*. *See* IONIZATION CHAMBER.

Cryogenic detectors. Another class of detectors consists of very low temperature devices that detect the heat deposited in an absorber when an x-ray is stopped. These devices promise improved energy resolution and efficiency over a broad range of energies. They all work at temperatures close to absolute zero, so that keeping the detector cold is a major problem. *See* TELESCOPE. Stephen S. Murray

Bibliography. G. Burbidge and A. Hewitt, *Telescopes for the '80s*, 1981; G. W. Fraser, *X-ray Detectors in Astronomy*, 1989; L. Golub (ed.), *X-ray Instrumentation in Astronomy II*, 1989; R. B. Hoover and A. B. Walker (eds.), *X-ray Optics, Instruments, and Missions II*, 1999; P. Joss (ed.), *High-Energy Astrophysics in the Twenty-First Century*, 1990; O. H. Siegmund and K. A. Flanagan (eds.), *EUV, X-ray, and Gamma-ray Instrumentation for Astronomy*, vol. 10, 1999.

X-ray tube

An electronic device used for the generation of x-rays. X-rays are produced when electrons that have been accelerated by a large potential are stopped in a target, which is usually made of metal. The x-rays radiate in all directions from the spot on the target where the electrons hit. X-rays are produced by two mechanisms in a metal target: bremsstrahlung (or braking radiation in German), which produces a continuous energy spectrum of x-rays; and characteristic x-rays, which are produced by the creation and subsequent filling of a vacancy in an inner shell of the target atoms (**Fig. 1**). *See* BREMSSTRAHLUNG; X-RAYS.

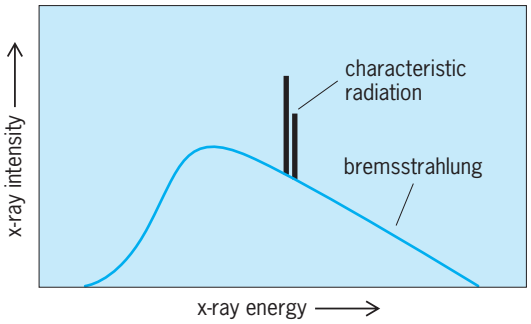

Fig. 1. X-ray spectrum from an x-ray tube.

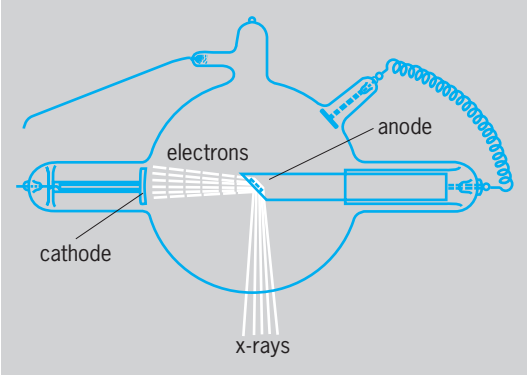

Fig. 2. Commercial model of gas x-ray tube.

Gas x-ray tubes. In gas tubes, electrons are freed from a cold cathode by positive ion bombardment. For the existence of the positive ions a certain gas pressure is required without which the tube will allow no current to pass. In the earliest gas tube the electrons liberated by positive ion bombardment from a flat aluminum cathode were emitted in a direction perpendicular to the cathode surface. The electrons traveled in straight lines until they impinged upon the glass end wall of the tube, where x-rays were generated.

Many designs of gas tubes have been built for useful application, particularly in the medical field. Metals, such as platinum and tungsten, have been placed in the path of the electron beam to replace glass as the target. Concave metal cathodes are used to focus the electrons on a small area of the metal target and increase the sharpness of the resulting shadows on the fluorescent screen or the photographic film. **Figure 2** shows one form of commercial gas x-ray tube.

The size of the bulb in commercial tubes was increased as the power input was increased, to reduce the pressure change in operation and to reduce the local heating of the glass envelope resulting from bombardment of electrons reflected from the focal spot. The aluminum cathodes were made heavier to withstand the increased positive-ion bombardment. The thin metal targets were also replaced by a heavier mass of metal. Targets now consist of two main parts: a refractory metal face, such as platinum, to take the direct impact of the electron beam, and a heavy backplate of a good heat-conducting material, such as copper, to conduct the heat away from the focal spot and store it temporarily.

Early in the development of x-ray tubes metals of high atomic weight were known to be the most efficient x-ray generators. W. C. Roentgen had used both aluminum and platinum as targets. W. D. Coolidge

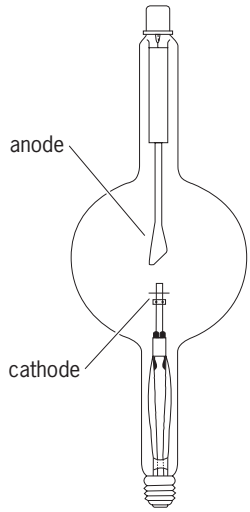

Fig. 3. Early commercial model of single-section, hot-cathode, high-vacuum x-ray tube.

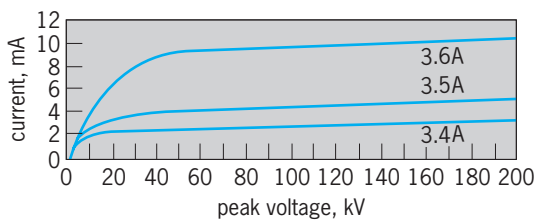

Fig. 4. Curves showing relation of current to voltage in hot-cathode, high-vacuum x-ray tube. Filament currents of the three curves are in amperes.

and others later used thorium and uranium with increased efficiency. The principal properties desired for a suitable target material are (1) high atomic number to give best x-ray production efficiency, (2) high melting point and high thermal conductivity to permit maximum electron power to be put into the focal spot, and (3) low vapor pressure to reduce the rate of evaporation of the metal on the walls of the glass envelope. Ductile tungsten as developed by Coolidge was found to combine these desired properties of an x-ray target to the greatest degree.

High-vacuum x-ray tubes. The operational difficulties and erratic behavior of gas x-ray tubes are inherently associated with the gas itself and the positive ion bombardment that takes place during operation. The high-vacuum x-ray tube eliminates these difficulties by using other means of emitting electrons from the cathode.

Coolidge made a high-vacuum x-ray tube with a hot tungsten-filament cathode and a solid tungsten target. This hot-cathode high-vacuum type of x-ray tube permitted stable and reproducible operation with relatively high voltages and large masses of metals. The vacuum was so good that positive ions did not play either an essential or a harmful role in the tube operation. The earliest commercial form of the hot-cathode high-vacuum type of x-ray tube (**Fig. 3**) utilizes a solid tungsten target. The independent relation of the current to the impressed voltage in a tube of this type is shown in **Fig. 4**. The different curves are for different filament temperatures and show that, over the operating range of x-ray voltages, the discharge current is practically independent of voltage. One typical form of a modern commercial hot-cathode high-vacuum x-ray tube (**Fig. 5**) is built with a liquid-cooled, copper-backed tungsten target, which operates over a wide range of energy ratings and is capable of rectifying its own current.

The inherent advantages of the hot-cathode high-vacuum tube over the gas tube are (1) flexibility—the voltage and current may be varied independently; (2) stability—this permits more accurate reproducibility of results; (3) small size; (4) operation—it can be operated directly from a transformer, making possible a very simple unit; and (5) long life.

High-voltage x-ray tubes. The operating voltage of a single-section hot-cathode x-ray tube is limited by the field current emitted from the cathode by the high electrostatic field. Field emission may cause erratic and uncontrollable operating behavior of an x-ray tube. Single-section x-ray tubes must preferably be operated at voltages below which field currents cannot be produced. *See* FIELD EMISSION.

There are, however, some highly specialized applications of x-ray tubes for which the tube is designed specifically to produce and use field emission currents to generate x-rays. To build x-ray tubes that operate stably and continuously at very high voltages, the multisection principle is usually employed. These tubes are made with many intermediate sections between the cathode and anode sections. The voltage applied across each section of this multisection tube is always less than that at which field currents will be produced. The electrons emitted from the cathode are accelerated by the voltage applied across each intermediate section. The sum of all of these sectional voltages determines the voltage rating of the multisection x-ray tube. By this procedure, x-ray tubes can be built to generate x-rays at several millionelectronvolts energy. **Figure 6** is a commercial 2-MeV x-ray tube built according to this principle.

The toroidal electromagnetic type of hot-cathode high-vacuum x-ray tube, first successfully built by D. W. Kerst, permits the generation of x-rays at energy levels of many megaelectronvolts without encountering the high-voltage insulation problem that

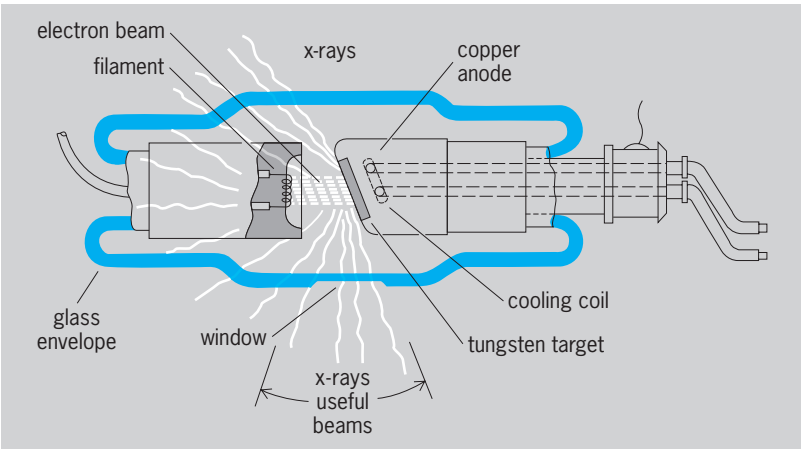

Fig. 5. Single-section, hot-cathode, high-vacuum x-ray tube.

is present in more conventional types of x-ray tubes. This x-ray tube consists of a toroidal high-vacuum envelope housing the hot-cathode and the x-ray target. This tube is designed to guide and accelerate electrons in a circular orbit in a machine called the betatron. In this toroidal-shaped betatron tube (**Fig. 7**) the electrons are injected and accelerated to an energy of many millions of electronvolts before they strike the target to produce x-rays. The average electron beam current is very small, but x-ray tubes of this type have been built to generate x-rays at energies ranging from 1 to over 100 MeV.

A synchrotron is a particle accelerator which can accelerate and store electrons at very high energies. X-rays are emitted when the direction of motion of relativistic stored electrons is changed, as by bending in a magnetic field. Synchrotron storage rings produce exceptionally bright beams of x-rays over a wide spectrum. *See* PARTICLE ACCELERATOR; SYNCHROTRON RADIATION.

The multisection linear accelerator type of x-ray tube is another hot-cathode high-vacuum device that is capable of generating intense x-rays over a wide range of energy levels from 1 to many megaelectronvolts. It consists (**Fig. 8**) of an electron gun and a copper pipe waveguide in which the electrons are accelerated by the axial electric field produced by a high-frequency oscillator. At the output end of the waveguide, the electrons strike a target and generate x-rays. X-ray tubes of this type have been built for operation at 5–15 MeV and higher, and with average currents of the order of 25 microamperes or more.

Applications. Many special forms of x-ray tubes, following the basic patterns already outlined, have been built for application in medicine, industry, and fundamental science. They vary from a few centimeters to many meters in length. X-ray tubes are used as an aid in medical diagnosis and in therapeutic treatment. They are used in the nondestructive testing of materials throughout industry. In both fundamental and applied science they are widely used in crystal-diffraction work, chemical analysis by x-ray spectra and absorption, and research on atomic structure. *See* MEDICAL IMAGING; NONDESTRUCTIVE EVALUATION; RADIOGRAPHY; RADIOLOGY; X-RAY CRYSTALLOGRAPHY; X-RAY DIFFRACTION; X-RAY FLUORESCENCE ANALYSIS; X-RAY POWDER METHODS.

To provide maximum use in these wide fields of application, particular design features are built into the tubes to meet the special requirements. Some of these special requirements are low or high currents, low- or high-voltage x-rays, large or small focal spot, high x-ray beam intensity for short or long intervals, and choice of x-ray target material to generate a particular quality of x-ray spectra. X-ray and electrical protection must be ensured.

The overall capacity of an x-ray tube is determined principally by the target material, the area of the focal spot, the duration of the energy applied, and the temperature of the target during the time the electron energy is applied. The modern x-ray tube is a precision tool of great stability and flexibility, capable of controlled operation with currents and voltages of any desired magnitude, and generating x-rays over a wide range of energy and intensity. Ernest E. Charlton

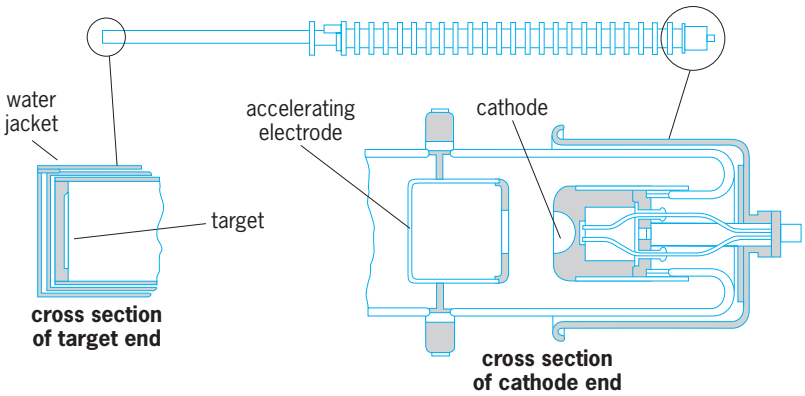

Fig. 6. A 2-MeV multisection, hot-cathode, high-vacuum x-ray tube.

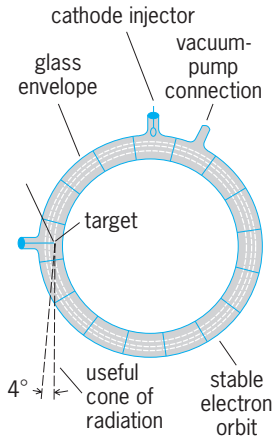

Fig. 7. Betatron "doughnut" high-vacuum x-ray tube. (*General Electric Co.*)

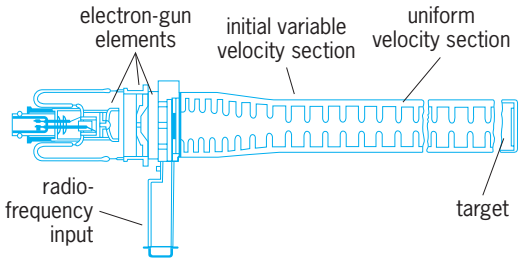

Fig. 8. Multisection linear accelerator type of hot-cathode, high-vacuum x-ray tube.

Bibliography. M. Y. Chen, T. L. Pope, and D. J. Ott, *Basic Radiology*, 2d ed., McGraw-Hill, 2004; R. Halmshaw, *Industrial Radiology: Theory and Practice*, 2d ed., Springer, 1995; J. C. Malott and J. Fodor, *The Art and Science of Medical Radiography*, 7th ed., Mosby-Year Book, 1993; K. N. Prasad, *Handbook of Radiology*, 2d ed., CRC Press, 1995; J. Selman, *Fundamentals of X-ray and Radium Physics*, 8th ed., C. C. Thomas, 1994.

X-rays

X-rays, or Roentgen rays, are electromagnetic waves; they are the same as visible light, except that they have shorter wavelengths (higher photon energies). Thus x-rays, visible light, ultraviolet, infrared, microwaves, and radio waves are all electromagnetic radiation in different wavelength (energy) spectral regions. X-rays are generated when fast-moving electrons slow down and stop in matter, when an inner-shell vacancy in an atom is filled by another electron, and when electrons moving at relativistic speeds (speeds near the speed of light) change their direction of motion in space. *See* ELECTROMAGNETIC RADIATION.

Roentgen's findings. X-rays were discovered by W. C. Roentgen in 1895. This discovery came about by accident. Roentgen was studying gas discharges in his laboratory when he noticed that unknown radiation from a gas discharge could induce fluorescence in certain materials. In his first communication, Roentgen described the properties of these rays as follows: They were invisible; moved in straight lines; were unaffected by electric or magnetic fields, and hence not electrically charged; passed through matter opaque to ordinary light (since they penetrated through the black cardboard around his cathode-ray tube); were differentially absorbed by matter of different densities or of different atomic weights; affected photographic plates; produced fluorescence in certain chemicals, such as in the barium platinocyanide screen with which the initial discovery was made and in the wall of his glass tube opposite the cathode; produced ionization in gases; and were evidently produced at the anode by the beam of rays (identified by J. J. Thomson in 1897 as electrons) issuing from the cathode in his vacuum tube. *See* CATHODE RAYS; FLUORESCENCE.

Along with all these definitive characteristics of the rays, however, other crucial experiments designed to establish similarity or differences from visible light were clearly called for. The fundamental optical properties of visible light were well established in 1895: reflection from mirrors; refraction in prisms (change in direction in passing from air into glass, for example), by means of which a beam of white light could be spread out into a rainbow or spectrum of colors; diffraction by narrow slits or ruled gratings, also a method of producing spectra; and polarization, or constraint of the electric field of the light wave to a single direction. In spite of the best efforts of Roentgen, no evidence of any of these four optical phenomena could be found. Hence the designation "x"—unknown—was assigned by Roentgen. Many theories were proposed to account for the apparently unique quality of x-rays, which seemed to be so closely similar and yet so greatly different from visible light. *See* DIFFRACTION; LIGHT; POLARIZATION OF WAVES; REFLECTION OF ELECTROMAGNETIC RADIATION; REFRACTION OF WAVES.

Later discoveries. Other scientists studying x-rays found the essential experimental conditions to prove that x-rays can be polarized (C. Barkla, 1905, by scattering from carbon); diffracted by crystals (M. von Laue, W. Friedrich, and P. Knipping, 1912); refracted in prisms and in crystals; reflected by mirrors; and diffracted by ruled gratings (A. Compton, 1921–1922). Instead of being refracted in passing from a less dense medium (air) to a more dense medium (a glass prism or a crystal) in the same direction as light (the index of refraction for visible light is always greater than 1), x-rays are deviated in the opposite direction by a very small amount: the index of refraction is less than 1 by an amount as small as 10^{-6}. Total reflection from mirrors is observed only when the beam impinges at a very small angle (grazing angle), a necessary condition understandably missed by Roentgen. Similarly, the beam must be incident at a very small angle on a ruled diffraction grating if diffraction (a spectrum, or spatial dispersion by wavelength) is to be observed. *See* X-RAY OPTICS.

From 1895 to 1912 there seemed to be no analyzer capable of dispersing an x-ray beam into a spectrum. The spectacular Laue diffraction pattern of a zinc sulfide crystal in 1912 proved the electromagnetic wave nature of x-rays and the ordered structure of crystals, with atoms lying on families of planes to constitute three-dimensional diffraction gratings, all governed by the simple Bragg law $n\lambda = 2d \sin \theta$ (which must be corrected for refraction in extremely accurate work). Here n is an integer indicating the order of the spectrum, λ the wavelength, d the crystal lattice spacing of one set of planes, and θ the angle between the incident ray and this set of planes. *See* CRYSTAL STRUCTURE; X-RAY DIFFRACTION.

The wavelength range of x-rays in the electromagnetic spectrum, as excited in x-ray tubes by the bombardment of a target by electrons under a high accelerating potential, overlaps the ultraviolet range on the order of 100 nanometers on the long-wavelength side, and the shortest-wavelength limit moves downward as voltages increase. An accelerating potential of 10^9 V, now readily generated, produces a λ of 10^{-6} nm. An average wavelength used in research is 0.1 nm, or about 1/6000 the wavelength of yellow light. *See* ULTRAVIOLET RADIATION; X-RAY TUBE.

Quantum mechanics. X-rays (and visible light) can be considered as an electromagnetic wave. X-rays can also be considered as discontinuous bundles of energy, or quanta, in accordance with the laws first enunciated by M. Planck and extended by A. Einstein early in the twentieth century. In diffraction, refraction, polarization, and interference phenomena, x-rays, together with all other electromagnetic radiation, appear to act as waves and λ has a real significance. There is duality, meaning that light and x-rays have both wave and particle properties, although these are generally not observed at the same time in a given experiment. Beams of electrons and neutrons also have wave properties, and are diffracted in appropriate media. In other phenomena—such as the appearance of sharp spectral lines, a definite short-wavelength limit λ_0 of the continuous "white" spectrum [defined by $\lambda_0 = hc/eV$, where h is Planck's constant, c the velocity of electromagnetic radiation (including light and x-rays), e the charge of electron,

and V the accelerating voltage], the shift in wavelength of x-rays scattered by electrons in atoms (Compton effect), and the photoelectric effect—the energy is propagated and transferred in quanta (called photons) defined by values of $h\nu$, where the frequency ν is c/λ. *See* COMPTON EFFECT; ELECTRON DIFFRACTION; NEUTRON DIFFRACTION; PHOTOEMISSION; PHOTON; QUANTUM MECHANICS.

Applications. X-rays are a valuable probe of matter, as they interact selectively with electrons; electrons in matter account for most of the significant properties of matter (excepting nuclear properties). Medical uses abound, as x-rays can penetrate matter, and are selectively absorbed by atoms containing many electrons—elements with a high atomic number. Thus dental x-rays show fillings as dark, teeth and bone in grey, and are only lightly absorbed in the cheek. X-rays in high intensities are also used for therapy, to treat certain cancers. X-rays are also used in many industrial processes, as in checking the integrity of welds. Tomographic techniques allow three-dimensional images to be obtained, which are invaluable in medical diagnosis.

X-rays are used in research for the study of both the electronic structure of matter and its spatial structure. X-ray microscopes and microprobes have been developed which allow imaging with high spatial resolution while obtaining contrast with different elemental discrimination, distinguishing chemical bonds, and, by use of circularly polarized x-rays, the orientation of magnetization. X-rays are used in protein crystallography to determine the spatial structure of proteins, viruses, and other objects that can be made into crystals but that otherwise cannot be seen. *See* COMPUTERIZED TOMOGRAPHY; HISTORADIOGRAPHY; MEDICAL IMAGING; MICRORADIOGRAPHY; NONDESTRUCTIVE EVALUATION; RADIOGRAPHY; RADIOLOGY; X-RAY CRYSTALLOGRAPHY; X-RAY FLUORESCENCE ANALYSIS; X-RAY MICROSCOPE; X-RAY POWDER METHODS. George L. Clark; Alfred S. Schlachter

Bibliography. J. G. Brown, *X-rays and Their Applications*, Plenum, 1975; J. E. Cullinan and A. M. Cullinan, *Illustrated Guide to X-ray Technics*, 2d ed., Lippincott, 1980; N. A. Dyson, *X-rays in Atomic and Nuclear Physics*, 2d ed., Cambridge University Press, 1990; H. Haken and H. C. Wolf, *The Physics of Atoms and Quanta: Introduction to Experiments and Theory*, 7th ed., Springer, 2005; J. Selman, *Fundamentals of X-ray and Radium Physics*, 8th ed., C. C. Thomas, 1994.

Xanthophyceae

A class of plants, comprising the yellow-green algae, in the chlorophyll *a–c* phylectic line (Chromophycota). Alternate names are Tribophyceae, derived from *Tribonema*, a filamentous member of the class, and Heterokontae, referring to the presence of two kinds of flagella on each motile cell. *See* ALGAE; CHROMOPHYCOTA.

Characteristics. The primary photosynthetic pigments are chlorophyll *a* and *c*. Carotenoids include β-carotene and several xanthophylls, among which heteroxanthin and vaucheriaxanthin ester are unique to the class, while diatoxanthin and diadinoxanthin are found also in diatoms and dinoflagellates. As in other Chromophycota, the chloroplasts of yellow-green algae are bounded by a double-membrane envelope surrounded by a chloroplast endoplasmic reticulum, the outer membrane of which is continuous with the outer membrane of the nuclear envelope. Each photosynthetic lamella comprises three thylakoids. Some chloroplasts have a multilamellar internal pyrenoid. Nonpigmented xanthophytes are rare.

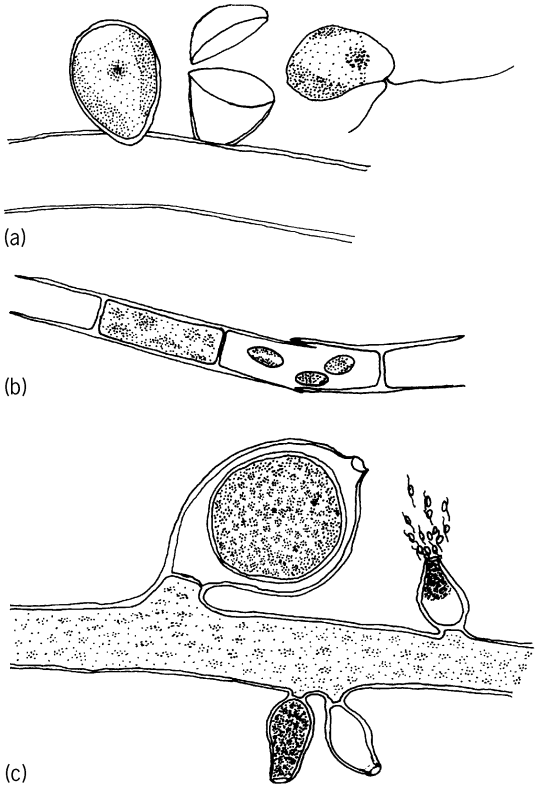

Xanthophytes. (*a*) *Lutherella*, a coccoid xanthophyte epiphytic on another alga, showing vegetative cell, an empty two-piece cell wall, and a zoospore. (*b*) *Tribonema*, a filamentous xanthophyte, showing immature zoospores and H-shaped wall segments. (*c*) *Vaucheria*, a siphonous xanthophyte; monoecious species with mature oogonium and three stages of development of antheridium, the mature one releasing antherozoids.

The cell wall is often visibly lamellate, but its chemical composition is poorly known. Cellulose, pectic compounds, and silica are often suggested. Cellulose is known for certain to be the chief component in *Vaucheria*. In many filamentous and some large coccoid forms, the wall of each cell is composed of two overlapping pieces (in filaments, H-shaped in optical section) [**illus.** *a* and *b*]. Endogenous cysts (statospores) also have walls with two equal or unequal pieces. The chemical composition of the food reserves is also poorly known, but leucosin (a β-1,3 glucan) and lipids are usually reported. The absence of starch is an important characteristic in distinguishing yellow-green algae from green algae, which they may resemble superficially in form and color.

Motile cells have two unequal flagella borne apically or laterally (illus. *a*). One (usually the longer) is directed forward and is pleuronematic (hairy); the other is directed backward and is acronematic (smooth). The compound zoospore (synzoospore) of *Vaucheria* is exceptional in having numerous pairs of slightly unequal smooth flagella. There are usually two chloroplasts in each motile cell, the ventral one with a reddish refractile eyespot at its anterior end. One or two contractile vacuoles occur at the anterior end of a motile cell.

Cells in most Xanthophyceae are uninucleate, but in some genera they are multinucleate. The mitochondria have tubular cristae. Each vegetative cell contains one to many yellow-green, laminate or discoid, parietal chloroplasts.

Reproduction. Formation of autospores (nonmotile spores that are miniatures of the cells that produce them), and less often zoospores, is the most common reproductive process. Statospores (internally formed resting stages) are formed by some genera. These germinate to form zoospores or ameboid cells. Sexual reproduction is rarely observed except in *Vaucheria*, which is oogamous. Isogamy and anisogamy are known for a very few genera.

Classification. About 600 species in 115 genera are recognized and distributed among six orders based on the form of the thallus. By far the largest order is Mischococcales, which comprises nonmotile coccoid unicells (illus. *a*). Families and genera within this order are distinguished primarily on the basis of habit (whether the cells are solitary or colonial, attached or free-living, sessile or stipe-borne on a supporting stalk, and with or without mucilage), cell shape, and cell wall ornamentation. Filamentous yellow-green algae constitute the order Tribonematales. *Tribonema* (illus. *b*) is a common inhabitant of standing waters, especially in spring.

Siphonous yellow-greens are placed in the order Vaucheriales. *Vaucheria* (illus. *c*) is the best-known xanthophyte, commonly forming extensive growths on moist soil or mud. Its thallus consists of branched coenocytic filaments with numerous nuclei and chloroplasts concentrated in a peripheral cytoplasmic layer. Septa are formed only to set apart the reproductive structures. Each oogonium contains an egg, while an antheridium produces numerous small colorless antherozoids. Each antherozoid has a pair of laterally inserted flagella, the anteriorly directed one being the shorter, contrary to the usual situation in xanthophyte motile cells. An antherozoid swims through a pore in the oogonial wall to fertilize the egg. The resulting zygote secretes a thick wall and remains in place as a zygospore until the oogonial wall decays. Following a period of dormancy, the zygospore germinates meiotically to produce haploid filaments. Asexual reproduction is effected by multinucleate structures set apart by septa. In terrestrial forms these structures are released as aplanospores, while in aquatic forms they develop a clothing of flagella before being released as compound zoospores.

The orders Chloramoebales, Rhizocloridales, and Heterogloeales comprise flagellate unicells (monads), pseudopodial forms, and palmelloid forms, respectively.

Distribution. Most Xanthophyceae occur in freshwater, especially soft water. Frequently, they are epiphytic on aquatic plants. The chief exception is *Vaucheria*, which has freshwater, brackish-water, and marine species. Xanthophyceae also occur in and on soil and mud, in snow and ice, and on tree trunks and damp walls. Paul C. Silva; Richard L. Moe

Bibliography. E. R. Cox (ed.), *Phytoflagellates*, pp. 243–271, 1980; H. Ettl, J. Gerloff, and H. Heynig (eds.), *Süsswasserflora von Mitteleuropa*, vols. 3 and 4, 1978.

Xenacanthidae

A family of Paleozoic elasmobranch sharks that are characterized by teeth with large divergent lateral cusps (diplodont crowns), an elongate dorsal fin preceded by a distinctive spine, and paired fins with an unusually well-developed skeletal axis. The xenacanths are one of the most clearly defined and stable groups to emerge from the early radiation of cartilaginous fishes (Chondrichthyes), with body fossils known from the early Carboniferous (*Diplodoselache*) through the late Triassic (*Xenacanthus*) [see **illustration**]. Numerous articulated specimens have been discovered, predominantly from the upper Carboniferous and Permian of Europe and North America, wherein xenacanths are a classic feature of faunas associated with brackish and freshwater lagoons and coal swamps. These xenacanths include the most frequently depicted genera, such as *Orthacanthus* and *Pleuracanthus*, with eellike body forms of up to 3 m (10 ft) or more in length.

Skeletal anatomy. The exceptional abundance and quality of xenacanth fossil material, relative to other Paleozoic chondrichthyans, has allowed their skeletal anatomy to be studied in unusual detail. This has yielded numerous shared, specialized features that ally this group with the lineage leading to modern sharks and rays rather than ratfish (chimaeroids). For a long time these were the only early sharks in which the braincase anatomy was known in any significant detail; likewise for the complex articulations between the jaws and braincase, which are far less mobile than those of modern elasmobranchs. The skeletons of the paired fins in more derived (evolutionarily advanced) members of the Xenacanthidae are noted for their convergent resemblance to the paired fins of lungfish (Dipnoi). In both examples, an elongate, segmented axial rod extends to the fin extremity, with pre- and postaxial branches of the fin skeleton creating a treelike pattern.

Diplodont teeth. The fossil record of xenacanth teeth exceeds that of the less resilient cartilaginous skeleton. Diplodont teeth are among the very earliest examples of sharklike teeth known, and include *Leonodus* from the Early Devonian of Europe. However, the attribution of these teeth to xenacanths is

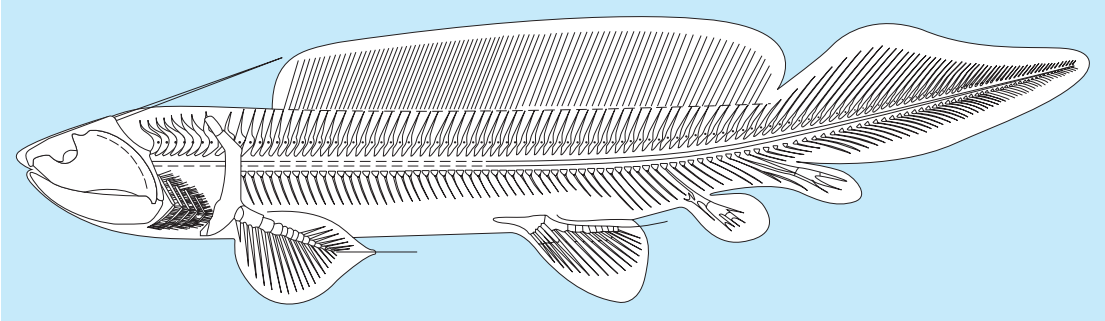

***Xenacanthus meisenheimensis*, from the Permo-Carboniferous Saar-Nahe Basin, southwestern Germany. (*After U. H. J. Heidtke, 2003*)**

uncertain; it is likely that simple, twin-cusped teeth have a wide distribution among primitive sharklike fishes, extending well beyond the boundaries of the Xenacanthidae. *See* CHONDRICHTHYES; ELASMOBRANCHII. Michael Coates

Bibliography. J. A. Long, *The Rise of Fishes*, Johns Hopkins University Press, Baltimore, 1995; J. G. Maisey, *Discovering Fossil Fishes*, Westview Press, New York, 1996.

Xenolith

A rock fragment enclosed in another rock, and of varying degrees of foreigness. Cognate xenoliths, for example, are pieces of rock that are genetically related to the host rock that contains them, such as pieces of a border zone in the interior of the same body. Included blocks of unrelated rocks are more deserving of the xenolith label. Such foreign rocks help establish the once fluid and hence molten condition of invading magma capable of incorporating and mixing an assemblage of unrelated rock inclusions, as at Hutton's Rock near Edinburgh, Scotland. Dark inclusions, commony called enclaves, up to many feet in length can be found in granitic plutons and silicic lava domes and flows. In many cases, these have fine-grained borders against the enclosing rock and are interpreted to be lenses of formerly fluid basaltic magma which crystallized as a result of contact with cooler surrounding silicic magma.

Xenoliths tend to react with the enclosing magma, so that their constituent minerals become like those in equilibrium with the melt. Reaction is rarely complete, however. Even completely equilibrated xenoliths may be conspicuous because the equilibration process does not require either the texture or the proportions of the minerals in the xenolith to match those in the enclosing rock. *See* PLUTON.

Xenoliths may be angular to round, millimeters to meters in diameter, aligned or haphazard, and sharply or gradationally bounded (**Figs. 1** and **2**). Xenoliths are present in most bodies of igneous rock, although they are rare in many igneous rocks. Many rocks are inferred to be igneous because they contain xenoliths, the incorporation of which is taken as evidence of fluidity of the host. Ambiguity arises in the case of some sedimentary rocks such as conglomerate and tillite which may become so thoroughly recrystallized that the sedimentary rock fragments appear to be xenoliths surrounded by crystalline magma.

Some remarkable xenoliths include the metallic iron-rich rock fragments found in some basaltic rocks which have intruded carbon-rich sedimentary rocks and the peridotite xenoliths found in diamond pipes and in some basaltic rocks. Scientists infer some of the peridotite xenoliths to be actually pieces of the

Fig. 1. Xenoliths in granodiorite, Sierra Nevada, California, appear soaked and drawn out parallel to the hammer handle. Hammer is 10 in. (25 cm) long. (*W. B. Hamilton, USGS*)

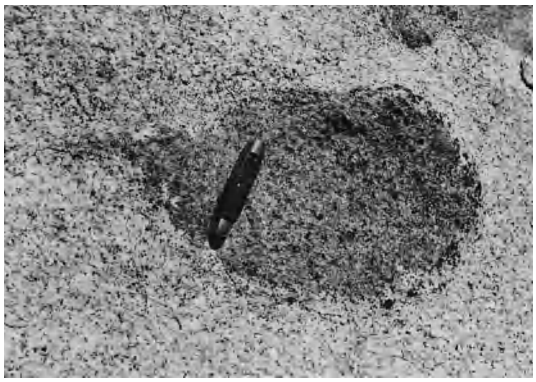

Fig. 2. Xenolith in granite, Sierra Nevada, California. Note sharper contact on one side and gradational contact against contaminated granite on other. (*W. B. Hamilton, USGS*)

Earth's mantle. *See* IGNEOUS ROCKS; LAVA; MAGMA; PERIDOTITE. Alfred T. Anderson, Jr.

Bibliography. J. E. Nielson and J. K. Nakata, Mantle Origin and Flow Sorting of Megacryst-Xenolith Inclusions in Mafic Dikes of Black Canyon, Arizona, *USGS Prof. Pap.*, no. 15, 1994; S. Shi et al., Xenolith evidence for lithospheric melting above anomalously hot mantle beneath the northern Canadian Cordillera, *Contrib. Mineral. Petrol.*, 131:39–53, 1998.

Xenon

A chemical element, Xe, atomic number 54. It is a member of the family of noble gases, group 18 in the periodic table. Xenon is colorless, odorless, and tasteless; it is a gas under ordinary conditions (see **table**). *See* INERT GASES; PERIODIC TABLE.

Xenon is the only one of the nonradioactive noble gases which forms chemical compounds that are stable at room temperature. Xenon also forms weakly bonded clathrates with such substances as water, hydroquinone, and phenol. *See* CLATHRATE COMPOUNDS.

The three fluorides, XeF_2, XeF_4, and XeF_6, are thermodynamically stable compounds at room temperature, and they may be prepared simply by heating mixtures of xenon and fluorine at 300–400°C (570–750°F).

The reaction of XeF_6 with water gives $XeOF_4$; if the reaction is allowed to continue, XeO_3 is formed. XeO_3 is a colorless, odorless, and dangerously explosive white solid of low volatility. Gaseous xenon tetroxide, XeO_4, is formed by the reaction of sodium perxenate, Na_4XeO_6, with concentrated H_2SO_4. The vapor pressure of XeO_4 is about 3.3 kPa at 0°C (32°F). It is unstable and has a tendency to explode.

Xenon is produced commercially in an air-separation plant. The air is liquefied and distilled. The oxygen is redistilled; the least volatile portion contains small amounts of xenon and krypton, which are adsorbed on silica gel directly from the liquid oxygen. The crude xenon and krypton thus obtained are separated and further purified by distillation and selective absorption, at controlled low temperatures, on activated carbon. Remaining impurities are removed by passing the xenon over hot titanium, which reacts with all but the inert gases. *See* AIR SEPARATION.

Physical properties of xenon

Property	Value
Atomic number	54
Atomic weight (atmospheric xenon only)	131.30
Melting point (triple point)	−111.8°C (−169.2°F)
Boiling point at 1 atm pressure	−108.1°C (−162.6°F)
Gas density at 0°C and 1 atm pressure, g/liter	5.8971
Liquid density at its boiling point, g/ml	3.057
Solubility in water at 20°C, ml xenon (STP) per 1000 g water at 1 atm partial pressure of xenon	108.1

Xenon is used to fill a type of flashbulb used in photography and called an electronic speed light. These bulbs produce a white light with a good balance of all the colors in the visible spectrum, and can be used 10,000 times or more before burning out.

A xenon-filled arc lamp gives a light intensity approaching that of the carbon arc; it is particularly valuable in projecting motion pictures. *See* VAPOR LAMP.

An important development in high-energy physics was the detection of nuclear radiation, such as gamma rays and mesons, by bubble chambers, in which a liquid is kept at a temperature just above its boiling point. Nucleation by the radiation results in bubble formation along the path of the particle. The tracks made by the particles are then photographed. Liquid xenon is one of the liquids used in these bubble chambers.

Xenon is used to fill neutron counters, x-ray counters, gas-filled thyratrons, and ionization chambers for cosmic rays; it is also used in high-pressure arc lamps to produce ultraviolet radiation.

Between 3 and 5% of the fissions in a nuclear reactor using uranium as fuel lead to the formation of xenon-135. Arthur W. Francis

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; M. Ozima and F. A. Podosek, *Noble Gas Geochemistry*, 2001.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb
actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Xenophyophorida

An order of Protozoa in the subclass Granuloreticulosia. The group includes deep-sea forms which are multinucleate at maturity. They develop as discoid to fan-shaped or algalike branching forms (**illus.** *a* and *b*) covered with a hyaline organic layer and sometimes measuring 0.08 in. (2 mm) or more overall. The aggregate contains many tubes (illus. *c*) which vary in diameter and form netlike or branching patterns. The more delicate tubes contain multinucleate protoplasm; the coarser ones, dark brownish clumps which are possibly waste

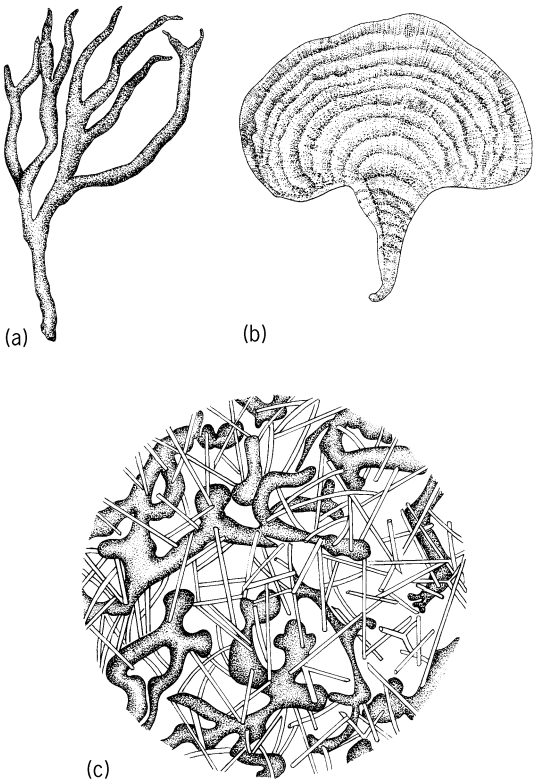

Xenophyophorids. (*a*) *Stannoma dendroides*. (*b*) *Stannophyllum zonarium*. (*c*) *Psammetta erythrocytomorpha*, section showing interlacing tubules and dark inclusions of uncertain nature.

materials. Foreign particles, such as radiolarian skeletons or foraminiferid tests, are often embedded in the nets or among the branches. A feltlike material also is found between adjacent tubes in some species. The lack of detailed information makes the taxonomic status of these organisms somewhat uncertain. The genera included are *Psammetta*, *Stannoma*, and *Stannophyllum*. *See* GRANULORETICULOSIA; PROTOZOA; RHIZOPODEA; SARCODINA; SARCOMASTIGOPHORA.

Richard P. Hall

Xenoturbella

A simple marine worm whose natural habitat is in soft mud bottoms at a depth of around 60 m (200 ft) off the coasts of Sweden and Norway. *Xenoturbella* was first discovered in 1915 by the Swedish biologist Sixten Bock, but was not described until 30 years later by Einar Westblad. Due to its simplicity of form, Westblad initially classified *Xenoturbella* as a primitive turbellarian, a class of ciliated nonparasitic flatworms, and named it accordingly (*Xenoturbella* = strange flatworm). However, recent analyses using deoxyribonucleic acid (DNA) have found *Xenoturbella* to be a primitive member of the deuterostomes, the group of animals that includes the vertebrates.

Morphology. *Xenoturbella* is a small worm up to 3 cm (1.2 in.) in length with a mouth opening into a gut cavity (**Fig. 1**). It has no anus and lacks defined excretory structures, body cavities, and reproductive organs. It has a gravity-sensing organ called a statocyst and two sensory grooves with a dense underlying nerve plexus. *Xenoturbella* has neither brain nor condensed nerve cord; its nervous system consists simply of a nerve net (a diffuse network of neurons) under the epidermis. Externally, it is completely covered in hairlike structures called cilia, which it uses to glide along the muddy seafloor. The reproduction and embryology of Xenoturbella are still a mystery.

Phylogeny. *Xenoturbella* has been the subject of controversy for almost a century. Despite numerous attempts at assigning *Xenoturbella* a position in the evolutionary tree of life, this animal has remained enigmatic. The problem with finding out its phylogenetic affinities (its relationship to other groups of organisms) has been its extreme morphological simplicity.

Following its initial classification, *Xenoturbella* was believed to be the most primitive bilaterally symmetrical animal. Another hypothesis supported a relationship with the hemichordates (acorn worms) and echinoderms (starfishes and sea urchins), based on ultrastructural similarities of the epidermis, the presence of a nerve net, and similarities between the statocysts of *Xenoturbella* and those of a group of echinoderms, the sea cucumbers.

In 1997, a molecular study based on the analysis of ribosomal DNA and eggs found inside the animal seemed to suggest a relationship to bivalve mollusks of the genus *Nucula*. The fact that nuculid bivalves are abundant in the same habitat as *Xenoturbella* suggested instead that these results might be contaminated with food ingested by *Xenoturbella* in the form of mollusk eggs and embryos.

More recent molecular analyses have revealed that *Xenoturbella* belongs within the deuterostomes. The deuterostomes contain the chordates (the phylum to which vertebrates and humans belong), the hemichordates (acorn worms), and the echinoderms (starfishes). Further analyses of the genetic code and the order of genes on the mitochondrial genome of *Xenoturbella* suggest a placement at the base of hemichordates and echinoderms (**Fig. 2**). These analyses suggest that *Xenoturbella* is a member of a separate phylum, the *Xenoturbellida*.

Because of the pivotal phylogenetic position of *Xenoturbella* as the closest sister group to the hemichordates and echinoderms and an outgroup to

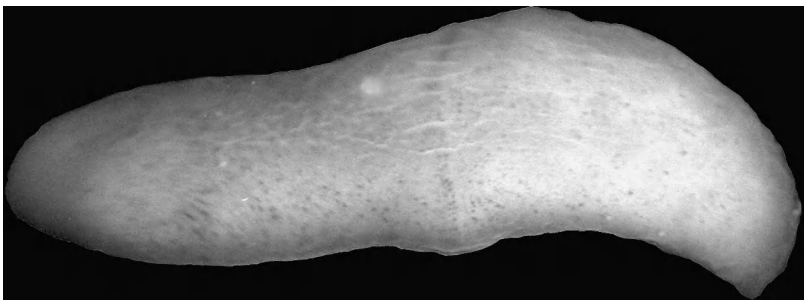

Fig. 1. Photograph of a live specimen of *Xenoturbella* (1–2 cm), showing the sensory circumferential groove around its middle.

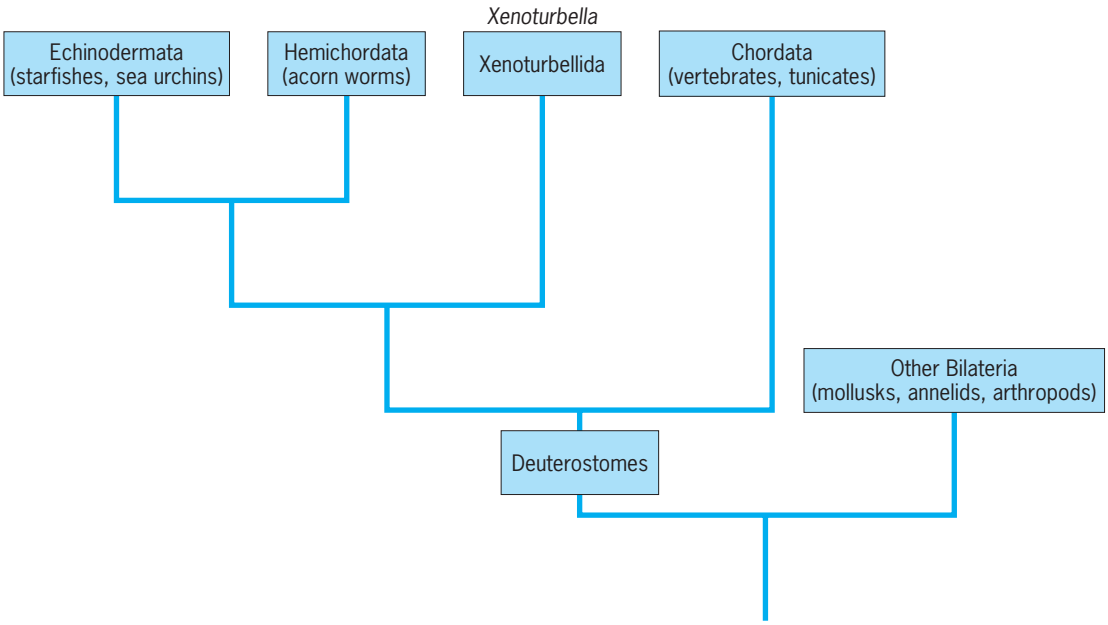

Fig. 2. Phylogenetic tree based on the 18s ribosomal DNA gene, showing the position of *Xenoturbella* at the base of the hemichordates and echinoderms.

the chordates, future studies of the genetics, morphology, and embryology of *Xenoturbella* have the potential to provide great insight into the evolution of the deuterostomes. *See* ANIMAL EVOLUTION; ANIMAL SYSTEMATICS; DEUTEROSTOMIA; ECHINODERMATA; GENE AMPLIFICATION; HEMICHORDATA; PHYLOGENY. Sarah J. Bourlat

Bibliography. S. J. Bourlat et al., *Xenoturbella* is a deuterostome that eats molluscs, *Nature*, 424:925–928, 2003; H. Gee, You aren't what you eat, *Nature*, 424:885–886, 2003.

Xenungulata

Xenungulates are large, primitive, digitigrade, hoofed, tapirlike mammals with relatively short, slender limbs and five-toed feet with broad, flat phalanges. They are restricted to the Paleocene deposits of Brazil and Argentina. Only one genus (*Carodnia*) constitutes the single family (Carodniidae) in the order. *See* EUTHERIA.

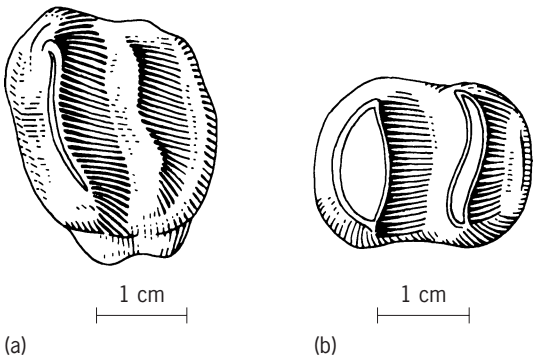

Molars of *Carodnia vieiral* from the Paleocene of Brazil: (*a*) upper and (*b*) lower.

The dentition (I 3/3 C 1/1 Pm 4/4 M 3/3) is complete with strong, procumbent, chisel-shaped incisors, strong sharp-pointed canines, and low-crowned cheek teeth with bilophodont molars (see **illus.**). *See* DENTITION.

The affinities of the Xenungulata remain uncertain. The combination of the dental characters is unique among the South American ungulates. Affinities with the Dinocerata are strongly supported by the dental characteristics, but the structure of the tarsus suggests that the xenungulates had common ancestry with the Pyrotheria, to which they were tentatively assigned originally. *See* DINOCERATA; PYROTHERIA. Gideon T. James; Everett C. Olson

Bibliography. M. J. Benton, *Vertebrate Paleontology*, 1991; R. Carroll, *Vertebrate Paleontology and Evolution*, 1988.

Xiphosurida

An order (xiphosurids, or horseshoe crabs) of the class Xiphosura (phylum Arthropoda). These "sword-tailed" arthropods are a classic example of a "living fossils" group that has seemingly changed little over a period of many millions of years.

Xiphosurids have three distinct body divisions: a prominent horseshoe-shaped, fused head and thorax, called the prosoma; a fused "abdomen," or opisthosoma; and a spikelike terminal segment, called the telson (**Fig. 1**). Xiphosurida is a low-diversity group that has a moderately good fossil record in some paleoenvironments, especially marginal-marine environments. Present-day xiphosurid species are tolerant of wide ranges of salinity, from marine to brackish water, and occasionally even venture into freshwater; ancient species may have

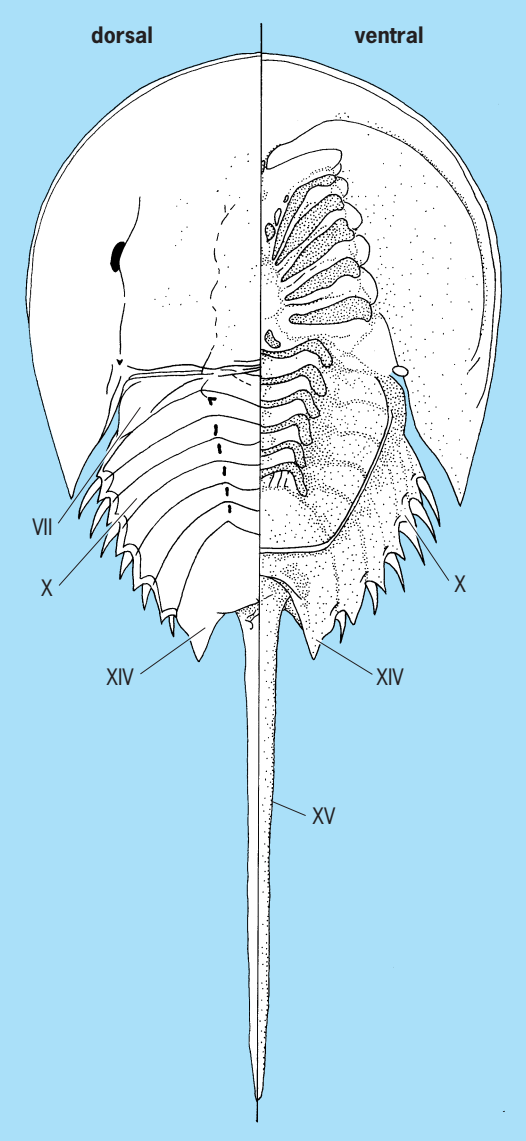

Fig. 1. Segmentation of *Limulus polyphemus*. Inconclusive evidence indicates that there are 15 segments (6 in the prosoma, 9 in the opisthosoma). The telson (XV) may be a terminal spine that has become articulated and has incorporated the anus into its base.

had similar salinity tolerances.

The order Xiphosurida is a monophyletic group of arachnomorph arthropods, a large group that includes nektaspids, trilobites, eurypterids, arachnids, and synziphosurines. Two xiphosurid suborders are recognized: the more primitive Bellinurina, all of which are now extinct, and the more derived Limulina, which includes all the four living species of the group (**Fig. 2**). The most primitive group of xiphosurids (suborder Bellinurina) probably evolved from a stem-group xiphosuran, called a synziphosurine, during the Late Devonian (about 365 million years ago), and became extinct in the Permian Period. In the synxiphosurines, which are the sister group to the bellinurine xiphosurids, the segments of the opisthosoma articulate freely and the telson is short and usually wide. Some species lack compound eyes. Evolution of the bellinurines involved fusion of the opisthosoma into a single plate sometimes called a thoracetron, and an increase in the length of the telson. Other developments include the addition of true lateral spines on the opisthosoma, and development of a small node bearing light-sensitive eyes (ocelli) in front of the axial, or cardiac, lobe on the prosoma. The Limulina evolved from a bellinurine ancestor in the early Carboniferous (about 335 million years ago). The most important evolutionary innovation in the limulines is the change in shape of the small lobe at the rear of the opisthothorax from triangular (the condition in bellinurines) to trapezoidal. Limulines are also characterized by having lost the axial part of the opercular tergite (the eighth tergite) and by the formation of conical projections bearing the genital pores on the back of the operculum. *See* MEROSTOMATA.

There are four extant species of limulines: *Limulus polyphemus*, which occurs on the eastern coast of North America, and three species distributed in the Indo-Pacific region (*Tachypleus tridentatus, T. gigas*, and *Carcinoscorpius rotundicauda*). Phylogenetic studies based on mitochondrial genes suggest that *L. polyphemus* is the sister group to all of the Indo-Pacific species, and that *T. gigas* and *C. rotundicauda* form a monophyletic group that has the less derived *T. tridentatus* as its sister group.

Fossil record. Xiphosurids include 13 genera known only from fossils, two genera known only from living species, and one genus that is known from both fossil and living species. The suborder Bellinurina includes the Late Paleozoic genera *Bellinurus, Liomesaspis, Euproops, Bellinuroopsis*, and *Rolfia*. The suborder Limulina includes the Late Paleozoic genera *Xaniopyramus, Paleolimulus, Panduralimulus*, and *Valloisella*; the Mesozoic genera *Limulitella, Psammolimulus*, and *Mesolimulus*; the Late Mesozoic-Cenozoic genus *Limulus*; and the exclusively Cenozoic genera *Tachypleus* and *Carcinoscorpius*.

Xiphosurids seem to have an anomalous fossil record. Their exoskeletons, like those of most arthropods, are nonbiomineralized; they are composed of chitin, which is a polysaccharide that does not normally fossilize. Moreover, fossil xiphosurids are unusual because they are commonly found complete, retaining the prosoma, the opisthothorax, and the telson. In some instances, ventral appendages and book gills are also attached (**Fig. 3**).

Studies on the fossilization history (taphonomy) of modern *Limulus* have elucidated some of the reasons that xiphosurids have such an unusual fossil record, making them a model for the preservation of other types of nonmineralizing arthropods. *Limulus* disarticulates slowly after death. As long as the bodies remain in water, smaller specimens in particular tend to become pliable, and subject to wrinkling and compression when buried under sediment. Normally, it takes at least 2 weeks for the telson to separate from the opisthosoma, and 4 weeks or more for the opistosoma to separate from the prosoma. The slow rate of separation of the major body divisions helps account

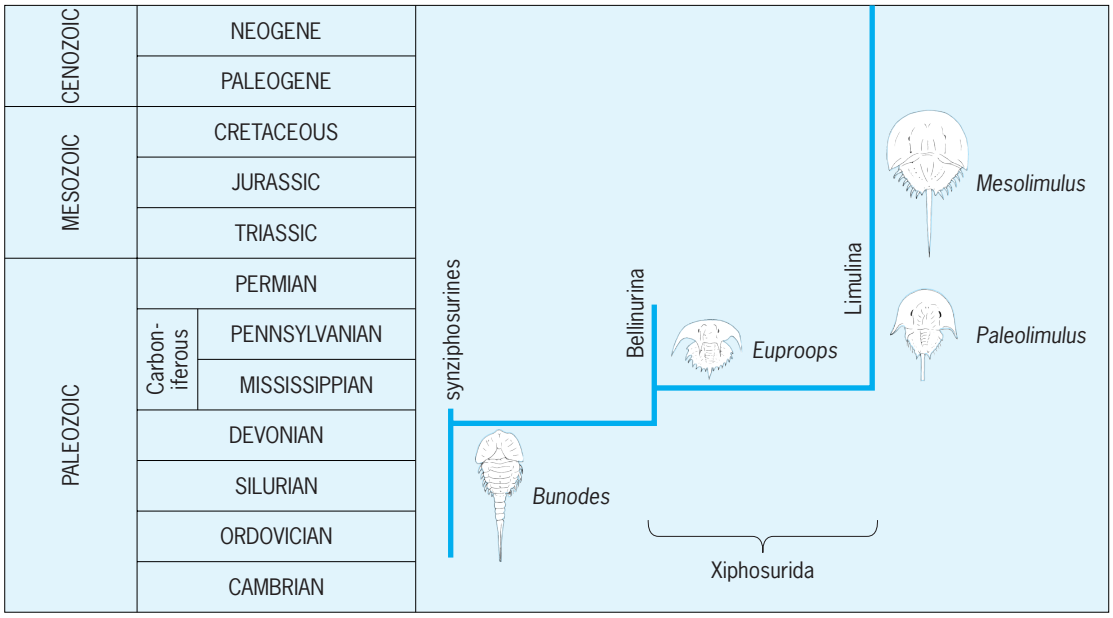

Fig. 2. Phylogeny and representive species of the class Xiphosura, including the order Xiphosurida and its sister group, synziphosurines (a paraphyletic stem group).

for the common occurrence of complete horseshoe crab fossils.

Within a few days after death, modern horseshoe crabs become enveloped in a thin film or "halo" composed of microscopic fungi and bacteria. The microbes apparently aid in decay of the chitinous exoskeleton, but at some point bacteria seem to initiate the process of fossilization of the exoskeleton by causing precipitation of minerals such as calcite (calcium carbonate) or siderite (iron carbonate) on the surface of the horseshoe crab. Precipitation of minerals on the exoskeleton apparently can begin within a few weeks of death, so a slow disarticulation rate coupled with a rapid fossilization rate contributes to the unusually good fossil record of these animals.

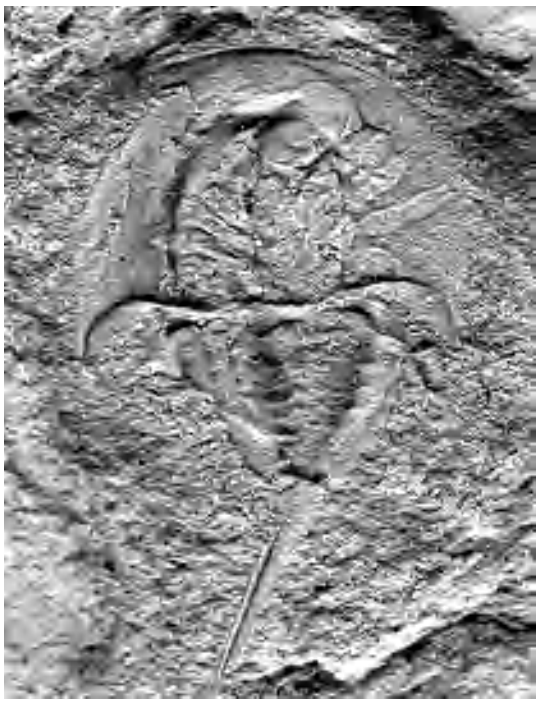

Fig. 3. *Paleolimulus signatus*, a limuline xiphosurid from the Permian of Kansas, showing appendages and book gills. Length of specimen: 1.25 cm (0.5 in.).

It is also possible that factors related to habitat played an important role in the fossilization of xiphosurids. Often their remains become buried in estuarine, tidal flat, or lacustrine environments. Salinity fluctuations, and possibly desiccation, in these settings may have limited the number of scavengers and sediment burrowers that could destroy xiphosurid remains. *See* TAPHONOMY.

Distribution. Distribution of fossilized and living species suggests that limuline xiphosurids have always been tropico-temperate species, occurring mostly in the Northern Hemisphere. None of the extant species has a continuous distribution; each has an intermittent distribution in coastal waters. Occasionally, individuals are found well upstream from coastal estuaries.

Limulus polyphemus is found along the Atlantic coast of North America from Maine to the Yucatán Peninsula, Mexico. The largest adults and largest populations occur in the middle part of the species range, from Delaware Bay to Georgia.

The ranges of the three species endemic to the Indo-Pacific region overlap in part. *Tachypleus tridentatus* occurs along the western and southern shores of Japan, along the southeastern coast of China to southern Vietnam, and along the western islands of the Philippines. *Tachypleus gigas* occurs along the shores of the Bay of Bengal, and from Indochina to northern Vietnam, Borneo, and the Celebes Islands of Indonesia. *Carcinoscorpius rotundicauda* extends from the western shore of the Bay of Bengal to the southern coast of the Philippines. In addition to geographic ranges of

comparable extent, the four species have similar ecologies, morphologies (**Fig. 4**), and serologies.

Life history. *Limulus polyphemus* is the most studied species of extant xiphosurid. Its common names include horseshoe crab (preferred), horsefoot, pan crab, piggyback crab, and king crab (not to be confused with *Paralithodes camtschatica*, the true king crab).

The characteristics, including behavior, of any one population are not exactly like those of another. There is evidence of genetic variability and, for whatever reason, obvious adjustments to local environments, such as the zone of the beach where eggs are laid: near the high-tide mark where the tidal level varies about 1 m (3.3 ft), as along the Gulf of Mexico coast of Florida; throughout the upper portion of the beach when the range between low and high tide is about 2 m (6.6 ft), as in Delaware Bay; and in the middle portion of the beach when the tidal amplitude exceeds 3 m (10 ft), as in the Cape Cod Bay area.

The life history of *L. polyphemus* in Delaware Bay, where the largest population abounds, is generally characteristic of the species. Adults spend the colder months in deeper waters within the bay and on the continental shelf. The geographic extent of this population is not known, but it may be far-ranging; other *Limulus* to the south, off North Carolina, have been dredged from depths of 66 m (218 ft) some 56 km (35 mi) off the coast. Warming waters and increasing daylight stimulate the adults to migrate toward the sandy bay beaches to spawn. Most follow the path of least resistance and follow the strong flood tide currents into the eastern side of Delaware Bay. Fewer horseshoe crabs gather along the western shore, except when rough waves strike Cape May beaches during prevailing northwesterlies and drive the animals back into the deeper water. Then they may venture westward toward the wind-protected shore.

On new moon, and especially full moon, and during high tides in May and June, thousands of *L. polyphemus* adults congregate on the beaches to spawn. Males "patroling" along the shore grasp the females heading for the beach. In an excavation dug beneath her body, each female lays several thousand eggs that are fertilized when sperm are washed into the nests by the waves. After making one or more nests, the animals leave the beach when the tide ebbs. Horseshoe crabs that do not leave before the tide recedes are stranded on the tidal flats, where they dig in to await the next high water.

Minnows swarm along the beaches at night and feed upon the eggs washed out of the nests by wave action. During the day sharp-eyed shore birds, mostly migrants such as plovers and sandpipers, dig into the nests and gorge themselves on eggs.

Horseshoe crab embryos begin to develop within opaque, tough-coated eggs. In a few days each egg cover ruptures and a thin, transparent sphere, formed during embryonic development, balloons out to assume a diameter larger than the original egg (**Fig. 5***a*). In this transparent showcase, growth of the legs, their first feeble attempts to move, and then

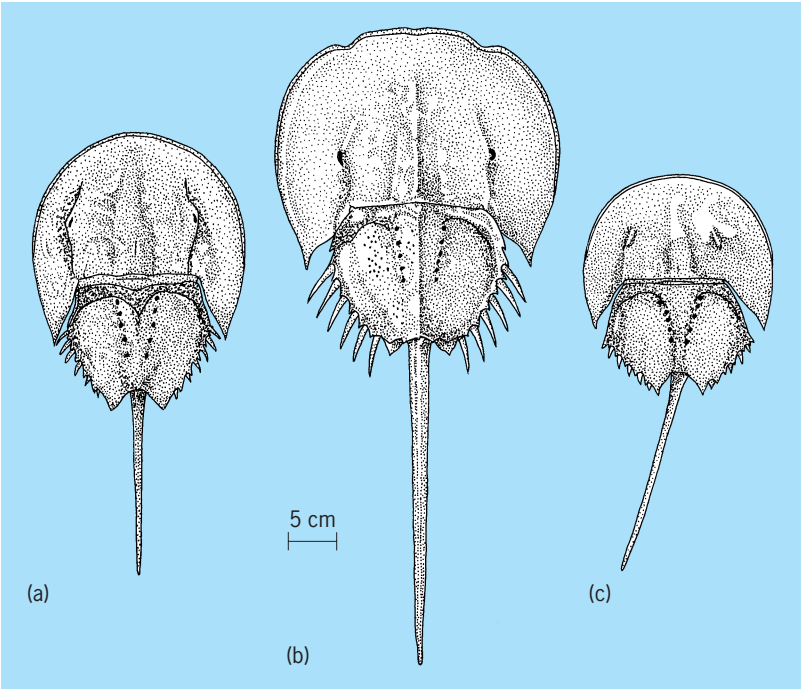

Fig. 4. Carapaces of adult male Limulina. (*a*) *Limulus polyphemus*. (*b*) *Tachypleus tridentatus*. (*c*) *Carcinoscorpius rotundicauda*.

their almost incessant movement are clearly visible (Fig. 5*b*).

Within 2 weeks some of the tiny animals reach the larval stage and are ready to hatch, although development and hatching may be delayed for several weeks. If the waves churn up a nest, the moving sand grains rupture the enclosing membrane and liberate the larvae (Fig. 5*c*). From then on, until they reach maturity in some 9 to 11 years and 15 or more stages later, their life is essentially seeking food and growing. Food consists mainly of marine worms, such as *Nereis*, and soft-shelled mollusks, such as *Mya arenaria*. The larger the immature animal, the more it moves toward the deeper bay water, from which, as an adult, it begins the spawning migration to renew the life cycle.

The newly hatched larva, which is referred to as the trilobite or Euproops stage, is in many respects still an embryo. Its incompletely formed digestive tract and nervous and circulatory systems develop further during early subsequent stages. The yolk-filled midgut serves as food for the larvae while they learn to right themselves from the upside-down position in which gravity held them during development in the egg (Fig. 5*c*). Horseshoe crabs of all ages swim on their backs, however, and usually near the water surface. *See* INVERTEBRATE EMBRYOLOGY.

Molting. Molting occurs several times during the first 2 or 3 years of life, but rarely more than once thereafter until adult size is reached. It is preceded by the formation of the new skin, recognizable externally in the larger animals by a deeper olive-green coloration with yellow margins of the shell. The new skin forms pleats that unfold as the animal expands with the uptake of water during emergence. These

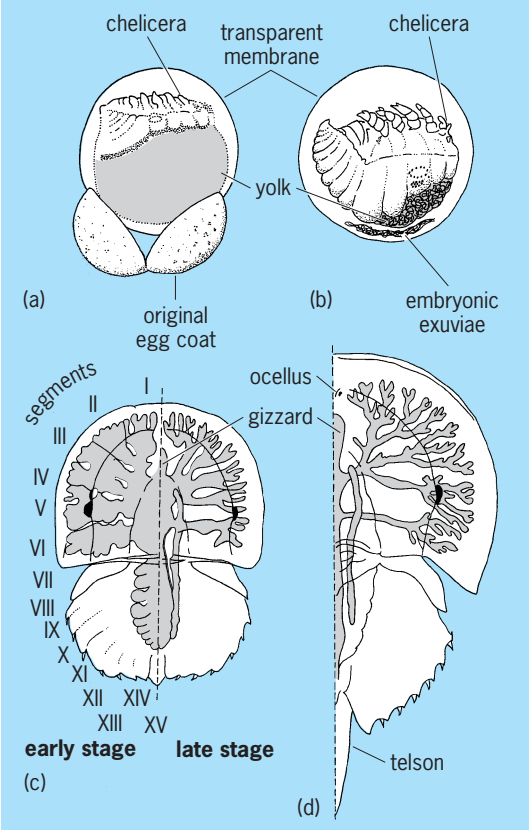

Fig. 5. Embryonic and larval development of *Limulus polyphemus.* (*a*) Early embryo, side view. (*b*) Late embryo. (*c*) Yolk-filled digestive tract of larva. Note the segmentation of yolk in the newly hatched stage and the six lobes of digestive gland. (*d*) First tailed stage.

pleats in the premolt skin give the hardened exoskeleton its dark-lined, mosaic appearance. A circumferential splitting, just within the ventral margin of the prosoma, is the first outward evidence of molting. Exuviation (casting off the shell) then proceeds slowly until the animal emerges one-third of the way and is loosened from the old shell. Emergence then proceeds more rapidly. The first tailed stage emerges within an hour, but an animal 10 cm (4 in.) wide may require more than 24 h to shed its old shell. Young specimens molt earlier than older ones, with cast shells appearing in sequence during July and August on tidal flats and in beach wrack. Instinctively, the larger, annually molting animals seek the areas along the edges of tidal flats and dig in prior to molting. Burying protects them from their enemies: birds, fishes, and crabs. The moist substrate supplies water at all times.

Nutrition. In feeding, the pincer-tipped legs grasp and bring prey to the heavy bases of the pushers, the fifth pair of legs, where shells of mollusks are crushed. A pair of paddlelike appendages and spines on the leg bases then push and knead the prey forward toward the mouth, which is surrounded by the leg bases. Chelipeds in front of the upper lip push food backward. Chitinous ridges of the muscular gizzard, with the aid of ingested sand and shell, grind up prey into pulp that is squeezed into the midgut, where it is partially digested. The resulting chyle is forced into the many-branched digestive gland, where digestion within the cells completes the process. Large pieces of shell and pebbles are regurgitated from the gizzard, but fine sand and smaller shell fragments may be passed with undigested residues in soft, slime-enveloped, fecal cylinders.

Anatomy. The carapace (test or shell) is composed of chitin, strengthened internally in the adults by bridgelike struts within the shell margins (**Fig. 6**). Sensory structures on the carapace include chemoreceptors for taste on the bases of the legs and a pair of simple eyes, or ocelli. The multilensed compound eyes form mosaic images and are sensitive to polarized light.

Adult males are easily distinguished from the larger females by the higher anterior arch of the prosoma, by the pair of "fish-and-thumb" claspers used in mating, and by the hard, conical projections bearing the genital pores on the back of the operculum (eighth appendage). The female apertures are larger, soft, and oval.

The organ systems, except the peripheral blood vessels and nerves, digestive, and reproductive glands, which branch extensively into the lateral portions of the prosoma, are mainly within the prominent axial region (Fig. 6*c*). This region contains the musculature of the paired appendages, hinge, and telson, the digestive tube, the coxal (excretory) gland, and the respiratory, central circulatory, and nervous structures. The dorsally situated tubular heart lies within a sinus that receives venous blood, the hemolymph, from the book gills (the respiratory structures) and the body. The central nervous system and the major nerve branches are sheathed by arteries. A plate of "cartilaginous" chitin, the endocranium, forms a roof over the brain. Flat, elliptical, primitive white blood cells, the amebocytes, circulate in the hemolymph. When exposed to air, the amebocytes send out pseudopods and fragment to form a clot. The blood pigment is hemocyanin. *See* RESPIRATORY PIGMENTS (INVERTEBRATE).

Importance. Early colonists from Europe saw Native Americans use the tails of horseshoe crabs for spearheads and the animals for fertilizer. In the Delaware Bay area, prior to 1960, horseshoe crabs were ground up to form a meal having a protein content of 46% and used in fertilizer. Quartered animals make good eel bait and have been used as food for hogs and poultry. The roe reportedly increases growth of broilers and egg laying, although a fishy flavor is imparted to both flesh and eggs. Eel fishery has made heavy inroads into some *L. polyphemus* populations.

Limulus polyphemus is particularly useful in biomedical research. The nerve and muscle tissue of the heart are easily separable anatomically. All the nerve cells lie in compact masses on the surface of the heart and can be removed without damaging the muscle tissue. Studies on the heart of horseshoe crabs have played an important role in the understanding of cardiac function in vertebrates. *Limulus polyphemus* has been used extensively in research on the visual system. The compound eye polarizes light, and its crystalline cones concentrate light

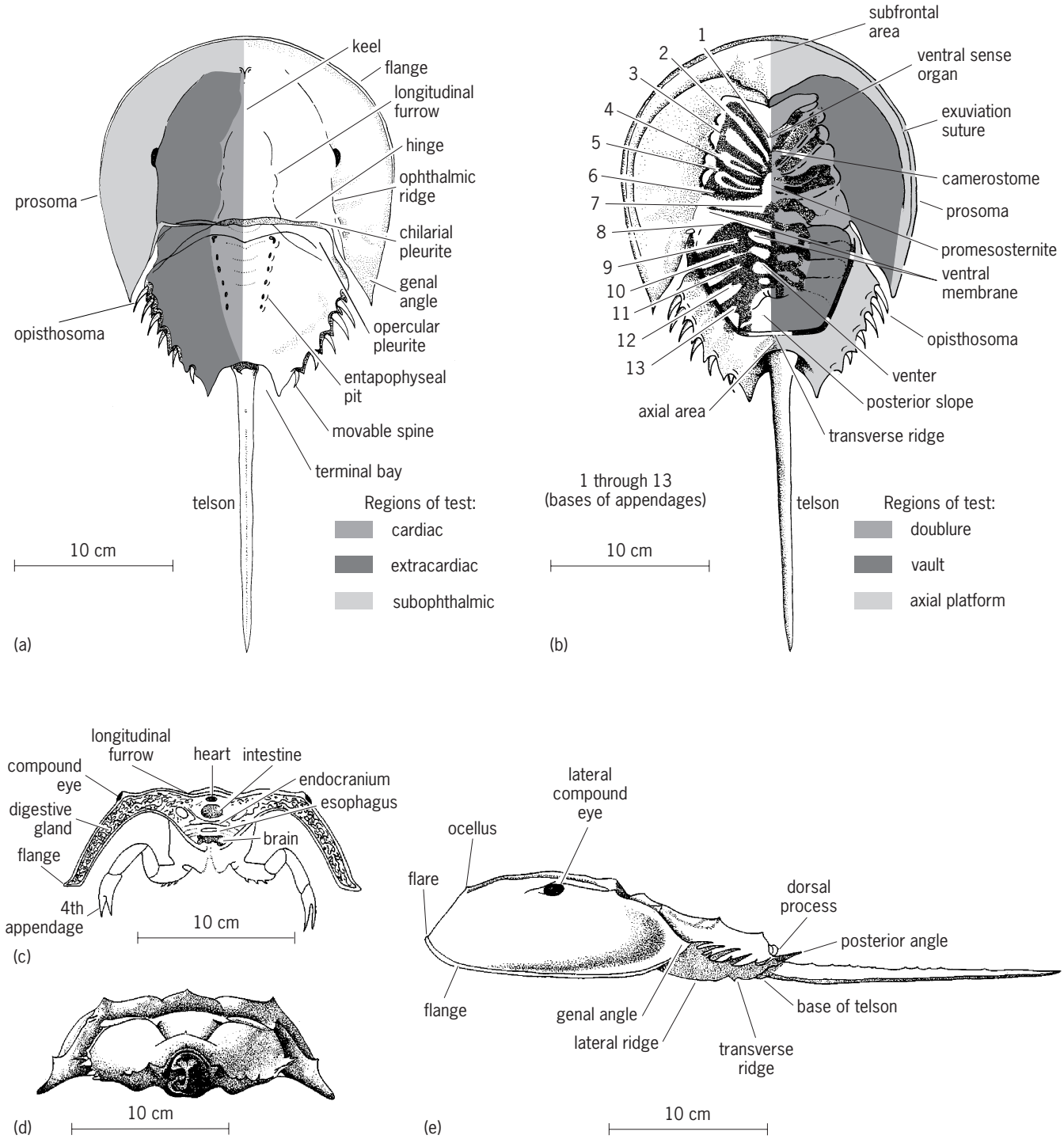

Fig. 6. Morphology of adult male *Limulus polyphemus.* (*a*) Dorsal view. (*b*) Ventral view, appendages not shown. (*c*) Cross section through prosoma showing internal anatomy. (*d*) Posterior view. (*e*) Lateral view.

tenfold. Just how the animals use these capabilities is unknown; perhaps they are used in sensing the time for spawning and for guiding the annual migration from deep water to the beaches where spawning takes place. The curved cone shape of the eye lenses has been incorporated into a solar energy collector. The major biomedical applications of limulines are the detection of bacterial endotoxin, using a lysate from the amebocytes, and the probing of cellular membrane constituents, using agglutinins (lectins) from the hemolymph. Limulines have some interesting medical uses. For example, they have been used to produce surgical sutures and wound dressings.

In some areas, as in New England waters and off New Jersey where shellfish resources are heavily harvested (the softshell clam, *Mya arenaria*, and the surf clam, *Spisula solidissima*, respectively), predation by *L. polyphemus* is a problem, and horseshoe crabs have been hunted and destroyed.

The taking of horseshoe crabs for various human uses and to prevent predation has seriously impacted

some *L. polyphemus* populations and, in conjunction with destruction of estuarine habitats, has raised concerns about the conservation of the species.

Carl N. Shuster, Jr.; Loren E. Babcock

Bibliography. L. I. Anderson and P. A. Selden, *Lethaia*, 30:19–31, 1997; L. E. Babcock, D. F. Merriam, and R. R. West, *Lethaia*, 33:129–141, 2000; P. S. Borkow, and L. E. Babcock, *The Sedimentary Record*, 1(3):4–7, 2003; D. C. Fisher, in N. Eldredge and S. M. Stanley (eds.), *Living Fossils*, 1984; J. H. Lochhead, in F. A. Brown (ed.), *Selected Invertebrate Types*, 1950; S. M. Manton, *The Arthropoda: Habits, Functional Morphology, and Evolution*, 1977; S. P. Parker (ed.), *Synopsis and Classification of Living Organisms*, 2 vols., 1982; C. N. Shuster, Jr., *The Circulatory System and Blood of the Horseshoe Crab, Limulus polyphemus L.: A Review*, 1978; C. N. Schuster, Jr., R. B. Barlow, and H. J. Brockmann (eds.), *The American Horseshoe Crab*, 2004; R. E. Snodgrass, *A Textbook of Arthropod Anatomy*, 1952; L. Størmer, in R. C. Moore (ed.), *Treatise on Invertebrate Paleontology*, pt. P2, 1955.

Xylem

The principal water-conducting tissue and the chief supporting system of higher plants. This tissue and the associated phloem constitute the vascular system of vascular plants. Xylem is composed of various kinds of cells, living or nonliving. The structure of these cells differs in their functions, but characteristically all have a rigid and enduring cell wall that is well preserved in fossils.

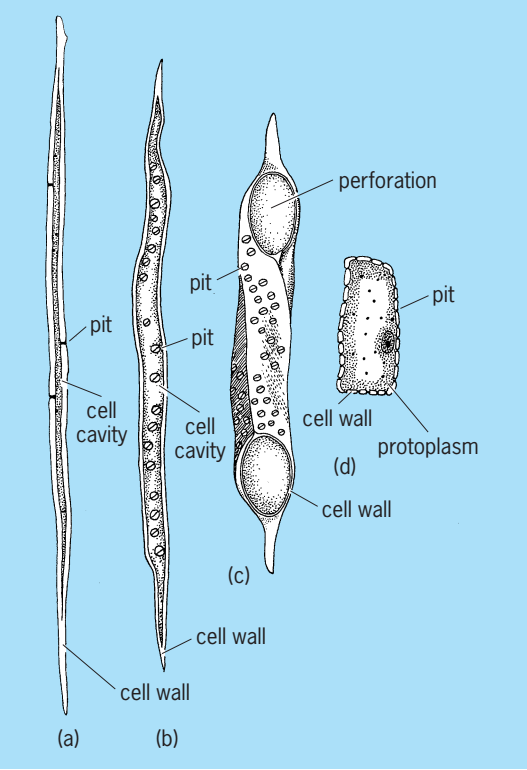

Fig. 1. Xylem cell types. (*a*) Wood (xylem) fiber. (*b*) Tracheid. (*c*) Vessel member. (*d*) Xylem parenchyma cell. All can vary widely in structure. (*After H. J. Fuller and O. Tippo, College Botany, rev. ed., Holt, 1954*)

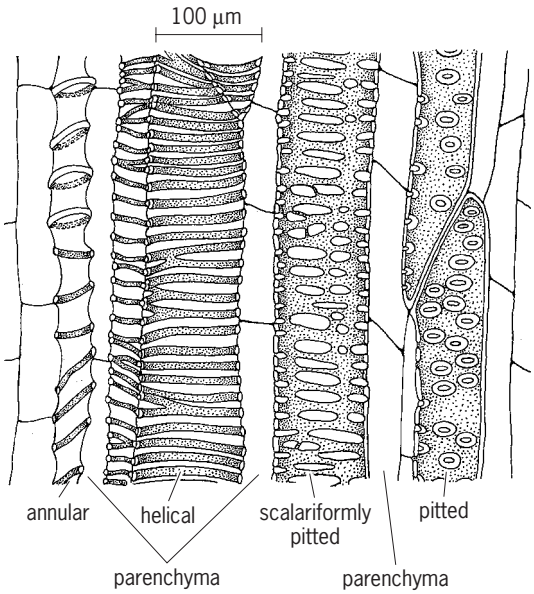

Fig. 2. Parts of tracheary elements and associated parenchyma from stem of *Aristolochia*, as seen in longitudinal section. The secondary walls are stippled. The earliest part of the xylem is at the left. (*After K. Esau, Plant Anatomy, 2d ed., John Wiley and Sons, 1965*)

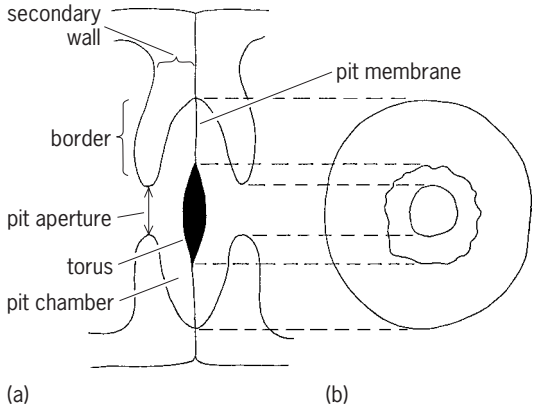

Fig. 3. Diagrams of bordered pit pairs of the genus *Pinus*. (*a*) Sectional view. (*b*) Face view. The torus is formed by thickening of primary walls. (*After K. Esau, Plant Anatomy, 2d ed., John Wiley and Sons, 1965*)

Kinds of xylem cells. In terms of their functions, the kinds of cells in xylem are those related principally to conduction and support, tracheids; to conduction, vessel members; to support, fibers; and to food storage, parenchyma (**Fig. 1**). Vessel members and tracheids are often called tracheary elements. The cells in each of the four categories vary widely in structure.

Tracheids. Tracheids are elongate, spindle-shaped cells, lacking protoplasm at maturity, and having secondary walls laid in various thicknesses and patterns over the primary wall. The secondary wall of tracheids and other conducting cells may consist of stretchable annular or helical (spiral) bands, or of a

nonextensible continuous expanse interrupted only at the bordered pits (**Fig. 2**).

Vessels. Vessel members differ from tracheids in having perforations, usually in the end walls common to superposed cells. A longitudinal series of vessel members constitutes a vessel. Although each varies markedly, tracheids are usually two to three times as long as vessel members, for example, 0.004–0.006 in. (1–1.5 mm) versus 0.02 in. (0.5 mm) in secondary wood. Vessels, however, may be from many centimeters to even meters in length.

In wood, pits are interruptions or recesses in the secondary wall, and are usually paired in common walls of contiguous cells. In bordered pits (**Fig. 3**) the secondary wall overarches the pit membrane (intercellular substance and two primary walls of contiguous cells); the overarch is absent in simple pits. The borders may be well developed or barely perceptible. The pit apertures (openings into the cell lumen) in bordered pits may be round to slitlike openings. Vertical series of transversely elongated pits are described as scalariform pitting. Round or oval pits may appear singly, or in horizontal (opposite pitting) or oblique (alternate pitting) series.

Perforations result from disintegration of the primary wall in the ends of vessel members. A perforation plate (**Fig. 4**) is the wall area bearing one or more perforations. The most primitive (scalariform) perforation plates are strongly oblique in position and have numerous transversely placed perforations (over 200). The most specialized have a single large (simple) perforation in a transverse wall.

Fibers. Fibers differ from tracheids in being more slender, in having simple pits and thicker walls in proportion to diameter, and in often retaining protoplasm. Fibers are thought to have evolved from tracheids, and intergrading forms between them are called fiber tracheids. Divisions of the protoplast of living fibers result in septate fibers.

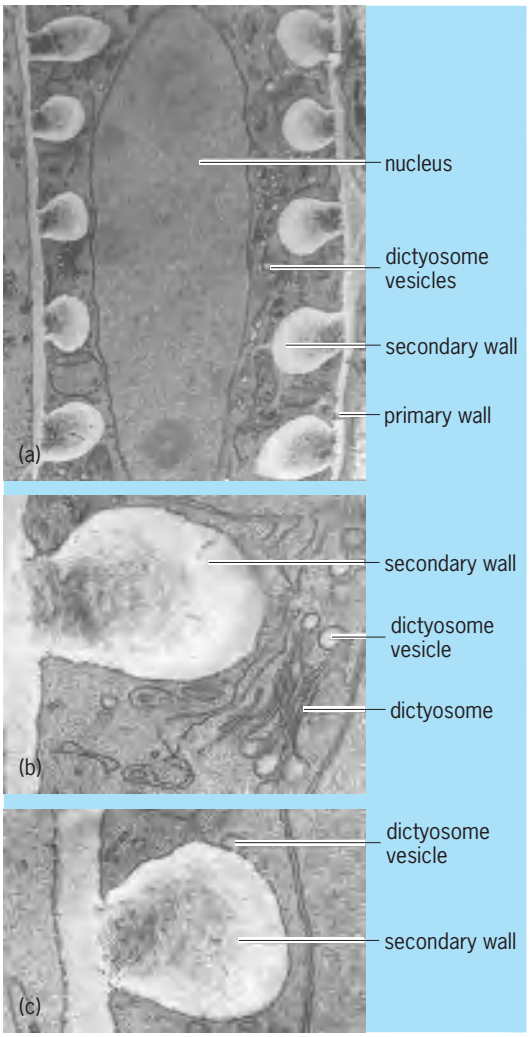

Fig. 5. Electron micrographs of longitudinal sections of vessel members with helical secondary wall. (*a*) Nucleus and numerous dictyosomes with vesicles. (*b*) One dictyosome near secondary wall. (*c*) Dictyosome vesicle attached to secondary wall after junction with plasmalemma coating the wall.

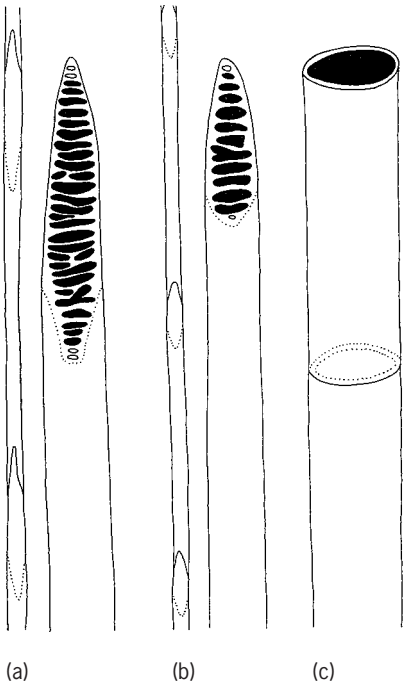

Fig. 4. Complete vessel members and parts of single-vessel members with perforation plates on end walls (perforations in black). (*a*) Long scalariform perforation plate. (*b*) Shorter scalariform. (*c*) Simple.

Parenchyma. The term parenchyma usually refers to the vertically arranged living cells of the xylem, although rays are also composed chiefly of parenchyma. Depending upon their origin, parenchyma cells may occur singly or as members of a strand. The cells may or may not have secondary walls, typically have simple pits, usually store starch, and may contain crystals, tannins, or other substances.

Development. Electron microscopy has shown that in the differentiating conducting cells the formation of secondary walls is associated with an intense activity of dictyosomes (Golgi bodies). These produce numerous vesicles which migrate toward the wall and fuse with the plasmalemma (plasma membrane) in

such a way that their contents appear to be released toward the secondary wall and their membranes become part of the plasmalemma (**Fig. 5**). The nucleus, all organelles, and the membrane systems degenerate as the conducting cell matures. If this cell is a vessel member, the perforation is formed before the protoplast dies.

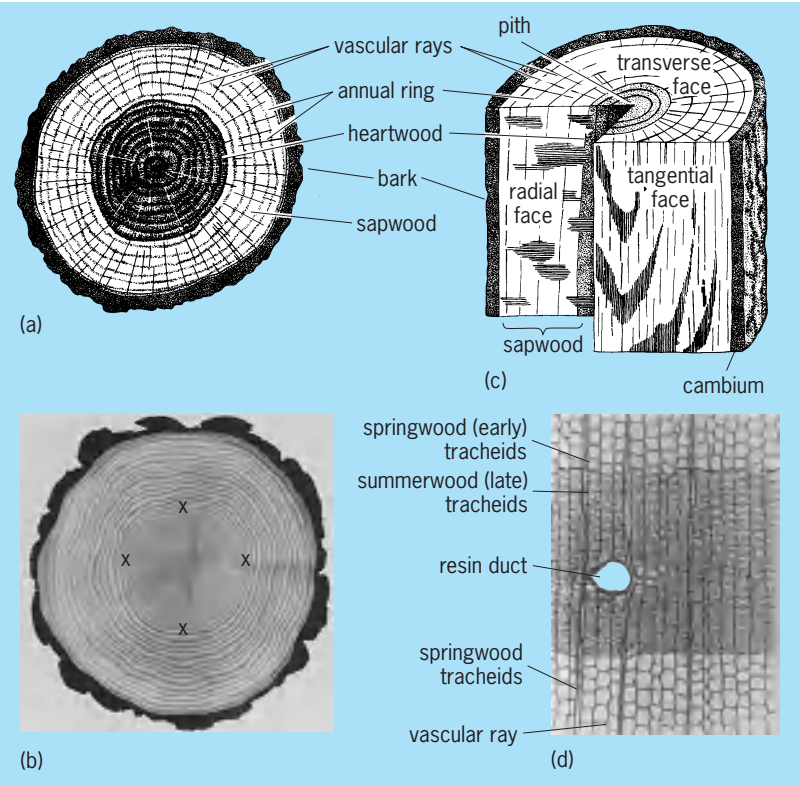

Fig. 6. Growth layers. (*a*) Cross section of oak branch. (*b*) Cross section of loblolly pine trunk showing effect of light on annual ring thickness. The annual rings inside the x's were formed while the young tree was densely shaded by older trees. The annual rings outside the x's were formed after the older trees had been removed and the young pine tree received abundant sunlight. (*c*) Three-dimensional appearance of a log. (*d*) Highly magnified part of transverse section of pinewood with parts of two annual rings. (*After H. J. Fuller and O. Tippo, College Botany, rev. ed., Holt, 1954*)

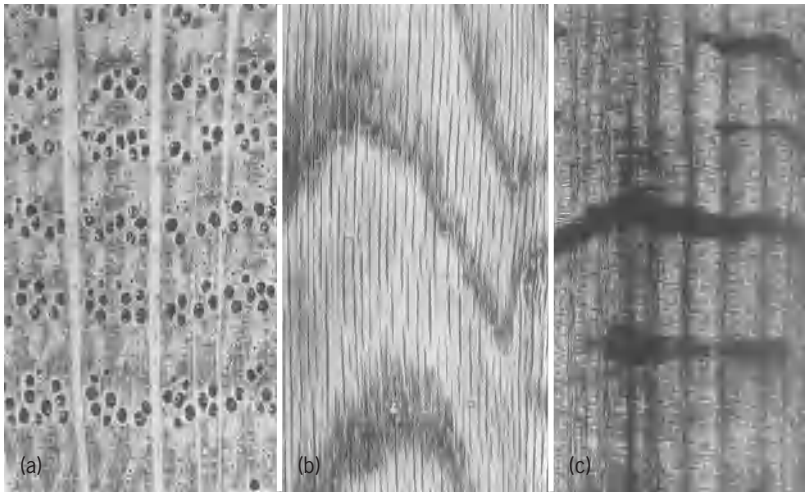

Fig. 7. Red oakwood (a ring-porous hardwood). The earlywood vessels are much larger than the latewood vessels. (*a*) Transverse section. (*b*) Tangential section. (*c*) Radial section. (*After H. J. Fuller and O. Tippo, College Botany, rev. ed., Holt, 1954*)

Primary and secondary xylem. Xylem tissues arise in later stages of embryo development of a given plant and are added to by differentiation of cells derived from the apical meristems of roots and stems. Growth and differentiation of tissues derived from the apical meristem provide the primary body of the plant, and the xylem tissues formed in it are called primary. Secondary xylem, when present, is produced by the vascular cambium. *See* LATERAL MERISTEM.

Primary. Protoxylem is that part of the primary xylem formed while the organ is still elongating and generally includes cells with extensible secondary wall thickening. These cells are frequently torn in vigorously growing plants. Metaxylem is formed after growth in length is nearly completed, and has cells mostly with nonextensible secondary walls.

Parenchyma cells, with thin or thick walls, occur in both protoxylem and metaxylem, and, depending upon the species, vessels, tracheids, and fibers all may be present.

Secondary. Secondary xylem is composed of two interpenetrating systems, the horizontal (ray) and vertical (axial). Wide variation occurs in secondary wood, mostly because of inherent differences among species and partly because of environmental influences. The differences in rate of evolutionary specialization of the different kinds of cells as seen in any one species make possible the positive identification of nearly all kinds of wood. Variations include differences in types of rays and their distribution; kinds, numbers, arrangement, and sizes of vessels, tracheids, and fibers; types of growth layers; character of sapwood and heartwood; kinds of knots; distribution of parenchyma; and other features, such as odor. Many of these variations bear on the strength and usefulness of wood.

Rays may be one to many cells in width as seen in transections of the organ in which they occur, may be composed of cells alike or unlike in form, are variable in length radially, and differ in height, for example, high in oak, low in maple. Evolutionary specialization has involved height and width of rays, even loss of rays, and changes in dimensions of the ray cells themselves. Activities of ray cells include the formation of tyloses—protrusions of parenchyma cells into the lumen of tracheary elements—and the secretion of variable quantities of gummy substances into neighboring cells.

A growth layer is the increment of secondary xylem produced in one growing season (**Fig. 6**). It is called an annual layer—or annual ring when seen in transection—if a single increment is formed each year. The boundaries between successive growth rings are identifiable largely because the xylem cells in the later part of one season (latewood or summerwood) are much smaller than the first ones produced in the next season's growth (earlywood or springwood). In ring-porous woods, the vessels are much larger in the earlywood than in the latewood of a given growth layer. The vessels are of approximately equal size throughout in diffuse-porous woods (**Fig. 7**).

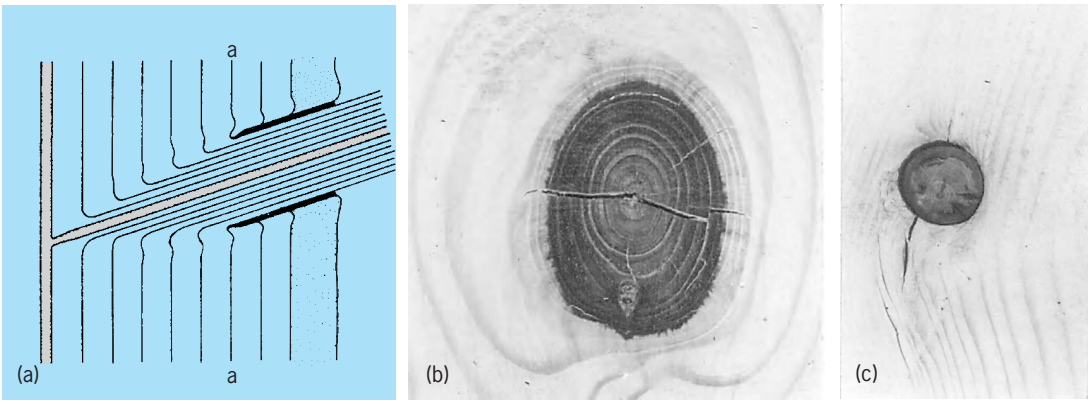

Fig. 8. Wood knots. (*a*) Diagram illustrating burial of base of a branch by secondary growth. The portion of the branch to the left of the line *a-a* was alive when buried, and its growth rings are continuous with those of the trunk. The portion to the right of *a-a* was dead when buried, and the growth rings of branch and trunk are not continuous. Pith finely stippled, bark coarsely stippled. (*b*) Hard pine board showing tight knot. (*c*) Hard pine board showing loose knot. (*After C. L. Wilson and W. E. Loomis, Botany, rev. ed., Dryden, 1957*)

Most knots are bases of twigs or branches buried, either alive (tight knot) or dead (loose knot), by cambial activity of the mother branch (**Fig. 8**). Their size depends largely upon the age at which they were buried. In lumber, round knots represent transections of the encased branch, and spike knots, transections of longitudinal sections.

Softwood and hardwood. In the trade, softwood is a name for xylem of gymnosperms (conifers) and hardwood for xylem of angiosperms. The terms do not refer to actual hardness of the wood. Woods of gymnosperms are generally composed only of tracheids, wood parenchyma, and small rays, but differ in detail (**Fig. 9**). Resin ducts are present in many softwoods. Woods of angiosperms show extreme variation in both vertical and horizontal systems, but with few exceptions have vessels (**Fig. 10**).

The grain, texture, and figure of wood reflect

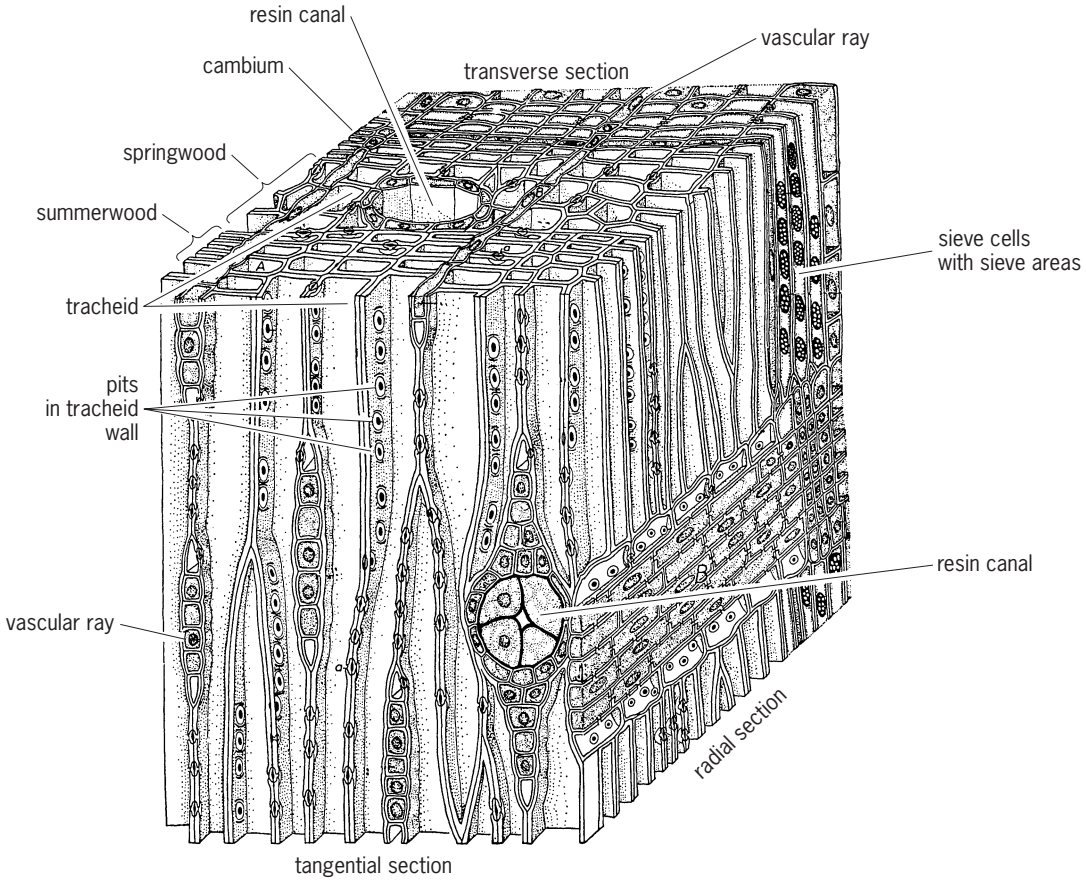

Fig. 9. Microscopic structure of pinewood (gymnosperm, three-dimensional). Gymnosperms are known as softwoods. (*After H. J. Fuller and O. Tippo, College Botany, rev. ed., Holt, 1954*)

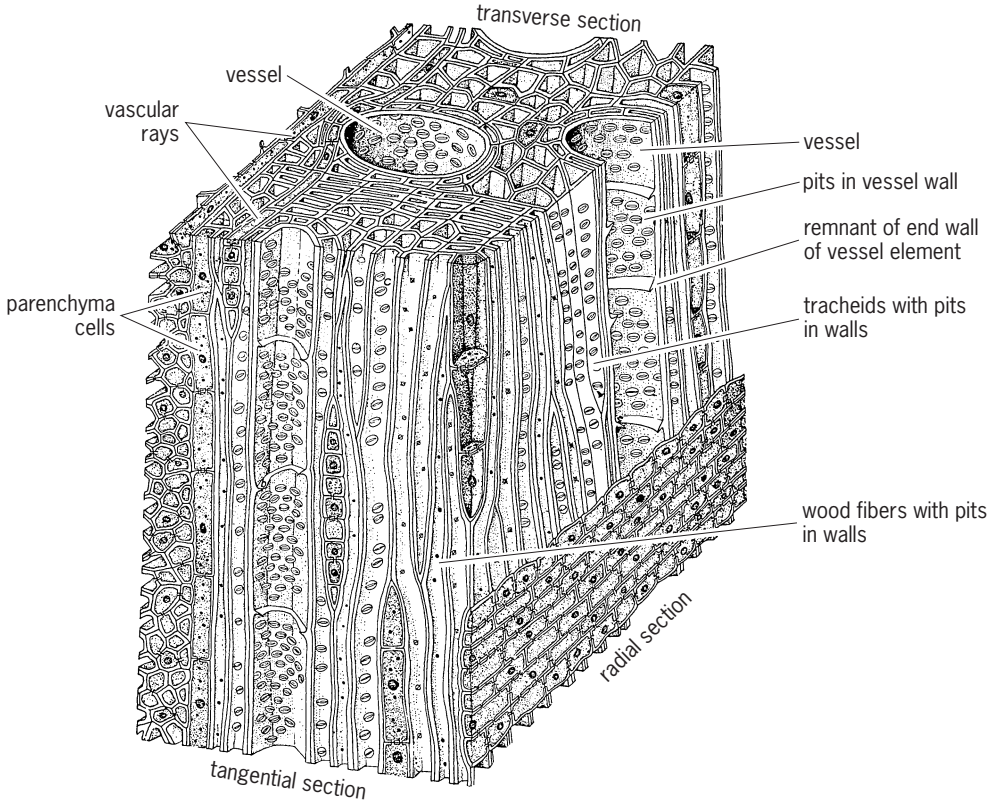

Fig. 10. Microscopic structure of oakwood (angiosperm, three-dimensional). Angiosperms are known as hardwoods. (*After H. J. Fuller and O. Tippo, College Botany, rev. ed., Holt, 1954*)

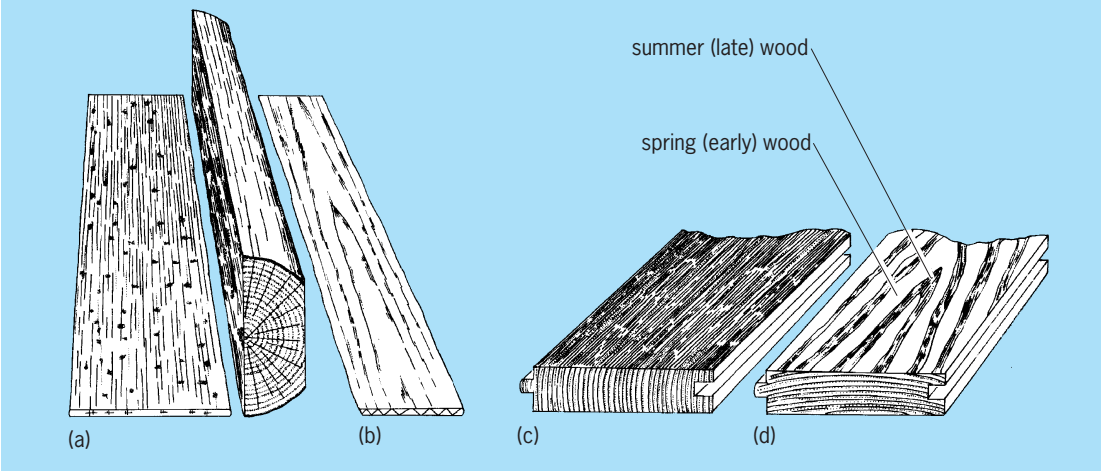

Fig. 11. Wood identification. (*a*) Quarter-sawed (radially sawed) and (*b*) plain-sawed (tangentially sawed) boards cut from a log (*after H. J. Fuller and O. Tippo, College Botany, rev. ed., Holt, 1954*). (*c*) "Edge-grain" flooring lumber quarter-sawed from fir trees. (*d*) Plain-sawed fir flooring, in which large uneven areas of springwood cause uneven wear and rapidly splintering surface (*after G. M. Smith et al., A Textbook of General Botany, 5th ed., Macmillan, 1953*).

gross differences in the arrangement, orientation, and kinds of cells in xylem and often are useful in identification (**Fig. 11**). These differences affect both beauty and strength of wood. Specific gravity also influences strength, as do the size, arrangement, and number of vessels, fibers, parenchyma cells, and rays. Heartwood, the older wood, is no stronger than the younger sapwood, although often it is more durable. Transforming sapwood into heartwood involves partial loss of water, death of living cells, production of various substances, often colored, and often the appearance of tyloses. *See* FOREST AND FORESTRY; PLANT TISSUE SYSTEMS; WOOD ANATOMY; WOOD PRODUCTS. Vernon I. Cheadle; Katherine Esau

Y-delta transformations — Yttrium

Y-delta transformations

Relationships between electrically equivalent networks with three terminals, one being connected internally by a Y configuration and the other being connected internally by a Δ configuration. These relationships are also known as star-delta transformations. More complicated relationships can also be derived between a many-armed star and the equivalent many-sided mesh.

Before considering these networks, consider the simpler two-terminal network. If the path consists of a network, however complicated, of passive linear elements, the input impedance is the ratio of the transform or phasor of input voltage to the transform or phasor of entering current. It is usually not difficult to find the input impedance of any passive, linear, two-terminal network, commonly by series and parallel combination of impedances. The input impedance will in general be a function of frequency. At one particular frequency it is possible to design a single element to have the same impedance, and this element can be said to be equivalent to the given network at that frequency, in the sense that the terminal relations of voltage and current will be the same (at that frequency) for either. In certain particular networks the equivalence is valid at all frequencies.

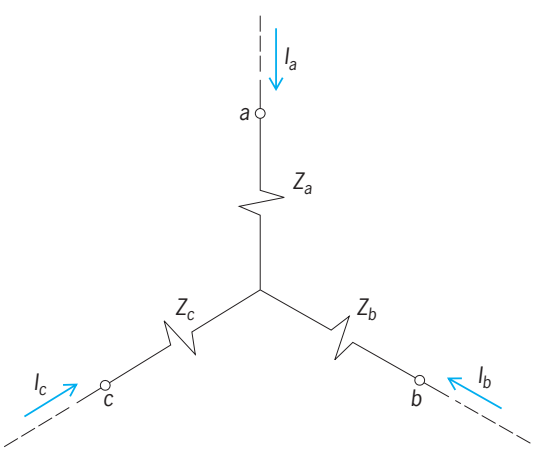

Fig. 1. Three-terminal star, or Y.

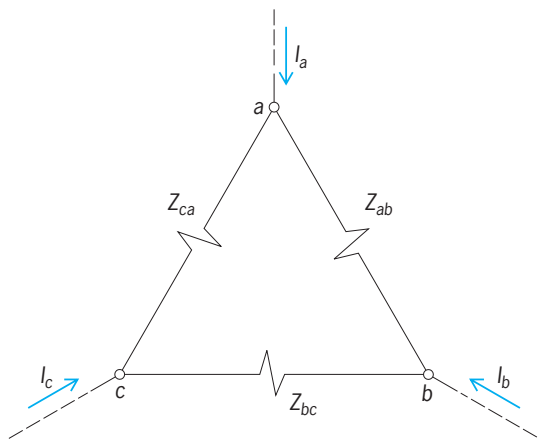

Fig. 2. Three-terminal mesh, or Δ.

If a network connects three terminals with one another, it is called a three-terminal network. The simplest configurations of three-terminal networks are the Y or star (**Fig. 1**) and the Δ or mesh (**Fig. 2**). [The star of three elements may also be called a T, and the mesh of three elements a π (**Fig. 3**).]

If a given three-terminal network, passive and linear, has a Y configuration, it is possible to design an equivalent Δ network that could be substituted for the Y without changing the relations of voltage and current at the network terminals, or elsewhere external to the network. Similarly, if a Δ network is given, an equivalent Y network can be found. Impedances of equivalent networks are usually functions of frequency, and realization of these impedances by use of physically possible elements is usually limited to a single frequency.

Derivation. If terminal-to-terminal voltages V_{ab}, V_{bc}, and V_{ca} are taken to be equal in the Y and Δ of Figs. 1 and 2, then the terminal currents I_a, I_b, and I_c are equal in the two networks if and only if the

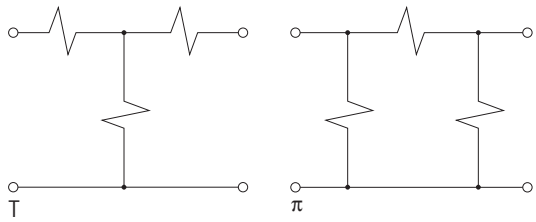

Fig. 3. Three-terminal star and mesh drawn as T and π.

impedance relations in Eqs. (1)–(3) are satisfied.

$$Z_a = \frac{Z_{ab}Z_{ca}}{Z_{ab} + Z_{bc} + Z_{ca}} \quad (1)$$

$$Z_b = \frac{Z_{bc}Z_{ab}}{Z_{ab} + Z_{bc} + Z_{ca}} \quad (2)$$

$$Z_c = \frac{Z_{ca}Z_{bc}}{Z_{ab} + Z_{bc} + Z_{ca}} \quad (3)$$

These are derived by expressing the three currents of each configuration in terms of the three voltages, equating, and solving simultaneously. Only one of these three equations is needed, for the subscripts a, b, and c are arbitrary (**Fig. 4**).

The same equivalence can be expressed in admittances instead of impedances by Eq. (4). Again, only

$$Y_a = \frac{Y_{ab}Y_{bc} + Y_{bc}Y_{ca} + Y_{ca}Y_{ab}}{Y_{bc}} \quad (4)$$

one equation is needed, for the subscripts are arbitrary.

To change Y to Δ, a simultaneous solution of Eqs. (1), (2), and (3) gives a set of relationships, Eqs. (5)–(7), to be used if the impedances or admittances of a Y are known and those of an equivalent Δ are wanted.

$$Z_{ab} = \frac{Z_aZ_b + Z_bZ_c + Z_cZ_a}{Z_c} \quad (5)$$

$$Y_{ab} = \frac{Y_aY_b}{Y_a + Y_b + Y_c} \quad (6)$$

$$Z_a = \frac{(20)(20)}{20 + 20 + j\omega 5} = \frac{400}{40 + j\omega 5} \text{ ohms} \quad (7)$$

For admittances, the equivalence is expressed by Eq. (6).

Example. The following example, though rather artificial, shows several interesting aspects of equivalence. Given the Δ (**Fig. 5**), with two sides purely resistive and the other side purely inductive, the problem is to find an equivalent Y. Given values are 20 ohms in each of two sides, and $L = 5$ henrys or $Z = j\omega 5$ ohms in the third side. From Eq. (1), relation (7) can be written. Similarly, from Eq. (2), relation (8) can be written. In this example $Z_c = Z_b$.

$$Z_b = \frac{(20)(j\omega 5)}{20 + 20 + j\omega 5} = \frac{j\omega 100}{40 + j\omega 5} \text{ ohms} \quad (8)$$

The impedances of the equivalent Y can be computed at any particular frequency; for example, at 60 cycles per second, or 60 Hz, $\omega = 377$. At low frequency, Z_a approaches 10 ohms of resistance, but for all finite frequencies it is somewhat less than 10 ohms and is more or less capacitive. Z_b and Z_c are equal because of symmetry in the original Δ network; at low frequency they have low inductive values, approaching zero, whereas at high frequencies each approaches 20 ohms of resistance.

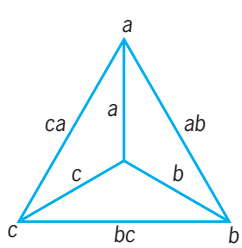

Fig. 4. Diagram to show the pattern of subscripts used in equations.

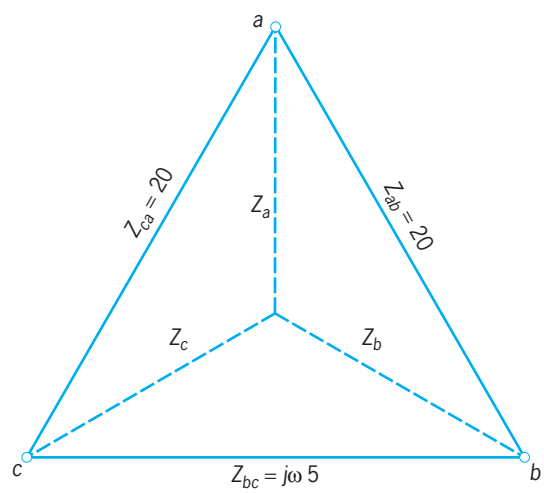

Fig. 5. Example of a Y that is equivalent to a given Δ.

At any particular frequency, each of the Y impedances can be physically realized as resistance in series with either a capacitance or an inductance, but physical realization is not simple over any range of frequencies.

Difficulty with physical realization, however, does not invalidate the mathematical equivalence, and the Y network that has here been determined can be substituted for the given Δ network for all analytical or computational purposes. *See* ALTERNATING-CURRENT CIRCUIT THEORY; NETWORK THEORY.

Hugh H. Skilling

Bibliography. R. DeCarlo and P.-M. Lin, *Linear Circuit Analysis*, 2d ed., 2001; J. D. Irwin and R. M. Nelms, *Basic Engineering Circuit Analysis*, 8th ed., 2005.

Yagi-Uda antenna

An antenna in which the gain of a single dipole element is enhanced by placing a reflector element behind the dipole (the driver) and one or more

director elements in front of it (see **illus.**). Invented in 1926 by H. Yagi, and S. Uda in Japan, the Yagi-Uda antenna is a long-time favorite of amateur radio enthusiasts. The gain is slightly increased by the reflector and further enhanced by the first director element. Additional director elements further increase the gain and improve the front-to-back ratio, up to a point of diminishing returns. This type of antenna has traditionally been used for local television reception. Its variants have found applications in the more modern communication systems at higher frequencies and smaller sizes, and have even been adapted to printed-circuit techniques in some applications. The same electromagnetic induction principle used in such linear elements can be applied to loop and disk elements as well with similar results. *See* DIRECTIVITY; ELECTROMAGNETIC INDUCTION; GAIN; PRINTED CIRCUIT.

Although the resonant half-wave dipole driver exhibits bandwidth limitation, the Yagi-Uda antenna can be considered as a traveling-wave antenna, and therefore various combinations of element lengths, separations, and other factors can achieve a range of phase velocities. The traveling-wave feature of the Yagi-Uda antenna can often provide a higher usable bandwidth than the equivalent resonant antenna design. The primary design objective is therefore to select the number of directors and the length and separation of the driver, the directors, and the reflector so as to optimize the antenna gain and improve the front-to-back ratio. This task has been rather tedious, but computer-aided techniques have been developed that can help provide an optimal design for a given gain and frequency. These techniques include method of moment (MOM) and finite difference time domain (FDTD).

The physical dimensions for the driver are normally about a half wavelength or less. The reflector is usually slightly larger than the driver and positioned about a quarter of a wavelength behind it. The directors are slightly shorter than the driver and positioned at distances of approximately a quarter of a wavelength in front. The design for one frequency can usually be scaled up or down for lower and higher frequencies. For comparison, a typical television antenna can be measured in feet (1 ft = 0.3 m), while the same antenna at 2.4 GHz, the frequency of the wireless local-area network, measures in inches (1 in. = 2.5 cm). The **table** shows some typical performance data for the traditional Yagi-Uda antenna. *See* LOCAL-AREA NETWORKS.

Typical performance characteristics of a Yagi-Uda antenna

Number of elements	Gain dBi*	Beam width in E-plane†	Beam width in H-plane†
2	5	>66°	>84°
3	9	66°	84°
4	10	52°	60°
5	11	58°	66°

*Power gain, in decibels, in the direction of maximum radiation, compared to the radiation from an isotropic antenna in free space receiving the same amount of power.
†Full width of the primary lobe of the radiation pattern in the specified plane at half the maximum power.

Since these antennas can be made highly directive with good radiation efficiency, they have found new applications and new manufacturing techniques with miniaturization. They can be printed on microwave circuit substrates with high dielectric constants, which reduces their size even further. The parasitic electromagnetic coupling demonstrated in the Yagi-Uda antenna has been adapted to many new types of miniaturized antennas applicable to mobile communication devices in wide use, and will be used in future wireless Internet devices. *See* ANTENNA (ELECTROMAGNETISM); INTERNET; MICROWAVE SOLID-STATE DEVICES; MOBILE RADIO.

Jamal S. Izadian

Bibliography. R. C. Johnson, *Antenna Engineering Handbook*, 3d ed., McGraw Hill, New York, 1993; W. L. Stutzman and G. A. Theile, *Antenna Theory and Design*, 2d ed., Wiley, New York, 1997.

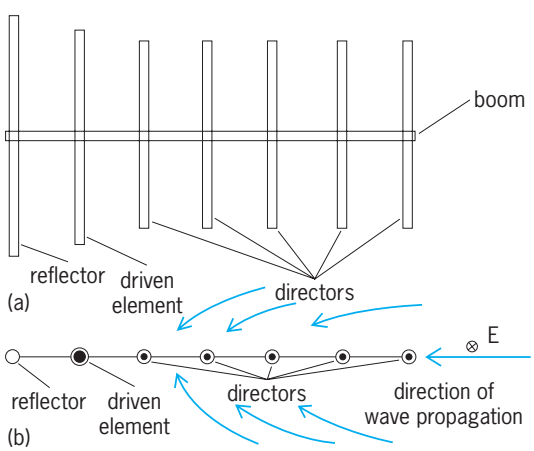

Yagi-Uda antenna. Typical radiation patterns are shown. (*a*) Side view with radiation pattern in the E plane (the plane containing the electric field). (*b*) Top view with radiation pattern in the H plane (the plane containing the magnetic field).

Yak

A member of the mammalian cattle family Bovidae in the order Artiodactyla. It is the least-known species of wild cattle alive today. Yaks (*Bos grunniens*) have been domesticated for at least several hundred years. They are powerful yet docile. Above 2000 m (6600 ft), they are the most useful domestic mammal. They are used for meat and milk production, as a mount, and as a beast of burden. The hair is used for making cloth and tents, the dung is used as fuel, and the tail serves as a flybrush.

Morphology. The yak, a relative of the bison, is a blackish-brown bovid with large black horns that grow out from the sides of the head and curve upward (see **illustration**). The horns, which are present in both sexes, are composed of bony cores attached to the frontal bones of the skull and hard

Adult male domestic yak (*Bos grunniens grunniens*). (*Copyright © Brent Huffman, 2000*)

sheaths of horny material. The massive body is covered by long guard hairs that almost reach the ground and a dense underfur of soft hair. The guard hairs form a shaggy fringe of coarse hair around the lower part of the body, the shoulders, flanks, and thighs. The shaggy muzzle may be white. The tail also consists of long hairs. The shoulders are high and humped, with a broad, drooping head. The limbs are short and sturdy with broad hooves and large dewclaws as an adaptation to mountainous environments. Males may attain a head and body length up to 325 cm (10.8 ft), a shoulder height up to 200 cm (6.6 ft), and a weight of 305–820 kg (670–1805 lb). Females are smaller and may weigh only one-third as much. Domesticated animals are considerably smaller than wild yaks and may be red, brown, or black in coloration.

Habitat and life history. Yaks are found in Tibet, China, India, Kashmir, and Nepal at elevations of 4000–6000 m (12,800–19,200 ft). They inhabit desolate steppe regions of alpine tundra and cold desert regions with swamps and moors. They are surefooted and are expert climbers. They spend a portion of the summer in the lower plains, but as the temperature rises they move to the higher plateaus. They swim and bathe in icy cold lakes and rivers. Yaks feed on grasses, herbs, mosses, and lichens. Since vegetation is very sparse, yaks must travel great distances from one grazing site to another. For most of the year, bulls live either a solitary existence or they live in small groups, whereas females and their young live in large herds of up to 200 animals. Breeding occurs in September and October. Following a gestation period of approximately 9 months, a single calf is born in June. Females produce a calf every other year because the nursing period lasts an entire year. Yaks reach their adult size in 6 to 8 years and live for about 25 years in the wild. Enemies include humans and packs of Tibetan wolves, which hunt solitary animals and calves that are separated from the herd.

The wild yak is classified as vulnerable by the World Conservation Union, is listed as endangered by the United States Department of the Interior, and is on Appendix 1 of CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora). Yaks are officially protected in China but are declining due to uncontrolled hunting. Organizing protective measures in the inaccessable mountain regions of Tibet is almost impossible; thus, there is little hope of saving the remaining wild yak populations. *See* ARTIODACTYLA; BISON.

Donald W. Linzey

Bibliography. *Grzimek's Encyclopedia of Mammals*, vol. 5, McGraw-Hill, 1990; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Yaws

An infectious disease of humans caused by the spirochete *Treponema pertenue*. It is also known as frambesia and is largely confined to the tropics. Anecdotal information indicates that yaws still exists in parts of central Africa, South America, India, and Indonesia. Usually yaws is contracted in childhood by direct contact or from small flies feeding in succession on infected lesions and open wounds. No race or age possesses natural immunity.

Treponema pertenue is morphologically similar to *T. pallidum*. The general course of yaws resembles that of syphilis, but the individual yaws lesions tend to be larger and more persistent. The primary lesion or mother yaw occurs commonly on the lower extremities. Generalized lesions are often large and fungating. Destructive bone lesions are common, as are painful incapacitating lesions of the soles of the feet (crab yaws). Central nervous system and cardiovascular involvement occurs rarely if at all. Serological tests for syphilis are also positive in yaws, and the disease responds well to penicillin.

There are indications that while *T. pertunue* and *T. pallidum* are biologically and immunologically related, significant and persistent differences occur to justify distinct names for the two diseases. The relationship of yaws to other treponematoses, such as pinta, endemic syphilis, and bejel, has not been clearly determined. Mass survey and treatment campaigns directed primarily to the reduction of infectious cases have been conducted with success in many countries. The wide availability of penicillin since World War II has encouraged public health programs aimed at the reduction of the incidence of yaws, although much of the evidence remains anecdotal. Improvement in the quality of the environment in many tropical areas has also contributed to a reduction in incidence, although some foci of endemic yaws persist there. *See* SEROLOGY; SYPHILIS.

Thomas B. Turner

Bibliography. A. B. Koff and T. Rosen, Non-venereal treponematoses: Yaws, endemic syphilis and pinta, *J. Amer. Acad. Dermatol.*, 29(4):519–535, 1993; V. N. Sehgal et al., Yaws control/eradication, *Int. J. Dermatol.*, 33(1):16–20, 1994; T. B. Turner and D. W. Hollander, *Biology of the Treponematoses*, 1957.

Year

The period of the Earth's revolution around the Sun. This period can be measured in several ways.

Types of years. The most commonly used type of year, known as the tropical year, is the mean period for the Sun to go around the sky from one vernal equinox to the next. The tropical year varies slightly over time. Tropical year 2000 was approximately 365.242190 mean solar days, or 365 d 5 h 48 m 45.2 s. Seasons repeat with intervals of a tropical year.

The sidereal year is the average period of the Earth's revolution around the Sun until the Earth-Sun line returns to the same direction in space as measured by the star background. A sidereal year is approximately 365.25636 mean solar days long, or 365 d 6 h 9 m 10 s.

The anomalistic year is the average period of the Earth's return to perihelion in its orbit around the Sun. An anomalistic year is approximately 365.25964 days, or 365 d 6 h 13 m 53 s.

An eclipse year is the duration between alternate eclipse seasons, measured as the interval between the Moon passing alternate nodes. An eclipse year lasts only 346.62005 days, or 346 d 14 h 52 m 52 s.

A Gaussian year, defined as the period derived from Kepler's law for 1 astronomical unit, is 365.25690 days, or 365 d 6 h 9 m 56 s. *See* ASTRONOMICAL UNIT; KEPLER'S LAWS.

A Besselian year is based on the Besselian elements used to compute eclipses. *See* EARTH ROTATION AND ORBITAL MOTION; ECLIPSE.

Counting of years. The counting of years used through most of the world incorporates A.D. and B.C. for Anno Domini and Before Christ (or BCE, for Before Christian Era). This practice dates from the sixth-century scholar Dionysius Exiguus (Dennis the Small). The date of the birth of Jesus, however, is now believed to be several years before A.D. 1. The Venerable Bede, in the eighth century, defined 1 B.C. as the year before A.D. 1 and then counted backward from 1 B.C. This system, however, omits a year 0, so care must be taken in computing intervals that span dates from A.D. to B.C. The problem led to the invention in 1583 by Joseph Justus Scaliger of Julian Dates (J.D.), counting individual days from 4713 B.C. Julian Dates begin at noon, so half-integral Julian Dates correspond to midnight, when other calendar dates begin. January 0.5 in Julian year 2000 corresponds to J.D. 2451544.0. Besselian year 2000.0 corresponds to J.D. 2451544.533. Astronomical catalogs are now given for the year 2000, called J2000.0 to signify centuries of exactly 36525 days counted from January 0.5, 1900, corresponding to January 1.5, 2000, or J.D. 2451545.0.

The Julian year of exactly 365.25 days leads to a drift in the calendar, which led to the Gregorian calendar. This Gregorian calendar introduced the system of leap years that is still in use. *See* CALENDAR.

Extraterrestrial years. The term "year" is also used for the period of the orbit of other planets, both inside the solar system and around other stars. For example, a Jupiter year is 5.2 Earth years long. One of the planets around a distant star orbits in only 2 Earth days, so its year is only 2 Earth days long. *See* TIME.

Jay M. Pasachoff

Bibliography. P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomical Almanac*, 1992.

Yeast

A collective name for those fungi which possess, under normal conditions of growth, a vegetative body (thallus) consisting, at least in part, of simple, single cells. The cells making up the thallus occur in pairs, in groups of three, or in straight or branched chains consisting of as many as 12 or more cells. Vegetative reproduction is characterized by budding or fission. Sexual reproduction also occurs in yeast, and is differentiated from that of other fungi by sexual states that are not enclosed in a fruiting body. Yeasts are a phylogenetically diverse group of organisms that occur in two divisions of fungi (Ascomycotina and Basidiomycotina) and 100 genera. The 700 or more species that have been described possibly represent only 1% of the species in nature, so the majority of the yeasts have yet to be discovered. Yeast plays a large part in industrial fermentation processes such as the production of industrial enzymes and chemicals, food products, industrial ethanol, and malt beverage and wine; in diseases of humans, animals and plants; in food spoilage; and as a model of molecular genetics. *See* DISTILLED SPIRITS; MALT BEVERAGE; WINE.

Characteristics

The shape and size of the individual cells of some species vary slightly, but in other species the cell morphology is extremely heterogeneous. The shape of yeast cells may be spherical, globose, ellipsoidal, elongate to cylindrical with rounded ends, more or less rectangular, pear-shaped, apiculate or lemon-shaped, ogival or pointed at one end, or tetrahedral (**Fig. 1**). The diameter of a spherical cell may vary

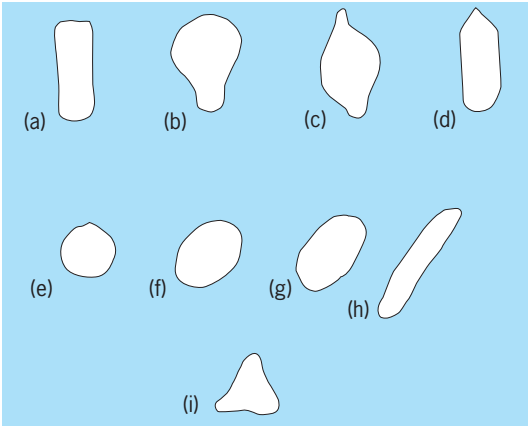

Fig. 1. Cell shapes of yeasts are extremely varied. (*a*) Rectangular. (*b*) Pear-shaped. (*c*) Apiculate. (*d*) Ogival. (*e*) Spherical. (*f*) Globose. (*g*) Ellipsoidal (oval). (*h*) Elongate. (*i*) Tetrahedral.

from 2 to 10 micrometers. The length of cylindrical cells is often 20–30 μm and, in some cases, even greater.

The asexual multiplication of yeast cells occurs by a budding process, by the formation of cross walls or fission, and sometimes by a combination of these two processes. Yeast buds are sometimes called blastospores or blastoconidia. When yeast reproduces by a fission mechanism, the resulting cells are termed arthrospores or arthroconidia. Some yeasts reproduce by cells known as ballistoconidia, which form on the parent cell at the tips of spear-shaped projections known as sterigmata. The ballistoconidia are forcefully ejected by a surface-tension catapult mechanism at an acceleration of 25,000 *g*, which is about 10,000 times the acceleration during the launch of the space shuttle. Whereas the shuttle consumes 50% of its weight in fuel in the first 2 minutes of flight, ballistoconidia discharge is fueled by carbohydrates that represent only 1% of the mass of the conidia.

In some genera of yeast, true septate mycelium is formed, for example, in some species of *Candida* and always in *Trichosporon*. In *Candida*, the mycelium, if present, remains intact; that is, it does not break up at the septa. In *Trichosporon*, the mycelium breaks apart into arthrospores which at first are rectangular but often develop rounded ends when mature (**Fig. 2***a*). Many yeasts produce a type of thallus which is termed a pseudomycelium. In a pseudomycelium the component cells are formed by elongation of buds and not by the formation of septa. Usually a group of characteristically arranged blastospores develops at the junctions of cells making up a pseudomycelium (Fig. 2*b*).

Yeasts are also characterized by certain macromorphological features. These are studied as slant cultures and as giant colonies. Slant cultures are made by inoculating the yeast as a thin line on the center of an agar slope in a medium which may be malt extract or synthetic. After several weeks of growth the culture may be described with respect to color (for example, cream, pink, yellow, brown, black, white, and intermediate shades); surface (smooth, wrinkled, warty, glossy, dull); texture (pasty, slimy or mucilaginous, tough); cross section (flat, raised, hemispherical, extent of spreading); and border (entire, hairy, irregular). Giant colonies are usually inoculated on malt gelatin, by a light, pinpoint inoculation. Macromorphological features usually develop more characteristically on gelatin than on agar media (**Fig. 3**). The description of the growth after about 3 weeks is similar to that used for the description of streak cultures.

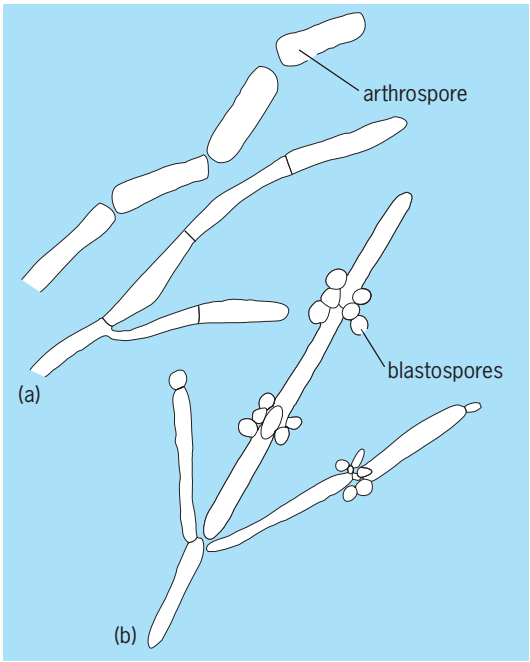

Fig. 2. Yeast mycelia. (*a*) True septate showing break to form arthrospore. (*b*) Pseudomycelium.

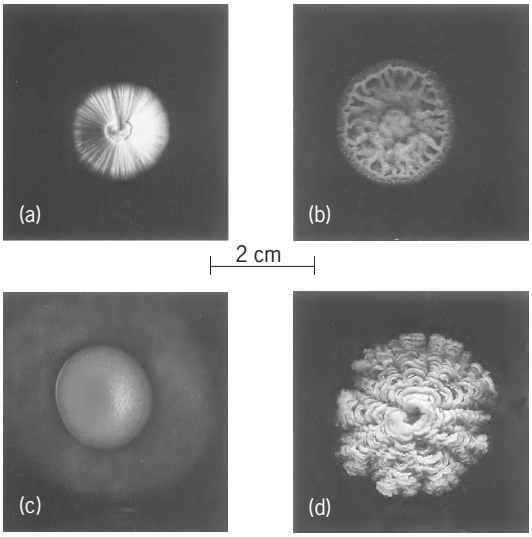

Fig. 3. Giant yeast colonies: (*a*) *Zygosaccharomyces rouxii*, (*b*) *Sporobolomyces pararoseus*, (*c*) *Rhodotorula glutinis*, and (*d*) *Saccharomyces cerevisiae*, all after 6 weeks of growth on wort gelatin at 68°F (20°C).

Cytology. Yeast cells are surrounded by a wall, which in the case of bakers' yeast and many other species consists of polysaccharides such as glucan and mannan, a small amount of chitin, protein, lipids, and minerals. The filamentous yeasts have a higher chitin content than the budding yeasts. Some species contain as yet unidentified components. With the electron microscope, bud scars can be observed in the walls of yeast. Successive buds, 20–40 from one bakers' yeast cell, are always formed at different places on the cell surface. A cell also contains a birth scar, which differs in appearance from a bud scar. Yeast cells are generally uninucleate. The cytoplasm may contain one or more vacuoles. In addition, the cytoplasm may contain lipid globules, volutin (polyphosphate) granules, mitochondria, and submicroscopic particles (microsomes). When yeast cells are treated with iodine, they usually stain deep brown, because of their high glycogen content.

Sexual reproduction. Yeasts are categorized into two groups, based on their methods of sexual reproduction: the ascomycetous (Division Ascomycotina)

and basidiomycetous (Division Basidiomycotina) yeasts. *See* ASCOMYCOTA; BASIDIOMYCOTA.

Ascospores. The sexual spores of the ascomycetous yeasts are termed ascospores, which are formed in simple structures, often a vegetative cell. Such asci are called naked asci because of the absence of an ascocarp, which is a more complex fruiting body found in the higher Ascomycetes. If the vegetative cells are diploid, a cell may transform directly into an ascus after the $2n$ nucleus undergoes a reduction or meiotic division. *See* MEIOSIS.

In the case of haploid yeasts, a conjugation or fusion between two cells normally occurs. The fusion of the cytoplasmic contents of the two cells is called plasmogamy. This is followed by fusion of the two haploid nuclei, or karyogamy. Meiosis then follows. Each of the haploid nuclei resulting from meiosis gathers cytoplasmic material around it and finally develops into a characteristically shaped ascospore. The spores in such a process are formed by "free cell formation," in contrast to spore delineation by cleavage of the cytoplasm. The ascospores may be spherical, oval, kidney- or crescent-shaped, hat-shaped, helmet-shaped, hemispherical, needle-shaped, walnut-shaped, or saturn-shaped (**Fig. 4**).

The number of ascospores per ascus varies with the species. Some contain a single spore, others one or two, two to four, or four to eight. In rare cases, more than eight spores are formed. In some yeasts, the asci rupture soon after maturation, and the spores are liberated into the medium. In others, the spores remain in the ascus until they germinate. The heat resistance of ascospores is only slightly greater than that of the vegetative cells. The number and shape of the ascospores are important characters in generic descriptions. The species of most genera have only one type of ascospore (such as the smooth spheroidal type in *Saccharomyces*), but sometimes more than one shape is found among the species of a genus, for example, in *Kluyveromyces*, *Hanseniaspora*, and *Saccharomycopsis*.

The ease with which yeasts form ascospores varies enormously with the species. In some, an abundance of ascospores is formed on nearly all media of growth. In others, special sporulation media must be used, and often the percentage of asci is small. Another complication is that certain yeasts lose the ability to sporulate upon prolonged cultivation in pure culture.

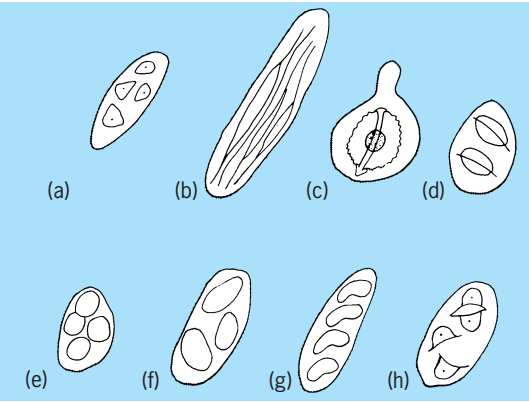

Fig. 4. Shapes of ascospores of various yeasts. (*a*) Helmet. (*b*) Needle. (*c*) Walnut. (*d*) Saturn. (*e*) Spherical. (*f*) Oval. (*g*) Kidney or crescent. (*h*) Hat.

Anamorphic and teleomorphic yeasts. Certain yeasts have been shown to be heterothallic; that is, sporulation occurs when strains of opposite mating type (usually indicated by a and α) are mixed on sporulation media. However, some strains may be homothallic (self-fertile), and reduction division and karyogamy take place during formation of the sexual spore. Yeasts that produce sporogenous cells represent the teleomorphic form of the life cycle. In cases, in which sexual cycles are unknown, the yeast represents the asexual or anamorphic form. A species of yeast may be originally discovered in the anamorphic form and named accordingly; subsequently, the sexual state may be found and a name applied to represent the teleomorph. Consequently, the anamorphic and teleomorphic names will differ. For example, the human pathogen *Cryptococcus neoformans* was originally found in the anamorphic form; several decades later the sexual cycle was discovered, and the teleomorphic form *Filobasidiella neoformans* was described. However, not all teleomorphic species have anamorphic counterparts. For example, the teleomorphic beer and bread yeast *Saccharomyces cerevisiae* was described in 1883, and anamorphic names for it are generally not recognized.

Among the majority of the yeast species, the teleomorphic forms are unknown; therefore, a separate classification scheme was created for the anamorphic yeasts, and orders such as Cryptococcales and Sporobolomycetales were erected. However, with the advent of molecular sequence analysis and the phylogenetic interpretation of the data, the relationships between anamorphic and teleomorphic species are becoming established. Consequently, the anamorphic species are being classified with the related teleomorphs, and the use of anamorphic orders is being eliminated.

Basidiospores and teliospores. These sexual spores are produced in the three classes of basidiomycetous yeasts: Urediniomycetes, Hymenomycetes, and Ustilaginomycetes. Sexual reproduction and life cycle in these yeasts is typical of other basidiomycetes in that it can include both unifactorial (bipolar) and bifactorial (tetrapolar) mating systems. The Urediniomycetes include species of widespread occurrence in the teleomorphic genera (*Leucosporidium*, *Rhodosporidium*, and *Sporidiobolus*) and the anamorphic genera (*Bensingtonia*, *Rhodotorula*, and *Sporobolomyces*). Reproduction in this class can include the formation of teliospores (**Fig. 5**). Conjugation between strains of opposite mating types (for example A1B1 × A2B2) results in dikaryotic mycelia with clamp connections. Karyogamy occurs in the heavy-walled teliospore, which germinates to produce a two- to four-celled septate metabasidium. Meiosis takes place during germination, and the resulting basidiospores are haploid.

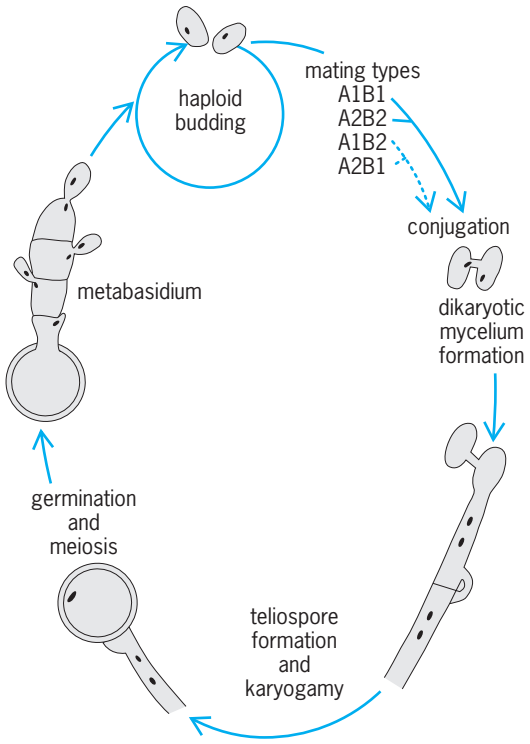

Fig. 5. Heterothallic (bifactorial) life cycle of the basidiomycetous yeast *Leucosporidium scottii*.

The Hymenomycetes include genera such as *Cryptococcus*, *Trichosporon*, *Bullera*, *Filobasidiella*, and *Filobasidium*. Some of these species are related to the jelly fungi, and others are well-known human pathogens. Sexual cycles among these yeasts are similar to the sexual cycles of the Urediniomycetes, except that most species do not produce teliospores, and the metabasidium is often an elongate single cell with an enlarged circular terminal head that bears the basidiospores.

The Ustilaginomycetes include the smut fungi. Yeasts in this class include species of *Malassezia*, *Pseudozyma*, and *Rhodotorula*. Sexual cycles among these yeasts are not well known.

Physiology. Some yeasts have the ability to carry out an alcoholic fermentation. Other yeasts lack this property. In addition to the fermentative type of metabolism, fermentative yeasts as a rule have a respiratory type of metabolism, whereas nonfermentative yeasts have only a respiratory, or oxidative, metabolism. Fermentation may be represented by reaction (1), respiration of glucose by reaction (2).

$$C_6H_{12}O_6 \rightarrow 2C_2H_5OH + 2CO_2 \qquad (1)$$

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O \qquad (2)$$

Both reactions produce energy, with respiration producing by far the most, which is used in part for synthetic reactions, such as assimilation and growth. Part is lost as heat. In addition, small or sometimes large amounts of by-products are formed, including organic acids, esters, aldehydes, glycerol, and higher alcohols. When a fermenting yeast culture is aerated, fermentation is suppressed and respiration increases. This phenomenon is called the Pasteur effect.

Fermentation. Yeasts are best known for their ability to ferment mono- or disaccharides. In the case of disaccharides, trisaccharides, and polysaccharides, hydrolysis by hydrolytic enzymes, or hydrolases, must precede fermentation. There is evidence that some hydrolytic enzymes are located in the peripheral layer of the cell. Such hydrolases may be secreted into the medium as extracellular enzymes. Species of yeast are often differentiated on the basis of the various hydrolases which they possess. *See* CARBOHYDRATE.

Some simple rules of fermentation have been formulated. If a yeast cannot ferment glucose, it cannot ferment any other sugar. If a yeast can ferment glucose, it can also ferment fructose and mannose (although often at different rates), but not always galactose. If a yeast ferments maltose, it usually cannot ferment lactose and vice versa. *See* FERMENTATION.

Another mechanism of alcoholic fermentation has been discovered with the pentose sugar D-xylose as a substrate. Pentose sugars used to be thought of as nonfermentable, but it is now known that some yeast species (for example, *Pachysolen tannophilus* and *Candida tropicalis*) can convert D-xylose aerobically by a complex process, involving isomerization and phosphorylation, to fructose-6-phosphate which is fermentable by the classical pathway. The alcohol yields, however, are considerably lower than those obtained from glucose by conventional yeasts (*Saccharomyces cerevisiae*).

Respiration and assimilation. Yeasts differ greatly in their ability to respire and to assimilate organic substrates. Depending on the species, yeasts can utilize such compounds as pentoses (D-xylose, D-arabinose, L-arabinose, D-ribose), methyl pentoses (L-rhamnose), sugar alcohols (mannitol, sorbitol, dulcitol, erythritol, adonitol), organic acids (lactic, acetic, succinic, citric, gluconic, 2-ketogluconic, 5-ketogluconic, and many other acids), polysaccharides (soluble starch, inulin), *i*-inositol, and even such compounds as phenol and related ones.

The source of nitrogen may be organic or inorganic. All yeasts, with rare exceptions, can use ammonium ion for the synthesis of proteins. Others can utilize nitrate, and still others can use nitrite but not nitrate. Yeasts also can absorb intact amino acids from the medium. There is considerable variation in the ability of yeasts to deaminate individual amino acids. Lysine, as a single nitrogen source, is used by few yeasts. Glutamic acid is utilized by nearly all yeasts. Sulfur is ordinarily supplied as sulfate, although some yeasts grow better when the sulfur is supplied as cysteine or methionine. The following mineral salts and concentrations are required by yeasts in a medium: monobasic potassium phosphate, KH_2PO_4, 0.1%; magnesium sulfate, $MgSO_4$, 0.05%; sodium chloride, NaCl, 0.01%; and calcium chloride, $CaCl_2$, 0.01%. In addition, the following trace elements should be added to a synthetic medium: boron, copper, zinc, iron, manganese, molybdenum, and iodine.

Special synthetic products are carotenoid pigments, produced by species of *Rhodotorula* and *Sporobolomyces* (β-carotene is responsible for a yellow color; torulin and torularhodin are responsible for pink colors). Some species of the genus *Cryptococcus* produce a starchlike polysaccharide in acidic media. In addition, species of *Cryptococcus*, *Rhodotorula*, and some other yeasts produce capsules, consisting of various polysaccharides.

Ecology. Yeasts are ubiquitous in nature. They exist on plants and animals; in waters, sediments, and soils; and in terrestrial, aquatic, and marine habitats. Yeasts require oxygen for growth and reproduction; therefore they do not inhabit anaerobic environments such as anoxic sediments. Many species have highly specific habitats, whereas others are found on a variety of substrates in nature.

Soil. Soil may be considered as a reservoir of yeasts. Active growth in soil occurs only under favorable conditions, such as those found in fruit orchards and meadowlands. However, small numbers of a great variety of yeasts are found in many types of soil. Yeasts belonging to the genera *Lipomyces* and *Schwanniomyces* have been isolated only from soil.

Trees. Other yeasts have their natural habitat in the exudates or slime fluxes of trees. The yeasts found in exudates of coniferous trees generally differ from those in exudates of deciduous trees. Species of *Nadsonia* have been found only in tree exudates; certain species of *Hansenula*, *Saccharomycodes*, *Pichia*, *Saccharomyces*, and *Endomyces* are specifically associated with certain trees.

Insects and mammals. Since insects, such as *Drosophila*, frequently use sap exudates for breeding and sometimes for feeding purposes, yeasts can be found in appreciable numbers in the intestinal tract of such insects. Bark beetles, which live in the cambium layer of coniferous as well as deciduous trees, and the wood-boring ambrosia beetles also have been shown to have a specific yeast flora associated with them. Many species of yeast have been found in nature only in association with certain insects. Yeasts are rapidly digested by insects, and in this way they form an important part of their diet. The nectar of flowers is another habitat for yeasts. Those flowers, which are much frequented by bees, are especially likely to contain yeast. The intestinal tract of warm-blooded animals is also a habitat for yeasts. In some cases the yeast is so dependent on its host that it has lost the ability to grow at room temperature and in the common media used for propagation of yeasts. An example is the monospecific genus *Cyniclomyces*, which has been found only in the intestinal tract of rabbits. *Cyniclomyces guttulatus* grows only at about 99°F (37°C) and in very complex media. Interestingly, in the presence of about 15% free carbon dioxide the minimum temperature for growth is lowered to 86°F (30°C).

Pathogens. Whereas most yeasts associated with warm-blooded animals are nonpathogenic, a number of at least potentially pathogenic yeasts are known. With the expanding use of antibiotics, the natural balance of microbes in the gastrointestinal tract of mammals can be disturbed, and yeast infections are on the increase. The most common locations of infections by yeasts are the skin, especially the mucosa, and the respiratory tract, but occasionally systemic infections occur. *See* MEDICAL MYCOLOGY.

Plant pathogenic yeasts are limited to a few species. These are *Nematospora coryli*, a yeast with needle-shaped ascospores, and a yeastlike fungus, *Ashbya gossypii*, producing similar ascospores. The diseases which result from infection by *N. coryli* are yeast spot in lima beans and soybeans; stigmatomycosis of the fruits of tomatoes, citrus fruits, and others; and discoloration of cotton bolls. The cotton plant is also the principal host for *A. gossypii*. The yeast is transmitted by species of a number of genera of insects with piercing-sucking mouthparts as found in the bug group. Lastly, aquatic species of *Metschnikowia* are known as pathogens of the fresh-water crustacean *Daphnia magna* and the brine shrimp *Artemia salina*. *Metschnikowia bicuspidata* is the principal species that has caused economic losses in the brine shrimp industry. Infection and virulence may be related to the forceful ejection of the needle-shaped ascospores that pierce the intestinal wall of the host. Unidentified species of *Metschnikowia* also have been implicated in diseases of copepods during certain seasons of the year. *See* PLANT PATHOLOGY.

Food spoilage. Yeasts are important as spoilage organisms of a great variety of foods. Specific types occur, depending primarily on the composition of the food. Sugar-tolerant, or osmophilic, yeasts occur on dried fruits, syrups, honey, and the like. Examples are *Zygosaccharomyces rouxii*, *Schizosaccharomyces octosporus*, and the yeastlike fungus *Eremascus albus*. Species of *Debaryomyces* are highly salt-tolerant and occur in brines and on processed meats, such as bacon, hams, and sausage. Dairy products may be spoiled by lactose-fermenting yeasts. Some species of *Torulopsis* and of *Zygosaccharomyces* (haploid forms of *Saccharomyces*) may spoil salad dressing and similar products because of their great tolerance to vinegar. *Saccharomycopsis fibuligera* forms starch-splitting enzymes and can cause spoilage of cereal products and grains. *See* FOOD MICROBIOLOGY.

Jack W. Fell; Herman J. Phaff

Industrial Yeast

Yeasts are the most exploited group of industrial microorganisms. *Saccharomyces cerevisiae* (baker's or brewer's yeast) has been used for centuries in the production of fermented beverages and baked foods. In recent years, the industrial importance of *S. cerevisiae*, together with non-*Saccharomyces* yeasts, has extended beyond traditional food fermentation applications into the production of a wide range of important industrial commodities. **Table 1** summarizes the applications of several yeast species in the pharmaceutical, chemical, food, and environment sectors of biotechnology. There is great potential to further exploit the industrial power of yeasts,

TABLE 1. Some industrially important yeasts

Yeasts	Uses
Candida spp (for example, *C. utilis, C. maltosa, C. shehatae*)	Production of single-cell protein, vitamins, and organic acids.
Hansenula polymorpha	Expression of foreign genes for production of pharmaceutical proteins.
Kluyveromyces marxianus	Rich source of food-grade enzymes (lactase, pectinase, protease); ethanol production (from cheese whey fermentation).
Phaffia rhodozyma	Production of additive colorant (red-pink astaxanthin pigments) for farmed salmon feed.
Pichia pastoris	Production of single-cell protein and therapeutic proteins (from genetically manipulated strains).
Saccharomyces cerevisiae	Production of bread, alcoholic beverages, yeast extracts for food flavorings, fuel alcohol (bioethanol), glycerol, animal feed/growth factor, and biopharmaceutical protiens (through recombinant DNA technology).
Schizosaccharomyces pombe	Millet beer and rum production; wine deacidification; in vitro toxicity testing of mutagenic chemicals; potential in cloning technology.
Schwanniomyces occidentalis	Conversion of starch and inulin (a fructose polymer) into ethanol.
Yarrowia lipolytica	Production of protein and lipid material and industrial enzymes (lipases); it citric acid fermentation.
Zygosaccharomyces rouxii	Food-spoilage (being osmotolerant); production of Japanese soy sauce and miso.

particularly in the health care sector where genetically manipulated yeasts can be used to produce therapeutic proteins for combating human disease.

Fermented beverages. The production of ethyl alcohol (ethanol) by yeast fermentation is as old as human history. Yeast-derived ethanol now represents the premier global biotechnological commodity on both a quantitative and economic basis. Current worldwide production stands at around 30 billion liters per year and includes both potable alcohol (beverages such as beer, wine, cider, sake, whiskey, gin, rum, cognac, vodka, and tequila) and nonpotable alcohol (for fuel and industrial use, often referred to as bioethanol). In both cases, production relies on the ability of yeast cells (principally *S. cerevisiae*) to readily metabolize sugars into ethanol and carbon dioxide under anaerobic conditions using a multienzyme-catalyzed process known as fermentation.

The cereal-based fermented beverages include beer and whiskey (derived from barley) and sake (derived from rice). These products require fermentable sugars to be released from the cereal plant starch following the action of amylase enzymes, which degrade the starch (a polymer of glucose which is not fermented by yeast) into maltose, glucose, and other sugars which are fermented. For beer and whiskey production, this is achieved by the malting process (in which the barley grains are partially germinated to allow amylases to be synthesized) and the mashing process (in which crushed malt infused with warm water allows the amylases to break down the starch). In sake production, release of fermentable sugars from rice is achieved by allowing amylases from fungi such as *Aspergillus oryzae* to facilitate starch hydrolysis. For fruit-based beverages such as wine and cider, fermentable sugars (glucose and fructose) are released by simply crushing the fruit before fermentation proceeds. Yeasts responsible for wine production may be associated with the natural microflora of the grape skin surface (as in traditional winemaking) or specially selected starter cultures of *S. cerevisiae* (as in modern, large-scale winemaking). *See* ALCOHOL FUEL; ETHYL ALCOHOL; FOOD FERMENTATION.

Technological improvements in the production of fermented beverages include both genetic and physiological improvement of yeast strains as well as advances in fermenter design (for example, use of immobilized yeast cells). Genetic improvement of brewing and winemaking yeasts is possible using recombinant deoxyribonucleic acid (DNA) technology, and some strains have approval for food use. However, such yeasts are not currently being used due to adverse public reaction to "genetically manipulated" foods. *See* GENETIC ENGINEERING.

Baker's yeast. Yeast fermentation of sugars produces carbon dioxide in addition to ethanol. Carbon dioxide from yeast is exploited in the leavening of cereal doughs in the breadmaking process. In order to meet the yeast demands of the baking industry, specially selected strains of *S. cerevisiae* are propagated on a large scale. In fact, baker's yeast propagation represents the largest global bulk production of a single-celled microorganism, with several million tons produced annually. Prior to the development of separate baker's yeast industries around a hundred years ago, bakers used to rely on the availability of excess yeast (barm) from breweries. However, yeast from these sources is not suited for modern bakeries, and today baker's yeast is produced in large-scale multistage processes using molasses (sugarcane and sugarbeet refining residues) as a feedstock.

Molasses is a rich source of readily utilized sucrose, vitamins, and minerals for yeast growth. Other nutrients such as nitrogen (in the form of ammonia) and phosphorus (in the form of dibasic sodium phosphate) are added to the molasses prior to dilution, clarification, and flash sterilization (for example, 136°C for 15–30 seconds). Baker's yeast is propagated under fully aerobic conditions (at around 30°C and pH 4), and the rate of molasses feeding into the yeast propagators is carefully controlled to encourage respiratory metabolism and prevent the cells from fermenting and producing ethanol. In this

way, energy generation in the form of adenosine triphosphate (ATP) from respiration is maximized, and the production of yeast biomass is enhanced. At the end of the molasses feeding regime (referred to as fed-batch cultivation), aeration in the yeast propagators continues in order to condition or ripen the cells. This results in an increase in the levels of trehalose (a glucose disaccharide) within the yeast cells, which improves their storage stability and general stress tolerance. Yeast biomass is finally harvested by centrifugation and dried by rotary-drum vacuum filtration (to around 70% moisture) prior to packaging as fresh-pressed baker's yeast. Active-dried baker's yeast (at around 8% moisture) is prepared by spray-drying prior to packaging. *See* ADENOSINE TRIPHOSPHATE (ATP); CARBOHYDRATE METABOLISM.

Other applications of yeast biomass. In addition to their direct food use in the baking industry, yeasts may be specially grown to provide sources of protein and vitamins for human and animal nutrition. Several yeast species have been considered over the years for the production of single-cell protein, including *Candida utilis*, *C. tropicalis*, *Kluyveromyces marxianus*, *K. lactis*, and *S. cerevisiae*. These yeasts have the ability to grow on cheap, readily available carbon sources, such as straight-chain alkanes, corn steep liquor, cheese whey, sulfite waste liquor, and molasses, to produce high-quality protein. Although the use of yeast single-cell protein for direct human consumption has been discontinued by several companies, yeasts grown for use as animal feed and as a livestock growth factor are widespread. The use of live cultures of *S. cerevisiae* in ruminant animal feed has been shown to stabilize the rumen bacteria and improve nutrient availability to increase animal growth (and milk yields in the case of cattle). In human dietary supplements, yeast provides a source of B-complex vitamins, proteins, essential amino acids, and minerals. *See* AMINO ACIDS; PROTEIN; VITAMIN.

Some foods benefit directly from the presence of yeast extracts (whole yeast cells extracted with acids, enzymes, or salt) by imparting characteristic savory flavors and aromas. These compounds derive from the hydrolysis of yeast protein and ribonucleic acid (RNA) to yield taste-enhancing chemicals (for example, glutamic acid, peptides, and ribonucleotides). Some yeast extracts are prepared from excess brewer's yeast, but most are produced from strains of *S. cerevisiae* specially propagated for food applications. Yeast extracts are used in processed convenience foods to impart a meaty flavor (for example, dried soups, gravy granules, and flavored potato snacks). When yeast extracts are prepared from whole cells, residual cell wall material (referred to as yeast hulls or yeast glycan) is generated. This material has potential in the food industry (as a stabilizer, nonnutritive bulking agent, and emulsifier), the environment (as heavy-metal biosorbent material), and in medicine (as immunomodulators to protect against bacterial and fungal infections).

Other industrial uses of yeast biomass are as biotherapeutic agents (for example, *S. boulardii* has been reported to combat harmful intestinal bacteria) and biocatalysts (for example, baker's yeast is used by organic chemists in certain catalytic conversions such as reductions of aldehydes and ketones). Several yeasts have potential use as biocontrol agents to combat fungal crop disease. Another application is in bioremediation, in which yeasts are used to sequester heavy metals from industrial effluents. Immobilized yeast cells are used as biosensors for monitoring levels of certain aquatic pollutants. *See* BIODEGRADATION; INDUSTRIAL MICROBIOLOGY.

Industrial enzymes and chemicals. Although yeasts are not as rich a source of useful enzymes as certain bacteria or fungi, several enzymes prepared from yeast fermentations have applications in the food processing industries. Some of these enzymes are listed, along with several industrially useful yeast-derived chemicals, in **Table 2.** Genetic engineering of yeasts has the potential to increase the range of industrial enzymes and other commodities currently produced. For example, *S. cerevisiae* can be transformed by recombinant DNA technology to synthesize certain hydrolytic enzymes that enable this yeast to break down polymers, such as starch. Such technology would be very desirable in the conversion of starchy wastes into ethanol. *See* ENZYME; GENETIC ENGINEERING.

TABLE 2. Industrially useful enzymes and chemicals produced by yeasts

Commodity	Examples	Yeasts employed	Applications
Enzymes	Invertase (sucrose hydrolysis)	*Saccharomyces cerevisiae*	Confectionery, invert sugar
	Lactase (lactose hydrolysis)	*Kluyveromyces marxianus*	Lactose-reduced milk
	Lipase (triglyceride hydrolysis)	*Yarrowia lipolytica*	Infant foods, flavor modifier
	Protease (protein hydrolysis)	*Kluyveromyces lactis*	Cheese making
	Pectinase (pectin hydrolysis)	*Kluyveromyces* spp.	Wine/juice clarification
Organic acids	Citric acid	*Yarrowia lipolytica*	Foods, pharmaceuticals
	Malic acid	Food acidulant	
Fatty acids	Stearic acid	*Cryptococcus curvatus*	Cocoa butter equivalents
Amino acids	Lysine, tryptophan	*Candida, Saccharomyces*	Dietary supplements, therapy
Vitamins	Thiamin, riboflavin, biotin, pyridoxin, inositol, choline, nicotinc and folic acids	*Saccharomyces cerevisiae, Candida utilis*	Therapeutic uses
Sterols	Ergosterol	*Saccharomyces cerevisiae*	Precursor of vitamin D_2 (feeds)
Polysaccharides	Phosphomannans	*Pichia, Pachysolen, Hansenula*	Food emulsifiers
Polyols	Glycerol	*Saccharomyces cerevisiae*	Nitroglycerine explosives (historical)

Pharmaceutical products. Yeast preparations were used in early medicine to treat bacterial infections and heal wounds. More recently, yeasts have been administered as probiotics to combat harmful intestinal bacteria, and baker's yeast has been considered useful in the treatment of acne and premenstrual stress. The use of genetically manipulated yeasts in modern biotechnology for the production of novel pharmaceuticals is on the increase. Yeasts possess distinct advantages over bacteria for expression of foreign genes. The fact that yeasts are eukaryotic (as opposed to bacteria, which are prokaryotes) means that they have a cellular makeup similar to mammalian cells and can synthesize and process many pharmacologically useful proteins for the treatment and prevention of human disease. Several foreign genes have already been successfully expressed in yeast cells. Examples of therapeutic proteins produced from yeast include viral vaccines (prevention of hepatitis B), hormones (insulin for diabetics), blood proteins (human serum albumin for use in transfusions), growth factors (epidermal growth factor for wound healing), and interferons (beta-interferon for treatment of multiple sclerosis). Although *S. cerevisiae* has been considered a good host for foreign gene expression in this field, other yeasts such as *Pichia pastoris*, *Hansenula polymorpha*, and *Kluyveromyces lactis* have great potential in the production of natureauthentic, bioactive therapeutic proteins using recombinant DNA technology. *See* BIOCHEMICAL ENGINEERING; BIOTECHNOLOGY.

Graeme M. Walker

Bibliography. J. R. Dickinson and M. Schweizer (eds.), *The Metabolism and Molecular Physiology of Saccharomyces cerevisiae*, Taylor and Francis, London, 1998; W. M. Ingledew, Alcohol production by *Saccharomyces cerevisiae*: A yeast primer, in K. A. Jacques et al. (eds.), *The Alcohol Textbook*, 3d ed., Nottingham University Press, 1999; C. P. Kurtzman and J. W. Fell, *The Yeasts: A Taxonomic Study*, 4th ed., Elsevier, Amsterdam, 1998; N. P. Money, More g's than the Space Shuttle: Ballistospore discharge, *Mycologia*, 90(4):547–558, 1998; H. J. Phaff et al., *The Life of Yeasts*, rev. ed., 1978; G. Reed (ed.), *Prescott and Dunn's Industrial Microbiology*, 4th ed., 1982; F. A. Skinner et al. (eds.), *Biology and Activities of Yeasts*, 1981; J. F. T. Spencer and D. M. Spencer (eds.), *Yeasts in Natural and Artificial Habitats*, Springer-Verlag, Berlin, 1997; J. F. T. Spencer et al. (eds.), *Yeast Genetics: Fundamental and Applied Aspects*, 1983; J. F. T. Spencer and D. M. Spencer, *Yeast Technology*, 1989; A. Vaughan and A. Martini, Facts, myths and legends on the prime industrial microorganism, *J. Ind. Microbiol.*, 14:514–522, 1995; G. M. Walker, *Yeast Physiology and Biotechnology*, Wiley, New York, 1998; A. E. Wheals, A. H. Rose, and J. S. Harrison (eds.), *The Yeasts*, 6 vols., 1969–1995; K. Wolf (ed.), *Non-conventional Yeasts in Biotechnology: A Handbook*, Springer-Verlag, Berlin, 1996; F. K. Zimmermann and K. D. Entian (eds.), *Yeast Sugar Metabolism*, Technomic Publishing AG, Basel, Switzerland, 1997.

Yellow fever

An acute, febrile, mosquito-borne viral disease characterized in severe cases by jaundice, albuminuria, and hemorrhage. Inapparent infections also occur.

Infectious agent. The agent is a flavivirus, an arbovirus of group B. The virus multiplies in mosquitoes, which remain infectious for life. After the mosquito ingests a virus-containing blood meal, an interval of 12–18 days (called the extrinsic incubation period) is required for it to become infectious. *See* ANIMAL VIRUS; ARBOVIRAL ENCEPHALITIDES.

The virus enters the body through a mosquito bite and multiplies in lymph nodes, circulates in the blood, and localizes in the liver, spleen, kidney, bone marrow, and lymph glands. The severity of the disease and the major signs and symptoms which appear depend upon where the virus localizes and how much cell destruction occurs. The incubation period is 3–6 days. At the onset, the individual has fever, chills, headache, and backache, followed by nausea and vomiting. A short period of remission often follows. On about the fourth day, the period of intoxication begins with a slow pulse relative to a high fever and moderate jaundice. In severe cases, there are high levels of protein in the urine, and manifestations of bleeding appear; the vomit may be black with altered blood; and there is an abnormally low number of lymphocytes in the blood. When the disease progresses to the severe stage (black vomit and jaundice), the mortality rate is high. However, the infection may be mild and go unrecognized. Regardless of severity, individuals either die or recover completely.

Diagnosis. Diagnosis is made by isolation of the virus from the serum obtained from an individual as early as possible in the disease and inoculated intracerebrally in mice, or by the rise in serum antibody. The antibody response can be either of two types, depending upon the individual's previous experience with group B arboviruses. In a primary infection, the antibodies, particularly the complement-fixing and neutralizing antibodies, are quite specific. In a secondary infection, the individual's serum may react so broadly with other group B viruses that a specific diagnosis is sometimes impossible. *See* ANTIBODY; COMPLEMENT-FIXATION TEST; NEUTRALIZATION REACTION (IMMUNOLOGY).

Epidemiology. There are two major epidemiological cycles of yellow fever: classical or urban epidemic yellow fever, and sylvan or jungle yellow fever. Urban yellow fever involves person-to-person transmission by *Aedes aegypti* mosquitoes in the Western Hemisphere and West Africa. This mosquito breeds in the accumulations of water that accompany human settlement. Mosquitoes remain close to houses and become infected by biting a viremic individual. Urban yellow fever is perpetuated in areas where there is a constant influx of susceptible persons, some cases of yellow fever, and *A. aegypti.* With the use of intensive measures for mosquito abatement, the incidence of urban yellow fever has been markedly reduced in

South America, even though 200–400 cases are recognized annually, mainly in persons occupationally exposed in forested areas. The disease is probably underreported in Africa; epidemics involving forest mosquito vectors affect tens of thousands of humans at intervals of a few years, but only a few cases are officially reported.

Jungle yellow fever is primarily a disease of monkeys. In South America and Africa, it is transmitted from monkey to monkey by arboreal mosquitoes (*Haemagogus* and *Aedes* species) that inhabit the moist forest canopy. The infection in animals ranges from severe to inapparent. Persons, such as woodcutters or road builders, who come in contact with these mosquitoes in the forest can become infected. Jungle yellow fever may also occur when an infected monkey visits a human habitation and is bitten by *A. aegypti*, which then transmits the virus to a human.

Large numbers of inapparent infections occur. Yellow fever has never been reported in India or the East, even though the vector, *A. aegypti*, is widely distributed there.

New outbreaks continue to occur in other regions. In Bolivia, 145 cases of jungle yellow fever, with a mortality rate over 50%, were reported in 1975; and a large epidemic in Africa occurred in 1986 in Nigeria, with total of 3291 cases and 623 deaths. However, the actual incidence has been estimated to have been nearly three times the reported figure, and the number of deaths tenfold higher. *Aedes africanus* appears to have been the most likely vector in this epidemic. Yellow fever is a zoonosis that is difficult to control, and is capable of causing unpredictable epidemics in human populations. Yellow fever in the Americas continues to present epidemiological features typical of its jungle cycle: most cases are in males aged 15–45 years and engaged in agricultural or forestry activities.

Control and prevention. Vigorous mosquito abatement programs have virtually eliminated urban yellow fever. In the United States, the last reported outbreak of yellow fever occurred in 1905. However, with the speed of modern air travel, the threat of a yellow fever outbreak exists where *A. aegypti* is present. Most countries insist upon proper mosquito control on airplanes, and vaccinations of all persons at least 10 days before arrival in or from an endemic zone. However, the yellow fever vaccination requirement for travelers entering the United States was eliminated in 1972.

An excellent attenuated live-virus vaccine is available, in the 17D strain. Vaccine is prepared in fertile chicken eggs and dispensed as a dried powder. It is a live virus which is rehydrated just before use, and must be kept cold. *See* VACCINATION.

Joseph L. Melnick

Bibliography. P. L. Bres, A century of progress in combatting yellow fever, *Bull. WHO*, 4:775–786, 1986; A. S. Evans (ed.), *Viral Infections of Humans: Epidemiology and Control*, 4th ed., 1997; T. D. Monath, Flaviviruses, *Fields Virology*, 4th ed., 2001; T. P. Monath, Yellow fever: A medically neglected disease, *Rev. Infect. Dis.*, 9:165–175, 1987; M. Theiler and H. H. Smith, Use of yellow fever modified by an in vitro cultivation for human immunization, *J. Exp. Med.*, 65:787–800, 1937.

Yersinia

A genus of bacteria in the Enterobacteriaceae family. The bacteria appear as gram-negative rods and share many physiological properties with related *Escherichia coli*. Of the 11 species of *Yersinia, Y. pestis, Y. enterocolitica*, and *Y. pseudotuberculosis* are etiological agents of human disease. *Yersinia pestis* causes flea-borne bubonic plague (the black death), an extraordinarily acute process believed to have killed over 200 million people during human history. Enteropathogenic *Y. pseudotuberculosis* and *Y. enterocolitica* typically cause mild chronic enteric infections. The remaining species either promote primary infection of fish (*Y. ruckeri*) or exist as secondary invaders or inhabitants of natural environments (*Y. aldovae, Y. bercovieri, Y. frederiksenii, Y. intermedia, Y. kristensenii, Y. mollaretii*, and *Y. rohdei*).

The three yersiniae pathogenic to humans share about 70-kilobase plasmids (termed pCD or pYV) encoding cytotoxins termed Yops and anti-inflammatory activities that are delivered to host cells via a type III secretion mechanism. *Yersinia pestis* possesses an additional approximately 100-kb plasmid (pMT) that facilitates survival in the flea vector plus an approximately 10-kb plasmid (pPCP) encoding a plasminogen activator that promotes tissue invasiveness associated with acute disease. Antibodies directed against antigens encoded by the plasmids [for example, pCD-encoded LcrV (V antigen) and pMT-encoded capsular (F1 antigen)] promote immunity against plague. The siderophore yersiniabactin is a chromosomally encoded virulence determinant that facilitates assimilation of iron in vivo by *Y. pestis, Y. pseudotuberculosis*, and highly virulent serotypes of *Y. enterocolitica*.

Plague in humans is typically diagnosed on the basis of clinical symptoms such as fever, septicemia, and formation of buboes (inflamed lymph nodes), and is confirmed upon biopsy by antibody examination or specific agglutination. In the absence of specific diagnostic reagents, *Y. pestis* can still be determined by the nature of its source, morphology, gram reaction, growth habit, and lack of motility as well as certain enzymes shared by other yersiniae. Fortunately, the pneumonic form of human plague seldom occurs at the immediate outset of an epidemic; thus, individuals exhibiting the initial symptoms of this almost uniformly fatal form of the disease are typically identified by the alerted physician and are subjected to prompt treatment with an effective antibiotic (for example, chloramphenicol, but not penicillin).

In North America, the phenomenon of sylvatic plague reflects the occurrence of a widespread reservoir of infected rodents of many species west of the Mississippi River. Contact between humans and infected fleas in these rural areas typically results in a modest number of fatalities almost every year

because of delay in seeking treatment, incorrect diagnosis, or use of ineffective antibiotics. *See* MEDICAL BACTERIOLOGY; PLAGUE. Robert R. Brubaker

Bibliography. R. R. Brubaker, Factors promoting acute and chronic diseases caused by yersiniae, *Clin. Microbiol. Rev.*, 4:309–324, 1991; T. Butler, *Plague and Other Yersinia Infections*, Plenum Press, 1983; G. R. Cornelis et al., The virulence plasmid of *Yersinia*, an antihost genome, *Microbiol. Mol. Biol. Rev.*, 62:1315–1352, 1998; R. D. Perry and J. D. Fetherston, *Clin. Microbiol. Rev.*, 10:35–66, 1997.

Yew

A genus of evergreen trees and shrubs, *Taxus*, with a fruit containing a single seed surrounded by a scarlet, fleshy, cuplike envelope (aril). The leaves are flat and acicular (needle-shaped), green below, with stalks extending downward on the stem. The only native American species of commercial importance is the rather uncommon Pacific yew (*T. brevifolia*), a medium-sized tree of the Pacific Coast and northern Rocky Mountain regions. Its wood is sometimes used for poles, paddles, bows, and small cabinetwork.

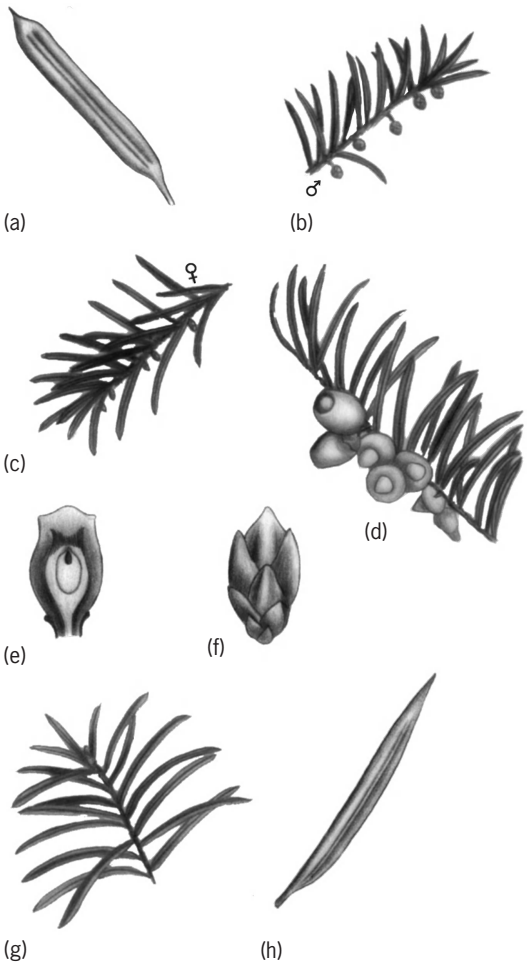

Identifying characteristics of yews. (*a*) Japanese yew (*Taxus cuspidata*) leaf, (*b*) male branchlet with staminate flower buds, (*c*) female branchlet showing ovule-bearing flowers, (*d*) female branchlet with various stages of development of fleshy aril around seed, and (*e*) lengthwise section of berry showing seed in aril. (*f*) English yew (*T. baccata*) bud, showing obtuse, scarcely keeled scales, (*g*) branchlet, and (*h*) leaf.

The English yew (*T. baccata*), native in Europe, North Africa, and northern Asia, and the Japanese yew (*T. cuspidata*) are much cultivated in the United States as evergreen ornamentals. Both are small trees when mature, but often in cultivation they are pruned to lower dimensions. In the English yew the leaves taper gradually to a point and are glossy, whereas in the Japanese yew the leaves are abruptly pointed and of duller appearance (see **illus.**). There are many cultivated forms of these two species.

The Canada yew (*T. canadensis*), also known as ground hemlock, is a low, straggling shrub of the forests of the northeastern quarter of the United States. It is distinguishable from a true hemlock by the absence of white lines on the underside of the leaves. *See* FOREST AND FORESTRY; TREE.

Arthur H. Graves; Kenneth H. Davis

Ylide

A variety of organic compounds which contain two adjacent atoms bearing formal positive and negative charges, and in which both atoms have full octets of electrons. Heteroatoms most commonly utilized as the positive atom are phosphorus, nitrogen, sulfur, selenium, and oxygen, and the negative atom usually involves carbon, nitrogen, oxygen, or sulfur. The ylide may be in an alicyclic or cyclic environment, and in the latter the ylide function may be endocyclic or exocyclic. Ylides most useful in organic synthesis are those containing phosphorus or sulfur and an adjacent carbanion.

Alicyclic and some cyclic ylides are most frequently prepared by deprotonation of the corresponding salt with a strong base. Thus, a phosphorus ylide (also called a phosphorane) is obtained from the phosphonium salt with *n*-butyllithium, the salt itself being prepared from the phosphine and an alkyl halide as shown in the reaction below, where

$$\underset{}{(Ph)_3P} + \underset{\text{Alkyl halide}}{XCHR'R''} \rightarrow \underset{\text{Phosphonium salt}}{(Ph)_3\overset{\oplus}{P}CHR'R''\overset{\ominus}{X}} \xrightarrow{n\text{-BuLi}}$$

$$\underset{\text{Phosphorus ylide}}{(Ph)_3P{=}CR'R'' \leftrightarrow (Ph)_3\overset{\oplus}{P}{-}\overset{\ominus}{C}R'R''}$$

Ph = phenyl group, Bu = butyl group, and X = halogen. *See* REACTIVE INTERMEDIATES.

Ylides of this type are highly reactive, and when R′ and R″ are electron-withdrawing groups, the ylides are more stable. Their principal chemical application is reaction with an aldehyde or ketone, which may be alicyclic, aromatic, or heterocyclic, so that the resultant product contains a double bond at the site of the carbonyl group (the Wittig reaction). *See* ALDEHYDE.

Sulfur-containing ylides (π-sulfuranes) are of two types, sulfonium ylides and oxosulfonium ylides, which differ in having an oxygen atom attached to the sulfur in the latter. Both are usually prepared by deprotonation of the corresponding sulfonium

salts prepared by alkylation of the sulfide or sulfoxide with an active alkyl halide. Limitations on access to the sulfonium salts determine the availability of the ylides which may be stabilized by electron-withdrawing substituents on the carbon atom.

The addition of a carbene to a sulfide provides the most direct route to sulfur ylides and has given rise to stabilized ylides. Stabilized oxosulfonium ylides are specially suited to this approach. Structures (**1**)–(**3**) are examples of the sulfur-containing ylides.

(**1**) (**2**) (**3**)

As with the phosphoranes, the main applications of sulfonium and oxosulfonium ylides are their reactions with aldehydes and ketones which lead to epoxides, and with α,β-unsaturated carbonyl compounds which provide a convenient method of introducing a cyclopropane ring.

The ylide function may also be incorporated into heterocyclic systems. For example, the meso-ionic sydnone molecule (**4*a***) contains an azomethine imine ylide (**4*b***), and the mesomeric betaine (**5*a***) derived from 3-hydroxypyridine contains an azomethine ylide (**5*b***). Both these ylides (often referred to

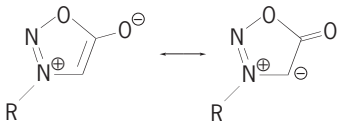

(**4*a***) (**4*b***)

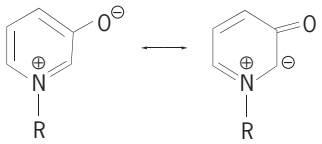

(**5*a***) (**5*b***)

as masked ylides) readily undergo cycloaddition reactions formally analogous to the Diels-Alder reaction. *See* DIELS-ALDER REACTION; MESO-IONIC COMPOUND.

Other heterocyclic ylides are of the exocyclic type and usually are formed by proton abstraction from a heterocyclic salt. Structures (**6**) and (**7**) show two general types.

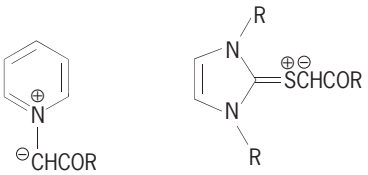

(**6**) (**7**)

Kevin T. Potts

Bibliography. F. A. Carey and R. J. Sundberg, *Advanced Organic Chemistry*, pt. A: *Structures and Mechanisms*, 4th ed., 2000; F. R. Hartley, *The Chemistry of Organophosphorus Compounds*, vol. 3: *Phosphonium Salts, Ylides, and Phosphoranes*, 1994; J. March, *Advanced Organic Chemistry: Reactions, Mechanisms, and Structures*, 5th ed., 2001; A. Padwa (ed.), *1,3-Dipolar Cycloaddition Chemistry*, 1984.

Yolk sac

An extraembryonic membrane which extends through the umbilicus in vertebrates. In some elasmobranchs, birds, and reptiles, it is laden with yolk which serves as the nutritive source of embryonic development. Developed in the splanchnopleure, it is composed of an endodermal layer and a mesodermal layer in which blood vessels are formed. *See* GERM LAYERS.

Chick embryo. The yolk sac develops as a membrane, from extraembryonic splanchnopleure, that grows over the surface of the yolk in embryos such as the chick. The yolk sac connects directly with the gut of the embryo at the midgut region. Yolk is broken down by enzymes present in the inner lining (endoderm) of the yolk sac membrane. Once digested, yolk passes through the yolk sac into blood vessels that are formed in the outer (mesoderm) layer of the yolk sac. Circulation carries the digested yolk to all parts of the developing embryo. As a chick approaches hatching, the yolk sac and remaining yolk become enclosed in the embryo proper. This yolk provides a newly hatched chick with an important source of nutrition.

Blood circulation. The yolk sac is the first source of blood and blood vessels in the chick embryo. Blood islands appear in the yolk sac and quickly differentiate into blood channels enclosing blood cells. The forming vessels in the blood islands extend and join, forming a vitelline plexus which joins larger vessels, the vitelline veins and arteries. These vessels join others and connect to the heart, which begins to pulsate, allowing circulation between the yolk sac and the embryo proper.

The circulation of blood from the yolk sac to the heart and embryo proper is relatively simple. Blood circulates in the small vessels of the yolk sac, picking up yolk components that have been liquefied by the digestive enzymes in the yolk sac endoderm. The blood eventually reaches one of the vitelline veins, which join the omphalomesenteric veins. These veins fuse into a single vein that enters the heart of the embryo proper. Blood from various sources mixes in the heart and leaves through the ventral aorta, then travels through the aortic arches into the dorsal aorta, and finally some of the blood returns to the yolk sac through the vitelline arteries.

Red blood cell production. The development of hemoglobin-rich red blood cells, which carry oxygen to different parts of the body, begins in the yolk sac. In the mouse embryo, for example, an embryonic form of hemoglobin begins to appear in the red blood cell precursors in the blood islands of the yolk sac by the eighth day of gestation. By the ninth day,

these cells are beginning to enter circulation. Later, the liver and then the spleen and bone marrow take over red blood cell production.

Mammals. In mammals, as in birds, the yolk sac generally develops from extraembryonic splanchnopleure, and extends beneath the developing embryo. Just as in the chick, a blood vessel network develops in the mammalian yolk sac lining. Though these blood vessels are empty, they play an important role in absorbing nourishing food and oxygen from the mother. Thus, although the yolk sac in higher mammals may be considered an evolutionary vestige from its yolky-egged ancestors, it still serves important functions in the young embryo. As the embryo ages, the yolk sac shrinks in size, and the allantois takes over the role of nutrition. *See* ALLANTOIS.

Role in immune system. The stem cells that give rise to the cells of the immune system in the chick embryo are also derived from the yolk sac wall. These cells, characterized by darkly staining basophilic cytoplasm, enter the two glands that are associated with the early development of the immune system, the thymus gland and the bursa of Fabricius. In mammalian embryos, such as the mouse, the stem cells of the immune system also originate in the yolk sac wall. These cells travel from the yolk sac to the fetal liver and then finally to the bone marrow.

Primordial germ cell origin. In some organisms, the yolk sac wall is also the site of origin of the primordial germ cells, those cells that eventually form the sperm and eggs. In mammals, the primordial germ cells appear to originate in the endoderm of the yolk sac. These cells contain large amounts of the enzyme alkaline phosphatase and therefore can be observed by using a special stain for this enzyme. The primordial germ cells migrate from the yolk sac wall, through adjoining dorsal mesentery, and into the genital ridges that form the gonads. *See* EMBRYOLOGY; FETAL MEMBRANE. Steven B. Oppenheimer

Bibliography. H. Butler and B. H. Juvrlink, *An Atlas for Staging Mammalian and Chick Embryos*, 1987; B. M. Carlson, *Patten's Foundations of Embryology*, 6th ed., 1996; S. B. Oppenheimer and E. J. Carroll, *Introduction to Embryonic Development*, 4th ed., 2004; T. W. Sadler, *Langman's Medical Embryology*, 8th ed., 2000.

Young's modulus

A constant designated E, the ratio of stress to corresponding strain when the material behaves elastically. Young's modulus is represented by the slope $E = \Delta S/\Delta\epsilon$ of the initial straight segment of the stress-strain diagram. More correctly, E may be represented as the slope of the tangent or the slope of the secant connecting two points of the stress-strain curve. The modulus is then designated as tangent modulus or secant modulus at stated values of stress. The modulus of elasticity applying specifically to tension is called Young's modulus. Many materials have the same value in compression. *See* ELASTICITY; HOOKE'S LAW; STRESS AND STRAIN. W. J. Krefeld; W. G. Bowman

Ytterbium

A chemical element, Yb, atomic number 70, and atomic weight 173.04. Ytterbium is a metal element of the rare-earth group. There are 7 naturally occurring stable isotopes. *See* PERIODIC TABLE.

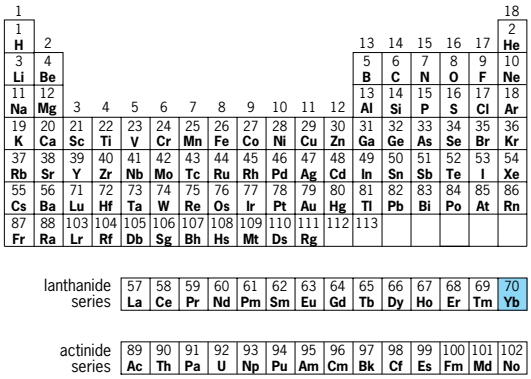

The common oxide, Yb_2O_3, is colorless and dissolves readily in acids to form colorless solutions of trivalent salts which are paramagnetic. Ytterbium also forms a series of divalent compounds. The divalent salts are soluble in water but react very slowly with water to liberate hydrogen.

The metal is best prepared by distillation. It is a silvery soft metal which corrodes slowly in air and resembles the calcium-strontium-barium series more than the rare-earth series. For a discussion of the properties of the metal and its salts *See* RARE-EARTH ELEMENTS. Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; K. A. Gschneidner Jr., J.-C. Bünzli, and V. K. Pecharsky (eds.), *Handbook on the Physics and Chemistry of Rare Earths*, 2005.

Yttrium

A chemical element, Y, atomic number 39, and atomic weight 88.905. Yttrium resembles the rare-earth elements closely. The stable isotope ^{89}Y constitutes 100% of the natural element, which is always found associated with the rare earths and is

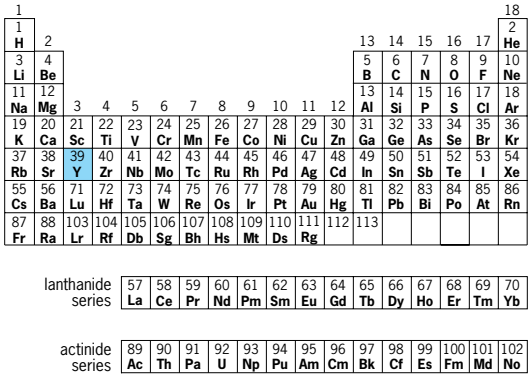

frequently classified as one. *See* PERIODIC TABLE.

Yttrium metal absorbs hydrogen, and in alloys up to a composition of YH_2 they resemble metals very closely. In fact, in certain composition ranges, the alloy is a better conductor of electricity than the pure metal.

Yttrium forms the matrix for the europium-activated yttrium phosphors which emit a brilliant, clear-red light when excited by electrons. The television industry uses these phosphors in manufacturing television screens.

Yttrium is used commercially in the metal industry for alloy purposes and as a "getter" to remove oxygen and nonmetallic impurities in other metals. For properties of the metal and its salts *see* RARE-EARTH ELEMENTS

Frank H. Spedding

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; M. F. Lappert, *Comprehensive Organometallic Chemistry II: Scandium, Yttrium, Lanthanides and Actinides, and Titanium, Zirconium, and Hafnium*, vol. 4, 1995.

Z

Z transform — Zygophyllales

Z transform

The preferred operational-calculus tool for analysis and design of discrete-time systems. (It should not be confused with the z transformation.) The role of the z transform with regard to discrete-time system is similar to that of the Laplace transform for continuous systems. In fact, the Laplace transform is a specialized case of the z transform. The z transform is by far the more insightful tool, and the Laplace transform is just the limiting case of the z transform in a practical as well as a conceptual way. *See* CONTROL SYSTEMS; DIGITAL FILTER; LAPLACE TRANSFORM; LINEAR SYSTEM ANALYSIS.

Definition. It is useful to consider a band-limited real continuous signal, $x(t)$, with no significant amount of energy above a frequency f_c. This signal is sampled at uniformly spaced intervals of time, T, $2T$, $3T, \ldots, nT, \ldots$ (see **illus.**), where the sampling interval, T, and the sampling frequency, f_s, are reciprocals of one another, and where $f_s > 2f_c$, a condition necessary for unambiguous interpretation of the sampled signal. If the common shorthand notation $x(nT) = x_n$ is used, the definition of the z transform of $x(t)$ is given by the equation below. The coefficient of

$$X(z) = x_0 + x_1 z^{-1} + x_2 z^{-2} + x_3 z^{-3} + \cdots$$

z^{-p} is therefore the value of the pth sample of the time signal. This gives a great deal of physical insight to the use of the z transform, a feature not shared by the Laplace transform. The Laplace transform can be found by evaluating the limit of the z transform as T approaches zero. *See* ANALOG-TO-DIGITAL CONVERTER; INFORMATION THEORY.

Poles, zeros, and the z plane. The z-transform expressions are polynomials which may be written in powers of either z or z^{-1}. The choice is a matter of personal preference. The quantity z^{-1} has the physical significance of representing one sample-period of delay. When the z-transform polynomials are

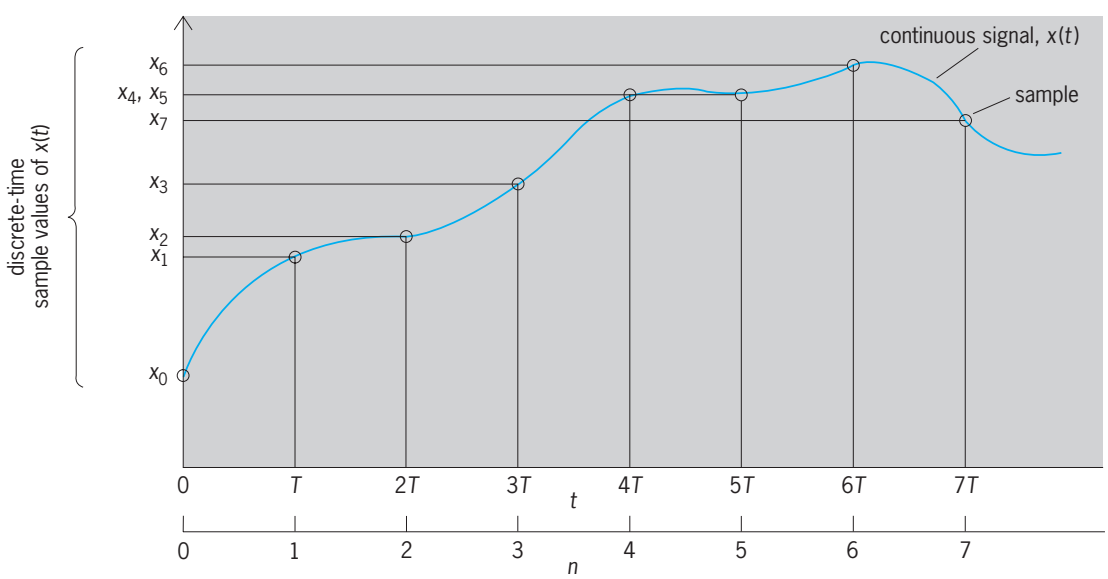

Extraction of discrete-time sample values from a continuous signal.

expressed in powers of z, the process of factoring and finding the roots is straightforward. The roots of the numerator and denominator are called zeros and poles, respectively, and are often mapped in the complex z plane. The vertical and horizontal axes of the complex z plane are the imaginary and real, respectively. Each pole location is traditionally marked by an x, and each zero location is marked by an o.

If the magnitudes of all poles are less than one (that is, they all lie within the unit circle), the corresponding time function will decrease to zero with increasing time. If there are one or more distinct poles whose magnitudes are one (that is, that sit on the unit circle), the corresponding time function will not go to zero with increasing time, but will be periodic and bounded. If any pole is greater than one (that is, if any one or more lie outside the unit circle), or if there are repeated (that is, identical) poles whose magnitudes are one (that is, those that sit on the unit circle), the corresponding time function will grow without bound as time increases. The locations of the zeros have no such effect.

The region of convergence in the z plane is that area in which poles may lie corresponding to a time function that decays to zero. For all positive time, the region of convergence is the area within the unit circle. *See* COMPLEX NUMBERS AND COMPLEX VARIABLES.

Stanley A. White

Bibliography. R. V. Churchill and J. W. Brown, *Complex Variables and Applications*, 6th ed., 1995; L. C. Ludeman, *Fundamentals of Digital Signal Processing*, 1986; A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, 2d ed., 1999; A. V. Oppenheim and A. S. Willsky, *Signals and Systems*, 2d ed., 1996.

Zebra

One of three species of striped, horselike mammals belonging to the family Equidae and found wild in Africa. Zebras are classified in the order Perissodactyla (odd-toed ungulates), a group of mammals in which the middle toe is functional and the second and fourth digits are vestigial. All equids walk on the tips of their toes (unguligrade locomotion). Zebras stand 1.2 to 1.5 m (4–5 ft) high at the withers. Adults weigh 106–202 kg (235–450 lb). The dental formula is I 3/3, C 1/1, PM 3/3, M 3/3 × 2 for a total of 40 teeth.

Grevy's zebra, *Equus grevyi*; Samburu National Reserve, Kenya. (*Gerald and Buff Corsi; copyright © California Academy of Sciences*)

Large numbers of zebras once lived over most of the eastern part of Africa, from southern Egypt to the Cape of Good Hope. They were killed for their meat, which is said to have an excellent taste, and their hides, which are usd to make a tough leather. Some species are nearly extinct, but others are numerous.

Zebras differ from all other members of the horse family because of their startling color pattern. They have parallel black or dark brown stripes on a whitish background arranged in exact designs. These stripes run all over the body, meeting diagonally down the sides of the head. The lines may appear even on the zebra's long ears, short thick mane, and down its tail to the tuft of hair at the tip. Researchers have rejected the hypotheses that stripes serve as camouflage for zebras, that they visually confuse predators and pests, or that they assist in regulating body temperature through heat absorption. Instead, it is thought that stripes facilitate group cohesion and socialization.

Zebras are sociable grazing animals and may be found in association with gnu, deer, and ostriches. Some live on open grassy plains, whereas others live in rough mountains. They live in small bands or family groups, each of which is led by a stallion. Additional stallions live in small stallion groups. Zebras are difficult to tame and train to work.

Reproduction. Births may occur in all months, with gestation periods ranging from 360 days to 13 months. A single foal is the norm. Foals are up and about within an hour of their birth. They start grazing within a few weeks, but are generally not weaned for 8 to 13 months. Females come into heat 7–10 days after giving birth; thus mating and birth occur during the same season. Although females can breed annually, the interbirth interval ranges from 18 to 36 months because of the strain of rearing the foal. Captive zebras have survived for 40 years.

Species and distribution. The plains, or common, zebra (*Equus burchellii*) inhabits the grasslands, light woodlands, open scrub, and savannahs of East Africa south of the Sahara Desert. The largest population (over 250,000) occurs in the Serengeti of Tanzania. This species has a sleek body with broad vertical black and white stripes running onto the belly and becoming horizontal on the haunches. There are three subspecies. The mountain zebra (*E. zebra*) is confined to the mountainous grasslands of southwest Africa. It has narrower stripes than the plains zebra, broad black stripes on the rump, and a white belly. A dewlap (fold of skin) is present on the underside of the neck. The two subspecies, both endangered and close to extinction, are now found only in small numbers on protected reserves. Grevy's zebra (*E. grevyi*, see **illustration**), is the largest

living wild equine, reaching a height of 1.5 m (5 ft) at the withers. It inhabits subdesert steppe and arid bushy grasslands in Ethiopia, Somalia, and northern Kenya. It has narrow, vertical black and white stripes on its slender body, which curve upward on the haunches. The belly is white and unstriped, the mane is prominent and erect, and the legs are long. The head is long and narrow with prominent broad ears giving it a mulelike appearance. This species is also endangered.

The quagga (*E. quagga*) is an extinct southern African mammal that resembled a zebra. The last known individuals died in captivity in Berlin in 1875 and in Amsterdam in 1883. Some observers considered the quagga to be most closely related to the horse based on analyses of mainly cranial characters. Others thought it was a distinct species of zebra related to the three living species. Still others felt it was merely the southern end of a cline (graded series of characters) and a subspecies of the plains zebra. Both DNA and protein analyses of samples from quagga skin confirmed that it was, indeed, related to the plains zebra (*E. burchelli*). A breeding program is now under way in an attempt to "recreate" the quagga by repeated inbreeding of the most quaggalike plains zebras. Molecular studies may thus prove responsible, if only indirectly, for the return of one extinct subspecies (at least one that superficially resembles the quagga, since we can never be sure of the evolutionary pathways that created it) to its native habitat.

Zebra numbers continue to decline, especially on unprotected lands where humans bring settlement, crops, and livestock ranching. Natural predators include lions and hyaenas. *See* AFRICA; PERISSODACTYLA; EQUIDAE. Donald W. Linzey

Bibliography. D. W. Linzey, *Vertebrate Biology*, McGraw-Hill, 2000; D. Macdonald (ed.), *The Encyclopedia of Mammals*, Andromeda Oxford, 2001; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Zebu

A breed of humped, domestic cattle native to India and belonging to the family Bovidae in the order Artiodactyla. Also known as Brahman, they are probably descended from the wild auroch (*Bos taurus*), which is considered to be the ancestor of domestic cattle.

Zebus (*Bos indicus*) have been domesticated in Asia for approximately 6500 years. They were imported by breeders into the United States in 1849 and have been crossed with local breeds of beef cattle to produce strains better adapted to the hot humid Gulf states. The breed has also contributed to beef production though cross breeding with European cattle, such as Hereford and Angus. These cattle exhibit hybrid vigor, that is, they generally exhibit growth and reproductive rates greater than either of the parental types. Several new breeds of cattle have been developed in the United States based on Brahman-European crosses, some important ones being the Beefmaster (Brahman combined with Shorthorn cattle and Hereford cattle); Brangus (Brahman combined with Angus cattle); charbray (Brahman combined with Charolais cattle); and Santa Gertrudis (Brahman combined with Shorthorn). The Santa Gertrudis was the first distinct breed of cattle produced in the United States. It was developed in the 1920s and 1930s at the King Ranch in Kingsville, Texas, by a cross of Shorthorn with Brahman, and was first recognized as a distinct breed in 1940. Large numbers of them live in the humid areas of the South and Southwest. The calves grow rapidly and mature into large cattle. Santa Gertrudis have been exported to Africa and several Latin American countries.

Brahman (*Bos indicus*).

Brahman cattle have a very distinctive appearance with a large, fleshy, fatty hump (or sometimes double hump) over the shoulders, a rounded forehead, loose skin under the throat (dewlap), large drooping ears, and white legs (see **illustration**). The hump is formed from two overdeveloped muscles plus fatty tissue and probably represents an energy reserve for emergencies. These cattle are generally light to medium gray in color, but domestic animals may also be red, brown, or black. Brahman cattle have short hair and well-developed sweat glands that enable them to withstand heat and humidity, ticks, and insects.

Breeding may occur throughout the year, and a single offspring is born after a gestation of 277–290 days. Females attain sexual maturity at about 18 months and remain fertile for about 12 years. The life span may be 20+ years.

Zebus are protected as sacred cows by the Hindus in India, who allow them to roam freely through the streets and villages. In India, they are used as draft animals as well as for milk. Zebus spread from India to parts of China and as far west as East Africa. Some strains were developed for saddle riding. Others were bred for draft purposes or as pack animals. As a result of such selective breeding, zebus vary in weight and in size. Some may be as big as the largest ox. *See* ARTIODACTYLA; BEEF CATTLE PRODUCTION; BOVIDAE; CATTLE. Donald W. Linzey

Bibliography. *Grzimek's Encyclopedia of Mammals*, vol. 5, McGraw-Hill, 1990; R. M. Nowak, *Walker's Mammals of the World*, 6th ed., Johns Hopkins University Press, 1999.

Zeeman effect

A splitting of spectral lines when the light source being studied is placed in a magnetic field. Discovered by P. Zeeman in 1896, the Zeeman effect furnishes information of prime importance in the analysis of spectra. Each kind of spectral term has its characteristic mode of splitting, and the types of terms are most definitely identified by this property. Furthermore, the effect allows an evaluation of the ratio of charge to mass of the electron and an evaluation of its precise magnetic moment.

Normal Zeeman effect. This is a splitting into two or three lines, depending on the direction of observation, as shown in **Fig. 1**. The light of these

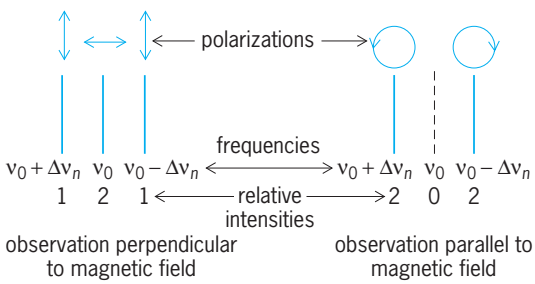

Fig. 1. Triplet observed in normal Zeeman effect.

components is polarized in ways indicated in the figure. The normal effect is observed for all lines belonging to singlet systems, those for which the spin quantum number $S = 0$. The change of frequency $\Delta\nu_n$ of the shifted components can be evaluated on classical electromagnetic principles as follows. Assume that the electron of charge e revolves in a circular orbit of radius r and circular frequency ω radians per second (**Fig. 2**). If a magnetic field H is applied perpendicular to the plane of the orbit, the electron will be speeded up or slowed down because of the changing flux through its orbit, just as in the electron accelerator known as the betatron.

Denoting the centripetal force holding the electron on its orbit before application of the field by f_0 (Fig. 2*b*), and the additional force due to the motion of the electron across the field by f_H (Fig. 2*a* and *c*), one has Eqs. (1). In Fig. 2*a*, where the two forces are

$$f_0 = m\omega^2 r \qquad f_H = He(\omega \pm \Delta\omega)r \tag{1}$$

in the same direction, one may equate their sum to

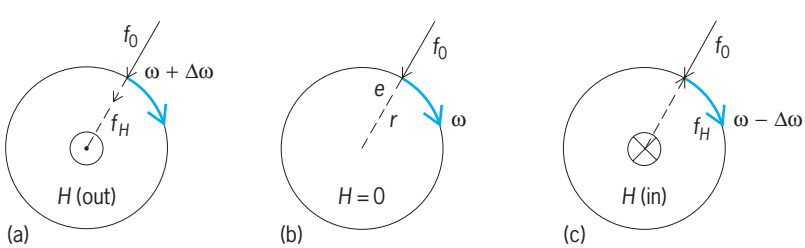

Fig. 2. Effect of a magnetic field *H* applied perpendicular to a circular electron orbit. Diagrams *a*–*c* are explained in the text.

the centripetal force on an electron of frequency $\omega + \Delta\omega$, obtaining Eq. (2). Solution of this equation for

$$\begin{aligned} f_0 + f_H &= m\omega^2 r + He(\omega + \Delta\omega)r \\ &= m(\omega + \Delta\omega)^2 r \end{aligned} \tag{2}$$

$\Delta\omega$, under the assumption that it is small compared to ω itself, yields Eqs. (3), the latter expression fol-

$$\Delta\omega = \frac{eH}{2m} \qquad \Delta\nu_n = \frac{eH}{4\pi m} \tag{3}$$

lowing from the fact that $\nu = \omega/2\pi$. This relation, although derived for a special case, is generally valid for any system of particles having a particular value of e/m and moving under the action of a central force. *See* LARMOR PRECESSION.

On substitution of the radio of charge to mass of an electron, one obtains Eq. (4). Conversely, from

$$\Delta\nu_n = 1.3996 \times 10^6 H \text{ s}^{-1} \tag{4}$$

the observed spectroscopic splitting $\Delta\nu_n$ and measurement of the field strength, the value of e/m for the electron has been evaluated as 1.7572 ± 0.0007 emu/g. This is in good agreement with the figure determined by other methods.

Anomalous Zeeman effect. This effect is a more complicated type of line splitting, so named because it did not agree with the predictions of classical theory. It occurs for any spectral line arising from a combination of terms of multiplicity greater than one. As examples, **Fig. 3** gives diagrams of the theoretical patterns for the yellow lines of sodium, belonging to a doublet system, while **Fig. 4** shows some actual patterns observed for doublets and quartets in rhodium.

Since multiplicity in spectral lines is caused by the presence of a resultant spin vector S of the electrons, the anomalous effect must be attributed to a nonclassical magnetic behavior of the electron spin. While classical theory associates with the vector L of the orbital angular momentum a magnetic moment as in Eq. (5), it is necessary, in explaining the

$$\mu_L = \frac{e\hbar}{4\pi mc} L \tag{5}$$

anomalous Zeeman effect, that the magnetic moment corresponding to S be as in Eq. (6). Thus the

$$\mu_S = \frac{e\hbar}{2\pi mc} S \tag{6}$$

spin generates twice as much magnetic moment, relative to its angular momentum, as does the orbital motion. In an atom for which both L and S are finite, the effective magnetic moment may be written as Eq. (7), where J measures the total angular momen-

$$\mu_J = g\frac{e\hbar}{4\pi mc} J = g\mu_0 J \tag{7}$$

tum (in LS coupling, the resultant of L and S), and μ_0 is the Bohr magneton, $e\hbar/4\pi mc$. *See* PARAMAGNETISM.

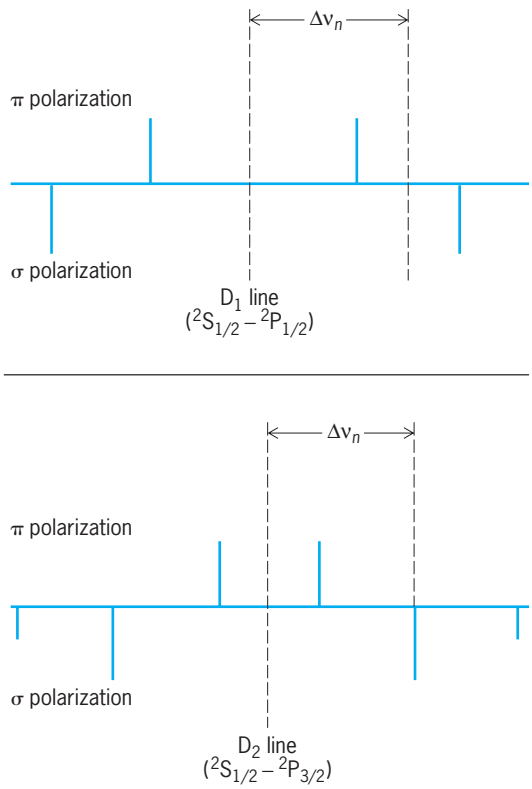

Fig. 3. Anomalous Zeeman effect of the sodium lines. $\Delta\nu_n$ denotes the normal Zeeman splitting, while π and σ refer to polarizations like those of the central and outer components in the normal effect illustrated in Fig. 1. The heights of the lines indicate their relative intensities.

Theory gives values of g, the Landé g factor, which are characteristic of the type of spectral term. In LS coupling, the value is given by Eq. (8).

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \qquad (8)$$

For a classical electron orbit $g = 1$, which yields the normal Zeeman effect. When the spin S is present, however, the changes in energy produced by the magnetic field, which are proportional to μ_J, are just g times as great, and this fact is responsible for the anomalous Zeeman effect. It also should be mentioned that both theory and experiment now show that the g factor for the electron is not exactly 2, but 2.00229.

The component of μ_J in the field direction is $g\mu_0 M$, where M is the quantized component of J in this direction. The energy terms become $T = T_0 + g\mu_0 M$, where T_0 is the term value with no field. This magnetic quantum number M has only the $2J + 1$ values, $J, J - 1, J - 2, \ldots, -J$, and the allowed transitions between energy terms must obey the selection rule $\Delta M = 0, \pm 1$. To explain the statement that the normal Zeeman effect is observed for all lines in singlet systems, recall that S is then zero, $J = L$, and hence $g = 1$.

Quadratic Zeeman effect. The quadratic effect, which depends on the square of the field strength, is of two kinds. The first results from the second-order terms that were neglected in the preceding derivation, and the second, from the diamagnetic reaction of the electron when revolving in large orbits.

Inverse Zeeman effect. This is the Zeeman effect of absorption lines. It is closely related to the Faraday effect, the rotation of plane-polarized light by matter situated in a magnetic field. *See* ATOMIC STRUCTURE AND SPECTRA; FARADAY EFFECT.

Zeeman effect in molecules. This effect is, in general, so small as to be unobservable, even for molecules which have a permanent magnetic moment. Each level with a total angular momentum J splits into $2J + 1$ components, as in the case of atoms.

The component of the magnetic moment along the direction of the external field is small, however, because the rotation of the molecule, which carries the magnetic moment along with it, causes the principal part of the magnetic moment to average out to zero. The consequence is that the magnetic levels have an extremely narrow spacing except for cases where the molecule has either very little rotation or none at all. An exception occurs for some light molecules where the magnetic moment is coupled so lightly to the frame of the molecule that it can orient itself freely in the magnetic field just as for atoms (**Fig. 5**).

Zeeman effect in crystals. A clear Zeeman effect also can be observed in many crystals with sharp spectrum lines in absorption or fluorescence. Such crystals are found particularly among the salts of the rare earths. In these cases the internal electric field

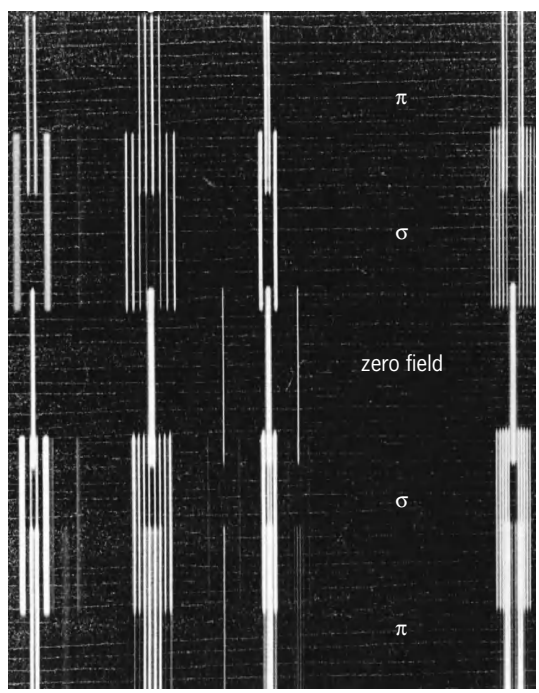

Fig. 4. Zeeman effect of the rhodium spectrum in the wavelength range 347.9–346.2 nanometers. Field strengths are 70,000 oersteds (lower exposure) and 90,500 (upper exposure). (*From G. R. Harrison and F. Bitter, Massachusetts Institute of Technology*)

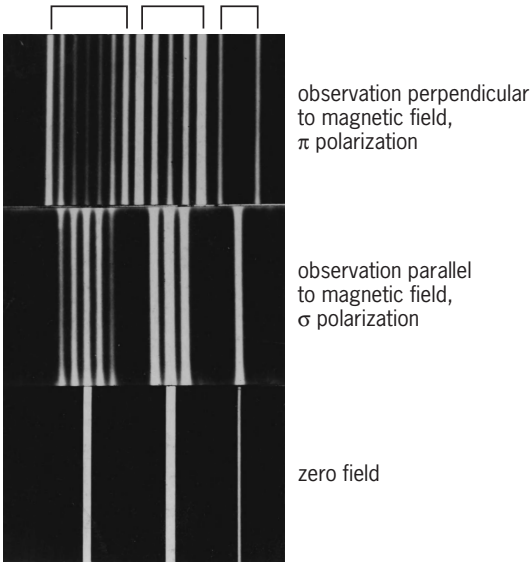

Fig. 5. Zeeman effect of three adjacent rotational lines of the hydrogen molecule. (*Johns Hopkins University*)

in the crystal splits and shifts the level of the free ion. When the number of electrons is even and the crystal symmetry low, this electric splitting is complete. No degeneracy remains, and there can be no further splitting by a magnetic field. If the number of electrons is odd, or if for an even number there is high crystal symmetry, the levels occur in degenerate pairs which are split by a magnetic field. Each line is then split into four components (**Fig. 6**). For cubic crystal symmetry and when angular momentum is due only to electron spin, splitting into more than four components may occur.

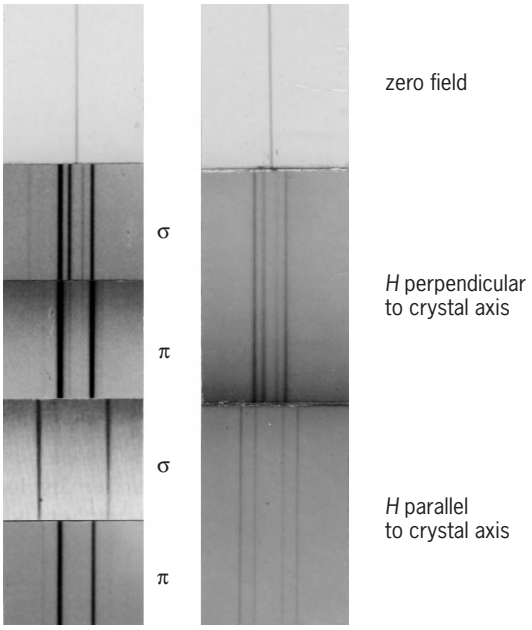

Fig. 6. Zeeman effect of an absorption line of neodymium chloride, with polarization (left) and without polarization (right). The splitting is quite different, depending on whether the trigonal crystal axis is parallel or perpendicular to the magnetic field. *H* = 35,000 oersteds. (*Johns Hopkins University*)

Nuclear Zeeman effect. The magnetic moment of the nucleus causes a Zeeman splitting in atomic spectra which is of an order of magnitude a thousand times smaller than the ordinary Zeeman effect. This Zeeman effect of the hyperfine structure usually is modified by a nuclear Paschen-Back effect, first studied by E. Back and S. A. Goudsmit for the spectral lines of bismuth. *See* PASCHEN-BACK EFFECT.

A strong magnetic field actually may modify the intensity and selection rules so that usually absent lines may appear. For example, the *J* selection rule is no longer valid in a magnetic field. *See* SELECTION RULES (PHYSICS). F. A. Jenkins; G. H. Dieke; W. W. Watson

Bibliography. R. D. Cowan, *The Theory of Atomic Structure and Spectra*, 1981; G. Herzberg, *Atomic Spectra and Atomic Structure*, 2d ed., 1944; F. A. Jenkins and H. E. White, *Fundamentals of Optics*, 4th ed., 1976; S. Svanberg, *Atomic and Molecular Spectroscopy*, 3d ed., 2001; L. Szasz, *The Electronic Structure of Atoms*, 1991; M. R. Wehr, J. A. Richards, and T. W. Adair, *Physics of the Atom*, 4th ed., 1984.

Zeiformes

The order of teleost fishes collectively called dories. The order has representatives in all seas of the world from the tropics to Antarctica; however, they are most common in tropical and temperate zones, in waters of moderate depth to at least 1800 m. Most species are mesopelagic or bathypelagic. Structurally intermediary to the beryciform and the perciform fishes, zeiforms are characterized by a body varying from moderately elongate to very deep and compressed; jaws usually greatly protractile; colors predominantly silvery or silver gray; absence of an orbitosphenoid bone; a simple posttemporal bone rigidly united to the skull; a pelvic fin each with one spine or none and from two to ten soft rays; a first dorsal fin with five to ten spines, usually strong; and an anal fin with one to four spines. Most species are of small size and of minor economic importance.

The order consists of five families. The Oreosomatidae have a very deep and compressed body, a superior and protractile mouth, and small cycloid or ctenoid scales; the young bear conical scutes on parts of body. The Parazenidae have a moderately deep body, highly protractile premaxillaries, weakly ctenoid scales that are not vertically elongate, two lateral lines that are fused behind the soft dorsal fin, and thoracic pelvic fins. This family consists of a single species, *Parazen pacificus*, known only from Japan and halfway around the world in waters off Cuba. Species of Zeniontidae (formerly Macrurocyttidae) have an elongate body, an extremely protractile upper jaw, large eyes, and a pelvic fin that may consist of a single large serrated spine preceding two inconspicuous soft rays and strong and serrated dorsal spines, with one quite larger than the others. Grammicolepididae have a very deep body, a small mouth with nearly vertical jaws, and vertically elongate scales. The Zeidae have a large, almost vertical mouth, small scales (if present) that

are not vertically elongate, and spines or bucklers (enlarged body scales) at the base of dorsal and anal fins. Herbert Boschung

Bibliography. P. C. Heemstra, A revision of the zeid fishes (Zeiformes: Zeidae) of South Africa, *Ichthyol. Bull. J. L. B. Smith Inst. Ichthyol*, 40:17, 1980; P. C. Heemstra, Zeniontidae, p. 441, in M. M. Smith and P. C. Heemstra (eds.), *Smith's Sea Fishes*, Springer-Verlag, Berlin, 1986; P. C. Heemstra, Zeidae. Dories, pp. 1207–1209, in K. E. Carpenter (ed.), *FAO Species Identification Guide for Fishery Purposes. The Living Marine Resources of the Western Central Atlantic*, vol. 2: *Bony Fishes, part 1 (Acipenseridae to Grammatidae)*, U.N. Food and Agricultural Organization (FAO), 2002.

Zener diode

A two-terminal semiconductor junction device with a very sharp voltage breakdown as reverse bias is applied. The device is used to provide a voltage reference. It is named after C. Zener, who first proposed electronic tunneling as a mechanism of electrical breakdown in insulators. *See* SEMICONDUCTOR; TUNNELING IN SOLIDS.

In the Zener diode, the normal rectifying characteristic is of no interest. Both the normal diode and the Zener diode voltage-current characteristics are shown in **Fig. 1**. When the Zener diode is reverse biased and is fed with a current from a higher potential source, a well-defined voltage is developed which mainly depends on the stability of the flowing current and the stability of the device temperature.

A classic circuit to define a very stable current uses an operational amplifier and three stable resistors (**Fig. 2**). The voltage across the Zener itself defines a higher level from which the current is drawn. Thus, a stable noise-free Zener defines its own stable noise-free current. *See* OPERATIONAL AMPLIFIER.

The effect of temperature on the breakdown voltage can be nulled by having a second forward-biased junction, which has a small negative temperature coefficient, in series with the Zener junction. Such a device is called compensated Zener and has a breakdown voltage of 6.2 V rather than the normal 5.6 V (for the smallest possible temperature coefficient). Alternatively, a Zener junction can be part of an integrated circuit which adds a whole temperature controller to keep the silicon substrate at a constant temperature. For the very best performance, only four components are integrated into the silicon: the Zener, a heater resistor, a temperature-sensing transistor, and a current-sensing transistor. A separate selected dual operational amplifier then completes the current-and-temperature-control circuit. Such a circuit sets the chip temperature at, say 122°F (50°C), and the junction condition is then largely independent of ambient temperature. *See* INTEGRATED CIRCUITS.

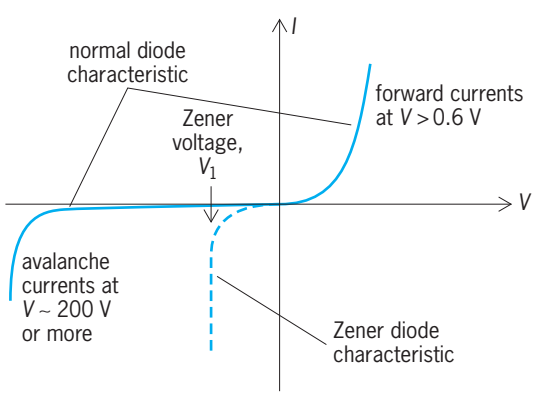

Fig. 1. Diode voltage-current (*V-I*) characteristics.

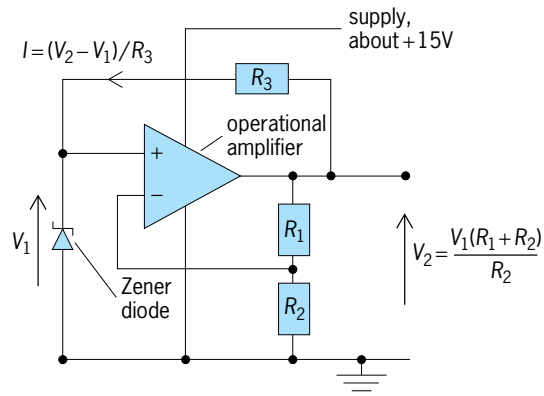

Fig. 2. Zener diode circuit to define a very stable current, *I*, using an operational amplifier and three stable resistors, R_1, R_2, and R_3. The voltage V_1 across the Zener itself defines a higher level V_2 from which the current *I* is drawn.

The *p* and *n* regions of the diode are often deposited under the silicon surface by ion bombardment. Then any surface contaminants will not affect the actual junction region. They would make the Zener voltage more noisy or less stable with time. *See* ION IMPLANTATION.

With the precautions mentioned to control the current and the temperature, the Zener voltage is still found to change by about 0.0002% per year. This is thought to be a function of the junction doping profiles slowly changing by diffusion. This rate of change of Zener voltage approximately doubles for each 18°F (10°C) that the device temperature is raised. Thus, the chip temperature must not be raised excessively to achieve temperature control.

The compact, robust Zener diode, with the circuits described to set its current and temperature, makes a fine portable voltage standard. This is used to disseminate the voltage level from national or accredited calibration laboratories to industry and to research laboratories. *See* JUNCTION DIODE; SEMICONDUCTOR DIODE; VOLTAGE MEASUREMENT; VOLTAGE REGULATOR. Peter J. Spreadbury

Bibliography. J. J. Brophy, *Basic Electronics for Scientists*, 5th ed., McGraw-Hill, 1990; P. Horowitz and W. Hill, *The Art of Electronics*, 2d ed., Cambridge University Press, 1989; J. Millman and A. Grabel, *Microelectronics*, 2d ed., McGraw-Hill, 1987; A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, 4th ed., Oxford University Press, 1997; P. J. Spreadbury, The Ultra-Zener: A portable replacement for the Weston cell?, *IEEE Trans. Instrum. Meas.*, 40:343–346, 1991.

Zenith

The point in the sky directly above an observer. If the observer is at a pole, then the observer's zenith is a celestial pole; but for observers at midlatitudes the zenith is a point in the sky that corresponds to a changing right ascension but a constant declination as the sky rotates overhead. The point directly below the observer is the observer's nadir, and is 180° in longitude and in latitude from the observer's zenith.

The zenith distance of an object in the sky is the angle across the sky from the zenith to the object. The Sun is never overhead for observers outside the equatorial zone marked by the Tropic of Cancer and the Tropic of Capricorn, $\pm 23^1/_2{}^\circ$, so the zenith distance is never less than $21^1/_2{}^\circ$ for an observer at a latitude of 45°. The zenith distance is 90° minus the altitude of an object above the horizon. *See* ASTRONOMICAL COORDINATE SYSTEMS; CELESTIAL SPHERE; TROPIC OF CANCER; TROPIC OF CAPRICORN. Jay M. Pasachoff

Bibliography. J. Mitton, *The Penguin Dictionary of Astronomy*, 3d ed., Penguin, New York, 2000; J. M. Pasachoff and A. Filippenko, *The Cosmos: Astronomy in the New Millenium*, 3d ed., Brooks/Cole Publishing, Davis, California, 2007; P. K. Seidelmann (ed.), *Explanatory Supplement to the Astronomomical Almanac*, University Science Books, Mill Valley, CA, 1992.

Zeolite

Any mineral belonging to the zeolite family of minerals and synthetic compounds characterized by an aluminosilicate tetrahedral framework, ion-exchangeable large cations, and loosely held water molecules permitting reversible dehydration. The general formula can be expressed as $X_y^{1+,2+}Al_x^{3+}Si_{1-x}^{4+}O_2 \cdot nH_2O$. Since the oxygen atoms in the framework are each shared by two tetrahedrons, the (Si,Al):O ratio is exactly 1:2. The amount of large cations (X) present is conditioned by the aluminum/silicon (Al:Si) ratio and the formal charge of these large cations. Typical large cations are the alkalies and alkaline earths such as sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), strontium (Sr^{2+}), and barium (Ba^{2+}). The large cations, coordinated by framework oxygens and water molecules, reside in large cavities in the crystal structure; these cavities and channels may even permit the selective passage of organic molecules. Thus, zeolites are extensively studied from theoretical and technical standpoints because of their potential and actual use as molecular sieves, catalysts, and water softeners. *See* MOLECULAR SIEVE; SILICATE MINERALS; WATER SOFTENING.

Occurrence. Zeolites are low-temperature and low-pressure minerals and commonly occur as late minerals in amygdaloidal basalts, as devitrification products, as authigenic minerals in sandstones and other sediments, and as alteration products of feldspars and nepheline. Phillipsite and laumontite occur extensively in sediments on the ocean floor. Stilbite, heulandite, analcime, chabazite, and scolecite are common as large crystals in vesicles and cavities in the basalts of the Minas Basin Region, Nova Scotia; West Paterson, New Jersey; the Columbia River Plateau; Berufjord, Iceland; Poona, India; and many other localities.

Zeolites are usually white, but often may be colored pink, brown, red, yellow, or green by inclusions; the hardness is moderate (3–5) and the specific gravity low (2.0–2.5) because of their rather open framework structures. Their habits are highly variable and depend on the atomic arrangement; some zeolites are fibrous and woolly while others are platy and micaceous and a few are equant in development and lacking in cleavage.

Crystal structure. As a result of detailed crystal structure analysis, the atomic arrangements of most zeolites are known. Important features are loops of 4-, 5-, 6-, 8-, and 12-membered tetrahedral rings which further link to form channels and cages. On the basis of the kinds of loops and channels present, the zeolite minerals can be classified according to groups (see **table**). The volume-per-framework oxygen atom is a measure of the packing efficiency of the structure, and zeolites have very high values, usually greater than 2.5 cubic nanometers (nm^3) per oxygen atom. This can be compared with a typical dense-packed silicate structure, such as olivine, which has a value of about 1.8 nm^3.

Knowledge of the crystal structure also explains the sieving properties of zeolites. The minimum width of a channel is an approximate measure of the maximum diameter of a molecule which can pass through. Thus, chabazite, with a channel of 0.39-nm minimum width, allows the passage of straight-chain hydrocarbons but not those with branched chains. Faujasite, with a 12-membered ring and channel diameter of 0.9 nm, even permits the passage of benzene rings. This substance, once a mineralogical curiosity, is now of key importance in the catalytic cracking of petroleum. Large quantities of faujasite and other zeolites are made synthetically, and their desirable qualities can be enhanced by tailoring their compositions.

Zeolites are usually synthesized hydrothermally, starting with gels of appropriate composition in an alkaline environment. The temperature of synthesis may range from 450 to 100°C (840 to 212°F) with the more open framework structures occurring at the lower temperatures. The appearance of zeolites in a metamorphic rock indicates a rock of lowest grade; such assemblages are referred to as the zeolite facies. *See* CRYSTAL STRUCTURE; FACIES (GEOLOGY). Paul B. Moore

Applications. The porous, yet crystalline nature of zeolites has been exploited commercially in three main areas—as sorbents, as cation-exchange materials, and as catalysts.

Sorbents. Most aluminosilicate zeolites have an extremely high (yet reversible) affinity for water and are widely used as desiccants. Zeolite A, in its potassium-exchanged form (K-A), is the material of choice for intensive drying of unsaturated hydrocarbon gas streams and polar liquids, where water contents as

Zeolite family of minerals

Group	Name	Formula	Crystal system	Volume/framework oxygen
Analcime	Analcime*	$Na(AlSi_2O_6) \cdot H_2O$	Cubic, pseudocubic	26.8
	Wairakite	$Ca(AlSi_2O_6) \cdot H_2O$	Monoclinic	
	Pollucite	$Cs(AlSi_2O_6) \cdot xH_2O$	Tetragonal	
Sodalite	Sodalite*	$Na_4(Al_3Si_3O_{12})Cl$	Cubic	30.0
	Linde A	$Na_{12}(Al_{12}Si_{12}O_{48}) \cdot 27H_2O$	Cubic	35.1
	ZK-5	$Na_{24}(Al_{24}Si_{72}O_{192}) \cdot 90H_2O$	Cubic	34.0
	Faujasite	$(Na_2,Ca)_{30}((Al,Si)_{192}O_{384}) \cdot 260H_2O$	Cubic	39.2
Chabazite	Chabazite*	$Ca_2(Al_4Si_8O_{24}) \cdot 13H_2O$	Rhombohedral, pseudorhombohedral	34.2
	Gmelinite	$Na_2(Al_2Si_4O_{12}) \cdot 6H_2O$	Hexagonal	33.8
	Erionite, offretite	$Ca_{4.5}(Al_9Si_{27}O_{72}) \cdot 27H_2O$	Hexagonal	31.9
	Levynite	$Ca(Al_2Si_4O_{12}) \cdot 6H_2O$	Rhombohedral	32.5
Natrolite	Natrolite*	$Na_2(Al_2Si_3O_{10}) \cdot 2H_2O$	Orthorhombic	28.3
	Scolecite	$Ca(Al_2Si_3O_{10}) \cdot 3H_2O$	Monoclinic	28.6
	Mesolite	$Na_2Ca_2(Al_2Si_3O_{10})_3 \cdot 8H_2O$	Monoclinic	
	Edingtonite	$Ba(Al_2Si_3O_{10}) \cdot 3H_2O$	Orthorhombic	30.1
	Thomsonite	$NaCa_2(Al_5Si_5O_{20}) \cdot 6H_2O$	Orthorhombic	28.3
	Gonnardite	$(Ca,Na)_{6-8}((Si,Al)_{20}O_{40}) \cdot 12H_2O$	Orthorhombic	
Phillipsite	Phillipsite	$(K,Na)_5(Al_5Si_{11}O_{32}) \cdot 10H_2O$	Orthorhombic	31.3
	Harmotome	$Ba_2(Al_4Si_{12}O_{32}) \cdot 12H_2O$	Monoclinic	30.9
	Gismondine	$Ca(Al_2Si_2O_8) \cdot 4H_2O$	Monoclinic	32.5
	Garronite	$NaCa_{2.5}(Al_6Si_{10}O_{32}) \cdot 13H_2O$	Tetragonal?	
Mordenite	Mordenite	$Na(AlSi_5O_{12}) \cdot 3H_2O$	Orthorhombic	29.2
	Dachiardite	$(Na_2,Ca)_2(Al_4Si_{20}O_{48}) \cdot 12H_2O$	Monoclinic	28.2
Other	Heulandite*	$Ca(Al_2Si_7O_{18}) \cdot 6H_2O$	Monoclinic	29.3
	Brewsterite	$Sr(Al_2Si_6O_{16}) \cdot 5H_2O$	Monoclinic	28.2
	Epistilbite	$Ca(Al_2Si_6O_{16}) \cdot 5H_2O$	Monoclinic	27.8
	Stilbite*	$Na_2Ca_4(Al_{10}Si_{26}O_{72}) \cdot 28H_2O$	Monoclinic	30.1
	Yugawaralite	$Ca_4(Al_7Si_{20}O_{54}) \cdot 14H_2O$	Monoclinic	29.2
	Laumontite	$Ca(Al_2Si_4O_{12}) \cdot 4H_2O$	Monoclinic	28.7
	Ferrierite	$Na_4Mg_2(OH)_2(Al_6Si_{30}O_{72}) \cdot 18H_2O$	Orthorhombic	28.2
	Paulingite	$(K,Ca)_{120}((Al,Si)_{580}O_{1160}) \cdot 690H_2O$	Cubic	37.5

*See separate article.

low as 35 parts per billion (ppb) can be achieved. This behavior exploits the high water affinity of zeolites coupled with the sieving properties of their molecularly sized pores. For example, K-A zeolite has pores of ~0.3-nanometer diameter and so will admit only those molecules that are smaller (water, ammonia, and so forth).

The discrimination between molecules on the basis of their size as compared to the zeolite pore size is the basis of several extremely important separation processes in a procedure known as molecular sieving. Separations such as linear-form branched hydrocarbons for octane enhancement of fuels [using a zeolite in its calcium-exchanged form (Ca-A);5A], and *p*-xylene from *o*-, *m*-isomers and ethylbenzene (using large-pore zeolites such as faujasite), are operated on commercial scale. In addition to separations of molecules on the basis of their size, zeolites can also separtate like-sized molecules on the basis of their polarity. An important example is the passage of air through a bed of zeolites of low Si/Al ratio. Nitrogen is preferentially absorbed because of its higher polarity, and an oxygen-enriched effluent may be obtained for use in oxidation-combustion catalysis, medical applications, and so forth. The bed may be regenerated by brief evacuation, and this pressure-swing process provides an economical alternative to cryogenic separation of nitrogen and oxygen. *See* CHEMICAL SEPARATION TECHNIQUES; CRYOGENICS; MOLECULAR SIEVE.

Ion exchange. The major commodity use of zeolites takes advantage of their cation-exchange ability and leads to their use in low-phosphate detergents. Sodium cations of zeolite Na-A are exchanged into solution to replace divalent cations, for example, calcium present in hard water, thereby softening it. Mineral zeolites have found utility in agricultural and wastewater treatment applications, where they either ion-exchange harmful metal ions out of the stream being treated or absorb ammonia by reacting, as their acidic form, to produce the absorbed ammonium ion. *See* ION EXCHANGE.

Catalysis. The most important and expanding area of application for zeolites is as heterogeneous catalysts. By combining the propeties of excellent thermal stability (>800°C or 1500°F), a pore size of molecular dimensions, and the ready introduction of a wide assortment of cations via ion exchange, numerous very selective catalysts can be prepared. Much of the important catalysis relies on the generation of highly acidic sites within the zeolite pores. Thermal decomposition of the ammonium ion form of many zeolites leads to superacidic porous solids that can catalyze a wide range of organic molecule transformations via protonation and carbocation formation. More than 90% of catalytic cracking catalysts used in the United States are zeolite based and are used to break down heavier molecular components of crude oil to lighter fractions used in fuels and gasoline. *See* CRACKING; REACTIVE INTERMEDIATES.

Lanthanum-cerium-exchanged faujasties are 10,000 times more active than the amorphous aluminosilicate catalysts used previously. The synthetic zeolite ZSM-5 has been developed commercially for the catalytic conversion of methanol to gasoline via successive dehydration and polymerization of nascent methylene (CH_2) fragments. This step provides the last link in a chemical chain from coal or biomass to liquid fuels. ZSM-5 is also used for alkylation of toluene (with methanol) and for isomerization of mixed xylenes to produce the desired pure *p*-xylene isomer used in polyester manufacture. These latter chemistries rely on the shape-selective nature of the pore system of the zeolite structure. Since almost all of the reactive surface area and attendant acid sites of the zeolite are located within the internal pore structure (>99%), the catalytic chemistries occur within the pores and are, therefore, being performed in a very restricted space. This physical space constraint limits the outcome of the chemistry to those products that either can fit into the available space or can migrate from the site of their formation to the exterior of the zeolite crystalline and hence escape to be collected.

In addition to the acid-catalyzed chemistries, metal-catalyzed processes can be effected by using metal ion–exchanged zeolites where the zeolite imprints a shape or size selectivity onto the chemistry performed by the metal sites. Platinum (Pt), ruthenium (Ru), and nickel (Ni) hydrogenation and dehydrogenation-dehydrocyclization catalysts have been developed that exhibit useful selectivities toward organic molecules that are explained on the basis of the sieving effect of the zeolite host. Other developments involve the production of crystalline frameworks substituted with transition-metal ions such as titanium (Ti), chromium (Cr), manganese (Mn), and iron (Fe). The titanosilicates show excellent activity in selective oxidation transformations such as phenol to hydroquinone, using hydrogen peroxide as oxidant. Many completely nonsilicate zeolite frameworks have begun to appear as the field matures and synthetic techniques and understanding develop. Structures based on tin chalcogenides, aluminophosphates, and other tetrahedral-ion frameworks have become available. *See* HETEROGENEOUS CATALYSIS.

Finally, several nontraditional uses of zeolites have been introduced where the zeolite structure is viewed either as a nano-sized reaction vessel or as a nanocontainer for novel optical, electronic, or polymeric materials. The concept of the shape-selective pore structure of a zeolite as an inorganic replacement for the tertiary structure of an enzyme (zeozyme) has been explored, and examples demonstrating oxygen transport, electron transfer, photosynthesis, and selective oxygenation have been reported. Novel composites can be prepared where the pores of the zeolite are filled with a second phase of a polymer, semiconductor, metal, and so forth. Different physical properties (such as nonlinear optical behavior or superconductivity) become apparent because of the extremely small size domains of the included material imposed by the zeolite structure. Any physical property that depends on the collective properties of atoms or molecules (such as conductivity or optical behavior) becomes profoundly affected by being forced to reside in this nanocrystalline regime. *See* COMPOSITE MATERIAL; NANOCHEMISTRY; NANOSTRUCTURE.

Norman Herron

Bibliography. D. W. Breck, *Zeolite Molecular Sieves,* 1984; Ch. Baerlocher, W. M. Meier, and D. Olson (eds.), *Atlas of Zeolite Framework Types*, 5th ed., 2001; D. R. Corbin and N. Herron, Designing zeolite catalysts for size and shape selective reactions, *J. Mol. Catal.*, 86:343–369, January 1994; J. C. Jansen (ed.), *Advanced Zeolite Science and Applications*, 1994; H. Van Bekkum, E. M. Flanigen, and J. C. Jansen (eds.), *Introduction to Zeolite Science and Practice,* 2d ed., 2001.

Zero

In mathematics, the concept zero is used in two ways: as a number and as a value of a variable. The positional system of number notation, developed first by the Babylonians with the base 60, and later by the Hindus and the Chinese with the base 10, required for greater clarity a special marker of the empty, nonoccupied position. As such, the symbol for zero was introduced by the Babylonians (about 500 B.C.), and a millennium later it appeared in the Indian system of notation, which came to the West through the Arabs as the arabic number system.

The zero as a number, however, is a new concept, introduced by the Hindus and Chinese about the same time (sixth century). Brahmagupta (born A.D. 598) remarked that the number 0 has special properties: $a \pm 0 = a$, and $a \cdot 0 = 0$, where a may be any number (integer). Using zero as a denominator, he stated that $a/0 \pm b = a/0$, where $a/0$ thus appears as a number of a new kind, which cannot be changed by addition or subtraction. (In the modern way of thinking, division by zero is therefore not a "permissible" operation.) In the acceptance of zero as a number, the Orient is centuries ahead of the Occident. Leonardo Fibonacci (*Liber Abaci*, 1228) was the first Westerner to treat the zero. It is of interest that negative numbers appear earlier in the history of mathematics than does the number zero; they were already known to the Chinese mathematician Liu Hui in the third century of the Christian era. The historically late understanding of the number zero is witnessed by the fact that the Julian calendar (established 46 B.C.), from which the present-day Gregorian calendar is derived, has no year 0.

In a modern way, zero can be called the identity element of the infinite Abelian additive group of integers. If in an integral domain (all the more certainly in a field) a product is equal to zero, then at least one factor of the product is zero.

In the second concept zero is the value of a variable for which a function is equal to zero. For example, "A polynomial of degree n has n zeros," or "The

Riemann zeta function $\zeta(s)$ has all its complex zeros in the strip $0 <$ real part $s < 1$." *See* NUMBER THEORY.

Hans Rademacher; Emil Grosswald

Zinc

A chemical element, Zn, atomic number 30, and atomic weight 65.38. Zinc is a malleable, ductile, gray metal. Because of chemical similarities among zinc, cadmium, and mercury, these three metals are classed together in a transition-elements subgroup of the periodic table. *See* METAL; PERIODIC TABLE; TRANSITION ELEMENTS.

1																	18
1 **H**	2											13	14	15	16	17	2 **He**
3 **Li**	4 **Be**											5 **B**	6 **C**	7 **N**	8 **O**	9 **F**	10 **Ne**
11 **Na**	12 **Mg**	3	4	5	6	7	8	9	10	11	12	13 **Al**	14 **Si**	15 **P**	16 **S**	17 **Cl**	18 **Ar**
19 **K**	20 **Ca**	21 **Sc**	22 **Ti**	23 **V**	24 **Cr**	25 **Mn**	26 **Fe**	27 **Co**	28 **Ni**	29 **Cu**	30 **Zn**	31 **Ga**	32 **Ge**	33 **As**	34 **Se**	35 **Br**	36 **Kr**
37 **Rb**	38 **Sr**	39 **Y**	40 **Zr**	41 **Nb**	42 **Mo**	43 **Tc**	44 **Ru**	45 **Rh**	46 **Pd**	47 **Ag**	48 **Cd**	49 **In**	50 **Sn**	51 **Sb**	52 **Te**	53 **I**	54 **Xe**
55 **Cs**	56 **Ba**	71 **Lu**	72 **Hf**	73 **Ta**	74 **W**	75 **Re**	76 **Os**	77 **Ir**	78 **Pt**	79 **Au**	80 **Hg**	81 **Tl**	82 **Pb**	83 **Bi**	84 **Po**	85 **At**	86 **Rn**
87 **Fr**	88 **Ra**	103 **Lr**	104 **Rf**	105 **Db**	106 **Sg**	107 **Bh**	108 **Hs**	109 **Mt**	110 **Ds**	111 **Rg**	112	113					

lanthanide series	57 **La**	58 **Ce**	59 **Pr**	60 **Nd**	61 **Pm**	62 **Sm**	63 **Eu**	64 **Gd**	65 **Tb**	66 **Dy**	67 **Ho**	68 **Er**	69 **Tm**	70 **Yb**

actinide series	89 **Ac**	90 **Th**	91 **Pa**	92 **U**	93 **Np**	94 **Pu**	95 **Am**	96 **Cm**	97 **Bk**	98 **Cf**	99 **Es**	100 **Fm**	101 **Md**	102 **No**

Fifteen isotopes of zinc are known, of which five are stable, having atomic masses of 64, 66, 67, 68, and 70. About half of ordinary zinc occurs as the isotope of atomic mass 64. The half-lives of the radioactive isotopes range from 88 s for ^{61}Zn to 244 days for ^{65}Zn.

Zinc is a fairly active metal chemically. It can be ignited with some difficulty to give a blue-green flame in air and to discharge clouds of zinc oxide smoke. Zinc ranks above hydrogen in the electrochemical series, so that metallic zinc in an acidic solution will react to liberate hydrogen gas as the zinc passes into solution to form dipositively charged zinc ions, Zn^{2+}. This reaction is slow with very pure zinc, but the presence of small amounts of impurities, addition of a trace of copper sulfate, or contact between the zinc surface and such metals as nickel or platinum facilitates formation of gaseous hydrogen and speeds the reaction. The combination of zinc and dilute acid is often used to generate small quantities of hydrogen in the laboratory. Zinc also dissolves in strongly alkaline solutions, such as sodium hydroxide, to liberate hydrogen and form dinegatively charged tetrahydroxozincate ions, $Zn(OH)_4^{2-}$, sometimes written as ZnO_2^{2-} in the formulas of the zincate compounds. Zinc also dissolves in solutions of ammonia or ammonium salts. The common soluble zinc compounds undergo to some extent the process of hydrolysis, which makes their solutions slightly acidic. The ion Zn^{2+} is colorless, so that the relatively few zinc compounds that are not colorless in large crystals, or white as powders, receive their color through the influence of the other constituents. Some of the atomic and ionic properties of zinc are shown in the **table**. *See* ELECTROCHEMICAL SERIES; HYDROLYSIS.

Atomic and ionic properties of zinc

Property	Value
Electronic configuration	$1s^2, 2s^2, 2p^6, 3s^2, 3p^6, 3d^{10}, 4s^2$
Ionization potentials	
1st electron loss	9.39 eV
2d electron loss	17.9 eV
Ionic radius, Zn^{2+}	0.072 nm
Covalent radius (tetrahedral)	0.131 nm
Oxidation potentials	$Zn \rightleftharpoons Zn^{2+} + 2e^-$, $E^\circ = 0.76$ V $Zn + 4OH^- \rightleftharpoons ZnO_2^{2-} + 2H_2O + 2e^-$, $E^\circ = 1.22$ V

Zinc also forms many coordination compounds. The zincates are actually coordination compounds, or complexes, in which hydroxide ions, OH^-, are bound to the zinc ions. Ammonia, NH_3, forms complexes with zinc, such as the typical tetrammine zinc ion, $[Zn(NH_3)_4]^{2+}$. Zinc cyanide, usually given the simple formula $Zn(CN)_2$, is a coordination compound in which many alternating zinc and cyanide ions are three-dimensionally bound together in a very large molecule. This compound is still widely used in zinc plating, but concern over environmental pollution has led to increasing use of zinc chloride plating baths. In most coordination compounds of zinc, the fundamental structural unit is a central zinc ion surrounded by four coordinated groups arranged spatially at the corners of a regular tetrahedron. *See* COORDINATION CHEMISTRY; COORDINATION COMPLEXES.

Pure, freshly polished zinc is bluish-white, lustrous, and moderately hard (2.5 on Mohs scale). Moist air brings about a superficial tarnish to give the metal its usual grayish color. Pure zinc is malleable and ductile enough to be rolled or drawn, but small amounts of other metals present as contaminants may render it brittle. Malleability of even pure zinc is improved by heating zinc to 100–150°C (212–300°F). If heated zinc is mechanically worked; it does not embrittle on cooling. Zinc melts at 420°C (788°F) and boils at 907°C (1665°F). Its density is 7.13 times that of water, so that 1 ft^3 (0.028 m^3) of zinc weighs 445 lb (200 kg).

As a conductor of heat and of electricity, zinc ranks fairly high. However, its electrical resistivity (5.92 microhm-cm at 20°C or 68°F) is almost four times that of silver, the best conductor. As a conductor of heat, zinc is likewise only about one-fourth as efficient as silver. At 0.91 K zinc is an electrical superconductor. Pure zinc is not ferromagnetic, but the alloy compound $ZrZn_2$ displays ferromagnetism below 35 K.

The most important uses of zinc are in its alloys and as a protective coating on other metals. Coating iron or steel with zinc is known as galvanizing, and it may be done by immersing the article in melted zinc (hot-dip process), depositing zinc electrolytically onto the article in a plating bath (electrogalvanizing), exposing the article to powdered zinc near

its melting point (sherardizing), or spraying the article with melted zinc (metallizing). The mere physical presence of the zinc coat prevents corrosion of iron, and even if breaks in the coat expose portions of the iron, the greater chemical activity of the zinc causes it to be consumed in preference to the iron. Adding small amounts of other metals to galvanizing baths has been found to improve the adhesion and weathering qualities of the coating. *See* ELECTROPLATING OF METALS; GALVANIZING.

Even such nonstructural materials as cardboard can be zinc-coated by low-temperature flame spraying. Other important uses of zinc are in brass and zinc die-casting alloys, in zinc sheet and strip, in electrical dry cells, in making certain zinc compounds, and as a reducing agent in chemical preparations.

A so-called tumble-plating process coats small metal parts by applying zinc powder to them with an adhesive, then tumbling them with glass beads to roll out the powder into a continuous coat of zinc. Rechargeable nickel-zinc batteries offer higher energy densities than conventional dry cells. Foamed zinc metal has been suggested for use in lightweight structures such as aircraft and spacecraft. Some other uses of zinc are in dry cells, roofing, lithographic plates, fuses, organ pipes, and wire coatings. Zinc dust, a flammable material when dry, is used in fireworks and as a chemical catalyst and reducing agent. Radioactive ^{65}Zn is used medically in the study of metabolism of zinc, and also in determining rates of wear for zinc-containing alloys. *See* METAL COATINGS; ZINC ALLOYS. William E. Cooley

Zinc is believed to be needed for normal growth and development of all living species, including humans; actually, life without zinc would be impossible. Zinc is a common element that is present in virtually every type of human food, and zinc deficiency is therefore not considered to be a common problem in humans. Zinc is a trace element; that is, it is present in biological fluids at a concentration below 1 ppm, and only a small amount (normally <25 mg) is required in the daily diet. (The recommended daily allowance for zinc is 15 mg/day for adults and 10 mg/day for growing children.) It is relatively nontoxic, without noticeable side effects at intake levels of up to 10 times the normal daily requirement. James F. Riordan; Kenneth H. Falchuk

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; D. R. Lide, *CRC Handbook Chemistry and Physics*, 85th ed., CRC Press, 2004; W. Maret (ed.), *Zinc Biochemistry, Physiology and Homeostasis: Recent Insights and Current Trends*, 2002; S. W. Morgan, *Zinc and Its Alloys and Compounds*, 1985.

Zinc alloys

Combinations of zinc with one or more other metals. If zinc is the primary constituent of the alloy, it is a zinc-base alloy. Zinc also is commonly used in varying degrees as an alloying component with other base metals, such as copper, aluminum, and magnesium. A familiar example of the latter is the association of varying amounts of zinc (up to 45%) with copper to produce brass, the third largest use of zinc in the United States. *See* BRASS; COPPER ALLOYS.

Uses. Zinc-base alloys have two major uses: for casting and for wrought applications. Casting includes both die casting and gravity casting, which differs from die casting primarily in that no pressure is applied, except the force of gravity, in forcing the molten metal into the mold. *See* METAL CASTING.

Zinc alloys have historically been used for die casting mainly because of their relatively low melting temperatures. This means lower fabrication costs and longer tool life. Zinc-alloy die castings also have good strength and dimensional tolerance and exceptional castability and are easily plated and machined.

Alloys for casting. The zinc alloys used for die casting are well standardized and have been covered by American Society for Testing and Materials (ASTM) specifications in the United States since 1931. Equivalent specifications are in effect by similar organizations throughout the world, including the British Standards Institution (BSI), the Canadian Standards Association (CSA), and the International Standards Organization (ISO).

The presence of impurities above a certain specified maximum is deleterious in all types of alloys. Zinc die-casting alloys, like other alloys, are sensitive to small variations of certain impurities. The alloys, however, are not difficult to make, and most large die casters do their own alloying.

Lead, tin, and cadmium must be avoided as impurities in zinc, or as contaminants in the other alloying ingredients used. These impurities are high-density metals and have low melting points and, when present in amounts beyond the specification limits, render the alloys susceptible to intergranular corrosion, particularly on exposure to warm moist environments. To avoid these problems Special High-Grade Zinc, which is 99.99% pure, is used for die-casting applications. The great bulk of zinc die casting currently is carried out with two alloys, commonly identified as alloys 3 and 5, which have been in use for several decades. A third alloy was added and is commonly designated as alloy 7. The entire group in the United States is often spoken of by its original trade name as Zamak alloys, and in England as the Mazak alloys. A fourth alloy that is especially suited for cold-chamber die-casting machines was introduced to the industry in the mid-1960s. The composition and the mechanical and other properties of these alloys are given in **Table 1**, along with corresponding ASTM and Society of Automotive Engineers (SAE) designations.

Aluminum, the major alloying constituent in alloys 3 and 5, is added in amounts of about 4%, which has proved to be the optimum composition from the standpoint of strength, ductility, and stability. Loss of strength and stability may be as high as 50% of the original values if the aluminum content is reduced to 2%. The aluminum content also sharply reduces the rate of attack of molten zinc on iron containers. Aluminum also avoids "soldering," or sticking

TABLE 1. Composition and properties of zinc alloy die castings

ASTM designation B86-64:	AG40A	AC41A	–	–
SAE designation:	903	925	903	–
General designation:	3	5	ZAMAK 7	ILZRO 16
Composition % by weight				
Copper	.25 max	.75–1.25	.25 max	1.0–1.5
Aluminum	3.5–4.3	3.5–4.3	3.5–4.3	.01–.04
Magnesium	.020–.05	.03–.08	.005–.02	.01
Iron max.	.100	.100	.075	–
Lead max.	.005	.005	.0030	.004
Cadmium max.	.004	.004	.0020	.003
Tin max.	.003	.003	.0010	.002
Nickel	–	–	.005–.020	–
Titanium	–	–	–	.15–.25
Zinc (99.99 + % purity)	Remainder	Remainder	Remainder	Remainder
Mechanical properties				
Charpy impact strength, ft-lb (N·m), 1/4 × 1/4 in.				
(6.35 × 6.35 mm) bar, as cast	43 (58)	48 (65)	40 (54)	–
after 10 years of indoor aging	41 (56)	40 (54)	41 (56)	–
after 20 years of indoor aging	39 (53)	20 (27)	–	–
Tensile strength, lb/in.^{2} (MPa), after 20 years of indoor aging	33,000 (228)	36,000 (248)	–	–
Elongation, % in 2 in. (50.8 mm), after 20 years of indoor aging	20	12	–	–
Tensile strength, lb/in.^{2} (MPa), as cast	41,000 (283)	47,600 (328)	41,000 (283)	33,000 (228)
after 10 years of indoor aging	35,000 (241)	39,300 (271)	35,000 (241)	–
Elongation, % in 2 in. (50.8 mm), as cast	10	7	10	5–6
after 10 years of indoor aging	16	13	16	–
Expansion (growth), in. per in. (cm per cm), after 10 years of indoor aging	.0001	.0001	.0001	–
Other properties (as cast)				
Brinell hardness	82	91	76	75–77
Compression strength, lb/in.^{2} (MPa)	60,000 (414)	87,000 (600)	60,000 (414)	–
Electrical conductivity, mhos/cm^3 at 20°C	157,000	153,000	157,000	119,000
Melting point, °C	386.6	386.1	386.6	418
°F	727.9	727.0	727.9	785
Modulus of rupture, lb/in.^{2} (MPa)	95,000 (655)	105,000 (724)	95,000 (655)	–
Shear strength, lb/in.^{2} (MPa)	31,000 (214)	38,000 (262)	31,000 (214)	–
Solidification point, °C	380.6	380.4	380.6	416
°F	717.1	716.7	717.1	780
Solidification shrinkage, in./ft (cm/m)	.14 (1.2)	.14 (1.2)	.14 (1.2)	–
Specific gravity	6.6	6.7	6.6	7.1
Specific heat, cal/g/°C [J/(g·°C)]	0.10 (0.42)	0.10 (0.42)	0.10 (0.42)	–
Thermal conductivity, cal/(s·cm·°C) [J/(s·cm·°C)] at 18°C	0.27 (1.13)	0.26 (1.09)	0.27 (1.13)	0.205 (0.86)
Transverse deflection, in.	0.27	0.16	0.27	–
Thermal expansion per °C	.0000274	.0000274	.0000274	.000027
per °F	.0000152	.0000152	.0000152	.000015
Specific weight, lb/in.^{3} (g·cm^3)	.24 (6.6)	.24 (6.6)	.24 (6.6)	.26 (7.2)

of the casting to the die. This permits the casting of these zinc alloys in the more productive hot-chamber (plunger) type of die-casting machine. This type of machine injects the molten metal into the die cavity automatically, unlike the cold-chamber machine which requires the molten metal to be poured, by ladling, into the cylinder. *See* ALUMINUM.

Alloys 3 and 5 differ primarily in their copper content. Copper additions increase strength and hardness and improve corrosion resistance. The copper content of alloys is held to a specified maximum of 1.25% (in alloy 5) to avoid making the alloy unstable through aging and to avoid reducing the impact strength to a very low value. Dimensional stability also is affected by high levels of copper, so that a 1% maximum is usually recommended. Changes in casting dimensions, however, are minimal and always occur within a few weeks after casting; consequently, unless the castings are large or require very close tolerances, the dimensional changes can usually be considered insignificant. In those cases where very high dimensional accuracy is required, alloy 3 is usually recommended.

Alloy 7 is considered a modification of alloy 3. It has a lower magnesium content, and iron, lead, tin, and cadmium are also held to lower levels. A small amount of nickel is added. When the alloys are compared, alloy 7 exhibits improved casting properties, making it easier to secure high-quality surfaces for finishing (mostly chromium plating) and to obtain higher production rates. The improved castability is due primarily to the low level of magnesium.

The alloy ILZRO 16 has a fairly wide permissible composition range, but in practice it is recommended that it be taken as titanium 0.15–0.25%, chromium 0.01–0.20% (Ti plus Cr should range between 0.30 and 0.40%), copper 1.0–1.5%, and aluminum 0.10–0.04%, the remainder being zinc. This alloy has a superior creep resistance and is designed for applications required to resist deformation under a sustained stress, especially at elevated temperatures. A comparison of the creep properties of ILZRO

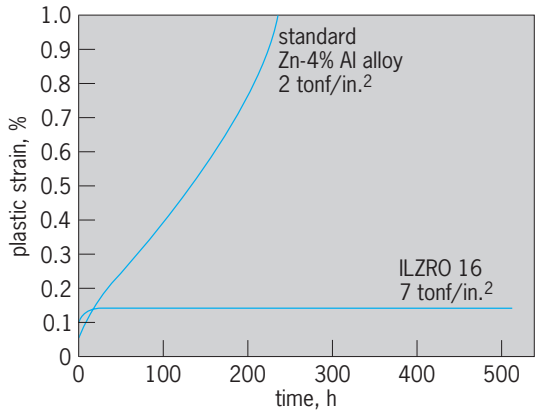

Fig. 1. Creep of two pressure-die-cast zinc alloys at 167°F (75°C). 1 ton-force/in.^{2} = 13.8 MPa.

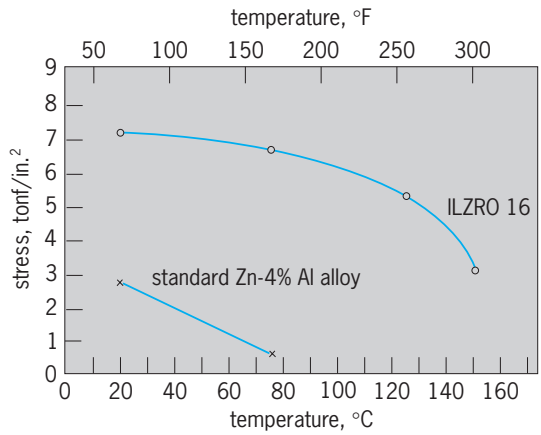

Fig. 2. Variation with temperature and stress to produce 0.1% plastic strain in 500 h for two zinc die-casting alloys. 1 tonf/in.^{2} = 13.8 MPa.

16 and the conventional 4% aluminum alloy is shown in **Fig. 1**. The effect of temperature on the stress that is required to produce 0.1% plastic strain in 500 h is shown in **Fig. 2**.

An important gravity-casting alloy, known as ILZRO 12, was introduced to the industry in 1967. This foundry alloy can be cast in sand, plaster, or permanent molds, is economic for general foundary work, and has found considerable application for prototypes. Its nominal composition is 10.5–11.5% aluminum, 0.5–1.25% copper, 0.01–0.03% tin (maximum), and balance zinc by difference. Principal properties of ILZRO 12 are shown in **Table 2**. There is some improvement in impact resistance as the aluminum content is reduced. Reducing the content to around 10% improves ductility, but tensile strength is lower and impact is about the same as in 11% aluminum alloys.

This new foundry alloy has high tensile strength, excellent castability, and other advantages inherent in zinc die-casting alloys. One of its most important characteristics is that its as-cast properties closely relate to those of the conventional zinc alloys when die cast. The rate of cooling has no apparent effect on its mechanical properties. The following are among its other advantages: foundry scrap can be recycled repeatedly; gating of dies is simple; it is not sensitive to superheating, stewing, or remelting; it has low shrinkage; and gas porosity is minimized. In addition, it can be plated or otherwise finished using conventional finishing processes.

Wrought zinc alloys. In the wrought zinc area, numerous compositions and alloys are used, depending on ultimate product requirements. Alloying metals can be used to improve various properties, such as stiffness, for special applications. The zinc-copper-titanium alloy has become the dominant wrought-zinc alloy for applications demanding superior performance. Composition of this alloy is 0.4–0.8% copper, 0.08–0.16% titanium, 0.3% max. lead, 0.015% max. iron, 0.01% max. manganese, 0.01% max. cadmium, 0.02% max. chromium, and the balance Special High-Grade Zinc.

The conventional grades of slab or ingot zinc are

TABLE 2. Properties of ILZRO 12

Property	Value
Yield strength, lb/in.^{2} (MPa)	
Sand cast	30,000 (207)
Chill cast	31,000 (214)
Tensile strength, lb/in.^{2} (MPa)	
Sand cast	43,500 (300)
Chill cast	53,500 (369)
Impact strength, ft-lb (N·m)	20 (27)
Shear strength, lb/in.^{2} (MPa)	32,250 (222)
Elongation, % in 2 in. (50.8 mm)	
Sand cast	3 1/2
Chill cast	5
Hardness (Brinell–500 kg)	
Sand cast	110
Chill cast	115
Compressive 0.1% proof stress, lb/in.^{2} (MPa)	21,000 (145.5)
Creep strength to give 0.1% strain at 500 h, lb/in.^{2} (MPa)	
68°F (20°C)	8000* (55)
122°F (50°C)	3000 (21)
Fatigue endurance limit, 10^8 cycles, lb/in.^{2} (MPa)	8000 (55)
Properties after 200 days at 212°F (100°C)	
Tensile strength, lb/in.^{2} (MPa)	30,130 (207.7)
0.1% proof stress, lb/in.^{2} (MPa)	17,800 (122.7)
0.2% proof stress, lb/in.^{2} (MPa)	21,160 (145.9)
Elongation, % in 2 in. (50.8 mm)	6
Impact strength, ft-lb (N·m)	2 1/2 (3.4)
Specific gravity	6.0
Density, lb/in.^{3} (g/cm^3)	0.216 (6.0)
Steam test	Resistant
Dimensional stability at 68°F (20°C)	0.0001 in./in./year (0.001 cm/cm/year) Shrinkage during first year
Thermal expansion at	
68–212°F	1.55×10^{-5} per °F
20–100°C	2.79×10^{-5} per °C
68–392°F	1.61×10^{-5} per °F
20–200°C	2.90×10^{-5} per °C
Thermal conductivity	0.21–0.22 cal/(s · cm · °C) 0.88–0.92 J/(cm · °C · s)

*Similar to zinc–4% aluminum pressure-die-casting alloy.

rolled into sheets, strips, ribbons, foils, plates, or rods for a wide variety of uses. The zinc alloy also may be continuous-cast into rod or bar. Because of its properties, rolled zinc can be easily worked into various shapes and forms by common fabricating methods, including stamping, forming, and spinning. It can be polished and lacquered to retain its natural bright luster, or it can be plated or painted for other finishing effects. It can also be left to weather naturally, in which case it forms an attractive, nonstaining patina, making it especially suited for architectural applications.

The zinc-copper-titanium wrought-zinc alloy has all of these properties and on a comparative basis is stronger and more dent-resistant than some other metals of the same thickness. Tensile strength is 29,000 lb/in.^{2} (200 megapascals) with the grain and 40,000 lb/in.^{2} (280 MPa) across the grain. Elongation over 2 in. (5 cm) is 26% with the grain and 14% across the grain. Rockwell hardness is 60–76. It is easily soldered, is nonmagnetic, and has an electrical conductivity equal to that of brass. It is produced in coils ranging 0.003–0.250 in. (0.008–0.635 cm) thick and 0.25–19 in. (0.635–48 cm) wide, and in sheets ranging 0.020–0.125 in. (0.051–0.318 cm) thick, 20–60 in. (51–153 cm) wide, and to 120 in. (305 cm) long. *See* ZINC; ZINC METALLURGY. Adolph L. Ponikvar

Bibliography. S. Gupta, *The World Zinc Industry*, 1981; C. H. Matthewson (ed.), *Zinc: The Science and Technology of the Metal, Its Alloys and Compounds* ACS Monogr. 142, 1959; S. W. Morgan, *Zinc and Its Alloys*, 1977.

Zinc metallurgy

Separating and extracting zinc from ores, refining it, and preparing it into usable forms. Ordinarily, after being mined, ores must first be separated into a concentrated mineral and a waste rock. This concentrate then is reduced to the metal in a metallurgical works. Finally, the metal may be further refined and alloyed to commercially usable form. Commonly, these three extractive metallurgical operations are conducted at separate locations and are broadly categorized as concentrating, smelting, and refining, respectively. At each operation there are several process steps necessary to accomplish the overall purpose. The **illustration** shows the various interrelated processes of zinc metallurgy. *See* METALLURGY; ZINC.

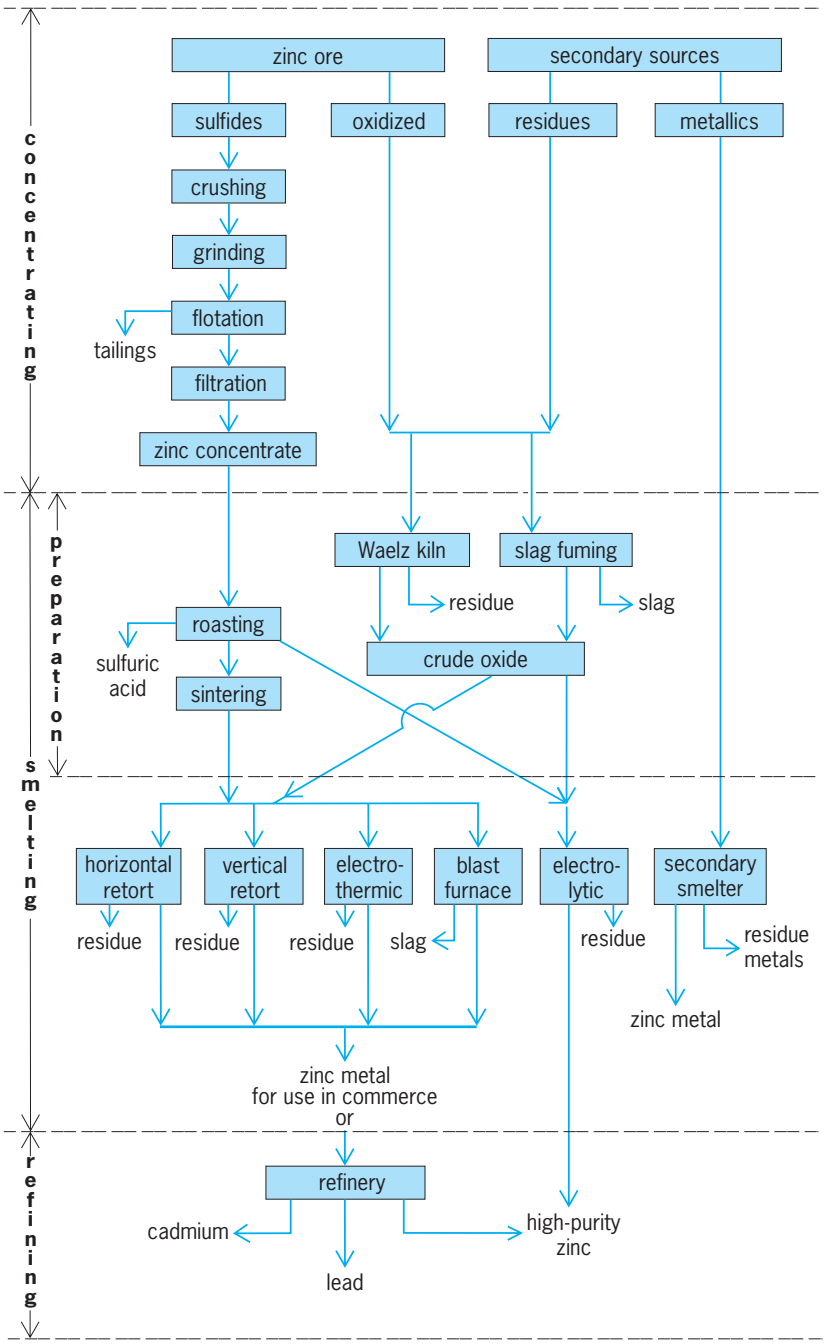

Schematic flow sheet for zinc metallurgy.

Zinc-bearing ores of commercial value occur predominantly in sulfide form (for example, sphalerite, ZnS). Some minor sources are known as oxidized ores and occur as zinc oxide, zinc silicate, zinc carbonate, and some complex combinations of these and other elements. Usual impurities in zinc ores will be one or more of the following: lead, cadmium, and iron; frequently copper, silver, gold, arsenic, and antimony; and occasionally tin, indium, germanium, and gallium. The primary objective of extractive metallurgical operations, then, is to separate the zinc from these impurities. Of major importance is the by-product recovery of some of these impurities in economic form during the processing for zinc extraction. *See* SPHALERITE.

Concentration. Zinc sulfide ores are usually beneficiated (concentrated or ore-dressed) adjacent to the mine site. First, the ore must be crushed and ground in order to free the mineral lattices from those of the waste rock (gangue). Next, the finely divided ore is mixed into a slurry with water and the mineral and gangue particles are separated utilizing the effect of gravity. The most common means of separation is the froth flotation process, wherein a small amount of pine oil is added to the slurry, followed by air agitation. A thin oil film entraps a small bubble of air which is then attracted to the fine mineral sulfide

particle. As they float to the top of the vessel, these bubbles (froth) carrying the mineral are skimmed from the overflow and collected.

The gangue does not attract the bubble and is removed from the bottom of the vessel to be discarded as waste (tailings). Further selective (or differential) flotation steps may be performed, using suitable chemical reagents, to separate much of the lead sulfide and iron sulfide minerals from the zinc sulfide, in producing a purer zinc concentrate. The final step in beneficiation consists of merely separating the mineral slurry into solids and water by means of a filtration process. Resulting zinc sulfide concentrates typically contain 50–64% zinc, depending upon the extent of other impurities which remain to be eliminated in ensuing metallurgical processing operations. Oxidized zinc ores and other zinc-bearing secondary materials are generally not susceptible to concentration at their source, this function usually being performed in conjunction with the smelting operation. *See* ORE DRESSING.

Preparation for smelting. This second operation of zinc metallurgy involves first the chemical or physical changes required in the source material to prepare it into a crude zinc oxide form and a specified particle size (depending upon the particular smelting method contemplated). Such process steps variously entail the operations of roasting, sintering, or pyroconcentration.

Roasting. Zinc sulfide concentrates must be converted to the oxide form prior to smelting. This is accomplished by heating the sulfide and burning it with oxygen (usually from air) to form crude zinc oxide and gaseous sulfur dioxide. The latter is usually converted to sulfuric acid with suitable equipment, and it yields an economic credit for the roasting operation. Production of sulfuric acid is mandatory in many locations to abate an air-pollution problem which would otherwise result.

Several types of roasting plants are in satisfactory operation, including rabbled single-hearth and multiple-hearth furnaces; suspension (flash) roasters in which a showering stream of finely divided sulfide concentrate is burned in a combustion chamber; and fluidized-bed roasters, in which the combustion air is pressurized, being admitted through very small tuyere holes in a steel plate forming the base of the combustion chamber and burning the sulfur from the sulfide concentrates as the gas stream flows upward through the hot bed. Most modern zinc roasters are of the fluid-bed type for reasons of economy and metallurgical efficiency.

Sintering. Most pyrometallurgical extraction methods require that the roasted or pyroconcentrated oxides (and sometimes a portion of concentrates) be further treated to increase their density and particle size prior to feeding into the smelter. Suitable control of temperature, airflow, and reducing agents in this step also serves to remove impurities from the crude zinc oxide during sintering. These are driven off chiefly as volatilized compounds of cadmium, lead, silver, and indium which are carried over in the combustion gas stream and captured in a bag house or electrostatic precipitator for subsequent recovery of these metals. *See* SINTERING.

Equipment for performing the sintering operation is usually a Dwight-Lloyd moving grate machine or a rotary kiln.

Pyroconcentration. Oxidic zinc ores of low grade, secondary zinc-bearing materials, and certain residues from other chemical and metallurgical operations are often concentrated economically in a rotary kiln (Waelz) operation or alternatively by means of the so-called slag-fuming process. In either case, the zinc-bearing material is mixed with sufficient coal to reduce the oxide zinc to its elemental form. Heat required for this endothermic chemical reaction and for vaporization of the zinc is provided partially by burning of natural gas or powdered coal inside the chamber and partially by oxidation of the carbon monoxide and zinc vapor by excess air in the traveling gas stream. Resultant zinc oxide (of much higher concentration than in the feed) is carried away from the reaction zone in the gas stream and is thence collected in a bag-house (filter) and forwarded to the smelter. Nonvolatile impurities, such as iron, copper, and silica, remain in the slag residue and are usually discarded as waste.

Smelting methods. The other basic part of smelting is the reduction step, wherein the zinc is reduced from its oxide to its elemental form. Several very successful and varied processes accomplish this function, including horizontal retort, vertical retort, electrothermic furnace, and blast furnace (all of which are pyrometallurgical and use carbon as a reducing agent). The electrolytic process uses the passage of electric current for reduction to metal from a liquid bath (hydrometallurgical). Used zinc is also reclaimed in secondary smelters. In all pyrometallurgical processes the basis for this step is the requirement for sufficient heat generation to raise the temperature of the zinciferous material and reduction fuel (usually coal and coke mixture) to at least 1832°F (1000°C), so that the carbon will reduce the zinc oxide to zinc metal. This is above the boiling point of zinc, thus it is in the vapor phase. Carbon above that required for zinc reduction must be present to shift the equilibrium away from formation of carbon dioxide and toward formation of carbon monoxide; otherwise the former will reoxidize the zinc vapor as it passes out of the reaction chamber into the condensation vessel. *See* PYROMETALLURGY, NONFERROUS.

Most impurities remaining from the preparation steps described in the previous topic are eliminated to the furnace residues, including cadmium and much of the lead.

The hydrometallurgical (electrolytic) process for zinc smelting, rather than using heat for the reduction, relies on electrodeposition of the metal from a zinc sulfate solution prepared from the crude zinc oxide and sulfuric acid. Virtually all impurities remaining from the preparation step are eliminated in this process. *See* HYDROMETALLURGY.

Horizontal (Belgian) retort process. This process is the earliest known method of zinc smelting and employs vast, honeycomblike batteries of fire clay or silicon-carbide retorts (800 or so in a block), 8–10 in. (20–25 cm) in diameter and 5–6 ft (1.5–1.8 m) long, set

horizontally in a gas- or coal-fired furnace. Retorts are supported at each end and set in rows of 15–20, 4–5 high. The mixed feed of sintered concentrate and coal is usually charged by mechanical means into one set of retorts at a time. As the reaction temperature is attained, a carbon monoxide gas and zinc vapor mixture issues from the mouth of the retort, the zinc being condensed in a clay vessel which has been luted to the open retort end. The mouth of this condenser is stuffed with a porous coke mixture, allowing the noncondensable carbon monoxide to pass through and burn while the molten zinc is collected inside. Four or five times during the 24–48-h firing cycle, zinc is tapped from these individual condensers into a transfer ladle, thence poured into ingot molds.

At cycle end, the condensers are taken down from an entire furnace and the residues cleaned out by machines to make the retorts ready for receiving new charge. Development of charging and cleanout machines has significantly reduced physical labor required in this process. Further labor-saving efforts have been made to centralize condensation and metal tapping by replacing the hundreds of individual condensers on a furnace with an airtight metal shield across the face of the furnace, but conclusive results have not yet been established. The number of operating horizontal retort zinc smelters has declined markedly due to the difficulty of preventing particulate emissions, the stringencies in natural gas supply, and the high labor requirements.

Vertical (New Jersey) retort process. In this process a vertical retort of silicon carbide brick, having a horizontal rectangular dimension of 1 ft (0.3 m) by 6–8 ft (1.8–2.4 m) and a vertical heated height of up to 35 ft (10.7 m), is employed. Gas heats the external walls and a briquetted charge of mixed roasted concentrate and coal is put through a coking process, then admitted through the top at frequent intervals. Residue discharges continuously from the bottom. The mixed zinc vapor–carbon monoxide gas stream is constantly drawn off near the top of the retort and the molten zinc collected in a special condenser, while the carbon monoxide is withdrawn for use as fuel in the external heating. Typical daily zinc production from one such vertical retort may be 10 tons (9 metric tons) whereas that for one entire block (of 800) horizontal retorts might be only 20 tons (18 metric tons).

Electrothermic (St. Joseph) retort process. An 8-ft-diameter (2.4-m), upright firebrick retort of up to 50 ft (15 m) high is the furnace in this process. Sintered concentrate and coke mixture is charged constantly into the top of the furnace where heat of reaction is supplied by passage of heavy electrical current through the resistance of the charge. Sixteen large-diameter graphite electrodes, eight near the top and eight near the bottom, transmit the current to the descending charge. Residue is continuously withdrawn from the bottom. The mixture of zinc vapor and gas is drawn through a bath and condensed into metal, or it may be alternately burned into zinc oxide and collected in a bag house for use in the rubber or paint industries. A single furnace produces 80 tons (72 metric tons) or more of zinc per day.

Blast furnace (Imperial Smelting) process. Design of this furnace is taken from the process ordinarily used in smelting of lead. In fact, the ISP furnace smelts zinc and lead concomitantly. The firebrick vertical shaft is of a rectangular cross section, pinched in at the bottom to accommodate several tuyeres for injection of blast air. Heat of reaction is generated internally by burning the carbon content of the mixed charge. The latter is admitted in frequent batches at the top of the shaft and consists of sintered lead, zinc concentrate, and suitable fluxes, to produce a slag which is tapped at shaft bottom. Certain impurities are carried out as waste in the slag; valuable lead bullion and copper-bearing matte are separated from the slag at this stage. Carbon monoxide gas carrying the zinc vapor is drawn through a spray of molten lead, where the zinc is shock-cooled and conveyed by the molten stream to settling basins. Zinc, being less dense than lead, is withdrawn from the top of the bath and cast. Because of the unique method of collection, zinc produced from this furnace carries high lead content.

Electrolytic process. An increasing percentage of the world's zinc metal is produced by this method. In contrast with the pyrometallurgical processes, the electrolytic reduction of zinc concentrate is leached with sulfuric acid (largely regenerated as spent electrolyte in the cells).

The resulting zinc sulfate solution is rigorously purified in two or more stages before being pumped into the electrolytic cells. Lead-silver alloy anodes and aluminum cathodes are used, the pure zinc deposited on the latter being stripped periodically, melted, and cast into ingots. Virtually all impurities entering with the zinc concentrate are removed as residues in the closely controlled roasting, leaching, and purification steps, resulting in production of very high-purity zinc, directly. *See* ELECTROMETALLURGY.

Secondary smelting. Of increasing importance is the reclaiming of used zinc contained in chemical and metallurgical residues and from scrappage of appliances, automobiles, and aircraft dies. The nonmetallic forms are treated in primary and secondary smelters much as are oxidized ores, requiring carbon reduction. Metallic forms are frequently melted in muffle furnaces or special retorts, and the zinc is selectively vaporized from higher-boiling impurities, being suitably condensed and cast for reuse.

Usually, the metal reclaimed by the secondary smelting process will not meet specifications for the higher grades of zinc; however, careful preselection of the scrap to be melted may result in production of up to the High-Grade purity.

Refining. Purity of zinc produced by the various carbon reduction processes is largely dependent upon controls and operating procedures practiced during the smelting preparation steps. The quality produced is variously suitable for hot-dip galvanizing, continuous-line galvanizing, and in some cases for brass manufacture and rolled (wrought) zinc; however, for the sizable usage in die-casting alloys, output from this type of smelter must

undergo a refining step. Fractional distillation in reflux refining columns is the leading method of upgrading the lower-purity zinc metal.

Typically, a three-column unit is operated with two lead columns followed by a cadmium column. The lead column concentrates lead, iron, and other high-boiling impurities in the lower end, driving zinc, cadmium, and other lower-boiling metals over the top and into the cadmium column. Cadmium is then condensed off the top of this column while purified zinc is withdrawn from the bottom. The process is capable of producing zinc that is 99.995% pure.

Five commercial grades of zinc are recognized by the American Society for Testing and Materials, ranging downward in purity from Special High-Grade (99.99% minimum), through High-Grade, Intermediate, Brass Special, to the lowest grade, Prime Western (98.3% minimum), with maximum impurity limits being placed on each grade with respect to content of lead, iron, cadmium, aluminum, tin, or all of these. *See* METAL AND MINERAL PROCESSING; ZINC ALLOYS.

Carl H. Cotterill

Bibliography. S. Gupta, *The World Zinc Industry*, 1981; C. H. Matthewson (ed.), *Zinc: The Science and Technology of the Metal, Its Alloys and Compounds*, ACS Monogr. 142, 1959; S. W. Morgan, *Zinc and Its Alloys*, 1977.

Zincite

A mineral with composition ZnO (zinc oxide). It crystallizes in the hexagonal system with a wurtzite-type structure. Thus its principal axis is polar and different forms appear at top and bottom of crystals. Such crystals are rare and the mineral is usually massive. Cleavage is prismatic. Its hardness is 4 and its specific gravity 5.6. The mineral has a subadamantine luster and a deep-red to orange-yellow color. The color of the mineral is believed to result partly from the presence of manganese and partly from defects in the crystal structure. Zincite is rare except at the zinc deposits at Franklin and Sterling Hill, New Jersey. There, associated with franklinite and willemite, it is mined as a valuable ore of zinc. *See* ZINC.

Cornelius S. Hurlbut, Jr.

Zingiberales

An order of flowering plants, division Magnoliophyta, in the monocots, consisting of eight families and about 1800 species. Zingiberales (also known as Scitamineae or Scitaminales) are morphologically well defined and clearly circumscribed in DNA sequence analyses. The largest families are Zingiberaceae (about 1000 species), Marantaceae (about 400 species), Costaceae (about 150 species), and Heliconiaceae (about 100 species). Zingiberales are most closely related to Commelinales, and the enigmatic *Hanguana* from southeast Asia is difficult to place with certainty in either order. Zingiberales are herbs or scarcely branched trees or shrubs with pinnately veined leaves and irregular flowers that have well-differentiated sepals and petals, an inferior ovary, septal or septa-derived nectaries, and usually either one or five functional stamens. The endosperm is starchy, and the stomates have two or (usually) more subsidiary cells. The economic crops, ginger (*Zingiber officinale*) in Zingiberaceae and banana (*Musa*) in Musaceae, and the ornamentals, bird of paradise flower (*Strelitzia*) in Strelitziaceae and *Canna* in Cannaceae, are familiar members of Zingiberales. *See* BANANA; COMMELINALES; FLOWER; FRUIT; GINGER; LILIOPSIDA; MAGNOLIOPHYTA.

Michael F. Fay

Zingiberidae

A subclass of Liliopsida (Monocotyledons) of the division Magnoliophyta (Angiospermae), the flowering plants, containing two orders (Bromeliales and Zingiberales), with nine families and about 3800 species. The subclass has been associated with the Commelinidae and the Liliidae, sharing some features with each but aberrant in either, and it differs from both in usually having four or more supporting cells around each stomate. The flowers are usually epigynous, usually bisexual, and often irregular, and they are usually showy and adapted to pollination by insects or other animals. The pistil consists of three united carpels, often with septal nectaries opening at the top of the ovary. The seeds usually have well-developed, mealy or starchy endosperm. *See* BROMELIALES; LILIOPSIDA; MAGNOLIOPHYTA; PLANT KINGDOM; SECRETORY STRUCTURES (PLANT); ZINGIBERALES.

T. M. Barkley

Zircon

A mineral with the idealized composition $ZrSiO_4$, one of the chief sources of the element zirconium. Trace amounts of uranium and of thorium are often present and the mineral may then be partly or entirely metamict. The name cyrtolite is applied to an altered type of zircon. Structurally, zircon is a nesosilicate, with isolated SiO_4 groups. It is isostructural with the thorium silicate thorite and the yttrium phosphate xenotime. *See* RADIOACTIVE MINERALS; SILICATE MINERALS; ZIRCONIUM.

Zircon is tetragonal in crystallization. It often occurs as well-formed crystals, which commonly are square prisms terminated by a low pyramid (see **illus.**). The color is variable, usually brown to reddish brown, but also colorless, pale yellowish, green, or blue. The transparent colorless or tinted varieties are popular gemstones. Hardness is $7^1/_2$ on Mohs scale; specific gravity is 4.7, decreasing in metamict types. *See* METAMICT STATE.

Because of its chemical and physical stability, zircon resists weathering and accumulates in residual deposits and in beach and river sands, from which it has been obtained commercially in Florida and in India, Brazil, and other countries. It is widespread

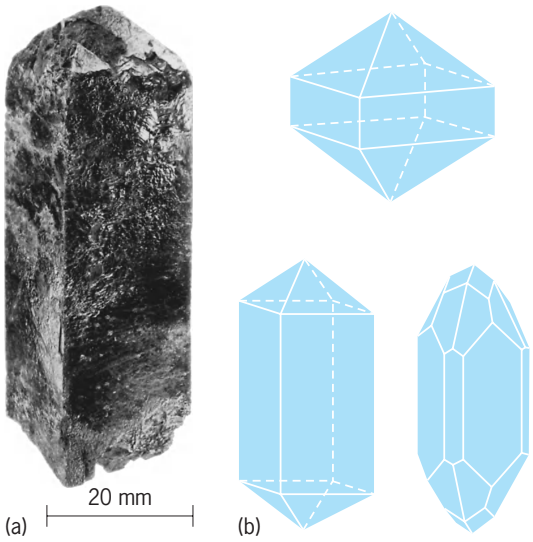

Zircon. (*a*) Crystal from Eganville, Ontario, Canada (*American Museum of Natural History specimen*). (*b*) Crystal habits (*after C. Klein and C. S. Hurlbut, Jr., Manual of Mineralogy, 21st ed., John Wiley and Sons, 1993*).

in small amounts as an accessory mineral in granitic and syenitic igneous rocks, and occurs more abundantly in pegmatite deposits associated with such rocks. Zircon also is a minor constituent in many types of gneissic and schistose metamorphic rocks. *See* HEAVY MINERALS.

Clifford Frondel

Zirconium

A chemical element, Zr, atomic number 40, atomic weight 91.22. Its naturally occurring isotopes are 90, 91, 92, 94, and 96. Zirconium is one of the more abundant elements, and is widely distributed in the Earth's crust. Being very reactive chemically, it is found only in the combined state. Under most conditions, it bonds with oxygen in preference to any other element, and it occurs in the Earth's crust only as the oxide, ZrO_2, baddeleyite, or as part of a complex of oxides as in zircon, elpidite, and eudialyte. Zircon is commercially the most important ore. Zirconium and hafnium are practically indistinguishable in chemical properties, and occur only together. *See* HAFNIUM; PERIODIC TABLE; ZIRCON.

1																	18
1 H	2											13	14	15	16	17	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg	3	4	5	6	7	8	9	10	11	12	13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	71 Lu	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	103 Lr	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112	113					

lanthanide series	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb

actinide series	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No

Most of the zirconium used has been as compounds for the ceramic industry: refractories, glazes, enamels, foundry mold and core washes, abrasive grits, and components of electrical ceramics, The incorporation of zirconium oxide in glass significantly increases its resistance to alkali. The use of zirconium metal is almost entirely for cladding uranium fuel elements for nuclear power plants. Another significant use has been in photo flashbulbs.

Zirconium is a lustrous, silvery metal, with a density of 6.5 g/cm^3 (3.8 oz/in.^{3}) at 20°C (68°F). It melts at about 1850°C (3362°F). Estimates of the boiling point from appropriate data have commonly been of the order of 3600°C (6500°F), but observations suggest about 8600°C (15,500°F). The free energies of formation of its compounds indicate that zirconium should react with any nonmetal, other than the inert gases, at ordinary temperatures. In practice, the metal is found to be nonreactive near room temperature because of an invisible, impervious oxide film on its surface. The film renders the metal passive, and it remains bright and shiny in ordinary air indefinitely. At elevated temperatures it is very reactive to the non-metallic elements and many of the metallic elements, forming either solid solutions or compounds.

Zirconium generally has normal covalency of 4, and commonly exhibits coordinate covalencies of 5, 6, 7, and 8. Zirconium is at oxidation number 4 in nearly all of its compounds, Halides in which its oxidation numbers are 3 and 2 have been prepared. While zirconium is often part of cationic or anionic complexes, there is no definite evidence for a monatomic zirconium ion in any of its compounds.

Most handling and testing of zirconium compounds have indicated no toxicity. There has generally been no ill consequence of contact of zirconium compounds with the unabraded skin. However, some individuals appear to have allergic sensitivity to zirconium compounds, characteristically manifested by appearance of nonmalignant granulomas. Inhalation of sprays containing some zirconium compounds and of metallic zirconium dusts have had inflammatory effects.

Warren B. Blumenthal

Bibliography. F. A. Cotton et al., *Advanced Inorganic Chemistry*, 6th ed., Wiley-Interscience, 1999; J. Hala (ed.), *Halides, Oxyhalides and Salts of Halogen Complexes of Titanium, Zirconium, Hafnium, Vanadium, Niobium and Tantalum*, 1989; M. F. Lappert, *Comprehensive Organometallic Chemistry II: Scandium, Yttrium, Lanthanides and Actinides, and Titanium, Zirconium, and Hafnium*, vol. 4, 1995.

Zoanthidea

An order of the cnidarian subclass Hexacorallia. Sometimes known popularly as "mat anemones" and technically also as Zoanthiniaria (the name Zoantharia is an alternative for subclass Hexacorallia), this order contains relatively few species. *See* ANTHOZOA; CNIDARIA; HEXACORALLIA.

Morphology. The column of one of these sedentary, anemonelike anthozoans is typically encrusted with sand grains, sponge spicules, tests of formaniferans, and other detritus that invades the mesoglea, thus acting as a sort of adventitious skeleton. Zoanthid musculature is poorly developed. The basal end of most is somewhat or deeply embedded in a mass of tissue that connects members of the colony; in solitary zoanthids, a pedal disk is absent. The simple, unbranched tentacles are arrayed in two cycles. As in some other hexacorallians, the mesenteries are paired and coupled, but otherwise the mesenterial arrangement is distinctive, differing slightly between two groups of zoanthids. In both groups—suborders Brachycnemina and Macrocnemina—the mesenteries that connect with the single siphonoglyph (termed the sulcal mesenteries, conventionally considered ventral in position) are complete, and the asulcal pair (those diametrically opposite the siphonoglyph, on the "dorsal" side of the animal) are incomplete. However, in the suborder Brachycnemina each nondirective pair (that is, each pair not connected to the siphonoglyph) consists of one complete and one incomplete mesentery (thus an animal has only one pair of complete mesenteries), whereas in the suborder Macrocnemina, both members of the second pair of mesenteries on each side of the asulcal pair are complete (thus an animal has three pairs of complete mesenteries). Mesenteries are added only in the exocoels on either side of the sulcal mesenteries; the larger mesentery of each pair bears filaments and is gametogenic.

Distribution. Members of Brachycnemina are abundant in warm, shallow waters, including on coral reefs, and typically harbor zooxanthellae; they comprise the families Neozoanthidae, Sphenopidae, and Zoanthidae. The Macrocnemina is characteristic of the deep sea, where colonies of some species encrust objects such as sea whips, and others associate with hermit crabs, forming a carcinoecium around the crustacean's abdomen; the families of this group are Epizoanthidae, Gerardiidae, and Parazoanthidae.

Ecology. Like all cnidarians, zoanthids are carnivorous. Bits of food dropped by their hosts may be important to those living with hermit crabs. Zooxanthellate zoanthids may derive much of their energy from their photosymbionts. Zoanthids of at least one species are bioluminescent. One species of *Palythoa* forms a nonproteinaceous toxin that was ostensibly used by Hawai'ian warriors to tip their spears; unlike other cnidarian toxins, palytoxin is not associated with nematocysts, but seems to be produced within the gastrovascular cavity, perhaps by symbiotic microorganisms, and functions to protect the eggs from predation.

Reproduction. Some zoanthids are hermaphroditic (individuals are both male and female) and some gonochoric (individuals are either male or female); most broadcast their gametes but some brood. Unique to zoanthids are Semper's larvae, of which there are two types: the zoanthina, with a girdle of long cilia and 12 mesenteries, is characteristic of the genera *Isaurus* and *Zoanthus*; the zoanthella, with a ventral band of long cilia extending from the oral to the aboral end, is characteristic of the genera *Palythoa*, *Protopalythoa*, and *Sphenopus*. Other zoanthids produce pelagic larvae other than Semper's. Daughter polyps of a colony arise by budding from the stolon or the polyp base, or by longitudinal or fission; transverse fission occurs in the solitary zoanthid *Sphenopus*.

Phylogeny. The phylogenetic position of the Zoanthidea is obscure. These animals have some features reminiscent of the Paleozoic rugosan corals, but because nothing is known about the soft-tissue morphology of the latter, and the former do not form skeletons, which are all that is known of rugosans, it is impossible to determine if this resemblance is phylogenetically significant or just coincidence. Molecular data support the Antipatharia (black corals) as the nearest extant relative of zoanthids. *See* ANTHOZOA.

Daphne G. Fautin

Bibliography. H. Erhardt and D. Knop, Order Zoantharia, *in Corals: Indo-Pacific Field Guide*, pp. 269–278, Ikan, Frankfurt, 2005; J. P. Hoover, Zoanthids (colonial anemones), in *Hawai'i's Sea Creatures: A Guide to Hawai'i's Marine Invertebrates*, pp. 42–45, Mutual Publishing, Honolulu, 1998; R. L. Manuel, Order Zoantharia Gray, 1870, *in British Anthozoa*, pp. 70–83, Linnaean Society, London, 1870.

Zodiac

The band of sky through which the Sun, Moon, and planets apparently move in the course of the year. The Babylonians, about 2500 years ago, divided the zodiac into 12 parts, which correspond to constellations. These zodiacal constellations, in order around the sky, are Aries, the Ram; Taurus, the Bull; Gemini, the Twins; Cancer, the Crab; Leo, the Lion; Virgo, the Virgin; Libra, the Scales; Scorpius, the Scorpion; Sagittarius, the Archer, Capricornus, the Sea Goat; Aquarius, the Water Carrier; and Pisces, the Fish. These constellations are based on Greek myths.

Because of the precession of the equinoxes, the positions of the constellations in the sky have drifted from the dates of the year with which they were associated thousands of years ago. Thus the popular astrological "signs" of the zodiac are not actually those that currently correspond to the sky. Though the vernal equinox is often called the first point of Aries, precession has moved it into Pisces, not far from Aquarius. Contemporary astrologers claim not to base detailed horoscopes on these outmoded zodiacal dates; nonetheless there is no basis for astrology, and tests have shown that astrological predictions are no more valid than chance or than profiles based on psychology dealing with those receiving the horoscopes. *See* PRECESSION OF EQUINOXES.

The Sun actually passes through parts of 13 constellations, as currently defined. Also, if the zodiac is defined as the region within latitudes $\pm 8^\circ$, which accommodates the eight planets through Neptune,

it contains all or part of 24 constellations. *See* CONSTELLATION; ECLIPTIC. Jay M. Pasachoff

Bibliography. M. Lachièze and J.-P. Luminet, *Figures du Ciel*, Bibliothèque Nationale de France, Seuil, 1998; J. Meeus, *More Mathematical Morsels*, William-Bell, 2002; J. Mosley, The real, real constellations of the zodiac, *The Planetarian*, 28(4):5, December 1999; J. M. Pasachoff, *Field Guide to the Stars and Planets*, 4th ed., Houghton Mifflin, Boston, rev. 2003; C. Stott (ed.), *Images of the Universe*, Cambridge University Press, 1991.

Zodiacal light

A diffuse, night-sky luminosity easily seen at low to middle geographic latitudes in the absence of moonlight. It is caused by sunlight scattered and absorbed by interplanetary (solar system) dust particles. Zodiacal light extends over the entire sky, but it is brightest toward the Sun and in the zodiacal band. (In the ecliptic at 30° from the Sun its visual brightness is three times that of the brightest part of the Milky Way.) It is best seen in the west after evening twilight and in the east before morning twilight (**Fig. 1**), when the ecliptic is close to the vertical. In the Northern Hemisphere this corresponds to spring evenings and autumn mornings. *See* ECLIPTIC; INTERPLANETARY MATTER.

The visual zodiacal light brightness decreases monotonically with elongation (angular distance from the Sun) to a relatively flat minimum in the ecliptic at 120 to 140°, after which it gradually increases out to the antisolar point at 180°. The corresponding polarization reaches a maximum of approximately 20% at 70° elongation, falling to zero at 160°. The polarization changes direction between 160 and 180°, where it is again zero. Such characteristics of brightness and polarization, when combined with information on color and angular dependence in and out of the ecliptic can be used to provide information on the optical properties and spatial distribution of the dust. Ground and space observations have shown the zodiacal light to have approximately the same color as the Sun. It has been shown to be remarkably constant (to 1%) over periods as long as a solar cycle (11 years). Exceptions are small seasonal variations due to the motion of the observer with respect to the plane of the dust and occasional plasma radiation, mostly light associated with coronal mass ejections (CMEs). The density of particles responsible for the zodiacal light falls off somewhat faster than $1/R$, where R is heliocentric distance. The particles are primarily in the size range of tens to hundreds of micrometers in diameter, and are believed to originate primarily from comets. Band structure seen in infrared measurements of thermal emission from the dust may be due to particles arising from collisions in and near the asteroid belt. Ultraviolet and, especially, infrared observations from space are providing information that will lead to a better understanding of the overall physics of the interplanetary dust. *See* ASTEROID; COMET; INFRARED ASTRONOMY; SOLAR WIND; ULTRAVIOLET ASTRONOMY.

Fig. 1. Morning zodiacal light as it typically appears at a low, northern latitude site, Mount Haleakala, Hawaii, at an altitude of 10,000 ft (3300 m). The outcropping of rock at the base of the zodiacal cone is approximately 20° from the Sun. (*P. B. Hutchison*)

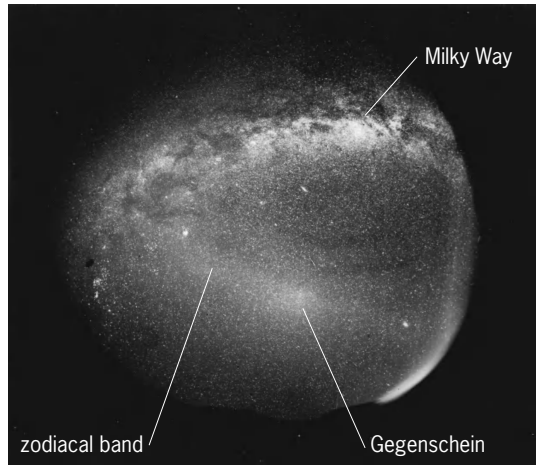

Fig. 2. All-sky photograph, taken from Mount Komagatake, Japan, showing the zodiacal band and Gegenschein, and the Milky Way.

Gegenschein. The enhanced brightness near the antisolar point is called the Gegenschein or counterglow (**Fig. 2**). It is barely above the visible threshold and is described by visual observers as being oval in appearance, 6 by 10° or larger, with the long axis in the ecliptic. Space observations have shown the Gegenschein to be an intrinsic part of the zodiacal light, rather than a separate phenomenon due to the Earth's atmosphere or to a concentration of dust opposite the Earth from the Sun; that is, the zodiacal dust particles have an increased scattering efficiency near backscattering.

Fraunhofer corona. The solar corona has two principal parts: a K-corona, which is due to Thomson scattering by free electrons located very close to the Sun, and an F-corona (also referred to as Fraunhofer or false corona or inner zodiacal light), which arises from scattering by zodiacal dust seen at small elongations. Although the Earth's atmosphere complicates attempts to observe zodiacal light closer than approximately 30° to the Sun (equivalent to seeing no particles closer than 0.5 astronomical unit to the

Sun), eclipse and space observations have shown the brightness to increase smoothly all the way in to the F-corona. The F-corona or inner zodiacal light and the "primary" zodiacal light seen at larger elongations come primarily from dust particles located relatively close to the Sun. Near-Sun multicolor (ultraviolet to visual to infrared) observations from space can be expected to add significantly to knowledge of the physics, dynamics, and chemistry of dust near the Sun and in the transition to interplanetary space. *See* SOLAR CORONA; SUN. J. L. Weinberg

Bibliography. E. Grün et al. (eds.), *Interplanetary Dust*, Springer, 2001; C. Leinert et al., The 1997 reference of diffuse night sky brightness, *Astron. Astrophys. Suppl. Ser.*, 127:1–99, 1997; Proceedings of Zodiacal Cloud Sciences, Kobe, Japan, September 1997, *Earth Planets Space*, 50:463–610, 1998.

Zone refining

One of a number of techniques used in the preparation of high-purity materials. The technique is capable of producing very low impurity levels, namely, parts per million or less in a wide range of materials, including metals, alloys, intermetallic compounds, semiconductors, and inorganic and organic chemical compounds. In principle, zone refining takes advantage of the fact that the solubility level of an impurity is different in the liquid and solid phases of the material being purified; it is therefore possible to segregate or redistribute an impurity within the material of interest. In practice, a narrow molten zone is moved slowly along the complete length of the specimen in order to effect the impurity segregation. *See* SOLUTION; SOLVENT.

Impurity atoms either raise or lower the melting point of the host material. There is also a difference in the concentration of the impurity in the liquid phase and in the solid phase when the liquid and solid exist together in equilibrium. In zone refining, advantage is taken of this difference, and the impurity atoms are gradually segregated to one end of the starting material. To do this, a molten zone is passed from one end of the impure material to the other, as in **Fig. 1***a*, and the process is repeated several times. The end to which the impurities are segregated depends on whether the impurity raises or lowers the melting point of the pure material; a lowering of the melting point is more common, in which case impurities are moved in the direction of travel of the molten zone.

The extent or effectiveness of the impurity distribution is controlled by the distribution coefficient, k, defined as the ratio of the impurity concentration in the solid to that in the liquid at a given temperature. The effect of multiple zone passes (in the same direction) on impurity content along the material is illustrated in Fig. 1*b*.

Zone leveling is a particular form of zone refining. Here zone passes are made in alternate directions so that a uniform (homogeneous) distribution of impurities (that is, alloying) is produced along the material. In this way desired levels of deliberately added impurities may be achieved in this form of zone refining.

In commercial application the object is to optimize the variables in order to produce the required amount of material having the necessary level of purity with a minimum of time and expense. Primary variables are zone length, interzone spacing in a multiple zone batch unit, number of zone passes, and travel rate of the zone along the material. The choice of zone length is governed by practical considerations and by the physical properties of the material; a narrow molten zone with a sharp demarcation between liquid and solid phases is desired. Small numbers of zone passes reduce the total time required but at the expense of the usable length of material. For large numbers of zone passes, the best segregation occurs for small zone lengths. Increasing the travel speed decreases the time per pass, but there is a concomitant decrease in the efficiency of segregation. Typical conditions for the zone refining of aluminum to a purity better than 99.9995% are a zone speed of 1 in./h (2.5 cm/h), 30 zone passes, and a zone length to material lenght ratio of 0.1.

The many designs of zone-refining apparatus center on factors such as the form of heating (or cooling), the form of container (material or shape), the travel mechanism, stirring of the liquid in the molten zone, and protective atmosphere. Common modes of heat input for material in a container are resistance heating and induction heating (3 Hz to 6 MHz in the 5- to 50-kW power output

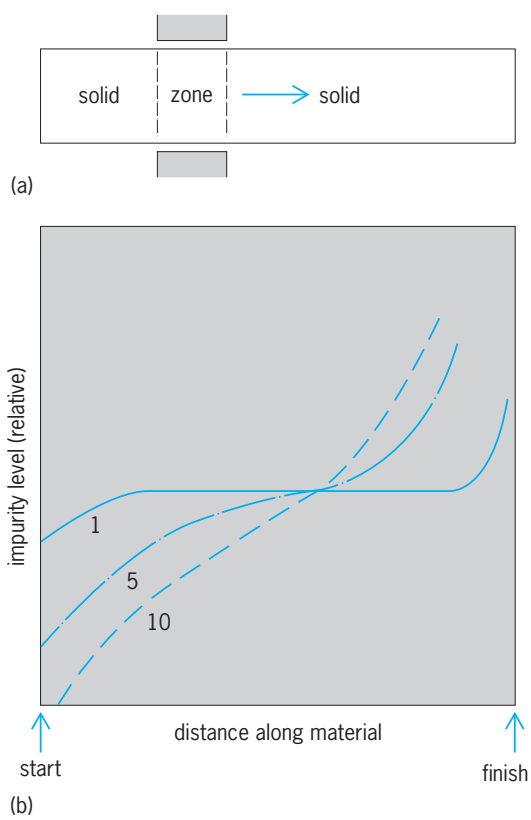

Fig. 1. Zone refining. (*a*) Passage of a molten zone along the material to be purified. (*b*) Effect of 1, 5, and 10 zone passes on the impurity distribution along the material.

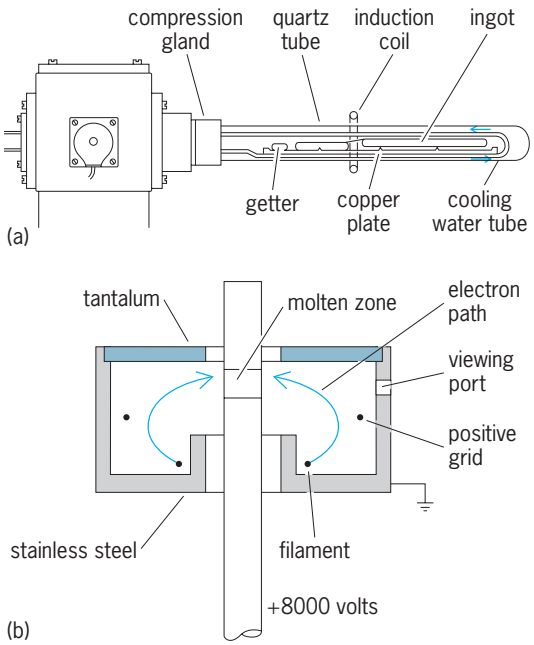

Fig. 2. Industrial techniques. (*a*) Zone refining on a water-cooled hearth. (*b*) Electron-beam zone-refining arrangement. (*After R. F. Bunshah, ed., Techniques of Metals Research, vol. 1, John Wiley and Sons, 1968*)

range). When the material to be purified melts below room temperature, it is necessary to zone-refine in a refrigerated atmosphere or to traverse a solid zone along the bar or column of liquid material. Container materials in use are plastics, metals, glass, graphite, and oxides of aluminum, zirconium, magnesium, silicon, and beryllium. The choice becomes more restrictive as the temperature of the molten zone increases. The upper limit of melting point for metals and alloys to be zone-refined in a container is ~2700°F (1500°C), since excessive reaction occurs with the container above this temperature. Low-melting-point organic compounds are usually purified in glass containers or plastics. High-melting-point or reactive metals, compounds, and semiconductors may be zone-refined in a container if the latter is water-cooled, as in **Fig. 2***a*. *See* RESISTANCE HEATING.

In floating zone refining, a molten zone is held in place by its own surface tension between two vertical colinear rods. Since no container is needed, contamination problems are avoided. The most common form of heating is electron bombardment (Fig. 2*b*), or induction heating. In this way refractory metals and compounds can be zone-refined. *See* INDUCTION HEATING.

The importance of the control and improvement of purity lies in the fact that, in most cases, the behavior of a material is a direct function of impurity content; the material behavior is manifested in mechanical, electrical, chemical, thermal, and optical properties. *See* METALLURGY. Alan Lawley

Bibliography. A. Choudhury, *Vacuum Metallurgy*, 1990; *Kirk-Othmer Encyclopedia of Chemical Technology*, 4th ed., 1997; W. G. Pfann and D. T. Hawkins, *Zone Melting*, 1966, reprint 1978.

Zooarcheology

Zooarcheology or archeozoology is the study of animal remains from archeological sites. Most such remains derive from people's meals. In other words, zooarcheology is essentially the study of ancient garbage, mainly bones and teeth of mammals, as well as birds, fish, mollusks, and even insects. Zooarcheology helps to provide a more complete picture of people's environment and way of life, especially their economy, as reflected in the relationship between people and animals 100,000 or even just 100 years ago.

Methods. The first task of the zooarcheologist is the taxonomic and anatomical identification of the faunal remains, ideally with the aid of a reference collection of modern identified specimens. Bone measurements help identify species size changes in the archeological record, and they can help to separate closely related species, such as sheep and goats, and domestic animals and their wild ancestors. Age-at-death in many mammals can be determined by the state of eruption and wear of the teeth and fusion of long-bone epiphyses.

A zooarcheological study is limited by the number of remains. Thus, 10 bones will reveal which species were exploited; 100 bones can reveal roughly in what proportion they were exploited; and 1000 bones may provide intraspecific information, such as the proportion of different age groups and sexes culled from the herd. In a sample of 10,000 bones, there should be enough mandibles to plot a continuous age distribution of the herbivores.

Bones are usually damaged from butchery, trampling, cooking, and gnawing in antiquity. Many have been subject to varying degrees of erosion in the soil. (These environmental alterations are included in the subject of taphonomy.) An archeological assemblage of bones, therefore, comprises numerous broken limb bone ends, isolated teeth, and unidentifiable fragments. The method of excavation can also have an effect, especially when the soil is not sieved, leading to the poor recovery of smaller bones and isolated teeth.

Environmental reconstruction. Animal remains often indicate global climate patterns of the past. For instance, in cold periods (such as the Ice Ages), cold-weather species were distributed in what are temperate areas today. Given sufficiently large samples in a multiperiod site, varying frequencies of species can reflect climate changes. At Combe-Grenal, France, Francois Bordes and Prat interpreted geologic levels with large numbers of reindeer as having formed during cold periods, layers with more red deer and wild cattle as having formed during warmer periods. At Boomplaas cave in South Africa, Richard Klein studied the animal remains and found that the lowermost and uppermost levels date back to the last Interglacial and present Interglacial (Holocene), respectively. These levels contain a predominantly woodland and bush fauna with hyrax and baboons. The intermediate levels date between 20,000 B.C. and 10,000 B.C. and were

found to contain a fauna consisting predominantly of zebras and wildebeest. These are more typical of grassland and, according to Klein, reflect cool, dry conditions which prevailed in South Africa during the late glacial period. An increase or decrease in species diversity may reflect milder or harsher environmental conditions, respectively. *See* CAVE.

The size of mammal bones and teeth can indicate temperature change. Today, mammals are often larger in colder areas (Bergmann's rule). In several zooarcheological assemblages (including South Africa and Israel) spanning the end of the last Ice Age and Holocene period (12,000 years ago), many mammals became smaller as the temperature increased. *See* CLIMATE HISTORY.

Hunting. It is clear that toward the end of the Pleistocene, people were effective hunters of big game. They occupied most of the continental landmass and subsequently oceanic islands as well. The arrival of hunting people in Australia, the Americas, and many islands coincided with the extinction of many large animals such as the giant kangaroo, giant wombat, and diprotodon in Australia; the ground sloth, giant beaver, horse camel, mammoth, and mastodon in North America; the moas in New Zealand; and the pygmy elephant and hippotamus on some Mediterranean islands. This catastrophic loss of fauna has been called "megafaunal overkill."

Zooarcheology seeks evidence of an association between early humans invaders and the extinction of the animals. In the United States, there are a number of localities where, for example, mammoth remains are associated with human artifacts dated between 9200 and 9000 B.C. However, such evidence is scarce, presumably because extinctions happened rapidly and the coexistence of people and megafauna was brief. Although it is debated whether climate change or overhunting (or both) was responsible for the extinctions in America, overhunting was probably the main cause.

In Africa and Eurasia, where humans and animals had coevolved for a long time, extinction was less abrupt. However, by 9000 years ago in the Near East, zooarcheological studies indicate that people exploited fewer large animals—which had presumably become scarce or even extinct. People exploited more small mammals such as hares, fish, and birds. The mortality curves of gazelles hunted in the Levant between 50,000 and 10,000 years ago show that by the end of this period, greater numbers of young were being hunted. This and the shift to hunting of smaller animals may reflect increased pressure on the resources by increased numbers of people. Continued demographic pressure subsequently forced people to husband animals (and plants) and colonize new territories such as oceanic islands. *See* PLEISTOCENE.

Domestication. One challenge for zooarcheologists is to try and discover when animals were first domesticated and why. The shift from hunting to husbanding may be recognized zooarcheologically in four ways:

1. *Size change.* Species size changes can occur due to human selective breeding practices. For example, contempoary dogs, cattle, sheep, and pigs are smaller than their wild ancestors: wolves, aurochs (wild cattle), mouflon sheep, and wild boar. However, domestic guinea pigs are larger than wild guinea pigs. Care is needed to avoid confusing size variation due to environmental change with size changes caused by human intervention. Size can also vary with sex and age, so these variables need to be controlled when considering measurements to distinguish between bones of wild and domestic mammals.

2. *Importation of exotic species.* Before the sixth millennium B.C., there were no sheep in Europe. Sheep bones found on sixth-millennium archeological sites in Europe must, therefore, derive from introduced and thus, domesticated animals. A similar example is the shipment of sheep, goats, cattle, cats, and dogs to Cyprus in the eighth millennium B.C.

3. *Species spectrum change.* In several parts of the Near East, a shift from gazelle to sheep and goats occurred in the seventh millennium B.C. At Franchthi cave in Greece, Sebastian Payne noticed a shift from horse, deer, wild boar, and ibex to sheep (previously unknown in Greece), goats, pigs, and cattle in the sixth millenium B.C. Both Near Eastern and Greek faunal shifts probably signify the change from hunting to husbanding.

4. *Cultural signs.* Several burials of dogs with humans have been excavated in Israeli Natufian (an Epipaleolithic culture) sites. They date back to the tenth millennium B.C., and they are the earliest known burials of this kind. At a Mesolithic site in Switzerland, Louis Chaix interpreted a pathological deformation between the first and second molar teeth of a brown bear as caused by a leather thong, indicating that it was probably kept as a pet. Shed caprine (goat) milk teeth (first teeth) in early Neolithic Greek and French sites suggest penning of sheep and goats.

The earliest evidence for domesticated animals in the Old World is of dogs in the Near East about 10,000 B.C., followed one or two thousand years later by the food animals such as sheep, goat, cattle, and pig. The origins of domesticated animals in eastern Asia (such as the pig and chicken) and the New World (including the llama, alpaca, and guinea pig) are less well known. *See* ARTIODACTYLA; DOGS; NEOLITHIC; PALEONTOLOGY.

Secondary uses of animals. Patterns of exploitation of animals for uses other than meat—for example, sheep and goats for their wool and milk—is more difficult to identify since wool and milk are rarely preserved. Thus a more indirect approach is adopted. Since in a meat-producing economy animals are slaughtered in their prime, high numbers of juvenile bones and teeth are found in an archeological assemblage. If wool is the prime objective, slaughter may be delayed until animals are 5 or more years of age, when teeth are well worn and epiphyses fused. There is some evidence, again from the Near East, of a shift in sheep and goat culling patterns in the fifth or fourth millennium B.C., perhaps reflecting the beginning of wool or milk exploitation.

Power is another secondary use of animals. Equids, camelids, and cattle have long been an important

source of power for transport and plowing. Excessive strain in life can cause joint disease, sometimes recognized as bony outgrowths and fusions of foot bones, though old age or soft soil may also induce the same arthropathies. Properly controlled samples of modern animals used for traction are needed to test this assumption.

Other uses of animals. Animal remains can sometimes provide other kinds of useful information about the human inhabitants of a site. Remains of migratory birds (for examples, winter visitors) can provide information such as the season an animal resource was exploited. Animals or specific parts of them were often sacrificed and used ritually. For example, the body of a man buried in Bronze Age Northamptonshire, England (Irthlingborough), was covered with 185 cattle skulls and the mandibles, scapulae, and pelves of 20 or 30 cattle. No other parts of the animal skeletons were present. Bones were sometimes used as construction materials. For example, in the Ice Age of Ukraine where wood was scarce, mammoth bones and tusks were used to construct the scaffolding for shelters, and presumably then covered with skins.

Overview. Zooarcheology is a multidisciplinary endeavor requiring knowledge of anatomy and biometry as well as an appreciation of the archeological questions that need to be addressed. Unlike most paleontological collections, zooarcheological collections are usualy well dated and comprise large numbers of bones. They provide excellent opportunities to study microevolution. However, much remains to be explored in this relatively new science. *See* ANTHROPOLOGY; ARCHEOLOGY; FOSSIL HUMANS; ZOOLOGY.

Simon J. M. Davis

Bibliography. S. J. M. Davis, *The Archaeology of Animals*, Yale University Press, New Haven, 1987; R. G. Klein and K. Cruz-Uribe, *The Analysis of Animal Bones from Archaeological Sites*, University of Chicago Press, 1984; P. S. Martin and R. G. Klein (eds.), *Quaternary Extinctions: A Prehistoric Revolution*, University of Arizona Press, 1984; E. J. Reitz and E. S. Wing, *Zooarchaeology*, Cambridge University Press, 1999; B. Wilson, C. Grigson, and S. Payne (eds.), *Ageing and Sexing Animal Bones from Archaeological Sites*, BAR British Series, 109, Oxford, 1982.

Zoogeography

The subdivision of the science of biogeography that is concerned with the detailed description of the distribution of animals and how their past distribution has produced present-day patterns. Scientists in this field attempt to formulate theories that explain the present distributions as elucidated by geography, physiography, climate, ecological correlates (especially vegetation), geological history, the canons of evolutionary theory, and an understanding of the evolutionary relationships of the particular animals under study. Zoogeographical theories are then tested by new data from all germane fields to amplify, verify, or falsify the constructs. In this sense, zoogeography is an integrative science that synthesizes data from other disciplines to apply to the realities of animal distribution.

Realities. The field of zoogeography is based upon five observations and two conclusions. The observations are as follows. (1) Each species and higher group of animals has a discrete nonrandom distribution in space and time (for example, the gorilla occurs only in two forest areas in Africa). (2) Different geographical regions have an assemblage of distinctive animals that coexist (for example, the fauna of Africa south of the Sahara with its monkeys, pigs, and antelopes is totally different from the fauna of Australia with its platypuses, kangaroos, and wombats). (3) These differences (and similarities) cannot be explained by the amount of distance between the regions or by the area of the region alone [for example, the fauna of Europe and eastern Asia is strikingly similar although separated by 6900 mi (11,500 km) of land, while the faunas of Borneo and New Guinea are extremely different although separated by a tenth of that distance across land and water]. (4) Faunas strikingly different from those found today previously occurred in all geographical regions (for example, dinosaurs existed over much of the world in the Cretaceous). (5) Faunas resembling those found today or their antecedents previously occurred, sometimes at sites far distant from their current range (for example, the subtropical-warm temperate fauna of Eocene Wyoming, including many fresh-water fishes, salamander, and turtle groups, is now restricted to the southeastern United States).

The conclusions are as follows. (1) There are recognizible recurrent patterns of animal distribution. (2) These patterns represent faunas composed of species and higher groups that have evolved through time in association with one another.

Approaches. Two rather different approaches have dominated the study of zoogeography since the beginning of the nineteenth century: ecological and historical. Ecological zoogeography attempts to explain current distribution patterns principally in terms of the ecological requirements of animals, with particular emphasis on environmental parameters, physiological tolerances, ecological roles, and adaptations. The space and time scales in this approach are narrow, and emphasis is upon the statics and dynamics of current or very recent events.

Historical zoogeography recognizes that each major geographical area has a different assemblage of species, that certain systematic groups of organisms tend to cluster geographically, and that the interaction of geography, climate, and evolutionary processes over a long time span is responsible for the patterns or general tracks. Emphasis in this approach is upon the statics and dynamics of major geographical and geological events ranging across vast areas and substantial time intervals of up to millions of years. The approach is based on concordant evolutionary association of diverse groups through time.

Ecological zoogeography is the study of animal distributions in terms of their environments; historical zoogeography is the study of animal distributions in terms of evolutionary history.

Comparison of three grassland biomes

Ecological roles	North American	African	Australian
Top predator	Wolf (dog)	Lion (cat)	Tasmanian wolf (marsupial)
Large herbivore	Bison (cattle)	Zebra (horse)	Red kangaroo (marsupial)
Small predator-scavenger	Coyote (dog)	Hyena (hyena)	Tasmanian devil (marsupial)

Ecological zoogeography. Ecological zoogeography is rooted within the discoveries of nineteenth-century plant ecologists and physiologists. They found that under similar conditions of temperature and moisture, terrestrial plants develop similar growth forms, regardless of their evolutionary relationships, to produce one of the following vegetation forms: forest, woodland, savanna, grassland, or scrub. Subsequently it was demonstrated that these units showed differentiation associated with major soil types and broad latitudinal climatic regions. Each of these main kinds of vegetation is called a formation type (such as tropical lowland evergreen forest), and its geographical subdivisions are formations (such as South American tropical lowland evergreen forest; African tropical lowland evergreen forest).

Later investigators realized that within each formation a series of animals had evolved to undertake homologous ecological functions in the dynamics of the community, so that the concept of biome-type (vegetation and associated animals) was developed. The biome-type is a series of major geographical climatic regions characterized by similar ecological adaptations in plants and animals. For example, the grassland biome-type may be typified by the comparisons in the **table**. The biome-types are distinctive from the zoological point of view in that ecological equivalents (such as top predator) in the different biomes are usually from phylogenetically nonrelated stocks. M. D. F. Udvardy provides the best recent summary of the general distribution of world terrestrial biomes.

The concept of biomes is a method of generalizing the distribution of animals by major environments on a latitudinal basis. The ecological zoogeographer is also interested in vertical (altitudinal and bathymetric) zonation as a feature of animal distribution. The idea of zonation is based upon the recognition that there are ecoclinal gradients in the parameters of the environment (especially temperature) with an increase in altitude or in ocean depth. The composition of species distributions along these ecoclines, because of differences in ecological requirements, produces recognizable and characteristic life zones. As an example, **Fig. 1** shows the principal life zones in the ocean. These zones are divided into two groups: benthos or substrate zones and pelagic or free-swimming zones. Altitudinal life zones on land have been similarly described. *See* BIOME; LIFE ZONES.

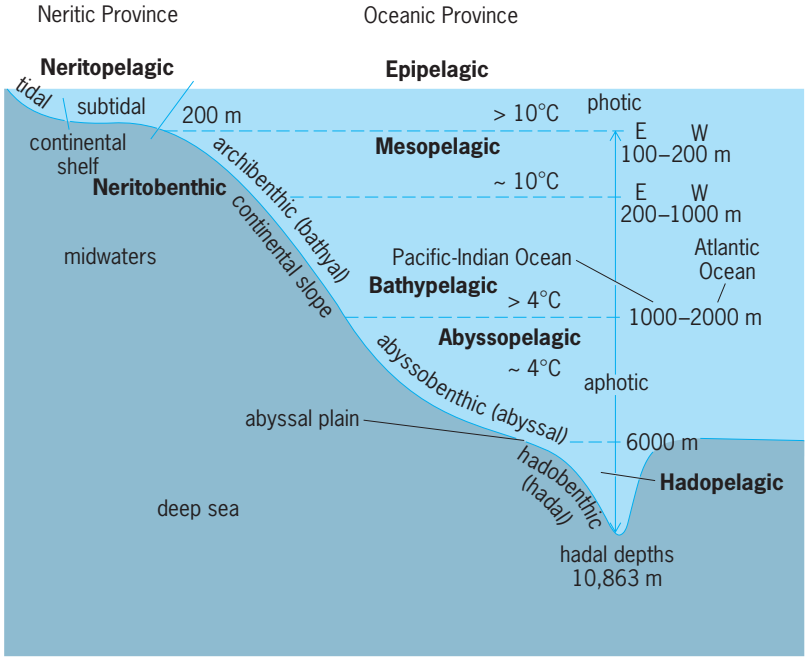

Fig. 1. Life zones in the sea. 1 m = 3.3 ft; °F = (°C × 1.8) + 32.

Historical zoogeography. This approach has its origins in systematic biology. Workers in this field recognized very early in the nineteenth century that different geographical regions support different faunas and that representatives of these faunas occur in a wide variety of environments (for example, apes and monkeys in Africa in rainforest to desert environments). In addition, they noted that different systematic groups tend to cluster geographically (kangaroos in Australia and anteaters in South America). With the development of the canons of evolutionary theory later in the century, a framework for understanding the evolution of faunas through time became the backbone of historical zoogeography.

The raw data of historical zoogeography are the distributions or tracks of individual species of animals in space (geographical ecology) and time. Because each species has its own set of peculiar ecological requirements and its own unique evolutionary history, each has a discrete nonrandom ecogeographical distribution. As a consequence, no species is universally present, and many species have small or unique tracks. The first level of generalization in zoogeography is based on the recognition that, in spite of the unique nature of individual species distributions, many tracks are concordant or show a common pattern. Determination of patterns involving the coincident distribution of species or monophyletic groups (genera, families, and so on) of species (generalized tracks) is the fundamental step in zoogeographical analysis.

The second level of generalization in this process

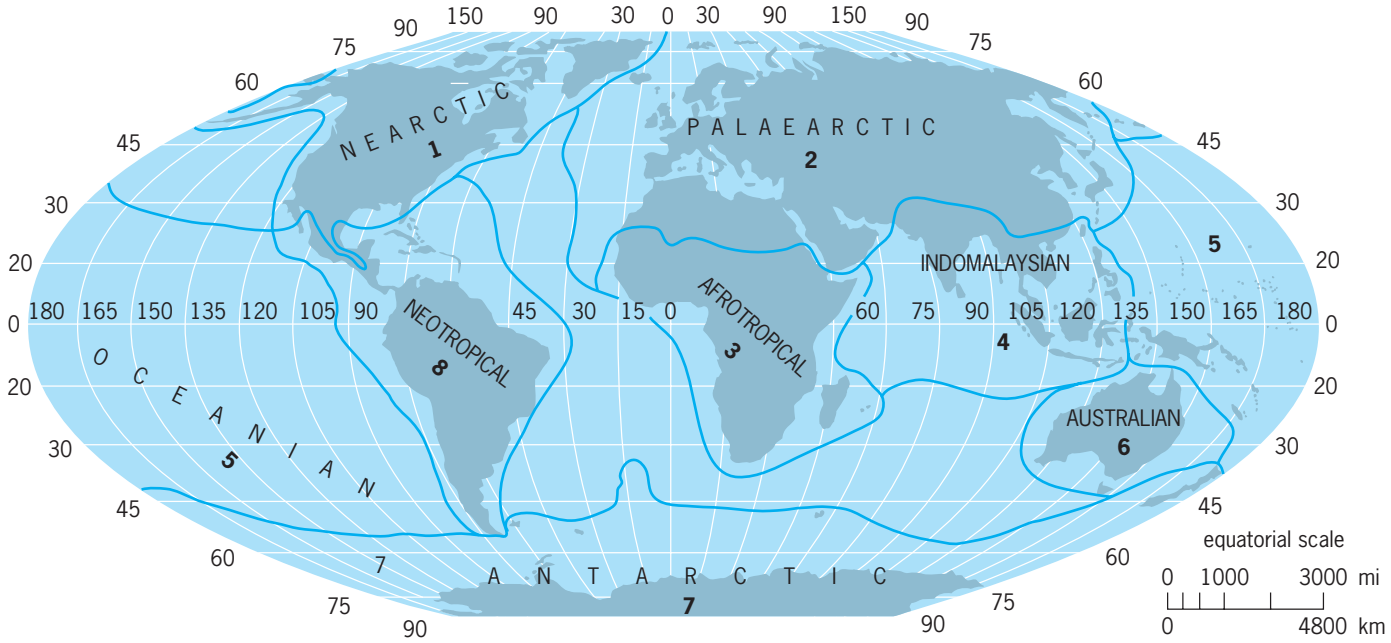

Fig. 2. Terrestrial biogeographic realms of the world. (*After M. D. F. Udvardy, A classification of the biogeographical provinces of the world, Int. Union Conserv. Nat. Occas. Pap., 18:1–49, 1975*)

is to cluster the strongly recurrent generalized tracks involving extensive geographical areas, whose components are then regarded as the major modern faunas. A third level of generalization attempts to tentatively identify the historical source units (ancestral faunas) that have contributed to the modern patterns.

Zoogeographical patterns. The early workers in zoogeography, especially P. Sclater in 1858 (birds) and A. Wallace in 1876, developed a system for classifying major patterns of terrestrial distributions, which in modified form (**Fig. 2**) is still applicable today. This system reflects the long-term isolation of major parts of the Earth's surface and the consequent divergent evolution of the fauna in each isolate. In some cases, subsequent reconnections of formerly isolated areas have blurred the distinctiveness of the core faunas in an area of transition. The major units in this system are called biogeographic realms; a series of biogeographic provinces form subdivisions within the larger divisions.

The eight recognized realms (which correspond to what were called zoogeographic regions in the Sclater-Wallace system and by many subsequent authors), with the number of provinces recognized for each realm indicated in parentheses are briefly characterized below. These units are based primarily on comparisons for the best-known animal groups. For this reason the several classes of vertebrates are used to characterize the realms below, although terrestrial and fresh-water invertebrates follow similar patterns but are less well known.

Nearctic Realm (22). North America, north of the edge of the Mexican-Guatemala highlands, although many workers would place the boundary with the Neotropical Realm at the edge of the Mexican plateau, west of the Isthmus of Tehuantepec. In terms of vegetation the northern areas support transcontinental zones of tundra and needle-leaf coniferous forests (taiga). The eastern portion of the continent to the south of the taiga zone was originally covered with broadleaf deciduous forests, and the central region by extensive grasslands. Coniferous forests predominate along the northwest coast and in the Rocky Mountains, Sierra Nevada, and upper regions of the Sierra Madre of Mexico. Much of the southwestern United States and adjacent Mexico are covered by oak savannas, scrub, and desertic vegetation. In terms of the fauna, a great many groups and species are shared with the Palearctic (that is, they have a holarctic distribution), and in more southern areas with the Neotropical Realm.

Distinctive vertebrate groups of the Nearctic include bowfin, gar-pike, bullhead catfishes, panfish, and basses; hellbenders, mud puppies, amphiumas and sirens, the bell-frog (*Ascaphus*), spadefoot toads, snapping turtles, the land tortoises (*Gopherus*), the Gila monster and its ally the Mexican beaded lizard (*Heloderma*), the night lizards, fence lizards and their allies, glass lizards, worm lizard (*Rhineura*), rattlesnakes, and alligator; flycatchers, vireos, orioles, hummingbirds (all shared with tropical America); racoons, pocket gophers (also tropical), and pronghorn. Many characteristic birds and mammals (woodpeckers, wood warblers, tanagers, pikas, flying squirrels, beaver, jaguar, and cougar) belong to wide-ranging groups, with endemic species in the New World.

Palearctic Realm (44). The Eurasian landmass from southern China and the southern slopes of the Himalaya Mountains northward and westward from the Indus Valley to include the Arabian Peninsula and northern Africa. As with the Nearctic Realm, transcontinental zones of tundra and taiga occur in

the northernmost areas. Broadleaf deciduous forest originally occurred over much of western (Europe) and eastern (China) portions of the realm, and the central areas of the landmass were steppe, grassland, or desert. Scrub-to-desert vegetation is typical of the Mediterranean area east to the Indus. The Palearctic fauna is a peculiar mixture of widespread groups with strong affinities to the Nearctic and others from the Old World tropics. For this reason, there is very little endemism at the major group level, and the distinctive groups are ones with a major radiation of species in the Palearctic. Distinctive vertebrate groups include a vast radiation of cyprinid fishes and numerous salmonids; many salamanders, lacertid lizards, slowworms, vipers; warblers, hedge sparrows (endemic); dormice, jerboas, the two-humped camel, and pandas.

Africotropical Realm (29). Africa south of the Sahara Desert and Madagascar. The realm is mostly tropical with broadleaf evergreen forests along the margins of the Gulf of Guinea and inland through the Congo Basin and on the eastern coast of Madagascar. Deserts cover the Kalahari, Namib, and the Somali Coast. Much of the rest of the continent was originally covered by tropical broadleaf deciduous forests which degraded to scrub-forest or savanna-grassland under human use. Montane evergreen forest occurs on the mountains of east Africa.

Although the realm shares many groups with tropical Asia and the Palearctic, it has one of the most easily recognizable faunas. Distinctive vertebrate groups for the continent include bichirs (lungfishes); caecilians, clawed frogs (*Xenopus*), reed frogs; side-neck turtles, chameleons, monitor lizards, sungazers, pythons, cobras, vipers, crocodiles; ostrich, secretary bird, hammerhead, touracos, mousebirds, helmet shrikes; otter shrew, golden mole, elephant shrews, monkeys, chimpanzees, gorilla, hyenas, aardvark, hyraxes, hippopotamuses, and spring haas. There are many genera and species of big game: the cats, rhinoceroses, buffalo, antelope, and pigs are African endemics. Madagascar lacks many important Continental groups but has a unique radiation of microhylid and reed frogs, specialized iguanids and boas, the endemic elephant birds (now extinct), flightless mesoenatids, endemic rollers, philapittas, and vangas; tenrecs, and lemurs, unique endemic mammals.

Indomalasian Realm (27). The Asian landmass from the west margin of the Indus Valley and south of the Himalaya Mountains and southern China and adjacent islands including Taiwan, the Philippines, and the Lesser Sundas. Most of the realm originally supported tropical forests with evergreen forests in western peninsular India, much of Burma, southeast Asia, southern China, the Malay Peninsula, and the Indo-Malayan Archipelago. Deciduous vegetation covered most of India and some areas of southeast Asia. Most of the latter has been converted to human use and degraded into scrub, grassland, or desert.

Distinctive groups of the Indo-Malayan include a high diversity of cyprinid fishes, loaches, labyrinth fishes; caecilians, rhacophorid tree frogs, microhylids; many endemic emydid turtles; gavial; pythons, pipe-snakes, cobras, sea snakes, pit vipers. pheasants, fairy bluebirds and leaf birds; hairy hedgehog, flying lemur, and spiny dormice. Endemics include monkeys, gibbons, orangutan, tarsiers, an elephant, a tapir, two rhinoceroses, and many antelopes.

Oceanian Realm (7). The Pacific islands of Melanesia, Micronesia, Polynesia, Hawaii, New Caledonia, and New Guinea. These islands are covered with tropical vegetation and show strong affinities with the Indo-Malayan unit. Many workers include New Guinea in the Australian Realm because of its great similarity in vertebrates to the latter. The remaining islands have a high degree of endemism, very few distinctive major groups, and a depauperate vertebrate fauna. There are no truly fresh-water fishes, frogs occur only on the islands near New Guinea and on the Fijis, snakes occur naturally only as far east as Fiji, while geckos and skinks range over the entire area and native flightless mammals occur only out to the Solomons.

Australian Realm (13). For these purposes, Australia and Tasmania, although many zoogeographers would include New Guinea and adjacent islands to the Solomons. Most of Australia is covered by grasslands, scrub, and desert, with temperate forests along the eastern and southeastern lowlands (including Tasmania) and tropical forests in the northwest. The Australian Realm has a unique fauna including these distinctive groups: an endemic lungfish, an osteoglossid (other fresh-water fishes are all marine-derived), myobatrachid frogs, pelodryadid tree frogs, side-neck turtles, a vast radiation of geckos and agamid lizards, pygopodids, pythons, many elapid snakes; cassowaries, emus, megapodes, owlet frogmouths, lyrebirds, scrubbirds, flowerpeckers, honey-eaters, bell magpies, magpie larks, bower birds, and birds of paradise; spiny anteater, platypus, marsupials—Tasmanian wolf, Tasmanian devil, anteater, bandicoots, Australian opossums, koala, wombats, kangaroos, wallabies, and wallaroos.

Antarctic Realm (4). The Antarctic continent, New Zealand, and the subantarctic islands. New Zealand has extensive temperate forests of mixed evergreen and deciduous trees over much of the North and South islands, with grassland predominating on parts of South Island. The subantarctic islands are relatively barren with vegetation primarily of grasses, sedges, and annual composites. The Antarctic continent is covered by ice and snow. Only New Zealand is of interest here because of its depauperate but unique fauna. It has no truly fresh-water fishes. A single amphibian, the frog *Leiopelma*, and the tuatara (*Sphenodon*), the only survivor of a Cretaceous order of reptiles, occur there. There are no turtles or snakes (except marine ones), but several endemic genera and species of geckos and skinks. Unique birds include the extinct moas, kiwis, an extinct flightless goose, flightless rails, a flightless parrot, and a predatory parrot (the kea). Native land mammals are two bats.

Neotropical Realm (47). Lowland Mexico, Central America, South America, and the Antilles. Evergreen

forests predominate along the Atlantic slope of Mexico, Central America, Panama, northwestern South America, and the Amazon Basin. Extensive deciduous forests formerly covered the lowlands of western Mexico and Central America, the Greater Antilles, and much of southern Brazil. Grassland and scrub vegetation cover extensive areas in central South America south of the Amazon Basin, while desert conditions occur along the west coast of Chile-Peru. The high mountains of Central America, the Andes, and the planaltino of Peru-Bolivia support a wide variety of montane vegetation. Temperate rainforest occurs in southern Chile, while steppe conditions prevail over the Patagonian region of southern Argentina and extreme southern Chile.

It is no wonder that, with this diversity of habitats, the Neotropical Realm has the richest fauna of any region. The fresh-water fish fauna is extremely diverse, but represented by rather primitive stocks, characins, gymnotid eels, several endemic catfish families as well as endemic osteoglossids, a lungfish, and synbranchid eels; distinctive are caecilians, Surinam toads, and leptodacylid and microhylid frogs; reptiles are side-neck turtles, caimans, many iguanid and microterid lizards, boas, coral snakes, and pit vipers; birds include rheas, tinamous, New World vultures, screamers, cracids, the hoatzin, seriamas, trumpeters, the limpkin, the sun bittern, seed snipes, potoos, motmots, hummingbirds, puffbirds, jacamars, toucans, manakins, and cotingas; distinctive mammals are phalangers, armadillos, anteaters, sloths, tapirs, peccaries, porcupines, capybara, paca, agouti, marmosets, and New World monkeys.

A similar biogeographical scheme has been established for marine areas; J. C. Briggs has provided an up-to-date summary of the system.

Zoogeographical dynamics. The early workers in zoogeography concentrated their effort on the description of patterns of animal distribution and the clustering of patterns into large units of distribution by major environment or major fauna. These static approaches, while useful in phrasing and structuring data, have proved inadequate for understanding scientific zoogeography since they address only the first of four key elements: (1) recognizing common patterns of distribution; (2) analyzing these patterns to determine common ecological or evolutionary processes that produced the patterns; (3) using the patterns and processes as a prediction of patterns for as yet unstudied groups and (4) for as yet undiscovered geographical and evolutionary events.

The static approach ultimately had a stultifying effect on the development of zoogeography, since by the middle of the twentieth century descriptive and narrative zoogeography had run out of new ideas. Fortunately since the 1960s the science has been revitalized by a resurgence of interest in the dynamics of animal distribution. This interest has led to the development of two major new theories of zoogeographical explanation, one ecological (the dynamic equilibrium theory) and one historical (the vicariance theory). The development of the latter theory has forced a reexamination of the previously dominant historical zoogeographical explanation (the dispersion theory), most effectively expounded earlier in the century by W. D. Matthew and G. G. Simpson.

Dynamic equilibrium theory. This theory of zoogeography was developed during the 1960s primarily by R. H. MacArthur and E. O. Wilson. It originally was aimed at developing predictive mathematical models that would explain the differences in numbers of species on islands of differing sizes and differing distances from the closest mainland source areas. Since almost all habitats on the mainland are patchy in distribution as well (meadows, lakes, mountaintops, and so forth), the theory can also be applied to any discontinuous (insular) segments of the same environment type.

The central axiom of the theory is that the fauna of any disjunct ecological area is a dynamic equilibrium between immigration of new species into the area and extinction of species already present. Species number is thus constant over ecological time, while evolution will act gradually over geological time to increase the equilibrium number of species. From this base it is possible to construct equilibrium models that predict the interactions of distance from the source area, areal extent of the disjunct area, and immigration versus extinction (**Fig. 3**). Data from well-known insular faunas supported the value of these models. Subsequently, controlled defaunization experiments on small islands and study of the immigration process through time confirmed the predictive power of the theory.

Several conclusions may be derived from these models and have been confirmed in the field: (1) Distant disjunct areas will have fewer species than those close to a source area. (2) Small disjunct areas will have fewer species than large ones. (3) Distant disjunct areas take longer to reach equilibria when originally sterile or defaunated than do those close to the source area. (4) The smaller the area and the closer the area source to the area, the higher the turnover rate (change in species composition). The essential insight of equilibrium theory is that local extinctions and immigrations are relatively frequent events. Development of more complex equilibrium models and extension of the approach to other areas of zoogeography are actively under way and promise to produce exciting new views of distribution events.

Dynamic historical biogeography. This field is currently undergoing a major revolution of thought stimulated by new knowledge of global tectonics based on the now generally accepted theory of continental drift as outlined by R. S. Dietz and J. C. Holden. Simply stated, theorists now believe that the major continents were formerly welded together as a supercontinent (Pangaea) that began to rift apart in the Triassic, about 190 million years before present. By the Early Jurassic, northern (Laurasia) and southern (Gondwanaland) landmasses had drifted apart. During the Cretaceous these masses fragmented further, and in the Cenozoic several southern segments became attached to the northern continents (Africa to Eurasia, India to Eurasia, and South America to North America). *See* PLATE TECTONICS.

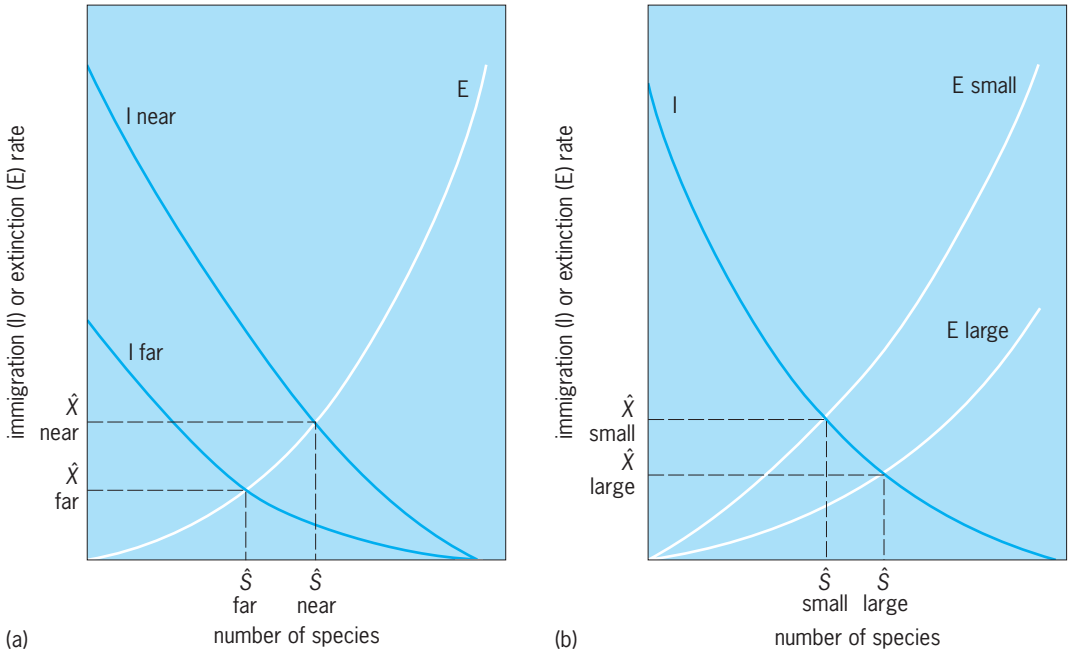

Fig. 3. A disjunct fauna is an equilibrium in ecological time between immigration of new species and extinction of those already present. (*a*) Distance effect; a near island has large equilibrium number of species ($\hat{S}$) and turnover rate ($\hat{X}$). (*b*) Area effect; a large island has larger $\hat{S}$ and smaller $\hat{X}$. (*After D. S. Simberloff, Equilibrium theory of island biogeography and ecology, Annu. Rev. Ecol. Syst., 5:161–182, 1974*)

Previously, historical zoogeography had been dominated by the ideas of Matthew and Simpson, who believed in the permanency of the ocean basins and continents. These authors developed the idea that major groups originated on the northern continents and dispersed southward. P. J. Darlington developed a slightly different point of view and suggested that major groups arose in the Asian tropics and dispersed elsewhere on the continental masses across land bridges or marine barriers by island hopping. According to these kinds of ideas, the present-day distribution of lungfishes (tropical Africa, South America, and Australia) involved origin in the Old World (Asia?) and immigration across land bridges or by island hopping to the southern landmasses.

A group of zoogeographers, L. Croziat, G. Nelson, and D. E. Rosen, much influenced by the role of continental drift in geography, have developed and extended some ideas originally put forward by Croziat into a new theoretical construct. This theory focuses not on the dispersion of organisms from centers of origin, but on the idea that current zoogeographical patterns are the result of the fragmentation of previously continuous tracks by major physiographical change. Thus the present distribution of lungfishes is the result of the fragmentation of a previously continuous track by the breakup of Gondwanaland in the Cretaceous Period.

Essentially, this last development creates a controversial dichotomy in zoogeographical thought. One view emphasizes the active movement (dispersal) of animals as the principal agent responsible for patterns; the other regards dispersion as unimportant and emphasizes the movement and fragmentation of the landmasses and the relative immobility of animals as responsible for patterns. The latter position is called the vicariance theory, in distinction to the dispersion theory. An active, ongoing, vigorous interchange of ideas, with serious reexamination and critique of both schools, is the present theme of zoogeography.

The essential features of the dispersion and vicariance theories are as follows.

Dispersion theory: (1) A monophyletic group arises at a center of origin. (2) Each group disperses from this center. (3) Substantial numbers of monophyletic groups followed the same dispersal route at about the same time to contribute to the composition of a modern fauna. (4) A generalized track corresponds to a dispersion route. (5) Each modern fauna represents an assemblage derived from one to several historical source units. (6) Direction of dispersal may be deduced from tracks, evolutionary relations, and past geodynamic and climatic history. (7) Climate or physiographical change provides the major impetus or opportunity for dispersal. (8) Faunas were shaped by dispersion across barriers and subsequent evolution in isolation. (9) Dispersion is the key to explaining modern patterns: related groups separated by barriers have dispersed across them.

Vicariance theory: (1) Vicariants (allopatric species) arise after barriers separate parts of a formerly continuous population. (2) Substantial numbers of monophyletic groups are simultaneously affected by the same vicariating events (geographical barrier formation). (3) A generalized track estimates the faunal composition and geographical distribution of an ancestral biota before it subdivided (vicariated) into descendant faunas. (4) Each generalized track represents a historical source unit. (5) Sympatry of generalized tracks reflects geographical overlap of

different faunas due to dispersal. (6) The primary vicariating events are changes in world geography (geodynamics) that subdivided ancestral faunas. (7) Faunas evolve in isolation after barriers arise. (8) Vicariance is of primary significance in understanding modern patterns: related groups separated by barriers were fragmented by the appearance of the barriers.

J. Savage has presented a critique of vicariance theory and provides a dispersal-vicariance model that combines the best explanatory features of both approaches. *See* ECOLOGY. Jay M. Savage

Bibliography. H. G. Andrewartha and L. C. Birch, *The Ecological Web: More on the Distribution and Abundance of Animals*, 1986; J. C. Briggs, *Biogeography and Plate Tectonics*, 1987; J. H. Brown and A. C. Gibson, *Biogeography*, 2d ed., 1998; P. J. Darlington, *Zoogeography*, 1980; R. S. Dietz and J. C. Holden, The breakup of Pangaea, *Sci. Amer.*, 223(4):30–41, 1970; J. W. Hedgepeth, Classification of marine environments, *Mem. Geol. Soc. Amer.*, 67(1):17–27, 1957; G. Nelson and N. Platnik, *Systematics and Biogeography: Cladistics and Vicariance*, 1981; G. Nelson and D. E. Rosen, *Vicariance Biogeography*, 1981; J. M. Savage, The enigma of the Central American herpetofauna, *Ann. Missouri Bot. Gard.*, 69(3):444–556, 1983; D. S. Simberloff, Equilibrium theory of island biogeography and ecology, *Annu. Rev. Ecol. Syst.*, 5:161–182, 1974; G. G. Simpson, *The Geography of Evolution*, 1965.

Zoological nomenclature

The system of naming animals that was adopted by zoologists and detailed in the International Code of Zoological Nomenclature. The present system is founded on the 10th edition of C. Linnaeus's *Systema Naturae* (1758) and has evolved through international agreements culminating in the Code adopted in 1985. The primary objective of the Code is to promote the stability of the names of taxa (groups of organisms) by providing rules concerning name usage and the activity of naming new taxa. The rules are binding for taxa ranked at certain levels and nonbinding on taxa ranked at other levels. *See* ANIMAL SYSTEMATICS.

Zoological nomenclature is built around four basic features. (1) The correct names of certain taxa are either unique or unique combinations. (2) These names are formed and treated as Latin names and are universally applicable, regardless of the native language of the zoologist. (3) The Code for animals is separate and independent from similar codes for plants and bacteria. (4) No provisions of the Code are meant to restrict the intellectual freedom of individual scientists to pursue their own research.

There are four common reasons why nomenclature may change. (1) New species are found that were once considered parts of other species. (2) Taxonomic revisions may uncover older names or mistakes in identification of types. (3) Taxa may be combined, creating homonyms that require replacement. (4) Concepts of the relationships of animals change. Stability is subservient to progress in understanding animal diversity.

Code provisions. The 1985 Code applies to both living and extinct animals. It contains a series of interrelated articles governing the formation and use of names that form a very complex document. It specifies the conditions for determining whether a name is validly published, the publication date, and the authorship of the name. Validly published names are available for use. Of the several available names, the Code helps the taxonomist determine which is correct for a particular taxon: the Code uses the principle of priority that the oldest available name is the correct one unless unusual circumstances obtain. Coupled with the principle of priority is the concept of the first reviser, who is the first author to have treated a particular taxon after it was described. The decisions of the first reviser regarding names are usually followed except in unusual circumstances.

An International Commission on Zoological Nomenclature is empowered by the General Assembly of the International Union of Biological Sciences to suppress names that would normally be available if this is in the best interest of stability. The Commission also interprets the various articles and publishes opinions that are binding on all matters relating to the Code. The Code includes a number of appendices, such as a code of ethics and guides for transliteration and latinization of names.

The articles in the Code are directed toward the names of taxa at three levels. The family group includes taxa ranked as at the family and tribe levels (including super- and subfamilies). The genus group includes taxa below subtribe and above species. The species group includes taxa ranked as species or subspecies. Taxa above the family group level are not specifically treated, and their formation and use are not strictly regulated. For each group, provisions are made that are either binding or recommended.

Binominal nomenclature. The basis for naming animals is binominal nomenclature, that is, a system of two-part names. The first name of each species is formed from the generic name, and the second is a trivial name, or species epithet. The two names agree in gender unless the specific epithet is a patronym (named for a person). The combination must be unique; no other animal can have the same binominal. The formal name of a species also includes the author, so the formal name for humans is *Homo sapiens* Linnaeus. The genus, as a higher taxon, may have one or many species, each with a different epithet. *Homo* includes *H. sapiens, H. erectus, H. habilis*, and so forth. One feature of the Linnaean binominal system is that species epithets can be used over and over again, so long as they are used in different genera. *Tyrannosaurus rex* is a large dinosaur, and *Percina rex* is a small fresh-water fish. The epithet *rex* is not the species name of *Percina rex* because all species names are binominal in form. It is recommended that names of genera and species be set in a different typeface from normal text; italics is conventional. Different names

for the same species are termed synonyms, and the senior synonym is usually correct (principle of priority). If two species have the same name, they are considered homonyms, and the older name is correct. Homonyms are usually produced when two or more genera are combined, but homonymy is a complex subject and those interested are encouraged to consult the Code. Modern species descriptions are accompanied by a description that attempts to show how the species is different from others, and the designation of one or more type specimens. The type system is frequently viewed as pedantic "stamp collecting" by other biologists, but it is actually one of the foundations of nomenclature. The existence of actual specimens that can be examined by future generations provides physical evidence of what the author actually meant when describing the species.

Higher taxa. All higher taxonomic names have one part (uninominal) and are plural. Names of taxa of the family and genus groups must be unique. The names of genera are in Latin or latinized, are displayed in italics, and may be used alone. The names of the family group are formed by a root and an ending specific to a particular hierarchical level: family Hominidae (root + idae), subfamily Homininae (root + inae). The endings of superfamilies (root + oidea) and tribes (root + ini) are recommended but not mandated. The endings of taxa higher than the family group vary. For example, orders of fishes are formed by adding -iformes to the root (Salmoniformes), while in insects the ending is usually -ptera (Coleoptera). The principles of types, synonymy, and homonymy apply to the generic and family names, but not to higher taxa. *See* CLASSIFICATION, BIOLOGICAL; TAXONOMIC CATEGORIES. E. O. Wiley III

Bibliography. International Trust for Zoological Nomenclature, *International Code of Zoological Nomenclature adopted by the XX General Assembly of the International Union of Biological Sciences*, 1985; E. Mayr and P. D. Ashlock, *Principles of Systematic Zoology*, 2d ed., 1991; E. O. Wiley, *Phylogenetics, The Theory and Practice of Phyolgenetic Systematics*, 1981.

Zoology

The science that deals with knowledge of animal life. Together with botany, the science of plants, it forms biology, the science of living things. With the great growth of information about animals, zoology has been much subdivided. Some major fields are anatomy, which deals with gross and microscopic structure; physiology, with living processes in animals; embryology, with development of new individuals; genetics, with heredity and variation; parasitology, with animals living in or on others; natural history, with life and behavior in nature; ecology, with the relation of animals to their environments; evolution, with the origin and differentiation of animal life; and taxonomy, with the classification of animals. *See* ANATOMY, REGIONAL; DEVELOPMENTAL BIOLOGY; GENETICS; PARASITOLOGY; PHYLOGENY; PLANT EVOLUTION; TAXONOMY. Tracy I. Storer

Bibliography. C. E. Bond, *Biology of Fishes*, 2d ed., Saunders College Publishers, 1996; J. L. Gould and W. T. Keeton, *Biological Science*, 6th ed., W. W. Norton & Co., 1996; M. Hildebrand and G.E. Gaslow, Jr., *Analysis of Vertebrate Structure*, 5th ed., Wiley, 2001; K. V. Kardong, *Vertebrates: Comparative Anatomy, Function, Evolution*, 2d ed., McGraw-Hill, 1998; D. W. Linzey, *Vertebrate Biology*, McGraw-Hill, 2001; L. Sherwood, H. Klandorf, and P. H. Yancy, *Animal Physiology: From Genes to Organisms*, Thomson, Brooks/Cole, 2005; C. Starr and R. Taggart, *Biology: The Unity and Diversity of Life*, 10th ed., Thomson, Brooks/Cole, 2004; G. J. Tortora and B. Derrickson, *Principles of Anatomy and Physiology*, 11th ed., Wiley, 2006; T. A. Vaughan, J. M. Ryan, and N. J. Czaplewski, *Mammalogy*, 4th ed., Thomson, Brooks/Cole, 2000.

Zoomastigophorea

A class of protozoa of the subphylum Sarcomastigophora. Zoomastigophorea, also known as Zoomastigina, are flagellates which have few or no characters relating them to the pigmented forms. Some are simple, some are specialized; some have pseudopodia besides flagella, others have no pseudopodia. One group engulfs solid food at any body point; another shows localized ingestive areas. All are colorless. None produce starch or paramylum, and lipids and glycogen are assimilation products. Cells are naked or have delicate membranes. Colony formation is common. Colonies differ greatly in form and may be amorphous, linear, spherical, arboroid, or plane. Flagella vary from none to many. Nuclei generally have endosomes; frequently a rhizostyle extends from the nucleus to the flagellum base. Nutrition is holozoic, saprozoic, or parasitic, and encystment is common. Ecologically Zoomastigophorea are adapted to waters of relatively high organic content, but they occur in the clearest fresh waters and in the high seas. Nine orders are included in this class: the Choanoflagellida, Bicosoecida, Rhizomastigida, Kinetoplastida, Retortomonadida, Diplomonadida, Oxymonadida, Trichomonadida, and Hypermastigida. See articles on these orders. *See* PROTOZOA; SARCOMASTIGOPHORA. James B. Lackey

Zoonoses

Diseases that are transmitted from animals to humans. Zoonotic agents are found among all major categories of microbes, including bacteria, fungi, parasites, and viruses (see **table**). Although anyone can acquire a zoonotic infection, young children, older adults, and immunocompromised individuals are most at risk. Currently, there are large numbers of people that are considered immunocompromised

Some significant zoonoses*

Disease by category	Etiologic agent
BACTERIAL AND RICKETTSIAL	
Anthrax	*Bacillus anthracis*
Bordetella bronchiseptica pneumonia	*Bordetella bronchiseptica*
Brucellosis	*Brucella* sp.
Capnocytophaga	*Capnocytophaga canimorsus*
Cat-scratch, bacillary angiomatosis	*Bartonella henselae, Afipia felis*
Chlamydiosis/psittacosis	*Chlamydia psittaci*
Ehrlichiosis (monocytic and granulocytic)	*Ehrlichia chaffeensis, E. phagocytophilia*
Enteric disease	*Campylobacter jejuni, Clostridium perfringens, Escherichia coli, Salmonella* sp., *Vibrio parahemolyticus, Yersinia enterocolitica*
Leptospirosis	*Leptospira* sp.
Listeriosis	*Listeria monocytogenes*
Lyme disease	*Borrelia burgdorferi*
Murine typhus	*Rickettsia typhi*
Pasteurellosis	*Pasteurella multocida*
Plague	*Yersinia pestis*
Q fever	*Coxiella burnetti*
Rat bite fever	*Streptobacillus moniliformis*
Rickettsialpox	*Rickettsia akari*
Rhodococcosis	*Rhodococcus equi*
Scrub typhus	*Orienta tsutsugamushi*
Spotted fever, Rocky Mountain	*Rickettsia rickettsii*
Tuberculosis	*Mycobacterium bovis*
Tularemia	*Francisella tularensis*
MYCOTIC	
Cryptococcosis	*Cryptococcus neoformans*
Dermatomycosis or ringworm	*Microsporum canis, Trichophyton mentagrophytes*
PARASITIC	
African trypanosomiasis, sleeping sickness	*Trypanosoma brucei rhodesiense*
American trypanosomiasis or Chagas' disease	*Trypanosoma cruzi*
Anisakiasis	*Anisakis* sp., *Psuedoterranova* sp., *Contracaecum* sp.
Balantidiasis	*Balantidium coli*
Clonorchiasis	*Clonorchis sinensis*
Coccidiosis	*Cryptosporidum parvum, Sarcocystis* sp.
Cutaneous larva migrans	*Ancylostoma canium, A. braziliense*
Diphyllobothriasis or fish tapeworm	*Diphyllobothrium latum*
Dipylidiasis or dog tapeworm	*Dipylidium caninum*
Dirofilariasis	*Dirofilaria immitus, D. repens*
Giardiasis	*Giardia lamblia*
Hydatid disease or echinococcosis	*Echinococcus* sp.
Paragonimiasis	*Paragonimus westermani*
Leishmaniasis	*Leishmania donovani, L. tropica,* and other species
Piroplasmosis, babesiosis	*Babesia microti* and other species
Taeniasis or cysticercosis	*Taenia saginatum, T. solium*
Toxoplasmosis	*Toxoplasma gondii*
Trichinosis	*Trichinella* sp.
Visceral larva migrans	*Toxocara canis, T. cati, Baylisasearis procyonis*
VIRAL	
Ebola and Marburg hemorrhagic fevers	RNA Filovirus
Encephalitis (La Cross, California)	RNA Bunyavirus
Equine encephalitis (Eastern, Western, Venezuelan)	RNA Togavirus
Hantavirus pulmonary syndrome	RNA Hantavirus
Hemorrhagic fevers: Argentine, Bolivian, Brazilian, Venezuelan	RNA Arenavirus
Hemorrhagic fever: Crimean-Congo	RNA Nairovirus
Hemorrhagic fever: renal	RNA Hantavirus
Hendra or Nipah diseases	RNA Paramyxovirus
Hepatitis	RNA Hepatitis A
Herpes B	DNA Herpes simiae
Lassa fever	RNA Arenavirus
Lymphocytic choriomeningitis	RNA Arenavirus
Monkey pox, cowpox	DNA Orthopoxviruses
Orf	DNA Parapoxvirus
Rabies	RNA Lyssavirus
Rift Valley fever	RNA Bunyavirus
Vesicular stomatitis	RNA Vesiculovirus
West Nile fever	RNA Flavivirus

*Based on data from William T. Hubbert, Ann Marie Nelson, and Brett Saladino.

(for example, people receiving chemotherapy or those infected with agents such as HIV).

Animals can also serve as reservoirs for zoonotic agents; that is, they are a source of the microorganism that allows transmission to the human population. Domesticated animals can often serve as a bridge for zoonotic agents that are commonly found in feral animal populations. However, sometimes animals serve only as indicators that an organism is present in the environment. In these cases, they become infected from contaminated soil, water, or air, as do humans. Additionally, there are diseases, called anthroponoses, that are transmitted from humans to animals.

Nervous system infections. Important zoonotic agents that can invade the nervous system in humans include *Toxoplasma gondii* and *Cryptococcus neoformans.*

Toxoplasma. Thirty to forty percent of adult humans in the United States are seropositive for *T. gondii*. Feline species are the only definitive hosts for the organism. Cats harbor sexual stages of the parasite in their gastrointestinal tract and shed infectious oocysts in their stool. The organism has been found in all areas of the world, in over 200 species of birds, and in many species of warm-blooded animals, which act as intermediate hosts. Intermediate hosts become infected by ingesting sporulated oocysts from cats. Humans can contract the disease by ingesting inadequately cooked meat of the intermediate hosts with infectious bradyzoites and tachyzoites, or by ingesting oocysts from soil contaminated with cat feces. In one European study, more than 50% of the meat of pigs and sheep carried toxoplasmosis. Humans can be infected by ingesting any source of meat with *Toxoplasma*, but lamb, mutton, and pork are most commonly infected. *See* TOXOPLASMIDA.

Cryptococcus. Cryptococcus neoformans is an important pathogen in immunosuppressed patients and occurs commonly as a premonitory sign of acquired immune deficiency syndrome (AIDS), where it causes severe life-threatening meningitis. *Cryptococcus* is taken up by the human host via the respiratory tract. The organism has been found in soil, plants, bird feces, raw milk, fruit juices, and from the oral cavity, gastrointestinal tract, and skin of healthy humans. It is commonly found in pigeon roosts and in soil contaminated by pigeon feces, and has been isolated from droppings of other bird species. *Cryptococcus neoformans* has an ecological relationship with pigeons in that it selectively assimilates creatinine in pigeon urine. *See* ACQUIRED IMMUNE DEFICIENCY SYNDROME (AIDS).

Respiratory tract infections. *Bordetella bronchiseptica* is a secondary bacterial invader that occurs commonly in the respiratory tracts of dogs, cats, and pigs, and less commonly in horses. It complicates the management of viral respiratory disease in the dog, and has been associated with clinical signs of fever, anorexia, coughing, and nasal discharge. Immunocompromised patients should avoid exposing themselves and their dogs to environments with sources of *B. bronchiseptica* such as dog shows, kennels, or other settings where dogs are housed closely together.

Gastrointestinal tract infections. Agents causing primary gastrointestinal disease that have been documented to be transmitted to humans from an animal source include *Salmonella, Campylobacter, Mycobacteria, Cryptosporidia*, and *Giardia*. Agents that have been implicated as zoonotic agents include *Entamoeba, Isospora, and Microsporidia*.

Salmonella. Salmonella is an important pathogen in humans. Salmonellosis is a major communicable disease problem in humans in the United States, with the infection rate estimated as high as 2 million cases per year. Poultry and beef products are the largest sources of nontyphoid *Salmonella* in the United States. Contaminated animals and animal products are a major source of *Salmonella* for humans. The most common source is from contaminated foods. Strict hygiene in food preparation is needed to avoid contaminating cold-served food products (such as salads and coleslaw) with raw meat. It is estimated that the rate of infection of *Salmonella* in domestic animals is 1–3%. Transmission of antibiotic-resistant *S. typhimurium* to humans has been reported due to contact with sick calves shedding the organism.

Salmonellosis is frequently diagnosed in dogs and cats, and these animals can also be a source of infection for humans. Confining dogs and cats in the house as much as possible minimizes their contact with fecal material of other animals. Avoidance of feeding pets raw meat prevents their contracting the disease from a well-known source of salmonellosis. Animals with diarrhea should be presented promptly to a veterinarian for diagnosis. Treatment of pets with salmonellosis is usually not recommended unless clinical signs of bacteremia are evident, because treatment may prolong the period of shedding of the organism. Pets should not be returned to the household of immunosuppressed individuals until two fecal cultures taken 24 hours apart are negative for *Salmonella*. It is important that these fecal cultures be taken after any antibiotic treatment has ended. *Salmonella* is also carried commonly by cold-blooded animals. *See* SALMONELLOSES.

Campylobacter. Campylobacter is another common gastrointestinal and blood infection that can be transmitted from animals to humans. *Campylobacter* is a significant cause of enteritis in humans all over the world; it has been targeted as one of four bacterial pathogens for elimination from human food supplies by United States agencies. *Campylobacter jejuni* is the most important species, although other species of *Campylobacter* can cause diarrhea in humans. Transmission of *C. jejuni* to humans occurs commonly from contaminated food and water. Meat products, especially poultry, are frequently contaminated with this bacteria. *Campylobacter jejuni* is found generally in water sources, where it can live for several weeks at low temperature. Many animals, including cattle, sheep, swine, fowl, dogs, and cats, excrete *Campylobacter* spp. Household pets, especially kittens and puppies, frequently shed these organisms, with an incidence of isolation of

Campylobacter in feces as high as 35–42%. To prevent infection with *Campylobacter*, people should cook meat products thoroughly and wash cooking utensils well. Immunocompromised individuals should be aware of the risks of owning or coming in contact with puppies and kittens and their excreta. Animals with diarrhea should be presented promptly to their veterinarian, and a fecal sample should be taken for testing for *Campylobacter*. If *Campylobacter* spp. are isolated from dogs or cats, treatment with erythromycin is recommended to eliminate shedding of the organism in feces.

Cryptosporidium. *Cryptosporidium* is a protozoan parasite that is only 1–6 micrometers in size. It lacks host specificity and infects the gastrointestinal tract of a wide range of mammals.

Cryptosporidium parvum is the most important species in humans and domesticated animals. It has been shown to occur naturally in mice, rabbits, guinea pigs, cats, dogs, squirrels, raccoons, horses, pigs, sheep, goats, cattle, and several species of nondomesticated ruminants. Isolates from humans have proven infective to cats, cattle, dogs, goats, mice, pigs, rats, sheep, and other humans. Transmission of *C. parvum* to humans from cats, cattle, and pigs has also been reported. *Cryptosporidium parvum* is a common cause of diarrhea in young calves who acquire the organism from their mothers and the environment; it is commonly shed in the feces of calves with no clinical signs of disease. Fecal contamination of food and water contribute to the high prevalence of cryptosporidial infection in persons residing in areas with poor sanitary conditions. Contaminated food, fomites, flies, and water are sources of infection with *Cryptosporidium*. Oocysts have proven resistant to most chemical disinfection at approved concentrations, and disinfection of drinking water by common methods is not effective.

Giardia. *Giardia* is a flagellate protozoan parasite that infects the gastrointestinal tract of many species of animals including dogs, cats, horses, pigs, cattle, sheep, goats, and many other mammals. It naturally infects humans, and is endemic in many developing countries where the prevalence in children is as high as 30%. Pets are potential sources of *Giardia* for humans. Studies in New Jersey have shown infection rates of 35% and 2.5% in stray dogs and stray cats, respectively. Other studies found high infection rates in cattle, beavers, and coyotes. Treatment for giardiasis is difficult in companion animals. Infected animals should be isolated from other animals in the household, and treated and bathed to remove cysts from the fur. Strict personal and food hygiene minimizes the risk of *Giardia* infection from human sources. *See* GIARDIASIS.

Parasitic diseases of animals. Some parasitic diseases of animals have zoonotic potential. The dog heartworm, *Dirofilaria immitis*, is endemic in many areas of North America and the rest of the world, requiring daily chemical prophylaxis in pet dogs during the mosquito season. Humans can be accidentally infected by *D. immitis* microfilaria-carrying mosquito vectors due to the high prevalence of *D. immitis*, although the disease is rare. The disease in humans is often diagnosed when a solitary lung lesion or coin lesion is noted on a chest radiograph and confirmed by biopsy.

Toxocara. *Toxocara canis* and *T. cati* are round worms of the dog and cat. Humans acquire infection with these parasites by ingesting embryonated eggs shed in the feces of animals. Visceral larva migrans occurs when eggs hatch, penetrate the intestine, and migrate through body tissue. Migration can result in multiple abscesses, eosinophilic granulomas, hepatomegaly, and pneumonitis with eosinophilia. Symptoms include coughing, nausea, vomiting, and dyspnea (difficulty in breathing). Larva migrans can also affect the eyes. Ocular larva migrans causes inflammation, fibrosis, and loss of vision. The disease most often occurs in children less than 3 years of age, but can also occur in adults. The infant and toddler age group is most predisposed due to the tendency to ingest soil. Eggs of *Toxocara* are extremely resistant in soil and remain infectious for years. Routine yearly fecal exams and treatment of positive animals with anthelminthics is recommended to decrease the potential for this problem.

Ancylostoma caninum and *A. brasiliense* are hookworms of dogs and cats which rarely cause cutaneous larva migrans in humans. Hookworm eggs are shed in feces into the environment, hatch, and undergo two larval molts to produce infectious, third-stage larvae. Hookworms infect their definitive hosts by penetrating the skin. In the correct host, larvae migrate from the skin to the lungs, ascend the respiratory tree, and are swallowed to resume development in the intestine where they mature into adult hook worms. Accidental infection of animal hookworms in humans usually occurs through larval penetration of the extremity when walking barefoot through infected soil. After skin penetration, the larvae migrate in the germinal layer of the epidermis, causing sinuous tunnels with associated itching and inflammatory response. The larvae can survive in the skin for several weeks or months. Although some human infections may be associated with pneumonitis and eosinophilia when the larvae invade the lung or other organ tissue, most cases are limited to the skin.

Echinococcus. *Echinococcus granulosus* and *E. multilocularis* are tapeworms that cause hydatid disease or echinococcosis. *Echinococcus granulosus* is a tapeworm whose life cycle involves dogs, coyotes, wolves, and dingos as the primary hosts and sheep, swine, cattle, moose, caribou, and humans as intermediate hosts. This parasite is endemic in North and South America, England, Africa, the Middle East, Australia, and New Zealand. In the United States, it occurs mainly in the western states. *Echinococcus multilocularis* is a tapeworm that cycles between dog, cat, and fox as definitive hosts, and between vole, lemming, cattle, horse, swine, and humans as intermediate hosts. *Echinococcus multilocularis* is endemic in north-central Europe, Alaska, Canada, and the central United States far south as Illinois and Nebraska. Dogs and cats that ingest wild game and rodents can be infected with

E. multilocularis and can be a risk to humans. Humans act as intermediate hosts for both of these parasites. Humans ingest *Echinococcus* eggs from feces of the definitive hosts, which hatch in the intestines releasing a larval tapeworm that migrates and develops in tissues. Hydatid cysts are formed which are found in the liver, lung, nervous system, bone, spleen, kidney, and heart. Symptoms range from inapparent to severe and relate to the site of the cyst and whether it ruptures. Prevention of infection involves avoidance of contact with fecal matter from pets and wild carnivores, good personal and food hygiene, keeping pets indoors, and routine fecal exams of pets for parasites.

Fleas. Fleas are common, difficult to eliminate, and capable of transmitting several important zoonotic diseases. However, in the United States fleas are mostly a nuisance problem causing pruritis and hypersensitivity in animals and humans. Control of fleas must focus on the environment of the pet, such as the house and the yard. Pets may be treated for fleas weekly, but treatment must coincide with treatment of the environment.

Dipylidium caninum is a common tapeworm of dogs and cats. Cats and dogs are infected during grooming by ingesting the flea intermediate host containing an infective cysticercoid. Cysticercoids mature into adult tapeworms in the intestine of their normal host. Tapeworm eggs that are shed in the feces of an infected dog or cat are eaten by flea larvae, where they hatch and develop to the cysticercoid in the body cavity of the flea. Humans, especially children, are accidentally infected when they swallow a flea while crawling on the floor or while playing with pets.

Ticks. Animals are exposed to ticks when they roam fields and wooded lots. Ticks transmit many zoonotic diseases, including babesiosis, tick fever, rickettsial diseases, Lyme disease, Q-fever, relapsing fever, typhus, ehrlichiosis, and tularemia. Humans are exposed to ticks by similar means, but in some cases can acquire ticks from animals. Pets kept indoors are less likely to acquire ticks. Humans can use tick repellants, check themselves for ticks after a walk in the woods, and avoid tick-infested areas. *See* LYME DISEASE; RICKETTSIOSES; TULAREMIA.

Systemic disease. *Bartonella* spp., including *B. henselae* (formerly *Rochalimaea*), a member of the rickettsia family, have been implicated as causative agents in cat scratch disease and bacillary angiomatosis (vascular tumors of the skin, internal organs, bone marrow, and lymph nodes). In humans, cat scratch disease is characterized by persistent regional lymphadenopathy in the lymphatic drainage area of the site of a recent cat bite or cat scratch. In immunocompetent hosts, catch scratch disease tends to be a disease of children, usually self-limited and localized to regional lymph nodes. It rarely causes systemic or widespread infections. In HIV-infected hosts, the disease usually affects young adults and manifests itself as bacillary angiomatosis. Manifestation of bacillary angiomatosis includes cutaneous nodules resembling Kaposi sarcoma, encephalopathy (disease of the brain), lymphadenopathy, pneumonitis (inflammation of lung tissue), bone lesions, hepatosplenomegaly (lenlargement of the liver and the spleen), and hypotension (low blood pressure) with metabolic acidosis. The disease in HIV-infected individuals is widespread, but responds to antibiotic therapy. *Bartonella henselae* is widespread in the cat population, with isolation from the blood of 41–46% of cats in studies involving household, stray, and pound cats. Since arthropods are known vectors for the transmission of rickettsias, prevention should be focused on flea control and minimizing direct transmission by cats. Specific measures include declawing of pet cats, avoiding bites or scratches from cats, and strict flea control. *See* BARTONELLOSIS; CAT SCRATCH DISEASE.

Pasteurella multocida is an important pathogen in animal bites and a documented life-threatening pathogen in immunocompromised patients. *Pasteurella multocida* has been found in the oropharynx of 60–90% of healthy cats and in 55% of dogs. Transmission occurs via traumatic wounds such as bites and scratches by the respiratory route via aerosol, and by ingestion. Dissemination and septicemia (persistence and multiplication of living bacteria in the blood) have been noted in patients on chemotherapy, with cirrhosis, or with hematologic malignancy. Two cases of *P. multocida* have been documented in an HIV-infected individual. One patient undergoing peritoneal dialysis developed peritonitis, and another patient developed sinusitis and pneumonia apparently from aerosol transmission by a cat. To minimize transmission of *P. multocida*, one should avoid being bitten or scratched by pet dogs and cats. *See* PASTEURELLA.

There are many sources of information on zoonotic diseases, including nonprofit organizations and the United States Centers for Disease Control. Overall, the risk of contracting zoonotic diseases by susceptible individuals is likely to be small, but physicians and veterinarians can decrease this risk by educating patients on the transmission and prevention of these diseases. *See* ANIMAL VIRUS; MEDICAL BACTERIOLOGY; MEDICAL PARASITOLOGY.

Linda S. Mansfield

Bibliography. F. J. Angulo et al., Caring for pets of immunocompromised persons, *J. Amer. Vet. Med. Ass.*, 205(12):1711–1718, 1994; E. U. Canning, Cryptosporidia, in J. P. Kreier (ed.), *Parasitic Protozoa*, Academic Press, New York, 1993; Centers for Disease Control and Prevention, *USPHS/IDSA Guidelines for the Prevention of Opportunistic Infections in Persons Infected with Humans Immunodeficiency Virus: A Summary*, MMWR 44(No. RR–8):1–34, 1995; J. P. Dubey, Toxoplasmosis, *J. Amer. Vet. Med. Ass.*, 205(11):1593–1598, 1994; I. Nachamkin, M. J. Blaser, and L. S. Tompkins, *Campylobacter jejuni: Current Status and Future Trends*, American Society for Microbiology, Washington, DC, 1992; C. R. Sterling and M. J. Arrowood, Cryptosporidia, in J. P. Kreier (ed.), *Parasitic Protozoa*, Academic Press,

New York, 1993; P. G. Wall et al., Transmission of multi-resistant strains of *Salmonella typhimurium* from cattle to man, *Vet. Rec.*, 136:591–592, 1995.

Zooplankton

Animals that inhabit the water column of oceans and lakes and lack the means to counteract transport currents. Zooplankton inhabit all layers of these water bodies to the greatest depths sampled, and constitute a major link between primary production and higher trophic levels in aquatic ecosystems. Many zooplankton are capable of strong swimming movements and may migrate vertically from tens to hundreds of meters; others have limited mobility and depend more on water turbulence to stay afloat. All zooplankton, however, lack the ability to maintain their position against the movement of large water masses.

Zooplankton can be divided into various operational categories. One means of classification is based on developmental stages and divides animals into meroplankton and holoplankton. Meroplanktonic forms spend only part of their life cycles as plankton and include larvae of benthic worms, mollusks, crustaceans, echinoderms, coral, and even insects, as well as the eggs and larvae of many fishes. Holoplankton spend essentially their whole existence in the water column. Examples are chaetognaths, pteropods, larvaceans, siphonophores, and many copepods.

Size is another basis of grouping all plankton. A commonly accepted size classification scheme includes the groupings: picoplankton (<2 micrometers), nanoplankton (2–20 μm), microplankton (20–200 μm), mesoplankton (0.2–20 mm), macroplankton (20–200 mm), and megaplankton (>200 mm).

Systematic composition. Nearly every major taxonomic group of animals has either meroplanktonic or holoplanktonic members. Some of the more common zooplankton groups (see **illustration**) are described below.

Protists are single-celled eukaryotes that are vital to planktonic systems. Photosynthetic protists form the base of planktonic food webs as the major primary producers in the plankton. Animallike protists (protozoa) constitute the largest proportion of the nano- and microzooplankton and include amebas, nonphotosynthetic flagellates, and ciliates. Some of the ameboid Foraminifera and Actinopoda with tests (rigid skeletal supports) or hard skeletons are larger and reach sizes exceeding 2 mm. The skeletons of foraminiferans and actinopods form an important part of the deep-sea sediments, and foraminiferans are used as markers in oil exploration. Nanoplanktonic flagellates are usually the most important predators of bacteria; however, some dinoflagellates can capture and digest larger organisms using protoplasmic nets. Both photosynthetic and nonphotosynthetic dinoflagellates have been implicated in noxious red tides or fish kills. Various ciliates species capture bacteria, phytoplankton, and other protists. Many protists that do not fall neatly into a classification as animallike (such as protozoa) or plantlike organisms are termed mixotrophs. Mixotrophs bridge the plant/animal division because they have both photosynthetic and phagotrophic capabilities. Some flagellates with chloroplasts ingest bacteria, phytoplankton, or other protists. Conversely, some ciliates retain functional chloroplasts from phytoplankton that are ingested and otherwise digested. *See* FORAMINIFERIDA; PROTOZOA; RADIOLARIA.

The phylum Cnidaria contains a number of groups which are important in the marine plankton, including scyphozoans, the true jellyfish. Both scyphozoans and the colonial siphonophores are carnivores with tentacles bearing stinging nematocyst cells. They are found at all depths in the ocean but are most common in the upper waters. Cnidaria are rare in fresh water.

The ctenophores, or comb jellies, were once grouped with Cnidaria but are now treated separately because they lack nematocysts. All are exclusively carnivorous and are important predators of many zooplankton, especially copepods. *See* COELENTERATA.

Rotifers probably arose in fresh water; only about 5% of the approximately 2500 described species are found in marine and brackish waters. Although about a hundred species are holoplanktonic, most are primarily sessile. Nevertheless, they are often a numerically important fraction of the zooplankton. Most rotifers feed on bacteria, detritus, and algae, although some are predators of protozoa and other rotifers. Under favorable conditions, rotifers exhibit parthenogenesis (egg production without fertilization leading to female-only populations). Sexual reproduction does occur, but only after male eggs are produced as a response to environmental stresses. Cyclomorphosis (seasonal changes in morphology within a species) is common among rotifers. *See* ROTIFERA.

Chaetognatha is a carnivorous marine group with worldwide distribution. They are mostly holoplanktonic and are found at all depths, although a given species may be restricted to certain water masses and depths. The elongate body is transparent with lateral fins and is usually less than 3 cm in length. *See* CHAETOGNATHA.

Veliger larvae of many benthic mollusks are frequently seen in coastal plankton. There are also marine gastropods, including heteropods and pteropods, which are adapted to a holoplanktonic life style. Another group represented in shallow marine waters by larvae is Polychaeta. This annelid class also contains a few holoplanktonic families. *See* MOLLUSCA; POLYCHAETA.

Copepods are almost always the most numerous members of the macrozooplankton community in marine systems, and are important in freshwater as well. They often migrate to deeper water in daytime and move toward the surface at night, using the first antennae and the thoracic appendages to

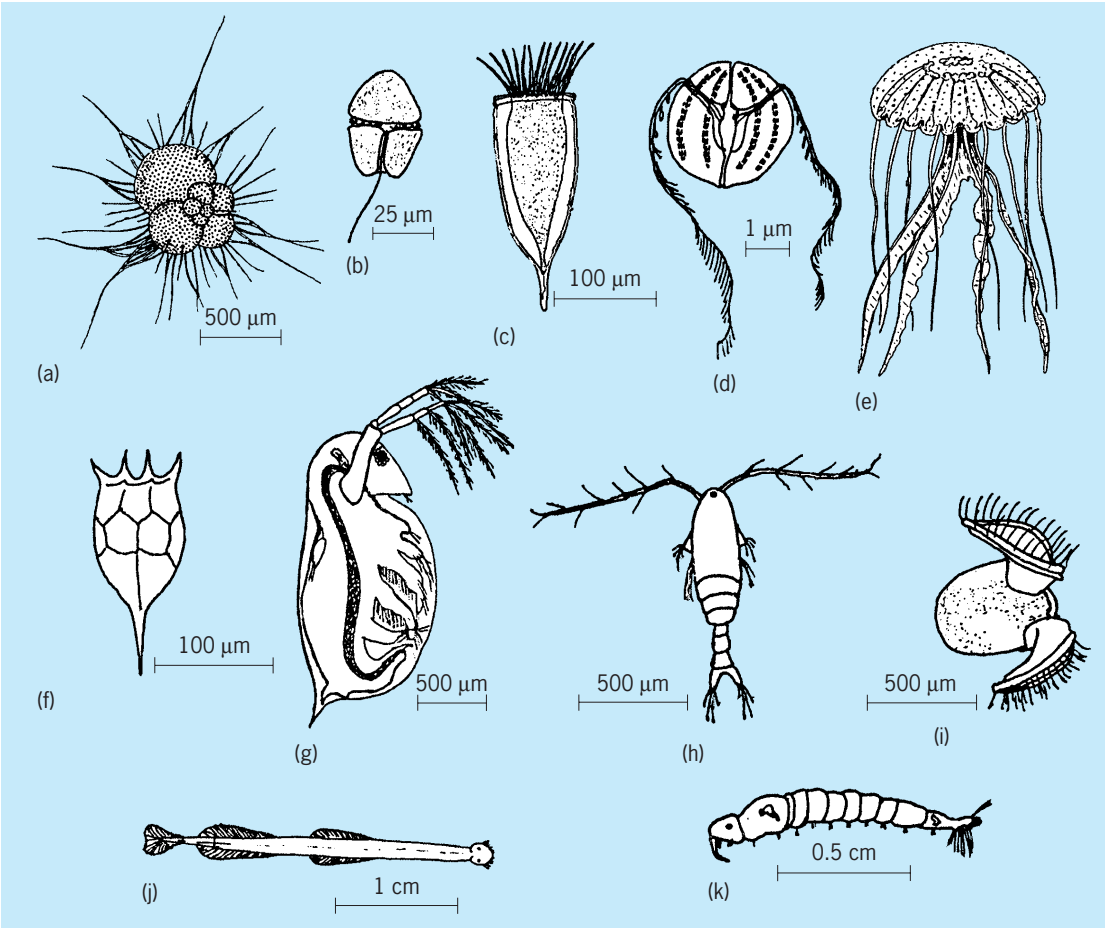

Representative zooplankton. (*a*) Foraminifera. (*b*) Dinoflagellate. (*c*) Tintinnid ciliate. (*d*) Ctenophore. (*e*) Cnidarian (scyphozoan jellyfish). (*f*) Rotifer. (*g*) Cladoceran. (*h*) Copepod. (*i*) Gastropod veliger larva. (*j*) Chaetognath. (*k*) Insect larva (*Chaoborus*). (*Parts a and e from L. H. Hyman, The Invertebrates, vol. 3: Protozoa through Ctenophora, McGraw-Hill, 1940*)

swim. Some are primarily suspension feeders and use appendages with hairlike setules to remove phytoplankton and detritus from the water. Others are primarily raptorial and feed on larger particles, including smaller zooplankton. *See* COPEPODA.

Cladocera are a second major group of planktonic crustaceans. They often outnumber copepods in fresh water and occasionally are abundant in coastal marine waters. Most have bodies covered by a folded carapace that gives an appearance of being bivalved. The enlarged second antennae are the primary swimming appendages. For most Cladocera the primary food source is phytoplankton, which are filtered from water passing through appendages within the carapace. A few species are predaceous on other zooplankton. Like copepods, cladocerans are well known for diel (diurnal) vertical migrations. Their occurrence tends to be seasonal, with a resting egg formed between periods of rapid parthenogenic reproduction. Cyclomorphosis is common.

Euphausids or krill, are relatively large marine crustaccans (up to 3cm) that are found worldwide at all depths. Individual species have more limited distribution. Some species, especially coldwater forms, demonstrate swarming behavior, and in high latitudes they are the major food source of baleen whales. The most important food of polar euphausids appears to be diatoms, while carnivorous species are more common in warmer waters. *See* EUPHAUSIACEA.

Most aquatic insect species tend to be associated with the bottoms, shores, or surface layers of streams and lakes. The only major insect member of the plankton is the larvae of the phantom midge *Chaoborus*. This larvae is able to resist anoxia (lack of oxygen) and is often found at the bottom during day and in the plankton at night. *Chaoborus* species are important predators of other zooplankton in lakes.

Tunicates are all marine. Many are benthic and produce planktonic larvae, but two classes, Appendicularia and Thaliasia, are holoplanktonic. Appendicularia are neotenic and build a mucous house which acts as a filter to capture small food particles. The Thaliacea, including salps and doliolids, are also filter feeders. *See* TUNICATA.

Adaptations. A problem faced by all plankton is to maintain position in the water column. Flattened bodies and numerous lateral spines or plumose setae which increase surface-to-volume ratios are common in various zooplankters. This increases resistance to the passage of water and thus slows

sinking. Other adaptations to a floating existence are positive buoyancy mechanisms, such as oil droplets, gas-filled floats, or regulation of ionic balance by replacement of heavy ions with lighter ones. Some zooplankton have gelatinous sheaths that appear to slow the sinking rate. For example, the gelatinous sheath produced by the cladoceran *Holopedium gibberum* reduces its sinking rate by about 50%.

Life in open water exposes zooplankton to heavy predation by visual predators, especially fishes. Many zooplankters are nearly transparent, which affords them some protection. Others, such as fresh-water mites, have conspicuous coloration to advertise their noxious taste, so that visual predators learn to avoid them. In the blue light that penetrates to deep water, reddish hues appear black and thus invisible; it is not surprising that many bathypelagic zooplankton are red. Another antipredation adaptation is cyclomorphosis. Kairomones (chemical messengers that are beneficial to the recepient) released by predators appear to induce cyclomorphosis in succeeding generations of rotifers and cladocerans. These changes in morphology, such as increased spine lengths or head shields, reduce predation. Some adaptations have multiple advantages; the gelatinous sheath of *Holopedium* reduces predation by invertebrates in addition to slowing the sinking rate. *See* ADAPTATION (BIOLOGY).

Vertical migration. Most major zooplankton groups have at least some species that display diel migrations which usually consist of downward movement during the day and upward movement at night. The distance traveled can be hundreds of meters. The same species may display the classical day-down, night-up movement in one area and display no migratory behavior in another location. The sex and age of the zooplankton, as well as the season, can affect their vertical position in the water column and the degree of migration observed.

Diurnal light variation is the most likely mechanism triggering vertical migration. Many planktonic animals are positively phototactic at low light intensities and negatively phototactic at high intensities. Although the general consensus is that light is a major stimulus for the timing of migration, there are many explanations for its purpose. Zooplankton may sink to depths where illumination is insufficient for detection by visual predators in daytime, while at night, when visual predators do not hunt, zooplankton can return to the surface to feed on phytoplankton. Another explanation is that zooplankton remain in deeper, colder water during the day to reduce their metabolism, and return to the surface at night to feed on phytoplankton. The energetic advantage gained by remaining in colder water must exceed that expended in migration and lost due to lack of continuous feeding at the surface. Additionally, the potentially reduced food quantity and the cold temperature of the deeper water result in slower egg development and thus lower reproductive rates. Differential migration by ecologically similar species could also reduce competition for resources. It is unlikely that a single factor can explain all vertical migration; the above factors and others probably interact, or are important at different times.

Communities. Zooplankton, like all organisms, have a range of environmental conditions to which they are adapted. The optimum environment for one species may be barely tolerable to another. Physical and chemical boundary conditions, including turbulence, light, temperature, and salinity gradients, are important in determining species makeup of a zooplankton community. For example, plankton on the two sides of the Gulf Stream or within the upper and lower waters of a lake differ considerably. Some zooplankton have such clear-cut ecological demands that the presence of particular species can indicate the origin of the water mass. Narrow temperature tolerances limit some taxa to tropical or polar waters or to certain periods of the year in temperate zones. Differential salinity tolerances are reflected by changes in the composition of a community as estuaries become increasingly brackish downstream. Thus zooplankton populations are not distributed homogeneously, but tend toward both vertical and horizontal patchiness.

In addition to the physical and chemical boundary conditions mentioned above, biological interactions such as predation, food availability, reproduction, and social behavior affect the distribution of zooplankton and the resulting community structure. Trophic interactions, that is, "what eats what," are of major importance in structuring zooplankton communities. The presence of fishes that feed selectively on larger zooplankton can limit the species composition to smaller-bodied animals. Lakes without size-selective vertebrate predators usually have a higher proportion of larger invertebrate plankton. Likewise, selective feeding by zooplankton on various algae and protozoa will alter the community structure of these groups.

Seasonal breeding of holoplanktonic organisms and entrance of larval meroplankton into the water column play a large part in changing community composition. Food supply, temperature, and other factors interact to determine seasonal breeding patterns in zooplankton. Increasing light and temperature in temperate and boreal areas leads to spring phytoplankton blooms, when many zooplankters release their young. Other groups have maximum abundances in summer or during a secondary peak of primary production in the fall. *See* ECOLOGICAL COMMUNITIES.

Productivity. Productivity is the amount of biomass generated per unit time. There is extensive geographic variation in the pattern and magnitude of productivity of the phytoplankton that are the base of planktonic food webs. This variation in primary production dictates a similar difference in the secondary productivity of zooplankton in both lakes and marine systems. Zooplankton tend to grow faster or produce more young when high-quality food is readily available. However, organisms are not 100% efficient at converting food to biomass. Some

ingested food is not assimilated into the animals' bodies, and some assimilated food is lost to maintenance metabolism and respiration; only what is left becomes available for secondary production (growth and reproductive output). Although conversion efficiencies as high as 90% have been determined for individual zooplankton species feeding on high-quality foods in laboratory experiments, 10–30% efficiencies are more common. Zooplankton production is difficult to determine for natural populations because of the diversity of zooplankton species, the differential digestibility and availability of their food sources, and temperature-induced alterations in their metabolism, among other reasons. *See* BIOMASS.

Planktonic food web. The classic description of the trophic dynamics of plankton is a food chain consisting of algae grazed by crustacean zooplankton which are in turn ingested by fishes. This model may hold true to a degree in some environments such as upwelling areas, but it masks the complexity of most natural food webs. Zooplankton have an essential role in linking trophic levels, but several intermediate zooplankton consumers can exist between the primary producers (phytoplankton) and fish. Thus, food webs with multiple links to different organisms indicate the versatility of food choice and energy transfer and are a more realistic description of the planktonic trophic interactions.

Size is of major importance in planktonic food webs. Most zooplankton tend to feed on organisms that have a body size smaller than their own. However, factors other than size also modify feeding interactions. Within a range of ingestible sizes, dependent on species and age, zooplankton can track changes in the particle-size spectrum of phytoplankton and graze the most abundant size classes. Some phytoplankton are noxious and are avoided by zooplankton, and others are ingested but not digested. Furthermore, zooplankton frequently assume different feeding habits as they grow from larval to adult form. They may ingest bacteria or phytoplankton at one stage of their life cycle and become raptorial feeders later. Other zooplankton are primarily herbivorous but also ingest heterotrophic protists and can opportunistically become carnivorous. Consequently, omnivory, which is considered rare in terrestrial systems, is a relatively common trophic strategy in the plankton. In all food webs, some individuals die without being consumed and are utilized by scavengers and ultimately by decomposers (bacteria and fungi). *See* ECOLOGY; ECOSYSTEM; MARINE ECOLOGY; PHYTOPLANKTON.

Robert W. Sanders

Bibliography. W. Lampert and U. Sommer, *Limnoecology*, 1997; M. Omori and I. Tsutomu, *Methods in Marine Zooplankton Ecology*, 1992; U. Sommer (ed.), *Plankton Ecology*, 1989; T. Tamminen and H. Kuosa (eds.), *Eutrophication in Planktonic Ecosystems: Food Web Dynamics and Elemental Cycling*, 19XX; J. Thorp and A. Covich (eds.), *Ecology and Classification of North American Freshwater Invertebrates*, 1991; C. D. Todd, M. S. Laverack, and G. Boxshall, *Coastal Marine Zooplankton*, 1996.

Zoraptera

An order of minute insects superficially resembling termites (Isoptera) or booklice (Psocoptera). Although species nest in decaying wood, they are of no economic importance. Interest in the group principally derives from their unique biology, anatomy, and evolutionary relationship to other insects.

Morphology and development. Zorapterans are generally less than 3 mm (0.11 in.) in total body length and can be distinguished by their characteristic two-segmented tarsi; unsegmented cerci; nine-segmented antennae (which is reduced to eight segments in two fossil species from the Cretaceous); dehiscent, paddle-shaped wings with uniquely reduced pattern of venation; greatly enlarged hind femora bearing rows of stiff spines along their undersurfaces; asymmetrical male genitalia; and vestigial ovipositor. Individuals of each species occur in two morphs (sometimes referred to as "castes" but not to be confused with the actual castes of truly social insects): eyed and winged forms that shed their wings after dispersal (becoming what are called "dealates"), or blind and wingless forms that predominate in colonies (see **illustration**).

Development progresses through a series of nymphal stages. The relatively soft integument of zoopterans is pale in nymphal stages and typically reddish brown in adults.

Ecology. Zoraptera are gregarious, living in small colonies of 15–120 individuals in crevices or under bark of moist, decaying logs. Zorapterans will sometimes settle in human-made sawdust piles, but such colonies tend not to last as long as those in logs or stumps. Species prefer rotting wood that has decomposed to the point that logs can be easily torn through by hand or with an ordinary garden tool. Colonies harbor in naturally formed spaces in logs that are not reached by light. Once the wood is disturbed, individuals quickly scatter to avoid detection. Species feed principally on fungal hyphae and spores but can also be generalist scavengers or predators, victimizing nematodes, mites, or other tiny arthropods.

Distribution. Zoraptera are principally distributed pantropically, with only four species occurring north of the Tropic of Cancer (that is, North of 23.5°N): two species in North America and two in Tibet. Fossil records for Zoraptera are all from ambers formed in warm, tropical paleoclimates. Although species have at times been considered highly endemic, due to their poor dispersal abilities, they are increasingly being discovered to have larger geographic ranges than previously understood and are apparently capable dispersers. The presence of zorapterans on distantly isolated islands of relatively recent geological age attests to their dispersal capabilities. In fact, the "rarity" of Zoraptera may be more the result of poor collecting than any actual scarcity. While some species are assuredly uncommon, the actual abundance of zorapterans does not live up to its reputation for scant occurrence.

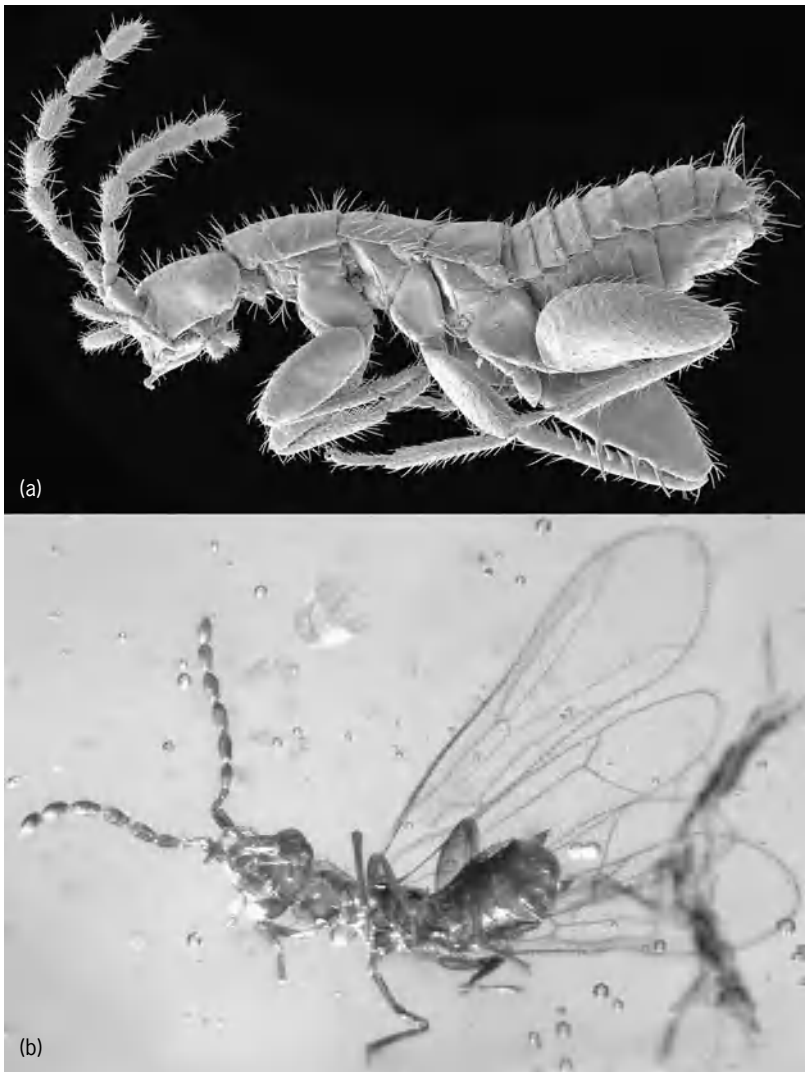

Representative zorapterans. The greatest length of both individuals is 2.5 mm. (*a*) Adult female of *Zorotypus hubbardi* (blind and wingless morph). (*b*) Adult female of *Zorotypus goeleti* (winged morph). (*Modified from D. Grimaldi and M. S. Engel, Evolution of the Insects, Cambridge University Press, 2005; photos © M. S. Engel*)

Classification. Zorapterans are notable for having only 32 modern species and six fossil species known, these classified into two genera (*Zorotypus* and *Xenozorotypus*) in a single family. Once believed to be related to termites or booklice, current hypotheses associate the zorapterans with webspinners (Embiodea, Embiidina, Embioptera), as the two groups share a unique pattern of musculature in the hind legs, a reduced anal region in the wings, occurrence of wingless morphs, and a gregarious biology.

Fossil history. Zorapterans are clearly of ancient origin, perhaps diverging from their webspinner relatives during the Late Triassic or Early Jurassic (approximately 200 million years ago). Relatively modern-looking species of Zoraptera are documented in amber from as long ago as the Early Cretaceous, and fossils from this time already exhibit the development of dual morphs within species. *See* INSECTA; ISOPTERA; PSOCOPTERA. Michael S. Engel

Bibliography. J. C. Choe, Courtship feeding and repeated mating in *Zorotypus barberi* (Insecta: Zoraptera) *Anim. Behav.*, 49:1511–1520, 1995; M. S. Engel and D. A. Grimaldi, A winged *Zorotypus* in Miocene amber from the Dominican Republic (Zoraptera: Zorotypidae), with discussion on relationships of and within the order, *Acta Geol. Hisp.*, 35:149–164, 2000; M. S. Engel and D. A. Grimaldi, The first Mesozoic Zoraptera (Insecta), *Amer. Mus. Novit.*, 3362:1–20, 2002; D. Grimaldi and M. S. Engel, *Evolution of the Insects*, 2005; A. B. Gurney, A synopsis of the order Zoraptera, with notes on the biology of *Zorotypus hubbardi Caudell, Proc. Entomol. Soc. Wash.*, 40:57–87, 1938.

Zosterophyllopsida

An extinct class of the division Lycophyta, comprising primitive vascular land plants that evolved from the rhyniophytes and reached a relatively brief peak in the late Early Devonian, when they were an ecologically important and widely distributed element of the exclusively herbaceous terrestrial flora.

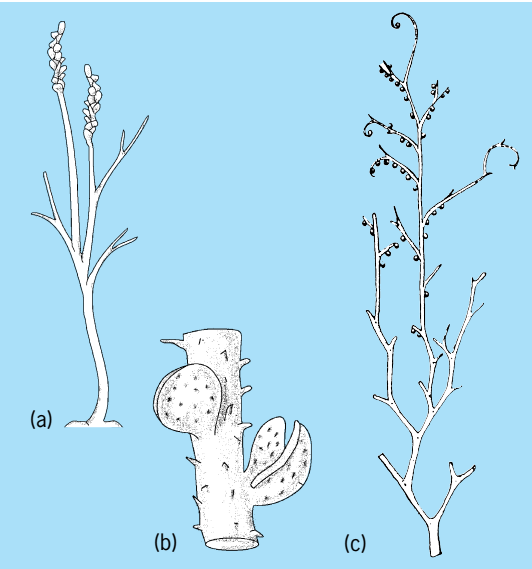

Examples of Devonian Zosterophyllopsida. (*a*) Reconstruction of the Early Devonian *Zosterophyllum rhenanum*, here depicted as an aquatic plant (*after R. Kräusel and H. Weyland, Neue Pflanzenfunde im rheinischen Unterdevon, Palaeontographica B, 80:170–190, 1935*). (*b*) Fertile axis of the Middle Devonian *Sawdonia acanthotheca*, showing the unequally valved sporangia and spinelike enations (*after P. G. Gensel, A new species of Sawdonia, Amer. J. Bot., 136:50–62, 1975*). (*c*) Reconstruction of the more derived, Early-Middle Devonian *Gosslingia breconensis* (*after D. Edwards, Fertile Rhyniophytina from the Lower Devonian, Palaeontology, 13:451–461, 1970*).

The Zosterophyllopsida consists of approximately 18 genera, although few of these include species that have been fully reconstructed from their disarticulated organs (see **illus.**). All genera exhibited rhizomatous growth. The rhizomes produced adventitious roots and upright aerial branches, which were typically densely crowded. The group is united by the key lycophyte characters of exarch protoxylem maturation and laterally borne, vascularized, kidney-shaped sporangia with multilayered walls and complete distal dehiscence.

Variation shown by the group is best considered by briefly describing a few of the species most successfully reconstructed from disarticulated organs. *Zosterophyllum llanoveranum* had naked axes up to 6 in. (15 cm) tall, which branched sparsely and somewhat unequally. Its main axes terminated in a reproductive zone that bore many radially arranged sporangia, which terminated short lateral branches. The geographically and stratigraphically more widespread *Sawdonia ornata* was somewhat taller and unequally branched. Globose sporangia were borne on short stalks and were concentrated toward the tips of the more distal lateral branches in approximately linear files. *Gosslingia breconensis* possessed the unornamented axes of *Zosterophyllum* and the size, branching habit, circinate apices, and sporangial arrangement of *Sawdonia*. The sporangia were horizontally oriented, and the stele was elliptical throughout the axes.

The most primitive generally accepted lycophyte is *Zosterophyllum*. *Sawdonia* and *Gosslingia* are more advanced zosterophyllopsids, possessing characters that are absent from the lycopsids. Although briefly successful as dense stands on floodplains and channel margins, these more sophisticated zosterophyllopsids were virtually extinct by the end of the Devonian. *See* LYCOPHYTA; LYCOPSIDA.

Richard M. Bateman; William A. DiMichele

Bibliography. F. M. Hueber, Thoughts on the early lycopsids and zosterophylls, *Ann. Mo. Bot. Gard.*, 79:474–499, 1992; K. J. Niklas and H. P. Banks, A reevaluation of the Zosterophyllophytina, *Amer. J. Bot.*, 77:274–283, 1990; W. N. Stewart and G. W. Rothwell, *Paleobotany and the Evolution of Plants*, 1993.

Zygnematales

A large order of green algae (Chlorophyceae) that is characterized by the lack of flagellate cells. Sexual reproduction is effected by the fusion of ameboid or passive gametes in conjugation tubes, which give rise to the alternate name Conjugales. Biochemical and ultrastructural features indicate that the Zygnematales are more closely related to charophytes than they are to most other green algae. *See* CHLOROPHYCEAE.

Zygnematales exhibit two somatic types—unicells (desmids) and unbranched uniseriate filaments. The filaments are usually free-floating, but they may be attached by hapteroid outgrowths. The cell wall, which may have pores, is composed of inner fibrillar cellulose layers and an outer slippery pectic layer. All cells are uninucleate. Sexual reproduction is isogamous or anisogamous, with the zygote developing in the conjugation tube or in one of the parents, respectively. The zygote secretes a thick wall that is usually ornamented or sculptured, and undergoes a period of dormancy. Meiosis occurs at some time (variable, depending upon the species) between conjugation and germination.

Members of this order are essentially restricted to fresh-water and subaerial habitats. They are sometimes placed in their own class, with each family elevated to the rank of order. In the present treatment, the order Zygnematales is divided into two suborders, Zygnematineae and Desmidiineae.

Zygnematineae. In this suborder the cells are cylindrical, of uniform diameter, and solitary or in uniseriate filaments. The vegetative cell wall lacks pores or ornamentation. The chloroplast complement of the cell is one of three types: one or more parietal, usually spirally twisted ribbons extending the length of the cell, each with several pyrenoids; an axile plate extending the length of the cell, without pyrenoids or with few to several pyrenoids; and an axile star or disk lying on either side of the central nucleus, each with a pyrenoid. This suborder includes about 18 genera and 650 species.

All filamentous forms are placed in the family Zygnemataceae. In most genera the protoplast of any cell can function directly as a gamete, but in some

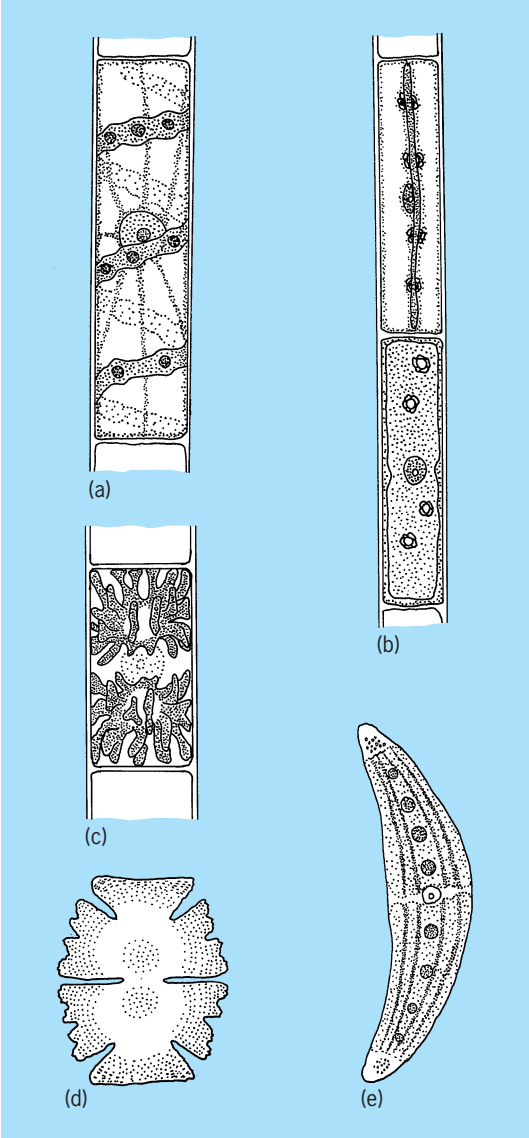

Representative genera of Zygnematales. (*a*) *Spirogyra* (family Zygnemataceae). Cell of filament showing spiral chloroplast with pyrenoids and central nucleus. (*b*) *Mougeotia* (Zygnemataceae). Two cells of filament showing axile ribbonlike chloroplast with pyrenoids, edge view in upper cell, face view in lower cell, and central nucleus. (*c*) *Zygnema* (Zygnemataceae). Cell of filament showing pair of stellate chloroplasts, one on either side of central nucleus. (*d*) *Micrasterias* (Desmidiaceae). A placoderm desmid with incised semicells joined by narrow isthmus. (*e*) *Closterium* (Peniaceae). A placoderm desmid without a median constriction.

genera a division prior to conjugation produces a pair of cells, one of which remains vegetative. Conjugation may be between gametes derived from cells of one filament (lateral) or adjacent filaments (scalariform). The entire protoplast of a cell may function as a gamete, or there may be a cytoplasmic residue. Gametes may be isogamous, meeting halfway in a conjugation tube, or anisogamous, with one gamete ameboid and the other passive. Vegetative multiplication by fragmentation is common.

Spirogyra (**illus.** *a*) has spiral chloroplasts and anisogamous conjugation. In *Mougeotia* (illus. *b*) each cell usually has a single axile laminate chloroplast, and conjugation is isogamous. In *Zygnema* (illus. *c*) each cell usually has a pair of stellate chloroplasts, one on either side of the nucleus, and conjugation is isogamous or anisogamous, depending upon the species. *Zygogonium* has a pair of cushion-shaped chloroplasts, at times connected by a bridge. Conjugation is isogamous and involves cytoplasmic residues.

Quiet bodies of fresh water often contain an abundance of the slippery skeins of the common genera of Zygnemataceae, especially *Spirogyra* and *Zygnema*. *Zygogonium* characteristically grows on damp soil.

Members of the Mesotaeniaceae, called saccoderm desmids, are essentially unicellular Zygnemataceae. Following cell division, daughter cells usually separate immediately, but they may remain united temporarily. Spiral chloroplasts are found in *Spirotaenia,* a single axile laminate chloroplast in *Mesotaenium,* and a pair of stellate chloroplasts in *Cylindrocystis*. Conjugation, which usually involves the formation of a tube, is isogamous. Representatives of this family are found in moist subaerial habitats as well as in quiet bodies of fresh water.

Desmidiineae. Members of this suborder, the placoderm desmids, have cells which are variously shaped but always composed of mirror-image semicells demarcated by a median constriction or incision. During division a cell cleaves transversely, and a new semicell is formed by each of the parental semicells. The nucleus lies at the junction of the semicells (where often a sinus constricts the desmid to form an isthmus). There is usually a chloroplast in each semicell, but the form and number of chloroplasts vary in different genera. The cell wall, which is often impregnated with iron salts, has pores and is frequently ornamented. In most genera the cells are solitary, but in some genera they are united in amorphous colonies or unbranched filaments. Locomotion, known in many placoderm desmids, is effected by localized secretion of gelatinous material through pores at one end of the cell. Conjugation takes place within a gelatinous envelope investing the two ameboid protoplasts or, in filamentous forms, within a tube. The zygote secretes a thick wall that is usually ornamented. Placoderm desmids are abundant in fresh-water habitats, especially those that are slightly acid. They often form a film on submerged plants. This suborder comprises about 34 genera and 4000 species.

All desmids with semicells demarcated by a median constriction or incision are included in the family Desmidiaceae. The genera are distinguished primarily on their somatic organization (solitary, colonial, or filamentous), shape and ornamentation of the semicell, and chloroplast features. The most common genera of solitary desmids are *Cosmarium, Euastrum, Micrasterias* (illus. *d*), *Staurastrum, Staurodesmus*, and *Xanthidium*. The intricate lobing of the semicell of *Micrasterias,* coupled with the unusual type of cell division characteristic of the Desmidiaceae, has made this alga a favorite subject for studies in morphogenesis.

In a small group of placoderm desmids, constituting the family Peniaceae, the cells are cylindrical or lunate, and are composed of semicells not demarcated by a median constriction. Regeneration of a daughter semicell may be followed by elongation of one or both semicells, the zones of elongation being visible in the mature desmid as girdle bands. *Closterium* (illus. *e*) and *Penium* are the most common of the four genera in this family.

Paul C. Silva; Richard L. Moe

Bibliography. H. C. Bold and M. J. Wynne, *Introduction to the Algae: Structure and Reproduction*, 2d ed., 1997; R. E. Lee, *Phycology,* 3d ed., 1999; G. M. Smith, *The Fresh-Water Algae of the United States,* 1950.

Zygomycetes

A class of terrestrial fungi in the phylum Zygomycota, reproducing sexually with a zygospore and asexually by sporangiospore or conidium. The Zygomycetes are one of two classes of the phylum Zygomycota, the other being the arthropod-associated Trichomycetes. Zygomycetes can be distinguished by the production of sporangia (sacs) that undergo internal cleavage to produce multiple sporangiospores (see **illustration**) as opposed to the production of conidiospores as observed in most other groups of fungi (Ascomycota and Basidiomycota). A notable exception is the order Entomophthorales, which reproduces asexually by conidium formation. Zygomycetes are all microscopic with the exception of *Endogone* and relatives that produce a zygospore containing structure called a sporocarp. Zygomycetes are typically studied by first isolating them into pure culture on nutrient media. Sexual reproduction results in the formation of a thick-walled zygosporangium following the conjugation of morphologically relatively similar, undifferentiated gametangia (single-cell structures functioning as gametes). The gametangia may be derived from different (heterothallism) or the same (homothallism) thallus, or genetic individual. *See* EUMYCOTA; FUNGI; ZYGOMYCOTA.

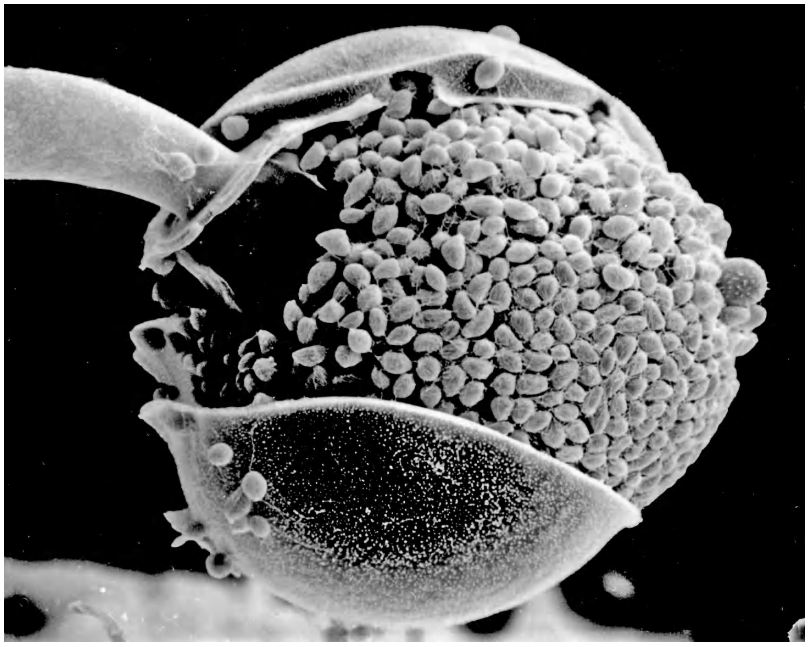

Scanning electron micrograph of dehisced sporangium of *Gilbertella persicaria* releasing sporangiospores. (*Reprinted with permission from K. O'Donnell, Zygomycetes in Culture, 1979; © Department of Botany, University of Georgia*)

The Zygomycetes comprise six orders: Dimargaritales, Endogonales, Entomophthorales, Kickxellales, Mucorales, Zoopagales. These orders span a great diversity of ecological habitat. The Dimargaritales and Zoopagales contain many species which parasitize other fungi (mycoparasites). The Entomophthorales are specialized as pathogens primarily of insects. The Kickxellales contains both mycoparasitic species and saprobic species (especially found on soil and animal dung). Endogonales are unique in that some species associate with plant roots as ectomycorrhizae in what is likely to be a mutualistic relationship. The Mucorales is the largest order and are the most commonly encountered species. These species are primarily saprobic (live on decaying organic matter) in dung and soil but are notorious for spoilage of fruits during storage. The Mucorales include some of the fastest-growing fungi, including the famous "hat-thrower" *Pilobolus. Pilobolus* produces sporangia on dung that are forcibly discharged up to 2 m. These sporangia may land on a plant leaf whereupon they are ingested by herbivores and remain viable until being expelled into fresh substrate upon defecation by the animal.

Zygomycetes do not appear to be a natural group. Recent evidence suggests that the Trichomycetes evolved from within the class Zygomycetes, with a strong relationship between the Trichomycetes and Kickxellales.

Timothy James

Bibliography. G. L. Benny, R. A. Humber, and J. B. Morton, Zygomycota: Zygomycetes, in D. J. McLaughlin, E. G. McLaughlin, and P. A. Lemke (eds.), *The Mycota*, vol. 7A: *Systematics and Evolution*, pp. 113–146, Springer Verlag, Berlin, 2001; P. M. Kirk et al. (eds.), *Ainsworth & Bisby's Dictionary of the Fungi*, 9th ed., CAB International, Wallingford, UK, 2001; C. W. Mims, M. M. Blackwell, and C. J. Alexopoulos, *Introductory Mycology*, 4th ed., 1995.

Zygomycota

A phylum of filamentous, microscopic fungi that undergo sexual reproduction by the formation of a zygospore. The Zygomycota are often referred by the common names "pin molds" or "sugar molds" and are ubiquitous fungi though rarely observed by humans because of the diminutive size of their reproductive structures. Zygomycota grow as a network of

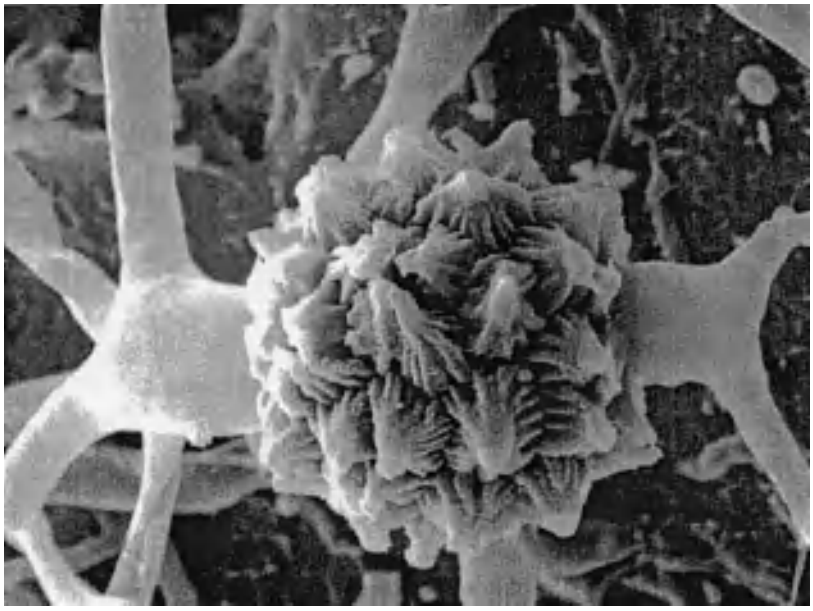

Scanning electron micrograph of ornamented zygosporangium of *Micotypha africana*. (*Reprinted with permission from K. O'Donnell, Zygomycetes in Culture, 1979; copyright Department of Botany, University of Georgia*)

interconnected filaments (hyphae) that grow on or inside their substrate and acquire nutrients through absorptive, heterotrophic nutrition. Zygomycota usually have coenocytic (multinucleate) hyphae lacking septa (cross-walls) between cells, but may also be regularly septate. The Zygomycota have a large diversity of ecological roles: as saprophytes on various substrates, especially sugary substrates and dung, as parasites of animals or plants, as commensualistic (both nonpathogenic and nonbeneficial) associates of arthropods, as parasites of other fungi, and as mycorrhizal symbionts with plant roots. *See* MYCORRHIZAE.

As with other fungi, members of the Zygomycota reproduce through spores. Asexual reproduction appears to be the dominant mode of reproduction in most species; sexual reproduction may be rare, or lacking altogether. Asexual reproduction occurs through the production of sporangiospores produced by the internal cleavage of saclike structures termed sporangia. Sporangia are usually borne at the end of a stalk (the sporangiophore) and are dispersed by air or water following the dehiscence (bursting open) of the sporangia wall. Entomophthorales are an exception and reproduce through the production of forcibly discharged conidia. Zygomycetes are haploid throughout nearly all of their life cycle. Sexual reproduction can occur between compatible individuals (heterothallism) or by self-fertilization of a single individual (homothallism). Sex involves the attraction of compatible mates via the hormone trisporic acid. Compatible hyphae grow toward each other in response to a hormonal gradient established by trisporic acid precursors and form specialized cells called gametangia upon contact. The gametangia undergo plasmogamy (fusion) followed by the development of a zygosporangium within which karyogamy (nuclear fusion) takes place, forming the only diploid cell type. Within the zygosporangium, a single zygospore develops. Zygosporangia are typically thick-walled and often contain extensive ornamentation (see **illustration**). Meiosis occurs within the zygosporangium or at the time of germination.

The Zygomycota may be divided into two classes: the Zygomycetes and the Trichomycetes. The Trichomycetes are distinguished from the Zygomycetes by the production of a specialized spore, the trichospore, which is a sporangiospore with an appendage to aid in aquatic dispersal. The Trichomycetes comprise those Zygomycota which have a commensual relationship with arthropods. Until 2000, the Zygomycota also included the arbuscular mycorrhizal fungi which form mutualistic associations with many plant roots. These fungi are now in their own phylum, the Glomeromycota. *See* EUMYCOTA; FUNGI; ZYGOMYCETES.

Bibliography. G. L. Benny, R. A. Humber, and J. B. Morton, Zygomycota: Zygomycetes, in D. J. McLaughlin, E. G. McLaughlin, and P. A. Lemke (eds.), *The Mycota*, vol. 7A, *Systematics and Evolution*, pp. 113–146, Springer, Berlin, 2001; R. W. Lichtwardt, *The Trichomycetes, Fungal Associates of Arthropods*, Springer, New York, 1986; C. W. Mims, M. M. Blackwell, and C. J. Alexopoulos, *Introductory Mycology*, 4th, ed., 1995.

Zygophyllales

An order of flowering plants in the eurosid I group of the rosid eudicots. This order comprises just two small families of distant relationship (as evidenced by DNA sequence studies) to all other rosid eudicots: Krameriaceae and Zygophyllaceae. In previous systems of classification, both families

were placed in the highly unnatural order Polygalales or Sapindales, the families of which share with them only highly zygomorphic flowers. Krameriaceae (25 species) are photosynthetic parasitic plants ("hemiparasites") found from the southwestern United States to South America, whereas Zygophyllaceae (240 species) are free-living and widely distributed in arid to semiarid zones throughout the warmer regions of the world. Some species of the latter family produce especially hard woods such as lignumvitate (*Bulnesia* and *Guaiacum*), and a few others produce medicinal compounds, such as the creosote bush (*Larrea*). *See* FLOWER; LIGNUMVITAE; MAGNOLIOPHYTA; MAGNOLIOPSIDA; POLYGALALES; ROSIDAE; SAPINDALES.

Mark W. Chase